Lecture Notes in Geoinformation and Cartography

Thomas Blaschke · Stefan Lang ·
Geoffrey J. Hay (Eds.)

Object-Based Image Analysis

Spatial Concepts for Knowledge-Driven
Remote Sensing Applications

Springer

Editors

Prof. Thomas Blaschke
Universität Salzburg
Zentrum für Geoinformatik
Hellbrunner Str. 34
5020 Salzburg
Austria
thomas.blaschke@sbg.ac.at

Dr. Stefan Lang
Universität Salzburg
Zentrum für Geoinformatik
Schillerstr. 30
5020 Salzburg
Austria
stefan.lang@sbg.ac.at

Dr. Geoffrey J. Hay
University of Calgary
Foothills Facility for Remote
Sensing & GIScience
2500 University Dr. NW.
Calgary T2N 1N4
Earth Sciences Bldg.
Canada
gjhay@ucalgary.ca

Preface

This book brings together a collection of invited interdisciplinary perspectives on the recent topic of *Object-based Image Analysis* (OBIA). Its content is based on select papers from the 1st OBIA International Conference held in Salzburg in July 2006, and is enriched by several invited chapters. All submissions have passed through a blind peer-review process resulting in what we believe is a timely volume of the highest scientific, theoretical and technical standards.

The concept of OBIA first gained widespread interest within the GIScience (Geographic Information Science) community circa 2000, with the advent of the first commercial software for what was then termed 'object-oriented image analysis'. However, it is widely agreed that OBIA builds on older segmentation, edge-detection and classification concepts that have been used in remote sensing image analysis for several decades. Nevertheless, its emergence has provided a new critical bridge to spatial concepts applied in multiscale landscape analysis, Geographic Information Systems (GIS) and the synergy between image-objects and their radiometric characteristics and analyses in Earth Observation data (EO).

Over the last year, a critical online discussion within this evolving multidisciplinary community – especially, among the editors – has also arisen concerning whether or not *Geographic space* should be included in the name of this concept. Hay and Castilla argue (in chapter 1.4) that it should be called "Geographic Object-Based Image Analysis" (GEOBIA), as only then will it be clear that it represents a sub-discipline of GIScience. Indeed, the term *OBIA* may be too broad; for it goes without saying for Remote Sensing scientists, GIS specialist and many 'environmental' based disciplines, that 'their' image data represents portions of the Earth's surface. However, such an association may not be taken for granted by scientists in disciplines such as Computer Vision, Material Sciences or Biomedical Imaging that also conduct OBIA. Because this name debate remains ongoing, we have chosen for this book to build on key OBIA concepts so as to lay out generic foundations for the continued evolution of this diverse community of practice. Furthermore, by incorporating a GEOBIA chapter in

this volume, we pave the road ahead for the GEOBIA 2008 conference at the University of Calgary, Alberta, Canada.

Our primary goal in this book is to unveil the concept of OBIA as applied within a broad range of remote sensing applications. Consequently, the first five chapters focus on fundamental and conceptual issues, followed by nine chapters on multiscale representation and object-based classification. These nine chapters include specific aspects such as the incorporation of image-texture, key pre-processing steps and quality assessment issues. The latter being a hot research topic that is repeatedly tackled within the application centric contributions, as well as in the last section on research questions. Since most members of this community are already actively engaged either in OBIA method development or operationalization, we only briefly address the theoretical scientific discourse regarding whether or not OBIA should be considered a paradigm shift according to Kuhn's definition.

The contributions in the first two sections explore and guide application driven development by explaining this new technological and user driven evolution in remote sensing image analysis as it moves from pixels to objects, and the software and infrastructure required to generate and exploit them. Notwithstanding this message, we suggest that the ultimate aim of OBIA should not be to focus on building better segmentation methods, but rather to incorporate and develop *geographic-based intelligence* i.e., appropriate information within a geographical context, and all that this implies to achieve it.

Another critical topic is the automation of image processing. Strongly related to the advent of high-resolution imagery, papers in these sections discuss automatic object delineation. Automated object-recognition is certainly an end goal. Realistically, it is at the moment mainly achieved stepwise, either with strongly interlinked procedures building workflows or with clear breaks in these workflows. In both cases the steps involve addressing various multiscale instances of related objects within a single image (e.g., individual tree crowns, tree clusters, stands, and forests). Several contributions also deal with object- and feature recognition and feature extraction which, though intrinsically tied to OBIA – in the majority of applications – are not an end in itself.

The 18 chapters of Sections 3, 4, 5 and 6 are dedicated to automated classification, mapping and updating. This wide range of applications is structured through four main fields, namely (*i*) forest, (*ii*) environmental resource management and agriculture, (*iii*) land use / land cover, and (*iv*) urban applications. The final two sections are more technical / methodological. The seven chapters of Section 7 cover developments of new

methodologies while the book closes with another four chapters on critical research questions, research needs and an outlook to the future.

This volume was planned as a coherent whole. The selection of submitted contributions was based on their quality, as well as on the overall design and story we wanted to present. Of course, due to the rapidly evolving nature of OBIA, this tome cannot be considered truly 'complete'. While there are certainly technical and methodological issues as well as application fields which have not been addressed, this book does represent the first comprehensive attempt to synthesize OBIA from an international and interdisciplinary perspective without bias to a specific type of image-processing software or EO data type.

Finally, this book covers an extremely challenging topic: the Earth's surface. This complex system can be represented by numerous multiscale image-objects extracted from a plethora of different Earth Observation data types, and yet such remote sensing imagery only indirectly provides clues to its underlying patterns and processes, each of which change with different scales of perception. Yet this linking – between imagery, patterns, process and scale — is exactly what is needed for effective environmental policy support. Only when the complex fabric of our planet can be segmented 'appropriately' and in a transparent and repeatable way, will we achieve 'geo-intelligence'. This latter term is currently dismissed widely outside North America since it is associated with the (U.S.A) homeland security concept. However 'geographic intelligence' is a potential term to describe what OBIA really aims for: using Earth Observation data to delineate and explore the multiscale spatial relationships of appropriately defined image-objects and related ancillary information as they model real-world geographic objects, and provide us new insight to better understand this planet and its function.

Acknowledgements

Thomas Blaschke
Stefan Lang
Geoffrey J. Hay

Contents

External Reviewers

Alzir Felippe B. Antunes
Universidade Federal do Paraná
Setor de Ciências da Terra -
Departamento de Geomática
Curitiba, Paraná, Brasil

Thomas Bauer
Universität für Bodenkultur
Wien
Department für Raum,
Landschaft und Infrastruktur
Institut für Vermessung,
Fernerkundung und
Landinformation
Wien, Österreich

Michael Bock
Deutsches Zentrum für Luft- und
Raumfahrt (DLR)
Deutsches
Fernerkundungsdatenzentrum
Umweltsysteme
Oberpfaffenhofen-Wessling,
Deutschland

Matthias Holger Braun
Zentrum für Fernerkundung der
Landoberfläche
Rheinische Friedrich-Wilhelms
Universität Bonn
Bonn, Deutschland

Adrijana Car
Zentrum für Geoinformatik
Paris-Lodron-Universität
Salzburg
Salzburg, Österreich

Alexandre Carleer
Université Libre de Bruxelles
Bruxelles, Belgique

Piermaria Corona
Università degli Studi della
Tuscia
Dipartimento di Scienze
dell'Ambiente Forestale e delle
Sue risorse
Viterbo, Italia

Baudouin Desclée
Unité d'Énvironnemétrie et de
Géomatique
Faculté d'ingénierie biologique,
agronomique et
environnementale
Louvain-la-Neuve, Belgique

Lucian Dragut
Zentrum für Geoinformatik
Paris-Lodron-Universität
Salzburg
Salzburg, Österreich

Marcelle Grenier
Geomatic and Remote Sensing
Environment Canada
Canadian Wildlife Service,
Québec
Ecosystem conservation
Sainte-Foy, Québec, Canada

Görres Grenzdörffer
Universität Rostock
Agrar- und
Umweltwissenschaftliche
Fakultät
Professur für Geodäsie und
Geoinformatik
Rostock, Deutschland

Andrew T. Hudak
USFS Rocky Mountain Research
Station
Moscow, ID, USA

Barbara Koch
Abteilung für Fernerkundung
und
Landschaftsinformationssysteme
Fakultät für Forst- und
Umweltwissenschaften
Albert-Ludwigs-Universität
Freiburg, Deutschland

Filip Langar
Abteilung für Fernerkundung
und
Landschaftsinformationssysteme
Fakultät für Forst- und
Umweltwissenschaften
Albert-Ludwigs-Universität
Freiburg, Deutschland

Richard G. Lathrop
Walton Center for Remote
Sensing & Spatial Analysis
Cook College, Rutgers
University
New Brunswick, NJ, USA

Durieux Laurent
Maison de la Télédétection
Montpellier, France

Wolfgang Lück
Council for Scientific and
Industrial Research
Satellite Application Centre
Earth Observation Service
Centre
Pretoria, Gauteng, South Africa

Matthias Möller
Austrian Academy of Sciences
Geographic Information Science
Salzburg, Österreich

Mathias Schardt
Institut für Digitale
Bildverarbeitung
Joanneum Research
Graz, Österreich

Mustafa Türker
Hacettepe Üniversitesi
Mühendislik Fakültesi
Jeodezi ve Fotogrametri
Mühendisliği Bölümü
Beytepe, Ankara, Turkey

Jacek Urbanski
Instytut Oceanografii
Uniwersytet Gdański,
Gdynia, Poland

Wolfgang Wagner
Institut für Photogrammetrie und
Fernerkundung
Technische Universität Wien
Wien, Österreich

Section 1
Why object-based image analysis

Chapter 1.1

Object-based image analysis for remote sensing applications: modeling reality – dealing with complexity

S. Lang [1]

[1] Centre for Geoinformatics (Z_GIS), Paris-Lodron University Salzburg, Schillerstr. 30, 5020 Salzburg, Austria; stefan.lang@sbg.ac.at

KEYWORDS: GMES, monitoring, geon, class modeling, cognition networks, object assessment

ABSTRACT: The advancement of feature recognition and advanced image analysis techniques facilitates the extraction of thematic information, for policy making support and informed decisions. As a strong driver, the availability of VHSR data and the ever increasing use of geo-information for all kinds of spatially relevant management issues have catalyzed the development of new methods to exploit image information more 'intelligently'. This chapter highlights some of the recent developments from both technology and policy and poses a synthetic view on an upcoming paradigm in image analysis and the extraction of geo-spatial information. It starts from a review of requirements from international initiatives like GMES (Global Monitoring of Environment and Security), followed by a discussion the possible answers from OBIA including a detailed portrait of the methodological framework of class modeling. The chapter closes with a short reflection on the required adaptation of standard methods of accuracy assessment and change detection, as well as on the evaluation of delineated and classified objects against the ultimate benchmark, our human perception.

1 Monitoring needs in a dynamic world

There is an ever increasing demand for regularly updated geo-spatial in-formation combined with techniques for rapid extraction and targeted pro-vision of relevant information. The need for timely and accurate geo-spatial information is steadily increasing, trying to keep pace with the changing requirements of the society at a global dimension. International initiatives strive for standardized solutions, as examples like cooperative effort of Global Earth Observation System of Systems (GEOSS) or the European initiative Global Monitoring for Environment and Security (GMES) impressively show. These initiatives strive to provide holistic, yet operational answers to global conventions or trans-national directives and agendas to halt uncontrollable change of physical parameters or loss of lives both potentially human-induced by unlimited growth (e.g. UN Framework Convention on Climate Change, FCCC or the UN Convention on Biological Diversity, CBD or the UN Convention to Combat Desertifi-cation, CCD[1]; EC Water Framework Directive, WFD or the EC Flora-Fauna-Habitat-Directive, FFH).

GMES: various applications – one approach

Beyond these more environmental aspects, the EC-ESA conjoint initiative GMES[2] follows the idea, that environmental integrity and societal stability both play together and may reinforce each other under certain conditions. The 'S' stands for security, and – next to environmental risks or potential hazards – the scope of observed treaties and conventions also comprises terrorism, critical infrastructure, refugees and weapons of mass destruc-tion, to name just a few topics. Both GEOSS and GMES rely on remote sensing (RS) technology as a powerful and ubiquitous data provider, and both initiatives promote the integration of RS with in-situ data technology for the development of operational monitoring systems and integrated ser-vices, based on earth observation (EO) data. The scope is wide covering areas ranging from environmental integrity to human security, and the idea of serving this range of applications with a profound, ubiquitous set of pooled data and adaptive methods is compelling. And more than this, the approach is concept-wise sustainable, both in terms of its scientific strength and its operational capability. Operational services, delivered as

---

[1] http://www.unccd.int/
[2] www.gmes.info

fast-track core services (FTCS)[3] and such in preparation[4] provide status observations of highly complex systems with relevance to treaties and political agreements of different kinds. However, dealing with such advanced integrated tasks may no longer keep valid the monitoring of single compartments, but an integrated high-level approach (see 2).

Monitoring systems as required in the GMES context (Zeil et al., in press; Blaschke et al., 2007; Tiede & Lang; 2007) need to be capable of transforming complex scene content into ready-to-use information. The advancement in feature recognition and advanced image analysis techniques facilitates the extraction of thematic information, for policy making support and informed decisions, irrespective of particular application fields. The availability of such data and the increasing use of geo-information for sustainable economic development and protection of the environment have catalyzed the development of new methods to exploit image information more efficiently and target-oriented. Global commitments, directives and policies with their pronounced demand for timely, accurate and conditioned geo-spatial information, ask for an effective answer to an ever increasing load of data collected from various monitoring systems. It is obvious, yet maybe not consciously thought of, that – along with ever improved sensor technology – a technically and spatially literate user community asks for ever more advanced geo-spatial products, and expresses their needs accordingly. With an increased level of consciousness of prevailing problems the need for targeted information is rising double, it seems. The remote sensing community has to react and must deliver. When industry primarily highlights achievements in sensor developments, the efforts taken to analyze these data and to generate added value from these can hardly be underlined too much.

The upcoming paradigm of object-based image analysis (OBIA) has high potential to integrate different techniques of processing, retrieval and analyzing multi-resolution data from various sensors. By bridging technical and sector-oriented domains from remote sensing and geoinformatics we may significantly enhance the efficiency of data provision for policy making and decision support.

New data and increasing complexity: OBIA as the answer?

Recent years' advances in sensor technology and digital imaging techniques, along with ever increasing spatial detail, have challenged the re-

---

[3] FTCS on land, sea, and emergency

[4] Pre-operational services on security and atmosphere

mote sensing community to strive for new methods of exploiting imaged information more intelligently. The word 'intelligence' in this context has several facets: (1) an advanced way of supervised delineation and categorization of spatial units, (2) the way of how implicit knowledge or experience is integrated, and (3) the degree, in which the output (results) are contributing to an increase of knowledge and better understanding of complex scene contents.

New earth observation (EO) techniques and concepts from GIScience have led to the emerging field of OBIA[5]. The main purpose of OBIA in the context of remote sensing applications is to provide adequate and automated methods for the analysis of very high spatial resolution (VHSR) imagery by describing the imaged reality using spectral, textural, spatial and topological characteristics. OBIA offers a methodological framework for machine-based interpretation of complex classes, defined by spectral, spatial and structural as well as hierarchical properties (Benz et al., 2004; Burnett & Blaschke, 2003; Schöpfer et al., in press; Niemeyer & Canty, 2001; Hay et al., 2003). OBIA has been pushed by the introduction of fine resolution image data that for a broad range of application domains provides an h-res situation (Strahler et al., 1986). In h-res situations, the pixel size is significantly smaller than the average size of the object of interest. In this constellation, segmentation as a means of regionalization is an efficient means of aggregation the high level of detail and producing usable objects. Therefore, segmentation is a crucial methodological element in OBIA, but not an exclusive or isolated one (see 3).

VHRS satellite imagery is widely available now and gained popularity in research, administration and private use. If not the 'real' data, so at least the 'natural color' products can be easily accessed through web-based virtual globes like Google Earth, NASA World Wind, MS Virtual Earth and the like. Globes have penetrated daily life information exchanges, and satellite data in a 'this-is-my-house-resolution' have become the staple diet to feed people's greed for immediate contextual spatial information[6].

From a scientific point of view satellite-mounted sensors and air-borne scanners have now reached the level of detail of classical aerial photography. For decades, fine-scaled (i.e. larger than 1:10,000) environmental as well as security-related applications were relying on aerial photography and visual inspection as a primary means to extract relevant information.

---

[5] The scientific community discusses to use the term GEOBIA to emphasize (1) the strong contribution of GIScience concepts and (2) the focus on space related applications (see Castilla & Hay in this volume).

[6] See Tiede & Lang (2007) for a discussion how the presence of familiar spatial context can be utilizing for communicating complex analytical contents.

On the other hand, medium to low resolution satellite sensors were mainly used for coarse-scaled mapping and multi-spectral classification, with a probabilistic, yet limited set of classes being targeted at (Lang, 2005). This quite dichotomous gap has been separating two application domains and the respective methods applied, e.g. in landscape-related studies which had to choose between either of them (Groom et al, 2004).

Closing this gap, but embarking on another class of problem: with the advent of digital data from airborne and satellite-borne sensors we return to the very challenge of air-photo interpretation: how do we deal with the enormous detail? Looking back to several decades of computer technology we trust in automated methods for analysis and interpretation, even of complex imaged contents. While a several problems remain a challenge, a range of tangible solutions have been developed by successfully combining GIS and remote sensing techniques for reaching closer at the photo-interpreter's capacity.

As briefly mentioned before, the need for highly accurate and regularly updated geo-spatial information cannot be met by advancements of sensor technology alone. New sensors and new kinds of data may do provide a wealth of information, but this 'information overload' needs to be conditioned, in order to fit the communication channels of the addressees. Thus, advanced methods are required to integrate single information packages. It is necessary to both synchronize technologies and harmonize approaches. The first is related to the acquisition, pre-processing, and retrieval of multi-sensor, multi-spectral, multi-resolution data from various sensors. The second deals with the integration of spatial analysis techniques into image processing procedures for addressing complex classes in a transparent, transferable and flexible manner.

The guiding principle of OBIA is likewise clear as it is ambitious: to represent complex scene content in such a way that the imaged reality is best described and a maximum of the respective content is understood, extracted and conveyed to users (including researchers). The realization, therefore, is not trivial, as the ultimate benchmark of OBIA is human perception (see 3.3). This, our visual sense of the environment is a common experience, easy to share yet difficult to express in words or even rule sets. Indeed, the challenge is to make explicit the way how we deal with imaged information in various scales, how we manage to relate recognized objects to each other with ease, how we understand complex scene contents readily. To this end, OBIA utilizes concepts from spatial thinking, which again is influenced by cognitive psychology.

2 A plurality of solutions – conditioned information and geons

An increasing detail of data and complex analysis tasks opens the door for a plurality of solutions. Often, there is no longer a single valid choice of e.g. a distinct land cover class. Rather, there is a user-driven set of classes; not necessarily restricted to extractable features, but expressed according to the very demand. Fine-scaled representations of complex real world phenomena require means for modeling the underlying complexity, for mapping the dynamics and constant changes. Automated techniques making effective use of advanced analysis methods help understanding complex scene contents and try to match the information extraction with our world view.

But OBIA is more than feature extraction (see chapter 3). It provides a unifying framework with implications for policy-oriented delivery of conditioned information. By this, it also enables monitoring of system-driven meta-indicators like vulnerability or stability.
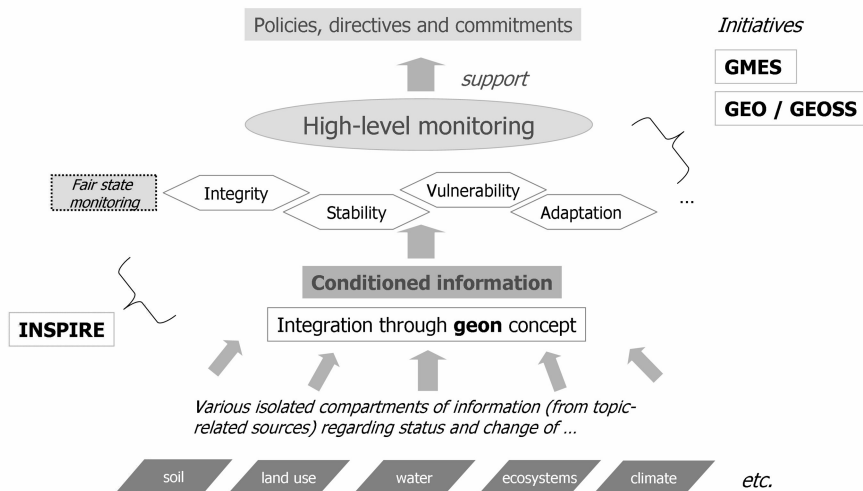
Fig. 1. High-level indicators for monitoring

From this point of view, a broad concept of manageable units makes sense. The author proposes the term *geon* (Lang & Tiede, 2007) to de-

scribe generic spatial objects that are derived by regionalization and homogenous in terms of a varying spatial phenomenon under the influence of, and partly controlled by, policy actions. A geon[7] (from Greek *gé* = Earth and *on* = part, unit) can be defined as a homogenous geo-spatial referencing unit, specifically designed for policy-related spatial decisions. The geon concept (see figure 2) can be seen as a framework for the regionalization of continuous spatial information according to defined parameters of homogeneity. It is flexible in terms of a certain perception of a problem (policy relevance/scope). It employs a comprehensive pool of techniques, tools and methods for (1) geon generation (i.e. transformation of continuous spatial information into discrete objects by algorithms for interpolation, segmentation, regionalization, generalization); (2) analyzing the spatial arrangement, which leads to emergent properties and specific spatial qualities; and (3) monitoring of modifications and changes and evaluation of state development. The latter, characterizing spatio-temporal variability require appropriate means to distinguish noise or slight modifications from real changes. In addition, there is the possibility of recovering objects in the presence of 'occlusions[8]' (i.e. data errors, measure failures, lack of data, mismatch of data due to bad referencing).

---

[7] The term geon (for geometric ions) was initially used by Biederman (1987), who hypothizes that cognitive objects can be decomposed into basic shapes or components. Geons in Biederman's view are basic volumetric bodies such as cubes, spheres, cylinders, and wedges. The concept used here is related, but not identical to this view.

[8] This term, again, is borrowed from Biederman (1987)
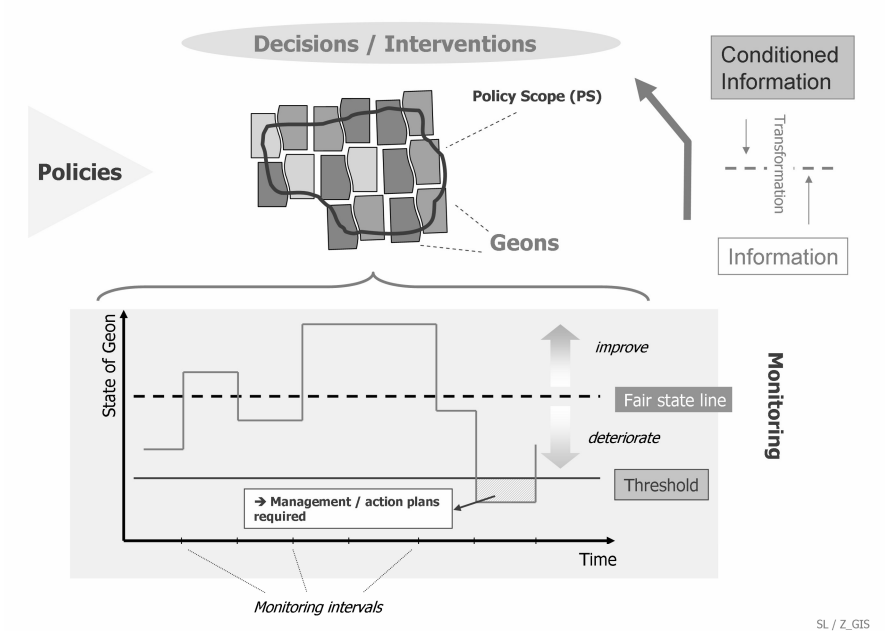
Fig. 2. The Geon Concept

Within the spatial extent in which a certain policy applies (policy scope, PS), a group of geons makes up a spatially exhaustive set (geon set). PS comprises the spatio-temporal extent in which a certain policy is valid. This extent usually coincides with administrative units, but not necessarily does: in the context of the EC Water Framework Directive, catchments function as reference units. In cases when PS is defined by legal boundaries, the spatial limit of the geon set, as derived functionally, may not fully coincide with PS. As policies address various scale domains and their implications apply to particular domains, a geon set is scale-specific and adapted to the respective policy level. Several geon sets may exist, altogether forming in a spatial hierarchy. Using geons, we are capable of transforming singular pieces of information on specific systemic components to policy-relevant, conditioned information. Geons are of dynamic nature. Monitoring the spatio-temporal development of geons is critical for assessing the impact of policies and the compliance with obligations or commitments attached to those. The 'fate' of a geon is monitored using the fair state concept, which takes into account the natural dynamics (improvement or deterioration in regard to an optimum or average state). Management actions have to be taken, when a certain threshold is reached.

Irrespective of the very concept applied for naming delineated units, and irrespective of the different fields of use, OBIA aims at the delineation and the classification of relevant spatial units. The way to perform this task is an integrated, cyclic one, and in the following section this will be discussed, under the heading 'class modeling'.

3 Class modeling

From a methodological point of view, one may observe a convergence of various techniques from formerly distinct GIS and remote sensing embankments; aiming at the aforementioned purpose, OBIA is trying to bridge these. OBIA rests upon two interrelated methodological pillars, i.e. (1) segmentation / regionalization for nested, scaled representations; (2) rule-based classifiers for making explicit the required spectral and geometrical properties as well as spatial relationships for advanced class modeling. We speak of 'image analysis' and not merely of 'image classification', since the process of OBIA is iterative rather than a linear and strict subsequent one. The process of OBIA is a cyclic one. It is usually not enough to think of (a) delineation and (b) labeling[9]. By its iterative nature, the process is highly adaptive and open for accommodating different categories of target classes, from specific domains, with different semantics, etc. Class modeling (Tiede et al., 2006; Tiede & Hoffmann, 2006) enables operators to tailoring transformation of scene contents into ready-to-use information according to user requirements. It supports data integration and the utilization of geo-spatial data other than images (e.g. continuous data like altitude or data representing administrative units).

[9] To underline this, Baatz et al. (this volume) propose the term "object-oriented" to be used instead of "object-based", because the former is more target-oriented, teleological, whereas the latter may be misleading and implying a more static concept. However, since the term "object-oriented" is strongly occupied by computer scientists, the author stays with "object-based".
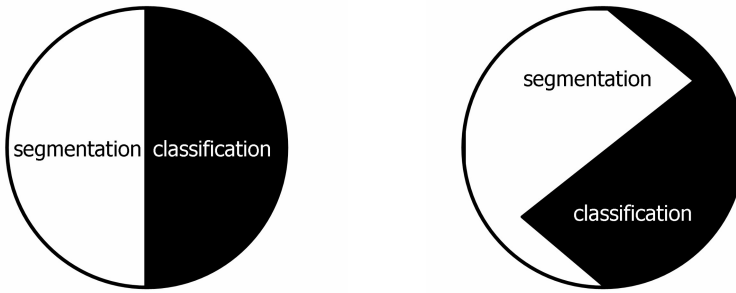
Fig. 3. Class modeling: the symbology marks the move from distinct realms of segmentation and classification towards an interlinked concept

Class modeling (as for example realized by the modular programming language CNL, cognition network language), provides flexibility in providing problem-oriented solutions for advanced analysis tasks. Examples are scene-specific high-level segmentation and region-specific multi-scale modeling (Tiede et al., this volume) or the composition of structurally defined classes as proposed by Lang & Langanke (2006). The latter was successfully realized in a study on semi-automated delineating biotope complexes (Schumacher et al., 2007, see figure 9a).
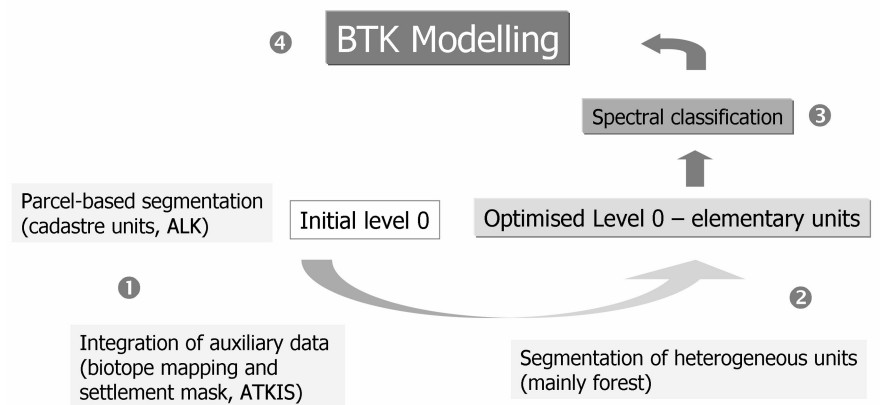
Fig. 4. Class modeling – a cyclic process. The example of modeling biotope complexes

From pixels (via image regions) to image objects

In object-based image analysis, the 'image object' is the central methodological element and as an object of investigation, it resides somewhere between application-driven plausibility and technology-driven detectability. To this end, we conjoin image segmentation with knowledge-based classification. Image segmentation decreases the level of detail, reduces image complexity, and makes image content graspable. Segmentation produces image regions, and these regions, once they are considered 'meaningful', become image objects; in other words an image object is a 'peer reviewed' image region; refereed by a human expert. A pixel as a technically defined unit can be interpreted in terms of its spectral behavior, in terms of the aggregation of spectral end-members, or in terms of its neighborhood. A pixel cannot be assigned a valid corresponding real-world object, but an image object can. Overcoming the pixel view and providing image objects that 'make sense' opens a new dimension in rule-based automated image analysis; image objects can be labeled directly using a range of characteristics, including spatial ones, or they can be used for modeling complex classes based on their spatial relationships. Coupled with e.g. a rule-based production system we can make expert knowledge explicit by the use of rules (see below).

Hierarchical, multi-scale segmentation

Multi-scale segmentation has often been linked with hierarchy theory (Lang, 2002). This is an appealing concept, and the comparison seems obvious as both hierarchy theory and multi-scale segmentation deal with hierarchical organization. Still we need to be careful: hierarchy theory proposes the decomposability of complex systems (Simon, 1973), but imagery is just *a* representation of such systems. An imaged representation is in several aspects reductionism: it is a plane picture only revealing reflection values. So hierarchy theory forms a strong conceptual framework, rather than a methodological term for multiple segmentation cycles. What we delineate, needs to be assigned a function first (image regions – image objects – (geons) – classes, see above). We should be aware that hierarchical segmentation at first hand produces regions of increasing average size (or number of pixels, respectively). But hierarchy theory is not about size, it deals with increasing degree of organization (Laszlo, 1972; Szaramovicz, 2004). What makes a strong link to hierarchy theory is not multiple segmentation alone, but the way we approach complexity, how we model and decompose it, and how we integrate our knowledge about it.

When fitting image complexity into hierarchical levels, it does not happen independently from human perception (Lang 2005). Allen & Starr (1982) point out that "discrete levels need to be recognized as convenience, not truth" and levels would emerge "as device of the observer" (ibid.). While drastically expressed, we should be aware that instead of questioning the ontological truth of scaled representations, we should rather focus on their epistemological relevance for target objects. Human perception is a complex matter of filtering relevant signals from noise, a selective processing of detailed information and, of course, experience. To improve automated object extraction we therefore seek for mimicking the way how human perception works (see 3.3 and Corcoran & Winstanley, this volume).

Is there one single set of multiple segmentations applicable 'horizontally' over the entire scene? The multi-scale option does not always lead to satisfying results. This applies in scenes, where multiple scales occur in different domains. Tiede et al. (this volume) discuss an application of regionalized segmentation (Lang, 2002).

Nested scaled representations need to consider scale. While this statement is somewhat tautologically, there is no single solution to this and different approaches exist to address this. One, aiming at a strict hierarchical representation, performs multi-scale segmentation with coinciding boundaries. In other words: a super-object gets assigned exactly n sub-objects (Baatz & Schäpe, 2000). The advantage is a clear 1:n relationship between super- and sub-object. On the other hand, since boundaries are 'cloned' up the scaling ladder (Wu, 1999), boundaries will not be generalized. It is scale-adapted, but not scale-specific. On the other hand, scaled representations are scale-specific, if there is – as in visual interpretation – a match between average size and generalization of boundaries. This is for example realized in the software SCRM (Castilla, 2004). The drawback is, however, boundaries do not coincide and cannot be used for 'direct' modeling (but see Schöpfer et al., Weinke et al., this volume).

Knowledge representation and cognition networks

Knowledge plays a key role in the interpretation-focused, value-adding part of the remote sensing process chain (Campbell, 2002). We have at our disposal a huge store of implicit knowledge and a substantial part of it is used in image interpretation (ibid.). By training we subsequently complement implicit knowledge with a more formalized knowledge obtained through formal learning situations (e.g. the specific spectral behavior of stressed vegetation) and experience. From an artificial intelligence (AI)

perspective two components of knowledge can be distinguished, procedural and structural knowledge. Procedural knowledge is concerned with the specific computational functions and is therefore represented by a set of rules. Structural knowledge implies the way of how concepts of e.g. a certain application domain are interrelated: in our case that means, in how far links between image objects and 'real world' geographical features is established. Structural knowledge is characterized by high semantic contents and it is difficult to tackle with. A way to organize structural knowledge is the use of knowledge organizing systems (KOS), either realized by graphic notations such as semantic networks (Ibrahim, 2000; Pinz, 1994; Liedtke et al., 1997; Sowa, 1999) or and more mathematical theories like formal concept analysis (FCA, Ganter & Wille, 1996). Within image analysis semantic nets and frames (Pinz, 1994) provide a formal framework for semantic knowledge representation using an inheritance concept (*is part of, is more specific than, is instance of*). As semantic nets need to be built manually, they allow for controlling each and every existing connection once being established. With increasing complexity the transparency and operability will reach a limit. Bayesian networks are manually built, but the weighting of the connections can be trained, though it has to be trained for every connection.
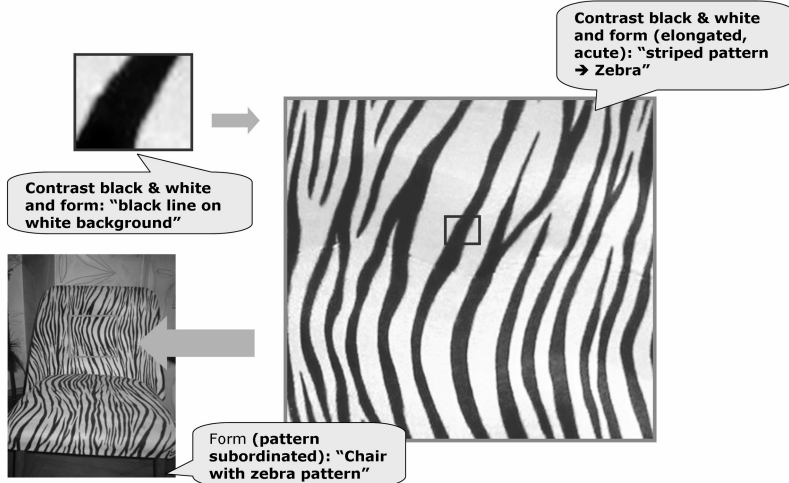
Fig. 5. Colour, form and arrangement evoke certain parts of our knowledge and experience (from Lang, 2005)

Over several decades techniques were developed to group pixels into spectrally similar areas and link spectral classes – if possible – to information classes. The pixel based approach is considered intrinsically limited (Blaschke & Strobl, 2001; Benz et al., 2004; Schöpfer et al., in press), since only spectral properties of (geographic) features are taken into account. The 'picture element' as a technical driven smallest unit integrates signals but does not reflect spatial behavior in a sufficient manner. Even if direct neighborhood is considered by applying kernel-based techniques, the 'environment' remains a square or any other predefined regular geometric shape. Modeling in this case is confined to spectral characteristics and related statistical behavior (texture, etc.). Spatial explicit characteristics are left aside.

The process of OBIA is supported by the use of so-called cognition networks (Binnig et al., 2002) or related concepts of KOS that provides the framework for modeling user-oriented target classes and their composition by spatial components. A cognition network controls the system of target classes and the class definitions as well as the mode of representation (e.g. regionalized segmentation or one-level vs. multi-level representation, see Tiede et al., this volume; Weinke et al., this volume). It provides the basis for a rule-based production system, which is controllable and transferable, as well transparent to both operators and users. Even if there are means and techniques to formalize knowledge and to encapsulate it into rule bases, the vast intuitive knowledge of a skilled interpreter operative for many years is hard to encapsulate in a rule system. Transferring existing experience effectively into procedural and structural knowledge remains a challenge of AI systems, especially taking into consideration the user-oriented plurality of solution, as discussed above. Cognition Network Language (CNL, Baatz et al., this volume), the meta-language to set up rule sets in Definiens software[10] offers a range of tools and algorithms to even address complex target classes. Establishing a rule set is often time-, labor- and therefore cost-intensive. But once a system is set up, and proved to be transferable, the effort pays off. The application of it to like scenes does not imply linear effort, as visual interpretation does. Therefore, in terms of operability one needs to distinguish between establishing a cognition network, and its, mostly scene-depending, parameterization.

OBIA can play a key role for image understanding (Lang & Blaschke, 2006). The entire process of image analysis is characterized by the transformation of knowledge. Finally, a scene description representing the image content should meet the conceptual reality of an interpreter or user. By establishing a body plan for the classes targeted at, knowledge is stepwise

---

[10] See www.definiens-imaging.com

adapted through progressive interpretation and class modeling. By this, knowledge is enriched through analyzing unknown scenes and transferring knowledge will incorporate or stimulate new rules. A particularity with spatial characteristics is related to the notion that parameters for spatial characteristics are difficult to determine. Whereas spectral reflectance can be measured directly by spectrometers on the ground, spatial or hierarchical properties are often ill-defined, less explicit and therefore difficult to be used as distinctive rules. Fuzzification of ranges (e.g. "rather small in size" or "rather compact") is one way to tackle this problem, but it is often not the range as such that is ambiguous, but the very spatial property itself.

Pro-active classification

Operators applying methods of object-based image analysis must be ready for taking over responsibilities. Isn't that contradictory? In most publications we read about 'automated' or at least 'semi-automated' analysis. Automation, of course, is the overall aim of using this approach – like with any other computer-based technique. However, with increasing complexity of the scenes and the classes or features to extract, we need to carefully feed the system with our experience, in a usable form. The approach towards this goal must be interdisciplinary: modeling complex target classes using spatial and structural characteristics not only requires computational skills, but also a wealth of knowledge about the area and the composition of the imaged setting.

Maybe it sounds provokingly, but it may be stated that a standard supervised multi-spectral classification is mechanistic. The machine is fed with samples, which we assume to be correct, and then provides a corresponding result. The process involves a certain level of probability but highly controllable. Class modeling, in contrast, is not mechanistic, but systemic. It not only requires 'supervision', but pro-active engagement from the operator. The combination of existing techniques incorporate know-how and existing experience in the different areas: (1) the modeling stage of the classes relies on expert knowledge that can build upon manual interpretation skills; (2) users that practice pixel-based statistical approaches can utilize their understanding in machine based classifications; (3) experiences in semi-automatically detecting and delineating features on high resolution data can be used for the classification process as such. Therefore object-based methods will integrate the existing remote sensing know how rather than replacing prevailing classification practices.
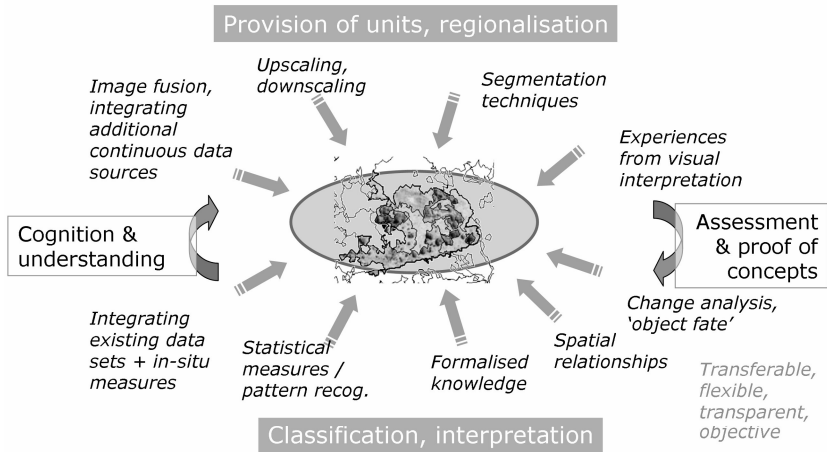
Fig. 6. OBIA as an integrated approach (from Lang, 2005, modified)

4 Object assessment and evaluation

Introducing new ways of image analysis is not free of challenges in terms of adapting both core and complementary methods (Lang & Blaschke, 2006). Among others, this applies for object-based accuracy assessment and object-based change detection. 'Object-based' in this case means that accuracy or changes are assessed in such a way that spatial properties are unambiguously reflected. Both spatial-explicit assessment techniques are facing similar challenge of comparing data sets according to criteria, which are hardly reducible to binary decisions of true or false. Possible solutions are discussed in this volume and preceding studies under different aspects (Schöpfer et al, this volume; Schöpfer & Lang, 2006; Lang et al., in press; Castilla & Hay, 2006; Weinke et al; this volume; Weinke & Lang, 2006). Thus, in the following two sections only key aspects are summarized.

Then, touching briefly at the strands of cognitive psychology, the chapter concludes with a note on object evaluation, for our results need to be proven at the ultimate benchmark, our human perception.

Object-based accuracy assessment

Quantitative site-specific accuracy assessment (Congalton & Green, 1999) using error matrices and specific assessment values (such as error of commission, error of omission, kappa ^) is widely used for evaluating the probability of correct class assignments and the overall quality or reliability of a classification result. Within OBIA, point-based accuracy assessment only gives indication on thematic aspects (labeling). Thematic assessment can be checked by generating random points within objects and comparing the labels against a ground truth layer. Alternatively, a set of objects can be selected in advance and be used as reference information. The decision of being thematically correct may not be a binary one: in fuzzy-based systems, the assessment of class membership is rather qualified by a confidence interval. Still, thematic accuracy assessment may be dubbed '1-dimensional': there is one specific label, i.e. the most likely classification result, to be checked on one specific site. Note that also checking an entire object in terms of its label is a point-based approach[11]. Spatial accuracy instead requires different ways of assessment. There are at least two aspects to be considered: (1) the appropriateness of an object's delineation (match of extend and shape with real situation) and (2) the precision of boundary delineation (match with scale applied).

In smaller test areas with a limited number of larger objects, every single object may be assessed individually: classified image objects can be visually checked against manual delineation (e.g. Koch et al., 2003). Still, a quantitative assessment requires at least some basic GIS overlay techniques. But performing hard intersections implies operational problems of producing sliver polygons and the like. A solution of 'virtual overlay' has been proposed by Schöpfer & Lang (2006) (see also Schöpfer et al., this volume), looking at specific object fate. This comprises object transition (fate in time) and object representation (fate in terms of different ways of delineation). Generally speaking, we encounter difficulties in performing object-based accuracy assessment, which satisfies the needs as being discussed by Congalton & Green (1999): (1) a 100% geometrical fit between reference and evaluation data is usually not given due to the issue of scale and the different ways of delineation; (2) the thematic classes are not mutually exclusive when using fuzzified rule bases. In other words, the accuracy is also a matter of geometrical and semantic agreement (Lang, 2005).

---

[11] This can be considered a special case of ecological fallacy (Openshaw, 1984), or better: individualistic fallacy, as we assume correctness for an entire area based on point-based observation. See also the discussion about polygon heterogeneity Castilla & Hay, 2006.

Object-based change detection and analysis

Monitoring is about detecting, quantifying and evaluating changes. Without proper change assessment, any decision making process is vague, and any management measure to be taken is ill-posed. Changing objects (or geons, see the discussion about fair state above) do not only change in terms of label, but also – and usually more often – in terms of their spatial properties. A habitat under observation may not have changed its class over time, given that is was measured on a specific point located e.g. in the centre of the habitat area, where there have no changes occurred. On the other hand, it may have been substantially shrinked through activities around it. So its function of providing living space for certain organisms may not be further fulfilled. In terms of its spatial component, object-based change detection faces a similar class of problems as object-based accuracy assessment. Common image-to-image or map-to-map comparisons (Singh, 1989) are site-, but not object-specific, i.e. they refer to pixels. Any aggregated measure based on this, becomes spatially implicit. Object-based change analysis needs to specifically compare corresponding objects. Methodological frameworks have been discussed by Blaschke, 2005; Niemeyer & Canty, 2001; Straub & Heipke, 2004; Schöpfer & Lang, 2006; Schöpfer et al., this volume. Like with object-based accuracy assessment, vector overlays (intersections) produce very complex geometry, which is later on difficult to handle and critical in terms of post-processing. Visual inspection is subjective and time-intensive and therefore of limited operational use.

In GIScience there are generic concepts available to describe the spatial relationships among spatial objects (e.g. Mark, 1999 or Hornsby and Egenhofer, 2000) which are built upon sets of spatial relationships. The relationships among spatial objects are built on basic topological principles like containment, overlap, and proximity. In addition to topology, the increase or decrease in size are two further qualities describing temporal dynamics, with presence or absence of objects form the extremes. While these basic categories for describing mutual spatial relationships are conceptually clearly distinguishable, in reality a combination or any kind of transition may occur. The problem is usually reinforced by the effects of generalization. Spatial changes may get completely averaged by applying smaller scales of representations (figure 7).
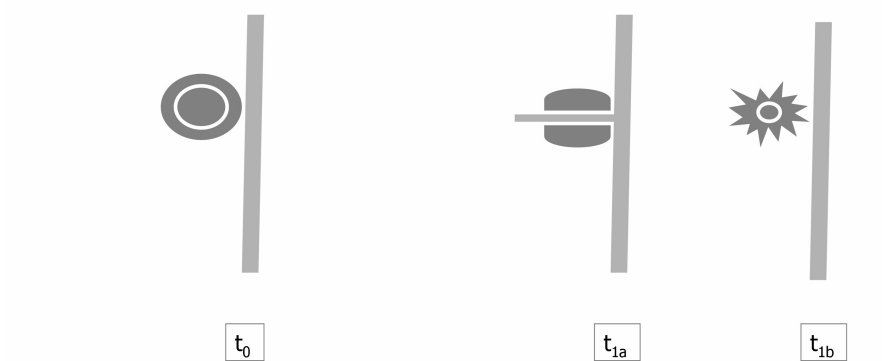
Fig. 7. Substantial changes that may only occur in fine-scaled representation. Left: Habitat next to a road in state t_0. Right: state t_{1b} - the habitat is split by a small road that (1) is not reflected in coarser scale and (2) only slightly decreases habitat area; state t_{1b}: the habitat has been influenced in terms of its boundaries; functional properties have changed substantially (the core area, indicated in white, has shrinked), but area has remained the same.

Object evaluation – the ultimate benchmark

It is recommendable, in the entire process of OBIA, not to forget the ultimate benchmark, i.e. our (visual) perception. The machine is supportive – it reacts on parameters, though the expert has to decide (figure 8).
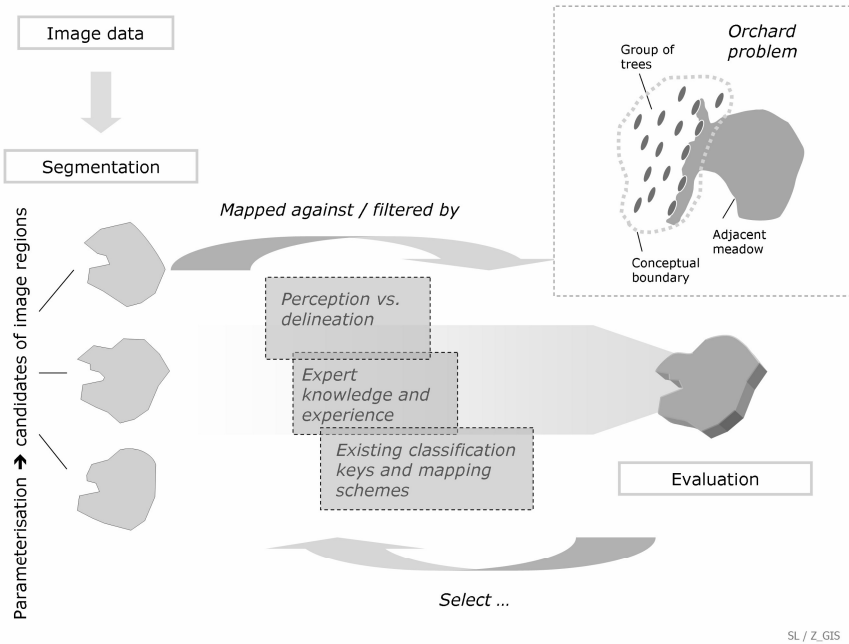
Fig. 8. Object delineation: the expert filter and the orchard problem (from Lang, 2005; Lang & Langanke, 2006, modified)

Region-based segmentation algorithms, like the name indicates, produce regions according to a certain criterion of homogeneity (spectral similarity, compactness, etc.). Due to their bottom-up nature, they are limited in providing delineations of aggregates that consists of high contrast, but regularly appearing (sub-)objects. These kinds of structural arrangements, such as an orchard (Lang & Langanke, 2006) or a mire complex with pools and hummocks (Burnett et al., 2003), are readily delineated by humans, though hard to grasp by a machine. This is a different kind of homogeneity: regularity in structure (repetitive patterns) or conformity (i.e. constancy) in change.

The *orchard problem* and related problems (Lang & Langanke, 2006) arises when addressing geographical features that exhibit conceptual boundaries rather than 'real' ones. Consider an orchard, which is delineated on an aerial-photograph with ease, because of the specific arrangement of fruit trees in a matrix of grass. Even, if the orchard is situated next to a meadow with the same spectral behavior as the matrix, humans would draw the boundary line somewhere in between. Reasons for that can be found in the human way to deal with heterogeneity according to the gestalt

laws (Wertheimer, 1925) and other principles of human vision. Both the factor of good gestalt and the factor of proximity and continuation explain why humans would delineate an object on a higher level. As found out by Navon (1977), a scene is rather decomposed than built-up: if segmentation routine start from pixels, it can hardly mimic the way of visual processing, namely to start from a global analysis of the overall pattern and then to proceed subsequently to finer structures.

The limitations as pointed out above may require coping strategies like the introduction of hybrid techniques. This means in this context the complementary usage of machine-based automated delineation of basic units and high-level aggregation by a human interpreter (see figure 9).

Fig. 9. Delineation of habitat complexes: full automated class modeling vs. hybrid approach.

Recent advances in visual intelligence research have found further explanatory rules for the interpretation of geometrical (or spatial) structures (Hofman, 1998), and some of these provide valuable hints for the shortcomings of the difficulties we are facing when trying to automate the interpretation of complex scenes. There are rules concerning the way how something is interpreted in a constructive way, e.g. how lines in 2D are interpreted in 3D (straight lines, coinciding lines, collinear lines). But when dealing with satellite image data or air-photos these rules are less important, since the perspective is always vertical and requires abstraction, anyway. Others make us favor constellations which likely exist, and neglect theoretical, unlikely ones. This implies utilizing certain 'world views', e.g. the rule of regular views, which excludes some constellations to be very unlikely, is based on the concept of regularity; and this is a matter of experience. The Gestalt law of symmetry, though being powerful in explain-

ing this phenomenon in part, is not capable to cover all cases (ibid.). But subjectively perceived structures based on constructed boundaries can be very complex.

6 Conclusion

This chapter has highlighted 'drivers' for object-based image analysis as well as some of the 'responses' as they became key characteristics of OBIA. The aim of the opening chapter for the first book section asking "Why object-based image analysis?" was to put forward profound ideas and basic concepts of this new approach, as it was to discuss the tasks challenging it. Both motivation and requirements for OBIA were presented in the light of a world of increasing complexity to be addressed by multiple solutions for global monitoring issues. Conceptual elements of this new approach were discussed considering spatial as well as perceptual aspects. Drawing from those, methodological implications have been pointed out in terms of adaptations and further development of traditional methods, empowered by a successful integration of GIS and remote sensing techniques. The subsequent chapters of this section will complement these views with many additional aspects presented from various angles and backgrounds.

Acknowledgements

References

Allen TFH & TB Starr (1982) Hierarchy. University of Chicago Press, Chicago.

Baatz M & A Schaepe (2000) Multi-resolution segmentation. An optimization approach for high-quality multi-scale image segmentation. In: J. Strobl, T. Blaschke & G. Griesebner (eds.), Angewandte Geographische Informationsverarbeitung XII, Wichmann Verlag, Heidelberg, pp. 12-23.

Benz U, P Hofmann, G Willhauck, I Lingenfelder & M Heynen (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-

ready information. ISPRS Journal of Photogrammetry and Remote Sensing, vol. 58, pp. 239-258.

Biederman I (1987) Recognition-by-Components: A Theory of Human Image Understanding. Psychological Review 94(2):115-147.

Binnig G, M Baatz J Klenk & G Schmidt (2002) Will Machines start to think like humans? Europhysics News, vol. 33 (2)

Blaschke T (2005) A framework for change detection based on image objects. In: Erasmi, S. et al. (eds.), Göttinger Geographische Abhandlungen 113, pp. 1-9.

Blaschke T, P Zeil, J Strobl, S Lang, D Tiede, M Möller, G Triebnig, C Schiller, M Mittlböck, B Resch (2007) GMES: From research projects to operational environmental monitoring services. ISPRS Workshop High-Resolution Earth Imaging for Geospatial Information, May 29 – June 1, 2007, Hannover.

Blaschke T & J Strobl (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS, vol. 6, pp. 12-17.

Burnett C & T Blaschke (2003) A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecological Modelling, pp. 233-249.

Burnett C, K Aavikso, S Lang, T Langanke & T Blaschke (2003) Mapping mire forests as hierarchy of objects. In: IUFRO conference "decision support for multiple purpose forestry, 23.-25.4. 2003. Vienna. CD-Rom

Campbell JB (2002) Introduction to Remote Sensing. The Guilford Press, New York.

Castilla, G & G Hay (2006) Uncertainties in land use data. Hydrology and Earth System Sciences, 3, 3439-3472.

Castilla GA Lobo & J Solana (2004) Size-constrained region merging (SCRM): a new segmentation method to derive a baseline partition for object-oriented classification. Proceedings of SPIE, 5239, pp. 472-482.

Congalton RG & K Green (1999) Assessing the accuracy of remotely sensed data: principles and practices. Lewis Publishers, Boco Raton.

Egenhofer MJ (1994) Pre-Processing queries with spatial constraints. Photogrammetric Engineering & Remote Sensing, 60 (6), pp. 783-790.

Ganter B & R Wille (1996) Die Formale Begriffsanalyse. Springer, Berlin.

Groom G, CA Mücher, M Ihse, T Wrbka (2004) Remote sensing in landscape ecology. Landscape Ecology, pp. 1-18.

Hay G, T Blaschke, DJ Marceau & A Bouchard (2003) A comparison of three image-object methods for the multiscale analysis of landscape structure. Photogrammetry and Remote Sensing, vol. 1253, pp. 1-19.

Hofmann DD (1998) Visual Intelligence. How we create what we see. W. W. Norton & Company, Inc.

Hornsby K & MJ Egenhofer (2000) Identity-based change: a foundation for spatio-temporal knowledge representation. International Journal of Geographical Information Science, 14 (3), pp. 207-224.

Ibrahim AE (2000) An intelligent framework for image understanding. Available at: http://www.engr.uconn.edu/~ibrahim/publications/image.html (May, 2004).

Koch B, M Jochum, E Ivits & M Dees (2003) Pixelbasierte Klassifizierung im Vergleich und zur Ergänzung von pixelbasierten Verfahren. Photogrammetrie Fernerkundung Geoinformatik, vol. 3, pp. 195-204.

Lang S & D Tiede (2007) Conditioned Information For GMES Applications. Presentation at the Definiens GMES Research Award 2007, March 29.

Lang S & T Blaschke (2006) Bridging remote sensing and GIS - what are the most supportive pillars? In: S Lang, T Blaschke & E Schöpfer (eds.): International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42.

Lang S & T Langanke (2006) Object-based mapping and object-relationship modeling for land use classes and habitats. In: Photogrammetrie, Fernerkundung, Geoinformation, 1/2006, 5-18

Lang S (2005) Image objects vs. landscape objects. Interpretation, hierarchical representation and significance, Salzburg, (unpublished PhD thesis).

Lang S (2002) Zur Anwendung des Holarchiekonzeptes bei der Generierung regionalisierter Segmentierungsebenen in höchstauflösenden Bilddaten. In: T. Blaschke (ed.), Fernerkundung und GIS: Neue Sensoren – innovative Methoden, Wichmann Verlag, Heidelberg, pp. 24-33.

Laszlo E. (1972) The Systems View of the World. George Braziller, New York.

Liedtke CE, J Bückner, O Grau, S Growe & R Tönjes (1997). AIDA: a system for the knowledge based interpretation of remote sensing data. Third International Airborne Remote Sensing Conference, 7-10 July 1997, Copenhagen, Denmark, Copenhagen.

Navon EJ (1977) Forest before trees: the precendence of global features in visual perception. Cognitive Psychology, vol. 9, pp. 353-383.

Niemeyer I & MJ Canty (2001) Object-oriented post-classification of change images. Proceedings of SPIE's International Symposium on Remote Sensing, Toulouse, 17-21 September 2001, vol. 4545, pp. 100-108.

Openshaw S (1984) The modifiable areal unit problem, Concepts and Techniques in Modern Geography 38: 41.

Pinz A. (1994) Bildverstehen. Springer, Wien.

Schöpfer E. & S Lang (2006) Object fate analysis - a virtual overlay method for the categorisation of object transition and object-based accuracy assessment. In: S Lang, T Blaschke & E Schöpfer (eds.): International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42.

Schöpfer E., S Lang & J Strobl (in press) Segmentation and object-based image analysis. In: C. Juergens & T. Rashed (eds.), Remote sensing of urban and suburban areas, Kluwer, Amsterdam. pages pending.

Schumacher J, S Lang, D Tiede, J Rietzke, J Trautner & D. Hölbling (2007) Einsatz von GIS und objekt-basierter Analyse von Fernerkundungsdaten in der regionalen Planung: Methoden und erste Erfahrungen aus dem Biotopinformations-und Managementsystem (BIMS) Region Stuttgarts. In: Angewandte Geografische Informationsverarbeitung XIX, pages pending.

Simon HA (1973) The Organization of Complex Systems. In: H.H. Pattee (ed.), Hierarchy Theory. The Challenge of Complex Systems, George Braziller, New York, pp. 1-28.

Singh A (1989) Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing, 10, pp 989-1003.

Sowa JF (1999) Knowledge representation: logical, philosophical, and computational foundations. Brooks Cole Publishing Co, Pacific Grove.

Strahler AH, CE Woodcock & J Smith (1986) Integrating per-pixel classification for improved land cover classification. Remote Sensing of the Environment, vol. 71, pp. 282-296.

Straub BM & C Heipke (2004) Concepts for internal and external evaluation of automatically delineated tree tops. International Archives of Photogrammetry and Remote Sensing, vol. 26 (8/W2), pp. 62-65.

Szaramovicz M (2004) "Hierarchy" und "Scale" - Interessante Denkmuster für die Landschaftsplanung?. Naturschutz und Landschaftsplanung, 36 (1), 21-26.

Tiede D & S Lang (2007) Analytical 3D views and virtual globes - putting analytical results into spatial context. ISPRS Joint Workshop Visualization and Exploration of Geospatial Data, June 27-29, Stuttgart.

Tiede D & Ch. Hoffmann (2006) Process oriented object-based algorithms for single tree detection using laser scanning data. EARSeL-Proceedings of the Workshop on 3D Remote Sensing in Forestry, 14th-15th Feb 2006, Vienna.

Tiede D, S Lang & C.Hoffmann (2006) Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. In: S. Lang, T. Blaschke and E. Schöpfer (eds.): International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42.

Wertheimer M (1925) Drei Abhandlungen zur Gestaltpsychologie. Verlag der Philosophischen Akademie, Erlangen.

Wu J (1999) Hierarchy and scaling: extrapolating information along a scaling ladder. Canadian Journal of Remote Sensing, vol. 25 (4), pp. 367-380.

Zeil P, S Lang, E Schöpfer, D Tiede, T Blaschke (in press) Monitoring for Human Security – reflecting the ‚S' in GMES, EARSeL Workshop GMES, Bolzano, 7-8 June, 2007.

Chapter 1.2

Progressing from object-based to object-oriented image analysis

M. Baatz [1], C. Hoffmann[1], G. Willhauck [1]

[1] Definiens AG, Munich, Germany;
{mbaatz, choffmann, gwillhauck}@definiens.com

ABSTRACT: This research describes an advanced workflow of an object-based image analysis approach. In comparison to the existing two-staged workflow where typically a segmentation step is followed by a classification step, a new workflow is illustrated where the objects themselves are altered constantly in order to move from object primitives in an early stage towards objects of interest in a final stage of the analysis. Consequently, this workflow can be called "object-oriented," due to the fact that the objects are not only used as information carriers but are modelled with the continuous extraction and accumulation of expert knowledge. For better demonstration, an existing study on single tree detection using laser scanning data is exploited to demonstrate the theoretical approach in an authentic environment.

1 Introduction

Recent developments in remote sensing made it possible to obtain data of a very high spatial resolution which allows extraction, evaluation, and monitoring of a broad range of possible target features. At the same time, the demand to automate image analysis in operational environments is constantly growing. However, the variety and number of different features to

be extracted, add challenges specifically in terms of modelling and auto-adaptive procedures.

The advantage of a spatial resolution with pixel sizes significantly smaller than the average size of the object of interest comes with the disadvantage of an abundance of spatial detail and the accordingly huge amount of data to be processed. To overcome this drawback, the object-based image analysis approach has proven to be an alternative to the pixel-based image analysis and a large number of publications suggest that better results can be expected (Baatz and Schäpe 2000, Willhauck et al. 2000, Hay et al. 2005, Kamagata et al. 2005, Manakos et al. 2000, Whiteside et al. 2005, Yan et al. 2006).

The object-based approach suggests a two-staged approach. In the first step pixels are merged to object clusters, possibly in a multi-level object hierarchy, which then will be analysed and classified in the second step. This means that, the created objects influence the classification result to a large extent although they might not represent the final objects of interest (i.e. single buildings, trees, etc.) already. Because the objects remain unchanged once they are created, and subsequently serve as basis for the actual analysis, this workflow can be called "object-based image analysis". A successful object-based image analysis results in the correct labelling / classification of regions rather than extracting final objects of interest for instance like trees, acres, buildings or roads in their final shape.

In comparison to the "object-based" workflow, this paper describes an alternative, more advanced workflow which not only uses object clusters as the basis for a classification analysis but brings the objects themselves and the shaping of the objects in the focus of the analysis.

This alternative workflow starts with creating object clusters and aims to produce desired objects of interest with correct shape and correct labelling. Why is this required? The accuracy and the significance of the final measurements, numbers, and statistics directly and actually critically depend on the quality of segmentation. Relevant information such as numbers, shapes or other statistics per unit is only accessible if trees are not only correctly labelled as "tree area" but also are correctly extracted tree by tree as "tree objects" or "tree units".

Typically, the correct extraction and shaping of objects of interest requires more advanced models, domain knowledge and semantics, in order to cope with the specific characteristics of the structure and to sort out ambiguities that often occur. The more or less simple and knowledge-free segmentation procedures used to produce object clusters or object primitives almost never succeeds in extracting objects of interest in a robust and reliable manner. Furthermore, different types of target objects also need different strategies for their extraction. In order to support this, decisions

need to be made throughout the process that classifies different regions and thus make them accessible to different extraction approaches.

2 Methodology

2.1 Object-oriented image analysis workflow

The workflow described here starts with object primitives as well. However, in contrast to the object-based workflow, it uses these objects not only as information carriers but also as building blocks for any further shape modification, merging, or segmentation procedures. In a whole sequence of processing steps, where segmentation steps alternate with evaluation and classification steps, these object primitives are constantly altered until they become the desired objects of interest. Because this strategy aims for correct shaping and classification of objects, and objects are also used at every step during the procedure as the central processing unit, serving both as information provider and as building blocks, this workflow consequently can be called "object-oriented image analysis".

This approach can be realised by using Cognition Network Technology (CNT), distributed by Definiens AG, with its included modular programming language CNL (Cognition Network Language). Amongst typical programming tasks like branching, looping and variable definition, CNL enables to build and perform specific analysis tasks based on hierarchical networks of objects and essentially supports an object-oriented image analysis workflow.

This workflow can be described best with a spiral and is illustrated in Fig. 1. The entire process is iterative and starts in the first step with the creation of object primitives using any (knowledge-free) segmentation algorithm. The next step uses the object primitives in order to perform a first evaluation and classification, thus introducing semantics. Building on this result, the subsequent step allows refinement or improvement of the segmentation locally for a specific class. Thus, the whole process alternates iteratively between local object modification on the one hand and local object evaluation and classification on the other. By using such an iterative approach, different object classes can be addressed with different object modification strategies.

During this process, the objects are altered from stage to stage until they represent the final target objects. Only the combination of correct shaping and correct classification characterizes the target objects; otherwise the results will be insufficient. The spiral in Figure 1 represents this alternation between segmentation and classification in the object-oriented approach.

As the analysis progresses, classification detail and accuracy are growing together with segmentation detail and accuracy. Whereas the process starts with rather simple and knowledge-free segmentation steps, more and more expert and domain knowledge is introduced and used in later steps. The more closely the process approximates the final target objects, the higher is the abstraction from the original image information.
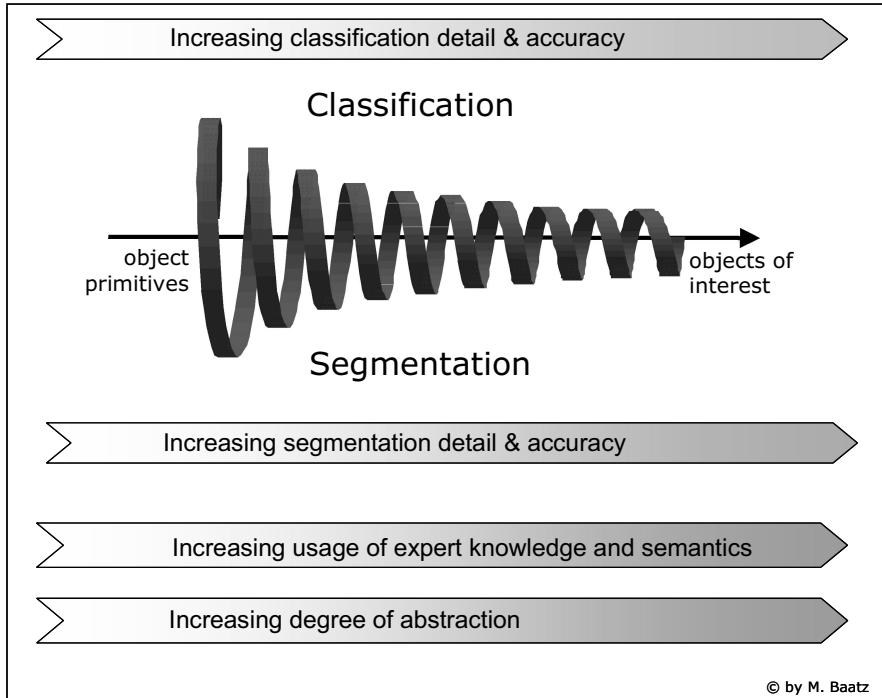
Fig. 1. Object-oriented image analysis: the generic procedure

2.2 The Object Domain

A central concept in the object-oriented workflow is the Object Domain. It defines for each algorithm the specific subset of objects to which the algorithm—independent if segmentation or classification—will be applied. It therefore enables implementation of locally specific strategies.

The Object Domain characterizes a subset of objects in the object hierarchy through the hierarchical object level, the classification of objects and/or specific attributes of objects.

The concept can be further differentiated. Starting with objects of a specific Object Domain, subdomains can be defined operating over the networked neighbourhood. In a hierarchical object network, subdomains of an object can for instance be neighbour objects, subobjects or a superobject in a defined distance, with a specific classification and/or with specific attributes.

In the continuously alternating object-oriented image analysis workflow, the Object Domain is the essential link between segmentation and classification.

2.3 Evaluation and classification aspects

During the course of the image analysis, localized evaluation and classification is essential. Before objects can be distinguished into different types by classification, object attributes must be evaluated. These attributes can be intrinsic to the object—such as shape, size or spectral characteristics— or they can be derived from operations over the networked context of an object. Contrasts, embedding, relative location, and composition are good examples.

Independent of the individual classification method it always can be applied to a specific subset of objects defined through an Object Domain. Thus very specific measurements and decisions can be performed locally.

If the evaluation is more complex, intermediate results can be stored in variables. Whereas object variables hold information specific to an individual object global variables hold information specific to the whole scene.

Making decisions that relate to measurements stored in variables is an important tool for auto-adaptive strategies which play a crucial role in making solutions robust over any expected variability of object characteristics. As mentioned before, the attributes derived from objects and used for classification critically depend on the way the object has been processed.

2.4 Segmentation aspects

The term "Segmentation" is used here as the summary of all procedures that build, modify, grow, merge, cut or shrink objects. In principal, segmentation techniques can be distinguished depending on if they are used to produce initial object primitives starting with pixels (A) or if they further process already existing objects (B). Typically, an object oriented image analysis process starts with (A) and performs all further object modification with (B).

Independent of the specific algorithm the modification (B) can be applied to a specific subset of objects defined through an Object Domain. Thus locally very specific modifications can be performed.

A number of segmentation methods are working with seed objects as a starting point and candidate objects in the networked neighbourhood which are used to modify (merge, grow, shrink, cut) the seed unit. In these cases, both types – seeds and candidates – can be defined through different Object Domains in order to constrain the algorithm and make it more selective.

2.5 Fractal structure of the workflow

The described workflow of alternating segmentation and classification steps is inherently fractal.

In many cases, a specific segmentation step needs preparation in form of an evaluation and a classification of objects which shall be modified or which contribute to a specific modification. A good example for this is the object-oriented watershed algorithm described in the case study below.

Symmetrically, image objects are not always directly in the appropriate state to provide the relevant information needed to do a certain decision. In theses cases preparation is needed through an adequate segmentation of objects. For instance, if the size of water bodies matters and at the current state of processing water bodies are represented through correctly classified object primitives then merging those primitives into complete water body units allows to access the needed information.

Fig. 2 illustrates how subprocesses themselves show the alternation of segmentation and classification.
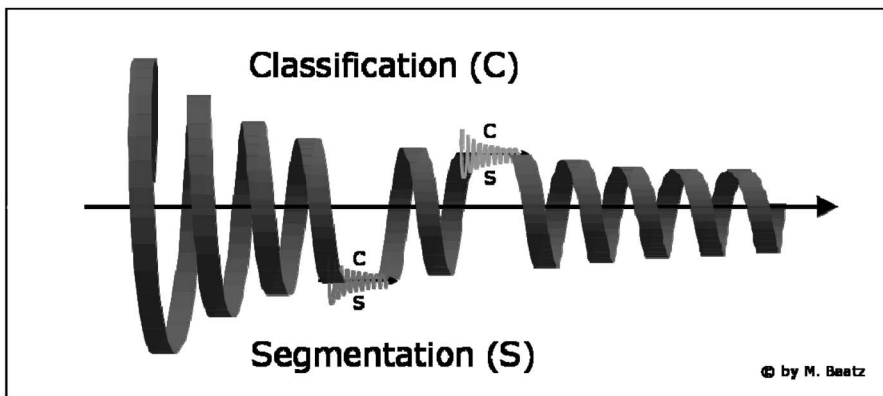
Fig. 2. Fractal approach of the object-oriented image analysis workflow

2.6 Modular Structure

Typically, an overall analysis procedure consists of a number of sub modules that each addresses a certain target class. Each individual module is an encapsulated object-oriented image analysis task. Fig. 3 shows an example for a sequential modular approach.
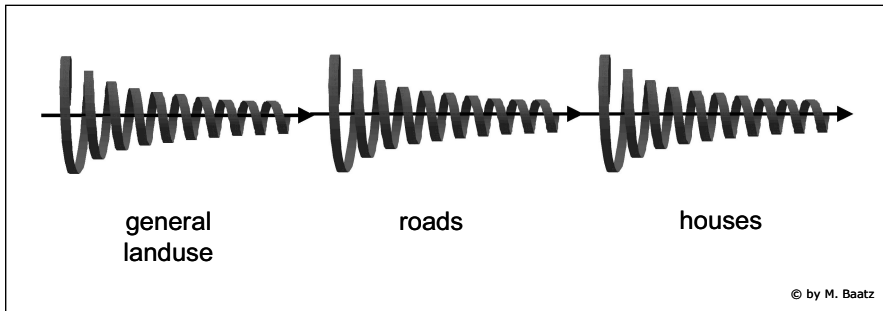
Fig. 3. Modular Structure of the object-oriented image analysis workflow

3 Case study – single tree detection

Individual object-based tree detection has been discussed in a number of publications as well as the use of the local maxima approach (Pitkänen et al. 2001, Pitkänen et al. 2004, Tiede et al. 2004, Tiede et al. 2005).

The following case study exemplarily demonstrates an object-oriented image analysis workflow. It is a solution for single tree detection using airborne laser scanning data and was carried out by Tiede and Hoffmann (2006).

Starting with object primitives the objects are constantly evaluated and altered until the target objects in form of tree units are found. The approach used can be called an object-oriented and knowledge-based watershed algorithm.

Fig. 4 gives an overview over the workflow as described in detail in the following chapters.
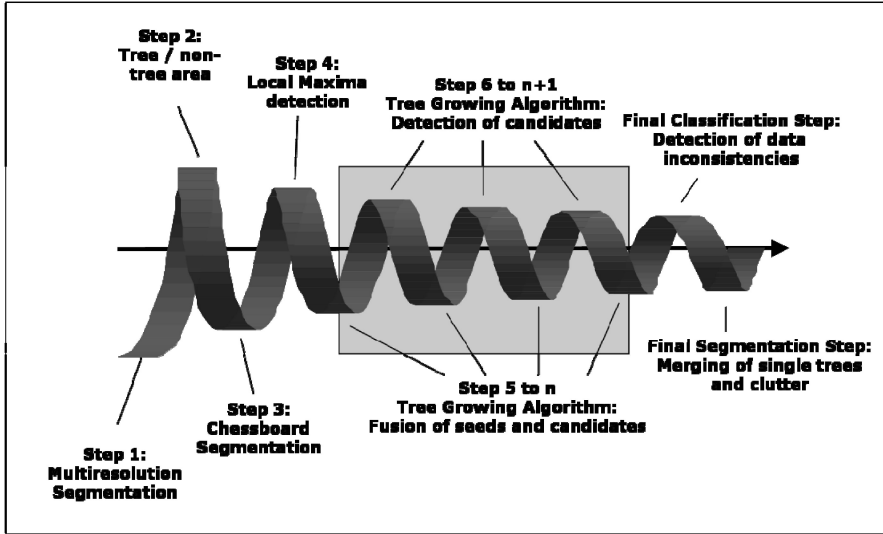
Fig. 4. Single tree detection workflow in the context of object-oriented image analysis. The tree growing algorithm consists of a number of iterations itself

3.1 Used data

Fig. 5 shows the used data derived from a digital surface model (DSM) of an airborne laser scanning dataset with a ground resolution of 0.5 meters.
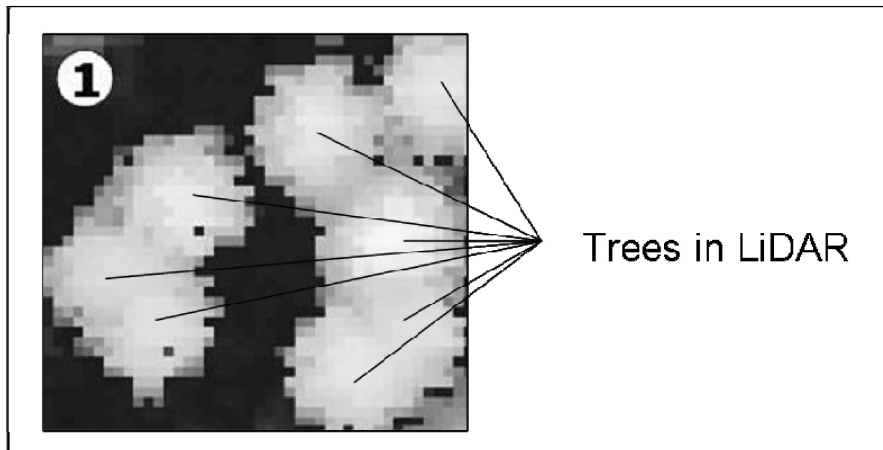
Fig. 5. Crown Model from airborne laser scanning. In this small subset, eight trees can be seen represented by high (bright) values. Tiede & Hoffmann (2006), edited

3.2 Tree / non-tree area

The first step is to distinguish the background from the tree area. It turns out that object primitives best suited to provide information for this decision can be created with Multiresolution Segmentation. Theses primitives might each cover a number of trees but they are small enough to reliably separate tree area from background.

The subsequent classification step distinguishes areas which are elevated and, thus, potentially represent tree crowns from the background. All further processing steps will build on this basic decision and use the tree area as the Object Domain.
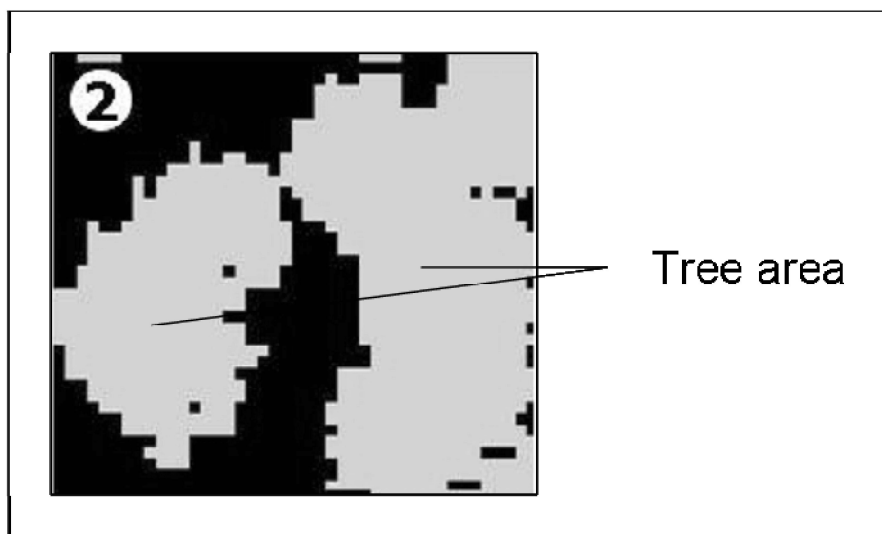
Fig. 6. Step 1: Multiresolution Segmentation and Classification of tree / non tree area. Tiede and Hoffmann (2006), edited

3.3 Local maxima

The model assumes that using a digital surface model a tree can be extracted starting with its tree top working as a seed point.

The currently available object primitives are by far too large to support this analysis. In order to achieve the needed fine granularity all objects in the Object Domain "tree area" are cut down to single pixels using a Chessboard segmentation. Thus, the region "tree area" remains unchanged; however, it now consists of small pixel-sized object primitives.

Now all object primitives that represent a local elevation maximum are classified within the tree area. A classification step operating on the Object Domain "tree area" compares the elevation of each object with that of its neighbours in a certain distance and if it is higher than all others it classifies the object as local maximum (Fig. 7).
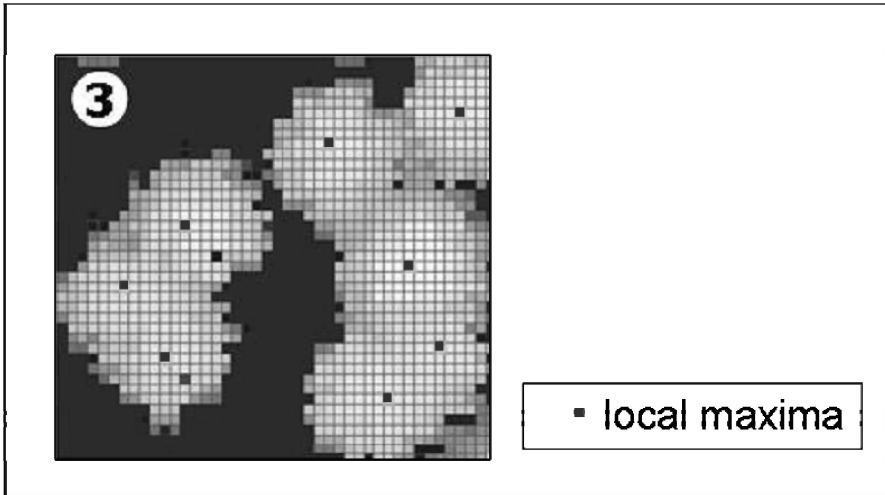
Fig. 7. Step 2: Domain-based break-down of tree area into pixel-sized objects and application of local maxima algorithm. Each maximum represents the highest point of a single tree. Tiede and Hoffmann (2006), edited

3.4 Tree building algorithm

Fig. 8 shows one step within the growing process of the single trees. This growing process is done by a simultaneous merging procedure of the neighbour objects of the already existing tree tops which work as "seeds".

Each tree is grown by using contextual information (similar height) to decide whether the neighbourhood objects are belonging to the currently growing tree. Stop criteria are used to prevent the trees from growing excessively or in an unwanted direction. Since the growing of the trees is done for all trees simultaneously, the trees can only grow until they reach another tree.
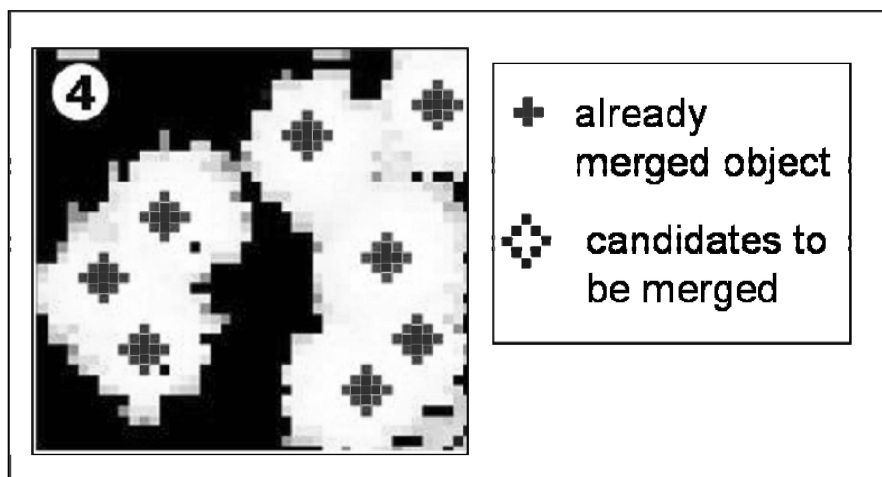
Fig. 8. Step 3: Growing algorithm around tree seeds by similarity. If surrounding objects are matching, they are merged. The growing algorithm itself consists of iterations until the final tree objects are found. Tiede and Hoffmann (2006), edited

The tree growing loop is an example for a sub-procedure which as well can be thought of as a spiral. In each iteration of the loop, there is a classification and a merging step. First, all appropriate neighbour objects are classified as candidates whom are going to be merged with the current tree object. In the second step, these classified objects are then merged with the currently processed tree object to form one large tree object.

Since the number of iterations is defined by a stop criteria and the growing of other trees, the number of iterations is not fixed. It will continue as long as candidate objects exist which can be merged with the growing tree objects.

3.5 Result of the growing procedure

After the tree growing loop, single tree objects exist which are almost representing the desired objects of interest. Fig. 9 shows the result of the tree growing loop. Because of data inconsistency of the laser scanning data, individual pixels can be found with no data or values that are not fitting. These pixels can be removed in a last refinement step, where context information is used again to remove these data errors.

It is important to approximate the tree objects as good as possible to the shape of naturally looking trees because in a potential analysis of the trees

in a later stage, parameters like the tree crown diameter might be important for any analysis.
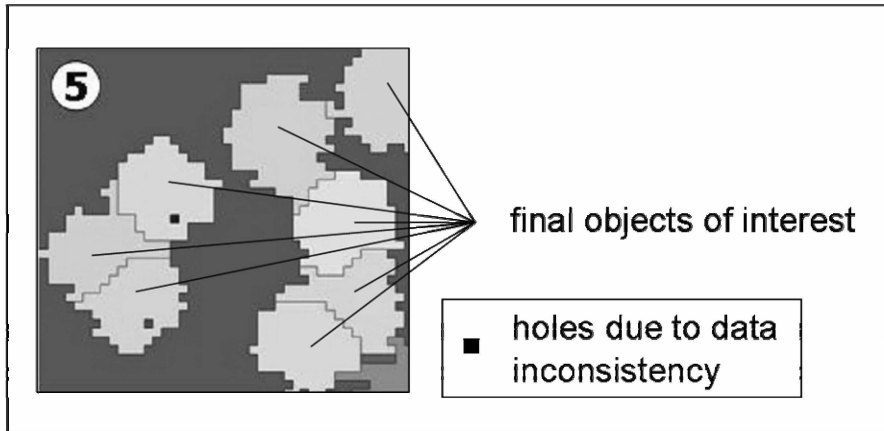
Fig. 9. Result after the tree growing algorithm is finished. Holes due to data inconsistency are still existing. Tiede and Hoffmann (2006), edited

3.6 Remove data inconsistencies

Fig. 10 shows the result after removing data inconsistencies. According to the theoretical workflow shown in Fig. 1, at the end of the alternating segmentation and classification steps, the final objects of interest are modelled in terms of their shape as well as their classification.
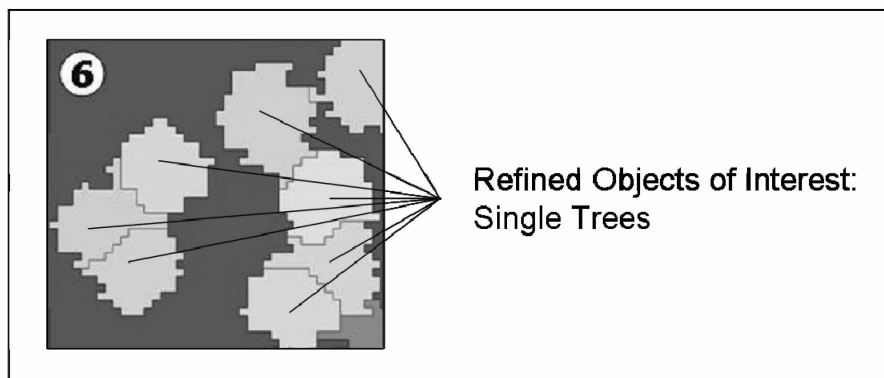
Fig. 10. Final result for single tree detection. Holes were removed by the use of context information. Tiede and Hoffmann (2006), edited

4 Discussion

Considering the complexity and often ambiguities of image analysis problems it is clear, that in most cases more complex models and semantics are needed. The spiral process turns out to be a very efficient paradigm for supporting more complex analysis problems which include modelling, semantics or deal with ambiguities. The object-oriented workflow has proven in many applications (both in Earth and Life Sciences) to be useful to extract objects of interest in an automated manner which is not supported by the object-based approach. Not only if objects of interest are to be extracted but also if only a correct labelling of regions is requested a pure object-based approach is often limited.

In the spiral which defines the object-oriented approach, each step builds on the results of the previous. There is a mutual dependency between segmentation and classification: the attributes and the quality of attributes used to evaluate and classify image objects directly depend on the objects and how they were formed before. The precise formation of objects on the other hand needs specific semantics, models or knowledge how to sort out ambiguities. Since local classification and local segmentation interact and mutually depend on each other in manifold ways through this process, they can be described in a colloquial sense as the Yin and Yang of object-oriented image analysis.

In other words, the object-oriented image analysis workflow overcomes a commonly known problem which can be called the "hen and egg problem" in image segmentation: for a successful classification of a certain feature class, the object-primitives need to exist already in a form at least very near, sometimes even identical to the final target objects. On the other hand, in order to achieve such objects, local evaluation and semantics provided by classification are needed during the segmentation process itself. For that reason, the workflow suggests a step-wise approximation from object primitives with a coarser classification and shape in an early stage to objects of interest and the according detailed classification in the final stage.

5 References

Baatz M and Schäpe A (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Angewandte Geographische Informationsverarbeitung XII, Strobl J, Blaschke T, Griesebner G (eds), Wichmann, Heidelberg, Germany, pp 12-23.

Hay GJ, Castilla G, Wulder MA, Ruiz JR (2005) An automated object-based approach for the multiscale image segmentation of forest scenes. In: International Journal of Applied Earth Observation and Geoinformation, Vol. 7, pp 339-359.

Kamagata N, Akamatsu Y, Mori M, Qing Li Y, Hoshinoy Y, Hara K (2005) Comparison of pixel-based and object-based classifications of high resolution satellite data in urban fringe areas. In: Proceedings of the 26th Asian Conference on Remote Sensing. Hanoi, Vietnam. 7–11 November 2005.

Manakos I, Schneider T, Ammer U (2000) A comparison between the ISODATA and the eCognition classification methods on basis of field data . In: ISPRS, Vol. XXXIII, Amsterdam, 2000.

Pitkänen J, Maltamo M, Hyyppä J, Wei Yu X (2004) Adaptive methods for individual tree detection on airborne laser based canopy height model. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-8/W2, Freiburg, pp 187-191.

Pitkänen J (2001) Individual tree detection in digital aerial images by combining locally adaptive binarization and local maximal methods. Canadian Journal of Forest Research, Vol. 31, pp 832-844.

Tiede D, Burnett C, Heurich M (2004) Objekt-basierte Analyse von Laserscanner- und Multispektraldaten zur Einzelbaumdelinierung im Nationalpark Bayerischer Wald. In: Strobl S, Blaschke T, Griesebner G (eds), Angewandte Geoinformatik 2004, Wichmann, Heidelberg, Germany, pp 690–695.

Tiede D, Hochleitner G, Blaschke T (2005) A full GIS-based workflow for tree identification and delineation using laser scanning. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, Part 3/W24, Vienna, pp 9 – 14.

Tiede D, Hoffmann C (2006) Process oriented object-based algorithms for single tree detection using laser scanning data. In: EARSeL-Proceedings of the Workshop on 3D Remote Sensing in Forestry, 14th-15th Feb 2006, Vienna, pp 151-156.

Whiteside T, Ahmad W (2005) A comparison of object-oriented and pixel-based classification methods for mapping land cover in northern Australia. Proceedings 2005. Melbourne: Spatial Sciences Institute. ISBN 0-9581366-2-9.

Willhauck G, Schneider T, De Kok, R Ammer U (2000) Comparison of object oriented classification techniques and standard image analysis for the use of change detection between SPOT multispectral satellite images and aerial photos . In: ISPRS, Vol. XXXIII, Amsterdam, 2000.

Yan G, Mas JF, Maathuis BHP, Xiangmin Z, Van Dijk PM (2006) Comparison of pixel-based and object-oriented image classification approaches–a case study in a coal fire area, Wuda, Inner Mongolia, China. International Journal of Remote Sensing. Volume 27, Number 18 / 20 September 2006, 4039–4055.

Chapter 1.3

An object-based cellular automata model to mitigate scale dependency

D. J. Marceau[1], N. Moreno[1, 2]

[1] Geocomputing Laboratory, Department of Geomatics Engineering, University of Calgary, Canada, marceau@geomatics.ucalgary.ca

[2] Centro de Simulación y Modelos, Facultad de Ingeniería, Universidad de Los Andes, Venezuela

KEYWORDS: Scale dependency, raster-based cellular automata model, vector-based raster-based cellular automata model, land-use change

   ABSTRACT: Cellular automata (CA) are individual-based spatial models increasingly used to simulate the dynamics of natural and human systems and forecast their evolution. Despite their simplicity, they can exhibit extraordinary rich behavior and are remarkably effective at generating realistic simulations of land-use patterns and other spatial structures. However, recent studies have demonstrated that the standard raster-based CA models are sensitive to spatial scale, more specifically to the cell size and neighborhood configuration used for the simulation. To mitigate cell size dependency, a novel object-based CA model has been developed where space is represented using a vector structure in which the polygons correspond to meaningful geographical entities composing the landscape under study. In addition, the proposed object-based CA model allows the geometric transformation of each polygon, expressed as a change of state in part or in totality of its surface, based on the influence of its respective neighbors. The implementation and testing of this model on real data reveals that it generates spatial configurations of landscape

patches that are more realistic than the conventional raster-based CA model.

1 Introduction

Scale dependency is inherent to the study of any geographical phenomena and refers to the sensitivity of the results to the spatial units used to conduct such studies. While it has been largely documented in social and natural sciences over the last five decades, it still remains a challenge in most scientific investigations involving the use of spatial units for analysis and modeling of geographical/environmental phenomena. With the proliferation of spatial data acquisition and analysis technologies, such as remote sensing and geographic information systems (GIS) as well as environmental models, the issue of scale dependency became more acute and scientists started to explore potential solutions to mitigate the problem. Among the proposed solutions are scale sensitivity analyses, multi-scale analyses, and vector- or object-based approaches.

This paper focuses on scale dependency that manifests itself in the context of a relatively new environmental modeling approach called cellular automata (CA). Cellular automata are individual-based spatial models increasingly used to simulate the dynamics of natural and human systems and forecast their evolution. In CA models, space is represented as a matrix of regular cells having a state value that evolves based on transition rules applied at each time step of the simulation. These rules dictate how each cell might change state based on the state configuration present in its neighborhood. Despite their simplicity, CA models can exhibit extraordinary rich behavior and are remarkably effective at generating realistic simulations of land-use patterns and other spatial structures. However, recent studies have demonstrated that raster-based CA models are sensitive to spatial scale, more specifically to the cell size and neighborhood configuration used for the simulation.

The main objective of this paper is to describe a novel vector-based geographic CA model (VecGCA) that has been developed to mitigate scale dependency. With VecGCA, space is represented using a vector structure in which the polygons correspond to meaningful geographical entities composing the landscape under study. In addition, the proposed object-based CA model allows the geometric transformation of each polygon, expressed as a change of state in part or in totality of its surface, based on the influence of its respective neighbors. The implementation and testing of this model on real data reveals that it generates spatial configurations of

landscape patches that are more realistic than the conventional raster-based CA model, while eliminating the necessity of conducting a sensitivity analysis to identify the appropriate cell size for the modeling of the landscape.

The remaining of the paper is organized into two main sections. Section 2 presents an overview of the numerous studies conducted over the last five decades addressing the issue of scale dependency in a variety of geographical contexts, with a description of solutions that have been proposed to mitigate the problem. In Section 3, the vector-based CA model is fully described with its application for land-use modeling of two landscapes of various complexity. The simulation results are compared with those obtained with a raster-based CA model. The conclusion highlights the advantage of an object-based approach over the conventional raster-based approach in CA modeling as well as some challenges that remain to be addressed.

2 Scale dependency in spatial analysis and modeling

Scale dependency, and more specifically spatial scale dependency, refers to the fact that data collected from spatial units and by extension analysis results derived from those data are dependent upon these spatial units. It is an inherent property of geographic phenomena; if a geographic phenomenon under study varies with scale, it is considered scale-dependent (Cao and Lam, 1997). Scale dependency has been largely investigated in the social and natural sciences over more than half a century with the pioneer studies of Gehlke and Biehl (1934), Yule and Kendall (1950), McCarthy et al. (1956), Kershaw (1957), Usher (1969), Clarke and Avery (1976), and the seminal work of Openshaw (1977, 1978, 1984) on the Modifiable Areal Unit Problem (MAUP) (see Marceau, 1999 for an exhaustive review). In social sciences, these early studies illustrated the sensitivity of traditional statistical analysis to scale and were followed by a series of others also demonstrating the severity of the scale dependency in location-allocation modeling (Bach, 1981), spatial interaction modeling (Putman and Chung, 1989), multivariate analysis (Fotheringham and Wong, 1991), and principal axis factor analysis (Hunt and Boots, 1996). In natural sciences, the studies first focused on acknowledging the existence of *natural scales* at which ecological processes and physical attributes occur within the landscape. Later, scientists conducted numerous studies to investigate the impact of scale on the analysis of landscape (Meetemeyer and Box, 1987; Turner et al. 1989;

Bian and Walsh, 1993; Moody and Woodcock, 1995, Benson and McKenzie, 1995; Qi and Wu, 1996; Kok and Veldkamp, 2001; Wu, 2004). This work gave rise to the concepts of *domain of scale* and *scale threshold* (Meetemeyer, 1989; Wiens, 1989), *scaling* (Jarvis, 1995), and the application of *hierarchy theory* (Allen and Star, 1982; O'Neill et al., 1986) and the *hierarchical patch dynamics paradigm* (Wu and Loucks, 1995) as frameworks to describe how spatio-temporal heterogeneity, scale and hierarchical organization influence the structure and dynamics of ecological systems.

Spatial scale dependency also became largely manifest with the proliferation of remote sensing images at various resolutions and spatial analytical tools offered by GIS (Geographic Information Systems). A series of studies conducted to assess the effects of spatial resolution on the ability to classify land-cover/land-use types using digital remote sensing techniques demonstrated that a change in spatial resolution could significantly affect classification accuracies (Sadowski et al., 1977; Latty and Hoffer, 1981; Markham and Townshend, 1981; Ahern et al., 1983; Irons et al., 1985; Cushnie, 1987). In 1992, Marceau undertook a review encompassing two decades of thematic mapping studies involving remotely-sensed imagery, and revealed that within a given classification scheme of several land-cover/land-use types, considerable inconsistencies in the results obtained from one class to another were typically displayed. Searching for an explanation, Marceau hypothesized that a parallel exists between such inconsistency in the classification results and the use of arbitrarily-defined spatial areal data known as the MAUP (Openshaw, 1984). In a latter study, Marceau et al. (1994) conducted an empirical investigation to verify the impact of spatial resolution and aggregation level on classification accuracy of remotely-sensed data. Their results indicated that per-class accuracies were considerably affected by changing scale and aggregation level, which led to the conclusion that remote sensing data are not independent of the sampling grid used for their acquisition. They also noted that there is no unique spatial resolution appropriate for the detection and discrimination of all geographical entities composing a complex natural scene, and advocated that classification based on the use of a unique spatial resolution should be replaced by a multi-scale approach (Marceau and Hay, 1999).

Further investigations were conducted to evaluate how sensor spatial resolution and data scale used when performing GIS spatial analysis affect the relationship between land surface attributes. As an example, studies have illustrated that the derivation of certain terrain attributes such as slope and aspect was greatly affected by the spatial resolution of the digital elevation model (DEM) used to extract these attributes (Chang and Tsai,

1991; Gao, 1997). Other studies have established the scale dependency of the relationship between various land characteristics such as vegetation biomass and topography (Bian and Walsh, 1993), vegetation index, cover types and elevation (Walsh et al., 1997), vegetation index, leaf area index and the fraction of absorbed photosynthetically-active radiation (Friedl, 1997), and of several land surface attributes required for hydrological modeling (Band and Moore, 1995).

The proliferation of environmental models over the last fifteen years also raised the issue of scale dependency. Bruneau et al. (1995) demonstrated the sensitivity of a hydrological model to both the spatial and temporal scales at which the model was run. Turner et al. (1996) showed significant differences in the outputs of a spatially distributed biogeochemical model (Forest-BGC) with the change of scale. Similarly, McNulty et al. (1997) investigated the impact of data aggregation on a forest-process model at three scales and found considerable differences in the performance of the model. Friedl (1997) reported a series of studies designed to examine how a change of spatial resolution of remotely-sensed data used as inputs into biophysical models can affect the model outputs and found that a change of scale may introduce significant bias to modeled land surface fluxes.

These studies have largely contributed to the recognition of a series of fundamental characteristics of human and ecological systems. It is now acknowledged that ecosystems are spatially heterogeneous and composed of different processes that operate dominantly at different scales and that are often non-linear. Therefore, relationships established at one scale might not be applicable at another. Processes operating at different scales affect each other. As described by hierarchy theory, processes operating at small and fast scales are constrained by those operating at slow and large scales, which in turn are constructed by the interactions of a large number of small fast processes. These interactions create feedback mechanisms. In addition, ecosystems are dynamic. Some processes can abruptly re-organize themselves through time in response to external stimuli, while emergent properties can arise from the local spatial interactions of small-scale components (Zhang et al., 2004). One of the greatest challenges in environmental modeling today is to identify the dominant scales at which various processes operate and find appropriate scaling methods to transfer information from one scale to another. It is somehow surprising that despite the clarity of the statements made sixty years ago and the considerable numbers of studies that have been since conducted on the subject, scientists dealing with spatial data analysis and modeling are still struggling with the scale dependence issue in almost its entirety.

2.1 Scale dependency in cellular automata modeling

Recently, a new class of environmental models, known as cellular automata (CA) (Wolfram, 1984), has considerably attracted the attention of the scientific community. A *cellular automaton* is a dynamic model in which space is represented as a matrix made of a regular arrangement of cells having the same dimension and shape. Each cell has a state value and evolves in time through simulation characterized by discrete time steps. Transition rules, applied at each time step, dictate how the different cell states will react to state configurations present in their neighborhood. They are often applied uniformly and synchronously to all cells, but can also be applied non-uniformly to reflect the heterogeneity of the territory, and they may incorporate a distance-based weighting function. The neighborhood can be local but extended neighborhoods are also commonly used to take into account regions of influence of different sizes. Finally, in some CA, the dynamics of the system is constrained by imposed requirements that specify the number of cells allowed to change state at each time step of the simulation (Torrens and O'Sullivan 2001; Straatman et al. 2004).

Primarily over the last 10 years, geographers and other environmental scientists have become increasingly aware of the potential offered by CA models to study the spatio-temporal dynamics of natural and human systems and to forecast their evolution in the context of management decision (see Benenson and Torrens 2004 for an excellent review). They have been used to simulate a wide range of phenomena including fire propagation (Favier et al., 2004), vegetal succession (Rietkerk et al., 2004), and rangeland degradation (Li and Reynolds, 1997). The dominant field of application however is land-use/cover change and urban development (Batty et al. 1999; White et al. 2000; Li and Yeh 2002; Wu, 2002; Almeida et al., 2003; Lau and Kam, 2005; Dietzel and Clarke, 2006; Ménard and Marceau, 2007). These studies have demonstrated that CA models are remarkably effective at generating realistic simulations of land-use patterns and other spatial structures. They are dynamic and can explicitly represent spatial processes; they are rule-based, highly adaptable and can capture a wide range of processes; they are simple and computationally efficient, and despite their simplicity, they can exhibit extraordinary rich behavior. Recently, their potential has been recognized in impact assessment, land-use planning, and social policy, and several CA models have been designed as prototypes of spatial decision-support systems for urban and regional planning (White and Engelen 2000). CA models are progressively seen not only as a framework for dynamic spatial modeling, but as a paradigm for thinking about complex spatio-temporal

phenomena and an experimental laboratory for testing ideas (Li and Yeh, 2000).

So far, the vast majority of these applications have been done using a discrete space representation and a regular tessellation of cells of same size and shape similar to the GIS raster model. However, recent studies have demonstrated that such raster-based CA models are sensitive to spatial scale. Chen and Mynett (2003) revealed that the choice of a particular cell size and neighborhood configuration has a clear effect on the resulting spatial patterns in their CA based prey-predator model. Jeanerette and Wu (2001) tested two cell sizes in their CA model developed to simulate the urbanization process in Phoenix, Arizona. They obtained reliable results with the coarser resolution, but poor results with the finer resolution. Jantz and Goetz (2005) showed the influence of cell size on the SLEUTH urban growth model; their results demonstrated that the model is able to capture the rate of growth reliably across different cell sizes, but differences in the ability to simulate growth patterns were substantial. In an exhaustive study undertaken to assess the spatial scale sensitivity in a land-use change CA model, Ménard and Marceau (2005) also revealed that the choice of cell size and neighborhood configuration has a considerable impact on the simulation results in terms of land-cover area and spatial patterns. Kocabas and Dragicevic (2006) reached similar conclusions when evaluating the sensitivity of an urban growth CA model to neighborhood size and configuration.

2.2 Solutions to mitigate scale dependency

With the increasing awareness of the importance of scale dependency in a large variety of contexts, solutions have been proposed and are currently implemented in attempts to mitigate its impact. The first one is scale sensitivity analysis, which consists in systematically assessing the effect of spatial scale on analysis and modeling results. Such an approach, quite potentially tedious, is easy to implement and represents a simple attempt at identifying the spatial units that best capture the essential characteristics of the geographical phenomenon under investigation. It might reveal the range of scales at which results do not significantly vary as well as critical thresholds indicating abrupt changes in the relationships between variables or in the modeling results. An excellent example of such an approach is provided by Bruneau et al. (1995) who performed a sensitivity analysis to space and time resolution of a hydrological model. They clearly identified a domain of spatial and temporal scale where the efficiency of the model was fairly constant and thresholds at which a strong decrease of modeling

efficiency was observable. Similarly, Ménard and Marceau (2005) were able to identify combinations of cell size and neighborhood configurations in a CA model that produced similar simulation results, while the use of different combinations resulted in significantly decreased model efficiency at generating realistic land-use patterns.

The major drawback of scale sensitivity analysis is that it is often practically impossible to test all the combinations of scale components that are relevant in the study. Therefore, an arbitrary selection of a set of discrete spatial scales or resolutions is often made, and critical values corresponding to meaningful patterns and processes are often missed in the process. This observation led to the development of a more sophisticated approach that attempts to capture the key entities or dominant patterns in a landscape as they emerge as their characteristic scale. This is known as the multi-scale approach.

Multi-scale refers to the multiple spatial dimensions at which entities, patterns and processes can be observed and measured (Hay et al., 2005). The rationale behind multi-scale analysis lies in the acknowledgement that landscapes exhibit distinctive spatial patterns associated to processes at a continuum of spatial scales and in the necessity of capturing adequate and complete information about this vertical structure of landscapes (Hay and Marceau, 2004). Numerous computational techniques have been developed over the years to generate multi-scale representation, including fractals (Mandelbrot, 1967), quadtrees (Klinger, 1971), spectral analysis (Platt and Denman, 1975), pyramids (Klinger and Dyer, 1976), wavelets (Daubechies, 1988), beamlets (Donoho and Huo, 2000) and scale space (Lindeberg, 1994). Illustrations of recent applications of such techniques are provided by Durieux et al. (2006) who applied a wavelet multi-resolution technique to improve continental scale land cover mapping from remote sensing imagery and by Hay et al. (2002) who used scale space to generate a multi-scale representation of a landscape using remote sensing data.

A third solution to mitigate scale dependency, as suggested several years ago by Fotheringham (1989) and Visvalingam (1991), is the use of meaningful geographical entities rather than arbitrary selected areal units in geographic investigations. This provides a solution to the MAUP since the product of aggregation within the spatial units relates directly to the objects of interest at the scale at which they can be observed and measured. This approach is known as entity-, or object-, or vector-based, depending on the context of its application.

In GIS, it has been shown that the object-oriented framework offers a richer semantics compared to the relational data model and permits a more direct intuitive representation of complex real-world entities and

phenomena in a way that simulates the human knowledge construction process (Makin et al., 1997; Wachowicz, 1999; Frihida et al., 2002; Shi et al., 2003). In remote sensing, object-based image analysis is increasingly recognized as the most valuable approach to get rid of the arbitrariness of pixel size when aiming at detecting significant entities and patterns from the images (Hay and Castilla, 2006; Kim and Madden, 2006). When combined to multi-scale analysis, it offers a powerful conceptual and methodological framework to automatically detect entities and patterns as they emerge at their characteristic scales (Hall and Hay, 2003; Hay et al., 2005). Object-based models, in which ecosystems are represented at fine scale as a collection of interacting components, are recognized as the obvious method to reproduce the complexity of such systems (Parrott and Kok, 2000). This conception has led to the proliferation of new classes of environmental models known as individual-based models (Grimm, 1999), multi-agent systems (Marceau, 2007) and cellular automata models (Benenson and Torrens, 2004).

2.3 The object-based approach in CA modeling

While CA models are considered individual-based models, they traditionally rely on a rather arbitrarily selected grid of cells to represent the landscape under investigation and therefore suffer from scale dependency. A solution to address the problem is the implementation of an object-based CA model.

The idea of using a vector tessellation to define space in CA modeling is not totally new. Recently, some researchers have begun to implement irregular space through the use of Voronoi polygons (Shi and Pang, 2000; Flache and Hegselmann, 2001). However, this space definition is automatically generated by the model and does not necessarily correspond to real-world entities. In other studies, the GIS vector format is used to define space, where each polygon represents a real-world entity (Hu and Li, 2004; Stevens et al., 2007). In the first model, the neighborhood is still represented using Voronoi boundaries, which does not allow the explicit definition of the neighborhood relationships. In the second model (Stevens et al., 2007), both space and neighborhood are composed of a collection of irregular cadastral land parcels and the transition rules dictate the level of development of a parcel. However, all these models present a common drawback: they do not allow the change of shape of the objects, only their change of state. This is an important limitation since changes of size and shape continuously occur in the real world and should be taken into account.

To overcome this limitation and mitigate scale dependency, a novel vector-based cellular automata model, called VecGCA, has been recently developed (Moreno et al., 2007; Moreno and Marceau, 2007). In this model, space is defined as a collection of irregular geographic objects corresponding to meaningful entities composing a landscape. Each object has a geometric representation (a polygon) that evolves through time according to a transition function that depends on the influence of its neighbors. A geometric transformation procedure allows the change of shape and size of a geographic object to more realistically reflect how land-use changes occur in a landscape. The detailed description of VecGCA and its implementation with real data is provided in the following section.

3 The Vector-based Geographic Cellular Automata Model (VecGCA)

VecGCA is an extension of the classical CA model, where the space, the neighborhood and the transition rules are redefined in order to represent objects of irregular shape that evolve through time, corresponding to real entities composing a study area.

3.1 Space, neighborhood and transition function definition

Space is defined as a collection of geo-referenced geographic objects of irregular shape. A geographic object is the representation of an entity of the real world, such as a lake, an urban land parcel, a forested patch, or a city. It is defined by its state and its geometry, which is associated to an irregular polygon that can change state and shape due to the influence of its neighbors. Each geographic object has its proper behavior and can define its neighborhood and evolution through time. It is connected to others through adjacent sides composing the whole geographic space corresponding to the study area.

VecGCA defines the neighborhood as a region of influence around each geographic object, implemented as an external buffer (Fig. 1). All the geographic objects inside this region are considered neighbors of a specific geographic object. According to this definition, non-adjacent objects can be neighbors. The size of this region of influence is expressed in length units and selected by the user.

Fig. 1. Neighborhood as defined in VecGCA. For *a*, the neighborhood delineated by *r* includes the objects *b*, *d* and *g*

Each neighbor exerts an influence on the geographic object and can produce a change of shape executed through a geometric transformation procedure. This influence is calculated for each neighbor according to a continuous function limited between 0 and 1, where 0 indicates no influence and 1 the highest influence (Eq. 3.1). The parameters of this function represent factors that are considered responsible for increasing or decreasing the influence value and can vary with the landscape under investigation.

$$g_{ab} = g\left(A(t)_a, P_{b \to a}, d_{ab}\right) \tag{3.1}$$

where

g_{ab} is the influence of the neighbor *a* on the object *b*,

$A(t)_a$ is the area of the neighbor *a* within the neighborhood of the object *b* at time *t*,

$P_{b \to a}$ is the transition probability to change from the state of the object *b* to the state of the object *a*,

d_{ab} is the distance between the centroid of the neighbor *a* and the centroid of the object *b*.

The transition function is used to calculate the area of the geographic object that changes state for the state of its neighbor when this neighbor exerts an influence whose value is higher than a threshold value corresponding to its resistance to change. This function is limited between 0, when there is no area to change, and the total area of the geographic object when the whole object changes state.

This function is specific to the study area and depends on the neighbor's area within the neighborhood and its influence. It is given by Eq. 3.2:

$$f_b(t+1) = f\left(A(t)_a, g_{ab}\right)$$

(3.2)

where

f_b is the transition function of object b.

The transition function is evaluated for each neighbor, which indicates that each neighbor can take a portion of the area of a geographic object. How and where a portion of the geographic object is removed and added to a neighbor is described in the geometric transformation procedure.

3.2 Geometric transformation procedure

This procedure executes the change of shape on a geographic object. For each neighbor, which influence is higher than a threshold value (Fig. 2a), this procedure is called and a portion (defined in the transition function) is removed from the geographic object and added to its neighbor. The area is removed from the nearest region to the corresponding neighbor. The procedure consists in building an adjustable external buffer around the neighbor, which is intersected with the geographic object (Fig. 2b). When the intersection area is approximately the area calculated in the transition function (at least 90%), this intersection area is removed from the geographic object and added to the neighbor (Fig. 2c).

(a)   (b)   (c)
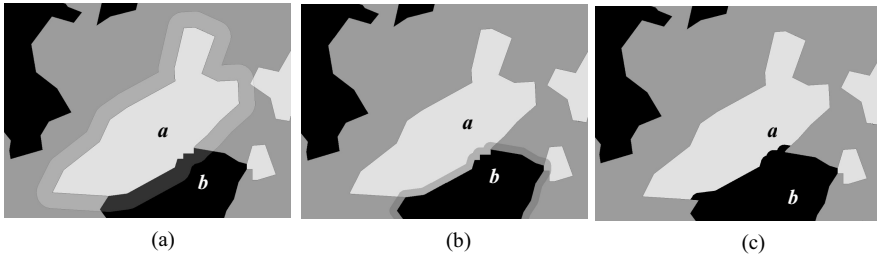
Fig. 2. **(a)** Buffer around *i* to calculate the influence of each neighbor. **(b)** The influence of *j* on *i* is higher than the threshold value; an external buffer around *j* is built and intersected with *i*. **(c)** The intersection area is removed from *i* and joined to *j*.

This procedure is called *n* times, one for each neighbor, and the order of execution is descendant from the neighbor with the highest influence to the neighbor with the lowest influence.

3.3 Modeling land-use/land-cover changes

VecGCA was tested using real data to simulate land-use/land-cover changes in two regions of different complexity in Canada. The first study area is the Maskoutains region located in Southern Quebec that covers an extent of 1312 km^2. In this region, two predominant land use/land covers can be observed: agriculture and forest. The landscape is characterized by small forested patches within a large agriculture matrix. Two land-use/land-cover maps originating from Landsat Thematic Mapper images, acquired at a spatial resolution of 30 m in 1999 and 2002 (Soucy-Gonthier et al. 2003), were used in this study.

The second study area is the Elbow river watershed in Southwest Alberta, which covers approximately 1238 km^2. A greater variety of land uses/land covers can be observed in this region, such as forest, agriculture, vegetated land (shrubs and other vegetation different from agricultural land), construction and open area, urban land, a portion of the Alberta's Rocky Mountains, a portion of the Tsuu Tina Nation reserve, and others. Data used for the study include two land-use maps generated from Landsat Thematic Mapper images acquired in the summer of 1996 and 2001, at 30 m spatial resolution.

The original land use/land cover maps (for both regions) were transformed into a vector format using the function Raster to Polygons of ArcMap (ESRI 2005).

3.3.1 Components of the land use/land cover VecGCA model

Several components must be defined in VecGCA: the space, including the set of states, the neighborhood size and influence function, and the transition function. In addition, the transition probabilities and threshold values must be calculated.

For both study areas, space is defined as a collection of patches of different land uses/land covers. Each patch corresponds to a polygon in the vector land-use/land-cover map of the study area. The vector land use/land cover maps for 1999 and 1996 were used as initial conditions for the Maskoutains region and the Elbow river watershed, respectively. In the Maskoutains region, there are four possible states for each object: forest, agriculture, water and road. Only two changes of states are considered,

namely forest to agriculture and agriculture to forest. The dynamics in the Elbow river watershed is more complex; there are nine possible states (forest, agriculture, vegetation, park/golf, urban land, forest and Tsu Tina reserve, developed land in Tsu Tina reserve, undeveloped land in Tsu Tina reserve and other), and 36 possible changes of state.

Transition probabilities were calculated from the comparison between the two land-use maps of different dates according to Eq. 3.3.

$$P_{X \to Y} = \frac{A_{X \to Y}}{\sum_{i=1}^{4} A_{X \to i}}$$

(3.3)

where

$P_{X \to Y}$ is the transition probability from state X to state Y

$A_{X \to Y}$ is the total area that changes from state X to state Y

The influence function was defined using Eq. 3.4, the variables being the same as in Equation 1. The influence value is proportional to the neighbor's area within the neighborhood, and inversely proportional to the distance between the centroids of the objects.

$$g_{ab} = 1 - e^{-P_{X_b(t) \to X_a(t+1)} * A(t)_a / d_{ab}}$$

(3.4)

The transition function that determines the area of change of each geographic object is given by Eq. 3.5.

$$f_b = \begin{cases} A(t)_a * g_{ab} & \text{if } g_{ab} \geq \lambda_{ab} \\ 0 & \text{other case} \end{cases}$$

(3.5)

where λ_{ab} is a threshold value that represents the resistance of the geographic object b to change its state for the state of its neighbor a. This value can be defined as the probability that a geographic object does not change its state from the state X to the state Y although all its neighbors are in state Y. The same transition function was applied to define the dynamics of both study areas.

3.3.2 Definition of the raster-based CA model

A stochastic raster-based CA model was implemented for each study area to compare its results with the simulation outcomes of the VecGCA

models. For the Maskoutains region, space was defined as a regular rectangular grid of 409 columns and 508 rows and a cell size of 100 m. This size was chosen based on the results previously obtained in a scale sensitivity analysis conducted by Ménard and Marceau (2005), which indicates that 100 m is the cell size that best captures the dynamics of the study area. For the Elbow river watershed, the grid defining space has 2450 columns and 1864 rows and a cell size of 30 m (corresponding to the original resolution of the land-use data). The initial conditions correspond to the 1999 and the 1996 raster-land use maps for the Maskoutains region and the Elbow river watershed, respectively.

For both models, a Moore neighborhood was chosen to represent the influence of the adjacent cells on a central cell. Probabilistic rules were calculated from the comparison between two land-use maps (1999 and 2002 for the Maskoutains region, and 1996 and 2001 for the Elbow river watershed), according to the procedure described in Ménard and Marceau (2005), where a cell in the state X that has n cells in the state Y in its neighborhood has a probability of changing to the state Y equal to the number of cells in the state X with n neighbors in the state Y that have changed to the state Y between t_1 and t_2, divided by the total number of cells in the state X with the same neighborhood in t_1.

To account for a temporal resolution of one year, the probabilistic rules were adjusted using the exponential method presented by Yeh and Li (2006) where the transition probability P calculated for a time step t is substituted by P^n for a time step T where $T = n*t$.

3.3.3 Model simulations

While it is hypothesized that the use of polygons to define space rather than cells of arbitrary sizes will mitigate the cell size sensitivity when using VecGCA, the potential sensitivity of the model to the neighborhood configuration remains. To address this issue, four simulations, from 1999 to 2002, were performed for the Maskoutains region and four others, from 1996 to 2001 were conducted for the Elbow river watershed. Each simulation was associated to a different neighborhood size: 10 m, 30 m, 60 m and 120 m. The results were compared to the land-use/land-cover maps of 2002 and 2001 for the Maskoutains region and the Elbow river watershed, respectively.

To compare the simulation outcomes produced by the VecGCA models and the raster-based CA models, a landscape analysis using Fragstats 3.3. (McGarigal et al. 1995) was done on the raster-based CA results. The number of patches for each land use was calculated on the raster map generated by the raster-based CA model and compared to the number of

polygons produced in the VecGCA model for each study area. In addition, for the Maskoutains region, an overlay of the 2002 vector and 100 m raster land-use/land-cover maps with the simulation outcomes of VecGCA and the raster-based CA models, respectively, was performed to determine the correspondence between the results of the models and the real state of the study area. The same procedure was executed for the Elbow river watershed using the 2001 vector and 30 m raster land-use/land-cover maps and the corresponding simulation outcomes.

3.4 Simulation results

The results obtained with VecGCA (using a neighborhood size of 30 m) for both regions were compared with the results obtained with a raster-based CA model applied for each study area. This comparison revealed that for both models the proportions of land use in the Maskoutains region are very similar to the proportions observed in the 2002 original land-use map (Table 1). A decrease of the total number of patches/polygons was observed in both models due to the disappearance of forested patches absorbed by large agricultural patches. In the raster-based CA model, this behavior is associated to the high transition probabilities from forest to agriculture when the number of agricultural neighbors is higher than 5. In the VecGCA, the disappearance of forested patches absorbed by large agricultural patches is explained by the high pressure that receives a forested object with an agricultural neighbor that covers its entire neighborhood. The raster-based CA model generates good simulation results because the cell size used (100 m) was determined from a previous sensitivity analysis. In comparison, the results of VecGCA were obtained from the original spatial distribution of the land-use map.

Table 1. Proportion of land-use and number of patches/polygons in the outcomes of the VecGCA model, the raster-based CA model and the 2002 raster land-use map in the Maskoutains region

	Initial condition Land-use map 1999		Land-use map 2002		Simulation outcomes for 2002	
	Vector format	Raster format	Vector format	Raster format	VecGCA	Raster-based CA
Forest[%]	16.57	16.58	14.83	14.86	14.92	14.86
Agriculture [%]	80.70	80.70	82.45	82.41	82.35	82.41
Total number of patches / polygons	5247	3118	5702	1371	4666	1569

A spatial overlay analysis shows that the land-use distribution generated by VecGCA for 2002 coincides as an average in 99% with the 2002 land-use map, whereas the distribution produced by the raster-based CA coincides in 89%, although an appropriate cell size has been used (Table 2). 100 % of the forest polygons produced by VecGCA coincide with the real forested patches in the study area, in comparison to only 77% generated by the 100 m raster-based CA model. These results can by explained by the capacity of VecGCA to reproduce the evolution of the objects by a change of shape, whereas the patches produced by the raster-based CA model are created by the agglomeration of individual cells changing state. In addition, when looking closely at the simulation maps, one can observe that the landscape generated by VecGCA is characterized by large patches of well-defined boundaries, in comparison to the diffuse boundaries and the high level of landscape fragmentation produced by the raster-based CA model (Fig. 3).

Table 2. Proportion of simulated area that coincides with the state of the system in 2002 for each land use in the Maskoutains region

Land uses	Proportion of simulated land uses [%]	
	VecGCA	Raster-based CA
Forest	100.00	77.47
Agriculture	99.00	93.04
Other	98.54	97.86
Average	99.18	89.45

Raster land-use map 2002
(resolution 30 m)

VecGCA result in 2002
(neighborhood 30 m)

Raster-based CA result in 2002
(cell size 100 m)

Forest Agriculture Others
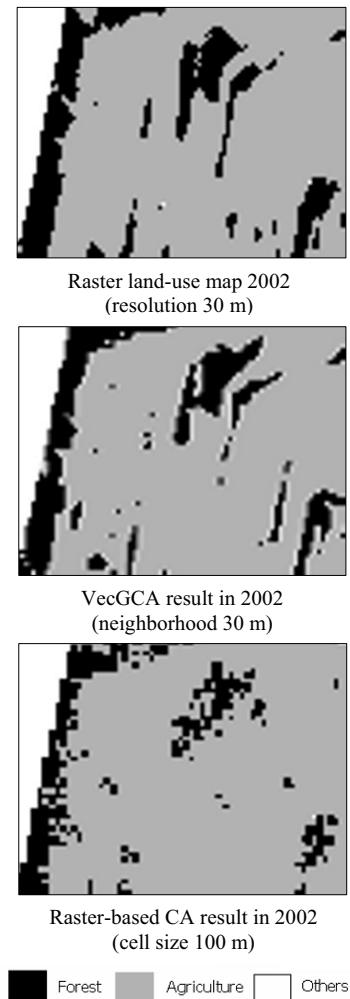
Fig. 3. Detail on the polygons' boundaries produced by the VecGCA model and the raster-based CA model in the Maskoutains region

Different results were obtained for the Elbow river watershed. In this case, VecGCA generated a land-use distribution that is more similar to the 2001 land-use map in comparison to the results produced by the raster-based CA (Table 3; Fig. 4). The proportions of the land-use classes vegetation, park/golf and developed land in the Tsu Tina Reserve are

underestimated by the raster CA model, while other classes such as urban and undeveloped land in the Tsu Tina Reserve are overestimated.

Table 3. Proportion of land-use and number of patches/polygons in the outcomes of the VecGCA model, the raster-based CA model and the 2001 raster land-use map in the Elbow river watershed

	Initial condition Land-use map 1996		Land-use map 2001		Simulation outcomes for 2001	
	Vector format	Raster format	Vector format	Raster format	VecGCA	Raster-based CA
Forest [%]	44.45	44.45	44.34	44.31	45.57	44.21
Agriculture [%]	15.39	15.39	13.47	13.43	15.29	15.02
Vegetation [%]	1.98	1.98	1.68	1.71	1.82	0.28
Park/golf [%]	0.87	0.87	0.82	0.83	0.54	0.23
Urban [%]	1.56	1.56	2.30	2.29	2.09	3.81
Forest in Tsu Tina Reserve [%]	6.80	6.80	5.87	5.87	6.99	5.90
Developed land in Tsu Tina Reserve [%]	3.04	3.04	3.22	3.23	3.11	1.40
Undeveloped land in Tsu Tina Reserve [%]	2.88	2.88	3.52	3.51	2.79	5.42
Total number of patches/polygons	7195	5837	8306	6686	2986	4321

A spatial overlay analysis of the land-use maps generated by VecGCA and the raster-based CA model with the 2001 land-use maps in vector and raster format reveals that the results obtained with VecGCA highly coincide with the land-use spatial distribution in the study area, whereas the results obtained with the raster-based CA model largely differ for most land-use classes (Table 4). The cell size used for the raster-based CA model was 30 m, which is the same as the resolution of the original land-use data. In that case, no previous sensitivity analysis was done to determine the best cell size to be used, and we might hypothesize that it is not the most appropriate cell size for this study area either.

Raster land-use map 2001
(resolution 30 m)

VecGCA result in 2001
(neighborhood 30 m)

Raster-based CA result in 2001 (cell size 30 m)
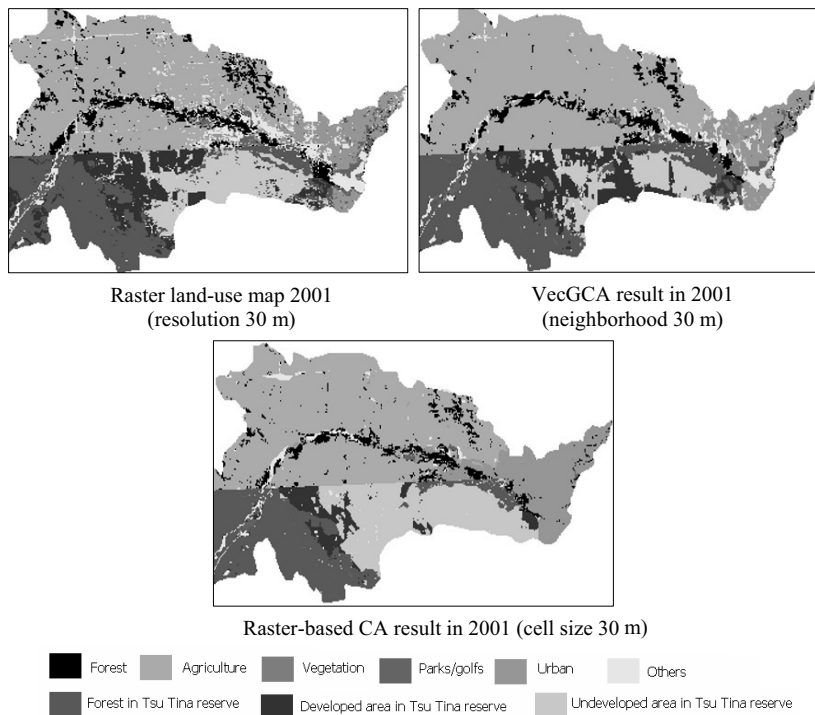
Fig. 4. Sub region of the Elbow river watershed showing the land-use spatial distribution generated by the two models

Table 4. Proportion of simulated area that coincides with the state of the system in 2001 for each land use in the Elbow river watershed

Land uses	Proportion of simulated land uses [%]	
	VecGCA	Raster-based CA
Forest	93.79	88.96
Agriculture	98.67	85.67
Vegetation	83.13	21.19
Park/golf	100.00	60.93
Urban	80.80	54.45
Forest in Tsu Tina Reserve	92.69	87.33
Developed land in Tsu Tina Reserve	96.45	73.90
Undeveloped land in Tsu Tina Reserve	100.00	60.42
Average	91.34	66.60

When varying the neighborhood size in VecGCA, the results for the Maskoutains region reveal that for the neighborhood sizes of 10 m and 30 m, the simulated proportion of forested and agricultural land for 2002 differs in less than 2% of the proportion calculated from the 2002 land-use map; for the neighborhoods of 60 m and 120 m, the difference might exceed 8% (Table 5). However, for the Elbow river watershed the variation of the neighborhood size does not produce significant variation in the simulation outcomes (Table 6). These results can be explained by the fact that the influence function (Equation 4) and the transition function (Equation 5) are directly proportional to the neighbors' area within the neighborhood and this area varies with the neighborhood size and the landscape configuration. In the Maskoutains region, the majority of the objects are small forested patches having only one agricultural neighbor (Fig. 5a). The influence of this neighbor and the area to change from forest to agriculture increase when the neighborhood size increases. In the Elbow river watershed, the geographic objects have several neighbors of different states (Fig. 5b). In this case, an increase of the neighborhood size produces a small increase of the neighbors' area within the neighborhood in comparison to the objects that have only one neighbor. In addition, when the neighborhood increases, new neighbors appear which influence and area to change are not significant because they are distant geographic objects separated by other objects. Therefore, in this landscape configuration the simulation outcomes are less sensitive to the neighborhood size.

Table 5. Proportion of land-use/land-cover (%) in the Maskoutains region using different neighborhood sizes

Land uses	Neighborhood	1999	2000	2001	2002	[a]
	10 m	16.57	16.38	16.30	16.22	1.40
	30 m	16.57	16.03	15.44	14.59	0.23
Forest[%]	60 m	16.57	16.02	14.90	12.16	2.66
	120 m	16.57	12.40	9.43	6.37	8.46
	Original	16.57	-	-	14.83	-
	10 m	80.70	80.89	80.98	81.05	1.40
	30 m	80.70	81.25	81.84	82.68	0.24
Agriculture[%]	60 m	80.70	81.25	82.38	85.11	2.67
	120 m	80.70	84.88	87.84	90.91	8.46
	Original	80.70	-	-	82.45	-

[a] Variation between the simulation outcomes and the 2002 land-use/land-cover map in the Maskoutains region
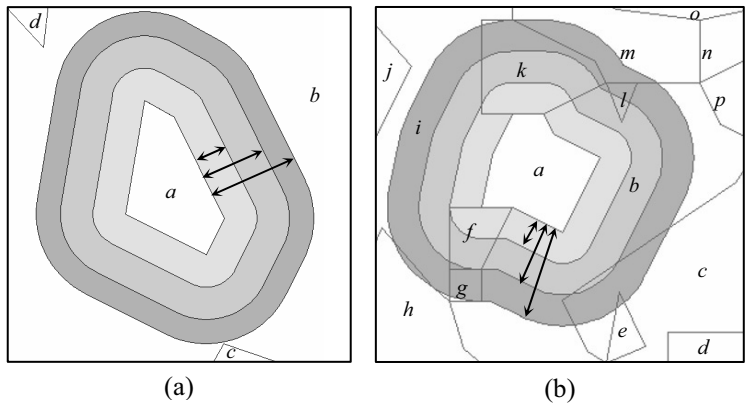
Fig. 5 (a) Schematic representation of the landscape configuration of the Maskoutains region where the objects have only one neighbor for different neighborhood sizes. **(b)** Schematic representation of the Elbow river landscape configuration where the objects have several neighbors; when the neighborhood size increases the number of neighbors and the neighbors' area within the neighborhood also increase

Table 6. Proportion of land-use/land-cover (%) in the Elbow river watershed using different neighborhood sizes

	Neighborhood	1996	1997	1998	1999	2000	2001	[a]
Forest	10 m	44.45	44.62	44.58	44.44	44.42	44.39	0.04
	30 m	44.45	44.49	44.36	44.15	44.11	43.46	0.89
	60 m	44.45	44.97	44.84	44.72	43.62	43.58	0.77
	120 m	44.45	45.53	45.35	45.08	44.86	43.53	0.82
	Original	44.45					44.35	
Agriculture	10 m	15.39	15.33	15.28	15.26	15.24	15.21	1.73
	30 m	15.39	15.41	15.30	15.40	15.36	15.35	1.87
	60 m	15.39	15.58	15.51	15.40	15.35	15.32	1.84
	120 m	15.39	15.74	15.77	15.60	15.56	15.56	2.09
	Original	15.39					13.48	
Vegetation	10 m	1.98	1.85	1.82	1.82	1.80	1.79	0.11
	30 m	1.98	1.61	1.48	1.45	1.43	1.54	0.14
	60 m	1.98	1.32	1.19	1.18	1.39	1.36	0.32
	120 m	1.98	0.82	0.48	0.28	0.25	0.33	1.35
	Original	1.98					1.68	
Parkrs	10 m	0.87	0.81	0.78	0.75	0.73	0.74	0.08
	30 m	0.87	0.74	0.67	0.62	0.61	0.60	0.23
	60 m	0.87	0.72	0.67	0.66	0.64	0.56	0.27
	120 m	0.87	0.71	0.66	0.62	0.62	0.61	0.22
	Original	0.87					0.82	

Urban	10 m	1.56	1.60	1.61	1.63	1.63	1.67	0.63
	30 m	1.56	1.66	1.69	1.71	1.72	1.72	0.58
	60 m	1.56	1.67	1.69	1.69	1.70	1.75	0.55
	120 m	1.56	1.66	1.68	1.70	1.70	1.70	0.60
	Original	1.56					2.30	
Forest in TTNR[b]	10 m	6.80	6.78	6.75	6.74	6.72	6.73	0.86
	30 m	6.80	6.79	6.66	6.61	6.60	6.59	0.72
	60 m	6.80	6.79	6.71	6.64	6.63	6.62	0.75
	120 m	6.80	6.81	6.77	6.72	6.50	6.46	0.59
	Original	6.80					5.87	
Developed land in TTNR[b]	10 m	3.04	3.03	3.01	2.99	2.99	2.97	0.25
	30 m	3.04	3.01	3.07	3.07	3.06	3.05	0.17
	60 m	3.04	3.03	2.95	2.97	2.96	2.97	0.26
	120 m	3.04	3.00	2.94	2.94	2.78	2.64	0.58
	Original	3.04					3.22	
Undeveloped land in TTNR[b]	10 m	2.88	2.92	2.96	2.99	3.01	3.05	0.47
	30 m	2.88	2.92	3.00	3.05	3.07	3.09	0.43
	60 m	2.88	2.91	3.07	3.11	3.13	3.14	0.38
	120 m	2.88	2.92	3.01	3.07	3.44	3.63	0.11
	Original	2.88					3.52	

[a] Variation between the simulation outcomes and the 2002 land-use/land-cover map in the Elbow river watershed
[b] TTNR: Tsu Tina Nation Reserve

4. Conclusion

Despite the numerous studies conducted on the subject over the last five decades, scale dependency remains a critical issue in spatial analysis and modeling. Lately, scientists have recognized its manifestation in the context of a new class of environmental models, known as cellular automata, which are increasingly used to simulate the dynamics of natural and human systems and forecast their evolution. This scale sensitivity is due to the arbitrary selection of a regular grid of cells to represent space and the neighborhood in CA modeling. In attempts to overcome this problem, a vector-based CA model (VecGCA) has been developed and compared to a raster-based model to simulate land-use changes in two study areas of varying complexity. VecGCA offers two main advantages over the conventional raster model. First, space is defined as a collection of irregular geographic objects using polygons that correspond to meaningful entities composing the landscape. Second, the model allows a change of shape and size of the polygons in addition to a change of state. The results reveal that VecGCA generates a more realistic representation of the evolution of the landscape compared to the conventional raster model. However, while VecGCA mitigates the sensitivity to cell size, it

remains sensitive to the neighborhood delineation, which is done using a buffer of arbitrary size. Work is currently in progress to implement the concept of dynamic neighborhood based on semantic relationships between the geographic objects composing the landscape.

Acknowledgements

References

Ahern, F.J. D.N.N. Horler, J. Cihlar, W.J. Bennett, and E. MacAulay. 1983. "Digital Processing to Improve Classification Results at Resolutions of 5 to 50 Meters". In *Proceedings of the SPIE Symposium on Techniques for Extraction from Remotely Sensed Images*, Rochester, NY, 16-19 August, pp. 153-170.

Allen, T.F.H., and T.B. Starr. 1982. *Hierarchy Perspective for Ecological Complexity*. University of Chicago Press, Chicago.

Almeida, C. M. d., M. Batty, A.M.V. Monteiro, G. Câmara, B.S. Soares-Filho, G.C. Cerqueira, and C.L. Pennachin. 2003. "Stochastic cellular automata modeling of urban land use dynamics: empirical development and estimation", *Computers, Environment and Urban Systems*, Vol. 27, pp. 481-509.

Bach, L. 1981. "The Problem of Aggregation and Distance for Analysis of Accessibility and access Opportunity in Location-Allocation Models", *Environment and Planning A*, Vol. 13, pp. 955-978.

Band, L. E. and I. D. Moore. 1995. "Scale: landscape attributes and geographical information systems", *Hydrological Processes*, Vol. 9, pp. 401-422.

Batty, M., Y. Xie, and A. Sun. 1999. "Modeling urban dynamics through GIS-based cellular automata", *Computers, Environment and Urban Systems*, Vol. 23, pp. 205-233.

Benenson, I., and P.M. Torrens. 2004. "Geosimulation: object-based modeling of urban phenomena", *Computers, Environment and Urban Systems*, Vol. 28, pp. 1-8.

Benson, B.J., and M.D. McKenzie. 1995. "Effects of Sensor Spatial Resolution on Landscape Structure Parameters", *Landscape Ecology*, Vol. 10, No. 2, pp. 113-120

Bian, L., and S.J. Walsh. 1993. "Scale Dependencies of Vegetation and Topography in a Mountainous Environment of Montana", *The Professional Geographer*, Vol. 45, No. 1, pp. 1-11.

Bruneau, P., C. Gascuel-Odoux, P. Robin, Ph. Merot, and K. Beven. 1995. "Sensitivity to Space and Time Resolution of a Hydrological Model Using Digital Elevation Data", *Hydrological Processes*, Vol. 9, pp. 69-81.

Cao, C., and N.S.-N. Lam. 1997. "Understanding the Scale and Resolution Effects in Remote Sensing and GIS". In *Scale in Remote Sensing in GIS*, Quattrochi, D.A. and M.F. Goodchild, eds., pp. 57-72.

Chang, K. T. and B. W. Tsai. 1991. "The effect of DEM resolution on slope and aspect mapping", *Cartography and Geographic Information Science*, Vol. 18, No. 1, pp. 69-77.

Chen, Q., and A.E. Mynett. 2003. "Effects of cell size and configuration in cellular automata based prey-predator modeling", *Simulation Modeling Practice and Theory*, Vol. 11, pp. 609-625.

Clarke, W.A.V., and K.L. Avery. 1976. "The Effects of Data Aggregation in Statistical Analysis", *Geographical Analysis*, Vol. 28, pp. 428-438.

Cushnie, J.L. 1987. "The Interactive Effect of Spatial Resolution and Degree of Internal Variability within Land-Cover Types on Classification Accuracies", *International Journal of Remote Sensing*, Vol. 8, No. 1, pp.15-29.

Daubechies, I. 1988. "Orthonormal bases of compactly supported wavelets", *Communications on Pure and Applied Mathematics*, Vol. 41, pp. 906-966.

Dietzel, C., and K. Clarke, K. 2006. "The effect of disaggregating land use categories in cellular automata during model calibration and forecasting", *Computers, Environment and Urban Systems*, Vol. 30, No. 1, pp. 78-101.

Donoho, D.L., and X. Huo. 2000. "Beamlets and multiscale image analysis", In *Lecture Notes in computational Science and Engineering: Multiscale and Multiresolution Methods*, Springer, pp. 149-196.

Durieux, L., J. Kropacek, G.D. De Grandi and F. Achard. 2006. "Object-oriented wavelet multi-resolution image analysis of the Siberia GBFM radar mosaic combined with MERIS imagery for continental scale land cover mapping". *First International Conference on Object-Based Image Analysis (OBIA 2006)*, Salzburg University, Austria, July 4-5.

ESRI. (2005). "ArcGIS 9. Desktop GIS. <http://www.esri.com/software/arcgis/about/desktop.html>."

Favier, C., J. Chave, A. Fabing, D. Schwartz, and M.A. Dubois. 2004. "Modelling forest-savanna mosaic dynamics in man-influenced environments: effects of fire, climate and soil heterogeneity", *Ecological Modelling*, Vol. 171, pp. 85-102.

Flache, A., and R. Hegselmann. 2001. "Do irregular grids make a difference? Relaxing the spatial regularity assumption in cellular models of social dynamics", *Journal of Artificial Societies and Social Simulation*, Vol. 4, No. 4, http://jasss.soc.surrey.ac.uk/4/4/6.html.

Fotheringham, A.S. 1989. "Scale-Independent Spatial Analysis", In *Accuracy of Spatial Databases*, Goodchild, M., S. Gopal, Eds, Taylor and Francis, pp. 221-228.

Fotheringham, A.S., and D.W.S. Wong. 1991. "TheModifiable Areal Unit Problem in Multivariate Statistical Analysis", *Environment and Planning A*, Vol. 23, pp.1025-1044.

Friedl, M.A. 1997. "Examining the Effects of Sensor Resolution and Sub-pixel Heterogeneity on Spectral Vegetation Index: Implications for Biophysical Modeling", In *Scale in Remote Sensing in GIS*, Quattrochi, D.A. and M.F. Goodchild, eds., pp.113-139.

Friedl, M.A. 1997. "Examining the Effects of Sensor Resolution and Sub-Pixel Heterogeneity on Spectral Vegetation Indices: Implication for Biophysical Modeling". In *Scale in Remote Sensing in GIS*, Quattrochi, D.A. and M.F. Goodchild, eds., pp. 113-139.

Frihida, A., D.J. Marceau, and M. Thériault. 2002. "Spatio-temporal object-oriented data model for disaggregate travel behaviour", *Transactions in GIS*, Vol. 6, No. 3, pp. 277-294.

Gao, J. 1997. "Resolution and accuracy of terrain representation by grid DEMs at a micro-scale", *International Journal of Geographical Information Science*, Vol. 11, No. 2, pp. 199-212.

Gehlke, C.E., and K. Biehl. 1934. "Certain Effects of Grouping Upon the size of the Correlation Coefficient in Census Tract Material", *Journal of American Statistical Association Supplement*, Vol. 29, pp. 169-170.

Grimm, V. 1999. "Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future", *Ecological Modelling*, Vol. 115, pp. 129-148.

Hall, O. and G. Hay. 2003. "A multiscale object-specific approach to digital change detection", *International Journal of Applied Earth Observations and Geoinformation*, Vol. 4, pp. 311-327.

Hay, G. and G. Castilla. 2006. "Object-based image analysis: strengths, weaknesses, opportunities and threats (SWOT)". *First International Conference on Object-Based Image Analysis (OBIA 2006)*, Salzburg University, Austria, July 4-5.

Hay, G.J. and D.J. Marceau. 2004. "Multiscale object-specific analysis (MOSA): An integrative approach for multiscale analysis". In: *Remote Sensing Image Analysis: Including the Spatial Domain*, S. de Jong and F. van der Meer, eds, Kluwer Academic Publishers, pp. 71-92.

Hay, G., G. Castilla, M. Wulder and J. R. Ruiz. 2005. "An automated object-based approach for the multiscale image segmentation of forest scenes", *International Journal of Applied Earth Observations and Geoinformation*, Vol. 7, pp. 339-359.

Hay, G.J., P. Dubé, A. Bouchard, and D.J. Marceau. 2002. "A Scale-Space Primer for Exploring and quantifying Complex Landscape", *Ecological Modelling*, Vol. 153, Nos. 1/1, pp. 27-49.

Hu, S., and D. Li. 2004. "Vector Cellular Automata Based Geographical Entity", *12th International Conference on Geoinformatics - Geoespatial Information Research: Bridging the Pacific and Atlantic.*, University of Gävle, Sweden.

Hunt, L., and B. Boots. 1996. "MAUP Effects in the Principal Axis Factoring Technique", *Geographical Systems*, Vol. 3, pp. 101-121.

Irons, J.R., B.L. Markham, R.F. Nelson, D.L. Toll, D.L. Williams, R.S. Latty, and M.L. Stauffer. 1985. "The Effects of Spatial Resolution on the Classification

of Thematic Mapper Data", *International Journal of Remote Sensing*, Vol. 6, No. 8, pp. 1385-1403.

Jantz, C. A., and S.J. Goetz. 2005. "Analysis of scale dependencies in an urban land-use-change model", *International Journal of Geographical Information Science*, Vol. 19, No. 2, pp. 217-241.

Jarvis, P.G. 1995. "Scaling Processes and Problems", *Plant, Cell and Environment*, Vol. 18, pp. 1079-1089.

Jenerette, G. D., and J. Wu. 2001. "Analysis and simulation of land-use change in the central Arizona-Phoenix region, USA", *Landscape Ecology*, Vol. 16, pp. 611-626.

Kershaw, K.A. 1957. "The Use of Cover and Frequency in the Detection of Pattern in Plant Communities", *Ecology*, Vol. 38, pp. 291-299.

Kim, M. and M. Madden. 2006. "Determination of optimal scale parameter for alliance-level forest classification of multispectral IKONOS images". *First International Conference on Object-Based Image Analysis (OBIA 2006)*. Salzburg University, Austria, July 4-5.

Klinger, A. 1971. "Pattern and search statistics". In *Optimizing Methods in Statistics*, J.S. Rustagi, eds., Academic Press, pp. 303-339.

Klinger, A. and C.R. Dyer. 1976. "Experiments in picture representation using regular decomposition", *Computer Graphics Image Processing*, Vol. 5, pp. 65-105.

Kocabas, V., and S. Dragicevic. 2006. "Assessing cellular automata model behavior using a sensitivity analysis approach", *Computers, Environment and Urban Systems*, Vol. 30, No. 6, pp. 921-953.

Kok, K. and A. Veldkamp. 2001. "Evaluating impact of spatial scales on land use pattern analysis in Central America", *Agriculture, Ecosystems and Environment*, Vol. 85, pp. 205-221.

Latty, R.S. and R.M. Hoffer. 1981. "Computer-Based Classification Accuracy Due to the Spatial Resolution Using Per-Point versus Per-Field Classification Techniques". Symposium of Machine Processing of Remotely Sensed Data, pp. 384-392.

Lau, K. H., and B.H. Kam. 2005. "A cellular automata model for urban land-use simulation", *Environment and Planning B: Planning and Design*, Vol. 32, pp. 247-263.

Li, H. and F. Reynolds. 1997. "Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: a combined Markov and cellular automaton approach". *Scale in Remote Sensing and GIS*. D. Quattrochi and M. F. Goodchild, eds, CRC Press, pp. 211-230.

Li, X., and A. G.-o. Yeh. 2000. "Modelling sustainable urban development by the integration of constrained cellular automata and GIS", *International Journal of Geographical Information Science,* Vol. 14, pp. 131-152.

Li, X., and A. G.-o. Yeh. 2002. "Neural-network-based cellular automata for simulating multiple land use changes using GIS", *International Journal of Geographical Information Science*, Vol. 16, No. 4, pp. 323-343.

Lindeberg, T. 1994. "Scale-space theory: A basic tool for analyzing structures at different scales", *Journal of Applied Statistics*, Vol. 21, No. 2, pp. 225-270.

Makin, J., R. G. Healey, and S. Dowers. 1997. "Simulation modelling with object-oriented GIS: a prototype application to the time geography of shopping behaviour", *Geographical Systems*, Vol. 4, No. 4, pp. 397-429.

Mandelbrot, B.B. 1967. "The Fractal geometry of nature", *Science*, Vol. 156, pp. 636-642.

Marceau, D.J., 2007. "What can be learned from multi-agent systems?", In: *Monitoring, Simulation and Management of Visitor Landscapes*, R. Gimblett, ed., University of Arizona Press. (submitted).

Marceau, D.J. 1992. *The Problem of Scale and Spatial Aggregation in Remote Sensing: An Empirical Investigation Using Forestry Data*. Unpublished Ph.D. Thesis, Department of Geography, University of Waterloo, 180 p.

Marceau, D.J., 1999. "The Scale Issue in the Social and Natural Sciences". *Canadian Journal of Remote Sensing*, Vol. 25, No. 4, pp. 347-356.

Marceau, D.J., and G.J. Hay, 1999. "Remote sensing contributions to the scale issue", *Canadian Journal of Remote Sensing*, Vol. 25, No. 4, pp. 357-366.

Marceau, D.J., D.J. Gratton, r. Fournier, and J.P. Fortin. 1994. "Remote Sensing and the Measurement of Geographical Entities in a Forested Environment; Part 2: The Optimal Spatial Resolution", *Remote Sensing and Environment*, Vol. 49, No. 2, pp. 105-117.

Marceau, D.J., P.J. Howarth, D.J. Gratton. 1994. "Remote Sensing and the Measurement of Geographical Entities in a Forested Environment; Part 1: The Scale and Spatial Aggregation Problem", *Remote Sensing and Environment*, Vol. 49, No. 2, pp. 93-104.

Markham, B.L., and J.R.G. Townshend. 1981. "Land Cover Classification Accuracy as a Function of Sensor Spatial Resolution". *Proceedings of the Fifteenth International Symposium on Remote Sensing of Environment*, Ann Arbor, Michigan, pp. 1075-1090.

McCarthy, H.H., J.C. Hook, and D.S. Knos. 1956. *The Measurement of Association in Industrial Geography*. Department of Geography, State University of Iowa, Iowa City.

McGarigal, K., B. Marks, E. Ene, and C. Holmes. 1995. "Fragstats". University of Massachusetts.
http://www.umass.edu/landeco/research/fragstats/fragstats.html.

McNulty, S.G., J.M. Vose, and W.T. Swank. 1997. "Scaling Predicted Pine Forest Hydrology and Productivity Across the Southern United States". In *Scale in Remote Sensing in GIS*, Quattrochi, D.A. and M.F. Goodchild, eds., pp. 187-209.

Meetemeyer, V., and E.O. Box. 1987. "Scale Effects in Landscape Studies". In *Landscape Heterogeneity and Disturbance*, M.G. Turner, ed., Springer-Verlag, pp. 15-34.

Meetemeyer. V. 1989. "Geographical Perspectives of Space, Time, and Scale", *Landscape Ecology*, Vol. 3, Nos. 3/4, pp. 163-173.

Ménard, A., and D.J. Marceau. 2005. "Exploration of spatial scale sensitivity in geographic cellular automata", *Environment and Planning B: Planning and Design*, Vol. 32, pp. 693-714.

Ménard, A., and D.J. Marceau. 2007. "Simulating the impact of forest management scenarios in an agricultural landscape of southern Quebec, Canada, using a geographic cellular automata", *Landscape and Urban Planning*, Vol. 79, Nos. 3/4, pp. 253-265.

Moody, A., and C.E. Woodcock. 1995. "The Influence of Scale and the Spatial Characteristics of Landscapes on Land-Cover Mapping Using Remote Sensing", *Landscape Ecology*, Vol. 10, No. 6, pp. 363-379.

Moreno, N. and D.J. Marceau, 2006. "A vector-based cellular automata model to allow changes of polygon shape", In *Proceedings of The 2006 SCS International Conference on Modeling and Simulation - Methodology, Tools, Software Applications*, R. Huntsinger, H. Vakilzadian and T. Ören, eds, July 31 - August 2, 2006, Calgary, Canada, pp. 85-92.

Moreno, N. and D.J. Marceau, 2007. "Modeling land-use changes using a novel vector-based cellular automata model", In *Proceedings of the 18th IASTED International Conference on Modelling and Simulation*, May 30 - June 1, 2007, Montreal, Quebec.

O'Neill, R.V., D.L. De Angelis, J.B. Waide, and T.F.H. Allen. 1986. *A Hierarchical Concept of Ecosystems*. Princeton University Press, Princeton, New Jersey.

Openshaw, S. 1977. "A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning and Spatial Modelling", *Institute of British Geographers*, Transactions, New Series, Vol. 2, pp. 459-472.

Openshaw, S. 1978. "An Empirical Study of Some Zone-Design Criteria", *Environment and Planning A*, Vol. 10, pp.781-794

Openshaw, S. 1984. *The Modifiable Areal Unit Problem*. Concepts and Techniques in Modern Geography (CATMOG), no. 38.

Parrott, L. and R. Kok. 2000. "Incorporating complexity in ecosystem modeling", *Complexity International*, Vol. 7, pp. 1-19.

Platt, T., and K.L. Denman. 1975. "Spectral analysis in ecology", *Annals of Rev. Ecological System*, Vol. 6, pp. 189-210.

Putman, S.H., and S.–H. Chung. 1989. "Effects of Spatial System Design on Spatial Interaction Models, 1: The Spatial System Definition Problem", *Environment and Planning A*, Vol. 21, pp. 27-46.

Qi, Y., and J. Wu. 1996. "Effects of Changing Spatial Resolution on the Results of Landscape Pattern Analysis Using Spatial Autocorrelation Indices", *Landscape Ecology*, Vol. 11, pp. 39-49.

Rietkerk, M., S.C. Dekker, P.C.D. Ruiter, and J.V.D Koppel. 2004. "Self-Organized Patchiness and Catastrophic Shifts in Ecosystems", *Review*, Vol. 305, pp. 1926-1929.

Sadowski, F.G., W.A. Malila, J.E. Sarno, and R.F. Nalepka. 1977. "The Influence of Multispectral Scanner Spatial Resolution on Forest Feature Classification". *Proceedings of the 11th International Symposium on Remote Sensing of Environment*, 25-29 April, Ann Arbor, pp. 1279-1288.

Shi, W., B. Yang, and Q. Li. 2003. "An object-oriented data model for complex objects in three-dimensional geographical information systems", *International Journal of Geographical Information Science*, Vol. 17, No. 5, pp. 411-430.

Shi, W., and M.Y.C. Pang. 2000. "Development of Voronoi-based cellular automata - an integrated dynamic model for Geographical Information Systems", *International Journal of Geographical Information Science*, Vol. 14, No. 5, pp. 455-474.

Soucy-Gonthier, N., Marceau, D., Delage, M., Cogliastro, A., Domon, G., and Bouchard, A. (2003). "Détection de l'évolution des superficies forestières en Montérégie entre juin 1999 et août 2002 à partir d'images satellitaires Landsat TM." *Agence forestière de la Montérégie*.

Stevens, D., S. Dragicevic, S., and K. Rothley. 2007. "iCity: A GIS-CA modelling tool for urban planning and decision making", *Environmental Modelling & Software*, Vol. 22, pp. 761-773.

Straatman, B., R. White, and G. Engelen. 2004. "Towards an automatic calibration procedure for constrained cellular automata", *Computers, Environment and Urban Systems*, Vol. 28, pp. 149-170.

Torrens, P. and D. O'Sullivan. 2001. "Cellular automata and urban simulation: where do we go from here?", *Environment and Planning B*, Vol. 28, pp. 163-168.

Turner, D.P., R. Dodson, and D. Marks. 1996. "Comparison of Alternative Spatial Resolutions in the Application of a Spatially Distributed Biogeochemical Model over Complex Terrain", *Ecological Modeling*, Vol. 90. pp. 53-67.

Turner, M.G., R.V. O'Neil, R.H. Gardner, and B.T. Milne. 1989. "Effects of Changing spatial Scale on the analysis of Landscape Pattern", *Landscape Ecology*, Vol. 3, pp. 153-162.

Usher, M.B. 1969. "The Relation Between Mean Square and Block Size in the Analysis of Similar Patterns", *Journal of Ecology*, Vol. 57, pp. 505-514.

Visvalingam, M. 1991. "Areal Units and the Linking of Data: Some Conceptual Issues", In *Spatial Analysis and Spatial Policy Using Geographic Information Systems*, L. Worrall, ed., Belhaven Press, pp. 12-37.

Wachowicz, M. 1999. *Object-oriented design for temporal GIS*, Taylor and Francis.

Walsh, S.J. A. Moody, T.R. Allen, and D.G. Brown. 1997. "Scale Dependence of NDVI and its Relationship to Mountains Terrain". In *Scale in Remote Sensing in GIS*, Quattrochi, D.A. and M.F. Goodchild, eds., pp. 27-55.

White, R., and G. Engelen. 2000. "High-resolution integrated modelling of the spatial dynamics of urban and regional systems", *Computers, Environment and Urban Systems,* Vol. 24, pp. 383-400.

White, R., G. Engelen, and I. Uljee. 2000. "Modelling land use change with linked cellular automata and socio-economic models: a tool for exploring the impact of climate change on the island of St Lucia". In *Spatial Information for Land Use Management,* M.J. Hill, and R.J. Aspinall, eds., pp. 189-204.

Wiens, J.A. 1989. "Spatial Scaling in Ecology", *Functional Ecology*, Vol. 3, pp. 385-397.

Wolfram, S. 1984. "Cellular automata as models of complexity", *Nature*, Vol. 311, pp. 419-424.

Wu, 2004

Wu, F. 2002. "Calibration of stochastic cellular automata: the application to rural-urban land conversions", *International Journal of Geographical Information Science*, Vol. 16, No. 8, pp. 795-818.

Wu. J., and O.L. Loucks. 1995. " From Balance of Nature to Hierarchical Patch Dynamics: A Paradigm Shift in Ecology", *The Quarterly Review of Biology*, Vol. 70, pp. 439-466.

Yule, G.U., and M.G. Kendall. 1950. *An Introduction to the Theory of Statistics*. Griffin.

Zhang, X., N. A. Drake, and J. Wainwright. 2004. "Scaling issues in environmental modeling". *Environmental Modeling: Finding simplicity in complexity*. J. Wainwright and M. Mulligan, eds., Wiley, pp. 319-334.

Chapter 1.4

Geographic Object-Based Image Analysis (GEOBIA): A new name for a new discipline

G. J. Hay, G. Castilla

Department of Geography, University of Calgary, 2500 University Dr. NW, Calgary, AB T2N 1N4. Contact email gjhay@ucalgary.ca

> *"If you do not know where you are going, any road will take you there."*

Sterling Holloway (1905 - 1992)

ABSTRACT: What is Geographic Object-Based Image Analysis (GEOBIA)? To answer this we provide a formal definition of GEOBIA, present a brief account of its coining, and propose a key objective for this new discipline. We then, conduct a SWOT[1] analysis of its potential, and discuss its main tenets and plausible future. Much still remains to be accomplished.

1 Introduction

Like the 'famous' singer whom after 20 years of hard work - 'overnight' - becomes an international success, a relatively recent *paradigm*[2] *shift* in remote sensing image analysis has been stealthy taking place over the last two decades that promises to change the way we think about, analyze and use remote sensing imagery. With it we will have moved from more than

---

[1] Strengths, Weaknesses, Opportunities and Threats (SWOT)

[2] *Paradigm* refers to the generally accepted perspective of a particular discipline at a given time.

20 years of a predominantly pixel-spectra based model[3] to a dynamic multiscale object-based contextual model that attempts to emulate the way humans interpret images. However, along this new path from pixels, to objects, to intelligence and the consolidation of this new paradigm, there are numerous challenges still to be addressed. We suggest - using the terminology of Thomas Kuhn (1962) - that this *shift* corresponds to a scientific *revolution (in this context it is more appropriately – an evolution)*, that is due to a change in the basic assumptions within the ruling theory, resulting in new ideas becoming prevalent. We observe this state now, as this new technological and user driven evolution in remote sensing image analysis moves from pixels to objects and the necessary infrastructure required to generate and exploit them. To hasten a consolidation of this new paradigm, an *ontology*[4] needs to be created with a common language and understanding. By building upon previous work (Hay and Castilla, 2006), we formally propose *Geographic Object-Based Image Analysis* (GEOBIA - pronounced *ge-o-be-uh*) as the name of this new paradigm. We further propose that a worldwide GEOBIA community needs to be fostered so as to rapidly facilitate the scrutiny and dissemination of new and evolving related principles, methods, tools and opportunities.

The proceeding sections provide a GEOBIA definition and a brief account of the coining of this term, along with a recommendation for a key discipline objective. This is followed by a SWOT[5] analysis of GEOBIA's potential, and a discussion regarding its main tenets and plausible future. We note that this is but a start towards developing GEOBIA as a robust international community of practice which like the UKGEOforum[6] and newly initiated NAGeoForum[7] we envision as being vendor and software neutral. Much remains to be done.

---

[3] Here the term *model* refers to the principles, methods and tools behind traditional (i.e., prior to object-based) digital remote sensing image analysis.

[4] Here we draw upon the definition of (computer science) *ontology* which represents a rigorous and exhaustive organization of some knowledge domain that is usually hierarchical and contains all the relevant entities and their relations.

[5] Strengths, Weaknesses, Opportunities and Threats (SWOT)

[6] http://www.ukgeoforum.org.uk/

[7] North American GeoForum – 1st Meeting Sept, 2007. This is not a policy and advocacy group, but rather a structure to increase awareness, share information, improve communication and promote all that is geospatial without getting involved in national/state policies or legislation (*per.com* M.Madden, 2007 President Elect of the ASPRS).

2 What is GEOBIA? A definition

In simple terms, GEOBIA is object-based analysis of Earth remote sensing imagery. More specifically,

> *Geographic Object-Based Image Analysis* (GEOBIA) is a sub-discipline of *Geographic Information Science (GIScience)* devoted to developing automated methods to partition remote sensing imagery into meaningful image-objects, and assessing their characteristics through spatial, spectral and temporal scales, so as to generate new geographic information in GIS-ready format.

Here, GIScience refers to the science behind Geographic Information technology[8]. Since GEOBIA relies on RS (remote sensing) data, and generates GIS (Geographic Information Systems**)** ready output, it represents a critical bridge[9] between the (often disparate) raster domain of RS, and the (predominantly) vector domain of GIS. The 'bridge' linking both sides of these domains is the generation of *polygons* (i.e., *classified image-objects*) representing geographic objects. See Castilla and Hay (this book) for a detailed account of *geo-objects* and *image-objects*.

At its most fundamental level, GEOBIA requires image segmentation, attribution, classification and the ability to query and link individual objects in space and time. In order to achieve this, GEOBIA incorporates knowledge and methods from a vast array of disciplines involved in the generation and use of geographic information (GI). Indeed, it is this unique emphasis and dependency on RS and GI – and the challenges that accompany them[10] - that distinguishes GEOBIA from *object-based image analysis* (OBIA) as used in related disciplines such as Computer Vision and Biomedical Imaging, where outstanding research exists that may significantly contribute to GEOBIA.

---

[8] GIScience - http://www.ncgia.ucsb.edu/giscc/units/u002/

[9] A good example is the recent news that Definiens – the developer of the first commercial object-oriented image analysis software for remote sensing imagery (circa 2002) – has joined with ESRI – the undisputed leader in GIS software – to developed its Definiens Extension for ArcGIS, to better integrate both GIS and RS data and information (Definiens, 2007).

[10] This is due (in part) to the inherent complexity of RS imagery resulting from differences in sensor platforms, geometry, and resolution along with a host of physical based characteristics ranging from shadow, and atmosphere to target and look-angle variability, among others.

3 Why GEOBIA instead of OBIA?

Now that a new name and definition have been proposed, where and why did they originate? During July 4-5, 2006 the first international OBIA conference was held at the University of Salzburg, Austria. In total, 175 authors contributed to the conference from 24 different countries (Blaschke and Lang, 2006). As an active researcher in this topic, Dr G.J.Hay was invited as a guest speaker. While preparing his topic, he realized that there were numerous important unanswered questions that needed to be addressed. For example:

- In 2006, if you *Googled* the word OBIA, your top three matches would be: (1) *Offshore Biologically Important Area (2) Ontario Brain Injury Association and (3) Oregon Building Industry Association*. What then is OBIA? Do we have a formal definition, a set of objectives, an ontology to follow, a road map to the future?
- Why is OBIA? Object-based image research has been going on for several decades in labs throughout the world, but why now do we have the first international conference? What are the drivers behind this?
- Is there an OBIA community? If so who are we, where are we, what are we working on, how are we related to other communities, and how can we collaborate and build upon the strengths and experience of others?

In an effort to raise these concerns and to provide answers to these and other questions, Hay and Castilla (2006) presented a formal definition of OBIA and conducted a SWOT Analysis. They also committed to create a Wiki to facilitate further international discussion and development. A Wiki is a kind of 'open' website that allows users to add, remove, or edit all content very quickly and easily (Wikipedia, 2007). Within five days of the conference concluding (July 12, 2007) an OBIA Wiki[11] was created and 'open for business'. Since this time, there have been more than 6000 views of this page.

During several months of early interactive wiki discussions, a key concern was expressed. Specifically, the problem was that the term OBIA encompassed techniques used in many different disciplines, i.e., Biomedical Imaging, Astronomy, Microscopy, Computer Vision and others, yet *our* main interest – the majority of conference participants - focused on Remote Sensing and Geographic Information Systems. On July 18, 2006 in the OBIA Wiki Discussion section a user (Prashanth) suggested that by

---

[11] http://wiki.ucalgary.ca/page/OBIA (last accessed September 03, 2007)

separating the application specific components of OBIA into RS/GIS, we could create our *own* identity. He then proposed GOBIA (Geo-Object-Based Image Analysis) as the name for this new field. A number of alternative names were also suggested by others, including *OARS* (Object-based Analysis of Remote Sensing images), *GIOA* (Geographic Image-Object Analysis) and *OBARSI* (Object-Based Analysis of Remotely Sensed Imagery). After debating these and other names, on Oct 27, 2006 G.J.Hay posted to the wiki the name GEOBIA as an acronym for *this discipline*, as it builds upon OBIA roots (see Lang and Blaschke, 2006 for a brief history), while also placing – with the *GEO* pseudo prefix- an emphasis on the Geographic components this community is involved in. The more Hay and Castilla considered this name and discussed it with colleagues, the more they became convinced that it was an appropriate title for this community's identity, even if – like the other proposed acronyms - it already had an alternative meaning[12].

Hay and Castilla's argument for this new name was relatively straight forward: If the name of a discipline is intended to identify a specific community and define what they do, GEOBIA does this for our community, whereas OBIA does not. Specifically, it is unreasonable to claim the generic OBIA name only for RS/GIS applications, since there are many other communities of practice with very different objectives, data and application domains that use these techniques. To facilitate this discrimination, the term 'geographic' has been adopted as a qualifier, because it simply and elegantly distinguishes RS/GIS OBIA from these different areas. Furthermore, as a sub-discipline of GIScience (Hay and Castilla, 2006), this area of research and application requires its own unique name. Thus, based on these arguments, the acronym *GEOBIA* was selected as the heir to the successful OBIA '06 international conference: *GEOBIA, 2008 – Pixels, Objects, Intelligence. GEOgraphic Object-Based Image Analysis for the 21st Century*[13].

4 GEOBIA: A key objective

Though much of the current OBIA literature describes the use of new and or improved segmentation algorithms (see other chapters in this book), we suggest that the primary objective of GEOBIA is not that of tool building,

---

[12] *Geobia* is a genus of predator land planarians. The other proposed acronyms also had existing meanings, e.g., *Gobia* is a city in Ivory Coast, *oars* are used to propel a water craft and *obarsi* means 'origin' in Romanian.

[13] http://www.ucalgary.ca/GEOBIA/

but rather the generation of geographic information (from remote sensing imagery), from which intelligence can be obtained. Here, *intelligence* refers to geographic information that enables users to effectively perceive, interpret and respond to some specific issue[14], such as global climate change, natural resource management, Landuse/Landcover mapping, and others.

Building on these ideas, we propose that the primary objective of GEOBIA as a discipline is to develop theory, methods and tools sufficient to replicate (and/or exceed experienced) human interpretation of RS images in automated/semi-automated ways. This will result in more accurate and repeatable information, less subjectivity, and reduced labor and time costs. In return, we envision that new opportunities will be developed within emerging GI markets. For example, Wade Roush (2007) describes how *virtual* and *mirror worlds* (such as *Second Life* and *Google Earth*, respectively) will merge into what is being called the *Metaverse*, which will look like the real earth and will ….[function] as the agora, laboratory, and gateway for almost every type of information-based pursuit. In order to perform multiscale analysis and queries of geographical features and places – which are part of the fabric of this Metaverse - GEOBIA will certainly find its way here as it delineates and partitions RS images of the planet based on predefined criteria. This will be especially relevant for generating (new) temporally sensitive geographic information/intelligence. Essentially, GEOBIA provides a way to move from simply collecting images of our planet, to creating geo-intelligence[15] (as defined above).

5 Why is GEOBIA?

Now that a definition and key objective have been proposed, let's step back and examine possible reasons for its emergence? Hindsight reveals that GEOBIA exists in response to a series of drivers that have appeared over the last two decades. These include, but are not limited to:

- A change in US space policy in the early 1990's and more recently (2003) with an emphasis on fostering commercial remote sensing policy (Hitchings, 2003). This has lead to a dramatic increase in commercially available high-spatial resolution remote sensing imagery

---

[14] That is, '…geoinformation within a specific user context.'

[15] We note that as defined here, this concept is not explicitly related to *geointelligence* as specified for security purposes, though it can be used as such.

(< 5.0 m) and the need to develop new value-added markets from these multi-billion dollar investments.

- The daily generation of Terabytes of Earth Observation (EO) data, together with post September 11/2001 security issues, have provided an impetus for new (automated and semi-automated) tools to analyze/mine such voluminous data.
- An ever-growing sophistication of user needs and expectations regarding GI products.
- Recognition of limitations with pixel-based image approaches (i.e., that current remote sensing image analysis largely neglects the spatial photointerpretive elements (i.e., texture, context, shape, etc), and that increased variability implicit within high-spatial resolution imagery confuses traditional pixel-based classifiers resulting in lower classification accuracies).
- Increasingly affordable, available and powerful computing tools.
- Increasing awareness that object-based methods can make better use of neglected spatial information implicit within RS images, and provide greater integration with vector based GIS.
- Recognition of the need for multiscale approaches in the monitoring, modeling and management of our environment, for which object-based methods are especially suited.
- Recognition that object-based approaches, represent viable solutions to mitigate the *modifiable areal unit problem* (MAUP, Openshaw, 1984), since they focus analysis on meaningful geographical objects rather than arbitrary defined spatial units i.e., individual pixels.
- GEOBIA concepts and tools also have the potential to be used in the operationalization of existing ecological theories (Burnett and Blaschke, 2003) such as the *Hierarchical Patch Dynamics Paradigm* (HPDP, Wu, 1999), which provides a conceptual framework for guiding and explaining the hierarchical/multiscale structure of landscapes.

6 GEOBIA SWOT

In this section we undertake a SWOT analysis to provide insight into the current state of GEOBIA, and to outline potential strategies to achieve the stated key objective (see *Section 4*). A SWOT Analysis is (one of many possible strategic planning tools) used to evaluate the *Strengths*, *Weakness*, *Opportunities* and *Threats* involved in a project, or any other situation requiring a decision. Our objective here is to apply this method of planning early in the discipline-life cycle of GEOBIA, so that concepts described

here can be used to strengthen and guide this emerging paradigm. In practice, once an objective has been established, a multidisciplinary team representing a broad range of experiences and perspectives should carry out a SWOT analysis; which is typically presented in the form of a matrix (see-Table. 1). Thus, we invite interested individuals to share their comments by participating in the recently developed GEOBIA *Wiki*[16], so as to further facilitate this discussion.

S.W.O.T	Helpful to achieving the objective	Harmful to achieving the objective
Internal (attributes of the organisation)	**S**trengths	**W**eaknesses
External (attributes of the environment)	**O**pportunities	**T**hreats

Table 1. SWOT matrix

SWOT's are defined based on the following criteria:

- *Strengths* are *internal* attributes of the organization that are helpful to the achievement of the objective.
- *Weaknesses* are *internal* attributes of the organization that are harmful to the achievement of the objective.
- *Opportunities* are *external* conditions that are helpful to the achievement of the objective.
- *Threats* are *external* conditions that are harmful to the achievement of the objective.

---

[16] The GEOBIA wiki (http://wiki.ucalgary.ca/page/GEOBIA) was created on February 03, 2007, and has received over 2000 page views (since September, 30, 2007).

In theory, SWOTs are used as inputs to the creative generation of possible strategies, by asking and answering the following four questions numerous times:

- How can we Use each Strength?
- How can we Stop each Weakness?
- How can we Exploit each Opportunity?
- How can we Defend against each Threat?

To reap the full benefits of a SWOT analysis it is important to use this tool correctly. In particular, it is most beneficial to look at the strengths and weaknesses originating within (i.e., internal to) the discipline or organization. For example, what do we do better than anyone else, what/where could we improve, what are others likely to see as weakness? Conversely, opportunities and threats should be externally focused i.e., what trends could you take advantage of, how can you turn your strengths into opportunities, what trends could do you harm? (MindTools, 2006).

The following sections represent a number of SWOTs identified as we considered the past, present and future of GEOBIA. They are by no means the only possible items, and in several cases – depending on one's perspective – individual items could exist in more than one category.

6.1 GEOBIA Strengths

- Partitioning an image into objects is akin to the way humans conceptually organize the landscape to comprehend it.
- Using image-objects as basic units reduces computational classifier loads by orders of magnitude, and at the same time enables the user to take advantage of more complex techniques.
- Image-objects exhibit useful features (e.g., shape, texture, contextual relations with other objects) that single pixels lack.
- Image-objects are less sensitive to MAUP than units that do not keep a correspondence with the structure of the phenomenon under study.
- Image-objects can be more readily integrated into a vector GIS than pixel-wise classified raster maps.
- The number of both free and commercially available GEOBIA software is increasing steadily (see Neubert et al., – this book).

6.2 GEOBIA Weaknesses

▪ Under the guise of 'flexibility' some commercial object-based software provides overly complicated options, resulting in time-consuming analyst 'tweaking'.

▪ There are numerous challenges involved in processing very large datasets. Even if GEOBIA is more efficient than pixel-based approaches, segmenting a multispectral image of hundreds or thousands of mega-pixels is a formidable task, thus efficient tiling/multiprocessing solutions are necessary.

▪ Segmentation is an *ill-posed problem*, in the sense that it has no unique solution, e.g., (i) changing the bit depth of your heterogeneity measure can lead to different segmentations. (ii) Even human photo-interpreters will not delineate exactly the same things.

▪ There is a lack of consensus and research on the conceptual foundations of this new paradigm, i.e., on the relationship between image-objects (segments) and landscape-objects (patches). For example, (i) what is the basis to believe that segmentation-derived objects are fine representations of landscape structural-functional units? (ii) How do you know when your segmentation is good? (iii) Is there a formally stated and accepted conceptual foundation?

▪ There exists a poor understanding of scale and hierarchical relations among objects derived at different resolutions. Do segments at coarse resolutions really 'emerge' or 'evolve' from the ones at finer resolutions? Should boundaries perfectly overlap (coincide) through scale? Operationally it's very appealing, but what is (if any) the ecological basis for this, and is such a basis necessary?

6.3 GEOBIA Opportunities

▪ *Object-Oriented* (OO) concepts and methods have been successfully applied to many different problem domains, not only computer languages, and they can be beneficially adapted to GEOBIA. This integration not only includes OO programming, but all the corpus of methods and techniques customarily used in biomedical imaging and computer vision (among others) that remain unknown to most of the remote sensing community.

▪ There are new information technology tools (e.g., Wikis) that may accelerate consensus and cohesion of a GEOBIA community.

▪ There is a steadily growing community of RS/GIS practitioners that currently use image segmentation for different GI applications. Thus,

as GEOBIA matures, new commercial/research opportunities will emerge to tailor object-based solutions for specific user needs i.e., forestry, habitat and urban mapping, mineral exploration, transportation, security, etc.

- Symmetric multiprocessing, parallel processing and grid computing are recent technologies that GEOBIA methods may build upon to tackle problems related to the analysis of large datasets.
- Adopting existing open GIS programming standards like Open Source GIS[17], and guidelines and methods from the Open Geospatial Consortium[18] along with provisioning to include Semantic Web[19] standards within current and new GEOBIA tools, will allow for re-use and integration between different platforms and data types, and opportunities for a web-wide dissemination and evaluation of image-object semantics. This in return will provide value-added opportunities for the sharing and generation of new GI, and the ability to build on expertise from different user communities throughout the globe.

6.4 GEOBIA Threats

- The visual appeal of image-objects, their easy GIS-integration and their enhanced classification possibilities and information potential have attracted the attention of major RS image processing vendors, who are increasingly incorporating new segmentation tools into their packages. This provides a wider choice for practitioners, but promotes confusion (among different packages, options, syntax, etc) and makes it more difficult to develop a cohesive GEOBIA community. Will a lack of protocols, formats, and standards lead to a segmentation of the field rather than a consolidation? Castilla and Hay (this book) refer to this critical GEOBIA threat as *The Tower of Babel problem* - where every user group develops different terminology than every other group for the same meaning, or the same term with different meanings

---

[17] http://opensourcegis.org/

[18] http://www.opengeospatial.org/

[19] At its core, Tim Berners-Lee's Semantic Web comprises a philosophy, a set of design principles, collaborative working groups, and a variety of enabling technologies. Some elements of the semantic web are expressed as prospective future possibilities that have yet to be implemented or realized. (http://en.wikipedia.org/wiki/Semantic_Web)

resulting in confusion, isolation, a fighting over standards, and a negative or flat progression of the discipline[20].

- Trying to make distinct GEOBIA from other OO concepts and methods (e.g., by using terms such as 'object-based' instead of 'object-oriented') may contribute to insulation (of users in an esoteric world of 'objects') and isolation (of the concept) rather than to consolidation.

- GEOBIA is far from being an established paradigm, yet many users of commercial segmentation software do not recognize this fundamental fact. GEOBIA <u>is not</u> one specific research or commercial software. Much still remains to be solved and discovered.

7 GEOBIA Tenets

Based on these SWOT items, we offer the following GEOBIA tenets as fundamental components of what we currently see this discipline as, and as guides to what it could become.

GEOBIA is…

- *Earth centric* – its data sources originate from the surface of this planet.
- *Multi-source capable* – its methods provide for the inclusion of multiple different digital data types/sources within a common geographic referent and for the flow of information and intelligence from pixel-based RS data to GIS ready polygons.
- *Object-based* – meaningful image-object delineation is a prerequisite of this approach, from which relevant intelligence can be generated.
- *Multiscale* – a scene is often composed of objects of different size, shape and spatial location, thus multiscale analysis both within a hierarchical level and between levels is essential. Because GEOBIA is multiscale, potential exists to model alternative 'multiscale' realities based on selective user defined aggregations of fine scale segments and or their attributes.
- *Contextual* – it has the ability to *incorporate* or *integrate* 'surrounding' information and attributes. When processing RS data, this in-

---

[20] Tim Berners-Lee's development of the Semantic web represents a significant effort to resolve this kind of issue at a global level, but at a local level, semantic and ontological standards need to be developed, shared, and agreed upon within specific application domains, and tied to semantic web formats.

cludes mechanisms to quantify an object's photointerpretive elements i.e., colour (hyperspectral), tone, size, shape, pattern, location, and texture. By adding time (multitemporal imagery), as well as other attributes such as height (Lidar) and heat (Thermal) into the 'contextual pool', there will be a greater information potential for each image-object than ever possible for individual pixels.

- *Adaptive* – it allows for the inclusion of human semantics and hierarchical networks – whether through experts systems, or expert interpreters, so that analysis may be tailored to specific user needs. However, to be fully adaptive, GEOBIA tools need to build on existing Open GIS standards and provide mechanisms to integrate user and domain specific ontologies into a semantic web so as to globally facilitate improved sharing, integration and generation of new synergistic GI and the development of their associated markets. For example, given a multispectral image, one user may derive a general Landuse classification. This information may then be shared/sold across the web to another who builds upon this information to single-out patches of Amazon Acai palm trees for lumber and thatching materials. This new GI layer may then be shared or sold to another user in a different part of the world, and when combined with their expertise and knowledge could provide opportunities to harvest Acai berries and explore them as a ground-breaking leukemia cure (based on a true scenario). In this case, one initial dataset, combined with a sharing of semantic information results in myriad different GI products and markets.

8. Conclusion

GEOBIA is a sub-discipline of GIScience devoted to developing automated methods to partition remote sensing imagery (of our planets surface) into meaningful image-objects, and assessing their characteristics through scale. Its primary objective is the generation of geographic information (in GIS-ready format) from which new intelligence can be obtained.

In this paper we have formally defined *Geographic Object-Based Image Analysis* (GEOBIA); provided a justification for this new name; outlined a key objective of this new discipline; identified a number of Strengths, Weakness, Opportunities and Threats (SWOT) that GEOBIA faces; and proposed a number of GEOBIA tenets. It is important to note that a key issue faced by this new discipline is to ensure that an integrative, well understood, and easily defined ontology is developed and incorporated within

the research and commercial software that is currently being built and used. A way to construct and promulgate such ontology is by creating a living document – a GEOBIA guide book - to which practitioners can contribute and turn to for understanding and direction. This could further be facilitated by adopting established Open GIS standards and semantic web protocols, so that geoinformation could be shared and integrated more easily. While this tome on OBIA represents a positive beginning, we propose that the existing GEOBIA *Wiki* – with its world wide accessibility - is an ideal vehicle to develop such a guide, and cordially invite all interested parties to participate in building a stronger GEOBIA community of practice.

Acknowledgements

References

Blaschke T, Lang S (2006) Object-Based Image Analysis for Automated Information Extraction – A Synthesis. Proceedings of MAPPS/ASPRS. November 6-10. San Antonio, Texas

Burnett C, Blaschke T (2003) A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecological Modelling 168(3): 233-249

Castilla G, Hay GJ (2007) Image-objects and geographic-objects. In: Blaschke T, Lang S, Hay GJ (Eds) Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications. Springer-Verlag, Berlin

Definiens (2007) Definiens Bridges the Gap between Remote Sensing and GIS. Press release, Munich, Germany/San Diego, CA, USA – June 18th. URL: http://www.definiens.com/news_dtls.php?id=117&link_id=27&sublink_id=30&PHPSESSID=82f5213ef3acbfd320224e4a90900dcc

Hay GJ, Castilla G (2006) Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT). International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. No. XXXVI-4/C42

Hitchings S (2003) Policy assessment of the impacts of remote-sensing technology. Space Policy 19:119-125

Kuhn TS (1962) The Structure of Scientific Revolutions. The Chicago University Press, Chicago

Lang S, Blaschke T (2006) Bridging remote sensing and GIS – what are the main supporting pillars? International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42

MindTools (2006) SWOT Analysis. URL:
    http://www.mindtools.com/pages/article/newTMC_05.htm

Neubert M, Herold H, Meinel G (2007) Evaluation of remote sensing image segmentation quality. In: Blaschke T, Lang S, Hay GJ (Eds) Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications. Springer-Verlag, Berlin

Openshaw S (1984) The Modifiable Areal Unit Problem. *Concepts and Techniques in Modern Geography* (CATMOG), vol. 38,. GeoBooks, Norwich, pp. 40

Roush, W (2007) Second Earth. Technology Review July/August Issue. MIT Press.URL:
    http://www.technologyreview.com/printer_friendly_article.aspx?id=18888

Schiewe J, Tufte L, Ehlers M (2001) Potential and problems of multi-scale segmentation methods in remote sensing. Geographische Informationssysteme 6: 34-39

Wikipedia (2007) Wiki. URL: http://en.wikipedia.org/wiki/Wiki

Wu J (1999) Hierarchy and scaling: extrapolating information along a scaling ladder. Can J Remote Sensing 25(4): 367-380

Chapter 1.5

Image objects and geographic objects

G. Castilla, G. J. Hay

Department of Geography, University of Calgary, 2500 University Dr. NW, Calgary, AB T2N 1N4. Contact email: gcastill@ucalgary.ca

"A map is not the territory it represents, but if correct, it has a similar structure to the territory, which accounts for its usefulness"

Alfred Korzybski, Science and Sanity (1933)

ABSTRACT: Object-Based Image Analysis (OBIA) has gained considerable impetus over the last decade. However, despite the many newly developed methods and the numerous successful case studies, little effort has been directed towards building the conceptual foundations underlying it. In particular, there are at least two questions that need a clear answer before OBIA can be considered a discipline: i) What is the definition and ontological status of both image objects and geographic objects? And ii) How do they relate to each other? This chapter provides the authors' tentative response to these questions.

1 Introduction

The inability of traditional pixel-based methods to cope with recent very high resolution satellite imagery (VHR, < 5 m pixel size) is fostering the widespread adoption of OBIA methods. This trend may be interpreted as a *paradigm shift* (Khun 1962) that can be explained by placing it within a historical perspective. Civilian spaceborne *remote sensing* (RS) of land-

scapes[1] began in 1972 with the launch by NASA of the ERTS-1 (Earth Resources Technology Satellite, later renamed *Landsat-1*). The spatial resolution (80 m) provided by the *Multi Spectral Scanner* (four bands) on board this satellite, though a technological milestone, was insufficient to individually resolve the recurrent elements (such as trees or buildings) characterizing common landcover classes[2]. The way to circumvent this was to assume that different landcover classes behaved like distinct surface materials capable of being analyzed with a spectrometric approach[3]. Thus it was natural to treat each pixel as a sample introduced in a desktop spectrometer, and as a result, the individual pixel was enshrined as the basic unit of analysis. Several digital classification methods (e.g., the *maximum likelihood classifier*) were developed based on this approach and soon after became the accepted paradigm in the analysis of RS imagery. The fact that *pixels do not come isolated but knitted into an image full of spatial patterns* was left out of the paradigm, since the spatial structure of the image could only be exploited manually by human interpreters. Despite this shortcoming, the pixel-based paradigm remained unchallenged for almost three decades, until a critical question was finally posed: *why are we so focused on the statistical analysis of single pixels, rather than on the spatial patterns they create?* (Blaschke and Strobl 2001).

There were two triggers to questioning the pixel-based paradigm (Lang and Blaschke 2006), namely the advent of VHR civilian satellites, and the debut in 2000 of the *eCognition* software (see Benz et al. 2004). The availability of this first commercial OBIA software provided worldwide access to tools that previously existed only in research labs. On the other hand, in VHR imagery, individual pixels are too small to be representative of the settings to which common landcover classes refer. That is, pixel-based classification requires a pixel footprint large enough to encompass a representative number of the recurring elements defining each class (Woodcock and Strahler 1987). Since this requirement cannot be satisfied for most classes using VHR imagery, a different approach is required. OBIA has emerged as an alternative to the traditional pixel-based paradigm, and is

---

[1] In this paper the term *landscape* is used exclusively to refer to a portion of solid Earth surface on the order of 1 to 1000s of km², and not to the view of that area. This distinction is important for reasons that will be explained in section 4.

[2] Note that a higher spatial resolution would have implied vast data volumes and faster downlinks, better optics and electronics, and improved platform control, which were not then feasible (Landgrebe 1999).

[3] Where the *digital numbers* (DNs) of each pixel in each band form a profile or *spectral signature* that is classified according to its similarity to the typical signatures (joint reflectance profile from representative areas) of the classes of interest.

based on the idea that by shifting the basic units from pixels to image objects, we can emulate (or exceed) visual interpretation, making better use of spatial information implicit within RS images while also providing greater integration with vector based GIS (Hay and Castilla 2006).

OBIA assumes that we can identify entities in remote sensing images that can be related to real entities in the landscape. A first step in this kind of analysis is *segmentation*, the partitioning of an image into a set of jointly exhaustive, mutually disjoint regions that are more uniform within themselves than when compared to adjacent regions. These regions (a.k.a. *segments*) are later related to geographic objects (such as forests and lakes) through some object-based classification. During the last decade, OBIA techniques for grey-level images have significantly improved as a result of research efforts in the fields of *Computer Vision* and *Biomedical Image Analysis* (Kartikeyan et al. 1998). However, these results have only partially been transferred to remote sensing, mainly due to the lack of typical shape (and even crisp boundaries) of the objects of interest, and to the multi-band and multi-scale nature of the images (Schiewe et al. 2001). These special characteristics, and the particular portion of reality towards which the analysis is oriented (i.e., the geographical domain), make necessary to draw a distinction between OBIA for biomedical imagery and OBIA for RS imagery. Hence we recommend the acronym GEOBIA (*Geographic Object-Based Image Analysis*) as the name of this field of expertise (Hay and Castilla, this volume), which we note has already been adopted for the 2nd International Conference[4] on this subject. Despite this recommendation, we will continue to use the 'old' acronym here, so as to keep consistency with the other chapters of this book.

OBIA has recently been conceptualized as a new sub-discipline of *GIScience* that uses as basic units computer-delineated regions derived from remote sensing imagery (Hay and Castilla 2006). As such, OBIA faces what Barry Smith (2003) calls the *Tower of Babel problem*. Being a burgeoning field, there are many groups in the world that actively work on OBIA. Each group may be using different terms than other groups for the same meaning, or the same term with different meanings. As ever more diverse groups are involved in OBIA, the problems of consolidating this field into a single system increase exponentially (Smith 2003). Consequently, the sooner a consensus is reached on key terms, their meaning and their relationships (between each other and with the world), the easier this problem can be tackled. In order words, **we urgently need to build an OBIA ontology**.

---

[4] Calgary, Alberta, 6-8 August 2008 (www.ucalgary.ca/GEOBIA)

The goal of this chapter is to recommend a definition of two key terms within OBIA, namely *image-object* and *geographic-object* (or *geo-object*), and to investigate their relationship with one each other and with the real world. They are the basic units of this approach, one belonging to the image domain and the other to the geographic domain. As such, they must be operationally defined in order for OBIA to be considered a discipline.

2 Image-objects

We tentatively define '**image-object**' as *a discrete region of a digital image that is internally coherent and different from its surroundings*. We note that image-objects have being defined by the developers of *eCognition* as '*contiguous regions in an image*' (Benz et al. 2004). The latter definition has the merit of simplicity. However, it grants the status of image-object to arbitrary regions that keep no correspondence with the spatial structure of the image. For example, the pieces of a jigsaw puzzle, being contiguous regions in a picture, would qualify as image-objects. Such possibility is at odds with what an image-object should intuitively be, i.e., a portion of an image that could be seen as a distinct entity if delineated. Therefore, a more sound definition has to include some perceptual constraints. More specifically, there are three traits that a region must possess in order to qualify as an image-object: (1) *discreteness*; (2) (internal) *coherency*; and (3) (external) *contrast*, which are discussed in the next subsections.

2.1 Discreteness

An image-object requires explicitly defined limits in order for a computer to be able to manipulate it. This might not be required for cases where the membership of a pixel to an image-object is *fuzzified*. However we strongly discourage such treatment for reasons explained in section 3.2, and hence will not consider it. In a raster environment, *discreteness* implies that a new image has to be created out of the original one, where each new pixel has as *digital number* (DN) the numeric identifier of the image-object within which it is included. Since in the new image all the pixels within an image-object have the same DN, the data volume required to represent this partition can be considerably reduced by converting it to (polygon) vector format. In fact, one of the advantages of OBIA is that it facilitates the integration of the result of RS image analysis into vector GIS.

Note that an image-object may contain in its interior some smaller image-objects. This would be the case of a 1 m resolution image representing a plan view of an orchard surrounded by a corn field. In this image, there would be two large image-objects, one corresponding to the corn field and another to the orchard soil. In addition, there would be as many small image-objects as there are trees in the orchard. In this case, it is preferable to conceptualize the small image-objects as gaps or holes rather than parts of the larger (soil) image-object. The reason is that image-objects are perceptual entities that have no meaning attached until they are classified. Hence the smaller and larger image-objects cannot be considered part of the same semantic unit (orchard) until they have been classified respectively as (representations of) trees and soil.

Finally, a related question is whether the region corresponding to an image-object has to be *unitary*, i.e., can the region consist of several disjoint parts? Or contrariwise, do all the pixels belonging to it have to be connected? In principle, the region has to be unitary to be considered an image-object. However, in a multiscale framework, adjacency and connectedness are scale-dependent, so this condition can be relaxed.

2.2 Coherency

Coherency means that for a region to qualify as an image-object, the pixels inside it have to perceptually *stick together*. Note that this does not necessarily imply homogeneity (low variance). A region that shows coarse texture or some geometric pattern may contain high variance and yet be coherent, providing the texture or pattern is uniform throughout its extent. These recurrent patterns can be encoded into a *structural signature* (Lang and Laganke 2006), i.e., a set of attributes and formal relations characterizing spatial patterns observable at certain scales that can be used in conjunction with *spectral signatures* (typically, the mean in each band of the pixels within an image-object) to discriminate between different classes of objects.

2.3 Contrast

Not only should there be some sense of uniformity within the region corresponding to an image-object, but there also should be some *contrast* with the surroundings so that the region can be perceived as a distinct entity. Contrast may be produced by a difference in color, tone or texture between the interior of the region and its exterior, or, when the latter has the same appearance, by the existence of an edge separating the region. Such would

be the case of an image-object representing an agricultural parcel having the same crop as the adjacent parcels, but that is separated from them by a thin strip of barren soil.

2.4 Image-segments as image-objects

Since OBIA aims to emulate human interpretation of RS images through computer means, the term *image-object* has to be explicitly related to the output of some digital procedure. Since image segmentation is the customary technique used to derive initial units within OBIA, a possible solution is to equate image segments to image-objects. Indeed, segments can be considered candidate image-objects, at least from the point of view of the segmentation algorithm: they are discrete; they are internally coherent since they have passed the homogeneity criteria of the algorithm; and they are different to some degree with their surroundings, otherwise they would have being merged with some neighbor. However, a human observer may judge their coherency and contrast differently, and not just because perceived chromatic differences are not isometric to the usual metrics employed as dissimilarity measures[5].

There is a two-sided problem in image segmentation that describes these judgmental differences. Specifically, (1) *Oversegmentation* (**Fig. 1a**) refers to a situation where, in the opinion of the perceiver, the contrast between some adjacent segments is insufficient and should be merged into a single image-object. (2) *Undersegmentation* (**Fig. 1.b**) refers to the existence of segments that in the opinion of the perceiver lack coherency and should be split into separate image-objects. In general, oversegmentation is less serious a problem than undersegmentation, since aggregating segments *a posteriori* is easier than splitting them. Also, since there is no straightforward relationship between similarity in the image and semantic similarity, it is preferable to err on the side of oversegmentation and relax the external contrast requirement. In short, a *good* segmentation is one that shows little oversegmentation and no undersegmentation, and a *good* segmentation algorithm is one that enables the user to derive a good segmentation without excessive fine tuning of input parameters. Thus, informally, *image-objects*

---

[5] The lack of isometry is due, on the one hand, to the varying sensitivity of the human eye to different wavelengths within the visible part of the spectrum; and on the other hand, to the different contrast enhancements that may be applied to display the image. In the case of multi-band images, there will also be differences in judgment between the interpreter and the algorithm, since the former can only visualize three bands at a time. However, this can be partially circumvented by visualizing different band combinations.

could be redefined as *the segments derived from a good segmentation algorithm*. The goodness of such algorithms has to be established on the basis of expert consensus. In this respect, we note there is already an initiative devoted to create such a basis (Neubert et al. 2007, this volume).
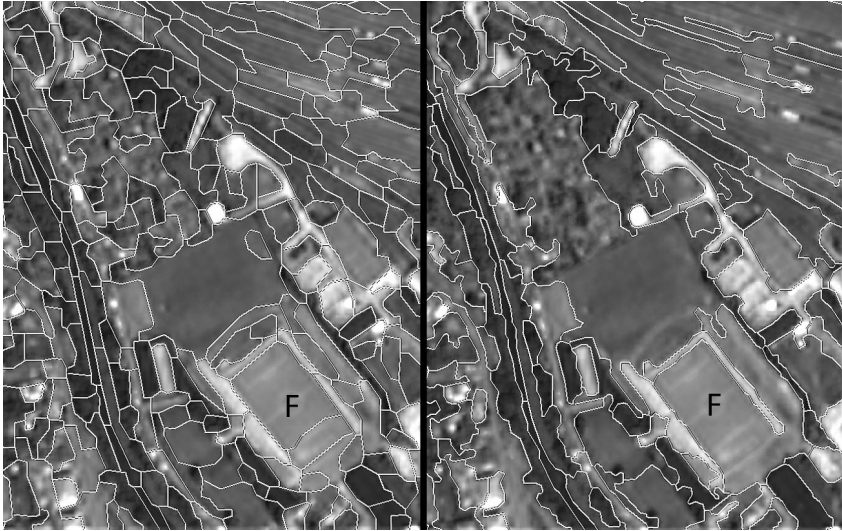
Fig. 1. A 1 m resolution satellite image representing an urban area with two segmentation levels overlaid: left (**Fig. 1a**), finer; and right (**Fig.1b**), coarser. The soccer field (F) has been oversegmented in A and undersegmented in B

2.5 Do image-objects exist independent of the viewer?

Regarding the ontological status of image-objects, it has to be noted that they do not exist *autonomously* within digital images. Rather, they are created during the segmentation as we will see. A digital image is an array of numbers that, when mapped to a planar display device such as a computer monitor, produces a visual phenomenon consisting of spatial patterns of various colors or tones and shapes. These patterns can in turn be mapped back to the array so as to establish the location, in terms of 2D coordinates (columns and rows in the array), of the pixels yielding each pattern. This can indeed be viewed as the goal of segmentation. If a given segment corresponds to a pattern that is distinct and unitary, then it can be considered an image-object. However, each demarcated image-object is dependent on the particular process used to single it out, so it owes its existence to an external decision made by a human or by a machine programmed by a hu-

man. In other words, image-objects are *fiat objects* (Smith 2001), they are created by human cognition. Strong evidence in support of this view is the multiplicity of 'good-looking' partitions that can be derived from the same image using different segmentation algorithms or even using the same algorithm and slightly different parameters.

2.6 Is information extracted or produced from images?

It is important to note that the result of the segmentation is the imposition of a simpler, hopefully meaningful structure upon the intricate array of numbers, i.e., a formal representation (a model) of the spatial structure of the image. This model is dependent not only on the data alone, but also on the chosen similarity criteria and on the intended level of detail (which can be encapsulated by the size distribution of segments) of the representation. Different choices during the segmentation will yield different representations (and hence different information) from the same data set. So *information is not extracted from the image*, as the common misconception states, *but produced during the analysis* (Castilla 2003). Actually the word *information* comes from the Latin verb *informare* ('to give form'), which implies the realization of structure upon some material. Thus, the image can be seen as the raw material upon which the model is carved out. The final form of the model is dependent not only on the material, but also on the analyst, be it human or machine. However, this creative freedom is by no means unrestricted. The model has to resemble the piece of reality it refers to. That is, the partition has to capture the spatial structure of the image displayed on the monitor. Indeed it is this correspondence which accounts for the usefulness of the resulting segments. We note that this is a separate issue from the correspondence between the image and the imaged landscape, which will be treated in section 4.1.

3 Geo-objects

Intuitively, a geographic object, or *geo-object*, is an object of a certain minimum size (as to allow representation in a map) on or near the surface of the Earth, such as a city, a forest, a lake, a mountain, an agricultural field, a vegetation patch and so on (Smith and Mark 1998). Here, the term 'object' refers to a discrete spatial entity that has many permanent properties which endow it with an enduring identity and which differ in some way or another from the properties of its surroundings. If we embrace *realism* (i.e., there exists a single physical reality which can be truthfully ob-

served via our *sensorium* or other artificial apparatus, and whose existence is independent of human cognition), we have to agree that the interiors of these objects are '*autonomous portions of autonomous reality*' (Smith 2001). However, many of these objects are delimited, at least in some portions of their perimeter, by boundaries that exist only in virtue of human *fiat* and that may not correspond to any observable discontinuity on the ground. Therefore we will consider that most geo-objects are again *fiat-objects sensu* Smith (2001).

Imagine we are mapping the distribution of landcover in a rural area that includes a forest surrounded by an agricultural field to the East, while its eastern end gives way to sparse woodland. There would be little dispute about where the boundary between the forest and the field should be placed. However, in the eastern part, the boundary will not be self-evident and will depend on how we define 'forest' and 'sparse woodland'. Notwithstanding, if the final boundary of the forest has been delineated in an agreeable way, there would be a consensus that, given the definition of 'forest' and the local configuration of the landscape in that area, the region enclosed by that boundary is a proper instance of the type of object 'forest'.

Thus, a '**geo-object**' can be better defined within OBIA as *a bounded geographic region that can be identified for a period of time as the referent[6] of a geographic term* (Castilla 2003). The latter is a noun or expression that refers to anything that can be represented in a map, such as 'glacier' or 'mineral extraction site'. The adjective 'geographic' implies that the map necessarily represents a portion of the earth surface. Under this view, typical geographic terms are those included in map legends, such as 'forest', 'sparse woodland', and 'lake', which are more conformant to what a geographer would consider a geographic object. However, terms included in the definition of some entry of the legend may also be considered *geographic terms,* providing they can be mapped at a finer scale, even if their referents will not be represented individually in that particular map.

For example, the term 'tree' is included in the definition of both 'sparse woodland' (e.g., *an area covered by shrubs and scattered trees*) and 'forest' (e.g., *an area densely covered by trees*). Since a tree crown can indeed be mapped at fine scales, 'tree' qualifies as a geographic term under our definition, and therefore trees are also geo-objects. We acknowledge that

---

[6] In *semiotics*, a *referent* is the 'real world' object to which a word or sign refers, i.e., an instance of a type of object. We prefer using 'referent' instead of 'instance' because the former makes clearer the distinction between a real object (a geo-object of a certain type) and its representation (the image-object classified as a representational instance of this type of geo-object, see section 4).

from a geographer's standpoint, it is arguable whether these smaller objects can be considered geographic objects, as in general they do not have location as an ontological mark. However, within OBIA, we need to account for them, as (1) they are 'visible' in VHR imagery; and (2) in order to identify a large image-object as a (representational) instance of e.g. the class 'orchard', we need first to identify a number of smaller image-objects as (representational) instances of a type of *sub-object* called 'tree'. A sub-object in this context is a referent of a term that is part of the definition of a type of geo-object listed in the map legend, where this term (e.g., 'tree') is itself not present in the legend as a separate entry.

Turning back to the definition of geo-object, it is important to note that the expression '*for a period of time*' is included not only to stress that as the landscape evolves, the region to which the term refers may change; but also to convey the idea that the region and the object are not the same thing. That is, a forest can change its shape, shrink or even disappear, but a geographic region necessarily has the shape and size it has (Casati et al. 1998). This temporal bound also accounts for mobile objects such as cars and ships, which for a period of time (at least during the image acquisition time) occupy a given region of geographic space, and in this sense they are also geo-objects.

Tangible objects such as cars, trees and buildings differ from conventional geographic objects not only in size, but also in the fact that they are *bona fide* objects, i.e., they have boundaries that are independent of human cognition, and hence they are ontologically autonomous entities (Smith 2001). In contrast, larger geographic objects such as forests and cities have boundaries that contribute as much to their ontological make-up as do the constituents composing their interior (Smith and Mark 1998). Therefore their existence as objects is dependent on the cognitive activity of demarcating their boundaries. Notwithstanding, this relative mind-dependence does not entail a serious ontological problem. To put it plainly, the fact that your hometown is a *fiat object* does not entail that your life is a sort of *Truman show*[7]. From an epistemological perspective, it neither diminishes at all their usefulness. On the contrary, it is far more economical thinking in terms of separate wholes (objects) with distinct enduring properties rather than in a *continuum* of connected plots or raster cells whose content has to be determined cell by cell every time (Frank 2001).

---

[7] This 1998 movie (directed by Peter Weir, written by Andrew Niccol, and starring Jim Carrey) accounts the life of a man who does not know that his entire life is a constructed reality designed for the purpose of a television show.

3.1 Geo-objects and the hierarchical structure of landscapes

Geo-objects can serve as nested structural units to model a complex system such as the landscape. Under this view, the latter can be represented as a nested hierarchical system having both vertical structure –composed of several discrete *levels of detail*[8] (LOD); and horizontal structure –given by the different geo-objects existing at each LOD. Note that the vertical decomposability is limited by *size*, since depending on the cartographic scale (selected for representing of the finest LOD), objects below a certain extent cannot qualify as geo-objects even if they are on the Earth surface. However this threshold is application dependent, being smaller e.g. for archeologists than for geographers. Thus the term 'geo-object' is subject to what Bennett (2001) called *sorites vagueness*[9]. It should also be noted that the idea of considering the landscape as a nested hierarchical system is by no means new (Woodcock and Harward 1992; Wu and Loucks 1995). However we envision that OBIA can provide a powerful framework to make this idea operational, as some have already attempted (Burnett and Blaschke 2003).

There are two primitive relations that can be used to construct such hierarchical models: (1) the *part_of* relation, which operates in the geographical domain and creates a *partonomy* providing *encapsulation* (where sub-objects belonging to a finer LOD are hidden); and (2) the *kind_of* relation, which operates in the categorical domain and creates a *taxonomy*, providing *inheritance* (where subclasses inherit the features of their parent class). The partonomy is constructed by identifying regions of the geographic space that can be seen as *referents* of the terms included in the taxonomy, or stated in the reverse direction (from mind to world), it is built by *projecting* these terms onto the landscape.

We note that both taxonomies and partonomies are *granular partitions* (Smith and Brogaard 2002), where the adjective 'granular' refers to the

---

[8] We prefer the term 'level of detail' to 'hierarchical level' because the former includes both semantic and cartographic (representational) aspects, and therefore conveys better the fact pointed out by Wiens (1995) that hierarchies are human constructs. The passage from one level to the next involves complex operations of both semantic and cartographic generalization, which are beyond the scope of this chapter.

[9] Referring to the sorites (from *soros*, 'heap' in greek) paradox, also known as *little by little arguments*. It was one of a series of puzzles attributed to Eubulides of Miletus (IV B.C.): Would you describe a single grain of wheat as a heap? No. And two? No. And three? No... You must admit the presence of a heap sooner or later, so where do you draw the line? For a detailed treatment of this problem in context with GIS, see Fisher (2000).

possibility of identifying objects without having to recognize all their constituent parts. Granular partitions are cognitive devices in the form of rooted graphs without cycles and without upward (in the direction from leaves to root) bifurcations within which the objects of interest are nested. They help us to capture in a synthetic way landscape complexity by dividing it into meaningful 'chunks' that we call *geo-objects*. Since pattern (structure) and process (function) are intimately related, these chunks also provide a key to understanding landscape dynamics. In particular, their boundaries can be seen as phase transitions in a physical system, since they are deemed to be the places where the qualities of the landscape change in some relevant respect. Consequently, these boundaries also act as filters, changing the intensity and/or frequency of the interactions between the different agents dwelling at each side of them.

3.2 The problems of demarcating geo-objects: boundary indeterminacy and vagueness of geographic terms

In order to make operational our definition of geo-object, we need to be able to demarcate regions of the territory that can be seen as referents of some geographic terms. However this is problematic, mainly for two reasons. (1) The landscape, especially in natural areas, is structured into continuously varying patterns that often defy crisp boundary placement at a fixed scale (Burnett and Blaschke 2003). That is to say that the width of the transition zone between two neighboring landscape settings for which we have a name (e.g., 'forest' and 'sparse woodland') is often several times greater than the width (at 1:1 scale) of the cartographic line that separate the objects representing the two settings[10]. (2) The terms themselves are vague, and this vagueness gives way to a multiplicity of candidate regions that could stand as the legitimate referent of say the term 'forest' in that area.

With regard to this, we strongly recommend that OBIA methods adopt Varzi's (2001) view of vagueness as *de dicto* ('belonging to the words'), and consequently reject the *de re* ('belonging to the things') view of *fuzzy*[11]

---

[10] Note that a line drawn in a map cannot be infinitely thin, so it will have a certain width when transferred to the ground.

[11] Fuzzy set theory (Zadeh 1965) permits the gradual assessment of the membership of elements in relation to a set with the aid of a membership function that maps the degree of belonging onto the interval [0,1]. Note that our recommendation applies only to geographic space. That is to say, categorical space may well be modeled using fuzzy membership functions between attributes and object classes.

methods. That is to say, *vagueness should be treated not as an inherent property of geographic objects but as a property of the terms we use to refer to them*. Vagueness, rather than a defect of language, is both an economic and epistemic need. On the one hand, economy of language facilitates communication without cumbersome additions required to achieve precision. On the other hand, '*precision decreases the certainty of propositions*'[12], so a certain amount of vagueness is required in order to ensure the truth of our statements about the world. The reason for choosing the *de dicto* view is that the other (*de re*) requires further ontological commitments on the nature of fuzzy objects, complicating their topological relations. Besides, the computation of membership functions for each cell of a grid representing the territory is an overly intensive task, and leads to statements (e.g., 'this point of the Himalayas is 40% part of Mount Everest, 35% Mount Lhotse and 25% part of the valley') that are at odds with our entity view of geographical phenomena (Bittner and Smith 2001).

The way to tackle the first problem (*indeterminacy of boundaries*) is to embed the analysis in a multiscale framework. On the one hand, there is no unique spatial resolution appropriate for the discrimination of all geographical entities composing a landscape (Marceau et al. 1994); and on the other, boundary saliency is scale dependent (Hay et al. 1997). By observing the landscape at several scales, some areas where the phenomenon of interest (e.g., vegetation) does not show a distinct pattern at a given scale may become apparent at finer or coarser scales. In addition, this approach allows for the modeling of hierarchical relationships between nested objects, e.g., a tree that is part of an urban park that is part of a city. In order to identify the geo-object 'urban park', we not only have to identify a large number of trees in its interior, but we also have to realize that the object is surrounded by clusters of buildings. Since the typical size of tree crowns and building roofs differs considerably, their correct identification clearly requires different observational scales.

In order to tackle the second problem (*de dicto vagueness*), the definition of each geographic term included in the map legend needs to be more precise than the one from the dictionary. In this way the multiple candidate regions that could stand as suitable referents of each term in a given area will be more coincident than when using a broader definition. In other words, the penumbra cast by the projection of a given term (i.e., the annular region that encompasses the area of intersection between all reasonable instantiations of a particular geo-object that can be described using this

---

[12] This is known as *Duhem's Law of Cognitive Complementarity* (Duhem 1906).

term) would be reduced. If we take the *egg-yolk representation*[13] (Cohn and Gotts 1996) as the framework to describe the indeterminacy of boundaries, then what we are saying is that using precise definitions of the items in the map legend leads to a shrinkage of the white of the egg that encompasses all suitable delineations of a given geo-object.

4 Linking image-objects to geo-objects

Having defined both image-objects and geo-objects, we can investigate how they relate to each other. The first thing that must be noted is that this relationship is not one of identity. Image-objects are at best representations of geo-objects. However we usually neglect this, because we tend to *reify* the image, i.e., we equate it to a window opened in the floor of a balloon gondola from which we contemplate the real scene, and therefore we assume that what we see in the image are real objects instead of representations of real objects. This delusion was masterly captured some eighty years ago by Belgian surrealist artist René Magritte in his painting *The Treachery Of Images* (**Fig. 2**). As Magritte himself commented: "*Just try to stuff it with tobacco! If I were to have had written on my picture 'This is a pipe' I would have been lying*". This distinction has an important implication. To *analyze* literally means 'to break down into the constituent parts'. Since geo-objects are not constituent parts of remote sensing images, you cannot *base* their analysis on them; at most you can *orient* (i.e., direct) the analysis towards obtaining representations of them. Therefore, it is important to note that the new acronym GEOBIA refers to *GEographic OBIA* (where 'geographic' is a qualifier restricting the discipline to the geographic domain) rather than to *GEo-Object-Based Image Analysis*. Furthermore, the name OBIA presupposes that image-objects 'exist' within images. As we have seen, this is only true if we are talking of ('good') segmented images. Therefore, strictly speaking, segmentation would be a pre-processing step within OBIA, with the analysis taking place immediately after it. A more correct name, that would include segmentation as part of the analysis, would be *Object Oriented Image Analysis* (OOIA), but because of the likely confusion with computer *Object Oriented Program-*

---

[13] A formal representation of regions with indeterminate boundaries where each region has two concentric boundaries, an outer one representing the maximum possible extent of the region, and an inner one with the minimum possible extent of the region. This representation is akin to a fried egg, where the egg white is the region encompassed by these two boundaries. Any acceptable 'precise' version (crisping) of the region must lie within the white of the egg.

ming paradigm, the previous name is preferable, providing users are aware of the implications.

Fig. 2. René Magritte's The Treachery Of Images (© René Magritte Estate/Artists Rights Society (ARS), New York/ADAGP, Paris)

4.1 Meaningful image-objects

Having clarified what the relationship between these two kinds of objects is not, we can now proceed to describe it. We deal with two domains, the image and the geographic, which have to be linked. Fortunately, there is an *isomorphism*[14] between the image and the imaged landscape that enables us to link them. Not only is there a conspicuous resemblance between the spatial patterns that we perceive when the image is displayed on a monitor and the ones that we would get if we observed the same scene from a balloon suspended at an equivalent altitude. If the image has been ortho-rectified, the topological relations are also accurately preserved, as the cartographic transformation used to represent the earth surface is an exact function. For example, if we sequentially record the map-coordinates of the pixels defining the boundary of an image-object representing a lake in a 10 m resolution ortho-rectified image, store them as waypoints in a GPS, and reproduce the path on the ground, we can reasonably expect to be walking along the shore of the lake. Therefore, if the image is partitioned into perceptually coherent pieces, there is some basis to believe that

---

[14] Strictly speaking, this is a mathematical term for an exact correspondence between both the elements of two sets and the relations defined by operations on these elements. Obviously here the correspondence refers only to the elements of the landscape that are observable at the spatial resolution of the image.

the counterparts of these pieces on the ground may also be semantically coherent. In other words, image-objects have the potential to correspond on the ground to some entity for which we have a name. When an image-object can be seen as a proper representation of an instance of some type of geo-object, then we can say it is a ***meaningful image-object***, that is, ***a representation of a geo-object***. However, in practice, most initial image-objects (i.e., segments) will not qualify as 'meaningful image-objects', as explained in the next section. Therefore 'meaningful image-object' should be considered a different term than just 'image-object', the former describing a semantic unit and the latter a perceptual unit. Although this may lead to some confusion, we have preferred to avoid coining completely new terms.

4.2. Object-based classification

The goal of OBIA is to partition RS images into meaningful image-objects as defined above. Since some degree of oversegmentation is desirable (see below), most meaningful image-objects will consist of aggregates of segments, formed via *object-based classification*. The latter is the process of associating initial image-objects (segments) to *geo-object classes*, based on both the internal features of the objects and their mutual relationships. Ideally, there should be a one-to-one correspondence between image-segments and meaningful image-objects. However, this is hardly attainable, since a semantic model (i.e., a classification scheme) of the landscape cannot be projected with complete success onto regions (segments) that have been derived from a data-driven process (segmentation)[15]. There are at least two reasons for this shortcoming.

The first reason is that the relationship between radiometric similarity and semantic similarity is not straightforward. For example, there can be conspicuous differences in the image, such as those created by shadows, which have no meaning or no importance within the classification scheme, or conversely, the legend may include classes whose instances can barely be differentiated from each other in the image. In other cases, the way in which humans perceptually compose objects in the image, especially when

---

[15] This statement refers to conventional segmentation algorithms. There are of coarse some automated procedures, such as *template matching* or *cellular automata*,which use external information to partition the image (and therefore are not data-driven). Integrating such *model-driven* techniques into OBIA should constitute a priority line of research.

some *Gestalt principles*[16] intervene, may be too complex to be captured by a conventional dissimilarity metric. Therefore the possibility exists that the boundaries of the classified (aggregates of) image-objects do not lead to an agreeable representation of geo-objects. This implies that there will be some classified image-objects that need to be split or reshaped in part of their perimeter. The less clear the relation between the two similarities, the more likely this possibility. A way to tackle this problem is to stop the segmentation at a stage where these errors are less frequent or easier to fix. This is why some degree of oversegmentation is desirable. In any case, the process of endowing image-objects with meaning is a complex one, likely requiring several cycles of mutual interaction between segmentation and classification (Benz et al. 2004), regardless of what segmentation method is used.

The second reason, which has already been implicitly suggested, is that given a segmentation derived from an image of a certain spatial resolution, and even assuming that each segment corresponds to some recognizable entity, it is highly unlikely that all delineated objects belong to the same hierarchical level. For example, in a segmented 1 m resolution image, there can be segments that correspond to individual trees and cars, but also to whole meadows and lakes. While 'meadow' and 'lake' may have been included as separate classes in the legend, trees and cars are probably sub-objects of some class like 'urban park' and 'parking lot'. Therefore, some aggregation of image-objects according to semantic rules will be required.

In addition, for reasons explained in section 3.2, it might be necessary to upscale the image and segment it at a coarser resolution, since some geo-objects (e.g., a city) have boundaries that are indistinguishable at fine scales. The latter operation involves a further complication: the linkage of image-objects derived at different resolutions, since in general they will not nest seamlessly. At the moment there is no consensus on how to make this linkage. However, we note that a related problem has been addressed in *Computer Vision* regarding the tracking of moving objects across video frames, referred to as the *correspondence problem* (Cox 1993). This is akin to matching ambiguities between image-objects in successive images of increasingly larger pixel size.

---

[16] Basic principles describing the *composing* capability of our senses, particularly with respect to the visual recognition of objects and whole forms instead of just a collection of individual patterns (Wertheimer 1923).

5 Summary

OBIA is a relatively new approach to the analysis of RS images where the basic units, instead of being individual pixels, are image-objects. An *image-object* is a discrete region of a digital image that is internally coherent and different from their surroundings, and that potentially represents – alone or in assemblage with other neighbors- a *geo-object*. The latter is a bounded geographic region that can be identified for a period of time as the referent of a geographic term such as those used in map legends. Within OBIA, image-objects are initially derived from a segmentation algorithm and then classified using both their internal features and their mutual relationships. Image-objects are devoid of meaning until they are formally recognized as representational instances of either constituents parts of geo-objects or of whole geo-objects. Recognition involves the projection of a semantic model (*taxonomy*) onto a representation (the segmented image) of the landscape, where the result is a *partonomy*, i.e., a partition into *meaningful image-objects* (which mostly are aggregates of segments, i.e., the initial image-objects), each representing a single geo-object. The instantiation of image-objects as (representations of) geo-objects is a complex process for which we need to develop, and agree upon a comprehensive ontology. Here we have proposed definitions and relationships for two basic terms of it: 'image-object' and 'geo-object'.

Acknowledgements

References

Bennett B (2001) What is a Forest? on the vagueness of certain geographic concepts. Topoi 20:189-201

Benz UC, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-

ready information. ISPRS Journal of Photogrammetry and Remote Sensing 58: 239-258

Bittner T, Smith B (2001b) Vagueness and Granular Partitions. In: Welty C, Smith B (eds) Formal Ontology and Information Systems. ACM Press, New York, pp 309-321

Blaschke T, Strobl J (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GeoBIT/GIS 6:12-17

Burnett C, Blaschke T (2003) A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecological Modelling 168(3): 233-249

Casati R, Smith B, Varzi A (1998) Ontological Tools for Geographic Representation. In: Guarino N (ed) Formal Ontology in Information Systems, IOS Press, pp 77-85

Castilla G (2003) Object-oriented Analysis of Remote Sensing Images for Land Cover Mapping: Conceptual Foundations and a Segmentation Method to Derive a Baseline Partition for Classification. PhD Thesis, Polytechnic University of Madrid, URL:
http://oa.upm.es/133/01/07200302_castilla_castellano.pdf

Cohn AG, Gotts NM (1996) The 'Egg-Yolk' Representation of Regions with Indeterminate Boundaries. In: Burrough PA, Frank, AU (eds) Geographic Objects with Indeterminate Boundaries. Taylor and Francis, pp 171-188

Cox IJ (1993) A Review of Statistical Data Association Techniques for Motion Correspondence. International Journal of Computer Vision 10(1):53-66

Duhem P (1906) La theorie physique: son objet, et sa structure. Chevalier and Riviere, Paris. Translated by Philip P. Wiener, The Aim and Structure of Physical Theory (Princeton University Press, 1954)

Fisher P (2000) Sorites paradox and vague geographies. Fuzzy Sets and Systems 113 (1):7-18

Frank AU (2001) Tiers of ontology and consistency constraints in geographic information systems. Int. J Geographical Information Science 15(7):667-678

Hay GJ, Castilla G (2007) Geographic Object-Based Image Analysis (GEOBIA). In: Blaschke T, Lang S, Hay GJ (eds) Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications. Springer-Verlag, Berlin

Hay GJ, Castilla G (2006) Object-Based Image Analysis: Strengths, Weaknesses, Opportunities and Threats (SWOT). In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. No. XXXVI-4/C42, Salzburg

Hay GJ, Niemann KO, Goodenough DG (1997) Spatial Thresholds, Image-Objects and Upscaling: A Multiscale Evaluation. Remote Sensing of Environment 62:1-19

Kartikeyan B, Sarkar A, Majumder KL (1998) A Segmentation Approach to Classification of Remote Sensing Imagery. International Journal of Remote Sensing 19(9):1695-1709

Korzybski A (1933) Science and Sanity. An Introduction to Non-Aristotelian Systems and General Semantics. International Non-Aristotelian Library Publishing Co., Lancaster

Kuhn TS (1962) The Structure of Scientific Revolutions. The Chicago University Press, Chicago

Lang S, Blaschke T (2006) Bridging remote sensing and GIS – what are the main supporting pillars? International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences vol. XXXVI-4/C42

Lang S, Langanke T (2006) Object-based mapping and object-relationship modeling for land use classes and habitats. PFG - Photogrammetrie, Fernerkundung, Geoinformatik 1:5-18

Landgrebe D (1999) Information Extraction Principles and Methods for Multispectral and Hyperspectral Image Data. In: Chen CH (ed) Information Processing for Remote Sensing. The World Scientific Publishing Co, New Jersey

Marceau DJ, Howarth PJ, Gratton DJ (1994) Remote Sensing and the Measurement of Geographical Entities in a Forested Environment. Part 1: The Scale and Spatial Aggregation Problem. Remote Sensing of Environment 49(2):93-104

Neubert M, Herold H, Meinel G (2007) Evaluation of remote sensing image segmentation quality. In: Blaschke T, Lang S, Hay GJ (eds) Object-Based Image Analysis - Spatial concepts for knowledge-driven remote sensing applications. Springer-Verlag, Berlin

Schiewe J, Tufte L, Ehlers M (2001) Potential and problems of multi-scale segmentation methods in remote sensing. Geographische Informationssysteme 6: 34-39

Smith B (2003) Ontology. In: Luciano Floridi (ed), Blackwell Guide to the Philosophy of Computing and Information. Blackwell, Oxford, pp 155-166

Smith B (2001) Fiat Objects. Topoi 20:131-148

Smith B, Brogaard B (2002) Quantum Mereotopology. Annals of Mathematics and Artificial Intelligence 35:153-175

Smith B, Mark DM (1998) Ontology and Geographic Kinds. Proc. 8th Int. Symp. on Spatial Data Handling (SDH'98), pp 308-320

Varzi A (2001) Vagueness in Geography. Philosophy and Geography 4(1): 49-65

Wertheimer M (1923) Laws of Organization in Perceptual Forms. Translation published in Ellis, W. (1938). A source book of Gestalt psychology (pp 71-88). Routledge, London

Wiens JA (1995) Landscape mosaics and ecological theory. In: Hansson L, Fahrig L, Merriam G (eds) Mosaic Landscapes and Ecological Processes (pp 1-26). Chapman and Hall, London

Woodcock CE, Strahler AH (1987) The factor of scale in remote sensing. Remote Sensing of Environment 21:311-332

Woodcock CE, Harward VJ (1992) Nested-hierarchical scene models and image segmentation. International Journal of Remote Sensing 13:3167-3187

Wu J, Loucks OL (1995) From balance-of-nature to hierarchical patch dynamics: a paradigm shift in ecology. Quarterly Review of Biology 70: 439-466

Zadeh L (1965) Fuzzy sets. Information and Control 8(3):338-353

Section 2

Multiscale representation and object-based classification

Chapter 2.1

Using texture to tackle the problem of scale in land-cover classification

P. Corcoran, A. Winstanley

National Centre for Geocomputation & Department of Computer Science, National University of Ireland Maynooth, Maynooth, Co. Kildare, Ireland. padraigc@cs.nuim.ie

ABSTRACT: Object-Based Image Analysis (OBIA) is a form of remote sensing which attempts to model the ability of the human visual system (HVS) to interpret aerial imagery. We argue that in many of its current implementations, OBIA is not an accurate model of this system. Drawing from current theories in cognitive psychology, we propose a new conceptual model which we believe more accurately represents how the HVS performs aerial image interpretation. The first step in this conceptual model is the generation of image segmentation where each area of uniform visual properties is represented correctly. The goal of this work is to implement this first step. To achieve this we extract a novel complementary set of intensity and texture gradients which offer increased discrimination strength over existing competing gradient sets. These gradients are then fused using a strategy which accounts for spatial uncertainty in boundary localization. Finally segmentation is performed using the watershed segmentation algorithm. Results achieved are very accurate and outperform the popular Canny gradient operator.

1 Introduction

To overcome the failings of classic pixel-based classification techniques at providing accurate land-use classification, many researchers are developing a new form of remote sensing known as OBIA (Benz et al. 2004).

OBIA attempts to model the HVS which can interpret aerial imagery quite easily. The OBIA strategy can be divided loosely into two stages. The first is image segmentation of the given scene into a set of objects. These objects then serve as input to a supervised classifier which utilizes spatial properties such as shape and object context to classify land-use (Blaschke 2003). Although OBIA outperforms traditional pixel-based remote sensing techniques, it does not offer the same level of accuracy as the HVS. We believe one of the reasons for this is due to many OBIA implementations being based upon an inaccurate conceptual model of the HVS.

The data used in this research consists of scanned aerial photography with a 0.25 m ground sample distance of Southampton city which is located on the south coast of England. This data was obtained from the Ordnance Survey of Great Britain. The imagery was initially in RGB format but was converted to grey-scale before any analysis was performed. This conversion was performed using the Matlab function rgb2grey (Gonzalez et al. 2003). Fig. 1 displays an example of the original RGB data following the conversion to greyscale. From this example it is obvious that the visual system can still accurately interpret aerial imagery even when colour information has been removed. Greyscale imagery contains only one channel compared to three for colour imagery. This reduces the computational complexity in terms of both feature extraction and fusion, and was one of the reasons for the conversion. Modelling how the HVS interprets colour information is much more complex in terms of theory than modelling how it interprets greyscale information (Palmer 1999) (p. 94). This was also a major factor which influenced the conversion to greyscale.

Fig. 1 The original RGB image is converted to greyscale before any analysis is performed. Ordnance Survey Crown Copyright. All rights reserved

The remainder of this chapter is organized as follows. First we propose an alternative conceptual model of the HVS which we believe agrees more

with current theories of visual perception than the current model underlying many OBIA implementations. This is followed in section 3 by details of how the first step in this model may be implemented. Results are presented in the final part of this section. Finally section 4 discusses our conclusions from this work and possible future research directions.

2 A conceptual model of aerial photo interpretation

It is generally accepted that the HVS is in some form object-based. Watt (1995) argues that this provides an alternative to computing a full representation of point based spatial relationships and allows the utilization of spatial properties such as shape. Many of the current implementations of OBIA attempt to define objects using segmentation by utilizing solely the visual cue of intensity/colour (Baatz and Schape 2000). This would be an appropriate strategy if each object of interest was of uniform intensity/colour, but this is not the case. Most will contain small scale texture which is another valuable cue used by the HVS to detect boundaries. Attempting to generate segmentation using exclusively intensity/colour based features in the presence of texture will lead to over- and under-segmentation. In an attempt to lessen such deficiencies some OBIA implementations run intensity/colour segmentation at multiple scales in the hope that each of the required objects will be defined correctly at some scale (Benz et al. 2004). All segmentation scales are then merged to generate the object-hierarchy. This procedure of running intensity/colour based segmentation at multiple scales followed by integrating to generate the object-hierarchy, we believe is not an accurate model of the HVS as it does not concur with current theories in visual perception. Therefore we now attempt to define a more accurate conceptual model of the processes involved in the human interpretation of aerial imagery.

In this research we define primitive-objects as connected regions of uniform visual properties, which in turn we define as areas of uniform intensity or texture. We argue that the initial step in an accurate conceptual model of the HVS is the segmentation of a given scene into a set of primitive-objects. This would be a single segmentation where each individual primitive-object is represented correctly. The following theories of visual perception support this view.

Perceptual organization is defined as how all the bits and pieces of visual information are structured into the larger units of perceived objects and their interrelations (Palmer 1999) (p. 255). The principle of uniform connectedness states that connected regions of uniform visual properties tend

to be perceived initially as single objects and correspond to entry level units of perceptual organization (Palmer and Rock 1994). Julesz renowned for his work on the visual perception of texture seems to support the principle of uniform connectedness stating that "*the aperture of attention changes its spatial scale according to the size of the feature being sought*" (Julesz 1983). Therefore when an area of uniform visual properties is viewed our aperture is adjusted to view it correctly as a single primitive-object. In theories of object recognition such as Biederman's recognition-by-components (Biederman 1987), complex objects are described by the spatial arrangement of basic component parts. The first step in this object recognition process is the segmentation of the given scene into areas of uniform visual properties. These segments are then matched against a collection of primitive-components called *Geons* which represent the individual parts of the object to be recognized. Theories of visual attention, which are generally considered to be object-based on some level, also argue that areas of uniform visual properties are initially represented as single objects. In Treisman feature integration theory of visual attention (Treisman and Gelade 1980), before visual attention is focused, the visual system generates separate representations of each visual stimulus. These separate features are then conjoined to form a single segmentation of the scene where each primitive-object is represented correctly. Drawing from the above theories, a corresponding implementation of early vision segmentation would be a segmentation algorithm which segments each area of uniform visual properties correctly within a single segmentation. If segmentation is defined exclusively using raw intensity/colour features no such segmentation can be achieved.

According to the principle of uniform connectedness, an early vision segmentation process defines primitive-objects upon which grouping and parsing operate to generate different levels or scales in the object-hierarchy (Palmer 1999) (p. 268). Grouping is the process of aggregating individual primitive-objects to form segmentation at a larger scale. On the other hand, parsing involves dividing primitive-objects into separate parts to form segmentation at a smaller scale. An important distinction in the processing of visual information is its metaphorical 'direction' of bottom-up or top-down (Palmer 1999) (p. 84). Top-down processing, also called hypothesis-driven or expectation-driven processing, is influenced by our prior knowledge, desires and expectations. On the other hand, bottom-up processing is unaffected by these influences and is purely driven from knowledge extracted from the scene. In all the above theories which we have drawn from, the early vision segmentation which defines primitive-objects is a bottom-up process. The Gestalt laws of perceptual organization are based on the assumption that the grouping and parsing of these primitive-objects to form

the object-hierarchy is also a bottom-up process (Palmer 1999) (p. 257). Vecera and Farah (1997) showed this not to be the case, and that in fact this merging and parsing of primitive-objects in the HVS to produce the object-hierarchy is a top-down process. Pylyshyn (1999) supports this claim arguing that such top-down influences play a major role in visual perception. An example of where such top-down influences are present could be for example in the aggregation of two primitive-objects corresponding to building roof and chimney to form a single building object. Although these two primitive-objects may have considerable different visual properties, our prior knowledge informs us that they should be aggregated at a larger scale in the object-hierarchy. In many current OBIA implementations such top-down influences are not considered. The object-hierarchy is generated in a completely bottom-up manner by simply running intensity/colour segmentation at multiple scales (Baatz and Schape 2000; Benz et al. 2004).

For the reasons just discussed we believe many current implementations of the OBIA paradigm to be founded on an imprecise conceptual model of the HVS. We therefore propose a new conceptual model for the human interpretation of aerial imagery which we feel agrees more with current theories in visual perception. This conceptual model contains the following three stages:

1. First a single segmentation where each primitive-object is segmented correctly is performed. This segmentation is a bottom-up process.
2. Primitive-objects are then merged and parsed to form the object-hierarchy. This process is both influenced by top-down and bottom-up factors. Both individual object properties and context relationships between objects play a major role in this step.
3. Finally land-use classification of the object-hierarchy is performed. This is very top-down driven process.

Most theorists believe that the early stages of human visual processing are strictly bottom-up. The point at which top-down processes begins to augment bottom-up process is a controversial issue. Some believe this point occurs early in visual processing while others believe it occurs later (Palmer 1999) (p. 85). In our conceptual model the point at which top-down processes begin to augment bottom-up processes is after the definition of primitive-objects. This choice was motivated by the principle of uniform connectedness which is believed to be a bottom-up process.

The HVS is extremely sophisticated and we do not claim our conceptual model to be entirely accurate. Nevertheless we deem it to be a more accurate conceptual model than the one underlying many current implementa-

tions of the OBIA paradigm. The next section in this chapter presents details of how the first step in this conceptual model may be implemented.

3 Methodology

Remotely sensed images contain both texture and intensity boundaries which must be exploited effectively if accurate primitive-object segmentation is to be realized. Most segmentation strategies can be classified as a region or boundary based approach. In region based methods grouping of homogenous areas is performed to produce segmentation. In contrast, boundary based methods attempt to extract the boundaries between these homogenous areas. The watershed transform combines both region and boundary based techniques (Soille 2002). Pixels are grouped around the regional minima of a gradient image and boundaries are located along the crest lines of this image. In order to achieve accurate segmentation using the watershed transform an accurate boundary gradient image must first be extracted. This section presents a strategy for computing such a gradient image accurately.

Gradients extracted from raw texture and intensity features are inaccurate and do not harmonize with each other. A gradient operator applied to a raw intensity image will not only respond to intensity boundaries but also the intensity variation due to object texture resulting in a significant number of false positives.

Texture being a spatial property can only be described by features calculated within a neighbourhood. This process of integrating information over a neighbourhood leads to what we refer to as the texture boundary-response problem, where an unique response is observed at primitive-object boundaries (Corcoran and Winstanley 2007). This response is due to features being extracted from a mixture of textures and/or an intensity boundary between primitive-objects. Segmentation performed using these raw texture features will result in the generation of unwanted classes along object boundaries. Also a gradient operator applied to such feature images will not give an accurate gauge of the texture boundary magnitude. Two measures of primitive-object to boundary-response gradient will result not the desired single measure of primitive-object to primitive-object gradient.

For the above reasons a sophisticated strategy needs to be employed to fuse the visual properties of texture and intensity. Most current segmentation methods which attempt to fuse texture and intensity feature gradients are based on a competing gradient set. These strategies are founded on the assumption that each primitive-object boundary is predominately a inten-

sity or texture boundary, and only the corresponding feature gradient should be used to define that boundary. The authors in (Malik et al. 2001; Chaji and Ghassemian 2006) use a measure of texturedness to modulate both texture and intensity gradients. Intensity gradients in the presence of texture are inhibited reducing the number of gradient false positives due to texture intensity variation. While texture gradients in the absence of texture are suppressed removing gradient false positives due to boundary-responses at pure intensity boundaries. O'Callaghan and Bull (2005) only modulate the intensity gradients and remove texture boundary responses by separable median filtering of the texture feature images.

Most boundaries in remotely sensed images consist of both an intensity and texture boundary. Therefore since boundary strength is only measured in terms of a single visual cue, either texture or intensity, all the above competing gradient strategies suffer reduced discrimination strength. To curb this shortcoming we propose to use a complementary gradient set. This allows the calculation of a more accurate boundary gradient in terms of a sum of individual texture and intensity gradients. This is achieved by performing post-processing of both raw texture and intensity feature images removing all gradients except those due to primitive-object boundaries. Our paper also reports a novel fusion strategy for such gradient images which offers a number of benefits over existing strategies. We describe our early vision segmentation algorithm under the following headings:

1. Complementary intensity gradient extraction.
2. Complementary texture gradient extraction.
3. Fusion of gradient images and segmentation.

3.1 Complementary intensity gradient extraction

A schematic of the complementary intensity gradient extraction process is displayed in Fig. 2.
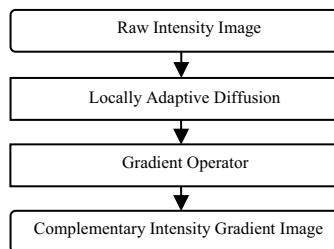

Fig. 2 Schematic of complementary intensity gradient extraction strategy

Removing intensity gradients due to texture while maintaining those due to primitive-object intensity boundaries can be achieved using a smoothing process (Deng and Liu 2003). Smoothing an image using a linear or non-linear diffusion process will cause each primitive-object to approach its average intensity value; this is known as the property of grey level invariance (Deng and Liu 2003). If every primitive-object in a given scene has a significant different average intensity value then accurate primitive-object segmentation, while maintaining accurate boundary localization, can be achieved using the diffused image.

Linear diffusion or Gaussian smoothing although probably the most common smoothing technique used, has a number of drawbacks. Namely it blurs boundaries giving mixed classes, suffers from a loss of boundary localization and applies equal smoothing to all locations. Non-linear diffusion was first introduced in (Perona and Malik 1990) to overcome the weaknesses of linear diffusion. In this strategy the amount of diffusion performed at any location is controlled by the use of an edge-stopping function of gradient magnitude. Black et al. (1998) introduced a diffusion process known as robust anisotropic diffusion which preserves sharper boundaries than previous formulations and improves termination of the diffusion. This diffusion process applied to image I is given by:

$$\frac{\partial I(x, y, y)}{\partial t} = div\big[g\big(\|\nabla I\|\big)\nabla I\big] \tag{1.1}$$

where $\|\nabla I\|$ is the gradient magnitude and g is a decreasing edge stopping function. Within robust anisotropic diffusion, Tukey's biweight function is used for g:

$$g(x, \sigma) = \begin{cases} \frac{1}{2}\big[1 - (x/\sigma)\big] & , |x| < \sigma \\ 0 & , otherwise \end{cases} \tag{1.2}$$

where σ is the diffusion scale parameter. Following a finite number of iterations of the diffusion process, all edges with a gradient magnitude of less than σ will be completely smoothed while all others will remain.

Using a single diffusion scale parameter for all spatial locations is undesirable. If an edge of a given gradient has a large percentage of neighbouring edges with similar gradients then the edge in question is probably the result of texture intensity variation. On the other hand if an edge with a given gradient does not have a large percentage of neighbouring edges with similar gradients then this edge is probably the result of an intensity primitive-object boundary. Therefore an edge with a specific gradient in one location should be classified as an edge due to texture intensity varia-

tion while in a different location should be classified as an edge due to a primitive-object boundary. This effect is shown graphically in Fig. 3.
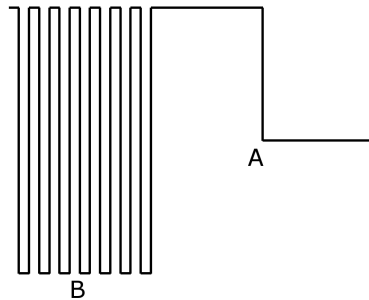

Fig. 3 The gradient of edge A is significantly greater than the median of its neighbouring gradients. Therefore this edge represents a boundary and will be preserved by the diffusion process. The gradient of edge B, although greater than that of A, is not significantly greater than the median of its neighbouring gradients. As a result this edge will not be preserved by the diffusion process

To accommodate this spatially varying property, (Black and Sapiro 1999) proposed to calculate a spatially varying diffusion scale parameter. In the original paper, Black and Sapiro used the Median Absolute Deviation (MAD) of local gradient values as a measure of local diffusion scale, where MAD is a robust measure of variance. It is important to use a quantification from the area of robust statistics. The neighbourhood of an edge resulting from a primitive-object intensity boundary will contain a certain percentage of edges of similar gradient magnitude resulting from the same boundary. The statistic used must therefore be robust to such outliers. Using a measure of variance such as MAD does not encapsulate information regarding the actual magnitude of neighbouring gradients. Therefore we propose to use the median of gradient values within a neighbourhood multiplied by a user specified parameter as a local measure of diffusion scale. Fig. 4 displays the result of this diffusion process applied to a remotely sensed image. The proposed complementary intensity gradients are extracted from this image using a Sobel gradient operator (Gonzalez et al. 2003).
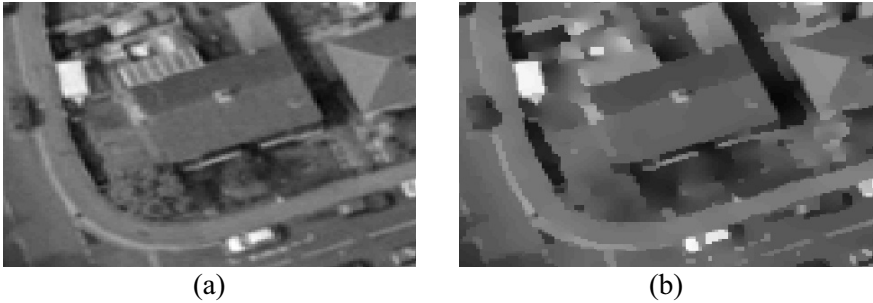
(a)   (b)

Fig. 4 The locally adaptive diffusion process is applied to image (a) with the result shown in (b). Significant intensity variation due to texture is removed while primitive-object intensity boundaries remain. Each primitive-object approaches its average grey value. The building roof and tree located just below in (a) have different textures but similar average intensity values. The diffused image in (b) displays little or no intensity boundary between these two primitive-objects. Therefore this boundary must be defined using primitive-object texture gradients. Ordnance Survey Crown Copyright.

3.2 Complementary texture gradient extraction

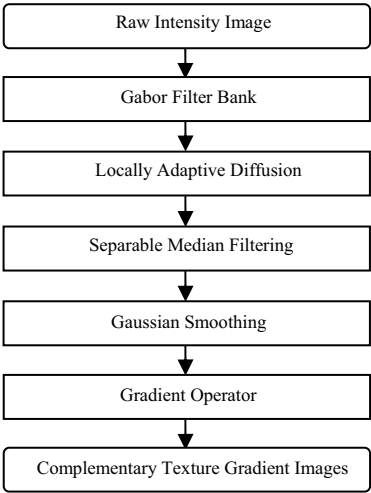A schematic of the complementary texture gradient extraction process is displayed in Fig. 5.

Fig. 5 Schematic of complementary texture gradient extraction strategy

When working with intensity based images, a large number of neighbouring primitive-objects will have similar average intensity values. Following the application of a diffusion process these will have a very similar appearance leading to under-segmentation. Therefore when attempting to derive accurate segmentation of intensity images it is important to model the texture within these images and integrate with intensity features in an intelligent manner. This will allow boundaries between primitive-objects having similar average intensity values but different textures to be detected. An example of such a primitive-object boundary can be seen in Fig. 4.

A bank of Gabor filters are utilized for texture feature extraction. Gabor filters are an attractive form of texture feature extraction because they have been shown to model the simple cells in the primary visual cortex quite accurately and minimize the uncertainty principle (Clausi and Jernigan 2000). Spatially, a Gabor function is a Gaussian modulated sinusoid. An example of a Gabor feature image is displayed in Fig. 6.
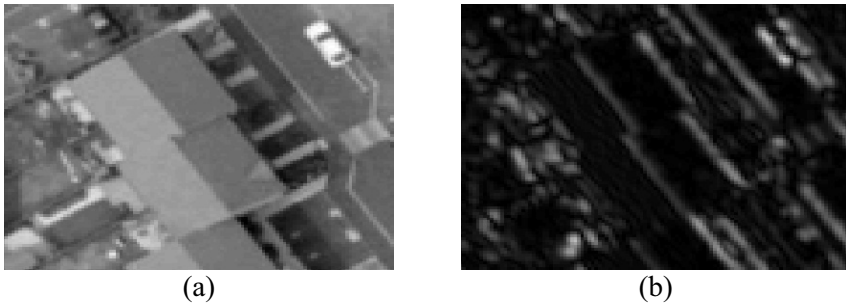

(a)   (b)

Fig. 6 Gabor features of a low spatial frequency extracted from (a) are shown in (b).

Clausi and Jernigan (2000) demonstrated the need for smoothing of the raw Gabor filter outputs to counter an effect known as leakage which results in over-segmentation. Following feature extraction all feature images are smoothed using the non-linear locally adaptive diffusion process discussed in the previous section.

To overcome the issues introduced by texture boundary responses post processing of feature images must be performed to remove these responses before a segmentation algorithm can be applied. A practical boundary response removal technique should have the following properties:

– It should prevent the generation of boundary classes.
– Post application of a gradient operator should give an accurate gauge of texture boundary magnitude.

In a recent paper we proposed such a technique which composes separable median filtering followed by smoothing with a relatively small Gaussian (Corcoran and Winstanley 2007). Following this gradient are then calculated with a Sobel gradient operator. This strategy will extract boundary gradients with both of the above required properties and are localized accurately to a spatial scale equal to half that of the corresponding feature extraction used. Fig. 7 shows an example result of the proposed complementary texture gradient extraction strategy.
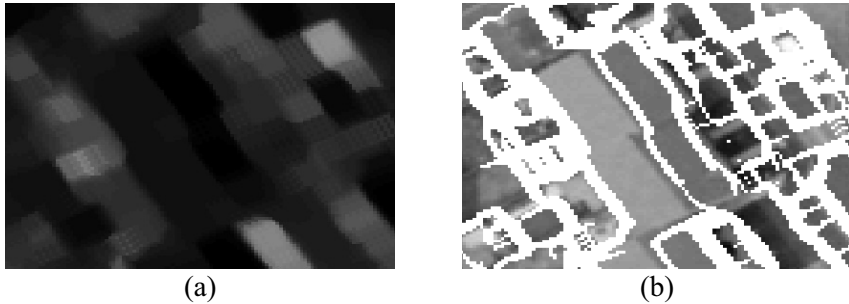

(a)   (b)

Fig. 7 The result of applying a diffusion process followed by texture boundary response removal to Fig. 6 (b) is shown in (a). Texture gradients are then extracted from (a). Those above a certain threshold are represented by white and superimposed on the original image in (b). These boundaries are localized to half the scale of feature extraction and therefore in some cases are not positioned entirely correctly. Ordnance Survey Crown Copyright. All rights reserved

3.3 Fusion of gradient images and segmentation

Given the resulting complementary texture and intensity gradient images, these must be fused to form a single gradient image before application of the watershed transform. A schematic of proposed fusion process is displayed in Fig. 8.
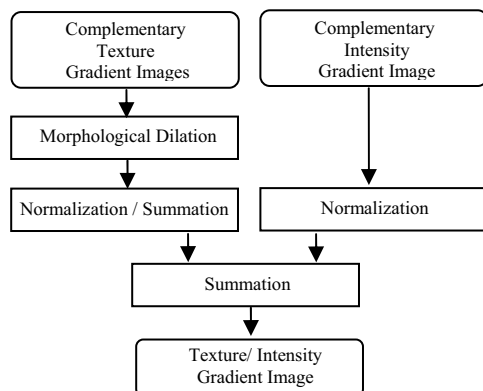
Fig. 8 Schematic of the complementary texture and intensity gradient image fusion process

Intensity is a point property therefore using a non-linear diffusion process which preserves boundary localization we can locate intensity boundaries accurately to the pixel scale. Texture on the other hand is a spatial property and therefore any features used to describe it must be calculated within a neighbourhood. This makes it difficult to accurately localize texture boundaries. In fact the approach we employed will only localize texture boundaries accurately to half the spatial scale of the corresponding feature extraction used. In Fig. 7 (b) we can see that some building texture boundaries are not located entirely correctly (Corcoran and Winstanley 2007).

Simply performing a summation of texture and intensity gradient images as is standard in previous approaches without taking this uncertainty into account introduces two drawbacks (O'Callaghan and Bull 2005; Chaji and Ghassemian 2006). If corresponding boundaries are located in different positions, a sum of their gradient values will give a double peaked effect, with neither peak being a true measure of the actual boundary gradient. This double peak effect will also result in the segmentation algorithm returning a false boundary object. These effects are demonstrated in Fig. 9.
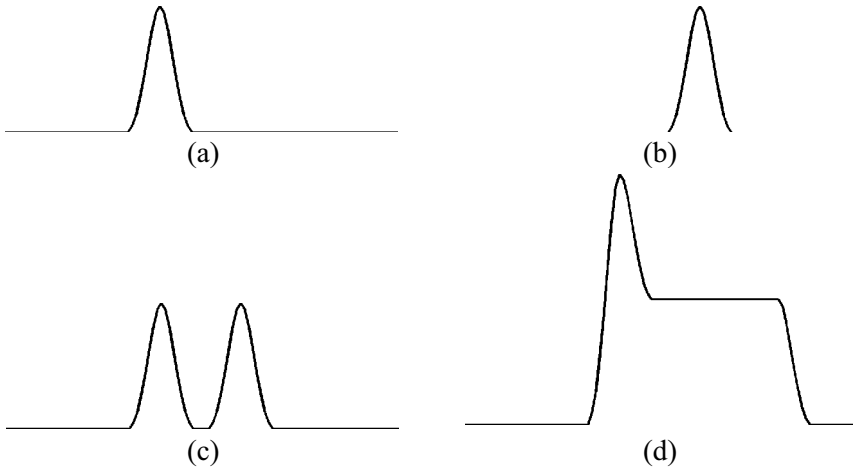
Fig. 9 A primitive-object boundary has both an intensity and texture gradient response shown in (a) and (b) respectively. Although both gradients correspond to the same boundary they are located in different positions. Simply performing a summation will result in a double peaked effect in the resulting gradient image as shown in (c). Neither peak is a true measure of actual boundary gradient. Also segmentation will return a false boundary object. We propose to dilate the texture gradient before summation with the result of this summation shown in (d). The peak of (d) is an accurate value of boundary gradient magnitude. The watershed segmentation algorithm will locate the boundary at this peak which is the location of the intensity boundary therefore minimizing spatial uncertainty

To overcome these downfalls we add a requirement of texture gradients which are later fused with intensity gradients in this manner. This requirement states: *the boundary response should be uniform over a scale greater than the scale of localization*. To achieve this we dilate the texture gradient images using a disk shaped structuring element (Soille 2002). The structuring element used has a radius equal to the boundary localization scale of the original gradients. Our original texture gradient extraction algorithm extracts gradients accurate to half the scale of the feature extraction performed therefore we dilate with a similar scale structuring element. This increases the area over which the gradient operator responds to a texture boundary fulfilling the above requirement. Fig. 9 displays visually the result of applying this strategy.

Before a summation of individual gradient images is performed, each must be normalized to weigh its contribution. Shao and Forstner (1994) proposes to normalize by variance or maximum filter response. O'Callaghan and Bull (2005) normalizes by maximum filter response followed

by normalization by the sum of filter response. There is no strong evidence to suggest one normalization technique is superior so we choose to normalize each gradient magnitude image by its maximum response. This scales each image to the range [0 1]. We then sum the individual texture gradient images and normalize this again by the maximum value giving a single texture gradient image. We then divide the intensity gradient image by four times its median value and the texture gradient image by its median value. This step aligns the noise floor of each function (O'Callaghan and Bull 2005). Finally the single intensity and texture gradient images are summed to form a final gradient image. All further segmentation analysis is performed on solely this image. Example complementary intensity, complementary texture and fused complementary intensity/texture gradient images are displayed in Fig. 10 (a), (b) and (c) respectively.
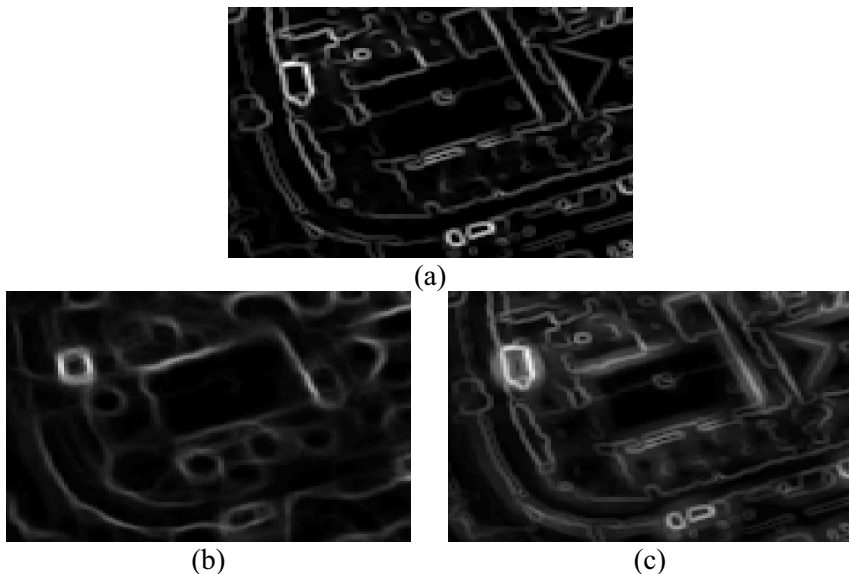

(a)

(b)   (c)

Fig. 10 Intensity and texture gradient images for Fig. 3 (a) are displayed in (a) and (b) respectively. These are fused in (c) to form a single gradient image

To perform segmentation we use the marker-controlled watershed transform (Soille 2002). In practice, direct computation of the watershed algorithm results in over-segmentation due to the presence of spurious minima. To overcome this, the gradient image is first filtered using a marker function, in this case the H-minima transform, to remove all irrelevant minima (Soille 2002). Intensity being a point property is not affected by the uncertainty principle and therefore offers superior boundary localization com-

pared to texture. If a boundary has both an intensity and texture boundary, the crest of the gradient magnitude image will be located at the intensity boundary. The watershed algorithm will therefore locate boundaries using an intensity boundary as opposed to a texture boundary where possible and minimize spatial uncertainty. This effect is illustrated in Fig. 9.

3.4 Segmentation evaluation

Primitive-object segmentation evaluation using ground-truth captured by human interpretation is difficult. This is owing to the fact that when humans are asked to generate the required ground-truth, they return segmentation at a larger scale than required with a large percentage of individual primitive-objects merged. This point is illustrated in a forthcoming publication where a cognitive experiment was performed on a number of subjects. Each subject was asked to generate the corresponding ground-truth for a number of remotely sensed images. All ground-truths returned where of a larger scale than required with significant merging of primitive-objects evident. Capturing of ground truth where each primitive-object is represented would require a great effort on the interpreter's part and may not be even possible. This is due to the fact that we as humans tend to merge primitive-objects unconsciously even though this merging a top-down process. It is well known that top-down factors can influence visual processes without us being conscious of it. For example, it turns out that how people identify letters depends strongly on whether those letters are part of known words or meaningless letter strings. For this to happen, there must be top-down influences from some higher-level representation of known words acting on a lower-level representation of letters (Palmer 1999) (p. 85). Therefore a comparison against ground-truth is pointless due to the fact that any ground-truth captured would be inaccurate. Due to this inability of humans to capture ground-truth at the required scale, segmentation evaluation is performed in a qualitative manner using visual inspection and a quantitative manner using an accurate unsupervised segmentation performance metric which does not require ground-truth.

   Some segmentation results are displayed in Fig. 11. These results are close to the scale we desire where each primitive-object, for example trees and building roofs, is segmented correctly. From these images we see that primitive-object boundaries which have both an intensity and texture boundary are localized very accurately. This demonstrates the ability of our algorithm to minimize spatial uncertainty in boundary localization. Examples of this include road and building boundaries. For boundaries which do not have a strong intensity boundary, for example some bounda-

ries between tree tops, localization is not so accurate. Some over- and under-segmentation is also evident in highly textured regions.
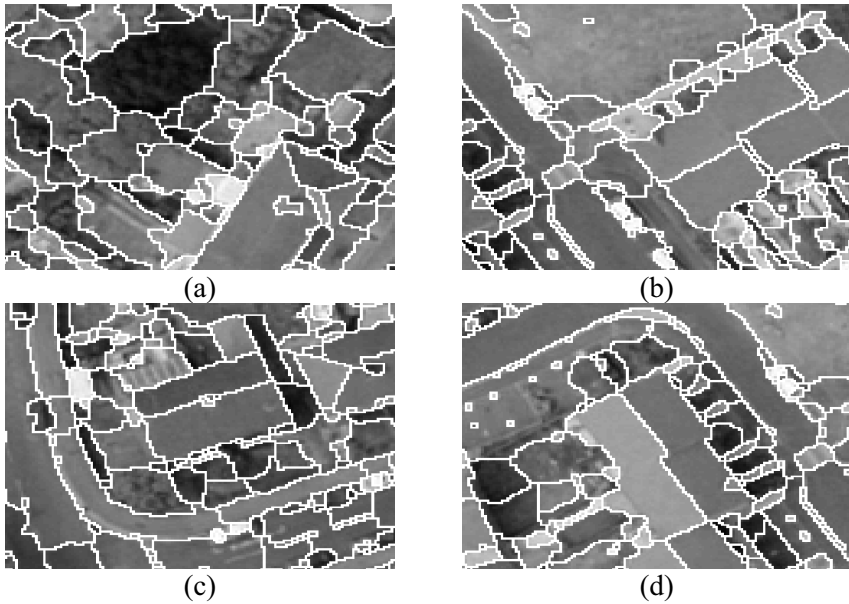

(a)	(b)
(c)	(d)

Fig. 11 Segmentation results obtained from the segmentation procedure described. Primitive-object boundaries are represented by the colour white. Ordnance Survey Crown Copyright. All rights reserved

We performed quantitative performance evaluation in an unsupervised manner, which does not require ground-truth, using an unsupervised performance metric know as the SU metric (in forthcoming publication). Evaluation of remotely sensed data segmentation has been performed in such an unsupervised manner in the past (Pal and Mitra 2002; Mitra et al. 2004). The SU metric calculates a ratio of the individual segments texture and intensity feature separation to cohesion, with higher values indicating better segmentation performance. This metric has been shown to be an accurate performance metric having a high correlation with an accurate supervised metric on a synthetic dataset, containing noisy uniform and textured regions, where accurate ground-truth is known. This correlation is significantly greater than previous attempts to establish a relationship between supervised and unsupervised metrics on the same data (in forthcoming publication). We evaluated our segmentation algorithm against the marker-controlled watershed transform applied to gradients extracted using

the Canny gradient detector (Canny 1986). These gradients were calculated by smoothing with a Gaussian of sigma 1.5 followed by application of the Sobel gradient operator. A set of 60 images of size 256x256 pixels was used for evaluation. This set was divided into 20 training images to optimize the segmentation scale parameter H for each algorithm and 40 test images. Our proposed segmentation algorithm achieved an average SU metric value of 0.43 on the test data set. This result outperformed the marker-controlled watershed transform applied to the Canny gradients which accomplished an average SU metric value of 0.38 on the same data.

4 Conclusions and Future Work

Generating accurate land-use classification in urban areas from remotely sensed imagery is a challenging problem. The ultimate goal of OBIA is to model the HVS, which can perform this task quite easily. We argue that some previous implementations of the OBIA paradigm are based on an inaccurate conceptual model of the HVS. We therefore proposed a new conceptual model which we feel overcomes this weakness and will hopefully lead to more accurate land-use classification systems in the future.

We attempted to implement the first step in this conceptual model by performing segmentation where each area of uniform visual properties is segmented correctly. To achieve this we proposed a segmentation algorithm which involves the computation and fusion of a novel complementary texture and intensity gradient set. The segmentation results achieved are visually very accurate, although some over- and under-segmentation is evident. Using an unsupervised performance metric we showed that the proposed algorithm quantitatively outperforms the marker-controlled watershed transform applied to gradients extracted using the popular Canny gradient operator.

The slight over- and under-segmentation evident in the results may be due to gradient images being based on a local measure of difference which does not encode information regarding the difference between the interior of regions (O'Callaghan and Bull 2005). A potential solution may be post-processing of the segmentation in which regions are merged based on their interior region properties. Some primitive-object boundaries suffer from poor localization. This is due to texture gradients only being localized to the spatial scale of feature extraction. Future work will attempt to localize these gradients to a finer spatial scale.

References

Baatz M, Schape A (2000), Multiresolution segmentation - an optimization approach for high quality multi-scale image segmentation, J. et al. (eds.): Angewandte Geographische Informationsverarbeitung XIII, Wichmann, Heidelberg, pp 12-23

Benz UC, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004), Multiresolution, object-orientated fuzzy analysis of remote sensing data or GIS-ready information, ISPRS Journal of Photogrammetry & Remote Sensing, vol 58, pp 239-258

Biederman I (1987), Recognition-by-components: a theory of human image understanding, Psychol Rev, vol 94, pp 115-147

Black M, Sapiro G (1999) Edges as outliers: anisotropic smoothing using local image statistics. International Conference on Scale-Space Theories in Computer Vision, Corfu, Greece, pp 259-270

Black MJ, Sapiro G, Marimont DH, Heeger D (1998), Robust anisotropic diffusion, IEEE Trans on Image Process, vol 7, no 3, pp 421-432

Blaschke T (2003) Object-based contextual image classification built on image segmentation. IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, pp 113-119

Canny JF (1986), A computational approach to edge detection, IEEE Trans Pattern Anal and Mach Intell, vol 8, no 6, pp 679-698

Chaji N, Ghassemian H (2006), Texture-gradient-based contour detection, EURASIP Journal on Applied Signal Processing, vol 2006, pp 1-8

Clausi DA, Jernigan ME (2000), Designing Gabor filters for optimal texture separability, Pattern Recogn, vol 33, no 11, pp 1835-1849

Corcoran P, Winstanley A (2007) Removing the texture feature response to object boundaries. International Conference on Computer Vision Theory and Applications, Barcelona, pp 363-368

Deng H, Liu J (2003), Development of anisotropic diffusion to segment texture images, J Electron Imag, vol 12, no 2, pp 307-316.

Gonzalez R, woods R, Eddins S (2003) Digital image processing using matlab, Prentice Hall

Julesz B (1983), Textons, the fundamental elements in preattentive vision and perception of textures, Bell System Technical Journal, vol 62, pp 1619-1645

Malik J, Belongie S, Leung T,Shi J (2001), Contour and texture analysis for image segmentation, Int J of Comput Vis, vol 43, no 1, pp 7-27

Mitra P, Shankar BU, Pal SK (2004), Segmentation of multispectral remote sensing images using active support vector machines, Pattern Recogn Lett, vol 25, no 2004, pp 1067-1074

O'Callaghan RJ, Bull DR (2005), Combined morphological-spectral unsupervised image segmentation, IEEE Trans Image Process, vol 14, no 1, pp 49-62

Pal SK, Mitra P (2002), Multispectral image segmentation using the rough-set-initialized EM algorithm, IEEE Trans Geosci Rem Sens, vol 40, no 11, pp 2495-2501

Palmer SE (1999) Vision science: photons to phenomenology, The MIT Press.

Palmer SE, Rock I (1994), Rethinking perceptual organization: The role of uniform connectedness, Psychonomic Bull Rev, vol 1, no 1, pp 29-55

Perona P, Malik J (1990), Scale-space and edge detection using anisotropic diffusion, IEEE Trans Pattern Anal Mach Intell, vol 12, no 7, pp 629-639

Pylyshyn ZW (1999), Is vision continuous with cognition? The case for cognition impenetrability of vision perception, Behav Brain Sci, vol 22, no 3, pp 341-365

Shao J, Forstner W (1994) Gabor wavelets for texture edge extraction. ISPRS Commission III Symposium on Spatial Information from Digital Photogrammetry and Computer Vision, Munich, Germany, pp 745-752

Soille P (2002) Morphological image analysis: principles and applications, Springer-Verlag Berlin and Heidelberg GmbH & Co. K

Treisman A, Gelade G (1980), A feature integration theory of attention, Cognit Psychol, vol 12, pp 97-136

Vecera SP, Farah MJ (1997), Is visual image segmentation a bottom-up or an interactive process?, Percept Psychophys, vol 59, no 8, pp 1280-1296

Watt RJ (1995) Some speculations on the role of texture processing in visual perception. In: T. V. Papathomas, C. Chubb, A. Gorea and E. Kowler (eds) Early Vision and Beyond, The MIT Press, pp 59-67

Chapter 2.2

Domain-specific class modelling for one-level representation of single trees

D. Tiede [1], S. Lang [1], C. Hoffmann [2]

[1] Centre for Geoinformatics (Z_GIS), Salzburg University, Austria; dirk.tiede@sbg.ac.at; stefan.lang@sbg.ac.at

[2] Definiens AG, Munich, Germany; choffmann@definiens.com

KEYWORDS: Segmentation, LiDAR, object-based, airborne laser scanning, tree crown delineation, object generation, OLR, region-specific

ABSTRACT: As a synthesis of a series of studies carried out by the authors this chapter discusses domain-specific class modelling which utilizes *a priori* knowledge on the specific scale domains of the target features addressed. Two near-natural forest settings served as testing environment for a combined use of airborne laser scanning (ALS) and optical image data to perform automated tree-crown delineation. The primary methodological aim was to represent the entire image data product in a single, spatially contiguous, set of scale-specific objects (one-level-representation, OLR). First, by high-level (broad-scale) segmentation an initial set of image regions was created. The regions, characterised by homogenous spectral behaviour and uniform ALS-based height information, represented different image object domains (in this case: areas of specific forest characteristics). The regions were then treated independently to perform domain-specific class modelling (i.e. the characteristics of each region controlled the generation of lower level objects). The class modelling was undertaken using Cognition Network Language (CNL), which allows for addressing single objects and enables supervising the object generation process through the provision of programming functions like branching and looping. Alto-

gether, the single processes of segmentation and classification were coupled in a cyclic approach. Finally, representing the entire scene content in a scale finer than the initial regional level, has accomplished OLR. Building upon the preceding papers, we endeavoured to improve the algorithms for tree crown delineation and also extended the underlying workflow. The transferability of our approach was evaluated by (1) shifting the geographical setting from a hilly study area (National Park Bavarian Forest, South-Eastern Germany) to a mountainous site (Montafon area, Western Austria); and (2) by applying it to different data sets, wherein the latter differ from the initial ones in terms of spectral resolution (line scanner RGBI data vs. false colour infrared orthophotos) and spatial resolution (0.5 m vs. 0.25 m), as well as ALS point density, which was ten times higher in the original setting. Only minor adaptations had to be done. Additional steps, however, were necessary targeting the data sets of different resolution. In terms of accuracy, in both study areas 90 % of the evaluated trees were correctly detected (concerning the location of trees). The following classification of tree types reached an accuracy of 75 % in the first study area. It was not evaluated for the second study area which was nearly exclusively covered by coniferous trees.

1 Introduction

Very high spatial resolution (VHSR) optical sensors and airborne laser scanning (ALS) technology, especially in combination, provide information on a broad range of possible target features in human-related scale domains. Increased spatial resolution along with additional continuous information (such as derived from ALS) serves to improve our ability to detect individual features. But the rich information content of the 'H-res' situation (Strahler et al., 1986), where many pixels make up each object, an excess of spatial detail must be dealt with (as in Lang, chapter1.1, Hoffmann et al., chapter 1.2, Castilla and Hay, chapter 1.5). Region-based segmentation techniques are an intuitive, yet empirically improvable means for re-aggregating this detailed information and thus reducing scene complexity. Multi-scale segmentation (Baatz and Schäpe, 2000) in general provides a hierarchical set of scaled representations, adaptable to the required level of detail. In forest applications we usually aim at scaled representations of the entire scene content, and not at solely feature extraction (i.e. Tiede et al. 2004a). However, for a full representation of the scene content, one single adapted layer produced through segmentation may be desired as an inappropriate trade-off between over- and under-

segmentation. As presented in Lang (2002), specific structural arrangements of target features may require class- (or domain-)specific multi-scale segmentation. Assessing relevant scales may be conducted in an analytical way, such as by performing scale-space analysis and deriving characteristics on object behaviour while moving through scale (e.g. Hay et al., 2005). The approach discussed in this chapter is complementary and knowledge-driven: we discuss a supervised approach for the process of segmentation and object generation utilizing human *a priori* knowledge on the specific scale domain of the target features.

A key premise of our approach is that the result should capture the entire scene content in a spatially contiguous one-level representation (OLR, Lang and Langanke, 2006). In addition, the approach should be transferable to different geographical settings and to data sets with different spatial or spectral resolution. We therefore selected two different study areas, transferred the workflow, and applied the approach with only minor adaptations.

2 Study Areas and Data sets

Study areas

The first study area is located in the National Park Bavarian Forest (NPBF) in South-Eastern Germany along the border with the Czech Republic (Fig. 1). The NPBF study area covers almost 270 hectares (ha) of near-natural forest, with an elevation between 780 and 1,020 meters above sea level (ASL) and slopes up to 25 degrees. Different forest structures occur with both open and closed forests, and multiple tree canopy layers with varying tree species, ages and sizes. In the 1990s, mountain spruce stands faced severe attacks from spruce bark beetle (*Ips typograficus*) especially, mainly triggered by major storm events.

The second study site is situated in the Montafon area in the federal state of Vorarlberg in Western Austria (see Fig. 1). The study area in the Montafon area is characterized by west-facing slopes ranging from 1,400 to 1,800 meters ASL with slopes ranging between 25 and 40 degrees. 22 ha (460 m x 480 m) in size, the area is dominated by old spruce stands partly thinned out due to windfall caused by heavy storms in the year 1990. Tree patterns typical for mountain forests occur, such as clusters or smaller groups of trees and gaps in between (in German: *Rotten*). This study site also includes a spruce-pole stand in the north-western part of the

area. Forest coverage on this slope helps prevent hotels and houses in the valley below from damages caused by rock-fall or avalanches. Automatically derived forest structure parameters from both ALS and optical data support a forest rehabilitation project carried out in this area to preserve the protection function on a long-term basis (Maier 2005).
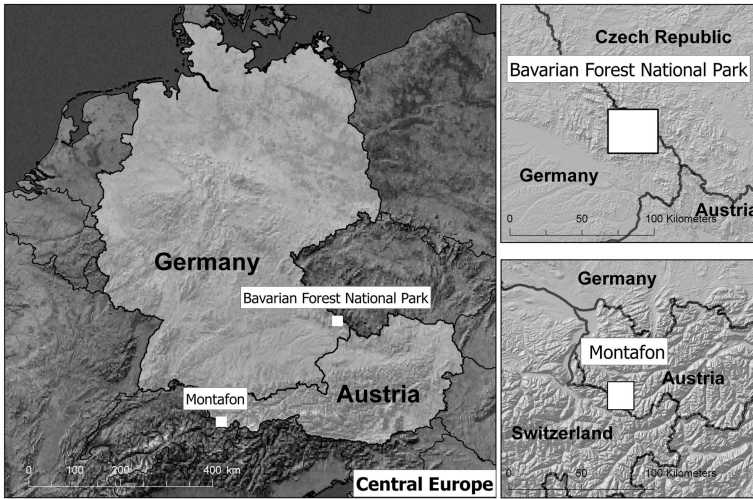

Fig. 1. Study areas in South-Eastern Germany (National Park Bavarian Forest) and Western Austria (Montafon area)

Data sets and pre-processing

For the NPBF study area data from the Toposys Falcon system (cf. Wehr and Lohr, 1999, Schnadt and Katzenbeisser, 2004) was available. Surveying of the study area was done at three different dates: two leaf-off flight campaigns in March 2002 and May 2002, and a leaf-on campaign in September 2002. Both first and last returns were collected during the flights with a pulse repetition rate of 83 kHz, a wavelength of 1560 nm and an average point density of 10 pts per m² in common. An average flight height around 850 m and a maximum scan angle of 14.3 degrees resulted in a swath width of about 210 m and a high overlap due to the three different flights.

The resulting data sets were pre-processed by Toposys using TopPit (TopoSys Processing and Imaging Tool) software. The derived surface model (DSM) and terrain model (DTM) were subtracted from each other to create a normalised crown model (nCM) with 1 m ground sample dis-

tance (GSD). The nCM served as a basis for the single tree-crown delineation. Simultaneously to the ALS measurements, image data were recorded using the line scanner camera of TopoSys. The camera provides four bands (RGBI) at a GSD of 0.5 m: blue (440-490 nm), green (500-580 nm), red (580-660 nm), and NIR (770-890 nm).

For the study area in Montafon ALS data were acquired in December 2002 under leaf-off canopy conditions by TopScan (Germany) on request from the Vorarlberg government. The Optech Airborne Laser Terrain Mapper (ALTM 1225) was used to collect first and last returns. The pulse repetition rate was 25 kHz with a wavelength of around 1000 nm and an average point density on ground of 0.9 points per m². The mean flying height was 1000 m and the maximum scan angle 20 degrees. As a result the swath width was about 725 m and the overlap between flight lines was 425 m (Wagner et al. 2004, Hollaus et al. 2006). The raw data were processed at the Institute of Photogrammetry and Remote Sensing (IPF, TU Vienna). Again, both DTM and a DSM with 1 m GSD were produced. The software package SCOP++ developed by IPF was used, based on the hierarchic robust filtering approach (Kraus and Pfeiffer, 1998). An nCM with 1 m GSD was derived. In addition, a set of false colour infrared (FCIR) aerial photos from 2001, recorded separately with a GSD of 0.25 m, were available.

In both study areas visual interpretation was used for validation purposes.

3 Methodology

The approach was realised by developing rule sets in Cognition Network Language (CNL) within the Definiens Developer Environment. CNL, similar to a modular programming language, supports programming tasks like branching, looping, and defining of variables. More specifically, it enables addressing single objects and supports manipulating and supervising the process of generating scaled objects in a region-specific manner (cf. Tiede and Hoffmann 2006). By this, the process steps of segmentation and classification can be coupled in a cyclic process; this we call class modelling. It provides flexibility in designing a transferable workflow from scene-specific high-level segmentation and classification to region-specific multi-scale modelling – in this case of single tree crowns.

High-level segmentation and classification generating a priori information

For initial segmentation (high-level segmentation), we used a region-based, local mutual best fitting segmentation approach (Baatz and Schäpe, 2000). The 'scale parameter' (which controls the average size of generated objects) in the two study areas was different due to differences in spatial resolution (see Tab. 1). In the NPBF study, the initial segmentation was built upon the five available dimensions of the input data layers (optical and nCM data). In the Montafon study the segmentation was based only on optical information. This step resulted in a rough delineation of domains with different forest characteristics (see below), and likewise non-vegetated areas (such as roads, larger clearance areas, etc.). The initial regions are characterised by homogeneous spectral behaviour and uniform height information. Accordingly they were assigned to image object domains, and provided '*a priori*' information for the subsequent delineation of single trees.

Table 1. High-level segmentation settings (L = Level, SP = scale parameter, SW = shape weighting, CPW = compactness weighting)

Study area	L	SP	SW	CPW	Remarks
NPBF	1	100	0.5	0.5	nCM was weighted three times higher than RGBI
Montafon	1	300	0.5	0.5	only FCIR were deployed

The Normalized Difference Vegetation Index (NDVI) was used for separating coniferous, deciduous, dead trees, and non-vegetated areas (cf, Lillesand et al., 2004, Wulder 1998). The standard deviation of the nCM data per object served as indicator for the respective forest structure. In the NPBF study we distinguished between five domains: a. coniferous open, b. coniferous closed, c. deciduous open, d. deciduous closed, and e. mixed forest. The coniferous and deciduous classes were further differentiated based on nCM values, introducing two more sub-categories, namely forests below 20 m height and forest above 20 m height. Note that we use the terms 'open' and 'closed' not in a strict methodological sense; we utilize nCM standard deviation per object as a proxy to differentiate roughly between open and closed forested areas to control the algorithms in the following single tree crown delineation.

In the Montafon study only two different domains were identified: a. coniferous open and b. coniferous closed. Reasons are the prevailing natural conditions which only allows for limited occurrence of deciduous trees.

Bare ground caused by wind throw, gaps or outcrop formed an additional, merged class. Figure 2 shows the results in subsets of both study areas.
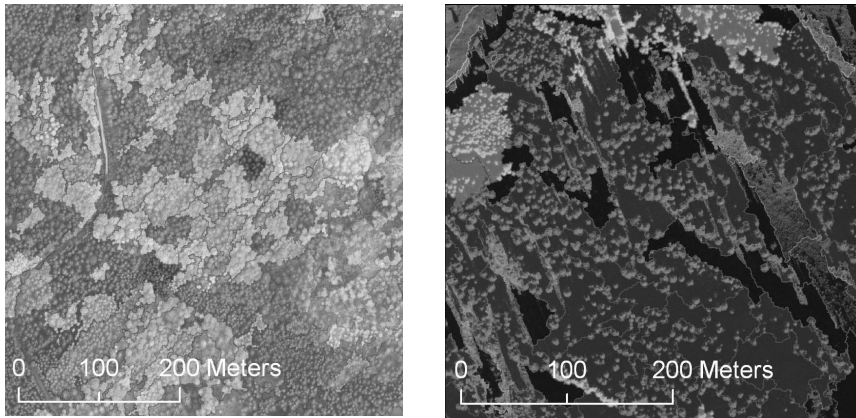

Fig. 2. High-level segmentation and initial classification of different domains in subsets of the study areas in the NPBF (left) and the Montafon area (right). Bright values indicate domains with a higher percentage of deciduous trees, dark values are indicative of predominant coniferous trees. Results are regions for supervised multi-scale class modelling.

Optimized multi-scale class modelling - object generation in a domain specific hierarchy

Segmentation algorithms based on homogeneity criteria, like the one used for the initial high-level segmentation, were found not suitable for delineating complex and heterogeneous canopy representations in VHSR optical data or ALS data (cf. Tiede et al. 2004b, Burnett et al. 2003). New developments help overcome these limitations through the use of scalable segmentation algorithms. These specific object generation algorithms can be adapted to the prevailing scale domains or even to the actual object. Object delineation is therefore controlled by user-specified parameters. The domains and their spatial instances, the regions, provide *a priori* information for the domain-specific embedded scalable algorithms, and they set the spatial constraints for its application. By this, we accomplish an optimized multi-resolution segmentation for single tree delineation.

One crucial element of this approach is to again break down the regions into pixels ('pixel-sized objects') within a region's boundary. These pixel-objects are used to generate objects in a supervised manner. Specific rule-sets used were developed for single tree crown delineation on ALS data by

Tiede and Hoffmann (2006) and have now been adapted to the specific conditions in the study areas and the data sets used. In the underlying rule sets a region growing segmentation algorithm is programmed using a continuity constraint starting from tree tops (local maximum) as seed points.

Table 2 gives an overview of the parameterisation controlled by the pre-classified domains. According to Tiede et al. (2006) the following region-specific parameters were controlled: (1) The search radius for the local maximum method needs to be adapted for each region depending on the assigned domain: taller deciduous trees require a bigger search radius to avoid detecting false positives due to the flat and wide crown structure; dense coniferous stands require a smaller search radius to detect close standing tree tops. Therefore the search radius varies between 1 m and 4 m (cf. Wulder et al. 2000). (2) The stopping criterion for the region-growing process depends on the underlying nCM data. Candidate objects are taken into account, as long as differences in height between the respective objects not exceed a certain limit. The limits are variable in terms of different tree height and tree types, ranging between 2 m and 9 m height difference. (3) A maximum crown width is used for preventing uncontrolled growth of tree crown objects and merging with other potential tree crowns. This may happen, if a local maximum was not recognized correctly, for example in dense deciduous stands due to fairly planar tree surface or missing tree top representations in the ALS data. In comparison to Tiede et al. (2006) this criterion was further adapted by introducing a crown width parameter. In addition to the *a priori* information being used, this parameter is now directly linked to the individual tree height value derived from the ALS data (cf. Pitkänen et al., 2004; Kini and Popescu, 2004; Koch et al., 2006). Crown width limits range from 4 to 17 m for coniferous trees and 5 to 20 m for deciduous trees, corresponding to the crown width of open grown trees in Austria dependent of tree heights according to Hasenauer (1997).

Table 2. Overview of domain-specific controlled differences in object generation processes. Plus (+) and minus (-) indicate higher or lower values, LMR = Local maximum search radius; SENS = Sensitivity of the stopping criterion value based on the nCM data (continuity criterion); CWL = Crown width limit influenced by the initial region and the local maximum height; further details see text above.

Parameter	Finer scale – smaller trees	Coarser scale –larger trees
LMR	- (to detect tree tops in closed stands)	+ (to avoid false positives)
SENS	+ (small coniferous trees)	-- (large coniferous trees)
	++ (for small deciduous trees)	- (large deciduous trees)
CWL	-- (small coniferous trees)	+ (large coniferous trees)
	- (small deciduous trees)	++ (large deciduous trees)

The objects resulting from this procedure are expected to correspond with single tree crowns. However, in both study areas smaller holes occur between the delineated tree crown objects. Causes for this are mainly due to limitations of the available ALS data with only 1 m GSD. Therefore we applied an additional algorithm, which uses object neighbourhood information to fill these holes up to a certain size with the respective surrounding object.

Figure 3 shows the complete workflow. Note that for the Montafon study area it was extended. Because of the given resampling method of the data sets used and the different spatial resolutions, the local maximum method was biased. To overcome this problem, the two major steps of the process, i.e. (1) high-level segmentation and (2) region-specific tree crown delineation were separated in two different project settings. The *a priori* information was integrated by importing the information as a thematic layer to be addressed via cognition network language in the new project setting. That means, re-loading FCIR data was not necessary in this step of the workflow.
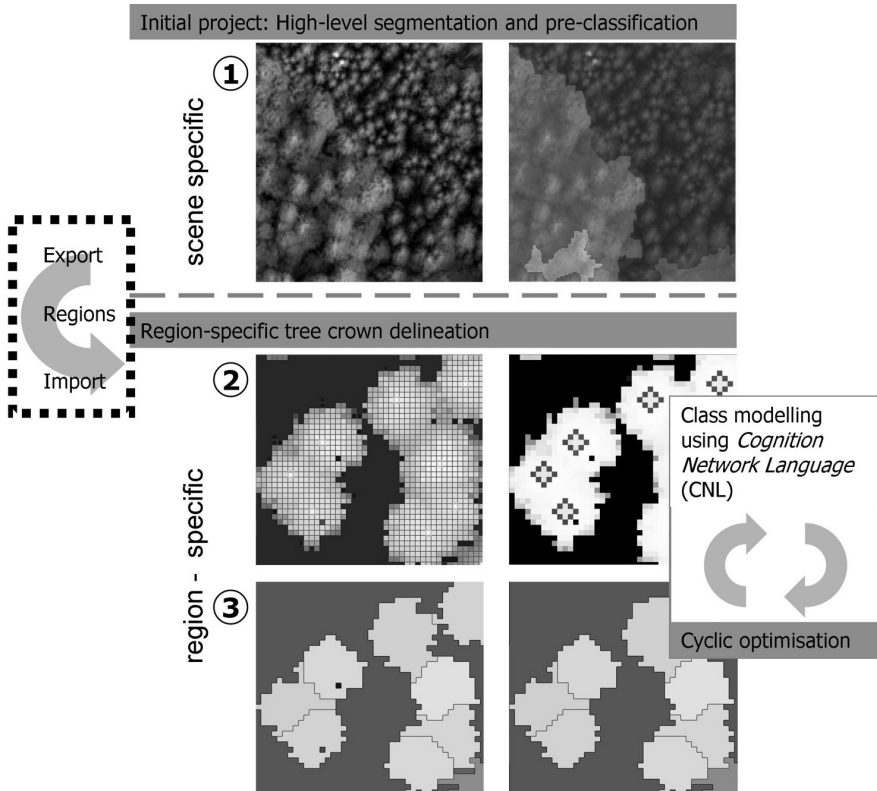
Fig. 3. Workflow after Tiede and Hoffmann 2006, extended: (1) High-level segmentation and classification of domains with different forest characteristics in an initial phase. The arrow and the dotted line around it indicate an extension of the workflow in the Montafon case. Export and re-import of regions are only necessary when multispectral and ALS data have different spatial resolution. (2) Breakdown of imported pre-classified forest type domains to small objects (here: pixel objects) and extraction of local maxima. Generation of domain-specific objects using a region growing algorithm (local maxima taken as seed points). (3) Extracted single tree crowns (holes mainly caused by limitations of available ALS data); cleaning up single tree crown objects using neighborhood information.

For the final classification of single tree objects we used NDVI values and nCM height values. In the NPBF study five classes were differentiated: coniferous trees, deciduous trees, dead trees, non-vegetated area and understorey. The latter comprises objects which could not be allocated to a tree crown according to the defined stopping criteria but which still show values indicating vegetated areas. In the spruce dominated Montafon study

area we only differentiated between coniferous trees (i.e. spruce), non-vegetated area, and understorey. Reasons for the use of these classes are different requirements in the study areas: In NPBF there is a need for differentiation between coniferous trees (mainly spruce), deciduous trees and dead tress, mainly to quantify and predict spruce bark beetle attacks (Ochs et al., 2003). In the Montafon area the main concern is the protection function of the forest. Therefore location, density and height of trees are important (Maier, 2005).

The validation of the classification process was conducted visually by on-screen digitizing of tree-tops and comparing the results with the automatically derived tree crown polygons (i.e. point in polygon analysis). Ground truth data were only available for the Montafon area but the small subset of measured trees did not contain a representative amount. Therefore we relied on visual interpretation carried out in top-view by experts. Visual inspection was considered more suited and more effective for evaluating this approach than any other quantitative method; reason being the required quality control not only of the class assignment but also the way of delineation. For the delineation of tree crowns from the given ALS data with a GSD of 1 m, a minimum tree crown size of several pixels is required (cf. Coops et al., 2004). According to Maltamo et al., (2004) and Pitkänen (2001) the local maxima method is mainly suited to find dominant trees (dominant tree layer according to Assmann, 1970). Due to the open conditions of the old Montafon spruce stands, we considered trees higher than 5 m above the shrub layer (Schieler and Hauk, 2001) for validation. In NPBF trees of the dominant layer and the middle layer were validated. Because of the more closed forest conditions, understorey (trees smaller than 10 m) was not taken into consideration. 'Understorey' was used according to Schieler and Hauk (2001) and IUFRO (1958), who defined understorey as the lowest third of the forest total top height or dominant height (h100 = average height of the 100 thickest trees per ha, herein replaced by the 100 highest trees per ha (see Assman, 1970, Zingg, 1999 and Hasenauer, 1997)).

4 Results and Discussion

In the NPBF study altogether more than 73,600 trees were extracted and tree crowns delineated, out of which 75 % were classified as coniferous trees, 19 % as deciduous trees and 6 % as dead trees.

In the Montafon study 2,344 single trees were detected in total, for almost all of which crowns were delineated, even in the dense pole forest in

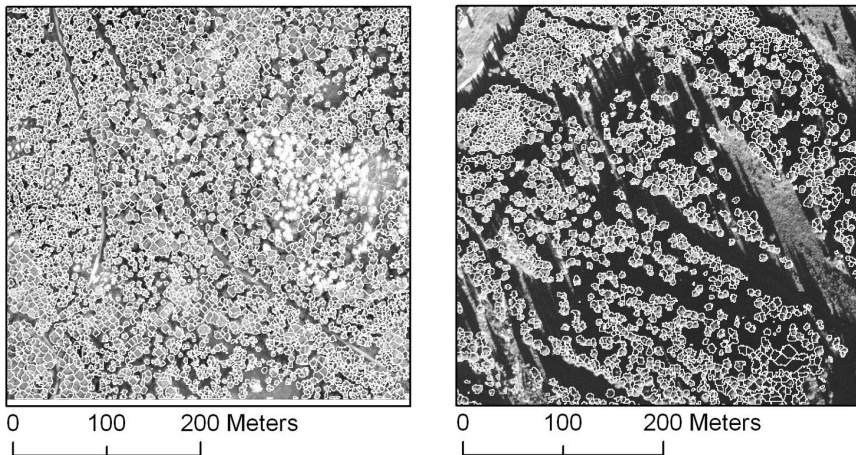the north-western part of the study area. Screenshots of the delineated tree crowns are shown in Figure 4.

Fig. 4. Result of the single tree delineation for two subsets in the study areas. Outline of tree crowns are underlain with RGBI data (left, NPBF area) and FCIR data (right, Montafon area).

In both study areas a few local maxima were detected without subsequently delineating a tree crown (NPBF: 0.7 %; Montafon Area; 2 %). Reasons are mainly the occurrences of small trees or dead trees, where the given point density of the ALS data fails to represent the whole crown.

Due to the size of the NPBF study area a reference data set was created by digitizing dominant trees in 20 randomly selected 50 m x 50 m cells (see Figure 5). Three different types of trees were distinguished ending up in a point data set representing 847 coniferous trees, 302 deciduous and 138 dead trees. Table 3 shows the results of the validation: 730 coniferous trees, 132 deciduous trees and 109 dead trees were classified correctly. Taking into account all detected trees, this results in an accuracy of approximately 75 %, ranging between 86 % for coniferous and 79 % for dead trees, but dropping to 44 % for deciduous trees. Taking into account trees which were detected but not classified correctly, the detection rate for delineated trees increases to 90 %.

Table 3. Results of the validation for 20 randomly selected cells in the NPBF study area

	Coniferous	Deciduous	Dead trees	Overall
Number (#) of manually digitized trees	847	302	138	1287
# of automatically detected trees	1017	206	142	1365
# of correctly detected trees (including misclassified trees belonging to other classes)	878	147	138	1163
# of correctly classified trees	730	132	109	971
% of correctly classified trees	~86 %	~44 %	~79 %	~75 %
# of misclassified but detected trees	148	15	29	192
# of false positives	139	59	4	202
% of false positives	~16%	~19.5%	~3%	~16%
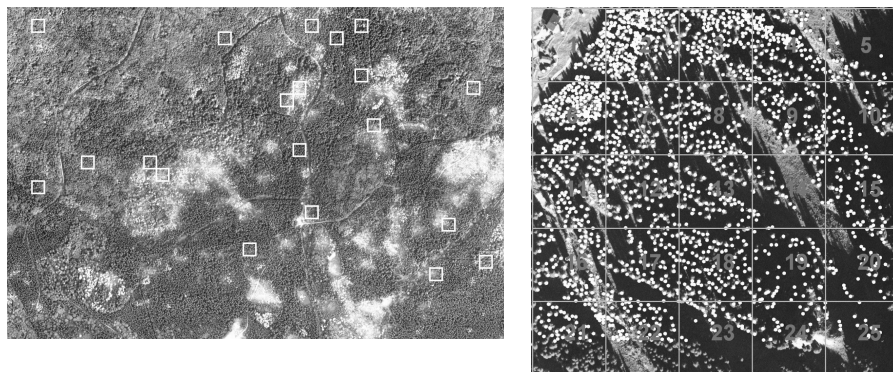
Fig. 5. Left: randomly selected cells (50 m x 50 m) for validation purpose in NPBF; right: digitized tree tops for the Montafon study area, 100 m x 100 m cells are overlain for statistical analysis.

In the Montafon study site the visual validation was conducted for the entire area. Altogether, 1,908 tree-tops were digitized, almost all of them coniferous trees. For a spatially disaggregated result the automatically detected tree-tops and the digitized ones were summarized in regular 100 x 100 m raster cells (see Figure 5, above, and Figure 6). The point-in-

polygon analysis revealed the results shown in Table 4. 1,721 trees were detected correctly, corresponding to an accuracy of approximately 90 %.

Table 4. Results of visual validation in the Montafon study area

# of manually digitized trees	1908
# of automatically classified trees	1721
% of automatically classified trees	~90 %
# of false positives	630
% of false positives	~33 %

In both study areas a high overestimation (i.e. high number of false positives) was recorded: 202 trees in NPBF and 630 trees in Montafon were detected by the algorithm, but were not be detected visually. Reasons for that are known limitations for the local maxima method to find tree tops as seed points for crown delineation (cf. Maltamo et al., 2004; Wulder et al. 2000; Tiede and Hoffmann, 2006): Especially in dense and highly structured forest different problems appear. Next to the fact that the laser scanner might have missed the highest point and that only nCM data of 1 m GSD were available, these problems are also related to the specific shape and structure of the trees: (1) crowns (especially deciduous tree crowns) with more than one local maximum because of a flat, rather than conical shape; (2) understorey trees which are only partly visible or double crowns (especially the *Rotten* structure in the Montafon area) are complicating the identification of only one local maximum per tree crown.

In the Montafon study another fact accounts for the higher number of automatically delineated trees, which – in this case – may be rather considered to visual underestimation: the shadowed conditions of the aerial photographs hampered visual detection. Consequently, most overestimated areas occurred in the steepest and most shady south-eastern part of the study site. Figure 6 shows a per-cell comparison of manually and automatically extracted tree tops. The diagram shows high congruence (correlation coefficient: 0.95), even in the denser pole forest in the North-western part of the area (cells #2 and #6).
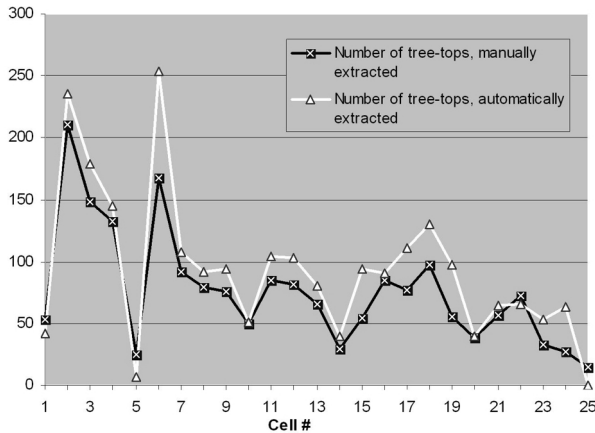
Fig. 6. Cell-by-cell comparison of manually and automatically extracted tree tops in the Montafon study area

5 Conclusions

The idea of domain-specific hierarchical representations discussed in this paper is based on the notion of hierarchically scaled compositions of the classes to be addressed (Lang, 2002). Domains exhibit similar hierarchical characteristics in their spatial instances, the regions. We have delineated these regions by a coarse segmentation, which integrates the underlying forest specific heterogeneity. Then, these classified regions control domain-specific segmentation. The initial segmentation and classification of different domains is crucial for the effectiveness of the single tree algorithms. Especially in heterogeneous mixed forests it is difficult to delineate initial regions representing more or less even-scaled areas. In this case a certain level of generalisation is required for successfully applying the algorithms tailored to a specific scale. This is where expert knowledge and user demand meet, and this is why we consider it a supervised approach. Object generation utilizes *a priori* knowledge about the specific scale domain of the target features. Considering transferability of the workflow, we have demonstrated that only minor adaptations were required to accomplish satisfying results. An improvement of the workflow compared to its initial version was accomplished for addressing data sets of different resolution. For future applications this adaptation is strengthening the utility and transferability of the workflow and enables higher flexibility with regard to different resolutions (spatial and spectral) of remotely sensed data.

Looking more generally at the transferability of rule-sets this approach provides a step towards operability by controlling linked algorithms through an initial high-level segmentation and classification. This facilitates development of an algorithm library for scale-specific target features, different scenes or even different data sets. High-level classification determines which features may be expected within the specific region, which then aids in determining which algorithms fit best in terms of data sets and features present.

In both studies we applied site-specific accuracy assessment (Congalton and Green, 1999), modified to a point-in-polygon analysis for usage in an object-based environment. By comparing with visually interpreted trees we assessed both the thematic correctness of the class assignment and the identification of the trees. In this respect we went one step further as compared to classical map-to-map comparison, since also the tree detection as such has been assessed. But object-specific accuracy assessment, as the name indicates, would even go beyond that (Lang and Langanke, 2006). In general, assessing the accuracy of polygonal features using point samples is a kind of reductionism within the object-based approach (Schöpfer and Lang, 2006). In this specific case we did not assessed, in how far the tree-crown delineation was appropriate. Generally stated: when dealing with spatial objects, the geometric accuracy of an object boundary also needs to be assessed – beyond assessing the locational and thematic agreement. This, however, implies consideration of scale: whatever method we apply for object delineation, be it automated, visually interpreted or surveyed on the ground, we produce representations which related to a particular scale. Issues and approaches for performing object-based accuracy assessment considering the issue of scaled representations is further discussed by Schöpfer et al., this volume.

References

Assmann, E (1970) The principles of forest yield study: studies in the organic production,structure, increment and yield of forest stands. Pergamon Press, Oxford, UK.

Baatz M and Schäpe A (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Angewandte Geographische Informationsverarbeitung XII, Strobl, J., Blaschke, T., Griesebner, G. (eds.), Wichmann: Heidelberg, 2000, pp. 12-23.

Burnett C, Heurich M, Tiede D (2003) Exploring Segmentation-based Mapping of Tree Crowns: Experiences with the Bavarian Forest NP Lidar/Digital Image

Dataset. Poster presented at ScandLaser 2003 International Conference and Workshop, Umeå , Sweden, September 2-4, 2003.

Congalton RG and Green K (1999) Assessing the accuracy of remotely sensed data: principles and practices. CRC/Lewis Press, Boca Raton, FL. 137 p.

Coops N. Wulder M, Culvenor D, and St-Onge B (2004) Comparison of forest attributes extracted from fine spatial resolution multispectral and lidar data. Canadian Journal of Remote Sensing. 30(6), pp. 855-866.

Hasenauer H (1997) Dimensional relationships of open-grown trees in Austria. Forest Ecology and Management, 96 (1997), pp. 197-206

Hay GJ, Castilla G, Wulder MA and Ruiz JR (2005) An automated object-based approach for the multiscale image segmentation of forest scenes. International Journal of Applied Earth Observation and Geoinformation, Vol. 7, pp 339-359.

Hollaus M, Wagner W, Eberhöfer C and Karel W (2006) Accuracy of large-scale canopy heights derived from LiDAR data under operational constraints in a complex alpine environment, ISPRS Journal of Photogrammetry and Remote Sensing, 60(5), 323-338.

IUFRO (1958) Empfehlungen für die Baumklassenbildung und Methodik bei Versuchen über die Wirkung von Waldpflegemaßnahmen IUFRO, 12. Kongreß Oxford, 1956, Bd 2 London

Kini A, Popescu SC (2004) TreeVaW: a versatile tool for analyzing forest canopy LIDAR data: A preview with an eye towards future. In: CD-ROM Proceedings, ASPRS 2004 Fall Conference, Kansas City, Missouri , Sept. 12-16, 2004

Koch B, Heyder U and Weinacker H (2006) Detection of Individual Tree Crowns in Airborne Lidar Data. In: Photogrammetric Engineering & Remote Sensing PE&RS Vol. 72, No. 4, April 2006, pp. 357 -363.

Lang S (2002) Zur Anwendung des Holarchiekonzeptes bei der Generierung regionalisierter Segmentierungsebenen in höchstauflösenden Bilddaten. – In: Blaschke, T. (ed.), Fernerkundung und GIS: Neue Sensoren – innovative Methoden, Wichmann Verlag, Heidelberg, pp 24-33.

Lang S and Langanke T (2006) Object-based mapping and object-relationship modeling for land use classes and habitats. PFG - Photogrammetrie, Fernerkundung, Geoinformatik, Vol. 1, pp 5-18.

Lillesand TM, Kiefer RW and Chipman JW (2004) Remote Sensing and Image Interpretation. 5th ed. New York: John Wiley & Sons, Inc.

Maier B (2005) Analyse von LiDAR-Baumkronen-Modellen mit Filtertechniken. Unpublished term paper, Salzburg University, Austria.

Maltamo M, Mustonen K, Hyyppä J, Pitkänen J and Yu X (2004) The accuracy of estimating individual tree variables with airborne laser scanning in boreal nature reserve. Canadian Journal of Forest Research, Vol. 34, pp.1791-1801

Ochs T,. Sschneider T., Heurich M and Kennel E (2003) Entwicklung von Methoden zur semiautomatisierten Totholzinventur nach Borkenkäferbefall im Nationalpark Bayerischer Wald. In: Strobl, J., Blaschke T. & Griesebner, G. (Hrsg.): Angewandte Geographische Informationsverarbeitung XV. Beiträge zum AGIT-Symposium Salzburg 2003, H. Wichmann Verlag, Heidelberg, pp. 336-341.

Pitkänen J (2001) Individual tree detection in digital aerial images by combining locally adaptive binarization and local maximal methods. Canadian Journal of Forest Research, Vol. 31, pp. 832-844.

Pitkänen J, Maltamo M, Hyyppä J, Wei Yu X (2004) Adaptive methods for individual tree detection on airborne laser based canopy height model. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-8/W2, Freiburg, pp. 187-191.

Schieler K, Hauk E (2001) Instruktion für die Feldarbeit. Österreichische Waldinventur 2000/2002: Fassung 2001. Forstliche Bundesversuchsanstalt, Wien, 209 p

Schöpfer E and Lang S (2006) Object fate analysis – a virtual overlay method for the categorisation of object transition and object-based accuracy assessment. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. No. XXXVI-4/C42, Salzburg, Austria.

Schnadt K, Katzenbeisser R (2004) Unique airborne fiber scanner technique for application-oriented LiDAR products. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-8W2, Freiburg, pp. 19-24.

Strahler AH, Woodcock CE and Smith J (1986). Integrating per-pixel classification cover classification. Remote Sensing of the Environment, Vol. 71, pp. 282-296.

Tiede D, Lang S and Hoffmann C (2006) Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. No. XXXVI-4/C42, Salzburg, Austria.

Tiede D and Hoffmann C (2006) Process oriented object-based algorithms for single tree detection using laser scanning data. EARSeL-Proceedings of the Workshop on 3D Remote Sensing in Forestry, 14th-15th Feb 2006, Vienna, 151-156.

Tiede D, Hochleitner G and Blaschke T (2005) A full GIS-based workflow for tree identification and delineation using laser scanning. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVI, Part 3/W24, Vienna, pp. 9 – 14

Tiede D, Heurich M and Blaschke T (2004a) Object-based semi automatic mapping of forest stands with Laser scanner and Multi-spectral data. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-8W2, Freiburg, pp. 328-333.

Tiede D, Burnett C and Heurich M (2004b) Objekt-basierte Analyse von Laserscanner- und Multispektraldaten zur Einzelbaumdelinierung im Nationalpark Bayerischer Wald. In: S. Strobl, T. Blaschke, G. Griesebner (Hrsg.), Angewandte Geoinformatik 2004, Wichmann Verlag, Heidelberg, Germany, pp. 690–695.

Wehr A and Lohr U (1999) Airborne Laser scanning - an introduction and overview. ISPRS Journal of Photogrammtry & Remote Sensing, 54, pp. 68-82

Wulder M, Niemann KO and Goodenough DG (2000) Local Maximum Filtering for the Extraction of Tree Locations and Basal Area from High Spatial Resolution Imagery. Remote Sens. Environ., 73, pp. 103-114

Wulder M (1998) Optical remote sensing techniques for the assessment of forest inventory and biophysical parameters. Progress in Physical Geography; 22(4), pp. 449-476.

Zingg, A., (1999) Genauigkeit und Interpretierbarkeit von Oberhöhen. Cent.bl. gesamte Forstwes. 116, 1 / 2, pp. 25-34.

Acknowledgments

Chapter 2.3

Object recognition and image segmentation: the Feature Analyst® approach

D. Opitz, S. Blundell

Visual Learning Systems, Missoula, MT 59803. USA. David.Opitz@overwatch.com.

ABSTRACT: The collection of object-specific geospatial features, such as roads and buildings, from high-resolution earth imagery is a time-consuming and expensive problem in the maintenance cycle of a Geographic Information System (GIS). Traditional collection methods, such as hand-digitizing, are slow, tedious and cannot keep up with the ever-increasing volume of imagery assets. In this paper we describe the methodology underlying the Feature Analyst automated feature extraction (AFE) software, which addresses this core problem in GIS technology. Feature Analyst, a leading, commercial AFE software system, provides a suite of machine learning algorithms that learn on-the-fly how to classify object-specific features specified by an analyst. The software uses spatial context when extracting features, and provides a natural, hierarchical learning approach that iteratively improves extraction accuracy. An adaptive user interface hides the complexity of the underlying machine learning system while providing a comprehensive set of tools for feature extraction, editing and attribution. Finally, the system will automatically generate scripts that allow batch-processing of AFE models on additional sets of images to support large-volume, geospatial, data-production requirements.

1 Introduction

High-resolution satellite imaging of the earth and its environment repre-
sents an important new technology for the creation and maintenance of
geographic information systems (GIS) databases. Geographic features such
as road networks, building footprints, vegetation, etc. form the backbone
of GIS mapping services for military intelligence, telecommunications, ag-
riculture, land-use planning, and many other vertical market applications.
Keeping geographic features current and up-to-date, however, represents a
major bottleneck in the exploitation of high-resolution satellite imagery.
The Feature Analyst software provides users with a powerful toolset for
extracting object-specific, geographic features from high-resolution pan-
chromatic and multi-spectral imagery. The result is a tremendous cost sav-
ings in labor and a new workflow process for maintaining the temporal
currency of geographic data.

Until recently there were two approaches for identifying and extracting
objects of interest in remotely sensed images: *manual* and *task-specific
automated*. The manual approach involves the use of trained image ana-
lysts, who manually identify features of interest using various image-
analysis and digitizing tools. Features are hand-digitized, attributed and
validated during geospatial, data-production workflows. Although this is
still the predominant approach, it falls short of meeting government and
commercial sector needs for three key reasons: (1) the lack of available
trained analysts; (2) the laborious, time-consuming nature of manual fea-
ture collection; and (3) the high-labor costs involved in manual production
methods.

Given these drawbacks, researchers since the 1970s have been attempt-
ing to *automate* the object recognition and feature extraction process. This
was commonly effected by writing a task-specific computer program
(McKeown 1993; Nixon and Aguado 2002). However, these programs
take an exceedingly long time to develop, requiring expert programmers to
spend weeks or months *explaining*, in computer code, visual clues that are
often trivially obvious to the human eye. In addition, the resulting hand-
crafted programs are typically large, slow and complex. Most importantly,
they are operational only for the specific task for which they were de-
signed; typically failing when given a slightly different problem such as a
change in spatial resolution, image type, surface material, geographic area,
or season. Developing such programs is complicated by the fact that user
interest varies significantly. While some task-specific automated ap-
proaches have been successful, it is virtually impossible to create fully

automated programs that will address all user needs for every possible future situation.

The Feature Analyst approach to object-recognition and feature extraction overcomes these shortcomings by using inductive learning algorithms (Mitchell 1997, Quinlan 1993, Rumelhart et al 1986) and techniques to model the object-recognition process. Using Feature Analyst, the user provides the system (computer program) with several examples of desired features from the image. The system then automatically develops a model that correlates known data (such as spectral or spatial signatures) with targeted outputs (i.e., the features or objects of interest). The resulting learned model classifies and extracts the remaining targets or objects in the image. Feature models can be cached in a repository, known as the Feature Model Library, for later use. The accompanying workflow and metadata (information on spectral bandwidth, date and time stamp, etc.) can be used to quickly compose new models for changing target conditions such as geographic location or hour of day.

2 Learning Applied to Image Analysis

An *inductive learner* is a system that learns from a set of labeled examples. A teacher provides the output for each example, and the set of labeled examples given to a learner is called a *training* set. The task of inductive learning is to generate from the training set a concept description that correctly predicts the output of all future examples, not just those from the training set. Many inductive-learning algorithms have been previously studied (Quinlan 1993; Rumelhart 1986). These algorithms differ both in their concept-representation language, and in their method (or *bias*) of constructing a concept within this language. These differences are important because they determine the concepts that a classifier induces.

Nearly all modern vision systems rely on hand-crafted determinations of which operators work best for an image and what parameter settings work best for those operators (Maloof 1998; McKeown 1996). Such operators not only vary across the desired object to be recognized, but also across resolutions of the same image. Learning in object recognition tasks works by (a) acquiring task-specific knowledge by watching a user perform the tasks and (b) then refining the existing knowledge based on feedback provided by the user. In this approach, the parameters for these objects are tuned by the learning algorithm "on-the-fly" during the deployment of the algorithm. It is not surprising, therefore, that visual learning (Evangelia

2000; Nayar and Poggio 1996) can greatly increase the accuracy of a visual system (Burl 1998; Kokiopoulou and Frossard 2006).

In addition to increasing overall accuracy, visual learning can yield object-recognition systems that are much *easier* and *faster* to develop for a particular problem and resolution (Opitz and Blundell 1999). A model can be trained for one task, and then used to "seed" the development of a similar problem. This supports the immediate deployment of a new problem with the ability to fine-tune and improve itself though experience. By having a learning system at the core of the object recognition task, one can easily transfer pertinent knowledge from one problem to another, even though that knowledge may be far from perfect. This approach overcomes prior research on visual learning that primarily consisted of hard-coded, problem-specific programs.

3 Feature Analyst

In 2001 Visual Learning Systems, Inc. (VLS) developed Feature Analyst as a commercial off-the-shelf (COTS) feature extraction extension for ESRI's ArcGIS™ software in response to the geospatial market's need for automating the production of geospatial features from earth imagery. Feature Analyst is based on an inductive learning approach to object recognition and feature extraction. Feature Analyst was developed as a plug-in toolset for established GIS and remote sensing software packages (ArcGIS, IMAGINE, SOCET SET, GeoMedia, and RemoteView) in order to integrate the AFE workflow into traditional map production environments. In the Feature Analyst system, the image analyst creates feature extraction models by simply classifying on the computer screen the objects of interest in a small subset of the image or images (Opitz and Blundell, 1999). This approach leverages the natural ability of humans to recognize complex objects in an image. Users with little computational knowledge can effectively create object-oriented AFE models for the tasks under consideration. In addition, users can focus on different features of interest, with the system dynamically learning these features.

Feature Analyst provides a paradigm shift to AFE and distinguishes itself from other learning and AFE approaches in that it: (a) incorporates advanced machine learning techniques to provide unparalleled levels of accuracy, (b) utilizes spectral, spatial, temporal, and ancillary information to model the feature extraction process, (c) provides the ability to remove clutter, (d) provides an exceedingly simple interface for feature extraction, (e) automatically generates scripts of each interactive learning process,

which can be applied to a large set of images, and (f) provides a set of clean-up and attribution tools to provide end-to-end workflows for geospatial data production.

3.1 Feature Analyst Learning Approach

Feature Analyst does not employ a single learning algorithm, but rather uses several different learning approaches depending on the data. The base learning algorithms for Feature Analyst are variants of artificial neural networks (Rumelhart 1986), decision trees (Quinlan 1993), Bayesian learning (Mitchell 1997), and K-nearest neighbor (Mitchell 1997); however, the power of Feature Analyst is believed to be largely due to the power of ensemble learning (Opitz 1999).

Research on ensembles has shown that ensembles generally produce more accurate predictions than the individual predictors within the ensemble (Dieterich 2002; Opitz and Maclin 1999). A sample ensemble approach for neural networks is shown in Fig 1, though any classification method can be substituted in place of a neural network (as is the case with Feature Analyst). Each network in the ensemble (networks 1 through N) is trained using the training instances for that network. Then, the predicted output of each of these networks is combined to produce the output of the ensemble (Ô in Fig. 1). Both theoretical research (Opitz and Shavlik 1999; Dieterich 2002) and empirical work (Opitz and Maclin 1999; Opitz and Shavlik 1996) have shown that a good ensemble is one in which (1) the individual networks are accurate, and (2) any errors that these individual networks make occur in different parts of the ensemble's input space.

Much ensemble research has focused on how to generate an accurate, yet diverse, set of predictors. Creating such an ensemble is the focus of Feature Analyst ensemble algorithm. Feature Analyst searches for an effective ensemble by using genetic algorithms to generate a set of classifiers that are accurate and diverse in their predictions (Opitz 1999). The Feature Analyst approach has proven to be more accurate on most domains than other current state-of-the-art learning algorithms (including Bagging and Boosting) and works particularly well on problems with numerous diverse inputs, such as high-resolution, multi-spectral and hyperspectral images (Opitz 1999).
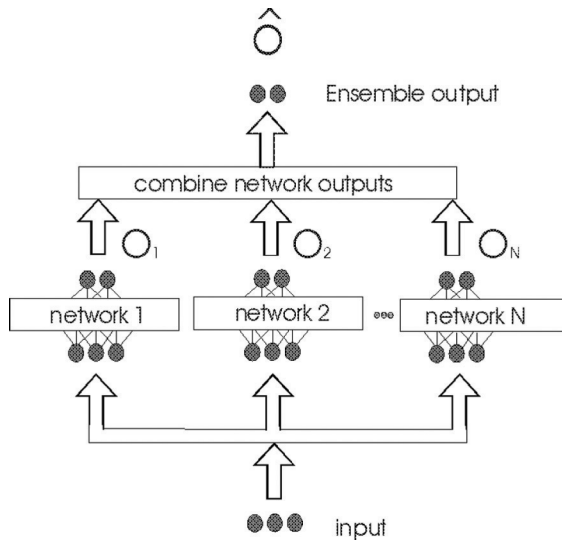
Fig. 1. Predictor ensemble for classifying data

3.2 Spatial and Ancillary Information in Learning

When classifying objects in imagery, there are only a few attributes accessible to human interpreters. For any single set of imagery these object recognition attributes include: shape, size, color, texture, pattern, shadow, and association. Traditional image processing techniques incorporate only color (spectral signature) and perhaps texture or pattern into an involved expert workflow process; Feature Analyst incorporates all these attributes, behind the scenes, with its learning agents. The most common (and intended) approach is to provide the learning algorithm with a local window of pixels from the image. The prediction task for the learner is to determine whether or not the center pixel is a member of the current feature theme being extracted. Fig. 2 demonstrates this approach. The center pixel for which the prediction is being made is represented in black. In this case, 80 surrounding pixels are also given to the learning algorithm (i.e., there is a 9x9 pixel window). The learner's task is to develop a model between the 81 inputs and the one output (whether or not the center pixel is part of the feature). This approach has distinct advantages: It works well; it is general purpose and applies to any object-recognition task; and it can easily accommodate any image transformation (e.g., edge detection) by simply supplying the pixels of the transformed image. The main *draw-*

backs to this approach are (1) its inability to take into account the spatial context of the larger image space, which can lead to false classifications, and (2) the large amount of information (i.e., pixels) to be processed, which taxes computer memory and slows down processing.

As a result of these drawbacks, Feature Analyst offers various input representations based upon the concept of "foveal" vision (Opitz and Bain 1999), shown in Fig. 3. With a foveal representation, a learning algorithm is given a region of the image with high spatial resolution at the center (where the prediction is being made) and lower spatial resolution away from the center. Such an approach mimics the visual process of most biological species, including humans (i.e., peripheral vision). Foveal representation provides contextual spatial information to the learning algorithm while not overwhelming it when making a local decision (e.g., Is the center pixel part of an armored vehicle?).
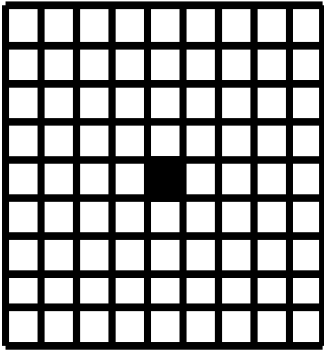

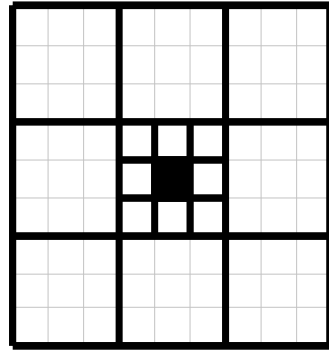
Fig. 2. Traditional Representation   **Fig. 3.** Foveal Representation

In Fig 3, foveal representation provides only 17 inputs to the learner when considering a 9x9 pixel region. Each outer 3x3 region gives the average of the 9 pixels as one input to the learning algorithm. The analyst can widen the range of foveal vision by making the next outer layer an average of a 9x9 region and so on. Thus a 27x27 region would provide only 25 inputs to the learner. Having the Learner concentrate on the center pixels, while taking into account the gist of the outer pixels, represents a great strength in using spatial context for object-recognition tasks. Foveal and other analogous input representations provided by Feature Analyst (such as Bullseye) are making a major breakthrough in automated feature extraction, as they greatly reduce the amount of data given to the Learner— especially important when extracting targets from cluttered scenes.

Figs. 4-6 show the value of using object-recognition attributes and spatial context in feature extraction tasks. Fig. 4 is the sample image. Here the object is to extract white lines on airport runways. Using only spectral information, the best an analyst can do is shown in Fig. 5; the results show all materials with similar "white" reflectance extracted. Fig. 6 shows the Feature Analyst results for extracting white lines using both spectral values and spatial parameters. In this case, the knowledge of the adjacent pavement or grass pixels is included when extracting the white line. This example illustrates the need to take into account spatial information when conducting object-recognition tasks with imagery.

Fig. 4. Original image: the objective is to extract only the thin white lines on the runway

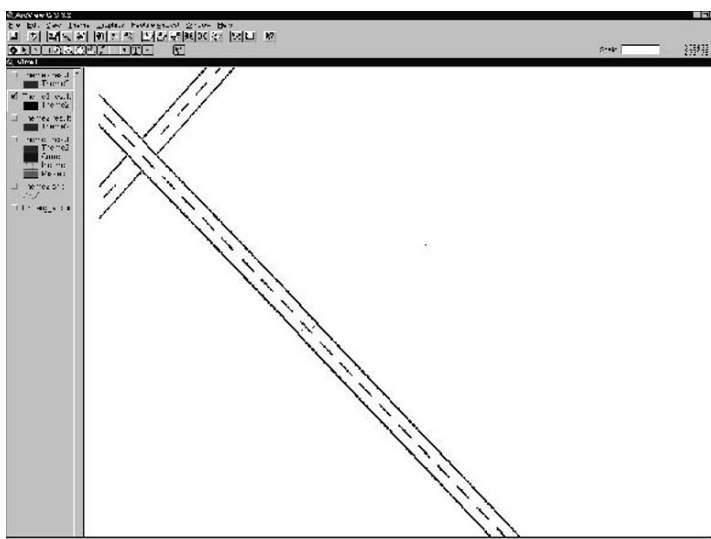**Fig. 5.** Extraction results without the use of spatial attributes

Fig. 6. Feature Analyst classification using spatial attributes to extract only the targeted white lines on the runway

3.3 Hierarchical Learning

The Feature Analyst hierarchical workflow consists of the following steps:

1. User digitizes a few examples of the target. Note that in the previous extraction example, the user only had to digitize 3 or 4 small examples for the learning algorithm to extract all of the features correctly.
2. User selects the feature type from the graphical user interface automatically setting all of the learning parameters behind the scenes.
3. User extracts features using a One-Button approach.
4. User examines results and, if required, provides positive and negative examples to remove clutter using Hierarchical Learning.
5. User refines the first-pass predictions, removing clutter with another pass of learning (or removing shape characteristics using the Feature Analyst, Remove Clutter by Shape tool).
6. The user repeats steps 4 and 5 as necessary.

Clutter is the most common form of error in feature extraction. The objective of clutter mitigation is to remove false positives. Thus, the learning task is to distinguish between false positives and correctly identified positives. The user generates a training set by labeling the positive features from the previous classification as either positive or negative. The trained learner then classifies only the positive instances from the previous pass. The negative instances are considered correct in clutter mitigation and are thus masked out.

Hierarchical learning is necessary for learning complex targets in high-resolution imagery. The overall process iteratively narrows the classification task into sub-problems that are more specific and well defined. The user begins the hierarchical process the same as any baseline inductive learning classification, i.e., select labeled examples for the feature being extracted, train the learner, and then classify every pixel in the image based on the learner's prediction. At this point, if not satisfied with the results, the user can apply a hierarchy of learners to improve the classification. The classification is improved in passes where each new pass is designed to remove one form of error from the results of the previous pass.

3.4 Automation of Feature Extraction Tasks

Automated feature extraction has been the long-term goal of geospatial data production workflows for the past 30 years. The challenge is developing a flexible approach for transferring domain knowledge of a feature extraction model from image-to-image that is capable of adapting to chang-

ing conditions (image resolution, pixel radiometric values, landscape seasonal changes, and the complexity of feature representation). In 2006, VLS introduced the Feature Modeler and the Feature Model Library (FML) tools. These additions to Feature Analyst serve to automate the feature extraction workflow process. Feature Modeler provides users with a comprehensive set of tools for examining and refining feature models created with Feature Analyst (Fig. 7). Feature models, designated as AFE models, comprise the parameter settings for a classifier to extract particular features, including setting for spatial context and hierarchical learning passes. Benefits of this approach include the following:

- Analysts can create, edit and refine the inner workings of an AFE model, including pixels used for classification of a feature with spatial processing, priority of input bands, rule extraction from a complex learned model and the parameter settings for a learning algorithm.
- Technicians can access AFE models to run in an interactive mode or in a silent batch mode. In interactive mode, the technician does not need to be concerned with creating the proper workflow or setting parameters for the learning algorithm; but rather, need only provide a labeled set of examples. In batch mode the process is completely automated. A single AFE model, or multiple AFE models, can be run against a single image or a directory of images.
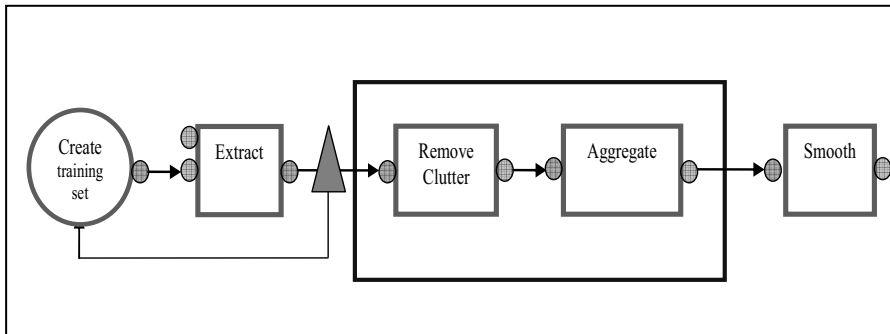

Fig. 7. A simple AFE model showing the five steps used in processing

The Feature Model Library resides within a relational database and is used for storing AFE models to support enterprise-wide geospatial processing. Analysts can search and retrieve AFE models for use in batch-mode processing to extract features from imagery without any training sets. The Feature Modeler application allows users to import AFE models,

examine and adjust parameter settings, or deploy a learning model in a batch-processing mode.

3.5 Cleanup and Attribution Tools

Object-recognition and feature extraction represent steps in a chain of process used in geospatial data production to collect features for a GIS database. Other steps in the process after feature extraction include feature editing (clean-up), feature generalization, feature attribution, quality control checks and then storage in the GIS database. In almost every instance of feature collection, the designated object needs to be stored as a *vector* feature to support GIS mapping and spatial analyses. Vector features, commonly stored in Shapefile format, can be stored as points, lines, polygons or TINs. One of the defining characteristics of geospatial vector feature data is the ability to define topology, store feature attributes and retain geopositional information.

Feature Analyst provides tools for the majority of these tasks, with an emphasis on semi-automated and automated vector clean-up tools and feature attribution. Feature representation of a road network requires that the road feature be collected as either a polygon or line feature or both. In either case, tools are required to adjust extracted vector features to account for gaps due to occlusion from overhanging trees, to eliminate dangles or overshoots into driveways, to fix intersections and to assist with a host of other tasks. As object-recognition systems evolve there is an ever-increasing expectation on the part of the user for a complete solution to the feature extraction problem.

5 Conclusions

Feature Analyst provides a comprehensive machine learning based system for assisted and automated feature extraction using earth imagery in commercial GIS, image processing and photogrammetry software. The AFE workflow, integrated with the supporting application tools and capabilities, provides a more holistic solution for geospatial data production tasks. The Feature Analyst user interface supports a simple feature extraction workflow whereby the user provides the system with a set of labeled examples (training set) and then corrects the predicted features of the learning algorithm during the clutter removal process (hierarchical learning). Benefits of this design include:

- **Significant time-savings in the extraction of 2-D and 3-D geo-spatial features from imagery**. O'Brien (2003) from the National Geospatial-Intelligence Agency (NGA) conducted a detailed study that indicated Feature Analyst is 5 to 10 times faster than manual extraction methods and more accurate than hand-digitizing on most features (Fig. 8).
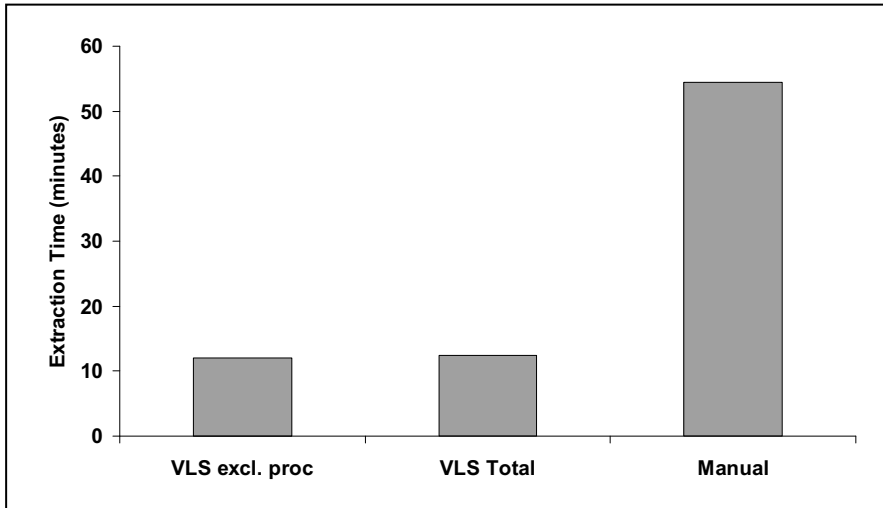

Fig. 8. NGA AFE test & evaluation program timing comparisons (O'Brien, 2003)

- **Significant increases in accuracy.** Feature Analyst has been shown to be more accurate than previous AFE methods and more accurate than hand digitizing on numerous datasets (Brewer et al 2005, O'Brien 2003).
- **Workflow extension capabilities to established software**. Analysts can leverage Feature Analyst within their preferred workflow on their existing ArcGIS, ERDAS IMAGINE, SOCET SET and soon Remote View systems, increasing operator efficiency and output.
- **A simple *One-Button* approach** for extracting features using the Feature Model Library, as well as advanced tools for creation of geospecific features from high resolution MSI, radar, LiDAR and hyperspectral data.
- **Open and standards-based software architecture** allowing third-party developers to incorporate innovative feature extraction algorithms and tools directly into Feature Analyst.

- **Interoperability amongst users** on different platforms. Expert analysts can create and store AFE models in the Feature Model Library, while other analyst can use these models for easy one-button extractions.
- **A simple workflow and user interface** hides the complexity of the AFE approaches.
- **High accuracy** with state-of-the-art learning algorithms for object recognition and feature extraction.
- **Post-processing cleanup tools** for editing and generalizing features, to providing an end-to-end solution for geospatial data production.
- **AFE modeling tools** for capturing workflows and automating feature collection tasks.

References

Burl M (1998) Learning to recognize volcanoes. Machine Learning Journal. vol 30 (2/3). pp 165-194.

Brewer K, Redmond R, Winne K, Opitz D, and Mangrich M. (2005) Classifying and mapping wildfire severity, Photogramm Eng and Remote Sensing. vol 71, no 11. pp 1311-1320.

Evangelia MT (2000) Supervised and unsupervised pattern recognition. CRC Press. Boca Raton, FL.

Dietterich TG (2002) Ensemble learning. The Handbook of Brain Theory and Neural Networks. vol 2. pp 110-125.

Kokiopoulou E and Frossard R (2006) Pattern detection by distributed feature extraction. IEEE International Conference on Image Processing. pp 2761-2764.

McKeown et al., (1993) Research in automated analysis of remotely sensed imagery. Proc of the DARPA Image Understanding Workshop. Washington DC.

McKeown D (1996) Top ten lessons learned in automated cartography. Technical Report CMU-CS-96-110. Computer science department. Carnegie Mellon University. Pittsburgh, PA.

Mitchell T (1997) Machine learning. McGraw Hill. New York, NY.

Nayar S, Poggio T (1996) Early visual learning. Oxford University Press. New York, NY.

Nixon MS, Aguado AS (2002) Feature extraction and image processing. Elsevier. Amsterdam.

Opitz D (1999) Feature selection for ensembles. Proc of the 16th National Conference on Artificial Intelligence. pp. 379-384.

Opitz D, Bain W (1999) Experiments on learning to extract features from digital images. IASTED Signal and Image Processing.

Opitz D, Blundell S (1999) An intelligent user interface for feature extraction from remotely sensed images. Proc. of the American Society for Photogrammetry and Remote Sensing. pp. 171-177.

Opitz D, Maclin R (1999) Popular ensemble methods: an empirical study. Journal of Artificial Intelligence Research. vol 11 (1). pp. 169-198.

Opitz D, Shavlik J (1996) Generating accurate and diverse members of a neural-network ensemble. Advances in Neural Information Processing Systems. vol 8. pp. 535-541.

Opitz D, Shavlik J (1999) Actively searching for an effective neural-network ensemble. Springer-Verlag Series on Perspective in Neural Computing. pp. 79-97.

O'Brien M (2003) Feature extraction with the VLS Feature Analyst System. ASPRS International Conference. Anchorage, AK.

Quinlan J (1993) C4.5: Programs for machine learning. Morgan Kaufmann. San Mateo, CA.

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-prapagation errors. Nature.

Chapter 2.4

A procedure for automatic object-based classification

P.R. Marpu[1], I. Niemeyer[1], S. Nussbaum[2], R. Gloaguen[1]

[1] Freiberg University of Mining and Technology, 09599, Freiberg, Germany, prashanthreddym@gmail.com

[2] Research Center Juelich, 52425 Juelich, Germany

KEYWORDS: Feature extraction, optimal features, threshold, separability

ABSTRACT: A human observer can easily categorize an image into classes of interest but it is generally difficult to reproduce the same result using a computer. The emerging object-based methodology for image classification appears to be a better way to mimic the human thought process. Unlike pixel-based techniques which only use the layer pixel values, the object-based techniques can also use shape and context information of a scene texture. These extra degrees of freedom provided by the objects will aid the identification (or classification) of visible textures. However, the concept of image-objects brings with it a large number of object features and thus a lot of information is associated with the objects. In this article, we present a procedure for object-based classification which effectively utilizes the huge information associated with the objects and automatically generates classification rules. The solution of automation depends on how we solve the problem of identifying the features that characterize the classes of interest and then finding the final distribution of the classes in the identified feature space. We try to illustrate the procedure applied for a two-class case and then suggest possible ways to extend the method for multiple classes.

1 Introduction

Object-based image classification is a promising methodology as it is close to human perception. A typical object-based classification system starts with segmenting the image into smaller homogeneous regions (or image objects). These objects correspond to approximations of real-world objects (Benz et al. 2004, Baatz et al. 2004). Every object is characterized by several features defined based on layer values, texture, shape and context of the object. Generally, the objects are classified using a defined rule base (Benz et al. 2004). This is where the possibility to automate the classification process becomes difficult. With a few input samples for every class and using the enormous object feature-space to our advantage, it is possible to automatically generate a rule base. However, the pre-condition obviously is a good segmentation result. We used the "Multi-resolution segmentation" of Definiens Professional[1] software in the present work.

The essential issue is to manage the huge information given by the color, shape, texture and context of the object. Only a few object features characterize a class and as the degrees of freedom increase, it gets increasingly difficult to identify the optimum features among the huge number of available features. In this article we summarize a few issues related to the object feature-space and thus make an attempt to provide a solution towards automation in object-based classification. It is possible to identify the optimum features based on a separability measure which quantifies the distance between the two random distributions and a simple Bayes' rule identifies the threshold of separation (Bhattacharya 1943, Nussbaum et al. 2005, Richards and Jia 1999).

An automatic classification procedure is prepared in this effect to minimize human involvement in classification steps. Such a procedure speeds up the process of classification when huge data is to be dealt with, however at the expense of accuracy. We first used this procedure to make algorithms for extracting morphology of geological faults using remote sensing data and in identifying fission tracks in microscopic images. In both the cases, there were only two classes i.e., class of interest and background. We then brief some possibilities of extending the methodology to the case of multiple classes. It is also interesting to see how the notion of the object helps in the post-classification analysis of objects.

---

[1] http://www.definiens.com

2 Theoretical background

2.1 Distance between random distributions

A popular measure of distance between two random distributions is the Bhattacharya distance measure (B). For two random distributions given by probability density functions $p_1(x),\ p_2(x),\ B$ can be given as (Bhattacharya 1943, Richards and Jia, 1999):

$$B = -\ln\left(\int\sqrt{p_1(x)p_2(x)}dx\right). \qquad (1)$$

For a discrete case, we can approximate Eq. 1 as

$$B = -\ln\left(\sum_i \sqrt{p_1(x_i)p_2(x_i)}\Delta x\right), \qquad (2)$$

x_i refers to the discrete points and Δx refers to the sampling interval or typically the width of the bins in the histogram. The discrete probability density function is obtained by normalizing the histogram with respect to the total area under the histogram.

If we assume that the classes are normally distributed, then for classes C_1, C_2 of size n_1, n_2 with means m_1, m_2 and standard deviations σ_1, σ_2,

$$B = \frac{1}{8}(m_1 - m_2)^2 \frac{2}{\sigma_1^2 + \sigma_2^2} + \frac{1}{2}\ln\left(\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1\sigma_2}\right). \qquad (3)$$

The range of B falls in half-closed interval $[0,\infty)$. This range is transformed into the closed interval $[0, 2]$ by using a simple transformation leading to so called Jeffries - Matusita distance measure (J),

$$J = 2(1 - e^{-B}). \qquad (4)$$

$J = 0$ implies that the two distributions are completely correlated and $J = 2$ implies that the distributions are completely uncorrelated. For every feature, we can calculate the separability between the two classes using J. The features which have very high J value are the optimum features which characterize the classes.

2.2 Threshold of separation

We can distinguish between two random distributions by using a threshold of separation. For two classes C_1 and C_2 and an observation x_j, using the Bayes' rule we get,

$$p(C_1|x_j) = \frac{p(x_j|C_1)p(C_1)}{p(x_j)}, \tag{5}$$

and similarly for C_2. We then have

$$p(C_1|x_j) + p(C_2|x_j) = 1. \tag{6}$$

These are the only possibilities and the total probability should be equal to 1. The best decision threshold then is given by the relation

$$p(C_1|x_j) = p(C_2|x_j). \tag{7}$$

On rearranging the terms using the above equations, we have

$$p(x_j|C_1)p(C_1) = p(x_j|C_2)p(C_2). \tag{8}$$

For discrete functions, we can find the threshold as

$$\left| p(x_j|C_1)p(C_1) - p(x_j|C_2)p(C_2) \right| \approx 0. \tag{9}$$

For the case of normally distributed classes as mentioned in Sec. 2.1, we have the threshold, T as:

$$T = x_j = \frac{m_2\sigma_1^2 - m_1\sigma_2^2 \pm \sigma_1\sqrt{m_1 - m_2^2 + 2A(\sigma_1^2 - \sigma_2^2)}}{(\sigma_1^2 - \sigma_2^2)},$$

$$\text{where } A = \log\left(\frac{\sigma_1}{\sigma_2} \times \frac{n_2}{n_1}\right). \tag{10}$$

The degree of misclassification depends on the separability. However, we can overcome this limitation to some extent by carefully shifting the threshold as will be discussed in Sec 3.1.

3 Towards automation

The solution towards automation in object-based classification depends on how we extract the necessary information from the huge information associated with the objects.

3.1 The procedure

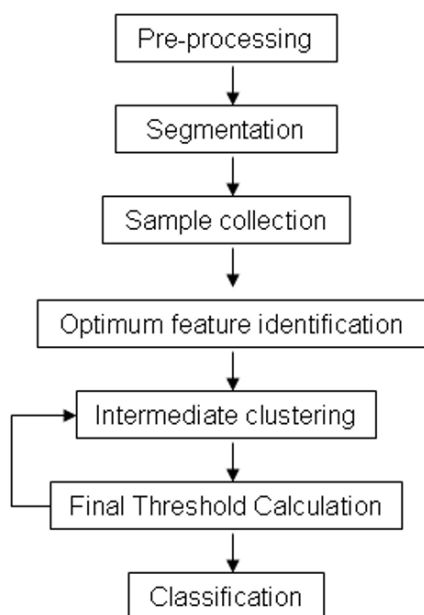The method for classifying the objects of a class of interest in an image is illustrated in Fig. 1.

Fig. 1. The procedure for automatic object-based classification

The goal of *pre-processing* is to increase the homogeneity of the subsequently extracted objects. Different image scenarios will require different pre-processing steps. Sometimes the pre-processing step can be crucial because the basic step of object-based classification is to first generate approximate real world objects and it requires that these object regions in the image are homogeneous.

After segmenting the image into primitive objects, few samples are collected. This can be done by selecting samples in the image based on hu-

man interpretation or by selecting specific regions in the concerned image histograms. For example if the objects of class of interest are characterized by bright regions compared to other objects in the image, then 2-5% of the objects in the image histogram which have high mean values are taken as samples of class of interest and 2-5% of image objects which have low mean values are assigned as samples to the background class.

When we have samples of the classes, we can automatically identify the *optimal features* using the distance, J defined in the eqn. (4). Since the samples cannot give the information about what kind of probability distribution the class has, we first assume that the distribution is Gaussian. However in the next step we try to get an approximate distribution of the classes and thus reassess the feature space. The optimum features will have high J value.

We now try to identify the approximate probability distributions of the classes and try to validate if the features we identified in the previous step are the best features defining that class. From the samples and features obtained in the earlier stages we cluster the objects into two classes. This will give the approximate distribution of the classes. Any clustering technique such as minimum distance, Fuzzy clustering, etc., can be used. However, to represent all the features on a common scale (in this case [0,1]), the feature values are normalized before clustering. For every object feature value F of a particular feature,

$$F_1 = F - F_{\min}$$

$$F' = \frac{F_1}{F_{1\max}} \tag{11}$$

F_{min} is the minimum value of the object feature values of that feature; F_{1max} is the maximum of values F_1 obtained in first step. F' is the transformed feature value of F. Again, any clustering algorithm can be used at this stage and some non-linear clustering schemes might not require that the data is normalized.

After clustering, we again check for the separability of features and hence find the threshold of separation for the features having high J value.

The thresholds are found based on the Bayes' rule defined in eqn (8). However there will be some misclassifications if we just classify the objects based on these thresholds. We can sometimes solve this problem by moving the threshold away from the mean of the class of interest (see Fig. 2).
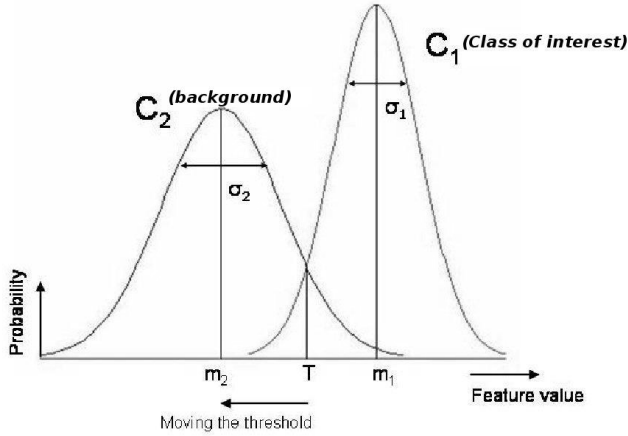
Fig. 2. Finding the threshold

This can be achieved by defining a simple criterion using the separability measure. If *J < 0.5*, the feature is ignored.

Otherwise,

$$T' = m_2, \qquad \text{for } 0.5 \leq J \leq 1.25$$

$$T' = (T + m_2)/2, \qquad \text{for } 1.25 < J < 1.75$$

$$T' = T, \qquad \text{for } J \geq 1.25. \qquad (12)$$

T' is the modified value of the threshold T and m_2 is the mean of the background class.

The above empirical criterion is based on observations using random data. It has been observed that this criterion is suitable when more features with a better *J* value exist.

The objects are finally classified using these thresholds. The classification is based on an 'AND' operation which is equivalent to sequentially eliminating the background objects from the class of interest using the optimum features.

3.2 Extending it to the multiple classes

We can extend the procedure defined in the previous section to the case of multiple classes in several ways. To give some examples the following methods can be employed.

1. We can identify the optimum features and use any clustering algorithm.
2. We can sequentially extract one class at a time using the procedure defined in section 3.1. But this can lead to objects with multiple classifications. We can reclassify the objects with multiple classifications using the minimum distance to the distributions.
3. We can use a fuzzy rule base to classify the objects. This method can efficiently minimize the number of multiple classifications.
4. A neural network classifier can be designed based on the samples and the optimum features.

The identification of the optimum features characterizing the classes is thus an important step.

4 Case studies

4.1 Extracting the morphology of normal faults

4.1.1 Background

Normal faults are formed in the rifts. A *rift* is a region where the Earth's crust and lithosphere are under extensional strain, hence forming a series of Horst and Graben segments. The fractures generated in this process of rifting are normal faults. Fig. 3 shows the geometry in a rift. The faults grow in time and form a network of faults which is fractal.
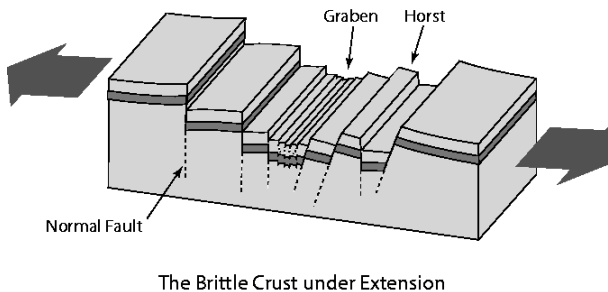
Fig. 3. The fault geometry in a rift

A DEM can be used for identification of the faults as the faults are characterized by steep slopes. In this context an algorithm is prepared using the defined procedure to process the DEM and automatically classify the faults in the region described by the image. The algorithm is applied on the data

of Lake Magadi area in South Kenya. This area is chosen as it is an active fault zone. Such a fault map is essential to analyze the deformation of the region. For example, studying the fault statistics to calculate the fractal dimension gives an indirect measure to quantify the deformation of the region. The concept of the object thus helps in determining the statistics immediately after classifying as the resulting map contains objects of faults.

4.1.2 Algorithm and results

In the pre-processing step two different approximations of the gradient (Sobel gradient, slope) of a DEM are used as image layers. This is done to emphasize the faults while segmenting the image. Faults are characterized by high derivative values in the DEM. For the same resolution in the images, the segmentation parameters will be the same. In the present case we used a scale parameter of 30, shape factor of 0.2 and compactness of 0.5.

Samples are then selected from the concerned histograms of the means of the objects. Fig. 4 shows the derivative of the DEM from Lake Magadi region in south Kenya. Fig. 5 shows the result of automatically classifying the DEM.
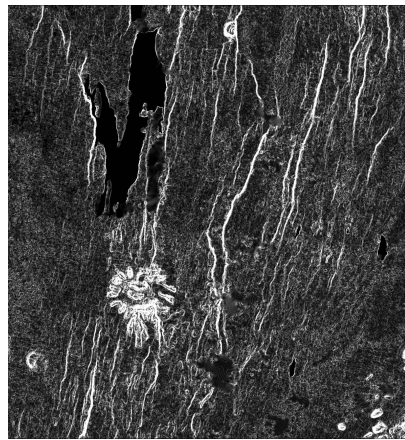

Fig. 4. Derivative of the DEM near Lake Magadi, South Kenya
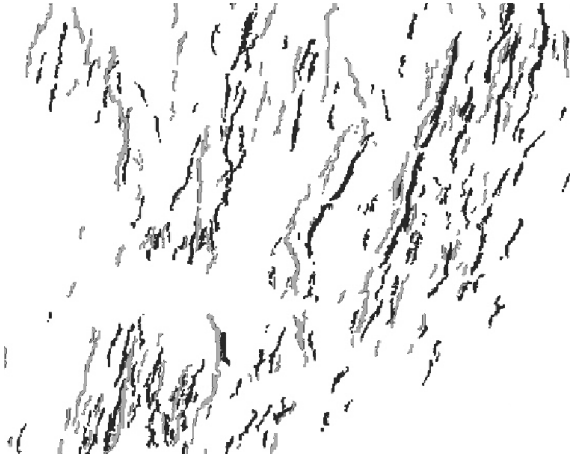
Fig. 5. Faults in the DEM (unwanted areas like volcanoes are masked)

As mentioned earlier, the notion of the object helps in the image analysis after the classification. After the classification every real world fault is a fault object in the image. So, it is easy to estimate the statistics of the faults using the faults objects in the image. The shape of the object can be used to find the approximate width and length of the faults. This statistics help in determining the fractal dimension of the faults in that region, which has many applications in geosciences (Turcotte 1997). Fig. 6 shows the length distribution and the corresponding fractal dimension calculated using that information. The linearity of the curve in the log-log plot of the cumulative histogram of lengths of faults also validates the method used in the present work. Such post- classification applications of objects make the object-based methodology a powerful tool in image analysis.
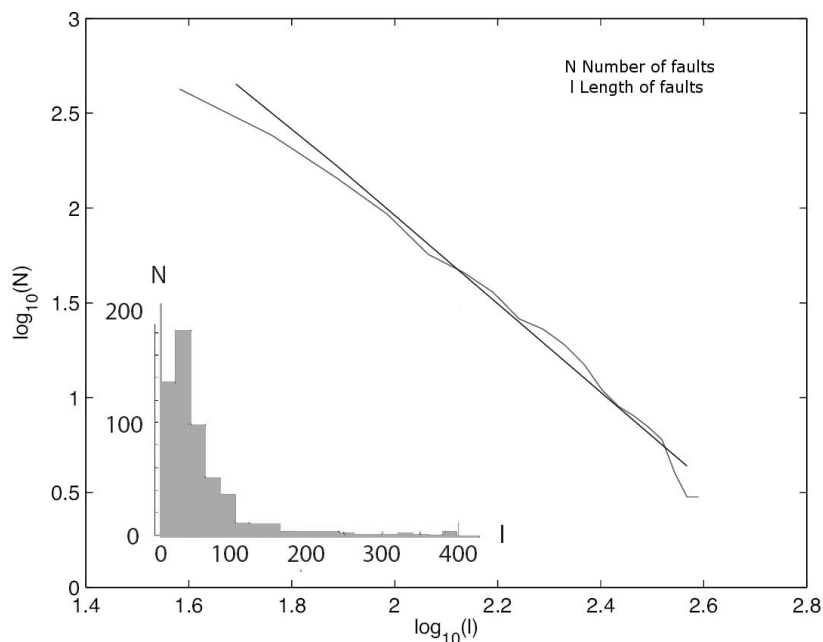
Fig. 6 The histogram of the faults lengths and the corresponding cumulative histogram curve in logarithmic scale to find fractal dimension.

4.2 Identifying fission-tracks in microscopic images

4.2.1 Background

The fission-track dating method is now commonly used in geological research but hindered by time consuming track counts and length measurements. Attempts at automation using conventional image analysis techniques on digital data have hitherto proved of limited practical use (Morgan et al, 2003). To automate the process of counting the tracks, we try to mimic human thinking procedures. We first identify all the tracks in the image and then count the number of tracks by accounting for the number of intersections. The first step in doing that is to automatically classify the tracks and define *tracks* objects. When the tracks are identified, then we can work on identifying algorithms to count the individual tracks by counting the intersections. We used the procedure described in this article to develop an algorithm for first identifying the *tracks* objects in the image.

4.2.2 Algorithm and results

The pre-processing step involves image enhancement and morphological closing to generate homogeneous regions of tracks. It has been observed in several of the acquired images that the standard deviation of the pixel values of the fission-tracks pixels is very high compared to that of the background. It is therefore the objective of the pre-processing step to generate homogeneous areas by reducing the standard deviation of the fission-tracks pixels. The following transformation is applied to the image

$$I_1 = \sqrt{I^2_{max} - I^2} \qquad\qquad (13)$$

I_{max}, is the maximum value in the image I. The image I_1 is then inverted so as to represent the tracks pixels in dark as the above transformation inverts the image histogram by compressing the lower part of the histogram. Then, on the resulting image, morphological closing operation is performed so as to generate homogeneous regions. Fig. 7b shows the result of processing the image shown in Fig. 7a using the described method.
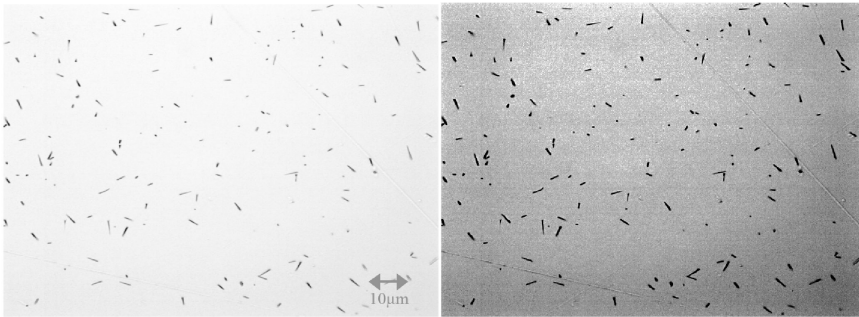

Fig. 7. (a) A microscopic image showing tracks to be identified **(b)** Processed image using image enhancement and morphological closing
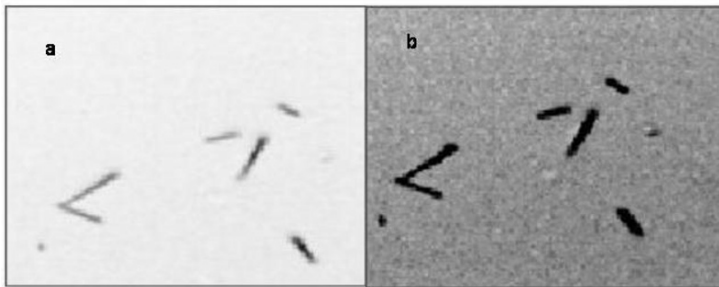

Fig. 8. A comparison using a close-up. It can be observed that the tracks are more homogeneous

The rest of the algorithm is the same as in the case of faults. The same distance measure is used to find the optimal features. The segmentation parameters are to be identified separately for different images based on lens magnifications. For a set of images acquired with the same camera on the same microscope using the same lens magnification, the same set of segmentation parameters will apply. So, it is necessary to fist define standards for image acquisition so as to use the same segmentation parameters. In the present case we used a scale parameter of 20, shape factor of 0.2 and compactness of 0.5.
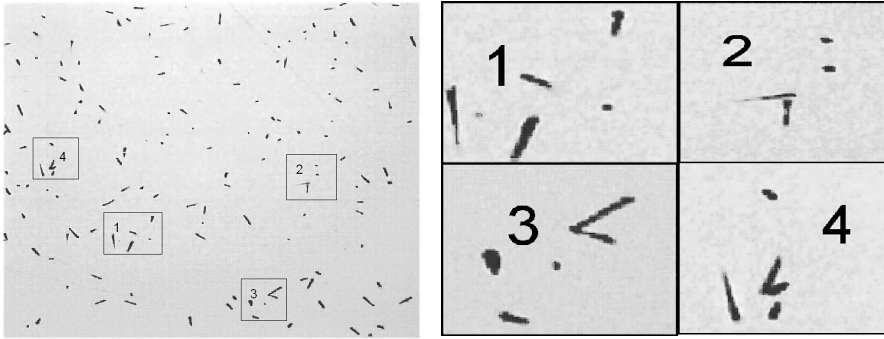

Fig. 9. (a) Classified Tracks **(b)** Zooming the regions 1,2,3,4 in the classified image

Tracks are characterized with low pixel values in the image. So, objects at the lower end of the histogram of the object means are selected as samples for object class of interest i.e., *tracks*. And, the objects at the higher end of the histogram are taken as samples of background objects class. In the present case a microscopic image acquired with a lens magnification of 20 is used. The result of classification is shown in Fig. 9. The object-based classification extracts the shape of the tracks with a good overall accuracy based on manual verification of five images. The classified objects are then ready for the next stage where the shape of the object is analyzed to detect the intersections so as to count the exact number of tracks in the image. The concept of using objects for classification thus helps in using the object information directly in the next stages where the tracks are counted.

4.3 Classifying Multiple Classes

We provide a simple example to give an idea of the potential of the proposed procedure in a multiple class problem. A small part of pan-

sharpened high resolution image of 61 cm resolution (acquired by Quick-bird satellite in 2003) of Esfahan nuclear facility, Iran (Fig. 10a) is used to classify for various classes such as *buildings, shadows, roads* and *background*. A scale parameter of 100 and shape factor of 0.2 was used for segmentation. We tried to extract the classes based on the features identified as best features using the separability measure when one class is compared with all the other classes. Fig. 10 shows the graphical user interface (GUI) which displays the separability (i.e., Jeffries-Matusita distance, *J*) and the threshold of separation between pairs of classes for all the features.
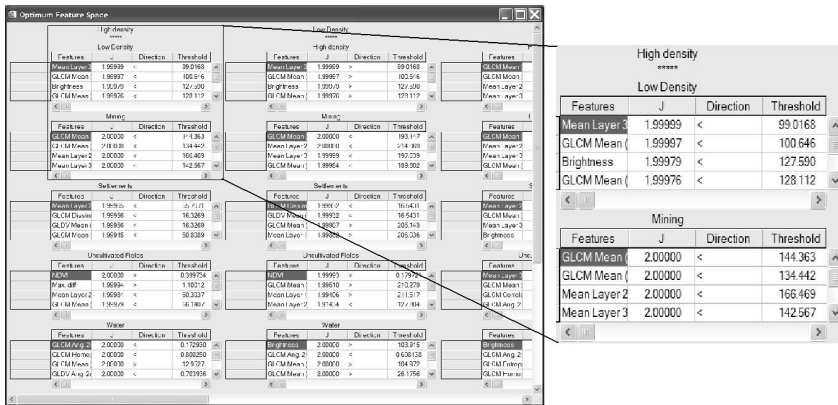

Fig. 10. The GUI for displaying separability and thresholds

The features which have high separability and are prominent for majority of the classes are selected as optimum features for classification. Fig.11b shows the result of classifying Fig.11a using a minimum-distance clustering method after identifying the suitable features. The best feature for every class combination is used in this example.

Fig. 11. (a) A part of the Esfahan Nuclear facility, Iran. **(b)** Classification result using minimum distance clustering

However, different classes are to be extracted at different scales and it is difficult to identify at which scale we have to extract each of the classes. This is a classical problem in object-based image analysis which can be partly overcome using a multi-resolution analysis. In the present example, most of the buildings and roads are classified properly for the segmentation parameters used here. We intend to extend the present work to classify an image into multiple classes by trying to identify the level of extraction for different classes. A semi-automatic method in the same lines has already been implemented successfully by Nussbaum et al (Nussbaum et al. 2005).

5. CONCLUSION

Object-based classification comes close to human perception. However, the huge information associated with the objects, hinders the proper utilization of the strength of image-objects. We have shown that the optimum features which characterize a class can be identified. This separates the relevant information from the feature space. We then demonstrated how we can step towards a solution for automatic image classification using image objects.

References

Baatz et al (2004), eCognition Userguide. URL: http://www.definiens.com

Benz UC, Hoffmann P, Willhauck G, Lingenfelder I and Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information, ISPRS Journal of Photogrammetry & Remote sensing, Vol. 58, 239-258.

Bhattacharyya A (1943) On a measure of divergence between two statistical populations defined by probability distributions, Bull. Calcutta Math. Soc., 35 pp. 99-109

Morgan L, Meehan Q and Stewart R (2003) Fission-track dating using digital image analysis and optical reflected light, Geological Society of America Abstracts with Programs, 35/6, 73.

Nussbaum S, Niemeyer I and Canty MJ (2005) Feature Recognition in the Context of automated Object-Oriented Analysis of Remote Sensing Data monitoring the Iranian Nuclear Sites, Proceedings of Optics/Photonics in Security & Defence, SPIE.

Richards JA and Jia X (1999) Remote sensing digital image analysis. An introduction, Springer, Berlin.

Turcotte DL (1997) Fractals and Chaos in Geology and Geophysics, 2nd Edition. Cambridge University Press

Chapter 2.5

Change detection using object features

I. Niemeyer[1], P.R. Marpu[1], S. Nussbaum[2]

[1] Freiberg University of Mining and Technology, D-09599 Freiberg, Germany, irmgard.niemeyer@tu-freiberg.de

[2] Research Center Juelich, D-52425 Juelich, Germany

KEYWORDS: object changes, change detection, object-based, clustering, high-resolution satellite imagery

ABSTRACT: For the detection of changes, several statistical techniques exist. When adopted to high-resolution imagery, the results of traditional pixel-based algorithms are often limited. We propose an unsupervised change detection and classification procedure based on object features. Following the automatic pre-processing of the image data, image objects and their object features are extracted. Change detection is performed by the multivariate alteration detection (MAD), accompanied by the maximum autocorrelation factor (MAF) transformation. The change objects are then classified using the fuzzy maximum likelihood estimation (FMLE). Finally the classification of changes is improved by probabilistic label relaxation.

1 Introduction

Change detection is the process of identifying and quantifying temporal differences in the state of an object or phenomenon (Singh 1989). When using remotely sensed imagery from two acquisition times, each image pixel or object from the first time can be compared with the corresponding pixel or object from the second time in order to derive the degree of

change between the two times. Most commonly, differences in pixels' values, such as reflected radiance or spectral reflectance, have been taken as the measure of change in pixel-based change detection studies.

A variety of digital change detection techniques has been developed in the past three decades. Basically, the various algorithms can be grouped into the following categories: algebra (differencing, rationing, and regression), change vector analysis, transformation (e.g. principal component analysis, multivariate alteration detection, and Chi-square transformation), classification (e.g. multi-temporal clustering, post-classification comparison, expectation maximization algorithm, and neural networks) and hybrid methods. When applied on the image pixel level, these statistical techniques use the changes of the spectral, texture or transformed values, or transitions of the class membership (see Figure 1) as the measure of change.

Reviews on the most commonly used techniques are given by i.e. Singh 1989, Lunetta and Elvidge 1999, Coppin et al. 2004, Lu et al. 2004. Many of the algorithms used for analyzing temporal changes are indeed not restricted to change detection. In summary, there is a wide variety of techniques with varying degrees of robustness, flexibility and significance, and only a few studies provided a comparison of change detection methods for different applications (Mas 1999, Liu and Zhou 2004, Liu et al. 2004).

However, the results of the pixel-based approaches are often limited when used with very high spatial (VHR) imagery. Due to the increased information density of the VHR image data (Benz et al. 2004), too many changes are detected which may not be of interest for the particular application. This problem is also known as the "salt and pepper" effect when applying pixel-based classifiers to VHR imagery. Many contributions in this book demonstrate the advantages of object-based procedures for image classification. Also, with regard to change detection, object-based approaches could be more suitable than the pixel-based ones. Object-based procedures can employ change measures described above, and also exploit the changes of shape values, such as border length or size, or changes of object relations (see Figure 2). Hence, thematic, geometric and topologic changes of objects could be analyzed.

Some interesting object-based approaches have been introduced in the last years: McDermid et al. (2003) compared an object-oriented approach to change labelling with a more traditional pixel-based routine using the enhanced wetness difference index derived from a multi-temporal sequence of Landsat TM/ETM+ data. Hall and Hay (2003) presented an object-specific multiscale digital change detection approach that incorporates multi-temporal SPOT panchromatic data, object-specific analysis, object-specific up-scaling, marker-controlled watershed

segmentation and image differencing change detection. Desclée et al. (2006) combined the advantages of image segmentation, image differencing and stochastic analysis of the multispectral signal. Based on a chi-square test of hypothesis, abnormal values of reflectance differences statistics were identified and the corresponding objects were labelled as change. Niemeyer and Nussbaum (2006) used a combination of pixel- and object-based approaches by firstly pinpointing the significant change pixels by statistical change detection, object extraction and subsequently post-classifying the changes based on a semantic model of object features.

Im et al. (2007) introduced change detection based on object/neighbourhood correlation image analysis and image segmentation techniques. Gamanya et al. (2007) adopted a hierarchical image segmentation approach and applied a standardized, object-oriented automatic classification method.
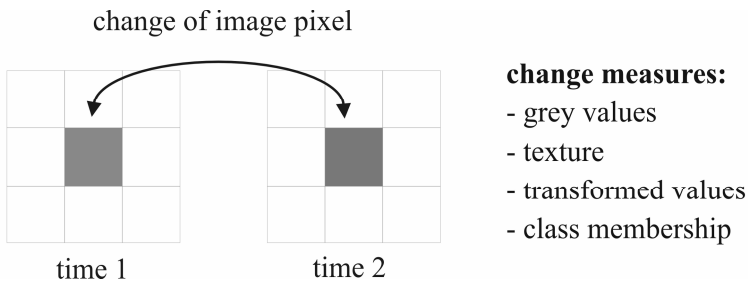

change of image pixel

change measures:
- grey values
- texture
- transformed values
- class membership

time 1   time 2

Fig. 1. Change measures used in pixel-based approaches

change of image object

Change measures:
- layer features
  (mean, stdev, ratio, texture,…)
- shape features
  (area, direction, position,…)
- relations among neighbouring,
  sub- and super-objects
- object class membership
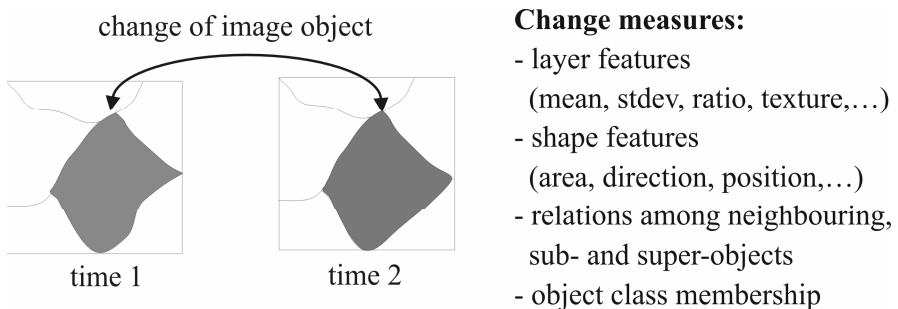
time 1   time 2

Fig. 2. Change measures used in object-based approaches

In this study, we propose an object-based unsupervised procedure for detecting and classifying infrastructure changes by VHR imagery. In the following part of this chapter the approach will be elaborated and illustrated by a case study on nuclear facilities.

2 Methodology

We propose an object-based unsupervised procedure for detecting and classifying infrastructure changes by VHR imagery. The proposed change analysis procedure consists of different steps: pre-processing, including pan-sharpening, geometric and radiometric correction, object extraction and feature extraction, statistical change/no-change detection, and a final change analysis by clustering changed objects. Figure 3 presents the workflow of the procedure that will be explained in the following.

A bi-temporal set of QuickBird images acquired over a nuclear facility was used for demonstration. QuickBird is a VHR commercial earth observation satellite, owned by DigitalGlobe and launched in 2001, which collects panchromatic (black & white) imagery at 60-70 centimeter resolution and multispectral (color) imagery at 2.4- and 2.8-meter resolutions.
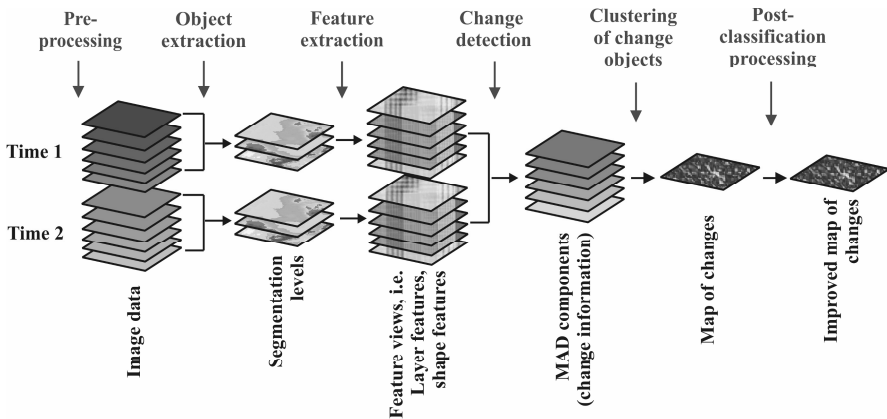

Fig. 3. Workflow of the change analysis procedure

2.1 Image pre-processing

This study mainly emphasized the detection and analysis of changed objects, but it is necessary to address which image preprocessing methods were utilized.

First of all, the two QuickBird scenes were pan-sharpened by wavelet transformation (Ranchin and Wald 2000). Pan-sharpening provides the fusion of the four color bands of the two QuickBird scenes with the respective panchromatic band. The pan-sharpened color bands with a spatial resolution of 0.6m do not involve additional information but have proved to be advantageous in the subsequent segmentation step compared to multispectral and panchromatic bands in their original spatial resolution.

Moreover, the aim of pre-processing is to correct temporal changes caused by variations in solar illumination, atmospheric conditions and sensor performance, and geometric distortion at the two image acquisition times. For geometric rectification, we performed the image-to-image registration by image correlation (Lehner 1986). Radiometric variations were finally reduced by means of radiometric normalization using no-change pixels (Canty et al. 2004).

For high-resolution imagery, the sensor's off-nadir viewing requires also an orthorectifying procedure to remove sensor and terrain-related distortions. Unfortunately no high-resolution digital elevation model was available within the study, thus orthorectification was not possible. The procedure continues with the extraction of objects primitives by segmentation.

2.2 Object and feature extraction

For image data taken at two different acquisition times, the image segmentation could be performed in three different ways, depending on the input data (Figure 4):

a) On the basis of the bi-temporal data set, i.e. using a data stack that consists of both scenes;
b) based on the image data of one acquisition time; the generated object boundaries are then simply assigned to the image data of the second acquisition time without segmentation;
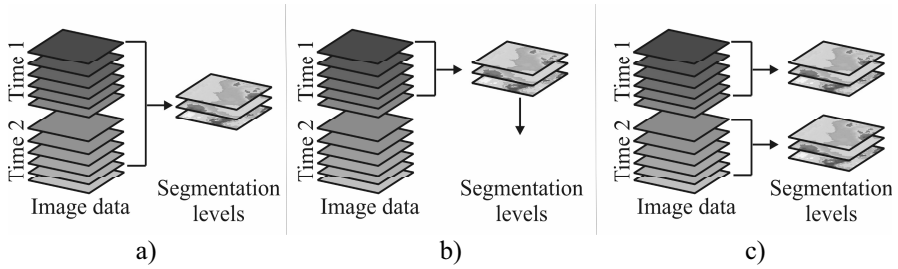c) separately for the two times, i.e. the two data sets are segmented independently.

Fig. 4. Object extraction of image data taken at two different acquisition times using different input data: **(a)** data stack consisting of both scenes, **(b)** image data of one time, **(c)** separately for the two times. Please see text for more information

When using a segmentation as suggested in (a) and (b), the produced image objects have the same geometric properties at the two times. With regard to change detection, the image objects have attribute values, which are either time-invariant, such as shape and size, or could vary in time, i.e. the layer values. Thus, changes of objects between the two image data sets could only be detected based on a limited number of the available object features.

In case (c) segmentations are independently performed for the two scenes. Consequently, the image object geometry will vary in time. In this case, all available object features could be used for object change detection.

An object-based change detection analysis requires a segmentation technique that similarly extracts objects that have not changed their shape and size between the two acquisition times. The multiresolution segmentation technique, implemented in Definiens Professional Version 5.0 (Baatz and Schaepe, 2000, Definiens, 2006), uses homogeneity criteria based on color and shape, and a scale parameter in combination with local and global optimization techniques. Thus, applying the same segmentation parameters to both scenes hardly produces similar objects in image regions with no or negligible changes (no-change objects), if other parts of the image have changed significantly. In fact, only for two identical images (i.e. a scene as a copy of the other one), the results of image segmentations would be the same. Discrepancies in the extraction of no-change objects due to overall alterations could probably be feasible. However, further methodological developments are needed with regard to the segmentation of multi-temporal objects. The variation of objects shape features is in any case an important indicator for real object changes and needs to be taken into account for object-based change detection.

In the study, object changes could be identified based on of the layer features, such as mean or standard deviation of the single image bands, shape and texture features. In an unsupervised approach presented here, class-related object features, such as the border of an image object shared with neighboring objects of a defined class or the existence of an image object assigned to a defined class in a certain perimeter around the image object concerned, were not considered. Also scene-related features, like the existence of a defined image layer, the number of image objects, or the minimum pixel value of a selected layer, were ignored so far. The methods of change detection and analysis will be outlined now.

2.3 Change detection and analysis

The Multivariate Alteration Detection (MAD) transformation (Nielsen 2007) was used for change detection. This method has been originally developed for change detection within the multispectral feature space of image pixels and was applied to this object-based change analysis study within the multidimensional and multivariate feature space of the objects. The MAD procedure is based on a classical statistical transformation referred to as canonical correlation analysis to enhance the change information in the difference images. The procedure is briefly described as follows:

If multispectral images of a scene acquired at times t_1 and t_2 are represented by random vectors X and Y, which are assumed to be multivariate normally distributed, the difference D between the two images is calculated by

$$D = \boldsymbol{a}^{\mathrm{T}} X - \boldsymbol{b}^{\mathrm{T}} Y \qquad (1)$$

In analogy to a principal component transformation, the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are sought subject to the condition that the variance of D is maximized and subject to the constraints that $\mathrm{var}(\boldsymbol{a}^T X) = \mathrm{var}(\boldsymbol{b}^T Y) = 1$.

As a consequence, the difference image D contains the maximum spread in its pixel intensities and, provided that this spread is due to real changes between t_1 and t_2, therefore maximum change information. Determining the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ in this way is a standard statistical procedure which considers a generalized Eigenvalue problem. For a given number of bands N, the procedure returns N Eigenvalues, N pairs of eigenvectors and N orthogonal (uncorrelated) difference images, referred to as to the MAD components. If the eigenvectors are ordered according to their decreasing Eigenvalues, the MAD components will be sorted in accordance with increasing variance.

The MAD components represent different categories of changes. Since relevant changes of man-made structures will generally be uncorrelated with seasonal vegetation changes or statistic image noise, they expectedly concentrate in different MAD components. If the components are sorted according to the increasing variance, higher order MAD components represent the small-scale changes whereas lower order components contain the overall or wide-area changes. Furthermore, the calculations of the MAD components are invariant under affine transformation of the original image data. For this reason, the MAD transformation can be qualified as a robust change detection technique.

The sum of squares of standardized variates

$$\chi^2 = \left(\frac{D_1}{\sigma_{D_1}}\right)^2 + \cdots + \left(\frac{D_N}{\sigma_{D_N}}\right)^2 \tag{2}$$

is approximately chi-square distributed with N degrees of freedom. Supposing that no-change pixels have a chi-square distribution with N degrees of freedom, N being the number of MAD components, the change-probability can be derived for each pixel or object.

In order to improve the spatial coherence of the change components, a maximum autocorrelation factor (MAF) transformation was applied to the MAD components (Nielsen et al. 1998). Assuming that image noise is estimated as difference between intensities of neighboring pixels, the MAF transformation is equivalent to a minimum noise fraction (MNF) transformation, which generates image components with maximum signal to noise ratio (Canty and Nielsen 2006).

Decision thresholds for change pixels could be set manually by standard deviations of the mean for each MAD or MAF/MAD components, i.e. by defining that all pixels in a MAD component with intensities within ±2σMAD are no-change pixels. Besides, the thresholds could be automatically estimated using a Bayesian procedure proposed by Bruzzone and Prieto (2000). This method is based on an Expectation-Maximization algorithm, which automatically determines probability density functions for change and no-change pixels given by the MAD or MAF/MAD components and iteratively calculates optimal decision thresholds for discriminating change and no-change pixels.

The MAD transformation based on N image layers results in N MAD components, as mentioned earlier. If 20 different feature layers are used as input data, the change detection has 20 MAD components as output. However, only change information included in three MAD or MAF/MAD components can be displayed at one time. Comprehensive visualization

and labeling of change objects becomes difficult. For this reason, a clustering procedure was subsequently applied to the change objects described in the next section.

2.4 Unsupervised classification of changes

Canty and Nielsen (2006) and Canty (2006) suggested to cluster the change pixels provided by the MAD components. Thus, any number of MAD or MAF/MAD components could be analyzed simultaneously, and the number of change categories could be fixed by the number of clusters. We adopted this approach in the study for subsclustering change objects given by the MAD components.

Canty and Nielsen (2006) recommended the use of the fuzzy maximum likelihood estimation (FMLE) introduced by Gath and Geva (1989), for clustering change pixels. The FMLE was also used in this study for classifying the change objects. Unlike conventional clustering algorithms, FMLE takes an advantages of forming also elongated clusters and clusters with widely varying memberships, which applies for change clusters in the MAD or MAF/MAD feature space. The fuzzy cluster membership of an object calculated in FMLE is the a-posteriori probability $p(C|f)$ of an object (change) class C, given the object feature f.

According to Richards and Jia (1999), misclassifications of an image pixel can be corrected by taking the spatial information into account. Canty and Nielsen (2006) adapted a method presented by Richards and Jia (1999) known as probabilistic label relaxation. This approach could also be applied to image objects. In the study, we evaluated the membership probabilities of an object with the following post-classification processing.

2.5 Post-classification processing

Using a neighborhood function $Q_m(k)$ for objects m, the technique examines for each object the membership probabilities of neighboring objects and corrects its membership probability

$$u'_{km} = u_{km} \frac{Q_m(k)}{u_{jm} \sum_{j=1}^{K} Q_m(k)}, \qquad k = 1 \ldots K \tag{3}$$

The modified membership probabilities can then be used for generating an improved classification. Moreover, analyzing the correlation between the MAD components and the original object features can help in the physically/spectrally founded labeling of the change clusters.

3 Case study

The proposed procedure will be illustrated and discussed for a bi-temporal, high-resolution satellite imagery. As an example we monitor a nuclear site in Iran. At the time when the images were taken some of the facilities of the site were still under construction, in operation or shut down.

Both Definiens Professional 5.0 for segmentation and feature extracting and ENVI4.3/6.3 including the ENVI extensions for pre-processing, change detection and classification[1], provided by Canty (2006) were used.

Two geo-referenced QuickBird scenes[2], acquired over the Esfahan Nuclear Technology Centre, Iran, at July 24, 2002 and July 9, 2003, were used in the study. The two image images contain a number of changes with different spatial and spectral dimensions, as the visual comparison of Figure 5 and 6 shows. Following section 2.1, a pan-sharpening method based on wavelet transformation was employed to create color bands with the spatial resolution of the higher resolution panchromatic band. Then, the 2003 scene was registered and radiometrically normalized to the 2002 scene using no-change pixels. Figure 5 and 6 show the pre-processed data for an image extract of 3500 pixels and 3000 lines, i.e. for an area of 2.1 km by 1.8 km.

Object extraction was conducted in Definiens Professional according to Figure 4c) by applying the same segmentation parameters to both scenes at three different scales. The image objects in the third segmentation level, generated with an identical scale parameter (200) and identical homogeneity criterions (color=0.5, compactness=1) were chosen for the subsequent change detection and classification steps. Figures 7 and 8 indicate 1229 objects for the 2002 scene and 1368 objects for the 2003 scene in the third level. The difference in the object's number could either be due to real object changes, the missing orthorectification or the quality of the segmentation procedure.

---

[1] http://www.fz-juelich.de/ste/remote sensing

[2] Both images were available in the context of the EU FP6 Global Monitoring for Security and Stability (GMOSS) Network of Excellence, 2004-2008. Credit: Digital Globe.

Since Definiens Professional neither supplies tools for the statistical change detection nor enables users to implement additional procedures in terms of computer programming, all object features needed to be exported and analyzed within the ENVI/IDL environment for change detection.

Fig. 5. Pan-sharpened QuickBird image acquired in July 2002 over the study site

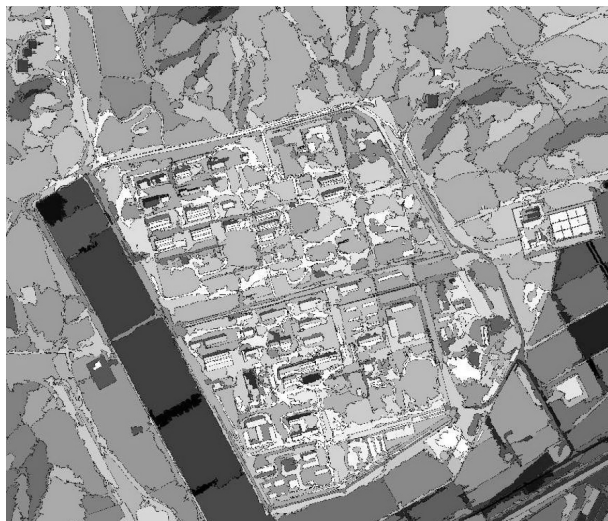**Fig. 6.** Pan-sharpened and radiometrically normalized QuickBird image acquired in July 2003 over the same site

Fig. 7. 1229 objects in segmentation level 3 for the 2002 scene

Fig. 8. 1368 objects in segmentation level 3 for 2003 scene

Two categories of object features were considered: layer features, in particular brightness and the mean values of the four multispectral bands, and shape features, i.e. border index, roundness, compactness, shape index

and length/width. Theoretically, other layer, shape or texture features could also have been included.

For detecting the object's changes, a MAD transformation was carried out based performed both on the layer and shape features, followed by a MAF transformation on the processed MAD components. The computed MAF/MAD components were ordered according to the increasing signal-to-noise-ratio.

Figures 9 and 10 show the MAF/MAD variates 3, 4 and 5 for the layer and the shape features with automatically determined thresholds, the medium grey displays the range between the lower and the upper threshold, i.e. no-change. Different categories of changes are represented by different colors. As expected, the shape features differ significantly due to the different segmentations results, as described in section 2.2.

Whether these changes are related to real changes, due to the missing orthorectification or caused by process of the segmentation itself, needs to be examined.

The visual comparison of Figure 6 and Figure 7 yields some indications: The changes indicated by one of the MAF/MAD components are mainly connected to segmentation variation of the sandy, non-vegetation surface.

The classification of the change objects was only performed for the layer features. A FMLE procedure was applied to the MAD components in order to group the change information in five clusters. A probabilistic label relaxation, described in section 2.5, was used to optimize the membership probabilities of the objects.

Final classification results are given in Figure 11. Classes 1 and 2 indicate no-change classes. Class 3 involves all changes related to high color variations, i.e. completion of the street, different shadow formation and others. Class 4 presents (among other changes) the roof covering of two buildings until 2003 in the upper part of the site, whereas class 5 highlights changes of the roof color of individual buildings due to different illumination conditions.
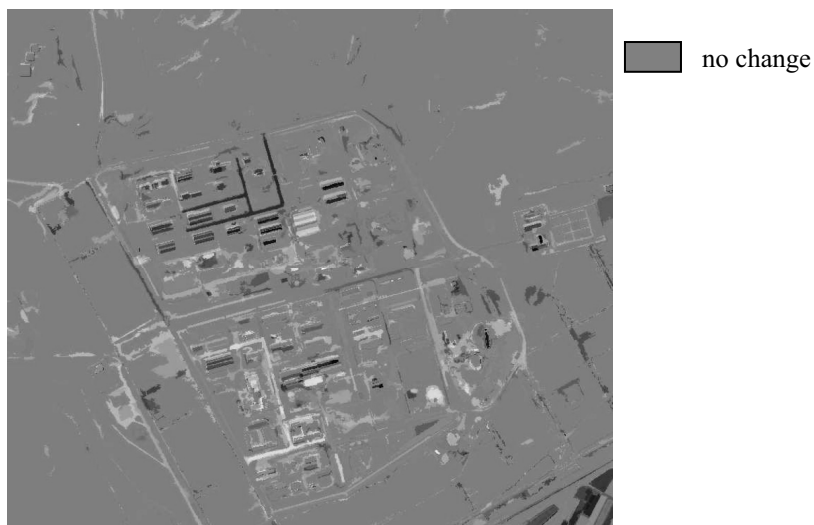
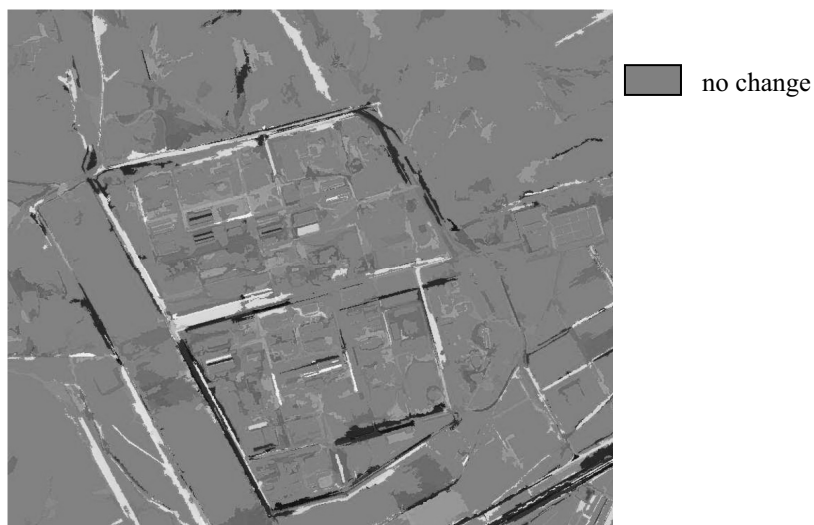**Fig. 9.** Changes of layer features, given in the MAF/MAD components 3, 4 and 5

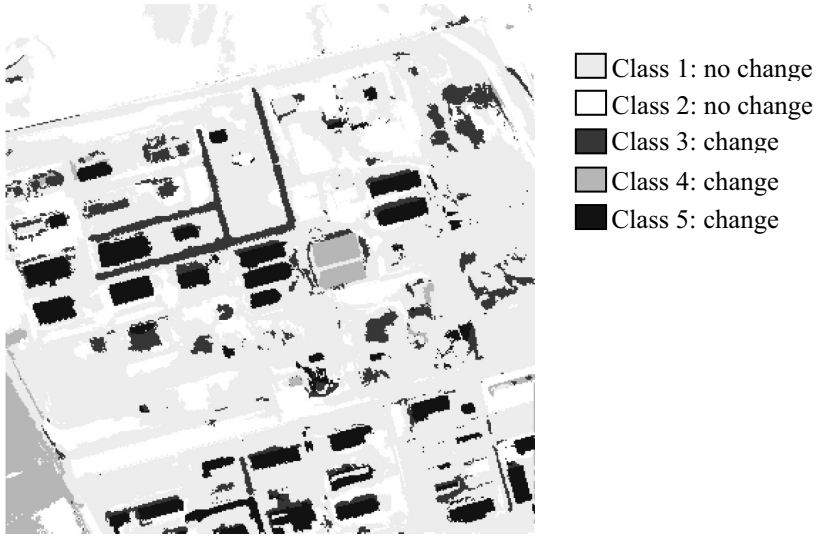**Fig. 10.** Changes of shape features, given in the MAF/MAD components 3,4 and 5

Class 1: no change
Class 2: no change
Class 3: change
Class 4: change
Class 5: change

Fig. 11. Final FMLE Clustering of the MAD change information on the layer features

4. Conclusions and future work

This chapter introduces an unsupervised approach to statistically analyze changes on the basis of various object features. Preliminary results show some promises. However, the procedure still needs notable improvement. Future work has particularly to meet the following challenges:

1. Assessing changes on the basis of shape features presumes a segmentation algorithm that is able to similarly extract no change objects, i.e. objects with unvarying shapes. The multiresolution segmentation provided with the Definiens software does not fulfill this condition, as it considers spectral information at least for the first segmentation level and applies global optimization techniques. Even if the initial random seed is being fixed for the two (or more) co-registered image data sets, the object boundaries will differ also for objects with constant shape.
2. The accuracy of the change analysis may be improved by applying the procedure only to the change pixels. As suggested by Canty and Nielsen (2006) the chi-square distribution calculated during the MAD

transformation could be used to estimate the change-probability for each pixel.

3. Analyzing (small-scale) three-dimensional objects, such as buildings, using high resolution imagery requires an accurate orthorectification of the data. Otherwise, different parallaxes due to varying off-nadir views will cause false alarms in the change detection process.

4. Finally, the inclusion of other object features and the integrated analysis needs to be investigated and evaluated.

References

Baatz M, Schaepe A (2000) Multiresolution segmentation - an optimization approach for high quality multi-scale image segmentation. In: Strobl J., Blaschke T, Griesebner G (ed), Angewandte Geographische Informations-Verarbeitung XII, Wichmann, Karlsruhe, pp 12–23

Benz U, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004): Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing 58, 239-258.

Bruzzone L, Prieto DF (2000) Automatic analysis of the difference image for unsupervised change detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 11(4), pp 1171–1182

Canty M, Nielsen AA (2006) Visualization and unsupervised classification of changes in multispectral satellite imagery. International Journal of Remote Sensing 27(18), pp 3961–3975

Canty MJ (2006) Image Analysis, Classification and Change Detection in Remote Sensing, with Algorithms for ENVI/IDL. Taylor, Francis, CRC Press, Boca Raton

Canty MJ, Nielsen AA, Schmidt M (2004) Automatic radiometric normalization of multispectral imagery. Remote Sensing of Environment 91, pp 441–451

Desclée B, Bogaert P, Defourny P (2006) Forest change detection by statistical object-based method. Remote Sensing of Environment 102, pp 1–11

Gamanya R, De Maeyer P, De Dapper M (2007) Object-oriented change detection for the city of Harare, Zimbabwe. Expert Systems with Applications (in print)

Gath I, Geva AB (1989) Unsupervised optimal fuzzy clustering. IEEE Transactions on Pattern Analysis and Machine Intelligence 3(3), pp 773–781

Hall O, Hay G.J. (2003) A Multiscale Object-Specific Approach to Digital Change Detection. International Journal of Applied Earth Observation and Geoinformation 4, pp 311–327

Im J, Jensen JR, Tullis, JA (2007) Object-based change detection using correlation image analysis and image segmentation. International Journal of Remote Sensing (in print)

Lehner M (1986) Triple stereoscopic imagery simulation and digital image correlation for MEOSS project. Proc. ISPRS Commision I Symposium, Stuttgart, pp 477–484

Liu H, Zhou Q (2004) Accuracy analysis of remote sensing change detection by rule-based rationality evaluation with postclassification comparison. International Journal of Remote Sensing 25(5), pp. 1037–1050

Liu Y, Nishiyama S, Yano T (2004) Analysis of four change detection algorithms in bi-temporal space with a case study. International Journal of Remote Sensing 25(11), pp 2121–2139

Lu D, Mausel P, Brondizio E, Moran E (2004) Change detection techniques. International Journal of Remote Sensing 25(12), pp 2365–2407

Lunetta RS, Elvidge CD (ed) (1999) Remote sensing change detection. Environmental monitoring methods and applications. Taylor, Francis, London

Mas JF (1999) Monitoring land-cover changes: a comparison of change detection techniques. International Journal of Remote Sensing 20(5), pp 139–152

McDermid GJ, Pape A, Chubey MS, Franklin SE (2003) Object-oriented analysis for change detection. Proceedings, 25th Canadian Symposium on Remote Sensing

Nielsen AA (2007) The Regularized Iteratively Reweighted MAD Method for Change Detection in Multi- and Hyperspectral Data. IEEE Transactions on Image Processing 16(2), pp 463–478

Nielsen AA, Conradsen K, Simpson JJ (1998) Multivariate alteration detection (MAD) and MAF processing in multispectral, bitemporal image data: New approaches to change detection studies. Remote Sensing of Environment 64, pp 1–19

Niemeyer I, Nussbaum S (2006) Change detection - the potential for nuclear safeguards. In: Avenhaus R, Kyriakopoulos N, Richard M, Stein G (ed), Verifying Treaty Compliance. Limiting Weapons of Mass Destruction and Monitoring Kyoto Protocol Provisions, Springer, Berlin, pp 335–348

Ranchin T, Wald L (2000) Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. Photogrammetric Engineering and Remote Sensing 66, pp 49–61

Richards JA, Jia X (1999) Remote Sensing Digital Image Analysis. An Introduction. 3 edition, Springer, Berlin

Singh A (1989) Digital change detection techniques using remotely-sensed data. International Journal of Remote Sensing 10(6), pp 989–1002

Chapter 2.6

Identifying benefits of pre-processing large area QuickBird imagery for object-based image analysis

T. Lübker, G. Schaab

Faculty of Geomatics, Karlsruhe University of Applied Sciences, Moltkestr. 30, D–76133 Karlsruhe, Germany, gertrud.schaab@hs-karlsruhe.de

ABSTRACT: Remote sensing and geoinformation systems (GIS) play an important role in biodiversity research and management, strengthened by the increasing availability of very high resolution (VHR) satellite imagery. The interdisciplinary research project BIOTA East Africa has acquired five QuickBird scenes of early 2005 covering Kakamega Forest (western Kenya) and surrounding farmland. Pre-processing steps applied to the imagery included the correction of atmospheric and orographic influences as well as a mosaicing procedure. This study demonstrates the benefits to object-based image analysis: after a thorough pre-processing, a) objects of the same kind have more consistent characteristics, b) segments do not orientate themselves to the join of adjacent image swaths. A rule-based multi-resolution classification scheme is presented for deriving landscape structures in the agricultural matrix surrounding the forest. The high complexity of the farmland, with extremely small structured parcels, demands analysis strategies that make use of image derivates obtained from cadastral maps as well as ground truth information. Finally, the classification result is compared to one based on the unprocessed imagery, emphasizing implications of pre-processing for object-based image analysis.

1 Introduction

Biodiversity conservation can not only be considered an essential part of ecology but "a common concern of humankind", as stated in the Convention on Biological Diversity (Secretariat of the Convention on Biological Diversity 2000 p 8), which was agreed on the Earth Summit in Rio de Janeiro in 1992. Remote sensing and geoinformation systems (GIS) play an important role in supporting the convention's aims with regard to the spatial context of biodiversity research and management (Schaab et al. 2002). For example, spatially explicit modelling of planning scenarios for the farmland can help to determine the potential for additional uses thus minimizing the pressure on forests (e.g. the collection of fire wood). Satellite imagery, especially such of very high resolution (VHR), offers a fundamental base for deriving the current situation of landscape patterns (e.g. Ivits et al. 2005).

Users usually obtain satellite imagery without much pre-processing applied by the data vendors. In order to improve image segmentation quality and to diminish differences in object characteristics, caused by illumination and atmospheric conditions, a thorough pre-processing should be applied. This is particularly needed if the imagery covering the study area is not taken in one overflight, but split between different swaths. Whilst pre-processing steps have been applied prior to classification in various studies (e.g. Lewinski 2006, Niemeyer et al. 2005, Hofmann 2001), their impacts on segmentation and classification seem not to have been investigated on their own.

For the interdisciplinary research project BIOTA East Africa, funded by the German Federal Ministry of Education and Research (BMBF), QuickBird satellite data were acquired by sub-project E02 in early 2005 (for information on E02 see e.g. Schaab et al. 2004). These offer four multispectral (MS) and one panchromatic band with a spatial resolution of 2.4 and 0.6 m, respectively. With a total area of 631 km² the imagery covers Kakamega Forest and surrounding farmland in western Kenya (Fig. 1). Due to the required size, five QuickBird scenes were recorded during two overflights on February 21st (eastern swath) and March 6th 2005 (western swath). Hence, varying geometrical recording conditions, different atmospheric conditions, as well as minor changes in vegetation cover have to be accepted.

Kakamega Forest is situated in the moist western part of Kenya (40 km north-east of Kisumu on Lake Victoria) and is known for its unique biodiversity (Mutangah et al. 1992). Gold exploitation in the 1930s and severe logging activities until the 1980s, however, has led to significant changes

in forest structure and diversity (Mitchell 2004). With a population density of about 600 persons/km² (Blackett 1994) the surrounding farmland, characterised by small-scale subsistence farming, is one of the most densely populated rural areas on earth. Jätzold and Schmidt (1982) state that farms are typically between 1 and 7 ha in size.

This work aims to derive farmland landscape elements through object-based image analysis of the acquired QuickBird imagery. This will help to provide a typology of the complex-structured farmland surrounding Kakamega Forest as a basis for land use planning. Here, we concentrate on the benefits of pre-processing (not including fusion methods or spectral transformations) for subsequent image analysis and a first classification strategy.

2 The pre-processing applied

QuickBird imagery obtained as 'standard' product is already processed with regard to radiometric, geometric and sensor corrections (DigitalGlobe 2004). However, certain additional pre-processing steps should be applied prior to image classification (see e.g. Schmidt 2003). Therefore, and in order to achieve better comparability between the scenes, as well as a homogeneous image mosaic, corrections of atmospheric and orographic influences were performed.

A correction of atmospheric influences is especially valuable if multi-temporal series are compared, e.g. for retrieving land use/cover changes, or if imagery from different sensors is compared (Richter 1996). When scenes from several swaths are used, as in the case of the available imagery, absolute calibration becomes less important than in the above stated applications, but still greatly enhances image quality and classification results. Within a single scene, corrections of orographic influences are of higher importance where reflection values of slopes being differently illuminated are adjusted. The resulting images were visually more appealing and pre-processing proved beneficial for mere visual interpretations.

'Atmospheric Correction' (ATCOR, see Richter 1996) presents the only available tool for an interactive atmospheric correction of VHR imagery (Neubert 2006). Implemented as version 3 in ERDAS Imagine it includes orographic correction. For counteracting atmospheric influences, the required input parameters are either taken from the recording conditions (e.g. sun and satellite angles, date, mean terrain height) or estimated by the implemented tool Spectra (atmosphere type, aerosol type, scene visibility). Orographic effects are modelled using a digital elevation model (DEM)

and its derivatives, slope and aspect. Set in relation to the sun's position they lead to local illumination angles. Additionally, a correction of effects described by the bidirectional reflectance distribution function is included (see Lübker 2005 for the settings of BRDF correction considering off-nadir and relative azimuth angles). Values for scene visibility were individually approximated for the two swaths in a test series. Although it may be questioned if real reflectance values could be obtained (Lübker 2005), it has been concluded that the atmospheric and orographic correction carried out with ATCOR 3 greatly enhanced the image quality and comparability between the two swaths (Lübker and Schaab submitted). Orographic effects in particular could be minimized (Fig. 2).

Adjacent image swaths were joined together in order to represent the large area under investigation as a coherent mosaic, i.e. without a distinctly visible join line. Whilst the mosaicing of scenes originating from the same swath was straight forward, difficulties were encountered when joining the two swaths. Due to the lack of high quality reference data, the georeferencing as performed and delivered by the vendor had to be accepted. This resulted in a mean horizontal aberration of approx. 5 m (or 8 pixels in the panchromatic band) for the two swaths relative to each other. For mosaicing, the geometrical adjustment of the two swaths was performed within a narrow band on either side of the join. Since the aberration was not linear, a procedure was especially elaborated for this purpose by making use of ERDAS Imagine 8.7, Leica Photogrammetry Suite, and Excel (for a detailed description of the procedure and the difficulties encountered see Lübker and Schaab submitted).

The final mosaic obtained was convincing. However, differences due to recording in two image swaths were still present. Here, differences that cannot be counteracted, like shadow lengths (see Lübker and Schaab 2006) and variations caused by different vegetation stages or differences in ground use, have to be distinguished from those where major enhancements have been made. However, differences in haziness as well as in reflectance behaviours caused by varying recording angles still exist. Also geometrical inaccuracies could not be improved due to the lack of adequate reference data. Nevertheless, the thorough pre-processing was worth the effort as shown in the next chapter.

3 Benefits of pre-processing

Two test areas were chosen to investigate the benefits of in-depth pre-processing of large-area VHR QuickBird imagery for object-based image

analyses. Using Definiens eCognition 4 software, improvements on segmentation and classification quality through a) the mosaicing in the stitching area of the two swaths and b) the correction of terrain shading are demonstrated. Effects of the atmospheric correction are not investigated on their own; they are included in the test results for the two study areas.

Test site A is situated north of Kakamega Forest towards Malava Forest (see Fig. 1) in the overlapping area of the two swaths. For comparison with the result of the in-depth pre-processing, the swaths were also mosaiced without any further treatment as well as with just a histogram matching applied. The latter method simulates thus a common methodology (e.g. Repaka et al. 2004). Fig. 3 shows the results of segmentation making use of the three different mosaicing approaches. In the simply stitched image (left) a large difference in spectral values between the eastern and the western part of the image can be observed. Here, segments strongly orientate themselves to the join which divides the subset. For the mosaic with an applied histogram matching (middle) the difference has become less apparent, but many segments still orientate themselves to the join. In the sophisticatedly processed mosaic (right) spectral differences have become inconspicuous. Segments only seldomly show an affinity towards the join. When looking at the spectral object characteristics (in the NIR-Red-Green composite segments are half-transparently coloured according to their mean spectral values) it becomes evident that a classification of the simply mosaiced image is very likely to lead to major errors.

Test site B is situated north-east of Kakamega Forest (see Fig. 1) covering parts of Kambiri Hill, a small edge of Kisere Forest and farmland including riverine vegetation and hedges. Fig. 2 visualizes the effects of orographic correction. In the original image (left) the western slope is noticeably shadowed while the eastern slope is brightened. In the processed image (right) the hill appears more uniform. However, it reveals that the two slopes of Kambiri Hill are not exactly equal in their vegetation cover. As shown in the processed image, the western slope is covered by denser shrub vegetation in the north and exhibits burning activities towards the south. The central part of the eastern slope is partly used for agriculture.

In order to evaluate the impact of the orographic correction on object-based image analysis, segmentation was accomplished for the atmospherically and orographically corrected example as well as the original image subset of test area B. For each side of the hill, image segments of similar vegetation (grassy land cover with sparse shrubs) were selected by means of visual interpretation. A comparison of their object properties reveals that the objects have very similar characteristics in the processed image. When drawn in two-dimensional feature space plots, the object's spectral

properties for mean Red, NIR and Soil Adjusted Vegetation Index (SAVI) overlap or group in close vicinity (Fig. 4, right). In the graph, black circles represent the vegetation on the eastern slope while the white circles represent vegetation on the western slope. In the original image the object characteristics of the selected vegetation type vary notably. Here, in the feature space plot, the object's spectral properties form two separate clusters according to the slopes (Fig. 4, left). Objects on the eastern slope (black circles) show higher values in the NIR band and slightly higher values in the Red band. This comparison points out that orographic influences are transferred from the imagery to the characteristics of image objects obtained through segmentation. In order to restrain the influences from affecting classification results, the original image should be orographically corrected.

4 Deriving landscape patterns in the agricultural matrix

As a first approach towards a farmland typology of the agricultural matrix surrounding Kakamega Forest, a classification scheme was elaborated employing a 1.6 km² test area (site C, see Fig. 1). This site is located near Isecheno forest station and covers farmland, forest, and the so-called 'tea belt' along the forest edge. In comparison to Jätzold and Schmidt (1982, see Introduction) the farms in this study area tend to be even smaller with a mean size of 0.8 ha, ranging from 0.1 to 4.7 ha, as delineated from cadastral maps. Farms are further subdivided into fields for growing tea, napier grass, sugar cane, mixed cultivations of maize and beans, bananas, cassava, sweet potatoes, other vegetables, and others, sometimes covering only a few hundred square meters. Parcel edges are particularly hard to distinguish when next to shrub and tree vegetation or when between fields cultivated with the same crop or of a similar preparation stage. Due to the complexity of the landscape, i.e. very different to the often industrialised farms in Western Europe and North America, comparable results to classifications as achieved e.g. for a Bavarian test site (Fockelmann 2001) are unlikely to be achieved.

As an additional input for segmentation cadastral maps in form of diazo copies (approximate scale: 1 : 2,500, dating from 1973/74 with in parts updates up to 2005) are available for major parts of test area C. The map sheets were scanned, georeferenced to the QuickBird imagery, offering the best positional accuracy available, and digitised adjusting the lines towards structures visible in the image. Subsequently a rasterisation was carried out and used as an input layer in the segmentation process to allow for deriva-

tions from the fictional sharp lines. In test area C, ancillary ground truthing information for the time of recording was collected by BIOTA-E14 project partners serving as evidence for land cover and subdivisions of parcels. In addition, serving as an image derivate (cf. Wezyk and de Kok 2005) an edge image was deduced from the panchromatic image using a Prewitt filter as well as the soil adjusted vegetation index (SAVI) from the multispectral image. Finally, a rule-based classification was carried out with Definiens Developer 6. Within the process tree three main cycles were developed applying multi-resolution and classification-based segmentation.

In the *first cycle* (segmentation level 3: scale factor 240, shape/compactness 0.7/0.7, only MS channels, no weighting) a rough separation into the basic land cover types of forest, agriculture, and tea was made (Fig. 5, top; urban settlement would be a fourth type but is not present in the test area). Classification was primarily based on mean values and standard deviations of the MS channels and the SAVI. In order to aggregate segments to large homogeneous regions further, rules were set up by means of the class-related features 'relative border to' and 'relative area of'. Conditions were then combined with logical terms. In this way tea plantations as well as larger tree lots on parcels could be assigned to the land cover type 'agriculture'. A smaller forest patch in the southern central part of the test area was undesirably classified 'agriculture'.

The *second cycle* (segmentation level 1: scale factor 70, shape/compactness 0.5/0.7, all available layers, adjusted weighting) aims at extracting the structure elements roads, houses/huts, trees/shrubs, and shadows of trees/shrubs within the main land cover type 'agriculture' (Fig. 5, middle; rivers would be a further element but are not present in the test area). Hence, segments classified as 'agriculture' in the first cycle were merged together and segmentation was only applied to this region. The rule set in this cycle is more complex, e.g. for the class 'houses/huts' three groups of membership functions are combined in the logical term 'and (min)': besides the generic shape feature 'length/width' (< 2.3) two further shape features, 'area' (< 290 m²) and 'border length' (< 90 m) combined with 'mean (arithm.)', as well as spectral characteristics of mean MS 4, mean SAVI, and standard derivation of MS 1 combined with 'and (min)' build the class description. While trees/shrubs and their shadows could be classified to full satisfaction within these preliminary tests, further refinement of segmentation quality and classification rules would be needed for detecting also narrow lanes and for distinguishing single huts and groups of houses/huts more clearly.

In the *third cycle* (segmentation level 2: scale factor 90, shape/compactness 0.75/0.85, all available layers, adjusted weighting) the remaining agricultural land not classified in cycle two and representing the actual

fields was further investigated concerning a distinction of its cover. During classification-based segmentation high values for the shape factor and compactness were chosen in compliance with the field structure. In layer weighting, the cadastral boundary layer and the SAVI image play an important role. Rule sets are based on spectral characteristics, namely mean values for MS 3, MS 4 and SAVI. In the classification only few distinctions in land cover/use could be made: tea, fields recently prepared for planting, fields at an intermediate stage or with young plants, and fields harvested and not yet prepared for planting (Fig. 5, bottom; sugar cane would be a further land cover but is not present in the test area). This can be ascribed to the fact that image recording had fallen within the period shortly before the start of the rainy season, usually around the end of March (Jätzold and Schmidt 1982) with the planting period being due to start (see Fig. 6). As can be seen when comparing the imagery of the two swaths in the overlapping area, soil digging activities took place within the two weeks that fall between the recordings, due to rainfall after a period of three weeks without any rain. A recording at a more advanced planting stage would have been advantageous for identifying real agricultural land use. But then higher humidity, or actual clouds, would certainly have resulted in poorer image quality.

In order to assess the influence of pre-processing on the above described object-based classification, the classification was additionally carried out on the imagery that was not pre-processed, apart from histogram matching to the pre-processed imagery. Border values for most of the membership functions had to be adjusted manually using the feature view tool so that similar objects to those in the original classification could be selected.

A comparison of the two classification results (Fig. 7) shows that the first cycle provides very similar results, even if the tea zone extends into the forest. But the more detailed classification contains large differences. Trees and shrubs as well as their shadows are underestimated in the western part and overestimated in the eastern part of the classification when not having undergone thorough pre-processing. This can be ascribed to the large difference in brightness of the MS channels. Roads are detected less well throughout the test area while houses and huts could be detected almost as precisely. Also the field/parcel classification differs significantly. In the unprocessed image segments are clearly orientated according to the join line.

The complex picture of structural elements making up the agricultural landscape next to Kakamega Forest, although not the agricultural land use per field, could be delineated by the method described (based on the pre-processed imagery). A further refinement of the approach is nevertheless needed. The retrieval of all the current hedges and fences along field

boundaries will serve as background for estimates on the potential for additional cultivations when also making use of field boundaries. In this context, it would be desirable to also extract information on land ownership. Due to the complex structure of the farmland, however, it seems that the grouping of fields per owner directly from the imagery is difficult to achieve, if not impossible. Here, knowledge obtained from areas with cadastral maps available in conjunction with a house/hut classification will hopefully enable an extrapolation to areas where cadastral maps are not at hand.

5 Summary and outlook

It was demonstrated that thorough pre-processing of large-area VHR satellite imagery is beneficial not only for more appealing visual results, but also for subsequent object-based image analyses. Corrections of influences caused by the orography and atmospheric conditions should be applied in order for objects of the same kind to have consistent characteristics. Elaborative mosaicing methods prevent segments from orientating themselves to imagery join lines and are thus preferable to the methods commonly applied.

The described classification scheme shows promising results as a first test in the development of a farmland typology for the Kakamega area. Multi-resolution segmentation based on classification results allows the user during the object-based classification to focus on the parts of the image relevant to the task, i.e. the farmland. Although complex membership rules have been set up for the agricultural fields, only a few classes could be separated due to the season and thus the planting stage at which the imagery was recorded. Overall, segmentation and classification was complicated by the intricate and small-structured composition of the agricultural matrix.

For a further refinement of the approach, the role of the available cadastral boundaries for segmentation needs to be examined. The potential of the edge detecting Canny filter (Canny 1986) should be tested with regards to field boundary delineation from the panchromatic imagery. An enhancement concerning segment boundaries is hoped for via a field-based classification as proposed by Türker and Kok (2006). Tests will show if this approach is beneficial for such complex structural elements.

As soon as the current structure of the agricultural landscape surrounding Kakamega Forest is obtained through object-based image analysis, scenarios of rural livelihood as demanded by Diwani and Becker (2005)

will be modelled in close cooperation with other BIOTA East sub-projects. Agronomists and land management planners in collaboration with biologists will provide landscape planning recommendations considering socioeconomic impacts on land use and landscape elements affecting biodiversity. The scenarios will be spatially explicit and simulate possible future landscapes (alternative futures, see Baker et al. 2004). The results will form an important contribution in support of biodiversity research in Eastern Africa.

Note

All satellite imagery used in the figures: © 2005 by Digital Globe TM, USA (distributed by Eurimage, Italy).

References

Baker J, Hulse D, Gregory S, White D, Van Sickle J, Berger P, Dole D, Schumaker N (2004) Alternative futures for the Willamette River Basis, Oregon. Ecol Appl 14(2):313-324

Blackett H (1994) Forest inventory report no 3: Kakamega. Kenya Indigenous Forest Conservation Programme, Nairobi, Kenya

Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8(6):679-698

Digital Globe (2004) QuickBird imagery products: Product guide (rev 4.3). Digital Globe, Longmont, CO

Diwani T, Becker M (2005) Characterization and classification of agricultural land use systems in Kakamega, Kenya: Implications on soil fertility, productivity and biodiversity. In: Proc of Tropentag 2005, The Global Food & Product Chain: Dynamics, Innovations, Conflicts, Strategies, 11-13 October 2005, Stuttgart, Germany, p 223

Fockelmann R (2001) Agricultural parcel detection with eCognition 2.0. Definiens eCognition Application Note 2(10), pp 1-2

Hofmann P (2001) Detecting informal settlements from Ikonos image data using methods of object oriented image analysis – an example from Cape Town (South Africa). In: Proc of 2nd International Symposium on Remote Sensing of Urban Areas, 22-23 June 2001, Regensburg, Germany, pp 107-118

Ivits E, Koch B, Blaschke T, Jochum M, Adler P (2005) Landscape structure assessment with image grey-values and object-based classification at three spatial resolutions. Int J Rem Sens 26(14):2975-2993

Jätzold R, Schmidt H (1982) Farm management handbook of Kenya: Vol II/a West Kenya, Nyanza and Western Provinces. Kenya Ministry of Agriculture, Nairobi, Kenya

Lewinski S (2006) Applying fused multispectral and panchromatic data of Landsat ETM+ to object oriented classification. In: Proc of 26th EARSeL Symposium New Developments and Challenges in Remote Sensing, 29 May-2 June 2006, Warsaw, Poland, pp 1-7

Lübker T (2005) Auswertung von QuickBird-Satellitenbilddaten im Gebiet Kakamega Forest (Westkenia) mittels Methoden der multispektralen Bildverarbeitung sowie der objektorientierten Segmentierung. Unpublished diploma thesis at Faculty of Geoinformation, Karlsruhe University of Applied Sciences, Germany

Lübker T, Schaab G (2006) Large area QuickBird imagery for object-based image analysis in western Kenya: Pre-processing difficulties, workarounds and resulting benefits as well as (first) segmentation strategies. In: Proc of 1st International Conference on Object-based Image Analysis (OBIA 2006): Bridging Remote Sensing and GIS, Int Arch Photogram Rem Sens Spatial Inform Sci XXXVI (4/C42), on CD, 4-5 July 2006, Salzburg, Austria

Lübker T, Schaab G (submitted) Prozessierung von 631 km^2 QuickBird-Satellitenbilddaten für das Gebiet Kakamega Forest (Westkenia): Erfahrungen mit Atmosphärenkorrektur, Mosaikbildung und Pan-sharpening. In: Photogrammetrie, Fernerkundung, Geoinformation

Mitchell N (2004) The exploitation and disturbance history of Kakamega Forest, western Kenya. Bielefelder Ökologische Beiträge 20, BIOTA East Report 1

Mutangah J, Mwangangi O, Mwaura J (1992) Kakamega Forest: A vegetation survey report. Kenya Indigenous Forest Conservation Programme, Nairobi, Kenya

Neubert M (2006) Bewertung, Verarbeitung und segmentbasierte Auswertung sehr hoch auflösender Satellitenbilddaten vor dem Hintergrund landschaftsplanerischer und landschaftsökologischer Anwendungen. Fernerkundung und Angewandte Geoinformatik 1, Rhombos, Berlin, Germany

Niemeyer I, Nussbaum S, Canty M (2005) Automation of change detection procedures for nuclear safeguards-related monitoring purposes. In: Proc IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2005), 25-29 July 2005, Seoul, South Korea, pp 2133-2136

Repaka S, Truax D, Kolstad E, O'Hara C (2004) Comparing spectral and object-based approaches for classification and transportation feature extraction from high resolution multispectral imagery. In: Proc Am Soc Photogram Annu Meet, 23-28 May 2004, Denver, CO

Richter R (1996) A spatially-adaptive fast atmospheric correction algorithm. Int J Rem Sens, 17(6):1201-1214

Schaab G, Kraus T, Strunz G (2004) GIS and remote sensing activities as an integrating link within the Biota-East Africa project. In: Proc of Sustainable use and conservation of biological diversity: a challenge for society, 1-4 December 2003, Berlin, Germany, pp 161-168

Schaab G, Hörsch B, Strunz G (2002) GIS und Fernerkundung für die Biodiversitätsforschung. In: Proc of 19. DFD-Nutzerseminar, 15-16 October 2002, Oberpfaffenhofen, Germany, pp 159-170

Schmidt M (2003) Development of a fuzzy expert system for detailed land cover mapping in the Dra catchment (Morocco) using high resolution satellite images. PhD thesis, University of Bonn, Germany, URN: http://nbn-resolving.de/urn:nbn:de:hbz:5n-02097 (accessed 2.4.2007)

Secretariat of the Convention on Biological Diversity (2000) Sustaining life on earth: How the convention on biological diversity promotes nature and human well-being. Montreal, Canada

Türker M, Kok E (2006) Developing an integrated system for the extraction of sub-fields within agricultural parcels from remote sensing images. In: Proc of 1st International Conference on Object-based Image Analysis (OBIA 2006): Bridging Remote Sensing and GIS, Int Arch Photogram Rem Sens Spatial Inform Sci XXXVI (4/C42), on CD, 4-5 July 2006, Salzburg, Austria

Wezyk P, de Kok R (2005) Automatic mapping of the dynamics of forest succession on abandoned parcels in south Poland. In: Proc of AGIT-Symposium 2005, Angewandte Geographische Informationsverarbeitung XVII, 6-8 July 2005, Salzburg, Austria, pp 774-779
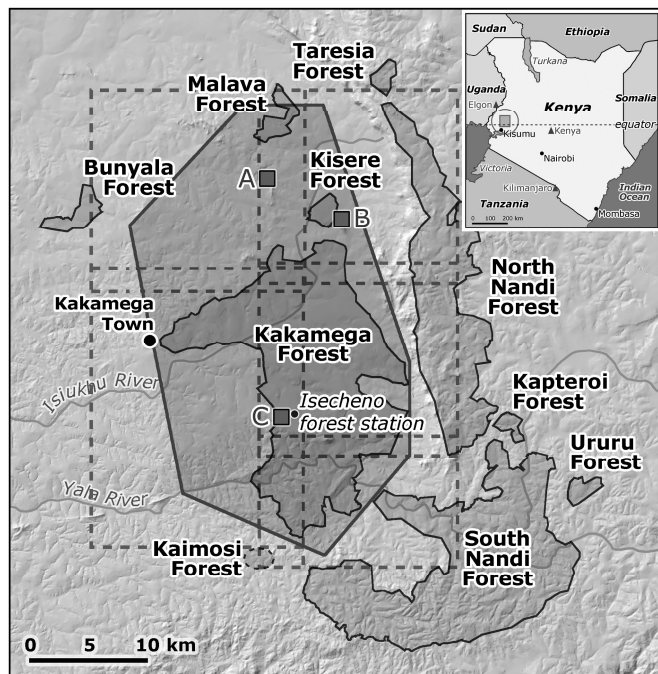
Fig. 1. Location of Kakamega Forest and its associated forest fragments with coverage of the acquired QuickBird scenes and location of the three test sites as referred to in this paper
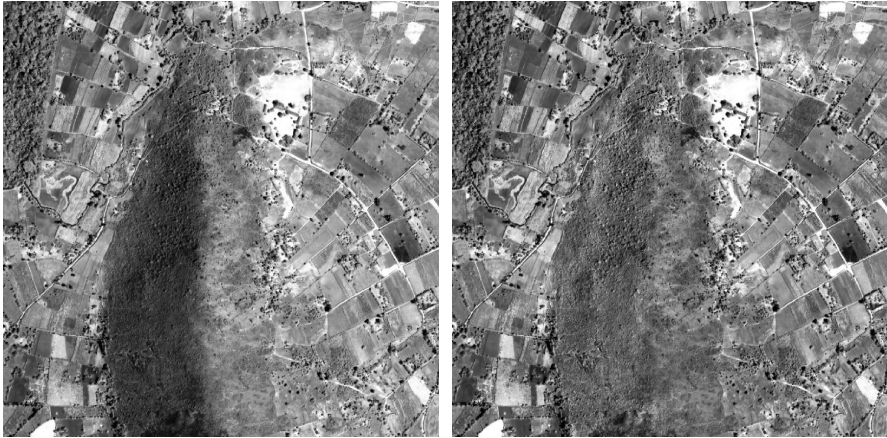
Fig. 2. Effects of orographic correction at Kambiri Hill (test site B): Original image (left) as compared to the processed image (right). Images are pan-sharpened and displayed as NIR-Red-Green composites (same to Fig 3 in Lübker and Schaab 2006)
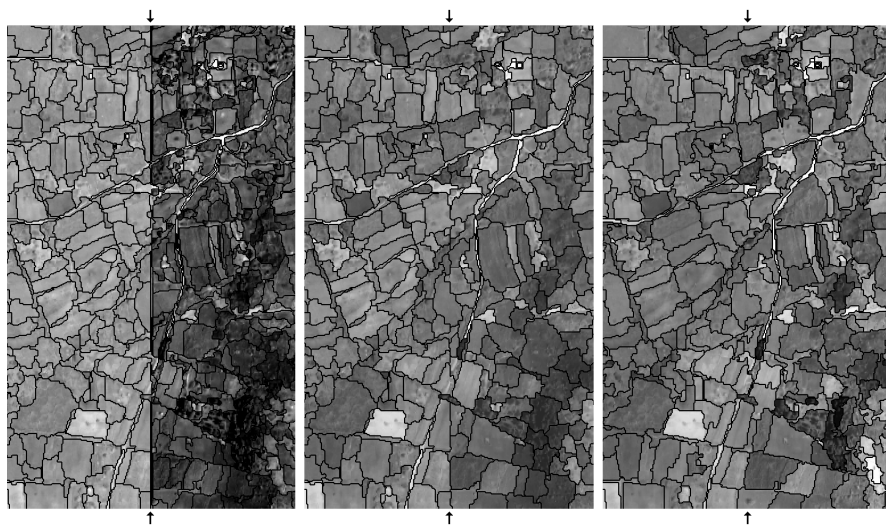
Fig. 3. Influence of different mosaicing approaches on segmentation results (test site A): simply stitched image (left) mosaic with just a histogram matching applied (middle), and mosaic based on a thorough pre-processing (right). Segments are half-transparently coloured according to their mean spectral values in the NIR-Red-Green composite
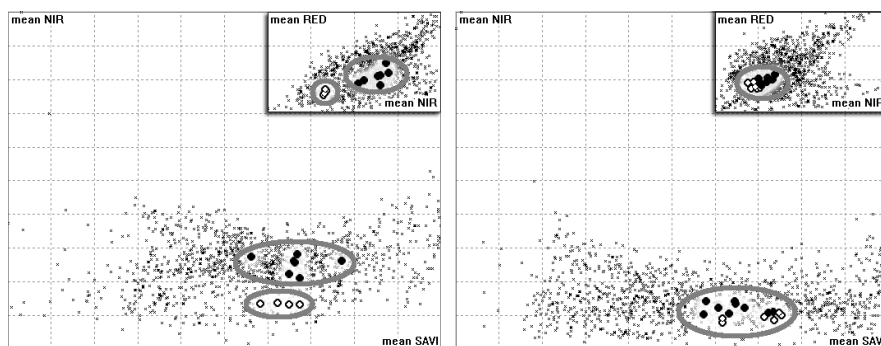
Fig. 4. Two-dimensional feature space plots demonstrating object characteristics for similar vegetation (grassy land cover with sparse shrubs) on the western slope (white circles) and eastern slope (black circles) of Kambiri Hill. Graphs relate to test site B before (left) and after pre-processing (right)
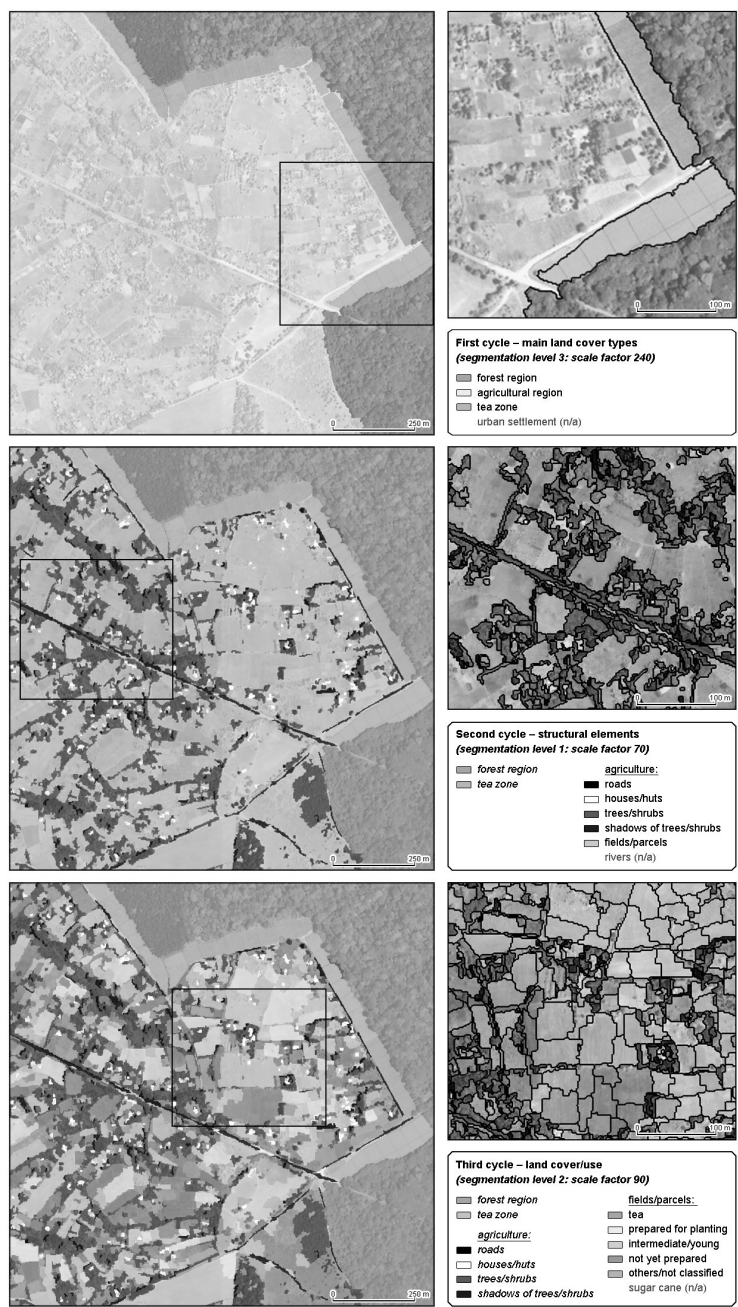
Fig. 5. Classification results of a rule-based multi-resolution classification for test site C near Isecheno forest station
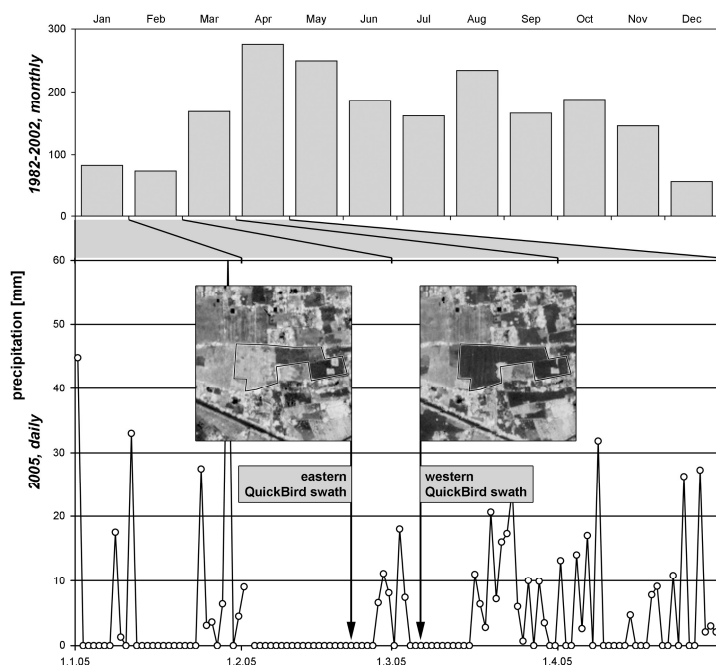
Fig. 6. Monthly (1982–2002, top) and daily (January–April 2005, bottom) precipitation [mm] at Isecheno forest station (compiled by E11, E02) in comparison to the two recording dates (21st of February, 6th of March) explaining differences in soil reflectances (SAVI)

Fig. 7: Classification results with (left, see Fig. 5 bottom) and without (right) pre-processing applied to the imagery. For a legend see Fig. 5

Chapter 2.7

A hybrid texture-based and region-based multi-scale image segmentation algorithm

A. Tzotsos, C. Iosifidis, D. Argialas

Laboratory of Remote Sensing, Department of Surveying, School of Rural and Surveying Engineering, National Technical University of Athens, Greece

ABSTRACT: The objective of this research was the design and development of a region-based multi-scale segmentation algorithm with the integration of complex texture features, in order to provide a low level processing tool for object-oriented image analysis. The implemented algorithm is called *Texture-based MSEG* and can be described as a region merging procedure. The first object representation is the single pixel of the image. Through iterative pair-wise object fusions, which are made at several iterations, called *passes*, the final segmentation is achieved. The criterion for object merging is a homogeneity cost measure, defined as object heterogeneity, and computed based on spectral and shape features for each possible object merge. An integration of texture features to the region merging segmentation procedure was implemented through an Advanced Texture Heuristics module. Towards this texture-enhanced segmentation method, complex statistical measures of texture had to be computed based on objects, however, and not on rectangular image regions. The approach was to compute grey level co-occurrence matrices for each image object and then to compute object-based statistical features. The Advanced Texture Heuristics module, integrated new heuristics in the decision for object merging, involving similarity measures of adjacent image objects, based on the computed texture features. The algorithm was implemented in C++ and was tested on remotely sensed images of different sensors, resolutions and complexity levels. The results were satisfactory since the produced primi-

tive objects, were comparable to those of other segmentation algorithms. A comparison between the simple algorithm and the texture-based algorithm results showed that in addition to spectral and shape features, texture features did provide good segmentation results.

1 Introduction

1.1 Recent developments in Remote Sensing

Recently, remote sensing has achieved great progress both in sensor resolution and image analysis algorithms. Due to very high resolution imagery, such as IKONOS and Quick Bird, traditional classification methods, have become less effective given the magnitude of heterogeneity appearing in the spectral feature space of such imagery. The spectral heterogeneity of imaging data has increased rapidly, and the traditional methods tend to produce "salt and pepper" classification results. Such problems occur also to medium resolution satellite data, such as Landsat TM, SPOT etc.

Another disadvantage of traditional classification methods is that they do not use information related to shape, site and spatial relation (context) of the objects of the scene. Context information is a key element to photo-interpretation, and a key feature used by all photo-interpreters because it encapsulates expert knowledge about the image objects (Argialas and Harlow 1990). Such knowledge is not explicit and needs to be represented and used for image analysis purposes. Shape and texture features are used extensively as context descriptors in photo-interpretation.

1.2 Texture-based Image Segmentation and Object-based Image Analysis

Texture initially was computed as standard deviation and variance. Haralick proposed texture features computed from co-occurrence matrices (Haralick et al. 1973, Haralick 1979). These second order texture features were used in image classification of remote sensing imagery with good results (Materka and Strzelecki 1998). Recently, even more complex texture models were used for classification and segmentation, such as Hidden Markov Models, Wavelets and Gabor filters (Materka and Strzelecki 1998) with very good results in remote sensing and medical applications. Several methods were proposed for texture-based image segmentation, taking advantage of the latest texture modeling methods (Liapis et al. 1998, Havlicek and Tay 2001, Chen et al. 2002, Fauzi and Lewis 2003). At the same

time, image classification moved towards artificial intelligence methods (Sukissian et al. 1994, Benz et al. 2004).

During the last few years, a new approach, called Object-Oriented Image Analysis, integrated low level image analysis methods, such as segmentation procedures and algorithms (Baatz & Schäpe 2000), with high level methods, such as Artificial Intelligence (fuzzy knowledge-based systems) and Pattern Recognition methods. Within this approach, the low level image analysis produces primitive image objects, while the high level processing classifies these primitives into meaningful domain objects (Benz et al. 2004).

In order to extract primitive objects from a digital image, a segmentation algorithm can be applied. Various segmentation algorithms and methods have been proposed over the last decades, with promising results (Pal and Pal 1993, Sonka et al. 1998). In remote sensing, a multi-scale image segmentation algorithm is aiming not to the extraction of semantic objects, but to the extraction of image primitives (Baatz & Schäpe 2000).

1.3 Research Objectives

The main objective of this research was the integration of complex texture features into an object-oriented image segmentation algorithm (Tzotsos and Argialas 2006) to be used as a low level processing part of an object-oriented image analysis system so that to be applied at multiple image resolutions and to produce objects of multiple scales (sizes), according to user-customizable parameters.

Another objective was the ability of the produced algorithm to be generic and produce satisfying and classification-ready results to as many remote sensing data as possible. Remote sensing data with complex texture and spectral information are, in general, difficult to process. Therefore, there was a need for the algorithm to be able to handle texture information and context features in order to produce better segmentation results.

This research aimed to further develop recent technologies and provide integration of new features to object-based image analysis methodology.

Motivation for this research is to provide an Object-Based Image Analysis system in the form of Free and Open-Source Software.

2 Methodology

2.1 MSEG algorithm overview

The MSEG algorithm (Tzotsos and Argialas 2006) was designed to be a region merging technique, since region merging techniques are fast, generic and can be fully automated (without the need of seed points) (Pal and Pal 1993, Sonka et al. 1998). Given that commercial Object-Oriented Image Analysis system (eCognition User Guide 2005) has used such methods was also a strong argument for the effectiveness of the region merging techniques.

Like all algorithms of this kind, MSEG is based on several local or global criteria and heuristics, in order to merge objects in an iterative procedure, until no other merges can occur (Sonka et al. 1998). In most cases, a feature of some kind (mean spectral values, texture, entropy, mean square errors, shape indices etc.) or combination of such features computes the overall "energy" of each object. Then, the merging algorithm uses heuristics and "trial and error" methods in order to minimize the overall energy of the segmentation that is produced. In other words, it is typical to select a cost function, to define how good and stable an object is after a merging procedure, or even to make the decision regarding that merge.

Various definitions of homogeneity (energy minimization measures within an object) have been defined (Pal and Pal 1993, Sonka et al. 1998). Recently, a very successful segmentation algorithm, embedded in the Object Oriented Image Analysis Software eCognition (Baatz & Schäpe 2000), implemented such measures of spectral and spatial homogeneity, for making the merging decision between neighboring objects, with very good results. In the proposed segmentation algorithm, similar homogeneity measures were used, and then complex texture features were implemented in later stages.

The proposed algorithm was initialized through the application of an image partitioning method to the dataset resulting into rectangular regions of variable dimensions, called *macroblocks*. Image partitioning was applied for computation of local statistics and starting points. It should be pointed that starting points were not used as seed points (as in region growing techniques) but were used to keep track of the order in which all pixels were processed initially. Having this order computed and stored for each macroblock, the whole segmentation algorithm can be reproduced accurately and provide 100% identical results for the same parameters and image.

There are two basic methods implemented within MSEG for starting point estimation (Tzotsos and Argialas 2006). The first is a statistical

method, producing one starting point per macroblock. The second method was based on dithering algorithms, transforming the original image into a binary image through statistical procedures (Ulichney 1987). The binary image was used to determine the starting points within each macroblock.

In order for the MSEG algorithm to provide primitive objects, several steps of region merging (*passes*) were followed. The initial objects of the image are the single pixels. The purpose of a first segmentation pass was to initialize the image objects and to provide the first over-segmentation, in order for the algorithm to be able to begin region merging at following stages. During first pass, the algorithm merged single pixels-objects pair wise. The criterion for object merging was a homogeneity cost measure, defined as object heterogeneity, and computed based on spectral and shape features for each possible object merge. The heterogeneity was then compared to a user defined threshold, called scale parameter, in order for the decision of the merge to be determined.

For the following pass of the algorithm, the objects created by the previous pass were used in a new pair wise merging procedure. The merging strategy included finding the best match for each object, and then checking if there was a mutual best match in order to merge the two objects (Tzotsos and Argialas 2006). Passes were executed iteratively until the algorithm converged. The algorithm was considered finished, when during the last pass no more merges occurred. Then, the objects were exported and marked as final primitives.

In order to extend the basic elements of the region merging segmentation procedure, a multi-scale algorithm was designed to give to the MSEG algorithm the capability to create multiple instances of segmentations for an image, each with different scale parameters. The problem when dealing with multiple segmentations is the compatibility between scales, in order to combine information and objects. One simple way to deal with this problem is to create a multi-level representation, and incorporate the multiple segmentations within this representation, hierarchically. A single-level hierarchy is not flexible, when dealing with remote sensing classification problems (Argialas and Tzotsos 2004). A multi-level hierarchy approach or a branch-based hierarchy model can represent more complex spatial relations. Thus, in the present multi-scale algorithm, every new level depends only from the nearest (scale-wise) super-level or the nearest sub-level, or both. More details on MSEG implementation can be found in (Tzotsos and Argialas 2006).

2.2 Advanced Texture Heuristics

The basic objective of the Advanced Texture Heuristic module was to build upon MSEG algorithm, in order to improve segmentation results. Since texture is a key photo-interpretation element, it was decided to use more complex texture features, than standard deviation (used in eCognition), variance, or other first order texture features.

Since second order texture features have been used as good and practical classification features (Haralick et al. 1973, Materka and Strzelecki 1998), there was a need to test those measures for segmentation purposes and specifically as an add-on to the region merging multi-scale algorithm that was previously developed.

The basic idea was that when making a merging decision between adjacent image objects, there should be a texture similarity measure, provided by complex texture computations that could help this decision. This way, primitive objects with similar texture could be merged, even if color or shape criteria are not in favor of this merge.

Given that MSEG is a region merging algorithm, not all state of the art methods for modeling texture are compatible for a hybrid segmentation solution. The recent literature has shown that Markov Random Fields, wavelets and Gabor filters, have great potential for texture analysis (Materka and Strzelecki 1998). Their disadvantage is that they are very complex and time consuming to use with a merging procedure, computing thousands of virtual merges during a full object merging search. At the same time, wavelets and Gabor filters are computationally inefficient to be used locally, within the boundaries of a single – and sometimes very small - primitive object. Markov Random Fields are easier to adopt for region-based texture segmentation, but they were found incompatible with the current merging search method, since they are based on Bayesian reasoning.

A traditional method for modeling texture, which was proved to be very good for practical purposes in supervised classification (Haralick et al. 1973, Schroder and Dimai 1998), is based on the Grey Level Co-occurrence Matrix (GLCM) features. GLCM is a two dimensional histogram of grey levels for a pair of pixels that are separated by a fix spatial relationship. The Grey Level Co-occurrence Matrix approximates the joint probability distribution of this pair of pixels. This is an insufficient approximation for small windows and a large number of grey levels. Therefore the image data have to be pre-scaled to reduce the number of grey levels in the image. Directional invariance can be obtained by summing over pairs of pixels with different orientations (Schroder and Dimai 1998).

From the GLCM, several texture measures can be obtained, such as homogeneity, entropy, angular second moment, variance, contrast etc (Haralick et al. 1973). To compute the GLCM, several optimization methods have been introduced. Most applications of GLCM for remote sensing images, at pixel-level, included computation of the co-occurrence matrix less often for the whole image, and more often for a predefined image sliding window of fixed size.

On a pixel-based texture analysis with the use of GLCM, for each direction (0, 45, 90, 135 degrees) a different co-occurrence matrix is formed by adding co-occurrences to the grey level pair position (Figure 1). If Ng is the number of grey levels after the grey level reduction, then each co-occurrence matrix will be of size Ng x Ng.
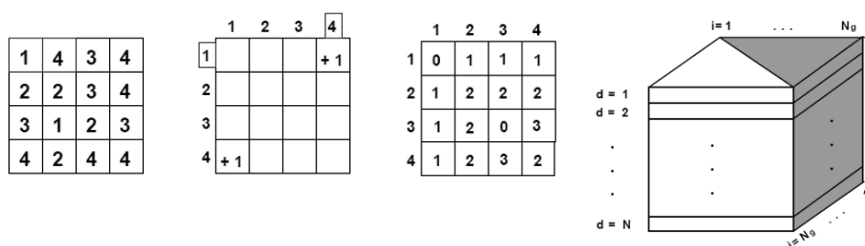
Fig. 1. Left: An example of GLCM computation at 0 degree angle for a 4x4 window. The empty GLCM gets filled by adding co-occurrences symmetrically. Right: A 3-dimensional representation of the Co-occurrence matrices that have to be computed for a given orientation. Ng is the number of grey levels and N is the total number of primitive image objects

In order to use the second order texture features into MSEG, four GLCMs should be computed for each primitive image object, during the merging procedure. The computation of so many GLCMs can be extremely intensive, and would significantly slow down the performance of the region merging algorithm. Furthermore, as multiple virtual merges occur, before an object merge could be decided, the GLCM computations can actually be much more than theoretically expected. Thus, a decision was made to optimize the use of GLCM features only for the initial objects and not for the virtual merged objects, so that to limit the co-occurrence matrix computation to a maximum of 4N (where N is the number of objects).

The Advanced Texture Heuristic module aimed to implement texture similarity measures in order to contribute to object merging search. Haralick states as good texture similarity measures, the Homogeneity

(Equation 1) and the Angular Second Moment (Equation 2) features (Haralick 1979). Both were implemented in the module.

$$Homogeneity = \sum_{i,j=0}^{N_g-1} \frac{GLCM_{i,j}}{1+(i-j)^2} \tag{1}$$

$$ASM = \sum_{i,j=0}^{N_g-1} GLCM_{i,j}{}^2 \tag{2}$$

When an object was treated from the MSEG algorithm during a *pass*, the texture features were computed and the mutual best match search procedure compared neighbor objects to the selected one. Before the color and shape heterogeneity criteria were computed and involved to the scale parameter comparison, texture heterogeneity was computed, as the difference of the values of the texture similarity features. These values, one for each direction, were then compared with a threshold called *texture parameter* which is defined by the user. If the two objects were compatible by the texture parameter, then the computation of the spectral and shape heterogeneity took place, in order to fulfill the mutual best match criterion, and the merge to occur.

The described heuristic, uses the texture parameter, to reduce the heterogeneity computations. This means that, when activated, the Advanced Texture Heuristic module has greater priority than the scale parameter, but cannot perform any merging, without color and shape compatibility of image objects. If one wishes to perform segmentation using only texture features, the scale parameter can be set to a very large value, so not to constrain the merging by the color and shape criteria.

In the following section, an optimization procedure for the GLCM computation is described.

2.3 Implementation

Having to compute thousands of co-occurrence matrices, during a region merging segmentation procedure can be computationally intense. Optimization algorithms for the computation of GLCMs have been proposed in literature (Argenti et al. 1990) but only for the pixel-based case. In order to tackle this problem, the GLCM computation should be optimized to be used with objects, rather than pixels. A modification to the traditional methods was performed, so that to make the procedure faster but still accurate.

At first, image band selection took place. If the computation of the GLCM was to be performed for each band separately, the whole segmentation process would not be optimal for performance. So, instead of using all bands, the Advanced Texture Heuristic module can use the intensity band of the HSI colorspace, or the Y band of the YCbCr colorspace (used as default), or a principal component band of the image, or finally a single image band. After the band selection, a grey level reduction was performed at the selected band. The final number of grey levels can be selected by the user, with a quantizer parameter. The default value, as used by many other GLCM implementations, was set to 32 grey levels.

It was determined that the optimal procedure to compute the GLCMs was to perform some kind of global initialization, so that to speed up the inter-object GLCM computation. For each of the image pixels, a direction search was performed, to evaluate the grey level pair co-occurrences. For the 4 different directions, a vector was designed to hold the overall co-occurrence information and was used in a way similar to a database index. Thus, no direction search was performed twice during the pass stages. Each time an object co-occurrence matrix had to be used, it was computed very fast within the object boundaries.

This procedure was not tested for algorithmic complexity, but was compared to a simple GLCM implementation and was found more stable and faster. The implementation of the Advanced Texture Heuristic module was performed in C++. The modified version of the algorithm was called *Texture-based MSEG*.

3 Discussion of Results

The implemented version of the MSEG algorithm was tested on a variety of image data, in order to assess the quality of the results, its generalization and speed. Evaluating the results of a segmentation algorithm does not depend on the delivery of semantic objects, but rather on the generation of good object primitives useful to further classification steps.

The algorithm was designed to provide (a) over-segmentation so that merging of segments, towards the final image semantics, to be achieved by a follow up classification procedure (Tzotsos and Argialas 2007) and (b) boundary distinction and full-scene segmentation. Since eCognition (eCognition User Guide 2005) is greatly used for object oriented image analysis purposes, the evaluation of results was mainly based on comparison with outputs from eCognition. Also, a comparison was made to results of

simple MSEG, to show how the texture features perform with region merging segmentation.

For the evaluation of the algorithms a Landsat TM image was used. For all tests, the color criterion was used with a weight of 0.7 and the shape criterion with weight 0.3. The eCognition software was used to provide segmentation with scale parameter 10. Then, the simple MSEG was used to provide segmentation with scale parameter 400 (through trial and error) to simulate the mean object size of eCognition's results. It should be noted that scale parameters are implementation dependent. The results are shown in Figure 2. In Figure 3 the results from the *texture-based MSEG* for scale parameter 400 are shown.
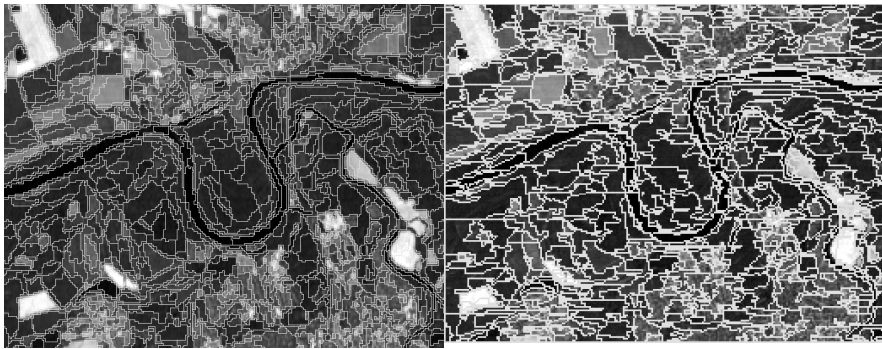
Fig. 2. Segmentation result as provided by eCognition for scale parameter 10 (left) and MSEG for scale parameter 400 (right)

Comparing results between Figures 2 and 3, shows that similar sized object can be obtained by all 3 segmentation algorithms. For the scale parameter of 400, the MSEG seems to be more sensitive to spectral heterogeneity than the eCognition results with scale parameter 10. Both algorithms keep good alignment with the image edges and both provide usable over-segmentations of the initial image (Tzotsos and Argialas 2007). The *texture-based MSEG* also provides good segmentation of the image, improving the simple MSEG result by creating texturally homogenous regions, but at the same time, working against the shape criterion, providing less compact or smooth boundaries for objects.

Further to its better vectorization, eCognition has better shaped boundaries, which is a clue that the shape weights of this algorithm are valued unequally or that additional embedded heuristics are involved. In both systems, MSEG (with or without texture) and eCognition, the thematic category boundaries are well respected by the segmentations.
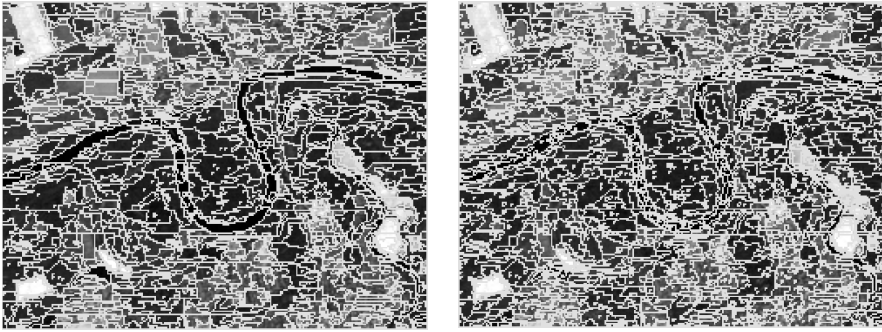
Fig. 3. Segmentation result as provided by *texture-based MSEG* for scale parameter 400 and texture parameter 2.0 (left) and for scale parameter 400 and texture parameter 1.0 (right)

In a further step of evaluation (Figure 4), the result of eCognition for the scale value of 20 is comparable to the result provided by the texture-based MSEG when a very large (2500) scale parameter was used (so that the scale parameter would not significantly interfere with the final mean object size) and the texture parameter was set to 3.0. The *texture-based MSEG* result is very good especially inside the urban areas, where there are complex texture patterns. There, it merged the object primitives in such a way, so that to provide larger homogenous objects in comparison to eCognition or the simple MSEG.
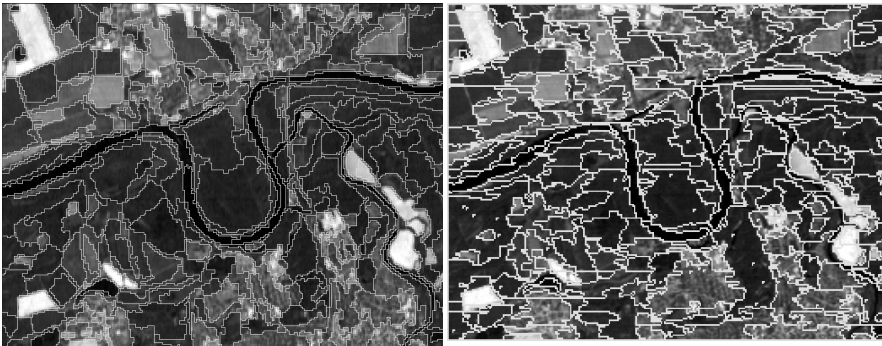
Fig. 4. Segmentation result by eCognition for scale parameter 20 (left) and texture-based MSEG for scale parameter 2500 and texture parameter 3.0 (right)

Also, in Figure 3, the difference of segmentation results is shown, when only the texture parameter is changed. Smaller texture parameter provides

smaller primitive objects. It should be noticed that texture free objects, like the very bright white areas in the image, don't get affected by the texture parameter change, as expected.

A final step of evaluation included testing the segmentation algorithms with a very high resolution remotely sensed image. A digital multispectral image from an aerial scanner (Toposys GmbH) with resolution of 0.5m was used. This image was selected because outperforms in resolution all commercial satellite data available today.

In Figure 5, the results of the eCognition algorithm with scale parameters 15 and 25 are presented. These results are very good, especially across the road outlines. The areas of interest in this test were those with complex texture, e.g. the mixed grasslands. The results of eCognition for these areas are the most over-segmented.

Fig. 5. Segmentation result by eCognition for scale parameters 15 (left) and 25 (right)

In Figure 6, the simple MSEG results are shown for scale parameters 400 and 700. As before, the results are good especially across the building and road edges. Scale parameter 400 provides very over-segmented result and seems to be the most sensitive to color heterogeneity. For scale parameter 700 the result is better, especially in the complex texture areas, but eCognition has better results on road edges. It can be concluded after extensive testing of both algorithms, that the color and shape criteria are not optimized in the same way, thus many differences occur in the shapes of primitive objects.
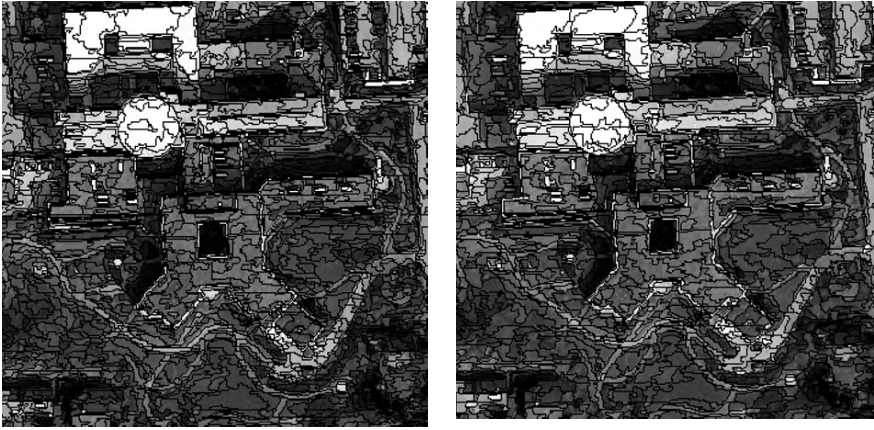
Fig. 6. Segmentation result by simple MSEG for scale parameters 400 (left) and 700 (right)
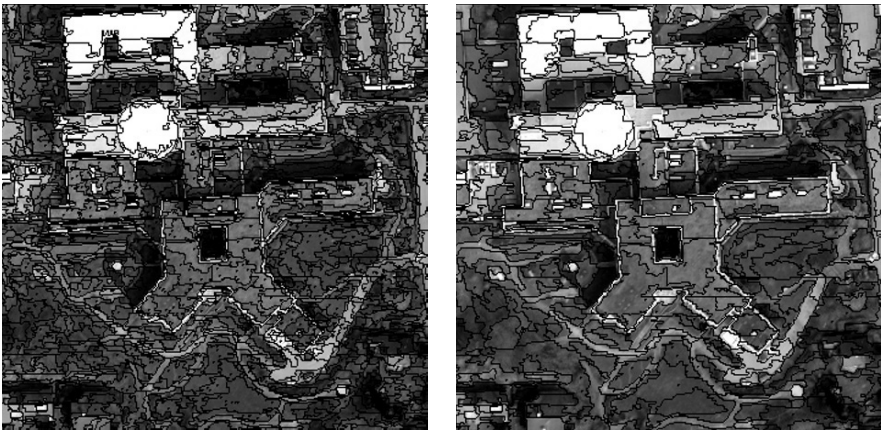

Fig. 7. Segmentation result by *Texture-based MSEG* for scale parameters 400 (left) and 2500 (right)

In Figure 7, the Texture-based MSEG results are shown for scale parameters of 400 and 2500. For the scale parameter 400 test, a strict texture parameter of 1.0 was used, so that to demonstrate the capability of the algorithm to locate even the smallest differences in texture. This can be observed in the left-center area of the image, where there is a mixture of bare land with grassland. Thus a small texture parameter provides significant over-segmentation of the region. For the scale parameter 2500 test, this high value was used in order to provide freedom to the algorithm to reach

to a result that was based only to texture parameter (3.0). The results were better than those of eCognition and of the simple MSEG in the complex textured areas. Again, the eCognition segmentation algorithm outperforms MSEG on the road edges, but results are very close.

Table 1. Segmentation Parameters

	Algorithm	Scale Parameter	Color	Compact-ness	Smooth-ness	Texture Parameter
Fig.2	eCognition	10	0.7	0.6	0.4	-
	MSEG	400	0.7	0.6	0.4	-
Fig.3	Tex.MSEG	400	0.7	0.6	0.4	2.0
	Tex.MSEG	400	0.7	0.6	0.4	1.0
Fig.4	eCognition	20	0.7	0.6	0.4	-
	Tex.MSEG	2500	0.7	0.6	0.4	3.0
Fig.5	eCognition	15	0.7	0.6	0.4	-
	eCognition	25	0.7	0.6	0.4	-
Fig.6	MSEG	400	0.7	0.6	0.4	-
	MSEG	700	0.7	0.6	0.4	-
Fig.7	Tex.MSEG	400	0.7	0.6	0.4	1.0
	Tex.MSEG	2500	0.7	0.6	0.4	3.0

4 Conclusions and future work

Overall, the designed image segmentation algorithm, gave very promising segmentation results for remote sensing imagery. With the addition of the Advanced Texture Heuristic module, it was shown to be a good and generic segmentation solution for remote sensing imagery. The boundaries of the primitive objects extracted in each case were compatible with those of the semantic objects. Thus, for object oriented image analysis, the *texture-based MSEG* is qualified as a successful low level processing algorithm.

MSEG has however some disadvantages that have to be further investigated. Its shape heterogeneity criteria are not quite effective and must be further tested and optimized to provide better results.

Future developments for the MSEG include integration of the algorithm with higher level artificial intelligence and pattern recognition methods for classification. Part of this integration with Support Vector Machines is presented in (Tzotsos and Argialas 2007).

Acknowledgements

References

Argenti F., Alparone L., Benelli G. (1990). Fast Algorithms for Texture Analysis Using Co-Occurrence Matrices. IEEE Proceedings, 137, F, 6, 1990, 443-448.

Argialas D., and Harlow C. (1990). Computational Image Interpretation Models: An Overview and a Perspective. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No 6, June, pp. 871-886.

Baatz M. and Schäpe A. (2000). Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. *In: Strobl, J. et al. (eds.): Angewandte Geographische Infor-mationsverarbeitung XII. Wichmann, Heidelberg, pp. 12-23.*

Benz U., Hoffman P., Willhauck G., Lingenfelder I., Heynen M., (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing 58* pp. 239-258.

Chen, J., Pappas, T.N., Mojsilovic, A., Rogowitz, B. (2002). Adaptive image segmentation based on color and texture. *Proceedings International Conference on Image Processing. Evanston, IL, USA, 777- 780 vol.3*

eCognition User Guide (2005), Definiens, Munchen. http://www.definiens-imaging.com

Fauzi M. FA. and Lewis, P. H. (2003). A Fully Unsupervised Texture Segmentation Algorithm. Harvey, R. and Bangham, J. A., Eds. *Proceedings British Machine Vision Conference 2003*, pages 519-528.

Haralick R.M., Shanmugan K., Dinstein I. (1973). Textural features for image classification. IEEE Trans. On Systems, Man and Cybernetics, 3(6):610-621, Nov. 1973.

Haralick R.M (1979). Statistical and Structural Approaches to Texture. Proceedings of the IEEE, Vol. 67, No. 5, May 1979, pp. 786-804.

Havlicek, J. P. and Tay, P.C. (2001). Determination of the number of texture segments using wavelets. *Electronic Journal of Differential Equations*, Conf. pp 61–70.

Liapis, S., Alvertos, N., Tziritas, G. (1998). Unsupervised Texture Segmentation using Discrete Wavelet Frames. *IX European Signal Processing Conference*, Sept. 1998, pp. 2529-2532

Materka A., Strzelecki M. (1998). Texture Analysis Methods – A Review, Technical University of Lodz, Institute of Electronics, COST B11 report, Brussels 1998

Pal, N. R. and Pal, S. K. (1993). A review on image segmentation techniques. *Pattern Recognition*, vol. 26, pp. 1277-1294.

Sonka, M., Hlavac, V. Boyle, R., (1998). *Image Processing, Analysis, and Machine Vision* - 2nd Edition, PWS, Pacific Grove, CA, 800 p., ISBN 0-534-95393-X.

Schroder M., Dimai A. (1998). Texture Information in Remote Sensing Images: A Case Study. Workshop on Texture Analysis, WTA 1998, Freiburg, Germany.

Sukissian L., Kollias S., Boutalis Y. (1994). Adaptive Classification of Textured Images Using Linear Prediction and Neural Networks. Signal Processing, 36, 1994, 209-232.

Tzotsos A. and Argialas D. (2006). MSEG: A generic region-based multi-scale image segmentation algorithm for remote sensing imagery. In: Proceedings of ASPRS 2006 Annual Conference, Reno, Nevada; May 1-5, 2006.

Tzotsos A. and Argialas D. (2007). Support Vector Machine Classification for Object-Based Image Analysis. In: Object-Based Image Analysis – Spatial Concepts For Knowledge-Driven Remote Sensing Applications. Springer 2007.

Ulichney, R. (1987). *Digital Halftoning*. The MIT Press, Cambridge, MA.

Chapter 2.8

Semi-automated forest stand delineation using wavelet based segmentation of very high resolution optical imagery

F.M.B. Van Coillie, L.P.C. Verbeke, R.R. De Wulf

Laboratory of Forest Management and Spatial Information Techniques, Ghent University, Belgium, frieke.vancoillie@ugent.be

ABSTRACT: Stand delineation is one of the cornerstones of forest inventory mapping and a key element to spatial aspects in forest management decision making. Stands are forest management units with similarity in attributes such as species composition, density, closure, height and age. Stand boundaries are traditionally estimated through subjective visual air photo interpretation. In this paper, an automatic stand delineation method is presented integrating wavelet analysis into the image segmentation process. The new method was developed using simulated forest stands and was subsequently applied to real imagery: scanned aerial photographs of a forest site in Belgium and ADS40 aerial digital data of an olive grove site in Les Beaux de Provence, France. The presented method was qualitatively and quantitatively compared with traditional spectral based segmentation, by assessing its ability to support the creation of pure forest stands and to improve classification performance. A parcel/stand purity index was developed to evaluate stand purity and the expected mapping accuracy was estimated by defining a potential mapping accuracy measure. Results showed that wavelet based image segmentation outperformed traditional segmentation. Multi-level wavelet analysis proved to be a valuable tool for characterizing local variability in image texture and therefore allowed for the discrimination between stands. In addition, the proposed evaluation measures were found appropriate as segmentation evaluation criteria.

1 Introduction

Forest stands are the basic units of management and are generally defined as spatially continuous units of uniform species composition, stem density, crown closure, height and age (Leckie et al. 2003). Correct tree species identification for example is essential for forest management and in applications such as species-specific growth models. The calculation of timber volumes is also usually species specific. Traditionally stand boundaries have been estimated through air photo interpretation. Visual interpretation however is subjective and can be ameliorated by numerical interpretation through automated image processing (Haara and Haarala 2002, Wulder et al. 2007). New mapping techniques are subject of research in terms of improved speed, consistency, accuracy, level of detail and overall effectiveness (Leckie et al. 2003). Several techniques have been developed but most of them are designed for automated tree isolation e.g. Gougeon (1995a,b), Culvenor (2002), Larsen (1997) and Warner et al. (1998). Subsequent stand delineation based on individually outlined trees is less developed but has been extensively studied by Leckie et al. (2003). Another possibility is the automatic delineation of stands based on image segmentation. Hay et al. (2005) introduced MOSS (Multiscale Object-specific Segmentation) and showed that it can be used to automatically delineate a range of objects from individual tree crowns to forest stands.

This chapter presents a method aiming at forest stand[1] delineation by wavelet based image segmentation. In its development stage, the method uses artificially simulated images representing forest stands that differ in species composition, stem density, crown closure, height and age class. The proposed method addresses tree stand delineation and not the extraction of other stand attributes. The method is applied to ADS40 digital aerial photographs with a spatial resolution of 50cm and to scanned color-infrared aerial photographs with a resolution of 1m.

2 Artificial imagery

The use of simulated remotely sensed imagery derived from artificially generated tree stands offers several advantages. First, artificially generated stands serve as initial test cases with complete freedom of design in terms

---

[1] Throughout this paper, the term stand will be used to identify an ensemble of trees, independent of the stand type. A stand thus represents as much a (semi) natural forest stand as a homogeneous area covered by olive or fruit trees.

of species composition, stem density, crown closure, height and age. Secondly, the need for validation data is alleviated since stand attributes are known in advance. Thirdly, there is total control over illumination conditions, viewing geometry and spatial resolution. Finally, the resulting images are not disturbed by atmospheric or sensor noise. As a consequence, keeping spatial resolution, illumination conditions and viewing geometry constant, errors in stand delineation are exclusively due to algorithm performance.

The generation of artificial remote sensing data is a two step process. First, a three dimensional artificial forest is generated. Afterwards this 3D model is illuminated using a ray-tracing model, and the sensor response is determined. The basic units of artificial forests or stands are artificial trees. Trees are characterized by a number of randomly distributed variables. Table 1 and Fig. 2.1 provide an overview of the random variables that control artificial tree generation. Crown projections are modeled as composed quarters of ellipses, and the crown itself is constructed by rotating an arbitrary function connecting the crown projection envelop with the tree top around the central axis of the tree.

2.1 Generation of artificial stands

Before stand generation starts, a number of artificial tree types are defined by choosing specific values for the parameters of the random distributions and the crown shapes shown in Table 1 and Fig. 2.1. A half-ellipsoid (dilated sphere) was used to model a generic broadleaf tree type, and a semicone was applied to model a generic conifer tree type. Additionally, it is necessary to define the allowed overlap or intersection between trees. In this study, the admitted overlap is a fixed fraction of the area of intersection between two crown projections. This parameter controls crown cover and, together with the crown radii, determines stand density (number of trees/ha).

Table 1. Overview of the random variables and the associated probability distributions that control artificial tree generation

Variable	Distribution	Symbol in Fig. 2.1
Stem height	normal	H_s
Crown height	normal	H_c
Stem diameter	normal	D_s
Crown shape	half sphere / semi-conical	
Crown radii	normal	d1, d2, d3, d4
Stem position	uniform	(x,y)
Orientation	uniform	a
Spectral characteristics	normal	
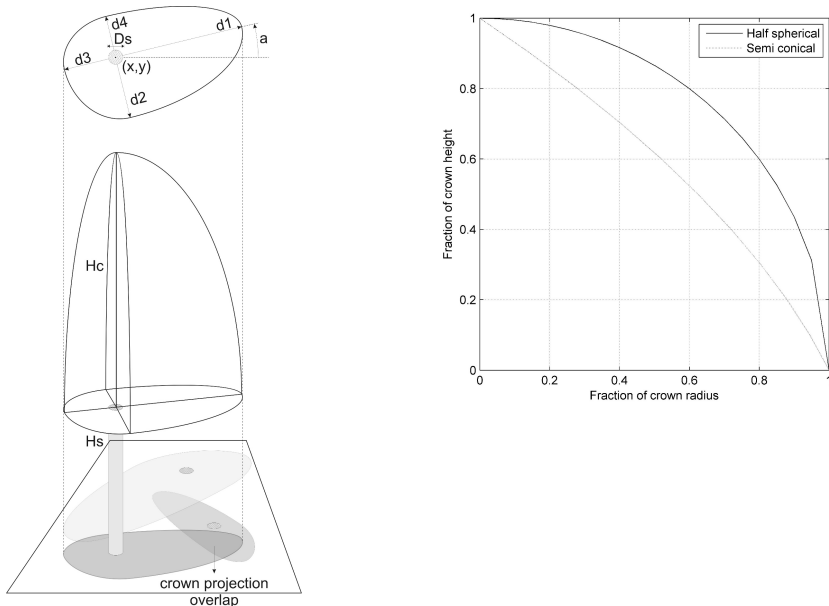
Fig. 2.1 Right: crown shapes used to model artificial tree types. Left: the artificial tree model

Once the artificial tree types are defined, a random grid is generated. Every cell in the grid is assigned a random value, corresponding to one of the generated artificial tree types. This grid is used to spatially constrain the uniformly distributed stem position of the trees. Stand/forest generation starts by randomly selecting an artificial tree type. A random candidate tree is then drawn from the probability distributions defined by the random

tree type (Table 1). If the position of the random tree is within one of the grid cells corresponding to the selected tree type, it is checked whether the candidate tree satisfies the overlap constraint by calculating the area of crown projection overlap with every tree already in the forest. If no overlap constraints are violated, the tree is added to the artificial stand/forest. This process continues until a fixed number of consecutive candidate trees fails to meet the positional or overlap constraints.

2.2 Illumination of artificial stands

Once an artificial forest is generated, it needs to be converted to a simulated remotely sensed image. This is achieved by using a naïve ray-tracing method. The sun's position is set to its position on June 15, at 11h30 (azimuth=145°, elevation=59°) in Brussels, Belgium. The sensor's viewing angle is fixed at 20.5°. The image is then processed using the POV-Ray Persistence of Vision Pty. Ltd. (2004) software. The spectral characteristics are chosen in such a way that the rendered RGB images correspond to respectively the near-infrared, red and green band of an IKONOS image. Spectral characteristics are derived from ex-situ measurements of *Pinus nigra.* Even though the generated scenes are assumed to contain both broadleaves and conifers, all trees are assigned the *Pinus nigra* spectral profile. This way, artificial tree types can only be distinguished based on structural characteristics.

3 Wavelets transforms

Wavelets have been used in a variety of remote sensing applications ranging from image fusion (Park and Kang 2004), over noise and speckle reduction (Sgrenzaroli et al. 2004), data compression (Zeng and Cumming 2001), sub-pixel mapping/sharpening (Mertens et al. 2004) to the analysis of image texture (Dekker 2003, Li 2004, Kandaswamy et al. 2005). Wavelet transforms reconstruct or decompose signals (i.e. images) by using a superposition of translated, dilated and scaled versions of certain basis functions (Mallat 1999). Although there exist a large number of wavelet functions, in this study wavelets from the Daubechies family (Daubechies 2004) are used. The fast discrete wavelet transform presented by Mallat (1999) is implemented. In the algorithm, a discrete signal is decomposed in a lower scale *approximation* signal A and *detail* signal D. Extending this one-dimensional case to two dimensions (by considering rows and columns consecutively), four new images are obtained: a single approxima-

tion image and a horizontal, vertical and diagonal detail image, all at coarser scales. The applied wavelet transform is described in detail in Verbeke et al. (2006). For a comprehensive discussion of wavelet analysis, the reader is referred to Daubechies (2004) and Mallat (1999).

The fundamental idea behind wavelets is to analyze according to scale. As multi-resolution wavelet transforms decompose an image into a set of approximation and detail images at coarser scales, they are able to characterize local variability within an image at different spatial resolutions. Intuitively, stand boundaries are expected to be "highlighted" in several coarser detail images as they represent high local variability. Using this local variability, represented by the wavelet coefficients as a basis for image segmentation, wavelet analysis is expected to allow for discrimination between forest stands.

4 Materials

The wavelet based segmentation method was developed using artificial forest stands generated as described in Section 2. Several artificial forests were rendered with a spatial resolution of 20cm. Both grid sizes and artificial tree types varied. Small 3x1 grids (containing 3 grid cells) were used, as well as large 10x10 grids (comprising 100 grid cells). The number of artificial tree types varied, according to the grid size, from 3 to 10.

Next the method was applied to real imagery. The first dataset consists of a digital aerial image acquired by the ADS40 Airborne Digital Sensor of an olive grove site in Les Baux de Provence, France. A patchy landscape of olive tree stands and wineyards typifies this area. The ADS40 captures imagery seamlessly along the flown strip, eliminating tedious mosaicing of numerous individual images. The ADS40 dataset comprises panchromatic and multispectral (blue/green/red/near-infrared) data at a spatial resolution of 50cm. The second image dataset covers a forest site in Flanders, Belgium. The site is characterized by a mixture of soft- and hardwood stands. Seven color-infrared aerial photographs (scale 1:5000, acquisition date: October 1987) were scanned, ortho-rectified and mosaiced yielding an image dataset with a very high spatial resolution of 20cm. The dataset was degraded to a spatial resolution of 1m to cut down computational time.

5 Method

5.1 Wavelet based image segmentation

The new method is an image segmentation procedure consisting of a segmentation and a merging step. The method starts with a three-level wavelet decomposition (using the Daubechies 4 wavelet) of the three-band artificial input image, for each spectral band resulting in four new images (a single approximation and three detail images) at three different scale levels (2, 4 and 8, corresponding to spatial resolutions of resp. 40, 80 and 120cm) (Fig 5.1).
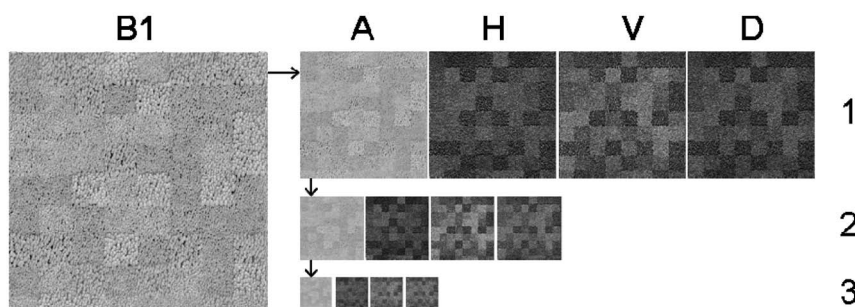

Fig. 5.1 Visual impression of the derived layers from the wavelet transformation of the first band (B1) of an artificial RGB image (with a spatial resolution of 20cm). The 3-level wavelet decomposition results in approximation (A) images, horizontal (H), vertical (V) and diagonal (D) detail coefficients at spatial resolutions of resp. 40, 80 and 120cm

Based on 9 (3 bands x 3 scale levels) approximations and 27 (3 bands x 3 details x 3 scale levels) detail coefficients, an image segmentation is performed. The applied segmentation algorithm is the one introduced by Baatz and Schäpe (2000), which is implemented in the eCognition software tool (eCognition 2000). It is a bottom-up region merging technique and is therefore regarded as a region-based algorithm. The algorithm starts by considering each pixel as a separate object. Subsequently, pairs of objects are merged to form larger segments. Throughout this pairwise clustering process, the underlying optimization procedure minimizes the weighted heterogeneity nh of resulting image objects, where n is the size of a segment and h an arbitrary definition of heterogeneity. In each step, the pair of adjacent image objects which stands for the smallest growth of the defined heterogeneity is merged. If the smallest growth exceeds a user-

defined threshold (the so-called scale parameter s), the process stops. The procedure simulates an even and simultaneous growth of the segments over a scene in each step and the algorithm guarantees a regular spatial distribution of the treated image objects. In this segmentation step, the algorithm utilizes spectral and shape information to extract spatially continuous, independent and homogeneous regions or image objects. The overall heterogeneity h is computed based on the spectral heterogeneity h_{color} and the shape heterogeneity h_{shape} as follows

$$h = w \cdot h_{color} + (1 - w) \cdot h_{shape} \tag{5.1}$$

where w is the user defined weight for color (against shape) with $0 \geq w \geq 1$. Spectral heterogeneity is defined as

$$h_{color} = \sum_c w_c \left(n_m \cdot \sigma_c^m - \left(n_{obj_1} \cdot \sigma_c^{obj_1} + n_{obj_2} \cdot \sigma_c^{obj_2} \right) \right) \tag{5.2}$$

where w_c are the weights attributed to each channel and σ_c are the standard deviations of the spectral values in each channel. The standard deviations themselves are weighted by the object sizes n.

Shape heterogeneity consists of two subcriteria for smoothness and compactness

$$h_{shape} = w_{cmpct} \cdot h_{cmpct} + (1 - w_{cmpct}) \cdot h_{smooth} \tag{5.3}$$

Change in shape heterogeneity caused by a merge m is evaluated by calculating the difference between the situation after and before the merge, with

$$h_{cmpct} = n_m \cdot \frac{l_m}{\sqrt{n_m}} - \left(n_{obj_1} \cdot \frac{l_{obj_1}}{\sqrt{n_{obj_1}}} + n_{obj_2} \cdot \frac{l_{obj_2}}{\sqrt{n_{obj_2}}} \right) \tag{5.4}$$

and

$$h_{smooth} = n_m \cdot \frac{l_m}{b_m} - \left(n_{obj_1} \cdot \frac{l_{obj_1}}{b_{obj_1}} + n_{obj_2} \cdot \frac{l_{obj_2}}{b_{obj_2}} \right) \tag{5.5}$$

For segmentation, w is initially set to 0.8, w_{cmpct} receives a value of 0.1. All 36 input channels are weighted equally by setting each individual layer weight w_c to a value of 1. The scale parameter s is set to a rather

small value resulting in an over-segmentation (the ensuing segments are referred to as sub-objects). The idea is to create homogeneous objects in terms of color and shape (the latter by reduction of the deviation from a compact or smooth shape).

Next, merging operations are performed. In this step, sub-objects are merged into larger image objects based on the defined heterogeneity criterion. This time only color information is included by setting w to a value of 1. Based on wavelet approximation and detail images at three scale levels (2, 4 and 8), sub-objects are merged gradually resulting in image segments corresponding to different forest stands.

5.2 Evaluation

The presented method is qualitatively and quantitatively evaluated against traditional image segmentation i.e. segmentation based on the images' spectral information. Segmentation parameters were optimized for both approaches (wavelet based versus spectral based) enabling a qualitative and quantitative evaluation of the methods' best results. Qualitative evaluation consists of visual inspection, whereas a quantitative assessment is performed using two segmentation evaluation measures: the Purity Index (PI) and the Potential Mapping Accuracy (PMA). Both measures are introduced below.

Stand Purity Index

As stand delineation is the objective, image segments should be smaller or equal to stands. Segments should not exceed parcel[2] boundaries as they then contain information of neighboring stands. A simple measure to assess stand purity is therefore introduced and is called the Purity Index (PI). Pure parcels feature a PI of 1, impure parcels have a PI close to 0. The purity of a parcel p is defined as:

$$PI(p) = \sum_{i=1}^{S} \frac{a_{ip}^2}{a_{i.} \times a_{.p}} \qquad (6.1)$$

---

[2] A parcel is a unit of land under unified ownership. The Purity Index is introduced to measure the purity of parcels in general. When a parcel contains trees, PI evaluates the stand purity.

with S the number of segments, a_{ip} the area of segment i intersecting parcel p, $a_{i.}$ the total area of segment i, and $a_{.p}$ the total parcel area. An example of purity calculation is given in Fig. 6.1.
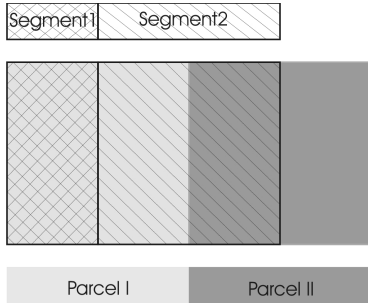
$$PI(p_1) = \frac{5^2}{5 \times 10} + \frac{5^2}{10 \times 10}$$
$$= 0.5 + 0.25 = 0.75$$

Fig. 5.2 An example of PI calculation. Both segments 1 and 2 intersect parcel I. Suppose parcel I, segments 1 and 2 consist of respectively 10, 5 and 10 pixels. According to Eq. 6.1, the purity index of parcel I equals 0.75, representing a rather impure parcel

Potential Mapping Accuracy

Subsequent to segmentation, image segments are typically classified into categorical maps, representing forest/non-forest cover, forest type or, if possible, tree species composition. Apart from supporting the delineation of pure forest stands, it is also important that segmentation facilitates or corroborates future image classification. To evaluate and quantify the value of a segmentation result for future classification, the Potential Mapping Accuracy (PMA) is introduced. This measure assesses the highest feasible overall classification accuracy using a reference dataset:

$$PMA = \frac{C_{max}}{T} \tag{7.1}$$

with T the total number of reference pixels and

$$C_{max} = \sum_{i=1}^{S} \max(n_{i1}, n_{i2}, ..., n_{in_c}) \tag{7.2}$$

the maximum number of correctly classified pixels with S the number of segments, n_{ij} the number of pixels within segment i of class j and n_c the number of classes.

6 Results and discussion

6.1 Artificial imagery

Qualitative evaluation

Fig. 6.1 (left) shows the expected stand delineation for a 10x10 grid filled with 6 artificial tree types. Note that the squared tree stands cut through existing tree crowns which is directly related to the tree stand generating process described in Section 2.1. Artificial trees are added to a grid cell based on their stem positions and taking into account the crown overlap constraints. As a consequence tree crowns may overlap the squared stand boundaries. On the right, the result of the developed method is depicted. The wavelet layers used in the segmentation process are visualized in Fig. 5.1. Wavelet based segmentation first produces a huge amount of sub-objects that are subsequently merged, forming homogeneous objects corresponding to the desired forest stands, as outlined by the predefined grid. Fig. 6.2 shows the discrimination between 4 artificial forest stands in more detail.

Running the segmentation process based on spectral information instead of wavelet texture leads to poor stand delineation results (Fig. 6.3). As all artificial trees were assigned the spectral profile of *Pinus nigra,* all artificial forest stands had similar spectral characteristics and could therefore only be distinguished based on structural features. By incorporating wavelet analysis into the segmentation method, the high local variability represented by stand boundaries is fully exploited at different spatial resolutions (wavelet scale levels). Even when the artificial forest stands are not spectrally discernable, wavelet based image segmentation leads to a (at least visually) satisfactory stand delineation.

Quantitative evaluation

To evaluate the segmentation results quantitatively, the stand purity index PI was calculated for all 66 artificial stands. Mean PIs for all segments, per segmentation level, were computed and the results are shown in Fig. 6.4. It is concluded that segmentation based on the wavelet coefficients produces on average purer forest stands compared to segmentation established with the image's spectral characteristics.
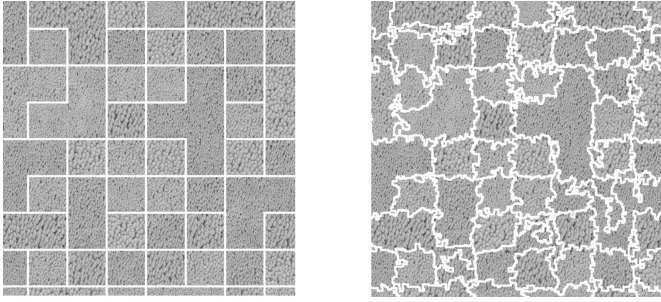
Fig. 6.1. Left: part of an artificially generated forest comprising 66 stands (10x10 grid) of 6 different artificial tree types. The white grid represents the expected stand delineation. Right: wavelet based image segmentation resulting in 64 artificial forest stands
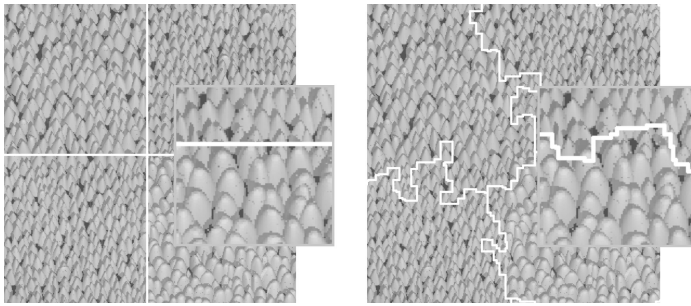

Fig. 6.2. Detail of stand delineation. Left: expected result, right: obtained result
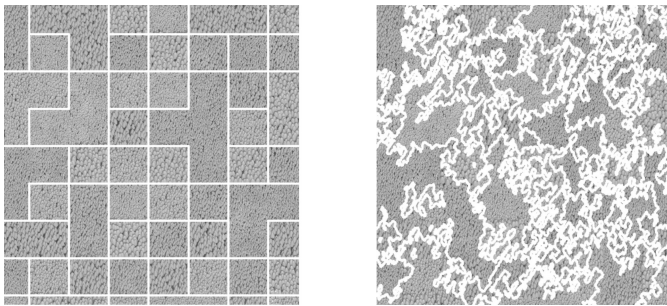

Fig. 6.3. Left: part of the artificially generated forest. The white grid represents the expected stand delineation. Right: segmentation based on image bands' spectral information

In its initial stage, the method results in a large amount of sub-objects, which are further merged through several consecutive merging steps. As

initial sub-objects are small and do not cross forest stand boundaries, both wavelet and spectral based stand purities are high. However, as merging proceeds (and the number of segments decreases), wavelet based stand purities are markedly higher than spectral based PIs. In other words, throughout the merging steps, wavelet based image segments keep within forest stands without exceeding stands boundaries too much. Consequently, resultant segments respect the stand edges, tending to purer forest stands.

Next to supporting pure forest stands, the segmentation method is evaluated on its ability to induce high classification performances. Does the underlying segmentation represent the ideal starting situation for classifying all segments into 6 forest stand types? Potential Mapping Accuracies (PMA) are therefore computed reflecting the highest feasible overall classification accuracy. Fig. 6.5 (left) shows that wavelet based segmentation results in higher potential mapping accuracies compared to spectral based segmentation, especially when merging proceeds, that is, when the number of segments decreases.
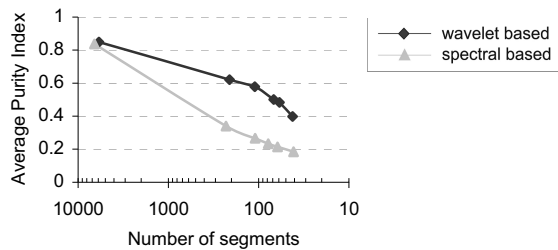
Fig. 6.4 Average stand purity index in function of the number of image segments for wavelet and spectral based segmentations of the artificial imagery
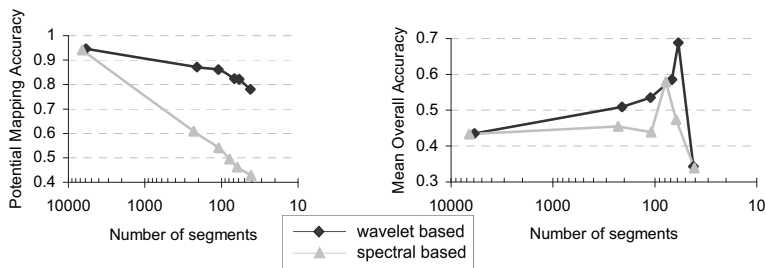
Fig. 6.5 Left: Potential Mapping Accuracy as function of the number of segments. Right: Mean overall accuracies for subsequent merging steps

To check whether this holds for actual classifications, the artificial scene (10x10 grid) is classified using a feed-forward backpropagation neural network. Segments are classified based on the per segment mean and standard deviations of the spectral input bands (green, red and near-infrared). In total 6 artificial tree/stand types are discriminated. Hence, the applied network features a three-layered architecture with 6 input neurons, 10 hidden units and 6 output neurons. A hyperbolic tangent activation function is used. As network initialization is a critical issue and influences network performance, neural network classification is repeated 20 times. Mean overall accuracies are computed on a per pixel basis. Fig. 6.5 (right) shows the overall accuracies for wavelet based and spectral based classifications.

From comparison of the left and right graphs of Fig. 6.5, it is obvious that mean overall accuracies are lower than PMAs, both for the wavelet and spectral based approaches. The PMAs reflect the highest feasible mapping accuracy. As stated in section 2 all trees were assigned the *Pinus nigra* spectral profile. Consequently, the artificial tree types can only be distinguished based on structural characteristics, and therefore mean overall accuracies are expectedly low. The right side of Fig. 6.5 shows that, throughout the merging process, the wavelet based mean overall accuracies are roughly higher than the spectral based mean accuracies. However, when merging proceeds and results in only tens of rather large segments, the number of training samples drastically drops beyond a critical level and the neural network classifier is unable to learn the problem. At that point, the mean overall accuracy is low for both the wavelet and spectral based approaches.
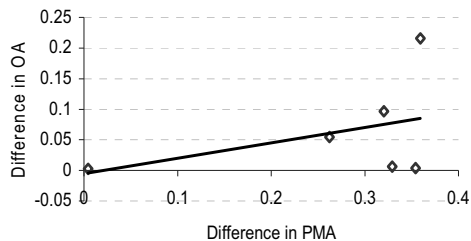

Fig. 6.6 The relation between the difference (wavelet based vs. spectral based) in PMA and the difference in mean overall accuracy for the artificial image

Classifier performance is affected by several processes: the accuracy is a function of the training set size, the quality of the segmentation and the separability of the dataset. So we do not expect the mean overall accuracy to behave exactly as the potential mapping accuracy that only depends on

the segmentation quality. Still, the accuracy curves are expected to move away from each other. Ignoring the utmost right point, this trend is roughly discernable. The tendency however is not very clear because of the difficulty of the classification problem. This is illustrated in the scatter plot shown in Fig. 6.6. It is assumed that the relation between the difference in PMA and the observed difference in overall accuracy will be stronger for easier classification problems.

6.2 Real Imagery

The method reported above is applied to two real datasets: one covers an olive grove site in Les Baux de Provence, France, the second dataset addresses a forest site in Flanders, Belgium.

Qualitative evaluation

Although trees in both image datasets are spectrally different, the (sub)meter resolution induces stands with high and sometimes similar spectral heterogeneity. Because of the plant distance of olive trees (which is higher than the image spatial resolution), olive tree stands feature especially high spectral heterogeneity. Spectral based segmentation of such imagery is likely to fail and leads to branched segments with a irregularly-shaped borderline traversing several stands (Fig. 6.7 top). In such cases multi-level wavelet decomposition is advantageous (Fig. 6.7 bottom) since variability is characterized locally.

Quantitative evaluation

Fig. 6.8 clearly shows that wavelets are well-suited for approximating data with sharp discontinuities like stand boundaries. For both real datasets, resultant segments consequently respect more the stand edges tending to purer stands.

Both spectral and wavelet based segmentation methods are evaluated on their power to guarantee high classification performances. Because of the lack of reference data for the forest site in Belgium, potential mapping accuracies are only computed for the olive grove site in France. Fig. 6.9 (left) shows that wavelet based segmentation results in higher potential mapping accuracies compared to spectral based segmentation. The quality of the wavelet based segmentation is therefore assessed higher.
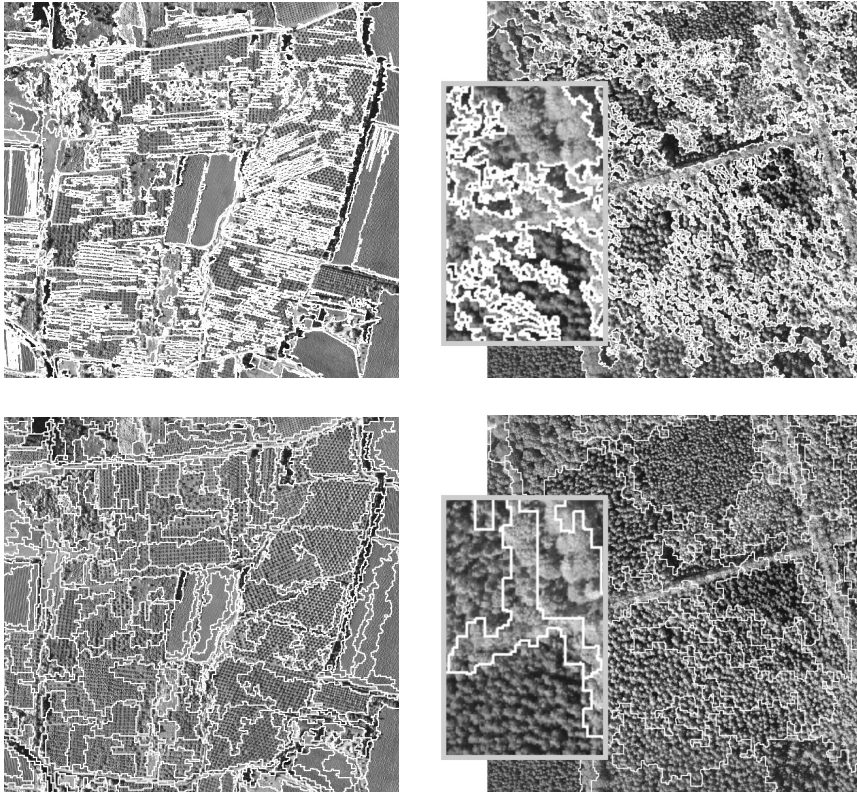
Fig. 6.7 Top: detail of spectral based segmentation of the olive grove site (left) and the forest site (right). Bottom: detail of wavelet based segmentation of the same sites
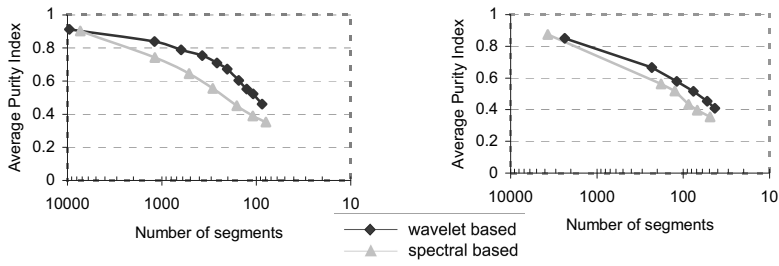

Fig. 6.8 Average purity indexes for the olive grove site (left) and the forest site (right)
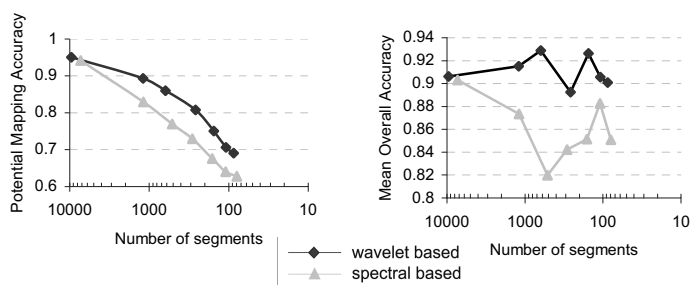
Fig. 6.9 Potential mapping accuracy and mean overall classification accuracy

To test whether PMA can be used to predict actual classification accuracy, 20 neural networks are trained to classify the scene into olive/non-olive parcels. This time the networks have 6 inputs, 10 hidden units and 2 output neurons. In Fig. 6.9, it can be observed that the mean overall accuracies are generally lower than PMAs, both for the wavelet and spectral based approaches. Hence, the PMAs can be interpreted as reflecting the highest possible mapping accuracy. Again the mean overall accuracy curves are expected to move away from each other. Here this trend is clearer than in the case of the artificial imagery, as indicated by the scatter plot shown in Fig. 6.10. It is assumed that this is due to the less complicated classification problem.
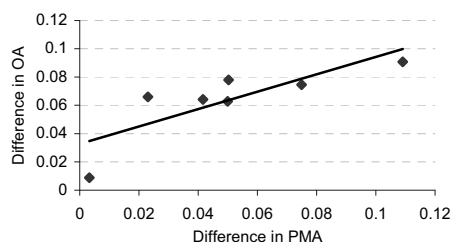
Fig. 6.10 The relation between the difference in PMA and the difference in mean overall accuracy for the olive grove site

7 Conclusion

This paper presents a new segmentation method to delineate stands from very high resolution optical imagery. The method was developed using ar-

tificially generated forest stands and was applied in a later stage to real imagery. The proposed method starts with multi-level wavelet decomposition. Based on the wavelet approximation and detail images, a segmentation is performed. The results indicated that wavelet based image segmentation outperforms traditional segmentation, i.e. segmentation based on spectral properties, when forest stands are similar in terms of spectral characteristics, but different in terms of textural properties. Multi-level wavelet analysis proved to be a valuable tool for characterizing local variability in image texture and therefore allows for the discrimination between stands. To evaluate the segmentation results, two measures were developed. The stand purity index was introduced to evaluate whether the segmentation method supports pure stands. The potential mapping accuracy, on the other hand, was presented to evaluate the segmentation method on its power to guarantee high classification performances. Both measures were found to be appropriate as segmentation evaluation criteria.

Acknowledgements

References

Baatz M, Schäpe A (2000) Multiresolution segmentation - an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Baschke T, Griesebner J (eds) Angewandte Geographische Informationsverarbeitung, Wichmann-Verlag, Heidelberg, pp. 12–23

Culvenor D (2002) TIDA: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. Computers and Geosciences 28, pp. 33–44

Daubechies I (2004) Ten lectures on wavelets. CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61. Society for Industrial and Applied Mathematics, Philadelphia.

Dekker RJ (2003) Texture analysis and classification of ERS SAR images for map updating of urban areas in the netherlands. IEEE Transactions on Geoscience and Remote Sensing 41(9), pp. 1950–1958

eCognition (2000) Definiens Imaging GmbH, Munchen, Germany

Gougeon F (1995a) Comparison of possible multispectral classification schemes for tree crowns individually delineated on high spatial resolution MEIS images. Canadian Journal of Remote Sensing 21, pp. 1–9

Gougeon F (1995b) A crown-following approach to the automatic delineation of individual tree crowns in high spatial resolution aerial images. Canadian Journal of Remote Sensing 21, pp. 274–284

Haara A, Haarala M (2002) Tree species classification using semi-automatic delineation of trees on aerial images. Scandinavian Journal of Forest Resources 17, pp. 556–565

Hay GJ, Castilla G, Wulder MA, Ruiz JR (2005) An automated object-based approach for the multiscale image segmentation of forest scenes. International Journal of Applied Earth Observation and Geoinformation 7(4), pp. 339–359

Kandaswamy U, Adjeroh DA, Lee AC (2005) Efficient texture analysis of SAR imagery. IEEE Transactions on Geoscience and Remote Sensing 43(9), pp. 2075– 2083

Larsen M (1997) Crown modeling to find tree top positions in aerial photographs. In: Proc. Third International Airborne Remote Sensing Conference and Exhibition, Copenhagen, Denmark, Vol. II, pp. 428–435

Leckie D, Gougeon F, Walsworth N, Paradine D (2003) Stand delineation and composition estimation using semi-automated indiviual tree crown analysis. Remote Sensing of Environment 85, pp. 355–369

Li J (2004) Wavelet-based feature extraction for improved endmember abundance estimation in linear unmixing of hyperspectral signals. IEEE Transactions on Geoscience and Remote Sensing 42(3), pp. 644–649.

Mallat S (1999) A wavelet tour of signal processing. Academic Press, London.

Mertens K, Verbeke L, Westra T, De Wulf R (2004) Subpixel mapping and subpixel sharpening using neural network predicted wavelet coefficients. Remote Sensing of Environment 91, pp. 225–236

Park JH, Kang MG (2004) Spatially adaptive multiresolution multispectral image fusion. International Journal of Remote Sensing 25(23), pp. 5491–5508

Persistence of Vision Pty. Ltd. (2004) Persistence of vision (tm) raytracer. Technical report, Persistence of Vision Pty. Ltd., Williamstown, Victoria, Australia. http://www.povray.org

Sgrenzaroli M, Baraldi A, De Grandi GD, Eva H, Achard F (2004) A novel approach to the classification of regional-scale radar mosaics for tropical vegetation mapping. IEEE Transactions on Geoscience and Remote Sensing 42(11), pp. 2654–2669

Verbeke LPC, Van Coillie FMB, De Wulf RR (2006) Object-based forest stand density estimation from very high resolution optical imagery using wavelet-based texture measures. In: Proceedings of the 1st International Conference on Object-based Image Analysis (OBIA 2006), http://www.commission4.isprs.org/obia06/Papers/13_Automated%20classific ation%20IC%20I%20-%20Forest/

Warner TA, Lee JY, McGraw JB (1998) Delineation and identification of individual trees in the eastern deciduous forest. In: Proc. International Forum: Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Canadian Forest Service, Pacific Forestry Centre Victoria, British Columbia, p. 8191

Wulder MA, White JC, Hay GJ, Castilla G (2007) Pixels to objects to information: Spatial context to aid in forest characterization with remote sensing. OBIA book chapter paper.

Zeng ZH, Cumming IG (2001) SAR image data compression using a tree-structured wavelet transform. IEEE Transactions on Geoscience and Remote Sensing 39(3), pp. 546–552

Chapter 2.9

Quality assessment of segmentation results devoted to object-based classification

J. Radoux, P. Defourny

Department of Environmental Sciences and Land Use Planning, Université catholique de Louvain, Belgium

ABSTRACT: Object-based image analysis often uses image segmentation as a preliminary step to enhance classification. Object-based classification therefore relies on the quality of the segmentation output. This study evaluates the relevance of quantitative segmentation quality indices to object-based classification. Image segmentation is expected to improve the thematic accuracy of classification but the counterpart is an increased chance of boundary artefacts. Goodness indices were used to assess the former while discrepancy indices evaluated boundary quality. Inter-class Bhattacharyya distance was used to test the relevance of the goodness indices. The results showed that the use of global goodness indices, which did not require *a priori* information about the study area, was relevant in the case of object-based classification. In this context, the goodness index based on intra-class standard deviation was more useful than the one based on mean object size. On the other hand, it was shown that object size improved class discrimination but this could deteriorate the boundary quality. The use of complementary discrepancy indices is therefore required in the case of frequent under-segmentation.

1 Introduction

Image segmentation is a major low level task in the context of object-based image analysis (Soille 2000; Marr 1982). The complementary high

level tasks are either object recognition or object-based image classification. In the former, the algorithm identifies groups of pixels belonging to specific objects (e.g. Inglada 2005; Leckie et al. 2003). In the latter, meaningful groups of pixels are generated in the entire image with the aim of classifying images beyond pixel-based classification (e.g. Wang 2002).

Previous studies on object-based image classification highlighted the key advantages of this approach compared with pixel-based classification (e.g. Almeida-Filho and Shimabukuro 2002; Lobo et al. 1998). First, it avoids the salt-and-pepper effect likely to occur in classifications of textured images (Wulder et al. 2004). Second, it improves the classification results thanks to more robust descriptors (e.g. mean of spectral values and texture indices) (e.g. Pekkarinen 2002). Third, it suits object reasoning and the use of new characteristics such as shape, topology and context (e.g. Hay et al. 2003). However, image segmentation is an additional step which has its drawbacks (Munoz et al. 2003): Object boundaries can be irregular or incorrectly located, and one group of pixels may envelop different land cover due to under-segmentation.

Considering the large number of existing segmentation algorithms and their versatility (e.g. Guigues et al. 2006; Baatz and Schäpe 2000; Jung 2007; Hay et al. 2003; Pal and Pal 1993; Zhang 1997), the choice of an appropriate segmentation algorithm must rely on objective methods to assess segmentation quality. The performance of segmentation comparison methods was evaluated by Zhang (1996) in the context of object recognition. He identified three types of quality assessment methods, namely analytical comparison, empirical goodness and empirical discrepancy, and concluded that empirical discrepancy indices were the most effective. In remote sensing, recent studies assessed object recognition applications (building or field detection) using object-based indices instead of pixel based (Zhan et al. 2005; Möller et al. 2007). Möller et al. (2007) used over- and under-segmentation indices to find an optimal segmentation parameter. Over- and under-segmentation indices were also used to assess image segmentation results in the context of object-based classification of land cover maps (Carleer et al. 2005), but the relevancy of the metrics was not assessed in this case.

This paper assesses quantitative indices for the comparison of segmentation results in the context of object-based classification. According to the advantages and the disadvantages of object-based classification, one group of indices is used to evaluate how much the classification could be improved thanks to segmentation and a second group assessed the positional quality of the boundaries of the delineated land cover classes. The method was illustrated on real segmentation of a Quickbird image using a combination of three parameters of the same segmentation algorithm.

2 Segmentation quality indices

Section 2.1 aims at quantifying how much the segmentation step could enhance a standard classification (e.g. maximum likelihood, kNN) compared with pixel-based approaches. Global indices are proposed because comprehensive reference data is not available in practice for mapping application. The sensitivity of these indices to segmentation parameters is evaluated and their reliability as indicator of class discrimination is tested thanks to a pair wise class distance metric using *a priori* information. On the other hand, complementary indices are proposed in section 2.2. Those are discrepancy indices that estimate the errors along the edges.

2.1 Quality of object characteristics

In object-based classification studies, it is expected from object characteristics to be more reliable than pixel characteristics. Indeed, it has been shown that replacing the values of pixels belonging to the same cluster by their means reduced the variance of the complete set of pixels (See Huygens' theorem in Edwards and Cavalli-Sforza 1965). In other words, large groups of pixels in an image contribute to the reduction of the variance. Assuming that small intra-class variance improves parametric classification, indices based on the mean object size, already used to assess over- and under- segmentation in object recognition studies, are good candidates for object-based classification. The Inverse of the Number of Object (INO) was used as a first quantitative goodness index.

The second global goodness index was based on the same assumption but was a direct measure of the class uniformity. The Normalized Post Segmentation Standard deviation (NPSS) was computed after replacing each pixel value by the mean values of its parent object and normalized by the pixel-based standard deviation (equation 1) in order to be bounded by zero (single object) and one (pixel level).

$$NPSS = \sqrt{\left(\sum_{i=1}^{no} n_i \left(\overline{x} - \overline{x}_i\right)^2\right) \bigg/ \left(\sum_{1}^{N} \left(\overline{x} - x\right)^2\right)} \qquad (1)$$

where $\overline{x}_i$ is the object mean, no the number of objects and n_i the number of pixels included in the ith object.

As a matter of fact, the reduction of the global variance does not guarantee the success of a classification. Actually, global variance reduction does not impel a variance reduction in all classes. The global NPSS was there-

fore compared with class based NPSS. Moreover, the assumption that a small intra-class variance improves the classification result is not always verified: large variance difference between two classes (for instance, bimodal versus unimodal distributions) can improve classification, and a shift of the mean class values may reduce their separability. A dissimilarity metric, the Bhattacharyya Distance (BD), was used in this study together with a comprehensive reference dataset to test the relevance of the proposed goodness indices. Bhattacharyya Distance was chosen because it can give a bound on classification errors (Webb 2002) and has a term able to compare the covariance matrices. Furthermore, it can be used with any number of characteristics and takes the correlation between each characteristic into account. Assuming the normal distribution of the variables (equation 2), it is interesting to note that it simplifies to exactly 1/8 of the Mahalanobis distance when the covariance matrices are equal:

$$
\begin{aligned}
BD = {}& \frac{1}{8} (\mu_1 - \mu_2)^T \left(\frac{\Sigma_1 + \Sigma_2}{2} \right)^{-1} (\mu_1 - \mu_2) \\
& + \frac{1}{2} \log \left[\left(\frac{|\Sigma_1| + |\Sigma_2|}{2} \right) / \sqrt{|\Sigma_1||\Sigma_2|} \right]
\end{aligned}
\tag{2}
$$

where μ_i is the vector of mean values for class I and Σ_i the corresponding covariance matrix, both estimated after weighting each object characteristics by the object size in the class i, like in the case of NPSS. The Normalized Bhattacharyya Distance (NBD) was computed by dividing BD_{object} by BD_{pixel}.

2.2 Quality of object boundaries

The main drawbacks of segmentation algorithms are the artifact along the boundaries and the missing boundaries. These artifacts along the boundaries were quantified using discrepancy indices in order to assess the final map precision. The positional quality of the edges was evaluated based on the concepts of accuracy and precision as described in Mowrer and Congalton (2000). Accuracy and precision were estimated respectively by the bias and the mean range of the distribution of boundary errors. Based on the method presented in Radoux and Defourny (2005), these values were derived after intersecting the image objects with the visually updated vector reference dataset from the NGI (fig.1). The area of the non-matching polygons was counted as negative when it induced an underestimate of the reference object area (omission error) and as positive in the

other case (commission error). The bias was calculated by adding the signed areas for each interface type in order to identify systematic commission or omission errors. The mean range was defined as the sum of the unsigned area, minus the absolute value of the bias, to measure the occurrence and the size of deviations around the biased boundary location.
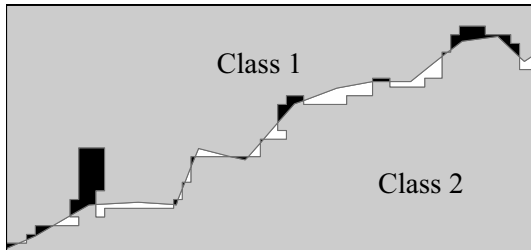

Fig. 1. Errors along the edges of a segmentation output. Black polygons are omissions (-) and white polygons are commissions (+) with respect to class 1

These two estimators were normalized by the total interface length in order to be expressed in map unit (i.e. meters) and to be independent of object area and boundary length. Each bias and mean range was estimated from more than 3 000 polygons, giving a high reliability.

3 Case study

The sensitivity and the robustness of the quantitative indices were tested on 60 image segmentation results. The raw data was a Quickbird image which was segmented using combinations of three parameters of the multiresolution segmentation algorithm implemented in Definiens.

3.1 Data and study area

A study area located in Southern Belgium has been chosen for its fragmented landscape. It is a rural area including patches of crop fields, pastures, sparsely urbanized areas and forests. This landscape is also characterized by the presence of hedges and isolated trees.

The 16 km² test site was covered by a cloud free multi-spectral Quickbird image of July 2006 and by a reference 1/10 000 vector database from the Belgian National Geographic Institute (NGI) updated on the field in 2004. The Quickbird "Basic" image was orthorectified with ENVI soft-

ware. This was based on 13 ground control points from the NGI vector database and on a digital elevation model built with 5 m contour lines. The root mean square error of the model was 2.4 m and the image was resampled at 2.56 m using a cubic convolution.

In order to have a reference land cover database that matches the image everywhere, NGI boundaries were visually updated on image boundaries using a geographic information system software. A constant scale of 1/2 500 was used at this purpose. Furthermore, a new "shadow" class (Fig. 2) was added to the existing classes, which were "deciduous forest", "mixed forest", "coniferous forest", "arable lands, meadows and pasture", "lake" and "impervious area". Finally, arable lands that were not covered by vegetation on the acquisition date were classified separately and a "miscellaneous" class was created to include all elements that did not belong to any of the above listed classes. The updated reference database was thus composed of 9 categories, and included 586 polygons leaving out the "miscellaneous". Each object was labeled in one of these 9 classes using a majority rule according to the updated reference database.
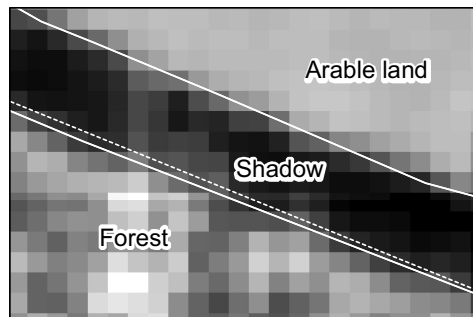
Fig. 2. Manually delineated shadow illustrated on a subset of the Quickbird image. The dashed line corresponds to the NGI database

3.2 Segmentation

The segmentation algorithm implemented in Definiens® software (Baatz and Schäpe 2000) was used to illustrate the proposed quality indices. It uses 3 parameters that allow users to tune the segmentation outputs. This algorithm is based on region merging. Each pixel is first considered as a separate object. The merging decision is based on the weighted sum of local color and shape homogeneity criteria: the pairs of objects with the smallest increase in heterogeneity are merged iteratively. The process ends when the smallest increase is above a user-defined threshold (scale parameter). The weight given to the shape constraint in the heterogeneity cri-

terion is called shape parameter. A third parameter, the compactness parameter, constrains segments' shape to be compact when it is large and smooth when it is small.

Sixty combinations of the 3 parameters (scale = {10, 20, 30, 40, 50, 60}, shape = {0, 0.2, 0.4, 0.6} and compactness = {0, 0.5, 1}) were used for the segmentation of the multi-spectral image, including near-infrared, red and green bands. Obviously, a null shape parameter disabled the choice for a compactness parameter. Beside this, scale parameter values over 60 were tested in preliminary study but were not assessed because the under-segmentation was too large. For each segmentation results, the mean reflectance values in the three spectral bands were computed.

4 Results

All reported below were computed for the entire dataset excepting the miscellaneous class. The number of polygons produced by image segmentation ranged between 900 and 44 000, so that the ratio between the number of regions in the segmented image and the number of objects in the reference was between 0.65 and 0.01. The largest number of polygon corresponded to the smallest scale and shape parameters. Section 4.1 displays the quality of object boundaries and section 4.2 includes the details about object characteristics. The models tested were limited to second order polynomials of multiple linear regression.

4.1 Goodness indices

The two goodness indices provided meaningful information about the segmentation results. However, global NPSS was more correlated than INO with the average of NBD on all the pairs of classes ($R^2 = 0.97$ against 0.87). As shown in figure 3, the ranking between two segmentation results was only affected in the case of small differences in the case of global NPSS. These good results could be explained because the mean class values were not modified by the segmentation, except for water bodies due to the inclusion of small islands.

The global NPSS index was sensitive on the segmentation parameters. As expected, object size was the major parameter: it explained more than 80 % of the variance. The shape parameter also strongly contributed to the reduction of class variance. A multi-linear regression model ($R^2 > 0.93$), was built based on shape parameter, scale parameter and the product of these two parameters (figure 4). On the other hand, the type of shape con-

straint had no significant influence (t-test, $\alpha = 0.05$) in the main range of scale and shape parameters.
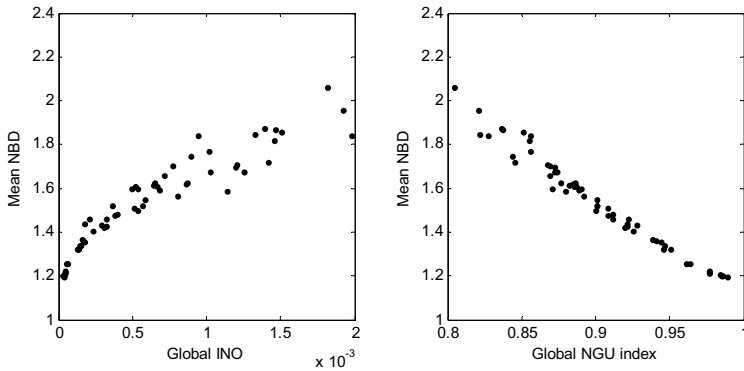

Fig. 3. Relationship between the goodness indices and the average NBD. Global INO (left) shows larger dispersion than global NPSS (right)
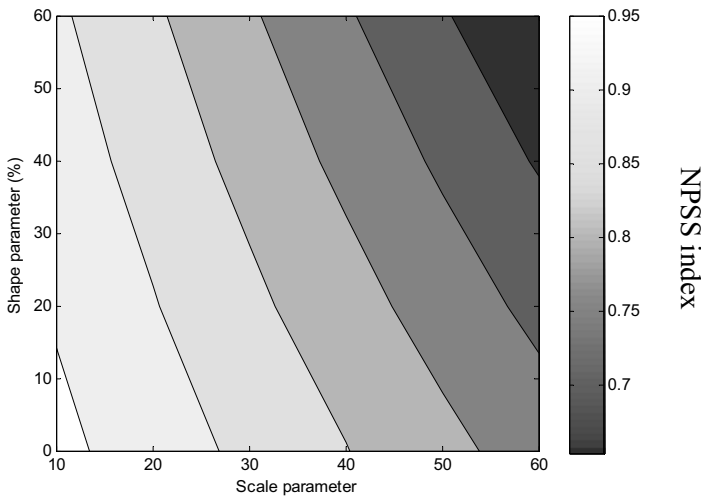

Fig. 4. Effect of scale and shape parameters on the global normalized post segmentation standard deviation (NPSS) illustrated with the near-infrared

Compared with the class based NPSS, the global NPSS proved to be robust with R^2 from 0.95 to 0.99 except in the case of water bodies and impervious surface were it was close to 0.6. On the other hand, the effect of

segmentation on the per class NPSS could vary (figure 5), which shows a sensitivity of the segmentation algorithm to the land cover class.
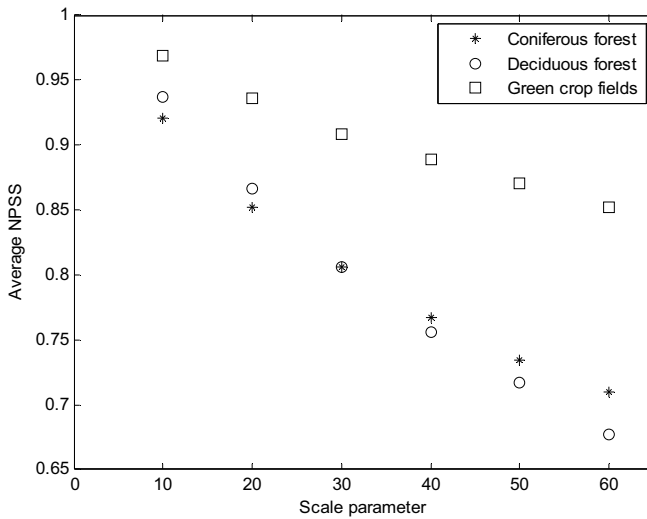
Fig. 5. Impact of land cover type on the evolution of NPSS in the near-infrared.

The results for BD between the main classes are illustrated together with the corresponding NBD (figure 6). This highlight some advantages of the segmentation that cannot be seen with NPSS: hardly separable classes achieve the best improvement compared with pixel-based classification, and this improvement was observable even for small scale parameters.
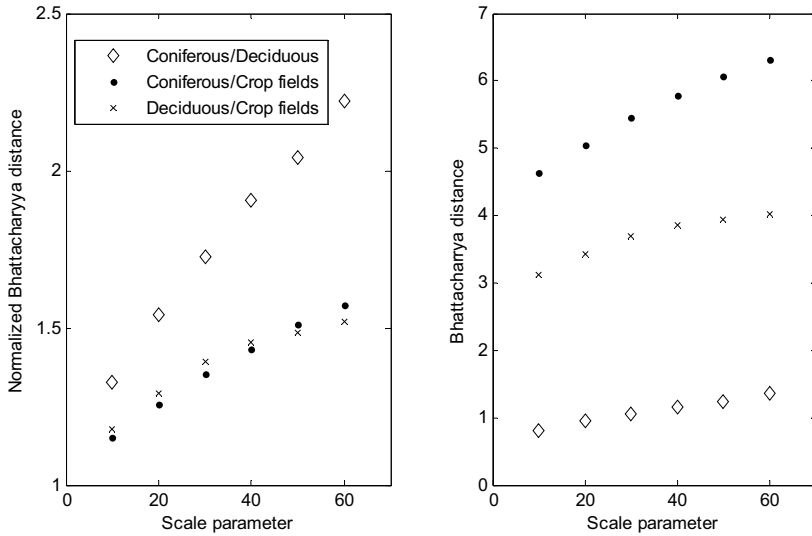
Fig. 6. Evolution of normalized (left) and absolute (right) Bhattacharyya distances illustrated for three pairs of land cover classes

4.2 Quality of object boundaries

The absolute boundary error was sensitive to the under-segmentation that mainly occurred for large scale parameters. The absolute boundary error index was also able to detect artifacts along the class boundaries. This resulted in a high correlation ($R^2 > 0.94$) between the shape parameter and the absolute boundary error but there was no significant effect of the shape parameter. On the other hand, compactness parameter did not lead to linear effect, but it can be seen on table 1 that compactness induced on average a small reduction of the absolute errors. In fact, as could also be seen by visual assessment, the compactness parameter was only effective for large shape parameters.

Table 1. Average of absolute errors (in meters) on boundary position between deciduous and coniferous forests, for combination of segmentation parameters, i.e. scale parameter between 10 and 60 and compactness, illustrated for forest/arable land interfaces.

Scale	10	20	30	40	50	60	Mean
Smooth	2.3	3.4	4.1	4.7	5.3	5.7	4.3
Mixed	2.2	3.2	3.8	4.3	4.8	5.3	3.9
Compact	2.1	3.0	3.7	4.2	4.7	5.1	3.8

In any case, there was a strong effect of the class boundary type on the indices, as illustrated for two types of class boundaries in figure 7. A significant bias in the delineated edge position was also highlighted. This bias was always below the pixel size but its contribution to the absolute errors was up to 30% in the case of large scale parameters.
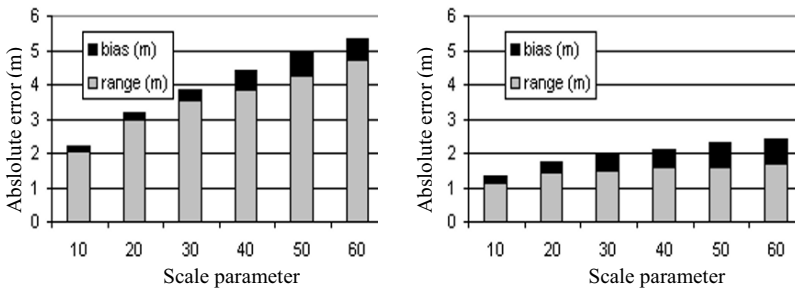

Fig. 7. Boundary errors for two types of boundaries with respect to scale parameter. Deciduous/coniferous are on the left and forest/arable land on the right. The sums of bias and mean range correspond to the absolute error

5   Discussion

Global NPSS was more relevant than INO for the assessment of segmentation results. The high correlation between NPSS and BD indices corroborates the robustness of the former despites its simple formulation. However, global NPSS should be used in an appropriate frame to keep its relevance. Actually, it was designed to globally assess segmentation results used for multi-class classification of land cover/land use maps or the like. The effect of the segmentation on uniformity is different for each class. For instance, there is a small variance inside crop fields but there are different phenological states between fields, so that the reduction of intra-class variance thanks to the segmentation is limited. On contrary, coniferous forests are heterogeneous but the stands are similar across the all study area. For the main classes, the ranking between segmentation results is not affected by these differences. However, when looking at a specific class covering a small percentage of the entire study area, e.g. water bodies, or composed of objects close to the spatial resolution of the image, e.g. buildings, the indices developed for object recognition assessment (e.g. Zhan 2005) are more appropriate. Another restriction on the use of global indices alone is the presence of large delineation errors (i.e. missing or irregu-

lar boundaries). In this case, the use of a complementary discrepancy index is necessary.

The proposed boundary quality indices measured the discrepancy along the border and were sensitive to under-segmentation. It was normalized in order to be expressed in mapping units and therefore allowed diverse segmentation results to be compared on different images of the same spatial resolution. Again, it was designed for large objects and object-based discrepancy indices are more appropriate when the objects are composed of few pixels. Furthermore, it is important to consider the natural uncertainty of some inter-class boundaries when interpreting the boundary quality indices. Some boundaries, such as ecotones, are indeed inherently less precise than other (Hay et al. 2001) or biased due to their definition. For instance, forest boundaries are drawn as straight lines at trunk location on the NGI reference map while the satellite captures the uneven crown boundaries. While boundary quality should be a major concern in object-based image analysis, the degree of line generalization of these boundaries should thus be considered during the quality control.

The quality of object boundaries and object characteristics were often negatively correlated: a) The scale parameter had a strong positive effect on object characteristics but decreased the boundary at the same time, b) the shape parameter did not affect the boundary quality and helped to improve the discrimination between classes, c) and the compactness was not significant although compactness had a major impact on boundary quality and class discrimination for large scale and shape parameters (figure 8).
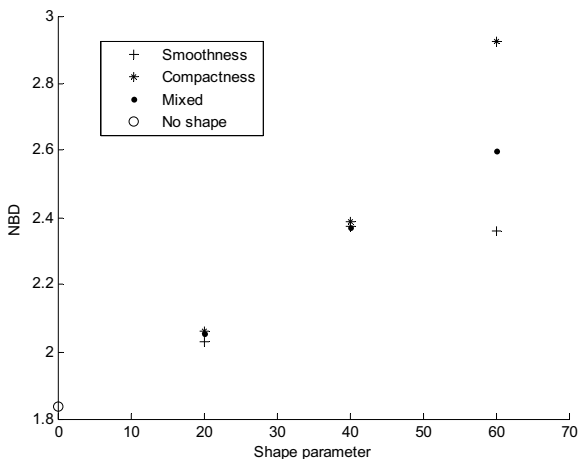

Fig. 8. Evolution of the NBD between the coniferous and the deciduous classes for a scale parameter of 60. The effect of the compactness parameter is only significant for the largest shape parameter

In any case, large scale mapping is limited to a given boundary quality (e.g. USGS standard). This boundary constraint has led to several post-processing methods to enhance the final edge quality (Munoz et al 2003). On the other hand, figure 6 suggests that discrimination between two close classes can be improved even with a small scale parameter, keeping good boundary quality. Moreover, object-based classification algorithms often use object characteristics (e.g. shape, context…) unavailable to segmentation algorithms. Classification/segmentation cycles could thus help to discriminate between classes whilst keeping small boundary errors.

Conclusion

The quality indices used in this study proved to be relevant to assess segmentation results devoted to object-based classification. The best goodness index was the normalized post segmentation standard deviation (NPSS). It was sensitive to the segmentation parameters and provided coherent rankings of the segmentation results. Furthermore, NPSS is easy to calculate as it does not require *a priori* information and it assists the comparison between object-based and pixel-based classification performance. However, it did not account for under-segmentation and positional errors along the boundaries. The use of a complementary boundary discrepancy index was therefore necessary to assess map precision.

Acknowledgement

References

Almeida-Filho R, Shimabukuro YE (2002) Digital processing of a Landsat-TM time series for mapping and monitoring degraded areas caused by independent gold miners, Roraima state, Brazilian Amazon. Remote Sensing of Environment, 79:42-50.

Baatz M, Schäpe A (2000) Multi-resolution segmentation – an optimization approach for high quality multi-scale segmentation. In: Strobl, J et al. (Hrsg.): Angewandte Geographische Informationsverarbeitung XII, 12-23. Beiträge zum AGIT Symposium Salsburg 2000, Karlruhe, Herbert Wichmann Verlag.

Carleer AP, Debeir O & Wolff E (2005) Assessment of very high resolution satellite image segmentations. Photogrammetric Engineering and Remote Sensing, 71(11), 1284-1294.

Definiens Imaging GmbH (2002) eCognition user guide. Definiens Imaging GmbH, Munich, Germany.

Edwards AWF, Cavalli-Sforza, LL (1965) A method for cluster analysis, Biometrics, 21: 362-375

Foody GM (2002) Status of land cover classification accuracy assessment. Remote Sensing of Environment, 80:185-201.

Guigues L, Cocquerez J-P, Le Man H (2006) Scale-set image analysis. International Journal of Computer Vision, 68(3):289-317

Hay GJ, Marceau DJ, Bouchard A, Dubé P (2001) A multiscale framework for landscape analysis: object-specific up-scaling. Landscape Ecology, 16: 1-19

Hay GJ, Blaschke T, Marceau DJ, Bouchard A (2003) A comparison of three image-object methods for the multiscale analysis of landscape structure. ISPRS Journal of Photogrammetry & Remote Sensing, 57:327-345.

Inglada J (2005) Use of pre-conscious vision and geometric characterizations for automatic man-made object recognition. Proceedings of the International Geoscience And Remote Sensing Symposium 2005. Seoul, Korea.

Jung CR (2007) Combining wavelets and watersheds for robust multiscale image segmentation. Image and Vision Computing, 25:24-33

Leckie DG, Gougeon FA, Walsworth N, Paradine D (2003) Stand delineation and composition estimation using semi-automated individual tree crown analysis. Remote Sensing of Environment, 85:355-369

Lobo A, Moloney K, Chiariello N (1998) Fine-scale mapping of a grassland from digitized aerial photography: an approach using image segmentation and discriminant analysis. International Journal of Remote Sensing, 19(1): 65-84

Marr D (1982) Vision: a computational investigation into the human representation and processing of visual information. WH Freeman and Company, NY pp 29-61

Möller M, Lymburner L, Volk M (2007) The comparison index: A tool for assessing the accuracy of image segmentation. International Journal of Applied Earth Observation and Geoinformation 9:311–321

Mowrer HT, Congalton RG(Eds) (2000) Quantifying Spatial uncertainty in natural resources : Theory and application for GIS and remote sensing, Ann Arbor Press, Chelsea, Michigan, 234 p

Munoz X, Freixenet J, Cufi X, Marti J (2003) Strategies for image segmentation combining region and boundary information. Pattern recognition letters 24:375-392

Pal NR, Pal SK (1993) A review on image segmentation techniques. Pattern recognition 26, 1277-1294

Pekkarinen A (2002) Image segment-based spectral features in the estimation of timber volume. Remote Sensing of Environment, 82: 349-359

Radoux J, Defourny P (2005) Accuracy assessment of forest stand delineation using very high spatial resolution satellite images. ForestSAT 2005 Proceedings, Boras, pp 84-89

Soille P (2000) Morphological image analysis: Principles and applications. Springer, Berlin Heidelberg New York

Stehman SV (1997) Selecting and interpreting measures of thematic classification accuracy. Remote Sensing of Environment, 62: 77-89

Webb A (2002) Statistical pattern recognition. Wiley and sons, West Sussex, England

Wulder MA, Skakun RS, Kurz WA, White JC (2004) Estimating time since forest harvest using segmented Landsat ETM+ imagery. Remote Sensing of Environment, 93:179-187

Zhan Q, Molenaar M, Tempfli K, Shi W (2005) Quality assessment for geo-spatial objects derived from remotely sensed data. International Journal of Remote Sensing 26:14 2953-2974

Zhang YJ (1996) A survey on evaluation method for image segmentation. Pattern recognition 29(8): 1335-1346

Zhang YJ (1997) Evaluation and comparison of different segmentation algorithms. Pattern recognition letters 18: 963-974

Section 3

Automated classification, mapping and updating: forest

Chapter 3.1

Object-based classification of QuickBird data using ancillary information for the detection of forest types and NATURA 2000 habitats

M. Förster, B. Kleinschmit

Berlin University of Technology, Department of Geoinformation
Straße des 17. Juni 145 (EB 5), 10623 Berlin, Germany
michael.foerster@tu-berlin.de, birgit.kleinschmit@tu-berlin.de
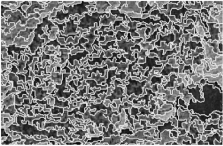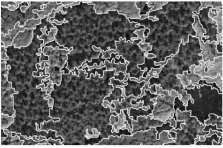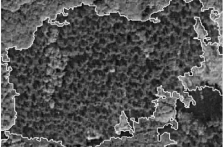
KEYWORDS: QuickBird, fuzzy-logic, NATURA 2000

ABSTRACT: The detection of forest types and habitats is of major importance for silvicultural management as well as for the monitoring of biodiversity in the context of NATURA 2000. For these purposes, the presented study applies an object-based classification method using VHR QuickBird data at a test site in the pre-alpine area of Bavaria (southern Germany). Additional geo-data and derived parameters, such as altitude, aspect, slope, or soil type, are combined with information about forest development and integrated into the classification using a fuzzy knowledge-base. Natural site conditions and silvicultural site conditions are considered in this rule-base.

The results of the presented approach show higher classification accuracy for the classification of forest types using ancillary information than can be reached without additional data. Moreover, for forest types with very distinctly defined ecological niches (e. g. alluvial types of forest), a better characterisation and integration of rules is possible than for habitats with very wide ecological niches. Hence, classification accuracies are significantly higher when these rules are applied. In a second step NATURA 2000 habitat types and selected habitat qualities are derived from the classified forest types. However, the share of habitat qualities varies with an altering scale. This difficulty should be addressed in further research of NATURA 2000 monitoring.

1 Introduction

With the development of a standardised and pan-European geo-data-infrastructure (Craglia et al. 2005), remote-sensing applications which integrate GIS information will become increasingly important. Therefore, various studies of combining additional data and knowledge into classification processes (Maselli et al. 1995; Stolz 1998) were undertaken. However, the integration of additional geo-data into very high spatial resolution (VHSR) imagery remains a challenging task.

Fig. 1 shows an exemplary overview of multi-scale segmentation for different forest scales with the corresponding levels of ancillary data and knowledge. Additional information can be differentiated into two categories (see Fig. 1). Firstly, spatially explicit knowledge is available. For forestry applications a broad range of this kind of data sources can be used, namely the simulation of geo-data (e.g. Disney et al. 2006; Verbeke et al. 2005), the usage of altitude information, especially with LIDAR techniques (e.g. Diedershagen et al. 2004), and the integration of silvicultural maps (e.g. Förster et al. 2005b) as well as soil and hydrology maps into classification procedures.

	Object-Based Information	Ancillary Spatial Data	Ancillary Information (partly used in non-spatial context)
Single Tree / Small Tree Group Level (1)		Samples on Test-Sites Crown Maps Soil samples	Information about Single Trees Growing behaviour of single trees (e.g. shadowing)
Tree Group Patch Level (2)		High Spatial Density Samples National Forest Inventory	Information about Forest Structures Silvicultural pratices
Combined Patch Level (3)		Area-wide available Data* Silvicultural Site Map* Forest Site Map* Digital Terrain Model* Soil Maps*	Information about Forest Development Potential natural vegetation* Land-use history

Data and information types used in the context of this case-study

Fig. 1. Exemplary overview of multi-scale dependence of object-based information and ancillary GIS-Data and knowledge for a forestry application

Complementary knowledge about processes of the forested landscapes is abundantly available and recorded, i.e. individual tree growth by measurement of crown diameter. Information about the land-use history, silvicultural practices (Pretzsch 2002), and the potential natural vegetation (Walentowski et al. 2004) is of high relevance.

Moreover, similar to delineated objects of satellite information (Burnett and Blaschke 2003), ancillary information can be grouped in different scales, which depend on the landscape level the data was assessed or observed. While some information is available on a site-specific scale other knowledge exists only as interpolated data for larger landscape patches.

The objective of this study is to develop a method of integrating different types of geo-data into an object-based classification process. The presented approach combines spectral and textural information of a QuickBird scene with ancillary data for the identification of forest structures and habitats. Because the knowledge of woodland development is often on purpose expressed ambiguously, a fuzzy-logic-based rule set is applied. This example is especially suitable to show chances and challenges of data-integration techniques, because long-term information about silvicultural practices and ecological woodland development in the study area (Bavaria, Germany) are available.

2 Data and Methods

For the presented study the satellite data are processed with a multi-scale segmentation method (Benz et al. 2004) by using an object-oriented approach with the software eCognition (Baatz and Schäpe 2000). The segmentation levels of different resolution are delineated and assigned to hierarchically organised groups of objects, such as forest habitats, crown combinations and crown types of single-tree species.

The segments are classified as different forest types with and without ancillary information and the results subsequently compared. Additional sources of information are combined using a fuzzy knowledge-base (Stolz and Mauser 1996). Since expert knowledge of the test area is partly available as verbal description, which often contains cognitive uncertainties and is imprecise, fuzzy logic represents a possibility to express these vague statements in a mathematical framework as a degree of membership to a fuzzy set (Zadeh 1983).

In order to show the practical applicability of the classification, the results are employed to derive NATURA 2000 forest habitat types and qualities. NATURA 2000 is a European directive designed to protect the most

seriously threatened habitats and species. The NATURA 2000 network consists of more than 20,000 sites and covers almost a fifth of the EU territory. The directive requires a standardised monitoring of the habitat types and a reporting every six years. From the German point of view, standards for mapping NATURA 2000 forest sites have been developed (Burkhardt et al. 2004).

2.1 Study Site

As test area the forested NATURA 2000 site "Angelberger Forst" in the pre-alpine area of Bavaria, Germany, was chosen. It covers approximately 650 ha. The test site was selected because a terrestrial mapping of NATURA 2000 sites had been completed. Moreover, a broad variety of different semi-natural mixed forest types exists.

Relatively large areas of this submontane forest are vegetated with Beech, mainly with two special habitats, named Luzulo-Fagetum (17.9 %) and Asperulo-Fagetum (3.5 %). Periodically moist locations can be covered by Stellario-Carpinetum (1.3 %) but with less frequent occurrence. Very moist habitats, mostly along streams, are also vegetated with Alluvial forests with Alnus and Fraxinus (Alder and Ash; 2.0 %). Additionally, small sites are covered with Larch and Sycamore. However, the influence of silvicultural practices can be estimated a high proportion of coniferous forest (mainly Spruce) in parts of the area.

The natural allocation mostly depends on soil moisture and acidity, but can equally rely on the relief or anthropogenic influences.

2.2 Data

For the presented investigation a QuickBird scene was acquired at 11th August of 2005. The sensors panchromatic band collects data with a 61 cm resolution at nadir while the multispectral (visible and near infrared) ground sampling distance is 2.44 m at nadir (DigitalGlobe 2006).

A selection of geo-data was supplied by the Bavarian State Institute of Forestry (see table 1). A Digital Terrain Model (DTM 5 and DTM 25) is used for relief information. The parameters slope, aspect, curvature, and altitude are derived from this data source. Furthermore, a Conceptual Soil Map is used. This map is generated by the Bavarian State Institute of Geology from available data sources (e.g. geological map) and mapped soil samples. The Conceptual Soil Map is available for large parts of Bavaria. In spite the detailed spatial specification (scale of 1 : 25,000), the attributes for the soil types are often summarised and imprecise (an example of a

class is: "soil complex of gleys and valley sediments"). As a further important data source a Forest Site Map is utilised. This map contains very detailed information about forest soils and is often used for silvicultural purposes.

For training of the classes and validation of the results a Silvicultural Map is used. This map is used for forest inventory and management. It consists of information about main tree types and tree-type mixtures per forest stand. As training sites a set of 527 sample polygons is selected in non-mixed forest stands with a stratified order for occurring tree types. The sites are allocated randomly within the areas of specific forest stands. The chosen samples are confirmed using current (2003) true colour aerial photographs. In a few indistinct cases, samples are terrestrially mapped together with local forest experts. The knowledge base to build up rule sets for potential forest types was available from a previous project in cooperation with the Bavarian State Institute of Forestry (Kleinschmit et al. 2006). These rules are complemented by silvicultural rules attained from local forest rangers and silvicultural literature (Walentowski et al. 2004).

Table 1. List of existing spatial data and derived parameters of the study site

Type of geo-data	Relevant attributes and derived parameter	Scale / Availability
Digital Terrain Model	Slope Aspect Curvature Altitude	DTM 5 and DTM 25 / parts of Germany
Conceptual Soil Map	Soil type	1 : 25,000 / parts of the federal state Bavaria
Forest Site Map	Availability of nutrients Availability of water Soil substrate	1 : 5,000 / digital for state forests
Silvicultural Site Map (for training and validation purposes)	Main tree type per stand Tree mixture per stand Age Usage	1 : 5,000 / digital for state forests
Terrestrial Map of Habitat Types (for validation purposes)	NATURA 2000 type Conservation status	1 : 1,000 to 1 : 5,000 / for large parts of NATURA 2000 areas

True colour aerial pho- tographs (for training)		1 : 12,400

2.3 Segmentation and Classification

After a geometric correction and the pan-sharpening of the original data to a resolution of 0.61 m with the algorithm of Zhang (2002), all spectral bands of the QuickBird scene are segmented at three landscape scales. These levels are named Single Tree / Small Tree Group Level or level 1 (Scale Parameter[1] (SP) 15, shape factor 0.1, compactness 0.5), Tree Group Patch Level or level 2 (SP 40, shape factor 0.1, compactness 0.5), and Combined Patch Level Structure or level 3 (SP 150, shape factor 0.1, compactness 0.5). The segmentation of all levels is mainly (to 90 %) based on the spectral values of the satellite scene because in heterogeneous structured mixed forests the shape of objects can be obscured by neighbouring trees. The scale parameter of level 1 is chosen to define objects of the size of a tree diameter (mean size of objects: 10.5 m²), while level 2 is selected to characterise patches of tree groups (mean size of objects: 27.7 m²). The scale of level 3 is generated to delineate forest habitats similar in size to terrestrially mapped NATURA 2000 areas (mean size of objects: 103.7 m²).

Previous to the forest classification non-forest land uses, such as agriculture or urban area are masked, based on thresholds for shape, texture and spectral mean value of these classes. The classification of the forest types Beech (two age classes), Black Alder, Larch, Spruce (two age classes), Sycamore, and the forest-related objects Clearances and Afforestation is performed on single tree / small tree-group level (1) as nearest neighbour classification of the mean spectral values of the segments. A 30 per cent share of different tree types is defined as mixed forest within an object. Therefore, the results of level 1 were aggregated to the tree-group patch level (2), where an object is assigned to a single species class if 70 per cent of the sub-objects are classified by one species. Mixed stands are assigned to a newly introduced group "Mixed deciduous" and "Mixed". The third level (Combined Patch Level) is used to improve the classifications of the sub-levels and to derive potential NATURA 2000

---

[1] The named specifications are parameters of the multiresolution segmentation method of eCognition. The Scale Parameter determines the maximum allowed heterogeneity within an object. The shape factor defines the textural homogeneity of the objects. Compactness of objects is one of the parameters (the other is smoothness) for the shape factor (Definiens 2006).

habitat types, such as Beech habitats, Alluvial forest habitats, or Stellario-Carpinetum habitats.

Shadowed areas are separately classified using a NDVI-threshold. To reduce the share of segments classified as shadow the class is differentiated. A second NDVI threshold is introduced, indicating shadowed vegetation. Then the shadowed segment is assigned to the class of the neighbouring forest object with the longest shared border.

2.4 Integration of Class Rules via Fuzzy Logic

The occurrence of different forest habitats depends on specific ecological and anthropogenic influences. These conditions allow or prevent species and habitats to exist. In the following these factors are referred to as silvicultural site conditions and natural site conditions. They can be related to geo-factors, which describe and attribute the ecological quality of a specific location (see tab. 1). The probability of assignment for each object (values from 0 to 1) classified with the nearest neighbour (see 2.3) is combined with a fuzzy knowledge-base, which consists of silvicultural and natural site conditions.

According to the fuzzy logic theory, a fuzzy set for each class concerning each geo-factor is defined, containing input and output membership function. For each parameter a set of possible verbal descriptions (linguistic terms) such as "very steep" or "flat" for the variable "slope" have to be defined and formalised by fuzzy input membership functions. Subsequently, linguistic variables for all possible forest type classes are defined via membership output functions (e.g. "possible occurrence" or "limited occurrence" for the class "Sycamore"). Furthermore, fuzzy rules have been developed describing the relationship between each linguistic term of each variable and the degree of possibility of the class. As a result of this process, defuzzicated membership functions are derived for the geo-factors.
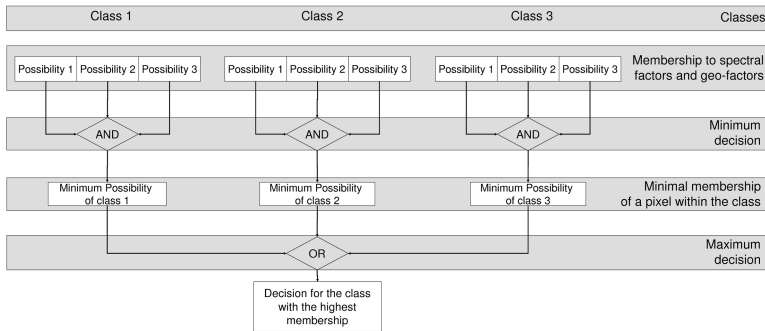
Fig. 2. Schematic application of a fuzzy class decision with spectral classification and geo-factor possibilities

In combining the fuzzy sets and the hierarchical classification results the approach uses the minimum (AND-) rule, which specifies that the most unacceptable factor is the critical value for the forest type to occur. In a next step the minimum possibility of each possible class will be compared. The class with the highest membership will be assigned to the object (maximum – OR – rule, see Fig. 2). Consequently, the lowest membership and the class assignment can be defined solely by a geo-factor.

2.4.1 Natural Site Conditions

For habitat types which can possibly exist in the study area the list of geo-factors is used, consisting of slope, aspect, curvature, and altitude, soil type from a Conceptual Soil Map, and available water, soil substrate, and availability of nutrients from a Forest Site Map (see tab. 1). For all main tree types the natural site conditions are developed based on knowledge of local experts and literature about natural forest associations in Bavaria (Walentowski et al. 2004; Walentowski et al. 2005). Especially for this kind of information the integration via fuzzy logic is useful, because there are often no sharp thresholds. The statements of local experts and literature sources are made in linguistic terms, such as: "*Sycamore can be found in higher pre-alpine regions, especially at very steep slopes*".

As schematically shown in Fig. 2, each classified object consists of at least three different classification outputs. Normally, the class with the highest probability will be assigned to the object. The extended fuzzy-classification includes the possibility of occurrence of forest types within certain natural conditions into this classification process. An example for this classification procedure is an object which is classified as Beech "with a high probability". However, there is still a smaller probability to be as-

signed to Alder or Sycamore. These probabilities are combined with the natural site conditions. If the slope of the object is very steep, it is very unlikely for Beech and Alder to exist with these natural site conditions. In contrast, Sycamore is adapted to these factors and has a high possibility to exist. Of all compared classification probabilities and possibilities of natural site conditions, the lowest value (on a scale from 0 to 1) is assigned to the possibility of existence for the class, because the most unacceptable factor is the critical value for the forest type to occur. In the case of a very steep object it is likely that the possibility of existence for Beech forest is lower after including ancillary data than the spectral probability of Sycamore and therefore the latter forest type is assigned to the object. Because the natural site conditions give very explicit information about species with narrow ecological niches, which cannot be distinguished by spectral values, these forest types are better recognisable with ancillary data.

2.4.2 Silvicultural Site Conditions

To include forestry practices, two approaches are used. Firstly, silvicultural preferred mixture types of forest stands in Germany were taken from literature (Jansen et al. 2002) and extracted from Silvicultural Maps (e.g. 65 % Spruce, 25 % Beech, 5 % Sycamore, 5 % Birch). Statistics of classified tree-type compositions are taken at segmentation level 3 (e.g. 60 % Spruce; 30 % Beech; 10 % Beech - young). If the dominant species and the second dominant species of such a classified mixture type differ not more than 5 % to a silvicultural preferred mixture type, the sub-level (level 2) included these possibilities for the tree species of the mixture as one ancillary layer (e.g. possibility Spruce = 0.6, possibility Beech = 0.3) and included it in the minimum (AND) decision.

Another approach was undertaken to improve the classification accuracy of Elder Spruce (from 120 years) stands. In classification level 1 small clearances are classified with the nearest neighbour approach. The existence of clearances is used in level 2. If an amount of more than 10 % of the classes clearance and old Spruce was detected in the sub-object, the possibility to assign the class to "old Spruce". Vice versa, if "old Spruce" is detected in level 2, the possibility of clearances in level 1 rises.

3 Results

For validation 121 objects of level 1, which cover 1.5 ha (0.23 %) of the study site, are chosen in a random-stratified order of occurring classes. These segments are compared to the Silvicultural Map. This is done on

pixel basis (0.6 m resolution), because the classified object can possibly intersect with validation polygons. The Silvicultural Map is additionally compared to recent (2003) true colour aerial photographs, in case of occurring errors in this map. Level 2 includes two additional summarised classes ("Mixed" and "Deciduous Mixed"), which cover 12 % of the area. Because the process of taking samples is carried out in level 1, these classes are assessed in a separate evaluation realised by visually interpreting aerial photographs.

Firstly, in this chapter the results of the forest type classification of level 1 and level 2 are presented (3.1). In a second step it is shown that NATURA 2000 habitat types and the habitat qualities can be derived from the classified forest types (level 3).

3.1 Classification of Forest Types

The results of the accuracy assessment are shown in table 2. Significantly higher classification accuracies can be reached with instead of without additional data. Especially the detection of species with a small ecological niche is improved. With a multispectral-based classification, a forest type such as Black Alder is not distinguishable from other deciduous forest while showing the highest classification accuracy with ancillary data. This is due to the influence of the natural site conditions, especially the geo-data and rules for the available water from the Forest Site Map and the curvature derived by the DTM. Other decisive factors can be the substrate (Larch) and the slope (Sycamore). Between level 1 and level 2 there is no advance in classification accuracy due to the introduction of mixture classes. As shown in the separate evaluation of these classes in table 2, a better overall classification result could be expected, if level 2 were assessed collectively for all classes, especially for mixed forest.

Table 2. Accuracy assessment for forest types (level 1 and level 2)

Forest Type	Multispectral Classifca-tion with Ancillary Data		Purely Multispectral-Based Classification	
	Level 1	Level 2	Level 1	Level 2
Beech	0.81	0.68	0.75	0.76
Beech – young	0.32	0.14	0.15	0.14
Spruce	0.74	0.97	0.74	0.81

Spruce – old	0.42	0.65	0.32	0.61
Black Alder	0.98	0.89	0.17	0.69
Afforestation	0.97	0.85	0.95	0.82
Larch	0.96	0.68	0.59	0.68
Sycamore	0.88	0.72	0.68	0.70
Overall Accuracy	0.77	0.75	0.64	0.70
Mixed Forest	-	0.90	-	0.85
Deciduous Mixed F.	-	0.55	-	0.45

Between the multispectral classification and the classification with an-cillary information 18.6 % of the objects changed the class assignment in level 1. The decisive possibility of class assignment changed for 36.4 % of the objects when additional data are applied. Therefore, it can be assumed that for the classification with ancillary data an amount between 18.6 % and 36.4 % of the objects are based on the factors of the natural site condi-tions. However, the interaction of classification results and the rule-base are complex as first tests of evaluating the influence of single natural site conditions have shown.

A further improvement could be made without differentiation of age levels (e.g. Beech – young). Natural site conditions are not suited for sepa-ration of the same species. Therefore, rule sets for silvicultural site condi-tions and age classes, similar to the usage of clearances, could be used.

3.2 Derivation of NATURA 2000 Habitat Types

For the classification levels 2 and 3 the attempt to obtain and assess NATURA 2000 habitat types and their qualities was undertaken. At the moment, this information is terrestrial mapped by field survey and com-bined to forestry management plans. For mapping forest habitat types, ob-jective mapping guides with defined rules are available in Germany (Burkhardt et al. 2004). Within these rules parameters of habitat structures, such as number of forest development phases, number of biotope trees per ha, number of dead wood per ha, or percentage of natural forest types are available for different habitat qualities. For the percentage of natural forest types the habitats are identified for:

- excellent quality (A) ≥ 90 % natural forest types,

- good quality (B) ≥ 80 % natural forest types, and
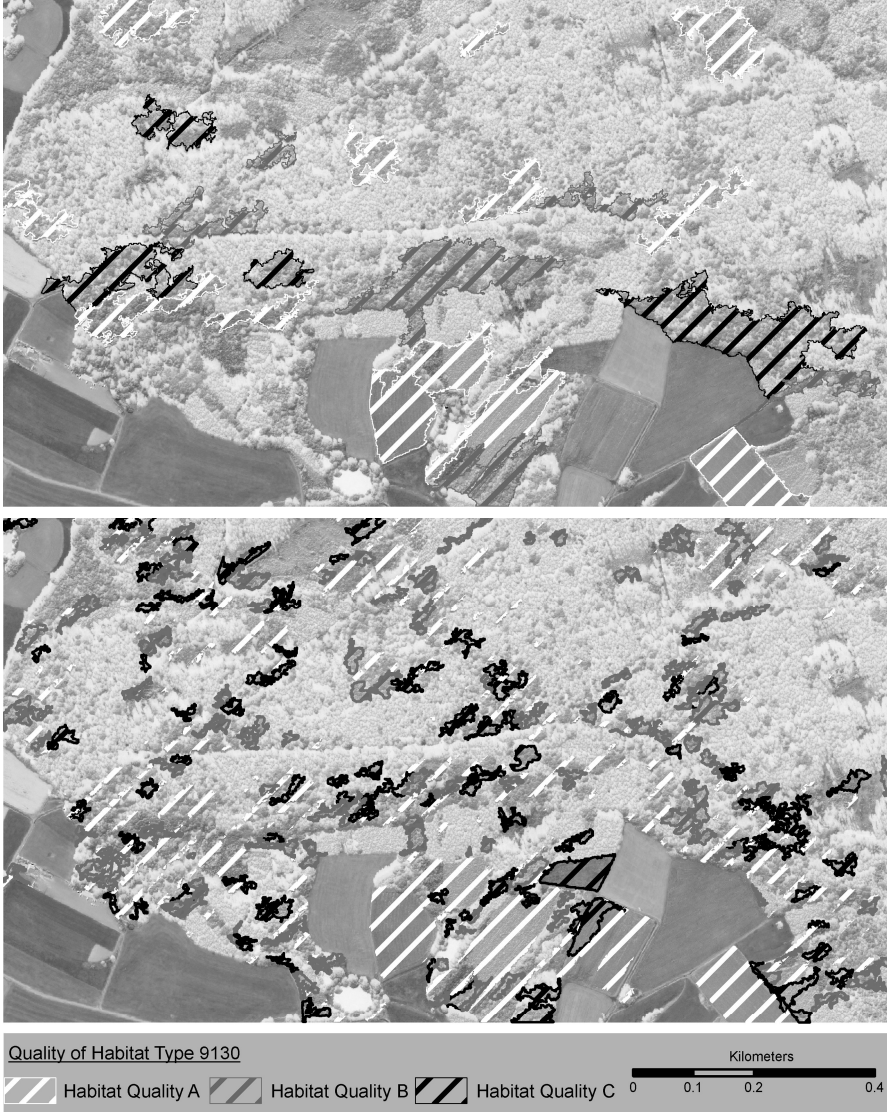
- medium quality (C) ≥ 70 % natural forest types.

Fig. 3. Derivation of different habitat type qualities for level 3 (upper figure) and level 2 (lower figure) for the special Beech habitat type Asperulo-Fagetum (NATURA 2000 nomenclature: 9130)

Fig. 3 shows the result of this approach exemplary for one type of Beech forest. Unfortunately it is only possible to a limited degree to derive habitat types automatically, due to parameters which cannot be detected by remote sensing (e.g. understorey vegetation), local knowledge of mapping, and political decisions (Förster et al. 2005a).

Moreover, the so-called orchard problem (Lang and Langanke 2006) arises. Because of different segmentation scales, altered mixtures of areas and qualities for the habitats are available. Smaller scale parameters tend to define more fragmented areas of a good and excellent quality, while with a scale parameter of 150 larger and coherent areas of good and medium quality will occur.

Both the results and the segmentation levels shown in fig. 3 have certain advantages and disadvantages, but the approach shows that the mapping guide for German forest disregards the need for a spatial reference unit such as landscape levels or a minimum mapping unit. An analysis of the correlation between terrestrial mapped habitat types and different landscape levels could be helpful to formulate a more exact definition of habitat-type quality.

4 Discussion and Outlook

The presented study shows that the classification accuracy of the investigated forest types is higher with ancillary information integrated by fuzzy logic than with a pure multispectral classification. In comparing the results with and without additional data an average increase of 13 % of the overall accuracy is detected. Especially the identification of forest types with narrow ecological niches, such as Black Alder, is significantly increased. However, forest types with wide ecological niches, such as Spruce or Beech, are classified with similar result (difference of not more than 10 %), when comparing the methods. Moreover, silvicultural site conditions are integrated into the classification process. However, two tested approaches have only a limited influence on the classification success (less than 1 % of the objects are classified differently). It can therefore be stated that natural site conditions are more relevant for the classification success than silvicultural site conditions.

In a second step, NATURA 2000 habitat types and the habitat quality "share of natural forest types" are derived from the forest-type classification. It is shown that the share of habitat qualities varies with an altering scale. For the two segmentation scales used in this study (level 2 and level 3) the fine scale defines smaller objects of a higher habitat quality,

while the coarse scale defines larger areas of a medium habitat quality. At present there is no standard defining a spatial reference size (e.g. minimum mapping units) for the quality of biodiversity. This question should be addressed to ecologists and included in mapping guidelines. If a certain habitat requires a coherent large area a larger segmentation scale should be applied, while small-sized habitats should be classified with a finer object size.

Performing a classification using additional GIS-data provokes the question for consistent availability of these data. Within the nationally owned forests of Germany a good data basis already exists, especially with the information from the Forest Site Map. However, in private forests in Germany, geo-data is often not even available as analogous map.

A further improvement of this study can be obtained by a careful analysis of the dependence of accuracy on natural site conditions, integration of other additional data (such as LIDAR data), and a more efficient usage of silvicultural site conditions, e.g. by incorporating stand-density data. The presented method has to be transferred to forest types and habitat types of other areas, such as north-east Germany, to validate the reliability of the technique and become more generally applicable. Moreover, a comparison of different techniques of integrating geo-data into classifications, such as neural networks, could be useful for a quality assessment of integration techniques (Baltsavias 2004).

References

Baatz M, Schäpe A (2000) Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Blaschke T, Griesebner G (eds) AGIT – Salzburg. Wichmann, Heidelberg, pp 12-23

Baltsavias E P (2004) Object extraction and revision by image analysis using existing geodata and knowledge: current status and Stepps towards operational systems. ISPRS Journal of Photogrammetry & Remote Sensing 58: 129-151.

Benz U, Hofmann P, Willhauck G, Lingenfelder I, Heynen M, (2004) Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing 58:239-258

Burkhardt R, Robisch F, Schröder E (2004) Umsetzung der FFH-Richtlinie im Wald - Gemeinsame bundesweite Empfehlungen der Länderarbeitsgemeinschaft Naturschutz (LANA) und der Forstchefkonferenz (FCK). Natur und Landschaft 79(7):316-323

Burnett C, Blaschke T (2003) A multi-scale segmentation / object relationship modelling methodology for landscape analysis. Ecological Modelling 168:233-249

Craglia M, Fullerton K, Annoni A (2005) INSPIRE: an Example of Participative Policy-making in Europe Openness, Participation and Transparency. Geoinformatics 9:43-47

Definiens (2006) Definiens Professional 5 Reference Book. Definiens. München

Diedershagen O, Koch B, Winacker H (2004) Automatic Segment and Characterisation of Forest Stand Parameters using Airborne LIDAR data, Multispectral and Fogis Data. Proceedings of the ISPRS working group VIII/2 - Freiburg, pp 208-212

DigitalGlobe (2006) Product information, imaging systems, ordering, and pricing. http://www.digitalglobe.com (March 2006)

Disney M, Lewis P, Saich P (2006) 3D modelling of forest canopy structure for remote sensing simulations in the optical and microwave domains. Remote Sensing of Environment 100:114-132

Förster M, Kleinschmit B, Walentowski H (2005a) Comparison of three modelling approaches of potential natural forest habitats in Bavaria, Germany. Waldökologie Online (2):126-135

Förster M, Kleinschmit B, Walentowski H (2005b) Monitoring NATURA 2000 forest habitats in Bavaria by the use of ASTER, SPOT5 and GIS data – an integrated approach. In: Olsson H (ed) ForestSat. Swedish National Board of Forestry, Borås, pp 21-25

Jansen M, Schulz R, Konitzer A, Sloboda B (2002) GIS based investigations of effects of the LÖWE program in the Harz mountains. In: Jansen M, Judas M, Saborowski J (eds) Spatial Modeling in Forest Ecology and Management Springer, Berlin Heidelberg New York, pp 177-193

Kleinschmit B, Walentowski H, Förster M, Fischer M, Seitz R, Kraft P, Schlutow A, Bock M (2006) Erfassung von Wald-Lebensraumtypen in FFH-Gebieten - Fernerkundung am Taubenberg und im Angelberger Forst. LWF Wissen 51:1-39

Lang S, Langanke T (2006) Object-based mapping and object-relationship modeling for land use classes and habitats. PFG(1):5-18

Maselli F, Conese C, Filippis T, Romani M (1995) Integration of ancillary data into a maximum likelihood classifier with nonparametric priors. Journal of Photogrammetry and Remote Sensing 50(2):2-11

Pretzsch H (2002) Grundlagen der Waldwachstumsforschung. Parey, Berlin

Stolz R (1998) Die Verwendung der Fuzzy Logic Theorie zur wissensbasierten Klassifikation von Fernerkundungsdaten. Münchner Geographische Abhandlungen B 26:1-177

Stolz R, Mauser W (1996) A fuzzy approach for improving landcover classification by integrating remote sensing and GIS data. In: Parlow E (ed) Progress in Environmental Remote Sensing Research and Applications. Balkema, Rotterdam, pp 33-41

Verbeke L, Van Coillie F, De Wulf R (2005) A directional variant of the local maximum filter for stand density estimation from IKONOS imagery. In: Olsson H (ed) ForestSat. Swedish National Board of Forestry, Boras, pp 49-53

Walentowski H, Ewald J, Fischer A, Kölling C, Türk W (2004) Handbuch der natürlichen Waldgesellschaften Bayerns. Geobotanica, Freising

Walentowski H, Fischer M, Seitz R (2005) Fir-dominated forests in Bavaria. waldökologie online(2):18-39

Zadeh LA (1983) The Role of Fuzzy Logic in the Management of Uncertainty in Expert Systems. Fuzzy Sets and Systems 11:199-227

Zhang Y (2002) Problems in the fusion of commercial high-resolution satellite images as well as Landsat 7 images and initial solutions. Int. Arch. Photogram. Remote Sens. 34(4)

Chapter 3.2

Estimation of optimal image object size for the segmentation of forest stands with multispectral IKONOS imagery

M. Kim[1], M. Madden[2], T. Warner[3]

[1] Center for Remote Sensing and Mapping Science (CRMS), Department of Geography, University of Georgia, Athens, GA, U.S.A., mhkim73@uga.edu

[2] CRMS, Department of Geography, University of Georgia, Athens, GA, U.S.A., mmadden@uga.edu

[3] Department of Geology and Geography, West Virginia University, Morgantown, WV, U.S.A., tim.warner@mail.wvu.edu

   ABSTRACT: The determination of segments that represents an optimal image object size is very challenging in object-based image analysis (OBIA). This research employs local variance and spatial autocorrelation to estimate the optimal size of image objects for segmenting forest stands. Segmented images are visually compared to a manually interpreted forest stand database to examine the quality of forest stand segmentation in terms of the average size and number of image objects. Average local variances are then graphed against segmentation scale in an attempt to determine the appropriate scale for optimally derived segments. In addition, an analysis of spatial autocorrelation is performed to investigate how between-object

correlation changes with segmentation scale in terms of over-, optimal, and under-segmentation.

1 Introduction

Conventional pixel-based classification approaches have limitations that should be considered when applied to very high spatial resolution (VHR) imagery (Fisher 1997; Townshend et al. 2000; Ehlers et al. 2003; Brandtberg and Warner 2006). The increased within-class spectral variation of VHR images decreases classification accuracy when used with the traditional pixel-based approaches (Shiewe et al. 2001). Object-based image analysis (OBIA), which became an area of increasing research interest in the late 1990s, is a contextual segmentation and classification approach that may offer an effective method for overcoming some of the limitations inherent to traditional pixel-based classification of VHR images. Particularly, the OBIA can overcome within-class spectral variation inherent to VHR imagery (Yu et al. 2006). In addition, it can be used to emulate a human interpreter's ability in image interpretation (Blaschke and Strobl 2001; Blaschke 2003; Benz et al. 2004; Meinel and Neubert 2004).

Although the OBIA scheme seems to hold promise for solving classification problems associated with VHR imagery, it also has an important related challenge, namely, the estimation of the desired size of image objects that should be obtained in an image segmentation procedure. Unfortunately, there is currently no objective method for deciding the optimal scale of segmentation, so the segmentation process is often highly dependent on trial-and-error methods (Meinel and Neubert 2004). Yet, the size of image objects is one of the most important and critical issues which directly influences the quality of the segmentation, and thus the accuracy of the classification (Blaschke 2003; Dorrren et al. 2003; Meinel and Neubert 2004).

In this paper, we employ a case study that builds on the results of Kim and Madden (2006) to investigate agreement between a manual interpretation and image segmentation at a variety of scales, and the pattern of segment variance and autocorrelation associated with those segmentation scales. Kim and Madden (2006) performed a research to examine the relationship between segmentation scale of general forest types (i.e., deciduous, evergreen, and mixed forests) and classification. For this follow-on study, we examine if an understanding of changes associated with segment variance and autocorrelation might provide image analysts with a method of determining optimal size of image objects in image segmentation for forest stand mapping.

2 Local variance, spatial autocorrelation and image objects associated with forest stand map

A number of previous studies have investigated how image properties change with pixel resolution (Cao and Lam 1997). One common method for understanding how image spatial structure changes with pixel size is the graph of average local variance, used by Woodcock and Strahler (1987). This approach has been used to determine the optimal spatial resolution for identifying thematic information of interest in the context of pixel-based image classification. In Woodcock and Strahler's (1987) approach, the image is degraded to a range of pixel size. Variance in spectral reflectance or brightness is then computed a 3×3 moving window, and then the average for the entire scene is graphed as a function of the associated pixel size. Woodcock and Strahler (1987) found that the local variance was low if the spatial resolution was considerably finer than objects on an image. When the pixel size was approximately one half to three quarters of the size of the objects in the image, the local variance was found to reach a maximum. On the other hand, if the pixel size was larger than the objects in the image, the local variance decreases once again. Thus, the graph of local variance against pixel size is one method that can be helpful for understanding spectral heterogeneity and the scale of the objects in the image.

Building on this idea of linking variance and scale, we hypothesize that the average variance of image objects, graphed as a function of image segment size, may provide insight as to the optimal scale of image objects for image segmentation. We define the optimal scale as one that is not over-segmented, with an excessive number of segments that are on average too small, and also not under-segmented, with too few segments that are on average too large.

This definition of an optimal scale is useful for considering the relationship between image object variance and scale. As the segmentation becomes coarser, each segment will tend to incorporate a wider range of image brightness values. Therefore, a general trend of increasing average variance of the segments is expected with coarser scale (and decreasing number of segments). However, we hypothesize that with mixed forest stands, as the segments become too large (i.e., reach a stage of under-segmentation) each segment will tend to include more pixels from pure forest stands. This inclusion would lower the variance of image objects corresponding to mixed forest stands. Therefore, we suggest that the optimal segmentation actually would occur at the scale just before a flattening of the graph. Our second hypothesis is that the optimal scale generates the

least positive, and potentially even negative, autocorrelation between the average brightness values of the segments. In other words, we assume that an optimal scale most clearly brings out contrasting average brightness values in the segmentation.

This autocorrelation hypothesis draws in part on the concept that ideal image enhancement should maximize autocorrelation at the pixel level (Warner 1999). Likewise, image enhancements (Warner and Shank 1997) and classifications (Warner et al. 1999) can be ranked based on their information content as indicated by autocorrelation at the pixel level. However, for segmentation, we suggest that the optimal pattern is obtained when the adjacent *segments* are the *least* similar in brightness values. In an over-segmented image, we would expect that the adjacent segments are on average somewhat similar, and thus the segments will tend to be autocorrelated. On the other hand, in an under-segmented image, the segments are too large, and lose their spectral homogeneity. In this instance, the average brightness of the adjacent segment will tend to converge on relatively similar mixtures, and once again the autocorrelation of the segments is relatively high. Therefore, we suggest that in a graph of scale versus autocorrelation of the segments, the optimal scale should be indicated by the scale associated with the least autocorrelation between segments.

3 Study area and data sources

The study area for the case research is the Guilford Courthouse National Military Park (GUCO) located in Greensboro, North Carolina, U.S.A. (Fig.1). The 1-km^2 park lies in one of the most rapidly developing portions of the city and provides increasingly important green space for recreational activities and wildlife refuge (Hiatt 2003).

The park was initially mapped by the Center for Remote Sensing and Mapping Science (CRMS), Department of Geography at the University of Georgia, in conjunction with the National Park Service (NPS) and NatureServe, as part of the NPS/U.S. Geological Survey (USGS) National Vegetation Mapping Program (Welch et al. 2002; Madden et al. 2004). The mapping of the park was based on manual interpretation of 1:12,000-scale leaf-on color-infrared (CIR) aerial photographs using the National Vegetation Classification System (NVCS) (Grossmann et al. 1998). Fig.1 illustrates the manually-interpreted association-level forest stands in GUCO.

In our study, we assume that the manually produced GUCO vegetation database/map can be used as a reference for determining the optimal scale of the segmentation. Although the manual map does not necessarily repre-

sent an *objective* optimal scale and classification of forest stands, it represents vegetation communities of the type, level of detail and scale that the resource managers require for management decisions. Furthermore, human interpreters bring to bear extensive local (i.e., field-based data) and expert knowledge of forest stands and associated remote sensing signatures. GUCO was visited several times throughout the course of the project by CRMS and NatureServe botanists to collect plot-level data of overstory and understory tree species, as well as numerous quick plots to identify the NVCS class of observation points geolocated with GPS (Madden et al. 2004). An independent accuracy assessment of the completed association-level vegetation and forest stand database conducted by NatureServe resulted in an overall accuracy of 83 % and a Kappa of 0.81 (NatureServe 2007). Therefore, we can assume that the manually interpreted database represents the best approximation of the optimal scale and classification of GUCO forest stands.

  A multispectral IKONOS image of 4-m pixel resolution acquired on July 6, 2002 by Space Imaging, Inc. (now GeoEye, Inc.) was used for this research with special attention to the near infrared band (NIR) that is crucial to vegetation studies. The image was georeferenced and co-registered to a scanned CIR aerial photograph that was acquired October 20, 2000 at 1:12,000 scale and rectified based on horizontal control from a 1998 USGS Digital Orthophoto Quarter Quadrangle (DOQQ). The study area has a very flat terrain, so the co-registration of the CIR aerial photograph and IKONOS image could be achieved to a root-mean-square-error of ±3 m with this process. The GUCO park includes non-forest areas such as open pasture, roads, cemeteries and homesites. These non-forest areas were masked out before image segmentation of forest stands.

4 Methodology

Multiresolution Segmentation, implemented in Definiens Professional Version 5.0 (formerly eCognition), was utilized for the OBIA. The segmentation is based on a region growing technique which places seed pixels over an entire image and groups surrounding pixels to the local seeds, if they meet specific criteria. The size and homogeneity of image objects are important parameters in the segmentation. The scale parameter, i.e., segmentation scale, controls the average size of the image objects (Definiens 2004). The homogeneity criterion determines the importance or weight of three heterogeneity attributes: color, smoothness and compactness. The color criterion minimizes the standard deviation of spectral values within

image objects and also considers the shape of image objects. The shape criterion itself is comprised of a smoothness criterion defined as how smooth the boundary of each image object is in a segmented image and a compactness criterion minimizing the deviation from the ideal compact form (Definiens 2004).

For this study, a series of segmentations was conducted using the IKONOS image with the segmentation scale varying from 2 to 30 in steps of 1 to examine the segmentation quality of image objects representing more detailed (association-level) forest stands. These values were chosen so as to encompass a range of scales that were observed to extend from over-segmentation to under-segmentation. In addition, the scale step of 1 was made possible by the relatively small size of the park (1 km^2) and therefore minimal time for processing iterations. An additional analysis was conducted to assess the repeatability of segmentation processes. The entire range of segmentation scales from 2 to 30 was repeated 5 times each and the resulting segments were visually compared. This thorough analysis of this size range of segmentation scale was needed to determine the optimal object size for object-based forest stands segmentation.

The color and shape ratios are inversely proportional to each other in the homogeneity criterion. The shape of each image object approaches that of a circle as the shape value approaches a value of 1.0. Therefore, if the segmentation scale is consistent (e.g., 18) and the color to shape ratio decreases, forest segment boundaries became less complex and segment sizes became more uniform (see Fig.2) when compared to forest stands in the manually interpreted CRMS database (see Fig.1). Since a value of 0.9 for this ratio was found to produce a pattern most like that of the manual map, this value was used for the entire subsequent analysis. All segmentation procedures were performed using only brightness values of the four spectral bands.

Segmented images across entire segmentation scales were exported to polygon vector files of ArcView Shapefile format to compute average local variance and Moran's I indices. The standard deviations of image objects' spectral reflectance and an average for all the segments were computed from the brightness values of NIR band. These average variances were then graphed as a function of segmentation scale. Moran's I indices were computed from the mean values of NIR band of the segmented images in order to examine between-object spatial autocorrelation at each scale. The contiguity matrix, used for computing the Moran's I, was calculated from the squared inverse Euclidean distance between segment centroids.

5 Results and discussion

The image segmentation was performed across the range of segmentation scales (2 to 30). This produced a range of sizes and numbers of image objects (Table 3.2.1). The segmentation across the entire range of scales (2 to 30) was repeated 5 times, with no discernable difference in the number and average size of segments, as well as the calculation of local variance and Moran's I autocorrelation.

Table 1. Average sizes and numbers of image object produced by the image segmentation

Scale	Average size [m^2]	No. of image objects	Scale	Average size [m^2]	No. of image objects
2	50	13,182	17	6,207	107
3	125	5,321	18	6,991	96
4	228	2,907	19	8,407	79
5	378	1,756	20	9,488	70
6	565	1,175	21	9,767	68
7	829	801	22	11,451	58
8	1,094	607	23	12,299	55
9	1,466	453	24	13,023	51
10	1,965	339	25	13,283	50
11	2,424	274	26	13,283	50
12	2,913	228	27	15,095	45
13	3,477	191	28	15,445	43
14	3,907	170	29	16,199	41
15	4,677	142	30	16,604	40
16	5,535	120			

The average sizes of image objects across the range of scales were then compared to the average size of association-level forest stands derived from the manual interpretation to identify scales associated with under- and over-segmentation (Fig.3). Since the average size of the CRMS forest stands was 7002 m^2 and the number of forest stands was 94, it can be inferred that the segmentation scale of 18 produced segmentation results most closely resembling manually interpreted forest stands in shape, size and location.

We visually assessed segmentation quality associated with over-, optimal, and under-segmentation in comparison with manually interpreted forest stands. Fig.4a and Fig.4b indicate over-segmentation of association-level forest stands compared with manual interpretation. At a scale of 6, the study area is clearly segmented to excess with over 5 times the desired number of segments. At the other extreme, small stands of pure evergreen forest, indicated by dashed circles on Fig.4a, do not appear in the under-segmented image as shown in Fig.4c. It is important to note, however, that even at the apparently optimal scale of 18 (Fig.4d), some small areas of disagreement do exist between the manual and automated segmentations. Forest stands, shown in the dashed-circle areas in Fig.4d, were divided into several image objects that do not match the manual forest stands.

Two different analyses of object local variance and spatial autocorrelation provided evidence of 18 and 19 being strong candidates for the optimal segmentation scales in this study. The graph of average local variance of image objects over the range of segmentation scale is shown in Fig.5. Average local variance increases from a low value of 2.0 at the minimum scale of 2, and levels off at the scale of 20. As discussed earlier, we anticipated that the under-segmentation would occur at the scale where the graph of local variance began to level off and the optimal scale would actually come just before this point. Therefore, Fig.5 supports a scale of 19 being optimal for forest objects. Additional support for determining the optimal segmentation scale comes from Table 3.2.1. This table indicates the number and average size of image objects is most similar to those of the manually produced forest stands map at the segmentation scale of 18. The graph of spatial autocorrelation, as calculated by Moran's I, versus segmentation scale also shows a distinct overall bowl shape with minima at scale of 14 and 18 (Fig.6). The lowest minimum at scale 18 coincides with the scale that produced the segmentation most similar to the manually produced map. It should also be noted that between the scale of 14 and 21, the autocorrelation is negative, indicating dissimilarity between adjacent values. The graph confirms the expectation that excessive over- and under-segmentation is associated with positive autocorrelation, while an optimal segmentation should be associated with the lowest autocorrelation.

6 Summary and conclusion

In summary, object-based forest stand segmentation was performed using a 4-m multispectral IKONOS image. Visual comparison was made between a vegetation database compiled by manual interpretation and seg-

mented images to examine segmentation quality associated with over-, optimal, and under- segmentation. Local variance and spatial autocorrelation were utilized in an attempt to estimate the optimal scales related to sizes of image objects for forest stand segmentation.

Average local variance was graphed as a function of segmentation scales. The average variance was found to increase with the magnitude of the segmentation scale, leveling off at the scale of 20 and therefore, indicating an optimal scale of 19. We also expected that the optimal segmentation scale would result in image objects similar in number and size to those of the manual forest stands before the scale where the graph leveled off. This expectation was validated at a scale of 18 which produced the number/size of image objects closet to those of the manual interpretation. The average size and number of image objects at the scale of 18 (6,991 m^2 and 96, respectively) were very close to those of manually interpreted association-level forest stands in the park (7,002 m^2 and 94, respectively).

The analysis of spatial autocorrelation indicated that there was high positive correlation between segmentation scales and Moran's I indices calculated for image objects with excessive over- and under-segmentation. In contrast, between-object autocorrelation was lowest, and indeed negative, when the scale was 18. This supports the scale at which the average size and number of segmented image objects were similar to those of manually interpreted forest stands.

In conclusion, three types of analyses (i.e., number/average size of objects, local variance, and spatial autocorrelation) all confirmed that segmentation scales of 18 to 19 are optimal for obtaining segmented image objects that most closely resemble those of manual interpretation of forest stands by vegetation experts. Although the analyses did not agree on the exact same optimal segmentation scale, they narrowed the wide range of possible segmentation scales to 18 or 19. Indeed, users often do not know the order of magnitude to begin for the determination of segmentation scale (e.g., 10 or 100). By 1) comparing segmentation results to objects in a dataset of known accuracy completeness, and 2) analyzing measures of image variance heterogeneity and spatial autocorrelation vs. segmentation scale, it is now possible to propose an image analysis methodology that may be useful for identifying optimal segmentation scales. Researchers, for example, could perform a rough cut of segmentation at a few scales between a wide range of (e.g., 5, 10, 15, 20) and then graph local variance and Moran's I. The shapes of these graphs will reveal scale ranges most likely to be associated with optimal image object sizes. The local variance graph will level off and Moran's I will dip to negative values. OBIA researchers can then target specific scales (e.g., 10-15) and avoid wasting time for segmentation at non-optimal scales. It is hoped that these results

will lead to more automated procedures of segmentation for the extraction of high quality features from very high resolution digital images.

Acknowledgements

References

Benz, UC, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing 58:239-258

Blaschke T (2003) Object-based contextual image classification built on image segmentation. Advances in Techniques for Analysis of Remotely Sensed Data 2003 IEEE Workshop:113-119

Blaschke T, Strobl J (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS-Zeitschrift für Geoinformationssysteme 6:12-17

Brandtberg T, Warner T (2006) High resolution remote sensing. In: G. Shao and K. M. Reynolds (eds) Computer Applications in Sustainable Forest Managment, Springer Verlag, Dordrecht, Netherlands, pp 19-41

Cao C, Lam NS (1997) Understanding the scale and resolution effects in remote sensing and GIS. In: D.A. Quattrochi and M.F. Goodchild (eds) Scale in Remote Sensing and GIS, CRC Press Inc., U.S., pp 57-72

Definiens (2004) eCognition User Guide 4, Definiens AG, Germany

Dorren LKA, Maier B, Seijmonsbergen AC (2003) Improved Landsat-based forest mapping in steep mountainous areas using object-based classification. Forest Ecology and Management 183:31-46

Ehlers M, Gähler M, Janowsky R (2003) Automated analysis of ultra high resolution remote sensing data for biotope type mapping: new possibilities and challenges. ISPRS Journal of Photogrammetry and Remote Sensing. 57:315-326

Fisher P (1997) The pixel: a snare and a delusion. International Journal of Remote Sensing 18(3):679-685

Grossman DH, Faber-Langendoen D, Weakley AS, Anderson M, Bourgeron P, Crawford R, Goodin K, Landaal S, Metzler K, Patterson KD, Payne M, Reid M, Sneddon L (1998) International Classification of Ecological Communities: Terrestrial Vegetation of the United States. vol I. The National Vegetation Classification System: Development, Status and Applications. The Nature Conservancy, Arlington, Virginia, p 126

Hiatt J (2003) Guilford Courthouse National Military Park: Cultural Landscape Report. National Park Service, Southeast Regional Office, Cultural Resources Division, http://www.nps.gov/guco/pphtml/documents.html, Accessed last on March 22, 2006

Kim M, Madden M (2006) Determination of optimal scale parameter for alliance-level forest classification of multispectral IKONOS image. Commission IV, WG IV/4 on Proceeding of 1st OBIA Conference, Salzburg, Austria

Madden M, Welch R, Jordan T, Jackson P, Seavey R and Seavey J (2004) Digital Vegetation Maps for the Great Smoky Mountains National Park, Final Report to the U.S. Dept. of Interior, National Park Service, Cooperative Agreement Number 1443-CA-5460-98-019, Center for Remote Sensing and Mapping Science, The University of Georgia, Athens, Georgia, p 112

Meinel G, Neubert M (2004) A comparison of segmentation programs for high resolution remote sensing data. Commission VI on Proceeding of 20th ISPRS Congress in Istanbul, www.isprs.org/istanbul2004/comm4/papers/506.pdf, Accessed on October 16, 2005

NatureServe (2007) Accuracy Assessment: Guilford Courthouse National Military Park, Final Report to the National Park Service, Durham, North Carolina

Shiewe J, Tufte L, Ehlers E (2001) Potential and problems of multi-scale segmentation methods in remote sensing. GIS-Zeitschrift für Geoinformationssysteme 6:34-39

Townshend JRG, Huang C, Kalluri SNV, Defries RS, Liang S, Yang K (2000) Beware of per-pixel characterization of land cover. International Journal of Remote Sensing 21(4):839-843

Warner T (1999) Analysis of spatial pattern in remotely sensed data using multivariate spatial autocorrelation. Geocarto International 14(1):59-65

Warner T, Shank M (1997) Spatial autocorrelation analysis of hyperspectral imagery for feature selection. Remote Sensing of Environment 60:58-70

Warner T, Steinmaus K, Foote H (1999) An evaluation of spatial autocorrelation-based feature selection. International Journal of Remote Sensing 20(8):1601-1616

Welch R, Madden M, Jordan T (2002) Photogrammetric and GIS techniques for the development of vegetation databases of mountainous areas: Great Smoky Mountains National Park. ISPRS Journal of Photogrammetry and Remote Sensing 57(1-2): 53-68

Woodcock CE, Strahler AH (1987) The factor of scale in remote sensing. Remote Sensing of Environment 21:311-332

Yu QP, Gong N, Clinton G, Biging MK, Shirokauer D (2006) Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. Photogrammetic Engineering and Remote Sensing 72(7):799-811
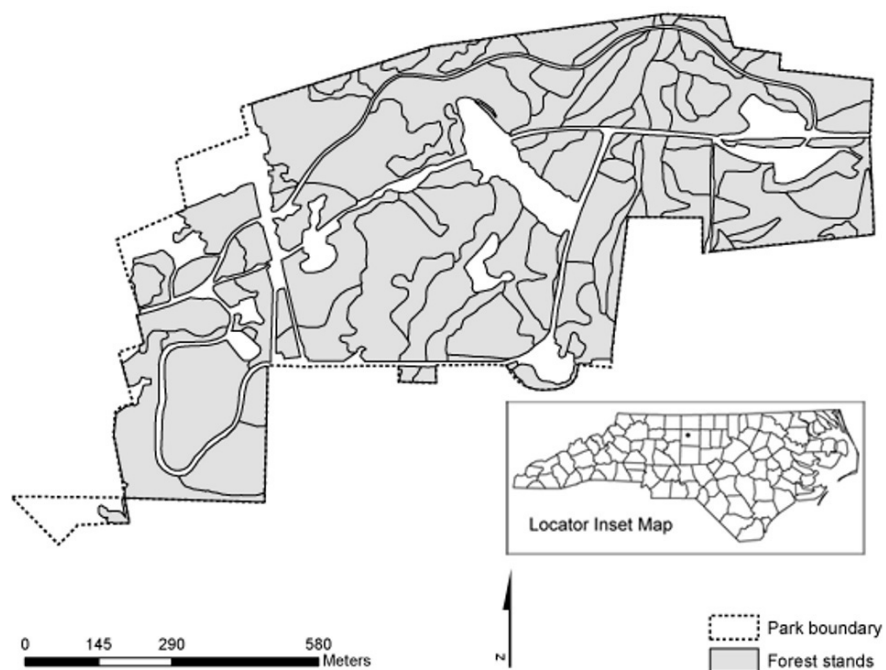
Fig. 1. Forest stands of GUCO park from University of Georgia CRMS-NPS vegetation database produced by the manual interpretation of aerial photographs
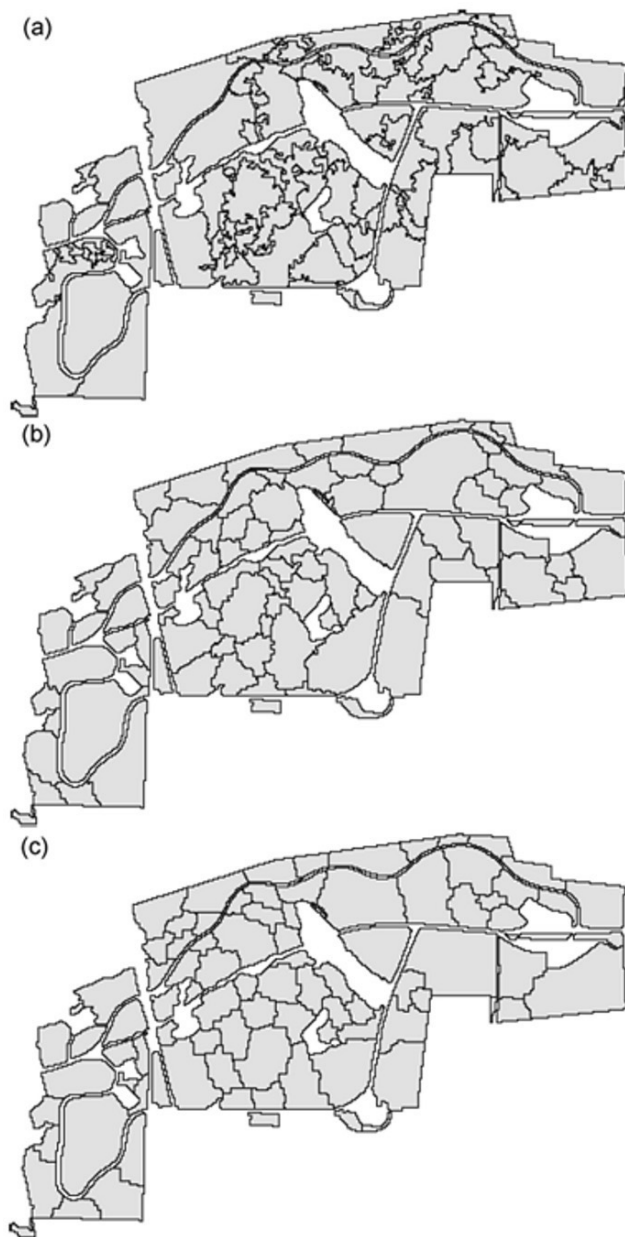
Fig. 2. Different shapes of image objects with a constant segmentation scale of 18, but decreasing ratios of color to shape. Image objects derived from the color criterion of 0.9 (a), 0.5 (b) and 0.1 (c)
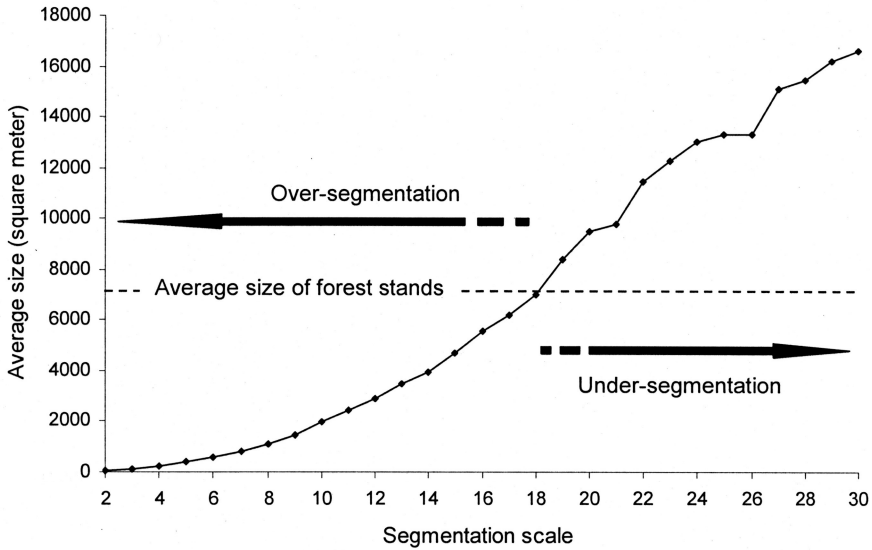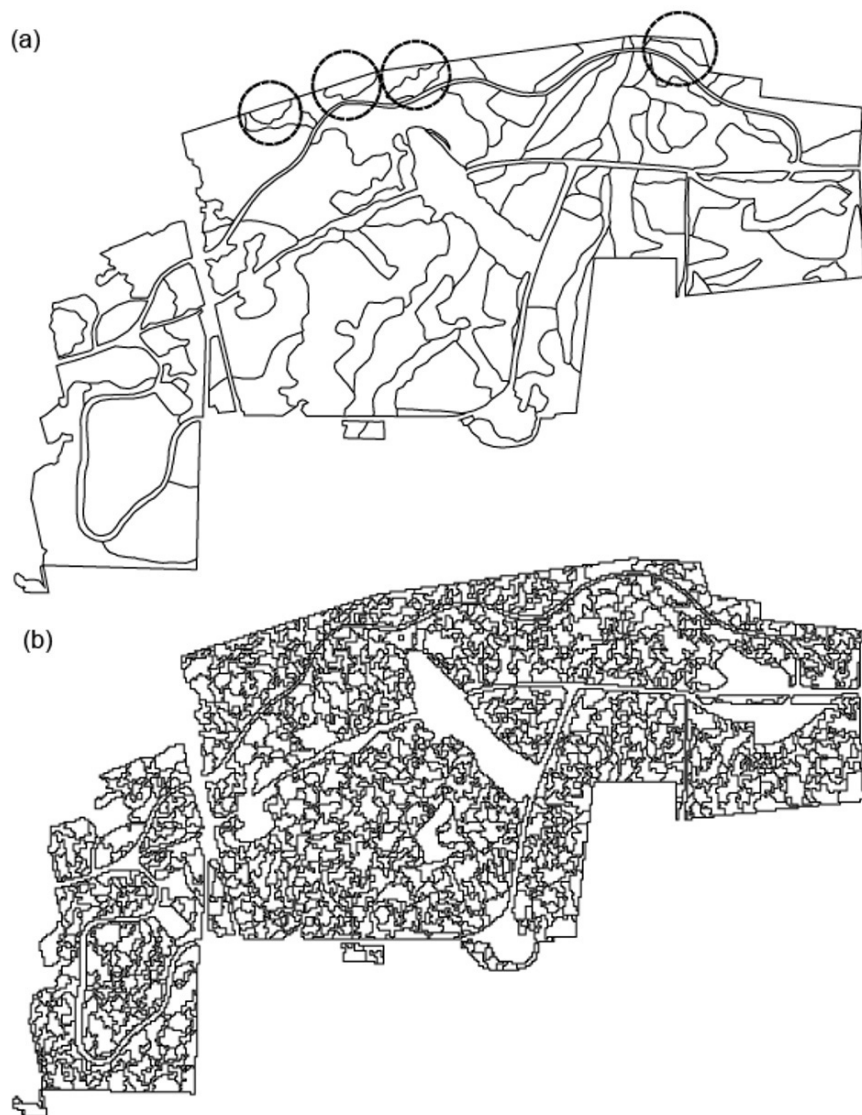
Fig. 3. Average sized image objects resulting from a range of segmentation scales to determine under- and over-segmentation of image objects in comparison to the manually produced map (7002 m^2)

(a)

(b)
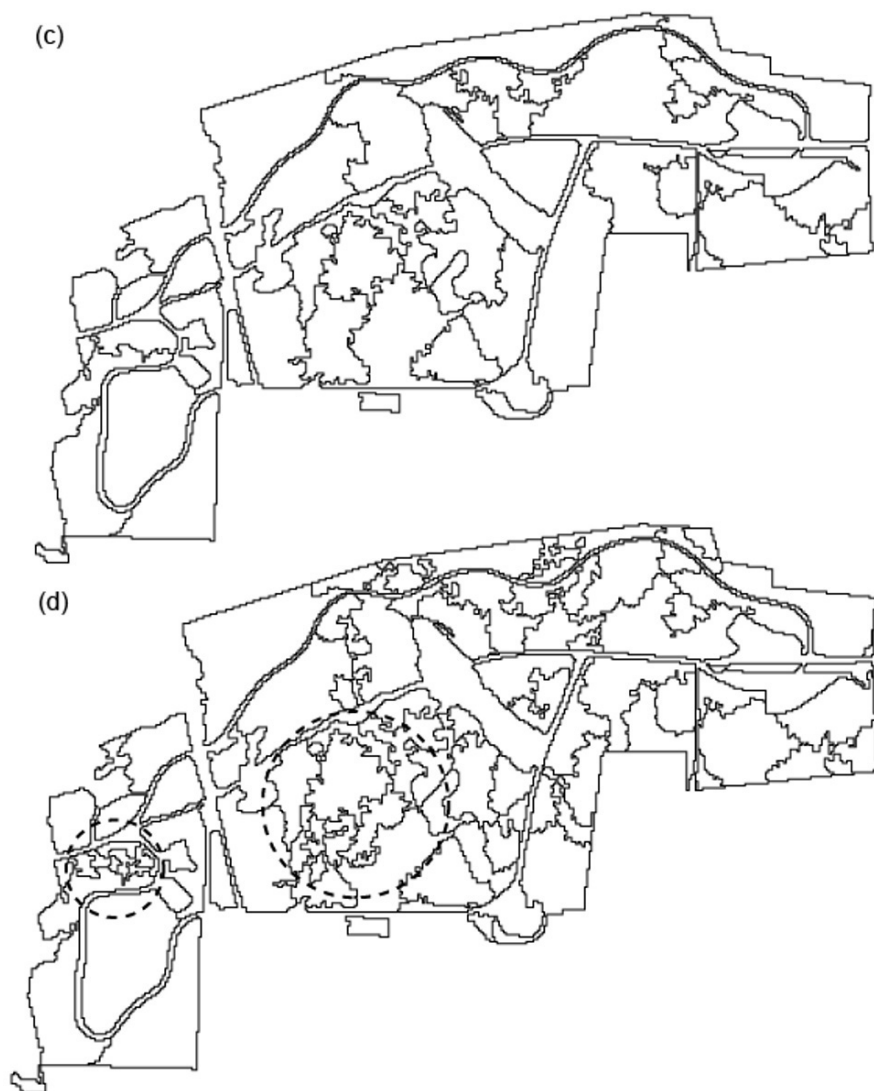
Fig. 4. Segmentation quality compared with manual interpretation: (a) manually produced vegetation associations map of GUCO, (b) over-segmentation at the scale of 6, (c) under-segmentation at the scale of 25, (d) optimal segmentation at the scale of 18
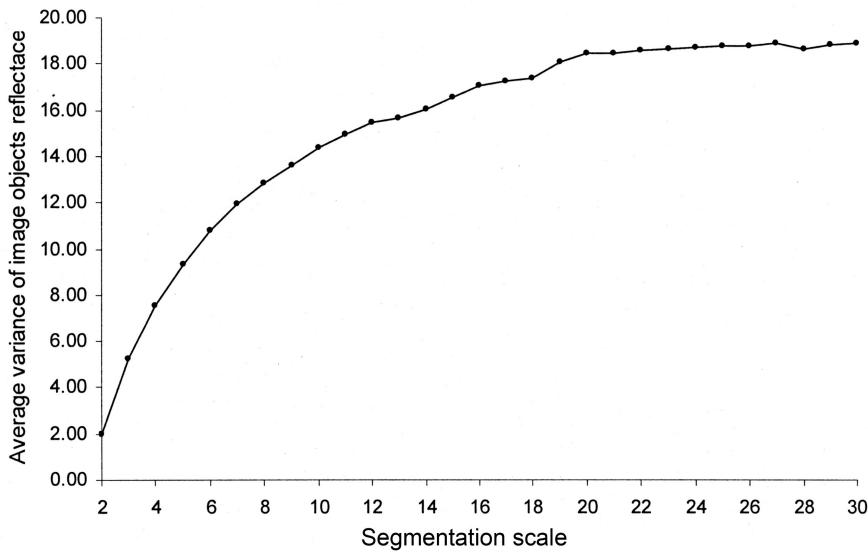
Fig. 5. Graph of local variance as a function of segmentation scale
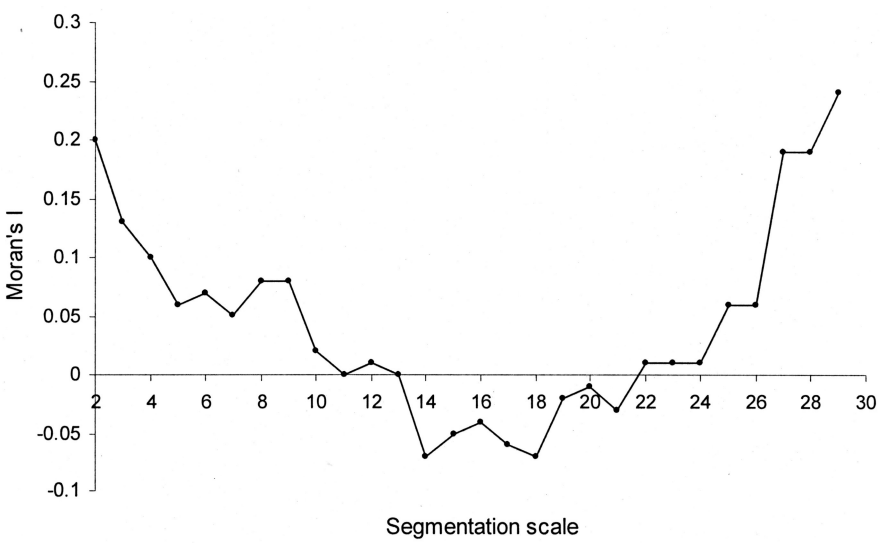
Fig. 6. Moran's I indices as a function of segmentation scale

Chapter 3.3

An object-based approach for the implementation of forest legislation in Greece using very high resolution satellite data

G. Mallinis, D. Karamanolis, M. Karteris, I. Gitas

School of Forestry and Natural Environment, Aristotle University of Thessaloniki, Greece, gmallin@for.auth.gr

ABSTRACT: The possibility to extract forest areas according to the criteria included in a legal forest definition was evaluated, in a mountainous area in the Northern-central part of Greece, following an object-based image analysis approach. While a lot of studies have focused on the estimation of forest cover at regional scale, no particular emphasis has been given so far to the delineation of forest boundary line at local scale.

The study area presents heterogeneity and it is occupied by deciduous and evergreen forest species, shrublands and grasslands. One level of fine scale objects was generated through the Fractal Net Evolution Approach (FNEA) from Quickbird data. Logistic regression statistical analysis was used to predict the existence or not of tree canopy for each image object. The classified objects were subject to a second classification process using class and hierarchy related information to quantify the criteria of the Greek Forest law. In addition, the usefulness of a fusion procedure of the multispectral with the panchromatic component of the Quickbird image was evaluated for the delineation of forest cover maps.

Logistic regression classification of the original multispectral image proved to be the best method in terms of absolute accuracy reaching around 85% but the comparison of the accuracy results based on the Z statistic indicated that the difference between the original and the fused image was non-significant.

Overall the object-based classification approach followed in our study, seems to be promising in order to discriminate in a more operational manner and with decreased subjectivity the extent of forest areas according to forest legal definitions.

1 Introduction

In the recent decades, an appreciation of the multi-functional role that forest ecosystems can serve raised the need for their protection, both from global organizations and national authorities. These goals can be achieved by introducing and enforcing measures and legislation rules which will at least ensure the conservation of the status and the area extent of forest resources. Therefore, there is more than ever a need to obtain information on various scales, regarding forest ecosystems to guide the decisions of policy makers (Franklin and Wulder 2003).

The relatively recent availability of data coming from very high resolution (VHR) spatial sensors (Tanaka 2001) has enabled the extraction of detailed information about forest ecosystems, which is helpful for local and national authorities and managers. Most of all, accurate knowledge of area designated as forest within each nation, apart from reporting to international conventions and mechanisms, is necessary in order to support environmental and development planning of its own and mapping of forest inventory measures and forest biophysical parameters.

Several research efforts have been made regarding the compliance of forest cover maps, mostly at regional scale, using low or medium resolution imagery. Kennedy and Bertolo (2002) using low resolution satellite data derived forest-non forest maps of the European Union and Haapanen et al. (2004) used a kNN classifier forest/nonforest stratification in developing forest area estimates for the USDA Forest Service's Forest Inventory and Analysis (FIA). McRoberts (2006) developed a logistic regression model using forest inventory plot data and Landsat TM imagery, to predict the probability of forest for several study areas in USA.

Trias-Sanz and Boldo (2005), developed a method for discriminating forest areas in high resolution aerial images, using a single level of regions and classifiers based on distances on probability distributions and whole-region probabilities. Also, Thiel et al. (2006), generated forest cover maps using radar remote sensing data and one level of segments produced with the commercial software eCognition.

However, delineating the extent of forest areas at local scales using remote sensing imagery is a more complex task, because most of the legal

definitions of forest worldwide usually include thresholds for such things as minimum area, percentage of tree canopy cover, tree height, and strip width, and it will specify any exclusion (Lund 2002). Therefore, the incorporation of criteria related to form, the size and the context of forest areas in the classification process, is necessary for a result to be consistent with legislation specifications.

The Greek forest law in particular, contains certain criteria for the characterization of land either as "forest" or "woodland". These criteria concern the canopy cover, the vertical structure of the forest canopy and the shape of the area.

Specifically, in order for an area to be initially characterized as "forest" or "woodland", its size should exceed 0.3 hectares or if not, it should be at least 30 meters wide or in close interdependence and interaction with other "forest" areas. Furthermore, the forest canopy cover has to exceed 30 percent of the area.

A further distinction between the two categories is based on criteria related to the existence of strata of forest canopy and the canopy cover of each layer. In general, woodlands are characterised mainly by areas covered with a discontinuous tree layer, shrubs and herbaceous vegetation.

So far, this discrimination has been made on the basis of manual photo interpretation of aerial photographs on a scale of 1:30000. However, visual interpretation presents deficiencies because it is affected by the skills of the interpreter based on their education, training, perceptual ability and experience (Hoffman et al. 2001). Also, manual delineation and interpretation of forest entities is considered as a labor-intensive and costly process (Wulder et al. in press).

Therefore, there is an obvious need to introduce new, more objective approaches in order to delineate forest areas, especially in land under pressure for residential expansion.

Object-based analysis of remotely sensed data which has been successfully applied in the past (Woodcock and Harward 1992, Gitas et al. 2004, Hay et al. 2005, Chubey et al. 2006) for forest characterization and mapping, could be a viable solution for forest cover maps generation at local scales.

Moreover, Kosaka et al. (2005), indicated that object-based analysis of a pan-sharpened Quickbird image (despite the already rich spatial information available in the multispectral component of a VHR satellite imagery) for forest types classification, raised considerably the overall as well as the individual class accuracies compared with the use of the original multispectral image. Image fusion has been extensively used in remote sensing studies related to vegetation, land-use and urban studies (Ranchin and Wald 2000). There is a number of potential benefits resulting from the fu-

sion procedure (Pohl and Genderen 1998), one of these might be the improvement in the classification accuracy, both in case of very high (López-Galoca et al. 2004) or medium spatial resolution data (Lu and Weng 2005).

The aim of this work was to develop a semi-operational approach, in order to accurately discriminate forest from natural non-forest areas (woodlands, shrublands and grasslands) following legal definition criteria. The study was part of a research programme which aimed to address the operational application of remote sensing in order to develop a National Forest GeoDatabase.

The specific objectives of the study were:

− To develop a methodology for discriminating forest areas on a basis of legal definition, with the use of very high spatial resolution Quickbird data.
− To assess the utility of a fusion procedure in order to improve the classification result.

2 Study area

The 280 hectares study site is part of the Aristotle University's Forest in Northern Greece (Fig. 1). The altitude of the University Forest ranges from 320 to 1200 meters. The most abundant forest species are oak (*Quercus conferta*), pine (*Pinus nigra*) and beech (*Fagus moesiaca*).

Fig. 1. Location of the study area

The study site was selected because it presents the maximum variability in terms of the anaglyph as well as in terms of species composition and stand mixture conditions. Closed forest stands, intermingled with patches of shrublands and grasslands, indicate a significantly fragmented landscape.

3 Materials and methodology

The overall methodology followed in the study is presented in Fig. 2.
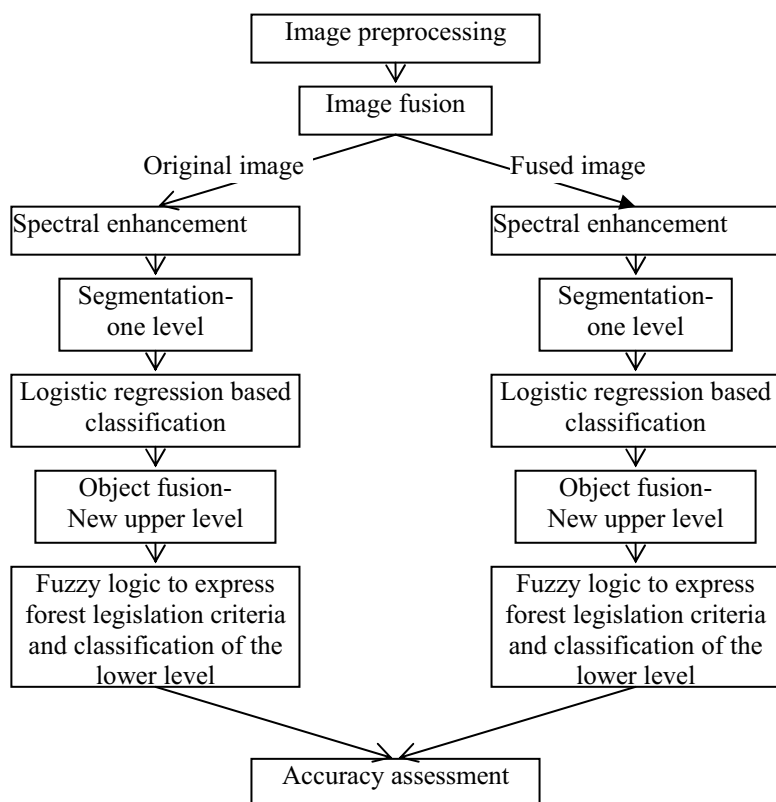
Fig. 2. The flowchart of the methodology

3.1 Data acquisition and pre-processing

A Quickbird satellite image was acquired on June 2004. A Digital Elevation Model of the area was extracted from existing stereo pairs of air photos and several ground control points were identified to support the

orthorectification process. The total RMS error did not exceed 1.1 meters. Small agricultural fields were masked out of the data, since the National Forest GeoDatabase will be a continuum of a currently developing National Cadastre system upon which accurate delineated geographic limits between natural and non-natural areas will be available.

3.2 Spatial and spectral enhancement

The Gram-Schmidt method of image fusion (Laben and Brower 2000), included in the commercial software package RSI ENVI 4.3, was used for the fusion of the multispectral with the panchromatic image, in order to enhance the spatial information present in the former one. Consideration of the additional spatial detail in an object-based classification process, could presumably affect both delineation of the limits as well as the internal structure (texture) of objects.

The fusion method adopted in this study proved to be more suitable for the specific forest scene (Mallinis 2006) considering spectral and spatial quality indicators (Wald 1997, Zhou et al. 1998), in comparison to the modified IHS (Bretschneider and Kao 2000), the PCA and the wavelet based (King and Wand 2001) methods of fusion.

Also, in order to overcome the limited spectral resolution of the original imagery (Key et al. 2001), the IHS transformation was applied and the Normalized Difference Vegetation Index (NDVI) was calculated and evaluated as means to aid the spectral discrimination process (Goetz et al. 2003).

3.3 Segmentation

The segmentation algorithm followed in this work is part of the FNEA concept (Baatz and Shape 2000), embedded within the commercial software Definiens Professional (eCognition ver. 4). It supports a multi resolution segmentation of an image producing different segmentation levels where each one is focused on different scale, leading to an arbitrary developed semantic hierarchy (Hay et al. 2003). Therefore, prior to the segmentation of a scene following the FNEA concept, the user has to consider four important issues:

- What are the processes observed in the scene under work?
- What is the operational scale of the objects of interest?
- Which scales/segmentation levels have to be produced and linked in order to produce the desired output objects?

- Is it feasible to construct this hierarchical network of linked objects given the resolution and the extent of the available imagery?

Clarifying the above issues is necessary in order to avoid the development of complex hierarchies which would create confusion in the information extraction process.

In our study, having already overcome the problem of segregating non-natural areas, all the candidate segmentation levels should presumably relate to phenomena and objects found within natural areas. Based on the findings of a field survey it was determined that pixel's size in the original image was larger than the 1/2 to 3/4 of the mean canopy size (Woodcock and Strahler 1987), making thus individual tree crown recognition not feasible.

Consequently, the scale of the segmentation was adjusted in order to extract clumps of crowns and/or very large individual trees (deciduous). The development of a second level of segments was not considered to be necessary, because no other forest situations were recognized to operate within the extent of the scene under investigation.

3.4 Classification

Within eCognition, various features are available for providing information about the spectral, shape, textural and class membership attributes of derived objects.

Multiple logistic regression modeling can be used to select the most appropriate explanatory variables for predicting a binary dichotomous response (Hosmer and Lemeshow 1989).

In terms of the probabilities for either of the two outcomes of the event, the equation for the logistic regression model can be expressed as:

$$p_i = \frac{\exp\left(a + \sum_{g=1}^{m} \beta_g x_g\right)}{1 + \exp\left(a + \sum_{g=1}^{m} \beta_g x_g\right)} \tag{1}$$

where the parameters α and β_g (g= 1,2,3..m) of the x_g variables, are usually computed using the maximum likelihood method.

The significance of a logistic regression model is indicated by a statistical significant reduction in the log likelihood value (-2LL), while the goodness of fit is assessed by statistical measures such as the Hosmer and

Lemeshow's value, the classification matrices and the Nagelkerke measure (Hair et al. 1998).

For the implementation of this method, the segments delineated were exported to a GIS software environment. Samples were selected through photo-interpretation for the two categories of interest (453 segments covered by tree canopy and 402 segments with no tree canopy). Due to the fine scale segmentation adopted no intermediate situations were observed (segments half-covered with canopy). The most appropriate variables were selected following a forward stepwise procedure based on the likelihood ratio statistic. Following the implementation of the logistic models, the resulted probabilities images (for the original and the fused Quickbird imagery) were converted to binary images adopting a cut-off threshold value. The appropriate thresholds were determined in order to maximize the observed overall proportion of correct classifications. These values were selected by considering the so called specificity which equals the proportion of correct classifications for the subset of data with y=0 and sensitivity, which represents the proportion of correct classifications for the subset of data with y=1 (Hadjicostas 2006).

The classified segments were subject to an object fusion procedure. This process aimed to unify adjacent segments which were parts of the same semantic entity (Definiens 2004). The output of the classification based fusion was a new level of objects located above the original one in the hierarchy.

The newly derived objects were classified considering the logistic-based classification of the first level (class of sub-objects) and the inclusion and exclusion criteria specified in the legal forest definition. Different fuzzy rules (membership functions) were formulated for each class to express the form, the size and the vicinity of each object with other classification categories.

3.5 Accuracy assessment

In order to assess the classification results of both the original and the fused image, manual photo interpretation procedures were followed using both air photos of the study area and the satellite imagery.

Two independent operational photo-interpreters digitized manually the forest and non-forest polygons within the study area. Based on these maps, 600 random points were selected automatically and used for the accuracy assessment procedure. Furthermore, an overlay procedure was followed to compare the results.

Finally, a pairwise test statistic was also performed, based on the Khat coefficient of agreement, in order to statistically compare the classification results of the two images (Cohen 1960):

$$Z = \left| \text{Khat}_1 - \text{Khat}_2 \right| / \sqrt{\text{varKhat}_1 + \text{varKhat}_2} \tag{2}$$

where Khat1 and Khat2, denote the estimates of the kappa statistic and VarKhat1 and VarKhat2, the corresponding estimates of their variance.

At the 95% confidence level, the critical value for determining if two error matrices are significantly different, is 1,96.

4 Results and discussion

4.1 Segmentation

Based on a "trial and error" procedure, the red band of the original image as well as the hue and saturation bands were used to segment the image (Fig. 3), whilst the scale parameter was set to 70 and no shape heterogeneity was considered. The positive contribution made by the former two bands can be explained by the fact that they are the principal containers of the original spectral information during the IHS transformation (Koutsias et al. 2000).
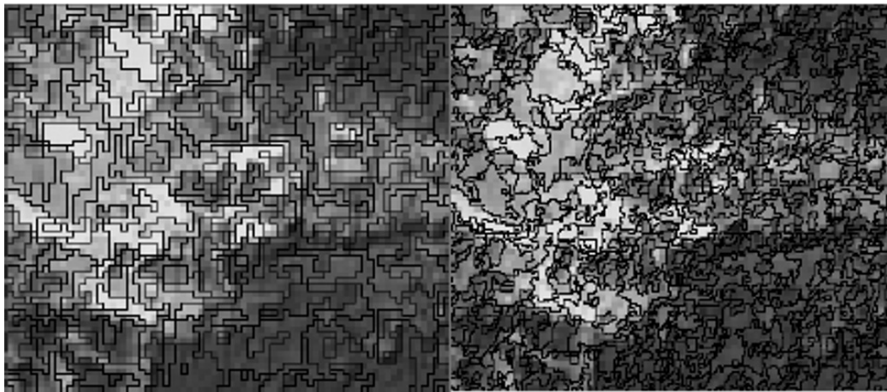
Fig. 3. Segmentation results for a portion of the original image (left) and the fused image (right)

Respectively, for the fused image, its red and near-infrared bands were used and the scale parameter was set to 40. The resulted segments (Fig. 3)

had an overall smoother appearance in comparison to the ones resulted from the original image.

4.2 Classification

Both spectral and textural features were considered to be potential variables for enhancing the discrimination among the categories.

In the case of the original image, the mean segment value (calculated from the pixels contained in each segment) for the green, blue and red bands of the imagery, the ratio of Qb4 as well as the mean segment value for the saturation band were selected.

The mean and ratio features of band Qb4, as well as the mean and standard deviation features of saturation band were included in the logistic model constructed for the segments of the fused image.

Table 1. Statistical measures for assessing overall fit of the logistic models.

Measure	Logistic model/ original image		Logistic model/ fused image	
	Value	Signif.	Value	Signif.
Hosmer and Lemeshow's χ^2	8.601	0.377	7.198	0.515
-2LL	124.528		62.121	
Cox & Snell R Square	0.672		0.712	
Nagelkerke R Square	0.897		0.951	

Goodness of fit for the logistic models proved to be satisfactory (Table 1) since both of them resulted in small values of -2LL.
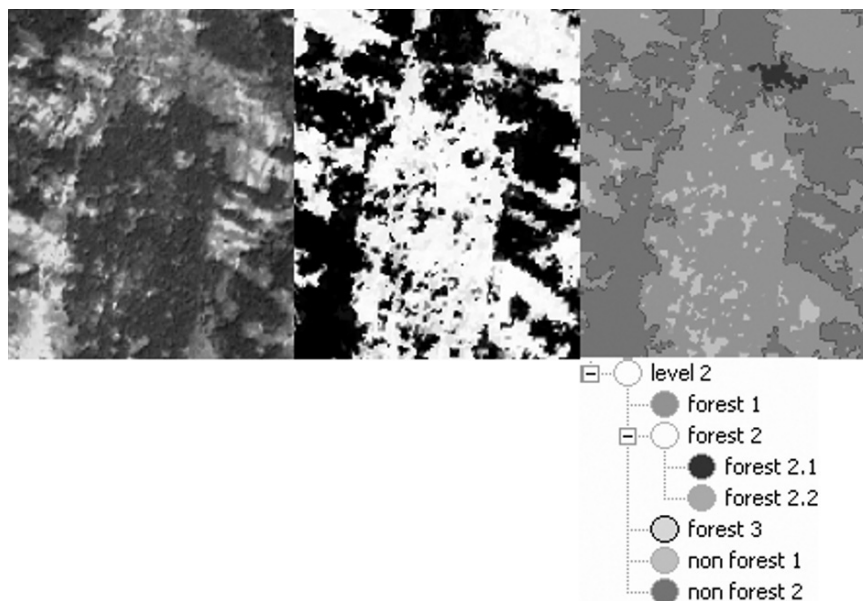
Fig. 4. Subset of the original image (left) the probabilities image (middle) and the classification result (right) of the second level. The lower right image depicts the class hierarchy. Lighter tones in the probabilities image indicate greater probability for a segment to be assigned to the category tree canopy

The Hosmer and Lemeshow's statistical measure indicated a good model fit considering the non-significant chi-square values obtained. Furthermore, the Negelkerke value was 0.897 for the original image and 0.951 for the fused image. This measure has a range of 0 to 1 with higher values indicating better fit. Also, the high values of the Cox and Snell asserted the satisfactory fit of the models.

Following the implementation of the logistic models, two images (Fig. 4-middle image) were derived respectively predicting the existence of tree canopy. The value that coincided with the intersection of the corresponding curves of sensitivity and specificity was 0.6 for both images (Fig. 5).
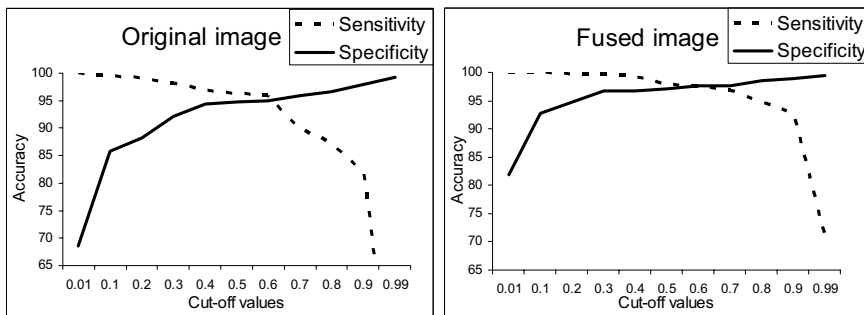
Fig. 5. Specification of the cut off values, adopted to convert the probabilities images for the original (left) and the fused data (right) to binary images.

The classification based fusion generated a second level of the objects and classification categories of this level were described using fuzzy rules in accordance to the criteria of the forest law (Table 2 and Fig. 4).

Table 2. Classification scheme of the second level according to the criteria of the legal forest definition.

Class name		Class description
forest 1		Objects with sub-objects classified as «tree canopy», with area exceeding 0.3 hectares
forest 2		Objects with sub-objects classified as "tree canopy" but with area smaller than 0.3 hectares
	forest 2.1	Objects wide or with interdependence with category *forest 1*
	forest 2.2	Objects without close interdependence with category forest 1 (therefore characterized as NonForest)
forest 3		Objects with sub-objects classified as "tree canopy", with area smaller than 0.3 hectares, surrounded from objects belonging to class *non forest 2*, which are encompassed from objects classified as *forest 1*
non forest 1		Objects with sub-objects classified as "no tree canopy"
non forest 2		Objects with sub-objects classified as "no tree canopy", with small area and elongated shape surrounded by *forest 1* (therefore characterized as Forest)

One important point of the approach was the specification of threshold values for the features used. Specifically, setting the threshold for the minimum patch to be characterized as forest was a straightforward process since it is specified in the legal definition (0.3 hectares). On the other hand, setting the distance i.e. denoting interdependence between adjacent forest

patches, was a subjective decision based on the expert knowledge of local forest managers. Since this distance is related not only to the forest species present but also to the overall landscape configuration, it is likely that transfer of the classification approach to other areas within the same country would require minor adjustments of the corresponding thresholds

After classifying this second level of objects, the final classification-based fusion took place, in which objects or parts of objects belonging to the same semantic entity (forest or no forest) were merged (Fig. 6).
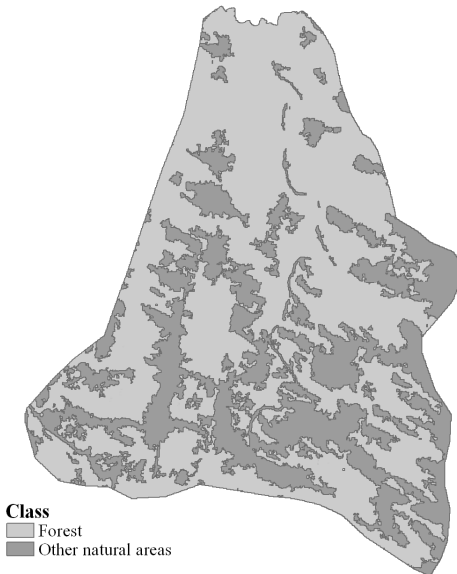
Fig. 6. Final classification map of the original image

The logistic regression based classification of the original image proved to be the most accurate (Table 3) in terms of overall accuracy, according to both photo-interpreters. According to the results of the pairwise test (Table 3), none of the matrices were found to be significantly different.

It appears that the use of the fused Quickbird image did not improve the classification result. A possible explanation could be the non-existence in the area of trees having a canopy small enough in the area to be properly segmented in the image.

Table 3. Overall accuracy for the three classification approaches

		Original	Fused	Z value
1st photointerpreter	Overall accuracy	85.67	81.00	
	Khat	0.639	0.589	0.489
2nd photointerpreter	Overall accuracy	84.17	82.33	
	Khat	0.656	0.629	0.578

The developed classification approach allowed quantification of the legal definition criteria with semantic knowledge. Development of a semi automated operational method, is advantageous, in a sense that, future modifications in definition of forest in the frame of a worldwide or European Union "harmonization", could be easily integrated to such a model. The adoption of this approach operationally seems promising, but accuracy should be further improved beforehand.

A potential solution might be the coupling of the current approach with individual tree delineation methods, an approach similar to that found in Burnett and Blaschke (2003). However, the implementation of automated methods for individual tree crown recognition seems to be complicated in areas covered with deciduous forest species (Wulder 1998). Probably the most hopeful approach in an operational context seems to be the incorporation of satellite remotely sensed data providing information about the vertical structure of the forest stands into the classification process.

5 Conclusions

An object-based methodology for forest areas discrimination in Greece has been developed in this work. Overall, the results obtained indicate that the developed approach has the potential to be adopted for forest cover assessment at local scale according to criteria contained in legal forest definition. Development of a set of rules applied to objects could minimize the discrepancies resulting from the judgment of different photo interpreters. Finally the fusion of the multispectral with the panchromatic component of the Quickbird image proved to be of no use for the classification result, considering also the additional processing time needed.

6 Acknowledgements

References

Baatz M, Benz U, Dehghani S, Heynen M, Höltje A, Hofmann P, Lingenfelder I, Mimler M, Sohlbach M, Weber M, Willhauck G (2004) eCognition User Guide 4. Definiens Imaging, Munchen

Baatz M, Schäpe M (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Blaschke T, Greisebener G (eds) Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT Symposium Salzburg, vol. 200. Herbert Wichmann Verlag, Karlsruhe, pp. 12–24

Bretschneider T, Kao O (2000) Image fusion in remote sensing. In: 1st Online Symposium of Electronic Engineers

Burnett C, Blaschke T (2003) A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecol. Model. 168:233-249

Chubey MS, Franklin SE, Wulder MA (2006) Object-based analysis of IKONOS-2 imagery for extraction of forest inventory parameters. Photogram. Eng. Remote Sens. 72:383-394

Cohen J (1960) A coefficient of agreement for nominal scales. Educ. Psychol. Meas. 20: 37-40

Franklin SE, Wulder M (2003) Remote Sensing of Forest Environments, Challenges and Opportunities; In Wulder M, and Franklin SE (eds) Methods and Applications for Remote Sensing of Forests: Concepts and Case Studies. Kluwer Academic Publishing, Dordrecht, pp. 511-514

Gitas I, Mitri G, Ventura G (2004) Object-oriented image analysis for burned area mapping using NOAA AVHRR imagery in Creus Cape, Spain. Remote Sens. Environ. 92:409-413

Goetz SJ, Wright RK, Smith AJ, Zinecker E, Schaub E (2003) IKONOS imagery for resource management: Tree cover, impervious surfaces, and riparian buffer analyses in the mid-Atlantic region. Remote Sens. Environ. 88: 195-208

Haapanen R, Ek AR, Bauer ME, Finley AO (2004) Delineation of forest/nonforest land use classes using nearest neighbor methods. Remote Sens. Environ. 89: 265-271

Hadjicostas P (2006) Maximizing proportions of correct classifications in binary logistic regression. J Appl. Stat. 33:629-640

Hair JFJr, Anderson RE, Tatham RL, Black WC (1998) Multivariate Data Analysis (5th edn). Prentice-Hall, New Jersey

Hay GJ, Blaschke T, Marceau DJ, Bouchard A (2003) A comparison of three image-object methods for the multiscale analysis of landscape structure. ISPRS J Photogramm. 57:327–345

Hay GJ, Castilla G, Wulder MA, Ruiz JR (2005) An automated object-based approach for the multiscale image segmentation of forest scenes. Int. J. Appl. Earth Obs. Geoinform. 7:339-359

Hoffman RR, Markman AB, Carnahan WH (2001) Angles of Regard: Physiology meets Technology in the Perception and Interpretation of Nonliteral Imagery. In: Hoffman RR, Markman AB (Eds) Interpreting Remote Sensing Imagery. Human Factors, Lewis Publishers, Boca Raton FL, pp. 11-58

Hosmer DW, Lemeshow S (1989) Applied Logistic Regression. Wiley and Sons, New York

Kennedy P, Bertolo F (2002) Mapping sub-pixel forest cover in Europe using AVHRR data and national and regional statistics. Can J Remote Sens 28: 302-321

Key T, Warner TA, McGraw JB, Fajvan MA (2001) A comparison of multispectral and multitemporal information in high spatial resolution imagery for classification of individual tree species in a temperate hardwood forest. Remote Sens. Environ. 75: 100-112

King RL, Wang, J (2001) A wavelet based algorithm for pan sharpening Landsat 7 imagery. Proc. IGARSS 2001, pp 849 –851.

Kosaka N, Akiyama T, Tsai T, Kojima T (2005) Forest Type Classification Using Data Fusion of Multispectral and Panchromatic High-Resolution Satellite Imageries. Proc. IGARSS 2005, pp 2980-2983

Koutsias N, Karteris M (1998) Logistic regression modelling of multitemporal Thematic Mapper data for burned area mapping. Int. J. Remote Sens. 19:3499-3514

Koutsias N, Karteris M, Chuvieco E (2000) The Use of Intensity-Hue-Saturation Transformation of Landsat-5 Thematic Mapper Data for Burned Land Mapping. Photogram. Eng. Remote Sens. 66:829-839

Laben CA, Brower BV (2000) Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. U.S. Patent 6 011 875

López-Caloca, AA, Mora F, Escalante-Ramírez B (2004) Mapping and characterization of urban forest in Mexico City. Proc. SPIE 5239:522-531

Lu D, Weng Q (2005) Urban classification using full spectral information of Landsat ETM+ imagery in Marion County, Indiana. Photogram. Eng. Remote Sens. 71:1275-1284

Lund HG (2002) When is a forest not a forest?. J. Forest. 100:21–28

Mallinis G (2006) A remote sensing approach for the generation of a National Forest Geodatabase. Phd thesis Aristotle University of Thessaloniki, Thessaloniki, Greece.

McRoberts RE (2006) A model-based approach to estimating forest area. Remote Sens. Environ. 103: 56-66

Phol C, van Genderen JL (1998) Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications. Int. J. Remote Sens. 19:823-854

Ranchin T, Wald L (2000) Fusion of high spatial and spectral resolution images: The ARSIS concept and its implementation. Photogram. Eng. Remote Sens. 66:49-61

Tanaka S, Sugimura T (2001) Cover: A new frontier of remote sensing from IKONOS images. Int. J. Remote Sens. 22:1-5

Thiel C, Drezet P, Weise C, Quegan S, Schmullius C (2006) Radar remote sensing for the delineation of forest cover maps and the detection of deforestation Forestry 79: 589-597

Trias-Sanz R., Boldo D (2005) A high-reliability, high-resolution method for land cover classification into forest and non-forest. Lecture Notes in Computer Science 3540: 831-840

Woodcock CE, Harward VJ (1992) Nested-Hierarchical Scene Models and Image Segmentation. Int. J. Remote Sens. 13:3167-3187

Woodcock CE, Strahler A (1987) The factor of scale in remote sensing. Remote Sens. Environ. 21:311-322

Wulder M (1998) Optical remote-sensing techniques for the assessment of forest inventory and biophysical parameters. Prog. Phys. Geog. 22:449-476

Wulder MA, White JC, Hay GJ, Castilla G (2007) Pixels to objects to information: Spatial context to aid in forest characterization with remote sensing. In: Blaschke T, Lang S, Hay G (eds.) Object-Based Image Analysis Spatial concepts for knowledge-driven remote sensing applications, Springer-Verlag (under publication)

Wald L, Ranchin T, Mangolini M (1997) Fusion of Satellite Images of Different Spatial Resolutions: Assessing the Quality of Resulting Images. Photogram. Eng. Remote Sens. 63: 691-699

Zhou J, Civco DL, Silander JA (1998) A wavelet transform method to merge Landsat TM and SPOT panchromatic data. Int. J. Remote Sens. 19: 743-757

Chapter 3.4

Object-based classification of SAR data for the delineation of forest cover maps and the detection of deforestation – A viable procedure and its application in GSE Forest Monitoring

Ch. Thiel, Ca. Thiel, T. Riedel, C. Schmullius

Friedrich-Schiller-University Jena, Inst. of Geography, Dept. for Earth Observation, Grietgasse 6, 07743 Jena, Germany

ABSTRACT: In this chapter forest cover and forest cover change mapping basing on image objects is discussed. Change relates to recent complete harvesting and reestablishment, degradation or thinning is not considered. For the change maps two different strategies are proposed. The first one derives the changes by means of previously classified images of a multitemporal dataset and is thus referred to as "post-classification change detection". For increasing the accuracy of the change maps a knowledge based change detection approach is introduced. The second strategy considers all scenes of the multitemporal dataset simultaneously. This method is referred to as "multidate classification".

Generally any kind of Earth Observation (EO) data allowing the grey value based separation of forest and non-forest can be applied with both strategies. In this study, JERS-1 (Japanese Earth Resource Satellite) SAR data are used for method development. The feasibility assessment of both object-based mapping strategies is performed at five test sites: Germany (Thuringia), UK (Kielder), Sweden (Remningstorp and Brattåker) and Russia (Chunsky). Due to the specific data requirements (broad multitemporal dataset) the first approach could only be successfully implemented at the Thuringia site. It was also tested at Kielder, but with deficient results.

The chapter concludes with the successful realisation of the approach at the Russian service case of GSE FM. Because of the given time frame (1990-recent) other EO data sources had to be implemented. As historical EO data source LANDSAT TM was selected, recent information is derived from ASAR APP.

1 Introduction

1.1 General Instructions

The potential of Synthetic Aperture Radar (SAR) data for forestry applications is substantiated by a huge number of publications. SAR data is widely used for forest cover mapping (Leckie & Ranson 1998, Schmullius et at. 2001), forest disturbance mapping (e.g. logging, forest fire, and wind damage) (Kasischke et al. 1992, Yatabe & Leckie 1995, Rignot et al. 1997) and forest biomass assessment (Dobson et al. 1992, Israelsson et al. 1994, Santoro et al. 2006). Lower radar frequencies proved to be of particular adequacy. Whereas current spaceborne SAR missions of the European Space Agency (ESA) are operating at C-band, the Japan Aerospace Exploration Agency (JAXA) provides L-band data (JERS-1, PALSAR). Despite the Japanese Earth Resource Satellite (JERS: 1,275 GHz, HH, 35° off nadir angle, 12,5 x 12,5 m² pixel spacing) program and NASA Shuttle Imaging Radar (SIR) missions, the experience with L-band applications especially in Europe is not as well developed as for C-band. Thus, the first issue of this chapter is to illustrate the capabilities of JERS SAR data for the forest cover mapping and the detection of clear-cutting.

The second issue of this chapter deals with the development of a reliable and above all reasonable approach for operational large area forest cover and clear-cut mapping with high accuracy. Degradation or thinning is not considered. For the generation of those maps two different object-based classification strategies were developed. The first one maps changes with (previously classified) images of a multitemporal dataset and is thus referred to as "post-classification change detection". The second strategy considers all scenes of the multitemporal dataset simultaneously and is referred to as "multidate classification". Although this chapter is rather dedicated to SAR data it can be stated, that generally any kind of EO data allowing the spectral separation of forest and non-forest can be applied.

Eventually the approach is implemented at a real service case of GSE FM (GMES Service Element Forest Monitoring). At that point the whole procedure had to reach operational level as the delivery of standardised

high quality products is core of GSE FM. The provided products of this service case include a forest area map, a clear-cut/burned area map and a forest area change map considering a minimum mapping unit of 1 ha. The acceptability threshold of the thematic mapping accuracy is 90% for non-change maps and 85% for change maps, respectively.

1.2 GSE Forest Monitoring – Intension and Accomplishment

GSE FM is one element of the GMES (Global Monitoring for Environmental and Security, see Lang in this book) initiative of the ESA Earthwatch Programmes. GSE FM stage 2 started in 2005 with three year duration. The extending international consortium is led by GAF AG and started with 18 Service Providers and 25 End User. Main goal is to deliver customised and policy-relevant and information mainly based on EO data in ready-to-use packages in the field of Climate Change, Sustainable Forest Management as well as in Environmental Issues and Natural Protection. The supplied products and services are validated and standardised to support decision-making and improved policies that enable cost effective sustainable forest management in various countries. To guarantee the GSE FM standards all products including their specifications and instructions for production are collected within the Service Portfolio Specifications.

2 JERS Test sites and data

To identify suitable test sites, the following criteria were adopted: Coverage of different vegetation zones, access to a JERS-1 SAR data time-series and the availability of ground data. Hence, the following sites were chosen (Fig. 1): Kielder, Thuringia, Brattåker, Remningstorp, and Chunsky. Tab. 1 summarises the available JERS SAR and reference data.

Kielder forest is located in northern England and is managed by the state funded Forest Enterprise agency. This hilly area covers about 86 km², of which more than 50% is forested. Productive coniferous species are predominant in these forests and a harvesting yield of some 400,000 tonnes of timber per year is sustained. The main tree species are Sitka spruce (75%), Norway spruce (12%) and Lodgepole pine (9%).

The Thuringia test site is part of the Thuringian Forest. The relief causes significant variations in backscattering intensity. The test site covers 1,000 km² and 58% of the area is tree covered. Tree species are 86% spruce, 7.5% pine, 3.1% beech and 3.5% others. Clear-cutting is generally

not practised; thinning is the more common logging technique. However, disturbances as bark beetle attacks and wind damage occur.

Brattåker is a forest research site (60 km²) managed by the Swedish forest company. The prevailing tree species are Norway spruce and Scots pine, but some deciduous tree species, e.g., birch, are also present. This test site represents rather intensively managed boreal forest, compared with other areas in the northern part of Sweden. The other Swedish site, Remningstorp, comprehends about 12 km² forested land. About 10% of the area is forested peatland. The main tree species are Norway spruce, Scots pine and birch. A few stands are dominated by oak and beech.

The Chunsky forest territory covers almost 400 km² and includes more than 1,200 forest compartments, of which about 90% can be denoted as natural stands. Birch and aspen are the major broad-leaved species, while pine and larch are the dominant coniferous species.

Fig. 1. Location of JERS test sites and Irkutsk Oblast

Table 1. Compendium of available EO and reference data

Test site	JERS scenes (Path, Row)	Reference data
Kielder	11.07.1993 (P325, R2492)	Forest cover map
	16.07.1996 (P326, R2492)	Clear cutting between 1996-1998
	02.08.1998 (P325, R2492)	
Thuringia	19.12.1992 (P308, R215)	Comprehensive forest inventory
	18.03.1993 (P298, R2588)	data
	05.03.1994 (P309, R215)	
	07.01.1995 (P309, R215)	
	11.01.1998 (P298, R2588)	
Brattåker	06.06.1993 (P290, R2312)	Location of clear cuts (dateless)
	24.07.1996 (P307, R192)	

	15.05.1998 (P290, R2312)	
Remning-storp	30.04.1993 (P297, R2432)	Forest inventory data for a small
	14.07.1994 (P297, R2432)	subsections of the site
	22.05.1998 (P297, R2432)	Location of clear cuts (partly dateless)
Chunsky	16.01.1996 (P157, R203)	Forest inventory data
	27.06.1997 (P157, R203)	Location of clear cuts (dateless)
	28.07.1998 (P157, R203)	

3 Methodology

The SAR data pre-processing involved, besides calibration to σ^0 and orthorectification, the removal of the typical topography induced distortions of backscattering intensity using the SAR processing algorithms for relief calibration and angular backscatter correction by Stussi et al. (1995).

Man-made clear-cutting is characterised by regular geometric patterns (see Wulder et al. in this book). Due to their dissimilar scattering behaviour forested and deforested areas can typically be distinguished with SAR data. Segmentation procedures are capable to recognise and frame the borders of adjoining areas with differing backscatter. Hence, the regular geometric patterns of clear-cuts can be transformed into image segments which can be used for classification purposes. This principle is of particular interest for SAR data classification as the disturbing statistical effects of speckle is diminished - only the average backscatter of each segment is considered by the classifier. Thus, the data analysis and classification are based on image objects (segments), where segments are identified using a multiresolution segmentation algorithm (Baatz & Schäpe 2000, Benz et al. 2004). The borders of the segments do not necessarily represent the forest compartments, but in general identify homogeneous patches of land. The segment size is determined by a scale parameter and can range from single pixels to the entire scene. An adequate choice of the segment size must amongst others account for the desired minimum mapping unit (1 ha is feasible using JERS).

Fig. 2 shows an example of a segmented image. Bright segments are forested and dark segments unforested. Clear-cuts or young forest stands are evident within the forested area (middle grey). The borders of differing adjoining forest partitions are clearly framed. For the segmentation the whole JERS time series was always embedded (multi-image segmentation). Thus, the segments are the same in all images.

For the creation of forest cover maps and the detection of forest cover changes, two different approaches were considered. The first approach dis-

cussed here is referred to "post-classification change detection". Thresholds on segmented σ^0 values are used to divide each scene into forest and non-forest (Fig. 3.). This threshold varies between acquisitions as a result of altered tree properties, weather conditions and acquisition system parameters, and it must be adapted for each scene individually to achieve optimum classification. GIS based forest stand information was used for sample selection and thus threshold definition.
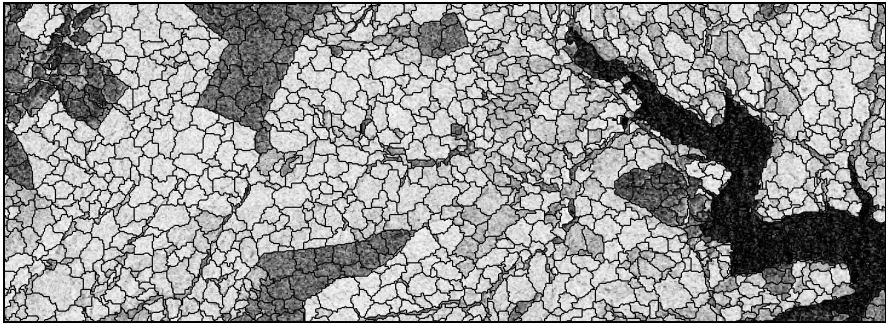
Fig. 2. Example for segmented JERS SAR-image (Kielder)
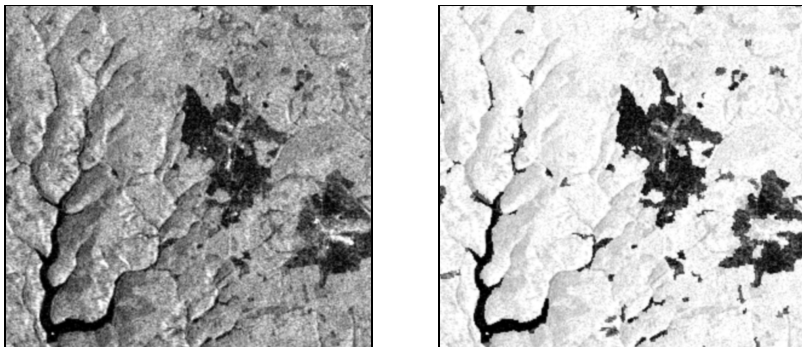
Fig. 3. Left: Original σ^0 image. Right: Example of forest cover map derived by threshold based classification (part of Thuringia test site, subset size ca. 7 x 8 km)

Time series of classified images are used to delineate forest cover changes. The delineation was executed with GIS procedures. Each image segment was analysed temporally with respect to its membership to forest or non-forest respectively. To reduce the impact of misclassifications for each particular time of recording expert knowledge was integrated into the change analysis. E.g. if the classification result of one image object was

fluctuating over the whole time series like 1992: forest, 1993: non-forest, 1994: non-forest, 1996: forest, 1998: non-forest etc. it is not interpreted as change, but as ambiguous (for manual review). This is basing on the fact, that it takes a much longer time span to grow and to harvest a forest. A variation of the classification of one image object was interpreted as change, if it was assigned stable to one class for the consecutive years before the change (e.g. 1992: forest, 1993: forest, 1994: forest) and to the opposing class past the change (1995: non-forest, 1998: non-forest). Based on such semantics the time series of classified images was translated into a forest cover change map. It is important to state that the statistical soundness of this translation and thus the quality of the change product is depending on the number of classified images. If there are less then two images before and after the actual forest stand specific change, this knowledge based translation method can not substantiate this change.

The second approach considers all the scenes simultaneously. The classification incorporates forest, non-forest and forest change classes. This method is referred to as "multidate classification" (Fig. 4). At the multitemporal composite (Fig. 4, left) permanent forested (tree covered) areas appear in bright grey, permanent non-forested areas in dark grey and changes within the forested area in medium shades of grey. The change map (Fig. 4, right) is based on a supervised (nearest neighbour) classification of σ^0 values. To create class signatures, an appropriate set of training samples is created for each class. During the classification process each image segment is then assigned to the appropriate class signature. By means of this classification result forest cover maps for each year of the time series can be derived.
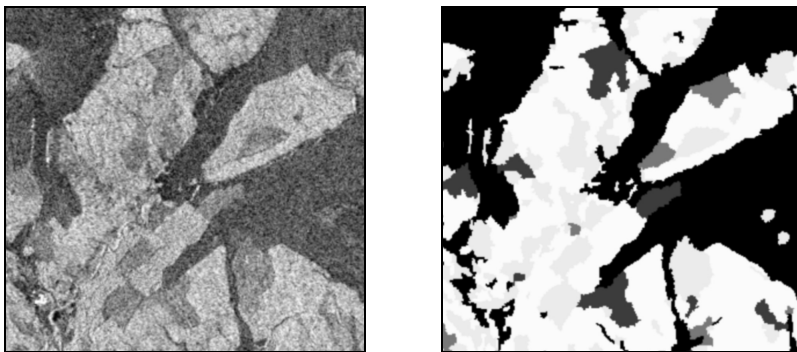
Fig. 4. Left: Multitemporal (1998/1996/1993) σ^0 composite (7 x 8 km subsection of Kielder site). Right: Change map: black: non-forest, white: forest, dark grey: clear-cuts 1993-1996, light grey: clear-cuts 1996-1998

4 Results

4.1 Kielder (UK)

At the Kielder site both mapping strategies were tested. The results obtained with the "multidate classification" are presented here and are compared with the "post-classification change detection" results in section 5. Three suitable JERS SAR images were available. In accordance with the existing ground data, the validation of the clear-cut map was carried out only for the 1996-1998 period.

To estimate the separability of the classes a signature plot for the training sample set was created (Fig. 5). For the second and the third acquisition the clear-cut segments experienced a decrease in backscatter. Although the σ^0 values of the clear-cuts lie in the range of the standard deviation of forest, most of the clear-cuts could be detected (see Fig. 6). The apparent temporal variability in overall backscatter was investigated in another study (Thiel et al. 2004) and was related to JERS calibration problems and Faraday rotation (Freeman & Saatchi 2004).
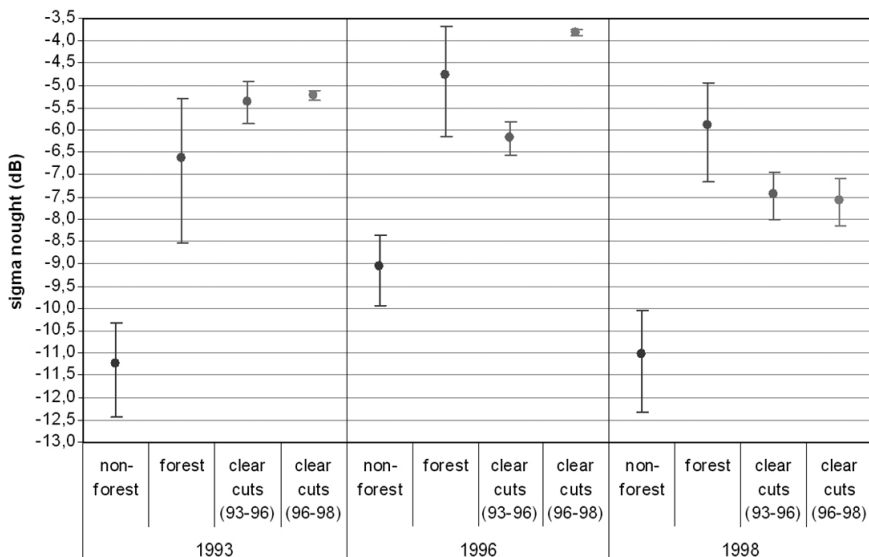
Fig. 5. Signature plot for Kielder site: Note the decrease of σ^0 with clear-cutting (clear cut (93-96) means, clear-cut was established between 1993 and 1996)

The spatial and temporal correspondence of the SAR and GIS data between 1996 and 1998 allows direct comparison of the results. Out of the 42

clear-cuts within the test site 37 could be recognised. A clear-cut was considered correctly detected if the centres of the SAR polygons were located within the ground data GIS polygon, even if the borders were somewhat different. In some cases discrepancies can be attributed to the different geometric properties (e.g. vector vs. raster, geometric errors, impact of speckle on segmentation). Seven "SAR-detected" clear-cuts could not be confirmed by the GIS data, which might be even due to incomplete clear-cut information in the ground data. In a few cases it could also be a consequence of wind damage, which in the case of Kielder Forest is a frequent occurrence that is not embodied in the clear-cut GIS data.

Fig.6 shows that SAR data analysis severalfold underestimates the clear-cut areas. Five of the clear-cuts are not detected. One reason could be that the felling data may include areas of cleared shrubs, dead forest, or failed (unestablished) forest in addition to mature forest that has been cleared in preparation for replanting. The change from these "forest types" to clear-cut is hardly detectable with JERS SAR data. Moreover those changes are regularly not treated as deforestation (IPCC 2003).
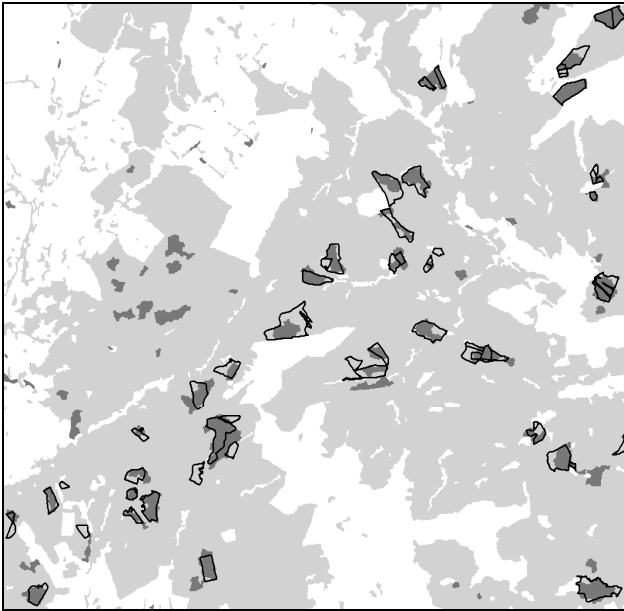
Fig. 6. Logging at Kielder site (1996-1998); Light grey: forest, white: non-forest, dark grey: SAR detected logging, polygons: in-situ logging GIS data

4.2 Thuringia (Germany)

For the creation of the forest cover change map of Thuringia the "post-classification change detection" approach was implemented. Five JERS SAR scenes were incorporated. The forest cover maps were formed by thresholding σ^0 for each particular SAR image. A sample of the classification for 1992 is given in Fig. 7. To estimate the separability of the classes, signature plots were generated as above (Fig. 5). The class means and the standard deviations of the training areas indicate a distinct separability of forest and non-forest.

The entire area of the 1992 SAR based forest cover map was compared (pixel wise overlay) to the inventory data. The overall accuracy of the delineated forest cover map is 90%. Image segments where forest was incorrectly detected amount to 7.5% (of total area). These misclassifications mainly occur at small settlements and some agricultural areas with high biomass crops such as maize. Both these land cover types can produce very high backscatter that extends into the range of the forest signatures (see discussion section). Converse misclassifications (detection of non-forest instead of forest) amount to 2.7% of total area and occur mainly for young forest stands with low backscatter. An inventory based verification of the SAR based forest cover maps for acquisitions after 1992 was not possible because of the lack of forest inventory, however similar accuracy would be expected for these maps.

Fig. 7. Forest cover map Thuringia (1992); grey: forest, white: non-forest, polygons: forest stands taken from forest inventory data

The SAR-based forest cover change map combines the information from the temporal sequence of forest cover maps as described above. Even

though clear-cutting is not officially exercised in Thuringia, deforestation is evident in Fig. 8. Clear-cutting may be applied as an exception to remove diseased (e.g. bark beetle) or damaged (e.g. storm damage) forest, but such activities cannot be inferred from the inventory data.
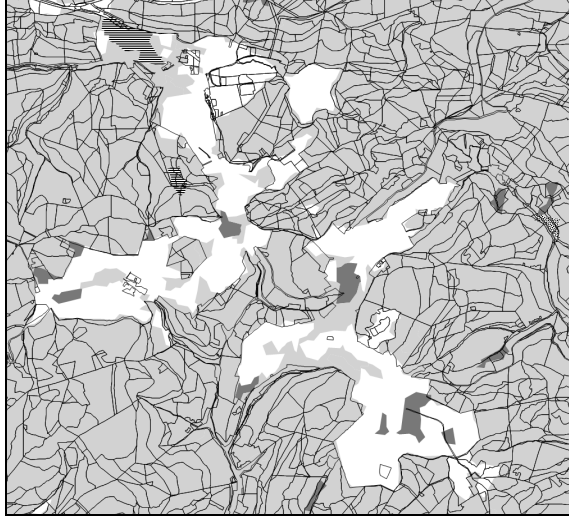
Fig. 8. Subset of forest cover change map for Thuringia; light grey: forest, white: non-forest, hatched: logging, dotted: replanting, dark grey: ambiguous

4.3 Remningstorp, Brattåker, and Chunsky

For the remaining sites only the "multidate classification" was applied, as the SAR data set was not appropriate for the other approach (short time span, only three good quality images). Unfortunately for Remningstorp, Brattåker, and Chunsky the in situ information was not adequate for thorough accuracy assessment. Thus this task was conducted using the SAR images and the confusion matrix method as proposed by IPCC (2003). Forest cover changes were detected by means of a visual interpretation of all considered SAR images. Validation areas for each class were selected basing on the visual analysis. These areas are taken as the actual "truth" data in the confusion matrix. For each class, the numbers of correctly classified pixels and confused pixels resulted in an overall accuracy of ca. 90% for the three sites which is comparable to Kielder and Thuringia.

5 Discussion of JERS Results and Methodology

The results demonstrate that JERS SAR data in combination with an object-based classification approach can be used to create reasonably accurate forest cover and forest cover change maps. The overall accuracy of the (by thresholding) derived forest cover maps is ~90% by area. Moreover, ~90% of the clear-cuts could be detected (multidate classification).

For the detection of forest cover changes and the delineation of forest maps, two differing methods were applied. The "multidate classification" uses a multi-parameter set of SAR data and delivers more precise results compared to the "post-classification change detection". For practical reasons a supervised classification algorithm such as "Nearest Neighbour" is preferable (against simple thresholding) to cope with the classification.

"Post-classification change detection" is a simple method for separating forest from non-forest by thresholding. It is more transparent and thus assumed rather transferable to other sites (preconditions: sound sensor calibration and comparable environmental conditions) and only one SAR image is required. On the other hand, change analysis requires more than one image and the classification errors in each individual forest cover map propagate into the forest cover change map. These errors can be minimised by integrating knowledge based rules, e.g. detecting fluctuations in classification during the time-series which cannot occur in real forests. However, the accuracy of "post-classification change detection" was found to be far below the "multidate classification" approach at the Kielder test site. Only 65% of the detected clear-cuts agree with the in situ data (65% of the centres of the SAR polygons were found within the in-situ GIS polygons). In addition, the number and area of felled forest stands was overestimated. The reason for these results can be explained by the limited number of available SAR images (see methodology section). Thus, the available data set (3 scenes) does not allow for a proper knowledge rule based "post-classification change detection".

For operational purposes the choice between both approaches is required to take account for practical issues such as data availability and investigation objectives. Typically the life span of one sensor such as JERS-1 SAR is short compared to the study period of a change detection survey (e.g. GSE FM Russia). That fact prevents optimal data conditions and thus can complicate the application of the "multidate classification" approach. Occasionally no EO data is available and the only source of historical information is a thematic map. Thus, e.g. Kyoto or GSE FM applications are concentrating on "post-classification change detection".

6 Implementation of the Object-Based Classification Approach at the Russian Service case of GSE FM

6.1 Irkutsk Oblast – The Russian Service Case

The Russian Service Case relates to the Irkutsk Oblast (Fig. 1). The Service Case is based on agreement with the Forest Agency of Irkutsk General Survey of Natural Resources (FA of GSNR) of Russian Ministry of Natural Resources (MNR). Current forest monitoring is accomplished by the conduction of periodical forest inventories, generally every 10 to 15 years. Due to frequent forest fires as well as intensive human activities such as clear-cutting and cultivation the Irkutsk territory is characterised by large area changes of forests. Therefore, management and monitoring issues better oblige an annual inventory cycle. Without the use of EO data this challenge is almost impossible to solve as the conduction of ground inventories within this region is rather complicated due to the large size of the region, insufficient infrastructure, and inaccessibility of many areas. Moreover, ground based forest inventory is costly and time consuming.

The Oblast is located in central Siberia (52°-64° N, 96°-118° E) and comprises 739,000 km². The Middle Siberian Plateau in the southern part of the territory is characterised by hills up to 1,700 m. The northern part is plain with heights up to 500 m. Taiga forests (birch, aspen, pine, larch etc.) dominate the Irkutsk Oblast and cover about 82% of the region.

6.2 Implementation of the Object-Based Classification

The GSE FM service case in Russia has a large influence on effective forest monitoring and inventory at regional scale. Recent information on forest extent and changes therein are currently generated using high-resolution ENVISAT ASAR APP IS7 (HV/HH) (5.331 GHz, 43,8° off nadir angle, 12.5 x 12.5 m² pixel spacing) data because of their availability and qualification (forest/non-forest sufficiently separable).

The provided products of this service case include a forest area map, a clear-cut/burned area map and a forest area change map. The forest area map is derived from recently acquired ASAR data. Analysis of backscatter intensities of a one-year time series have resulted in a preferred image acquisition during the thawing season around April (Thiel et al. 2007). For the generation of the other maps archived LANDSAT TM data around year 1990 are utilised. Specifications of the products require a geometric accuracy of an RMS < 30 m and a minimum mapping unit of 1 ha. Both requirements can be fulfilled. The acceptability threshold of the thematic

mapping accuracy is 90% for non-change maps and 85% for change maps, respectively. All products will be implemented into the forest inventory of the FAI of GSNR and are produced within three years for regions of rapid change in the Irkutsk Oblast, comprising a total area of 200,000 km². In the first year an area of about 50,000 km² has already been processed, the service area will be extended by ~20% each year.

The methodology chain comprises pre-processing (calibration, orthorectification, topographic normalisation, and ratio computation), classification, post-classification refinement, manual separation of clear-cuts and burned areas with the forest area map as origin and change map production. For the creation of forest cover maps the principle of the above discussed "multidate classification" has been applied. It only differs by the fact, that not multidate, but multipolarisation data sets were used. One single ASAR scene (three channels: HH/HV/ratio[HH/HV]) provides the input for detecting forest and non-forest areas, whereas non-forest areas form a mixed class consisting of clear-cuts and fire scars. Several samples (15-25, depending on size of forest enterprise and site characteristics) for each class are identified before conducting the supervised classification (nearest neighbour) based on σ^0 in dB. Manual post-classification refinement is necessary to fulfil the requested mapping accuracy. Basing on the forest area map which comprises the mixed class "clear-cuts and burned areas" clear-cut and burned area maps are produced. The separation is based on the object shapes, which differ for fire scars (irregular, rather large) and clear-cuts (squared, rather small) as well as on fire event data. For the production of forest area change maps archived LANDSAT TM scenes from years around 1990 are classified using the same procedure as described above. The classification results are then combined with the forest area maps to create the forest area change maps (Fig. 9).

The quality assessment of the GSE FM products is performed by the Service Provider (SP) and by the End User (thus by FAI of GSNR in that case), whereby the End User is not expected to compute quality measures but to check the product consistency and to carry out some real field comparisons. The mapping accuracy (by SP) was found to be 92.2% (prod. acc. 94.1%, users acc. 90.3%) in average for the non-change maps and 86.1% (prod. acc. 86.9%, users acc. 85.3%) for the change maps. For this assessment high resolution optical and forest inventory data were applied and the stratified random sampling strategy was chosen. All products and specifications have been approved by the End User.
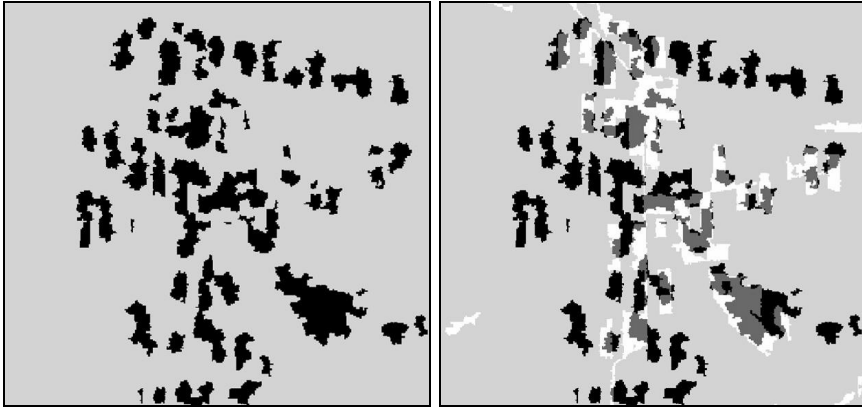
Fig. 9. Left: Forest area map (2006); grey: forest, black: non-forest. Right: Forest area change map (1990-2006); grey: forest, white: regrowth, black: new clear-cuts, dark grey: non-forest (subset size ca. 10 x 11 km)

7 Conclusions and Outlook

Two differing object-based classification and change detection strategies proved being appropriate image analysis approaches for forest cover and forest cover change mapping. In summary, JERS SAR images provide sufficient information for the detection of clear-cutting. Due to the limited time span of seven years, JERS data permitted neither the detection of afforestation nor reforestation (in the sense of the Kyoto Protocol, see IPCC 2003). However, some forest stands with recently planted trees are apparent at most of the JERS test sites. As their backscatter lies close to the threshold that separates forest from non-forest, these stands cannot be clearly assigned to either class (forest/non-forest). The usage of polarimetric L-band SAR data is expected to remedy this deficiency (Israelsson et al. 1994, Rignot et al. 1997). Hence, the effect of this data type has to be investigated in further studies. Cross checks of both classification strategies shall be provided in these investigations.

   One of the developed object-based classification and change detection strategies could be successfully realised at the Russian service case of GSE Forest Monitoring. Amongst others GSE FM aims at the operationalisation of Earth Observation data based services and is already characterised by a high operational level. The object-based classification strategy fits well in this intention. The product list of the Russian service case includes a forest area map, a clear-cut/burned area map and a forest area change map. Re-

cent forest area maps are created using ASAR APP IS 7 (HV/HH) data. To generate historical maps LANDSAT TM data from the 1990s were used. Some manual correction had to be accomplished to meet the demanded mapping accuracy. This in particular refers to the classification of the ASAR data. Recent L-band data such as provided by PALSAR will minimise the manual intervention effort and will be implemented at phase 3 of GSE FM. Products derived from PALSAR can be expected to be even more accurate than those from JERS. This is primarily because of the availability of multiple polarisations and multiple incidence angles. Additionally, the improved geometric resolution allows the detection of smaller-scale forestry activities.

References

Baatz M, Schäpe A (2000) Multiresolution segmentation - an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Blaschke T, Griesebner G (eds) Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000, Herbert Wichmann Verlag, Heidelberg, pp 12–23

Benz UC, Hofmann P, Willhauck G, Langenfelder I, Heynen M (2004) Multiresolution, objectoriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS J. of Photogrammetry and Remote Sensing 58: 239-258

Dobson MC, Ulaby FT, Le Toan T, Beaudoin A, Kasischke ES, Christensen N (1992) Dependence of radar backscatter on coniferous forest biomass. IEEE Trans. Geoscience and Remote Sensing 30(2): 412-415

Freeman A, Saatchi SS (2004) On the Detection of Faraday Rotation in Linearly Polarized L-Band SAR Backscatter Signatures. IEEE Trans. Geosc. Remote Sensing 42 (8): 1607–1116

Intergovernmental Panel on Climate Change (IPCC) (2003) Good practice guidance for land use, land-use change and forestry, IPCC National Greenhouse Gas Inventories Programme UNEP. Penman J, Gytarsky M, Hiraishi T, Krug T, Kruger D, Pipatti R, Buendia L, Miwa K, Ngara T, Tanabe K, Wagner F (eds) Institute for Global Environmental Strategies, Hayama

Israelsson H, Askne J, Sylvander R (1994) Potential of SAR for forest bole volume estimation. Int. J. Remote Sensing 15 (14): 2809-2826

Kasischke E, Bourgeau-Chavez L, French N, Harrell P, Christensen N (1992) Initial observations on using SAR to monitor wildfire scars in boreal forests. Int. J. Remote Sensing 13(18): 3495-3501

Leckie DG, Ranson KJ (1998) Forestry applications using imaging radar. In: Henderson FM, Lewis AJ (eds) Principles and applications of imaging radar, 3rd edn, Wiley, New York, pp 435-510

Rignot E, Salas WA, Skole DA (1997) Mapping deforestation and secondary growth in Rondonia, Brazil, using imaging radar and thematic mapper data. Remote Sensing Env 59: 167-179

Santoro M, Eriksson L, Askne J, Schmullius C (2006) Assessment of stand-wise stem volume retrieval in boreal forest from JERS-1 L-band SAR backscatter. Int. J. Remote Sensing 27: 3425-3454

Schmullius C (ed), Baker J, Balzter H, Davidson M, Eriksson L, Gaveau D, Gluck M, Holz A, Le Toan T, Luckman A, Marschalk U, Mc Callum I, Nilsson S, Orrmalm S, Quegan S, Rauste Y, Roth A, Rozhkov V, Sokolov V, Shvidenko A, Sirro L, Skuding V, Strozzi T, Tansey K, Utsi R, Vietmeier J, Voloshuk L, Wagner W, Wegmüller U, Westin T, Wiesmann A, Yu JJ (2001) SIBERIA - SAR imaging for boreal ecology and radar interferometry applications, Final Report, Friedrich-Schiller-University, Jena

Stussi N, Beaudoin A, Castel T, Gigord P (1995) Radiometric correction of multi-configuration spaceborne SAR data over hilly terrain. Proc. CNES/IEEE Int. Symp. on the Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications, 10-13 October, Toulouse: 347-467

Thiel Ca, Thiel Ch, Riedel T, Schmullius C (2007) Analysis of ASAR APP Time Series over Siberia for Optimising Forest Cover Mapping - A GSE Forest Monitoring Study. Proc. ENVISAT Symposium 23-27 April, Montreux.

Thiel C, Weise C, Eriksson L (2004) Demonstration of Forest Change Monitoring. Final Report on the ESA Study: Demonstration of L-Band Capabilities Using JERS Data, ESTEC Contract No. 18311/04/NL/CB, 109 pp.

Yatabe SM, Leckie DG (1995) Clearcut and forest-type discrimination in satellite SAR imagery. Canadian J. Remote Sensing 21(4): 456-467

Acknowledgements

Chapter 3.5

Pixels to objects to information:
Spatial context to aid in forest characterization
with remote sensing

M.A. Wulder[1*], J.C. White[1], G.J. Hay[2], G. Castilla[2]

[1]Canadian Forest Service (Pacific Forestry Centre), Natural Resources Canada, Victoria, British Columbia, Canada

[2]Foothills Facility for Remote Sensing and GIScience, Department of Geography, University of Calgary, Calgary, Alberta, Canada

[*]Corresponding author:
Mike Wulder
506 West Burnside Rd., Victoria, BC V8Z 1M5;
Phone: 250-363-6090; Fax: 250-363-0775; Email: mwulder@nrcan.gc.ca

ABSTRACT: Forest monitoring information needs span a range of spatial, spectral and temporal scales. Forest management and monitoring are typically enabled through the collection and interpretation of air photos, upon which spatial units are manually delineated representing areas that are homogeneous in attribution and sufficiently distinct from neighboring units. The process of acquiring, processing, and interpreting air photos is well established, understood, and relatively cost effective. As a result, the integration of other data sources or methods into this work-flow must be shown to be of value to those using forest inventory data. For example, new data sources or techniques must provide information that is currently not available from existing data and/or methods, or it must enable cost efficiencies. Traditional forest inventories may be augmented using digital

remote sensing and automated approaches to provide timely information within the inventory cycle, such as disturbance or update information. In particular, image segmentation provides meaningful generalizations of image data to assist in isolating within and between stand conditions, for extrapolating sampled information over landscapes, and to reduce the impact of local radiometric and geometric variability when implementing change detection with high spatial resolution imagery. In this Chapter, we present application examples demonstrating the utility of segmentation for producing forest inventory relevant information from remotely sensed data.

Introduction

Information needs for forest characterization span across a range of spatial scales - both in terms of areal coverage and level of detail required - and across information needs, ranging from resource management and conservation, through to land use planning. The areal extent over which information are required varies from tens to millions of hectares. Remotely sensed data is often seen as a practical and cost effective means to represent forest conditions. An ever expanding suite of air- and space-borne sensors are available that collect data across a wide range of spatial and spectral resolutions. High spatial resolution data may be selected for fine scale characterizations. Alternately for wide-area studies, lower spatial resolution data may be utilized to provide a more generalized perspective. The image extent is typically tied to a given spatial resolution, with smaller pixels forming images with smaller extents (or footprints). For example, Landsat-5 Thematic Mapper (TM) and Landsat-7 Enhanced Thematic Mapper Plus (ETM+) data have a 185 by 185 km image footprint with six 30 m (multispectral; MS) bands and one 15 m (panchromatic; PAN) band; IKONOS has up to a 10 x 10 km footprint with four 4 m (MS) bands and one 1 m (PAN) band. The reader is referred to Franklin et al. (2002) for a more detailed discussion of the relationship between image extent, spatial resolution, and data cost. Trade-offs are typically required between image spatial and temporal resolution that have implications for data selection. Generally, medium resolution satellites such as Landsat, will revisit the same location once every 16 days, while high spatial resolution imagery such as IKONOS and QuickBird have longer revisit period (144 days for Ikonos for true nadir), albeit it can be shortened to 1 to ≈4 days due to the off-nadir acquisition capabilities (up to 30°) of these satellites (Coops et al. 2006).

Forest characterization will often require generalization of detailed large-scale data to produce maps that depict large areas for management purposes. Air photo interpretation involves the delineation of polygons representing areas homogenous in terms of the relevant forest attributes, and is the most common approach for initially generating a forest inventory. The ability to develop meaningful units for management from mid-scale (1:25,000 to 1:50,000) photography provides a cost effective and established means for spatial data generation. Forest inventory is supported by field data collection and calibration of the units and attributes interpreted from the photos. Shortcomings in this interpretive approach include inconsistencies in the interpretation process (both object delineation and attribution), time to produce products, and update cycle. For example, in 1:10,000 to 1:20,000 scale photography, accuracies are generally 70–85% for main species in a stand, but can be lower (Leckie et al. 2003). In addition, inventory data volumes are often enormous, with manual interpretation typically completing 5–15 photos per day. This represents approximately one photo per hour, or 400 new hectares and 10–15 stands per hour (Leckie et al. 1998; data of Leckie and Gillis 1995), with similar numbers for mid-scale photography. Shortcomings notwithstanding, no other currently available data source can provide the same combination of affordability and level of detail.

Development of digital remote sensing techniques that are supportive and complementary to current operational forest inventory approaches are desired. For instance, remotely sensed data can provide information that augments the forest inventory with attributes of interest, produces update (change) information, or enables increased automation of expensive and inconsistent elements of the forest inventory production process. Image segmentation is one such technique capable of supporting forest inventory.

Image segmentation is the partitioning of a digital image into a set of jointly exhaustive and mutually disjoint regions that are more uniform within themselves than when compared to adjacent regions (uniformity being evaluated by a dissimilarity measure with which the partition itself is constructed). There are hundreds of image segmentation algorithms that have been developed, not only for remote sensing applications, but also for computer vision and biomedical imaging. The techniques are so varied that there is no up-to-date review in the literature; the last comprehensive attempt was made more than a decade ago (Pal and Pal 1993). However, in the context of forest inventory, only a few methods have been applied, which are mostly based on *region merging* (e.g. Hagner 1990; Woodcock and Harward 1992; Baatz and Schape 2000; Pekkarinen 2002). This process consists of the sequential aggregation of adjacent regions according to their similarity until a stop criterion is reached, which can be based on ei-

ther a similarity or a size threshold. So far, only one of these methods, the multi-resolution segmentation of Baatz and Schape (2000), has been implemented into commercial software (eCognition; Definiens Imaging GmbH 2002). The segmentations produced with eCognition are typically visually appealing; although, a disadvantage is the need for users to iteratively deduce useful segmentation levels through trial and error (Hay et al. 2003). Further, the algorithms employed in eCognition are proprietary, and are described only thematically. Feature Analyst is an additional commonly used commercially available software package for image segmentation; it too has some subjective and iterative training requirements[1]. A more recently developed algorithm, Size-Constrained Region Merging (SCRM), enables the user to explicitly control the size of the output regions and the complexity of their boundaries (Castilla et al. In Press).

The objective of this Chapter is to demonstrate the utility of image segmentation within a forest inventory context. Consequently, in this Chapter, we present examples of applications where image segmentation routines have been applied to aid in the generalization and attribution of remotely sensed forest data. A comparison of the strengths and weaknesses of differing algorithms is beyond the scope of this communication. Applications include estimating time since forest disturbance, extending a sample of Light Detection and Ranging Data (LIDAR) data across a broader area for the purposes of forest inventory height update, and a segment based comparison of forest change over time. Segmentation 'opportunities' and 'issues' related to individual tree level change detection are also presented, followed by potential uses of segmentation as part of the forest inventory process. Finally, we summarize our findings and indicate future directions and opportunities for segmentation to aid in meeting forest characterization information needs.

Applications

Lidar and spectral data fusion to enable extension of forest inventory attributes

Optical remotely sensed imagery is well suited for capturing horizontally distributed conditions, structures, and changes (Wulder 1998), while LIDAR data are more appropriate for capturing vertically distributed elements of forest structure and change (Lefsky 2002). The integration of optical remotely sensed imagery and LIDAR data provides improved oppor-

---

[1] http://www.featureanalyst.com

tunities to fully characterize forest canopy attributes and dynamics. Medium resolution remotely sensed data such as Landsat is relatively inexpensive to acquire over large areas (Franklin and Wulder 2002), whereas LIDAR covers small areas, at a high cost per unit area (Lim et al. 2002). As a result, these two data types may be combined to generate estimates of stand height over large areas at a reasonable cost (Hudak et al. 2002).

Forest inventories in Canada are typically updated on a 10 year cycle (Gillis and Leckie 1993). Applications requiring up-to-date estimates of height must often use growth and yield modeling to predict changes to height over time, based on a number of other inventory attributes. Wulder and Seemann (2003) presented an approach where image segments generated from large-extent Landsat-5 Thematic Mapper (TM) data were used to extend height estimates from samples of LIDAR data collected with the *Scanning LIDAR Image of Canopies by Echo Recovery* (SLICER) instrument. SLICER records data on canopy height, vertical structure, and ground elevation, collecting 5 full waveform footprints, typically resulting in a narrow-transect (< 50 m). Image segments were generated from Landsat-5 TM bands 1 to 5 and 7 using eCognition's segmentation algorithm (Definiens Imaging GmbH 2002).

Initially, Wulder and Seemann (2003) examined the empirical relationship between LIDAR height estimates and the height attribute of the corresponding forest inventory polygon. This method used the mean height of all the LIDAR hits within each forest inventory polygon and resulted in a regression model with an R^2 of 0.23 and a standard error of 4.15 m. Subsequently, Wulder and Seemann (2003) decomposed the image-based segments generated from the Landsat-5 TM data to the forest inventory polygons (Wulder and Franklin 2001). The result was that each forest inventory polygon was partitioned into spectrally homogenous sub-units, similar to the examples provided in Figure 1. The mean LIDAR height within each of these polygon sub-units was determined, and then an area-weighted mean height was generated for each forest inventory polygon. A regression model built using this area weighted mean LIDAR height, calculated from the within stand image segments, resulted in a R^2 of 0.61 and a standard error of 3.15 m. 80% of the forest inventory polygons had LIDAR height estimates within ± 6 m of the existing inventory height. Independent validation data was used to subsequently test the model, generating a R^2 of 0.67 and a standard error of 3.30 m. In this application, remotely sensed imagery, LIDAR, and forest inventory were combined to estimate stand height over a 7000 km^2 area, based on a 0.48% LIDAR sample that covered 33.9 km^2. The image-based segments were critical for developing an appropriate regression model, and in turn, for extrapolating the LIDAR estimates across the larger area.

Time since disturbance estimation

Modeling of forest carbon dynamics requires precise information regarding when a disturbance occurred and the age of the subsequent regenerated forest. As mentioned earlier, forest inventory data is not always up-to-date. In order to provide information required for modeling purposes, Wulder et al. (2004a) used Landsat-7 ETM+ data and derived tasseled cap transformation (TCT) (Crist and Cicone 1984) values to estimate the age of lodgepole pine (*Pinus contorta*) stands from approximate time of disturbance to 20 years of regeneration. Image segmentation was employed to aid the removal of pixels representative of residual forest and other non-characteristic stand conditions within forest inventory polygons, thereby improving the relationship between Landsat spectral values and the age of the forest stand.

Forest inventory polygons capture a homogenous assemblage of forest attributes (Gillis and Leckie 1993); however, the spectral response within any given forest inventory polygon is highly variable (Wulder 1998) and the disparity in the relationships between spectral values and forest inventory attributes has been demonstrated (Wulder et al. 1999). Thus, when the Landsat-7 ETM+ imagery is segmented, groups of pixels with similar spectral values are generated; when these groups are then combined with the forest inventory polygons using polygon decomposition, the inventory polygons are stratified into sub-units representing the variability in the distribution of trees within the inventory polygon. For example, a stand, as represented by a forest inventory polygon, may have been harvested; however, some residual forest may have been left behind to serve a specified management function (e.g. a wildlife tree patch). This patch of residual forest will likely have spectral properties different from the harvested area surrounding it, and as a result would be segmented into a separate sub-unit within the inventory polygon. In a study by Wulder et al. (2004a), a total of 1305 segments were created, 65% more than the number of polygons in the inventory (809). Examples of four inventory polygons and their associated segments are provided in Figure 1. The thicker vectors in Figure 1 indicate the boundary of the original forest inventory polygon, while the thinner vectors represent the polygon segments.
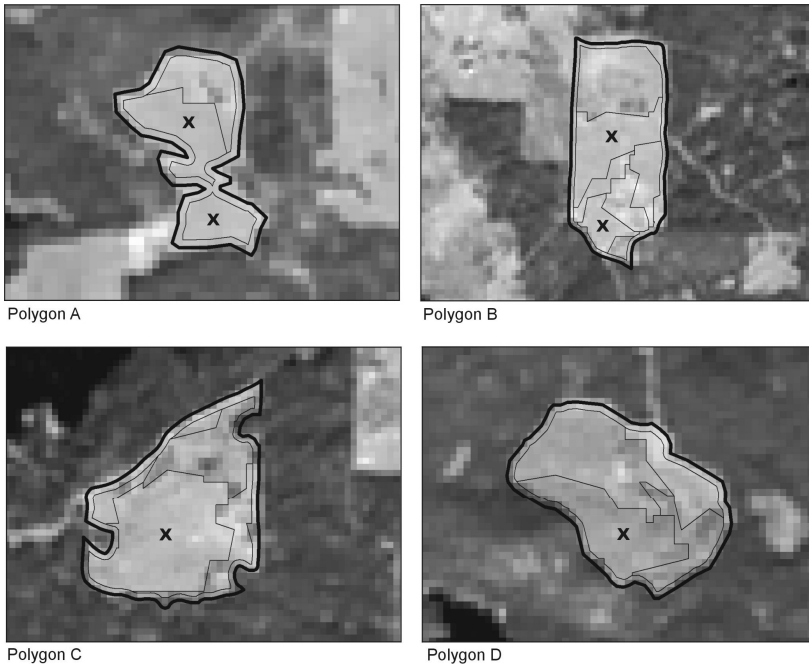
Fig. 1. Illustrative examples of polygon segmentation for four forest inventory polygons. The Landsat ETM+ image used to generate the spectrally homogenous segments is shown in the background

Wulder et al. (2004a) then applied a set of rules to determine which of the polygon sub-units actually represent harvest conditions, and which represent residual stand structure or were not indicative of post-harvest conditions. Landsat bands 3, 4, 5 and 7, along with TCT components brightness, greenness, and wetness were used as independent variables (for background, see Healey et al. 2005, 2006). Stand age, as estimated in the forest inventory, was used as the dependent variable.

For comparative purposes, age estimates were generated using two different approaches. In the first approach, estimates of stand age were based on the mean spectral value of all the pixels in the polygon, including any residual forest and roads or landings within the polygons. In the second approach, the image pixels representing these features and other non-characteristic stand conditions were removed and the mean value of only the remaining segments was used to represent the age of the stand; a segment-based area weighted estimate of age was then generated for the remaining segments in each polygon. In Figure 1, only those segments marked with an 'x' would have been used in height estimation.

The inventory stand age was found to be negatively correlated with Landsat-7 ETM+ spectral response. The strongest correlation to stand age was found using the TCT wetness component (R = 0.56 for the inventory polygons and R = 0.78 for the segments within the polygons). The use of the polygon segments was more successful at generating an estimate of stand age, as indicated by the larger correlation coefficient values and the univariate and multivariate regression models: a step-wise multivariate regression procedure resulted in a R^2 of 0.68 with a standard error of 2.39 years. By comparison, the regression model generated using just the forest inventory polygons had a R^2 of 0.46 with a standard error of 2.83 years. From a carbon modeling perspective, the ability to estimate stand age to within 3 years, where no other current age information is available, provides a useful option for model input. In this application, the image segments served to stratify the forest inventory polygons and improved the relationships between Landsat-7 ETM+ spectral response and inventory age. This experiment adds a new *segment-* or *object-based* dimension to the literature which confirms the validity of empirical modeling, or regression based approaches, to estimate stand age.

Capture of large area forest dynamics

Forest dynamics are characterized by both continuous (i.e., growth) and discontinuous (i.e., disturbance) changes. Wulder et al. (2007) used a combination of LIDAR and remotely sensed imagery to characterize both the horizontal and vertical forest structure over a large area in the boreal forest of Canada. Their study uses two SLICER LIDAR transects, each approximately 600 km in length. The first transect was collected in 1997 and the second transect was collected in 2002. Image segments were generated from Landsat-7 ETM+ imagery using the same procedure as Wulder and Seemann (2003). The image segments provided a spatial framework within which the attributes and temporal dynamics of the forest canopy were estimated and compared.

The detection and mapping of forest disturbance (discontinuous), especially stand replacing disturbances such as fire and harvesting, are operationally and reliably captured with optical remote sensing systems (Cohen and Goward 2004). Conversely, continuous change, such as growth, is more difficult to detect with optical imagery. The use of LIDAR presents opportunities for capture and characterization of this subtle continuous change. Two different approaches for estimating changes in forest attributes were used. The first one was a global approach that emphasized the overall trend in forest change along the entire transect. It was found that

globally, the key canopy attributes, including average canopy height, canopy profile area, and standing stock, were stable, indicating that transect characteristics following the five-year period did not differ. A two-sample t-test confirmed the inseparability of the 1997 and 2002 canopy height ($p =$ 0.08).

The local analysis approach examined segment-based changes in canopy attributes, providing spatially explicit indications of forest growth and depletion. As expected, the difference in the magnitude of the changes was greater for depletions than it was for growth, but was less spatially extensive; 84% of segments intercepted by both LIDAR transects either had no change or a very small average decrease in canopy height. While growth tends to occur incrementally over broad areas (Figure 2), depletions are typically dramatic and spatially constrained (Figure 3). The fact that depletions occurred over a smaller spatial extent, but with a greater magnitude than growth, nullified the detection of positive growth that occurred in smaller increments over a much larger area. As a result, the conclusion reached with the global approach - that no change occurred between 1997 and 2002 - is misleading. This result suggests that a local approach is preferable for characterizing change from LIDAR transects, particularly given the complexity and natural variability inherent in most forest ecosystems.

In this application, segments were used as a proxy for homogenous vegetation units (Wulder et al. 2007) and facilitated the reporting and comparison of forestry canopy attributes derived from the LIDAR transects. Figures 2 and 3 represent examples of canopy height profiles of segments experiencing forest growth and forest depletion, respectively. If forest inventory polygons are available, they can serve as the spatial unit; however, for this study, the inventory was not readily available, and in many parts of northern Canada, no spatially extensive forest inventory has ever been conducted.
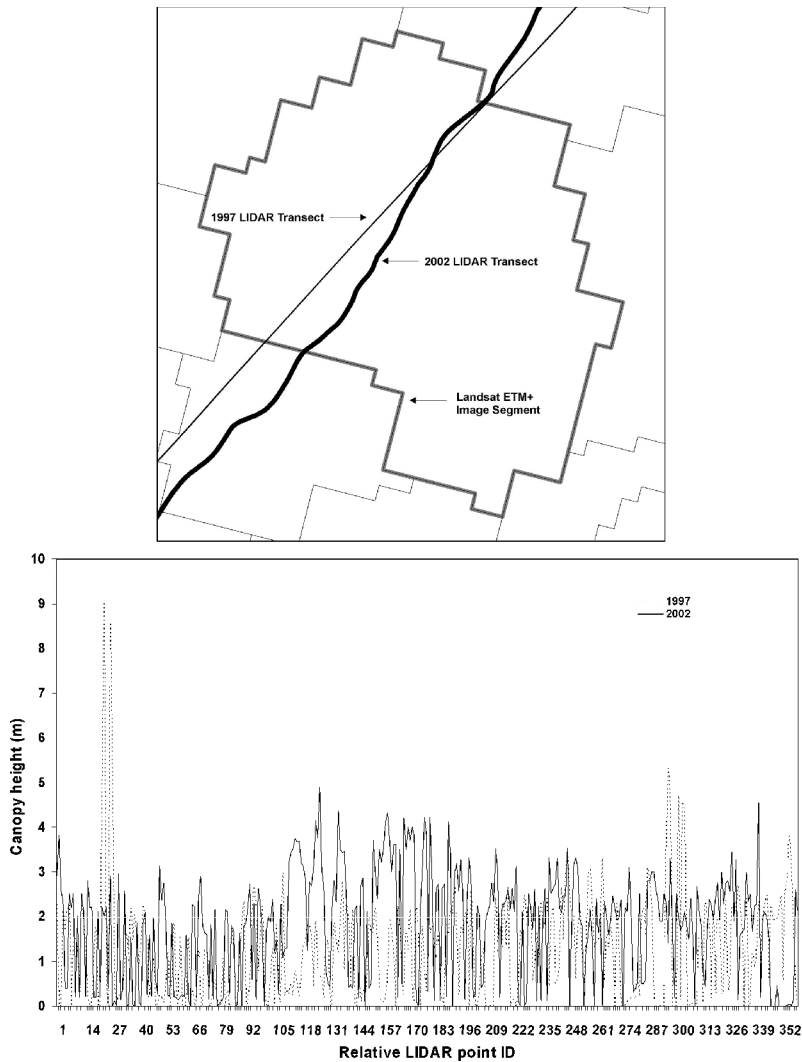
Fig. 2. Segment-based canopy height comparison (growth case) (scene: 4119, segment: 111786; greatest between-transect distance: 60 m)

The image-based segments provided a spatial context to account for the disparity in the transect flight paths, which were not exactly coincident. Furthermore, the segments facilitated the investigation of the spatial properties of the changes in forest attributes, which is critical for determining the scale of the changes, and for inferring the processes that cause the

changes. The segments are therefore essential for capturing the variability in change events along the LIDAR transects.

Fig. 3. Segment-based canopy height comparison (depletion case) (scene: 4222, segment: 80081; greatest between-transect distance: 100m)

Discussion

The applications presented demonstrate the utility of image segmentation for providing spatial context, which in turn facilitates analysis of the spatial properties of forest attributes and change over time. Segments can be used to stratify image data into spectrally homogenous units, which are subsequently used as surrogates for forest polygons, on the assumption that they will also be relatively homogeneous in terms of the relevant attributes. The findings presented in this communication suggest that in the absence of better alternatives, segments derived from optical imagery may be used to constrain analysis, reduce data variability, and aid in extrapolating relationships across larger areas.

There are two other potential areas where segmentation could prove useful. The first is in the context of high spatial resolution change detection (see also Hall and Hay 2003). Since high spatial resolution space-borne sensors have only been in commercial operation since 1999, the uncommon activity of acquiring a satellite-based high spatial resolution temporal sequence has seldom been possible due to limited archiving, cloud cover, and high cost. Furthermore, due to the occurrence of a large number of small objects all creating high contrast in a high spatial resolution satellite image (i.e., trees and shadows), as well as off-nadir look angles increasing horizontal distortions (buildings, bridges), traditional pixel-based change metrics fail to operate successfully (Im and Jensen 2005; Niemeyer and Canty 2003). This is especially true in forests, as automated crown delineation using either high spatial resolution satellite imagery (Johansen and Phinn 2006) or small-format aerial photography (Key et al. 2001) has been met with limited success with the exception of very high spatial resolution (<30cm) images (Pouliot et al. 2002). Figure 4 illustrates the impact that different viewing azimuths can have on forest image-objects. Here, QuickBird scenes collected over the same area of mature lodgepole pine forest in four consecutive years, each exhibits a unique satellite acquisition viewing geometry As a result, trying to identify the same "tree object" from these four dates of imagery is a non-trivial, if not impossible task.

Figure 5 demonstrates how image segmentation could provide a potential solution to the problem of high spatial resolution change detection. In this figure, we see how the multispectral bands of QuickBird can be used to derive a robust (e.g. meaningful in a vegetation context) image segmentation. These segments can then be overlaid onto the higher spatial resolution panchromatic QuickBird band, and local maxima filtering can be used to identify individual trees (Wulder et al. 2004b) within each segment. While difficulties are found in tracking individual tree crowns over time,

the organization of the tree-crown objects within upper-level segments provides a means for generalization and monitoring.
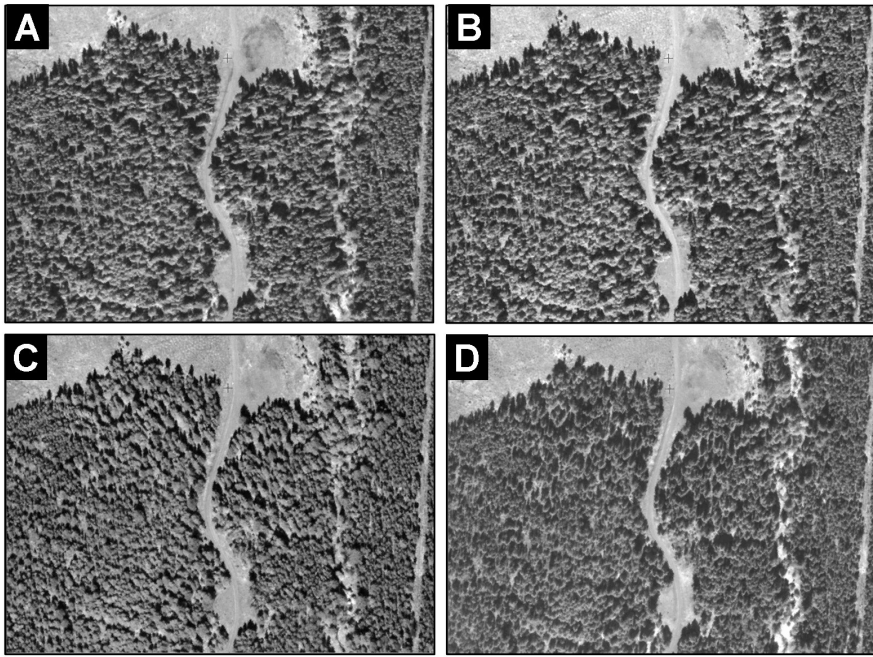
Fig. 4. Four dates of QuickBird imagery collected with different satellite azimuth angles: A. 2003, 51°; B. 2004, 44°; C. 2005, 129°; D. 2006, 353°
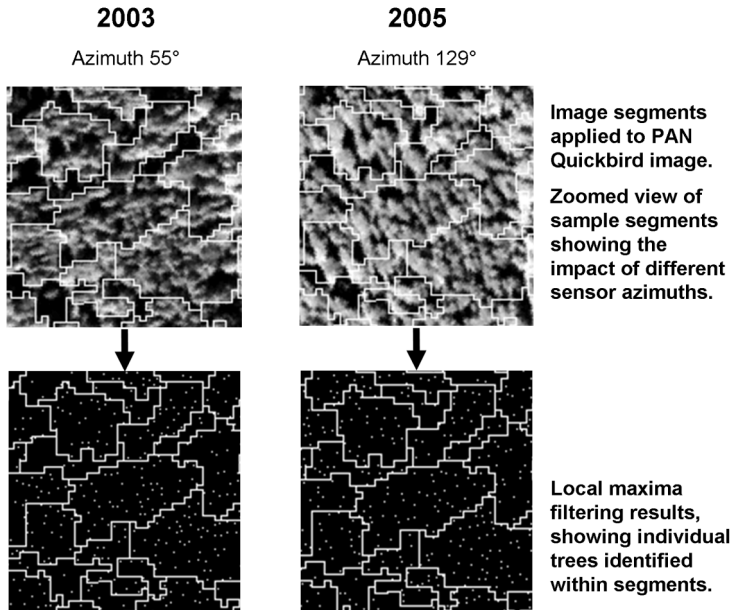
Fig. 5. Multispectral QuickBird imagery (representing 2003 conditions) is segmented, the segments (illustrated here) are transferred to the panchromatic images, with local maxima filtering used to identify individual trees

The second potential use of segmentation is in the context of delineation for forest inventory. As mentioned earlier in this Chapter, air photo delineation and interpretation are the foundation of forest inventory. This process is largely manual and is labor-intensive. There is an opportunity for automated segmentation to augment manual interpretation methods and provide more timely, consistent, and accurate delineations (Hay et al. 2005). Table 1 generalizes some of the advantages and disadvantages of manual and automated methods. Figure 6 illustrates an example of an IKONOS image which has been segmented using the SCRM algorithm mentioned earlier. The polygon in the centre (A) is an example of an effective automated delineation – a homogenous stand, with consistent tone and texture. The objective of automated segmentation in the short-term is not to replace human interpreters, but to aid them - since the interpretation of inventory attributes necessitates human expertise. As Leckie et al. (1998) advocate, any new tools designed to aid photo interpretation of forests must be simple to apply, not require expensive equipment, not substantially alter the production process, nor involve inordinate fine-tuning by the interpreter.

Table 1. Comparative advantages and disadvantages of manual delineation and automated segmentation.

Method	Advantages	Disadvantages
Manual delineation	• Established standard • Delineation and attribution can be achieved in a single step	• Slow • Costly • Scarcity of skilled interpreters • Prone to show inconsistencies between different interpreters or even the same interpreter at different times • Highly subjective and hence no good for monitoring
Segmentation	• Faster • Cheaper • More consistent • Less subjective and more repeatable and hence better for monitoring.	• Likely to require manual correction of some arcs/polygons • Efficiency rapidly decreasing with the amount of manual tweaking required • Undesired results in areas with low contrast or where different appearance does not imply different meaning
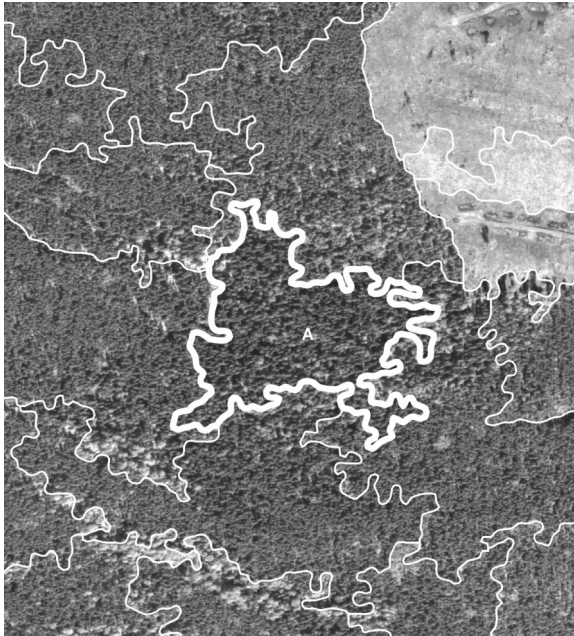
Fig. 6. Automated segmentation of an IKONOS panchromatic image

Conclusion

This chapter has demonstrated the utility of image segments to support a broad range of forestry and forest monitoring applications. It is our view that segmentation can offer a means to augment established forest inventory data sources and processes such as air photos and delineation, and also as a tool for integrating new data sources such as digital remotely sensed data and LIDAR into existing forest inventory contexts. The conversion from pixels to objects serves to provide us with the strengths from remote sensing, GIS, and image processing to produce valuable and useful forestry information.

References

Baatz M, Schape A (2000) Multiresolution segmentation : An optimization approach for high quality multi-scale image segmentation. In: Strobl J Blaschke T (Eds) Angewandte Geographische Informationsverarbeitung XII, Beiträge

zum AGIT-Symposium Salzburg 2000. Wichmann-Verlag, Heidelberg, pp 12-23

Castilla G, Hay GJ, Ruiz, JR (In Press) Size-constrained region merging (SCRM): an automated delineation tool for assisted photointerpretation. Photogrammetric Engineering and Remote Sensing. Accepted for publication June 16, 2006

Cohen W, Goward S (2004) Landsat's role in ecological application of remote sensing. Bioscience 54:535-545

Coops NC, Wulder MA, White JC (2006) Identifying and describing forest disturbance and spatial pattern : Data selection issues and methodological implications. In: Wulder MA, Franklin SE (Eds) Understanding forest disturbance and spatial pattern : Remote sensing and GIS approaches. Taylor and Francis, Boca Raton, Florida, USA, pp 31-62

Crist EP, Cicone RC (1984) Application of the tasseled cap concept to simulated Thematic Mapper data. Photogrammetric Engineering and Remote Sensing, 51:327-331

Definiens Imaging GmbH (2002) eCognition User's Guide. Version 2.1. Definiens Image GmbH, München, Germany. 237p

Franklin SE, Wulder MA (2002) Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas. Progress in Physical Geography 26:173-205

Franklin SE, Lavigne MB, Wulder MA, Stenhouse GB (2002) Change detection and landscape structure mapping using remote sensing. Forestry Chronicle 78:618-625

Gillis M, Leckie D (1993) Forest inventory mapping procedures across Canada. Petawawa National Forestry Institute (PNFI), Forestry Canada, Chalk River, Ontario Information Report PI-X-114, 79p

Hagner O (1990) Computer aided forest stand delineation and inventory based on satellite remote sensing. Proceedings of SNS/IUFRO workshop on the usability of remote sensing for forest inventory and planning. Swedish University of Agricultural Sciences, Umeå, Sweden.

Hall O, Hay GJ (2003) A Multiscale Object-specific Approach to Digital Change Detection. International Journal of Applied Earth Observation and Geoinformation 4(4): 311-327

Hay G, Blaschke JT, Marceau DJ, Bouchard A (2003) A comparison of three image-object methods for the multiscale analysis of landscape structure. Photogrammetry and Remote Sensing 57:327-345

Hay GJ, Castilla G, Wulder MA, Ruiz JR (2005) An automated object-based approach for the multiscale image segmentation of forest scenes. International Journal of Applied Earth Observation and Geoinformation 7:339-359

Healey SP, Cohen WB, Zhiqiang Y, Krankina ON (2005) Comparison of tasseled cap-based Landsat data structures for use in forest disturbance detection. Remote Sensing of Environment 97:301-310

Healey SP, Cohen WB, Zhiqiang Y, Kennedy RE (2006) Remotely sensed data in the mapping of forest harvest patterns. Chapter 3. In: Wulder MA Franklin SE (Eds) Understanding forest disturbance and spatial pattern: Remote sensing and GIS approaches, Taylor and Francis, Boca Raton, Florida, USA, 264p

Hudak AT, Lefsky MA, Cohen W, Berterrectche M (2002) Integration of lidar and Landsat ETM+ for estimating and mapping forest canopy height. Remote Sensing of Environment 82 :397-416

Im J, Jensen, JR (2005) A change detection model based on neighborhood correlation image analysis and decision tree classification. Remote Sensing of Environment 99:326-340

Johansen K, Phinn S (2006) Mapping structural parameters and species composition of riparian vegetation using IKONOS and Landsat ETM+ Data in Australian tropical savannahs. Photogrammetric Engineering and Remote Sensing 72:71-80

Key T, Warner TA, McGraw JB, Fajvan MA (2001) A Comparison of Multispectral and Multitemporal Information in High Spatial Resolution Imagery for Classification of Individual Tree Species in a Temperate Hardwood Forest. Remote Sensing of Environment 75:100-112

Leckie DG, Gillis MD (1995) Forest inventory in Canada with emphasis on map production. Forestry Chronicle 71(1):74–88

Leckie DG, Gillis MD, Gougeon F, Lodin M, Wakelin J, Yuan X (1998) Computer-assisted photointerpretation aids to forest inventory mapping: some possible approaches. In: Hill DA, Leckie DG (Eds) Proceedings of the International Forum: Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, Victoria, BC, February 10–12. Canadian Forest Service, Pacific Forestry Center, Victoria, BC, pp. 335–343

Leckie DG, Gougeon FA, Walsworth N, Paradine D (2003) Stand delineation and composition estimation using semi-automated individual tree crown analysis. Remote Sensing of Environment 85:355–369

Lefsky MA, Cohen WB, Parker GG, Harding DJ (2002) Lidar remote sensing for ecosystem studies. BioScience 52:19-30

Lim K, Treitz P, Wulder M, St-Onge B, Flood M (2002) LIDAR remote sensing of forest structure. Progress in Physical Geography 27:88-106

Neimeyer I, Canty MJ (2003) Pixel-based and object-oriented change detection using high-resolution imagery. Online: http://www.definiens-imaging.com/documents/publications/esarda03_nieca.pdf/

Pal NR, Pal SK (1993) A review on image segmentation techniques. Pattern Recognition 26:1277-1294

Pekkarinen A (2002) A method for the segmentation of very high spatial resolution images of forested landscapes. International Journal of Remote Sensing 23:2817-2836

Pouliot DA, King DJ, Bell FW, Pitt DG (2002) Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. Remote Sensing of Environment 82:322-334

Woodcock CE, Harward VJ (1992) Nested-hierarchical scene models and image segmentation. International Journal of Remote Sensing 13:3167-3187

Wulder M (1998) Optical remote sensing techniques fro the assessment of forest inventory and biophysical parameters. Progress in Physical Geography 22:449-476

Wulder, MA; Magnussen S; Boudewyn P, Seemann, D (1999) Spectral variability related to forest inventory polygons stored within a GIS. IUFRO Conference on Remote Sensing and Forest Monitoring; June 1-3, 1999, Rogow, Poland. pp. 141-153.

Wulder MA, Franklin SE (2001) Polygon decomposition with remotely sensed data: Rationale, methods, and applications. Geomatica 55:11-21

Wulder MA, Seemann D (2003) Forest inventory height update through the integration of LIDAR data with segmented Landsat imagery. Canadian Journal of Remote Sensing 29:536-543

Wulder MA, Skakun R, Kurz W, White JC (2004a) Estimating time since forest disturbance using segmented Landsat ETM+ imagery. Remote Sensing of Environment 93:179-187

Wulder MA, White JC, Niemann KO, Nelson T (2004b) Comparison of airborne and satellite high spatial resolution data for the identification of individual trees with local maxima filtering. International Journal of Remote Sensing, 25 :2225-2232

Wulder MA, Han T, White JC, Sweda T, Tsuzuki H (2007) Integrating profiling LIDAR with Landsat data for regional boreal forest canopy attribute estimation and change characterization. Remote Sensing of Environment 110: 123-137.

Section 4

Automated classification, mapping and updating: environmental resource management and agriculture

Chapter 4.1

Object-oriented oil spill contamination mapping in West Siberia with Quickbird data

S. Hese, C. Schmullius

Friedrich-Schiller University Jena, Institute of Geography, Department for Earth Observation, Löbdergraben 32, 07743 Jena,
Soeren.hese@uni-jena.de

KEYWORDS: Oil spills, context classification, OSCaR, pollution, contamination, Siberia

ABSTRACT: The work presented is part of the OSCaR pilot study (Oil Spill Contamination and mapping in Russia) and is co-financed by the International Office of the Federal Ministry of Education and Research (BMBF) Germany as part of the Core-to-Core activities on "The Symptoms of Environmental Change in the Siberian Permafrost Region" with the Japan Society of the Promotion of Science (JSPS). This paper presents concepts for an object-based mapping and classification system for terrestrial oil spill pollution in West-Siberia using Quickbird data. An object oriented classification system is created to map contaminated soils and vegetation using spectral information, shape and context information. Due to the limited spectral resolution of Quickbird data context information is used as an additional feature. The distance to industrial land use and infrastructure objects is utilized to increase the classification accuracy. Validation of the results is done with field data from the Russian partners at the Yugra State University in Khanty-Mansiyskiy.

1 Introduction

Research on the application of Earth observation data and image processing methods for oil spill detection concentrated in the past on marine pollution scenarios. For marine and coastal applications various methods and

results have been published using SAR data to monitor extent, type and drift of oil pollution (e.g. in Pedersen et al. 1996, Wismann et al. 1998, Espedal and Wahl 1999, Jones 2001, Fiscella et al. 2000, Lu 2003, Brekke and Solberg 2005). Terrestrial oil spill pollution did not receive very much attention. This is due to the regional and small scale character of terrestrial oil spill contaminations often also complicated by mixed spectral signatures of recovering vegetation, dead vegetation and contaminated soils. Hörig et al. (2001) analysed the spectral properties of oil contaminated soils and sands using hyperspectral Hymap data and found specific absorption features in the SWIR region of the spectrum. Salem and Kafatos (2005) investigated the potential of spectral linear unmixing for delineating oil contamination types in AVIRIS hyperspectral data.

With increasing demand on the global markets for crude oil it can be expected that the environmental impact for areas with intensive production of oil and gas will become a major issue in the near future. Earth observation can deliver precise information about the state and change of the ecosystem in these regions.

The presented work of the OSCaR project (Oil Spill Contamination mapping in Russia) concentrates on high spatial resolution data to detect small scale contaminations and later to precisely date the oil spill events with multitemporal data. The methods utilised in this work concentrate on object oriented image processing techniques.

The Khanty-Mansiyskiy area in West Siberia is one of the most important territories for the Russian oil and gas production with 58% of the total Russian oil production and being on the 3rd place with its national gas production (IWACO 2001). The Russian Federation belongs to the top 5 energy producers with Germany and other European countries being the major importers. West Siberia is the oldest oil and gas region mainly exploited by Russian privatised companies (LUKoil, Surneftegaz, Yukos, Sidanko, Tatneft, Tyumen Oil (TNK)). Large areas are polluted by oil and waste water from pipeline leakages with heavy direct impact on underground and surface water quality, ecological conditions and quality of living (**Fig. 1.**). The IWACO Report (from 2001) states that about 700 000 to 840 000 hectares in West Siberia are oil polluted – a much larger area than indicated by the government or oil company statistics.

The region is largely covered by taiga and tundra forest with sub arctic to continental climate and areas of permafrost with annual precipitation between 400 to 500 mm. The area has low nutrient peat based soils with long biological recovery times and includes important habitats for endangered fauna. Due to low evaporation and low temperatures, lack of drainage and small infiltration rates large wetlands are formed. Thermokarst occurs due to melting ground ice (pattern ground) and thaw lakes and thermokarst pits

can occur. The geological structure of the West Siberian basin is dominated by lower cretaceous and jurassic sections. The oil and gas resources are found in stratigraphic and structural traps that extend into the Kara Sea region.

Population is concentrated in a number of urban and industrial areas, which have developed over the last 50 years due to a strong immigration.

Fig. 1. Oil spill polluted tundra in the Samotlor oil field, West Siberia (S. Cejchan, BFH Hamburg)

The major environmental and social impacts come from activities like:

1.   pipeline breaks, spills, and pipeline accidents,
2.   deposition of oily mud, drilling and production wastes,
3.   chemical waste disposal and leaking storages,
4.   emissions of hydrocarbons and greenhouse gases from flaring and venting of gases and oily waste burning and
5.   inadequate emergency planning and under-developed awareness of environmental impact and remediation measures.

In the Khanty Mansiysk district more than 62000 oil wells have been drilled and according to sources from IWACO (2001) 64000 km of pipelines have been constructed.

The magnitude of the oils spills is very difficult to calculate. Accurate and updated data on recent numbers is very hard to obtain. According to different sources about 2% of the total oil quantity produced is spilled into

the tundra. The average oil spill loses about 2 tons of oil and covers about 1000 sqm.

2. The OSCaR pilot study (Oil Spill Contamination Mapping in Russia)

The OSCaR pilot study project was initiated in 2005 as part of BMBF financed permafrost degradation proposal preparation meetings in Russia (Challenges of Permafrost Degradation of Siberian Soils to Science and Technology) and Core2Core activities with the Japanese JSPS programme in 2005.

The international office of the German Ministry for Education and Science co-financed OSCaR with funding for Earth observation data of the Khanty Mansiysk area (Landsat and Quickbird data). The main goal of OSCaR was to test spatial very high resolution multispectral data for oil spill contamination mapping with advanced image processing algorithms. The methodological focus is on high spatial resolution data analysis. Object oriented image analysis has been carried out to link the spectral characteristics of oil spill objects to secondary contextual image object features that have a relation to oil spills (infrastructure, drilling platforms, pipelines, waste water reservoirs or drilling mud reservoirs). Post classification analysis of specific objects has to be performed to identify the structural identity of oil spills and the related objects.

In a second stage in OSCaR the development of the region is analysed using multi temporal data with lower spatial but with high temporal resolution. Changes will be mapped starting in the early 80s and with 2-3 year steps between the datasets until 2005. The main interests are 1. to identify area and position of larger oil spills and 2. to map changes in industrial structures and infrastructure (increase of oil wells, construction of new pipelines etc.). Identification of oil spills in multitemporal data is important for dating of spill events and large oil pipeline leakages.

3 Data

For the OSCaR project Quickbird data was selected covering an approximate 20x16 km subset of a region in the Khanty Mansiysk area north of Surgut in West Siberia. A full "Basic Set" Quickbird data take was ordered from the archive from 2004 (acquisition date 2004-09-27) including the full resolution multispectral information (2.72 m) and a panchromatic data

layer with 0.68 m spatial resolution (Table 1, **Fig. 2.**). The data was imported and georeferenced using the RPCs (Rational Polynomial Coefficients) provided by Digital Globe without applying ortho correction.
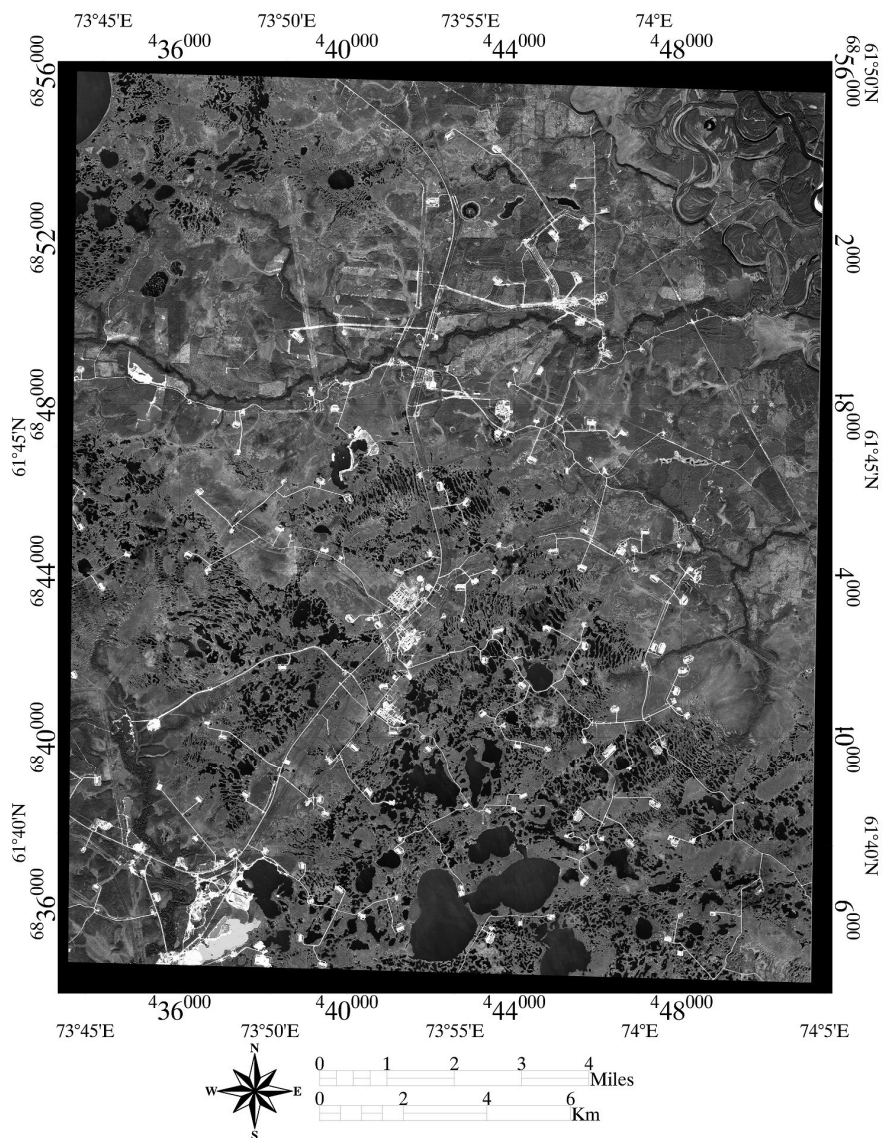
Fig. 2. Quickbird data set (Basic Set product, Digital Globe Catalog ID: 101001000348E202 (27. September 2004), reprojected to UTM43, WGS84
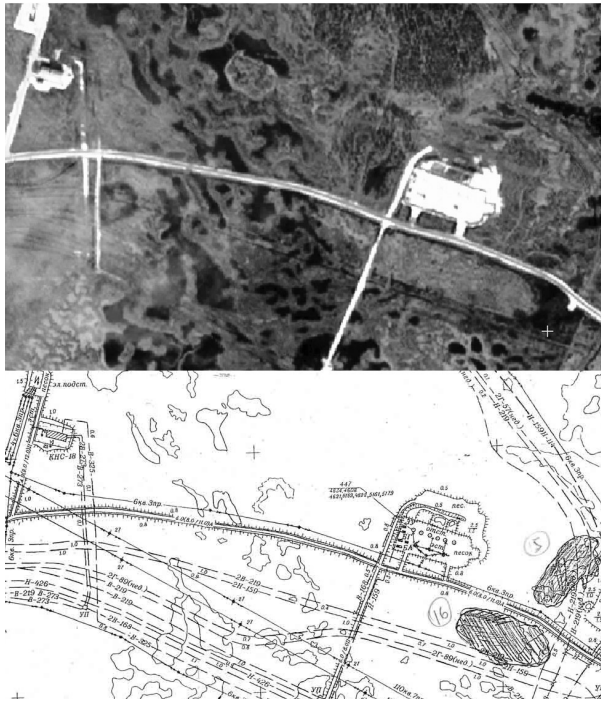
Fig. 3. Analogue information (expert identification) about position and extent of oil pollution compared with (above) Quickbird data

The data was interpolated to the appropriate spatial resolution with a bilinear interpolation to UTM 43 (WGS84) in 16 Bit radiometric resolution (nearest neighbour interpolation was not performed as no pixel based analysis was planned and smooth object geometry was a priority).

Landsat ETM+ (Enhanced Thematic Mapper) and Landsat TM5 (Thematic Mapper) data were ordered for the path/row sets 156/17 and 157/17 with a temporal coverage for the years 1987, 1988, 1990, 1995, 2000, 2001 and 2003. Ground information was provided by the Russian partners from the Yugra State University in Khanty-Mansiyskiy as digital maps with indicated dates and extents of oil spill events and digitized information about infrastructure and the position of oil wells. Comparing the expert interpretation with Quickbird data a mismatch was evident (Fig. 3.). The ground information was therefore used as an orientation and not digitized for training and validation.

It should be noted that precise ground data is very difficult to obtain for intensively managed industrial areas in this region. The companies that are active in this region are not directly interested in external validations of the

gas- and oil production related soil contamination by non-russian institutions. The limited ground data was provided as a cooperation by the University of Khanty-Mansiyskiy.

Table 1. Quickbird data

Quickbird data	Digital Globe
Date	2004 09 27
Cloud Cover	1 %
Catalog ID	101001000348E202
Spatial Resolution	0.68 / 2.72 m
Off-Nadir	19 degrees

4 Methods

The object-based strategy for data classification (Baatz and Schäpe 1999 and Benz et al. 2004) uses as a first stage a segmentation into different scales of image object primitives according to spatial and spectral features. The segmentation is a bottom up region merging technique starting with one pixel sized objects. In numerous subsequent steps smaller objects are merged into bigger objects minimizing the weighted heterogeneity of resulting objects using the size and a parameter of heterogeneity (local optimization procedure) (Benz et al. 2004). This concept has the advantage to account for contextual information using image objects instead of the pixel based concept used frequently as the basic element in Earth observation image analysis. In a second stage rule-based decisions can be used to classify the image objects. Class based feature definitions (integrating a post classification analysis) are possible as well as the inheritance of class descriptions to form a class hierarchy. Image processing tasks can be performed using vector shape and vector characteristics, increasing the flexibility of the image processing concept and integrating GIS-like data queries in an attribute database directly into the image processing and analysis approach. New attributes like object shape or structural characteristics, e.g. distance to other objects can be used.

Various approaches applied object oriented methods to urban applications (Damm et al. 2005; Grenzdörfer 2005, Argialas and Derzekos 2003), biotope type classification (Leser 2002) and forest applications: Mitri and Gitas (2002) developed an object oriented classification model for burned area mapping. Flanders et al. (2003) tested the object oriented approach for forest cut block delineation. Hese et al. (2005) used contextual information to classify forest cover change patterns and Chubey et al. (2006) analyzed

object oriented procedures for forest inventory parameters from Ikonos data.

Advantages over pixel-based approaches have been published mainly using very high spatial resolution airborne or orbital Earth observation data. The primary advantage of reducing the spectral variability in very high spatial resolution data sets (spatial resolution better 5 m) is however only one aspect of object oriented image analysis.

In this work object oriented image analysis has been carried out to link the spectral characteristics of oil spill objects to secondary image object features that have a contextual relation to oil spills (e.g. infrastructure objects, oil well objects, pipeline objects and waste water reservoir objects or drilling mud reservoir objects). This is done to overcome the limitations of the spectral resolution with Quickbird data. The classification of these secondary objects is done in different segmentation scales. Post classification analysis of specific objects has to be performed to identify the structural identity of oil spills and related objects.
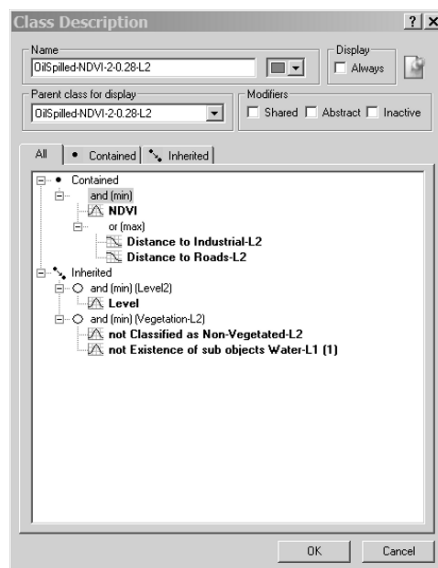
Fig. 4. Class description for oil contaminated objects using NDVI thresholds and class related distance functions to industrial objects and road objects

The dataset was segmented into two different layers (a third layer was used experimentally). A hierarchical class description was build that classified water bodies and vegetation cover types in the finest level and infrastructure, roads and industrial sites in the coarser segmentation level (Fig.

6 and Fig. 7). Road objects and industrial classes were differentiated with object shape features and with spectral characteristics. Oil contaminated areas were mapped with a thresholding of NDVI calculations into three different vegetation sub classes ranging from healthy vegetation to heavily polluted vegetation. The correlation between NDVI and oil spills is based on the reduced amount of healthy vegetation on oil contaminated soils. This is clearly visible in the NIR with a reduced amount of reflectance in polluted areas. Water bodies and non-vegetated areas have been masked out to avoid the overlap with non-vegetated areas through inverted expressions (Fig. 4.).

To increase the accuracy of the classification class related feature sets were designed that introduce distance in relation to the class infrastructure and industrial areas as a characteristic object property of oil spill objects. Oil polluted areas are only classified in a specific distance range of 0 to 500 m to road network objects, oil production platforms objects or other industrial objects. A typical example of oil polluted vegetation in direct neighbourhood of oil wells is shown in Fig. 5. The main problem also discussed in chapter 5 is the changing status of regenerating vegetation and soils after the pollution event.

Quickbird MS   500 m   Quickbird Pan
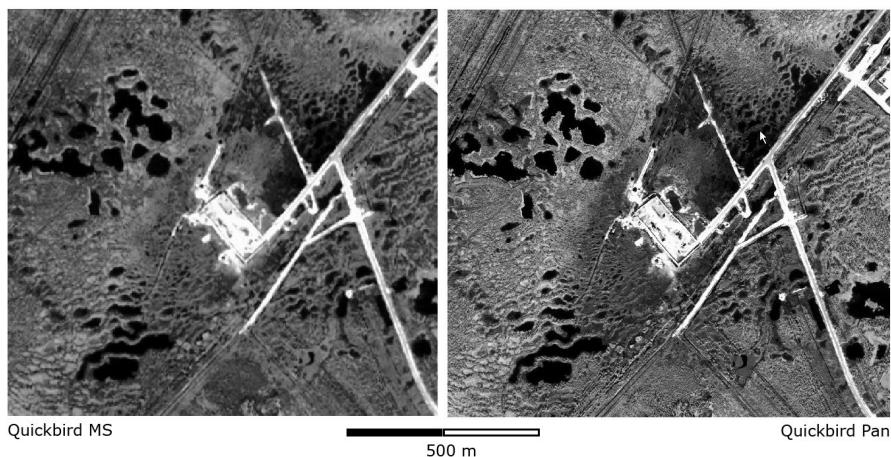
Fig. 5. Quickbird multispectral (2.5 m, RGB 4-3-2) and panchromatic (0.6 m) data subsets of an oil spill area connected to a drilling platform area and infrastructure image objects (West Siberia, Khanty Mansiysk district)
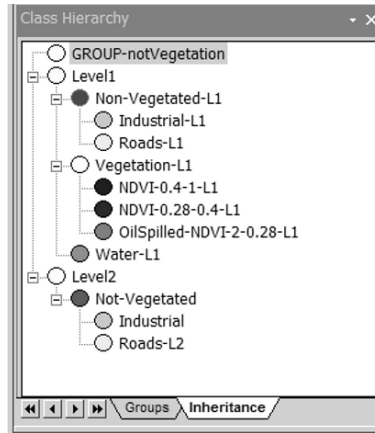
Fig. 6. Class hierarchy for the object oriented classification of oil spills using two segmentation levels and grouping of classes (groups hierarchy)

Validation of the classification results was done using information from expert identifications on analogue topographic maps (compare with Fig. 3). The analogue maps were scanned and used as Jpeg tiles for a more accurate delineation in the georeferenced Quickbird data. The original expert identification was spatially imprecise and was therefore only used as an indication for a contaminated object. For the final validation a pixel based evaluation scheme was set up in a raster GIS environment. A confusion mask was created for all final child classes of the class hierarchy (table 2). An evaluation of the classification accuracy using (segmented) objects as test areas showed even higher accuracy values compared with the pixel based accuracy assessment. The object-based validation (comparing classified objects with test objects) however was rejected for this analysis because the test areas for the evaluation were found not identical with the segmented image objects and using image objects introduces artificial agreements between ground reference information and mapping results as both data layers are based on the primary segmentation.

Table 2. Pixel-based validation of classification results

Validation Class	Pixels	Unclassified	Roads	Industrial	Tundra	Oil spilled	Water
Roads	8092	0.4	60.0	37.2	0.5	2.0	0.0
Industrial	34466	0.8	2.2	89.6	0.0	5.7	1.7
Tundra	37049	0.0	0.0	0.6	99.3	0.1	0.0
Oilspilled	37285	0.0	0.0	0.3	4.5	94.9	0.2
Water	17918	0.0	0.0	0.0	0.0	0.0	100
% Area		1,2	1,35	3,89	61,76	11.37	20,44
Average:88.76%		Kappa: 0.9122		Confidence Level: 95%: 0.912 +/- 0.00174			
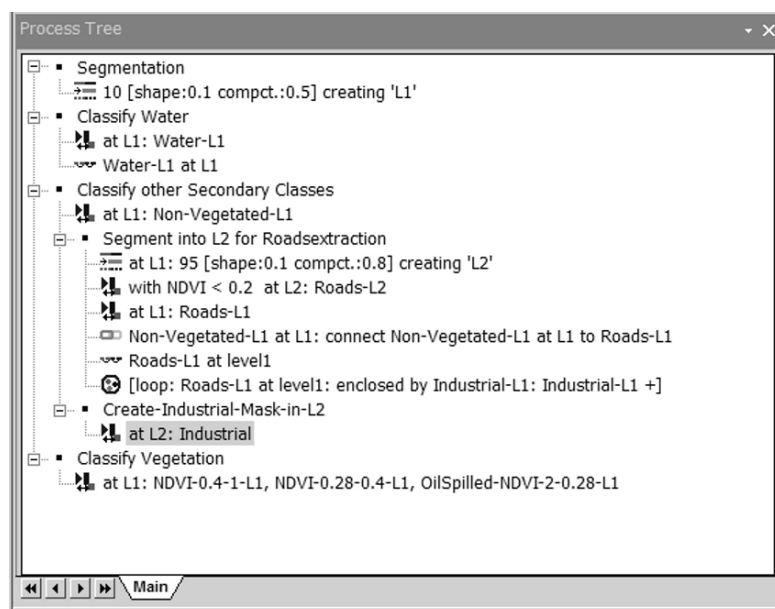
Fig. 7. The Process Tree information for the primary object definition, the re-segmentation of objects (object fusion, creation of the secondary segmentation level) and for the classification usi*ng the class hierarchy (implemented in De-finiens Developer Software ver 6)*

5 Results and Discussion

For this study (using only a small subset of the available Quickbird data) a high percentage (more than 10 %) of the area was found to be oil spill con-taminated. Of the 25856 ha of the subset about 1298 are occupied by in-

dustrial objects, 3120 ha are oil spilled and 6166 ha are open water bodies. The amount of water areas indicates that also water bodies are probably affected by oil spill pollution. The detection of water oil spills has been neglected in this work. The developed class hierarchy will be refined and applied to the complete dataset.

The classification results were validated with reference information from local experts. The classification did not differentiate various regeneration stages of the vegetation on contaminated soils. The exact dates of the contamination events would have been needed for a more detailed differentiation of the spectral signal. In this work a threshold was defined for contaminated surface types and the selection of this threshold was dominated by the ground information available. The comparably high accuracy in table 2 can be explained with the simple classification hierarchy with only five different child classes in the final hierarchical level. Reference information for contaminated soils and vegetation surfaces was not sub differentiated in age classes. This results in a broad class "Oil spilled". Some classification errors were detected between road objects and industrial objects. This confusion (user- and producer accuracy is reduced) is explained with very similar spectral properties. The differentiation was done on the basis of object shape features but for the context classification (of oil spilled objects) the differentiation into roads and industrial objects was not directly needed.

Ground data is difficult to get for this kind of analysis but information from experts from the Yugra State University in Khanty-Mansiyskiy was used as reference information in this work. Some of the indicated contaminated areas in the analogue reference maps did not fit spatially to the Quickbird data. The shape and the position of the proposed areas did not correlate with objects in the Quickbird data set. Automatic integration of the identified oil spilled areas (from the reference data) into the classification/validation system was therefore not possible. The identified areas were manually marked on the basis of the reference information using the most probable location in the Quickbird data.

The features that were used in this work are not directly based on the spectral properties of hydrocarbons. The classification uses indirect features that originate from the destructive pollution event (reduced reflectance in all visible bands, reduced vitality of vegetation with reduced NDVI values) or features that include class-related context information and class related dependencies. Recent oil contaminations with very low NDVI values on unvegetated soils were partly classified as water. The OSCaR project identified some methodological gaps that originate from this indirect mapping/image analysis approach. Results from this pilot project lead to an extended project proposal in 2006 that utilises hyperspectral

orbital data sets together with very high spatial resolution sub-meter data (the "OILSPILL" proposal). However - the performed analysis in OSCaR is one of the first approaches to set up an automatic oil spill contamination mapping system on the basis of very high spatial resolution optical remote sensing data that could lead to an operational environmental monitoring system for large areas. Improvements are possible with spatially more accurate ground reference information and information about contamination event dates. Also knowledge about regeneration stages of vegetation (and their related spectral properties) would help to differentiate pollution events.
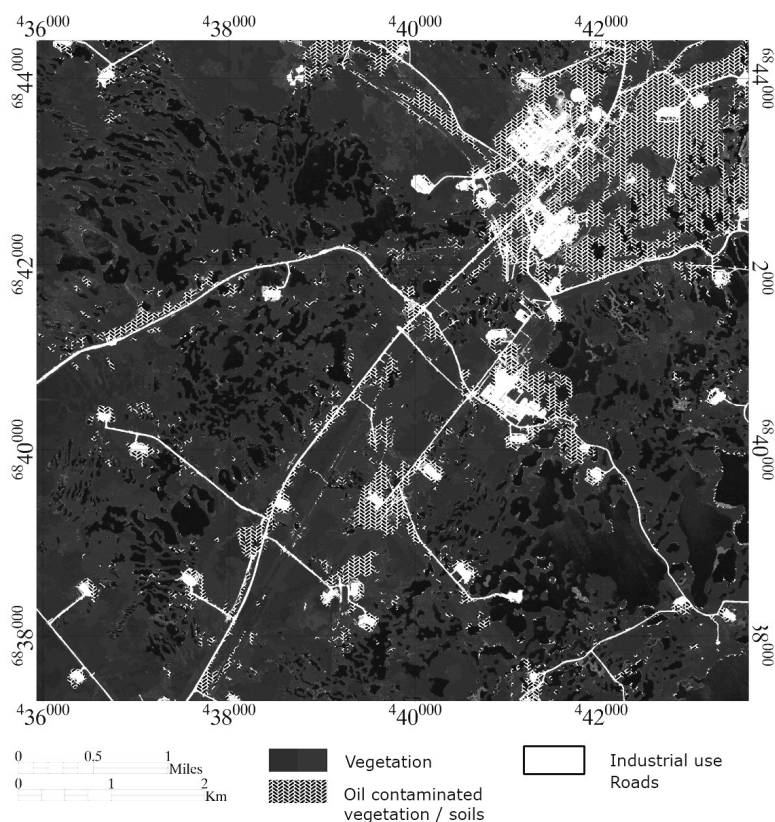
Fig. 8. Results of an object-based classification of oil contaminated objects using object shape, spectral information and object context information (class related features). Quickbird multispectral data with 2.5 m resolution. Water bodies were masked out in this analysis in an earlier step (Fig. 7)

6 Summary

This paper presents results of the BMBF (IB) co-financed OSCaR (Oil-Spill Contamination mapping in Russia) project for terrestrial oil spill classification in West Siberia with very high spatial resolution Quickbird Earth observation data and object oriented image processing methods. The developed class and process hierarchy (developed in Definiens Developer Software) for a test area in the Khanty Mansiysk district in West Siberia classified oil spills using spectral information, object shape information and class related features. The classification procedure used indirect features to differentiate contaminated soils and contaminated vegetation objects from uncontaminated objects. As reference information analogue ground truth data from Russian experts was used. Results show that class related information can be applied successfully to utilise the structural image object information of oil spilled land surface objects.

Final conclusions however also identified the need for mapping of specific spectral characteristics of hydrocarbon substances with very high spectral resolution data using calibrated and corrected hyperspectral information to complement the spatially very high resolution object oriented analysis with spectral signature and absorption feature analysis.

Acknowledgements

References

Argialas D, Derzekos, P (2003) Mapping urban green from IKONOS data by an object-oriented knowledge-base and fuzzy logic. In: Proc. SPIE Vol. 4886, p. 96-106, Remote Sensing for Environmental Monitoring, GIS Applications, and Geology II; Manfred Ehlers; Ed. 22-27 September, Aghia Pelagia, Crete

Baatz M, Schäpe A (1999) Object-oriented and multi-scale image analysis in semantic networks, in: proceedings of the 2nd International Symposium: Operationalization of Remote Sensing, 16-20 August, ITC, NL.

Benz UC, Hofmann P, Willhauck G, Langenfelder I, Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-

ready information, *ISPRS Journal of Photogrammetry and Remote Sensing*, 58 (2004), 239-258.

Brekke C, Solberg AHS (2005) Review: Oil spill detection by satellite remote sensing, *Remote Sensing of Environment*, No 95, 2005, pp. 1-13.

Chubey M, Franklin S, Wulder M (2006) Object-based Analysis of Ikonos-2 Imagery for Extraction of Forest Inventory Parameters. *PE & RS*, April 2006.

Damm A, Hostert P, Schiefer S (2005) Investigating Urban Railway Corridors with Geometric High Resolution Satellite Data, Urban Remote Sensing 2005, Berlin Adlershof.

Espedal HA, Wahl T (1999) Satellite SAR oil spill detection using wind history information, *Int. J. Remote Sensing*, 1999, Vol. 20, No. 1, pp. 49-65.

Fiscella B, Giancaspro A, Nirchio F, Pavese P, Trivero P (2000) Oil Spill Detection using marine SAR images. *Int. J. Remote Sensing*, 2000, Vol. 21, No. 18, pp. 3561-3566.

Flanders D, Hall-beyer M, Pereverzoff J (2003) Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. In: *Canadian Journal of Remote Sensing*, Vol. 29, No. 4, pp. 441–452, August 2003.

Folkman M, Pearlman J, Liao L, Jarecke P (2000) EO-1 Hyperion hyperspectral imager design, development, characterization, and calibration. SPIE, Vol. 4151, 2000.

Hese S, Schmullius C (2005) Forest Cover Change in Siberia - Results from the Siberia-II Project. International Conference on Remote Sensing of Environment, Conference Proceedings, St. Petersburg, Russia.

Hörig B, Kühn F, Oschütz F, Lehmann F (2001) HyMap hyperspectral remote sensing to detect hydrocarbons. *Int. J. Remote Sensing,* 2001, Vol. 22, No. 8, pp. 1413-1422.

IWACO Report (2001) West Siberia Oil Industry Environmental and Social Profile. Final Report, edited by M. Lodewijkx, V. Ingram, R. Willemse, June 2001.

Jones B (2001) A comparison of visual observations of surface oil with Synthetic Aperture Radar imagery of the Sea Empress oil spill. *Int. J. Remote Sensing*, 2001, Vol. 22, No. 9, pp. 1619-1638.

Leser C (2002) Operationelle Biotoptypenkartierung mit HRSC-Daten – Probleme und Lösungsansätze. In: Blaschke, T. (Hrsg.): GIS und Fernerkundung: Neue Sensoren – Innovative Methoden. Wichmann Verlag, Heidelberg: 88-97.

Lu J (2003) Marine oil spill detection, statistics and mapping with ERS SAR imagery in south-east Asia. *Int. J. Remote Sensing*, 2003, Vol. 24, No. 15, pp. 3013-3032.

Mitri GH, Gitas I (2002) The development of an object-oriented classification model for operational burned area mapping on the Mediterranean island of Thasos using LANDSAT TM images. in Viegas X. (ed.) Forest Fire Research & Wildland Fire Safety, 2002 Millpress, Rotterdam, ISBN 90-77017-72-0.

Pedersen JP, Seljev LG, Srom GD, Follum OA, Andersen JH, Wahl T, Skolev A (1995) Oil spill detection by use of ERS SAR data—from R&D towards pre-

operational early warning detection service. Proceedings of the 2nd ERS Applications Workshop, London, 6–8 December 1995, pp. 181–185.

Salem F, Kafatos M, El-Ghazawi T, Gomes R, Yang R (2005) Hyperspectral image assessment of oil-contaminated wetland. *Int. J. Remote Sensing*, Vol. 26, No.4, 20 February 2005, pp. 811-821.

Wismann V, Gade M, Alpers W, Hühnerfuss H (1998) Radar signatures of marine mineral oil spills measured by an airborne multi frequency radar. *Int. J. Remote Sensing*, 1998, Vol. 19, No. 18, pp. 3607-3623.

Chapter 4.2

An object-oriented image analysis approach for the identification of geologic lineaments in a sedimentary geotectonic environment

O. Mavrantza[1], D. Argialas[2]

[1] Department of Data Quality Control, KTIMATOLOGIO S.A., Athens, Greece, omavratz@ktimatologio.gr

[2] Laboratory of Remote Sensing, School of Rural and Surveying Engineering, National Technical University of Athens, Greece, argialas@central.ntua.gr

ABSTRACT: In this work, an object-oriented image analysis methodological framework employing edge extraction and image processing techniques was proposed and applied on a LANDSAT-ETM+ image of Alevrada, Greece to derive lineament classes. For the design of the knowledge base, the input data layers to the lineament identification system were (1) an edge map from the EDISON edge extraction algorithm performed on band 5 of the LANDSAT-ETM+ image, (2) the geologic layers derived from the geologic map at the pre-processing stage, (3) the initial ETM+ image of the study area and its derived thematic products using remote sensing methods (such as NDVI, PCA and ISODATA unsupervised classification) for the discrimination of the land cover classes. Segmentation was performed based on the multi-scale hierarchical segmentation algorithm in eCognition for the extraction of primitive objects of the input data. Finally, intrinsic spectral and geometric attributes, texture, spatial context and association were determined for the designed object classes / sub-classes on each segmentation level, and fuzzy membership functions and Nearest

Neighbor Classification were employed for the assignment of primitive objects into the desired thematic classes combining all participating levels of hierarchy. The output results of the system were classification maps at every hierarchy level as well as the final lineament map containing the geological lineaments of the study area (possible faults) and the non-tectonic lineaments.

1 Introduction

Geological lineament mapping is important in engineering problem solving, especially, in site selection for construction (dams, bridges, roads, facilities, etc.), seismic and landslide risk assessment, mineral exploration, hot spring detection, hydrogeological research, etc.

Photointerpretation of geologic lineaments (e.g. faults, fractures and joints) is very subjective and highly dependent on the photointerpreter's skills (Sabins 1997). It is also time-consuming and expensive, as it requires well-trained photointerpreters. Therefore, an effort towards automation of the lineament interpretation process is well justified.

Research efforts reported for the semi-automatic or automatic lineament detection and extraction can be categorized as following:

1. Semiautomatic and automatic lineament extraction, such as edge following, graph searching (Wang and Howarth 1990), novel edge tracing algorithms (Segment Tracing Algorithm (STA)) (Koike et al. 1995), etc.
2. Optimal edge detectors (e.g. the algorithms of Canny (1986), Rothwell et al. (1994), the EDISON algorithm (Meer and Georgescu 2001), etc.) were assessed for lineament extraction and provided quite promising results in terms of one-pixel thickness, efficient length and pixel connectivity (Argialas and Mavrantza 2004).
3. The design of a knowledge-based system that could take into account the measurable information (length, aspect) of geological lineaments (faults and drainage net segments) from a DEM (Morris 1991), the design of an expert system exploiting the descriptive geologic information relative to lineaments, without however interacting with an image (Masuda et al. 1991), as well as the embedding of geological knowledge into small autonomous programming routines (Rasco 1999).

1.1 Motivation and aim

For the discrimination of geologic (e.g., possible faults), topographic (ridges, drainage segments) and non-geologic lineaments (e.g., manmade lineaments and agricultural boundaries) and their automated identification, additional geological and geomorphological knowledge is required beyond the image information. In this context, semi- or even fully automated procedures of edge extraction aren't adequate for discriminating the type of the extracted edges and "isolating" in a way the linear features of "non-interest".

Therefore, an integrated knowledge-based methodological framework was designed for the identification and classification of geologic lineaments. It was based on the concept of low-level to high-level vision and processing, including three main processing stages. The designed framework included the following stages: **(1)** Remote sensing and edge extraction procedures, **(2)** Multi scale segmentation, and **(3)** Object-oriented knowledge base design and classification.

2 Methodology

2.1 Study area and data used – Image pre-processing

This work was conducted on a selected part (subset) of Alevrada District, in the extended geographic area of Amphilochia, Aitoloakarnania Prefecture, Central Greece. The region of Alevrada is characterized of sedimentary terrain with many faults with horizontal displacement, joints, as well as syncline and anticline folds. The study area belongs to the Gavrovo geotectonic zone and consists of Cretaceous and Upper Paleocenic – Eocenic successions of limestone, Upper Eocenic – Oligocenic flysch, and Quaternary alluvia (I.G.M.E. 1989). From the observation of the structural map and the corresponding satellite image it was derived that the majority of the faults in the study area of Alevrada lie in the boundaries of change of the lithological layers as it is presented in Figure 1(a) and the dominant structural directions are the NW-SE and the NE-SW.

The following initial and derived datasets were employed:

- The LANDSAT-ETM+ satellite image of the Alevrada District, (August, 6, 1999).
- A lineament map derived from the EDISON edge extraction algorithm, containing the lineaments to be classified.

- The structural and the lithological layers derived from the pre-processing phase (Mavrantza and Argialas 2002). These layers contained the information of the actual position of the dominant geologic lineaments of the scene (faults) and the lithological unit boundaries. Lineaments to be identified as faults appeared on the borders of the lithological units created by successive sedimentation.

At the pre-processing stage, the satellite image of the study area was geodetically transformed into the Transverse Mercator Projection and the Hellenic Geodetic Datum (HGRS87). The geological map of the study area of Alevrada with scale 1:50.000 was scanned. The structural (faults, fractures, joints and folds) and the lithological information layers were created by on-screen digitizing of the geological map.

Geometric registration of the LANDSAT-ETM+ satellite image with the digital geological map followed. This image was then radiometrically corrected by subtracting the path radiance of the visible bands.

From the interpretation of all ETM+ bands, it was observed that the ETM-4 and the ETM-5 bands provided the best visualization of the lineaments. In Figure 1(b), the digitized structural map features (faults and joints) as an overlay of a subset of a LANDSAT-ETM+ image with a size of 338x518 pixels are shown.

(a)   (b)
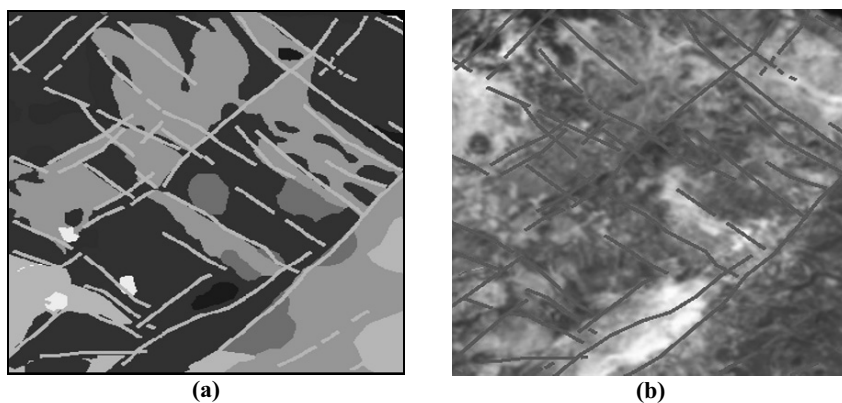
Fig. 1. (a) Grayscale representation of the structural and the lithological information layer digitized on the geologic map (Courtesy of I.G.M.E., 1989) (left), and **(b)** Digitized structural map features (faults and joints) as an overlay on the ETM-5 band – District of Alevrada – Pixel size 338x518 (right) (Mavrantza and Argialas 2002)

2.2 Image processing: Remote Sensing / Edge extraction methods and techniques

Digital Remote Sensing and edge extraction methods and techniques were applied in order to produce proper thematic and edge maps respectively, to be employed as input data to the knowledge representation model of the Lineament Identification System (LIS). Remote sensing methodology was applied for information extraction of the inherent land cover classes on the satellite image that surround the geologic lineaments, such as vegetated lithological classes (such as limestone, shale, etc.), as well as barren terrain.

The conceptual framework for the construction of the knowledge-based lineament identification system was initially based on the available satellite image and the geological map. The satellite image was processed in order to provide adequate spectral information related to vegetation, lithological sub-classes of successive sedimentation in areas containing lineaments. In turn, the geological map provided (a) the exact location of ground-verified faults and (b) the presentation of extended lithological units of diverse geological eras. The aforementioned information was represented in terms of rules, classes and attributes during the design of the object-oriented lineament identification system. The knowledge-based system of Alevrada only identified geologic faults and not topographic lineaments due to lack of 3D information (e.g., DEM).

The final purpose of the investigation of applying Remote Sensing and edge extraction methods was the identification of the following features:

- **Vegetation covering lithological sub-classes**: For the visual (qualitative) identification of the vegetation that covers the surficial lithological entities (and their distinction from barren soil), as well as the vegetation covering tectonic features, the color composites RGB-432 and RGB-543 were created from the satellite image (Sabins 1997). These composites were introduced into the knowledge-based system for constructing rules for extracting the classes of interest using Fuzzy Membership Functions (FMFs) related to spectral mean values of classes in the selected spectral bands. In addition, the application of the vegetation index NDVI was considered as a quantitative measure of vegetation content and vigor. The NDVI output was used as input data for the construction of rules using FMFs for the accentuation of land cover classes of different vegetation content and vigor, as well as for the discrimination among vegetation, bare sediments and very bright areas (roads). Vegetation is an indicator of certain lithological sub-classes, and is interrelated with water penetration and susceptibility to erosion. For example, sedimentary areas covered by sandstone and shale might be covered by forest species

and intense cultivation, but on the other hand, in humid areas, limestone (particularly, dolomitic limestone) is less penetrated by water and therefore, not covered with vegetation (Lillesand and Kiefer 2000).

- **Lithological sub-classes forming extended lithological units contained on the geological map**: The localization of the lithological sub-classes was performed by creating an RGB-132 color composite of the Principal Components that were derived from the PCA method (a grayscale representation of the RGB-PCA composite is illustrated in Figure 2). In this composite, spectral sub-classes were discriminated, which correspond to extended lithological units. ISODATA classification led to the derivation of the spectral sub-classes, which were recognized by photointerpretation of the PCA output and further introduced into the knowledge-based lineament identification system (a grayscale representation of the ISODATA classification output map is presented in Figure 3). In particular the application of PCA and ISODATA enabled the spectral discrimination of additional sub-classes of quaternary shale and cretaceous limestones indicated with thick arrows in Figures 2 and 3, The output maps of the latter two methods were introduced to the LIS for providing additional spectral information (in form of rules) for the discrimination of the existing land cover classes.
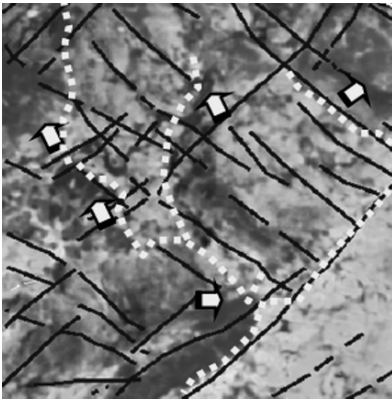
Fig. 2. Grayscale representation of the RGB-132 color composite of the PC1, PC2 and PC3 for the study area of Alevrada
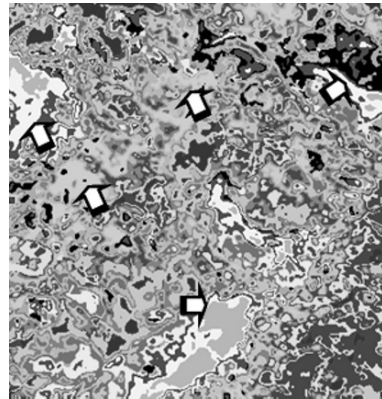
Fig. 3. Grayscale representation of the thematic map of the same area derived using the ISODATA unsupervised classification with 16 spectral classes. It should be noted that there is a correspondence in the identification of spectral classes with those of Fig. 2, presented with thick arrows

Edge extraction and evaluation was also performed. Various optimal edge detectors, such as the algorithms by Canny, Rothwell, the EDISON algorithm, etc. were investigated and applied on the satellite image, in order to derive the most suitable edge map. This edge map served as input to the designed LIS, as being the basis map containing edges to be classified as lineament types. At this stage, the edges cover all possible categories, including also non-significant lineaments (e.g. road segments, etc.).

The evaluation of the applied optimal edge detectors was conducted using qualitative (visual consistency) and quantitative criteria (use of the Abdou and Pratt (1979) and the Kitchen and Rosenfeld (1981) evaluation metrics) in combination.

At this stage, the EDISON edge map was selected in order to be introduced to the LIS because the application of the *EDISON* algorithm fulfilled the criteria of good edge extraction (good edge localization, edge connectivity and edge response) and sufficient interpretation ability of the semantic content (according to the qualitative (visual) and quantitative (measurement) criteria set).

2.3 Multi-scale hierarchical segmentation

The first stage of the object-oriented approach is image segmentation (Baatz and Schäpe 1999). Image segmentation is designed for the extraction of primitive objects to be classified during the object-oriented fuzzy classification process. The selection of the parameters was dependent from the type of the data input, so that the extracted segments could be minimum and meaningful. The color criterion had a greater weight than the shape criterion in the thematic maps, because the segmentation was based mostly on color. A scale parameter of 6 (in a tested range from 3 to 10) was adequate for extracting sufficient and meaningful segments in a "reasonable" processing time.

The conceptualization behind the design of the final structure of the participating hierarchical levels for the study area of Alevrada was based on data availability. The geologic map was decomposed into two different information layers, which in turn were used as two separate segmentation levels of tectonics (faults), and lithology (classes containing extended lithological units of different geological periods). The ETM+ image was used for the identification of the diverse inherent spectral classes of land cover (geological sub-classes and vegetation). The information of lithological units was derived from the lithology layer of the geological map. The information related to the lithological sub-classes was derived from the LANDSAT-ETM+ image, while the edge map containing edges to be

identified as geological lineaments was derived from the application of the EDISON algorithm. The **In_region lineaments** are lineaments that lie inside the lithological sub-classes and were detected by using the EDISON edge map and the LANDSAT-ETM+ image in the same segmentation layer. On the other hand, the **Border_lineaments** are those lineaments that lie on the borders of the lithological units. The border lineaments were detected by using the EDISON edge map and the lithology layer at the same segmentation level. The border lineaments in this study area were verified as faults. In Figure 4, the semantic network for the representation of the knowledge of the study area is presented.
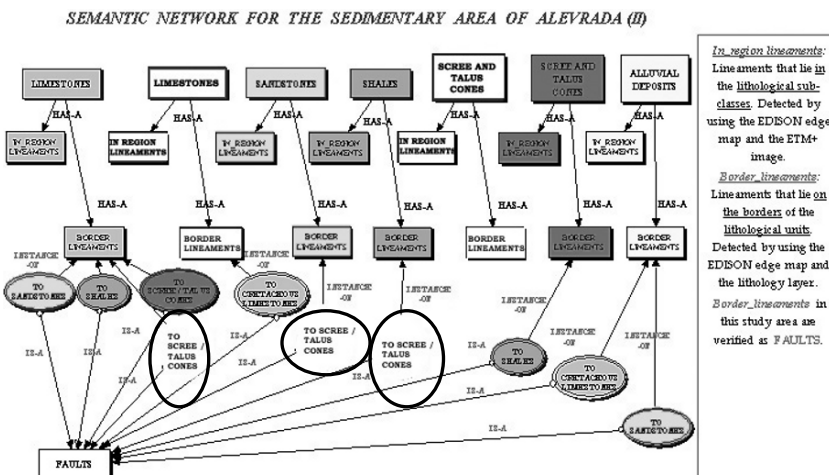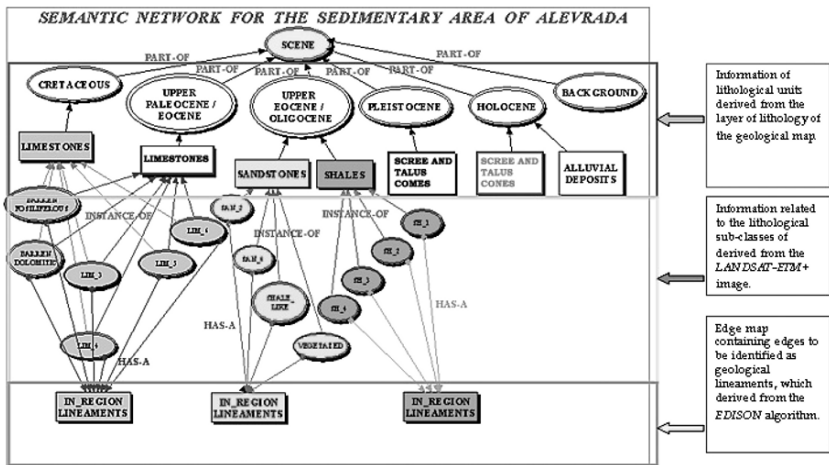
Fig. 4. The semantic network for the representation of knowledge for the study area of Alevrada

The semantic network is an alternative form of class representation, which presents in what manner the inherent lithological categories of different geological eras (e.g. limestones (Cretaceous or Upper Paleocene – Eocene) are connected with their sub-classes and how lineaments are spatially related to these classes, which under certain conditions indicate faulting.

Due to the diversity of the information format (raster – vector), the designed level organization for performing segmentation was the result of a trial-and-error process (horizontal and vertical parameter combinations).

2.4 Object-oriented knowledge base design and classification

After the stage of segmentation of digital geodata layers into object primitives, follows the stage of the creation of the object-oriented knowledge base and the classification of the segmented objects into meaningful semantic classes that represent the study area. During the design of the knowledge base the following points were taken into consideration:

- The determination of the appropriate object classes and sub-classes in every segmentation level.
- The determination of class attributes. These attributes were based on the spectral and geometric features of each class, and on the spatial relations and context of the information. The attributes to be inherited by the sub-classes followed the logic "from general-to-specific". In addition, these attributes were properly connected with the AND, OR, MEAN relations, according to the "weight" of the criterion that assigns each object to a specific class.
- The determination of the use of the Fuzzy Membership Functions (type of function and membership values), alone or in combination with the Nearest Neighbor Method (NNM).

The classification was conducted using six classification levels in the followed order:

1. **LEVEL 6**: At the hierarchy Level 6, based on the segmentation of the *EDISON* map, the objects were assigned into 2 classes, namely the class **ED-lineaments** and the class **ED-non-lineaments.** The distinction of those classes was based on the creation of the FMF of the spectral mean value of the participating edge map.
2. **LEVEL 2**: At the hierarchy Level 2, based on the segmentation of the Tectonic layer, the objects were assigned into 3 classes, namely the class **Structural Map_non-Faults,** the class **Structural Map_Faults**

and the class **Structural Map_Joints**. The distinction among those classes was based on the creation of the FMFs of the spectral mean values for R, G, B of the participating map and their logical AND linking.

3. **LEVEL 3:** At the hierarchy Level 3, based on the segmentation of the Lithology layer, the objects were assigned into 8 classes concerning the classes of the lithological units from Paleocene to Quaternary (which are interrelated with fault localization), namely the classes: Shales - Upper Eocene-Oligocene **(Flysch of Aitoloakarnania Syncline (L3),** Scree and Talus cones – **Pleistocene (L3),** Scree and Talus cones – **Holocene (L3),** Sandstones - Upper Eocene-Oligocene **(Flysch of Aitoloakarnania Syncline) (L3),** Limestones - Upper Paleocene-Oligocene **(Gavrovo Zone) (L3),** Limestones - Cretaceous **(Gavrovo Zone) (L3),** Background **(L3)** and Alluvial deposits – Holocene **(Quaternary) (L3)**. The distinction among those classes was based on the creation of the FMFs of the spectral mean values for R, G, B of the participating lithology map (grayscale representation of the coloured map) and their logical AND linking.

4. **LEVEL 4**: At the hierarchy Level 4, the classification into lithological sub-classes at the boundaries of the extended lithological units was performed. Lithological sub-classes were identified with the assistance of the ISODATA classification output representing the extended lithological formations of Level 3. In particular, Level 4 contained the following classes: (1) **Shale:** This class contains 4 sub-classes**: shale_1, shale_2, shale_3,** and **shale_4**. (2) **Scree and talus cones:** This class contains 3 sub-classes**: Scree and talus cones_1, Scree and talus cones_2** and **Scree and talus cones_3**. (3) **Sandstone:** This class contains 4 sub-classes: **Sandstone_shale-like, Sandstone_2, Vegetated Sandstone** and **Sandstone_4.** (4) **Limestones:** This class contains 6 sub-classes: **fossiliferous limestone (barren)**, **dolomitic limestone (barren)**, **limestone_3**, **limestone_4**, **limestone_5** and **limestone_6**. (5) **Background** and (6) **Alluvial deposits**. The distinction among those classes was based on the creation of the FMFs and the use of the NNM with spectral value criteria and relations of existence in different inheritance levels.

5. **LEVEL 5**: At the hierarchy Level 5, the interrelation among the edges of the edge map and their position (at the borders of lithological units or inside the lithological sub-classes) was examined. For the creation of Level 5, the image of Level 4 had to be classified into further nominal classes inside the lithological sub-class / unit and at the border of the lithological units (e.g. **in_f.sh_lineaments (L5)** and **border_of_f.sh_lineaments (L5)).** The distinction among those classes / sub-classes was based on the creation of the FMFs using spatial rela-

tions (neighborhood, distance) and existence relations to other (upper or lower) levels of hierarchy.

6. **LEVEL 1**: At the hierarchy Level 1 the final assignment of the edges on the edge map into **faults** (**faults** and **Structural_faults** (verified)) and **non-faults**, was performed, based on the information of the tectonic layer, which was used for verification due to lack of additional information (ground-truth data and DEM information). The distinction among those classes / sub-classes was based on the creation of the FMFs using spatial relations (neighborhood, distance) and existence relations to other (upper or lower) levels of hierarchy using AND / OR logical relations.

Concerning the determination of proper classification order of each segmentation level, it should be indicated that the classification order was not determined by the segmentation level order, but from the determination of the FMFs and the bi-directional (top-down and bottom-up) linking of attributes at different hierarchy levels in order to assign objects to each class.

In Table 1, object classes / sub-classes, their attributes, the FMFs as well as the value range are indicatively presented for Level 2.

Table 1: Object classes / sub-classes, their attributes, the FMFs and the value range for hierarchy Level 2

CLASS / SUB-CLASS	ATTRIBUTE	FMF	LEFT LIMIT	RIGHT LIMIT
Structural Map_Faults	Mean lineaments-to-big-plaisio-mapegsa.img (1)		-1	1
	Mean lineaments-to-big-plaisio-mapegsa.img (2)		10	129
	Mean lineaments-to-big-plaisio-mapegsa.img (3)		-1	1
Structural Map_Joints	Structural Map_Faults			
	Structural Map_non_Faults			
Structural Map_non_Faults	Mean lineaments-to-big-plaisio-mapegsa.img (1)		-1	1
	Mean lineaments-to-big-plaisio-		-1	1

mapegsa.img (2)			
Mean lineaments-to-big-plaisio-mapegsa.img (3)		**-1**	**1**

3 Results and Discussion

In this section, indicative derivatives of the knowledge-based system are presented, from all stages of processing (Figures 5-6 – grayscale representation).
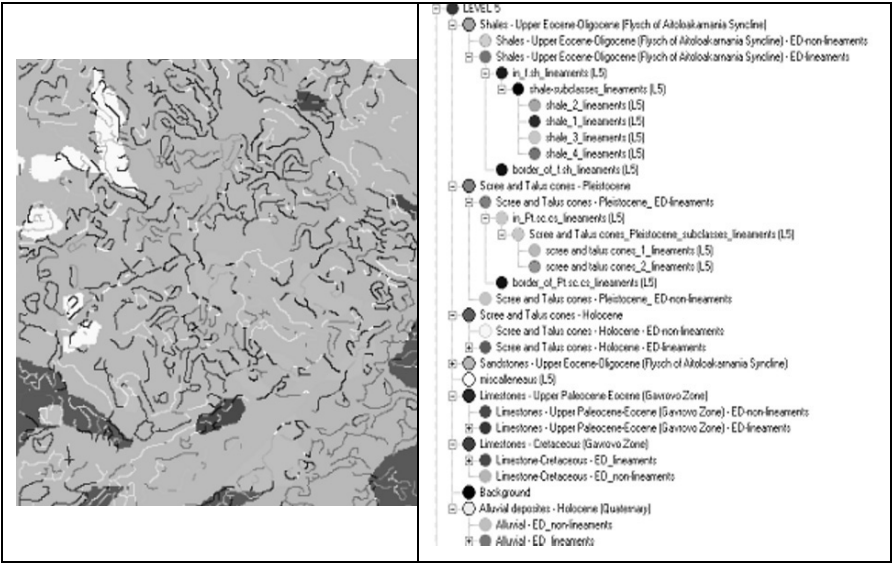
Fig. 5. Grayscale representation of the Level 5 classification output and the corresponding class hierarchy
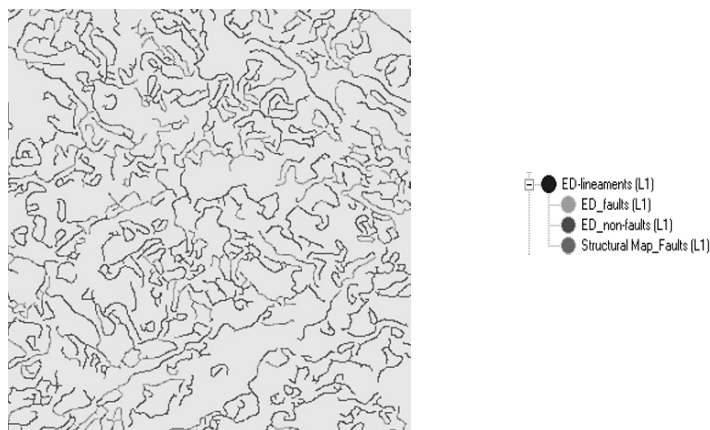
Fig. 6. Grayscale representation of the Level 1 classification output and the corresponding class hierarchy – Final lineament map, where lineaments correspond to faults

In Figure 7 the classification stability map is presented. It should be noted that dark gray edges on the stability map, which indicate points of **minimum stability** correspond **to common edges** on the tectonic map and the edge map, and lie on the major tectonic directions of the study area of Alevrada, as appeared in Figure 8.

Finally, for the study area of Alevrada, from the classification stability map it was inferred that there was a high coincidence of the results of the knowledge-based system and the tectonic map of the area. In all expected directions edge pixels have been correctly identified and the classification stability map served as an additional, qualitative indicator of adequacy of the classification.
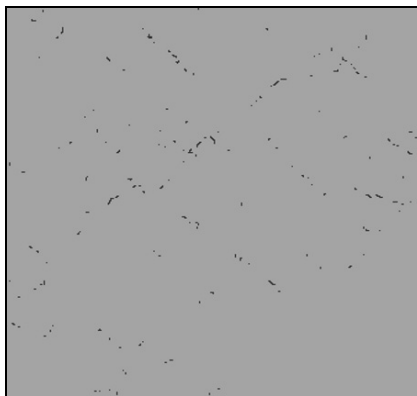
Fig. 7. Classification stability map for Level 1
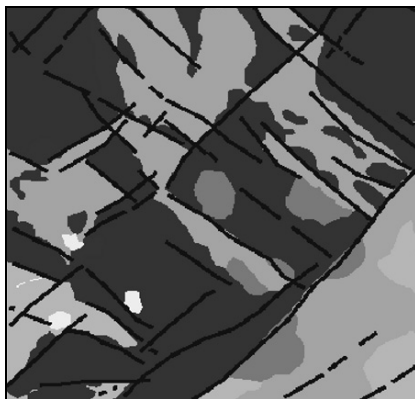
Fig. 8. Overlay of the tectonic layer (black lines) on the lithology layer

4 Conclusions

In the present work, a knowledge-based system was designed for the identification and classification of geologic lineaments. The output result of the system was a classified lineament map containing the geological lineaments of the study area (faults) and the lineaments, which were not identified as faults (non-interest lineaments).

The designed knowledge-based methodological framework, which combines multiple geodata types and techniques in the joint domain of Remote Sensing, Computer Vision and Knowledge-based systems enabled the derivation of final thematic maps of features of topographic / geological interest, by taking into consideration the limitations of the study area, the spectral behavior of the satellite geo-data, as well as the production accuracy of the input data, which were created with automated procedures.

Nevertheless, the designed knowledge base is strongly task and data-dependent and cannot be "generalized" to cover all geologic cases, but generally. However, the generalized methodological approach is transferable. The knowledge base is not transferable because it is totally data-dependent and also task-dependent, because every geotectonic regime is connected with different surficial expression of geologic features, while they are also differently originated, and therefore knowledge shall be "adapted" to the corresponding geological information.

References

Argialas DP, Mavrantza OD (2004) Comparison of edge detection and Hough transform techniques in extraction of geologic features. Proc. XXth ISPRS Congress of the International Society of Photogrammetry and Remote Sensing Turkey IAPRS-XXXV, pp 790-795.

Abdou IE, Pratt WK (1979) Quantitative Design and Evaluation of Enhancement / Thresholding Edge Detectors. Proc. IEEE 67, pp 753-763.

Baatz M, Schäpe A (1999) Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In: Angewandte Geographische Informationsverarbeitung, XI. Beiträge zum AGIT-Symposium, Salzburg, Karlsruhe, pp 12-23.

Canny JF (1986) A computational approach to edge detection. IEEE Transactions. on Pattern Analysis and Machine Intelligence, vol. 8, pp 679-714.

Institute of Geological and Mineral Exploration (I.G.M.E.) (1989) Geologic Map Sheet "Alevrada", Scale 1: 50.000, Athens, Greece.

Kitchen L, Rosenfeld A (1981) Edge Evaluation using Local Edge Coherence. IEEE Transactions on Systems, Man and Cybernetics, vol. 11, pp 597-605.

Koike K, Nagano S, Ohmi M (1995) Lineament analysis of satellite images using a Segment Tracing Algorithm (STA). Computers & Geosciences, vol. 21, pp 1091-1104.

Lillesand TM, Kiefer RW (2000) Remote Sensing and Image Interpretation, Fourth Edition, John Wiley & Sons, New York.

Masuda S, Tokuo T, Ichinose T, Otani K, Uchi T (1991) Expert System for Lineament Extraction from Optical Sensor Data. Geoinformatics, vol. 2, Japanese Society of Geoinformatics, pp 195-200.

Mavrantza OD, Argialas DP (2002) Implementation and evaluation of spatial filtering and edge detection techniques for lineament mapping - Case study: Alevrada, Central Greece. Proc. SPIE International Conference on Remote Sensing: Remote Sensing for Environmental Monitoring, GIS Applications, and Geology II, M. Ehlers (editor), SPIE Press, Bellingham, WA, SPIE-4886, pp 417-428.

Meer P, Georgescu B (2001) Edge detection with embedded confidence. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-23, pp 1351-1365.

Morris K (1991) Using knowledge-base rules to map the three-dimensional nature of geological features. Photogrammetric Engineering and Remote Sensing, vol. 57, pp 1209-1216.

Rasco HP (1999) Multiple Data Set Integration And GIS Techniques Used To Investigate Linear Structure Controls In Southern Powder River Basin, Wyoming. M.Sc. Thesis, Department of Geology and Geography, West Virginia University, West Virginia, USA, pp 1-95.

Rothwell Ch, Mundy J, Hoffman B, Nguyen V (1994) Driving Vision by Topology. TR-2444, INRIA, pp 1-29.

Sabins FF (1997) Remote Sensing: Principles and Interpretation. W. H. Freeman and Company, New York.

Wang J, Howarth PJ (1990) Use of the Hough transform in automated lineament detection. IEEE Transactions on Geoscience and Remote Sensing, vol. 28, pp 561-566.

Classification of linear environmental impacts and habitat fragmentation by object-oriented analysis of aerial photographs in Corrubedo National Park (NW Iberian Peninsula)

R.A.D. Varela,[1] P.R. Rego,[1] M.S.C. Iglesias[2]

[1]Botanic & Biogeography Laboratory. Department of Botany. I.B.A.D.E.R. University of Santiago de Compostela. Escola Politécnica Superior, Campus Universitario s/n, E27002 Lugo, Spain. E-mail: mon@lugo.usc.es, dibader@lugo.usc.es

[2]Department of Agroforestry Engineering, University of Santiago de Compostela. Escola Politécnica Superior, Campus Universitario s/n, 27002 Lugo, Spain E mail: scalvo@lugo.usc.es

KEYWORDS: Vegetation monitoring, orthoimage classification, fuzzy logic, object oriented analysis, context, linear features

ABSTRACT: The increase of tourism in coastal protected areas is a potential driver for ecosystem fragmentation, as it frequently involves damages to their biodiversity values. One of the key issues in the management of protected areas is their access regulation in order to avoid negative effects of trampling on natural and semi-natural habitats. Protection measures need to be carefully designed to achieve an effective protection of biodiversity while minimizing as possible the constraints to tourism development. The accurate design and monitoring of performance of such measures needs effective methods to evaluate fragmentation.

In this work, we aimed at the development of a consistent method for the automatic recognition of linear environmental impacts (paths and tracks) on natural and semi-natural habitats using colour aerial photo-

graphs. The method is based on a multi-scale segmentation of the images. Brightness, shape and connectivity criteria were implemented at several scales of analysis by means of a fuzzy knowledge base, allowing the recognition of linear elements corresponding to networks of trails. We used as study case a protected area in the NW coast of the Iberian Peninsula with a complex vegetation pattern. The method allowed the general discrimination of linear artificial features, causing environmental impacts and human-driven habitat fragmentation, against other linear elements with similar brightness and shape, but different thematic meaning regarding their conservation implications.

1 Introduction

Environmental planning in coastal protected areas needs taking into account the pressure on biodiversity due to the increase of leisure and tourism activities in these locations. In fact, tourism pressure may cause important damages to biodiversity values in ecosystems with high vulnerability and low resilience. For instance, the effects of trampling on habitats like coastal sand dunes may remove and destabilize the vegetation cover and cause its fragmentation, particularly when creating linear features like footpaths.

Although several definitions may be found in the literature, habitat fragmentation is often regarded as a process in which "a large expanse of habitat is transformed into a number of smallest patches of smaller total area, isolated from each other by a matrix of habitats unlike the original" (Wilcove et al. 1986). Linear infrastructures are acknowledged to be one of the main causes of fragmentation of habitats and involving several kinds of environmental impacts (Geneletti 2004). The reduction of overall amount of habitat and mean patch size, the increase of edges, the decrease of core area and the isolation of habitat patches are some of the principal fragmentation effects on landscape structure. Due to their particular ecological requirements, not all the species and ecosystems show the same response to such effects, but it is possible to state an overall negative effect of fragmentation on biodiversity (Fahrig 2003).

Tourism pressure and particularly non-regulated circulation of people and vehicles, is one of the main reasons of sand dune ecosystem fragmentation in coastal areas. The management of such areas requires establishing a control of the public access to avoid irreversible damages to ecosystems derived from the unplanned rising and enlargement of footpaths and trails. These measures have to be carefully designed in order to constraint as few

as possible the tourism development and to contribute efficiently to the protection of biodiversity.

Assuming the existence of a relationship between the landscape pattern and processes taking place on it, an accurate characterization of the landscape structure contributes significantly to the understanding of ecological processes and eventually in the identification of threats on the long term maintenance of biodiversity (Turner 1989). Indeed it is necessary to have accurate and updated spatial data about the composition and configuration of the landscape, in order to evaluate its state and response to perturbations and for the eventual design and monitoring of protection measures. Besides, the retrieving of information about past condition of the habitats is also necessary for a better understanding of their present state and trends.

In the current work, we aimed at the development of a consistent method for the automatic recognition and mapping of linear environmental impacts, namely trails and footpaths, on natural and semi-natural habitats in a protected area of the NW of Spain. Due to their informal and in some cases even temporal character depending on the tourism affluence, they are not included in the official topographic cartography and therefore they have to be retrieved from other data sources.

Remote sensing data and analysis techniques show advantages over traditional mapping techniques for cost effective generation of environmental data, such as the exhaustive and systematic covering of the territory, its periodical data acquisition or the possibility of automation of analyses. Nevertheless, the availability of remote sensed data sources with enough resolution (temporal, spectral and spatial) may be quite limited to identify certain environmental features, particularly when historical data records are demanded. In this regard, collections of aerial photographs may be a valuable source of information, despite of their frequent shortcomings in quality, homogeneity and spectral information. Thus, the basic dataset for the study was a mosaic of ortho-rectified aerial photographs with a spatial resolution that allows the identification of narrow linear features. In this particular case, the target is the classification of roads, trails as well as footpaths resulting from tourist impact in a complex coastal sand dune system. Since it is foreseen the analysis of historical aerial photographs in further work, one of the premises of the research is to develop a methodology valid for both colour and black and white images. In view of the limited availability of spectral information in the images, the method was designed to account as less as possible on spectral information and taking profit of other information such as morphology and spatial context of image objects.

2 Material and methods

2.1 Study area

The study area is located inside the Natural Park of "Complexo dunar de Corrubedo e lagoas de Carrexal e Vixán" located in the NW coast of the Iberian Peninsula (Figure 1). The Park comprises a complex mosaic of environments hosting a great variety of fauna and flora (García-Bobadilla Prósper et al. 2003). Its environmental importance has been recognised with its declaration as Natural Park, International Importance Wetland (Ramsar Wetland), Special Area of Conservation (SAC) of the Natura 2000 network in the Atlantic region as well as Special Protection Area (SPA) for Birds.

For the purpose of this study a rectangle of 65 ha within the limits of the Natural Park was used as test area. Its size results of a compromise between the computational costs of high resolution images and the need of exemplifying the problem of fragmentation in coastal sand dune ecosystems. The study area comprises a complex pattern of ecosystems near the shore, including wetlands, sand beaches, white dunes (i.e. no stable with bare and loose sand) and grey dunes (more stable, with higher amount of organic matter). Other types of vegetation like agricultural fields, grasslands, meadows and hedgerows are mainly distributed further in the inland. Rocky slopes and different kind of scrublands are distributed throughout the area. Finally, the study area has an important presence of human-made features such as infrastructures (mainly narrow roads, tracks and footpaths) as well as some buildings.
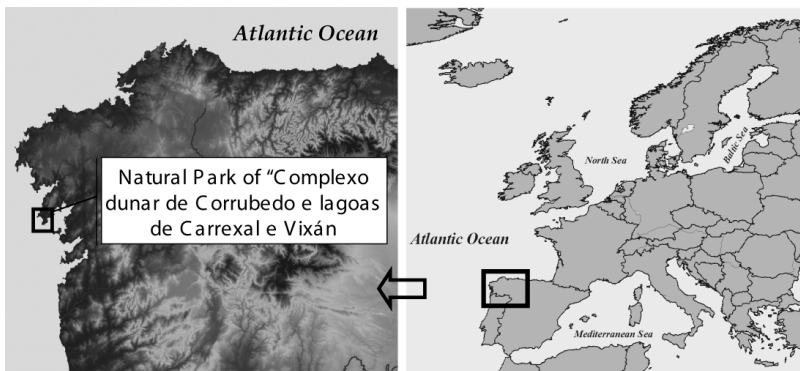
Fig. 1. Location of the study area in the European context

2.2 Datasets

We used a subset of an 8-bit orthorectified real colour aerial photograph with a spatial resolution 0.20 m, meaning an image size of 3695 by 3536 pixels. The original photographs were recorded in 2002-2003 with an original scale of 1:18000 (FOGGA 2003). We took information about the current distribution of habitats and path network by means of fieldwork and visual aerial photograph interpretation. Existing habitat maps with a scale of 1:5000 (Consellería de Medio Ambiente 2005) were also used as ancillary data in the classification process.

2.3 Classification methods

For the design of the method, we worked from the basic premise that the target elements (i.e. the network of paths and roads) accomplish the following conditions:

– To show a linear morphology
– To have high brightness values (i.e. to have low vegetation coverage)
– To take part in a connected network

These basic rules usually perform well when typical anthropogenic features like roads or paths are the targets, as they are built according to certain standards and therefore are relatively straightforward to model (Lang and Blaschke 2003). However, in the current work, this basic statement is likely to present exceptions as some parts of the path network could be partially colonized by vegetation (i.e. might show relatively low brightness values). Besides, some natural and seminatural habitats might show morphology and brightness values similar to paths and roads, while in other cases some sections of the paths might not be connected with the general infrastructures network (particularly in the case of temporal unplanned footpaths). Therefore the targets of the classification may show an important variation in their main characteristics as often being the result of unplanned human activities and consequently, the difficulty of their modelling increases. In order to take into account such variability, we designed a flexible system based on a fuzzy logic approach for the recognition of artificial linear features. Classificatory analyses were done following an object-oriented approach, since it shows advantages comparing to pixel-oriented approach when features such as shape, texture and topology of objects are to be used (Blaschke et al. 2001; Giakoumakis et al. 2002). Analyses were based on a multi-scale segmentation and subsequently a rule-based classification of the resulting image objects by means of a fuzzy

knowledge base. The multi-scale approach allowed the recognition of complex features resulting from the integration of basic elements (i.e. objects) at different levels of analysis.

We used the software eCognition ProfessionalTM V4.0 both for the segmentation of the image and for the implementation of the knowledge base. In the segmentation process, the image is divided or segmented in a number of spatial clusters likely to represent objects belonging to the same class. The method is based on a bottom-up region growing technique that starts from the objects with the lowest spatial entity in the image (i.e. individual pixels) and iteratively builds spatial clusters meeting a criterion of minimum heterogeneity according to certain parameters of spatial scale and shape (see Baatz et al. 2004 for further details). We implemented the classification knowledge-base by means of decision rules based on spectral, morphological and topological information of the image objects combined via fuzzy logic operators at different levels of analysis. Instead of individual information for the RGB channels of the colour photograph, we used the overall brightness of the image. This will ease the replication of the method for the identification of linear features in other datasets, such as historical black and white aerial photographs. We also integrated information of shape (linear vs. non-linear morphology of objects) and context by mean of topological information (connectivity between certain classes).

Parameters and function profiles for the fuzzy decision rules were based on trial-error procedures using as reference well known elements belonging to the target classes. For almost all rules, we avoided sharp boundaries for the membership functions and adopted sigmoid shaped functions so as to enhance the advantages of fuzzy operators as allowing certain flexibility in the fuzzy decision space (Baatz et al., 2004; Bock et al., 2005).

A basic level (Level 2) of image objects was derived using segmentation parameters that allowed the discrimination of elements likely to belong to the trail network (c.f. table 1). Several trials followed by a visual comparison of results against known target features on the aerial photograph were done until satisfactory image segmentation was achieved. The scale parameter was adjusted for allowing the discrimination of path stretches against other land cover types (even if it meant a certain degree of over-segmentation). As the classification was not based exclusively on spectral features, the shape parameter was set to a relatively high value. Compactness and smoothness parameters were equally weighted.

Table 1. Parameters for the basic level of the segmentation (Level 2)

Scale	Shape	Compactness
80	0.4	0.5

A lower hierarchic level (Level 1) was derived for the purpose of line analysis. Each Level 2 object was subdivided in sub-objects that maximize the border length to the outer environment which were used for the generation of shape information about the linear character of their respective super-object (see Baatz et al. 2004 for further details).

Level 2 objects were classified in several categories according its brightness and shape. Overall brightness mean and object asymmetry at Level 2 and width and width/length ratio at Level 1 (sub-object) were the criteria used in the classification (for a thorough description of these features see Baatz et al. 2004).

At this stage, linear elements with a low vegetation coverage and therefore likely to form part of trails were classified. Nevertheless, linear elements corresponding with natural and seminatural vegetation with inherent low vegetation coverage (e.g. long and narrow rocky slopes) were also assigned to the former class, even when they did not take part that in the trail network. On the other hand, vegetated and linear elements that form part of the trail network (e.g. temporal footpaths partially colonised by vegetation) along with other elements corresponding with hedgerows or narrow strips of natural or seminatural vegetation (conceptually very different to the trail network) were assigned to the same class. Therefore, additional connectivity criteria were introduced in the classification to allow the semantic discrimination between objects belonging to the trail network against other elements spectrally and morphologically similar but not trails, and vice versa. Criteria had to be implemented for all the classes likely to take part of the trail network evaluating the existence of other neighbouring elements likely to form part of the same network. To achieve this, rules of connectivity were designed at two hierarchical levels:

Level 1 objects likely to be trails were labelled in two classes according their relation with Level 2 elements for its further integration in the decision rules:

– "Possible trail" (Level 1), when the objects were contained in any Level 2 super-object belonging to any of the classes likely to be trails
– "Near possible trail" (Level 1), when the objects fell within close distance of an element of the possible trail at Level 1

Level 2 objects covered by vegetation, with linear shape and contacting linear non-vegetated elements were classified as "contacting non-vegetated and linear vegetation objects" (see figure 2).

In the case of absence of this contact, they were classified both as "neighbouring possible trail" or "non-neighbouring possible trail" according to the distance to elements presumably belonging to the trail network. The first class corresponds to linear objects with some vegetation coverage

and close to linear unvegetated elements very likely to be trails. The second class corresponds to elements at a certain distance to linear unvegetated elements or to the previously defined class "contacting non-vegetated and linear vegetation objects class". For the definition of this class, the Level 1 categories were used as intermediate step (see figure 2). At this point, a circular reference arose, since the classification of one element in one of the classes likely to be part of a trail depends on the classification of the neighboring and sub-objects as belonging to one of these classes. Accordingly, the classification of the neighboring objects in one of these classes depends on the final assignation of the targeted segment. For solving this, we applied the option of class-related iterative classification of eCognition software.

Next step in the analysis consisted in the generation of a higher hierarchical level (Level 3) based on the Level 2 classification. The main objective of this level was to identify concatenations of Level 2 objects belonging to classes semantically affine and contiguous in the space (see figure 3).
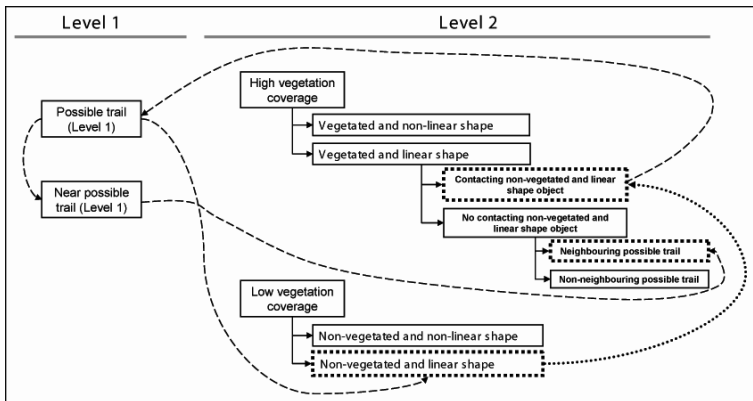
Fig. 2. Hierarchical classification scheme at Level 1 and 2. Classes likely to form part of the trail network are marked with dot rectangles. Curved dot line indicates relations used for the definition of connectivity criteria
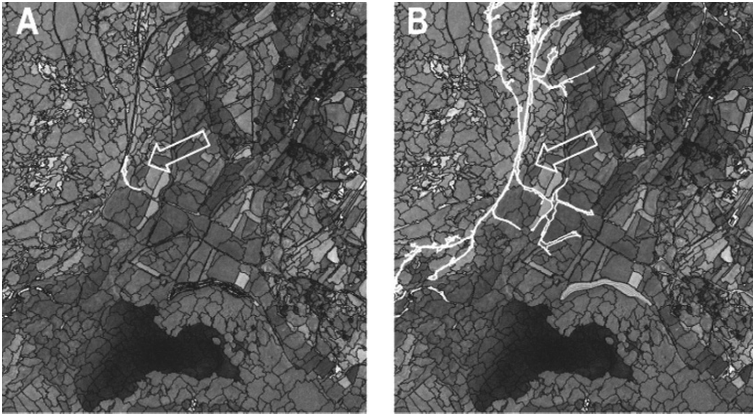
Fig. 3. Level 3 classification based segmentation. In A, an example of level 2 object is underlined in white. In B is shown this object assembled with other contacting objects likely to belong to a connected trail network

Thus, a spatial assemblage of the objects considered conceptually more likely to be part of the trail networks was done. Besides, a new abstract group ("not possible trail") put together the classes more unlikely to be part of the trail networks at Level 3. This abstract group was created in order to facilitate the joint integration of the classes conceptually different to trails in the decision rules.

A new set of decision rules were developed to recognize topological relationships between the elements assigned to classes with different probability of belonging to the trail networks and eventually, to discriminate which of the complex Level 3 objects were real trails (see figure 4). We used as connectivity and length of the Level 3 objects as primary decision criteria. Long linear segments with a presumed connection with the trail network outside the image limits or with a certain degree of vicinity with possible paths were classified as trails. Conversely clearly isolated and short composed segments were classified as "non vegetated linear".
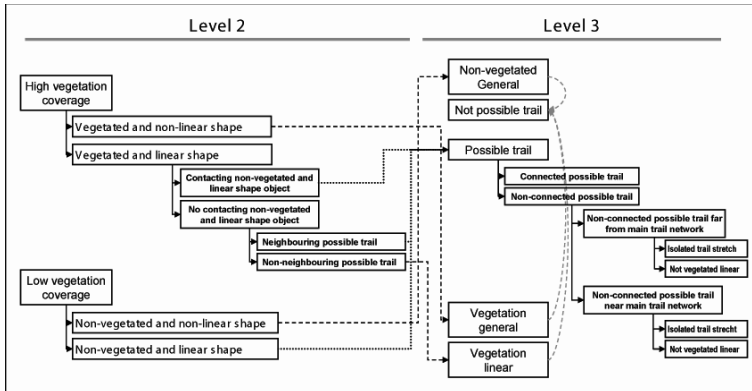
Fig. 4. Hierarchical classification scheme at Level 2 and Level 3. Dash square lines represent the direct correspondence between Level 2 and Level 3 classes. Dot square lines indicate the spatial assemblage at Level 3 of classes more likely to be trails at Level 2. Dash curved arrows indicate the assemblage at Level 3 of the classes more unlikely to be trails.

Final Level 3 classification was re-labeled to a definitive output classification scheme of three classes. All elements corresponding with vegetation covered ground, were assigned to the output class "vegetation". Elements corresponding with bare ground and not belonging to any kind of classes of the trail network were assigned to the output class "no vegetation", Finally all the classes semantically belonging to the trail network were labeled as "linear impacts".

2.4 Accuracy assessment

We tested the classification accuracy by means of a contrast against independent reference data obtained by the visual analysis of aerial photograph and ancillary data. In order to ensure the statistical coherence of the validation, verification points were extracted from the image by means of a simple random sampling (Congalton, 1991). Sample size was determined according the expression for multinomial test and variables (Tortora, 1978) for 5 variables (number of classes), probability of Type 1 error of 0.05, 95 % of precision and for the worst case scenario of class frequency distribution. A final sample size of 665 verification points was computed, but since some of the classes showed low frequencies, a final sample size of 1000 points was used so as to avoid sub-sampling these classes.

Confusion matrix and descriptive accuracy indices were computed from verification points (Congalton, 1991). Omission and commission errors

and conditional kappa were computed at class level. Overall accuracy and Kappa index as well as Kappa confidence intervals and Z values were computed at overall classification level.

3 Results and discussion

Level 2 classification allowed the discrimination of elements according their degree of vegetation coverage, quantified by overall brightness for the three channels of the RGB photograph, At this stage of the classification, the knowledge base allowed the discrimination of terrestrial and aquatic vegetation (low brightness values) from bare ground (high brightness values). From the semantic point of view, bare ground can corresponded both to trails or to other elements with low vegetation coverage, as some natural and semi-natural habitats, ploughed fields or houses. Therefore it was necessary to design rules to discriminate between them. First step for this discrimination consisted in splitting between linear elements, likely to form part of the trail network, and non-linear elements. To achieve this, we used lineal shape as classification criteria, quantified by the width and ratio length/width.

Objects classified as "non-vegetated and linear shape" correspond in most of the cases with stretches of trails. Nevertheless other vegetation types were recorded in this class, as occurred with some objects corresponding to strips of pioneer vegetation on rocky slopes or bare sand. Therefore, this class could not be directly matched with a thematic category corresponding only with trails, and some refinement of the classification was needed.

On the other hand, in some cases trails can be partially colonized by vegetation, showing intermediate to low values of brightness but having certain effects in terms of habitat fragmentation. It was therefore necessary to identify such elements and label them as part of the trail network. For this purpose, vegetation elements were spited up in two sub-groups, discriminating between linear and non-linear vegetation patches. The class of vegetation linear elements was again splitted in sub-classes according their degree of connectivity with other linear elements likely to take part of the trail network. This allowed the distinction of elements with high probability of being stretches of the trail network partially colonized by vegetation against other linear and vegetated elements like hedgerows or gallery forests semantically different from the class of linear impacts.

Level 3 classification was used mainly for refinement of the classification of linear elements. Here, decision rules aimed at identifying concate-

nations of stretches forming a theoretical topologically coherent network of trails.

Once re-labeled the Level 3 classification, the final output followed a legend of three land cover classes, discriminating between the network of trails (identified with linear impacts) vegetation coverage and non-vegetation coverage (fig 5). Most frequent class was "vegetation", covering almost 89 % of the area, corresponding to different types of more or less dense vegetated natural and seminatural habitats, artificial grasslands and crops (cf. table 2). The class "no vegetation" covers less than 7 % of the area, comprising natural and seminatural habitats with an inherent low coverage of vegetation (rocky slopes, boulders or sand dunes), along with anthropocentric elements such as arable land or recently mowed meadows. Finally, less than 5 % of the area was classified as "linear impacts", corresponding to roads, trails or footpaths.

Table 2. Class area distribution in ha. and percentages

Class	Area [ha]	Area [%]
Vegetation	58.2	88.6
No vegetation	4.5	6.8
Linear impacts	3.0	4.6
Total	65.7	100.0
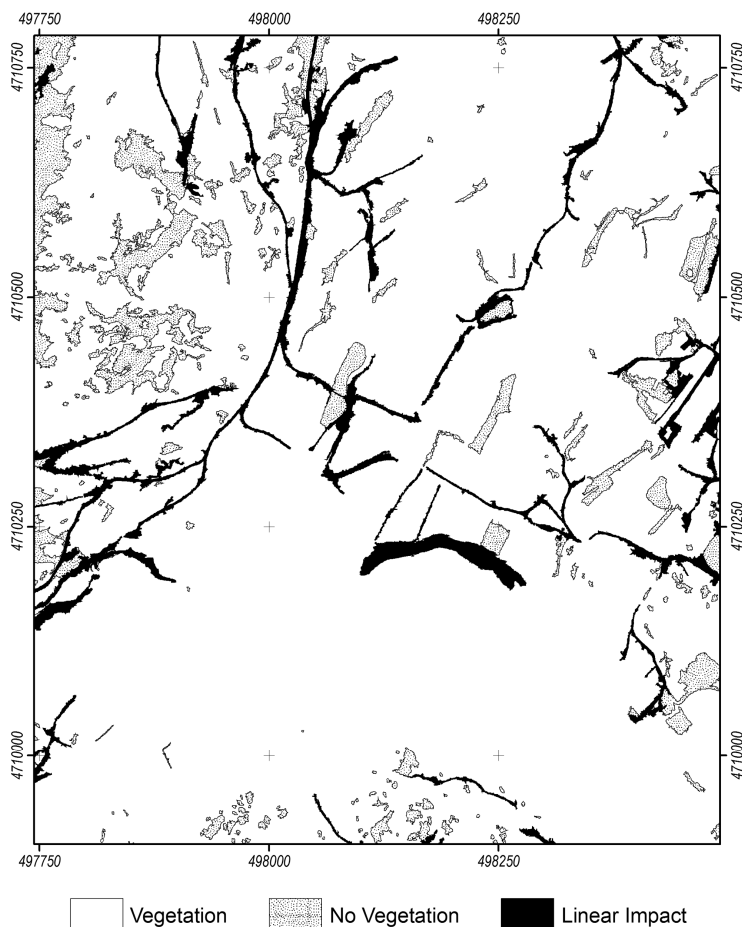
Fig. 5. Final classification. Grid coordinates in meters, UTM projection Zone 29 T North, Datum European 1950

Classification achieved an overall accuracy of 95 % (cf. table 3). Nevertheless, the clear dominance of one of the classes in the map, may have caused an overestimation of the accuracy by this index (Fung and LeDrew, 1988). Estimation of Kappa index is regarded as a more reliable indication of the overall accuracy, achieving a value of 0.80, also indicating an almost perfect agreement between the classification and reference values according the scale of Landis and Koch (1977) or very good agreement according Monserud and Leemans (1992). High value of the Kappa Z statistic also confirms clearly the statistical difference with a random classification.

Table 3. Overall accuracy indices. Confidence intervals for kappa value were computed for 95 % of confidence level.

Overall accuracy [%]	Kappa	Z	Kappa confidence interval	
			Upper	Lower
96.10	0.80	26.75	0.74	0.86

Per class accuracy indices also showed high values (cf. table 4) with Kappa values indicating very good or substantial agreement with reference data (Landis and Koch, 1977; Monserud and Leemans, 1992). For the class linear impacts, producers accuracy value was lower than users, pointing out a potential too narrow definition of the class. Nevertheless for the potential uses of the method in protected areas management and planning, certain sub-estimation of the linear impacts may have less social effect than an over-estimation, for not inducing unnecessary severe regulations. On the other hand, some tuning of the classification parameters may easily be implemented in the knowledge base in order to minimize the omission errors and ensure a more strict protection of the space.

Table 4. Per-class accuracy. Contingency matrix with reference data and per-class accuracy indices. Reference data in columns, classification data in rows.

Class	1	2	3	Total	Producers Accuracy [%]	Users Accuracy [%]	Kappa
1: Vegetation	881	6	10	897	98,66	98,22	0,83
2: No vegetation	7	51	5	63	80,95	80,95	0,80
3: Linear impacts	5	6	29	40	65,91	72,50	0,71
Total	893	63	44	1000			

An assessment of the spatial distribution of errors allowed the identification of some potential limitations for generalization of the knowledge-base classification to other scenarios. As an example, commission errors revealed how a plot of land of recently mowed reed, with a spatial sequence alternating linear strips of mowed reed and remains of vegetation, was recognized as part of the trail network. Although the objects of this plot fulfilled the requirements for belonging to the class of linear impacts, they conceptually corresponded to a single element of partially unvegetated land. Most of the remaining errors occurred on transition areas in the margins of the trail network or in complex patterns of trails and linear shaped plots of vegetation. A manual refining of the classification in such areas or improvements in the method by including ancillary information are potential solutions for these misclassifications.

4 Conclusions

The integration in a knowledge base of geometric and contextual data, in addition to spectral information, constituted an optimal strategy for the identification of linear infrastructures, even when complex arranged landscapes are studied (for other examples of successful applications see Nobrega et al. contribution in the book).

We developed a reliable method for the discrimination of artificial features, corresponding with environmental impacts and human-driven habitat fragmentation, against other linear elements with similar brightness and shape, but with different thematic signification regarding their biodiversity conservation implications. These anthropogenic features were classified despite of their complexity as they showed an important variation in width, shape of the borders and connectivity as a result of an unplanned action.

Even when the method achieved high accuracy values, some isolated errors were identified, resulting from the occurrence of very complex patterns of interspersed natural and anthropogenic features with linear shape occurred. In other cases, the intrinsic characteristics of particular features, as happened with strips of mowed reed alternating with strips with some remains of vegetation, led to confusion between non-linear elements and complex linear elements.

In spite of the good results for the test area, new experiments are foreseen, in order to test potential shortcomings in the generalization of the method to other areas.

5 References

Baatz M, Benz U, Dehghani S, Heynen M, Höltje A, Hofmann P, Lingenfelder I, Mimler M, Sohlbach M, Weber M, Willhauck G (2004) eCognition Professional User Manual 4. Definiens Imaging, München

Blaschke T, Conradi M, Lang S (2001) Multi-scale image analysis for ecological monitoring of heterogeneous, small structured landscapes. Proceedings of SPIE, Toulouse, pp 35-44

Bock M, Xofis P, Mitchley J, Rossner G, Wissen M (2005) Object-oriented methods for habitat mapping at multiple scales - Case studies from Northern Germany and Wye Downs, UK. Journal for Nature Conservation 13(2-3): 75-89

Congalton RG (1991) A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37(1): 35-46

Consellería de Medio Ambiente (2005) Plan de conservación de las Zonas de Especial Protección de los Valores Naturales de Galicia. IBADER. Universidade de Santiago de Compostela. Dirección Xeral de Conservación da Natureza. Xunta de Galicia, Lugo (8 Tomos)

Fahrig L (2003) Effects of habitat fragmentation on biodiversity. Annual Review of Ecology, Evolution, and Systematics 34: 487-515

FOGGA (2003) SIXPAC. Xunta de Galicia. Consellería de Medio Rural. Available in internet, URL: http://www.xunta.es/visor/. Data last access: 22/06/2006

Fung T, LeDrew E (1988) Application of Principal Component Analysis to Change Detection. Photogrammetric Engineering & Remote Sensing 53(12): 1649-1658

García-Bobadilla Prósper F, Lago García JM, Juliani Aguado C, Callejo Rey A, Ramil Rego P, Izco Sevillano J (2003). Humedales de Galicia. DXCN. Consellería de Medio Ambiente. Xunta de Galicia, Santiago de Compostela

Geneletti D (2004) Using spatial indicators and value functions to assess ecosystem fragmentation caused by linear infrastructures. International Journal of Applied Earth Observation and Geoinformation 5: 1-15.

Giakoumakis NM, Gitas IZ, San-Miguel J (2002) Object-oriented classification modelling for fuel type mapping in the Mediterranean, using LANDSAT TM and IKONOS imagery- preliminary results. In: Viegas X (ed) Forest Fires Research & Wildland Fire Safety. Millpress, Rotterdam

Landis JR, Koch, GG (1977) The measurement of observer agreement for categorical data. Biometrics 33: 159-174

Lang S, Blaschke T (2003) Hierarchical Object Representation–Comparative Multiscale Mapping of Anthropogenic and Natural Features. ISPRS Archives, Vol. XXXIV, Part 3/W8, Munich, pp. 17-19

Monserud RA, Leemans R (1992) Comparing global vegetation maps with the Kappa statistic. Ecological Modelling 62(4): 275-293

Nobrega RA, O'Hara CG, Quintanilha JA (2007) An object-based approach to mine road features for informal settlements near Sao Paulo, Brazil. In: Blaschke T, Lang S, Hay G (eds) Object-Based Image Analysis – Spatial concepts for knowledge-driven remote sensing applications. Springer, pp 595-613

Tortora R (1978) A note on sample size stimation for multinomial populations. American Statistician 32(3): 100-102.

Turner MG (1989) Landscape Ecology: The Effect of Pattern on Process. Annual Review of Ecology and Systematics 20: 171-197

Wilcove DS, McLellan CH, Dobson AT (1986) Habitat fragmentation in the temperate zone. In: Soulé ME (ed) Conservation Biology: Science of Scarcity and Diversity. University of Michigan, Sunderland, MA, pp 237-256

Acknowledgements

Chapter 4.4

Multi-scale functional mapping of tidal marsh vegetation using object-based image analysis

K. Tuxen, M. Kelly

Geospatial Imaging and Informatics Facility, College of Natural Resources, University of California-Berkeley, Berkeley, California, USA, karin@nature.berkeley.edu

ABSTRACT: Nearly half of the world's natural wetlands have been destroyed or degraded, and in recent years, there have been significant endeavors to restore wetland habitat throughout the world. Detailed mapping of restoring wetlands can offer valuable information about changes in vegetation and geomorphology, which can inform the restoration process and ultimately help to improve chances of successful restoration. We performed an object-based image analysis using color infrared aerial photography, which maps specific wetland functions at multiple scales. The combined results of our work highlight important trends and management implications for monitoring wetland restoration using remote sensing, and will further enable restoration ecologists to use remote sensing for tidal marsh monitoring. Restoration objectives, ecosystem function, and scale can be integrated into monitoring techniques using remote sensing for improved restoration monitoring.

1 Introduction

1.1 Tidal marsh mapping

Effective tidal marsh restoration necessitates (1) the statement of specific restoration goals and objectives; (2) an understanding of how marshes will

evolve as they are restored, which can inform the restoration process; and (3) monitoring during and after restoration (Williams and Faber 2001). Restoration goals and objectives include improvements to the habitat and ecosystem services that tidal marshes provide (Baird 2005), which include biologically rich and diverse habitat, as measured by vegetation composition, configuration, and diversity, and the provision of adequate habitat for endangered species (Philip Williams & Associates Ltd. and Faber 2004).

Monitoring changes in ecosystem function after restoration can help managers understand how a marsh is evolving, which can inform the restoration process and ultimately improve changes of successful restoration. Effective monitoring should be objective, cost-effective, and as automated and non-invasive as possible (Andresen et al. 2002). Remote sensing technologies meet these requirements, and have been used to monitor wetlands around the world, allowing for the detailed mapping of restoring wetlands. The mapping of tidal marsh vegetation and habitat is a necessary and important part of wetland restoration monitoring. The heterogeneous landscape provides habitat for numerous organisms as well as the productive basis for the estuarine food web. Monitoring the pattern of vegetation communities over time can help track changes in both the pattern of a wetland landscape and the underlying biological processes to which it contributes, and thus it is important to accurately map vegetation for measurement of cover, complexity, and habitat use.

There are numerous ways to map and delineate continuous and complex landscapes using remotely sensed imagery (Blaschke and Hay 2001). The traditional mapping process stresses the production of the best single map with the highest overall accuracy, but given that ecosystems provide services or functions differently at different scales, the scale at which a landscape is mapped ideally should depend on the end use or purpose of the map (Burnett and Blaschke 2003). For example, mapping habitat for two different species across a wetland might require two separate mapping efforts, as a bird utilizes a wetland differently than a mouse. We aim here to develop a more flexible, multi-scale approach at wetland mapping that integrates hierarchical patch relationships across multiple ecosystem functional scales.

In this chapter, we utilize object-based approaches to map a restoring tidal brackish marsh; the method embraces a multi-scalar approach to mapping, and is thus more synergistic with our concept of wetland habitat functioning. Object-based image analysis (OBIA) has the potential to map tidal marsh ecosystems in such a way where objects are extracted at multiple scales from one or more images, and hierarchically linked with each other. The objects that are extracted, as well as the hierarchical relationships that are created, depend upon the map purpose (Blaschke and Hay

2001), such as a target species, usually endangered or threatened, for which the marsh is being restored. We mapped a brackish marsh using OBIA with two target ecosystem functions, both operating at different scales: (1) song sparrow habitat, and (2) salt marsh harvest mouse habitat.

1.2 The need for OBIA in wetland mapping

1.2.1 Past image analysis methods

Remote sensing is highly effective for analyzing estuaries and coastal systems (Phinn et al. 2000; Klemas 2001; Yang 2005), and has been used to map, monitor, detect and predict change in wetlands (Zhang et al. 1997; Jensen 2000). Remote sensing is ideal for monitoring restoring wetlands because it is cost-effective, time-efficient, and non-invasive. It allows for a high intensity of measurements in relatively inaccessible and sensitive sites, without the potential invasiveness that traditional field methods present to delicate habitat conditions, bird nesting territories, or endangered species habitat (Shuman and Ambrose 2003). It also allows for broad-scale estimation of many parameters valuable to ecologists, including land cover, vegetation structure, biophysical characteristics, and habitat areas (Wulder et al. 2004).

Visual interpretation, or manual delineation, of remotely sensed imagery is still a common method in wetland mapping (Andresen et al. 2002), as it allows for the delineation of accurate boundaries around objects and the production of maps that are visually appealing. While effective for some, visual interpretation can be expensive and time-intensive to achieve detailed classification results since these manual methods are not automated, and maps by different interpreters or at different time periods can produce variable results that are not comparable across space or time (Blaschke and Hay 2001). Automated (computer-assisted) image analysis approaches are becoming more common for wetland restoration projects, including those that involve pixel-based and object-based methods. Automated pixel-based classifiers have been gaining in popularity over the past decade because computational power has made them more operational. The use of automated image classification reduces inconsistencies and error introduced through visual photo-interpretation of imagery, and classifications between sites and between time periods are more consistent. In addition, automated methods have been found to be more cost-effective than visual delineation and classification (Thomson et al. 2003).

Unsupervised classification like Iterative Self-Organizing Data Analysis (ISODATA) and K-means clustering are commonly used in wetland mapping (Ramsey and Laine 1997; Ramsey et al. 1998; Everitt et al. 1999; Lu-

netta and Balogh 1999; Everitt et al. 2004), due to their objectivity, reduced user input (Thomson et al. 1998), and lower demand for knowledge of ground information. However, the lack of analyst input can produce clustering of different species within the same class, with less control by the analyst to how pixels are assigned a class. Supervised methods afford more control to the analyst, because training data are used to correlate the ground data in known locations with spectral properties, and then vegetation classes are extrapolated to the entire image. Several algorithms are commonly used in the wetland classification, including Maximum Likelihood Classifier (Scarpace et al. 1981; Jensen et al. 1984; Munyati 2000; Thomson et al. 2004; Isacch et al. 2006), Spectral Angle Mapper (Artigas and Yang 2005; Belluco et al. 2006), and Spectral Mixture Analysis (Underwood et al. 2003; Rosso et al. 2005). However, because wetland species are easily spectrally confused and brackish marshes exhibit high species diversity in a spatially complex arrangement, supervised classifications can also potentially result in unsatisfactory results. A hybrid approach between classification methods is often the most accurate of all the pixel-based methods (e.g. Kelly et al. 2004), as it decreases the time and amount of data needed with supervised classification because it can separate easily distinguishable classes to isolate more spectrally-conflicted vegetation types.

1.2.2 The case for OBIA

There are several reasons why OBIA holds promise as a method for classifying tidal marsh vegetation with high-resolution (less than one-meter pixel size) imagery. First, in contrast to traditional pixel-based methods, OBIA allows for the segmentation of one image into segments at multiple scales (Schiewe et al. 2001), allowing scales to be linked together, and to model the hierarchical nature of complex systems, such as tidal brackish marshes. Homogeneous objects are extracted at multiple scales from a single image and are linked across scales with rules of inheritance, enabling multi-scale hierarchical analyses. Pixels, on the other hand, are uni-scale and represent a fixed area on the ground (Benz et al. 2004). While pixel-based classification methods essentially cluster these pixels into "objects," they are usually non-hierarchical and single-valued. Manual delineation methods also "segment" an image, but can depict only the one scale perceived by the image interpreter (Burnett and Blaschke 2003).

A second reason why OBIA holds promise for tidal marsh mapping is because detailed wetland mapping necessitates high-resolution imagery in order to capture the small patches that make up the heterogeneous landscape. Coarse imagery does not effectively capture the fine-scaled features;

for example, Dobson et al (2001) used Landsat to monitor coastal wetland in the US, but could only map very broad areas of wetland, and could provide no within-site detail. Pixel-based image analysis and classification are limited with high-resolution data and can render unsatisfactory results. Object-based approaches are especially good for high spatial resolution data because neighboring pixels more likely belong to same class (Blaschke and Hay 2001; Schiewe et al. 2001). Variability between neighboring pixels becomes advantageous information because it now defines the internal heterogeneity, or texture, of an object, thus expressing texture more explicitly than with pixel-based approaches (Blaschke and Hay 2001). Using objects instead of individual pixels does not produce the "speckled" or "salt-and-pepper" results that are common in pixel-based classifications (Yu et al. 2006; Guo et al. 2007), so no post-classification filtering or smoothing is needed.

Third, ecosystem mapping is more ecologically sound with OBIA because principles of landscape ecology are maintained. Patches are based on homogeneous objects that incorporate the inter-patch variability rather than pixels (Andresen et al. 2002), which in turn make up the entire landscape in all its heterogeneity. Thus, object-based methods follow ecological phenomena more closely than traditional pixel-based methods (Blaschke and Strobl 2001), which analyze each pixel independently without taking into account spatial concepts like neighborhood, proximity, and homogeneity (Burnett and Blaschke 2003). Likewise, the hierarchical relationships between objects at multiple scales represent the multi-scale nature of complex ecosystems such as tidal marshes, with patterns and processes interacting across multiple scales. Therefore, a major goal with OBIA is to segment out patches that represent meaningful objects based on a specific level of homogeneity.

In summary, OBIA combines the advantages of visual interpretation and pixel-based methods in that patches are delineated into homogeneous areas that are both accurate and visually appealing, and the methods are objective, automated, and repeatable. OBIA allows for more semantic, intuitive, and human-conceived object shapes (Blaschke and Strobl 2001; Schiewe et al. 2001) that are based on user knowledge (Hay et al. 2003). Shape and context are taken into account as well as color or spectral quality of the patch (Schiewe et al. 2001). In addition, soft classifications, or fuzzy modeling, that are based on user knowledge can be integrated into the analyses. In this way, an object is assigned to multiple classes at varying degrees of membership, in order to reflect the ecotonal nature of many systems.

In the past, representing pattern and processes in this multi-scale manner was difficult because hierarchical linking of scales was not well defined (Hay et al. 2003). Geographic Information Science (GIS) was often con-

sidered the closest thing to multi-scale mapping because different scales could be analyzed separately, but GIS had some problems representing more than one scale at a time. OBIA has the potential to model landscapes in a multi-scale manner because a single image represents a variety of scales and levels of abstraction (Hay et al. 2003). Object segmentation uses the same object boundaries across scales (Hay et al. 2003), so topological relationships can be utilized. In this way, OBIA can integrate GIS analyses while representing the hierarchical scaling of real-world ecosystems which humans can mentally move between easily, but previously could not adequately model in a GIS (Blaschke and Strobl 2001).

Wetland ecosystems often consist of small isolated wetland patches, with complex pattern. The use of high spatial resolution imagery is necessary to capture the detail. There are several studies published using object-based techniques for mapping wetlands. Some have used OBIA with coarse-scale Landsat imagery (Bock 2003; Dorren et al. 2003; Stankiewicz et al. 2003; Yoon et al. 2003; Hurd et al. 2006), while others have used it with high spatial resolution images, such as CIR aerial photography (0.2 – 1 m) (Ivits et al. 2002; Burnett et al. 2003), Quickbird satellite imagery (60 cm – 2.4 m) (Wang et al. 2004a; Hurd et al. 2005), and IKONOS satellite imagery (Blaschke and Hay 2001; Hall et al. 2004; Wang et al. 2004a).

Multiple studies have compared OBIA with pixel-based methods. Wang et al (2004b) compared the two methods together and found that object-based methods were more efficient at differentiating spectrally mixed vegetation classes, but over-generalized species diversity in areas where spectral differentiation was clear. Therefore, they integrated both methods by applying object-based methods only to those classes that were spectrally similar, and pixel-based methods to all other classes (Wang et al. 2004b). Andresen et al. (2002) applied OBIA to IKONOS satellite imagery to map aquatic vegetation and found that they could successfully map vegetation both for inventory purposes using hierarchical segmentation and classification, and for monitoring purposes by measuring incremental patch change. Stankiewicz et al (2003) found that while shrub classes could be successfully delineated, many wetland patches could not be identified with coarse-scale Landsat imagery. Hurd et al. (2005) discovered that results from the object-based multi-scale methods on Quickbird imagery were smoother and reduced errors of commission and omission. Harken and Sugumaran (2005) found object-based results from eCognition software to be nearly 30% more accurate overall than pixel-based results from Spectral Angle Mapper (SAM). Lathrop et al (2006) found that traditional per-pixel based multi-spectral classification approaches were not as effective as object-based approaches, due to the varying image radiometric conditions, in part caused by water depth, turbidity, and background bot-

tom sediment. The majority of applied OBIA studies found used the commercially available software Definiens Professional (formerly known as "eCognition") made by Definiens, Inc. (1995-2006). Using this software, studies have implemented multi-scale object-based segmentation and classification of mires (Burnett et al. 2003). Others have integrated multiple data sources (Stankiewicz et al. 2003; Li and Chen 2005), such as elevation data, to help further discriminate classes. Some have experimented with hyperspectral data, such as Greiwe and Ehlers (2005), who found an almost 20% accuracy improvement when they used hyperspectral information rather than just high spatial resolution data. All of these studies demonstrated an increase in accuracy over pixel-based image analysis, but none discuss the multi-scale mapping of ecosystem function. OBIA has the ability for powerful mapping, modeling, and visualization across multiple scales for many future applications, including restoration. In this chapter, we incorporate multi-scale restoration goals into tidal marsh vegetation and habitat mapping.

1.3 Multi-scale nature of brackish marshes

The multi-scale nature of brackish marshes make them complex systems as defined by Hay et al (2001), in that they exhibit multi-scale patterns and processes, have many interactive components which are non-linear in nature, and has a high level of spatial heterogeneity (Hall et al. 2004). They are the ideal ecosystem in which to apply object-based image analysis, due to the multiple and hierarchically-linked scales involved in wetland functioning.

Brackish marshes have biological and physical components at multiple scales, arranged in a hierarchy and interacting in multiple ways. For example, fine-scale micro-topography causes highly variable salinity and inundation regimes that produce patches with high levels of plant species diversity. These patches are nested in larger patches, forming species assemblages, or vegetation community types, which exhibit larger-scale landscape heterogeneity. Many organisms inhabit brackish marshes, each viewing and using the marsh from a different perspective, or scale. For example, the song sparrow (*Melospiza melodia)* and the salt marsh harvest mouse (*Reithrodontomys raviventris*) have different habitat requirements, both defined differently by the composition, configuration, patch size, landscape context, and spatial extent of vegetation species (Phinn et al. 1996). In the next two sections, we give information about the two target ecosystem functions in this study, song sparrow habitat and salt marsh

harvest mouse habitat, that will be used to determine rule sets used in the classification and habitat mapping.

1.3.1 Song sparrow

Three subspecies of song sparrow are endemic to the San Francisco Estuary Region (Alameda Song Sparrow, Samuel's Song Sparrow, and Suisun Song Sparrow) and are Species of Special Concern for California due to loss of wetland habitat (Nur and Spautz 2002). Restoration efforts are underway to increase the breeding and nest success of all song sparrows subspecies. Sparrow habitat is related to vegetation diversity, productivity, and species composition, and includes areas near channels or levees, especially areas with high channel density (Nur and Spautz 2002). Sparrows are very particular about where they make their nests and need areas where they can watch for predators. One major factor is vegetation species, as song sparrows prefer to perch and often nest in shrub cover, especially gumplant (*Grindelia stricta* var. *stricta*) and coyote brush (*Baccharis pilularis*) (Nur and Spautz 2002), in which they can conceal their nests. Gumplant and coyote brush are normally on levees (both natural and human-made) or other elevated areas in salt and brackish marshes (Traut 2005). Song sparrow nest predation also has been linked to proximity of upland edge (Stralberg and Chan 2002); however, since the site for this study is an island, the upland levee edge is surrounded by water, hindering predation. Also, song sparrows have been found to be negatively affected by rushes (*Juncus* spp.) (Nur and Spautz 2002).

1.3.2 Salt marsh harvest mouse

The salt marsh harvest mouse is endemic to the salt and brackish marshes of San Francisco Estuary. The species was listed as federally endangered in 1970 and as state endangered the following year, mainly due to habitat loss or modification (Geissel et al. 1988). The salt marsh harvest mouse is cover-dependent on pickleweed (*Sarcocornia pacifica*), a succulent halophyte common to San Francisco Estuary and brackish marshes. Due to its high salinity tolerance, pickleweed occurs in areas along channels where natural berms have formed from sedimentation and in high marsh plains, both of which have saline soils due to high evaporation levels in the summer months and relatively limited flooding during high tides (Onuf 2006). Salt marsh harvest mice prefer dense canopy cover of pickleweed of 30-60 cm in height (Shellhammer et al. 1982; Bias and Morrison 2006), with nearby areas of high elevation where they can move to during high tides, but only if those areas are densely covered with low-lying vegetation,

similar to pickleweed, such as alkali heath (*Frankenia grandifolia*) (Shell-hammer 1989).

   The mice can disperse locally, but only if there are small vegetated distances of 20m or less between the preferred pickleweed patches (Shell-hammer et al. 1982; Geissel et al. 1988). Therefore, small patches can serve as stepping-stones for species dispersal or recolonization and can protect scattered individuals, although larger areas of pickleweed tend to host the species in more numbers than small patches (Shellhammer et al. 1982). While the mice mostly prefer pickleweed habitat, they have been known to exist in dense canopy of mixed species, as long as pickleweed is a co-dominant species. They avoid pure stands of certain species, including salt grass (*Distichlis spicata*) and alkali bulrush (*Bolboschoenus maritimus*) (Shellhammer et al. 1982). Also, the context of adjacent vegetation types, such as upland/levee areas, can allow the salt marsh harvest mice an escape from high tides (Shellhammer et al. 1982).

2 Methods

2.1 Study site

Roughly 90-95% of tidal wetlands in the San Francisco Estuary in California, USA, have been lost or altered since European settlement. Bull Island on the Napa River in California, USA is an intertidal brackish marsh about 40 hectares in size, located 16 kilometers upstream from the mouth of the Napa River in the San Francisco Estuary, California (Figure 1). The site was naturally restored to tidal influence in 1954 when a flood breached the levee. Today, it serves as habitat for organisms such as birds, mammals, and fish, increased productivity for the Napa River and San Pablo Bay, and provides recreation activities for Bay Area residents. The site is completely vegetated and surrounding by a levee except for in the southwest corner of the site, where the main drainage basin with an extensive channel network connects with the Napa River.

   The SF Bay-Delta Region includes a particular habitat structure that is supported by complex heterogeneous patterns of diverse vegetation communities common to Pacific Coast brackish tidal marshes (García et al. 1993; Callaway and Sabraw 1994; Zedler et al. 1999). While salty tidal marshes have distinct vegetation zones, brackish and freshwater tidal marshes are characterized by diverse species mixtures, usually composed of several co-dominant plant species with numerous sub-dominants, and configured in a heterogeneous mosaic of vegetation community types with sometimes poorly defined boundaries between patch types. Dominant

vegetation at Bull Island consists of alkali bulrush (*Bolboschoenus mari-timus*), perennial pickleweed (*Sarcocornia pacifica*), cattail (*Typha* spp.), Common tule/bulrush (*Schoenoplectus acutus* var. *occidentalis* and *Schoe-noplectus californicus*), and gumplant and coyote brush on the levees.
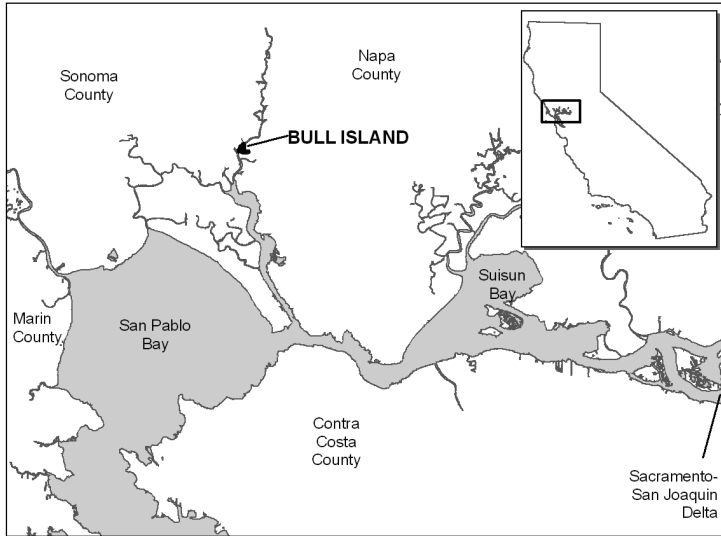
Fig. 1. Map of study site vicinity, Bull Island, Napa County, USA

2.2 Data

Color-infrared (CIR) aerial photography was acquired on October 24, 2003 as part of the Integrated Regional Wetland Monitoring Project (IRWM Project 2003). The three-band (near infrared, red, and green) photo was scanned at 1,200 dots per inch (dpi) at a scale of 1:9,600, to render a pixel size of 0.2m on a side (0.04m^2 pixel area) (IRWM Project 2003), fine-scaled enough for effective mapping of highly diverse patches in a hetero-geneous pattern. The image was orthorectified using four ground control points spread out in different corners of the photo, in order to maximize the accuracy of the rectification process. The targets were surveyed hori-zontally using a sub-meter-accuracy Trimble® GeoXT^{TM} mapping grade Global Positioning System (GPS) unit (Trimble Inc. 2005). The images were orthorectified, to account for camera tilt, topographic displacement, and lens distortion. To account for the local terrain's effect on image dis-tortion, digital elevation models (DEMs) were used in the rectification process, using publicly available USGS 10m DEMs for the relatively to-pographically-homogenous sites. The accuracy standards employed for or-

tho-rectification were such that approximately 90% of all control points on the photos are within two meters of their corresponding ground coordinates (which, as stated above, are of sub-meter accuracy—generally within 0.3-0.75 m of the nominal XY coordinate) determined using the GPS unit (IRWM Project 2003). Imagery is freely-available at http://www.irwm.org/.

Because elevation is a strong determinant of vegetation zonation and landscape pattern in tidal marshes, fine-scaled topography imagery was integrated into the object-based classification. Light Detection and Ranging (LiDAR) data was acquired for the Napa River Basin in May and October 2003 by the University of Florida in partnership with UC Berkeley (NCALM 2005). The LiDAR data was acquired with an Optech 1233 laser, which collects data at a frequency of 33 kHz. Data density (unfiltered, last return) for Bull Island was 1.5 points per square-meter, which was interpolated to a 1-meter LiDAR raster file. The RMSE for the vertical accuracy was 0.15 cm.

2.3 Field data collection

We collected ground reference data to train the classification of the image into vegetation classes and to assess accuracy of the final map. The field effort was carried out in two phases. During the first phase, ground data were collected for use in choosing sample objects to train the classification. We collected data at: (1) points that were randomly-generated in a GIS prior to field visits, and (2) points that were chosen during the field visit at the discretion of the field samplers, to better train the image classification. This was necessary since we had little prior knowledge of the vegetation that occurred in the sites. During the second field phase, we collected data at randomly-generated points to be used to assess the accuracy of the classification. The points were stratified by vegetation classes and weighted by the area of the vegetation class so that larger classes consisted of more points.

At all points for both field phases, we recorded absolute percent cover of each species within a circular plot with a three-meter radius, and the dominant species was used for classification training and accuracy assessment. Percent cover was determined with ocular estimates, using a seven-category Daubenmire ranking system (1:0-2%; 2:2-5%; 3:5-25%, 4:25-50%; 5:50-75%; 6:75-95%; and 7:95-100%) (Daubenmire 1959). Two-hundred twenty-eight point locations were visited; of those, 68 were used to train the classification and 160 were used for the accuracy assessment. Handheld GPS units with an average recorded location error of five meters

were used to navigate to each point. Due to delicate conditions at some of the sites, including soft sediments and bird nesting territories, samplers had the choice to observe the sample point from a few meters away. This practice was performed on less than 10% of the data points.

2.4 Image analysis for vegetation type mapping

We used a five-step multi-scale segmentation/object relationship modeling (MSS/ORM) approach described by Burnett and Blaschke (2003) including: (1) GIS building, or the collection of georeferenced geographic information to assist in the analyses; (2) segmentation, or the partitioning of the imagery into multiple segmentation levels or scales; (3) object relationship model building, or the assignment of linkages between the hierarchical levels; (4) visualization, or the production of the classified image(s) for a specific goal; and (5) quality assessment, or the measurement of accuracy. This methodology provides an effective organized structure with which to analyze complex systems. Mapping a function is challenging because functions operate across multiple scales. However, by combining information from different resolutions, the scale can be optimized for each target function (Burnett and Blaschke 2003). Based on knowledge about each target function (described in Section 1.3), we created rule sets that built the object relationship model.

2.4.1 GIS building

The CIR aerial photo image and the LiDAR image were the primary data used for the analysis. In addition, we created an image with three principle component (PC) bands derived from the raw image bands. The first PC band represented the highest variation within the data, while the second and third represented most of the remaining variation (ERDAS 1999). We found that the second and third PC bands depicted obvious patterns in the landscape, representing the vegetation communities important to our target functions. We also created a Normalized Difference Vegetation Index (NDVI) image to indicate presence of vegetation and as a proxy for primary production (Jensen 2000). We equalized the histograms for both the second and third bands of the principal components image and the NDVI image. Many have found that it is very helpful to integrate vegetation indices and image transformations into the analysis and classification process to aid in vegetation mapping (Eastwood et al. 1997; Zhang et al. 1997; Thomson et al. 2004), and some studies have used image transformations such as Principal Components Analysis (Munyati 2004) with success. All

six layers were inputted into Definiens Professional software: the near infrared, red, and green bands of the CIR photo, the LiDAR image, and the PCA and NDVI images (both with their histogram equalized).

2.4.2 Segmentation

Segmentation is critical to multi-scale object relationship modeling (Burnett and Blaschke 2003), and the first step to accurate mapping using OBIA. The segmentation used in this study was the result of numerous trials using various segmentation parameters. The final segmentation levels we used rendered desirable objects for both the target ecological scales in this study: (1) a medium patch, or class level, scale for mapping salt marsh harvest mouse habitat (depicted by the second segmentation level), and (2) a small patch scale for mapping song sparrow habitat (depicted by the third segmentation level). Definiens Professional version 5.0 was used for all object-based image analysis in this study. Definiens Professional contains two groups of classification tools: nearest neighbor (NN) classifiers and knowledge-based rule sets. Our image analysis used both the data-driven nearest neighbor classification, as well as the knowledge-driven crisp and fuzzy membership functions.

For this study, multi-scale segmentation was carried out in three major steps in a top-down manner. The first segmentation was performed using both the second and third PC bands. We used a scale parameter of 100, with the following homogeneity parameters: color versus shape (0.2 and 0.8 respectively), and compactness versus smoothness (0.2 and 0.8 respectively). Weighting shape more than color uses less of the spectral values of the pixels and gives more influence to shape (whether that is defined as compactness or smoothness). We chose to weight smoothness over compactness to prevent too much complexity in the object boundaries, and to render more appropriate habitat patches. Smaller patches were pulled out here, which mainly utilized the variable pattern that the second PC band depicted. The purpose of this large segmentation level was to delineate the large vegetation groups out, mainly differentiated by plant vigor, or level of chlorophyll in the leaves, of which the patterns in the PC bands illustrated. This top segmentation rendered objects that effectively separated water and vegetation, but objects were not small enough to be used for mapping vegetation types or targeted functions. As PCA aims to transform the data to produce axes of greatest variation, using these bands for segmentation allows for effective partitioning of the variation between objects, better than with the CIR image.

Both the second and third segmentations were performed with the three CIR bands: the near infrared, red, and green bands. We used a scale pa-

rameter of 50 for the third (medium patch) segmentation and a scale parameter of 25 for the fourth (smallest patch) segmentation, and the following homogeneity parameters for both: color versus shape (0.9 and 0.1 respectively), and compactness versus smoothness (0.5 and 0.5 respectively). The second and third segmentation were used for the habitat mapping of the salt marsh harvest mouse and song sparrow, respectively. The LiDAR image was not used in segmentation because its pixel size (1 m) was much larger than the pixel size of the multispectral image (0.2 m). Instead, the LiDAR image was only used in the classification, following recommended strategies for segmentation by Definiens (1995-2006).

2.4.3 Object relationship model building

In this step, we set the rules that defined the relationships between the multiple scales in the object hierarchy. The class hierarchy is the first step in building object relationships, and represents two ways in which the different land cover classes are related: (1) by inheritance, or the way in which they possess similar spectral or contextual characteristics, and (2) by semantic group, or the ecologically-meaningful way we want to depict the land cover groups at various levels of detail. For example, a patch of bulrush and a patch of coyote brush might share similar spectral traits and therefore fall under similar inheritance structure, but they belong in two different semantic groups (the former in the low marsh and the latter in the high marsh). In this study, we used the semantic organization for visualization of the marsh functions; therefore, we describe this in the visualization step. First, we describe the inherited relationships between the land cover classes that are used to build the functional maps.

The class hierarchy for the vegetation type map was designed with several tiers of inheritance in order to optimize separation between classes. At the top tier, we differentiated between water and vegetation, by applying two rule sets: to NDVI and to LiDAR elevation (Figure 2). At the second tier, we further separated vegetation objects into two groups based on low and high NDVI values (Figure 2). The reason for this was certain vegetation types had low NDVI values at the time of image acquisition (i.e. cattail), while others had high NDVI values (i.e. pickleweed), observable in the image. The third tier further divided each of the two NDVI groups into levee and non-levee groups, based on LiDAR elevation. The fourth tier further divided the two levee groups into vegetation types (with the levee group having one type and the non-levee group having eight types). At this tier, objects were assigned into each individual vegetation types based on training samples and standard nearest neighbor classification using object means of the NIR, red, and green image layers.
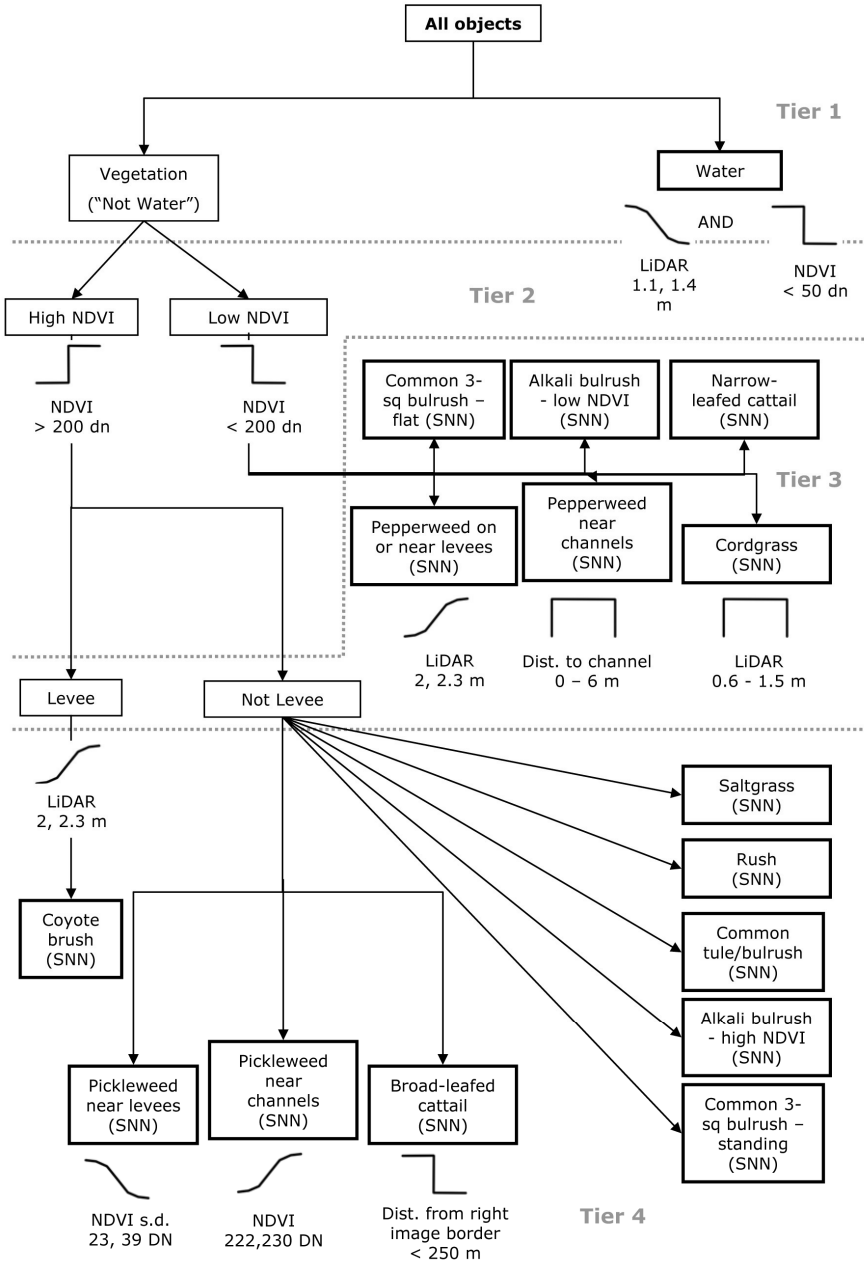
Fig. 2. Hierarchical object-based classification methodology. All Standard Nearest Neighbor classification algorithm was performed using object means of the NIR, red, and green image layers. The gray dashed lines depict each tier of the classification

In addition to the separations at the four tiers just described, we also employed several other rule sets that improved the classification of certain vegetation types. The reasons for this were because many objects were still confused between two or more vegetation types even at the fourth tier. Two classes that often had confusion were "pickleweed" and "common tule," as both these classes share similar NIR reflectance in the CIR image. After investigation, we found that the feature describing the average standard deviation of NDVI separated pickleweed objects from common tule objects very well, and assigned a fuzzy positive relationship to the pickleweed class (Figure 2). Another spectrally confused pair was "cattail" and "pepperweed (near channels)." In order to separate these two classes, we used a distance to channel feature to distinguish pepperweed, since it nearly always bordered the channels. Therefore, we assigned a crisp negative relationship of pepperweed to distance to water with a threshold of 6 meters. A third confused pair were the two main types of cattail: narrow-leafed cattail and broad-leafed cattail. These two species are so taxonomically similar that we did not key to species in the field. They also have very similar spectral characteristics, but have different reflectance in the NIR and red bands, therefore having different NDVI values. The broad-leafed cattail species were showing up very similar to other vegetation types in the southeastern part of the site; therefore a contextual feature was added to the classification that forced that cattail in the part of the site that was 250 m from the edge of the image to be classified as broad-leafed cattail. Finally, the last knowledge-based rule set applied was for Pacific cordgrass, which was getting confused with numerous species and getting overestimated throughout the site, even when it only existed at certain elevations and next to channels. Therefore, we assigned a crisp relationship that assigned any LiDAR values (with the cordgrass spectral characteristics) that fell within the range of 0.6 and 1.5 m to be classified as cordgrass.

In order to effectively map our two target habitats, we prioritized marsh features that described optimal conditions for each. In this section, we describe features that improve or reduce habitat quality for song sparrows and salt marsh harvest mice. Using these features, we created rule sets that we used to build the object relationship models for each function.

Song sparrows prefer gumplant and coyote brush shrub species for nest sites, and may also exist in other shrub species where their nests are concealed. In Bull Island, all shrub species occur on or near the levee, or upland, areas. Some gumplant and coyote brush plants exist in the areas of the perennial pepperweed (*Lepidium latifolium*) surrounded by pickleweed in the areas by the levee where there are large patches of pickleweed and other high marsh species. Song sparrow habitat is negatively affected by

rushes (*Juncus* spp.), due to its simple structural complexity, low height, and sometimes low density. In addition, Song sparrow habitat is positively correlated with distance to water (Stralberg, unpublished data). From characteristics, we created three rule sets that modeled Song sparrow habitat:

- Patches that are either gumplant or coyote brush;
- Patches that are levee; and
- Patches that are not in rush (*Juncus* spp.) patches.

The salt marsh harvest mouse is cover-dependent on pickleweed, and only moves between patches if there is a narrow band of dense vegetation. They have been found to move between patches that are up to 20 meters apart. While the mice will exist in areas with saltgrass mixed with pickleweed, they do not prefer pure stands of saltgrass (*Distichlis spicata*), most likely due to the low lying nature of saltgrass without a dense canopy. In addition, they less frequently inhabit areas near water than away from water, preferring to be near levees or some upland area to which they can retreat during the highest tides. From these characteristics, we created two rule sets that modeled salt marsh harvest mouse habitat:

- Patches that are pickleweed; and
- Patches that are not dominantly saltgrass.

2.4.4 Visualization

In addition to representing the class hierarchy with inherited relationships, Definiens Professional software also allows for the organization of classes into semantic groups, which aid in more ecologically meaningful groupings. We organized the classes to reflect the rule sets for the functional maps. For the salt marsh harvest mouse habitat, we assigned a class ("SMHM habitat") to encompass both the pickleweed habitat near channels and in the areas near the levee. For the song sparrow habitat, we assigned two new classes. One was "Song sparrow habitat" and was placed in the non-levee class under the high productivity areas. The other class added was rush. Rush was not part of any other mapping purpose due to the fact that its patches were too small to be captured in any other segmentation level. Both these classes were assigned to a nearest neighbor classifier using training object samples. The scales are linked explicitly because they use the same information from the original image and call on one another for rule sets.

We performed a classification at the two target functional scales: the smallest segmentation level (scale parameter of 25) for the Song sparrow

habitat, and the medium segmentation level (scale parameter of 50) for the salt marsh harvest mouse habitat.

2.4.5 Quality assessment

An accuracy assessment was performed on the vegetation map and on both the habitat maps (song sparrow and salt marsh harvest mouse) by comparing mapped class to ground reference vegetation at all ground reference points.

3 Results

The final vegetation map is depicted in Figure 3, and the multi-scale functional maps are depicted for song sparrow habitat in Figure 4, and for salt marsh harvest mouse in Figure 5. The overall accuracy of the final vegetation map was 85%. The error matrix for the accuracy assessment of the final vegetation map is shown in Figure 6. The overall accuracies of the two habitat maps were 95% for song sparrow habitat and 94% for salt marsh harvest mouse habitat.

Results show that song sparrow habitat mainly exists along the levees where the shrub species, including coyote brush and gumplant, are located. Likewise, salt marsh harvest mouse habitat exists along channels and near the levees where pickleweed cover exists either alone or in combination with other suitable species.
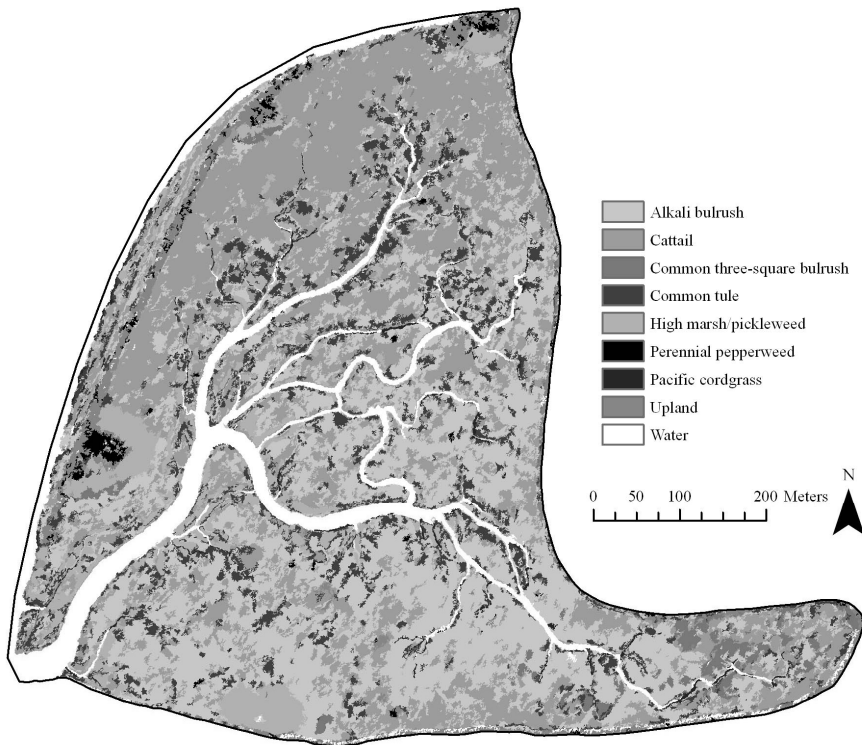
Fig. 3. Map of vegetation types at Bull Island, produced at a scale parameter of 25
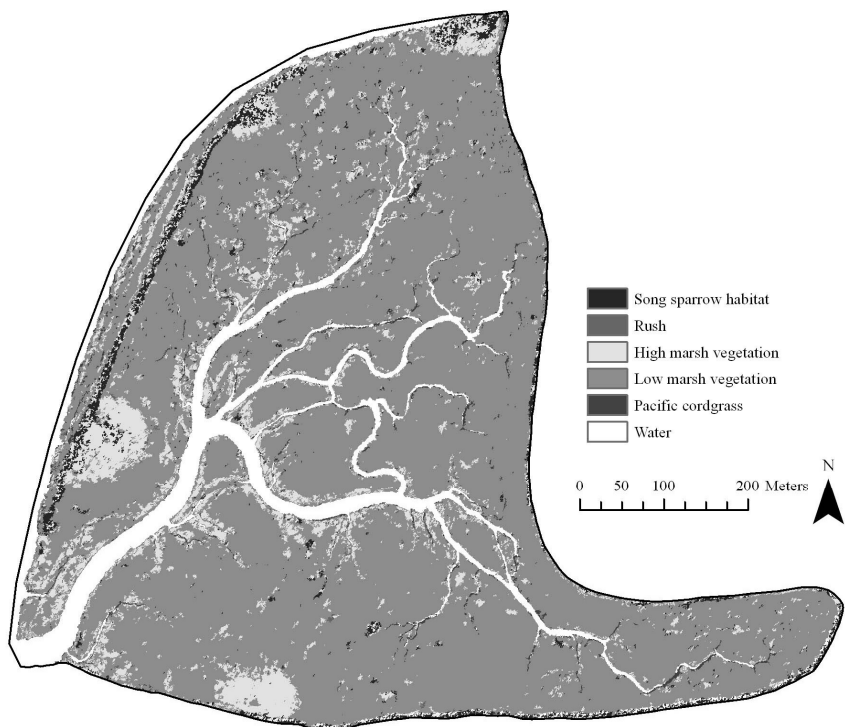
Fig. 4. Map of Song Sparrow habitat at Bull Island, produced at a scale parameter of 25.
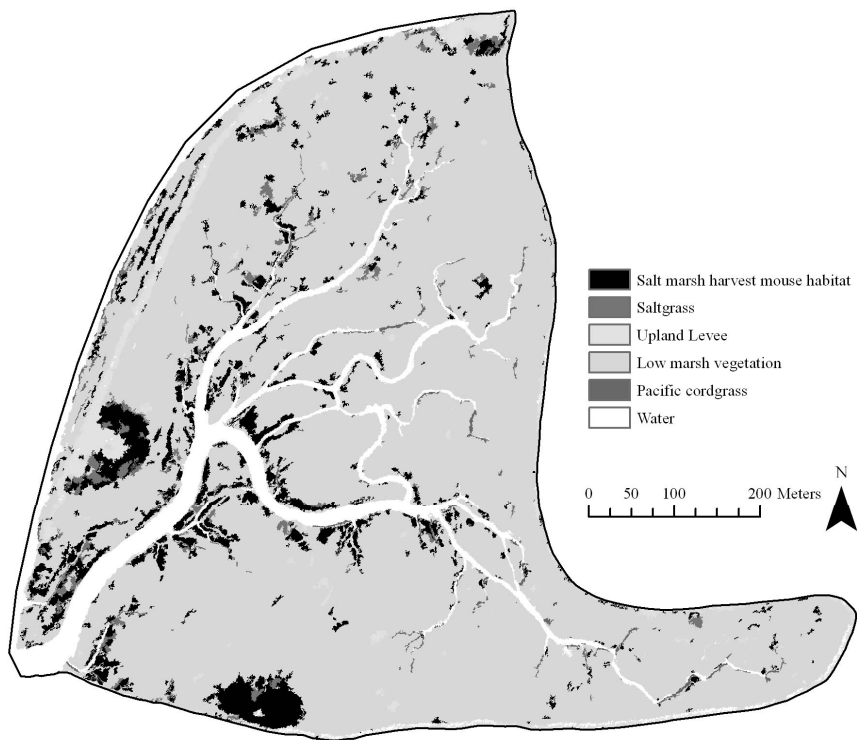
Fig. 5. Map of Salt Marsh Harvest Mouse habitat at Bull Island, produced at a scale parameter of 50.

		REFERENCE CLASS								
		1	2	3	4	5	6	7	8	9	
1	39	1	4	1	2				1	48
2	2	45	3	3						53
3			5							5
4				14		1				15
5		2			17	1				20
6						3				3
7										0
8			2		1			7		10
9									6	6
	41	48	14	18	20	5	0	7	7	**160**

OVERALL ACCURACY = 136/160 = 85.0%

Fig. 6. Error matrix for the accuracy assessment for the final vegetation map.

4 Discussion

The results of the multi-scale functional mapping were useful, but have limitations in their use. The habitat maps provide means for useful inventory or baseline to monitor change after restoration; however, they do not necessarily map habitat suitability or weighted preference across a landscape. For example, a higher numbers of salt marsh harvest mice may be found in the pickleweed bordering the levee than in the pickleweed near the channel, due to their preference for the former. However, both land cover classes are important for the mice and need to be mapped. The mapped products that resulted from this study are the integration of habitat preferences into a vegetation mapping process for the combined approach of mapping both land cover and function, which will aid in monitoring and management decisions. An added benefit to this approach is that the accuracy of habitat maps has the potential of being higher than a detailed vegetation map as it was in this study. This is because multiple vegetation type classes are often combined in one habitat class, potentially lumping together vegetation types that might be confused with each other and therefore misclassified in the detailed vegetation map. This study was a habitat mapping study that integrates monitoring goals into vegetation type mapping to aid in restoration management. It is not a suitability modeling study, but rather a method for including habitat traits into map making, nested in the overall object-oriented approach, for ease of detection of change in habitat over time. As we have classified fine scale patches of homogeneous vegetation type, and then aggregated them based on the rule sets, the accuracy assessment for the functional habitat maps are essentially validating the accuracy of our aggregated fine scale vegetation map. Careful thought should go into deciding which layers to use for segmentation and classification. Some data can be effective at partitioning pixels into meaningful objects while being ineffective when used in the classification, and vice versa. For example, we did not use the LiDAR layer in the segmentation process, because the spatial resolution was coarser than the other datasets. However, it was used extensively in the classification process for the creation of multiple rule sets. On the other hand, the PCA bands were used for segmentation because they highlighted meaningful patterns and variability in the landscape, but were not used in the classification, as their values were not easily interpretable to spectral, spatial, or contextual characteristics.

   The vast majority of image analysis for wetland mapping uses pixel-based classification for distinguishing between types of land cover. Object-base image analysis has the ability to improve the mapping of vegetation,

habitat, and productivity for restoring wetlands. However, there is still much research that needs to be done, as far as investigating which bands are useful for classifying natural habitats that exhibit extremely complex and ecotonal boundaries between patches.

In complex systems like tidal marshes, multi-scale mapping is possible, but care must be taken to map for each scale, as certain training objects, thresholds, and fuzzy relationships are not often transferable across segmentation levels. In addition, segmentation is still an iterative and manual approach (Schiewe et al. 2001). While the user must know scales of interest, and OBIA requires exploratory work to define appropriate segmentation levels (Hay et al. 2003), multi-scale segmentation and classification can provide powerful results.

Finding the appropriate focal level, or segmentation level with the optimal object size and shape for what you are mapping, can be tricky and time-consuming. The Definitions Professional User Guide (1995-2006) recommends a trial-and-error process, but OBIA application could benefit from more studies to find systematic ways of choose segmentation parameters for specific types of ecosystems. There has, however, been several generalizations made that offer effective rules of thumb. For example, Blaschke et al (2002) recommend that the scale of image objects to be detected must be significantly bigger than the scale of image noise relative to texture.

Using the Definiens Professional software we were able to integrate multiple sources of data: in this case we integrated very high resolution (20 cm) multispectral data and high resolution (1 m) LiDAR elevation data. The software also enables the application of rule sets to the classification process incorporating our knowledge of the site and the vegetation, and performs fuzzy classification, which allow an object patch to be assigned to several classes but with varying membership.

Bull Island, California was chosen as the site for this study for three reasons. First, it contains all of the target vegetation species needed for Song sparrow and harvest mouse habitat. Second, the aerial photography was acquired in one tile due to its relative small site compared to the other sites. Therefore, effects of brightness gradients and other radiometric issues from airplane tilt were less problematic. Third, the software used in this study is limited in its processing speeds of imagery of certain sizes. The image for Bull Island was approximately 5000 x 6000 pixels, which although large, was manageable for a study of this magnitude.

OBIA usually requires extensive knowledge of the study site and its relationship to the data that is available for analysis. The determination of the most adequate spectral, spatial, and contextual features that tell each

class apart requires a high level of familiarity with both the vegetation types and the imagery.

5 Conclusions

In many ways, estuarine and coastal environments are a "last frontier" in remote sensing analysis, primarily because the problem of scale has yet to be solved (Cracknell 1999). Most satellite-based imagery does not have the spatial resolution required to distinguish small yet ecologically important wetlands from other land cover types, nor to distinguish and map the complex vegetation patterning and mixed species typically found within (particularly Pacific coast) wetlands with adequate accuracy (Cracknell 1999). Moreover, temporal resolution has typically been inadequate to capture important change in estuarine and coastal land, especially in the time intervals need to monitor ever-increasing human development or small changes in a restoring wetland.

With increases in spatial, temporal, and spectral resolutions produced by newer sensors in the past several years, and with even more expected in the future, wetland mapping and monitoring is expected to improve *if* we can use the data and imagery effectively and in innovative ways. Object-based image analysis offers methods to accommodate high spatial, spectral, and temporal resolution data, by reducing small pixels into meaningful objects, accommodating multiple spectral band ratios and relationships, and by offering important change detection methods that allow for the understanding of incremental change in vegetation type, size, shape, and context. Thus OBIA methods can be a benefit to understanding tidal wetlands in a restoration context.

6 References

Andresen, T., C. Mott, S. Zimmermann, T. Schneider and A. Melzer. 2002. Object-oriented information extraction for the monitoring of sensitive aquatic environments. IEEE 3083-3085.

Artigas, F. J. and J. S. Yang. 2005. Hyperspectral remote sensing of marsh species and plant vigour gradient in the New Jersey Meadowlands. International Journal of Remote Sensing **26**:5209-5220.

Baird, R. C. 2005. On sustainability, estuaries, and ecosystem restoration: The art of the practical. Restoration Ecology **13**:154-158.

Belluco, E., M. Camuffo, S. Ferrari, L. Modenese, S. Silvestri, A. Marani and M. Marani. 2006. Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing. Remote Sensing of Environment **105**:54-67.

Benz, U. C., P. Hofmann, G. Willhauck, I. Lingenfelder and M. Heynen. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. Journal of Photogrammetry and Remote Sensing **58**:239-258.

Blaschke, T. and J. Strobl. 2001. What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GeoBIT/GIS **6**:12-17.

Blaschke, T. and G. J. Hay. 2001. Object-oriented image analysis and scale-space: Theory and methods for modeling and evaluating multiscale landscape structure. International Archives of Photogrammetry and Remote Sensing **34**:22-29.

Bock, M. 2003. Remote sensing and GIS-based techniques for the classification and monitoring of biotopes. Journal for Nature Conservation **11**:145-155.

Burnett, C. and T. Blaschke. 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecological Modelling **168**:233-249.

Burnett, C., K. Aaviksoo, S. Lang, T. Langanke and T. Blaschke 2003. An object-based methdology for mapping mires using high resolution imagery. Ecohydrological Processes in Northern Wetlands, Tallinn. June 30 - July 4, 2003.

Daubenmire, R. F. 1959. Canopy coverage method of vegetation analysis. Northwest Science **33**:43-64.

Definiens Inc., 1995-2006. URL http://www.definiens-imaging.com/ [accessed on 17 January 2007]

Dobson, J. E., E. A. Bright, R. L. Ferguson, D. W. Field, L. L. Wood, K. D. Haddad, H. I. III, J. R. Jensen, V. V. Klemas, R. J. Orth and J. P. Thomas 2001. NOAA Coastal Change Analysis Program (C-CAP): Guidance for Regional Implementation.NOAA Technical Report National Oceanographic Atmospheric Association.139.

Dorren, L. K. A., B. Maier and A. C. Seijmonsbergen. 2003. Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. Forest Ecology and Management **183**:31-46.

Eastwood, J. A., M. G. Yates, A. G. Thomson and R. M. Fuller. 1997. The reliability of vegetation indices for monitoring saltmarsh vegetation cover. International Journal of Remote Sensing **18**:3901-3907.

Everitt, J. H., C. Yang, D. E. Escobar, C. F. Webster, R. I. Lonard and M. R. Davis. 1999. Using remote sensing and spatial information technologies to detect and map two aquatic macrophytes. Journal of Aquatic Plant Management **37**:71-80.

Everitt, J. H., C. Yang, R. S. Fletcher, M. R. Davis and D. L. Drawe. 2004. Using aerial color-infrared photography and Quickbird satellite imagery for mapping wetland vegetation. Geocarto International **19**:15-22.

Greiwe, A. and M. Ehlers 2005. Combined analysis of hyperspectral and high resolution image data in an object oriented classification approach. American

Society for Photogrammetry and Remote Sensing (ASPRS), Baltimore, MD. March 7-11, 2005.

Guo, Q. C., M. Kelly, P. Gong and D. Liu. 2007. An object-based classification approach in mapping tree mortality using high spatial resolution imagery. GIScience and Remote Sensing **44**:24-47.

Hall, O., G. J. Hay, A. Bouchard and D. J. Marceau. 2004. Detecting dominant landscape objects through multiple scales: An integration of object-specific methods and watershed segmentation. Landscape Ecology **19**:59-76.

Harken, J. and R. Sugumaran. 2005. Classificatino of Iowa wetlands using an airborne hyperspectral image: a comparison of the spectral angle mapper classifier and an object-oriented approach. Canadian Journal of Remote Sensing **31**:167-174.

Hay, G. J., T. Blaschke, D. J. Marceau and A. Bouchard. 2003. A comparison of three image-object methods for the multiscale analysis of landscape structure. ISPRS Journal of Photogrammetry & Remote Sensing **57**:327-345.

Hurd, J. D., D. L. Civco, M. S. Gilmore, S. Prisloe and E. H. Wilson 2005. Coastal marsh characterization using satellite remote sensing and *in situ* radiometry data: Preliminary results. American Society of Photogrammetry and Remote Sensing, Baltimore, MD. March 7-11, 2005.

Hurd, J. D., D. L. Civco, M. S. Gilmore, S. Prisloe and E. H. Wilson 2006. Tidal wetland classification from Landsat imagery using an integrated pixel-based and object-based classification approach. American Society of Photogrammetry and Remote Sensing, Reno, NV. May 1-5, 2006.

IRWM Project 2003. 2003 Imagery Report.h. w. i. o. Integrated Regional Wetland Monitoring Pilot Program (IRWM).

Isacch, J. P., C. S. B. Costa, L. Rodríguez-Gallego, D. Conde, M. Escapa, D. A. Gagliardini and O. O. Iribarne. 2006. Distribution of saltmarsh plant communities associated with environmental factors along a latitudinal gradient on the south-west Atlantic coast. Journal of Biogeography **33**:888-900.

Ivits, E., B. Koch, T. Blaschke and L. Waser 2002. Landscape connectivity studies on segmentation based classification and manual interpretation of remote sensing data. eCognition User Meeting, München, Germany. October 2002. 1-10.

Jensen, J. R., E. J. Christensen and R. Sharitz. 1984. Nontidal wetland mapping in South Carolina using airborne multispectral scanner data. Remote Sensing of Environment **16**:1-12.

Jensen, J. R. 2000. Remote Sensing of the Environment: An Earth Resource Perspective. 2nd edition. Prentice Hall, Upper Saddle River, NJ.

Kelly, M., D. Shaari, Q. Guo and D. Liu. 2004. A comparison of standard and hybrid classifier methods for mapping hardwood mortality in areas affected by sudden oak death. Photogrammetric Engineering and Remote Sensing **70**:1229-1239.

Klemas, V. V. 2001. Remote sensing of landscape-level coastal environmental indicators. Environmental Management **27**:47-57.

Lathrop, R. G., P. Montesano and S. Haag. 2006. A multi-scale segmentation approach to mapping seagrass habitats using airborne digital camera imagery. Photogrammetric Engineering & Remote Sensing **72**:665-675.

Li, J. and W. Chen. 2005. A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. International Journal of Remote Sensing **26**:5051-5069.

Lunetta, R. S. and M. E. Balogh. 1999. Application of multi-temporal Landsat 5 TM imagery for wetland identification. Photogrammetric Engineering & Remote Sensing **65**:1303-1310.

Munyati, C. 2000. Wetland change detection on the Kafue Flats, Zambia, by classification of a multitemporal remote sensing image dataset. International Journal of Remote Sensing **21**:1787-1806.

Munyati, C. 2004. Use of principal component analysis (PCA) of remote sensing images in wetland change detection on the Kafue Flats, Zambia. Geocarto International **19**:11-22.

Philip Williams & Associates Ltd. and P. M. Faber 2004. Design guidelines for tidal wetland restoration in San Francisco Bay.Oakland, CA,The Bay Institute and California State Coastal Conservancy.83.

Phinn, S. R., C. Menges, G. J. E. Hill and M. Stanford. 2000. Optimizing remotely sensed solutions for monitoring, modeling, and managing coastal environments. Remote Sensing of Environment **73**:117-132.

Ramsey, E. W., III and S. Laine. 1997. Comparison of Landsat Thematic Mapper and high resolution photography to identify change in complex coastal wetlands. Journal of Coastal Research **13**:281-292.

Ramsey, E. W., III, G. A. Nelson and S. K. Sapkota. 1998. Classifying coastal resources by integrating optical and radar imagery and color infrared photography. Mangroves and Salt Marshes **2**:109-119.

Rosso, P. H., S. L. Ustin and A. Hastings. 2005. Mapping marshland vegetation of San Francisco Bay, California, using hyperspectral data. International Journal of Remote Sensing **26**:5169-5191.

Scarpace, F. L., B. K. Quirk, R. W. Kiefer and S. L. Wynn. 1981. Wetland mapping from digitized aerial photography. Photogrammetric Engineering & Remote Sensing **47**:829-838.

Schiewe, J., L. Tufte and M. Ehlers. 2001. Potential and problems of multi-scale segmentation methods in remote sensing. GeoBIT/GIS **6**:34-39.

Shuman, C. S. and R. F. Ambrose. 2003. A comparison of remote sensing and ground-based methods for monitoring wetland restoration success. Restoration Ecology **11**:325-333.

Stankiewicz, K., K. Dąbrowska-Zielińska, M. Gruszczyńska and A. Hościło 2003. Mapping vegetation of a wetland ecosystem by fuzzy classification of optical and microwave satellite images supported by various ancillary data. SPIE, Remote Sensing for Agriculture, Ecosystems, and Hydrology IV.

Thomson, A. G., R. M. Fuller and J. A. Eastwood. 1998. Supervised versus unsupervised methods for classification of coasts and river corridors from airborne remote sensing. International Journal of Remote Sensing **19**:3423-3431.

Thomson, A. G., R. M. Fuller, M. G. Yates, S. L. Brown, R. Cox and R. A. Wadsworth. 2003. The use of airborne remote sensing for extensive mapping of intertidal sediments and saltmarshes in eastern England. International Journal of Remote Sensing **24**:2717-2737.

Thomson, A. G., A. Huiskes, R. Cox, R. A. Wadsworth and L. A. Boorman. 2004. Short-term vegetation succession and erosion identified by airborne remote sensing of Westerschelde salt marshes, The Netherlends. International Journal of Remote Sensing **25**:4151-4176.

Trimble Inc., 2005. URL http://www.trimble.com/ [accessed on 17 January 2007]

Underwood, E., S. Ustin and D. DiPietro. 2003. Mapping nonnative plants using hypersepctral imagery. Remote Sensing of Environment **86**:150-161.

Wang, L., W. P. Sousa, P. Gong and G. S. Biging. 2004a. Comparison of IKONOS and QuickBird images for mapping mangrove species on the Caribbean coast of Panama. Remote Sensing of Environment **91**:432-440.

Wang, L., W. P. Sousa and P. Gong. 2004b. Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. International Journal of Remote Sensing **25**:5655-5668.

Williams, P. and P. Faber. 2001. Salt marsh restoration experience in San Francisco Bay. Journal of Coastal Research **27**:203-211.

Wulder, M. A., R. J. Hall, N. C. Coops and S. E. Franklin. 2004. High spatial resolution remotely sensed data for ecosystem characterization. BioScience **54**:511-521.

Yang, X. 2005. Remote sensing and GIS applications for estuarine ecosystem analysis: an overview. International Journal of Remote Sensing **26**:5347-5356.

Yoon, G.-W., S. I. Cho, S. Jeong and J.-H. Park 2003. Object oriented classification using Landsat images. American Society for Photogrammetry and Remote Sensing (ASPRS), Anchorage, Alaska. May 5-9, 2003.

Yu, Q., P. Gong, N. Clinton, M. Kelly and D. Schirokauer. 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. Photogrammetric Engineering and Remote Sensing **72**:799-811.

Zhang, M., S. L. Ustin, E. Rejmankova and E. W. Sanderson. 1997. Monitoring Pacific coast salt marshes using remote sensing. Ecological Applications **7**:1039-1053.

Chapter 4.5

A Local Fourier Transform approach for vine plot extraction from aerial images

C. Delenne[1], S. Durrieu[1], G. Rabatel[2], M. Deshayes[1]

[1] UMR TETIS, Maison de la Télédétection, Montpellier

[2] UMR ITAP, Cemagref, Montpellier

ABSTRACT: An automatic tool for vineyard mapping and monitoring is expected by numbers of managers and winegrower cooperatives in winegrowing regions. A frequency analysis using the Fourier Transform on aerial images has been developed to meet this need. This results in vine plot detection, delineation and characterization with very precise estimation of inter-row width and row orientation. To foster large-scale applications, tests and validation have been carried out on standard very high spatial resolution remotely sensed data. On the study area, only 8% of vine plots have not been detected (corresponding to small plots with only 1.5% of the total vineyard surface) and more than 61% have been correctly detected (other mainly concerns partial detection).

1 Introduction

The considerable increase in digital technologies enables to automatically analyze images but also to understand them by providing high-level information on their content. Concurrently, such a considerable increase is observed in the availability of very high spatial resolution (VHSR) remotely-sensed data. As a consequence, a lot of new potential applications are now

made possible since the shape or the spatial structure of observed objects is becoming more distinguishable, providing greater possibilities for discrimination and characterization. In the agricultural domain, various types of vegetation can thus be distinguished according to their spatial patterns (cereal crops, forests, orchards…). In this context, automatic analysis methods could be developed to build or update geographical databases for land management. However, because they deal with spatial structures or shape, these new applications also require new image processing methods. Several shape-model based approaches can thus be found in the literature, especially for building detection (Garcin et al. 2001; Segl et al. 2003) or isolated trees detection (Barbezat et al. 1996). For forest identification, various textural approaches based on co-occurrence matrices are proposed such as (Franklin et al. 2000) or (Moskal 2002).

We address the issue of vineyard detection and characterization from VHSR images for inventory and management purposes. Indeed accurate and up to date digital mapping of vineyards could be used by vinegrower cooperatives, e.g. to improve the monitoring of quality compliance in areas registered in the list of Controlled Origin Denomination. The management of pollution, erosion and flood risks is another field that can take advantage of these maps. Indeed, these risks, depending on culture and soil surface condition, are worsened by mechanization and intensive cropping practices.

Most vineyard related studies based on remotely-sensed data aim to characterize previously delimited plots (training mode (Wassenaar et al. 2002), foliar density (Hall et al. 2003)…). They emphasize the relevance of textural analysis applied to submetric spatial resolution images. Indeed a very important feature concerning vineyard is the spatial periodicity of the pattern resulting from the spatial arrangement of plants (often in lines or grid), which can be perceptible with a spatial resolution that is at least twice as small as the pattern period. Because of this periodic organization, a vine pattern can roughly be assimilated to a local planar wave of a given spatial frequency and orientation. Therefore, frequency analysis appears as a suitable approach for vine detection. The wavelet analysis presented in (Ranchin et al. 2001) is applied to 25 cm resolution images for vine/non-vine pixel classification. Using a plot based validation, 78 % of plots were accurately classified; but this approach is complex and needs significant user intervention. Wassenaar et al. (2002) successfully used a Fourier Transform analysis for vine/non-vine classification and for characterization of already delimited plots on 25 cm resolution images. This method also gave a very precise estimation of interrow width and row orientation. Delenne et al. (accepted for publication) showed the superiority of a frequency analysis on a textural analysis using co-occurrence matrices, concerning vine-

yard detection (pixel classification) in VHSR images, without any parcel plan availability.

We address here the OBIA problem of vineyard segmentation, using this frequency analysis approach. To foster large-scale applications, the process has been applied to standard VHSR aerial images in natural colors and with a 50cm spatial resolution. In the following part, the theoretical aspects of the method are presented. Then results obtained on a 200 ha study area, presenting a large set of vineyard types and conditions, are given and discussed.

2 Method

2.1 Fourier Transform for vineyard detection

Fourier theory states that almost any signal, including images, can be expressed as a sum of sinusoidal waves oscillating at different frequencies. The Fourier Transform amplitude (or Fourier spectrum) of an image I can be represented in the frequency domain as another image, FI. In the conventional representation, this image is symmetric with respect to its center, which contains the average of I, i.e. the amplitude of the null frequency $F0$. Each pixel position corresponds to a particular spatial frequency f ranging from $f = 0$ at the center to $f = \pm 0.5$. Its value codes the amplitude of Fourier spectrum, which depends on the presence of the corresponding frequency in the original image I.

In most of vine-growing regions, two main patterns can be observed on aerial images according to the vine training mode used:

- Grid pattern: about a quarter of the vineyards considered in this study is trained in 'goblet'. This old method of vine training involves no wires or other system of support: vine stocks are planted according to a grid pattern (often on a square basis) with approximately 1.5 m x 1.5 m spacing in the study area but sometimes up to 3 m spacing in dry regions.
- Parallel lines pattern: most of the recent vineyards are trained using horizontal wires to which the fruiting shoots are tied. The distance between two wires is larger than between vine stocks guided by the same wire (often 1 m x 2.5 m spacing in the study area). More adapted to mechanization, this training mode named trellis, is mainly used.

Since vineyard patterns on aerial images are periodical and oriented, they induce very located peaks of amplitude in Fourier spectrum (Fig. 1).

(a)   (b)
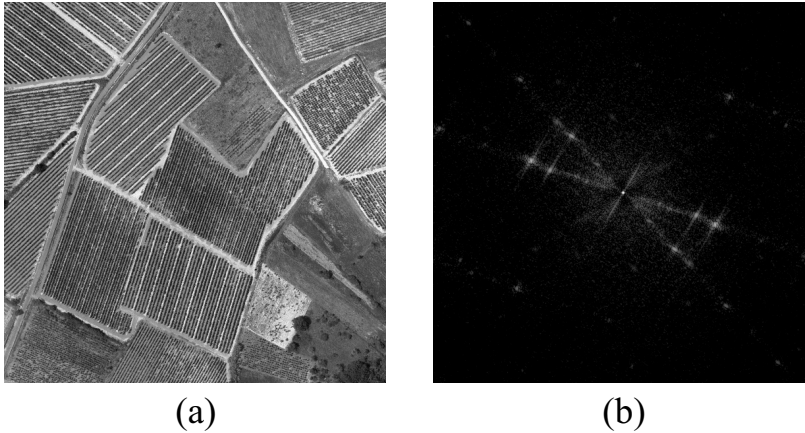
Fig. 1. Fourier spectrum of a vineyard area: **(a)** original image; **(b)** Fourier spectrum on which several peaks can be seen (highest intensity values)

Three characteristics can be deduced from the value and positions of these peaks:

1. Peak value can be seen as an estimation of the vine presence in the original image (the higher the peak, the more contrasted the pattern).
2. The angle formed by *(center, peak)* vector with the horizontal line, determines the wave direction in a polar coordinate system, which is perpendicular to the pattern direction i.e. the vine row orientation.
3. The distance r between one peak and the spectrum centre, is the frequency f of the corresponding wave ($f \in [0, 0.5]$). This value is directly linked to the pattern period T in pixel *i.e.* the vine interrow width, by $f = 1/T$.

To ensure the presence of a unique vine pattern in the analyzed part of the original image, FFT is computed on a small sliding window (included in a plot), which size will be discussed in the next section. When this window contains vineyard, two symmetric peaks are present on the Fourier spectrum image and two other peaks at 90° for the grid pattern of a goblet vine. On the contrary, only the central peak is present for the non-periodic patterns of other land covers (Fig. 2).
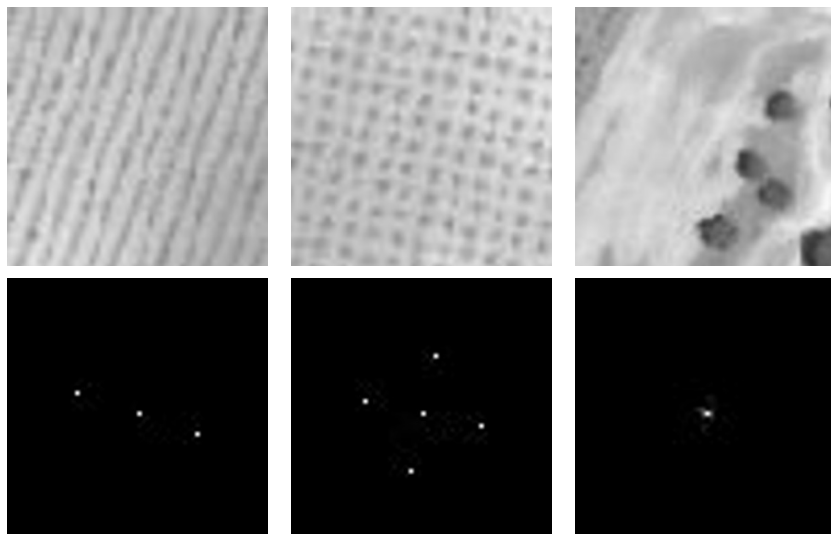
Fig. 2. Examples of Fourier spectrum (down) for 3 sub-images of size 61x61 pixels (up) extracted from a trellis vine (left), a goblet vine (middle), and a non-vine (right)

A normalization step can be applied to increase the visibility of these peaks. We use here a linear transformation on a sliding window, which normalizes the standard deviation mean and ensures that any input image will have approximately the same amplitude of luminance variation from row to interrow in a vine plot.

We then use the following method for vine plot detection:

1. The Fourier Transform is computed on the sliding window.
2. The Fourier spectrum maximum is searched for in an annular ring corresponding to potential vineyard frequencies (linked to interrow width).
3. The interrow width and row orientation are deduced from the maximum position in the Fourier spectrum.

Then, three characteristics are affected to the sliding window central pixel: the maximum amplitude, the interrow width and the row orientation. At the end of this step, 3 pseudo-channels are available, on which a segmentation will be performed.

2.2 Segmentation procedure

Row orientation and interrow width are supposed to be constant within a same vine plot. The homogeneity h of these characteristics can thus be used with the amplitude of Fourier spectrum to compute a 'vine index', V. Many combinations - linear or not – of these data have been tested. The aim of getting a discriminating vine index, in which most of non-vine pixels are null, has been reached with a multiplication of the three pieces of data. For each pixel p, $V(p)$ is then defined as:

$$V(p) = \left(a(p) \cdot h_\theta(p) \cdot h_T(p)\right)^2 \tag{1}$$

where a is the maximum of Fourier Transform amplitude in the sliding window centered on p; θ and T are respectively the corresponding orientation and period and the homogeneity h is defined by:

$$h_F(p) = 1 - \mathrm{var}(F(p)) \qquad F = \theta, T \tag{2}$$

where var(F) is the normalized variance of F, computed on a sliding window of size 3x3 centered on p.

This vine index enables to discriminate vineyard textures from other ones that have generally null values. We then use a simple "contour" function, which create a vector object for each clump.

Although this function segments vineyards area, many plots that are spatially close can be grouped into a unique object. To overcome this issue, the segmentation procedure is reiterated on the vine rows orientation and interrow widths "channels" on area identified as vineyard by the first segmentation. This corresponds to a colored image segmentation, which enables the delineation of each vine plots in a unique object, except for those that are spatially close and have the same characteristics, which can remain grouped.

The general scheme of the method is represented in figure 3. At the end of the process the Fourier transform is computed again on each segmented plot in order to precisely characterize vine row orientation and interrow width.
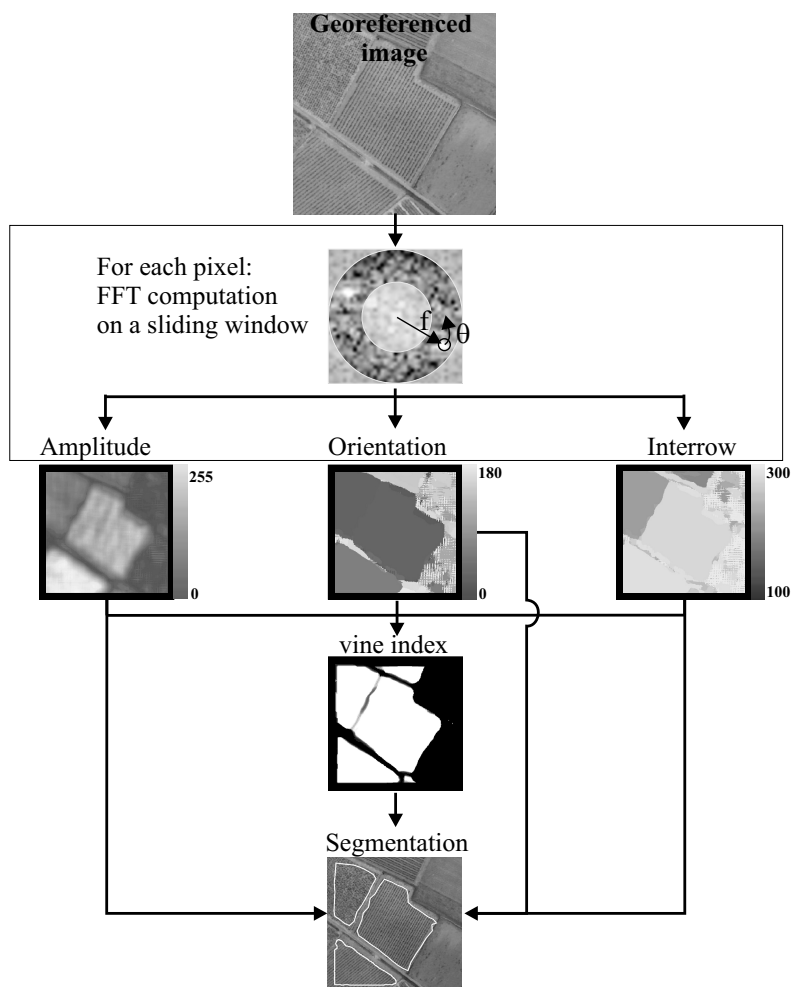
Fig. 3. Simplified representation of the whole process

2.3 Validation method

The main goal of the presented process is the plot contours delineation. Therefore, a plot based validation has been performed, comparing automatic and manual segmentation (considered as reference) according to their overlapping rate. Eight different cases have been defined:

1. Good segmentation: the overlapping surface of reference and automatically segmented plots is higher than 70%.

2. Over-segmentation: several plots are automatically segmented in one reference plot.
3. Under-segmentation: one automatically segmented plot includes several reference plots.
4. Partial segmentation: only one part of the plot is detected.
5. Large segmentation: the automatically segmented plot over-flows onto other plots.
6. Missing segmentation: undetected reference plot.
7. Extra segmentation: non-vine automatically segmented as a vine plot.
8. Other cases.

3 Results

3.1 Study area and data

To assess the global detection process, a study area of 200 ha has been chosen, which is a subset of the La Peyne watershed (110km²) located in the Languedoc-Roussillon region, France (Fig. 4).
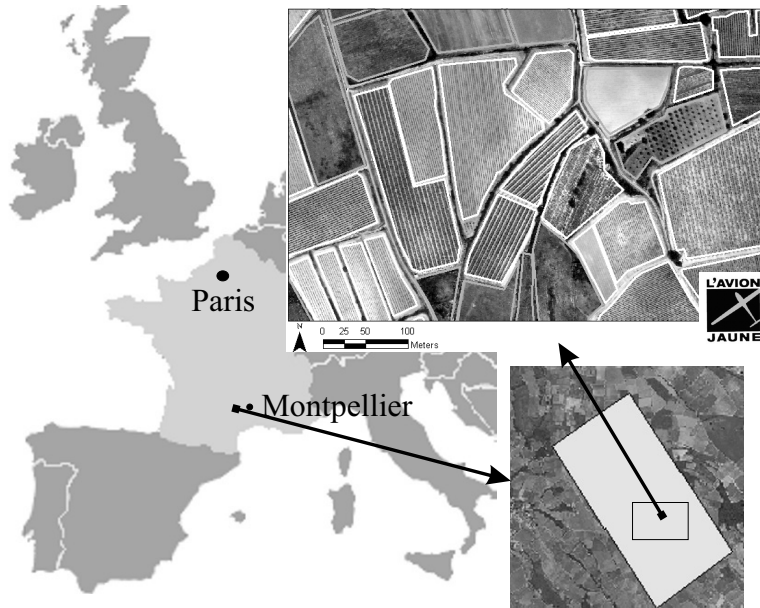
Fig. 4. Location and subset of the study area

Data acquisition was carried out during the first week of July 2005, when foliar development was such that both vine and soil were visible on

aerial images, providing contrasted periodical patterns. Photographs in natural colors were acquired using a digital camera aboard an Ultra Light Motorized (ULM) by the company l'Avion Jaune®. They were geometrically corrected, mosaicked and resampled for a 50 cm resolution.

Since the original image required for FFT computation must be in gray levels, the three channels have been tested separately. Best results, presented thereafter, have been obtained with the red one, since it provides higher contrast between vine and soil (even covered by grass).

For validation purpose, ground-truth information was collected for the whole study area the same day as image acquisition. For example: land use; crop pattern (e.g. grid or line); soil surface condition; estimation of vineyard age or mortality rate. Interrow width and orientation were obtained by precise on-screen measurements.

3.2 Sensitivity analysis to the window's size

The main parameter of the method is the sliding window size. A sensitivity of the results to this size has been carried out since accuracy of detection and characterization depends on the number of pixels in the window. On the one hand, this window must be large enough to take into account the repetition of row or grid patterns, so a large window provides more precise information when located inside a plot. On the other hand, besides increasing calculation times, a larger window decreases classification results near plots boundaries as it can contain several patterns at the same time. Eight window sizes have been tested from 5×5 m to 40×40 m. Results become acceptable for a 20×20 m window size, for which the good segmentation (case 1) rate is the higher. However, the lowest rate of non-detection (case 6) is reached for 30×30 m. Extending window size up to 40×40 m does not improve results while considerably increasing computational time. We do prefer a better result in term of non-detection than good segmentation, because a further step can be developed to improve plot contours. Consequently, the best trade-off for the window size is about 30×30 m, which can contain from 12 to 15 vine rows in the study area.

3.3 Segmentation results

Quantitative results obtained with a 30×30 m sliding window are given in table 1 according to the previous classification and are illustrated in figure 5.

Table 1. Segmentation results with comparison to reference plots

Segmentation type	Number of reference plots (except for 7)	
1. Good	54	(47.5%)
2. Over	3	(2.5%)
3. Under	16	(14%)
4. Partial	12	(10.5%)
5. Larger	6	(5%)
6. Missing	9	(8%)
7. Extra	5	
8. Other	14	(12.5%)
Total (except 7)	114	(100%)
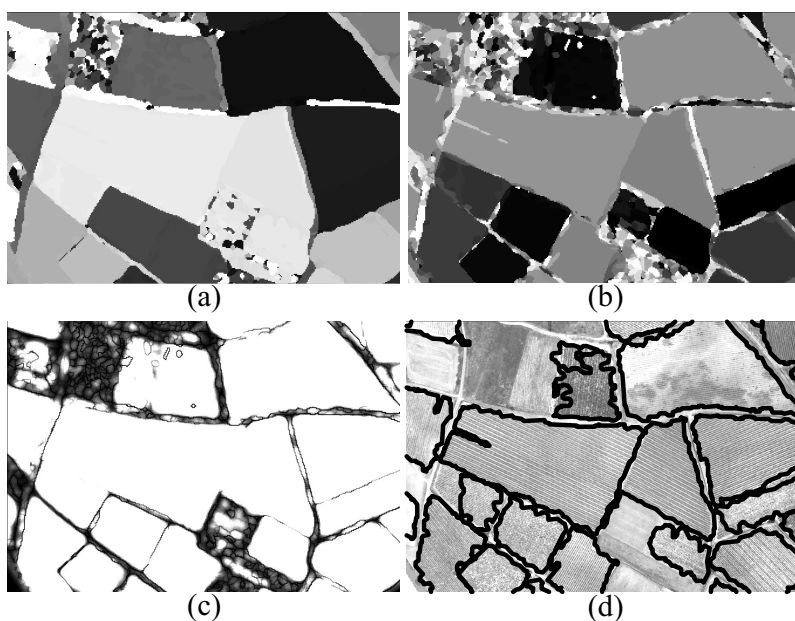
Fig. 5. Zoom on the results: **(a)** pseudo-channel representing vine-row orientation; **(b)** pseudo-channel representing interrow width; **(c)** pseudo-channel representing vine index; **(d)** segmentation results drawn above image red channel

The most represented types of results are: good segmentation (47.5%), under segmentation (14%), partial segmentation (10.5%) and other (12.5%).

The missing plot ratio is relatively weak (8%) and mainly concerns small plots. Indeed, eight out of the nine non-detected vines are smaller

than 0.2ha and their surfaces correspond to only 1.15% of the total vine-yard surface.

Under-segmentation typically correspond to the grouping of neighboring plots that have the same row orientation and inter-row width and are only separated by a narrow road or a ditch. Some of them are not spatially separated and only differ by the soil surface condition between rows or by some characteristics undistinguishable in aerial images such as age or height (Fig. 6a). This kind of segmentation error can nevertheless be considered as good detection and is a side effect of a non-sufficient accuracy in plot edge definition. This is inherent to any segmentation method relying on spatial pattern detection: a minimal neighborhood is required to detect patterns, leading to a limited spatial resolution of the segmentation.

A significant amount of plots are only partially segmented, mainly because of internal heterogeneities (*e.g.* Fig. 6b).

The "other" case often corresponds to plots that are at the same time partially and over or under segmented (Fig. 6c).
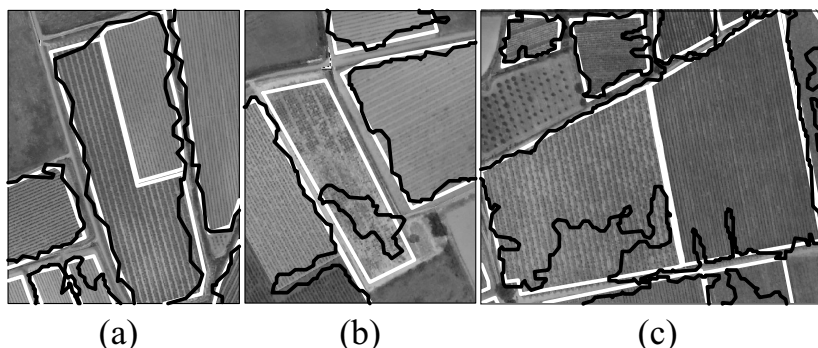
Fig. 6. Examples of results: **(a)** partially segmented plot, **(b)** under-segmented plots, **(c)** under and partially segmented plots (manual segmentation in white, automatic one in black)

Finally, four 'plots' have been wrongly segmented. One of them is due to a high longitudinal gray-level transition, which has generated an amplitude peaks in the search range. The others are located in non-cultivated plots, but which have been recently ploughed and therefore present a parallel structure. This last kind of 'plot' could probably be eliminated e.g. using a complementary radiometric analysis.

3.5 Characterization results

Thanks to the amplitude peak position in the Fourier spectrum, vine row orientation and interrow width can be precisely characterized.

Concerning orientation, on-screen measurement and automatic characterization are nearly identical (Fig. 7), with an absolute average difference δ lower than 1°.

A bigger (but still small) gap between on-screen measurement and automatic characterization can be observed concerning interrow width (Fig. 8), with an absolute average difference δ = 3.3cm.
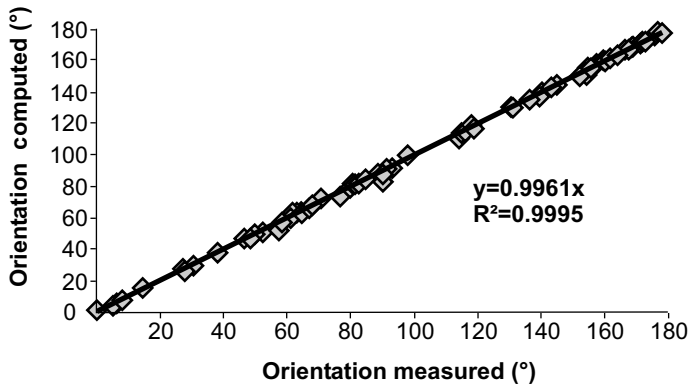
Fig. 7. Characterization of vine row orientation
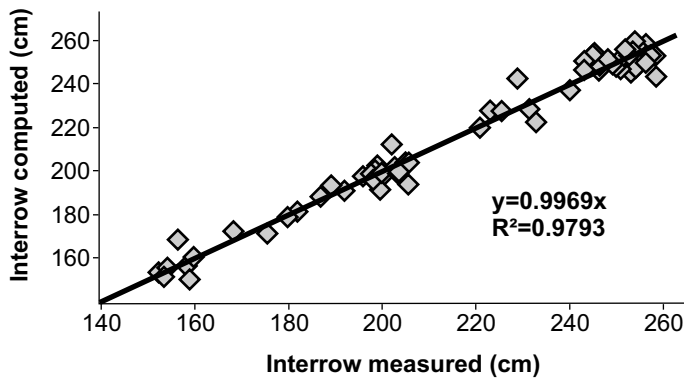
Fig. 8. Characterization of interrow width

Allowing to Wassenaar et al. (2002), a twice-higher resolution (0.25cm) enables a twice-better interrow width characterization (δ = 1.6cm), but does not improve the precision of orientation estimations.

4 Conclusion and discussion

The proposed approach has proved its efficiency for vineyard detection and characterization in many ways. While most of detection studies provide a pixel classification, the main originality of this method is that it performs both detection and delineation in polygons. Another significant advantage is that good results are obtained with the red channel, present in widely available images in natural colors. And since the appropriate spatial resolution is linked to the local pattern period, a coarser one (e.g. issued from Ikonos or Quickbird satellites) could be used in many vine-growing regions, especially dry ones such as in Spain where inter-row widths are up to 3 m (providing that the periodic pattern is still contrasted).

Moreover, this kind of approach, based on the periodical crop pattern, enables a limited risk of commission error compared to spectral analysis. Indeed, only crops with the same range of interrow width could lead to confusion. Then, orchards, for example, will not be detected as vine.

In terms of performance, the present study has nevertheless shown some limitations concerning plot contours shape and localization accuracy. To overcome this issue, we do believe that a further step based on individual vine row analysis and adjustment would probably be sufficient (presently under development). Moreover, this could for example decrease the rate of partially delimited plots and correct some erroneous detection due to high longitudinal transitions in the image.

The cases of missing detection are more problematic as no further improvement step can be envisaged. However, as seen above, they mainly concern small plots, with very few rows, and this kind of plot tends to disappear with mechanization in most vineyard regions.

Acknowledgments

References

Barbezat V, Kreiss P, Sulzmann A, Jacot J (1996) Automated recognition of forest patterns using aerial photographs. SPIE: Optics in Agriculture, Forestry, and Biological Processing II, Bellingham (Washington).

Delenne C, Durrieu S, Rabatel G, Deshayes M, Bailly J-S, Lelong C, Couteron P, Textural approaches for vineyard detection and characterization using very high spatial remote-sensing data, International Journal of Remote-Sensing (accepted for publication).

Franklin S E, Hall R J, Moskal L M, Maudie A J, Lavigne M B (2000) Incorporating texture into classification of forest species composition from airborne multispectral images. International Journal of Remote Sensing 21, pp. 61–79.

Garcin L, Descombes X, Zerubia J, Le Men H (2001) Building detection by markov object processes and a MCMC algorithm. Technical Report 4206, INRIA.

Hall A, Louis J, Lamb D (2003) Characterising and mapping vineyard canopy using high-spatial-resolution aerial multispectral images. Computers and Geosciences 29, pp. 813–822.

Moskal L (2002). Investigating texture inversion in high resolution multispectral imagery; implications for forest classification. In: Annual conference and FIG XXII Congress, Washington, D. C.

Ranchin T, Naert B, Albuisson M, Boyer G, Astrand P (2001) An automatic method for vine detection in airborne imagery using wavelet transform and multiresolution analysis. Photogrammetric Engineering and Remote Sensing 67(1), pp. 91–98.

Segl K, Roessner S, Heiden U, Kaufmann H (2003) Fusion of spectral and shape features for identification of urban surface cover types using reflective and thermal hyperspectral data. ISPRS Journal of Photogrammetry and Remote Sensing 58, pp. 99–112.

Wassenaar T, Robbez-Masson J-M, Andrieux P, Baret F (2002) Vineyard identification and description of spatial crop structure by per-field frequency analysis. International Journal of Remote Sensing 23(17), pp. 3311–3325.

Section 5

Automated classification, mapping and updating: land use / land cover

Chapter 5.1

Object-based classification of IKONOS data for vegetation mapping in Central Japan

N. Kamagata[1], K. Hara[1], M. Mori[2], Y. Akamatsu[2], Y. Li[3] and Y. Hoshino[4]

[1] Graduate School of Informatics, Tokyo University of Information Sciences, Japan, h05002nk@edu.tuis.ac.jp

[2] Kokusai Kogyo Co. Ltd., Japan

[3] Japan Space Imaging Corporation, Japan

[4] Faculty of Agriculture, Tokyo University of Agriculture and Technology, Japan

KEYWORDS: Vegetation mapping, Object-based classification, Texture, IKONOS, Feature selection

ABSTRACT: Vegetation mapping using IKONOS data was implemented at a countryside study area in central Japan, where small patches of various plant communities are mixed together in a complicated mosaic pattern. Pixel-based and object-based classifications using only spectral features were implemented and their accuracies were compared. In addition, the object-based classification was also performed on a combination of spectral and textural features, with a stepwise regression model used in the discriminate analysis to select the most relevant features. Classifications were implemented at four levels, the highest of which used seven vegetation categories. The object-based classification proved more accurate than

the pixel-based classification. In addition, the addition of textural features generated significant improvements in accuracy. The overall classification accuracy and Kappa coefficients at the highest level were 52.8% and 0.373 for the pixel-based classification; 58.9% and 0.458 for the object-based with spectral features only; and 65.0% and 0.542 for the object-based with additional features. Some problems with misclassification remained, but the overall results demonstrate that object-based classification of very high resolution satellite images using additional features is a practical tool for vegetation mapping in Japan.

1 Introduction

Plant communities all over Japan are changing rapidly due to shifts in land management policies, and are also becoming fragmented by urbanization and development. Continuously updated vegetation maps, showing the current distribution of the various communities, are required for proper management and conservation. Representative vegetation maps are those produced by the Ministry of the Environment (formerly the Environment Agency) since 1973. These maps, however, rely on detailed analysis of aerial photographs coupled with substantial field research, and as such require great amounts of time and labor. In addition, the accuracy of the maps varies with the skill and experience of the researchers. The need for a more efficient, less labor-intensive methodology has become even more urgent since the Ministry switched from 1/50,000 scale maps to 1/25,000 scale for the 6th National Survey on the Natural Environment (1999~2004).

New technologies, employing Geographic Information Systems (GIS) and remote sensing technology, are being examined as tools for efficient vegetation mapping (e.g., Alexander and Millington 2000). The Japanese Ministry of the Environment is also considering the use of remote sensing in future vegetation mapping projects (Biodiversity Center of Japan 2005). To date, however, most remote sensing attempts have employed Landsat or SPOT data. In many regions of Japan, small vegetation patches are mixed together in a complicated mosaic pattern, and these medium resolution data are insufficient for vegetation mapping at 1/25,000 or finer scale.

Recently, however, very high resolution satellite data, such as IKONOS, have become widely available, and have increased expectations for enhanced discrimination of cover types (Mumby and Edwards 2002; Wang et al. 2004). In Japan, remote sensing using very high resolution IKONOS

data shows promise as a tool for classifying vegetation elements even in areas with exceptionally complicated patterns of vegetation.

Technical problems, however, such as misinterpretation of shadows and variable-shaped individual crown trees, still arise in pixel-based classification using very high resolution remote sensing data (Hay et al. 1996; Laliberte et al. 2007). Recent research (e.g., Wang et al. 2004; Marçal et al. 2005) has thus focused on object-based classification programs as a means of overcoming these problems.

Very high spatial resolution imagery provides 1 to 4m resolution multispectral data, which is sufficient to estimate forest stand characteristics (Tuominen and Pekkarinen 2005). Herold et al. (2003), however, have noted that very high resolution satellite data are limited to four multispectral bands, and thus may have limitations with regard to detailed mapping. In addition to spectral features, however, IKONOS also provides 1m resolution panchromatic images, which enable classification using texture patterns. Carleer and Wolff (2006), Puissant et al. (2005) and Laliberte et al. (2007), have reported that a gain in classification accuracy can be obtained by using a combination of both spectral and textural features.

In this research, the practicality of vegetation mapping by object-based classification using IKONOS very high resolution remote sensing data was examined and compared with two types of conventional pixel-based classification methods. In addition, the object-based classification was repeated using a combination of spectral and textural features. These results were then compared to those obtained by using spectral features alone.

2 Object-based classification in vegetation mapping

The conventional vegetation mapping process is outlined in Figure 1. To begin with, aerial photographs are interpreted, and boundaries of the major vegetation classes are mapped out, producing a draft physiognomical vegetation map. Next, on site field research is implemented to correct and verify the boundaries drawn on this draft map. Data on vertical and horizontal distribution density of individual species is also collected at this time.

```
┌ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ─ ┐
│  ┌─────────────────────────────┐ │
│  │ Draft of Physiognomical Vegetation Map │ │
│  │    from Aerial Photographs    │ │
│  └─────────────────────────────┘ │
└ ─ ─ ─ ─ ─ ─ ─ ↓ ─ ─ ─ ─ ─ ─ ─ ─ ┘
   ┌─────────────────────────────┐
   │     Initial Phase Field Work     │
   └─────────────────────────────┘
                  ↓
   ┌─────────────────────────────┐
   │   Completion of Physiognomical   │
   │         Vegetation Map          │
   └─────────────────────────────┘
                  ↓
   ┌─────────────────────────────┐
   │        Compilation of Data        │
   │   Extraction of Plant Communities  │
   └─────────────────────────────┘
                  ↓
   ┌─────────────────────────────┐
   │     Second Phase Field Work     │
   └─────────────────────────────┘
                  ↓
   ┌─────────────────────────────┐
   │ Completion of Phytosociological Map │
   └─────────────────────────────┘
```
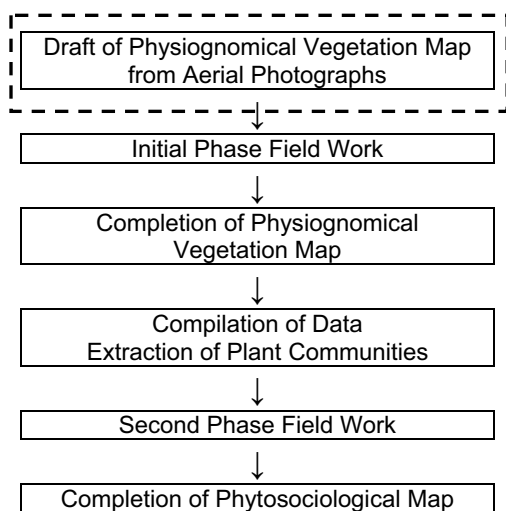
Fig. 1. Steps in production of phytosociological vegetation map (Suzuki et al. 1985). Remote Sensing can contribute to the first step, encased in a dotted line.

The field data is then compiled and analyzed to identify precise plant communities. Finally, these results, combined with additional fieldwork, are used to produce a phytosociological vegetation map based on associations of particular species (Braun-Blanquet 1932; Suzuki et al. 1985).

To our knowledge, to date there have been few or no attempts to apply remote sensing directly to vegetation mapping. Remote sensing, which is limited in its ability to identify understory vegetation, can not be expected to generate or update the completed phytosociological vegetation map. Classification of remotely-sensed data, however, should have a potential for streamlining and standardizing the initial phase of vegetation mapping, which involves drafting the physiognomical map.

To illustrate how remote sensing can contribute to vegetation mapping, Figure 2 shows a schematic representation of the differences among manual interpretation using aerial photographs (2a); pixel-based classification of remotely sensed data (2b) and object-based classification of remotely sensed data (2c). As shown in the figure, the pixel-based classification approach, which does not consider the relationship of each pixel data with its adjacent units, tends to divide the area up into finer divisions than is called for. Small gaps or shadows in a relatively homogeneous forest, for example, show up as different vegetation categories. In the object-based approach, on the other hand, the program can be set to filter out these minor inconsistencies, producing results very similar to those generated by the manual interpretation. Thus object-based classification should have potential for streamlining and standardizing the initial phase work of vegetation mapping.
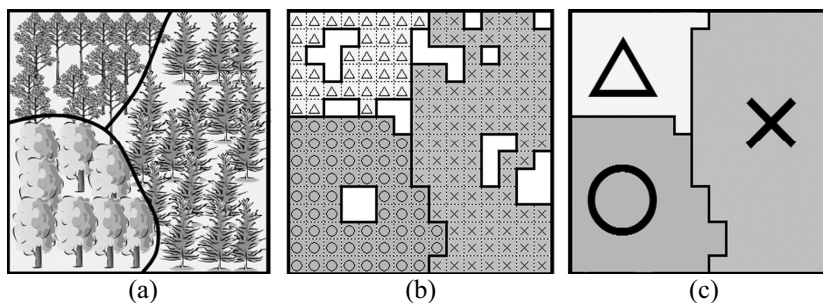
Fig. 2. Classification of vegetation types (Modified from Kamagata et al. 2006)
(a): based on visual interpretation
(b): pixel-based (over classification)
(c): object-based (results similar to visual interpretation)

3 Methods

3.1 Study Area and Data

This research was implemented on a 2 km by 2 km test area, located in an agricultural area of Sosa City, located in eastern Chiba Prefecture, central Honshu (Figure 3). The study area consists of narrow, highly-branched alluvial valleys, called 'Yatsu', which are cut deeply into flat-topped, plateau-like uplands. The height of the uplands is only 30-40 meters, limiting the effect of terrain on the analysis. Forest communities in the study area consist of deciduous broadleaved woodlands (*Quercus*, *Carpinus*); evergreen broad-leaved woodlands (*Castanopsis*, *Quercus*); bamboo groves (*Phyllostachys*) and conifer plantations (*Cryptomeria*, *Chamaecyparis*). These vegetation types border each other in a complicated patchwork pattern. In addition, some deciduous broadleaved woodlands have been invaded by bamboo or dwarf bamboo (*Pleioblastus*).

The research utilized multi-spectral, 4m resolution and panchromatic, 1m resolution remotely sensed data from IKONOS (Japan Space Imaging), obtained on 1 April 2001 with 0% cloud coverage (Figure 4). The five spectral bands are 0.45-0.90µm (Panchromatic), 0.45-0.52µm (Blue), 0.52-0.60µm (Green), 0.63-0.69µm (Red) 0.76-0.90µm (NIR). Also employed was the Normalized Difference Vegetation Index (NDVI), calculated as follows (Eq.3.1).

$$NDVI = (\, NIR - Red \,) \, / \, (\, NIR + Red \,) \qquad (3.1)$$
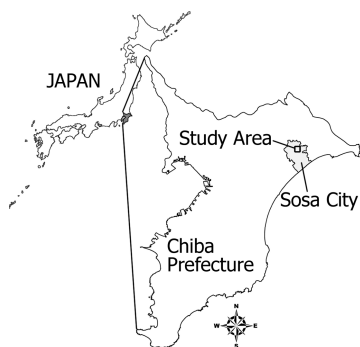
Fig. 3. Location of study area
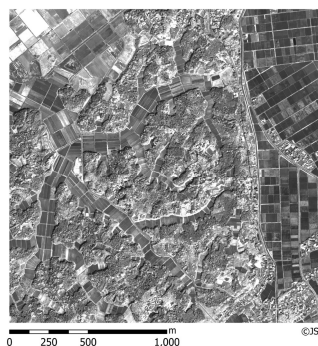
Fig. 4. IKONOS image of study area (2000m x 2000m)

The classification results were compared against a master vegetation map, which was prepared based on standards used by the Ministry of the Environment in their physiognomical vegetation maps. Only vegetated areas were considered. The classification scheme employed in this research follows that of the Ministry of Environment's Actual Vegetation Maps. The vegetation community categories were: Evergreen broad-leaved forest; Deciduous broad-leaved forest; Secondary Grassland; Wetland vegetation; Coniferous plantation; Other plantation; Bamboo grove.

3.2 Classification Categories and Classification Levels

Based on the composition of forests in the study area, classification was implemented at four levels (Table 1). At the first, or roughest level, the differentiation was only between forest and grassland. At level 2, grasslands were divided into secondary and wetland types, and forests into evergreen and deciduous. At Level 3, the evergreen forests were divided into wood and bamboo; and at the finest level, evergreen forest (wood) were subdivided into evergreen broad-leaved, coniferous plantation, and other plantation. A total of 7 vegetation classes were thus obtained at Level 4. The ultimate goal of the research was Level 4, but classifications were implemented at all levels to provide examples and to determine the practical limits of the classification system being tested.

Table 1. Vegetation classes used at four levels of classification

Level 1	Level 2	Level 3	Level 4
Forest c1	Evergreen c3 forest	Evergreen c7 forest (wood)	Evergreen c9 broad-leaved forest
			Coniferous c10 plantation
			Other planta- c11 tion
		Bamboo c8 grove	
	Deciduous c4 broad-leaved forest		
Grassland c2	Secondary c5 grassland		
	Wetland ve- c6 getation		

3.3 Image Segmentation and Accuracy Assessment

The object-based classification employed eCognition software (Definiens). Initial segmentation was a multi resolution, bottom-up system based on the method of Baatz and Schape (2000). In object-based classification, object size, shape and other parameters can be adjusted to fit the needs of the research. In this case, the same segmentation parameters were used for classifications at all levels, with the scale parameter set at the level of the plant communities as noted above. Texture and color of the image data were used to classify each unit, and integration of areas was accomplished by increasing the scale parameters. The study area was divided by segmentation processing.

For evaluating the accuracy of the classifications, the master map was converted from vector data to 4m resolution raster data, and stratified random sampling was employed to choose sampling points from each category. A minimum of 50 points were selected for each category, with a total of 1000 points for each image. Using the master map as a base, producer accuracy, user accuracy and Kappa coefficients were calculated for each vegetation category in each of the classification methods.

3.4 Comparison of Object-based and Pixel-based Classification Methods

Satellite data for the study area was classified using object-based and pixel-based (maximum likelihood and ISODATA) methods. Only spectral features (Blue, Green, Red and NIR) were used for this part of the research. Level 4 vegetation maps were generated for each method, and the results of the classifications were compared. In the object-based classification, Segmentation was implemented using eCognition Version 4 software; and classification by minimum distance classifier by the Mahalanobis distance method using discriminate analysis with SPSS Version 12 software.

The master vegetation map and aerial photographs were used as training data for the image objects generated by segmentation. In many cases, vegetation categories showed individual variation in tree crown shapes based on age and composition of forest. In particular, evergreen broad-leaved categories in the study area included mature forests of chinquapin (*Castanopsis sieboldii*), as well as younger forests of chinquapin mixed with evergreen oaks (*Quercus myrsinaefolia*, *Q. acuta*, etc). Coniferous plantations also showed variation in crown density. In these cases, the different age and composition patches were originally classified separately, but were consolidated into their main categories (either evergreen broad-leaved forest or coniferous plantation) in the final analysis.

The same vegetation categories were employed in the pixel-based maximum likelihood and ISODATA classifications, which were implemented for purposes of comparison. The training data used in the object-based classification were also used in the maximum likelihood classification. In the ISODATA classification, the data was divided into 40 clusters and labeled.

3.5 Classification using texture features

In addition to spectral features, textural and morphological features can be effectively employed in vegetation classification. Carleer and Wolff (2006), for example, performed classifications utilizing 4 relevant features selected from a total of 33 spectral, textural and morphological features. In other cases, however, a greater number of relevant features may be preferable. This is especially so in cases, such as the study area in this research, where the vegetation pattern includes a mixture of objects with very similar characteristics.

Using a stepwise regression in the discriminate analysis is one method for choosing the best number of relevant features for a particular classifica-

tion (von Eye and Schuster, 1998). This method attempts to find the "best" regression without examining all possible regressions (Levine et al. 2001), and is thus especially well suited for use in Japan, where the vegetation categories vary widely from region to region, and where many small patches of vegetations types are mixed together in a complicated mosaic pattern.

In this research, a total of 28 spectral and textural features (textural features based on Haralick et al. 1973) were established (Table 2), and stepwise regression, relying on the minimum distance (Mahalanobis distance), was employed within the discriminate analysis to select those features that were relevant for statistical classification. The features selected were then classified by minimum distance classifier (Mahalanobis distance), using discriminate analysis with SPSS Version 12 software. The Mahalanobis distance between sample i and class j in both the feature selection and classification stages was calculated as follows (Eq.3.2).

$$d^2_{ij} = (x_i - \mu_j) \, {}^t\Sigma_j^{-1} \, (x_i - \mu_j) \tag{3.2}$$

Where x denotes vector of raw data, μ denotes mean vector of population and Σ^{-1} denotes inverse covariance matrix of population.

When selecting features in the stepwise regression, the F-value was employed to statistically assess the usefulness of an explanatory feature for discriminating class. The F-value was calculated as follows (Eq.3.3).

$$F = \{b_i / SE(b_i)\}^2 \tag{3.3}$$

Where $SE(b_i)$ denotes standard error of the partial regression coefficient (b_i). When the critical (F-to-enter) F-value is set at 1.0, a maximum number of explanatory factors are selected. To prevent over selection, setting the critical F-value at 2.0 has been suggested (Haga and Hashimoto 1980, Okuno et al. 1981).

On the other hand, setting the value this high might result in loss of some useful features. In this research, the regression and subsequent classification were thus run at both critical F-value 1.0 and 2.0, and the results were compared. In addition, to determine the extent to which the results can be applied to actual vegetation mapping, the regressions and classifications were run at Level 1 through Level 4.

Table 2. 28 features used in study

Spectral features		Textural features	
Mean of Blue band	s1	Homogeneity on NIR band	t1
Mean of Green band	s2	Homogeneity on PAN band	t2
Mean of Red band	s3	Contrast on NIR band	t3
Mean of NIR band	s4	Contrast on PAN band	t4

Mean of PAN band	s5	Dissimilarity on NIR band	t5
Mean of NDVI band	s6	Dissimilarity on PAN band	t6
St. Dev. of Blue band	s7	Entropy on NIR band	t7
St. Dev. of Green band	s8	Entropy on PAN band	t8
St. Dev. of Red band	s9	Angular second moment on NIR band	t9
St. Dev. of NIR band	s10	Angular second moment on PAN band	t10
St. Dev. of PAN band	s11	Mean on NIR band	t11
St. Dev. of NDVI band	s12	Mean on PAN band	t12
		St. Dev on NIR band	t13
		St. Dev on PAN band	t14
		Correlation on NIR band	t15
		Correlation on PAN band	t16

Textural features on GLCM (Gray Level Co-occurrence Matrix).

4 Results and Discussion

4.1 Comparison of Object-based and Pixel-based Classification Methods

The image results generated by the three classification methods are shown in Figure 5; and the classification accuracy of those results are shown respectively in Tables 3. Only spectral features were utilized, and the classifications were performed at the fourth level (seven vegetation classes).

In terms of overall classification accuracy, the object-based results (58.9%) scored higher than both the maximum likelihood (52.8%) and the ISODATA (49.3%). In terms of overall Kappa coefficients as well, object-based (0.458) outscored maximum likelihood (0.373) and ISODATA (0.309).

In the ISODATA results, compared to the master map, the pixels classified as coniferous plantation are exceedingly high. Some sections are shown as completely covered by this category. This result can be attributed to clustering of Coniferous plantation and Evergreen broad-leaved forest into the same category. There is thus a large difference in the producer and user accuracies for these two categories, and the Kappa coefficients is low as well. For example, even though the producer accuracy for Coniferous plantation is 88%, this figure results from the large number of pixels classified in this category. Also, this classification method proved unable to distinguish Other plantation from different forest types, and simply labeled entire sections based on the dominant type found within.

The maximum likelihood method produced better results, in terms of both classification accuracy and Kappa coefficients, than the ISODATA method. On the other hand, this method suffered from the same problem of being unable to distinguish between Coniferous plantation and Evergreen

broad-leaved forest. As a result, large areas are classified as continuous Coniferous plantation. Compared to the ISODATA method, the maximum likelihood method proved more accurate at extracting Bamboo grove, but was unable to distinguish between Secondary grassland and Wetland vegetation. This resulted from similarities in the spectral characteristics of the plume grass (*Miscanthus sinensis*) and dwarf bamboo (*Pleioblastus chino.*) in the Secondary grasslands, and the reeds (*Phragmites australis*) in the Wetland vegetation.

| (a) | (b) | (c) |

Evergreen broad-leaved forest   Deciduous broad-leaved forest
Secondary grassland   Wetland vegetation   Coniferous plantation
Other plantation   Bamboo grove
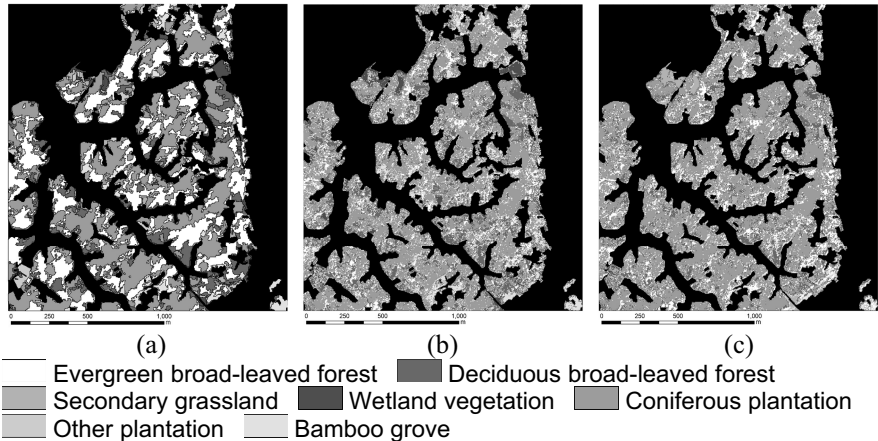
Fig. 5. Results of classification
(a): Minimum distance classifier: Maharanobis distance (object-based)
(b): Maximum likelihood classifier (pixel-based)
(c): ISODATA classifier (pixel-based)

Table 3. Classification accuracy for object-based and pixel-based classification methods

	Object-based		Maximum likelihood		ISODATA	
	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy
c9	60.2	44.2	19.9	52.2	19.5	34.8
c4	44.6	48.7	57.0	33.0	47.1	39.9
c5	26.7	71.9	10.5	50.0	62.8	44.3
c6	78.0	75.0	94.0	50.5	20.0	66.7
c10	73.6	70.5	84.6	59.4	87.9	56.9
c11	42.6	48.9	13.0	58.3	---	---
c8	44.7	74.5	48.2	69.5	7.1	25.0
Overall accuracy	58.9%		52.8%		49.3%	
Kappa coefficients	0.458		0.373		0.309	

In the object-based classification as well, there were some difficulties in distinguishing among various forest types. As the classification method also takes into consideration the relationships among adjacent units, these errors were mitigated, and the classification accuracy and Kappa coefficients are thus higher than for the two pixel-based classifications.

Also, in both the maximum likelihood and ISODATA classifications, there are a large number of tiny 'salt and pepper' units scattered throughout. These methods proved incapable of producing the distinct boundary lines seen on the master map. Comparing the object-based results with the master map, however, although there was some mislabeling, the boundaries can be seen to be fairly accurate. As is clear from the above discussions and comparisons, the results of this study show that object-based classification is better suited for vegetation mapping than pixel-based classification.

4.2 Classification using texture features

The results of object-based classification using additional features selected by stepwise regression for Level 1 through Level 4 are shown in Tables 4, 5, 6, and 7. The actual features selected by stepwise regression are shown at the bottom of the Tables.

For Level 1, the overall accuracy using only spectral features was 93.9%, which increased to 95.3 % for both critical F-values. The overall Kappa coefficients also increased with use of textural features. These results showed that for this level some increase in accuracy and Kappa coefficients can be attained by incorporating textural features.

For Level 2, the increase in overall accuracy was 3.7% for critical F-value at 2.0, and 4.7% at 1.0. The Kappa coefficients also increased, rising 0.11 for critical F-value 2.0 and 0.13 for 1.0. At this level, the improvement in performance obtained with the textural features was even greater than for Level 1.

For Level 3, the overall accuracy increased 4.6% when using additional features with the critical F-value set at 1.0. Kappa coefficients also increased more than 0.1. These results are more than acceptable for this level of classification.

At Level 4, the level of finest classification, overall accuracy and Kappa coefficients decreased substantially when compared to the lower levels. Still, the addition of 1.0 critical F-value textural features resulted in an improvement of 6% in overall accuracy and 0.1 in Kappa coefficients.

Overall accuracy ranged from 95.3% for Level 1 to 65.0% for Level 4. Overall Kappa coefficients using textural features for the classifications

ranged from 0.771 for Level 1 to 0.542 for Level 4. From experience, a Kappa coefficient of less than 0.40 is considered indicative of low correspondence, from 0.40 to 0.80 as medium level matching, and over 0.80 as high (Landis and Koch 1977).

Table 4. Classification accuracy for Level 1

	Spectral only		2.0 critical F-value		1.0 critical F-value	
	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy
c1	98.8	94.4	99.7	95.1	99.8	95.0
c2	62.5	89.5	67.7	96.8	66.9	97.6
Overall accuracy	93.9%		95.3%		95.3%	
Kappa coefficients	0.703		0.771		0.769	
selected features (critical F-value=2.0): s1,s7,s8,s11,t4,t5,t6,t13,t16						
selected features (critical F-value=1.0): s1,s7,s8,s11,t4,t5,t6,t11,t13,t16						

Table 5. Classification accuracy for Level 2

	Spectral only		2.0 critical F-value		1.0 critical F-value	
	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy
c3	94.1	87.3	94.2	90.3	95.4	90.1
c4	33.9	39.4	48.8	43.7	48.8	52.2
c5	33.7	65.9	52.3	93.8	60.5	85.3
c6	78.0	76.5	82.0	97.6	70.0	89.7
Overall accuracy	80.8%		84.5%		85.5%	
Kappa coefficients	0.503		0.614		0.632	
selected features (critical F-value=2.0): s1,s3,s7,s8,s9,s10,s11,t1,t4,t6,t12,t13,t14						
selected features (critical F-value=1.0): s1,s2,s3,s5,s6,s7,s8,s9,s10,s11						
t1,t3,t4,t6,t8,t9,t12,t13,t14,t15,t16						

Table 6. Classification accuracy for Level 3

	Spectral only		2.0 critical F-value		1.0 critical F-value	
	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy
c7	92.6	81.1	89.7	82.4	93.3	84.1
c4	33.9	40.2	46.3	43.1	49.6	52.2
c5	32.6	57.1	54.7	61.8	55.8	73.9
c6	78.0	76.5	64.0	97.0	68.0	94.4
c8	41.0	74.5	44.7	84.4	49.4	77.8
Overall accuracy	75.2%		76.3%		79.8%	

Kappa coefficients	0.486	0.527	0.591

selected features (critical F-value=2.0): s1,s2,s4,s5,s6,s7,s8,s10,s11,s12
t4,t6,t7,t8,t11,t12,t13,t14,t16

selected features (critical F-value=1.0): s1,s2,s3,s4,s5,s6,s7,s8,s9,s10,s11,s12
t3,t4,t6,t7,t8,t9,t12,t13,t14,t15,t16

Table 7. Classification accuracy for Level 4

	Spectral only		2.0 critical F-value		1.0 critical F-value	
	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy
c9	60.2	44.2	64.3	50.5	68.1	53.6
c4	44.6	48.7	47.1	53.8	49.6	57.1
c5	26.7	71.9	52.3	95.8	54.7	79.7
c6	78.0	75.0	82.0	91.1	68.0	91.9
c10	73.6	70.5	70.8	68.9	71.9	70.9
c11	42.6	48.9	51.9	43.1	59.3	52.5
c8	44.7	74.5	55.3	82.5	61.2	81.3
Overall accuracy	58.9%		63.0%		65.0%	
Kappa coefficients	0.458		0.515		0.542	

selected features (critical F-value=2.0): s1,s2,s3,s5,s6,s7,s8,s9,s10,s11,s12
t3,t4,t5,t6,t13,t14,t16

selected features (critical F-value=1.0): s1,s2,s3,s4,s5,s6,s7,s8,s9,s10,s11,s12
t1,t3,t4,t5,t6,t7,t8,t9,t11,t12,t13,t14,t15,t16

As can be seen from the above tables, the results of this research agree with previous studies, such as those by Carleer and Wolff (2006) and Puissant et al (2005), that report a gain in classification accuracy by using both spectral and textural features.

Still, even using this system, in some cases acceptable results may be difficult to obtain. In this study area, for example, secondary or immature evergreen broad-leaved forests and irregular coniferous plantations presented a special challenge. Mature evergreen broad-leaved forests show dense canopies, and the chinquapins and evergreen oaks have consistently shaped crowns. These patches can usually be extracted successfully. Well tended, mature conifer plantations also show good crown consistency, and can usually be extracted. Younger evergreen broad-leaved forests, however, have more open canopies, and the trees show more varied crowns. Poorly tended or disrupted coniferous plantations also have open canopies and irregular crowns. There is thus very little, in terms of both spectral and textural features, to distinguish between immature evergreen broad-leaved forests and untended coniferous plantations.

In addition, the images were acquired on April 1, when the deciduous trees had not yet leafed out. As a result, dwarf bamboo and evergreen shrubs in the understory were visible, causing some of these patches to be mistakenly classified as coniferous plantation. This problem, however, can be eliminated by acquiring images later in the year, when the deciduous trees are in leaf.

5 Conclusions

The results of this research show that object-based classification using very high resolution satellite remote sensing data can be an effective tool for improving efficiency and consistency in generating vegetation maps. Object-based classification proved better than pixel-based methods at extracting boundaries among vegetation patches, and at eliminating over classification. The results obtained from the object-based classifications were much closer to the master maps than those achieved by the pixel-based. At the coarser levels of classification, spectral features alone produced excellent results. The highly accurate results produced at these levels can be used for distinguishing between forest and grassland and among some basic forest and grassland types. Addition of textural features improved overall accuracy and Kappa coefficients at all levels. The gain in accuracy obtained by addition of textural features proved especially vital at the finer levels of classification, which are close to the actual level employed in vegetation mapping.

In this research, a stepwise regression was used in the discriminate analysis to select the most appropriate number of features for the classification. A total of 21 features out of a possible 28 were selected for the Level 2 classification, and 23 and 26 respectively for Level 3 and Level 4. Although in this case the regression selected a high proportion of the available features, in other cases, especially when the vegetation patterns are simpler or the number of categories smaller, the number of features selected might be considerably fewer. We thus believe that the feature selection process is an effective tool for improving classification results.

Some problems remained with misclassification, but object-based classification using both spectral and textural features proved capable of generation classification accuracy ranging from 65-95%, even in regions like the Japanese countryside, where small patches of different vegetation types are distributed in a complex mosaic pattern.

This research clearly indicates that very high resolution remote sensing data has a high potential for practical use in producing and updating the

vegetation maps required for proper conservation and management of countryside landscapes. In the future, classification accuracy can be improved even further using advanced classification methods such as segmentation optimization and multi-temporal data.

Acknowledgements

References

Alexander R, Millington AC (2000) Vegetation Mapping. John Wiley & Sons Ltd, Chichester

Baatz M, Schape A (2000) Multiresolution Segmentation–an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Blaschke T (ed) Angewandte Geographische Informations Verarbeitung XII, Wichmann-Verlag, pp.12-23

Biodiversity Center of Japan (2004) The 6th National Survey on the Natural Environment; Survey of Vegetation, Japan. (in Japanese)
    http://www.vegetation.jp/6thgaiyou/sakusei.html (accessed 15 Mar. 2005)

Braun-Blanquet J (1932) Plant sociology: The study of plant communities. McGraw-Hill, New York and London

Carleer AP, Wolff E (2006) Urban land cover multi-level region-based classification of VHR data by selecting relevant features. Int. J. Remote Sensing 27:1035-1051

Haga T, Hashiomto S (1980) Regression analysis and principal component analysis. JUSE Press, Ltd., Tokyo (in Japanese)

Haralick RM, Shanmugam K, Dinstein I (1973) Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics 3:610-621

Hay GJ, Niemann KO, McLean GF (1996) An object-specific image-texture analysis of H-resolution forest imagery. Remote Sensing of Environment 55:108-122

Herold M, Gardner ME, Roberts DA (2003) Spectral resolution requirements for mapping urban areas. IEEE Transactions on Geoscience and Remote Sensing 41: 1907-1919

Kamagata N, Hara K, Mori M, Akamatsu Y, Li Q, Hoshino Y (2006) A new method of vegetation mapping by object-based classification using high resolution satellite data. Journal of the Japan Society of Photogrammetry and Remote Sensing 45:43-49 (in Japanese)

Laliberte AS, Fredrickson EL, Rango A (2007) Combining decision tree with hierarchical object-oriented image analysis for mapping arid rangelands. Photogrammetric Engineering & Remote Sensing 73:197-207

Landis J, Koch G (1977) The measurement of observer agreement for categorical data. Biometrics 33:159-174

Levine DM, Ramsey PP, Smidt RK (2001) Applied statistics for engineers and scientists. Prentice Hall, Upper Saddle River

Marçal ARS, Borges JS, Gomes JA, Costa JF, Pinto D (2005) Land cover update by supervised classification of segmented ASTER images. Int. J. Remote Sensing 26:1347-1362

Mumby PJ, Edwards AJ (2002) Mapping marine environments with IKONOS imagery: Enhanced spatial resolution can deliver greater thematic accuracy. Remote Sensing of Environment 82:248-257

Okuno T, Haga T, Kume H, Yoshizawa T (1981) Multivariate analysis revised edition. JUSE Press, Ltd., Tokyo (in Japanese)

Puissant A, Hirsch J, Weber C (2005) The utility of texture analysis to improve per-pixel classification for high to very high spatial resolution imagery. Int. J. Remote Sensing 26:733-745

Suzuki H, Ito S, Toyohara G (1985) Method of Vegetation Survey II: Phytosociological Approach. Kyoritsu shuppan, Ltd, Tokyo (in Japanese)

Tuominen S, Pekkarinen A (2005) Performance of different spectral and textural aerial photograph features in multi-source forest inventory. Remote Sensing of Environment 94:256-268

von Eye A, Schuster C (1998) Regression Analysis for Social Sciences, ACADEMIC PRESS, California

Wang L, Sousa WP, Gong P (2004) Integration of object-based and pixel-based classification for mapping mangroves with IKONOS imagery. Int. J. Remote Sensing 25:5655-5668

Chapter 5.2

Structural biodiversity monitoring in savanna ecosystems: Integrating LiDAR and high resolution imagery through object-based image analysis

S.R. Levick, K.H. Rogers

Centre for Water in the Environment, University of the Witwatersrand, Private Bag 3, WITS, 2050, South Africa, shaun@shaunlevick.com

ABSTRACT: Savannas are heterogeneous systems characterized by the coexistence of grasses and woody trees. Growing recognition of the importance of the structural component of biodiversity has highlighted the need to understand the spatial distribution and temporal dynamics of woody plant structural diversity. Advances in LiDAR technology have enabled three dimensional information of vegetation to be obtained remotely over large areas. Whilst the use of LiDAR has gained considerable momentum in forested areas there has been limited application to savanna systems. We explore the applicability of LiDAR and object-based image analysis to the monitoring of woody structural diversity in a savanna system. We demonstrate how an object-based approach to image analysis significantly improves the accuracy of woody layer classification form in a heterogeneous landscape. Furthermore we illustrate how standard approaches to LiDAR derived canopy models suffer from interpolation artifacts in savannas, due to the heterogeneity of the woody layer. By integrating LiDAR with high resolution aerial photography, through object-based analysis, these artifacts can be removed to produce a robust canopy model. The object-based integration of LiDAR with aerial imagery holds immense potential for structural diversity monitoring in savannas.

1 Monitoring structural biodiversity in savanna ecosystems

Savannas are heterogeneous environments driven by a wide range of factors at multiple scales. A key characteristic of savanna landscapes is the co-dominance of two life forms – grasses and woody trees (Scholes and Walker 1993). The spatial structure and composition of savannas is controlled primarily at the broad scale by climate and geology, whilst rainfall, topography, soil type, fire and herbivory influence structure at a range of finer scales (Pickett et al. 2003, Gillson 2004a, Sankaran et al. 2005). In addition to being spatially heterogeneous, savannas are highly dynamic over time (Gillson 2004b). The variability of these systems presents challenges to their management and conservation.

Management of savanna systems has historically taken place under a balance of nature/homogeneity paradigm (Rogers 2003). The growing recognition of savanna heterogeneity has led to changes in the management of certain savanna systems. In the Kruger Park (South Africa), for example, management has adopted a heterogeneity paradigm that 'aims to maintain biodiversity in all its facets and fluxes' (Braack 1997). This paradigm shift reflects a holistic view of biodiversity which incorporates the composition, structure and function of ecological systems at multiple scales (Noss 1990). Given that heterogeneity is considered to be the ultimate source of biodiversity (Pickett et al. 2003), monitoring system heterogeneity should be of high management priority within savanna systems.

1.1 Monitoring savanna heterogeneity remotely

Monitoring of savanna vegetation has traditionally taken place through aerial photographic analyses and field surveys. Ground based field monitoring can provide detailed information of changes in vegetation structure over time, but is very time intensive and can only feasibly be conducted over small spatial scales. Fixed point photography can reveal changes in the three-dimensional structure of vegetation, but it suffers the same constraints as field measurements. Extrapolating results obtained at small spatial scales to larger scales is difficult in heterogeneous systems like savannas. Managers need to be able to monitor large spatial areas to in order to encompass system variability. Remote sensing techniques at broader scales therefore need to be employed.

Savannas have historically presented numerous challenges to the field of remote sensing. Given their proximity to the tropics, and the regular occurrence of thunderstorms in summer months, cloud free days are rarely

found during the growing season. Synthetic Aperture Radar (SAR), however, holds a lot of potential for monitoring vegetation communities in savanna landscapes as it has the ability to penetrate cloud cover (Menges et al. 1999). SAR can therefore provide valuable insight into temporal changes in ecosystems by enabling land cover monitoring at all times of the year. SAR has also shown potential for broad scale biomass estimation, although accuracy has been shown to fall with increasing biomass and leaf area index (Waring et al. 1995). Fine scale three-dimensional representation of vegetation requires the use of Light Detection and Ranging (LiDAR).

In recent years LiDAR has become commercially available and provides a robust means of measuring the three-dimensional structure of terrain and vegetation surfaces remotely. LiDAR has been utilized extensively in forestry applications and has been shown to reliably return ground elevation and tree height data in forested systems (Lefsky et al. 2002). LiDAR has experienced limited use in savanna landscapes, although Dowling and Accad (2003) and Lovell et al. (2003) have explored its potential in Australian savannas.

Given that savannas are heterogeneous systems at multiple scales, interpretation of remotely sensed data should be conducted in hierarchical manner which accounts for spatial variation across the landscape. Object-based image analysis provides a means for achieving this objective.

1.2 The object-based approach to image analysis

Object-based image analysis arose through the realization that image objects hold more real world value than pixels alone (Blaschke and Strobel 2001). The software eCognition 4.0, developed by Definiens Imaging, adopts an object-based image analysis approach and provides a platform for incorporating contextual and ancillary data in image classification. The first step in the analysis is the mutiresolutional segmentation of an image into areas of homogeneity. Homogeneity criteria are based on both spectral and shape properties. A bottom-up region merging technique is employed where smaller objects are merged into larger ones based on the criteria set. The approach allows for segmentation at different scales, which is used to construct a hierarchical network of image objects representing the image information in different spatial resolutions simultaneously (Laliberte et al. 2004). The image objects have relationships to both adjacent objects on the same level and objects on different hierarchical levels.

Classification is then performed on the image objects, not the pixels, at the desired scale.

In this chapter we demonstrate how object-based image analysis of both aerial imagery and LiDAR can aid the structural diversity monitoring process in savanna landscapes. Our first example explores the classification of black and white aerial photography, which is an important record of historical woody cover and is a critical step in temporal change analyses. The second example examines the integration of both LiDAR and color aerial photography for representing the three-dimensional structure of savanna vegetation. The research was conducted in the Shingwedzi Catchment of the northern Kruger Park, South Africa (Figure 1).
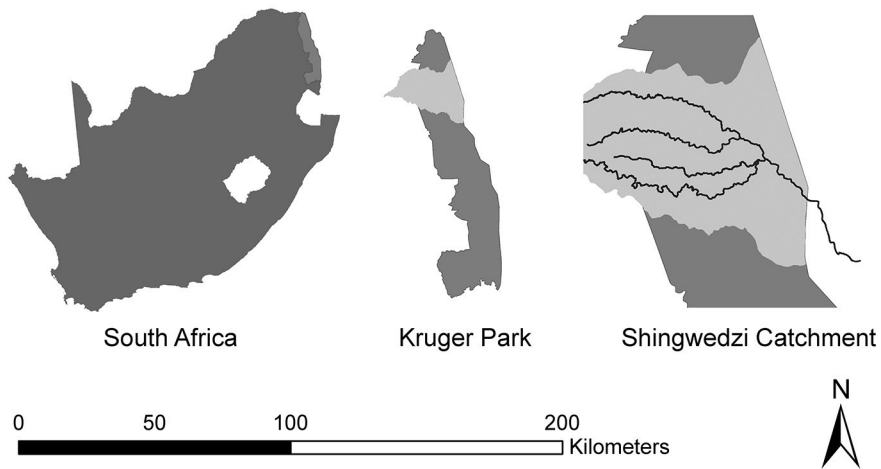
Fig. 1. Study location – Shingwedzi Catchment, Kruger Park, South Africa

2 Woody canopy delineation from black and white aerial photographs

Black and white aerial photographic records provide a means of exploring vegetation changes over fairly large spatial areas. These records provide valuable evidence of changes in woody vegetation cover over time, but accurately extracting the woody layer has proved difficult. Much of the difficulty in extracting woody cover from aerial photographs stems from the heterogeneity inherent in savanna landscapes.
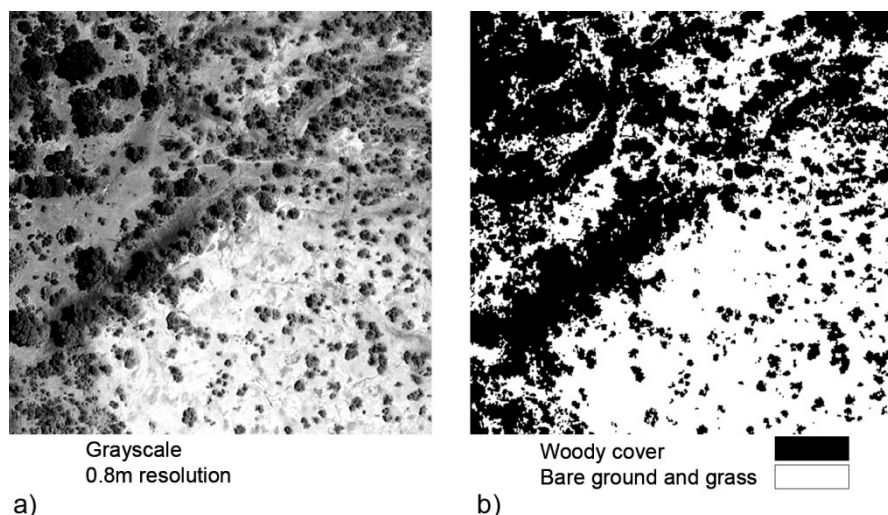
Fig. 2. a) Aerial photograph of a heterogeneous savanna landscape. **b)** Pixel-based classification of woody vegetation resulting in an overestimate of woody cover and 68% accuracy

Figure 2a, for example, depicts an aerial view of the riparian fringe adjacent to the Phugwane river of the Shingwedzi Catchment. The image was taken in May 2001. A split between dark basaltic soils (top left-hand corner) and white alluvial soils (bottom right-hand corner) runs diagonally through the image. This variation in soil color presents challenges when trying to extract woody cover from the image. The dark basaltic soils are of similar brightness to some of the woody vegetation types.

A traditional pixel based classification of woody cover (Figure 2b) fails to extract only the woody plants and overestimates woody coverage immensely. This is primarily the result of some dark soil areas being classified as woody canopy.

Four 1km X 1km sites were selected along each of the four rivers of the Shingwedzi catchment for classification validation purposes. Field based validation was not possible due to the historical nature of the photographs. Visual validation was therefore performed whereby 50 woody cover and 50 bare ground points where digitized onscreen for each of the 16 sites. An error matrix was constructed to access woody cover classification accuracy at the 800 validation point locations. A maximum likelihood pixel based approach only achieved 68% accuracy when compared against the validation data. This is clearly not acceptable for monitoring purposes. To im-

prove this classification, multi-scaled object-based image analysis is needed.

2.1 Object-based processing of historical aerial photographs for woody canopy extraction

eCognition 4.0 was used to conduct multi-resolutional segmentation and classification on Figure 3a. Prior to segmentation the image was filtered with a 3 X 3 low-pass filter to remove excessive variation. Smoothing of layers prior to segmentation helps produce fewer, and more homogeneous image objects, so that individual trees are represented by fewer polygons (Laliberte et al. 2004). A fine level of segmentation was initially chosen to ensure that image objects were small enough to represent individual trees (Figure 3b). Larger scale segmentation was then conducted to group areas of similar vegetation/soil type units together (Figure 3c). The primary aim of this broader segmentation is to provide some spatial context for the smaller 'tree' objects at the lower level. Laliberte et al. (2004) used this technique to successfully extract shrubs from aerial photographs of arid rangelands in southern New Mexico.

Although there is little difference in brightness between some of the woody trees and the basalt soils in Figure 3a, the difference between the mean of image objects in Figure 3b and the larger image objects in Figure 3c can be used to differentiate trees from soil. During the classification process, first level objects were considered woody cover if they had a mean brightness value of between 0 and 90, as well as a ratio of between 0 and 0.95 relative to their super object. The resulting classification (Figure 3d) was 97% accurate when tested against the validation data. By adopting the object-based approach, contextual and ancillary data can be included in the classification process to produce more robust results.

a) Grayscale 0.8m resolution

b) Scale parameter 3 / Shape factor 0.2 / Smoothness 0.8

c) Scale parameter 250 / Shape factor 0.5 / Smoothness 0.5

d) Woody cover / Bare ground and grass
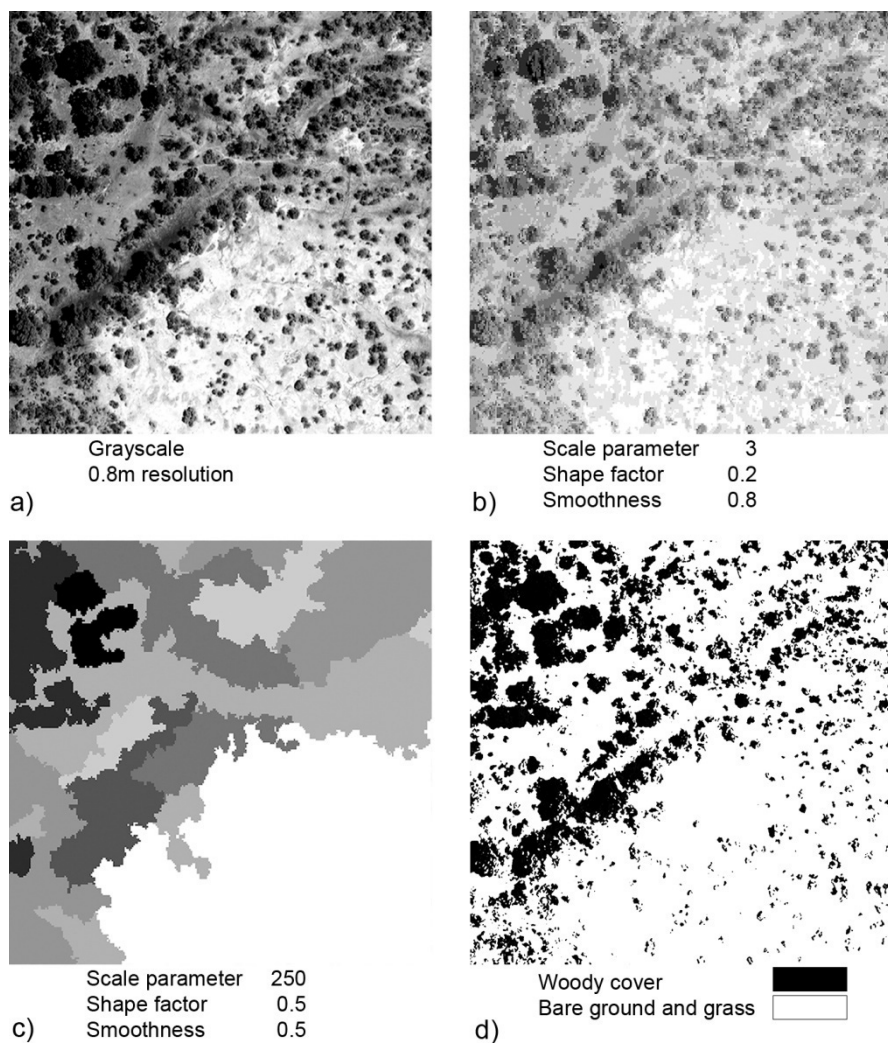
Fig. 3. a) Aerial photograph of heterogeneous landscape. **b)** Fine scale segmentation **c)** Large scale segmentation **d)** Object-based classification of woody cover resulting in 97% accuracy.

3 Extracting woody vegetation structural attributes from LiDAR and high resolution aerial photography

Aerial photography records are a valuable resource for monitoring changes in vegetation cover over time. They are, however, limited in their ability to depict three-dimensional changes in woody structure. Managers concerned with changes in the vertical structure of trees need to look for alternate monitoring techniques. LiDAR, in conjunction with high resolution aerial photography, provides a remote sensing solution for monitoring vegetation structural diversity in savanna landscapes. The instrument used in this example was a first/last pulse ALTM 1225 (Optech Inc., Canada) with a pulse frequency of 25kHz. The flight path focused on the four major rivers of the Shingwedzi Catchment (Figure 4).
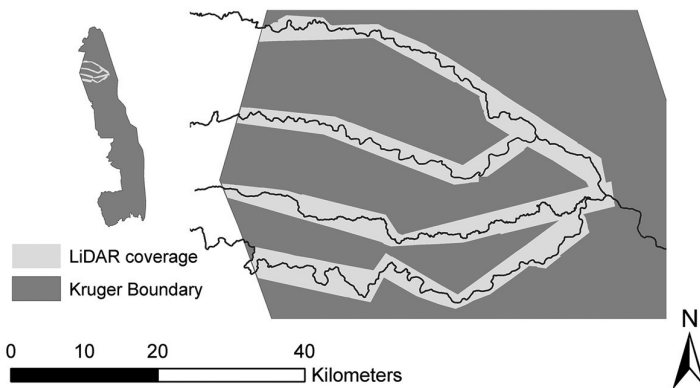
Fig. 4. LiDAR coverage of the four major rivers of the Shingwedzi Catchment

3.1 Normalized vegetation canopy model (nVCM) construction

A key advantage of using LiDAR, from a vegetation monitoring point of view, is the ability to create a normalized vegetation canopy model (nVCM). An nVCM is a spatial raster representation of above ground tree height. Normalized vegetation canopy models are widely used in forestry and can be utilized for monitoring changes in tree height and biomass over time. An nVCM is typically constructed from first creating a digital terrain model (DTM) from the LiDAR ground returned points and a digital surface model (DSM) from the total LiDAR points. The DTM can then be subtracted from the DSM to create a surface of above ground elevation. In natural landscapes this generates an nVCM. This system is widely used in

forested systems and has been shown to return reliable estimates of above ground vegetation height (Maier et al. 2006, Tiede et al. 2006).
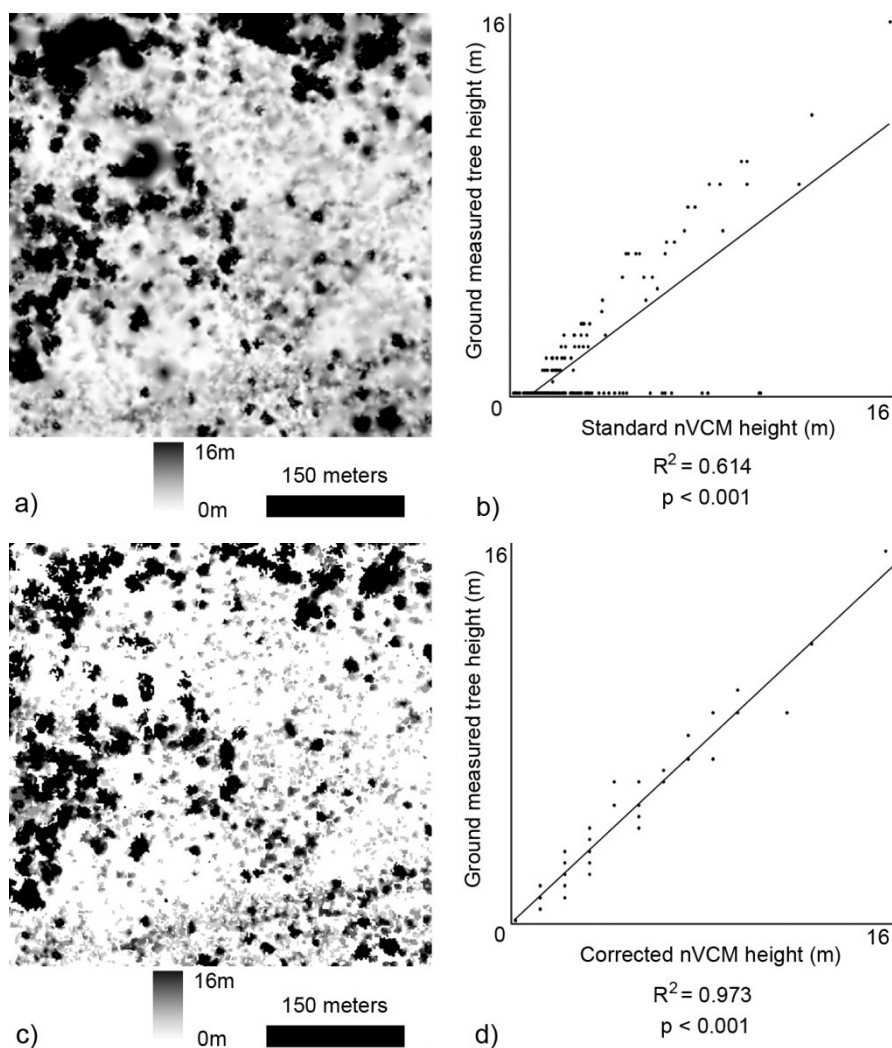
Fig. 5. a) Subset of a standard nVCM derived from DSM and DTM subtraction. **b)** Regression of field measured woody canopy height against standard nVCM derived from LiDAR. **c)** nVCM corrected for discontinuous canopy structure through object-based analysis. **d)** Regression of field measured woody canopy height against the corrected nVCM.

We constructed an nVCM by standard DSM-DTM subtraction techniques for a subset of the Shingwedzi dataset. The resulting model produces a visually realistic representation of woody canopy height (Figure 5a). When this subset of the total LiDAR coverage was tested against 200 ground validated data points, however, it is clear that the model overestimates vegetation height in areas of bare ground and grass cover (Figure 5b). Ground validation was performed at 200 stratified random points. Points were located with a differential GPS and canopy height was measured with a Vertex III hypsometer.

This artifact is due to the discontinuous nature of savanna canopy layers. In forested systems with continuous canopies, standard nVCM calculations are sufficient, but gaps between savanna trees result in interpolation artifacts between trees which results in an overestimation of tree height in the canopy gaps. In order to address this issue the actual woody canopy coverage needs to be extracted from high resolution aerial imagery. We achieved this by combining the LiDAR data with high-resolution color aerial photography through object-based image analysis. Building on from the black and white photograph workflow, the image was segmented at both a fine and broad scale. The fine scale segmentation (scale parameter = 3, shape factor = 0.2, smoothness = 0.8) delineated individual tree objects and the broad scale segmentation (scale parameter = 250, shape factor = 0.5, smoothness = 0.5) provided context for woody canopy classification. Once the woody canopy layer was clearly defined, it was used as a mask on the standard nVCM to eliminate artifacts created by the gaps in the canopy and to produce an nVCM corrected for discontinuous canopies. The corrected nVCM (Figure 5c) is more robust when tested against the ground validated data (Figure 5d). Interpolation errors between the tree canopies are removed through the object-based image analysis workflow that was followed (Figure 6).
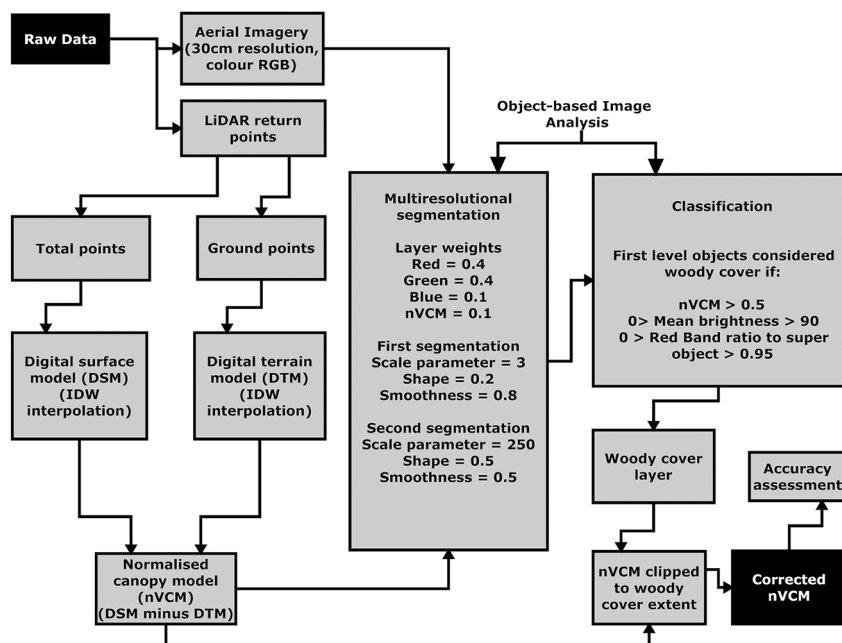
Fig. 6. Workflow model for developing a robust normalized vegetation canopy model (nVCM) for landscapes with discontinuous vegetation cover

The cross-sectional profile in Figure 7 runs through both the standard nVCM (grey) and the corrected nVCM (black) and highlights the differences between the two approaches. There is very little difference between the standard and corrected nVCM in areas where trees are present, it is in the gaps between trees that the standard nVCM overestimates above ground canopy height. The cross-section through the standard nVCM returns a mean canopy height of 5.17m with a coefficient of variation equal to 78.975. The corrected nVCM, however, returns a mean of 3.74m and a coefficient of variation equal to 121.16. This has important implications for the monitoring of vegetation structure and diversity. Without applying the object-based approach in savannas, managers may greatly overestimate the above ground canopy height and standing biomass of the system, and underestimate the level of structural diversity.
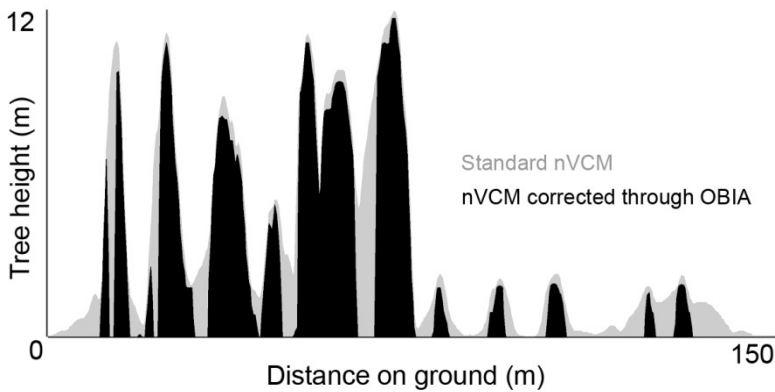
Fig. 7. Cross-section through the standard nVCM (grey) and the nVCM corrected through object-based image analysis (black)

This also has important implications for the ground validation process. The random approach taken to our ground point selection enforced the sampling of both trees and the bare ground between them. If validation points had only been collected at locations where trees were present, as is often the case, the interpolation artifacts and overestimation errors would not have been detected.

After refining the workflow on the subset of the LiDAR, the technique was applied across the entire LiDAR coverage. A stratified random sampling technique was used to select 500 canopy and a 500 inter-canopy points across the LiDAR coverage. Points were located with a differential GPS and canopy height was measured with a Vertex III hypsometer. The resulting nVCM (Figure 8) was validated against 1000 ground control points and returned and R-squared value of 0.851 (p < 0.0001).
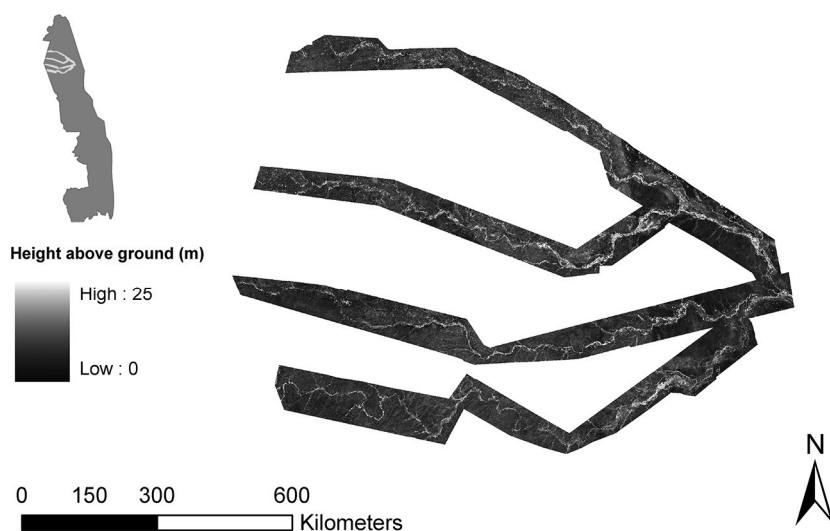
Fig. 8. Corrected normalized vegetation canopy model for the entire study site (R^2=0.851, p < 0.0001, 1000 ground control points)

4 Implications for the monitoring of savanna structural diversity

The heterogeneity of complex systems at different scales proves problematic for traditional pixel based classification techniques. The object-based approach, however, produces an accurate representation of woody cover from both black and white historical aerial photographs and high resolution color aerial photography. By using a multi-resolutional segmentation approach, and grouping homogeneous objects together at different scales, contextual and hierarchical information can be incorporated into the classification process. This procedure returns reliable woody cover classifications despite the complex heterogeneity of savanna systems.

Combining elevation data from LiDAR with high resolution digital color imagery through object-based image analysis greatly enhances the structural description of a landscape by adding the three-dimensional height component. The corrected normalized canopy model provides a more realistic representation of vegetation height distribution than standard DSM-DTM subtraction approaches. This is primarily due to the fact it does not assume continuous vegetation cover and accounts for the spatial heterogeneity of savanna woody cover.

Comparisons of remotely measured tree height with ground validated points indicate that structural attributes of woody vegetation can be reliably obtained from fusing LiDAR and imagery through object-based image analysis in a savanna system. This holds significant implications for vegetation management in savannas by providing a tool for monitoring vegetation structure remotely.

Understanding the patterns of spatial and temporal heterogeneity of a system is fundamental to its successful management. If we consider the reciprocal relationship between pattern and process in ecological systems (Turner 1989), understanding where structural changes occur spatially in the landscape can help elucidate the drivers of vegetation change. The fusion of LiDAR and imagery in an object-based image analysis environment provides the means for generating this spatio-temporal understanding.

The multi-scale, contextual approach inherent in object-based image analysis provides managers with a powerful tool for monitoring changes in vegetation structural diversity across heterogeneous landscapes at different scales.

6 References

Blaschke T and Strobel J (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS - Zeitschrift fur Geoinformationssysteme 6:12-17

Braack L (1997) An objectives hierarchy for the management of the KNP. A revision of parts of the management plan for the Kruger National Park.Volume VII

Dowling R and Accad A (2003)Vegetation classification of the riparian zone along the Brisbane River, Queensland Australia, using light detection and ranging (lidar) data and forward looking digital video. Can. J. Remote Sensing **29**:556-563.

Gillson L (2004a) Evidence of hierarchical patch dynamics in an East African savanna? Landscape Ecology 19:883-894.

Gillson L(2004b) Testing non-equilibrium theories in savannas: 1400 years of vegetation change in Tsavo National Park, Kenya. Ecological Complexity 1:281-298

Laliberte AS, Rango A, Harvstad KM, Paris JF, Beck RF, McNeely R, Gonzalez AL (2004) Object-oriented image analysis for mapping shrub encroachment from 1937 to 2003 in southern New Mexico. Remote Sensing of Environment 93:198-210

Lefsky MA, Cohen BW, Parker GG, Harding DJ (2002) Lidar remote sensing for ecosystem studies. Bioscience 52:19-30

Lovell JL, Jupp DLB, Culvenor DS and Coops NC (2003. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. Canadian Journal of Remote Sensing **29**:607-622.

Maier B, Tiede D, Dorren L (2006) Assessing mountain forest structure using airborne laser scanning and landscape metrics. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI-4/C42

Menges CH, Van Zyl JJ, Ahmad W and Hill GJE (1999) Classification of vegetation communities in the tropical savannas of northern Australia using AirSAR data. Pacific Rim AIRSAR Significant Results and Planning Workshop, Maui, Hawaii, 24-26 August

Noss RF (1990) Indicators for monitoring biodiversity: A hierarchical approach. Conservation Biology 4:355-364

Pickett STA, Cadenasso ML, Benning TL (2003) Biotic and Abiotic Variability as Key Determinants of Savanna Heterogeneity at Multiple Spatiotemporal Scales. In: du Toit JT, Rogers KH, Biggs H (ed) The Kruger Experience: Ecology and Management of Savanna Heterogeneity. Island Press, Washington, DC,pp 23-40

Rogers KH (2003) Adopting a heterogeneity paradigm: Implications for management of protected savannas. In: du Toit JT, Rogers KH, Biggs H (ed) The Kruger Experience: Ecology and Management of Savanna Heterogeneity. Island Press, Washington, DC,pp 41-58

Sankaran M, Hanan NP, Scholes RJ, Ratnam J, Augustine DJ, Cade BS, Gignoux J, Higgins SI, Le Roux X, Ludwig F, Ardo J, Banyikwa F, Bronn A, Bucini G, Caylor KK, Coughenour MB, Diouf A, Ekaya W, Feral CJ, February EC, Frost PGH, Hiernaux P, Hrabar H, Metzger KL, Prins HHT, Ringrose S, Sea W, Tews J, Worden J, Zambatis N (2005) Determinants of woody cover in African savannas. Nature 438:846-849

Scholes RJ and Walker BH (1993) An African savanna: a synthesis of the Nylsvley study. Cambridge University Press

Tiede D, Lang S, Hoffmann C (2006) Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences Vol. No. XXXVI-4/C42

Turner MG (1989) Landscape Ecology: The effect of pattern on process. Ann. Rev. Ecol. Syst. 20:171-197

Waring RH, Way J, Hunt ER Jr, Morrissey L, Ranson KJ, Weishampel JF, Oren R and Franklin SE (1995) Imaging radar for ecosystem studies. BioScience 45: 715-723

Chapter 5.3

Fusion of multispectral optical and SAR images towards operational land cover mapping in Central Europe

T. Riedel, C. Thiel, C. Schmullius

Department of Geography, Friedrich-Schiller-University, Germany,
tanja.riedel@uni-jena.de

KEYWORDS: Land cover mapping, automation, texture analysis, image fusion, optical & SAR

ABSTRACT: The availability of up-to-date and reliable land cover maps is of great importance for many earth science applications. For the generation of operational and transferable land cover products the development of semi- and fully-automated classification procedures is essential. The aim of this paper is to present a strategy for the generation of basic land cover products using both optical and SAR data. The study area is located in Northern Thuringia, Germany, with mainly forested regions of the eastern part of the Harz mountains and intensively used agricultural areas to the south. From April to December 2005 optical and SAR data were acquired continuously to generate a comprehensive time series.

The main objective of this work was to develop a working flow with high potential for automation. The proposed classification procedure is composed of three main stages. The first processing step comprises the segmentation of the optical EO-data. Next, potential training sites are being selected automatically by applying a decision tree with flexible, scene-specific thresholds calculated based on expert knowledge and histogram analyses. Finally, as the third step, training samples are being used as input to a supervised classification. Here, three classification methods were compared: nearest neighbor, fuzzy logic and a combined pixel-/object-based maximum likelihood classification. Best overall performance was achieved for the pixel-/object-based approach. In order to improve the product qual-

ity and accuracy, the classification was performed several times using randomly varying subsets of all potential training samples as input. The classification accuracy was improved significantly through the integration of textural features, especially for urban areas. Further, the advantage of applying the rarely used grey level dependency matrix is demonstrated.

1 Introduction

Remote sensing represents a cost-efficient method for large-area land cover mapping which provides spatially consistent and multitemporal information of the Earth's surface. The availability of reliable and up-to-date land cover information is required for a multitude of applications ranging from regional to global scales such as land cover change studies, ecological monitoring, map updating, management and planning activities or the implementation and control of national and international treaties (Franklin and Wulder 2002, Jensen 2000).

The development of robust, transferable, semi-automated and automated approaches is of particular interest for operational applications in order to save time and manpower. In regions with frequent cloud cover such as Central Europe the number of suitable optical data is often limited. The all-weather capability is one major advantage of radar systems with respect to optical systems. Furthermore, radar sensors provide information that is complementary to that of visible to infrared imagery. In the optical range of the electromagnetic spectrum the information depends on reflective and emissive characteristics of the Earth's surface, whereas the radar backscatter coefficient is primarily determined by structural and dielectric attributes of the surface target. The benefit of combining optical and Synthetic Aperture Radar (SAR) data for improved land cover mapping was demonstrated in several studies (Alparone et al. 2004, Amarsaikhan and Douglas 2004, Hegarat-Mascle 2000). In general, the data fusion process can be performed on the pixel, feature or decision level (see Ehlers et al. in this book, Pohl and Van Genderen 1998). With the availability of multifrequency and high-resolution spaceborne SAR data, such as provided by the TerraSAR-X and ALOS PALSAR missions, an increased interest in tools to exploit the full information content of both data types is arising.

Emphasis of this study was to develop a robust and transferable methodology for the generation of basic land cover products including a limited number of optical data, and exploiting the information content of multitemporal SAR data. The algorithm was developed in a study area in Thuringia and was successfully applied to data acquired at different seasons and

years. The main processing steps were programmed in IDL and the script application required no user interaction. Thus, the proposed procedure comprises a high potential for automation, which is a necessary prerequisite for operational applications. Additionally, minor adaptations to the processing chain can be implemented in an easy and straight forward way. This is advantageous for the integration of different EO data sources, the application to other topics of interest or the transferability to other regions in Europe or globally.

2 Study Area and Experimental Data

The study site is located in the northern part of Thuringia, Germany (Fig. 1). It encompasses an area of 31 km x 39 km covered mainly by forest at the higher elevations of the Harz mountains in the North and intensively used agricultural areas in the "Goldene Aue" plains in the southern part. From April to December 2005 Landsat-5 TM and HH/HV-polarized ASAR APP data (swath 2) were acquired continuously over the test site to generate a consistent time series (Fig. 2). For validation archived EO data from 2003 were purchased.
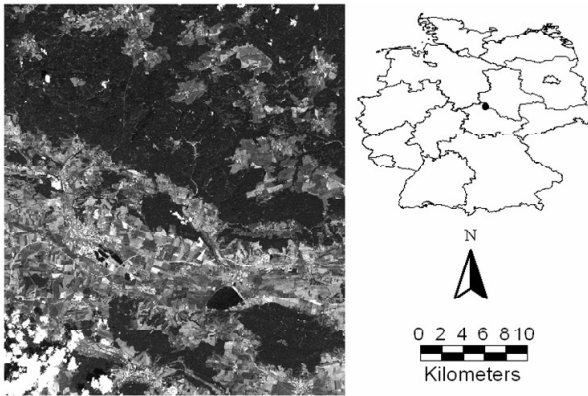
Fig. 1. Location of the test site – Landsat-5 TM image acquired on July 10, 2005, green channel
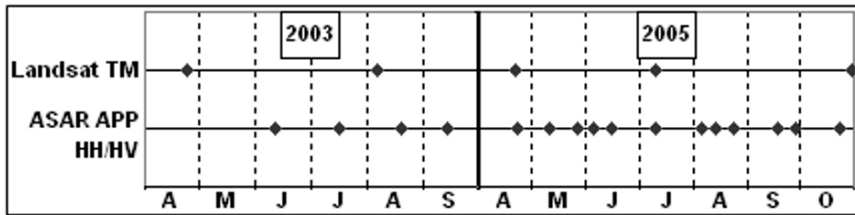
Fig. 2. EO-data base

3 Methodology

The basic land cover maps were generated by applying the object-based classification scheme described in this chapter and illustrated in Fig. 3. Before the classification procedure, all EO data were pre-processed on the basis of standard techniques. The normalization procedure introduced by Stussi et al. (1995) was applied to all SAR data since parts of the test site are characterized by rough terrain. The pre-processing of the multi-spectral optical images included atmospheric correction and orthorectification using the C-band SRTM DEM.

After pre-processing, image objects were delineated in the optical data with the eCognition software using its segmentation algorithm with a scale parameter of 15. The next processing step consisted of the selection of potential training samples for each land cover class based on the segment features in the optical, SAR and texture channels with an object-based decision tree. This step is implemented in IDL and runs fully automatically. As fixed thresholds for the training samples would fail in an operational procedure, these values have to be adapted to each EO scene separately. To collect representative image characteristics for each land cover type the available time series were analyzed in a systematic manner. Additionally, published information from the literature (e.g. European RAdar-Optical Research Assemblage library - ERA-ORA) was explored. By combining this expert knowledge about typical target characteristics (e.g. low reflectance of water bodies in the near infrared) and histogram analyses, it was possible to produce scene-specific, fixed threshold values. Another procedure to automatically identify optimal image features and the corresponding class thresholds is introduced by Marpu et al. in this book. However, the application of this algorithm requires the collection of training samples. In the next stage of the proposed classification scheme the identified training sites were used as input to supervised classification. Three classification techniques were compared: nearest neighbor, fuzzy logic and a com-

bined pixel-/object-based classification. The classification procedures were performed using standard software packages. If needed however, the implementation of the classification step in the existing IDL-routine is not a complex task. The classification routines were performed several times with randomly varying subsets of all training sets to improve product quality and accuracy. The final land cover class of each image segment was assigned by simple majority voting. Finally, postclassification procedures were applied involving simple GIS-analysis such as the re-coding of "island segments" within residential areas (these majority filters are also known as sieve functions).

Texture information was included as an additional input layer to the training and classification procedures mainly to improve the detection of urban areas. To use textural features jointly with spectral and backscatter information is a common approach to this problem. The advantages of applying texture information from medium resolution earth observation data was demonstrated by many studies (e.g. Dekker 2003, Dell' Acqua and Gamba 2003, Forsythe and Waters 2006, Ndi Nyoungui et al. 2002).In the framework of this study the potential of various texture measures such as standard deviation, data range and parameters derived based on the grey-level co-occurrence matrix (GLCM) were investigated. Additionally, the neighborhood grey level dependency matrix (NGLD) was calculated for the SAR data (Franke 2006). The NGLD algorithm – which is not available in standard image processing packages - was implemented using the IDL programming language, whereby a short processing time was realized. This is of great importance as long processing times often hinder the integration of textural parameters in the classification process. The NGLD matrix is invariant with respect to texture direction. The measure was calculated for moving windows of different sizes. For every pixel of the kernel the relation to all neighboring pixels at a specified distance d is investigated and the number of pixel pairs satisfying the relation is counted. The NGLD matrix is a squared matrix describing the pixel value a in one direction and the number of pixel pairs fulfilling the given relation n_r in the other direction. Each element Q_d of the matrix describes the number of occurrences of every possible combination of a and n_r. Analogous to the widely used co-occurrence matrix, several texture measures could be extracted from the NGLD matrix such as entropy, energy, minimal and maximal emphasis. The relation used for the computation of the NGLD matrix in this study is a difference of 0.1 between neighboring pixels at a distance of two pixels. This relation was selected based on statistical analyses of various test areas indicating a high proportion of neighboring pixels with a difference of at least 0.1 in urban areas (mostly greater than
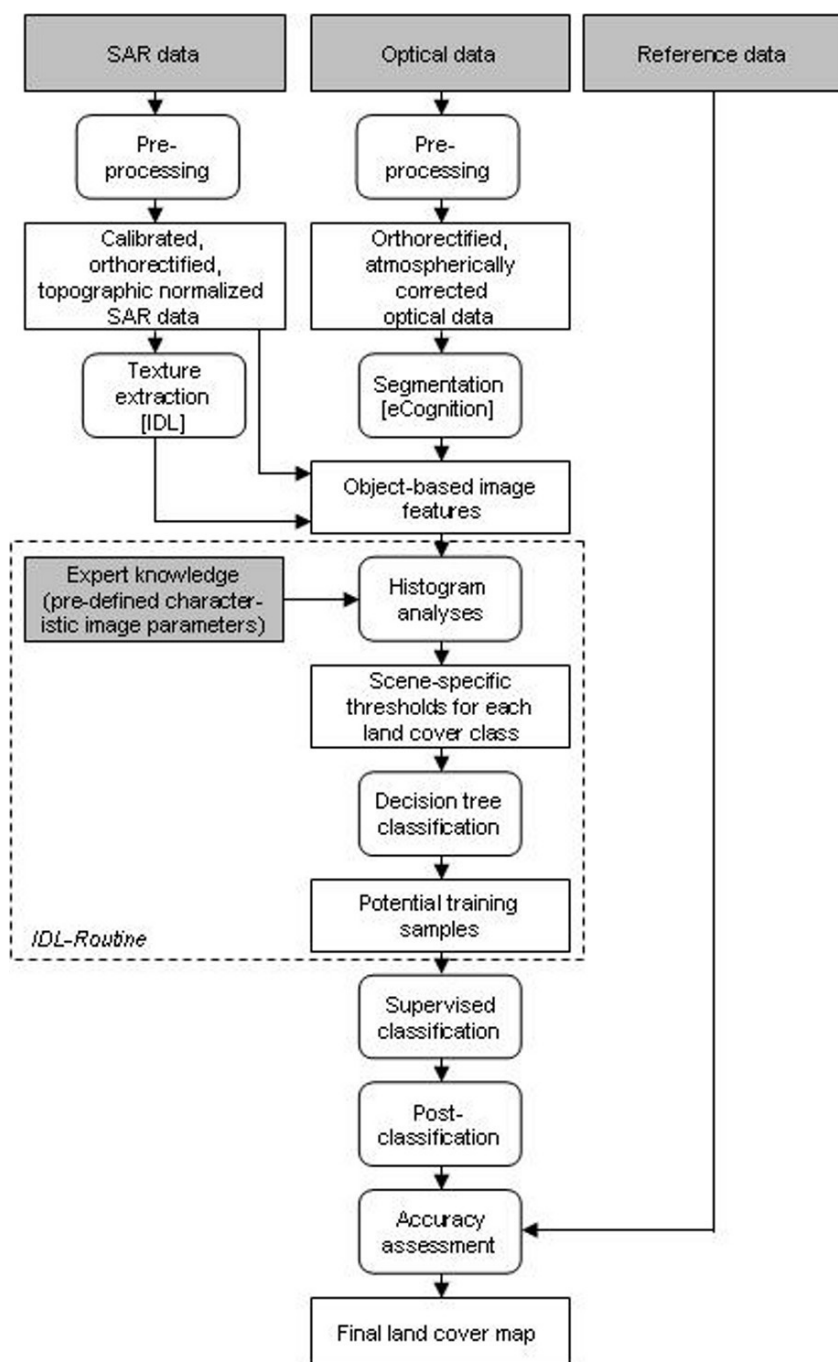
Fig. 3. Processing chain

40%). Special focus was on small villages and low density residential areas such as districts with single-family houses and gardens since the classification of these structures is most problematic. Besides a large difference in the radar signal between neighboring pixels, urban areas are characterized by a high radar backscatter at co-polarized C-band data. To account for this behavior, the following parameter, hereinafter called *urban area texture measure* UATM, was calculated based on the NGLD matrix:

$$UATM = \sum_{a=0}^{gz-1} \sum_{n_r=0}^{N} Q_d(a, n_r) * n_r * a \qquad (3.1)$$

The potential of the various investigated texture parameters was evaluated using separability analyses including the interpretation of signature plots and calculation of the Jeffries-Matusita distance (JM) as well as classification accuracy assessments. The JM-distance is a widely used measure in image processing to determine the statistical distance between two multivariate, Gaussian distributed signatures. Its values range from 0 to 1414, whereas 0 signifies no and 1414 a very high separability (Swain and Davis 1978, Richards 1999). Few features used in this study are not Gaussian, but JM is still considered to be a reasonable tool for assessing the potential of second-order measures for urban area mapping (Fukunaga 1990 and Ndi Nyoungui et al. 2002).

Thematic map accuracy of the final land cover map was assessed by calculating the confusion matrix and kappa coefficient for one hundred randomly distributed reference points per land cover category. The class membership of each reference target was determined on high resolution optical data (Quickbird satellite scenes and Hymap airborne data), air photographs, and official land surveying information as GIS layers. To test the robustness and stability of the proposed methodology, the classification scheme was also applied to EO data acquired two years earlier, in 2003.

4 Results

4.1 Texture measures for urban mapping

Several texture measures were investigated while focusing on synergy effects of texture in optical and SAR data for urban area mapping. The potential of the investigated texture features was quantified based on separability analyses and the achieved map accuracies. This chapter focuses on texture scenes only to emphasize the suitability and information content of each single texture measure for urban applications. Only simple object-

based decision tree classifications with hard thresholds were applied. The channels with spectral reflectances or backscattering information were not considered in this context.

A critical issue for texture extraction is the choice of the kernel size being used for computation. The impact of the window size on class separability is illustrated in Figure 4 for the co-occurrence measure entropy derived from HH-polarized SAR data. A significant increase in separability arises with increasing kernel size. The same trend was observed for all other investigated SAR texture features except the NGLD-matrix approach

Fig. 4. Separability of low density urban areas as a function of window size - texture feature co-occurrence entropy

which showed best performance for a window size of 5x5 pixels. For the optical data, a kernel size of 11 x 11 pixels was found to be a good choice. The obtained relationship between product quality and kernel size was confirmed by the corresponding classification results. As an object-based classification strategy is followed, the loss of spatial resolution with increasing kernel size is largely compensated.

Figure 5 illustrates selected texture parameters derived from HH-polarized ASAR APP data acquired on April 22, 2005 and Landsat-5 TM data from April 21, 2005. For SAR data, the visual interpretation of the texture images as well as the achieved classification results indicated the potential of using the standard deviation computed from a 23 x 23 window size and the NGLD matrix approach to map urban areas. Misclassifications, however, between urban and forested areas do appear in regions with rough topography, especially in direct neighborhood to radar shadows and layover. Best overall performance was found for the NGLD matrix approach (Table 1). Applying a SAR speckle filter prior to image texture extraction did not improve the achieved classification results significantly, except for the NGLD matrix approach. These findings are partly in disagreement with those reported by Ndi Nyoungui et al. (2002). The authors observed an enhanced map accuracy when using filtered SAR scenes, whereby texture

features derived from 2nd- and 3rd-order histograms performed best. These conflicting results (effect of speckle suppression, choice of texture measures) may be explained in part by the different speckle filters and classification algorithms used as well as by differences in the SAR image structure with respect to the distribution of land use classes investigated. For the optical EO data best performance was achieved using the GLDV entropy of the blue channel. Especially bare soils, fields covered by sparse and low vegetation or crop residues were misclassified as residential areas. The use of the investigated optical texture parameters for urban area mapping varies significantly between seasons. None of the textural features investigated showed a high potential for all Landsat-5 TM scenes available and thus cannot be used in automated procedures. For example, the GLDV entropy calculated from the blue channel data of the Landsat-5 TM scene acquired on October 30, 2005 is not appropriate to distinguish urban and agricultural areas which results in a very low user accuracy (Table 1).

As outlined by Dell' Acqua and Gamba (2003) and Forsythe and Waters (2006), the combination of different texture features improves the quality of the urban area map. The use of SAR texture features only is recommended, as the results from optical data or optical and SAR data are less stable in time. The investigations demonstrated the suitability of the UATM information in conjunction with the standard deviation for the extraction of residential areas (Table 1). Note that only the texture measures were used for classification in order to select the parameters with the highest potential and information gain for urban area mapping. By incorporating the spectral and backscattering information in the classification process a significant improvement was achieved. When using multitemporal images the results became more accurate and stable.

Table 1. Classification accuracies of urban area maps derived based on one single texture channel only

Texture feature	TM 21.04.05 SAR 22.04.05		TM 10.07.05 SAR 10.07.05		TM 30.10.05 SAR 23.10.05	
	PA[a]	UA[b]	PA	UA	PA	UA
C-HH standard deviation	73.6	57.6	78.3	69.5	83.9	59.8
C-HH UATM	80.0	83.3	82.0	73.2	82.0	58.6
C-HH standard derivation and UATM	80.4	84.1	82.6	79.3	80.4	80.4
TM ch1 GLCM homogeneity	80.5	53.8	74.2	75.9	71.0	38.0
TM ch1 GLCM entropy	75.0	50.9	75.9	77.6	73.8	20.9
TM ch1 standard deviation	70.0	52.0	85.5	60.9	72.0	59.0

[a]*PA:* Producer's accuracy [%].
[b]*UA:* User's accuracy [%].
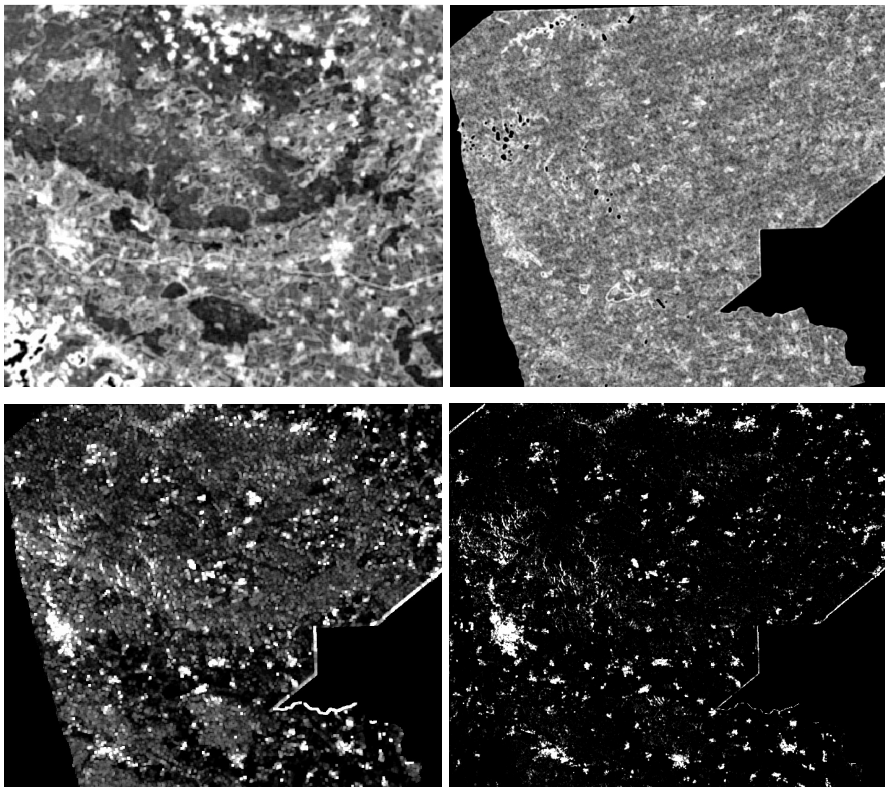
Fig. 5. Selected texture features derived from Landsat-5 TM (blue band) and HH-polarized ASAR APP data acquired in April 2005 – top: GLDV entropy for TM (left) and SAR (right); bottom: standard derivation (left) and UATM (right), SAR

4.2 Automatic selection of training samples

As outlined in chapter 3, the second step in the proposed classification chain is the automatic selection of potential training sites based on a decision tree with scene-specific thresholds. The absolute values at each node of the decision tree were estimated stepwise for every input scene by using expert knowledge in combination with histogram analyses. In Table 2 the characteristic image parameters for each land cover class are listed. As stated above, these image features were selected based on literature and an EO data library review, separability analyses, and the comparison of classification performances. The land use class "Agricultural areas" is characterized by a high spectral variability in space and time as it includes bare

soil as well as fields covered by crops in different growing stages. The application of some classification approaches such as the maximum likelihood classifier used in this study requires a Gaussian distribution of the training data. To achieve this requirement, the image segments belonging to agricultural areas were subdivided into sub-classes with similar spectral characteristics by histogram analysis of all optical input bands.

As an example, the threshold estimation process is described in detail for the land cover class "Coniferous forest". The analyses of the available time series and EO libraries all confirmed that coniferous forest areas are generally characterized by a low near (NIR) and middle infrared reflectance (MIR). Firstly, segments probably belonging to the "Water bodies" class were excluded by unselecting all image objects with a reflectance in NIR below the threshold for water plus 20 (8-bit data). In the next step, an initial threshold for coniferous forest in MIR is defined as the minimum histogram value with a segment frequency greater than five plus twenty. By this process mainly coniferous forest and - depending on growth stage - agricultural crops are selected. Both classes could be separated analyzing the corresponding histogram in the NIR channel (Fig. 6). The peak with lower reflectances represents coniferous forest segments and the higher reflectance values agricultural fields. The threshold was defined as the mean between the local maximum and the adjacent local minima considering a minimum object frequency of 5 segments and 10 % of the maximum histogram value, respectively. Applying the threshold calculated for the NIR channel, the final threshold in MIR could be computed by histogram analysis. Finally, since sometimes a small number of urban objects were selected, the mean NDVI ± 0.05 is computed and used for the selection of the final training samples of coniferous forest.

With this proposed methodology a large number of potential training sites were selected (up to 55 % of all image objects). For validation issues

Table 2. Characteristic image parameters used for the automatic selection of potential training samples from expert knowledge

Class	Image Parameter
Water bodies	NIR
Coniferous forest	NIR, MIR, NDVI
Decid. / mixed forest - leafoff	Green, NDVI, texture – HH-pol.
Decid. / mixed forest - leafon	Green, NIR, MIR, NIR - MIR
Unvegetated areas	Minimum red, HH, HV
Urban areas	NIR, MIR, texture – HH-pol., Min. HV-pol.
Grassland	NDVI, HV-pol.
Agricultural areas	Green, NIR, HV-pol., texture – HH-pol.
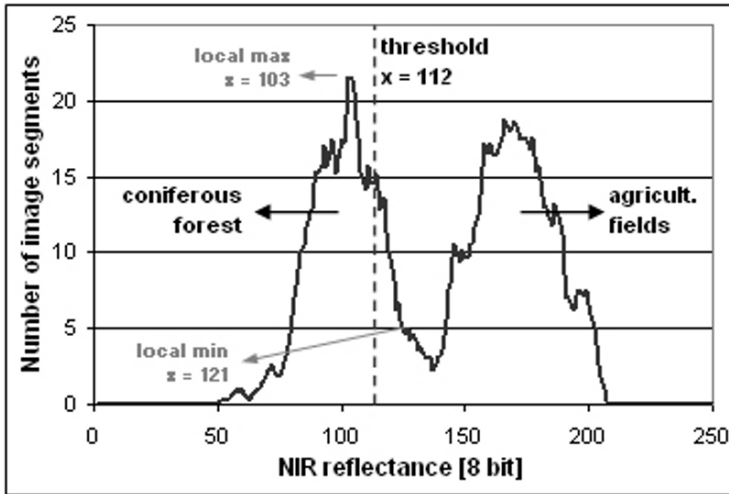
Fig. 6. Threshold estimation by histogram analysis – example coniferous forest

the algorithm was tested with six Landsat TM/ETM scenes acquired over the Thuringian test site in 2003 and 2005. The quality of the training samples selected was assessed by calculating the confusion matrix based on the same reference points used to estimate the classification accuracy of the final land cover maps. The user accuracy, i.e. the probability that a training sample is in agreement with the reference data, usually exceeds 95%. However, especially for the classes "Agriculture", "Grassland" and "Unvegetated areas" problems in finding correct training samples arise when one optical scene is available only.

4.3 Classification results

For classification three different approaches were compared considering the training sites and threshold values specified by the methodology described in the previous chapter. First, a simple nearest neighbor classification was performed using all potential training sites as input data. The second classification procedure used the estimated class thresholds to set up an object-based fuzzy classification rule. Thereby the membership function for all image parameters specified in Table 1 was set to 1 for data values within the interval defined from the estimated thresholds. For values outside the specified ranges the membership function decreases linearly. The upper and lower limit characterized by a membership function equal zero was widened stepwise until all image segments were assigned to one land

cover class. Finally, a combined pixel-based/object-based approach was applied. The potential training sites were used as input samples for a supervised pixel-based maximum likelihood classification. The finally assigned land cover class of each image object corresponds to the most frequent class per image segment. The classification accuracies obtained from the two Landsat-5 TM scenes acquired on April 21 and July 10, 2005 as well as from the multitemporal SAR data are listed in Table 3. The achieved classification accuracies differ significantly, whereby the best overall performance was found for the combined pixel-/object-based classification. Further analysis showed that the low classification accuracies for the object-based nearest neighbor classification could be partly attributed to a stronger effect of ambiguous training samples.

As outlined in the previous chapter, a large number of training sites is selected by the proposed methodology. Regarding the application of different classifiers such as artificial neural networks this is a critical issue which can result in an overtraining of the classifier and thus in reduced map accuracy (Kavzoglu & Mather 2003). The maximum likelihood classifier used in this study is based on the class mean vectors and covariance matrices estimated from the training samples. With an increasing number of training samples the mean values and the standard deviations show less variation and converge to those values obtained for all training samples. In consequence, the achieved map accuracies remain nearly unchanged by enlarging the number of training samples. To demonstrate this, classifications were performed using randomly selected subsets of all potential training data as input. The sample size per land cover category was set to 10, 20, 30, 50, 80 and 100 segments. For each sample size, 10 classification runs were conducted. The average product accuracies achieved for the different sample sizes did not exceed those obtained when using all training samples at once. This confirms that for the maximum likelihood classifier overtraining is not critical.

However, randomly selected subsets of all training samples were found to be useful in order to minimize the classification errors and to generate more stable land cover maps. The classification procedure was performed 20 times with varying training sets and a sample size of 20 segments per land cover category. The overall classification accuracy of each single map varies between 83.9 % and 91.1 %. For each of the 20 training sample sets the most frequent class as well as the corresponding class fraction were computed for all image segments. The final land cover class of each object was assigned to the most frequent class of the ten training sets characterized by the highest class fractions. The overall product accuracy achieved accounts for 90.8 %. Especially for urban areas a significant improvement of the classification results was found compared to the land cover map

generated by simultaneously using all selected potential training sites (Table 4).

Table 3. Achieved classification accuracies using all potential training samples

Class	Nearest neighbor classification		Fuzzy classification approach		Combined pixel-/object-based approach	
	PA [%]	UA [%]	PA [%]	UA [%]	PA [%]	UA [%]
Water bodies	96.0	100.0	95.8	100.0	96.0	100.0
Coniferous forest	87.8	100.0	91.3	100.0	100.0	92.5
Decid./mixed forest	96.0	88.9	100.0	89.1	92.0	95.8
Grassland	60.0	96.8	69.4	85.0	96.0	88.9
Unvegetated areas	38.0	100.0	48.9	92.0	74.0	97.4
Urban areas	87.8	76.8	75.0	83.7	84.0	72.4
Agriculture	100.0	50.5	97.7	53.8	89.6	89.6
Overall accuracy	80.6		82.5		90.2	
Kappa coefficient	0.774		0.796		0.886	

Table 4. Classification accuracies achieved by the combined pixel-/object-based approach using all potential training samples at once (a) or randomly selected training data subsets (b)

Class	Single classification / all training samples		Multiple classification / random training data subsets	
	PA [%]	UA [%]	PA [%]	UA [%]
Water bodies	96.0	100.0	92.0	100.0
Coniferous forest	100.0	92.5	100.0	96.1
Decid./mixed forest	92.0	95.8	94.0	92.2
Grassland	96.0	88.9	88.0	93.2
Unvegetated areas	74.0	97.4	82.0	89.1
Urban areas	84.0	72.4	86.0	82.7
Agriculture	89.6	89.6	93.8	83.3
Overall accuracy	90.2		90.8	
Kappa coefficient	0.886		0.892	

Postclassification procedures include the re-coding of agricultural fields misclassified as urban area. These areas are characterized by a multitemporal minimum below -19 dB in HV-polarized C-band SAR data. This procedure requires at least two cross-polarized radar scenes from the beginning of the growing season in April and shortly after the main harvest period (in this test area mid of August). Additionally, image segments (except water bodies) completely surrounded by residential areas were reclas-

sified. For example small-sized objects in industrial parks were sometimes assigned to the land cover class "Unvegetated areas".

4.4 Validation

For validation issues the proposed methodology was tested using different sets of input data including images acquired in 2003 (Table 5). The classification accuracies are listed in Table 6. Here, the obtained land cover maps are less accurate than the maps generated from multitemporal optical data (see previous chapter). If monotemporal (i.e. single date) optical data are available only, increased misclassifications occur between (1) "Unvegetated areas" and bare fields, (2) "Grassland" and agricultural crops and (3) "Grassland" or agricultural crops and "Deciduous/mixed forest". In consequence, the producer and user accuracies of the corresponding land use categories decline.

Table 5. Test image data sets used for validation

	Landsat-TM/ETM	ASAR APP IS2 – HH/HV		
		Date 1	Date2	Date3
Set 1	21.04.05	22.04.05	10.07.05	24.08.05
Set 2	10.07.05	22.04.05	14.08.05	18.09.05
Set 3	06.08.03	11.06.03	20.08.03	14.09.03
Set 4	22.04.03	16.07.03	20.08.03	14.09.03

Table 6. Classification accuracies for the four test scenarios – WA: water bodies, CF: coniferous forest, D/MF: deciduous/mixed forest, GL: grassland, UA: unvegetated areas, UR: urban areas, AG: agriculture

	Set1		Set2		Set3		Set4	
	PA [%]	UA [%]	PA [%]	UA [%]	PA [%]	UA [%]	PA [%]	UA [%]
WA	96.0	96.0	94.0	100.0	98.0	100.0	96.0	98.0
CF	100.0	94.2	100.0	94.2	95.9	92.2	95.4	95.4
D/MF	94.0	97.9	92.0	79.3	92.0	93.9	96.0	92.3
GL	56.0	93.3	54.0	87.1	92.0	76.7	84.0	63.6
UA	68.0	94.4	62.0	81.6	72.0	83.7	64.0	94.1
UR	78.0	86.7	82.0	74.6	80.0	81.6	83.0	78.0
AG	97.9	54.7	91.7	66.7	81.3	84.8	58.3	63.6
OA[a]	84.2		82.1		87.3		82.3	
Kappa	0.81		0.79		0.85		0.80	

[a]OA: Overall accuracy.

5 Discussion

The objective of this study was to develop a classification procedure for the generation of basic land cover maps with a high potential for automation. In order to achieve this, the necessary training samples, required as input for any supervised classification, were selected automatically. However, a requirement for this approach is the presence of coniferous forest stands in the imaged area. Furthermore, the optical data must include a channel in the MIR range of the electromagnetic spectrum. Since the thresholds of the decision tree used to delineate potential training sites are flexible and estimated for each input scene separately, an atmospheric correction of the optical EO data is not necessarily required.

The proposed classification procedure combines the complementary information provided by optical and SAR data. For both steps, the selection of potential training samples and as input layer for the supervised classification, radar scenes are required. Including the C-band texture and backscatter information during supervised classification, the overall accuracy of the final land cover map increases by approximately 3 %. A significant improvement was specifically achieved for the classes "Urban areas" and "Grassland" as well as for forest stands.

Generally, the quality of the final land cover map significantly depends on the available EO data. Stable classification accuracies were only obtained for the land cover classes "Water bodies", "Deciduous/mixed forest", "Coniferous forest" and "Urban areas". Regarding the SAR data, best acquisition dates for the HH/HV-polarized ASAR APP data are at the early beginning of the growing season and after the main harvest and tillage period. Bare fields and fields covered by crops in an early growing stage are characterized by low radar backscatter values. Thus, using the multitemporal minimum in HV-polarisation will reduce misclassifications between agricultural fields/grassland and urban and forested areas, respectively. To map "Unvegetated areas" and "Grassland", multitemporal optical data are required. Optimal acquisition dates are the beginning and mid of growing season as well as the time span after the main harvest period. The first and the last acquisition date is needed to improve grassland mapping, whereas optical scenes during full crops development are suitable to map permanently unvegetated areas.

6 Conclusions and Outlook

A combined object-/pixel-based classification scheme for the generation of basic land cover maps providing a high potential for automation has been presented. Since often multispectral data are not available due to cloudy weather conditions, emphasis of this study was to develop a working flow including a limited number of optical data only but exploiting the information content of multitemporal SAR data. The proposed classification procedure is divided into three main processing stages. In a first step image segments are delineated on base from optical EO data. Then, potential training sites will be selected using an object-based decision tree with flexible thresholds. The absolute threshold values are calculated for each EO scene separately by combining expert knowledge and histogram analyses. The selected training samples were used as input to a supervised, pixel-based maximum-likelihood classification. The final land cover class of each image segment corresponds to the most frequent class of the pixel-based classification result. Textural features were incorporated in the classification procedure in order to improve the mapping accuracy for urban areas. The use of SAR textural features is recommended as the results obtained from optical data or combined optical and SAR data are less stable in time. This research demonstrated the suitability of the UATM information in conjunction with the standard deviation to extract successfully residential areas. For validation issues the methodology was applied to different sets of input data acquired over the Nordhausen test site, Thuringia, in 2003 and 2005. The proposed processing chain requires the existence of coniferous forests in the image area. Regarding the SAR data, the use of at least two Envisat ASAR APP scenes in HH/HV-polarization acquired at the beginning of the growing season in April and after the main harvest and tillage period is recommended.

The algorithm was developed for a project area in Northern Thuringia and successfully tested with EO data acquired in different seasons and years. Especially for developing countries and large inaccessible areas an urgent need for reliable and up-to-date basic land cover products exists. The proposed methodology could be transferred to other regions by an adaptation of the desired land cover classes and the corresponding characteristic image parameters used for the selection of potential training samples (expert knowledge). The implementation of these adaptations is rather unproblematic as main processing steps were realized in IDL. Furthermore, the proposed methodology could be applied to other topics of interest. For example, at the Nordhausen test site the mapping of rape seed fields with a high accuracy was possible. Other useful adjustments could imply the in-

tegration of satellite data acquired by different sensors or the application of enhanced supervised classification approaches such as artificial neural networks or support vector machines (see Tzotsos and Argialas in this book).

Further improvement is expected for both, the automatic selection of potential training samples as well as the final map accuracy, with the availability of polarimetric and high-resolution spaceborne X-and L-band SAR data as provided by the TerraSAR-X and ALOS PALSAR missions. For example, L-band SAR data are known to provide an excellent database for forest cover mapping, i.e. the misclassification between forested areas, agricultural crops and grassland is expected to decrease (see Thiel et al. in this book). Regarding urban area mapping several studies emphasized the potential of X- und L-band data. Furthermore, due to the very high spatial resolution, texture measures extracted from TerraSAR-X and PALSAR data will most probably improve the generation of urban area maps.

Acknowledgement

References

Alparone L, Baronti S, Garzelli A, Nencini F (2004) Landsat ETM+ and SAR image fusion based on generalized intensity modulation. IEEE Transactions on Geoscience and Remote Sensing, 42(12): 2832 – 2839

Amarsaikhan D, Douglas T (2004) Data fusion and multisource image classification. International Journal of Remote Sensing 25(17): 3529-3539

Dekker RJ (2003) Texture analysis and classification of ERS SAR images for map updating of urban areas in The Netherlands. IEEE Transactions on Geoscience and Remote Sensing, 41(9):1950 – 1958

Dell' Acqua F, Gamba P (2003) Texture-based characterization of urban environments on satellite SAR images. IEEE Transactions on Geoscience and Remote Sensing, 41(1): 153 – 159

Forsythe KW, Waters NM (2006) The utilization of image texture measures in urban change detection. PFG 4: 287 – 296

Franke KH (2006) Grundlagen der digitalen Bildverarbeitung und Mustererkennung. Schriftenreihe des ZBS e.V., ISSN 1432-3346, Ilmenau

Franklin SE, Wulder MA (2002) Remote sensing methods in medium spatial satellite data land cover classification of large areas. Progress in Physical Geography, 26(2): 173 – 205

Fukanaga K (1990) Introduction to statistical pattern recognition. Academic Press, San Diego

Hegarat-Mascle SL, Quesney A, Vidal-Madjar D, Taconet O, Normand M, Loumagne C (2000) Land cover discrimination from multitemporal ERS images and multispectral Landsat images: a study case in an agricultural area in France. International Journal of Remote Sensing, 21: 435 – 456

Jensen JR (2000) Remote sensing of the environment – an earth resource perspective. Prentice Hall, New Jersey

Kavzoglu T, Mather PM (2003) The use of backpropagating artificial neural networks in land cover classification. . International Journal of Remote Sensing, 24: 4907 – 4938

Ndi Nyoungui AD, Tonye E, Akono A (2002) Evaluation of speckle filtering and texture analysis methods for land cover classification from SAR images. International Journal of Remote Sensing, 23(9): 1895 – 1925

Pohl C, Van Genderen JL (1998) Multisensor image fusion in remote sensing: concepts, methods and applications. International Journal of Remote Sensing, 19(5): 823 – 854

Richards JA (1999) Remote sensing digital image analysis. Springer, Berlin

Stussi N, Beaudoin A, Castel T, Gigord P (1995) Radiometric correction of multi-configuration spaceborne SAR data over hilly terrain. Proc. Int. Symp. on Retrieval of Bio- and Geophysical Parameters from SAR Data for Land Applications, Toulouse, France, 10-13 October 1995, pp. 469-478

Swain PH, Davis SM (1978) Remote sensing: the quantitative approach. McGraw Hill Book Company, New York

Chapter 5.4

The development of integrated object-based analysis of EO data within UK national land cover products

G.M. Smith

Section for Earth Observation, Centre for Ecology and Hydrology, Monks Wood, Abbots Ripton, Huntingdon. PE28 2LS United Kingdom.
gesm@ceh.ac.uk

KEYWORDS: Land cover, National, Object-based, Generalised cartography

ABSTRACT: The United Kingdom (UK) undertakes regular surveys of its countryside which are accompanied by national land cover maps derived from Earth Observation data which have exploited the leading edge analysis methods of the day. The Land Cover Map of Great Britain of 1990 was a relatively simple pixel-based classification while the Land Cover Map 2000 adopted an object-based approach. The objects, or land parcels, were derived by automated segmentation of the input image data and had a minimum mapping unit of 0.5 ha. Both of the above land cover products have been extremely successful, with in excess of 300 users. There have of course been problems with these products and these are mainly associated with the data models which were somewhat abstract from reality. Preparations are now underway for a further update of the UK national land cover product which will again be object-based, but this time it is planned that digital cartography will be adapted to give an object structure which more accurately reflect the true structure of the landscape. A feasibility study has demonstrated the key processes required to achieve the generalisation. The use of such a spatial structure will deliver a world leading land cover product which will increase the potential user community and possibilities for integration.

1 Background

The United Kingdom (UK) undertakes an assessment of its landscape at intervals of 8 to 10 years known as the Countryside Survey (CS) (Haines-Young et al. 2000). The main component of the CS in 2000 was a field survey where approximately 560 1 km squares were visited for detailed ground-based measurements. The last two CSs have been accompanied by national land cover maps derived from Earth Observation (EO) data. These maps have developed over time exploiting leading edge analysis methods while maintaining a focus on the operational requirements of a national mapping exercise.

The first of these, the Land Cover Map of Great Britain (LCMGB) in 1990, was a relatively simple pixel-based classification using Landsat Thematic Mapper (TM) data (Fuller et al. 1994). Multi-temporal TM data were used to maximise the amount of land cover discrimination that could be achieved, as certain land covers change on a seasonal basis. The data were classified with a conventional per-pixel implementation of a maximum likelihood algorithm and low level knowledge-based corrections were applied using simple raster masks.

This chapter reviews the development of object-based approaches to land cover mapping associated with the UK land cover products between the early 1990s and mid-2000s. In so doing, it reports the key issues, advances and outputs while referring the reader to the appropriate literature for a more in depth analysis.

2 Object-based land cover mapping

The LCMGB was extremely successful, but the pixel-based approach imposed an arbitrary grid pattern on the product and its depiction of the landscape. The landscape is not divided into a grid of square cells equivalent to image pixels and thus this sampling scheme failed to address the actual landscape structure. The pixel-based approaches also incorporated noise and unwanted natural variation into the classification output resulting in a product with a speckled, 'salt and pepper', appearance and with little if any information on the relationships between land cover patches (Fuller et al. 1998). This situation encouraged the team responsible for the UK national land cover maps to develop object-based approaches which analysed the EO data in units representative of real world features.

The Classification of Environment with Vector and Raster-Mapping (CLEVER-Mapping) project in the late 1990s developed an object-based

classification procedure (Figure 1). The procedure relied on the availability of some form of land parcel data set which subdivided the area of interest into a series of relatively homogenous land cover units and thus provided a landscape structure. The speckled appearance of conventional per-pixel approaches was partly caused by mixed pixels at the edge of each land cover patch which were often misclassified due to their mixed spectral signatures from adjacent land cover types. The object-based approach avoided these mixed pixels by shrinking the land parcel geometry, or boundary, and only considering the core pixels for further analysis (Dean and Smith 2003). The spectral response in each image band was averaged for the core pixels only within each object to minimise noise and unwanted natural variation. The averaged spectral responses for the objects were then applied to a standard maximum likelihood algorithm and the resulting classification attached to the object as a whole. The exclusion of edge pixels and the use of average spectral response for the classification was found to increase the confidence of the results (Dean and Smith 2003).

Attribute	Value
Parcel ID	Field 1
Area	78.4
Perimeter	30.3
Elevation	110.0
Soil type	Clay
Ownership	Farm
Blue resp	30.8
Green resp	55.3
Red resp	23.5

- - - Parcel boundary

   ■ Pure pixels
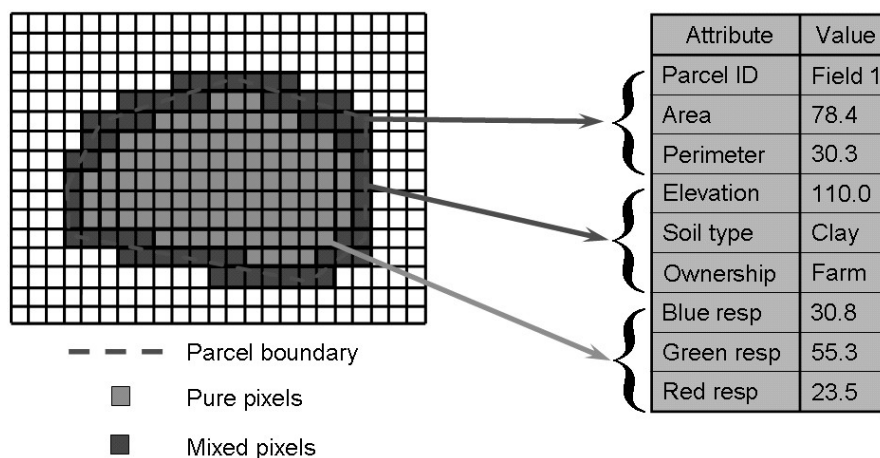
   ■ Mixed pixels

Fig. 1. Schematic of the object-based classification approach developed during the CLEVER-Mapping project. The land parcel boundary is used to exclude the mixed edge pixels by extracting spectral information from the core area only.

By obtaining the land parcel objects from a data source other than the EO data used for classification, the area sampling and spatial structure of the land parcels were unrelated to the grid pattern and spatial resolution of the EO data. The use of an independent object-based structure then allowed different EO data types to be combined and a broad range of non-EO data to be included as attributes on the land parcels. The later were

used to perform complex knowledge-based enhancements (Fuller et al. 2004). For instance, there is often spectral confusion between coastal saltmarsh and upland shrub habitats due to their similar vegetation structure. Therefore land parcels classified as saltmarsh, but with elevations greater than a few metres derived from a digital elevation model could be reclassified to upland shrub and visa versa. In the UK there is a requirement to subdivide semi-natural grasslands by the substrate on which they are found. These distinctions cannot be made reliably from EO data alone therefore soil type data sets could be used to refine the semi-natural grassland types recorded. The topologically structured objects also allowed advanced spatial context enhancements to be applied. For example, it is likely that small patches of arable completely surrounded urban are incorrect and bare ground in a coniferous forest context is more likely to be felled forest. By assessing the spatial context of each land parcel these confusions can be corrected.

2.1 Land Cover Map 2000

An update of LCMGB was produced between 1998 and 2001, referred to as Land cover Map 2000 (LCM2000), which adopted an object-based approach (Fuller et al. 2002a) and initially it was intended to fully implement the above approach. Unfortunately, a suitable object, land parcel, data set was not available nationally in the UK at the time of production so image segmentation procedures were applied to the EO data to generate a set of objects, or segments. The EO data again consisted of multi-date imagery, but this time from the Landsat 5 TM, the Landsat 7 Enhanced Thematic Mapper and the Indian Research Satellite Linear Imaging Self-Scanner-III. The resulting segments, or spectrally-based land parcels, related in the main to fields, woods, lakes etc. and had a minimum mapping unit (MMU) of 0.5 ha and a minimum feature width (MFW) of 25 m (Figure 2). The image segmentation had been particularly challenging in that it had to provide a reasonably consistent subdivision of the landscape over areas in excess of 150 km and between EO data sets that could have considerable differences (timing of summer winter composites, combinations of sensors). The final data set was deemed suitable for the task at hand and contained 6.6 million objects covering the ~240 000 km^2 of the UK (Fuller et al. 2002b).
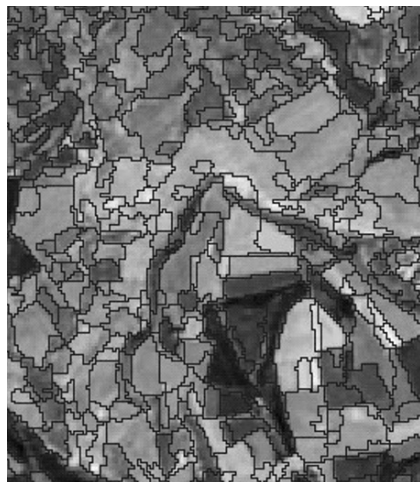
Fig. 2. A 3.5 by 4.0 km example of the object-based structure of LCM2000 derived from image segmentation compared to a Landsat TM false colour (red: band 4, green: band 5 & blue: band 3) composite image

The LCM2000 land parcels were classified (Figure 3) using procedures similar to those developed during CLEVER-Mapping, except this time landscape structure and the spectral information for the thematic classification came form the same EO data source. This approach generates relatively homogeneous segments, but the segments only represent aspects of landscape structure which can be detected by EO data. The maximum likelihood classification and knowledge-based enhancements mapped the widespread examples of the UK Biodiversity Action Plan Broad Habitats (Jackson, 2000). Each land parcel carried a rich set of parcel-level metadata as well as the resulting Broad Habitat label as part of a hierarchical land cover class scheme with up to 72 detailed classes (Fuller et al. 2002b; Fuller et al. 2005).
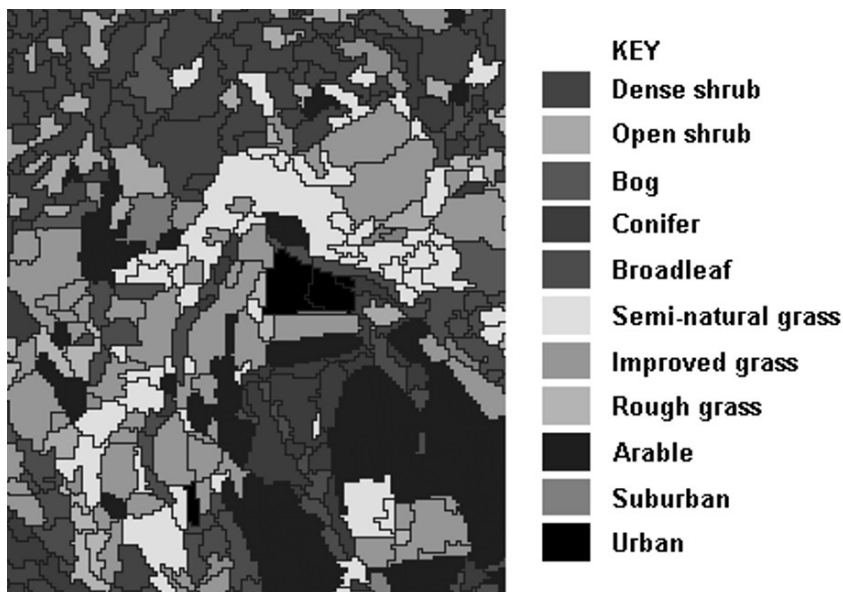
Fig. 3. A 3.5 by 4.0 km example of the completed LCM2000 map labelled with a simplified version of the UK Broad Habitats

2.2 Issues with previous products

Both of the land cover products described above have been extremely successful, with in excess of 300 user for each one in all sectors from academic research to commercial consultancy. They have been used in a broad range of applications from studies of national carbon budgets and the distribution of species to the locating of telecommunications equipment.

There have of course been problems and criticisms with these products considering the broad user community they aimed to address. One of the main criticisms was associated with the data models which were somewhat abstract from reality.

The image pixel-based spatial model used in the LCMGB was an arbitrary grid and thus unrelated to the actual landscape structure. The segment-based spectral land parcels of LCM2000, although an improvement on the LCMGB, reflected the spectral structure of the landscape rather than the presence of true boundary features related to the habitats that were to be mapped. For instance, in LCM2000, two adjacent wheat fields could be combined into a single object by the segmentation process even though they may be owned by different farmers or managed differently in subse-

quent years. Even if a boundary feature existed between them it would need to be spatially and spectrally significant at the spatial resolution of the image data to cause the segmentation algorithm to initiate a new object. Conversely, single fields may contain natural and acceptable variability which causes the segmentation algorithm to erroneously initiate a new object, giving multiple objects per field. For instance, a crop may progressively come into flower across a field and over time and the pattern of flowering could be captured by the image data and then recorded as spurious objects. In some instances, even quite major boundaries, such as rivers, were not fully recognized by the segmentation algorithm as the presence of bridges could cause the resulting segments to bleed into adjacent areas. Finally, the pixelated nature of the objects was also found to cause problems when comparing other data sets which represented linear diagonal boundaries in a more conventional manner.

The above considerations suggested that a new approach to the creation of the objects, land parcels, or spatial framework should be developed which avoided the use of image data as far as possible and aimed to capture the true landscape structure as real world objects.

2.3 Development of real world object approach

The real world object approach would derive the majority of its landscape structure from the use of existing digital cartography which had been captured by organization such as national mapping agencies. This approach had already been developed partially, as during the CLEVER-Mapping project a large scale prototype mapping exercise had been undertaken by producing a land cover map for the island of Jersey in 1997 (Smith and Fuller, 2001). The island government had digital cartography available for an area of approximately 215 km^2, but this was too detailed to integrate with standard EO data sets with a spatial resolution of around 25 m (Figure 4) and only formatted as individual lines, with not true land parcels present. It was therefore necessary to first generalise the digital cartography to a point where the boundaries it contained would represent objects mappable from the EO data. For instance, roads were often composed of a centre line, two edge lines for the metalled surface and possibly also further lines for boundary features such as hedges, ditches, walls or pavements. From this set of near parallel linear features it was only necessary to keep the road centerline, but unfortunately such simple rules could not be automated and applied across the whole data set as a large number of feature combinations could indicate a boundary. Finally, the lines remaining after generalization were built into the required land parcels before the

object-based classification could be applied. Unfortunately, at the time, the only means creating the land parcels was to perform a generalisation by manually editing the line work and using standard polygon building functionality to produce objects from the disconnected lines.
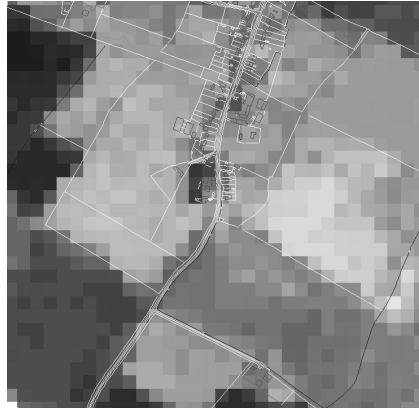
Fig. 4. An 800 m by 800 m example of detailed digital cartography (vectors) overlaid on a 25 m spatial resolution Landsat TM false colour (red: band 4, green: band 5 & blue: band 3) composite image

The resulting product (Figure 5) was of exceptional spatial quality compared to pixel-based and segment-based equivalents. The thematic accuracy was improved above that of the pixel-based approach (Table 1) with an increase correspondence to independent validation data of around 20 % (Smith and Fuller 2001). The relationship of the land parcels to existing cartography improved the usability and opportunities for integration with other data sets and within existing business systems. For instance, it was possible to link the land cover data set to the States of Jersey agricultural census as each field had a centroid which tied the census information to a geographical location. Unfortunately, the process to build the land parcels took around 2 person months for 215 km^2 and was therefore impractical for larger areas such as the UK and thus not a viable solution for the need to improve the land parcel structure of the national land cover products.
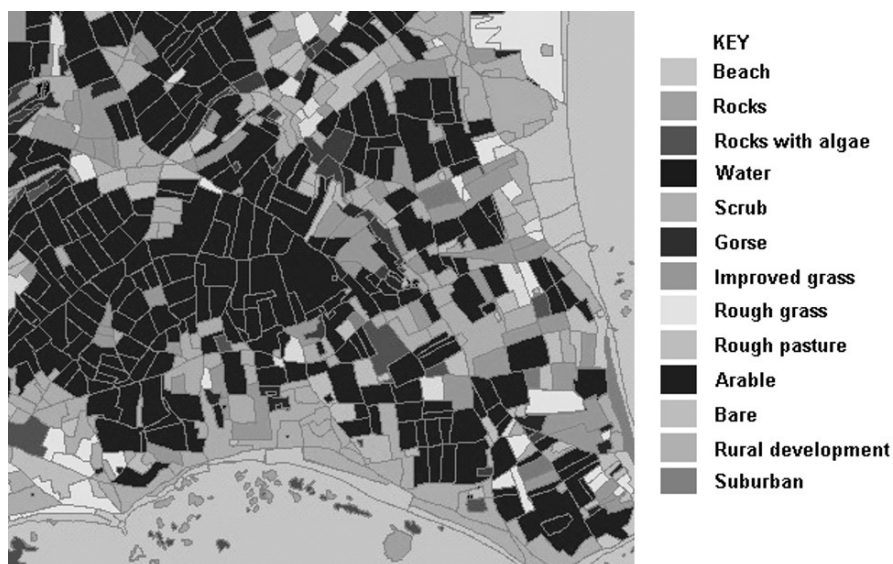
Fig. 5. A 2 km by 2 km section of the 1997 Land Cover Map of Jersey

Table 1. Values of percentage correspondence for the 1997 Land Cover Map of Jersey showing the improvement of the CLEVER-Mapping approach and the use of knowledge-based enhancements.

	Classification procedure	Percentage correspondence
Conventional	Per-pixel	61
CLEVER-	Per-parcel	79
Mapping	Plus knowledge-based enhancements	85

Preparations are now underway for a further update of the UK national land cover product with a target summer of 2007. This product will again be object-based, but this time it is hoped that digital cartography can be adapted to give an object structure which more accurately reflect the true structure of the landscape. The realization of this aim relies on developments in both digital cartography and spatial analysis technology.

2.4 LCM2007 feasibility study

Since the release of LCM2000, the Ordnance Survey (OS) of Great Britain, the national mapping agency, have produced a digital cartography product called MasterMap (MM) (OS, 2007) by topologically structuring

existing digital line work (similar to the data used in Jersey). The structuring of the data produces land parcels / real world objects rather than disconnected line work. This dataset is still far too detailed for effective integration with EO data with a 25 m spatial resolution, but is suited to automated generalization. It is therefore proposed to base the spatial structure of the next UK land cover product on a spatially generalized version of MM.

It was also necessary to develop an automated approach to the generalisation, so that large areas could be processed cost effectively and in a timely manner. A feasibility study was undertaken with the help of a spatial database technology company, Laser Scan (now called 1Spatial) of Cambridge, UK. The study developed and demonstrated the key processes required to achieve the generalisation. The specification for the generalised MM was based on the LCM2000 spatial specification of 0.5 ha MMU and 25 m MFW. The first step of the generalisation was to classify the MM objects by their geometric characteristics (Figure 6) into the following categories:

- A) objects are less than the MMU and simple;
- B) objects are less than the MMU but complex (e.g. fail MFW rules);
- C) objects are larger than the MMU but complex;
- D) objects are larger than the MMU and MFW but elongated; and
- E) objects are larger than the MMU and simple.

Using this geometric classification scheme as a guide the objects are merged and split iteratively until the data set only contains objects that are larger than the MMU and with relatively simple shapes which fit the specification (category E). Figure 7 shows an example of the generalised MM data compared to an aerial photograph and Figure 8 shows it compared to a 25 m spatial resolution satellite image. An assessment of the quality and utility of the results was carried out by aerial photograph interpreters who digitized test areas of the aerial photography independently and then visually compared the results. This qualitative assessment confirmed the utility of the generalised MM land parcels, particularly as they were to be used with 25 m spatial resolution EO data, not aerial photography. There were a few minor errors and ambiguities, but these were either to be corrected in the next version of the procedure or highlighted for an operator to correct manually. In comparison with the satellite image data it can be seen that the generalised MM is fully aligned with the needs of an object-based analysis procedure. The area of woodland shows where the MM may lack some important boundaries and in the arable landscape where farm practices, rather than actual boundary features dictate the pattern of crops this

is also a problem. Finally, in the uplands semi-natural habitat boundaries are not always mapped by national mapping agencies. These additional boundary features can be obtained from either external digital cartography data sets (e.g. forestry maps or agricultural land parcel data sets) or by within object segmentation.
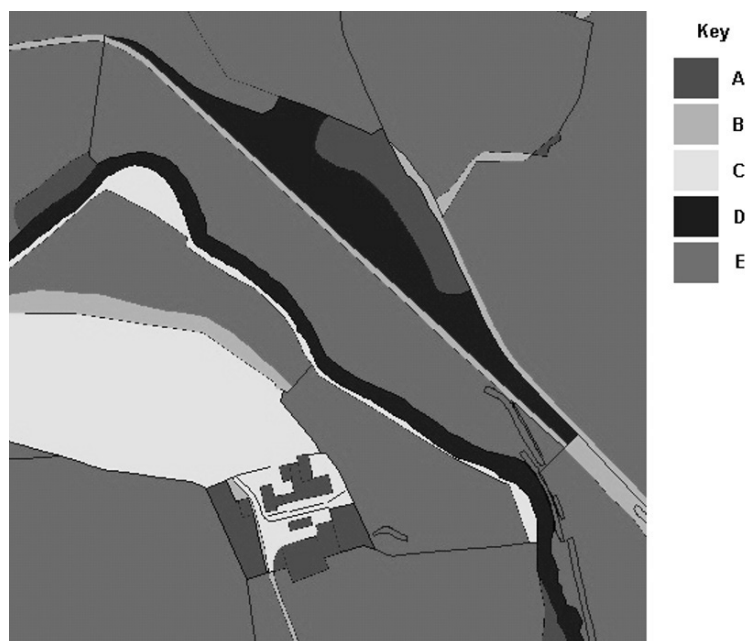
Fig. 6. A 1 km tile of OS MasterMap objects classified by their geometric characteristics

Fig. 7. A 1 km tile of generalised OS MasterMap data (dark lines, compare with Figure 8) compared to true colour aerial photography.
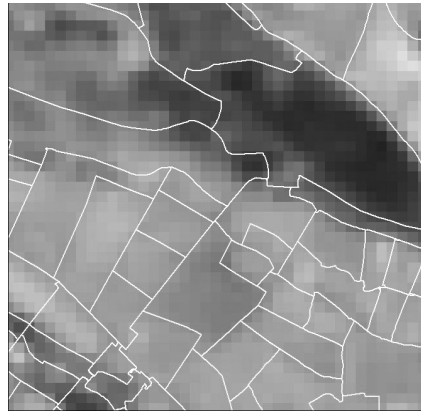
Fig. 8. A 1km tile of generalised OS MasterMap data (white lines) compared to a 25 m spatial resolution Landsat TM false colour (red: band 4, green: band 5 & blue: band 3) composite image

The work reported here was undertaken on 400 km^2 area centred on the town of Ripon in the UK which included semi-natural, rural and urban landscapes, but the procedure has since been rolled out to areas in excess of 3600 km^2. The generalisation approach appears to work well in all landscapes and has been tested at other sites across the UK in preparation for LCM2007 production. The generalisation has also been assessed in terms of its scalability and it can be used in a parallel approach on cluster computer systems by dividing large areas into more manageable chunks. It has been established that the approach can be applied to the whole of the UK

in an automated fashion and reasonable timescale to provide nationally consistent generalised digital cartography to support land cover mapping with EO data.

The results obtained when classifying the generalised MM (Figure 9) are very similar to those achieved in Jersey. Real world objects are clearly being mapped and can be interpreted easily in their landscape context and against other data sets. When compared to the LCM2000 data for the same area the likely improvements in quality and usability are obvious (Figure 10). Although the LCM2000 is recording the general land cover structure of the areas reasonably well, delineation of individual land parcels is poor, especially in the small fields in the southern half of this area. The segmentation algorithm has caused real world objects to bleed into each other, most noticeably in the north east where the arable fields have bled into the woodland.

Fig. 9. A 1 km tile of generalised OS MasterMap data classified as Broad Habitats (the same scheme as used in LCM2000).
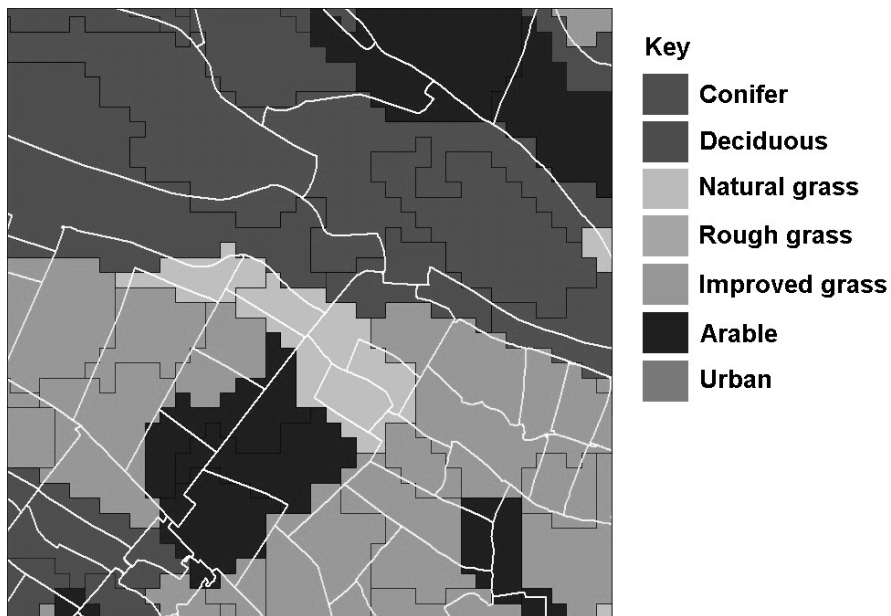
Fig. 10. A 1 km tile of generalised OS MasterMap data (white lines) compared to LCM2000 data (dark pixilated lines) of the same area

3 Summary

This paper described the background to the UK land cover products and the developments that have kept them at the leading edge of integrated object-based analysis of EO data. Through this work it has been possible to develop an approach to land cover mapping which has been flexible and adaptable as new datasets and technologies come on line. The object-based classification procedure can be summarized into following components:

- Obtain a suitable set of land parcel objects (by image segmentation and / or from digital cartography generalisation).
- Extract EO data based on the land parcels using a shrunken geometry.
- Perform a mutli-spectral analysis / classification on the extracted EO data.
- Write results back to the original land parcel.
- Attach ancillary data sets where necessary.
- Apply knowledge-based enhancements.
- Finalise land parcel level results and meta data.

This work toward the production of LCM2007 will represent a major step forward for object-based land cover mapping. The use of a spatial structure based on generalised digital cartography makes a direct relationship to real world objects and breaks the link to the transient and often 'fuzzy' EO data based image segments. These developments will increase the potential user community and possibilities for integration of the next UK land cover product.

4. References

Dean AM, Smith GM (2003) An evaluation of per-parcel land cover mapping using maximum likelihood class probabilities. International Journal of Remote Sensing 24: 2905-2920.

Fuller RM, Cox R, Clarke RT, Rothery P, Hill RA, Smith GM, Thomson AG, Brown NJ, Howard DC, Stott, AP (2005) The UK Land Cover Map 2000: planning, construction and calibration of a remotely sensed, user-oriented map of broad habitats. International Journal of Applied Earth Observation and Geoinformation 7: 202-216.

Fuller RM, Groom GB, Jones AR (1994) The Land-Cover Map of Great-Britain - an Automated Classification of Landsat Thematic Mapper Data. Photogrammetric Engineering and Remote Sensing 60: 553-562.

Fuller RM, Smith GM, Sanderson JM, Hill RA, Thomson AG (2002a) The UK Land Cover Map 2000: Construction of a parcel-based vector map from satellite images. Cartographic Journal 39: 15-25.

Fuller RM, Smith GM, Sanderson JM, Hill RA, Thomson AG, Cox R, Brown NJ, Clarke RT, Rothery P, Gerard FF (2002b) Countryside Survey 2000 Module 7: Land Cover Map 2000 Final Report. At http://www.cs2000.org.uk/Final_reports/M07_final_report.htm [10/08/07].

Fuller RM, Smith GM, Thomson AG (2004) Contextual Analyses of Remotely Sensed Images for the Operational Classification of Land Cover in United Kingdom. In: de Jong SM, van der Meer FD (eds) Remote Sensing and Image Analysis: Including the Spatial Domain, Kluwer Academic Publishers, The Netherlands, pp271-290.

Fuller RM Wyatt BK, Barr CJ (1998) Countryside Survey form ground and space: different perspectives, complimentary results. Journal of Environmental Management 54: 1001-126.

Haines-Young RH, Barr CJ, Black HIJ, Briggs DJ, Bunce RGH, Clarke RT, Cooper A, Dawson FH, Firbank LG, Fuller RM, Furse MT, Gillespie MK, Hill R, Hornung M, Howard DC, McCann T, Morecroft MD, Petit S, Sier ARJ, Smart SM, Smith GM, Stott AP, Stuart RC, Watkins, JW (2000) Accounting for nature: assessing habitats in the UK countryside, DETR, London ISBN 1 85112 460 8

Jackson DL (2000) JNCC Report No. 307 Guidance on the interpretation of the Biodiversity Broad Habitat Classification (terrestrial and freshwater types): definitions and the relationships with other habitat classifications. Joint Nature Conservation Committee, Peterborough, UK.

OS   (2007)   OS   MasterMap   Topography   Layer   User   guide.   At http://www.ordnancesurvey.co.uk/products/osmastermap/userguides/docs/OS MMTopoLayerUserGuide.pdf [10/08/07].

Smith GM, Fuller RM (2001) An integrated approach to land cover classification: an example in the Island of Jersey. International Journal of Remote Sensing 22: 3123-3142.

5. Acknowledgements

Section 6

Automated classification, mapping and updating: urban applications

Chapter 6.1

Detecting informal settlements from QuickBird data in Rio de Janeiro using an object-based approach

P. Hofmann[1], J. Strobl[1], T. Blaschke[1], H. Kux[2]

[1] Centre for Geoinformatics, Salzburg University, Austria,
hofmann@ipi.uni-hannover.de, josef.strobl@sbg.ac.at,
thomas.blaschke@sbg.ac.at

[2]Divisão de Sensoriamento Remoto, INPE, Brazil
hermann@ltid.inpe.br

ABSTRACT: Informal settlements behave very dynamical over space and time and the number of people living in such housing areas is growing worldwide. The reasons for this dynamical behavior are manifold and are not matter of this article. Nevertheless, informal settlements represent a status quo of housing and living conditions which is from a humanitarian point of view in the most cases below acceptable levels. Therefore, reliable spatial information about informal settlements is vital for any actions of improvement of these living conditions. Since remote sensing data is a well suited data source for mapping and monitoring we demonstrate a methodology to detect informal settlements (favelas) from QuickBird data using an object-based approach. On the one hand we therefore use experiences and adapt them which were already made by Hofmann, P. (2001) regarding the image segmentation of an IKONOS scene of Cape Town. On the other hand we resort to a general ontology of informal settlements which we then transfer to a fuzzy-logic rule base which acts as basic classifier of the generated segments. This basic rule base is than extended in a way that features of segregation given by the ontology (namely neighbor-

hood) are applied to the extraction method as an iterative process (i.e. a knowledge based region growing). Finally, we assess the results of the simple and iterative method by comparing them with the results of a manual mapping.

1 Introduction

Informal settlements are usually a phenomenon which mostly occurs in developing and newly industrializing countries. Although different definitions of *informal settlement* do exist, *slum, favella, squatter settlement* or *shanty town* are commonly used synonyms for this special type of settlement. Sub-standard sanitary situations and high crime rates are only a few of attributes which go aside with the phenomenon informal settlement. Nevertheless, the UN (UNSTAT 2005) define *informal settlements* as:

„1. areas where groups of housing units have been constructed on land that the occupants have no legal claim to, or occupy illegally; 2. unplanned settlements and areas where housing is not in compliance with current planning and building regulations (unauthorized housing)."

Both definitions are obviously emphasizing the illegal character of informal settlements. In contrast, the definition of Mason and Fraser (1998) takes the environmental, socio-economic and living conditions more into account. They describe *informal settlements* as:

"... dense settlements comprising communities housed in self-constructed shelters under conditions of informal or traditional land tenure ... . They are a common feature of developing countries and are typically the product of an urgent need for shelter by the urban poor. As such they are characterised by a dense proliferation of small, makeshift shelters built from diverse materials (such as plastic, tin sheeting and wooden planks), by degradation of the local ecosystem (for example, erosion and poor water quality and sanitation) and by severe social problems."

However, there is obviously a strong need to transform informal into formal settlements and to gain more control about the actual spatial development of informal settlements: According to UN-HABITAT (2006a) the number of people living in slums, favellas or shanty towns worldwide will grow from approx. 1.0 Billion in 2005 to 1.2 Billion in 2010 and 1.5 Billion in 2020. From the perspective of an urban or regional planner, as well as from the perspective of local or regional authorities, informal settlements might become a more and more challenging problem in the years to come. One reason is the rising number of people living in informal settle-

ments. Another is the threats that go aside with unplanned settlement, such as the destruction of nature reserves, contamination of soil or the settlement in areas with a high risk of natural hazards. Programs, such as the *Global Campaign For Secure Tenure* (UN-HABITAT 2006) are emphasizing that informal settlements are becoming a challenge.

In order to face these challenges, obtaining up-to-date spatial information about informal settlements is vital for any (re-)actions in terms of urban or regional planning. But due to their informal character, reliable and accurate data about informal settlements and their inhabitants is rarely available. Data sources, such as maps, statistics or even GIS data are usually obsolete, not available, not as accurate as needed or do not hold the information needed. Obtaining the data needed by surveying is expensive and time consuming. Because of the very dynamical spatio-temporal behavior of informal settlements the spatial information mapped by surveying might be already obsolete when it is consolidated. Consequently, reliable procedures for detecting and monitoring the spatial behavior of informal settlements are needed. Methods of remote sensing and image analysis can certainly contribute to an enhanced process of information acquisition.

From a methodological point of view, the challenge of detecting informal settlements lies in having appropriate methods to detect and monitor their spatio-temporal behavior reliably (Lemma et. al. 2005, Mason et al 1998, Dare and Fraser 2001, Kuffer 2003, Radnaabazar et al. 2004). Regarding available data sources, remotely sensed imagery from satellites in comparison to other data sources is of advantage, because:

- Remote sensing data in principle show a true image of the spatial situation on the ground in almost real-time.
- Satellite images are taken within constant time intervals. The length of intervals varies depending on the repetition rate of the platform, or other technical specifications (i.e. side-looking-modes and so on).

A disadvantage of remote sensing data lies in the complexity of methods necessary to extract the spatial information needed, which is directly or indirectly imaged in the data. In many cases experts in remote sensing, image processing and image analysis are needed to extract this information. Hence, in order to benefit from the advantages of remote sensing and to obtain the information as needed, adequate methods for analyzing the imagery are necessary. In an ideal case, these methods can be applied without the need of expert knowledge and human interaction. In practice easiness of use and the degree of automation for information extraction from im-

agery depends on the data used and the phenomena to be extracted from the image.

2 Methods and general methodologies

Because of the relatively high inner-structural heterogeneity of informal settlements their typical pattern in remote sensing data is in many cases hard describable. In general, this hampers the generation of an automated detection process which is easy to use. Nevertheless, in (Hofmann 2001) first results of detecting informal settlements from IKONOS data in Cape Town showed the feasibilities by using an object-oriented approach in principle. The results were promising but seamed to be very dependent on the data used. Especially the relative complex class-hierarchy turned out not to be flexible enough for being applied successfully to other scenes. Applying the developed extraction methods to the QuickBird-Scene used here showed that several adoptions towards the segmentation and class-descriptions were necessary.

Since the class-hierarchy described in (Hofmann 2001) is more data driven than driven by a model of informal settlements, we decided to completely redesign the class-hierarchy, which takes more the phenomena and their ontologies into account. As a basic difference of both rule bases a noticeably reduced number of classes and sub-classes can be observed. Additionally in the rule base used here, the use of class descriptions by nearest neighbor was completely renounced, which led to classes only described by fuzzy membership functions and their combinations.

Nevertheless, in both cases the basic strategy and the initial segmentation were similar: starting with eCognition's multi resolution segmentation[1] as the initial segmentation a classification of the generated image objects follows and ends in an iterative process of knowledge-based object enhancement and (re-) classification (see Fig. 1).

---

[1] as implemented in Version 4.0
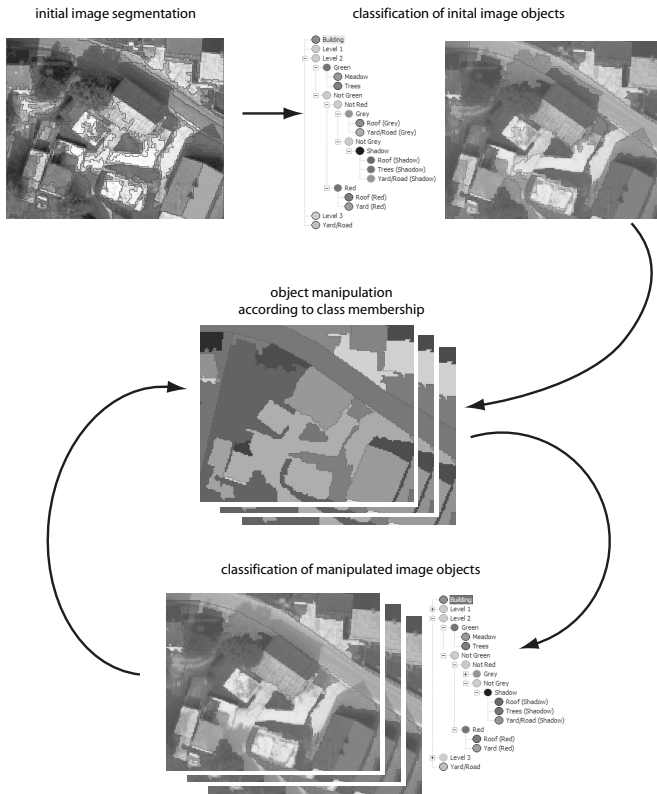
Fig. 1. Iterative process of (knowledge-based) image segmentation and classification

2.1 The ontology of informal settlements - what are informal settlements?

The methodological approach demonstrated here is mainly taking into account the ontology of the objects to be detected. In order to have a common understanding of "what is an ontology", a short introduction about ontologies in general and some possible descriptions of the ontology of informal settlements in special is given.

Besides the two basic definitions already given in the first chapter of this article, to describe what an informal settlement is depends strongly on the context respectively the point of view. Without going deep into details about theoretical aspects of ontologies, describing "what is an informal settlement" is a special case of describing phenomena of the (real) world. The

description of a phenomenon of the (real) world is usually expressed by describing its ontology. In general, this is done by describing knowledge about the phenomenon from a certain point of view, e.g. as like the two definitions stated in the beginning of this article and by using a certain language. Hence, the ontological description of "informal settlement" can be understood as the representation of knowledge about "informal settlement" from a certain point of view using a defined language. Thereby, the language has to follow certain rules which are described more detailed for example in Guarino (1998). However, there can be as many knowledge representations of a phenomenon as points of view (Fonseca 2001). These points of view are commonly named *domains*. In the context of image analysis, there are two basic domains - namely *image domain* and *real world domain* - identifiable. Both are interacting with each other as Fig. 2 illustrates.
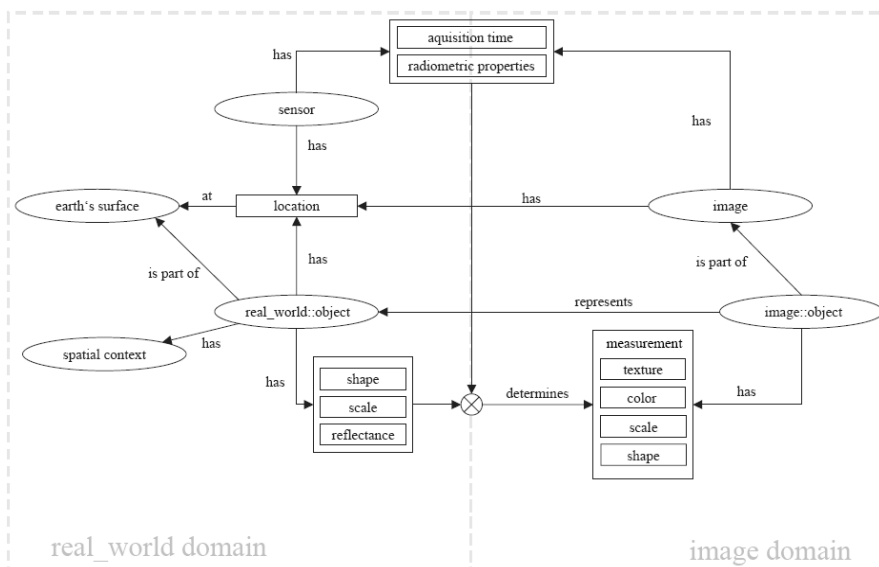
Fig. 2. Relationships between objects of the real world domain and the image domain.

Descriptions of a phenomenon from the point of view of the *real world domain* describe its general observable properties, i.e. what is typical or even unique for the phenomenon (here: informal settlements) in general. From the point of view of the *image domain*, these properties have to be measurable (i.e. detectable) in the image. These properties are commonly understood as the signature or pattern of a phenomenon. In the first in-

stance, the signature or pattern is independent of possible methods of image processing which are to quantify the signature or pattern of the phenomenon. Nevertheless, in order to separate different phenomena by their patterns a quantification of their patterns is unavoidable.

Using a taxonomical description in terms of "informal settlement is a special case of settlement" helps to identify unique properties of informal settlements and common properties with other types of settlement. Describing the spatial relationships between objects of different classes goes beyond the description of patterns but is nevertheless part of the phenomena's ontologies. I.e. these are descriptions of the phenomena from the real_world domain's point of view. Within the context of informal settlements, such relationships have to take into account the dense structure of housing and infrastructure, i.e. the informal character of these sub-elements as illustrated in Fig. 3 and the effect of segregation (see Fig. 4) which leads to distinct and spatially separated types of housing.
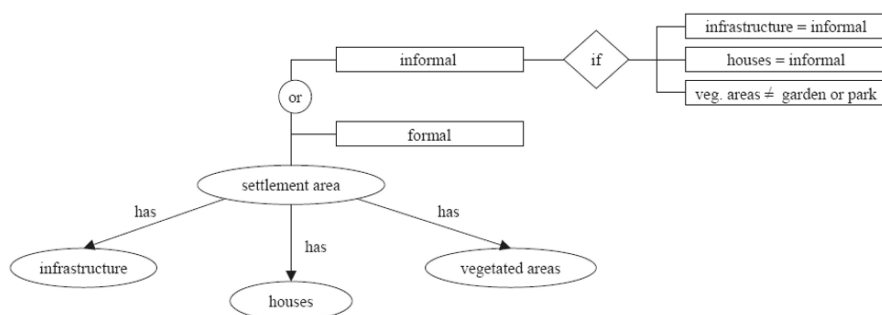
Fig. 3: General description of settlement area and informal settlement as a special case of it
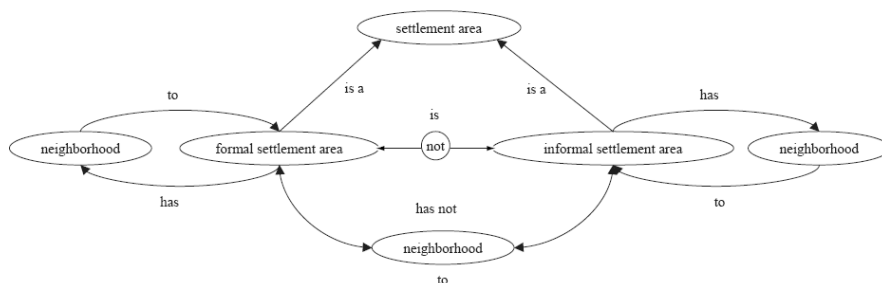
Fig. 4: Description of segregation

When referencing objects of other classes, especially when describing a spatial context relation, the ontologies of these other objects have to be described as well: analogous to the description of settlement areas and infor-

mal settlements, the ontologies of the sub-elements must be described sufficiently. I.e. it has to be described what is typical for an informal house, informal infrastructure or vegetation in the *real world domain* and in the *image domain* (see Fig. 5 and Fig. 6).
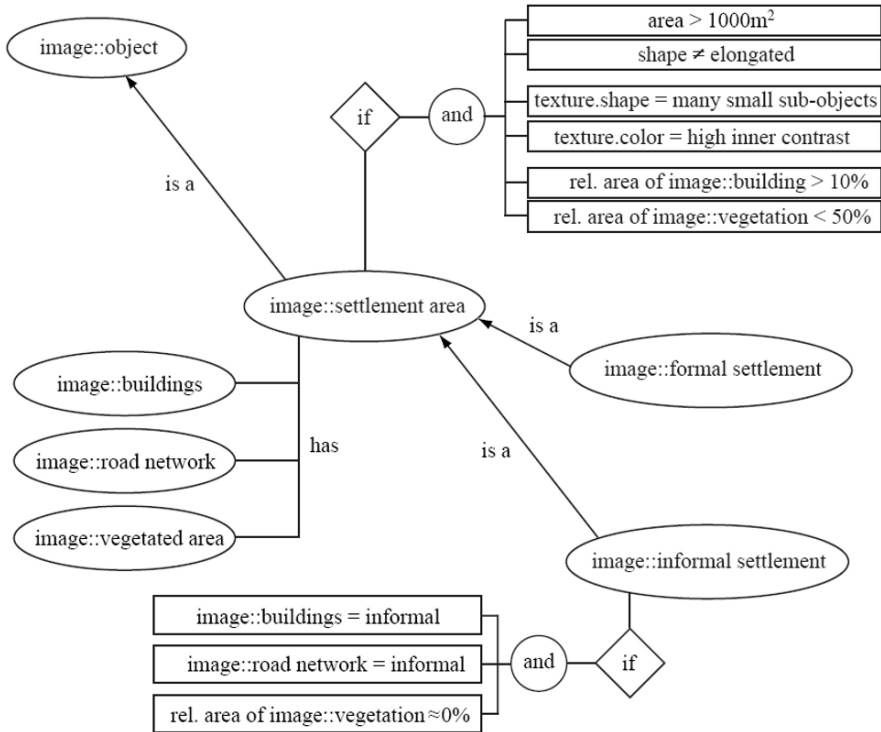
Fig. 5: Ontology description for informal settlement in the image domain
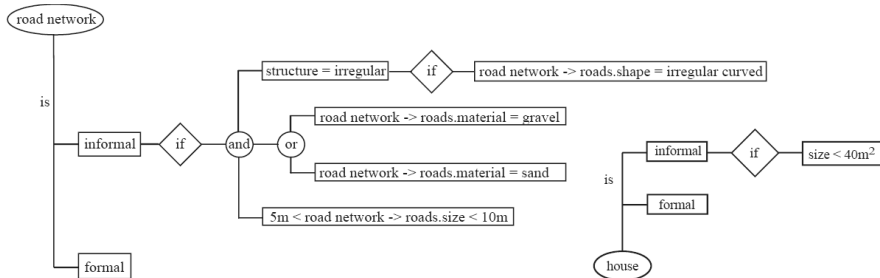
Fig. 6: Ontology description of road network and house in the real world domain

2.2 Image segmentation

Regarding the ontology of informal settlements, especially the spatial relationships between (*informal) settlement* and its sub-objects *buildings, road network and vegetation*, it is obvious that the image segmentation has to generate image objects representing settlement areas and the named sub-objects. Using eCognition with its multi-resolution segmentation method for this purpose, means generating two segmentation levels, wherein the top level more or less represents *settlement areas* and other objects of comparable size, while the base level holds image objects which coincide with *small houses*, *small road segments* and *small vegetation areas.* The image objects of both segmentation levels are linked to each other in terms of a hierarchical net of objects which is depicted by a tree-structure (see Hofmann 2001, Baatz et al 2004, Benz et al 2004). This makes it possible in principle to describe later on spatial relationships between *settlement areas* and *small houses*, *small road segments* and *small vegetation areas* as well as the described neighborhood relations reflecting aspects of segregation.

Regarding the segmentation parameters, according to Baatz et al (2004) the image objects created by the initial segmentation should best suit the image analysis purposes. I.e. the image segmentation should lead to image objects which best suit the ontologies of the desired classes. In practice this leads to many trial and error tasks in order to find the best suited segmentation parameters for the initial segmentation. Thereby, usually not all desired objects will be outlined semantically perfect, i.e. some objects are over-segmented, while others are under-segmented. Thus, the optimum segmentation parameters are those which obviously generate the least over- or under-segmentations. In the scenes used here and in Hofmann (2001) color contrast in informal settlement areas is relatively low. Thus, the segmentation of small houses respectively single shacks succeeded in both scenes only partially. However, Hofmann (2001) has already demonstrated that it is not necessary to identify each individual house to classify different types of settlement areas. Moreover, it turned out that most of the houses' shadows as well as roofs with higher contrast to their environment (mostly houses with red roofs) could be segmented well enough in order to identify different housing structures by the generated segments. These different housing structures were used later on successfully to differentiate informal settlement areas from other areas. For the ontology of informal settlements, from the point of view of the image domain, the spatial relationship between *(informal) settlement* and *(small) houses* must be expressed indirectly by these detectable indicators (i.e. roofs and shadows) and their respective properties (see Fig. 7).
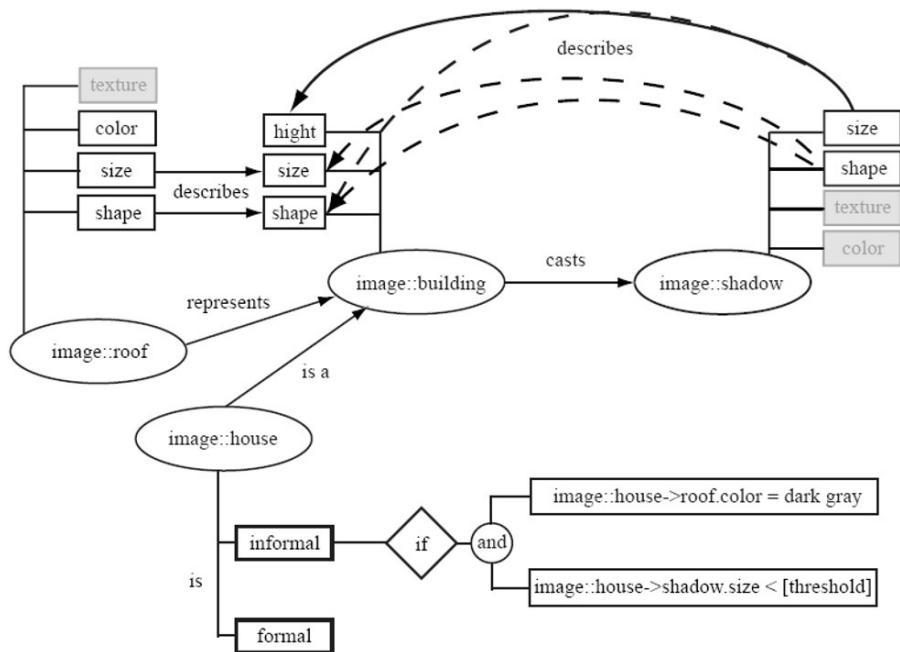
Fig. 7: Relationship between building properties and describing properties of its indicating sub-elements

Although the segmentation parameters used in Hofmann (2001) led to well usable image objects with the above described restrictions and although there is only little difference between the image properties of the scenes (see Table 1), using identical parameters for segmenting the Quick-Bird-scene did not lead to as useful image objects as in the IKONOS scene. In fact the average object size in the IKONOS scene was in the top level at approx. $3,165m^2$ and in the base level at approx. $49m^2$. Within informal settlements the average object size was at $14,608m^2$ and $34m^2$ respectively. In the QuickBird scene using identical segmentation parameters the average object size in the top level was at $1,045m^2$ and at $25m^2$ in the base level. Within informal settlements the respective average object sizes were $1,273m^2$ and $15m^2$. Regarding the objects' properties, only slight differences of the spectral properties are observable, but properties describing the objects' shape and structure vary (see Table 2).

While it was not quite clear which differences in the image properties finally led to these differences in the object properties, we assumed that one key feature is the higher resolution of the QuickBird scene. Hence, we decided to multiply the 'scale parameter' according to the ratio of the resolutions of both sensors and to segment the QuickBird-scene with 'scale pa-

rameters' of 144 instead of 100 (top-level) and 14 instead of 10 (base-level). Although this approach follows more heuristic assumptions compared to provable calculations, the objects generated by the adapted segmentation meet the ontological strings better than those generated by a non adapted segmentation and are visually more comparable (see Fig. 8). As Table 1 demonstrates, especially the shape features *area*, *length*, *width* and *border length* became more comparable by adapting the scale parameter according to the ratio of the resolution of the images.

Tab. 1: Properties of the image data used in Hofmann 2001 and in the research work described here

Scene		IKONOS Cape Town	QuickBird Rio de Janeiro
Location		Cape Town (Nyanga/Crossroads)	Rio de Janeiro (Ilha do Governador)
Spatial resolution			
Channel	pan	1.0m	-
	blue	4.0m	0.69m
	green	4.0m	0.69m
	red	4.0m	0.69m
	nir	4.0m	0.69m
Radiometric resolution			
Channel	pan	11bit (16bit)	11bit (16bit)
	blue	11bit (16bit)	11bit (16bit)
	green	11bit (16bit)	11bit (16bit)
	red	11bit (16bit)	11bit (16bit)
	nir	11bit (16bit)	11bit (16bit)
Band widths			
Channel	pan	450-900nm	450-900nm
	blue	450-520nm	450-520nm
	green	510-600nm	520-600nm
	red	630-700nm	630-690nm
	nir	760-850nm	760-900nm
pre-processing by provider		reclassification to 11 classes	pan-sharpening
pre-processing by customer		pan-sharpening	reclassification to 11 classes

Tab. 2: Feature-statistics for objects generated with different scale parameters in the scenes

	complete scene			inside informal settlements		
	IKONOS	QuickBird	QuickBird adapted	IKONOS	QuickBird	QuickBird adapted
	base-level					
Objects	1.037.038	1.485.830	812.729	6.326	7.628	5.779
Object Features	mean values					
Mean blue	464,73	520,84	522,74	460,23	606,52	599,59
Mean green	532,72	469,08	471,04	517,70	547,90	542,59
Mean red	470,72	504,12	506,64	457,53	586,06	581,50
Mean nir	502,65	553,94	552,92	441,15	506,99	495,60
Stddev blue	39,68	41,65	49,82	46,52	56,74	65,59
Stddev green	59,82	48,46	57,98	70,04	65,40	76,35
Stddev red	63,24	51,70	61,86	73,56	67,46	78,60
Stddev nir	54,13	70,07	83,64	59,34	87,17	100,58
Ratio blue	0,24	0,26	0,26	0,25	0,28	0,28
Ratio green	0,27	0,22	0,22	0,28	0,24	0,24
Ratio red	0,23	0,24	0,24	0,24	0,26	0,26
Ratio nir	0,26	0,28	0,28	0,23	0,22	0,22
Area [m²]	48,92	24,62	44,72	33,80	15,19	27,20
Length [m]	11,21	7,73	10,94	9,27	6,47	9,05
Width [m]	6,45	4,61	6,26	5,68	3,78	5,19
Border length [m]	36,90	25,57	36,55	31,49	21,56	30,82
Asymmetry	0,66	0,55	0,56	0,53	0,56	0,57
	top-level					
Objects	16.028	34.982	16.094	105	58	39
Object Features	mean values					
Mean blue	469,73	581,64	590,29	472,87	572,93	583,61
Mean green	540,66	535,42	545,26	539,07	513,42	524,39
Mean red	478,51	575,37	585,57	473,57	553,07	563,57
Mean nir	508,88	572,89	589,19	475,26	476,52	478,55
Stddev blue	71,29	101,79	114,63	111,78	138,29	166,63
Stddev green	108,74	118,38	133,07	170,69	161,60	193,60
Stddev red	115,16	125,20	140,70	179,84	165,20	195,53
Stddev nir	104,53	157,86	176,06	157,96	191,89	216,83
Ratio blue	0,24	0,26	0,26	0,24	0,27	0,27
Ratio green	0,27	0,23	0,23	0,27	0,24	0,24
Ratio red	0,23	0,25	0,25	0,24	0,26	0,26
Ratio nir	0,26	0,26	0,26	0,24	0,22	0,22
Area [m²]	3.165,36	1.045,52	2.268,07	14.608,33	1.273,12	3.616,53
Length [m]	129,97	69,42	104,84	185,25	75,39	120,65
Width [m]	53,81	29,96	44,23	116,08	38,54	60,56
Border length [m]	501,52	273,33	427,18	793,22	376,37	605,43
Avrg. mean diff to neighbors of sub-objects nir (1)	111,08	172,78	198,61	175,29	216,57	246,74
Number of subobjects	64,70	42,47	50,86	427,38	83,64	137,05
Asymmetry of sub-objects mean (1)	0,57	0,57	0,59	0,58	0,56	0,57
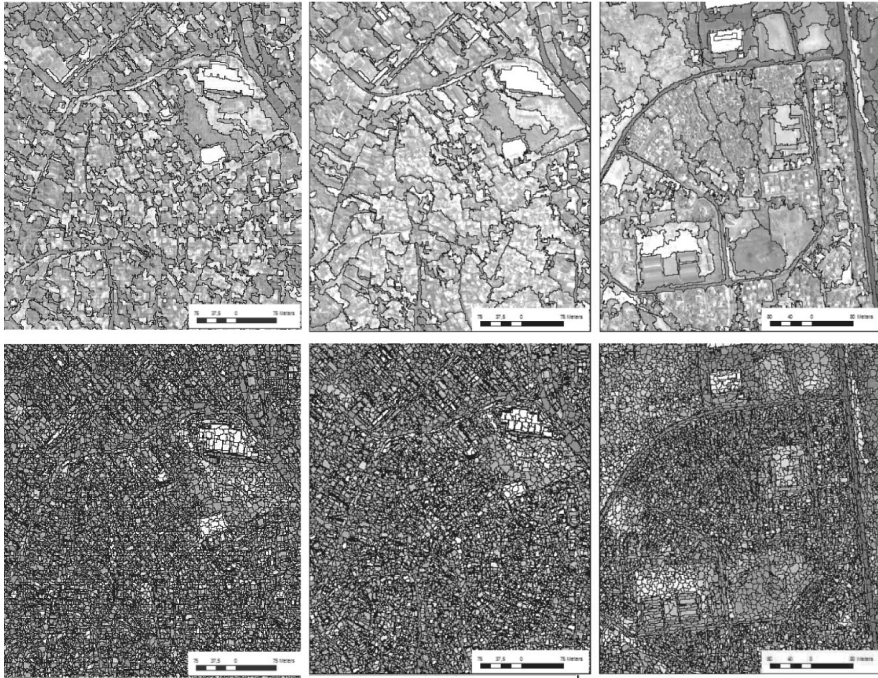
Fig. 8. Different segmentation results in subset areas of the IKONOS (right) and QuickBird (middle and left) scene, using identical segmentation parameters (left) and incurred parameters (middle). Top: top level, bottom: bottom level

2.3 From ontology to rule-base – generating an adequate knowledge description

While in Hofmann (2001) the design of the knowledge-base was more or less orienting on the data used, which led to a relatively complex class hierarchy with likewise complex class descriptions, the rule base developed here is more orienting to the phenomena and their ontologies. Thus, comparing both class-hierarchies, the newly generated one is more pruned, focusing on relevant properties according to the ontology and consequently more transparent. Having in mind to apply the class-hierarchy to several comparable scenes, the improved transparency simplifies necessary incurring and maintenance tasks.

Although the complexities of the class-hierarchies are very different, the base strategy of beginning with so-called "level super-classes" applies to both hierarchies. With these super-classes objects belonging to the base- or top-segmentation level are separated semantically by using a crisp membership function describing the feature *Level*. Following the spatial rela-

tionships of sub- and super-objects in the ontology of informal settlements, the classes *settlement area, formal settlement* and *informal settlement* are sub-classes of *top level*, whereas *formal settlement* and *informal settlement* are sub-classes of *settlement*. This way, properties which are typical for settlements in general are described in the class *settlement* and inherited to *formal settlement* and *informal settlement*, i.e. these properties are common to both types of settlement. The hierarchical structure reflects the "is_a" relations of the ontologies described before.

The classes *red roofs, small shadows/dark objects, bright small roofs/objects* and *vegetation* are sub-classes of *base level*. Simultaneously, the classes of the base level are acting as indicators for (*informal*) *settlement* as described in the chapter before. Since it was not possible to segment and identify single houses with informal character the density of the indicators *red roofs, small shadows/dark objects, bright small roofs/objects* was used to identify settlements and informal settlements. This was done by using and combining membership functions for the features *Asymmetry* and *Area of Sub-Objects* with the features *Rel. area of bright small roofs/objects, Rel. area of small shadows/dark objects* and *Rel. area of red roofs*. The feature *Avrg. mean diff to neighbors of sub-objects* in the nir-channel was used to describe the relatively high spectral heterogeneity within settlement areas. *Asymmetry* was used to differentiate settlement areas from other elongated objects like rail roads or roads.

Informal settlements then could be differentiated from other types of settlement by a smaller *Area of Sub-Objects*, an explicit lower *Rel. area of red roofs* combined with an explicit lower *Rel. area of bright small roofs/objects*, but an explicit higher *Rel. area of small shadows/dark objects*. The informal character of the road network within informal settlements was expresses by a relative low value for *Asymmetry of sub-objects: mean*. Since a higher value for this feature indicates more elongated objects present (e.g. regular road segments), a lower value indicates the opposite. The class *formal settlement* finally is simply described by the fuzzy-logical negation (inversion) of *informal settlement* combined with a fuzzy-limiting value for *Area* of more than 1800-1900m^2. Since *formal settlement* is a sub-class of *settlement*, it inherits al its feature descriptions and is simultaneously the 'opposite' of *informal settlement*, i.e. all settlement areas which are not informal (see Fig. 9 and 10).
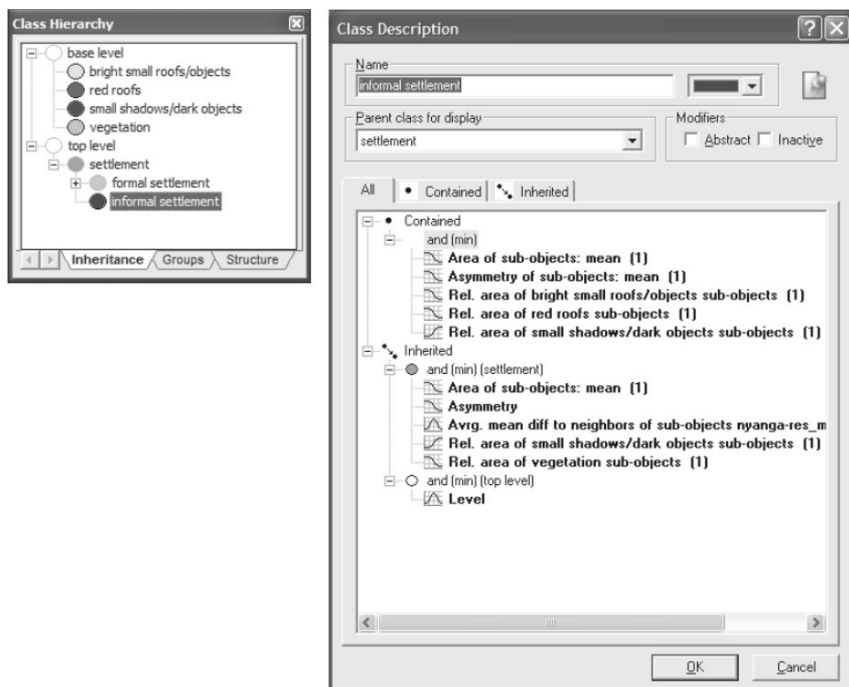
Fig. 9. Class hierarchy used for the detection of (informal) settlements (left) and class description of the class informal settlement as a sub-class of settlement (right)
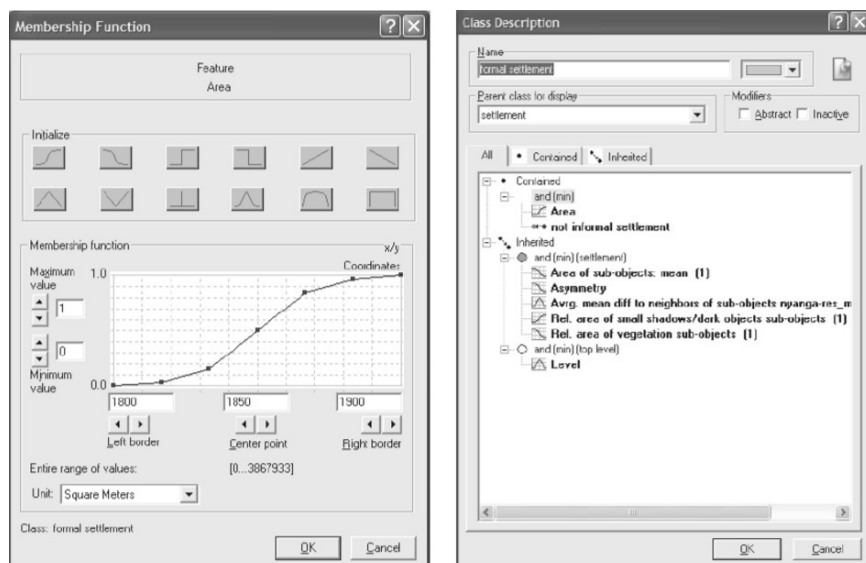
Fig. 10. Description of formal settlement as sub-class of settlement and the inverse class of informal settlement (right) and a limiting area of more than 1900m^2 (left)

Applying this class-hierarchy to the segmented image leads to a classification result which holds formal and informal settlement areas, whereas only those informal settlement areas are extracted, which fulfill the criteria for *informal settlement* to at least 50%. Due to the structure of the class hierarchy - especially the inheritance relation between *settlement*, *formal settlement* and *informal settlement* – there are some *informal settlement* areas, which are wrongly classified as *formal settlement*. I.e. *settlement* areas could be extracted relatively well, but to distinguish between formal and informal settlement it was necessary to take into account further properties of (informal) settlement. Regarding the ontology of (informal) settlement, taking aspects of segregation into account, seams to be a sensible approach.
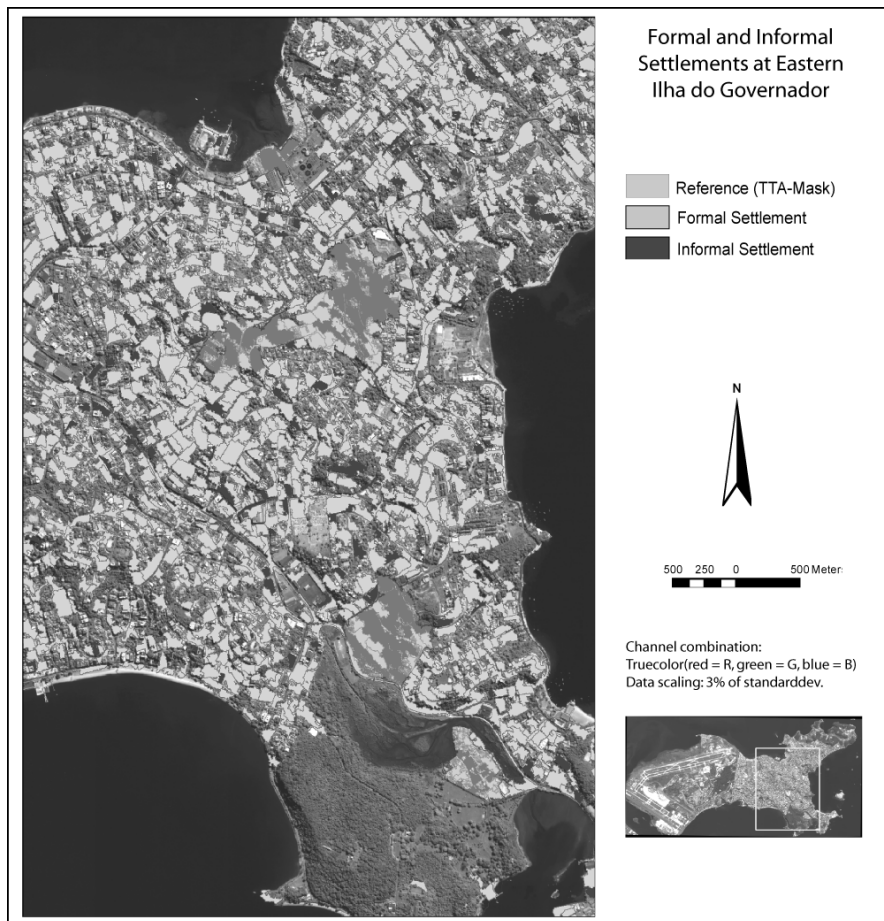
Fig. 11. Classification result without taking aspects of segregation into account, i.e. prior to the application of a knowledge-based iterative segmentation

2.4 Knowledge based iterative segmentation

The class descriptions presented until now were mainly focusing on describing aspects of color, shape, texture and structure, i.e. the physical properties of (informal) settlements. As was shown already in this article, one typical property of settlement areas – especially in urban areas – is to be segregated according to different features which are not all directly detectable from remote sensing data. Nevertheless, regarding spatial neighborhood relationships between different types of housing, it is very unlikely that formal settlement areas are completely surrounded by structures of informal settlements. Notably the inverse constellation is very likely so that preferably within transition zones from formal to informal settlements such mixed areas do occur. However, the classification result in the chapter above illustrate some misclassified *formal settlement* objects which are either very close to *informal settlement* objects or even embedded by such areas. Although the objects fulfill the criteria of *settlement*, they do not for some reason for *informal settlement* and are thus classified as *formal settlement* (the inverse of *informal settlement*). These circumstances seam to infringe the 'rules' of segregation, i.e. it seams to be very likely, that these objects are truly misclassified and should be assigned to *informal settlement*. In order to identify *formal settlement* objects which are embedded by *informal settlement* objects, we used the feature *Rel. area of [class] neighbor-objects*, whereas *informal settlement* was used for *[class]*. The distance was set to 0.0, which means we are looking for objects at direct neighborhood. In order to obtain more realistic values for measuring embedding, we fused all neighboring objects classified as *informal settlement* by the so-called 'classification based segmentation'. This way, the higher the *Rel. area of informal settlement neighbor-objects* is, the more it is embedded.

To realize the measurement after the object fusion, a new class was created as a sub-class of *formal settlement,* but with the additional property *Rel. area of informal settlement neighbor-objects 0.0* fuzzy-more than 0.3 – 0.5. This means, all objects classified as *formal settlement* and whose share of area of informal settlement in the direct neighborhood is fuzzy-more than 30% - 50% are assigned to the new class *formal surrounded by informal* (see Fig. 12)*. Since we stated before that these objects are actually misclassified - i.e. they have the physical properties of *formal settlement*, but should be regarded semantically as *informal settlement* - the class *formal surrounded by informal* was assigned to the semantic group (Baatz et al 2004) *informal settlement* (see Fig. 12).
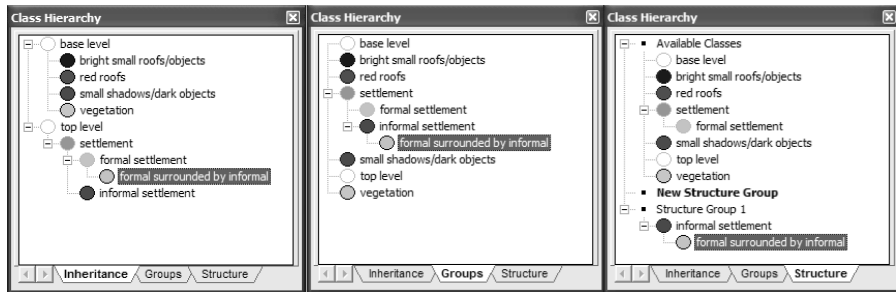
Fig. 12. Extended class-hierarchy to detect wrongly classified embedded formal settlement areas

This semantic assignment in consequence raises the number of *informal settlement* objects and of course the global area of this class. For the *formal settlement* neighbors of the prior *formal surrounded by informal* neighbors this means: they now have a direct neighborhood to *informal settlement*. If their value for *Rel. area of informal settlement neighbor-objects 0.0* is fuzzy-more than 0.3 – 0.5 they are now embedded by *informal settlement* too, i.e. they are now belonging to the class *formal surrounded by informal*. Now, the fusion and reclassification process can be started again until the misclassified "holes" inside *informal settlement* areas are merged to contiguous *informal settlement* objects (see Fig. 13).
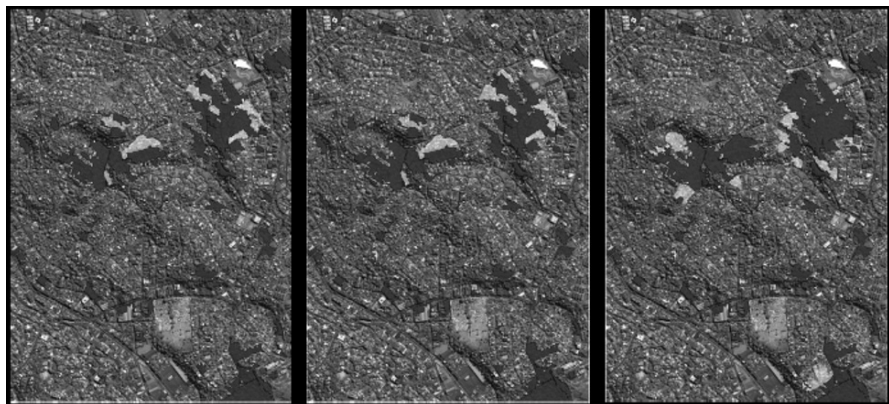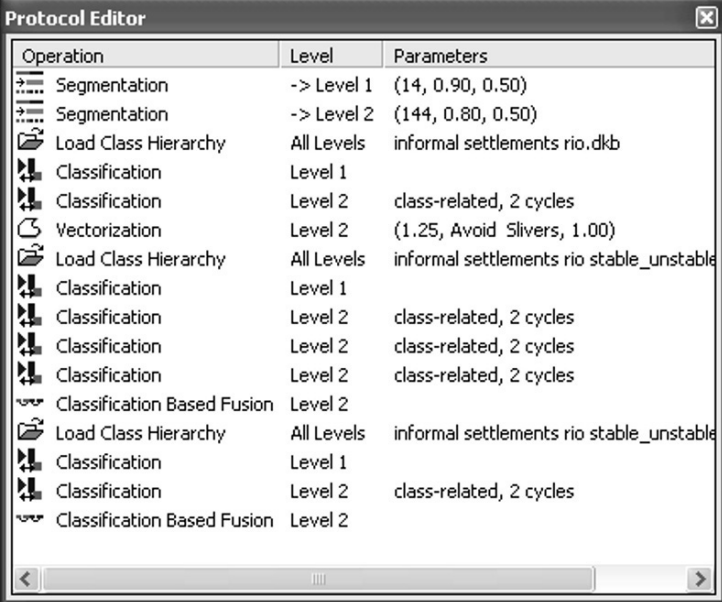
Fig. 13. Sequence of classification, knowledge-based fusion and re-classification of merged image objects. Left: initial classification. Middle: re-classification after merging. Right: final result after merging again

Repeating this sequence of reclassification and fusion leads to an iterative process that will stop as soon as there are no further *formal settlement*

objects present which are embedded by *informal settlement*, or the process is aborted by the operator. The former case of course needs some expert knowledge about the settlement structures in the scenes of matter. E.g. objects classified in the last step as *formal surrounded by informal* are in the most cases transition zones between formal and informal settlement areas, i.e. they are a mixed type of both forms of settlement. Saving the sequence of (re-)classification and fusion in a so-called 'protocol' leads to a program-like knowledge-based region-growing procedure. In this form it is possible to re-apply the procedure or modify it wherever necessary (see Fig. 14 and 15).

Operation	Level	Parameters
Segmentation	-> Level 1	(14, 0.90, 0.50)
Segmentation	-> Level 2	(144, 0.80, 0.50)
Load Class Hierarchy	All Levels	informal settlements rio.dkb
Classification	Level 1	
Classification	Level 2	class-related, 2 cycles
Vectorization	Level 2	(1.25, Avoid Slivers, 1.00)
Load Class Hierarchy	All Levels	informal settlements rio stable_unstable
Classification	Level 1	
Classification	Level 2	class-related, 2 cycles
Classification	Level 2	class-related, 2 cycles
Classification	Level 2	class-related, 2 cycles
Classification Based Fusion	Level 2	
Load Class Hierarchy	All Levels	informal settlements rio stable_unstable
Classification	Level 1	
Classification	Level 2	class-related, 2 cycles
Classification Based Fusion	Level 2	

Fig. 14. Sequence of knowledge based iterative process saved as protocol
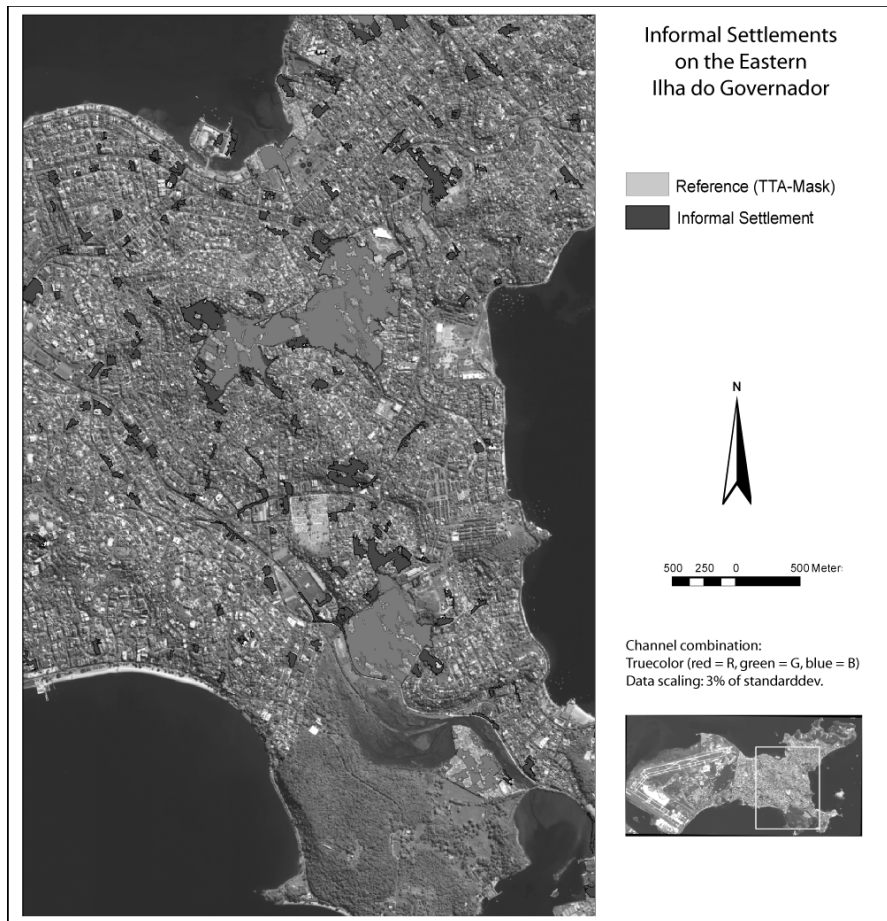
Fig. 15. Final result after applying the iterative process of knowledge-based classification and fusion without transition zones

3 Accuracy Assessment

To assess the accuracy of the classification process we used the results of a manual classification as reference areas (see Fig. 11 and 15). In this manual reference only larger favelas were outlined manually as contiguous areas. Detecting smaller favelas is even hard when performing it manually/visually. Thus, we could only assess the accuracy by comparing the automatic classifications with the manual reference mapping of larger contiguous favelas. I.e.: we could determine how many pixels of the reference mapping were also classified by the automatic classifiers which then lead

to a percentage of agreement between both classifications and an error of omission respectively. As Table 3 demonstrates, for the classification before applying the iterative approach 47% of the pixels of the reference mapping were also detected by the automatic classifier, i.e. the error of omission was at 53%. After applying the iterative process the ratio of agreement between automatic and manual classification could be raised to 68%, i.e. the error of omission could be reduced to 32%. Nevertheless, it has to be mentioned, that after controlling the automatic results visually – especially those of the iterative approach – we discovered, that some of the smaller favelas were detected by the automatic classifier probably correctly, i.e. they were probably omitted by the manual (reference) mapping (see Fig. 15). However, only a ground truth could give absolute evidence for these assumptions.

Tab. 3: Comparison of manual and automatic classification of favelas expressed by agreed and omitted pixels.

	before iterative process		after iterative process	
	no. of pixels	in %	no. of pixels	in %
ommited	1067350	53,13	634836	31,60
agreed	941534	46,87	1374048	68,40
sum	2008884	100,00	2008884	100,00

Regarding the measures 'Best Classification Result' and 'Classification Stability' offered by eCognition[2] (see Baatz et al 2004, pp. 160 - 163), there are also slight enhancements observable, when applying the iterative process: The mean values for 'Best Classification Result' and 'Classification Stability' within the reference areas could be raised from 0.92 to 0.94 after applying the iterative process. These values indicate, that most of the detected *informal settlement* objects are described relatively well by the developed class descriptors and that the classifier is relatively well suited to distinguish *informal settlement* from other classes.

4 Conclusion and outlook

In the paper present we have demonstrated how informal settlements can be extracted from VHR satellite imagery using an object-based approach. In contrast to former approaches the image segmentation and knowledge description was more driven by the ontologies of the desired objects than by the data used. Nevertheless, some data related adaptations of the seg-

---

[2] Version 4.0

mentation parameters were necessary. However, the class hierarchy and the class descriptions could be simplyfied which finaly leads to a more transperent and easier maintainable extraction process. Further, by formulating and applying image-independent knowledge about the spatial behaviour of the desired objects (here: segregation) led to an iterative process which produced enhanced objects and classification results with higher accuracy. Additionally, we assume that due to the pruned class-hierarchy, i.e. the more ontology-driven class descriptions the class-hierarchy is more transferable than the data driven class-hierarchy. One hint for this assumption is given by the few necessary adaptions of the class-hierarchy and class descriptions we observed while re-applying the approach to the Cape-Town-Scene. How to tackle the transferabilty of rule bases in object-based image analysis is currently under investigation.

References

Baatz M. et al. (2004): eCognition 4 User Guide, 4th Ed., Munich, 2004.

Benz U. et al. (2004): Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. In: *ISPRS Journal of Photogrammetry & Remote Sensing* 58 (2004), pp. 239-258.

Dare P.M., Fraser C.S. (2001): Mapping informal settlements using high resolution satellite imagery. In: *Int. J. of Remote Sensing*, Vol. 22, No. 8 pp. 1399-1301.

Fonseca T.F. (2001): Ontology-driven Geographic Information Systems, University of Maine, Diss., 2001.

Guarino N. (1998): Formal Ontology and Information System. In: *Proceedings of the International Conference on Formal Ontology in Information Systems*, IOS Press, Amsterdam, 1998, pp. 3-15.

Hofmann P. (2001): Detecting Informal Settlements from IKONOS Image Data Using Methods of Object Oriented Image Analysis – An Example from Cape Town (South Africa). In: *Jürgens, C. (Ed.): Remote Sensing of Urban Areas / Fernerkundung in urbanen Räumen*. Regensburg: University Regensburg, Germany, 2001, pp. 41-42.

Kuffer M. (2003): Monitoring the Dynamics of Informal Settlements in Dar Es Salaam by Remote Sensing: Exploring the Use of SPOT, ERS and Small Format Aerial Photgraphy. In: *Schrenk, M. (Ed.): Proceedings CORP 2003*, pp. 473-483.

Lemma T. et al. (2005): A Participatory Approach to Monitoring Slum Conditions. http://www.itc.nl/library/papers_2005/conf/sliuzas_par.pdf (accessed 2006).

Mason O.S., Fraser C. (1998): Image sources for informal settlement management. In: *Photogrammetric Record*, Vol. 92 (1998), No. 16, pp. 313-330.

Mason S.O. et al. (1998): Spatial Decision Support Systems for the Management of Informal Settlements. In: *Comput., Environ. and Urban Systems*, Vol. 21, No ¾, pp. 189-208.

Radnaabazar G. et al. (2004): Monitoring the development of informal settlements in Ulanbaatar, Mongolia. http://www.schrenk.at/CORP_CD_2004/archiv/papers/CORP2004_RADNAABAZAR _KUFFER_HOFSTEE.PDF (accessed 2006)

UN-HABITAT (2006): http://www.unhabitat.org/campaigns/tenure/cws.asp (accessed 2006).

UN-HABITAT (2006a): http://www.unchs.org/programmes/guo/statistics.asp (accessed 2006).

UNSTAT (2005): http://unstats.un.org/unsd/environmentgl/gesform.asp?getitem=665 (accessed 2006).

Chapter 6.2

Opportunities and limitations of object-based image analysis for detecting urban impervious and vegetated surfaces using true-colour aerial photography

M. Kampouraki, G. A. Wood, T. R. Brewer

School of Applied Sciences, Department of Natural Resources, Cranfield University, Bedfordshire, MK43 0AL, UK (m.kampouraki.s04, g.a.wood, t.brewer)@cranfield.ac.uk

   ABSTRACT: Monitoring soil sealing in urban environments is of great interest as a key indicator of sustainable land use. Many studies have attempted to automatically classify surface impermeability by using satellite or aerial imagery. Air photo interpretation (API) has been used as a method to verify their accuracy. However, independent accuracy assessments of API have not been widely reported. The aims of this research are, firstly, to investigate independent accuracy assessments of API. Secondly, to determine whether object-based image analysis could replace manual interpretation for the detection of sealed soil and vegetated surfaces at the residential garden plot level. Four study areas, representing the industrial, commercial and residential parts of Cambridge, UK were manually digitised and classified by API. The same areas were automatically segmented and manually classified with the use of eCognition. The two methods were compared and the average overall mapping agreement was estimated to be 92%. The disagreement was qualitatively analysed and the advantages and disadvantages of each method were discussed. The very high agreement between the two methods in conjunction with the benefits of the automated method led to the conclusion that automated segmentation using eCognition could replace the manual boundary delineation when true-colour aerial

photography is used. Future work will examine automated image classification methods, using eCognition, as a replacement for normal image interpretation methods.

1 Introduction

Urban development presents the greatest driver of soil loss due to sealing-over by buildings, pavement and transport infrastructure. Soil sealing is recognised as one of the major threats to soil. The ability to monitor the rates, types and geo-spatial distribution of soil sealing is crucial to understanding the severity of pressure on soils and their impact on European and global socio-economic and environmental systems (Wood et al., 2006).

1.1 Monitoring soil sealing by remote sensing

There are few internationally recognised definitions of soil sealing (Burghardt et al. 2004). The European Union accepts that "soil sealing refers to changing the nature of a soil such that it behaves as an impermeable medium and describes the covering or sealing of the soil surface by impervious materials" (EEA glossary 2006). Remotely sensed data cannot directly measure whether a surface is permeable but it can monitor cover types (e.g. concrete or tarmac) and infer permeability. Grenzdörffer (2005) categorised sealed areas simply as either built-up or non-built-up areas.

Arguably, the most detailed mapping of soil sealing was carried by the Office for Urban Drainage Systems in Dresden, Germany. They used orthorectified aerial photography (1:50,000 scale) and digitized soil sealing values for the whole city by air photo interpretation (API). The degree of sealing was estimated for each housing plot and given a soil sealing value, e.g. roofs were 100% sealed; green roofs, 50%; concrete-asphalt 100%; semi-permeable areas (paving stone) 70%; water-absorbing areas like gravel, 50% and residual areas, 0% (Meinel and Hering 2005).

Recently, a variety of projects have been undertaken in Europe to develop more automated methods for detecting soil sealing at European, national or regional scales such as the SoilSAGE project, the GMES Urban Services (GUS) project, the GMES Service Element (GSE) Land monitoring project, the Monitoring Urban Dynamics (MURBANDY) project and the Monitoring Land Use/Cover Change Dynamics (MOLAND) project. Soil sealing has also been investigated by the Technical Working Groups (TWG) of the Soil Thematic Strategy described by Burghardt et al. (2004) in two reports. Most of these projects have used remote sensing image

classification techniques based on pixel procedures.

The argument for using object-based image analysis over pixel-based methods will not be repeated here (see Blaschke and Strobl 2001; Caprioli and Tarantino 2003; Yuan and Bauer 2006). Suffice it to say that real-world objects are not characterized by single, square pixels. In the case of high and very high resolution imagery, groups of individual pixels are more likely to represent what would normally be interpreted as recognisable land cover features. Object-based image analysis is based on sensible pixel groupings and is, therefore, more representative of the systematic process carried out in API. Delineating image objects by 'segmentation' in the digital domain is analogous to API boundary delineation.

Automatic segmentation is not new (Blaschke and Strobl 2001). Existing algorithms include texture segmentation, watershed information and mean shift, but none of them have proved to be a robust, operational approach (Zhou and Wang, 2006). More recently, with the introduction of eCognition software, from Definiens Imaging GmbH, homogeneous image object extraction, over a range of image object sizes, is now possible. In contrast to pixel approaches, image objects produced using eCognition contain spectral, shape and texture information but also a whole network of relations which connects image objects and incorporates contextual information. The objects extracted during the segmentation process are then later classified.

Many studies have attempted to extract urban features and classify urban land cover and land use by using eCognition, e.g. Hoffman (2001); Herold et al. (2003); Wang et al. (2004); Frauman and Wolf (2005); Blaschke et al. (2005). Mittelberg (2002) attempted to analyse the urban environment by using aerial photography and very high resolution IKONOS data. Hodgson et al. (2003) used aerial photography along with elevation data (Lidar) to identify urban imperviousness. The data were compared with visual interpretation of aerial photography which was segmented using eCognition. Cothren and Gorham (2005) analysed QuickBird images to detect impervious and permeable surfaces. Grenzdorffer (2005) used a combination of satellite (Landsat TM and SPOT) images with high resolution aerial photographs to identify urban land use change. Yuan and Bauer (2006) investigated digital classification techniques (both pixel and object-based) for mapping urban impervious surfaces using QuickBird images. In most cases, the classification results were compared with an air photo interpretation of ortho-rectified aerial photography.

API is considered *de facto* as the most accurate procedure for mapping land cover and none of the studies cited determines the appropriateness of using API methods to assess the accuracy of boundary delineation for landcover mapping. API is also subjective, time consuming, expensive, la-

bour intensive, and requires skilled operators.

This paper aims to evaluate manual classification of aerial photography by comparing it with results produced using eCognition. The work specifically focuses on the segmentation stage of the process, where API will be compared with semi-automated object-based procedures. The scope is to investigate whether object-based image analysis could replace the traditional way of manual digitising and visual labelling for the detection of sealed soil and vegetated surfaces at the residential garden plot level.

2 Data and methods

The study area is the city of Cambridge, UK. The data source acquired for the analysis is ortho-rectified aerial photography, taken in June-July 2003 at 0.125 m spatial resolution and scanned to an 8 bit RGB format. The image provided, was an already geometrically corrected mosaic where the stereo pairs were unavailable. Four study areas of 250 by 250 m were selected as representative land covers of the built environment: two types of residential, one commercial and an industrial part of Cambridge (Fig. 2.1).

Fig. 2.1 (i) 1960's semi-detached residential area with large gardens, **(ii)** Densely built Victorian terrace house residential area with small, narrow gardens **(iii)** City centre commercial area with tall buildings, densely built, predominantly sealed, **(iv)** Industrial area mainly sealed with little green space. Cities Revealed® copyright by The GeoInformation® Group, 1996 and Crown Copyright© All rights reserved

2.1 Aerial Photo Interpretation (API)

The four study areas were manually segmented by on-screen digitising using ArcGIS® software (Figs. 2.2i, 2.2ii) at 1:200 scale for two main reasons: (a) a 2 m minimum mapping unit ($4m^2$ area) was deemed to provide a good threshold for extracting urban features found in the built environment of Cambridge, and (b) a larger scale than 1:200 revealed a higher degree of pixelation which was difficult to interpret. Features smaller than $4\ m^2$ on the ground were ignored even though they could be seen (i.e. small individual trees, narrow footpaths in back gardens, or small areas of shadow). Vegetated surfaces were equated to unsealed soil, and non-vegetated surfaces were equated to sealed soils. Only shadow cast on ground surfaces were digitised as 'shadow'; sides of buildings in shadow, visible due to relief displacement, were interpreted as 'sealed'. Seven land cover classes were used (Fig. 2.2iii): sealed surfaces, vegetation, trees,

shadow, rail tracks, bare soil and temporary features. Shadow was further classified as 'sealed surface in shadow', 'grass in shadow', 'tree in shadow', and 'mixed or unclassified shadow' when it was impossible to identify the kind of objects in the shadow.

Residential area description

☐ 1= sealed surfaces
☐ 2= vegetation
☐ 3= trees
☐ 4= shadow
☐ 5= rail tracks
☐ 6= bare soil
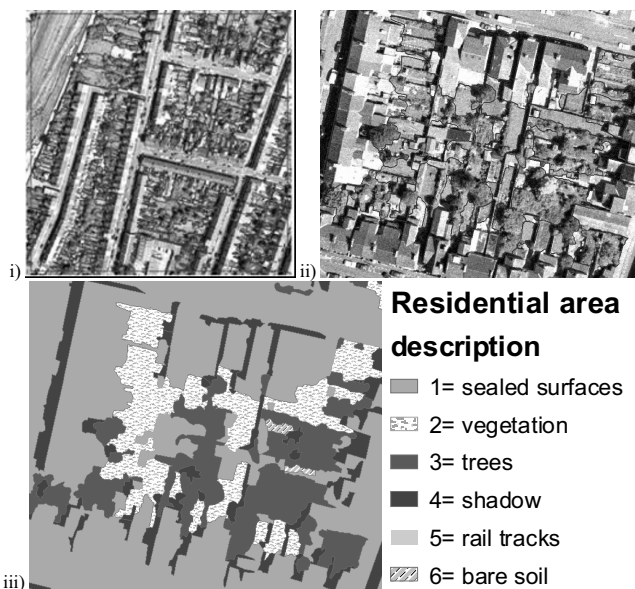
Fig. 2.2 (i) On-screen manual digitising of the densely built residential study area **(ii)** Delineation of feature detail, **(iii)** Manual classification. No temporary features were identified in this example

2.2 Semi-automated object-based classification approach

The four study areas were automatically segmented with the use of eCognition software, Definiens Professional version 5. The multi-resolution segmentation, which generates objects resembling ground features very closely (Definiens User Guide 2006), was used for this study. As a first step, eCognition links pixels to produce image objects by extracting homogeneous areas. The outcome of the segmentation is dependent on several parameters, such as, scale, colour, shape, compactness and image layer weights. These parameters are defined manually by the user. The scale parameter determines the maximum allowable heterogeneity for the resulting image objects and, consequently, their size.

A new feature of eCognition v.5 is the process editor. A single process represents an individual operation of an image analysis routine and defines

an algorithm which is executed on a specific image object domain. The image object domain describes the area of interest where the algorithm will be executed in the image hierarchy (Definiens User Guide 2006) and can either be the raw data at the pixel level, all of the image objects in a specific level of the hierarchy, or a specific object class from any level. The flexibility of the new version gives the ability to apply rules in a specific class domain of the class hierarchy that suits local conditions in an image. This affords similar flexibility that an interpreter has during manual API.

A general rule of thumb for a meaningful segmentation is to create image objects as large as possible and as small as necessary (Definiens User Guide 2006). The image must be segmented at such a scale so as to identify the smallest feature of interest. It is very important to use as many object levels, at different scales, as necessary until all image objects explicitly represent the classes to be assigned for the classification procedure. For a detailed description of image segmentation using eCognition and how the parameters affect the image analysis, see Baatz and Schape (2000), Benz et al (2003) and Definiens User Guide (2006).

The first study area examined was the industrial area. After several empirical trials the image was segmented into two levels. At the upper level, the best values for each parameter were found to be: scale=225, shape= 0.3, compactness= 0.7, weight of red band (layer 1) = 2, weight of green band (layer 2) = 4 and weight of blue band (layer 3) = 1. Using these parameters, the image sample was segmented into the coarse classes of sealed, unsealed and shadowed areas. Features that were not extracted at that scale were identified later when a new level with a smaller scale value was applied. High values of compactness along with larger weights in the red and green wave bands resulted in a better discrimination between trees (highly compacted objects) and other vegetated surfaces. The image was then manually classified with the manual editing tool. The manual classification followed the same pattern and criteria used with the API. The classes identified were: sealed surfaces, vegetation, trees, shadow, rail tracks and mixed areas. The 'mixed areas' class represents the regions of the image which are not satisfactorily segmented and require a lower scale parameter in order to create meaningful objects. Consequently, the 'mixed areas' class was re-segmented with a scale parameter value equal to 40; all other parameters remained the same (Fig. 2.3). Every individual neighbour polygon assigned to the same class was merged at both segmentation levels. The whole process rule set was saved and used for the image analysis of the remaining study areas. The same methodology and identical parameter values (rules) were used in the three other urban areas, in order to determine the transportability of the rules to areas where the urban land cover is different.

i)   ii)
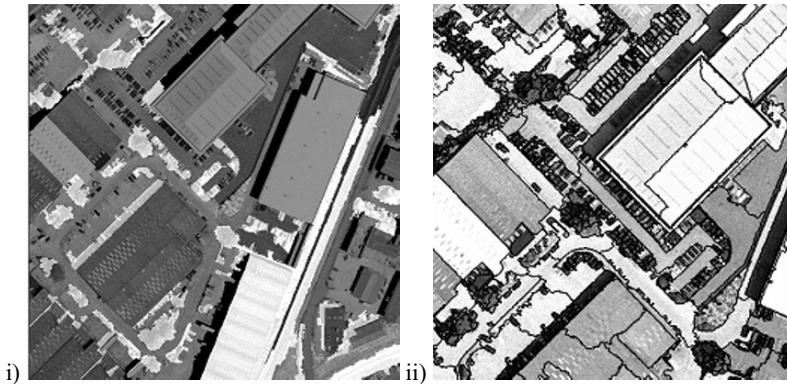
Fig. 2.3 (i) Industrial area segmentation at scale 225 and manually classified. **(ii)** The bright areas are the mixed areas which were later re-segmented at scale 40

2.4 Accuracy assessment

The results from eCognition were exported to ArcGIS® as smoothed polygons in vector format. The accuracy of the results was quantitatively assessed by comparison with the visual interpretation of the ortho-rectified aerial photography by cross-tabulation. The maps were also qualitatively analysed in order to understand any differences between the two approaches.

3. Results and discussion

3.1 Quantitative analysis

Initially, all the maps produced by the two methods were in vector format. The data from each method in a study area were merged together by a union function and the attributes of the new map were exported to a spreadsheet for the production of confusion matrices. The results showed very low agreement between the two methods. All four areas had accuracies between 28-29% which can be explained by the fact that although the two segmentations look very similar at the small scale, the boundaries of each polygon do not match perfectly (Fig. 3.1). Many insignificant 'sliver' polygons were produced when the maps were combined. This is due to the fact that eCognition follows a pixel pattern while the interpreter digitizes with smoother lines. The very small sliver polygons carry the same weight in

the cross-tabulation as the larger polygons of interest, which introduces bias and leads to an artificial underestimate of the overall mapping accuracy.
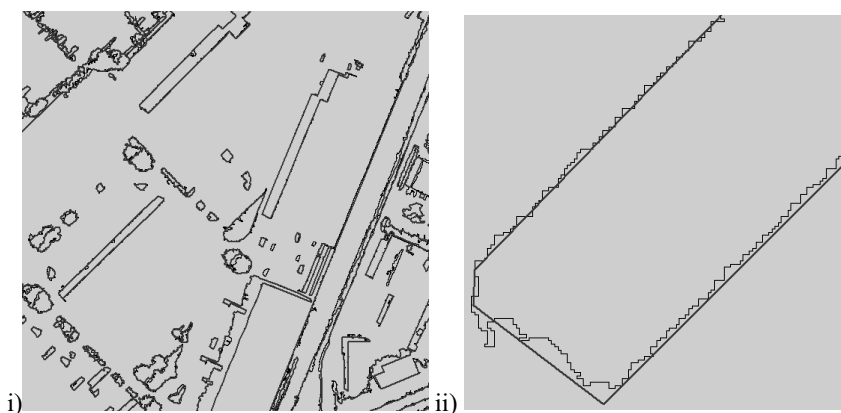
Fig. 3.1 (i) The two maps produced by API and eCognition overlaid for comparison – the differences are negligible at this scale. **(ii)** Enlargement of a polygon reveals the minor discrepancies

To solve the boundary problem and to eliminate the majority of sliver polygons, the vector files were converted to a raster format with a cell size of 0.125 m (equivalent to the pixel resolution of the aerial photography). The cross-tabulation was repeated. The accuracies obtained for each study area were: industrial area 96.3%, commercial area 94%, residential area (semi-detached houses) 89% and residential area (dense terrace houses) 90%. Because the polygons were attributed manually in both methods, the high agreement was expected (92% on average). Perhaps of greater interest is the 7-8% disagreement. Although this is not significant in the example of Cambridge, the causal factors may have greater predominance in other urban areas or at different times, and so must be understood.

3.2 Qualitative analysis

Close scrutiny of the factors causing the differences between manual delineation and eCognition's segmentation identified many examples of misclassification that were due to human error during digitisation. When the manual digitising was compared with the automated segmentation, some surface objects greater than the proposed 4 m^2 were found to have been

omitted. These same objects were successfully recognized in eCognition.

The omitted features were predominantly shadow or individual trees, either in back gardens or along streets. Sometimes, especially for trees, the reason they were not digitised was because they had low contrast with neighbouring objects and were not distinct enough to be easily identified. At 1:200 scale, it is sometimes difficult for the human eye to distinguish a tree when it is next to grass. This depends on the type of tree and the condition of the grass, as they affect the intrinsic contrast between object tones. It also depends on the levels of illumination and quality of the photograph or scan. A smaller scale can help to overcome this, but the delineation of the boundary is less precise and difficult to digitise. Consequently, there is a high probability that such features will be missed by API. eCognition automatically recognised these cased due to subtly different textures (internal object tonal variability) in the tree canopies; trees are very compact with a 'rough' texture, while grass is more monotone. Again, this may vary depending on image quality.

A big advantage of using eCognition is that the image can be segmented at scales larger than 1:200. In many examples, eCognition has extracted features that were lost in the API due to this threshold. If automated segmentation is used for boundary delineation, a fixed scale is not necessary and the image can be analysed in greater detail compared to API.

Conversely, there are cases where eCognition has not satisfactorily identified or separated features, for example, the fusion of individual trees with the shadow next to them and also the misidentification of smooth textured trees with bushes or grass. But this discrimination was also difficult during the API and therefore this misclassification can occur in the manual digitising. Examples of these misclassifications are infrequent and can be considered less important for mapping sealing, since both these classes are indicators of unsealed soil.

A significant difference between the API and the image segmentation is that eCognition identifies image objects, which are not always real world objects. A very good example is shown in Figure 3.2. The interpreter has the intelligence to identify the two different types of shadow, one produced by the building and one by the chimneys, and can ignore the latter by digitising the shadow polygon along the roof edge. But eCognition has created one polygon with both shadow types included in it. This may cause problems during automated classification if different features are classified with this type of shadow polygon.

The automated segmentation has the advantage of being much less time consuming, especially if rule sets can be universally applied. Half of the amount of time was needed when the boundary delineation was made by using eCognition in comparison to manual digitising. Manual classification

in eCognition was also faster but the interpreter was familiar with the area as the classification was repeated during API. The advantages and disadvantages of each method are summarised in Table 1.
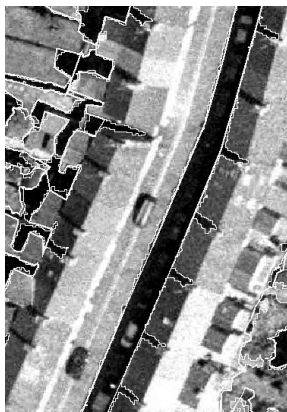
Fig. 3.2 eCognition's segmentation of shadows cast by buildings and chimneys

Table 1. The advantages and disadvantages of using either API or eCognition to delineate real-world objects from remotely sensed imagery

	Advantages	**Disadvantages**
API	♦ Interpretation of real objects ♦ Identification of complex patterns and complex situations ♦ Ability to include or ignore features intelligently ♦ Multi-scale representation ♦ Use of shape, context, neighbourhood relationships	♦ Subjective ♦ Time consuming ♦ A fixed scale is necessary ♦ Inconsistency in the use of a steady scale to the whole image ♦ Human error ♦ Imprecise boundary delineation
eCognition	♦ Objective (the rules and chosen parameters are subjective but the rules are applied to the whole image objectively) ♦ Multi-scale representation ♦ Hierarchical connection between multi-scales ♦ Use of shape, context, neighbourhood relationships ♦ Transferable rules: boundaries reproduced automatically across different data sets ♦ Quick method	♦ Identification of image objects, not real objects ♦ Inability to include or ignore features intelligently ♦ Fusion of real objects due to spectral confusion

4. Conclusions

In this paper, two approaches for mapping urban land cover (for the purposes of identifying sealed soils) using true colour ortho-rectified aerial photography have been presented. The traditional technique of aerial photo interpretation (API) has been used to compare against new automated methods for boundary delineation, with the use of eCognition software. A quantitative analysis showed a very high agreement between the two methods across a range of different UK urban land use types: 1960s residential, Victorian residential, commercial, and industrial.

Both methods identify features at multi-scales and use shape, context and proximity information. The great benefit of eCognition is that once the user finds the appropriate parameters for a satisfactory segmentation and classification then these can instantly be applied in other areas with similar land cover. Consequently, the automated analysis is objective and quick in contrary to the subjective and very time consuming API.

eCognition's main disadvantage is that it cannot interpret an image as intelligently as a manual interpreter would, mainly because it does not recognise real objects, but identifies image objects, which can be spectrally confused. This can be overcome to an extent by applying fuzzy rules during the classification stage.

The ability to work flexibly on specific parts of the image, allows the user to analyse the image in a way that replicates API. The benefits of the automated approach in conjunction with the high agreement that the quantitative analysis showed, led to the conclusion that eCognition can replace the manual method of on-screen digitisation of aerial photography.

This research study will continue by exploring eCognition's automated classification using the 'membership function' approach. This will prove whether API can be replaced by a completely automated method. In the future automated classification, shadow will be reclassified and assigned to a specific land cover class. This procedure has already been done manually during API but has not been used as only segmentation (and not classification) was compared in this paper. If the work was to be repeated again, the most complicated sample area should be used first in order to find the appropriate values of the segmentation's parameters. In this way, fewer 'mixed areas' will need to be manually classified in order to run the segmentation again in a smaller level by using object domains (applicable in eCognition Professional v5).

References

Baatz M, Schape A (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. AGIT-Symposium Salzburg 2000, 12–23. http://www.definiens.com/documents/publications_earth2000.php

Benz UC, Hofmann P, Willhauk G, Lingenfelder I, Heyen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing, 58, 239-258.

Blaschke T, Strobl J (2001) What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GeoBIT/GIS, 6, 12-17. http://www.definiens.com/documents/publications_earth2001.php

Blaschke T, Lang S, Moller M (2005) Object-based analysis of remote sensing data for landscape monitoring: Recent developments. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiania, Brasil, 16-21 April 2005, http://www.definiens-imaging.com/documents/reference.htm:2879-2885.

Burghardt W, the Working Group Urban Soils of the German Soil Science Society (2004) Soil Monitoring Instruction on Sealed Areas in the European Union, Contribution to the EU Soil Thematic Strategy. http://www.uni-essen.de/bodenkunde/links/eusoil/sealingmonitoring2703.pdf

Burghardt W, Banko G, Hoeke S, Hursthouse A, de L' Escaille T, Ledin S, Marsan FA, Sauer D, Stahr K (2004) TG 5 –SOIL SEALING, SOILS in URBAN AREAS, LAND USE and LAND USE PLANNING. http://www.uni-essen.de/bodenkunde/links/eusoil/EUUrbansoilsapr04.pdf

Caprioli M, Tarantino E (2003) Urban features recognition form VHR satellite data with an object-oriented approach. International Symposium of Remote Sensing (ISPRS) Commission IV Joint Workshop "Challenges in Geospatial Analysis, Integration and Visualization II" http://www.igf.uni-osnabrueck.de/mitarbeiter/schiewe/papers/24.pdf

Cothren J, Gorham B (2005) Automated Feature-Extraction: Software Advances Extract Impervious Surfaces from Satellite Imagery. Earth Imaging Journal, 2, 32-34.

Definiens Professional 5 User Guide (2006) Definiens AG

European Environment Agency (EEA) glossary web-site (2006) http://glossary.eea.europa.eu/EEAGlossary/S/soil_sealing

Frauman E, Wolf E (2005) Segmentation of very high spatial resolution satellite images in urban areas for segments-based classification. In: Anon., ISPRS WG VII/1 "Human Settlements and Impact Analysis" 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas (URBAN 2005) and 5th International Symposium Remote Sensing of Urban Areas (URS 2005), Tempe, AZ, USA, 14-16 March 2005, http://www.definiens-imaging.com/documents/reference.htm:

Grenzdorffer GJ (2005) Land use change in Rostock, Germany since the reunification - a combined approach with satellite data and high resolution aerial im-

ages. In: Anon., ISPRS WG VII/1 "Human Settlements and Impact Analysis" 3rd International Symposium Remote Sensing and Data Fusion Over Urban Areas (URBAN 2005) and 5th International Symposium Remote Sensing of Urban Areas (URS 2005), Tempe, AZ, USA, 14-16 March 2005, http://www.definiens-imaging.com/documents/reference.htm:

Herold M, Guenther K, Clarke C (2003) Mapping urban areas in the Santa Barbara South Coast using IKONOS data and eCognition. http://www.definiens-imaging.com/documents/an/sb.pdf.

Hodgson ME, Jensen JR, Tullis JA, Riordan KD, Archer CM (2003) Synergistic use of Lidar and color aerial photography for mapping urban parcel imperviousness. Photogrammetric Engineering and Remote Sensing, 69, 973-980.

Hofmann P (2001) Detecting urban features from IKONOS data using an object-oriented approach. Fist Annual Conference of the Remote Sensing & Photogrammetry Society, 12-14 September 2001, 28-33. http://www.definiens.com/documents/publications_earth2001.php

Meinel G, Hernig A (2005) Survey of soil sealing on the basis of the ATKIS basic DLM − feasibilities and limits. 10th International Conference on Information & Communication Technologies (ICT) in Urban Planning and Spatial Development and Impacts of ICT on Physical Space, 22-25 February 2005. http://www.schrenk.at/corp/archiv/papers/2005/CORP2005_MEINEL_HERNIG.pdf

Mittelbeg B (2002) PIXEL VERSUS OBJECT: A method comparison for analysing urban areas with VHR data. http://www.definiens-imagining.com

Wang Z, Wei W, Zhao S, Chen X (2004) Object-oriented Classification and Application in Land Use Classification Using SPOT-5 PAN imagery. IEEE, 3158-3160.

Wood GA, Kampouraki M, Braganza S, Brewer TR, Harris JA, Hannam J, Burton RJ (2006) The application of remote sensing to identify and measure changes in the area of soil prevented from carrying out functions of sealing. Technical report of GIFTSS project BNSC/ITT/54. A report prepared for the Department of the Environment, Food and Rural Affairs.

Yan F, Bauer M (2006) Mapping impervious surface are using high resolution imagery: a comparison of object-based and pixel classification. American Society of Photogrammetry and Remote Sensing (ASPRS) Annual Conference, Reno, Nevada, 1-5 May 2006. http://www.definiens.com/documents/publications_earth.php

Zhou Y, Wang YQ (2006) Extraction of impervious surface area using orthophotos in Rhode island. American Society of Photogrammetry and Remote Sensing (ASPRS) Annual Conference, Reno, Nevada, 1-5 May 2006. http://www.definiens-imaging.com/documents/reference.htm:

Chapter 6.3

Object-based Image Analysis using *QuickBird* satellite images and GIS data, case study Belo Horizonte (Brazil)

H. J. H. Kux, E. H. G. Araújo

INPE – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brazil; {hermann, araujo}@dsr.inpe.br

KEYWORDS: *QuickBird* satellite, High-resolution satellite data, Object-based image analysis, Urban planning, GIS, Belo Horizonte.

ABSTRACT: Brazil, like many other developing countries is demanding modern IT tools that would allow to monitor, to analyze and to intervene in the city planning process. Urban areas, especially metropolitan regions like Belo Horizonte (Minas Gerais State Capital, Brazil) are highly complex and its' diagnosis is quite difficult. With the advent of high-resolution satellite systems, such as *QuickBird*, object-based image classification techniques became efficient procedures for the analysis and mapping of land use/land cover. This new image classification paradigm uses context relations, hierarchy and fuzzy logic. A multi-temporal analysis, considering also ancillary data, allows spatial inferences generating information to subsidize urban planning. In this study, spatial inferences were made for two quarters of Belo Horizonte using *QuickBird* ORStandard scenes from 2002 and 2004, working with object-based image classification techniques and considering geological, geotechnical and urban legislation data on a GIS. At both scenes ortho-rectification was done using a rigorous model. Following information were generated considering the growth of two quarters (Belvedere and Buritis) in the timeframe considered: a land use/land cover map, detection of irregular land occupation, of areas with risks for slope slipping/erosion as well as an analysis of potential damage to population and property.

1 Introduction and problem setting

Remote Sensing and GIS (Geographic Information System) due to its cost/benefit ratio and to its advanced technology, are being increasingly used to generate relevant information to support decision-taking for a wide range of urban applications. *QuickBird* is one of the new high resolution satellites, whose data have quality and precision for urban applications (Yang 2003). In order to use the full potential of orbital images with high and very high spatial resolution, an adequate mathematical model or a tri-dimensional interpolation function, based on the sensor geometry and orientation, is necessary (Büyüksalih et al. 2004). According to Toutin (2004), images without geometric correction, contain a large and significant amount of distortions that would not allow a direct superposition to cartographic data in a GIS. Considering the requirement of geometric correction, an ortho-rectification procedure (Cheng et al. 2003 and Digital Globe 2004) was used for the two satellite scenes analyzed. Considering the RMSE found after the ortho-rectification (1,05 and 0,86 m respectively for the scene of 2002 and 2004), the scale of work of this study is 1:5,000, and these results are in accordance with the requirements established by the Brazilian Cartographic Society, referring to the Cartographic Exactness Standard (PEC-A) for this scale (Brasil, 1984).

To explore the richness of data delivered by high resolution sensors, a "bridge" is necessary between those approaches already established for visual interpretation that join the hierarchic relations of the interpretation from basic image elements with digital image processing (Herold et al. 2003). Object-oriented approaches, based on multi-resolution segmentation, hierarchical nets, fuzzy membership functions, as well as cognition elements, showed high efficiency to discriminate a large amount of targets represented in high-resolution images. The use of the "Object" concept is essential in such an approach of image analysis, because one assumes that the semantic information needed for the interpretation of an image is not present in the pixel, but rather in the image objects and its' interrelations (Definiens, 2004). Being so, the contextual information can be described and used mainly in two forms: (1) in a spatial context where the neighbor entities are described in a tree with horizontal or vertical direction, and (2) in a semantic context which allows to group those classes which have similar semantic characteristics (Hofmann and Reinhardt 2000; Thomas et al. 2003).

The main characteristic of segmentation based on multiple resolutions, is the option to segment the same image at different scales related to each other, forming a hierarchical net which is the knowledge basis for the object classifications (Baatz and Schäpe 2000; Benz et al. 2004). The object-

based classification allows the user to define complex rules based on spectral characteristics and in inherited spatial relations. Objects can be defined and classified by the structure and behavior of similar objects (Blaschke et al. 2000; Definiens 2004; Kux and Araújo 2006). The classification based on fuzzy functions transforms the values of attributes from an arbitrary interval to a defined interval between 0 and 1, indicating the membership of an object to a specific class. So each object can have a fuzzy membership to more than one class, expressed by its degree of adequateness to the descriptors of these classes (Blaschke et al. 2000, Definiens 2004). The major advantage of this approach is the expression of uncertainness of membership and knowledge (Bock et al. 2005).

As a contribution to urban planning and to sustainable development, the objective of this study is to evaluate (1) the performance and characteristics of object-based image classifications, using two *QuickBird* OR Standard scenes of two quarters from Belo Horizonte (Belvedere and Buritis), (2) spatial inferences based on the integration of these classifications and other data within a GIS.

2 Brief description of test sites

Taking into account that both scenes available for the study are from recent dates within a short timeframe (2002 and 2004), the choice of quarters Belvedere and Buritis was done mainly because both suffered very strong changes in this period. Besides that, these places have a high occupation density and a net of streets which are unable to allow good traffic conditions due to the rising number of vehicles. In both quarters there is an increase of soil imperviousness as a consequence of constructions growth in parallel to the reduction of green areas, changing the original landscape. These problems indicate the need of a constant monitoring as well as an analysis on the evolution of the occupation from these quarters, considering both technical and legal aspects of relevance for urban planning.

Concerning the geological/geotechnical and relief aspects, both quarters are located on a succession of meta-sedimentary rocks (phyllites, schists) with general NE-SW strike and 45° SE dip (Silva et al. 1995). At the slopes dipping to SE, the geotechnical conditions are poor, especially for shallow foundations of constructions. Referring to earth cuts, there is a great probability of earth slipping along the schistosity plan. Due to these reasons, those slopes which coincide with the direction, and dip or schistosity angle of the geological layers, present a lower stability, especially in areas of phyllite outcrops. At Buritis these issues are aggravated due to

strong relief and high altimetric amplitude, favoring earth slipping. Silva et al. (1995) recommend a preferential settlement model at slopes turned to NW, and buildings interspaced with large empty spaces. Nevertheless, these recommendations have not been followed and one verifies presently a high population density with very close buildings.

3 Object-based image classifications

After the geometric correction and fusing, images were cut and imported to the classification software used (eCognition). An urban cadastre was also imported and used as an auxiliary thematic layer. The following procedures were considered: correction/adjustment of images and cartographic data used as a base, definition of land cover classes, image segmentation, determination of class hierarchy, definition of membership rules and evaluation of classification results.

The definition of land cover classes started with the visual interpretation of the satellite images and aerial ortho-photos available. Interpretation keys were elaborated from the elements color, texture, shape, size, shadow, height, pattern and localization. The classes defined for mapping and its' main characteristics are at Table 1.

Table 1. Classes defined for the study

CLASSES	DESCRIPTIVE CHARACTERISTICS
Asphalt	Possibility to be mapped with the urban cadastre.
White	Highly bright. Constituting materials not discernible. Various forms.
Gray Cover	Impervious. Covers of high buildings. Many variations in tone and brightness.
Flare	Quantization level close to 2048.
Swimming pools	High response in the blue channel, sometimes in the green, regular geometric forms.
Ceramic tile	Linear borders. Pattern according to legislation. Variable Geometry. Response in the red channel. Large tonal variation.
Bare soil	Modified terrain (earth works). No vegetation. Response in the red channel. Irregular forms. Tonal Variations.
Shadow	Low brightness. Closeness to high buildings and arboreal vegetation.
Arboreal vegetation	High response in near infrared. Texture caused by shadow of different height of trees. Variable forms. Sometimes covering streets.
Grass vegetation	High response in near infrared. Uniform texture. Response in the red channel due to bare soil presence. Variable forms.

The analysis of characteristics from each class and the respective interpretation keys indicate initially the attributes to be used in the classification by fuzzy membership rules.

3.1 Segmentation

The segmentation is the first step for the classification, and the objects are created from it. The first segmentation, afterwards called "Level 2", separated different objects of interest from the urban environment, specifically buildings. Here color and smoothness were prioritized in detriment of form and compactness. So objects with smoother linear borders, such as houses and buildings were considered. The results of this segmentation were considered to perform the classification that was used in the following procedures (spatial inferences). It is recommended to create the level in which to perform the classification first and those levels used to help the classification afterwards. This is important to avoid the risk of having the most important object level negatively influenced by sub- or super-objects (Definiens, 2004).

At the second segmentation, later designed as "Level 3", parameters which prioritize the form were used, and large objects were searched, like those of the polygons of the thematic data from the urban cadastre. At Level 1, the third segmentation performed, the scale threshold was reduced in order to obtain smaller objects than those of interest, which allowed the creation of descriptors based on texture. Levels 1 and 3 were established to help on the definition of these rules and class limits, allowing that the objects of interest recognize its' super-objects (Level 3) and its' under-objects (Level 1). After the classification, the objects of these auxiliary levels were used for the definition of classes of interest.

3.2 Hierarchical net

The hierarchical net allows the relation between classes and its groups. So a net was developed where those objects which are more easily distinguished have got a higher position than those presenting higher confusion. Figure 1 shows the generic hierarchical net, where one can find the classes of interest as gray rectangles. Other classes were introduced in the process as support to the generation of membership rules. The objects were divided, *a priori,* as belonging or not to the classes of higher separability. Doing so, the creation of rules was facilitated and the characteristics of just one class were investigated. The hierarchy was developed as a consequence of non-association of objects to higher classes. This was done so that classes presenting higher

confusion remain at the lowest levels of hierarchy, reducing the classification error and evidencing the limitation of the process.

One of the most significant advantages of the hierarchical net is the fact that it can be used in other classifications or similar studies. This saves a lot of time and allows to focus in the membership rules (Item 3.3 – Table 2) that usually are different for each project. In this study, for instance, the hierarchical net developed for the first classification was used at the other classifications with just a few changes.
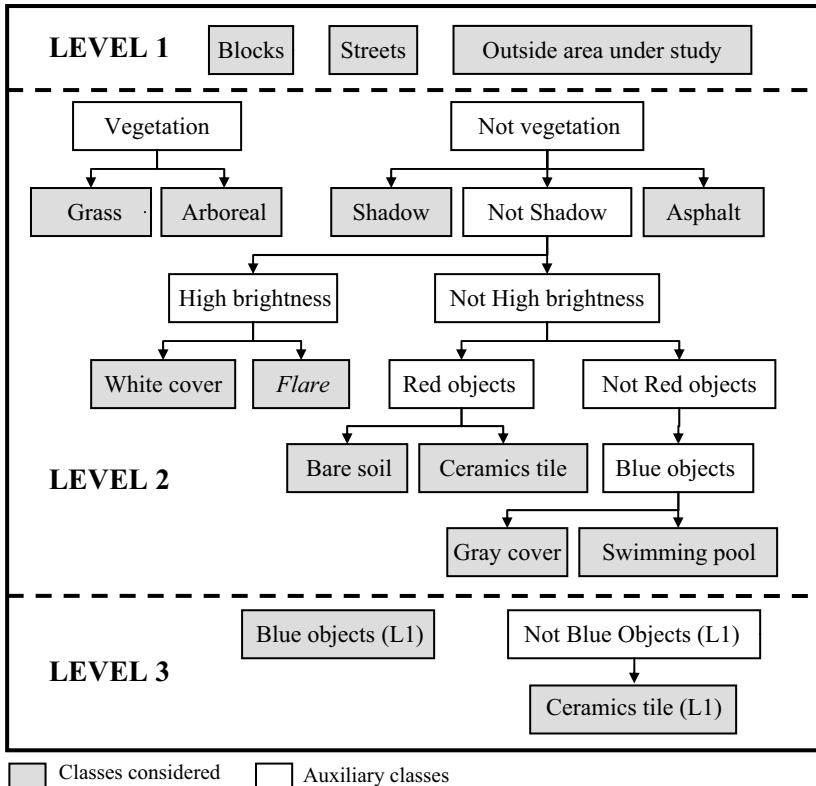
Fig. 1. Generic hierarchical net applied.

3.3 Membership rules

The definition of membership rules has a high weight at the performance of a classification, because here the attributes used by the classifier are defined. Initially the behavior of objects was visualized, and related to a certain spatial attribute by means of a gray level image generated with the

eCognition package. This spatial viewing was done for several attributes, and so it was possible to identify which of these attributes could be used for the separation of classes. After this investigation and in order to identify the limits of separation among the classes with higher accuracy for each membership rule considered, samples from each class were collected. Histograms from these samples were used to limit the curve of each attribute for the classes represented and to separate them with this attribute. We attributed the least possible number of rules for each class in order to avoid the increase of confusion for the classifier using the most efficient rule capable to discriminate this class. Table 2 presents the description of the rules defined for the classes of interest at Level 2 (Fig. 1) for the classification of quarter Belvedere at the 2004 image. Further more detailed information about the rules and the hierarchical net of other classifications are found in Araújo (2006).

Table 2. Membership rules used for the classification of the 2004 image, quarter Belvedere.

CLASSES	RULES APPLIED
Arboreal vegetation	VEGETATION; Brightness, Vegetation index; Texture: average of sub-objects at band 4.
Grass vegetation	VEGETATION; Brightness; Vegetation index; Texture: average of sub-objects at band 4; Inexistence of "Asphalt" as neighbor objects.
Asphalt	NON-VEGETATION; Brightness; existence of super-objects "Streets" (from cadastre).
Shadow	NON-VEGETATION; Brightness; Nonexistence of super-objects "out of area".
White roofs	NON-SHADOW; HIGH BRIGHTNESS; Brightness.
Flare	NON-SHADOW; HIGH BRIGHTNESS; Brightness.
Ceramic tile	NO HIGH BRIGHTNESS; RED; Vegetation Index; It is not "Asphalt"; Area relationship with sub-objects "Ceramic tile n1" (*)
Bare soil	NO HIGH BRIGHTNESS; RED; It is not a tile; Bright; It is not Asphalt; Ratio of band 3 by all others; Vegetation Index.
Gray Cover	NOT RED; BLUE; Bright; Vegetation Index; Relation of area with sub-objects "blue (n1)"; It is not Asphalt; It is not a Swimming pool.
Swimming pool	NOT RED; BLUE; Bright; It is not Asphalt; NDVI; Ratio of bands 3 and 1 by all others; Area relation of sub-objects "blue (n1)".

(*) Objects classified as Ceramic tile in "Level 1".

The detached classes (HIGH CASE) at Table 2 were introduced in the process to compose a hierarchy (Fig. 1) that contains some rules inherited from hierarchically lower classes. So the lower classes have two sets of rules: those developed specifically for that class and those inherited from upper classes. Furthermore Table 2 shows the importance of previous visual interpretation keys. As mentioned before, it was this analysis that gave us

the knowledge and direction for further procedures: spatial viewing of attributes and definition of the membership rules and its curve limits. It is important to mention that although the hierarchy used was almost the same for all classifications, the membership rules had to be adapted or edited in order to optimize the results for the classification of each image. Different off-nadir angles, acquisition dates or atmospheric conditions change the characteristics and the Digital Numbers (DN) of the image, also for scenes of the same area. In that case, the fuzzy interval used in the rules also changes and significant modifications have to be made in the attributes of classes.

3.4 Classifications

In this section the most significant results obtained with the object-based classification, as well as its' peculiarities, challenges and perspectives are emphasized. To evaluate the classification quality, the *Kappa* index was used. For this analysis, 30 samples per class were collected and considered as ground truth. Only class "Flare" had a reduced number of test samples due its low representation. Besides that, aiming to determine the uncertainness of classification, a map of stability was generated, considering the difference among the highest and the second highest level of membership of an object. It must be emphasized that an object that was classified with a high level of ambiguity does not mean that its' classification is erroneous, but that it does not belong specifically to a certain class.

The first classification was done from quarter Belvedere, using the 2004 image, which became the base for all further image classifications. Due to that, and aiming to evaluate the possibilities of automation of the process, many procedures of this classification were repeated in the following ones. Major difficulties were found for the discrimination between Arboreal and Grass vegetation and between Ceramic tile and Bare soil. In order to reduce the confusion between the two classes at the first case, a rule was introduced which differentiates both classes with the number of sub-objects included in the targets (objects) of interest (Fig.2). It is referred here to a texture attribute of possible implementation due to the introduction of a lower level as that one of interest. Frequently arboreal vegetation has a higher number of sub-objects by the presence of shadow between the leaves, allowing for more texture. Figure 2 illustrates visually, at the indicated spots, the progress of separation among arboreal and grass vegetation, using the texture rule.
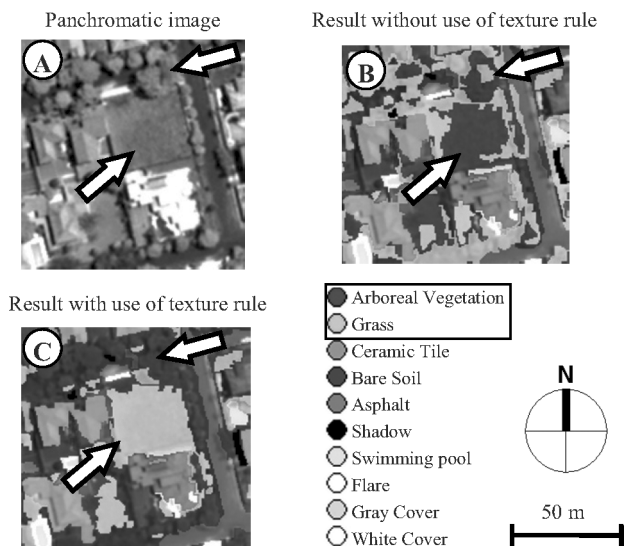
Fig. 2. Example of classification improvement by adding a texture rule showing the Panchromatic image (A), Classification without the rule of texture (B) and the classification with texture rule (C)

The use of texture, as shown in the example above, is an important step to separate arboreal from grass vegetation, particularly in tropical countries like Brazil, where vegetation grows very fast.

As for the discrimination between classes "Bare Soil" and "Ceramic tile" it is noteworthy that both are constituted by clay materials with very similar spectral responses as those in the visible part of the spectrum. During the analysis of these classes, we found out that in average, the samples of class "Ceramic tile" presented higher vegetation index values (ratio of band 4 by 3) than those of class "Bare soil". This can be due to a small musk-like vegetation on the older tiles. Taking into account that a large amount of the houses with this type of cover is older than a few years, it is quite probable that this type of vegetation installed itself there. Furthermore, the satellite scenes used are from the rainy period, which contributes to the proliferation of this tiny vegetation type. Since the bare soil of this region is a product of earth works, all vegetation was eliminated for the construction of buildings. Then a rule was set up that limited the response of the objects, referring to the vegetation index, which reduces the confusion among the two classes.

The results of the *Kappa* values per class for all four classifications are presented at Table 3. According to Table 3, the results obtained are very

good. The values of *Kappa* total index for the four classifications vary between "Very Good" and "Excellent", in accordance with the reference table developed by Landis & Koch (1977).

Table 3. Results of total and per class *Kappa* indices for the classifications of both quarters

Kappa	Asphalt	White	Gray	Flare	S. pool	Ceramics	Soil	Shadow	Arboreal	Grass
Belvedere 2004 image classification										
Class	1	1	0,96	0,60	0,93	1	0,78	0,89	0,53	0,77
Total										0,86
Belvedere 2002 image classification										
Kappa	Asphalt	White	Gray	Flare	S. pool	Ceramics	Soil	Shadow	Arboreal	Grass
Class	1	1	0,60	0	0,78	0,93	0,56	0,96	0,42	0,61
Total										0,75
Buritis 2004 image classification										
Kappa	Asphalt	White	Gray	Flare	S. pool	Ceramics	Soil	Shadow	Arboreal	Grass
Class	0,95	0,98	0,89	1	0,63	0,49	0,93	1	0,84	0,84
Total										0,84
Buritis 2002 image classification										
Kappa	Asphalt	White	Gray	Flare	S. pool	Ceramics	Soil	Shadow	Arboreal	Grass
Class	1	0,82	0,92	1	0,60	0,52	0,70	1	0,63	0,77
Total										0,77

At the Belvedere 2004 classification, the *Kappa* values per class are very high, except for classes "Arboreal vegetation", "Grass vegetation" and "Flare". As for "Grass vegetation" and "Arboreal vegetation" the low result is due to the confusion among these classes, although it became smaller with the addition of the texture rule mentioned. Referring to class "Flare", a tenuous difference was verified among some objects with high brightness and other saturated ones. This occurs because in the neighborhood of saturated pixels (Flare) there were always high brightness pixels, influenced by them, making its separation more difficult. It is important to notice that class "Asphalt" got 100% correctness due to the shared information about the urban cadastre contained at Level 3. If no ancillary data (urban cadastre) and/or no geometric correction would be available, the results for "Asphalt" would be significantly worse because it is difficult to separate this class from dark cement and asbestos. Referring to the analysis of uncertainness, the map on stability of classified objects indicated that the less stable objects are those related to vegetation classes ("Arboreal" and "Grass"). This was already expected because of the great confusion among these classes.

At the Belvedere mapping with the 2002 image, Table 3 indicates that the individual and total values, even though very good, are smaller than those obtained with the 2004 image. This occurred mainly among those classes with higher confusion, namely at "Ceramic tile" and "Bare soil". Only class

"Asphalt" maintained the same value. Other results probably would have been obtained if other thresholds and attributes were tested to increase the class separation. It is noteworthy that class "Flare" presented result zero at the individual *Kappa* index. Besides the factors mentioned at the analysis of this class at the 2004 image, here it is possible that there are no objects with the same characteristics like those mapped at the 2004 scene. The analysis of classification stability of this image came close to the 2004 one, with a larger amount of non-classified objects. It indicates once more the need to test specific fuzzy functions intervals for each classification.

The classification evaluation of the 2004 image from the Buritis quarter showed a better discrimination of the vegetation classes when compared to the results obtained for Belvedere, due to the inclusion of a new attribute (ratio of band 4 by 3) for the distinction between arboreal and grass vegetation. One observes also a decrease of values at class "Swimming pool", probably because this class is not strongly represented in this quarter. The stability map presented a lower number of ambiguous objects when compared to the results of the Belvedere images. This occurred because here the membership rules defined more clearly the objects "arboreal" and "grass".

The analysis of classification results from the 2002 Buritis image shows a lower performance when compared to 2004, similarly to Belvedere. This result indicates that those rules elaborated for an image of a certain date do not apply necessarily to other images. The stability map for this classification confirms that the uncertainness of vegetation classes was reduced with the addition of a further rule to discriminate them in relation to the Belvedere scenes. As for the other classes, it is noteworthy that although the exactness is a little lower than at the 2004 image, the stability of its objects had no significant change.

As a general observation of the image classification procedure, we verified that the hierarchical net developed for the 2004 image of Belvedere, can be used at further classifications with small changes. Nevertheless the membership rules have to be adapted or even modified. This shows that the interferences from the radiation arriving at the sensor as well as climatic factors (rainy season) and the off-nadir incidence angle must be considered, especially in multi-temporal studies like this one.

4 Spatial inferences

From the information available related to local Geology/Geo-Techniques, Geological risks and Legislation, together with the land cover mapping presented earlier in this study, some spatial inferences were made for both

quarters under study. The objective of these inferences is to point out important questions of interest to local population, and to contribute to urban planning of these quarters and others with similar problems.

4.1 Urban expansion for the timeframe 2002-2004

This analysis was done to quantify the increase of soil imperviousness, in square meters for the timeframe 2002-2004 (Table 4), using the classifications of both dates. Those classes corresponding to impervious areas were calculated for 2004 and the result subtracted from those of 2002. Similarly the green areas were calculated.

Table 4. Areas of land cover classes obtained from 2002 and 2004 image classifications from Belvedere and Buritis

Classes	Areas of land cover classes – m^2			
	Buritis 2002	Buritis 2004	Belvedere 2002	Belvedere 2004
Asphalt	122242,68	134768,53	169990,93	166739,05
White roof	45323,28	52157,16	52663,32	30020,40
Gray Cover	412826,79	444673,83	203189,77	274347,02
Flare	2331,36	631,80	93,60	611.64
Swimming pool	471,60	626,76	2994,12	4590,00
Shadow	58006,08	92256,48	42384,60	52450,56
Bare soil	125866,09	88027,56	242991,73	182394,73
Ceramic tile	9463,32	4576,32	87817,32	99083,88
Arboreal vegetation	430268,43	501247,48	199428,85	179812,81
Grass	551191,36	437422,71	165793,33	181056,25
Total	**1757990,99**	**1756388,63**	**1170494,70**	**1167347,57**
2004 – 2002		**1602,36**		**-3147,13**
Total impervious	**589856,07**	**636175,84**	**513661,34**	**570190,35**
2004 – 2002		**46319,77**		**56529,01**
Total Vegetation	**981459,79**	**938670,19**	**365222,18**	**360869,06**
2004 – 2002		**-42789,60**		**-4353,12**

Analyzing Table 4, one verifies that there is a difference of the totals of classes. This is due to the non-classified objects which vary at each image. The amount of non-classified objects is an important issue due its relationship with the membership rules. It is more likely to have greater non-classified objects when there are several attributes used to discriminate a class. The object has to adequate itself to each rule in order to belong to a certain class.

Referring to the increase of imperviousness of urban soil, demonstrated by the difference between the 2002 and 2004 classes "Asphalt", "White roofs", "Cover" and "Ceramics tile", the results are a proof of densification from these quarters in the timeframe under analysis. It is important to note that the 2004 image presents a larger area of Shadow apparently due to the following reasons: 1) the 2002 image was obtained with a lateral viewing angle so that part of the shadow is covered by the distorted object geometry, 2) there was an increase on the number of high buildings identified in the more recent image. This fact could also explain the reduction of class "Asphalt" in 2004 at Belvedere. Shadow is a factor that changes considerably the results, and it is an element to be considered in urban analysis with remote sensing data.

A process of substitution of one-family houses by vertical buildings (whose cover is normally made of metal, asbestos or cement) due to the high cost of the lot can explain the reduction of the area from "Ceramics tile" at Buritis. Another possibility is that the membership rules created to discriminate classes "Ceramic tile" and "Bare soil" did not reduce the confusion among them. The reduction of the green area and the increase of impervious covers are clear evidences that these quarters are growing disregarding quality of life for the population. This reinforces the need to decision taking by the authorities at the planning agency to reduce the urban densification.

4.2 Irregular occupation

This analysis aimed to verify which constructions were built on inadequate places according to the municipal legislation and to the variable slope steepness. In order to perform this task, the maps of land cover and slope steepness, as well as the information of the present legislation for the city of Belo Horizonte (Law on Parceling, Land Use and Occupation, Belo Horizonte 1996 and 2000) were put together in a Boolean operator using a program developed in SPRING software package (INPE, 2005). Those buildings located in areas with slope steepness above 47% and/or which were positioned within a ZPAM (Zone for Environmental Protection), were pointed out in the inference. Using this fast forward procedure of data integration in a GIS, it is possible to identify built-up areas which disagree with any legal restriction.

4.3 Risks of slope slipping

For this analysis data from geology, slope orientation and steepness were included in a GIS. The last two datasets were derived from a digital terrain

model. A geological map (Silva et al. 2005) and geotechnical descriptions were used to identify the most important characteristics from these themes at both test sites. A Boolean operation was defined and its result detached those areas with slopes directed to SE with high steepness (Figure 3), since according to Silva et al. (1995) these areas are most prone to slope slipping.
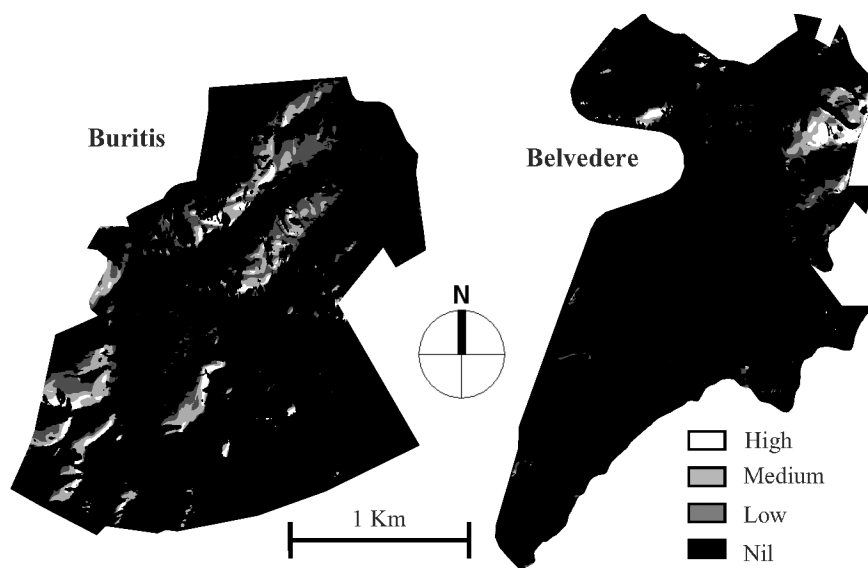
Fig. 3. Map of risks of slope slipping at Belvedere and Buritis quarters

Comparing these maps one verifies that Buritis has got a higher probability to slope slipping due to the predominance of relief characterized by steep slopes and high altimetry amplitudes. This is an important alert for future actions referring to the land use/occupation of such environmentally sensitive areas.

4.4 Hazards to population

In this analysis a large amount of variables and procedures was put together. Four categories of data were incorporated in an analytic-hierarchic process (AHP): critical areas, land use classification, specific legislation and slope steepness. Major importance was given to the category containing the critical areas because, from its information, the hazards to population and property were derived. The Map of Critical Areas is based on 3 maps obtained at PRODABEL (Belo Horizonte City Planning Agency), which contain areas

of geological risks related to excavations, slope slippery and erosion. This information was grouped in just one map with places subject simultaneously to one, two or three geological risks mentioned.

Each one of the four categories incorporated to the AHP process was compared to the others, establishing a hierarchy and using a decision support tool, to attribute weights to each information source. After defining a hierarchy among the categories, a program in LEGAL (a general spatial computer language used in SPRING software package – INPE, 2005) was generated and edited with the values of the weights from each of the classes proposed. After edited and executed, the program produced a weighted numeric map. From this numeric grid an image was generated, afterwards associated to a variation in gray levels, which represent those places with low (black) and high (white) susceptibility of hazards to population (Fig. 4).
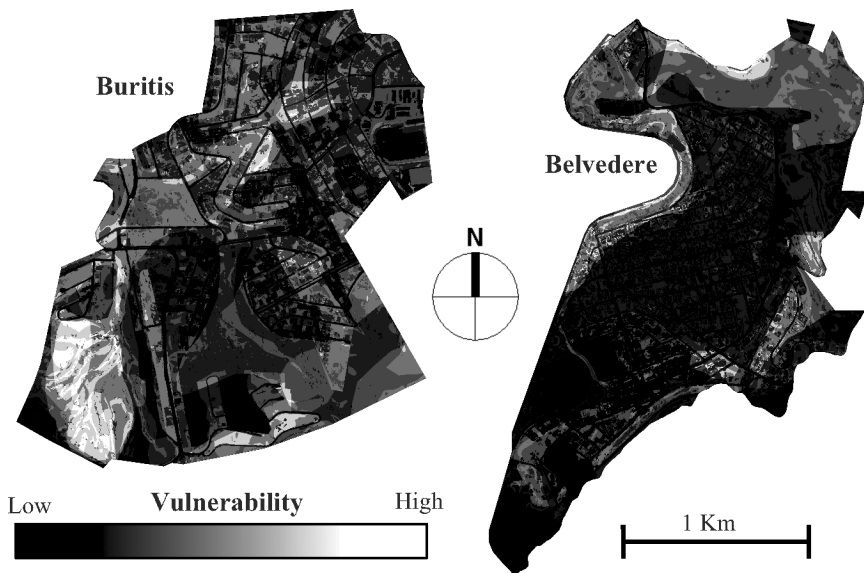
Fig. 4. Map of environmental vulnerability for population and property

The places of interest in this inference are those represented by gray levels at the fourth quartile of the gradation bar (Fig.4). Since these areas have a less restrictive legislation (which is an incentive to construct high buildings), unfavorable geotechnical characteristics and high slope steepness, there is an increase of risks to the population. The "Geological risk" here is considered as an association of several activities which allow the occurrence of events that would cause damage to population and property, caused or not by human interference (Silva et al. 1995). Nevertheless this risk is subject to events such as strong rainfall. At Figure 4, one observes

that Buritis has a major number of risk areas than Belvedere, due to its geotechnical conditions and to a less restrictive legislation. Areas with high slope steepness and geological risks are included in a zoning of preferential densification. This is an alert for urban planners, in order to change or create limitations to urban expansion in high risk areas.

5 Conclusions and Perspectives

From the results found we conclude that:

- The evaluation of classifications indicated the possibility to use the land cover map in the following steps of the study, especially for the spatial inferences;
- The study showed the possibility to map areas with tendencies of growth and risk, besides being a reliable information source for urban planning;
- The object-based image analysis methodology is a valid approach for the classification of such large datasets as the *QuickBird* images used in this study. It is advisable to improve this approach by all means, so it becomes an operational tool for e.g. urban planning agencies, as indicated in previous application examples.
  To improve this methodology specifically for urban applications, we strongly recommend the following further studies:
- Use a DSM (Digital Surface Model) for the ortho-rectification of objects and buildings. One should investigate the influence and use of DSM in areas of occlusion due to the satellite incidence angle;
- Evaluate the application of high-resolution images, obtained at several incidence angles. Use preferentially complementary angles, so it is possible to generate stereoscopy to obtain information of land/soil cover in occlusion areas;
- Test the efficiency on the use of the DSM for the distinction of bare soils and ceramic tiles;
- Investigate descriptors that are able to discriminate classes of gray cover like asbestos, metallic roofs and cement, quite frequent in Brazil;
- Map objects of several classes which are in the shadow.

Last but not least, it is absolutely necessary to improve the OBIA approach, considering the "flood" of data which will become available from the large number of high-resolution satellite systems to be launched in the next few years (Ehlers, 2007).

References

Araújo EHG (2006) Análise multi-temporal de cenas do satélite *Quickbird* usando um novo paradigma de classificação de imagens e inferências espaciais: estudo de caso Belo Horizonte (MG). 159 pp., Dissertação (Mestrado em Sensoriamento Remoto) – INPE 2006. (INPE-13956-TDI/1062) available at: http://mtc-m13.sid.inpe.br/col/sid.inpe.br/MTC-m13@80/2006/07.24.19.43/doc/paginadeacesso.html

Baatz M, Schäpe A (2000) Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation. In: STROBL, J. & BLASCHKE, T. (Eds.): Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000. Karlsruhe. (Herbert Wichmann Verlag), pp 12-23

Belo Horizonte (2000). Lei Nº 8137/00. Available at: www.pbh.gov.br/mapas/leiuso/lei-8137.htm

Belo Horizonte (1996). Lei Nº 7165/96. Available at: http://pbh.gov.br/siga/procuradoria/pgmlegis.htm

Benz UC, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing, vol. 58, n.3-4, pp 239-258

Blaschke T, Lang S, Lorup E, Strobl J, Zeil P (2000) Object-oriented image processing in an integrated GIS/Remote Sensing environment and perspectives for environmental applications. vol. 2, pp 555-570, Environmental Information for Planning

Bock M, Xofis P, Mitchley J, Rossner G, Wissen M (2005) Object-oriented methods for habitat mapping at multiple scales – case studies from Northern Germany and Wye Downs, UK, Journal for Nature Conservation, vol. 13 (Nr. 2-3), pp 75-89

Brasil.(1984) Decreto Nº 89.817/84 Regulating Instructions of Technical Standards for National Cartography. Available at: http://www.concar.ibge.gov.br/index7a0.html?q=node/41

Büyüksalih G, Murat O, Jacobsen, K, (2004) Precise georeferencing of rectified high-resolution space images. In:XXth ISPRS Congress. Proceedings Istanbul: International Society for Photogrammetry and Remote Sensing vol. 35, pp 184-188.

Cheng P, Toutin T, Zhang,Y, Wood M (2003) QuickBird - geometric correction, path and block processing and data fusion. EOM, vol. 12, nr. 3 pp 24-30

Definiens (2004) E-Cognition: User Guide 4. Germany pp 486

DigitalGlobe Inc (2004) QuickBird imagery products: product guide. Longmont, Colorado

Ehlers, M (2007) Sensoriamento Remoto para Usuários de SIG –Sistemas e Métodos: entre as Exigências do Usuário e a Realidade. In: Blaschke, T. &

Kux, H. (eds) Sensoriamento Remoto e SIG avançados, 2nd edition. Oficina de Textos Ltda, São Paulo, pp 19-38

Herold M, Liu X, Clarke KC (2003): Spatial Metric and Image texture for mapping urban land use. Photogrammetric Engineering & Remote Sensing, 69, pp 999-1001.

Hofmann P, Reinhardt W (2000) The extraction of GIS features from high resolution imagery using advanced methods base don additional contextual information – first experiences. Proceedings, vol. XXXIII, pp. 51-58, Amsterdam International Archives of Photogrammetry and Remote Sensing

Instituto Nacional de Pesquisas Espaciais (INPE) (2005) SPRING: Sistema de Processamento de Informações Georreferenciadas. São José dos Campos, Available at: http://www.dpi.inpe.br/spring/portugues/download.html.

Kux HJH, Araújo EHG (2006) Multi-temporal object-oriented classifications and analysis of QuickBird scenes at a metropolitan area in Brazil (Belo Horizonte, Minas Gerais State), 1st International Conference on Object-based Image Analysis (OBIA, 2006), Conference Proceedings [CD], vol. Nr. XXXVI – 4/C42, Salzburg/Austria

Landis JR, Koch GG (1977) The measurement of observer agreement for categorial data. Biometrics, vol. 33 n.1, pp. 159-174

Silva AB, Carvalho ET, Fantinel LM, Romano AW, Viana CS (1995) Estudos geológicos, hidrogeológicos, geotécnicos e geoambientais integrados no município de Belo Horizonte. Belo Horizonte, MG, Secretaria Municipal de Planejamento

Silva AB, Carvalho ET, Fantinel LM, Romano AW, Viana CS (1995) Mapa Geológico de Belo Horizonte. Belo Horizonte, MG, PMBH, (Escala 1:25.000).

Thomas N, Hendrix C, Congalton RG (2003) A comparison of urban mapping methods using high-resolution digital imagery. Photogrammetric Engineering & Remote Sensing, vol 69 (9), pp 963-972

Toutin T (2004) Review article: Geometric processing of remote sensing images: models, algorithms and methods. International Journal of Remote Sensing, vol 25 n. 10 pp 1893-1924

Yang X (2003) Remote Sensing and GIS for urban analysis: an introduction. Photogrammetric Engineering and Remote Sensing, vol 69, nr 9, pp 937-939

Chapter 6.4

An object-based approach to detect road features for informal settlements near Sao Paulo, Brazil

R. A. A. Nobrega[1,2], C. G. O'Hara[1], J. A. Quintanilha[2]

[1] GeoResearch Institute, Mississippi State University, Starkville-MS, USA.

[2] Polytechnic School of Engineering, University of Sao Paulo, Sao Paulo, Brazil. rodrigo.nobrega@poli.usp.br

ABSTRACT: Continuous mapping efforts have been required to monitor intense urbanization processes of large cities of developing countries. The uncontrolled sprawl occurring in the vicinity of Sao Paulo, Brazil since the 70's illustrates this scenario. Considering that urban sprawl causes changes to road networks, monitoring new roads as well as the changes along existing roads can provide significant information for urban management. Due to the lack of coverage in historical and accurate aerial image and map products, existing or even outdated image data are unavailable for planning urban land use with substantial relevance. The recent availability of high resolution satellite images, beginning with the IKONOS II in 1999, has enabled urban applications of Remote Sensing. Unfortunately, traditional techniques employed to detect land cover information based on per-pixel analysis have yielded unsatisfactory results in urban application of high resolution satellite images. In this sense, enhanced capabilities and successful applications of object-based classification have stimulated research to develop new methodologies to provide geoinformation. To this end, road extraction research has been formulated to segment object-primitives from images and to use the resultant informa-

tion to devise enhancements to improve the road detection and classification process.

This chapter reports the use of object-based image classification applied to road detection in informal settlements areas. An 11-bit IKONOS image was employed as the primary remote sensing data for classification. Principal components and semi-homogeneous segmented area products (segmentation products) were computed and used to define custom features. Auxiliary data were calculated from spectral information in combination with geometric information extracted from segments. Contextual information was also employed to support the implementation of a classification rule base. The classification rule base eliminated vegetated areas and then considered impervious surface and bare soil areas, as well as the width, length, asymmetry and the neighborhood relationship for the objects to detect road features. Comparisons between the automatic approach results and manually extracted road feature areas delivered insight regarding omission and commission error by area counting as well as metrics employed to determine completeness and correctness of extracted road features by linear correspondence analysis attest to the efficiency of the methodology. Results indicate that the methodology produces significant information and offers improvements over traditional pixel-based methods of road extraction and classification.

1 Introduction

Investigations delivering insight as to changes in the urban landscape are of extreme importance for land planning and land management. The uncontrolled sprawl that has occurred in large cities of developing countries presents an obstacle to urban management. (Nobrega 2007). Currently, the Sao Paulo metropolitan region has a population estimated in excess of 17 million people (Barros 2004). The intense urbanization process experienced since the 70's has caused major problems in regional land use. Traditional mapping efforts utilizing aerial imagery and periodic updates could potentially provide the best technical solution, however these approaches carry with them very high effective cost due to the continuous map updates necessitated by the emergence of buildings, roads as well as new settlements surrounding these cities. However, the recent developments in satellite remote sensing and sensors provide data products that deliver frequent revisits as well as improvements in spatial, spectral and radiometric resolution; thereby stimulating the formulation of new methodologies for land cover and land use classification which employ

these satellite-delivered data products. Coupled with improved classification methodologies, this next generation of commercial satellite image data products, the promise of precise monitoring and mapping locations of urban development on Earth's surface can be achieved, including the mapping of new settlements surrounding large cities.

According to Quintanilha & Silva (2005), the large amount of details present at high spatial resolution imagery has created many possibilities to deliver needed geoinformation. On the other hand, the high internal variance of these images becomes a problem for traditional per-pixel classifiers. Considering that information present in a remote sensing scene are fractal in nature (Blaschke & Strobl 2001), the more characteristics - geometric, spectral and topologic- for these objects, the more realistic the classification can be. The extraction of the object-primitives based on segmentation has supported the best results of classification for high resolution images. The object-primitives provide a wide range of information to discriminate different land cover types that greatly exceed information available through the consideration of pixel attributes by themselves.

Comparisons between traditional pixel-based and object-based methodologies have demonstrated the effectiveness of the new technology. Regarding spectral analyses, the separability of urban features remains a challenge due to the large diversity their spectral patterns. Unknown spectral patterns, as well as the high heterogeneity of the urban environment requires the combination of pixel and object information to effectively classify land cover as well as land use (Shackelford & Davis 2003 ; Hofmann 2001 ; Tso & Mather 2001).

This chapter presents research and results describing the development of strategies for detecting and classifying roads in informal settlements by using object information extracted from high resolution satellite images. In these strategies, it is proposed that general classification rules may be created, but it is also clear that these rule-bases must be adapted to the geographic locale due to potential differences in image characteristics as well as particularities of the land cover and land use. Regardless of how easily roads may be visualized from high resolution images, their automatic detection remains a challenge task. Irregular building materials and patterns, occlusions by overhanging objects and shadows, differences in width and length, discontinuities and the high degree of spatial heterogeneity in terms of artificial and natural land cover all contribute to the considerable challenges to automating the process. These difficulties still exist for object-based classification, however, more information and tools are likewise available in object-based classification to address these issues.

Indeed, the results obtained indicate a large potential to provide information about the unplanned urban sprawl from multi-spectral high resolu-

tion images. Considering the current demand for urban planning, these roads can be used as a layer of information on GIS efforts.

2 Study area, data and tools

The research presented was conducted near the urbanized extend of Sao Paulo, in the southeast region of Brazil, close to the latitude 23^0 27' S and longitude 46^0 41' W. The area is characterized by recent and unplanned occupation, where dense urbanization is found intermixed with preserved natural areas. These areas of preserved nature comprise dense tropical forest and moderated to rugged terrain (see Fig. 1) provide excellent conditions for groundwater recharge and conserving the environmental balance, and are being threatened due to excessive deforestation and uncontrolled land use.

The study area is located on the north part of Sao Paulo, where informal settlements have removed the natural coverage and significantly aggravating environmental conditions leading to increased erosion, flooding, and water quality problems. Due to adverse human and environmental conditions in such areas, there is an urgent need to update land use/cover information, conduct environmental impact assessments and map road networks as reported in Repaka et al (2004).
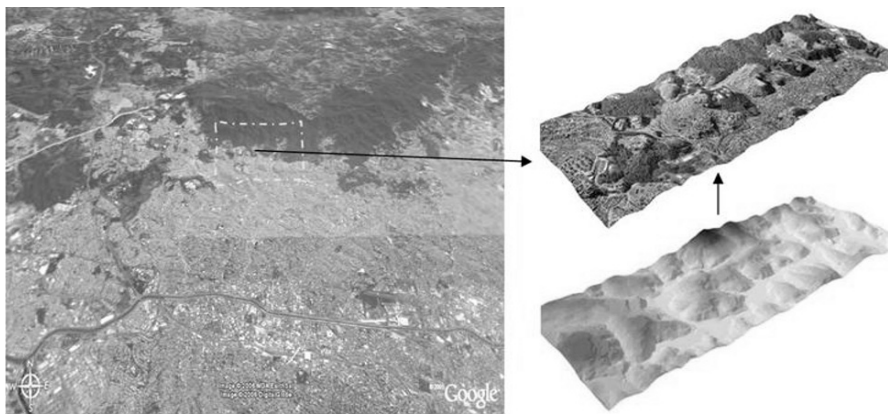
Fig. 1. Overview of north part of Sao Paulo city and the study area. Details show the dense occupation among the remaining natural areas and the rugged terrain

An 11-bit IKONOS image, recorded in 2002 over the north region of the town was used on this research. It is a CARTERRA Ortho Precision prod-

uct compiled from a digital terrain model (DTM) provided by a local agency for mapping. Also, resolution merging of the multi-spectral bands of Ortho Precision products with the panchromatic band resulted in a sharpened one meter resolution multi-spectral image, processed by Space Imaging Corporate.

The technologies and methodologies employed in this research utilized three significant remote sensing and GIS software tools which follow:

- Erdas Imagine, to subset areas of interest and computing the principal components;
- Definiens eCognition, for the segmentation and object-based classification;
- Esri ArcGIS, for mapping as well as comparative analysis.

As for the linear correspondence analysis, Matlab and L-CAT were employed.

3 Methodology

With the objective of detecting roads among many other urban features, the development of this research required sequential steps of image processing including data matching, Principal Components Analysis (PCA), image segmentation, and class definition. With the classes, their strategies of use as well as the classification rules were customized. In short, feature information was mined from segments providing characteristics that included shape, context and spectral characteristics and these characteristics were employed to develop rule bases that were applied in the classification process (Fig.2).
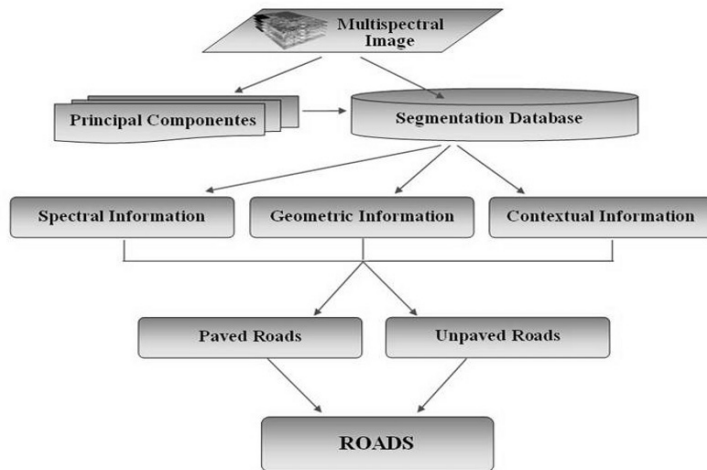
Fig. 2. Generalized workflow scheme used to extract and classify (detect) road features

3.1 Preprocessing

The first steps included stacking layers comprising the multi-spectral bands of individual images and mosaicking the adjacent IKONOS images to provide continuous coverage of the entire study area. Next, the multi-spectral mosaic was subset and reprojected to Corrego Alegre datum. These procedures are fully reported on Nobrega (2007).

Creating image products from the multi-spectral data included processing of principal components for including these results in the image segmentation and subsequent object-based classification. As a mathematical technique to reduce the redundancy of the spectral bands through axis-rotation and translation on the feature space, results of principal components analysis (PCA) have been extensively used in remote sensing to enable the best use of multi-spectral images. In practice, the content of information decreases according to the order of the PC layers. In this work, the first three layers were computed by Erdas Imagine and employed in the classification rules. Note that the PC layers correspond to index images, the range of which is normally greater than the 2048 levels of original 11-bit image.

When conducting object-based image classification, the segmentation process requires special attention. The segmentation process results in homogeneous segments of area which are treated as bounded objects. Con-

sidering roads and related urban features as objects of interest, several trial segmentations were performed to deliver an ideal fit between the derived segments and the targeted urban features. The segmentation used all multi-spectral bands with the same weight. Emphasizing shape more than color characteristics, as described by Pinho (2005), and using a scale parameter of 30 produced objects that generally corresponded to the narrow streets (around 4-6 meters width), typical found in informal settlement areas.

Table 1. Parameter adjusted to segment the images in eCognition.

Band	Scale parameter	Color/Shape	Compactness/Smoothness
R, G, B and NIR	30	0.1 / 0.9	0.1 / 0.9

3.2 Classes definition

With the systematic classification of a road system being the overall goal of the classification, significant challenges arise to accurately detect and classify these target features among many other urban features, especially over informal settlement areas. The lack of standardized spatial distribution of roads and buildings, as well as the presence of many different materials covering the land causes enormous spectral confusion during the image classification, even for visual inspection.

To minimize mistakes, the solution employed several classes. Some classes were used to mask out areas not in consideration, reducing the amount of remaining information, other classes were used during the contextual analysis to further define areas of interest. In short, the main auxiliary classes are: VEGETATION; SHADOW; BARE SOIL; UNPAVED ROAD; IMPERVIOUS and PAVED ROAD. Additional classes which contributed to improve the accuracy of the classification were defined for CERAMIC ROOF and METALIC ROOF classes, which used shape information and relative border to shadow. Finally, GAP and TRACK classes were defined using shape information to assist in deriving continuous road features.

3.3 Computing auxiliary layers

The capability for each derived segment to compute and employ object information extracted from the original pixel-based data as well as other layers (such as PCA products or elevation data) is one of the most important concepts related to object-based image analysis. Therefore, using the im-

age data and each object segment, the mean Digital Number (DN) was automatically computed for each original band as well as the PC layers. These new data support the extraction of auxiliary layers as well as the definition of custom features such as vegetation indices.

Due to the difficulty of detecting road features, a first step is suggested that identifies general impervious and bare soil areas to provide candidate areas that may be further classified to extract actual roads. In this way, a masking strategy was implemented that step-by-step decreased the amount of information that was considered for road classification from the full scene.

Initially, vegetated areas were detected though a custom NDVI feature. After masking out the vegetation, shadowed areas were also detected using a custom feature in eCognition. Removing both vegetation and shadow enabled the detection of bare soil areas with high accuracy, as presented by Nobrega et al (2006b). On the other hand, the urbanized areas were estimated by the PC-2 that indicated most part of the impervious surfaces in the study area. In short, impervious areas and bare soil areas were used as reference to detect paved roads and unpaved road, respectively. Table 2 summarizes the customized functions. The auxiliary layers used during the contextual analysis (bare soil and impervious) are shown on Fig. 3.

Table 2. eCognition´s customized functions to detected the auxiliary layers.

Description	Customized Function	Bands/Layers	Threshold
NDVI	$NDVI = \left(\dfrac{\mu NIR - \mu RED}{\mu NIR + \mu RED} \right)$	NIR, Red	up to 0.18
Shadow General Indicator	$SGI = \mu Blue - \mu PC1$	Blue, PC-1	-450 to -400
Bare Soil General Indicator	$BGI = \left(\dfrac{\mu PC3 * \mu Blue}{-100000} \right)$	PC-3, Blue	0.8 to 1.5
Imprevious areas		PC-2	-350 to -320
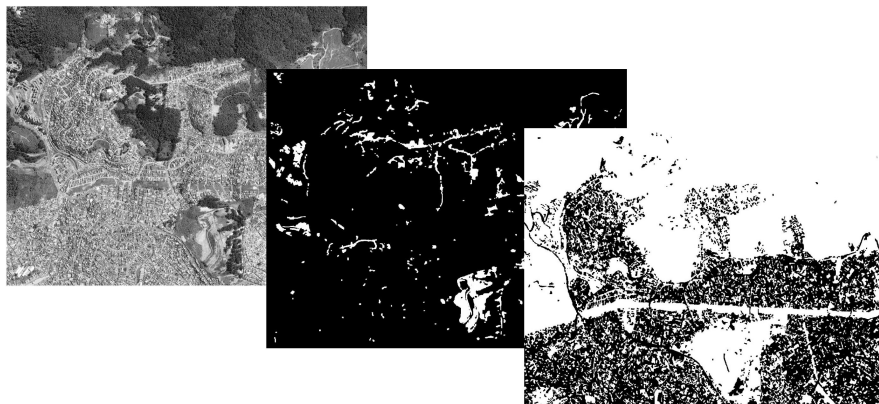
Fig. 3. IKONOS image illustrating the study area (left); bare soil areas in white (middle); and impervious areas in black (right)

3.4 Strategies used on mining roads from images

Considering that almost all features present in the urban landscape are man-made features, information generated from their geometry can be used in the classification process. Since the bare-soil-layer includes the unpaved roads and the impervious layer includes the paved roads, the extraction of road features requires further classification of these parent classes (bare soils and impervious areas) and appropriate use of information mined from the shape of the segments.

Parameters such as area, width, length and asymmetry, plus relative border to other objects where employed, as illustrated on Fig 4. Table 3 presents a combination of shape and contextual parameters used to develop the classification rules.
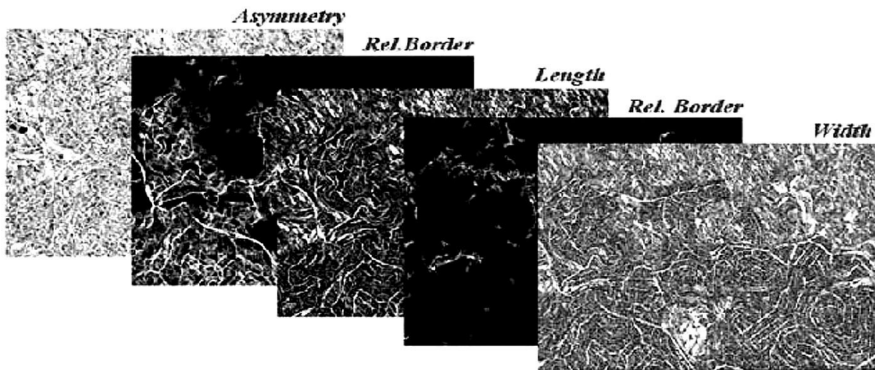
Fig. 4. Parameters extracted from object-shape used on the classification scheme

Table 3. Features used to detect road features.

Class	Parameter	Function	Threshold
PAVED ROAD	Border length		140 to 150 m
	length		25 to 50 m
	length/width		2 to 3
	asymmetry		0.8 to 1
	similarity to IMPERVIOUS		
UNPAVED ROAD	length/width		2 to 3
	asymmetry		0.8 to 1
	similarity to BARE SOIL		
TRACK	area		100 m^2 or more
	rectangular fit		0.6 to 1
	asymmetry		0.7 to 1
	similarity to BARE SOIL		
	relative border to VEGETATION		50% or more
GAP	border length		140 to 150 m
	length/width		1 to 2
	asymmetry		0.7 to 1
	similarity to IMPERVIOUS		
	relative border to IMPERVIOUS		50% or more

Fig. 5 illustrates the general flowchart of the methodology. The classification process required sequential steps to detect and mask out unwanted information. Similarity and non-similarity criteria were used in combination with geometric, spectral and contextual information. The order of the classification reveals the strategy created.
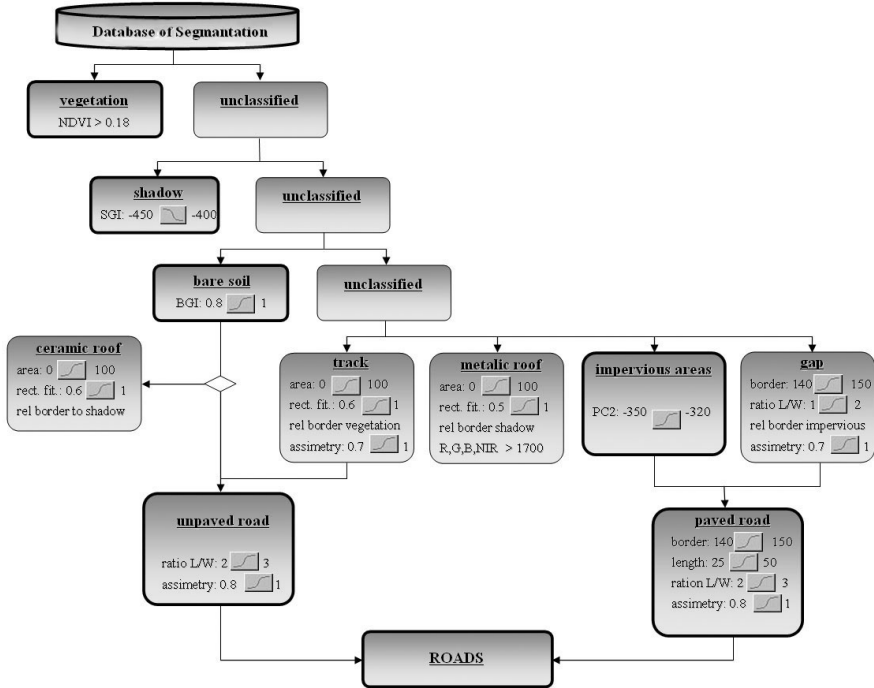
Fig. 5. General classification scheme

4 Results

At first, the results were refined and analyzed by visual inspection to determine the general similarity of classified and observed roads, observed in IKONOS data during the classification stage. When the detected roads and observed roads presented a good matching, the results were exported to a shape file. The extracted results were complemented by manually extracted referent polygons and lines which were analyzed to determine agreement in a GIS program. The detected roads are illustrated in Fig. 6.

Fig. 6. The resultant road system, formed by paved and unpaved streets

In a general view, the road system detected in this work presented a considerable high agreement of match with the actual road system. By draping the detected roads over the very high resolution orthophotos, the agreement was checked through visual inspection for the whole study area. Esri ArcScene was used. This procedure was used as basis for the qualitative analysis of results.
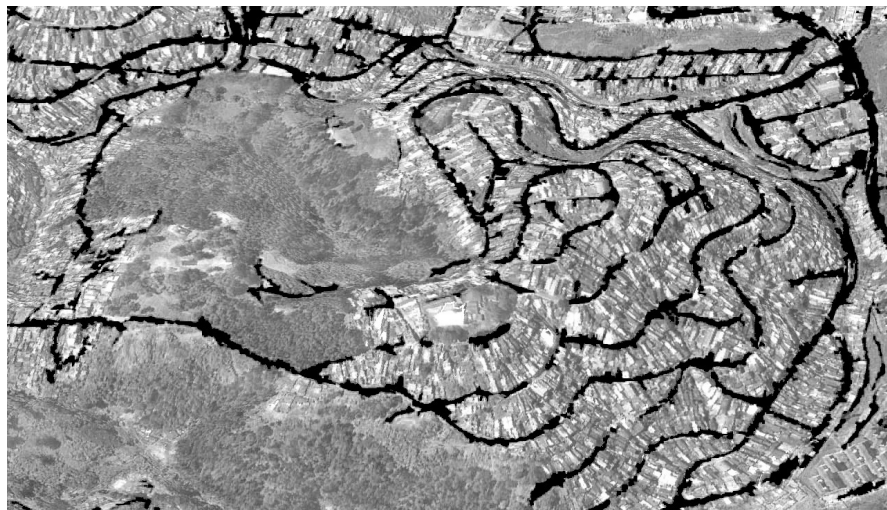
Fig. 7. Detected roads draped over an orthophoto during the qualitative analysis. In detail, the detection of narrow and unstandardized streets formed over the hills

5 Quantitative Analyses

Although the observed results suggest good quality, the quantitative accuracy of the process must be computed so that qualified comparisons may be drawn between the results of this procedure and other procedures as well as to measure improvements that may be obtained by further methodological enhancements. To quantify the accuracy of results, area-based analyses were conducted utilizing feature polygon-comparisons of the detected manually extracted reference roads. The overall accuracy affirmed the expectations but produced significant commission error. To overcome this shortcoming, linear comparative analyses were employed, producing parameters that included completeness and correctness.

5.1 Accuracy, omission and commission errors

In Nobrega et al (2006a) the detected roads were compared with reference ones. Accuracy, omission and commission errors were calculated through GIS logic operations such as *intersect* and *erase* (Fig. 8). The overall accuracy of 64.5% was considered satisfactory due the complexity of the scenario. Omission error explained the 35.5% of roads not detected possibly due to such factors as occluded roads, dirty pavement, trees and cars. The statistical indices used are presented in the equations below.

Resultant Data	Logic Operation from Primary Data
CORRECT	REFERENCE *INTERSECT* EXTRACTED
MISSING	REFERENCE *ERASE* EXTRACTED
WRONG	EXTRACTED *ERASE* REFERENCE
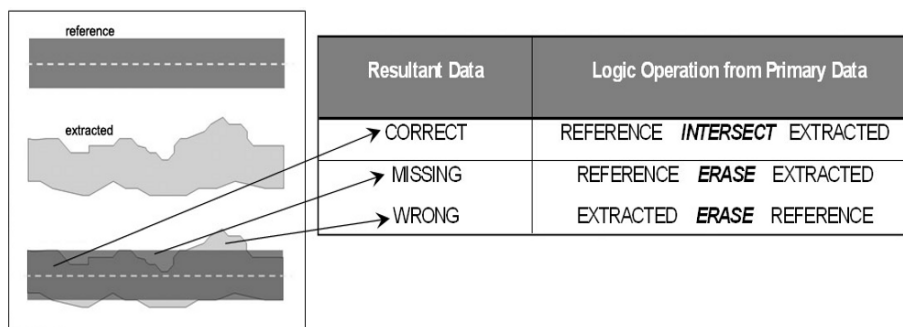
Fig. 8. Spatial logic operations used to compute the accuracy of the extraction. Adapted from Nobrega (2007)

$$ACCURACY = \frac{CORRECT}{REFERENCE} \quad = 64.5\% \qquad (1)$$

$$OMMISION = \frac{REFERENCE - CORRECT}{REFERENCE} \quad = 35.5\% \qquad (2)$$

$$COMMISSION = \left(\frac{EXTRACTED - CORRECT}{REFERENCE} \right) \quad = 76.5\% \qquad (3)$$

Unfortunately, misclassified features contributed to large commission error (76.5%). Disregarding the fact that some of the commission error include features like buildings and parking lots, the largest part of the false-positive roads (commission areas) were found to be sidewalks and overrun features (such as road shoulders) which surround the roads.

5.2 Linear correspondence analyses

Despite the substantially accurate result, the large commission error indicates that area-based methods such as pixel-counting are not an appropriate metric for fully quantifying the accuracy of object-based road classification. To overcome this shortcoming we introduced the linear comparative analysis as presented in Wiedemann (2003) and Seo & O'Hara (2004). The linear comparative analysis is a process to find correspondence between linear features in one data set (as reference) and linear features in the other data set (as extracted). This method typically employs a buffering criterion to establish maximum distance for considering objects that may correspond.

Two main accuracy parameters are produced: completeness and correctness. Completeness is a metric to measure the quality of extraction based on its coverage of extraction over the reference data. Correctness is the percentage of parts of extracted features which are considered as correct in the sense that those parts correspond to the reference feature within a given tolerance (Seo & O'Hara 2004). Fig. 9 (on top) shows the coverage of a reference feature by a buffered region around an extracted feature, used to compute the completeness ; and (on bottom) shows the coverage of an extracted feature by a buffer generated around a reference feature, as used to compute the correctness.
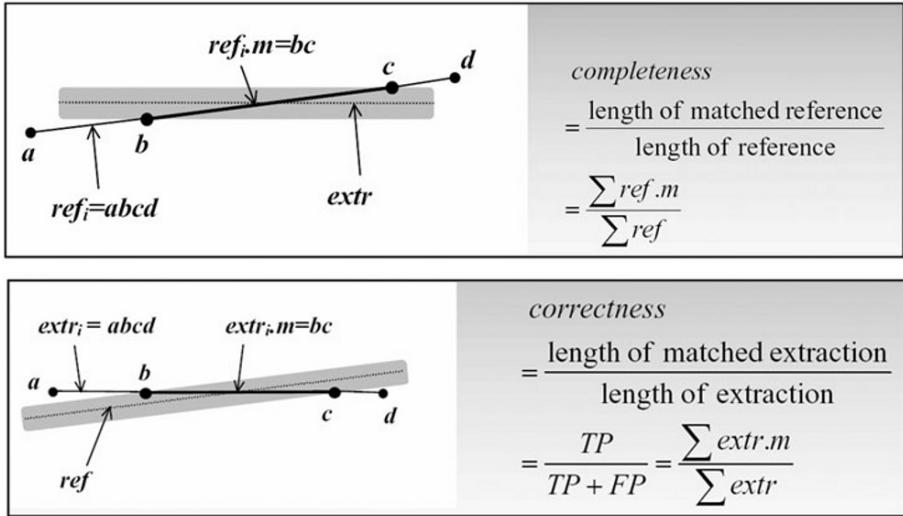
Fig. 9. Precision metrics used to compute the linear comparative analysis. Adapted from Seo and O'Hara (2004)

In this study, the detected roads were converted in lines before computing the linear comparative analysis. Then, the roads resultants from image classification were exported to an image file and imported into MatLab. Since the image presents basically the roads and background, some noise was removed and roads were thinned in an interactive process using mathematic morphology tools present on MatLab's image processing toolbox. Fig. 10 illustrates some roads detected over an informal settlement area and their respective thinned centerlines.
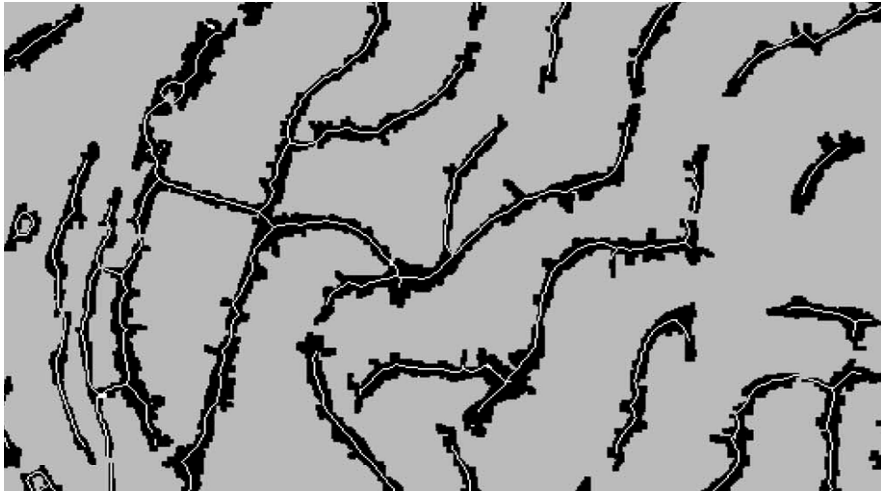
Fig. 10. Roads resultant from the classification process (black) and their respective thinned linear features. Comparing lines prevented excessive commission error caused by differences in width of the real and the detected road

Also, a reference linear database was manually extracted from the orthophotos as road centerlines, due to the lack of up to dated official database for the study area. So, considering the roads centerlines as reference lines and the thinned roads as extracted lines, the linear correspondence analysis was conducted using the Linear Correspondence Analysis Toolkit – LCAT (Seo & O'Hara 2004).

Because the analysis uses buffers, different combinations of buffers were computed ranging from 2 meters as minimum road width to 24 meters as maximum one.

Considering that the average road width is 9 metes in the study area, buffers were adjusted to 9 meters either. Using these 9-meters buffers, linear correspondence analysis computed completeness equals to 67.86% and the correctness equals to 64.48%. Also, the comparative analysis provided another parameter: the gap statistic, which corresponds to percentage of reference roads without correspondence matching. The resultant gap statistics pointed to 36.8 % but was computed under a conservative way, considering 8 meters buffers width due to limitation of the software.

6 Conclusions

The study area presents enormous heterogeneity in terms of spatial and spectral distribution of urban features over the informal settlement areas, as reported in Quintanilha & Silva (2005). Schiewe & Tufte (2005) emphasize the difficulties that arise with increasing internal classes' variability due to increases in the spatial resolution of remote sensing images. With increased spatial resolution images, actual heterogeneities that exist on the surfaces and along edges of objects are no longer "averaged out" due to course resolution pixels. In multi-spectral remote sensing images, increases in the spatial resolution of the image (smaller pixel size or reduced ground space distance) are accompanied by increases in the within-object variability of spectral data for a given feature. Not only are spectral values within the object prone to increased variability, but also along the borders of features of interest significant variability of spectral response occurs due to such complexities as shadow effects, mixed pixels, or stair-step patterns along what should be a distinct border. These increases in variability of the spectral response of a given feature complicate the task of separating desired feature classes. The hierarchic classification methods available in object-based classification technologies allow consideration of feature characteristics that exceed those available through considering spectral characteristics alone. Therefore, an object-based approach was selected to develop the road detection in this work.

Despite some difficulties, the final result demonstrated high agreement between the extracted and the actual roads observed for the study area. Two distinct approaches were used to compute the quantitative analysis: 1) commission and omission errors; and 2) linear comparative analysis. Results produced from both methods are similar in terms of precision of the detected roads (64.5% and 64.48% respectively). Both methods used to quantify classification accuracy produced metrics that closely agreed when considering omission errors for areas where the procedure failed to detect actual roads. Per pixel omission errors were computed to be 35.5 percent by area while linear correspondence produced a 36.8 percent gap statistic both showing a good indication and high agreement for areas where the classification methodology failed to classify complete road systems.

In closing, the expansion of urban roads with no planning toward the limits of urbanized areas illustrates real human problems in sprawl that are ongoing, especially in large-size cities of developing countries. Monitoring this expansion using high resolution images requires automatic approaches due to the amount of data. The methodology reported herein provides significant results in terms of road detection and cost-benefit that may be

achieved in comparison to a manual approach. The results are encouraging and offer a pathway to methods that may be enhanced and further refined in future research. Refinements may possibly be achieved through refinements in custom features and classifiers as well as the use of supplemental data such as LIDAR or IFSAR terrain data (or other remote sensing raster data sets that may be used in object-based analysis) to add additional information enabling improved ability to classify roads and separate them from surrounding features.

Acknowledgements

References

Barros, M. T. L. (2004). Gerenciamento integrado de bacias hidrográficas em áreas urbanas, CNPq-EPUSP, Sao Paulo.

Blaschke, T.; Strobl, J. (2001). What's Wrong With Pixels? Some Recent Developments Interfacing Remote Sensing and GIS. In GIS-Zeitschrift fur Geoinformationsysteme, Heft 6, p 12-17.

Hofmann, P. (2001). Detecting informal settlements from Ikonos image data using methods of object oriented image analysis – an example from Cape Town (South Africa). In Jurgens, C. (editor): Remote Sensing of Urban Areas / Fernerkundung in Urbanen Raumen. (Regensburger Geographische Schriften, Heft 35). Regensburg: 41-42.

Nobrega, R. A. A. (2007) Detecção da malha viária na periferia urbana de São Paulo utilizando imagens de alta resolução espacial e classificação orientada a objetos. Tese de Doutorado, Escola Politécnica da Universidade de São Paulo, São Paulo, 166 p. http://www.teses.usp.br/teses/disponiveis/3/3138/tde-10082007-171032/

Nobrega, R. A. A. ; Quintanilha, J. A. ; O'Hara, C. G. (2006). Detecting roads in informal settlements surrounding Sao Paulo city by using object-based classification. In: OBIA'06 - 1st International Conference on Object-Based Image Analysis, Salzburg - Austria. OBIA'06 - 1st International Conference on Object-Based Image Analysis.
http://www.commission4.isprs.org/obia06/papers.htm

Nobrega, R. A. A.; O'Hara, C. G.; Olson, G.; Kim, S.; Vijayaraj, V.; Quintanilha, J. A.; Barros, M. T. L. (2006) Extracting and classifying bare soil erosion risk areas in an urban basin using object-oriented technologies, high resolution im-

agery and elevation data. Geographic Information Systems and Water Resources IV - AWRA Spring Specialty Conference, Houston, Texas. CD-ROM http://www.gri.msstate.edu/information/pubs/docs/2006/AWRA_2006.pdf

Pinho, C. M. D (2005) Análise orientada a objetos de imagens de satélite da alta resolução espacial aplicada à classificação da cobertura do espaço intra-urbano: o caso de São José dos Campos-SP. Dissertação de Mestrado, Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 179 p.

Quintanilha, J. A.; Silva, O. F. (2005). Identification of urban objects through Ikonos Images. In: Anais XII Simpósio Brasileiro de Sensoriamento Remoto. Goiânia, INPE, p. 4265-4268.

Repaka, S.; Truax, D.; Kolstad, E.; O'Hara, C. G. (2004). Comparing Spectral and Object-Based approaches for Classification and Transportation Feature Extraction from High Resolution Multispectral Imagery. Proceeding for ASPRS 2004 Annual Conference. CD-ROM.

Schiewe, J.; Tufte, L. (2005). O potencial de procedimentos baseados em regiões para avaliação integrada de dados de SIG e sensoriamento remoto. In: Blaschke, T. Kux, H. Sensoriamento Remoto e SIG avançados: novos sistemas sensores e métodos inovadores. Ed. Oficina de Textos, São Paulo, p. 51-60. CD-ROM   http://www.gri.msstate.edu/information/pubsearch.php

Seo, S.; O'Hara, C. G. (2004). Quantifying Linear Feature Extraction Performance. Proceedings of the American Society of Photogrammetry and Remote Sensing, Denver. CD-ROM.

Skackelford, A. K.; Davis, C. H. (2003). A combined fuzzy pixel-based and object-based approach for classification of high resolution multispectral data over urban areas. IEEE Transactions on Geosciences and Remote Sensing, v41, n. 10, p. 2354-2363.

Tso, B.; Mather, P. M. (2001). Classification Methods for Remotely Sensed Data. Taylor and Francis, London.

Wiedemann, C. (2003). External evaluation of road network. In: ISPRS Workshop Photogrammetry and Remote Sensing 2003, Munich. Proceedings. v. 34, part 3/W8, p. 93-98.

Section 7

Development of new methodologies

Chapter 7.1

Object-oriented analysis of image and LiDAR data and its potential for a dasymetric mapping application

F. Kressler, K. Steinnocher

Austrian Research Centers GmbH – ARC, Donau-City-Strasse 1, 1220 Wien, Austria

KEYWORDS: urban land use, building height, census data, population disaggregation

ABSTRACT: In this chapter, object-oriented image analysis, on the basis of eCognition, is used to classify an IKONOS-2 image together with data derived from an aerial laser scanner. After the initial classification of different land cover types, a subset containing only buildings was integrated with zoning as well as population data. Zoning data was used to assign the classified buildings to different land use types, thus allowing the identification of buildings designated for residential use. On these were mapped population data, available on a 250 m grid basis while taking into account the height of each building, allowing a differentiated presentation of where people live within in each grid cell. Comparison with reference data suggest a high potential not only for further dasymetric mapping applications, but other applications as well.

1 Introduction

The classification of optical images has always been hampered by the missing third dimension. With the increased availability of LiDAR (<u>Li</u>ght <u>De</u>tection <u>A</u>nd <u>R</u>anging) data classification can be significantly improved both in terms of accuracy as well as automation. Object-oriented classification is an obvious choice as a tool for an analysis with its ability of inte-

grating and processing data with very different properties. In the first part of this paper an IKONOS-2 image will be analysed together with a normalised Digital Surface Model (nDSM) and a difference model between first- and last pulses (ΔH_{FL}) derived from airborne laser scanner data. The second part looks at an application that can now be performed due to the superior quality of the classification and the available height information. This application examines an improved dasymetric mapping method for disaggregating population data to houses based on building volume. Dasymetric mapping may be defined as a kind of areal interpolation that uses ancillary data to support an areal interpolation process (Mennis, 2003). Zoning information is used to identify those buildings belonging to residential areas, making a move from land cover to land use. This is followed by an outlook examining how the integration of different data bases can support the creation of further improved land use maps.

LiDAR data in combination with optical remote sensing images have already been used for classification in urban/suburban area in a number of studies. Zeng, et al. (2002) use a maximum likelihood classifier (MLC) after creating three different datasets for high, low and shadow areas. Haala and Walter (1999) use a maximum likelihood classification to separate buildings, trees, roads and grassland from a colour aerial image with a nDSM as an additional band. The classification result could be considerably improved by this method. Syed et al. (2005) compare MLC to object-oriented classification. MLC resulted in a "salt and pepper" and limits in successfully classifying shadows, a problem not encountered with object-oriented classification. Hofmann (2001) examined the detection of buildings and roads from an IKONOS-2 image and height data, using object oriented classification, confirming its suitability for the data used.

Linking remote sensing with demographic data from the census has been subject of a number of studies. Harvey (2002) estimated populations from Landsat TM imagery in Australia using a recursive algorithm to recalculate population from the census zone according to a pixel's spectral vector. The correlation between census dwelling data and residential densities derived from remote sensing were analysed in Chen (2002) combining classified image data and population from census tracts. Spatial disaggregation of population based on housing densities on a regional scale was discussed in Steinnocher et al. (2006). The relationship between population density and vegetation cover in urban areas was investigated in Pozzi & Small (2005) in order to improve differentiation of urban land use classes. Liu et al (2006) compared population density with texture measures from IKONOS images. All these analyses have in common that they link population with (1) spectral or textural image features, (2) index images such as NDVI or (3) thematic classes such as residential areas or housing densities.

In the present paper mapping will be taken one step further, not only going on a building or building block level but by including height information as well.

2 Data and Study Area

The study area lies in the province of Upper Austria, in the north of the City of Linz, covering approximately 3.7 x 3.5 km (see Fig. 1). It features mixed land cover/land use with large parts being covered by fields and forest. Residential areas comprise both single houses as well as large building blocks. In addition some commercial districts and a motorway are situated in the south and a university complex in the east.
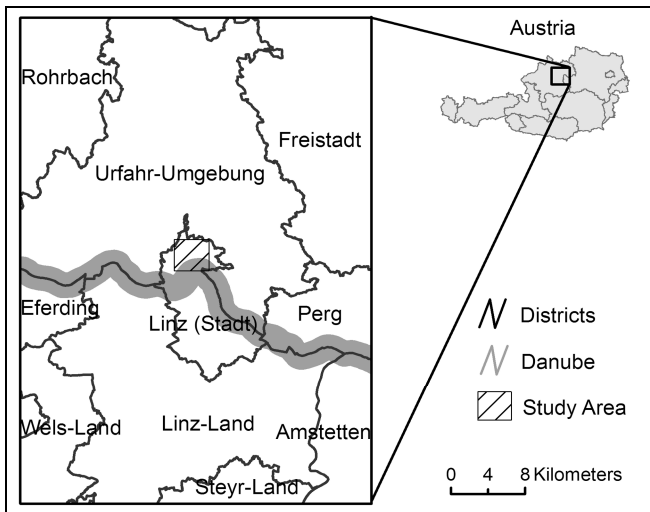
Fig. 1. Study area

Two remote sensing data sets were used for the study (see Fig. 2). One is subset of an IKONOS-2 scene, recorded on 15th June, 2002. It is a fusion product of the four multispectral bands (4 m spatial resolution) and the panchromatic band (1 m spatial resolution), resulting in four multispectral band with 1 m spatial resolution. The multispectral bands cover blue, green, red and part of the near-infrared of the electromagnetic spectrum.

The second data set is from an airborne laser scanner (ALS), acquired on 24th March, 2003, at a flying height above ground of 1,000 m and an average point density of 1 point/m². From this two different models were

generated. One is a normalised digital surface model (nDSM), containing height information above ground (see Fig. 3, right - lighter tones representing greater heights). Buildings and forests can be clearly seen in this image while flat areas such as water bodies, roads, meadows, and so forth cannot be separated. The other is a difference model between first and last pulses (ΔHFL), which has a height of zero for sealed areas and a height similar to that of the nDSM in forested areas, thus aiding the identification of trees.
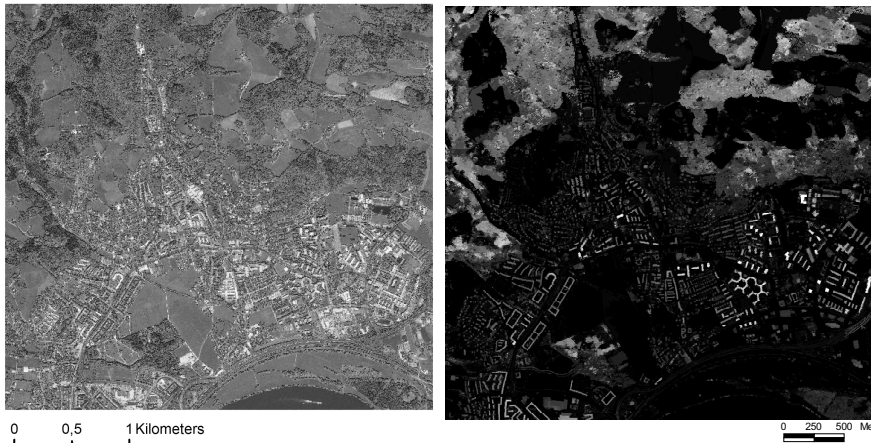
Fig. 2. False-colour-infrared IKONOS-2 image (left) and normalised digital surface model (right) of study area

For the integration of ancillary data, two data sets were available (see Fig. 3). One is the generalized zoning plan, as defined by the upper Austrian Land Use Planning law. The generalized zoning plan covers 5 classes (residential areas, mixed use, industrial areas, commercial areas, green areas), aggregated from its original 15 classes.

The other data set contains population data collected by the Statistik Austria[1], the leading provider of statistical information in Austria. The data are from the census 2001, aggregated to a raster with a cell size of 250 x 250 m. Its advantage to the original representation on the level of census tracts is that the cell size is fixed and does not change over time. In addition these cells are generally smaller than census tracts, especially outside core urban areas. In the future an improved version will be available with a cell size of 125 x 125 m.

---

[1] http://www.statistik.at/

It can be seen that some of the squares show population outside of those areas designated as buildings zones, especially in the Northeast. This is mainly due to individual farms that are not part of the zoning plan.
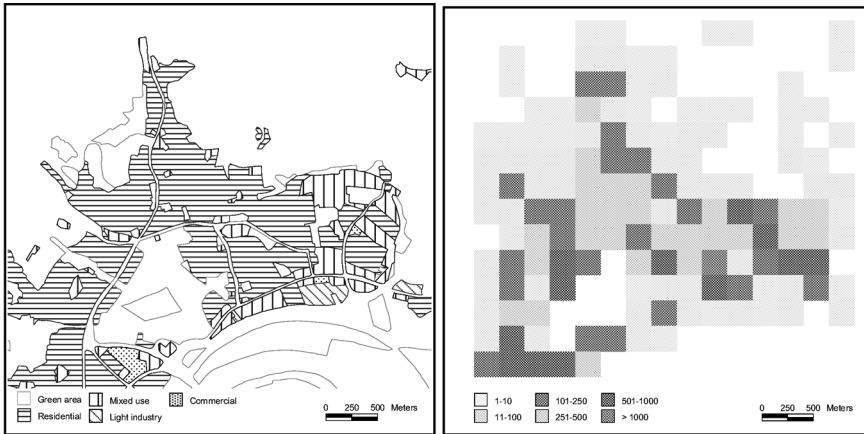
Fig. 3. Aggregated zoning data (left) and population data on 250 m grid (right) of study area

3 Methodology

For the classification of the satellite image and the ALS data an object-oriented approach as implemented in eCognition 4.0[2] was used. For the analysis two processing steps can be distinguished. One is the segmentation of the data into homogenous segments; the other is the assignment of these segments to discrete classes. These steps can also be used alternately, i.e. the classification results of one processing step can be input for a subsequent segmentation. In the following sections, image segmentation and classification will be examined followed by the integration of ancillary data.

3.1 Image Segmentation

Aim of this processing step is to identify homogenous image segments. Depending on the data and goal of the analysis, different degrees of homogeneity can be desirable. Segmentation is controlled by the parameters

---

[2] www.definiens-imaging.com

scale, colour and shape, selected by the user (Baatz and Schäpe, 2000). Scale controls the maximum allowed heterogeneity per segment. Leaving all other parameters constant a larger scale will result in larger segments. The parameters colour and shape sum to one and determine how much colour and shape information is used for the segmentation process. The shape parameter is further divided into compactness and smoothness. A high value for compactness leads to smaller and very compact segments, more suitable for manmade objects, while a high value for smoothness leads to segments optimised to have smooth borders, more suitable for natural objects.

Different data sets such as laser and satellite images can be segmented simultaneously and their influence to the segmentation process can be adjusted by weights. Nevertheless, a stepwise approach was chosen here due to the very different information content as well as different scaling of the data, allowing more control over the segmentation process. An initial segmentation was carried out on the basis of the satellite image as its information content is larger than that of the laser image. Discrimination of different land cover types, especially of vegetation and built-up areas can be done very easily using these data. A second segmentation was performed one level below the initial segmentation and here only the laser information was used. This allows a separation of both built-up areas and vegetation on the basis of height, improving both the identification of roads as well as that of shrubs and grassland. In a third segmentation level between level 1 and 2 a segmentation was performed based only the absolute height difference while at the same time staying within the boundaries defined by the original segmentation of the satellite image.

3.2 Image Classification

For each segment, identified in the previous processing steps, a large number of features is available for the classification. In order to combine the relevant information and establish the necessary classification rules, a class hierarchy has been set up. At this stage eight classes were identified (water, shadow, meadow, shrub, tree, soil, road and buildings). In addition a class for unclassified segments was introduced. Four types of features were used for the classification: mean, ratio, brightness and normalised difference vegetation index (NDVI). Mean refers to the mean value of all pixels present in a segment, e.g. mean nDSM is the mean height of all pixels within a segment. Ratio refers to the mean spectral values of one band divided by the sum of all spectral values, e.g. ratio blue is the ratio of the IKONOS blue band to the sum of all other IKONOS bands. Brightness is

the mean of all spectral bands and the NDVI is an index calculated from the red and the near-infrared bands, giving an indication as to the vegetation intensity (Lillesand and Kiefer, 1994).

Each class was described by one or more of these features. Thresholds for each feature are given in the form of fuzzy functions. Classification is performed in a hierarchical manner, examining each segment whether its membership is high enough for a classification or not. If it is, then the segment is assigned to the appropriate class, all remaining segments are examined for the next class until all segments are classified. Table 1 shows an overview of the features used for each class. A combination of spectral as well as height information was used to successfully assign the different classes.

Table 1. Features used for classification

Class	Features used
Water	Mean nDSM
	Ratio infrared
Not water	Not water
Shadow	Brightness
Not shadow	Not shadow
Vegetation	NDVI
	Brightness
	Ratio blue
Meadow	Mean nDSM
Shrub	Not shrub
	Mean nDSM
Tree	Not meadow
	Not shrub
	Mean ΔH_{FL}
Not vegetation	Not vegetation
Sealed-flat	Mean nDSM
Soil	NDVI
Veg_flat	NDVI
Road	Not soil
	Not road_veg
Not Sealed-flat	Not sealed-flat
Building	Ratio blue
Unclassified	Not building

3.3 Integration with Ancillary Data

Of the classification result only the identified building objects are used for the integration with ancillary data. In a first step the buildings are intersected with the zoning data, making a first move from land cover to land use, assigning each building to one of the five zoning classes. These were then edited to further separate buildings with public use and those that belong to allotments (typically in a green zone).

With this added information the next step, the integration of the population data, can be greatly improved. Population data from statistical sources, whether on a raster or on a census tract basis, does not provide information on a building block level. This can be approximated by disaggregating the population data only on those buildings designated as residential or mixed use based on the area covered by them. An even better approximation can be derived by considering building height as well. In order to perform this dasymetric mapping approach, the volume of each building is calculated. Based on this volume the population of each grid cell is mapped onto the buildings belonging to that cell. If a building occupies more than one grid cell then the population is averaged, arriving at one population value for the whole building. This is necessary as the whole population of a building is always assigned to only one grid cell in the census data.

4 Results

4.1 Segmentation and Classification

Before a classification can be performed the data must be segmented. The initial segmentation was carried out using only the IKONOS-2 scene (scale: 20, color/shape: 0.5, smoothness/compactness 0.5). A second layer was created based on a spectral difference of 40. The result is the separation of spectrally distinct land cover types. Before the next segmentation can be carried out, the first layer must be deleted. Below the remaining layer a new segmentation is performed only on the basis of nDSM data (scale 3, colour 1). Here no shape information is used for the segmentation. The borders of the segments, derived in the segmentation of the satellite image limit how much a segment may grow. From this follows that spectrally similar image objects can be separated on the basis of height (e.g. roof and road) but objects of similar height will not be combined if spectrally different (e.g. a tree next to a house). The final segmentation level is created between level one and two, again only based on nDSM data and

using a spectral difference value of 1, again within the boundaries defined by the segmentation of the satellite image.

Fig. 4. Final segmentation

The result of the segmentation (see Fig. 4 for a subset) is basically a product of two separate segmentation procedures, where the first limits the extent as how far a segment may grow in the later segmentation. This assumes that basic land cover types can be separated successfully in the first segmentation stage and only need refining, based on the LiDAR data within each land cover type. Examples are sealed areas, which in most cases can be separated very well from non-sealed areas. A differentiation of different types of sealing is much more difficult and often impossible using only optical data. Here LiDAR gives us the opportunity to identify land cover types such as roads and building with a high degree of accuracy.

Classification is performed on the final segmentation level using the features defined in Table 1. Seven land cover classes were differentiated plus shadow and one for unclassified segments (see Fig. 5).
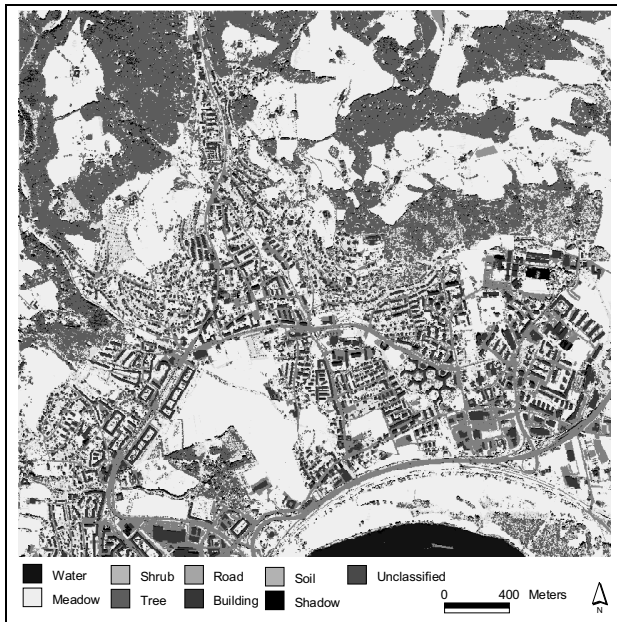
Fig. 5. Classification result

Vegetation and sealed areas were differentiated relying only on spectral information derived from IKONOS-2. Vegetation was further separated using nDSM information for meadow (light green) and shrub (green). Trees were identified both on the basis of height as well as first-last pulse information (dark green). Sealed areas were divided into roads (grey) and buildings (red). Here the nDSM information greatly improves their differentiation, the quality being mainly dependent on the resolution of the data but not on any ambiguity within the data. The only time where a manual editing was necessary was in the case of road bridges, which were classified as buildings during the initial classification (sealed areas with high nDSM). With the use of a road network this problem could certainly be overcome to a large extent, if the degree of automation were to be increased.

4.2 Integration

For the data integration, only the identified buildings or building blocks were used. The first involves the integration with zoning data. This allows the separation of buildings based on the use allowed in a particular zone.

This is an important step to facilitate the following integration of the population data, as then the data can be projected only on those buildings which, according to the zoning plans, are actually residential areas. The buildings were assigned to one of five classes. Light industry (magenta), residential (red) and mixed use (orange) were taken directly from the zoning data. Public use (cyan) such as schools and university buildings was manually assigned based on local knowledge, although in the future, automation for this step can be foreseen as well. Allotments were defined as houses built in green zones (see Fig. 6).

For each building or building block a mean height is available, allowing the calculation of volume. This assumes that all buildings have a flat roof, which is accepted for reasons of simplification. The population of each grid cell can then be disaggregated to the single buildings, weighted either by area or the building's volume (see Fig. 7). Buildings assigned light industry or public use are excluded from the disaggregation, as they are not likely to be used for residential purposes.

Fig. 6. Buildings classified according to zoning plan
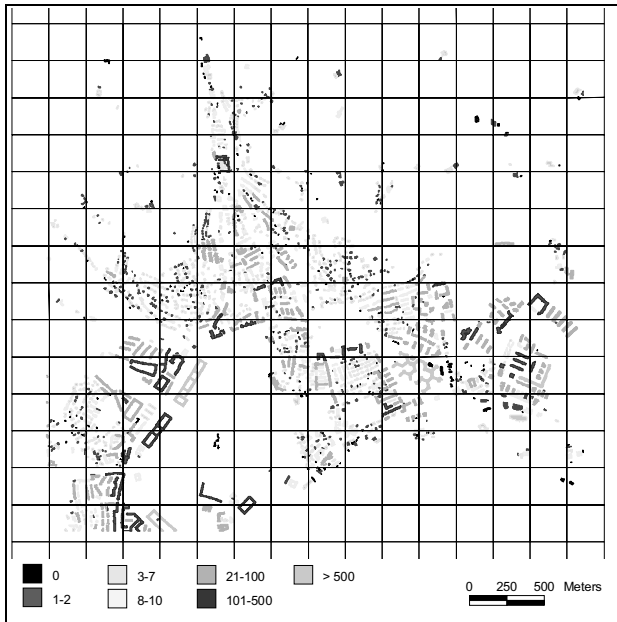
Fig. 7. Population data mapped onto residential and mixed use buildings

Reference data, i.e. population per building, is not available publicly due to data privacy. However, the results of the disaggregation were evaluated by Statistik Austria internally and evaluation results provided anonymously in terms of error measures. For a sample of 1712 buildings the rms error equalled 18,4 for the areal approach and to 13,7 for the volume based approach. Average deviations came up to 5,8 and 3,7 respectively. It is obvious that the volume based approach improves the results significantly compared to the areal approach.

The remaining error might be due to time difference of the data sets (census from 2001, IKONOS-2 from 2002 and from LiDAR in 2003) and to the limitation of the zoning plan in terms of single building use. Also the assignment of the population of a building to one grid cell of the population data, irrespective of how many cells the building spans, is critical. This could be improved by using addresses points as spatial reference for buildings as they represent a clear link to corresponding grid cells.

5 Conclusion and Outlook

In this paper object-oriented classification of LiDAR (nDSM and ΔH_{FL}) and multispectral IKONOS-2 data (pansharpened to 1 m spatial resolution) was examined. The use of both data sets allowed an easy separation of the basic land cover types present in the study area. LiDAR data are especially beneficial for the separation of flat sealed areas from buildings, while optical data allow a good separation of vegetation and sealed areas. Manual editing of the land cover classification could be limited to correcting road bridges, which were classified as buildings.

The next step was the integration of the classification results with statistical data and the information provided by LiDAR. Aim was the mapping of population data, available on a 250 m grid basis, not only on the basis of areas covered by the buildings but also on the height of these buildings. As not all buildings are residential they building mask was first intersected with a zoning map of the area. This allowed the identification of those buildings which are not used for residential purposes. Using local knowledge this was further improved by manually identifying those buildings which lie in mixed use areas but are not used as residential buildings (e.g. churches, university, etc.). On the remaining buildings the population data was mapped, allowing a more detailed view of how many people live in each building. Comparison with reference data proved that with this method population disaggregation leads to higher accuracies compared to when only the area of a building is considered for dasymetric mapping.

With the integration of the zoning data a fist step had been made from land cover to land use. The next will be the linking of the identified building blocks to address data, thus allowing a mapping of any kind of data, indexed by address. This way the identification of different uses of a building can be further improved. Depending on the accuracy of the address data and the various data bases used, very detailed land use maps could be created.

6 References

Baatz, M. and Schäpe A., 2000. Multiresolution segmentation – an optimization approach for high quality multi-scale image segmentation, In: Strobl, Blaschke & Greisebener (Edts): *Angewandte Geographische Informationsverarbeitung XI. Beiträge zum AGIT-Symposium Salzburg* 2000, Karlsruhe. Herbert Wichmann Verlag, 2000, pp. 12-23.

Chen K., 2002. An approach to linking remotely sensed data and areal census data. Int. J. Remote Sensing, 23(1), pp. 37-48.

Halla, N. and Walter, V., 1999. Automatic classification of urban environments for database revision using LiDAR and color aerial imagery. *International Archives of Photogrammetry and Remote Sensing*, 32(Part 7-4-3 W6), pp. 641-648.

Harvey J.T., 2002, Estimating census district populations form satellite imagery: some approaches and limitations. Int. J. Remote Sensing, 23(10), pp. 2071-2095.

Hofmann, P., 2001. Detecting buildings and roads from IKONOS data using additional elevatio information. *GeoBIT/GIS*, 6, pp. 28-33.

Lillesand T.M., and Kiefer R.W. (1994), Remote Sensing and Image Interpretation, John Wiley and Sons, Inc.

Liu, X.H., Clarke, K. and Herold, .M., 2006. Population Density and Image Texture: A Comparison Study. *Photogrammetric Engeneering and Remote Sensing*, 72(2), pp. 187-196

Mennis, J., 2003. Generating surface models of population using dasymetric mapping. *The Professional Geographer*, 55(1): 31-42.

Pozzi F. and Small C., 2005. Analysis of Urban Land Cover and Population Density in the United States. *Photogrammetric Engeneering and Remote Sensing*, 71(6), pp. 719-726.

Steinnocher K., Weichselbaum J. and Köstl M., 2006. Linking remote sensing and demographic analysis in urbanised areas. In (P. Hostert, A. Damm, S. Schiefer Eds.): First Workshop of the EARSeL SIG on Urban Remote Sensing "Challenges and Solutions", March 2-3, 2006, Berlin, CD-ROM.

Syed, S., Dare, P. and Jones, S., 2005. Automatic classification of land cover features with high resolution imagery and LiDAR data: an objectoriented approach, Proceedings of SSC2005 Spatial Intelligence, Innovation and Praxis: The national biennial Conference of the Spatial Sciences Institute, September, 2005, Melbourne, Australia.

Zeng, Y., Zhang, J., Wang, G. and Lin, Z., 2002. Urban land-use classification using integrated airborn laser scanning data and high resolution multispectral imagery, Pecora 15/Land Satellite Information IV/ISPRS Commssion I/FIEOS 2002.

Acknowledgements

Chapter 7.2

Characterising mountain forest structure using landscape metrics on LiDAR-based canopy surface models

B. Maier[1,]*, D. Tiede[2], L. Dorren[3]

[1] Stand Montafon Forstfonds, Montafonerstr. 21, 6780 Schruns, Austria, bernhard.maier@stand-montafon.at

[2] Salzburg University, Centre for Geoinformatics, Hellbrunnerstr. 34, 5020 Salzburg, Austria, dirk.tiede@sbg.ac.at

[3] Cemagref Grenoble, 2 rue de la Papeterie, BP76, 38402 St. Martin d`Heres cedex, France, luuk.dorren@cemagref.fr

ABSTRACT: Forest structure is a key element to determine the capacity of mountain forests to protect people and their assets against natural hazards. LiDAR (Light Detection And Ranging) offers new ways for describing forest structure in 3D. However, mountain forest structures are complex and creative methods are therefore needed to extract reliable structural information from LiDAR. The objective of this study was to investigate if the application of landscape metrics to a normalised canopy model (nCM) allows an automatic characterisation of forest structure. We used a generic automated approach that created height class patches based on a segmented nCM. Two multi-resolution segmentations were carried

---

* Corresponding author

out: level 1 objects represented tree crowns and collectives of tree crowns, level 2 objects represented forest stands. Level 1 objects were classified into four height classes and subsequently overlaid with level 2 stands in order to calculate the metrics 1) canopy density and vertical layering of the forest, 2) forest gap distribution and 3) canopy roughness using the Division Index (DIVI). Canopy density values of each height class allowed the classification of the vertical layering. Distinguishing between single- and multi-layered stands, 82% of all the sample plots were correctly classified. The DIVI calculated on gaps proved to be sufficient to describe the spatial arrangement of patches and distinguish between stands with many small gaps from stands with only a few but larger gaps. Canopy roughness could not satisfactorily be described using the DIVI based on the validation plots. With the approach presented, resource and natural hazard managers can assess the structure of forest stands and as such more easily take into account the protective effect of forests.

1 Introduction

"Mountain forests provide the most effective, the least expensive and the most aesthetic protection against natural hazards" recalls the first paragraph of the Mountain Forest Protocol of the Alpine Convention. Without mountain forests, the costs of building and maintaining technical protective constructions against rapid mass movements in the Alps would be unaffordable. Forest structure is a key element that determines the protective capacity of mountain forests (Dorren et al. 2004). It can be characterised by the horizontal distribution of trees, the vertical layering and the tree species mixture.

Structures of mountain forests differ greatly from those in the lowlands. Mountain forests contain relatively few species, tend to be quite open and consist of a mosaic of tree clusters and gaps (Schönenberger 2001). In mountain forests, particularly, structure is closely related to stand stability, i.e. resistance against and resilience after disturbances such as storm and snow loads (Brang 2001). Other characteristics that determine structure are crown closure and tree density. These directly influence release probability of forest avalanches and the capacity to stop falling rocks (Dorren et al. 2005). Consequently, assessing forest structure enables forest managers and natural risk engineers to evaluate if a forest can fulfil its protective function or not. Reliable and area-extensive data on forest structure is thus a prerequisite for effective resource and risk management in mountainous regions.

Traditional methods for assessing forest structure comprise field inventories (Herold and Ulmer 2001) and aerial photo interpretation (Bebi 1999). The drawback of inventories is that they cannot provide spatially continuous information over a large area. The usefulness of photo interpretation im mountainous terrain is hampered by different illumination and shading effects. Small footprint airborne LiDAR, however, allows deriving detailed digital terrain (DTM) and surface (DSM) models. Subtracting these two models of a forested area results in a so-called normalised canopy model (nCM), which is spatially continuous and not hampered by shading effects - typical for optical remote sensing data. This facilitates assessing forest structure in 2.5D. Various studies show that it is possible to derive a variety of single structural attributes such as tree height, basal area, crown size and above-ground biomass from LiDAR data (Hyyppä et al. 2006, Hall et al. 2005, Naesset 2004, Maltamo et al. 2004, Tiede et al. 2004, Lim et al. 2003). Some studies focus on tree height variance as an indicator of vertical forest structure (Blaschke et al. 2004, Zimble et al. 2003). Until now, little attention, however, has been paid to the application of landscape metrics to LiDAR-based canopy surface models.

Landscape metrics, representing a view "from above", allow quantifying the link between landscape structure and function. Usually, they are derived for the following three levels: 1) the patch level which corresponds to individual objects, 2) the class level, i.e., characteristics of all patches of the same type and 3) landscape level which integrates all patch types or classes across the extent of the data. For some applications a fourth level, the region level, is introduced. This level indicates a sub-area of the landscape. If we translate the landscape metrics nomenclature into the forest structure context of this study, landscape refers to the whole forest, region refers to stand, class to tree height class and patch to tree height patch. The use of landscape metrics within forests mainly focuses on forest fragmentation and biodiversity, and their changes over time (Traub & Klein 1996, Venema et al. 2005).

In this study, landscape metrics are used for another purpose, namely to describe fragmentation of the canopy surface height in terms of variability or structuring. Landscape metrics are expected to link forest structure with the protective capacity of a mountain forest. The general idea is that the use of landscape metrics leads to more objective, transparent and repeatable results compared to visual interpretation by a human interpreter. Our overall goal is to derive stand structure information as a basis for protection forest planning, management and monitoring. Therefore, the objective of this study was to investigate if the application of landscape metrics to a nCM allows an automatic characterisation of forest structure.

2 Study area and data

2.1 Study area

The study area covers 120 ha of spruce-dominated protection forest on a west facing slope near the Austrian village of Gaschurn in the Montafon valley (see Fig. 1). This forest provides essential protection against natural hazards such as landslides, avalanches and rockfall. The study area ranges from 1000 m altitude in the valley floor up to 1800 m at the tree line. The study area is dominated by steep rugged terrain with rock faces, gullies and torrents. Old-growth forests with different structure types dominate. The structural differences are due to various factors. Multi-layered structures with short-distance height variation in the canopy, so called fine-grained canopy patterns, grow on shallow soils that cover glacially eroded bedrock.
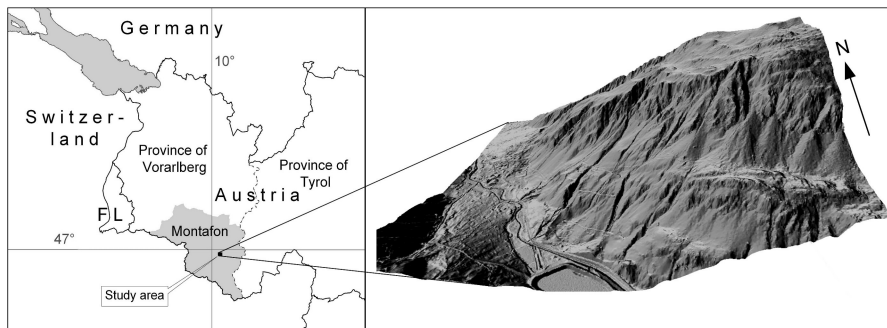
Fig. 1. Location (left) and DTM of the study area viewed from Southwest (right)

Homogenous pole stands occur in areas that used to be meadows in the lower parts of the slope. Open structures caused by windthrow are found below the rock faces in the southeastern part of the study area. Another type of open structures is found close to the tree line. These, however, consist of tree clusters (collectives), which are typical for high altitude forests. Along the avalanche tracks, tree regeneration and young stands can be found. The presence of all these different structure types makes this area well-suited for this study.

2.2 Data

The ALS data used in this study were acquired on the 10th of December 2002 under leaf-off canopy conditions. The instrument applied was a first/last pulse Airborne Laser Terrain Mapper (ALTM 1225) produced by Optech Inc. (Canada). The pulse repetition frequency of the ALTM is 25 kHz that resulted in a point density of 0.9 points m^{-2} at an average flight height of 1000 m above ground level. With a laser beam divergence of 0.3 mrad, the theoretical footprint on the ground was about 0.30 m. The average ground swath width was about 725 m, the maximum scanning angle 20º (Wagner et al. 2004).

The data obtained by the ALTM have been processed and interpolated by the TU Vienna using the *hierarchical robust filtering* approach within the SCOP++ software package (Kraus and Pfeifer 1998). As a result a DTM and a DSM, both with a resolution of 1 m × 1 m, were created. By subtracting the DTM from the DSM we obtained a nCM which describes an estimate of the forest height.

Thirty-three circular inventory plots of 314 m² served for validation. Within these plots, we recorded for each tree: diameter at breast height (1.3 m), position (azimuth and distance), height and species. In addition, we determined the vertical layering and the stand type following Maier (2007a).

3 Methodology and Implementation

3.1 Segmentation and height classification

Our method combines object-based multi-resolution segmentation and classification with GIS analyses. The rule sets for segmentation and classification are developed using CNL (Cognition Network Language), a modular programming environment available in the Definiens Developer environment. Segmentation is done using a region-based, local mutual best fitting segmentation approach. It is a pair-wise clustering process, starting with randomly selected one-cell objects. In an iterative process these objects are merged into larger ones. This segmentation process is determined by three parameters which can be set by the user. The first one is *scale* which is a stop criterion for the object growing process and thus defines the maximum segment, or object size. The other two parameters *shape (vs. colour)* and *compactness (vs. smoothness)* define the weights for the allowed heterogeneity of image values (here: height values) and shape complexity (Baatz and Schäpe 2002, Benz et al. 2004).

Segmentation was used to create two object levels which corresponded to 1) single and clustered tree crowns and to 2) forest stands. To delineate single and clustered (collectives) tree crowns the nCM was segmented into objects using a scale parameter of 5. The ancillary parameters *shape* and *compactness* were both set to 0.5. Since single tree crowns comprise different height classes, a small *scale* parameter led to onion-like concentric objects only representing parts of a tree crown. A big *scale* parameter, however, levelled out the different tree heights and resulted in a loss of structural complexity. Segmentation tests with known trees showed that a scale parameter of 5 was a good compromise. The resulting objects in this level represent homogenous tree height patches. These canopy objects were then classified according to their mean height using the height classification schema presented in Fig. 2.
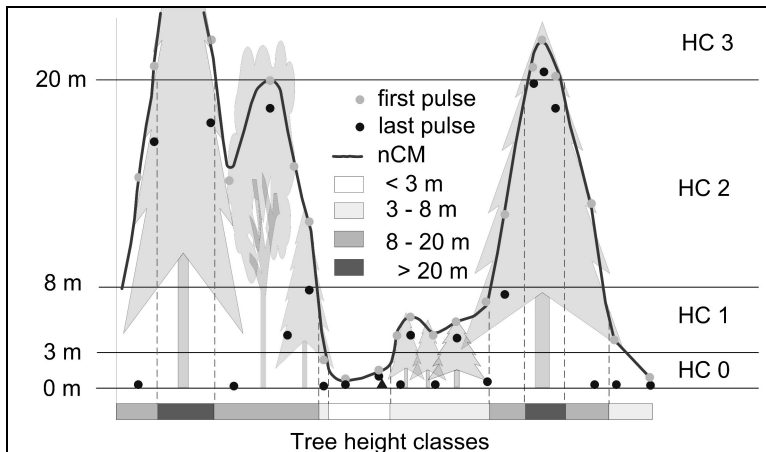
Fig. 2. Tree height classification schema

The height classification schema follows the different forest development stages that are defined in the Manual for the Aerial Photo Interpretation of the Swiss Forest Inventory (Ginzler et al. 2005). Four height classes (HC) are introduced: HC 0 comprises all segments with a mean vegetation height below 3 m. Those are regarded as unstocked as the differentiation accuracy of the laser allows no distinction between surface roughness (lying dead wood, rocks, stumps or low vegetation) and young trees. HC 1 (3-8 m) corresponds to young trees and HC 2 covers the range from 8-20 m. This class contains mainly pole forests and timber forest. HC 3 consists of tree crowns higher than 20 m, which are usually thicker trees, or old growth forest. The classified objects were subsequently dissolved into ho-

mogenous height class patches and exported into a GIS readable vector format (Polygon Shape file).
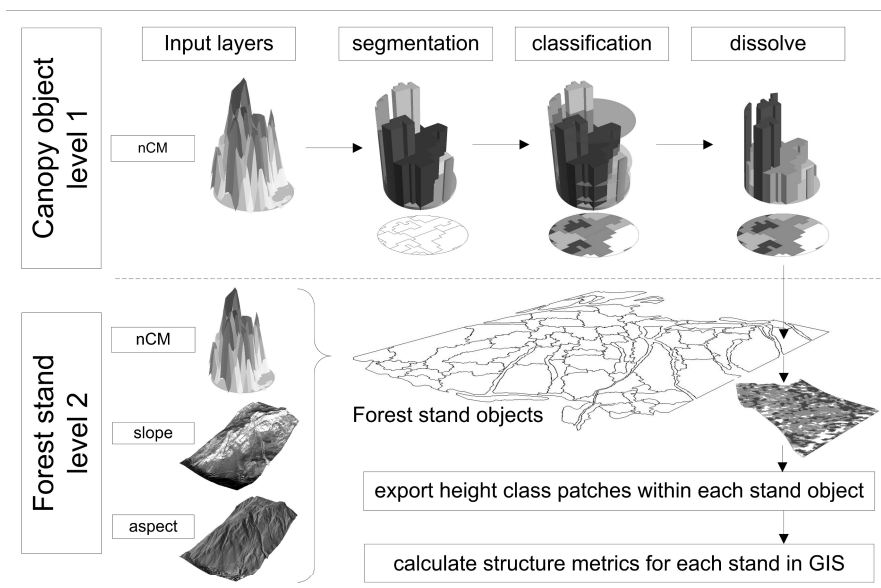
Fig. 3. Flow diagram of study methodology

To delineate forest stands, input rasters were segmented into objects using a *scale* parameter of 100, as well as a *shape* and *compactness* parameter of 0.3 and 0.4 respectively. Because terrain features strongly influence forest growth and development in relief-rich mountainous terrain, the input rasters consisted of the nCM, but also the mean slope gradient, and the aspect. We assumed that forest stands are largely homogenous in terms of age, species distribution and developmental stage and that they reflect similar physiographical conditions. The segmentation weights of the three input rasters nCM, slope and aspect were set to 1, 0.5 and 0.5 respectively. The setting of the above mentioned parameters provided the forest stand map that respected best existing stand borders observed in the field and on CIR orthophotos.

Following the segmentation and classification, a GIS was used to characterise the patch-structure (level 1) within each stand object (level 2) using landscape metrics for canopy density, vertical layering, gap distribution and canopy roughness of a forest stand. All metrics were calculated with the V-late extension for ArcGIS (Lang and Tiede 2003). Fig. 3 summarises the complete method flow of this study.

3.2 Landscape metrics calculation

The structural characteristics considered relevant for this study were: canopy density and vertical layering of the forest, forest gap distribution and canopy roughness. In the following paragraphs their calculation using landscape metrics will be explained in detail.

3.2.1 Canopy density and vertical layering of the forest

Canopy density (CD) is defined as the percentage of the area which is occupied by the horizontal projection of tree crowns. Multiple coverage in super-imposed tree layers (height classes) is not taken into account (Keller 2005). The landscape metric we used for describing CD is called CD_{total} and corresponds to the total area of all stocked height patches of HC 1, 2 and 3, in percent of the total stand area. Maier (2007b) shows that there is a good correspondence ($R^2 = 0.7$) between canopy density measured in the field by mapping crown radius and the canopy density derived from LiDAR-based canopy models.

$$CD = P_i = \frac{\sum_{j=1}^{n} a_{ij}}{A} (100) \qquad \text{(eq. 1)}$$

P_i = proportion [%] of the stand (without gaps) occupied by height class i

a_{ij} = area [m²] of patch ij
A = stand area [m²] without gaps (HC 0)

Vertical layering or stratification can be referred to as the vertical distribution of foliage (Parker and Brown 2000). When calculating CD for the three different HC (eq. 1, see McGarigal and Marks 1995), the vertical layering of the forest can be assessed. It is important to note that super-imposed tree layers cannot be detected using a 2.5D nCM. Therefore, the layering assessment is restricted to the different height levels which can be seen from above. According to the Austrian Forest Inventory (AFI), a single-layered stand has only one distinct canopy layer. Two-layered structures have in addition to the dominant canopy layer a second layer, which exhibits considerable growth potential. Here, a separate forest layer needs to have a total CD of at least 30% (Schieler and Hauk 2001). According to the Swiss National Forest Inventory, single layered stands consist of one layer that covers more than 80%. Our rules for assessing the vertical layer-

ing were principally based on these guidelines, but adapted for an nCM-based analysis (Table 1).

Table 1. Rule set for vertical layering according to canopy density (CD) per height class (HC)

single-layered	one HC with CD > 60%, no other HC > 35%
two-layered	two HC with CD > 30% or one HC > 50% and one HC >20%
multi-layered	three HC with CD > 20%

For validation, we compared the vertical layering classification with the vertical structures recorded in the 33 field plots.

3.2.2 Forest gap distribution

Gaps are unstocked patches in a forest matrix. They can be formed by harvesting trees (silvicultural intervention), by natural succession or disturbances. When trees die, they create gaps that allow the sunlight reaching the ground, which could initiate forest regeneration. Windthrow, snow break and subsequent bark beetle infestations belong to the natural gap-making processes and disturbances. Small gaps are distinct and common features of mountain forests, and do not pose problems to the protective function of the forest against natural hazards. Large openings in the forest cover, however, do. They are potential avalanche release areas or can accelerate rockfall.

Therefore, the size and spatial distribution of gaps influence the protective capacity of a forest. The bigger and longer a gap is in the slope direction, the more likely forest avalanche release or falling rocks acceleration is. Gap length can be calculated by defining the longest flow path in slope direction. These values can then be compared with existing protection forest management guidelines (e.g. Frehner et al. 2005) in order to assess the need for silvicultural intervention or tending.

$$DIVI_{HC0} = 1 - \sum_{i=1}^{n} \left(\frac{a_i}{A} \right)^2 \qquad (eq.\ 2)$$

where a_i = area [m²] of gap patch i
              A = here: total area [m²] of gaps in one stand (HC 0)

In order to assess the spatial distribution of gaps, we calculated the Division Index (DIVI, see eq. 2). The DIVI is defined as the probability that two randomly selected locations do not occur within the same patch in the forest (Jaeger 2000). Although this index originates from ecology where it

refers to the likeliness that two organisms will meet within the same patch, it can also refer to the degree of gap fragmentation, or "gappiness" of a forest stand. In other words, it expresses if gaps occur as single big unstocked areas or highly fragmented small gaps throughout the stand.

3.2.3 Canopy roughness

Finally, we tried to quantify the graininess or roughness of the forest canopy by measuring the overall fragmentation of all height classes and interpret it as a degree of structuring. This is done by calculating the Division Index on all height classes on a stand level (eq. 3). This measure of fragmentation describes if the canopy is vertically homogenous or heterogeneously structured, following McGarigal and Marks (1995), who used this index to differentiate between coarse- and fine-grained landscape structures. Within the conventional repertoire of forest inventory parameters, the parameter *StandType* (Schieler and Hauk 2001) which defines uniform, irregular and single tree structures (see Table 2) appeared the only one to represent canopy roughness (CR) as a structural parameter.

$$DIVI = 1 - \sum_{i=1}^{n} \left(\frac{a_i}{A} \right)^2 \qquad \text{(eq. 3)}$$

where
a_i = area [m²] of patch i
A = here: total area [m²] of stand

Table 2. Definition of stand type according to the Austrian Forest Inventory (Schieler and Hauk 2001)

Uniform	Homogenous, more or less closed canopy with uniformly distributed trees or tree groups
Irregluar	Stands consisting of tree groups of similar height with a distinct green mantle forming inner forest edges
Single trees	Dissolved stand structure with single trees and a canopy density below 30%; trees loosely distributed over the whole area

4 Results

4.1 Canopy density and vertical layering

Canopy density results are given in Fig. 4. The stands with prevailing open or light canopy density dominate in the uppermost parts close to the tree line as well as near the valley floor (stands with $CD_{total} < 60\%$).
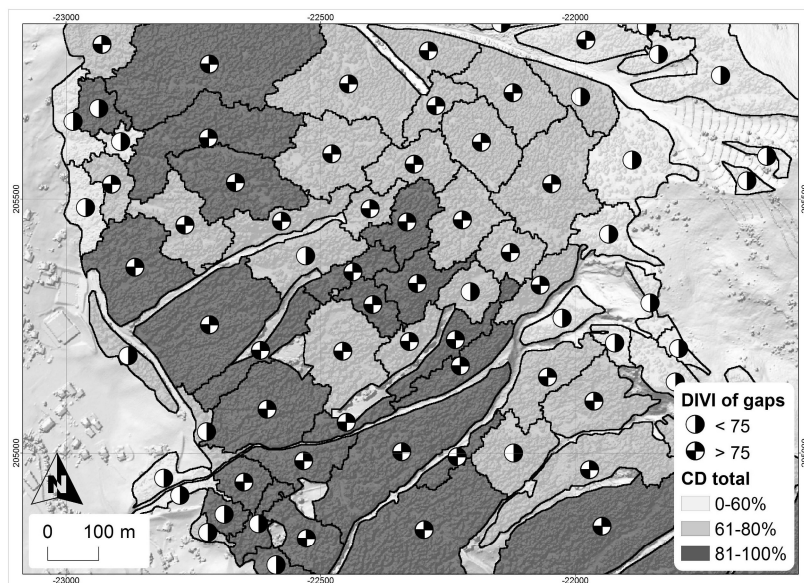
Fig. 4. Canopy density (CD_{total}) and division index (DIVI) of gaps

The latter represent the mixed and broadleaved stands which appear open due to the lack of crown reflection in the leaf-off season. The forests in the central part of the study area exhibit light structures due to the numerous rock fall channels just below the massive rock faces in that area. The open structures indicated in the upper parts correspond to a mosaic with gaps typical for high altitude forests.

Results of the layering assessment show that the homogenous pole stands in the lower part of the study area are classified as single and two-layered forests. Multi-layered structures, however, can be found in sporadic stands mainly along and in between the avalanche tracks. A comparison of the automatic layering-classification with the assessment carried out in the field results in 73% correct detections. This number increases to 82% if two- and multi-layered structures are put together into one category. Most mismatches can be found in the high-altitude forests with tree

clusters. Whilst they appear multi-layered in the field, their vertical variation falls within one height class during automatic classification. Half of the false detected sample plots are characterised by low CD values. Fig. 6 shows six examples of validation plots with layering assessments and different CD_{total}.

4.2 Gap distribution

The Division Index (DIVI) calculated on gaps (HC 0) quantifies the distribution of gaps within a stand. In order to get an idea of the total size of the unstocked area, it is necessary to combine the DIVI with the total canopy density (CD_{total}). As shown in Fig. 4 this combination very well distinguishes stands with many small gaps from others with a few large openings. In the central part of the study area a stand with a large windthrow opening is characterised by a low CD_{total} and a DIVI below 75, indicating low fragmentation. The forest stands in the northwestern part exhibit DIVI-values above 75 describing a high degree of fragmentation meaning that the unstocked area is divided into numerous small gaps.
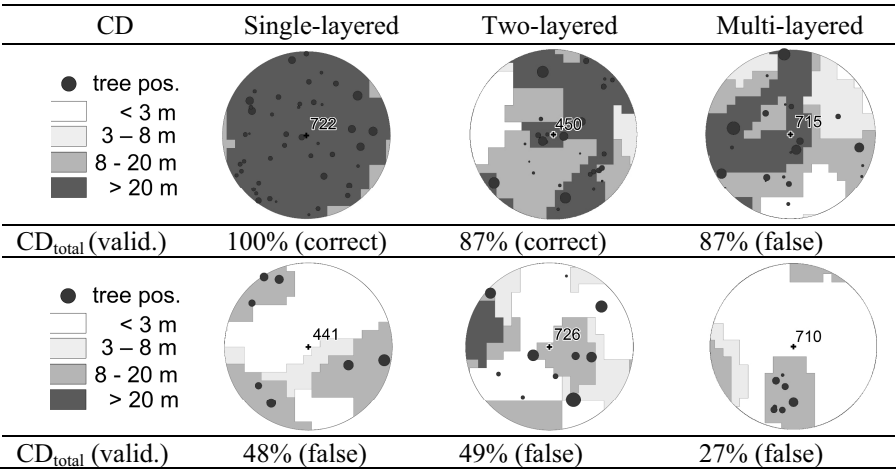
Fig. 5 Selected field validation plots (radius = 10 m) with layering and CD_{total}

The DIVI-threshold of 75 is deduced from a series of gap fragmentation patterns shown in Fig. 6. These DIVI-values are calculated on a basic gap share of 30% with increasing fragmentation.
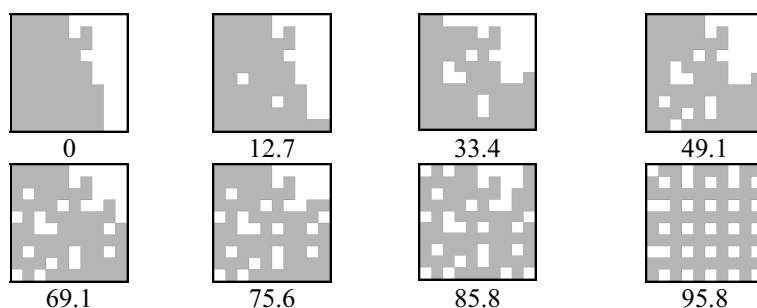
Fig. 6 Series of gap fragmentation patterns indicating the range of the Division Index (DIVI)

4.3 Canopy roughness

We regard the overall fragmentation of the height class patch mosaic as a measure for structural variability, including both the vertical as well as the horizontal structural view. Fig. 7 presents the distribution of DIVI values among the *StandType* attributes. According to a mean comparison Scheffé test (Bahrenberg et al. 1992) categories "uniform" and "irregular" are significantly different from each other (p=0.05). The category "single trees" does not significantly differ from the other two categories which may be due to the low number of samples in that class (see Fig. 7).
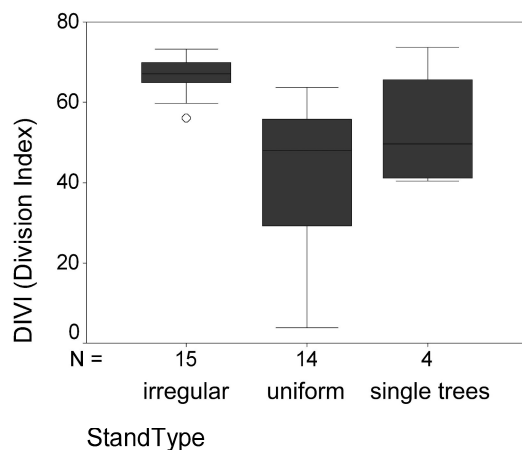
Fig. 7. Distribution of DIVI values among *StandType* categories

According to Fig. 7 we define the threshold between uniform and irregular stands at a DIVI value of 65. Using this threshold, only the open structured stands close to the tree line and the stands near the valley floor can be classified as uniform. All together only 1.5% of all stands can be classified uniform, the rest is irregularly structured.

5 Discussion

The objective of this study was to investigate if the use of landscape metrics on height class patches of the nCM allows an automatic characterisation of forest structure. For that purpose we segmented and classified the nCM and transformed it from a continuous raster surface into a polygon mosaic of vegetation height patches. This segmentation was done at the level of tree crowns and tree groups (level 1) and proved to be a straightforward way to simplify the complex canopy surface. Attempts to incorporate different morphometric derivatives such as slope and curvature of the nCM in the segmentation process did not improve the results. Thus the nCM served as the only input for that segmentation.

The segmentation of forest stands (level 2) turned out to be more difficult. It was difficult to balance the three segmentation parameters in a way that the size and shape of the resulting stand objects correlated to natural stand borders which might have appeared obvious to an expert when visually inspecting the nCM. Many stand borders looked fissured or indented, but as soon as the compactness factor was increased, narrow elongated stands (e.g. regeneration along the avalanche tracks) were dissected. Due to the iterative process the eventual parameter setting of stand segmentation must be regarded as a specific adaptation for the study area. Therefore this segmentation setting cannot be transferred to other areas.

The study showed that the automatic detection of the vertical layering corresponded fairly well with assessments in the field. If distinguishing between single- and multi-layered stands only, 82% of all the sample plots were correctly classified. Zimble et al. (2003) reached 97% correspondence between field- and LiDAR-based classification. This might be due to the fact, that they used a method based on tree-height variability and applied it to forests with canopy closures less than 60%.

As we worked with an already interpolated canopy surface model (2.5D), it was not possible to detect a forest layer which is completely covered by an dominant crown level. But because of the distinct structural characteristics of mountain forests, this restriction did not really hamper vertical layering assessment. Usually, mountain forests exhibit a relatively

open canopy and have almost no regeneration below dense layers (Ott et al. 1997). Furthermore, mountain forests are layered in tree groups, e.g. a mixture of groups with large trees and gaps filled by smaller trees. These characteristics of subalpine forersts together with the high resolution of the nCM allowed the use of already interpolated surfaces to quantify the vertical layering.

Since the perception of layering strongly depends on the scale of consideration (Parker and Brown 2000), we used forest stands as the level for detecting the vertical layering. This study showed that assessing the layering on plots with 10 m radius is possible but becomes problematic as soon as CD is very low. It turned out that the layering typology used in field inventories cannot directly be adopted to a LiDAR-nCM. The high resolution nCM we used made all the small gaps visible, which generally could not be distinguished so clearly in the field. This is why we used lower CD thresholds in our rule set for the vertical layering than suggested for field inventories.

Layering assessment very much depends on the height classification schema applied. We applied a discrete schema following the Manual for the Aerial Photo Interpretation within the Swiss Forest Inventory because it is suitable for remote-sensing based datasets and is related to practical forest management. The basic threshold of 3 m e.g. corresponds to a tree height, above which trees are regarded more stable against natural hazards and ungulate browsing. As an alternative, a height classification schema based on relative height limits for the lower, middle and upper storeys could be used. In such a schema each of the three height storeys referred to one third of the top height of a stand. By applying such a classification schema we would be able to accomodate different site and growth conditions, but at the same time it would make comparisons between stands much more complex and difficult to interpret.

The results of the study showed that forest gaps could easily be detected by means of a segmented nCM. The DIVI calculated on patches with HC 0 proved to be sufficient to describe the spatial arrangement of gaps. It was highly correlated with the distribution of gaps and helped to distinguish between stands with many small gaps and stands with only a few but larger gaps. Gap structure and distribution changes during the transition from mature to old-growth stands (Lertzmann et al. 1996). Old-growth forests are usually dominated by many small gaps and mature forests by fewer larger gaps. Generally, this developmental gap-sequence also occurs in mountain forests but is enforced by natural disturbances and the altitudinal gradient. Windthrow is the main gap-forming disturbance in mature and old-growth forests. Snowbreak mainly occurs in young pole stands. Generally, windthrow forms larger openings and snowbreak smaller gaps Gaps from both

disturbances were correctly detected in the study area. Whereas in low-altitude forests gaps occur as holes in a matrix of forest, this phenomenon reverses in high-altitude forests, where forest patches become islands in a matrix of gaps. This was also reflected in our results given in Fig. 4.

Enhancing structural diversity in mountain forests is widely recognised as a general objective of silvicultural interventions aiming at higher resistance and resilience to natural disturbances (Ott et al. 1997). In order to quantify this structural diversity or canopy roughness we used the overall DIVI on all height classes to describe if the canopy is vertically homogenous or heterogeneously structured. Whereas the concept of forest fragmentation usually implies the breakup of contiguous forest habitats by the development of settlements, roads and clearings, we were interested in the fragmentation of the nCM in terms of height variability or structuring. Within the validation plots the DIVI yielded statistically significant distinctions between uniform and irregular canopy structures. But using this threshold to distinguish between uniform and irregular types on the stand level led to almost exclusively irregular stands which did not reflect the structural reality. This might be due to the fact that the validation plots are too small to derive a suitable DIVI-threshold for stands that are on average 50 times larger than the validation plots. One would need much larger validation plots to overcome this problem.

6 Conclusions

Our object-based image analysis approach for assessing mountain forest structure confirmed the capacity of landscape metrics applied on a nCM to quantify forest structure. The selected metrics offered useful assessments of the canopy density and vertical layering as well as gap distribution and isolation. Canopy roughness could not satisfactorily be described using the DIVI based on the validation plots.

The advantage of structure assessment using landscape metrics is that it can be carried out in a transparent and easily repeatable way. But the metrics need to be calibrated with field assessments in order to link their values with local situations. Generally, this approach works particularly well in spruce-dominated mountain forests, as conifers possess well-shaped crowns and the forests are usually open and the top layer of trees is not closed. Automated structure assessment can be used in the course of protection forest planning, management and monitoring. Such an approach will and should not replace detailed field investigations, but it will help to assess structure in an area-extensive and efficient manner.

Future research will have to focus on testing different static and dynamic height classification schema. Furthermore, it might be helpful to include local maxima detection to explicitly consider tree clusters as structure types and stability features. In order to quantitatively assess the performance of such an approach, further calibration with existing structure assessments on a larger scale should be conducted. To assess scale dependency of the various metrics, corresponding sensitivity tests should be conducted prior to further application.

With this approach, resource and natural hazard managers can easily assess the structure of different forests or the same forest at different times or under different management alternatives. In the light of increasing pressure to pay attention to the protective effect of forests in natural hazard management, this forest structure assessment approach can be considered a highly valuable contribution.

References

Baatz M and Schäpe A (2000) Multiresolution segmentation—an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Blaschke T, Griesebner, G. (Eds) Angewandte Geographische Informationsverarbeitung XII. Beiträge zum AGIT-Symposium Salzburg 2000, Herbert Wichmann Verlag, Kahrlsruhe, pp. 12–23.

Bahrenberg G, Giese E, Nipper J (1992) Statistische Methoden in der Geographie. Band 2 Multivariate Statistik. B.G. Teubner Stuttgart.

Bebi P (1999) Strukturen im Gebirgswald als Beurteilungsgrundlage ausgewählter Waldwirkungen. Dissertation ETH-Zürich, 125 pp.

Benz UC, Hofmann P, Willhauk G, Lingenfelder I, Heynen M (2004) Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry & Remote Sensing 58, pp. 239-258.

Blaschke T, Tiede, D and Heurich M (2004) 3D landscape metrics to modelling forest structue and diversity based on laser scanning data. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-8/W2, Freiburg, pp. 129-132.

Brang P (2001) Resistance and elasticity: the management of protection forests, Forest Ecology and Management 145, pp. 107–117.

Dorren LKA, Berger F, Imeson A, Maier B and Rey F (2004) Integrity, stability and management of protection forests in the European Alps. Forest Ecology and Management 195, pp. 165-176.

Dorren LKA, Berger F, Le Hir C, Mermin E, Tardif P (2005) Mechanisms, effects and management implications of rockfall in forests. Forest Ecology and Management. 215, 1–3, pp 183–195.

Frehner M, Wasser B and Schwitter R (2005) Nachhaltigkeit und Erfolgskontrolle im Schutzwald - Wegleitung für Pflegemassnahmen in Wäldern mit Schutzfunktion. BUWAL, Bundesamt für Umwelt, Wald und Landschaft, Bern.

Ginzler C, Bärtschi H, Bedolla A, Brassel P, Hägeli Hauser M, Kamphues H, Laranjeiro L, Mathys L, Uebersax D, Weber E, Wicki P, Zulliger D (2005) Luftbildinterpretation LFI3 – Interpretationsanleitung zum dritten Landesforstinventar. Eidgenössische Forschungsanstalt für Wald, Schnee und Landschaft WSL. Birmensdorf. 87 pp.

Hall SA, Burke IC, Box DO, Kaufmann MR and Stoker JM (2005) Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests. Forest Ecology and Management 208, pp. 189-209.

Herold A and Ulmer U (2001) Stand stability in the Swiss National Forest Inventory: assessment technique, reproducibility and relevance. Forest Ecology and Management 145, pp. 29-42.

Hyyppä J, Yu X, Hyyppä H and Maltamo M (2006) Methods of Airborne Laser Scanning for Forest Information Extraction. In: Koukal T and Schneider W (Eds.) Proceedings of International Workshop on 3D Remote Sensing in Forestry. Vienna, pp. 63-78.

Jaeger JAG (2000) Landscape division, splitting index, and effective mesh size: new measures of landscape fragmentation. Landscape Ecology 15, pp. 115-130.

Keller M (Red) 2005 Schweizerisches Landesforstinventar. Anleitung für die Feldaufnahmen der Erhebung 2004-2007. Birmensdorf, Eidgenössische Forschungsanstalt WSL. 393 p.

Kraus K and Pfeifer N (1998) Determination of terrain models in wooded areas with airborne laser scanner data. ISPRS Journal of Photogrammetry & Remote Sensing 53, pp. 193-203.

Lang S and Tiede D (2003) vLATE Extension für ArcGIS – vektorbasiertes Tool zur quantitativen Landschaftsstrukturanalyse, ESRI Anwenderkonferenz 2003 Innsbruck. CDROM

Lertzmann KP, Sutherland GD, Inselberg, A, Saunders, SC (1996) Canopy gaps and the landscape mosaic in a coasteal temperate rain forest. Ecology 77(4) pp. 1254–1270.

Lim K, Treitz P, Wulder M, St-Onge B and Flood M (2003) LiDAR remote sensing of forest structure. Progress in Physical Geography 27, pp. 88-106.

Maier B (2007a) Zählen Messen Schätzen – Waldinventur im Standeswald. In: Malin H, Maier B, Dönz-Breuss M. (Ed.) Montafoner Standeswald – Beiträge zur Geschichte und Gegenwart eines kommunalen Forstbetriebes. Montafoner Schriftenreihe 18 pp. 43-66.

Maier B (2007b) Characterising Mountain Forest structure using Airborne Laser Scanning. Thesis. Department for Geography and Geology. Salzburg University

Maltamo H, Eerikäinen K, Pitkänen J, Hyyppä J and Vehmas M (2004) Estimation of timber volume and stem density based on scanning laser altimetry and

expected tree size distribution functions. Remote Sensing of Environment 90, pp. 319-330.

McGarigal K and Marks BJ (1995) FRAGSTATS: spatial pattern analysis program for quantifying landscape structure. USDA For. Serv. Gen. Tech. Rep. PNW-351.

Naesset E (2004) Practical Large-scale Forest Stand Inventory Using a Small-footprint Airborne Scanning Laser. Scandinavian Journal of Forest Research 19:164-179.

Ott E, Frehner M, Frey H-U, Lüscher P (1997) Gebirgsnadelwälder – Ein praxis-orientierter Leitfaden für eine standortgerechte Waldbehandlung. Paul Haupt Bern.

Parker GG and Brown MJ (2000) Forest Canopy Stratification - is it Useful? The American Naturalist 155, pp. 473–484

Schieler K and Hauk E (2001) Instruktion für die Feldarbeit – Österreichische Waldinventur 2000/2002. Forstliche Bundesversuchsanstalt FBVA. Wien.

Schönenberger W (2001) Structure of mountain forests: Assessment, impacts, management, modelling. Forest Ecology and Management 145, pp. 1-2.

Tiede D, Heurich M and Blaschke T (2004) Object-based semi automatic mapping of forest stands with Laser scanner and Multi-spectral data. In: The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVI-8W2, Freiburg, pp. 328-333.

Traub B, Klein Ch (1996) Quantitative Charakterisierung von Waldlfächenformen. Allg. Forst- u. J.-Ztg. 168(2), pp. 30-40.

Venema HD, Calamai PH, Fieguth P (2005) Forest structure optimization using evolutionalry programming and landscape ecology metrics. European Journal of Operational Research 164, pp. 423-439.

Wagner W, Eberhöfer C, Hollaus M, Summer G (2004) Robust Filtering of Airborne Laser Scanner Data for Vegetation Analysis, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences. XXXVI (Part 8/W2), pp. 56-61.

Zimble DA, Evans DL, Carlson GC, Parker RC, Grado SC, Gerard PD (2003) Characterizing vertical forest structure using small-footprint airborne LiDAR. Remote Sensing of Environment 87, pp. 171-182.

Chapter 7.3

Object detection in airborne laser scanning data - an integrative approach on object-based image and point cloud analysis

M. Rutzinger[12], B. Höfle[12], N. Pfeifer[3]

[1] alpS, Centre for Natural Hazard Management, Austria

[2] Institute of Geography, University of Innsbruck, Austria, martin.rutzinger@uibk.ac.at, bernhard.hoefle@uibk.ac.at

[3] Institute of Photogrammetry and Remote Sensing, Vienna University of Technology, Austria, np@ipf.tuwien.ac.at

ABSTRACT: In recent years object-based image analysis of digital elevation models acquired by airborne laser scanning gained in importance. Various applications for land cover classification (e.g. building and tree detection) already show promising results. Additionally to elevation rasters the original airborne laser scanning point cloud contains highly detailed 3D information. This paper introduces an integrative approach combining object-based image analysis and object-based point cloud analysis. This integrative concept is applied to building detection in the raster domain followed by a 3D roof facet delineation and classification in the point cloud. The building detection algorithm consists of a segmentation task, which is based on a fill sinks algorithm applied to the inverted digital surface model, and a rule-based classification task. The 340 buildings of the test site could be derived with 85% user's accuracy and 92% producer's

accuracy. For each building object the original laser points are further investigated by a 3D segmentation (region growing) searching for planar roof patches. The finally delineated roof facets and their descriptive attributes (e.g. slope, 3D area) represent a useful input for a multitude of applications, such as positioning of solar-thermal panels and photovoltaics or snow load capacity modeling.

1 Introduction

High-resolution imaging in satellite and airborne remote sensing led to the development of Object-Based Image Analysis (OBIA), in order to overcome problems of noise and misclassification, which occur in classification results if conventional pixel-based approaches are applied. Concerning Airborne Laser Scanning (ALS) data, OBIA has mainly been restricted to approaches using rasterized ALS data, often in combination with multispectral data. First and last echo Digital Surface Models (DSM) and Digital Terrain Models (DTM), Digital Intensity Models (DIM), as well as their derivatives from ALS, provide important additional information, such as surface geometry, topography and surface reflectance, which open new possibilities in land cover classification (e.g. Brennan and Webster 2006).

ALS systems in operational use measure X, Y, Z coordinates and intensity values for several reflections stored as 3D point cloud with intensity attribute. While the platform position is recorded with the Global Positioning System (GPS), deviations in position and angular attitude between linearly connected GPS positions are determined with the records of the Internal Measurement Unit (IMU). The point coordinates on the reflecting surface are measured taking into consideration all flight parameters and the travel time of the laser beam (Wehr and Lohr 1999).

For common applications, the ALS point cloud (raw data) is converted into regular cells so that common image analysis algorithms as well as object-based image analysis workflows can be applied. The rasterization of data can be seen as a first abstraction in the sense of the multi-resolution representation of objects. The degree of abstraction depends on the relation between point density and raster resolution (see Fig. 1). Additionally, the method of aggregation (e.g. minimum point height) and interpolation applied in the rasterization step are essential for the resulting raster. Furthermore, raster data can model only 2.5D surfaces, but not 3D objects, such as a bridge. The rasterization of the point cloud is an irreversible processing step and is therefore accompanied by a loss of information. The fact that the DSM and the DTM as well as the raw data point cloud are often deliv-

ered to the customer, and the fact that there are several efficient algorithms to classify rasterized data, lead to the development of an integrative analysis approach, combining the advantages of OBIA and Object-Based Point Cloud Analysis (OBPA) handling the high level of detail in the ALS raw data point cloud.
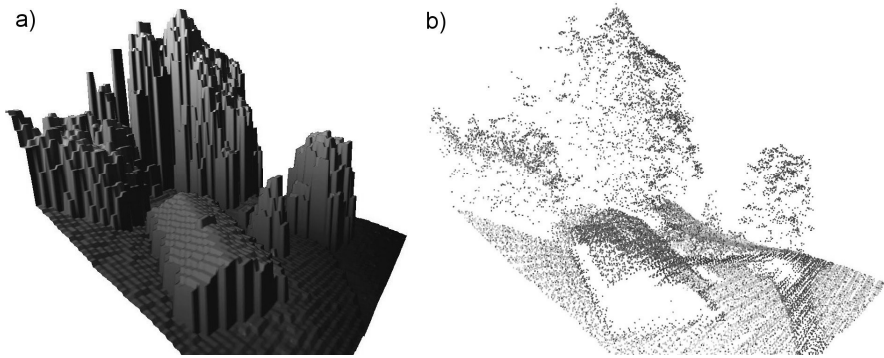
Fig. 1. Comparison of level of detail for **(a)** 2.5D DSM (1 m cell size) and **(b)** 3D raw data point cloud (ca. 8 shots/m²)

In Sect. 2 related work will be presented and Sect. 3 explains the methodology for the integrated OBIA-OBPA approach. The concept of the integrated approach of OBIA and OBPA is applied to 3D roof classification (Sect. 4), which has several applications. The results can be used to determine potential areas for solar-thermal panels and photovoltaics (Vögtle et al. 2005). The roof type of a building defined by the extracted roof segments can be used for the reconstruction of buildings in 3D city modeling (Brenner 2005). In natural hazard research the roof type is one input for estimating the snow load capacity of buildings in damage potential calculations. On the one hand, the snow load capacity depends on the building construction, roof area and slope; and on the other hand on the snow pack-properties and meteorological conditions.

2 Related work

2.1 Object-based image analysis on airborne laser scanning data

Several authors show the advantage of using ALS data in addition to other rasterized remote sensing images in order to classify the data applying an

OBIA approach. Most studies from different research areas use the multi-resolution segmentation and fuzzy classification approach, implemented in the commercial software eCognition (Benz et al. 2004).

Asselen and Seijmonsbergen (2006) use this OBIA approach to classify geomorphologic landforms such as alluvial fan, fluvial terrace, talus slope, incised channel, rock cliff and glacial landform. The segmentation is based on a slope map derived from a rasterized DTM. The rule-base is built by statistical values such as mean, standard deviation, minimum and maximum elevation derived from the DTM for geomorphologic units in a training dataset.

Hollaus et al. (2005) combine a normalized Digital Surface Model (nDSM) from ALS with a Normalized Difference Vegetation Index (NDVI) from a high resolution Composite Infrared (CIR) image to define different land cover classes, which are then used to apply hydrological roughness values for hydrological flood modeling.

Maier et al. (2006) classify forest units according to their height by segmenting single tree crowns and forest stands respectively. The inputs are an nDSM and an aspect and slope map derived from a DTM representing the varying forest growth conditions. They derive different landscape metric indices in order to describe forest structure and stability.

Tóvári and Vögtle (2004) present a method to classify buildings and high vegetation. An nDSM calculated from ALS data is segmented. The segmentation starts at a seed point and includes all objects defined by a certain minimum object height threshold. Subsequently for these segments the features first-/last echo difference, area, shape index, height texture, gradients on segment borders and intensity are calculated for each object. Then they compare a fuzzy rule-base, and a maximum likelihood classification.

Further applications of OBIA using ALS data with focus on natural hazard management are described in Rutzinger et al. (2006b).

2.2 Segmentation and classification of 3D point clouds

Concepts to segment and classify 3D point clouds collected by either Terrestrial Laser Scanning (TLS) or ALS follow the idea of an object-based analysis workflow. For each field of application only the most recent references are cited.

Tóvári and Pfeifer (2005) present a segmentation-based method to separate high object points from terrain points in order to generate a DTM. The method works partly in 3D (point cloud segmentation) and in 2.5D (ground surface interpolation). The points are first grouped by a region

growing segmentation using random seed points. Then a surface is interpolated for all points. The single segments are weighted by the mean least square distance from the interpolated surface. Then a new surface is calculated in several iterations, taking into account, that segments with a low weight contribute less to the surface calculation, and are therefore high objects.

Sithole (2005) presents different algorithms to segment and classify ALS data. In the sense of a multi-resolution approach, he distinguishes between micro (local surface roughness) and macro objects (predominant landscape objects) and classifies man-made (buildings, bridges) and natural objects (vegetation and bare Earth).

Filin and Pfeifer (2006) present a segmentation algorithm based on a clustering approach, taking into account features calculated on the basis of an adaptive neighborhood. The neighborhood parameters depend on point density, measurement accuracy, and horizontal and vertical point distribution. The features used are the point coordinates and the surface normal vector in each point, estimated from the neighboring points. The segmentation distinguishes noisy points representing vegetation and plane areas like vertical walls, bare Earth and roof facets.

Vosselman et al. (2005) show the possibilities for 3D mapping using ALS data. The software eCognition is used to segment and classify buildings with a rasterized DSM and nDSM for change detection analysis. The classified buildings are compared to a vector ground-plan in order to detect newly constructed, demolished and unchanged buildings. Furthermore a 3D city model is derived using ground-plan information to select the point cloud and to extract the roof surface type. The final model contains buildings, trees, and streets with the corresponding height from the ALS point cloud.

Reitberger et al. (2006) classify tree species using full-waveform ALS data. Full-waveform systems detect not only the first and last reflection of a laser shot, but record the whole backscattering cross-section. This gives insight into the inner structure of vegetation, as well as additional information to derive features for the classification of 3D objects. The working steps are the segmentation of individual tree crowns, the derivation of describing features, and the final classification of tree species.

An application of OBPA for glacier surface classification is described by Höfle et al. (2007). After a region growing segmentation in the 3D point cloud based on corrected intensity values the segment outlines are modeled using alpha shapes. The classes snow, firn, ice, and surface irregularities (mainly crevasses) are assigned to the segments using a supervised rule-based classification.

3 Methodology

3.1 Airborne laser scanning point cloud data management

3.1.1 LISA system

ALS delivers a dense 3D point cloud of the scanned surface (typically >1 pt/m²). For large areas (e.g. country-wide acquisition campaigns) a high amount of single point measurements arises. The LISA (LIdar Surface Analysis) system incorporates the strength of a GIS and the strength of an external data model, developed specifically for ALS data, implemented in an external spatial database (Höfle et al. 2006). Exclusively Open Source components are in use, which is enabled as many sophisticated free available programs with well-defined interfaces and open data formats already exist. LISA integrates management, processing and visualization of large ALS datasets.

3.1.2 Point cloud data model

Each area-wide ALS flight campaign typically consists of many single flight strips. Each flight strip holds the corresponding single point measurements, which are described by a timestamp, 3D coordinates (X, Y, Z) and a measure for the strength of the received signal (signal intensity). The developed data model accounts for this structure of ALS data. Additionally the data model supports unlimited multi echoes (reflections) per laser shot, as well as the storage of the air plane trajectory, which can be linked to each measurement. This allows the reconstruction of the scan geometry, which is, for example, required for the correction of signal intensities (Höfle and Pfeifer 2007).

Data management is designed to provide a fast and simple retrieval of data subsets. For large data volumes, performance is guaranteed by spatial indexing of the single flight strips (i.e. data partitioning by flight strip), as well as the spatial indexing of the laser points belonging to a strip (Bartunov and Sigaev 2007). Consequently, query performance mainly depends on the data volume, which is requested/selected, and only to a small portion on the total size of the database.

3.1.3 System components and implementation

The LISA framework consists of two main components: i) the geographical information system GRASS (GRASS Developer Team 2007) and ii) the object-relational database management system (DBMS) PostgreSQL (PostgreSQL Global Development Group 2007) with its spatial add-on

PostGIS (Refractions Research Inc. 2007). Workflows, i.e. processing and analysis algorithms, are written in the scripting language Python (Python Software Foundation 2007), which offers a great variety of scientific libraries (Jones et al. 2001) and database connectors (Cain 2007). The single programs are provided as GRASS GIS commands. The integration of all commands into GRASS GIS offers i) a good usability, ii) full access to the standard functionality of GRASS and iii) many interfaces, such as GDAL (Warmerdam 2007), to common spatial data formats.

As described above, the laser scanning point cloud is managed using PostGIS, but can be easily accessed from the GRASS GIS client side. The DBMS offers a client-server-architecture; therefore the large datasets are stored only once. The laser scanning data can be either retrieved directly, i.e. as 3D vector points, or indirectly by prior rasterization of the point cloud.

3.2 Concept of an integrative approach

The requirement to use the synergy of vector and raster algorithms, as well as the different level of detail in the point cloud and the additional information of the rasterized ALS derivatives, lead to the demand of an object-based analysis approach, which can handle both raster cells and vector points. Due to the complexity (e.g. definition of neighborhood; cf. Filin and Pfeifer 2005) and computational effort (e.g. neighbor searching, feature derivation) of area wide point cloud analysis, the combination of raster and vector analysis is the method of choice. In the sense of a "classification-based segmentation" the target class is derived in the raster domain, while the OBPA, offering higher resolution, is done only for those selected objects. The workflow presented here consists of a modular architecture for segmentation, feature calculation and error assessment running iteratively until an acceptable classification result is reached.

3.2.1 Workflow

The object-based analysis workflow comprises two main traces. The first one is the raster-based object derivation in 2.5D and the second one is the point-based object derivation using the planimetrically defined objects to derive 3D objects in higher detail (see Fig. 2). However, both follow the same basic processing chain of segmentation, feature calculation and classification. The specific selection of the parameters and features for the single processing steps is described in Sect. 4.2.
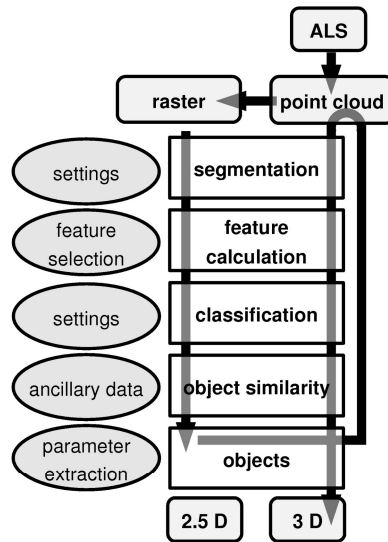
Fig. 2. Integrated object-based analysis workflow for raster and point cloud

The first derivatives from the point cloud are digital elevation models such as first echo and last echo DSMs. A first/last echo difference model (FLDM), which is an additional input in the raster segmentation, is calculated per laser shot (not per cell) using the possibility to reconstruct the flight geometry in the DBMS. The segmentation returns either polygon or line segments, depending on the segmentation method. For these segments topological, shape, raster or point cloud based features are calculated. The selected features are the input for the object classification. The classification can be done either supervised or unsupervised with the built-in classifiers of GRASS GIS or rule-based. For the latter value ranges are applied to a set of object features to define the target class. All segments within the value ranges are labeled with a weight. In the end the weights for each segment are summed up in order to get a quality measurement specifying the degree of membership of a segment to the target class. Additionally, a tool for object comparison and similarity is used either to compare multi-temporal datasets or to apply error assessment using ancillary data. The comparison of classified objects considers changes in spatial distribution, shape, and topology. Ancillary data or digitized objects in a training area are used to optimize the parameters of the raster-based workflow (Pfeifer et al. 2007).

The objects derived from the raster models are primary results, which can be used in further applications (e.g. building polygons in urban planning or line objects for semi-automatic mapping) (Rutzinger et al. 2006b).

In the integrative approach the objects derived in the raster domain are used to individually extract the original laser points for each object (i.e. building) from the spatial database. Then these points are grouped into segments. Each set of points generated in the segmentation is considered as one object primitive. Based on this, 3D point set features like slope, aspect, area, etc. are computed. These features are more accurate than features computed on the basis of rasterized images, because mixed pixels along the boundary deteriorate the results. With typical pixel sizes of 1 m, the number of boundary pixels makes a significant contribution. Another advantage is that the influence of the method of aggregation (i.e. taking the highest, lowest, etc. point per grid cell) does not influence the result. From the segmented points the outline of the 3D polygons is constructed and the final objects are classified by their 3D features.

3.2.2 Segmentation and classification in the raster domain

The raster-based segmentation is carried out independently for every target class and thus differs for specific applications. Due to the different object properties, individual segmentation approaches are used for the object of interest.

In the case of building classification, the best segmentation and classification settings are estimated for a representative training area with reference data. The parameters are estimated in a brute-force manner within user-defined intervals. Then these estimated parameters are used to segment and classify the whole data set. It is crucial that the reference data set is complete and equivalent to the object representation in the ALS data set. In case of buildings it is advisable to use manual digitized areas rather than polygons of a Digital Cadastral Map (DCM) as training data because of the fact that buildings derived from ALS contain roof overhangs, which are not present in a DCM (cf. Pfeifer et al. 2007). In an ideal case digitized training areas and an independent reference data set with the same object representation are used for parameter estimation and error assessment respectively. The segmentation method used for building detection is based on a hydrological approach implemented in GRASS GIS (Arge et al. 2003, Rutzinger et al. 2006a). In a first step high vegetation is masked out using the information on first/last echo difference. Then a fill sink algorithm is applied to the inverted DSM in order to segment the remaining high objects. The final segment outline is defined by a height threshold (segHeight), which suppresses low non-building objects. Furthermore the

segment shape is adapted by morphological opening (segOpening, radius of the structure element) and finally very small segments considered as noise are erased by a minimum area criterion (segArea).

The user selected features (standard deviation of height, shape index, area, mean first/last echo difference, etc.) are first derived for the training data set to estimate the classification rules. Then the rules are adjusted automatically in order to keep errors between training and classified polygons as small as possible. A quality map is generated from the applied weights to each feature layer. Only those objects reaching a particular, defined quality criterion are used for the final, crisp classification. The other areas could be used in a second classification with adjusted settings in the next iteration. Finally, the classified objects are used to select the corresponding points in the spatial database.

3.2.3 Object generation in 3D

Point cloud segmentation
Segmentation is a process to aggregate information originating in digital image processing. Here, it is applied to point sets, which require that concepts applicable to images are transferred to point clouds (Haralick and Shapiro 1992).

The purpose of segmentation is to split a set of spatially distributed measurement icons (originally pixels) into disjoint sets. The union of the generated sets is the original. Grouping of the original measurements is done on the basis of a homogeneity criterion (e.g. gray value for pixels, or belonging to the same plane for points, etc.) requiring that such a set is spatially connected. Additionally, there is a constraint that joining two neighboring sets will violate the homogeneity criterion. This requires that neighborhood is defined, which is trivial for images using the 4- or 8-neighborhood. For point clouds it can be the set of fixed distance neighbors or the k nearest neighbors. A detailed discussion on neighborhoods in point clouds of airborne laser scanning can be found in Filin and Pfeifer (2005). The definition above normally requires the additional definition of a rejection set, i.e., a set of points, which do not belong to any other segment. Of course, the points in that set are spatially not connected.

The segmentation method applied is based on region growing and uses the neighborhood definition of k nearest neighbors for each point. We chose the latter because it adapts to variations in point density and for reasons of computational efficiency. These neighbors are used in the first preprocessing step to estimate the normal vector for each point. This is done by least squares fitting of a plane (regression) to each point and its neighbors, which can be interpreted as estimating the tangent plane of that

surface point. This plane's normal is the normal vector to each point. Therefore the inputs to the region growing are the points and their normal vectors.

The region growing randomly picks a seed point and then examines the k nearest neighboring points, termed the candidate points. Random seed point selection is justified because it can be assumed that all points belong to the same target class due to the pre-processing and selection of the OBIA approach (see Sect. 3.2.2). Furthermore the parameters are updated during region growing (see below). A number of criteria are checked. During region growing an adjusting plane is also estimated for the points of a segment. This is basically the same estimation process as before, however now each point that is accepted as belonging to the segment is used to renew the estimated normal vector. Plane estimation can only be initiated after a segment has grown to at least three points, so first the plane is instantiated by the seed point and its normal vector. If a candidate point is accepted as belonging to the segment, its k nearest neighbors become new candidate points.

Candidate points will be connected to the segment if they fulfill three criteria:

- similarity of normal vectors (a)
- distance of candidate point to the adjusting plane (r)
- distance between current point and candidate point (d)

The first criterion (similarity of the normal vectors) determines that the angle difference between the segment normal and the point normal should be under a predefined threshold (e.g. 10°). The second criterion, the distance of the candidate point to the segment plane, must also fall below a threshold (e.g. 15 cm). Finally, each point and the candidate points must not be further apart than a defined distance, measured in 3D. Growing continues until no further points, which fulfill the criteria, can be found.

This means that segments will constitute planar patches. They are spatially connected (distance d), and points on rounded edges are excluded because of the normal vector similarity requirement.

This process is governed by four parameters: k, a, r, and d. The density and noise of the data can be considered by setting the number of neighbors and the maximum distance for accepting points.

Points in the vegetation, on cars, traffic signs, overhead cables, chimneys, etc. typically form one point segments and are put in the rejection set. Of course, this is also a question of point density, and the above holds for point sets with approximately one point per square meter.

Triangulation

The corresponding points of each segment are triangulated separately in iterations over all segments. A triangulation of a point set results in a convex hull and therefore information about the segment shape is lost. A shape adjustment is applied to every segment by removing the longest edges of the triangulated segment. As shown in Fig. 3a, the minimum length of the triangle edge to be removed must be larger than the average point distance. If it is smaller, holes within the segment appear. If the chosen value for the minimum length is too high, the segment area is overestimated and the shape at concave outline sections is not well defined. After the shape adjustment, the remaining triangles are merged to the final segment (see Fig. 3b). The triangulation is done only in 2D, but the vertices of the final segment outline are 3D points.
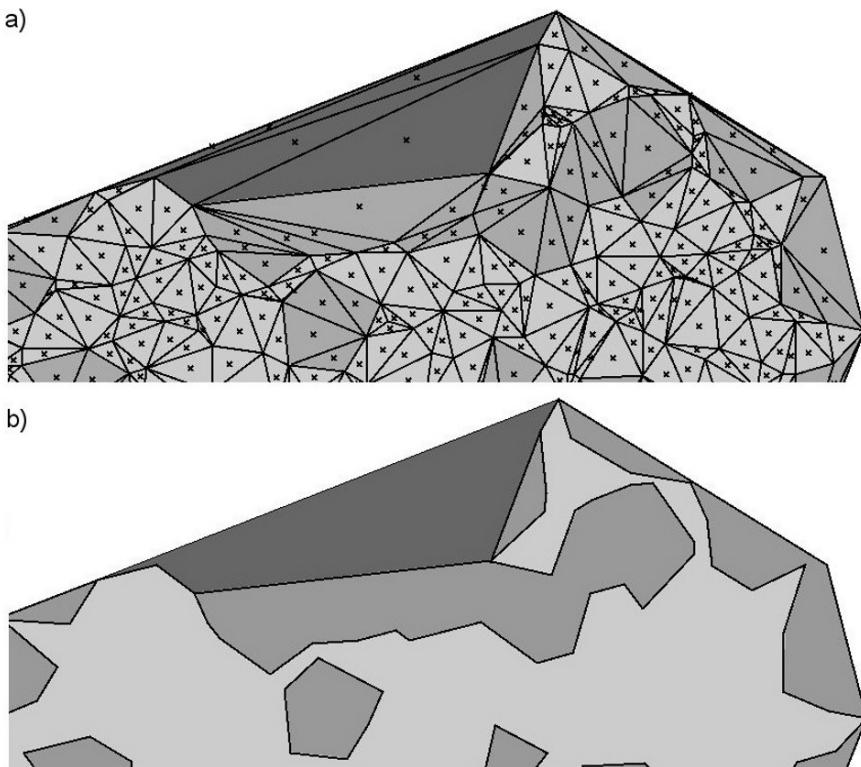
Fig. 3. (a) Object outline determined by the minimum triangle edge length (dark gray – convex hull, medium gray – 5m, light gray 1m), **(b)** final segments from merged triangles

The features for the classification are calculated on the basis of the 3D point cloud used in the triangulation. These are the average point distance to the estimated plane, slope and aspect, and area in 3D (projected on the segment plane) and 2D respectively.

4 Application: Classification of roof facets

4.1 Test site

The tested data is a subset of the ALS acquisition campaign of the city of Innsbruck (Tyrol/Austria), which was carried out with an ALTM 3100 Optech scanner in 2005. An area of 269.6 km² is covered by 110 flight strips with an average point density of 4 shots/m². From the raw data point cloud, which is stored in the LISA system, a last echo DSM and a FLDM, both in 1 m resolution, are used as input elevation models in further processing.

4.2 Parameter settings

The parameters for the detection of building footprints (see Sect. 3.2.2) are estimated for a test area where a DCM is available (segHeight 0.5 m, segArea 50 m², segOpening 2). For these building footprints, the last echoes are selected from the spatial database in order to segment the point cloud by region growing. The chosen parameters for the roof segmentation (k 70, a 20°, r 2 m, d 7 m) are explained in Sect. 3.2.3. Additionally, the features slope, aspect, noise, 3D and 2D area are calculated for every point segment. The optimal shape representation of the roof facets is produced with a minimum triangle edge length of 5 m. Finally, the roof facets are classified by their slope into five classes (from 0 to 40° in 10° steps and one class for more than 40°).

4.3 Results

Figure 4 and 5 show the results of the integrative approach for two selected samples. Figure 4c shows that the shape adjustment also works for extreme concave objects such as the X-shaped building on the lower left site. The comparison of the DCM (Fig. 4a) and the detected building outlines (Fig. 4b) show that it is possible to use the OBIA approach on ALS data for map updating. The W-shaped building on the upper left could be detected, even though it was not included in the DCM. The different object representation

of buildings in the DCM and in ALS data becomes apparent at the left roof edge of the largest building. The single 3D roof facets from the OBPA on the basis of the OBIA are seen in Fig. 4c and 4d.
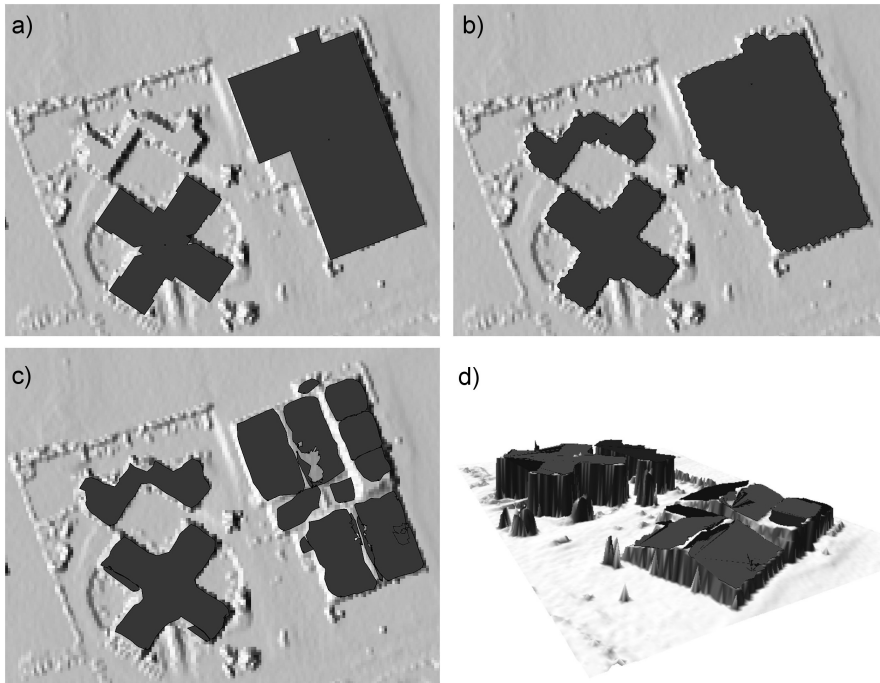
Fig. 4. (a) DCM, **(b)** buildings from OBIA, **(c)** roof facets from OBPA, **(d)** 3D view of roof facets

For a test site containing 340 buildings located in the city of Innsbruck (see Sect. 4.1) error assessment was calculated. A subsample is shown in Fig. 5. Compairing the total building areas between reference data (DCM) and buildings derived by the OBIA approach (see Sect. 3.2.2) a user's accuracy of 84.89% and a producer's accuracy of 91.87% could be reached. These error values still suffer from uncertainties caused by the different object representation in ALS data and DCM (e.g. roof overhangs or missing reference buildings). The 3D roof facets and the derived slope classes look very promising but could not be evaluated because of lacking reference data.
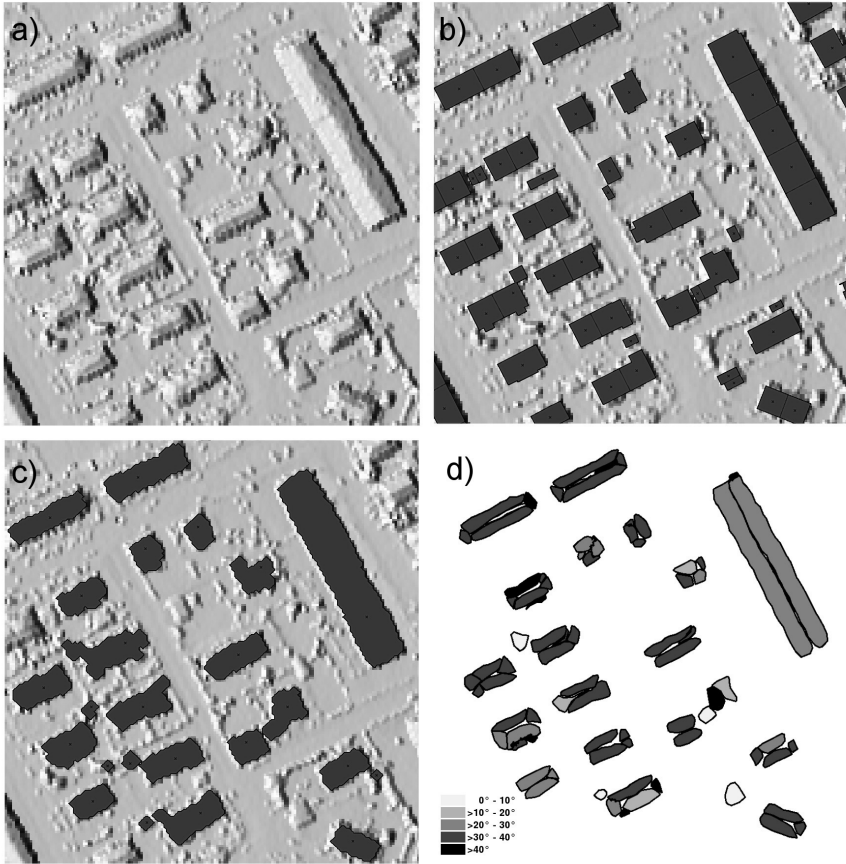
Fig. 5. (a) last echo DSM, **(b)** DCM, **(c)** buildings from OBIA, **(d)** roof facets from OBPA classified by slope

5 Conclusion

This paper demonstrates the demand to extend the OBIA approach to 3D data sources like ALS data. The OBIA approach must not be limited to 'image' analysis, but must also handle irregular distributed point clouds (OBPA). The combination of image analysis tools and GIS lead to new innovative solutions in handling high resolution data, and in classifying and determining objects in 3D. This enables the use of highly accurate products in several applications. The point cloud segmentation by region growing is already showing convincing results, although its application and settings are still subject to further research.

The core function of the integrative approach is the iterative access to the original point cloud, making this approach a flexible solution.

Vosselman et al. (2005) state "the future's information society will require up-to-date object-oriented three-dimensional geo-information." ALS, as well as TLS, with different data structures, demands a development in the concept of OBIA in order to provide methods for irregular distributed point data handling and 3D object classification.

New sensors in the field of laser scanning, such as full-waveform analysis, require new innovative concepts where integrative approaches can contribute to design efficient strategies for full- and semi-automatic object analysis.

References

Arge L, Chase J, Toma L, Vitter J, Wickremesinghe R, Halpin P, Urban D (2003) Efficient flow computation on massive grid terrain datasets. GeoInformatica, vol 7 (4), pp 283-313

Asselen S, Seijmonsbergen AC (2006) Expert-driven semiautomated geomorphological mapping for a mountainous area using a laser DTM. Geomorphology, vol 78 (3-4), pp 309-320

Bartunov O, Sigaev T (2007) GiST for PostgreSQL, http://www.sai.msu.su/~megera/postgres/gist/, last accessed: 9.3.2007

Benz UC, Hofmann P, Willhauck G, Lingenfelder I, Heynen M (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. ISPRS Journal of Photogrammetry and Remote Sensing, vol 58 (3-4), pp 239-258

Brennan R, Webster TL (2006) Object-oriented land cover classification of lidar-derived surfaces. Canadian Journal of Remote Sensing, vol 32 (2), pp 162-172

Brenner C (2005) Building reconstruction from images and laser scanning. International Journal of Applied Earth Observation and Geoinformation, vol 6 (3-4), pp 187-198

Cain DJM (2007) PyGreSQL – PostgreSQL module for Python, http://www.pygresql.org, last accessed: 9.3.2007

Filin S, Pfeifer N (2005) Neighborhood systems for airborne laser scanner data. Photogrammetric Engineering and Remote Sensing, vol 71 (6), pp 743-755

Filin S, Pfeifer N (2006) Segmentation of airborne laser scanning data using a slope adaptive neighborhood. ISPRS Journal of Photogrammetry and Remote Sensing, vol 60 (2), pp 71-80

GRASS Development Team (2007) Geographic Resources Analysis Support System (GRASS) Software, ITC-irst, Trento Italy, http://grass.itc.it, last accessed: 9.3.2007

Haralick RM, Shapiro LG (1992) Computer and Robot Vision. Addison-Wesley Longman Publishing Co., Boston Massachusetts

Höfle B, Geist T, Rutzinger M, Pfeifer N (2007) Glacier surface segmentation using airborne laser scanning point cloud and intensity data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Espoo, Finland, on CD

Höfle B, Pfeifer N (2007) Correction of laser scanning intensity data: data and model-driven approaches. ISPRS Journal of Photogrammetry and Remote Sensing, doi:10.1016/j.isprsjprs.2007.05.008, accepted for publication

Höfle B, Rutzinger M, Geist T, Stötter J (2006) Using airborne laser scanning data in urban data management - set up of a flexible information system with open source components. In: Fendel E, Rumor M (eds) Proceedings of UDMS 2006: 25th Urban Data Management Symposium, Aalborg, on CD

Hollaus M, Wagner W, Kraus K (2005) Airborne laser scanning and usefulness for hydrological models. Advances in Geosciences, vol 5, pp 57-63

Jones E, Oliphant T, Peterson P et al. (2001-): SciPy: Open Source Scientific Tools for Python, http://www.scipy.org, , last accessed: 9.3.2007

Maier B, Tiede D, Dorren L (2006) Assessing mountain forest structure using airborne laser scanning and landscape metrics. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol XXXVI (4/C42), on CD

Pfeifer N, Rutzinger M, Rottensteiner F, Mücke W, Hollaus M (2007) Extraction of building footprints from airborne laser scanning: comparison and validation techniques. Proceedings of the Urban Remote Sensing Joint Event 2007: 4th IEEE/GRSS ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas 6th International Symposium of Remote Sensing over Urban Areas, Paris, on CD

PostgreSQL Global Development Group (2007) PostgreSQL 8.1 Documentation, http://www.postgresql.org/docs/manuals/, last accessed: 9.3.2007

Python Software Foundation (2007) Python programming language, http://www.python.org, last accessed: 9.3.2007

Refractions Research Inc. (2007): PostGIS: Geographic Objects for PostgreSQL, PostGIS Manual, http://postgis.refractions.net/docs/, last accessed: 9.3.2007

Reitberger J, Krzystek P, Stilla U (2006) Analysis of full waveform lidar data for tree species classification. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol XXXVI (3), pp 228-233

Rutzinger M, Höfle B, Geist T, Stötter J (2006a) Object-based building detection based on airborne laser scanning data within GRASS GIS environment. In: Fendel E, Rumor M (eds) Proceedings of UDMS 2006: 25th Urban Data Management Symposium, Aalborg, on CD

Rutzinger M, Höfle B, Pfeifer N, Geist T, Stötter J (2006b) Object-based analysis of airborne laser scanning data for natural hazard purposes using open source components. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol XXXVI (4/C42), on CD

Sithole G (2005) Segmentation and classification of airborne laser scanner data. Publications on Geodesy of the Netherlands Commission of Geodesy, Delft

Tóvári D, Pfeifer N (2005) Segmentation based robust interpolation – a new approach to laser data filtering. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol XXXVI (3/W19), pp 79-84

Tóvári D, Vögtle T (2004) Classification methods for 3D objects in laserscanning data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol XXXV (B3), on CD

Vögtle T, Steinle E, Tóvári D (2005) Airborne laserscanning data for determination of suitable areas for photovoltaics. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, vol XXXVI (3/W19), pp 215-220

Vosselman G, Kessels P, Gorte B (2005) The utilisation of airborne laser scanning for mapping. International Journal of Applied Earth Observation and Geoinformation, vol 6 (3/4), pp 177-186

Warmerdam F (2007) GDAL - Geospatial Data Abstraction Library. http://www.gdal.org, last accessed: 9.3.2007

Wehr A, Lohr U (1999) Airborne laser scanning - an introduction and overview. ISPRS Journal of Photogrammetry and Remote Sensing, vol 54 (2-3), pp 68-82

Chapter 7.4

Support Vector Machine classification for Object-Based Image Analysis

A. Tzotsos, D. Argialas

Laboratory of Remote Sensing, Department of Surveying, School of Rural and Surveying Engineering, National Technical University of Athens, Greece

KEYWORDS: Machine Learning, Computational Intelligence, Multi-scale Segmentation, Remote Sensing

ABSTRACT: The Support Vector Machine is a theoretically superior machine learning methodology with great results in pattern recognition. Especially for supervised classification of high-dimensional datasets and has been found competitive with the best machine learning algorithms. In the past, SVMs were tested and evaluated only as pixel-based image classifiers. During recent years, advances in Remote Sensing occurred in the field of Object-Based Image Analysis (OBIA) with combination of low level and high level computer vision techniques. Moving from pixel-based techniques towards object-based representation, the dimensions of remote sensing imagery feature space increases significantly. This results to increased complexity of the classification process, and causes problems to traditional classification schemes. The objective of this study was to evaluate SVMs for their effectiveness and prospects for object-based image analysis as a modern computational intelligence method. Here, an SVM approach for multi-class classification was followed, based on primitive image objects provided by a multi-resolution segmentation algorithm. Then, a feature selection step took place in order to provide the features for classification which involved spectral, texture and shape information. After the feature selection step, a module that integrated an SVM classifier and the segmentation algorithm was developed in C++. For training the SVM, sample image objects derived from the segmentation procedure were used. The proposed classification procedure followed, resulting in the

final object classification. The classification results were compared to the Nearest Neighbor object-based classifier results, and were found satisfactory. The SVM methodology seems very promising for Object-Based Image Analysis and future work will focus on integrating SVM classifiers with rule-based classifiers.

1 Introduction

1.1 Knowledge-based image classification and Object Oriented Image Analysis

In recent years, research has progressed in computer vision methods applied to remotely sensed images such as segmentation, object oriented and knowledge-based methods for classification of high-resolution imagery (Argialas and Harlow 1990, Kanellopoulos et al. 1997). In Computer Vision, image analysis is considered in three levels: low, medium and high. Such approaches were usually implemented in separate software environments since low and medium level algorithms are procedural in nature, while high level is inferential and thus for the first procedural languages are best suitable while for the second an expert system environment is more appropriate.

New approaches were proposed, during recent years in the field of Remote Sensing. Some of them were based on knowledge-based techniques in order to take advantage of the expert knowledge derived from human photo-interpreters (Argialas and Goudoula 2003, Yooa et al 2005). Especially within an Expert System environment, the classification step can be implemented through logic rules and heuristics, operating on classes and features, which are implemented by the user through an object-oriented representation (Moller-Jensen 1997, De Moraes 2004).

Very recently a new methodology called Object Oriented Image Analysis was introduced, integrating low-level, knowledge-free segmentation with high-level, knowledge-based fuzzy classification methods. This new methodology was implemented through commercial software, eCognition, which included an object-oriented environment for the classification of satellite imagery (Baatz and Shape 2000, Benz et al. 2004).

1.2 Computational Intelligence methods in Remote Sensing

Other fields of Artificial Intelligence have been developed and applied recently. Computers became more capable of performing calculations and

thus, a new field of A.I. called Computational Intelligence and Machine Learning evolved. In this field, techniques like Neural Networks, Fuzzy Systems, Genetic Algorithms, Intelligent Agents and Support Vector Machines (Negnevitsky 2005) are included. Machine learning is an integral part of Pattern Recognition, and its applications such as classification (Theodoridis and Koutroumbas 2003). Digital remote sensing used in the past pattern recognition techniques for pixel classification purposes, while recently modern machine learning techniques have been implemented to achieve superior classification results (Brown et al 2000, Fang and Liang 2003, Huang et al 2002, Theodoridis and Koutroumbas 2003, Foody and Mathur 2004).

The Support Vector Machine (SVM) is a theoretically superior machine learning methodology with great results in classification of high-dimensional datasets and has been found competitive with the best machine learning algorithms (Huang et al 2002, Foody and Mathur 2004). In the past, SVMs were tested and evaluated only as pixel based image classifiers with very good results (Brown et al 2000, Huang et al 2002, Foody and Mathur 2004, Melgani and Bruzzone 2004).

Furthermore, for remote sensing data, it has been shown that Support Vector Machines have great potential, especially for hyperspectral data, due to their high-dimensionality (Mercier and Lennon 2003, Melgani and Bruzzone 2004). In recent studies, Support Vector Machines were compared to other classification methods, such as Neural Networks, Nearest Neighbor, Maximum Likelihood and Decision Tree classifiers for remote sensing imagery and have surpassed all of them in robustness and accuracy (Huang et al 2002, Foody and Mathur 2004).

1.3 Research objectives

The objective of this study was to evaluate the effectiveness and prospects of SVMs for object-based image analysis, as a modern computational intelligence method.

A secondary objective was to evaluate the accuracy of SVMs compared to a simpler and widely used classification technique in object-based image analysis such as the Nearest Neighbor. Also, the computational efficiency and training size requirements of SVMs were addressed.

2 Methodology

2.1 Multi-scale segmentation

Image segmentation is an integral part of Object-Based Image Analysis methodology (Benz et al 2004). The digital image is no longer considered as a grid of pixels, but as a group of primitive and homogeneous regions, called image objects. The object oriented representation provides the classification process with context and shape information that could not be derived from single pixels. These are very important factors to photo-interpretation and image understanding (Biederman 1985, Lillesand and Kiefer 1987, Sonka et al 1998). Objects can be more intelligent than pixels, in a sense of knowing their "neighbours" and the spatial or spectral relations with and among them (Baatz and Shape 2000).

In order to perform object-based classification, a segmentation algorithm is needed to provide knowledge-free primitive image objects. For this research effort the MSEG multi-scale segmentation algorithm was used (Tzotsos and Argialas 2006, Tzotsos et al. 2007). The main reason for this choice was that it has an open architecture to implement new features in C++. For evaluation purposes, the Multiresolution Segmentation algorithm from eCognition was also used (Baatz and Shape 2000).

2.2 Support Vector Machines

Recently, particular attention has been dedicated to Support Vector Machines as a classification method. The SVM approach seeks to find the optimal separating hyperplane between classes by focusing on the training cases that are placed at the edge of the class descriptors (Fig. 1). These training cases are called support vectors. Training cases other than support vectors are discarded. This way, not only an optimal hyperplane is fitted, but also less training samples are effectively used; thus high classification accuracy is achieved with small training sets (Mercier and Lennon 2003). This feature is very advantageous, especially for remote sensing datasets and more specifically for Object-based Image Analysis, where object samples tend to be less in number than in pixel based approaches.

A complete formulation of Support Vector Machines can be found in a number of publications (Cortes and Vapnik 1995, Vapnik 1995, 1998, Theodoridis and Koutroumbas 2003). Here, the basic principles will be presented and then their implementation and application to Object-Based Image Analysis will be evaluated.

Let us consider a supervised binary classification problem. If the training data are represented by $\{x_i, y_i\}$, $i = 1, 2, ..., N$, and $y_i \in \{-1, +1\}$, where N is the number of training samples, $y_i = +1$ for class ω_1 and $y_i = -1$ for class ω_2. Suppose the two classes are linearly separable. This means that it is possible to find at least one hyperplane defined by a vector w with a bias w_0, which can separate the classes without error:

$$f(x) = w \cdot x + w_0 = 0 \tag{1}$$

To find such a hyperplane, w and w_0 should be estimated in a way that $y_i(w \cdot x_i + w_0) \geq +1$ for $y_i = +1$ (class ω_1) and $y_i(w \cdot x_i + w_0) \leq -1$ for $y_i = -1$ (class ω_2). These two, can be combined to provide equation 2:

$$y_i(w \cdot x_i + w_0) - 1 \geq 0 \tag{2}$$

Many hyperplanes could be fitted to separate the two classes but there is only one optimal hyperplane that is expected to generalize better than other hyperplanes (Fig.1). The goal is to search for the hyperplane that leaves the maximum margin between classes. To be able to find the optimal hyperplane, the support vectors must be defined. The support vectors lie on two hyperplanes which are parallel to the optimal and are given by:

$$w \cdot x_i + w_0 = \pm 1 \tag{3}$$

If a simple rescale of the hyperplane parameters w and w_0 takes place, the margin can be expressed as $\dfrac{2}{\|w\|}$. The optimal hyperplane can be found by solving the following optimization problem:

$$\text{Minimize } \frac{1}{2}\|w\|^2 \tag{4}$$

$$\text{Subject to } y_i(w \cdot x_i + w_0) - 1 \geq 0 \quad i = 0, 1, ... N$$

Using a Lagrangian formulation, the optimal hyperplane discriminant function becomes:

$$f(x) = \sum_{i \in S} \lambda_i y_i (x_i x) + w_0 \tag{5}$$

where λ_i are the Lagrange multipliers and S is a subset of training samples that correspond to non-zero Lagrange multipliers. These training samples are called *support vectors*.

In most cases, classes are not linearly separable, and the constrain of equation 2 cannot be satisfied. In order to handle such cases, a cost function can be formulated to combine maximization of margin and minimization of error criteria, using a set of variables called slack variables ξ (Fig. 1). To generalize the above method to non-linear discriminant functions, the Support Vector Machine maps the input vector x into a high-dimensional feature space and then constructs the optimal separating hyperplane in that space. One would consider that mapping into a high dimensional feature space would add extra complexity to the problem. But, according to the Mercer's theorem (Vapnik 1998, Theodoridis and Koutroumbas 2003), the inner product of the vectors in the mapping space, can be expressed as a function of the inner products of the corresponding vectors in the original space.

The inner product operation has an equivalent representation:

$$\Phi(x)\Phi(z) = K(x, z) \qquad (6)$$

where K(x,z) is called a kernel function. If a kernel function K can be found, this function can be used for training without knowing the explicit form of Φ .
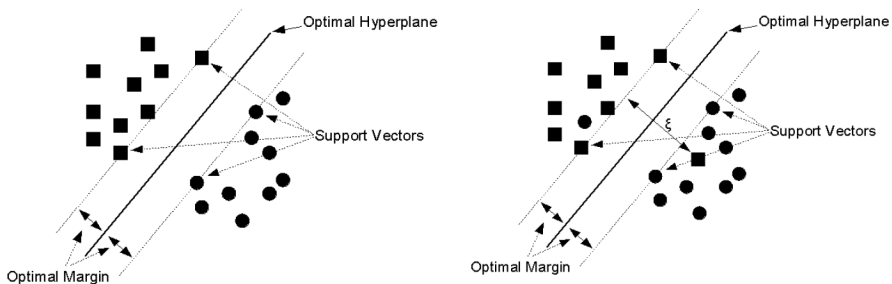
Fig. 1. Left: The case of linear separable classes. Right: The case of non linear separable classes. ξ measures the error of the hyperplane fitting

2.3 SVM Multi-class Classification

The SVM method was designed to be applied only for two class problems. For applying SVM to multi-class classifications, two main approaches have been suggested. The basic idea is to reduce the multi-class to a set of binary problems so that the SVM approach can be used.

The first approach is called "one against all". In this approach, a set of binary classifiers is trained to be able to separate each class from all others.

Then each data object is classified to the class for which the largest decision value was determined (Hsu and Lin 2002). This method trains N SVMs (where N is the number of classes) and there are N decision functions. Although it is a fast method, it suffers from errors caused by marginally imbalanced training sets.

The second approach is called "one against one". In this, a series of classifiers is applied to each pair of classes, with the most commonly computed class kept for each object. Then a max-win operator is used to determine to which class the object will be finally assigned. The application of this method requires N(N-1)/2 machines to be applied. Even if this method is more computationally demanding than the "one against all" method, it has been shown that can be more suitable for multi-class classification problems (Hsu and Lin 2002), thus it was selected for SVM object-based image classification.

2.4 Implementation

In order to apply the SVM methodology for Object-Based Image Analysis, it was necessary to perform a segmentation task. The MSEG algorithm was selected to perform segmentation at multiple scales (Tzotsos and Argialas 2006, Tzotsos et al. 2007) and to produce primitive image objects to be used for SVM classification.

For the primitive objects to be usable by a classification algorithm there was a need to implement an interface between image objects and the classifier. An extra module was implemented into the MSEG core library to add the functionality of selecting sample objects. Because a comparison was to be made with the Nearest Neighbor classifier used in eCognition, a TTA Mask (eCognition user guide 2005) import module was also implemented, so that the training object selection process would be as transparent and objective as possible. For the object feature representation, XML was selected, so that open standards are followed.

A widely used SVM library called LIBSVM (Chang and Lin 2001) was then modified to be able to handle XML files as well as training samples from the MSEG algorithm. A classifier module was then implemented as a modified version of LIBSVM.

The proposed Object-based Image Analysis system worked as following. A segmentation procedure was carried out with parameters like scale, color and shape. The properties of the primitive objects were then computed and exported to XML. A TTA Mask file along with its attribute table was imported to the system and training object samples were defined. A

training set of feature vectors was exported from the MSEG algorithm and was used for training the SVM module.

The SVM module is capable to use 4 types of kernels for training and classification: Linear, Polynomial, Radial Basis Function and Sigmoid. All the above kernels follow Mercer's theorem and can be used for mapping the feature space into a higher dimensional space to find an optimal separating hyperplane. In literature, there have been many comparison studies between the most common kernels (Huang et al 2002, Mercier and Lennon 2003). For pixel-based classification of remotely sensed data, it has been shown that local kernels such as RBF can be very effective and accurate. Also, the linear kernel is a special case of the RBF kernel, under specific parameters (Hsu and Lin 2002). Based on the above, for the current study only RBF kernels were used.

For the training of the SVM classifier, an error parameter C and a kernel parameter γ had to be obtained. In order to find the optimal parameters for the RBF kernel function a cross-validation procedure was followed. First the training set was scaled to the range of [-1, +1] to avoid features in greater numerical ranges dominating those in smaller ranges (Negnevitsky 2005). Then, the training set was divided to many smaller sets of equal size. Sequentially each subset was tested using the classifier trained by the remaining subsets. This way each image object was predicted once during the above process. The overall accuracy of the cross-validation was the percentage of correctly classified image objects.

After the cross-validation delivered the optimal parameters for the SVM classifier, the training set was used to train the SVM. Then the classifier was provided with all image primitive objects so to derive the final object-based classification. The output of the above procedure was a classification map as well as an updated XML representation of the segmentation level.

3 Discussion of Results

For the evaluation of the above methodology, initially a Landsat TM image was used. For comparison purposes, an object-based classification of the same image was performed in eCognition. The training samples in both cases were the same (a TTA mask file) and were obtained by the eCognition user guide (2005) for objective evaluation. The original Landsat TM image and the training samples are presented in Figure 2. A reference dataset was also derived from photo-interpretation and was used to compute confusion matrices.

First, the training samples were applied to small primitive objects that were derived by eCognition with scale parameter 10 and by MSEG with scale parameter 100. Scale parameters are implementation dependent (Tzotsos and Argialas 2006). For the export of training samples, the minimum overlap for sample object was set to 50% (eCognition User Guide 2005). The overall accuracy of the Nearest Neighbor (NN) method, based on the reference dataset was 85.6%. The overall accuracy of the object-based SVM classification was 90.6% (Figure 3, Tables 1 and 2).

○ Woodland General
○ Grassland General
○ Impervious General
● Waterbodies
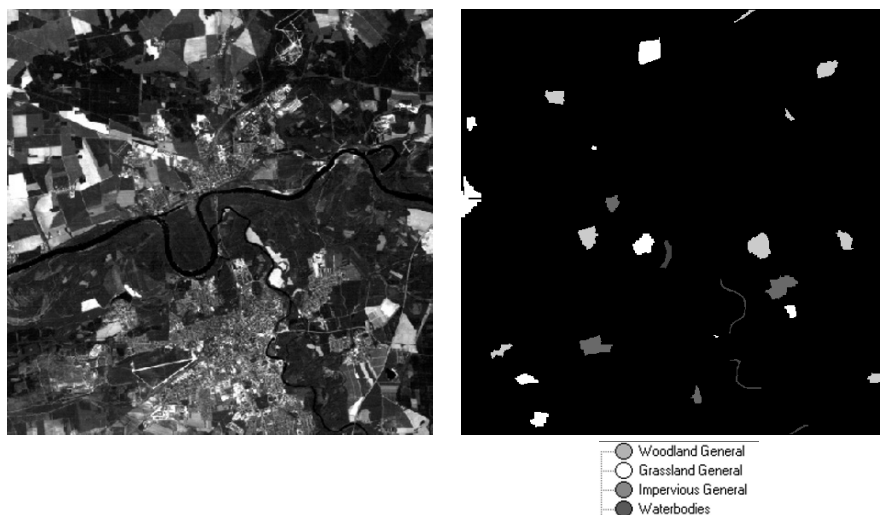
Fig. 2. Left: the original Landsat TM image (source: eCognition User Guide 2005). Right: The training set of class samples

Then, in order to test the generalization ability of both classifiers, an error was introduced into the training samples, in the form of not using a minimum overlap restriction for sample object selection. This way, more training objects were selected with errors derived from the segmentation procedures. An interesting observation was that the SVM behaved better than the NN at the second training set and provided better classification results (Tables 3 and 4) giving an overall accuracy of 86.0% against 84.1% for the NN. Both classification results are presented in Figure 4.

The evaluation of the methodology also included testing on very high resolution data (Toposys GmbH). An aerial scanner image with resolution of 0.5m was used for classification purposes. This image was selected because outperforms in resolution all commercial satellite data available today. Again, the SVM classification method was compared to the Nearest

Neighbour classifier of eCognition software. In both cases 4 basic land cover classes were selected and the same samples were used to train the classifiers. In Figure 5, the original dataset is presented. A reference data-set was also created through photo-interpretation in order to compute con-fusion matrices (Figure 5).

Woodland General
Grassland General
Impervious General
Waterbodies
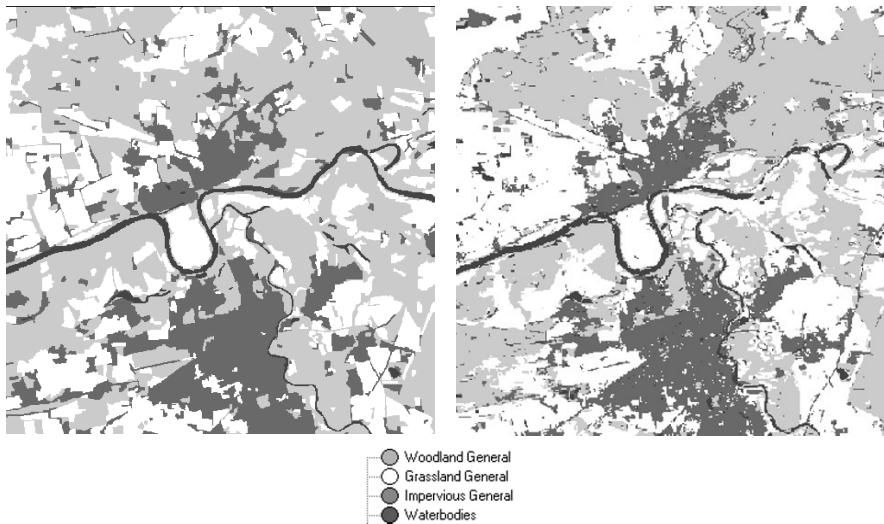
Fig. 3. Left: eCognition classification result with Nearest Neighbor. Right: MSEG classification result with SVM.
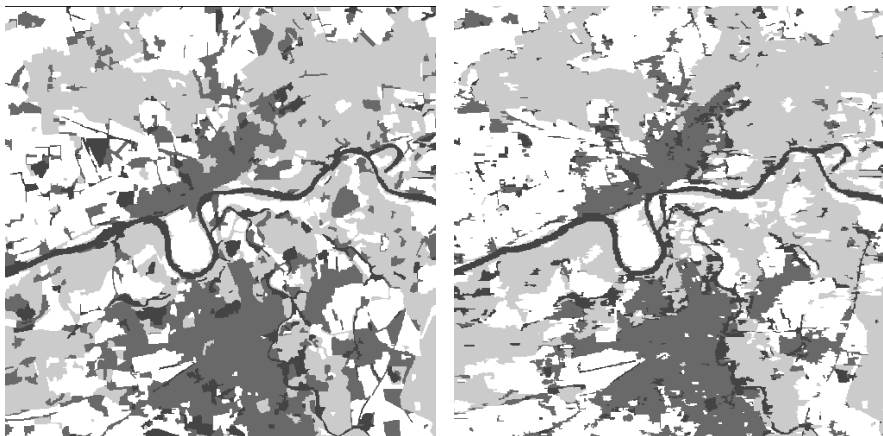
Fig. 4. Left: eCognition classification result with Nearest Neighbor. Right: MSEG classification result with SVM. In both classification results, errors have been in-troduced to the training sets for generalization evaluation. (Legend as in Fig3)

Table 1. Nearest Neighbor confusion matrix. The overall accuracy was 85.6%

	Woodland	Grassland	Impervious	Water
Woodland	17922	3381	280	0
Grassland	2578	12854	195	0
Impervious	139	770	8539	0
Water	80	0	0	4740
				85.6%

Table 2. SVM confusion matrix. The overall accuracy was 90.6%

	Woodland	Grassland	Impervious	Water
Woodland	17846	2088	45	740
Grassland	767	15937	210	91
Impervious	231	215	8305	263
Water	180	13	10	4537
				90.6%

Table 3. Nearest Neighbor confusion matrix. The overall accuracy was 84.1%

	Woodland	Grassland	Impervious	Water
Woodland	16080	1470	0	0
Grassland	2195	13891	195	0
Impervious	899	314	8605	0
Water	1545	1330	214	4740
				84.1%

Table 4. SVM confusion matrix. The overall accuracy was 86.0%

	Woodland	Grassland	Impervious	Water
Woodland	16816	3458	207	238
Grassland	1262	15506	178	59
Impervious	249	325	8315	125
Water	349	755	1	3635
				86.0%

The eCognition software provided primitive objects through the segmentation algorithm with scale parameter 15. Then the sample data were imported to the NN classifier and the classification took place. The minimum overlap between samples and primitive objects was set to 50%. The classification result is presented in Figure 6. After the classification step, a confusion matrix was computed (Table 5) and the overall accuracy was 87.4%.
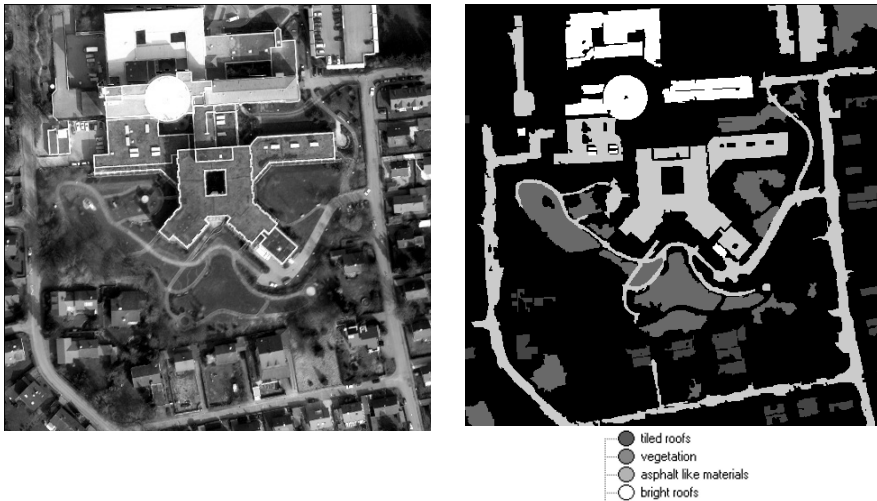
Fig. 5. Left: the original aerial scanner image (source: Toposys GmbH). Right: The ground-truth dataset used to evaluate results
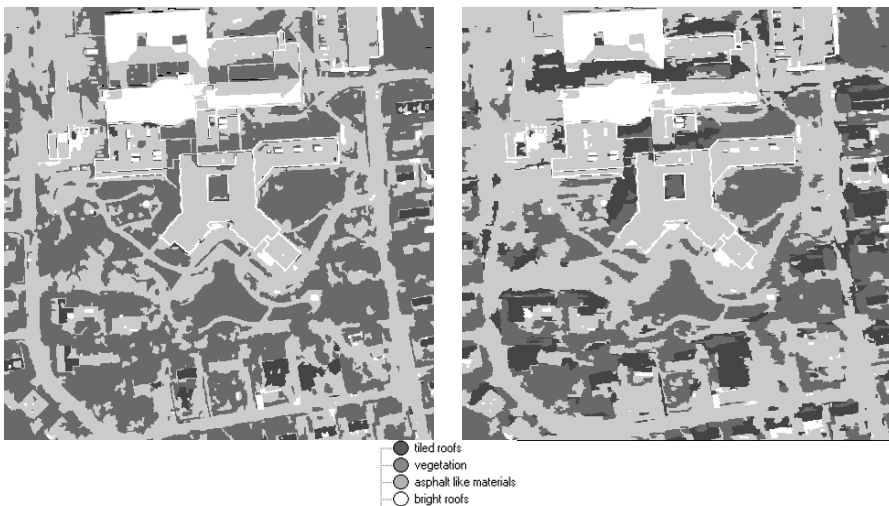
Fig. 6. Left: eCognition classification result with Nearest Neighbor. Right: MSEG classification result with SVM. Training sample overlap with objects set to 50%

The MSEG algorithm provided primitive objects for testing the object-based SVM classifier. The scale parameter was set to 300. The sample data were given as input to the SVM classifier and a cross validation procedure was followed to provide the best C and γ parameters for the SVM classi-

fier. After the classification step, a confusion matrix was computed (Table 6) The overall accuracy of the object-based SVM classification was 87.6%.

Although Table 5 and 6 overall accuracies seem to imply similar classification results, Figure 6 shows that there is not great similarity. A closer look at the confusion matrices reveals that both classifiers include classification errors, but in different class combinations. This behaviour is caused by the fact that object-based NN is treating internally each sample as a separate class to calculate distances, while SVM uses only the objects that form the support vectors for learning.

Table 5. Nearest Neighbor confusion matrix. The overall accuracy was 87.4%

	Veg.	Tile R.	Bright R.	Asphalt
Vegetation	15708	1528	0	887
Tile Roofs	0	2672	0	238
Bright Roofs	0	12	8408	387
Asphalt Like	2078	1706	2019	34847
				87.4%

Table 6. SVM confusion matrix. The overall accuracy was 87.6%

	Veg.	Tile R.	Bright R.	Asphalt
Vegetation	14760	39	0	2987
Tile Roofs	36	4229	480	1173
Bright Roofs	0	18	8377	2032
Asphalt Like	45	1493	421	34400
				87.6%

4 Conclusions

Overall, the SVM classification methodology was found very promising for Object-Based Image Analysis. It has been shown that it can produce better results than the Nearest Neighbor for supervised classification.

The computational efficiency of SVM was great, with only a few seconds of runtime necessary for training. This was theoretically expected but also, the implementation in C++ is extremely fast. This performance result occurred on test images up to 1000x1000 pixels size. However, very large remote sensing datasets were not tested.

A very good feature of the SVM is that only a small sample set is needed to provide very good results, because only the support vectors are of importance during training.

Future work will include comparison of many SVM kernels for object oriented image classification. Also, an integration of SVM classifiers with rule-based classifiers will be implemented for context-based classification.

Acknowledgements

References

Argialas D. and Goudoula V. (2003). Knowledge-based Land Use Classification from IKONOS Imagery for Arkadi, Crete, Greece. Remote Sensing for Environmental Monitoring, GIS Applications, and Geology II, Proceedings of SPIE Vol. 4886.

Argialas D., and Harlow C. (1990). Computational Image Interpretation Models: An Overview and a Perspective. *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No 6, June, pp. 871-886.

Baatz M. and Schäpe A. (2000). Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. *In: Strobl, J. et al. (eds.): Angewandte Geographische Infor-mationsverarbeitung XII. Wichmann, Heidelberg, pp. 12-23.*

Benz U., Hoffman P., Willhauck G., Lingenfelder I., Heynen M. (2004). Multiresolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing 58* pp. 239-258.

Biederman I. (1985). Human image understanding: Recent research and a theory. Computer Vision, Graphics, and Image Processing, 32, 29–73.

Brown M., Lewis H.G., Gunn S.R. (2000). Linear Spectral Mixture Models and Support Vector Machines for Remote Sensing. IEEE Transactions On Geoscience And Remote Sensing, Vol. 38, No. 5, September 2000

Chang, C.-C. and Lin C.-J. (2001). LIBSVM: a library for support vector machines. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm. (accessed 15/06/2006)

Cortes C. and Vapnik V. (1995). Support-vector network. Machine Learning, 20:273–297, 1995.

De Moraes R.M. (2004). An Analysis Of The Fuzzy Expert Systems Architecture For Multispectral Image Classification Using Mathematical Morphology Operators (Invited Paper) International Journal of Computational Cognition (http://www.YangSky.com/yangijcc.htm) Volume 2, Number 2, Pages 35–69, June 2004.

eCognition User Guide, (2005). Definiens, Munchen. http://www.definiens.com (accessed 15/06/2006)

Fang H. and Liang S. (2003). Retrieving Leaf Area Index With a Neural Network Method: Simulation and Validation IEEE Transactions On Geoscience And Remote Sensing, Vol. 41, No. 9, September 2003.

Foody G.M. and Mathur A. (2004). A Relative Evaluation of Multiclass Image Classification by Support Vector Machines. IEEE Transactions On Geoscience And Remote Sensing, Vol. 42, No. 6, June 2004.

Hsu C.-W. and Lin C.-J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Networks*, vol. 13, pp. 415–425, Mar. 2002.

Huang C., Davis L. S., Townshend J. R. G. (2002). An assessement of support vector machines for land cover classifiocation," *Int. J. Remote sensing*, vol. 23, no. 4, pp. 725–749, 2002.

Kanellopoulos I., Wilkinson G., Moons T. (1999). *Machine Vision and Advanced Image Processing in Remote Sensing*. Springer Verlag.

Lillesand, T.M., Kiefer, R.W. (1987). Remote-Sensing and Image Interpretation. Wiley, New York.

Melgani F. and Bruzzone L. (2004). Classification of Hyperspectral Remote Sensing Images With Support Vector Machines. IEEE Transactions On Geoscience And Remote Sensing, Vol. 42, No. 8, August 2004.

Mercier G. and Lennon M. (2003). Support vector machines for hyperspectral image classification with spectral-based kernels. in *Proc. IGARSS*, Toulouse, France, July 21–25, 2003.

Moller-Jensen L. (1997). Classification of Ubrban Land Cover Based on Expert Systems, Object Models and Texture. Comput. Environ and Urban Systems, Vol.21, No. 3/4, pp. 291-302, 1997.

Negnevitsky M. (2005). Artificial Intelligence, a Guide to Intelligent Systems. Pearson Education, p.440, 2005.

Sonka, M., Hlavac, V. Boyle, R. (1998). *Image Processing, Analysis, and Machine Vision* - 2nd Edition, PWS, Pacific Grove, CA, 800 p., ISBN 0-534-95393-X.

Theodoridis S. and Koutroumbas K. (2003). Pattern Recognition. Second Edition. Elsevier Academic Press, 2003.

Tzotsos A. and Argialas D. (2006). MSEG: A Generic Region-Based Multi-Scale Image Segmentation Algorithm For Remote Sensing Imagery. Proceedings of ASPRS 2006 Annual Conference, Reno, Nevada; May 1-5, 2006.

Tzotsos A., Iosifidis C., Argialas D. (2007). A Hybrid Texture-Based and Region-Based Multi-Scale Image Segmentation Algorithm. In: Object-Based Image Analysis – Spatial Concepts For Knowledge-Driven Remote Sensing Applications. Springer 2007.

Vapnik V.N. (1998). Statistical Learning Theory. John-Wiley and Sons, Inc.

Vapnik, V.N. (1995). The Nature of Statistical Learning Theory. New York, NY: Springer-Verlag.

Yooa H.W., Park H.S., Jang D.S. (2005). Expert system for color image retrieval. Expert Systems with Applications 28 (2005) 347–357

Chapter 7.5

Genetic adaptation of segmentation parameters

G. A. O. P. Costa[1], R. Q. Feitosa[1, 3], T. B. Cazes[1], B. Feijó[2]

[1] Department of Electrical Engineering, Catholic University of Rio de Janeiro (PUC-Rio), Brasil, (gilson, raul, tcazes)@ele.puc-rio.br

[2] Department of Informatics, Catholic University of Rio de Janeiro (PUC-Rio), Brasil, bruno@inf.puc-rio.br

[3] Department of Computer Engineering, Rio de Janeiro State University (UERJ), Brasil

KEYWORDS: Object-based classification, optimization, genetic algorithm

ABSTRACT: This work presents a method for the automatic adaptation of segmentation parameters based on Genetic Algorithms. An intuitive and computationally simple fitness function, which expresses the similarity between the segmentation result and a reference provided by the user, is proposed. The method searches the solution space for a set of parameter values that minimizes the fitness function. A prototype including an implementation of a widely used segmentation algorithm was developed to assess the performance of the method. A set of experiments with medium and high spatial resolution remote sensing image data was carried out and the method was able to come close to the ideal solutions.

1 Introduction

The key step in object-based image interpretation is the segmentation of the image (Blaschke and Strobl 2001). In fact, the performance of the whole interpretation strongly depends on the segmentation quality. Therefore, proper segmentation parameters must be chosen before starting the classification process. The relation between the parameter values and the corresponding segmentation outcome, however, is in most cases far from being obvious, and the definition of suitable parameter values is usually done through a troublesome and time consuming trial and error process.

Many semiautomatic approaches have been proposed to reduce the burden of parameter adaptation, starting from simple graphic support tools (e.g. Schneider et al. 1997), going through interactive systems (e.g. Matsuyama 1993) in which the user is asked to rate the result after each adaptation iteration (Crevier and Lepage 1997), up to nearly automatic solutions that require a minimum of human intervention.

The automatic adaptation of segmentation parameters involves two main issues: the selection of an objective function that expresses adequately the quality of the segmentation (Bhanu et al. 1995); and the choice of the optimization method for the search of parameter values that maximize the objective function. In supervised methods the quality measure reflects the similarity among the segmentation output and reference segments usually produced manually by a photo-interpreter (Zhang 1996). Unsupervised methods, on the contrary, use no references and do not consider human induced subjectivity or application particularities (Espindola et al. 2006).

Generally the relationship among the segmentation parameter values and the quality measure can not be formulated analytically. In such cases calculus based optimization methods cannot be used. Genetic Algorithms (GAs) do not require any explicit model of the underlying process (Davis 1990) and can work with virtually any objective function (Bhanu et al. 1991; Bhanu and Lee 1994; Everingham et al. 2002).

The present work addresses these topics and proposes an automatic GA-based adaptation method. The method uses an intuitive and computationally uncomplicated fitness function that expresses the agreement between a set of user defined segment samples and the automatic segmentation outcome.

A software prototype of the automatic adaptation method was built for performance assessment. Although the method can be easily applied to a variety of segmentation algorithms the experiments described here were limited to the algorithm proposed by Baatz and Schäpe (2000) and used in the eCognition software package (Definiens 2004).

2 Genetic Algorithms

2.1 Basic concepts and terminology

Genetic Algorithms are a computational search technique to find approximate solutions to optimization problems. They are based in the biological evolution of species as presented by Charles Darwin (Darwin 1859). The main principle of the Darwin's Theory of Evolution is that individual characteristics are transmitted from parents to children over generations, and individuals more adapted to the environment have greater chances to survive and pass on particular characteristics to their offspring.

In evolutionary computing terms an *individual* represents a potential solution for a given problem, and its relevant characteristics with respect to the problem are called *genes*.

A *population* is a set of individuals in a particular *generation*, and individuals in a population are graded as to their capacity to solve the problem. That capacity is determined by a *fitness function*, which indicates numerically how good an individual is as a solution to the problem (Michalewicz 1994).
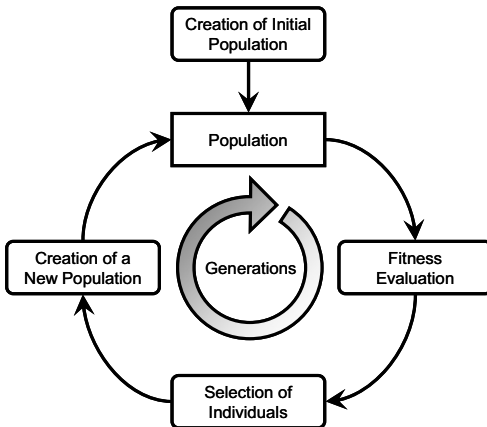
Fig. 1. Basic GA processes

GAs propose an evolutionary process to search for solutions that maximize or minimize a fitness function. This search is performed iteratively over generations of individuals. For each generation the less fitted individuals are discarded and new individuals are generated by the reproduction of the fittest. The creation of the new individuals is done by the use of *genetic operators*.

2.2 Genetic operators

A genetic operator represents a rule for the generation of new individuals. The classical genetic operators are *crossover* and *mutation*. Mutation change gene values in a random fashion, respecting the genes' search spaces. Mutation is important to introduce a random component in the search of a solution in order to avoid convergence to local minima.

Crossover operators act by mixing genes between two individuals to create new ones that inherit characteristics of the original individuals. The general idea is that an individual's fitness is a function of its characteristics, and the exchange of good genes may produce better fitted individuals depending on the genes inherited from their parents. Although less fitted individuals can also be generated by this process, they will have a lower chance of being selected for reproduction.

Other genetic operators can be found in the literature (Michalewicz 1994). Most of them are variants of crossover and mutation, adapted for specific types of problems.

3 Adaptation of segmentation parameters using a genetic algorithm

3.1 Processing scheme

In the devised GA each individual consists of a set of segmentation parameter values; each parameter is represented by a gene. The fitness of each solution (individual) is calculated by comparing the segmentation produced by the solution with a reference segmentation (Fig. 2).

The parameter values (genes) of the initial set of solutions (initial population) are generated randomly. As the evolutionary process advances, the best solutions (fittest individuals) are selected and new solutions (generations) are created from them (reproduction).

The selection of individuals for reproduction takes the fitness values into consideration, so that the fittest individuals have a larger probability of being selected. Furthermore, the best individuals from one generation are kept in the next generation. The evolutionary process stops after a fixed number of generations, and the gene values of the fittest individual are taken as the final (adapted) segmentation parameter values.

For computational efficiency, segmentation may be restricted to a small window around each target segment. This considerably reduces the processing time in comparison to segmenting the whole image at each fitness evaluation.
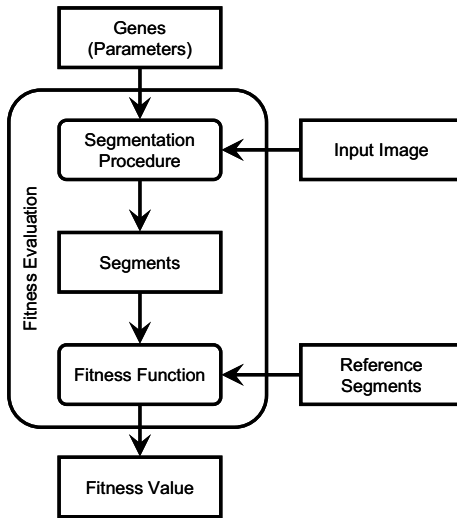
Fig. 2. Fitness evaluation

3.2 Reproduction procedure

As stated before, the initial population is created by setting random values for the genes of each individual. After fitness evaluation a new population is created by replacing the *M* worst individuals of the prior population, being *M* a positive integer value smaller than the population size.

The new individuals are created by genetic operations over selected individuals of the prior population. The selection of individuals is done by a roulette mechanism, which takes into consideration normalized fitness values (Davis 1990).

The following genetic operators were used (Davis 1990; Michalewicz 1994). *One point crossover*: two individuals exchange genes; *arithmetic crossover*: a linear combination of a set of genes of two individuals is performed; *mutation*: the value of a gene is substituted by a random value; two types of *creep mutation*: gene values are adjusted (added or subtracted) by a small or large randomly generated value.

The selection of the reproduction operation is also done by a roulette mechanism, considering a predefined probability value for each operator. To help preventing convergence to local minima, the operators' application probabilities are interpolated during the evolution process (Davis 1990), decreasing crossover probability while increasing mutation and creep probabilities.

3.3 Fitness evaluation

The fitness of an individual should indicate how good the segmentation result in relation to the reference segmentation is. In mathematical terms, given a set of reference segments S and a parameter vector P, a fitness function $F(S, P)$ that appropriately expresses the goodness of a segmentation outcome must be defined. Once the fitness function F is chosen, the task of the GA consists in searching for the parameter vector P_{opt}, for which the value of F is minimum:

$$P_{opt} = arg_P(min[F(S,P)])$$ (1)

The fitness function devised in this work is defined as follows. Let S_i denote the set of pixels belonging to the ith segment of the set S, and $O_i(P)$ the set of pixels belonging to the segment with the largest intersection with S_i among the segments produced by using P as parameter values of the segmentation algorithm. The fitness function is then given by the equation below, in which '-' represents the set difference operator; '#( )' is the cardinality function; and n is the number of segments in the set S.

$$F(S,P) = \frac{1}{n}\sum_{i=1}^{n}\frac{\#(S_i - O(P)_i) + \#(O(P)_i - S_i)}{\#(S_i)}$$ (2)

Figure 3 shows graphically the components of the proposed fitness function. The region enclosed by the solid contour represents a reference segment S_i and its area corresponds to the denominator in Eq. 2. The region with the dashed contour represents $O_i(P)$. The shadowed area corresponds to the numerator in Eq. 2. Notice that $F=0$ indicates a perfect match between the reference and the output segmentation.

It is important to point out that S does not need to represent a complete segmentation of the input image, such that every pixel of the image would belong to a segment in S. In fact, in the experiments presented here, S contains only 5 to 10 segments.
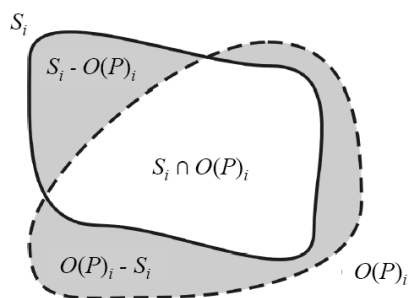
Fig. 3. Graphical representation of the fitness function

4 Segmentation procedure

The segmentation procedure used in this work is based on the region growing algorithm proposed in (Baatz and Schäpe 2000). The algorithm is a stepwise local optimization procedure that minimizes the average heterogeneity of the image segments.

Segments grow from single pixels, merging to neighbouring segments. In each processing step a segment can be merged to the neighbour that provides for the smallest growth of global heterogeneity. The merging decision is based on minimizing the resulting segment's *weighted heterogeneity*, an arbitrary measure of heterogeneity weighted by segment size.

The heterogeneity measure has a spectral and a spatial component. Spectral heterogeneity is defined over the spectral values of the pixels belonging to the segment, it is proportional to the standard deviation of the pixels' spectral values weighted by arbitrary spectral band weights.

The spatial heterogeneity component is based on the deviation of the segment's shape from a compact and a smooth shape. Compactness is defined as the ratio of the perimeter of the segment and the square root of its area. Smoothness is defined as the ratio of the segment's perimeter and the length of its bounding box (parallel to the image borders).

To simulate the parallel growth of the segments, they are selected for merging only once in each iteration, in an evenly distributed fashion.

The merging decision mechanism is of key importance to this work, as it is where the external parameters of the segmentation procedure are employed. A *fusion factor* is calculated for each neighbour of the selected segment. The neighbour for which this factor is minimum will be merged, but only if the fusion factor is smaller than a certain threshold defined as

the square of the so called *scale parameter*. The procedure stops when no more segments can be merged.

As shown in Eq. 3 the fusion factor *f* contains a spectral heterogeneity component h_{color} and a spatial heterogeneity component h_{shape}. The relative importance of each type of heterogeneity is set by the color weight w_{color}.

$$f = w_{color} \cdot h_{color} + (1 - w_{color}) \cdot h_{shape} \tag{3}$$

The spatial component of the fusion factor has again two components (Eq. 4), a compactness component h_{cmpct} and a smoothness component h_{smooth}. The relative importance of each component is set by the weight w_{cmpct}.

$$h_{shape} = w_{cmpct} \cdot h_{cmpct} + (1 - w_{cmpct}) \cdot h_{smooth} \tag{4}$$

Throughout the segmentation procedure segments grow based on an adjustable criterion for heterogeneity. This adjustment can be made by setting the values of the scale parameter (s), the spectral band weights (w_c), the color weight (w_{color}) and the compactness weight (w_{cmpct}).

Adjusting the scale parameter influences the overall object size: the larger its value, the bigger the resulting segments. Additionally, the influence of each spectral channel, the influence of shape against color, and of compactness against smoothness in shapes can be set.

Given a particular image's spectral and spatial characteristics and the land use/land cover characteristics of the investigated site, the values of those parameters can change considerably. And finding a good set of parameter values for each case is by no means a trivial task.

5 Experiments

In order to evaluate the performance of the proposed method a software prototype was developed in C++. The prototype includes the implementation of the algorithm described by Baatz and Schäpe (2000).

In the proposed method a human interpreter delimits manually a set of polygons, which represents his expectation regarding the result of the automatic segmentation. These polygons are used as the reference segmentation for fitness evaluation, and should capture the subjectivity of the interpreter in the evaluation process.

The cognitive processes applied by the human interpreter is very complex and cannot be fully represented in the segmentation algorithm as they lie almost totally below our threshold of consciousness (Crevier and Lepage 1997). Therefore, there is no guaranty that a set of parameter val-

ues that produces a perfect match with the reference exists. That considered, the eventual GA incapacity to converge to a satisfactory solution can result from the limitations of the segmentation procedure or from an eccentric choice of reference segments by the interpreter.

Two sets of experiments were devised in order to investigate the issues just mentioned. In the first experiments, subjective factors were eliminated from the analysis, so that only the capacity of the method to converge to optimal solutions could be tested. In the second set of experiments the capacity of the method to capture the interpreter's subjective perception about the quality of the segmentation was evaluated.

5.1 Convergence analysis

The purpose of this series of experiments was to test the capacity of the GA to converge to a solution that minimizes the fitness function. The optimal solution is known beforehand and achievable by the segmentation algorithm.

Image data of two different sources were used: pan-sharpened ETM Landsat and IKONOS scenes covering areas of the City of Rio de Janeiro, taken respectively in 2001 and 2002. From each scene two 256 by 256 pixel image subsets, over sites with different land cover characteristics, were produced. The input images for the experiments are identified in this section as follows: image 1 – subset of the Landsat scene over a dense urban area; image 2 – subset of the Landsat scene over a forest; image 3 – subset of the IKONOS scene over a sparse residential area; image 4 – subset of the IKONOS scene over a dense urban area.

For each experiment the input images were segmented using predefined values for the parameters. Sample segments were selected manually from the resulting segmentation to be used as the input reference segmentation for fitness evaluation (see Fig. 2). The basic criterion for the selection of samples was to end up with a evenly spatially distributed set of segments that did not intercept the borders of the images. For each experiment between 5 and 10 segments were selected.

In these experiments, therefore, the optimal solution is known a priory and it is expected that the GA converges to the predefined parameters, producing exactly the same reference segments

All the results reported in this work were produced by the GA operating with the following configuration: population size of 50 individuals; 40 generations; 90% of the individuals changed from one generation to the next. The GA performed well with this configuration throughout our experiments for all images of our data set. This is an evidence for the robust-

ness of the method, i.e., that the GA parameters does not need to be adapted for each application.

Table 1. Parameters and results of the first set of experiments

exp	image	s	w_{color}	w_{cmpct}	w_1	w_2	w_3	fitness
1	1	*30.0*	*0.80*	*0.50*	*0.6*	*0.3*	*0.1*	0.09
		27.7	**0.65**	**0.33**	**0.6**	**0.3**	**0.1**	
2	1	*30.0*	*0.80*	*0.50*	*0.3*	*0.1*	*0.6*	0.00
		25.4	**0.69**	**0.32**	**0.3**	**0.1**	**0.6**	
3	1	*60.0*	*0.80*	*0.50*	*0.6*	*0.3*	*0.1*	0.02
		44.2	**0.82**	**0.57**	**0.6**	**0.3**	**0.1**	
4	1	*60.0*	*0.80*	*0.50*	*0.3*	*0.1*	*0.6*	0.01
		44.2	**0.71**	**0.32**	**0.3**	**0.1**	**0.6**	
5	2	*30.0*	*0.80*	*0.50*	*0.6*	*0.3*	*0.1*	0.01
		27.8	**0.76**	**0.38**	**0.6**	**0.3**	**0.1**	
6	2	*30.0*	*0.80*	*0.50*	*0.3*	*0.1*	*0.6*	0.02
		29.0	**0.70**	**0.35**	**0.3**	**0.1**	**0.6**	
7	2	*60.0*	*0.80*	*0.50*	*0.6*	*0.3*	*0.1*	0.06
		41.6	**0.56**	**0.14**	**0.6**	**0.3**	**0.1**	
8	2	*60.0*	*0.80*	*0.50*	*0.3*	*0.1*	*0.6*	0.00
		53.7	**0.79**	**0.47**	**0.3**	**0.1**	**0.6**	
9	3	*30.0*	*0.80*	*0.50*	*0.1*	*0.3*	*0.6*	0.00
		29.7	**0.77**	**0.42**	**0.1**	**0.4**	**0.5**	
10	3	*30.0*	*0.80*	*0.50*	*0.3*	*0.6*	*0.1*	0.01
		30.2	**0.81**	**0.50**	**0.3**	**0.6**	**0.1**	
11	3	*60.0*	*0.80*	*0.50*	*0.1*	*0.3*	*0.6*	0.02
		58.8	**0.80**	**0.51**	**0.3**	**0.2**	**0.5**	
12	3	*60.0*	*0.80*	*0.50*	*0.1*	*0.3*	*0.6*	0.11
		57.6	**0.74**	**0.28**	**0.1**	**0.2**	**0.7**	
13	4	*30.0*	*0.80*	*0.50*	*0.1*	*0.3*	*0.6*	0.03
		26.2	**0.62**	**0.25**	**0.1**	**0.3**	**0.6**	
14	4	*30.0*	*0.80*	*0.50*	*0.3*	*0.6*	*0.1*	0.01
		26.8	**0.70**	**0.24**	**0.3**	**0.6**	**0.1**	
15	4	*60.0*	*0.80*	*0.50*	*0.3*	*0.6*	*0.1*	0.21
		44.5	**0.65**	**0.24**	**0.3**	**0.6**	**0.1**	
16	4	*60.0*	*0.80*	*0.50*	*0.1*	*0.3*	*0.6*	0.26
		53.4	**0.79**	**0.57**	**0.1**	**0.3**	**0.6**	

The experiment results are stated in Table 1. For each experiment it shows two lines with parameter values. On the top line, in italic, the parameter values used to create the reference segmentation are shown. They represent the optimal solution. The values of the fittest individuals found by the GA in five runs of each experiment are shown in bold style on the bottom line. The column *fitness* shows the fitness value of that individual.

The fitness value achieved for the Landsat images (images 1 and 2) were all very close to zero, the ideal value. For the IKONOS images (images 3 and 4), slightly worse results were obtained in the experiments 15 and 16. Those results may be explained by the greater complexity of the shapes of the sample segments used in those experiments, related to the particular choice of band weights.

It is interesting to notice the slight deviations from the parameter values used in the reference segmentation. The largest deviation for the compactness weight can be explained by the high color weight value (0.8) chosen for the reference segmentation. As the relative importance of shape characteristics in the region growing process is small, variations of the compactness weight have little influence on the output of the segmentation.

In experiments 2, 8 and 9, the segments generated with the adapted parameter values are identical to the references. However, the parameters values found by the GA were not exactly the same as the ones used to produce the reference segments. This indicates that the segmentation algorithm used here can produce similar results from (moderately) different sets of parameter values. Other experimental results, not showed in Table 1, have demonstrated that the GA is able in these cases to find more than one optimal solution (with fitness value equal to zero).

5.2 Manually defined references

The second series of experiments had two objectives. The first was to investigate the capacity of the method to find solutions from which meaningful objects can be produced. Unlike the previous set of experiments, the polygons that represent the reference segments were drawn manually by a photo-interpreter. In this case there is, in principle, no ideal solution (that can generate segments identical to the reference).

The first objective of this set of experiments was, in other words, to verify if the achieved solution, although not optimal, is good enough considering the subjective interpreter's point of view. The second objective was to investigate if the desired segmentation output, for all objects of interest in the image, can be achieved by using a small set of segment samples.

The images used in this set of experiments (Figure 4) were produced from different sensors, over areas with different land covers. Image 1 was extracted from an aerial photograph taken over a residential area in the City of Rio de Janeiro. The other two images were obtained from public resources on the Internet. Image 2 shows the parking area of a bus company, also situated in Rio de Janeiro, and image 3 shows storage tanks of an industrial plant in the City of Duque de Caxias, in Rio de Janeiro

State. The three images have 400 by 400 pixels, RGB format with 24 bits (8 per band).
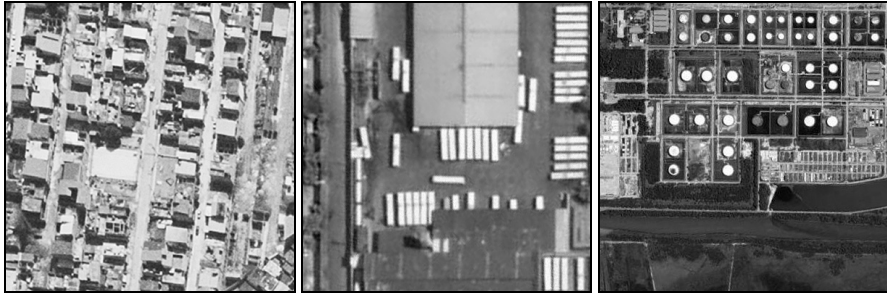
Fig. 4. Input images: image 1 (left); image 2 (center); image 3 (right)

The experiments were organized as follows. Over each of the three images the photo-interpreter has drawn segments delimiting different image objects: roof tops of ceramic material; buses; and storage tanks. For each image the delimited segments were organized in three groups (A, B and C), each group containing approximately the same number of segments. The segments were assigned to the groups randomly.

Segments representing only one class of objects (roofs, busses, tanks) were considered in each experiment. One of the segment groups was selected to serve as the reference for the parameter evolution. The selected group was regarded as the training set. The solution obtained by using the training set as reference segments was then applied to the whole image, and the fitness evaluation was performed using as reference the segments of the other two groups, the validation set. Three experiments were performed for each image, using different groups of segments for the training and validation sets.

Table 2 shows the best results obtained for the three executions of the GA for each input image. The column *image* indicates the image used in the experiment. The column *training group* identifies the group used as the training set. These groups are represented with different gray levels in figure 5 (group A: black; group B: dark gray; group C: light gray). The columns w_{color}, w_{cmpct}, w_1, w_2 and w_3 show the values of the parameters found by the GA. The last two columns contain the fitness values calculated for the fittest individual using the training set and the validation set as references.

Table 2. Results of the second set of experiments

exp	image	training group	s	W_{color}	W_{cmpct}	w_1	w_2	w_3	training (fitness)	validation (fitness)
1	1	A	13.8	0.15	0.70	0.4	0.2	0.4	0.21	0.45
2	1	B	16.6	0.48	0.57	0.5	0.4	0.1	0.28	0.33
3	1	C	14.6	0.24	0.64	0.5	0.1	0.4	0.19	0.33
4	2	A	64.4	0.41	0.59	0.1	0.2	0.7	0.48	0.68
5	2	B	66.3	0.51	0.50	0.1	0.7	0.2	0.51	0.56
6	2	C	75.3	0.44	0.66	0.4	0.6	0.0	0.45	0.76
7	3	A	58.1	0.63	0.66	0.0	0.1	0.9	0.38	0.65
8	3	B	36.6	0.56	0.63	0.5	0.1	0.4	0.26	0.49
9	3	C	41.1	0.37	0.63	0.1	0.7	0.2	0.39	0.83

The analysis of the fitness values indicates that the GA reached consistent solutions in all the experiments. As expected, the fitness values are higher than those obtained in the previous set of experiments since, as discussed earlier, it is not likely that the segmentation procedure can generate segments identical to the ones drawn manually.

Figures 5 to 7 permit a visual inspection of the method's performance. Figure 5 shows the segments drawn by the interpreter. Figures 6 and 7 show the segments generated from the solutions found by the GA. These images were produced from the complete segmentation of the image, but contain only the segments with largest intersection with the ones defined by the interpreter. Figure 6 shows the worst results (respectively experiments 2, 5 and 9), while figure 7 shows the best results (respectively experiments 3, 4 and 8).

A large similarity can be perceived among the manually drawn segments and those produced by the segmentation procedure with the parameter values found by the GA. Even for the first image, in which the reference segments are more heterogeneous, the results are visually consistent.

It should be noted that the results obtained were achieved with 40 generations of 50 individuals. A better output could be obtained if the GA worked for more generation cycles or with a larger population size. There is, however, no guaranty that this investment in processing time would bring a correspondent improvement in the final result. This uncertainty is typical of stochastic optimization methods such as GA.
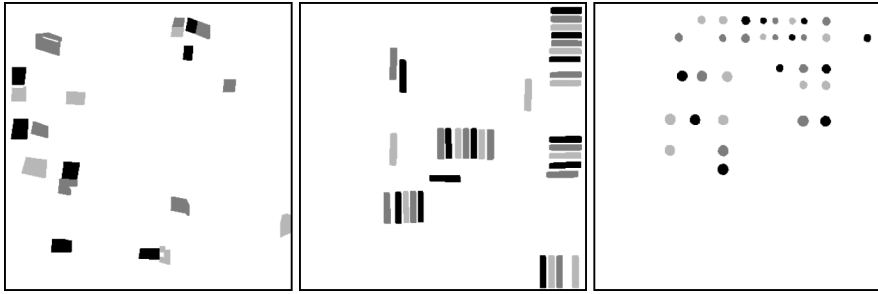
Fig. 5. Segments drawn over: image 1 (left); image 2 (center); image 3 (right)
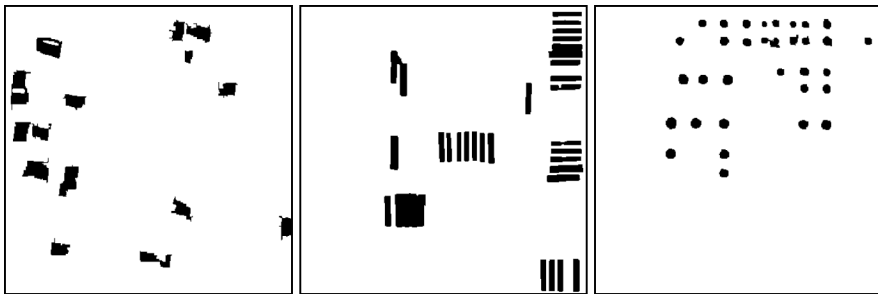
Fig. 6. Worst results for: image 1 (left); image 2 (center); image 3 (right)
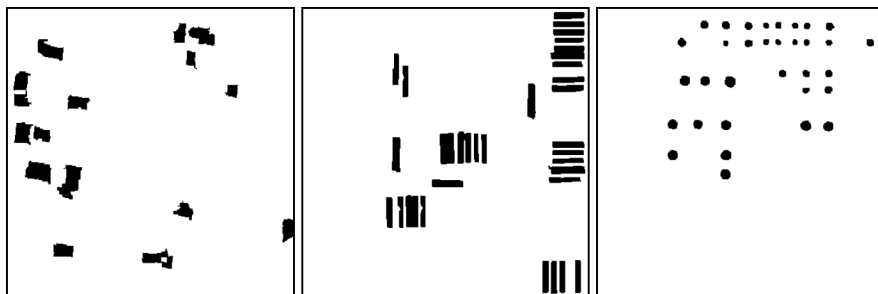
Fig. 7. Best results for: image 1 (left); image 2 (center); image 3 (right)

It is important to mention that the performance of the GA was compared to random search. The same number of solutions (individuals) generated in the evolution process was created using random values for the segmentation parameters. The fitness values of the best individuals created during

the random search were always higher (worse in our case) than the fitness values of the solutions found by the GA.

6 Conclusions and future work

We have presented a method for the automatic adaptation of segmentation parameter values based on Genetic Algorithms. In the proposed method segmentation parameters are coded into genes of the individuals of a GA, and fitness evaluation is a measure of the similarity between a user defined set of segment and the segmentation result. Two groups of experiments were implemented to evaluate the performance of the method.

In the experiments where the optimal solution was known a priori, the GA converged to optimal or near optimal solutions. In some cases the GA was able to find more than one solution that delivers segments identical or nearly identical to the reference. The performance was similar for images of different sensors, over areas with different land covers, what indicates the robustness of the method. It should also be noted that the parameters of the GA were the same in all experiments, and that the solutions produced were always superior to those found by random search.

The results of the second set of experiments indicate that the method is able to express numerically, therefore in an objective way, the subjective perception of the human interpreter about the segmentation quality. Through a small set of manually drawn segments the interpreter could indicate the expected output of the segmentation procedure, and in all experiments the adaptation method was able to produce visually consistent solutions for different types of image objects.

Alternatives to increase computational performance of the developed prototype are being considered. Currently, with a standard Pentium 4 processor, each experiment was executed in approximately 3.5 hours. Parallel processing of the individuals of a generation is a possibility. Improvements in the architecture of the optimization procedure can also help, as the utilization of concepts such as cultural algorithms (Becerra and Coello 2005) and co-evolution (Davis 1990). In any case it is certain that the progress of the hardware and software technology will also contribute for the reduction of processing time.

The software prototype implemented for this work is available upon request to the corresponding author.

Acknowledgments

References

Baatz M, Schäpe A (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Blaschke T, Griesebner G (eds) Angewandte Geographische Informationsverarbeitung XII. Wichmann-Verlag, Heidelberg, pp 12-23

Blaschke T, Strobl J (2001) What is wrong with pixels? Some recent developments interfacing remote sensing and GIS. GIS-Zeitschrift für Geoinformationssysteme 6:12-17

Becerra RL, Coello CAC (2005) Use of domain information to improve the performance of an evolutionary algorithm. In: Proceedings of the 2005 workshops on Genetic and Evolutionary Computation, pp 362-365

Bhanu B, Ming JC, Lee S (1991) Closed-loop adaptive image segmentation. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '91), pp 734-735

Bhanu B, Lee S (1994) Genetic learning for adaptive image segmentation. Kluwer Academic Publishers, Boston

Bhanu B, Lee S, Das S (1995) Adaptive image segmentation using genetic and hybrid search methods. IEEE Trans Aero Electron Syst 31:1268-1291

Crevier D, Lepage R (1997) Knowledge-based image understanding system: A survey. Comput Vis Image Understand 67:161-185

Davis L (1990) Handbook of Genetic Algorithms. Van Nostrand Reinhold Company, New York

Darwin C (1859) On the origin of species by means of natural selection, or the preservation of favored races in the struggle for life. John Murray, London

Definiens (2004) eCognition Professional: User Guide 4, Definiens Imaging, Munich

Espindola GM, Câmara G, Reis IA, Bins LS, Monteiro AM (2006) Parameter selection for region growing image segmentation algorithms using spatial autocorrelation. Int J Rem Sens 27:3035-3040

Everingham MR, Muller H, Thomas BT (2002) Evaluating image segmentation algorithms using the pareto front. In: Proceedings of the 7th European Conference on Computer Vision (ECCV2002), part IV (LNCS 2353), pp 34-48

Matsuyama T (1993) Expert system for image processing, analysis, and recognition: declarative knowledge representation for computer vision. Adv Electron Electron Phys 86:81-171

Michalewicz Z (1994) Genetic Algorithms + Data Structures = Evolution Programs. Springer-Verlag, Berlin Heidelberg New York

Schneider W, Eckstein W, Steger C (1997) Real-time visualization of interactive parameter changes in image processing systems. In: Grinstein GG, Erbacher RF (eds) Proceedings of the SPIE: Visual Data Exploration and Analysis IV, vol 3017. SPIE, San Jose, pp 286-295

Zhang YJ (1996) A survey on evaluation methods for image segmentation. Pattern Recogn, vol 29, 8:1335-1346

Chapter 7.6

Principles of full autonomy in image interpretation. The basic architectural design for a sequential process with image objects

R. de Kok[1], P. Wezyk[2]

[1]www.ProGea.pl

[2] Lab. of GIS & RS, AUC, Poland

KEYWORDS: Autonomic classification, process design, transferability.

ABSTRACT: Full autonomic image interpretation methods require a detailed description on the predictable behaviour of image objects. This standard behaviour allows to make a shift from scene depending properties towards object depending attributes. Those standard features must be put under scrutiny by peer experts. In this paper a sequential architectural process design for a full autonomic image analysis is presented in it's initial stage. Also the introduction of predictable behaviour of edge objects is clearified. Not the existence of edge objects, but their predictable behaviour is presented as a new contribution to OBIA. A test case is shown and can be tried by expert user. The aim is to reach maturity for standardized architectural process-design. Due to their predictable behaviour, edge objects can function as a start for full automatic image interpretation. Image objects do not need to be spectral homogeneous to be predictable. Also a spectral mean of a class is not an absolute necessary attribute for class predictability. Both conditions: spectral homogenity and spectral mean of a class in feature space where core assumptions in traditional remote sensing but lose their importance in OBIA.

Introduction

Typical in OBIA is the large variety of spectral and spatial attributes for image objects after segmentation. This makes it possible to start image analysis without defining a spectral mean and standard deviation of a class. This implies that image sampling by visual interpretation becomes obsolete. Instead, histogram extremes of a selection of classes now come into focus. The predictability of certain classes in the extremes of their histograms allows for automatic establishment of a sample population. Those samples from the histogramm extremes form the basis for spectral calibration of the rest of the representative object-class population. This is not completely new as in standard remote sensing, for example vegetation coverage, could always be expected in the upper part of an NDVI band. Now we take such principles to their extreme.

The initial understanding of the imagery started from local pixel populations or object primitives that should show normal distribution in feature space and are spectral homogeneous. This was the underlying concept in the pioneer work on ECHO of Kettig & Landgrebe (1976). However, this still does not show the complete picture. An important source of object-primitives with spectral variability in OBIA analysis are edge objects that do not behave as homogeneous spectral image objects and should be treated separately. The explicit registration and classification of edge objects offers new opportunities. Due to a sequential process architecture, as offered in a new approach in definiens developer, a tailor made sequence of processes can be constructed for the classes within the legend. The sequences can be discussed with fellow experts and in this way maturity can be reached for a standardized process flow for standardized classes. Extending the knowledge on class behavior is a necessity. Not only in the spectral domain of the feature space but moreover the increment of knowledge on behavior in spatial relationships, which are often unique, stable and transferable. This makes full autonomy in the process of image analysis possible.

Homogeneous versus non-homogeneous objects

Pioneer work on automatic image classification was focused on the apparent spectral similarity of neighbouring pixels in optical remote sensing satellite data (after Kettig & Landgrebe, 1976). With increasing spatial resolution, many neighbouring pixels in an earth observation satellite image

are registered over the identical class of land cover type. This is apparent in a per parcel classification (Janssen, 1994)

Here many neighbouring pixels were inside an agricultural field under homogeneous treatment. These pixels are registered above crops with almost identical vegetation development stage as well as biomass, moisture conditions and chlorophyll activity. Because parcels often exist inside a cadastral administration, the shape of a group of training pixels or object primitive inside a parcel have low priority. This is still the case for agricultural image objects. The shape of image objects primitives, which are subsets inside this agricultural parcel, does not add information to the system when its neighbours are nearly identical spectral objects. The shape does not reveal if this object is a maize or a rice field. Only the spectral properties of this object-primitive might reveal this information.

Edge objects

Many pixels from image objects cannot be appointed to a homogeneous spectral class but rather to a spatial class. Those pixels that respond to edge detection filters cannot be used to extract mean spectral values with relevance for spectral classes. Here the shape of a population of such pixels in the image domain reveals more about their object class.

Fig. 1. A Pure edge detection result from high resolution data (image courtesy CSIR, SA)

Shape characteristics are most often stable and transferable. The information content of pure edge imagery (see Fig. 1) is sufficient for a visual interpretor. This edge image reduces the normal complexity of 8 or 11 Bit data in an almost zero-to-one image (edge, non-edge). At the same time, edge imagery implicitly preserves existing details derived from original scenes. To stress the strenght of the human vision, remark that figure 1 contains sufficient data for a visual interpreter to digitize buildings, road-network and high-vegetation. Very few image interpretation software however are capable to assign a correct label to an image object such as "edge of a large building" only using figure 1 as input. The grey value of figure 1 has no meaning, it can be a 0;1 image or 0;255 or even a 0;65536 value range. This means, sampling the grey values in feature space does provide almost no information on the image content.

Although edge imagery is a classical result of standard remote sensing software, the classification of edges seems to be still an undiscovered territory. Because the human eye and the way the visual interpretation can work so easily with edge imagery alone, it would offer itself to computer analysis to reduce image interpretation first to a correct classification of an edge image and mimic exactly the human interpreter. A first test on the quality of the image interpretation software should be a standard test on the results of classifying only edges. Above all, the relative insensitivity of edge information on differences of acquisition date and sun angle as well as atmospheric conditions (except clouds) not to mention their stability and transferability over scale and sensor type make them unique image objects. Note that all black/dark areas it is non-edge-areas in figure 1 can be considered to represent homogeneous pixel populations with a Gaussian distribution. For those areas, traditional remote sensing methods do function efficiently as well as image-object analysis for so called "homogeneous pixels" and homogeneous image-objects which are the core of image analysis as developed by Landgrebe (Kettig & Landgrebe, 1976).

Sequential classification

The latest development in object-oriented classification is the construction of a workflow according to a sequence of processes. The original design of the hierarchical architecture of a class-tree now subordinates the design of a 'process-tree'. The sequential process flow changes the treatment of image objects in the scene, most notably in operations that previously used to be image-based operations. These image operations now have become part of object-based operations such as segmenting only a single object-class or

calculating a principle component for one class of objects only and not for a set of spectral bands using a whole image stack, a standard practice in hyperspectral image analysis. Hierarchical classification is now taken to the extreme allowing automatic sampling processes as well as developments towards standardizations. Hierarchical classification is not a new phenomenon in remote sensing. The traditional output of such a classification is a stratification of the thematic layers focused on (legend) classes. In the sequential process approach, such as applied in Definiens developer, the output goes beyond the stratification of thematic layers alone.

The traditional stratification contains per layer one or a few element(s) of the legend. Moreover, the various classes are initiated using comparable (statistical) procedures and the classes themselves were organized in a decision tree. When switching from the traditional hierarchy of classes to the hierarchy of processes, there is a need to develop a tailor made design for each of the individual classes. Although initializing the class itself remains important, the main focus stays with a correct transferability of the overall architecture of the process design. This is something very different than the architecture of a class tree design. A class previously used to incorporate its decision rules and the initialization of the class was not done step by step but instanteniously. Iterative processes such as isodata clustering still are considered intanteniously in the sense that a statistical process is chosen and applied to all classes while conditions are not set according to each class separatively between the iterations. The correct sequence of classification now becomes a crucial factor to allow further autonomy of the classification process. This is a new phenomenon in remote sensing with also new consequences. The consequences of using the correct process sequence makes it difficult to mix various classification methods or maintain mixing class hierarchies incorporating (fuzzy-) rule sets together with process hierarchies. The total procedure should have a logical internal flow as well as reaching maturity after discussion by fellow experts. This could imply that a compromise on class accuracy is acceptable in a trade off for an enormous gain in speed and quantity of imagery to be processed. The meaning of this can be clearified with the availability of a single process architecture used for all IKONOS and QuickBird imagery over Europe using a single classification process-protocol. Such a process-protocol might be less accurate than a visual sampling for each scene but it requires no further operator interventions for many thousand images. Meanwhile expecting processes becoming mature and reach or surpass the accuracy of traditional hierarchical classification schemes.

Representative populations

The large variety of process operations now make pure statistical methods like cluster analysis or minimum distance to mean algorithms redundant. These statistical methods function very well for spectral classes with a predictable mean value in feature space but not for all object classes within a VHR satellite image frame. In statistically dominated analysis, the initiation of the thematic layer is achieved using a mean and standard deviation of the samples. Traditionally, the operator had to select the correct pixels or objects in the image domain. In the new approach, the operator must now explicitly formulate expert knowledge into the expert process without the need for sampling in the image domain. The knowledge on the object class (predictable) behavior in feature space has to be considerable. The operator has to differentiate classes with an easy statistical description (water) from classes, which are easy to discriminate using spatial attributes (buildings). The end-user might continue 'Click & Classify', but to initiate the class of interest, the 'Click' operation should not be focused on a class inside the image domain returning co-ordinates, but instead, the end-user should simply click inside the legend.

Peer review will be the referee on the discussion over the conditions, which are deciding factors being crucial for representative populations. Because much of these standard features are encountered from empirical data and few can be predicted from theory on image understanding alone. The latter procedure is of course preferable as empirical definitions are merely a rule of thumb and should be backed up by sound theoretical knowledge on the cause of predictable behavior of image objects. A typical example is the low standard deviation of the Red band over forest areas (in SPOT/Landsat data type). The shadow component as well as crown activity both reduces Red band albedo values over forest. However, aging forest stands seems to be reducing in the variability of the Red band more than young forests. A theoretical explanation on the bio-physical properties of aging forest stands and the theoretical reason for the relationship between forest age and low standard deviation values in the Red band are more valueable than the simple statement, 'use low standard deviations in red' to classify a forest mask.

Template matching

The strategy of predefining features in sequential processing is not uncommon and resembles in its assumptions the traditional template match-

ing technique. In template matching, no sampling is required. The traditional template functions within the image domain. The process-template however does not function inside the image domain but in the feature space domain. The decision on the population that is regarded representative for a class thus no longer lies solely with the visual interpreter of the image. The selection of a sample population is initiated from a number of rules that describe the predictable behaviour of the sample set which is to be regarded representative. Therefore the sample selection is not left to chance by the operator to encounter a set of samples in the image which 'seems' to be a good start but to follow explicit rules about the must have conditions for samples to be ensigned candidates for the representative population. As in traditional template matching, there is a strict expectation towards the object population within its domain. The success lies in the correct treatment of histogram extremes for the crucial classes. This is possible because certain crucial classes behave predictable within the histogram extremes of a selected group of features. The template in traditional techniques with a fixed description in pixel/vector size and shape is virtually similar to a standardized set of process-rules in the process-tree (see Fig. 3).

Categorization

Categorization can precede class construction. The categorization is using feature based category descriptions preceding final class-names from the legend. As an example, a category name "*high textured, high NDVI value objects containing shadow-sub objects*" is used instead of the legend class name '*broadleaved forest*'. The categories behave differently towards accuracy assessment compared to the legend classes. The final classes in the legend must compromise in accuracy due to generalization issues, where mere categories simply follow process decision rules and should always be correct.

Self-adapting

The processes can be made self adapting towards the changes of the spectral values in each varying image scene of the same mosaic due to the ability to let the process measure a spectral and/or spatial value on a selected (anchor-) object population and assign the measurement to a process variable. The variable changes with each image scene because the anchor object values which are absolute histogram values change. The relative posi-

tion of the anchor object within their histogram however does not change. Thus traditional template matching works in the image domain in cases where the image domain is predictable. Hence the new process approach works in feature space because the histogram shape is predictable. By completing the list between crucial classes and those features which are predictable, robotic vision becomes apparent.

Central role of edges

Edges are a 'must have' condition. Basically people separate nature from their dominion by explicitly constructing a physical edge and a connecting infrastructure (wall, fence, city and road) or by their treatment of nature (agricultural parcel, forest-stands) with of course exceptions like the Scythian burial mount (Kurgan), which cast almost no shadows. Despite the very rare exceptions, physical edge construction is assumed here to be a standard phenomena in populated environments and crucial to any class description containing anthropomorphic features and the basis for topographic maps. The pixels registered on the edges of real world objects are responsible for much of the spectral confusion within an image. The pure spectral mean of an edge has therefore little importance. Separating the edges from other more homogeneous surfaces and using spectral image analysis for non-edge objects alone is preferable. While the object-population of contrasting edges is a must-have condition for anthropogenic objects, they contain unique as well as non-unique features. Their behavior towards the standard deviation in the Red band is unique. Their NDVI values are not unique. They share similar NDVI mean values (not NDVI standard deviation values!) to all other non-vegetative areas. Although it has been stated that the pure spectral mean of an edge object has little importance as there is no predictable mean value in feature space of the class 'edges' in the NIR or Red band, the Ratio value such as the NDVI, still contains meaningful spectral information.

Anchor objects

The contrasting edges can be defined as a candidate for an anchor objects. Be aware that such edges are only one of many potential anchor objects. A famous candidate for anchor objects are the 'High NDVI value objects' which are normally reserved for vegetation classes. But there are many more. The anchor objects should be automatically detectable and representing a much larger class of image objects. The extreme contrasting edges with high standard deviation of the red band will always belong to

non-vegetative areas, if available in the scene. After automatically measuring their NDVI mean value in each scene, they can be used as automatically detected samples and applied for mapping all other objects without vegetation. The contrasting edges are just one example of a set of crucial objects each with their transferable attributes. A list of crucial objects per class would suffice to indicate its class potential to be fully automatically mapped. The main aim becomes now to define standard rules, which are used for crucial classes to establish a population of anchor objects within the scene. The initiation of these anchor objects allows measuring spectral characteristics of this particular scene and assigning these measured values to a process variable. These values from the process variables are applied to detect the thresholds where a critical class separates itself from all other classes. If such as unique separation can be found, the mean and standard deviation of the class is not a decisive factor anymore, because overlap in feature space can be expected to be negligible.

VHSR analysis

The casestudie shows a universal set-up for VHSR imagery such as IKONOS and QuickBird scenes. A pre-processing step is essential, where a 'texture image' splits all objects in high and low textured areas. Texture image derivatives are known for their large variability as well as their redundancy. From experiences with Haralick textural features (Schleicher et al., 2003) our research has moved towards a texture map based upon edge detection filters. Due to the redundancy, this texture derivative is very similar to the Haralick homogeneity feature (known as IDM in Steinnocher, 1997; compare also Fig. 2A and Fig. 2B). Now the texture analysis on panchromatic images has evolved from texture calculations in a fixed window size towards texture analysis per object (Schleicher et al. 2003; Wezyk et al., 2004). Continuing with the advantage of the texture analysis applications in image analysis over the years, a new variance on IDM has been developed using the central cause of an important element of the texture map namely the mixed pixels that share various spectral distributions of its neighbouring surfaces and can be found exactly at the edge of two different distributions, preferably only 1 pixel thick. The pre-advantage of this texture map is the ability to calculate the object feature "contrasting edge". This texture map is produced by dividing an intensity image by the edge image(s). Here the Pan band was used for intensity.
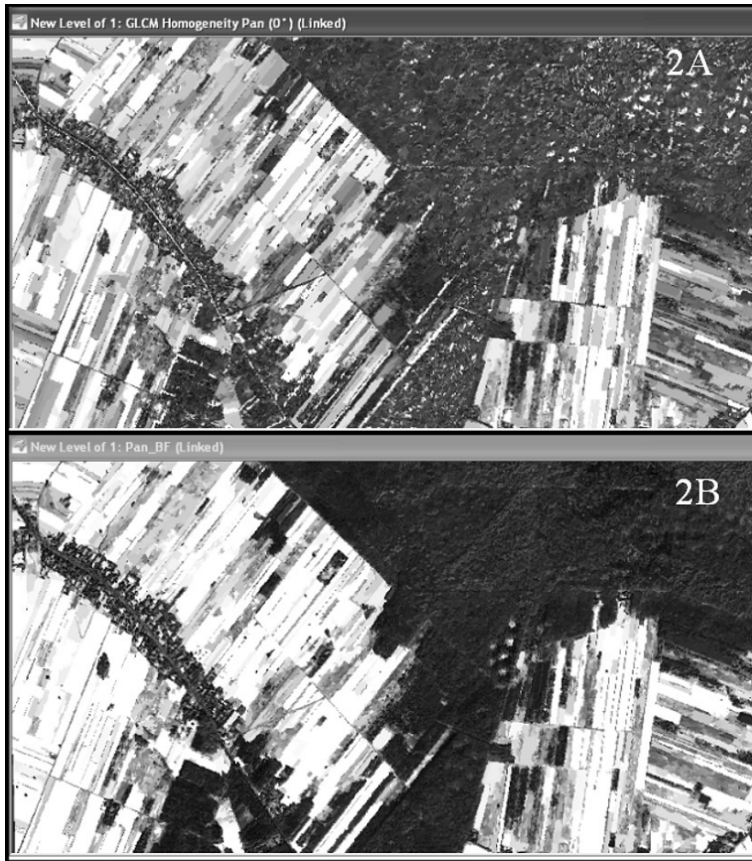
Fig. 2. a) Homogeneity or IDM per object in accordance with Steinnocher IDM analysis (1997). Calculated for the panchromatic band (here QuickBbird, image data from Wezyk et al. 2005); **b)** The Pan band divided by the sum of two edge detection filter images Border and Frame. Border and Frame are images derived from the Lee-Sigma difference filter (Wezyk et al. 2005)

This textural/edge image used to produce Fig. 2B in combination with the Pan intensity is the sum of Border and Frame (Wezyk & de Kok, 2005), which are edges created by the difference of the original Pan band minus the Lee-Sigma filter result. Pixels on edges are dominantly responsible for low IDM values, as they do not belong to a single Gaussian distribution in the image domain. This visualization approaches IDM results (Steinnocher, 1997) as shown in Fig. 2A. This textural image can be calculated with much less computation time than Fig. 2A while containing almost similar information on texture.

A process tree
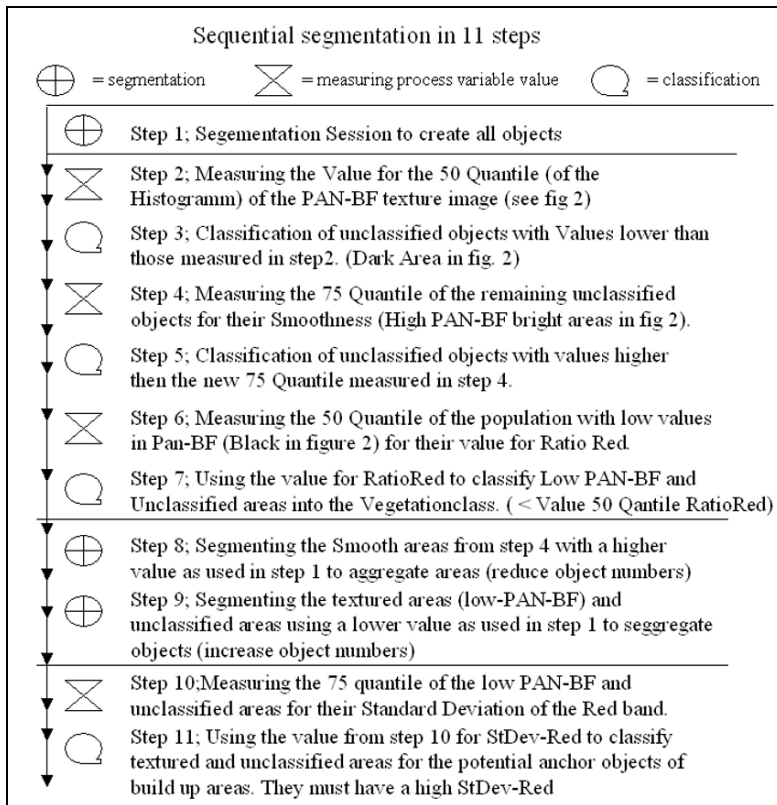
Fig. 3. The process tree in 11 steps for initial VHR analysis

Eleven standard steps: Like any O.O. process we start with initial segmentation with scale factor 35, using Pan with weight=5 and Border Frame weight=3. (See Fig. 3, step 1). In step 2, we measure the lowest 50 Quantile of the texture Histogram. In step 3 - this 50 Quantile is classified as Low-Pan_BF. These objects have a low value for the textured image which means low grey values for Panchromatic band divided by the sum of Border and Frame (Wezyk & de Kok, 2005). This category with alias name Low-PAN_BF always contains artificial areas and forests in any European and North American VHSR imagery. The Fig. 2B visualize this features representing artificial areas and forests in black/dark-grey tones similar to IDM or "Homogeneity" from Haralick Grey-Level-Co-Occurrence matrix analysis. The opposite category, Smooth with the more

than 75 Quantile values of Pan-BF (steps: 4 and 5) always contains mainly agricultural areas and water. In step 6, the highly textured category including forest is measured for their lower value of the Ratio of Red and thus classified (step 7). Here Red is divided by Blue+Green+Red+NIR+PAN. This makes a better Ratio than the NDVI, as we can incorporate the effects of the high resolution of the Pan band. Where normally the NDVI is bright, the Ratio of Red has low values and vice versa. Remember that in step 7 the vegetation part is changed into VEG_1, therefore Low-Pan_BF contains after step 7 mainly artificial areas. The agricultural areas within the category Smooth are generalized in step 8 - to reduce the total amount of objects. In step 9, after reduction of the amount of objects, we are able to increase again the total amount of a more interesting object group by applying a re-segmentation with a lower scale factor of 15 only for both the categories: artificial areas and the remaining unclassified regions. Note that the larger objects of the category Smooth are left untouched. Step 9 has become possible in the new process approach and was not part of previous Definiens developer versions (eCognition), where always an image layer was segmented. The smaller objects in the category artificial areas allow to measure (in step 10) the highest values for the small edges around build up areas which are above the 75 Quantile for the standard deviation of the red band normalized for size (here StDev_Red/(Area^(0,5)) and after that classified (step 11). With these 11 steps we have now automatically derived training areas in 11 steps for the categories Build up Areas, Forests as well as Agricultural areas.

Standardization and outlook

Sampling by a single operator for the purpose of establishing a mean and a standard deviation for a class is a subjective task. Discussing the sample selection with fellow experts is useful for an individual scene but not applicable for all other scenes. The sequential process workflow however allows a discussion of each step within the process with fellow experts. A good example is the case study. Here each of the 11 steps can be meticulously debated by fellow experts and transformed in a standard set for splitting any IKONOS or QuickBird image into a texture class as a standard protocol for extracting automatic samples for any land cover classification. This discussion would lead to objective process-protocols. If the evaluation is positive, the line (set) will reach maturity and allows for standardization of the process (lines). They will become applicable for standard classes of each sensor type. This opens the way to full robotic

land cover classification or full robotic counting as part of an automatic feature detection such as counting: ships, cars, trees etc., which, in this initial stage, can already be demonstrated (de Kok & Wezyk, 2006)

The quest for image understanding

The step-by-step approach of the processes allows task sharing between expert groups. One being responsible for the automatic cloud detector, another for the forest-type mask or settlement masks. Standard libraries for automatic image analysis have the potential to evolve in similar structures like libraries for numerical recipes. Hence reviving the old dream of signal separability of classes that was underlying the ECHO concept ('The Extraction and Classification of Homogeneous Objects') by Kettig and Landgrebe 30 years ago. This would imply a continuous effort for a 'Broader, more fundamentally based research on understanding the signal response' of material under laboratory conditions as well as in the field (after Landgrebe, 1997). Following Landgrebes advice we continue expanding the signal response research. This ongoing work has shown intermediary success.

References

Janssen LLF (1994) Methodology for updating terrain object data from remote sensing data. PhD thesis DLO Staring Centre Wageningen.

Kettig RL, Landgrebe DA (1976) Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects. Institute of Electrical and Electronics Engineers. Reprinted from Transactions on Geoscience Electronics, Vol. GE-14, No. 1, pp. 19-26.

De Kok R, Wezyk P (2006) Process development and sequential image classification for automatic mapping using case studies in forestry. Poster Session. In Proceedings International Workshop. 3D Remote Sensing in Forestry Vienna, 14th-15th Feb. 2006 organised by EARSeL SIG Forestry co-sponsored by ISPRS WG VIII/11. Tatjana Koukal, Werner Schneider (eds.) pp. 380-386.

Landgrebe D (1997) The Evolution of Landsat Data Analysis. Photogrammetric Engineering and Remote Sensing, Vol. LXIII, No. 7, July 1997, pp. 859-867.

Schleicher S, Kammerer P, De Kok R, Wever T (2003) Extrahierung von stabilen Strukturen aus Satellitenbildern zur Klassifizierung von Mosaiken. In: Strobl, J, Blaschke T & Griesebner G (Hrsg.) Angewandte Geographische Informationsverarbeitung XV. Beiträge zum AGIT-Symposium Salzburg 2003, H. Wichmann Verlag, Heidelberg, pp. 431-436.

Steinnocher K (1997) Texturanalyse zur Detektion von Siedlungsgebieten in hochauflösenden panchromatischen Satellitenbilddaten Proceedings AGIT IX. Salzburger Geographische Materialien, Heft 24, pp. 143-152.

Wezyk P, De Kok R, Zajaczkowski G (2004) The role of statistical and structural texture analysis in VHR image analysis for forest applications - A case study on Quickbird data in Niepolomice Forest. In: Strobl J, Blaschke T & Griesebner G (Hrsg.) Angewandte Geoinformatik 2004. Beiträge zum 16. AGIT-Symposium Salzburg 2004, H. Wichmann Verlag, Heidelberg, S. 770-775.

Wezyk P, De Kok R (2005) Automatic mapping of the dynamics of forest succession on abandoned parcels in south Poland. In: In: Strobl J, Blaschke T & Griesebner G (Hrsg.) Angewandte Geoinformatik 2005. Wichman Verlag. Heidelberg. pp. 774-779.

Chapter 7.7

Strategies for semi-automated habitat delineation and spatial change assessment in an Alpine environment

E. Weinke, S. Lang, M. Preiner

Centre for Geoinformatics (Z_GIS), Salzburg University, Schillerstr. 30, 5020 Austria; elisabeth.weinke@sbg.ac.at

ABSTRACT: Driven by vegetation and environmental changes caused by global warming and an ongoing loss of biodiversity, there is an increasing demand for updated image data in a high time rate. These requirements, as well as economic reasons, ask for automated image analysis techniques. This article presents a synthesis of two complementary, preceding studies (Preiner et al., 2006; Weinke and Lang, 2006) on semi-automated habitat delineation. The work has been carried out in two neighboring study sites, which are situated in the Berchtesgaden National Park, in the Alpine region of south-eastern Germany. Both test sites represent a mountainous area, characterized by a high bio- and habitat diversity and a cliffy relief. The NPB has been involved in the European Interreg-IIIb project HABITALP (Alpine Habitat Diversity), which came up with harmonized methods for trans-Alpine monitoring of habitat diversity and environmental changes by standardized, comparative habitat interpretation on color infrared (CIR) aerial photographs. The work presented evaluates the potential of object-based image analysis (OBIA) for partly automising this process. In both studies we used pan-sharpened data from the same QuickBird scene. First, two different approaches, i.e. iterative one-level representation (OLR) strategy and multi-scale segmentation/object relationship modeling (MSS/ORM), are described. Both are meant for dealing with high spectral and spatial variability. They are discussed in the light of

specific settings: whereas iterative OLR was used for delineating homogeneous, but differently scaled single habitat types, MSS/ORM proved suitable for classifying hierarchically structured patches of mountain pine (*Pinus mugo*). Second, based on the results from iterative OLR, a spatial overlay method ('virtual overlay', cf. Schöpfer and Lang, 2006) is applied to characterize habitat fate. The aim is to tell 'real' object changes from differences originating from different ways of delineation, i.e. visual interpretation in the year 2003 vs. automated delineation in 2005. The results showed that both strategies in combination with object-oriented image analysis software are suited to delineate habitats similar to the habitats as addressed by the HABITALP project.

1 Introduction

The Berchtesgaden National Park (NPB) is one of eleven protected areas within the Alps which were participating in the Interreg IIIb project HABITALP (Alpine Habitat Diversity[1]). One aim of this four-year project was to monitor habitats and long-term environmental changes on the basis of color-infrared (CIR) aerial photographs. In the HABITALP project, based on previous experiences with CIR photos, a standardized habitat interpretation key (HIK) has been developed for land use types in protected alpine areas. The HIK-key is designed to be transferred to other high mountain landscapes outside the Alps in the future (HABITALP 2003). NPB has been using CIR photographs since 1980, for carrying out manual delineations and habitat mapping, approximately in an interval of five years. Whereas CIR-aerial photographs are widely used for a variety of natural resource mapping, there is a pronounced demand for using high spatial resolution satellite imagery for this purpose (Neubert and Meinel, 2002; Lang and Langanke, 2006). Besides this general trend more specifically we try to meet reporting obligations as prescribed from legal frameworks such as the Natura-2000. There is an increasing demand on frequently updated image data for monitoring the conservation status of habitats (Langanke et al., 2004). Satellite systems of the new generation with very high spatial resolution (VHSR) of around 1m ground sample distance (GSD) open new perspectives, but likewise challenges (Lang and Langanke, 2006). Data originating from the QuickBird sensor launched in 2001 with a repetition rate of one to three days, depending on the latitude (DigitalGlobe 2006), provide a very high spatial resolution of 0.61 m in

---

[1] www.habitalp.org (2003-2006)

the panchromatic spectral range. The four multispectral bands have a reso-lution of approximately 2.4 m, resolved radiometrically with 11 bit. By applying resolution merge, the spatial resolution of this type of imagery is close to the resolution of operationally used (i.e. downsampled) air-photos. Therefore, for particular tasks, satellite-based data are even more qualified as an alternative to aerial-based data (Lang and Langanke 2006).

Among the numerous image analysis procedures developed in recent years, the conventional per-pixel approaches has limitations when working with VHSR data. In comparison object-based classification schemes have been largely utilized in remote sensing applications in recent years, espe-cially with VHSR data. The object-based classification approach is divided into two major steps: 1) image segmentation and 2) image classification (Kim and Madden 2006). Whereas these steps seem to be distinct, they are in fact mutually interlinked in an altogether cyclic process (Lang, this vol-ume). According to Gorte (1998), the segmentation process is an important step for image analysis forming the conceptual link to human perception and is considered essential for image understanding (Lang and Langanke 2006). Image segmentation subdivides an image into spatially coherent and spectrally homogeneous regions with the aim of a reduction and ag-gregation of data by distilling its inherent information. Through this proc-ess an image is subdivided into its constituent parts and these parts of in-terest (objects) are extracted (Zhang 1996) and classified. But furthermore, segmentation can be used to obtain scaled sets of image primitives as rep-resentatives landscape objects (Lang and Langanke, 2006; Burnett and Blaschke 2003). A range of segmentation strategies exists, including histo-gram-, edge- and region-based segmentation approaches. In the domain of landscape analysis the group of region-based approaches is promising and well established (Lang et al., submitted). Image objects which originate from region-based segmentation techniques can be assessed according to their spatial characteristics as being meaningful. In other words, generated objects exhibit a specific shape which can be considered adequate or not (Blaschke and Strobl 2001). Multiscale approaches try to mimic human perception capacity by hierarchical scalable segmentation and object char-acterization by semantic knowledge (Lang, this volume). A critical issue of these approaches is that the applied 'scale', i.e. the average size of objects may directly influence the quality of classification (Kim and Madden 2006).

The software Definiens (formerly eCognition) follows a multi-scale ap-proach and provides hierarchical levels with image objects in different scales obtained by multi-resolution region-based segmentation (Baatz and Schäpe 2000).

In this study the potential of two strategies, based on the use of the image analysis software Definiens, for representing alpine habitats in two different settings is evaluated. Definiens, or eCognition respectively, was used because its evaluation in previous works (e.g. Meinel and Neubert 2004) showed high qualitative segmentation results and only limited over-segmentation as compared to other software products. We applied two different strategies in this work which were developed especially for this software product and which were declared as suitable to segment scenes of high spatial and spectral variability by Lang and Langanke (2006). The first one, called iterative one-level representation (iterative OLR, Weinke and Lang 2006) was used to extract a given set of habitat types in specific scales. The second strategy, multi-scale segmentation/object relationship modeling (MSS/ORM, Burnett and Blaschke 2003) has been used for extracting patches with varying percentages of mountain pine (*Pinus mugo*). Whereas the first study mainly aims at an optimized delineation of certain target habitats, the latter study involves delineating, modeling and classification of a specific habitat type, i.e. mountain pine. The first set of habitat geometries, originating from iterative OLR, was compared with visually delineated outlines by a human interpreter; both representing geometries at a different time and derived by different technique (cf. Schöpfer et al., this volume). These two geometries have non-congruent object outlines, as manual interpretation extracts image objects, according to the very scale being set, in a generalized manner. Contrarily, the image analysis software being used utilizes a scale parameter for object delineation, yet not generalized (cf. Lang, this volume).

2 Geographical Settings

The two different strategies were applied in two partly overlapping test sites (see Figure 1), situated in the Berchtesgaden National Park, which lies in the Alpine region of South Eastern Germany (47°36'N, 13°00'E). The first study site **A** (*Klausbachtal*) comprised an area of about 16 km², situated in the south-western Klausbachtal. The highest point is located on the western mountain flank of the Hochkalter massif with an elevation of 2460 m a.s.l, whereas the north-easternmost point of the Klausbach river represents the lowest elevation of about 825 m a.s.l. The second test site **B** (*Hochkalter*) is a 29 km² sized area the peak of the Hochkalter massif, rising up to 2607 m a.s.l. Both sites represent a typical Alpine mountainous area, characterized by high biodiversity and habitat diversity and a steep and rough terrain.

The entire NPB has been mapped for fine-scaled biotope and land-use types using eight out of the main HIK categories, each divided in several subclasses. The first test site comprises the following five main HIK types: (1) agricultural land and perennial forb communities, (2) immature soil sites and dwarf-shrub plant community, (3) trees, field trees or shrubs and groups of shrubs, (4) forest and (5) settlement and traffic. Exemplary habitats of these five main types were analyzed in the first study. The second study focused on mountain pines (*Pinus mugo*), which is of European community interest and belongs to HIK type 4 (forest).

Most of the NPB area obviously belongs to the core zone in which natural processes and dynamics are hardly influenced by human interference. The core zone is surrounded by a permanent and temporary transition zone with anthropogenic use or at least influence. Here, managed forests of former times (like saline-spruce forests) are being transformed into nature-orientated mixed forests (Berchtesgaden National Park 2005). In the first test site the permanent and temporary transition zone together makes up about 49%. Accordingly, the management of the national park is interested in an automated habitat monitoring procedure of this area due to its high changes in habitat dynamic. Instead, most of the second study site is located in the core zone. But still there is high interest from the management as the mountain pine is priority habitat type within the Natura-2000 network and, moreover, highly sensitive to climate changes.
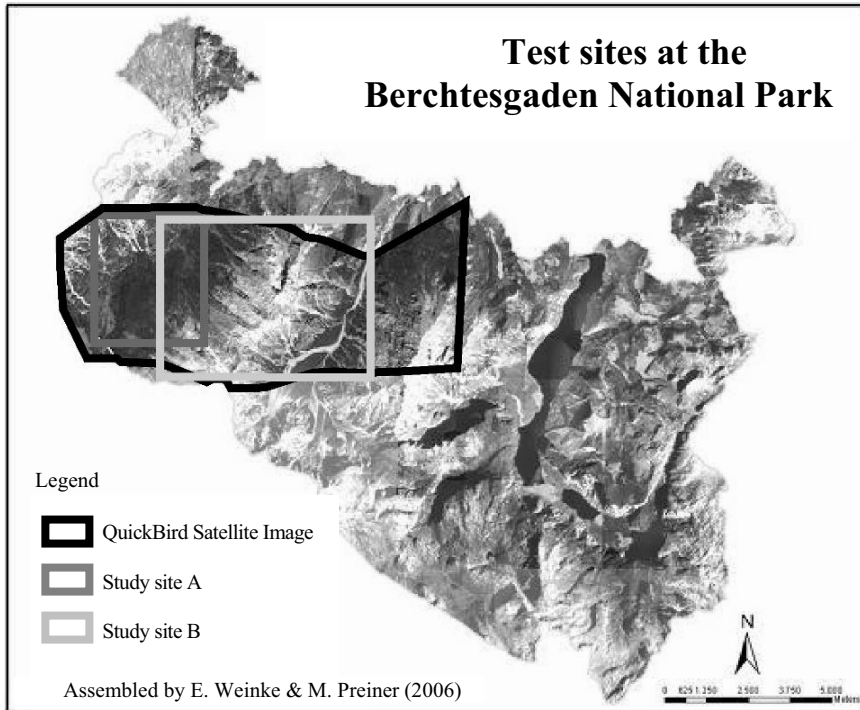
Fig. 1. Study sites at the Berchtesgaden National Park

3 Data and Data pre-processing

VHSR data from a QuickBird (QB) scene was used, recorded on August 2005. The QB-scene was ordered as a standard ortho-ready bundle-product including panchromatic and multispectral dataset (see table 1). Data were pre-projected in UTM-33 (WGS-84), with a low geo-location accuracy of +/- 23 m (Eurimage 2007). Therefore it was necessary to post-orthorectify this QB-product. Because the multispectral and the panchromatic sensor are offset from each other on the QB-focal plane, the same point on the ground is imaged at a slightly different view angles (Hitachi Software Engineering 2004). Accordingly both datasets were separately orthorectified using the rational polynomial coefficients (RPC) orthorectification model. To derive an orthorectified image based on this model, several components were required: a) each of the datasets, b) their rational polynomial coefficients, and c) a digital elevation model (DEM) of 10 m spatial resolution.

The orthorectification process of data representing rough terrain can be improved by specifying accurate, well-distributed ground control points (GCPs) (Smith 2005). The entire study site consists to a lesser degree of anthropogenic objects like streets, larger buildings or rectangular squares. We therefore focused on rooftops of huts, crossroads of paths and some larger singular rocks, which all have a high recognizability and stability (Weinke and Lang 2006). To combine multispectral information with the higher spatial resolution provided by the panchromatic band, we performed two different resolution merge operations, according to the requirements of the two studies. For the Klausbachtal study a pan-sharpening-method after Liu (2000) was used. This process results in a multispectral dataset ground-resolution of 0.6 m. The pan-sharpening procedure is optimized to maintain the original spectral values to a large extent (> 90%), in comparison to methods based on principal component analysis (ibid.). The Hochkalter study used a resolution merge-technique based on principal component analysis to obtain a better differentiation among the vegetation coverage.

Table 1 Spectral, Spatial and Radiometric Resolution of QuickBird Satellite Imagery

Band	Wavelength Region [ym]	Spatial Resolution [m]	Radiometric Resolution [bit]
1	0.45 - 0.52 (blue)	2.44	11
2	0.52 - 0.60 (green)	2.44	11
3	0.63 - 0.69 (red)	2.44	11
4	0.76 - 0.89 (NIR)	2.44	11
PAN	0.45 - 0.90 (PAN)	0.61	11

4 Methods

With regard to object-based analysis of VHSR scenes representing natural environment settings two segmentation strategies (i.e. one-level representation, OLR vs. multi-scale segmentation / object-relationship modeling, MSS/ORM) have been contrasted by Lang and Langanke (2006). We applied both strategies accordingly for identifying Alpine habitats of different scales and composition (see 4.1 and 4.2). For the image analysis process we used the software eCognition 4.0 with its implemented region-based local mutual best fitting algorithm (Baatz and Schäpe 2000). The resulting image objects are controllable by parameterization of color and

shape homogeneity. To identify changes we applied habitat fate analysis in study site A as described in section 4.3.

4.1 Iterative one-level representation (test site A, Klausbachtal)

The test site **A** entails boundary delineation of a given set of habitat types via iterative OLR (see figure 2). In this case, OLR means that through iterative segmentation entire habitats are delineated at one single target level which reflects the appropriate scale domain (Lang 2002). In a strict hierarchical multiscale segmentation approach (see Lang, this volume) an image is subdivided in non-overlapping regions so that image objects on coarser levels topologically coincide with objects in finer sub-levels. Applying OLR in test site **A**, several subsets were selected and single habitats were segmented on specific target levels. The term target level describes the level at which one single habitat like e.g. meadow or mountain pine is completely segmented with specific parameters within a specific subset. Therefore each target level may consist of several (temporary) sublevels (Lang and Langanke, 2006, see also Figure 2). The quality of the segmented objects is compared with reference habitat boundaries from a visual interpretation of CIR air photos, using overlay approach as described in chapter 4.3. Then the specific target levels of single habitats were transferred to the entire test site. By this, at some levels, habitats were identified in the test site as a result of the same iteration steps and parameters. A user-defined number of target levels can be extracted from the entire test scene via iterative OLR. Habitat geometries of neighboring target levels should not be disregarded during the process (Weinke and Lang 2006). At last the single levels are reassembled on one super level (Preiner et al. 2006).
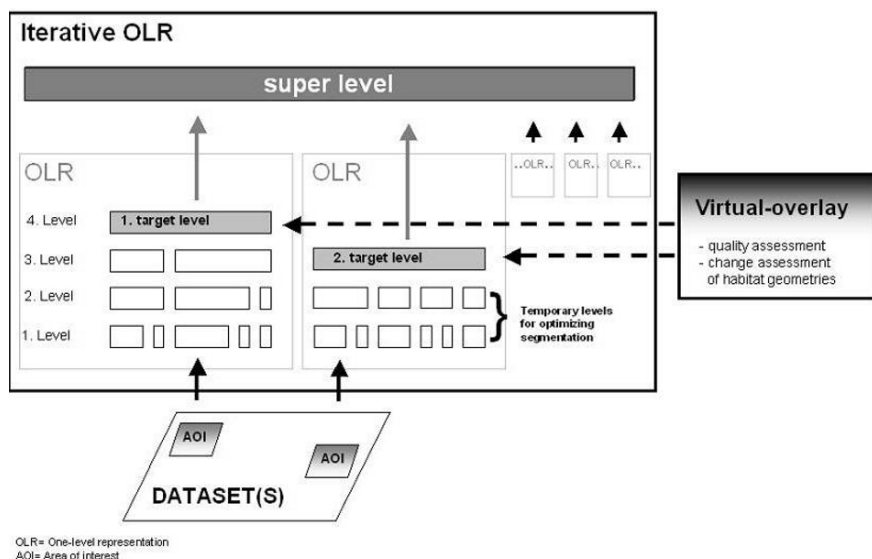
Fig. 2. Scheme of iterative OLR in combination with the virtual-overlay approach (see chapter 4.3, Weinke and Lang 2006, modified by E. Weinke 2007)

4.2 Multi-scale segmentation/object relationship modeling (test site B, Hochkalter)

In test site **B**, *Hochkalter*, we applied the MSS/ORM approach (Burnett and Blaschke 2003). In contrast to the OLR strategy image objects are represented in at least two scale-levels and relationships are built among them (see Figure 3). Hierarchically arranged levels of image objects are generated by strict hierarchical segmentation (multiscale segmentation, MSS). By this, besides spectral and form related features of the generated image objects also their horizontal and vertical spatial relations are known. The generated image objects are the basis for the subsequent object relationship modeling (ORM). For realizing ORM first the decisive features of the generated image object are analyzed according to the study purpose. As a second step a hierarchy of classified objects is constructed with a set of semantic rules utilizing fuzzy-based membership functions. Finally, this process leads to user-demanded objects, aggregated from the former building block-like image objects. Since usually this process is not a linear one, recent literature uses the term 'class modeling' to indicate its cyclic character (Tiede et al. 2006, Lang, this volume, Tiede et al., this volume)

The main target class of the Hochkalter case study was mountain pine shrubbery, a subtype of the FFH[2] habitat type Bushes with *Pinus mugo* and *Rhododendron hirsutum (Mugo-Rhododendretum hirsuti)*.

Within this test site, areas covered by *pinus mugo* are characterized by a heterogeneous arrangement (Preiner et al. 2006). Thus they were represented on up to five different scale levels and addressed by a rather complex class modeling process. Several sub-classes were discerned, such as *Pure mountain pine stand, Mountain pine with single trees; Mountain pine with trees: Mountain pine with rock, debris or pebbles; Mountain pine with alpine meadow and/or rock, debris or pebbles; Forest with mountain pine, Alpine meadow with mountain pine; Rock, debris or pebbles with mountain pine*. For the internal differentiation of the classes a number of additional auxiliary classes (*Forest; Alpine meadow; Rock, debris or pebbles; Firn- or Snowfield,…*) were introduced.

To reach the objectives of a spatially exhaustive classification the single objects (e.g. tree, bush, single rock …) and small components (e.g. shaded trees sunlit trees) were defined according to their spectral and form features on the finest segmentation level (near pixel level). Afterwards and contrary to this, the heterogeneous composite objects of the higher scale levels (e.g. *Mountain pine with single trees*; *Alpine meadows with mountain pine)* were identified using additional semantic information and topological object relations, such as 'number of trees-objects on a lower segmentation level'.

A special refinement, with an extra process cycle was applied, only concerning the mixed areas and transition zones between mountain pine and wood, alpine grassland or rocks, which were extracted not satisfyingly. To this end, all mixed classes (e.g. *Alpine meadow with mountain pine; Mountain pine with single trees; Forest with mountain pine,…*) were merged into one single class, exported to ArcGIS and re-imported as a thematic layer, which was then exclusively segmented and classified without affecting the already established classes (*Forest; Alpine meadow; Pure mountain pine stand; Rock, debris or pebbles; Firn- or Snowfield,…*).

---

[2] Flora-Fauna-Habitat Directive (92/43/EEC)
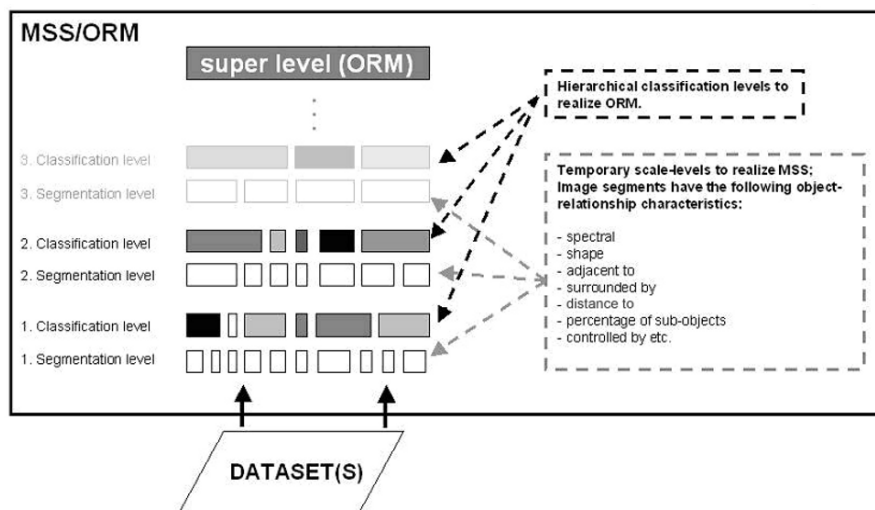
Fig. 3. Scheme of MSS/ORM (Lang and Langanke 2006, modified by E. Weinke 2007)

4.3 Change assessment: habitat fate

Change assessment was examined in the study site **A** (*Klausbachtal*). After the extraction of single habitat geometries via OLR, the segmented geometries of the QB scene 2005 were compared with visually delineated and manually digitized geometries on CIR orthofotos from the year 2003. The approach ('virtual overlay', Schöpfer and Lang, 2006) investigates spatial relationships among corresponding objects (see Schöpfer et al, this volume, for further discussion). Object correspondence can be seen as a product of spatial object change over time or as a result of different object delineations (ibid.). In this study, two different geometries with non-congruent object outlines have been introduced by two different systems, even if working in the same scale domain. Those systems were: (1) the visual human-perceptive system and (2) the digital image analyses system. In addition, virtual overlay has also been used to compare habitat outlines of two different time slots and to perform spatial change assessment. Implemented in a tool called LIST (*Landscape Interpretation Support Tool*, Lang et al. in press), an extension for ESRI's ArcGIS 9, two vector layers are analyzed in terms of their specific object fate (ibid.) through time. In this case the analyzed objects correspond to habitats; thus we consequently use the term 'habitat fate' (Weinke and Lang 2006). Habitat fate of two comparable geometries was examined by setting a threshold value for

'spatial overlay strictness (SOS)' to cope with overlaps of corresponding, but not fully congruent boundaries. By the threshold value a specific buffer zone is introduced. The buffer size controls the allowed spatial difference of two habitat geometries under consideration; the SOS-threshold reflects the degree of overlaps, expressed by a percentage value. Using virtual overlay altering geometries can be analyzed over time without modifying geometries (Schöpfer and Lang 2006). Habitat fate is investigated by overlaying geometries from habitat fate time t_0 and t_1. In our case, the respective status at t_0 is taken as reference habitat reflecting the visual delineation on orthophotos. Instead, habitat fate time t_1 represents the segmented outlines provided by eCognition 4.0. A habitat at t_0 can change over time; consequently, the same habitat at t_1 can be characterized by the following four expressions: (1) 'good habitat': t_1-habitat is completely within the buffered outline of a t_0-habitat; (2) 'expanding habitat': t_1-habitat is exceeding the buffered outline of a t_0-habitat but its centroid is within the delineated outline of habitat t_0; (3) 'invading habitat': t_1-habitat has its centroid outside the origin or buffered t_0-object, therefore usually this object type in not directly related to object t_0 (Lang et al. in press). Finally (4) 'emerged habitat': t_1-habitat which is completely within the original outline of a t_0-habitat (see figure 4). A special case of habitat fate occurs when habitat t_1 behaves like an expanding habitat but represents an area which is completely new developed (such as a clear-cut forest). In this case LIST will identify this habitat as an expanding habitat. In our study we only calculated habitat fate types 1, 2 and 4, because we determined the geometries of one habitat at two different times (t_1 and t_0). That means the analyzed habitats either can be a good, expanding or emerged (Weinke and Lang 2006). Note that the concept of object relationships has been further extended by Schöpfer et al. (this volume).
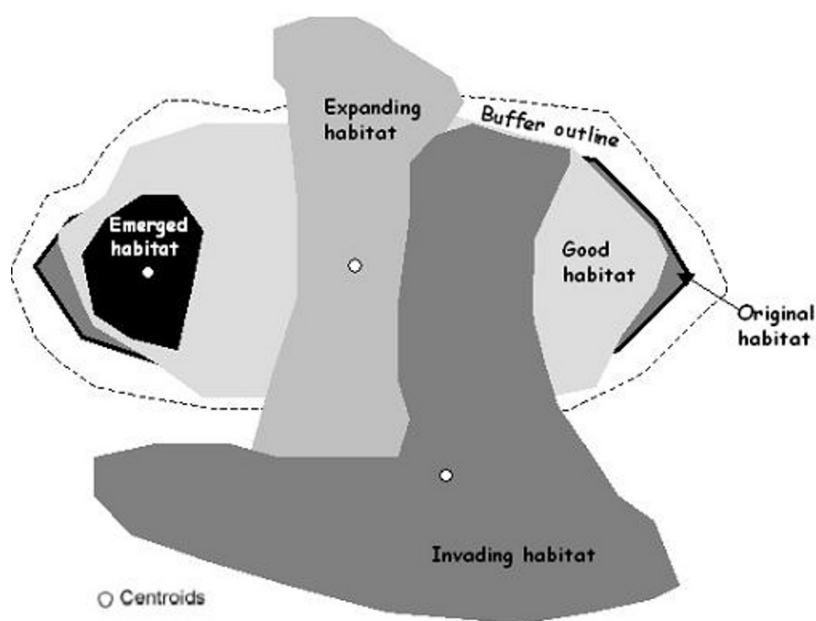
Fig. 4. 'Habitat fate' categories implemented in LIST (Schöpfer and Lang 2006, modified by Weinke and Lang 2006)

5 Results

5.1 Test site A

5.1.1 Iterative OLR

Using OLR, habitats were segmented at single representation levels which produced habitat boundaries similar to the delineations of the human interpreter. The quality of the habitats introduced at different target levels depends upon the defined iteration steps and parameters used in eCognition 4.0.

Altogether, forty-four single habitats (see Figure 5) were extracted at eight target levels. At each representation level a specific number of habitats with different textural features, different habitat type (see table 2) and size were identified. These results need to consider the high number of heterogeneously composed habitats and the high variation of habitat size in the entire test site. Although after Gonzales and Woods (2002) and Sonka et al. (1999) a formal definition of texture hardly exists, texture comprises

instinctive characteristics like smoothness, coarseness, graininess and regularity in a general understanding. Obviously, also habitats represented on a satellite imagery can exhibit these characteristics. A meadow for example has a homogeneous smooth texture, in comparison to forest habitat, the texture of which is heterogeneous. The extracted habitats represent five of the eight main habitat types of the standardized HIK. Although different habitat types were segmented at each target level, some target levels are more suitable to extract a specific habitat type (HABITALP Interpretation key 2003). As an example, in target level 4 most of the extracted habitats are of type immature soil site (HIK: 5000) which is made up by a combination of coniferous and debris areas. To reach this target level at the first iteration steps we selected a very high color and compactness factor (see table 3). Only in the last iteration step 'shape' was weighted higher. These results show that by applying the iterative OLR segmentation strategy it is possible to obtain one specific habitat or more single habitats at a specific target level, which have a similar texture and represent a specific main HABITALP-mapping unit.

Table 2 Number of segmented superior HABITALP-mapping units per each target level (Weinke and Lang 2006)

target levels	HIK 4000[a]	HIK 5000[b]	HIK 6000[c]	HIK 7000[d]	HIK 9000[e]	num. of habitats/ target level[f]
1. target level	3			1	5	9
2. target level	5	1		7		13
3. target level	2	1	2	1	5	11
4. target level	1	4		1		6
5. target level	1	1				2
6. target level				1		1
7. target level				1		1
8. target level				1		1
					noH[g]	44

[a]HIK 4000: agricultural land, perennial forb communities.
[b]HIK 5000: immature soil sites, dwarf-shrub plant community.
[c]HIK 6000: trees, field trees or shrubs, groups of shrubs.
[d]HIK 7000: forest.
[e]HIK 9000: settlement, traffic.
[f]num. of habitats/ target level: number of habitats per each target level.
[g]noH: number of habitats

Table 3 Color- and Compactnessparameters of each hierarchical segmentation level on target level 4

Parameters	SL1 [a]	SL 2 [a]	SL 3[a]	SL 4 [a]
Scale	30	90	180	300
Color / Shape	1 / 0	0.9 / 0.1	0.9 / 0.1	0.6 / 0.4
Compactness / Smoothness	0.5 / 0.5	0.9 / 0.1	0.9 / 0.1	0.9 / 0.1

[a]SL 1,2,3,4: Hierarchical Segmentation levels which are needed to extract habitats on target level 4 .

Number of habitats of all target levels

44

Assembled by E. Weinke (2007)

N

0   200   400   800   1.200   1.600
Meters

Fig.5. Extracted habitats originated from the iterative OLR segmentation strategy in the test site Klausbachtal

5.1.2 Habitat boundary coincidence and change assessment

As stated before, due to the fact that two different systems were used for producing habitat outlines, we are dealing with non-congruent object outlines. In table 4, one example for each of the main HABITALP-mapping units with non-congruent habitat outlines is shown. Applying the virtual-overlay approach, habitat fate of each habitat is identified. The habitat of the first example shows an instance of the main HIK category 'agricultural land, perennial forb communities' with rather homogeneous texture and no changes between 2003 and 2005. As a consequence a small SOS-factor of 5% was used to identify this habitat as a good (i.e. stable) habitat (refer to chapter 4.3. to recall the how SOS-factor and habitat fate are related). The second example consists of a heterogeneous texture of debris and coniferous areas and is an instance of the main HIK category 'immature soil sites and dwarf-shrub plant community'. A SOS-factor of 9.4% was used to identify good 'habitat fate' of this habitat. In this case the high buffer distance is justified because the interpreter perceptually included single features (like trees) in an otherwise homogeneous debris habitat matrix, whereas the software strictly segments the homogeneous debris area. In the sixth example an SOS-factor of 4% was applied to ensure good habitat fate. This case shows that the software extracts the entire house, which belong to the main HIK category 'settlement and traffic', because it works along pixels, whereas a human interpreter will generalize. The habitat in the third example has a high SOS-factor of 14.2% and is an expanding habitat. The software apparently cannot differentiate between tree crown and tree shadow. Consequently the expanding tree crown in the south-eastern part of the habitat could not be identified. The habitat in the fourth example also shows an expanding habitat with a high SOS-factor of 15%. This habitat is truly expanding because spectral changes could be identified in the south-east of this habitat. About half of the test site is situated in the temporary and permanent transition zone of the national park. Thus many habitats as the habitat in the fifth example could be identified. This represents a clear-cut forest, as an example for the main HIK category 'special sub mapping unit of forest' and an emerged habitat (Weinke and Lang 2006).

Table 4 Examples for each main HABITALP-mapping unit which can be found in the Klausbachtal (Weinke and Lang 2006, modified by E. Weinke 2007)

1) Agricultural land, perennial forb communities (HIK:4000)	Characteristics of segmented habitats with eCognition 4.0: Target level: 1 Size: ~2100m² Habitat fate: good SOS-factor: 5% Texture: homogeneous	2) Immature soil sites, dwarf-shrub plant community (HIK: 5000)	Characteristics of segmented habitats with eCognition 4.0: Target level: 5 Size: ~3800m² Habitat fate: good SOS-factor: 9,4% Texture: heterogeneous
HIK-boundary overlays CIR-orthophoto 2003	change assessment (QuickBird-imagery 2005)	HIK-boundary overlays CIR-orthophoto 2003	change assessment (QuickBird-imagery 2005)
3) Trees, field trees or shrubs, groups of shrubs (HIK: 6000)	Characteristics of segmented habitats with eCognition 4.0: Target level: 3 Size: ~6800m² Habitat fate: expanding SOS-factor: 14,2% Texture: heterogeneous	4) Forest (HIK: 7000)	Characteristics of segmented habitats with eCognition 4.0: Target level: 7 Size: ~13300m² Habitat fate: expanding SOS-factor: 15% Texture: heterogeneous
HIK-boundary overlays CIR-orthophoto 2003	change assessment (QuickBird-imagery 2005)	HIK-boundary overlays CIR-orthophoto 2003	change assessment (QuickBird-imagery 2005)
5) Special sub mapping unit of Forest (HIK:7000)	Characteristics of segmented habitats with eCognition 4.0: Target level: 2 Size: ~3900m² Habitat fate: emerged SOS-factor: 0% Texture: heterogeneous	6) Settlement, traffic (HIK: 9000)	Characteristics of segmented habitats with eCognition 4.0: Target level: 1 Size: ~258m² Habitat fate: good SOS-factor: 4% Texture: heterogeneous
HIK-boundary overlays CIR-orthophoto 2003	change assessment (QuickBird-imagery 2005)	HIK-boundary overlays CIR-orthophoto 2003	change assessment (QuickBird-imagery 2005)

Legend: ▭ Buffer zone   ▭ Habitat border which results of a segmentation process   ▭ Visual border delineation of CIR-aerial photographs

5.2 Test site B- MSS/ORM

According to the objectives of the Hochkalter study mountain pine areas were identified and classified. Especially the mixed classes, the ecologically relevant transition zones between wood and mountain pines as well as mountain pines and alpine grassland or rock were well recognized (see figure 6). Problems due to topographic illumination effects (shadowed objects, similarities of different classes, spectral differences within one class) have been eliminated mostly satisfyingly by using semantic rules and topological relations.

Processing times increased substantially while processing the entire subset size of 29 km^2 (Preiner et al. 2006).

(a)   (b)

Apparent classes:
A) Pure mountain pine stand
B) Mountain pine with alpine meadow and/or rock, debris or pebbles
C) Mountain pine with trees
D) Mountain pine with single trees
E) Forest with mountain pine
F) Heterogeneous class with mountain pine
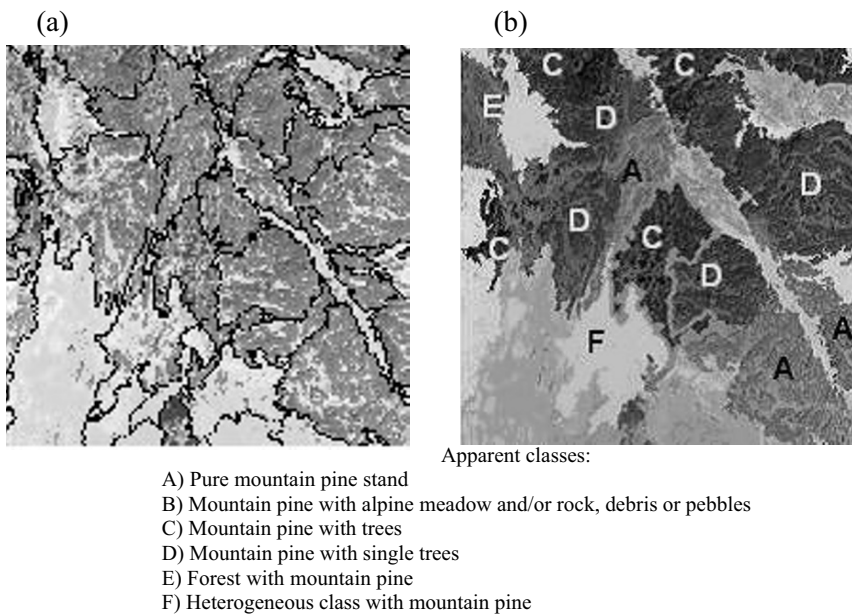
Fig. 6. Extracting process of qualitative sub-classes of mountain pine vegetation within the transition zone. **(a)** Initially segmented image object composition **(b)** Final classification result with user-defined modified useful image objects (Preiner et al. 2006)

6 Discussion

In this article we showed that very high spatial resolution (VHSR) Quick-Bird imagery are well suited for automated habitat delineation. QB satellite data, as compared to CIR aerial photographs, also provide a blue channel and a separate panchromatic dataset. Using a suitable resolution merge technique, the spatial resolution of the multispectral dataset can be combined with the high spatial resolution of the panchromatic dataset, maintaining most of the high radiometric resolution of the multispectral bands. Approximately, pan-sharpened QB-data have similar spatial resolution as CIR-aerial orthophotos, so segmentation results are usually more appropriate on a merged dataset. Leaving technical limits aside, we assume that satellite data providing multispectral bands with sub-meter resolution will improve segmentation results even more. Further advantages are the temporal resolution and revision periods of satellite sensors for scanning the same region of the earth at similar illumination conditions. If looking at the cost side, habitat mapping using QuickBird imagery, with standard pre-processing steps and semi-automated analysis will be altogether more cost-efficient than mere visual interpretation (Lang and Langanke 2006).

In this work we evaluated the potential of OBIA using the software eCognition 4.0 for habitat mapping, partly automating this process. Since the last years a large number of tools for analyzing remote sensing data appeared. Meinel and Neubert (2004) and Neubert et al. (2006) evaluated some of these tools focusing on the quality of the results of the segmentation process, which maybe considered the crucial and basic step of OBIA. Visual inspection and a qualitative assessment of these results showed that beside eCognition also SegSAR 1.0, SEGEN and SPRING 4.0 provide satisfying results. Further work will therefore focus on investigating the potential of such alternatives for habitat delineation. However, in terms of operational use eCognition provides a user-friendly graphical user interface, integrates both segmentation and classification, and supports efficient data interchange.

The results of the two studies show the considerable potential of the multi-resolution region-based segmentation in combination with the iterative OLR and MSS/ORM-strategy for automated habitat delineation and classification. These results have important implication for habitat monitoring and management, particularly in areas that are undergoing significant land use/land cover change. Depending on thematic aims and further planned applications either of the strategies being discussed can provide the desired benefits for extracting (alpine) habitats. Change assessment was evaluated for the segmented habitats resulting from iterative OLR.

This method can be used for assessing or updating respective data layers from previous dates, even if being derived by a different method. This is important since as it has happened in the past, techniques for delineation may change even further. Still, telling apart slight modifications from real changes, is a challenge. Appropriate representation in a geodata model is another challenge. One possibility to address this problem would be to connect a certain segmentation algorithm with a database to perform automated segmentation using habitat specific parameters, which can be applied and adapted to identify changes.

Further applications, in particular for further assessment of complex target classes (like the *Pinus mugo* transition zones), could be done via landscape metrics as implemented in the V-LATE tool (Lang and Tiede 2003). The spatial characterization and quantitative description concerning the very arrangement and composition of the aggregates may be linked to some crucial ecological parameters of the respective habitats.

Acknowledgements

References

Baatz M, Schäpe A (2000) Multiresolution Segmentation- an optimization approach for high quality multi-scale image analysis. In: Strobl J, Blaschke T (eds) Angewandte geographische Informationsverarbeitung, vol. XII., pp. 12 - 23 – Heidelberg

Berchtesgaden National Park (2005) Umweltschutz- Planung und Zonierung. http://www.nationalpark-berchtesgaden.de/html/naturschutz.html (accessed 24.10.2005)

Blaschke T, Strobl J (2001) What's wrong with pixels ?- Some recent developments interfacing remote sensing and GIS. GIS- Zeitschrift für Geoinformationssysteme 6/2001, pp. 12-17

Blaschke T, Lang S, Möller M (2005) Object-based analysis of remote sensing data for landscape monitoring: Recent Developments. Anais XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16-21, INPE, p. 2879-2885

Blaschke T, Conradi M, Lang S (2001) Multi-scale image analysis for ecological monitoring of heterogeneous, small structured landscapes. Proceedings of SPIE, Toulouse, p. 35-44

Burnett C, Blaschke T (2003) A multi-scale segmentation/object relation-ship modeling methodology for landscape analyses. Ecological Modeling, 168, pp. 233-249

DigitalGlobe (2006) QuickBird-QuickBird Specification. http://www.digitalglobe.com/about/quickbird.html (accessed 23.5.2006)

Eurimage (2007) Multi-Mission Satellite Data. http://www.eurimage.com/faq/faq.html (accessed 17.8.2007)

Gorte B (1998) Probabilistic Segmentation of Remotely Sensed Images. ITC Publication Series No. 63

Gonzales RC, Woods RE (2002) Digital Image Processing. Prentice Hall, Upper Saddle River, NY

HABITALP (2003) http://www.habitalp.de/ (accessed 10.3.2006).

HABITALP Interpretation key (2003) http://www.habitalp.org/ (accessed 26.8.2005)

Hitachi Software Engineering (2004) Quickbird Imagery Products FAQ. http://www.hgiis.com/documents/en/FAQ_03-01-2004_en.pdf   (accessed 12.8.2005)

Kim M, Madden M (2006) Determination of optimal scale parameter for alliance-level forest classification of multispectral IKONOS images. In: Lang S, Blaschke T, Schöpfer E (eds) Proceedings of the 1st International Conference on Object-based Image Analysis, July 4-5, 2006

Lang S (2002) Zur Anwendung des Holarchiekonzeptes bei der Generierung regionalisierter Segmentierungsebenen in höchst-auflösenden Bilddaten. In: Blaschke T (ed) Fernerkundung und GIS: Neue Sensoren – innovative Methoden, pp. 24-33 – Wichmann, Heidelberg

Lang S (2005) Image objects vs. landscape objects. Interpretation, hierarchical representation and significance. Unpublished PhD thesis, Salzburg University.

Lang S, Langanke T (2006) Object-based mapping and object-relationship modeling for land use classes and habitats. PFG - Photogrammetrie, Fernerkundung, Geoinformatik, vol. 6: 5-18

Lang S, Tiede D (2003) vLATE Extension für ArcGIS – vektorbasiertes Tool zur quantitativen Landschaftsstruktur-analyse, Proceedings ESRI User Conference 2003 Innsbruck

Lang S, Schöpfer E, Langanke T (in press) Combined object-based classification and manual interpretation – Synergies for a quantitative assessment of parcels and biotopes. Geocarto International

Liu JG (2000) Smoothing Filter-based Intensity Modulation- a spectral preserve image   fusion   technique   for   improving   spatial   details. http://www3.imperial.ac.uk/pls/portallive/docs/1/11657.PDF   (accessed 12.9.2005)

Langanke, T., T. Blaschke & S. Lang (2004) An object-based GIS/remote sensing approach supporting monitoring tasks in European-wide nature conservation. .

First Mediterranean Conference on Earth Observation (Remote Sensing), April 21-23, 2004, Belgrade, pp. 245-252

Meinel G, Neubert M (2004) A comparison of segmentation programs for high resolution remote sensing data. Intern. Archives of Photogrammetry and Remote Sensing, Vol. XXXV, Part B, Commission IV, (Istanbul)

Neubert M, Herold H, Meinel G (2006) Evaluation of remote sensing image segmentation quality - Further results and concepts In: Lang S, Blaschke T, Schöpfer E (eds) Proceedings of the 1st International Conference on Object-based Image Analysis, July 4-5, 2006

Neubert, M. & G. Meinel (2002) Segmentbasierte Auswertung von IKONOS-Daten. Anwendung der Bildanalyse Software eCognition auf unterschiedliche Testgebiete. In: T. Blaschke (ed.), Fernerkundung und GIS: Neue Sensoren – innovative Methoden, Wichmann Verlag, Heidelberg, pp. 108-117

Preiner M, Weinke E, Lang S (2006) Two structure-related strategies for automatically delineating and classifying habitats in an alpine environment. In: Lang S, Blaschke T, Schöpfer E (eds) Proceedings of the 1st International Conference on Object-based Image Analysis, July 4-5, 2006

Schiewe J, Tufte L (2002) Potential und Probleme multiskalarer Segmentierungsmethoden der Fernerkundung. In: Blaschke T (ed): Fernerkundung und GIS: Neue Sensoren – Innovative Methoden. Wichmann, Heidelberg, p. 42-51

Schöpfer E, Lang S (2006) Object fate analysis - a virtual overlay method for the categorisation of object transition and object-based accuracy assessment. In: Lang S, Blaschke T, Schöpfer E (eds) Proceedings of the 1st International Conference on Object-based Image Analysis, July 4-5, 2006

Smith RB (2005) Tutorial- Orthorectification using Rational Polynomials with TNTmips. http://www.microimages.com/getstart/pdf_new/rpcortho.pdf (accessed 1.11.2005)

Sonka M, Hlavac V, Boyle R (1999) Image processing, analysis and machine vision. PWS Publishing, NY

Tiede D, Lang S, Ch. Hoffmann (2006) Supervised and forest type-specific multi-scale segmentation for a one-level-representation of single trees. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. No. XXXVI-4/C42, Salzburg, Austria

Weinke E, Lang S (2006) Automated boundary delineation and spatial change assessment of alpine habitats (Berchtesgaden National Park, Germany). In: Kaufmann V, Sulzer W (eds) Proceedings of the 9th International Symposium on High Mountain Remote Sensing Cartography, September 14-15, 2006

Zhang YJ (1996) A survey of evaluation methods for image segmentation. Pattern Recognition 29(8), p. 1335-1346

Section 8
Burning research questions, research needs and outlook

Chapter 8.1

On segment based image fusion

M. Ehlers, D. Tomowski

University of Osnabrueck, Institute for Geoinformatics and Remote Sensing, Germany
mehlers@igf.uni-osnabrueck.de, dtomowski@igf.uni-osnabrueck.de

KEYWORDS: Integration of GIS and remote sensing, NDVI segmentation, decision and feature based fusion techniques

ABSTRACT: The new generation of satellite and aircraft sensors provides image data of very and ultra high resolution which challenge conventional aerial photography. The high-resolution information, however, is acquired only in a panchromatic mode whereas the multispectral images are of lower spatial resolution. The ratios between high resolution panchromatic and low resolution multispectral images vary between 1:2 and 1:8 (or even higher if different sensors are involved). Consequently, appropriate techniques have been developed to merge the high resolution panchromatic information into the multispectral datasets. These techniques are usually referred to as pansharpening or data fusion. The methods can be classified into three levels: pixel level (iconic) fusion, feature level (symbolic) fusion and decision level fusion. Much research has concentrated on the iconic fusion because there exists a wealth of theory behind it. With the advent of object or segment oriented image processing techniques, however, feature based and decision based fusion techniques are becoming more important despite the fact that these approaches are more application oriented and heuristic. Within this context, the integration of GIS based information can easily be accomplished. The features can come from a specific segmentation algorithm or from an existing GIS database. Within the context of feature and decision based fusion, we present two exemplary case studies to prove the potential of decision and feature based fusion. The examples include:

- Decision based integration of panchromatic high resolution data with multispectral images for the identification of settlement areas;
- Rapid image enhancement merging GIS and multispectral satellite data; and
- NDVI based segmentation.

1 Introduction

The availability of remote sensing data that are needed for global, regional and local monitoring has greatly increased over the recent years. While the increase in spatial resolution for digital images has been hailed as a significant progress, methods for their automated analyses (i.e. land cover mapping, change analysis, GIS integration) are still in the process of being developed. Object (or segment) based preprocessing techniques seem to be an adequate methodology because inter-class variances can be minimized and the image interpretation techniques of the human eye be mimicked. However, the question of appropriate data fusion techniques within this context has hardly been addressed.

Over the last years, image fusion techniques have gained a renewed interest within the remote sensing community. The reason for this is that in most cases the new generation of remote sensors with very high spatial resolution records image datasets in two separate modes: the highest spatial resolution is obtained for panchromatic images whereas multispectral information is associated with lower spatial resolution. The ratios between panchromatic and multispectral imaging modes of one sensor vary between 1:2 and 1:8. For multisensor fusion, ratios can exceed 1:20 (e.g. Ikonos and SPOT merge). Consequently, for analyses that require both, high spatial and spectral information, fusion techniques have to be developed to extract 'the best of both worlds'. The term fusion is used by the image community to address the problem of sensor fusion, where images from different sensors are combined. The term is also used by the database community for parts of the interoperability problem. Generally, fusion exists in different forms in different scientific communities (see, for example, Edwards and Jeansouline 2004).

Usually, the imaging community uses it to address the problem of sensor fusion, where images from different sensors (or different modes of one sensor) are combined. They can be classified into three levels: pixel level (iconic), feature level (symbolic) and knowledge or decision level (Pohl and van Genderen 1998).

Until now, of highest relevance for remote sensing data processing and analysis have been techniques for iconic image fusion for which many different methods have been developed and a rich theory exists. Unfortunately, for many fusion techniques we experience more or less significant color shifts which, in most cases, impede a subsequent automated analysis (Ehlers & Klonus 2004, Ehlers et al. 2007). Even with a fusion technique that preserves the original spectral characteristics, automated techniques often do not produce the desired results because of the high spatial resolution of the fused datasets (Tomowski et al. 2006).

For this purpose, feature based or decision based fusion techniques are employed that are usually based on empirical or heuristic rules. Because a general theory is lacking, these fusion algorithms are usually developed for certain applications and datasets. To discuss the advantages and disadvantages on segment based image fusion techniques, we introduce three examples for segment based fusion methods in this paper (decision based data fusion, GIS information integration, and 'NDVI cookie cutter' fusion).

2 Decision based fusion

As a basis for the decision based fusion process, we selected high and medium spatial resolution satellites data to develop, implement, and test a method for the automated detection of settlement areas. The high resolution satellite datasets were comprised of panchromatic images from SPOT-5 (Fig. 1) with 5 m GSD and KOMPSAT-1 with 6.6 m GSD (Fig. 2). Medium resolution multispectral data were obtained from Landsat ETM and Aster datasets with 30 m and 15 m resolution, respectively. Our method was applied to two randomly selected test areas (25 km^2 each), using panchromatic and multispectral satellite data. For the first area, data from SPOT (recording date 16 March 2003) and Landsat (recording date 26 June 2001) were used, and for the second, KOMPSAT-1 (recording date 20 May 2004) and Aster data (recording date 3 August 2003).

Fig. 1. Panchromatic SPOT-5 image (5 m pixel size)

Fig. 2. Panchromatic KOMPSAT-1 image (6.6 m pixel size)

The aim was to produce a binary mask with the classes 'settlement' and 'non-settlement'. Settlement is understood as a sum of real estates, traffic surfaces, commercial areas, sport and recreation facilities as well as parks

and cemeteries (Apel & Henckel, 1995). The images were rectified to ground coordinates but otherwise left in their original format. Parameters such as texture and shape were extracted from the high resolution panchromatic data, vegetation information from the multispectral images (Fig. 3).
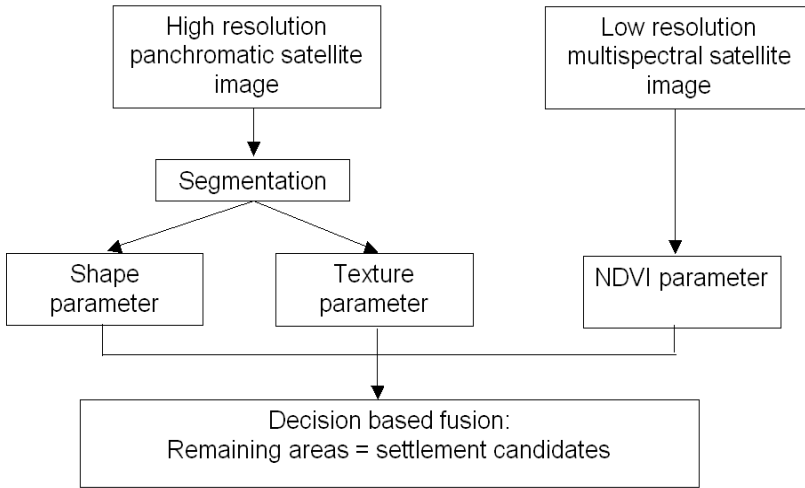
Fig. 3. Decision based data fusion process

Using an adaptive threshold procedure, the information from the image datasets was fused and formed a binary mask for the areas 'settlements candidates' and 'definitely no settlements'. This process was repeated at a hierarchy of differently sized segments with a set of different threshold parameters at each level. The hierarchical network of segments consisted of three levels (Fig. 4).

The size of the segments decreases from level 3 (coarse) to level 1 (fine). The segmentation in eCognition was applied solely to the panchromatic data. The classification algorithm starts at the third level. For each segment of the newly generated class 'settlement', texture and form parameters as well as an average normalized difference vegetation index (NDVI) are calculated (Jensen 2005). The 'gray level co-occurence' (GLC) matrices that examine the spectral as well as the spatial distribution of gray values in the image form the basis for the texture calculation (Haralick et al. 1973). The method is described in detail in Tomowski et al. (2005).

A GLC matrix describes the likelihood of the transition of the gray value i to the gray value j of two neighboring pixels. For the differentiation of 'settlement' and 'not-settlement' we use the inverse distance moment (IDM) derivative from the GLC matrix. With the application of the IDM, it is possible to distinguish between heterogeneous and partially homogeneous non-settlement areas (Steinnocher 1997).
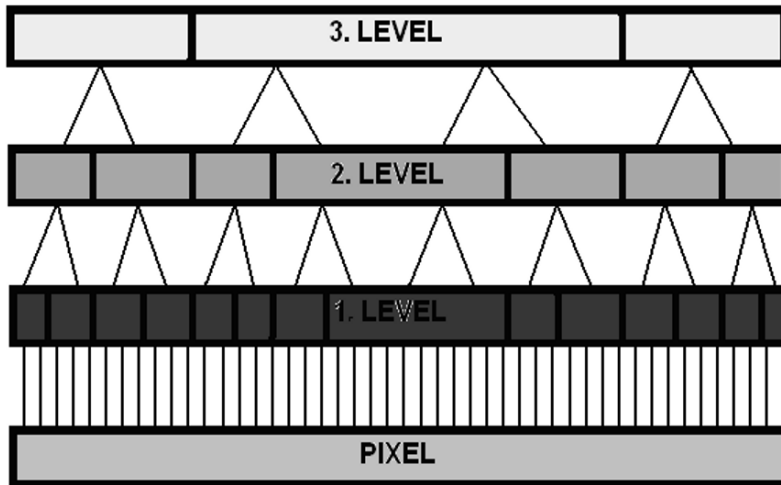
Fig. 4. Hierarchical network of segments for the decision based fusion

The next step of the method starts at segmentation level 2, in which the threshold values for the classification characteristics (texture, form and NDVI) are increased. Additionally, the classification characteristics are only calculated for the settlement areas (so-called filial segments) that are part of a non-excluding area at the third level (Ehlers et al. 2006). At the finest segmentation level 1, the classification rules are again applied but with highest restriction parameters. Finally, the settlement segments are merged and cleaned by automated filter procedures to eliminate small remaining agriculture segments and to include urban parks and lakes in the settlement areas. The result is a binary mask containing the classes 'settlement' and 'non-settlement' (endlevel). More details on this algorithm can be found in Ehlers et al. (2006) and Tomowski et al. (2006).

Despite the differences between the used datasets, the results are very similar (see Fig. 5 and 6). Contiguous settlement areas (conurbation areas) are detected with a high accuracy. For both test areas the borders between

'settlement' (red) and 'non settlement' (no color) are represented with a low level of generalization (borders are not smooth).

It is evident that only a few vegetated areas such as playgrounds or parks are missing and small houses or farms outside the kernel settlements are not completely included. To analyze the final accuracy, settlement areas were manually digitized and compared to the results of the hierarchical processing at each level (Table 1). For both combinations, results are almost identical and exceed 95% user accuracy at the final level.

Fig. 5. Binary mask for the SPOT 5 study site (settlement in red)

Table 1: Users' accuracy for the detection of settlement areas.

Hierarchical Level	SPOT-5 & Landsat ETM	KOMPSAT 1 & Aster
3	13.6%	45.3%
2	70.0%	84.2%
1	86.9%	95.0%
Final	96.3%	97.3%

Fig. 6. Binary mask for the KOMPSAT 1 study site (settlement in red)

3 GIS and NDVI based image enhancement

Image enhancement techniques are usually applied to remote sensing data to improve the appearance of an image for human visual analysis. Enhancement methods range from simple contrast stretch techniques to filtering and image transforms. Image enhancement techniques, although normally not required for automated analysis techniques, have regained a significant interest in recent years. Applications such as virtual environments or battlefield simulations require specific enhancement techniques to create 'real life' environments or to process images in near real time.

Problems with standard fast enhancement techniques such as contrast stretch or histogram equalization are that they are usually optimized for whole images and might not prove appropriate for selected features. This affects especially coastal areas that contain land, water and beach classes. Using global image enhancement techniques, the image will be transformed in a way that would produce a compromise for the different classes. Water is usually dark (especially in CIR display), beach will be very bright with little discernible structure (similar for urban classes), and other land classes (e.g. vegetation) will not make use of the full possible range of spectral values. Also, different features might require different band combinations for optimum display. This cannot be done using conventional enhancement and display strategies. Water, for example, may re-

veal more information in a true color display whereas vegetation requires a false color infrared approach. The indicated problems will only get worse with the availability of hyperspectral data where possible combinations of narrow bandwidth spectral bands can differ for land and water features.

The proposed methods make use of existing GIS information, if available, and/or image preprocessing such as NDVI calculations. Using this approach, it is possible to design a procedure for completely automated image enhancement that works in an optimized way for the selected features.

The study site is located southeast of Jacksonville, North Carolina, USA. It presents one of the largest US Marines sites for which an extensive amount of ground truth, GIS, and remote sensing data is available. The datasets consisted of Landsat, SPOT, IKONOS and Qickbird images as well as GIS landuse/landcover data in shape format (Ehlers et al. 2004, Fig. 7).

Fig. 7. IKONOS multispectral image (2048 x 2048 subset) of the Camp Lejeune study site overlaid with vector GIS information

3.1 Methodology

Selected stretching especially for regions of low contrast is nothing new in the analysis of remotely sensed data. Usually, this is done interactively by the analyst either by selecting a box or digitizing a certain area of interest in the image. This area is then enhanced using standard image processing

techniques (e.g., histogram equalization or linear contrast stretch). The subset is then displayed separately to highlight certain features that would have been impossible to discern in a global enhancement mode.

The goal of this study was to develop automated procedures for feature based image enhancement techniques for rapid display purposes, especially of high resolution remote sensing images (Ehlers 2004). Feature based enhancement means that different feature classes in the image require different procedures for optimum display. The procedures do not only encompass locally varying enhancement techniques such as histogram equalization or contrast stretch but also the selection of different spectral bands. The image class water, for example, may be best displayed in a true color mode whereas for the feature class vegetation a false color infrared display is more appropriate. It is envisioned that this technique could be implemented in a near–realtime environment making use of a priori information.

There are two main sources for this kind of information: (a) storage of a priori knowledge in a GIS, and (b) context based image information that can be extracted through a segmentation process. Both techniques can also be applied for optimum feature class selection. For many areas in the world, there exists a wealth of a priori information in existing spatial databases, digital maps or previous analyses of remotely sensed data. Usually, this type of information is stored in a raster or vector based GIS. With the progress in the integration of remote sensing and GIS software, many commercial systems allow the simultaneous display and use of GIS and image layers. Usually, GIS vector layers have to be converted to raster data for a joint GIS/image analysis.

The case study for our research was conducted in an integrated ArcGIS/ERDAS environment. The developed procedure, however, is system independent and can work in any integrated GIS/remote sensing environment. The procedure consists of five steps involving either GIS based or image context based masking (Fig. 8). Despite the fact that no general theory for segment based fusion exists the procedure outlined in Fig. 8 can be viewed as a general principle, the appropriate segmentation and GIS laxer selection as well as the enhancement procedures have to be chosen in accordance with the actual application.
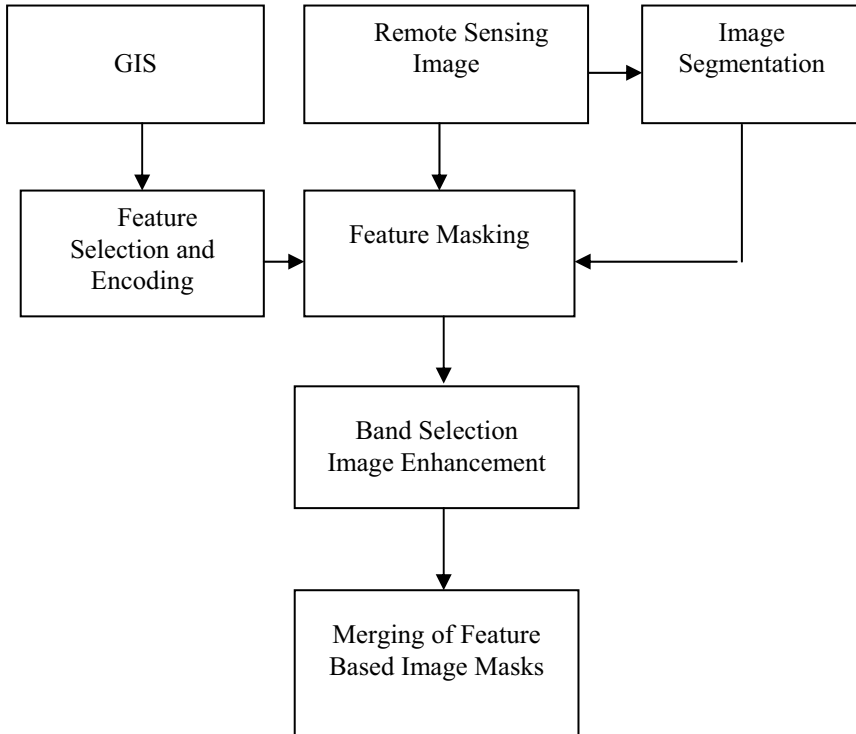
Fig. 8. General concept for GIS and context based image enhancement

3.2 GIS based enhancement

First, all image and GIS layers have to be registered to a common coordinate system, e.g., UTM. GIS layers should be displayed as polygons or raster boundaries overlaid on the remotely sensed image to check for inaccuracies or geometric and semantic inconsistencies of the data sources. In a second step, GIS information to be used as feature masks for local enhancement is analyzed and merged into meaningful classes. If, for example, vegetation classes are to be evaluated, all non-vegetation classes can be recoded into one mask. Other GIS masks that can be used for local image enhancement may separate water from land or built-up areas from open fields. The GIS layers can be overlaid on the image data for visual inspection. With this, editing can be performed if required (Fig. 9).

The third step is the creation of separate image layers that are based on the selected feature classes. After recoding, the GIS layers form 0/1 input masks (0 = area outside the feature class, 1 = area inside the feature class)

to segment the image into independent layers (Fig. 10). Each spectral band of the image is multiplied with the individual GIS masks to form separate multispectral image layers with have the original nonzero pixel values only inside the selected GIS masks. The last image to be created contains the complement mask to all selected feature classes. Using this procedure, it is assured that for each pixel location only one of the image layers contains the original image value. All the others will have a zero value at this location. Using the 'union' operator, a simple overlay of all separate image layers recreates the original image.

Fig. 9. Selected and recoded GIS classes for 'Water/Ocean', 'Water/River', 'Beach', and 'Open Field/Roads/Built-Up' overlaid on multispectral Ikonos data

In a fourth step, each layer is processed separately. This step does include the selection of an appropriate enhancement algorithm and the choice of suitable bands for display or printing purposes. In our study, we worked with 4-band remote sensing data. This step, however, will become more important if it involves hyperspectral images. For example, water information is usually displayed with a higher lever of detail if the blue band is included. Narrow band widths will make it possible to select spectral bands that depict physical phenomena such as turbidity or sediment content. Vegetation, on the other hand, is displayed best in standard false color infrared band combination due to the high reflectance in the near infrared domain.

The user can interactively be involved in this process or can leave the display and contrast enhancement to the default options. The default dis-

play options are established based on user experience and standard image processing literature. For water classes, the standard bands to be displayed are near infrared, green and blue (or for Ikonos and Quickbird bands 4, 2, 1). For all other areas, the standard display is near infrared, red, green (or bands 4, 3, 2, respectively). For image enhancement, we selected a contrast stretch based on $\pm 2.5\sigma$. This means that the digital numbers (DNs) for each band are stretched so that the values $[\mu - 2.5\sigma, \mu + 2.5\sigma]$ are mapped to $[0,255]$ (μ being the mean value of the input image band). Values outside the selected range are mapped to 0 and 255, respectively. This contrast stretch usually produces better visual results than the histogram equalization process with often too saturated areas of less discernible level of detail.
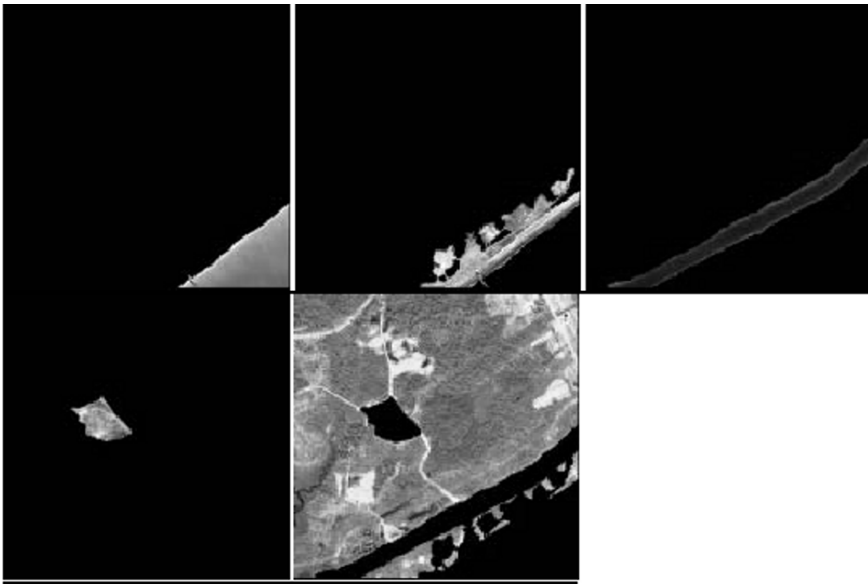
Fig. 10. Separate image layers for the selected GIS classes 'Water/Ocean' (top left), 'Beach' (top center), 'Water/River' (top right), 'Open Field' (bottom left) and the complement class (mostly vegetation) (bottom center)

The last step involves merging of the separate image layers into a single image file using standard GIS overlay procedures. As the image masks do not overlap, the procedure is based on a simple union process. Fig. 11 shows the result of the GIS based local image enhancement process compared to the standard full image enhancement option. The GIS layers selected from the database were 'Water/Ocean', 'Water/River', 'Beach',

'Open Field/Roads/Built-Up', and 'Vegetation'. The GIS based enhanced image shows more detail in all parts of the study area. There are almost no areas that are too bright or too dark to convey any information as is the case in the globally enhanced image which represents a compromise over the different spectral reflectance distribution for the image (see Fig. 12 for a close-up).

Fig. 11. GIS based enhancement of the Ikonos image.

The process can be modeled in a flow chart or script language environment and thus be applied to other images and geographic regions. It has to be noted that the choice of suitable feature classes is still an interactive process and has to be performed by the user.
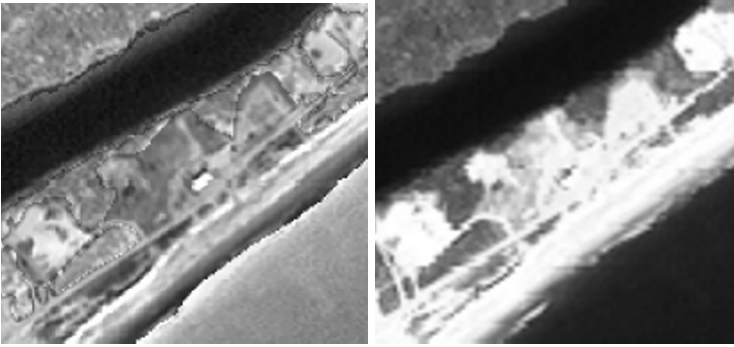
Fig. 12. The subset of the GIS enhanced image (left) shows a higher level of detail compared to the globally enhanced image (right) (contrast stretch with $\pm 2.5\sigma$)

3.3 NDVI based enhancement

Often, there is not enough a priori information for a GIS based feature selection process or the information is not accurate and/or outdated. In this case, the information contained in the image itself is the most reliable source for feature based enhancement. The image has to be segmented into meaningful features classes which are again mutually exclusive and can be used as masks to create independent image layers. When vegetation is present in the images, we can make use of indices that are sensitive to the near-infrared band of the image sensor. We selected the NDVI as a means for segmentation. Using this index, the difference between vegetation and non-vegetation is emphasized. Reflectance values for vegetation have their maximum in the near infrared and a minimum in the red spectral domain. High values of the NDVI indicate lush vegetation, values around 0 non-vegetated land areas and negative values are usually associated with water. Fig. 13 and 14 show the NDVI as gray value display and the NDVI histogram for the study site. For this investigation, we used a Quickbird image of the same area.
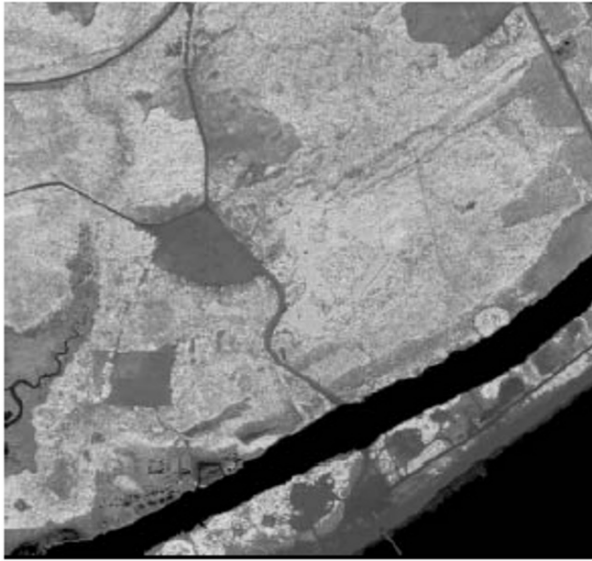
Fig. 13. NDVI for the Quickbird image of the study site. Bright values indicate high level of vegetation, intermediate values open fields, and dark values water
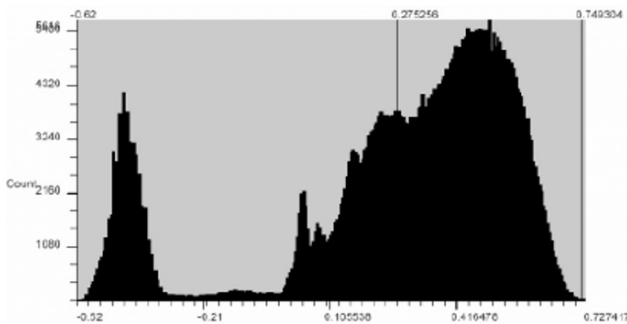
Fig. 14. Histogram of the NDVI image. The separations between potential classes (local minima) are clearly visible

The only interaction that is still required is the NDVI threshold selection for the required feature classes. Although some procedures exist to develop automated threshold selection based on local minima detection, we decided to use an interactive process so that the user can immediately see the selected classes on top of the image data. Mistakes and erroneous thresholds can be interactively corrected. Once the selected NDVI classes have been

verified by visual analysis, the image is separated into independent layers and processed similar to the previous process based on the GIS input.

Table 2 presents the NDVI thresholds and the respected feature classes. Fig. 15 shows the NDVI masks and Fig. 16 the result of the feature based enhancement process. The steps are the same as with the GIS based enhancement. It should be noted that for the water areas, another band combination (3, 2, 1) was employed for better feature separation.

Table 2: Selected enhancement classes with NDVI values.

Class	NDVI Value
Water	$NDVI \leq -0.12$
Open/Beach	$-0.12 < NDVI \leq 0.00$
Open/Inland/Built-up	$0.00 < NDVI \leq 0.19$
Vegetation	$0.19 < NDVI$
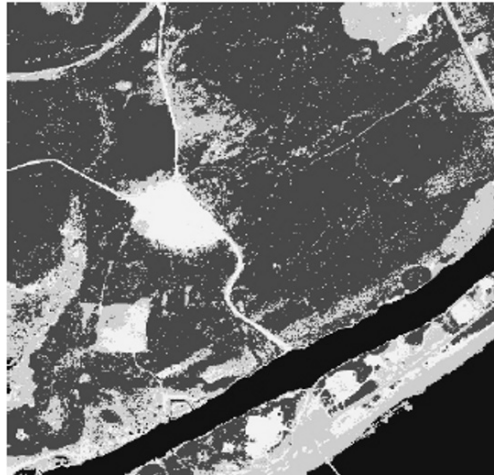
Fig. 15. Selected NDVI classes (pseudo color coded).

For a better comparison, Fig. 17 presents the same subset as shown in Fig. 12. Again, the level of detail demonstrates the superiority of the local enhancement procedure.

Fig. 16. NDVI enhanced Quickbird image
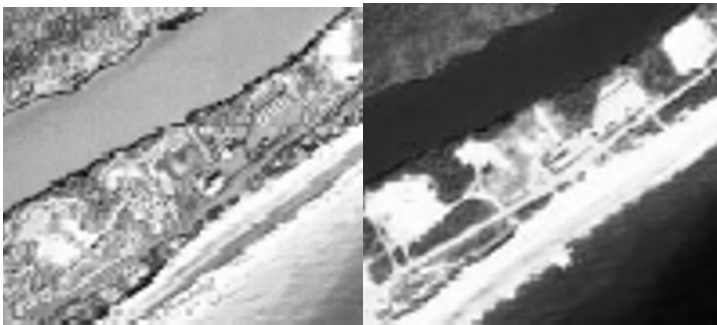
Fig. 17. The subset of the NDVI enhanced image (left) shows a higher level of detail compared to the globally enhanced image (right)

The result of the NDVI based enhancement seems almost better than the one that is based on GIS information. The reason is that the selected contrast enhancement is based on the underlying image information. At large magnifications, however, discontinuities in the selected classes become visible. In contrast to GIS feature classes, NDVI classes are not contiguous and may contain single pixels and small pixel groups that are differently enhanced than their neighbors. This can result in image noise in certain areas. Image processing such as filtering or blow and shrink operations may be employed to create more contiguous image masks. At standard resolu-

tions, however, this method shows results that prove the validity of the presented approach.

4 Conclusion

Several aspects of image fusion and GIS/remote sensing integration were investigated in this paper. All presented fusion techniques use the benefits of a combination of high spatial and high spectral resolution remote sensing. For the detection of settlement areas, we employed a decision based fusion technique for a set of panchromatic high-resolution and multispectral low resolution images. Through a hierarchical segment based approach it was possible to improve the classification results at each level. Furthermore, this procedure works equally well with different multisensor satellite data without altering the procedure or the employed parameters steps.

Information stored in GIS databases can be used for rapid image enhancement with real-time potential. GIS selected image classes were separately enhanced and fused at the final steps. If no GIS information is available, use can be made of segments such as those that are produced by NDVI values. Both procedures work well for the display of multispectral images. As individual band selection can be incorporated in this enhancement process, the extension to rapid hyperspectral image display is possible. Known optimum band selections can be combined with spectral enhancement in this procedure. The method can also be automated to a large degree using a flow chart environment or scripting language. With more investigations in the future, some of the interactive steps can be replaced by default values.

In comparison to pixel based classification procedures (like the maximum likelihood method) it is evident that the introduced feature ('cookie cutter') and decision based fusion techniques are significant improvements for the design of future automated processing modules. Through the adoption of object-based image processing methods and data fusion techniques it is possible to avoid the salt-and-pepper effect of pixel based analyses and to enhance the classification accuracies. In our opinion, a feature and/or decision based fusion seems to be the most promising technique for the improvements of classification accuracy.

References

Apel, D., Henckel, D. (1995): Flächen sparen, Verkehr reduzieren – Möglichkeiten zur Steuerung der Siedlungs- und Verkehrsentwicklung, in: Deutsches Institut für Urbanistik (Eds.) Difu-Beiträge zur Stadtentwicklung, Berlin: 29-40

Edwards, G., Jeansoulin, R. (2004): Data fusion - from a logic perspective with a view to implementation. International Journal of Geographical Information Science, 18(4): 303-307.

Ehlers, M. (2004): Remote Sensing for GIS applications: New sensors and analysis methods, in: Ehlers, M., Kaufmann, H.J., Michel, U. (Eds.) Remote Sensing for Environmental Monitoring, GIS Applications, and Geology III, Proceedings of SPIE Vol. 5239, Bellingham, WA: 1-13.

Ehlers, M., Klonus, S. (2004): Erhalt der spektralen Charakteristika bei der Bildfusion durch FFT basierte Filterung. Photogrammetrie-Fernerkundung-Geoinformation, 6: 495-506.

Ehlers, M., Welch, R., Ling, Y. (2004): GIS and context based image enhancement, Proceedings of the XXth International Congress of ISPRS, Istanbul, Turkey, IAPRS XXXV/B4: 397-402.

Ehlers, M., Michel, U., Bohmann, G., Tomowski, D. (2006): Decision based data fusion techniques for the analysis of settlement areas from multisensor remote sensing data, Proceedings of ASPRS 2006 Annual Convention "Prospecting for Geospatial Integration", Reno, Nevada (CD Publication): 8 pp.

Ehlers, M., Klonus, S., Astrand, P.J. (2007): Spectral Change Analysis for Multi-Date Multi-Sensor Image Fusion (in preparation)

Haralick, R.M., Shanmugam, K., Dinstein, I. (1973): Textural features for image classification, IEEE Transactions on Systems, Man, and Cybernetics, SMC-3: 610-621.

Jensen, J.R., 2005, Introductory Digital Image Processing: A Remote Sensing Perspective, Upper Saddle River, NY, Prentice Hall.

Pohl, C., van Genderen, J. (1998): Multisensor image fusion in remote sensing: concepts, methods and applications. International Journal of Remote Sensing, 19: 823–854.

Steinnocher, K. (1997): Texturanalyse zur Detektion von Siedlungsgebieten in hochauflösenden panchromatischen Satellitenbilddaten. Salzburger Geographische Materialien 26: 143-152.

Tomowski, D., Ehlers, M., Michel, U., Bohmann, G. (2005): Objektorientierte Klassifikation von Siedlungsflächen durch multisensorale Fernerkundungsdaten, gi-reports@igf 3, pdf document (http://elib.ub.uni-osnabrueck.de/publications/ELibD131_gi-reports_igf3.pdf)

Tomowski, D., Ehlers, M., Michel, U., Bohmann, G. (2006): Decision based data fusion techniques for settlement area detection from multisensor remote sensing data, Proceedings, 1st Workshop of the EARSeL Special Interest Group Urban Remote Sensing, "Urban Remote Sensing: Challenges and Solutions", Berlin-Adlershof (CD Publication): 8 pp.

Chapter 8.2

Modelling uncertainty in high resolution remotely sensed scenes using a fuzzy logic approach

J. Schiewe[1], M. Gähler[2]

[1] HafenCity University Hamburg, Department Geomatics,
jochen.schiewe@hcu-hamburg.de

[2] German Aerospace Center (DLR), monika.gaehler@dlr.de

ABSTRACT: This contribution concentrates on the determination of thematic uncertainty after the classification process. It is shown that in this context –particularly when evaluating remotely sensed scenes showing high spatial resolution – severe problems arise due to indeterminate boundaries in both reference data and classification results. This effect mainly occurs between natural objects or are due to blurred or overlapping definitions of classes or related attributes in a given classification scheme. Based on some approaches in the literature and unsatisfactory tools in existing software packages, we propose the introduction of a new characteristic value, the Fuzzy Certainty Measure (FCM), that is able to model and quantify uncertainty as indeterminate boundaries in reference data and classification results.

1 Introduction

Remotely sensed scenes have a steadily increasing impact on the description, modelling and simulation of landscape structures and processes. Nowadays, new applications or classical applications at larger scales can be addressed due to technical developments of the sensing systems concerning their spatial, spectral and radiometric resolutions. In reaction to the

new properties of the data, appropriate processing methods have been developed and are widely applied – in particular region-based methods, i.e. segmentations, as well as follow-up object-based and eventually fuzzy classification approaches.

With this increased potential, a stronger integration of the derived geo data in binding planning and decision processes takes place. In this context, users need reliable information about the thematic and geometric uncertainty of the results that have been derived by interpreting the remotely sensed scenes. In the area of research and development, measures of the classification accuracy are also mandatory, for instance, for the evaluation of the potential of (new) sensing systems or processing methods.

Concentrating on a post classification accuracy assessment, generally well known standard methods are applied. Those compare reference data ("ground truth") and the classification result from which error matrices and related measures like overall accuracy or Kappa coefficient can be derived. By doing this, one assumes discrete boundaries between regions of a scene for which one and only one topographical object is attached and which is not subject to temporal changes. Furthermore, the reference data are assumed to be error-free, which is obviously not the case with most applications. Instead, we have to deal with some effects of fuzziness, i.e. indeterminate boundaries between neighbouring objects. These effects are even amplified with the use of spatial high resolution data like those from the above mentioned new digital airborne or satellite systems.

Section 2 will elaborate on the just indicated problems in uncertainty, determination from which the motivation arises to introduce an integrated and fuzzy uncertainty measure. Section 3 gives an overview of such measures from the literature, but also looks at the available tools within the software package *Definiens Enterprise Image Intelligence™ Suite* (formerly: *eCognition*). From this survey it can be concluded that the existing characteristic values do not fulfil all demands with respect to uncertainties in reference data as well as to fuzziness in both reference and classification results. Hence, our goal is to develop and to test a more profound methodology to determine the classification uncertainty. The resulting integrated and fuzzy measure is presented in section 4. Section 5 summarizes these results and presents recommendations for further developments.

2 Problems in uncertainty determination

When accuracy is known objectively then it can be expressed as *error*. Where it is not, the term *uncertainty* applies (Hunter and Goodchild 1993).

Thus, uncertainty covers a broader range of doubt or inconsistency and includes errors. In the following we concentrate on *thematic uncertainty* which shall be determined after the classification process. The corresponding evaluation, which assesses the quality of the input data and the classification process as such, seems to be a standard task. Quantitative methods compare reference data ("ground truth") and the classification result from which error matrices and related measures like overall, producer's and user's accuracy or Kappa coefficient can be derived. However, in the case of using spatial high resolution data, some of the general problems related to this procedure are even amplified and need even more attention compared with the use of lower resolution data. The underlying reasons, which will be briefly discussed in the following, can be grouped into geometric and semantic aspects.

From a **geometric point of view,** the smaller pixel sizes lead to the fact that a suitable reference with appropriate positional accuracy as well as little model and cartographic generalisation is more difficult to find. Furthermore, an adoption of the number and size of sample units has to take place. In particular, the conventional acquisition on a per pixel basis is not suitable anymore due to excessively small elements and neglect of the neighbourhood. In analogy to the object-based interpretation approach, a per-object sampling seems to be necessary in order to define training and test elements. However, due to missing methodology, such an approach is hardly applied in practise.

It is also well known that we have to handle indeterminate boundaries or spatial transition zones between mostly natural objects (e.g., between different forest types, or forest and meadow), which are in some cases also a function of time (e.g., the boundary between beach and water). On the other hand, we have also to consider blurred or overlapping definitions of classes or related attributes in a given classification scheme. Taking now high resolution data into account, the absolute number of pixels describing spatial transition zones – and with that the effect of fuzziness – increases.

From a **semantic point of view**, high resolution data allow for the extraction of more thematic details and object classes. With that a more complex classification scheme becomes necessary, which on the other hand inherits a greater chance of overlapping definitions of attribute value ranges. As a consequence, this may lead to errors or ambiguous assignments during the visual or automatic interpretation process. The greater number of possible classes also makes more sampling units necessary. Like with geometric properties, it is also difficult to find a suitable reference with appropriate thematic details and semantic accuracy. It has to be kept in

mind that very often a reference data set is nothing other than another classification result based on another, eventually lower resolution data set.

Finally, the spatial variance within regions representing a topographical object is increased in high resolution imagery which leads to more objects, more mixed objects (e.g., forest consists of trees, bare soil and others) and more boundaries. With the latter, the number of indeterminate boundaries, in other words the effect of fuzziness, is again increased.

Therefore due to the increasing importance of remotely sensed data with high spatial resolution on one hand, and the above described problems on the other hand, there is a significant necessity to develop uncertainty measures that consider uncertainties in reference data as well as indeterminate boundaries in both reference data and classification results.

3 Previous work

3.1 Consideration of reference accuracy

It is obvious that with the improved information content and better geometrical accuracies of new high resolution remote sensing systems, the demands for the reference data increase. Very often, however, appropriate data that match the respective criteria (e.g., Lunetta et al. 2001, Plourde and Congalton 2003, or Stehman 2004) are not available. As an example, investigations of Lunetta et al. (2001) demonstrate the huge variation of results during a visual interpretation of aerial images which are often used as a reference. The authors determine the repetition accuracy (of different operators) which is furthermore a function of the object classes under investigation. The authors conclude that "absolute ground reference is an illusive goal that is difficult (if not impossible) to achieve". Vice versa, this demonstrates the necessity for considering the inherent uncertainty in the reference data by using an **integrated** approach, particularly when using high resolution data.

3.2 Consideration of fuzzy boundaries

An approach for describing uncertainty in classification results based on probability theory is meaningful if random variations in the class categorization occur. This theoretical basis is missing if the classification shows **vague effects** (Zhang and Goodchild 2002), which can be due to several effects:

- Blurred boundaries between classes (e.g., along the forest border) make an unique allocation subjective if not impossible. Chrisman (1982) showed in an empirical study that 18% of the test area belonged to such regions.
- The categorization into discrete classes depends strongly on the underlying object model (classification scheme) which varies with different applications and is rather subjective (Zhang and Goodchild 2002). As a consequence, vague categories (such as "urban" vs. "suburban") exist. Using new sensors a stronger thematic depth can and shall be achieved so that this effect of vagueness is even strengthened.
- Because of the use of high resolution systems the spectral variance within single objects is increased, as is the number of mixed elements and non unique class categorizations (e.g., forest has to be characterized as a mixture of trees, bare ground, etc.). Wang (1990) shows that the introduction of additional mixed classes can not solve this problem.

While for the application of conventional, statistically founded methods a variety of papers exist (e.g., Thomas et al. 2003; Foody 2004), approaches for the determination of fuzziness have been considered rather rarely. One approach for modelling transition zones is to introduce the so called ε–**bands**, as defined by Blakemore (1994; cited after Ehlers and Shi 1997). Here, the different chances of a point-in-polygon-relation are described by five *qualitative* measures ("definitively in", etc.). Ehlers and Shi (1997) propose to use a probabilistic model in order to give a *quantitative* description which also allows for the combination with values of thematic uncertainty. They apply the so called S-band model of positional and thematic uncertainty values are linked by using the product rule. Other options to handle indeterminate boundaries (e.g., least squares polynomial fitting, corridor techniques, etc.) are treated by Edwards and Lowell (1996).

The application of **fuzzy set theory** for the determination of classification accuracy has been demonstrated by Fisher (2000). Wang (1990) also proposes the derivation of a fuzzy partition matrix which contains the membership of a pixel to each of the classes under consideration. Gopal and Woodcock (1994) added certainty values on a linguistic scale ("absolutely safe", etc.) to their visually classified elements. Those linguistic values can be combined using fuzzy logic theory for a better understanding of the resulting map. Similar approaches are reported by Wang and Hall (1996), Townsend (2000) or Lowell et al. (2005).

Edwards and Lowell (1996) concentrate on the definition of a membership function which in their case describes spatial uncertainties. For this purpose they introduce fuzzy transition zones whose widths are defined for

all pairs of object classes ("twins"). In this case, the corresponding zone width values had been derived from the mean deviations resulting from multiple digitizations in aerial images. The authors also found that not only the thematic class memberships, but also the area sizes of the polygons under consideration, have a significant influence on the width of the transition zone (the smaller the area, the larger the fuzziness). Similar approaches for combined thematic and geometric uncertainties are treated, for example, by Wang and Hall (1996), Binaghi et al. (1999), or Townsend (2000).

3.3 Implementation under Definiens Enterprise Image Intelligence™ Suite

It is not surprising that due to the relatively sparse theoretical basis for using fuzzy logic methods for the determination of classification accuracy, only limited implementations in commercial software products can be observed. As an example, the software package *Definiens Enterprise Image Intelligence™ Suite* (formerly: *eCognition*, Definiens 2006) offers conventional methods for the determination of classification accuracy by using reference data and the classification result (here called „error matrix based on TTA mask", or "error matrix based on samples"). Besides that we also find the option for an evaluation of classification results based on fuzzy set methods. The concept of the „advanced classification evaluation based on fuzzy classification" assumes that the larger the deviation of all membership values (for all classes, for one pixel or segment), the more uncertain the classification is. With that, uncertainties in the classification scheme (and indirectly also measurement errors) can be addressed. Meanwhile the acquisition and processing methods themselves cannot be assessed, because uncertainties of the reference data are not considered.

In this context, the following characteristic values, which are only derived from the classification result (and the corresponding membership values), can be taken into account with that system:

- „Classification stability": For each pixel the difference between the largest and second largest membership values is computed and summed for the entire class (or even the entire scene) – in the literature also known as *ambiguity*.
- „Best classification result": This gives the visual representation of the corresponding largest membership value for each pixel. The mean value of all membership values can be interpreted as an indicator of the total uncertainty.

- For each pixel one can determine the standard deviation of the membership values which again can be summed for the entire scene.

4 Fuzzy certainty measure

Based on the above outlined problems and the current state of implementations as presented in the previous section, our goal is to develop and to test a more profound methodology to determine the *a posteriori* classification accuracy considering uncertainty in reference data as well as indeterminate boundaries in the reference and the classification result. We will end up with the new, so called *Fuzzy Certainty Measure (FCM)*.

In order to focus on the above mentioned aspects, our approach starts with the – theoretical – **assumptions** that the classification schemes between reference and classification are identical and that no discrepancies occur due to different pixel sizes or temporal changes. Furthermore, we assume that an appropriate sampling procedure has been taken into account.

In order to demonstrate the overall process as described above, we have applied our proposed method to a **data set** of the digital airborne camera system HRSC-AX which delivers image and elevation data in very high spatial resolution (in our case ground pixel size equals to 15 cm). The HRSC image data consist of a panchromatic nadir channel and four multi-spectral bands (blue, green, red, near infrared), each at a radiometric resolution of 8 bits. For further information about the camera system refer to Ehlers et al. (2006). Figure 1 also shows the reference data and the classification result.
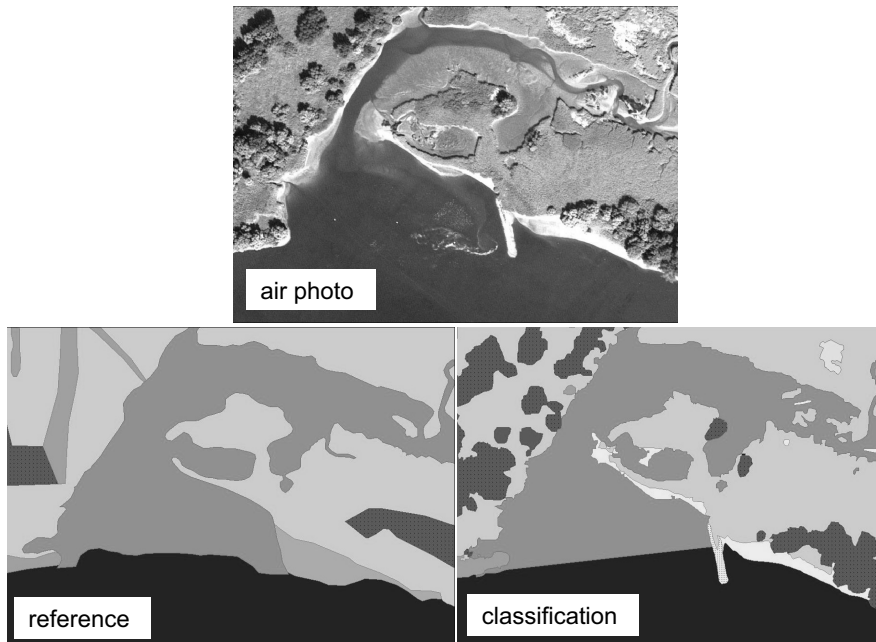
Fig. 1. Top: Given digital air photo (HRSC-AX, 15 cm ground pixel size), bottom left: reference from field survey, bottom right: image classification result (Ehlers et al. 2006). Note: Class descriptions (and with that colours) are not of importance at this stage, just the diversity of classes

In order to model the inherent fuzziness, **transition zones** are now introduced. Those are defined *a posteriori* using a buffering zone along a given boundary between two topographical objects. Based on the investigations by Edwards and Lowell (1996) the width of this zone depends on the combination of objects (e.g., the transition zone between forest and meadow is obviously larger than those between building and road) as well as on the size of the object areas (figure 2). Within these zones a **membership function** (in our case presently a linear function) is applied perpendicular to the boundary for each class *c*. This procedure is performed for both the classification result (leading to membership values $\mu_{CLASS}(c)$) and the reference data ($\mu_{REF}(c)$). Figure 3 demonstrates the important additional value of the fuzzy approach: While a crisp method would have evaluated the classification result as clearly wrong, now the realistic chance of a correct categorization is no longer neglected.
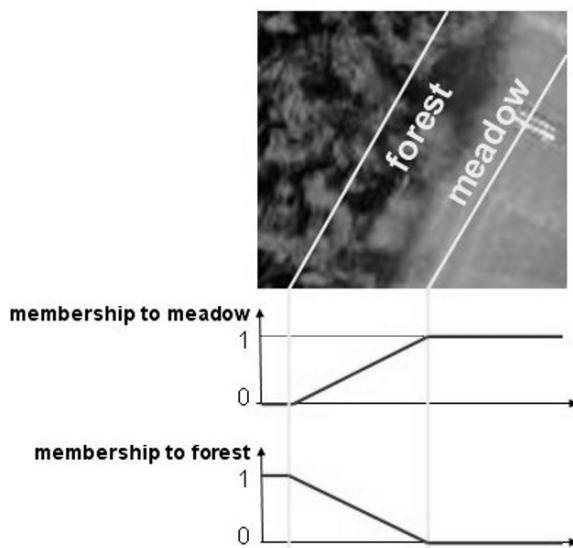
Fig. 2. Principle of building transition zones and applying membership functions to a pair of neighbored topographical objects
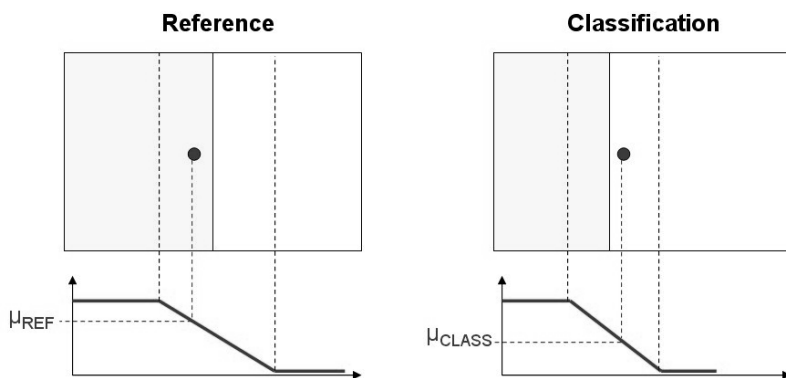
Fig. 3. Principle of deriving Fuzzy Certainty Measure FCM: Building transition zones (dashed lines) in reference (left) and classification result (right), determining membership values of a selected location to shaded class in reference (μ_{REF}) and classification (μ_{CLASS})

In order to derive one or more characteristic values, our approach consults the respective membership values for the same spatial elements (i.e., pixels

or regions) in the reference as well as in the classification result. Those are compared separately for each topographical class and for those elements for which a possibility of existence (or membership values larger than 0, respectively) is present in reference and classification. The resulting **Fuzzy Certainty Measure FCM(c) per class c** is determined as follows:

$$FCM(c) = 1 - \frac{1}{n}\sum_{i=1}^{n} |\mu_{i,REF}(c) - \mu_{i,CLASS}(c)| \tag{1}$$

$$\forall i \,|\, \mu_{i,REF} > 0 \wedge \mu_{i,CLASS} > 0$$

where:

$\mu_{REF}(c)$: membership value of a pixel or region for class c in reference data
$\mu_{CLASS}(c)$: membership value of a pixel or region for class c in classification result
n: number of pixels or regions under consideration

The FCM(c) values vary between 0 and 1 – the larger the coincidence between reference and classification is, the larger the FCM(c) value becomes. Figure 4 visualizes the derivation of the FCM for a selected class within our test data set.
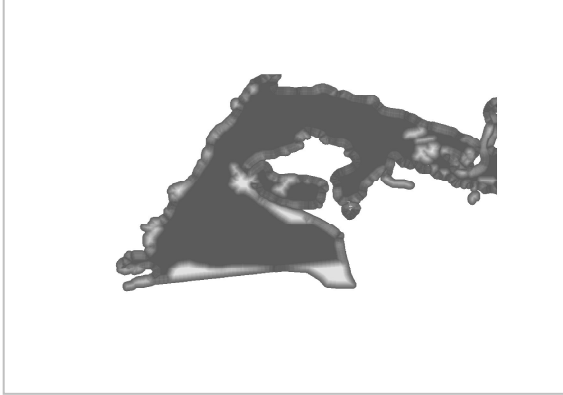
Fig. 4. Visualization of derived FMC values for class KPS (tidal creek / tideland) (compare to figure 1) – the darker the area, the higher the respective value (white = 0)

For a more thorough description of the classification uncertainty, the confusion between different classes (in the following named A and B) also is of interest. When introducing fuzzy membership values, the major difference to a conventional confusion matrix based on crisp boundaries is

that the membership areas (i.e., those regions for which the membership values are larger than 0) can overlap each other so that their class-specific values generally sum up to a value larger than 100%. And while for the computation of off-diagonal elements for conventional confusion matrices the number of misclassified elements is counted, in the fuzzy case it makes no sense just to compare the membership values for class A in the reference and class B in the classification results because the resulting difference might be correct in reality. Instead of that we consider the change in ambiguity between classes A and B in the reference compared to that in the classification as follows:

$$FMC(c_{AB}) =$$

$$\frac{1}{2}\sum_{i=1}^{n} |[(\mu_{i,REF}(c_A) - \mu_{i,REF}(c_B)) - (\mu_{i,CLASS}(c_A) - \mu_{i,CLASS}(c_B))]| \qquad (2)$$

Also for this confusion measure (with values ranging from 0 to 1), it holds that the larger the confusion is, the larger the $FCM(c_{AB})$ value becomes. Table 1 summarizes the certainty measures for the given example (diagonal elements represent $FMC(c)$ values, off-diagonal elements $FCM(c_{AB})$ values).

Table 1. Fuzzy Certainty Measures for object classes under consideration (class names according to specific object catalogue, Ehlers et al. 2006).

	Classes		Reference				
			KPS	BAT	FWR	FZT	WWT
Classi-fication	Tidal Creek / Tideland	KPS	0.90	0.05	0.09	0.05	0.06
	Shrubs (Salix)	BAT		0.80	0.12	0.04	0.30
	Reed (Phragmites)	FWR			0.81	0.07	0.23
	Tidal River	FZT				0.93	0.05
	Willow Forest, Salix	WWT					0.66

So far, with $FCM(c)$ and $FCM(c_{AB})$, respectively, the certainty of all elements belonging to a certain class in a given scene is computed at once. However, from this the uncertainty of individual objects, as well as variations or outliers among different objects of the same class cannot be evaluated. Hence, an *object-specific* value is desired. There exist already some

definitions by Zhang et al. (2005) for per-object measures (like: thematic, location, or shape similarity) based on *crisp* boundaries. On the other hand, it is no problem to transfer the above introduced FCM to an **object-specific FCM (OFCM).** If the data model does not support a distinction of such individual objects, the necessary separation is performed either by introducing α-cuts or by applying a spatial segmentation (i.e., summing up membership values μ only as long as there is a N8-neighbour with $\mu>0$).

5 Summary and future work

Various basic research work (e.g., Hunter and Goodchild 1993, 1995; Ehlers and Shi 1997; Congalton and Green 1999) has pointed out that it is not possible to define a generally valid or optimal model for determining the classification accuracy based on remotely sensed scenes. In fact, a variety of parameters such as available data sources, systems or processes which shall be described, user demands, etc., have to be taken into account for every specific case.

In this overall context, our contribution concentrates on thematic uncertainty which shall be determined after the classification process. Here, we addressed the problems of indeterminate boundaries in both reference and classification results, which occur between mostly natural objects or are due to blurred or overlapping definitions of classes or related attributes in a given classification scheme. These problems are even amplified with the use of remotely sensed data showing high spatial resolution as given with the new digital airborne or spaceborne systems.

We propose the introduction of a new characteristic value, the *Fuzzy Certainty Measure (FCM)* which considers the influence of the reference data and gives an indication of the quality of the classification procedure as such. With that, comparisons between different classifications (with respect to different methods, time stamps etc.) can be evaluated more reliably. The procedure can be characterized as flexible and quite simple to apply. In summary, it can be expected that further developing and applying these methods a more complete and thorough modeling of uncertainty can be accomplished.

Our future work is concerned with a sensitivity analysis of the parameters (in particular with the width of transition zones based on object class combination and area sizes). Furthermore, empirical investigations will be performed for the combination of fuzzy with additional probabilistic measures. Finally, the extension towards a change analysis will be taken

into consideration by introducing thresholds for FCM values for the classifications of different time stamps.

References

Binaghi E, Brivio PA, Ghezzi P, Rampini A (1999) A fuzzy set-based accuracy assessment of soft classification. Pattern Recognition Letters, (20): 935-948.

Blakemore M (1984): Generalization and error in spatial data bases. Cartographica, (21):131-139.

Chrisman NR (1982) A Theory of Cartographic Error and Its Measurement in Digital Databases. AUTOCARTO, (5): 159-168.

Congalton R, Green K (1999) Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. CRC/Lewis Press, Boca Raton, FL. 137 p.

Definiens (2006) www.definiens.com.

Edwards G, Lowell KE (1996) Modeling Uncertainty in Photointerpreted Boundaries. Photogrammetric Engineering & Remote Sensing. 62(4): 337-391.

Ehlers M, Shi W (1997) Error Modelling for Integrated GIS. Cartographica. 33(1): 11-21.

Ehlers M, Gähler M, Janowsky R (2006) Automated Techniques for Environmental Monitoring and Change Analyses for Ultra High-resolution Remote Sensing Data. Photogrammetric Engineering & Remote Sensing. 72(7): 835-844.

Fisher P (2000) Sorites paradox and vague geographies. Fuzzy Sets and Systems, (113): 7-18.

Foody GM (2004) Thematic Map Comparison: Evaluating the Statistical Significance of Differences in Classification Accuracy. Photogrammetric Engineering and Remote Sensing. 70(5): 627-633.

Gopal S, Woodcock C (1994) Theory and methods for Accuracy Assessment of Thematic Maps Using Fuzzy Sets. Photogrammetric Engineering and Remote Sensing. 60(2): 181-188.

Hunter GJ, Goodchild MF (1993) Mapping Uncertainty in Spatial Databases, Putting Theory into Practise. Journal of Urban and Regional Information Systems Association. 5: 55-62.

Hunter GJ, Goodchild, MF (1995) Dealing with Error in Spatial Databases: A Simple Case Study. Photogrammetric Engineering & Remote Sensing. 61(5): 529-537.

Lowell K, Richards G, Woodgate P, Jones S, Buxton L (2005) Fuzzy Reliability Assessment of Multi-Period Land-cover Change Maps. Photogrammetric Engineering & Remote Sensing. 71(8): 939-945.

Lunetta RJ, Iiames J, Knight R, Congalton R, Mace T (2001) An assessment of reference data variability using a "virtual field reference database". Photogrammetric Engineering & Remote Sensing. 67 (6): 707-715.

Plourde L, Congalton R (2003) Sampling method and sample placement: How do they affect the accuracy of remotely sensed maps? Photogrammetric Engineering & Remote Sensing. 69(3): 289-297.

Stehman SV (2004) A Critical Evaluation of the Normalized Error Matrix in Map Accuracy Assessment. Photogrammetric Engineering and Remote Sensing. 70(6): 743-751.

Thomas N, Hendrix C, Congalton R (2003) A comparison of urban mapping methods using high-resolution digital imagery. Photogrammetric Engineering and Remote Sensing. 69(9): 963-972.

Townsend PA (2000) A Quantitative Fuzzy Approach to Assess Mapped Vegetation Classifications for Ecological Applications. Remote Sensing of Environment. 72: 253-267.

Wang F (1990) Improving Remote Sensing Image Analysis through Fuzzy Information Representation. Photogrammetric Engineering & Remote Sensing. 56(8): 1163-1169.

Wang F, Hall GB (1996) Fuzzy representation of geographical boundaries in GIS. Int. J. Geographical Information Systems. 10(5): 573-590.

Zhang J, Goodchild MF (2002) Uncertainties in Geographical Information, Taylor and Francis, London, UK.

Zhang Q, Molenaar M, Tempfli K, Shi W (2005) Quality assessment for geospatial objects derived from remotely sensed data. International Journal of Remote Sensing, 26(14): 2953-2974.

Chapter 8.3

Assessing image segmentation quality – concepts, methods and application

M. Neubert, H. Herold, G. Meinel

Leibniz Institute of Ecological and Regional Development (IOER), Weberplatz 1, 01217 Dresden, Germany, (m.neubert, h.herold, g.meinel)@ioer.de

KEYWORDS: Comparison, evaluation methods, software, remote sensing, IKONOS

ABSTRACT: Image segmentation is a crucial step within the remote sensing information retrieval chain. As a step prior classification the quality of the segmentation result is of fundamental significance. This contribution gives an overview of existing methods for the evaluation of image segmentation quality. Furthermore, seven recent programs for remote sensing imagery are introduced and their results based on very high resolution IKONOS data are evaluated using an empirical discrepancy method.

1 Introduction and related work

The qualitative and quantitative assessment of segmentation results is very important for further image analysis as well as for choosing the appropriate approach for a given segmentation task. Since segmentation is a processing step prior classification, the accuracy of the recognition and classification process of image objects is significantly affected. Thus, this research is also related to the topic of object-based accuracy assessment (see Radoux and Defourny, chapter 2.8, and Schöpfer et al., chapter 8.4).

Evaluation studies either intend to compare various segmentation approaches (e.g. Estrada and Jepson 2005) or different parameterisations of one algorithm (e.g. Palus and Kotyczka 2001). Only very few studies employ their evaluation on remote sensing data, e.g. Carleer et al. (2005),

Karantzalos and Argialas (2003). Mostly natural colour images or artificially generated images are used.

Similar to the segmentation theory itself, there is no established standard procedure for the evaluation of its results. In literature there exists a multitude of very different approaches. A general classification of evaluation methods has been proposed by Zhang (1996), categorising three variants: analytic methods, empirical goodness methods, and empirical discrepancy methods. In recent studies, empirical goodness methods are also referred to as unsupervised evaluation methods, empirical discrepancy methods are denoted as supervised or stand-alone evaluation methods (e.g. Zhang et al. 2005). While most existing approaches are supervised methods using discrepancy measures between a reference and the segmentation, recently much effort is put into the development of empirical goodness methods, which do not require any reference (a priori knowledge). However, when comparing different approaches, these methods immanently show a strong bias towards a particular algorithm (Everingham et al. 2002). For this reason an empirical discrepancy method using the relative ultimate measurement accuracy (Zhang 1996) has been applied in this evaluation.

2 Evaluated segmentation software

2.1 Overview

There is a large variety of implemented segmentation algorithms using very different concepts. They are distributed commercially, are freely available for scientific use or are in-house developments of research organisations. For this evaluation only approaches were considered that are able to perform a full (so-called multi-region) image segmentation in an operational way. Furthermore, the choice of approaches was based on the suitability to segment remote sensing imagery. Programs which only perform selective image segmentation are not considered in this study.

In addition to the results presented in Meinel and Neubert (2004) and Neubert et al. (2006), the following new algorithms and programs or new releases of previously tested software respectively were included in the comparison: EDISON (Edge Detection and Image Segmentation system), EWS 1.0 (Extended WaterShed), Definiens Developer 7.0, HalconSEG 1.1, InfoPACK 2.0, RHSEG 1.30 (Recursive Hierarchical Segmentation) and SCRM (Size-Constrained Region Merging) (see table 1 for details).

Table 1. Outline of evaluated segmentation software

Software	EDISON	EWS 1.0	Definiens Developer 7	HalconSEG 1.1	InfoPack 2.0	RHSEG 1.30	SCRM
Developer	Rutgers Univ., Robust Image Understanding Lab	P. Li and B. Xiao	Definiens AG	TU Munich/ IOER Dresden	InfoSAR Ltd.	NASA, Goddard Space Flight Center	G. Castilla
Website	www.caip.rutgers.edu/riul/	-	www.definiens.com	www9.informatik.tu-muenchen.de/research/	www.infosar.co.uk	http://ipp.gsfc.nasa.gov/RHSEG/	homepages.ucalgary.ca/~gcastill/SCRM.htm
Algorithm	Mean shift based image segmenter	Multi-channel watershed transformation	Multiresolution segmentation	Hybrid (edge/region oriented)	Simulated annealing	Hierarchical region growing	Watershed and region merging
Field of application	Colour images	Remote sensing	Remote sensing	Colour images, mobile systems	Remote sensing, esp. radar data	Remote sensing	Remote sensing, assisted photo-interpretation
Fundamental reference	Comanicu and Meer 2002	Li and Xiao (in press)	Baatz and Schäpe 2000	Lanser 1993, Herold 2005	Cook et al. 1996	Tilton 2003	Castilla 2003, Castilla et al. (in press)
State of development	04/2003	02/2006	07/2007	03/2007	06/2006	07/2007	02/2007
Operating system	Win, Linux, Unix	Win	Win	Win, Linux, Unix	Linux, Win	Win, Unix (Solaris)	Win, Linux, Unix
System environment	Stand-alone	Stand-alone	Stand-alone	HALCON 8.0	Stand-alone	Stand-alone	IDL, IDL-VM, ENVI
Number of parameters	3	2	3	5 (3)[a]	17(6)[a, b]	16 (4)[a]	4

Table 1. (cont.)

Ca. runtime [min][c, d]	< 5	3 - 5	1	< 2	10	3 - 5	< 2
Reproducibility[e]	No	No	Yes	No	No	No	No
Classification support	No	Yes (Object-based classification via overlay)	Yes (Fuzzy Logic, Near. Neighbour)	No	Yes (Maximum Likelihood)	Yes (HSEG Viewer)	Polygons with attributes
Max. image size [ca. Pixel]	3,000 x 3,000	No limitations	No limitations	5,000 x 5,000	No limitations	8,000 x 8,000	2,000 x 2,000 (IDL limitation)
Max. bit depth [bit]	16	16	32	8	64	16	64
Input formats	Raster (TIFF, BMP, PNG, JPEG)	Raster (ENVI)	Raster, Vector (Shapefile)	Raster (TIFF, RAW)	NetCDF[f]	Raster (RAW)	Raster (Geo-TIFF, JPEG, ENVI)
Vector output format	No (external conversion)	No (external conversion)	Shapefile	ASCII (GEN)	No (external conversion)	No (external conversion)	Shapefile
Use of external data	No	No	Yes	No	Yes	No	No
Availability	Freeware	On request	Commercial	On request	Commercial	Evaluation copy	On request (pre-commercial)

[a] Parenthessised number: especially relevant segmentation parameters. [b] Software is able to estimate defaults from the data. [c] Specification heavily depends on system resources. [d] Specifications for the used imagery (2,000 x 2,000 Pixel). [e] When image extend is modified. [f] Convertible from divers raster formats.

2.2 Optimised segmentation algorithm HalconSEG

While all algorithms were tested in their implemented version, the proposed HalconSEG was developed on the basis of the segmentation approach described in Lanser (1993). The algorithm is a combination of an edge-detection and a region-growing procedure. It was originally designed for the research on segmenting natural images on mobile devices. This approach was adapted and optimised in order to handle and process high resolution remote sensing data by adding various parameterisation opportunities (e.g. for the edge detection filters and the morphological operators), a region-merging algorithm, a hierarchical extension and a GIS-interface proposed in Herold (2005).

A further extension allows importing various manually generated segments. The algorithm automatically optimises the parameterisation to fit the result best to the reference (internal evaluation). It is a contribution to minimise the time consuming process to find optimal segmentation parameter settings. Furthermore, it is a step towards a model-based segmentation. An extension visualises the so-called uncertainty of segmentation (the Boundary Stability Index) which is proposed in Lucieer (2004).

3 Evaluation methods

3.1 Approaches to quantitative segmentation evaluation

As stated before, there is a variety of additional concepts and methods for evaluating image segmentation results. Here, a brief introduction to some prevailing algorithms is presented. The vast majority of the quantitative approaches are basically empirical discrepancy methods, analysing the number of misclassified pixels in relation to reference segmentations. In contrast, other algorithms directly address over- and under-segmentation by considering the number of segments, e.g. the Fragmentation Index FRAG (Strasters and Gerbrands 1991), the Area-Fit-Index AFI (Lucieer 2004) and the Precision/Recall Measure (Estrada and Jepson 2005).

Similar to the evaluation employed in this paper Yang et al. (1995) used shape features to quantify the differences between segmentations and reference regions. Mezaris et al. (2003) presented a distance weighted error measure for misclassified pixels. A Hausdorf-distance-based evaluation method for arc-segmentation algorithms is proposed by Liu et al. (2001). A map-algebra-based evaluation approach is introduced in Hirschmugl (2002). An intersection image of the segmentation result and the morphologically dilated binary reference segmentation is used to quantify the

number of misclassified pixels. A combined vector-raster-based procedure for assessing the precision of cadastral data using fractal box dimension is introduced by Schukraft and Lenz (2003). Assuming a given full cover reference segmentation, an adapted version of this algorithm is a promising approach to evaluate segmentation quality.

Other evaluation approaches are designed to minimise or exclude the a priori knowledge and the subjective (human) bias added to the evaluation by manually created references. Instead of using reference segmentations various objective evaluation criteria such as the intra-regional uniformity of segments are introduced (unsupervised evaluation methods). Cavallaro et al. (2002) present a perceptual spatio-temporal quality measure which allows an automated and objective evaluation by considering human perception criteria. However, it only applies to video sequence segmentation. Borsotti et al. (1998) presented an evaluation function which uses the colour uniformity within the segmented regions as criterion. Zhang et al. (2004) introduced a new entropy based evaluation approach, which leads to a very stable assessment measure using different segmentations.

An approach that comprises both analytical and empirical criteria is presented in Everingham et al. (2002) by defining a multidimensional fitness-cost-space instead of a single discrepancy-parameter-space. A promising co-evaluation framework which combines the results of various evaluation approaches using a machine learning approach is proposed in Zhang et al. (2005).

Further and partially older approaches to quantitative evaluation of segmentation results can be found in Yasnoff (1977), Levine and Nazif (1985), Haberäcker (1995), Yang et al. (1995), Schouten and Klein Gebbinck (1995), Zhang (1996), and Letournel et al. (2002). Table 2 provides an overview of recently proposed quantitative evaluation methods within the classification framework given in Zhang (1996).

3.2 Applied evaluation method

All segmentations as well as the delineation of reference objects were performed based on pan-sharpened multi-spectral IKONOS data (4 channels, 16 bit, 1 m ground resolution, principle component algorithm) of two test areas. Each test area has a size of about 2,000 by 2,000 pixels, representing an urban and a rural landscape. The procedure was aimed at the extraction of relevant land cover/use object boundaries.

Table 2. Approaches to quantitative evaluation of segmentation results

Evaluation Approach/Reference	Method Type[a]	Equation	Addressed issue
Fragmentation *(FRAG)* Strasters and Gerbrands (1991)	ED	$FRAG = \dfrac{1}{1 + p \cdot |T_N - A_N|^q}$ where T_N is the number of objects in the image and A_N the number of regions in the reference; p and q are scaling parameters	(1) Over-/under-segmentation by analysing the number of segmented and reference regions
Area-Fit-Index *(AFI)* Lucieer (2004)	ED	$AFI = \dfrac{A_{\text{reference object}} - A_{\text{largestsegment}}}{A_{\text{reference object}}}$	(2)
Geometric features Circularity Yang et al. (1995)	ED	$Circularity = \dfrac{4\pi A}{P}$ where A is the area and P is the perimeter	(3) Shape conformity between segmentation and reference regions (scaling invariant shape feature)
Geometric features Shape Index Neubert and Meinel (2003)	ED	$ShapeIndex = \dfrac{P}{4\sqrt{A}}$ where A is the area and P is the perimeter	(4)
Empirical Evaluation Function Borsotti et al. (1998)	EG	$Q(I) = \dfrac{1}{1000(N \cdot M)} \sqrt{R} \sum_{i=1}^{R} \left[\dfrac{e_i^2}{1 + \log A_i} + \left(\dfrac{R(A_i)}{A_i} \right)^2 \right]$ where $N \cdot M$ is the size of the image I, e_i is the colour error of the region i and $R(A)$ the number of regions of the size A	(5) Uniformity within segmented regions (colour deviation)
Entropy-based evaluation function and a weighted disorder function Zhang et al. (2004)	EG	$E = H_l(I) + H_r(I)$ where H_l is the layout entropy and H_r is the expected region entropy of the image I	(6) Uniformity within segmented regions (luminosity) using the entropy as a criterion of disorder within a region
Fitness function Everingham et al. (2002)	A, ED	Probabilistic hull, Potential accuracy $f(a,I)$ Multidimensional fitness-cost-space	(7) Multiple criteria and parameterisations of algorithms by a probabilistic Fitness/Cost Analysis

[a] According to Zhang (1996): Analytical (A), Empirical Goodness, unsupervised (EG), Empirical Discrepancy, supervised (ED).

According to the procedure proposed and applied in Neubert and Meinel (2003) firstly all results came under an overall visual survey. General criterions, like the delineation of varying land cover types (e. g. meadow/forest, agriculture/meadow, etc.), the segmentation of linear objects, the occurrence of faulty segmentations and a description of the overall segmentation quality were in the focus of this first step.

Furthermore, a detailed comparison based on visual delineated and clearly separable reference areas each representing a single land cover type was carried out. Therefore 20 different areas (varying in location, form, area, texture, contrast, land cover type etc.) were selected and each was visually and geometrically compared with the segmented pendants. The geometrical comparison as a combination of morphological features (area A_i, perimeter P_i, and Shape Index SI_i, see table 2, Eq. 3) is performed using an automated GIS procedure. The Shape Index comes from Landscape Ecology and addresses the polygon form. For all features the variances to the reference values were calculated. Contrary to previous studies the maximum variances have been capped at 100 % to avoid an overvaluation of single object variances. As partial segments all polygons with at least 50 % area within a reference object were counted. The number of partial segments addresses the over-segmentation. However, this issue can be treated as only conditionally prejudicial, since a further visual interpretation, segmentation level or automated classification can tackle the object fragments by dissolving or merging them. The much more critical problem is undersegmentation. Thus, for the evaluation process the partial segments become merged. Additionally, the quality of segmentation was visually rated (0 poor, 1 medium, 2 good). A good segmentation quality is reached, when the overall differences of all criteria between the segmentation results and the associated reference objects are as low as possible.

4 Results and discussion

4.1 Visual quality assessment and software specifics

4.1.1 EDISON

The mean shift based segmentation algorithm of EDISON produces well delineated representations of the image objects (figure 1). This is represented by the low area and perimeter deviations between the result and the according reference objects. A drawback is the extensive over-segmentation especially of forested and urban areas which leads to a minor visual quality rating.
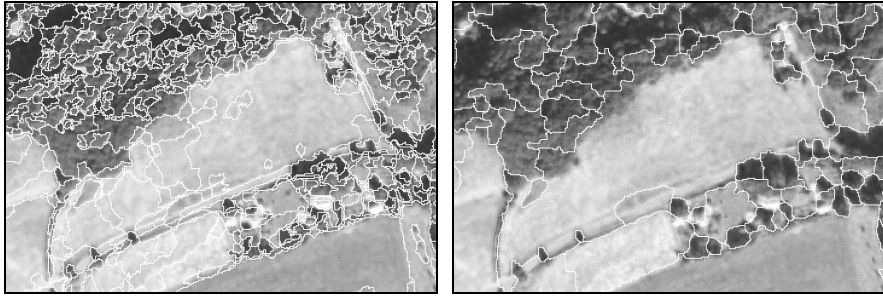
Fig. 1. Subsets of the segmentation results of EDISON (left) and EWS (right)

4.1.2 EWS 1.0

The results of EWS fit most of the object classes well, especially in the urban scene (figure 1). The algorithm tends to form more compact objects, what leads to some under-segmentation, even in the case of high object contrast (e.g. representation of linear objects or at field and forest boundaries). At the moment the software works only on Chinese operating systems and has a Chinese GUI only. Difficulties using projected imagery may occur.

4.1.3 Definiens Developer 7.0

In comparison to earlier evaluations (Neubert and Meinel 2003) the results obtained by Developer 7.0 using the same parameters are almost identical to version 3.0 (figure 2). Only at the image borders some minor differences occur. The results still contain some irregular or ragged delineated segments, especially at seam-forming boundaries and in forest areas. In areas of low contrast the occurrence of faulty segmentations is possible. Large homogenous image objects are divided arbitrarily sometimes. Since release 3.0 and some minor adjustments in version 5.0 Definiens uses an algorithm which enables a reproducible result not depending on image size. This is an important improvement because often parameters are tested on small subsets. By exporting the results into a vector output, they can be smoothened. Altogether Definiens has a high potential due to its multi-scale segmentation and the fuzzy logic based image classification capabilities. Because of the various interfaces to other GIS and remote sensing software systems important user requirements are complied.
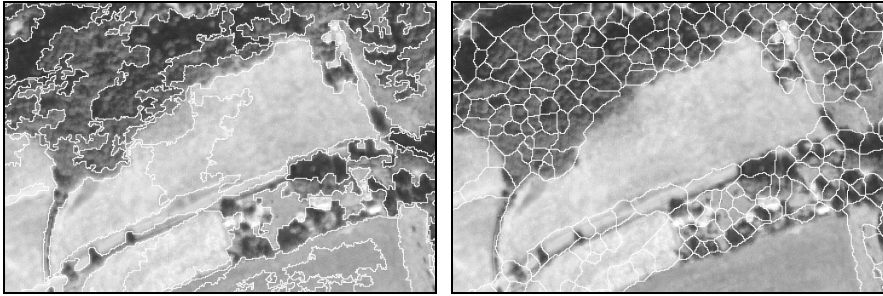
Fig. 2. Subsets of the segmentation results of Definiens Developer (left) and HalconSeg (right)

4.1.4 HalconSeg 1.1

The proposed adapted segmentation algorithm for the image processing software HALCON offers satisfying results (figure 2). As observed with all approaches the result is highly depending on the parameter settings. A significant influence to the segmentation quality could be seen for the parameterisations of the applied edge detection filter and the morphological operators. Linear structures are represented very well in the segmentation result, while it tends to forming compact objects some areas.

4.1.5 InfoPACK 2.0

The result of InfoPACK 2.0 does not reach the release 1.0 quality (figure 3). Nevertheless, it shows a good delineation for most of the objects, but strongly tends to over-segmentation. Homogeneous areas are thereof less affected and are adequately represented. In particular especially forested and built-up areas were much partitioned. At land cover transitions often interfering seam-forming segments were created. Generally low contrasted boundaries were segmented correctly. For processing scenes of any size the software uses an implemented tiling algorithm. Indeed, this leads to additional segment boundaries at the tile transitions. Furthermore, margin effects can yield to different results on both sides of the tile boundary. Like Definiens the software contains additional classification tools. Thus, a merging of similar classified and neighbouring segments is possible and this reduces the number of elements significantly. It must be pointed out, that InfoPACK has been developed to analyse very noisy radar data. Hence, the segmentation of optical data could be suboptimal.
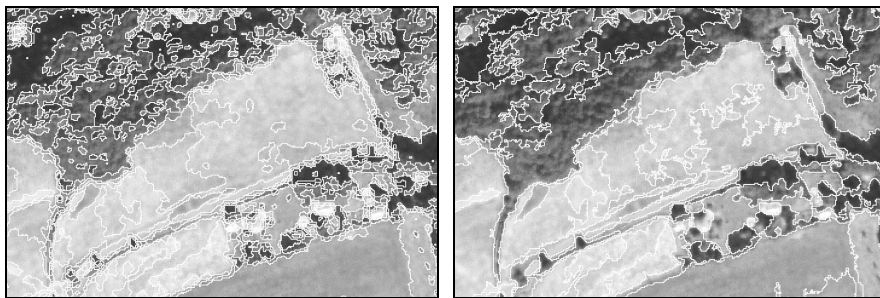
Fig. 3. Subsets of the segmentation results of InfoPACK (left) and RHSEG (right)

4.1.6 RHSEG 1.30

The RHSEG software produces a set of segmentations as a result of different hierarchy and resolution levels (figure 3). Additionally it allows the most extensive parameter settings of all programs beside InfoPACK. The results show both over- and under-segmentation within the same segmentation. Well-contrasted boundaries between main land cover classes were correctly represented. Areas of low contrast were often not reproduced properly. Due to the multitude of parameter settings there is still a need for optimisation regarding the usability.

4.1.7 SCRM

The segmentation result of SCRM yields to a good overall quality and a mostly good representation of the main object outlines (figure 4). Predominantly compact objects are delineated. Thus, some effects of under-segmentation occur, mainly by segmenting linear objects.
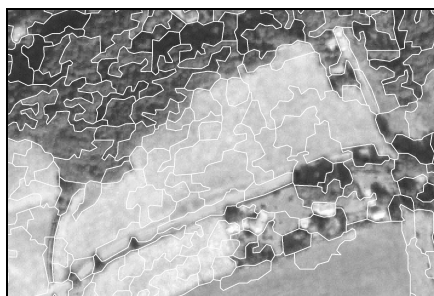
Fig. 4. Subset of the segmentation result of SCRM

Furthermore, homogeneous regions are often over-segmented (e.g. rivers, forests, meadows). The resulting delineation does not correspond with the pixel outlines due to an internal smoothing procedure. Like RHSEG the SCRM algorithm has the ability to run from batch files.

4.2 Comparison based on reference areas

In addition to the visual assessment, all segmentations were quantitatively (objectively) evaluated by means of 20 reference areas. The overall results are cumulated and compared in table 3. It can be recognised that the values of the tested algorithms are almost in the same range. The highest conformity concerning the average area is achieved by EDISON, Definiens as well as InfoPACK. A high average perimeter and region shape congruency (low Shape Index differences) is reached by EDISON, EWS, Definiens and HalconSEG. The low perimeter deviations of HalconSEG are the result of smoothed segment boundaries due to the application of a morphological image processing. The segmentations of Definiens, EWS and SCRM yield to a minor over-segmentation. Definiens and SCRM show the visual best results. The objective quantitative measures correspond with the subjective visual rating.

Table 3. Cumulated results of all 20 reference areas

Segmentation program	EDI-SON	EWS 1.0	Definiens Dev. 7.0	Halcon-SEG 1.1	Info-PACK 2.0	RH-SEG 1.30	SCRM
Av. diff. of area [%]	11,5	24,6	15,9	21,0	17,0	19,0	22,1
Av. diff. of perimeter [%]	13,8	18,1	17,2	16,6	29,6	21,6	20,0
Av. diff. of Shape Index [%]	12,4	15,4	16,2	14,9	46,0	23,5	19,0
Av. no. of partial segments	13,4	3,8	1,8	8,2	21,4	41,3	5,1
Av. quality, visual rated [0…2][a]	0,8	0,7	0,9	0,7	0,6	0,6	0,9

[a] 0 - poor, 1 - medium, 2 - good.

5 Conclusions

This contribution has presented an overview and some theoretical background on image segmentation quality assessment. It has been shown that there is no established standard evaluation method. However, there exist various ad hoc approaches. For this study an empirical discrepancy method was used to compare recently available segmentation programs. Due to the diversity of implemented algorithms the segmentation results are varying remarkably. The appropriateness of each program is still highly depending on the specific segmentation task. Beside a suboptimal segmentation quality some of the implemented algorithms still face technical issues such as a lack of process stability and robust import routines concerning image size and format, radiometric resolution, data structure and projection parameters. For this reason, almost all algorithms are still under development. Another optimisation aspect is the minimisation of segmentation parameters. As it could have been noticed, the algorithms are very sensitive to slightly differing parameterisations. Their number should be diminished in order to reduce the effort needed to obtain an optimal segmentation (mostly by *trial-and-error*). Implemented evaluation methods (like proposed in HalconSEG) could iteratively support the user to reach the optimal scene-depending settings (see also Costa et al., chapter 7.5). For the extraction of geo-information from segmentation results, integrated segment-based classification methods are desirable. Referring to this most of the programs tested offer simple to high-end resources.

Despite suboptimal results, segmentation offers an important approach to semi-automated image analysis. Particularly in combination with presented evaluation methods and existing GIS-data image segmentation algorithms already are indispensable resources to retrieve geo-information from remote sensing imagery.

In combination with the previous studies, in total 21 segmentation programs or its releases have been evaluated. The results of all segmentations are displayed at the website www.ioer.de/segmentation-evaluation. Is has been shown, that there is more than one interesting approach in this dynamic field of research. The evaluation will be continued, e.g. using new algorithms like Feature Analyst 4.1, SAGA GIS 2.0 and other implementations. Furthermore, it is planned to extend the quality assessment procedure itself by some of the presented evaluation methods as well as by spatially explicit outline delineation quality measures. The evaluation is open to further algorithms.

References

Baatz M, Schäpe A (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Strobl J, Griesebner G (eds) Angewandte Geographische Informationsverarbeitung XII. Wichmann, Heidelberg, pp. 12-23

Borsotti M, Campadelli P, Schettini R (1998) Quantitative evaluation of color image segmentation results. Pattern Recogn Lett 19(8):741-747

Carleer AP, Debeir O, Wolff E (2005) Assessment of very high spatial resolution satellite image segmentations. Photogramm Eng Rem Sens 71(11):1285-1294

Castilla G (2003) Object-oriented analysis of Remote Sensing images for land cover mapping: conceptual foundations and a segmentation method to derive a baseline partition for classification. Ph.D. Thesis, Polytechnic University of Madrid   [http://www.montes.upm.es/Servicios/biblioteca/tesis/GCastillaTD_Montes.pdf – 10. April 2007]

Castilla G, Hay GJ, Ruiz JR (in press) Size-constrained region merging (SCRM): an automated delineation tool for assisted photointerpretation. Photogramm Eng Rem Sens

Cavallaro A, Gelasca ED, Ebrahimi T (2002) Objective Evaluation of Segmentation Quality Using Spatio-temporal Context. In: Proc IEEE Int Conf Image Process, 4 p. [http://www.elec.qmul.ac.uk/staffinfo/andrea/papers/cavallaro_icip02.pdf – 10. April 2007]

Comanicu D, Meer P (2002) Mean shift: A robust approach toward feature space analysis. IEEE Trans Pattern Anal Machine Intell 24:603-619

Cook R, McConnell I, Stewart D, Oliver CJ (1996) Segmentation and simulated annealing. In: Franceschetti G, Rubertone FS, Oliver CJ, Tajbakhsh S (eds.) Microwave Sensing and Synthetic Aperture Radar. Proc. SPIE 2958:30-35

Estrada FJ, Jepson AD (2005) Quantitative evaluation of a novel image segmentation algorithm. Comput Vis Pattern Recogn 2:1132-1139

Everingham M, Muller H, Thomas B (2002) Evaluating image segmentation algorithms using monotonic hulls in fitness/cost space. In: Cootes T, Taylor C (eds.) Proc 12th Br Mach Vis Conf. pp. 363-372

Haberäcker P (1995) Praxis der digitalen Bildverarbeitung und Mustererkennung. Hanser Verlag, München, 350 p.

Herold H (2005) Beiträge zum Vergleich von Segmentierungsprogrammen fernerkundlicher Bilddaten. Diploma Thesis, Inst Geogr, Tech Univ Dresden, 125 p. (unpublished)

Hirschmugl M (2002) Automatisierte Bildsegmentierung von Nutzungsstrukturen in der Fernerkundung. Diploma Thesis, Inst Geogr Reg Stud, Univ Graz, 115 p. (unpublished)

Karantzalos KG, Argialas DP (2003) Evaluation of Selected Edge Detection Techniques in Remote Sensing. In: Serpico SB (ed.) Image and Signal Processing for Remote Sensing VIII, Proc SPIE 4885, pp. 102-110

Lanser S (1993) Kantenbasierte Farbsegmentierung im CIE-Lab Raum. In: Pöppl H (ed.) Mustererkennung. Informatik aktuell, Springer, Heidelberg, pp. 639-646

Letournel V, Sankur B, Pradeilles F, Maître H (2002) Feature Extraction for Quality Assessment of Aerial Image Segmentation. ISPRS Tech Comm III Symp, 6 p. [http://www.isprs.org/commission3/proceedings02/papers/paper105.pdf – 10. April 2007]

Levine M, Nazif A (1985) Dynamic Measurement of Computer Generated Image Segmentations. IEEE Trans Pattern Anal Mach Intell 7(2):155-164

Li P, Xiao B (in press) Multispectral image segmentation by a multichannel watershed-based approach. Int J Rem Sens [http://www.informaworld.com/ 10.1080/01431160601034910 – 10. September 2007]

Liu W, Zhai J, Dori D, Tang L (2001) A System for Performance Evaluation of Arc Segmentation Algorithms. Proc CVPR Workshop Empir Eval Comput Vis, 11 p. [http://www.cs.cityu.edu.hk/~liuwy/ArcContest/NoiseModels.pdf – 10. April 2007]

Lucieer A (2004) Uncertainties in Segmentation and Their Visualisation. PhD Thesis, Utrecht University, ITC Dissertation 113, Enschede, 174 p. [http:// www.itc.nl/library/Papers_2004/phd/lucieer.pdf – 10. April 2007]

Meinel G, Neubert M (2004) A comparison of segmentation programs for high resolution remote sensing data. Int Arch Photogram Rem Sens Spatial Inform Sci XXXV-B4:1097-1102 [http://www.isprs.org/istanbul2004/comm4/papers/ 506.pdf – 10. April 2007]

Mezaris V, Kompatsiaris I, Strintzis MG (2004) Still Image Segmentation Tools for Object-based Multimedia Applications. Int J Pattern Recogn 18(4):701-725

Neubert M, Meinel G (2003) Evaluation of segmentation programs for high resolution remote sensing applications. In: Schroeder M, Jacobsen K, Heipke C (eds.) Proc Joint ISPRS/EARSeL Workshop High Resolution Mapping from Space 2003, 8 p. [http://www.ipi.uni-hannover.de/html/publikationen/2003/ workshop/neubert.pdf – 10. April 2007]

Neubert M, Herold H, Meinel G (2006) Evaluation of Remote Sensing image Segmentation Quality – Further Results and concepts. Int Arch Photogram Rem Sens Spatial Inform Sci XXXVI-4/C42, 6 p. [http://www.commission4. isprs.org/obia06/Papers/10_Adaption and further development II/OBIA2006_ Neubert_Herold_Meinel.pdf – 10. April 2007]

Palus H, Kotyczka T (2001) Evaluation of Colour Image Segmentation Results. Proc 7th Workshop Farbbildverarbeitung 2001, 4 p. [http://kb-bmts.rz.tu-ilmenau.de/gcg/html/Vortr_01_pdf/Palus_Bild.pdf – 10. April 2007]

Schouten TE, Klein Gebbinck MS (1995) Quality Measures for Image Segmentation Using Generated Images. Proc Image Signal Process Rem Sens II:411-422

Schukraft A, Lenz R (2003) Ermittlung der Lagegenauigkeit von Vektordaten mit Hilfe der fraktalen Boxdimension. GeoBIT/GIS 3:15-19

Strasters K, Gerbrands J (1991) Three-dimensional segmentation using a split, merge and group approach. Pattern Recogn Lett 12:307-325

Tilton JC (2003) Analysis of hierarchically related image segmentations. Proc IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data:60-69

Yang L, Albregtsen F, Lønnestad T, Grøttum P (1995) A Supervised Approach to the Evaluation of Image Segmentation Methods. Proc. CAIP, Lect Notes Comput Sc 970:759-765

Yasnoff WA, Mui JK, Bacus J. W. (1977) Error measures for scene segmentation. Pattern Recogn 9(4):217-231

Zhang H, Fritts JA, Goldman SA (2004) An entropy-based objective segmentation evaluation method for image segmentation. SPIE Electronic Imaging - Storage and Retrieval Methods and Applications for Multimedia:38-49

Zhang H, Fritts JA, Goldman SA (2005) A Co-Evaluation Framework for Improving Segmentation Evaluation. SPIE Signal Processing and Target Recognition XIV:420-430

Zhang YJ (1996) A survey on evaluation methods for image segmentation. Pattern Recogn 29(8):1335-1346

Acknowledgments

Chapter 8.4

Object-fate analysis: Spatial relationships for the assessment of object transition and correspondence

E. Schöpfer, S. Lang, F. Albrecht

Centre for Geoinformatics (Z_GIS), Salzburg University, Schillerstr. 30, 5020 Salzburg, Austria - (elisabeth.schoepfer, stefan.lang, florian.albrecht)@sbg.ac.at

ABSTRACT: In the near future several new highest resolution, next generation satellites will be launched with panchromatic half-meter resolution imagery, e.g. WorldView 1 and 2. The ever increasing supply of high-resolution imagery seeks for adequate, i.e. more effective, more automated and reliable methods for image processing and interpretation. At the interface of geographic information science and remote sensing, object-based image analysis methodologies provide a solid basis for exploiting imagery more intelligently. Working with image objects enables (1) single feature, specific information extraction, (2) performing complex classifications and multi-scale representations and (3) spatial analysis and modeling. However, deriving image objects from various sources and in different scales implies the problem of generating inconsistent boundaries. To specifically address this challenge, a tool called LIST (*landscape interpretation support tool*) is used, which, based on a straight-forward principle, analyses the spatial relationships of image objects, i.e. their correspondence and their changes over time. The chapter presents a methodological discussion and preliminary results from an ongoing study on 'object fate analysis' (OFA). OFA means the investigation of object transition (change over time) or object correspondence (different delineations or representations). The concept and the application of OFA are illustrated by two case-studies representing both aspects. The first one carried out in medium scale uses Landsat TM and ETM imagery and shows an example of performing change assessment as well as object-based accuracy assessment. The sec-

ond fine-scale study is based on SPOT 5 scenes and demonstrates how object correspondence can be assessed on two different object representations, machine-based segmentation and manual delineation.

Introduction

Landscape changes through time

With increasing awareness of sustainable usage of land, the preservation of cultural landscapes and the protection of the environment, both detecting and assessing changes become ever more important. European landscapes, shaped and managed over time, reveal a complex mosaic of land uses from cultivated to natural lands over a large geographic area. Changes in these landscapes are driven by inter-annual climate variability or long-term climatic trends, vegetation succession, natural disturbances, and of course, human land use (Schöpfer, 2005). In many areas of the world, socioeconomic and political factors induce and influence prevailing land use and land cover structures (Croissant, 2004). On satellite imagery, resulting changes are often acute and recognizable, though sometimes subtle and difficult to tell from noise. Analyses of changes in the landscape patterns are based on where the changes occur, the type of changes, and the degree and rates of these (Southworth et al., 2004). To this end, change detection maps are important at first, but the understanding of temporal and spatial dynamics of the landscape is likewise crucial (Turner, 1990).

A way to characterize landscape patterns is to quantify their composition (i.e. percentage of classes) and configuration (i.e. spatial arrangement of patches) (Turner et al., 2001). So called landscape metrics have been successfully introduced as quantitative structural measures for monitoring landscape patterns and relating ecological meanings to these structures (Wickham et al., 1997; Petit and Lambin, 2002; Hudak et al., 2004; Narumalani et al., 2004; Corry and Nassauer, 2005; Lang and Blaschke, 2007). Landscape metrics are also used as structural indicators to highlight pattern-related changes in the landscape, which are considered to cause a substantial shift of underlying processes (e.g. a bog falling dry due to water subtraction, cf. Langanke and Lang, 2004).

Spatial change detection

Numerous techniques have been developed for detecting changes on images or other raster representations of the landscape. It was shown that dif-

ferent types of landscapes to be represented require respective methods (e.g. pre- versus post-classification change detection techniques, Coppin et al., 2003). For simplicity, though, most change detection studies focus on spatial implicit measures like the percentage of changes. The latter is a basic indicator and may reveal a general trend, but fine-scaled applications such as the distribution of urban green or presence of sensitive habitats in mountainous areas, the spatial pattern formed by landscape elements and the change of this very arrangement is essential (Lang and Blaschke, 2007). Thus, methods need to be capable of assessing spatial-explicit changes within the arrangement of landscape units (or image objects, respectively). But tools which handle object-specific change detection in terms of a specific 'fate' of an object are rare (Lang et al., in press). Common ways of performing change assessment, i.e. map-to-map comparisons using raster overlay techniques, are site-, but not object-specific (i.e. they refer to pixels). Their aggregate is spatial implicit. Contrarily, vector overlay operations such as intersection, identity and union, are spatial explicit, but usually produce complex geometry with sliver polygons. Finally, visual comparisons, though maybe more accurate or appropriate than those made by a machine, rely on subjective interpretation and are rather time-consuming.

In geographical information science (GIScience), sets of spatial relationships are used to characterize how objects are related to each other in space. These topological relationships allow for performing complex spatial modeling and analysis. This class of spatial relationships is based on principles like adjacency (*what adjoins what*), containment (*what encloses what*), and proximity (*how close something is to something else*). Spatial relations in general have been discussed within GIScience since several decades. Profound theoretical papers have been presented by Mark (1999) or Hornsby and Egenhofer (2000). However, within the remote sensing community, very few attempts have been made to utilize spatial relationships for change assessment. Attempts to characterizing changes through topological relationships between corresponding patches were used by e.g. Raza and Kainz (2001) and the idea was conceptually presented by Blaschke (2005).

A simple typology for changes of object geometries may be based on four categories, namely existence-related, size and shape-related and location-related changes (Schöpfer and Lang, 2006). Whereas in theory these categories sound quite distinct, in reality a combination of all of these basic types of geometric changes needs to be tackled. In this context, it needs to be highlighted that this article (in continuation of preceding publications of the authors) is built upon the notion that object-based change assessment is not limited to temporal changes only. In an abstracted way,

'changes' may also refer to different object representations. In this sense, 'object fate' (Lang et al., in press) may be attributed to both the change of an object over time and the specific way of how it has been delineated. In this general understanding, the object-fate approach being used throughout, t_0 and t_1 either indicate time slot 1 and time slot 2 or they stand for any given representation A and representation B.

Test sites and data preprocessing

The research took place in two study sites in Europe (see figure 1). Site selection was performed by focusing on high structural dynamic areas. The first area under investigation (**AUI_1**) is situated within the border zone of the former Iron Curtain between Austria and Hungary. The landscape is characterized by a unique pattern along the border before the fall of the Iron curtain. Recently the transition zone became increasingly homogenized, i.e. extensive changes in land ownership and fragmentation of agricultural fields due to land reforms (Csaki, 2000). Land use and land cover change is dominated by land abandonment at unprecedented rates, and the conversion of farm land to grassland and forests (Kuemmerle et al., 2006).

Landsat TM (August 17, 1985) and Landsat ETM+ (August 2, 2000) were co-registered to a ground sample distance (GSD) of 30 meter in UTM zone 33N (WGS-84) by using a second degree polynomial transformation and nearest neighbor resampling with an RMS error of 0.47. Three representative sub-sites of 5000 m by 5000 m along the border were selected to perform the studies on changes in the land use pattern bearing typical cross-border characteristics. Image segmentation was performed to extract field boundaries for both time slots. For both images the same average object size (scale parameter) has been used, optimized to capturing the large fields in 1985 and the small heterogeneous parcels in 2000.

The second study area (**AUI_2**) is located in southern Germany in the region of Stuttgart, federal state of Baden-Wurttemberg. The test site, cropped for this purpose, comprises an area of 25.87 km² which includes three representative areas (22 ha, 34 ha, 117 ha) to perform object correspondence analysis. The area is characterized by intensive agricultural use on the one hand, and high settlement growth on the other hand. The use of high resolution satellite data is of critical importance to analyze the fine-scaled dynamics of the landscape (Schumacher et al., 2007). Land use changes in such areas are – besides the expansions of human settlements – mainly related to crop rotation, crop type changes, or crop density/maturity differences between the two dates.
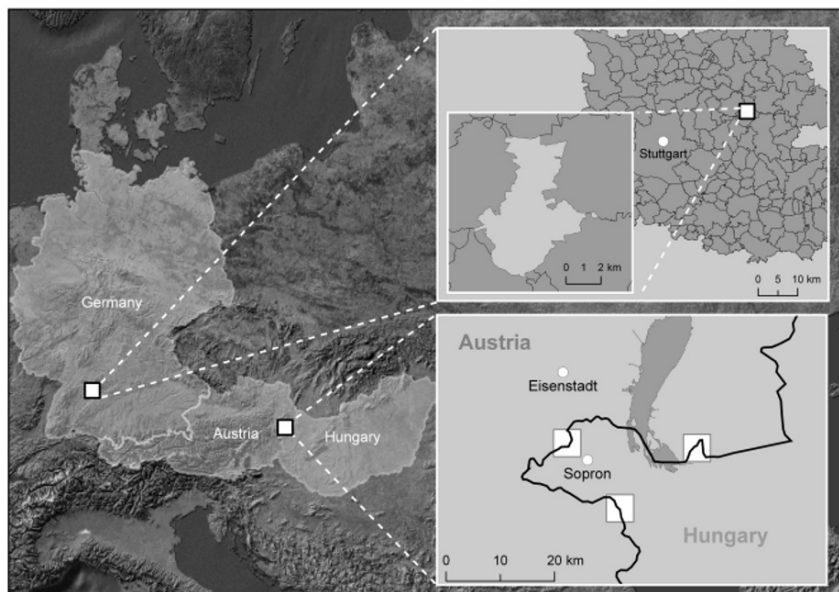
Fig. 1. Study areas situated along the border between Austria and Hungary (AUI_1) and in
southern Germany (AUI_2).

   In this study pan-sharpened SPOT imagery with a ground sample dis-
tance (GSD) of 5 m, co-registered (UTM-33) to an existing orthophoto
mosaic (0.25 m GSD) and orthorectified using parameterized orbital
pushbroom model, have been used (Tiede et al., submitted). Data were
segmented using an adaptive per-parcel approach (ibid.). Digital cadastre
data were used as pre-defined boundaries to perform parcel-based segmen-
tation. Within heterogeneous objects, region-based, local mutual best fit-
ting segmentation (Baatz and Schäpe, 2000) was used to split the respec-
tive objects into homogenous units (ibid.). For performing object
correspondence analysis, results from a visual interpretation were used car-
ried out on the orthophoto mosaic.

Object-fate analysis

Studies on imagery with high spatial variability should pay more attention
on the spatial arrangements rather than merely on (spatial implicit) re-
cordings of changes in discerned classes (Crews-Meyer, 2004). We present

and discuss a method for analyzing object-specific changes by investigating spatial relationships among corresponding objects. As outlined earlier, this can be seen both as a product of object transition (change over time) or as an outcome of different object representations or delineations. The latter allows for transferring the concept to object-based accuracy assessment (Schöpfer and Lang, 2006). Since spatial relations are usually a combination of basic relationship types, there is a demand for ready-to-use solutions which are able to categorize these.

The concept of object fate analysis is implemented in a tool called LIST (*Landscape Interpretation Support Tool*). LIST is developed and programmed as an extension for ESRI's ArcView 3 and ArcGIS 9 (Lang et al. (in press); Schöpfer and Lang, 2006; Weinke et al., chapter 8.5). The tool provides object quantification, complements visual interpretation and includes a method for object-based change analysis and object-based accuracy assessment. Following the concept of parent and child relationships between layers two vector themes are used to represent the specific 'fate' of corresponding objects (Lang et al., in press). 'Object fate' may reflect different data captured at various points of time (change analysis) or different methods for object generation (i.e. different segmentation algorithms, heterogeneous data material, visual vs. machine-based interpretation, reference data sets from other sources, etc.). Thus the comparison of two data sets can also be utilized for object-based accuracy assessment, as generated objects can be compared with manually delineated ones. When comparing two automatically segmented images, the corresponding boundaries of delineated image objects do not necessarily coincide due to differences in data material or segmentation algorithms (Schöpfer & Lang, 2006.). Although change may not have occurred, spatial boundaries may not be fully congruent. LIST utilizes a virtual overlay for investigating spatial relationships. To overcome the fact of spatial uncertainty in image object generation (spatial overlay strictness, SOS), an uncertainty measure in form of buffer, is introduced. The size of the buffer, either specified by the user or dynamically according to object size, controls the allowed spatial difference of spatially coinciding and corresponding objects. SOS also controls the degree of overlap of child-objects (t_1) between neighboring parent-objects (t_0), expressed by percentage values.

LIST investigates spatial relationships for three generalized states of transition (see figure 2).
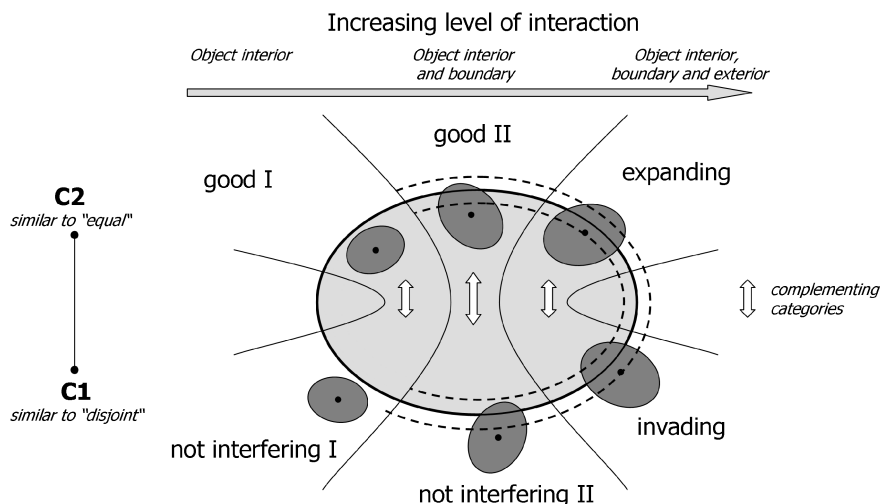
Fig. 2. Types of object-fate implemented in LIST. A buffer zone (spatial overlay strictness, SOS) is introduced (dashed line) in order to take into account inconsistent boundaries due to various sources and in different scales

It is assumed that t_0 (*before / A-representation*) objects are larger than t_1 (*after / B-representation*). The spatial relationships are divided into object categories by two decision steps. The first one decides if the t_1 object is more associated with the interior or the exterior of the t_0 object. This division is adopted in an analogy to an approach of object comparison (Straub and Heipke 2004) that groups the spatial relations into the clusters 'similar to disjoint' (C1) and 'similar to equal' (C2). The decision criterion divides the relations by the placement of the t_1 object's centroid within or outside of the t_0 object. The second decision specifies the spatial overlay strictness SOS and determines whether a t_1 object only touches the boundary, which has been softened by the SOS, or if it interferes both with the interior and the exterior of the t_0 object. Based on these decision rules object fate can be expressed by three special types of object transition:

(1) 'good I' objects inside a t_0 object and 'good II' objects covering the buffered outline of a t_0 object; A special case of 'good' objects occurs, when only one 'good' object is recorded, i.e. the t_0-object remains stable;

(2) 'expanding' objects exceeding the original boundary to a certain degree;

(3) 'invading' objects where the object's centroid lays outside of t_0-boundary, but a certain part of the object which can be specified in percent, lays within the t_0-object.

This concept allows both to characterize the development of a t_0-object and to categorize objects being produced in t_1. Finally the category (4) 'not interfering' was introduced to complete the status of transition. Again, two sub-categories are used to consider the buffer zone.

Within the concept there are complementing categories which are displayed in figure 3. The 'good' objects are in relationship to 'not interfering' whereas 'expanding' objects are completing 'invading' objects.
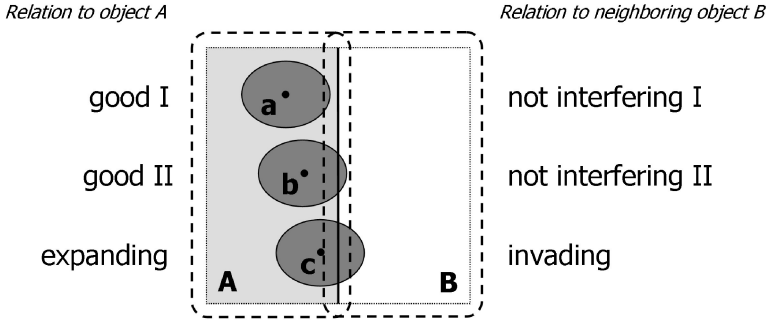
Fig. 3. Complementing categories within the LIST concept

The overall concept is a straight-forward solution to characterizing the development of each t_0-object ('parent object') and additionally enables unique categorization of every 'child object'. To characterize overall object stability two object-specific measures were introduced by Schöpfer and Lang, 2006. The first index, 'offspring loyalty' (*OL*), is calculated by:

$$OL = \frac{n_{good}}{n_{good} + n_{exp}} \qquad (1)$$

where n_{good} = number of good objects
       n_{exp} = number of expanding objects

A value of 1 indicates that no expanding objects are among the t_0 object. The second index, 'interference' (*I*) is defined by:

$$I = \frac{n_{inv}}{n_{all}} \qquad (2)$$

where n_{inv} = number of invading objects
       n_{all} = number of all intersecting t_1 objects

The smaller the value, the smaller the number of invading objects that interfere with the t_0 object. Both indices *OL* and *I* can be aggregated on landscape level (Schöpfer and Lang, 2006).

In the context of the two studies (**AUI_1** and **AUI_2**), the following steps were performed in particular. In **AUI_1** a dynamic buffer was created and 10% overlap for the definition of invading objects was used. This means, in case a t_1 object overlaps less than 10% with the t_0 object, it will be neglected. Usually, the overall number of expanding and good objects will be equal to the number of child objects. However, if working with stairs-like boundaries induced by the Landsat pixels, some centroids may fall exactly onto the field boundary between two neighboring fields. Centroid ambiguity applies if the sum of the numbers of t_1-objects in each category exceeds the actual number of t_1-objects (ibid.).

In **AUI_2**, in order to overcome stepped vector lines the t_o-theme and t_1-theme were generalized, by using 'smooth' and 'simplify' commands, implemented in the Data Management Tools in ArcGIS 9.2. Afterwards the lines were converted back to a polygon theme. The buffer was defined manually for each of the three sub study sites.

Results and discussion

Assessment of object transition

Object fate analysis has been performed in the **AUI_1** study area by using both offspring loyalty (*OL*) and interference (*I*) (see figure 4). *OL* ranges between 0 and 1 in all three study sites. Taking into account the t_0 objects with at least one good child object, the respective number of child objects are shown in three specific ranges of *OL* (< 0.6, 0.6-0.8 and >0.8) (see figure 5). The other measure, *I*, ranges from 0 to 1 as well, but with a majority of values between 0 and 0.6. Thus, we consider Interference moderate for all three study sites with high occurrences in the range >0-0.6 in sub-sites 1 and 3 as opposed to sub-site 2.
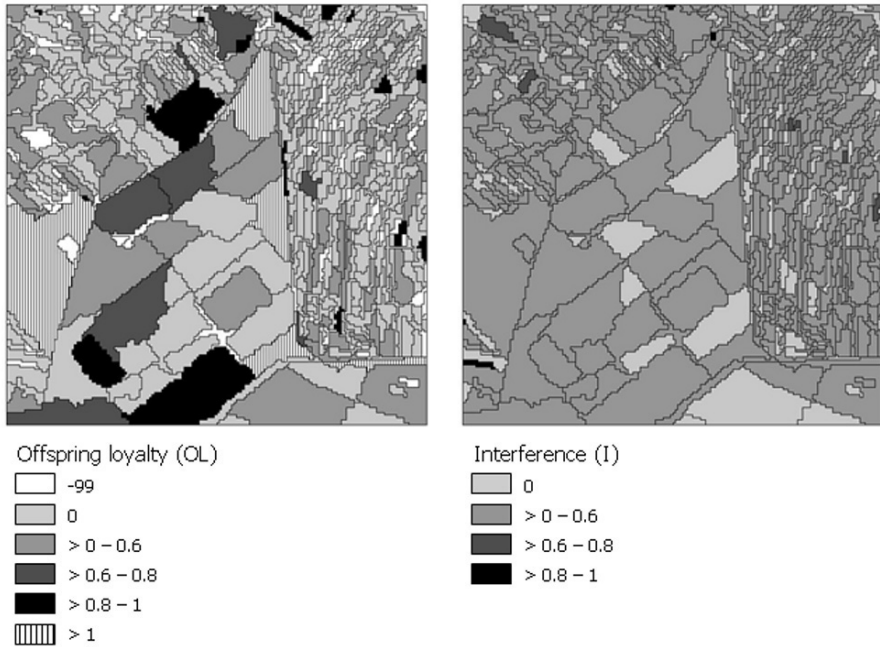
Fig. 4. Spatial distribution of two measures calculated for each object t_0 in study area 1. Left: Offspring loyalty (*OL*). Values of -99 indicate cases with neither good nor expanding objects; values higher than 1 indicate centroid ambiguity. Right: Interference (*I*)
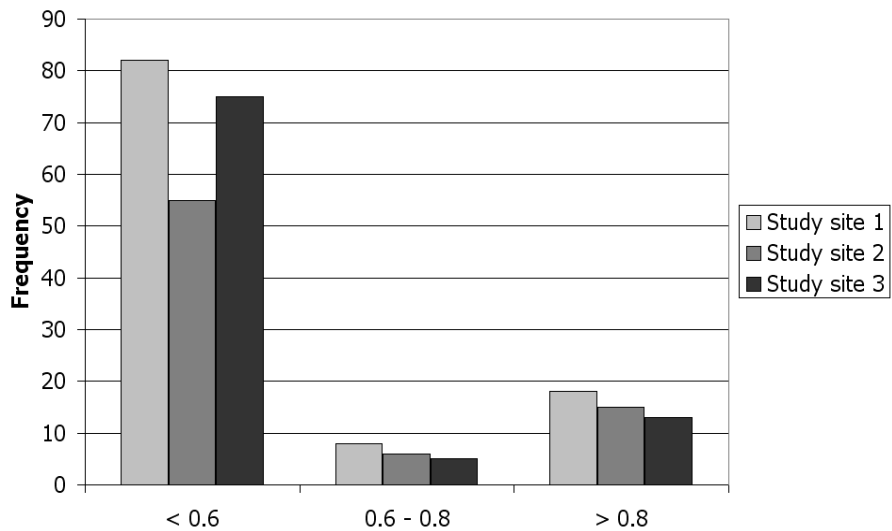
Fig. 5. Histogram of offspring loyalty in three ranges (< 0.6, 0.6 - 0.8 and > 0.8) for the three study sites

Object correspondence

Object correspondence was analyzed in the test site **AUI_2** on vector layers provided by manual interpretation and automated image segmentation. Table 1 shows the number of different categories of t_1 objects for the corresponding t_0 object.

Table 1. Assessment of object correspondence on two different object representations (machine-based segmentation and manual delineation) in the test site **AUI_2**; SOS-2 refers to the buffer around the object (meter) whereas SOS-3 regulates the percentage of the area covered by an invading object.

Id	Data status	n_all	n_good	n_exp	n_inv	stable	SOS-2	SOS-3
13	Original	7	1	0	0	yes	9	1
13	Generalized	5	1	0	0	yes	8	1
4	Original	16	7	0	0	no	4	3
4	Generalized	14	7	0	0	no	3.5	3
3	Original	9	1	2	4	no	11	1
3	Generalized	9	1	2	4	no	4.5	1
5	Original	9	1	1	2	no	11	1
5	Generalized	9	1	1	2	no	4.5	1

For each investigated t_0 object a version with generalized outlines was also compared. Due to the generalization some of the t_1 objects do not interfere with the corresponding t_0 object any more. So, for the t_0 objects with the ID 13 and ID 4 the number of n_all is lower for the comparison of the generalized data than for that of the original data. Fig. 6 shows two 'Not Interfering I'-objects that changed their category through generalization. For all the investigated objects the generalization allowed to use a smaller buffer size (SOS-2) which is very significant for the t_0 objects with the ID 3 and Id 5 (see table 1 and figure 8). The t_0 object with the ID 13 is a stable object because it has only one 'Good' t_1 object and no invading or expanding objects (see figure 6). The t_0 object with the ID 4 has several 'Good' t_1 objects that can be assigned directly (see figure 7). The t_0 object with the ID 3 has invading and expanding objects as well as good II objects that, in the t_0 object with the Id 5, belong to the complementing object category.
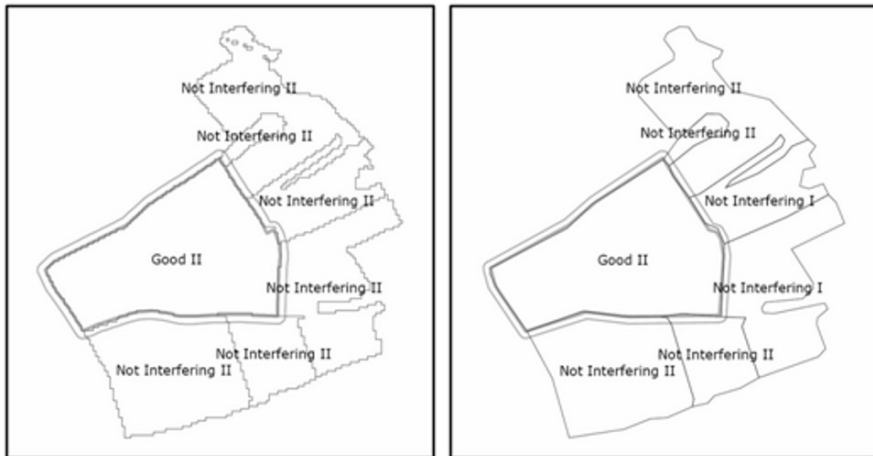
Fig. 6. T_0 object ID 13; original data (left) and generalized data (right).
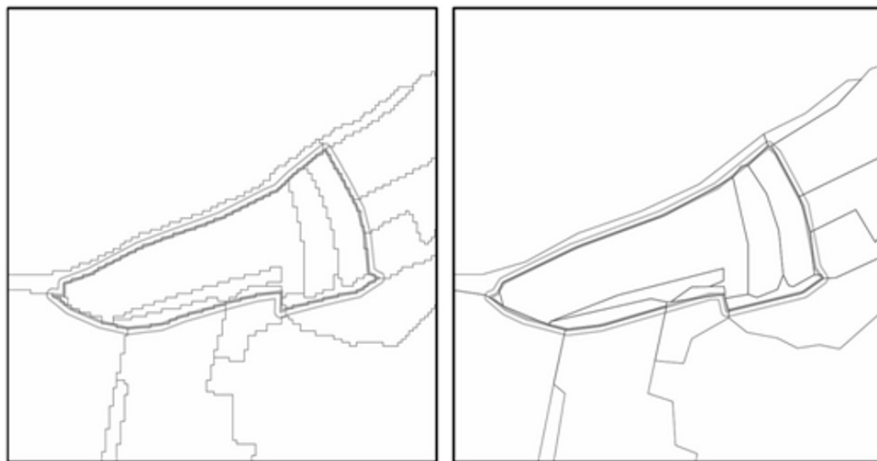
Fig. 7. T_0 object ID 4; original data (left) and generalized data (right).

Fig. 8. T_0 object ID 3 (above) and ID 5 (below); original data (left) and generalized data (right).

Conclusions

In order to fulfill the requirements for new approaches to tackle with temporal and spatial complexities, the authors further developed their method on object fate analysis. The initial set of three different spatial relationships of object transition has been extended to four types: 'good', 'expand-

ing', 'invading' and 'interfering' objects There is a demand to come up with a ready-to-use solution that generalizes and categorizes the variety of spatial relations, which appear in different combinations over time. In general, the integration of GIS concepts of representing relationships as discussed by Langran (1992), Egenhofer (1994), Mark (1999) and Hornsby and Egenhofer (2000) is challenging to be adapted to change detection applications. Remote sensing methods need to integrate spatial concepts (Lang & Blaschke, 2006). Whereas the pixel-based neighborhood concept is limited to immediate or fixed neighborhood, and again depending on absolute pixel size, the object-based spatial relationship concept is based on polygons and thus suitable for adaptive spatial analysis.

Object fate analysis relies on a straightforward and easily operable concept of spatial relationships among corresponding image objects. These objects can be obtained from any two different sources, such as automatically segmented images, manually digitized ones, corresponding, but time wise differing ones, etc. Our experience shows that the exact comparison of such data sets, different in scale and representation, is not trivial. A scale-dependent SOS concept, i.e. a buffer, has been introduced to consider uncertainty in delineation. The parameterization of SOS is important for distinguishing real changes from data-induced one. Still, limitations occur, because a buffer is usually calculated with a fixed size for the entire object. If there are small outcrops along an object, the buffer needs to be increased, respectively (see Weinke et al., chapter 8.5). Thus this analysis approach depends on the quality of the generated objects, which in turn depends on the specific data material and segmentation algorithms. Ongoing studies attempt to assess particularly by looking at each vertex of an object's outline, at which portion of the object either changes have occurred or the delineation is different.

In general, the integration of GIS concepts representing spatial relationships offers a new dimension of change interpretation for land use/land cover related studies. The spatial explicit comparison of objects may overcome the limitations of a mere comparison of classification results based upon pixel classification.

The presented object-fate analysis can also be used for object-based accuracy assessment as discussed by Schöpfer and Lang (2006). Next to a point-based collection of correctly classified features, the very delineation of any given object may be even more important for validating the overall accuracy of the classification. This especially applies for high resolution imagery. Spatial explicit change analysis considers the high spatial variability of nature. How the pattern specifically changes, may be additionally characterized by spatial measures from the toolbox of landscape metrics.

By this, and with complementary in situ measure, we can not only assess, but also evaluate these changes.

Acknowledgements

References

Baatz M and Schäpe A (2000) Multiresolution Segmentation – an optimization approach for high quality multi-scale image segmentation. In: Angewandte Geographische Informationsverarbeitung XII, Strobl, J., Blaschke, T., Griesebner, G. (eds.), Wichmann: Heidelberg, 2000, pp. 12-23.

Blaschke T (2005) A framework for change detection based on image objects. In: Erasmi, S. et al. (eds.), Göttinger Geographische Abhandlungen 113, pp. 1-9.

Coppin P R, Jonckheere I G and Lambin E F (2003). Digital change detection in ecosystem monitoring: a review. International Journal of Remote Sensing 25 (9). pp. 1565-1596.

Corry R C and Nassauer J I (2005) Limitations of using landscape pattern indices to evaluate the ecological consequences of alternative plans and designs. Landscape and Urban Planning, 72, pp. 256-280.

Crews-Meyer K A (2004) Agricultural Landscape Change and Stability: Historical Patch-Level Analysis. Agriculture, Ecosystems and Environment, 101, pp. 155-169.

Croissant C (2004) Landscape Patterns and Parcel Boundaries: An Analysis of Composition and Configuration of Land Use and Land Cover in South-Central Indiana. Agriculture, Ecosystems and Environment, 101, pp. 219–232.

Csaki C (2000) Agricultural reforms in Central and Eastern Europe and the former Soviet Union - Status and perspectives. Agricultural Economics 22, pp. 37–54.

Egenhofer M J (1994) Pre-Processing queries with spatial constraints. Photogrammetric Engineering & Remote Sensing, 60 (6), pp. 783-790.

Hornsby K and Egenhofer M J (2000) Identity-based change: a foundation for spatio-temporal knowledge representation. International Journal of Geographical Information Science, 14 (3), pp. 207-224.

Hudak A T, Fairbanks D H K and Brockett B H (2004) Trends in fire patterns in a southern African savanna under alternative land use practices. Agriculture, Ecosystems and Environment, 101, pp. 307-325.

Kuemmerle T, Radeloff, V C, Perzanowski K and Hostert P (2006) Cross-border comparison of land cover and landscape pattern in Eastern Europe using a hybrid classification technique. Remote Sensing of Environment, 103, pp. 449-464.

Lang S, Schöpfer E and Langanke T, in press. Combined object-based classification and manual interpretation – Synergies for a quantitative assessment of parcels and biotopes. Geocarto International.

Lang S and Blaschke T (2007) Landschaftsanalyse mit GIS. UTB-Reihe. Stuttgart: Eugen-Ulmer-Verlag, 420 pages.

Lang S and Blaschke T (2006) Bridging remote sensing and GIS - what are the most supportive pillars? In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. No. XXXVI-4/C42, Salzburg, Austria.

Langanke T and Lang S (2004) Strukturelle Indikatoren zur Beurteilung von Habitatqualität im europäischen Naturschutz. In: C. Dormann et al. (Ed.), Habitatmodelle – Methodik, Anwendung, Nutzen. Leipzig: UFZ-Berichte 9/2004, pp. 141-145.

Langran G (1992) Time in geographic information system. London: Taylor and Francis.

Mark D M (1999) Spatial Representation: A Cognitive View. In: Maguire, D. J. et al. (Ed.), Geographical Information Systems: Principles and Applications. New York: Wiley, pp. 81-89.

Narumalani S, Mishra D R and Rothwell R G (2004) Change detection and landscape metrics for inferring anthropogenic processes in the greater EFMO area. Remote Sensing of Environment, 91, pp. 478-489.

Petit C C and Lambin E F (2002) Impact of data integration technique on historical land-use/land-cover change: Comparing historical maps with remote sensing data in the Belgian Ardennes. Landscape Ecology, 17, pp. 117-132.

Raza A and Kainz W (2001) An Object-Oriented Approach for Modeling Urban Land-Use Changes. URISA Journal, 14 (1), pp. 37-55.

Schöpfer E (2005) Change Detection in Multitemporal Images utilizing Object-based Image Analysis unpublished PhD thesis, Universität Salzburg.

Schöpfer E and Lang S (2006) Object fate analysis – a virtual overlay method for the categorization of object transition and object-based accuracy assessment. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. No. XXXVI-4/C42, Salzburg, Austria.

Schumacher J, Lang S, Tiede D, Hölbling D, Rietze J and Trautner J (2007) Einsatz von GIS und objekt-basierter Analyse von Fernerkundungsdaten in der regionalen Planung: Methoden und erste Erfahrungen aus dem Biotopinformations- und Managementsystem (BIMS) Region Stuttgart. In: Strobl J, Blaschke, T and Griesebner, G (eds.): Angewandte Geoinformatik 2007. Heidelberg: Wichmann, pp. 703-708.

Southworth J, Munroe D and Nagendra H (2004) Land cover change and landscape fragmentation – comparing the utility of continuous and discrete analyses for a western Honduras region. Agriculture, Ecosystems and Environment, 101, pp. 185-205.

Straub BM, Heipke C (2004) Concepts for internal and external evaluation of automatically delineated tree tops. IntArchPhRS. Vol. No. XXXVI-8/W2, Freiburg, pp. 62-65.

Tiede D, Moeller M, Lang S and Hoelbling D (2007) Adapting, splitting and merging cadastral boundaries according to homogenous LULC types derived from SPOT 5 data.- Proc. of the ISPRS Workshop Photogr. Image Analysis, Munich 2007.

Turner M G (1990) Spatial and temporal analysis of landscape patterns. Landscape Ecology, 4, pp. 21-30.

Turner M G, Gardner R H and O'Neill R V (2001) Landscape Ecology in Theory and Practice: Pattern and Process. New York: Springer-Verlag.

Wickam J D, O'Neill R V, Riitters K H, Wade T G and Jones K B (1997) Sensitivity of selected landscape pattern metrics to land-cover misclassification and differences in land-cover composition. Photogrammetric Engineering & Remote Sensing, 63 (4), pp. 397-402.

Index