

# Comments on “On Approximating Euclidean Metrics by Weighted $t$ -Cost Distances in Arbitrary Dimension”

M. Emre Celebi

*Department of Computer Science  
Louisiana State University, Shreveport, LA, USA  
ecelebi@lsus.edu*

Hassan A. Kingravi

*School of Electrical and Computer Engineering  
Georgia Institute of Technology, Atlanta, GA, USA  
kingravi@gatech.edu*

Fatih Celiker

*Department of Mathematics  
Wayne State University, Detroit, MI, USA  
celiker@math.wayne.edu*

---

## Abstract

Mukherjee (Pattern Recognition Letters, vol. 32, pp. 824–831, 2011) recently introduced a class of distance functions called weighted  $t$ -cost distances that generalize  $m$ -neighbor, octagonal, and  $t$ -cost distances. He proved that weighted  $t$ -cost distances form a family of metrics and derived an approximation for the Euclidean norm in  $\mathbb{Z}^n$ . In this note we compare this approximation to two previously proposed Euclidean norm approximations and demonstrate that the empirical average errors given by Mukherjee are significantly optimistic in  $\mathbb{R}^n$ . We also propose a simple normalization scheme that improves the accuracy of his approximation substantially with respect to both average and maximum relative errors.

---

## 1. Introduction

The Minkowski ( $\mathcal{L}_p$ ) metric is inarguably one of the most commonly used quantitative distance (dissimilarity) measures in scientific and engineering applications. The Minkowski distance between two vectors  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  in the  $n$ -dimensional Euclidean space,  $\mathbb{R}^n$ , is given by

$$\mathcal{L}_p(\mathbf{x}, \mathbf{y}) = \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{1/p}. \quad (1)$$

Three special cases of the  $\mathcal{L}_p$  metric are of particular interest, namely,  $\mathcal{L}_1$  (city-block metric),  $\mathcal{L}_2$  (Euclidean metric), and  $\mathcal{L}_\infty$  (chessboard metric). Given the general form (1),  $\mathcal{L}_1$  and  $\mathcal{L}_2$  can be defined in a straightforward fashion, while  $\mathcal{L}_\infty$  is defined as

$$\mathcal{L}_\infty(\mathbf{x}, \mathbf{y}) = \max_{1 \leq i \leq n} |x_i - y_i|.$$

The Minkowski metric enjoys the property of being translation invariant, i.e.,  $\mathcal{L}_p(\mathbf{x}, \mathbf{y}) = \mathcal{L}_p(\mathbf{x} + \mathbf{z}, \mathbf{y} + \mathbf{z})$  for all  $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ . Since in many applications the data space is Euclidean, the most natural choice of metric is  $\mathcal{L}_2$ ,

10 which has the added advantage being isotropic (rotation invariant). For example, when the input vectors stem  
 11 from an isotropic vector field, e.g., a velocity field, the most appropriate choice is to use the  $\mathcal{L}_2$  metric so that all  
 12 vectors are processed in the same way, regardless of their orientation [1]. However,  $\mathcal{L}_2$  has the drawback of a high  
 13 computational cost due to the multiplication and square root operations. As a result,  $\mathcal{L}_1$  and  $\mathcal{L}_\infty$  are often used as  
 14 alternatives. Although these metrics are computationally more efficient, they deviate from  $\mathcal{L}_2$  significantly.

15 Due to the translation invariance of  $\mathcal{L}_p$ , it suffices to consider  $D_p(\mathbf{x}) = \mathcal{L}_p(\mathbf{x}, \mathbf{0})$ , i.e., the distance from the point  
 16  $\mathbf{x}$  to the origin. Therefore, in the rest of the paper, we will consider approximations to  $D_p(\mathbf{x})$  rather than  $\mathcal{L}_p(\mathbf{x}, \mathbf{y})$ .

17 Let  $\tilde{D}$ , defined on  $\mathbb{R}^n$ , be an approximation to  $D_2$  (Euclidean norm). We assume that  $\tilde{D}$  is a continuous and  
 18 absolutely homogeneous function. Recall that  $\tilde{D}$  is called absolutely homogeneous (of degree one) if  $\tilde{D}(\lambda\mathbf{x}) =$   
 19  $|\lambda|\tilde{D}(\mathbf{x}) \quad \forall \lambda \in \mathbb{R}, \forall \mathbf{x} \in \mathbb{R}^n$ .

20 We note that all variants of  $\tilde{D}$  we consider in this paper satisfy these assumptions. As a measure of the quality  
 21 of the approximation of  $\tilde{D}$  to  $D_2$  we define the maximum relative error (MRE) as

$$\varepsilon_{\max}^{\tilde{D}} = \sup_{\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}} \frac{|\tilde{D}(\mathbf{x}) - D_2(\mathbf{x})|}{D_2(\mathbf{x})}. \quad (2)$$

22 Using the homogeneity of  $D_2$  and  $\tilde{D}$ , (2) can be written as

$$\varepsilon_{\max}^{\tilde{D}} = \sup_{\mathbf{x} \in S_2^{n-1}} |\tilde{D}(\mathbf{x}) - 1|, \quad (3)$$

23 where  $S_2^{n-1} = \{\mathbf{x} \in \mathbb{R}^n : D_2(\mathbf{x}) = 1\}$  is the unit hypersphere of  $\mathbb{R}^n$  with respect to the Euclidean norm. Furthermore,  
 24 by the continuity of  $\tilde{D}$ , we can replace the supremum with maximum in (3) and write

$$\varepsilon_{\max}^{\tilde{D}} = \max_{\mathbf{x} \in S_2^{n-1}} |\tilde{D}(\mathbf{x}) - 1|. \quad (4)$$

25 We will use (4) as the definition of MRE throughout.

26 Mukherjee [2] recently introduced a class of distance functions called weighted  $t$ -cost distances that generalize  
 27  $m$ -neighbor [3], octagonal [4], and  $t$ -cost [5] distances. He proved that weighted  $t$ -cost distances form a family of  
 28 metrics and derived an approximation for the Euclidean norm in  $\mathbb{Z}^n$ . Here we briefly review the  $t$ -cost norm.

29 The  $t$ -cost norm [5] defines two points in the rectangular grid as neighbors when their respective hypercubes  
 30 (or hypervoxels) share a hyperplane of any dimension. The cost associated with these points can be at most  $t$ ,  
 31  $1 \leq t \leq n$ , such that if two consecutive points on a shortest path share a hyperplane of dimension  $r$ , the distance  
 32 between them is taken as  $\min(t, n - r)$ . There are  $n$  distinct  $t$ -cost norms defined by

$$D_t(\mathbf{x}) = \sum_{i=1}^t x_{(i)}, \quad 1 \leq t \leq n$$

33 where  $x_{(i)}$  is the  $i$ -th absolute largest component of  $\mathbf{x}$ , i.e.,  $(x_{(1)}, x_{(2)}, \dots, x_{(n)})$  is a permutation of  $(|x_1|, |x_2|, \dots, |x_n|)$   
 34 such that  $x_{(1)} \geq x_{(2)} \geq \dots \geq x_{(n)}$ . The MRE of this norm is given by [5]

$$\varepsilon_{\max}^{D_t} = \max \left( \sqrt{t} - 1, 1 - \frac{t}{\sqrt{n}} \right).$$

35 Mukherjee generalized the  $t$ -cost norm as follows [2]:

$$D_M(\mathbf{x}) = \max_{1 \leq t \leq n} \{w_t D_t(\mathbf{x})\},$$

where  $w_t$ 's are non-negative real constants. Based on this weighted norm, he then derived an approximation for  $D_2$  using the following weight assignment:  $w_t = 1/\sqrt{t}$  for  $1 \leq t \leq n$ . Note that  $D_M$  consistently underestimates  $D_2$  and the corresponding MRE is given by [2]

$$\varepsilon_{\max}^{D_M} = 1 - \frac{1}{\sqrt{\sum_{i=1}^n (\sqrt{i} - \sqrt{i-1})^2}}. \quad (5)$$

In a recent study [6], we examined various Euclidean norm approximations in detail and compared their average and maximum errors using numerical simulations. Here we show that two of those approximations, namely Barni *et al.*'s norm [1, 7] and Seol and Cheun's norm [8], are viable alternatives to  $D_M$ .

Barni *et al.* [1, 7] formulated a generic approximation for  $D_2$  as

$$D_B(\mathbf{x}) = \delta \sum_{i=1}^n \alpha_i x_{(i)},$$

where  $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n)$  and  $\delta > 0$  are approximation parameters. Note that a non-increasing ordering and strict positivity of the component weights, i.e.,  $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_n > 0$  is a necessary and sufficient condition for  $D_B$  to define a norm [7].

Barni *et al.* showed that the minimization of (4) is equivalent to determining the weight vector  $\boldsymbol{\alpha}$  and the scale factor  $\delta$  that solve the following minimax problem:

$$\min_{\boldsymbol{\alpha}, \delta} \max_{\mathbf{x} \in V} |D_B(\mathbf{x}) - 1|,$$

where  $V = \{\mathbf{x} \in \mathbb{R}^n : x_1 \geq x_2 \geq \dots \geq x_n \geq 0, D_2(\mathbf{x}) = 1\}$ . The optimal solution and its MRE are given by

$$\alpha_i^* = \sqrt{i} - \sqrt{i-1}, \quad \delta^* = \frac{2}{1 + \sqrt{\sum_{i=1}^n \alpha_i^{*2}}}, \quad \varepsilon_{\max}^{D_B} = 1 - \delta^*. \quad (6)$$

Note the striking similarity between (5) and (6). Interestingly, a similar but less rigorous approach had been published earlier by Ohashi [9]. It should also be noted that several authors approached the problem from a Euclidean distance transform perspective and derived similar approximations for the 2- and 3-dimensional cases, see for example [10] and [11]. Furthermore, computation of weighted (Chamfer) distances in arbitrary dimensions on general point lattices is discussed in [12].

More recently, Seol and Cheun [8] proposed an approximation of the form

$$D_{a,b}(\mathbf{x}) = aD_\infty(\mathbf{x}) + bD_1(\mathbf{x}), \quad (7)$$

where  $a$  and  $b$  are strictly positive parameters to be determined by solving the following  $2 \times 2$  linear system

$$\begin{aligned} aE(D_\infty^2) + bE(D_\infty D_1) &= E(D_2 D_\infty), \\ aE(D_\infty D_1) + bE(D_1^2) &= E(D_2 D_1), \end{aligned}$$

56 where  $E(\cdot)$  is the expectation operator.

57 Seol and Cheun estimated the optimal values of  $a$  and  $b$  using 100,000  $n$ -dimensional vectors whose components  
 58 are independent and identically distributed, standard Gaussian random variables. In [6], we demonstrated that a  
 59 fixed number of samples from the unit hypersphere gives biased estimates for the MRE. The basic reason behind  
 60 this is the fact that a fixed number of samples fail to suffice as the dimension of the space increases.

61 It is easy to see that  $D_B$  and  $D_{a,b}$  fit into the general form

$$\tilde{D}(\mathbf{x}) = \sum_{i=1}^n w_i x_{(i)},$$

62 which is a weighted  $D_1$  norm. For  $D_B$  the weights are  $w_1 = \delta^*$  and  $w_{i \neq 1} = \delta^* \alpha_i^*$ , whereas for  $D_{a,b}$  they are  
 63  $w_1 = a + b$  and  $w_{i \neq 1} = b$ . Clearly,  $D_B$  has a more elaborate design in which each component is assigned a weight  
 64 proportional to its ranking (absolute value). However, this weighting scheme also presents a drawback in that a full  
 65 ordering of the component absolute values is required.

66  $D_B$  and  $D_{a,b}$  can also be written as linear combinations of the  $D_1$  and  $D_\infty$  norms, as in (7).  $D_1$  overes-  
 67 timates the  $D_2$  norm, whereas  $D_\infty$  underestimates it [13]. Therefore, it is natural to expect a suitable linear  
 68 combination of  $D_1$  and  $D_\infty$  to give an approximation to  $D_2$  better than either of them [14]. Note that Rosen-  
 69 feld and Pfaltz [15] obtained a 2-dimensional approximation by combining  $D_1$  and  $D_\infty$  nonlinearly as follows:  
 70  $\tilde{D}(\mathbf{x}) = \max(\lfloor 2(D_1(\mathbf{x}) + 1)/3 \rfloor, D_\infty(\mathbf{x}))$ .

## 71 2. Comparison of the Euclidean Norm Approximations

72 Due to their formulations, the MREs for  $D_M$  and  $D_B$  can be calculated analytically using (5) and (6), respec-  
 73 tively. In Figure 1 we plot the theoretical errors for these norms for  $n \leq 100$ . It can be seen that  $D_B$  is not only  
 74 more accurate than  $D_M$ , but also it scales significantly better.

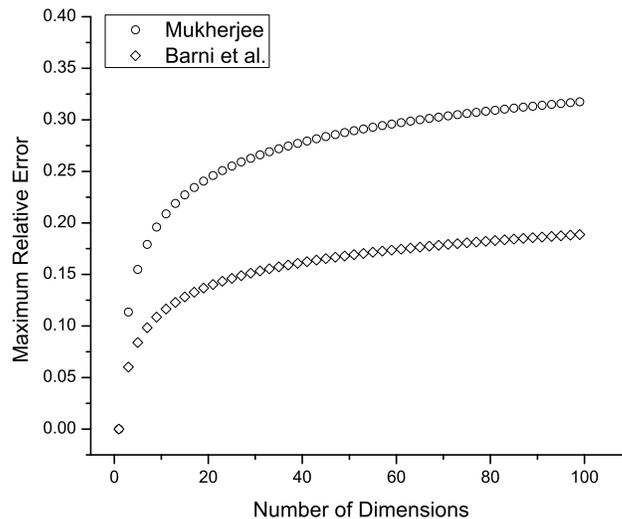


Figure 1: Maximum relative errors for  $D_M$  and  $D_B$

Table 1: Operation counts for the norms

Norm	ABS	COMP	ADD	MULT	SQRT
$D_\infty$	$n$	$n - 1$	0	0	0
$D_1$	$n$	0	$n - 1$	0	0
$D_2$	0	0	$n - 1$	$n$	1
$D_B$	$n$	$\mathcal{O}(n \log n)$	$n - 1$	$n$	0
$D_{a,b}$	$n$	$n - 1$	$n$	2	0
$D_M$	$n$	$\mathcal{O}(n \log n)$	$n - 1$	$n$	0

75 The operation counts for each norm are given in Table 1 (**ABS**: absolute value, **COMP**: comparison, **ADD**:  
76 addition, **MULT**: multiplication, **SQRT**: square root). The following conclusions can be drawn:

77  $\triangleright D_B$  and  $D_M$  have the highest computational cost due to the fact that they require sorting of the absolute  
78 values of the vector components.

79  $\triangleright D_{a,b}$  has the lowest computational cost among the approximate norms. A significant advantage of this norm  
80 is that it requires only two multiplications regardless of the value of  $n$ .

81  $\triangleright D_{a,b}$  can be used to approximate  $D_2^2$  (squared Euclidean norm) using an extra multiplication. On the other  
82 hand, the computational cost of  $D_B$  ( $D_M$ ) is higher than that of  $D_2^2$  due to the extra absolute value and  
83 sorting operations involved.

84 In Table 2 we display the percentage average and maximum errors for  $D_{a,b}$ ,  $D_B$ , and  $D_M$  for  $n \leq 8$ . Average  
85 relative error (ARE) is defined as

$$\varepsilon_{\text{avg}}^{\tilde{D}} = \frac{1}{|S|} \sum_{\mathbf{x} \in S} |\tilde{D}(\mathbf{x}) - 1|,$$

86 where  $S$  is a finite subset of the unit hypersphere  $S_2^{n-1}$ , and  $|S|$  denotes the number of elements in  $S$ . An efficient  
87 way to pick a random point on  $S_2^{n-1}$  is to generate  $n$  independent Gaussian random variables  $x_1, x_2, \dots, x_n$  with  
88 zero mean and unit variance. The distribution of the unit vectors

$$\left\{ \mathbf{y} = (y_1, y_2, \dots, y_n) : y_i = x_i / \left( \sum_{j=1}^n x_j^2 \right)^{1/2}, \quad i = 1, 2, \dots, n \right\}$$

89 will then be uniform over the surface of the hypersphere [16]. For each approximate norm, the ARE and MRE  
90 values were calculated over an increasing number of points,  $2^{20}, 2^{21}, \dots$  (that are uniformly distributed on the  
91 hypersphere) until the error values converge, i.e., the error values do not differ by more than  $\epsilon = 10^{-5}$  in two  
92 consecutive iterations.

93 In Table 2, the error values under the column “ $D_M(\mathbb{R}^n)$ ” were obtained using the aforementioned iterative  
94 scheme, whereas those under the column “ $D_M(\mathbb{Z}^n)$ ” are taken from [2]. Motivated by the fact that  $D_M$  consistently  
95 underestimates  $D_2$ , we also experimented with a normalized form of this approximate norm given by  $D_{\hat{M}}(\mathbf{x}) =$   
96  $D_M(\mathbf{x})/\delta^*$ . Note that  $\delta^* < 1$  for  $n \geq 2$  (6).

97 Note that for  $D_M$  and  $D_B$ , two types of maximum error were considered: empirical maximum error ( $\text{MRE}_e$ ),  
98 which is calculated numerically over  $S$  and the theoretical maximum error ( $\text{MRE}_t$ ), which is calculated analytically  
99 using (5) and (6), respectively.

Table 2: Percentage average and maximum errors for the approximate Euclidean norms

	$D_{a,b}$		$D_B$			$D_{\widehat{M}}$		$D_M$				
								$\mathbb{R}^n$		$\mathbb{Z}^n$		
$n$	ARE	MRE <sub>e</sub>	ARE	MRE <sub>e</sub>	MRE <sub>t</sub>	ARE	MRE <sub>e</sub>	ARE	MRE <sub>e</sub>	ARE	MRE <sub>e</sub>	MRE <sub>t</sub>
2	2.00	5.25	2.41	3.96	3.96	2.48	4.12	2.55	7.61	2.40	7.61	7.61
3	2.39	9.98	3.00	6.02	6.02	2.97	6.40	4.14	11.35	3.63	11.35	11.35
4	2.57	13.64	3.44	7.39	7.39	3.28	7.97	5.21	13.75	4.29	13.75	13.75
5	2.68	16.59	3.77	8.39	8.39	3.53	9.16	5.98	15.47	4.65	15.46	15.49
6	2.73	18.88	4.01	9.19	9.19	3.73	10.12	6.55	16.80	4.85	16.79	16.83
7	2.76	20.67	4.18	9.84	9.84	3.92	10.91	7.00	17.90	5.00	17.86	17.92
8	2.77	21.92	4.31	10.39	10.39	4.10	11.59	7.35	18.78	5.04	18.75	18.82

By examining Table 2, the following observations can be made regarding the maximum error:

- ▷ The most accurate approximation is  $D_B$ . This is because this norm is designed to minimize the maximum error.
- ▷ The proposed normalization is quite effective since the resulting norm,  $D_{\widehat{M}}$ , is, on the average, only 8.6% less accurate than  $D_B$ , whereas both  $D_M(\mathbb{R}^n)$  and  $D_M(\mathbb{Z}^n)$  are, on the average, about 85% less accurate than  $D_B$ .
- ▷ The least accurate approximations are  $D_M$  and  $D_{a,b}$  for  $n \leq 4$  and  $n > 4$ , respectively.
- ▷ As  $n$  is increased, the error increases in all approximations. However, as can also be seen in Fig. 1, the error grows faster in some approximations than others.
- ▷ For  $D_B$ , the empirical and theoretical errors agree almost perfectly in all cases, which demonstrates the validity of the presented iterative error calculation scheme. As for  $D_M$ , the agreement in each case is close, but not as close as that observed in  $D_B$ . We have confirmed that using a smaller convergence threshold ( $\epsilon$ ) alleviates this problem at the expense of increased computational cost.

On the other hand, with respect to average error we can see that:

- ▷  $D_{a,b}$  is the most accurate approximation. This is because this norm is designed to minimize the average error.
- ▷  $D_M(\mathbb{R}^n)$  and  $D_M(\mathbb{Z}^n)$  are the least accurate approximations. Furthermore, the errors given by Mukherjee are lower than those that we obtained (over  $\mathbb{R}^n$ ), and the discrepancy between the outcomes of the two error calculation schemes increases as  $n$  is increased. The optimistic average error values given by Mukherjee are due to the fact that his approximation was primarily intended for use in digital geometry and hence the calculations were performed in  $\mathbb{Z}^n$  (rather than  $\mathbb{R}^n$ ) using a very small number of points ranging from 32 to 512 [2]. In fact, Mukherjee used progressively fewer points with increasing  $n$  to calculate the error values. In [6], we demonstrated that more points are required in higher dimensions to obtain unbiased error estimates.

In the calculation of  $D_{\widehat{M}}$ , we assumed that the optimal scaling factor for  $D_M$  is the same as that of  $D_B$ , i.e.,  $\delta^*$ .

In order to check this assumption, we performed a one-dimensional grid search over  $[\delta^*, 1]$  for each  $n$  value. The results are shown in Table 3. It can be seen that:

Table 3: Percentage average and maximum errors for  $D_{\widehat{M}}$

$n$	$D_{\widehat{M}}^{\delta^*}$			$D_{\widehat{M}}^{\widehat{\delta}}$		
	ARE	MRE <sub>e</sub>	$\delta^*$	ARE	MRE <sub>e</sub>	$\widehat{\delta}$
2	2.48	4.12	0.960434	2.41	3.96	0.961971
3	2.97	6.40	0.939809	2.79	6.02	0.943192
4	3.28	7.97	0.926150	2.99	7.39	0.931336
5	3.53	9.16	0.916059	3.13	8.40	0.922654
6	3.73	10.12	0.908117	3.23	9.18	0.915927
7	3.92	10.91	0.901603	3.31	9.84	0.910619
8	4.10	11.59	0.896101	3.40	10.39	0.905850

- 125  $\triangleright D_{\widehat{M}}^{\widehat{\delta}}$  is significantly more accurate than  $D_{\widehat{M}}^{\delta^*}$  with respect to both ARE and MRE.
- 126  $\triangleright D_{\widehat{M}}^{\widehat{\delta}}$  and  $D_B$  have almost identical MREs. Since  $D_B$  is analytically optimized for the maximum error it can
- 127 be concluded that  $D_{\widehat{M}}^{\widehat{\delta}}$  can reach the same optimality by means of a suitable scaling factor.
- 128  $\triangleright$  Interestingly,  $D_{\widehat{M}}^{\widehat{\delta}}$  is more accurate than  $D_B$  with respect to ARE. This could be due to the fact that the two
- 129 approximations take different paths towards minimizing the MRE.

### 130 3. Conclusions

131 In this paper, we examined the weighted  $t$ -cost norm recently proposed by Mukherjee [2] with respect to its  
 132 ability to approximate the Euclidean norm in  $\mathbb{R}^n$ . We evaluated the average and maximum errors of this norm using  
 133 numerical simulations and compared the results to those of two other well-known Euclidean norm approximations.  
 134 The results demonstrated that, because it was designed for digital geometry applications in  $\mathbb{Z}^n$ , the original weighted  
 135  $t$ -cost norm is not particularly suited to approximate the Euclidean norm in  $\mathbb{R}^n$ . It is also shown, however, that  
 136 when normalized with an appropriate scaling factor, Mukherjee’s norm becomes competitive with an analytically  
 137 optimized approximation with respect to both average and maximum relative errors.

### 138 4. Acknowledgments

139 This work was supported by grants from the Louisiana Board of Regents (LEQSF2008-11-RD-A-12) and US  
 140 National Science Foundation (0959583, 1117457). The authors are grateful to the anonymous reviewers for their  
 141 insightful suggestions and constructive comments that improved the quality and presentation of this paper.

### 142 References

143 [1] M. Barni, F. Bartolini, F. Buti, and V. Cappellini, “Optimum Linear Approximation of the Euclidean Norm to  
 144 Speed up Vector Median Filtering,” Proceedings of the 2nd IEEE International Conference on Image Processing,  
 145 pp. 362–365, 1995.

146 [2] J. Mukherjee, “On Approximating Euclidean Metrics by Weighted  $t$ -Cost Distances in Arbitrary Dimension,”  
 147 Pattern Recognition Letters, vol. 32, no. 6, pp. 824–831, 2011.

- 148 [3] P.P. Das, P.P. Chakrabarti, and B. Chatterji, “Generalized Distances in Digital Geometry,” *Information Sci-*  
 149 *ences*, vol. 42, no. 1, pp. 51–67, 1987.
- 150 [4] P.P. Das, P.P. Chakrabarti, and B. Chatterji, “Distance Functions in Digital Geometry,” *Information Sciences*,  
 151 vol. 42, no. 2, pp. 113–136, 1987.
- 152 [5] P.P. Das, J. Mukherjee, and B.N. Chatterji, “The t-Cost Distance in Digital Geometry,” *Information Sciences*,  
 153 vol. 59, no. 1, pp. 1–20, 1992.
- 154 [6] M.E. Celebi, F. Celiker, and H. Kingravi, “On Euclidean Norm Approximations,” *Pattern Recognition*, vol.  
 155 44, no. 2, pp. 278–283, 2011.
- 156 [7] M. Barni, F. Buti, F. Bartolini, and V. Cappellini, “A Quasi-Euclidean Norm to Speed up Vector Median  
 157 Filtering,” *IEEE Transactions on Image Processing*, vol. 9, no. 10, pp. 1704–1709, 2000.
- 158 [8] C. Seol and K. Cheun, “A Low Complexity Euclidean Norm Approximation,” *IEEE Transactions on Signal*  
 159 *Processing*, vol. 56, no. 4, pp. 1721–1726, 2008.
- 160 [9] Y. Ohashi, “Fast Linear Approximations of Euclidean Distance in Higher Dimensions,” in P. Heckbert (Ed.),  
 161 *Graphics Gems IV*, Academic Press, 1994.
- 162 [10] G. Borgefors, “Distance Transformations in Digital Images,” *Computer Vision, Graphics, and Image Processing*,  
 163 vol. 34, no. 3, pp. 344–371, 1986.
- 164 [11] B.J.H. Verwer, “Local Distances for Distance Transformations in Two and Three Dimensions,” *Pattern Recog-*  
 165 *niton Letters*, vol. 12, no. 11, pp. 671–682, 1991.
- 166 [12] C. Fouard, R. Strand, and G. Borgefors, “Weighted Distance Transforms Generalized to Modules and Their  
 167 Computation on Point Lattices,” *Pattern Recognition*, vol. 40, no. 9, pp. 2453–2474, 2007.
- 168 [13] D. Chaudhuri, C.A. Murthy, and B.B. Chaudhuri, “A Modified Metric to Compute Distance,” *Pattern Recog-*  
 169 *niton*, vol. 25, no. 7, pp. 667–677, 1992.
- 170 [14] F. Rhodes, “On the Metrics of Chaudhuri, Murthy and Chaudhuri,” *Pattern Recognition*, vol. 28, no. 5, pp.  
 171 745–752, 1995.
- 172 [15] A. Rosenfeld and J.L. Pfaltz, “Distance Functions on Digital Pictures,” *Pattern Recognition*, vol. 1, no. 1, pp.  
 173 33–61, 1968.
- 174 [16] M.E. Muller, “A Note on a Method for Generating Points Uniformly on N-Dimensional Spheres,” *Communi-*  
 175 *cations of the ACM*, vol. 2, no. 4, pp. 19–20, 1959.