ORIENTATION ESTIMATION BASED ON GAUSS-NEWTON METHOD AND IMPLEMENTATION OF A QUATERNION COMPLEMENTARY FILTER

D. Comotti

Abstract

This work will show new algorithms developed during the month following our project release (*http://code.google.com/p/9dof-orientation-estimation/*). Indeed, when report [1] was released, there were still a lot of work to do, especially for Complementary Filter and observation estimation algorithms. In this direction, an alternative method to the Gradient Descent algorithm to get quaternions from accelerometer and magnetometer measurements will be presented in this report. Moreover, an implementation of a quaternion based Complementary Filter will be shown. The whole work is integrated with results.

Part I

Observation Estimation Based on Gauss-Newton Method

In this application context quaternions have an important role; rotations can be expressed simply by mean of a vector and without trigonometric functions. However, it is not trivial to get the right quaternion configuration from the sensor measurements. Using gyroscope datas, the quaternion variation can be computed as follows [2]:

$$\dot{\vec{q}_t} = \frac{1}{2} \cdot \vec{q}_{t-1} \otimes \vec{\omega}_t \tag{1}$$

where $\vec{q}_{t-1}$ is the latest computed quaternion and $\vec{\omega}_t = \begin{bmatrix} 0 & \omega_x(t) & \omega_y(t) & \omega_z(t) \end{bmatrix}^T$ is the angular rates vector at the current time.

Computing quaternions from accelerometer and magnetometer measurements need a different approach, because datas are not directly related to angles. In [2, 1], it is shown how to compute angles minimizing a function by mean of Gradient Descent method. In this section, a minimization algorithm briefly discussed in [3] and based on Gauss-Newton method, will be presented.

Let $\vec{z} = \begin{bmatrix} A_x & A_y & A_z & M_x & M_y & M_z \end{bmatrix}^T$ be the accelerometer and magnetometer measurements vector. It can be expressed in the Earth frame, $\vec{z_0} = \begin{bmatrix} 0 & 0 & 1 & \tilde{M_x} & \tilde{M_y} & \tilde{M_z} \end{bmatrix}^T$ (it is assumed that the configuration of "zero" position has accelerometer z axis parallel with gravity vector and upward), where $\tilde{M_i}$ is the magnetic component after the magnetic compensation process in "zero" position, and with respect of the body (IMU) frame, $\vec{z_t} = \begin{bmatrix} A_x(t) & A_y(t) & A_z(t) & M_x(t) & M_y(t) & M_z(t) \end{bmatrix}^T$. Now, if we could have the real and correct angles values of the IMU, then:

$$\epsilon = \vec{z_0} - R_t \times \vec{z_t} = 0 \tag{2}$$

where R_t is the Direction Cosine Matrix (DCM) at the current time. Actually, since quaternions we trat with are just an estimation, the best we can do is to minimize ϵ using the latest available angles or quaternion. In quaternion case (as our case), R_t matrix is expressed as follows:

$$R_t = \begin{bmatrix} q_4^2+q_1^2-q_2^2-q_3^2 & 2\cdot(q_1\cdot q_2-q_3\cdot q_4) & 2\cdot(q_1\cdot q_3+q_2\cdot q_4) \\ 2\cdot(q_1\cdot q_2+q_3\cdot q_4) & q_4^2+q_2^2-q_1^2-q_3^2 & 2\cdot(q_2\cdot q_3-q_4\cdot q_1) \\ 2\cdot(q_1\cdot q_3-q_2\cdot q_4) & 2\cdot(q_3\cdot q_2+q_1\cdot q_4) & q_4^2+q_3^2-q_1^2-q_2^2 \end{bmatrix} \quad (3)$$

Note that quaternion convention is the same of [3], $\vec{q} = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 \end{bmatrix}$ with q_4 the real component, and it is different from the convention used in [1, 2].

Since we are considering two contributions to orientation estimation, privided by accelerometer and magnetometer, ϵ computation must be rewritten as:

$$\epsilon = \vec{z_0} - M_t \times \vec{z_t} \quad (4)$$

where M_t is the current time DCM which rotates both the measurements vectors:

$$M_t = \begin{bmatrix} R_t & 0 \\ 0 & R_t \end{bmatrix} \quad (5)$$

As shown in [4], given a vector of n functions in m variables, $\mathbf{f}(x_1, \ldots, x_m)$, Gauss-Newton method consists in the following optimization step:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \left[J_k^T \cdot J_k\right]^{-1} \cdot J_k^T \cdot \mathbf{f} \quad (6)$$

where J_k is the Jacobian of $\mathbf{f}$ calculated in $\mathbf{x}_k$.
In our case the function $\mathbf{f}$ is $\epsilon = \mathbf{f}\left(\vec{q}(t-1), \vec{z_0}, \vec{z_t}\right)$ in case of first iteration, where $\vec{q}(t-1)$ is the latest iteration quaternion provided in the previous time $(t-1)$ by the same algorithm, and $\epsilon = \mathbf{f}\left(\vec{q}_{k-1}(t), \vec{z_0}, \vec{z_t}\right)$ in case of iterations following the first (k is iteration index). Then the Jacobian of the $k-th$ iteration is the following:

$$J_t(\vec{q_k}(t)) = \frac{\delta(\epsilon)}{\delta \vec{q_k}(t)} = -\left[\left(\frac{\delta M}{\delta q_1}\cdot\vec{z_t}\right) \left(\frac{\delta M}{\delta q_2}\cdot\vec{z_t}\right) \left(\frac{\delta M}{\delta q_3}\cdot\vec{z_t}\right) \left(\frac{\delta M}{\delta q_4}\cdot\vec{z_t}\right) \right] \quad (7)$$

which, after performing the derivatives and matrix products, can be rewritten as follows:

$$J_t(\vec{q_k}(t)) =$$

$$-\begin{bmatrix} (2q_1A_x+2q_2A_y+2q_3A_z) & (-2q_2A_x+2q_1A_y+2q_4A_z) & (-2q_3A_x-2q_4A_y+2q_1A_z) & (2q_4A_x-2q_3A_y+2q_2A_z) \\ (2q_2A_x-2q_1A_y-2q_4A_z) & (2q_1A_x+2q_2A_y+2q_3A_z) & (2q_4A_x-2q_3A_y+2q_2A_z) & (2q_3A_x+2q_4A_y-2q_1A_z) \\ (2q_3A_x+2q_4A_y-2q_1A_z) & (-2q_4A_x+2q_3A_y-2q_2A_z) & (2q_1A_x+2q_2A_y+2q_3A_z) & (-2q_2A_x+2q_1A_y+2q_4A_z) \\ (2q_1M_x+2q_2M_y+2q_3M_z) & (-2q_2M_x+2q_1M_y+2q_4M_z) & (-2q_3M_x-2q_4M_y+2q_1M_z) & (2q_4M_x-2q_3M_y+2q_2M_z) \\ (2q_2M_x-2q_1M_y-2q_4M_z) & (2q_1M_x+2q_2M_y+2q_3M_z) & (2q_4M_x-2q_3M_y+2q_2M_z) & (2q_3M_x+2q_4M_y-2q_1M_z) \\ (2q_3M_x+2q_4M_y-2q_1M_z) & (-2q_4M_x+2q_3M_y-2q_2M_z) & (2q_1M_x+2q_2M_y+2q_3M_z) & (-2q_2M_x+2q_1M_y+2q_4M_z) \end{bmatrix}$$

Note that the quaternion vector $\vec{q_k}(t)$ refers to the $k-th$ iteration at the current time t (during each time an acquisition from the IMU is performed, and more than one iteration per time could be computed, in order to faster reach the minimum point).
Finally, the next iteration quaternion is computed as follows:

$$\vec{q}_{k+1}(t) = \vec{q_k}(t) - \left[J_t(\vec{q_k}(t))^T \cdot J_t(\vec{q_k}(t))\right]^{-1} \cdot$$

$$\cdot J_t(\vec{q_k}(t))^T \cdot (\vec{z_0} - M_t(\vec{q_k}) \cdot \vec{z_t}) \quad (8)$$

With regards to experimental results and comparison with Gradient Descent algorithm, as theory claims, figure 1 and figure 2shows that Gauss-Newton is faster, it doesn't require many iterations to reach the optimum point and it doesn't introduce the "zig-zag" effect.

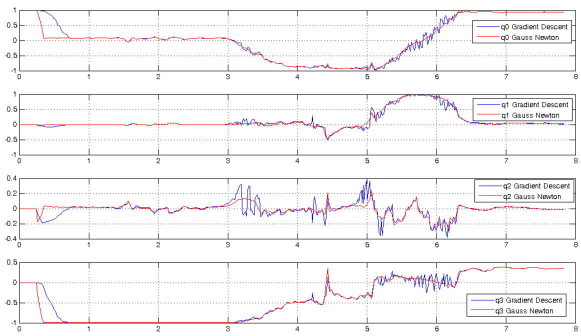

Figure 1: Guass-Newton and Gradient Descent methods comparison. In this example, 3 iterations per acqusition are computed for Gauss-Newton method, while 10 iterations per acquisition are executed for Gradient Descent method.

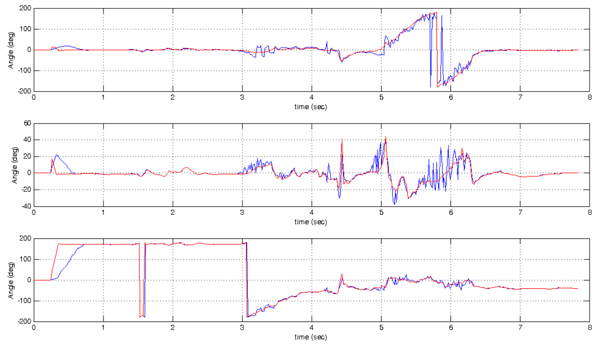

Figure 2: Angles related to the quaternions shown in figure 1

Part II

Quaternion Complementary Filter

Complementary Filter is an easy way to estimate orientation of an IMU using its sensors datas. As shown in [1], the usage of Euler angles can leads to some problems, so it would be nice to perform the filter using quaternions [1]. In this section, the convention used to represent quaternion is the following: $\vec{q} = \begin{bmatrix} q_0 & q_1 & q_2 & q_3 \end{bmatrix}$, where q_0 is the real component.

Complementary Filter fuses the gyroscope based estimation with the accelerometer and magnetometer one using two gain factors which sum to 1, chosen in such a way that lacks of each fusion component are avoided. With regards to the gyroscope, current time quaternion is computed by the following equation:

$$\vec{q}_{Gyro}(t) = \vec{q}_{Gyro}(t-1) + \dot{\vec{q}}_{Gyro}(t) \cdot \delta t \qquad (9)$$

where $\dot{\vec{q}}_{Gyro}(t)$ is obtained from equation 1.

Measurements from accelerometer and magnetometer are used to get the quaternion $\vec{q}(t)$, by mean of the Gradient Descent or the Gauss-Newton methods. In order to prevent the estimation from wrong values, magnetic compensation can be performed; using the latest quaternion computed, $\vec{q}(t-1)$, the magnetic refence in the earth frame is obtained as follows:

$$^{E}\vec{h}_t = \vec{q}(t-1) \otimes^{S} \vec{m}_t \otimes \vec{q}(t)^*$$

where $\vec{q}(t)*$ indicates the conjugate of $\vec{q}(t)$ and $\vec{m}_t$ is the magnetometer measurements at time t. Since the magnetic reference doesn't have a component aligned with y axis, it is derived as follows:

$$^{E}\vec{b}_t = \begin{bmatrix} 0 & \sqrt{h_x^2 + h_y^2} & 0 & h_z \end{bmatrix}$$

This will be the magnetic reference passed to the algorithm (Gradient Descent or Gauss-Newton) which compute the quaternion from accelerometer and magnetometer measurements (the accelerometer reference is just the $\begin{bmatrix} 0 & 0 & 1 \end{bmatrix}$ vector).

When quaternions of both the Complementary Filter fusion components are available, the filtered quaternion is obtained by the following equation:

$$\hat{\vec{q}}_t = K \cdot \vec{q}_{Gyro}(t) + (1-K) \cdot \vec{q}(t) \qquad (10)$$

where K is the filter gain, $0 \leq K \leq 1$.

[1] I suggest you to have a look to report [1] to understand Complementary Filter at all. Although that report is written in italian language, block diagrams and equations could be useful.

In order to have a feedback of the filtered quaternion, equation 9 and 10 must be substituted with the following ones:

$$\vec{q}_{GyroFilt}(t) = \vec{q}_{Filt}(t-1) + \dot{\vec{q}}_{Gyro}(t) \cdot \delta t \qquad (11)$$

$$\hat{\vec{q}}_t = K \cdot \vec{q}_{GyroFilt}(t) + (1-K) \cdot \vec{q}(t) \qquad (12)$$

In this way, divergences due to gyroscope bias are avoided, while dynamic errors related to magnetometer and accelerometer during IMU movements are prevented with a large enough value of K. As shown in [5], the time constant of the system is obtained by the following relation:

$$\tau = \frac{K \cdot \delta t}{1-K}$$

where δt is the sampling period. With respect to accelerometer and magnetometer datas, τ represents the time constant of a LP filter, while it acts quite like an HP filter for gyroscope measurements.

During the experimental analisys both the Gradient Descent and Gauss-Newton algorithm has been considered to estimate quaternions from accelerometer and magnetometers measurements. In figure 3, 4, 5 and 6 are shown quaternions and angles trends. The chosen value for K has been 0.98, as in [1].

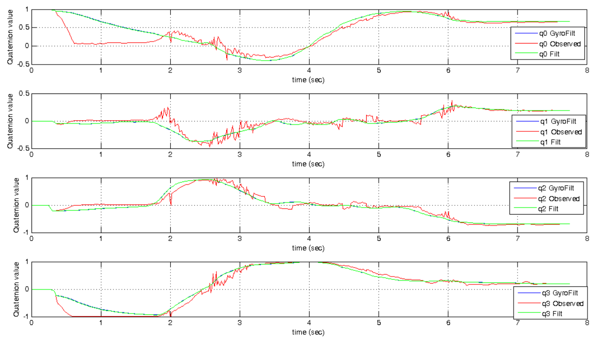

Figure 3: Quaternion Complementary Filter with observation estimation based on Gradient Descent: trend of quaternions. As can be seen, the filtered values take a little time to reach the convergence to initial value. This is due to the time constant τ.

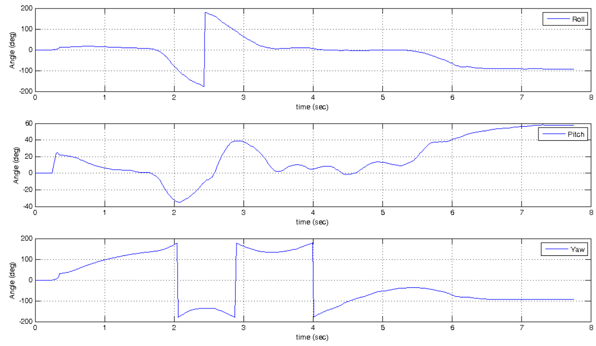

Figure 4: Quaternion Complementary Filter with observation estimation based on Gradient Descent: angles estimated.

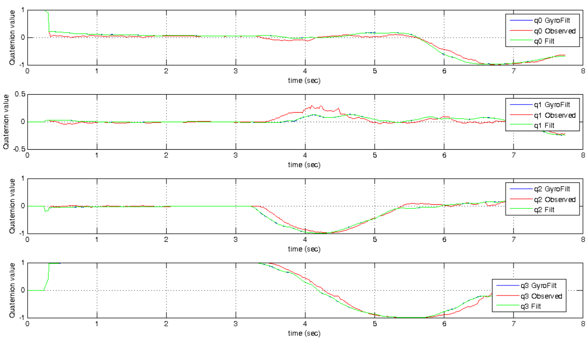

Figure 5: Quaternion Complementary Filter with observation estimation based on Gauss-Newton method: trend of quaternions. As can be seen, since Gauss-Newton reach the optimum point faster than Gradient Descent, during some $(2 \div 5)$ initial acquisition the values of $\vec{q}(t)$ could also be assigned to $\vec{q}_{GyroFilt}(t)$ and $\hat{\vec{q}}_t$, in order to have a good convergence to initial values.

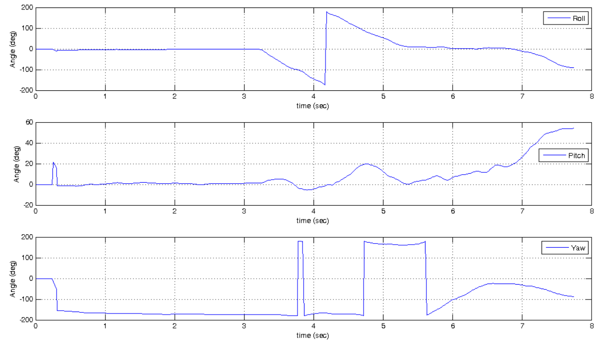

Figure 6: Quaternion Complementary Filter with observation estimation based on Gauss-Newton method: angles estimated.

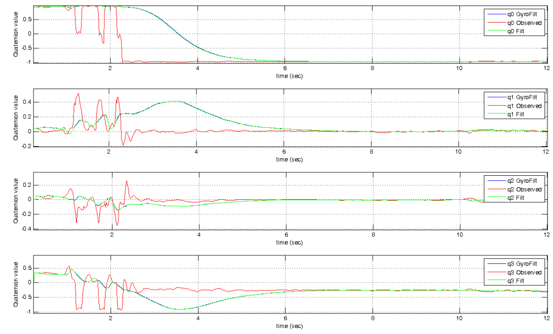

Figure 7: Trend of quaternions during an experiment with high acceleration perturbations. As can be seen, quaternions, after the perturbation, change configuration to the dual one.

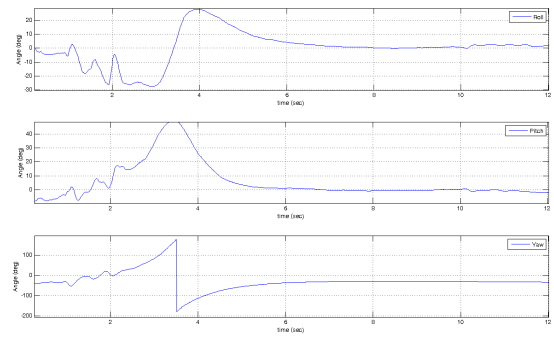

Figure 8: Trend of angles during the same experiment of figure 7.

It has to be noticed that quaternions estimation by mean of minimization algorithms suffers of some problems when the IMU undergoes high acceleration (i.e. strong vibrations). Indeed, since a quaternion has a dual representation, it may happens that, after a strong movement (i.e. acceleration stronger than $2\,g$ amplitude) during which observation estimation is not reliable (an high value of K prevent for bad estimations during this stage), new quaternions provided by the observation estimation algorithm are dual respect the previous ones, as shown in figures 7 and 8. This lack introduce a time, proportional to K, during which the estimation is not correct. The higher is K, the longer will take the algorithm to reach the convergence.

A solution to the problem above can be simply achieved by using the latest filtered quaternion, $\hat{\vec{q}}(t-1)$, in the observation estimation algorithm (Gradient Descent or Gauss-Newton) instead of the latest quaternion provided by the same algorithm, $\vec{q}(t-1)$. Then, during first step of the observation estimation algorithm, minimization function becomes $\mathbf{f}\left(\hat{\vec{q}}(t-1), \vec{z}_0, \vec{z}_t\right)$, which could be a common sense assumption. Figure 9 and figure 10 shows system response against high perturbations; as can be seen, quaternions don't change configuration, avoiding temporary divergences.

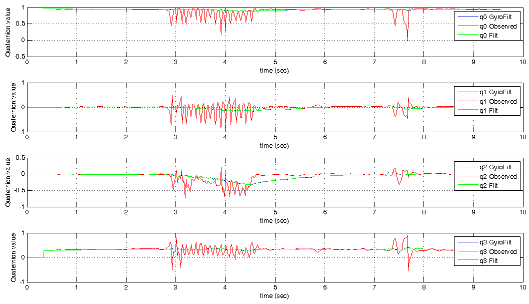

Figure 9: Trend of quaternions during an experiment with high acceleration perturbations. As can be seen, quaternions doesn't change configuration since the filtered signal is used as reference in observation estimation algorithm.

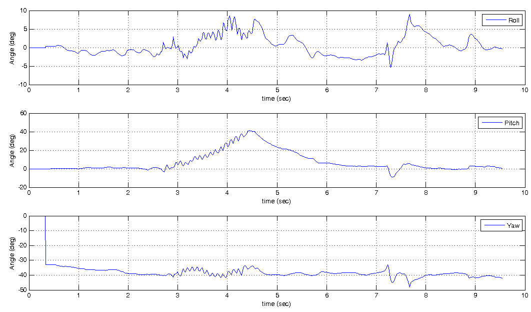

Figure 10: Trend of angles related to quaternions shown in figure 9. Angles estimated aren't the best, but of course are better than the ones shown in figure 8.

Part III

Complementary Filter and Kalman Filter Comparison

In [1] it is shown a solution to the orientation estimation problem using Kalman filter as estimation algorithm. This kind of solution is not a trivial one and requires more tuning operations than Complementary filter in order to fit the algorithm to the IMU. In this section a brief comparison between Kalman filter and Complementary filter will be presented. Two main experiments has been played:

- The first experiment takes in account system response against common rotations and movements;
- The second experiment treats system response against high external perturbations.

With regards to first experiment, figures 11 and 12 show the results.

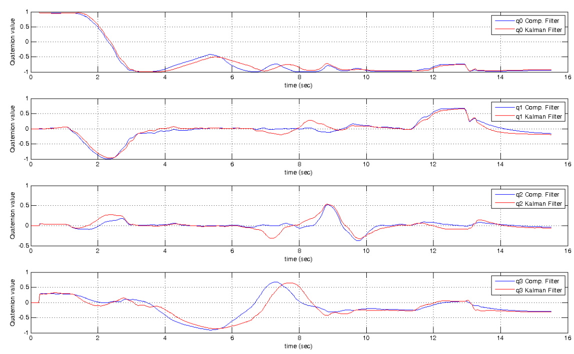

Figure 11: Trend of Complementary filter (blue) and Kalman filter (red) quaternions against normal movements of the IMU.

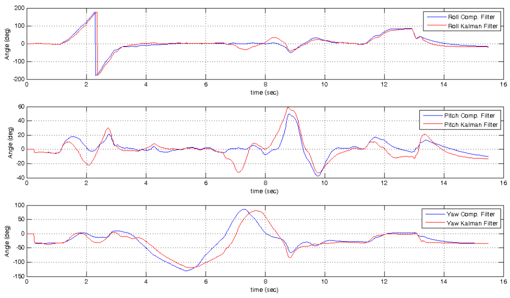

Figure 12: Trend of the angles related to quaternions shown in figure 11.

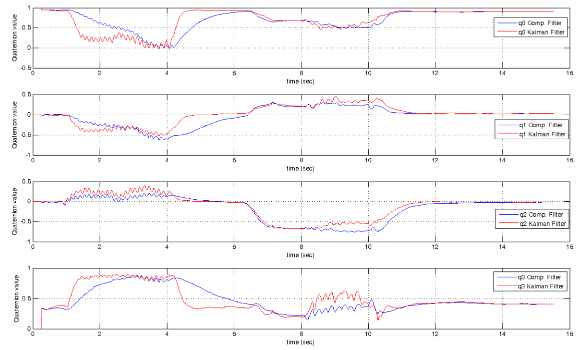

Figure 13: Trend of Complementary filter (blue) and Kalman filter (red) quaternions against high perturbations on the IMU.

As can be seen, Kalman filter response against fast movements is faster than Complementary filter one (during last 4 seconds of the experiments fast rotation has been played). This is due to the static gain of Complementar filter, that allows a great response when the IMU moves while introduces a time to reach the steady state (gyroscope doesn't provide measures anymore, so the dynamic component is related to $(1-K)$ gain). The lower is K the faster will be the steady state convergence, but the swings around the steady state will increase too (less expectation is based on gyroscope measurements); this is the reason why a value of $K = 0.98$ has been chosen. Kalman filter, through a dynamic gain inversely proportional to variances and covariances matrix of the observations, R, and in some way related to the variance of prediction error, allows the filter to adjust to IMU movements. This behavior leads to a faster response to high perturbations, as shown in figures 13 and 14. Indeed, while Complementary filter takes a while to diverge, Kalman filter change almost immediately state.

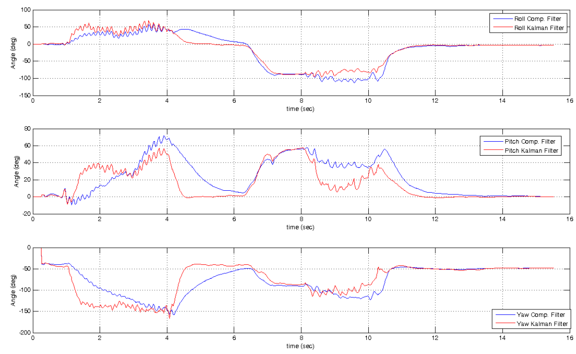

Figure 14: Trend of the angles related to quaternions shown in figure 13.

Conclusions

Continuing our work on the project published on google code (*http://code.google.com/p/9dof-orientation-estimation/*), some improvement has been achieved during the month following the project release. There is still much work to do, so during the next months we'll try to get new results.

References

[1] D. Comotti, M. Ermidoro, "Relazione del Corso di Progetto di Microelettronica".

[2] S. O. H. Madwick, "An Efficient Orientation Filter for Inertial and Inertial/Magnetic Sensor Arrays"

[3] J. L. Marins, X. Yun, E. R. Bachmann, R. B. McGhee, and M. J. Zyda, "An Extended Kalman Filter for Quaternion-Based Orientation Estimation Using MARG Sensors".

[4] http://en.wikipedia.org/

[5] Shane Colton, "The Balance Filter A Simple Solution for Integrating Accelerometer and Gyroscope Measurements for a Balancing Platform".