Notes on Quaternion-based EKF Orientation Estimation for a Mobile Phone

Jorhabib Eljaik Gomez

1 Initial Considerations

In all of the followig versions of the filter, we will assume the following:

- The process to be estimated is random and can be modeled as: $x_{k+1} = \phi_k x_k + w_k$.
- The measurement (observation) of the process respects the linear relationship $z_k = H_k x_k + v_k$.
- An initial estimate of the process at some time t_k is available. This prior estimate is denoted $\hat{x}_k^-$.
- The error covariance matrix associated with $\hat{x}_k^-$ is known, which corresponds to $P_k^- = E[e_k^- e^{-T}_k]$ where e_k^- is the **estimation error** defined as $e^-{}_k = x_k - \hat{x}_k^-$.

Where:

variables $x_k(n \times 1)$ process state vector at time t_k

variables $\phi_k(n \times n)$ matrix relating x_k to x_{k+1} in the absence of a forcing function (no input).

variables $w_k(n \times 1)$ noise vector with known covariance structure.

variables $z_k(m \times 1)$ vector measurement at time t_k

variables $H_k(m \times 1)$ matrix giving the (noiseless) linear relationship between the measurement and the state vector at time t_k

variables $v_k(m \times 1)$ measurement error

Covariance matrices for the w_k and v_k vectors at time t_k are given by Q_k and R_k.

1.1 Kalman Filter

Here goes a summarized explanation of the Kalman Filter.

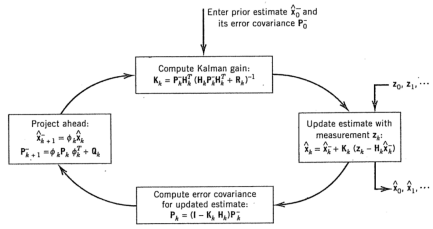

Figure 1: Kalman filter algorithm. Extracted from Introduction to Random Signals and Applied Kalman Filtering.

1.2 Sensors Calibration

This is a very delicate and important issue that needs to be addressed. Good summarized information can be found in [2] and the selected choice described here.

2 Simplified Version (Gyroscope and Accelerometer)

In this simplified version we have the following description of the process:

- **State Space Model:** The state is the orientation of the sensor local frame {S} with respect to the earth-fixed world reference frame {W} corresponding to the east-north-up system (y-axis pointing north, z-axis pointing up). This orientation is what we want to estimate, and to do so, we will use the angular velocity provided by the gyroscope as input to the system, while the measurement equation will come from the model of the accelerometer such that:

$$x_k = q_{s_k}^w \tag{1}$$
$$u_k = \omega_k^s \tag{2}$$
$$z_k = f(a_k^s) \tag{3}$$

  In the previous equations $q_{s_k}^w$ stands for the quaternion in rad at time k, representing the orientation between the sensor and the world reference frame (quaternions in this document have the following structure: $q = [q_{\text{real}} \quad q_{\text{vec}}]$), ω_k^s the angular velocity in rad/s in the sensor reference frame while $f(a_k^s)$ is actually the rotated gravity acceleration in {W} felt by the accelerometer in m/s^2. Recall that the accelerometer measures the acceleration acting **on** itself in its local reference frame {S}. We will assume also that the gravity acceleration **felt** by the accelerometer at rest in {W} is constant and equal to $g^w = [0 \quad 0 \quad 9.81]^T$ and that the body does not move too fast.

- **System Dynamics:** Since our state vector consists only of the quaternion $q_{s_k}^w$, its dynamics are given by the derivate of the quaternion, i.e.

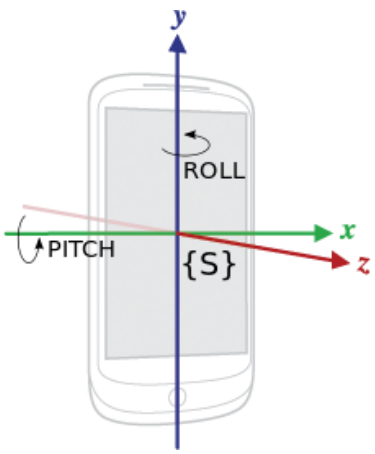

Figure 2: Local sensor frame {S}

$\dot{q}_s^w = \frac{1}{2}\Omega(\omega^s)q_s^w$ in continuous time. Assuming that the angular velocity remains constant over the integration time, the previus equation can be transformed into a discretized update equation of the form $q_{s_{k+1}}^w = \Phi_k q_{s_k}^w$ where $\Phi_k = e^{F\Delta T}$ and F the transition matrix $\frac{1}{2}\Omega(\omega^s)$. Including the gyro noise w with probability distribution $\mathcal{N}\ (0, Q_c)$ and following a Taylor series expansion we get:

$$q_{s_{k+1}}^w = e^{\frac{1}{2}\Omega(\omega_k^s + w_k)\Delta T} q_{s_k}^w \tag{4}$$

$$= (I_{4\times4} + \frac{1}{2}\Omega(\omega_k^s)\Delta T)q_{s_k}^w + \frac{\Delta T}{2}\Xi(q_{s_k}^w)w_k \tag{5}$$

Where the operators Ω and Ξ are:

$$\Omega(\omega_k^s) = \begin{bmatrix} 0 & -(\omega_k^s)^T \\ \omega_k^s & -S(\omega_k^s)z \end{bmatrix} \tag{6}$$

$$\Xi(q_{s_k}^w) = \begin{bmatrix} [-q_{s_k}^w]_{vec}^T \\ [q_{s_k}^w]_{real} I_{3\times3} + S([q_{s_k}^w]_{vec}) \end{bmatrix} \tag{7}$$

$\Delta T = t_{k+1} - t_k$ and $S(v)$ is the skew symmetric operator:

$$S(v) = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix} \tag{8}$$

- **Sensor Fusion Model:** Let us recall the discrete Kalman filter structure given by the process and sensor equations as:

$$x_{k+1} = \phi_k x_k + w_k \tag{9}$$

$$z_k = H_k x_k + v_k \tag{10}$$

Given the system dynamics previously described, the continuous state model was:

$$\dot{x} = \underbrace{\frac{1}{2}S(\omega_k + w_k)}_{F} x \tag{11}$$

Where F is the transition matrix.

The accelerometer can be modeled using the following equation,

$$y_{k_a} = Q^T(q_k)(g^o + F_k) + v_k^a \tag{12}$$

Where g^0 is the nominal gravity vector and measurement noise v_k^a

While the magnetometer can be modeled using the following equation,

$$y_{k_m} = Q^T(q_k)m^0 + v_k^m \tag{13}$$

Where m^0 is the earth magnetic field in world coordinates, or nominal magnetic field and given by $m^0 = (0 \quad \sqrt{m_x^2 + m_y^2} \quad m_z)^T$

Then, following discretization the Kalman model can be written as:

$$x_{k+1} = \underbrace{\left(I + \frac{1}{2}S(\omega_k)\Delta T\right)}_{\phi_k} x_k + \underbrace{\frac{\Delta T}{2}\bar{S}(q_k)w_k}_{w'_k} \tag{14}$$

$$z_k = Q^T(q_k)(g^o + F_k) + v_k^a \tag{15}$$

This equation is also known as the Zeroth Order Quaternion Integrator, which assumes constant angular velocity over the integration period ΔT. A very detailed derivation can be found in Section 1.6.1 in [6].

The measurement model contains so far only the accelerometer. Also, assuming quaternions of the form $q = [q_{real} \quad q_x \quad q_y \quad q_z]^T$.

$$S(\omega) = \begin{bmatrix} 0 & -\omega_x & -\omega_y & -\omega_z \\ \omega_x & 0 & \omega_z & -\omega_y \\ \omega_y & -\omega_z & 0 & \omega_x \\ \omega_z & \omega_y & -\omega_x & 0 \end{bmatrix} \tag{16}$$

and

$$\bar{S}(q) = \begin{bmatrix} -q_1 & -q_2 & -q_3 \\ q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{bmatrix} \tag{17}$$

Where F_k will be assumed to be zero. The measurement equation however must be linearized! Once linearized, we can obtain also H_k. The measurement equation for the accelerometer can be linearized by taking its Jacobian and evaluating at the most recent estimate. However, as done in [3], the linearized measurement matrix H_k can be obtained as:

$$H_k = \Psi(q_k, g^0) = \frac{\partial}{\partial q_k}\left(q_k \otimes \bar{g}^0 \otimes q_k^{-1}\right) \tag{18}$$

$$= \bar{S}_R(q_k)\bar{S}_R(g^0)^T + \bar{S}_L^T(q_k)\bar{S}_L(g^0)^T \begin{bmatrix} -I_3 & 0 \\ 0 & 1 \end{bmatrix} \tag{19}$$

Where

$$S_L = \begin{bmatrix} q_4 & q_3 & -q_2 & q_1 \\ -q_3 & q_4 & q_1 & q_2 \\ q_2 & -q_1 & q_4 & q_3 \\ -q_1 & -q_2 & -q_3 & q_4 \end{bmatrix} \tag{20}$$

$$S_R = \begin{bmatrix} q_4 & -q_3 & q_2 & q_1 \\ q_3 & q_4 & -q_1 & q_2 \\ -q_2 & q_1 & q_4 & q_3 \\ -q_1 & -q_2 & -q_3 & q_4 \end{bmatrix} \tag{21}$$

And corresponding measurement noise covariance matrix

$$R_k = \Sigma_a \tag{22}$$

Where $\Sigma_a = I_4\sigma_v^2$.

In this way, we finally obtain the linearized measurement equations such that

$$z_k = H_k q_k \tag{23}$$

There is also another possibility, which consists in using accelerometer and magnetometer readings as inputs to the QUEST [5] algorithm to produce what will be called the *computed quaternion* which is then presented to the Kalman Filter as measurements, greatly simplifying its implementation [5]. A method using QUEST will be presented later in Section 4.

NOTE: DON'T FORGET TO CHECK THAT ALL THE QUATERNION FUNCTIONS THAT I'M USING CONSIDER THE QUATERNION IN THE SAME WAY. THE REAL PART COULD BE AT THE BEGINNING OR AT THE END OF THE VECTOR.

3 Estimating Orientation and Gyro Bias

4 Simplified Version Using the QUaternionESTimator algorithm (QUEST)

An alternative to the linearization of the accelerometer measurement equation is that of fusing the accelerometer and magnetometer measurement into a single one through an efficient statistical attitude determination method called QUEST, which stands for QUaternion ESTimator. The basic idea is that if two orientation observations are available, this information can be joint together

using a statistical method. Suppose we have a set of of N unit vectors $\hat{z}_k$, for which we have the mathematical model that relates their sensor measurement in the inertial i frame to the body frame b through a rotation matrix R_i^b as:

$$z_{kb} = R_i^b z_{ki} \tag{24}$$

This set of equations is overdetermined when having two or more readings. For this reason, we want to find a solution for R_i^b that minimizes the overall error for all vectors. The classical way to state this problem is as an optimization one that minimizes the sum of the squared errors for each vector measurement. Thus, the loss function can be expressed as:

$$J(R_i^b) = \frac{1}{2}\sum_{k=1}^{N} w_k \left| z_{kb} - R_i^b z_{ki} \right|^2 \tag{25}$$

Where J is the loss function to be minimized, k is the index for observations (measurements) z_k is the kth vector being measured, z_{kb} is the matrix of measured components in the body frame and z_{ki} the matrix of components in the inertial frame [4]. The previous loss function is also known in literature as Wahba's loss function.

Depending on how noisy measurements are, the harder it becomes to minimize this loss function.

QUEST is just one efficient way to perform this minimization and will be described next. Many spacecraft attitude systems use this method to fuse the unit vector giving the direction to the sun or a star and the unit vector in the direction of the Earth's magnetic field. In this particular implementation we intend to do the same, with the gravity vector given by the accelerometer and the Earth's magnetic field direction provided by the magnetometer.

4.1 QUEST

5 Multiplicative Kalman Filter

The Kalman Filter explained in Section 1.1 is known as a *Direct Kalman Filter*

6 Bayesian Interpretation of the Kalman Filter

The Kalman Filter can be seen as a special case of the sequential state estimation problem in iterative Bayesian filtering. In recursive Bayesian estimation, two assumptions are used to derive the recursive Bayesian filter: i) The state follows a first-order Markov process, meaning $p(x_n|x_{0:n-1}) = p(x_n|x_{n-1})$, ii) The observations are independent of the given states. From Bayes rule it can be obtained that the **conditional pdf** of x_n is:

$$p(x_n|y_{0:n}) = \frac{p(y_n|x_n)p(x_n|y_{0:n-1})}{p(y_n|y_{n-1})} \tag{26}$$

In the previous equation the posterior probability is described by three terms:

- **Prior:** $p(x_n|y_{0:n-1})$ and is also a function of the transition probability $p(x_{n+1}|x_n)$ (or $p(x_n|x_{n-1})$) which is further characterized by the transition matrix F from the formulation of the discrete dynamic state-space problem.
- **Likelihood:** $p(y_n|x_n)$ which is characterized by a measurement equation $y_n = g(x_n, v_n)$ where v_n is measurement noise (or its measurement matrix H_n).
- **Evidence:** $p(y_n|y_{0:n-1})$

Kalman Filtering can be reduced to a Maximum A Posteriori (MAP) solution to

$$\left.\frac{\partial \log p(x_n|y_{0:n})}{\partial x_n}\right|_{x_n=\hat{x}^{MAP}} = 0 \tag{27}$$

after calculating the means and covariances of the previously listed terms, i.e.

$$\begin{aligned} \mathrm{E}[y_n|x_n] &= G_n x_n \\ Cov[y_n|x_n] &= \Sigma_v \\ \mathrm{E}[x_n|y_{0:n-1}] &= F x_{n-1} \\ Cov[x_n|y_{0:n-1}] &= Cov[x_n - \hat{x}_{n|n-1}] \end{aligned} \tag{28}$$

And pdfs:

$$p(y_n|x_n) = A_1 \exp(-\frac{1}{2}(y_n - G_n x_n)^T \Sigma_v^{-1}(y_n - G_n x_n)) \tag{29}$$

$$p(x_n|y_{0:n-1}) = A_2 \exp(\frac{1}{2}(x_n - \hat{x}_{n|n-1})^T \times P_{n,n-1}^{-1}(x_n - \hat{x}_{n|n-1})) \tag{30}$$

$$\tag{31}$$

We can now obtain an expression for the posterior $p(x_n|y_{0:n})$ as:

$$\begin{aligned} p(x_n|y_{0:n}) \propto A \exp\Bigg(&\frac{1}{2}(y_n - G_n x_n)^T)\Sigma_v^{-1}(y_n - G_n x_n) \\ &-\frac{1}{2}(x_n - \hat{x}_{n|n-1})^T P_{n,n-1}^{-1}(x_n - \hat{x}_{n|n-1}\Bigg) \end{aligned} \tag{32}$$

Where $A_1 = 2\pi^{-N_y/2}|\Sigma_v|^{-1/2}$, $A_2 = 2\pi^{-N_x/2}|P_{n,n-1}|^{-1/2}$ and $A = A_1 A_2$. For more details on these derivations and further references please refer to [1].

References

[1] Zhe Chen. Bayesian filtering: From kalman filters to particle filters, and beyond. *Statistics*, 182(1):1–69, 2003. Used for basic understanding of the Bayesian interpretation of the Kalman and Particle filter.

[2] Lucian Cucu. Applying kalman filtering on a quadruped robot. Technical report, Ecole Polytechnique Federale de Lausanne, 2012. Used for gyroscope calibration methods. Derivation of an adaptive MEASUREMENT COVARIANCE MATRIX is found here.

[3] Angelo Maria Sabatini Gabriele Ligorio. Extended kalman filter-based methods for pose estimation using visual, inertial and magnetic sensors: Comparative analysis and performance evaluation. 2013. Linearized version of the measurement equation.

[4] C. D. Hall. *[Notes] Spacecraft Attitude Dynamics and Control. Chapter 4.* 2003.

[5] M.D. Shuster and S.D. Oh. Three-axis attitude determination from vector observations. *J. Guid. Control*, 4:70–77, 1981. This is the original paper in which the QUEST algorithm is presented.

[6] Nikolas Trawny and Stergios I. Roumeliotis. Indirect kalman filter for 3d attitude estimation. a tutorial for quaternion algebra. Technical report, University of Minnesota, 2005. Quaternion Algebra explained.