

Unscented Filtering in a Unit Quaternion Space for Spacecraft Attitude Estimation

Yee-Jin Cheon*

*KOMPSAT-3 System Engineering & Integration Dept., Korea Aerospace Research Institute, Daejeon 305-333, Korea
(Tel: +82-42-860-2757; Fax: +82-42-860-2007; Email:yjcheon@kari.re.kr)

Abstract: A new approach to the straightforward implementation of the unscented filter in a unit quaternion space is proposed for spacecraft attitude estimation. Since the unscented filter is formulated in a vector space and the unit quaternions do not belong to a vector space but lie on a nonlinear manifold, the weighted sum of quaternion samples does not produce a unit quaternion estimate. To overcome this difficulty, a method of weighted mean computation for quaternions is derived in rotational space, leading to a quaternion with unit norm. A quaternion multiplication is used for predicted covariance computation and quaternion update, which makes a quaternion in a filter lie in the unit quaternion space. Since the quaternion process noise increases the uncertainty in attitude orientation, modeling it either as the vector part of a quaternion or as a rotation vector is considered. Simulation results illustrate that the proposed approach successfully estimates spacecraft attitude for large initial errors and high tip-off rates, and modeling the quaternion process noise as a rotation vector is more optimal than handling it as the vector part of a quaternion.

Keywords: Unscented Kalman Filter, Unit Quaternion Space, Quaternion Mean, Attitude Estimation, Rotation Vector

1. Introduction

Attitude estimation is one of the basic problems and an important issue in many spacecraft applications, and has been studied extensively in the past [5][10][14][15][16][17][18]. There are several parameterizations available for spacecraft attitude representation, such as Euler angles, quaternions, modified Rodrigues parameters, and even the rotation vector. The detailed discussion on each parametrization for attitude estimation is found in [5]. The extended Kalman filter (EKF) is widely used for attitude estimation [14][15][18]. The unscented Kalman filter (UKF) is an extension of the classical Kalman filter to nonlinear process and measurement models and an alternative to the EKF for a variety of estimation and control problems [5][11][13][19][20]. For nonlinear systems the UKF uses a deterministically chosen minimal set of sample points to capture the probability distribution more accurately than the linearization of the EKF, resulting in faster convergence from inaccurate initial conditions in attitude estimation problems.

In [11], the experiment which compares UKF and EKF for human head and hand orientation data represented with quaternions is presented. However, the estimated mean value is simply derived using an average sum of the quaternions by treating the quaternion as an element in vector space, which leads the quaternion estimate to depart from the unit sphere. The quaternion is normalized by brute force and the linearized model of quaternion kinematics is used, which does not take full advantage of the UKF capabilities. To directly utilize UKF formulation built in vector space and to avoid a mean computation problem of quaternions, a modified method employing the combination of the quaternion with the generalized Rodrigues parameters is used in [5]. This method converts a quaternion to a Rodrigues parameter and converted Rodrigues parameters are used to compute the predicted mean, the covariance and cross-relation matrix since Rodrigues parameters are unconstrained, i.e., lying in

vector space. Rodrigues parameters are converted back to quaternion because this method uses the inertial quaternion for state propagation. It was shown that the performance of UKF outperforms that of the standard EKF for large initialization errors [5].

Generally, the problem in calculating the mean of a set of rotations lies in the fact that rotations do not belong to a vector space, but lie on a nonlinear manifold. If we choose to represent the rotation by a quaternion, the proper rotations are constrained to the three-dimensional unit sphere of a four-dimensional Euclidian space. In the original concept of UKF, the predicted mean is simply calculated using an averaged sum of predicted states because state vector elements lie in a vector space. However, the weighted sum of unit quaternion points, i.e., a quaternion of unit norm, does not lie on the unit sphere because unit quaternion is not mathematically closed for addition and scalar multiplication. There is no guarantee that the mean of quaternions lies on the nonlinear manifold with allowable rotations by directly using UKF with quaternion parameterization. Due to this fact, it seems that the direct implementation of the UKF with quaternions is undesirable. There are several research works for computing means and averaging in the group of rotations [7][8][9]. The method of averaging rotations for rotation matrices is proposed in [8]. Providing that the inner product of the quaternions is non-negative, an average of two rotations is a normalized arithmetic mean of the unit quaternions associated with the two rotations [9]. Very often the barycenter of the quaternions that represents the rotations is used as an estimate of the mean rotation. In [7], a method of averaging rotations was derived from natural approximations to the Riemannian metric to show how the barycenter mean of rotations approximates the mean based on Riemannian metric. By using this method, the mean rotation, i. e., the optimal quaternion, is exactly the barycenter of rotations on the curvature of the nonlinear manifold with renormalization.

In this paper, a new approach to the straightforward im-

plementation of UKF with quaternion in a unit quaternion space is proposed, where no parameter conversion is required, such as transformation between quaternions and Rodrigues parameters. To ensure that a weighted sum of quaternions lies in unit quaternion space, we derive a quaternion-based weighted mean computation method by expanding the method given in [7]. This method can be also used for the predicted observation computation when measurement output is given by the quaternion. To evenly disseminate quaternion samples on the unit sphere around the current quaternion estimate, both error quaternion and its inverse quaternion are used to generate sigma point quaternions. A multiplicative quaternion-error is used for predicted covariance computation of the quaternion because it represents the distance from the predicted mean quaternion, and the unit quaternion is not closed for subtraction, guaranteeing that the quaternion in filter lies in unit quaternion space. To resolve the problem that a simple linear correction for quaternion update makes the quaternion depart from constraint unit sphere, the quaternion is updated by multiplication. Since quaternion process noise finally results in an increase of the uncertainty in attitude orientation, we treat the quaternion process noise as a vector part of the quaternion or as a rotation vector.

This paper is organized as follows. A brief review of UKF is provided in Section 2. In Section 3, a summary of attitude kinematics and the gyro model is described. The weighted mean computation method of quaternions is derived, which guarantees that the resulting quaternion has unit norm. The direct implementation of UKF with a quaternion in unit quaternion space is presented in Section 3. In Section 4, numerical simulations for the new filter are presented using simulated three-axis magnetometer and gyro measurements. Concluding remarks follow in Section 5.

2. Unscented Kalman Filtering

The unscented Kalman filter represents a derivative-free, i.e., no need for Jacobian and Hessian calculation, alternative to the extended Kalman filter, and builds on the principle that it is easier to approximate a Gaussian distribution than it is to approximate an arbitrary nonlinear function. In the UKF, a minimal set of sample points are deterministically chosen and propagated through the original nonlinear system to capture the posterior mean and covariance of a random variable accurately to the 2nd order Taylor series expansion for any nonlinearity. More details are found in [1][2][3][4]. The general formulation of the UKF is summarized for a discrete-time nonlinear system. Let us consider the following system model of discrete-time nonlinear equations:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, k) + \mathbf{w}_k \quad (1a)$$

$$\mathbf{y}_k = \mathbf{h}(\mathbf{x}_k, k) + \mathbf{v}_k \quad (1b)$$

where $\mathbf{x}_k \in \mathbb{R}^{n \times 1}$ is the state vector, $\mathbf{y}_k \in \mathbb{R}^{m \times 1}$ is the observation vector, \mathbf{f} and \mathbf{h} are the nonlinear system dynamic models. It is assumed that the process noise \mathbf{w}_k and measurement noise vector \mathbf{v}_k are uncorrelated white Gaussian processes with zero means and covariances given by \mathbf{Q}_k and

\mathbf{R}_k , respectively. Then, given an $n \times n$ covariance matrix \mathbf{P}_k , the set of $(2n+1)$ sigma points $\mathcal{X}_k \in \mathbb{R}^{2n+1}$ is computed as follows:

$$\mathcal{X}_{0,k} = \hat{\mathbf{x}}_k \quad (2a)$$

$$\mathcal{X}_{i,k} = \hat{\mathbf{x}}_k + \left(\gamma \sqrt{(\mathbf{P}_k + \mathbf{Q}_k)} \right)_i \quad \text{for } i = 1, \dots, n \quad (2b)$$

$$\mathcal{X}_{i,k} = \hat{\mathbf{x}}_k - \left(\gamma \sqrt{(\mathbf{P}_k + \mathbf{Q}_k)} \right)_i \quad \text{for } i = n+1, \dots, 2n \quad (2c)$$

where $\gamma = \sqrt{n+\lambda}$ and $\lambda = \alpha^2(n+\kappa) - n$ is a composite scaling parameter. The constant α determines the spread of the sigma points around $\hat{\mathbf{x}}_k$ and is usually set to a small positive value ($0 \leq \alpha \leq 1$). The constant κ is a secondary scaling parameter which is usually set to 0 or $3-n$ and provides an extra degree of freedom for fine tuning of the higher order moments. The matrix square-root $\sqrt{(\mathbf{P}_k + \mathbf{Q}_k)}$ can be calculated by a lower triangular Cholesky factorization. From Eq. (2) a matrix \mathcal{X}_k of $2n+1$ sigma vectors $\mathcal{X}_{i,k}$ is formed as

$$\mathcal{X}_k = \begin{bmatrix} \hat{\mathbf{x}}_k & \hat{\mathbf{x}}_k + \gamma \sqrt{(\mathbf{P}_k + \mathbf{Q}_k)} & \hat{\mathbf{x}}_k + \gamma \sqrt{(\mathbf{P}_k + \mathbf{Q}_k)} \end{bmatrix} \quad (3)$$

The transformed sigma vectors are propagated through the following nonlinear function:

$$\mathcal{X}_{i,k+1} = \mathbf{f}(\mathcal{X}_{i,k}, k) \quad (4)$$

where $\mathcal{X}_{i,k}$ is the i th column of \mathcal{X}_k . The predicted state vector $\hat{\mathbf{x}}_{k+1}^-$ and its predicted covariance \mathbf{P}_{k+1}^- are computed using a weighted sample mean and covariance of the posterior sigma point vectors, as follows:

$$\hat{\mathbf{x}}_{k+1}^- = \sum_{i=0}^{2n} W_i^{(m)} \mathcal{X}_{i,k+1} \quad (5)$$

$$\mathbf{P}_{k+1}^- = \sum_{i=0}^{2n} W_i^{(c)} \{ \mathcal{X}_{i,k+1} - \hat{\mathbf{x}}_{k+1}^- \} \{ \mathcal{X}_{i,k+1} - \hat{\mathbf{x}}_{k+1}^- \}^T \quad (6)$$

with the following weights W_i [4]:

$$W_0^{(m)} = \lambda / (n + \lambda) \quad (7a)$$

$$W_0^{(c)} = \lambda / (n + \lambda) + (1 - \alpha^2 + \beta) \quad (7b)$$

$$W_i^{(m)} = W_i^{(c)} = 1 / \{ 2(n + \lambda) \} \quad i = 1, \dots, 2n \quad (7c)$$

where β is used to make further higher order effects to be incorporated by adding the weighting of the *zeroth* sigma point of the calculation of the covariance (for Gaussian distributions $\beta = 2$ is optimal).

Similarly, the predicted observation vector $\hat{\mathbf{y}}_{k+1}^-$ is also calculated as

$$\hat{\mathbf{y}}_{k+1}^- = \sum_{i=0}^{2n} W_i^{(m)} \mathcal{Y}_{i,k+1} \quad (8)$$

where

$$\mathcal{Y}_{i,k+1} = \mathbf{h}(\mathcal{X}_{i,k+1}, k+1) \quad (9)$$

The predicted output covariance \mathbf{P}_{k+1}^{yy} is given by

$$\mathbf{P}_{k+1}^{yy} = \sum_{i=0}^{2n} W_i^{(c)} \{\mathcal{Y}_{i,k+1} - \hat{\mathbf{y}}_{k+1}^-\} \{\mathcal{Y}_{i,k+1} - \hat{\mathbf{y}}_{k+1}^-\}^T \quad (10)$$

then the innovation covariance \mathbf{P}_{k+1}^{vv} is calculated by

$$\mathbf{P}_{k+1}^{vv} = \mathbf{P}_{k+1}^{yy} + \mathbf{R}_{k+1} \quad (11)$$

The filter gain \mathcal{K}_{k+1} is computed by

$$\mathcal{K}_{k+1} = \mathbf{P}_{k+1}^{xy} (\mathbf{P}_{k+1}^{vv})^{-1} \quad (12)$$

and the cross correlation matrix is obtained as

$$\mathbf{P}_{k+1}^{xy} = \sum_{i=0}^{2n} W_i^{(c)} \{\mathcal{X}_{i,k+1} - \hat{\mathbf{x}}_{k+1}^-\} \{\mathcal{Y}_{i,k+1} - \hat{\mathbf{y}}_{k+1}^-\}^T \quad (13)$$

The estimated state vector $\hat{\mathbf{x}}_{k+1}^+$ and updated covariance \mathbf{P}_{k+1}^+ are given by

$$\hat{\mathbf{x}}_{k+1}^+ = \hat{\mathbf{x}}_{k+1}^- + \mathcal{K}_{k+1} (\mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1}^-) \quad (14)$$

$$\mathbf{P}_{k+1}^+ = \mathbf{P}_{k+1}^- - \mathcal{K}_{k+1} \mathbf{P}_{k+1}^{vv} \mathcal{K}_{k+1}^T \quad (15)$$

3. Unscented Filtering in a Unit Quaternion Space

The difficulty of filtering unit quaternion data stems from the nonlinear nature of unit quaternion space. In other words, the problem in calculating the mean of a set of quaternions lies in the fact that rotations represented by the quaternion do not belong to a vector space, but lie on a nonlinear manifold and the quaternions are constrained to the three-dimensional unit sphere of a four-dimensional Euclidean space. The weighted sum of unit quaternion points is generally not a quaternion of unit norm because the unit quaternion is not mathematically closed for addition and scalar multiplication. With reference to Eq. (5), there is no guarantee that the resulting quaternion will have unit norm. Due to this fact, the straightforward implementation of the UKF with quaternions seems to be undesirable. In the original concept of UKF, the predicted mean is simply calculated using an averaged sum of predicted states, i.e., sum over all the elements of the set divided by the number of elements. This is true when the state vector elements lie in a vector space. However, if the state vector is quaternion, the predicted quaternion mean should be calculated in the rotational space to preserve the nonlinear nature of unit quaternion. A UKF in a unit quaternion space is possible when a weighted mean of quaternions produces a unit quaternion estimate, and the predicted covariance computation and quaternion guarantees that quaternion in filter lies on unit sphere.

In this section, the quaternion attitude kinematics and gyro model are briefly reviewed and a UKF in unit quaternion space for attitude estimation is derived by preserving the nonlinear nature of unit quaternion space.

The attitude of spacecraft can be described by means of quaternion parameterization, defined as

$$\mathbf{q} \equiv \begin{bmatrix} q_{13} \\ q_4 \end{bmatrix} \quad (16)$$

with

$$\mathbf{q}_{13} \equiv \begin{bmatrix} q_1 \\ q_2 \\ q_3 \end{bmatrix} = \hat{\mathbf{n}} \sin(\Phi/2) \quad (17a)$$

$$q_4 = \cos(\Phi/2) \quad (17b)$$

where $\hat{\mathbf{n}}$ is a unit vector corresponding to the rotation axis and Φ is the angle of rotation. The quaternion satisfies the following normalization constraint

$$\mathbf{q}^T \mathbf{q} = \mathbf{q}_{13}^T \mathbf{q}_{13} + q_4^2 = 1 \quad (18)$$

The inverse quaternion is defined as

$$\mathbf{q}^{-1} \equiv \begin{bmatrix} -\mathbf{q}_{13} & q_4 \end{bmatrix}^T = \begin{bmatrix} -q_1 & -q_2 & -q_3 & q_4 \end{bmatrix}^T \quad (19)$$

The quaternion kinematic equations of motion are obtained by using angular velocity ($\boldsymbol{\omega}$) as follows

$$\dot{\mathbf{q}} = \frac{1}{2} \Omega(\boldsymbol{\omega}) \mathbf{q} \quad (20)$$

with

$$\Omega(\boldsymbol{\omega}) \equiv \begin{bmatrix} -[\boldsymbol{\omega} \times] & \vdots & \boldsymbol{\omega} \\ \vdots & \ddots & \vdots \\ -\boldsymbol{\omega}^T & \vdots & 0 \end{bmatrix} \quad (21)$$

where the matrix cross product operator $[\boldsymbol{\omega} \times]$ represents the skew-symmetric matrix, defined by

$$[\boldsymbol{\omega} \times] = \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \quad (22)$$

The measured angular rate output $\tilde{\boldsymbol{\omega}}$ is modeled as [6]

$$\tilde{\boldsymbol{\omega}} = \boldsymbol{\omega} + \mathbf{b} + \boldsymbol{\eta}_1 \quad (23a)$$

$$\dot{\mathbf{b}} = \boldsymbol{\eta}_2 \quad (23b)$$

where $\boldsymbol{\omega}$ is the true angular rate of the spacecraft, $\boldsymbol{\eta}_1$ and $\boldsymbol{\eta}_2$ are white Gaussian processes with zero means and \mathbf{b} is a gyro bias error. Given a post-update estimates $\hat{\mathbf{b}}_k^+$, the post-update of $\boldsymbol{\omega}$ and \mathbf{b} are given by

$$\hat{\boldsymbol{\omega}}_k^+ = \tilde{\boldsymbol{\omega}} - \hat{\mathbf{b}}_k^+ \quad (24a)$$

$$\hat{\mathbf{b}}_{k+1}^- = \hat{\mathbf{b}}_k^+ \quad (24b)$$

Given post-update estimates $\hat{\boldsymbol{\omega}}_k^+$ and $\hat{\mathbf{q}}_k^+$, the quaternion can then be propagated in the discrete-time equivalent form of Eq. (20), given by [5]

$$\hat{\mathbf{q}}_{k+1}^- = \Omega(\hat{\boldsymbol{\omega}}_k^+) \hat{\mathbf{q}}_k^+ \quad (25)$$

with

$$\Omega(\hat{\boldsymbol{\omega}}_k^+) \equiv \begin{bmatrix} \cos(0.5 \|\hat{\boldsymbol{\omega}}_k^+\| \Delta t) \mathbf{I}_{3 \times 3} & \hat{\boldsymbol{\psi}}_k^+ \\ -\hat{\boldsymbol{\psi}}_k^{+T} & \cos(0.5 \|\hat{\boldsymbol{\omega}}_k^+\| \Delta t) \end{bmatrix} \quad (26)$$

where $\hat{\boldsymbol{\psi}}_k^+ = \sin(0.5 \|\hat{\boldsymbol{\omega}}_k^+\| \Delta t) \hat{\boldsymbol{\omega}}_k^+ / \|\hat{\boldsymbol{\omega}}_k^+\|$ and Δt is the sampling interval of the gyro measurement.

The normalization constraint Eq. (18) deprives the quaternion of one degree of freedom and simplifies the handling of the quaternion. However, the three degrees of freedom, i.e., vector part of the quaternion, are sufficient to represent any spatial rotation because \mathbf{q} and $-\mathbf{q}$ yield identical rotations and the quaternion elements are no longer independent of each other. Given the vector part \mathbf{q}_{13} of the quaternion, there will be positive and negative values of q_4 that satisfy the quaternion normalization constraint. We choose the following positive value of q_4 throughout the filter formulation:

$$q_4 = \sqrt{1 - \mathbf{q}_{13}^T \mathbf{q}_{13}} \quad (27)$$

The state vector of the filter is defined as

$$\hat{\mathbf{x}}_k^+ \equiv \begin{bmatrix} \hat{\boldsymbol{\sigma}}_k^+ \\ \hat{\mathbf{b}}_k^+ \end{bmatrix} \quad (28)$$

where $\hat{\boldsymbol{\sigma}}_k^+$ is the vector part of the quaternion and $\hat{\mathbf{b}}_k^+$ is the gyro bias part of the state vector.

Therefore we choose the process noise \mathbf{w}_k in Eq. (1a) with a 6-dimensional noise vector.

The set of transformed sigma points will be of the form:

$$\mathcal{X}_k \equiv \begin{bmatrix} \mathcal{X}_k^\sigma \\ \mathcal{X}_k^b \end{bmatrix} \quad (29)$$

where \mathcal{X}_k^σ corresponds to the first three components of the quaternion and \mathcal{X}_k^b is the gyro bias part. To forwardly propagate quaternions using Eq. (25), sigma points corresponding to the quaternion have to be expanded to a 4×1 dimensional vector. The scalar part of the quaternion is simply calculated using Eq. (27) from $\hat{\boldsymbol{\sigma}}_k^+$, i.e.,

$$\hat{q}_k^+ = \begin{bmatrix} \hat{\boldsymbol{\sigma}}_k^+ \\ \sqrt{1 - \hat{\boldsymbol{\sigma}}_k^{+T} \hat{\boldsymbol{\sigma}}_k^+} \end{bmatrix}^T.$$

Let denote $\gamma\sqrt{\mathbf{P}_k + \mathbf{Q}_k}$ in Eq. (3) as

$$\boldsymbol{\zeta} \equiv \begin{bmatrix} \boldsymbol{\zeta}_k^\sigma \\ \boldsymbol{\zeta}_k^b \end{bmatrix} \quad (30)$$

where $\boldsymbol{\zeta}_{i,k}^\sigma$ are the first 3 elements of the quaternion, and $\boldsymbol{\zeta}_{i,k}^b$ are the next 3 elements corresponding to the gyro bias. Since the role of process noise for the quaternion in Eq. (1a) increases the uncertainty in attitude orientation, the process noise for the quaternion can be modeled either as a vector part of a quaternion or as a rotation vector, both leading to orientation uncertainty in the attitude. Since the quaternion process noise is a 3-dimensional noise vector, it has to be expanded into a 4×1 dimensional unit quaternion to apply the multiplicative quaternion-error approach.

When the quaternion process noise is treated as a vector part of the quaternion, the noise vector is simply expanded into quaternion parametrization using Eq. (27), as follows:

$$\delta \mathbf{q}_{i,k}^+ = \begin{bmatrix} \boldsymbol{\zeta}_{i,k}^\sigma \\ \sqrt{1 - \boldsymbol{\zeta}_{i,k}^{\sigma T} \boldsymbol{\zeta}_{i,k}^\sigma} \end{bmatrix}^T. \quad (31)$$

To find a quaternion from a rotation vector, compute the angle from the magnitude of the rotation vector, then find the rotation axis by the dividing rotation vector by the angle, and then convert the angle and the rotation axis to the

quaternion. Modeling $\boldsymbol{\zeta}_{i,k}^\sigma$ as a rotation vector leads to the following 4×1 error quaternion:

$$\delta \mathbf{q}_{i,k}^+ = \begin{bmatrix} \boldsymbol{\zeta}_{i,k}^\sigma \sin(|\boldsymbol{\zeta}_{i,k}^\sigma|/2)/|\boldsymbol{\zeta}_{i,k}^\sigma| & \cos(|\boldsymbol{\zeta}_{i,k}^\sigma|/2) \end{bmatrix}^T \quad (32)$$

Since the unit quaternion is not closed for addition and subtraction, the transformed sigma points of the quaternion are not simply constructed while the sigma points for the gyro bias are simply calculated by Eq. (3). The transformed sigma points for the quaternion are also quaternions satisfying the normalization constraints and should be scattered around the current quaternion estimate on the unit sphere. Therefore, the transformed quaternion sigma points are generated by multiplying the error quaternion by the current quaternion estimate. To evenly disseminate quaternion samples on the unit sphere around the current quaternion estimate, both the error quaternion and the inverse of error quaternion, i.e., $\delta \mathbf{q}_{i,k}^+$ and $(\delta \mathbf{q}_{i,k}^+)^{-1}$, are used. The sigma point quaternions are computed by the following equations:

$$\mathcal{X}_{0,k}^q = \hat{\mathbf{q}}_k^+ \quad (33a)$$

$$\mathcal{X}_{i,k}^q = \delta \mathbf{q}_{i,k}^+ \otimes \hat{\mathbf{q}}_k^+ \quad \text{for } i = 1, \dots, 6 \quad (33b)$$

$$\mathcal{X}_{i,k}^q = (\delta \mathbf{q}_{i,k}^+)^{-1} \otimes \hat{\mathbf{q}}_k^+ \quad \text{for } i = 7, \dots, 12 \quad (33c)$$

where \otimes is the quaternion multiplier [21]. Then the set of 13 sigma points is constructed as

$$\mathcal{X}'_k \equiv \begin{bmatrix} \mathcal{X}_k^q \\ \mathcal{X}_k^b \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{q}}_k^+ & \delta \mathbf{q}_k^+ \otimes \hat{\mathbf{q}}_k^+ & (\delta \mathbf{q}_k^+)^{-1} \otimes \hat{\mathbf{q}}_k^+ \\ \hat{\mathbf{b}}_k^+ & \hat{\mathbf{b}}_k^+ + \boldsymbol{\zeta}_k^b & \hat{\mathbf{b}}_k^+ - \boldsymbol{\zeta}_k^b \end{bmatrix} \quad (34)$$

Now these transformed quaternions are propagated forward to $k+1$ by Eq. (25) as follows:

$$\mathcal{X}_{i,k+1}^q = \Omega(\hat{\boldsymbol{\omega}}_{i,k}^+) \mathcal{X}_{i,k}^q \quad (35)$$

where the estimated angular rates are given by

$$\hat{\boldsymbol{\omega}}_{i,k}^+ = \tilde{\boldsymbol{\omega}}_k - \mathcal{X}_{i,k}^b \quad (36)$$

The weighted sum of unit quaternions is generally not a quaternion of unit norm because the unit quaternion is not mathematically closed for addition and scalar multiplication. Therefore, the predicted quaternion $\hat{\mathbf{q}}_{k+1}^-$ is not simply computed using weighted sample mean in Eq. (5). To resolve this problem, we derive a quaternion-based weighted mean computation method based on the formulation in [7] by considering weights W_i . For two rotations represented by quaternions, \mathbf{q} and \mathbf{q}' , we define a distance metric $d(\mathbf{q}, \mathbf{q}')$ that represents the distance along a curve on the manifold because the vectorial distance $\|\mathbf{q} - \mathbf{q}'\|$ takes no account of the curvature of manifold. The distance metric between two rotations represented by quaternions is defined using the quaternion multiplication as follows:

$$d(\mathbf{q}, \mathbf{q}') = \theta(\mathbf{q}^{-1} \otimes \mathbf{q}') \quad (37a)$$

$$= \theta((q_4 \mathbf{q}'_{13} - \mathbf{q}'_4 \mathbf{q}_{13} - \mathbf{q}_{13} \times \mathbf{q}'_{13}, \mathbf{q} \cdot \mathbf{q}')) \quad (37b)$$

$$= 2 \arccos(\mathbf{q} \cdot \mathbf{q}') \quad (37c)$$

where $\theta(\cdot)$ is the angle of rotation produced by the quaternions. From classical geometry, the geodesics on a sphere

are formed by the great circles and the distance between two points on the unit sphere is therefore the angle between the corresponding vectors [7]. Ignoring the difference in scale, the distance between quaternions is exactly given by Eq. (37). As shown in Eq. (37), the distance between two rotations represented by quaternions is nonlinear function of vectorial representation. We linearize Eq. (37) to compute a closed form approximation of the mean. The following Taylor expansion of cosine is used:

$$\cos\left(\frac{\theta}{2}\right) = 1 - \frac{1}{8}\theta^2 + o(\theta^4) \quad (38)$$

Rearranging this expression yields

$$\theta^2(\mathbf{q}^{-1} \otimes \mathbf{q}') = 8 \left(1 - \cos\left(\frac{\theta(\mathbf{q}^{-1} \otimes \mathbf{q}')}{2}\right) \right) + o(\theta^4) \quad (39a)$$

$$\approx 8(1 - \mathbf{q} \cdot \mathbf{q}') \quad (39b)$$

The mean rotation is now the following quaternion \mathbf{q} that minimizes the quadratic cost function [7]:

$$E^2 = \sum_{i=1}^{N_e} \theta^2(\mathbf{q}^{-1} \otimes \mathbf{q}_i) \quad (40)$$

where N_e is the number of quaternion samples. Using the approximation in Eq. (39b), the quaternion mean $\bar{\mathbf{q}}$ with weights W_i is determined as

$$\bar{\mathbf{q}} = \arg \min_{|\mathbf{q}|=1} \sum_{i=1}^{N_e} W_i \theta^2(\mathbf{q}^{-1} \otimes \mathbf{q}_i) \quad (41a)$$

$$\approx \arg \min_{|\mathbf{q}|=1} \sum_{i=1}^{N_e} 8W_i(1 - \mathbf{q} \cdot \mathbf{q}_i) \quad (41b)$$

$$= \arg \max_{|\mathbf{q}|=1} \sum_{i=1}^{N_e} W_i \mathbf{q} \cdot \mathbf{q}_i = \arg \max_{|\mathbf{q}|=1} \mathbf{q} \cdot \sum_{i=1}^{N_e} W_i \mathbf{q}_i \quad (41c)$$

$$= \frac{\sum_{i=1}^{N_e} W_i \mathbf{q}_i}{\left| \sum_{i=1}^{N_e} W_i \mathbf{q}_i \right|} \quad (41d)$$

The quaternion mean computed by Eq. (41d) is the barycentric mean with renormalization. It should be noted that renormalization is obtained by the true optimization of an approximated cost function and not an ad hoc repair of an approximated optimization.

Now the predicted quaternion $\hat{\mathbf{q}}_{k+1}^-$ is computed using Eqs. (7a), (7c), (35) and (41d) as follows:

$$\hat{\mathbf{q}}_{k+1}^- = \frac{\sum_{i=0}^{12} W_i \mathcal{X}_{i,k+1}^q}{\left| \sum_{i=0}^{12} W_i \mathcal{X}_{i,k+1}^q \right|} \quad (42)$$

Since the quaternion samples are not distributed over all the unit sphere but scattered around the mean quaternion, the predicted quaternion produces a good approximation compared to the mean based on the Riemannian metric.

For $\zeta_{i,k}^\sigma$ with small values the quaternion process noise, i.e., error quaternion given by Eq. (32), is approximated by

$$\delta \mathbf{q}_{i,k}^+ = \begin{bmatrix} \zeta_{i,k}^\sigma / 2 & \sqrt{1 - \zeta_{i,k}^{\sigma T} \zeta_{i,k}^\sigma / 4} \end{bmatrix}^T \quad (43)$$

Comparing this result with Eq. (31), the orientation uncertainty induced by the process noise modeled as the rotation

vector is less than that produced by modeling the process noise as the vector part of a quaternion. In other words, the transformed sigma points for the process noise modeled as the vector part of a quaternion are more widely scattered around the current estimate than that for the process noise treated as a rotation vector. The propagated quaternion samples using Eq. (25) produce the same situation, i.e., the process noise of the rotation vector yields samples with smaller distances from the mean compared to the case where the process noise is modeled as the vector part of a quaternion. Since the approximation is used to derive the quaternion mean given by (41d), samples with smaller distances from the mean yield better approximation of the true mean. Therefore the process noise modeled as the rotation vector will produce better results over the case where it is modeled as the vector part of a quaternion. The performance comparison between both process models is made in Section 4. The predicted covariance of posterior quaternions can not be directly computed using Eq. (6) because the unit quaternion is not closed for subtraction. We can use the error quaternion for predicted covariance computation because it represents the distance from the predicted mean quaternion $\hat{\mathbf{q}}_{k+1}^-$. The error quaternion for predicted covariance computation is given by

$$\Lambda_i = \mathcal{X}_{i,k+1}^q \otimes [\hat{\mathbf{q}}_{k+1}^-]^{-1} \quad i = 0, 1, \dots, 12 \quad (44)$$

Since state vector of the filter includes the vector part of the quaternion, the predicted covariance for the quaternion is then computed as

$$\mathbf{P}_{k+1}^{q-} = \sum_{i=0}^{12} W_i^{(c)} \Lambda_i' \Lambda_i'^T \quad (45)$$

where Λ_i' are the first three components of Λ_i .

The weighted quaternion averaging method can be used again for the predicted observation computation when the measurement model is given by the quaternion, such as a star tracker quaternion output. The predicted measurement covariance is computed for such an output measurement model using Eq. (45) with Λ_i . The cross correlation matrix for the quaternion is given by

$$\mathbf{P}_{k+1}^{qy} = \sum_{i=0}^{12} W_i^{(c)} \Lambda_i' \{ \mathcal{Y}_{i,k+1} - \hat{\mathbf{y}}_{k+1}^- \}^T \quad (46)$$

Simple linear correction applied to quaternion update as in Eq. (14) makes the quaternion depart from the constraint unit sphere. To solve this problem, the quaternion is updated by multiplication. The estimated quaternion $\hat{\mathbf{q}}_{k+1}^+$ is given by

$$\hat{\mathbf{q}}_{k+1}^+ = \delta \mathbf{q}_{k+1}^+ \otimes \hat{\mathbf{q}}_{k+1}^- \quad (47)$$

where $\delta \mathbf{q}_{k+1}^+ \equiv \begin{bmatrix} \delta \mathbf{q}_{13}^+ & \delta \mathbf{q}_4^+ \end{bmatrix}^T$ is represented by

$$\delta \mathbf{q}_{13}^+ = \mathcal{K}_{k+1}(\mathbf{y}_{k+1} - \hat{\mathbf{y}}_{k+1}^-) \quad (48a)$$

$$\delta \mathbf{q}_4^+ = \sqrt{1 - \delta \mathbf{q}_{13}^{+T} \delta \mathbf{q}_{13}^+} \quad (48b)$$

4. Numerical Simulations

In this section, numerical simulations for the new filter are presented using simulated three-axis magnetometer (TAM) and gyro measurements of KOMPSAT-1 (Korea Multi-Purpose Satellite-1) which was launched in 1999 and is still flying in space. The mission orbit of KOMPSAT-1 is a sun-synchronous circular orbit with altitude 685.13km. Orbit inclination is 98.13 deg and eccentricity is 0 to 0.001. The satellite operates in a nominal 10:50 AM $\pm 10/-15$ minutes local time of ascending node. Since the main mission of KOMPSAT-1 is earth observation, the nominal mission mode is earth-pointing requiring a rotation once per orbit about the spacecraft's pitch axis, i.e., y axis. Here we only considered TAM and gyro measurements for attitude estimation. Inertial magnetic field vectors were modeled using the 10th-order International Geomagnetic Reference Field model [22]. Measurement noise of the TAM was modeled by a zero mean white Gaussian process with a standard deviation of 100 nT. The gyroscopic measurements were simulated with an initial bias of 0.1 deg/h for each axis. The gyro and TAM measurements were both sampled at a 1 second interval. The simulations were performed for both process noise cases, i.e., vector part of a quaternion and rotation vector, under the same situations.

In the first simulation, we considered the nominal earth pointing mode. Errors of -180, -60, and 180 deg for roll, pitch, and yaw axis, respectively, were added into the initial condition of the attitude estimate, with an initial bias estimate of 0 deg/h for each axis. Figure 1 depicts the norm of attitude errors for this simulation case. The process noise modeled as a vector part of a quaternion took almost 8 hours to reach 0.1 deg, whereas a rotation vector treated as the process noise reached this value in 1.5 hours. Finally it converged to a value of 0.001 deg within 2.5 hours. When there is large uncertainty in the initial estimates, the results indicate that the process noise as a rotation vector yields much better performance over the case where it is handled as a vector part of quaternion.

In the second simulation, a tumbling situation after separation from launch vehicle was considered. Attitude estimation is not so easy in tumbling situation because initial attitude errors are generally large and angular velocity of each axis is much higher than that of the earth pointing mode. The maximum allowable tip-off rate for KOMPSAT-1 was 2 deg/sec about an arbitrary axis. Tip-off rates of 2, -2, and 2 deg/sec, respectively, were set with initial bias of 0.1 deg/h for each axis. A plot of the norm of attitude errors of spacecraft's tumbling case is shown in Figure 2. The attitude errors did not reach the value obtained in the nominal earth pointing mode because angular velocity of each axis is much larger than that of the earth pointing mode, leading to rapid attitude changes. However, we can see that the process noise modeled as a rotation vector produces better performance over the case for which it is formulated as a vector part of quaternion. The filter treating rotation vector as the process noise reduces the attitude error of 0.2 deg, whereas the attitude error for the process noise handled as a vector part

of quaternion remains within error bounds of 1 deg. Comparing to the result of the first simulation for the process noise as vector part quaternion, the attitude error of 0.2 deg is acceptable in the tumbling mode. This simulation result illustrates that attitude estimation in tumbling mode is achieved with some degree of error bounds, and the process noise in the form of rotation vector produces better result compared to the case where it is modeled as vector part of quaternion.

From simulations, it can be said that in the presence of large uncertainties in the initial conditions and in the situation of high tip-off rates, the process noise as a rotation vector is more optimal than modeling it as a vector part of quaternion.

5. Conclusions

In this paper a new approach to the straightforward implementation of UKF for spacecraft attitude estimation in unit quaternion space was presented. Since the UKF formulation is built in vector space, weighted mean computation for quaternions did not produce an estimate in unit quaternion space. To solve this problem, weighted mean computation method for quaternions in rotational space was derived. The quaternion multiplication was used for predicted covariance computation and quaternion update, which makes quaternion in the filter lie in unit quaternion space. To evenly scatter the quaternion samples on the unit sphere around the current quaternion estimate, both error quaternion and its inverse quaternion were used to construct sigma point quaternions. Numerical simulation results showed that the proposed approach successfully estimates attitude of spacecraft for large initial errors and high tip-off rates. Since the process noise for quaternion resulted in increase of the uncertainty in attitude orientation, either modeling it as a vector part of quaternion or as a rotation vector was considered. From simulation results, modeling the quaternion process noise as a rotation vector is more optimal than treating it as a vector part of quaternion.

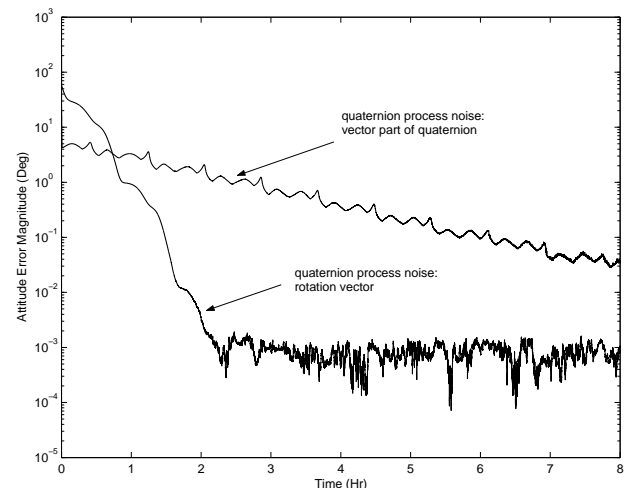


Fig. 1. Attitude Estimation Errors: Earth-Pointing Mode.

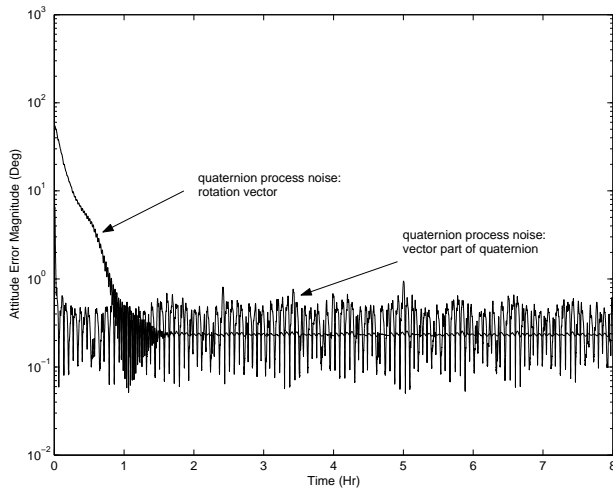


Fig. 2. Attitude Estimation Errors: Tumbling Mode.

References

- [1] S. J. Julier, "The Scaled Unscented Transformation," *Proceedings of the American Control Conference*, American Automatic Control Council, Evanston, IL, 2002, pp. 1108–1114.
- [2] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A New Method for the Nonlinear Transformation of Means and Covariances in Filters and Estimators," *IEEE Transaction on Automatic Control*, Vol. AC-45, No. 3, 2000, pp. 477–482.
- [3] S. J. Julier, and J. K. Uhlmann, "A New Extension of the Kalman Filter to Nonlinear Systems," *Signal Processing, Sensor Fusion, and Target Recognition VI*, edited by I. Kadar, Proceedings of the SPIE, Vol. 3068, Society of Photo-Optical Instrumentation Engineers, Bellingham, WA, 1997, pp. 182–193.
- [4] E.A. Wan, and R. van der Merwe, "The Unscented Kalman Filter," *Kalman Filtering in Neural Networks*, edited by S. Haykin, Wiley, New York, 2001, Chap. 7.
- [5] J. L. Crassidis, and F. L. Markley, "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control and Dynamics*, Vol. 26, No. 4, 2003, pp. 536–542.
- [6] R. L. Farrenkopf, "Analytic Steady-State Accuracy Solutions for Two Common Spacecraft Attitude Estimators," *Journal of Guidance, Control and Dynamics*, Vol. 37, No. 1, 1978, pp. 282–284.
- [7] C. Gramkow, "On Averaging Rotations," *International Journal of Computer Vision*, Vol. 42(1/2), 2001, pp. 7–16.
- [8] M. Moakher, "Means and Averaging in the Group of Rotations," *SIAM Journal on Matrix Analysis and Applications*, Vol. 24, No. 1, 2002, pp. 1–16.
- [9] W. D. Curtis, A. L. Janin, and K. Zikan, "A Note on Averaging Rotations," *Virtual Reality Annual International Symposium*, IEEE, Sept. 1993, pp. 377–385.
- [10] J.-W. Kim, J. L. Crassidis, S.R. Vadali, and A.L. Dershowitz, "International Space Station Leak Localization Using Attitude Disturbance Estimation," *IEEE Aerospace Conferences*, Big Sky, Montana, 2003, #01-3475.
- [11] J. J. LaViola, "A Comparison of Unscented and Extended Kalman Filtering for Estimating Quaternion Motion", *In the Proceedings of the 2003 American Control Conference*, IEEE Press, June 2003, pp. 2435–2440.
- [12] F. D. Busse, "Demonstration of Adaptive Extended Kalman Filter for Low Earth Orbit Estimation Using DGPS", *Institute of Navigation GPS Meeting*, Sept. 2002.
- [13] D.-J. Lee, and K. T. Alfriend, "Adaptive Sigma Point Filtering for State and Parameter Estimation", *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, Providence, Rhode Island, August 16–19, 2004.
- [14] F. L. Markley, "Attitude Error Representation for Kalman Filtering," *Journal of Guidance, Control and Dynamics*, Vol. 26, No. 2, 2003, pp. 311–317.
- [15] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control and Dynamics*, Vol. 5, No. 5, 1982, pp. 417–429.
- [16] J. L. Crassidis, and F. L. Markley, "Attitude Estimation Using Modified Rodrigues Parameters," *Proceedings of the Flight Mechanics/Estimation Theory Symposium*, NASA Goddard Space Flight Center, Greenbelt, MD, May 1996, pp. 71–83.
- [17] N. J. Kasdin, "Satellite Quaternion Estimation from Vector Measurement with the Two-Step Optimal Estimator," *American Astronautical Society*, AAS Paper 02-002, Feb. 2002.
- [18] M. L. Psiaki, "Attitude-Determination Filtering via Extended Quaternion Estimation," *Journal of Guidance, Control and Dynamics*, Vol. 23, No. 2, 2000, pp. 206–214.
- [19] S. J. Julier, and H. F. Durrant-Whyte, "Navigation and Parameter Estimation of High Speed Road Vehicles," *Robotics and Automation Conference*, 1995, pp. 101–105.
- [20] B. Stenger, P. R. S. Mendonca, and R. Cipolla, "Model-Based Hand Tracking Using an Unscented Kalman Filter," *Proceedings of the British Machine Vision Conference*, Sept. 2001, pp. 63–72.
- [21] M. D. Shuster, "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, 1993, pp. 439–517.
- [22] C. E. Barton, "International Geomagnetic Reference Field: The Seventh Generation," *Journal of Geomagnetism and Geoelectricity*, Vol. 49, 1997, pp. 123–148.