Spacecraft Attitude Estimation using Huber-based Kalman Filter

Wei Li, Meihong Liu, Deren Gong and Dengping Duan
Institute of Aerospace Science and Technology, Shanghai Jiao Tong University, China

Abstract: A Huber-based Kalman Filter (HKF) is presented to non-Gaussian random measurement errors. The measurement noise uncertainty is tackled at each filter step by minimizing a criterion function that is original from Huber technique. A recursive algorithm is also provided to solve the criterion function. The proposed HKF algorithm has been tested in attitude estimation using gyroscopic and star tracker sensors for a single spacecraft in flight simulations. Simulation results demonstrate the superior performance of the proposed filter as compared to the standard Kalman Filter (KF) in the presence of non-Gaussian measurement noise.

Key words: Huber-based Kalman filter, spacecraft, attitude estimation

INTRODUCTION

The attitude estimation system consisted of gyroscope and star tracker sensors are widely used in the spacecraft with precision pointing requirement. The measurements of the gyroscope and the star tracker sensors are usually combined with a Kalman Filter (KF) (Lefferts et al., 1982; Crassidis et al., 2007). However, The Kalman filter is a recursive weighted least squares or minimum l_2 norm estimation (sample mean) procedure and is a maximum likelihood technique assuming that the error statistics follow the Gaussian density distributions. It is well known that the performance of the Kalman filter can rapidly degrade when the true error statistics follow non-Gaussian probability distributions, particularly those with much thicker tails than the Gaussian probability distribution (Tukey, 1960; Huber, 1972). Since the typical sensor systems are known to exhibit non-Gaussian characteristics (Kuhl, 2005; Oshman and Carmi, 2006) as well as uncertain noise statistics (Lam et al., 2002), it is essential to develop filtering techniques that are robust with respect to deviations from the assumption of Gaussianity.

Huber (1972, 1964) has developed a robust filter approach to deal with the unknown measurement noise which is a combined l_1 and l_2 norm estimator and it has been successfully applied to relative navigation filter design for robust rendezvous in elliptic orbit and tracking problems (Karlaard, 2006; Karlgaard and Schaub, 2007). The minimum l_1 norm estimator (sample median) is a maximum likelihood estimator assuming that the error statistics follow the Laplacian probability distribution. The Huber estimator is a combine of the two estimators that seeks to utilize the best of both methods, in particular, the robustness of the sample median and the efficiency of the sample mean.

By using the similar method provided by Huber (1972, 1964), a Huber-based Kalman filter is developed to apply to spacecraft attitude estimation problem. Numerical simulation results indicate that the proposed HKF filter performs better in robustness and accuracy compared with typical Kalman filter algorithm in the presence of non-Gaussian measurement errors.

HUBER-BASED KALMAN FILTER

Typical Kalman filter: This section provides a brief summary of the typical Kalman filter. The class of systems considered here is given by:

\[ x_{k+1} = F_k x_k + g_k w_k \]  \hspace{1cm} (1)
\[ y_k = H_k x_k + v_k \]  \hspace{1cm} (2)

where, \( x_k \in \mathbb{R}^n \) is the state vector, \( y_k \in \mathbb{R}^p \) is the output vector, \( w_k \) and \( v_k \) are, respectively, the disturbance and sensor noise vector which are assumed to be independent, identically distributed Gaussian white noise with:

\[ w_k \sim N(0, Q_k) \quad v_k \sim N(0, R_k) \]  \hspace{1cm} (3)

The Kalman filter is a recursive scheme that propagates a current estimate of a state and the error covariance matrix of the state forward in time. The filter optimally blends the new information introduced by the

Corresponding Author: Deren Gong, Institute of Aerospace Science and Technology, Shanghai Jiao Tong University, China
Tel: +86 1880196485Fax: +86 021 34207170

1234
measurements with old information embodied in the prior state with a Kalman gain matrix. The gain matrix balances uncertainty in the measurements with the uncertainty in the dynamics model (Bar-Shalom et al., 2001). Since the common knowledge of Kalman filter derivation, here we omit it for concise.

**Huber-based Kalman filter:** For spacecraft attitude estimation problem, the measurements of the typical sensors are well known to exhibit non-Gaussian characteristics (Kuhl, 2005; Oshman and Carmi, 2006) and uncertain noise statistics (Lam et al., 2002). Some modification should be augmented to the filtering procedure to deal with unknown measurement noise.

Huber (1972, 1964) has developed a recursive method for handling measurement noise uncertainty by a weighted linear combination of the predicted state and the actual measurement. Bonelet and Dickinson (1983) has proposed a robust filtering algorithm at each measurement update step as a sequence of linear regression problem between the predicted state and the observed quantity. In addition, Karlgaard (2006) and Karlgaard and Schaub (2007) has successfully applied this filtering technique to a robust navigation filter design in lunar orbit and tracking problems.

In order to apply this method, it is imperative to recast the measurement update as a regression problem between the predicted state and observed quantity. If the true value of the state is \( x_0 \), then the state error can be expressed as \( \delta x = x_0 - \bar{x}_0 \). The linear regression problem then has the following form:

\[
\begin{bmatrix}
    y_1 \\
    x_1
\end{bmatrix} = \begin{bmatrix}
    H_k & 0 \\
    0 & 1
\end{bmatrix} \begin{bmatrix}
    x_1 \\
    \delta x
\end{bmatrix} + \begin{bmatrix}
    v_1 \\
\end{bmatrix}
\]

By defining the following quantities:

\[
T_k = \begin{bmatrix}
    R(k) & 0 \\
    0 & P(k)
\end{bmatrix}
\]

\[
M_k = T_k^{-1/2} \begin{bmatrix}
    H_k \\
    1
\end{bmatrix}
\]

\[
z_k = T_k^{-1/2} \begin{bmatrix}
    y_1 \\
    x_1
\end{bmatrix}
\]

\[
\bar{z}_k = T_k^{-1/2} \begin{bmatrix}
    \bar{y}_1 \\
    \bar{x}_1
\end{bmatrix}
\]

then the linear regression problem is transformed to the form as:

\[
z_k = M_k \bar{z}_k + \xi_k
\]

where, the covariance of \( \xi_k \) is simply the identity matrix. This regression problem can be solved by using Huber’s generalized maximum likelihood method and the solution is attained by minimizing the cost function:

\[
J(\xi_k) = \sum_{i=1}^{n} u(\xi_i)
\]

where, \( \xi_i \) is the ith component of the residual vector \( \xi = (M_k \bar{x}_k - \bar{z}_k) \), \( n \) is the dimension of the residual \( \xi \) and the function \( u \) is known as the cost function, written as:

\[
u(\xi_i) = \begin{cases}
    \frac{1}{2} \xi_i^2, & |\xi_i| < \gamma \\
    \frac{\gamma}{2} |\xi_i| - \frac{\gamma^2}{4}, & |\xi_i| \geq \gamma
\end{cases}
\]

where, \( \gamma \) is a tuning parameter. This \( u \) function is a blend of the \( l_1 \) and \( l_2 \) norm functions and estimates obtained from the use of this \( u \) function achieved desirable robustness. If this function \( u \) is differentiable which is the case for Eq. 11 and then the solution to this linear regression problem is given implicitly by the derivative of the cost function:

\[
\frac{\partial J(\xi_k)}{\partial \xi} = \frac{\partial u(\xi_k)}{\partial \xi} \begin{bmatrix}
    \xi_1 \\
    \vdots \\
    \xi_n
\end{bmatrix} = \begin{bmatrix}
    \frac{\partial u(\xi_k)}{\partial \xi_k} |_{|\xi_k| < \gamma} \\
    \frac{\partial u(\xi_k)}{\partial \xi_k} |_{|\xi_k| \geq \gamma}
\end{bmatrix} = 0
\]

Note:

\[
u(\xi_k) = \frac{\partial u(\xi_k)}{\partial \xi_k} \begin{bmatrix}
    \xi_1 \\
    \vdots \\
    \xi_n
\end{bmatrix} = \begin{bmatrix}
    \frac{\partial u(\xi_k)}{\partial \xi_k} |_{|\xi_k| < \gamma} \\
    \frac{\partial u(\xi_k)}{\partial \xi_k} |_{|\xi_k| \geq \gamma}
\end{bmatrix}
\]

The solution of Eq. 12 is typically solved by using one of two approaches: a Newton-Raphson iteration, or the iteratively reweighted approach which is developed by Beaton and Tukey (1974). In this study, the latter algorithm is used. The approach may be derived as follows. First, defining the function \( \Psi(\xi_k) = u(\xi_k)/\xi_k \), the matrix \( \Psi = \text{diag} \{ \Psi(\xi_k) \} \) and substituting \( \xi_k = (M_k \bar{x}_k - \bar{z}_k) \), into equation to yield the following matrix:

\[
M_k^T \Psi (M_k \bar{x}_k - \bar{z}_k) = 0
\]

Equation 14 may be expanded to yield

\[
M_k^T \Psi M_k \bar{x}_k = M_k^T \Psi \bar{z}_k
\]

and the iterative solution can be expressed as:

\[
x^{(j+1)} = (M_k^T \Psi M_k)^{-1} M_k^T \Psi \bar{z}_k
\]

where the superscript \( j \) is the iteration index. This approach can be initialized by using the least squares solution \( x^{(0)} = (M_k^T M_k)^{-1} M_k^T \bar{z}_k \). The converged value from the iterative procedure is taken as the a posteriori state.
estimate following a measurement update $\hat{x}_e$. Finally, the a posterior state estimate error covariance matrix is calculated from:

$$\hat{P}(k) = (M^T \Psi M)^{-1}$$  \hspace{1cm} (16)

using the final value of $\Psi$ corresponding to the converged state estimate, with the residual that adopts the following form:

$$\zeta = T_k \left( \frac{y_k - x_k}{\sigma^2_k} \right)$$  \hspace{1cm} (17)

where, the superscript (0) is the initial state estimate calculated from the typical Kalman filter measurement update.

Notice that as $\gamma \to \infty$, the Huber-based Kalman filtering technique reduces to the least squares estimator, particularly, when $\gamma \to 0$, the matrix $\Psi - I$, this filter measurement updates is identical to the typical Kalman filter. Also notice that as $\gamma \to 0$, this filtering technique reduces to the absolute value estimator. This blend of estimation approaches makes the filter more robust against deviations from Gaussian distributed random measurements errors. The robustness arises from the $\Psi$ matrix and that all residuals are not weighted equally. Specifically, large residuals are down-weighted in the iterative process by the inverse of the magnitude of the residual (Hampel, 1974; Hampel et al., 1986).

APPLICATION TO SPACECRAFT ATTITUDE ESTIMATION

Attitude kinematics: The attitude dynamics of spacecraft are well understood and can be expressed in many forms. Maybe the most attractive form for attitude estimation is the quaternion representation (Hamilton, 1866) which is the smallest globally non-singular attitude parameterization. The quaternion is a four-dimensional vector, written as:

$$q = \begin{bmatrix} \theta \\ q_u \end{bmatrix}$$  \hspace{1cm} (18)

with:

$$\theta = \sqrt{q_u^2} = \sin \left( \frac{\phi}{2} \right)$$  \hspace{1cm} (19)

$$q_u = \cos \frac{\phi}{2}$$  \hspace{1cm} (20)

where, $\phi$ is the Euler axis and $\phi$ is the rotation angle. Since a four-dimensional vector is adopted to describe three dimensions, the quaternion components can’t be independent of each other.

The quaternion kinematic equations is given as (Schaub and Junkins, 2003):

$$\dot{q} = \frac{1}{2} \Omega(q) q$$  \hspace{1cm} (21)

where, $\omega$ is the angular velocity of the body frame with respect to the inertial frame and:

$$\Omega(q) = \begin{bmatrix} -[\omega \times] & \omega \\ -\omega & 0 \end{bmatrix}$$  \hspace{1cm} (22)

also, $[\omega \times]$ is the cross-product matrix defined by:

$$[\omega \times] = \begin{bmatrix} 0 & -\omega_z & \omega_y \\ \omega_z & 0 & -\omega_x \\ -\omega_y & \omega_x & 0 \end{bmatrix}$$  \hspace{1cm} (23)

One major advantage of the use of the quaternion is that the kinematics equation is linear in the quaternion and is also free of singularities. The other advantage of the quaternion is that successive rotations can be tackled using quaternion multiplication. In this paper, the rule of Lefferts et al. (1982) is utilized, where the quaternions are multiplied in the same order as the attitude matrix multiplication, in contrast to the usual convention proposed by Hamilton (1866). A successive rotation is given using:

$$A(q) A(q') = A(q' \circ q)$$  \hspace{1cm} (24)

The composition of the quaternions is bilinear, with:

$$q' \circ q = [\Psi(q') q] q = [\Xi(q') q] q'$$  \hspace{1cm} (25)

$$\Psi(q') = \begin{bmatrix} q_{33}^2 - [\theta^2] \\ -\theta^2 \end{bmatrix}$$  \hspace{1cm} (26)

$$\Xi(q') = \begin{bmatrix} q_{33}^2 + [\theta^2] \\ -\theta^2 \end{bmatrix}$$  \hspace{1cm} (27)

and the inversion rule is defined by:

$$q' = \begin{bmatrix} -\theta \\ q_u \end{bmatrix}$$  \hspace{1cm} (28)

Gyroscope model: The commonly used sensor that measures angular rate is rate integrating gyroscope such as fibre optic gyros which can be represented.
mathematically by employing Farrenkopf's model (Farrenkopf, 1978). For this sensor, the measured angular velocity is expressed as the true angular velocity with an additive bias and white noise. The bias item is a slowly varying parameter driven by Gaussian white noise. The model is given by:

\[ \dot{\omega} = \omega + \beta + \eta \beta = \eta \]

(29)

where, \( \dot{\omega} \) is the measured inertial angular velocity, \( \omega \) is the true inertial angular velocity, \( \beta \) is the gyroscope bias and \( \eta \) is independent zero-mean Gaussian white noise processes with:

\[
\begin{align*}
E\{ \mathbf{h}(\mathbf{t}) \mathbf{q}^{\top}(\mathbf{t}) \} &= \mathbf{0} + \sigma_\beta^2 \delta(t - t) \mathbf{I}_{3x3} \\
E\{ \mathbf{h}(\mathbf{t}) \mathbf{q}^{\top}(\mathbf{t}) - \sigma_\beta^2 \delta(t - t) \mathbf{I}_{3x3} \}
\end{align*}
\]

(30)

where, \( E\{\mathbf{\cdot}\} \) denotes expectation and \( \delta(t-t) \) represents the Dirac delta function. A more general gyroscope model contains scale factors and misalignments which can also be estimated in real time (Pittelkau, 2001; Pandiyar et al., 2004). For simulation purposes, discrete-time gyroscope measurements can be generated according to the model proposed by Farrenkopf using the following equations:

\[
\begin{align*}
\mathbf{q}_k &= \mathbf{q}_k + \Delta t (\mathbf{q}_k + \mathbf{b}_k) + \frac{1}{2} \left( \frac{\mathbf{q}_k^2 + \mathbf{q}_k^2}{\Delta t} \right) \mathbf{I}_3 \\
\mathbf{b}_k &= \mathbf{b}_k + \sigma_\beta \Delta t \mathbf{I}_3
\end{align*}
\]

(31)

where, the subscript \( k \) denotes the time instant \( k \), \( \Delta t = t_{k+1} - t_k \) is the gyroscope sampling interval, \( \eta \) and \( \beta \) are zero-mean Gaussian white noise processes with covariance each given by the identity matrix.

**Star tracker model:** It is assumed that a star tracker or some other attitude sensor is available to provide corrections to the attitude estimates formed by direct integration of the angular velocity measurements which suffer error build-up due to integrating errors in the estimated bias and the random errors. The star tracker is assumed to output an estimated quaternion that associates the orientation of the body with the inertial frame.

A star tracker model is given by Pittelkau (2002):

\[ \mathbf{q}_n = \mathbf{h}_n (\Delta \mathbf{s}) \mathbf{q} \]

(32)

where, \( \mathbf{q}_n \) denotes star tracker output quaternion, \( \mathbf{q} \) is the quaternion that represents the true orientation and \( \Delta \mathbf{s} \) is an error quaternion parameterized by a random angular error \( \Delta \mathbf{s} \). Continuously measured error quaternion vector, \( \delta \mathbf{q} = \eta \), is modelled as an independent zero-mean Gaussian white noise process of the measurement noise vector with:

\[ E\{ \eta(t) \eta^{\top}(t) \} = \sigma_\eta^2 \delta(t-t) \mathbf{I}_{3x3} \]

(33)

**Attitude filter implementation:** The filtering of the attitude quaternion is complicated by the unit norm constraint. One common method to deal with the constraint is to employ a multiplicative approach for the measurement update. In this approach, the attitude quaternion is replaced by an error quaternion, \( \Delta \mathbf{q} \) which allows for the processing of the error quaternion as a three component vector of small angles, denoted by \( \delta \). In filtering process, the quaternion estimate is propagated to the star tracker measurement at \( k_0 \) time-step using Eq. 38. Then, an error quaternion relating the star tracker measurement quaternion to the predicted quaternion based on propagation from the previous measurement is written as (Markley, 2001):

\[ \delta \mathbf{q}_{\text{in}} = \mathbf{q}_n \mathbf{\otimes} \mathbf{q}^{-1}_n \]

(34)

The first three components of this error quaternion are processed as independent small angles with the measurement matrix \( \mathbf{H} = [I_{3x3} \ 0_{3x3}] \).

The small angle error and gyroscope bias estimate error dynamics are given as follows:

\[ \begin{bmatrix} \dot{\mathbf{\delta} s} \mathbf{\delta s} \mathbf{\delta s} \mathbf{v} \mathbf{n} \end{bmatrix} \]

(35)

By defining \( \mathbf{x} = [\delta \mathbf{s} \ \delta \mathbf{\beta}]^{\top} \), then these differential equations are in a suitable form for filter implementation corresponding to section 2. The measurement update of quaternion and gyroscope bias estimates are given by:

\[ \begin{align*}
\mathbf{\delta s} &= \mathbf{\delta s} \otimes \mathbf{\delta s} \\
\mathbf{\delta \beta} &= \mathbf{\delta \beta} + \mathbf{\delta \beta}
\end{align*} \]

(36)

with:

\[ \mathbf{\delta s} \]

(37)

State propagation in between star tracker updates is based on the assumption that the angular velocity is constant over each gyroscope sample interval. First, the estimated angular velocity of the spacecraft at time instant \( k \), denoted by \( \dot{\omega}_k \), can be written as \( \dot{\omega}_k = \dot{\omega}_k - \dot{\beta}_k \), where, \( \dot{\beta}_k \) is the estimated gyroscope measurement bias at time instant \( k \). If the gyroscope sampling rate is high enough,
then the angular velocity can reasonably be assumed to be constant over the sampling interval. Based on this assumption, the quaternion kinematic equations become constant coefficient ordinary differential equations and so the solution can be obtained from the matrix exponential. Obviously, the solution to the predicted quaternion at time instant $k+1$ is given by Kane (1973):

$$
\hat{q}_k = \int e^{\omega_k \Delta t_2} \left( \frac{\Delta \hat{q}}{2} - \frac{\hat{q}_k}{2} \hat{\Omega}(\hat{q}_k) \right) dt
$$

(38)

where $\hat{q}_k$ is the estimated quaternion at time $k$, $\hat{q}_{k+1}$ is the predicted quaternion at time $k+1$ and $|\hat{\omega}_k|$ is the assumed constant angular velocity between time $k$ and $k+1$.

**SIMULATIONS**

This section provides the results of the application of the filtering technique discussed in previous sections to the spacecraft attitude estimation problem using gyroscope and star trackers. In the simulation examples, the true spacecraft angular velocity is $\omega = [\omega_x \ \omega_y \ \omega_z]^T$, for $\omega = 10^{-3}$ rad sec$^{-1}$. The star tracker and gyroscope measurements are both sampled at 1 sec interval. The sensor uncertainty of gyroscope and star tracker are simulated using Eq. 30 and 33 with $\sigma_{\omega} = 2.25 \times 10^{-11}$ rad$^2$ sec$^{-1}$, $\sigma_{\omega} = 2.25 \times 10^{-19}$ rad$^2$ s$^{-3}$ and $\sigma_{\omega} = 3.81 \times 10^{-12}$ rad$^3$. The initial bias on each axis is set to 0.1$^\circ$ h$^{-1}$. The tuning parameter $\gamma$ is set to 1.345 (Karlgaard, 2006).

This section discusses the results of a 500 case Monte-Carlo simulation of the proposed HKF and KF with the specified noise variances provided in the above paragraph. Random measurement errors for the gyroscope and star tracker measurements are drawn from a mixture model, defined by the probability density function:

$$
p(\xi) = 1 - \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{\xi^2}{2} \right) - \frac{1}{2b} \exp \left( -\frac{\xi}{b} \right)
$$

(39)

where $b$ is set to $\xi/\sqrt{2}$ and $\xi$ is a perturbing parameter that denotes error model contamination.

For the simulation, both HKF and the KF are performed with same initial errors; the perturbing parameter $\xi$ is set to 0.5, for this case in which the random measurement errors are highly non-Gaussian and the 2-norm of the average attitude errors are shown in Fig. 1. The red solid line ("KF" in legend) represents the estimated attitude errors using the typical Kalman filter. The black dashed line ("HKF" in legend) denotes the attitude errors using the Huber-based Kalman filter for unknown measurement noise introduced in section.

**CONCLUSION**

For improving the performance of typical Kalman filter with unknown measurement noise, a Huber-based Kalman filter is proposed. The measurement noise uncertainty is dealt with in real time by solving a criterion function that is original from Huber technique, also, a recursive algorithm is provided for solving the criterion function. The proposed Huber-based Kalman filter has been tested in attitude estimation system using gyroscope and star tracker sensors for single spacecraft in flight simulations. Numerical simulation results indicate that the proposed filter performs better than Kalman filter algorithm in terms of accuracy and robustness in the presence of non-Gaussian measurement noise.
ACKNOWLEDGMENT

This study was partially supported by the national natural science foundation of china [grant number 11272205] and china postdoctoral science foundation [grant No. 2012M520899].

REFERENCES


