Attitude Estimation for Large Field-of-View Sensors

Yang Cheng*, John L. Crassidis† and F. Landis Markley‡

Abstract

The QUEST measurement noise model for unit vector observations has been widely used in spacecraft attitude estimation for more than twenty years. It was derived under the approximation that the noise lies in the tangent plane of the respective unit vector and is axially symmetrically distributed about the vector. For large field-of-view sensors, however, this approximation may be poor, especially when the measurement falls near the edge of the field of view. In this paper a new measurement noise model is derived based on a realistic noise distribution in the focal plane of a large field-of-view sensor, which shows significant differences from the QUEST model for unit vector observations far away from the sensor boresight. An extended Kalman filter for attitude estimation is then designed with the new measurement noise model. Simulation results show that with the new measurement model the extended Kalman filter achieves better estimation performance using large field-of-view sensor observations.

Introduction

Attitude determination is the identification of a proper orthogonal rotation (attitude) matrix that maps sensed vectors from a reference frame into the sensor frame. If all the measured and reference vectors are error free, then the rotation matrix is the same for all sets of observations. However, if measurement errors exist, then a least-squares type approach must be used to determine the attitude. Several attitude sensors exist, including: three-axis

*Postdoctoral Research Fellow, Department of Mechanical & Aerospace Engineering, University at Buffalo, State University of New York, Amherst, NY 14260-4400. E-mail: cheng3@eng.buffalo.edu.

†Associate Professor, Department of Mechanical & Aerospace Engineering, University at Buffalo, State University of New York, Amherst, NY 14260-4400. E-mail: johnc@eng.buffalo.edu.

‡Aerospace Engineer, NASA Goddard Space Flight Center, Guidance, Navigation and Control Systems Engineering Branch, Greenbelt, MD 20771. E-mail: Landis.Markley@nasa.gov.

magnetometers, sun sensors, Earth-horizon sensors, global positioning system (GPS) sensors and star trackers. Reference [1] provides detailed descriptions of each of these sensors. The specific choice for the complement of onboard attitude sensor hardware is mostly driven by the individual requirements of the spacecraft mission. For example, for low accuracy requirements, such as a few degrees, a three-axis magnetometer can be used solely to determine three-axis attitude coupled with gyroscopes or a dynamic model [2]. The Solar, Anomalous, and Magnetospheric Particle Explorer (SAMPEX) is an example of a highly successful mission that employs only a three-axis magnetometer and a coarse (0.25 degree) sun sensor coupled with a dynamic model [3].

For missions with tight attitude knowledge requirements, the primary means to determine attitude is the star tracker. The technology behind star trackers has changed much over the years. Evolving from gimballed to fixed-head, the latest star trackers now use charge-coupled devices (CCD) for imaging, which offer high accuracy [4]. By using star image centroiding, accuracies of approximately 1/10 the size of a pixel can be achieved. For small field-of-view (FOV) star trackers, this leads to off-boresight attitude knowledge on the order of 10 to 20 arcsec. Star trackers fall into the category of line-of-sight (LOS) sensors because they measure the direction of a celestial body. In particular, the angle of that body is measured from the sensor boresight in two mutually orthogonal planes [5]. With the advent of low-cost CCD arrays and powerful processors, the use of star trackers is more common today.

For LOS sensors the observation equations are given by the well-known collinearity equations [6], which relate image plane coordinates to object plane coordinates through an attitude rotation. For stellar applications the light sources can be treated as infinite distance points, so that the only unknown, once a star is identified, is the attitude matrix. All attitude sensors, including star trackers, contain noise in their measurements however. This noise includes both systematic errors and random errors. Systematic errors are reduced through calibration procedures, which can even be done on-orbit [7]. Random errors are usually treated as zero-mean Gaussian white-noise processes with known covariance. A realistic covariance matrix takes into account an increase in the errors away from the boresight

due to radial distortions and contains correlated terms. A frequently-used covariance model for the noise added to the collinearity truth equations is given by Eq. (10) of Ref. [8]. This model is referred to here as the "focal-plane model."

The collinearity equations are usually cast in vector form because the attitude matrix appears linearily in this form. This unit vector form, also called the "LOS measurement model," is the most widely used observation equation in attitude determination [9]. Recent research has also shown that using the unit vector form produces better results in a filter design over the standard collinearity equations form for the observation equations [10]. This is due to lower nonlinearity of the LOS measurement model and boundedness of the LOS measurements. Unfortunately, the measurement noise is also transformed when the collinearity equations are converted into the unit vector form. A simple covariance model that is valid for small FOVs has been developed by Shuster and Oh [9], called the "QUEST measurement model," which is a singular matrix that arises for the unit-normalization of the observations. The beauty of this model is in its simplicity, which is evident in its use in the extended Kalman filter (EKF). In particular, Shuster [8] has shown that the singular covariance matrix can effectively be replaced with a nonsingular isotropic matrix, thereby providing practical use in the EKF.

In this paper the QUEST measurement model is replaced with a general model that is valid for large FOVs. New sensors are evolving that incorporate wider FOVs, which may lead to degraded performance when using the QUEST measurement model. One such sensor is the vision-based navigation (VISNAV) system [11], which comprises an optical sensor of a new kind combined with specific light sources (beacons) in order to achieve a selective or "intelligent" vision. The sensor is made up of a Position Sensing Diode (PSD) placed in the focal plane of a wide angle lens, which yields a 100 degree FOV. When the rectangular silicon area of the PSD is illuminated by energy from a beacon focused by the lens, it generates electrical currents in four directions that can be processed with appropriate electronic equipment to estimate the energy centroid of the image. The new measurement model is derived using a first-order Taylor series expansion approach of the observation noise

model. This makes the assumption that the noise is "small" compared to the signal. It will be shown that this model can produce more accurate results than the QUEST measurement model for the VISNAV sensor. Unfortunately, as is the case with the QUEST measurement model, the new covariance is also singular, which causes a problem in the computation of the Kalman gain. To overcome this problem, two solutions are presented. The first is based on a matrix decomposition of the new covariance and the second is based on a rank-one update.

The organization of this paper proceeds as follows. First, the collinearity equations are summarized, followed by the introduction of the focal-plane covariance model. Then, the QUEST measurement model is reviewed. Next, a new covariance model is derived that is valid for large FOV sensors. Its implementation in an EKF design is then shown by using a measurement transformation approach as well as a rank-one update approach. Finally, simulation results are shown that compare EKF results with the QUEST model versus the new model using synthetic VISNAV sensor observations.

Overview

In this section the collinearity equations are summarized for close range photogrammetry applications using the VISNAV sensor. A covariance model for the focal-plane equations is also summarized and the QUEST measurement model is shown.

Collinearity Equations

Photogrammetry is the technique of measuring objects (2D or 3D) from photographic images or LOS measurements. Photogrammetry can generally be divided into two categories: far range photogrammetry with camera distance settings to infinity (commonly used in star trackers), and close range photogrammetry with camera distance settings to finite values. In general close range photogrammetry can be used to determine both the position and attitude of an object, while far range photogrammetry can only be used to determine attitude. The VISNAV system comprises an optical sensor of a new kind combined with specific light

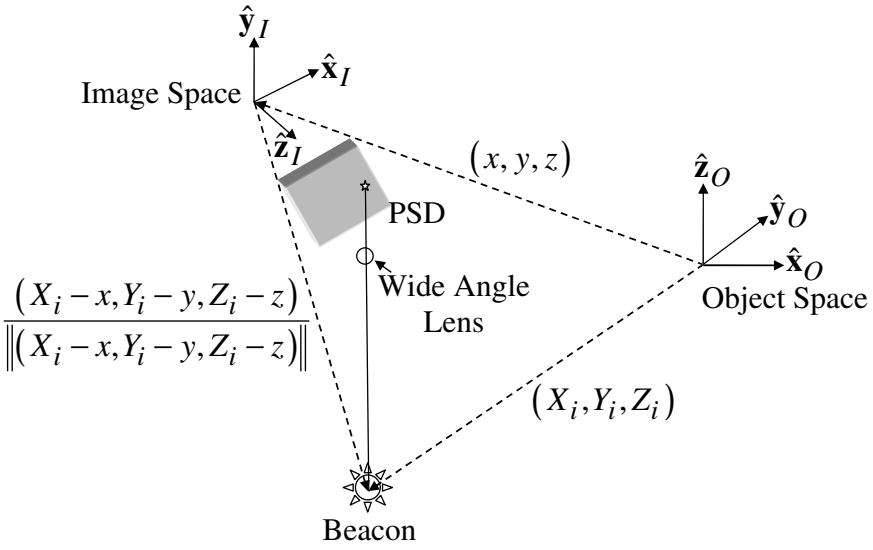


Figure 1. Vision Based Navigation System

sources (beacons), which can be used for close range photogrammetry-type applications. The relationship between the position/attitude and the observations used in photogrammetry involves a set of collinearity equations, which are reviewed in this section. Figure 1 shows a schematic of the typical quantities involved in basic photogrammetry from LOS measurements, derived from light beacons in this case. If we choose the z-axis of the sensor coordinate system to be directed outward along the boresight, then given object space and image space coordinate frames (see Fig. 1), the ideal object to image space projective transformation (noiseless) can be written as follows [6]:

$$\alpha_i = -f \frac{A_{11}(X_i - x) + A_{12}(Y_i - y) + A_{13}(Z_i - z)}{A_{31}(X_i - x) + A_{32}(Y_i - y) + A_{33}(Z_i - z)}, \quad i = 1, 2, \ldots, N \tag{1a}$$

$$\beta_i = -f \frac{A_{21}(X_i - x) + A_{22}(Y_i - y) + A_{23}(Z_i - z)}{A_{31}(X_i - x) + A_{32}(Y_i - y) + A_{33}(Z_i - z)}, \quad i = 1, 2, \ldots, N \tag{1b}$$

where N is the total number of observations, (α_i, β_i) are the image space observations for the i^{th} LOS, (X_i, Y_i, Z_i) are the known object space locations of the i^{th} beacon, (x, y, z) are the unknown object space location of the sensor, f is the known focal length, and A_{jk} are the unknown coefficients of the attitude matrix, A. In general the observations can be given by α_i/f and β_i/f, so we can assume $f = 1$ without loss in generality, which is done for

the remainder of this paper. The goal of the *inverse problem* is given observations (α_i, β_i) and object space locations (X_i, Y_i, Z_i), for $i = 1, 2, \ldots, N$, determine the attitude (A) and position (x, y, z). Note that if the beacons are "infinitely away" then Eq. (1) reduces down to the standard stellar collinearity equations.

Denoting α_i and β_i by the 2×1 vector $\boldsymbol{\gamma}_i \equiv [\alpha_i \;\; \beta_i]^T$, then the measurement model follows

$$\tilde{\boldsymbol{\gamma}}_i = \boldsymbol{\gamma}_i + \mathbf{w}_i \tag{2}$$

where $\mathbf{w}_i$ is a zero-mean Gaussian noise process. A frequently used covariance for $\mathbf{w}_i$ with $f = 1$ is given by [8]

$$R_i^{\text{FOCAL}} = \frac{\sigma^2}{1 + d\,(\alpha_i^2 + \beta_i^2)} \begin{bmatrix} (1 + d\,\alpha_i^2)^2 & (d\,\alpha_i\beta_i)^2 \\ \\ (d\,\alpha_i\beta_i)^2 & (1 + d\,\beta_i^2)^2 \end{bmatrix} \tag{3}$$

where d is on the order of one and σ is assumed to be known. Note that as α_i or β_i increases then the individual components of R_i^{FOCAL} increase, which realistically shows that the errors increase as the observation moves away from the boresight. Also, as stated in Ref. [8], the covariance model is a function of the true variables α_i and β_i, which are never available in practice. However, using the measurements themselves or estimated quantities in the EKF leads to only second-order error effects.

Unit Vector Form

The observation can be reconstructed in unit vector form as

$$\mathbf{b}_i = A\mathbf{r}_i, \quad i = 1, 2, \ldots, N \tag{4}$$

where

$$\mathbf{b}_i \equiv \frac{1}{\sqrt{1+\alpha_i^2+\beta_i^2}} \begin{bmatrix} -\alpha_i \\ -\beta_i \\ 1 \end{bmatrix} \tag{5a}$$

$$\mathbf{r}_i \equiv \frac{1}{\sqrt{(X_i-x)^2+(Y_i-y)^2+(Z_i-z)^2}} \begin{bmatrix} X_i - x \\ Y_i - y \\ Z_i - z \end{bmatrix} \tag{5b}$$

When measurement noise is present, Shuster and Oh [9] have shown that nearly all the probability of the errors is concentrated on a very small area about the direction of $\mathbf{b}_i$, so the sphere containing that point can be approximated by a tangent plane, characterized by

$$\tilde{\mathbf{b}}_i = A\mathbf{r}_i + \boldsymbol{\upsilon}_i, \quad \boldsymbol{\upsilon}_i^T \mathbf{b}_i = 0 \tag{6}$$

where $\tilde{\mathbf{b}}_i$ denotes the i^{th} measurement and the sensor error $\boldsymbol{\upsilon}_i$ is approximately Gaussian, which satisfies

$$E\{\boldsymbol{\upsilon}_i\} = \mathbf{0} \tag{7a}$$

$$R_i^{\text{QUEST}} \equiv E\left\{\boldsymbol{\upsilon}_i \boldsymbol{\upsilon}_i^T\right\} = \sigma^2\left(I_{3\times 3} - \mathbf{b}_i \mathbf{b}_i^T\right) \tag{7b}$$

where $E\{\,\}$ denotes expectation and $I_{3\times 3}$ denotes a 3×3 identity matrix. Equation (7b) is known as the *QUEST measurement model*. Note that Eq. (7b) is a also function of the unknown true values in $\mathbf{b}_i$. However, the advantage of using the QUEST measurement model is that the measurement covariance in the EKF formulation can effectively be replaced by a nonsingular matrix, given by $\sigma^2 I_{3\times 3}$, which does not contain the unknown true values (see Ref. [8] for more details).

New Model

To derive a covariance for the actual unit vector measurement, the true values for α_i and β_i must be replaced with the measured ones in Eq. (5). Performing this replacement does not explicitly yield the form given by Eq. (6) because the actual model cannot separate $A\mathbf{r}_i$ from the noise. Hence, the actual noise model contains nonlinear terms coupled with non-Gaussian components. In order to derive a covariance, the new measurement model is based on a first-order Taylor series expansion of the unit vector model in Eq. (6). Note that this approach does not make the small FOV assumption, but rather it makes the assumption that the measurement noise is "small" compared to the signal, which is valid for every star tracker and for the VISNAV sensor as well. The Jacobian of Eq. (5a) is given by

$$J_i \equiv \frac{\partial \mathbf{b}_i}{\partial \boldsymbol{\gamma}_i} = \frac{1}{\sqrt{1+\alpha_i^2+\beta_i^2}} \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} - \frac{1}{1+\alpha_i^2+\beta_i^2} \mathbf{b}_i \begin{bmatrix} \alpha_i & \beta_i \end{bmatrix} \tag{8}$$

The new covariance is now given by

$$R_i^{\text{NEW}} = J_i \, R_i^{\text{FOCAL}} \, J_i^T \tag{9}$$

Clearly, R_i^{NEW} is a singular matrix, just as Eq. (7b). It will be shown that the eigenvector associated with the zero eigenvalue of R_i^{NEW} is $\mathbf{b}_i$, which is exactly the same eigenvector associated with the zero eigenvalue of R_i^{QUEST}. Since R_i^{QUEST} has two repeated eigenvalues, σ^2, then the associated eigenvectors, which are always in the plane perpendicular to $\mathbf{b}_i$, are not unique. Therefore, without loss in generality it can be assumed that R_i^{QUEST} has the same eigenvectors as R_i^{NEW}. Thus, the only differences between these two covariances are their nonzero eigenvalues. The covariance R_i^{QUEST} can also be written using J_i as

$$R_i^{\text{QUEST}} = J_i \, \mathbb{R}_i^{\text{FOCAL}} \, J_i^T \tag{10}$$

where

$$\mathbb{R}_i^{\mathrm{FOCAL}} \equiv \sigma^2(1+\alpha_i^2+\beta_i^2)\begin{bmatrix} 1+\alpha_i^2 & \alpha_i\beta_i \\ \\ \alpha_i\beta_i & 1+\beta_i^2 \end{bmatrix} \tag{11}$$

The covariances are identical, i.e. $R_i^{\mathrm{NEW}} = R_i^{\mathrm{QUEST}}$, when $R_i^{\mathrm{FOCAL}} = \mathbb{R}_i^{\mathrm{FOCAL}}$, which agrees with the corrected result shown in Ref. [12]. This shows that in order to recover R_i^{QUEST}, the errors also should not be constant over the FOV. Note the differences between Eq. (3) and Eq. (11) though. Performing a series expansion for the small FOV condition $|\alpha_i| \ll 1$ and $|\beta_i| \ll 1$ indicates that the differences between the two covariance matrices are on the order of $\alpha_i^2+\beta_i^2$ for the diagonal elements and $\alpha_i\beta_i$ for the off-diagonal elements. Since both covariances contain elements with comparable orders, Eqs. (3) and (11) only agree well with each other when they both agree equally well with $\sigma^2 I_{2\times 2}$.

We now prove that $R_i^{\mathrm{QUEST}} \geq R_i^{\mathrm{NEW}}$ using $0 \leq d \leq 1$ in R_i^{FOCAL}. The eigenvalues of $\mathbb{R}_i^{\mathrm{FOCAL}}$ are $\sigma^2(1+\alpha_i^2+\beta_i^2)$ and $\sigma^2(1+\alpha_i^2+\beta_i^2)^2$. So, $\mathbb{R}_i^{\mathrm{FOCAL}} \geq \sigma^2(1+\alpha_i^2+\beta_i^2)I_{2\times 2}$; the equality is given when $\alpha_i = \beta_i = 0$. Assuming $\alpha_i \geq \beta_i$ without loss in generality, the following equation is valid:

$$\begin{aligned} R_i^{\mathrm{FOCAL}} &\leq \frac{\sigma^2}{1+d\,(\alpha_i^2+\beta_i^2)}\begin{bmatrix} 0 & 0 \\ 0 & (1+d\,\alpha_i^2)^2-(1+d\,\beta_i^2)^2 \end{bmatrix} + R_i^{\mathrm{FOCAL}} \\ &= \frac{\sigma^2(1+d\,\alpha_i^2)^2}{1+d\,(\alpha_i^2+\beta_i^2)}\begin{bmatrix} 1 & \xi \\ \xi & 1 \end{bmatrix} \end{aligned} \tag{12}$$

where $\xi \equiv d^2\alpha_i^2\beta_i^2/(1+d\,\alpha_i^2)^2$. The eigenvalues of the 2×2 matrix involving ξ in Eq. (12)

are given by $1 \pm \xi$. This leads to

$$\begin{aligned} R_i^{\text{FOCAL}} &\leq \left[1 + \frac{d^2\alpha_i^2\beta_i^2}{(1+d\,\alpha_i^2)^2}\right]\left[\frac{\sigma^2(1+d\,\alpha_i^2)^2}{1+d\,(\alpha_i^2+\beta_i^2)}\right] I_{2\times 2} \\ &= \frac{\sigma^2}{1+d\,(\alpha_i^2+\beta_i^2)}\left[(1+d\,\alpha_i^2)^2 + d^2\alpha_i^2\beta_i^2\right] I_{2\times 2} \end{aligned} \tag{13}$$

Therefore, $R_i^{\text{FOCAL}} \leq \sigma^2(1+d\,\alpha_i^2+d\,\beta_i^2)I_{2\times 2}$, since $(1+d\,\alpha_i^2+d\,\beta_i^2)^2 \geq (1+d\,\alpha_i^2)^2 + d^2\alpha_i^2\beta_i^2$. Hence,

$$\mathbb{R}_i^{\text{FOCAL}} \geq \sigma^2(1+\alpha_i^2+\beta_i^2)I_{2\times 2} \geq \sigma^2(1+d\,\alpha_i^2+d\,\beta_i^2)I_{2\times 2} \geq R_i^{\text{FOCAL}} \tag{14}$$

which proves that $\mathbb{R}_i^{\text{FOCAL}} \geq R_i^{\text{FOCAL}}$. Now, using the definition of a positive semi-definite matrix, we have

$$(J_i^T\mathbf{x})^T(\mathbb{R}_i^{\text{FOCAL}} - R_i^{\text{FOCAL}})(J_i^T\mathbf{x}) \geq 0 \tag{15}$$

for any $\mathbf{x}$. This leads directly to $J_i(\mathbb{R}_i^{\text{FOCAL}} - R_i^{\text{FOCAL}})J_i^T \geq 0$. Using the definitions of R_i^{NEW} and R_i^{QUEST} from Eqs. (9) and (10), respectively, gives $R_i^{\text{QUEST}} \geq R_i^{\text{NEW}}$.

In the QUEST measurement model, all the LOS vector observations have the same noise level, characterized by σ^2. This requires the associated measurement noise on the focal plane to increase at a specific rate from the center to the edge of the FOV and be distributed according to a specific pattern, given by Eq. (11). When the assumed noise distribution on the focal plane does not match the true measurement noise characteristics, which is the case in this study, performance degradation is observed.

Extended Kalman Filter Implementation

In this section the new covariance is implemented in the EKF. Two approaches to overcome the singularity issue in the EKF are presented. The first is based on an eigenvalue/eigenvector decomposition of R_i^{NEW} and the second is based on a rank-one update of the covariance matrix. Only attitude estimation is considered here. The VISNAV sensor is capable of determining position as well, but this is an easy extension of the attitude estimation problem by simply augmenting the state vector. The EKF equations follow the

multiplicative quaternion approach of Ref. [13], which includes the three-component attitude error vector and gyro-bias errors. Unfortunately, straightforward implementation of the EKF with the new model is not possible. The Kalman gain for a single observation, written in terms of using the true values for now, is given by [14]

$$K = P^- H_i^T \left[H_i P^- H_i^T + R_i^{\text{NEW}} \right]^{-1} \tag{16}$$

where P^- is the propagated 6×6 covariance matrix and H_i is given by [13]

$$H_i = \begin{bmatrix} [A\,\mathbf{r}_i \times] & 0_{3\times 3} \end{bmatrix} \tag{17}$$

where $[A\,\mathbf{r}_i \times]$ is the standard cross product matrix (see Ref. [13]) and $0_{3\times 3}$ is a 3×3 matrix of zeros. We now investigate the properties of the matrix $Z_i \equiv H_i P^- H_i^T + R_i^{\text{NEW}}$, which is known as the innovation matrix. In the EKF formulation the estimated values will be used to form the matrices $H_i P^- H_i^T$ and R_i^{NEW}. Hence, for the analysis of the matrix Z_i we can set $\mathbf{b}_i = A\,\mathbf{r}_i$ without loss in generality for now. From the definition of the cross product matrix it is easy to see that $H_i^T \mathbf{b}_i = \mathbf{0}$. Performing the multiplication $J_i^T \mathbf{b}_i$ also gives $J_i^T \mathbf{b}_i = \mathbf{0}$. Hence, $Z_i \mathbf{b}_i$ is always zero, which means Z_i is always singular. Furthermore, since the rank of R_i^{FOCAL} is two, which means the rank of R_i^{NEW} is also two, then $\mathbf{b}_i$ is the eigenvector associated with the zero eigenvalue of R_i^{NEW}. The singularity problem always exists no matter how many measurements are used in the EKF. Therefore, the standard EKF cannot be executed.

The specific form for the EKF follows Murrell's version [15], which processes one unit vector observation at a time through a sequential approach. This approach reduces taking an inverse of a $3N \times 3N$ matrix for the Kalman gain to taking an inverse of a 3×3 matrix N times, which significantly reduces the computational load. A flow chart of Murrell's version for the update is shown in Fig. 2, where $\hat{\boldsymbol{\beta}}^-$ and $\hat{\boldsymbol{\beta}}^+$ are the propagated and updated gyro-bias estimates, respectively, P^+ is the updated 6×6 covariance matrix, $\Delta\hat{\hat{\mathbf{x}}}^-$ and $\Delta\hat{\hat{\mathbf{x}}}^+$ are the propagated and updated state corrections for the attitude and gyro bias, respectively,

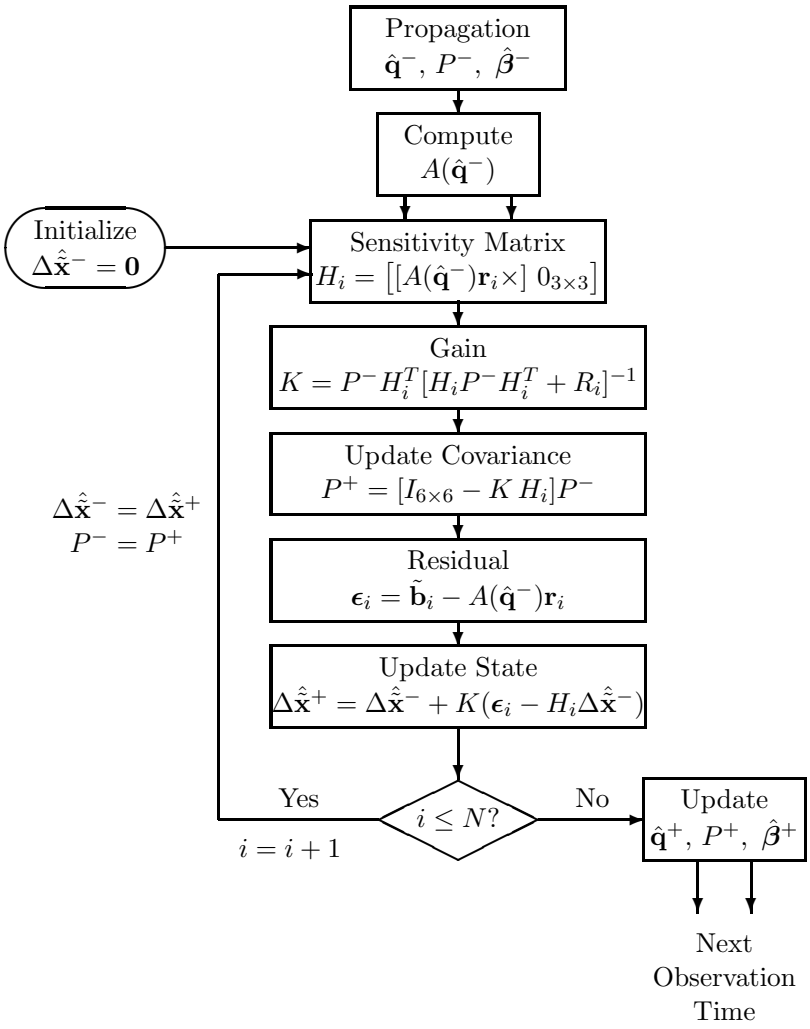


Figure 2. Computationally Efficient Attitude Estimation Algorithm of Murrell

$A(\hat{\mathbf{q}}^-)$ is the attitude matrix parameterized using the propagated quaternion estimate, $\hat{\mathbf{q}}^-$, and R_i is the measurement covariance. The first step involves propagating the quaternion, gyro bias and error-covariance to the current observation time. Then, the attitude matrix is computed. The propagated state vector is now initialized to zero. Next, the error-covariance and state quantities are updated using a single vector observation. This procedure is continued, replacing the propagated error-covariance and state vector with the updated values, until all vector observations are processed. Finally, the updated values are used to propagate the error-covariance and state quantities to the next observation time.

Decomposition Approach

The eigenvalue/eigenvector decomposition of R_i^{NEW} can be written in the form given by

$$\begin{aligned} R_i^{\text{NEW}} &= T_i \, E_i \, T_i^T \\ &\equiv \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 & \mathbf{t}_3 \end{bmatrix}_i \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & 0 \end{bmatrix}_i \begin{bmatrix} \mathbf{t}_1 & \mathbf{t}_2 & \mathbf{t}_3 \end{bmatrix}_i^T \end{aligned} \tag{18}$$

where $\mathbf{t}_1$, $\mathbf{t}_2$ and $\mathbf{t}_3$ are the eigenvectors, and λ_1 and λ_2 are the nonzero eigenvalues. A linear transformation of the measurement residual in Fig. 2 is now performed, giving a new residual $\mathbf{d}_i$:

$$\mathbf{d}_i \equiv \begin{bmatrix} \mathbf{e}_i \\ g_i \end{bmatrix} = T_i^T \boldsymbol{\epsilon}_i \tag{19}$$

where $\mathbf{e}_i$ is a 2×1 vector made up of the first two components of $\mathbf{d}_i$ and g_i is the third component of $\mathbf{d}_i$. Since $\mathbf{t}_3 = A(\hat{\mathbf{q}}^-)\mathbf{r}_i$, then the transformed sensitivity matrix has the form

$$T_i^T H_i = \begin{bmatrix} \mathcal{T}_i & 0_{2\times 3} \\ \mathbf{0}^T & \mathbf{0}^T \end{bmatrix} \tag{20}$$

where $\mathcal{T}_i$ is a 2×3 matrix, $\mathbf{0}$ is a 3×1 vector of zeros, and $0_{2\times 3}$ is a 2×3 matrix of zeros. Hence, only the vector $\mathbf{e}_i$ is needed to perform the updates. These equations are now given by

$$\mathcal{K} = P^- \mathcal{H}_i^T [\mathcal{H}_i P^- \mathcal{H}_i^T + \Lambda_i]^{-1} \tag{21a}$$

$$P^+ = [I_{6\times 6} - \mathcal{K}\,\mathcal{H}_i] P^- \tag{21b}$$

$$\Delta\hat{\hat{\mathbf{x}}}^+ = \Delta\hat{\hat{\mathbf{x}}}^- + \mathcal{K}\,(\mathbf{e}_i - \mathcal{H}_i \Delta\hat{\hat{\mathbf{x}}}^-) \tag{21c}$$

where $\mathcal{H}_i \equiv \begin{bmatrix} \mathcal{T}_i & 0_{2\times 3} \end{bmatrix}$ and Λ_i is a diagonal matrix made up of the nonzero eigenvalues of R_i^{NEW}:

$$\Lambda_i \equiv \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}_i \tag{22}$$

Note that the matrix inverse in the Kalman gain reduces from a 3×3 matrix to a 2×2 matrix. Also, since Λ_i is a diagonal matrix, then a sequential process can be used to process each component of $\mathbf{e}_i$ one at a time [14]. This reduces taking a 2×2 matrix inverse down to taking an inverse of a scalar twice, which further reduces the computational load. The quaternion and gyro-bias updates are given by

$$\Delta\hat{\tilde{\mathbf{x}}}^+ \equiv \begin{bmatrix} \boldsymbol{\delta\hat{\alpha}}^{+T} & \Delta\hat{\boldsymbol{\beta}}^{+T} \end{bmatrix}^T \tag{23a}$$

$$\hat{\mathbf{q}}^+ = \hat{\mathbf{q}}^- + \frac{1}{2}\Xi(\hat{\mathbf{q}}^-)\boldsymbol{\delta\hat{\alpha}}^+, \quad \text{re-normalize quaternion} \tag{23b}$$

$$\hat{\boldsymbol{\beta}}^+ = \hat{\boldsymbol{\beta}}^- + \Delta\hat{\boldsymbol{\beta}}^+ \tag{23c}$$

where $\boldsymbol{\delta\hat{\alpha}}$ is the angle correction, $\Delta\hat{\boldsymbol{\beta}}$ is the bias correction and $\Xi(\hat{\mathbf{q}}^-)$ is a 4×3 matrix given by

$$\Xi(\hat{\mathbf{q}}^-) \equiv \begin{bmatrix} \hat{q}_4^- I_{3\times 3} + [\hat{\boldsymbol{\rho}}^-\times] \\ \\ -\hat{\boldsymbol{\rho}}^{-T} \end{bmatrix} \tag{24}$$

with $\hat{\mathbf{q}}^- \equiv [\hat{\boldsymbol{\rho}}^{-T} \;\; \hat{q}_4^-]^T$. The propagation equations remain unchanged, which are not shown here for brevity (see Ref. [13] for more details). In order to further enhance the numerical properties of the algorithm, a *U-D* factorization [14] of the covariance update and propagation is employed. The main advantage of this approach is that the factorization is accomplished without taking square roots, unlike a square-root filter, and the formulation does not add much computational effort over the standard EKF. The new model uses the specific covariance shown in Eq. (3), but it is important to state that any covariance matrix can be used in this formulation. The eigenvalue/eigenvector decomposition can easily be performed

using numerical techniques.

Rank-One Update Approach

An alternative approach to overcome the problem in the EKF due to the singularity of the new measurement covariance matrix is shown. The main idea of this approach is to add an extra term, $c_i \mathbf{b}_i \mathbf{b}_i^T$ $(c_i > 0)$, to the singular measurement covariance matrix to ensure that the modified measurement covariance matrix is nonsingular. After this modification the standard EKF equations are used, as shown in Fig. 2. The EKF with the modified measurement covariance matrix will yield the identical measurement update result with the one employing size-reduced residuals and measurement covariances. This approach is a straightforward extension of Shuster's approach in Ref. [8] to overcome the problem with the singular QUEST measurement covariance matrix (and singular innovation matrix). In Ref. [8] the original QUEST measurement covariance matrix in the standard EKF, $R_i^{\text{QUEST}} = \sigma^2 \left(I_{3\times 3} - \mathbf{b}_i \mathbf{b}_i^T \right)$, is replaced by $\mathcal{R}_i^{\text{QUEST}} = \sigma^2 I_{3\times 3}$, with the modification to the QUEST measurement covariance given by

$$\mathcal{R}_i^{\text{QUEST}} = R_i^{\text{QUEST}} + \sigma^2 \mathbf{b}_i \mathbf{b}_i^T \tag{25}$$

Again note that $R_i^{\text{QUEST}} \mathbf{b}_i = \mathbf{0}$, but $\mathcal{R}_i^{\text{QUEST}}$ is always invertible.

For the new measurement model, we propose to modify the singular measurement covariance matrix, R_i^{NEW}, in a similar manner:

$$\mathcal{R}_i^{\text{NEW}} = R_i^{\text{NEW}} + c_i \mathbf{b}_i \mathbf{b}_i^T \tag{26}$$

with $c_i > 0$. For any R_i^{NEW} derived from R_i^{FOCAL}, it is also true that $R_i^{\text{NEW}} \mathbf{b}_i = \mathbf{0}$, which is a direct result of $J_i^T \mathbf{b}_i = \mathbf{0}$. The inverse of $\mathcal{R}_i^{\text{NEW}}$ is then given by

$$\left(\mathcal{R}_i^{\text{NEW}} \right)^{-1} = \left(R_i^{\text{NEW}} \right)^{\dagger} + \frac{1}{c_i} \mathbf{b}_i \mathbf{b}_i^T \tag{27}$$

where $\left(R_i^{\text{NEW}}\right)^\dagger$ may be interpreted as the pseudo-inverse of R_i^{NEW} or simply a convenient notation. Pre-multiplying both sides of Eq. (27) by H_i^T leads to

$$H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1} = H_i^T \left(R_i^{\text{NEW}}\right)^\dagger + H_i^T \cdot \frac{1}{c_i} \mathbf{b}_i \mathbf{b}_i^T = H_i^T \left(R_i^{\text{NEW}}\right)^\dagger \tag{28}$$

Note that because $H_i^T \mathbf{b}_i = \mathbf{0}$, the identity holds regardless of the value of $c_i > 0$. The limit of $H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1}$ as c_i approaches zero is $H_i^T \left(R_i^{\text{NEW}}\right)^\dagger$ as well. Similarly, for all c_i we have

$$H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1} H_i = H_i^T \left(R_i^{\text{NEW}}\right)^\dagger H_i \tag{29}$$

The inverse of $\mathcal{R}_i^{\text{NEW}}$ always exists because the eigenvalues obey [16]

$$\lambda_j(\mathcal{R}_i^{\text{NEW}}) = \lambda_j(R_i^{\text{NEW}}) + m_j \, c_i, \quad j = 1, \, 2, \, 3 \tag{30}$$

with $m_1 + m_2 + m_3 = 1$. In fact the eigenvalues and eigenvectors of $\mathcal{R}_i^{\text{NEW}}$ are identical to the ones of R_i^{NEW}, except that the zero eigenvalue of R_i^{NEW} is replaced with c_i. Since c_i is assumed to never be zero, then the inverse of $\mathcal{R}_i^{\text{NEW}}$ always exists.

The measurement update of the covariance matrix and the computation of the Kalman gain matrix with the original and modified measurement covariance matrices may be re-written as

$$\left(P^+\right)^{-1} = \left(P^-\right)^{-1} + H_i^T \left(R_i^{\text{NEW}}\right)^\dagger H_i \tag{31}$$

$$K = P^+ H_i^T \left(R_i^{\text{NEW}}\right)^\dagger \tag{32}$$

and

$$\left(P^+\right)^{-1} = \left(P^-\right)^{-1} + H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1} H_i \tag{33}$$

$$K = P^+ H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1} \tag{34}$$

From the two sets of equations and Eqs. (28) and (29), we can see that any $\mathcal{R}_i^{\text{NEW}}$ will lead to the same P^+ and K and thus to the same measurement update result as R_i^{NEW} does. Therefore, the use of $\mathcal{R}_i^{\text{NEW}}$ overcomes the singularity problem, but does not alter the result of the EKF. Finally, we note the following:

1. As far as the EKF or first-order approximation is concerned, there is no limitation on the parameter $c_i > 0$ in order to guarantee the invertibility of the measurement covariance matrix and the innovation matrix. Physically, c_i is small because the error along the true boresight of the effective unit vector measurement converted from the focal-plane measurement is much smaller than the errors along the other two directions (the first-order approximation of c_i is zero). For numerical purposes, however, c_i may be chosen to be

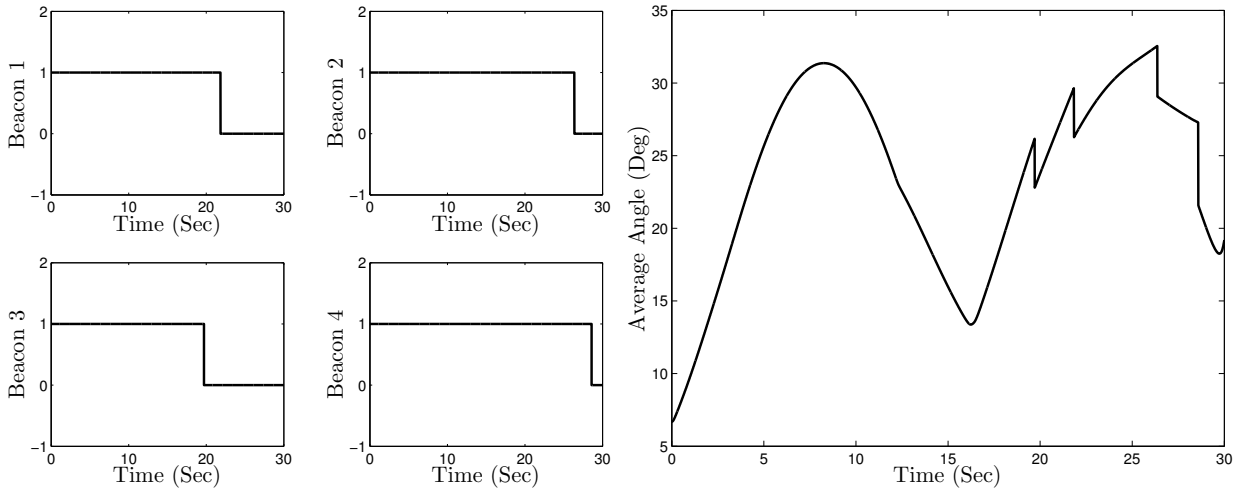
$$c_i = \frac{1}{2}\text{trace}(R_i^{\text{NEW}}) \tag{35}$$

   where trace denotes the trace of a matrix. That is, c_i is the average of the nonzero eigenvalues of R_i^{NEW}. Note that using Eq. (35) on R_i^{QUEST} instead of R_i^{NEW} gives $c_i = \sigma^2$, which yields Eq. (25).

2. The parameter c_i will change the property of the measurement covariance matrix, but not that of $H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1} H_i$. If $H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1} H_i$ is ill-conditioned, the approach does not help.

3. Since $H_i^T \left(\mathcal{R}_i^{\text{NEW}}\right)^{-1} H_i$ remains unchanged with the modified measurement covariance matrix, there is no information loss or gain in the EKF.

4. The approach overcomes the singularity problem based on the very special structure of H_i and R_i^{NEW}: $H_i^T \mathbf{b}_i = \mathbf{0}$ and $R_i^{\text{NEW}} \mathbf{b}_i = \mathbf{0}$. However, if the system does not have such a special structure, the innovation matrix in the EKF will not be singular and the problem no longer exists.



(a) **Availability of First Four Beacons** (b) **Average Angle from Boresight**

Figure 3. Availability of Beacons and Average Angle from Boresight

Simulation Results

In this section the simulation results using synthetic VISNAV measurements are presented. The VISNAV sensor is only used here to determine attitude for the simulation results. As previously stated, position can also be determined by augmenting the EKF state vector. Eight beacons are used to create the synthetic measurements. The beacon locations in meters are given by

$$X_1 = 5, \quad Y_1 = 5, \quad Z_1 = 0 \tag{36a}$$

$$X_2 = 5, \quad Y_2 = -5, \quad Z_2 = 0 \tag{36b}$$

$$X_3 = -5, \quad Y_3 = 5, \quad Z_3 = 0 \tag{36c}$$

$$X_4 = -5, \quad Y_4 = -5, \quad Z_4 = 0 \tag{36d}$$

$$X_5 = 0.2, \quad Y_5 = 0.2, \quad Z_5 = 0 \tag{36e}$$

$$X_6 = 0.2, \quad Y_6 = -0.2, \quad Z_6 = 0 \tag{36f}$$

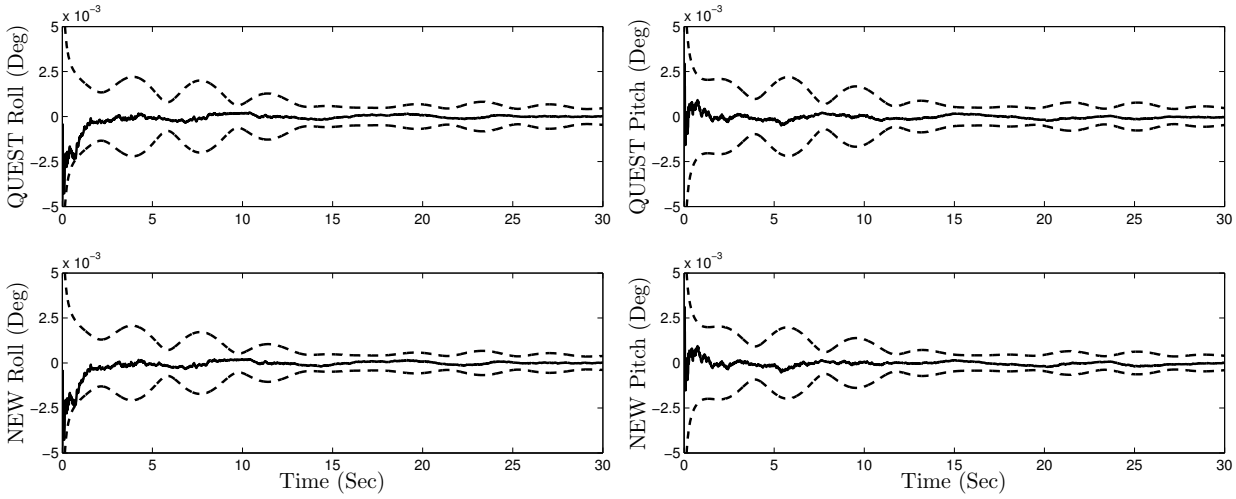
$$X_7 = -0.2, \quad Y_7 = 0.2, \quad Z_7 = 0 \tag{36g}$$

$$X_8 = -0.2, \quad Y_8 = -0.2, \quad Z_8 = 0 \tag{36h}$$

The first four beacons are spread out further than the last four beacons. A 30 second simulation is used with a sampling rate of 100 Hz for both the VISNAV and gyro measurements. A docking-type motion with a rotation is used in the simulation. The linear motions for both the x and y axes are zero for the entire simulation run. The vehicle performs a linear motion maneuver along the z axis starting 30 meters away at the initial time and ending at 0 at the final time. The initial attitude is given by the identity matrix and the angular velocity is given by $\boldsymbol{\omega} = [0.1\sin(t)\ \ 0.1\cos(t)\ \ (3 \times 360/30) \times (\pi/180)]^T$ rad/sec. The VISNAV sensor FOV is assumed to be 100 degrees. Therefore, any LOS measurement that is greater than 50 degrees from the boresight is not available. Since the last four beacons are "close in," they are available throughout the entire simulation. The availability of the first four beacons is shown in Fig. 3(a), where 1 indicates that the beacon is available and 0 indicates that it is not available. The first four beacons are all available until about 20 seconds into the simulation run. Synthetic measurements are generated using the covariance in Eq. (3) with $d = 1$ and $\sigma = (100/5000) \times (\pi/180)$ radians. These measurements are then converted into unit-vector form to be used in the EKF. The gyro noise parameters are given by $\sigma_u = \sqrt{10} \times 10^{-10}$ rad/sec$^{3/2}$ and $\sigma_v = \sqrt{10} \times 10^{-7}$ rad/sec$^{1/2}$. See Ref. [17] for a detailed explanation of the gyro model. The initial biases for each axis of the gyro are given by 0.1 deg/hr.

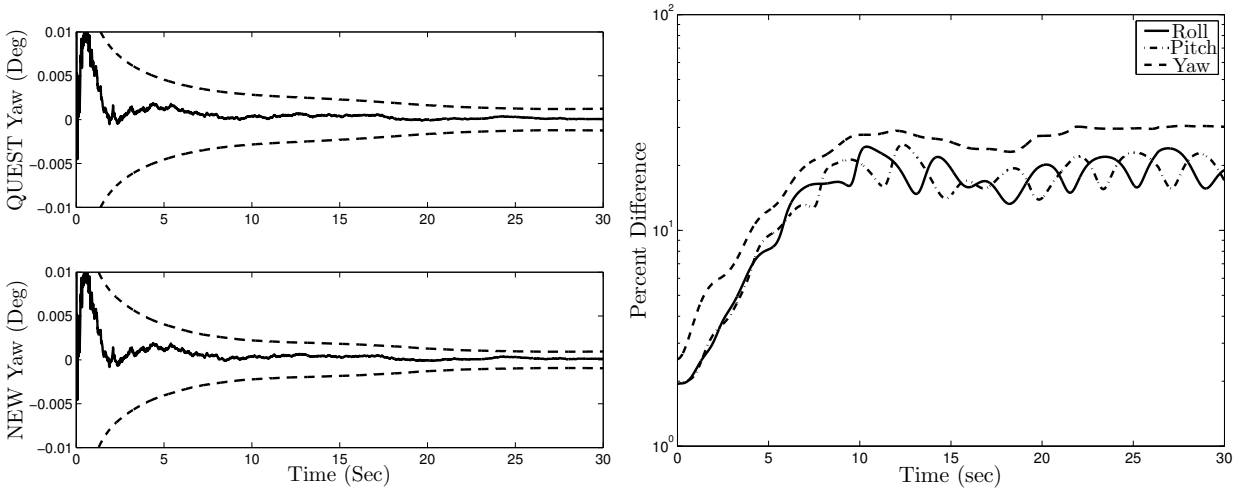
Two EKFs are executed. Both EKFs use the same exact synthetic LOS and gyro measurements, and both use Murrell's version shown in Fig. 2. The first EKF uses the QUEST model approach for the measurement covariance with $R_i = \sigma^2 I_{3\times3}$. The second EKF uses the new measurement covariance with $R_i = \mathcal{R}_i^{\text{NEW}}$. For the new model, the simulation results using the decomposition approach with the U-D factorization are identical to the rank-one update approach results. A plot of the average angle from the boresight for all the available beacons is shown in Fig. 3(b). The discontinuities are due to a loss of a beacon at that specific time.

Plots of the roll, pitch and yaw errors for a typical simulation run, with their respective 3σ bounds computed from the error covariance, are shown in Figs. 4(a)-4(c), respectively. The top plot of each figure corresponds to using the QUEST model in the EKF and the bottom



(a) **Roll Errors and 3σ Bounds**

(b) **Pitch Errors and 3σ Bounds**



(c) **Yaw Errors and 3σ Bounds**

(d) **Percent Difference Between QUEST and NEW Model**

Figure 4. Extended Kalman Filter Results

plot corresponds to using the new model approach. At first glance it seems as though both approaches yield nearly identical results, but closer inspection shows that this is not true. To investigate the performance aspects of both filters, the percent difference between the 3σ attitude bounds is computed. The results are shown in Fig. 4(d). Using the QUEST model approach gives respective 3σ bounds that are always larger than the new model approach. Figure 5 shows the norm of the percent differences taken from Fig. 4(d) versus the average

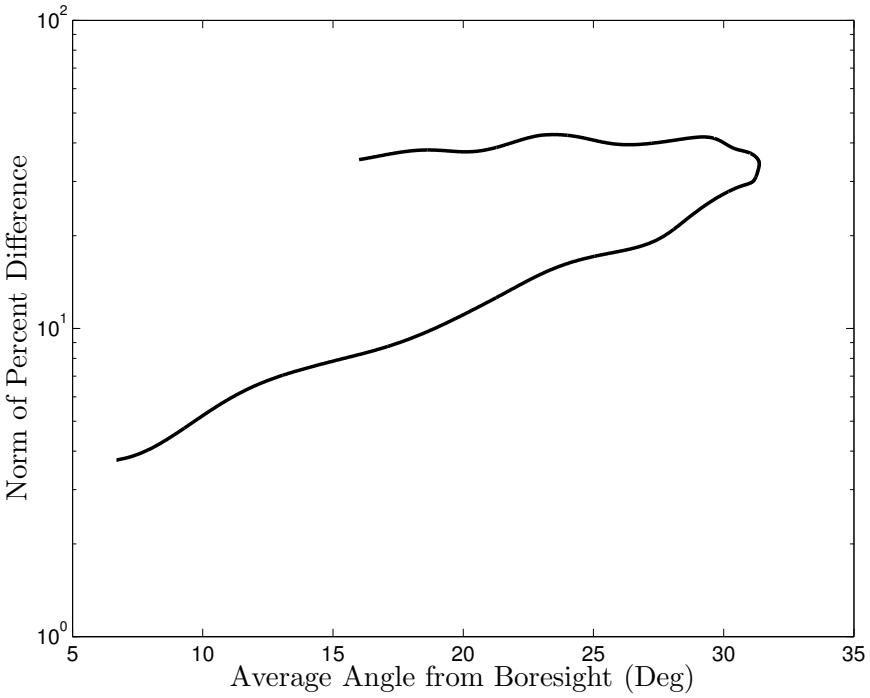
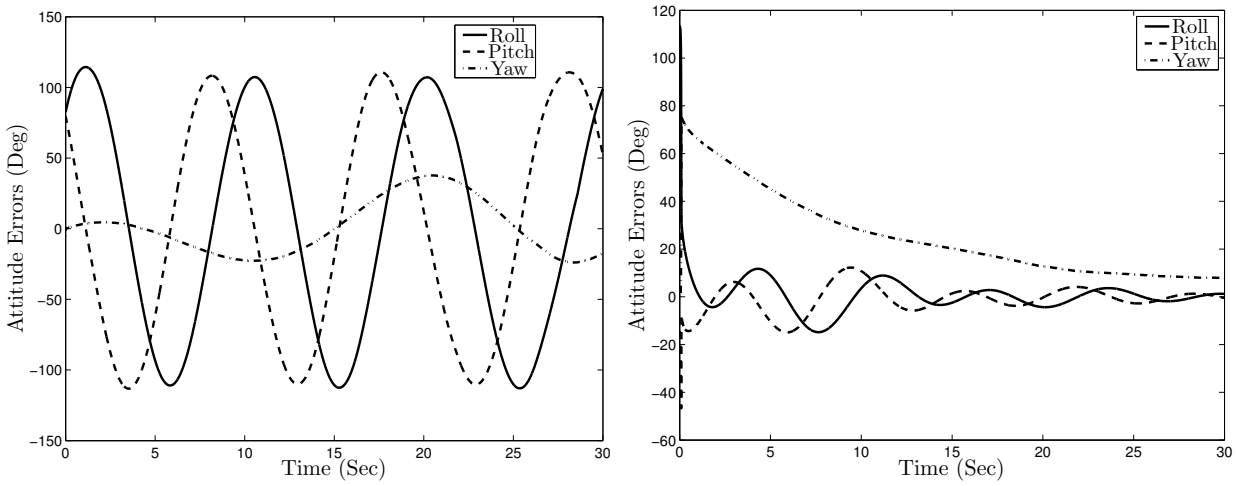


Figure 5. Norm Percent Difference Versus Average Angle

angle from boresight for the first 15 seconds of the simulation run. At the beginning of the simulation run, when the average boresight angle is small, both models produce nearly identical results. The errors then begin to increase when the average boresight angle becomes larger, but the errors still remain large even as the average boresight angle decreases. This may be due to the "memory" nature of the EKF in its error covariance. From Figs. 4(a)-4(c), we see that the transients in the EKF error covariance have significantly reduced after about 8 seconds. Even though the average boresight angle begins to decrease after this time, the gain in the EKF does not vary much which leads to large differences between the two models, as seen by the 30 percent difference in the two model responses. Having accurate 3σ bounds is a paramount issue, because they are often used to develop error budgets for the overall attitude knowledge. The new model provides more realistic and lower bounds for large FOV sensors than the QUEST model. This is especially useful for missions with very tight attitude requirements.

An augmented filter that estimates attitude as well as position has also been developed. The state vector is augmented by six states, which model the three accelerations as zero-



(a) **Decomposition Approach** (b) **Rank-One Update Approach**

Figure 6. Extended Kalman Filter Results for Large Roll Initial Condition Error

mean Gaussian white-noise processes. The position improvements using the new covariance are nearly the same as the attitude improvements, and thus are not shown here for brevity. Overall, the simulation results clearly show that the new covariance model can provide better results over the QUEST model for large FOV sensors.

A comparison between the covariance decomposition approach and the rank-one update approach has also been done. The decomposition approach is found to be sensitive to numerical ill-conditioning effects. In fact large deviations from the rank-one approach results, shown in Fig. 4, are present when the U-D factorization is not used, i.e. running the standard EKF equations. The convergence properties of each approach has also been studied. A roll error of 180 degrees is introduced at the initial condition with the variance for that part of the initial covariance of the EKF set to $[(220/3) \times (\pi/180)]^2$ rad^2 for both approaches. For this error both approaches do not converge well. But, for this test the decomposition approach uses the U-D factorization, while the rank-one update approach uses the standard EKF equations. A U-D factorization for the rank-one update approach has also been designed. The results of the U-D factorization EKF version for both the decomposition and rank-one update approaches are shown in Fig. 6. Clearly, the rank-one update approach pro-

duces better convergence behaviors than the decomposition approach. Several other initial condition errors have also been tested, as well as a number of Monte Carlo-type simulations. When using the U-D factorization EKF version for both approaches, the rank-one update approach always shows better convergence behaviors than the decomposition approach. Further improvements may be obtained by using an Unscented filter [18].

Conclusions

In this paper a new measurement covariance model was derived that can be used for large field-of-view sensors. The new model was derived using a first-order Taylor series expansion of the line-of-sight measurement model. Due to the singular nature of the new covariance, straightforward implementation of the extended Kalman filter was not possible. This was overcome by using one of two approaches. The first approach involves decomposing the covariance matrix into its eigenvalues and eigenvectors, and then performing a transformation of state so that a reduced nonsingular matrix is used in the extended Kalman filter. The second approach adds a simple term, i.e. a rank-one update, to the covariance matrix that does not alter the overall Kalman gain. Simulation results indicate that the new model can yield better estimation results over using the small field-of-view QUEST measurement model in the filter design. Also, a comparison of the two approaches showed that the rank-one update approach can provide better convergence properties than the decomposition approach for large initial condition errors. Also, the rank-one update approach allows a designer to easily implement the extended Kalman filter, even if the QUEST measurement model is used for the measurement covariance. For these reasons, the rank-one update approach is preferred over the decomposition approach.

Acknowledgement

The authors wish to thank Dr. John L. Junkins from Texas A&M University for numerous discussions on the VISNAV sensor. The second author was supported by NASA Goddard

Space Flight Center Grant NAG5-12179, under the supervision of Mr. Richard R. Harman. This author greatly appreciate the support.

References

[1] WERTZ, J. R., *Mission Geometry; Orbit and Constellation Design and Management*, chap. 3, Microcosm Press, El Segundo, CA and Kluwer Academic Publishers, Dordrecht, Netherlands, 2001, pp. 152–167.

[2] PSIAKI, M. L., MARTEL, F., and PAL, P. K., "Three-Axis Attitude Determination via Kalman Filtering of Magnetometer Data," *Journal of Guidance, Control, and Dynamics*, Vol. 13, No. 3, May-June 1990, pp. 506–514.

[3] CHALLA, M., NATANSON, G., and WHEELER, C., "Simultaneous Determination of Spacecraft Attitude and Rates Using Only a Magnetometer," *AIAA/AAS Astrodynamics Conference*, San Diego, CA, July 1996, AIAA-1996-3630.

[4] JU, G. and JUNKINS, J. L., "Overview of Star Tracker Technology and its Trends in Research and Development," *Advances in the Astronautical Sciences, The John L. Junkins Astrodynamics Symposium*, Vol. 115, 2003, pp. 461–478, AAS-03-285.

[5] SHUSTER, M. D., "Maximum Likelihood Estimation of Spacecraft Attitude," *The Journal of the Astronautical Sciences*, Vol. 37, No. 1, Jan.-March 1989, pp. 79–88.

[6] LIGHT, D. L., "Satellite Photogrammetry," *Manual of Photogrammetry*, edited by C. C. Slama, chap. 17, American Society of Photogrammetry, Falls Church, VA, 4th ed., 1980.

[7] GRIFFITH, D. T., SINGLA, P., and JUNKINS, J. L., "Autonomous On-orbit Calibration Approaches for Star Tracker Cameras," *AAS/AIAA Space Flight Mechanics Meeting*, San Antonio, TX, Jan. 2002, AAS-02-102.

[8] SHUSTER, M. D., "Kalman Filtering of Spacecraft Attitude and the QUEST Model," *The Journal of the Astronautical Sciences*, Vol. 38, No. 3, July-Sept. 1990, pp. 377–393.

[9] SHUSTER, M. D. and OH, S. D., "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, Jan.-Feb. 1981, pp. 70–77.

[10] CHENG, Y. and CRASSIDIS, J. L., "Particle Filtering for Sequential Spacecraft Attitude Estimation," *AIAA Guidance, Navigation, and Control Conference*, Providence, RI, Aug. 2004, AIAA-04-5337.

[11] JUNKINS, J. L., HUGHES, D. C., WAZNI, K. P., and PARIYAPONG, V., "Vision-Based Navigation for Rendezvous, Docking and Proximity Operations," *22nd Annual AAS Guidance and Control Conference*, Breckenridge, CO, Feb. 1999, AAS-99-021.

[12] SHUSTER, M. D., "Erratum: Kalman Filtering of Spacecraft Attitude and the QUEST Model," *The Journal of the Astronautical Sciences*, Vol. 51, No. 3, July-Sept. 2003, pp. 359.

[13] LEFFERTS, E. J., MARKLEY, F. L., and SHUSTER, M. D., "Kalman Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 5, No. 5, Sept.-Oct. 1982, pp. 417–429.

[14] CRASSIDIS, J. L. and JUNKINS, J. L., *Optimal Estimation of Dynamic Systems*, chap. 5, Chapman & Hall/CRC, Boca Raton, FL, 2004.

[15] MURRELL, J. W., "Precision Attitude Determination for Multimission Spacecraft," *Proceedings of the AIAA Guidance, Navigation, and Control Conference*, Palo Alto, CA, Aug. 1978, pp. 70–87.

[16] WILKINSON, J. H., *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, England, 1965, pp. 94–97.

[17] FARRENKOPF, R. L., "Analytic Steady-State Accuracy Solutions for Two Common Spacecraft Attitude Estimators," *Journal of Guidance and Control*, Vol. 1, No. 4, July-Aug. 1978, pp. 282–284.

[18] CRASSIDIS, J. L. and MARKLEY, F. L., "Unscented Filtering for Spacecraft Attitude Estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 26, No. 4, July-Aug. 2003, pp. 536–542.