AAS/GSFC International Symposium on Space Flight Mechanics, Greenbelt, Maryland, May 11–15, 1998
Advances in the Astronautical Sciences, Vol. 100, 1998, pp. 407–419

AAS 98-333

AN EFFICIENT ALGORITHM FOR SPACECRAFT ATTITUDE DETERMINATION WITH OPTICAL SENSORS

Malcolm D. Shuster*

An earlier algorithm for multiple sensors is extended to provide three-axis attitude from multiple line-of-sight observations with a single optical sensor, typically a star camera. The algorithm, called SCAD, is simpler computationally than either the QUEST or FOAM algorithms and, although suboptimal, suffers only imperceptible loss of accuracy for typical star cameras with limited fields of view. A complete analysis of variance is presented.

INTRODUCTION

A central problem in Spacecraft Attitude Determination has been that of determining the three-axis attitude which minimizes the cost function

$$J(A) = \frac{1}{2} \sum_{k=1}^{N} a_k \, \left| \hat{\mathbf{W}}_k - A\hat{\mathbf{V}}_k \right|^2 , \tag{1}$$

where A is the direction-cosine matrix[1], $\hat{\mathbf{W}}_k$, $k = 1, \ldots , N$, are directions (lines of sight, observation vectors) observed in the spacecraft body frame, $\hat{\mathbf{V}}_k$, $k = 1, \ldots , N$, are the corresponding directions known in an inertial frame (the reference vectors) and a_k, $k = 1, \ldots , N$, are a set of positive weights. A caret in this work will be used to denote a unit vector. This cost function was first proposed by G. Wahba[2] in 1965 and has been the starting point of many algorithms, of which the most popular has been the QUEST algorithm[3], although other attractive algorithms exist[4--8].

Of particular importance is the fact that the Wahba cost function can be derived from maximum-likelihood estimation[9] provided one assumes the following measurement model[10], which has been called the QUEST model, because it was used in an early accuracy study of the QUEST algorithm[3],

$$\hat{\mathbf{W}}_k = A\,\hat{\mathbf{V}}_k + \Delta\hat{\mathbf{W}}_k , \tag{2}$$

with the measurement error $\Delta\hat{\mathbf{W}}_k$ having first and second moments[†]

$$E\{\Delta\hat{\mathbf{W}}_k\} = \mathbf{0} , \tag{3}$$

$$E\{\Delta\hat{\mathbf{W}}_k\,\Delta\hat{\mathbf{W}}_k^T\} = \sigma_k^2\,[I - (A\,\hat{\mathbf{V}}_k)\,(A\,\hat{\mathbf{V}}_k)^T] , \tag{4}$$

This version has been retypeset by the author with slightly different fonts and may differ in line and page breaks from the original. Minor typographical and punctuation errors have been corrected. Some factual errors have not been corrected.

*Professor, Department of Aerospace Engineering, Mechanics and Engineering Science, University of Florida, Gainesville, FL 32611-6250. email: m.shuster@ieee.org.

†In fact, because of the unity constraint on the norm of $\hat{\mathbf{W}}_k$, the mean of $\Delta\hat{\mathbf{W}}_k$ will have a small non-vanishing part[10] equal to $-\sigma_k^2\hat{\mathbf{W}}_k$. This may be safely neglected in our discussion.

where $E\{\cdot\}$ denotes the expectation, I is the 3×3 identity matrix, and one chooses the weights to be

$$a_k = \frac{\sigma_{\text{tot}}^2}{\sigma_k^2}\,, \tag{5}$$

with

$$\frac{1}{\sigma_{\text{tot}}^2} \equiv \sum_{k=1}^{N} \frac{1}{\sigma_k^2}\,. \tag{6}$$

The common constant in the numerators of Eq. (5) is arbitrary, of course, but the choice of Eq. (6) makes

$$\sum_{k=1}^{N} a_k = 1\,. \tag{7}$$

One defines the attitude covariance matrix $P_{\theta\theta}$ (Refs. 3, 10) as the covariance of the attitude error vector, which is the rotation vector[1] of the small rotation carrying the true attitude into the estimated attitude. Assuming the QUEST model for the measurements, this leads to the following expression for the attitude covariance matrix

$$P_{\theta\theta}^{-1} = \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left(I - \hat{\mathbf{W}}_k^{\text{true}} \hat{\mathbf{W}}_k^{\text{true}\,T} \right)\,, \tag{8}$$

and

$$\hat{\mathbf{W}}_k^{\text{true}} \equiv A\,\hat{\mathbf{V}}_k\,, \tag{9}$$

In actual computations we must replace $\hat{\mathbf{W}}_k^{\text{true}}$ by $\hat{\mathbf{W}}_k$, because the former is not known in general. Since we will be interested in calculating quantities only to lowest nonvanishing order in $\Delta\hat{\mathbf{W}}_k$ this replacement will not lead to important errors in general.

In a previous work[11] a method was presented which simplified the attitude estimation process using data from an Earth albedo sensor. In that work, an approximate measurement for the direction of the Earth albedo centroid was determined by taking an average of the centroid of the directions of individual elements of the Earth albedo sensor weighted by the measured intensity, which was compared with a simulated model centroid. The effective vector measurement was combined with a measurement of the Sun direction and used as input to the TRIAD algorithm[3]. It could equally well have been used as input to the QUEST algorithm, but the minuscule improvement in accuracy was not justified by the additional computational burden. Brozenec and Bender[12] used a similar averaging of multiple star directions in a star camera to generate a reduced set of measurements for the QUEST algorithm. In the present work we present a method for retaining full three-axis attitude information from multiple data from a single optical sensor, typically a star camera. In addition, rather than relying on heuristic arguments, we will develop the algorithm in a rigorous manner. We call the algorithm SCAD (Star Camera Attitude Determination).

CONSTRUCTION OF A SUBOPTIMAL COST FUNCTION

Let us reexamine the Wahba cost function, which we write in the form of the data-dependent part of the negative-log-likelihood function[9,10], assuming that the measurement model of Eqs. (2) through (4) is valid, namely,

$$J(A) = \frac{1}{2} \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left| \hat{\mathbf{W}}_k - A\hat{\mathbf{V}}_k \right|^2\,. \tag{10}$$

Let us now introduce vectors $\overline{\mathbf{W}}$ and $\overline{\mathbf{V}}$ into the cost function as

$$J(A) = \frac{1}{2}\sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left| \left[\overline{\mathbf{W}} - A\overline{\mathbf{V}}\right] + \left[(\hat{\mathbf{W}}_k - \overline{\mathbf{W}}) - A(\hat{\mathbf{V}}_k - \overline{\mathbf{V}})\right] \right|^2 , \tag{11}$$

and expand the cost function as

$$\begin{aligned} J(A) = {} & \frac{1}{2}\frac{1}{\sigma_{\text{tot}}^2}|\overline{\mathbf{W}} - A\overline{\mathbf{V}}|^2 \\ & + (\overline{\mathbf{W}} - A\overline{\mathbf{V}})^T \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \Big((\hat{\mathbf{W}}_k - \overline{\mathbf{W}}) - A(\hat{\mathbf{V}}_k - \overline{\mathbf{V}}) \Big) \\ & + \frac{1}{2}\sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left| (\hat{\mathbf{W}}_k - \overline{\mathbf{W}}) - A(\hat{\mathbf{V}}_k - \overline{\mathbf{V}}) \right|^2 . \end{aligned} \tag{12}$$

If now $\overline{\mathbf{W}}$ and $\overline{\mathbf{V}}$ are chosen to have the values

$$\overline{\mathbf{W}} = \sum_{k=1}^{N} a_k \hat{\mathbf{W}}_k , \qquad \text{and} \qquad \overline{\mathbf{V}} = \sum_{k=1}^{N} a_k \hat{\mathbf{V}}_k , \tag{13}$$

with a_k, $k = 1, \ldots, N$, given by Eq. (5), then the second line of Eq. (12) will vanish identically, and the third line will be a minimum (for given A) leaving

$$L(A) = \frac{1}{2}\frac{1}{\sigma_{\text{tot}}^2}|\overline{\mathbf{W}} - A\overline{\mathbf{V}}|^2 + \frac{1}{2}\sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left| (\hat{\mathbf{W}}_k - \overline{\mathbf{W}}) - A(\hat{\mathbf{V}}_k - \overline{\mathbf{V}}) \right|^2 \tag{14}$$

For a focal-plane sensor with a field of view of ±0.1 rad per axis (roughly ±6 deg per axis), we anticipate that the effective contribution of the second summation in Eq. (14) will be roughly $(0.1)^2$ or one per cent of the first. Thus, the estimation of the spacecraft attitude will be "dominated" by the first term. The second term, which could be discarded if a another vector sensor were present [11,12], is not unimportant, however, if data from this sensor alone must be used to construct the three-axis attitude.* Minimizing only the first term is not sufficient to determine the spacecraft attitude. If A_o minimizes the first term, then so does $R(\hat{\overline{\mathbf{W}}}, \psi)A_o$, where $R(\hat{\overline{\mathbf{W}}}, \psi)$ denotes the direction-cosine matrix for a rotation through an arbitrary angle ψ about the direction $\hat{\overline{\mathbf{W}}}$,

$$\hat{\overline{\mathbf{W}}} = \frac{\overline{\mathbf{W}}}{|\overline{\mathbf{W}}|} . \tag{15}$$

It is the second term of Eq. (14) which provides the information on ψ.

Since the overall weight of the first summation in Eq. (14) will be so much greater than that of the second term, we can determine an approximate value for the optimal attitude by writing

$$A = R(\hat{\overline{\mathbf{W}}}, \psi)\, A_o , \tag{16}$$

and seeking first the value A_o^* which minimizes

$$L'(A_o) \equiv \frac{1}{2}\frac{1}{\sigma_{\text{tot}}^2} \left| \overline{\mathbf{W}} - A_o\overline{\mathbf{V}} \right|^2 , \tag{17}$$

*In fact, in the illustrative example of Ref. 12, the second sensor is an Earth horizon scanner, whose data is sufficiently poor that attitude accuracy about the star-camera boresight is worsened by an order of magnitude by the averaging procedure. The algorithm of Ref. 12 will not lead to a loss of attitude estimation accuracy, however, if applied to the case of two noncollinear star trackers, or if the second sensor is a precise Sun sensor.

and then the value ψ^* which minimizes

$$L''(\psi) \equiv \frac{1}{2} \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left| (\hat{\mathbf{W}}_k - \overline{\mathbf{W}}) - R(\widehat{\overline{\mathbf{W}}}, \psi)\, A_o^*(\hat{\mathbf{V}}_k - \overline{\mathbf{V}}) \right|^2 . \tag{18}$$

Given these A_o^* and ψ^*, we anticipate that

$$A^* \equiv R(\widehat{\overline{\mathbf{W}}}, \psi^*)\, A_o^* , \tag{19}$$

will be a good approximation for the optimal direction-cosine matrix which minimizes the cost function of Eq. (10).

SIMPLIFICATION OF THE COST FUNCTIONS

We can simplify the two cost functions, $L'(A_o)$ and $L''(\psi)$, without loss of accuracy. Examine first $L'(A_o)$. Defining

$$\epsilon \equiv \frac{|\overline{\mathbf{W}}| - |\overline{\mathbf{V}}|}{|\overline{\mathbf{W}}|} , \tag{20}$$

we write

$$|\overline{\mathbf{V}}| = (1 - \epsilon)\, |\overline{\mathbf{W}}| , \tag{21}$$

and we can recast $L'(A_o)$ accordingly in the form

$$L'(A_o) = \frac{1}{2} \frac{|\overline{\mathbf{W}}|^2}{\sigma_{\text{tot}}^2} \left| \widehat{\overline{\mathbf{W}}} - (1 - \epsilon) A_o \widehat{\overline{\mathbf{V}}} \right|^2 . \tag{22}$$

The optimizing value of A_o will cause $A_o \widehat{\overline{\mathbf{V}}}$ to be parallel to $\widehat{\overline{\mathbf{W}}}$ independently of the value of ϵ, Thus, we will achieve the identical value of A_o^* if we discard ϵ in Eq. (22).

Likewise, substituting Eq. (21) into Eq. (18) leads to

$$L''(\psi) = \frac{1}{2} \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left| \left(\hat{\mathbf{W}}_k - R(\widehat{\overline{\mathbf{W}}}, \psi)\, A_o^* \hat{\mathbf{V}}_k \right) - \left(\overline{\mathbf{W}} - R(\widehat{\overline{\mathbf{W}}}, \psi)\, A_o^* \overline{\mathbf{V}} \right) \right|^2 \tag{23a}$$

$$= \frac{1}{2} \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left| \left(\hat{\mathbf{W}}_k - R(\widehat{\overline{\mathbf{W}}}, \psi)\, A_o^* \hat{\mathbf{V}}_k \right) - \epsilon \overline{\mathbf{W}} \right|^2 . \tag{23b}$$

Separating the terms in the argument of the vector norm which are parallel and perpendicular to $\widehat{\overline{\mathbf{W}}}$ leads further to

$$L''(\psi) = \frac{1}{2} \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left\{ \left| \left(I - \widehat{\overline{\mathbf{W}}}\widehat{\overline{\mathbf{W}}}^T \right) \left(\hat{\mathbf{W}}_k - R(\widehat{\overline{\mathbf{W}}}, \psi) A_o^* \hat{\mathbf{V}}_k \right) \right|^2 \right. \\ \left. + \left| \widehat{\overline{\mathbf{W}}}^T \left(\hat{\mathbf{W}}_k - A_o^* \hat{\mathbf{V}}_k \right) - \epsilon |\overline{\mathbf{W}}| \right|^2 \right\} \tag{24a}$$

$$= \frac{1}{2} \sum_{k=1}^{N} \frac{1}{\sigma_k^2} \left\{ \left| \hat{\mathbf{W}}_k - R(\widehat{\overline{\mathbf{W}}}, \psi) A_o^* \hat{\mathbf{V}}_k \right|^2 - \left| \widehat{\overline{\mathbf{W}}}^T \left(\hat{\mathbf{W}}_k - A_o^* \hat{\mathbf{V}}_k \right) \right|^2 \right. \\ \left. + \left| \widehat{\overline{\mathbf{W}}}^T \left(\hat{\mathbf{W}}_k - A_o^* \hat{\mathbf{V}}_k \right) - \epsilon |\overline{\mathbf{W}}| \right|^2 \right\} \tag{24b}$$

The last two terms of Eq. (24b) are independent of ψ and may be discarded, therefore, from the cost function. Thus, we may write finally,

$$L'(A_o) = \frac{1}{2}\left|\widehat{\overline{\mathbf{W}}} - A_o\widehat{\overline{\mathbf{V}}}\right|^2 , \tag{25a}$$

$$L''(\psi) = \frac{1}{2}\sum_{k=1}^{N}\frac{1}{\sigma_k^2}\left|\hat{\mathbf{W}}_k - R(\widehat{\overline{\mathbf{W}}}, \psi)\, A_o^*\hat{\mathbf{V}}_k\right|^2 , \tag{25b}$$

Note that the simplification of Eqs. (17) and (18) to obtain Eqs. (25) did not rely on any approximation for the value of ϵ/ Note also that we have discarded an interesting factor in Eq. (25a).

We determine the suboptimal attitude by minimizing the two cost functions of Eq. (25), $L'(A_o)$ and $L''(\psi)$, in sequence.

CONSTRUCTION OF THE SUBOPTIMAL ATTITUDE

The cost function of Eq. (25a) can be made to vanish exactly for a continuum of solutions A_o^*. Except for the special cases $\widehat{\overline{\mathbf{W}}} = \widehat{\overline{\mathbf{V}}}$, for which an A_o^* may be found trivially, a suitable A_o^* is given by[13]

$$A_o^* = (\hat{\mathbf{W}}_1 \cdot \hat{\mathbf{V}}_1)\, I + \frac{(\hat{\mathbf{W}}_1 \times \hat{\mathbf{V}}_1)(\hat{\mathbf{W}}_1 \times \hat{\mathbf{V}}_1)^T}{1 + \hat{\mathbf{W}}_1 \cdot \hat{\mathbf{V}}_1} + [[\,\hat{\mathbf{W}}_1 \times \hat{\mathbf{V}}_1\,]] \tag{26a}$$

$$= I + [[\,\hat{\mathbf{W}}_1 \times \hat{\mathbf{V}}_1\,]] + \frac{1}{1 + \hat{\mathbf{W}}_1 \cdot \hat{\mathbf{V}}_1}[[\,\hat{\mathbf{W}}_1 \times \hat{\mathbf{V}}_1]]^2 . \tag{26b}$$

corresponding to the quaternion[1]

$$\bar{q}_o^* = \sqrt{\frac{1 + \hat{\mathbf{W}}_1 \cdot \hat{\mathbf{V}}_1}{2}}\begin{bmatrix}\left(\dfrac{\hat{\mathbf{W}}_1 \times \hat{\mathbf{V}}_1}{1 + \hat{\mathbf{W}}_1 \cdot \hat{\mathbf{V}}_1}\right) \\ 1\end{bmatrix} , \tag{27}$$

and Rodrigues vector[1]

$$\boldsymbol{\rho}_o^* = \frac{\hat{\mathbf{W}}_1 \times \hat{\mathbf{V}}_1}{1 + \hat{\mathbf{W}}_1 \cdot \hat{\mathbf{V}}_1} . \tag{28}$$

The particular A_o^* that we chose is of no consequence, provided that it satisfy

$$A_o^*\widehat{\overline{\mathbf{V}}} = \widehat{\overline{\mathbf{W}}} . \tag{29}$$

It remains only to find the angle ψ^* which minimizes the cost function of Eq. (25b).

To determine ψ^* we rewrite $L''(\psi)$, using techniques developed by Davenport[14] which have become part of the development of the QUEST algorithm[3], as

$$L''(\psi) = \frac{1}{\sigma_{\text{tot}}^2} - \sum_{k=1}^{N}\frac{1}{\sigma_k^2}\hat{\mathbf{W}}_k^T R(\widehat{\overline{\mathbf{W}}}, \psi)\, A_o^*\hat{\mathbf{V}}_k \tag{30a}$$

$$= \frac{1}{\sigma_{\text{tot}}^2} - \operatorname{tr}\left[B^T R(\widehat{\overline{\mathbf{W}}}, \psi)\right] , \tag{30b}$$

where tr(·) denotes the trace operation, and

$$B = \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2} \hat{\mathbf{W}}_k \hat{\mathbf{V}}_k^T \right) A_o^{*T} \equiv C A_o^{*T} . \tag{31}$$

Writing Euler's formula[1] in the form

$$R(\widehat{\widehat{\mathbf{W}}}, \psi) = \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T + \sin\psi\, [[\widehat{\widehat{\mathbf{W}}}]] + \cos\psi \left(I - \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T \right) , \tag{32}$$

where

$$[[\,\mathbf{v}\,]] \equiv \begin{bmatrix} 0 & v_3 & -v_2 \\ -v_3 & 0 & v_1 \\ v_2 & -v_1 & 0 \end{bmatrix} , \tag{33}$$

we have

$$L''(\psi) = \frac{1}{\sigma_{\text{tot}}^2} - \text{tr}[B] - \sin\psi\, \text{tr}\left[B^T [[\widehat{\widehat{\mathbf{W}}}]] \right] - \cos\psi\, \text{tr}\left[B^T \left(I - \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T \right) \right] \tag{34a}$$

$$\equiv \frac{1}{\sigma_{\text{tot}}^2} - \text{tr}[B] - s\, \sin\psi - c\, \cos\psi , \tag{34b}$$

with

$$\mathbf{Z} \equiv \begin{bmatrix} B_{23} - B_{32} \\ B_{31} - B_{13} \\ B_{12} - B_{21} \end{bmatrix} , \qquad s \equiv \left(\mathbf{Z}^T \widehat{\widehat{\mathbf{W}}} \right) , \qquad \text{and} \qquad c \equiv \left(\text{tr}\,[B] - \widehat{\widehat{\mathbf{W}}}^T B \widehat{\widehat{\mathbf{W}}} \right) . \tag{35}$$

Minumization of $L''(\psi)$ leads straightforwardly to

$$-s\, \cos\psi^* + c\, \sin\psi^* = 0 , \tag{36}$$

or

$$\psi^* = \text{arctan}_2(s,\, c) . \tag{37a}$$

Here $\text{arctan}_2(s, c)$ is the function which returns the arc tangent of s/c in the correct quadrant. In the FORTRAN language this function is called ATAN2. The angle ψ^* will be indeterminate if both s and c vanish. This is possible, however, only if all of the $\hat{\mathbf{W}}_k$ are identical.

The parallelism of the calculation of ψ^* in the present algorithm with that of $\bar{q}^*$ in the QUEST algorithm is apparent. However, this methods are applied to a single angle variable and not to the four components of the quaternion of rotation. The computational burden is therefore much smaller, particularly since the need to compute the overlap eigenvalue has been eliminated.

COVARIANCE ANALYSIS OF THE ALGORITHM

A simple approximate expression for the covariance matrix of the SCAD algorithm can be obtained if we neglect correlations between the two steps. In that case, we effectively treat the estimation of A_o and ψ as separate maximum likelihood estimation problems and can obtain an approximate estimate error covariance matrix from the Fisher information matrix associated with each of the estimation steps. Clearly, the direction-cosine matrix of Eqs. (26) causes the cost function $L'(A_o^*)$ to vanish identically. Therefore, the coefficient in Eq. (25a) has no direct connection to the covariance matrix of the suboptimal attitude estimate, once we have embarked on our two-step optimization sequence.

To compute the attitude covariance matrix, $P_{\theta\theta}$, we must begin with the correct measurement model characterizing $\widehat{\overline{\mathbf{W}}}$.

From the definition of $\overline{\mathbf{W}}$ and $\overline{\mathbf{V}}$ we may write

$$\overline{\mathbf{W}} = A_o \overline{\mathbf{V}} + \Delta\overline{\mathbf{W}}, \tag{38}$$

where

$$\Delta\overline{\mathbf{W}} = \sum_{k=1}^{N} a_k \, \Delta\hat{\mathbf{W}}_k \, . \tag{39}$$

$\Delta\overline{\mathbf{W}}$ will have vanishing expectation (to order σ_{tot}) and covariance matrix

$$R_{\overline{\mathbf{W}}} = \sigma_{\text{tot}}^2 \sum_{k=1}^{N} a_k \left(I - \hat{\mathbf{W}}_k \hat{\mathbf{W}}_k^T \right) \, . \tag{40}$$

From Eq. (38) it follows that

$$\widehat{\overline{\mathbf{W}}} = A_o \widehat{\overline{\mathbf{V}}} + \Delta\widehat{\overline{\mathbf{W}}}, \tag{41}$$

with

$$\Delta\widehat{\overline{\mathbf{W}}} = \frac{1}{|\overline{\mathbf{W}}|} \left(I - \widehat{\overline{\mathbf{W}}}\widehat{\overline{\mathbf{W}}}^T \right) \Delta\overline{\mathbf{W}} \, . \tag{42}$$

Thus, $\Delta\widehat{\overline{\mathbf{W}}}$ also has vanishing expectation and covariance matrix

$$R_{\widehat{\overline{\mathbf{W}}}} = \left(I - \widehat{\overline{\mathbf{W}}}\widehat{\overline{\mathbf{W}}}^T \right) R_{\overline{\mathbf{W}}} \left(I - \widehat{\overline{\mathbf{W}}}\widehat{\overline{\mathbf{W}}}^T \right) \, . \tag{43}$$

The data-dependent part of the negative-log-likelihood[9] corresponding to the measurement of model of Eq. (41) is, therefore,[9,10]

$$J(A) = \frac{1}{2} \left(\widehat{\overline{\mathbf{W}}} - A\widehat{\overline{\mathbf{V}}} \right)^T R_{\widehat{\overline{\mathbf{W}}}}^{\#} \left(\widehat{\overline{\mathbf{W}}} - A\widehat{\overline{\mathbf{V}}} \right) \tag{44a}$$

$$= \frac{1}{2} \left(\widehat{\overline{\mathbf{W}}} - (I - [[\Delta\boldsymbol{\theta}]]\, \widehat{\overline{\mathbf{W}}}^{\text{true}} \right)^T R_{\widehat{\overline{\mathbf{W}}}}^{\#} \left(\widehat{\overline{\mathbf{W}}} - (I - [[\, \Delta\boldsymbol{\theta}\,]]\, \widehat{\overline{\mathbf{W}}}^{\text{true}} \right)^T \tag{44b}$$

$$= \frac{1}{2} \left(\widehat{\overline{\mathbf{W}}} - \widehat{\overline{\mathbf{W}}}^{\text{true}} + [[\, \widehat{\overline{\mathbf{W}}}^{\text{true}} \,]]\, \Delta\boldsymbol{\theta} \right)^T R_{\widehat{\overline{\mathbf{W}}}}^{\#} \frac{1}{2} \left(\widehat{\overline{\mathbf{W}}} - \widehat{\overline{\mathbf{W}}}^{\text{true}} + [[\, \widehat{\overline{\mathbf{W}}}^{\text{true}} \,]]\, \Delta\boldsymbol{\theta} \right) , \tag{44c}$$

where

$$\widehat{\overline{\mathbf{W}}}^{\text{true}} = A\, \widehat{\overline{\mathbf{V}}}, \tag{45}$$

and $\Delta\boldsymbol{\theta}$ is the attitude error vector. The Hessian matrix (i.e., the matrix of second-order partial derivatives of the negative-log-likelihood function of Eq. (44) with respect to $\Delta\boldsymbol{\theta}$) is then the contribution of the effective measurement $\widehat{\overline{\mathbf{W}}}$ to $P_{\theta\theta}^{-1}$. Thus,

$$(P_{\theta\theta}^{-1})' = [[\, \widehat{\overline{\mathbf{W}}}^{\text{true}} \,]]\, R_{\widehat{\overline{\mathbf{W}}}}^{\#}\, [[\, \widehat{\overline{\mathbf{W}}}^{\text{true}} \,]]^T \tag{46}$$

and the single prime denotes that this is the contribution to the attitude covariance arising from $\widehat{\overline{\mathbf{W}}}$.

To determine the information matrix associated with the degree of freedom expressed by ψ let us write

$$\hat{\mathbf{U}}_k \equiv R(\widehat{\widehat{\mathbf{W}}},\ \psi_{\text{true}})\, A_o^*\, \hat{\mathbf{V}}_k\ , \tag{47}$$

where ψ_{true} is the true value of ψ. (Since the errors associated with A_o^* are expected to be very small compared to the errors associated with ψ^* because of the geometric dilution of precision associated with the limited field of view of the star tracker, we expect $\hat{\mathbf{U}}_k$ to be very close to $\hat{\mathbf{W}}_k^{\text{true}}$. Thus, we now seek the value of the infinitesimal $\Delta\psi$ which minimizes

$$L''(\Delta\psi) = \sum_{k=1}^{N} \frac{1}{\sigma_k^2}\, |\hat{\mathbf{W}}_k - R(\widehat{\widehat{\mathbf{W}}},\ \Delta\psi)\, \hat{\mathbf{U}}_k|^2 \tag{48a}$$

$$= \sum_{k=1}^{N} \frac{1}{\sigma_k^2}\, |\hat{\mathbf{W}}_k - \hat{\mathbf{U}}_k - [[\, \Delta\psi \widehat{\widehat{\mathbf{W}}}\,]]\, \hat{\mathbf{U}}_k|^2 \tag{48b}$$

$$= \sum_{k=1}^{N} \frac{1}{\sigma_k^2}\, |\hat{\mathbf{W}}_k - \hat{\mathbf{U}}_k + \Delta\psi\, \widehat{\widehat{\mathbf{W}}} \times \hat{\mathbf{U}}_k|^2 \tag{48c}$$

$$= \sum_{k=1}^{N} \frac{1}{\sigma_k^2}\, \left[|\hat{\mathbf{W}}_k - \hat{\mathbf{U}}_k|^2 + 2(\hat{\mathbf{W}}_k - \hat{\mathbf{U}}_k)^T (\hat{\mathbf{W}}_k \times \hat{\mathbf{U}}_k)\, \Delta\psi + |\hat{\mathbf{W}}_k \times \hat{\mathbf{U}}_k|^2 |\Delta\psi|^2 \right] \tag{48d}$$

The information for $\Delta\psi$ is just the second derivative of this quantity with respect to $\Delta\psi$, and, since this angle is about the direction $\widehat{\widehat{\mathbf{W}}}$, its contribution to the attitude information is just

$$\left(P_{\theta\theta}{}^{-1}\right)'' = \sum_{k=1}^{N} \frac{1}{\sigma_k^2}\, |\widehat{\widehat{\mathbf{W}}} \times \hat{\mathbf{U}}_k|^2\, \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T .$$

Thus, the inverse of the covariance matrix of our new algorithm is

$$\left(P_{\theta\theta}^{\text{SCAD}}\right)^{-1} = [[\, \widehat{\widehat{\mathbf{W}}}\,]]\, R_{\widehat{\widehat{\mathbf{W}}}}^{\#}\, [[\, \widehat{\widehat{\mathbf{W}}}\,]]^T + \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2}\, |\widehat{\widehat{\mathbf{W}}} \times \hat{\mathbf{W}}_k|^2 \right) \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T . \tag{49}$$

In anticipation of practical application we have made the replacements

$$\widehat{\widehat{\mathbf{W}}}^{\text{true}} \to \widehat{\widehat{\mathbf{W}}}, \qquad \text{and} \qquad \hat{\mathbf{U}}_k^* \to \hat{\mathbf{W}}_k\ . \tag{50}$$

The geometric dilution of precision is manifest in the factors $|\widehat{\widehat{\mathbf{W}}} \times \hat{\mathbf{W}}_k|^2$.

The computation of the pseudo-inverse is easily accomplished. Let $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ be any two unit vectors such that $\{\hat{\mathbf{u}},\ \hat{\mathbf{v}},\ \widehat{\widehat{\mathbf{W}}}\}$ form a right-handed orthonormal set. Since the singularity of $R_{\widehat{\widehat{\mathbf{W}}}}$ is solely the manifestation that $\widehat{\widehat{\mathbf{W}}}$ is a null vector of the covariance matrix, it follows that

$$R_{\widehat{\widehat{\mathbf{W}}}} = F F^T R_{\widehat{\widehat{\mathbf{W}}}}\, F F^T , \tag{51}$$

where

$$F \equiv [\, \hat{\mathbf{u}} \,\vdots\, \hat{\mathbf{v}} \,] \tag{52}$$

is a (3) × 2 matrix labeled by its columns. Note that

$$F\,F^T = I - \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T . \tag{53}$$

The 2 × 2 matrix R',

$$R' = F^T R_{\widehat{\widehat{\mathbf{W}}}} F , \tag{54}$$

will be nonsingular if the attitude is observable. Hence in the only cases of interest

$$R^{\#}_{\widehat{\widehat{\mathbf{W}}}} = F \left(F^T R_{\widehat{\widehat{\mathbf{W}}}} F \right)^{-1} F^T . \tag{55}$$

The verification that R' is nonsingular would be a routine step in the computation of the attitude to verify observability. A similar step occurs in implementation of the QUEST algorithm in which the rank of the QUEST information matrix is tested[15].

We may simplify the expression for the SCAD inverse covariance matrix further by noting that

$$[[\widehat{\widehat{\mathbf{W}}}]]\,F \equiv G = [\,-\hat{\mathbf{v}} \,\vdots\, \hat{\mathbf{u}}\,] . \tag{56}$$

Hence, we have finally

$$\left(P^{\text{SCAD}}_{\theta\theta}\right)^{-1} = G \left(F^T R_{\widehat{\widehat{\mathbf{W}}}} F \right)^{-1} G^T + \left(\sum_{k=1}^{N} \frac{1}{\sigma_k^2} |\widehat{\widehat{\mathbf{W}}} \times \hat{\mathbf{W}}_k|^2 \right) \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T . \tag{57}$$

Note that the second term in Eq. (57) may be written as

$$\widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T \left(P^{\text{QUEST}}_{\theta\theta}\right)^{-1} \widehat{\widehat{\mathbf{W}}}\widehat{\widehat{\mathbf{W}}}^T$$

Note that the QUEST covariance matrix will be a good approximation for the SCAD covariance matrix when the field of view of the sensor has a small diameter. The comparison of these expressions will be carried out in the next section.

If ρ is the root-mean-square arc length of the individual star observations from $\widehat{\widehat{\mathbf{W}}}$, then we anticipate that the first term in Eq. (57) will differ from the corresponding term of Eq. (49) by fractional errors of order ρ^2. By this same token, we observe that the contribution of the second term in Eq. (49) or (57) to the total inverse covariance matrix will be smaller than that of the first term by a factor of ρ^2.

MODEL COVARIANCE ANALYSIS

It follows from the Cramér-Rao Theorem[9] that

$$P^{\text{QUEST}} \le P^{\text{SCAD}} . \tag{58}$$

The important question is how large is the difference between the two attitude covariance matrices. To answer this question, we examine the two covariance matrices (rather than the two inverse covariances matrices) in a simple model, in which the star camera is assumed to have a circular field of view of angular radius ρ, and the stars are assumed to be distributed uniformly over the field of view of the sensor. We will assume for convenience that the star camera has its boresight along the spacecraft z-axis. We assume in addition that the covariance matrix of every line-of-sight observation is characterized by the same standard deviation σ^2.

In the limit that N is large we may replace the summation over the observations by an integral. Thus, if $f(\hat{\mathbf{W}})$ is any function of an observed direction, we may write

$$\sum_{k=1}^{N} f(\hat{\mathbf{W}}_k) \to \frac{N}{\Omega} \int_0^{2\pi} \int_0^{\rho} f\big(\hat{\mathbf{W}}(\vartheta, \varphi)\big) \sin\vartheta \, d\vartheta \, d\varphi \,, \tag{59}$$

with Ω the solid angle subtended by the star camera field of view,

$$\Omega = 2\pi(1 - \cos\rho) \,, \tag{60}$$

and

$$\hat{\mathbf{W}}(\vartheta, \varphi) = \begin{bmatrix} \sin\vartheta \, \cos\varphi \\ \sin\vartheta \, \sin\varphi \\ \cos\vartheta \end{bmatrix} \,. \tag{61}$$

With substitutions the inverse covariance matrix for each star camera using the QUEST algorithm for computing the attitude is

$$\left(P_{\theta\theta}^{\text{QUEST}}\right)^{-1} = \frac{N}{\sigma^2} \operatorname{diag}(a, a, b) \,, \tag{62}$$

where

$$\operatorname{diag}(a, b, c) \equiv \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix} \,, \tag{63}$$

and

$$a = (4 + \cos\rho + \cos^2\rho)/6 \,, \tag{64a}$$

$$b = (2 - \cos\rho - \cos^2\rho)/3 \,. \tag{64b}$$

Note that as $\rho \to 0$ we have that $a \to 1$ and $b \to 0$.

For the SCAD algorithm, we note first that

$$\overline{\mathbf{W}} = \left(\frac{1 + \cos\rho}{2}\right) \hat{\mathbf{z}} \,, \tag{65a}$$

$$R_{\overline{\mathbf{W}}} = \frac{\sigma^2}{N} \operatorname{diag}(a, a, b) \,. \tag{65b}$$

$$R_{\widehat{\overline{\mathbf{W}}}} = \frac{\sigma^2}{N} \left(\frac{2}{1 + \cos\rho}\right)^2 \operatorname{diag}(a, a, 0) \,. \tag{65c}$$

From these results we may compute the inverse covariance matrix for the SCAD algorithm given in Eq. (49) to obtain

$$\left(P_{\theta\theta}^{\text{SCAD}}\right)^{-1} = \frac{N}{\sigma^2} \left\{ \left(\frac{1 + \cos\rho}{2}\right)^2 \operatorname{diag}(1/a, 1/a, 0) + \operatorname{diag}(0, 0, b) \right\} \,. \tag{66}$$

The two covariance matrices are both diagonal in the model case examined.

We note first that the standard deviation about the boresight is identical for this example for both the QUEST and the SCAD algorithms,

$$\frac{\sigma_b^{\text{SCAD}}}{\sigma_b^{\text{QUEST}}} = 1 \,, \tag{67}$$

TABLE 1
COMPARISON OF THE SCAD AND QUEST ALGORITHMS

Field of View	$\sigma_t^{\mathrm{SCAD}}/\sigma_t^{\mathrm{QUEST}}$	$\sigma_b^{\mathrm{SCAD}}/\sigma_b^{\mathrm{QUEST}}$
± 6 deg	1.000004	1.000000
± 12 deg	1.00007	1.000000
± 30 deg	1.003	1.000000
± 60 deg	1.06	1.000000
± 90 deg	1.33	1.000000

where the subscript b stands for "boresight." Thus, not only do we recover the information on the attitude about the boresight, we recover it completely.

The ratio of the standard deviation of the SCAD algorithm to that of the QUEST algorithm for attitude errors about axes normal to $\widehat{\overline{\mathbf{W}}}$ is

$$\frac{\sigma_t^{\mathrm{SCAD}}}{\sigma_t^{\mathrm{QUEST}}} = \frac{a}{|\overline{\mathbf{W}}|} = \frac{2}{1+\cos\rho}\,\frac{4+\cos\rho+\cos^2\rho}{6}\,, \tag{68}$$

where t stands for "transverse." Since we are interested in this algorithm primarily for a sensor of limited field of view, we define

$$\delta \equiv 1 - \cos\rho\,. \tag{69}$$

Then

$$\frac{\sigma_t^{\mathrm{SCAD}}}{\sigma_t^{\mathrm{QUEST}}} = \frac{2}{2-\delta}\,\frac{6-3\delta+\delta^2}{6} = 1 + \frac{1}{6}\,\frac{\delta^2}{1-\delta/2}\,. \tag{70}$$

Thus, for this simple example, the standard deviation of attitude errors for axes perpendicular to the boresight is approximately

$$\frac{\sigma_t^{\mathrm{SCAD}}}{\sigma_t^{\mathrm{QUEST}}} \approx 1 + \rho^4/24 + O(\rho^6/96)\,. \tag{71}$$

For limited fields of view, the relative loss in accuracy compared to the QUEST algorithm is imperceptible. Table 1 gives the relative loss of accuracy for several fields of view. Note that because of the rotational symmetry of our example about the star-camera boresight, The cross covariance matrix between $\Delta\psi$ and $\Delta\theta$ will vanish. Thus, the errors introduced by our approximate treatment of the attitude estimate covariance are completely suppressed.

REFERENCES

1. SHUSTER, M. D. "A Survey of Attitude Representations," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, pp. 439–517, October–December 1993.
2. WAHBA, G., "Problem 65–1: A Least Squares Estimate of Spacecraft Attitude," *Siam Review*, Vol. 7, No. 3, July 1965, p. 409.
3. SHUSTER, M. D., and OH, S. D. "Three-Axis Attitude Determination from Vector Observations," *Journal of Guidance and Control*, Vol. 4, No. 1, pp. 70–77, January-February 1981.
4. MARKLEY, F. L. "Attitude Determination using Vector Observations and the Singular Value Decomposition," *Journal of the Astronautical Sciences*, Vol. 36, No. 3, pp. 245–258, July–Sept. 1988.

5. MARKLEY, F. L., "Attitude Determination Using Vector Observations: A Fast Optimal Matrix Algorithm," *Journal of the Astronautical Sciences*, Vol. 41, No. 2, pp. 439–517, April–June 1993.

6. MORTARI, D., "Energy Approach Algorithm for Attitude Determination from Vector Observations, *Journal of the Astronautical Sciences* (to appear).

7. MORTARI, D., "Euler-q Algorithm for Attitude Determination from Vector Observations," submitted to the *Journal of Guidance, Control and Dynamics*.

8. MORTARI, D., "A Closed-Form Solution to the Wahba Problem," submitted to the *Journal of the Astronautical Sciences*.

9. SORENSON, H. W., *Parameter Estimation,* Marcel Dekker, New York, 1980.

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

11. FISHER, H. L., MUSSER, K. L., and SHUSTER, M. D., "Coarse Attitude Determination from Earth Albedo Measurements," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 29, No. 1, Jan.–Feb. 1993, pp. 22-26.

12. BROZENEC, THOMAS F., and BENDER, DOUGLAS J., "A Simple Suboptimal Least-Squares Algorithm for Attitude Determination with Multiple Sensors," *Proceedings, Flight Mechanics/Estimation Theory Symposium,* NASA Goddard Space Flight Center, Greenbelt, Maryland, May 17–19, 1994, pp. 529–543.

13. SHUSTER, M. D., "Attitude Estimation from the Measurement of a Direction and an Angle," *Revista Brasileira de Ciências Mecânicas*, Vol. 16, Special Issue, pp. 19–23, 1994.

14. DAVENPORT, P. B. (unpublished).

15. SHUSTER, M. D., (unpublished).

16. SHUSTER, M. D., and NATANSON, G. A., "Quaternion Computation from a Geometric Point of View," *Journal of the Astronautical Sciences*, Vol. 41, No. 4, pp. 545–556, October–December 1993.

17. SHUSTER, M. D., 'Kalman Filtering of Spacecraft Attitude and the QUEST Model," *Journal of the Astronautical Sciences* Vol. 38, No. 3, pp. 377–393, July–September, 1990.

Appendix: Implementation of SCAD

The following are the steps for computing the optimal attitude using the SCAD algorithm:

- From the input data, $\hat{\mathbf{W}}_k$, $k = 1, \ldots, N$, the corresponding reference vectors, $\hat{\mathbf{V}}_k$, $k = 1, \ldots, N$, and the sensor variances, σ_k^2, $k = 1, \ldots, N$, compute: (1) σ_{tot}^2 according to Eq. (6); (2) the weights a_k, $k = 1, \ldots, N$, according to Eq. (5); (3) $\overline{\mathbf{W}}$ and $\overline{\mathbf{V}}$ according to Eq. (13); and (4) the matrix C according to Eq. (31).

- From these quantities compute the unit vectors $\widehat{\overline{\mathbf{W}}}$ and $\widehat{\overline{\mathbf{V}}}$ according to Eq. (15) for $\widehat{\overline{\mathbf{W}}}$ and similarly for $\widehat{\overline{\mathbf{V}}}$.

- Compute $R_{\widehat{\overline{\mathbf{W}}}}$ according to Eqs. (40) and (43).

- Compute the matrices F and G according to Eqs. (54) and (56).

- Compute R' from Eq. (54) and verify that it is full rank. If R' is not full tank, the attitude is not observable from the data and the computation is terminated. If R' is full rank, continue.

- Compute A_o^* according to the following method:

  - If $\widehat{\widehat{\mathbf{W}}} \cdot \widehat{\mathbf{V}} > 1 - \varepsilon$ for some predetermined value of ε, set $A_o^* = I$. (The value of ε will be a function of the machine precision and the accuracy of the data.)
  - If $\widehat{\widehat{\mathbf{W}}} \cdot \widehat{\mathbf{V}} < -1 + \varepsilon$ for some predetermined value of ε, set $A_o^* = R(\hat{\mathbf{n}}, \pi)$, where $\hat{\mathbf{n}}$ is the representation of the sensor coordinate axis for which $|\hat{\mathbf{n}} \times \widehat{\widehat{\mathbf{W}}}|$ is largest.
  - Otherwise, use any of Eqs. (26a) through (28) to generate A_o^* either directly or via the quaternion or Rodrigues vector.

- Compute B according to Eq. (31), and $\mathbf{Z}$, s, and c according to Eq. (35).

- Compute ψ^* according to Eq. (37) and A^* according to Eq. (19).

- Compute $P_{\theta\theta}^{\text{SCAD}}$ according to Eq. (57). o

This completes the SCAD algorithm.

The above implementation was given with a mind to generating the direction-cosine matrix as final output. If it is desired to generate instead either the quaternion or the Rodrigues vector as final output, one requires the formulae:

$$\cos(\psi^*/2) = \sqrt{\frac{1+c}{2}}\,, \qquad \text{and} \qquad \sin(\psi^*/2) = \frac{s}{2\cos(\psi^*/2)}\,, \tag{A1}$$

whence

$$\bar{q}(\widehat{\widehat{\mathbf{W}}}, \psi^*) = \begin{bmatrix} \sin(\psi^*/2)\,\widehat{\widehat{\mathbf{W}}} \\ \cos(\psi^*/2) \end{bmatrix}, \qquad \text{and} \qquad \boldsymbol{\rho}(\widehat{\widehat{\mathbf{W}}}, \psi^*) = \frac{\sin(\psi^*/2)}{\cos(\psi^*/2)}\,\widehat{\widehat{\mathbf{W}}}\,, \tag{A2}$$

and combining these directly with $\bar{q}_0^*$ and $\boldsymbol{\rho}_o^*$ according to the prescriptions in Ref 1.