JOURNAL OF SPACECRAFT AND ROCKETS
Vol. 29, No. 1, January–February 1992

Shape Determination for Large Flexible Satellites via Stereo Vision

D. N. C. Tse* and G. R. Heppler†
University of Waterloo, Waterloo, Ontario, N2L 3G1, Canada

The use of stereo vision to determine the deformed shape of an elastic plate is investigated. The quantization error associated with using discrete charge coupled device camera images for this purpose is examined. An upper bound on the error is derived in terms of the stationary configuration parameters. An expression for the average (root mean square) error is also developed. The issue of interpolating the shape of the plate through erroneous data is addressed. The vibratory mode shapes are used as interpolation functions and two cases are considered: the case when the number of interpolation points (targets) is the same as the number of modes used in the interpolation, and the case when the number of targets exceeds the number of modes used. Error criteria are established for both cases and they provide a means of establishing the best fit to the measured data.

Nomenclature

A	= area of the plate
b_i	= column vectors of Φ^{-1}
$\hat{b}_i$	= the columns of $(\Phi^T\Phi)^{-1}$
b_{ij}	= an entry in Φ^{-1}
$\|C\|$	= the volume of C
$C(e)$	= the rectangular box in R^m centered at the origin and whose sides are of lengths $2e_i$, $i = 1, 2, \ldots, m$
c	= centroid position
d	= length of an error interval
E_h	= length of the error interval (horizontal configuration)
$(E_h)_{\text{rms}}$	= the root mean square error (horizontal configuration)
E_v	= length of the error interval (vertical configuration
$(E_v)_{\text{rms}}$	= the root mean square error (vertical configuration)
$\mathcal{E}_{\text{rms}}$	= average (rms) case error measure
$\mathcal{E}_w$	= worst case error measure
e	= vector of e_i
e_i	= semi-interval of uncertainty on w_i
$\mathcal{F}$	= frac$[(2fn_u\Delta)/X_0a]$
f	= focal length
$f(q)$	= the probability density function of q over Q
frac$(\cdot)$	= the fractional part of the real argument
$g(w)$	= the probability density function of w over W
h	= height of cameras
(I_l, J)	= the object image on the left camera screen in the (i,j) system
(I_r, J)	= the object image on the right camera screen in the (i,j) system
(i,j)	= screen integer coordinate system
$\mathcal{L}$	= locus of positions of a target
l_i	= lower bound on the ith target's deflection w_i
ℓ	= vector of l_i
$\ell_i(X_i, Y_i, w_i)$	= lower bound on the interval of uncertainty for w_i
m	= number of targets
n	= number of modes
n_u, n_v	= number of pixels in the x_c and y_c directions, respectively
P_1, P_2	= two points of intersection of U with $\mathcal{L}$ (the endpoints of an error interval)
$\mathcal{P}$	= the partition of $\mathcal{L}$ into error intervals
$\mathcal{P}_\ell$	= subsets of $\mathcal{P}$ associated wtih the left camera
$\mathcal{P}_r$	= subsets of $\mathcal{P}$ associated with the right camera
Q	= space containing q in R^n
Q_0	= the transformed version of Q that results from translating the centroid of Q to the origin
q	= vector of modal coordinates
q_{LS}	= least square solution
q_n	= modal coordinate
q_T	= true values
q_0	= an estimate of the vector of modal coordinates
q^*	= optimal estimate of the vector of modal coordinates
U	= region of uncertainty
U_i	= upper bound on e_i
u	= vector of u_i
u_i	= upper bound on the ith target's deflection w_i
$u_i(X_i, Y_i, w_i)$	= upper bound on the interval of uncertainty for w_i
u_L, u_R	= the abscissas of the true image of the target in the left and right cameras, respectively
(u_l, v_l)	= continuous left image plane coordinates of the image of a point object
u_{l1}, u_{l2}	= abscissas of the images of P_1 and P_2 in the left camera
(u_r, v_r)	= continuous right image plane coordinates of the image of a point object
u_{r1}, u_{r2}	= abscissas of the images of P_1 and P_2 in the right camera
$V(Q)$	= the set of vertices of Q
v	= a vertex of Q
v_1, v_2	= the ordinates of the error interval endpoints
W	= space containing w
$\|W\|$	= the volume of W
w	= vector of measured displacements
$\hat{w}$	= vector of approximation of the individual target positions ($m > n$ case)
$w(X, Y)$	= deflection of the plate
w_i	= deflection of the ith target
$\hat{w}_T$	= vector of true deflections
(X, Y, Z)	= global reference frame

Received Sept. 13, 1990; revision received Aug. 2, 1991; accepted for publication Aug. 5, 1991.

*Research Assistant, Systems Design Engineering; currently Research Assistant, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

†Associate Professor, Systems Design Engineering. Member AIAA.

(X_0, Y_0, Z) = the true position of the target in the global frame
(x_c, y_c, z_c) = position of a point object in the camera coordinate frame
$Z_L(X, Y)$ = known lower bound on the deflected shape $w(X, Y)$ of the plate
$\hat{z}$ = vector measuring the deviation of the true deflections from the midpoint of their corresponding error intervals
$\hat{z}_i$ = components of $\hat{z}$
$\| \| = \| \|_2$ = Euclidean norm
$\| \|_F$ = Frobenius norm
Δ = semibaseline separation distance
θ = angle of inclination of the cameras
Φ = matrix of mode shapes
ϕ_n = mode shape

Introduction

THIRD generation satellites are characterized by being both much larger than previous satellites and much more flexible. This latter characteristic brings with it, in addition to the usual requirements for attitude control, a requirement for the control of the flexible modes as well. An excellent example of the shape control requirement is the case of a satellite with a large phased array radar. It will be impossible to make the support structure for the antenna array sufficiently stiff for it to be considered rigid, but at the same time there will be constraints on the deformations of the support structure that are governed by radar performance considerations. A certain amount of deformation can be accommodated by the radar provided that the nature of the deformations are known. In order to be known, and possibly controlled, the flexible motions need to be sensed and the physical characteristics of the sensors are important, in that they need to be small and lightweight so that their presence does not significantly alter the dynamics of the satellite. Accelerometers would seem to be a natural choice for sensing the flexible modes because they may be of low mass, low power consumption, high resolution, and of relatively low cost. However, there are features of accelerometers that make them less than ideal[1]; they are prone to large biases that grow with time and can exceed the resolution of the accelerometer, thus destroying the accuracy of the sensor, and accelerometer signals may also be corrupted by disturbances to the satellite as well as to the accelerometer with the added difficulty that these two sources of disturbances are not independent in the case of accelerometers.[1] Hence, it is not entirely clear that accelerometers should be the only motion sensors used or even that they are the best choice. For the purposes of shape control or shape determination, a means of directly measuring the displacements may be advantageous.

The use of stereo charged coupled device (CCD) cameras for the purpose of sensing the deformed shape of an elastic structure, such as a third generation satellite, is very appealing because this method of deflection measurement would certainly not require the addition of any components of significant mass or size to the flexible structure and would in this sense be noninvasive. The implementation of this scheme would require that the structure be viewed by a stereoscopic pair of cameras (possibly more than one pair may be required if the structural dimensions exceed the cameras' field of view) that would look at a set of readily identifiable targets. These targets may be as simple as LEDs placed at suitable locations.

As a reference case, a large rectangular structure that can be treated as a thin plate is used. It is assumed to be made of a linear elastic material and to experience small deflections such that linearized strain-displacement relations may be used. For the case of small deformations and assuming bending to be the dominant mode of deformation, it is reasonable that the in-plane deformations be considered to be negligible relative to the out-of-plane deformations. This means that the targets can be considered to be constrained to move along a line that is perpendicular to the undeformed reference position of the plate. The cameras are idealized as perfect pinhole cameras, thus ignoring lens distortion and other optical nonlinearities such as Seidel aberrations. Although this is a restrictive assumption, it does allow the analysis to go forward without it being complicated by issues that are very difficult to deal with due to their inherent nonlinearities. Indeed, Blostein and Huang[2] made this assumption in their work as have Rodriguez and Aggarwal[3] and McVey and Lee.[4] Because of this assumption, we will be considering the errors associated with postcalibration pixel quantization error effects, but it should be recognized that distortion effects could play a significant role in practice. Because the dominant natural frequencies of the structures of interest are typically very much less than the currently available framing rates for CCD cameras, the effect of the finite CCD integration time has been assumed to be negligible.

Because of the large target to camera separation distances in this application, and under the previously stated assumptions, it is expected that pixel quantization error will dominate over other sources of error. It is the purpose of this correspondence to analyze the shape determination error when the image positions are corrupted by error of this nature. Although the actual error varies as the shape of the plate changes over time, we have derived upper bounds and an expected value of this error as a function only of the positions of the targets on the plate and the camera configuration. These can then be used as a criterion in placing the cameras and targets in the design stage.

Further, it is also assumed that the object of interest (target) is sufficiently small so that modeling it as a single point will not introduce a significant error.

The error associated with using stereo CCD cameras has been discussed previously.[2-6] Other relevant works may be found in the references of these papers. However, these papers deal either with the error in depth measurement[4,5] or with locating the three-dimensional position of a point,[2] or with navigation problems that involve both of these.[6] The analysis in these and other papers is either restricted to depth measurement, which does not concern us in this application, or is overly general and does not take into account the knowledge we have regarding the constraints on the positions of the targets that arise from the dynamics of the plate. For our purposes, we need to consider only those aspects of the error that relate to the problem at hand.

Frames of Reference and Problem Geometry

The camera arrangement is illustrated in Fig. 1, where the image planes of each camera are considered to be identical and arranged to be coplanar with their edges parallel. Each image plane has dimensions $a \times b$ with n_v pixels in the y_c direction and n_u pixels in the x_c direction. These should not be viewed as parameters that may be continuously varied as their values are restricted by the available technology and are often specified a

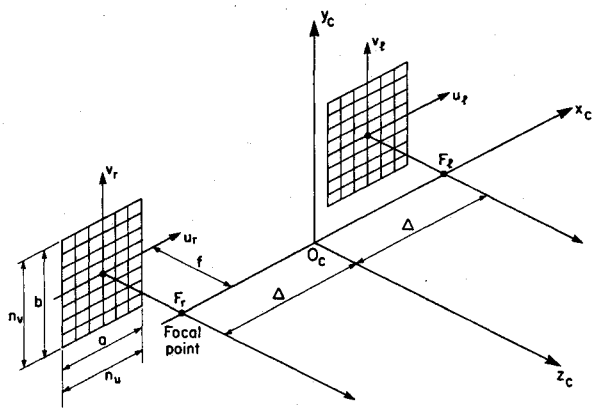

Fig. 1 Camera coordinate frame geometry.

priori. Note that the center of the camera screen is not necessarily at the origin of the image coordinate frame.

Let (x_c, y_c, z_c) be the position of a point object in the camera coordinate system, and let (u_l, v_l) and (u_r, v_r) be the positions of its images in the left and right cameras, respectively. Then, the transformation between these coordinates is

$$u_l = \frac{f(\Delta - x_c)}{z_c}, \qquad v_l = \frac{-fy_c}{z_c}$$
$$u_r = \frac{-f(\Delta + x_c)}{z_c}, \qquad v_r = \frac{-fy_c}{z_c} \tag{1}$$

Solving for (x_c, y_c, z_c) yields

$$x_c = \frac{\Delta(u_r + u_l)}{(u_r - u_l)}, \qquad y_c = \frac{2\Delta v_r}{(u_r - u_l)}, \qquad z_c = \frac{-2\Delta f}{(u_r - u_l)} \tag{2}$$

We also introduce a global reference frame (X, Y, Z) which is oriented such that the X-Y plane coincides with the reference plane of the plate and the position of the cameras is given relative to this frame. The horizontal camera configuration is illustrated in Fig. 2a. The vertical configuration is similar, but the camera frame is rotated 90 deg about the z_c axis (see Fig. 2b). To relate the position of a point as expressed in the global frame to its position as expressed in the camera frame, the usual trigonometric transformation matrices may be used.[7]

When a set of point targets is observed by both cameras, the correspondence between points in the two different images has to be determined. From Eqs. (1) and (2), it can be seen that the left and right images of a point are always on the same horizontal line. This can therefore be used as a criterion for matching. The only case where this does not work is when two or more objects have sufficiently close y_c/z_c values so that the images fall on the same horizontal row of pixels. In practice, this is an extremely unlikely event since the number of point targets is usually much smaller than the number of rows of pixels. Moreover, even if this does occur, because the range of motion of the targets is constrained by the plate dynamics and because the original placement of the targets is known, there is sufficient information to rule out fictitious objects from erroneous matches.

Error Analysis

Image quantization error occurs because the position of the image is not known exactly but only up to the pixel on which the image falls. The objective of this section is to obtain bounds and an expected value of this error for a single point target, as a function of its position (X_0, Y_0) on the plate but independent of its deflection at a particular time. The analysis will be first carried out for the vertical configuration.

Define an integer coordinate system on the camera screen where each point (i,j) corresponds to the center of a pixel. Then let the object image on the left and right cameras' screens be designated by (I_l, J) and (I_r, J), respectively, noting that their ordinates are always equal. The actual image positions (u_l, v_l) and (u_r, v_r), $(v_r = v_l = v)$, lie somewhere in the corresponding pixels but their exact values are unknown. Then, with reference to Eqs. (2), the position of the objects in the camera frame must lie in the set:

$$U = \left\{ \left[\frac{(u_r + u_l)\Delta}{(u_r - u_l)}, \frac{2v\Delta}{(u_r - u_l)}, \frac{-2f\Delta}{(u_r - u_l)} \right] : \begin{array}{l} (a/n_u)(I_r - \tfrac{1}{2}) < u_r < (a/n_u)(I_r + \tfrac{1}{2}) \\ (a/n_u)(I_l - \tfrac{1}{2}) < u_l < (a/n_u)(I_l + \tfrac{1}{2}) \\ (a/n_u)(J - \tfrac{1}{2}) < v < (a/n_u)(J + \tfrac{1}{2}) \end{array} \right\} \tag{3}$$

where (a/n_u) is the width of one pixel and U shall be referred to as the region of uncertainty.

The uncertainty in the position of the target can be further reduced by noting that the target is constrained to move along a locus $\mathcal{L}$, which is a straight line perpendicular to the reference plane, as the plate vibrates. For each target, there is one such locus. Assuming the size of the satellite is much larger than that of the regions of uncertainty, the targets are sufficiently far apart so that the region of uncertainty U of any one target can only intersect with its own locus. This allows identification of an observed target from its region of uncertainty, thus obtaining its exact locus. Once the target has been identified, the only measurement that has to be made is the position of the target on the locus, i.e., its current deflection. The error interval is therefore reduced to the intersection of the locus $\mathcal{L}$ with U, which is only one dimensional. The length of this interval measures the worst case error.

If (X_0, Y_0) is the position of the target in the reference plane, the locus of the target as expressed in the camera coordinate system is[7]

$$X_0 = z_c \cos\theta + x_c \sin\theta, \qquad Y_0 = -y_c \tag{4}$$

where θ is the angle of inclination of the cameras (see Fig. 1). Let P_1 and P_2 be the two points of intersection of $\mathcal{L}$ with U (i.e., the endpoints of the error interval). If $u_{r1}, u_{\ell 1}, u_{r2}, u_{\ell 2}$ are the abscissas of the images of P_1 and P_2 in the left and right cameras, respectively, the length of the error interval can be easily calculated as[7]

$$E_v = 2f|\csc\theta|\Delta \left| \frac{1}{u_{r2} - u_{\ell 2}} - \frac{1}{u_{r1} - u_{\ell 1}} \right| \tag{5}$$

Now suppose the true position of the target is (X_0, Y_0, Z) in the global frame, and u_L and u_R are the abscissas of the true image of the target in the left and right cameras, respectively. We can write

$$u_{r1} - u_{\ell 1} = u_R - u_L + \delta_1, \qquad u_{r2} - u_{\ell 2} = u_R - u_L + \delta_2 \tag{6}$$

where δ_1 and δ_2 are small, less than the width of one pixel, compared to $u_R - u_L$. Substituting Eq. (6) into Eq. (5) and neglecting second-order and higher terms in δ_1 and δ_2, the length of the error interval is

$$E_v \approx \frac{2f|\csc\theta|\Delta}{(u_R - u_L)^2}|\delta_2 - \delta_1| \tag{7}$$

By utilizing the image to camera frames transformation [Eqs. (1)] and the camera to global frames transformation [Eqs. (6)], the length of the error interval can be expressed in terms of global coordinates of the target[7]

$$E_v = \frac{[X_0 \cos\theta - (Z - h)\sin\theta]^2}{f} \left| \frac{u_{\ell 1} - u_{\ell 2}}{X_0 - \Delta \sin\theta} \right| \tag{8}$$

The range of view of the cameras is fixed by requirements that can be determined by finding bounds on the deflections of the parts of the satellite that are within the range of view of the cameras. With reference to Fig. 3, the range of view can be defined by the angles α and β, and for a given range of view, these angles are assumed to be fixed independently of the configuration variables. Although Fig. 3, for the sake of clarity, shows the horizontal configuration, there is no difficulty in arranging both cameras in the vertical configuration to have the same α and β values. Given a screen of fixed size, the range of view should always be projected onto the entire screen in order to use its full resolution. Hence, corresponding to different values of the configuration variable θ, the focal length f should be adjusted in order to maintain maximum resolution.

The relationship between f and θ is found to be

$$f(\theta) = \frac{a}{\tan(\theta - \alpha) + \tan(\beta - \theta)} \tag{9}$$

Using Eq. (9) in the previous expression for E_v yields

$$E_v = \frac{[X_0 \cos\theta - (Z-H)\sin\theta]^2}{a} \times [\tan(\theta - \alpha) + \tan(\beta - \theta)] \left| \frac{u_{\ell 1} - u_{\ell 2}}{X_0 - \Delta \sin\theta} \right| \tag{10}$$

E_v is the length of the error interval when the target is at a specific deflected position. To get an upper bound of this quantity for all deflected positions of the target, we first note that, since P_1 and P_2 lie in the region of uncertainty U, their images in any one of the cameras must be in the same pixel. Hence, it is deduced that $|u_{\ell 1} - u_{\ell 2}| \leq a/n_u$ (the width of a pixel). If one assumes a known lower bound $Z_L(X, Y)$ for the deflected shape $w(X, Y)$ of the plate and also that the camera is always above the plate (necessary to avoid occlusion), an upper bound on E_v can be readily deduced as[7]

$$E_v \leq \frac{\{X_0 \cos\theta - [Z_L(X_0, Y_0) - h]\sin\theta\}^2}{n_u |X_0 - \Delta \sin\theta|} \times [\tan(\theta - \alpha) + \tan(\beta - \theta)] \tag{11}$$

This should be a reasonably tight bound because the magnitude of the deflection is expected to be small, especially in relation to the distance from the camera. Consequently, the amount of inaccuracy that would arise from replacing Z with $Z_L(X_0, Y_0)$ is small.

Another useful measure is the expected length of the error interval over time. This is a more difficult estimate to obtain

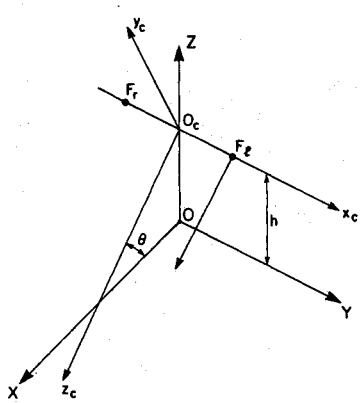

Fig. 2a Horizontal configuration.

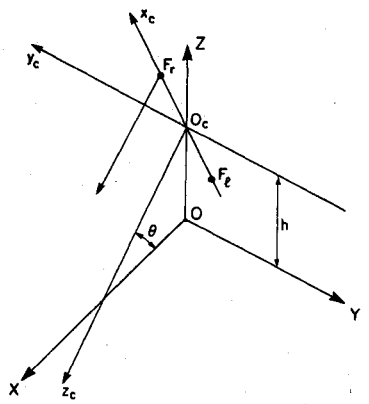

Fig. 2b Vertical configuration.

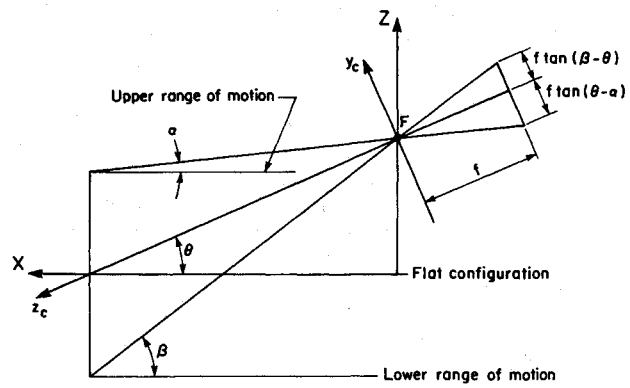

Fig. 3 Definitions of angles α and β for the horizontal configuration.

for a general configuration because the error is a function of the deflection w which in turn depends on the forces acting on the satellite, many of which cannot be known a priori. However, the expected error can be estimated quite reliably for the special case $\theta = 0$. It should be noted that finding the optimal choice of angle θ must take into consideration the fact that at $\theta \approx 0$ deg distinguishing the target trajectories may be very difficult and at $\theta \approx 90$ deg the trajectories will be unobservable. The problem of distinguishing the trajectories can be partially dealt with by suitable placement of the targets, which have known locations.

For the case where $\theta = 0$, the locus of the target in camera coordinates, Eq. (5), is

$$z_c = X_0, \qquad y_c = -Y_0 \tag{12}$$

The locus $\mathcal{L}$ of the target can be viewed as being partitioned, by the regions of uncertainty, into error intervals. This is because for each point on the locus there is exactly one region of uncertainty in which it lies, and the intersection of the locus with this region of uncertainty gives the unique error interval that contains the point. To find the expected value of E_v, it is necessary to consider the distribution of the lengths of the error intervals on $\mathcal{L}$.

The partition $\mathcal{L}$ into error intervals can be described by

$$\mathcal{P} = \{(C_1, -Y_0, X_0), (C_2, -Y_0, X_0), \ldots, (C_n, -Y_0, X_0)\} \tag{13}$$

where each point $(C_i, -Y_0, X_0)$ lies on $\mathcal{L}$ [in view of Eqs. (12)] and they are all the endpoints of the error intervals on $\mathcal{L}$ in the range of motion of the target.

Each of these points lies on a boundary of some region of uncertainty; hence, its image on either the right or the left camera lies on a pixel's boundary. By applying Eqs. (1), the equation of $\mathcal{L}$ in image coordinates may be shown to be

$$v_\ell = v_r = \frac{fY_0}{X_0} \tag{14}$$

so that it is parallel to the u axis. Hence, the image of a point in the set $\mathcal{P}$ is an intersection of the image of $\mathcal{L}$ with a pixel boundary that is parallel to the v axis in either the right or left camera. Conversely, for each such intersection, its corresponding point on $\mathcal{L}$ must belong to $\mathcal{P}$. This is because, if the target on $\mathcal{L}$ moves across such a point, its image on either the left or the right camera changes by one pixel so that the target moves into a different region of uncertainty and, hence, a different error interval. For the left camera, let these intersections be

$$\left\{\left(\frac{k_0 a}{n_u}, \frac{fY_0}{X_0}\right), \left[\frac{(k_0+1)a}{n_u}, \frac{fY_0}{X_0}\right], \ldots, \left(\frac{k_1 a}{n_u}, \frac{fY_0}{X_0}\right)\right\} \tag{15}$$

and similarly for the right camera

$$\left\{\left(\frac{m_0 a}{n_u}, \frac{fY_0}{X_0}\right), \left[\frac{(m_0+1)a}{n_u}, \frac{fY_0}{X_0}\right], \ldots, \left(\frac{m_1 a}{n_u}, \frac{fY_0}{X_0}\right)\right\} \tag{16}$$

where k_0, k_1, m_0, and m_1 are integers such that $k_0 < k_1$ and $m_0 < m_1$.

Using Eqs. (1), the points on $\mathcal{L}$ whose images are these points are

$$\begin{aligned}\mathcal{P}_\ell &= \left\{\left[\Delta - k_0\left(\frac{X_0 a}{fn_u}\right), -Y_0, X_0\right], \ldots, \right. \\ &\quad \left.\left[\Delta - k_1\left(\frac{X_0 a}{fn_u}\right), -Y_0, X_0\right]\right\} \\ \mathcal{P}_r &= \left\{\left[-\Delta - m_0\left(\frac{X_0 a}{fn_u}\right), -Y_0, X_0\right], \ldots, \right. \\ &\quad \left.\left[-\Delta - m_1\left(\frac{X_0 a}{fn_u}\right), -Y_0, X_0\right]\right\}\end{aligned} \tag{17}$$

with $\mathcal{P} = \mathcal{P}_\ell \cup \mathcal{P}_r$.

By noting that the separation between adjacent points in both sets is the same (i.e., $X_0 a/fn_u$), it can be concluded that consecutive points in $\mathcal{P}$ must be from different sets. Hence, the error intervals are of only two distinct lengths, say d, and $(X_0 a/fn_u) - d$ (see Fig. 4).

To find d, it is observed that the separation of a point in $\mathcal{P}_\ell$ from one in $\mathcal{P}_r$ is given by

$$\left|2\Delta - (k-m)\frac{X_0 a}{fn_u}\right| \tag{18}$$

where k and m are integers such that $k_0 \le k \le k_1$ and $m_0 \le m \le m_1$.

The lengths d and $(X_0 a/fn_u) - d$ are the two smallest values of Eq. (18) as k and m run through their respective ranges. The two smallest values of Eq. (18) are[7]

$$\frac{X_0 a}{fn_u}\,\mathrm{frac}\left(\frac{2fn_u\Delta}{X_0 a}\right), \qquad \frac{X_0 a}{fn_u}\left[1 - \mathrm{frac}\left(\frac{2fn_u\Delta}{X_0 a}\right)\right] \tag{19}$$

where frac($\cdot$) denotes the fractional part of the real argument. Thus, the consecutive error intervals are always of different lengths, and if the lengths of these error intervals are assumed to be small compared to the range of motion of the target, the probability of the target being in one kind of error interval can be estimated to be proportional to the length of that kind of interval. Hence,

$$\begin{aligned}\mathrm{Prob}\left(E_v = \mathfrak{F}\frac{X_0 a}{fn_u}\right) &= \mathfrak{F} \\ \mathrm{Prob}\left[E_v = (1-\mathfrak{F})\frac{X_0 a}{fn_u}\right] &= 1 - \mathfrak{F}\end{aligned} \tag{20}$$

where $\mathfrak{F} = \mathrm{frac}(2fn_u\Delta/X_0 a)$.

For reasons to be discussed in the next section, we are interested in the root mean square (rms) error for a measure of its average. This is given by

$$(E_v)_{\mathrm{rms}} = \sqrt{1 - 3\mathfrak{F} + 3\mathfrak{F}^2}\cdot\frac{X_0 a}{fn_u} \tag{21}$$

where f may be expressed in terms of α and β by employing Eq. (9) such that

$$(E_v)_{\mathrm{rms}} = \sqrt{1 - 3\mathfrak{F} + 3\mathfrak{F}^2}\cdot\frac{X_0}{n_u}(\tan\beta - \tan\alpha) \tag{22}$$

In some cases, this may not be a reliable estimate. For if $2\Delta fn_u/X_0 a$ is large, then $\mathfrak{F} = \mathrm{frac}(2\Delta fn_u/X_0 a)$ is very sensitive to the slight inaccuracies in the value of X_0 (the parameter value that is hardest to estimate precisely). In such situations, a more reliable error estimate can be obtained by modeling $\mathfrak{F}$ as a random variable uniformly distributed in [0,1]. In which case,

$$\begin{aligned}(E_v)_{\mathrm{rms}} &= \sqrt{\int_0^1 (1 - 3\mathfrak{F} + 3\mathfrak{F}^2)\,d\mathfrak{F}}\;\frac{X_0(\tan\beta - \tan\alpha)}{n_u} \\ &= \frac{X_0(\tan\beta - \tan\alpha)}{\sqrt{2}n_u}\end{aligned} \tag{23}$$

The corresponding analysis for the horizontal configuration is similar and only the results will be reported here. The length of the error interval for the horizontal configuration may be found to be[7]

$$\begin{aligned}E_h &= \frac{[X_0\cos\theta - (Z-h)\sin\theta]^2}{bX_0} \\ &\quad \times [\tan(\theta-\alpha) + \tan(\beta-\theta)]\,|v_1 - v_2|\end{aligned} \tag{24}$$

where v_1 and v_2 are the ordinates of the error interval endpoints and are camera independent.

In a similar manner to the vertical case, the upper bound on the error is given by[7]

$$\begin{aligned}E_h &\le \frac{\{X_0\cos\theta - [Z_L(X_0, Y_0) - h]\sin\theta\}^2}{n_v X_0} \\ &\quad \times [\tan(\theta-\alpha) - \tan(\beta-\theta)]\end{aligned} \tag{25}$$

For the special case of $\theta = 0$, the length of the error interval is

$$E_h\,|_{\theta=0} = \frac{X_0}{b}(\tan\beta - \tan\alpha)\,|v_1 - v_2| \tag{26}$$

In the horizontal configuration with $\theta = 0$, $|v_1 - v_2|$ is a constant independent of z with a width of one pixel such that

$$|v_1 - v_2| = \frac{b}{n_v} \tag{27}$$

which, when used in Eq. (26), yields the expression for the error that is constant and equal to the rms error. Therefore,[7]

$$(E_h)_{\mathrm{rms}} = \frac{X_0}{n_v}(\tan\beta - \tan\alpha) \tag{28}$$

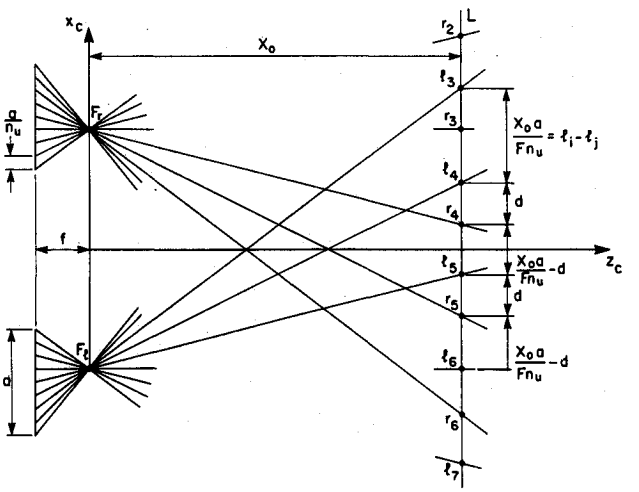

Fig. 4 Error intervals.

The upper bounds on the length of the error interval [Eqs. (11) and (25)] and the expected values [Eqs. (22), (23), and (28)] are the principal results of this section. An important question that arises is, "What implication do these error expressions have for the angle of view of the cameras?" This is important because the larger the area covered, the fewer pairs of cameras we need to view the entire plate.

The area to be viewed is spanned by the X and Y coordinates of the global frame, where the X coordinate corresponds to a depth of view and the Y coordinate corresponds to a breadth of view.

The depth of view depends on the angles α and β and, because each of the error expressions is proportional to $\tan(\theta - \alpha) + \tan(\beta - \theta)$, the error measures and the depth of view are dependent on one another. It may be readily observed that an increase in the depth of view will translate into an increase in the error.

The breadth of view is given by the angle of view; in the vertical configuration, this is given by $2\arctan(b/2f)$, and in the horizontal case, by $2\arctan(a/2f)$. In both the vertical and the horizontal configurations, the value of the error depends on the focal length f. Thus, the screen dimensions b or a may be increased as large as possible in order to increase the field of view without affecting the measurement accuracy. The only restriction is that the resulting error in measuring the X and Y position of the target should not be so large so that there is a risk of misidentifying targets. This is a loose restriction because the large area of the satellite means that the targets can be placed far apart.

Shape Estimation

Having established error measures for the deflection of a single point, we will now consider how to use deflection measurements for n such points to estimate the shape of the plate. This problem is different from the usual surface reconstruction problems[8-11] in that we have a need for an interpolation that reproduces the deformed shape of the plate with as much accuracy and fidelity as possible. The interpolation process must also consider the fact that the data upon which it is based have implicit errors due to the quantization error that arises from the CCD cameras. This is a consideration that has not been dealt with in other surface reconstruction methods.[8-11] Nor can we do a curve fit that ignores the structural and dynamic realities of the problem at hand. But by choosing, as interpolating functions, the natural model shapes of the plate, we will implicitly choose a basis that satisfies the governing Euler-Lagrange equations that arise from the application of Hamilton's (variational) principle to the structure of interest. It is assumed that the mode shapes are available. In the case of plates, mode shapes for a wide range of shapes and boundary conditions are available in Leissa.[12] For actual satellites, the structural identification methods needed to determine the mode shapes are beyond the scope of this paper, but certainly erroneous mode shapes could have a serious effect on the following.

Given the spatial coordinates of a finite number of points on the satellite, the overall shape has to be determined by interpolation. In this paper, the interpolating basis functions will be the free vibration mode shapes for the satellite, which in this case is assumed to be a thin rectangular plate. Although, mathematically, the shape of the plate can only be described by a linear combination of an infinite number of mode shapes, practically only a finite number of mode shapes need to be considered in order to satisfactorily approximate the shape. It is assumed that the shape approximation error incurred by choosing a finite set of mode shapes $\{\phi_1, \phi_2, \ldots, \phi_n\}$ is negligible compared to the quantization error discussed earlier.

Interpolation Through Erroneous Data

If at specific point (X_i, Y_i), $i = 1, 2, \ldots, m$, the deflection $w(X_i, Y_i)$ can be measured exactly, then the modal coordinates q_j, $j = 1, 2, \ldots, n$, can be determined exactly by solving the system of m equations

$$\sum_{j=1}^{n} q_j \phi_j(X_i, Y_i) = w(X_i, Y_i), \qquad i = 1, 2, \ldots, m \tag{29}$$

or in matrix vector notation, $\Phi q = w$, where $\Phi_{ij} = \phi_j(X_i, Y_i)$. But as shown earlier, w cannot be determined exactly but only within a known interval. Let the interval for the ith entry be

$$\ell_i(X_i, Y_i, w_i) \le w_i \le u_i(X_i, Y_i, w_i), \qquad i = 1, 2, \ldots, m \tag{30}$$

or

$$\ell(X, Y, w) \le w \le u(X, Y, w) \tag{31}$$

Then, the problem is to find q subject to $\ell \le \Phi q \le u$, where the inequality of vectors is taken to mean that a vector is less than another vector if and only if each component of the one is less than the corresponding component of the other.

The vector q is not unique and, therefore, an estimate q^* for q that will minimize some error criterion is sought. Expressing the error associated with q^* in terms of only the configuration parameters will provide a criterion for selecting the best configuration. Thus, the remaining tasks are to develop an error criterion for defining q^*, find an explicit way to calculate q^* given the error criterion, and find an expression for the error associated with q^* independent of the deformation.

Error Criteria for Parameter Estimation

The coefficient vector q is confined to a closed space Q in R^n and initially we require a suitable norm for Q. The L^2 norm is used in order that the shape can be approximated from the information supplied by the finite number of targets/sensors:

$$\|q\| = \sqrt{\int_A \left[\sum_{i=1}^{n} q_i \phi_i(X, Y)\right]^2 dX\, dY} \tag{32}$$

Recalling that the mode shapes are orthogonal on A, it may be shown that

$$\|q\| = \sqrt{\sum_{i=1}^{n} q_i^2} = \|q\|_2 \tag{33}$$

which is the L^2 norm of q.

Using this norm, two error criteria are considered: a worst case and an average case. Given an estimate q_0, the worst case measure is defined as

$$\mathcal{E}_w = \max_{q \in Q} \|q - q_0\| \tag{34}$$

And for the average case, an rms measure is used such that

$$\mathcal{E}_{\text{rms}} = \sqrt{\underset{q \in Q}{\text{expected}} \|q - q_0\|^2} \tag{35}$$

The procedure used to find q^* for each of these measurements depends on whether the number of targets (m) is the same or larger than the number of mode shapes.

Equal Number of Targets and Modes ($n = m$)

For this case, Φ is $n \times n$ and is assumed to be invertible. The space Q is, geometrically, an n-dimensional parallelepiped bounded by n pair of parallel hyperplanes, one pair for each of the n inequalities (e.g., when $n = 2$, Q is a parallelogram).

Considering the maximum error criterion first, the following intuitive result is useful.

Result I: For any point $q_0 \in R^n$, the point in Q farthest from q_0 must be in $V(Q)$, the set of vertices of Q.

Proof: See Tse and Heppler.[7]

Hence, the maximum error using a vector q_0 as the estimate can be written as

$$\mathcal{E}_w(q_0) = \max_{q \in V(Q)} \|q - q_0\|$$

By symmetry, it is intuitively clear that the estimate that minimizes the maximum error should be the midpoint between two opposite vertices in the parallelepiped Q. Thus, we have the following result.

Result II: The optimal estimate q^* for the maximum error criterion is the centroid of Q, i.e.,

$$q^* = \frac{1}{2}\Phi^{-1}\begin{bmatrix} \ell_1 + u_1 \\ \vdots \\ \ell_n + u_n \end{bmatrix}$$

Proof: See Tse and Heppler.[7]

The worst case error associated with using q^* as the estimate is given by

$$\mathcal{E}_w(q^*) = \max_{q \in V(Q)} \left\| \Phi^{-1}\left(\frac{\ell_1 + u_1}{2}, \ldots, \frac{\ell_n + u_n}{2}\right)^T - v \right\|$$

$$= \max_{q \in \{-1,1\}} \|\Phi^{-1}(\mu_1 e_1, \ldots, \mu_n e_n)^T\| \qquad (36)$$

where $e_i = (u_i - \ell_i)/2$.

To directly compute Eq. (36), the error for each element of $V(Q)$ (corresponding to a specific choice of $\mu_1, \mu_2, \ldots, \mu_n$) would have to be computed and compared to the others in order to find the largest. There are 2^n vertices, which for large n makes this procedure infeasible. Two computationally more efficient alternatives are presented in the following.

Recalling that $\Phi^{-1} = (b_{ij})$, Eqs. (36) become

$$\mathcal{E}_w(c) = \max_{\mu_i \in \{-1,1\}} \sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} b_{ij}\mu_j e_j\right)^2}$$

$$\leq \sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} |b_{ij}|e_j\right)^2} \qquad (37)$$

This estimate is a function of X_i, Y_i, and w_i ($i = 1, 2, \ldots, n$) because e_i depends on w_i. In order to use Eq. (37) to assess the configuration, the dependence on w_i must be removed. This is accomplished by using the upper bounds for each e_i as given in inequalities (11) and (25) for the vertical and horizontal configurations, respectively. Now suppose that the calculated upper bounds are such that

$$e_i(X_i, Y_i, w_i) \leq U_i(X_i, Y_i), \qquad i = 1, 2, \ldots, n \qquad (38)$$

and the resulting upper bound on $\mathcal{E}_w(c)$ is

$$\mathcal{E}_w(c) \leq \sqrt{\sum_{i=1}^{n}\left(\sum_{j=1}^{n} |b_{ij}|U_j\right)^2} \qquad (39)$$

As a more accurate but computationally more expensive alternative, consider the following. First, define Q_0 to be the transformed version of Q that results from translating the centroid of Q to the origin of R^n. Q_0 is given by

$$Q_0 = \{q : -e \leq \Phi q \leq e\} \qquad (40)$$

where $e = (e_1, e_2, \ldots, e_n)^T$. Now Q_0 depends only on e for a fixed Φ and so it will be denoted $Q_0(e)$. The size of Q_0 is totally ordered by e [i.e., $Q_0(e_1)$ is a subset of $Q_0(e_2)$ if $e_1 \leq e_2$]. Hence, the largest shape error $\mathcal{E}_w(c)$ increases monotonically with e. Now each e_i is bounded above by U_i and so it follows that

$$\mathcal{E}_w(c) \leq \max_{\mu_i \in \{-1,1\}} \|\Phi^{-1}(\mu_i U_i, \ldots, \mu_n U_n)^T\| \qquad (41)$$

This gives a bound on $\mathcal{E}_w(c)$ that is independent of the deflections w_i. To compute it, we can apply the modified simplex method[13] on the following quadratic program with linear constraints: maximize $\|q\|^2$ subject to $q \in Q$ (a convex polytope).

Having dealt with the worst case, now consider the average case where the rms error measure is used. By definition, the rms error of using an estimate q^* is

$$\mathcal{E}_{\text{rms}}(q^*) = \sqrt{\int_Q \|q^* - q\|^2 f(q)\, dq_1 \ldots dq_n} \qquad (42)$$

where $f(q)$ is the probability density function of q over Q. Because there is a one-to-one correspondence between points in Q and points in W, we can integrate over W instead. Thus,

$$\mathcal{E}_{\text{rms}}(q^*) = \sqrt{\int_W \|q^* - \Phi^{-1}w\|^2 g(w)\, dw_1 \ldots dw_n} \qquad (43)$$

where $g(w)$ is the probability density function of w over W. Without additional information, there is no reason to assume that w is more likely to be in one subset of W or another. Hence, a uniform distribution of w over W is assumed and

$$\mathcal{E}^2_{\text{rms}}(q^*) = \frac{1}{|W|}\int_W \|q^* - \Phi^{-1}w\|^2\, dw_1 \ldots dw_n \qquad (44)$$

where $|W|$ is the volume of W.

Result III: The centroid c of W or Q is the optimal estimate for the rms error criterion under the uniform distribution assumption.

Proof: See Tse and Heppler.[7]

In order to evaluate the rms error associated wtih the centroid c, begin by representing Φ^{-1} in terms of the column vectors $[b_1, b_2, \ldots, b_n] = \Phi^{-1}$. Noting that

$$c = \frac{1}{2}\Phi^{-1}\begin{bmatrix} \ell_1 + u_1 \\ \vdots \\ \ell_n + u_n \end{bmatrix}$$

it follows that

$$\mathcal{E}^2_{\text{rms}}(c) = \frac{1}{|W|}\int_W \left\| \sum_{i=1}^{n}\left(\frac{\ell_i + u_i}{2} - w_i\right) b_i \right\|^2 dw_1 \ldots dw_n \qquad (45)$$

which may be expressed as[7]

$$\mathcal{E}^2_{\text{rms}}(c) = \frac{1}{3}\sum_{i=1}^{n} e_i^2 \|b_i\|^2 \qquad (46)$$

where it may be recalled that the e_i is the semilength of the error interval associated with measurement w_i. Because of this association, the expected error is implicitly a function of the deflections w_i. In order to find the optimal rms error as a function of only the configuration variables, it is necessary to find the expected value of the error intervals e_i over time so that it is independent of the deflections. As in the previous section, a reliable estimate can only be found when the angle of inclination θ of the cameras to the horizontal is zero. For this particular case, we have derived the expected rms values in Eqs. (22) and (28) for the vertical and horizontal configurations, respectively. The rms error for the shape configuration can be expressed as

$$\mathcal{E}_{\text{rms}} = \frac{1}{\sqrt{3}}\|\Phi^{-1}A\|_F \qquad (47)$$

where

$$A \triangleq \begin{bmatrix} \sqrt{e_1^2} & & 0 \\ & \ddots & \\ 0 & & \sqrt{e_n^2} \end{bmatrix} \tag{48}$$

and the subscript F indicates the Frobenius matrix norm.

More Targets Than Modes ($m > n$)

Unlike the case when $m = n$, the region of uncertainty for q is not a parallelepiped. In general, it is an irregular convex polytope. Since each subset of n constraints bounds a parallelepiped, the polytope can be viewed as the intersection of $\binom{m}{n}$ such parallelepipeds. As a result, the vertices of the region are extremely difficult to compute, unlike the straightforward formulas for the case $m = n$. In view of the complicated geometry and lack of symmetry, finding an optimal estimate for both error criteria is much more difficult than in the $m = n$ case.

We, therefore, settle for a suboptimal estimate for q. Let $\hat{w} = (\ell + u)/2$ be the vector of approximation of the individual target positions. The estimate for q is taken to be the least square solution q_{LS} of the overdetermined system: $\Phi q = \hat{w}$, i.e.,

$$q_{LS} = (\Phi^T\Phi)^{-1}\Phi^T\hat{w} \tag{49}$$

No strong theoretical justifications exist for using this estimate compared to others. However, it is a practical estimate because, as a result of the coefficient matrix Φ being only a function of the positions where the targets are placed on the satellite, its pseudoinverse $(\Phi^T\Phi)^{-1}\Phi^T$ can be precalculated. Hence, the least squares estimate can be calculated quite efficiently by a matrix-vector multiplication, as in the $n = m$ case. Moreover, the rms error associated with q_{LS} can be explicitly calculated, as will be shown presently.

Suppose $\hat{w}_T$ is the vector of true deflections. Then the true parameter values q_T is given by

$$q_T = (\Phi^T\Phi)^{-1}\Phi^T\hat{w}_T \tag{50}$$

Because the shape of the uncertainty region Q is very much dependent on the particular measurements taken (ℓ and u), we shall not attempt to find $\mathcal{E}_{rms}(q_{LS})$, the rms error given a fixed Q(i.e., a fixed ℓ and u), as we did for the $n = m$ case Eq. (46). Because our main objective is to find the rms error independent of the instantaneous deflections, we will instead find the expected value of the squared error $\|q_{LS} - q_T\|^2$ for a fixed e [measurement error interval vector, $(u - \ell)/2$]. Since for a fixed e there are many possible different shapes for Q, we are finding the expected value over all of the shapes. Once we obtain this expected value, we can make use of the distribution of e to find the overall expected value independent of the deflections.

Define $\hat{\xi} \triangleq \hat{w} - \hat{w}_T$ to be the vector measuring the deviation of the true deflections from the midpoint of their corresponding error intervals. For a given $e = (e_1, e_2, \ldots, e_m)^T$, define $C(e)$ as the rectangular box in R^m centered at the origin and whose sides are of lengths $2e_i$, $i = 1, 2, \ldots, m$. Since each component of $\hat{w}_T$ must lie in its error interval, $\hat{\xi}$ must lie inside $C(e)$. If the range of the deflections and distance between targets are large compares to the error intervals, it is safe to assume that the components $\hat{\xi}_i$ of $\hat{\xi}$ are independent and uniformly distributed in $[-e_i, e_i]$.

Now, from Eqs. (47) and (43),

$$\|q_{LS} - q_T\| = \|(\Phi^T\Phi)^{-1}(\hat{w} - \hat{w}_T)\| = \|(\Phi^T\Phi)^{-1}\hat{\xi}\| \tag{51}$$

so the rms value of the error for a fixed e can be obtained by integrating over C:

$$\mathcal{E}^2_{rms}(e) = \frac{1}{|C|}\int_C \|(\Phi^T\Phi)^{-1}\Phi^T\hat{\xi}\|^2 \, d\hat{z}_1 \ldots d\hat{z}_m \tag{52}$$

Letting $\hat{b}_i$ be the columns of $(\Phi^T\Phi)^{-1}$ we can obtain[7]:

$$\mathcal{E}^2_{rms}(e) = \frac{1}{3}\sum_{i=1}^{n} e_i^2 \|\hat{b}_i\|^2 \tag{53}$$

Finally, to obtain the expected value independent of the deflections:

$$\mathcal{E}_{rms} = \frac{1}{\sqrt{3}}\|(\Phi^T\Phi)^{-1}\Phi^T A\|_F \tag{54}$$

where A is as defined in Eq. (47).

Summary and Conclusions

Stereo vision is a viable means of determining the real time shape of a flexible structure, but the very nature of CCD cameras leads to special considerations that must be addressed in designing and implementing the system.

In the paper, we have treated two limiting camera configurations, horizontal and vertical, and have determined explicit expressions for the size of the error interval for both of these cases. Upper bounds corresponding to these error intervals have also been found. These expressions confirm what would appear to be intuitively correct, that the error can be reduced by increasing the focal length of the cameras and by decreasing the size of the individual pixels; the error varies linearly with each of the parameters.

Given that the deflections of the targets as determined by the system under consideration will be in error, as would the deflections determined by any other means, the issue of interpolating the shape of the structure through the erroneous data has been addressed. Two criteria for evaluating the interpolation have been presented, a worst case criterion and an average case criterion, both based on the L^2 norm. The associated upper bounds have also been reported.

References

[1]Deitz, P. R. W., "Determination of the Efficacy of Using Accelerometers as Sensors for Third Generation Satellites," Univ. of Toronto Inst. for Aerospace Studies, Toronto, Ontario, Canada, TN 262, Dec. 1986.

[2]Blostein, S. D., and Huang, T. S., "Error Analysis in Stereo Determination of 3-D Point Positions," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 9, No. 6, Nov. 1987, pp. 752-765.

[3]Rodriguez, J. J., and Aggarwal, J. K., "Quantization Error in Stereo Imaging," *Proceedings of the Computer Science Conference on Computer Vision and Pattern Recognition,* Inst. of Electrical and Electronics Engineers, New York, June 1988, pp. 153-158.

[4]McVey, E. S., and Lee, J. W., "Some Accuracy and Resolution Aspects of Computer Vision Distance Measurements," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 4, No. 6, Nov. 1982, pp. 646-649.

[5]Verri, A., and Torre, V., "Absolute Depth Estimate in Stereopsis," *Journal of the Optical Society of America, A,* Vol. 3, No. 3, 1986, pp. 297-299.

[6]Matthies, L., and Shafer, S. A., "Error Modeling in Stereo Navigation," *IEEE Journal on Robotics and Automation,* Vol. 3, No. 3, June 1987, pp. 239-248.

[7]Tse, D. N. C., and Heppler, G. R., "Shape Identification for Flexible Structures via Stereo Vision," Dept. of Systems Design Engi-

neering, Univ. of Waterloo, TR 168-PR041089, Waterloo, Ontario, Canada, Aug. 1989.

[8]Grimson, W. E. L., "An Implementation of a Computational Theory of Visual Surface Interpolation," *Computer Vision, Graphics, and Image Processing,* Vol. 22, No. 1, 1983, pp. 39–69.

[9]Brady, M., and Horn, B. K. P., "Rotationally Symmetric Operators for Surface Interpolation," *Computer Vision, Graphics, and Image Processing,* Vol. 22, No. 1, 1983, pp. 70–94.

[10]Terzopoulos, D., "Multilevel Computational Processes for Visual Surface Reconstruction," *Computer Vision, Graphics, and Image Processing,* Vol. 24, No. 1, 1983, pp. 52–96.

[11]Terzopoulos, D., "Regularization of Inverse Visual Problems Involving Discontinuities," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 8, No. 4, 1986, pp. 413–424.

[12]Leissa, A. W., "Vibration of Plates," NASA SP-160, 1969.

[13]Bradley, S. P., Hax, A., and Magnanti, T., *Applied Mathematical Programming,* Addison-Wesley, Reading, MA, 1977.

Earl A. Thornton
Associate Editor