Motion and Structure From Point Correspondences with Error Estimation: Planar Surfaces

Juyang Weng, *Member, IEEE,* Narendra Ahuja, *Senior Member, IEEE,* and Thomas S. Huang, *Fellow, IEEE*

***Abstract*—We investigate in this paper the determination of motion and structure of a planar scene from monocular image sequences. Since algorithms for curved surfaces are not applicable to planar scenes, algorithms for planar surfaces need to be developed. In this paper, we present a new and simpler linear algorithm that gives closed-form solutions for motion and structure parameters using point correspondences between two images, assuming that the coplanar points undergo a rigid motion in 3-D space. A series of new analytical results is established. It is proved in this paper that any 3×3 matrix with a rank of no less than two is decomposable in the sense defined in this paper. This implies that our linear algorithm does not neglect any necessary constraint in solving the linear equations, and it optimizes a direct nonlinear, nonquadratic objective function. From two images, the two candidate solutions of the existing algorithms are both valid interpretations, although one of them does not agree with what actually happened (illusive solution). The existence of four object points, among which no three are collinear, is not only a sufficient condition for our linear algorithm to give at most two solutions, it is in fact a necessary condition for any algorithm. In other words, if this condition is not satisfied, no algorithm can reach at most two solutions from two views. There exists a class of so-called plane-perceivable surfaces, which includes planes as a spatial case. If the points lie in a noncoplanar surface in this class, one can still interpret the two images as the projections of coplanar points undergoing a rigid motion. If three images are available, the solution is generally unique, but not always. An approach to assessing the accuracy of the solutions is applied to this problem. The estimated errors give quantitative assessment of the reliability of the solutions and indicate any degenerate or nearly degenerate point configurations that cause failure of the motion analysis algorithm. This method of error estimation is applicable to other general least squares or minimum-norm least squares problems. Experimental results for real images are presented with automatically computed point correspondences.**

I. Introduction

UNDERSTANDING three-dimensional motion and recovering structure of a scene from image sequences may involve several steps. One of the steps is estimating the motion parameters between two consecutive image frames (called two-view motion parameters) and the structure of the scene. A subsequent step is understanding local motion (short-term motion that covers more image frames) based on some model of object dynamics [16]. This step makes it possible to predict the future motion and recover some missing motion information.

Existing approaches to motion analysis may be classified into two categories: continuous and discrete. With a continuous approach, small interframe motion and smooth image intensity functions are required, and the motion parameters are analyzed in terms of velocity. Typically, an optical flow field is computed and then the motion and structure parameters are derived from the optical flow field. With a discrete approach, interframe motions can be either large or small, and the motion parameters are analyzed in terms of displacement. The continuous approaches suffer from the unreliability of the optical flow generated by the available methods [13]. Since the motion has to be small, the pixel level errors can easily override the information in the data [20]. With the discrete approaches, it is difficult to match images due to the allowed large image disparities. However, the solutions of the corresponding motion analysis algorithms are more stable under relatively large interframe motions.

With a discrete approach, estimating motion and structure parameters using point correspondences generally involves three steps: first, extracting feature points, where corner or edge detectors are applied to each image to locate the feature points [1], [4], [9], [21]; second, establishing the point correspondences between two images through matching or tracking [1], [3], [8] (the point correspondences can also be provided by an image plane displacement field [18], and so, the above two steps can be unified into one); finally, computing the motion parameters and the 3-D positions of the points based on the computed point correspondences. This paper is devoted mainly to the last step.

Under perspective projection, the equations that relate the motion parameters and the image coordinates of the object points are nonlinear in the motion parameters. To solve these nonlinear equations, generally an iterative search has to be performed, which may often converge to a local extremum or even diverge. Therefore, a closed-form solution is highly desirable. Longuet-Higgins [5], and Tsai and Huang [12] developed linear algorithms that give closed-form solutions to the motion parameters and the structure of the object points. With the consideration

Manuscript received May 1, 1989; revised January 8, 1991. This work was supported in part by the National Science Foundation under Grants ECS-83-52408 and IRI-86-05400.

The authors are with the Beckman Institute and Coordinated Science Laboratory, University of Illinois, Urbana, IL 61801.

IEEE Log Number 9103198.

of noise in the image data, more stable linear algorithms have also been developed [2], [15].

However, if the object points are coplanar, i.e., they all lie on a 3-D plane, those algorithms fail. The configurations for which those algorithms fail are called degenerate configurations. Coplanar point configurations could arise from a planar scene or object. Aerial images of plain and sea, images of wall, ceiling, floor, and open road give some examples of planar scenes. Sometimes, the extracted and matched points can be coplanar, although the actual surface where the points are located is not a plane. For those coplanar point configurations, special algorithms are needed to solve the problem. In fact, the use of such a plane-based algorithm is not restricted to only planar scenes. If the point configuration is degenerate or nearly degenerate [7] to the algorithms for general surfaces, the problem may still be solved by first segmenting the scene into planar surfaces and then applying the plane-based algorithm to each planar surface. Of course, such a segmentation is not a trivial task.

Tsai and Huang [10] have developed a linear algorithm, with a discrete approach, that solves for the motion parameters of a planar patch using singular value decomposition. Longuet-Higgins [6] as well as Waxman and Wohn [14] studied this problem in terms of motion velocity (continuous approach). It is known that from two images, there are generally two candidate solutions. It is stated in [11] that if four point correspondences (four points in each image) are available, no three object points are collinear, the intermediate parameters can be determined up to a scaling factor from the linear equations.

In this paper, we investigate the related issues, in terms of the discrete approach, and give the following major new results: 1) A new and simpler algorithm is presented which is based on the eigenvalue decomposition of FF^t: $FF^t = U\Lambda^2U^t$ (instead of the singular value decomposition of F: $F = U\Lambda V^t$ used in [10]). The derivation of our algorithm provides new insight into the problem. 2) The complete case study of the problem is presented: including cases that have one solution, two solutions, and infinitely many solutions along with the corresponding necessary and sufficient conditions that result in those cases (Theorem 4 and Fig. 4). If the motion does not flip the plane to show the other side to the camera, then at most two solutions are possible (Corollary 2). 3) As other algorithms [6], [10], [14], in addition to the veridical solution (which agrees with the reality), our new algorithm generally gives one extra solution. It is pointed out and proved in this paper that this extra solution is not just a by-product of the specific algorithms, it is, in fact, another valid interpretation. That is, two solutions give the same pair of images (Theorem 3). 4) Necessary and sufficient conditions on the point configurations are presented that determine whether the solution of the intermediate motion parameter matrix F is unique up to a scaling factor from the linear equations (Theorem 1 and Corollary 1). The conditions are given in terms of both image plane points and object plane points. It is worth mentioning that those conditions are derived for an arbitrary number of points (not just 4), which is not trivial to the necessity part of the conditions. 5) It is pointed out and proved that general 3×3 matrices are decomposable in the sense defined in this paper (Theorem 5). This implies that, unlike the linear algorithms for general surfaces, our linear algorithm for planar surfaces does not neglect any constraint in solving for the intermediate parameters based on the linear equations. This result enables us to conclude that the conditons of point configuration mentioned above are also necessary to any algorithm (Theorem 6). It also leads to a direct optimality of our algorithm (expression (3.16)). 6) There is a class of plane-perceivable surfaces, including planes as special cases (Theorem 7). If the object points lie on a nonplanar surface in this class, one can still interpret the two given images as the projections of coplanar points undergoing a rigid motion. If and only if the object points do not lie in any plane-perceivable surface, one can detect that the two images do not correspond to a rigid motion of coplanar points. 7) The approach to assessing accuracy of the solutions in the presence of noise, which has been introduced in [15], is applied to the problem here.

The motion estimation algorithm and the associated analytical results are discussed in the next section. The inherent uniqueness of the problem is investigated in Section III. Section IV discusses error estimation. Section V presents the experimental results and Section VI summarizes the paper.

II. Motion and Structure of a Planar Patch

Let the coordinate system be fixed on the camera with the origin coinciding with the projection center of the camera, and the Z axis coinciding with the optical axis and pointing to the scene (see Fig. 1). Since the image and the focal length of a pinhole camera model can be scaled by any positive number without changing the projection line of a point, we scale the image and focal length by the reciprocal of the original focal length so that the new focal length is equal to one. The resulting model is called a normalized camera model, in which the image plane is located at $z = 1$ and the image size determines the field of view. Visible objects are always located in front (forward half space) of the camera, i.e., $z > 0$. $0 < z < 1$ may occur since the camera model is normalized.

A. Notation and Problem Statement

We first introduce some notation. $\triangleq$ is used in the equations to define new variables when the variables to be defined are obvious. Matrices are denoted by capital italics (e.g., R). Vectors are denoted by uppercase or lowercase boldface letters, (e.g., $\boldsymbol{T}$, $\boldsymbol{x}$). A vector with a top circle (e.g., $\mathring{\boldsymbol{T}}$) is used to denote the corresponding unit vector of the original vector. A vector with a tilde (e.g., $\tilde{\boldsymbol{T}}$) is

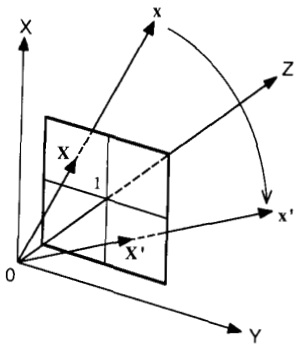


Fig. 1. Geometry of the normalized camera model.

used to denote the scaled (normalized) version of the original vector with the scaling factor defined where appropriate. The components of a column matrix (or vector) $\boldsymbol{V}$ is specified by $\boldsymbol{V} = (v_1, v_2, v_3)^t$, where ^{t} denotes the transposition. All the vectors are column vectors by default. A vector is also regarded as a column matrix, and so vector operations such as cross product ($\times$), and matrix operations such as matrix multiplication, are both applied to three-dimensional vectors. Matrix operations precede vector operations. $\|$ denotes "parallel," i.e., $\boldsymbol{X} \| \boldsymbol{Y}$ if and only if $\boldsymbol{X} \times \boldsymbol{Y} = 0$. A dot ("$\cdot$") denotes dot (inner) product. For vectors we have $\boldsymbol{X} \cdot \boldsymbol{Y} = \boldsymbol{X}^t\boldsymbol{Y}$. I_m denotes an $m \times m$ identity matrix. For a matrix $A = [a_{ij}]$, $\|A\|$ denotes the Euclidean norm of the matrix, i.e., $\|[a_{ij}]\| = \sqrt{\Sigma_{ij}\, a_{ij}^2}$. We define a mapping $[\,\cdot\,]_\times$ from a three-dimensional vector to a 3×3 matrix:

$$[(x_1, x_2, x_3)]_\times = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix}. \tag{2.1}$$

Using this mapping, we can express cross operation of two vectors by the matrix multiplication of a 3×3 matrix and a column matrix:

$$\boldsymbol{X} \times \boldsymbol{Y} = [\boldsymbol{X}]_\times \boldsymbol{Y}. \tag{2.2}$$

Consider a point P on the object which is visible before motion (at time t_1) and after motion (at time t_2). The following notation is used for the spatial vectors and the image vectors (see Fig. 1):

$\boldsymbol{x} = (x, y, z)^t$ spatial vector of P at time t_1,

$\boldsymbol{x}' = (x', y', z')^t$ spatial vector of P at time t_2,

$$\boldsymbol{X} = (u, v, 1)^t = \left(\frac{x}{z}, \frac{y}{z}, 1\right)^t$$

image vector of P at time t_1,

$$\boldsymbol{X}' = (u', v', 1)^t = \left(\frac{x'}{z'}, \frac{y'}{z'}, 1\right)^t$$

image vector of P at time t_2,

where (u, v) and (u', v') are the image coordinates of the point. Therefore, the spatial vector and image vector are related by

$$\boldsymbol{x} = z\boldsymbol{X}, \qquad \boldsymbol{x}' = z'\boldsymbol{X}'. \tag{2.3}$$

The third component of a spatial vector is called depth.

Let $\boldsymbol{R}$ and $\boldsymbol{T}$ be a rotation matrix and a translation vector, respectively, that represent a rigid motion from time t_1 to time t_2, i.e., the spatial vectors at the two time instants are related by

$$\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{T}. \tag{2.4}$$

In terms of image vectors we have

$$z'\boldsymbol{X}' = z\boldsymbol{R}\boldsymbol{X} + \boldsymbol{T}. \tag{2.5}$$

If $\|\boldsymbol{T}\| \neq 0$, from (2.5) we get

$$\frac{z'}{\|\boldsymbol{T}\|}\boldsymbol{X}' = \frac{z}{\|\boldsymbol{T}\|}\boldsymbol{R}\boldsymbol{X} + \mathring{\boldsymbol{T}} \tag{2.6}$$

where, using our notation,

$$\mathring{\boldsymbol{T}} = \frac{\boldsymbol{T}}{\|\boldsymbol{T}\|}. \tag{2.7}$$

Let the plane where the points are located at time t_1 be represented by

$$\boldsymbol{N}^t\boldsymbol{x} = 1 \tag{2.8}$$

where $\boldsymbol{N}$ is the normal vector of the plane. The distance d between the origin and the plane is $d = \|\boldsymbol{N}\|^{-1}$. By such a representation of object plane, we exclude the cases in which the plane goes through the origin, because in those cases the plane is invisible to the camera (the image of the plane is a straight line and the motion cannot be determined). Similarly, at time t_2, that plane is represented by

$$(\boldsymbol{N}')^t\boldsymbol{x}' = 1. \tag{2.9}$$

From (2.3), (2.8), and (2.9), we have

$$(\|\boldsymbol{T}\|\boldsymbol{N}^t)\frac{z}{\|\boldsymbol{T}\|}\boldsymbol{X} = 1, \qquad (\|\boldsymbol{T}\|(\boldsymbol{N}')^t)\frac{z'}{\|\boldsymbol{T}\|}\boldsymbol{X}' = 1. \tag{2.10}$$

We want to determine the normalized depths $z/\|\boldsymbol{T}\|$ and $z'/\|\boldsymbol{T}\|$. Equivalently, we determine the normalized normals

$$\tilde{\boldsymbol{N}} = \|\boldsymbol{T}\|\boldsymbol{N}, \qquad \tilde{\boldsymbol{N}}' = \|\boldsymbol{T}\|\boldsymbol{N}'. \tag{2.11}$$

From the normalized normal we can determine the normalized depth from (2.10).

Solving for $\boldsymbol{x}$ in (2.4) and substituting $\boldsymbol{x}$ in (2.8), we get the plane at time t_2

$$(\boldsymbol{R}\boldsymbol{N})^t\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{N} \cdot \boldsymbol{T} + 1. \tag{2.12}$$

From (2.9) and (2.12), the normal of the plane at time t_2 is expressed by

$$\boldsymbol{N}' = \frac{\boldsymbol{R}\boldsymbol{N}}{\boldsymbol{R}\boldsymbol{N} \cdot \boldsymbol{T} + 1}. \tag{2.13}$$

From (2.11) and (2.13) we get the equation for the normalized normal at time t_2:

$$\tilde{N}' = \frac{R\tilde{N}}{R\tilde{N} \cdot \mathring{T} + 1}. \tag{2.14}$$

Given n corresponding image vector pairs at two time instants, (X_i, X_i'), $i = 1, 2, \cdots, n$, which represent the projections of 3-D points on the image plane, the algorithm solves for the rotation matrix R. If the translation vector T does not vanish, the algorithm also solves for the translation direction (represented by a unit vector $\mathring{T}$) and the normalized normal $\tilde{N}$. The normalized depth $z_i/\|T\|$ and $z_i'/\|T\|$ for object points x_i and x_i', can then be determined from (2.10) and (2.14). By solution or interpretation, we always mean this normalized solution. The magnitude of the translation vector ($\|T\|$), and the absolute depths of the object points (z_i and z_i') cannot be determined by monocular vision. This is easy to be seen from (2.6), which still holds when $\|T\|$, z and z' are multiplied by any positive constant. That is, multiplying the point depths and $\|T\|$ by the same scaling factor does not change the images. Equivalently, the magnitude of the normal vectors $\|N\|$ and $\|N'\|$ cannot be determined since we have (2.10). Once any one of those original unknowns is available according to some additional knowledge, the scaling factor can be determined directly using the normalized version, and then all the rest of the unknowns can be determined completely. In summary, $\|T\|$ inherently cannot be determined from monocular vision. We defined the normalized version of the solution above that corresponds to a unit translation. In the remainder of this paper, we call the normalized solutions simply solutions.

B. Algorithm

To provide a guide for the derivation, we first present an outline of the algorithm in this subsection. This outline may be skipped in the first reading and, instead, be examined later for correctness when each step is discussed in the derivation. The early introduction of this outline is helpful to keep track of the derivation, as our derivation is interlaced with discussions of the associated properties.

We will define a 3×3 intermediate parameter matrix F, which is a function of the motion parameters and the normal of the object plane. The equations that relate the image coordinates of the points and the matrix F are linear in the elements of F. This allows us to solve for F based on these linear equations. Then the motion parameters R and $\mathring{T}$, and the normalized normal of the object plane $\tilde{N}$ are determined from the intermediate matrix F. It turns out that the two images allow two different interpretations in general. Therefore, two solutions are computed from F. There are some special cases in which the number of solutions is not exactly two. In the absence of noise, our algorithm gives exact solution. For the reliability in the presence of noise, the algorithm is also designed to utilize redundant information in the data.

The algorithm from two views, is as follows.

Step i) *Solve for intermediate parameter matrix* $F = R + TN^t$:

Given $X_i = (u_i, v_i, 1)^t$, $X_i' = (u_i', v_i', 1)^t$, $i = 1, 2, \cdots, n$, the corresponding image vectors of n ($n \geq 4$) points, let A be a $2n \times 9$ matrix such that

$$A = \begin{bmatrix} X_1^t & 0 & -u_1' X_1^t \\ 0 & X_1^t & -v_1' X_1^t \\ X_2^t & 0 & -u_2' X_2^t \\ 0 & X_2^t & -v_2' X_2^t \\ \vdots & \vdots & \vdots \\ X_n^t & 0 & -u_n' X_n^t \\ 0 & X_n^t & -v_n' X_n^t \end{bmatrix} \tag{Ag.1}$$

and h be a 9-dimensional vector

$$h = (h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9)^t. \tag{Ag.2}$$

We solve for unit vector h in the following:

$$\min_h \|Ah\|, \qquad \text{subject to: } \|h\| = 1. \tag{Ag.3}$$

The solution of h is a unit eigenvector of A^tA associated with the smallest eigenvalue. Then F_s is determined by

$$F_s = \begin{bmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{bmatrix}. \tag{Ag.4}$$

Let $H = [h_1 \quad h_2 \quad h_3]$ be a 3×3 orthogonal matrix such that

$$H^t F_s^t F_s H = \text{diag}\,(\gamma_1, \gamma_2, \gamma_3) \tag{Ag.5}$$

with $\gamma_1 \leq \gamma_2 \leq \gamma_3$, where diag $(\gamma_1, \gamma_2, \gamma_3)$ denotes the conventional diagonal matrix with the corresponding diagonal elements. Then

$$F = \frac{1}{\sqrt{\gamma_2}} F_s. \tag{Ag.6}$$

If

$$\sum_i X_i' \cdot FX_i < 0 \tag{Ag.7}$$

then $F \leftarrow -F$. The summation in (Ag.7) is over several values of i's to reduce the sensitivity to noise (usually three or four values of i suffice).

Step ii) *Solve for* R, $\mathring{T}$, *and* $\tilde{N}$ *from* F:

We have

$$H^t F^t F H = \text{diag}\,(\gamma_1/\gamma_2, 1, \gamma_3/\gamma_2) \triangleq \text{diag}\,(\lambda_1, \lambda_2, \lambda_3). \tag{Ag.8}$$

Case 1: $\lambda_1 < 1 < \lambda_3$ (iff $\boldsymbol{T} \times (\boldsymbol{RN}) \neq \boldsymbol{0}$):

There exist two solutions that give the same images. Let

$$\alpha = \sqrt{\frac{\lambda_3 - 1}{\lambda_3 - \lambda_1}}, \qquad \beta = \sqrt{\frac{1 - \lambda_1}{\lambda_3 - \lambda_1}}. \tag{Ag.9}$$

The first solution: Let

$$\boldsymbol{V}_1 = \alpha \boldsymbol{h}_1 + \beta \boldsymbol{h}_3, \qquad \boldsymbol{V}_2 = \boldsymbol{h}_2. \tag{Ag.10}$$

Then

$$R = [F\boldsymbol{V}_1 \quad F\boldsymbol{V}_2 \quad F\boldsymbol{V}_1 \times F\boldsymbol{V}_2][\boldsymbol{V}_1 \quad \boldsymbol{V}_2 \quad \boldsymbol{V}_1 \times \boldsymbol{V}_2]'. \tag{Ag.11}$$

Let

$$\mathring{\boldsymbol{N}} = \boldsymbol{V}_1 \times \boldsymbol{V}_2 \tag{Ag.12}$$

and

$$\tilde{\boldsymbol{T}} = F\mathring{\boldsymbol{N}} - R\mathring{\boldsymbol{N}}. \tag{Ag.13}$$

If

$$\sum_i (\boldsymbol{X}_i' \times R\boldsymbol{X}_i) \cdot (\tilde{\boldsymbol{T}} \times \boldsymbol{X}_i') < 0 \tag{Ag.14}$$

then $\mathring{\boldsymbol{N}} \leftarrow -\mathring{\boldsymbol{N}}$ and reevaluate (Ag.13) (changes the sign of $\tilde{\boldsymbol{T}}$). As in (Ag.7), Σ_i in (Ag.14) sums over several values of i to suppress noise. Finally, we get

$$\mathring{\boldsymbol{T}} = \frac{\tilde{\boldsymbol{T}}}{\|\tilde{\boldsymbol{T}}\|} \tag{Ag.15}$$

and

$$\tilde{\boldsymbol{N}} = \|\tilde{\boldsymbol{T}}\| \mathring{\boldsymbol{N}}. \tag{Ag.16}$$

The second solution: Change the sign of β we got in (Ag.9) (i.e., $\beta \leftarrow -\beta$) and keep the α unchanged. (Ag.10)-(Ag.16) give the second solution.

Case 2: $\lambda_1 = 1 < \lambda_3$ (iff $\boldsymbol{T} \| \boldsymbol{RN}$ and the absolute distance between the plane and the origin increases due to motion):

In this case, $\alpha = 1$ and $\beta = 0$. Equation (Ag.9)-(Ag.16) give the unique solution.

Case 3: $\lambda_1 < 1 = \lambda_3$ (iff $\boldsymbol{T} \| \boldsymbol{RN}$ and the absolute distance between the plane and the origin decreases due to motion):

In this case, $\alpha = 0$ and $\beta = 1$. Equations (Ag.9)-(Ag.16) give the unique solution.

Case 4: $\lambda_1 = 1 = \lambda_3$ (iff $\boldsymbol{T} \| \boldsymbol{RN}$ and the absolute distance between the plane and the origin does not change):

If det $(F) > 0$, report $\boldsymbol{T} = 0$. $R = F$. $\tilde{\boldsymbol{N}}$ cannot be determined.

det $(F) < 0$ occurs only if the backside of the plane faces the camera after motion, which is impossible for an opaque plane. If the plane is transparent and the points on the plane are visible on both sides, this case can happen. If so, the solutions are infinitely many. For any unit $\mathring{\boldsymbol{N}}$, the following is a solution:

$$R = F(I_3 - 2\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}') \tag{Ag.17}$$

$$\tilde{\boldsymbol{T}} = -2R\mathring{\boldsymbol{N}}. \tag{Ag.18}$$

Equations (Ag.15) and (Ag.16) give $\mathring{\boldsymbol{T}}$ and $\tilde{\boldsymbol{N}}$, respectively.

Note 1: The necessary and sufficient conditions for each case in step ii to occur are for the noise-free images. With noise, generally only case 1 can occur.

Note 2: Even if u_i, v_i, u_i', and v_i', $i = 1, 2, \cdots, n$, are contaminated by noise, R in (Ag.11) is still a rotation matrix, and step ii gives rotation matrix R, unit vector $\mathring{\boldsymbol{T}}$, and vector $\tilde{\boldsymbol{N}}$ such that the following exactly holds:

$$F_s = \pm\sqrt{\gamma_2}(R + \mathring{\boldsymbol{T}}\tilde{\boldsymbol{N}}') \tag{Ag.19}$$

where F_s is given in (Ag.4) and the sign is determined in (Ag.7).

Note 3: The condition that the points are coplanar can be checked in the following way. Without noise, if the rank of A in (Ag.1) is more than 8, the points are not coplanar. In the presence of noise, if

$$\lambda_1 > \epsilon \tag{Ag.20}$$

where ϵ is a threshold based on the error analysis of λ_1 discussed in Section IV, the points are expected to be noncoplanar.

C. Intermediate Parameter Matrix F

We first justify step i, where an intermediate parameter matrix F is determined using linear equations. From (2.4) and (2.8), we have

$$\boldsymbol{x}' = (R + \boldsymbol{TN}')\boldsymbol{x}. \tag{2.15}$$

We define the intermediate parameter matrix:

$$F = R + \boldsymbol{TN}'. \tag{2.16}$$

From (2.3), (2.15), and (2.16) we get

$$z'\boldsymbol{X}' = Fz\boldsymbol{X}. \tag{2.17}$$

Since $z > 0$ and $z' > 0$, we get the relation between the intermediate parameter matrix F and the image vectors:

$$\boldsymbol{X}' \times F\boldsymbol{X} = \boldsymbol{0}. \tag{2.18}$$

Let $F = [e_{ij}]$ and $\boldsymbol{h} \| (e_{11}, e_{12}, e_{13}, e_{21}, \cdots, e_{33})'$. Equation (2.18) can be rewritten as

$$\begin{bmatrix} \boldsymbol{X}' & \boldsymbol{0} & -u'\boldsymbol{X}' \\ \boldsymbol{0} & \boldsymbol{X}' & -v'\boldsymbol{X}' \\ v'\boldsymbol{X}' & -u'\boldsymbol{X}' & \boldsymbol{0} \end{bmatrix} \boldsymbol{h} = \boldsymbol{0}. \tag{2.19}$$

Only two of the above three scalar equations are independent (see (C.2) in Appendix C). The last can be derived from the first two. We keep the first two equations. Thus, each point correspondence gives two scalar equations and

n point correspondences give $2n$ such scalar equations. We get

$$\boldsymbol{A}\boldsymbol{h} = \boldsymbol{0} \tag{2.20}$$

where A is a $2n \times 9$ matrix defined in (Ag.1). If rank (A) is equal to 8, $\boldsymbol{h}$ can be determined up to a scaling factor using (2.20). n must be at least 4 to make rank $(A) = 8$ possible. The following theorem gives a necessary and sufficient condition for the rank of A to equal 8.

Theorem 1: Rank $(A) = 8$ if and only if there exists a set of four object points such that no image projections of any three points in this set are collinear in any of the two images.

Proof: See Appendix A.

The above theorem gives a necessary and sufficient condition in terms of the images of the points. It is then ready to get the following corollary which gives a necessary and sufficient condition in terms of 3-D object points.

Corollary 1: Rank $(A) = 8$ if and only if 1) there exists a set of four points in the object plane such that no three points in this set are collinear in the object plane, and 2) if the object plane is extended, it does not go through the projection center of the camera before and after motion.

Proof: See Appendix A.

Assuming rank $(A) = 8$, we can determine $\boldsymbol{h}$ (and consequently F) up to a scaling factor. In the presence of noise, we solve a unit vector $\boldsymbol{h}$ using (Ag.3). Since rank $(A) = 8$, we have $F_s = sF$ for some scaling factor s. The magnitude of this scaling factor can be determined by the second smallest eigenvalue of $F'F$, which should be equal to one as we show below in (2.30). In (Ag.5) the second smallest eigenvalue of $F_s'F_s$ is then $\gamma_2 = s^2$. So $F = \pm F_s/|s| = \pm F_s/\sqrt{\gamma_2}$, which accounts for the step in (Ag.6). The sign of the scaling factor of F can be determined from (2.17). Since $z > 0$, $z' > 0$, we have

$$\boldsymbol{X}' \cdot \boldsymbol{F}\boldsymbol{X} > 0. \tag{2.21}$$

With noise we use (Ag.7) to determine the sign of F. This step is stable in the presence of noise, since the decision is made based on the sign of the inner product of the two vectors which are in the same or opposite direction without noise. In (Ag.7) the summation over several points suppresses the cases where two noise-corrupted small vectors happen to be used, whose inner product is close to zero and the sign is unreliable.

D. Solving for Unknowns from F: General Case

In step ii, we solve for R, $\mathring{T}$ and $\mathring{N}$ from F. We first give a guideline. For the definition of F in (2.16), we can see that $\boldsymbol{F}\boldsymbol{V}_i = \boldsymbol{R}\boldsymbol{V}_i$ if $\boldsymbol{N}'\boldsymbol{V}_i = \boldsymbol{0}$. Geometrically, this implies that F is a rotation matrix for the vectors in the plane Π that passes through the origin and is parallel to the object plane. We can determine this plane Π by using the property that the length of vector is conserved under a rotation. What we first need to do is finding two different vectors $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ that belong to Π, which is done using the eigenvalue decomposition of $F'F$. The rotation matrix R is then determined by the constraint that R rotates V_i to FV_i, for $i = 1, 2$. $\mathring{N}$ is a unit vector that is orthogonal to both $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$, and finally $\mathring{T}$ is determined from R, $\mathring{N}$, and F.

Let $V = [\boldsymbol{V}_1 \quad \boldsymbol{V}_2 \quad \boldsymbol{V}_3]$, where $\boldsymbol{V}_3 = \mathring{\boldsymbol{N}} \triangleq \boldsymbol{N}/\|\boldsymbol{N}\|$, $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are such that V is an orthogonal matrix. Since $\boldsymbol{V}_1$ is orthogonal to $\boldsymbol{N}$, it follows that

$$\boldsymbol{F}\boldsymbol{V}_1 = (\boldsymbol{R} + \boldsymbol{T}\boldsymbol{N}^t)\boldsymbol{V}_1 = \boldsymbol{R}\boldsymbol{V}_1 + \boldsymbol{T}\boldsymbol{N}^t\boldsymbol{V}_1 = \boldsymbol{R}\boldsymbol{V}_1. \tag{2.22}$$

Similarly we have

$$\boldsymbol{F}\boldsymbol{V}_2 = \boldsymbol{R}\boldsymbol{V}_2. \tag{2.23}$$

Therefore FV_1 and FV_2 are orthonormal. We get

$$\boldsymbol{F}\boldsymbol{V}_3 = \boldsymbol{R}\boldsymbol{V}_3 + \boldsymbol{T}\boldsymbol{N}^t\boldsymbol{V}_3 = \boldsymbol{R}\boldsymbol{V}_3 + \|\boldsymbol{N}\|\boldsymbol{T}. \tag{2.24}$$

Since FV_1 and FV_2 are orthonormal, we obtain

$$V'F'FV = \begin{bmatrix} 1 & 0 & a \\ 0 & 1 & b \\ a & b & c \end{bmatrix} \triangleq U \tag{2.25}$$

where (using (2.22)–(2.24))

$$\begin{aligned} a &= \boldsymbol{F}\boldsymbol{V}_1 \cdot \boldsymbol{F}\boldsymbol{V}_3 = \boldsymbol{R}\boldsymbol{V}_1 \cdot (\boldsymbol{R}\boldsymbol{V}_3 + \|\boldsymbol{N}\|\boldsymbol{T}) \\ &= (\boldsymbol{R}\boldsymbol{V}_1 \cdot \boldsymbol{T})\|\boldsymbol{N}\| \end{aligned} \tag{2.26.a}$$

$$\begin{aligned} b &= \boldsymbol{F}\boldsymbol{V}_2 \cdot \boldsymbol{F}\boldsymbol{V}_3 = \boldsymbol{R}\boldsymbol{V}_2 \cdot (\boldsymbol{R}\boldsymbol{V}_3 + \|\boldsymbol{N}\|\boldsymbol{T}) \\ &= (\boldsymbol{R}\boldsymbol{V}_2 \cdot \boldsymbol{T})\|\boldsymbol{N}\| \end{aligned} \tag{2.26.b}$$

$$\begin{aligned} c &= \boldsymbol{F}\boldsymbol{V}_3 \cdot \boldsymbol{F}\boldsymbol{V}_3 = \|\boldsymbol{R}\boldsymbol{V}_3 + \|\boldsymbol{N}\|\boldsymbol{T}\|^2 \\ &= 1 + 2(\boldsymbol{R}\boldsymbol{N} \cdot \boldsymbol{T}) + \|\boldsymbol{T}\|^2\|\boldsymbol{N}\|^2. \end{aligned} \tag{2.26.c}$$

From (2.25), the characteristic polynomial of U is

$$\begin{aligned} &\det(\lambda I_3 - U) \\ &= \det\left(\begin{bmatrix} \lambda - 1 & 0 & -a \\ 0 & \lambda - 1 & -b \\ -a & -b & \lambda - c \end{bmatrix}\right) \\ &= (\lambda - 1)(\lambda^2 - (1 + c)\lambda - (a^2 + b^2 - c)). \end{aligned} \tag{2.27}$$

The eigenvalues of $F'F$ are obtained by solving for the roots of the characteristic polynomial in (2.27):

$$\lambda_2 = 1 \tag{2.28.a}$$

$$\lambda_1 = \frac{1 + c - \sqrt{(1 - c)^2 + 4(a^2 + b^2)}}{2} \tag{2.28.b}$$

$$\lambda_3 = \frac{1 + c + \sqrt{(1 - c)^2 + 4(a^2 + b^2)}}{2}. \tag{2.28.c}$$

First consider $a^2 + b^2 = 0$. From (2.28) we have

$$\lambda_1 = \min(1, c), \qquad \lambda_3 = \max(1, c) \quad \text{if } a^2 + b^2 = 0. \tag{2.29.a}$$

If $a^2 + b^2$ increases from zero, from (2.28) we see that λ_1 decreases and λ_3 increases. Therefore, it follows that

$$\lambda_1 < \min(1, c), \qquad \lambda_3 > \max(1, c)$$
$$\text{if } a^2 + b^2 \neq 0. \tag{2.29.b}$$

Noticing $\lambda_2 = 1$, we always have

$$\lambda_1 \leq \lambda_2 = 1 \leq \lambda_3. \tag{2.30}$$

The multiple eigenvalues result in special solutions, which we will discuss in the next subsection. We first assume that the eigenvalues are distinct.

Let Π be a plane spanned by all the vector orthogonal to $\boldsymbol{N}$, i.e., $\Pi = \{\boldsymbol{V} | \boldsymbol{V} \cdot \boldsymbol{N} = 0\}$. In addition, we define another plane Π_H that is transformed from Π by orthogonal matrix H', where H is defined in (Ag.5), i.e., $\Pi_H = \{H'\boldsymbol{V} | \boldsymbol{V} \cdot \boldsymbol{N} = 0\}$. For any vector $\boldsymbol{U} \in \Pi_H$, $\boldsymbol{U} = (u_x, u_y, u_z)$, there exists a vector $\boldsymbol{V}$ in Π such that

$$\boldsymbol{U} = H'\boldsymbol{V}. \tag{2.31}$$

From (Ag.8) and (2.31), we have

$$\|F\boldsymbol{V}\|^2 = \boldsymbol{V}'F'F\boldsymbol{V} = \boldsymbol{U}'H'F'FH\boldsymbol{U}$$
$$= \boldsymbol{U}' \operatorname{diag}(\lambda_1, 1, \lambda_3)\boldsymbol{U} = \lambda_1 u_x^2 + u_y^2 + \lambda_3 u_z^2. \tag{2.32}$$

On the other hand,

$$\|R\boldsymbol{V}\|^2 = \|\boldsymbol{V}\|^2 = \|H\boldsymbol{U}\|^2 = \|\boldsymbol{U}\|^2 = u_x^2 + u_y^2 + u_z^2. \tag{2.33}$$

Since $\boldsymbol{V}$ is in Π, we get

$$F\boldsymbol{V} = (R + \boldsymbol{T}\boldsymbol{N}')\boldsymbol{V} = R\boldsymbol{V}. \tag{2.34}$$

Therefore the last terms in (2.32) and (2.33) are equal:

$$\lambda_1 u_x^2 + u_y^2 + \lambda_3 u_z^2 = u_x^2 + u_y^2 + u_z^2 \tag{2.35}$$

or equivalently

$$\sqrt{1 - \lambda_1}\, u_x = \pm\sqrt{\lambda_3 - 1}\, u_z. \tag{2.36}$$

Each sign in (2.36) determines a plane. If $\lambda_1 \neq 1$ and $\lambda_3 \neq 1$, we get two planes:

$$\Pi_a = \{(u_x, u_y, u_z) | \sqrt{1 - \lambda_1}\, u_x = \sqrt{\lambda_3 - 1}\, u_z\} \tag{2.37.a}$$

$$\Pi_b = \{(u_x, u_y, u_z) | \sqrt{1 - \lambda_1}\, u_x = -\sqrt{\lambda_3 - 1}\, u_z\}. \tag{2.37.b}$$

Since all the vectors in the plane Π_H satisfy (2.35), we know $\Pi_H = \Pi_a$ or $\Pi_H = \Pi_b$.

Geometrically, (2.32) implies that F transforms a unit sphere to an ellipsoid. Equation (2.35) indicates that we intersect a unit sphere with this ellipsoid to find a plane Π from which F is a length-preserving transformation. Since F is a rotation for vectors in Π, such an intersection exists. However, since there are generally two planes whose intersection with the ellipsoid is a circle, two planes Π_a and Π_b are determined. Π_H is one of them.

First assume $\Pi_H = \Pi_a$. Choosing any two unit vectors $\boldsymbol{U}_1$ and $\boldsymbol{U}_2$ from Π_a, from (2.31) we get two unit vectors in Π: $\boldsymbol{V}_i = H\boldsymbol{U}_i$, $i = 1, 2$. In the algorithm, $\boldsymbol{U}_1 = (\alpha, 0, \beta)'$ and $\boldsymbol{U}_2 = (0, 1, 0)'$ are the two orthonormal vectors chosen from Π_a in (Ag.10), where α and β are defined in (Ag.9).

From (2.34) we have $R\boldsymbol{V}_i = F\boldsymbol{V}_i$, $i = 1, 2$. R is determined uniquely by the positions of the two nonparallel vectors. In fact, $R[\boldsymbol{V}_1 \quad \boldsymbol{V}_2 \quad \boldsymbol{V}_1 \times \boldsymbol{V}_2] = [F\boldsymbol{V}_1 \quad F\boldsymbol{V}_2 \quad F\boldsymbol{V}_1 \times F\boldsymbol{V}_2]$, which leads to (Ag.11).

Now suppose that the observed image points are contaminated by noise, and so F has error. Given any 3×3 matrix F which is scaled in step i so that $\lambda_2 = 1$, is R given in (Ag.11) still a rotation matrix? The answer is positive. In other words, $F\boldsymbol{V}_1$ and $F\boldsymbol{V}_2$ are still orthonormal vectors. In fact, from (Ag.10) we know that $\boldsymbol{V}_i$, $i = 1, 2$, can be expressed by $\boldsymbol{V}_i = H\boldsymbol{U}_i$. It follows from (Ag.8) that

$$(F\boldsymbol{V}_i) \cdot (F\boldsymbol{V}_j) = \boldsymbol{U}_i'H'F'FH\boldsymbol{U}_j$$
$$= \boldsymbol{U}_i \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)\boldsymbol{U}_j = \delta(i - j) \tag{2.38}$$

for $i = 1, 2$, where $\delta(k) = 1$ if $k = 0$ and $\delta(k) = 0$ otherwise. Therefore, in the presence of noise, we still found a rotation matrix R such that

$$R\boldsymbol{V}_i = F\boldsymbol{V}_i \tag{2.39}$$

exactly holds for $i = 1, 2$, where $\boldsymbol{V}_i$ and F are computed from noisy data.

Now consider the unit normal of the plane $\mathring{N}$. Since $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$ are two orthonormal vector in plane Π which is orthogonal to $\boldsymbol{N}$, we have

$$\mathring{N} = \pm\boldsymbol{V}_1 \times \boldsymbol{V}_2. \tag{2.40}$$

For the time being, we assume the positive sign is correct and get (Ag.12). Define

$$\tilde{\boldsymbol{T}} = \|\boldsymbol{N}\|\boldsymbol{T}. \tag{2.41}$$

We get $F\mathring{N} = R\mathring{N} = \tilde{\boldsymbol{T}}\mathring{N}'\mathring{N} = R\mathring{N} + \tilde{\boldsymbol{T}}$, which gives (Ag.13). From (2.5) we know that the solutions p and p' in

$$p'\boldsymbol{X}' = p R\boldsymbol{X} + \tilde{\boldsymbol{T}} \tag{2.42}$$

are positive. Equation (2.42) gives

$$\boldsymbol{0} = p\boldsymbol{X}' \times R\boldsymbol{X} + \boldsymbol{X}' \times \tilde{\boldsymbol{T}}. \tag{2.43}$$

That is, $\tilde{\boldsymbol{T}} \times \boldsymbol{X}'$ and $\boldsymbol{X}' \times R\boldsymbol{X}$ have the same directions. If and only if the sign of $\tilde{\boldsymbol{T}}$ is wrong, they have the opposite directions and (Ag.14) holds. On the other hand, from (Ag.13) we know that the sign of $\mathring{N}$ is wrong if and only if that of $\tilde{\boldsymbol{T}}$ is wrong. So we change the sign of $\mathring{N}$ (and that of $\tilde{\boldsymbol{T}}$ accordingly) if (Ag.14) is true.

Equations (Ag.15) and (Ag.16) immediately follow from (2.7), (2.11), and (2.41). Note that $\tilde{\boldsymbol{T}}$ in (Ag.13) is not a zero vector, since otherwise F is orthogonal from (2.39) and case 4 occurs.

If $\Pi_H = \Pi_b$, $U_1 = (\alpha, 0, -\beta)$ and $U_2 = (0, 1, 0)$ are the two orthonormal vectors chosen from Π_b. The remaining steps are the same as those in the case $\Pi_H = \Pi_a$.

Since we have either $\Pi_H = \Pi_a$ or $\Pi_H = \Pi_b$, there are two possible candidate solutions. The following theorem says that they are both valid solutions to $F = R + \mathring{T}\tilde{N}^t$.

Theorem 2: If the eigenvalues of F^tF are distinct, there are exactly two solutions for R, $\mathring{T}$, and $\mathring{N}$ to the equation

$$R + \mathring{T}\tilde{N}^t = F \tag{2.44}$$

with the constraints that R is a rotation matrix, $\mathring{T}$ a unit vector and the signs of $\mathring{T}$ and $\tilde{N}$ determined by (Ag.14).

Proof: By the above derivation we know that there are at most two possible solutions corresponding to $\Pi_H = \Pi_a$ and $\Pi_H = \Pi_b$, respectively. We now prove that these two possible solutions are both solutions of (2.44). Using (Ag.9)–(Ag.16) we get the first solution R, $\mathring{T}$ and $\tilde{N}$. Since $\tilde{N}$ is orthogonal to V_i, for $i = 1, 2$, we have

$$(R + \mathring{T}\tilde{N}^t)V_i = RV_i = FN_i. \tag{2.45}$$

From (Ag.13)–(Ag.16), it follows that

$$\begin{aligned}(R + \mathring{T}\tilde{N}^t)\mathring{N} &= R\mathring{N} + \mathring{T}\|\tilde{T}\| = R\mathring{N} + \tilde{T} \\ &= R\mathring{N} + (F\mathring{N} - R\mathring{N}) = F\mathring{N}.\end{aligned} \tag{2.46}$$

From (2.45) and (2.46) we get $R + \mathring{T}\tilde{N}^t = F$, since V_1, V_2, and $\mathring{N}$ are linearly independent vectors (in fact, they are orthonormal). Thus R, $\mathring{T}$, and $\tilde{N}$ give one solution of (2.44).

For the second solution, change the sign of β in (Ag.9) and keep α in (Ag.9) unchanged. We get another set of R, $\mathring{T}$, and $\tilde{N}$ from (Ag.10)–(Ag.16). Equations (2.45) and (2.46) still hold for this solution. Thus the second solution satisfies (2.44). □

Now we know that the two solutions both satisfy (2.44). One question arises as to whether they both are consistent with the observed images. As stated in the following theorem, they correspond to two planes (under different motions) that give the same images at the two time instants. Therefore, two solutions are valid interpretations.

Theorem 3: Given n point correspondences, (X_i, X_i'), $i = 1, 2, \cdots, n$. Let F be a matrix that satisfies (2.18) for every (X_i, X_i'). If there are rotation matrices R_a, R_b, unit vectors $\mathring{T}_a$, $\mathring{T}_b$ and vectors $\tilde{N}_a$ and $\tilde{N}_b$, such that

$$F = R_a + \mathring{T}_a\tilde{N}_a' = R_b + \mathring{T}_b\tilde{N}_b' \tag{2.47}$$

and $\tilde{N}_a \cdot X_i > 0$, $\tilde{N}_b \cdot X_i > 0$, $i = 1, 2, \cdots, n$, hold true, then there exist two planes with $\tilde{N}_a$ and $\tilde{N}_b$ as normals, respectively, such that if they undergo motions represented by $(R_a, \mathring{T}_a)$ and $(R_b, \mathring{T}_b)$, respectively, they render the same pair of images (with image vectors X_i and X_i' at the two time instants) and the corresponding 3-D points are all located in the forward half space before and after the motion.

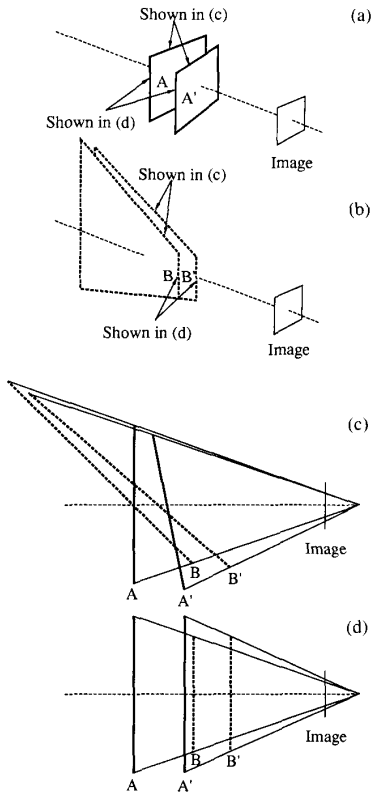


Fig. 2. Two solutions, veridical and illusive, render the same pair of images (see text). (a) Illustration of the setup and the veridical solution. (b) Illusive solution. (c) A combined top view. (d) A combined side view, with closer vertical edges shown only.

Proof: See Appendix B.

Fig. 2 shows an example where two solutions render the same pair of images with the positive depths for all the points. Four corners of a planar square board are used as feature points. Before motion, the board is orthogonal to the optical axis of the camera with the optical axis passing through its center. The motion consists of a rotation about the vertical axis by a small angle and a translation towards the camera. Fig. 2(a) shows the setup and motion of the board, where board A moves to board A' (both boards are shown with solid bold lines). In addition to the veridical interpretation (that agrees with what actually happened), another illusive interpretation (that does not agree with what actually happened) exists that is consistent with these two images. This illusive interpretation corresponds to board positions shown in Fig. 2(b) with dashed bold lines. That is, in Fig. 2(b) if board B moves to board B', we get exactly the same images as in Fig. 2(a). To show the relationships between these two interpretations more precisely, the top edges of the boards in these two interpretations are shown in Fig. 2(c) as a top view combining Figs. 2(a) and (b). Fig. 2(d) is the corresponding side view, where for each board only the vertical edge that is closer to the reader is shown for clarity. As can be seen from the figure, the image projections of

the corresponding corners of boards are the same for those two interpretations.

In some cases, we are able to reject the illusive solution on the ground that all the 3-D points should have positive depths. The illusive solution may have some recovered points in the back half space of the camera. In other words, one of the conditions $\tilde{N}_a \cdot X > 0$ and $\tilde{N}_b X > 0$ stated in Theorem 3 may be violated for some image points. However, it is proved in Theorem 3 that once these conditions for positive depths are satisfied before motion, positive depths are guaranteed after motion for the illusive solution as well. We have conducted a series of simulations, and we were often not able to reject the illusive solution based on the condition of positive depths. The example shown in Fig. 2 has positive depths for both solutions, and so the illusive solution cannot be rejected. If the plane of the illusive solution is so tilted that the intersections of the plane and the projection lines of some points are located in the back half space, as illustrated in Fig. 3, then the illusive solution can be rejected. □

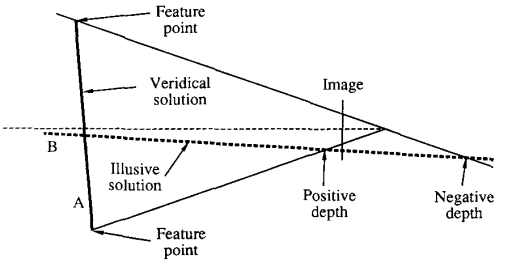


Fig. 3. If the illusive plane is very tilted, some of the recovered points are in the back half space.

D. Solving for Unknowns from F: Special Cases

Now let us turn to the cases where F^tF has multiple eigenvalues. We first derive the necessary and sufficient conditions for F^tF to have multiple eigenvalues, starting from the following lemma.

Lemma 1: Let the eigenvalues of F^tF be λ_1, λ_2, λ_3, with $\lambda_1 \le \lambda_2 \le \lambda_3$. Then

1) F^tF has multiple eigenvalues iff $\boldsymbol{T} \| \boldsymbol{RN}$,
2) $\lambda_1 = 1 < \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$ and $2\boldsymbol{RN} \cdot \boldsymbol{T} > -\|\boldsymbol{T}\|^2 \|\boldsymbol{N}\|^2$,
3) $\lambda_1 < 1 = \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$ and $2\boldsymbol{RN} \cdot \boldsymbol{T} < -\|\boldsymbol{T}\|^2 \|\boldsymbol{N}\|^2$,
4) $\lambda_1 = 1 = \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$ and $2\boldsymbol{RN} \cdot \boldsymbol{T} = -\|\boldsymbol{T}\|^2 \|\boldsymbol{N}\|^2$.

Proof: Notice that $\|\boldsymbol{N}\| \neq 0$, since $\boldsymbol{N}^t\boldsymbol{x} = 1$.

For 2: From (2.29), $\lambda_1 = 1 < \lambda_3$ iff $a^2 + b^2 = 0$ and $c > 1$. From (2.26), we have $a^2 + b^2 = 0$ and $c > 1$ iff $\boldsymbol{RV}_1 \cdot \boldsymbol{T} = 0$ and $\boldsymbol{RV}_2 \cdot \boldsymbol{T} = 0$ and $2\boldsymbol{RN} \cdot \boldsymbol{T} > -\|\boldsymbol{T}\|^2 \|\boldsymbol{N}\|^2$. But from (Ag.12), $\boldsymbol{RV}_1 \cdot \boldsymbol{T} = \boldsymbol{0}$ and $\boldsymbol{RV}_2 \cdot \boldsymbol{T} = \boldsymbol{0}$ iff $\boldsymbol{T} \| \boldsymbol{RN}$.

The proofs for cases 3 and 4 are similar to that of 2 and so are omitted. Case 1 immediately follows from 2–4. □

Lemma 1 gives algebraic conditions for the multiple eigenvalues to occur. Further investigation of these conditions leads to the following theorem.

Theorem 4. Let the eigenvalues of F^tF be λ_1, λ_2, λ_3, and $\lambda_1 \le \lambda_2 \le \lambda_3$. Then

1) F^tF has multiple eigenvalues iff $\boldsymbol{T} \| \boldsymbol{RN}$,
2) $\lambda_1 = 1 < \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$, and the absolute distance between the plane and the origin increases due to motion,
3) $\lambda_1 < 1 = \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$, and the absolute distance between the plane and the origin decreases due to motion,
4) $\lambda_1 = 1 = \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$, and the absolute distance between the plane and the origin does not change.

Before proving this theorem, we first discuss the corresponding geometrical meanings of this theorem. Remember that we represent the motion of the plane by a rotation followed by a translation. Fig. 4 shows the intermediate position of the plane that is rotated by the matrix $\boldsymbol{R}$. So the normal of this rotated plane is $\boldsymbol{RN}$. The final position (shown with dashed line in Fig. 4) is determined by the translation vector $\boldsymbol{T}$. If and only if the translation is aligned with the normal of the rotated plane, the multiple eigenvalues occur. The sign and the magnitude of the translation determine which case occurs. If the translation is such that the final absolute distance between the origin and the plane increases, case 2 occurs. The absolute distance can be increased in two ways, as shown in Fig. 4. The first is to translate the rotated plane further away from the origin, in which the translation and the normal of the rotated plane have the same directions. The second way is to translate the plane to the back of the camera until the absolute distance exceeds the original distance. Case 3 occurs if the absolute distance decreases, which can be done only by translating the camera towards the origin as shown in Fig. 4. Although the magnitude of translation can be so large that the plane passes the origin, the final absolute distance is less than the original distance in case 3. Finally, case 4 occurs if the absolute distance does not change. Only two situations are associated with this case: the translation vanishes or the plane is translated to its mirror position with respect to the origin.

Now we prove Theorem 4 from Lemma 1. According to the definition of $\boldsymbol{N}$ in (2.8) it follows that

$$\|\boldsymbol{N}\| = 1/d \tag{2.48}$$

where d is the absolute distance between the origin and the plane. Replacing $\|\boldsymbol{N}\|$ by $1/d$ in the inequalities in Lemma 1 and using the condition $\boldsymbol{T} \| \boldsymbol{RN}$ we can see that those inequalities are in fact constraints on the magnitude of translation in terms of d. From Lemma 1, what we need to prove are cases 2–4.

First consider case 2. $\boldsymbol{T} \| \boldsymbol{RN}$ includes two situations: $\boldsymbol{T}$ and $\boldsymbol{RN}$ have either the same directions or the opposite directions. a) Suppose the former is true, i.e., $\boldsymbol{RN} \cdot \boldsymbol{T} = \|\boldsymbol{N}\| \|\boldsymbol{T}\|$. The last inequality of case 2 in Lemma 1 im-

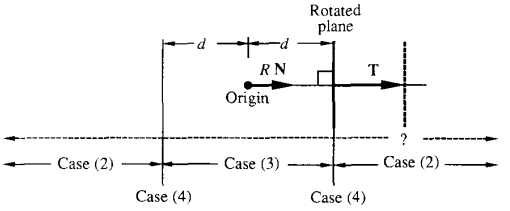


Fig. 4. Illustration of the necessary and sufficient conditions for the multiple eigenvalues to occur. Multiple eigenvalues occur if and only if $\boldsymbol{T}$ is parallel to $\boldsymbol{RN}$. The sign and the magnitude of $\boldsymbol{T}$ determine one of the cases 2 to 4 in Theorem 4.

plies $2 > -\|\boldsymbol{T}\|\,\|\boldsymbol{N}\|$, which is always true. b) Suppose $\boldsymbol{T}$ and $\boldsymbol{RN}$ have opposite directions, i.e., $\boldsymbol{RN} \cdot \boldsymbol{T} = -\|\boldsymbol{N}\|\,\|\boldsymbol{T}\|$. The last inequality of case 2 in Lemma 1 implies $2 < \|\boldsymbol{T}\|\,\|\boldsymbol{N}\|$, or equivalently $\|\boldsymbol{T}\| > 2d$. This means that if translation is in the direction towards the origin (see Fig. 4), the magnitude of translation must be large enough so that the absolute distance is increased. Summarizing these two situations, the condition $\boldsymbol{T} \| \boldsymbol{RN}$ and $2\boldsymbol{RN} \cdot \boldsymbol{T} > -\|\boldsymbol{T}\|^2\|\boldsymbol{N}\|^2$ in Lemma 1 is equivalent to the following condition: $\boldsymbol{T} \| \boldsymbol{RN}$ and the absolute distance between the plane and the origin increases due to motion.

Case 3 can be proved in a similar way: In case 3, $\boldsymbol{T}$ and $\boldsymbol{N}$ cannot have the same directions, otherwise the last inequality in case 3 of Lemma 1 leads to $\|\boldsymbol{T}\| < -2d$ which is impossible. If $\boldsymbol{T}$ and $\boldsymbol{N}$ have the opposite directions, the last inequality means $\|\boldsymbol{T}\| < 2d$. That is, the absolute distance must decrease.

Case 4 occurs if and only if either $\boldsymbol{T}$ vanishes (then conditions in Lemma 1 are trivially satisfied) or $\boldsymbol{T}$ does not vanish and $\boldsymbol{T} \| \boldsymbol{RN}$ and $2\boldsymbol{RN} \cdot \boldsymbol{T} = -\|\boldsymbol{T}\|^2\|\boldsymbol{N}\|^2$. In the latter situation with $\boldsymbol{T} \neq \boldsymbol{0}$, $\boldsymbol{T}$ and $\boldsymbol{N}$ cannot have the same directions since this leads to impossible equation $\|\boldsymbol{T}\| = -2d$. Now suppose that $\boldsymbol{T}$ and $\boldsymbol{N}$ have the opposite directions. The corresponding necessary and sufficient condition in Lemma 1 is then equivalent to $\|\boldsymbol{T}\| = 2d$. That is, the absolute distance does not change (see Fig. 4). □

In Fig. 4, if the final position of the rotated plane is located to the left of the origin, then the other side of the plane faces the camera after motion, since rotation does not change the side. Then from Theorem 4, we get the following corollary immediately.

Corollary 2: Assuming that the translation does not vanish and the same side of the plane faces camera before and after motion, only two cases are associated with multiple eigenvalues: a) $\lambda_1 = 1 < \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$, and $\boldsymbol{T}$ and $\boldsymbol{RN}$ have the same directions. b) $\lambda_1 < 1 = \lambda_3$ iff $\boldsymbol{T} \| \boldsymbol{RN}$, and $\boldsymbol{T}$ and $\boldsymbol{RN}$ have the opposite directions. □

Now we turn to the solutions for these special cases. If F^tF has an eigenvalue with multiplicity of 2 ($\lambda_1 = 1$ or $\lambda_3 = 1$ but not both), the two planes in (2.37) degenerate into one plane. This gives steps for cases 2 and 3, respectively, in the algorithm.

If F^tF has an eigenvalue of multiplicity of 3 ($\lambda_1 = \lambda_2 = \lambda_3$), case 4 occurs. Then it is clear from (Ag.8) that F is orthogonal. For $\boldsymbol{V}_i \in \Pi$, we still have $F\boldsymbol{V}_i = R\boldsymbol{V}_i$, $i = 1, 2$. Since F is orthogonal, we have

$$F\mathring{\boldsymbol{N}} = R\mathring{\boldsymbol{N}} + \boldsymbol{T}\|\boldsymbol{N}\| = \pm R\mathring{\boldsymbol{N}}. \tag{2.49}$$

If $\det(F) = 1$, $F\mathring{\boldsymbol{N}} = R\mathring{\boldsymbol{N}}$. Thus $R[\boldsymbol{V}_1 \;\; \boldsymbol{V}_2 \;\; \mathring{\boldsymbol{N}}] = F[\boldsymbol{V}_1 \;\; \boldsymbol{V}_2 \;\; \mathring{\boldsymbol{N}}]$. Therefore $R = F$. From $R\mathring{\boldsymbol{N}} = F\mathring{\boldsymbol{N}} = (R + \boldsymbol{T}\boldsymbol{N}^t)\mathring{\boldsymbol{N}} = R\mathring{\boldsymbol{N}} + \boldsymbol{T}\|\boldsymbol{N}\|$ and $\|\boldsymbol{N}\| \neq 0$, we get $\boldsymbol{T} = \boldsymbol{0}$. $\mathring{\boldsymbol{N}}$ cannot be determined.

If $\det(F) = -1$, $F\mathring{\boldsymbol{N}} = -R\mathring{\boldsymbol{N}}$. From (2.49) we have

$$\boldsymbol{T} = \frac{-2R\mathring{\boldsymbol{N}}}{\|\boldsymbol{N}\|}. \tag{2.50}$$

Then

$$F = R + \boldsymbol{T}\boldsymbol{N}^t = R - 2R\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t = R(I_3 - 2\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t). \tag{2.51}$$

Noticing $(I_3 - 2\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t)(I_3 - \mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t) = I_3$ for any unit vector $\mathring{\boldsymbol{N}}$, we get (Ag.17). From (2.41) and (2.50) we get (Ag.18).

Finally, we need to show that given any $\mathring{\boldsymbol{N}}$, R in (Ag.17) and $\tilde{\boldsymbol{T}}$ in (Ag.18) give a solution for $R + \tilde{\boldsymbol{T}}\mathring{\boldsymbol{N}}^t = F$. First, R in (Ag.17) is a rotation matrix, since $I_3 - 2\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t$ is a reflection with respect to a plane that has $\mathring{\boldsymbol{N}}$ as a normal and goes through the origin, and F is also a reflection. With R in (Ag.17) and $\tilde{\boldsymbol{T}}$ in (Ag.18), for any unit $\mathring{\boldsymbol{N}}$ we have

$$R + \boldsymbol{T}\boldsymbol{N}^t = R + \tilde{\boldsymbol{T}}\mathring{\boldsymbol{N}}^t = R - 2R\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t = R(I_3 - 2\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t)$$

$$= F(I_3 - 2\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t)(I_3 - 2\mathring{\boldsymbol{N}}\mathring{\boldsymbol{N}}^t) = F.$$

III. Inherent Uniqueness

In this section, we present the inherent uniqueness of the problem. The uniqueness investigated here is inherent, because it addresses uniqueness of the problem itself, not just for a particular algorithm. The uniqueness of the problem exists by itself no matter whether we have an algorithm to solve the problem or not.

Theorem 3 states that under two conditions, a) if the matrix F satisfies (2.18) for all the points, or equivalently (2.20), and b) the matrix F is decomposable as in (2.47), then there exists a corresponding plane and motion which render the same images, aside from possibly a few negative depths. In solving (2.20) the vector $\boldsymbol{h}$ is regarded as a free vector. Therefore, it is not clear up to now whether such a solution $\boldsymbol{h}$, using noisy image points, will always result in a decomposable F. This point is closely related to the inherent uniqueness of the problem.

A. Decomposability

The question is whether any 3×3 matrix, say F_s, is always decomposable, i.e., there exists a scale factor k, a rotation matrix R, a vector $\tilde{\boldsymbol{T}}$, and a unit vector $\mathring{\boldsymbol{N}}$ such

that

$$F_s = k(R + \tilde{T}\mathring{N}^t) \tag{3.1}$$

is exactly true (the matrix F_s can be considered as a matrix computed from $\boldsymbol{h}$ in (2.20)). In decomposition (3.1), one can further impose that the third component of $\mathring{N}$ is nonnegative by absorbing a possible sign into $\tilde{T}$. Note that the form of decomposition (3.1) can always interchange with the decomposition

$$F_s = k(R + \mathring{T}\tilde{N}^t). \tag{3.2}$$

Since when $\tilde{T} = 0$ in (3.1), $F = R$ and $\tilde{N}$ in (3.2) is replaced by a zero vector and $\mathring{T}$ can be any unit vector. It is also true from (3.2) to (3.1): a zero $\tilde{N}$ in (3.2) leads to a zero $\tilde{T}$ in (3.1).

It can be seen that there are exactly 9 degrees of freedom on the right-hand side of (3.1): one for k, three for R, three for $\tilde{T}$, and two for $\mathring{N}$. In other words, a decomposable matrix has the same number of degrees of freedom as any 3×3 matrix. This suggests that it might be true that any 3×3 matrix is decomposable. In fact, we have the following new theorem.

Theorem 5: Any 3×3 matrix F_s that has a rank no less than 2 is decomposable, i.e., there exists a scale factor k, a 3×3 rotation matrix R, a 3-D vector $\tilde{T}$, and a unit 3-D vector $\mathring{N}$ such that (3.1) holds exactly (similar for (3.2)).

Proof: Suppose a matrix F_s is given whose rank is no less than two. Then the second smallest eigenvalue γ_2 of $F_s^t F_s$ is positive. Let $k = \sqrt{\gamma_2}$, we define $F = k^{-1}F_s$. Therefore, the second smallest eigenvalue, λ_2, of F^tF is equal to 1. We have the following decomposition:

$$H^tF^tFH = \text{diag}\,(\lambda_1, 1, \lambda_3) \tag{3.3}$$

where H is an orthogonal matrix and $\lambda_1 \le 1 \le \lambda_3$. To prove the decomposability, what we need to prove is that we can define a rotation matrix R, a vector $\tilde{T}$, and a unit vector $\mathring{N}$ such that

$$F = R + \tilde{T}\mathring{N}^t \tag{3.4}$$

since this gives (3.1) immediately.

To do this, first consider the case where $\lambda_1 < \lambda_3$. From (3.3) define two numbers

$$\alpha = \sqrt{\frac{\lambda_3 - 1}{\lambda_3 - \lambda_1}}, \qquad \beta = \sqrt{\frac{1 - \lambda_1}{\lambda_3 - \lambda_1}} \tag{3.5}$$

from which define two orthonormal vectors: $\boldsymbol{U}_1 = (\alpha, 0, \beta)^t$ and $\boldsymbol{U}_2 = (0, 1, 0)^t$. Then, multiplying by the orthogonal matrix H, we get another pair of orthonormal vectors

$$\boldsymbol{V}_1 = H\boldsymbol{U}_1, \qquad \boldsymbol{V}_2 = H\boldsymbol{U}_2. \tag{3.6}$$

Determine a rotation matrix R by

$$R\boldsymbol{V}_i = F\boldsymbol{V}_i, \qquad i = 1, 2. \tag{3.7}$$

Is there a unique rotation matrix R that exactly satisfies (3.7)? The answer is positive. To prove this, it is equivalent to prove that $F\boldsymbol{V}_1$ and $F\boldsymbol{V}_2$ are still orthonormal vectors. It follows from (3.3) that

$$\begin{aligned}(F\boldsymbol{V}_i)^t(F\boldsymbol{V}_j) &= \boldsymbol{U}_i^tH^tF^tFH\boldsymbol{U}_j \\ &= \boldsymbol{U}_i^t\,\text{diag}\,(\lambda_1, 1, \lambda_3)\boldsymbol{U}_j = \delta(i - j)\end{aligned} \tag{3.8}$$

for $i = 1, 2$, where $\delta(k) = 1$ if $k = 0$ and $\delta(k) = 0$ otherwise. Therefore, there is a unique rotation matrix that rotates these two orthonormal vectors exactly to the other two orthonormal vectors so that

$$R\boldsymbol{V}_i = F\boldsymbol{V}_i \tag{3.9}$$

exactly holds for $i = 1, 2$.

Then we define a unit vector $\mathring{N}$ by

$$\mathring{N} = \pm\boldsymbol{V}_1 \times \boldsymbol{V}_2 \tag{3.10}$$

where the sign is such that the third component is nonnegative. The vector $\tilde{T}$ is defined by

$$\tilde{T} = F\mathring{N} - R\mathring{N}. \tag{3.11}$$

Equations (3.9) and (3.10) lead to

$$(R + \tilde{T}\mathring{N}^t)\boldsymbol{V}_i = R\boldsymbol{V}_i = F\boldsymbol{V}_i \tag{3.12}$$

for $i = 1, 2$. Equation (3.11) gives

$$(R + \tilde{T}\mathring{N}^t)\mathring{N} = R\mathring{N} + \tilde{T} = R\mathring{N} + (F\mathring{N} - R\mathring{N}) = F\mathring{N}. \tag{3.13}$$

Since $\boldsymbol{V}_1$, $\boldsymbol{V}_2$, and $\mathring{N}$ are orthonormal, it follows from (3.12) and (3.13) that

$$F = R + \tilde{T}\mathring{N}^t. \tag{3.14}$$

Therefore, F is decomposable.

The remaining case is that $\lambda_1 = \lambda_3 = 1$. From (3.3) we know that F is an orthogonal matrix. If F is a rotation matrix, $\det(F) = 1$, let $R = F$, $\tilde{T} = 0$, and (3.4) holds true for any unit $\mathring{N}$. Otherwise F is a reflection matrix. Pick any unit vector $\mathring{N}$ with a positive third component. Define

$$R = F(I - 2\mathring{N}\mathring{N}^t)$$

which is a rotation matrix (reflection times reflection), and

$$\tilde{T} = -2R\mathring{N}.$$

We have

$$R + \tilde{T}\mathring{N}^t = F(I - 2\mathring{N}N^t)(I - 2\mathring{N}\mathring{N}^t) = F$$

since $(I - 2\mathring{N}\mathring{N}^t)(I - 2\mathring{N}\mathring{N}^t) = I$. Thus we can always find a decomposition as in (3.4). □

B. The Rank Condition of A is Algorithm Independent

Theorem 5 leads to a very important conclusion: The condition rank $(A) = 8$ is not only a sufficient condition

for the algorithm to reach two solutions (except the special cases in Theorem 4), it is in fact an algorithm-independent necessary condition! We have the following new theorem.

Theorem 6: There exists the exact number of solutions as given by the linear algorithm, if and only if rank $(A) = 8$.

Proof: We know that rank $(A) = 8$ is a sufficient condition for the algorithm. So it is a sufficient condition for the problem itself. What we need to prove is that if rank $(A) \neq 8$ there exist infinitely many solutions that render the same images.

Certainly rank $(A) > 8$ is impossible for coplanar points undergoing a rigid motion since $\boldsymbol{h}$ in (2.20) is a nonzero vector. Suppose rank $(A) < 8$. Then in (2.20) there is at least a solution $\boldsymbol{h}$ that corresponds to the true motion, and another solution $\boldsymbol{h}'$, which is linearly independent of $\boldsymbol{h}$. For any real number k, $\boldsymbol{h}(k) \triangleq \boldsymbol{h} + k\boldsymbol{h}'$ is also a solution of (2.20). Let $F_s(k)$ be the matrix determined from $\boldsymbol{h}(k)$. $F_s(0)$ corresponds to the true solution and so it is decomposable, and the rank of $F_s(0)$ is no less than two. So, $F_s(0)$ has at least one 2×2 minor determinant that is not equal to zero. Since a determinant is continuous with respect to its elements, there exists a sufficiently small positive number $\delta > 0$, such that for any k satisfying $|k| < \delta$, the corresponding minor determinant of $F_s(k)$ is also nonzero. That is, for any k, $|k| < \delta$, the matrix $F_s(k)$ has a rank no less than two. Hence these infinitely many $F_s(k)$ are decomposable. On the other hand, there is no number $c(k)$ so that $F_s(k) = c(k)F_s(0)$ holds true, because $\boldsymbol{h}$ and $\boldsymbol{h}'$ are linearly independent. Therefore, there are infinitely many matrices $F(k)$, which are not scaled versions of $F_s(0)$, so that the median eigenvalue of $F(k)^t F(k)$ is a unit. Finally, from Theorem 3 we know that all these $F(k)$ give infinitely many solutions for motion and structure which render the same pair of images. □

C. The Fundamental Theorem

Now we have established that rank $(A) = 8$ is not only sufficient to the algorithm, it is in fact a necessary condition to determine at most two solutions (except for case 4 in Theorem 4). We summarize the results in the following fundamental theorem.

Fundamental Theorem: Given two perspective views of a rigidly moving planar surface, the minimum number of solutions of $(R, \mathring{T}, \tilde{N})$ falls into the following cases according to the true motion and true position of the surface:

Case 1: $\boldsymbol{T} \times \boldsymbol{RN} \neq \boldsymbol{0}$: the minimum number of solutions is 2.

Case 2: $\boldsymbol{T} \| \boldsymbol{RN}$, and the absolute distance between the plane and the origin increases due to motion: the minimum number of solutions is 1.

Case 3: $\boldsymbol{T} \| \boldsymbol{RN}$, and the absolute distance between the plane and the origin decreases due to motion: the minimum number of solutions is 1.

Case 4: $\boldsymbol{T} \| \boldsymbol{RN}$, and the absolute distance between the plane and the origin does not change, which includes two subcases:

Subcase 4a: $\boldsymbol{T} = 0$: The minimum number of solutions for $(R, \boldsymbol{T})$ is 1, but infinite for $\tilde{N}$.

Subcase 4b: The origin is located halfway between the rotated-only plane and the fully moved plane: the minimum number of solutions is infinite.

All the above-mentioned solutions render the same images. Some of the solutions may force some points to be reconstructed in the back half-space, in which case one may identify that the corresponding solution is illusive. Whenever the number of existing solutions reaches the above minimum, the linear algorithm guarantees to provide exactly these solutions. The necessary and sufficient conditions for the number of solutions to reach the above minimum can be given in the following three ways:

A) In terms of the rank of A: rank $(A) = 8$.

B) In terms of the n points in 2-D image plane: there exists a set of four object points such that no image projections of any three points in this set are collinear in any of the two images.

C) In terms of the n points in 3-D surface: there exists a set of four points in the object plane such that no three points in this set are collinear in the object plane and if the object plane is extended, it does not go through the projection center of the camera before and after motion.

Proof: Since our above discussion has established many related results, the proof of this fundamental theorem can be concise. Theorem 3 concludes that the solutions derived in Section II all satisfy (2.18) and that they all render the same images. Theorem 4 reveals under which configurations the different cases occur. Rank $(A) = 8$ is the sufficient condition for the linear algorithm to give these minimum numbers of solutions but Theorem 6 states that this condition is necessary, for the problem itself, to restrict the number of existing solutions to this minimum. The geometrical necessary and sufficient conditions given by Theorem 1 and Corollary 1 are then not only necessary and sufficient for rank $(A) = 8$ but also necessary and sufficient for the problem itself, independent of the specific algorithm. □

We call the above theorem the fundamental theorem because it addresses the intrinsic properties of the problem, not just of a particular algorithm. Before the establishment of Theorem 6, those properties were only applicable to the linear algorithm. The fundamental theorem not only answers the inherent uniqueness question of this problem, but also gives a complete list of all the cases, each of which is associated with the clear geometrical conditions. Best of all, we know that all can be solved by our linear algorithm. Therefore, if the linear algorithm fails to give these solutions, so does any algorithm.

D. Plane-Perceivable Surfaces

In this subsection, we investigate the problem from a different perspective. Previously, we assumed that the points are coplanar and they undergo a rigid motion. Now,

suppose that we do not know that the points are coplanar. A question arises as to how one can detect in which cases, from image points, the points are coplanar or not. The answer to this problem is made possible by our solution to the inherent uniqueness.

According to Theorems 3 and 5 we know that if rank $(A) \leq 8$, there exist solutions, each one corresponding to coplanar points undergoing a rigid motion, no matter whether the actual points are coplanar or not. (Sometimes the solution may result in some negative depths, but for the discussion in this subsection, we put this aside.) Therefore we have no solution only when rank $(A) = 9$. The following theorem gives a necessary and sufficient condition on the configuration of points to result in rank $(A) \leq 8$, or, equivalently, according to the fundamental theorem, there exists an interpretation that a planar patch undergoes a rigid motion.

Theorem 7: From two views, there exists an interpretation that a planar patch undergoes a rigid motion, if and only if there exists a 3×3 nonzero matrix K such that all the observed points lie in the following intersection of two quadratic surfaces before motion:

$$(\boldsymbol{x} - \boldsymbol{O}') \times K\boldsymbol{x} = \boldsymbol{0} \tag{3.15}$$

where $\boldsymbol{O}' = -R'\boldsymbol{T}$.

Proof: See Appendix C.

It is clear that the three scalar equations in (3.15) are quadratic in the components of $\boldsymbol{x}$. So each equation determines a quadratic surface. It is easy to prove that in (3.15) there are just two scalar equations that are independent. Therefore, (3.15) determines the intersection of two quadratic surfaces. Such an intersection gives a curve, in general, with plane as a special case. But we will still use the more general term "surface." The origin and $\boldsymbol{O}'$ are in this intersection. If we consider the object as stationary and the motion is due to that of the camera, the origin and $\boldsymbol{O}'$ are the positions of the camera at times t_1 and t_2, respectively.

We know that there is a solution that gives coplanar 3-D points if and only if rank $(A) \leq 8$. According to the above theorem, if and only if the points lie on a surface in the form of (3.15), the images can be interpreted as the projections of a set of coplanar points undergoing a rigid motion. Therefore we call the surfaces in (3.15) plane-perceivable surfaces. Given two images, if the 3-D points are from such a surface, one cannot tell if they are coplanar or not and can always interpret them as coplanar. If and only if they do not lie on a plane-perceivable surface in (3.15), one can detect that the points are not coplanar by checking rank $(A) > 8$. Fig. 5 illustrates the relations among planes, plane-perceivable surfaces, and the rest of the detectable nonplanar surfaces. We will see that plane-perceivable surfaces include many nonplanar surfaces. This is due to the fact that image projections do not provide complete information about what happened in 3-D.

Now we investigate how many points are needed so that it is possible for the points to not lie on any plane-perceivable surface. It is easy to see that the scalar equations in (3.15) are linear in the elements of K. Each point on the surface, other than the origin and $\boldsymbol{O}'$, gives at most two independent linear equations in the elements of K (since one scalar equation in (3.15) is redundant). Any set of 4 points on the surface gives 8 such linear equations for the elements in K, and a nonzero solution K exists. Therefore, four points always lie on some plane-perceivable surface in (3.15). Since four points are generally not coplanar, the set of plane-perceivable surfaces does include many noncoplanar surfaces (including curves). Similarly, we conclude that 5 points generally do not lie on any surface in the form of (3.15). Therefore, with 5 points it is possible to detect that they are not coplanar.

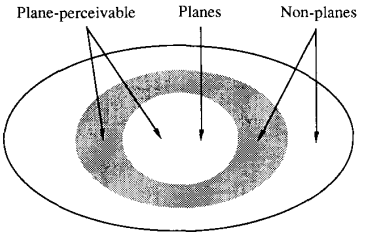


Fig. 5. The relations among planes, plane-perceivable surfaces, and nonplanar surfaces. Plane-perceivable surfaces include some of the nonplanar surfaces (including curves) and all the visible planes.

E. Optimality

Now we turn to the stability of the solution in the presence of noise. Theorem 5 is not only important to the inherent uniqueness, it is also important to understand the optimality of the linear algorithm. From Theorem 5, we know that a noise-contaminated version of F_s is decomposable, as long as the noise is not so large and coincidental that the contaminated matrix happens to degenerate into a rank-one matrix. Then the algorithm computes the solutions that fit this F_s exactly. In other words, no constraint needs to be considered in the solution for F_s, since the probability for the noise contaminated F_s to be a rank-one matrix is zero.

From (2.18)–(2.20) and Theorem 5, it immediately follows that our algorithm determines a rotation matrix R, a unit 3-D vector $\mathring{\boldsymbol{T}}$, and a 3-D vector $\tilde{\boldsymbol{N}}$ such that

$$\sum_{i=1}^{n} \| \boldsymbol{X}_i' \times (R + \mathring{\boldsymbol{T}}\tilde{\boldsymbol{N}}')\boldsymbol{X}_i \|_{z=0}^2 = \min \tag{3.16}$$

where $\| \cdot \|_{z=0}$ denotes the Euclidean norm of the projection onto x–y plane: $\|(a, b, c)\|_{z=0}^2 = a^2 + b^2$. If all three rows in (2.19) are included in A in (2.20) used for the algorithm, then $\| \cdot \|_{z=0}$ in (3.16) is replaced by the regular Euclidean norm $\| \cdot \|$. In the absence of noise, the minimum value in (3.16) is equal to zero. Theorem 3 implies that the images can be interpreted as a planar patch undergoing a rigid motion if and only if

$$\boldsymbol{X}_i' \times (R + \mathring{\boldsymbol{T}}\tilde{\boldsymbol{N}}')\boldsymbol{X}_i = \boldsymbol{0} \tag{3.17}$$

is satisfied for every image point pair $(\boldsymbol{X}_i, \boldsymbol{X}_i')$. Therefore, the objective function (3.16) uses all the constraint in the

problem, except the constraint on negative depth which is automatically satisfied for the veridical solution.

Since the variance of the equation residual in (3.17) varies from point to point, the minimum variance estimator requires that the objective function in (3.16) be weighted by the inverse of the residual variance. Let $E' = [e'_{ij}] \triangleq R + \mathring{T}\tilde{N}^t$, and $M = [m_{ij}]$ be a matrix constructed from E' so that $m_{11} = e'_{21}$, $m_{12} = -e'_{11}$, $m_{21} = e'_{22}$, $m_{22} = -e'_{12}$ and the remaining entries of M_i are zeros. Then the computed weight w_i^2 can be expressed concisely by $w_i^2 = \sigma^{-2}(\|E'X_i\|_{z=0}^2 + \|MX'_i\|^2)^{-1}$, assuming the error in different components of image plane points is uncorrelated and has a zero mean and a constant variance σ^2. The objective function in (3.16) can be replaced by the optimal one:

$$\sum_{i=1}^{n} w_i^2 \|X'_i \times (R + \mathring{T}\tilde{N}^t)X_i\|^2 = \min. \qquad (3.18)$$

However, the weight w_i^2 is a function of unknown motion parameters and structure. The solution of (3.18) cannot be solved in a closed form and so an iterative process needs to be used to minimize the left-hand side of (3.18). A more detailed treatment of optimal motion and structure estimation under general surfaces is presented in [17]. Compared to (3.18), the objective function in (3.17) is simpler and can be directly computed using the linear algorithm. Based on the properties of the linear algorithm we discussed, as well as the result of the simulation, the fast solution of (3.16) is expected to be good enough for most applications.

F. Three-View Problem

We know that two views allow two interpretations in general: one is veridical, the other is illusive. In this subsection, we investigate how to reject the illusive interpretation.

Suppose that in addition we have observed an extra point that we know *a priori* does not belong to the plane. Then, generally, we can reject the illusive solution as follows. Using the image points that correspond to coplanar points, we determine two solutions. If the solution $(R, \mathring{T})$ is veridical, from (2.5) we know that X', RX, and $\mathring{T}$ are coplanar for all the point pairs (X, X') including the extra point not in the plane. However, if the solution is illusive, this will not, in general, be true for the extra point. This fact can be used to reject the illusive solution. However, since it is unlikely for one to have *a priori* knowledge that a specified point is not coplanar with the rest, this method is not very useful in practice.

Fortunately, provided with three distinct images of coplanar points that undergo a rigid motion, in general, we can uniquely determine the motion paramaters and the normalized normal of the object plane. Suppose that we have three image frames taken at time instants t_1, t_2, and t_3, respectively. Consider the two motions: one from t_2 to t_1 and the other from t_2 to t_3. We solve for the motion parameters of these two motions and the unit normal $\mathring{N}$ of the object plane at time t_2. The veridical solutions for these two motions should have the same answer for the unit normal $\mathring{N}$ of the object plane at time t_2. According to the following theorem, the illusive solutions generally do not agree on the unit normal of the object plane. Therefore, we can determine which solution is veridical based on the agreement on the unit normal at time t_2.

Theorem 8: Let F and F' correspond, respectively, to two motions from one positon. Suppose there exist two solutions $(R_a, \tilde{T}_a, \mathring{N}_a)$ and $(R_b, \tilde{T}_b, \mathring{N}_b)$ such that

$$F = R_a + \tilde{T}_a\mathring{N}_a^t = R_b + \tilde{T}_b\mathring{N}_b^t. \qquad (3.19)$$

Let H be an orthogonal matrix with $\det(H) = 1$ such that

$$H^tF^tFH = \operatorname{diag}(\lambda_1, 1, \lambda_3) \qquad (3.20)$$

and $\lambda_1 < 1 < \lambda_3$.

Similarly, for F' we have two solutions such that

$$F' = R'_a + \tilde{T}'_a(\mathring{N}'_a)^t = R'_b + \tilde{T}'_b(\mathring{N}'_b)^t.$$

Then, both solutions agree on the unit normal, i.e.,

$$\mathring{N}_a = \mathring{N}'_a, \quad \text{and} \quad \mathring{N}_b = \mathring{N}'_b \qquad (3.21)$$

if and only if the ambiguity condition is satisfied, i.e., there exists an orthogonal matrix Q and a positive number k such that

$$F' = Q \operatorname{diag}\left(\sqrt{1 - k(1 - \lambda_1)}, \quad 1, \quad \sqrt{1 + k(\lambda_3 - 1)}\right)H^t \qquad (3.22)$$

with $1 - k(1 - \lambda_1) \geq 0$ or

$$F' = Q \operatorname{diag}(\sqrt{1 + k(1 - \lambda_1)}, \quad 1, \quad \sqrt{1 - k(\lambda_3 - 1)})H^t \qquad (3.23)$$

with $1 - k(\lambda_3 - 1) \geq 0$.

Proof: See Appendix D.

Generally, the ambiguity condition is not satisfied according to the following observation. We know that $F = R + \mathring{T}\tilde{N}^t$ has 8 degrees of freedom (3 in R, 3 in $\tilde{T}$, and 2 in $\mathring{N}$). With the fixed unit normal $\mathring{N}$ the degree of freedom of F' is $8 - 2 = 6$. The ambiguity condition restricts the degree of freedom for F', with fixed unit normal $\mathring{N}$, to 4 (3 in Q and one in k). In other words, the ambiguity condition reduces the degree of freedom for F' by two.

As an example, it is easy to see that the illusive solutions for the two motions agree on the unit normal if the motions correspond to (R, T) and (QR, QT), respectively, where Q is any rotation matrix. That is, there are many cases where three views do not give a unique solution. But if we consider all three possible interframe motions among three views and examine the agreement on the object normal at three time instants, more cases can have a unique solution [11].

Given a sequence of images, as long as there exist three consecutive frames that do not satisfy the ambiguity condition, all the veridical interframe motions through the sequence can be determined. In fact, we know that ve-

ridical solutions can be identified from two candidate solutions if either veridical motion is known or veridical unit normal of the object plane is known. Therefore, for example, the veridical unit normal at time t_1 can be used to identify the veridical motion from t_1 to t_2, which, in turn, can be used to identify the veridical unit normal at t_2, and so on. Such a propagation of the veridical solutions through the sequence determines the veridical solution of every interframe motion and the orientation of object plane at every time instant.

IV. Error Estimation

In the previous section, we presented our motion analysis algorithm. In the presence of noise, the accuracy or reliability of the solutions given by the algorithm is of great concern. In practice, the numerical solutions to a problem often cannot be used if it is not known how accurate they are. In this section, we briefly present our approach to error estimation.

A. Some Basic Concepts

In the area of numerical analysis, usually a bound on the worst case error is derived for a given numerical algorithm. However, a worst case bound is usually too large to be useful here, because the algorithm involves many complicated computations and considerably large measurement errors are present. In other words, a worst case bound is very large but is almost never reached. Therefore, statistical estimates of the errors are more useful than a worst case bound.

The sensitivity of solution to the errors in the measurement depends very much on the configuration. For example, under a nearly degenerate point configuration, small errors in the measured data will cause much larger solution errors compared to those under a stable point configuration. Formally, let the image coordinates of all the points be represented by I, and the errors in the image coordinates of these points be represented by a random variable ϵ. The error e in the estimated motion parameters is a function of I and ϵ. Denoting this function by f, we can write $e = f(I, \epsilon)$. The objective of this section is to estimate the error e given the images I. However, since ϵ is random, and consequently, so is e. One cannot determine e completely. It is, therefore, more appropriate to determine the expected error. If we can estimate the standard deviation of e (with ϵ as a random variable) given the noise-corrupted image I, we can use it to indicate the reliability of the solution. The images I corresponding to a degenerate or nearly degenerate spatial configuration should give large estimates of e, and that corresponding to a stable configuration should give small estimates. Although we have derived the necessary condition of degenerate configuration in Theorem 1, this condition cannot be directly used to quantitatively assess the accuracy of a solution. The method to be discussed below gives a systematic way of estimating errors.

For the sake of conciseness, we use the following notation: A matrix A without noise is denoted by A itself and its elements denoted by the corresponding small letters, a_{ij}, i.e., $A = [a_{ij}]$. The noise-corrupted A is denoted by $A(\epsilon)$, and the noise matrix of A with the same size as A is denoted by Δ_A: $A(\epsilon) = A + \Delta_A$. Similarly, for vectors, we use $\boldsymbol{\delta}$ with a corresponding subscript to denote the noise vectors: $\boldsymbol{X}(\epsilon) = \boldsymbol{X} + \boldsymbol{\delta}_x$. Γ with a subscript is used to denote the autocovariance matrix of the vector (if only the first-order errors are considered, the means of the errors are zero): $\Gamma_x = \boldsymbol{E}(\boldsymbol{\delta}_x \boldsymbol{\delta}_x^t)$ where $\boldsymbol{E}$ denotes expectation. A matrix $\boldsymbol{A} = [\boldsymbol{A}_1 \ \boldsymbol{A}_2 \ \cdots \ \boldsymbol{A}_n]$ is associated with a corresponding vector $\boldsymbol{A}$ as $\boldsymbol{A} = [\boldsymbol{A}_1^t, \cdots, \boldsymbol{A}_n^t]^t$. Similarly, Γ_A denotes the corresponding covariance matrix of the vector $\boldsymbol{A}$ associated with the matrix A. δ_A denotes the perturbation vector associated with the perturbation matrix Δ_A.

Assuming two variables a and b with small errors: $a(\epsilon) = a + \boldsymbol{\delta}_a$, and $b(\epsilon) = b + \boldsymbol{\delta}_b$, we have

$$a(\epsilon)b(\epsilon) = ab + \boldsymbol{\delta}_a b + a\boldsymbol{\delta}_b + \boldsymbol{\delta}_a \boldsymbol{\delta}_b \triangleq ab + \boldsymbol{\delta}_{ab}. \tag{4.1}$$

The errors in $a(\epsilon)b(\epsilon)$ are

$$\boldsymbol{\delta}_{ab} = \boldsymbol{\delta}_a b + a\boldsymbol{\delta}_b + \boldsymbol{\delta}_a \boldsymbol{\delta}_b \cong \boldsymbol{\delta}_a b + a\boldsymbol{\delta}_b. \tag{4.2}$$

In the last approximation we keep the linear terms of the error (first-order perturbation) and ignore the higher order terms. Later in this paper we use the sign $\cong$ for the equations that are equal in the linear terms ($\approx$ for the approximate equality in the usual sense).

B. Perturbation of Eigenvalues and Eigenvectors

In our problem, the noise or error arises from the image coordinates. The solutions are the motion parameters and the normalized normal calculated by our motion analysis algorithm, which involves the calculation of eigenvalues and eigenvectors of a symmetrical matrix. With small perturbation in the matrix, we need to know the corresponding perturbation in its eigenvalues and eigenvectors. We have the following theorem.

Theorem 9: Let $A = [a_{ij}]$ be an $n \times n$ symmetrical matrix and H be an orthonormal matrix such that

$$H^t A H = \text{diag}\,(\lambda_1, \lambda_2, \cdots, \lambda_n). \tag{4.3}$$

Let the eigenvalues be ordered according to increasing magnitudes. Without loss of generality, consider the eigenvalue λ_1. Assume λ_1 is a simple eigenvalue:

$$\lambda_1 < \lambda_2 \leq \lambda_3 \leq \cdots \leq \lambda_n. \tag{4.4}$$

Let

$$H = [\boldsymbol{h}_1 \ \boldsymbol{h}_2 \ \cdots \ \boldsymbol{h}_n] \tag{4.5}$$

and $\boldsymbol{X}$ be an eigenvector of A associated with λ_1. $\boldsymbol{X}$ is then a vector in span $(\boldsymbol{h}_1)$ (the linear space spanned by $\boldsymbol{h}_1$). Let $\boldsymbol{X}(\epsilon)$ be the eigenvector of the perturbed matrix $A(\epsilon) = A + \Delta_A$ associated with the perturbed eigenvalue $\lambda_1(\epsilon)$. $\boldsymbol{X}(\epsilon)$ can be written as

$$\boldsymbol{X}(\epsilon) = \boldsymbol{X} + \boldsymbol{\delta}_x \tag{4.6}$$

with $\delta_x \in \text{span}(\boldsymbol{h}_2, \boldsymbol{h}_3, \cdots, \boldsymbol{h}_n)$. Letting ϵ be the maximum absolute value of the elements in $\Delta_A = [\delta_{a_{ij}}]$, we have

$$\Delta_A = \epsilon B \tag{4.7}$$

where $B = [b_{ij}]$, with $b_{ij} = \delta_{a_{ij}}/\epsilon$. Therefore, $|b_{ij}| \le 1$, $1 \le i \le n$, $1 \le j \le n$. Then for sufficiently small ϵ, the perturbation of λ_1 can be expressed by a convergent series in ϵ:

$$\delta_{\lambda_1} \triangleq \lambda_1(\epsilon) - \lambda_1 = p_1\epsilon + p_2\epsilon^2 + p_3\epsilon^3 + \cdots \tag{4.8}$$

and the perturbation vector $\boldsymbol{\delta}_x$ can be expressed by a convergent vector series in the space span $(\boldsymbol{h}_2, \boldsymbol{h}_3, \cdots, \boldsymbol{h}_n)$. In other words, letting $H_2 = [\boldsymbol{h}_2, \boldsymbol{h}_3, \cdots, \boldsymbol{h}_n]$, then for sufficiently small positive ϵ, there exist $(n-1)$-dimensional vectors $\boldsymbol{g}_1, \boldsymbol{g}_2, \boldsymbol{g}_3, \cdots$ such that

$$\boldsymbol{\delta}_x = \epsilon H_2\boldsymbol{g}_1 + \epsilon^2 H_2\boldsymbol{g}_2 + \epsilon^3 H_2\boldsymbol{g}_3 + \cdots. \tag{4.9}$$

The linear term (in ϵ) in (4.9) is given by

$$p_1\epsilon = \boldsymbol{h}_1^t \Delta_A \boldsymbol{h}_1 \tag{4.10}$$

and the linear term (in ϵ) is given by

$$\epsilon H_2\boldsymbol{g}_1 = H \Delta H^t \Delta_A \boldsymbol{X} \tag{4.11}$$

where

$$\Delta = \text{diag}\,(0, (\lambda_1 - \lambda_2)^{-1}, \cdots, (\lambda_1 - \lambda_n)^{-1}). \tag{4.12}$$

That is, suppressing the second and higher order terms (considering first-order perturbation), for the eigenvalue we have

$$\delta_{\lambda_1} \cong \boldsymbol{h}_1^t \Delta_A \boldsymbol{h}_1 \tag{4.13}$$

and for the eigenvector:

$$\boldsymbol{\delta}_x \cong H \Delta H^t \Delta_A \boldsymbol{X}. \tag{4.14}$$

The proof of this theorem is presented in [19]. □

Using the above theorem, the first-order perturbation of the simple eigenvalue λ_1 and the eigenvector associated with λ_1 can be determined from the perturbation of the matrix. A similar result holds for other simple eigenvalues. For example, the first-order perturbation of the simple eigenvalue λ_2 is given by

$$\delta_{\lambda_2} \cong \boldsymbol{h}_2^t \Delta_A \boldsymbol{h}_2 \tag{4.15}$$

and the first-order perturbation vector of eigenvector $\boldsymbol{X}_2$ associated with the simple eigenvalue λ_2 is given by

$$\boldsymbol{\delta}_{x_2} \cong H \Delta \boldsymbol{H}^t \Delta_A \boldsymbol{X}_2 \tag{4.16}$$

where Δ should be

$$\Delta = \text{diag}\,((\lambda_2 - \lambda_1)^{-1}, 0, (\lambda_2 - \lambda_3)^{-1}, \cdots, (\lambda_2 - \lambda_n)^{-1}). \tag{4.17}$$

If the covariance of the random perturbation Δ_A can be estimated, the corresponding covariance of the perturbation in the eigenvalue and the eigenvectors of A can be estimated, up to the first-order perturbation, based on the above results.

C. Error Estimation for the Motion Analysis Algorithm

In steps i, and ii, the algorithm computes the eigenvalues and the eigenvectors of the corresponding matrices. What we need to do is to estimate the perturbation of the corresponding matrices from the perturbation in the image coordinates. We use the first-order approximation to estimate these perturbations in the matrices.

For Step i: Assume the components of the image vector $\boldsymbol{X}_i = (u_i, v_i, 1)$ and $\boldsymbol{X}_i' = (u_i', v_i', 1)$ have errors. (The third component 1 in each image vector is accurate.) Let u_i, v_i, u_i', and v_i' have additive errors δ_{u_i}, δ_{v_i}, $\delta_{u_i'}$, and $\delta_{v_i'}$, respectively, for $1 \le i \le n$. From (Ag.1) we get

$$\Delta_A^t = \begin{bmatrix} \delta_{u_1} & 0 & \cdots & 0 \\ \delta_{v_1} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & \delta_{u_1} & \cdots & \delta_{u_n} \\ 0 & \delta_{v_1} & \cdots & \delta_{v_n} \\ 0 & 0 & \cdots & 0 \\ -\delta_{u_1}u_1' - \delta_{u_1'}u_1 & -\delta_{u_1}v_1' - \delta_{v_1'}u_1 & \cdots & -\delta_{u_n}v_n' - \delta_{v_n'}u_n \\ -\delta_{v_1}u_1' - \delta_{u_1'}v_1 & -\delta_{v_1}v_1' - \delta_{v_1'}v_1 & \cdots & -\delta_{v_n}v_n' - \delta_{v_n'}v_n \\ -\delta_{u_1'} & -\delta_{v_1'} & \cdots & -\delta_{v_n'} \end{bmatrix}. \tag{4.18}$$

Assuming that the errors in different points and different components of the image coordinates are uncorrelated and have the same variance σ^2, we get

$$\Gamma_{A'} = \sigma^2 \,\text{diag}\,(P_1, P_2, \cdots, P_n) \tag{4.19}$$

where P_i, $1 \le i \le n$, is an 18×18 submatrix:

$$P_i = \begin{bmatrix} J & O & -u_i'J & O & v_i'J \\ O & O & O & O & O \\ -u_i'J & O & M_{1i} & -u_i'J & u_i'v_i'J \\ O & O & O & O & O \\ J & O & -u_i'J & J & -v_i'J \\ -v_i'J & O & u_i'v_i'J & -v_i'J & M_{2i} \end{bmatrix} \tag{4.20}$$

where O is a 3×3 zero matrix and

$$J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$M_{1i} = \boldsymbol{X}_i \boldsymbol{X}_i^t + (u_i')^2 J, \qquad M_{2i} = \boldsymbol{X}_i \boldsymbol{X}_i^t + (v_i')^2 J \tag{4.21}$$

Now, consider the error of $\boldsymbol{h}$ in (Ag.3). From Theorem 9 and (Ag.2), we have (note that $\boldsymbol{h}$ is an eigenvector of A^tA instead of A)

$$\begin{aligned} \boldsymbol{\delta}_h &\cong H\, \Delta H^t\, \Delta_{A'A} \boldsymbol{h} \\ &= H\, \Delta H^t [h_1 I_9 \quad h_2 I_9 \quad \cdots \quad h_9 I_9]\, \boldsymbol{\delta}_{A'A} \triangleq G_h\, \boldsymbol{\delta}_{A'A}. \end{aligned} \tag{4.22}$$

In the above equations, we have rewritten the matrices $\Delta_{A'A}$ by $\boldsymbol{\delta}_{A'A}$ and moved the perturbation to the right end of the expression. In this way, the perturbation of the eigenvector is then the linear transformation (by matrix G_h) of the perturbation vector $\boldsymbol{\delta}_{A'A}$. Since we have $\Gamma_{A'} (= \Gamma_A^t)$ in (4.18), we need to relate $\boldsymbol{\delta}_{A'A}$ in (4.22) to $\boldsymbol{\delta}_{A'}$. Using first-order approximation, we get

$$\Delta_{A'A} \cong A^t \Delta_A + \Delta_A^t A. \tag{4.23}$$

Letting $A^t = [a_{ij}]^t \triangleq [\boldsymbol{A}_1 \quad \boldsymbol{A}_2 \quad \cdots \quad \boldsymbol{A}_n]$ we write

$$\boldsymbol{\delta}_{A'A} \cong G_{A'A} \boldsymbol{\delta}_{A'} \tag{4.24}$$

where $G_{A'A}$ can be easily determined from (4.23): $G_{A'A} = [F_{ij}] + [G_{ij}]$ where $[F_{ij}]$ and $[G_{ij}]$ are matrices with $9 \times n$ submatrices F_{ij} and G_{ij}, respectively, $F_{ij} = a_{ji} I_9$, and G_{ij} is a 9×9 matrix with the ith column being the column vector $\boldsymbol{A}_j$ (see 3.30) and all other columns being zeros. From (4.22) and (4.24) we get

$$\boldsymbol{\delta}_h \cong G_h\, \boldsymbol{\delta}_{A'A} \cong G_h G_{A'A}\, \boldsymbol{\delta}_{A'} \triangleq D_h\, \boldsymbol{\delta}_{A'}. \tag{4.25}$$

Then $\Gamma_h \cong D_h \Delta_{A'} D_h^t$. Since $\boldsymbol{F}^t = \pm \boldsymbol{h} / \sqrt{|\gamma_2|}$, there exists a permutation matrix M such that $\boldsymbol{F} = \pm M \boldsymbol{h} / \sqrt{|\gamma_2|}$. We have

$$\boldsymbol{\delta}_F \cong \pm \frac{1}{\sqrt{|\gamma_2|}} M\, \boldsymbol{\delta}_h \cong \pm \frac{1}{\sqrt{|\gamma_2|}} M D_h\, \boldsymbol{\delta}_{A'} \triangleq \pm D_F\, \boldsymbol{\delta}_{A'} \tag{4.26}$$

where the sign is negative iff (Ag.7) holds.

Starting from the covariance matrix of the perturbation in A^t, we got the covariance matrix of the perturbation in the eigenvector of A^tA. Finally, we get the perturbation vector of F, $\boldsymbol{\delta}_F$. For the perturbation vectors of R, $\mathring{\boldsymbol{T}}$, and $\boldsymbol{N}$ we will get the linear expression in terms of $\boldsymbol{\delta}_F$. The corresponding covariance matrices then can be obtained. For example, if we get $D_{\mathring{T}}$ such that $\boldsymbol{\delta}_{\mathring{T}} \cong D_{\mathring{T}} \boldsymbol{\delta}_F$, it follows that $\Gamma_{\mathring{T}} \cong D_{\mathring{T}} \Gamma_F D_{\mathring{T}}^t$.

The solution of step i needs the eigenvector of A^tA associated with the smallest eigenvalue. The smallest eigenvalue is a simple zero eigenvalue when rank $(A) = 8$ (nondegenerate configuration). When rank $(A) < 8$ (i.e., when degenerate configurations occur), the solution $\boldsymbol{h}$ in step i is very sensitive to noise. As can be seen from (4.12), the second diagonal element of Δ is infinite when $\lambda_1 = \lambda_2$, which makes the estimated errors infinite.

However, in most real applications, we do not know the noise-free A. We only know the noise-corrupted A: $A(\epsilon)$. Therefore, $A(\epsilon)$ is used to estimate A. In the presence of noise, generally, the rank of $A(\epsilon)$ is full mathematically and the smallest eigenvalue of $A(\epsilon)^t A(\epsilon)$ is a small positive number. If noise is reasonably small, when rank $(A) < 8$ we have $\lambda_1 \approx \lambda_2$. Then large estimates of errors are still generated. From a slightly different point of view, we can regard A as a "noise-corrupted" matrix by adding $-\Delta_A$ to the matrix $A(\epsilon)$. Now the error is the deviation of the true solution from the noise-corrupted solution. This observation justifies our use of the noise-corrupted A to estimate errors.

For Step ii: To obtain the perturbation of R, $\mathring{\boldsymbol{T}}$ and $\tilde{\boldsymbol{N}}$, we need the perturbation of H in (Ag.5) and that of α and β in (Ag.9), which, in turn, need the perturbation of the eigenvalues and the eigenvectors of F^tF. Theorem 9 can be used again to give a first-order perturbation of the eigenvalues and eigenvectors (column vectors in H) of F^tF. According to Theorem 9, for $i = 1, 2, 3$, we have

$$\boldsymbol{\delta}_{\gamma_i} \cong \boldsymbol{h}_i^t\, \Delta_{F_s^t F_s} \boldsymbol{h}_i \cong \boldsymbol{d}_i^t\, \boldsymbol{\delta}_{F_s^t F_s} \tag{4.27}$$

where $\boldsymbol{d}_i = (h_{1i}h_{1i},\ h_{2i}h_{1i},\ h_{3i}h_{1i},\ h_{1i}h_{2i},\ h_{2i}h_{2i},\ h_{3i}h_{2i},\ h_{1i}h_{3i},\ h_{2i}h_{3i},\ h_{3i}h_{3i})$ and $H = [h_{ij}]$. Therefore,

$$\boldsymbol{\delta}_\gamma \triangleq \begin{bmatrix} \delta_{\gamma_1} \\ \gamma_{\gamma_2} \\ \gamma_{\gamma_3} \end{bmatrix} \cong \begin{bmatrix} \boldsymbol{d}_1^t \\ \boldsymbol{d}_2^t \\ \boldsymbol{d}_3^t \end{bmatrix} \boldsymbol{\delta}_{F_s^t F_s} \triangleq G_\gamma\, \boldsymbol{\delta}_{F_s^t F_s} = G_\gamma \gamma_2\, \boldsymbol{\delta}_{F'F} \tag{4.28}$$

since $F_s^t F_s = \gamma_2 F^t F$. So we need $\boldsymbol{\delta}_{F'F}$, which can be derived in terms of $\boldsymbol{\delta}_F$:

$$\Delta_{F'F} \cong F^t\, \Delta_F + \Delta_F^t F. \tag{4.29}$$

Letting $F^t = [e_{ij}]^t \triangleq [\boldsymbol{F}_1 \quad \boldsymbol{F}_2 \quad \boldsymbol{F}_3]$ we write

$$\boldsymbol{\delta}_{F'F} \cong D_{F'F}\, \boldsymbol{\delta}_F \tag{4.30}$$

where $D_{F'F}$ can be easily determined from (4.30): $D_{F'F} = \text{diag}\,(F^t, F^t, F^t) + [G_{ij}]$ where $\text{diag}\,(F^t, F^t, F^t)$ is a diagonal matrix whose elements are 3×3 submatrices, $[G_{ij}]$ is a matrix with 3×3 submatrices G_{ij}'s, G_{ij} is a 3×3 matrix with the jth row vector being the column vector $\boldsymbol{F}_i$ (see 3.39) and all other rows being zeros. From (4.29) and (4.31) we get

$$\boldsymbol{\delta}_\gamma \cong G_\gamma \gamma_2\, \boldsymbol{\delta}_{F'F} \cong G_\gamma \gamma_2 D_{F'F}\, \boldsymbol{\delta}_F \triangleq D_\gamma\, \boldsymbol{\delta}_F. \tag{4.31}$$

Now we discuss the perturbation of the eigenvectors of F^tF, i.e., the column vectors in H. Using Theorem 9, we have

$$\begin{aligned} \boldsymbol{\delta}_H \triangleq \begin{bmatrix} \boldsymbol{\delta}_{h_1} \\ \boldsymbol{\delta}_{h_2} \\ \boldsymbol{\delta}_{h_3} \end{bmatrix} &\cong \begin{bmatrix} H\, \Delta_1\, H^t [h_{11} I_3 & h_{21} I_3 & h_{31} I] \\ H\, \Delta_2\, H^t [h_{12} I_3 & h_{22} I_3 & h_{32} I] \\ H\, \Delta_3\, H^t [h_{13} I_3 & h_{23} I_3 & h_{33} I] \end{bmatrix} \boldsymbol{\delta}_{F'F} \\ &\triangleq G_H\, \boldsymbol{\delta}_{F'F} \cong G_H D_{F'F}\, \boldsymbol{\delta}_F \triangleq D_H\, \boldsymbol{\delta}_F \end{aligned} \tag{4.32}$$

where

$$\Delta_1 = \text{diag}\,(0, (\lambda_1 - \lambda_2)^{-1}, (\lambda_1 - \lambda_3)^{-1})$$
$$\Delta_2 = \text{diag}\,((\lambda_2 - \lambda_1)^{-1}, 0, (\lambda_2 - \lambda_3)^{-1})$$
$$\Delta_3 = \text{diag}\,((\lambda_3 - \lambda_1)^{-1}, (\lambda_3 - \lambda_2)^{-1}, 0). \quad (4.33)$$

From (Ag.11) we know that the perturbations in R depend on those in V_1, V_2, and F. From (Ag.10), the perturbations of V_1 and V_2 depend on those in α, β, and H. Now we consider the first-order perturbation in α and β. From (Ag.8) and (Ag.9) we have

$$\alpha = \sqrt{\frac{\lambda_3 - 1}{\lambda_3 - \lambda_1}} = \sqrt{\frac{\gamma_3 - \gamma_2}{\gamma_3 - \gamma_1}}$$
$$\beta = \sqrt{\frac{1 - \lambda_1}{\lambda_3 - \lambda_1}} = \sqrt{\frac{\gamma_2 - \gamma_1}{\gamma_3 - \gamma_1}}. \quad (4.34)$$

The noise corrupted α can be written as $\alpha(\delta_{\gamma_1}, \delta_{\gamma_2}, \delta_{\gamma_3})$, with the noise-free α being $\alpha(0, 0, 0)$. Expending $\alpha(\delta_{\gamma_1}, \delta_{\gamma_2}, \delta_{\gamma_3})$ at point $(0, 0, 0)$ using Taylor series for multiple variables, it follows that

$$\alpha(\delta_{\gamma_1}, \delta_{\gamma_2}, \delta_{\gamma_3}) \cong \alpha + \frac{\sqrt{\gamma_3 - \gamma_2}}{2\sqrt{(\gamma_3 - \gamma_1)^3}}\delta_{\gamma_1} - \frac{1}{2\sqrt{(\gamma_3 - \gamma_1)(\gamma_3 - \gamma_2)}}\delta_{\gamma_2} + \frac{\gamma_2 - \gamma_1}{2\sqrt{(\gamma_3 - \gamma_1)^3(\gamma_3 - \gamma_2)}}\delta_{\gamma_3}. \quad (4.35)$$

Similarly, we get the first-order approximation of $\beta(\delta_{\gamma_1}, \delta_{\gamma_2}, \delta_{\gamma_3})$. Writing the first-order perturbation of α and β in matrix form, we have

$$\begin{bmatrix}\delta_\alpha \\ \delta_\beta\end{bmatrix} \cong \frac{1}{\sqrt{(\gamma_3 - \gamma_1)^3(\gamma_2 - \gamma_1)(\gamma_3 - \gamma_2)}} \cdot \begin{bmatrix} (\gamma_3 - \gamma_2)\sqrt{\gamma_2 - \gamma_1} & -(\gamma_3 - \gamma_1)\sqrt{\gamma_2 - \gamma_1} & \sqrt{(\gamma_2 - \gamma_1)^3} \\ -\sqrt{(\gamma_3 - \gamma_2)^3} & -(\gamma_3 - \gamma_1)\sqrt{\gamma_3 - \gamma_2} & (\gamma_2 - \gamma_1)\sqrt{\gamma_3 - \gamma_2} \end{bmatrix} \begin{bmatrix}\delta_{\gamma_1} \\ \delta_{\gamma_2} \\ \delta_{\gamma_3}\end{bmatrix}$$
$$\triangleq G_{\alpha\beta}\,\boldsymbol{\delta}_\gamma \cong G_{\alpha\beta}D_\gamma\,\boldsymbol{\delta}_F \triangleq D_{\alpha\beta}\,\boldsymbol{\delta}_F. \quad (4.36)$$

Let $\boldsymbol{G}_\alpha^t$ and $\boldsymbol{G}_\beta^t$ be the row vectors of $D_{\alpha\beta}$, i.e., $G_{\alpha\beta} = [\boldsymbol{G}_\alpha \quad \boldsymbol{G}_\beta]^t$. From (Ag.10) we have the first-order perturbation of $\boldsymbol{V}_1$ and $\boldsymbol{V}_2$:

$$\boldsymbol{\delta}_{V_1} \cong \delta_\alpha \boldsymbol{h}_1 + \alpha\boldsymbol{\delta}_{h_1} + \delta_\beta \boldsymbol{h}_3 + \beta\delta_{h_3} = (\boldsymbol{h}_1\boldsymbol{G}_\alpha^t + \boldsymbol{h}_3\boldsymbol{G}_\beta^t + [\alpha I_3 \quad O \quad \beta I_3]D_H)\delta_F \triangleq D_{V_1}\,\boldsymbol{\delta}_F \quad (4.37)$$

$$\boldsymbol{\delta}_{V_2} = \boldsymbol{\delta}_{h_2} \cong [O \quad I_3 \quad O]D_H\,\boldsymbol{\delta}_F \triangleq D_{V_2}\,\boldsymbol{\delta}_F. \quad (4.38)$$

Now we find D_R such that $\boldsymbol{\delta}_R = D_R\,\boldsymbol{\delta}_F$. First, we consider the perturbation of $\boldsymbol{W}_i \triangleq F\boldsymbol{V}_i$, $i = 1, 2$. Letting $\boldsymbol{V}_i = (v_1, v_{2i}, v_{3i})$, we have

$$\boldsymbol{\delta}_{W_i} \cong \Delta_F \boldsymbol{V}_i + F\,\boldsymbol{\delta}_{V_i} = ([v_{1i}I_3 \quad v_{2i}I_3 \quad v_{3i}I_3] + FD_{V_i})\,\boldsymbol{\delta}_F \triangleq D_{W_i}\,\boldsymbol{\delta}_F. \quad (4.39)$$

For the conciseness of notations, we define a new vector $\boldsymbol{K}$:

$$\boldsymbol{\delta}_K \triangleq \begin{bmatrix}\boldsymbol{\delta}_{V_1} \\ \boldsymbol{\delta}_{V_2} \\ \boldsymbol{\delta}_{W_1} \\ \boldsymbol{\delta}_{W_2}\end{bmatrix} \cong \begin{bmatrix}D_{V_1} \\ D_{V_2} \\ D_{W_1} \\ D_{W_2}\end{bmatrix}\boldsymbol{\delta}_F \triangleq D_K\,\boldsymbol{\delta}_F. \quad (4.40)$$

Since $R = [\boldsymbol{W}_1, \boldsymbol{W}_2, \boldsymbol{W}_1 \times \boldsymbol{W}_2][\boldsymbol{V}_1, \boldsymbol{V}_2, \boldsymbol{V}_1 \times \boldsymbol{V}_2]^t$, evaluating the first-order perturbation as in (4.2), one can construct a matrix G_R such that

$$\boldsymbol{\delta}_R \cong G_R\,\boldsymbol{\delta}_K. \quad (4.41)$$

Then, (4.41) and (4.42) give

$$\boldsymbol{\delta}_R \cong G_R\,\boldsymbol{\delta}_K \cong G_R D_K\,\boldsymbol{\delta}_R \triangleq D_R\,\boldsymbol{\delta}_F. \quad (4.42)$$

Now we consider the perturbation of $\hat{\boldsymbol{N}}$. From (Ag.12), (4.38), (4.39) and the possible sign change according to (Ag.14) we have

$$\pm\boldsymbol{\delta}_{\hat{N}} \cong \boldsymbol{V}_1 \times \boldsymbol{\delta}_{V_2} + \boldsymbol{\delta}_{V_1} \times \boldsymbol{V}_2 = [\boldsymbol{V}_1]_\times\boldsymbol{\delta}_{V_2} - [\boldsymbol{V}_2]_\times\boldsymbol{\delta}_{V_1} = ([\boldsymbol{V}_1]_\times D_{V_2} - [\boldsymbol{V}_1]_\times D_{V_2})\,\boldsymbol{\delta}_F \triangleq \pm D_{\hat{N}}\,\boldsymbol{\delta}_F. \quad (4.43)$$

From (Ag.16), we get the perturbation of $\tilde{\boldsymbol{N}}$

$$\boldsymbol{\delta}_{\tilde{N}} \cong \|\tilde{\boldsymbol{T}}\|\,\boldsymbol{\delta}_{\hat{N}} \cong \|\tilde{\boldsymbol{T}}\|D_{\hat{N}}\,\boldsymbol{\delta}_F. \quad (4.44)$$

From (Ag.13), we obtain the first-order perturbation of $\tilde{\boldsymbol{T}}$:

$$\boldsymbol{\delta}_{\tilde{T}} \cong (F - R)\,\boldsymbol{\delta}_{\hat{N}} + (\Delta_F - \Delta_R)\mathring{\boldsymbol{N}}. \quad (4.45)$$

Letting $\mathring{\boldsymbol{N}} = (n_1, n_2, n_3)$, and $M_{\mathring{N}} \triangleq [n_1 I_3 \quad n_2 I_3 \quad n_3 I_3]$ we have $\Delta_F\mathring{\boldsymbol{N}} = M_{\mathring{N}}\,\boldsymbol{\delta}_F$ and $\Delta_R\mathring{\boldsymbol{N}} = M_{\mathring{N}}\,\boldsymbol{\delta}_R$. Using (4.43),

(4.44), and (4.46), we have

$$\delta_{\tilde{T}} = (F - R)D_{\tilde{N}}\,\delta_F + M_{\tilde{N}}^{\circ}\,\delta_F - M_{\tilde{N}}\,\delta_R$$
$$\cong ((F - R)D_{\tilde{N}} + M_{\tilde{N}} - M_{\tilde{N}}D_R)\,\delta_F \triangleq D_{\tilde{T}}\,\delta_F. \quad (4.46)$$

Then $\delta_{\hat{T}} \cong \|\tilde{T}\|^{-1}\,\delta_{\tilde{T}}$.

As in step i, we estimate in step ii the errors by using the perturbed F, R, and $\|\tilde{T}\|$, etc., to substitute for the noise-free ones.

In summary, the perturbation vectors of R, $\mathring{T}$, and $\tilde{N}$ are expressed in terms of linear transformation of perturbation in F. Using the covariance matrix of F in (4.27), the covariance matrix of R is then given by

$$\Gamma_R = D_R\Gamma_F D_R^t. \quad (4.47)$$

We can estimate the Euclidean norm of the perturbation matrix by the square root of the trace of the corresponding covariance matrix:

$$\|\Delta_R\| = \|\delta_R\| = \sqrt{\mathrm{tr}\,(\Gamma_R)}. \quad (4.48)$$

Similarly,

$$\|\delta_{\hat{T}}\| \approx \sqrt{\mathrm{tr}\,(\Gamma_{\hat{T}})} \quad (4.49)$$

where $\Gamma_{\hat{T}}$ is the covariance matrix of $\mathring{T}$. The Euclidean norm of the perturbation vector of $\tilde{N}$ can be estimated in a similar way.

V. Experimental Results

The experiments were designed to demonstrate the performance of the motion estimation algorithm and the associated error estimation. In this section, we first present the results of simulation. Then we show the experiments with real images.

A. Simulation

In the simulation, the object plane intersects the optical axis at $z = 11$. The image is a unit square with a unit focal length. The object feature points are generated randomly with a uniform distribution in the object plane. Only those points that are visible both before and after motion are used as feature points. The image coordinates of the points are digitized according to the resolution of the camera. In other words, the image coordinates are rounded off to the center of the nearest pixel before they are used by the motion estimation algorithm. Other additonal random noise, such as corner detection errors, can be simulated by a reduced image resolution.

The errors of the estimated R, $\hat{T}$, and $\tilde{N}$ are shown as relative errors, which are defined by the Euclidean norm of the error vector (or matrix) divided by the Euclidean norm of the original vector (or matrix).

We assume that the roundoff errors are uniformly distributed between plus half and minus half of the pixel size. So the variance of the errors in the image coordinates is $\sigma^2 = s^2/12$, where s is the spacing between the quantization levels in the images. This variance is used in the error estimation. Different motion parameters with different image resolutions are simulated. Both the results of simulation and the formulation of error estimation indicate that the error in solution is roughly proportional to the variance of the noise. Therefore, we need only to show here the results with a typical noise level. We choose image resolution to be 256 × 256 pixels here. Fig. 6 shows the results of a sequence of trials with 5 point correspondences. The normal of the object plane is (0.5, 0.4, 0.7) before motion. The rotation is about an axis (1, 1, 1) by an angle of 5°. The translation vector is (0.2, 0.3, −2). In Fig. 6, the results of 40 random trials (randomly generated points on the plane) are shown in the order of trials. Fig. 6(a) shows the relative errors of F. Figs. 6(b), (c), and (d) show the relative errors in the rotation matrix R, the translation T, and the normal N, respectively. As can be seen from these figures, the estimated errors are strongly correlated with the actual errors. The estimated errors are especially important to detect some relatively unreliable configurations.

The average performance of the error estimation as well as that of the motion estimation algorithm is presented in Fig. 7 with different numbers of point correspondences. The data shown are based on 40 random trials. The deviation of error estimation is defined as the average of the absolute difference between the estimated error and the actual error. The bias of error estimation is defined as the difference between the mean of the estimated errors and the mean of the actual errors. As can be seen from Fig. 7, the actual mean relative errors decrease very fast when the number of points increases beyond the required minimum of 4. This indicates that it is very effective to reduce the error by using a few more points in addition to the minimally required 4. In Fig. 7, the mean deviation between the estimated error and the actual error is about half of the actual error, except for the cases where the number of points is small (e.g., four points). When the number of point correspondences is equal to 4, there is a reasonably high probability for the randomly generated points to form a nearly degenerate configuration. When the point configuration is degenerate or nearly degenerate, the difference between the estimated error and the actual error is expected to be large. This is one of the reasons for the large deviations and bias in the 4-point case. Some individual simulations still show a good agreement between the estimated errors and the actual errors in the 4-point case.

B. Real Images

A CCD video camera with 480 × 500 pixels was used as the image sensor. The camera is calibrated but no correction has been made for lens distortion. Fig. 8(a)–(c) shows three views of a bulletin board, called image 1, image 2, and image 3, respectively. As we discussed, there are generally two solutions from two views, and adding a third view generally leads to a unique solution. Our image matching algorithm presented in [18] was used to establish the image plane displacement fields between

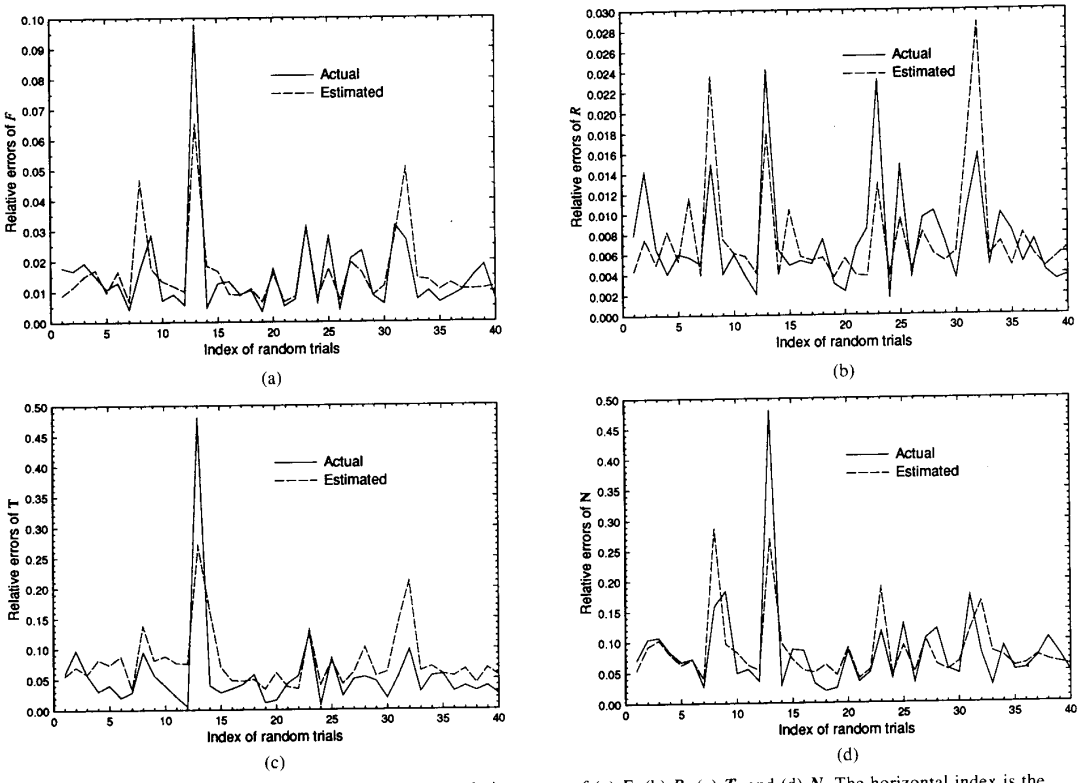


Fig. 6. The actual relative errors and the estimated relative errors of (a) F, (b) R, (c) T, and (d) N. The horizontal index is the order of random trials.

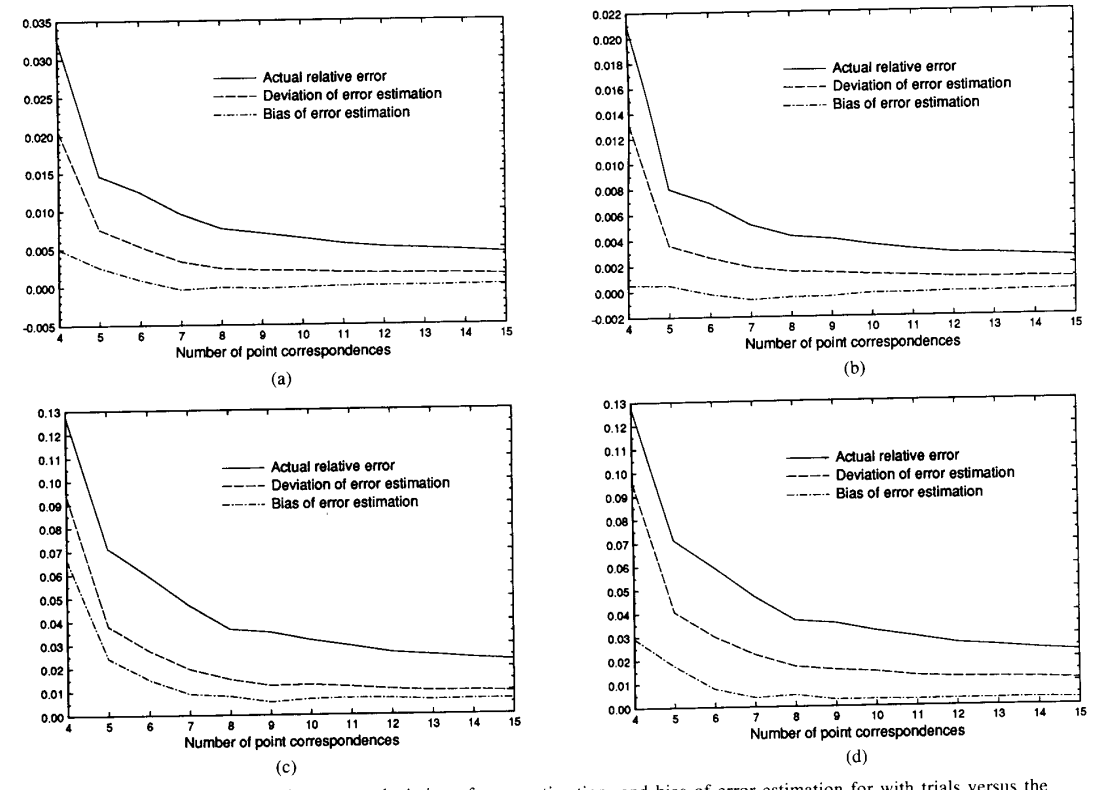
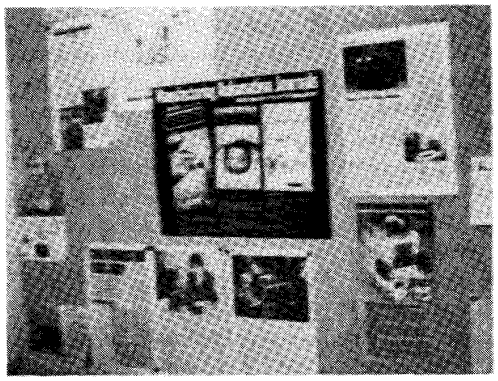
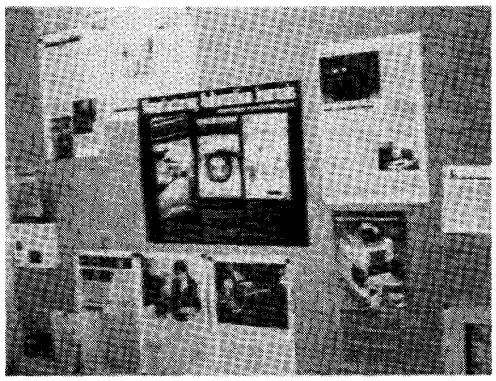


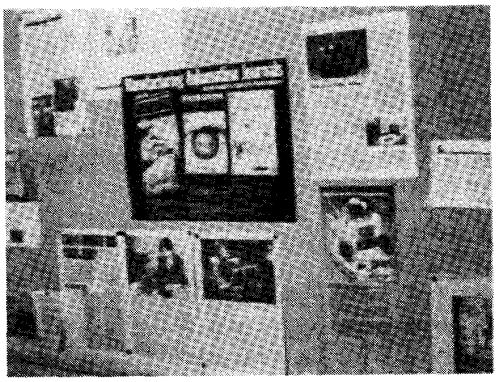
Fig. 7. The actual mean relative errors, deviation of error estimation, and bias of error estimation for with trials versus the number of point correspondences for (a) F, (b) R, (c) T, and (d) N.



(a)

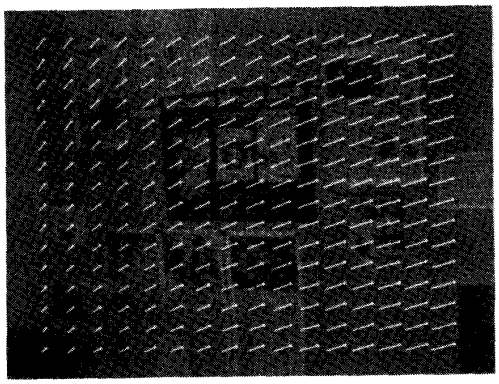


(b)

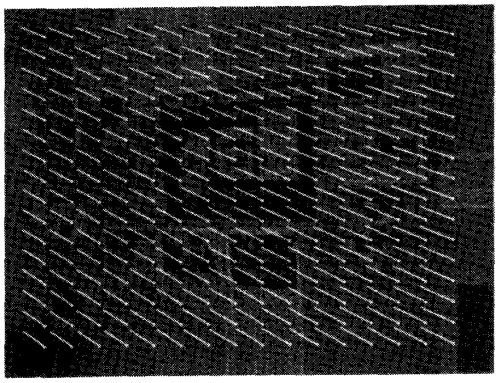


(c)

Fig. 8. Three views of a bulletin board. (a) Image 1, (b) image 2, (c) image 3.



(a)



(b)

Fig. 9. Samples of the computed displacement fields for the bulletin scene, superimposed on the blurred extended intensity image. (a) Image 1 to 2 (1-2); (b) image 1 to 3 (1-3).

TABLE I
RESULTS FOR BULLETIN BOARD IMAGES (1-2)

Solution 1			
Unit normal	−0.08	−0.966	0.24
Unit translation	0.04	0.58	0.81
Rotation axis	−0.95	0.29	0.08
Rotation angle		−1.8°	
E. R. error of normal		0.16	
E. R. error of translation		0.13	
E. R. error of rotation matrix		0.009	
Solution 2			
Unit normal	0.04	0.58	0.81
Unit translation	−0.08	−0.966	0.24
Rotation axis	0.986	0.13	0.09
Rotation angle		−1.8°	
E. R. error of normal		0.12	
E. R. error of translation		0.18	
E. R. error of rotation matrix		0.016	

two images (equivalent to point correspondences). Fig. 9 shows sampled displacement fields of the two motions: motion 1-2 (from image 1 to 2) and motion 1-3 (from image 1 to 3). Table I shows the estimated parameters and the corresponding estimated relative errors for motion 1-2, and Table II shows those for motion 1-3. For error estimation, the variance of the noise in image points is assumed to be equal to that of a uniformly distributed random noise ranging ± 1 pixel. Let $N_{i-j,k}$ denote the kth solution of the unit object normal for motion from image i to j. The correct solution for each motion is determined

TABLE II
RESULTS FOR BULLETIN BOARD IMAGES (1-3)

Solution 1			
Unit normal	0.12	0.35	0.930
Unit translation	0.28	−0.77	0.57
Rotation axis	0.61	0.79	0.02
Rotation angle		0.5°	
E. R. error of normal		0.13	
E. R. error of translation		0.22	
E. R. error of rotation matrix		0.013	
Solution 2			
Unit normal	0.27	−0.76	0.59
Unit translation	0.14	0.32	0.936
Rotation axis	−0.936	−0.30	0.17
Rotation angle		−2.7°	
E. R. error of normal		0.20	
E. R. error of translation		0.17	
E. R. error of rotation matrix		0.006	

by the consistency of an object normal at image 1:

$$\|N_{1\text{-}2,1} - N_{1\text{-}3,1}\| = 1.49$$
$$\|N_{1\text{-}2,1} - N_{1\text{-}3,2}\| = 0.53$$
$$\|N_{1\text{-}2,2} - N_{1\text{-}3,1}\| = 0.27$$
$$\|N_{1\text{-}2,2} - N_{1\text{-}3,2}\| = 1.37.$$

Since the solution 2 of motion 1-2 and solution 1 of motion 1-3 give the most consistent object normals, the corresponding solutions are what we needed. These solutions roughly agree with visual inspection (since the rotation angles are very small, the rotation axes are not important), although lens distortion is visible from images.

VI. CONCLUSIONS

From two perspective views of a planar scene which is undergoing a rigid motion, there are generally two (normalized) interpretations for motion parameters and the positions of the object plane. These two interpretations, one is veridical and the other illusive, are both valid in the sense that they render the same pair of images. We have identified all the special cases in which the number of interpretations is not two, and we have also derived necessary and sufficient geometrical conditions for those special cases to occur. The illusive solution can be rejected if some constructed points happen to fall in the backward half space.

Assuming that n coplanar points undergo a rigid motion with a nonvanishing translational component and that the plane where the points lie does not go through the projection center of the camera before and after motion, then the sufficient condition for our algorithm to be able to reach the above solution(s) is the existence of four extracted points among which no three are collinear in the object plane. However, we have further established a fundamental result which solves the inherent uniqueness of the problem: This sufficient condition is in fact also a necessary condition for existing the above number of solution(s). Therefore, the condition is both necessary and sufficient for the problem itself to be solvable. In terms of image plane points, the corresponding necessary and sufficient condition is that there exist at least four points in each image among which no three points are collinear in either image plane. More points are useful in practice to combat noise.

We have presented a new, simpler linear algorithm that gives the solution(s) in a closed form. Due to the decomposability of any 3×3 matrix, we have derived a somewhat surprising optimality for this linear algorithm. This optimality indicates that our linear algorithm does not neglect any constraint in solving the linear equations.

If it is not known that a rigidly moving scene is planar or not, then one may illusively interpret a nonplanar surface as planar so long as the surface belongs to the class of plane perceivables investigated in this paper.

Using three distinct perspective views, the interpretation is generally unique.

Our approach to error estimation is based on the first-order perturbation. The estimated errors provide quantitative assessment for the accuracy of the solutions. They also indicate degenerate or nearly degenerate configurations in the presence of noise.

APPENDIX A

We first prove a lemma which is needed for the proofs of Theorem 1 and Corollary 1.

Lemma 2: Assume the object plane does not go through the projection center (origin). The set of points in the plane lie on a straight line in the image plane if and only if they lie on a straight line in the object plane.

Proof: The equation of the object plane that does not go through the origin can be written as $ax + by + cz = 1$. If the points lie on a straight line in image plane, they lie on the intersection of $ax + by + cz = 1$ and another plane $ex + fy + gz = 1$ that does not go through the origin either (such a plane exists since the origin is not on the line). For the intersection to be a line, (a, b, c) is not parallel to (e, f, g). These two equations give $(a - e)x + (b - f)y + (c - g)z = 0$ or $(a - e)(x/z) + (b - f)(y/z) + c - g = 0$. That is, their images lie on a straight line $(a - e)u + (b - f)v + (c - g) = 0$ in the image plane provided $(a - e)^2 + (b - f)^2 \neq 0$. In fact, if $(a - e)^2 + (b - f)^2 = 0$, we have $c = g$, and thus $(a, b, c) \| (e, f, g)$, which is a contradiction.

Conversely, assume the image of the points lie on a straight line $eu + fv + g = 0$ in the image plane with $e^2 + f^2 \neq 0$. Their 3-D coordinates must satisfy $e(x/z) + f(y/z) + g = 0$, or $ex + fy + gz = 0$. The intersection of the object plane $ax + by + cz = 1$ and $ex + fy + gz = 0$ give a straight line in 3-D provided (a, b, c) is not

parallel to (e, f, g). In fact, (a, b, c) is never parallel to (e, f, g). If they are parallel, $ax + by + cz = 1$ and $ex + fy + gz = 0$ have no intersection, which is impossible since the object points are in the intersection. □

Theorem 1: Rank $(A) = 8$ if and only if there exists a set of four object points such that no image projections of any three points in this set are collinear in any of the two images.

Proof: For the sufficiency part, assume such a set of four points exists. We prove the corresponding 8 rows of A has a rank 8. We prove by contradiction. Suppose that these 8 rows are linearly dependent. We have two vectors $\boldsymbol{a} = (a_1, a_2, a_3, a_4)$ and $\boldsymbol{b} = (b_1, b_2, b_3, b_4)$ with $\boldsymbol{a} \neq \boldsymbol{0}$ or $\boldsymbol{b} \neq \boldsymbol{0}$, such that (see (Ag. 1))

$$\sum_{i=1}^{4} a_i \boldsymbol{X}_i = \boldsymbol{0} \tag{A.1}$$

$$\sum_{i=1}^{4} \boldsymbol{b}_i X_i = \boldsymbol{0} \tag{A.2}$$

$$\sum_{i=1}^{4} a_i v_i' \boldsymbol{X}_i + \sum_{i=1}^{4} b_i u_i' \boldsymbol{X}_i = \boldsymbol{0}. \tag{A.3}$$

Let $X = [\boldsymbol{X}_1 \quad \boldsymbol{X}_2 \quad \boldsymbol{X}_3 \quad \boldsymbol{X}_4]$ and $\boldsymbol{c} = (c_1, c_2, c_3, c_4)$, where $c_i = a_i u_i' + b_i v_i'$. Equations (A.1)-(A.3) give $X\boldsymbol{a} = \boldsymbol{0}$, $X\boldsymbol{b} = \boldsymbol{0}$ and $X\boldsymbol{c} = \boldsymbol{0}$. Without loss of generality assume $\boldsymbol{a} \neq \boldsymbol{0}$. Since no three points lie on a straight line, no a_i can be zero. We have rank $(X) \geq 3$. So, $\boldsymbol{a} \| \boldsymbol{b} \| \boldsymbol{c}$. There exist constants k and c such that $b_i = ka_i$ and $ca_i = c_i = (a_i u_i' + b_i v_i')(u_i' + kv_i')a_i$, $1 \leq i \leq 4$. Since $a_i \neq 0$, $1 \leq i \leq 4$, we have $u_i' + kv_i' = c$, $1 \leq i \leq 4$. This means that all four points lie in a straight line after motion, which is a contradiction. (From Lemma 2 we see that if such a set of four points exists before motion, the only way for the condition to be violated after motion is that the object plane goes through the projection center after motion.)

For the necessity part, first we prove that if all the image points are collinear in the image plane, then rank $(A) \leq 5$. Assume that all the points lie on a straight line (called common line). Let $\boldsymbol{v}$ be the unit normal of the plane that contains the common line and passes through the origin. We have $\boldsymbol{X}_i' \boldsymbol{v} = 0$, $i = 1, 2, \cdots, n$. Define four 9-dimensional vectors:

$$\boldsymbol{h}_1 = \begin{bmatrix} \boldsymbol{v} \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{h}_2 = \begin{bmatrix} 0 \\ \boldsymbol{v} \\ 0 \end{bmatrix}$$

$$\boldsymbol{h}_3 = \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{v} \end{bmatrix}, \quad \boldsymbol{h}_4 = \begin{bmatrix} \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \\ \boldsymbol{e}_3 \end{bmatrix} \tag{A.4}$$

where $[\boldsymbol{e}_1 \quad \boldsymbol{e}_2 \quad \boldsymbol{e}_3] = F^t$. From (2.20) and the structure of A in (Ag.1), we have $A\boldsymbol{h}_i = 0$, $i = 1, 2, 3, 4$. Furthermore, we prove that the four vectors in (A.4) are linearly independent. Otherwise, $\boldsymbol{h}_4$ is a linear combination of $\boldsymbol{h}_1$, $\boldsymbol{h}_2$, and $\boldsymbol{h}_3$, since it is clear that $\boldsymbol{h}_1$, $\boldsymbol{h}_2$, and $\boldsymbol{h}_3$ are linearly dependent. This means that the rank of F is at most one which is contradictory to (2.30). Since the equation $A\boldsymbol{h} = 0$ has at least four linearly independent solutions in (A.4), rank $(A) \leq 9 - 4 = 5$.

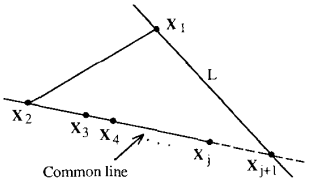


Fig. 10. Illustration for the proof of Theorem 1.

Now we assume that the condition of the theorem is not satisfied and prove rank $(A) < 8$. If $n < 4$ the theorem is trivially true. Now let $n \geq 4$. If any set of three points are collinear, it is clear that all points lie on a single straight line. Then rank of A is no more than 5 as we proved earlier. Now assume there exists a set of three points that are not collinear. Without loss of generality, assume they are $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, and $\boldsymbol{X}_3$. $\boldsymbol{X}_4$ must lie on a line that contains two of $\boldsymbol{X}_1$, $\boldsymbol{X}_2$, and $\boldsymbol{X}_3$ (otherwise the set of these four points violate our assumption). Without loss of generality, let $\boldsymbol{X}_2$, $\boldsymbol{X}_3$, and $\boldsymbol{X}_4$ lie on a straight line (call this common line, see Fig. 10).

We prove all the $\boldsymbol{X}_i$'s, $i > 4$, must lie on the common line. This is proved by the induction on i. Assume it is true for $i = j \geq 4$. Supose $\boldsymbol{X}_{j+1}$ is not on the common line. We will get a contradiction. The straight line L that passes through $\boldsymbol{X}_1$ and $\boldsymbol{X}_{j+1}$ can contain at most one point $\boldsymbol{X}_k$ on the common line, for some k, $2 \leq k \leq j$ (otherwise $\boldsymbol{X}_1$ would have been on the common line). Any two points other than $\boldsymbol{X}_k$ on the common line together with $\boldsymbol{X}_1$ and $\boldsymbol{X}_{j+1}$ form a set of four points that satisfy the condition, which is a contradiction. Therefore, $\boldsymbol{X}_{j+1}$ is on the common line.

Since all the points lie on a single line except $\boldsymbol{X}_1$, the rank of A is no more than $5 + 2 = 7$. □

Corollary 1: Rank $(A) = 8$ if and only if 1) there exists a set of four pionts in the object plane such that no three points in this set are collinear in the object plane; and 2) if the object plane is extended, it does not go through the projection center of the camera before and after motion.

Proof: If 1) and 2) are satisfied, by Lemma 2 the condition of Theorem 1 is satisfied. Thus rank $(A) = 8$. Conversely assume rank $(A) = 8$. Part 2 of Corollary 1 must be satisfied, otherwise all the points lie in a straight line and so rank $(A) \leq 5$ as proved in the proof of Theorem 1. Part 1 of Corollary 1 must be satisfied according to Lemma 2. □

Appendix B

Theorem 3: Given n point correspondences, $(\boldsymbol{X}_i, \boldsymbol{X}_i')$, $i = 1, 2, \cdots, n$. Let F be a matrix that satisfies (2.18)

for every $(\boldsymbol{X}_i, \boldsymbol{X}_i')$. If there are rotation matrices R_a, R_b, unit vectors $\mathring{\boldsymbol{T}}_a$, $\mathring{\boldsymbol{T}}_b$ and vectors $\tilde{\boldsymbol{N}}_a$ and $\tilde{\boldsymbol{N}}_b$, such that

$$F = R_a + \mathring{\boldsymbol{T}}_a \tilde{\boldsymbol{N}}_a^t = R_b + \mathring{\boldsymbol{T}}_b \tilde{\boldsymbol{N}}_b^t \tag{B.1}$$

and $\tilde{\boldsymbol{N}}_a \cdot \boldsymbol{X}_i > 0$, $\tilde{\boldsymbol{N}}_b \cdot \boldsymbol{X}_i > 0$, $i = 1, 2, \cdots, n$, hold true. Then there exist planes with $\tilde{\boldsymbol{N}}_a$ and $\tilde{\boldsymbol{N}}_b$ as normals, respectively, such that if they undergo motions represented by $(R_a, \mathring{\boldsymbol{T}}_a)$ and $(R_b, \mathring{\boldsymbol{T}}_b)$, respectively, they render the same pair of images (with image vectors $\boldsymbol{X}_i$ and $\boldsymbol{X}_i'$ at the two time instants) and the corresponding 3-D points are all located in the forward half-space before and after the motion.

Proof: For any point pair $(\boldsymbol{X}, \boldsymbol{X}')$ in $\{(\boldsymbol{X}_i, \boldsymbol{X}_i') \mid i = 1, 2, \cdots, n\}$, give its two positive depths z_a and z_b before motion by

$$z_a = \frac{1}{\tilde{\boldsymbol{N}}_a^t \boldsymbol{X}}, \qquad z_b = \frac{1}{\tilde{\boldsymbol{N}}_b^t \boldsymbol{X}}. \tag{B.2}$$

Their corresponding depths after motion, z_a' and z_b', are determined by (2.17):

$$z_a' \boldsymbol{X}' = z_a F \boldsymbol{X}, \qquad z_b' \boldsymbol{X}' = z_b F \boldsymbol{X}. \tag{B.3}$$

Equation (B.3) can be exactly satisfied because we have (2.18). Moreover, the solutions for depths in (B.3) are all positive since there exist true positive depths z and z' that satisfy (2.17) and positive depths before motion in (B.2) are the given conditions. Specifically, positive depths after motion are guaranteed for the illusive solution as well. We get the 3-D object points as follows:

$$\boldsymbol{x}_a = z_a \boldsymbol{X}, \qquad \boldsymbol{x}_a' = z_a' \boldsymbol{X}' \tag{B.4a}$$

$$\boldsymbol{x}_b = z_b \boldsymbol{X}, \qquad \boldsymbol{x}_b' = z_b' \boldsymbol{X}'. \tag{B.4b}$$

Using (B.1)–(B.4), we have

$$\begin{aligned} \boldsymbol{x}_a' = z_a' \boldsymbol{X}' &= F z_a \boldsymbol{X} = (R_a + \mathring{\boldsymbol{T}}_a \tilde{\boldsymbol{N}}_a^t) z_a \boldsymbol{X} \\ &= R_a z_a \boldsymbol{X} + \mathring{\boldsymbol{T}}_a (\tilde{\boldsymbol{N}}_a^t \boldsymbol{X}) z_a = R_a \boldsymbol{x}_a + \mathring{\boldsymbol{T}}_a. \end{aligned} \tag{B.5a}$$

Similarly we have

$$\boldsymbol{x}_b' = R_b \boldsymbol{x}_b + \mathring{\boldsymbol{T}}_b. \tag{B.5b}$$

For the two solutions corresponding to subscripts a and b, respectively, (B.2) guarantees all the points lie in the plane before motion:

$$\tilde{\boldsymbol{N}}_a^t \boldsymbol{x}_a = 1, \qquad \tilde{\boldsymbol{N}}_b^t \boldsymbol{x}_b = 1. \tag{B.6}$$

Equations (B.5a) and (B.5b) are equations of the rigid motion, so the points after motion also lie in a plane. Furthermore, (B.5) states that the two solutions have the motions corresponding to R_a, $\mathring{\boldsymbol{T}}_a$ and R_b, $\mathring{\boldsymbol{T}}_b$, respectively. Equation (B.4) ensures that both solutions render the same images. □

Appendix C

Theorem 7: Rank $(A) \leq 8$ if and only if there exists a 3×3 nonzero matrix K such that all the points lie in the following intersection of two quadratic surfaces before motion:

$$(\boldsymbol{x} - \boldsymbol{O}') \times K \boldsymbol{x} = \boldsymbol{0} \tag{C.1}$$

where $\boldsymbol{O}' = -R^t \boldsymbol{T}$.

Proof: First assume that the rank of A is less than 9. Then there exists a nonzero vector $\boldsymbol{h}$ such that (2.20) and (2.19) hold for all the points. Since

$$\begin{aligned} &[-v'\boldsymbol{X}^t \quad u'\boldsymbol{X}^t \quad 0] \\ &\quad = -v[\boldsymbol{X}^t \quad \boldsymbol{0} \quad -u'\boldsymbol{X}^t] + u[\boldsymbol{0} \quad \boldsymbol{X}^t \quad -v'\boldsymbol{X}^t] \end{aligned} \tag{C.2}$$

the third equation in (2.19) is redundant once the first two equations are given. Thus (2.19) gives (2.18). From (2.3) and (2.17) we obtain

$$\boldsymbol{x}' \times F \boldsymbol{x} = \boldsymbol{0}. \tag{C.3}$$

From (2.4) we have

$$(R\boldsymbol{x} + \boldsymbol{T}) \times F \boldsymbol{x} = \boldsymbol{0} \tag{C.4}$$

or

$$(\boldsymbol{x} + R^t \boldsymbol{T}) \times R^t F \boldsymbol{x} = \boldsymbol{0}. \tag{C.5}$$

Letting $K = R^t F$ and $O' = -R^t \boldsymbol{T}$, we get (C.1).

Conversely, assume (C.1) holds. Letting $F = RK$ ($F \neq O$) we have (C.5). Thus (C.4) and (C.3) hold. From (2.3) and the fact that the depths of the points are not zeros, we have (2.18) and then (2.19) and (2.20). Since $\boldsymbol{h} \neq 0$, the rank of A is less than 9. □

Appendix D

Theorem 8: Let F and F' correspond, respectively, to two motions from one position. Suppose there exist two solutions $(R_a, \tilde{\boldsymbol{T}}_a, \mathring{\boldsymbol{N}}_a)$ and $(R_b, \tilde{\boldsymbol{T}}_a, \mathring{\boldsymbol{N}}_b)$ such that

$$F = R_a + \tilde{\boldsymbol{T}}_a \mathring{\boldsymbol{N}}_a^t = R_b + \tilde{\boldsymbol{T}}_b \mathring{\boldsymbol{N}}_b^t. \tag{D.1}$$

Let H be an orthogonal matrix with $\det(H) = 1$ such that

$$H^t F^t F H = \operatorname{diag}(\lambda_1, 1, \lambda_3) \tag{D.2}$$

and $\lambda_1 < 1 < \lambda_3$.

Similarly, for F' we have two solutions such that

$$F' = R_a' + \tilde{\boldsymbol{T}}_a' (\mathring{\boldsymbol{N}}_a')^t = R_b' + \tilde{\boldsymbol{T}}_b' (\mathring{\boldsymbol{N}}_b')^t.$$

Then, both solutions agree on the unit normal, i.e.,

$$\mathring{\boldsymbol{N}}_a = \mathring{\boldsymbol{N}}_a', \quad \text{and} \quad \mathring{\boldsymbol{N}}_b = \mathring{\boldsymbol{N}}_b' \tag{D.3}$$

if and only if the ambiguity condition is satisfied, i.e., there exists an orthogonal matrix Q and a positive number k such that

$$\begin{aligned} F' = Q \operatorname{diag}(&\sqrt{1 - k(1 - \lambda_1)}, \quad 1, \\ &\sqrt{1 + k(\lambda_3 - 1)}) H^t \end{aligned} \tag{D.4a}$$

with $1 - k(1 - \lambda_1) > 0$ or

$$F' = Q \operatorname{diag}(\sqrt{1 + k(1 - \lambda_1)}, \quad 1 \quad \sqrt{1 - k(\lambda_3 - 1)}) H^t \tag{D.4b}$$

with $1 - k(\lambda_3 - 1) \geq 0$.

Proof: Specify the columns of H by $H = [\boldsymbol{h}_1' \ \boldsymbol{h}_2' \ \boldsymbol{h}_3']$. Let $H' = [\boldsymbol{h}_1', \boldsymbol{h}_2', \boldsymbol{h}_3']$ be orthogonal matrices such that

$$(H')^t(F')^tF'H' = \text{diag}\,(\lambda_1', 1, \lambda_3'). \tag{D.5}$$

If we change the signs of the columns of H or H', (D.2) and (D.5) still hold. So we can assume $\det(H) = \det(H') = 1$. Then the column vectors of H and H' form a right-hand orthonormal basis, respectively.

We first prove the necessity part. Assume (D.3) holds. We apply the algorithm to F and F'. From (Ag.12) and the sign determined by (Ag.14), we have (using (Ag.10))

$$\pm\mathring{\boldsymbol{N}}_a = \boldsymbol{V}_1 \times \boldsymbol{V}_2 = (\alpha\boldsymbol{h}_1 + \beta\boldsymbol{h}_3) \times \boldsymbol{h}_2$$
$$= \alpha(\boldsymbol{h}_1 \times \boldsymbol{h}_2) + \beta(\boldsymbol{h}_3 \times \boldsymbol{h}_2) = \alpha\boldsymbol{h}_3 - \beta\boldsymbol{h}_1 \tag{D.6a}$$

and for the second solution:

$$\pm\mathring{\boldsymbol{N}}_b = (\alpha\boldsymbol{h}_1 - \beta\boldsymbol{h}_3) \times \boldsymbol{h}_2$$
$$= \alpha(\boldsymbol{h}_1 \times \boldsymbol{h}_2) - \beta(\boldsymbol{h}_3 \times \boldsymbol{h}_2) = \alpha\boldsymbol{h}_3 + \beta\boldsymbol{h}_1. \tag{D.6b}$$

Similarly, we have the equations corresponding to (D.6a) and (D.6b) for F':

$$\pm\mathring{\boldsymbol{N}}_a' = \boldsymbol{V}_1' \times \boldsymbol{V}_2' = \alpha'\boldsymbol{h}_3' - \beta'\boldsymbol{h}_1' \tag{D.7a}$$

$$\pm\mathring{\boldsymbol{N}}_b' = \boldsymbol{V}_1' \times \boldsymbol{V}_2' = \alpha'\boldsymbol{h}_3' + \beta'\boldsymbol{h}_1'. \tag{D.7b}$$

If the first solution for F' gives $\mathring{\boldsymbol{N}}_a(\mathring{\boldsymbol{N}}_a' = \mathring{\boldsymbol{N}}_a)$, we change, if necessary, the sign of $(\boldsymbol{h}_1', \boldsymbol{h}_3')$ without change $\det(H')$ to make the signs of (D.6a) and (D.7a) the same. Then (D.6) and (D.7) give

$$\alpha\boldsymbol{h}_3 - \beta\boldsymbol{h}_1 = \alpha'\boldsymbol{h}_3' - \beta'\boldsymbol{h}_1'$$
$$\alpha\boldsymbol{h}_3 + \beta\boldsymbol{h}_1 = s_1(\alpha'\boldsymbol{h}_3' + \beta'\boldsymbol{h}_1') \tag{D.8.I}$$

where $s_1 \in \{-1, 1\}$.

Otherwise, the second solution for F' gives $\mathring{\boldsymbol{N}}_a(\mathring{\boldsymbol{N}}_b' = \mathring{\boldsymbol{N}}_a)$. Then we change, if necessary, the sign of $(\boldsymbol{h}_1', \boldsymbol{h}_2')$ without changing $\det(H')$ to make the signs of (D.6a) and (D.6b) the same. Then (D.6) and (D.6) give

$$\alpha\boldsymbol{h}_3 - \beta\boldsymbol{h}_1 = \alpha'\boldsymbol{h}_3' + \beta'\boldsymbol{h}_1'$$
$$\alpha\boldsymbol{h}_3 + \beta\boldsymbol{h}_1 = s_2(\alpha'\boldsymbol{h}_3' - \beta'\boldsymbol{h}_1') \tag{D.8.II}$$

where $s_2 \in \{-1, 1\}$. Now consider (D.8.I). Suppose $s_1 = 1$. Summing and subtracting the corresponding sides of the two equations in (D.8.I) we have

$$2\alpha\boldsymbol{h}_3 = 2\alpha'\boldsymbol{h}_3', \qquad 2\beta\boldsymbol{h}_1 = 2\beta\boldsymbol{h}_1'. \tag{D.9}$$

Since α and β are positive and H and H' are orthogonal matrices with unit determinants, we have

$$\alpha' = \alpha, \qquad \beta' = \beta, \qquad H = H'. \tag{D.10}$$

Similarly, we get the similar equations for $s_1 = -1$. We also consider the two cases in (D.8.II). Thus, in total we get four cases:

$$\text{Ia:}\begin{cases}\alpha' = \alpha\\ \beta' = \beta\\ H' = H\end{cases} \qquad \text{Ib:}\begin{cases}\alpha' = \beta\\ \beta' = \alpha\\ H' = H\,\text{rdiag}\,(1, 1, -1)\end{cases} \tag{D.11.I}$$

$$\text{IIa:}\begin{cases}\alpha' = \alpha\\ \beta = \beta\\ H' = H\,\text{diag}\,(-1, -1, 1)\end{cases}$$

$$\text{IIb:}\begin{cases}\alpha' = \beta\\ \beta' = \alpha\\ H' = H\,\text{rdiag}\,(-1, 1, 1)\end{cases} \tag{D.11.II}$$

where $\text{rdiag}\,(\lambda_1, \lambda_2, \lambda_3)$ denotes the (unconventional) diagonal matrix where the diagonal line is defined from top right to bottom left. In fact, cases Ia and IIa give (D.4a), and Ib and IIb give (D.4b). We prove that the case IIb gives (D.4b). The proofs for the other cases are similar and so are omitted. From the case IIb in (D.11.II), using (Ag.9) we have

$$\frac{\lambda_3' - 1}{\lambda_3' - \lambda_1'} = \frac{1 - \lambda_1}{\lambda_3 - \lambda_1}, \qquad \frac{1 - \lambda_1'}{\lambda_3' - \lambda_1'} = \frac{\lambda_3 - 1}{\lambda_3 - \lambda_1}. \tag{D.12}$$

Letting $k = (\lambda_3' - \lambda_1')/(\lambda_3 - \lambda_1)$ it follows that

$$\lambda_1' = 1 - k(\lambda_3 - 1), \qquad \lambda_3' = 1 + k(1 - \lambda_1). \tag{D.13}$$

Since $k > 0$, and $\lambda_1 < 1 < \lambda_3$, we know $\lambda_1' < 1 < \lambda_3'$. Then (D.5), (D.13), and (D.11.II) give

$$(F')^tF' = H'\,\text{diag}\,(\lambda_1', 1, \lambda_3')(H')^t$$
$$= H\,\text{rdiag}\,(-1, 1, 1)\,\text{diag}\,(\lambda_1', 1, \lambda_3')$$
$$\cdot\,\text{rdiag}\,(-1, 1, 1)^tH^t$$
$$= H\,\text{diag}\,(\lambda_3', 1, \lambda_1')H^t$$
$$= H\,\text{diag}\,(1 + k(1 - \lambda_1), 1, 1 - k(\lambda_3 - 1))H^t. \tag{D.14}$$

Let $C = \text{diag}\,(\sqrt{1 + k(1 - \lambda_1)}, 1, \sqrt{1 - k(\lambda_3 - 1)})H^t$. Noticing (D.5) and (D.14) we know that $F'H$ and CH are orthogonal matrices. From (D.13) we know that one can establish column correspondences, according to vector length, between $F'H$ and CH so that the corresponding column vectors have the same length. Therefore there exists an orthogonal matrix (rotation or reflection) such that $F'H = QCH$, i.e., $F' = QC$. This is (D.4b).

Now we prove the sufficiency part. If we assume (D.4a) or (D.4b) holds, we prove (D.3). We present the proof for the case where (D.4b) holds. The proof corresponding to (D.4a) is similar and is omitted. Letting $H' = [\boldsymbol{h}_3 \ \boldsymbol{h}_2 \ \boldsymbol{h}_1]$, we get (D.5), where $\lambda_1' = 1 - k(\lambda_3 - 1)$, $\lambda_3' = 1 + k(1 - \lambda_1)$. By the algorithm we have (we use

prime to denote the corresponding variables for F')

$$\alpha' = \sqrt{\frac{\lambda_3' - 1}{\lambda_3' - \lambda_1'}} = \sqrt{\frac{1 - \lambda_1}{\lambda_3 - \lambda_1}} = \beta,$$

$$\beta' = \sqrt{\frac{1 - \lambda_1'}{\lambda_3' - \lambda_1'}} = \sqrt{\frac{\lambda_3 - 1}{\lambda_3 - \lambda_1}} = \alpha \qquad \text{(D.15)}$$

$$\boldsymbol{V}_1' = \alpha' \boldsymbol{h}_1' + \beta' \boldsymbol{h}_3' = \beta \boldsymbol{h}_3 + \alpha \boldsymbol{h}_1 \qquad \text{(D.16)}$$

$$\boldsymbol{V}_2' = \boldsymbol{h}_2 \qquad \text{(D.17)}$$

and

$$\pm \mathring{\boldsymbol{N}}_a' = \boldsymbol{V}_1' \times \boldsymbol{V}_2' = \beta \boldsymbol{h}_3 \times \boldsymbol{h}_2 + \alpha \boldsymbol{h}_1 \times \boldsymbol{h}_2$$
$$= -\beta \boldsymbol{h}_1 + \alpha \boldsymbol{h}_3 = (\alpha \boldsymbol{h}_1 + \beta \boldsymbol{h}_3) \times \boldsymbol{h}_2 = \pm \mathring{\boldsymbol{N}}_a. \qquad \text{(D.18)}$$

From the definition of F in (2.16), we know that given an F the sign of the normal $\boldsymbol{N}$ can be coupled with the sign of $\boldsymbol{T}$. In other words, we can let

$$\mathring{\boldsymbol{N}}_a' = \mathring{\boldsymbol{N}}_a \qquad \text{(D.19)}$$

and according to Theorem 6, the algorithm determines R_a', $\tilde{\boldsymbol{T}}_a$ such that

$$F' = R_a' + \tilde{\boldsymbol{T}}_a \mathring{\boldsymbol{N}}_a'. \qquad \text{(D.20)}$$

Similarly, we have $\mathring{\boldsymbol{N}}_b' = \mathring{\boldsymbol{N}}_b$ and

$$F' = R_b' + \tilde{\boldsymbol{T}}_b \mathring{\boldsymbol{N}}_b'. \qquad \text{(D.21)}$$

□

ACKNOWLEDGMENT

The authors would like to thank X. Zhuang for helpful discussions.

REFERENCES

[1] L. Dreschler and H.-H. Nagel, "Volumetric model and 3-D trajectory of a moving car derived from monocular TV frame sequences of a street scene," *Comput. Graphics Image Processing*, vol. 20, pp. 199-228, 1982.

[2] O. D. Faugeras, F. Lustman, and G. Toscani, "Motion and structure from point and line matches," in *Proc. Int. Conf. Comput. Vision* (London, England), June 1987.

[3] W. K. Gu, J. Y. Yang, and T. S. Huang, "Matching perspective views of a 3-D object using composite circuits," in *Proc. 7th Int. Conf. Pattern Recognition*, July-August 1984.

[4] L. Kitchen and A. Rosenfeld, "Gray-level corner detection," *Patt. Recog. Lett.*, vol. 1, pp. 95-102, 1982.

[5] H. C. Longuet-Higgins, "A computer program for reconstructing a scene from two projections," *Nature*, vol. 293, pp. 133-135, Sept. 1981.

[6] H. C. Longuet-Higgins, "The visual ambiguity of a moving plane," *Proc. Roy. Soc. London*, B223, pp. 165-175, 1984.

[7] H. C. Longuet-Higgins, "The reconstruction of a scene from two projections—configurations that defeat the 8-point algorithm," in *Proc. 1st Conf. Artificial Intell. Appl.* (Denver, CO), Dec. 5-7, 1984, pp. 395-397.

[8] H.-H. Nagel and W. Enkelmann, "An investigation of smoothness constraints for the estimation of displacement vector fields from image sequences," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-8, no. 5, pp. 565-593, 1986.

[9] V. Torre and T. Poggio, "On edge detection," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-8, no. 2, pp. 147-163, 1986.

[10] R. Y. Tsai and T. S. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch, II: Singular value decomposition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, no. 4, pp. 525-534, 1983.

[11] R. Y. Tsai and T. S. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch, III: Finite point correspondences and the three-view problem," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, no. 2, pp. 213-220, 1984.

[12] R. Y. Tsai and T. S. Huang, "Uniqueness and estimation of 3-D motion parameters of rigid bodies with curved surfaces," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-6, pp. 13-27, 1984.

[13] A. Verri and T. Poggio, "Against quantitative optical flow," in *Proc. First Int. Conf. Comput. Vision* (London, England), June 1987, pp. 171-180.

[14] A. M. Waxman and K. Wohn, "Contour evolution, neighborhood deformation, and global image flow: Planar surfaces in motion," *Int. J. Robotics Res.*, vol. 4, pp. 95-108, 1985.

[15] J. Weng, N. Ahuja, and T. S. Huang, "Error analysis of motion parameter estimation from image sequences," in *Proc. 1st Int. Conf. Comput. Vision* (London, England), June 1987, pp. 703-707.

[16] J. Weng, T. S. Huang, and N. Ahuja, "3-D motion estimation, understanding, and prediction from noisy image sequences," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-9, no. 3, pp. 370-389, 1987.

[17] J. Weng, N. Ahuja, and T. S. Huang, "Closed-form solution and maximum likelihood: A robust approach to motion and structure estimation," in *Proc. IEEE Conf. Comput. Vision Pattern Recog.* (Ann Arbor, MI), June 5-9, 1988, pp. 381-386.

[18] J. Weng, N. Ahuja, and T. S. Huang, "Two-view matching," in *Proc. 2nd Int. Conf. Comput. Vision* (Florida), Dec. 1988, pp. 64-73.

[19] J. Weng, T. S. Huang, and N. Ahuja, "Motion and structure from two perspective views: Algorithm, error analysis, and error estimation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, no. 5, pp. 451-476, May 1989.

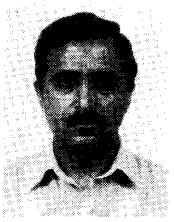
[20] J. Weng, T. S. Huang, and N. Ahuja, "Motion from images: Image matching, parameter estimation, and intrinsic stability," in *Proc. IEEE Workshop Visual Motion* (Irvine, CA), Mar. 20-22, 1989, pp. 359-366.

[21] O. A. Zuniga and R. M. Haralick, "Corner detection using the facet model," in *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, 1983, pp. 30-37.

Juyang Weng (S'85-M'88) received the B.S. degree in computer science from Fudan University, Shanghai, China, in 1982. He received the M.S. and Ph.D. degrees, both in computer science, from the University of Illinois, Urbana-Champaign, in 1985 and 1988, respectively.

From September 1984 to December 1988, he was a Research Assistant at the Coordinated Science Laboratory, University of Illinois, Urbana-Champaign. In the summer of 1987 he was employed at IBM Los Angeles Scientific Center, Los Angeles. Since January 1989 he has been a Researcher at Centre de Recherche Informatique de Montreal, Montreal, Que., Canada, and an adjunct Researcher at Ecole Polytechnique, University of Montreal, Montreal, Que., Canada. Since October 1990 he has been a Visiting Assistant Professor at the University of Illinois, Urbana-Champaign. His current research interests are in computer vision, image processing, object modeling and representation, parallel architecture for image processing, autonomous navigation, and artificial intelligence.

Dr. Weng is a member of Phi Kappa Phi and Sigma Xi.



Narendra Ahuja (S'79-M'79-SM'85) received the B.E. degree with honors in electronics engineering from the Birla Institute of Technology and Science, Pilani, India, in 1972, the M.E. degree with distinction in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1974, and the Ph.D. degree in computer science from the University of Maryland, College Park, in 1979.

From 1974 to 1975 he was Scientific Officer in the Department of Electronics, Government of India, New Delhi. From 1975 to 1979 he was at the Computer Vision Laboratory, University of Maryland, College Park. Since 1979 he has been with the University of Illinois at Urbana-Champaign. He has been a Professor in the Department of Electrical and Computer Engineering, the Coordinated Science Laboratory, and the Beckman Institute since 1988. His interests are in computer vision, robotics, image processing, and parallel algorithms. He has been involved in teaching, research, consulting, and organizing conferences in these areas. His current research emphasizes the integrated use of multiple image sources of scene information to construct three-dimensional descriptions of scenes, the use of the acquired three-dimensional information for object manipulation and navigation, and multiprocessor architectures for computer vision.

Dr. Ahuja was selected as a Beckman Associate in the University of Illinois Center for Advanced Study for 1990-1991. He received the University Scholar Award in 1985, the Presidential Young Investigator Award in 1984, the National Scholarship (1967-1972), and the President's Merit Award in 1966. He has coauthored the book *Pattern Models* (Wiley, 1983) with Bruce Schachter. He is Associate Editor of the journals IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, *Computer Vision, Graphics, and Image Processing*, and *Journal of Mathematical Imaging and Vision*. He is a member of the American Association for Artificial Intelligence, the Society of Photo-Optical Instrumentation Engineers, and the Association for Computing Machinery.

Thomas S. Huang (S'61-M'63-SM'76-F'79) received the B.S. degree in electrical engineering from National Taiwan University, Taipei, Taiwan, Republic of China, and the M.S. and Sc.D. degrees in electrical engineering from the Massachusetts Institute of Technology, Cambridge.

He was on the faculty of the Department of Electrical Engineering at M.I.T. from 1963 to 1973; and on the faculty of the School of Electrical Engineering and Director of its Laboratory for Information and Signal Processing at Purdue University from 1973 to 1980. In 1980, he joined the University of Illinois at Urbana-Champaign, where he is now Professor of Electrical and Computer Engineering and Research Professor at the Coordinated Science Laboratory. During his sabbatical leaves he was with the M.I.T. Lincoln Laboratory, the IBM Thomas J. Watson Research Center, and the Rheinishes Landes Museum in Bonn, Germany. He held Visiting Professor positions at the Swiss Institutes of Technology in Zürich and Lausanne; the University of Hannover in Germany; and INRS-Telecommunications of the University of Quebec in Montreal, Canada. He has served as a consultant to numerous industrial firms and government agencies, both in the U.S. and abroad. His interests lie in the broad area of information technology, especially the transmission and processing of multidimensional signals. He has written 10 published books, and over 200 papers in network theory, digital filtering, image processing, and computer vision.

Dr. Huang is a Fellow of the Optical Society of America. He received a Guggenheim Fellowship (1971-1972), an A. V. Humboldt Foundation Senior U.S. Scientist Award (1976-1977), and a Fellowship from the Japan Association for the Promotion of Science (1986). He received the IEEE Acoustics, Speech, and Signal Processing Society's Technical Achievement Award in 1987. He is an Editor for the international journal, *Computer Vision, Graphics, and Image Processing*; and is Editor of the Springer Series in Information Sciences, published by Springer-Verlag; and is Editor of the Research Annual Series on Advances in Computer Vision and Image Processing published by the JAI Press.