noise electron tubes, and superconductive parametric and storage devices. In 1965 he performed an experiment which first proved the existence of the Magnus force in superconductors and was a co-recipient of the RCA Research Award for the development of a superconductive parametric amplifier. In 1966 he became Director of Advanced Memory Devices at Electronic Memories, Inc., Hawthorne, CA, heading the development of plated wire memories. In 1969 he joined the School of Engineering and Applied Science, University of California, Los Angeles, where he is a Professor of Computer Science and Director of the Cognitive Systems Laboratory. His present interests lie in artificial intelligence, decision theory, signal interpretation, and complexity of algorithms. He has published over 60 technical papers in his fields of interest. He is on the Editorial Boards of the *Artificial Intelligence* journal, the journal of *Cybernetics and Systems*, and serves as a Chairman of the IEEE SMC Committee on Knowledge-Based Systems. He has edited *Search and Heuristics* (North-Holland, 1983) and is the author of *Heuristics: Intelligence Search Strategies* (Addison-Wesley, 1984).

Dr. Pearl is a member of the Association for Computing Machinery (SIGART) and the American Association of Artificial Intelligence.

Uniqueness and Estimation of Three-Dimensional Motion Parameters of Rigid Objects with Curved Surfaces

ROGER Y. TSAI, MEMBER, IEEE, AND THOMAS S. HUANG, FELLOW, IEEE

***Abstract*—Two main results are established in this paper. First, we show that seven point correspondences are sufficient to uniquely determine from two perspective views the three-dimensional motion parameters (within a scale factor for the translations) of a rigid object with curved surfaces. The seven points should not be traversed by two planes with one plane containing the origin, nor by a cone containing the origin. Second, a set of "essential parameters" are introduced which uniquely determine the motion parameters up to a scale factor for the translations, and can be estimated by solving a set of linear equations which are derived from the correspondences of eight image points. The actual motion parameters can subsequently be determined by computing the singular value decomposition (SVD) of a 3 × 3 matrix containing the essential parameters.**

***Index Terms*—Dynamic scene analysis, image sequence analysis, motion estimation, motion stereo, optical flow.**

I. Introduction

The importance of motion estimation in image sequence analysis has long been recognized, particularly in such fields as image coding, tracking, and robotic vision. Methods for two-dimensional motion estimation are relatively well known [1]-[8]. As for three-dimensional motion estimation from two image frames with a single camera, when the object surface is not restricted to be of any particular shape, past theoretical analyses and estimation schemes were unsatisfactory in the sense that, theoretically, it was not known how many image point correspondences are needed to ensure the uniqueness of the motion parameters (up to a scale factor for the translation parameters), and practically, almost all estimation schemes rely on the solution of nonlinear equations using iterative search [9]-[15], unless very special assumptions or simplifications are made. Examples are: the fixed axis assumption [15] which is satisfied only by those movements consisting of translations and rotations around a fixed axis over several frames; the parallel projection assumption [10], [15] which is approximately true only if the object is far away from the camera, and the pure translations assumption [16], [17]. It is important to observe that deriving the nonlinear equations relating the image coordinates to the object surface geometry and the motion parameters is not a difficult task since only geometry and simple algebraic manipulation are involved. However, once the nonlinear equations are obtained, two challenging problems remain. First, how many image point correspondences are sufficient to guarantee uniqueness of solution? And second, how can one avoid iterative search when computing the solutions of the nonlinear equation? For the uniqueness problem, almost all previous methods that do not make any assumptions or simplifications merely tried to get the number of equations to be equal to the number of unknowns [11]-[13], [18], [19]. In solving nonlinear equations iteratively, the computational complexity is almost in-

Manuscript received November 4, 1981; revised January 27, 1983. This work was supported in part by the Joint Service Electronics Program under Contract N00014-79-C-0424 and was performed in part while the authors were at INRS Telecommunications/Bell Northern Research, Montreal, P.Q., Canada.

R. Y. Tsai was with the Coordinated Science Laboratory and the Department of Electrical Engineering, University of Illinois, Urbana, IL 61801. He is now with the IBM T. J. Watson Research Center, Yorktown Heights, NY 10598.

T. S. Huang is with the Coordinated Science Laboratory and the Department of Electrical Engineering, University of Illinois, Urbana, IL 61801.

finite in general since global search is needed, unless the initial guess is good enough so that the local minimum obtained is also the global minimum. The following statement quoted from [10, p. 157] can best reveal the situation: "In any case, the general method was not really practicable, nor was it designed for efficient use." In contrast, our methods to be described in this paper do not require iterative search, and our proof for uniqueness is rigorous. The computational tasks in our methods are just computing the least-square solution of a system of n linear equations with 8 unknown (n depends on the number of point correspondences used, and is greater than or equal to 8) and the singular value decomposition (SVD) of a 3×3 matrix. If the normal equation approach is used in solving the linear least square problem, then forming the normal equation takes $0(8^2 n)$ operations and solving the 8 linear equations (the normal equations) with LU decomposition and forward-backward substitutions take $8^3/3 + 0(8^2) \approx 234$ operations [20]. Computing the SVD of a 3×3 matrix consists of two phases [21]. The first phase is the householder transformation which requires about $0(3^2)$ operations. The second phase is the implicit QR iterations or "chasing," which almost always converges cubically as shown by Wilkinson [22]. The complexity is roughly $0(3^2)$. Of past methods, only those that make some sort of simplifications or restrictions can provide rigorous proof of uniqueness and can avoid using iterative search in obtaining the solution. For those methods that treat the general situation, Nagel's [18] and Praxdny's [13], [14] methods are slightly different from the others in the sense that they cancelled the unknowns corresponding to the translation parameters in the system of nonlinear equations by algebraic manipulations, leaving fewer unknowns in the reduced system of nonlinear equations. However, they did not resolve the uniqueness problem, and their methods still require iterative search. Without rigorous proof of uniqueness, one cannot be sure of the conditions for the solutions of the nonlinear equations to be unique. Thus, Praxdny tested his method on a moving planar patch and regard the solution computed via iterative search as the only solution, while it was shown in [23]-[26] that given two views of a moving planar patch, there are actually two possible solutions (aside from a scale factor) for the motion and geometrical parameters. In [10], although the uniqueness problem is proved for the parallel projection case, no proof was given for the central projection case. Computer experiments in [10] showed that in the case of central projection five point correspondences are usually but not always sufficient to ensure uniqueness of solution. In this paper, we present new results for the uniqueness problem for the central projection case.

II. The E Matrix and the Eight Essential Parameters

The basic geometry of the problem is sketched in Fig. 1. Consider a particular point P on an object. Let

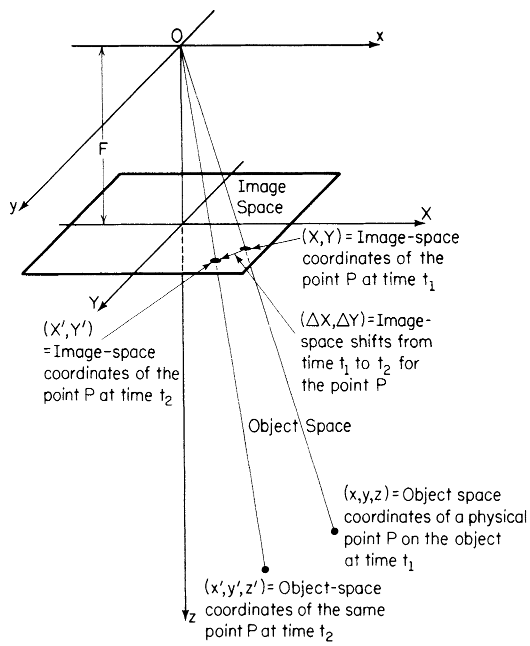


Fig. 1. Basic geometry for three-dimensional motion estimation.

(x, y, z) = object-space coordinates of a point P before motion
(x', y', z') = object-space coordinates of P after motion
(X, Y) = image-space coordinates of P before motion
(X', Y') = image-space coordinates of P after motion

The mapping $(X, Y) \rightarrow (X', Y')$ for a particular point is called an image point correspondence. It is well known [27] that any 3-D rigid body motion is equivalent to a rotation by an angle θ around an axis through the origin with directional cosines $n1, n2, n3$, followed by a translation $(\Delta x, \Delta y, \Delta z)$,

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} + T \tag{1}$$

where R is a 3×3 orthonormal matrix of the first kind (i.e., det $(R) = 1$)

$$R = \begin{bmatrix} n1^2 + (1 - n1^2)\cos\theta & n1 \cdot n2(1 - \cos\theta) - n3 \sin\theta & n1 \cdot n3(1 - \cos\theta) + n2 \sin\theta \\ n1 \cdot n2(1 - \cos\theta) + n3 \sin\theta & n2^2 + (1 - n2^2)\cos\theta & n2 \cdot n3(1 - \cos\theta) - n1 \sin\theta \\ n1 \cdot n3(1 - \cos\theta) - n2 \sin\theta & n2 \cdot n3(1 - \cos\theta) + n1 \sin\theta & n3^2 + (1 - n3^2)\cos\theta \end{bmatrix} \tag{2}$$

$$\triangleq \begin{bmatrix} r1 & r2 & r3 \\ r4 & r5 & r6 \\ r7 & r8 & r9 \end{bmatrix} \quad \text{and} \quad T \triangleq \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}.$$

Although the elements in R, namely, $r1, r2, \cdots, r9$, are nonlinear functions of the rotation parameters $n1$, $n2$, $n3$, and θ, throughout this paper, the uniqueness and computation of R rather than $n1, n2, n3$, and θ are discussed. The reason is two-fold. First, as will be seen later, to each possible R in (2), there corresponds exactly two sets of rotation parameters $n1$, $n2$, $n3$, θ with one set the negative of the other. Since these two solutions are physically indistinguishable, we may regard the relationship between R and the rotation parameters as one to one. The second reason is that once R is determined, the task of computing $n1$, $n2$, $n3$ and θ is straightforward, as can be seen in the following.

From (2), we have

$$R = S + K$$

where

$$S = \begin{bmatrix} n1^2 + (1 - n1^2)\cos\theta & n1n2(1 - \cos\theta) & n1n3(1 - \cos\theta) \\ n1n2(1 - \cos\theta) & n2^2 + (1 - n2^2)\cos\theta & n2n3(1 - \cos\theta) \\ n1n3(1 - \cos\theta) & n2n3(1 - \cos\theta) & n3^2 + (1 - n3^2)\cos\theta \end{bmatrix}$$

is symmetric and

$$K = \sin\theta \cdot \begin{bmatrix} 0 & -n3 & +n2 \\ +n3 & 0 & -n1 \\ -n2 & +n1 & 0 \end{bmatrix}$$

is skew-symmetric.

Since any matrix can be decomposed uniquely into a sum of a symmetric and a skew-symmetric matrix, we see that K is unique given R, and thus $n1$, $n2$, $n3$, θ are fixed up to a possible sign change. In fact, it is trivial to see that

$$K = \frac{1}{2}\begin{bmatrix} 0 & r2 - r4 & r3 - r7 \\ r4 - r2 & 0 & r8 - r6 \\ r7 - r3 & r8 - r7 & 0 \end{bmatrix}$$

or $n1 \cdot \sin\theta = (r8 - r6)/2$, $n2 \cdot \sin\theta = (r3 - r7)/2$, $n3 \cdot \sin\theta = (r4 - r2)/2$, which imply $\sin^2\theta(n1^2 + n2^2 + n3^2) = \sin^2\theta \cdot 1 = d/4$ or

$$\sin\theta = \pm d/2, \qquad n1 = \pm(r8 - r6)/d,$$
$$n2 = \pm(r3 - r7)/d, \qquad n3 = \pm(r4 - r2)/d$$

where $d = (r8 - r6)^2 + (r3 - r7)^2 + (r4 - r2)^2$. (If $d = 0$, then $\theta = 0$, $R = I$, and $n1$, $n2$, $n3$ can be anything since without rotation, the axis is meaningless.) Since $\sin\theta$ alone does not determine θ uniquely, we still need $\cos\theta$ to fix θ. From (2), $n1^2 + (1 - n1^2)\cos\theta = r1$,

$$\cos\theta = \frac{r1 - n1^2}{1 - n1^2} = \frac{r1 - \left(\frac{r8 - r6}{d}\right)^2}{1 - \left(\frac{r8 - r6}{d}\right)^2} = \frac{d^2 r1 - (r8 - r6)^2}{d^2 - (r8 - r6)^2}.$$

Therefore, θ, $n1$, $n2$, and $n3$ can be easily determined from R.

We now combine (1) with the following equations relating the object and image space coordinates:

$$X = x/z \qquad X' = x'/z'$$
$$Y = y/z \qquad Y' = y'/z' \tag{3}$$

to obtain

$$X' = \frac{(r1\,X + r2\,Y + r3)\,z + \Delta x}{(r7\,X + r8\,Y + r9)\,z + \Delta z} \tag{4a}$$

$$Y' = \frac{(r4\,X + r5\,Y + r6)\,z + \Delta y}{(r7\,X + r8\,Y + r9)\,z + \Delta z} \tag{4b}$$

where the focal length F is normalized to 1 for simplicity. From (4),

$$z = \frac{\Delta x - \Delta z \cdot X'}{X'(r7\,X + r8\,Y + r9) - (r1\,X + r2\,Y + r3)} \tag{5a}$$

and

$$z = \frac{\Delta y - \Delta z \cdot Y'}{Y'(r7\,X + r8\,Y + r9) - (r4\,X + r5\,Y + r6)}. \tag{5b}$$

Equating the right-hand sides of (5a) and (5b) gives

$$[X' \quad Y' \quad 1] \; E \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = 0 \tag{6}$$

where

$$E \triangleq \begin{bmatrix} \Delta z \cdot r4 - \Delta y \cdot r7 & \Delta z \cdot r5 - \Delta y \cdot r8 & \Delta z \cdot r6 - \Delta y \cdot r9 \\ \Delta x \cdot r7 - \Delta z \cdot r1 & \Delta x \cdot r8 - \Delta z \cdot r2 & \Delta x \cdot r9 - \Delta z \cdot r3 \\ \Delta y \cdot r1 - \Delta x \cdot r4 & \Delta y \cdot r2 - \Delta x \cdot r5 & \Delta y \cdot r3 - \Delta x \cdot r6 \end{bmatrix} \tag{7}$$

$$\triangleq \begin{bmatrix} e1 & e2 & e3 \\ e4 & e5 & e6 \\ e7 & e8 & e9 \end{bmatrix}. \tag{8}$$

Note that the equality of (6) will not be influenced by multiplying E with any scalar. Since each element of E is linear in Δx, Δy, and Δz, this simply means that there is a common scale factor for the translation parameters that cannot be determined. (This scale factor also influences z in (5a) and (5b), but not the rotation parameters.) For this reason, we can set $e9$ equal to some fixed number, say 1, without losing generality. We call the elements in E "essential parameters," because as we shall see presently the actual 3-D motion parameters can be determined uniquely given E, and can be computed by taking the SVD of E without having to solve nonlinear equations. Furthermore, given the image correspondences of eight object points in general positions, the E matrix can be determined uniquely by solving eight linear equations.

Before giving Theorem I (which concerns the uniqueness and the computation of motion parameters given the matrix E), let us first analyze the matrix E. From (7), we have

$$E = \begin{bmatrix} \Delta z & & \\ & \Delta x & \\ & & \Delta y \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} R - \begin{bmatrix} \Delta y & & \\ & \Delta z & \\ & & \Delta x \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} R$$

$$= \begin{bmatrix} & \Delta z & \\ & & \Delta x \\ \Delta y & & \end{bmatrix} R - \begin{bmatrix} & & \Delta y \\ \Delta z & & \\ & \Delta x & \end{bmatrix} R = G\,R \tag{9}$$

where

$$G \triangleq \begin{bmatrix} 0 & \Delta z & -\Delta y \\ -\Delta z & 0 & \Delta x \\ \Delta y & -\Delta x & 0 \end{bmatrix} \tag{10}$$

is skew-symmetric and contains only the translation parameters and R is the rotation matrix. It is well known in matrix theory [28] that any skew-symmetric matrix K must have even rank, say $2n$, and that K, if real, always assumes the following normal form:

$$K = Q^T \begin{bmatrix} 0 & \varphi_1 & & & & & & & & \\ -\varphi_1 & 0 & & & & & & & & \\ & & 0 & \varphi_2 & & & & & & \\ & & -\varphi_2 & 0 & & & & & & \\ & & & & \ddots & & & & & \\ & & & & & 0 & \varphi_n & & & \\ & & & & & -\varphi_n & 0 & & & \\ & & & & & & & 0 & & \\ & & & & & & & & 0 & \\ & & & & & & & & & \ddots \\ & & & & & & & & & & 0 \end{bmatrix} Q \tag{11}$$

where Q is some orthonormal matrix, not necessarily unique and the φ's are real constants. Since G in (10) is 3×3 skew-symmetric, we can see from the above the G must be singular, and that there exist a 3×3 orthonormal matrix Q and a real number φ such that

$$G = Q^T \begin{bmatrix} 0 & \varphi & \\ -\varphi & 0 & \\ & & 0 \end{bmatrix} Q. \tag{12}$$

Equation (12) will play an important role in the analysis hereafter.

Let $P = i\,E$ where $i = \sqrt{-1}$; then from (10), we have

$$P = i \cdot E = i \cdot G \cdot R = H \cdot R \tag{13}$$

where

$$H \triangleq i \cdot G = \begin{bmatrix} 0 & i \cdot \Delta z & -i \cdot \Delta y \\ -i \cdot \Delta z & 0 & i \cdot \Delta x \\ i \cdot \Delta y & -i \cdot \Delta x & 0 \end{bmatrix}.$$

Note that H is Hermitian. Therefore, (13) gives the polar decomposition [28] of P. Since the polar decomposition of any nonsingular matrix with distinct singular values is always unique, we can see that G and R would be unique if P should satisfy the conditions that it was nonsingular and the P^*P did not have multiple eigenvalues. (* denotes conjugate transpose.) However, we have seen that G is always singular, which implies that P is always singular. Furthermore, P always contains multiple singular values since

$$P^* \cdot P = R^* \cdot H^* \cdot H \cdot R \qquad (\text{* denotes conjugate transpose})$$
$$= R^* \cdot H^2 \cdot R = R^* \cdot (iG)\,(iG) \cdot R = -R^* \cdot G^2 \cdot R$$
$$= -R^* \left\{ Q^T \begin{bmatrix} 0 & \varphi & \\ -\varphi & 0 & \\ & & 0 \end{bmatrix} Q \right\} \cdot \left\{ Q^T \begin{bmatrix} 0 & \varphi & \\ -\varphi & 0 & \\ & & 0 \end{bmatrix} Q \right\} R$$
$$= -R^* \cdot Q \cdot \begin{bmatrix} -\varphi^2 & & \\ & -\varphi^2 & \\ & & 0 \end{bmatrix} \cdot Q \cdot R$$
$$= R \cdot Q^T \cdot \begin{bmatrix} \varphi^2 & & \\ & \varphi^2 & \\ & & 0 \end{bmatrix} \cdot Q \cdot R \tag{14}$$

and thus the eigenvalues of P^*P (or the square of the singular values of P) are φ^2, φ^2, and 0. However, we shall show in Theorem I that because of the special structure of G, once E is given, G and R are unique.

III-1A. Uniqueness and Estimation of Motion Parameters Given E: Theorem I

Theorem I: Let the SVD of E be given by

$$E = U \wedge V^T \tag{15}$$

Then, given E to within a scale factor, there are two solutions for the rotation matrix:

$$R = U \begin{bmatrix} 0 & -1 & \\ 1 & 0 & \\ & & s \end{bmatrix} V^T \tag{16}$$

or

$$= U \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & s \end{bmatrix} V^T \tag{17}$$

where $s = \det(U) \cdot \det(V)$ and one solution for the translation vector (up to a scale factor)

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \alpha \begin{bmatrix} \phi_1^T \phi_2 / \phi_2^T \phi_3 \\ \phi_1^T \phi_2 / \phi_1^T \phi_3 \\ 1 \end{bmatrix}$$

where ϕ^i is the ith row of E, $i = 1, 2, 3$, and α is some scale factor. Only one of the two solutions of R, together with the appropriate sign for α, will yield positive z and z'. Since the object must be in front of the camera, the solution is unique.

Proof: Let us first verify the uniqueness of

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}$$

given E, and give the computational formula for it.

From (9), we have

$$EE^T = GRR^TG^T = GG^T = -G^2$$

$$= \begin{bmatrix} \Delta z^2 + \Delta y^2 & -\Delta x \cdot \Delta y & -\Delta x \cdot \Delta z \\ -\Delta x \cdot \Delta y & \Delta z^2 + \Delta x^2 & -\Delta y \cdot \Delta z \\ -\Delta x \cdot \Delta z & -\Delta y \cdot \Delta z & \Delta x^2 + \Delta y^2 \end{bmatrix} \tag{18}$$

or

$$\Delta z^2 + \Delta y^2 = \phi_1^T \phi_1 \tag{19}$$

$$\Delta z^2 + \Delta x^2 = \phi_2^T \phi_2 \tag{20}$$

$$\Delta x^2 + \Delta y^2 = \phi_3^T \phi_3 \tag{21}$$

$$\Delta x \cdot \Delta y = -\phi_1^T \phi_2 \tag{22}$$

$$\Delta x \cdot \Delta z = -\phi_1^T \phi_3 \tag{23}$$

$$\Delta y \cdot \Delta z = -\phi_2^T \phi_3. \tag{24}$$

Equation (19) + (20) - (21) gives

$$2 \cdot \Delta z^2 = \phi_1^T \phi_1 + \phi_2^T \phi_2 - \phi_3^T \phi_3$$

or

$$\Delta z = \pm 1/\sqrt{2} (\phi_1^T \phi_1 + \phi_2^T \phi_2 - \phi_3^T \phi_3)^{1/2} \tag{25}$$

Similarly,

$$\Delta x = \pm 1/\sqrt{2} (-\phi_1^T \phi_1 + \phi_2^T \phi_2 + \phi_3^T \phi_3)^{1/2} \tag{26}$$

$$\Delta y = \pm 1/\sqrt{2} (\phi_1^T \phi_1 - \phi_2^T \phi_2 + \phi_3^T \phi_3)^{1/2}. \tag{27}$$

Therefore, given E, Δx, Δy, and Δz are fixed except for the signs. When a particular sign for Δz is chosen, the signs for Δx and Δy are determined from (23) and (24). Thus the translation vector

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}$$

is fixed except for the sign. Since, as mentioned twice before, multiplying E or G with any scalar does not alter the equality of (6),

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}$$

is unique up to a scale factor. Alternatively, since there is a common scale factor among the translations, Δx, Δy, and Δz, we have from (23) and (24),

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \alpha \begin{bmatrix} \phi_1^T \phi_2 / \phi_2^T \phi_3 \\ \phi_1^T \phi_2 / \phi_1^T \phi_3 \\ 1 \end{bmatrix}$$

where α is a scale factor. We now proceed to prove that given E, there are two solutions given in (17) and (18) for the rotation matrix R with only one among the two, together with the appropriate sign for α, yielding positive z and z'.

From (9), (12) and (15), we have

$$E = U \Lambda V^T = GR = Q^T \begin{bmatrix} 0 & \varphi & \\ -\varphi & 0 & \\ & & 0 \end{bmatrix} QR. \tag{28}$$

It is easy to see that Q^T is one of the singular vector matrices of E since $EE^T = GRR^TG^T + GG^T$

$$= Q^T \begin{bmatrix} 0 & \varphi & \\ -\varphi & 0 & \\ & & 0 \end{bmatrix} \begin{bmatrix} 0 & -\varphi & \\ \varphi & 0 & \\ & & 0 \end{bmatrix} Q = Q^T \begin{bmatrix} \varphi^2 & & \\ & \varphi^2 & \\ & & 0 \end{bmatrix} Q. \tag{29}$$

Thus (28) becomes

$$E = U \Lambda V^T = U \begin{bmatrix} 0 & \varphi & \\ -\varphi & 0 & \\ & & 0 \end{bmatrix} U^T R. \tag{30}$$

Premultiplying (30) by U^T gives

$$\begin{bmatrix} |\varphi| & & \\ & |\varphi| & \\ & & 0 \end{bmatrix} V^T = \begin{bmatrix} 0 & \varphi & \\ -\varphi & 0 & \\ & & 0 \end{bmatrix} U^T R$$

or by cancelling φ

$$\begin{bmatrix} \pm 1 & & \\ & \pm 1 & \\ & & 0 \end{bmatrix} V^T = \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} U^T R. \tag{31}$$

Premultiplying (31) by

$$\begin{bmatrix} 0 & -1 & \\ 1 & 0 & \\ & & 0 \end{bmatrix}$$

gives

$$\begin{bmatrix} 0 & \mp 1 & \\ \pm 1 & 0 & \\ & & 0 \end{bmatrix} V^T = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} U^T R. \tag{32}$$

Equation (32) indicates that the first and second rows of the orthonormal matrix $U^T R$ are the same as the second and first rows of V, respectively, except for a possible sign change. This implies that the third row of $U^T R$ must be equal to that of V up a possible sign difference. Therefore,

$$\begin{bmatrix} 0 & \mp 1 & \\ \pm 1 & 0 & \\ & & s \end{bmatrix} V^T = U^T R \tag{33}$$

where

$s = \pm 1.$

By taking the determinants of both sides of (33), we have

$s \det(V) = \det(U) \cdot 1.$

Thus,

$s = \det(U) \det(V).$

Equation (33) gives

$$R = U \begin{bmatrix} 0 & \mp 1 & \\ \pm 1 & 0 & \\ & & s \end{bmatrix} V^T. \tag{34}$$

One way to show that although U and V are not unique, (34) gives only two possible solutions for R, is described in the following.

Let U_1 and V_1 be some singular vector matrices for E. Let

$$R_1 = U_1 \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & s \end{bmatrix} V_1^T$$

$$R_2 = U_1 \begin{bmatrix} 0 & -1 & \\ 1 & 0 & \\ & & s \end{bmatrix} V_1^T. \tag{35}$$

Suppose U_2 and V_2 are any other singular vector matrices of E and let

$$R_3 = U_2 \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & s \end{bmatrix} V_2^T$$

and

$$R_4 = U_2 \begin{bmatrix} 0 & -1 & \\ 1 & 0 & \\ & & s \end{bmatrix} V_2^T.$$

Observe that (30) must be maintained whether U is equal to U_1 or U_2, irrespective of R. Were R_3 and R_4 be different from R_1 and R_2, then by replacing R by R_3 (or R_4) and U by U_1, controdiction would arise since (35) says R must be equal to R_1 and R_2, which are different from R_3. From this, we conclude that (34) gives the only two possible solutions for R, irrespective of which U and V are chosen.

It remains to be shown that only one among the two in (35) is valid.

Let (x_i, y_i, z_i) and (x_i', y_i', z_i'), for $i = 1, \cdots, n$, be the "true" 3-D coordinates of the n object points. Since

$$X_i' = \frac{x_i'}{z_i'} \text{ and } Y_i' = \frac{y_i'}{z_i'},$$

the image coordinates (X_i', Y_i') of the corresponding points will not be influenced by reversing the signs of (x_i', y_i', z_i') while keeping (x_i, y_i, z_i) unchanged. Also, the rigidity constraint obviously are not violated since this is just like rescaling the x, y, z axes by a common factor -1 without altering the right-hand rule. Note that for $n \geqslant 2$, it takes more than just translational motion to move object points from (x_i', y_i', z_i') to $(-x_i', -y_i', -z_i')$. Therefore, there are at least two distinct solutions for R, one for the motion $(x_i, y_i, z_i) \to (x_i', y_i', z_i')$, and other $(x_i, y_i, z_i) \to (-x_i', -y_i', -z_i')$, $i = 1, \cdots, n$. Since (35) says there are at most two possible solutions, it can be concluded that exactly one of the two solutions in (35) must correspond to the case when the object moves from the front to the back of the camera or vice versa. Finally, the requirement that $z_i \geqslant 0$, $z_i' \geqslant 0$ determines the sign of α.

III-1B. Estimation of E Given Eight Image Point Correspondences

Given eight image point correspondences $(Xi, Yi) \to (Xi', Yi')$, for $i = 1, \cdots, 8$, we have from (6),

$$\begin{bmatrix} X1'X1 & X1'Y1 & X1' & Y1'X1 & Y1'Y1 & Y1' & X1 & Y1 \\ X2'X2 & X2'Y2 & X2' & Y2'X2 & Y2'Y2 & Y2' & X2 & Y2 \\ X3'X3 & X3'Y3 & X3' & Y3'X3 & Y3'Y3 & Y3' & X3 & Y3 \\ X4'X4 & X4'Y4 & X4' & Y4'X4 & Y4'Y4 & Y4' & X4 & Y4 \\ X5'X5 & X5'Y5 & X5' & Y5'X5 & Y5'Y5 & Y5' & X5 & Y5 \\ X6'X6 & X6'Y6 & X6' & Y6'X6 & Y6'Y6 & Y6' & X6 & Y6 \\ X7'X7 & X7'Y7 & X7' & Y7'X7 & Y7'Y7 & Y7' & X7 & Y7 \\ X8'X8 & X8'Y8 & X8' & Y8'X8 & Y8'Y8 & Y8' & X8 & Y8 \end{bmatrix} \cdot \begin{bmatrix} e1 \\ e2 \\ e3 \\ e4 \\ e5 \\ e6 \\ e7 \\ e8 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix}. \tag{36}$$

Therefore, $e1, e2, \cdots, e8$ can be estimated by solving a system of linear equations expressed in (36). The conditions when the ei's are unique (or equivalently when the 8×8 matrix in (36) is nonsingular) are under investigation. In practice, given eight image point correspondences, one first substitutes the image point coordinates into the above 8×8 matrix and check its determinant. If it is nonzero, the matrix E can be determined by solving (36) for the ei's. Next the SVD of E is computed and used to calculate the actual motion parameters by the formulas given in Theorem I. Note that nonuniqueness of solutions to the ei's does not imply nonuniqueness of the motion parameters. Different E matrices may give the same set of values for the motion parameters.

IV. Restrictions on the Spatial Distribution of Object Points to Ensure Uniqueness: Lemma I and Theorem II

Multiplying (6) by z and z' gives

$$z'[X' \quad Y' \quad 1] \cdot E \cdot z \cdot \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = 0 \tag{37}$$

from (3) and (37),

$$[x' \quad y' \quad z'] \cdot E \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

or

$$[x' \quad y' \quad z'] \cdot G \cdot R \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0. \tag{38}$$

Let

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$

be transformed from

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

with some reference rotation matrix Ro and translation vector

$$To = \begin{bmatrix} \Delta xo \\ \Delta yo \\ \Delta zo \end{bmatrix},$$

i.e.,

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = Ro \begin{bmatrix} x \\ y \\ z \end{bmatrix} + To = Ro \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta xo \\ \Delta yo \\ \Delta zo \end{bmatrix}. \tag{39}$$

Let

$$Go = \begin{bmatrix} 0 & \Delta zo & -\Delta yo \\ -\Delta zo & 0 & \Delta xo \\ \Delta yo & -\Delta xo & 0 \end{bmatrix} \text{ and } Eo = Go \cdot Ro.$$

The purpose of this section is to investigate how many image point correspondences are needed to ensure that there are no other solutions to G and R as factors of E in (9) than the reference Go and Ro that can satisfy (37) [or (38)], and to state the conditions or restrictions on the spatial distribution of the object points under observation in order to ensure unique solutions.

Substituting (39) into (38) gives

$$([x \quad y \quad z] \cdot Ro^T + To^T) \cdot E \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= [x \quad y \quad z] \cdot Ro^T \cdot E \begin{bmatrix} x \\ y \\ z \end{bmatrix} + To^T \cdot E \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$= [x \quad y \quad z \quad 1] \begin{bmatrix} & & & 0 \\ & Ro^T \cdot E & & 0 \\ & & & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$+ \left\{ [x \quad y \quad z \quad 1] \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\} [To^T \cdot E \quad 0] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$= [x \quad y \quad z \quad 1] \left[\begin{array}{c|c} & 0 \\ Ro^T \cdot E & 0 \\ & 0 \\ \hline To^T \cdot E & 0 \end{array} \right] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$= [x \quad y \quad z \quad 1] \, C \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0 \tag{40}$$

where

$$C = \left[\begin{array}{c|c} & 0 \\ Ro^T \cdot E & 0 \\ & 0 \\ \hline To^T \cdot E & 0 \end{array} \right] = \left[\begin{array}{c|c} & 0 \\ Ro^T \cdot GR & 0 \\ & 0 \\ \hline To^T \cdot GR & 0 \end{array} \right]. \tag{41}$$

Note that if C is skew-symmetric, then (40) is always satisfied regardless of what x, y, z or X, Y are, since

$$2\cdot[x\ \ y\ \ z\ \ 1]\cdot C\cdot\begin{bmatrix}x\\y\\z\\1\end{bmatrix}=[x\ \ y\ \ z\ \ 1]\cdot C\cdot\begin{bmatrix}x\\y\\z\\1\end{bmatrix}+\left\{[x\ \ y\ \ z\ \ 1]\,C\begin{bmatrix}x\\y\\z\\1\end{bmatrix}\right\}^T$$

$$=[x\ \ y\ \ z\ \ 1]\cdot C\cdot\begin{bmatrix}x\\y\\z\\1\end{bmatrix}+[x\ \ y\ \ z\ \ 1]\cdot C^T\cdot\begin{bmatrix}x\\y\\z\\1\end{bmatrix}$$

$$=[x\ \ y\ \ z\ \ 1]\cdot C\cdot\begin{bmatrix}x\\y\\z\\1\end{bmatrix}+[x\ \ y\ \ z\ \ 1]\cdot(-C)\cdot\begin{bmatrix}x\\y\\z\\1\end{bmatrix}$$

$$=0.$$

It is to be stated in Lemma I that C is skew-symmetric if and only if $E = Eo$ (then according to Theorem I, the solution for the motion parameters is unique). The purpose of Theorem II is to prove that the matrix C in (41) has to be skew-symmetric if the object points under observation do not reside on two planes with one of the two planes containing the origin, nor do they lie on a cone containing the origin. We note that five or fewer points in space can always be traversed by two planes with one plane containing the origin, and that six or fewer points in space can always be traversed by a cone containing the origin. A minimum of seven points is needed to violate these two conditions. Therefore, it follows from Theorem II and Lemma I that seven points in general positions can ensure a unique solution for the motion parameters.

Lemma I: The necessary and sufficient conditions for C defined by (41) to be skew-symmetric is that

$$R = Ro \tag{42}$$

or

$$R = Q^T\begin{bmatrix}-1&&\\&-1&\\&&1\end{bmatrix}Q\,Ro \tag{43}$$

where Q is a 3×3 orthonormal matrix such that

$$G = Q^T\begin{bmatrix}0&1&\\-1&0&\\&&0\end{bmatrix}Q \tag{44}$$

and

$$\begin{bmatrix}\Delta x\\ \Delta y\\ \Delta z\end{bmatrix}=\alpha\begin{bmatrix}\Delta xo\\ \Delta yo\\ \Delta zo\end{bmatrix} \tag{45}$$

where α is some constant. (According to Theorem I, (43) and (45) are equivalent to $E = \alpha Eo$.) The proof of Lemma I is relegated to the Appendix.

Theorem II: If

$$[X'\ \ Y'\ \ 1]\,E\begin{bmatrix}X\\Y\\1\end{bmatrix}=0$$

is satisfied by the image point correspondences of a group of object points not lying on two planes with one plane containing the origin, nor on a cone containing the origin, then the C matrix in (41) has to be skew-symmetric.

Proof: From (40), which is the necessary condition of (6), we have

$$[x\ \ y\ \ z\ \ 1]\,C\begin{bmatrix}x\\y\\z\\1\end{bmatrix}+\left\{[x\ \ y\ \ z\ \ 1]\,C\begin{bmatrix}x\\y\\z\\1\end{bmatrix}\right\}^T=0$$

or

$$[x\ \ y\ \ z\ \ 1]\,(C+C^T)\begin{bmatrix}x\\y\\z\\1\end{bmatrix}=0. \tag{46}$$

From (41)

$$C+C^T=\left[\begin{array}{c|c}Ro^T\cdot E+E^T\cdot Ro & E^T\cdot To\\ \hline To^T\cdot G\cdot R & 0\end{array}\right]$$

$$C+C^T=\left[\begin{array}{c|c}Ro^T\cdot G\cdot R+R^T\cdot G^T\cdot Ro & R^T\cdot G^T\cdot To\\ \hline To^T\cdot G\cdot R & 0\end{array}\right].$$

Substituting (12) into the above gives

$$C + C^T = \left[\begin{array}{c|c} Ro^T \cdot Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot Q \cdot R - R^T \cdot Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot Q \cdot Ro & -R^T \cdot Q^T \cdot \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot Q \cdot To \\ \hline To^T \cdot Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot Q \cdot R & 0 \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} & R^T \cdot Q^T & & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c|c} Q \cdot R \cdot Ro^T \cdot Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} Q \cdot Ro \cdot R^T \cdot Q^T & - \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot Q \cdot To \\ \hline To^T \cdot Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} & 0 \end{array}\right]$$

$$\cdot \left[\begin{array}{ccc|c} & Q \cdot R & & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} & Q \cdot R & & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 & 1 \end{array}\right]^T \left[\begin{array}{c|c} M \cdot \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot M^T & - \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot To' \\ \hline To' \cdot \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} & 0 \end{array}\right] \left[\begin{array}{ccc|c} & Q \cdot R & & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \\ \hline 0 & 0 & 0 & 1 \end{array}\right] \tag{47}$$

where

$$M \triangleq Q \cdot R \cdot Ro^T \cdot Q^T$$

$$To' \triangleq Q \cdot To \triangleq \begin{bmatrix} t1 \\ t2 \\ t3 \end{bmatrix}.$$

Let

$$M \triangleq \begin{bmatrix} m1 & m2 & m3 \\ m4 & m5 & m6 \\ m7 & m8 & m9 \end{bmatrix}.$$

Then (47) becomes

$$C + C^T = \begin{bmatrix} QR & \\ & 1 \end{bmatrix}^T \begin{bmatrix} 2\,m4 & m5\,m1 & m6 & -t2 \\ m5 - m1 & -2\,m2 & -m3 & t1 \\ m6 & -m3 & 0 & 0 \\ -t2 & t1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} QR & \\ & 1 \end{bmatrix}. \tag{48}$$

Let the original coordinate system be rotated with $R \cdot Q$ such that

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = Q \cdot R \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix}. \tag{49}$$

Then from (46) and (48),

$$[x_c \ \ y_c \ \ z_c \ \ 1] \cdot J \cdot \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = 0 \tag{50}$$

where

$$J = \begin{bmatrix} 2\,m4 & m5 - m1 & m6 & -t2 \\ m5 - m1 & -2\,m2 & -m3 & t1 \\ m6 & -m3 & 0 & 0 \\ -t2 & t1 & 0 & 0 \end{bmatrix}. \tag{51}$$

Equation (50) gives

$$2[m4 \cdot x_c^2 + (m5 - m1)\,x_c \cdot y_c - m2 \cdot y_c^2 - t2 \cdot x_c + t1 \cdot y_c + (m6 \cdot x_c - m3 \cdot y_c)\,z_c] = 0 \tag{52}$$

or

$$z = [m4 \cdot x_c^2 + (m5 - m1)\,x_c \cdot y_c - m2 \cdot y_c^2 - t2 \cdot x_c + t1 \cdot y_c]/(m6 \cdot x_c - m3 \cdot y_c). \tag{53}$$

Unless J in (51) is identically zero, (52) indicates that all the points must lie on a quadric surface of some type containing the origin. It can be shown that unless J in (51) is a zero matrix, all the points must either lie on two planes with one plane containing the origin, or on a cone passing through the origin. (The detailed proof can be found in [29], [30].) But, as was defined in (48),

$$C + C^T = \left[\begin{array}{ccc|c} & & & 0 \\ & QR & & 0 \\ & & & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right]^T \cdot J \cdot \left[\begin{array}{ccc|c} & & & 0 \\ & QR & & 0 \\ & & & 0 \\ \hline 0 & 0 & 0 & 1 \end{array}\right].$$

Therefore,

$$C + C^T = 0$$

or

$$C = -C^T$$

which means that C has to be skew-symmetric. Q.E.D.

We now state five corollaries that can be readily derived from Lemma I and Theorem II. The proofs for these lemmas can be found in [29], [30].

Corollary I: Given the image correspondences of two planes not passing through the origin, the motion is unique.

Corollary II: Given the image correspondences of six points with four points on one plane not containing the origin, four points on the other plane also not containing the origin, and two points common to the above two groups of four points on the intersection of the two planes can ensure unique solutions for the motion parameters.

Corollary III: The image correspondences of four points on a plane not passing through the origin and two other points not on this plane determine the motion parameters uniquely.

Corollary IV: Given the image correspondences of seven or more points not traversable by two planes with one plane containing the origin, nor by a cone containing the origin, the motion parameters are unique.

Note that Corollary IV only gives a sufficient condition for uniqueness. Even if the seven points are traversable by two planes with one plane passing through the origin, or by a cone containing the origin, the motion parameters might still be unique in some situations. For example, if six among the seven points satisfy the condition stated in Corollary III, then the motion parameters are unique even if there may be two planes passing through these seven points with one plane containing the origin.

From (52), the criteria for whether there exists a cone containing the origin that passes through n points is whether the following $n \times 7$ rectangular matrix has full column rank or not:

$$\begin{bmatrix} x1^2 & x1 \cdot y1 & y1^2 & x1 & y1 & z1 \cdot x1 & z1 \cdot y1 \\ x2^2 & x2 \cdot y2 & y2^2 & x2 & y2 & z2 \cdot x2 & x2 \cdot y2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ xn^2 & xn \cdot yn & yn^2 & xn & yn & zn \cdot xn & zn \cdot yn \end{bmatrix}.$$

However, since only the image coordinates are given, the only useful criterion available is whether or not the 8×8 matrix in (36) is nonsingular or not. If it is nonsingular, one can solve for the E matrix, compute its SVD, and then use the formula in Theorem I to calculate the actual motion parameters. The following corollary states a sufficient condition for the 8×8 matrix in (36) to be singular.

Corollary V: Given the image correspondences of eight points among which more than six points are coplanar, the 8×8 coefficient matrix in (36) is singular.

V. Simulation Results

In the following computer simulation tests, the image coordinates at t_1 of n object points with randomly chosen object space coordinates (x_i, y_i, z_i), $i = 1, \cdots, n$, in 3-D space are obtained using (3). Next, the object points are rotated with some reference motion parameters, and (1) is used to compute (x_i', y_i', z_i'), $i = 1, \cdots, n$. Then the image coordinates of these n points at t_2, i.e., (X_i', Y_i'), $i = 1, 2, \cdots, n$, were computed using (4). These n simulated image point correspondences $(X_i, Y_i) \rightarrow (X_i', Y_i')$, $i = 1, \cdots, n$, are first perturbed to simulate the error in finding the point correspondences, and then substituted into the linear matrix equations (36), which is solved by least-squares for the eight parameters $e_1, \cdots, e_8$. The motion parameters are computed using the method described in Theorem I. The results are then compared with the reference motion parameters.

Test 1:

Perturbation of (X', Y'): 0 percent.

Number of Points: 8.

Reference Motion Parameters: ($n1 = \cos\alpha$, $n2 = \cos\beta$, $n3 = \cos\gamma$).

$$\theta_o = 78^\circ, \quad \alpha_o = 52^\circ, \quad \beta_o = 75^\circ, \quad \gamma_o = 42^\circ$$

$$\Delta x_o/\Delta z_o = 23, \quad \Delta y_o/\Delta z_o = -10.$$

Ideal E Matrix:

$$E_o = \begin{bmatrix} 0.467 & 1.868 & 1.439 \\ 0.483 & 4.297 & 3.411 \\ -5.916 & 0.004 & 1 \end{bmatrix}.$$

Estimated E Matrix:

$$E = E_o.$$

Two solutions for the motion and geometrical parameters:

⟨*Solution 1*⟩

Motion Parameters: Same as reference parameters.

Estimated z and z′: (Up to a scale factor)

$$z_1 = 5.00, \quad z_1' = 14.41; \quad z_2 = 32.30, \quad z_2' = 36.30;$$

$$z_3 = 7.00, \quad z_3' = 15.81; \quad z_4 = 18.00, \quad z_4' = 34.88;$$

$$z_5 = 22.20, \quad z_5' = 40.19; \quad z_6 = 4.00, \quad z_6' = 12.56;$$

$$z_7 = 6.50, \quad z_7' = 18.27; \quad z_8 = 35.80, \quad z_8' = 31.60.$$

⟨*Solution 2*⟩

Translation Parameters: Same as Solution I according to Theorem I.

Rotation Parameters:

$$\theta = 35.97°, \quad \alpha = 123.07°, \quad \beta = 40.49°, \quad \gamma = 110.60°.$$

Estimated z and z′:

$$z_1 = -9.01, \quad z_1' = 29.56; \quad z_2 = -12.80, \quad z_2' = 14.38;$$

$$z_3 = -8.35, \quad z_3' = 18.87; \quad z_4 = -14.87, \quad z_4' = 28.81;$$

$$z_5 = -23.33, \quad z_5' = 42.23; \quad z_6 = -4.36, \quad z_6' = 13.69;$$

$$z_7 = -9.36, \quad z_7' = 26.31; \quad z_8 = -12.75, \quad z_8' = 11.25.$$

Although there are two solutions for the rotation matrix, the z_i and z_i' in Solution 2 have opposite signs, which implies that the object either moves from the back to the front, or from the front to the back of the camera. Therefore, only Solution 1 is valid. For the rest of this section, we only present one of the two solutions that yield z and z' with the same sign.

Test 2:

Perturbation of (X', Y')*:* 0.1 percent.

Number of Points: 8.

Reference Motion Parameters: Same as Test 1.

Ideal E Matrix: Same as Test 1.

Estimated E Matrix:

$$E = \begin{bmatrix} 0.429 & 1.782 & 1.289 \\ 0.458 & 4.003 & 3.165 \\ -5.504 & -0.051 & 1 \end{bmatrix}.$$

Estimated Motion Parameters:

$$\theta = 77.52°, \quad \alpha = 51.29°, \quad \beta = 75.12°, \quad \gamma = 42.54°$$

$$\Delta x/\Delta z = 23.77, \quad \Delta y/\Delta z = -10.38.$$

From a series of simulation tests with different amounts of perturbation for (X', Y'), we obtain Table I.

TABLE I
ERROR VERSUS PERTURBATION OF (X', Y') USING EIGHT POINTS

Perturbation of (X', Y')	0	0.1%	0.5%	1%	2%
Error of E	0	7.30%	29.36%	47.13%	67.54%
Error of R	0	1.36%	6.96%	14.32%	30.28%
Error of $\Delta x/\Delta z$ and $\Delta y/\Delta z$	0	0.51%	20.66%	53.97%	94.63%

TABLE II
ERROR VERSUS NUMBER OF POINTS FOR 1 PERCENT PERTURBATION OF (X', Y')

Number of Points	8	9	20
Error of E	47.13%	18.74%	2.32%
Error of R	14.32%	3.69%	0.83%
Error of $\Delta x/\Delta z$ and $\Delta y/\Delta z$	53.97%	3.52%	10.09%

TABLE III
ERROR VERSUS NUMBER OF POINTS FOR 3 PERCENT PERTURBATION OF (X', Y')

Number of Points	8	20
Error of E	78.85%	26.73%
Error of R	47.10%	3.28%
Error of $\Delta x/\Delta z$ and $\Delta y/\Delta z$	101.8%	93.46%

Table I shows that even for perturbation of (X', Y') as small as 1 percent, the error is still significant. Table II shows how drastically the error can be reduced by using more points for the case of 1 percent perturbation on (X', Y').

Note that the error of translations using 20 points is even greater than that using 9 points. This indicates the difficulty of obtaining a very accurate estimation of translations unless the resolution of the camera is good enough and the frame rate is high enough so that the image points can be tracked effectively over several frames to obtain point correspondences with error less than 1 percent.

Although the error of translation using 20 points is greater than that using 9 points, it is much smaller than the error using 8 points. However, this will not happen unless the perturbation of (X', Y') is less than 3 percent. When the perturbation of (X', Y') is greater than 3 percent, using more than 8 points will still reduce the errors in R significantly, but not the errors in the translations. Table III reveals this phenomenon.

When the motion is small, (X_i', Y_i') is close to (X_i, Y_i), and therefore, the columns of the coefficient matrix in (36) become nearly dependent, causing the condition number of the matrix and the error for the essential parameters to be very large. This can be seen in Table IV.

The reference motion parameters for motion A and B are the following.

For motion A,

$$\theta_o = 78°, \quad \alpha = 52°, \quad \beta = 75°, \quad \gamma = 42°, \quad \Delta x_o/\Delta z_o = 23,$$

$$\Delta y_o/\Delta z_o = -10.$$

TABLE IV
ERROR VERSUS AMOUNT OF MOTION FOR 2 PERCENT PERTURBATION OF (X', Y') USING 20 POINTS

	Motion A (large)	Motion B (small)
Error of E	12.60%	97.94%
Error of R	1.79%	6.12%
Error of $\Delta x/\Delta z$ and $\Delta y/\Delta z$	22.84%	42.94%
$\epsilon \cdot \text{cond}(H)$	7.47	67.16

For motion B,

$$\theta_o = 1°, \quad \alpha = 52°, \quad \beta = 75°, \quad \gamma = 42°, \quad \Delta x_o/\Delta z_o = -0.1,$$
$$\Delta y_o/\Delta z_o = -0.1.$$

A detailed error analysis can be found in [30].

VI. CONCLUDING REMARKS

Several theorems and corollaries have been presented regarding the uniqueness and estimation of 3-D motion parameters of a rigid body. In summary, the following results have been established.

1) The fact that we can define eight parameters $e1, e2, \cdots, e8$ that contain all essential information one can possibly obtain given any number of image point correspondences.

2) The fact that given the E matrix consisting of the eight essential parameters, the actual motion parameters are unique, and can be computed by taking the singular value decomposition (SVD) of the 3×3 E matrix.

3) A method of determining the E matrix given eight image correspondences. This requires the solution of a set of linear equations only.

4) An operational criterion for the uniqueness of motion parameters: If the determinant of a certain 8×8 matrix containing only the image coordinates of eight point correspondences does not vanish, the uniqueness is assured.

5) A sufficient condition for the uniqueness of the motion parameters: Given seven or more image point correspondences, the motion parameters are uniquely determined if the seven object points do not lie on two planes with one plane passing through the origin or on a cone containing the origin.

The results in this paper should be of interest to numerous areas of research including image sequence analysis, tracking, image coding, stereo imaging, photogrammetry, and robotic vision. These results can, for example, be applied to the stereo imaging problems in photogrammetry and computer vision without assuming the relative orientation of the two cameras since pictures taken at two time instances can be regarded as taken by two cameras at one instance. After the motion parameters are computed using the formulas in Theorem I, the surface structure of the object can be determined up to a common scale factor by computing the z coordinates using (5a) or (5b).

In concluding the paper, we would like to mention that after we had submitted our original manuscript, it was brought to our attention by Prof. H. H. Nagel, Hamburg University, that results similar to some of ours were obtained independently by Longuet-Higgins [31]. In his elegant and concise paper, Prof. Longuet-Higgins derived the E matrix equation (36) and an algorithm for determining R and T from E using tensor and vector analysis. However, he did not consider the uniqueness question for the motion parameters, which we did in Section IV.

APPENDIX
PROOF OF LEMMA I

If C is skew-symmetric, then it is necessary from (41) that

$$Ro^T \cdot G \cdot R = -(Ro^T GR)^T \tag{54}$$

and

$$To^T \cdot G \cdot R = [0 \;\; 0 \;\; 0]$$

or

$$R^T \cdot G^T \cdot To = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \tag{55}$$

(54) gives

$$Ro^T \cdot G \cdot R = -R^T \cdot G^T \cdot Ro = R^T \cdot G \cdot Ro.$$

Substituting (44) into the above gives

$$Ro^T Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} QR = R^T Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} QRo. \tag{56}$$

Premultiplying (56) by QR and postmultiplying by $R^T \cdot Q^T$ give

$$QRRo^T Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} QRoR^T Q^T$$

or

$$L \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot L^T \tag{57}$$

where

$$L \triangleq QRRo^T Q^T \triangleq \begin{bmatrix} j1 & j2 & j3 \\ j4 & j5 & j6 \\ j7 & j8 & j9 \end{bmatrix}. \tag{58}$$

From (57) and (58)

$$\begin{bmatrix} -j2 & j1 & 0 \\ -j5 & j4 & 0 \\ -j8 & j7 & 0 \end{bmatrix} = \begin{bmatrix} j2 & j5 & j8 \\ -j1 & -j4 & -j7 \\ 0 & 0 & 0 \end{bmatrix}.$$

Thus

$$j7 = j8 = 0$$

and

$$j2 = -j2 \quad \text{or } j2 = 0$$
$$j4 = -j4 \quad \text{or } j4 = 0$$
$$j1 = j5.$$

Then L becomes

$$L = \begin{bmatrix} j1 & 0 & j3 \\ 0 & j1 & j6 \\ 0 & 0 & j9 \end{bmatrix}. \tag{59}$$

Equation (58) implies that L is orthonormal of the 1st kind since R, Ro, Qo are orthonormal and that $\det(L) = \det(Q)\det(R)\det(Ro)\det(Q) = (\det(Q))^2 = (\pm 1)^2 = 1$. Taking the inner product of the first and third rows of L in (59), and equating it to zero gives $j3 \cdot j9 = 0$. Since $j9 \neq 0$ (otherwise the third row of L would be zero), $j3 = 0$. Similarly, $j6 = 0$. With these and the fact that $\det(L) = 1$, we conclude that L can assume only the following forms:

$$L = \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \text{ or } L = \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix}.$$

From (58),

$$R = Q^T \cdot \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \cdot Q \cdot Ro = Ro$$

or

$$R = Q^T \cdot \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix} \cdot Q \cdot Ro.$$

Thus (42) and (43) are the necessary conditions for C to be skew-symmetric. The next thing is to verify (45).

Premultiplying (55) by R gives

$$G^T \cdot To = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

or

$$\begin{bmatrix} 0 & \Delta z & -\Delta y \\ -\Delta z & 0 & \Delta x \\ \Delta y & -\Delta x & 0 \end{bmatrix} \begin{bmatrix} \Delta xo \\ \Delta yo \\ \Delta zo \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

which gives

$$\Delta z \cdot \Delta yo - \Delta y \cdot \Delta zo = 0$$
$$-\Delta z \cdot \Delta xo + \Delta x \cdot \Delta zo = 0$$
$$\Delta y \cdot \Delta xo - \Delta x \cdot \Delta yo = 0.$$

Let $\alpha = \Delta z / \Delta zo$, then $\Delta y = \alpha \cdot \Delta yo$, $\Delta x = \alpha \cdot \Delta xo$. Hence,

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \Delta xo \\ \Delta yo \\ \Delta zo \end{bmatrix}$$

which is the same as (45). The E matrix then is equal to Eo (i.e., unique) up to a scale factor since

$$E = G \cdot R = \begin{bmatrix} 0 & \Delta z & -\Delta y \\ -\Delta z & 0 & \Delta x \\ \Delta y & -\Delta x & 0 \end{bmatrix} R = \alpha \cdot Go \cdot Ro = \alpha \cdot Eo$$

if (42) is used, or

$$= Go\, Q^T \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix} Q\, Ro = \alpha\, Q^T \begin{bmatrix} 0 & 1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} Q \cdot Q^T$$
$$\cdot \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix} Q\, Ro$$
$$= \alpha\, Q^T \begin{bmatrix} 0 & -1 & \\ 1 & 0 & \\ & & 0 \end{bmatrix} Q\, Ro = -\alpha\, Eo$$

if (43) is used.

Sufficiency Part: From the structure of C in (41), it is obvious that in order for C to be skew-symmetric, the row vector $To^T \cdot G \cdot R$ on the fourth row has to be equal to the negative of the transpose of the fourth column, which is a zero vector, and that the 3×3 matrix $Ro^T \cdot G \cdot R$ on the upper-left corner of C must be itself skew-symmetric. With (45), $To^T \cdot G \cdot R$ in (41) becomes

$$To^T \cdot G \cdot R = [\Delta xo \quad \Delta yo \quad \Delta zo] \begin{bmatrix} 0 & \Delta zo & -\Delta yo \\ -\Delta zo & 0 & \Delta xo \\ \Delta yo & -\Delta xo & 0 \end{bmatrix} R$$
$$= [-\Delta yo \cdot \Delta zo + \Delta zo \cdot \Delta yo$$
$$\Delta xo \cdot \Delta zo - \Delta zo \cdot \Delta xo$$
$$-\Delta xo \cdot \Delta yo + \Delta yo \cdot \Delta xo]\, R$$
$$= [0 \quad 0 \quad 0]\, R = [0 \quad 0 \quad 0]. \tag{60}$$

We now proceed to show that with R either given by (42) or by (43), the 3×3 submatrix $To^T \cdot G \cdot R$ in C has to be skew-symmetric.

With (42), $Ro^T\, G\, R$ in (41) becomes

$$Ro^T \cdot G \cdot R = Ro^T \cdot G \cdot Ro = Ro \cdot (-G^T) \cdot Ro$$
$$= -(Ro^T \quad G \quad Ro)^T. \tag{61a}$$

On the other hand, with (43), $Ro^T \cdot G \cdot R$ in (41) becomes

$$Ro^T \cdot G \cdot R = Ro^T \cdot G \cdot Q \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix} \cdot Q \cdot Ro$$

$$= Ro^T \cdot Q \cdot \begin{bmatrix} 0 & -1 & \\ -1 & 0 & \\ & & 0 \end{bmatrix} \cdot Q$$

$$\cdot Q \cdot \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix} \cdot Q \cdot Ro$$

$$= Ro^T \cdot Q^T \begin{bmatrix} 0 & -1 & \\ 1 & 0 & \\ & & 0 \end{bmatrix} Q\,Ro = -Ro^T \cdot G \cdot Ro.$$

Thus

$$(Ro^T \cdot G \cdot R)^T = (-Ro^T \cdot G \cdot Ro)^T = Ro^T \cdot G \cdot Ro$$
$$= -Ro^T \cdot G \cdot R. \qquad (61b)$$

Equation (61) shows that with either (42) or (43), $Ro^T \cdot G \cdot R$ is skew-symmetric. This fact, together with (60), imply that C in (41) is skew-symmetric. Q.E.D.

Acknowledgment

The authors gratefully acknowledge the encouragement and support of Dr. M. Blostein, Director, INRS.

References

[1] A. Netravali and J. Robbins, "Motion compensation TV coding: Part 1," *Bell Syst. Tech. J.*, vol. 58, pp. 631-670, 1979.

[2] R. Y. Tsai and T. S. Huang, "Moving image restoration and registration," in *Proc. IEEE Int. Conf. ASSP*, Denver, CO, Apr. 9-11, 1980.

[3] F. Rocca, "TV bandwidth compression utilizing frame-to frame correlation and movement compensation," in *Picture Bandwidth Compression*, T. S. Huang and O. J. Tretiak, Eds. London: Gordon and Breach, 1972.

[4] R. Y. Tsai and T. S. Huang, "Moving image restoration," presented at the IEEE Comput. Soc. Workshop on Comput. Anal. of Time Varying Imagery, Philadelphia, PA, Apr. 5-6, 1976.

[5] J. Limb and J. Murphy, "Estimating the velocity of moving images in TV signals," *Comput. Graphics Image Processing*, vol. 4, pp. 311-327, 1975.

[6] B.K.P. Horn and B. G. Schunck, "Determining optical flow," AI Memo. 572, MIT, Apr. 1980.

[7] C. Cafforio and F. Rocca, "Tracking moving objects in TV images," *Signal Processing*, vol. 1, pp. 133-140, 1979.

[8] R. J. Schalkoff, "Algorithms for a real-time automatic video tracking system," Ph.D. dissertation, Dep. Elec. Eng., Univ. Virginia, Charlottesville, 1979.

[9] T. S. Huang and R. Y. Tsai, "Image sequence analysis: Motion estimation," *Image Sequence Analysis*, T. S. Huang, Ed. Springer-Verlag, 1981, ch. 1.

[10] S. Ullman, *The Interpretation of Visual Motion.* Cambridge, MA: MIT Press, 1979.

[11] J. W. Roach and J. K. Aggarwal, "Determining the movement of objects from a sequence of images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, Nov. 1979.

[12] H. H. Nagel, "On the derivation of 3-D rigid point configurations from image sequences," in *Proc. IEEE Conf. Pattern Recognition and Image Processing*, Dallas, TX, Aug. 3-5, 1981, pp. 103-108.

[13] K. Praxdny, "Determining the instantaneous direction of motion from optical flow generated by a curvilinearly moving observer," in *Proc. IEEE Conf. Pattern Recognition and Image Processing*, Dallas, TX, Aug. 3-5, 1981, pp. 109-114.

[14] K. Praxdny, "Egomotion and relative depth map from optical flow," *Biol. Cybern.*, vol. 30, pp. 87-102, 1980.

[15] J. A. Webb and J. K. Aggarwal, "Structure from motion of rigid and jointed bodies," in *Proc. 7th Int. Joint Conf. Artificial Intell.*, Aug. 1981.

[16] D. N. Lee, "The optical flow field: The foundation of vision," *Phil. Trans. Royal Soc.*, London, vol. B, no. 290, pp. 169-179, 1980.

[17] W. F. Clocksin, "Perception of surface slant and edge labels from optical flow: A computational approach," *Perception*, vol. 9, no. 3, pp. 253-269, 1980.

[18] H. H. Nagel and B. Neumann, "On 3-D reconstruction from two perspective views," in *Proc. IJCAI 81*, vol. II, Aug. 1981.

[19] A. Zvi Meiri, "On monocular perception of 3-D moving objects," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, Nov. 1980.

[20] G. W. Stewart, *Introduction to Matrix Computations.* New York: Academic, 1973.

[21] G. H. Golub and C. Reinsch, "Singular value decomposition and least square solutions," *Math.*, vol. 14, pp. 403-420, 1970.

[22] J. Wilkinson, "Global convergence of tridiagonal QR algorithm with origin shifts," *Linear Algebra Appl.*, vol. 1, pp. 409-420, 1980.

[23] R. Y. Tsai and T. S. Huang, "Estimating 3-D motion parameters of a rigid planar patch," in *1981 Proc. IEEE Conf. Pattern Recognition and Image Processing*, Dallas, TX, Aug. 3-5, 1981.

[24] —, "Estimating three-dimensional motion parameters of a rigid planar patch, I," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 1147-1152, Dec. 1981.

[25] R. Y. Tsai, T. S. Huang, and W. L. Zhu, "Estimating three-dimensional motion parameters of a rigid planar patch, II: Singular value decomposition," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 525-534, Aug. 1982.

[26] R. Y. Tsai and T. S. Huang, "Estimating three-dimensional motion parameters of a rigid planar patch, III: Finite point correspondences and three-view problem," in *Proc. IEEE Int. Conf. ASSP*, Paris, May 3-5, 1982; see also *IEEE Trans. Acoust., Speech, Signal Processing*, to be published.

[27] D. F. Rogers and J. A. Adams, *Mathematical Elements for Computer Graphics.* New York: McGraw-Hill, 1976.

[28] F. R. Gantmacher, *The Theory of Matrices*, vol. 1. Chelsea, 1959.

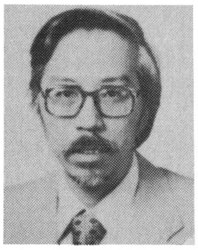
[29] R. Y. Tsai and T. S. Huang, "Uniqueness and estimation of 3-D motion parameters of rigid objects with curved surfaces," Univ. Illinois, Tech. Rep. R-921, UIUC-ENG-81-2252, Aug. 1981.

[30] R. Y. Tsai and T. S. Huang, *Three-Dimensional Motion and Structure from Image Sequences.* Springer-Verlag, 1983.

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

Roger Y. Tsai (M'82) was born in Taiwan, Republic of China, on May 10, 1956. He received the M.S. degree from Purdue University, West Lafayette, IN, and the Ph.D. degree from the University of Illinois, Urbana-Champaign, both in electrical engineering, in 1980 and 1981, respectively.

He was employed by Bell Northern Research/INRS Telecommunications, Montreal, Canada, for three months during the Summer of 1979 as a Visiting Scientist, working on moving image registration and enhancement. During the Summer of 1980, he was employed by the Signal Processing Group, EPFL, Lausanne, Switzerland, for three months working on 3-D time-varying scene analysis. In the Summer of 1981, he again visited BNR/INRS, Montreal, for three months, working on image sequence analysis and computer vision. He has been with IBM T. J. Watson Research Center, Yorktown Heights, NY, since the Spring of 1982. His major research interests include signal and image processing, robotics vision inspection, part acquisition and automated assembly, 3-D time-varying scene analysis, efficient algorithms, and parallel architecture for robotics vision problems. He has published a few dozen technical papers in these areas.



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

From 1963 to 1973, he was on the faculty of the Department of Electrical Engineering, MIT. During the academic year 1971–1972, he was on sabbatical leave visiting ETH-Zürich, Switzerland, as a Guggenheim Fellow. During the academic year 1972–1973, he was again on leave, working in the Optics Group at the MIT Lincoln Laboratory. From 1973–1980 he was Professor of Electrical Engineering and Director of the Laboratory for Information and Signal Processing at Purdue University. During the Summers of 1974, 1979, and 1981 he was a visiting Professor at INRS Telecommunications, University of Quebec. During the academic year 1976–1977, he was a recipient of the Humboldt Foundation U.S. Senior Scientist Award, working on the archaeological applications of image processing at the Rheinisches Landesmuseum, Bonn, West Germany. During the summer semester of 1978 he was a visiting Professor at the Technical University of Hannover, Germany. During the summer semester of 1980 he was a visiting Professor at the Swiss Institute of Technology, Lausanne. In August 1980, he joined the University of Illinois, Urbana-Champaign, where he is currently Professor of Electrical Engineering and Research Professor of the Coordinated Science Laboratory.

His professional interest lies in the broad area of information and communication technology but especially the transmission and processing of multidimensional signals. He has served as a consultant to numerous industrial firms and government agencies in the U.S. and abroad. He is author (with R. R. Parker) of the book *Network Theory: An Introductory Course* (Addison-Wesley, 1972); and Editor (with O. J. Tretiak) of the book *Picture Bandwidth Compression* (Gordon and Breach, 1972), and the books *Picture Processing and Digital Filtering* (1975, 2nd ed. 1979), *Two-Dimensional Digital Signal Processing I: Linear Filters* (1981), *Two-Dimensional Digital Signal Processing II: Transforms and Median Filters* (1981), and *Image Sequence Analysis* (1981), all published by Springer-Verlag.

Dr. Huang is an Editor of the international journal *Computer Graphics and Image Processing*, an Associate Editor of *Pattern Recognition*, an Associate Editor for signal processing of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, and on the overseas Editorial Board of *Signal Processing*, the official journal of the European Association for Signal Processing.

Curvature and Tangential Deflection of Discrete Arcs: A Theory Based on the Commutator of Scatter Matrix Pairs and Its Application to Vertex Detection in Planar Shape Data

IAN M. ANDERSON AND JAMES C. BEZDEK

***Abstract*—This paper introduces a new theory for the tangential deflection and curvature of plane discrete curves. Our theory applies to discrete data in either rectangular boundary coordinate or chain coded formats: its rationale is drawn from the statistical and geometric properties associated with the eigenvalue-eigenvector structure of sample covariance matrices. Specifically, we prove that the nonzero entry of the commutator of a piar of scatter matrices constructed from discrete arcs is related to the angle between their eigenspaces. And further, we show that this entry is—in certain limiting cases—also proportional to the analytical curvature of the plane curve from which the discrete data are drawn. These results lend a sound theoretical basis to the notions of discrete curvature and tangential deflection; and moreover, they provide a means for computationally efficient implementation of algorithms which use these ideas in various image processing contexts. As a concrete example, we develop the commutator vertex detection (CVD) algorithm, which identifies the location of vertices in shape data based on excessive cummulative tangential deflection; and we compare its performance to several well established corner detectors that utilize the alternative strategy of finding (approximate) curvature extrema.**

***Index Terms*—Chain codes, commutator, corner detection, covariance matrix, discrete curvature, polygonal approximation, scatter matrices, shape analysis, tangential deflection, vertex detection.**

Manuscript received June 3, 1982; revised August 30, 1982 and May 14, 1983. This work was supported in part by the National Science Foundation under Grants MCS80-02328 and MCS80-02856.

I. M. Anderson is with the Department of Mathematics, Utah State University, Logan, UT 84322.

J. C. Bezdek was with the Department of Mathematics, Utah State University, Logan, UT 84322. He is now with the Department of Computer Science, University of South Carolina, Columbia, SC 29208.

I. INTRODUCTION

MANY image processing tasks involve the shape analysis, matching, or clustering of data sets which represent the boundary of arbitrary planar figures. Data sets of this kind typically assume one of two formats: rectangular boundary coordinates (RBC's), each data point being the Cartesian coordinates of a point belonging to the boundary curve (and the data set itself ordered in a manner consistent with the orientation of the boundary curve); or chain coded (CC), wherein each data item is defined by the relative position of one boundary pixel to the next. Rectangular boundary coordinates arise naturally in various image processing contexts: they are generated whenever a continuous (analog) representation—for