Unbiased Estimation and Statistical Analysis of 3-D Rigid Motion from Two Views

Kenichi Kanatani

Abstract—**The problem of estimating 3-D rigid motion from point correspondences over two views is formulated as nonlinear least-squares optimization, and the statistical behaviors of the errors in the solution are analyzed by introducing a realistic model of noise described in terms of the covariance matrices of "N-vectors." It is shown that the least-squares solution based on the epipolar constraint is statistically biased. The geometry of this bias is described in both quantitative and qualitative terms. Finally, an unbiased estimation scheme is presented, and random number simulations are conducted to observe its effectiveness.**

Index Terms—**Error analysis, estimation, model of noise, statistical bias, 3-D motion, unbiased estimation.**

I. Introduction

MATHEMATICAL analysis of 3-D rigid motion estimation known as *shape* (or *structure*) *from motion* was initiated by Ullman [32], who presented a basic mathematical framework that had a long lasting influence over the subsequent computer vision research. Roach and Aggarwal [28] applied this framework to real images and obtained the solution by numerical search. Nagel [24] presented a semi-analytical formulation, reducing the problem to solving a single nonlinear equation. A complete analytical solution for eight feature points was independently given by Longuet-Higgins [21] and Tsai and Huang [31]. The solution of Longuet-Higgins was based on elementary vector calculus, whereas the solution of Tsai and Huang involved singular value decomposition. Zhuang *et al.* [37] combined them into a simplified eight-point algorithm. Zhuang [36] also discussed the uniqueness issue. All these algorithms first compute the *essential matrix* from the so-called *epipolar equation* and then compute the *motion parameters* from it.

Since the essential matrix has five degrees of freedom, 3-D interpretations can be determined in principle from five feature points. Using a numerical technique called the *homotopy method*, Netravali *et al.* [25] showed the existence of, at most, ten solutions. Arguing from the standpoint of projective geometry, Faugeras and Maybank [7] also showed that, at most, ten solutions can be obtained from five feature points. They reduced the problem to solving an algebraic equation of degree ten and solved it by symbolic algebra software. Using the quaternion representation of 3-D rotation, Jerian and Jain [11] reduced the problem to solving the resultant of degree 16 of a pair of polynomials of degree 4 in two variables and computed the solution by symbolic algebra software. Jerian and Jain [12] also reviewed known algorithms exhaustively and compared their performances for noisy data.

However, all these algorithms are constructed on the assumption that all data are exact. Hence, they are all *fragile* in the sense that inconsistencies arise in the presence of noise (e.g., the solution becomes different, depending on which of the theoretically equivalent relationships are used). A noise robust algorithm was presented by Weng *et al.* [34]. They estimated the essential matrix from the epipolar equation by least squares and then computed the motion parameters by least squares. Spetsakis and Aloimonos [29], on the other hand, applied direct optimization to the epipolar equation without computing the essential matrix. Spetsakis and Aloimonos [30] also considered optimization over more than two frames.

In order to apply 3-D motion analysis to real systems, statistical analysis of error behaviors becomes a key issue since noise is inevitable for real systems. Even if errors are inevitable, the knowledge of how reliable each computation is is indispensable in guaranteeing achievable performance of the systems that use such computations. In addition, statistical reliability estimation becomes vital when using multiple sensors and fusing the data (*sensor fusion*) because in order to fuse multiple data, they must be weighted so that reliable date contribute more than unreliable data.

The statistical approach to image processing and computer vision problems is not new. The best known among them is the Kalman filter, which is essentially linearized iterative optimization. Each time new data is added, the solution is modified linearly under Gaussian approximation so that it is optimal over the data observed thus far. The Kalman filter was originally invented for linear dynamic systems, but its various variations and related ideas have been extended and applied to many types of computer vision problems involving a sequence of image data [3]–[6], [20], [23], [27], [35].

However, the main emphasis in such studies is the "techniques" for computing robust and accurate solutions; there has not been much systematic study of statistical error behaviors. Although error analyses have been given to 3-D motion analysis by several researchers, most of the studies were empirical and qualitative, e.g., estimating approximate orders of errors and conducting simulations with noisy data [19], [26]. A notable exception is Weng *et al.* [34], who analyzed the perturbation of the essential matrix and the resulting motion parameters in detail. However, the method involving the essential matrix is not optimal in the presence of noise.

Manuscript received March 22, 1990; revised May 29, 1992. Recommended for acceptance by Associate Editor C. Brown.

The author is with the Department of Computer Science, Gunma University, Gunma, Japan.

IEEE Log Number 9203756.

0162–8828/93$03.00 © 1993 IEEE

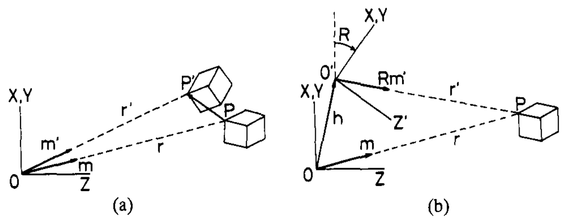


Fig. 1. Camera motion $\{\boldsymbol{R}, \boldsymbol{h}\}$ relative to a point P: (a) Description with respect to the camera coordinate system; (b) description with respect to the scene coordinate system.

The fact that the least-squares solution based on the epipolar equation are statistically biased has also been recognized [1], [29]. In this paper, we give a rigorous statistical analysis of direct optimization that does not involve the essential matrix.

We first establish a *statistical model* of noise in terms of the *covariance matrix* of the N-vector. This formulation is based on the theoretical framework Kanatani [17] called *computational projective geometry*. Then, the statistical bias of the solution is evaluated in quantitative terms, and the geometry of this bias is described in both qualitative and qualitative terms. Finally, an unbiased estimation scheme is presented, and random number simulations are conducted to observe its effectiveness.

II. General Formulation of 3-D Motion Estimation

Let $\{P_\alpha\}$, $\alpha = 1, \ldots, N$ be *feature points*, such as special markings and corner vertices, that can be clearly distinguished from other points. Let $\{\boldsymbol{m}_\alpha\}$ be unit vectors starting from the center of projection O (or the center of the lens of the camera), which we call the *viewpoint*. We call these vectors the N-vectors of the feature points. The image coordinates of the feature points play no role in obtaining 3-D interpretation other than in determining the N-vectors. Mathematically, the use of N-vectors is equivalent to considering a hypothetical spherical image surface of radius 1 centered at the viewpoint.

Since the 3-D motion of an object relative to a fixed camera is equivalent to the opposite 3-D motion of the camera relative to the object, we henceforth assume that the object we are viewing is fixed in the scene, relative to which the camera is rotated by $\boldsymbol{R}$ and translated by $\boldsymbol{h}$ (Fig. 1). Hence, the 3-D motion of the camera is specified by the *motion parameters* $\{\boldsymbol{R}, \boldsymbol{h}\}$.

If the camera is moved, feature points move into other positions on the image plane. Let $\{\boldsymbol{m}'_\alpha\}$ be their N-vectors viewed from the camera coordinate system after the motion (Fig. 1(a)). Viewed from the scene coordinate system, which we identify with the first camera coordinate system, they are $\boldsymbol{R}\boldsymbol{m}_\alpha$ because the second camera coordinate system is rotated by $\boldsymbol{R}$ relative to the first one (Fig. 1(b)). Hence, if r_α and r'_α are the *depths* (i.e., the distances from the viewpoint O) of feature point P_α before and after the motion, respectively, all we need to do is solve the following problem:

Problem 1: Given two sets of unit vectors $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}$, $\alpha = 1, \ldots, N$, compute the depths r_α and r'_α and the motion parameters $\{\boldsymbol{R}, \boldsymbol{h}\}$ that satisfy

$$r_\alpha \boldsymbol{m}_\alpha - r'_\alpha \boldsymbol{R}\boldsymbol{m}'_\alpha = \boldsymbol{h}. \tag{1}$$

From this, we immediately observe the following:

- The translation $\boldsymbol{h}$ and the depths r_α and r'_α are determined only up to scale. Namely, if r_α, r'_α and $\{\boldsymbol{R}, \boldsymbol{h}\}$ are a solution, so are kr_α, kr'_α and $\{\boldsymbol{R}, k\boldsymbol{h}\}$ for any nonzero k. This means that a large motion far from the viewer is indistinguishable from a small motion near the viewer. In order to eliminate this scale indeterminacy, we adopt the scaling $\|\boldsymbol{h}\| = 1$ whenever $\boldsymbol{h} \neq \boldsymbol{0}$.
- It is easy to check if $\boldsymbol{h} \neq \boldsymbol{0}$. If $\boldsymbol{h} = \boldsymbol{0}$, the motion is a pure 3-D rotation, and all N-vectors undergo the "camera rotation transformation" [13]–[16]:

$$\boldsymbol{m}_\alpha = \boldsymbol{R}\boldsymbol{m}'_\alpha. \tag{2}$$

  Such a rotation $\boldsymbol{R}$ is determined by the least-squares optimization

$$\sum_{\alpha=1}^{N} W_\alpha \|\boldsymbol{m}_\alpha - \boldsymbol{R}\boldsymbol{m}'_\alpha\|^2 \to \min, \tag{3}$$

  where W_α are positive weights. The solution is obtained by the method of singular value decomposition [2], [33], the method of polar decomposition [9], or the method of quaternion representation [8] (Appendix A). The residual of (3) measures the "goodness of fit," according to which a decision is made as to whether or not $\boldsymbol{h} = \boldsymbol{0}$. Assuming that $\boldsymbol{h} \neq \boldsymbol{0}$ has already been confirmed, we henceforth adopt the scaling $\|\boldsymbol{h}\| = 1$.
- There still remains indeterminacy of sign. Namely, if r_α, r'_α and $\{\boldsymbol{R}, \boldsymbol{h}\}$ are a solution, so are $-r_\alpha, -r'_\alpha$ and $\{\boldsymbol{R}, -\boldsymbol{h}\}$. Depths r_α and $-r_\alpha$ define a "mirror image pair." The correct sign is chosen by imposing the constraint that all feature points are "visible." Namely, the solution is signed so that $r_\alpha > 0$ and $r'_\alpha > 0$.
- The number of unknowns is three for $\boldsymbol{R}$ (3-D rotation is parameterized by three numbers [16]), two for $\boldsymbol{h}$ (unit vector), and $2N$ for r_α and r'_α, totaling $2N + 5$. Since the equations in (1) provide N vector equations, they assign $3N$ constraints; therefore, N must be such that $2N+5 \leq 3N$ or $N \geq 5$. Thus, the problem can be solved in principle if at least five feature points are observed, although multiple solutions may exist [7], [11], [25].

III. The Epipolar Equation

In this paper, $(\boldsymbol{a}, \boldsymbol{b})$ denotes the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$, and $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}|(= (\boldsymbol{a} \times \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{b} \times \boldsymbol{c}, \boldsymbol{a}) = (\boldsymbol{c} \times \boldsymbol{a}, \boldsymbol{b}))$, which is the scalar triple product of vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$.

Theorem 1: Two sets of unit vectors $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}$, $\alpha = 1, \ldots, N$, can be interpreted as resulting from a camera motion of motion parameters $\{\boldsymbol{R}, \boldsymbol{h}\}$ if and only if

$$|\boldsymbol{h}, \boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha| = 0. \tag{4}$$

Proof: Equation (1) states that vector $\boldsymbol{h}$ is expressed as a linear combination of unit vectors $\boldsymbol{R}\boldsymbol{m}'_\alpha$ and $\boldsymbol{m}_\alpha$. The depths

r_α and r'_α that satisfy this equation exist if and only if the three vectors $\boldsymbol{h}$, $\boldsymbol{m}_\alpha$, and $\boldsymbol{R}\boldsymbol{m}'_\alpha$ are coplanar for each $\alpha = 1, \ldots, N$. ■

We can also see from the above argument that the depths r_α and r'_α are unique (for a particular choice of the sign of $\boldsymbol{h}$) if and only if $\boldsymbol{m}_\alpha$ and $\boldsymbol{R}\boldsymbol{m}'_\alpha$ are nonparallel, or equivalently $(\boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha)^2 \neq 1$.

Equation (4) is identical to the *epipolar constraint* of converging stereo if the translation $\boldsymbol{h}$ is identified with the base-line vector. Hence, it makes sense to call (4) the *epipolar equation*. The epipolar equation (4) provides N constraints for five unknowns (two for $\boldsymbol{h}$ and three for $\boldsymbol{R}$). So again, the problem is solved if $N \geq 5$.

If the motion parameters $\{\boldsymbol{R}, \boldsymbol{h}\}$ are obtained, the depths r_α and r'_α are computed from (1). However, (1) is an *overspecification*, giving three equations for depths r_α and r'_α. Hence, there exist infinitely many ways to express the depths r_α and r'_α in terms of the motion parameters $\{\boldsymbol{R}, \boldsymbol{h}\}$. A robust approach is the least-squares optimization

$$\|r_\alpha \boldsymbol{m}_\alpha - r'_\alpha \boldsymbol{R}\boldsymbol{m}'_\alpha - \boldsymbol{h}\|^2 \rightarrow \min. \qquad (5)$$

Differentiating the left-hand side with respect to r_α and r'_α and setting the result to 0, we obtain

$$r_\alpha = \frac{(\boldsymbol{h}, \boldsymbol{m}_\alpha) - (\boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{h}, \boldsymbol{R}\boldsymbol{m}'_\alpha)}{1 - (\boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha)^2},$$

$$r'_\alpha = \frac{(\boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{h}, \boldsymbol{m}_\alpha) - (\boldsymbol{h}, \boldsymbol{R}\boldsymbol{m}'_\alpha)}{1 - (\boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha)^2}. \qquad (6)$$

The sign of the translation $\boldsymbol{h}$ is determined so that the depths r_α and r'_α become positive. If this is not possible in the presence of noise, a reasonable policy is to ask for the "majority vote": $\sum_{\alpha=1}^{N}(r_\alpha + r'_\alpha) > 0$. If $(\boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha)^2 = 1$, the depths are indeterminate.

IV. Least-Squares Optimization

In order to assure robustness of computation, we must apply some kind of optimization. The direct optimization based on the epipolar equation (4) is stated as follows:

Problem 2: Given two sets of unit vectors $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}, \alpha = 1, \ldots, N$, compute a rotation $\boldsymbol{R}$ and a unit vector $\boldsymbol{h}$ (up to sign) such that

$$\sum_{\alpha=1}^{N} W_\alpha |\boldsymbol{h}, \boldsymbol{m}_\alpha, \boldsymbol{R}\boldsymbol{m}'_\alpha|^2 \rightarrow \min. \qquad (7)$$

The sum of squares is rewritten as

$$\sum_{\alpha=1}^{N} W_\alpha (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha, \boldsymbol{h})^2 = (\boldsymbol{h}, \sum_{\alpha=1}^{N} W_\alpha (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top \boldsymbol{h}) \qquad (8)$$

where $\top$ denotes transpose. This is a quadratic form in unit vector $\boldsymbol{h}$. Hence, Problem 2 reduces to the following form:

Problem 3: Given two sets of unit vectors $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}, \alpha = 1, \ldots, N$, compute a rotation $\boldsymbol{R}$ that minimizes the smallest eigenvalue of the matrix

$$\boldsymbol{A}(\boldsymbol{R}) = \sum_{\alpha=1}^{N} W_\alpha (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top \qquad (9)$$

and compute the corresponding unit eigenvector $\boldsymbol{h}$.

This is a nonlinear optimization problem; therefore, we must resort to numerical search. In order to do iterations, we need to compute the gradient of the cost function. By means of the *quaternion representation* of 3-D rotation (Appendix A), any rotation $\boldsymbol{R}$ is expressed in terms of four real numbers q_0, q_1, q_2, and q_3 such that $q_0{}^2 + q_1{}^2 + q_2{}^2 + q_3{}^2 = 1$. Hence, the matrix $\boldsymbol{A}(\boldsymbol{R})$ can be regarded as a function of 4-D unit vector $\boldsymbol{q} = (q_0, q_1, q_2, q_3)^\top$. Its smallest eigenvalue λ_m is also regarded as a function of $\boldsymbol{q}$.

By applying the well known *perturbation theorem*, the gradient of λ_m with respect to $\boldsymbol{q}$ is obtained in the following form (Appendix B):

$$\frac{\partial \lambda_m}{\partial q_k} = (\boldsymbol{h}, \boldsymbol{T}_\kappa \boldsymbol{h}), \qquad \kappa = 0, 1, 2, 3. \qquad (10)$$

Here, $\boldsymbol{h}$ is the unit eigenvector of $\boldsymbol{A}(\boldsymbol{R})$ for the smallest eigenvalue λ_m and

$$\boldsymbol{T}_\kappa = \sum_{\alpha=1}^{N} W_\alpha [(\boldsymbol{m}_\alpha \times \boldsymbol{D}_\kappa \boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top + (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{D}_\kappa \boldsymbol{m}'_\alpha)^\top] \qquad (11)$$

where $\boldsymbol{D}_\kappa = \partial \boldsymbol{R}/\partial q_\kappa, \kappa = 0, 1, 2, 3$, (Appendix B).

Numerical search may be trapped into a local minimum if the initial guess is not close to the true solution. Fortunately, the problem can be solved analytically *if all data are exact* [21], [31], [34], [36], [37]. The analytic solution is obtained by first computing the *essential matrix* and then computing the motion parameters (the procedure of Weng *et al.* [34] is summarized in Appendix C). Although this solution is not optimal in the presence of noise, it can be used as the starting value of the optimization search. In the following, we analyze how the resulting optimal solution is affected if the original data $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}, \alpha = 1, \ldots, N$ are perturbed by noise.

V. Statistical Model of Noise

We extend the meaning of the term "noise." A digital image consists of discrete pixels, and the noise in the strict sense affects the electric signal that carries information about the gray levels of the pixels. The signal is then quantized and stored as image data. As a result, point and line data detected by applying image operations are not accurate. Consequently, N-vectors computed from them are not exact. Here, we regard such errors as caused by "noise." This means that the noise behavior is characterized not only by the camera and the memory frame system but also by the image operations involved—edge operators, thinning algorithms, etc.

Let $\boldsymbol{m}$ be the N-vector of a point in the image when there is no noise. In the presence of noise, a perturbed N-vector $\boldsymbol{m}' = \boldsymbol{m} + \Delta\boldsymbol{m}$ is observed. We regard the "noise" $\Delta\boldsymbol{m}$

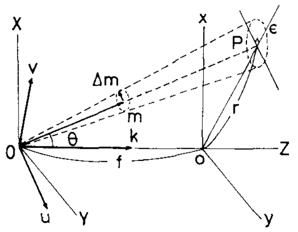


Fig. 2. Model of noise.

as a random variable. Namely, each observation is regarded as a "sample" from a "statistical ensemble." Consider the *covariance matrix*

$$V[\boldsymbol{m}] = E[\Delta\boldsymbol{m}\Delta\boldsymbol{m}^\top] \tag{12}$$

where $E[\cdot]$ denotes the expectation over the statistical ensemble. Assume that noise $\Delta\boldsymbol{m}$ is sufficiently small compared with $\boldsymbol{m}$ itself. Since $\boldsymbol{m}$ is a unit vector, noise $\Delta\boldsymbol{m}$ is always orthogonal to $\boldsymbol{m}$ to a first approximation; $\boldsymbol{m}' = \boldsymbol{m} + \Delta\boldsymbol{m}$ is again a unit vector to a first approximation. We immediately observe the following (Appendix D):

- The covariance matrix $V[\boldsymbol{m}]$ is symmetric and positive semi-definite.
- The covariance matrix $V[\boldsymbol{m}]$ is singular with $\boldsymbol{m}$ itself as the unit eigenvector for eigenvalue 0: $V[\boldsymbol{m}]\boldsymbol{m} = \boldsymbol{0}$.
- If ${\sigma_1}^2$, ${\sigma_2}^2$, and 0 ($\sigma_1 \geq \sigma_2 > 0$) are the three eigenvalues and if $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{m}\}$ is the corresponding orthonormal system of eigenvectors, the covariance matrix $V[\boldsymbol{m}]$ has the following spectral decomposition [18]:

$$V[\boldsymbol{m}] = {\sigma_1}^2\boldsymbol{u}\boldsymbol{u}^\top + {\sigma_2}^2\boldsymbol{v}\boldsymbol{v}^\top(+0\boldsymbol{m}\boldsymbol{m}^\top). \tag{13}$$

- The root mean square of the orthogonal projection of noise $\Delta\boldsymbol{m}$ onto orientation $\boldsymbol{l}$ (unit vector) takes its maximum for $\boldsymbol{l} = \boldsymbol{u}$ and its minimum for $\boldsymbol{l} = \boldsymbol{v}$. The maximum and minimum values are σ_1 and σ_2, respectively.
- The root-mean-square magnitude of $\Delta\boldsymbol{m}$ is $\sqrt{\mathrm{tr}V[\boldsymbol{m}]} = \sqrt{{\sigma_1}^2 + {\sigma_2}^2}$.

In intuitive terms, noise $\Delta\boldsymbol{m}$ is most likely to occur in orientation $\boldsymbol{u}$ (which is the unit eigenvector of $V[\boldsymbol{m}]$ for the largest eigenvalue ${\sigma_1}^2$) and least likely to occur in orientation $\boldsymbol{v}$ (which is the unit eigenvector of $V[\boldsymbol{m}]$ for the second largest eigenvalue ${\sigma_2}^2$). The magnitude $\|\Delta\boldsymbol{m}\|$ is σ_1 in orientation $\boldsymbol{u}$ and σ_2 in orientation $\boldsymbol{v}$ in the sense of root mean square.

We adopt the following model of noise: "Noise (in our extended sense) occurs at each point on the image plane placed at distance f from the viewpoint O and is equally likely in all orientations with the same root mean square ϵ" (Fig. 2). We call the constant f the *focal length* and ϵ (measured in pixels) the *image accuracy*. Let $\boldsymbol{k} = (0, 0, 1)^\top$. Our model implies the following:

Proposition 1: The covariance matrix $V[\boldsymbol{m}]$ of the N-vector of a point at distance r from the image origin is given by

$$V[\boldsymbol{m}] = \frac{\epsilon^2/f}{2(1 + r^2/f^2)}\left(\boldsymbol{u}\boldsymbol{u}^\top + \frac{1}{1 + r^2/f^2}\boldsymbol{v}\boldsymbol{v}^\top\right) \tag{14}$$

where

$$\boldsymbol{u} = \pm\sqrt{1 + \frac{f^2}{r^2}}\boldsymbol{m} \times \boldsymbol{k}, \qquad \boldsymbol{v} = \pm\boldsymbol{u} \times \boldsymbol{m}. \tag{15}$$

Proof: Let $\boldsymbol{m}$ be the N-vector of point P. The eigenvector $\boldsymbol{u}$ for the largest eigenvalue is orthogonal to the plane defined by $\boldsymbol{m}$ and $\boldsymbol{k}$ (Fig. 2). Hence, $\boldsymbol{u} = \pm N[\boldsymbol{m} \times \boldsymbol{k}] = \pm\boldsymbol{m} \times \boldsymbol{k}/\sin\theta$, where θ is the angle between $\boldsymbol{m}$ and $\boldsymbol{k}$. The eigenvector $\boldsymbol{v}$ for the second largest eigenvalue is orthogonal to both $\boldsymbol{m}$ and $\boldsymbol{u}$ and, hence, is given by $\boldsymbol{v} = \pm\boldsymbol{u} \times \boldsymbol{m}$. According to our model, noise is isotropic in the image, and the root mean square of the image error ϵ_1 in orientation $\boldsymbol{u}$ is equal to the root mean square ϵ_2 in the orientation perpendicular to it. Since the image accuracy is $\epsilon = \sqrt{{\epsilon_1}^2 + {\epsilon_2}^2}$, it follows that $\epsilon_1 = \epsilon_2 = \epsilon/\sqrt{2}$. Noise $\epsilon/\sqrt{2}$ at point P in orientation $\boldsymbol{u}$ causes a perturbation of $\boldsymbol{m}$ by $\epsilon/\sqrt{2}|OP|$, whereas in the orientation perpendicular to it, the perturbation is $\epsilon\cos\theta/\sqrt{2}|OP|$. This means that the covariance matrix $V[\boldsymbol{m}]$ is given by

$$\begin{aligned} V[\boldsymbol{m}] &= \frac{\epsilon^2}{2|OP|^2}\boldsymbol{u}\boldsymbol{u}^\top + \frac{\epsilon^2\cos^2\theta}{2|OP|^2}\boldsymbol{v}\boldsymbol{v}^\top \\ &= \frac{1}{2}\frac{\epsilon^2}{|OP|^2}(\boldsymbol{u}\boldsymbol{u}^\top + \cos^2\theta\boldsymbol{v}\boldsymbol{v}^\top). \end{aligned} \tag{16}$$

Substituting $\cos\theta = 1/\sqrt{1 + r^2/f^2}$, $\sin\theta = 1/\sqrt{1 + f^2/r^2}$, and $|OP| = f\sqrt{1 + r^2/f^2}$, we obtain (14). ∎

If the size of the image is small as compared with the focal length f, we can assume that $r \ll f$, and hence, $1/(1 + r^2/f^2) \approx 1$, which we call the *small image approximation.* In this approximation, the covariance matrix $V[\boldsymbol{m}]$ of (14) reduces to $(\epsilon^2/f^2)(\boldsymbol{u}\boldsymbol{u}^\top + \boldsymbol{v}\boldsymbol{v}^\top)/2$. If we put

$$\tilde{\epsilon} = \sqrt{E[\|\Delta\boldsymbol{m}\|^2]} \tag{17}$$

then $\tilde{\epsilon}^2 = \mathrm{tr}V[\boldsymbol{m}] = \epsilon^2/f^2$. Since $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{m}\}$ is an orthonormal system, we have $\boldsymbol{u}\boldsymbol{u}^\top + \boldsymbol{v}\boldsymbol{v}^\top + \boldsymbol{m}\boldsymbol{m}^\top = \boldsymbol{I}$ (the unit matrix). Hence, we have the following:

Corollary 1: In the small image approximation, the covariance matrix $V[\boldsymbol{m}]$ is given by

$$V[\boldsymbol{m}] = \frac{\tilde{\epsilon}^2}{2}(\boldsymbol{I} - \boldsymbol{m}\boldsymbol{m}^\top). \tag{18}$$

In our model, the true position of a point coincides with its expected position to a first approximation. This does not hold if higher order effects are considered. According to our model (Fig. 2), the image coordinates (x, y) and the corresponding N-vector $\boldsymbol{m} = (m_1, m_2, m_3)^\top$ are related in the form

$$x = f\frac{m_1}{m_3}, \qquad y = f\frac{m_2}{m_3}. \tag{19}$$

Conversely, the N-vector of a point (x, y) on the image plane is given by

$$\boldsymbol{m} = N\left[\begin{pmatrix} x \\ y \\ f \end{pmatrix}\right] \tag{20}$$

where $N[\cdot]$ denotes normalization into a unit vector. If noise $\Delta\boldsymbol{m}$ occurs, it can be shown (Appendix E) that

$$E\left[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}\right] = x\left(1 + \frac{\tilde{\epsilon}^2/2}{1 + r^2/f^2} + O(\Delta\boldsymbol{m})^3\right),$$

$$E[f\frac{m_2+\Delta m_2}{m_3+\Delta m_3}] = y\Big(1+\frac{\tilde{\epsilon}^2/2}{1+r^2/f^2}+O(\Delta \boldsymbol{m})^3\Big) \quad (21)$$

where (and hereafter) $O(\cdots)^n$ denotes a term of order n in $\cdots$. This means that the expected image coordinates do not agree with their unperturbed values. One way to remove this statistical bias is to define the N-vector $\boldsymbol{m}$ by not (20) but

$$\boldsymbol{m} = N\left[\begin{pmatrix} x \\ y \\ \hat{f} \end{pmatrix}\right] \quad (22)$$

where

$$\hat{f} = f\Big(1+\frac{\tilde{\epsilon}^2/2}{1+r^2/f^2}\Big). \quad (23)$$

We call this $\hat{f}$ the *effective focal length*. The following can be confirmed:

$$E\left[f\frac{m_1+\Delta m_1}{m_3+\Delta m_3}\right] = x + O(\Delta \boldsymbol{m})^3,$$
$$E\left[f\frac{m_2+\Delta m_2}{m_3+\Delta m_3}\right] = y + O(\Delta \boldsymbol{m})^3. \quad (24)$$

Hence, all we need to do is treat the correspondence between points in the image and their N-vectors unsymmetrically. When we define the N-vector of a "data point" (x, y), we use (22); when we interpret a "computed" N-vector $\boldsymbol{m}$, we use (19). The intuitive meaning of the effective focal length is as follows. If we define $\boldsymbol{m}$ by (20), the error model of Proposition 1 defines an error distribution symmetric with respect to the axes defined by $\boldsymbol{u}$ and $\boldsymbol{v}$. This means that when projected onto the image, noise is more likely to occur away from the image origin than toward it. Our unsymmetric treatment effectively displaces data points toward the image origin to cancel this bias. However, *this bias is a second-order effect and is extremely small.* If $\epsilon \approx 1 \sim 10$ (pixels) for $f = 1000$ (pixels), then $\tilde{\epsilon}^2 \approx 10^{-6} \sim 10^{-4}$, which can be negligible as compared with 1.

We introduce the following notations. Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be vectors, and let $\boldsymbol{A}$ be a matrix. We define $\boldsymbol{u}\times\boldsymbol{A}$ to be the matrix constructed by the vector product of $\boldsymbol{u}$ and each "column" of $\boldsymbol{A}$ and $\boldsymbol{A}\times\boldsymbol{v}$ as the matrix constructed by the vector product of each "row" and $\boldsymbol{v}$. To be precise, we define for vectors $\boldsymbol{u} = (u_i)$, $\boldsymbol{v} = (v_i)$, and matrix $\boldsymbol{A} = (A_{ij})$

$$\boldsymbol{u}\times\boldsymbol{A} = \begin{pmatrix} u_2A_{31}-u_3A_{21} & u_2A_{32}-u_3A_{22} & u_2A_{33}-u_3A_{23} \\ u_3A_{11}-u_1A_{31} & u_3A_{12}-u_1A_{32} & u_3A_{13}-u_1A_{33} \\ u_1A_{21}-u_2A_{11} & u_1A_{22}-u_2A_{12} & u_1A_{23}-u_2A_{13} \end{pmatrix}, \quad (25)$$

$$\boldsymbol{A}\times\boldsymbol{v} = \begin{pmatrix} A_{12}v_3-A_{13}v_2 & A_{13}v_1-A_{11}v_3 & A_{11}v_2-A_{12}v_1 \\ A_{22}v_3-A_{23}v_2 & A_{23}v_1-A_{21}v_3 & A_{21}v_2-A_{22}v_1 \\ A_{32}v_3-A_{33}v_2 & A_{33}v_1-A_{31}v_3 & A_{31}v_2-A_{32}v_1 \end{pmatrix}. \quad (26)$$

It can be easily shown (Appendix F) that

$$(\boldsymbol{u}\times\boldsymbol{A})\times\boldsymbol{v} = \boldsymbol{u}\times(\boldsymbol{A}\times\boldsymbol{v}) \quad (27)$$

which we simply write as $\boldsymbol{u}\times\boldsymbol{A}\times\boldsymbol{v}$.

VI. Statistical Analysis of 3-D Motion

In Problem 2, it is reasonable to measure the reliability of each N-vector $\boldsymbol{m}_\alpha$ by the mean squares $E[\|\Delta\boldsymbol{m}_\alpha\|] = \mathrm{tr}V[\boldsymbol{m}_\alpha]$ and $E[\|\Delta\boldsymbol{m}'_\alpha\|] = \mathrm{tr}V[\boldsymbol{m}'_\alpha]$; therefore, we can choose the weight

$$W_\alpha = \frac{\text{const.}}{\mathrm{tr}V[\boldsymbol{m}_\alpha]\mathrm{tr}V[\boldsymbol{m}'_\alpha]}. \quad (28)$$

Reliable data have small covariance matrices and are thereby assigned large weights, whereas unreliable data have large covariance matrices and are thereby assigned small weights. Since multiplication of W_α by a constant does not affect the solution, we adopt the scaling $\sum_{\alpha=1}^{N} W_\alpha = 1$. Define the *moment matrices* of $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}$ by

$$\boldsymbol{M} = \sum_{\alpha=1}^{N} W_\alpha \boldsymbol{m}_\alpha\boldsymbol{m}_\alpha^\top, \qquad \boldsymbol{M}' = \sum_{\alpha=1}^{N} W_\alpha \boldsymbol{m}'_\alpha\boldsymbol{m}'^\top_\alpha. \quad (29)$$

Consider Problem 3. In the presence of noise, the matrix $\boldsymbol{A}(\boldsymbol{R})$ is perturbed into

$$\tilde{\boldsymbol{A}}(\boldsymbol{R}) = \sum_{\alpha=1}^{N} W_\alpha((\boldsymbol{m}_\alpha+\Delta\boldsymbol{m}_\alpha)\times\boldsymbol{R}(\boldsymbol{m}'_\alpha+\Delta\boldsymbol{m}'_\alpha))$$
$$((\boldsymbol{m}_\alpha+\Delta\boldsymbol{m}_\alpha)\times\boldsymbol{R}(\boldsymbol{m}'_\alpha+\Delta\boldsymbol{m}'_\alpha))^\top. \quad (30)$$

Employing the small image approximation and assuming that the data are statistically independent, we now show that this perturbation is statistically biased.

Lemma 1:

$$E[\tilde{\boldsymbol{A}}(\boldsymbol{R})] = (1-\frac{\tilde{\epsilon}^2}{2})^2\boldsymbol{A}(\boldsymbol{R})$$
$$-\frac{\tilde{\epsilon}^2}{2}(1-\frac{\tilde{\epsilon}^2}{2})(\boldsymbol{M}+\boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top)+\tilde{\epsilon}^2\boldsymbol{I}. \quad (31)$$

Proof: If $\boldsymbol{m}_\alpha$ and $\boldsymbol{m}'_\alpha$ are perturbed by $\Delta\boldsymbol{m}_\alpha$ and $\Delta\boldsymbol{m}'_\alpha$, respectively, we have $E[\Delta\boldsymbol{m}_\alpha] = E[\Delta\boldsymbol{m}'_\alpha] = \boldsymbol{0}$, $E[\Delta\boldsymbol{m}_\alpha\Delta\boldsymbol{m}'^\top_\beta] = E[\Delta\boldsymbol{m}'_\alpha\Delta\boldsymbol{m}^\top_\beta] = \boldsymbol{0}$, and

$$E[\Delta\boldsymbol{m}_\alpha\Delta\boldsymbol{m}_\beta^\top] = \delta_{\alpha\beta}V[\boldsymbol{m}_\alpha],$$
$$E[\Delta\boldsymbol{m}'_\alpha\Delta\boldsymbol{m}'^\top_\beta] = \delta_{\alpha\beta}V[\boldsymbol{m}'_\alpha] \quad (32)$$

where $\delta_{\alpha\beta}$ is the Kronecker delta. Hence, the expectation of $\tilde{\boldsymbol{A}}(\boldsymbol{R})$ is

$$E[\tilde{\boldsymbol{A}}(\boldsymbol{R})] = \sum_{\alpha=1}^{N} W_\alpha((\boldsymbol{m}_\alpha\times\boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha\times\boldsymbol{R}\boldsymbol{m}'_\alpha)^\top$$
$$+E[(\Delta\boldsymbol{m}_\alpha\times\boldsymbol{R}\boldsymbol{m}'_\alpha)(\Delta\boldsymbol{m}_\alpha\times\boldsymbol{R}\boldsymbol{m}'_\alpha)^\top]$$
$$+E[(\boldsymbol{m}_\alpha\times\boldsymbol{R}\Delta\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha\times\boldsymbol{R}\Delta\boldsymbol{m}'_\alpha)^\top]$$
$$+E[(\Delta\boldsymbol{m}_\alpha\times\boldsymbol{R}\Delta\boldsymbol{m}'_\alpha)(\Delta\boldsymbol{m}_\alpha\times\boldsymbol{R}\Delta\boldsymbol{m}'_\alpha)^\top]). \quad (33)$$

Using the identities $\boldsymbol{u}\times(\boldsymbol{a}\boldsymbol{b}^\top)\times\boldsymbol{v} = (\boldsymbol{u}\times\boldsymbol{a})(\boldsymbol{b}\times\boldsymbol{v})^\top$ and $\boldsymbol{u}\times\boldsymbol{I}\times\boldsymbol{u} = \boldsymbol{u}\boldsymbol{u}^\top-\boldsymbol{I}$ for unit vectors $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{a}$, and $\boldsymbol{b}$ (Appendix

F), we obtain

$$E[(\Delta \boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\Delta \boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top] = -\boldsymbol{R}\boldsymbol{m}'_\alpha \times E[\Delta \boldsymbol{m}_\alpha \Delta \boldsymbol{m}_\alpha^\top] \times \boldsymbol{R}\boldsymbol{m}'_\alpha = -\boldsymbol{R}\boldsymbol{m}'_\alpha \times V[\boldsymbol{m}_\alpha] \times \boldsymbol{R}\boldsymbol{m}'_\alpha = -\frac{\tilde{\epsilon}^2}{2}\boldsymbol{R}\boldsymbol{m}'_\alpha \times (\boldsymbol{I} - \boldsymbol{m}_\alpha \boldsymbol{m}_\alpha^\top) \times \boldsymbol{R}\boldsymbol{m}'_\alpha = \frac{\tilde{\epsilon}^2}{2}(\boldsymbol{I} - (\boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{R}\boldsymbol{m}'_\alpha)^\top - (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top). \quad (34)$$

Similarly

$$E[(\boldsymbol{m}_\alpha \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha)^\top] = -\boldsymbol{m}_\alpha \times \boldsymbol{R}E[\Delta \boldsymbol{m}'_\alpha \Delta \boldsymbol{m}'^\top_\alpha]\boldsymbol{R}^\top \times \boldsymbol{m}'_\alpha = -\boldsymbol{m}_\alpha \times \boldsymbol{R}V[\boldsymbol{m}'_\alpha]\boldsymbol{R}^\top \times \boldsymbol{m}'_\alpha = -\frac{\tilde{\epsilon}^2}{2}\boldsymbol{m}_\alpha \times \boldsymbol{R}(\boldsymbol{I} - \boldsymbol{m}'_\alpha \boldsymbol{m}'^\top_\alpha)\boldsymbol{R}^\top \times \boldsymbol{m}'_\alpha = \frac{\tilde{\epsilon}^2}{2}(\boldsymbol{I} - \boldsymbol{m}_\alpha \boldsymbol{m}_\alpha^\top - (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top). \quad (35)$$

On the other hand, we have (36), which is at the bottom of this page, where (35) is used. Substituting (34)–(36) into (33), we obtain

$$E[\tilde{\boldsymbol{A}}(\boldsymbol{R})] = (1 - \frac{\tilde{\epsilon}^2}{2})^2 \sum_{\alpha=1}^N W_\alpha (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top - \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2}) \sum_{\alpha=1}^N W_\alpha (\boldsymbol{m}_\alpha \boldsymbol{m}_\alpha^\top + (\boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{R}\boldsymbol{m}'_\alpha)^\top) + \tilde{\epsilon}^2 \boldsymbol{I} = (1 - \frac{\tilde{\epsilon}^2}{2})^2 \boldsymbol{A}(\boldsymbol{R}) - \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2})\left(\left(\sum_{\alpha=1}^N W_\alpha \boldsymbol{m}_\alpha \boldsymbol{m}_\alpha^\top\right) + \boldsymbol{R}\left(\sum_{\alpha=1}^N W_\alpha \boldsymbol{m}'_\alpha \boldsymbol{m}'^\top_\alpha\right)\boldsymbol{R}^\top\right) + \tilde{\epsilon}^2 \boldsymbol{I} \quad (37)$$

from which follows (31). ∎

Theorem 2: Computation of the rotation $\boldsymbol{R}$ is statistically biased by a small positive angle about axis

$$\tilde{\boldsymbol{l}} \approx -N[\boldsymbol{h} \times \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top \boldsymbol{h}]. \quad (38)$$

Proof: To a first approximation, the bias of the rotation $\boldsymbol{R}$ and the bias of the translation $\boldsymbol{h}$ can be estimated independently. Hence, we can assume that the translation $\boldsymbol{h}$ has its true value $\bar{\boldsymbol{h}}$. According to the perturbation theorem (Appendix B), the smallest eigenvalue of $\tilde{\boldsymbol{A}}(\boldsymbol{R})$ is $\tilde{\lambda}_m(\boldsymbol{R}) = (\bar{\boldsymbol{h}}, \tilde{\boldsymbol{A}}(\boldsymbol{R})\bar{\boldsymbol{h}})$ to a first approximation. From Lemma 1, its expectation $E[\tilde{\lambda}_m(\boldsymbol{R})] = (\bar{\boldsymbol{h}}, E[\tilde{\boldsymbol{A}}(\boldsymbol{R})]\bar{\boldsymbol{h}})$ is

$$E[\tilde{\lambda}_m(\boldsymbol{R})] = (1 - \frac{\tilde{\epsilon}^2}{2})^2 (\bar{\boldsymbol{h}}, \boldsymbol{A}(\boldsymbol{R})\bar{\boldsymbol{h}}) - \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2})(\bar{\boldsymbol{h}}, (\boldsymbol{M} + \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top)\bar{\boldsymbol{h}}) + \tilde{\epsilon}^2 \boldsymbol{I}. \quad (39)$$

Let $\bar{\boldsymbol{R}}$ be the true rotation that minimizes the smallest eigenvalue of $\boldsymbol{A}(\boldsymbol{R})$, and put $\boldsymbol{R} = \bar{\boldsymbol{R}} + \Delta \boldsymbol{R}$. Since $(\bar{\boldsymbol{h}}, \boldsymbol{A}(\bar{\boldsymbol{R}} + \Delta \boldsymbol{R})\bar{\boldsymbol{h}}) = O(\Delta \boldsymbol{R})^2$, we have

$$E[\tilde{\lambda}_m(\bar{\boldsymbol{R}} + \Delta \boldsymbol{R})] = \tilde{\epsilon}^2 \boldsymbol{I} - \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2})(\bar{\boldsymbol{h}}, (\boldsymbol{M} + \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top)\bar{\boldsymbol{h}}) - \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2})(\bar{\boldsymbol{h}}, (\Delta \boldsymbol{R}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}}\boldsymbol{M}'\Delta \boldsymbol{R}^\top)\bar{\boldsymbol{h}}) + O(\Delta \boldsymbol{R})^2. \quad (40)$$

Since $\bar{\boldsymbol{R}} + \Delta \boldsymbol{R}$ is a rotation matrix, there exists a vector $\Delta \boldsymbol{l}$ such that $\Delta \boldsymbol{R} = \Delta \boldsymbol{l} \times \bar{\boldsymbol{R}}$ to a first approximation. Hence

$$(\bar{\boldsymbol{h}}, (\Delta \boldsymbol{R}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top + \bar{\boldsymbol{R}}\boldsymbol{M}'\Delta \boldsymbol{R}^\top)\bar{\boldsymbol{h}}) = 2(\bar{\boldsymbol{h}}, \Delta \boldsymbol{R}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{h}}) = 2(\bar{\boldsymbol{h}}, \Delta \boldsymbol{l} \times \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{h}}) = -2(\Delta \boldsymbol{l}, \bar{\boldsymbol{h}} \times \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{h}}). \quad (41)$$

Consequently

$$E[\tilde{\lambda}_m(\bar{\boldsymbol{R}} + \Delta \boldsymbol{R})] = \tilde{\epsilon}^2 \boldsymbol{I} - \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2})(\bar{\boldsymbol{h}}, (\boldsymbol{M} + \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top)\bar{\boldsymbol{h}}) + \tilde{\epsilon}^2(1 - \frac{\tilde{\epsilon}^2}{2})(\Delta \boldsymbol{l}, \bar{\boldsymbol{h}} \times \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{h}}) + O(\Delta \boldsymbol{R})^2. \quad (42)$$

This is minimized most steeply by

$$\Delta \boldsymbol{l} = -c\bar{\boldsymbol{h}} \times \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{h}} \quad (43)$$

where c is a sufficiently small positive constant. In other words, the rotation that minimizes the perturbed $\tilde{\lambda}_m(\boldsymbol{R})$ is biased by a small rotation about $N[\Delta \boldsymbol{l}] = -N[\bar{\boldsymbol{h}} \times \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top \bar{\boldsymbol{h}}]$. Omitting the bars, we obtain (38). ∎

$$E[(\Delta \boldsymbol{m}_\alpha \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha)(\Delta \boldsymbol{m}_\alpha \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha)^\top] = -E[\boldsymbol{R}\Delta \boldsymbol{m}'_\alpha \times (\Delta \boldsymbol{m}_\alpha \Delta \boldsymbol{m}_\alpha^\top) \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha] = -E[\boldsymbol{R}\Delta \boldsymbol{m}'_\alpha \times E[\Delta \boldsymbol{m}_\alpha \Delta \boldsymbol{m}_\alpha^\top] \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha] = -\frac{\tilde{\epsilon}^2}{2}E[\boldsymbol{R}\Delta \boldsymbol{m}'_\alpha \times (\boldsymbol{I} - \boldsymbol{m}_\alpha \boldsymbol{m}_\alpha^\top) \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha] = \frac{\tilde{\epsilon}^2}{2}\left(E[\|\Delta \boldsymbol{m}'_\alpha\|^2]\boldsymbol{I} - \boldsymbol{R}E[\Delta \boldsymbol{m}'_\alpha \Delta \boldsymbol{m}'^\top_\alpha]\boldsymbol{R}^\top - E[(\boldsymbol{m}_\alpha \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\Delta \boldsymbol{m}'_\alpha)^\top]\right) = \frac{\tilde{\epsilon}^4}{4}\left(2\boldsymbol{I} - \boldsymbol{R}(\boldsymbol{I} - \boldsymbol{m}'_\alpha \boldsymbol{m}'^\top_\alpha)\boldsymbol{R}^\top - (\boldsymbol{I} - \boldsymbol{m}_\alpha \boldsymbol{m}_\alpha^\top - (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top)\right) = \frac{\tilde{\epsilon}^4}{4}(\boldsymbol{m}_\alpha \boldsymbol{m}_\alpha^\top + (\boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{R}\boldsymbol{m}'_\alpha)^\top + (\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top). \quad (36)$$

Theorem 3: Computation of the translation $\boldsymbol{h}$ is statistically biased by

$$\Delta \boldsymbol{h} \approx C\Big((\boldsymbol{M} + \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top)\boldsymbol{h} - (\boldsymbol{h}, (\boldsymbol{M} + \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top)\boldsymbol{h})\boldsymbol{h}\Big) \tag{44}$$

where C is a small positive constant.

Proof: This time, we can assume that the rotation $\boldsymbol{R}$ has its true value $\bar{\boldsymbol{R}}$. Let $\bar{\boldsymbol{h}}$ be the true translation. If $\boldsymbol{A}(\bar{\boldsymbol{R}})$ is perturbed into $\tilde{\boldsymbol{A}}(\bar{\boldsymbol{R}})$ and if $\boldsymbol{h} + \Delta\boldsymbol{h}$ is the unit eigenvector for its smallest eigenvalue, the smallest eigenvalue is

$$\tilde{\lambda}_m(\bar{\boldsymbol{R}}) = (\bar{\boldsymbol{h}} + \Delta\boldsymbol{h}, \tilde{\boldsymbol{A}}(\boldsymbol{R})(\bar{\boldsymbol{h}} + \Delta\boldsymbol{h})). \tag{45}$$

From Lemma 1 and $(\bar{\boldsymbol{h}}, \boldsymbol{A}(\bar{\boldsymbol{R}})\bar{\boldsymbol{h}}) = 0$, its expectation is

$$E[\tilde{\lambda}_m(\bar{\boldsymbol{R}})] = -\frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2})(\bar{\boldsymbol{h}}, (\boldsymbol{M} + \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top)\bar{\boldsymbol{h}}) + \tilde{\epsilon}^2\boldsymbol{I}$$
$$- \tilde{\epsilon}^2(1 - \frac{\tilde{\epsilon}^2}{2})(\Delta\boldsymbol{h}, (\boldsymbol{M} + \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top)\bar{\boldsymbol{h}}) + O(\Delta\boldsymbol{h})^2. \tag{46}$$

This is most steeply minimized if $\Delta\boldsymbol{h}$ is in the direction of $(\boldsymbol{M} + \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top)\bar{\boldsymbol{h}}$, but since $\bar{\boldsymbol{h}} + \Delta\boldsymbol{h}$ must be a unit vector, the orientation of the steepest descent is the projection of $(\boldsymbol{M} + \bar{\boldsymbol{R}}\boldsymbol{M}'\bar{\boldsymbol{R}}^\top)\bar{\boldsymbol{h}}$ onto the plane perpendicular to $\bar{\boldsymbol{h}}$. Omitting the bar, we obtain (44). ∎

VII. Small Object Approximation

If the object we are viewing is small and if all the feature points are concentrated in a small region of depth $r \gg \|\boldsymbol{h}\|(= 1)$, we can approximate $\boldsymbol{m}_\alpha$ by $\bar{\boldsymbol{m}}$. Since $\boldsymbol{m}'_\alpha = \boldsymbol{R}^\top N[r_\alpha\boldsymbol{m}_\alpha - \boldsymbol{h}]$ (see (1)), we can also approximate $\boldsymbol{m}'_\alpha$ by $\boldsymbol{R}^\top N[\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h}]$. Hence

$$\boldsymbol{M} \approx \bar{\boldsymbol{m}}\bar{\boldsymbol{m}}^\top, \qquad \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top \approx N[\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h}]N[\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h}]^\top \tag{47}$$

which we call the *small object approximation*.

Proposition 2: In the small object approximation, the axis $\boldsymbol{l}$ of the statistical bias of the computed rotation $\boldsymbol{R}$ is

$$\boldsymbol{l} \approx \mathrm{sgn}((\bar{\boldsymbol{m}}, \boldsymbol{h}))N[\bar{\boldsymbol{m}} \times \boldsymbol{h}]. \tag{48}$$

Proof: Since $\|\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h}\|^2 = \bar{r}^2 - 2\bar{r}(\bar{\boldsymbol{m}}, \boldsymbol{h}) + 1$ and $r \gg 1$, the second equation in (47) implies

$$\boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top\boldsymbol{h} \approx \frac{(\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h})(\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h})^\top}{\bar{r}^2 - 2\bar{r}(\bar{\boldsymbol{m}}, \boldsymbol{h}) + 1}\boldsymbol{h}$$
$$= \frac{\bar{r}(\bar{\boldsymbol{m}}, \boldsymbol{h}) - 1}{\bar{r}^2 - 2\bar{r}(\bar{\boldsymbol{m}}, \boldsymbol{h}) + 1}(\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h}) \approx \frac{(\bar{\boldsymbol{m}}, \boldsymbol{h})}{\bar{r}}(\bar{r}\bar{\boldsymbol{m}} - \boldsymbol{h}). \tag{49}$$

Hence

$$\boldsymbol{h} \times \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top\boldsymbol{h} \approx (\bar{\boldsymbol{m}}, \boldsymbol{h})\boldsymbol{h} \times \bar{\boldsymbol{m}}. \tag{50}$$

From Theorem 2, we obtain (48). ∎

Proposition 3: In the small object approximation, the statistical bias $\Delta\boldsymbol{h}$ of the computed translation $\boldsymbol{h}$ is

$$\Delta\boldsymbol{h} \approx C(\bar{\boldsymbol{m}}, \boldsymbol{h})(\bar{\boldsymbol{m}} - (\bar{\boldsymbol{m}}, \boldsymbol{h})\boldsymbol{h}). \tag{51}$$

where C is a positive constant.

Proof: From (47) and (49), we have

$$(\boldsymbol{M} + \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top)\boldsymbol{h} \approx (\bar{\boldsymbol{m}}, \boldsymbol{h})\bar{\boldsymbol{m}} + (\bar{\boldsymbol{m}}, \boldsymbol{h})(\boldsymbol{m} - \frac{\boldsymbol{h}}{\bar{r}}). \tag{52}$$

Substituting this into (44) and renaming $2C$ as C, we obtain (51). ∎

From Propositions 2 and 3, we observe the following:

- The computed rotation $\boldsymbol{R}$ is biased by a small rotation about an axis approximately perpendicular to both the translation $\boldsymbol{h}$ and the viewing orientation $\bar{\boldsymbol{m}}$.
- The bias of the rotation is such that the viewing orientation $\bar{\boldsymbol{m}}$ is rotated closer to $\boldsymbol{h}$ if the camera approaches the object and away from $\boldsymbol{h}$ if the camera recedes from the object.
- The computed translation $\boldsymbol{h}$ is biased toward the viewing orientation $\bar{\boldsymbol{m}}$.
- The biases of the rotation $\boldsymbol{R}$ and the translation $\boldsymbol{h}$ are both minimum when the camera moves in the viewing orientation $\bar{\boldsymbol{m}}$ or perpendicular to it.

VIII. Unbiased Estimation of Motion Parameters

From Lemma 1, we can construct a scheme for estimating unbiased motion parameters.

Theorem 4: Given two sets of N vectors $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}$, the rotation $\hat{\boldsymbol{R}}$ that minimizes the smallest eigenvalue of

$$\hat{\boldsymbol{A}}(\boldsymbol{R}) = \sum_{\alpha=1}^{N} W_\alpha(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)$$
$$(\boldsymbol{m}_\alpha \times \boldsymbol{R}\boldsymbol{m}'_\alpha)^\top + \frac{\tilde{\epsilon}^2}{2}(\boldsymbol{M} + \boldsymbol{R}\boldsymbol{M}'\boldsymbol{R}^\top) - \tilde{\epsilon}^2\boldsymbol{I} \tag{53}$$

is an unbiased estimate of $\boldsymbol{R}$. The corresponding unit eigenvector $\hat{\boldsymbol{h}}$ is an unbiased estimate of $\boldsymbol{h}$.

Proof: If there is no noise ($\tilde{\epsilon} = 0$), the matrix $\hat{\boldsymbol{A}}(\boldsymbol{R})$ reduces to $\boldsymbol{A}(\boldsymbol{R})$. In the presence of noise, $\hat{\boldsymbol{A}}(\boldsymbol{R})$ becomes

$$\tilde{\boldsymbol{A}}(\boldsymbol{R}) + \frac{\tilde{\epsilon}^2}{2}(\tilde{\boldsymbol{M}} + \boldsymbol{R}\tilde{\boldsymbol{M}}'\boldsymbol{R}^\top) - \tilde{\epsilon}^2\boldsymbol{I} \tag{54}$$

where $\tilde{\boldsymbol{A}}(\boldsymbol{R})$ is the matrix given by (30). Matrices $\tilde{\boldsymbol{M}}$ and $\tilde{\boldsymbol{M}}'$ are given by

$$\tilde{\boldsymbol{M}} = \sum_{\alpha=1}^{N} W_\alpha(\boldsymbol{m}_\alpha + \Delta\boldsymbol{m}_\alpha)(\boldsymbol{m}_\alpha + \Delta\boldsymbol{m}_\alpha)^\top,$$
$$\tilde{\boldsymbol{M}}' = \sum_{\alpha=1}^{N} W_\alpha(\boldsymbol{m}'_\alpha + \Delta\boldsymbol{m}'_\alpha)(\boldsymbol{m}'_\alpha + \Delta\boldsymbol{m}'_\alpha)^\top. \tag{55}$$

We can easily see that

$$E[\tilde{\boldsymbol{M}}] = \sum_{\alpha=1}^{N} W_\alpha(\boldsymbol{m}_\alpha\boldsymbol{m}_\alpha^\top + E[\Delta\boldsymbol{m}_\alpha\Delta\boldsymbol{m}_\alpha^\top])$$
$$= \boldsymbol{M} + \sum_{\alpha=1}^{N} W_\alpha V[\boldsymbol{m}_\alpha]$$
$$= \boldsymbol{M} + \frac{\tilde{\epsilon}^2}{2}\sum_{\alpha}^{N} W_\alpha(\boldsymbol{I} - \boldsymbol{m}_\alpha\boldsymbol{m}_\alpha^\top)$$
$$= (1 - \frac{\tilde{\epsilon}^2}{2})\boldsymbol{M} + \frac{\tilde{\epsilon}^2}{2}\boldsymbol{I}. \tag{56}$$

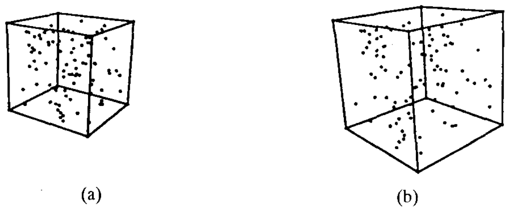


Fig. 3. Simulated 512×512-pixel images of 100 feature points randomly generated inside a cube before and after a small camera motion.

Similarly

$$E[\tilde{M}'] = (1 - \frac{\tilde{\epsilon}^2}{2})M' + \frac{\tilde{\epsilon}^2}{2}I. \tag{57}$$

From Lemma 1 and (56) and (57), the expectation of (54) is

$$\begin{aligned} &(1 - \frac{\tilde{\epsilon}^2}{2})^2 A(R) - \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2}) \\ &(M + RM'R^\top) + \tilde{\epsilon}^2 I \\ &+ \frac{\tilde{\epsilon}^2}{2}(1 - \frac{\tilde{\epsilon}^2}{2})(M + RM'R^\top) + \frac{\tilde{\epsilon}^4}{2}I - \tilde{\epsilon}^2 I \\ &= (1 - \frac{\tilde{\epsilon}^2}{2})^2 A(R) + \frac{\tilde{\epsilon}^4}{2}I \end{aligned} \tag{58}$$

which has the same eigenvectors as $A(R)$; if λ_m is the smallest eigenvalue of $A(R)$, the smallest eigenvalue of the above matrix is $\tilde{\lambda}_m = (1 - \tilde{\epsilon}^2/2)^2\lambda_m + \tilde{\epsilon}^4/2$. Hence, minimizing $\tilde{\lambda}_m$ is equivalent to minimizing λ_m. ■

IX. Random Number Simulations

Fig. 3(a) and (b) are simulated 512×512-pixel images of 100 feature points randomly generated inside a cube before and after a small camera motion. The focal length is set to $f = 500$ (pixels). A random noise obeying a normal distribution of standard deviation $\epsilon = 1.0$ (pixel) is added to the x and y coordinates of all the feature points independently before and after the motion.

The discrepancy between the computed rotation R and the true rotation $\bar{R}$ is measured by vector $\Delta l = \Delta\Omega l'$, where $\Delta\Omega$ and l' (unit vector) are, respectively, the angle and axis of relative rotation $R\bar{R}^\top$ (i.e., $R = \bar{R} + \Delta l \times R + O(\Delta l)^2$). Fig. 4(a) shows orthogonal projections of Δl onto the plane perpendicular to the axis $\tilde{l}$ predicted by Theorem 2 for 100 trials, each time using different noise. Vector $\tilde{l}$ points upward. White circles indicate upward orientations, and black circles indicate downward orientations. The large circle indicates the magnitude 0.5°. We see that the error of rotation is biased around the axis $\tilde{l}$. Fig. 4(b) shows orthogonal projections of the computed translation h onto the plane perpendicular to the true translation $\bar{h}$. The solid line indicates the orientation predicted by Theorem 3, and the large circle indicates the magnitude 0.2. We see that the error of translation is biased in the orientation predicted by Theorem 3. Fig. 5(a) is the histogram of the error $\Delta\Omega$ (in degrees), and Fig. 5(b) is the histogram of the angle $\Delta\theta = \cos^{-1}(\bar{h}, h)$ (in degrees). The root mean squares of $\Delta\Omega$ and $\Delta\theta$ are 3.73° and 4.83°, respectively.

Fig. 6(a) and (b), which, respectively, correspond to Fig. 4(a) and (b), are obtained by applying the unbiased scheme of Theorem 4. It is clearly seen that the error Δl is almost uniformly distributed in all directions. Fig. 6(b) also shows that the bias is removed and the magnitude of the error is reduced as well. Fig. 7(a) is the histogram of the error $\Delta\Omega$ (in degrees), and Fig. 7(b) is the histogram of the angle $\Delta\theta = \cos^{-1}(\bar{h}, h)$ (in degrees). The magnitude of error is drastically reduced for both rotation and translation. The root mean squares of $\Delta\Omega$ and $\Delta\theta$ are 0.47 and 0.68°, respectively. In all experiments, the effects of the effective focal length discussed in Section V are negligibly small, as expected.

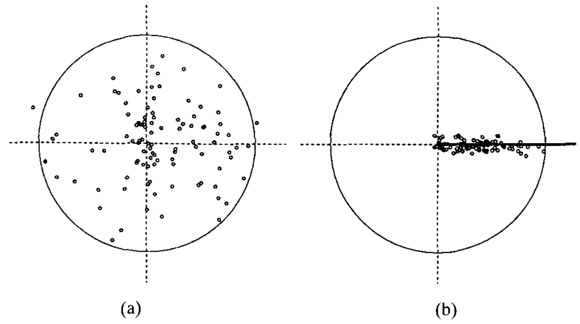


Fig. 4. (a) Projected errors of rotation and (b) projected errors of translation for 100 trials. The solid line in (b) indicates the predicted orientation of the bias.

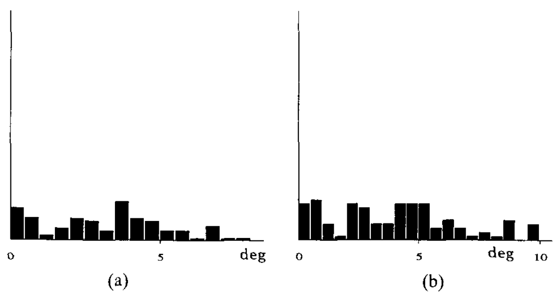


Fig. 5. (a) Histogram of errors of rotation and (b) histogram of errors of translation.

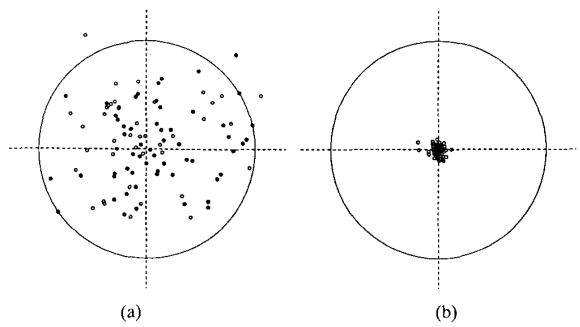


Fig. 6. (a) Projected errors of rotation and (b) projected errors of translation for 100 trials of unbiased estimates.

X. Concluding Remarks

In this paper, the problem of estimating 3-D rigid motion from two views was formulated as nonlinear optimization, and the statistical behaviors of the errors in the solution were

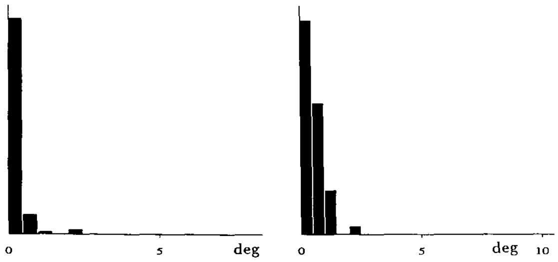


Fig. 7. (a) Histogram of errors of rotation and (b) histogram of errors of translation of unbiased estimates.

analyzed by introducing a realistic model of noise. It has been shown that the optimal solution based on the epipolar equation is "statistically biased." The expected bias was evaluated in analytical terms, and its geometry was described in both quantitative and qualitative terms by employing the small object approximation. Finally, an unbiased estimation scheme was presented, and random number simulations were conducted to observe its effectiveness.

In order to apply the unbiased scheme, the statistical behaviors of noise—in particular, the root mean square of the error in each feature point—must be known. Many approaches are conceivable for estimating them. For example, we can use an "*a posteriori* estimation"—we guess the error magnitude and compute a 3-D interpretation, according to the feature points that are matched on the image plane. Then, the amount of average mismatch is regarded as the magnitude of noise. This process can be iterated if necessary. Alternately, we can guess the magnitude of noise from the residual of the least-squares optimization. Detailed studies and comparisons of such error estimating techniques are beyond the scope of this paper and left to future research. The purpose of this paper it to establish the fact that accuracy can be greatly increased by correctly estimating the magnitude of noise and to present a new mathematical framework suitable for statistical error analysis.

APPENDIX A
3-D ROTATION FITTING

Consider the following problem:

Problem A.1: *Given two sets of vectors* $\{\boldsymbol{m}_\alpha\}$ *and* $\{\boldsymbol{m}'_\alpha\}$, $\alpha = 1, \ldots, N$, *compute a rotation* $\boldsymbol{R}$ *such that*

$$\sum_{\alpha=1}^{N} W_\alpha \|\boldsymbol{m}_\alpha - \boldsymbol{R}\boldsymbol{m}'_\alpha\|^2 \to \min \tag{59}$$

where W_α *are nonnegative weights.*

If the *correlation matrix* $\boldsymbol{K}$ between $\{\boldsymbol{m}_\alpha\}$ and $\{\boldsymbol{m}'_\alpha\}$ is defined by

$$\boldsymbol{K} = \sum_{\alpha=1}^{N} W_\alpha \boldsymbol{m}_\alpha \boldsymbol{m}'^{\top}_\alpha \tag{60}$$

Problem A.1 is restated as follows:

Problem A.2: *Given a correlation matrix* $\boldsymbol{K}$, *compute a rotation* $\boldsymbol{R}$ *such that*

$$\mathrm{tr}(\boldsymbol{R}^\top \boldsymbol{K}) \to \max. \tag{61}$$

Analytical procedures to solve this problem were proposed independently by Horn [8], using the "quaternion representation" of 3-D rotation by Arun *et al.* [2], using "singular value decomposition," and by Horn *et al.* [9], using "polar decomposition." The methods of Arun *et al.* [2] and Horn *et al.* [9] carry out minimization over orthogonal matrices. Umeyama [33] modified their methods so that minimization is carried out over rotations. We state these results in a more refined form without proofs (see [18] for the details).

The first method is to decompose the correlation matrix $\boldsymbol{K}$ into the form

$$\boldsymbol{K} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{U}^\top, \quad \boldsymbol{\Lambda} = \begin{pmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{pmatrix}, \quad \sigma_1 \geq \sigma_2 \geq \sigma_3 \geq 0 \tag{62}$$

where $\boldsymbol{V}$ and $\boldsymbol{U}$ are orthogonal matrices. This decomposition is called the *singular value decomposition*, and σ_1, σ_2, and σ_3 are the *singular values*, where the number of nonzero singular values are the rank of $\boldsymbol{K}$. The following theorem is mathematically equivalent to Umeyama's extension [33] of the method of Arun *et al.* [2].

Theorem A.1: *If* $\boldsymbol{K} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{U}^\top$ *is the singular value decomposition,* $\mathrm{tr}(\boldsymbol{R}^\top\boldsymbol{K})$ *is maximized over all rotations if*

$$\boldsymbol{R} = \boldsymbol{V}\begin{pmatrix} 1 & & \\ & 1 & \\ & & \det(\boldsymbol{V}\boldsymbol{U}^\top) \end{pmatrix}\boldsymbol{U}^\top. \tag{63}$$

The solution is unique if $\mathrm{rank}\boldsymbol{K} > 1$ *and* $\det(\boldsymbol{V}\boldsymbol{U}^\top) = 1$ *or if* $\mathrm{rank}\boldsymbol{K} > 1$ *and the minimum singular value is a simple root.*

The second method to solve Problem A.2 is to decompose the correlation matrix $\boldsymbol{K}$ into the form

$$\boldsymbol{K} = \boldsymbol{V}\boldsymbol{S} = \boldsymbol{S}'\boldsymbol{V} \tag{64}$$

where $\boldsymbol{V}$ is an orthogonal matrix, and $\boldsymbol{S}$ and $\boldsymbol{S}'$ are semi-positive definite symmetric matrices. This decomposition is known as the *polar decomposition*. The following theorem is an extension to the method of Horn *et al.* [9].

Theorem A.2: *If* $\boldsymbol{K} = \boldsymbol{V}\boldsymbol{S} = \boldsymbol{S}'\boldsymbol{V}$ *is the polar decomposition,* $\mathrm{tr}(\boldsymbol{R}^\top\boldsymbol{K})$ *is maximized over all rotations if*

$$\begin{aligned} \boldsymbol{R} &= \boldsymbol{V}(\boldsymbol{I} + (\det\boldsymbol{V} - 1)\boldsymbol{u}_m\boldsymbol{u}_m^\top) \\ &= (\boldsymbol{I} + (\det\boldsymbol{V} - 1)\boldsymbol{v}_m\boldsymbol{v}_m^\top)\boldsymbol{V} \end{aligned} \tag{65}$$

where $\boldsymbol{u}_m$ *and* $\boldsymbol{v}_m$ *are the unit eigenvectors of* $\boldsymbol{S}$ *and* $\boldsymbol{S}'$, *respectively, for the smallest eigenvalue. The solution is unique if* $\mathrm{rank}\boldsymbol{K} > 1$ *and* $\det\boldsymbol{V} = 1$, *or if* $\mathrm{rank}\boldsymbol{K} > 1$, *and the smallest eigenvalue of* $\boldsymbol{S}$*(and of* $\boldsymbol{S}'$*) is a simple root.*

The third method is based on the well-known fact that for any rotation matrix $\boldsymbol{R}$, there exist four numbers q_0, q_1, q_2, and q_3 such that $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$ and (66), which is shown at the bottom of the next page. Conversely, any four numbers q_0,

q_1, q_2, and q_3 such that $q_0^2+q_1^2+q_2^2+q_3^2 = 1$ define a rotation by this equation. This is known as the *quaternion representation* of 3-D rotation (see [16]). The following theorem summarizes the method of Horn [8].

Theorem A.3: *Given correlation matrix $\boldsymbol{K}$, define a 4-D symmetric matrix, which is shown in (67) at the bottom of this page. Let $\hat{\boldsymbol{q}}$ be the 4-D unit eigenvector of $\hat{\boldsymbol{K}}$ for the largest eigenvalue. Then, $\mathrm{tr}(\boldsymbol{R}^\top \boldsymbol{K})$ is maximized by the rotation represented by $\hat{\boldsymbol{q}}$. The solution is unique if the largest eigenvalue of $\hat{\boldsymbol{K}}$ is a simple root.*

Appendix B
Perturbation Theorem

The following is the well-known *perturbation theorem*. We omit the proof (see [18] for the details).

Proposition B.1: *Let $\boldsymbol{A}$ be an n-dimensional symmetric matrix having eigenvalues $\lambda_1, \ldots, \lambda_n$ with $\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ the corresponding eigenvectors forming an orthonormal system. If matrix $\boldsymbol{A}$ is perturbed into*

$$\boldsymbol{A}' = \boldsymbol{A} + \delta \boldsymbol{A} \tag{68}$$

each eigenvalue λ_i is perturbed into

$$\lambda_i' = \lambda_i + (\boldsymbol{u}_i, \delta \boldsymbol{A} \boldsymbol{u}_i) + O(\delta \boldsymbol{A})^2. \tag{69}$$

If eigenvalue λ_i is a simple root, the corresponding eigenvector $\boldsymbol{u}_i$ is perturbed into

$$\boldsymbol{u}_i' = \boldsymbol{u}_i + \sum_{k \neq i} \frac{(\boldsymbol{u}_k, \delta \boldsymbol{A} \boldsymbol{u}_i)}{\lambda_i - \lambda_k} \boldsymbol{u}_k + O(\delta \boldsymbol{A})^2. \tag{70}$$

If the "quaternion representation" of 3-D rotation (Appendix A) is used, every rotation $\boldsymbol{R}$ is expressed in terms of a 4-D unit vector $\boldsymbol{q} = (q_0, q_1, q_2, q_3)^\top$ in the form of (66). If each q_κ is perturbed into

$$q_\kappa' = q_\kappa + \delta q_\kappa \tag{71}$$

the rotation $\boldsymbol{R}$ of (66) is perturbed by

$$\delta \boldsymbol{R} = \sum_{\kappa=0}^{3} \boldsymbol{D}_\kappa \delta q_\kappa \tag{72}$$

where $\boldsymbol{D}_\kappa = \partial \boldsymbol{R} / \partial q_\kappa$, $\kappa = 0, 1, 2, 3$, are given by

$$\boldsymbol{D}_0 = 2 \begin{pmatrix} q_0 & -q_3 & q_2 \\ q_3 & q_0 & -q_1 \\ -q_2 & q_1 & q_0 \end{pmatrix},$$

$$\boldsymbol{D}_1 = 2 \begin{pmatrix} q_1 & q_2 & q_3 \\ q_2 & -q_1 & -q_0 \\ q_3 & q_0 & -q_1 \end{pmatrix},$$

$$\boldsymbol{D}_2 = 2 \begin{pmatrix} -q_2 & q_1 & q_0 \\ q_1 & q_2 & q_3 \\ -q_0 & q_3 & -q_2 \end{pmatrix},$$

$$\boldsymbol{D}_3 = 2 \begin{pmatrix} -q_3 & -q_0 & q_1 \\ q_0 & -q_3 & q_2 \\ q_1 & q_2 & q_3 \end{pmatrix}. \tag{73}$$

This perturbation of $\boldsymbol{R}$ in turn causes a perturbation of $\boldsymbol{A}$ of (9) to a first approximation by

$$\begin{aligned} \delta \boldsymbol{A}(\boldsymbol{R}) &= \sum_{\alpha=1}^{n} W_\alpha \Big((\boldsymbol{m}_\alpha \times \delta \boldsymbol{R} \boldsymbol{m}_\alpha')(\boldsymbol{m}_\alpha \times \boldsymbol{R} \boldsymbol{m}_\alpha')^\top \\ &\quad + (\boldsymbol{m}_\alpha \times \boldsymbol{R} \boldsymbol{m}_\alpha')(\boldsymbol{m}_\alpha \times \delta \boldsymbol{R} \boldsymbol{m}_\alpha')^\top \Big) \\ &= \sum_{\kappa=1}^{3} \sum_{\alpha=1}^{n} W_\alpha \Big((\boldsymbol{m}_\alpha \times \boldsymbol{D}_\kappa \boldsymbol{m}_\alpha')(\boldsymbol{m}_\alpha \times \boldsymbol{R} \boldsymbol{m}_\alpha')^\top \\ &\quad + (\boldsymbol{m}_\alpha \times \boldsymbol{R} \boldsymbol{m}_\alpha')(\boldsymbol{m}_\alpha \times \boldsymbol{D}_\kappa \boldsymbol{R} \boldsymbol{m}_\alpha')^\top \Big). \end{aligned} \tag{74}$$

Hence, if $\boldsymbol{T}_\kappa$, $\kappa = 0, 1, 2, 3$, is defined by (11), the perturbation theorem implies that the smallest eigenvalue λ_m of $\boldsymbol{A}(\boldsymbol{R})$ is perturbed to a first approximation by

$$\delta \lambda_m = \sum_{\kappa=0}^{3} (\boldsymbol{h}, \boldsymbol{T}_\kappa \boldsymbol{h}) \delta q_\kappa \tag{75}$$

where $\boldsymbol{h}$ is the unit eigenvector of $\boldsymbol{A}(\boldsymbol{R})$ for the smallest eigenvalue λ_m. This means that the gradient of λ_m with respect to q_κ is given by (10).

Appendix C
Analytical Solution of 3-D Motion

For motion parameters $\{\boldsymbol{R}, \boldsymbol{h}\}$, define the *essential matrix* $\boldsymbol{G}$ by

$$\boldsymbol{G} = \boldsymbol{h} \times \boldsymbol{R}. \tag{76}$$

(See (25) for the definition of this multiplication operation.) It can easily been shown that $\|\boldsymbol{G}\| = \sqrt{2}$, where $\|\boldsymbol{G}\| = \sqrt{\sum_{i,j=1}^{3} {G_{ij}}^2}$. It is easy to see that Problem 2 is split into the following two subproblems:

$$\boldsymbol{R} = \begin{pmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1 q_2 - q_0 q_3) & 2(q_1 q_3 + q_0 q_2) \\ 2(q_2 q_1 + q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2 q_3 - q_0 q_1) \\ 2(q_3 q_1 - q_0 q_2) & 2(q_3 q_2 + q_0 q_1) & {}^2_0 - q_1^2 - q_2^2 + q_3^2 \end{pmatrix}. \tag{66}$$

$$\hat{\boldsymbol{K}} = \begin{pmatrix} K_{11} + K_{22} + K_{33} & K_{32} - K_{23} & K_{13} - K_{31} & K_{21} - K_{12} \\ K_{32} - K_{23} & K_{11} - K_{22} - K_{33} & K_{12} + K_{21} & K_{31} + K_{13} \\ K_{13} - K_{31} & K_{12} + K_{21} & -K_{11} + K_{22} - K_{33} & K_{23} + K_{32} \\ K_{21} - K_{12} & K_{31} + K_{13} & K_{23} + K_{32} & -K_{11} - K_{22} + K_{33} \end{pmatrix}. \tag{67}$$

Problem C.1: *Given unit vectors* $\{\boldsymbol{m}_\alpha\}$ *and* $\{\boldsymbol{m}'_\alpha\}$, $\alpha = 1,\dots,N$, *compute a matrix* $\boldsymbol{G}$ *(up to sign) such that*

$$\sum_{\alpha=1}^{N} W_\alpha(\boldsymbol{m}_\alpha, \boldsymbol{G}\boldsymbol{m}'_\alpha)^2 \to \min, \qquad \|\boldsymbol{G}\| = \sqrt{2}. \tag{77}$$

Problem C.2: *For a given matrix* $\boldsymbol{G}$, *compute a unit vector* $\boldsymbol{h}$ *and a rotation* $\boldsymbol{R}$ *such that*

$$\|\boldsymbol{G} - \boldsymbol{h}\times\boldsymbol{R}\|^2 \to \min. \tag{78}$$

A matrix $\boldsymbol{G}$ is said to be *decomposable* if it is expressed in the form of (76) for a unit vector $\boldsymbol{h}$ and a rotation matrix $\boldsymbol{R}$. The minimization (77) should be carried out under the constraint that $\boldsymbol{G}$ be decomposable. With a small compromise, we carry out the minimization (77) without the constraint of (76). The resulting solution is expected to be an approximation to the true solution, or at least as a good initial guess for the optimization search. The following summarizes the procedure of Weng *et al.* [34].

Problem C.1 is easy. If $m_{\alpha(i)}$ and $m'_{\alpha(i)}$ are the ith components of vectors $\boldsymbol{m}_\alpha$ and $\boldsymbol{m}'_\alpha$, respectively, the sum of squares of (77) is written in elements as

$$\sum_{\alpha=1}^{N} W_\alpha(\boldsymbol{m}_\alpha, \boldsymbol{G}\boldsymbol{m}'_\alpha)^2 = \sum_{i,j,k,l=1}^{3}\left(\sum_{\alpha=1}^{N} W_\alpha m_{\alpha(i)} m'_{\alpha(j)} m_{\alpha(k)} m'_{\alpha(l)}\right) G_{ij}G_{kl}. \tag{79}$$

If tensor $\mathcal{M} = (M_{ijkl})$ is defined by

$$M_{ijkl} = \sum_{\alpha=1}^{N} W_\alpha m_{\alpha(i)} m'_{\alpha(j)} m_{\alpha(k)} m'_{\alpha(l)} \tag{80}$$

the minimization (77) is written in the form

$$(\boldsymbol{G}, \mathcal{M}\boldsymbol{G}) = \sum_{i,j,k,l=1}^{3} M_{ijkl}G_{ij}G_{kl} \to \min, \ \sum_{i,j=1}^{3} {G_{ij}}^2 = 2. \tag{81}$$

Define a 9-D vector $\hat{\boldsymbol{G}} = (\hat{G}_\kappa)$ by renaming indices (ij) of G_{ij} as $\kappa = 3(i-1)+j : (11) \to 1, (12) \to 2, \dots, (33) \to 9$. Similarly, define a 9-D matrix $\hat{\boldsymbol{M}} = (\hat{M}_{\kappa\lambda})$ by renaming two pairs of indices (ij) and (kl) of M_{ijkl} as $(\kappa\lambda) = (3(i-1)+j, 3(k-1)+1)$: $(1111) \to (11), (1112) \to (12), \dots, (3333) \to (99)$. The above minimization now reads

$$(\hat{\boldsymbol{G}}, \hat{\boldsymbol{M}}\hat{\boldsymbol{G}}) = \sum_{\kappa,\lambda=1}^{9} \hat{M}_{\kappa\lambda}\hat{G}_\kappa\hat{G}_\lambda \to \min, \ \|\hat{\boldsymbol{G}}\|^2 = \sum_{\kappa=1}^{9} {\hat{A}_\kappa}^2 = 2. \tag{82}$$

The minimum is attained by the 9-D eigenvector of norm $\sqrt{2}$ of matrix $\hat{\boldsymbol{M}}$ for the smallest eigenvalue. The computed 9-D vector $\hat{\boldsymbol{G}} = (\hat{G}_\kappa)$ is then rearranged into a 3-D matrix $\boldsymbol{G} = (G_{ij})$ by renaming the index $\kappa : i = (\kappa-1)\mathrm{div}3+1$ and $j = (\kappa - 1) \pmod 3 + 1$, where "div" indicates the integer part of the quotient and "mod" the remainder.

Problem C.2 is easy to solve analytically if the strict minimum over the $\boldsymbol{h}$ and $\boldsymbol{R}$ is not sought for (see [18] for a rigorous treatment). Since $\boldsymbol{G}^\top\boldsymbol{h} = \boldsymbol{0}$ in the absence of noise, we determine $\boldsymbol{h}$ so that $\|\boldsymbol{G}^\top\boldsymbol{h}\|^2 = (\boldsymbol{h}, \boldsymbol{G}\boldsymbol{G}^\top\boldsymbol{h}) \to \min$. The solution is given by the unit eigenvector of $\boldsymbol{G}\boldsymbol{G}^\top$ for the smallest eigenvalue. Then, Problem C.2 reduces to

$$\mathrm{tr}(\boldsymbol{R}^T\boldsymbol{K}) \to \max \tag{83}$$

where

$$\boldsymbol{K} = -\boldsymbol{h}\times\boldsymbol{G}. \tag{84}$$

This problem is solved by the method of singular value decomposition (Theorem A.1), the method of polar decomposition (Theorem A.2), or the method of quaternion representation (Theorem A.3), as shown in Appendix A.

The above procedure determines the translation $\boldsymbol{h}$ only up to sign. The sign of $\boldsymbol{h}$ compatible with the (arbitrarily chosen) sign of $\boldsymbol{G}$ is easily determined from the following fact (we omit the proof).

Proposition C.1: *If* $\boldsymbol{h}$ *is not parallel to* $\boldsymbol{m}_\alpha$, *then*

$$|\boldsymbol{h}, \boldsymbol{m}_\alpha, \boldsymbol{G}\boldsymbol{m}'_\alpha| > 0. \tag{85}$$

The following is the result of Huang and Faugeras [10] and Faugeras and Maybank [7] (we omit the proof):

Proposition C.2: *A matrix* $\boldsymbol{G}$ *is decomposable if and only if its singular values are* 1, 1, and 0.

Corollary: *A matrix* $\boldsymbol{G}$ *is decomposable if and only if*

$$\det\boldsymbol{G} = 0, \qquad \|\boldsymbol{G}\| = \|\boldsymbol{G}\boldsymbol{G}^T\| = \sqrt{2}. \tag{86}$$

Corollary: *If matrix* $\boldsymbol{G}$ *is decomposable, it can be decomposed in exactly two ways. If* $\{\boldsymbol{R}, \boldsymbol{h}\}$ *and* $\{\boldsymbol{R}', \boldsymbol{u}'\}$ *are the two decompositions, then*

$$\boldsymbol{h}' = -\boldsymbol{h}, \qquad \boldsymbol{R}' = \boldsymbol{I}_{\boldsymbol{h}}\boldsymbol{R} \tag{87}$$

where $\boldsymbol{I}_{\boldsymbol{h}} = 2\boldsymbol{h}\boldsymbol{h}^T - \boldsymbol{I}$ is a half-rotation about $\boldsymbol{h}$.

Thus, the true and the spurious rotations form a *twisted pair* [22]. We observe the following:

- The number of unknowns for the essential matrix $\boldsymbol{G}$ is eight (since the scale is indeterminate); therefore, at least eight terms of the form $(\boldsymbol{m}_\alpha, \boldsymbol{G}\boldsymbol{m}'_\alpha)^2$ must be minimized to solve Problem C.1. This means that the number N of the feature points to be observed must be $N \geq 8$.
- The essential matrix $\boldsymbol{G}$ admits exactly two decompositions:

$$\boldsymbol{G} = \boldsymbol{h}\times\boldsymbol{R} = (-\boldsymbol{h})\times(\boldsymbol{I}_{\boldsymbol{h}}\boldsymbol{R}). \tag{88}$$

However, the sign of the essential matrix $\boldsymbol{G}$ computed by Problem C.1 is indeterminate, and matrix $-\boldsymbol{G}$ also admits two decompositions

$$-\boldsymbol{G} = (-\boldsymbol{h})\times\boldsymbol{R} = \boldsymbol{h}\times(\boldsymbol{I}_{\boldsymbol{h}}\boldsymbol{R}). \tag{89}$$

Hence, *four* decompositions are obtained.

- Proposition C.1 reduces these four solutions to *two* if $\{\boldsymbol{R}, \boldsymbol{u}\}$ is the true solution; the other is $\{\boldsymbol{R}, -\boldsymbol{h}\}$. Thus, the translation $\boldsymbol{h}$ is determined up to sign, and the rotation $\boldsymbol{R}$ is determined uniquely.
- Finally, the sign of $\boldsymbol{h}$ is determined so that the depths r_α and r'_α given by (6) are positive.

Thus, the motion parameters $\{\boldsymbol{R}, \boldsymbol{h}\}$ are uniquely determined. This algorithm works for $N \geq 8$ unless the eight feature points are in a special configuration (i.e., unless they are on a *critical surface*; see [18] for the details).

Appendix D
Covariance Matrices

Proposition D.1:

1. If $\boldsymbol{m}$ is a unit vector, then $\boldsymbol{m} + \Delta\boldsymbol{m}$ is a unit vector to a first approximation if and only if $\Delta\boldsymbol{m}$ is orthogonal to $\boldsymbol{m} : (\boldsymbol{m}, \Delta\boldsymbol{m}) = 0$.
2. The covariance matrix $V[\boldsymbol{m}]$ is symmetric and positive semi-definite.
3. The covariance matrix $V[\boldsymbol{m}]$ is singular with $\boldsymbol{m}$ as the unit eigenvector for eigenvalue $0 : V[\boldsymbol{m}]\boldsymbol{m} = \boldsymbol{0}$.
4. If σ_1^2, σ_2^2, and 0 ($\sigma_1 \geq \sigma_2 >$) are the eigenvalues of $V[\boldsymbol{m}]$ and if $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{m}\}$ is the corresponding orthonormal system of eigenvectors, the root mean square of the orthogonal projection of noise $\Delta\boldsymbol{m}$ onto orientation $\boldsymbol{l}$ (unit vector) takes its maximum when $\boldsymbol{l} = \boldsymbol{u}$ and its minimum when $\boldsymbol{l} = \boldsymbol{v}$, and the maximum and the minimum values are given, respectively, by σ_1 and σ_2.
5. The root-mean-square magnitude of $\Delta\boldsymbol{m}$ is $\sqrt{\mathrm{tr}V[\boldsymbol{m}]}$.

Proof:

1. The assertion is obvious from

$$\|\boldsymbol{m} + \Delta\boldsymbol{m}\|^2 = \|\boldsymbol{m}\|^2 + 2(\boldsymbol{m}, \Delta\boldsymbol{m}) + \|\Delta\boldsymbol{m}\|^2 = 1 + 2(\boldsymbol{m}, \Delta\boldsymbol{m}) + O(\Delta\boldsymbol{m})^2. \tag{90}$$

2. The covariance matrix $V[\boldsymbol{m}]$ is obviously symmetric:

$$V[\boldsymbol{m}]^\top = E[\Delta\boldsymbol{m}\Delta\boldsymbol{m}^\top]^\top = E[\Delta\boldsymbol{m}\Delta\boldsymbol{m}^\top] = V[\boldsymbol{m}]. \tag{91}$$

Let $\boldsymbol{a}$ be an arbitrary vector. Then

$$(\boldsymbol{a}, V[\boldsymbol{m}]\boldsymbol{a}) = (\boldsymbol{a}, E[\Delta\boldsymbol{m}\Delta\boldsymbol{m}^\top]\boldsymbol{a}) = E[(\boldsymbol{a}, \Delta\boldsymbol{m})^2] \geq 0 \tag{92}$$

meaning that $V[\boldsymbol{m}]$ is positive semi-definite.

3. Since $(\boldsymbol{m}, \Delta\boldsymbol{m}) = 0$, we see that

$$V[\boldsymbol{m}]\boldsymbol{m} = E[\Delta\boldsymbol{m}\Delta\boldsymbol{m}^\top]\boldsymbol{m} = E[(\boldsymbol{m}, \Delta\boldsymbol{m})\Delta\boldsymbol{m}] = 0. \tag{93}$$

4. If $\boldsymbol{l}$ is a unit vector, then

$$E[(\boldsymbol{l}, \Delta\boldsymbol{m})^2] = \boldsymbol{l}^\top E[\Delta\boldsymbol{m}\Delta\boldsymbol{m}^\top]\boldsymbol{l} = (\boldsymbol{l}, V[\boldsymbol{m}]\boldsymbol{l}). \tag{94}$$

This is maximized (or minimized) by the unit eigenvector for the largest (or smallest) eigenvalue of $V[\boldsymbol{m}]$. The maximum (or minimum) value is the largest (or smallest) eigenvalue.

5. From the definition of the covariance matrix $V[\boldsymbol{m}]$,

$$E[\|\Delta\boldsymbol{m}\|^2] = E[(\Delta\boldsymbol{m}, \Delta\boldsymbol{m})] = \mathrm{tr}E[\Delta\boldsymbol{m}\Delta\boldsymbol{m}^\top] = \mathrm{tr}V[\boldsymbol{m}]. \tag{95}$$

This proves the assertion. ∎

Appendix E
Effective Focal Length

Proposition E.1:

$$E[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}] = x\left(1 + \frac{\epsilon^2/f^2}{2(1 + r^2/f^2)} + O(\Delta\boldsymbol{m})^3\right),$$
$$E[f\frac{m_2 + \Delta m_2}{m_3 + \Delta m_3}] = y\left(1 + \frac{\epsilon^2/f^2}{2(1 + r^2/f^2)} + O(\Delta\boldsymbol{m})^3\right). \tag{96}$$

Proof: Since

$$f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3} = f\frac{m_1}{m_3}\left(\frac{1 + \Delta m_1/m_1}{1 + \Delta m_3/m_3}\right) = x\left(1 + \frac{\Delta m_1}{m_1} - \frac{\Delta m_3}{m_3} + (\frac{\Delta m_3}{m_3})^2 - (\frac{\Delta m_1}{m_1})(\frac{\Delta m_3}{m_3}) + O(\Delta\boldsymbol{m})^3\right) \tag{97}$$

its expectation is

$$E[f\frac{m_1 + \Delta m_1}{m_3 + \Delta m_3}] = x\left(1 + \frac{E[\Delta m_3{}^2]}{m_3{}^2} - \frac{E[\Delta m_1 \Delta m_3]}{m_1 m_3} + O(\Delta\boldsymbol{m})^3\right). \tag{98}$$

If we put $\boldsymbol{k} = (0, 0, 1)^\top$ and $\boldsymbol{i} = (1, 0, 0)^\top$, we have $m_3 = (\boldsymbol{m}, \boldsymbol{k})$ and $m_1 = (\boldsymbol{m}, \boldsymbol{i})$. Hence, $E[\Delta m_3{}^2] = (\boldsymbol{k}, V[\boldsymbol{m}]\boldsymbol{k})$ and $E[\Delta m_1 \Delta m_3] = (\boldsymbol{i}, V[\boldsymbol{m}]\boldsymbol{k})$. Since $\boldsymbol{u}$ is orthogonal to the Z axis, we have $(\boldsymbol{u}, \boldsymbol{k}) = 0$. From (14), we have

$$(\boldsymbol{k}, V[\boldsymbol{m}]\boldsymbol{k}) = \frac{\tilde{\epsilon}^2(\boldsymbol{v}, \boldsymbol{k})^2}{2(1 + r^2/f^2)^2},$$
$$(\boldsymbol{i}, V[\boldsymbol{m}]\boldsymbol{k}) = \frac{\tilde{\epsilon}^2(\boldsymbol{v}, \boldsymbol{i})(\boldsymbol{v}, \boldsymbol{k})}{2(1 + r^2/f^2)^2} \tag{99}$$

where we put $\tilde{\epsilon} = \epsilon/f$. In (15), the signs of $\boldsymbol{u}$ and $\boldsymbol{v}$ are irrelevant; therefore, we can choose $\boldsymbol{u} = \sqrt{1 + f^2/r^2}\boldsymbol{m} \times \boldsymbol{k}$ and $\boldsymbol{v} = \boldsymbol{u} \times \boldsymbol{m}$. From Fig. 2, we observe that

$$\cos\theta = m_3 = \frac{1}{\sqrt{1 + r^2/f^2}},$$
$$\sin\theta = \sqrt{1 - m_3{}^2} = \frac{r/f}{\sqrt{1 + r^2/f^2}}. \tag{100}$$

Hence

$$(\boldsymbol{v}, \boldsymbol{k}) = (\boldsymbol{u} \times \boldsymbol{m}, \boldsymbol{k}) = (\boldsymbol{m} \times \boldsymbol{k}, \boldsymbol{u}) = \sqrt{1 + \frac{f^2}{r^2}}(\boldsymbol{m} \times \boldsymbol{k}, \boldsymbol{m} \times \boldsymbol{k}) = \frac{\sqrt{1 + r^2/f^2}}{r/f}(1 - (\boldsymbol{m}, \boldsymbol{k})^2) = \frac{r/f}{\sqrt{1 + r^2/f^2}} = \frac{r}{f}m_3, \tag{101}$$

$$(\boldsymbol{v}, \boldsymbol{i}) = (\boldsymbol{u} \times \boldsymbol{m}, \boldsymbol{i}) = (\boldsymbol{m} \times \boldsymbol{i}, \boldsymbol{u}) = \sqrt{1 + \frac{f^2}{r^2}}(\boldsymbol{m} \times \boldsymbol{i}, \boldsymbol{m} \times \boldsymbol{k}) = -\frac{\sqrt{1 + r^2/f^2}}{r/f}(\boldsymbol{m}, \boldsymbol{k})(\boldsymbol{m}, \boldsymbol{i}) = -\frac{f}{r}m_1. \tag{102}$$

Thus

$$(\boldsymbol{k}, V[\boldsymbol{m}]\boldsymbol{k}) = \frac{\tilde{\epsilon}^2 r^2 {m_3}^2/f^2}{2(1+r^2/f^2)^2},$$
$$(\boldsymbol{i}, V[\boldsymbol{m}]\boldsymbol{k}) = -\frac{\tilde{\epsilon}^2 m_1 m_3}{2(1+r^2/f^2)^2}. \quad (103)$$

Hence

$$E[f\frac{m_1+\Delta m_1}{m_3+\Delta m_3}] = x(1+\frac{\tilde{\epsilon}^2/2}{1+r^2/f^2} + O(\Delta \boldsymbol{m})^3). \quad (104)$$

The expression of $E[f(m_2+\Delta m_2)/(m_3+\Delta m_3)]$ is obtained similarly. ■

Proposition E.2: *If*

$$\boldsymbol{m} = N[\begin{pmatrix} x \\ y \\ \hat{f} \end{pmatrix}], \qquad \hat{f} = f\left(1+\frac{\tilde{\epsilon}^2/2}{1+r^2/f^2}\right) \quad (105)$$

then

$$E[f\frac{m_1+\Delta m_1}{m_3+\Delta m_3}] = x + O(\Delta \boldsymbol{m})^3,$$
$$E[f\frac{m_2+\Delta m_2}{m_3+\Delta m_3}] = y + O(\Delta \boldsymbol{m})^3. \quad (106)$$

Proof: Proceeding in exactly the same way as in the proof of Proposition E.1, we obtain

$$E[f\frac{m_1+\Delta m_1}{m_3+\Delta m_3}] = f\frac{m_1}{m_3}\left(1+\frac{\tilde{\epsilon}^2/2}{1+r^2/f^2} + O(\Delta \boldsymbol{m})^3\right)$$
$$= \hat{f}\frac{m_1}{m_3}(\frac{f}{\hat{f}})\left(1+\frac{\tilde{\epsilon}^2/2}{1+r^2/f^2} + O(\Delta \boldsymbol{m})^3\right) = x + O(\Delta \boldsymbol{m})^3. \quad (107)$$

The expression for y is obtained similarly. To be strict, the expression $1+r^2/f^2$ in the above equation must be $1+r^2/\hat{f}^2$. However, replacing it by $1+r^2/f^2$ introduces only a difference of $O(\tilde{\epsilon}^4)$ in the final result. ■

Appendix E
Vector and Matrix Identities

Proposition F.1:

1. For vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ and a matrix $\boldsymbol{A}$,

$$(\boldsymbol{u} \times \boldsymbol{A}) \times \boldsymbol{v} = \boldsymbol{u} \times (\boldsymbol{A} \times \boldsymbol{v}). \quad (108)$$

2. For vectors $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{a}$, and $\boldsymbol{b}$,

$$\boldsymbol{u} \times (\boldsymbol{a}\boldsymbol{b}^\top) \times \boldsymbol{v} = (\boldsymbol{u} \times \boldsymbol{a})(\boldsymbol{b} \times \boldsymbol{v})^\top. \quad (109)$$

3. For unit vector $\boldsymbol{u}$ and the identity matrix $\boldsymbol{I}$,

$$\boldsymbol{u} \times \boldsymbol{I} \times \boldsymbol{u} = \boldsymbol{u}\boldsymbol{u}^\top - \boldsymbol{I}. \quad (110)$$

Proof:

1. We have

$$\boldsymbol{u} \times \boldsymbol{I} = \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}. \quad (111)$$

It is easy to see that the matrices $\boldsymbol{u} \times \boldsymbol{A}$ and $\boldsymbol{A} \times \boldsymbol{u}$ of (25) and (26) are written as

$$\boldsymbol{u} \times \boldsymbol{A} = (\boldsymbol{u} \times \boldsymbol{I})\boldsymbol{A}, \qquad \boldsymbol{A} \times \boldsymbol{v} = \boldsymbol{A}(\boldsymbol{v} \times \boldsymbol{I}). \quad (112)$$

Hence

$$(\boldsymbol{u} \times \boldsymbol{A}) \times \boldsymbol{v} = \boldsymbol{u} \times (\boldsymbol{A} \times \boldsymbol{v}) = (\boldsymbol{u} \times \boldsymbol{I})\boldsymbol{A}(\boldsymbol{v} \times \boldsymbol{I}). \quad (113)$$

2. We see that $(\boldsymbol{v} \times \boldsymbol{I})^\top = -(\boldsymbol{v} \times \boldsymbol{I})$. Hence

$$\begin{aligned} &\boldsymbol{u} \times (\boldsymbol{a}\boldsymbol{b}^\top) \times \boldsymbol{v} \\ &= (\boldsymbol{u} \times \boldsymbol{I})\boldsymbol{a}\boldsymbol{b}^\top(\boldsymbol{v} \times \boldsymbol{I}) \\ &= ((\boldsymbol{u} \times \boldsymbol{I})\boldsymbol{a})((\boldsymbol{v} \times \boldsymbol{I})^\top\boldsymbol{b})^\top \\ &= -((\boldsymbol{u} \times \boldsymbol{I})\boldsymbol{a})((\boldsymbol{v} \times \boldsymbol{I})\boldsymbol{b})^\top \\ &= -(\boldsymbol{u} \times \boldsymbol{a})(\boldsymbol{v} \times \boldsymbol{b})^\top = (\boldsymbol{u} \times \boldsymbol{a})(\boldsymbol{b} \times \boldsymbol{v})^\top. \end{aligned} \quad (114)$$

3. We see that

$$\boldsymbol{u} \times \boldsymbol{I} \times \boldsymbol{u} = (\boldsymbol{u} \times \boldsymbol{I})\boldsymbol{I}(\boldsymbol{u} \times \boldsymbol{I}) = (\boldsymbol{u} \times \boldsymbol{I})^2. \quad (115)$$

It is easy to confirm that $(\boldsymbol{u} \times \boldsymbol{I})^2 = \boldsymbol{u}\boldsymbol{u}^\top - \boldsymbol{I}$. ■

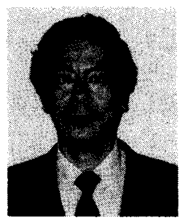
Acknowledgment

The author thanks M. Brady, A. Blake, and A. Zisserman of the University of Oxford, S. Maybank of the GEC Hirst Research Center (U.K.), and A. Rosenfeld, L. Davis, Y. Aloimonos, and R. Chellappa of the University of Maryland for various discussions. He would also like to thank T. Morijiri of Gunma University for doing numerical experiments.

References

[1] J. Aisbett, "An iterated estimation of the motion parameters of a rigid body from noisy displacement vectors," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 12, pp. 1092–1098, 1990.
[2] K. S. Arun, T. S. Huang, and S. D. Blostein, "Least-squares fitting of two 3-D point sets," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-9, pp. 698–700, 1987.
[3] T. J. Broida and R. Chellappa, "Estimation of object motion parameters from noisy images," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-8, pp. 90–99, 1986.
[4] ——, "Performance bounds for estimating three-dimensional motion parameters from a sequence of noisy images," *J. Opt. Soc. Amer.*, vol. A-6, pp. 879–998, 1989.
[5] ——, "Estimating the kinematics and structure of a rigid object from a sequence of monocular images," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, pp. 497–513, 1991.
[6] T. J. Broida, S. Chandrashekhar, and R. Chellappa, "Recursive estimation of 3-D motion from a monocular image sequence," *IEEE Trans. Aero. Electr. Sys.*, vol. 26, pp. 639–656, 1990.
[7] O. D. Faugeras and S. Maybank, "Motion from point matches: Multiplicity of solutions," *Int. J. Comput. Vision*, vol. 4, pp. 225–246, 1990.
[8] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *J. Opt. Soc. Amer.*, vol. A-4, pp. 629–642, 1987.
[9] B. K. P. Horn, H. M. Hilden, and S. Negahdaripour, "Closed-form solution of absolute orientation using orthonormal matrices," *J. Opt. Soc. Amer.*, vol. A-5, pp. 1128–1135, 1988.
[10] T. S. Huang and O. D. Faugeras, "Some properties of the E matrix in two-view motion estimation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, pp. 1310–1312, 1989.

[11] C. Jerian and R. Jain, "Polynomial methods for structure from motion," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 12, pp. 1150–1165, 1990.
[12] ——, "Structure from motion—A critical analysis of methods," *IEEE Trans. Sys. Man Cybern.*, vol. 21, pp. 572–588, 1991.
[13] K. Kanatani, "Camera rotation invariance of image characteristics," *Comput. Vision Graphics Image Process.*, vol. 39, pp. 328–354, 1987.
[14] ——, "Constraints on length and angle," *Comput. Vision Graphics Image Process.*, vol. 41, pp. 28–42, 1988.
[15] ——, "Transformation of optical flow by camera rotation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 10, pp. 131–143, 1988.
[16] ——, *Group-Theoretical Methods in Image Understanding*. Berlin: Springer, 1990.
[17] ——, "Computational projective geometry," *CVGIP: Image Understanding*, vol. 54, pp. 333–448, 1991.
[18] ——, *Geometric Computation for Machine Vision*, to be published by Oxford University Press.
[19] C. -H. Lee, "Time-varying images: The effect of finite resolution on uniqueness," *CVGIP: Image Understanding*, vol. 54, pp. 325–332, 1991.
[20] S. Lee and Y. Kay, "A Kalman filter approach for accurate 3-D motion estimation from sequence of stereo images," *CVGIP: Image Understanding*, vol. 54, pp. 244–258, 1991.
[21] H. C. Longuet-Higgins, "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol. 293, no. 10, pp. 133–135, 1981.
[22] ——, "Multiple interpretations of a pair of images of a surface," *Proc. Roy. Soc. Lond.*, vol. A-418, pp. 1–15, 1988.
[23] L. Matthies, T. Kanade, and R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *Int. J. Comput. Vision*, vol. 3, pp. 209–236, 1989.
[24] H. -H. Nagel, "Representation of moving rigid objects based on visual observations," *Comput.*, vol. 14, no. 8, pp. 29–39, 1981.
[25] A. N. Netravali, T. S. Huang, A. S. Krishnakumar, and R. J. Holt, "Algebraic methods in 3-D motion estimation from two-view point correspondences," *Int. J. Imaging Syst. Tech.*, vol. 1, pp. 78–99, 1989.
[26] J. Philip, "Estimation of three-dimensional motion of rigid objects from noisy observations," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, pp. 61–66, 1991.
[27] J. Porrill, "Fitting ellipses and predicting confidence envelopes using a bias corrected Kalman filter," *Image Vision Comput.*, vol. 8, pp. 37–41, 1990.
[28] J. W. Roach and J. K. Aggarwal, "Determining the movement of objects from a sequence of images," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 2, pp. 544–562, 1980.
[29] M. E. Spetsakis and Y. Aloimonos, "Optimal computing of structure from motion using point correspondences in two frames," in *Proc. Second Int. Conf. Comput. Vision* (Tampa, FL), Dec. 1988, pp. 449–453.
[30] ——, "A multi-frame approach to visual motion perception," *Int. J. Comput. Vision*, vol. 6, pp. 245–255, 1991.
[31] R. Y. Tsai and T. S. Huang, "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. PAMI-6, pp. 13–27, 1984.
[32] S. Ullman, *The Interpretation of Visual Motion*. cambridge, MA: MIT Press, 1979.
[33] S. Umeyama, "Least-squares estimation of transformation parameters between two point patterns," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 13, pp. 376–380, 1991.
[34] J. Weng, T. S. Huang, and N. Ahuja, "Motion and structure from two perspective views: Algorithms, error analysis, and error estimation," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, pp. 451–467, 1989.
[35] G. -S. J. Young and R. Chellappa, "3-D motion estimation using a sequence of noisy stereo images: Models, estimation, and uniqueness results," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 12, pp. 735–759, 1990.
[36] X. Zhuang, "A simplification to linear two-view motion algorithm," *Comput. Vision Graphics Image Process.*, vol. 46, pp. 175–178, 1989.
[37] X. Zhuang, T. S. Huang, and R. M. Haralick, "Two-view motion analysis: A unified algorithm," *J. Opt. Soc. Amer.*, vol. A-3, pp. 1492–1500, 1986.



Kenichi Kanatani was born in Okayama, Japan, in 1947 and received the B.Eng., M.Eng., and Ph.D. degrees in applied mathematics from the University of Tokyo, Japan, in 1972, 1974, and 1979, respectively.

In 1979, he joined Gunma University, Japan, where he is now Professor of Computer Science. He studied physics at Case Western Reserve University, Cleveland, OH, from 1969 to 1970. He was Visiting Researcher at the University of Maryland, College Park, from 1985 to 1986, at the University of Copenhagen, Denmark, in 1988, and at the University of Oxford, England, in 1991.

Dr. Kanatani is the author of *Group-Theoretical Methods in Image Understanding* (Berlin: Springer, 1990).