PAPER Special Issue on Computer Vision

Renormalization for Motion Analysis: Statistically Optimal Algorithm

Kenichi KANATANI†, *Member*

SUMMARY Introducing a general statistical model of image noise, we present an optimal algorithm for computing 3-D motion from two views without involving numerical search: (i) the *essential matrix* is computed by a scheme called *renormalization*; (ii) the *decomposability condition* is optimally imposed on it so that it exactly decomposes into motion parameters; (iii) image feature points are optimally corrected so that they define their 3-D depths. Our scheme not only produces a statistically optimal solution but also evaluates the reliability of the computed motion parameters and reconstructed points in quantitative terms.

key words: *error analysis, motion analysis, optimal estimation, renormalization, 3-D reconstruction*

1. Introduction

The study of 3-D rigid motion estimation from images, known as *structure from motion*, was initiated by Ullman [9], and an analytical solution for eight feature points was independently given by Longuet-Higgins [6] and Tsai and Huang [8]. Their algorithms are constructed on the assumption that all data are exact. Various types of statistical optimization for noisy data have been proposed by Weng et al. [10], [11]. Also, many other approaches have been proposed, each emphasizing different aspects [2], [4], [7].

In this paper, a general model of image noise is introduced, and a statistically optimal criterion is established for motion estimation. First, the statistical behavior of the theoretically optimal solution is analyzed, providing a theoretical bound on accuracy. Then, we construct a scheme for computing the optimal solution as closely as possible without involving numerical search: (i) the *essential matrix* is computed by a scheme called *renormalization*; (ii) the *decomposability condition* is optimally imposed on it so that it decomposes into motion parameters; (iii) image feature points are optimally corrected so that they define their 3-D depths.

Our scheme not only produces a statistically optimal solution but also quantitatively evaluates the reliability of the computed motion parameters and reconstructed points.

Manuscript received March 28, 1994.
Manuscript revised June 11, 1994.

†The author is with the Department of Computer Science, Gunma University, Kiryu-shi, 376 Japan.

2. Statistical Model of Image Noise and Epipolar Equation

The camera is associated with an XYZ coordinate system with origin O at the center of the lens and Z-axis along the optical axis. The plane $Z = 1$ is identified with the image plane, on which an xy image coordinate system is defined around the Z-axis such that the x- and y-axes are parallel to the X- and Y-axes, respectively. A point on the image plane with image coordinates (x, y) is represented by its *position vector* $\boldsymbol{x} = (x, y, 1)^\top$, where the superscript $\top$ denotes transpose.

Let $\bar{\boldsymbol{x}}$ be the position vector on the image plane when there is no noise. In the presence of noise, it is perturbed into $\boldsymbol{x} = \bar{\boldsymbol{x}} + \Delta\boldsymbol{x}$. The noise $\Delta\boldsymbol{x}$ is regarded as a random variable, and its *covariance matrix* is defined by

$$V[\boldsymbol{x}] = E[\Delta\boldsymbol{x}\Delta\boldsymbol{x}^\top], \tag{1}$$

where the symbol $E[\cdot]$ denotes expectation. This matrix is singular and in general has rank 2. Noise is assumed to occur at each point on the image plane independently, but its distribution can be different from point to point.

Suppose two cameras are positioned in the scene in such a way that the position of the second camera is obtained by translating the first camera by $\boldsymbol{h}$ and rotating it around the center of the lens by $\boldsymbol{R}$: we call $\{\boldsymbol{h}, \boldsymbol{R}\}$ the *motion parameters*.

Let O and O' be, respectively, the origins of the first and second camera systems. If p and p' are corresponding image points, they can be projections of a feature point P in the scene if and only if vectors $\vec{OP}$, $\vec{OO'}$, and $\vec{O'P}$ are coplanar (Fig. 1):

$$|\vec{Op}, \vec{OO'}, \vec{O'p'}| = 0. \tag{2}$$

In this paper, $(\boldsymbol{a}, \boldsymbol{b})$ denotes the inner product of vectors

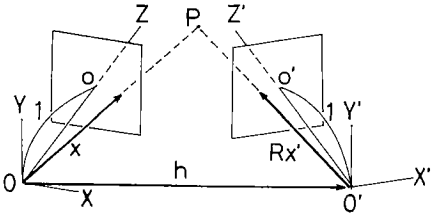


Fig. 1 Epipolar geometry.

$\boldsymbol{a}$ and $\boldsymbol{b}$, and $|\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}|$ $(= (\boldsymbol{a} \times \boldsymbol{b}, \boldsymbol{c}))$ denotes the scalar triple product of vectors $\boldsymbol{a}$, $\boldsymbol{b}$, and $\boldsymbol{c}$.

Let $\bar{\boldsymbol{x}}$ and $\bar{\boldsymbol{x}}'$ be the position vectors of p and p' on their respective image planes in the absence of noise. Since the second camera is rotated by $\boldsymbol{R}$ relative to the first camera, the position vector $\bar{\boldsymbol{x}}'$ with respect to the second camera is $\boldsymbol{R}\bar{\boldsymbol{x}}'$ with respect to the first camera. Since $\vec{OO}' = \boldsymbol{h}$, Eq. (2) is rewritten as

$$|\bar{\boldsymbol{x}}, \boldsymbol{h}, \boldsymbol{R}\bar{\boldsymbol{x}}'| = 0. \tag{3}$$

For a fixed $\bar{\boldsymbol{x}}'$, this equation describes a line, called the *epipolar* of p', passing through p on the first image plane, while for a fixed $\bar{\boldsymbol{x}}'$ it describes a line, called the *epipolar* of p, passing through p' on the second image plane. For this reason, Eq. (3) is called the *epipolar equation*. The problem is to determine $\{\boldsymbol{h}, \boldsymbol{R}\}$ that satisfy Eq. (3) from multiple corresponding data points $\{\boldsymbol{x}_\alpha\}$, $\{\boldsymbol{x}'_\alpha\}$, $\alpha = 1, ..., N$, in the presence of noise. However, the scale is indeterminate for the translation $\boldsymbol{h}$, meaning that a small camera motion near a small object is indistinguishable from a large camera motion far apart from a large object. Since whether $\boldsymbol{h} = \boldsymbol{0}$ or not is easily judged by a robust computation [4], we henceforth assume that $\boldsymbol{h} \neq \boldsymbol{0}$ and adopt the scaling $\|\boldsymbol{h}\| = 1$.

3. Theoretically Optimal Solution

Since $\{\boldsymbol{x}_\alpha, \boldsymbol{x}'_\alpha\}$, $\alpha = 1, ..., N$, do not necessarily satisfy Eq. (3) for common $\boldsymbol{h}$ and $\boldsymbol{R}$, we correct points $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ to $\hat{\boldsymbol{x}}_\alpha = \boldsymbol{x}_\alpha + \Delta\boldsymbol{x}_\alpha$ and $\hat{\boldsymbol{x}}'_\alpha = \boldsymbol{x}_\alpha + \Delta\boldsymbol{x}'_\alpha$ in such a way that Eq. (3) is satisfied for *some* $\boldsymbol{h}$ and $\boldsymbol{R}$. There exist infinitely many ways to do this. From among them, we choose the one that is statistically optimal. We measure the optimality by the *Mahalanobis* metric: we choose such $\Delta\boldsymbol{x}_\alpha$ and $\Delta\boldsymbol{x}'_\alpha$, $\alpha = 1, ..., N$, that

$$\sum_{\alpha=1}^{N} (\Delta\boldsymbol{x}_\alpha, V[\boldsymbol{x}_\alpha]^- \Delta\boldsymbol{x}_\alpha) + \sum_{\alpha=1}^{N} (\Delta\boldsymbol{x}'_\alpha, V[\boldsymbol{x}'_\alpha]^- \Delta\boldsymbol{x}'_\alpha) \to \min, \tag{4}$$

where $V[\boldsymbol{x}_\alpha]^-$ and $V[\boldsymbol{x}'_\alpha]^-$ are the generalized inverses of the covariance matrices $V[\boldsymbol{x}_\alpha]$ and $V[\boldsymbol{x}'_\alpha]$, respectively. Introducing a Lagrange multiplier and minimizing (4) with respect to $\Delta\boldsymbol{x}_\alpha$ and $\Delta\boldsymbol{x}'_\alpha$, we find that the problem reduces to the following optimization [5]:

$$J[\boldsymbol{h}, \boldsymbol{R}] = \frac{1}{N} \sum_{\alpha=1}^{N} W_\alpha(\boldsymbol{h}, \boldsymbol{R}) |\boldsymbol{x}_\alpha, \boldsymbol{h}, \boldsymbol{R}\boldsymbol{x}'_\alpha|^2 \to \min, \tag{5}$$

$$\begin{aligned} W_\alpha(\boldsymbol{h}, \boldsymbol{R}) = 1/((&\boldsymbol{h} \times \boldsymbol{R}\boldsymbol{x}'_\alpha, V[\boldsymbol{x}_\alpha](\boldsymbol{h} \times \boldsymbol{R}\boldsymbol{x}'_\alpha)) \\ &+ ((\boldsymbol{h} \times \boldsymbol{R})^\top \boldsymbol{x}_\alpha, V[\boldsymbol{x}'_\alpha](\boldsymbol{h} \times \boldsymbol{R})^\top \boldsymbol{x}_\alpha) \\ &+ (V[\boldsymbol{x}_\alpha](\boldsymbol{h} \times \boldsymbol{R}); \boldsymbol{h} \times \boldsymbol{R} V[\boldsymbol{x}'_\alpha])) \end{aligned} \tag{6}$$

In this paper, the product $\boldsymbol{u} \times \boldsymbol{A}$ of a vector $\boldsymbol{u}$ and a matrix $\boldsymbol{A}$ is defined by column-wise vector product. Weng et al. [11] recently proposed an optimization criterion which involves the depths Z_α and Z'_α, but their criterion should be theoretically equivalent to the above one.

Let $\{\hat{\boldsymbol{h}}, \hat{\boldsymbol{R}}\}$ be the optimal estimates, and $\{\bar{\boldsymbol{h}}, \bar{\boldsymbol{R}}\}$ be their true values. Let $\hat{\boldsymbol{h}} = \bar{\boldsymbol{h}} + \Delta\boldsymbol{h}$ and $\hat{\boldsymbol{R}} = \bar{\boldsymbol{R}} + \Delta\boldsymbol{R}$. The transformation from $\bar{\boldsymbol{R}}$ to $\hat{\boldsymbol{R}}$ is a small rotation. Hence, there exists a small vector $\Delta\boldsymbol{\Omega}$ such that $\hat{\boldsymbol{R}} = \bar{\boldsymbol{R}} + \Delta\boldsymbol{\Omega} \times \bar{\boldsymbol{R}}$ to a first approximation. The covariance matrices of $\{\hat{\boldsymbol{h}}, \hat{\boldsymbol{R}}\}$ are defined by

$$V[\hat{\boldsymbol{h}}] = E[\Delta\boldsymbol{h}\Delta\boldsymbol{h}^\top], \quad V[\hat{\boldsymbol{h}}, \hat{\boldsymbol{R}}] = E[\Delta\boldsymbol{h}\Delta\boldsymbol{\Omega}^\top],$$

$$V[\hat{\boldsymbol{R}}, \hat{\boldsymbol{h}}] = E[\Delta\boldsymbol{\Omega}\Delta\boldsymbol{h}^\top], \quad V[\hat{\boldsymbol{R}}] = E[\Delta\boldsymbol{\Omega}\Delta\boldsymbol{\Omega}^\top]. \tag{7}$$

Expanding $J[\boldsymbol{h}, \boldsymbol{R}]$ up to the second order in the neighborhood of $\bar{\boldsymbol{h}}$ and $\bar{\boldsymbol{R}}$ and analyzing the statistical behavior of the solution, we obtain the covariance matrices of $\hat{\boldsymbol{h}}$ and $\hat{\boldsymbol{R}}$ in the following form [5]:

$$\begin{pmatrix} V[\hat{\boldsymbol{h}}] & V[\hat{\boldsymbol{h}}, \hat{\boldsymbol{R}}] \\ V[\hat{\boldsymbol{R}}, \hat{\boldsymbol{h}}] & V[\hat{\boldsymbol{R}}] \end{pmatrix} = \begin{pmatrix} \sum_{\alpha=1}^{N} \bar{W}_\alpha \bar{\boldsymbol{a}}_\alpha \bar{\boldsymbol{a}}_\alpha^\top & \sum_{\alpha=1}^{N} \bar{W}_\alpha \bar{\boldsymbol{a}}_\alpha \bar{\boldsymbol{b}}_\alpha^\top \\ \sum_{\alpha=1}^{N} \bar{W}_\alpha \bar{\boldsymbol{b}}_\alpha \bar{\boldsymbol{a}}_\alpha^\top & \sum_{\alpha=1}^{N} \bar{W}_\alpha \bar{\boldsymbol{b}}_\alpha \bar{\boldsymbol{b}}_\alpha^\top \end{pmatrix}, \tag{8}$$

$$\bar{\boldsymbol{a}}_\alpha = \bar{\boldsymbol{x}}_\alpha \times \bar{\boldsymbol{R}}\bar{\boldsymbol{x}}'_\alpha,$$

$$\bar{\boldsymbol{b}}_\alpha = (\bar{\boldsymbol{x}}_\alpha, \bar{\boldsymbol{R}}\bar{\boldsymbol{x}}_{\alpha'})\bar{\boldsymbol{h}} - (\bar{\boldsymbol{h}}, \bar{\boldsymbol{R}}\bar{\boldsymbol{x}}'_\alpha)\bar{\boldsymbol{x}}_\alpha. \tag{9}$$

Here, $\bar{W}_\alpha$ is the value obtained by replacing $\boldsymbol{x}_\alpha$, $\boldsymbol{x}'_\alpha$, $\boldsymbol{h}$ and $\boldsymbol{R}$ by $\bar{\boldsymbol{x}}_\alpha$, $\bar{\boldsymbol{x}}'_\alpha$, $\bar{\boldsymbol{h}}$ and $\bar{\boldsymbol{R}}$, respectively, in Eq. (6). Eq. (8) gives an attainable bound on the accuracy of the computed motion parameters $\{\boldsymbol{h}, \boldsymbol{R}\}$ in statistical terms.

4. Renormalization

The optimization (5) is nonlinear, requiring numerical search. Here, we derive a semi-analytical scheme for computing the optimal solution as closely as possible without involving numerical search. Let ϵ be an appropriately defined "average magnitude" of image noise, and decompose the covariance matrices $V[\boldsymbol{x}_\alpha]$ and $V[\boldsymbol{x}'_\alpha]$ into the following form:

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha], \quad V[\boldsymbol{x}'_\alpha] = \epsilon^2 V_0[\boldsymbol{x}'_\alpha]. \tag{10}$$

We call ϵ the *noise level*, and $V_0[\boldsymbol{x}_\alpha]$ and $V_0[\boldsymbol{x}'_\alpha]$ the *normalized covariance matrices*. In many practical problems, the form of the covariance matrix of image noise is easy to predict (e.g., isotropic), while its absolute magnitude is difficult to estimate *a priori*. Hence, assuming that the normalized covariance matrices $V_0[\boldsymbol{x}_\alpha]$ and $V_0[\boldsymbol{x}'_\alpha]$ are known but the noise level ϵ is unknown, we construct a scheme for estimating ϵ *a posteriori*.

Define the *essential matrix*

$$\boldsymbol{G} = \boldsymbol{h} \times \boldsymbol{R}. \tag{11}$$

The constraint $\|\boldsymbol{h}\| = 1$ is equivalent to $\|\boldsymbol{G}\| = \sqrt{2}$ [3], where the matrix norm is defined by $\|\boldsymbol{G}\| = \sqrt{\sum_{i,j=1}^{3} G_{ij}{}^2}$. Then, the function $J[\boldsymbol{h}, \boldsymbol{R}]$ given in (5) can be regarded as a function of $\boldsymbol{G}$. Our strategy consists of two stages: (i) computing $\boldsymbol{G}$ that minimizes $J[\boldsymbol{h}, \boldsymbol{R}]$ as closely as possible; (ii) decomposing $\boldsymbol{G}$ into $\{\boldsymbol{h}, \boldsymbol{R}\}$ as closely as possible. This procedure is known as the *linearized algorithm* [3], [6], [8], [10] (for the direct optimization approach, see [2], [4], [7], [11]).

The essential matrix is optimally computed by the following procedure, which we call *renormalization* [5]:

1. Let $c = 0$ and $W_\alpha = 1, \alpha = 1, ..., N$.
2. Compute the *moment tensor* $\mathcal{M} = (M_{ijkl})$ and tensors $\mathcal{N}^{(1)} = (N^{(1)}_{ijkl})$ and $\mathcal{N}^{(2)} = (N^{(2)}_{ijkl})$ by

$$M_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha x_{\alpha(i)} x'_{\alpha(j)} x_{\alpha(k)} x'_{\alpha(l)}, \tag{12}$$

$$N^{(1)}_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha \Big(V_0[\boldsymbol{x}_\alpha]_{ik} x'_{\alpha(j)} x'_{\alpha(l)} + V_0[\boldsymbol{x}'_\alpha]_{jl} x_{\alpha(i)} x_{\alpha(k)} \Big), \tag{13}$$

$$N^{(2)}_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha V_0[\boldsymbol{x}_\alpha]_{ik} V_0[\boldsymbol{x}'_\alpha]_{jl}. \tag{14}$$

3. Compute the smallest 'eigenvalue' λ of the *unbiased moment tensor*

$$\hat{\mathcal{M}} = \mathcal{M} - c\mathcal{N}^{(1)} + c^2\mathcal{N}^{(2)}, \tag{15}$$

and the corresponding 'eigenmatrix' $\boldsymbol{G}$ of norm $\sqrt{2}$.

4. If $\lambda \approx 0$, return $\boldsymbol{G}$, c, and $\hat{\mathcal{M}}$. Else, update the constant c and the weights W_α as follows:

$$D = \Big((\boldsymbol{G}; \mathcal{N}^{(1)}\boldsymbol{G}) - 2c(\boldsymbol{G}; \mathcal{N}^{(2)}\boldsymbol{G})\Big)^2 - 8\lambda(\boldsymbol{G}; \mathcal{N}^{(2)}\boldsymbol{G}), \tag{16}$$

$$c \leftarrow c + \frac{(\boldsymbol{G}; \mathcal{N}^{(1)}\boldsymbol{G}) - 2c(\boldsymbol{G}; \mathcal{N}^{(2)}\boldsymbol{G}) - \sqrt{D}}{2(\boldsymbol{G}; \mathcal{N}^{(2)}\boldsymbol{G})}, \tag{17}$$

$$W_\alpha \leftarrow 1/((\boldsymbol{G}\boldsymbol{x}'_\alpha, V_0[\boldsymbol{x}_\alpha]\boldsymbol{G}\boldsymbol{x}'_\alpha) + (\boldsymbol{G}^\top\boldsymbol{x}_\alpha, V_0[\boldsymbol{x}'_\alpha]\boldsymbol{G}^\top\boldsymbol{x}_\alpha) + c(V_0[\boldsymbol{x}_\alpha]\boldsymbol{G}; \boldsymbol{G}V_0[\boldsymbol{x}'_\alpha])). \tag{18}$$

If $D < 0$, Eq. (17) is replaced by $c \leftarrow c + 2\lambda/(\boldsymbol{G}; \mathcal{N}^{(1)}\boldsymbol{G})$.

5. Go back to Step 2.

Here, $x_{\alpha(i)}$ and $x'_{\alpha(i)}$ denote the ith components of vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, respectively. The inner product of matrices is defined by $(\boldsymbol{A}; \boldsymbol{B}) = \sum_{i,j=1}^{3} A_{ij}B_{ij}$. The product $\mathcal{T}\boldsymbol{A}$ of tensor $\mathcal{T} = (T_{ijkl})$ and matrix $\boldsymbol{A} = (A_{ij})$ is a matrix whose (ij) element is $\sum_{k,l=1}^{3} T_{ijkl}A_{kl}$. A matrix $\boldsymbol{A}$ is said to be an *eigenmatrix* of tensor $\mathcal{T}$ for *eigenvalue* λ if $\mathcal{T}\boldsymbol{A} = \lambda\boldsymbol{A}$. The eigenmatrices and eigenvalues of a tensor are easily computed by rearranging a tensor into a nine-dimensional matrix and solving the usual eigenvalue problem [5].

The motivation that underlies the above procedure is as follows. Let $\bar{\mathcal{M}}$ be the tensor obtained by replacing $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ by $\bar{\boldsymbol{x}}_\alpha$ and $\bar{\boldsymbol{x}}'_\alpha$, respectively, in Eq. (12). The true value of $\boldsymbol{G}$ is the eigenmatrix of $\bar{\mathcal{M}}$ for eigenvalue 0. If the unbiased moment tensor $\hat{\mathcal{M}}$ is defined by Eq. (15), we have $E[\hat{\mathcal{M}}] = \bar{\mathcal{M}}$ for $c = \epsilon^2$. Renormalization finds a value of c such that $\hat{\mathcal{M}}$ has eigenvalue 0. At the same time, the weights W_α are optimally chosen.

After renormalization, an unbiased estimate of the squared noise level ϵ^2 is obtained in the form

$$\hat{\epsilon}^2 = \frac{c}{1 - 8/N}. \tag{19}$$

Its mean and variance are given by

$$E[\hat{\epsilon}^2] = \epsilon^2, \qquad V[\hat{\epsilon}^2] = \frac{2\epsilon^4}{N - 8}. \tag{20}$$

This result is obtained from the fact that Nc/ϵ^2 is a χ^2-variable of $N-8$ degrees of freedom under the Gaussian assumption [5].

The *covariance tensor* $\mathcal{V}[\boldsymbol{G}]$ of the computed essential matrix $\boldsymbol{G} = \bar{\boldsymbol{G}} + \Delta\boldsymbol{G}$ ($\bar{\boldsymbol{G}} = \bar{\boldsymbol{h}} \times \bar{\boldsymbol{R}}$) is defined by

$$\mathcal{V}[\boldsymbol{G}] = E[\Delta\boldsymbol{G} \otimes \Delta\boldsymbol{G}]. \tag{21}$$

Its $(ijkl)$ element is $E[\Delta G_{ij}\Delta G_{kl}]$. If we write $\mathcal{V}[\boldsymbol{G}] = \epsilon^2\mathcal{V}_0[\boldsymbol{G}]$, the *normalized covariance tensor* $\mathcal{V}_0[\boldsymbol{G}]$ is obtained from the unbiased moment tensor $\hat{\mathcal{M}}$ resulting from renormalization in the following form [5]:

$$\mathcal{V}_0[\boldsymbol{G}] = \frac{1}{N}\hat{\mathcal{M}}^-, \tag{22}$$

Here, $\hat{\mathcal{M}}^-$ is the generalized inverse of tensor $\hat{\mathcal{M}}$, which is computed by rearranging tensor $\hat{\mathcal{M}}$ into a six-dimensional symmetric matrix and computing its generalized inverse [5]. The unbiased moment tensor $\hat{\mathcal{M}}$ resulting from renormalization has rank 8, so $\mathcal{V}_0[\boldsymbol{G}]$ also has rank 8.

5. Optimal Correction

A matrix $\boldsymbol{G}$ is said to be *decomposable* if there exist a unit vector $\boldsymbol{h}$ and a rotation matrix $\boldsymbol{R}$ such that $\boldsymbol{G} = \boldsymbol{h} \times \boldsymbol{R}$. A matrix $\boldsymbol{G}$ is decomposable if and only if

$$\det \boldsymbol{G} = 0, \qquad \|\boldsymbol{G}\| = \|\boldsymbol{G}\boldsymbol{G}^\top\| = \sqrt{2}, \tag{23}$$

which is equivalent to saying that the singular values of $\boldsymbol{G}$ are 1, 1, and 0 [1],[3].

Since the essential matrix $\boldsymbol{G}$ computed by renormalization may not be exactly decomposable, we next corrected it optimally into $\boldsymbol{G}' = \boldsymbol{G} + \Delta\boldsymbol{G}$ so that Eqs. (23) are satisfied. We measure the optimality by the Mahalanobis metric:

$$(\Delta\boldsymbol{G}; \mathcal{V}_0[\boldsymbol{G}]^- \Delta\boldsymbol{G}) \to \min. \tag{24}$$

Introducing Lagrange multipliers, we obtain the following first order solution [5]:

$$\Delta\boldsymbol{G} = \lambda_1 \mathcal{V}_0[\boldsymbol{G}]\boldsymbol{G}^{\dagger\top} + \lambda_2 \mathcal{V}_0[\boldsymbol{G}](\boldsymbol{G}\boldsymbol{G}^\top \boldsymbol{G}). \tag{25}$$

Here, $\boldsymbol{G}^\dagger$ is the cofactor matrix of $\boldsymbol{G}$, and λ_1 and λ_2 are given by

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = \boldsymbol{W} \begin{pmatrix} -\det \boldsymbol{G} \\ (2 - \|\boldsymbol{G}\boldsymbol{G}^\top\|^2)/4 \end{pmatrix}, \tag{26}$$

$$\boldsymbol{W} = \begin{pmatrix} (\boldsymbol{G}^{\dagger\top}; \mathcal{V}_0[\boldsymbol{G}]\boldsymbol{G}^{\dagger\top}) \\ (\boldsymbol{G}\boldsymbol{G}^\top \boldsymbol{G}; \mathcal{V}_0[\boldsymbol{G}]\boldsymbol{G}^{\dagger\top}) \\ (\boldsymbol{G}^{\dagger\top}; \mathcal{V}_0[\boldsymbol{G}](\boldsymbol{G}\boldsymbol{G}^\top \boldsymbol{G})) \\ (\boldsymbol{G}\boldsymbol{G}^\top \boldsymbol{G}; \mathcal{V}_0[\boldsymbol{G}](\boldsymbol{G}\boldsymbol{G}^\top \boldsymbol{G})) \end{pmatrix}^{-1}. \tag{27}$$

This correction is iterated until the decomposability condition (23) is sufficiently satisfied.

In general, three equations $\det \boldsymbol{G} = c_1$, $\|\boldsymbol{G}\| = c_2$, and $\|\boldsymbol{G}\boldsymbol{G}^\top\| = c_3$ constrain $\boldsymbol{G}$ to be in a *six*-dimensional manifold in the nine-dimensional parameter space for $\boldsymbol{G}$. However, $c_1 = 0$, $c_2 = \sqrt{2}$, and $c_3 = \sqrt{2}$ are *critical values*, at which the six-dimensional manifold *degenerates* into five dimensions, admitting only *five* degrees of freedom to $\boldsymbol{G}$. This degeneracy lowers the speed of convergence of the correction.

A realistic computation should be $\boldsymbol{G}' = \sqrt{2}N[\boldsymbol{G} + \gamma\Delta\boldsymbol{G}]$ $(0 < \gamma < 1)$, where $N[\cdot]$ denotes normalization to unit norm. At each step, the normalized covariance tensor $\mathcal{V}_0[\boldsymbol{G}]$ is also updated, since its null space should be compatible with $\boldsymbol{G}$ as it changes. It is projected in the form

$$\mathcal{V}_0[\boldsymbol{G}]'_{ijkl} = \sum_{m,n,p,q=1}^{3} P_{ijmn} P_{klpq} \mathcal{V}_0[\boldsymbol{G}]_{mnpq}, \tag{28}$$

where

$$P_{ijkl} = \delta_{ik}\delta_{jl} - \frac{1}{2} G_{ij} G_{kl}, \tag{29}$$

and δ_{ij} is the Kronecker delta.

6. Decomposition into Motion Parameters

The decomposition of $\boldsymbol{G}$ into $\{\boldsymbol{h}, \boldsymbol{R}\}$ is done as follows [3],[4]:

1. Let $\boldsymbol{h}$ be the unit eigenvector of matrix $\boldsymbol{G}\boldsymbol{G}^\top$ for the smallest eigenvalue.
2. Adjust the sign of $\boldsymbol{h}$ so that $\sum_{\alpha=1}^{N} |\boldsymbol{h}, \boldsymbol{x}_\alpha, \boldsymbol{G}\boldsymbol{x}'_\alpha| > 0$.
3. Compute the matrix $\boldsymbol{K} = -\boldsymbol{h} \times \boldsymbol{G}$, and let $\boldsymbol{K} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{U}^\top$ be its *singular value decomposition* [3].
4. Compute $\boldsymbol{R} = \boldsymbol{V}\mathrm{diag}(1, 1, \det(\boldsymbol{V}\boldsymbol{U}^\top))\boldsymbol{U}^\top$.

Here, the symbol $\mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$ denotes the diagonal matrix whose diagonal elements are λ_1, λ_2, and λ_3 in that order.

If $\boldsymbol{G}$ is decomposable, simpler methods exist for this decomposition [3], but they do not produce an exact unit vector $\boldsymbol{h}$ and an exact rotation matrix $\boldsymbol{R}$ if $\boldsymbol{G}$ is not strictly decomposable. In contrast, the above procedure yields an exact unit vector $\boldsymbol{h}$ and an exact rotation matrix $\boldsymbol{R}$ in an optimal way whether $\boldsymbol{G}$ is decomposable or not [3]. Hence, the iterations for correcting $\boldsymbol{G}$ can be stopped at any time if accuracy is not required so much (the crudest decomposition could be obtained by skipping the correction altogether).

The *a posteriori* covariance matrices $V[\boldsymbol{h}]$, $V[\boldsymbol{h}, \boldsymbol{R}]$, and $V[\boldsymbol{R}]$ of the actually computed motion parameters $\{\boldsymbol{h}, \boldsymbol{R}\}$ can be evaluated by decomposing the covariance tensor $\mathcal{V}[\boldsymbol{G}]$ obtained after the modification (28). Since $\mathcal{V}[\boldsymbol{G}]$ has rank 8, it is first projected onto the five-dimensional subspace admitted to the perturbation $\Delta\boldsymbol{G}$ of $\boldsymbol{G}$ [5]. It is then decomposed into $V[\boldsymbol{h}]$, $V[\boldsymbol{h}, \boldsymbol{R}]$, and $V[\boldsymbol{R}]$. Let $\boldsymbol{g}_i$ be the ith column of $\boldsymbol{G}$. The covariance matrix $V[\boldsymbol{h}]$ is given by

$$V[\boldsymbol{h}] = \sum_{k,l=1}^{3} (\boldsymbol{h}, V[\boldsymbol{g}_k, \boldsymbol{g}_l]\boldsymbol{h}) \boldsymbol{g}_k \boldsymbol{g}_l^\top, \tag{30}$$

where matrix $V[\boldsymbol{g}_k, \boldsymbol{g}_l]$ is defined by

$$V[\boldsymbol{g}_k, \boldsymbol{g}_l]_{ij} = E[\Delta\boldsymbol{g}_k \boldsymbol{g}_l^\top] = \mathcal{V}[\boldsymbol{G}]_{ikjl}. \tag{31}$$

The covariance matrix $V[\boldsymbol{R}]$ is given by

$$V[\boldsymbol{R}] = \sum_{i=1}^{3} \left(\frac{\mathrm{tr} V[\boldsymbol{r}_i]}{2} \boldsymbol{I} - V[\boldsymbol{r}_i] \right), \tag{32}$$

where tr means trace, $\boldsymbol{I}$ is the unit matrix, and $\boldsymbol{r}_i = \boldsymbol{g}_i \times \boldsymbol{h} + \boldsymbol{g}_{i+1} \times \boldsymbol{g}_{i+2}$ (the index is added modulo 3). The matrix $V[\boldsymbol{h}, \boldsymbol{R}]$ is computed similarly. However, the covariance matrices thus computed are often overestimation, probably due to various approximations involved. According to our numerical experiment (see Sect. 9), the error behavior is characterized sufficiently well by the a priori covariance matrices computed from Eq. (8) by substituting the computed estimates for their true values. The error is approximately proportional to the noise level ϵ and the inverse square root of the number of corresponding pairs $1/\sqrt{N}$.

7. Computation of 3-D Positions

Even if $\{\boldsymbol{h}, \boldsymbol{R}\}$ are optimally estimated, the 3-D po-

sition of each feature point cannot be defined unless Eq. (3) is satisfied. Hence, each corresponding pair $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ is optimally corrected into $\hat{\boldsymbol{x}}_\alpha = \boldsymbol{x}_\alpha + \Delta\boldsymbol{x}_\alpha$ and $\hat{\boldsymbol{x}}'_\alpha = \boldsymbol{x}'_\alpha + \Delta\boldsymbol{x}'_\alpha$ so that Eq. (3) is exactly satisfied. We measure the optimality by the Mahalanobis metric:

$$J_\alpha = (\Delta\boldsymbol{x}_\alpha, V_0[\boldsymbol{x}_\alpha]^-\Delta\boldsymbol{x}_\alpha) + (\Delta\boldsymbol{x}'_\alpha, V_0[\boldsymbol{x}'_\alpha]^-\Delta\boldsymbol{x}'_\alpha) \to \min. \tag{33}$$

Introducing a Lagrange multiplier, we obtain the following first order solution [5]:

$$\begin{aligned}\Delta\boldsymbol{x}_\alpha &= -(\boldsymbol{x}_\alpha, \boldsymbol{G}\boldsymbol{x}'_\alpha)V_0[\boldsymbol{x}_\alpha]\boldsymbol{G}\boldsymbol{x}'_\alpha/V_\alpha,\\ \Delta\boldsymbol{x}'_\alpha &= -(\boldsymbol{x}_\alpha, \boldsymbol{G}\boldsymbol{x}'_\alpha)V_0[\boldsymbol{x}'_\alpha]\boldsymbol{G}^\top\boldsymbol{x}_\alpha/V_\alpha,\end{aligned} \tag{34}$$

$$V_\alpha = (\boldsymbol{x}'_\alpha, \boldsymbol{G}^\top V_0[\boldsymbol{x}_\alpha]\boldsymbol{G}\boldsymbol{x}'_\alpha) + (\boldsymbol{x}_\alpha, \boldsymbol{G}V_0[\boldsymbol{x}'_\alpha]\boldsymbol{G}^\top\boldsymbol{x}_\alpha). \tag{35}$$

In real computation, the corrections $\boldsymbol{x}_\alpha \leftarrow \boldsymbol{x}_\alpha + \Delta\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha \leftarrow \boldsymbol{x}'_\alpha + \Delta\boldsymbol{x}'_\alpha$ are iterated until Eq. (3) is sufficiently satisfied.

The residual of the optimization (33) is given by

$$\hat{J}_\alpha = (\hat{\boldsymbol{x}}_\alpha, \boldsymbol{G}\hat{\boldsymbol{x}}'_\alpha)^2/\hat{V}_\alpha, \tag{36}$$

where $\hat{V}_\alpha$ is the value obtained by replacing $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ by $\hat{\boldsymbol{x}}_\alpha$ and $\hat{\boldsymbol{x}}'_\alpha$, respectively, in Eq. (35). It can be shown that $\hat{J}_\alpha/\epsilon^2$ is a χ^2-variable with one degree of freedom under the Gaussian assumption [5]. Hence, $\hat{J}_\alpha$ is an unbiased estimator of the squared noise level ϵ for "individual" feature points. Its mean and variance are given by $E[\hat{J}_\alpha] = \epsilon^2$ and $V[\hat{J}_\alpha] = 2\epsilon^4$, respectively. Thus, the individual covariance matrices are estimated in the form $V[\boldsymbol{x}_\alpha] = E[\hat{J}_\alpha]V_0[\boldsymbol{x}_\alpha]$ and $V[\boldsymbol{x}'_\alpha] = E[\hat{J}_\alpha]V_0[\boldsymbol{x}'_\alpha]$.

Let $\boldsymbol{r}_\alpha$ and $\boldsymbol{r}'_\alpha$ be the 3-D positions of a feature point P_α with respect to the first and the second camera coordinate systems, respectively. The 3-D locations $\boldsymbol{r}_\alpha$ and $\boldsymbol{r}'_\alpha$ have the form

$$\boldsymbol{r}_\alpha = Z_\alpha\hat{\boldsymbol{x}}_\alpha, \qquad \boldsymbol{r}'_\alpha = Z'_\alpha\hat{\boldsymbol{x}}'_\alpha, \tag{37}$$

where the depths Z_α and Z'_α are given as follows:

$$\begin{aligned}Z_\alpha &= \frac{(\boldsymbol{h}\times\boldsymbol{R}\hat{\boldsymbol{x}}'_\alpha, \hat{\boldsymbol{x}}_\alpha\times\boldsymbol{R}\hat{\boldsymbol{x}}'_\alpha)}{\|\hat{\boldsymbol{x}}_\alpha\times\boldsymbol{R}\hat{\boldsymbol{x}}'_\alpha\|^2},\\ Z'_\alpha &= \frac{(\boldsymbol{h}\times\hat{\boldsymbol{x}}_\alpha, \hat{\boldsymbol{x}}_\alpha\times\boldsymbol{R}\hat{\boldsymbol{x}}'_\alpha)}{\|\hat{\boldsymbol{x}}_\alpha\times\boldsymbol{R}\hat{\boldsymbol{x}}'_\alpha\|^2}.\end{aligned} \tag{38}$$

Two solutions exist, since renormalization computes the essential matrix $\boldsymbol{G}$ up to sign. One solution gives $\{\boldsymbol{h}, \boldsymbol{R}\}$, Z_α, and Z'_α; the other gives $\{-\boldsymbol{h}, \boldsymbol{R}\}$, $-Z_\alpha$, and $-Z'_\alpha$ [3], [4]. The correct solution is chosen by imposing the condition

$$\sum_{\alpha=1}^{N}(\mathrm{sgn}(Z_\alpha) + \mathrm{sgn}(Z'_\alpha)) > 0. \tag{39}$$

It appears that we could alternatively impose $\sum_{\alpha=1}^{N}(Z_\alpha + Z'_\alpha) > 0$. However, this is dangerous because there exists a possibility that the depth of a feature point located far apart in front of the camera ($Z_\alpha \approx \infty$) can be computed, due to image noise, to be far apart *behind* the camera ($Z_\alpha \approx -\infty$), disrupting the judgment.

8. Reliability of 3-D Reconstruction

Once the 3-D position $\boldsymbol{r}_\alpha$ is reconstructed, its covariance matrix $V[\boldsymbol{r}_\alpha]$ is also computed. Here, two sources of error must be considered: (i) image noise; (ii) errors in $\{\boldsymbol{h}, \boldsymbol{R}\}$. Strictly speaking, errors in $\{\boldsymbol{h}, \boldsymbol{R}\}$ are correlated with errors in $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, since the motion parameters $\{\boldsymbol{h}, \boldsymbol{R}\}$ are computed from the feature points $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, $\alpha = 1, ..., N$. However, if we focus on an "individual" feature point, image noise can be regarded as approximately independent of errors in $\{\boldsymbol{h}, \boldsymbol{R}\}$ provided the number of feature points is large. Hence, the two sources of error can be analyzed separately.

If $\{\boldsymbol{h}, \boldsymbol{R}\}$ are correct, the covariance matrix of the position $\hat{\boldsymbol{x}}_\alpha = \boldsymbol{x}_\alpha + \Delta\boldsymbol{x}_\alpha$ corrected by Eqs. (34) is given as follows [5]:

$$V[\hat{\boldsymbol{x}}_\alpha] = V[\boldsymbol{x}_\alpha] - (V[\boldsymbol{x}_\alpha]\boldsymbol{G}\hat{\boldsymbol{x}}'_\alpha)(V[\boldsymbol{x}_\alpha]\boldsymbol{G}\hat{\boldsymbol{x}}'_\alpha)^\top/\hat{V}_\alpha. \tag{40}$$

The covariance matrices $V[\hat{\boldsymbol{x}}'_\alpha]$ and $V[\hat{\boldsymbol{x}}_\alpha, \hat{\boldsymbol{x}}'_\alpha]$ are computed similarly. Then, the covariance matrix of the position $\boldsymbol{r}_\alpha$ reconstructed by Eqs. (37) is given in the following form:

$$\begin{aligned}V[\boldsymbol{r}_\alpha] = {} & Z_\alpha^2 V[\hat{\boldsymbol{x}}_\alpha] + Z_\alpha(V[\hat{\boldsymbol{x}}_\alpha, Z_\alpha]\hat{\boldsymbol{x}}_\alpha^\top\\ & + \hat{\boldsymbol{x}}_\alpha V[\hat{\boldsymbol{x}}_\alpha, Z_\alpha]^\top) + V[Z_\alpha]\hat{\boldsymbol{x}}_\alpha\hat{\boldsymbol{x}}_\alpha^\top.\end{aligned} \tag{41}$$

The variance $V[Z]$ and the correlation $V[\hat{\boldsymbol{x}}_\alpha, Z]$ are easily computed from $V[\hat{\boldsymbol{x}}_\alpha]$, $V[\hat{\boldsymbol{x}}'_\alpha]$, and $V[\hat{\boldsymbol{x}}_\alpha, \hat{\boldsymbol{x}}'_\alpha]$ [5].

On the other hand, suppose the feature points $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ are accurate. If $\{\boldsymbol{h}, \boldsymbol{R}\}$ are perturbed, the true positions $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ are corrected into $\hat{\boldsymbol{x}}_\alpha$ and $\hat{\boldsymbol{x}}'_\alpha$, respectively, by Eqs. (34). The covariance matrix of the corrected position $\hat{\boldsymbol{x}}_\alpha$ is given by

$$V[\hat{\boldsymbol{x}}_\alpha] = -V[E_\alpha](V_0[\boldsymbol{x}_\alpha]\boldsymbol{G}\boldsymbol{x}'_\alpha)(V_0[\boldsymbol{x}_\alpha]\boldsymbol{G}\boldsymbol{x}'_\alpha)^\top/\hat{V}_\alpha, \tag{42}$$

where $E_\alpha = (\boldsymbol{x}_\alpha, \boldsymbol{G}\boldsymbol{x}'_\alpha)$. We obtain

$$\begin{aligned}V[E_\alpha] = {} & (\boldsymbol{a}_\alpha, V[\boldsymbol{h}]\boldsymbol{a}_\alpha) + 2(\boldsymbol{a}_\alpha, V[\boldsymbol{h}, \boldsymbol{R}]\boldsymbol{b}_\alpha)\\ & + (\boldsymbol{b}_\alpha, V[\boldsymbol{R}]\boldsymbol{b}_\alpha),\end{aligned} \tag{43}$$

where $\boldsymbol{a}_\alpha$ and $\boldsymbol{b}_\alpha$ are the values obtained by removing all the bars in Eqs. (9) [5]. The covariance matrices $V[\hat{\boldsymbol{x}}'_\alpha]$ and $V[\hat{\boldsymbol{x}}_\alpha, \hat{\boldsymbol{x}}'_\alpha]$ are computed similarly. The covariance matrix $V[\boldsymbol{r}_\alpha]$ is given by Eq. (41), and the variance $V[Z]$ and the correlation $V[\hat{\boldsymbol{x}}_\alpha, Z]$ are easily computed from $V[\hat{\boldsymbol{x}}_\alpha]$, $V[\hat{\boldsymbol{x}}'_\alpha]$, and $V[\hat{\boldsymbol{x}}_\alpha, \hat{\boldsymbol{x}}'_\alpha]$ [5].

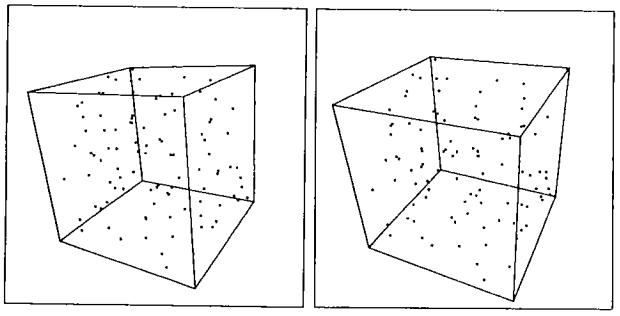


Fig. 2 Simulated images of feature points in the scene.

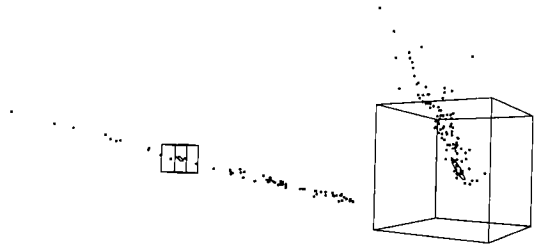


Fig. 3 Errors in motion parameters estimated by optimal least-squares. (a) Translation. (b) Rotation.

9. Numerical Example

Figure 2 shows simulated 512×512-pixel images of one hundred feature points randomly scattered in the scene (the surrounding box is for visual effect only). The focal length is set to 600 (pixels). To the x- and y-coordinates of each feature point in both frames is added independent Gaussian noise of mean 0 and standard deviation 1 (pixel). Hence, the covariance matrix of each point is $V[\boldsymbol{x}_\alpha] = V[\boldsymbol{x}'_\alpha] = \epsilon^2 \mathrm{diag}(1,1,0)$ and $\epsilon = 1/600$. In 3-D reconstruction, the value of ϵ is regarded as unknown.

In Figs. 3, 4, and 5, the translation error $\Delta\boldsymbol{h} = \boldsymbol{P}_{\bar{h}}(\boldsymbol{h} - \bar{\boldsymbol{h}})$ and the rotation error $\Delta\Omega\boldsymbol{l}$ are plotted in three-dimensions for one hundred trials, using different noise each time, where $\boldsymbol{P}_{\bar{h}} = \boldsymbol{I} - \boldsymbol{h}\boldsymbol{h}^\top$. Here, $\Delta\Omega$ and $\boldsymbol{l}$ (unit vector) are, respectively, the angle and axis of the relative rotation $\boldsymbol{R}\bar{\boldsymbol{R}}^{-1}$. The ellipse that indicates the theoretical standard deviation of $\Delta\boldsymbol{h}$ in each orientation orthogonal to $\bar{\boldsymbol{h}}$ is shown in Figs. 3 (a), 4(a), and 5(a); it is computed from the covariance matrix $V[\boldsymbol{h}]$ given by Eqs. (8). Similarly, the ellipsoid that indicates the theoretical standard deviation of $\Delta\Omega$ in each orientation $\boldsymbol{l}$ is shown in Figs. 3 (b), 4(b), and 5(b); it is computed from the covariance matrix $V[\boldsymbol{R}]$ given by Eqs. (8).

Figure 3 is for the optimally weighted least-squares method [4], [10], [11]; Fig. 4 is obtained by applying renormalization, and Fig. 5 is obtained by also adding the optimal correction. The scale of each figure is different, but the ellipses, ellipsoids, and the boxes shown there have the same absolute sizes. It is observed that the error behavior in Fig. 5 is well characterized by the theoretical bound, meaning that nearly optimal estimates are obtained.

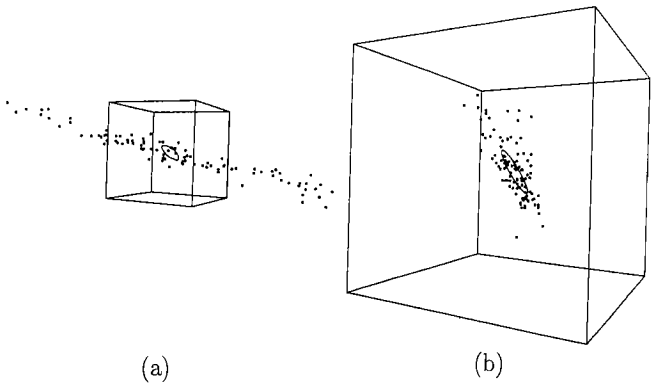


Fig. 4 Errors in motion parameters estimated by renormalization. (a) Translation. (b) Rotation.

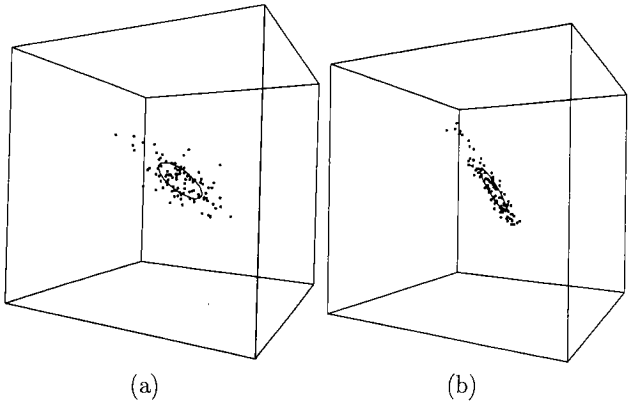


Fig. 5 Errors in motion parameters estimated by renormalization and optimal correction of the essential matrix. (a) Translation. (b) Rotation.

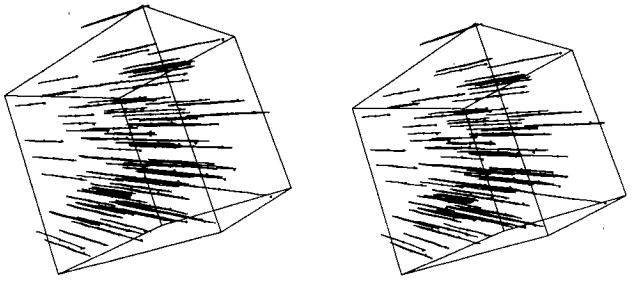


Fig. 6 Reliability of reconstructed points (stereogram).

Figure 6 is a stereogram of the ellipsoids that indicate the standard deviations of the errors in the reconstructed positions in each orientation; they are computed from the covariance matrices $V[\boldsymbol{r}_\alpha]$ obtained by the procedure in the preceding section. The true positions are indicated by dots. In the figure, all the ellipsoids are extremely elongated in the depth directions, showing that uncertainty occurs almost always in the depth direction.

10. Concluding Remarks

In this paper, we have introduced a general statistical model of image noise and presented an optimal scheme for computing 3-D motion from two views. We have also obtained a theoretical bound on accuracy. In order to avoid numerical search, we proposed a scheme called "renormalization" and an optimal correction of the essential matrix. Our method is characterized as the "most refined linearized algorithm". We also presented an optimal procedure for 3-D reconstruction and computed its reliability in quantitative terms.

The emphasis of this paper is the pursuit of *theoretical optimality*, so computational efficiency is not considered very much. Our procedure consists of several optimization stages, and a more efficient algorithm could be obtained by skipping some altogether and replacing some by cruder but more efficient approximations. The following are some examples:

- The effect of adding $c^2\mathcal{N}^{(2)}$ in Eq. (15) is very small. So, Eq. (15) could be replaced by $\hat{\mathcal{M}} = \mathcal{M} - c\mathcal{N}^{(1)}$; Eqs. (16) and (17) could be simply $c \leftarrow c + 2\lambda/(\boldsymbol{G}; \mathcal{N}^{(1)})$.
- The accuracy of the solution is not very sensitive to the weights W_α, so Eq. (18) could be skipped by fixing $W_\alpha = 1, \alpha = 1, ..., N$.
- As mentioned in Sect. 6, the computation of eigenvectors and singular value decompositions for decomposing $\boldsymbol{G}$ into $\{\boldsymbol{h}, \boldsymbol{R}\}$ could be replaced by the computation of simpler explicit algebraic expressions. The result is the same if decomposability is sufficiently imposed on $\boldsymbol{G}$.
- The optimal correction given by Eqs. (34) could be skipped by defining the lines of sight separately for $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ and estimating their intersection by least squares. The difference in accuracy is very small.
- Each of the iterative procedures could be stopped after one or two iterations. Further improvements are very small.

Put differently, the following is crucial:

- The renormalization procedure for $\boldsymbol{G}$ should be iterated at least once or twice to remove statistical bias.
- The optimal correction of $\boldsymbol{G}$ should be iterated at least once or twice to impose the decomposability condition.

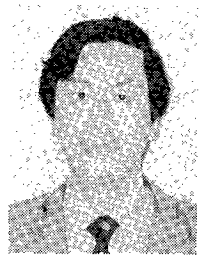
Evaluation of the tradeoff between accuracy and time requires a *bench-mark algorithm* that attains the highest accuracy without regard to efficiency, against which measures for efficiency are evaluated. Also, the accuracy of the bench-mark solution should be evaluated in quantitative terms. To present such an algorithm is the main purpose of this paper.

Acknowledgments

The author thanks Sachio Takeda and Toru Kawashima of Gunma University for conducting numerical simulations and implementing a 3-D representation system for this research.

References

[1] Huang, T.S. and Faugeras, O.D., "Some properties of the E matrix in two-view motion estimation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.11, pp.1310–1312, 1989.
[2] Jerian, C. and Jain, R., "Polynomial methods for structure from motion," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.12, pp.1150–1165, 1990.
[3] Kanatani, K., *Geometric Computation for Machine Vision*, Oxford Univ. Press, Oxford, 1993.
[4] Kanatani, K., "Unbiased estimation and statistical analysis of 3-D rigid motion from two views," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.15, no.1, pp.37–50, 1993.
[5] Kanatani, K., *Statistical Optimization for Geometric Computation*, Lecture Note, Department of Computer Science, Gunma University, Apr. 1994.
[6] Longuet-Higgins, H.C., "A computer algorithm for reconstructing a scene from two projections," *Nature*, vol.293, no.10, pp.133–135, 1981.
[7] Netravali, A.N., Huang, T.S., Krishnakumar, A.S. and Holt, R.J., "Algebraic methods in 3-D motion estimation from two-view point correspondences," *Int. J. Imaging Sys. Tech.*, vol.1, pp.78–99, 1989.
[8] Tsai, R.Y. and Huang, T.S., "Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.6, pp.13–27, 1984.
[9] Ullman, S., *The Interpretation of Visual Motion*, MIT Press, Cambridge, MA, 1979.
[10] Weng, J., Huang, T.S. and Ahuja, N., "Motion and structure from two perspective views: Algorithms, error analysis, and error estimation," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol.11, pp.451–467, 1989.
[11] Weng, J., Ahuja, N. and Huang, T.S., "Optimal motion and structure estimation," *IEEE Trans. Pat. Anal. Mach. Intell.*, vol.15, pp.864–884, 1993.



Kenichi Kanatani received his Ph.D. in applied mathematics from the University of Tokyo in 1979. He is currently Professor of computer science at Gunma University. He is the author of *Group-Theoretical Methods in Image Understanding* (Springer, 1990) and *Geometric Computation for Machine Vision* (Oxford Univ. Press, 1993).