6th Symposium on Sensing via Image Information (SII 2000), June 2000, Yokokahma, Japan.

Optimal Fundamental Matrix Computation: Algorithm and Reliability Analysis

Kenichi Kanatani
Department of Computer Science, Gunma University, Kiryu, Gunma 376-8515 Japan

Abstract This paper presents an optimal linear algorithm for computing the fundamental matrix from corresponding points over two images under an assumed model, which admits independent Gaussian noise that is not necessarily isotropic or homogeneous. We derive a theoretical bound and demonstrate by experiments that our algorithm delivers results in the vicinity of the bound. Simulated and real-image examples are shown to observe the performance of our algorithm and illustrate the reliability evaluation process.

1. Introduction

Computing the fundamental matrix from corresponding points over two images is one of the basic elements of computer vision. Already, many algorithms have been presented for this purpose [1, 3, 6, 7, 13, 18, 22, 23, 24, 25]. These are roughly classified into two approaches: the *bundle-adjustment* and the *linear algorithm*.

The bundle-adjustment is based on maximal likelihood estimation after reprojection: we assume all the parameter values, predict where the image data should ideally be observed, and determine the parameters by minimizing the Mahalanobis distance between the predicted and observed data positions. If the noise is Gaussian, the solution is known to satisfy the *Cramer-Rao lower bound* (*CRLB*).

The linear algorithm is based on the observation that the epipolar constraint is linear in the fundamental matrix. Hence, the solution can be computed in a single step by solving an eigenvalue problem. However, the resulting solution is known to have a large statistical bias [9].

In this paper, we demonstrate that the linear algorithm *can* achieve the accuracy equivalent to that of the bundle-adjustment if appropriate modifications are incorporated. We can also give a theoretical accuracy bound to our linear algorithm. We demonstrate by experiments that our algorithm indeed delivers results in the vicinity of that bound.

2. Fundamental Matrix

Let $\{(x_\alpha, y_\alpha)\}$ and $\{(x'_\alpha, y'_\alpha)\}$, $\alpha = 1, ..., N$, be image coordinates (in pixels) of two sets of N points on two different images; the image coordinate system is defined arbitrarily for each camera. Define vectors

$$\boldsymbol{x}_\alpha = \begin{pmatrix} x_\alpha/f_0 \\ y_\alpha/f_0 \\ 1 \end{pmatrix}, \quad \boldsymbol{x}'_\alpha = \begin{pmatrix} x'_\alpha/f_0 \\ y'_\alpha/f_0 \\ 1 \end{pmatrix}, \tag{1}$$

where f_0 is a scale factor[1] (in pixels) chosen so that x_α/f_0, y_α/f_0, x'_α/f_0, and y'_α/f_0 have an order 1. The two sets of points are said to satisfy the *epipolar constraint* if there exists a matrix $\boldsymbol{F}$ of determinant 0 such that

$$(\boldsymbol{x}_\alpha, \boldsymbol{F}\boldsymbol{x}'_\alpha) = 0, \quad \alpha = 1, ..., N, \tag{2}$$

where and throughout this paper we denote by $(\boldsymbol{a}, \boldsymbol{b})$ the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. The matrix $\boldsymbol{F}$ is called the *fundamental matrix*. Eq. (2) implies that the absolute scale of $\boldsymbol{F}$ is indeterminate, so we adopt the normalization $\|\boldsymbol{F}\| = 1$, where the norm of a matrix $\boldsymbol{A} = (A_{ij})$ is defined by $\|\boldsymbol{A}\| = \sqrt{\sum_{i,j=1}^{3} A_{ij}^2}$.

3. Statistical Model of Noise

We view the two sets of points $\{(x_\alpha, y_\alpha)\}$ and $\{(x'_\alpha, y'_\alpha)\}$ as perturbed from the points $\{(\bar{x}_\alpha, \bar{y}_\alpha)\}$ and $\{(\bar{x}'_\alpha, \bar{y}'_\alpha)\}$ that satisfy the epipolar constraint exactly. Let $\bar{\boldsymbol{x}}_\alpha$ and $\bar{\boldsymbol{x}}'_\alpha$ be the vector representations of the true positions $(\bar{x}_\alpha, \bar{y}_\alpha)$ and $(\bar{x}'_\alpha, \bar{y}'_\alpha)$, respectively. We write

$$\boldsymbol{x}_\alpha = \bar{\boldsymbol{x}}_\alpha + \Delta\boldsymbol{x}_\alpha, \quad \boldsymbol{x}'_\alpha = \bar{\boldsymbol{x}}'_\alpha + \Delta\boldsymbol{x}'_\alpha, \tag{3}$$

and regard $\Delta\boldsymbol{x}_\alpha$ and $\Delta\boldsymbol{x}'_\alpha$ as independent Gaussian random variables of mean $\boldsymbol{0}$ but not necessarily isotropic or homogeneous. Let $V[\boldsymbol{x}_\alpha]$ and $V[\boldsymbol{x}'_\alpha]$ be the covariance matrices of vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, respectively. In practice, we need not know their absolute values; it suffices to know them *up to scale*. So, we write

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha], \quad V[\boldsymbol{x}'_\alpha] = \epsilon^2 V_0[\boldsymbol{x}'_\alpha], \tag{4}$$

and assume that $V_0[\boldsymbol{x}_\alpha]$ and $V_0[\boldsymbol{x}'_\alpha]$, which we call the *normalized covariance matrices*, are known but the constant ϵ, which we call the *noise level*, is unknown. Since the third components of $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ are identically 1, the matrices $V_0[\boldsymbol{x}_\alpha]$ and $V_0[\boldsymbol{x}'_\alpha]$ are singular with third columns and third rows filled with zeros.

[1] For example, we can take this to be the size of the image frame.

The normalized covariance matrices can be estimated from Hessian of the residual surface of template matching [4, 16, 20, 21]. If the noise has the same isotropic distribution everywhere, we have $V_0[\boldsymbol{x}_\alpha] = V_0[\boldsymbol{x}'_\alpha] = \mathrm{diag}(1,1,0)$, where $\mathrm{diag}(\cdots)$ is the diagonal matrix with diagonal elements $\cdots$. We use this as the default value when no information is available about the noise behavior.

4. Theoretical Accuracy Bound

Let $\hat{\boldsymbol{F}}$ be an estimate of the fundamental matrix, and $\bar{\boldsymbol{F}}$ its true value. The uncertainty of the estimate $\hat{\boldsymbol{F}}$ is measured by its *covariance tensor*

$$\mathcal{V}[\hat{\boldsymbol{F}}] = E[\mathcal{P}\Big((\hat{\boldsymbol{F}} - \bar{\boldsymbol{F}}) \otimes (\hat{\boldsymbol{F}} - \bar{\boldsymbol{F}})\Big)\mathcal{P}^\top], \quad (5)$$

where $E[\cdot]$ denotes expectation. The operator $\otimes$ denotes tensor product: for matrices $\boldsymbol{A} = (A_{ij})$ and $\boldsymbol{B} = (B_{ij})$, the $(ijkl)$ element of their tensor product is $A_{ij}B_{kl}$. For tensors $\mathcal{P} = (P_{ijkl})$ and $\mathcal{T} = (T_{ijkl})$, the product $\mathcal{P}\mathcal{T}\mathcal{P}^\top$ is a tensor whose $(ijkl)$ element is $\sum_{m,n,p,q=1}^{3} P_{ijmn}P_{klpq}T_{mnpq}$. The $(ijkl)$ element of the tensor $\mathcal{P} = (P_{ijkl})$ in eq. (5) is given by

$$P_{ijkl} = \delta_{ik}\delta_{jl} - \bar{F}_{ij}\bar{F}_{kl}. \quad (6)$$

Here, δ_{ij} is the Kronecker delta, taking 1 for $i = j$ and 0 otherwise.

Invoking the general theory of statistical optimization [9], we can derive a (non-CRLB type) lower bound on the covariance tensor $\mathcal{V}[\hat{\boldsymbol{F}}]$.

Define the *moment tensor* $\bar{\mathcal{M}} = (\bar{M}_{ijkl})$ by

$$\bar{\mathcal{M}} = \frac{1}{N}\sum_{\alpha=1}^{N} \bar{W}_\alpha \bar{\boldsymbol{x}}_\alpha \otimes \bar{\boldsymbol{x}}'_\alpha \otimes \bar{\boldsymbol{x}}_\alpha \otimes \bar{\boldsymbol{x}}'_\alpha, \quad (7)$$

$$\bar{W}_\alpha = \frac{1}{(\bar{\boldsymbol{x}}'_\alpha, \bar{\boldsymbol{F}}^\top V_0[\boldsymbol{x}_\alpha]\bar{\boldsymbol{F}}\bar{\boldsymbol{x}}'_\alpha) + (\bar{\boldsymbol{x}}_\alpha, \bar{\boldsymbol{F}}V_0[\boldsymbol{x}'_\alpha]\bar{\boldsymbol{F}}^\top\bar{\boldsymbol{x}}_\alpha)}. \quad (8)$$

The accuracy bound is given in the form

$$\mathcal{V}[\hat{\boldsymbol{F}}] \succ \frac{\epsilon^2}{N}\Big(\mathcal{P}^{\mathcal{S}}\bar{\mathcal{M}}\mathcal{P}^{\mathcal{S}\top}\Big)_7^-. \quad (9)$$

where $\mathcal{T} \succ \mathcal{S}$ for tensors $\mathcal{T}$ and $\mathcal{S}$ means that $\mathcal{T} - \mathcal{S}$ is a positive semi-definite tensor, and the operation $(\cdot)_r^-$ denotes the (Moore-Penrose) generalized inverse of rank r (discussed later). The $(ijkl)$ element of the tensor $\mathcal{P}^{\mathcal{S}} = (P^{\mathcal{S}}_{ijkl})$ in eq. (9) is given by

$$P^{\mathcal{S}}_{ijkl} = \delta_{ik}\delta_{jl} - \frac{\bar{F}^\dagger_{ji}\bar{F}^\dagger_{lk}}{\|\bar{\boldsymbol{F}}^\dagger\|^2}, \quad (10)$$

where $\bar{\boldsymbol{F}}^\dagger$ is the cofactor matrix of $\bar{\boldsymbol{F}}$.

For a tensor $\mathcal{T} = (T_{ijkl})$, a matrix $\boldsymbol{A} = (A_{ij})$, and a scalar λ, we say that $\boldsymbol{A}$ is an *eigenmatrix* of $\mathcal{T}$ with *eigenvalue* λ if $\mathcal{T}\boldsymbol{A} = \lambda\boldsymbol{A}$, where the product $\mathcal{T}\boldsymbol{A}$ is a matrix whose (ij) element is $\sum_{k,l=1}^{3} T_{ijkl}A_{kl}$. The eigenmatrices and eigenvalues of a tensor can be computed by identifying a tensor and a matrix with a 9×9 matrix and an 9-dimensional vector [9].

A tensor $\mathcal{T} = (T_{ijkl})$ is said to be *symmetric* if $T_{ijkl} = T_{klij}$. A symmetric $3 \times 3 \times 3 \times 3$ tensor has nine real eigenvalues $\{\lambda_i\}$. The corresponding eigenmatrices $\{\boldsymbol{U}_i\}$ can be chosen to be an orthogonal system of matrices of unit norm, where the inner product of matrices $\boldsymbol{A} = (A_{ij})$ and $\boldsymbol{B} = (B_{ij})$ is defined by $(\boldsymbol{A}; \boldsymbol{B}) = \sum_{i,j=1}^{3} A_{ij}B_{ij}$. A symmetric tensor is *positive semi-definite* if its eigenvalues are all nonnegative.

Let $\lambda_1 \geq \cdots \geq \lambda_9$ (≥ 0) be the eigenvalues of a positive semi-definite symmetric tensor $\mathcal{T}$, and let $\{\boldsymbol{U}_1, ..., \boldsymbol{U}_9\}$ be the corresponding orthonormal set of eigenmatrices of unit norm. If $\lambda_r > 0$, the *(Moore-Penrose) generalized inverse* of $\mathcal{T}$ of rank r is computed as follows:

$$\mathcal{T}_r^- = \sum_{i=1}^{r} \frac{\boldsymbol{U}_i \otimes \boldsymbol{U}_i}{\lambda_i}. \quad (11)$$

The *root-mean-square error* of an estimate $\hat{\boldsymbol{F}}$ is defined by

$$\mathrm{rms}[\hat{\boldsymbol{F}}] = \sqrt{E[\|\mathcal{P}(\hat{\boldsymbol{F}} - \bar{\boldsymbol{F}})\|^2]}. \quad (12)$$

Since $\mathcal{P}$ is a projection, it follows that $0 \leq \mathrm{rms}[\hat{\boldsymbol{F}}] \leq 1$. From eq. (9), we may infer the lower bound

$$\mathrm{rms}[\hat{\boldsymbol{F}}] \geq \sqrt{\mathrm{tr}\mathcal{V}[\hat{\boldsymbol{F}}]}, \quad (13)$$

where the *trace* $\mathrm{tr}\mathcal{T}$ of a tensor $\mathcal{T} = (T_{ijkl})$ is defined by

$$\mathrm{tr}\mathcal{T} = \sum_{k,l=1}^{3} T_{klkl}. \quad (14)$$

5. Algorithm

Our algorithm consists of two stages. First, we optimally compute the matrix $\boldsymbol{F}$ without considering the constraint $\det \boldsymbol{F} = 0$ by applying a technique called *renormalization* [9], which computes the weighted least-squares solution and iteratively removes the statistical bias incurred. In the second stage, we correct $\boldsymbol{F}$ optimally with respect to the covariance tensor $\mathcal{V}[\hat{\boldsymbol{F}}]$ computed in the first stage.

5.1. Renormalization

1. Let $c = 0$, $W_\alpha = 1$, $\alpha = 1, ..., N$, and $J = \infty$. % In real computation, the symbol ∞ means a very large number, e.g., 10^{10}.
2. Compute the tensors $\mathcal{M} = (M_{ijkl})$ and $\mathcal{N} = (N_{ijkl})$ as follows:

$$M_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha x_{\alpha(i)}x'_{\alpha(j)}x_{\alpha(k)}x'_{\alpha(l)}, \quad (15)$$

$$N_{ijkl} = \frac{1}{N}\sum_{\alpha=1}^{N} W_\alpha (V_0[\boldsymbol{x}_\alpha]_{ik} x'_{\alpha(j)} x'_{\alpha(l)} + V_0[\boldsymbol{x}'_\alpha]_{jl} x_{\alpha(i)} x_{\alpha(k)}). \quad (16)$$

% $x_{\alpha(i)}$ and $x'_{\alpha(i)}$ are the ith components of $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, respectively, and $V_0[\boldsymbol{x}_\alpha]_{ij}$ and $V_0[\boldsymbol{x}'_\alpha]_{ij}$ are the (ij) elements of $V_0[\boldsymbol{x}_\alpha]_{ij}$ and $V_0[\boldsymbol{x}'_\alpha]_{ij}$, respectively.

3. Compute the nine eigenvalues $\lambda_1 \geq \cdots \geq \lambda_9$ of the tensor $\mathcal{M}$ and the corresponding orthonormal set $\{\boldsymbol{F}_1, ..., \boldsymbol{F}_9\}$ of eigenmatrices of unit norm.
4. Do the following computation:

- Update c as follows:

$$c \leftarrow c + \frac{\lambda_9}{(\boldsymbol{F}_9; \mathcal{N}\boldsymbol{F}_9)}. \quad (17)$$

- Compute W_α, $\alpha = 1, ..., N$, as follows:

$$W_\alpha = 1 \Big/ \Big((\boldsymbol{x}'_\alpha, \boldsymbol{F}_9^\top V_0[\boldsymbol{x}_\alpha] \boldsymbol{F}_9 \boldsymbol{x}'_\alpha) + (\boldsymbol{x}_\alpha, \boldsymbol{F}_9 V_0[\boldsymbol{x}'_\alpha] \boldsymbol{F}_9^\top \boldsymbol{x}_\alpha) \Big). \quad (18)$$

- Compute the tensors $\mathcal{M}$ and $\mathcal{N}$ by eqs. (15) and (16).
- Let

$$J' \leftarrow J, \quad J \leftarrow (\boldsymbol{F}_9, \mathcal{M}\boldsymbol{F}_9). \quad (19)$$

- If $J' < J$, let $J \leftarrow J'$.
  % The variable J stores the minimum residual.
- Else, compute the nine eigenvalues $\lambda_1 \geq \cdots \geq \lambda_9$ of the tensor

$$\hat{\mathcal{M}} = \mathcal{M} - c\mathcal{N} \quad (20)$$

and the corresponding orthonormal set $\{\boldsymbol{F}_1, ..., \boldsymbol{F}_9\}$ of eigenmatrices of unit norm.

5. Repeat the above computation until $J' \leq J$ or $|\lambda_9| \approx 0$.
   % This guarantees the iterations to converge.
6. Let $\boldsymbol{F}$ take the value $\boldsymbol{F}_9$, this being our estimate of the fundamental matrix.

5.2 Optimal correction

1. Estimate the squared noise level $\hat{\epsilon}^2$ as follows:

$$\hat{\epsilon}^2 = \frac{J}{1 - 8/N}. \quad (21)$$

2. Compute the normalized *covariance tensor* of $\boldsymbol{F}$ as follows:

$$\mathcal{V}_0[\boldsymbol{F}] = \frac{1}{N}\sum_{i=1}^{8} \frac{\boldsymbol{F}_i \otimes \boldsymbol{F}_i}{\lambda_i}. \quad (22)$$

3. Repeat the following computation until $\det \boldsymbol{F} \approx 0$.

- Update $\boldsymbol{F}$ as follows:

$$\boldsymbol{F} \leftarrow N[\boldsymbol{F} - \frac{(\det \boldsymbol{F}) \mathcal{V}_0[\boldsymbol{F}] \boldsymbol{F}^{\dagger\top}}{(\boldsymbol{F}^{\dagger\top}; \mathcal{V}_0[\boldsymbol{F}] \boldsymbol{F}^{\dagger\top})}]. \quad (23)$$

  % The operator $N[\,\cdot\,]$ denotes normalization of the norm to 1: $N[\boldsymbol{F}] = \boldsymbol{F}/\|\boldsymbol{F}\|$.
- Compute the projection tensor $\mathcal{P} = (P_{ijkl})$ as follows:

$$P_{ijkl} = \delta_{ik}\delta_{jl} - F_{ij}F_{kl}. \quad (24)$$

- Update the normalized covariance tensor $\mathcal{V}_0[\boldsymbol{F}]$ as follows:

$$\mathcal{V}_0[\boldsymbol{F}]_{ijkl} \leftarrow \sum_{m,n,p,q=1}^{3} P_{ijmn} P_{klpq} \mathcal{V}_0[\boldsymbol{F}]_{mnpq}. \quad (25)$$

  % Since errors of $\boldsymbol{F}$ should be orthogonal to $\boldsymbol{F}$, we project the error distribution onto the space orthogonal to $\boldsymbol{F}$.

5.3. Program package

This algorithm is implemented in C++ and placed in our Web page[2]. It outputs a solution $\hat{\boldsymbol{F}}$ along with its *standard deviation pair* $\{\boldsymbol{F}^{(+)}, \boldsymbol{F}^{(-)}\}$. These two matrices are the values in the parameter space that are separated from $\hat{\boldsymbol{F}}$ by the standard deviation in the two direction along which errors are the most likely to occur. Let $\lambda_{\max}$ be the maximum eigenvalue of the tensor on the right-hand side of eq. (9) computed after substituting the data and the estimate $\hat{\boldsymbol{F}}$ for their true values, and let $\boldsymbol{U}_{\max}$ be the corresponding eigenmatrix of unit norm. The primary deviation pair is defined by

$$\boldsymbol{F}^{(+)} = N[\hat{\boldsymbol{F}} + \sqrt{\lambda_{\max}}\boldsymbol{U}_{\max}], \quad \boldsymbol{F}^{(-)} = N[\hat{\boldsymbol{F}} - \sqrt{\lambda_{\max}}\boldsymbol{U}_{\max}]. \quad (26)$$

If $\boldsymbol{F}^{(+)}$ and $\boldsymbol{F}^{(-)}$ coincide up to, say, three significant digits, the solution $\hat{\boldsymbol{F}}$ is likely to have accuracy up to approximately three significant digits.

The fundamental matrix cannot be defined uniquely if the feature points are in a degenerate configuration. This occurs, for example, when the camera translation is zero or all the feature points are on a special quadric called a *critical surface*, a typical instance of which is a planar surface [8]. If $\lambda_{\max}$ in eqs. (26) is predicted to be approximately 1 in the course of computation, our program judges that degeneracy has occurred and stops the computation after issuing a warning message.

6. Accuracy Bounds on Epipoles

Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be the unit eigenvectors of $\boldsymbol{F}^\top$ and $\boldsymbol{F}$, respectively, with eigenvalue 0. The *epipoles* in

[2] http://www.ail.cs.gunma-u.ac.jp/~kanatani/e

the first and the second images are the points defined respectively by

$$\boldsymbol{x}_e = Z[\boldsymbol{u}], \quad \boldsymbol{x}'_e = Z[\boldsymbol{v}], \tag{27}$$

where $Z[\,\cdot\,]$ denotes normalization to make the third component 1. The epipole $\boldsymbol{x}_e$ indicates the position of the center of the lens of the second camera in the first image; the epipole $\boldsymbol{x}'_e$ indicates the position of the center of the lens of the first camera in the second image.

Since we have a theoretical lower bound on the covariance tensor $\mathcal{V}[\hat{\boldsymbol{F}}]$, we have the lower bound on the covariance matrix of $\boldsymbol{x}_e$. Invoking the *perturbation theorem* [9], we can obtain the form

$$V[\boldsymbol{x}_e] = \frac{\boldsymbol{Q}_{\bar{\mathbf{x}}_e} V[\boldsymbol{u}] \boldsymbol{Q}_{\bar{\mathbf{x}}_e}^\top}{(\boldsymbol{k}, \bar{\boldsymbol{u}})^2}, \tag{28}$$

where the bars denote the true values and $\boldsymbol{k} = (0,0,1)^\top$ is a unit vector in the direction of the optical axis of the first camera. For a vector $\boldsymbol{x}$, we define the associated projection matrices $\boldsymbol{Q}_{\mathbf{x}}$ by

$$\boldsymbol{Q}_{\mathbf{x}} = \boldsymbol{I} - \boldsymbol{x}\boldsymbol{k}^\top \tag{29}$$

In eq. (28), the covariance matrix $V[\boldsymbol{u}]$ of the vector $\boldsymbol{u}$ and $\boldsymbol{v}$ is given by

$$V[\boldsymbol{u}] = \epsilon^2 (\bar{\boldsymbol{F}}\bar{\boldsymbol{F}}^\top)_2^- \boldsymbol{G} (\bar{\boldsymbol{F}}\bar{\boldsymbol{F}}^\top)_2^-, \tag{30}$$

where we have defined the matrix $\boldsymbol{G} = (G_{ij})$ by

$$G_{ij} = \sum_{j,k=1}^{3} E_{ijkl} \bar{u}_k \bar{u}_l. \tag{31}$$

The tensor $\mathcal{E} = (E_{ijkl})$ is defined by

$$E_{ijkl} = \sum_{m,n=1}^{3} \Big(\mathcal{V}_0[\boldsymbol{F}]_{imjn} \bar{F}_{km} \bar{F}_{ln} + \mathcal{V}_0[\boldsymbol{F}]_{imln} \bar{F}_{km} \bar{F}_{jn} + \mathcal{V}_0[\boldsymbol{F}]_{kmjn} \bar{F}_{im} \bar{F}_{ln} + \mathcal{V}_0[\boldsymbol{F}]_{kmln} \bar{F}_{im} \bar{F}_{jn} \Big). \tag{32}$$

The bound on the epipole $\boldsymbol{x}'_e$ is obtained by replacing $\boldsymbol{u}$ and $\boldsymbol{F}$ in the above procedure by $\boldsymbol{v}$ and $\boldsymbol{F}^\top$, respectively.

7. Simulated Experiments

Fig. 1 shows two images of a simulated 3-D grid environment. They are supposedly captured by a moving camera with different focal lengths. Some of the epipolars are superimposed. Random Gaussian noise of mean 0 and standard deviation σ (pixels) was added to the x and y coordinates of each grid point independently, and the fundamental matrix was computed by using the default noise model. The renormalization converged after three or four iterations.

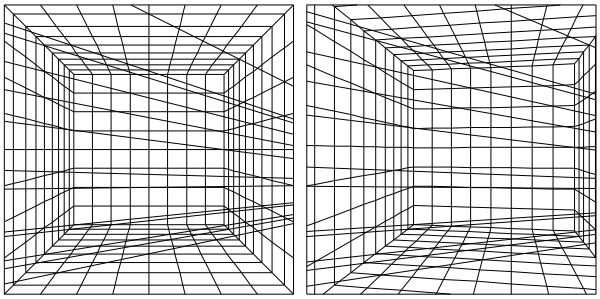


Figure 1: Simulated images of a 3-D scene and epipolars.

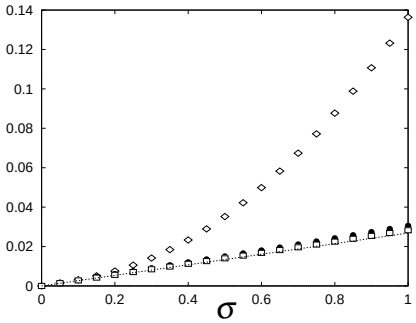


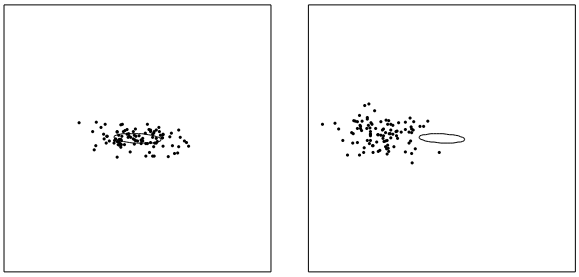
Figure 2: Root-means-squares error: our solutions (□), solutions without the optimal correction (●), and least-squares solutions (◇). The dotted line indicates the theoretical lower bound.

Fig. 2(a) shows a plot of the root-mean-squares error $\sqrt{\sum_{a=1}^{100} \|\mathcal{P}(\hat{\boldsymbol{F}}^a - \bar{\boldsymbol{F}})\|^2 / 100}$ over 100 trials for each σ using different noise each time, where $\hat{\boldsymbol{F}}^a$ is the ath estimate, $\bar{\boldsymbol{F}}$ is the true value, and $\mathcal{P}$ is the projection tensor defined in eq. (6). The symbol □ denotes solutions obtained via the method presented in this paper, and the dotted line indicates the theoretical lower bound derived from eq. (13). The symbol ● denotes renormalization solutions without applying the optimal correction of Section 5.2. The symbol ◇ denotes solutions computed by the naive linear algorithm, often referred to as the *least-squares method* or the *algebraic distance minimization* [6, 7], which minimizes the expression

$$\frac{1}{N} \sum_{\alpha=1}^{N} (\boldsymbol{x}_\alpha, \boldsymbol{F}\boldsymbol{x}'_\alpha)^2 \to \min, \tag{33}$$

obtained directly from the epipolar constraint (2). Applying this method, rather than our own, amounts to removing the iterations of Step 4 in the renormalization procedure of Section 5.1.

Theoretically, our method is proved to attain the theoretical accuracy bound in the first order (i.e., if higher order terms in the noise magnitude are omitted) [9]. Fig. 2, we see that the errors in our estimates practically fall on the theoretical lower bound, which is known to be attained by the bundle adjustment. This means that our linear algorithm achieves the accuracy of the bundle adjustment.



(a) (b)

Figure 3: Reliability of the epipoles. The ellipses indicate the theoretical lower bound on the standard deviation in each orientation. (a) Our solutions. (b) Least-squares solutions.

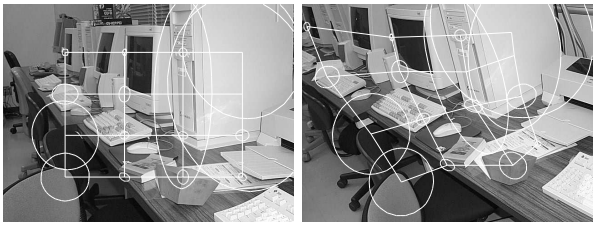


Figure 4: Real images of an indoor scene. Corresponding grid patterns are generated and matched by deformable template matching. The ellipses visualize the covariance matrices of the matched points.

Fig. 3(a) is an enlargement near the epipole (outside the image frame), in which we plot the epipoles computed 100 times using different noise each time for $\sigma = 0.5$ (pixels). The ellipse depicts the lower bound on the standard deviation in each orientation. It is computed from the covariance matrix $V[\boldsymbol{x}_e]$ given in eq. (28) and is centered on the true epipole $\bar{\boldsymbol{x}}_e$. Again, the indications are that the errors in our estimates conform to the theoretical lower bound. Fig. 3(b) shows the corresponding result for the least-squares solutions; it is evident that the distribution has a large statistical bias.

8. Real Image Examples

8.1 Covariance evaluation of image features

Fig. 4 shows a pair of real images of an indoor scene. We generated a grid pattern in each image and manually adjusted the second one so that its vertices approximately correspond to those of the first one. Then, each vertex was corrected with sub-pixel accuracy by deformable template matching (we omit the details). The ellipses drawn in the images visualize the covariance matrices (defined up to scale) of the individual vertices estimated from the gray level variations in their neighborhoods [4, 16, 20, 21] (we omit the details). As we can see, the covariance diverges and produces a huge ellipse if the neighborhood is of almost uniform gray levels; if the point is on an edge line, the ellipse is elongated along it.

From these 16 corresponding vertices we computed the fundamental matrix in three ways: the least square solution $\boldsymbol{F}_{LS}$, the optimal solution $\boldsymbol{F}_1$ for the default noise model, and the optimal solution $\boldsymbol{F}_2$ for the estimated covariance matrices. We also computed the fundamental matrix $\bar{\boldsymbol{F}}$ from 68 carefully selected feature points (not shown in the images) by using the default noise level. Regarding $\bar{\boldsymbol{F}}$ as the true value, we evaluated the errors in $\boldsymbol{F}_{LS}$, $\boldsymbol{F}_1$, and $\boldsymbol{F}_2$ as follows:

$$\|\boldsymbol{F}_{LS} - \bar{\boldsymbol{F}}\| = 0.0161,$$

$$\|\boldsymbol{F}_1 - \bar{\boldsymbol{F}}\| = 0.0151, \quad \|\boldsymbol{F}_2 - \bar{\boldsymbol{F}}\| = 0.0098. \quad (34)$$

Of course, this is merely an illustration. In fact, 16 points are not sufficient for reliable 3-D reconstruction, and this type of feature generation is unlikely to select good feature points. According to our experience, the use of the default noise model allows us sufficient accuracy if salient feature points are carefully positioned by hand. However, estimation of noise inhomogeneity or anisotropy becomes crucial if feature points are automatically or semi-automatically positioned by an image processing operation, as shown above.

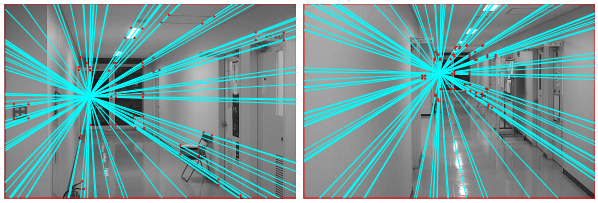


Figure 5: Real images of an indoor scene and the computed epipoles and epipolars.

8.2 Reliability evaluation of epipoles

Fig. 5 shows another pair of images of an indoor scene captured by a moving camera with varying zooming. Feature points were located manually, and the fundamental matrix was computed by using the default noise model. The computed epipoles and epipolars are superimposed in the images. The inherent reliability of the epipole is 3.44 pixels for the first image and 1.74 pixels for the second, which are obtained by computing $\sqrt{\mathrm{tr}V[\boldsymbol{x}_e]}$ and $\sqrt{\mathrm{tr}V[\boldsymbol{x}'_e]}$.

8.3 Reliability evaluation of 3-D reconstruction

If the principal point (i.e., the point corresponding to the camera optical axis in the image) is at the center of the frame and there is no image skew, we can self-calibrate the focal lengths of the two images by an analytical method [2, 5, 11, 19] and thereby reconstruct the 3-D Euclidean structure from two views.

Using the images shown in Fig. 6, we reconstructed the 3-D Euclidean shape from the 60 manually selected feature points (not shown) by using

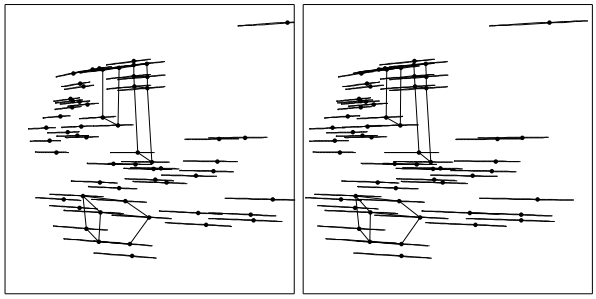


Figure 6: The reconstructed 3-D structure with uncertainty ellipsoids (stereogram).

the method described in [11]. Fig. 6 show a stereogram of the reconstructed 3-D shape. Wireframes consisting of some of the reconstructed points are shown.

On each reconstructed point is centered the ellipsoid that indicates its standard deviation in every direction; they are obtained by propagating the assumed image noise characteristics and expanding the accuracy bound on the fundamental matrix (we omit the details). All the ellipsoids look like thin needles, showing that the uncertainty is very large along the depth orientation. It can be shown that this uncertainty is mostly due to the uncertainty in the camera translation, not the object shape[3] [10]. Our algorithm enables us to do this type of reliability evaluation very easily.

9. Concluding Remarks

A linear algorithm has been presented for computing the fundamental matrix from corresponding points over two images in the presence of independent Gaussian noise that may be anisotropic and inhomogeneous. The computation is divided into two stages: estimation of the fundamental matrix by renormalization without considering the rank constraint and its correction in order to impose that constraint. Experiments indicate that the obtained estimates are in the vicinity of the theoretical accuracy bound, which can be used for evaluating the reliability the solution. The evaluation process has been illustrated by using simulated and real images. Our algorithm is implemented as a public-domain program in C++.

References

[1] M. Bober, N. Geogis and J. Kittler, On accurate and robust estimation of fundamental matrix, *Comput. Vision Image Understanding*, **72**-1 (1998), 39–53.

[2] S. Bougnoux, From projective to Euclidean space under any practical situation, a criticism of self calibration, *Proc. 6th Int. Conf. Comput. Vision.*, January 1998, Bombay, India, pp. 790–796.

[3] G. Csurka, C. Zeller, Z. Zhang and O. D. Faugeras, Characterizing the uncertainty of the fundamental matrix, *Comput Vision Image Understand.*, **68**-1 (1997), 18–36.

[4] W. Förstner, Reliability analysis of parameter estimation in linear models with applications to mensuration problems in computer vision, *Comput. Vision Graphics Image Process.*, **40** (1987), 273–310.

[5] R. I. Hartley, Estimation of relative camera position for uncalibrated cameras, *Proc. 2nd Euro. Conf. Comput. Vision*, May 1992, Santa Margherita Ligure, Italy, pp. 579–587.

[6] R. I. Hartley, In defense of the eight-point algorithm, *IEEE Trans. Patt. Anal. Mach. Intell.*, **97**-6 (1997), 580–593.

[7] R. I. Hartley, Minimizing algebraic error, *Phil. Trans. R. Soc. Lond.*, A, **3569** (1998), 1175–1192.

[8] K. Kanatani, *Geometric Computation for Machine Vision*, Oxford University Press, Oxford, U.K., 1993.

[9] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier, Amsterdam, The Netherlands, 1996.

[10] K. Kanatani, Gauge-based reliability analysis of 3-D reconstruction from two uncalibrated perspective views, *Proc. 15th Int. Conf. Patt. Recog.*, September 2000, Barcelona, Spain, to appear.

[11] K. Kanatani and C. Matsunaga, Closed-form expression for focal lengths from the fundamental matrix, *Proc. 4th Asian Conf. Comput. Vision*, January 2000, Taipei, Taiwan, pp. 128–133.

[12] K. Kanatani and D. D. Morris, Gauges and gauge transformations in 3-D reconstruction from a sequence of images, *Proc. 4th Asian Conf. Comput. Vision*, January 2000, Taipei, Taiwan, pp. 1046–1051.

[13] Q.-T. Luong and O. D. Faugeras, The fundamental matrix: Theory, algorithm, and stability analysis, *Int. J. Comput. Vision*, **17**-3 (1996), 43–7.

[14] Q.-T. Luong and O. D. Faugeras, Self-calibration of a moving camera from point correspondences and fundamental matrices, *Int. J. Comput. Vision*, **23**-3 (1997), 261–289.

[15] P. F. McLauchlan, Coordinate-frame independence in optimization algorithms for 3D vision, *IEEE Workshop on Vision Algorithms: Theory and Practice*, September 1999, Corfu, Greece.

[16] D. D. Morris and T. Kanade, A unified factorization algorithm for points, line segments and planes with uncertainty models, *Proc. 6th Int. Conf. Comput. Vision*, January 1998, Bombay, India, pp. 696–702.

[17] D. D. Morris, K. Kanatani and T. Kanade, Uncertainty modeling for optimal structure from motion, *IEEE Workshop on Vision Algorithm: Theory and Practice*, September 1999, Corfu, Greece.

[18] M. Mühlich and R. Mester, The role of total least squares in motion analysis, *Proc. 5th Euro. Conf. Comput. Vision*, June 1998, Freiburg, Germany, pp. 305–321.

[19] G. N. Newsam, D. Q. Huynh, M. J. Brooks and H.-P. Pan, Recovering unknown focal lengths in self-calibration: An essentially linear algorithm and degenerate configurations, *Int. Arch. Photogram. Remote Sensing*, **31**-B3-III, July 1996, Vienna, Austria, pp. 575–580.

[20] J. Shi and C. Tomasi, Good features to track, *Proc. IEEE Conf. Comput. Vision Patt. Recogn.*, June 1994, Seattle, WA, pp. 593–600.

[21] A. Singh, An estimation-theoretic framework for image-flow computation, *Proc. 3rd Int. Conf. Comput. Vision*, December 1990, Osaka, Japan, pp. 168–177.

[22] P. H. S. Torr and A. Zissermann, Performance characterization of fundamental matrix estimation under image degradation, *Mach. Vision Appl.*, **9** (1997), 321–333.

[23] P. H. S. Torr and A. Zisserman, Robust detection of degenerate configurations while estimating the fundamental matrix, *Comput. Vision Image Understanding*, **71**-3 (1998), 312–333.

[24] Z. Zhang, Determining the epipolar geometry and its uncertainty: A review, *Int. J. Comput. Vision*, **27**-2 (1998), 161–195.

[25] Z. Zhang, On the optimization criteria used in two-view motion analysis, *IEEE Trans. Patt. Anal. Mach. Intell.*, **20**-7 (1998), 717–729.

[3] This could be easily seen by applying a "gauge transformation", i.e., changing the "gauge" for normalizing indeterminacy from the camera-centered coordinate system to an object-centered one, or evaluating the covariances of "gauge invariants" [12, 15, 17].