PAPER

Direct Reconstruction of Planar Surfaces by Stereo Vision

Yasushi KANAZAWA† *and* Kenichi KANATANI††, *Members*

SUMMARY This paper studies the problem of reconstructing a planar surface from stereo images of multiple feature points that are known to be coplanar in the scene. We present a direct method by applying maximum likelihood estimation based on a statistical model of image noise. The significant fact about our method is that not only the 3-D position of the surface is reconstructed accurately but its reliability is also computed quantitatively. The effectiveness of our method is demonstrated by doing numerical simulation.

key words: *reliability of 3-D reconstruction, stereo vision, plane fitting, statistical optimization, noise estimation*

1. Introduction

Stereo vision is one of the most fundamental means of 3-D sensing from images and is widely used as a visual sensor for autonomous navigation of robots [1], [7]. In the past, the study of stereo vision has mainly focused on the *correspondence detection* between the two images. In fact, detecting correspondences is a very difficult task to automate efficiently, and many detection techniques have been proposed [1]. However, various other issues arise when we reconstruct 3-D from detected correspondences. First of all, the 3-D reconstruction should be accurate. Hence, we must apply an optimization technique that maximizes the accuracy of the reconstruction by considering the statistical characteristics of the image noise. At the same time, *the reliability of the reconstructed 3-D must be evaluated* [6]. If the errors involved in the reconstructed 3-D cannot be estimated, robots cannot take appropriate actions to archive given tasks effectively. This paper presents a new theory for reconstructing planar surfaces by stereo vision in a statistically optimal way and evaluating the reliability of the reconstruction in quantitative terms.

Many man-made objects have planar surfaces. Hence, 3-D reconstruction of planar surfaces is one of the most important tasks for autonomous robot operations. If the feature points we are observing are assumed to be coplanar in the scene, the parameters of the plane can be computed by the following *least-squares method*:

1. Correspondences of the feature points are established between the two stereo images.
2. Their 3-D positions are computed by triangulation.
3. A planar surface is fitted so that the sum of the squares of the perpendicular distances from the computed positions to the surface is minimized.

This method is very simple but indirect. In this paper, we propose a method for a direct reconstruction based on geometric constraints and a statistical model of image noise.

In order to reconstruct an optimal planar surface, we introduce the principle of *maximum likelihood estimation* and derive a scheme for nonlinear optimization. At the same time, we derive a *theoretical bound* on the attainable accuracy of the estimation in the form of the covariance matrix of the estimate. In order to compute the optimal solution, we use a numerical scheme called *renormalization* [3]. By numerical simulation, we show that the obtained solution almost attains the theoretical bound on accuracy. This means that we can quantitatively predict the reliability of the reconstructed surface. This fact has a great significance in robotics applications of stereo vision.

2. Camera and Stereo Model

Let $\{P_\alpha\}, \alpha = 1, ..., N$, be feature points on a planar surface in the scene. Let $\boldsymbol{n}$ be the unit normal to the surface, and d the distance of it from the origin O. We call $\{\boldsymbol{n}, d\}$ the *surface parameters.* As illustrated in Fig. 1, we take the first camera as the reference coordinate system and place the second camera in a position

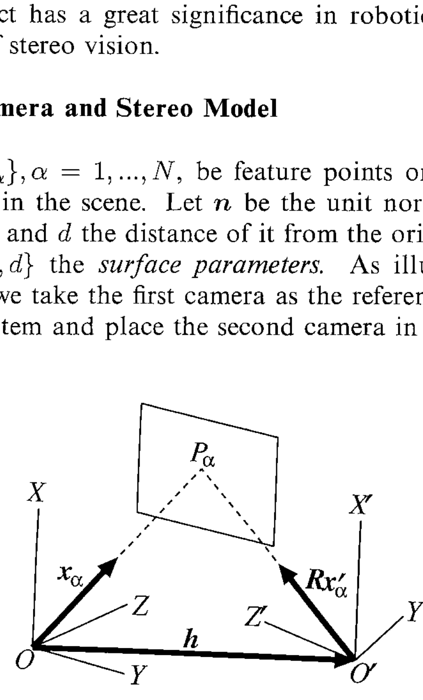


Fig. 1 The camera model and the coordinates systems.

Manuscript received September 10, 1994.
Manuscript revised January 27, 1995.

†The author is with the Department of Information and Computer Engineering, Gunma College of Technology, Maebashi-shi, 371 Japan.

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

obtained by translating the first camera by vector $\boldsymbol{h}$ and rotating it around the center of the lens by matrix $\boldsymbol{R}$. We call $\{\boldsymbol{h}, \boldsymbol{R}\}$ the *motion* (or *stereo*) *parameters*. The two cameras may have different focal lengths f and f'.

Let $\{(x_\alpha, y_\alpha)\}, \alpha = 1, ..., N$, be the image coordinates of the feature points projected on the image plane of the first camera, and $\{(x'_\alpha, y'_\alpha)\}, \alpha = 1, ..., N$, those for the second camera. We use the following three-dimensional vectors to represent them:

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

In the absence of noise, the vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, the motion parameters $\{\boldsymbol{h}, \boldsymbol{R}\}$, and the surface parameters $\{\boldsymbol{n}, d\}$ satisfy the following relation (we omit the derivation [2], [4]):

$$\boldsymbol{x}'_\alpha \times \boldsymbol{A}\boldsymbol{x}_\alpha = \boldsymbol{0}, \quad \boldsymbol{A} = \frac{\boldsymbol{R}^\top(\boldsymbol{h}\boldsymbol{n}^\top - d\boldsymbol{I})}{\sqrt{1+d^2}}. \tag{2}$$

Here, $\boldsymbol{a} \times \boldsymbol{A}$ denotes the matrix defined by the vector product of three-dimensional vector $\boldsymbol{a}$ and each column of 3×3-matrix $\boldsymbol{A}$, and the superscript $\top$ denotes transpose. Let $\boldsymbol{B}_\alpha$ and $\boldsymbol{\nu}$ be a 3×4-matrix and a four-dimensional vector, respectively, defined by

$$\boldsymbol{B}_\alpha = \begin{pmatrix} \boldsymbol{x}'_\alpha \times \boldsymbol{R}^\top\boldsymbol{h}\boldsymbol{x}_\alpha^\top & \boldsymbol{x}'_\alpha \times \boldsymbol{R}^\top\boldsymbol{x}_\alpha \end{pmatrix}, \tag{3}$$

$$\boldsymbol{\nu} = \frac{1}{\sqrt{1+d^2}}\begin{pmatrix} \boldsymbol{n} \\ -d \end{pmatrix}. \tag{4}$$

Then, Eqs. (2) can be rewritten in the following form:

$$\boldsymbol{B}_\alpha\boldsymbol{\nu} = \boldsymbol{0}. \tag{5}$$

From Fig. 1, we see that vectors $\boldsymbol{x}_\alpha$, $\boldsymbol{h}$ and $\boldsymbol{R}\boldsymbol{x}'_\alpha$ are coplanar. It follows that $\boldsymbol{R}^\top\boldsymbol{x}_\alpha$, $\boldsymbol{R}^\top\boldsymbol{h}$, and $\boldsymbol{x}'_\alpha$ are also coplanar. Consequently, vectors $\boldsymbol{x}'_\alpha \times \boldsymbol{R}^\top\boldsymbol{h}$ and $\boldsymbol{x}'_\alpha \times \boldsymbol{R}^\top\boldsymbol{x}_\alpha$ are collinear, so the matrix $\boldsymbol{B}_\alpha$ has rank 1.

3. Statistical Model of Image Noise

In the presence of noise, vectors $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ do not necessarily satisfy Eq. (5). Write

$$\boldsymbol{x}_\alpha = \bar{\boldsymbol{x}}_\alpha + \Delta\boldsymbol{x}_\alpha, \quad \boldsymbol{x}'_\alpha = \bar{\boldsymbol{x}}'_\alpha + \Delta\boldsymbol{x}'_\alpha, \tag{6}$$

where $\bar{\boldsymbol{x}}_\alpha$ and $\bar{\boldsymbol{x}}'_\alpha$ are the true values of $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$, respectively. We regard $\Delta\boldsymbol{x}_\alpha$ and $\Delta\boldsymbol{x}'_\alpha$ as random variables that have means 0 and covariance matrices $V[\boldsymbol{x}_\alpha]$ and $V[\boldsymbol{x}'_\alpha]$, respectively [6]. The absolute magnitude of the image noise is very difficult to estimate a priori. Let ϵ be its average magnitude, which is unknown. We call it the *noise level.* On the other hand, geometric characteristics of image noise such as uniformity and isotropy can be easily predicted, so we introduce the *normalized covariance matrices* $V_0[\boldsymbol{x}_\alpha]$ and $V_0[\boldsymbol{x}'_\alpha]$, which are assumed to be known, and express the covariance matrices in the following form:

$$V[\boldsymbol{x}_\alpha] = \epsilon^2 V_0[\boldsymbol{x}_\alpha], \quad V[\boldsymbol{x}'_\alpha] = \epsilon^2 V_0[\boldsymbol{x}'_\alpha]. \tag{7}$$

4. Maximum Likelihood Estimation

We apply *maximum likelihood estimation* for estimating an optimal value of $\boldsymbol{\nu}$. First, we optimally correct $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ in the form

$$\hat{\boldsymbol{x}}_\alpha = \boldsymbol{x}_\alpha - \Delta\boldsymbol{x}_\alpha, \quad \hat{\boldsymbol{x}}'_\alpha = \boldsymbol{x}'_\alpha - \Delta\boldsymbol{x}'_\alpha, \tag{8}$$

so that Eq. (5) is satisfied for a fixed value of $\boldsymbol{\nu}$. If the image noise has a Gaussian distribution, this correction is done for each α by the optimization based on the *Mahalanobis distance* [4] in the form

$$J_\alpha = (\Delta\boldsymbol{x}_\alpha, V_0[\boldsymbol{x}_\alpha]^-\Delta\boldsymbol{x}_\alpha) + (\Delta\boldsymbol{x}'_\alpha, V_0[\boldsymbol{x}'_\alpha]^-\Delta\boldsymbol{x}'_\alpha) \to \min, \tag{9}$$

where $V_0[\boldsymbol{x}]^-$ is the generalized inverse of $V_0[\boldsymbol{x}]$ and $(\boldsymbol{a}, \boldsymbol{b})$ denotes the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. The residual J_α obtained by substituting the resulting optimal values $\hat{\boldsymbol{x}}_\alpha$ and $\hat{\boldsymbol{x}}'_\alpha$ is a function of $\boldsymbol{\nu}$, so we rewrite it as $J_\alpha[\boldsymbol{\nu}]$ and seek an optimal value of $\boldsymbol{\nu}$ by the minimization

$$\frac{1}{N}\sum_{\alpha=1}^{N} J_\alpha[\boldsymbol{\nu}] \to \min. \tag{10}$$

This minimization can be rewritten in the form

$$J[\boldsymbol{\nu}] = \frac{1}{N}\sum_{\alpha=1}^{N}(\boldsymbol{B}_\alpha\boldsymbol{\nu}, \boldsymbol{W}_\alpha(\boldsymbol{\nu})\boldsymbol{B}_\alpha\boldsymbol{\nu}) \to \min, \tag{11}$$

where

$$\boldsymbol{W}_\alpha(\boldsymbol{\nu}) = \left(V_0[\boldsymbol{B}_\alpha\boldsymbol{\nu}]\right)_2^-. \tag{12}$$

The notation $(\cdot)_2^-$ means computing the generalized inverse after projecting the matrix to a matrix of rank 2 [4]. In Eq. (12), the matrix $V_0[\boldsymbol{B}_\alpha\boldsymbol{\nu}]$ is given in the following form [4]:

$$V_0[\boldsymbol{B}_\alpha\boldsymbol{\nu}] = \bar{\boldsymbol{x}}'_\alpha \times \boldsymbol{A}V_0[\boldsymbol{x}_\alpha]\boldsymbol{A}^\top \times \bar{\boldsymbol{x}}'_\alpha + (\boldsymbol{A}\bar{\boldsymbol{x}}_\alpha) \times V_0[\boldsymbol{x}'_\alpha] \times (\boldsymbol{A}\bar{\boldsymbol{x}}_\alpha) + [V_0[\boldsymbol{x}'_\alpha] \times \boldsymbol{A}V_0[\boldsymbol{x}_\alpha]\boldsymbol{A}^\top]. \tag{13}$$

Here, the vector product $\boldsymbol{A}\times\boldsymbol{a}$ of a 3×3-matrix $\boldsymbol{A}$ and a three-dimensional vector $\boldsymbol{a}$ is a 3×3-matrix defined by

$$\boldsymbol{A} \times \boldsymbol{a} = (\boldsymbol{a} \times \boldsymbol{A}^\top)^\top. \tag{14}$$

The exterior product $[\boldsymbol{A} \times \boldsymbol{B}]$ of 3×3-matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is a 3×3-matrix defined by

$$[\boldsymbol{A} \times \boldsymbol{B}]_{ij} = \sum_{k,l,m,n=1}^{3} \epsilon_{ikl}\epsilon_{jmn}A_{km}B_{ln}, \tag{15}$$

where ϵ_{ijk} is the *Eddington epsilon*, taking values 1, -1, and 0 if (ijk) is obtained from (123) by an even permutation, an odd permutation, and otherwise, respectively. The matrix $V_0[\boldsymbol{B}_\alpha\boldsymbol{\nu}]$ as defined by Eq. (13) should have rank 2 if the matrix $\boldsymbol{A}$ is computed from the true surface parameters $\{\boldsymbol{n}, d\}$ (see the second of Eqs. (2)). However, the values of $\{\boldsymbol{n}, d\}$ used in the course of optimization are not exact, so the matrix $V_0[\boldsymbol{B}_\alpha \boldsymbol{\nu}]$ no longer has rank 2. The operation $(\cdot)_2^-$ is applied to constrain the rank to be 2 in order to prevent numerical instability of the computation [4].

Let $\boldsymbol{\nu}$ be the optimal solution of the minimization (11) under the constraint $\|\boldsymbol{\nu}\| = 1$. It can be shown that the theoretical covariance matrix of the optimal solution $\boldsymbol{\nu}$ has the form

$$V[\boldsymbol{\nu}] = \epsilon^2 \left(\sum_{\alpha=1}^{N} \boldsymbol{P}_{\boldsymbol{\nu}} \boldsymbol{B}_\alpha^\top \boldsymbol{W}_\alpha(\boldsymbol{\nu}) \boldsymbol{B}_\alpha \boldsymbol{P}_{\boldsymbol{\nu}} \right)^-, \tag{16}$$

where $\boldsymbol{P}_{\boldsymbol{\nu}} = \boldsymbol{I} - \boldsymbol{\nu}\boldsymbol{\nu}^\top$. This covariance matrix gives a theoretical bound on attainable accuracy (we omit the proof [4]).

5. Renormalization

If $\boldsymbol{W}_\alpha(\boldsymbol{\nu})$ is replaced by a constant matrix $\boldsymbol{W}_\alpha$, the function $J[\boldsymbol{\nu}]$ in Eq. (11) can be written in the form

$$J[\boldsymbol{\nu}] = (\boldsymbol{\nu}, \boldsymbol{M}\boldsymbol{\nu}), \tag{17}$$

where $\boldsymbol{M}$ is the *moment matrix* defined by

$$\boldsymbol{M} = \frac{1}{N} \sum_{\alpha=1}^{N} \boldsymbol{B}_\alpha^\top \boldsymbol{W}_\alpha \boldsymbol{B}_\alpha. \tag{18}$$

The solution that minimizes Eq. (17) under the constraint $\|\boldsymbol{\nu}\| = 1$ is given by the unit eigenvector for the smallest eigenvalue of $\boldsymbol{M}$. It appears that the optimal solution of Eq. (11) can be obtained by letting $\boldsymbol{W}_\alpha = \boldsymbol{W}_\alpha(\boldsymbol{\nu}_0)$ for an appropriate estimate $\boldsymbol{\nu}_0$ and minimizing Eq. (17). Using the resulting solution $\boldsymbol{\nu}_1$, we can update the weight by letting $\boldsymbol{W}_\alpha = \boldsymbol{W}_\alpha(\boldsymbol{\nu}_1)$ and iterate this process. However, such iterations introduce *statistical bias* into the solution [3]. This is shown as follows.

Define 4×4-matrices $\boldsymbol{N}^{(1)}$ and $\boldsymbol{N}^{(2)}$ by

$$\boldsymbol{N}^{(1)} = \begin{pmatrix} \frac{1}{N}\sum_{\alpha=1}^{N}\left((\boldsymbol{h}, \boldsymbol{X}_\alpha \boldsymbol{h}) V_0[\boldsymbol{x}_\alpha] + (\boldsymbol{h}, \boldsymbol{Y}_\alpha \boldsymbol{h}) \boldsymbol{x}_\alpha \boldsymbol{x}_\alpha^\top\right) & \frac{1}{N}\sum_{\alpha=1}^{N}(V_0[\boldsymbol{x}_\alpha]\boldsymbol{X}_\alpha \boldsymbol{h} + (\boldsymbol{x}_\alpha, \boldsymbol{Y}_\alpha \boldsymbol{h})\boldsymbol{x}_\alpha) \\ \frac{1}{N}\sum_{\alpha=1}^{N}\left((V_0[\boldsymbol{x}_\alpha]\boldsymbol{X}_\alpha \boldsymbol{h})^\top + (\boldsymbol{x}_\alpha, \boldsymbol{Y}_\alpha \boldsymbol{h})\boldsymbol{x}_\alpha^\top\right) & \frac{1}{N}\sum_{\alpha=1}^{N}((V_0[\boldsymbol{x}_\alpha]; \boldsymbol{X}_\alpha) + (\boldsymbol{x}_\alpha, \boldsymbol{Y}_\alpha \boldsymbol{x}_\alpha)) \end{pmatrix},$$

$$\boldsymbol{N}^{(2)} = \begin{pmatrix} \frac{1}{N}\sum_{\alpha=1}^{N}(\boldsymbol{h}, \boldsymbol{Y}_\alpha \boldsymbol{h}) V_0[\boldsymbol{x}_\alpha] & \frac{1}{N}\sum_{\alpha=1}^{N} V_0[\boldsymbol{x}_\alpha]\boldsymbol{Y}_\alpha \boldsymbol{h} \\ \frac{1}{N}\sum_{\alpha=1}^{N}(V_0[\boldsymbol{x}_\alpha]\boldsymbol{Y}_\alpha \boldsymbol{h})^\top & \frac{1}{N}\sum_{\alpha=1}^{N}(V_0[\boldsymbol{x}_\alpha]; \boldsymbol{Y}_\alpha) \end{pmatrix}, \tag{19}$$

where $\boldsymbol{X}_\alpha$ and $\boldsymbol{Y}_\alpha$ are 3×3-matrices defined, respectively, by

$$\boldsymbol{X}_\alpha = \boldsymbol{R}(\boldsymbol{x}'_\alpha \times \boldsymbol{W}_\alpha \times \boldsymbol{x}'_\alpha)\boldsymbol{R}^\top, \tag{20}$$

$$\boldsymbol{Y}_\alpha = \boldsymbol{R}[\boldsymbol{W}_\alpha \times V_0[\boldsymbol{x}'_\alpha]]\boldsymbol{R}^\top, \tag{21}$$

and the inner product $(\boldsymbol{A}; \boldsymbol{B})$ of 3×3-matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined by

$$(\boldsymbol{A}; \boldsymbol{B}) = \sum_{i,j=1}^{3} A_{ij} B_{ij}. \tag{22}$$

It can be shown that the expectation of $\boldsymbol{M}$ has the form

$$E[\boldsymbol{M}] = \bar{\boldsymbol{M}} + \epsilon^2 \bar{\boldsymbol{N}}^{(1)} + \epsilon^4 \boldsymbol{N}^{(2)}, \tag{23}$$

where $\bar{\boldsymbol{M}}$ and $\bar{\boldsymbol{N}}^{(1)}$ are the values of $\boldsymbol{M}$ and $\boldsymbol{N}^{(1)}$ obtained by replacing $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ by their true values $\bar{\boldsymbol{x}}_\alpha$ and $\bar{\boldsymbol{x}}'_\alpha$, respectively, in their definitions. Since $E[\boldsymbol{M}]$ is biased from $\bar{\boldsymbol{M}}$ by $O(\epsilon^2)$, the expectation of the eigenvector $\boldsymbol{\nu}$ of $\boldsymbol{M}$ is also biased from its true value $\bar{\boldsymbol{\nu}}$ by $O(\epsilon^2)$ according to the well known *perturbation theorem* [3].

On the other hand, we can see from Eqs. (19) that

$$E[\boldsymbol{N}^{(1)}] = \bar{\boldsymbol{N}}^{(1)} + 2\epsilon^2 \boldsymbol{N}^{(2)}. \tag{24}$$

Define the *unbiased moment matrix* $\hat{\boldsymbol{M}}$ by

$$\hat{\boldsymbol{M}} = \boldsymbol{M} - \epsilon^2 \boldsymbol{N}^{(1)} + \epsilon^4 \boldsymbol{N}^{(2)}. \tag{25}$$

Then, we have $E[\hat{\boldsymbol{M}}] = \bar{\boldsymbol{M}}$. Hence, we can obtain an unbiased estimator of $\boldsymbol{\nu}$ if we use $\hat{\boldsymbol{M}}$ instead of $\boldsymbol{M}$. However, the noise level ϵ is unknown. If we overestimate or underestimate it, the resulting solution is still biased. In order to resolve this difficulty, we introduce a numerical scheme called *renormalization*, which treats ϵ^2 as an unknown parameter. The procedure for renormalization is described as follows [3], [4], [6]:

1. Let $c = 0$ and $\boldsymbol{W}_\alpha = \boldsymbol{I}$, $\alpha = 1, ..., N$.
2. Compute the moment matrix $\boldsymbol{M}$ defined by Eq. (18).
3. Compute the 4×4-matrices $\boldsymbol{N}^{(1)}$ and $\boldsymbol{N}^{(2)}$ defined by Eqs. (19), and compute the following 4×4-matrix
$$\hat{\boldsymbol{M}} = \boldsymbol{M} - c\boldsymbol{N}^{(1)} + c^2 \boldsymbol{N}^{(2)}. \tag{26}$$
4. Compute the smallest eigenvalue λ of $\hat{\boldsymbol{M}}$ and the corresponding unit eigenvector $\boldsymbol{\nu}$.
5. If $\lambda \approx 0$, return $\boldsymbol{\nu}$, c and $\hat{\boldsymbol{M}}$. Otherwise, update c

and $\boldsymbol{W}_\alpha$ as follows:

$$D = \Big((\boldsymbol{\nu}, \boldsymbol{N}^{(1)}\boldsymbol{\nu}) - 2c(\boldsymbol{\nu}, \boldsymbol{N}^{(2)}\boldsymbol{\nu})\Big)^2 - 4\lambda(\boldsymbol{\nu}, \boldsymbol{N}^{(2)}\boldsymbol{\nu}), \tag{27}$$

if $D \geqq 0$,

$$c \leftarrow c + \frac{(\boldsymbol{\nu}, \boldsymbol{N}^{(1)}\boldsymbol{\nu}) - 2c(\boldsymbol{\nu}, \boldsymbol{N}^{(2)}\boldsymbol{\nu}) - \sqrt{D}}{2(\boldsymbol{\nu}, \boldsymbol{N}^{(2)}\boldsymbol{\nu})},$$

$$\text{if } D < 0, \quad c \leftarrow c + \frac{\lambda}{(\boldsymbol{\nu}, \boldsymbol{N}^{(1)}\boldsymbol{\nu})}, \tag{28}$$

$$\boldsymbol{A} = \boldsymbol{R}^\top(\boldsymbol{h}(\nu_1, \nu_2, \nu_3) + \nu_4\boldsymbol{I}), \tag{29}$$

$$\boldsymbol{W}_\alpha \leftarrow \Big(\boldsymbol{x}'_\alpha \times \boldsymbol{A}V_0[\boldsymbol{x}_\alpha]\boldsymbol{A}^\top \times \boldsymbol{x}'_\alpha + (\boldsymbol{A}\boldsymbol{x}_\alpha) \times V_0[\boldsymbol{x}_\alpha] \times (\boldsymbol{A}\boldsymbol{x}_\alpha) + c[V_0[\boldsymbol{x}'_\alpha] \times \boldsymbol{A}V_0[\boldsymbol{x}'_\alpha]\boldsymbol{A}^\top]\Big)_2^-. \tag{30}$$

6. Go back to Step 2.

If the vector $\boldsymbol{\nu}$ is obtained, we can compute the parameters $\{\boldsymbol{n}, d\}$ of the fitted plane in the form

$$\boldsymbol{n} = N[\begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix}], \quad d = -\frac{\nu_4}{\sqrt{1-\nu_4^2}}, \tag{31}$$

where the symbol $N[\cdot]$ denotes normalization into a unit vector. An unbiased estimator of the squared noise level ϵ^2 is given in the following form [4]:

$$\hat{\epsilon}^2 = \frac{c}{1 - 3/2N}. \tag{32}$$

The covariance matrix $V[\boldsymbol{\nu}]$ given by Eq. (16) is approximated by

$$V[\boldsymbol{\nu}] \approx \frac{\hat{\epsilon}^2}{N}(\hat{\boldsymbol{M}})_3^-. \tag{33}$$

Thus, we can compute by renormalization not only an optimal estimate of $\boldsymbol{\nu}$ but also an estimate of the unknown noise level ϵ and the reliability of the computed estimate $\boldsymbol{\nu}$.

6. Back Projection of Feature Points

After the vector $\boldsymbol{\nu}$ is determined, we correct $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ by the criterion given by Eq. (9) so as to satisfy Eqs. (2). Then, we *back project* the corrected $\hat{\boldsymbol{x}}_\alpha$ and $\hat{\boldsymbol{x}}'_\alpha$ onto the estimated plane. This procedure is carried out as follows [4]:

1. Correct $\boldsymbol{x}_\alpha$ and $\boldsymbol{x}'_\alpha$ in the form

$$\hat{\boldsymbol{x}}_\alpha = \boldsymbol{x}_\alpha - (V_0[\boldsymbol{x}_\alpha]\boldsymbol{A}^\top \times \boldsymbol{x}'_\alpha)\boldsymbol{W}_{0\alpha}(\boldsymbol{x}'_\alpha \times \boldsymbol{A}\boldsymbol{x}_\alpha),$$
$$\hat{\boldsymbol{x}}'_\alpha = \boldsymbol{x}'_\alpha + (V_0[\boldsymbol{x}'_\alpha] \times (\boldsymbol{A}\boldsymbol{x}_\alpha))\boldsymbol{W}_{0\alpha}(\boldsymbol{x}'_\alpha \times \boldsymbol{A}\boldsymbol{x}_\alpha), \tag{34}$$

where the 3×3-matrix $\boldsymbol{W}_{0\alpha}$ is defined by

$$\boldsymbol{W}_{0\alpha} = \Big(\boldsymbol{x}'_\alpha \times \boldsymbol{A}V_0[\boldsymbol{x}_\alpha]\boldsymbol{A}^\top \times \boldsymbol{x}'_\alpha + (\boldsymbol{A}\boldsymbol{x}_\alpha) \times V_0[\boldsymbol{x}'_\alpha] \times (\boldsymbol{A}\boldsymbol{x}_\alpha)\Big)_2^-. \tag{35}$$

2. If $\hat{\boldsymbol{x}}'_\alpha \times \boldsymbol{A}\hat{\boldsymbol{x}}_\alpha \not\approx \boldsymbol{0}$, let $\boldsymbol{x}_\alpha \leftarrow \hat{\boldsymbol{x}}_\alpha$ and $\boldsymbol{x}'_\alpha \leftarrow \hat{\boldsymbol{x}}'_\alpha$, and go back to Step 1.
3. Compute the 3-D position $\boldsymbol{r}_\alpha$ by

$$\boldsymbol{r}_\alpha = \frac{d\hat{\boldsymbol{x}}_\alpha}{(\boldsymbol{n}, \hat{\boldsymbol{x}}_\alpha)}. \tag{36}$$

7. Experiment

7.1 Numerical Simulation

We illustrate the effectiveness of our method by doing numerical simulation. We place a planar grid in a 3-D scene and regard the grid points as feature points. The two cameras are assumed to have the same focal length $f = 600$ (pixels). After projecting the feature points onto the image planes, we add as image noise a Gaussian random number with standard deviation 3 (pixels) to each of the image coordinates independently. Hence, the noise level ϵ is $1/200$, and $V_0[\boldsymbol{x}_\alpha] = V_0[\boldsymbol{x}'_\alpha] = \text{diag}(1, 1, 0)$ (the diagonal matrix with 1, 1, 0 as the diagonal elements in that order). However, the value ϵ is regarded as unknown in the simulation. Figure 2 shows the generated stereo images. The surface reconstructed by our method is shown in Fig. 3. For the sake of comparison, we show the surface reconstructed by the least-squares method (as described in Sect. 1) in Fig. 4. We can observe that our method produces a better result than the least-squares method.

7.2 Analysis of Error Behavior

We define the *error vector* by

$$\Delta\boldsymbol{u} = \boldsymbol{P}_{\bar{\boldsymbol{n}}}(\boldsymbol{n} - \bar{\boldsymbol{n}}) + \frac{d - \bar{d}}{\bar{d}}\bar{\boldsymbol{n}}, \tag{37}$$

where we put $\boldsymbol{P}_{\bar{\boldsymbol{n}}} = \boldsymbol{I} - \bar{\boldsymbol{n}}\bar{\boldsymbol{n}}^\top$ and $\{\bar{\boldsymbol{n}}, \bar{d}\}$ are the true surface parameters. From the theoretical covariance matrix $V[\boldsymbol{\nu}]$ given by Eq. (16), the covariance matrix $V[\boldsymbol{u}]$

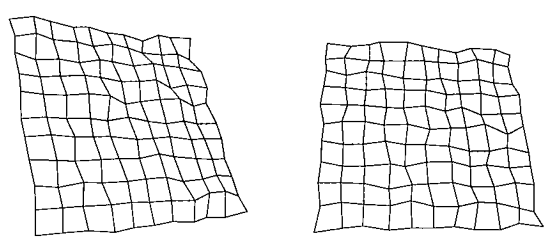
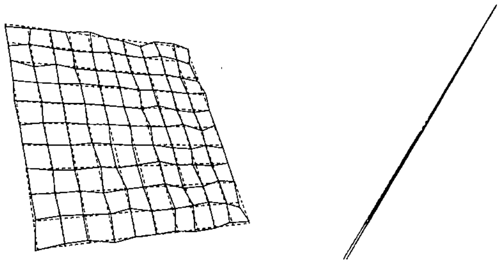
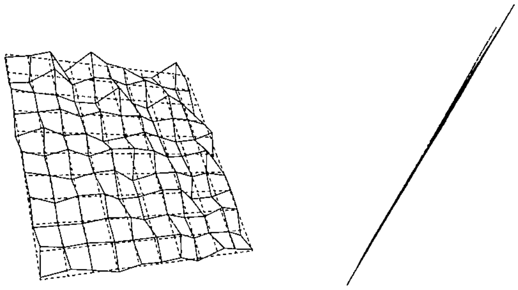


Fig. 2 Simulated stereo images with noise.



(a) (b)

Fig. 3 Planar surface reconstructed by our method: (a) a view from the left camera; (b) a side view.



(a) (b)

Fig. 4 Planar surface reconstructed by the least-squares method: (a) a view from the left camera; (b) a side view.

of the error vector is computed in the following form (we omit the derivation [4]):

$$V[\boldsymbol{u}] = V[\boldsymbol{n}] + \frac{1}{\bar{d}}(V[\boldsymbol{n}, d]\bar{\boldsymbol{n}}^\top + \bar{\boldsymbol{n}}V[\boldsymbol{n}, d]^\top) + \frac{1}{\bar{d}^2}V[d]\bar{\boldsymbol{n}}\bar{\boldsymbol{n}}^\top. \tag{38}$$

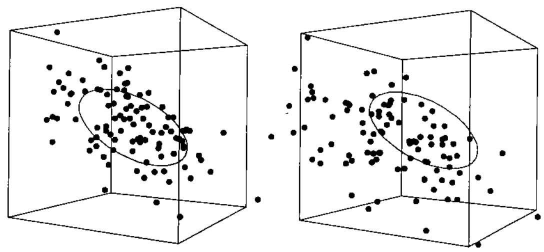
Here, $V[\boldsymbol{n}]$, $V[\boldsymbol{n}, d]$, and $V[d]$ are given as follows:

$$V[\boldsymbol{n}] = (1+\bar{d}^2) \times \boldsymbol{P}_{\bar{\boldsymbol{n}}} \begin{pmatrix} V[\boldsymbol{\nu}]_{11} & V[\boldsymbol{\nu}]_{12} & V[\boldsymbol{\nu}]_{13} \\ V[\boldsymbol{\nu}]_{21} & V[\boldsymbol{\nu}]_{22} & V[\boldsymbol{\nu}]_{23} \\ V[\boldsymbol{\nu}]_{31} & V[\boldsymbol{\nu}]_{32} & V[\boldsymbol{\nu}]_{33} \end{pmatrix} \boldsymbol{P}_{\bar{\boldsymbol{n}}},$$

$$V[\boldsymbol{n}, d] = -(1+\bar{d}^2)^2 \boldsymbol{P}_{\bar{\boldsymbol{n}}} \begin{pmatrix} V[\boldsymbol{\nu}]_{14} \\ V[\boldsymbol{\nu}]_{24} \\ V[\boldsymbol{\nu}]_{34} \end{pmatrix},$$

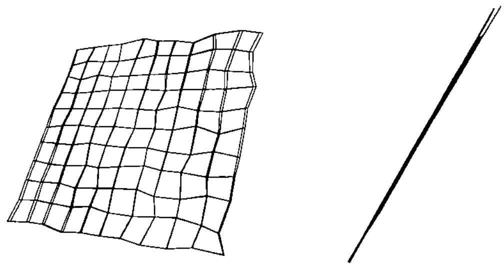
$$V[d] = (1+\bar{d}^2)^3 V[\boldsymbol{\nu}]_{44}. \tag{39}$$

We repeat the computation 100 times, each time using different noise, and plot the error vector three-dimensionally in Fig. 5. The ellipsoids in the figures indicate the theoretical standard deviation in each orientation computed from Eq. (16). We can observe that the solution computed by the least-squares method is statistically biased and that the bias is removed by renormalization. We can also see that the theoretical bound is almost attained; the number of points inside the ellipsoid has approximately the same percentage as theoretically predicted.



(a) (b)

Fig. 5 Distribution of errors: (a) our method; (b) least-squares method.



(a) (b)

Fig. 6 Primary deviation pair: (a) a view from an angle (60° from the optical axis of the first camera); (b) a side view.

7.3 Reliability of 3-D Reconstruction

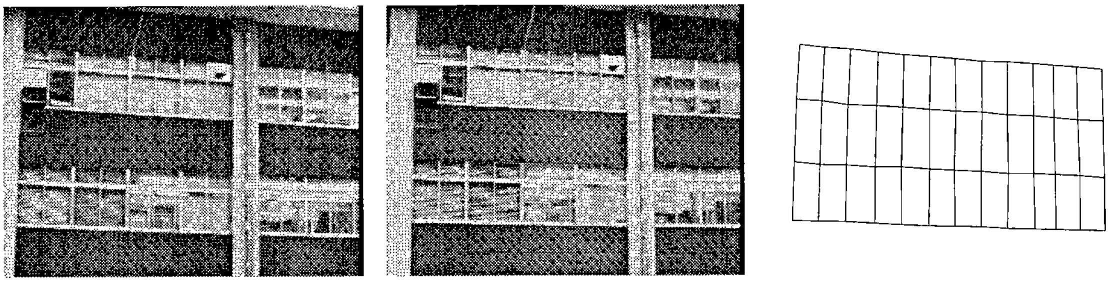
The unit eigenvector $\boldsymbol{\xi}_{\max}$ of the covariance matrix $V[\boldsymbol{\nu}]$ for the largest eigenvalue $\lambda_{\max}$ indicates the orientation of the most likely deviation of $\boldsymbol{\nu}$ from its true value, and $\sqrt{\lambda_{\max}}$ indicates the standard deviation in that orientation. Hence, we can visualize the reliability of the reconstructed planar surface by displaying the two planes represented by the two vectors

$$\boldsymbol{\nu}^+ = N[\hat{\boldsymbol{\nu}} + \sqrt{\lambda_{\max}}\boldsymbol{\xi}_{\max}],$$
$$\boldsymbol{\nu}^- = N[\hat{\boldsymbol{\nu}} - \sqrt{\lambda_{\max}}\boldsymbol{\xi}_{\max}]. \tag{40}$$

These two planes indicate the most likely deviation of the reconstructed planar surface. We call them the *primary deviation pair.* Since the covariance matrix $V[\boldsymbol{\nu}]$ can be computed by the approximation (33) from the data alone, the primary deviation pair can be computed without any knowledge of the true noise level. The primary deviation pair computed from Fig. 2 is shown in Fig. 6.

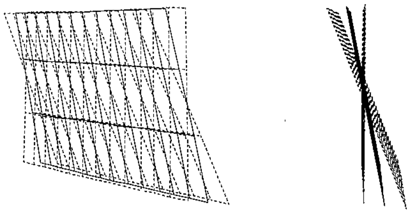
7.4 Real-Image Example

Figure 7 (a) shows real stereo images. Figure 7 (b) shows a grid pattern defined by feature points (corners of the windows) extracted from the left image. The motion parameters $\{\boldsymbol{h}, \boldsymbol{R}\}$ are computed by the *optimal camera calibration system* [5]. In Fig. 8, the reconstructed grid pattern is displayed in solid lines, and



(a) (b)

Fig. 7 (a) Real stereo images. (b) Feature points extracted from the left image.



(a) (b)

Fig. 8 Reconstructed planar surface and its primary deviation pair: (a) a view from an angle (75° from the optical axis of the first camera); (b) a side view.

the primary deviation pair is displayed in dashed lines. This example demonstrates that we can visualize the reliability of 3-D reconstruction without any knowledge of the magnitude of the image noise. In this experiment, the distance between the two cameras is very short as compared with the distance to the building surface (approximately 1/16). Since the noise level ϵ is estimated from the degree to which Eq. (5) is not satisfied, the error in the motion parameters $\{\boldsymbol{h}, \boldsymbol{R}\}$ is also treated as "image noise". Hence, the reliability of this 3-D reconstruction is very low.

8. Conclusion

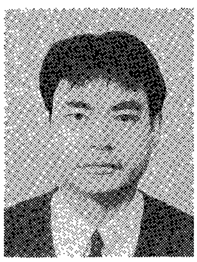
We have presented a direct reconstruction method for reconstructing a planar surface by stereo vision. Our method can not only reconstruct an optimal estimate but also allows us to evaluate the reliability of the computed estimate quantitatively. By doing numerical simulation, we have shown that the obtained solution almost attains the theoretical bound on accuracy. Our method can be applied to many applications in real environments, such as recognizing walls and ceilings in robotics workspaces. The ability to evaluate the reliability of 3-D reconstruction has a great significance in such applications.

Acknowledgment

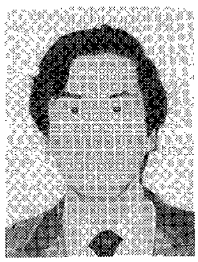
This work was in part supported by the Ministry of Education, Science, and Culture, Japan and under the Grant in Aid for Scientific Research B (No.07458067) and the Okawa Institute of Information and Telecommunication.

References

[1] Faugeras, O., *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press, Cambridge, MA, U.S.A., 1993.
[2] Kanatani, K., *Geometric Computation for Machine Vision*, Oxford Univ. Press, Oxford, U.K., 1993.
[3] Kanatani, K., "Renormalization for unbiased estimation," *Proc. 4th Int. Conf. Comput. Vision (ICCV'93)*, Berlin, F.R.G., pp.599–606, May 1993.
[4] Kanatani, K., *Statistical Optimization for Geometric Computation: Theory and Practice*, Lecture Note, Department of Computer Science, Gunma University, Apr. 1994.
[5] Kanatani, K. and Maruyama, T., "Optimal focal length calibration system," *Proc. IEEE/RSJ Int. Conf. Intell. Robots Sys. (IROS'93)*, Yokohama, Japan, pp.1816–1821, Jul. 1993.
[6] Kanatani, K., "Statistical analysis of geometric computation," *CVGIP: Image Understanding*, vol.59, no.3, pp.286–306, 1994.
[7] Weng, J., Huang, T.S. and Ahuja, N., *Motion and Structure from Image Sequences*, Springer, Berlin, F.R.G., 1993.



Yasushi Kanazawa received his B.E. and M.E. degrees in information and computer science from Toyohashi University of Technology, Aichi, Japan, in 1985 and 1987, respectively. He is currently Assistant Professor of information and computer engineering at Gunma College of Technology. His research interests include image processing and computer vision.



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).