Optimal Two-View Planar Scene Triangulation

Kenichi Kanatani and Hirotaka Niitsuma

Department of Computer Science, Okayama University, Okayama 700-8530, Japan
{kanatani,niitsuma}@suri.cs.okayama-u.ac.jp

Abstract. We present a new algorithm for optimally computing from point correspondences over two images their 3-D positions when they are constrained to be on a planar surface. We consider two cases: the case in which the plane and camera parameters are known and the case in which they are not. In the former, we show how observed point correspondences are optimally corrected so that they are compatible with the homography between the two images determined by the plane and camera parameters. In the latter, we show how the homography is optimally estimated by iteratively using the triangulation procedure.

1   Introduction

Computing the 3-D position of a point from its projection in two images is called *triangulation* and is a fundamental tool of computer vision [4]. The basic principle is to compute the intersection of the rays starting from the camera lens center and passing through the corresponding image points. However, point correspondence detection using an image processing operation incurs errors to some extent, and the two rays may not intersect. A naive solution is to compute the midpoint of the shortest segment connecting the two rays (Fig. 1(a)), but Kanatani [7] and Hartley and Sturm [5] pointed out that for optimal estimation the corresponding points should be displaced so that the rays meet in the scene (Fig. 1(b)) in such a way that the sum of the square displaced distances, or *reprojection error*, is minimized. For this, Hartley and Sturm [5] presented an algorithm that reduces to solving a 6th degree polynomial, while Kanazawa and Kanatani [13] gave a first approximation in an analytical form. Later, Kanatani et al. [11] showed that the first approximation is sufficiently accurate and that a few iterations lead to complete agreement with the Hartley-Sturm solution with far more efficiency. Lindstrom [14] further improved this approach.

The aim of this paper is to demonstrate that exactly the same holds when the points we are viewing are constrained to be on a planar surface. This is a common situation in indoor and urban scenes. If a 3-D point is constrained to be on a known plane, the corresponding points must be displaced so that the rays not merely intersect but also meet on that plane. We call this *planar triangulation* after Chum et al. [3]. A first approximation solution was given by Kanazawa and Kanatani [12], while Chum et al. [3] presented an algorithm that reduces to solving an 8th degree polynomial. The purpose of this paper is to demonstrate that a few iterations of the first approximation lead to an
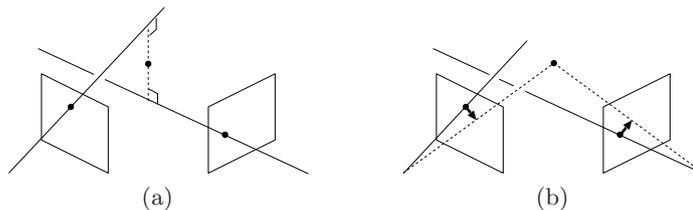
Fig. 1. Triangulation. (a) The midpoint of the shortest segment connecting the rays. (b) The points are optimally corrected so that their rays intersect.

optimal solution. We consider two cases: the case in which the plane and camera parameters are known, and the case in which they are not. The algorithm of Chum et al. [3] deals with the former. The latter case could be solved using the bundle adjustment approach, as demonstrated by Bartoli and Sturm [1] for an arbitrary number of images, but we show that a much simpler method exists. In fact, we show that our optimal triangulation procedure for the former case can easily be extended to the latter.

In Sec. 2, we summarize the fundamentals about planar projection and homographies. In Sec. 3 and 4, we present an iterative algorithm for optimal planar triangulation for known plane and camera parameters. In Sec. 5, we show that our optimal triangulation procedure can be straightforwardly extended to the case of unknown plane and camera parameters. In Sec. 6, we do numerical simulation and demonstrate that our algorithm and the first approximation of Kanazawa and Kanatani [12] give practically the same value.

2 Planar Surface and Homography

Consider a plane with a unit surface normal n at distance d from the origin of an XYZ coordinate system fixed to the scene (Fig. 2(a)). We take images of this plane from two positions. The ith camera, $i = 1, 2$, is translated from the coordinate origin O by t_i after rotated by R_i (Fig. 2(b)). We call $\{t_i, R_i\}$ the *motion parameters* of the ith camera. We assume that by prior camera
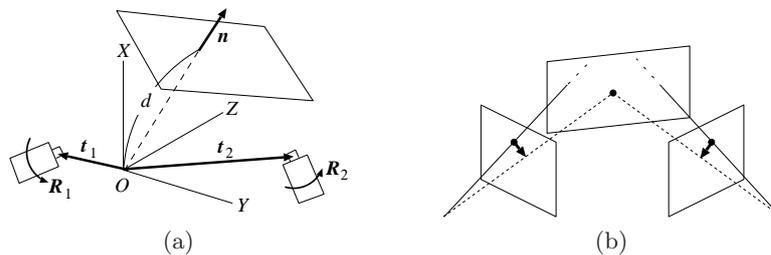
Fig. 2. (a) Plane and camera configuration. (b) The points are optimally corrected so that their rays intersect on the plane.
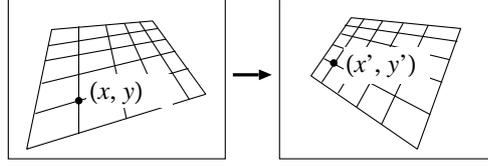
Fig. 3. Two images of the same planar surface are related by a homography.

calibration the image coordinate origin is placed at the principal point and that the aspect ratio is 1 with no image skew.

The image of the plane taken from the first position, call it the "first image", and the image taken from the second position, call it the "second image", are related by the following *homography* (Fig. 3) [4, 7]:

$$x' = f_0 \frac{h_{11}x + h_{12}y + h_{13}f_0}{h_{31}x + h_{32}y + h_{33}f_0}, \qquad y' = f_0 \frac{h_{21}x + h_{22}y + h_{23}f_0}{h_{31}x + h_{32}y + h_{33}f_0}. \qquad (1)$$

Here, f_0 is a scale factor of approximately the size of the image for stabilizing numerical computation with finite length. The 3×3 matrix $\boldsymbol{H} = (h_{ij})$ is determined by the parameter $\{\boldsymbol{n}, d\}$ of the plane, the motion parameters $\{\boldsymbol{R}_i, \boldsymbol{t}_i\}$ and the focal lengths f_i, $i = 1, 2$, in the following form [4, 7]:

$$\boldsymbol{H} = \text{diag}(1, 1, \frac{f_0}{f_2})\boldsymbol{R}_2^\top (\boldsymbol{I} - \frac{\boldsymbol{t}_2\boldsymbol{n}^\top}{d})(\boldsymbol{I} + \frac{\boldsymbol{t}_1\boldsymbol{n}^\top}{d - (\boldsymbol{t}_1, \boldsymbol{n})})\boldsymbol{R}_1\text{diag}(1, 1, \frac{f_1}{f_0}). \qquad (2)$$

Here, $\boldsymbol{I}$ is the unit matrix, and $\text{diag}(a, b, c)$ denotes the diagonal matrix with diagonal elements a, b, and c in that order. Throughout this paper, we denote the inner product of vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ by $(\boldsymbol{a}, \boldsymbol{b})$.

3 Triangulation for Known Plane and Cameras

We first consider the case in which we know $\{\boldsymbol{n}, d\}$, $\{\boldsymbol{R}_i, \boldsymbol{t}_i\}$, and f_i, $i = 1$, 2, hence the homography $\boldsymbol{H}$. In homogeneous coordinates, we can write (1) as follow [4, 7]:

$$\begin{pmatrix} x'/f_0 \\ y'/f_0 \\ 1 \end{pmatrix} \cong \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x/f_0 \\ y/f_0 \\ 1 \end{pmatrix}. \qquad (3)$$

The symbol $\cong$ denotes equality up to a nonzero constant. We can equivalently write (3) as

$$\begin{pmatrix} x'/f_0 \\ y'/f_0 \\ 1 \end{pmatrix} \times \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix} \begin{pmatrix} x/f_0 \\ y/f_0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \qquad (4)$$

The three components of this equation multiplied by f_0^2 are

$$(\boldsymbol{\xi}^{(1)}, \boldsymbol{h}) = 0, \qquad (\boldsymbol{\xi}^{(2)}, \boldsymbol{h}) = 0, \qquad (\boldsymbol{\xi}^{(3)}, \boldsymbol{h}) = 0, \qquad (5)$$
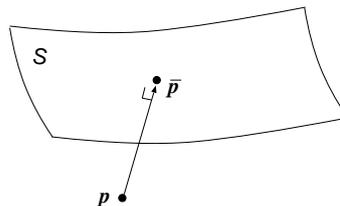
Fig. 4. The point $\boldsymbol{p}$ is orthogonally projected on to $\bar{\boldsymbol{p}}$ on $\mathcal{S}$ in the 4-D joint space.

where we define the 9-D vectors $\boldsymbol{h}$, $\boldsymbol{\xi}^{(1)}$, $\boldsymbol{\xi}^{(2)}$, and $\boldsymbol{\xi}^{(3)}$ by

$$\boldsymbol{h} = \left(h_{11} \; h_{12} \; h_{13} \; h_{21} \; h_{22} \; h_{23} \; h_{31} \; h_{32} \; h_{33} \right)^\top ,$$
$$\boldsymbol{\xi}^{(1)} = \left(0 \; 0 \; 0 \; -f_0 x \; -f_0 y \; -f_0^2 \; xy' \; yy' \; f_0 y' \right)^\top ,$$
$$\boldsymbol{\xi}^{(2)} = \left(f_0 x \; f_0 y \; f_0^2 \; 0 \; 0 \; 0 \; -xx' \; -yx' \; -f_0 x' \right)^\top ,$$
$$\boldsymbol{\xi}^{(3)} = \left(-xy' \; -yy' \; -f_0 y' \; xx' \; yx' \; f_0 x' \; 0 \; 0 \; 0 \right)^\top . \tag{6}$$

A corresponding point pair (x, y) and (x', y') can be identified with a point $\boldsymbol{p}$ $= (x, y, x', y')^\top$ in the 4-D $xyx'y'$ joint space. Each of the three equations in (5) defines a hypersurface in this 4-D joint space. However, the identity $x' \boldsymbol{\xi}^{(1)} + y' \boldsymbol{\xi}^{(2)} + f_0 \boldsymbol{\xi}^{(3)} = \boldsymbol{0}$ holds, so (5) defines a 2-D variety (algebraic manifold) $\mathcal{S}$ in the 4-D joint space.

In the presence of noise, the point $\boldsymbol{p}$ is not necessarily on $\mathcal{S}$. Optimal planar triangulation is to displace $\boldsymbol{p}$ to a point $\bar{\boldsymbol{p}}$ on $\mathcal{S}$ in such a way that the *reprojection error*

$$E = \|\boldsymbol{p} - \bar{\boldsymbol{p}}\|^2 , \tag{7}$$

is minimized subject to

$$(\boldsymbol{\xi}^{(k)}(\bar{\boldsymbol{p}}), \boldsymbol{h}) = 0, \qquad k = 1, 2, 3. \tag{8}$$

Geometrically, this means *orthogonally projecting $\boldsymbol{p}$ onto the variety $\mathcal{S}$ in the 4-D joint space* (Fig. 4). Once such a $\bar{\boldsymbol{p}} = (\bar{x}, \bar{y}, \bar{x}', \bar{y}')^\top$ is obtained, the corresponding 3-D position (X, Y, Z) is determined by solving

$$\begin{pmatrix} \bar{x}/f_0 \\ \bar{y}/f_0 \\ 1 \end{pmatrix} \cong \boldsymbol{P}_1 \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} \bar{x}'/f_0 \\ \bar{y}'/f_0 \\ 1 \end{pmatrix} \cong \boldsymbol{P}_2 \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}, \tag{9}$$

where $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ are the 3×4 projection matrices defined as follows [4, 11]:

$$\boldsymbol{P}_1 = \operatorname{diag}(1, 1, \frac{f_0}{f_1}) \boldsymbol{R}_1^\top \left(\boldsymbol{I} \; -\boldsymbol{t}_1 \right), \quad \boldsymbol{P}_2 = \operatorname{diag}(1, 1, \frac{f_0}{f_2}) \boldsymbol{R}_2^\top \left(\boldsymbol{I} \; -\boldsymbol{t}_1 \right). \tag{10}$$

Four linear equations in X, Y, and Z are obtained from (9), but because (8) is satisfied, a unique solution is obtained [11].

4 Optimal Planar Triangulation

We now present a new procedure for minimizing (7) subject to (8). While unconstrained triangulation [11] involves a single constraint describing the epipolar geometry, planar triangulation is constrained by three equations in the form of (8), not mutually algebraically independent. For this, the first approximation has already been presented by Kanazawa and Kanatani [12]. We modify their method so that an optimal solution is obtained by iterations. The procedure is as follows (the derivation is omitted; it is a straightforward extension of the unconstrained triangulation in [11]):

1. Let $E_0 = \infty$ (a sufficiently large number), and define the 4-D vectors

$$\boldsymbol{p} = \begin{pmatrix} x \\ y \\ x' \\ y' \end{pmatrix}, \qquad \hat{\boldsymbol{p}} = \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{x}' \\ \hat{y}' \end{pmatrix}, \qquad \tilde{\boldsymbol{p}} = \begin{pmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{x}' \\ \tilde{y}' \end{pmatrix}, \qquad (11)$$

   where we let $\hat{x} = x$, $\hat{y} = y$, $\hat{x}' = x'$, $\hat{y}' = y'$, and $\tilde{x} = \tilde{y} = \tilde{x}' = \tilde{y}' = 0$.

2. Compute the following 9×4 matrices $\boldsymbol{T}^{(1)}$, $\boldsymbol{T}^{(2)}$, and $\boldsymbol{T}^{(3)}$:

$$\boldsymbol{T}^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -f_0 & 0 & 0 & 0 \\ 0 & -f_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hat{y}' & 0 & 0 & \hat{x} \\ 0 & \hat{y}' & 0 & \hat{y} \\ 0 & 0 & 0 & f_0 \end{pmatrix}, \quad \boldsymbol{T}^{(2)} = \begin{pmatrix} f_0 & 0 & 0 & 0 \\ 0 & f_0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\hat{x}' & 0 & -\hat{x} & 0 \\ 0 & -\hat{x}' & -\hat{y} & 0 \\ 0 & 0 & -f_0 & 0 \end{pmatrix}, \quad \boldsymbol{T}^{(3)} = \begin{pmatrix} -\hat{y}' & 0 & 0 & -\hat{x} \\ 0 & -\hat{y}' & 0 & -\hat{y} \\ 0 & 0 & 0 & -f_0 \\ \hat{x}' & 0 & \hat{x} & 0 \\ 0 & \hat{x}' & \hat{y} & 0 \\ 0 & 0 & f_0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{12}$$

3. Compute the following $\boldsymbol{\xi}^{(1)*}$, $\boldsymbol{\xi}^{(2)*}$, and $\boldsymbol{\xi}^{(3)*}$:

$$\boldsymbol{\xi}^{(1)*} = \begin{pmatrix} 0 & 0 & 0 & -f_0\hat{x} & -f_0\hat{y} & -f_0^2 & \hat{x}\hat{y}' & \hat{y}\hat{y}' & f_0\hat{y}' \end{pmatrix}^\top + \boldsymbol{T}^{(1)}\tilde{\boldsymbol{p}},$$

$$\boldsymbol{\xi}^{(2)*} = \begin{pmatrix} f_0\hat{x} & f_0\hat{y} & f_0^2 & 0 & 0 & 0 & -\hat{x}\hat{x}' & -\hat{y}\hat{x}' & -f_0\hat{x}' \end{pmatrix}^\top + \boldsymbol{T}^{(2)}\tilde{\boldsymbol{p}},$$

$$\boldsymbol{\xi}^{(3)*} = \begin{pmatrix} -\hat{x}\hat{y}' & -\hat{y}\hat{y}' & -f_0\hat{y}' & \hat{x}\hat{x}' & \hat{y}\hat{x}' & f_0\hat{x}' & 0 & 0 & 0 \end{pmatrix}^\top + \boldsymbol{T}^{(3)}\tilde{\boldsymbol{p}}. \tag{13}$$

4. Compute the following 9×9 matrices $V_0^{(kl)}[\boldsymbol{\xi}]$:

$$V_0^{(kl)}[\boldsymbol{\xi}] = \boldsymbol{T}^{(k)}\boldsymbol{T}^{(l)\top}. \tag{14}$$

5. Compute the 3×3 matrix $\boldsymbol{W} = (W^{(kl)})$

$$\boldsymbol{W} = \begin{pmatrix} (\boldsymbol{h}, V_0^{(11)}[\boldsymbol{\xi}]\boldsymbol{h}) & (\boldsymbol{h}, V_0^{(12)}[\boldsymbol{\xi}]\boldsymbol{h}) & (\boldsymbol{h}, V_0^{(13)}[\boldsymbol{\xi}]\boldsymbol{h}) \\ (\boldsymbol{h}, V_0^{(21)}[\boldsymbol{\xi}]\boldsymbol{h}) & (\boldsymbol{h}, V_0^{(22)}[\boldsymbol{\xi}]\boldsymbol{h}) & (\boldsymbol{h}, V_0^{(23)}[\boldsymbol{\xi}]\boldsymbol{h}) \\ (\boldsymbol{h}, V_0^{(31)}[\boldsymbol{\xi}]\boldsymbol{h}) & (\boldsymbol{h}, V_0^{(32)}[\boldsymbol{\xi}]\boldsymbol{h}) & (\boldsymbol{h}, V_0^{(33)}[\boldsymbol{\xi}]\boldsymbol{h}) \end{pmatrix}_2^-, \tag{15}$$

   where $(\cdot)_r^-$ denotes pseudoinverse constrained to rank r (the smallest eigenvalue is replaced by 0 in its spectral decomposition).

6. Update $\tilde{\boldsymbol{p}}$ and $\hat{\boldsymbol{p}}$ as follows:

$$\tilde{\boldsymbol{p}} = \sum_{k,l=1}^{3} W^{(kl)}(\boldsymbol{\xi}^{(l)*}, \boldsymbol{h})\boldsymbol{T}^{(k)\top}\boldsymbol{h}, \qquad \hat{\boldsymbol{p}} \leftarrow \boldsymbol{p} - \tilde{\boldsymbol{p}}. \qquad (16)$$

7. Evaluate the reprojection error E by

$$E = \|\tilde{\boldsymbol{p}}\|^2. \qquad (17)$$

If $E \approx E_0$, then return $\hat{\boldsymbol{p}}$ as $\bar{\boldsymbol{p}}$ and stop. Else, let $E_0 \leftarrow E$ and go back to Step 2.

The use of the pseudoinverse constrained to rank 2 in (15) reflects the fact that only two of the three constraints in (5) are algebraically independent. This algorithm produces the same solution as that of Chum et al. [3], which solves an 8th degree polynomial. However, our algorithm involves only a few iterations of linear calculus without requiring any polynomial solver, which is sometimes inefficient and numerically unstable.

5   Triangulation for Unknown Plane and Cameras

Next, we consider the case in which $\{\boldsymbol{n}, d\}$, $\{\boldsymbol{R}_i, \boldsymbol{t}_i\}$, $i = 1$, 2, are unknown. As is well known [4, 7], these parameters can be estimated by computing the homography $\boldsymbol{H}$ between the two images, provided that the focal lengths f_i are known; we assume that they are given by prior camera calibration. Since the analytical procedure for computing the plane and camera parameters from a homography $\boldsymbol{H}$ has been described by many researchers in different forms [6, 15, 18–20], the problem reduces to computing the homography $\boldsymbol{H}$ from point correspondences over the two images. The simplest method is to minimize the algebraic distance, which is known by many names such as *least squares* and *DLT* (*Direct Linear Transformation*) [4], but the accuracy is low in the presence of noise. A method known to be very accurate is what is called *Sampson error minimization* [4], and an iterative scheme was presented by Scoleri et al. [17]. However, Sampson error minimization does not necessarily compute an exactly optimal solution in the sense of maximum likelihood [9].

We now show that by combining the Sampson error minimization with the optimal planar triangulation described in Sec. 4, we can obtain an exactly optimal $\boldsymbol{H}$. The basic principle is already presented by Kanatani and Sugaya [9], but their theory applies only for a single constraint, which is the case in ellipse fitting and fundamental matrix computation. Here, we extend their procedure to homographies constrained by multiple equations. Given N corresponding points (x_α, y_α) and (x'_α, y'_α), $\alpha = 1, ..., N$, we compute the 9-D vector $\boldsymbol{h}$ that encodes the homography $\boldsymbol{H}$. Omitting the derivation, we describe the procedure:

1. Let $E_0 = \infty$ (a sufficiently large number), and give an initial guess of $\boldsymbol{h}$ in (6) using any method, say least squares.

2. Let $\boldsymbol{p}_\alpha$, $\hat{\boldsymbol{p}}_\alpha$, and $\tilde{\boldsymbol{p}}_\alpha$ be the vectors in (11) for the αth pair (x_α, y_α), and (x'_α, y'_α), $\alpha = 1, ..., N$.

3. Let $\boldsymbol{T}^{(1)}_\alpha$, $\boldsymbol{T}^{(2)}_\alpha$, and $\boldsymbol{T}^{(3)}_\alpha$ be, respectively, the values of $\boldsymbol{T}^{(1)}$, $\boldsymbol{T}^{(2)}$, and $\boldsymbol{T}^{(3)}$ in (12) for the αth pair, $\alpha = 1, ..., N$.

4. Compute the 9×9 matrices $V_0^{(kl)}[\boldsymbol{\xi}_\alpha] = \boldsymbol{T}^{(k)}_\alpha \boldsymbol{T}^{(l)\top}_\alpha$ and the 3×3 matrices $\boldsymbol{W}_\alpha = (W^{(kl)}_\alpha)$ in (15) for the αth pair, and let the 9-D vectors $\boldsymbol{\xi}^{(1)*}_\alpha$, $\boldsymbol{\xi}^{(2)*}_\alpha$, and $\boldsymbol{\xi}^{(3)*}_\alpha$ be the values of $\boldsymbol{\xi}^{(1)*}$, $\boldsymbol{\xi}^{(2)*}$, and $\boldsymbol{\xi}^{(3)*}$ in (13) for the αth pair, $\alpha = 1, ..., N$.

5. Compute the 9-D unit vector $\boldsymbol{h}$ that minimizes

$$J = \frac{1}{N} \sum_{\alpha=1}^{N} \sum_{k,l=1}^{3} W^{(kl)}_\alpha (\boldsymbol{\xi}^{(k)*}_\alpha, \boldsymbol{h})(\boldsymbol{\xi}^{(l)*}_\alpha, \boldsymbol{h}). \tag{18}$$

6. Update $\tilde{\boldsymbol{p}}_\alpha$ and $\hat{\boldsymbol{p}}_\alpha$, $\alpha = 1, ..., N$, as follows:

$$\tilde{\boldsymbol{p}}_\alpha = \sum_{k,l=1}^{3} W^{(kl)}_\alpha (\boldsymbol{\xi}^{(l)*}_\alpha, \boldsymbol{h}) \boldsymbol{T}^{(k)\top}_\alpha \boldsymbol{h}, \qquad \hat{\boldsymbol{p}}_\alpha \leftarrow \boldsymbol{p}_\alpha - \tilde{\boldsymbol{p}}_\alpha. \tag{19}$$

7. Evaluate the reprojection error $E = \sum_{\alpha=1}^{N} \|\tilde{\boldsymbol{p}}_\alpha\|^2$. If $E \approx E_0$, then return $\boldsymbol{h}$ and stop. Else, let $E_0 \leftarrow E$ and go back to Step 3.

This procedure is identical to the optimal planar triangulation in Sec. 4 except for Step 5. The expression in (18) coincides with what is known as the *Sampson error* [4] if $\boldsymbol{\xi}^{(k)*}_\alpha$ and $\boldsymbol{\xi}^{(l)*}_\alpha$ on the right-hand side are respectively replaced by $\boldsymbol{\xi}^{(k)}_\alpha$ and $\boldsymbol{\xi}^{(l)}_\alpha$ (the values of (6) for the αth pair). It can be minimized by the scheme of Scoleri et al. [17], but here we use a much simpler reformulation of Kanatani et al. [16], which is a direct extension of the *FNS* (*Fundamental Numerical Scheme*) of Chojnacki et al. [2]. The procedure goes as follows:

1. Provide an initial value $\boldsymbol{h}_0$ for $\boldsymbol{h}$, e.g., by least squares.
2. Compute the matrices $\boldsymbol{M}$ and $\boldsymbol{L}$ as follows:

$$\boldsymbol{M} = \frac{1}{N} \sum_{\alpha=1}^{N} W^{(kl)}_\alpha \boldsymbol{\xi}^{(k)*}_\alpha \boldsymbol{\xi}^{(l)*\top}_\alpha, \quad \boldsymbol{L} = \frac{1}{N} \sum_{\alpha=1}^{N} \sum_{k,l=1}^{N} v^{(k)}_\alpha v^{(l)}_\alpha V_0^{(kl)}[\boldsymbol{\xi}_\alpha], \tag{20}$$

where we define

$$v^{(k)}_\alpha = \sum_{l=1}^{3} W^{(kl)}_\alpha (\boldsymbol{\xi}^{(l)*}_\alpha, \boldsymbol{h}). \tag{21}$$

3. Solve the eigenvalue problem

$$(\boldsymbol{M} - \boldsymbol{L})\boldsymbol{h} = \lambda \boldsymbol{h}, \tag{22}$$

and compute the unit eigenvector $\boldsymbol{h}$ for the smallest eigenvalue λ.

4. If $\boldsymbol{h} \approx \boldsymbol{h}_0$, return $\boldsymbol{h}$ and stop. Else, let $\boldsymbol{h}_0 \leftarrow \boldsymbol{h}$, and go back to Step 2.

For unconstrained triangulation, the optimal algorithm of Kanatani et al. [11], which assumes a given fundamental matrix, can be automatically converted to optimal fundamental matrix computation merely by inserting a Sampson error minimization step, as shown by Kanatani and Sugaya [8]. In contrast, the polynomial solving algorithm of Hartley and Sturm [5] cannot be so easily converted to optimal fundamental matrix computation. Similarly, the optimal planar triangulation in Sec. 4, which assumes a given homography, can be automatically converted to optimal homography computation merely by inserting a Sampson error minimization step. In contrast, the polynomial solving algorithm of Chum et al. [3] cannot be so easily converted to optimal homography computation.

6   Experiments

Figure 5(a) shows two images of a simulated grid. The image size is 500×500 pixels; the focal lengths are $f_1 = f_2 = 600$ pixels. We added Gaussian noise of mean 0 and standard deviation 1 pixel to the x and y coordinates of each of the N ($= 121$) grid points. Then, we reconstructed the 3-D position of each grid point by unconstrained triangulation [11] and by our planar triangulation. Figure 5(b) shows the 3-D positions of the grid points. We can see that by planar triangulation (in blue) all the points are on the specified plane but not by unconstrained triangulation (in red). For quantitative evaluation, we measured the root mean square reprojection error
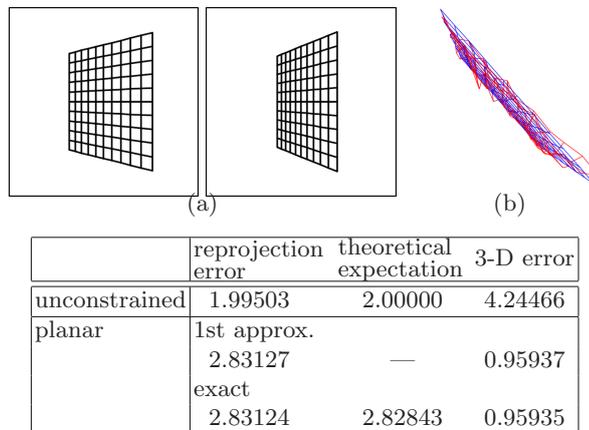
$$e = \sqrt{\frac{1}{N} \sum_{\alpha=1}^{N} \Big((\hat{x}_\alpha - x_\alpha)^2 + (\hat{y}_\alpha - y_\alpha)^2 + (\hat{x}'_\alpha - x'_\alpha)^2 + (\hat{y}'_\alpha - y'_\alpha)^2 \Big)}, \quad (23)$$

where $(\hat{x}_\alpha, \hat{y}_\alpha)$ and $(\hat{x}'_\alpha, \hat{y}'_\alpha)$ are the corrected positions of the observations (x_α, y_α) and (x'_α, y'_α), respectively. We also evaluated the 3-D reconstruction error
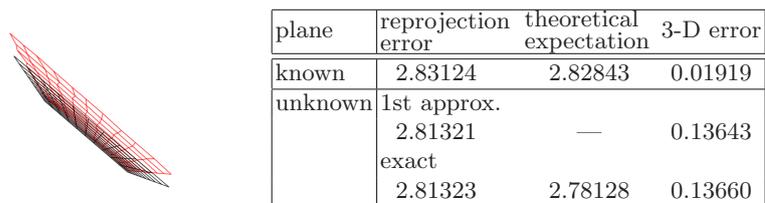
$$D = \sqrt{\frac{1}{N} \sum_{\alpha=1}^{N} \| \hat{\boldsymbol{r}}_\alpha - \bar{\boldsymbol{r}}_\alpha \|^2}, \quad (24)$$

where $\hat{\boldsymbol{r}}_\alpha$ is the reconstructed position of the αth point, and $\bar{\boldsymbol{r}}_\alpha$ its true position. The table in Fig. 5 lists the values for unconstrained triangulation [11], the first approximation of the planar triangulation (the iteration is terminated after the first round), which corresponds to the result of Kanazawa and Kanatani [12], and the exact values computed by our method.

From this table, we observe that the reprojection error e *increases* by assuming planarity. This is because the corresponding points need to be displaced so that the rays not simply intersect but also intersect on the specified plane. Statistical analysis [7] tells us that under maximum likelihood Ne^2/σ^2 is subject to a χ^2 distribution with N degree of freedom for unconstrained triangulation and with $2N$ degrees of freedom for planar triangulation. Hence, e should be approximately σ and $\sqrt{2}\sigma$, respectively. The values in the table in Fig. 5 are very

(a)   (b)

	reprojection error	theoretical expectation	3-D error
unconstrained	1.99503	2.00000	4.24466
planar	1st approx.		
	2.83127	—	0.95937
	exact		
	2.83124	2.82843	0.95935

Fig. 5. (a) Simulated images of a planar grid taken from different places. (b) 3-D position of the reconstructed grid. Points reconstructed by planar triangulation (in blue) are on the specified plane, but those reconstructed by unconstrained triangulation [11] (in red) are not necessarily on it. The table below lists the reprojection error, its theoretical expectation, and the average 3-D reconstruction error.

plane	reprojection error	theoretical expectation	3-D error
known	2.83124	2.82843	0.01919
unknown	1st approx.		
	2.81321	—	0.13643
	exact		
	2.81323	2.78128	0.13660

Fig. 6. The reconstructed grid by estimating the plane and the camera positions (in red) and its true position (in black). The table lists the reprojection error of planar triangulation by estimating the plane and the camera positions, its theoretical expectation, and the average 3-D reconstruction error.

close to the prediction. However, the increase in the reprojection error e does *not* mean the increase in the 3-D reconstruction error D. In fact, the 3-D reconstruction error D actually decreases with the knowledge of planarity. We also see that the exact values are very close to the first approximation of Kanazawa and Kanatani [12].

Next, we tested the case in which the plane and camera parameters are unknown. Since the absolute scale is indeterminate, we scaled the relative displacement between the cameras to unit length. Figure 6 shows the 3-D positions of the reconstructed grid (in red) and its true position (in black). Due to the error in estimating the plane, i.e., the homography, the computed position is slightly different from its true position. The table in Fig. 6 compares the reprojection error e and the 3-D reconstruction error D in the known and unknown plane cases. The values in the known plane case are the same as in the table in Fig 5 except the normalization $\|\boldsymbol{t}\| = 1$.

We observe that the reprojection error e is *smaller* in the unknown plane case than in the known plane case. This is because the parameters of the plane are estimated so that the reprojection error is minimized. Statistical analysis [7] tells us that under maximum likelihood Ne^2/σ^2 is subject to a χ^2 distribution with $2N - 8$ degrees of freedom and hence has expectation $2N - 8$. This reflects the fact that the homography constraint has eight degrees of freedom with codimension two [7]. Consequently, the reprojection error e should approximately be $\sqrt{2(1 - 4/N)}\sigma$. The value in the table in Fig 5 is very close to the prediction. We also see that the first approximation (using only a single Sampson error minimization step) and the exact maximum likelihood value are very close to each other, as generally predicted in [9]. Again, the smaller reprojection error does not mean more accurate 3-D reconstruction. Rather, the 3-D reconstruction accuracy deteriorates because of the error in estimating the plane, as shown in the table in Fig. 5.

Note that when the camera positions are unknown, the 3-D positions of the points cannot be reconstructed without the knowledge of planarity. If the points are in general position, their 3-D positions and the camera positions can be reconstructed from two views [10], but that computation fails if the points degenerate to be coplanar [4, 7].

7   Concluding Remarks

We have presented an optimal algorithm[1] for computing the 3-D positions of points viewed from two images by using the knowledge that they are constrained to be on a planar surface. This is an extension of the unconstrained triangulation of Kanatani et al. [11] without assuming planarity. Our algorithm automatically encompass the case in which the plane and camera parameters are unknown; they are estimated merely by inserting a Sampson error minimization step. As a result, an exact maximum likelihood estimate is obtained for the homography between the two images.

This is a complete parallel to the scheme of Kanatani and Sugaya [8] for computing an exact maximum likelihood estimate of the fundamental matrix between two images merely by inserting a Sampson error minimization step in the unconstrained triangulation of Kanatani et al. [11]. In contrast, the optimal triangulation of Hartley and Sturm [5], which solves a 6th degree polynomial, is not so easily converted to optimal fundamental matrix estimation. Similarly, the optimal planar triangulation of Chum et al. [3], which solves an 8th degree polynomial, is not so easily converted to produce an optimal homography.

We have also confirmed experimentally that the first approximation is very close to the exact maximum likelihood estimate. Thus, we conclude that our optimal scheme is not really necessary in practice. In fact, that the Sampson error minimization solution is known to coincide with the exactly optimal solution up to several significant digits in many problems [9]. In our experiment, we used the

---

[1] The code is available at `http://www.suri.cs.okayama-u.ac.jp/~kanatani/e/`.

least squares, also known as the DLT, to compute the initial value $\boldsymbol{h}$ to start the FNS procedure described in Sec. 5. We have observed that the Sampson error minimization iterations may not converge in the presence of extremely large noise and that the use of "HyperLS" or its "Taubin approximation" [16] can significantly extend the noise level range of convergence.

Our optimal homography computation does not reach a global minimum of the reprojection error if the Sampson error minimization in Step 5 does not return a global minimum of the Sampson error. In fact, the FNS procedure described in Sec. 5 is not theoretically guaranteed to return a global minimum $\boldsymbol{h}$ of the Sampson error J, although it usually does. However, if the Sampson error J can be globally minimized, e.g., using branch and bound, in each iteration, we can obtain a global minimum solution $\boldsymbol{h}$ in the end, because the Sampson error J coincides with the reprojection error when the iterations have converged [9].

References

1. Bartoli, A., P. Sturm, P.: Constrained structure and motion from multiple un-calibrated views of a piecewise planar scene, Int. J. Comput. Vis. 52(1), 45-2003 (2003)
2. Chojnacki, W., Brooks, M.J., van den Hengel, A., Gawley, D.: On the fitting of surfaces to data with covariances, IEEE Trans. Patt. Anal. Mach. Intell. 22(11), 1294–1303 (2000)
3. Chum, O., Pajdla, T., Sturm, P.: The geometric error for homographies, Comput. Vis. Im. Understand. 97(1), 86-102 (2002)
4. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. 2nd ed., Cambridge University Press, Cambridge (2004)
5. Hartley, R.I., Sturm, P.: Triangulation, Comput. Vision Image Understand. 68(2), 146–157 (1997)
6. Hay, J.C.: Optical motions and space perception—An extension of Gibson's analysis, Psych. Rev. 73(6), 550–565 (1966)
7. Kanatani, K.: Statistical Optimization for Geometric Computation: Theory and Practice. Elsevier, Amsterdam (1996); reprinted, Dover, York (2005)
8. Kanatani, K., Sugaya, Y.: Compact fundamental matrix computation, IPSJ Trans. Comp. Vis. Appl. 2, 59–70 (2010)
9. Kanatani, K., Sugaya, Y.: Unified computation of strict maximum likelihood for geometric fitting, J. Math. Imaging Vis. 38(1), 1–13 (2010)
10. Kanatani, K., Sugaya, Y., Kanazawa, Y.: Latest algorithms for 3-D reconstruction form two views. In: Chen, C.H. (ed.), Handbook of Pattern Recognition and Computer Vision, 4th ed., World Scientific, Singapore, pp. 201–234 (2009)
11. Kanatani, K., Sugaya, Y., Niitsuma, H.: Triangulation from two views revisited: Hartley-Sturm vs. optimal correction. In: Proc. 9th British Machine Vis. Conf., September 2008, pp. 173–182 (2008)
12. Kanazawa, Y. Kanatani, Y.: Direct reconstruction of planar surfaces by stereo vision, IEICE Trans. Inf. & Syst. E78-D(7), 917–922 (1995)
13. Kanazawa, Y., Kanatani, K.: Reliability of 3-D reconstruction by stereo vision, IEICE Trans. Inf. & Syst. E78-D(10), 1301–1306 (1995)

14. Lindstrom, P.: Triangulation made easy. In: Proc. IEEE Conf. Comput. Vis. Patt. Recog., June 2010 (2010)
15. Longuet-Higgins, H.C.: The reconstruction of a planar surface from two perspective projections, Proc. Royal Soc. Lond. B-227(1249), 399-410 (1986)
16. Niitsuma, H., Rangarajan, P., Kanatani, K.: High accuracy homography computation without iterations. In: Proc. 16th Symp. Sensing via Imaging Information, June 2010 (2010)
17. Scoleri, T., Chojnacki, W., Brooks, M.J.: A multi-objective parameter estimator for image mosaicing. In: Proc. 8th Int. Symp. Signal Processing and its Applications, August 2005, vol. 2, pp. 551–554 (2005)
18. Tsai, R.Y., Huang, T.S.: Estimation of three-dimensional motion parameters of a rigid planar patch, IEEE Trans. Acoustics Speech Signal Process. 29(6), 1147–1152 (1981)
19. Tsai, R.Y., Huang, T.S.: Estimation of three-dimensional motion parameters of a rigid planar patch III: Finite point correspondence and the three-view problem, IEEE Trans. Acoustics Speech Signal Process. 32(2), 213–220 (1984)
20. Tsai, R.Y., Huang, T.S., Zhu, W.-L.: Estimation of three-dimensional motion parameters of a rigid planar patch II: Singular value decomposition, IEEE Trans. Acoustics Speech Signal Process. 30(4), 525–544 (1982)