ELSEVIER

Image and Vision Computing 14 (1996) 379–388

image and vision COMPUTING

Statistical estimation for exterior orientation from line-to-line correspondences

Chung-Nan Lee[a], Robert M. Haralick[b]

[a]*Institute of Information and Computer Engineering, National Sun Yat-Sen University, Kaohsiung, Taiwan 80424, ROC*
[b]*Intelligent Systems Laboratory, Department of Electrical Engineering • FT-10, University of Washington, Seattle, WA 98195, USA*

Received 15 May 1995; revised 21 August 1995; accepted 25 August 1995

Abstract

This paper presents a statistical estimation from which a new objective function for exterior orientation from line correspondences is derived. The objective function is based on the assumption that the underlying noise model for the line correspondences is the Fisher distribution. The assumption is appropriate for 3D orientation, is different from the underlying noise models for k pixels positions, and allows us to do a consistent estimation of the unknown parameters. The objective function gives two important facts: its formulation and concept is different for that of previous work, and it automatically estimates six unknown parameters simultaneously. As a result, it provides an optimal solution and better accuracy. We design an experimental protocol to evaluate the performance of the new algorithm. The results of each experiment shows that the new algorithm produces answers whose errors are 10%–20% less than the competing decoupled least squares algorithm.

Keywords: Exterior orientation; Line-to-line correspondence; Objective functions

1. Introduction

The problem of determining the orientation and position of an object in a 3D world coordinate system relative to a 3D camera coordinate system is equivalent to solving the relation between the 3D object features and their corresponding perspective projection features. It is an important problem both in computer vision and in photogrammetry.

Most methods use point features to obtain the transformation function, which is governing the orientation and position of an object [1–8]. Besides point-to-point correspondences, line-to-line correspondences can be used to obtain the transformation function. Lowe [9], Kumar and Hanson [10, 11] and Liu et al. [12] use line-to-line correspondence to determine a transformation function. Related application using line information include: estimation of motion and structure [13, 14], object recognition [15] and exterior orientation [16].

The paper is organized in the following way. We first give a review on a traditional least square approach in Section 2. Section 3 discusses the statistical estimation of unknown parameters of position and orientation of a camera. We make the assumption that the underlying noise model is the Fisher distribution [18, 19] for the line correspondences based on the observation and characteristics of the line correspondences. In Section 4 we design an experimental protocol to carry out the performance characterization of the algorithms. Section 5 provides the results and discussion. Finally, Section 6 gives summary. Before we go to next section, we give a problem statement.

Problem definition

Suppose $l_1, l_2, \ldots, l_n$ are known n line segments in a 3D world coordinate system whose observed randomly perturbed 2D perspective projections are $l_{i1}, l_{i2}, \ldots, l_{in}$. The problem of exterior orientation using line correspondences is to determine a transformation function that governs the relationship between $l_1, l_2, \ldots, l_n$ and $l_{i1}, l_{i2}, \ldots, l_{in}$.

2. Decoupled least square approach

The decoupled least square approach is proposed in Ref. [12]. It decouples the transformation function into a rotation matrix and a translation vector which are

SSDI 0262-8856(95)01058-0

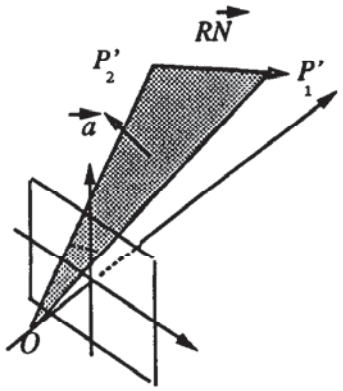

Fig. 1. Interpretation plane passing through the center of perspective, the 2D image line, and the 3D line after transformation.

successively determined to reduce the computational complexity. This approach has the advantage of computational efficiency, but sacrifices some numerical accuracy due to not estimating six unknown parameters simultaneously.

Fig. 1 shows the camera coordinate system, the 3D line after transformation, the center of perspective, and the 2D image line. As we can see the center of perspectivity, the 2D image line which is the perspective projection of the 3D line, and 3D line itself are on the same plane which is called the interpretation plane. Let $\vec{a}$ be a unit normal vector of the interpretation plane. Then $\vec{a}$ will be perpendicular to the 3D line. It gives

$$\vec{a}^t(Rp + T) = 0 \tag{2.1a}$$

$$\vec{a}^t R\vec{N} = 0 \tag{2.1b}$$

where

$$\vec{N} = \begin{pmatrix} l \\ m \\ n \end{pmatrix}$$

is the vector of direction cosines of the 3D line before transformation, p is an arbitrary point on the 3D line, R is a 3 by 3 orthonormal rotation matrix (i.e. $RR^t = I$) and T is a translation vector.

Actually Eq. (2.1a) and Eq. (2.1b) are not independent of each other. We can derive Eq. (2.1b) from Eq. (2.1a). In Fig. 1 we let O be the origin of the camera coordinate system, P_1' and P_2' be any two points on the 3D line after transformation and N the direction cosine of the 3D line. Without noise the vector $\vec{a}$ is perpendicular to $R\vec{N}$, $\vec{OP}_1$, $\vec{OP}_2$, i.e. Eqs. (2.1a) and (2.1b). We can represent $R\vec{N}$ as

$$R\vec{N} = \frac{\vec{OP}_1' - \vec{OP}_2'}{\|\vec{OP}_1' - \vec{OP}_2'\|}$$

Since $\vec{a}$ and the direction cosines of 3D line are known, Eq. (2.1a) is an implicit form for the three unknown parameters of R. Similarly, Eq. (2.1b) is an implicit form for the six unknown parameters of R and T.

A minimum of three correspondences is required to solve the problem for both Eqs. (2.1a) and (2.1b), because two points from each line can be used in Eq. (2.1b).

2.1. Estimation of unknown parameters

When the observation contains noise, Eqs. (2.1a) and (2.1b) are no longer satisfied. To infer R and T from n noisy observations, the traditional least squares approach makes an assumption that the noisy is the ideal $\vec{a}_i^t R\vec{N}_i$ plus additive independent identically distributed Gaussian distribution, then it minimizes the following error function (objective function):

$$F_1 = \epsilon^t \epsilon = \sum_{i=1}^{n} (\vec{a}_i^{*t} R\vec{N}_i)^2 \tag{2.2}$$

Let

$$\Psi = \begin{pmatrix} \omega \\ \phi \\ \kappa \end{pmatrix},$$

then $R(\Psi)$ is a nonlinear function of Ψ. Then, applying a linearization procedure, Eq. (2.2) becomes

$$F_1 = \epsilon^t \epsilon = \sum_{i=1}^{n} \left(\vec{a}_i^{*t} \left(R(\Psi^k) + \frac{\partial R(\Psi^k)}{\partial \omega} \Delta\omega + \frac{\partial R(\Psi^k)}{\partial \phi} \Delta\phi + \frac{\partial R(\Psi^k)}{\partial \kappa} \Delta\kappa \right) \vec{N}_i \right)^2 \tag{2.3}$$

where Ψ^k is the kth iteration of Ψ.

Taking a partial derivative of F_1 with respect to $\Delta\omega$, $\Delta\phi$ and $\Delta\kappa$ in Eq. (2.3) and setting it to zero, we obtain

$$A^t A \Delta\Psi = -A^t B \tag{2.4}$$

where A is an $n \times 3$ matrix

$$A = \begin{pmatrix} a_{11}^{*k} & a_{12}^{*k} & a_{13}^{*k} \\ a_{21}^{*k} & a_{22}^{*k} & a_{23}^{*k} \\ & \vdots & \\ a_{n1}^{*k} & a_{n2}^{*k} & a_{n3}^{*k} \end{pmatrix}$$

$\Delta\Psi$ is a 3×1 parameter vector

$$\Delta\Psi = \begin{pmatrix} \Delta\omega \\ \Delta\phi \\ \Delta\kappa \end{pmatrix}$$

B is an $n \times 1$ vector

$$B = \begin{pmatrix} a_{14}^{*k} \\ a_{24}^{*k} \\ \cdots \\ a_{n4}^{*k} \end{pmatrix}$$

and the elements of A and B are computed as follows:

$$a_{i1}^{*k} = \vec{a}_i^t \frac{\partial R(\Psi^k)}{\partial \omega} \vec{N}_i$$

$$a_{i2}^{*k} = \vec{a}_i^t \frac{\partial R(\Psi^k)}{\partial \phi} \vec{N}_i$$

$$a_{i3}^{*k} = \vec{a}_i^t \frac{\partial R(\Psi^k)}{\partial \kappa} \vec{N}_i$$

$$a_{i4}^{*k} = \vec{a}_i^t R(\Psi^k) \vec{N}_i$$

Eq. (2.4) can be solved by a singular value decomposition in the least squares sense. Once the estimated rotation matrix $\hat{R}$ is obtained, we can minimize the following error function:

$$F_2 = \epsilon^t \epsilon = \sum_{i=1}^{n} \sum_{j=1}^{2} \{\vec{a}_i^{*t}(RP_i^j + T)\}^2$$

Substituting the estimated rotation matrix for the above equation and taking derivative of F_2 with respect to T, it gives

$$CT = D$$

where $C = [a_1^* a_1^* a_2^* a_2^* \ldots a_n^* a_n^*]^t$ and $D = [-\vec{a}_1^{*t} RP_1^1 - \vec{a}_1^{*t} RP_1^2 - \vec{a}_2^{*t} RP_2^1 \ldots - \vec{a}_n^{*t} RP_n^1 - \vec{a}_n^{*t} RP_n^2]$. Again, the above equation can be solved by a singular value decomposition in the least squares sense.

2.2. Discussion

Although the estimation procedure mentioned above is commonly used, it has two problems: one is that the underlying noise for the objective function is assumed to come from a Gaussian distribution, but this is not necessarily correct; the other is that the unknown parameters are not estimated simultaneously; in other word, the estimation is not an optimal one, since the error in the rotation matrix calculation is propagated to the translation vector calculation.

3. Statistical estimation

Statistical estimation to unknown parameters of the exterior orientation involves the posterior distribution and the prior distribution of the unknown parameters and an underlying noise model. Many papers related to applications of the new algorithm include: image analysis [20, 21], pose estimation [22] and object recognition [23]. In this section, we first discuss the underlying noise model. Then, we obtain the posterior distribution of the unknown parameters. Finally, we estimate the unknown parameters by the maximum a posterior.

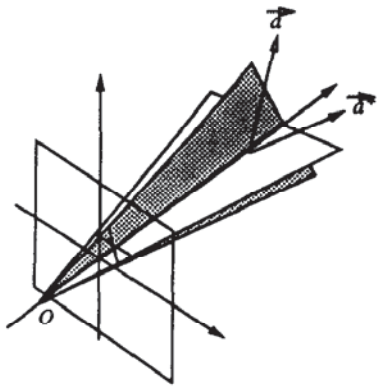

Fig. 2. Effect of the true unit normal vector $\vec{a}$ by the noise on the 2D image line segment.

3.1. Noise model

To estimate unknown parameters in a consistent way we require an underlying noise model. What is an appropriate one for line parameters? The chosen one should be as close as reasonably possible to the real observed noise. To find out what kind of observation noise we may encounter, we represent a 2D line as

$$\cos\xi u + \sin\xi v - d = 0 \tag{3.1}$$

where d is the perpendicular distance from the line to the origin and ξ is the angle between the perpendicular and the u-axis. The focal length is set to one. Then, the unit normal vector of the interpretation plane is

$$\vec{a} = \frac{1}{\sqrt{1+d^2}} \begin{pmatrix} \cos\xi \\ \sin\xi \\ d \end{pmatrix} \tag{3.2}$$

Noise may affect the values of both ξ and d. As a result, the observed unit normal vector $\vec{a}^*$ may deviate from the unit normal vector $\vec{a}$, as shown in Fig. 2, in any direction depending on how both ξ and d are changed. We can see that the observed unit normal vector $\vec{a}^*$ most likely appears in a conical region

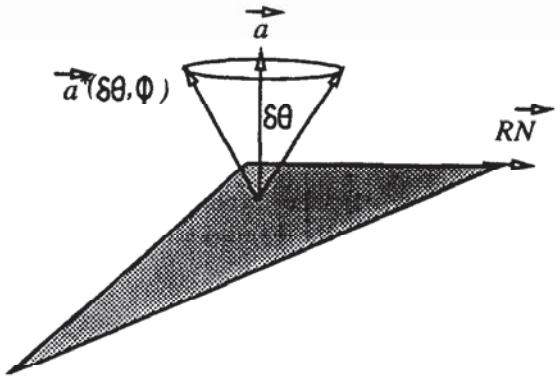

Fig. 3. Distribution of the perturbed $\vec{a}$ is most likely in a cone with the true unit normal vector $\vec{a}$ as the axis.

along the true unit normal vector as axis with solid angle $\delta\theta$ as shown in Fig. 3, which illustrates the observed normal $\vec{a}^*$, the solid angle, and a 2D line in the image plane and its corresponding interpretation plane. We expect that the probability density is high when the solid angle is small and is low when the solid angle is large.

Based on the above discussion, and the nature of a unit normal vector which is directional data, we make the assumption that the underlying noise model is the Fisher distribution for line-to-line correspondence.

The Fisher distribution is one kind of spherical distribution. The density function of the Fisher distribution with mean direction along an arbitrary vector (λ, μ, ν) in spherical coordinates is expressed as

$$g(l,m,n) = c(k_c)\, e^{k_c(l\lambda + m\mu + n\nu)}, \quad k_c > 0 \tag{3.3}$$

where (l,m,n) is defined on the surface of the sphere with unit radius and center at the origin and $c(k_c) = k_c/2\pi(2\sin hk_c)$. The parameter k_c is called as the 'concentration' parameter. For large k_c the distribution is clustered around the mean direction. If the vector (λ, μ, ν) is the polar axis we have

$$g(l,m,n) = c(k_c)\, e^{k_c\cos\theta}, \quad 0 < \theta < \pi,$$
$$0 < \phi < 2\pi, \quad k_c > 0 \tag{3.4}$$

Let colatitude angle θ and latitude angle ϕ be random variables on the surface of the unit sphere and

$$\begin{pmatrix} l \\ m \\ n \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix}$$

Because the differential element of integration on the unit sphere is $\sin\theta d\theta d\phi$, we can rewrite the probability density function with the differential element for the Fisher distribution as follows:

$$\Pr(\theta < \theta_1 < \theta + \delta\theta, \phi < \phi_1 < \phi + \delta\phi)$$
$$= c(k_c)\, e^{k_c\cos\theta}\sin\theta d\theta d\phi, \quad 0 < \theta < \pi, 0 < \phi < 2\pi$$

Taking the $d\theta d\phi$ as the probability measure, the probability density for the θ and ϕ is

$$g(\theta,\phi) = c(k_c) e^{k_c\cos\theta}\sin\theta,$$
$$0 < \theta < \pi, 0 < \phi < 2\pi, k_c > 0 \tag{3.5}$$

The density function described by the above equation for different values of k_c is shown in Fig. 4.

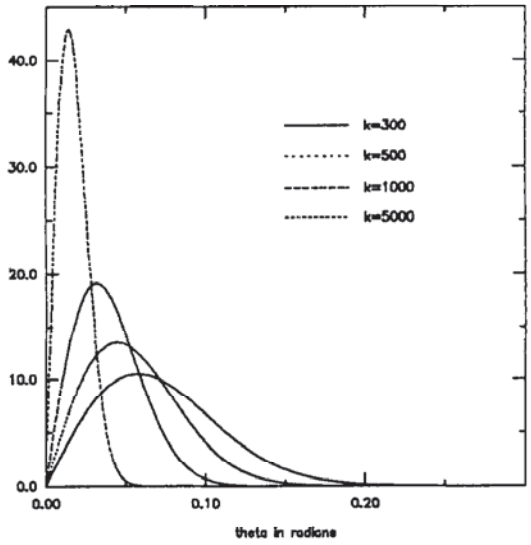

Fig. 4. The probability density of θ for the Fisher distribution for $k_c = 300, 500, 1000$ and 5000.

3.2. Posterior distribution

In this subsection we derive a posterior distribution for the unknown parameters. Let

$$\Phi = \begin{pmatrix} \omega \\ \phi \\ k \\ t_x \\ t_y \\ t_z \end{pmatrix}$$

be the unknown parameter vector. Now, given the observations $\{\vec{a}_i^* | i = 1, 2, \ldots, n\}$, whose relation with 2D lines are expressed in Eq. (3.1), and the 3D lines $\{l_i | i = 1, 2, \ldots, n\}$. We wish to find the most probable value of Φ,

$$P(\Phi|\vec{a}_1^*, \vec{a}_2^*, \ldots, \vec{a}_n^*, l_1, l_2, \ldots, l_n)$$
$$= \frac{P(\vec{a}_1^*, \vec{a}_2^*, \ldots, \vec{a}_n^*|\Phi, l_1, l_2, \ldots, l_n)P(\Phi, l_1, l_2, \ldots, l_n)}{P(\vec{a}_1^*, \vec{a}_2^*, \ldots, \vec{a}_n^*, l_1, l_2, \ldots, l_n)}$$

Because $\vec{a}_i^*$s are conditionally independent on Φ and l_i's, and Φ is also conditionally independent on l_i's, we have

$$P(\Phi|\vec{a}_1^*, \vec{a}_2^*, \ldots, \vec{a}_n^*, l_1, l_2, \ldots, l_n)$$
$$= \frac{\left[\prod_i P(\vec{a}_i^*|\Phi, l_i)\right] P(\Phi)}{P(\vec{a}_1^*, \vec{a}_2^*, \ldots, \vec{a}_n^*|l_1, l_2, \ldots, l_n)} \tag{3.6}$$

where $P(\Phi)$ is the prior distribution. Though the calculation of the exact value is very complicated, it is

a constant which does not depend on Φ. Thus we may rewrite equation (3.6) as

$$P(\Phi|\vec{a}_1^*, \vec{a}_2^*, \ldots, \vec{a}_n^*, l_1, l_2, \ldots, l_n) = c\left[\prod_i P(\vec{a}_i^*|\Phi, l_i)\right] P(\Phi)$$

where c is the constant.

3.3. Estimating the unknown parameter

After the posterior distribution obtained, we want to choose Φ to maximize the distribution, i.e. maximize *a posterior*,

$$\text{Maximize } P(\Phi|\vec{a}_1^*, \vec{a}_2^*, \ldots, \vec{a}_n^*, l_1, l_2, \ldots, l_n)P(\Phi) \tag{3.7}$$

or

$$\prod_i P(\vec{a}_i^*|\Phi, l_i)P(\Phi) \tag{3.8}$$

As stated earlier, our underlying noise model is the Fisher distribution. Therefore,

$$P(\vec{a}_i^*|\Phi, l_i) = c(k_c)e^{k_c \cos\delta\theta_i}, i = 1, 2, \ldots, n$$
$$0 < \delta\theta_i < \pi, 0 < \Phi < 2\pi, k_c > 0 \tag{3.9}$$

Using the relationships of trigonometric functions, Eq. (3.9) becomes

$$P(\vec{a}_i^*|\Phi, l_i) = c(k_c)e^{k_c(1-2\sin^2\frac{\delta\theta_i}{2})},$$
$$0 < \delta\theta_i < \pi, 0 < \phi < 2\pi, k_c > 0 \tag{3.10}$$

From Fig. 3 we have

$$\sin\frac{\delta\theta_i}{2} = \frac{1}{2}\|\vec{a}_i^* - \vec{a}\| \tag{3.11}$$

Substituting Eq. (3.11) into Eq. (3.10), it becomes

$$P(\vec{a}_i^*|\Phi, l_i) = c(k_c)e^{k_c(1-\frac{1}{2}\|\vec{a}_i^* - \vec{a}_i\|^2)} \tag{3.12}$$

Now we can rewrite Eq. (3.8) as

$$\prod_i P(\vec{a}_i^*|\Phi, l_i)P(\Phi) = \prod_i c(k_c)e^{k_c(1-\frac{1}{2}\|\vec{a}_i^* - \vec{a}_i\|^2)}P(\Phi) \tag{3.13}$$

Upon taking logarithms of the above equation there results

$$\ln\prod_i P(\vec{a}_i^*|\Phi, l_i) = \sum_{i=1}^n \times \left\{\ln c(k_c) + k_c\left(1 - \frac{1}{2}\|\vec{a}_i^* - \vec{a}_i\|^2\right)\right\} + \ln P(\Phi) \tag{3.14}$$

Since the prior distribution plays an important role in the estimation, we would like to give a reasonable assumption. In the exterior orientation problem the domain of unknown parameter vector, Φ is in $[0, 2\pi] \times [0, 2\pi] \times [0, \pi] \times R^3$. Usually we translate the object within a finite space. Hence, we can assume that the domain of Φ is in $[0, 2\pi] \times [0, 2\pi] \times [0, \pi] \times [-x, x] \times [-y, y] \times [-z, z]$; here x, y and z are arbitrary values. Without any preference, Φ should be uniformly distributed over the domain. Hence, we can always assume the prior distribution $P(\Phi)$ is a constant. Thus, maximizing Eq. (3.14) is equivalent to determining Φ to minimize

$$\sum_{i=1}^n \frac{k_c}{2}\|\vec{a}_i^* - \vec{a}_i(\Phi)\|^2 \tag{3.15}$$

Eq. (3.15) is quite different from Eq. (2.2) both in concept and formula. The object function in Eq. (2.2) minimizes the error of the dot product between the observed unit normal vector and the direction cosines of 3D line. However, the objective function in Eq. (3.15) is to minimize the norm distance between the observed unit normal vector and the true unit normal vector of the plane by which the 3D line lies. The objective function used in Eq. (2.2) just considers the rotation parameters only, and cannot give an optimal solution. On the contrary, the objective function in Eq. (3.15) solves the six parameters simultaneously and will give a global optimal solution.

To obtain the maximum likelihood estimation, we take the partial derivative of Eq. (3.15) with respect to Φ. This results in

$$\sum_{i=1}^n \sum_{j=1}^3 \frac{k_c}{2}(a_i^{j^*} - a_i^j)\frac{\partial a^j}{\partial\Phi} = 0 \tag{3.16}$$

where a_i^j is the jth component of $\vec{a}_i$ and is a function of Φ. We again apply a linearization procedure to linearize $a_i^j(\Phi^h + \Delta\Phi)$ and omit the higher order terms; we obtain

$$\sum_{i=1}^n \sum_{j=1}^3 \frac{k_c}{2}(a_i^{j^*} - a_i^j(\Phi^h) - a_i^j(\Phi^h)'^t\Delta\Phi)a_i^j(\Phi^h)' = 0 \tag{3.17}$$

where h denotes the hth iteration.

In matrix notation, Eq. (3.17) can be rewritten as

$$M^t M\Delta\Phi = M^t E \tag{3.18}$$

where M is a $3n$ by 6 matrix and E is a $3n$ by 1 matrix, and each of the elements in two matrices is represented as follows:

$$M = \begin{pmatrix} \frac{\partial a_1^1(\Phi^h)}{\partial\omega} & \frac{\partial a_1^1(\Phi^h)}{\partial\kappa} & \frac{\partial a_1^1(\Phi^h)}{\partial\phi} & \frac{\partial a_1^1(\Phi^h)}{\partial t_x} & \frac{\partial a_1^1(\Phi^h)}{\partial t_y} & \frac{\partial a_1^1(\Phi^h)}{\partial t_z} \\ \frac{\partial a_1^2(\Phi^h)}{\partial\omega} & \frac{\partial a_1^2(\Phi^h)}{\partial\kappa} & \frac{\partial a_1^2(\Phi^h)}{\partial\phi} & \frac{\partial a_1^2(\Phi^h)}{\partial t_x} & \frac{\partial a_1^2(\Phi^h)}{\partial t_y} & \frac{\partial a_1^2(\Phi^h)}{\partial t_z} \\ & \vdots & & \vdots & & \\ \frac{\partial a_n^3(\Phi^h)}{\partial\omega} & \frac{\partial a_n^3(\Phi^h)}{\partial\kappa} & \frac{\partial a_n^3(\Phi^h)}{\partial\phi} & \frac{\partial a_n^3(\Phi^h)}{\partial t_x} & \frac{\partial a_n^3(\Phi^h)}{\partial t_y} & \frac{\partial a_n^3(\Phi^h)}{\partial t_s} \end{pmatrix}$$

$$E = \begin{pmatrix} a_1^{1^*} - a_1^1 \\ a_1^{2^*} - a_1^2 \\ a_1^{3^*} - a_1^3 \\ a_2^{1^*} - a_2^1 \\ \vdots \\ a_n^{3^*} - a_n^3 \end{pmatrix}$$

We calculate the incremental value $\Delta\Phi$, and update $\Phi^{h+1} = \Phi^h + \Delta\Phi$ until the criterion is satisfied.

4. Experimental protocol

4.1. Generation of simulated data

To evaluate the algorithms, we arbitrarily generate the corresponding 3D and 2D line segments by giving segment midpoints, orientations and lengths uniformly over the image. If the length of each image side is s, the lengths are uniform over the range $[s/50, s/10]$. The orientation is uniform over the range $[0, 2\pi]$.

Each of the line segments backprojects to an infinite triangle in 3D. To determine the corresponding 3D line segment, we backproject the end points of image line segment. We randomly choose a magnitude between 30 and 70. The end point of the 3D line is obtained by multiplying the unit vector that is passing the center of perspectivity and the end point of image line segment by the magnitude. The focal length is taken to be one. Two end points determine the 3D line segment in the camera coordinates. To determine the line segments in the 3D object coordinate system, we generate a random rotation and translation (6 degrees of freedom). The range of Φ is in $[15°, 45°] \times [30°, 60°] \times [45°, 75°] \times [-20, 20] \times [-20, 20] \times [-20, 20]$ for the simulated data. The unit for the translation vector is the same as focal length. Then, we use the Φ to transform the 3D lines from the 3D camera coordinate system to the world coordinate system.

The noise is generated by using the transformation method [24] to obtain the Fisher distribution from a uniform distribution in [0,1]. We verify the correctness of the noise generator by plotting the probability density of the 100,000 generated data of θ for different k_c from the noise generator. The density of the simulated data, which is almost the replication of the theoretical density in Fig. 4, is shown in Fig. 5. The quantitative test gives the same result.

The procedure for the data generation is:

1. Generate corresponding image and world line segments.
2. Add noise to the image line segments by adding a small perturbation to the true unit normal vector.

4.2. Performance characterization

To evaluate the performance of the least squares solution and our algorithm, we run one thousand trials based on the two controlled parameters as follows:

1. The number of corresponding line pairs n. The number n is assigned values of 3, 6, 10, 15 and 30.
2. The concentration parameter k_c. The level of k_c is changed based on 100,000, 50,000, 10,000, 5000, 1000, 500 and 300. The larger value of k_c corresponds to the smaller noise level. The correspondences between the value of k_c and the estimated mean and variance have been shown in Table 1.

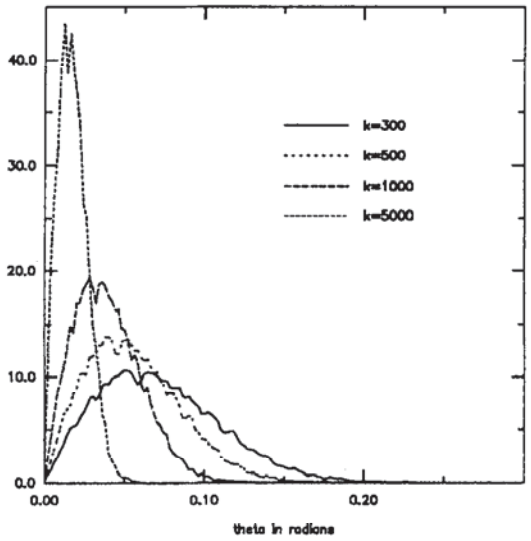

Fig. 5. The probability density of generated data of θ for the Fisher distribution for $k_c = 300$, 500, 1000 and 5000.

We define the error for the six transformation parameters of Φ as follows:

$$error_\omega = \min\{2\pi \times n + (\hat{\omega} - \omega)| : n \in Z\}$$

$$error_\phi = \min\{2\pi \times n + (\hat{\phi} - \phi)| : n \in Z\}$$

$$error_k = \min\{2\pi \times n + (\hat{\kappa} - \kappa)| : n \in Z\}$$

$$error_{t_x} = |\hat{t}_x - t_x|$$

$$error_{t_y} = |\hat{t}_y - t_y|$$

$$error_{t_z} = |\hat{t}_z - t_z|$$

and the average absolute error is defined as follows.

$$avg\_error\_rot = \tfrac{1}{3}(error_\omega + error_\phi + error_\kappa)$$

$$avg\_error\_t = \tfrac{1}{3}(error_{t_x} + error_{t_y} + error_{t_z})$$

Table 1
The correspondences between the value of k_c and the estimated mean and variance value of θ in degrees

k_c	300	500	1000	5000	10,000	50,000
Mean	4.16	3.21	2.28	1.016	0.720	0.322
Variance	4.74	2.85	1.42	0.285	0.141	0.028

The initial guess is generated within 20% of the true value. For example, the value of ω is generated from [15°,45°] and the initial value, for ω_i is generated from a uniform distribution satisfying $|\omega_i - \omega| < 0.2|\omega|$.

The estimation procedures are as follows.

1. Get the initial value for Φ.
2. Use Eq. (3.18).
3. Calculate the incremental $\Delta\Phi$ and then update the $\hat{\Phi}$.
4. Use the following performance criteria.
   Let $\hat{\Phi}$ be an estimate for the unknown parameters. We assume that the user desires that $|\Delta\Phi| < t_1$, be a fixed threshold. For instance, $t_1 = 10^{-7}$. Sometimes, due to the noise, the algorithm may not able to converge to the required accuracy t_1 after 25 iterations. Experiments show that if there is convergence, the magnitude of $\Delta\Phi$ is usually less than 10^{-7} after five iterations. When the magnitude of $\Delta\Phi$ fails to reach t_1, the error is usually very large in estimation. Hence, we define a rejection criterion by checking the magnitude of $\Delta\Phi$ at the maximum number of iterations allowed. If the magnitude of $\Delta\Phi$ is still larger than $t_d = 10^{-4} > t_1$ after 25 iterations, it will indicate that a large error is detected and the trial is rejected.

The same procedure is applied to Eq. (2.4) to obtain Ψ and T of the decoupled least squares approach.

Sometimes, the incremental value of $\Delta\Phi$ may be too big and overshoot, because of the translation vector component. To solve this problem we may multiply $\Delta\Phi$ by a constant which is determined by searching a value in [0,1] to minimize the objective function for each iteration. The search steps are ten, i.e. the step value is 0.1. By doing this it can prevent the divergence due to the overshoot of the translation vector component. In the decoupled least squares algorithm, the Euler angle in the rotation matrix is a periodical variable and is solved first. Therefore, the overshoot doesn't cause the problem.

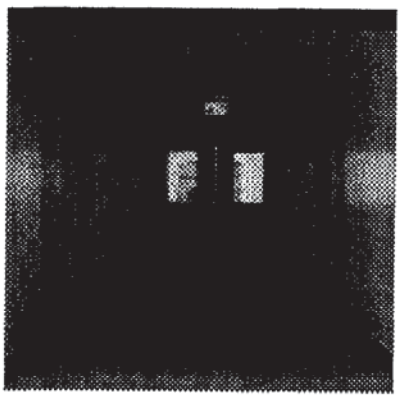

Fig. 6. The hallway image. The origin of the image plane is assumed to be at (242, 256).

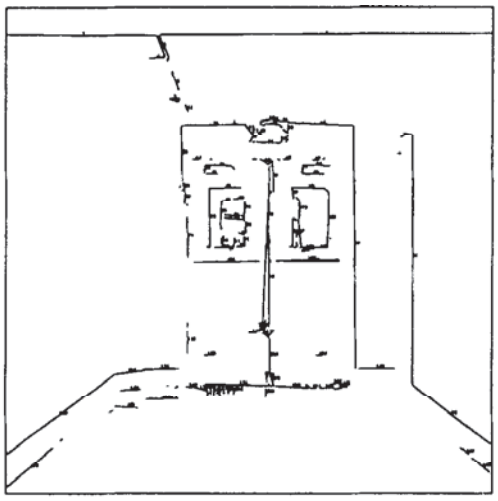

Fig. 7. The segmentation results of the hallway image.

4.3. Real image experiment

We apply the new algorithm to the hallway image obtained from the image library of the VISIONS group at the University of Massachusetts. The image and model parameters are described in [10]. We quote as follows. "The 3D model was built by measuring distance with a tape measure and is accurate to about 0.1 feet. The image was acquired using a SONY B/W camera mounted on a Denning Mobil Robotics vehicle tethered to GOULD frame grabber. 512 by 484 images are obtained, with a field of view of 24.0 degree by 23.0 degree." The focal length can be calculated from the field of view. One of the images used in this experiment is shown in Fig. 6. The segmentation image is shown in Fig. 7. We take nine line segments for the exterior orientation is shown in Fig. 8.

5. Results and discussions

In the first experiment we study the numerical stability of the new algorithm and the decoupled least squares

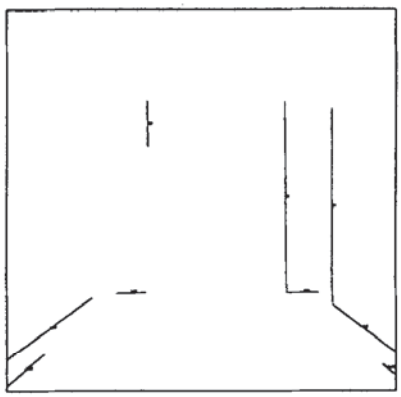

Fig. 8. The line segments used for the exterior orientation.

Table 2
The results of unperturbed observation with initial guess within 20% of the true value. The number of correspondences N changes based on 6, 10, 15 and 30

N	Method	Mean absolute error						No. of Rejects
		w	Φ	κ	t_x	t_y	t_z	
6	*LS*	0.0	0.0	0.0	6.8×10^{-7}	1.6×10^{-6}	6.1×10^{-7}	0
	New	0.0	0.0	0.0	0.0	0.0	0.0	0
10	*LS*	0.0	0.0	0.0	7.2×10^{-7}	1.3×10^{-6}	2.1×10^{-7}	0
	New	0.0	0.0	0.0	0.0	0.0	0.0	0
15	*LS*	0.0	0.0	0.0	1.2×10^{-6}	2.1×10^{-6}	2.8×10^{-7}	0
	New	0.0	0.0	0.0	0.0	0.0	0.0	0
30	*LS*	0.0	0.0	0.0	7.1×10^{-7}	1.2×10^{-6}	2.9×10^{-7}	0
	New	0.0	0.0	0.0	0.0	0.0	0.0	0

algorithm by applying both algorithms on the unperturbed observations. The initial guess is within 20% of the true value. The number of line correspondences is controlled to take one of the value based on 6, 10, 156 and 30. The results of the study are shown in Table 2. The computation resolution is 12 digits, that means the error for Φ in the new algorithm and Ψ (three Euler angles) in the decoupled least squares algorithm is less than 10^{-12}. However, the unnoticeable error in the Euler angles is propagated to the translation vector when the decoupled calculation is used. As a result, all three components in the translation vector have an error around 10^{-6}. This explains why we should estimate the six parameters simultaneously in order to obtain the optimal solution.

Because the initial guess usually affects the convergence of optimization, we study how the initial guess affects on both algorithms in the second experiment. In the experiment we use six line correspondences and $k_c = 1000$. The initial guess can be within 10%, 15%, 20% and 30% of the true values. The results are shown in Fig. 9. The results show that both the new algorithm and the decoupled least squares algorithm are not sensitive to the initial guess at that range. A slight change in rotation angles of the new algorithm may result from the local minima. Generally speaking, the decoupled least square approach is more stable to large initial guess errors due to the fact that its rotation parameters are periodical variables. The mean absolute error of the *avg_error_rot* for the new and the decoupled least squares approaches are 0.039 and 0.042 radians and that of *avg_error_rot* for the new and the decoupled least squares approaches are 2.161 and 2.516. In accuracy, the new algorithm is about 8% better for the rotation parameters and 16.4% better for the translation vector than the decoupled least squares approach.

In the third experiment we study how the number of line correspondences affects the accuracy of the estimation. The concentration parameter k_c is fixed at 1000 and the initial guess is within 20% of the true value. The number of line correspondences can be 6, 10, 20 and 30. As shown in Fig. 10, the increase of the number of line correspondences surely improves the estimation results for both techniques.

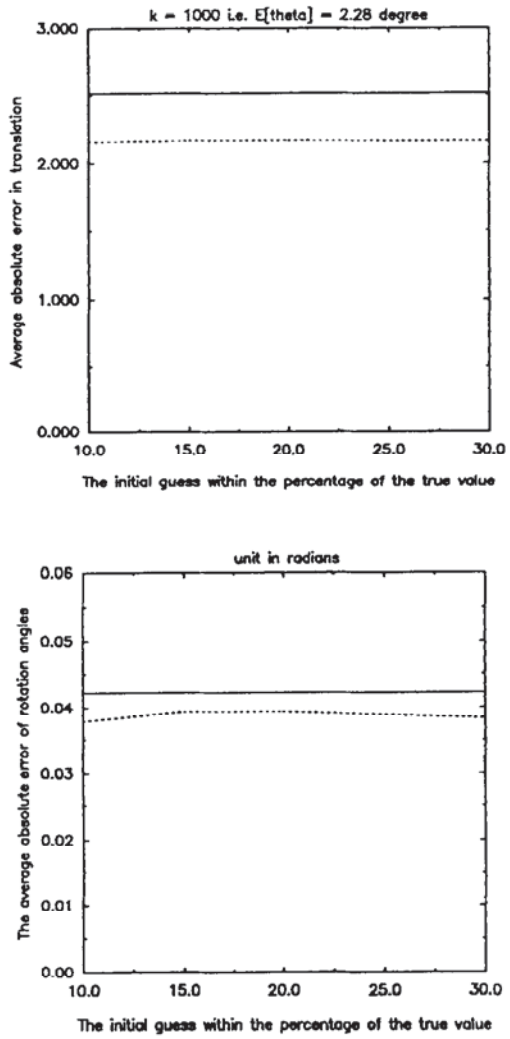

Fig. 9. The study of the initial guess affects on the estimation. The number of corresponding lines is six and $k_c = 1000$. – – – – new; —— least squares.

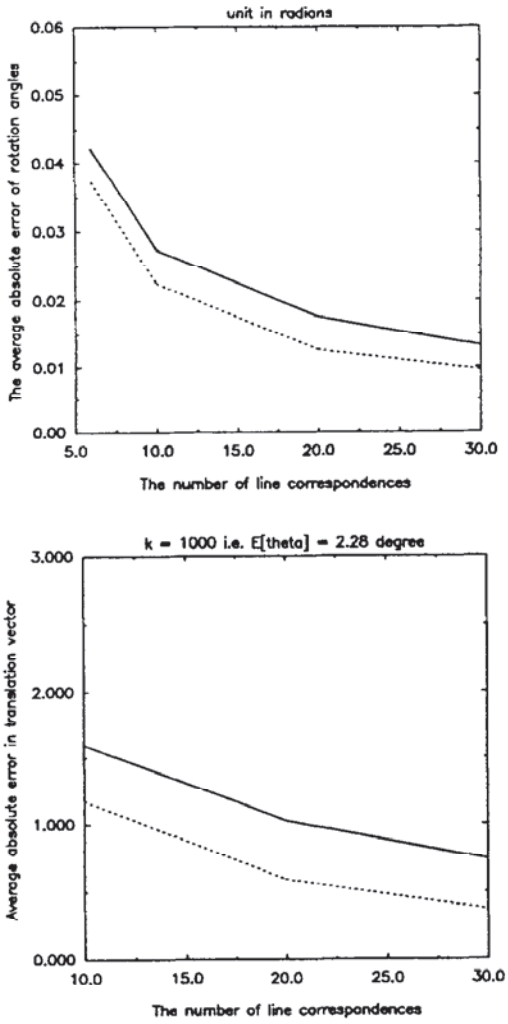

Fig. 10. The study of the number of corresponding lines affects on the estimation. The initial guess is within 20% of the true value and $k_c = 1000$. - - - - new; —— least squares.

The mean absolute error of translation vector is inversely proportional to the number of line correspondences. The mean absolute error of the three Euler angles is improved too.

Fig. 11. The backprojection of the 3D object model lines onto the 2D image.

Sometimes the mean absolute error of thousand trials may not give a fair look at the algorithms, because a few large error may bring up the mean. Hence, we show in brief from the histogram of the average absolute error of rotation components and the average absolute error of the translation for the six line correspondences and $k_c = 50000$ in Table 3.

In the hallway image we list the results in Table 4. The results show that the estimation is very close to the ground truth in x and z components and about 7% difference in the y component. The error in the y component may be caused by the bias along the horizontal axis of the image center. Since the true rotation is unknown, we backproject the 3D model lines onto the image in Fig. 6 by the estimated transformation. The result is shown in Fig. 11.

Though the new algorithm is better in accuracy, it pays a price in computing cost due to more

Table 3
The mean absolute error histogram of the three Euler angles and the translation vector

Mean absolute error		<0.001	[0.001, 0.01)	[0.01, 0.1)	[0.1, 1)	[1, 10)
Euler angler	LS	25	845	135	0	0
	B	32	871	97	0	0
Translation vector	LS	0	0	59	918	23
	B	0	2	100	881	17

Table 4
The estimation of six unknown parameters and the ground truth for the hallway image. NA means that data are not available. The translation x direction is parallel to hallway, the y direction is the horizontal direction, and the z-direction is the vertical direction

Parameter	ω	ϕ	κ	t_x (ft.)	t_y (ft.)	t_z (ft.)
True	NA	NA	NA	34.792	4.033	3.6
Estimation	0.00318	0.01207	−0.01047	34.533	3.740	3.545

complicated derivation of the unit normal vector $\vec{d}$. However, the number of iterations is about the same for both methods. More detailed results can be found in Ref. [25].

6. Summary

We conclude this paper by summarizing the results. We have developed a new algorithm for statistical estimation to estimate six unknown parameters simultaneously from which the optimal solution is achieved. The experiments verify that the new algorithm is better than the decoupled least squares approach in two aspects. First, the new algorithm gives exact solution in all six parameters when there is no noise. However, the decoupled least squares algorithm only gives the exact solution in three Euler angles only. This is because the numerical error in the rotation calculation propagates to the translation calculation. Second, the new algorithm gives about 10–20% better in accuracy than the least squares approach in all experiments.

Acknowledgements

The authors wish to thank the VISIONS group at the University of Massachusetts for providing the image data and information, and the reviewers for their helpful suggestions.

References

[1] H. Bopp and H. Krauss, Ein Orientierungs- und Kalibrierungsverfahren für nichttopographische Anwendungen der Photogrammetrie, ANV, 5 (1978) 182–188.

[2] M.A. Crombie and W.A. Baracat, Applying photogrammetry to real time collection of digital image data, U.S. Army Corps of Engineers ETL-0275, Fort Belvoir, VA, 1981, pp. 4–38.

[3] S.F. El-Hakim, A photogrammetric vision system for robots, Photogrammetric Engineering and Remote Sensing, 51(5) (1985) 545–552.

[4] W. Förstner, The reliability of block triangulation, Photogrammetric Engineering and Remote Sensing, 51(6) (August 1985) 1137–1149.

[5] R.M. Haralick, C.N. Lee, K. Ottenberg and M. Nölle, Analysis of the three point perspective pose estimation problem and solutions, IEEE Conf. Computer Vision and Pattern Recognition, Maui, Hawaii, June 1991.

[6] S. Linnainmaa, D. Harwood and L.S. Davis, Pose estimation of a three-dimensional object using triangle pairs, IEEE Trans. Patt. Analysis and Machine Intelligence, 10(5) (1988) 634–647.

[7] J.A. Pope, An advantageous, alternative parameterization of rotations for analytical photogrammetry, ESSA Tech. Rep., C and GS 39.

[8] E.H. Thompson, On exact linear solution of the problem of absolute orientation, Photogrammetria, 13(4) (1958) 163–178.

[9] D.G. Lowe, Perceptual Organization and Visual Recognition, Kluwer, Boston, 1985.

[10] R. Kumar and R. Hanson, Analysis of different robust methods for pose estimation, IEEE Workshop on Robust Computer Vision, Seattle, WA, October 1–3 1990.

[11] R. Kumar, Model dependent inference of 3D information from a sequence of 2D iamges, Coins TR 92-04, University of Massachusetts, 1992.

[12] Liu, Yuncai, T.S. Huang and O.D. Faugeras, Determination of camera location from 2-D to 3-D line and point correspondences, IEEE Trans. Patt. Analysis and Machine Intelligence, 12(1) (1990) 28–37.

[13] J. Weng, N. Ahuja and T.S. Huang, Optimal motion and structure estimation, Proc. IEEE Conf. Computer Vision and Pattern Recognition 1989, pp. 144–152.

[14] O.D. Faugeras, F. Lustman and G. Toscani, Motion and structure from point and line matches, Proc. Int. Conf. Computer Vision, London, England, June 1987.

[15] R. Horaud, New methods for matching 3-D objects with single perspective views, IEEE Trans. Patt. Analysis and Machine Intelligence, 9(3) (May 1987).

[16] T. Shakunaga, Pose estimation of jointed structures, Proc. IEEE Conf. Computer Vision and Pattern Recognition, 144–152, 1991.

[17] S.T. Barnard, Interpreting perspective images, Artif. Intelligence, 21 (1983) 435–462.

[18] R.A. Fisher, Dispersion on a sphere, Proc. Royal Soc. London A, 217 (1953).

[19] K.V. Mardia, Statistics of Directional Data, Academic Press, New York, 1972.

[20] J. Besag, On the statistical analysis of the dirty pictures (with discussion), J. Roy. Statist. Soc. B, 48 (1986) 259–302.

[21] S. German, Experiments in bayesian image analysis, in Bayesian Statistics, 159–172 (J.M. Bernardo, M.H. DeGroot, D.V. Lindley and A.F.M. Smith, eds), Oxford University Press, 1988.

[22] R.M. Haralick, Computer vision theory: the lack thereof, Computer Vision, Graphics, and Image Processing, 36 (1986) 372–386.

[23] W.M. Wells III, MAP model matching, Proc. IEEE Conf. Computer Vision and Pattern Recognition, 1991, 486–492.

[24] A. Papoulis, Probability, Random Variables, and Stochastic Process, 1984.

[25] C.N. Lee, Perspective projection expert, PhD Dissertation, University of Washington, 1992.