Available online at www.sciencedirect.com
ScienceDirect
Pattern Recognition Letters 29 (2008) 223–231
Pattern Recognition Letters
www.elsevier.com/locate/patrec

Pose determination and plane measurement using a trapezium

Fuqing Duan [a,b,*], Fuchao Wu [b], Zhanyi Hu [b]

[a] College of Information Science and Technology, Beijing Normal University, Beijing 100875, PR China
[b] National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, P.O. Box 2728, Beijing 100080, PR China

Received 20 September 2006; received in revised form 11 September 2007
Available online 17 October 2007

Communicated by J.A. Robinson

Abstract

In this paper, a new affine invariant of trapezia is introduced, and the projection of trapezia is deduced from this invariant. Known the lengths of the two parallel sides of a trapezium, pose estimation and plane measurement can be realized in a very simple way from the projection of the trapezium. Experiments on simulated and real images show that the approach is robust and accurate. Two parallel lines, which can determine a trapezium, are not rare in many structured scenes, the proposed method has wide applicability.

Keywords: Pose estimation; PnP; Affine invariant; Trapezium; 3D reconstruction

1. Introduction

Pose determination (Haralick et al., 1991; Quan and Lan, 1999; Haralick, 1989) is to estimate the position and orientation of one calibrated camera from a set of correspondences between 3D control points and 2D image points. It has important applications such as cartography, robotics, tracking and object recognition. Pose determination is also referred to as the perspective-n-point problem (PnP) (Horaud et al., 1989; Quan and Lan, 1999; Fishler and Bolles, 1981). Fishler and Bolles (1981) define the problem as: Given the relative spatial locations of n control points, and given the angle to each pair of control points from the perspective center, find the lengths of the line segments joining the perspective center to each control point.

Generally speaking, the PnP problem has a unique linear solution with $n \geqslant 6$ and can be solved by the direct linear transformation (DLT), while it has multiple solutions with $n < 6$ (Fishler and Bolles, 1981; Gao et al., 2003; Hu and Wu, 2002; Wu and Hu, 2001). Fishler and Bolles (1981) solve the P4P problem by finding solutions associated with subsets of three points and selecting the common solutions they have. Horaud et al. (1989) casts the P4P problem into a problem of solving a biquadratic polynomial equation with one unknown, and gives an analytic solution. Penna (1991) points out that, given four coplanar points with known coordinates in the object centered frame and their image points, camera's pose can be determined linearly and uniquely. Quan and Lan (1999) present a linear solution to pose estimation with $n \geqslant 4$. They decompose the problem into several P3P problems with a common solution. Using the classical Sylvester resultant, they convert the problem into a 4-degree polynomial equation system with one unknown, which can be viewed as a homogeneous linear equation system in each-order term. Recently, Ansar and Daniilidis (2003) present a similar linear approach. By considering the multiplication of each pair of unknowns as a new variable, they convert the constraint equation system into a high dimensional linear equation system. Both of the two linear approaches can give the unique solution for the reference points not lying on the critical configurations. Moreover, there are many

* Corresponding author. Address: College of Information Science and Technology, Beijing Normal University, Beijing 100875, PR China. Tel.: +86 10 58800442; fax: +86 10 58809444.
E-mail addresses: fqduan@gmail.com (F. Duan), fcwu@nlpr.ia.ac.cn (F. Wu), huzy@nlpr.ia.ac.cn (Z. Hu).

0167-8655/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2007.09.018

approaches based on iterative methods (Yuan (1989); Lowe, 1991; Liu and Wong, 1999). The iterative methods may suffer from the problems of initialization and convergence.

Vision based measurement is another important problem in computer vision. It has wide applications such as architectural measurement, reconstruction from paintings, forensic measurement and traffic accident investigation (Criminisi, 2001). Generally speaking, the methods of computer vision based measurements in the literature may be broadly divided into two categories. One is to reconstruct the metric structure of the scene from two or more images by stereo vision techniques (Liebowitz et al., 1999; Hartely and Zisserman, 2000), which is a hard task due to the problem of point correspondence in different views, the other is to directly use a single uncalibrated image under the knowledge of some geometrical scene constraints such as coplanarity, parallelism, perpendicularity, etc. (Criminisi, 2001; Criminis et al., 1999a; Criminis et al., 1999b). In (Criminisi, 2001; Criminis et al., 1999a), a key point based method to measure the Euclidean distance of two points on a world plane is proposed. In this method, at least four correspondences between 3D control points and 2D image points should be known to estimate the plane homography, and the points on the plane are reconstructed by the homography.

Invariants (Mundy and Zisserman, 1992; Mundy et al., 1994) are properties of geometric configurations, which remain unchanged under an appropriate class of transformations. Many visual tasks can be solved by invariants such as recognition in model based vision, shape descriptors for 3D objects and transference, characterizing unknown geometric structure. In this work, we introduce a new affine invariant of trapezia, and the projection of trapezia is deduced from this invariant. Using the projection of trapezia, we present a very simple solution of pose determination and plane measurement, in which we only need to know the lengths of the two parallel sides of a trapezium. In some applications such as face recognition, visual surveillance, we only need to estimate the camera's orientation. If the trapezium is an isosceles trapezium, we can estimate the rotation between the object frame and the camera frame without any metric information.

In the paper, a 3D point is denoted by $\boldsymbol{X} = [x, y, z]^{\mathrm{T}}$, and a 2D point is denoted by $\boldsymbol{m} = [u, v, 1]^{\mathrm{T}}$ (homogeneous coordinate). The camera is of the pinhole model, then under the camera coordinate system, a 3D point $\boldsymbol{X}$ is projected to its image point $\boldsymbol{m}$ by $\lambda \boldsymbol{m} = K\boldsymbol{X}$, where λ is a scale factor (projection depth of 3D point $\boldsymbol{X}$), K is the camera intrinsic matrix.

$$\boldsymbol{K} = \begin{bmatrix} f & s & u_0 \\ 0 & \alpha f & v_0 \\ 0 & 0 & 1 \end{bmatrix},$$

where f is the focal length, α is the aspect ratio of the image u and v axes; $\boldsymbol{p} = [u_0, v_0, 1]^{\mathrm{T}}$ is the principal point, and s the parameter describing the skew of the two image axes. We always assume that the camera intrinsic matrix is known and use the normalized coordinate on the image plane: $\tilde{\boldsymbol{m}} = K^{-1}\boldsymbol{m}$. In this case, the coordinate of a 3D point $\boldsymbol{X}$ under the camera coordinate system can be expressed as $\boldsymbol{X} = \lambda\tilde{\boldsymbol{m}}$.

The paper is organized as follows: In Section 2, a new affine invariant of trapezia is introduced, and the projection of trapezia is deduced from this invariant; Analytical solution of pose estimation is shown in Section 3; Section 4 discusses the plane measurement; Section 5 analyses degeneration of the system; experiments are reported in Section 6; conclusions are given at the end of this paper.

2. Invariants and projection of trapezia

2.1. Invariants of trapezia

Let $\{\boldsymbol{X}_i: i = 1, 2, 3, 4\}$ be the four vertices of a trapezium, we always assume $\overrightarrow{\boldsymbol{X}_1\boldsymbol{X}_2} \,//\, \overrightarrow{\boldsymbol{X}_3\boldsymbol{X}_4}$, $d_{12} = \|\boldsymbol{X}_2 - \boldsymbol{X}_1\|$, and $d_{34} = \|\boldsymbol{X}_4 - \boldsymbol{X}_3\|$ in this paper. Then, the parameter,

$$r = \frac{||\boldsymbol{X}_4 - \boldsymbol{X}_3||}{||\boldsymbol{X}_2 - \boldsymbol{X}_1||} \tag{1}$$

is an affine invariant of trapezia since an affine transformation preserves the length ratio of two parallel line segments. Next, we introduce a new affine invariant of trapezia, which is crucial in our work.

From $\overrightarrow{\boldsymbol{X}_1\boldsymbol{X}_2} \,//\, \overrightarrow{\boldsymbol{X}_3\boldsymbol{X}_4}$ and Eq. (1), we have $\boldsymbol{X}_4 - \boldsymbol{X}_3 = r(\boldsymbol{X}_2 - \boldsymbol{X}_1)$. Thus,

$$\boldsymbol{X}_4 = r\boldsymbol{X}_2 - r\boldsymbol{X}_1 + \boldsymbol{X}_3 = [\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3][-r, r, 1]^{\mathrm{T}}.$$

Let $X = [\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3]$, then

$$X^{-1}\boldsymbol{X}_4 = [-r, r, 1]^{\mathrm{T}}. \tag{2}$$

Since r is an affine invariant, Eq. (2) is also an affine invariant. We can give a direct proof as:

Let $\boldsymbol{X}' = A\boldsymbol{X} + \boldsymbol{t}$ be a 3D affine transformation, then

$$\begin{aligned} \boldsymbol{X}'_4 &= A\boldsymbol{X}_4 + \boldsymbol{t} = (-rA\boldsymbol{X}_1 + rA\boldsymbol{X}_2 + A\boldsymbol{X}_3) + \boldsymbol{t} \\ &= (-r\boldsymbol{X}'_1 + r\boldsymbol{t} + r\boldsymbol{X}'_2 - r\boldsymbol{t} + \boldsymbol{X}'_3 - \boldsymbol{t}) + \boldsymbol{t} \\ &= -r\boldsymbol{X}'_1 + r\boldsymbol{X}'_2 + \boldsymbol{X}'_3. \end{aligned}$$

Thus $X'^{-1}\boldsymbol{X}'_4 = [-r, r, 1]^{\mathrm{T}}$. Hence, Eq. (2) is an affine invariant of trapezia.

2.2. Projection of trapezia

In the paper, we always assume that the camera center is not on the plane defined by the trapezium. Using the affine invariant (2), we can obtain the projection of trapezia.

Proposition 1. *Given the image* $\{\tilde{\boldsymbol{m}}_j : j = 1, 2, 3, 4\}$ *of a trapezium* $\{\boldsymbol{X}_j: j = 1, 2, 3, 4\}$, *under the camera coordinate system, we have*:

$$\boldsymbol{X}_j = \lambda_4 r^{-1} q_j \tilde{\boldsymbol{m}}_j, \quad j = 1, 2; \quad \boldsymbol{X}_j = \lambda_4 q_j \tilde{\boldsymbol{m}}_j, \quad j = 3, 4, \tag{3}$$

where λ_4 *is the projection depth of the point* $\boldsymbol{X}_4$, $[q_1, q_2, q_3]^{\mathrm{T}} = [-\tilde{\boldsymbol{m}}_1, \tilde{\boldsymbol{m}}_2, \tilde{\boldsymbol{m}}_3]^{-1}\tilde{\boldsymbol{m}}_4$, $q_4 = 1$.[1]

[1] In the paper, we always assume $q_4 = 1$.

Proof. Under the camera coordinate system, we have

$$\boldsymbol{X}_j = \lambda_j \tilde{\boldsymbol{m}}_j, \quad j = 1,2,3,4. \tag{4}$$

Then, $X \triangleq [\boldsymbol{X}_1, \boldsymbol{X}_2, \boldsymbol{X}_3] = M\mathrm{diag}[\lambda_1, \lambda_2, \lambda_3]$, where $M = [\tilde{\boldsymbol{m}}_1, \tilde{\boldsymbol{m}}_2, \tilde{\boldsymbol{m}}_3]$.

Since the camera center is not on the trapezium plane, the matrix $\boldsymbol{X}$ is reversible. Thus, the matrix M is also reversible. Hence, we have: $X^{-1} = \mathrm{diag}(1/\lambda_1, 1/\lambda_2, 1/\lambda_3) M^{-1}$. Then

$$\begin{aligned} X^{-1}\boldsymbol{X}_4 &= \mathrm{diag}(1/\lambda_1, 1/\lambda_2, 1/\lambda_3) M^{-1}(\lambda_4 \tilde{\boldsymbol{m}}_4) \\ &= \mathrm{diag}(\lambda_4/\lambda_1, \lambda_4/\lambda_2, \lambda_4/\lambda_3) M^{-1}\tilde{\boldsymbol{m}}_4. \end{aligned}$$

From

$$\begin{aligned} M^{-1}\tilde{\boldsymbol{m}}_4 &= \mathrm{diag}(-1,1,1) \cdot [M\mathrm{diag}(-1,1,1)]^{-1}\tilde{\boldsymbol{m}}_4 \\ &= \mathrm{diag}(-1,1,1) \cdot [-\tilde{\boldsymbol{m}}_1, \tilde{\boldsymbol{m}}_2, \tilde{\boldsymbol{m}}_3]^{-1}\tilde{\boldsymbol{m}}_4 \\ &= [-q_1, q_2, q_3]^{\mathrm{T}}, \end{aligned}$$

we obtain $\boldsymbol{X}^{-1}\boldsymbol{X}_4 = [-\lambda_4 q_1/\lambda_1, \lambda_4 q_2/\lambda_2, \lambda_4 q_3/\lambda_3]^{\mathrm{T}}$.

By the affine invariant (2), we have $[-\lambda_4 q_1/\lambda_1, \lambda_4 q_2/\lambda_2, \lambda_4 q_3/\lambda_3]^{\mathrm{T}} = [-r, r, 1]^{\mathrm{T}}$. Thus,

$$\lambda_j = \lambda_4 r^{-1} q_j, \ \ j = 1,2; \quad \lambda_j = \lambda_4 q_j, \ \ j = 3,4.$$

Substituting them to Eq. (4), Eq. (3) holds. □

3. Pose estimation

3.1. Analytical solution for the P4P problem

Proposition 2. *Assume that the image of a trapezium $\{\boldsymbol{X}_j\}$ is $\{\tilde{\boldsymbol{m}}_j\}$, and the lengths of the two parallel sides, d_{12} and d_{34}, are known, then the distance from each vertex of the trapezium to the camera center is*

$$||\boldsymbol{X}_j|| = \frac{d_{12}||q_j \tilde{\boldsymbol{m}}_j||}{||q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3||}, \quad j = 1,2, \tag{5}$$

$$||\boldsymbol{X}_j|| = \frac{d_{34}||q_j \tilde{\boldsymbol{m}}_j||}{||q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3||}, \quad j = 3,4. \tag{6}$$

Proof. By Eq. (3), under the camera coordinate system,

$$\boldsymbol{X}_j = \lambda_4 r^{-1} q_j \tilde{\boldsymbol{m}}_j, \quad j = 1,2; \qquad \boldsymbol{X}_j = \lambda_4 q_j \tilde{\boldsymbol{m}}_j, \quad j = 3,4.$$

From $d_{34} = \|\boldsymbol{X}_4 - \boldsymbol{X}_3\|$, we have $d_{34} = \lambda_4 \|q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3\|$.

Then $\lambda_4 = \frac{d_{34}}{\|q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3\|}$. Since $r = \frac{\|\boldsymbol{X}_4 - \boldsymbol{X}_3\|}{\|\boldsymbol{X}_2 - \boldsymbol{X}_1\|} = \frac{d_{34}}{d_{12}}$, we can obtain

$$\boldsymbol{X}_j = \frac{d_{12} q_j}{||q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3||}\tilde{\boldsymbol{m}}_j, \quad j = 1,2, \tag{7}$$

$$\boldsymbol{X}_j = \frac{d_{34} q_j}{||q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3||}\tilde{\boldsymbol{m}}_j, \quad j = 3,4. \tag{8}$$

And thus, we obtain Eqs. (5) and (6). □

Remark 1. In the classical P4P problem, the distance between each pair of control points must be given. Proposition 2 shows that if the four control points are vertices of a trapezium, we only need to know the lengths of the two parallel sides. Especially, if the four points are vertices of a parallelogram, we only need to know the length of one side. In addition, we can also obtain the lengths of other two sides of the trapezium:

$$||\boldsymbol{X}_3 - \boldsymbol{X}_1|| = \sqrt{\frac{(d_{34}q_3\tilde{\boldsymbol{m}}_3 - d_{12}q_1\tilde{\boldsymbol{m}}_1)^{\mathrm{T}}(d_{34}q_3\tilde{\boldsymbol{m}}_3 - d_{12}q_1\tilde{\boldsymbol{m}}_1)}{(q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3)^{\mathrm{T}}(q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3)}} \tag{9}$$

$$||\boldsymbol{X}_4 - \boldsymbol{X}_2|| = \sqrt{\frac{(d_{34}\tilde{\boldsymbol{m}}_4 - d_{12}q_2\tilde{\boldsymbol{m}}_2)^{\mathrm{T}}(d_{34}\tilde{\boldsymbol{m}}_4 - d_{12}q_2\tilde{\boldsymbol{m}}_2)}{(q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3)^{\mathrm{T}}(q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3)}} \tag{10}$$

Corollary 1. *Given the image of the two non-parallel line segments and their middle points, and if known the lengths of the two line segments, then we can compute the distances from the end points of the line segments to the camera center.*

Proof. As shown in Fig. 1, let the images of the two line segments $\{\boldsymbol{X}_1\boldsymbol{X}_2, \boldsymbol{Y}_1\boldsymbol{Y}_2\}$ and their middle points $\{\boldsymbol{X}_{12}, \boldsymbol{Y}_{12}\}$ be $\{\tilde{\boldsymbol{m}}_1\tilde{\boldsymbol{m}}_2, \tilde{\boldsymbol{n}}_1\tilde{\boldsymbol{n}}_2\}$ and $\{\tilde{\boldsymbol{m}}_{12}, \tilde{\boldsymbol{n}}_{12}\}$ respectively. Then we can determine the vanishing points, $\tilde{\boldsymbol{v}}_1$ and $\tilde{\boldsymbol{v}}_2$, of the two line segments $\boldsymbol{X}_1\boldsymbol{X}_2$ and $\boldsymbol{Y}_1\boldsymbol{Y}_2$ on the image plane by cross ratio:

$$\begin{cases} (\tilde{\boldsymbol{m}}_1\tilde{\boldsymbol{m}}_2; \tilde{\boldsymbol{m}}_{12}\tilde{\boldsymbol{v}}_1) = -1, \\ (\tilde{\boldsymbol{n}}_1\tilde{\boldsymbol{n}}_2; \tilde{\boldsymbol{n}}_{12}\tilde{\boldsymbol{v}}_2) = -1. \end{cases}$$

Let $\boldsymbol{p}(\boldsymbol{q})$ be the line that passes through $\tilde{\boldsymbol{v}}_2$ and $\tilde{\boldsymbol{m}}_1(\tilde{\boldsymbol{m}}_2)$. The intersection of the line $\boldsymbol{p}(\boldsymbol{q})$ with one line that passes through $\tilde{\boldsymbol{v}}_1$ is denoted by $\tilde{\boldsymbol{m}}_3(\tilde{\boldsymbol{m}}_4)$. Then, $\{\tilde{\boldsymbol{m}}_1, \tilde{\boldsymbol{m}}_2, \tilde{\boldsymbol{m}}_3, \tilde{\boldsymbol{m}}_4\}$ must be the image of a parallelogram, and thus we can compute the distances from the end points of the line segment $\boldsymbol{X}_1\boldsymbol{X}_2$ to the camera center.

Similarly, we can compute the distances from the end points of $\boldsymbol{Y}_1\boldsymbol{Y}_2$ to the camera center. □

Remark 2. In Corollary 1, if the two non-parallel line segments are coplanar, we only need to know any one length of the two lengths for computing the distance from the end points of the line segments to the camera center by Corollary 2.

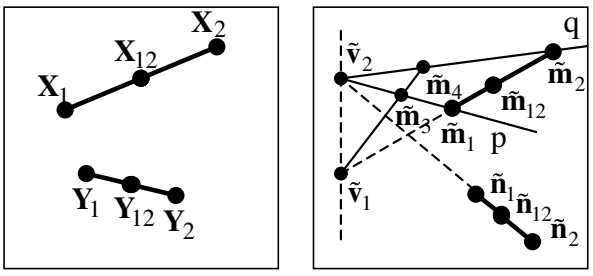

Fig. 1. The images of two non-parallel line segments and their middle points.

3.2. Analytical solution for the camera's position and orientation

Lemma 1. *Assume that the image of a trapezium $\{\boldsymbol{X}_j\}$ is $\{\tilde{\boldsymbol{m}}_j\}$, and the lengths of the two parallel sides, d_{12} and d_{34}, are known. Let $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Pi} = 1$ be the equation of the trapezium plane under the camera coordinate system. Then, we have*

$$\boldsymbol{\Pi} = M_\pi^{-1}\boldsymbol{b}_\pi, \tag{11}$$

where

$$M_\pi = \frac{d_{12}^2\sum_{j=1}^{2} q_j^2\tilde{\boldsymbol{m}}_j\tilde{\boldsymbol{m}}_j^{\mathrm{T}} + d_{34}^2\sum_{j=3}^{4} q_j^2\tilde{\boldsymbol{m}}_j\tilde{\boldsymbol{m}}_j^{\mathrm{T}}}{||q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3||},$$

$$\boldsymbol{b}_\pi = d_{12}\sum_{j=1}^{2} q_j\tilde{\boldsymbol{m}}_j + d_{34}\sum_{j=3}^{4} q_j\tilde{\boldsymbol{m}}_j.$$

Proof. From $\boldsymbol{X}_j^{\mathrm{T}}\boldsymbol{\Pi} = 1$, we have $\boldsymbol{X}_j\boldsymbol{X}_j^{\mathrm{T}}\boldsymbol{\Pi} = \boldsymbol{X}_j$, and then, $\left(\sum_{j=1}^{4}\boldsymbol{X}_j\boldsymbol{X}_j^{\mathrm{T}}\right)\boldsymbol{\Pi} = \sum_{j=1}^{4}\boldsymbol{X}_j$. So, we have:

$$\boldsymbol{\Pi} = \left(\sum_{j=1}^{4}\boldsymbol{X}_j\boldsymbol{X}_j^{\mathrm{T}}\right)^{-1}\sum_{j=1}^{4}\boldsymbol{X}_j. \tag{12}$$

Substituting Eqs. (7) and (8) into (12), we can obtain Eq. (11). □

We establish an object coordinate system with origin at the point $\boldsymbol{X}_1$ and axes as shown in Fig. 2. Where, the direction $\boldsymbol{e}_3$ of z-axis is the direction $\boldsymbol{n}_\pi$ of the trapezium plane, the direction of x-axis is the unit vector $\boldsymbol{e}_1 = \overrightarrow{\boldsymbol{X}_1\boldsymbol{X}_2}/\|\overrightarrow{\boldsymbol{X}_1\boldsymbol{X}_2}\|$, and thus the direction of y-axis is $\boldsymbol{e}_2 = \boldsymbol{n}_\pi \times \boldsymbol{e}_1$.

By Eq. (7) and Lemma 1, under the camera coordinate system, the coordinates of origin and three axis directions of the above object coordinate system are as follows:

$$\boldsymbol{X}_1 = \frac{d_{12}q_1\tilde{\boldsymbol{m}}_1}{||q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3||},$$

$$\boldsymbol{e}_1 = \frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{||q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1||},$$

$$\boldsymbol{e}_3 = \boldsymbol{n}_\pi = \boldsymbol{\Pi}/||\boldsymbol{\Pi}||,$$

$$\boldsymbol{e}_2 = \boldsymbol{e}_3 \times \boldsymbol{e}_1 = \frac{\boldsymbol{\Pi}}{||\boldsymbol{\Pi}||} \times \left(\frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{||q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1||}\right).$$

Hence, we can give the following Proposition 3.

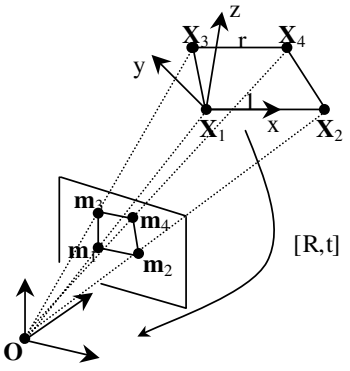

Fig. 2. The object coordinate system.

Proposition 3. *Assume that the image of a trapezium $\{\boldsymbol{X}_j\}$ is $\{\tilde{\boldsymbol{m}}_j\}$, and the lengths of the two parallel sides, d_{12} and d_{34}, are known. Then the camera's position and orientation can be expressed respectively as*

$$\boldsymbol{t} = \frac{d_{12}q_1\tilde{\boldsymbol{m}}_1}{\|q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3\|}, \tag{13}$$

$$R = \left[\frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{\|q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1\|}, \frac{\boldsymbol{\Pi}}{\|\boldsymbol{\Pi}\|} \times \left(\frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{\|q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1\|}\right), \frac{\boldsymbol{\Pi}}{\|\boldsymbol{\Pi}\|}\right]. \tag{14}$$

If the trapezium is an isosceles trapezium, we have the following proposition.

Proposition 4. *Assume that the image of an isosceles trapezium $\{\boldsymbol{X}_j\}$ is $\{\tilde{\boldsymbol{m}}_j\}$. Then we can determine the length ratio r of the two parallel segments and the camera's orientation as follows*:

$$r = \frac{(q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1)^{\mathrm{T}}(q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1)}{(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3)^{\mathrm{T}}(q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1)},$$

$$\boldsymbol{R} = \left[\frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{\|q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1\|}, \boldsymbol{e}_2, \left(\frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{\|q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1\|}\right) \times \boldsymbol{e}_2\right].$$

where

$$\boldsymbol{e}_2 = \frac{r(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3) - (q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1)}{\|r(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3) - (q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1)\|}.$$

Proof. Assume the middle points of the line segments $\boldsymbol{X}_1\boldsymbol{X}_2$ and $\boldsymbol{X}_3\boldsymbol{X}_4$ be $\boldsymbol{X}_{12}$ and $\boldsymbol{X}_{34}$ respectively. Then under the object coordinate system in Fig. 2, $\boldsymbol{e}_2 = \overrightarrow{\boldsymbol{X}_{12}\boldsymbol{X}_{34}}/\|\overrightarrow{\boldsymbol{X}_{12}\boldsymbol{X}_{34}}\|$ and $\boldsymbol{e}_3 = \boldsymbol{e}_1 \times \boldsymbol{e}_2$ since $\overrightarrow{\boldsymbol{X}_{12}\boldsymbol{X}_{34}} \perp \overrightarrow{\boldsymbol{X}_1\boldsymbol{X}_2}$.

From Eqs. (7) and (8), we have

$$\boldsymbol{X}_{12} = \frac{\boldsymbol{X}_1 + \boldsymbol{X}_2}{2} = \frac{d_{12}(q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1)}{2\|(q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3)\|},$$

$$\boldsymbol{X}_{34} = \frac{\boldsymbol{X}_3 + \boldsymbol{X}_4}{2} = \frac{rd_{12}(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3)}{2\|(q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3)\|}.$$

From $\overline{\boldsymbol{X}_{12}\boldsymbol{X}_{34}} \perp \overline{\boldsymbol{X}_1\boldsymbol{X}_2}$, we have $(r(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3) - (q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1))^{\mathrm{T}}(q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1) = 0$. Then,

$$r = \frac{\|\boldsymbol{X}_4 - \boldsymbol{X}_3\|}{\|\boldsymbol{X}_2 - \boldsymbol{X}_1\|} = \frac{(q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1)^{\mathrm{T}}(q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1)}{(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3)^{\mathrm{T}}(q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1)}.$$

Hence,

$$\boldsymbol{e}_2 = \frac{r(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3) - (q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1)}{\|r(q_4\tilde{\boldsymbol{m}}_4 + q_3\tilde{\boldsymbol{m}}_3) - (q_2\tilde{\boldsymbol{m}}_2 + q_1\tilde{\boldsymbol{m}}_1)\|},$$

$$\boldsymbol{R} = \left[\frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{\|q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1\|}, \boldsymbol{e}_2, \left(\frac{q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1}{\|q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1\|}\right) \times \boldsymbol{e}_2\right]. \qquad \square$$

4. Plane measurement

Proposition 5. *Assume that the image of a trapezium $\{\boldsymbol{X}_j\}$ is $\{\tilde{\boldsymbol{m}}_j\}$, and the lengths of the two parallel sides, d_{12} and d_{34}, are known. Then we have*

(1) *If the image of a point* $\boldsymbol{X}$ *on the trapezium plane is* $\tilde{\boldsymbol{m}}$, *under the camera coordinate system*,

$$\boldsymbol{X} = \frac{1}{\tilde{\boldsymbol{m}}^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}}. \tag{15}$$

(2) *If the images of two points* $\boldsymbol{X}, \boldsymbol{X}'$ *on the trapezium plane are* $\tilde{\boldsymbol{m}}, \tilde{\boldsymbol{m}}'$, *the distance between* $\boldsymbol{X}$ *and* $\boldsymbol{X}'$ *is*

$$\|\boldsymbol{X} - \boldsymbol{X}'\| = \left\| \frac{1}{\tilde{\boldsymbol{m}}^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}} - \frac{1}{\tilde{\boldsymbol{m}}'^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}}' \right\|. \tag{16}$$

(3) *If the images of two lines* $\boldsymbol{L}, \boldsymbol{L}'$ *on the trapezium plane are* $\tilde{\boldsymbol{l}}, \tilde{\boldsymbol{l}}'$, *the angle between* $\boldsymbol{L}$ *and* $\boldsymbol{L}'$ *is*

$$\cos\phi = \frac{(\tilde{\boldsymbol{l}}' \times \boldsymbol{\Pi})^{\mathrm{T}}(\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi})}{\|\tilde{\boldsymbol{l}}' \times \boldsymbol{\Pi}\| \cdot \|\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi}\|}. \tag{17}$$

(4) *If the images of two parallel lines* $\boldsymbol{L}, \boldsymbol{L}'$ *on the trapezium plane are* $\tilde{\boldsymbol{l}}, \tilde{\boldsymbol{l}}'$, *the distance between them is*

$$d = \left\| \frac{1}{\tilde{\boldsymbol{m}}^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}} - \frac{1}{\tilde{\boldsymbol{m}}'^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}}' \right\| \cdot \sqrt{1 - \cos^2\theta}, \tag{18}$$

where $\tilde{\boldsymbol{m}} \in \tilde{\boldsymbol{l}}$, $\tilde{\boldsymbol{m}}' \in \tilde{\boldsymbol{l}}'$, $\cos\theta = \frac{(\tilde{\boldsymbol{m}} \times \tilde{\boldsymbol{m}}' \times \boldsymbol{\Pi})^{T}(\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi})}{\|\tilde{\boldsymbol{m}} \times \tilde{\boldsymbol{m}}' \times \boldsymbol{\Pi}\| \cdot \|\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi}\|}$.

Proof

(1) $\boldsymbol{X}$ is the intersecting point of the line $\boldsymbol{X}(s) = s\tilde{\boldsymbol{m}}$ with the plane $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{\Pi} = 1$, then $s = \frac{1}{\tilde{\boldsymbol{m}}^{\mathrm{T}}\boldsymbol{\Pi}}$. Thus, $\boldsymbol{X} = \frac{1}{\tilde{\boldsymbol{m}}^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}}$.

(2) By Eq. (15), Eq. (16) holds.

(3) Since we use the normalized coordinate on the image plane, the vector $\boldsymbol{\Pi}$ is the coordinate of the vanishing line of the trapezium plane, i.e., the image of the infinite line on the trapezium plane. And thus, the vanishing points of lines $\boldsymbol{L}$, $\boldsymbol{L}'$ are $\boldsymbol{l} \times \boldsymbol{\Pi}$, $\boldsymbol{l}' \times \boldsymbol{\Pi}$ respectively. Hence, the angle between the two lines $\boldsymbol{L}$, $\boldsymbol{L}'$ is: $\cos\phi = \frac{(\tilde{\boldsymbol{l}}' \times \boldsymbol{\Pi})^{\mathrm{T}}(\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi})}{\|\tilde{\boldsymbol{l}}' \times \boldsymbol{\Pi}\| \cdot \|\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi}\|}$.

(4) Let $\tilde{\boldsymbol{m}} \in \tilde{\boldsymbol{l}}, \tilde{\boldsymbol{m}}' \in \tilde{\boldsymbol{l}}'$. Then, they are the images of two points $\boldsymbol{X} \in \boldsymbol{L}$ and $\boldsymbol{X}' \in \boldsymbol{L}'$ respectively. Let θ be the angle between two lines $\boldsymbol{X}\boldsymbol{X}'$ and $\boldsymbol{L}$. By Eq. (17),

$$\cos\theta = \frac{(\tilde{\boldsymbol{m}} \times \tilde{\boldsymbol{m}}' \times \boldsymbol{\Pi})^{\mathrm{T}}(\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi})}{\|\tilde{\boldsymbol{m}} \times \tilde{\boldsymbol{m}}' \times \boldsymbol{\Pi}\| \cdot \|\tilde{\boldsymbol{l}} \times \boldsymbol{\Pi}\|}.$$

Hence, by Eq. (16), we obtain the distance between $\boldsymbol{L}$ and $\boldsymbol{L}'$:

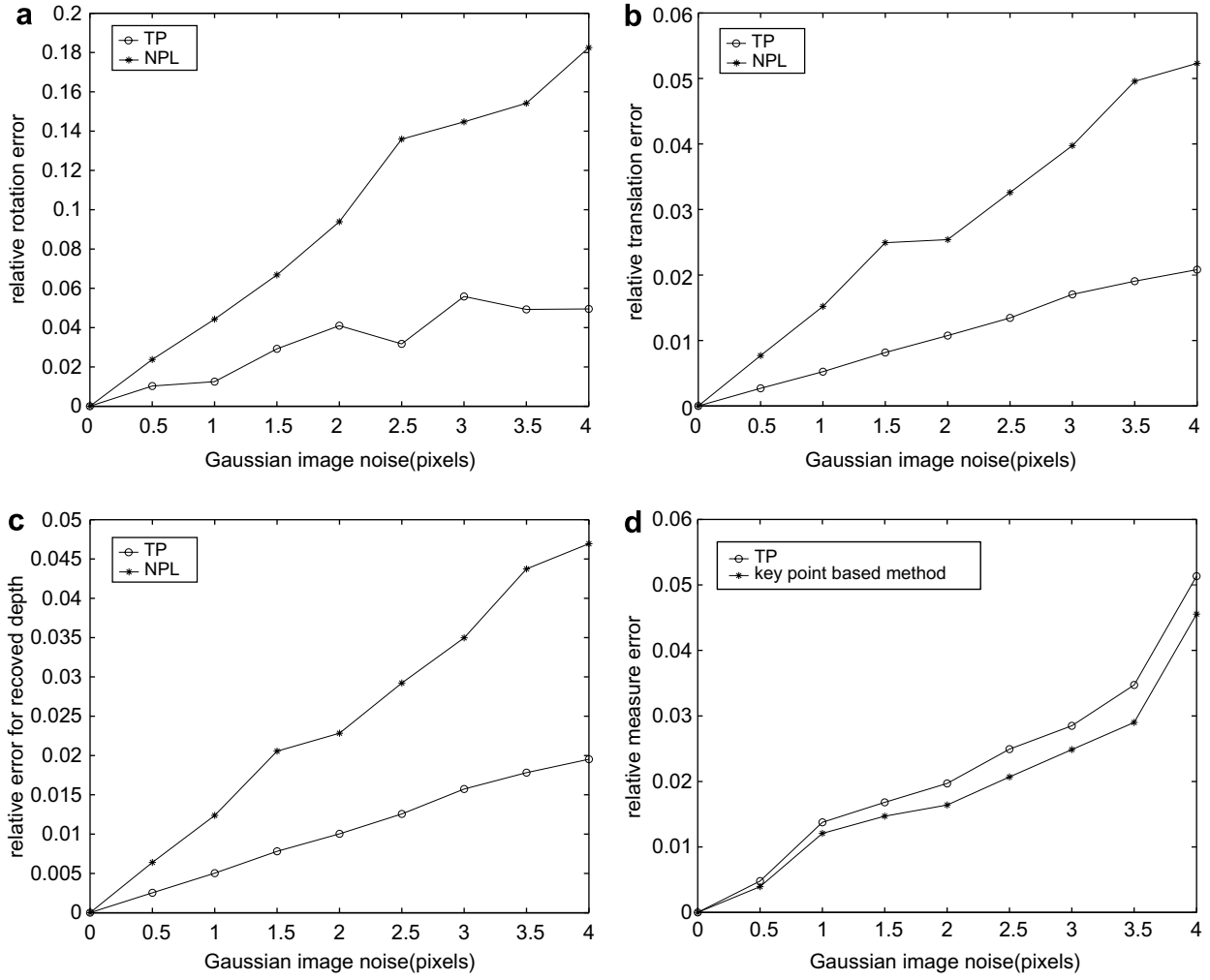

Fig. 3. Simulated result. (a) Rotation error; (b) translation error; (c) depth error and (d) measurement error.

$$d = \|\boldsymbol{X} - \boldsymbol{X}'\| \cdot \sin\theta = \left\| \frac{1}{\tilde{\boldsymbol{m}}^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}} - \frac{1}{\tilde{\boldsymbol{m}}'^{\mathrm{T}}\boldsymbol{\Pi}}\tilde{\boldsymbol{m}}' \right\| \cdot \sqrt{1-\cos^2\theta}.$$

Corollary 2. *If the length of one side of a parallelogram is known, from its image we can compute the length of the other side, the distances from an arbitrary point on the parallelogram plane to the camera center, and other metrical quantities on the parallelogram plane.*

5. Degenerations of the system

Since all conclusions in this paper are deduced from Proposition 2, we only discuss the degeneration of Proposition 2.

From Eqs. (5) and (6), we know the system will degenerate if and only if $\|q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3\| = 0$ or the matrix $\boldsymbol{M} = [-\tilde{\boldsymbol{m}}_1 \quad \tilde{\boldsymbol{m}}_2 \quad \tilde{\boldsymbol{m}}_3]$ is not reversible.

In the first case, we assume the matrix $\boldsymbol{M}$ is reversible.

From $[q_1, q_2, q_3]^{\mathrm{T}} = [-\tilde{\boldsymbol{m}}_1, \tilde{\boldsymbol{m}}_2, \tilde{\boldsymbol{m}}_3]^{-1}\tilde{\boldsymbol{m}}_4$, we have $\|q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3\| = \|q_2\tilde{\boldsymbol{m}}_2 - q_1\tilde{\boldsymbol{m}}_1\|$.

Since $\|q_4\tilde{\boldsymbol{m}}_4 - q_3\tilde{\boldsymbol{m}}_3\| = 0$, we have

$$q_4\tilde{\boldsymbol{m}}_4 = q_3\tilde{\boldsymbol{m}}_3,$$
$$q_2\tilde{\boldsymbol{m}}_2 = q_1\tilde{\boldsymbol{m}}_1.$$

However, it contradicts the condition that the matrix $\boldsymbol{M}$ is reversible. Therefore, the first case is impossible if the matrix $\boldsymbol{M}$ is reversible.

In the second case, the three image points $\boldsymbol{m}_1$, $\boldsymbol{m}_2$ and $\mathbf{m}_3$ are collinear due to the matrix $\boldsymbol{M}$ is not reversible. It means that the three 3D points $\boldsymbol{X}_1$, $\boldsymbol{X}_2$ and $\boldsymbol{X}_3$ are coplanar with the camera center.

In conclusion, the system degenerates only when the plane defined by the trapezium passes through the camera center.

6. Experimental results

The presented algorithms have been experimented both on simulated and real data.

6.1. Synthetic data

During the simulations, the camera's setting is $(f, \alpha f, s, u_0, v_0) = (1500, 1200, 0, 512, 512)$, and the image size is 1024×1024. To study the dependence on noise level, we vary the Gaussian noise from $\sigma = 0$ to 4. For each noise level, we generate 1000 random poses. For each pose, four vertices of a trapezium are generated randomly, and 30 random points on the trapezium plane are also generated. The Gaussian image noise is added to each image point of the 30 points and the trapezium vertices.

Fig. 4. Images of the calibration board.

D1=265.1678	D1=345.9553	D1=426.4786	D1=529.0063	D1=350.9497	D1=393.1261
D2=262.8897	D2=313.0262	D2=435.4758	D2=490.9453	D2=307.7708	D2=443.3357
D3=318.0753	D3=392.3210	D3=484.1185	D3=578.8203	D3=396.8248	D3=448.7529
D4=316.3655	D4=367.1188	D4=489.9169	D4=547.8773	D4=363.2397	D4=489.3346
D5=66.2144	D5=66.0942	D5=67.2406	D5=67.6053	D5=66.8242	D5=65.8508
D6=66.2021	D6=65.5752	D6=68.2211	D6=67.1181	D6=66.4520	D6=66.1465

Fig. 5. Images of the car and the estimated result for each pose. D1, D2, D3, D4 are the recovered depths of the trapezium vertexes. The real lengths of the two non-parallel sides, D5 and D6, are 67 cm.

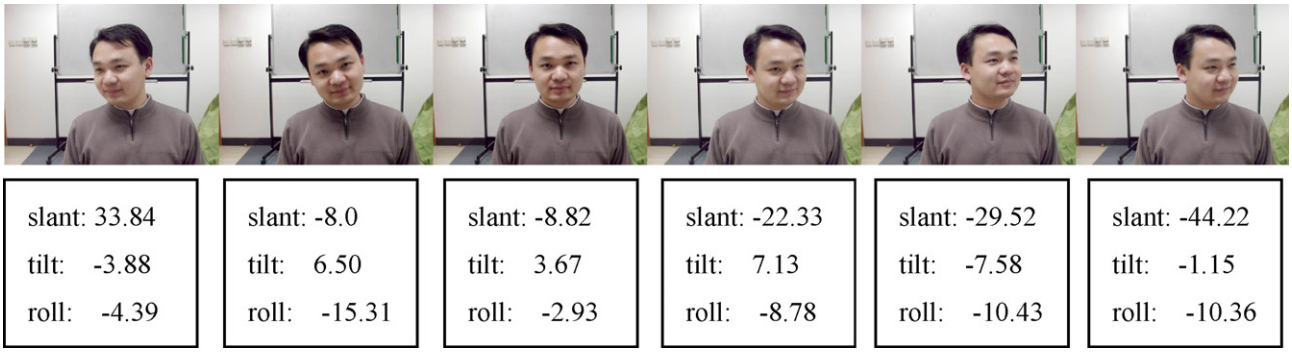

Fig. 6. Images of the face and the estimated rotational parameters for each pose.

For pose estimation, we compare our approach (referred to as TP) to the linear algorithm (NPL) proposed in (Ansar and Daniilidis, 2003). In NPL, the distance between each two vertices must be given. The depth of each vertex to camera center and the projection matrix of each pose are estimated. We report the relative error for rotation, translation and the depth estimation. Each rotation matrix is represented as a Rodrigue vector. The relative error is computed as $\|v_r - v\|/\|v\|$, where v_r is the recovered value, and v is the real value. The test results are shown in Fig. 3a–c, where, the value at each noise level is the mean over 1000 trails. From the figures, we can see that the proposed approach has a good performance, and the errors degrade gracefully with the increasing of the noise level. The proposed method outperforms NPL.

For plane measurement, our approach is compared to the key point based method in (Criminisi, 2001; Criminis et al., 1999a). In the key point based method, the 3D–2D correspondence of the four vertices must be given. For each pose, 30 pairs of points are selected randomly from the 30 points, and the distance of each pair is computed using the two methods. The mean relative error of these distances is shown in Fig. 3d. From the figure, we can see that the key point based method is better. Since the trapezium plane in 3D space is unknown in our method but known in the key point based method, the error of 3D reconstruction of the points is higher in our method, which affects the distance measurement.

6.2. Real images

All images are taken by a Nikon Coolpix 990 digital camera with resolution of 1024×768. We take images for a car, a calibration board and one person's face from six camera poses respectively. Fig. 4 shows the image set of the calibration board, Fig. 5 the car, and Fig. 6 the face. The used trapezium is plotted using white lines on the first image of each image set. The image points of the four vertices on each image are located in a manual way. We adopt

Table 2
Relative error for the test result in Table 1

	Rotation	Translation	D1	D2	D3
1	0.0138	0.0071	0.00005	0.0016	0.0065
2	0.0049	0.0134	0.00002	0.0019	0.0047
3	0.0030	0.0095	0.0031	0.0028	0.0081
4	0.0053	0.0105	0.0031	0.0022	0.0053
5	0.0105	0.0074	0.0022	0.0084	0.0045
6	0.0049	0.0099	0.0019	0.0023	0.0035
Mean	0.0071	0.0096	0.0017	0.0032	0.0054

Table 1
The test result for calibration board

	Method	Rotation			Translation			D1	D2	D3
1	Standard	1.97	1.82	0.50	−55.05	−27.57	277.16	170.00	63.25	117.04
	Our	1.98	1.82	0.53	−54.80	−27.28	279.14	170.09	63.15	116.29
2	Standard	−2.14	−2.19	0.11	−84.96	−46.73	323.78	170.00	78.10	164.01
	Our	−2.14	−2.19	0.10	−84.78	−46.49	328.29	169.98	77.96	164.78
3	Standard	1.98	1.83	0.51	−44.46	−43.68	383.35	170.00	101.98	152.64
	Our	1.97	1.83	0.51	−44.16	−43.48	387.04	170.52	102.26	153.88
4	Standard	2.12	2.04	0.22	−62.58	−42.19	377.02	170.00	70.71	152.64
	Our	2.11	2.03	0.23	−62.36	−41.95	381.03	170.53	70.86	153.45
5	Standard	−2.08	−2.19	0.24	−60.27	−42.59	414.41	170.00	90.00	121.66
	Our	−2.08	−2.19	0.22	−60.33	−42.21	417.51	169.63	90.75	122.21
6	Standard	2.11	2.08	0.15	−59.13	−39.69	475.93	170.00	130.00	156.52
	Our	2.11	2.08	0.14	−59.07	−39.30	480.69	169.67	129.70	155.98

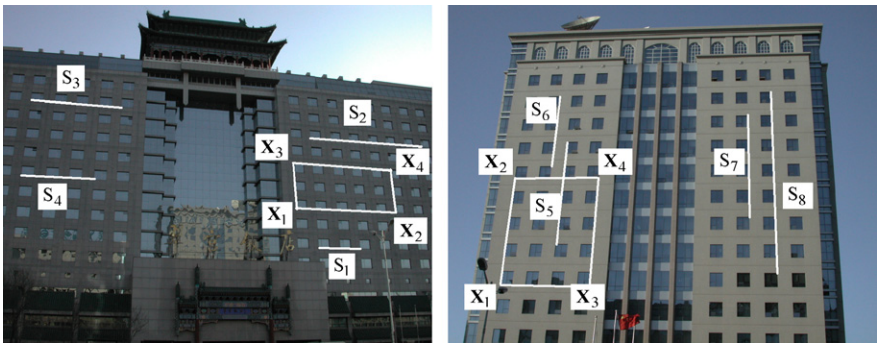

Fig. 7. Two images of the measurement.

Table 3
The measurement result

Line segment	S_1	S_2	S_3	S_4	S_5	S_6	S_7	S_8
Real distance	40	120	80	60	100	75	100	175
Measured distance	38.94	120.79	78.8	59.65	100.06	74.50	99.23	173.14
Relative error %	2.65	0.66	1.50	0.58	0.06	0.67	0.77	1.06

the camera calibration toolbox in (http://www.vision.caltech.edu/bouguetj/) for camera calibration (Zhang, 1991), where, the six images of the calibration board are used.

6.2.1. Experiment for calibration board

The lengths of two sides of the rectangle on the calibration board are 100 and 170 cm. For each pose, we recover the rotation and translation of the camera, the length (D1) of the other side of the rectangle. In addition, we pick up two pairs of random image points on each image to estimate their space distances (D2, D3). The test results are shown in Table 1. The calibration results using the camera calibration toolbox in (http://www.vision.caltech.edu/bouguetj/) are also shown in Table 1 (hereafter referred to as the standard value). For each pose in Table 1, the upper line shows the standard value, and the lower line shows the recovered value. Table 2 shows the relative error of the recovered pose and distances. From the test results, we can see the recovered distances are very close to the real values. The relative rotation errors of the poses 1 and 5, the relative translation errors of the poses 2 and 4 exceed 1%. However, we cannot conclude that the pose estimation results are not good since the standard values are also estimated values. Overall, the proposed approach performs well, and the two estimated results are very close.

6.2.2. Experiment for car

The lengths of sides of the trapezium on the car are measured manually with a scale in centimeter. The lengths of the parallel sides are 110 and 120 cm, and the length of the other sides is 67 cm. In Fig. 5, the recovered depths of the trapezium vertexes (D1, D2, D3, D4) and lengths of the two non-parallel sides (D5, D6) are shown under each image of the car for that pose. From Fig. 5, we can see that the result of the recovered depths is roughly coincident with the real scene and the precision of the estimated lengths is acceptable according to the measurement scale and the measurement error of the image points.

6.2.3. Experiment for face

We estimate the rotational parameters, slant, tilt and roll, by Proposition 4. The vertices of the isosceles trapezium are the outer corners of the two eyes and the mouth. In Fig. 6, the estimated parameters are shown under each image of the face for that pose. Although the true poses of the input faces are not available, we can see from Fig. 6 that the estimated parameters can approximately reflect the poses. However, we cannot expect the results are accurate due to a rough localization of the control points.

6.2.4. Experiment for plane measurement

Fig. 7 shows two images of the test set, where, the first building is Jade Palace Restaurant, and the second one is the satellite mansion. The used rectangles and line segments are plotted in the images. We take the window's size as reference. The real lengths of the line segments are computed according to the reference sizes. The measurement results are shown in Table 3. From the test results, we can learn that the precision of the measurements is acceptable.

7. Conclusions

In the paper, we introduce a new affine invariant of trapezia and deduce the projection of trapezia from this invariant. Using the projection of trapezia, we present a very simple solution of pose determination and plane measurement. For pose determination, if the lengths of the two parallel sides of a trapezium are known, we can obtain the analytical solutions of the distance from each trapezium vertex to the camera center, the camera's position and orientation. For plane measurement, under the same condition we can obtain the analytical solutions of the distance between any two points, the angle of any two lines and other metrical quantities on the trapezium plane. Experiments on simulated and real images show the proposed approach is robust and accurate.

Acknowledgements

We wish to thank the anonymous reviewers for their helpful comments. This work has been partially supported by the National Natural Science Foundation of China (Grant No.60575019 and No.60736008).

References

Ansar, Adnan, Daniilidis, Kostas, 2003. Linear pose estimation from points or lines. IEEE Trans. PAMI 25 (5), 578–589.

Criminisi, A., 2001. Accurate Visual Metrology from Single and Multiple Uncalibrated Images, Distinguished Dissertation Series. Springer-Verlag, London.

Criminis, A., Reid, I., Zisserman, A., 1999a. A plane measuring device. Image Vision Comput. 17 (8), 625–634.

Criminis, A., Reid, I., Zisserman, A., 1999b. Single view metrology. ICCV, 434–442.

Fishler, M.A., Bolles, R.C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartomated cartography. Commun. ACM 24 (6), 381–395.

Gao, X.S., Hou, X.R., Tang, J.L., Cheng, H., 2003. Complete solution classification for the perspective-three-point problem. IEEE Trans. PAMI 25 (8), 930–943.

Haralick, R.M., 1989. Determining camera parameters from the perspective projection of a rectangle. Pattern Recognition 22 (3), 225–230.

Haralick, R.M., Lee, C., Ottenberg, K., Nolle, M., 1991. Analysis and solutions of the three point perspective pose estimation problem. CVPR, 592–598.

Hartely, R., Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press.

Horaud, R., Conio, B., Leboulleux, O., 1989. An analytic solution for the perspective 4-point problem. CVGIP 47 (1), 33–44.

Hu, Z.Y., Wu, F.C., 2002. A note on the number solution of the non-coplanar P4P problem. IEEE Trans. PAMI 24 (4), 550–555.

Liebowitz, D., Criminisi, A., Zisserman, A., 1999. Creating architectural models from images. In: Proceedings of Eurographics, pp. 39–50.

Liu, M.L., Wong, K.H., 1999. Pose estimation using four corresponding points. Pattern Recognition Lett. 20 (1), 69–74.

Lowe, D.G., 1991. Fitting parameterized three-dimensional models to images. IEEE Trans. PAMI 13 (5), 441–450.

Mundy, J.L., Zisserman, A., 1992. Geometric Invariance in Computer Vision. MIT Press, Cambridge MA.

Mundy, J.L., Zisserman, A., Forsyth, D.A., 1994. Applications of Invariance in Computer Vision. In: LNCS 825. Springer, Berlin.

Penna, M.A., 1991. Determining camera parameters from the perspective projection of a quadrilateral. Pattern Recognition 24 (6), 533–541.

Quan, L., Lan, Z., 1999. Linear *N*-point camera pose determination. IEEE Trans. PAMI 21 (8), 774–780.

Wu, F.C., Hu, Z.Y., 2001. A study on the P5P problem. Chinese J. Software 12 (5), 768–775 (in Chinese).

Yuan, J.S.C., 1989. A general photogrammetric method for determining object position and orientation. IEEE Trans. Robot. Automat. 5 (2), 129–142.

Zhang, Zhengyou, 1991. Flexible camera calibration by viewing a plane from unknown orientations. ICCV, 666–673.