Available online at www.sciencedirect.com
SCIENCE DIRECT®
Pattern Recognition Letters 25 (2004) 1399–1410
Pattern Recognition Letters
www.elsevier.com/locate/patrec

A novel 3D planar object reconstruction from multiple uncalibrated images using the plane-induced homographies

H.L. Chou, Z. Chen *

Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu, Taiwan

Received 4 April 2003; received in revised form 27 December 2003
Available online 1 July 2004

Abstract

A computer vision method is proposed to determine all the visible 3D planar surfaces in a scene from uncalibrated images and locate them in a single 3D projective space. Most of the existing methods for reconstructing planar objects use point correspondences to estimate the fundamental matrix and derive the compatible projection equations, before they apply the standard triangulation technique to find the 3D points and fit the planes to the 3D points. This type of approaches is generally slow and less accurate because the 3D points are estimated separately, making them vulnerable to image error. We present a plane based reconstruction method to estimate the 3D projective structure using the planar homographies estimated from the plane features in the images. First, we estimate the homography for each visible plane, and then we use the homographies of two primary planes to compute an epipole. We proceed to represent the epipolar geometry for each image pair using the estimated homography and epipole, together with a specified reference plane coefficient vector. Next, we show that the 3D plane coefficient vector of any plane visible in each image pair can be determined with respect to the reference plane coefficient vector once its planar homography is found. Finally, the reconstruction results obtained in individual projective spaces are integrated within a common projective space. To this end, we use the homography and plane equation information of two planes and the epipole associated to derive the coordinate transformation matrix between two involved projective spaces. To evaluate the performance of our method, we apply our method to the synthetic images and real images. All the results indicate the method works successfully.

Keywords: Computer vision; 3D projective reconstruction; Plane-based projective reconstruction; Uncalibrated camera; Homography; Projective geometry; Reconstruction integration

1. Introduction

In the physical world (especially the man-made world) planar surfaces such as walls, windows, table, roof, road, and terrace can be found in the indoor as well as the outdoor scenes. Our task is to reconstruct the 3D planar surfaces in a scene from multiple uncalibrated images taken by a camera

* Corresponding author. Tel.: +886-3-573-1875; fax: +886-3-572-3148.
E-mail address: zchen@csie.nctu.edu.tw (Z. Chen).

0167-8655/$ - see front matter © 2004 Published by Elsevier B.V.
doi:10.1016/j.patrec.2004.05.018

placed at different viewpoints. In general, the methods for 3D projective or uncalibrated reconstruction (Mohr and Arbogast, 1991; Faugeras, 1992, 1993; Hartley et al., 1992, 1994; Beardsley et al., 1997) are point-based. They estimate the fundamental matrix from a sufficient number of corresponding point pairs first, and then derive the epipole and the canonical geometric representation for projective views using the fundamental matrix. Then, for each pair of corresponding points, they use a triangulation technique or bundle adjustment technique to compute the 3D point coordinates in the projective space. Finally, for the determination of the uncalibrated planar scene structure (Luong and Faugeras, 1993; Sawhney, 1994; Criminisi and Zisserman, 1998; Irani et al., 1998; Szeliski and Torr, 1998; Fradkin et al., 1999; Johansson, 1999; Zelnik-Manor and Irani, 2000), the 3D points found are fitted by planes. However, it is desirable to derive the 3D planar scene structure in terms of plane features in the images directly, for these features are more reliable than the point or line features (Luong and Faugeras, 1993). The estimation of the 3D projective planar structure based on the projected plane feature information exclusively has not yet received much attention, although it is known that the corresponding projected plane regions in a pair of stereo images induce a homography. It is also known that homographies are useful to many other practical applications including:

(a) Fundamental matrix estimation or canonical projective geometry representation (Luong and Vieville, 1996; Luong and Faugeras, 1993).
(b) 2D image mosaicing or view synthesis (Szeliski, 1996).
(c) Plane + parallax analysis (Irani et al., 1998; Criminisi and Zisserman, 1998; Sawhney, 1994).
(d) Planar motion estimation and ego-motion (Irani et al., 1997; Szeliski and Torr, 1998; Zelnik-Manor and Irani, 2000).

Recently, two methods have been proposed for the 3D projective reconstruction of planes and cameras. The first method assumes all planes are visible in all images and the second method assumes a reference plane is visible in all images (Rother et al., 2002, 2003). In practice, it is not realistic to have all planes or even one plane visible in all images unless a very large ground plane is available. When there is no reference plane visible in all images, the reconstruction problem cannot be formulated within a common projective space and the reconstruction results will be inevitably obtained in different projective spaces.

We shall recover the 3D scene planar structure from the uncalibrated images using the plane-induced homographies without assuming that all planes or one plane must be seen in all images. To obtain the homographies, we must locate the projected regions of planar surfaces in the images. There are methods for detecting regions corresponding to planar surfaces in the image (Sinclair and Blake, 1996; Hamid and Cipolla, 1997; Theiler and Chabbi, 1999). After the image regions of planar surfaces have been extracted, we use the Gabor filtering technique (Sun et al., 2002) to identify at least four point correspondences for every plane in the stereo images in order to obtain the initial value of the homography. Then we iteratively refine the homography based on a nonlinear minimization method given in (Szeliski, 1996). Next, we use two homographies to compute the epipole and to find the compatible projection equations in terms of the estimated homography and an assigned plane coefficient vector of a reference plane, together with the estimated epipole. With the projection equations thus derived we then prove that the 3D equation of any other plane visible in the stereo images can be computed with respect to the reference plane equation as long as its homography is determined. Finally, we merge or integrate all reconstructed plane equations found in individual projective spaces within a common space through the coordinate (or space) transformations. Again, each required coordinate transformation matrix is expressed by the homography and plane coefficient vector information of two planes visible in the involved image pairs. Fig. 1 shows the flow diagram of our method.

The remaining sections of the paper are organized as follows. Section 2 is the preliminaries and mathematical notations for the projective recon-

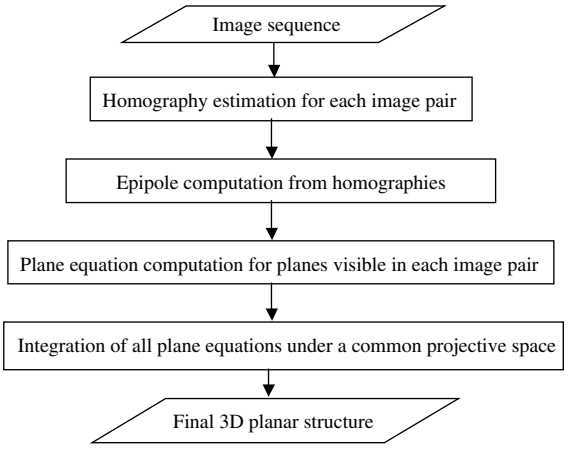

Fig. 1. The flow diagram of our reconstruction method.

struction. Section 3 shows how the 3D equations of all planar surfaces visible in the stereo images can be determined from their homographies. Section 4 presents the integration of the reconstruction results obtained in different projective spaces through the coordinate transformations. Section 5 shows the estimation of the plane-induced homographies and the related epipole. Section 6 reports the experimental results on both the synthetic and real images. Section 7 is the concluding remarks.

2. Preliminaries and mathematical notations for projective reconstruction

Consider any two consecutive images $(\boldsymbol{I}_i, \boldsymbol{I}_j)$ in an image sequence for reconstructing the visible planar surfaces. Let $\boldsymbol{R}_i$, $\vec{\boldsymbol{t}}_i$ be the extrinsic parameters and $\boldsymbol{M}_i$ be the 3×3 upper triangular intrinsic camera matrix of the ith camera. Then the coordinates of a 3D point $\vec{\boldsymbol{p}}_E = [\,x_E \quad y_E \quad z_E\,]^{\mathrm{T}}$ and its 2D projection point $\vec{\boldsymbol{u}}_i = [\,u_i \quad v_i\,]^{\mathrm{T}}$ in image $\boldsymbol{I}_i$ are related by a pinhole camera model (Hartley et al., 1992; Faugeras, 1993; Hartley and Zisserman, 2000):

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \cong \boldsymbol{M}_i[\boldsymbol{R}_i|\vec{\boldsymbol{t}}_i] \begin{bmatrix} x_E \\ y_E \\ z_E \\ 1 \end{bmatrix}.$$

To represent the point in the projective space or the homogeneous coordinate system, we use the vectors with a tilde to denote the homogeneous coordinates of the 3D points and its image projection point such that $\tilde{\boldsymbol{u}}_i = \begin{bmatrix} \vec{\boldsymbol{u}}_i^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}}$ and $\tilde{\boldsymbol{p}}_E = \begin{bmatrix} \vec{\boldsymbol{p}}_E^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}}$. The symbol $\cong$ indicates an equality up to a nonzero scale in the homogeneous coordinate system.

Assume the world coordinate system is chosen to be the ith camera coordinate system; namely, $\boldsymbol{R}_i = \boldsymbol{I}$ and $\vec{\boldsymbol{t}}_i = \vec{\boldsymbol{0}}$.

$$\tilde{\boldsymbol{u}}_i \cong \boldsymbol{M}_i \cdot [\boldsymbol{I}|\vec{\boldsymbol{0}}] \cdot \tilde{\boldsymbol{p}}_E \triangleq \boldsymbol{M}_i \cdot \vec{\boldsymbol{p}}_E. \tag{1}$$

Similarly, let $\boldsymbol{M}_j$, $\boldsymbol{R}_j$, $\vec{\boldsymbol{t}}_j$ be the camera parameters of the jth camera. For image $\boldsymbol{I}_j$, we have

$$\tilde{\boldsymbol{u}}_j \cong \boldsymbol{M}_j \cdot [\boldsymbol{R}_j|\vec{\boldsymbol{t}}_j] \cdot \tilde{\boldsymbol{p}}_E. \tag{2}$$

Since the epipole on image $\boldsymbol{I}_j$ is given by $\tilde{\boldsymbol{e}}_j \cong \boldsymbol{M}_j[\boldsymbol{R}_j|\vec{\boldsymbol{t}}_j][\,0 \quad 0 \quad 0 \quad 1\,]^{\mathrm{T}} \cong \boldsymbol{M}_j \cdot \vec{\boldsymbol{t}}_j$ or $\lambda_e\tilde{\boldsymbol{e}}_j = \boldsymbol{M}_j\vec{\boldsymbol{t}}_j$ (λ_e is the lens depth parameter), we rewrite Eq. (2) as

$$\tilde{\boldsymbol{u}}_j \cong [\boldsymbol{M}_j\boldsymbol{R}_j|\lambda_e\tilde{\boldsymbol{e}}_j] \cdot \tilde{\boldsymbol{p}}_E. \tag{3}$$

Consider a plane Π_{A}, which does not pass through the optical center of the ith camera (otherwise, its image will be degenerated into a line). Let its plane equation be $\vec{\boldsymbol{a}}_E^{\mathrm{T}}\vec{\boldsymbol{p}}_E + 1 = 0$ with $\vec{\boldsymbol{a}}_E^{\mathrm{T}} = [\,a_{E1} \quad a_{E2} \quad a_{E3}\,]^{\mathrm{T}}$. After eliminating the variable $\vec{\boldsymbol{p}}_E$ in the two projection Eqs. (1) and (3), we can obtain a homography $\boldsymbol{A}_{ij}$ as follows (Tsai and Huang, 1982; Luong and Vieville, 1996; Szeliski, 1996):

$$\tilde{\boldsymbol{u}}_j \cong \boldsymbol{A}_{ij}\tilde{\boldsymbol{u}}_i \quad \text{with } \boldsymbol{A}_{ij} \cong \left\{\boldsymbol{M}_j\boldsymbol{R}_j\boldsymbol{M}_i^{-1} - \lambda_e\tilde{\boldsymbol{e}}_j\vec{\boldsymbol{a}}_E^{\mathrm{T}}\boldsymbol{M}_i^{-1}\right\}.$$

The homography $\boldsymbol{A}_{ij}$ from image $\boldsymbol{I}_i$ to image $\boldsymbol{I}_j$ is said to be induced by plane Π_{A}. In Section 5 we shall show how to compute the homography $\boldsymbol{A}_{ij}$ from image pair $(\boldsymbol{I}_i, \boldsymbol{I}_j)$.

For an uncalibrated camera the intrinsic and extrinsic camera parameters in Eqs. (1) and (3) cannot be estimated. We need to replace these two equations by some new parameters that can be estimated. This is done as follows:

Let $\boldsymbol{A}_{ij}$ be rewritten as $\boldsymbol{A}_{ij} = \lambda_{\mathrm{A}}\{\boldsymbol{M}_j\boldsymbol{R}_j\boldsymbol{M}_i^{-1} - \lambda_e\tilde{\boldsymbol{e}}_j\vec{\boldsymbol{a}}_E^{\mathrm{T}}\boldsymbol{M}_i^{-1}\}$, then

$$\boldsymbol{M}_j\boldsymbol{R}_j = \frac{1}{\lambda_{\mathrm{A}}}\boldsymbol{A}_{ij}\boldsymbol{M}_i + \lambda_e\tilde{\boldsymbol{e}}_j\vec{\boldsymbol{a}}_E^{\mathrm{T}}.$$

Eq. (3) can be rewritten as

$$\begin{aligned}\tilde{\boldsymbol{u}}_j &\cong [\boldsymbol{M}_j\boldsymbol{R}_j \mid \lambda_e\tilde{\boldsymbol{e}}_j]\begin{bmatrix}\vec{\boldsymbol{p}}_E\\ 1\end{bmatrix}\\ &\cong \frac{1}{\lambda_{\mathrm{A}}}[\boldsymbol{A}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c}\boldsymbol{M}_i & \boldsymbol{0}\\ \hline \lambda_{\mathrm{A}}\lambda_e\vec{\boldsymbol{a}}_E^{\mathrm{T}} & \lambda_{\mathrm{A}}\lambda_e\end{array}\right]\begin{bmatrix}\vec{\boldsymbol{p}}_E\\ 1\end{bmatrix}\\ &\cong \frac{1}{\lambda_{\mathrm{A}}}[\boldsymbol{A}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c}\boldsymbol{I} & \boldsymbol{0}\\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right]\left[\begin{array}{c|c}\boldsymbol{I} & \boldsymbol{0}\\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right]^{-1}\\ &\quad\times\left[\begin{array}{c|c}\boldsymbol{M}_i & \boldsymbol{0}\\ \hline \lambda_{\mathrm{A}}\lambda_e\vec{\boldsymbol{a}}_E^{\mathrm{T}} & \lambda_{\mathrm{A}}\lambda_e\end{array}\right]\begin{bmatrix}\vec{\boldsymbol{p}}_E\\ 1\end{bmatrix}\end{aligned}$$

where $\tilde{\boldsymbol{a}}_{ij}$ (with $a_{ij4}\neq 0$) is assumed to be a nonnull column vector. Then

$$\begin{aligned}\tilde{\boldsymbol{u}}_j &\cong \frac{1}{\lambda_{\mathrm{A}}}[\boldsymbol{A}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c}\boldsymbol{I} & \boldsymbol{0}\\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right]\\ &\quad\times\left[\begin{array}{c|c}\boldsymbol{M}_i & \vec{\boldsymbol{0}}\\ \hline \dfrac{-\vec{\boldsymbol{a}}_{ij}^{\mathrm{T}}}{a_{ij4}}\boldsymbol{M}_i + \dfrac{\lambda_{\mathrm{A}}\lambda_e}{a_{ij4}}\vec{\boldsymbol{a}}_E^{\mathrm{T}} & \dfrac{\lambda_{\mathrm{A}}\lambda_e}{a_{ij4}}\end{array}\right]\begin{bmatrix}\vec{\boldsymbol{p}}_E\\ 1\end{bmatrix} \qquad (4)\\ &\cong \frac{1}{\lambda_{\mathrm{A}}}[\boldsymbol{A}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c}\boldsymbol{I} & \boldsymbol{0}\\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right]\begin{bmatrix}\vec{\boldsymbol{p}}_{ij}\\ p_{ij4}\end{bmatrix}.\end{aligned}$$

A similar formulation of Eq. (4) has been derived in (Luong and Vieville, 1996). In this way, the original Euclidean point $\tilde{\boldsymbol{p}}_E$ becomes point $\tilde{\boldsymbol{p}}_{ij} = [\vec{\boldsymbol{p}}_{ij}^{\mathrm{T}} \;\; p_{ij4}]^{\mathrm{T}}$ in the new projective space, denoted by $\{\tilde{\boldsymbol{p}}_{ij}\}$, which describes the projective geometry associated with images i and j.

The coordinate transformation from the Euclidean space $\{[[\vec{\boldsymbol{p}}_E^{\mathrm{T}} \;\; 1]^{\mathrm{T}}]\}$ to the projective space $\{[[\vec{\boldsymbol{p}}_{ij}^{\mathrm{T}} \;\; p_{ij4}]^{\mathrm{T}}]\}$ is given by

$$\vec{\boldsymbol{p}}_{ij} = \boldsymbol{M}_i\vec{\boldsymbol{p}}_E,$$

$$p_{ij4} = \left(\frac{\lambda_{\mathrm{A}}\lambda_e}{a_{ij4}}\right)\left(\vec{\boldsymbol{a}}_E^{\mathrm{T}}\vec{\boldsymbol{p}}_E + 1\right) - \frac{\vec{\boldsymbol{a}}_{ij}^{\mathrm{T}}}{a_{ij4}}\boldsymbol{M}_i\vec{\boldsymbol{p}}_E.$$

Also,

$$\begin{aligned}\vec{\boldsymbol{a}}_{ij}^{\mathrm{T}}\vec{\boldsymbol{p}}_{ij} + a_{ij4}p_{ij4} &= \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}}\vec{\boldsymbol{p}}_{ij} + (\lambda_{\mathrm{A}}\lambda_e)\left(\vec{\boldsymbol{a}}_E^{\mathrm{T}}\vec{\boldsymbol{p}}_E + 1\right) - \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}}\vec{\boldsymbol{p}}_{ij}\\ &= \lambda_{\mathrm{A}}\lambda_e\left(\vec{\boldsymbol{a}}_E^{\mathrm{T}}\vec{\boldsymbol{p}}_E + 1\right) = 0.\end{aligned}$$

It implies that $\vec{\boldsymbol{a}}_{ij}^{\mathrm{T}}\vec{\boldsymbol{p}}_{ij} + a_{ij4}p_{ij4} = 0$ is the new 3D equation of the plane in the projective space $\{\tilde{\boldsymbol{p}}_{ij}\}$. Since the projective structure can only be determined up to a 4×4 nonsingular projective matrix (Hartley and Zisserman, 2000), the new plane coefficient vector $\tilde{\boldsymbol{a}}_{ij} = [\vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} \;\; a_{ij4}]^{\mathrm{T}}$ can take on some general value, say, $[1 \;\; 1 \;\; 1 \;\; 1]^{\mathrm{T}}$ (more discussion on the values of $\tilde{\boldsymbol{a}}_{ij}$ is given in Section 6). In this new space the parameters including homography $\boldsymbol{A}_{ij}$, epipole $\tilde{\boldsymbol{e}}_j$ and plane coefficients $\tilde{\boldsymbol{a}}_{ij} = [\vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} \;\; a_{ij4}]^{\mathrm{T}}$ involved in Eq. (4) are now all known.

Next, we shall describe how to obtain the projective reconstruction for the other planes visible in the image pair $(\boldsymbol{I}_i, \boldsymbol{I}_j)$ in the newly defined projective space $\{\tilde{\boldsymbol{p}}_{ij}\}$.

3. Reconstruction of all visible planes from a given image pair

In the new projective space the projection equations become

$$\begin{aligned}\tilde{\boldsymbol{u}}_i &\cong \left[\boldsymbol{I}\middle|\vec{\boldsymbol{0}}\right]\begin{bmatrix}\vec{\boldsymbol{p}}_{ij}\\ \boldsymbol{p}_{ij4}\end{bmatrix},\\ \tilde{\boldsymbol{u}}_j &\cong \frac{1}{\lambda_{\mathrm{A}}}[\boldsymbol{A}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c}\boldsymbol{I} & \vec{\boldsymbol{0}}\\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right]\begin{bmatrix}\vec{\boldsymbol{p}}_{ij}\\ p_{ij4}\end{bmatrix}.\end{aligned} \qquad (5)$$

Similarly, for any other plane Π_{B} visible in $(\boldsymbol{I}_i, \boldsymbol{I}_j)$ the induced homography between the plane regions in image pair $(\boldsymbol{I}_i, \boldsymbol{I}_j)$ is expressed by

$$\boldsymbol{B}_{ij} = \lambda_{\mathrm{B}}\left\{\boldsymbol{M}_j\boldsymbol{R}_j\boldsymbol{M}_i^{-1} - \lambda_e\tilde{\boldsymbol{e}}_j\vec{\boldsymbol{b}}_E^{\mathrm{T}}\boldsymbol{M}_i^{-1}\right\} \qquad (6)$$

with the plane equation of Π_{B} being $\vec{\boldsymbol{b}}_E^{\mathrm{T}}\vec{\boldsymbol{p}}_E + 1 = 0$.

Next, we shall prove the fact that the relation between plane coefficient vectors of planes Π_{B} and Π_{A} is determined once their homographies $\boldsymbol{A}_{ij}$ and $\boldsymbol{B}_{ij}$ are found. From above we have

$$\begin{aligned}\boldsymbol{M}_j\boldsymbol{R}_j\boldsymbol{M}_i^{-1} &= \frac{1}{\lambda_{\mathrm{A}}}\boldsymbol{A}_{ij} + \lambda_e\tilde{\boldsymbol{e}}_j\vec{\boldsymbol{a}}_E^{\mathrm{T}}\boldsymbol{M}_i^{-1}\\ &= \frac{1}{\lambda_{\mathrm{B}}}\boldsymbol{B}_{ij} + \lambda_e\tilde{\boldsymbol{e}}_j\vec{\boldsymbol{b}}_E^{\mathrm{T}}\boldsymbol{M}_i^{-1}\end{aligned}$$

or

$$\boldsymbol{A}_{ij} = \frac{\lambda_A}{\lambda_B}\boldsymbol{B}_{ij} + \lambda_A\lambda_e\tilde{\boldsymbol{e}}_j\left[\vec{\boldsymbol{b}}_E^{\mathrm{T}} - \vec{\boldsymbol{a}}_E^{\mathrm{T}}\right]\boldsymbol{M}_i^{-1} \triangleq \frac{\lambda_A}{\lambda_B}\boldsymbol{B}_{ij} + \tilde{\boldsymbol{e}}_j\vec{\boldsymbol{\eta}}_{ij}^{\mathrm{T}}, \tag{7}$$

where $\vec{\boldsymbol{\eta}}_{ij} = \lambda_A\lambda_e[\vec{\boldsymbol{b}}_E - \vec{\boldsymbol{a}}_E]\boldsymbol{M}_i^{-1}$. We can apply the least-squares method to estimate the unknowns $\vec{\boldsymbol{\eta}}_{ij}$ and $\frac{\lambda_A}{\lambda_B}$ in a system of nine linear equations; here the epipole $\tilde{\boldsymbol{e}}_j$ can be determined in advance from the two homographies $\boldsymbol{A}_{ij}$ and $\boldsymbol{B}_{ij}$ based on the fact that $\boldsymbol{B}^{\mathrm{T}}[\tilde{\boldsymbol{e}}_j]_\times\boldsymbol{A}(\cong \boldsymbol{B}^{\mathrm{T}}[\tilde{\boldsymbol{e}}_j]_\times\boldsymbol{B})$ is skew symmetric.

Therefore,

$$\tilde{\boldsymbol{u}}_j \cong \frac{1}{\lambda_A}[\boldsymbol{A}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right]\left[\begin{array}{c}\vec{\boldsymbol{p}}_{ij}\\ \hline p_{ij4}\end{array}\right]$$

$$\cong \frac{1}{\lambda_B}\left[\boldsymbol{B}_{ij} \mid \tilde{\boldsymbol{e}}_j\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4}\end{array}\right]\left[\begin{array}{c}\vec{\boldsymbol{p}}_{ij}\\ \hline p_{ij4}\end{array}\right].$$

Substituting $\boldsymbol{A}_{ij} = \frac{\lambda_A}{\lambda_B}\boldsymbol{B}_{ij} + \tilde{\boldsymbol{e}}_j\vec{\boldsymbol{\eta}}_{ij}^{\mathrm{T}}$ into the above equation, we have

$$\frac{1}{\lambda_A}\left[\frac{\lambda_A}{\lambda_B}\boldsymbol{B}_{ij} + \tilde{\boldsymbol{e}}_j\vec{\boldsymbol{\eta}}_{ij}^{\mathrm{T}} \;\middle|\; \tilde{\boldsymbol{e}}_j\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right]$$

$$\cong \frac{1}{\lambda_B}\left[\boldsymbol{B}_{ij} \mid \tilde{\boldsymbol{e}}_j\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4}\end{array}\right].$$

Then

$$[\boldsymbol{B}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \frac{\lambda_B}{\lambda_A}\vec{\boldsymbol{\eta}}_{ij}^{\mathrm{T}} & \frac{\lambda_B}{\lambda_A}\end{array}\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right] \cong \left[\boldsymbol{B}_{ij} \mid \tilde{\boldsymbol{e}}_j\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4}\end{array}\right].$$

Since $a_{ij4} \neq 0$ and $b_{ij4} \neq 0$ (i.e., planes Π_A and Π_B do not contain the lens center), we obtain,

$$\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \frac{\lambda_B}{\lambda_A}\vec{\boldsymbol{\eta}}_{ij}^{\mathrm{T}} & \frac{\lambda_B}{\lambda_A}\end{array}\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} & a_{ij4}\end{array}\right] \cong \left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4}\end{array}\right].$$

Thus

$$\left[\frac{\lambda_B}{\lambda_A}\vec{\boldsymbol{\eta}}_{ij}^{\mathrm{T}} + \frac{\lambda_B}{\lambda_A}\vec{\boldsymbol{a}}_{ij}^{\mathrm{T}} \;\middle|\; \frac{\lambda_B}{\lambda_A}a_{ij4}\right] \cong \left[\vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} \middle| b_{ij4}\right].$$

In other words, the relationship between two plane coefficient vectors is given by

$$\tilde{\boldsymbol{b}}_{ij} \cong \left(\tilde{\boldsymbol{a}}_{ij} + \begin{bmatrix}\vec{\boldsymbol{\eta}}_{ij}\\ 0\end{bmatrix}\right). \tag{8}$$

Thus, $\tilde{\boldsymbol{b}}_{ij}$ can be determined with respect to $\tilde{\boldsymbol{a}}_{ij}$ once the planar homography $\boldsymbol{B}_{ij}$ is known.

4. Integration of planes reconstructed from different image pairs

Next, we consider the integration of reconstructed planes obtained from different image pairs $(\boldsymbol{I}_i, \boldsymbol{I}_j)$ and $(\boldsymbol{I}_j, \boldsymbol{I}_k)$, which contain the projections of two commonly visible planes. We shall use the plane-based coordinate transformation method for integrating the reconstruction results defined in different spaces.

Let the 4×4 coordinate transformation matrix $\boldsymbol{H}_{ijk}$, mapping the points in the projective space $\{\tilde{\boldsymbol{p}}_{ij}\}$ to the points in the projective space $\{\tilde{\boldsymbol{p}}_{jk}\}$, be defined by

$$\tilde{\boldsymbol{p}}_{jk} = \boldsymbol{H}_{ijk}\tilde{\boldsymbol{p}}_{ij}.$$

Then, the plane coefficient vectors $\tilde{\boldsymbol{c}}_{ij}$, $\tilde{\boldsymbol{c}}_{jk}$ of a common plane, which are respectively defined in the two different projective spaces $\{\tilde{\boldsymbol{p}}_{ij}\}$ and $\{\tilde{\boldsymbol{p}}_{jk}\}$, will be related by:

$$\tilde{\boldsymbol{c}}_{jk} \cong \boldsymbol{H}_{ijk}^{-\mathrm{T}}\tilde{\boldsymbol{c}}_{ij}. \tag{9}$$

Thus, it requires the information of five common planes in the two different projective spaces in order to solve for the transformation matrix $\boldsymbol{H}_{ijk}$. It is usually not very practical to find five common planes in the image pairs.

On the other hand, the two respective 3×4 projection matrices associated with image $\boldsymbol{I}_j$ defined in the two projective spaces $\{\tilde{\boldsymbol{p}}_{ij}\}$ and $\{\tilde{\boldsymbol{p}}_{jk}\}$ are related directly by the matrix $\boldsymbol{H}_{ijk}$ (Fitzgibbon and Zisserman, 1998). This relationship provides 11 linear equations in the 15 matrix elements in $\boldsymbol{H}_{ijk}$. Then, it is reduced to a need of two plane information to provide six additional linear equations to solve for the 15 unknowns. In the following we shall give a system of 24 linear equations using the information of two planes for solving for the 15 unknowns; the result will be more reliable.

From Eq. (4), we have

$$l_j\tilde{\boldsymbol{u}}_j = \frac{1}{\lambda_B}[\boldsymbol{B}_{ij} \mid \tilde{\boldsymbol{e}}_j]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4}\end{array}\right]\tilde{\boldsymbol{p}}_{ij}.$$

Combining this equation with a plane equation $\tilde{\boldsymbol{b}}_{ij}^{\mathrm{T}}\tilde{\boldsymbol{p}}_{ij}=0$, we have

$$\begin{bmatrix} l_j\tilde{\boldsymbol{u}}_j \\ 0 \end{bmatrix} = \frac{1}{\lambda_{\mathrm{B}}}\left[\begin{array}{c|c} \boldsymbol{B}_{ij} & \tilde{\boldsymbol{e}}_j \\ \hline \vec{\boldsymbol{0}}^{\mathrm{T}} & 1 \end{array}\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4} \end{array}\right]\tilde{\boldsymbol{p}}_{ij}.$$

Thus,

$$\tilde{\boldsymbol{p}}_{ij}=\lambda_{\mathrm{B}}\left(\left[\begin{array}{c|c} \boldsymbol{B}_{ij} & \tilde{\boldsymbol{e}}_j \\ \hline \vec{\boldsymbol{0}}^{\mathrm{T}} & 1 \end{array}\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4} \end{array}\right]\right)^{-1}\begin{bmatrix} l_j\tilde{\boldsymbol{u}}_j \\ 0 \end{bmatrix}.$$

Similarly, for the same point on plane Π_{B}, but represented as $\tilde{\boldsymbol{p}}_{jk}$ in the different projective space $\{\tilde{\boldsymbol{p}}_{jk}\}$, we can relate it to the same 4×1 vector $\begin{bmatrix} l_j\tilde{\boldsymbol{u}}_j^{\mathrm{T}} & 0 \end{bmatrix}^{\mathrm{T}}$ by

$$\tilde{\boldsymbol{p}}_{jk}=\left(\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{0}^{\mathrm{T}} & 1 \end{array}\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{0} \\ \hline \vec{\boldsymbol{b}}_{jk}^{\mathrm{T}} & b_{jk4} \end{array}\right]\right)^{-1}\begin{bmatrix} l_j\tilde{\boldsymbol{u}}_j \\ 0 \end{bmatrix}.$$

Then

$$\left(\left[\begin{array}{c|c} \boldsymbol{I} & \vec{\boldsymbol{0}} \\ \hline \vec{\boldsymbol{0}}^{\mathrm{T}} & 1 \end{array}\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{0} \\ \hline \vec{\boldsymbol{b}}_{jk}^{\mathrm{T}} & b_{jk4} \end{array}\right]\right)^{-1}\begin{bmatrix} l_j\tilde{\boldsymbol{u}}_j \\ 0 \end{bmatrix} = \boldsymbol{H}_{ijk}\,\lambda_{\mathrm{B}}\left(\left[\begin{array}{c|c} \boldsymbol{B}_{ij} & \tilde{\boldsymbol{e}}_j \\ \hline \vec{\boldsymbol{0}}^{\mathrm{T}} & 1 \end{array}\right]\left[\begin{array}{c|c} \boldsymbol{I} & \vec{0} \\ \hline \vec{\boldsymbol{b}}_{ij}^{\mathrm{T}} & b_{ij4} \end{array}\right]\right)^{-1}\begin{bmatrix} l_j\tilde{\boldsymbol{u}}_j \\ 0 \end{bmatrix}.$$

After some algebraic manipulation, this can be reduced to

$$\begin{bmatrix} \boldsymbol{I} \\ -\boldsymbol{b}_{jk}^{\mathrm{T}}/b_{jk4} \end{bmatrix}\tilde{\boldsymbol{u}}_{\boldsymbol{j}} = \lambda_{\mathrm{B}}\boldsymbol{H}_{ijk}\begin{bmatrix} \boldsymbol{B}_{ij}^{-1} \\ -\boldsymbol{b}_{ij}^{\mathrm{T}}\boldsymbol{B}_{ij}^{-1}/b_{ij4} \end{bmatrix}\tilde{\boldsymbol{u}}_{\boldsymbol{j}}.$$

Since this equality holds for all image points on plane Π_{B}, it further implies:

$$\begin{bmatrix} \boldsymbol{I} \\ -\boldsymbol{b}_{jk}^{\mathrm{T}}/b_{jk4} \end{bmatrix}\boldsymbol{B}_{ij} = \lambda_{\mathrm{B}}\boldsymbol{H}_{ijk}\begin{bmatrix} \boldsymbol{I} \\ -\boldsymbol{b}_{ij}^{\mathrm{T}}/b_{ij4} \end{bmatrix}.$$

This leads to a system of 12 linear equations in 16 unknowns: 15 from the matrix $\boldsymbol{H}_{ijk}$ plus one from λ_{B}. Therefore, we need another system of equations provided by a second visible plane, say, Π_{G}:

$$\begin{bmatrix} \boldsymbol{I} \\ -\boldsymbol{g}_{jk}^{\mathrm{T}}/g_{jk4} \end{bmatrix}\boldsymbol{G}_{ij} = \lambda_{\mathrm{G}}\boldsymbol{H}_{ijk}\begin{bmatrix} \boldsymbol{I} \\ -\boldsymbol{g}_{ij}^{\mathrm{T}}/g_{ij4} \end{bmatrix}.$$

Combining the above two systems of equations, we have a total of 24 linear equations in 17 unknowns. Here we give a least squared solution by placing the two systems of equations in the following form

$$\boldsymbol{H}_{ijk}\left[\begin{bmatrix} \boldsymbol{I} \\ -\vec{\boldsymbol{b}}_{ij}^{\mathrm{T}}/b_{ij4} \end{bmatrix}\middle|\middle|\begin{bmatrix} \boldsymbol{I} \\ -\vec{\boldsymbol{g}}_{ij}^{\mathrm{T}}/g_{ij4} \end{bmatrix}\right] \cong \left[\begin{bmatrix} \boldsymbol{I} \\ -\vec{\boldsymbol{b}}_{jk}^{\mathrm{T}}/b_{jk4} \end{bmatrix}\boldsymbol{B}_{ij}\middle|\frac{\lambda_{\mathrm{B}}}{\lambda_{\mathrm{G}}}\begin{bmatrix} \boldsymbol{I} \\ -\vec{\boldsymbol{g}}_{jk}^{\mathrm{T}}/g_{jk4} \end{bmatrix}\boldsymbol{G}_{ij}\right] \tag{10}$$

where the ratio of $\lambda_{\mathrm{B}}/\lambda_{\mathrm{G}}$ has been estimated during the plane reconstruction phase (see Eq. (7)). We can find the matrix $\boldsymbol{H}_{ijk}$ using the pseudo-inverse matrix of the 4×6 matrix on the left-hand side of the above equation.

5. Computation of homographies

We need to estimate $\boldsymbol{A}_{ij}$ from the image data associated with the planar surface Π_{A}. We shall use the region-based matching, instead of point-based matching, to find the homography. First of all, we use the Gabor filtering technique (Sun et al., 2002) to identify at least four point correspondences in order to obtain the initial solution of the homography. We then use the Levenberg–Marquardt iterative nonlinear minimization algorithm (Szeliski, 1996) to minimize the sum of the squared intensity differences of the transformed and original image points due to the plane Π_{A} in the image pair

$$E=\sum_k\left\{\boldsymbol{I}_j((u_k^i)',(v_k^i)')-\boldsymbol{I}_i(u_k^i,v_k^i)\right\}^2.$$

Here the transformed location $\left[(u_k^i)',(v_k^i)',1\right]^{\mathrm{T}}$ is obtained from the image point (u_k^i,v_k^i) using an estimated $\boldsymbol{A}_{ij}$, and $\boldsymbol{I}_j\left((u_k^i)',(v_k^i)'\right)$ is the intensity obtained by a bilinear interpolation from the original image $\boldsymbol{I}_j$. The intensity values of the image points in the common region of the two images are normalized to remove the possible illumination

difference. The above minimization method converges in a few iterations.

After finding two homographies, recall that we can compute the epipole $\tilde{\boldsymbol{e}}_j$ from the skew symmetry property of $\boldsymbol{B}^{\mathrm{T}}[\tilde{\boldsymbol{e}}_j]_{\times}\boldsymbol{A}$. Also, in turn, we can compute the fundamental matrix F using the epipole $\tilde{\boldsymbol{e}}_j$ as follows:

$$[\tilde{\boldsymbol{e}}_j]_{\times}\boldsymbol{A} \cong [\tilde{\boldsymbol{e}}_j]_{\times}\left(\left\{\boldsymbol{M}_j\boldsymbol{R}_j\boldsymbol{M}_i^{-1} - \lambda_e\tilde{\boldsymbol{e}}_j\vec{\boldsymbol{a}}_E^{\mathrm{T}}\boldsymbol{M}_i^{-1}\right\}\right) \cong [\tilde{\boldsymbol{e}}_j]_{\times}\boldsymbol{M}_j\boldsymbol{R}_j\boldsymbol{M}_i^{-1} = \boldsymbol{F}.$$

6. Experimental results

6.1. Experiment 1

In the first experiment we use a synthetic tower whose feature points and schematic diagram are given in Table 1 and Fig. 2. We take a sequence of six pictures to cover all aspects of the tower using a virtual camera looking down from the upper positions. The image resolution is 640×480 in pixel. Three consecutive images of the sequence, I_1, I_2, and I_3, are shown in Fig. 3. We apply the reconstruction process to this data set. We employ a linear least-squares method based on eight corresponding image point pairs available in the synthetic data to get the true homography for each of the five planes, Π_{Gr}, Π_{A}, Π_{B}, Π_{E}, Π_{F} visible in the pair (I_1, I_2). In addition, to handle the possible problems caused by data translation and scaling change, we also use the normalization transform proposed by Hartley (1997) to compute the homographies. We choose Π_{Gr} as the reference plane. During the reconstruction process, we find the plane coefficient vectors with respect to the reference plane Π_{Gr} vector designated as $[1 \quad 1 \quad 1 \quad 1]^{\mathrm{T}}$. To check the correctness of the final 3D projective reconstruction result, we convert the 3D camera centered projective space back to the 3D object centered Euclidean space using the 3D Euclidean data of the tower available in Table 1 to measure the reconstruction errors in the metric space. The computation times for estimating the plane coefficient vectors and the coordinate transformation matrix for space integration are within a second.

Table 1
3D object centered coordinates of the tower feature points

Point	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
X	−20	−20	−20	−20	−20	−20	−20	−20	−20	−20	−20	−20	−13.333	−13.333	−6.667	−6.667	0	0	−50	−70	−50	0	50	70
Y	0	0	120	120	50	50	70	70	80	80	110	110	140	140	160	160	180	0	0	0	0	0	0	0
Z	−20	20	20	−20	−10	10	10	−10	−15	15	15	−15	−13.333	13.333	−6.667	6.667	0	−70	−50	0	50	70	50	0

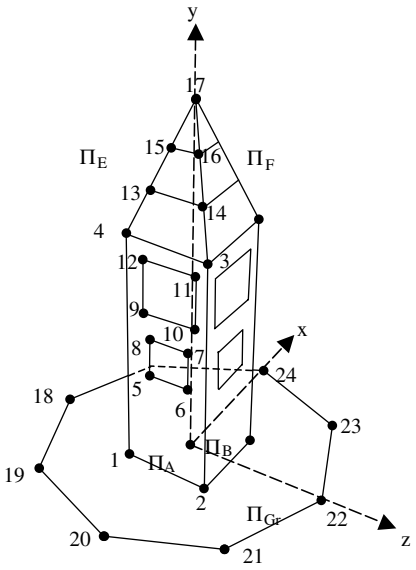

Fig. 2. The schematic diagram of the tower. The dimensions of the tower are 40 in. in depth (the x-direction), 40 in. in width (the z-direction) and 180 in. in height (the y-direction).

In what follows we assume the noise is uniformly distributed over the interval $[-R, R]$, where R indicates the noise strength or level. We generate 500 copies of noisy image data using the given noise model with $R = 0.5$, 1.0, 1.5 and 2.0 pixels, respectively. Then we find the 500 reconstruction results and compute the mean and standard deviation of the differences between the true and the estimated values of the 3D coordinates of the tower feature points. Table 2 lists the statistics of the relative distance errors into the x, y, and z components. The results indicate our reconstruction method is quite stable in the presence of the noise.

Since the rest of the planes visible in the images are estimated relative to the reference plane, we shall examine the effect of the assigned value of the reference plane coefficient vector on the reconstruction. Five hundreds of the reference plane coefficient vectors are uniformly generated from the range $[10^{-3}, 10^{2}]$; we also randomly select the

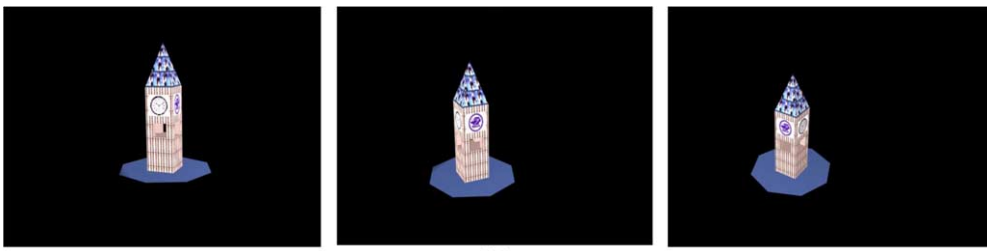

Fig. 3. Three distinct images I_1, I_2 and I_3 taken at a distance of about 500 in. The visible planes in the three images are Π_{Gr}, Π_A, Π_B, Π_E, Π_F in I_1 and I_2, and Π_{Gr}, Π_B, Π_F, Π_C, Π_G in I_3.

Table 2
The statistics of the distance errors of the reconstruction results

Error value			Noise level R (in pixels)				
			0	0.5	1	1.5	2
Error type	Distance mean error	x	6.624e−5	0.0547	0.1096	0.2108	0.2505
		y	2.492e−5	0.0896	0.1811	0.2944	0.3923
		z	7.341e−5	0.0761	0.1514	0.3379	0.3798
	Standard deviation	x	2.027e−4	0.0855	0.1707	0.3655	0.3850
		y	3.188e−5	0.1323	0.2673	0.4416	0.5745
		z	1.192e−4	0.1165	0.2320	0.6085	0.9093

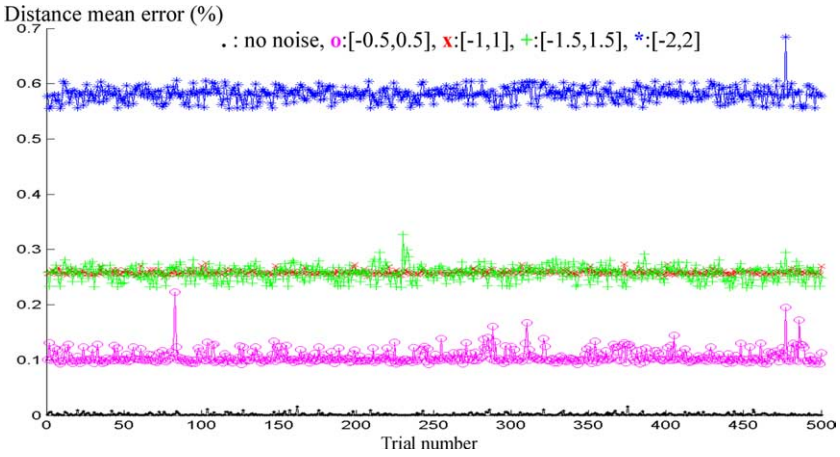

Fig. 4. The effects of the uniformly generated reference plane coefficient vectors and the noise at different levels on the reconstruction result.

positive or negative sign for the coefficients. Fig. 4 depicts the reconstruction results under the effects of the random selection of the reference plane coefficient vector and the noise at different levels. The horizontal axis indicates the trial number of the reconstruction process and the vertical axis indicates the resulting distance errors. The various marks "·", "o", "x", "+" and "∗" stand for mean errors of the computed relative distances associated with the uniform noise levels of $R = 0$, 0.5, 1, 1.5 and 2 pixels, respectively. The figure indicates the reconstruction results are virtually not affected by the random selection of the reference plane coefficient vector under the given specified noise. A remark is in order here. That is, we must avoid using $(1,1,1,0)^{\mathrm{T}}$ for the reference plane coefficient vector, since the camera origin $(0,0,0,1)$ is supposed not to lie on the plane.

6.2. Experiment 2

For a comparison between our method and the point-based method employing the fundamental matrix estimated from two arbitrary planes with the aid of hallucinated points (Szeliski and Torr, 1998), we use the same setup as in the previous experiment and run the experiment 500 times with a uniform distribution at different noise levels. The reconstructed Euclidean position errors are computed and tabulated in Table 3. The position errors are in the unit of inch.

From this table, we observe that in the noiseless cases where the noise level is 0, all the reconstruction results obtained by the two methods are almost equally good and very small; the errors are due to the rounding/truncation errors arising from numerical computations. As the noise level

Table 3
The mean errors of the reconstructed Euclidean point positions for different setups

Method	Noise level (in pixel)			
	0	0.2	0.5	1.0
Ours	0.0000301259	0.2030772297	0.4996064197	1.2803144642
$(4,2,2)^{a}$	0.0000781737	1.6271384793	14.5382797879	47.6350445089
$(4,4,1)^{a}$	0.0000125658	0.2113535985	0.5662224535	1.4202651215
$(4,4,2)^{a}$	0.0000134805	0.2128984139	0.5343441024	1.4456471793

[a] (n,m,p): n, m are the respective numbers of points on the two planes, p is the number of points hallucinated per plane.

increases from 0 to 1.0 for all the (n, m, p) cases, we notice that the homography estimation using the four noisy data points varies dramatically, and, thus, the fundamental matrix computation with the resulting noisy data points is bad. These lead to the final reconstruction results with large errors. For a more interesting comparison, we compute the homography using the four outmost data points and then use the estimated homography to generate the hallucinated points located inside the area surrounded by the four outmost points; we denote these hallucinated points as the p ($p = 1$ or 2) points. In our method since the homographies are iteratively estimated in a region-based way, so our reconstruction results are good even in the presence of image noise. In these simulations, the reconstruction results of the two methods are nearly equally good. Even so, our method is better in the sense that we can efficiently find each 3D plane without the need of computing 3D points, while the authors in the other method alternated a plane estimation stage with the point reconstruction stage. Thus, their method conducted two kinds of estimations: plane and points.

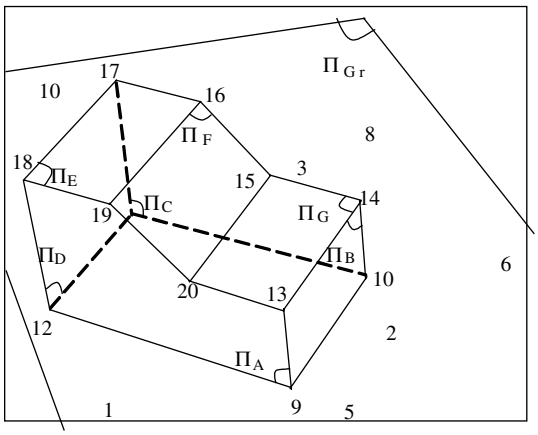

Fig. 5. The indices of the vertices and planes of the object.

6.3. Experiment 3

In this experiment the real images of a polyhedral, depicted in Fig. 5, are used to reconstruct the model of seven major planar surfaces. We go through the whole reconstruction process as we did in Experiment 1. The line parallelism and perpendicularity properties of the scene are used to

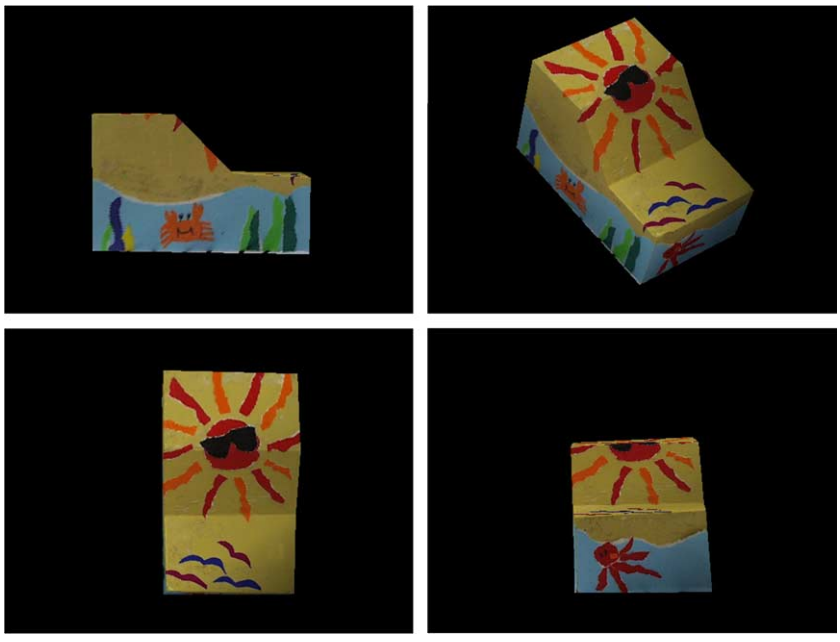

Fig. 6. New views of the reconstructed object with texture mapping.

Table 4
The estimated angles between the object planes and the ground plane

Planes form the angle	Angle	
	Estimated value	Actual value
(Π_{Gr}, Π_A)	92.21	90
(Π_{Gr}, Π_B)	92.69	90
(Π_{Gr}, Π_C)	97.09	90
(Π_{Gr}, Π_D)	91.67	90
(Π_{Gr}, Π_E)	1.69	0
(Π_{Gr}, Π_F)	45.31	45
(Π_{Gr}, Π_G)	1.66	0

compute the projective-to-Euclidean coordinate transformation matrix. First, the line parallelism is used to compute the plane at infinity which is then used to transform the reconstruction results from projective space to affine space. Secondly, the line perpendicularity is used to transform the reconstruction results from affine space to metric space. Further details can be found in (Daniilidis and Ernst, 1996; Zhang et al., 1998). Fig. 6 shows the four new views of the reconstructed model, which look like the real ones. Besides, the metric angles between individual object plane and the ground plane are shown in Table 4. The reconstructed object is found to be rather close to the true one.

7. Conclusions

An uncalibrated planar object reconstruction method has been described in which we rely on the plane information. We first estimate the homography for all planar surfaces using the region features of planar surfaces, and then we use the homographies induced by two planes to compute the epipole. We represent explicitly the compatible projection equations for the stereo images using the information of planar homographies and an assigned reference plane coefficient vector. We continue to derive the 3D equations of the planes visible in the stereo images with respect to the assigned reference plane once the planar homography is determined. Finally, to integrate the reconstruction results obtained from different image pairs under a unified projective space, we use the homography and the plane coefficient vector information of two planes to derive the coordinate transformation matrix. We then compute the new plane equation for the planes in the unified projective space. In the experiments we conduct the sensitivity analysis on our method by introducing image noise. We also consider the effect of assigning the different values of the reference plane coefficient vector on the reconstruction results. Experimental results on the synthetic and real images indicate the reconstruction method works quite successfully. In the future, we shall consider combining this plane based reconstruction method with other methods to determine the 3D structure of more complex objects.

References

Beardsley, P.A., Zisserman, A., Murray, D., 1997. Sequential updating of projective and affine structure from motion. Int. J. of Comput. Vis. 23 (3), 235–259.

Criminisi, R., Zisserman, A., 1998. Duality, rigidity and planar parallax. In: Proceedings of European Conference on Computer Vision. vol. 2. pp. 846–861.

Daniilidis, K., Ernst, J., 1996. Active intrinsic calibration using vanishing points. Pattern Recogn. Lett. 17, 1179–1189.

Faugeras, O.D., 1992. What can be seen in three dimensions with an uncalibrated stereo rig? In: Proceedings of European Conference on Computer Vision. pp. 563–578.

Faugeras, O.D., 1993. Three-Dimensional Computer Vision: A Geometric Viewpoint. MIT Press, Cambridge, MA.

Fitzgibbon, A.W., Zisserman, A., 1998. Automatic camera recovery for closed or open image sequences. In: Proceedings of European Conference on Computer Vision. pp. 311–326.

Fradkin, M., Roux, M., Maitre, H., Lelogu, U.M., 1999. Surface reconstruction from multiple aerial images in dense urban areas. In: Proceedings of International Conference on Computer Vision. pp. 262–267.

Hartley, R.I., Gupta, R., Chang, T., 1992. Stereo from uncalibrated cameras. In: Proceedings of International Conference on Computer Vision and Pattern Recognition. pp. 761–764.

Hartley, R.I., 1994. Projective reconstruction and invariants from multiple images. IEEE Trans. Pattern Anal. Mach. Intell. 16 (10), 1036–1041.

Hartley, R.I., 1997. In defense of the eight-point algorithm. IEEE Trans. Pattern Anal. Mach. Intell. 19 (6), 580–593.

Hartley, R.I., Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge, UK. Chapter 9.

Hamid, N.H.G., Cipolla, R., 1997. Identifying planar regions in a scene using uncalibrated stereo vision. In: Proceedings of British Machine Vision Conference.

Irani, M., Rousso, B., Peleg, S., 1997. Recovery of ego-motion using region alignment. IEEE Trans. Pattern Anal. Mach. Intell. 19 (3), 268–272.

Irani, M., Anandan, P., Weinshall, D., 1998. From reference frames to reference planes: multi-view parallax geometry and applications. In: Proceedings of European Conference on Computer Vision, vol. 2. pp. 829–845.

Johansson, B., 1999. View synthesis and 3D reconstruction of piecewise planar scenes using intersection lines between the planes. Proceedings of European Conference on Computer Vision. pp. 54–59.

Luong, Q.T., Vieville, T., 1996. Canonic representations for the geometries of multiple projective views. Comput. Vis. Image Und. 64 (2), 193–229.

Luong, Q.-T., Faugeras, O.D.,1993. Determining the fundamental matrix with planes: instability and new algorithms. In: Proceedings of Conference on Computer Vision and Pattern Recognition. pp. 489–494.

Mohr, R., Arbogast, E., 1991. It can be done without camera calibration. Pattern Recogn. Lett. 12 (1), 39–43.

Rother, C., Carlsson, S., Tell, D., 2002. Projective factorization of planes and cameras in multiple views. In: Proceedings of International Conference on Pattern Recognition. pp. 737–740.

Rother, C., 2003. Multi-View Reconstruction and Camera Recovery using a Real or Virtual Reference Plane. Doctoral Dissertation. ISBN 91-7283-422-6, KTH, Stockholm, Sweden.

Szeliski, R., 1996. Video mosaics for virtual environments. IEEE Comput. Graph. Appl. 6 (March), 22–30.

Szeliski, R., Torr, P.H.S., 1998. Geometrically constrained structure from motion: Points on planes. In: Proceedings of European Workshop on 3D Structure from Multiple Images of Large-Scale Environments. pp. 171–186.

Sinclair, D., Blake, A., 1996. Quantitative planar region detection. Int. J. Comput. Vis. 18 (1), 77–91.

Sawhney, H.S., 1994. 3D geometry from planar parallax. In: Proceedings of Conference on Computer Vision and Pattern Recognition. pp. 929–934.

Sun, S.K., Chen, Z., Chia, T.L., 2002. Invariant feature extraction and object shape matching using Gabor filtering. In: Proceedings of International Conference on Visual Information System, Hsinchu, Taiwan. pp. 95–104.

Tsai, R.Y., Huang, T.S., 1982. Estimating three-dimensional motion parameters of a rigid planar patch, II: singular value decomposition. IEEE Transactions on Acoustic, Speech and Signal Processing, vol. 30.

Theiler, L., Chabbi, H., 1999. Projective decomposition of planar facets. In: Proceedings of International Conference on Image Analysis and Processing. pp. 556–561.

Zelnik-Manor, L., Irani, M., 2000. Multi-frame estimate of planar motion. IEEE Trans. Pattern Anal. Mach. Intell. 22 (10), 1105–1116.

Zhang, Z., Isono, K., Akamatsu, S., 1998. Euclidean structure from uncalibrated images using fuzzy domain knowledge: Application to facial images synthesis. In: Proceedings of International Conference on Computer Vision. pp. 784–789.