A Comparative Study on Two Reconstruction Methods

HUANG Feng-rong, ZHAO Xun-po, HU Zhan-yi
(National Laboratory of Pattern RecognitionInstitute of Automation, Chinese Academy of Sciences, Beijing 100080)

Abstract The commonly used reconstruction method from images acquired by a calibrated camera but with unknown motion parameters is the essential matrix based one in the literature, i. e. , from the essential matrix to decompose the rotation matrix and translation vector at first, then to reconstruct the scene by the standard stereo method. A less well-known and conceptually rather more abstract method is the so-called absolute dual quadric based method. However, based on extensive experiments on simulated data as well as on real images, we showed that the absolute dual quadric based method consistently outperforms the essential matrix based one in terms of reconstruction accuracy and robustness, hence it is highly recommended to use the absolute dual quadric based method rather than the commonly used essential matrix based one in practice.

Keywords projective reconstruction, metric reconstruction, essential matrix, the absolute dual quadric

中图法分类号: TP391　**文献标识码:** A　**文章编号:** 1006-8961(2005)10-1320-09

两种 3 维重建方法的比较

黄凤荣　赵训坡　胡占义
(中国科学院自动化研究所模式识别国家重点实验室, 北京　100080)

摘　要　对基于本质矩阵和基于绝对二次曲面的两种重建方法进行了比较研究,仿真实验和真实图像实验结果表明,尽管基于绝对二次曲面的重建方法比较抽象,很多人对这种方法不太熟悉,但是该方法无论是在重建精度还是鲁棒性方面都优于基于本质矩阵的重建方法,这一发现对实际应用有非常重要的指导作用。

关键词　摄影重建　度量重建　本质矩阵　绝对二次曲面

1　Introduction

3D reconstruction which means recovering 3D shape of object from images has attracted much of the attention of the computer vision community over the last decade, and a lot of progress has been made in developing practical techniques from images, both with calibrated and uncalibrated cameras[1~3]. It is well known that if the camera is uncalibrated, it is only possible to reconstruct the scene up to an unknown projective transformation, called projective reconstruction. However, once the camera is calibrated, a projective reconstruction can be upgraded to an unknown similarity transformation, called metric reconstruction (i. e. Euclidean up to unknown scale). In practical applications, obtaining a metric reconstruction is always our desired goal, therefore we should either calibrate the camera beforehand or calibrate the camera with self-calibration methods in order to obtain a metric reconstruction. With calibrated 2D images at hands, a problem comes: How to

基金项目:国家自然科学基金项目(60303021,60375006)
收稿日期:2004-11-05;**改回日期:**2005-03-30
第一作者简介:黄凤荣(1969～　),女,博士。研究方向为模式识别、3 维重建与图像测量。E-mail: hfr6869@sina.com.cn

reconstruct accurate 3D model from these calibrated 2D images via triangulation? The scenarios commonly occur as: reconstruction from image sequences taken by a calibrated hand-held camera, or from images after camera self-calibration. The most familiar way to proceed for this problem in the literature is the so-called essential matrix based one, i. e. by decomposing the essential matrix, the appropriate rotation matrix and translation vector is obtained at first, then the reconstruction can be done by the standard stereo vision method. Another less well known reconstruction method is the absolute dual quadric based one, whose objective is firstly to determine the plane at infinity, then by which the projective reconstruction is upgraded to a metric one in combination with the camera intrinsic matrix. The goal of this work is to give a comparative study of these two methods. Surprisingly enough, based on our extensive experiments, we find the absolute dual quadric based method consistently outperforms the essential matrix based one in terms of both accuracy and robustness. This result seems original, and does not appear in other places.

2 Two reconstruction methods

If the camera is calibrated and image point correspondences are established between images, we can obtain a metric reconstruction by anyone of the two methods, i. e. the essential matrix based method and the absolute dual quadric based method.

2.1 The essential matrix based method

Given a pair of calibrated images, the metric reconstruction is equal to determining the correct pair of camera projection matrices, which can be done by the following steps.

(1) Determining the essential matrix

Because the camera is assumed calibrated and image point correspondences established, the essential matrix (see chapter 8 of [4] for detail information) may be computed either directly using normalized image points[4], or from the fundamental matrix $\boldsymbol{F}$ (see chapter 8 of [4] for detail information) as $\boldsymbol{E} = \boldsymbol{K}_2^{\mathrm{T}}\boldsymbol{F}\boldsymbol{K}_1$[4]. There is an abundant literature on the fundamental matrix determination, see chapter 10 of [4] for example.

(2) Extracting camera projection matrices from the essential matrix

We know that if $\boldsymbol{E} = \boldsymbol{U}diag(\delta,\delta,0)\boldsymbol{V}^{\mathrm{T}}$ is the SVD decomposition of the essential matrix[4], and by setting the first camera projection matrix as $\boldsymbol{P}_1 = \boldsymbol{K}_1[\boldsymbol{I}\quad \boldsymbol{0}]$, there exists a four-fold ambiguity for the second camera projection matrix $\boldsymbol{P}_2$ as[4]:

$$\boldsymbol{P}_2 = \boldsymbol{K}_2[\boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^{\mathrm{T}}\quad \delta\boldsymbol{u}_3]$$
$$\boldsymbol{P}_2 = \boldsymbol{K}_2[\boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^{\mathrm{T}}\quad -\delta\boldsymbol{u}_3]$$
$$\boldsymbol{P}_2 = \boldsymbol{K}_2[\boldsymbol{U}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}\quad \delta\boldsymbol{u}_3]$$
$$\boldsymbol{P}_2 = \boldsymbol{K}_2[\boldsymbol{U}\boldsymbol{W}^{\mathrm{T}}\boldsymbol{V}^{\mathrm{T}}\quad -\delta\boldsymbol{u}_3]$$

where $\boldsymbol{u}_3$ is the last column of $\boldsymbol{U}$, and

$$\boldsymbol{W} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since reconstructed points must all be in front of both cameras, the unique correct pair of camera projection matrices can be selected among the above four possible ones[4].

2.2 The absolute dual quadric based method

It is well known that if image point correspondences are known between two images, a projective reconstruction of the scene can be obtained such as[4]:

$$\begin{aligned} \boldsymbol{P}_1 &= [\boldsymbol{I}\quad \boldsymbol{0}] \\ \boldsymbol{P}_2 &= [\boldsymbol{e}\times\boldsymbol{F}\quad \boldsymbol{e}] \end{aligned} \qquad (1)$$

where $\boldsymbol{F}$ is the fundamental matrix of these two images, and $\boldsymbol{e}$ is the epipole in the second image, both entities can be computed from only image correspondences.

Given a projective reconstruction, if the first camera projection matrix is of the form in (1), then as shown in[3], the projective reconstruction can be upgraded to a metric reconstruction by an appropriate 4×4 matrix $\boldsymbol{H}$ as:

$$\boldsymbol{P}_{1e} = \boldsymbol{P}_1\boldsymbol{H}$$
$$\boldsymbol{P}_{2e} = \boldsymbol{P}_2\boldsymbol{H}$$

and $\boldsymbol{H}$ has the following form:

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{K}_1 & \boldsymbol{0} \\ -\boldsymbol{a}^{\mathrm{T}}\boldsymbol{K}_1 & 1 \end{bmatrix} \qquad (2)$$

where K_1 is the intrinsic matrix of the first camera, $(a^T \quad 1)^T$ is the plane at infinity in the projective frame. In the other words, the problem of metric reconstruction becomes how to determine the plane at infinity $(a^T \quad 1)^T$ in the projective frame.

(1) Determining the plane at infinity

As shown in [3], 3-vector a can be determined from the following constraints, which are generated by the absolute dual quadric:

$$sK_2K_2^T = P_2\begin{bmatrix} K_1K_1^T & -K_1K_1^Ta \\ -a^TK_1K_1^T & a^TK_1K_1^Ta \end{bmatrix}P_2^T \qquad (3)$$

where s is an unknown scale.

Since K_1, K_2, P_2 are known, by some simple manipulation, 3 constraints on 3-vector a can be derived. Among these three constraints, two are linear and one is quadratic. Hence a two-fold ambiguity could at most exist for the plane at infinity a. In other words, two possible pairs of the camera projection matrices could be obtained. Similarly as those in the essential matrix based method, this ambiguity can also be removed by reconstructing at least one space point, the pair by which the reconstructed space point is in front of both cameras should be the correct one.

(2) Algorithm outline

The absolute dual quadric based method can be outlined as:

① Obtain a projective reconstruction $P_1 = [I \quad 0]$, P_2 from image correspondences;

② Determine the plane at infinity 3-vector a from the constraints in equation (3). At most a two-fold ambiguity exists.

③ Take the visibility constraint, i.e., the reconstructed points must be in front of both the cameras, the correct pair of projection matrices can be obtained.

3 Experiments with simulated data

Reconstruction performance depends on both the camera configuration and the magnitude of image noise, hence the two reconstruction methods will be compared under typical camera configurations and under various noise level in this section.

3.1 Simulation setup

In our simulations, the space points, randomly distribute inside a sphere centered at $(0 \quad 0 \quad 500)$ of radius = 500, are used to generate synthetic image points. The first camera's setup is:

$$K_1 = \begin{bmatrix} 1\,200 & 0.1 & 512 \\ 0 & 960 & 384 \\ 0 & 0 & 1 \end{bmatrix}$$

the second camera's setup is:

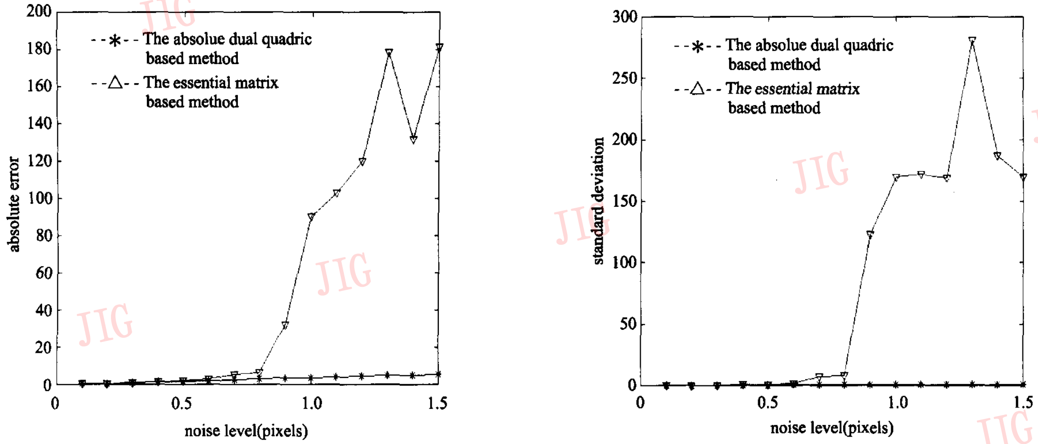
$$K_2 = \begin{bmatrix} 1\,300 & 0.2 & 512 \\ 0 & 1\,100 & 384 \\ 0 & 0 & 1 \end{bmatrix}$$

The image resolution is: 1 024 × 768 pixels. The fundamental matrix F is calculated using the normalized 8-point algorithm[4], the essential matrix is obtained by $E = K_2^TFK_1$. Furthermore, in order to ensure the comparability of the two methods, all tests are taken under same camera parameters and simulation data.

The reconstruction accuracy and robustness is measured by the average absolute errors of the reconstructed points and their standard deviation.

3.2 Noise influence

This experiment investigates the performance of two methods with respect to random noise. The parameters between two cameras are: rotation axis $r = [2 \quad 1 \quad 4]^T$, rotation angle $\vartheta = \pi/3$ and translation vector $t = [-10 \quad 15 \quad 250]^T$. In order to provide more statistically meaningful results, the Guassian noise with mean 0 and standard deviation ranging from 0 to 1.5 pixels is added to the image points. At each noise level, we randomly choose 200 space points and use their corresponding image points for the reconstruction by the two methods. The final result is the average of 100 independent trials. The average absolute error and their standard deviation under different noise level are shown in Figure 1. From Figure 1, we can see that the absolute dual quadric based method is much more robust and accurate than the essential matrix based one, especially when the noise level is larger than one pixel.



(a) Average absolute error　　(b) Standard deviation

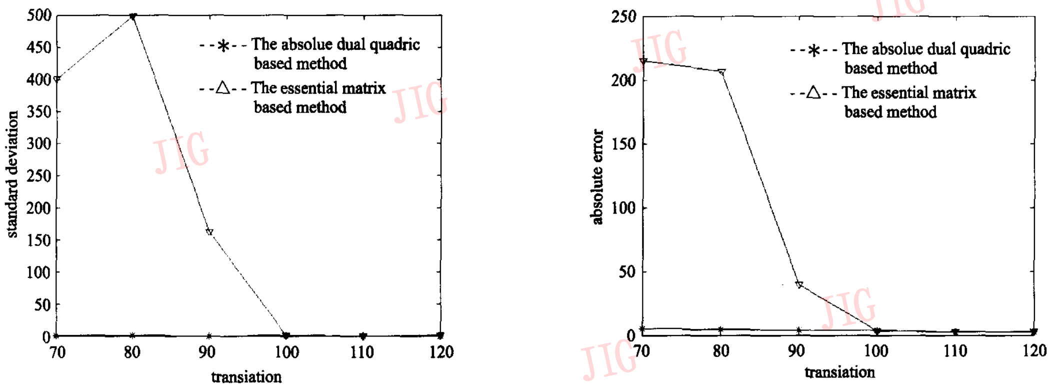
Fig. 1　Average absolute error and the standard deviation vs. the noise level

3.3 Influence of camera configurations

Lateral, axial, and circular are widely regarded as the three typical camera configurations, the two methods will be assessed with respect to these three configurations and under different noise level.

(1) Lateral configuration

By "lateral configuration", we mean the two cameras are related by only a pure lateral translation $t = [d \quad 0 \quad 0]^T$, i. e. under the standard stereo setup. As we know, the magnitude of baseline under the standard stereo configuration affects largely the reconstructed accuracy, parameter d is changed from 70 to 120 with the step of 10 in our simulation. At each step, we randomly choose 200 space points and use their corresponding image points perturbed by the Gaussian noise with mean 0 and standard deviation of 0.5 pixel for the reconstruction by the two methods. The results are shown in Figure 2, where the average absolute error and the standard deviation of 100 independent trials are reported. From Figure 2, we can see that the absolute dual quadric based method is more robust and accurate than the essential matrix based one, in especially for those cases where the baseline is relatively small.

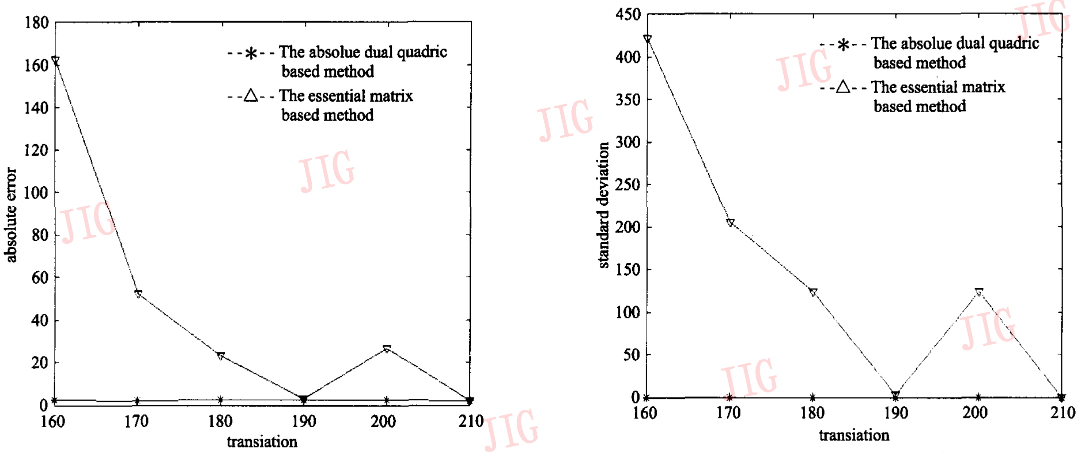


(a) Average absolute error　　(b) Standard deviation

Fig. 2　Absolute error and the standard deviation vs. the lateral translation

(2) Axial configuration

Simulation conditions are similar to the lateral configuration except that in this case, the translation is along the optical axis, i. e., $t = [0 \quad 0 \quad d]^T$ and parameter d is changed from 160 to 210 with the step of 10. The results are shown in Figure 3. From these results, it is evident that the absolute dual quadric based method consistently outperforms the essential matrix based one.

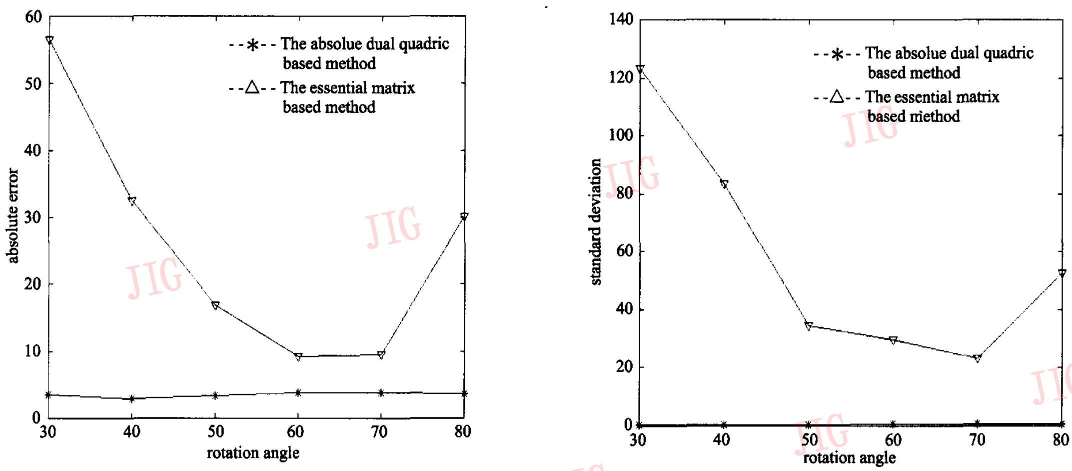


(a) Average absolute error (b) Standard deviation

Fig. 3 Average absolute error and the standard deviation vs. the axial translation

(3) Circular configuration

By "circular configuration", we mean the camera optical centers lie on a circle, and their optical axes are oriented to the circle's center. In our simulations, the radius of circle is set to 50. As in the previous two cases, 200 space points, Gaussian noise of 0.5 pixel, and 100 independent trials are used for the simulations. The rotation angle is changed from 30° to 80° with the step of 10°. The results are shown in Figure 4. Once again, we can see that the absolute dual quadric based method is better than the essential matrix based one.



(a) Average absolute error (b) Standard deviation

Fig. 4 Average absolute error and the standard deviation vs. the rotation angle

Remarks

(1) From Figure 2 and Figure 3, we can see that the absolute dual quadric based method largely outperforms the essential based one. We thought such a superiority could come from the fact that the absolute dual quadric based method is essential to determine the

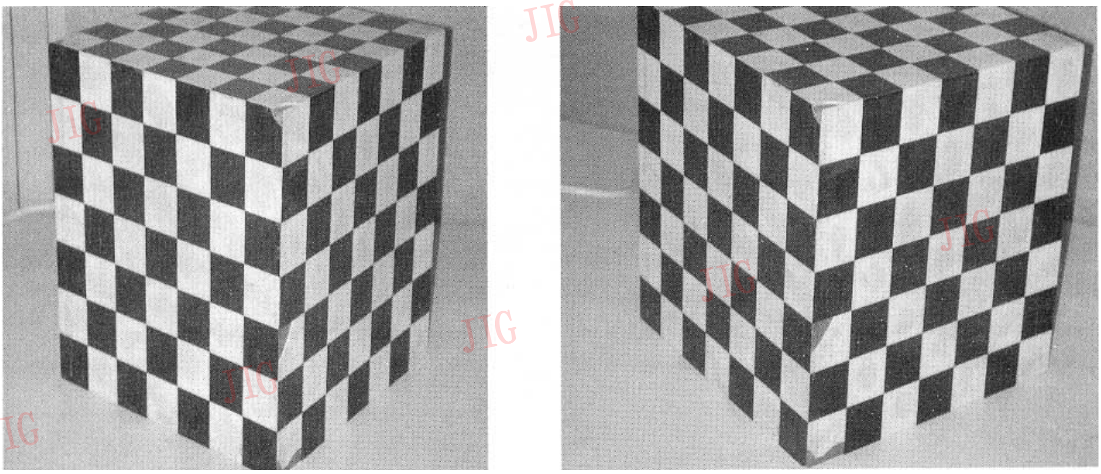
infinite homography, and this homography does not change under camera translations.

(2) In the essential matrix based method, a highly non-linear intermediate step, i. e. to compute rotation matrix and translation vector from the essential matrix, is involved, as a result, errors could be accumulated and propagated, which could be another reason of inferiority of this method.

4 Experiment with real images

In our real image tests, the images are obtained by an NIKON-COOLPIX990 CCD digital camera with resolution of 1 024 × 768. For each test, the camera is calibrated using the DLT method beforehand[5], and image point correspondences are established by our own software NLPR-CORRESP.

Test 1: We take two images of a calibration rig, shown in Figure 5, and reconstruct it using the two methods. The reconstruction result by the essential matrix based method is shown in Figure 6, and that by the absolute dual quadric based one is shown in Figure 7. The reconstructed angle of two perpendicular planes is shown in Table 1, the ratio of the maximum and the minimal reconstructed distance of grid is shown in Table 2.



(a) image1　　(b) image2

Fig. 5　Two test images of the calibration rig

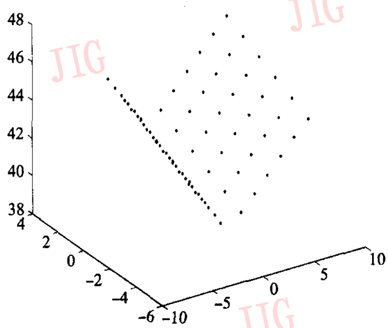


Fig. 6　The essential matrix based method

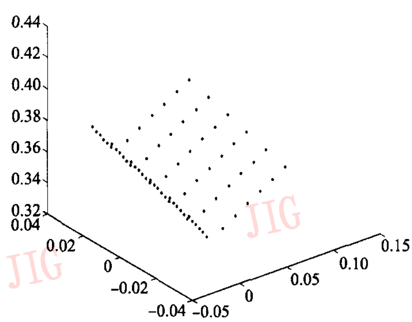


Fig. 7　The absolute dual quadric based method

From Figure 6 and Figure 7, we can see that the reconstructed points on each side of the calibration rig are indeed coplanar for both methods. However, from Table 1, we can see that the reconstructed angle of two

Tab. 1 The reconstructed angle of two perpendicular planes

The absolute dual quadric based method	The essential matrix based method
90.618 3°	91.657 4°

Tab. 2 The ratio of the maximum and the minimal reconstructed distance of grid

The absolute dual quadric based method	The essential matrix based method
1.083 9	1.107 0

perpendicular planes by the absolute dual quadric based method is 90.618 3°, closer to its ground truth of than that by the essential matrix based method, 91.657 4°. Furthermore, from Table 2, it is shown that the ratio of the maximum and the minimal reconstructed distance of the calibration grid by the absolute dual quadric based method is 1.083 9, much closer to its true value 1.0 than that by the essential matrix based one, 1.107 0. These show that for this pair of text images, the reconstruction results by the absolute dual quadric based method is more accurate than that by the essential matrix based one.

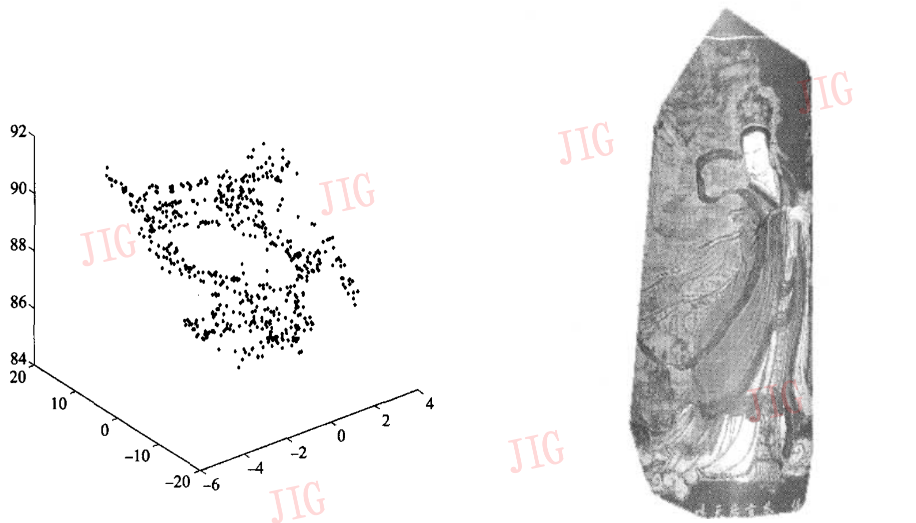
Test 2: We take two images of a tea caddy, shown in Figure 8, and reconstruct it using the two methods. The reconstruction results are shown in Figure 9 and 10 respectively.

Test 3: We take two images of a candy box, shown in Figure 11, and reconstruct it using the two methods. The reconstruction results are shown in Figure 12 and 13 respectively.

For Test 2 and Test 3, reconstructed results are similar, and both are satisfying.

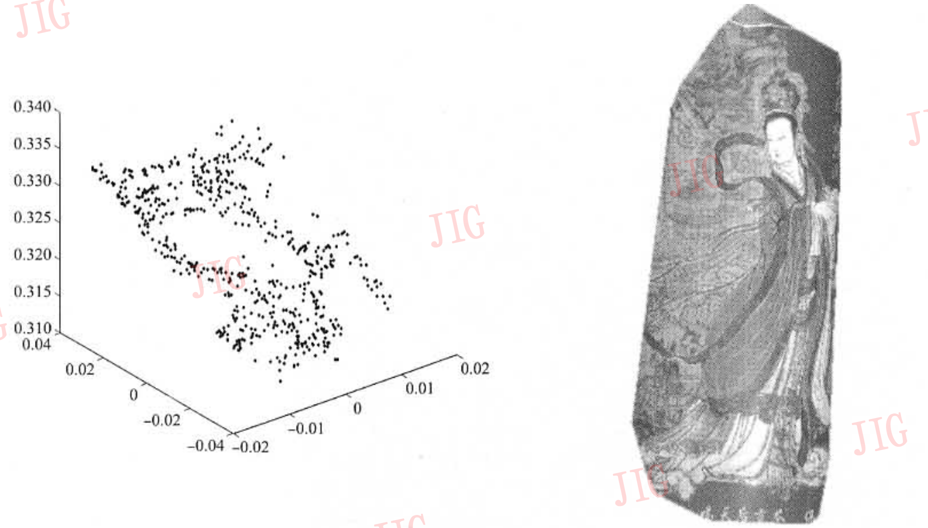
(a) image 1 (b) image 2

Fig. 8 Two test images of a tea caddy



(a) Reconstructed 3D points (b) Reconstruction results after texture matching

Fig. 9 The essential matrix based method

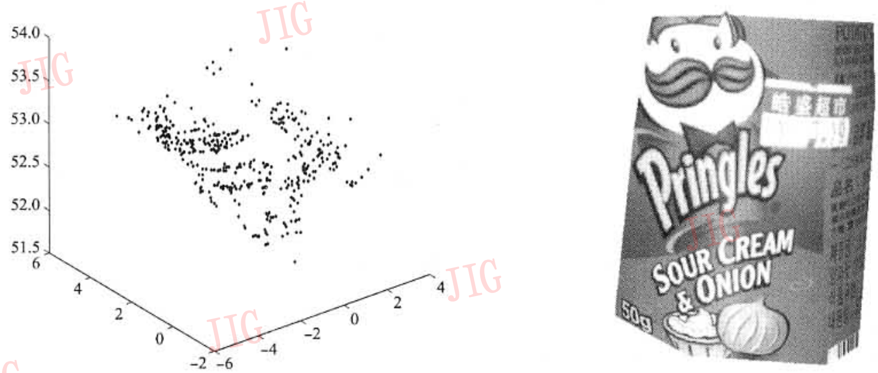


(a) Reconstructed 3D points　　(b) Reconstruction results after texture matching

Fig. 10　The absolute dual quadric based method

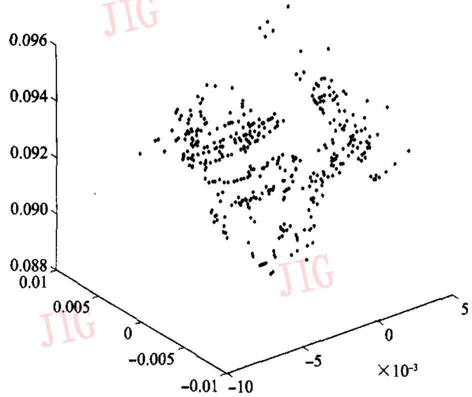
(a) image 1　　(b) image 2

Fig. 11　Two test images of a candy box



(a) Reconstructed 3D points　　(b) Reconstruction results after texture matching

Fig. 12　The essential matrix based method



(a) Reconstructed 3D points

(b) Reconstruction results after texture matching

Fig. 13 The absolute dual quadric based method

5 Conclusions

Extensive simulations and experiments with real images showed that the absolute dual quadric based method consistently outperforms the essential matrix based one in terms of reconstruction accuracy and robustness, hence it is highly recommendable in practice although it is conceptually rather abstract.

References

1 Faugeras O D. What can be seen in three dimensions with an unclibrated stereo rig? [A]. In: Proceedings of the 1992 European Conference on Computer Vision[C], Santa Margherita Liguere Italy, 1992: 563 ~ 578.

2 Hartley R I, Gupta R. Stereo from uncalibrated cameras[A]. In: Proceedings of the 1992 Computer Vision and Pattern Recognition [C], Urbana, Illinois, America, 1992: 579 ~ 587.

3 Pollefeys M, Koch R, VanGool L. Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters[A]. In: Proceeding of the 1998 International Conference on Computer Vision[C], Bombay, India, 1998:90 ~ 95.

4 Hartley R, Zisserman A. Multiple view geometry in computer vision [M]. London: Cambridge University Press, 2000.

5 Abdel-Aziz Y I, Karara H M. Direct linear transformation from comparator coordinates into object space coordinates[A]. In: ASP Symposium on Close-Range Photogrammetry[C], Urbana, Illinois, America, 1971:1 ~ 18.