Li Y, Wang YT, Liu Y. Fiducial marker based on projective invariant for augmented reality. JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY 22(6): 890~897 Nov. 2007

Fiducial Marker Based on Projective Invariant for Augmented Reality

Yu Li (李　玉), Yong-Tian Wang (王涌天), and Yue Liu (刘　越)

School of Information Science and Technology, Beijing Institute of Technology, Beijing 100081, China

E-mail: l_y_o@bit.edu.cn

Received October 26, 2006; revised July 10, 2007.

Abstract Fiducial marker based Augmented Reality has many applications. So far the inner pattern of the fiducial marker is always used to encode the markers. Thus a large portion of the fiducial marker image is used for encoding instead of providing corresponding feature points for pose accuracy. This paper presents a novel method which utilizes directly the projective invariant contained in the positional relation of the corresponding feature points to encode the marker. The proposed method does not require the region of pattern image for encoding any more and can provide more corresponding feature points so that higher pose accuracy can be achieved easily. Many related approaches such as cumulative distribution function, reprojection verification and robust process are proposed to overcome the problem of sensibility of the projective invariant. Experimental results show that the proposed fiducial marker system is reliable and robust, and can provide higher pose accuracy than that achieved by existing fiducial marker systems.

Keywords Augmented reality, fiducial marker, pose estimation, projective invariant, registration

1 Introduction

Augmented Reality (AR) combines the computer generated graphics and real environmental image to enhance the real scene[1,2]. AR is often used in medical care to overlap the MRI or CT images to the corresponding part of patients' body to assist diagnosis and operation. AR can also be used for the training of soldiers. In industry, three-dimensional "virtual" objects are embedded into workers' environment to guide the assembling and servicing of complex machines.

Registration method based on fiducial marker is the most mature solution for Augmented Reality application systems. Marker based Augmented Reality has various advantages: it is easy to use, highly robust, accurate, efficient and requires no initialization. There have been many well organized reusable programs which make the development of marker based Augmented Reality much easier and therefore marker based Augmented Reality has a large number of important applications in many areas.

1.1 Related Work

Circular markers are employed in some fiducial marker systems such as nested colored rings[3], 2D bar-coded round fiducials[4], and the TRIP system[5,6]. However, these circular markers are always considered as too complicated for pose estimation and hard to encode[7].

The most popular shape of fiducial marker is square as in ARToolKit system[8] (Fig.1). ARToolKit has been widely used but lacks accuracy in both pose estimation and marker recognition[9]. Rekimoto has improved the identification method by changing the inner image pattern into square shaped 2D barcode[10,11] (Fig.2).

Fig.1. ARToolKit marker.

Fig.2. CyberCode.

Short Paper

Supported by the National Grand Fundamental Research 973 Program of China (Grant No. 2002CB312104).

Owen, Xiao and Middlin conclude that the square shape is the best form for fiducial marker. In addition, the design of the identification method of template matching is also discussed[7].

Zhang, Navab and Liou have improved the accuracy of pose estimation in [12, 13]. In their SCR marker detection system, a form similar to the ARToolKit has been deployed (Fig.3). A comparison has been made, which indicates that SCR system has much better detection accuracy[9].

Instead of threshold and connected region detection method, an edge based approach is presented in ARTag fiducial marker system[14,15] (Fig.4).

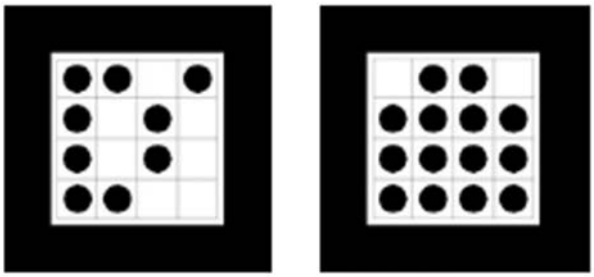

Fig.3. SCR marker.

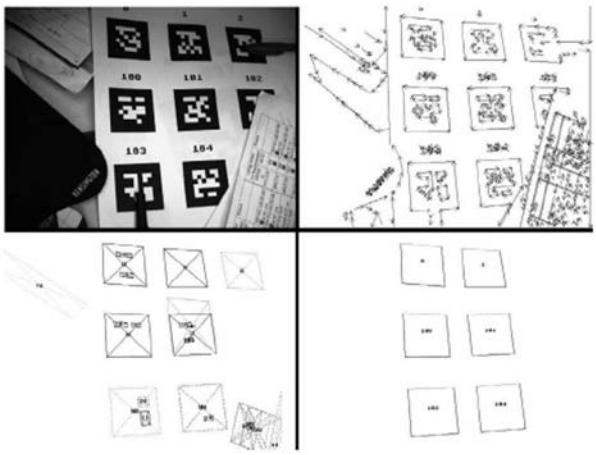

Fig.4. ARTag marker.

1.2 Weakness of Previous Approaches

The above-mentioned approaches could make Augmented Reality effective, accurate and robust, but they all require an image region for encoding or template matching inside the fiducial marker image to distinguish the marker and other square regions in the environment. The encoding image region always takes the biggest area of the fiducial marker but could not provide any correspondence feature to improve the accuracy of the pose estimation. The inner area must be large enough to be distinctly captured and cannot be occluded. Although the adoption of multiple markers can improve the accuracy and adapt to the situation of occlusion, a large area for multiple markers in the capture image is required.

If other kinds of information such as the related position of the correspondence feature points could be utilized to identify the markers, the inner encoding image region will not be required and the fiducial marker will become smaller and more flexible.

1.3 Fiducial Marker System Based on Projective Invariant (FPI)

A novel fiducial marker system based on projective invariant (FPI) is proposed in this paper (Fig.5), which is designed as a set of arbitrarily arranged triangles, whose vertices are used as the correspondence feature points to calculate the pose. In order to decrease the image area for the encoding or template matching, the markers are identified directly by calculating the projective invariants from the relative positions of these vertices. Actually, the projective invariant extracted from the point set is used to recognize the fiducial marker.

Fig.5. FPI marker.

To solve the problem that the projective invariant is very sensitive to errors[16~19], the cumulative distribution function and the reprojection verification are employed. In addition, a robust algorithm is presented to achieve the robust and reliable marker detection, which is demonstrated in the experiments.

1.4 Organization of the Paper

The remaining part of this paper is organized as follows. Section 2 discusses the characteristic of the projective invariant. Section 3 describes the design of FPI marker system and how projective invariant is used for marker identification. Section4 introduces the process of the robust marker recognition algorithm of the FPI marker system. In Section 5 the experimental result of the FPI system and the comparison with other marker systems are presented. Section 6 is the conclusion.

2 Projective Invariant

The invariants of simple feature sets (e.g., configurations of a few points or lines) are the frequently used tools in computer vision. Given a set of feature S and an operator g belonging to a transformation group, the invariant $I[\cdot]$ must satisfy $I[s] = I[gs]$ for any g in the group. If g is a projective transformation, then the projective invariants could be obtained. An extensive survey of the applications of projective invariants in computer vision can be found in [16, 19, 22∼24].

A configuration of five points A_i $(i = 1, \ldots, 5)$ is required in 2D to define the projective invariant which is known as 2D cross-ratios. The invariant cross-ratio is defined as:

$$\mu = \frac{(\Delta A_1 A_2 A_4)(\Delta A_1 A_3 A_5)}{(\Delta A_1 A_3 A_4)(\Delta A_1 A_2 A_5)}, \tag{1}$$

$$\nu = \frac{(\Delta A_1 A_2 A_4)(\Delta A_2 A_3 A_5)}{(\Delta A_2 A_3 A_4)(\Delta A_1 A_2 A_5)}, \tag{2}$$

where, $\Delta A_i A_j A_k = \begin{vmatrix} x_i & x_j & x_k \\ y_i & y_j & y_k \\ z_i & z_j & z_k \end{vmatrix}$.

The 2D cross-ratio is permutation-sensitive. There are 120 $(5! = 120)$ permutations for a configuration of five points, but only two invariants are independent; others can be expressed as the functions of these two invariants[22∼24]. Therefore, given two configurations of five points, if the two independent cross-ratios are equal separately, the two configurations are equal projectively, which is used in the FPI marker system to identify a configuration of five points.

Invariants p^2 (invariant to both projection and permutation) are derived from [23. 24], so that the problem of permutation could be ignored in the recognition of objects.

3 Fiducial Marker System Based on Projective Invariant

Fiducial marker image as shown in Fig.5 is used in FPI system. There is a set of triangles in arbitrary shapes with a white colored circle in every one in order to detect the order of vertices. Vertices of these triangles are extracted and used as feature points for pose estimation. Triangles in the captured image are corresponded to the reference points by examining the projective invariants. Changing the shape of triangles in the markers could change the corresponding projective invariants of the marker so that different marker set is produced. Different markers have different projective invariants and could be distinguished from each other according to their projective invariants. Therefore, multiple markers could be used in AR system.

3.1 Design

The vertices of triangles are designed as feature points since triangle is more proper than other polygon shapes. Firstly, concave polygon is hard to be detected and extracted. Secondly, it will be a kind of waste for the adoption of convex polygons which have more than 3 sides, since they will take up more inner area; at the same time, they lack the location accuracy because of the obtuse vertex angle. Therefore triangle would be the best choice in the FPI marker system.

As mentioned in Section 2, objects could be recognized when p^2 invariants is used even if the order of points is unknown. But for a fiducial maker after recognizing the object, the order of points is also needed for pose estimation. In addition, the p^2 invariant is very sensitive to the location error so the FPI system will not use p^2 invariants.

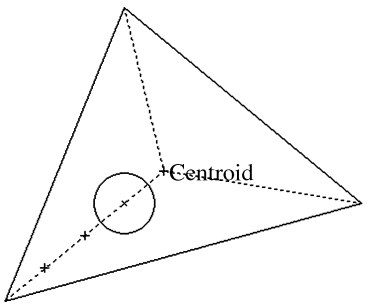

Fig.6. Design of the triangle fiducial.

Instead, a circle tag is put on every triangle (as shown in Fig.6), which is used to distinguish which vertex of a triangle is the first one so that the direction of the triangle is recognized. Then calculations of the two projective invariants could be performed without confusion. The center of the white circle resides in (x_0, y_0).

$$\begin{cases} x_0 = \dfrac{3x_c + x_1}{4}, \\ y_0 = \dfrac{3y_c + y_1}{4}, \end{cases} \tag{3}$$

where (x_i, y_i) is the i-th vertex of the triangle, $x_c = (x_1 + x_2 + x_3)/3$ and $y_c = (y_1 + y_2 + y_3)/3$ are the coordinates of the centroid of the triangle.

Affine translation will be used approximately here to find a way to detect which vertex is the nearest one to the circle tag. As the marker points will not differ too much in the dimension of depth, which makes the affine approximation easier to hold. Secondly, the circle tag needs not to be detected precisely. We only need to know which vertex of the triangle is the nearest one from the circle tag. Even though in some extreme cases the approximation turns too inaccurate, the reprojection verification step described in the later

section will filter error matches out. Experiments indicating that this approximation is reliable are given later.

Under affine translation approximation, we have: $\begin{bmatrix} x' \\ y' \\ 1' \end{bmatrix} = \boldsymbol{H}_A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$, where $\boldsymbol{H}_A = \begin{bmatrix} \boldsymbol{H}_{A1} \\ \boldsymbol{H}_{A2} \\ 0 \quad 0 \quad 1 \end{bmatrix}$ is the affine translation matrix, $\boldsymbol{H}_{A1}$ is the first row of $\boldsymbol{H}_A$, $\boldsymbol{H}_{A2}$ is the second row, and (x', y') is the affine translation of a point (x, y).

Intuitively, we have: $x' = \boldsymbol{H}_{A1}\boldsymbol{x}$, $y' = \boldsymbol{H}_{A2}\boldsymbol{y}$.

Multiplying $\boldsymbol{H}_{A1}$ and $\boldsymbol{H}_{A2}$ to both sides of (3), (4) could be obtained.

$$\begin{cases} x'_0 = \boldsymbol{H}_{A1}x_0 = \dfrac{3\boldsymbol{H}_{A1}x_c + \boldsymbol{H}_{A1}x_1}{4} = \dfrac{3x'_c + x'_1}{4} \\ y'_0 = \boldsymbol{H}_{A2}y_0 = \dfrac{3\boldsymbol{H}_{A2}y_c + \boldsymbol{H}_{A2}y_1}{4} = \dfrac{3y'_c + y'_1}{4} \end{cases} \tag{4}$$

To detect the direction of the triangle, 3 positions of the tags are calculated. Each position use one of the 3 corners of the triangle as the first one. Then the pixels around the 3 positions of the white tag (x'_0, y'_0) are examined. If white tag is discovered in one position, the first corner is recognized and the order of the triangle is determined as well.

In order to verify the effect of the above direction detection method, experiments are performed. A single triangle is printed out and captured by camera from several different angles from the optical axis in different distances. For every combination, experiments are taken for 20 times, correctness rate is recorded. As shown in Table 1, the method is feasible in our system.

Table 1. Correctness Rate of Direction Detection

Angle from Optical Axis	500mm Away (%)	750mm Away (%)	1000mm Away (%)	1250mm Away (%)
90°	100	100	100	100
75°	100	100	100	95
60°	100	100	95	90
45°	100	90	85	90

3.2 Triangle Detection

For an input image, first, a "Watershed adaptive threshold" (WAT) algorithm[25] is used to roughly separate the black regions, followed by a connected region detection by which all the marker regions could be obtained. Then triangles are extracted by triangle detection algorithm. Finally, the white tag in the specified position is examined to determine the order of three vertices of a triangle, and the triangles without a white tag inside are discarded.

3.3 Marker Identification by Projective Invariant

Firstly, for every pairs of two triangles in the marker, the projective invariants of the six vertices could be calculated and recorded to a reference table.

In the tracking phase, for all of the discovered triangles in the triangle detection step, the projective invariants of every combination of two triangles could be calculated and then the matching triangle pair in the reference table will be searched.

There are $6 \times 2 - 8 = 4$ independent cross-ratios for six 2D points[16]. When examining two triangles, as long as 2 configurations of five points are selected arbitrarily, then 4 independent projective invariants could be derived. The distance between these 4 projective invariants and the values in the reference table are used to compare the two triangles. If the distance is less than a specified threshold, they could be considered as the matched triangle pair.

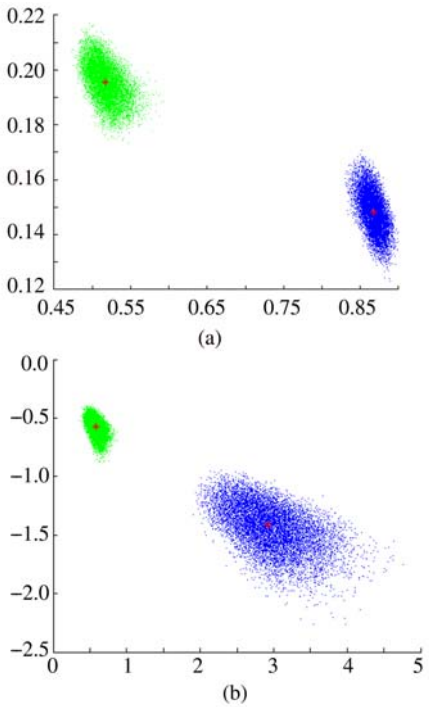

Fig.7. Error distribution comparison between cross-ratios and cumulative distributions function. (a) For cross-ratios. (b) For cumulative distributions function.

Note that the distance between two cross-ratios cannot be used directly because the cross-ratio is very sensitive to errors[16~19], the cross-ratios of two configurations may change greatly with their situated intervals. If located in different intervals, it is hard to distinguish a configuration by a fixed threshold of cross-ratio. For example as shown in Fig.7 (a), 5% random

noise has been added to the location of the 5 points, and 10 000 samples have been made. The cross-ratios have a large range of distribution around the real values.

The probability density function of the cross ratio X has been explained in [26, 27].

$$f_X(x) = \begin{cases} f_1(x) + f_3(x), & \text{if } x < 0, \\ f_3(x) + f_2(x), & \text{if } 0 \leqslant x \leqslant 1, \\ f_2(x) + f_1(x), & \text{if } x > 1, \end{cases} \quad (5)$$

where

$$f_1(x) = \frac{1}{3}\Big((2x-1)\ln\Big(\frac{x}{x-1}\Big) - 2\Big),$$
$$f_2(x) = \frac{1}{3}\frac{(x+1)\ln x + 2(1-x)}{(x-1)^3},$$
$$f_3(x) = \frac{1}{3}\frac{(x-2)\ln(1-x) - 2x}{(x-1)^3}.$$

Subsequently the cumulative distribution function could be expressed by (6).

$$F_X(x) = P(X < x),$$
$$= \begin{cases} F_1(x) + F_3(x), & \text{if } x < 0, \\ \dfrac{1}{3}, & \text{if } x = 0, \\ \dfrac{1}{2} + F_2(x) + F_3(x), & \text{if } 0 < x < 1, \\ \dfrac{2}{3}, & \text{if } x = 1, \\ 1 + F_1(x) + F_2(x) & \text{if } 1 < x. \end{cases} \quad (6)$$

where

$$F_1(x) = \frac{1}{3}\Big(x(1-x)\ln\Big(\frac{x-1}{x}\Big) - x + \frac{1}{2}\Big),$$
$$F_2(x) = \frac{1}{3}\Big(\frac{x - x\ln x - 1}{(x-1)^2}\Big),$$
$$F_3(x) = \frac{1}{3}\Big(\frac{(1-x)\ln(1-x) + x}{x^2}\Big).$$

The distance between cumulative distribution function of cross-ratios is:

$$d(x, y) = \min(|F(x) - F(y)|, 1 - |F(x) + F(y)|). \quad (7)$$

This distance is employed instead of the distance of cross-ratios so that the equal distance corresponds to the equal location error. The erroneous distribution on condition of 5% random locative noises is shown in Fig.7(b).

3.4 Reprojection Verification

Although the distance of cumulative distribution function is used, the projective invariant is very sensitive to the location errors. If the threshold is too big, then there will be more than one matching result; contrarily, if the threshold is too small, then it will fail to find a match. It is impossible to obtain the correct match if only using a threshold of cross-ratios. Therefore, a little big value should be chosen as the threshold and all matches obtained will be reprojected to the reference marker. The correct match will be recognized by the reprojection verification in the end. After that, several candidate triangle pairs that have correspondence with the reference marker are obtained. There are 6 corresponding points for every triangle pair and the related pair on the marker.

On the assumption that $\boldsymbol{P}_i^c = [x_i^c \ y_i^c \ 0 \ 1]^{\mathrm{T}}$ would be the points on the captured image and $\boldsymbol{Q}_i^m = [x_i^m \ y_i^m \ 0 \ 1]$ would be the points on the marker, then (8) can be obtained.

$$\boldsymbol{P}_i^c = \boldsymbol{P} \cdot \boldsymbol{Q}_i^m$$
$$\begin{bmatrix} x_i^c \\ y_i^c \\ 0 \\ 1 \end{bmatrix} = \lambda \boldsymbol{P} \cdot \begin{bmatrix} x_i^m \\ y_i^m \\ 0 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} x_i^c \\ y_i^c \\ 1 \end{bmatrix} = \lambda \boldsymbol{H} \cdot \begin{bmatrix} x_i^m \\ y_i^m \\ 1 \end{bmatrix},$$
$$i = 0, \ldots, 6, \quad (8)$$

where $\boldsymbol{P}$ is the projective matrix and $\boldsymbol{H}$ is the homography matrix. If the 6 points were given, then a linear equitation can be obtained and $\boldsymbol{H}$ can be computed by such linear methods as singular value decomposition. Then the reprojective error of other triangles detected in the captured image could be obtained from (9).

$$e_j = \sum_{i=1}^{3} |\boldsymbol{P}_i^{Tj} - \boldsymbol{H}\boldsymbol{Q}_i^{Tj}|, \quad (9)$$

where e_j is the reprojection error of the j-th triangle, $\boldsymbol{P}_i^{Tj}$ is the i-th vertex of the j-th triangle in the captured image and $\boldsymbol{Q}_i^{Tj}$ is the i-th vertex of the j-th triangle in the marker.

Pose estimation needs at least 4 corresponding points, so actually 2 triangles are enough to provide more than 4 points to ensure the correctness of recognition. This paper uses 3 triangles in the verification to guarantee the marker matching and bring more robustness to the system.

If $e_j < d_{\text{threshold}}$ holds, the j-th triangle is considered to have a correct match. If enough correctly matched triangles can be found in a captured image, the marker will be detected correctly and the vertices of all the matched triangles could be used in pose estimation.

4 Process of Robust Marker Recognition Algorithm

The process of the robust algorithm is summarized as follows.

1) Watershed Adapter Threshold (WAT).

2) Connected region detection.

3) Triangle extraction.

4) The white tag is examined to distinguish the order of the triangles. Triangles without a white tag are discarded.

5) 1D Canny edge detection.

6) Two triangles are chosen randomly.

7) Four independent cross-ratios are calculated to search for matching triangle pairs in the markers by the cumulative distribution function distance.

8) For all the matching triangle pairs, the reprojection error is computed to find the correct correspondence.

9) If no correct correspondent triangle pair is found, go back to Step 6) and repeat for N times. If it fails after the N times, it means that there is no FPI marker in the captured scene. N is chosen according to the number of triangles in the marker.

10) If FPI marker is identified, all the correct correspondent vertices verified by reprojection are outputted to the pose estimative algorithm.

5 Experiments

FPI marker system was implemented with C ++ and Microsoft DirectX library. The hardware platform in the experiments includes: Pentium 4 2.4GHz, 1GB memory, nVidia 5900 display adaptor card.

Due to some important improvements on the implementing coding and data structure, the system can achieve more than 30fps stably in the experiments.

Experiments have been conducted to investigate the performance of the FPI system. As described in the above sections, the most important advantage of the proposed algorithm is that it utilizes the relative position of features for identification, so that a lot of encoding area can be saved without losing the extraction accuracy of feature points as well. The proposed marker system could use less area to provide more correspondent feature points, so that higher pose accuracy could be achieved as shown in the following experiments.

5.1 Correspondent Features Dependent on Marker Area

ARToolKit is the most popular fiducial marker system and performs best compared with the other systems in the small region recognition experiment as demonstrated in Zhang's paper[9]. So ARToolKit system is selected as the reference in the experiments.

Both the ARToolKit and FPI markers are located 700mm away from the camera on the direction of the optical axis. The camera used is a WebEye USB 1.1 camera which has little distortion. Markers with 27 correspondent feature points for FPI system and 28 feature points for ARToolKit system are used separately (Fig.8). Markers in different sizes for two marker systems are tested to find out how many features could be detected.

Fig.8. Markers used in experiments.

Firstly, the experiments are conducted in the resolution of 640 × 480, the plan of the marker is perpendicular to the camera's optical axis, and the result is shown in Table 2.

Table 2. Recognition Rate Comparison for Two Marker Systems in Different Sizes in 640 × 480

Maker Size (mm^2)	ARToolKit (%)	FPI (%)
120 × 120	100.00	100.00
110 × 110	100.00	100.00
100 × 100	85.71	100.00
90 × 90	85.71	100.00
80 × 80	71.43	100.00
70 × 70	71.43	100.00
60 × 60	28.57	100.00
50 × 50	0.00	77.78

Secondly, the experiments are conducted with the resolution of 320 × 240, the plan of the marker is perpendicular to the camera's optical axis, and the result is shown in Table 3.

It can be concluded from Table 2 and Table 3 that in a rather small size, ARToolKit could not be recognized correctly, because it depends on the inner image to identify the marker.

At last, the performance when the marker is tilting from the camera's optical axis is tested in the resolution of 640 × 480. The results are shown in Table 4.

As demonstrated above, the FPI marker performs much better and is able to provide more correspondence feature points than ARToolKit marker with the

equal size and distance. Generally speaking, more correspondence points mean higher pose accuracy as shown in [28].

Table 3. Recognition Rate Comparison for Two Marker Systems in Different Sizes in 320 × 240

Maker Size (mm^2)	ARToolKit (%)	FPI (%)
130 × 130	100.00	100.00
120 × 120	85.71	100.00
110 × 110	71.43	100.00
100 × 100	71.43	88.89
90 × 90	57.14	77.78
80 × 80	28.57	66.67
70 × 70	0.00	55.56
60 × 60	0.00	44.45
50 × 50	0.00	33.34
40 × 40	0.00	0.00
30 × 30	0.00	0.00

Table 4. Recognition Rate Comparison When the Marker Pane is Tilting (In Resolution of 640 × 480)

Maker Size (mm^2)	Angle from Optical Axis	ARToolKit (%)	FPI (%)
120 × 120	80°	100.00	100.00
120 × 120	70°	100.00	100.00
120 × 120	60°	93.67	100.00
100 × 100	80°	83.50	100.00
100 × 100	70°	81.33	97.67
100 × 100	60°	77.67	98.50
80 × 80	80°	71.50	96.67
80 × 80	70°	68.33	93.33
80 × 80	60°	65.67	95.00

5.2 Accuracy of Features Extraction

Since there is no ground truth for the error evaluation of the marker detection, the investigation method in Zhang's paper[9] is employed. The corner detection implemented in OpenCV[29] library is used as the reference value to compute the location of the feature points. A comparison among the ARToolKit system, the system developed by Siemens Corporate Research (SCR) and FPI system is made. As the software of SCR marker system could not be obtained, the system for comparison in the experiments is implemented by the authors according to [12, 13]. The error in pixel compared to OpenCV in different tilt angles is listed in Table 5. The average and standard deviation of the error are given in every cell.

It can be concluded from the experimental results that the performance of FPI system and SCR are much better than ARToolKit system. The performance of FPI system and SCR system are similar for they use the similar feature extraction algorithms.

Table 5. Errors Related to OpenCV Corner Detection

Angle	ARToolKit	SCR	FPI
90°	1.55/0.32	0.59/0.14	0.58/0.15
80°	1.47/0.09	0.58/0.14	0.60/0.14
70°	1.43/0.11	0.62/0.14	0.59/0.14
60°	1.42/0.12	0.78/0.24	0.79/0.24
50°	1.40/0.12	0.56/0.12	0.56/0.12
Average	1.454/0.18	0.626/0.15	0.624/0.15

6 Conclusion

A novel Augmented Reality Fiducial marker system based on projective invariant is presented in this article. It is necessary for the current AR fiducial marker systems to employ inner pattern image to identify the marker, and it is a waste because this pattern area has no contribution to the accuracy of pose estimation. The marker system based on projective invariant uses the relation of the feature points' position directly to identify the makers and could provide more correspondent feature points. Projective invariant is very sensitive to the location errors so that the cumulative distribution function is used and a robust detection process is proposed to realize stable and reliable marker detection. Lots of tests have shown that this novel marker system could reduce the necessary size of marker and improve the accuracy of the Augment effectively.

There is a valuable direction for the future research: for the marker systems directly using inner encoding image it is easy to investigate the possible number of the marker set to be distinguished to each other. But for the presented marker system, it will provide new markers to change the shape of triangles in the marker. The number of markers is not determined by discrete quantity. It is not so intuitive to count the possible number of markers, more complicated effort is needed. So this problem is not solved in this paper. This result of possible marker number will be useful for the scenario using multiple markers.

References

[1] Azuma R. A survey in Augmented Reality. *Presence — Teleoperations and Virtual Environment*, Aug. 1997, 6(4): 355∼385.

[2] Ohta Y, Tamura H. Mixed Reality Merging Real and Virtual Worlds. Springer-Verlag, 1999, pp.379∼390.

[3] Youngkwan Cho, Jongweon Lee, Ulrich Neumann. A multi-ring fiducial system and an intensity-invariant detection method for scalable augmented reality. In *Proc. IEEE and ACM International Workshop on Augmented Reality*, San Francisco, California, 1998, pp.147∼165.

[4] Leonid N, Eric F. Circular data matrix fiducial system and robust image processing for a wearable vision-inertial self-tracker. In *Proc. IEEE International Symposium on Mixed*

and Augmented Reality (*ISMAR'02*), Los Alamitos, California, 2002, pp.27~36.
[5] Lopez de Ipina. D. TRIP: A low-cost vision-based location system for ubiquitous computing. *Personal and Ubiquitous Computing*, May 2002, 6(3): 206~219.
[6] Forsyth, D. Invariant descriptors for 3-d object recognition and pose. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 1991, 13(10): 971~991.
[7] C B Owen, Fan Xiao, P Middlin. What is the best fiducial? In *Proc. 1st IEEE International Workshop on Augmented Reality Toolkit*, Darmstadt, Germany, Sept. 2002, p.8.
[8] Kato H, Billinghurst M. Marker tracking and HMD calibration for a video-based Augmentnd Reality conferencing system. In *Proc. 2nd IEEE and ACM International Workshop on Augmented Reality* (*IWAR 99*), San Francisco, CA, 1999, pp.85~94.
[9] Zhang X, Fronz S, Navab N. Visual marker detection and decoding in AR systems: A comparative study. In *Proc. Int. Symp. Mixed and Augmented Reality*, Darmstadt, Germany, 2002, pp.97~106.
[10] Rekimoto J. Matrix: A realtime object identification and registration method for augmented reality. In *Proc. 3rd Asia Pacific Computer Human Interaction*, Hayama-machi, Kanagawa, Japan, July 15~17, 1998, pp.63~68.
[11] Rekimoto J, Ayatsuka Y. CyberCode: Designing Augmented Reality environments with visual tags. In *Proc. Designing Augmented Reality Environments*, Denmark, April 2000, pp.1~10.
[12] Zhang X, Navab N. Tracking and pose estimation for computer assisted localization in industrial environments. In *Proc. Applications of Computer Vision*, 4–6 Dec. 2000, pp.214~221.
[13] Zhang X, Navab N, Liou S. E-commerce direct marketing using augmented reality. In *Proc. IEEE Int. Conf. Multimedia Expo.*, New York, 2000, pp.88~91.
[14] Fiala M. Fiducial marker systems for augmented reality: Comparison between ARTag and ARToolkit. In *Proc. MIRAGE 2005: Computer Vision/Computer Graphics Collaboration for Model Based Imaging, Rendering, Image Analysis and Graphical Special Effects*, INRIA Rocquencourt, France, Mar. 2005.
[15] Fiala M. ARTag, a fiducial marker system using digital techniques. In *Proc. IEEE Computer Vision and Pattern Recognition* (*CVPR 2005*), San Diego, California, June 2005, pp.590~596.
[16] David F, Joseph M, Andrew Z, Chris C, Aaron H, Charles R. Invariant descriptors for 3D object recognition and pose. *IEEE Trans. Pattern Analysis and Machine Intelligence*, October 1991, 23(2): 116~128.
[17] Vassilios T, Konstantinos C, Panos T. Landmark-based navigation using projective invariants. In *Proc. 1998 IEEE/RSJ Int. Conference on Intelligent Robots and Systems*, Canada, October 1998, pp.342~347.
[18] George B, Michael G, Niels L. Learning geometric hashing functions for model-based object recognition. *IEEE Transactions on Neural Networks*, 1998, 9(5):560~570.
[19] Isaac W, Manjit R. Model-based recognition of 3D objects from single images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, February 2001, 23(2): 116~128.
[20] Robert L, Jurriaan M. Optical tracking using projective invariant marker pattern properties. In *Proc. IEEE Virtual Reality 2003* (*VR'03*), Los Angeles, California, March 2003, pp.191~198.
[21] Alberto S, Antoni L, Walter E. Sensibility, relative error and error probability of projective invariants of planar surfaces of 3d objects. In *Proc. 11th IAPR International Conference on Pattern Recognition*, Hague, Netherlands, 1992, pp.328~331.
[22] Mundy J, Zisserman A. Geometric Invariance in Computer Vision. Massachusetts: Cambridge, MIT Press, 1992.
[23] Lenz R, Meer P. Point configuration invariants under simultaneous projective and permutation transformations. *Pattern Recognition*, 1994, 27(11): 1523~1532.
[24] Meer P, Lenz R, Ramakrishna S. Efficient invariant representations. *Int. J. Computer Vision*, 1998, 26(2): 137~152.
[25] Beucher S, Lantuejoul C. Use of watersheds in contour detection. In *Proc. Int. Workshop on Image Processing, Real-Time Edge and Motion Detection/Estimation*, Rennes, 1979, pp.17~21.
[26] Gabriella C, Olivier F. Algebraic and geometric tools to compute projective and permutation invariants. *IEEE Trans. Pattern Analysis and Machine Intelligence*, January 1999, 21(1): 58~65.
[27] Aström K E, Morin L. Random cross ratios. In *Proc. Ninth Scandinavian Conf. Image Analyses*, 1995, Uppsala, Sweden, pp.1053~1061.
[28] Lu C, Hager G, Mjolsness E. Fast and globally convergent pose estimation from video images. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2000, 22(6): 610~622.
[29] OpenCV: Open source computer vision library. http://www.intel.com/research/mrl/research/opencv/.

Yu Li was awarded a B.S. degree in optical and electronic engineering from Beijing Institute of Technology in 2003, and is studying for the Ph.D. degree in virtual reality and augmented reality at Beijing Institute of Technology. His research interests include augmented reality, computer vision, and computer graphics.

Yong-Tian Wang was awarded his Ph.D. degree in physical engineering at Reading University, UK. He is a professor of Beijing Institute of Technology, director of the research center of optical electronic information technology, a member of the Commission of Science and Technology of the National Education Ministry. His research interests include: optical design and CAD, new optical instrument, image processing and virtual reality and augmented reality. He has won 1 third-class national prize from the Ministry of Science and Technology. He has also won the "Yantze River Scholar" prize in 2001.

Yue Liu was awarded a Ph.D. degree from Jilin University. His research interests include augmented reality, computer vision.