

Markerless Registration Based on Natural Features Tracking for Augmented Reality

Ziliang Jiao

School of Computer Engineering and Science, Shanghai University, No.149, Yan Chang Rd, Shanghai, China
jiao.wayne@yahoo.com

Abstract - In augmented reality system, registration is always a problem which is difficult to resolve effectively, especially for marker less registration. Due to the placed marks for marked visual registration method, it would cause visual pollution, and the registration often fails when the mark is partially occluded. Registered quality directly affects the fusion effects of the virtuality and reality. This paper proposed a marker less tracking registration method combined natural feature points with affine reprojection. Firstly it extracted natural feature points from real scene with SIFT algorithm, then exploited KLT algorithm to track and match the effective feature points in each image, carried out the homography matrix of each frame with affine reprojection technology, and computed the real scene 3D registration for each frame image, so as to implement the tracking registration effectively. This method is also applicable to those cases of plane feature points being occluded.

Index Terms - Augmented Reality, Markerless, Tracking, Registration

I. Introduction

Augmented reality is an important researching field in the development of contemporary computer technology, the superposition of virtual objects and real scene is the core area in augmented reality, and the 3D registration is very important to the augmented reality development and its application.

The 3D registration is to compute the relative position to real world in real time and precisely, and superimpose the virtual information onto the position of the real world. It is an important topic in the field of augmented reality research, and provides a strong guarantee for achieving augmented reality actual seamless integration. So far, researchers have put forward a variety of ways to register, such as registration based on sensors or based on computer vision, etc.

Compared with the sensor based registration approach, Vision based registration approach is dominating, which require simpler and cheaper equipment. Vision based registration approach can be divided into registration based on the mark and no mark.

Among the registration approaches based on mark, the ARToolkit registration technology is the most mature algorithm, it has several features such as fast computation speed, low complexity, stable and accurate registration, it is available to such equipments with limited compute resources. But this approach requires markers easy to be recognized to be placed in a real scene, which has brought visual pollution, and in the case of partial shade, easy to cause registration failure, which limits the motion range of the camera, so that limits the application range of AR.

So the research of marker less registration approach is becoming more and more popular. Since there is no mark placed in advance, the system requires a very practical and reliable solution to identify registration area. The approach of iterating homography matrix has been mentioned before, but the deviation will increase gradually during the superposition, so that it requires the feature points to be extracted again after a certain period of time, time consumption is not available.

In this paper, we use the affine coordinate system to complete the 3D registration, affine coordinate system was established based on the extraction of feature points, choose four feature points to construct coordinate system. The reprojection reconstruction is built based on tracking the feature points by KLT algorithm, and the image projection is obtained from selected feature points. After that we can get the camera extrinsic parameters and position, and then the rendering of virtual objects can be completed.

II. The basic principle

A. the camera imaging model and homography matrix

The 3D registration includes the shift and projection transformation among image plane coordinate, camera coordinate system and. In a pinhole camera imaging model, the homogeneous coordinate point of world coordinate system, its projection on image plane is .Their relationship is represented as follows:

$$X = P X_w = \lambda K M X_w = \lambda K \begin{bmatrix} R_1 & R_2 & R_3 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} X_w \quad (1)$$

When is located in the marked plane, then the formula (1) can be converted to formula (2):

$$X = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \lambda K \begin{bmatrix} R_1 & R_2 & R_3 & T \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 0 \\ 1 \end{bmatrix} = \lambda K \begin{bmatrix} R_1 & R_2 & T \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} = H_w \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \quad (2)$$

Where: K in formula (1) and (2) represents the internal camera parameter matrix, M is the outer camera parameter matrix.

The main task of 3D registration is to computer the internal and outer parameters, the K parameter matrix can be got after camera calibration, and it keeps invariant. The matrix M includes three rotating components (R1, R2, R3) and a shift component T. Based on formula (2), if we know the internal

parameter K and homography matrix, and the condition R1, R2 and R3 with the relationship as $R3=R1*R2$, so parameters of (R1, R2, R3) and T can be got, and further the outer parameter M also can be got.

Affine reconstruction and affine reprojection

The affine coordinate system definition: assume a point set as $\{P_1, P_2, \dots, P_n \in R^3 (n \geq 4)\}$, and at least four points are non-coplanar, an affine coordinate system can be defined by the four non coplanar points.

Two important properties of the affine coordinate system:

Through the affine reconstruction we can get the corresponding position of the space affine coordinate system of the points on the image plane, supposed that the two different image plane under perspective must be known so does the origin and affine basis points under different view angle. Assumed that $X=(U,V,W,1)$ is a point of projection coordinate system in the real space, its two image plane coordinate under two different view angle are $x^1=\{u^1, v^1, 1\}$ and $x^2=\{u^2, v^2, 1\}$, based on formula (3), we can find out its locations of affine coordinates system in real space $X=(U,V,W,1)$:

$$\begin{bmatrix} u^1 \\ v^1 \\ u^2 \\ v^2 \end{bmatrix} = \begin{bmatrix} u_1^1 - u_0^1 & u_2^1 - u_0^1 & u_3^1 - u_0^1 & u_0^1 \\ v_1^1 - v_0^1 & v_2^1 - v_0^1 & v_3^1 - v_0^1 & v_0^1 \\ u_1^2 - u_0^2 & u_2^2 - u_0^2 & u_3^2 - u_0^2 & u_0^2 \\ v_1^2 - v_0^2 & v_2^2 - v_0^2 & v_3^2 - v_0^2 & v_0^2 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} = R_{4 \times 4} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} \quad (3)$$

Where: $[u_j^i \ v_j^i]^T (i=1, 2; j=0, 1, 2, 3)$, is defined as affine origin u and the three affine basis points within the image plane under different view angle, $R_{4 \times 4}$ is affine reconstruction matrix.

The main purpose of affine reprojection is to get the projection on image plane, according to the specified point's true space and the four non-coplanar affine coordinates. Assumed the point is $X=(U,V,W,1)$, then its projection $x=\{u, v, 1\}$ can be got by formula(4):

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} u_1 - u_0 & u_2 - u_0 & u_3 - u_0 & u_0 \\ v_1 - v_0 & v_2 - v_0 & v_3 - v_0 & v_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} = M_{3 \times 4} \begin{bmatrix} U \\ V \\ W \\ 1 \end{bmatrix} \quad (4)$$

Where: $[u_j^i \ v_j^i]^T (i=1, 2; j=0, 1, 2, 3)$ is defined as the image plane coordinates of the affine origin and three affine basis points with different view angle, $M_{3 \times 4}$ is affine reprojection matrix.

III . the registration algorithm based on natural features

The registration method of 3D is mainly to get the natural features from real world, and tracking these feature points, projecting these points' 3D coordinates to observed 2D image coordinate system, and establish corresponding relation between a two-dimensional image plane and three-dimensional

coordinates in real world, compute the homography matrix of each frame image, estimate the camera pose information, so as to realize 3D registration.

Feature tracking algorithm includes two main steps: extracting natural feature point and feature point tracking. Feature extraction method based on gray image is the most widely used, more commonly used algorithm include KLT operator, Harris operator, Sift algorithm, etc. The principle of KLT operator is simple, easy to implement and less time consuming, But to KLT operator, it need to compute the second order derivative, and smoothed by Gaussian kernel, easily lead to low accuracy to the detected feature points, more sensitive to noise, due to the feature points matching accuracy drops, it can lead to problem of inaccurate tracking while the camera or target objects move fast.

Sift algorithm is proposed by Lowe in 1999, further completed in 2004. It mainly using multi-scale ideas to build Gaussian pyramid image, Gaussian differential pyramid image is obtained based on two neighboring Gaussian image subtraction under same scales.

In differential Gaussian pyramid image, to calculate the local minima as candidate extremum points, then remove the unstable and marginal feature points, which makes the feature points with strong stability and uniqueness, and these point are robust to light, noise, affine change. Therefore this article mainly use SIFT algorithm to extract the natural feature points in real scenario for tracking match. The algorithm is divided into five steps are shown in figure 1 as below.

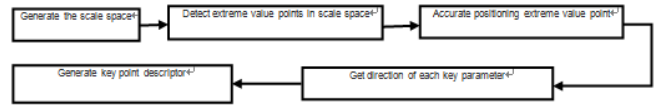


Fig. 1 Sift algorithm to extract the feature points

After get feature points, it is necessary to track match the feature points. Considering the KLT matching algorithm is the typical optimal estimated matching method, it is mainly to get the optimal matching position by estimating the square of gray scale differential and the minimum principle, this method does not need to carry on the exhaustive search, with less time-consuming, therefore it is widely applied to real-time tracking registered.

KLT matching algorithm is a kind of feature points matching method using interframe continuity information, the algorithm's theory is as follows:

Assumed $I(x, y, t)$ is the image frame at time t, $J(x+d)$ is the image frame at time $t+\tau$, the two frames meet the following formula (3-1):

$$\begin{aligned} I(x, y, t) &= J(x+d) \\ J(x+d) &= I(x+dx, y+dy, t+\tau) \end{aligned} \quad (5)$$

Within the window of w, all points shift a distance with $d=(dx, dy)$ along the same direction, so a point (x, y) at time t moved to the new position (x+dx, y+dy) at time t+τ, finally we get

the points (x', y') . So the matching problem can be got by calculating the minimum of formula (6):

$$\xi = \iint_w [J(x+d) - I(x)]^2 w(x) dx \quad (6)$$

Where, $w(x)$ is weighting function, usually $w(x)=1$. To carry on the Taylor expansion of (6), we can get

$$Zd = e \quad (7)$$

Where, Z is a matrix of $2*2$, e is a vector of $2*1$

$$Z = \iint_w g(x) g^T(x) w(x) dx$$

$$e = \iint_w [I(x) - J(x)] g(x) w(x) dx \quad (8)$$

$$g(x) = \begin{bmatrix} \frac{\partial(I+J)}{\partial x} \\ \frac{\partial(I+J)}{\partial y} \end{bmatrix}$$

To calculate the Newton iterative to formula (8), we can estimate the offset $d = (dx, dy)$.

In this paper, considering the advantages of Sift algorithm's strong stability to extract the feature points and the advantages of KLT algorithm's high precision matching with fast speed, this paper mainly utilize KLT matching algorithm to track the feature point based on the extracted the feature points by SIFT algorithm. Figure 2 shows feature extraction by SIFT and tracking by KLT algorithm respectively.

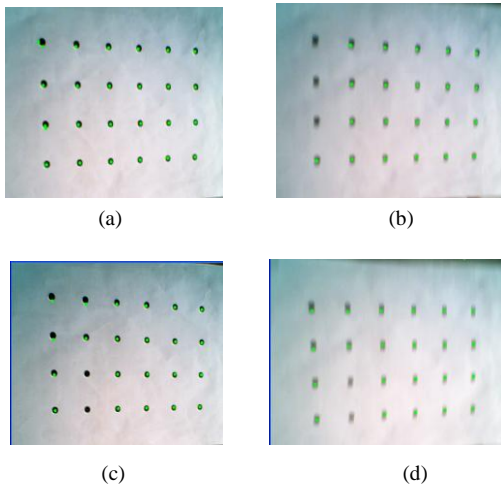


Fig. 2 Experimental contrast figure

IV. The registration algorithm

To achieve the registration of 3D objects, we only need to figure out the homography matrix H_w under knowing internal parameter of camera.

From formula (2), if given the coordinates of $P(x_i, y_i)$ in image plane and the $P(x_w, y_w, 1)$ in world coordinates system

the homography matrix H_w will be got.

In this paper, registration is realized based on the affine reprojection, firstly exploit the affine reconstruction principle, we can get the four points' coordinates in the real space system, which are located in two image plane under different view angle, then restore the four points' positions in the currently image plane according to reprojection principle, the points were specified by the user in the initial phase. The specific algorithm is as follows, whose flow chart is shown in Figure 3:

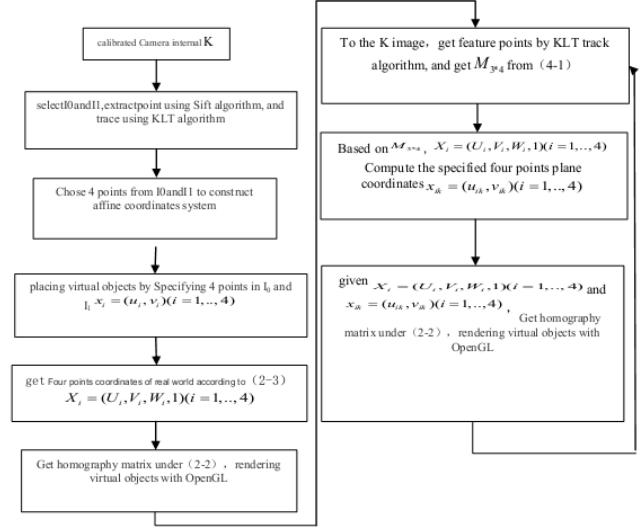


Fig. 3 flow chart of algorithm

- Through the camera calibration, the camera internal matrix K is determined.
- Select two images under different view angle in the real space as a reference, marked as I_0 and I_1 , then extract the feature points in image I_0 according to SIFT algorithm, and further trace the points by the KLT algorithm.
- Choose four points which are mutual matched with high stability to establish affine coordinate system in two figures, specify an affine origin and three affine basis points.
- Respectively specify four points in the selected two reference images, $x_i = (u_i, v_i) (i=1, \dots, 4)$, the four points are Coplanar, but are not collinear, according to the user needs, can be a feature point, also can be non-feature points, is mainly used to overlay virtual objects on it.
- Based on the affine reconstruction formula (3) and specify four points coordinates respectively in a real scenario of the affine coordinate system, $X_i = (U_i, V_i, W_i, 1) (i=1, \dots, 4)$, to establish the plane coordinates and the corresponding relation between the real coordinates
- For the K frame, based on affine origin and affine basis points by KLT algorithm, according to the following formula (9) to calculate the corresponding affine projection matrix $M_{3 \times 4}$ Where, $[u_i^k \ v_i^k \ 1]^T (i=1, 2; j=0, 1, 2, 3)$,

is the updated affine origin and affine basis points in the K image plane.

$$M_{3 \times 4} = \begin{bmatrix} u_1^k - u_o^k & u_2^k - u_o^k & u_3^k - u_o^k & u_o^k \\ v_1^k - v_o^k & v_2^k - v_o^k & v_3^k - v_o^k & v_o^k \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

- According to step 5 to get the four points coordinates under affine coordinates and step 6 affine projection matrix $M_{3 \times 4}$, then we can get the four points coordinates on the current frame image plane $x_{ik} = (u_{ik}, v_{ik}) (i=1, \dots, 4)$
- After given the four points coordinates in the image plane and the plane coordinates in the real world, according to the formula (2), get the homography matrix of current frame, so as to figure out the registration matrix. Then use OpenGL to render the virtual object and registry virtual object in the right position.

V. The experimental results

In this paper, we use OpenGL, Open CV and VS2008 to implement the proposed registered tracking algorithm. The hardware configuration is as follows, CPU of Intel Core Duo with frequency of 2.80 GHz and 2G memory, graphics card type is Geforce 9600 GT with a resolution of 1280 * 1280.

Experimental results are shown in figure 4 below, the figure (a) and (b) are the selected two frames of reference images under different Angle of view in the initialization phase, the red dot is specified, the square formed by 4 red points represent the location to place virtual object.

Figure (a) is the initial frame, the red dots stands for virtual object position to register, in figure (b) place a cube on the position specified by the red dots, represents the effectiveness of registration.

Figure (c) (d) are the registered effectiveness of virtual object under different perspective, Figure (e) and (f) register result on the specified registered location in the case of partial shaded, the result shows that even in the case of partial shaded, the algorithm can realize the registration well, greatly improved the registered feasibility than any algorithms based on the artificial marks.

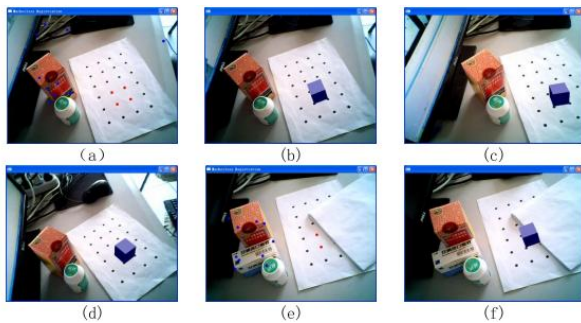


Fig. 4 experimental results

VI. Conclusions

In augmented reality, 3D registration is the most critical and most basic AR technology, and its registration effect can directly affect the fusion quality of real and virtual object. This paper discussed the approach of tracking based marker less registrations in details. Firstly introduced registration method of the augmented reality, including registered tracking approaches based on the marked and unmarked.

Second an improved feature point extraction and tracking algorithm is introduced in details, mainly considering the stability and accuracy to extract the feature points using SIFT algorithm, and KLT algorithm is a matching method based on the optimal estimation with high matching accuracy and fast speed, so the SIFT algorithm is adopted to extract the feature points, the KLT algorithm for tracking and matching feature points.

And the algorithm can be divided into two phases, the initialization and operation, during the initialization phase, its task is mainly to extract the feature points and build the affine coordinate system by Sift algorithm, with affine reconstruction technology, compute the real world coordinates of the four specified points to place virtual object,

During operation phase, calculated the affine projection matrix to restore the current image plane coordinates of the four points specified in initialization phase. And using the four points to calculate the homography relation between the current image plane and real world scenario, that is the homography matrix of each image. Further computed registration to overlay virtual objects on the real scenario corresponding to each image frame, so as to realize the stable and efficient tracking based marker less registration.

References

- [1] G. Klein and D. Murray, Parallel Tracking and Mapping for Small AR Workspaces, Proc. IEEE/ACM Int'l Symp. Mixed and Augmented Reality (ISMAR '07), 2007: 225-234.
- [2] Peng Ren, Bo kang, A New Tracking Method Based on Markers in Augmented Reality. Journal of system simulation. 2009, 21(2): 465-468.
- [3] Yang li, Chao Sun, Mingmin Zhang, Wenjie Zhang, Zhigeng Pan, parallel augmented reality registration method of Tracking and matching. journal of image and graphics. 2011.16(4):680-685.
- [4] Teng Fei, Liang Xiao-hui, He Zhi-ying, Hua Guo-liang. A Registration Method Based on Nature Feature with KLT Tracking Algorithm for Wearable Computers. In Proceedings of CW. 2008: 416-421.
- [5] Kiyoung Kim Vincent Lepetit WoontackWoo. Scalable real-time planar targets tracking for digilog books. Vis Comput, 2010,26,(6-8): 1145-1154.
- [6] Lee, T., Hollerer, T.: Multithreaded hybrid feature tracking for markerless augmented reality. IEEE Trans. Vis. Comput. Graph. 2008: 355-368.
- [7] Yuan M L, Ong S K, Nee A Y C. Registration using natural features for augmented reality systems. Visualization and Computer Graphics, IEEE Transactions on, 2006, 12(4): 569-580.