Journal of Theoretical and Applied Information Technology
31 August 2012. Vol. 42 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A REAL-TIME REGISTRATION METHOD OF AUGMENTED REALITY BASED ON SURF AND OPTICAL FLOW

[1]HONGBO LI, [2]MING QI AND [3]YU WU
[1, 2, 3]Institute of Web Intelligence, Chongqing University of Posts and Telecommunications,
Chongqing, 400065, China
E-mail:[1] lihongbo@cqupt.edu.cn , [2]qiming-163@163.com , [3]wuyu@cqupt.edu.cn

ABSTRACT

To solve the time-consuming and errors in traditional 3D registration that restrict the application of augmented reality technology, a real-time registration method of augmented reality based on natural features is presented. It includes two procedures, the image feature detecting and tracing for registration. Firstly, to improve arithmetic speed, SURF (Speeded up Robust Features) algorithm is adopted to replace traditional SIFT (Scale-invariant Feature Transform) algorithm. Since detecting image features from every frame reduces efficiency of 3D registration, a L-K (Lucas-Kanade) optical flow algorithm based on image pyramid is used to trace the detected features in the second procedure. Accumulation of registration error is avoided by adopting feature redetection strategy which maintains 3D registration stability. Experiments demonstrate that our approach improves the accuracy and real time capabilities of 3D registration.

Keywords: *Image Feature, Homography, Optical Flow, 3D Registration, Augmented Reality*

1. INTRODUCTION

Augmented Reality (AR) is an emerging technology appeared in the 1990s. It uses virtual world to augment people's perception of real world by combining computer-generated virtual information with real-world environment. AR technology has three main characteristics, including real-virtual combination, 3D registration and real-time interaction. 3D registration is a key technology which performs the position registration of virtual scene and the real world. It has two types, including technologies based on patterns and natural features [1]. The AR technology based on patterns has developed maturely. Since the AR technology based on natural features can be achieved without placing man-made patterns in real scene, it has a wide range of applications.

Currently, research of AR is more concentrated on registration method based on natural features. L. G. Hua used Harris Corner detection algorithm to extract image corners firstly, then adopted a correlation matching method to match the Gray correlation value of corners and got the initial set of matching corners [2]. However， it had low positioning accuracy and only pixel level corners were detected. The Scale-invariant feature transform (SIFT) algorithm was adopted to replace the Harris algorithm [3, 4]. Its superiority is that feature descriptor is invariant to scale, orientation, affine distortion, and partially invariant to illumination changes. The detected features are robust to changes in noise and minor changes in viewpoint. However, its computational efficiency is so low that reduces 3D registration efficiency. Kanade-Lucas-Tomasi (KLT) feature tracker can solve this problem, but the displacement between adjacent frames was limited in one pixel [5].

Since time-consuming and large errors of traditional 3D registration restrict the application of the AR technology, a real-time registration method based on natural features is presented. Because Speeded up Robust Features (SURF) features detection algorithm has the same advantages as SIFT but higher computing efficiency, we adopt SURF algorithm to replace traditional SIFT algorithm firstly. Since detecting image features from every frame reduce 3D registration efficiency, we use a Lucas-Kanade (L-K) optical flow algorithm based on image pyramid to trace the detected features. The accumulation of registration error is avoided by adopting features redetection strategy that maintains 3D registration stability.

2. RELATED METHODS

3D registration principle based on homograghy. The key to realize AR 3D registration is obtaining camera projection matrix

Journal of Theoretical and Applied Information Technology
31 August 2012. Vol. 42 No.2

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

which represents the relationship between world coordinate system and image coordinates. The general form of projection equation is

$$x = PX = K\left[R \mid t\right]X \quad (1)$$

where P is 3×4 projection matrix, x= $(x, y, w)^T$ is homogeneous coordinates of image features, X= $(X, Y, Z, 1)^T$ is homogeneous coordinates of feature points in world coordinates, and K is 3×3 camera intrinsic parameter matrix and is a known value here. R and t is rotation matrix and displacement vector of the camera relative to the world coordinate system respectively [6].

The world coordinate system can be chosen arbitrarily. To facilitate the calculation, we choose the XY plane of the world coordinate system coincides with the scene plane. Z coordinates of the points in this system relative to scene plane are all zero. The projection equation can be described as

$$x = K\left[r_1 \quad r_2 \quad r_3 \quad t\right]\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = K\left[r_1 \quad r_2 \quad t\right]\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = H\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (2)$$

Where r_i is column of the camera rotation matrix, X= $(X, Y, 0, 1)^T$ is homogeneous coordinate of points in the XY plane of world coordinate system, and x= $(x, y, w)^T$ is homogeneous coordinate of points in scene plane coordinates [7]. The projection matrix can be obtained by Eq. (2) to achieve 3D registration.

SURF features detection algorithm. SURF algorithm is based on multi-scale space theory, and the feature extractor is based on Hessian matrix. Since Hessian matrix has good performance in calculation time and accuracy, scale space can be achieved with the form of image pyramid. The values of features can be obtained by calculating the determinant of Hessian matrix of every pixel in image pyramid. For one point $\hat{x} = (x, y)$ in image I, Hessian matrix can be defined in scale σ as

$$H = \begin{bmatrix} L_{xx}(\hat{x}, \sigma) & L_{xy}(\hat{x}, \sigma) \\ L_{xy}(\hat{x}, \sigma) & L_{yy}(\hat{x}, \sigma) \end{bmatrix}. \quad (3)$$

Where L_{xx} is the convolution result of second derivative of Gaussian filter $\frac{\partial^2}{\partial x^2} g(\partial)$ and $I(x, y)$. $g(\partial) = \frac{1}{2\pi\sigma^2} e^{\frac{-(x^2+y^2)}{2\sigma^2}}$, L_{xy}, L_{yy} have similar meanings [8].

SURF algorithm uses box filters to replace the second order Gaussian filter approximately, and uses integral image to accelerate the convolution. It is more efficient than SIFT.

After obtaining location and scale information of the detected feature points, a unique direction and feature descriptor is assigned to every feature points. To form the four-dimensional vector

$$V_{sub} = (\sum d_x, \sum |d_x|, \sum d_y, \sum |d_y|) \quad (4).$$

In the 16 sub-regions near a feature point, the summation of the Haar wavelet response and its absolute values $\sum d_x$, $\sum d_y$, $\sum |d_x|$, and $\sum |d_y|$ must be calculated in the direction of x and y [7]. For every feature point, a descriptor vector of 64 dimensionalities is obtained. Since the vector must be normalized, this algorithm is robust to illumination change.

3. AR REGISTRATION BASED ON SURF AND OPTICAL FLOW

Proposed issues and system design. Accurate projection matrix can be obtained by using the features detection method for the initial registration. Since this method is time-consuming, tracing for registration is needed to improve the real time capabilities. Because the registration result of every frame is dependent on the previous frame, the registration error will be accumulated as the frame number increases. As the virtual objects have large changes of displacement, rotation and scale, the registration will fail.

We present a method to solve these conflicts and improve the accuracy and real time capabilities. It includes image feature detecting and tracing for registration. In the first procedure, the SURF algorithm is adopted to detect image feature points. These points are used to match with features of model image. If match is successful, the homography matrix and projection are obtained using Random Sample Consensus (RANSAC) algorithm. Then in the second procedure, if the number of matched features is less than the default, registration will return to the first procedure.

In the tracing for registration procedure, an L-K optical flow algorithm based on image pyramid is used to trace the features of two adjacent frames. If

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the number of traced feature points is greater than the present value, the homography is calculated with the tracing results, or the features detection phase restarts. To avoid the accumulation of registration error as the frame number increases, we adopt a features redetection strategy every 10 frames. So the accuracy and real time capabilities of 3D registration are all improved.

Features detection and initial registration. We use SURF algorithm to detect feature points in scene image, then coarse matching of features proceeds. The Best Bin First (BBF) algorithm based on k-dimensional (k-d) tree is used to coarsely match the descriptor of SURF features. But the coarse matching method will lead to some errors of matching points. In order to get an accurate homography, false matching set must be removed. RANSAC algorithm is used to remove the false matching and calculate the homography [9]. It divides all the matching points into inner and outer points according to an allowable error. The method takes full advantage of all the initial match points. The accurate point data can be used to parameter estimating and remove inaccurate matching points. Follows are the detailed steps of the homography calculation algorithm, where the current best estimate of the number of inner points S is 0, and the homography from the world coordinate system to the first frame coordinate system is H_w^1.

Step 1. Repeat N times of random sampling.

Step 2. Calculate the homography H_w^1 with the four pairs of matching points.

Step 3. Calculate Euclidean distance $d(x_i^{'}, Hx_i)$ of the corresponding point after this homography transformation for every matching point.

Step 4. Set a distance threshold T, and select the matched points which satisfy d<T as inner points.

Step 5. Compare with the number of current inner points with S. If the number is greater than S, the current inner points set will be the best estimation. Update S. At the same time, the current number of iterations is also dynamically estimated. If the number reaches N, H matrix and current inner points set will be saved and the iteration stops.

Step 6. Calculate the accurate homography H_w^1 with the current inner points set.

Tracing for registration. Since the time-cost of current features detection algorithm cannot be decreased below 50ms for the 640×480 images, features tracing method must be used in the subsequent registration to improve the real time capabilities. L-K optical flow algorithm based on image pyramid is adopted to trace the features because it is more stable than traditional KLT algorithm under large displacement. Images stratified by pyramid can also greatly improve the matching speed. The basic idea of this method is to construct a pyramid of an image sequences which is a series of progressively decreased images arranged like a pyramid.

Registration of the subsequent frames is achieved by tracing features of two adjacent frames, and the points traced are obtained from the initial registration phase. The homography H_{k-1}^k between frame k and frame k-1 can be got by tracing position of feature points. The homography H_w^k from the world coordinates to frame k coordinates can be obtained from the equation

$$H_w^k = H_{k-1}^k H_{k-2}^{k-1} ... H_1^2 H_w^1 \quad (5)$$

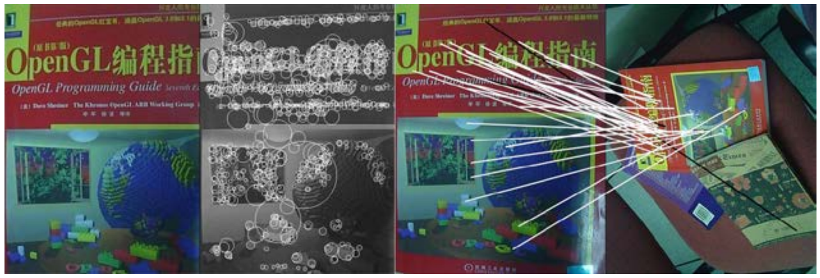
By the homography H and the orthogonality of R in Eq. (1), R, t and projection matrix P are calculated to complete the subsequent 3D registration [7].

4. EXPERIMENTAL RESULTS AND DISCUSSION

To achieve the AR 3D registration, a computer with Intel Core i3 2.93GHz CPU, 2GB memory, NVIDIA GeForce GTS 250 graphics card and common USB camera is used in the experiment. The OpenGL2.0 and OpenCV2.1 software library is used to realize the algorithm in Visual Studio 2005. The resolution of the collected image is 640×480 which is commonly used in network videos, and the frame rate is about 25 frames every second which can meet the requirement of real-time. After successful registration, a virtual teapot will be drawn in the scene image.

Features detection and matching experiments. We detect features in a general image using SURF algorithm. As shown in Figure 1, Figure 1 (a) is an original image that can be used as a reference template in registration. Figure 1 (b) shows the features detected using SURF algorithm from the original image. In this image 304 stable feature points are detected. After the detection and description of each feature point, a scale of feature points, coordinates of feature points, and a 64-dimensional neighborhood vector can be stored for the subsequent calculation.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



(a) (b) (c) (d)

Figure1 Template image and feature points (a, The template image; b, Feature points detected by SURF algorithm; Features matching and outer points rejection between the template image (c) and the scene image (d)).

The features matching experiment is based on the features detection. Firstly, the BBF algorithm based on k-d tree is used to the coarse matching. Secondly, RANSAC is used to the exact matching and remove the outer points. At last, the accurate homography and projection matrix will be calculated. Figure1 (c, d) shows the results of feature points matching. Finally we get 32 stable matching point pairs. Since this algorithm homography only needs four pairs of matching points, this method can obtain an accurate projection matrix for AR 3D registration.

Matching point pairs of two colors are all detected in the coarse matching process. The black feature point pairs are the false matching points which are removed by RANSAC algorithm, and white pairs are the final matched point pairs.

3D registration experiments. On the basis of the above experiments, we compare the time-cost of X. Y. Li's registration initialization method with our initialization method. X. Y. Li used the SIFT algorithm to detect the image feature points first, and then used transient chaotic neural network method to match feature points and achieve the registration initialization. Although the neural network method was used to optimize the process of matching, the overall efficiency of the initialization process was limited by SIFT algorithm. For the resolution used in this experiment, the average time-cost was 840ms. As the initialization phase was time-consuming, the features redetection strategy cannot be used. Although the features tracing can improve the efficiency of registration, the accumulated error would not be avoided. However, the average time-cost of our initialization is 180ms. So that even the features redetection strategy was used in the subsequent procedure, it would not decrease the efficiency.

The detailed test results in every phase of the experiment are recorded in Table 1. At the beginning of video capturing, there are not enough matching points to calculate the homography, so this stage is called unregistered stage. When the registration is successful, the frame rate can be maintained between 20-30 frames/sec because of the alternation of features detection and tracing. To validate the real time capability of the algorithm further, videos of different resolution are tested in the experiment. The results (Table 2) show that the presented approach improves real time capabilities of 3D registration.

Table 1 The test results statistics of every 3D registration stage

	Unregistered stage	Registration stage	
		Features detection	Features tracing
Number of features detected	183	272	-
Number of feature pairs matched	0	34	31
Time-cost [ms]	162	180	19
Frame rate [frames/sec]	6.2	5.5	52.6

Table 2 The frame rates of 3D registration in various resolutions

Resolution	Time-cost of features detection [ms]	Frame rate [frames/sec]
320×240	47	48
320×480	92	39
640×480	180	26

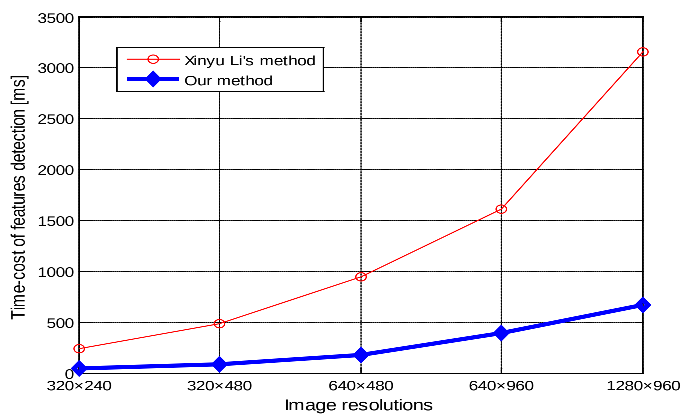
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The final experimental results are shown in Figure2. Experimental results demonstrate that the presented approach is robust to changes in rotation, illumination, scale and noise which can meet the requirement of AR applications.

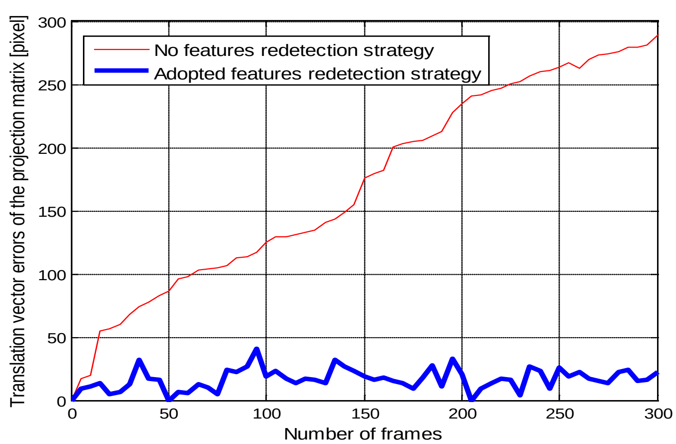
Finally, we test the efficiency of two methods in different resolutions. The impact of adopting the features redetection strategy to accumulated error is also tested. Figure 3(a) shows that the presented initial registration algorithm improves the efficiency of 3D registration greatly. Figure3 (b) shows that the translation vector error is not cumulative over time adopting the features redetection strategy. Results demonstrate that the presented approach improves the accuracy of 3D registration.

(a) (b) (c) (d)

Figure2 Registration results in various conditions (a, Rotation changes; b, Distance changes; c, Illumination changes, and d, Shelter).



(a)



(b)

Figure 3 Performance testing (a, Comparison of the time-cost of two different initial registration algorithms; b, Impact of adopting the features redetection strategy to the error of projection matrix).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. SUMMARY

A new method of AR real-time registration based on natural features is presented. It includes two procedures, which are image feature detecting and tracing for registration. In the first procedure, SURF algorithm is used to detect image features and calculate the initial homography matrix. The feature descriptor of SURF is invariant to scale, orientation, and affine distortion, and partially invariant to illumination changes like SIFT. The high computation speed of SURF can improve the efficiency of 3D registration. In the tracing for registration procedure, L-K optical flow algorithm based on image pyramid is adopted which improves the efficiency further. Besides, we also adopt a features redetection strategy every 10 frames to avoid the accumulation of registration error. Experimental results demonstrate that the presented approach improves both the accuracy and real time capabilities of 3D registration, and it has great practicality.

6. ACKNOWLEDGMENTS

This paper is supported by the Program for New Century Excellent Talents in University (NCET) of China.

REFERENCES

[1] C. Geiger, T. Schmidt and J. Stocklein, "Rapid development of expressive AR applications", *2004 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004)*, (Germany), 2004, pp. 292-293.

[2] G. L. Hua and B. He, "A Registration Method Using Natural Features for Augmented Reality", *Journal of System Simulation*, No.19, 2007, pp:273-277.

[3] J. Chen, Y. T. Wang, Y. Liu and D. D. Weng, "System initialization algorithm based on SIFT key points for markerless augmented reality applications", *Infrared and Laser Engineering*. No.36, 2007, pp:949-953.

[4] X. Y. Li and D. Y. Chen, "Augmented *reality in e-commerce with markerless tracking", 2010 second IEEE International Conference on Information Management and Engineering (ICIME 2010)*. (China), 2010, pp:609-613.

[5] Z. J. Jiang and H. R. Yi, "An Image Pyramid-Based Feature Detection and Tracking Algorithm", *Geomatics and Information Science of Wuhan University*, No.32, 2007, pp 680-683.

[6] A. I. Comport, E. Marchand, M. Pressigout and F. Chaumette, "Real-time markerless tracking for augmented reality: the virtual visual servoing framework", *IEEE Transactions on Visualization and Computer Graphics*. No.12, 2006, pp: 615-628.

[7] M. L. Yuan, S. K. Ong and A. Y. C. Nee , "Registration using natural features for augmented reality systems", *IEEE Transactions on Visualization and Computer Graphic, No.*12, 2006, pp:569-580.

[8] H. Bay, T. Tuytelaars and G. L. "Van, SURF: Speeded up robust features", *2006 9th European Conference on Computer Vision (ECCV 2006)*, No.3951, 2006, pp:404-417.

[9] S. W. Yang, C. Wang and C. H. Chang, "RANSAC matching: Simultaneous registration and segmentation", *2010 IEEE International Conference on Robotics and Automation (ICRA 2010)*. (Taiwan), 2010, pp: 1905-1912.