第30卷 第3期　　电子科技大学学报　　Vol.30 No.3
2001年6月　　Journal of UEST of China　　Jun. 2001

Resolving Occlusion Problem in Augmented Reality

Shen Jie　Wu Yonghong　Liu Haowu

(College of Hydroelectric Eng., Sichuan University　Chengdu　610065)

Abstract A new approach for resolving occlusion problem in augmented reality is presented in this paper. The heart of this algorithm is to track at least four fiducial points from frame to frame that construct the affine frame, thus the occlusion contours between virtual and real objects can be reconstructed and reprojected. Further, each contour is labeled as being "behind" or "in front of" in the first two frames, depending on whether it is in front of or behind the virtual object, so the occlusion is resolved. Two typical examples show that the algorithm is feasible for resolving occlusion.

Key words occlusion; augmented reality; affine invariants; contours; reconstruction

Augmented reality (AR) is a technology by which user's view of the real world is augmented with additional information from a computer model, thus more information spaces are shown [1]. Generally, such systems only blend virtual imagery with real images and attempt to minimize registration errors. However, this method is not effective when occlusions exist between virtual and real objects. This occlusion problem could easily be solved on condition that the model of 3D scene is given. Since little is usually known about the real world to be augmented, it becomes challenging to resolve occlusion in augmented reality. Theoretically, a dense map may be inferred from a stereo pair, so the depth between virtual and real objects can be compared [2,3]. As a matter of fact, this method lacks accuracy, and is difficult to use.

In this paper, we present a contour-based approach with 3D reconstruction. As it is known, there are several levels of reconstruction, including the projective, affine, and Euclidean reconstruction. Different vision tasks require different levels of reconstruction. Only an affine reconstruction is required for fixation point tracking [4]. First, four or more fiducial points from frame to frame which construct an affine basis are tracked. Second, the key points of occluding contours between virtual and real objects may be specified interactively according to epipolar and other constraints in the first two frames. Finally, since the four fiducial points have established the affine basis, the occluding contours across frames can be affinely reconstructed and reprojected according to the image coordinates of key points in the first two frames, and thus the occlusion problem can be resolved.

Section 1 introduces the affine reconstruction algorithm. Section 2 applies these results to the occlusion problem, and presents two practical examples. Section 3 briefly discusses this algorithm. Limited by spaces, we have only talked about the major algorithm of resolving occlusion, not referred to the tracking problem of fiducial points in video images [5].

1 Reconstruction Based on Affine Structure

In the past several years, with the development of computer vision, new advances in 3D reconstruction have been made, and many new practical models have come into being. These models imply that when 3D location and calibration parameters of camera are unknown, through the recognition

---

Received December 15, 2000

of fiducial points in the scene and by means of affine representation, the structure of any point in 3D scene can be recovered. With these models, occlusion problem may be resolved.

1.1 Affine Structure Based on Four Points

Based on the results of Koenderink and van Doorn[6] , the recovery of 3D affine structure from multiple images may be allowed, for an affine camera. Let O, Y_1, Y_2, Y_3 be four noncoplanar points in a 3D world, and O', Y_1', Y_2', Y_3' be corresponding coordinates from the second camera position, an object point of interest Y with respect to the basis OY_1, OY_2, OY_3 is shown as follows

$$Y = O + \alpha OY_1 + \beta OY_2 + \gamma OY_3 \tag{1}$$

where α, β, γ are affine coordinates.

Under parallel projection, the viewing transformation between two scene-cameras can be represented by an arbitrary affine transformation [7-9]. Therefore, the affine coordinates of Y remain fixed under the viewing transformation. At the second camera position, the corresponding Y' of point Y may be written as

$$Y' = O' + \alpha O'Y_1' + \beta O'Y_2' + \gamma O'Y_3' \tag{2}$$

these affine coordinates are held under such a transformation, so α, β, γ are affine invariants. Since the depth is lost in image coordinates, we have a similar relation (using lower case)

$$y = o + \alpha oy_1 + \beta oy_2 + \gamma oy_3 \tag{3}$$

$$y' = o' + \alpha o'y_1' + \beta o'y_2' + \gamma o'y_3' \tag{4}$$

According to Eq.(3) and Eq.(4), if the coordinates of point Y in two images are known, the affine coordinates of point Y can be derived from the affine basis. At last, the position of point Y in any frame may be reprojected according to Eq.(3) or Eq. (4).

1.2 Location of any Point in the Second Frame

The result of stereo vision is that the three-dimensional location of a point in the environment can be derived from two images taken at different locations of a camera. The main question is to fix its location in the second frame, after the location of a point in the first frame having been determined. As it is known, if Y's projection is specified in one image, its projection in the second image must lie on a line satisfying the epipolar constraint. At the same time, through point collinearity or coplanarity constraint, the location of any point in the second frame can be determined.

To determine the locations of any point in two images, interactive method may be adopted. Once the locations of a point Y are determined in two images, its affine coordinate can be computed from Eq. (3) or Eq.(4), and thus the location of reprojection in other frame can be fixed.

As shown in Fig. 1, Y is a point of 3D space, and X_1, X_2, X_3 are coplanar with images x_1, x_2, x_3 and x_1', x_2', x_3' in the first and second images respectively. The epipolar plane defined by the optical centers O, O' and point Y intersects the plane Π which is determined by lines YO and YO' respectively. From the result of affine geometry, three points determine a plane affine transformation such that $x_i' = T x_i$, i=1,2,3 . This transformation is used to transfer the point y to $y_1'=Ty$, and so point y_1' on the epipolar line can be determined. Under affine transformation, all epipolar lines are parallel. The direction of any epipolar line is simply determined by making the global affine coordinate frame aligned with the first image [10]. If B is the projection matrix of the first image, there is a non-singular 3×3 matrix S satisfying

$$BS = \begin{bmatrix} 1 & 0 & 0 \\ \\ 0 & 1 & 0 \end{bmatrix}$$

A simple solution for S is

$$S=\begin{bmatrix} A_1^{-1} & -A_1^{-1}d_1 \\ 0 \quad 0 & 1 \end{bmatrix}$$

where $\boldsymbol{A}_1$ is 2×2 and $\boldsymbol{d}_1$ is 2×1. If $\boldsymbol{B}$ is divided into $\boldsymbol{B}=[\boldsymbol{A}_1 \mid \boldsymbol{d}_1]$, and $\boldsymbol{B}_i=[\boldsymbol{A}_i \mid \boldsymbol{d}_i]$, $\boldsymbol{d}_i$ is the direction of the epipolar line. To get $\boldsymbol{d}_i$, we must get each projection matrix post-multiplying matrix $\boldsymbol{S}$.

Once a point and the direction of epipolar line are determined, the epipolar line can be drawn in images. For point $\boldsymbol{Y}$ of 3D space, not only is there an epipolar constraint, but also there is collinearity or coplanarity constraint. Through these constraints, the location of point $\boldsymbol{Y}$ in the second image can be determined.

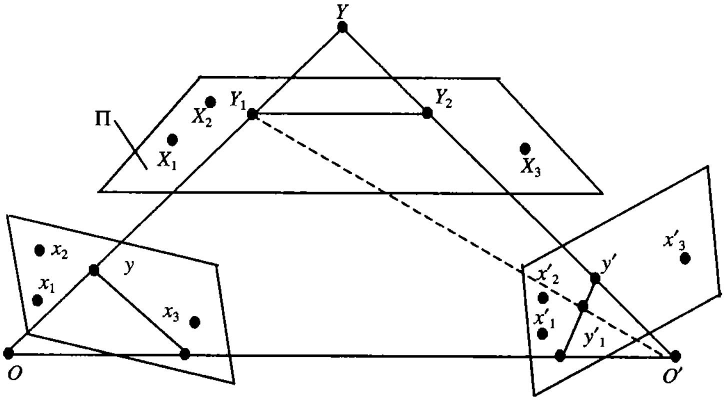


Fig.1 Epipolar geometry relationship

1.3 Resolving Occlusion in Augmented Reality

When a virtual object is added into the scene, the user needs to specify the relationship between them. For instance, we want to add the virtual house A behind the real house B (shown in Fig.2). In order to resolve occlusion, if it is feasible to track the contours consisted of key points 1, 2, 3, 4, the virtual house A can be drawn behind the contour. Since we have obtained the affine frame in every frame, when the image coordinates of four key points of the contour in two images are known, the reprojection of the contour can be drawn in each image. If the scene is complicated, we may label each contour point as being "behind" or "in front of", which depends on whether it is in front of or behind the virtual object.

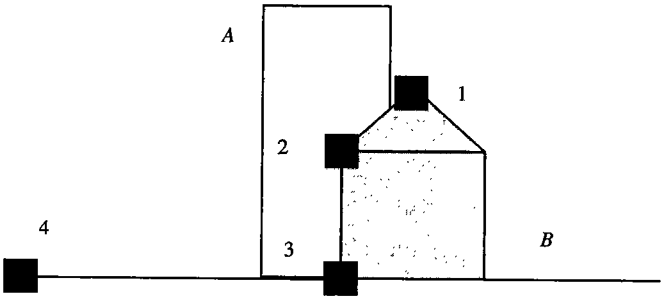


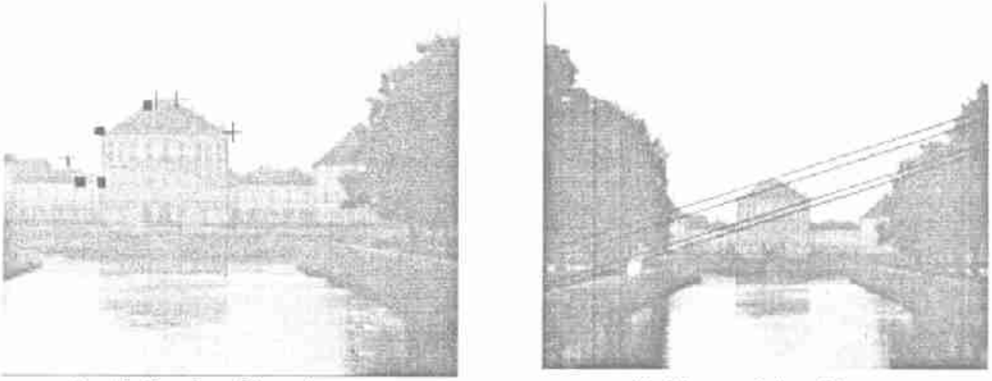
Fig.2 The occlusion between real and virtual objects

2 Examples

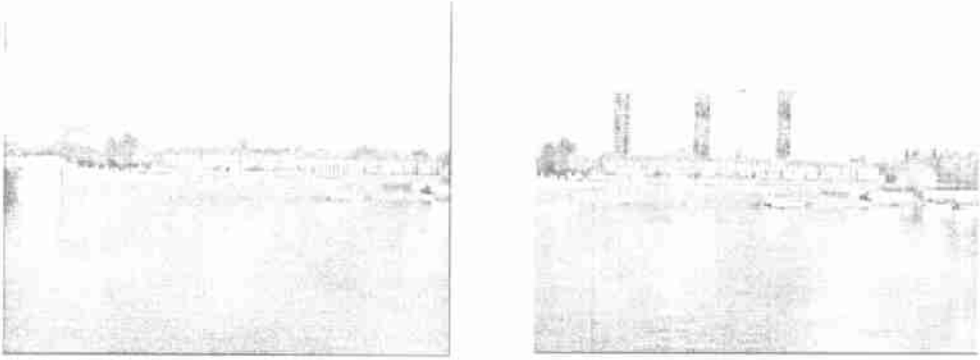
In order to demonstrate the effectiveness of this method, two typical examples were selected and briefly described in the following.

2.1 Example 1

The first example is to fuse the virtual building into real 3D scene. System hardware includes a digital camera and a Alpha-500 workstation. The software for this application is written in VC++. Fig.3 is the experiment environment. Now, suppose that a high building will be constructed in Fig.3a, the affine frame is constructed from the image coordinates of cross-center in Fig.3a, which are acquired by interactive method. Similarly, the image coordinates of key points on the occlusion contour are obtained in the first image (see four small black squares in Fig.3a). The epipolar lines of these points are computed and drawn in the second image (see four black lines in Fig.3b). Then, according to epipolar and collinearity or coplanarity constraints, the image coordinates of key points on the occlusion contour in Fig.3b are obtained by interactive method. Furthermore, according to affine reconstruction, the locations of the occlusion contour in other frames are determined. In Fig.3c and Fig.3d, the virtual building is fused into the 3D scene very well.

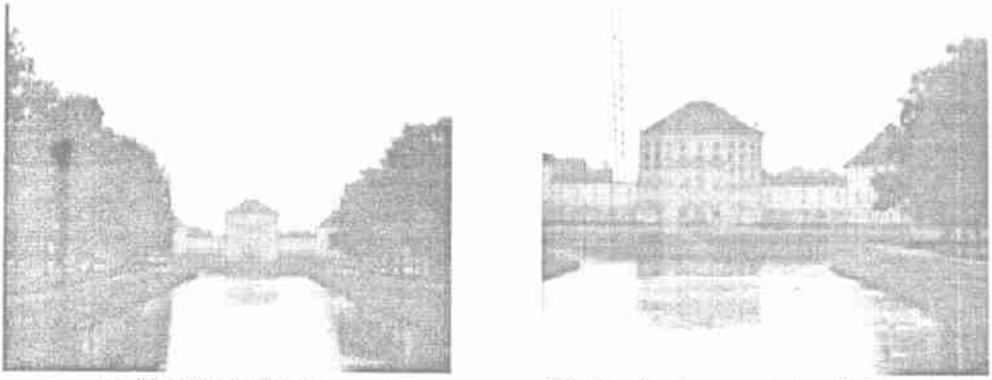


(a) The first view of the real scene (b) The second view of the scene



(c) The first view of augmented virtual building (d) The second view of augmented virtual building

Fig.3 Resolving occlusion with a simple occlusion contour



(a) The view of real scene (b) The view of augmented virtual building group

Fig.4 Resolving occlusion with a relatively complicated occlusion contour

2.2 Example 2

The second example (see Fig.4) is to fuse building group into a scenic spot. The process of resolving the occlusion problem is similar to that in example 1. The difference between these two examples is that the occlusion contour between virtual and real objects is relatively complicated. Of course, the more accurate the occlusion contours tracked are, the more key points on contours are needed.

3 Conclusion

This paper describes a contour-based approach with 3D reconstruction for dealing with the occlusion problem in augmented reality. Two examples in this paper demonstrate that this algorithm is feasible for augmented reality. To our knowledge, no methods can successfully handle all the problems in augmented reality. The algorithm in this paper has greatly enriched the tools for resolving the occlusion problem in augmented reality.

References

1 Azuma R T. A survey of augmented reality. Presence: Teleoperators and Virtual Environments, 1997,6(4): 355~385

2 Wloka M. Anderson B. Resolving occlusions in augmented reality. In Symposium on Interactive 3D Graphics Proceedings, New York, 1995: 5~12

3 Koch Reinhard. Automatic reconstruction of buildings from stereoscopic image sequences. In R J Hubbold and R Juan eds, Eurographics'93, Oxford, UK: Blackwell Publishers, 1993: 339~350

4 Reid I D, Murray D W. Active tracking of foveated feature clusters using affine structure. International Journal of Computer Vision, 1996,18(1): 41~60

5 Cho Y. Scalable fiducial-tracking augmented reality. [PhD thesis]. Los Angeles:University of Southern California, 1999

6 Koenderink J J, van Doorn A J. Affine structure from motion. J Opt Soc Am A,1991,8(2): 377~385

7 Ullman S, Basril R. Recognition by linear combinations of models. IEEE Trans on Pattern Anal and Mach Intell, 1991, 13(10): 992~1 006

8 Li Jiayan. Perpendicularity principle of projection in four-dimensional space. Journal of Electronic Science and Technology of China, 1989, 18(2): 177~183[梨家燕. 四维空间投影的垂直特性. 电子科技大学学报, 1989, 18(2): 177~183]

9 Zhang Fengli. Qin Zhiguang, Jing Wanjun. The research and realization of 3D graphic parallel projections. Journal of Electronic Science and Technology of China, 1994, 23(5):510~516[张凤荔, 秦志光, 敬万钧. 3D 图形平行投影的算法研究与实现. 电子科技大学学报, 1994, 23(5): 510~516]

10 Seitz S M , Dyer C R. Complete scene structure from four point correspondences. Proc Fifth Int'l Computer Vision, 1995: 330~337

(下转第249页)

5 Hu Haoquan, Yang Xianqing, Zhao Jiasheng. Research on mathematics models for electromagnetic interference prediction between radar systems. Journal of Electronic Science and Technology of China, 2001, 30(1): 37~40[胡皓全，杨显清，赵家升．雷达之间电磁干扰预测模型研究．电子科技大学学报，2001，30(1): 37~40]

Methods of Evaluating Electromagnetic Compatibility of Cluster Electronic Equipments

Yan Yijun
(Fifth Department, The 28th Electronic Research Institute of the Ministry of Information Industry　Nanjing　210014)
Yang Xianqing
(Dept. of Microwave Eng., UEST of China　Chengdu　610054)

Abstract The prediction of electromagnetic interference between the electronic equipments of a large information system is a complex problem. Furthermore, the prediction of the effect of interference on electromagnetic compatibility of the system is a random process, which causes much difficulty on practical application. Based on the results of electromagnetic interference analysis, three methods of evaluating the electromagnetic compatibility are presented. Evaluation model on certain electromagnetic environment is also discussed in this paper.

Key words electromagnetic compatibility;　system method;　cluster method;　electromagnetic environment;　electromagnetic interference

---

(上接第240页)

解决扩充现实中的遮挡问题

申　杰*　吴永红　刘浩吾
(四川大学水电学院　成都　610065)

【摘要】提出了一种新的解决扩充现实中遮挡问题的方法。该算法的核心是在每一帧图像中追踪至少四个基准点，构成仿射基，进而对虚拟物体与真实物体之间的遮挡轮廓进行重建和重投影，且在前两帧图像中交互指定这些阻挡轮廓相对于虚拟物体的前后位置，从而解决了扩充现实的阻挡问题。两个典型实例表明，此算法是可行的。

关 键 词　遮挡；　扩充现实；　仿射不变量；　轮廓；　重建

中图分类号　TP391

2000年12月15日收稿

* 男 29岁　在职博士生