

行政院國家科學委員會補助專題研究計畫成果報告

用於崁入虛擬物體於影像序列中的擴充實境技術

計畫類別： 個別型計畫 整合型計畫
計畫編號：NSC 89 - 2213 - E - 009 - 093 -
執行期間： 88 年 8 月 1 日至 89 年 7 月 31 日

計畫主持人：陳 稔 國立交通大學資工系教授
共同主持人：賈叢林 中正理工學院電機系教授

本成果報告包括以下應繳交之附件：

- 赴國外出差或研習心得報告一份
- 赴大陸地區出差或研習心得報告一份
- 出席國際學術會議心得報告及發表之論文各一份
- 國際合作研究計畫國外研究報告書一份

執行單位：交通大學資訊工程學系

中 華 民 國 八 十 九 年 九 月 十 日

行政院國家科學委員會專題研究計畫成果報告

用於嵌入虛擬物體於影像序列中的擴充實境技術

Calibration-Free Augmented Reality For Embedding Virtual Objects In A Real Image Sequence

計畫編號：NSC 89-2213-E-009-093

執行期限：88 年 8 月 1 日至 89 年 7 月 31 日

主持人：陳稔 國立交通大學資工系教授

協同主持人：賈叢林 中正理工學院電機系教授

E-mail: zchen@csie.nctu.edu.tw Tel:03-5731875

一、中文摘要

本計畫在一段現有影片中嵌入一件物品(稱虛擬物體),而使這融合後的影片如同原先就有該件物品存在而拍的。我們所發展之技術不同以往之處是我們不必假設攝影機在拍攝影片時已將相機在影片中的場景位置設定好,而事後要嵌入的人或物亦先規劃好其位置,然後再做嵌入工作。另外,我們也不限制相機的投影公式是某種近似公式(如 affine projection)。

為了逼真地將虛擬物體擺設到真實影像中,我們必須解決(一)空間的一致性(二)物理特性的一致性(三)真實場景和虛擬物體的互動三個問題。

在本計畫中,使用者能手拿攝影機拍一段影片回來後,便可以將虛擬物體放置到影片中。因為最後輸出的結果為 2D 影像,而空間一致性的要求並不須要計算出物體的 Euclidean space 3D 才有辦法達成,只要知道物體間次序(ordering)關係,便可以將虛擬物體正確的擺設到真實場景中,進而可以做重行投影。因此,我們提出利用投影幾何(Projective geometry),計算出實際場景在 projective space 下的三維座標,在此空間下,物體間的長度,角度,比例,平行性會失真,但次序是不變的,再利用 epipolar geometry 及物體 rigidity 的 constraints 便可將虛擬物體擺設到所要嵌入的影像中,進而計算出虛擬物體在 projective space 中的座標。如此,在其他影像中,只要計算出虛擬物體應投影的地方,並且計算此區域的 projective depth,比較虛擬物體和真實場景間的 depth 關係,便可以解決遮蔽(Occlusion)的問題。

關鍵詞：虛擬實境;影像三維資訊重建技術;虛擬物體的擺設;遮蔽問題。

Abstract

This project attempts to merge a 3D virtual object in an image sequence. The resultant augmented virtual object will look so natural that its appearance in the scene is seamless. Our method is different from

the existing methods in the literature in several aspects. First, we do not assume that the camera position in the scene is known in advance and that the position where the virtual object is to be placed is prearranged. Next, no approximate projection model of the camera such as an orthographic or affine one is assumed. Thus, our method is easy to use and it yields a more precise result.

For a realistic augmented reality result we shall consider three issues: (1) the spatial consistency, (2) the physical consistency, and (3) the interaction between the real scene and the virtual object.

In this project we plan to shot some outdoor or indoor scenes with a hand-held video recorder for a later combining of a virtual object which is constructed by a vision based technique with a photo-realistic texture. The final output is the augmented 2D scene images, so we do not need to use the Euclidean geometry to model the virtual object and real scene. Instead, we will work in the projective space, since we do not assume a calibrated camera. To begin with, we arrange a proper placement of the virtual object in two selected scene images using the epipolar and rigidity constraints and then derive the projective coordinates of the virtual object with respect to the scene. To solve the occlusion problem, we need to compute the projective depth of the part of the scene, which is overlapped with the virtual object, and compare the depth ordering between the scene and the virtual object.

Keywords: Virtual reality, Object placement, Occlusion.

二、緣由與目的

擴充實境(Augmented reality)[1]技術在探討如何將電腦圖學技術展現的影像逼真地嵌入到實際拍攝到的影像中,以供人們觀看。利用此一技術,我們可以將不存在的物體,或一般看不到的東西,加入所拍攝的影像中,如此可以豐富我們視覺資訊的接受。最常看到 Augmented reality 的展現,就屬影片工業,其稱為特效(Special effects)。從影片中,

我們可以看到慧星逼真的撞地球,看到恐龍再度復活,看到許許多多不存在的景物。但這些效果是要花許多人力物力慢慢的兜出來的。

為了逼真地將虛擬物體擺設到真實影像中,我們必須解決(一)空間的一致性(二)物理特性的一致性及(三)真實場景和虛擬物體的互動三個問題。第一個問題在於如何將物體擺設進去而不會有懸空,穿透別的真实物體,遮蔽的問題。其次則是在於打光的效果,重力的影響等等。最後則是碰撞等等的效果。而現今的研究重點多在如何達成空間一致性上。

目前國外有許多大學研究機構正致力研究 Augmented reality 並將其應用在不同領域上。在 MIT AI Lab[1], GE[1],及 UNC Chapel Hill[1],他們將此一技術應用在醫療診斷輔助上,整合 CT,MRI,超音波影像,醫生藉由 see-through HMD 可以同時看到患者身上腫瘤,胎兒的位置,避免一邊聽診一邊看照片的不便。在哥倫比亞大學[1],波音公司[1]則是利用 Augmented reality 來協助做機器的組裝及檢修。法國 INRIA[4]-[9]利用 Augmented reality 來做河邊,橋上燈光擺設的模擬。其他研究單位還有 CMU Robotics institute[3],及國內中研院[10]在做此相關研究。

前述的單位研究的方法多著重在校正好的相機,才能有效的將虛擬物體擺設到真實影像中,然在現實應用中,相機校正的工作常是煩雜的且不易做的。而在 Rochester 大學的 K. N. Kutulakos[11][12],試著將要相機校正工作的限制拿掉,他們假設相機投影模型為平行投影[15],因此利用兩張影像,可以建構出實際場景在 affine space 下的三維座標。而藉由操作者指定虛擬物體在二張影像的成像,可以計算出虛擬物體在此 affine space 的位置,如此可以逼真的將物體擺設到其他的影像中。然因其假設相機投影模型為平行投影,為一 approximation model,因此所計算出的三維資訊只能接近真實的,如此會影響電腦產生影像和實際拍攝影像的合成。

在本計畫中,我們希望使用者能手拿攝影機拍一段影片回來後,便可以將虛擬物體放置到影片中。因為最後輸出的結果為 2D 影像,而空間一致性的要求並不須要計算出物體的 Euclidean space 3D 才有辦法達成,只要知道物體間次序(ordering)關係,便可以將虛擬物體正確的擺設到真實場景中,進而可以做重行投影。因此,我們提出利用投影幾何[16]-[28](Projective geometry),便於計算出實際場景在 projective space 下的三維座標,在此空間下,物體間的長度,角度,比例,平行性會失真,但次序是不變的,再利用 epipolar geometry 及物體 rigidity 的 constraints 便可將虛擬物體擺設到所要嵌入的影像中,進而計算出虛擬物體在 projective space 中的座標。如此,在其他影像中,只要計算出虛擬物體應投影的地方,並且計算此區域的 projective depth,比較虛擬物體和真實場景間的 depth 關係,便可以解決遮蔽(Occlusion)的問題。

三、研究方法與成果

本計畫將採用電腦視覺技術來解決 Augmented reality 中虛擬物體的擺設(Object placement)及遮蔽(Occlusion)二個問題。而要處理上述二個問題,首先,我們必須能夠估測出現實場景的三維資訊,如此,才能指定虛擬物體要擺設在真實場景中的那個位置。在知道虛擬物體和真實場景的空間關係後,我們便能計算出虛擬物體和在真實場景所拍得的影像間的投影關係,如此便能將虛擬物體正確的投影在每一張影像中。而也因為知道虛擬物體和真實場景的空間關係,便可以用其間的景深關係來解決遮蔽問題。比較細節的方法及進行步驟如下:

步驟一: 影像三維資訊重建技術

我們擬採用投影幾何(Projective geometry)原理,發展出一套未校正相機三維場景重建的技術。其所牽涉到的工作有:

1. 影像中立體匹配(Stereo matching)的建立:由於影像的擷取是來自於連續的拍攝,因此若相機移動的速度不會很快,則兩張影像間的差異不會很大。我們偵測影像中基本的特徵,如角點(Corner),利用 correlation 方法,輔以空間關係,找出兩張影像間的點對應。
2. 兩張影像間投影幾何關係的推導:在兩張影像間,會存在一個 3×3 的矩陣,稱之為 Fundamental matrix, F 。此一關係即為此二張影像及其所對應之三維場景的 epipolar geometry。
3. 分解 Fundamental matrix 為二個投影矩陣 (Projection matrix), P_1 及 P_2 : Fundamental matrix 是由兩張影像的投影矩陣所合成。因此我們以第一張影像(二張影像選一即可)為參考座標,設定其投影矩陣 P_1 為 $[I|0]_{3 \times 4}$,則由 F 便可以分解出另一張影像的投影矩陣。而此二個投影矩陣分別為二張影像中,以第一張影像所定義出的參考投影空間(Projective space)的投影矩陣,其和 Euclidean space 中的投影矩陣將會相差一個轉換矩陣。
4. 投影空間中場景三維座標的估測:在得到兩張影像的投影矩陣後,只要給定兩張影像中的任一對點對應,利用三角定位方式,便可鎖定影像點對應在投影空間的三維點座標。

步驟二: 虛擬物體的擺設

欲將虛擬物體擺設到所有拍攝得的真實影像中,若是採用人工一張張的去調整修飾將是很煩雜的工作。我們發展出一套物件擺設的工具,協助使用者將虛擬物體擺設到其所選取的兩張影像中。透過之前談到的未校正相機三維場景重建技術,可以決定虛擬物體要落在真實場景中的那一個位置,以及其和其他實際物體的空間關係,如此,虛擬物體在其他影像的投影便可以藉由此影像和參考影像之間的 Fundamental matrix 關係,自動的產生出來。為協助使用者方便的擺設物體到影像中,我們完成下列工作:

1. 發展一套快速的 3D 投影程式。此系統允許使用者調整虛擬物體投影中的虛擬相機參數,並

將其疊合在真實影像中。使用者可以即時的看到虛擬物體在影像上的可能投影,並選擇一個其所認為最像的樣子。

2. 引入 epipolar geometry 及 rigidity constraints 到第二張影像的擺設工作中：當我們決定物體在第一張影像的成像後,便利用兩張影像間的 epipolar geometry constraint,使虛擬物體上各點座標在第二張影像的投影被限制在其對應的 epipolar line 上,以加快使用者決定虛擬物體在第二張的投影。
3. 由物體在二張影像的投影,決定其在 Projective space 的三維位置。

步驟三： 遮蔽(Occlusion)問題的解決

在得知虛擬物體在以某張影像為參考影像所定義出的 Projective space 三維座標後,其他影像便可以與此一參考影像做關連,找出其間的 Fundamental matrix,並設定參考影像的投影矩陣為 $[I|0]_{3 \times 4}$ 。如此,另一個求得的投影矩陣均是在同一個 Projective space 中,也因此,我們可以將之前算出的虛擬物體 Projective space 的三維座標重新投影,而獲得其在真實影像的成像。再根據這部份真實及虛擬物體相重疊地方,計算此區域真實影像所對應的三維點座標,並將這些點座標的景深和虛擬物體相比,便可判斷出誰在前,誰在後,解決了遮蔽的問題。

在實作過程中,為了簡化整個擴充實境系統之操作流程,我們提供一個目前在視窗作業系統下,常見之精靈模式,透過我們的 AR 操作精靈,使用者只需一步步依據畫面所顯示之指示進行,便能完成整個擴充實境之程序。在擺設物體進入第一張影像之即時繪製虛擬物體畫面,我們先透過 OpenGL 函式庫繪製影像於精靈畫面中,再將此結果疊合於第一張影像之上,讓使用者觀看其擺設之整體效果。

圖 1 至圖 4 是我們將檯燈擺進真實畫面中的結果圖。圖 5 至圖 8 為我們將一磁片盒擺進真實畫面中的結果圖,其中物體繪製方面包含了該磁片盒表面的材質資料,使得繪製出來的效果更逼真,而在光源方面,我們也透過光源設定,來儘量符合真實場景中之情形,最後並依據所設定之光源資訊及虛擬物體(磁片盒)所在平面,而繪製出磁片盒在桌面上的陰影圖,如此,整個完成的最後結果圖。圖 9 至圖 12 為多重擺設之結果圖,透過我們的擴充實境系統,將一奶粉罐接著擺設進入先前已經擺設磁碟盒的影像中,其中會發生遮蔽問題,經由我們遮蔽演算法自動加以處理之後,正確地決定出前後遮蔽的關係。

四、結論

在本計劃中,我們採用電腦視覺技術來解決 Augmented Reality 中虛擬物體的擺設(Object Placement)及遮蔽(Occlusion)二個問題。首先,我們透過虛擬物體之擺設程序,計算出虛擬物體和影像中真實場景間的關係,之後,透過未校正相機重建技術計算出來的影像之投影矩陣,便能將虛

擬物體正確的投影在每一張影像中。而也因為知道虛擬物體和真實場景的空間關係,便可以用其間的景深關係來解決遮蔽問題。

五、參考文獻

- [1] R. Azuma, "A Survey of Augmented Reality", Presence: Tele-operation and Virtual Environment, vol. 6, no. 4, 355-385, Aug. 1997.
- [2] P. Jancene, et al., "RES: computing the interactions between real and virtual objects in video sequences", Proceedings of Networked Realities Conference, 1995.
- [3] Takeo Kanade, et al., "Video-Rate Z Keying: A New Method for Merging Image for Merging Images", Technical Report 95-38, The Robotics Institute in CMU, 1995.
- [4] M.-O. Berger. G. Simon, S. Petitjean and B. Wrobel-Dautcourt, "Mixing Synthesis and Video Images of Outdoor Environments: Application to the Bridges of Paris", Proceedings of the 13th ICPR, 90-94, Vienna, 1996.
- [5] M.-O. Berger and G. Simon. "Robust Image Composition Algorithms for Augmented Reality", Proc. of the 3rd ACCV, Hong Kong.
- [6] M.-O. Berger. "Resolving Occlusion in Augmented Reality: a Contour Based Approach without 3D Reconstruction", Proc. of CVPR'97, 91-96, Puerto Rico, USA, 1997.
- [7] G. Simon and M.-O. Berger, "A two-stage robust statistical method for temporal registration from features of various type", Proc. of ICCV, 261-266, Bombay, India, 1998.
- [8] E. Boyer and M.-O. Berger, "Smooth Surface Reconstruction from Image Sequences", Proc. of ICIP'95, Washington, 1995.
- [9] G. Simon, V. Lepetit and M.-O. Berger, "Computer Vision Methods for Registration: Mixing 3D Knowledge and 2D Correspondences for Accurate Image Composition", International Workshop on Augmented Reality, San Francisco, 1998.
- [10] Chu-Song Chen, et al., "Integrating Virtual Objects into Real Images for Augmented Reality", ACM Symposium on Virtual Reality Software and Technology, Taipei, November 1998.
- [11] K. N. Kutulakos, J. R. Vallino, "Calibration-Free Augmented Reality", IEEE Transactions on Visualization and Computer Graphics, vol. 4, no. 1, 1-19, 1998.
- [12] K. N. Kutulakos and J. R. Vallino, "Affine Object Representations for Calibration-Free Augmented Reality", Proceedings of 1996 IEEE Virtual Reality Annual International Symposium, 25-36.
- [13] C. Chevier, "Handling Interactions between Real and Virtual Worlds", Computer Graphics

- International.
- [14] S. Feiner, B. MacIntyre, T. Hollerer, "A Touring Machine: Prototyping 3D Mobile Augmented Reality Systems for Exploring the Urban Environment", Proceedings of ISWC'97, pp. 74-81, Cambridge, MA, Oct. 13-14, 1997.
- [15] Jan J. Koenderink and Andrea J. van Doorn, "Affine structure from motion," J. Opt. Soc. Am. A., Vol. 8, No. 2, pp. 377-385, Feb. 1991.
- [16] Roger Mohr and Emmanuel Arbogast, "It can be done without camera calibration," Pattern Recognition Letters, Vol. 12, No. 1, pp. 39-43, Jan. 1991.
- [17] R. Szeliski, "Video Mosaics for Virtual Environments," IEEE Computer Graphics and Applications, Vol. 6, pp. 22-30, Mar. 1996.
- [18] T. Moons, L. Van Gool, M. Proesmans, and E. Pauwels, "Affine Reconstruction from Perspective Image Pairs with a Relative Object-Camera Translation in Between," IEEE Tran. PAMI, vol. 18, no. 1, pp. 77-82, Jan. 1996.
- [19] Q. T. Luong and T. Vieville, "Canonic Representations for the Geometries of Multiple Projective Views," in ECCV'94, pp. 589-599, 1994.
- [20] O. D. Faugeras, "What can be seen in three dimensions with an uncalibrated stereo rig?" in ECCV'92, pp. 563-578, 1992.
- [21] O. D. Faugeras and Luc Robert, "What can two images tell us about a third one?" International Journal of Computer Vision, 18, pp. 5-19, 1996.
- [22] Q. T. Luong and O. D. Faugeras, "A stability analysis of the fundamental matrix," in Proceeding of ECCV'94, pp. 577-588, 1994.
- [23] Zhengyou Zhang et al., "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry," Artificial Intelligence, 78, pp.87-119, 1995.
- [24] A. Shashua, "Projective Structure from Uncalibrated Images: Structure from Motion and Recognition," IEEE Tran. PAMI, Vol. 16, No. 8, pp.778-790, Aug. 1994.
- [25] A. Shashua, "Projective Depth: A Geometric Invariant For 3D Reconstruction From Two Perspective/Orthographic Views For Visual Recognition," in Proc. of Fourth ICCV, pp. 583-590.
- [26] R. Hartley and R. Gupta, "Computing Matched-epipolar Projections," in Proceeding of IEEE Conf. On CVPR, pp. 549-555, 1993.
- [27] R. Hartley, "Projective Reconstruction an Invariants from Multiple Images," IEEE Tran. PAMI, vol. 16, no. 10, pp. 1036-1041, Oct. 1994.
- [28] R. Hartley, "Euclidean Reconstruction from Uncalibrated Views," Application of Invariance in Computer Vision, pp. 237-256, 1994.



圖 1 擺設檯燈進入影像結果



圖 2 擺設檯燈進入影像結果



圖 3 擺設檯燈進入影像結果

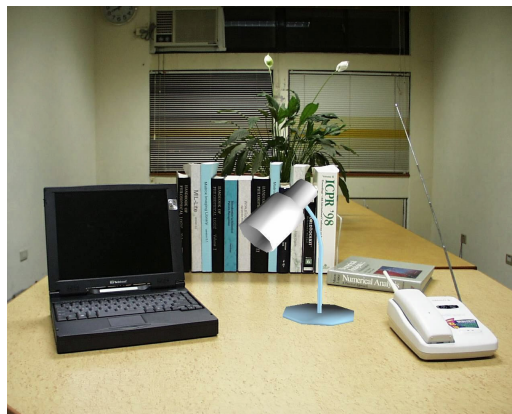


圖 4 擺設檯燈進入影像結果



圖 5 擺設磁片盒進入影像結果



圖 6 擺設磁片盒進入影像結果



圖 7 擺設磁片盒進入影像結果



圖 8 擺設磁片盒進入影像結果



圖 9 繼續擺奶粉罐進入影像結果



圖 10 繼續擺奶粉罐進入影像結果



圖 11 繼續擺奶粉罐進入影像結果



圖 12 繼續擺奶粉罐進入影像結果