

泛用型動態虛擬實境操控與運動復健輔助系統研發 成果報告(第一年度)

子計劃三: 動態 VR 運動復健輔助系統之互動式實景影像設計與 研發

A Design of Supporting Dynamic System with Interactive Real-image-based Scenes

計畫編號: NSC 90-2213-E-009-106

執行期限: 90 年 8 月 1 日至 91 年 7 月 31 日

主持人: 林昇甫 國立交通大學電機與控制研究所

計畫參與人員: 蘇建焜、張家銘 國立交通大學電機與控制研究所

一、中文摘要

本子計畫主要任務是應用立體實景影像成像技術和虛擬實境顯像技術，令使用者有身歷其境和賞心悅目的感受，以增加使用者的興趣，而達到提昇訓練的目的。本子計畫分三個年度進行，第一年度係配合動態跑步機操控與互動式實景影像展示系統之研發與設計，第二年度係配合動態划船操控與互動式實景影像展示系統之研發與設計，最後一個年度係配合動態騎馬操控與互動式實景影像展示系統之研發與設計。

基本上對每一個系統均是針對其個別特性和需求，進行下列項目之研究:首先確立適合主題之場景，並作實景記錄。其次利用實景影像建立環境模型與研究實景影像銜接合成技術，並由實景影像序列建立立體實景影像且應用於虛擬實境中。同時我們也進行操作者與捲動播放互動關係之研究，探討操作者前進方向與速度和六軸平台之運動姿態間之關係。最後，針對大量的影像資料研究其有效存取和播放的裝置。

總括來說，我們最終希望提供使用者一個栩栩如生的視覺環境，且操作者可與環境即時互動的影像系統。對於使用者而言，他可以選擇路徑與速率快慢，使用者的方向和速率變化由六軸平台透過適當感應器讀取，然後傳送至本子系統，本系統就會呈現出使用者所應看到的 180 度景觀

在使用者前方，並且把地形起伏的資料傳送給六軸平台以修正六軸平台的位置和提供適當的力回饋給使用者。

本報告的第一部份是敘述這一個階段研究計畫的目的，當中將從功能的角度，再將本子計畫做更精細的劃分。其次在第二部份將詳細說明研究發展及進行步驟，當中將介紹本系統之規格與使用之技術。最後的部份則是下年度的展望。

關鍵詞: 虛擬實境、六軸平台、影像系統、即時性、互動性。

Abstract

The goal of this subproject is to provide a vivid immersing virtual scene, which is synthesized by 3D image-based rendering technologies and virtual reality technologies. This project will last three years. During the first year, we are going to support the design and development of the "interactive dynamic jogging machine." In the second year, our task is to support the design and development of the "interactive dynamic rowing machine." And, in the last year, supporting the design and development of the "interactive dynamic rodeo machine."

Generally speaking, we do the following things for the specific characteristic and needs of each system. At first, we select a proper environment and record the images. Next, construct the environment model and study the image processing technologies

(image morphing, 3-D layer presumption method, ... etc.) by using image-based methods. Furthermore, reconstruct the 3D environment and the virtual world. At the same time, studying the interactive relation among the user's speed, the scenes, and the state of the Stewart platform is an important work. Finally, since the amount of data is very large, we need some fast and effective method for image storage and playing.

To sum up, the purpose of this subproject is to offer a user a vivid visual environment with real time and interactive requirement. The user can choose the path and control his jogging speed, and the sensors on the Stewart platform will pick up all the information of the user. After that, the Stewart platform transfer the user data to the image system, and the image system shows the user a 180 degree scene in front of him. At the same time, the image system transfer the environment data (altitude, slop, ...) to the Stewart platform for adjusting the Stewart platform and producing the feedback force.

Keywords: Virtual reality, Image-Based, Stewart platform, Interactive, Real-time.

二、本階段研究計劃之目的及成果

本研究報告為國科會整合性研究計劃「泛用型虛擬實境操控與運動復健輔助系統研發」中之子計劃三「動態 VR 運動復健輔助系統之互動式實景影像設計與研發」的第一年期末成果報告。

在第一年度之內，本計劃已達成上列預定的 (1)固定路線，固定速率前進，全平面地形，和 (2)固定路線，速率可變，全平面地形二項目標。圖(一)所示是真實影像擷取系統，為降低成本及符合一般大眾的真正情形所以採用三台市售的低價位 DV (SONY TRV 18) 架設於汽車頂，而這三台攝影機之距離和光軸方向是可以測量得知的，所以事後可以根据這些數據來進行 warping、mosaic、或 view morphing 等影像處理的工做，使得合成的畫面更加逼真。將 DV 架設於汽車之上除了運動方便之外，亦可利用汽車之避震系統減少攝影機之震動。

取景之後，再把三台攝影機所錄得的影像透過 *IEEE 1394* 介面以 *AVI* 格式儲存於電腦硬碟之中成為基本素材。然後使用 *AVIedit* (網路下載之 share ware) 把 *AVI* 影片轉為一連串的靜態影像 (BMP 檔)。圖(二)所示即是左、中、右三台攝影機在某一時間所攝得的三個畫面。假設三台攝影機之相關位置在拍攝之中不會改變，因此只要根據一組左、中、右的照片來研究左、中和中、右的相關點即可据此進行所有時間點照片接合。除了剪裁接合之外，因為使用的 DV 採用自動調整光圈的模式，所以從圖(二)可以看出來，三張影像的亮度有明顯的差異，所以必需把三張影像的亮度調成一致，接合之影像才不會有不自然的亮度變化。圖(三)是經過指定相關點，然後進行裁剪、調整亮度然後接合而成的廣角影像。可以明顯看出亮度的變化已經做了大幅度的修正。

經過裁剪、調整、接合等動作，把左、中、右三張影像接合成一張廣角影像，重覆這個動作可以得到一連串的廣角靜態影像，最後利用 *BMP2AVI*，把一連串的廣角靜態影像編輯成為一個 *AVI* 檔，再使用播放軟體，由 PC 送出訊號，透過單槍投影機投射於螢幕上，即可得到以實景為基礎的虛擬實境影像。關於速度的調整可以從兩方面著手，一種是在我們產生廣角靜態影像時，另外一種方法是在轉換成 *AVI* 檔時，透過 *Frame rate* 的改變可以離線控制速率。

本系統規格如下(1)影像擷取：三台 *SONY TRV18* 數位攝影機。(2)影像格式：靜態用 *BMP* 格式，動態儲存成 *AVI* 格式。(3)影像計算：一部 PC (*Pentium 4*, *1.8GHz*)。(4)影像播放：以單槍投影機投影至螢幕。

三、研究發展及進行步驟

本子計劃是採用以真實影像為基礎的方法，這種方法的優點是影像逼真，其主要缺點是如果要記錄的環境結複雜，則需要耗用相當大的記憶體儲存空間才能夠完全記錄下這個背景。M. Hirose[1]等人以 8 部攝影機來拍攝所需的真實影像序列，我們則決定使用三台一般用途的 DV 來拍

攝，如此可以簡化器材，降低成本，當然如果要達到相同的影像品質，我們需要使用更多的影像處理技術。影像格式目前一般的 DV 大多使用 *AVI* [2]格式和 *MPEG1* 格式，為了方便快捷編碼和解碼，我們選擇用 *AVI* 格式，來儲存動態影像。

三部攝影機同步取得動態影像之後，我們必需把這些相關的動態影像加以剪輯、接合，以及加上必要的影像處理技巧，最後再分割成三個畫面分別存在三台 PC 硬碟之中。在攝影的當時也必需同時記錄下地形的變化，這部份最簡單的方法是使用足夠精確的「全球衛星定位系統」或者利用立體影像重建的技巧，但是後者需要複雜的計算以及較長的時間。誠如上述所言，本子計畫需要使用大量的影像處理技巧，以下將對主要的部份加以介紹。

1、*AVI*格式:[2]

*AVI*格式是由美國微軟公司(*Microsoft*)所發展的動態影像儲存格式，它的原義是"Audio Video Interface"，也就是說一個 *AVI* 檔案中影像和語音的區塊是互相交錯存放的。這種動態儲存格式是目前 PC 系統中最常見的，它也是一種特殊的 *RIFF*(Resource Interchange File Format)

2、Elastic Registration(彈性註冊方法)[5][6]

當兩張影像要接合時，我們必需先找出分別位於兩張圖上的相關點。一般的註冊方法只適用於被拍攝物體沒有太大的失真或變形，其中以小波轉換的方式可以得到相當好的效果。但是通常在不同攝影角度所拍得的同一物體，往往伴有相當程度的幾何變化，這種情況就不適合使用一般應用於無形變的註冊方法[6]，而必需採用「彈性註冊法(Elastic Registration)」。*H. Maitre* 和 *Y. Wu*[5]使用動態規劃演算法以及自回歸模型發展出一種彈性註冊的方法可用於處理兩張幾何結構發生變化的影像註冊問題。

3、根本矩陣(Fundamental Matrice)之計算[3]

假設對同一物體在不同角度、位置分別攝得影像 I_0 和 I_1 ， P_0 是 I_0 上的一點而 P_1 是 P_0 在 I_1 上相對應的點，若 F 是 I_0 和 I_1 的根本矩陣，則 P_0 ， P_1 和 F 滿足下列方程式：

$$P_0^T F P_1 = 0 \quad ,$$

此處

$$P_0 = \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \quad , \quad P_1 = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \quad . \quad (1)$$

所以根本矩陣隱含了 I_0 和 I_1 內景物的幾何結構資訊，而這正是電腦視覺的關鍵，所以計算出相對應兩影像之根本矩陣一直是許多人研究的重要題目。如果把 F 看成是 3 乘 3 的矩陣，則方程式(1)可表為：

$$\begin{bmatrix} x_0 & y_0 & 1 \end{bmatrix} \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = 0 \quad (2)$$

或

$$x_0 x_1 f_{11} + y_0 x_1 f_{21} + x_1 f_{31} + x_0 y_1 f_{12} + y_0 y_1 f_{22} + y_1 f_{32} + x_0 f_{13} + y_0 f_{23} + f_{33} = 0 \quad (3)$$

在理想的情況下我們可以由測量 9 組 P_1 和 P_0 得出一個齊次線性系統，當中包含了 9 個類似方程式(3)的方程式，即

$$\begin{bmatrix} x_0^{(1)} x_1^{(1)} & y_0^{(1)} x_1^{(1)} & x_1^{(1)} & x_0^{(1)} y_1^{(1)} & y_1^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_0^{(9)} x_1^{(9)} & y_0^{(9)} x_1^{(9)} & x_1^{(9)} & x_0^{(9)} y_1^{(9)} & y_1^{(9)} \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \\ f_{33} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4)$$

方程式(4)要有不為 0 的解，其係數矩陣之秩數(rank)必需小於 9，也就是行列式值為 0。但在實際的測量中由於測量的

誤差，係數矩陣的行列式值不為 0，所以無法求出有意義的非零解的 F 。在 f_{33} 不為 0 的情況下系統(4)可簡化為

$$\begin{bmatrix} x_0^{(1)} x_1^{(1)} & y_0^{(1)} x_1^{(1)} & x_1^{(1)} & x_0^{(1)} y_1^{(1)} & y_1^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_0^{(8)} x_1^{(8)} & y_0^{(8)} x_1^{(8)} & x_1^{(8)} & x_0^{(8)} y_1^{(8)} & y_1^{(8)} \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \end{bmatrix} = - \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \quad (5)$$

此時只要 8 組 $P_0^{(i)}$ 和 $P_1^{(i)}$ 即可解出 F 。但是當 f_{33} 很小時所求出的根本矩陣就會因誤差太大而失去參考價值。

正因為根本矩陣如此重要又不易正確求出，所以一直到今日都有人不斷地提出更好的方法來估計根本矩陣。我們通常把這許多求根本矩陣的方法歸納為三大類：(1)線性的方法(Linear Methods) (2)疊代的方法(Iterative Methods) 和(3)強健性的方法(Robust Methods)。實驗結果顯示出，如果參考點適當地落在影像當中，而相關性的問題又獲得解決，則線性的方法可以達到相當好的效果。疊代法可處理點位置含高斯雜訊的情形，但這種方法對於不在影像當中的點會有困難。強健法則可以應付位置的矛盾與錯誤的配對。一般人的看法是使用特徵分析(Eigen Analysis)的正交最小平方法(Orthogonal Least-Square)會比傳統的最小平方(Least-Square)法有更好的結果。但當把矩陣限制在 rank 為 2 且移除不在兩影像中的點時最小中間值平方方法(Least Median Square)是所有方法中估計結果最正確的。

4. View Morphing

以影像為基礎的(Image-Based)虛擬實境系統，如果要儲存的影像太多則會有記憶體不足的問題產生。解決此問題的方法之一是僅拍攝部份必需的場景，其餘不同

角度、位置所見的場景再以 View Morphing[4]的方式由現有的影像產生，如此可以省下大量的記憶體空間，但需要付出的成本是計算時間，以及可能降低解析度的問題。View Morphing 是由 Steven Seitz 和 Charles R. Dyer 所提出的影像處理技術。以圖(四)為例，左邊的影像是照像機從左側拍攝所得，右邊的影像是照像機從右側拍攝得到的，我們可以利用這左、右兩張照片加上 View Morphing 的技術，產生由正中央拍攝時所應有的照片。

(a) Prewarp :

利用投影變換(Projective transformation)把既有的兩張影像(I_0 、 I_1)轉換成兩張光軸平行的影像(\hat{I}_0 、 \hat{I}_1)。

(b) Morph:

其次使用線性內差法產生所要的中間影像 \hat{I}_s 。

(c) Postwarp:

最後再用投影變換把中間影像 \hat{I}_s 轉變為所要的 I_s 。

5. Layered Depth Images:

如何利用 2D 的圖片建構出具有 3D 效果的圖片，也是本子計劃的重點之一，在諸多方法中，比較有效率的方法稱為 Layered Depth Images (LDI)[7]。

Layered Depth Images (LDI)是先將每一張影像中的每個 pixel 除了其原有資訊外再添加了深度資料，再由好幾張單層的 depth image 來投射到我們想要的視角所看到的平面上。

在 LDI 方法中。假如有兩個以上的 pixel 投射到同一個位置，我們便必須比較它們的深度值，若深度的差異超過一個上限，我們便增加一層新的 layer 來儲存；否則便以取平均的方式來儲存深度。有了這些 layer 的資訊後便可將 2D 的圖片以較具立體化的感覺呈現給使用者。

四、結論與未來展望

真實、即時、互動是本子計劃所要追求的三大目標，首先依照本報告所規劃的進度，配合動態跑步機拍攝完成(1)固定路線、固定速率前進全平面地形之場景和(2)

固定路線、速率可變全平面地形的場景。以目前的處理方式(即播放 *AVI* 影片),因為影片是拍攝實際場景,所以自然地有”真實”的感覺而且也可以在離線的情況去模擬跑步者速率的改變,算是可以做到離線狀況的互動,但無法做到即時的互動。為了達到即時的目標,必需變更架構,這將會是未來兩年最大的挑戰。

關於場景製作有許多仍有改善的空間。首先是顏色的調整,因為不同三台攝影機的取景除了亮度會不同,其顏色亦會有差異,故需加以調整。第二,由於攝影機之光軸不平行,所攝影的物體會有形變,另外同樣的物體成像大小亦不同,所以會出現兩張照片(左、中或中、右)無法完全接合的情況,這個問題可以利用 *Warping* 或 *View Morphing* 的問題加以解決。第三,可以加入背景音樂以增加臨場感,這在 *AVI* 的格式中可以很容易做到。第四,配合 *Projection Matrix* 的應用可以改變視線(View Point)營造出上坡,下坡所看到的景物情況。

總結第一年度的做法,它的優點是可以用最經濟的方法來製作栩栩如生的場景,並且達到某種程度的互動,其主要缺點是”即時性”的問題,這也是未來兩年計畫的核心。

第二年我們將更進一步做到路線,速率地形其中兩者可變的情況,而在最後一年去除所有限制,達到路線、速率、地形皆可即時互動的要求。此時由於要傳送、計算的資料量均非常龐大,單純以軟體的方法來進行將會很吃力,所以需要藉助硬體系統的改進,軟、硬體相輔相成來完成本計畫的最終目標。

五、參考文獻

- [1] Yuichi Ohta and Hideyuki Tamura, *Mixed Reality- Merging Real and Virtual Worlds*, Springer-Verlag, pp. 183-197, 1999.
- [2] John F. McGowan, *AVI Overview*,
<http://www.jmcgowan.com/>.
- [3] J. Salvi, X. Armangue and J. Page, “*A Survey addressing the Fundamental Matrix Problem*,”
IEEE
- [4] Steven M. Seitz and Charles R. Dyer, “View

Morphing,”

<http://www.cs.wisc.edu/~dyer/vision.html/>.

- [5] Henri Maitre and Yifeng Wu, “*A Dynamic Programming Algorithm for Elastic Registration of Distorted Pictures Based on Autoregressive Model*,” IEEE trans on Acoustic, Speech, and Signal Processing, Vol. 37, No. 2, pp. 288-297, February 1989.
- [6] Raymond J. Althof, Macro G. J. Wind, and James T. Dobbins, III, “*A Rapid and Automatic Image Registration Algorithm with Subpixel Accuracy*,” IEEE trans on Medical Imaging, Vol. 16, No. 3, pp.308-316, June 1997.
- [7] Jonathan Shade, Steven Gortler, Li-wei He, Richard Szeliski, “*Layered Depth Images*,” *SIGGRAPH98*, Computer Graphics proceedings, Annual Conference Series, pp231-242, July 19-24,1998.



圖(一)：影像擷取系統



(a)

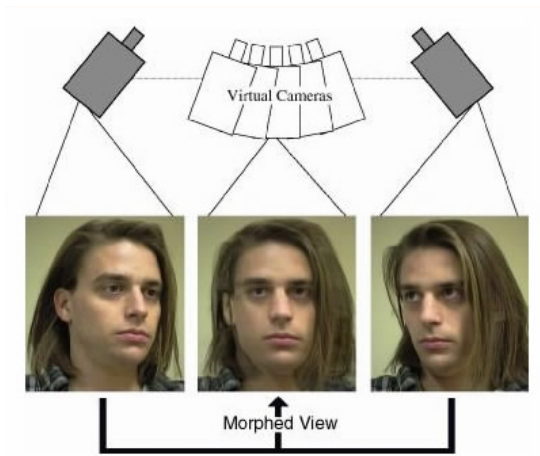
(b)

(c)

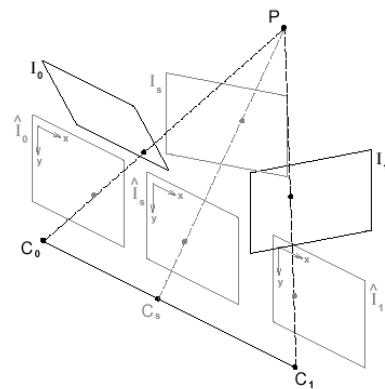
圖(二)：(a)左側攝影機影像、(b)中央攝影機影像、(c)右側攝影機影像、



圖(三)：接合並做亮度調整之廣角影像



圖(四)：View Morphing



圖(五)：Warping