

行政院國家科學委員會專題研究計畫成果報告

泛用型動態虛擬實境操控與運動復健輔助系統研發

計畫編號：NSC90-2213-E-009-103

執行期限：90年08月01日至91年07月31日

主持人：林進燈 教授 交通大學電控所

同主持人：李祖添、林昇甫、張志永、莊家峰教授 交通大學電控所

摘要

本計畫的目標是發展一套運動復健訓練的輔助控制系統，此系統主要是在一個由電腦操控的六軸動作平台上建構走步機或跑步機等運動訓練輔助機制，配合聲光效果建立之虛擬實境，提供使用者一個運動與訓練環境。其主要特色在於運用感測資訊及操作者的選擇，以電腦預測出使用者的體能與身心狀態之型態(Pattern)，進而改變輔助訓練的內容與場景，期可以鼓勵使用者，達到進行運動訓練的積極目的。此系統不但可作為一般人士的運動、健身器材，尤其適合提供給身心殘障人士一個更人性化及具輔助性的復健空間與設備，並達到職業重建的目的。

本年度此計畫共以五個子計畫完成各自負責的項目，以下就分別陳述各子計畫的進度狀況：

子計畫一：動態 VR 運動復健輔助系統之行為轉換與即時模擬研究

一、計畫概述

在本計畫研發的系統中，操控員透過操控界面下達指令，以操控系統中的虛擬的運動訓練的輔助設備或載具。此操控命令將被輸入所模擬之設備或載具的精確物理模型中，以求得真實情況下系統的反應。這些反應將透過本子計畫的行為轉換與控制模組，而由六軸運動平台、力迴饋模組及虛擬實境顯示器表現出來，以讓操控員獲得身歷其境的感受。這整個系統中，因為行為轉換模組

的設計扮演實際環境裡物體運動狀況以及虛擬環境中操控員身體感受之間橋樑，如何讓操控員有身歷其境的感受，有賴於在動態模擬系統發展過程中，適當的人(操控員)機(動態模擬器)溝通界面及迴饋學習技術，以獲得最符合人類感覺的行為轉換模式。本子計畫只要將相關之運動行為提供給顯示子系統即可。Stewart Platform 為一具有六自由度運動能力之機械平台，但其運動範圍受到其工作空間的限制，有些動作是無法達成的，因此為了在其有限空間下模擬出騎馬、飛機、輪船等交通工具時之運動感覺。所以一般是以沖淡演算法(Washout Algorithm)架構出對模擬器的力的感受。在第一年的研究中，我們探討了如下的四個課題：沖淡演算法級行為轉換演算法、函數型類神經網路、訓練者修正的加強式學習、以及模糊控制器穩定設計與探討。我們簡述於下。

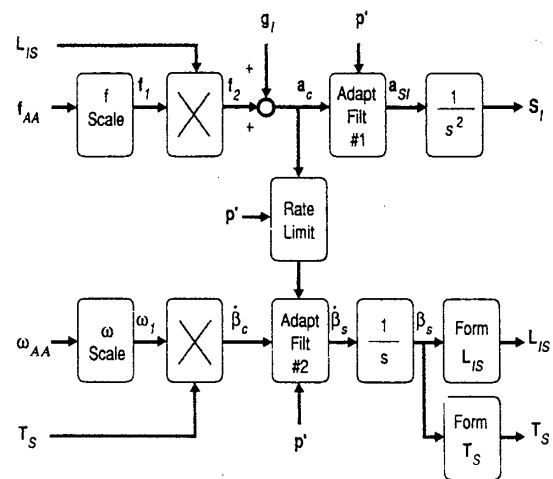
二、沖淡演算法級行為轉換演算法

本子計畫主要在於當虛擬系統接收到控制命令，設法產生應有的運動行為描述以及在模擬運動平台的運動行為，以達到虛擬實境的目的。目前文獻上對於如是的轉換是以 Washout 演算法或稱 Washout Filter 的方式來設計。Washout Filter 的輸入為由數學運動模型所求得的比例力與角速度。所謂的比例力(Specific Force)是指單位質量所受的非重力力。由向量的觀點來看的話，比例力即是慣性加速度與重力加速度之差，因此比例力可定義為 $\vec{f} = \vec{a} - \vec{g}$ 。在

文獻中也有提出調適式的 Washout filter 如圖一為一調適式之 Washout filter，其概念是根據線上取得的資訊去調整 Filter 中的增益。其主要是將平台的運動和被模擬之運動之間的差異及平台特性的考量來，架構出一性能指標。並經由一最佳化的運算，來找出增益調適的法則，由於指標函數及增益間之關係複雜及即時的需要。一般都是以最陡下降法來產生。如欲追隨之運動已知，則也可利用最佳控制的概念，來直接尋找最佳之線性 filter。此一調適演算法主要的功能，即是要減低錯誤動作方式及使平台儘量維持在原點。而在如圖一的架構中，其斜傾處理的計算也有所改變。首先其所利用的信號本身就不一樣。在調適式架中由於是針對慣性座標軸，因此有許多的運動的斜傾反應，也進入角度運動環路中了。因此其在調適的考慮上也變得非常複雜。

固定平台的虛擬實境主要是靠影像部分來達成，當運動為直線定速時，如是的模擬是相當不錯的，可是當運動路徑是多變的時候，在轉彎加速或上下起伏時，在固定平台上的感受就差很多。若使用動作平台時，由於虛擬系統能設法產生一些相對應的動作反應，使操作者能感受到類似的感受力，此時在輔以影像的虛擬，確實能使虛擬實境達到更逼真的效果。可是當動作平台產生誤動作時，整個系統的感受將是很不好的。因此如何去設計好的動作，便是虛擬實境系統中不可或缺。此一部分的設計一般統稱為 motion cue。在我們的國科會研究中(89-90 年度)，我們取得美新科技一組影片及對應的平台操控命令列，目前我們在探討以資料庫的方式來建立的 motion cue 的對應法則。有關 motion cue 方面，我們也分析了一段雲霄飛車的影片，企圖來了解一般在 motion cue 方面的設計概念。可是這樣的資料庫再使用上可能會很難，因為 generalization 的能力是相當缺乏的。而且此資料庫的型態是以函數為本質而不是點。在我們的研

究中我們將研究利用 Function Artificial Neural Network(FANN)來做 Motion cue 資料的建立。而在第一年的研究中，由於其運動體就是操作者，所以其所需要的動作平台的模擬反應較小，而且是可預測的。因此在研究上我們將考慮在我們先前研究中運動轉換資料庫的設計上，運動轉換資料庫主要的概念是針對已知的運動行程(如動態電影院)，直接設計平台動作的方式，由於此部分有些專家的經驗，因此針對如是的資料庫，我們可以累積經驗使之成為與被模擬機構及平台機構無關的運動轉換法則，而為了能達到知識粹取的過程，利用類神經方式的建模是值得探討的，除了其能利用 recognition without definition 的特質，其將知識參數化，也可提供未來進一步學習修正的依據。



圖一：調適式之 Washout filter

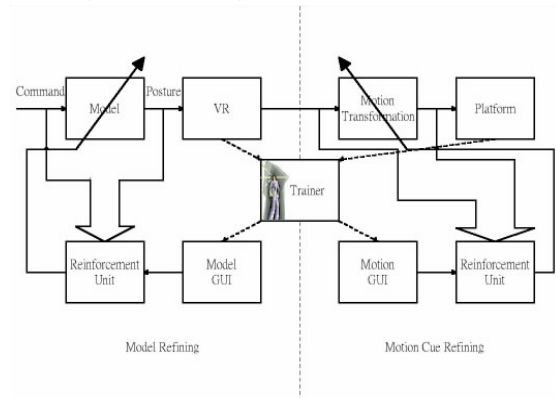
三、函數型類神經網路 (FFANN)

而在 motion cue 的記錄方面，我們考慮使用函數型類神經網路來記錄 motion cue。在傳統的類神經網路，通常以點對點的方式在時域當中進行系統的建構。而所建構的網路以適應性網路進行網路鍵結值的學習。近來，一種函數型類神經網路架構已發表出來，函數型網路架構使用數學的方式進行函數對函數的輸出輸入對訓練網路之鍵結值。因為函數型網路使用平行處理之數學計

算，可以快速估計出約略所需的輸出鍵結值。由於它優異的函數近似能力，使得我們可以將其運用在系統之頻域響應當中。藉由前置的離散訊號處理以頻率域替代傳統的時間域學習，並期望在噪音的環境之下由在頻率域作業的優點，而產生抑噪的功能；第二步，並將頻率分解，由 FANN 的近似能力來學習，進而追蹤期望訊號；第三步輸出並轉換為時間域。這個架構可結合訊號處理的優點（高速、抑噪），而達到我們想要的結果。目前 FANN 正處於萌芽時期，其中尚有以上一些缺點必須解決，對於週期性的輸入便無法辨識，還有當期望輸出上下限落差大，FANN 也無法學習，及鍵結值的學習速度也在改進的範圍之中。希望藉由我們提出的架構來改善以上所提之缺點，並提升神經網路的效能。FANN 以函數為輸出及輸入的類神經網路。FANN 的架構及其學習效果等，都是本研究所要探討的。長久以來，頻域的脈衝特性使得傳統性的點對點網路難以學習，在這當中還有頻域之中的複數特性（大小及相位）需要比較精確的比例才能使系統之輸出正確還原。在研究中，我們以函數型網路為網路主體，運用其良好函數近似能力去趨近受建構系統之頻率響應；進行系統分析之後，加上複數型學習網路機制，適當地調整複數鍵結值並且保持相位的正確。為使學習可以達到所期望的輸出，列舉三種複數網路學習機制，其中兩種傳統學習法並不能正確還原受建構系統時域輸出。因此我們依照系統所需還原及複數網路特性去打造另一種學習法，結合原有之函數型神經網路主體，為頻域中之混合式函數型類神經網路。Algorithm 4 則為 Algorithm 3 的另一種變行。經由模擬的結果可以明顯驗證函數型網路在頻域展現出比時域更優良的性能，加以學習機制的參與網路，頻域中混合式函數型網路表現出速度快、高近似度、以及強健度的效果。

四、訓練者修正的加強式學習

在本子計畫的研究中，我們也將從事有關訓練者修正的加強式學習，主要的兩個方向是在建模方面，另一則是 motion cue 的部分，主要的觀念乃是在虛擬實境上，主要是要人的感受。因此當操作者對於系統有認為不好的地方，我們希望能將這一類的感受用於修正系統，對建模部分而言，主要是對反應的不真實性，加以修正，若是能建立這樣的系統，則針對建模部分就可以只要建一個簡單大致的模式，而後透過加強式學習來使得模式反應更逼真，如是未來對不同的系統開發可以不需要每一次都要找一個所要開發系統的建模專家，而只要有一對該系統反應狀況較熟悉的人就可以了。同樣的，在 motion cue 設計的部分，也是如此，只要一感受靈敏的人，針對系統反應的評估，即可設計出較精準而感受逼真的動態模擬系統。



圖二：加強式學習系統架構

在加強式的研究中，初步的架構如圖二上，訓練者在運動平台上，針對平台的反應來修正而修正分為部分，其一為 Model 部分，而另一為 motion cue 部分。在做法上，我們要研究訓練者的評估在加強式學習中的效能。而系統對訓練者其評估的不確定性，也是研究所需要析的。這是因為對同樣的狀況，在不同的時候，由於人類感受評估的不精確性，其可能會有不同的評估值。而同時，針對訓練者評估息的定義，也不必然和事先設定的概念一致。因此如何的調適評估信息的歸屬函數，及其在加強式學習中的學習法則，都需要深入的探討與

研究。而目前我們加強式學習的架構是以 SONFIN 的方式來使得每一網路的法則數不需要那麼多，可是由於 SONFIN 的學習在使用 BP 下較慢，而使用 LS 方法其所欲之輸出不知道，這些在研究中我們都將加以探討。而在內部加強式信號的產生部分，目前我們是以傳統的 Temporal Difference (TD) 預估的方式來進行。由於最近的研究大都使用遺傳演算法，我們也會考慮使用，在未來的研究中我們將更進一步的探討分析其結果，同時也將納入類似 washout filter 的概念以及段落分割之學習方面。最後，由於如上二種加強式學習。在訓練者的感受上似乎有不可分的現象。目前的概念是先分別探討與學習，以便個別系統能正確做加強式學習，若在分析中能正確分割兩系統效應。則可據以做加強式信號回饋的依據，若不能分割，則可考慮將二系統直接合併。由於加強式學習本來就是可用於較複雜，而不容易數學化的系統而當系統參數化後，其利用 trial-and-error 的特質，也是可以架構出加強式學習系統的。

五、模糊控制器穩定設計與探討

In traditional, the stability of a fuzzy system can be assured by finding a common positive definite matrix in a Lyapunov function. That matrix and the corresponding controller can be simultaneously found by solving a set of linear matrix inequalities (LMIs). However, LMI-based fuzzy static output-feedback gains are not obtained by traditional approaches for nonlinear systems with uncertainties or with time delays in the literatures. In this paper, a new approach will be proposed to stabilize nonlinear interval time-delay systems via static output feedback. This approach does not need to find a common positive definite matrix to guarantee the system stability. The verification process is via computing the matrix measure for each separate

subsystem. In this paper, we shall derive the stability conditions based on the matrix measures of the system matrices of the subsystems of a fuzzy model. We derive a new sufficient condition to ensure the stability of the following T-S fuzzy model

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \sum_{i=1}^r h_i(\mathbf{y}(t))(\mathbf{A}_i \mathbf{x}(t) + \mathbf{D}_i \mathbf{x}(t-t)) \\ \mathbf{y}(t) &= \sum_{i=1}^r h_i(\mathbf{y}(t))(\mathbf{C}_i \mathbf{x}(t))\end{aligned}$$

The static output feedback gains \mathbf{F}_j , $j=1,2,\dots,r$, satisfy the following conditions

$$\gamma_2(\mathbf{A}_i + \mathbf{B}_i \mathbf{F}_j \mathbf{C}_k) < \chi_i, \quad i, j, k = 1, 2, \dots, r.$$

if and only if

$$\begin{aligned}(\mathbf{B}_i)_\perp (\mathbf{A}_i + \mathbf{A}_i^T - 2\chi_i \mathbf{I})(\mathbf{B}_i^T)_\perp &< \mathbf{0}, \quad i = 1, 2, \dots, r, \\ \text{and } (\mathbf{C}_i^T)_\perp (\mathbf{A}_i + \mathbf{A}_i^T - 2\chi_i \mathbf{I})(\mathbf{C}_i)_\perp &< \mathbf{0}, \quad i = 1, 2, \dots, r.\end{aligned}$$

In this research, we have proposed a new robust fuzzy static output feedback controller design methodology for nonlinear interval time-delay systems with norm-bounded uncertainties and time delays. Finding an admissible solution to the matrix measure assignment problem can solve the problem of stabilizing controller design via static output-feedback for nonlinear interval time-delay systems represented by a T-S fuzzy model. The designed controller can indeed tolerate norm-bounded parametric uncertainties and against the effect of time delays. In the paper, we have shown that the matrix measure assignment problem is equivalent to an LMI feasibility problem. A sufficient condition for the existence of the fuzzy static output-feedback gain is obtained. Our approach does not need to find a common positive definite matrix for all rules and the verification of stability is very easy. Simulation results have verified and confirmed the effectiveness of the new approach in controlling a nonlinear interval time-delay system.

子計畫二：動態 VR 運動復健輔助系統之智慧型感測與控制

一、計畫概述

本計畫為「泛用型動態虛擬實境操控與運動復健輔助系統研發」整合計畫之子計畫二，研發重點在於電動六軸動作平台之設計與分析、運動訓練輔助機制之感測與控制及操作者運動狀況之偵測與分析。以之前本團隊所開發之油壓式六軸動作平台的經驗為基礎，在本子計畫裡，我們將繼續鑽研以馬達驅動(電動式)之六軸動作平台的設計與分析，以為整個動態運動訓練輔助系統的根基平台。我們也將進行平台的硬體機構分析並解決運動學、奇異點、工作空間等問題，以達位置、速度及加速度控制等目的。本子計畫的第二個重點在設計、開發運動訓練輔助機制，以架設在動作平台上，而構成即時動態運動訓練系統的整合機構。我們將朝向泛用型機構設計，以涵括如跑步、划船、騎馬等運動訓練項目。我們也將研發此機構之感測與控制技術。本計畫的第三個研究主題是操作者運動狀況之偵測與分析，藉由非侵入式檢測方式，瞭解運動員身心狀況(如心跳、運動速度)，以適當控制平台及運動訓練輔助機制，達到輔助訓練效果。本子計畫也將發展即時計算環境的軟硬體，以達整體系統即時控制效果。

二、目前研究進度：

針對本子計畫所將陸續完成的三大研究主題，在今年度本子計畫目前朝向電動六軸動作平台的開發與分析、平台嵌入式即時硬體單板之設計與發展、VR跑步場景之開發、VR 投影技術之探討等。以下就分別針對此幾個方向，來加以說明目前進度狀況。

A. 電動六軸動作平台的開發與分析

在此部分的研究上，由於是要從無到有，所以今年度我們從電動伺服馬達的選定與控制、控制架構的設計、單軸平台的載重分析等，都是目前的研究重心，以下分述如下：

◆ 電動伺服馬達的選定

為得知該採用哪款的伺服馬達，我們假定所欲建構的電動六軸平台載重為1000公斤，再乘以2倍之安全係數得單軸所需載重為470公斤。經計算結果得知所需負載轉矩為 4.25 Nm 與慣性矩為 $18.6\text{ Kg}\cdot\text{cm}^2$ 。最後查閱三菱MELSERVO-J2_SUPER規格書，選定對應之馬達類型如圖三所示，其為HC_SFS_153B (2000rpm/min)。



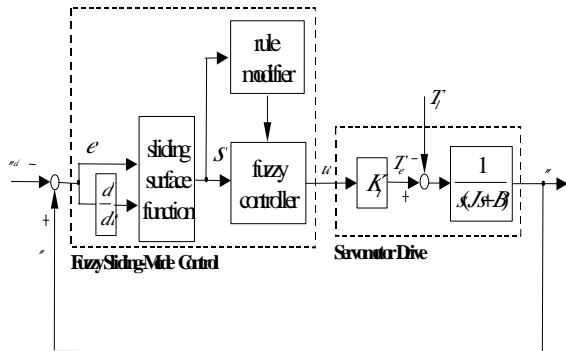
圖三：三菱 HC_SFS_153B 伺服馬達

◆ 電動伺服馬達的控制

在傳統的控制理論，控制器的設計都是基於數學描述和系統方程式模型，但是在現實世界中，當系統越來越複雜時就很難以數學模型來描述，所以要再使用傳統的控制理論去解決此類問題就顯得極為困難。近年來陸續有研究者將模糊控制器(Fuzzy Logic Control)運用到各種自動控制系統，使控制器加入了日常生活中的語意特質，其中的模糊規則為匯集專家的經驗和操控法則。六軸平台因其數學模式十分難以獲得，故利用傳統的控制理論進行設計控制器時將需花費很多時間調整控制器參數才能獲得不錯的系統響應。為了解決此問題，我們一開始選用利用模糊控制器為基本架構來進行設計。

此外我們還進一步提出模糊滑動模式控制(Fuzzy Sliding-Mode Control)設計方法，我們可藉由滑動表面(Sliding Surface)的提出使模糊規則數目得以減少，以利於設計與分析工作。這裡我們

使用模糊滑動模式控制後，將原先模糊控制所需的 49 條規則減少至 7 條規則。圖四展示所用之模糊滑動控制伺服馬達系統。



圖四：模糊滑動控制伺服馬達系統

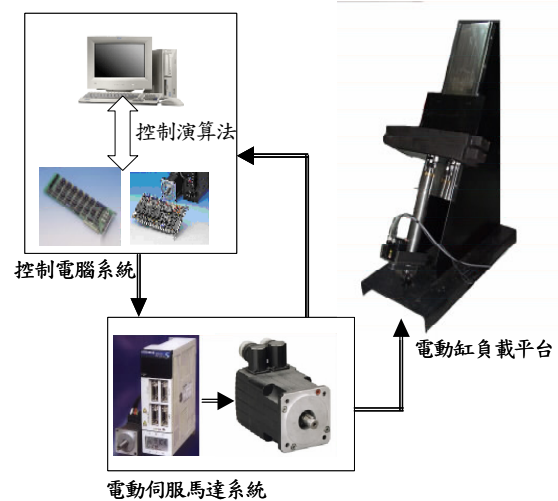
◆ 單軸電動缸載重測試

我們所發展之單軸電動缸控制系統架構圖，如圖五所示，控制演算法則經由控制電腦執行後，經過數位/類比伺服控制卡送出命令電流至驅動電路模組，由驅動電路器送出驅動信號至電動伺服馬達控制馬達正反轉與轉動角度，再經由皮帶帶動整個單軸電動缸測試平台轉換成直線型位置，最後檢測電動伺服馬達的實際位置，作成位置迴授至控制電腦以利閉迴路控制法則的執行。整個實驗結果為驗證所設計控制器之系統強健性，我們提供兩個測試條件如下：

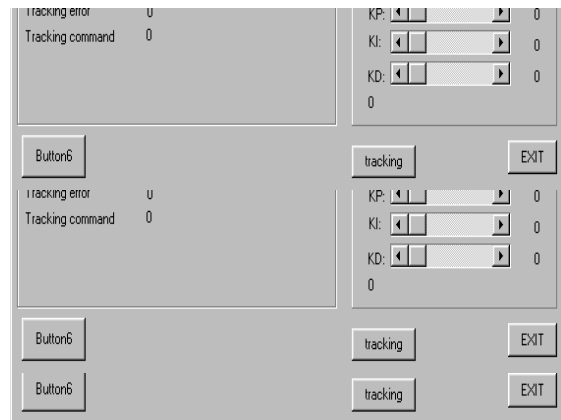
- (1) 測試條件一（空載條件）：將馬達轉動直接帶動整個電動缸。
- (2) 測試條件二（負載條件）：在電動缸平台上外加 375Kg 的負載。

這裡我們曾用業界中常用的 PID 控制器來控制整個系統，其實驗結果發現利用 PID 控制器時，不管測試條件一或二均並無法有效地精確追蹤到我們所下之控制命令。為了解決此問題，利用所設計之模糊滑動控制器 (Fuzzy Sliding-Mode Control) 做比較，其實驗結果發現利用模糊滑動控制器時，不管測試條件一或二均可精確地追蹤到我們所下之控制命令。此外為了方便隨時監控電動缸之控制情狀，本子題也發展了一適合且簡易的人機界面，透過這個人機

界面可以讓使用者更容易修改控制參數、儲存資料及察看結果，如圖六所示。



圖五：單軸電動缸控制系統架構圖



圖六：單軸電動缸之人機界面執行畫面

B. 嵌入式即時硬體單板之設計與發展

一般控制系統中大多採用 DSP 或 8 位元的單晶片微處理器，然而由於其無法支援作業系統的核心架構，所以用其低階語言編譯出來的控制程式，修改程序非常繁複，使系統發展性大為受限。因此我們採用嵌入式控制系統為核心，以 StrongARM 作為中央處理器，其可支援 open source Linux 作業系統。而開放式的程式碼不但可自由下載，並可依自己的需要修改原始碼並重新編譯，大大增進了整個系統發展應用軟體的效率。

目前我們採用以 Intel StrongARM 處理器為核心的 LART 嵌入式硬體，其

中 LART 是由全名 Linux Advanced Radio Terminal 的字首組合而來，為一套消耗不到 1W 功率即可達到 250MIPS 高效能的嵌入式系統硬體套件，其包括一塊主板、週邊裝置擴充板與網路卡，經由 LART 嵌入式硬體模組搭配 ARMLinux 嵌入式作業系統來取代舊有的工業電腦控制方式，此優點不僅可使控制系統體積大幅減少，同時也便於在作業系統上發展新的控制程式與未來控制程式發展更具彈性，而嵌入式系統週邊支援 TCP/IP 網路，使得先前發展的另一塊 Stand Alone 單板-DSP 控制卡無法連接 TCP/IP 網路的困境獲得解決。

接下來依序介紹在此嵌入式單板所發展的一些主題與成果：1. ADDA 傳輸介面。2. 嵌入式動態模擬控制系統。3. 整合成果與比較。

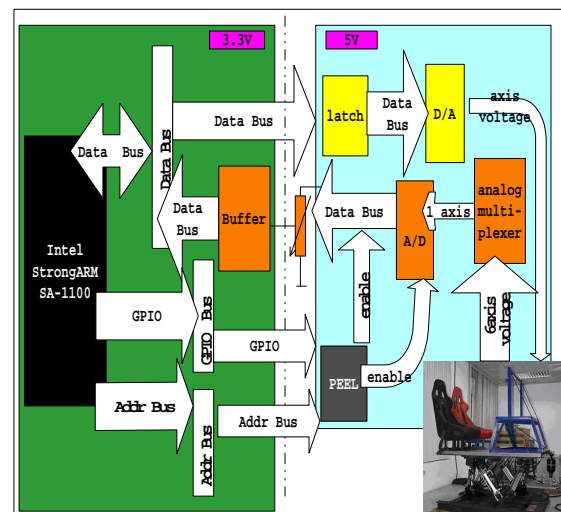
◆ ADDA 傳輸介面

圖五為 ADDA 電路系統方塊圖，由於六軸動作平台是利用類比電壓訊號來控制油壓伺服閥以決定軸伸長量，所以要利用 LART 嵌入式單板控制六軸平台，必須擴充設計六組的類比數位轉換的介面，讓 LART 的數位訊號與平台的類比訊號間做資料的傳遞。

整個 ADDA 電路系統需要提供 5V、+12V、-12V 的工作電壓，而 LART 主板的工作電壓則是在較低的 3.3V，在 5V 和 3.3V 間的訊號傳遞必須加入一個電壓準位轉換的機制才不會造成電壓過大系統損毀，或是電壓過小資料流失的問題。圖七中虛線右半部代表 LART 主板之工作電壓 3.3V，虛線左半部為 ADDA 電路系統之工作電壓 5V。考慮當 LART 訊號送到 D/A 轉換器時，因為 3.3V 到達 5V 系統之 V_{IH} ，資料門鎖器(Latch)會把 3.3V 的輸入數位訊號電壓拉到 5V 電壓準位送到 D/A 轉換器；AD 迴授訊號傳回 LART 時，則是將數位訊號用可變電阻分壓降到 3.3V 再由系統讀回。

數位轉類比的訊號是由六組 D/A 轉換器分別控制六軸電壓訊號，所以在 D/A

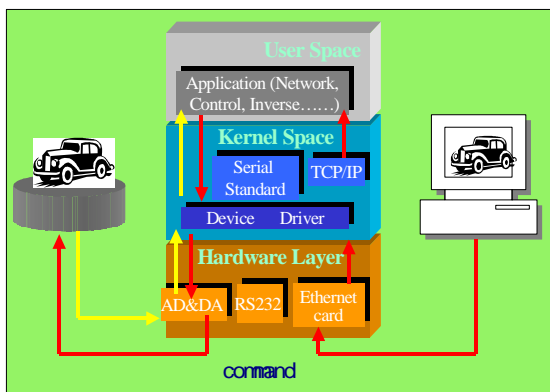
轉換器各需加一組獨立的資料門鎖器(Data Latch)來鎖住各軸所需的轉換資料。在類比轉數位電路裡，先用一顆類比多工器搭配 SA1100 上的三個 I/O port 來選擇讓第幾軸的類比訊號輸入轉換，因此只需一顆類比轉數位轉換器即可完成六軸電壓轉換動作。另外在這必須考慮的是，A/D 轉換完的輸出 12-bit 數位資料需經由 CPU 的 Data Bus 讀回系統處理，而 CPU 的 Data Bus 乃雙向匯流排，為了避免 CPU 的 Data Bus 在輸入與輸出的值之間沒有協調好互相衝突，所以在 A/D 轉換器和 LART 的 Data Bus 間加上一組三態緩衝器(Tri-State Buffer)，在系統尚未讀取資料前，AD 電路會被系統視為一個高組抗元件。



圖七：ADDA 電路架構圖

◆ 嵌入式動態模擬控制系統

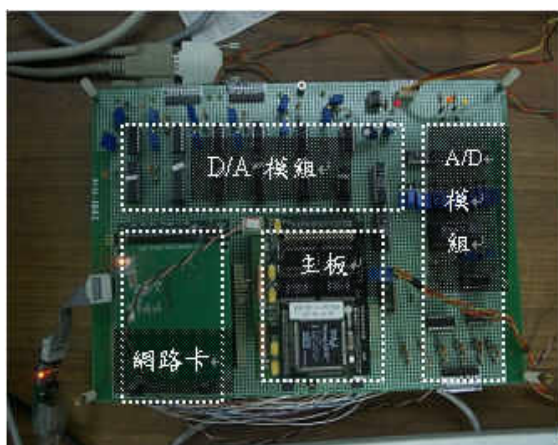
嵌入式動態模擬控制系統的發展主要分成三層的巨觀架構，包括作業系統控管的核心層(Kernel Space)、應用層(User Space)和硬體裝置所屬的硬體層，在數位類比介面卡硬體撰寫完成、網路軟硬機制設置完備後，接著的工作就是整合數位類比介面卡驅動程式並在應用層發展網路連線程式與各種平台控制程式(包括平台之姿態與腳長轉換的 Inverse Kinematics、發展模糊位置控制器、奇異點狀況的規避等)。



圖八：軟硬體整合系統架構

◆ 整合成果與比較

圖九展示所完成之嵌入式單板實體，而圖十則是進一步比較以前所用之 PC、DSP 單板以及目前所發展的 LART 嵌入式單板之優缺點。



圖九：LART 嵌入式硬體單板

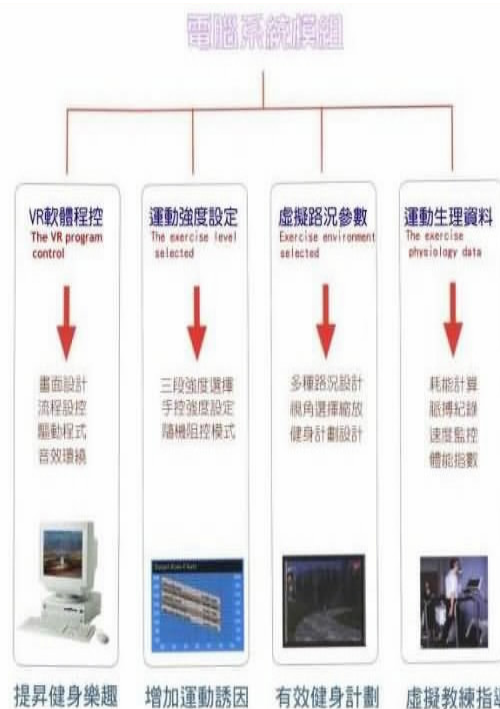
控制硬體	優點	缺點
工業電腦	◆多種作業系統支援 ◆豐富的通訊介面	◆體積龐大 ◆價格昂貴
DSP 控制卡	◆較小的體積 ◆電路全為自行設計有完整電路圖與 DSP 原始程式碼	◆與平台只能靠 RS232 傳輸 ◆程式全用組語發展，移植性不高 ◆記憶體容量受限
LART 嵌入式控制硬體	◆豐富的通訊介面 ◆功能強大的 ARM Linux 作業系統 ◆完整的硬體電路圖與自己設計的 ADDA 介面卡	◆記憶體容量受限 ◆無法直接在單板上編譯程式

圖十：控制硬體比較 C. VR 跑步場景之開發

隨著電腦科技的快速進步，虛擬實

境(VR)這個名詞已從當初人們對它的陌生發展演變而成現今炙手可熱的熱門技術。市場上也出現愈來愈多發展相關的軟、硬體，或者是利用此技術製作發展的商品。而運動器材與虛擬實境的搭配應用更是一項完美的結合，藉由虛擬實境的聲光影像使消費者身歷其境，再搭配軟體對於運動生理學及娛樂效果之輔助功能，讓室內健身運動不再是枯燥無味之運動，而是充滿著新奇、樂趣與資訊的一項新興運動。

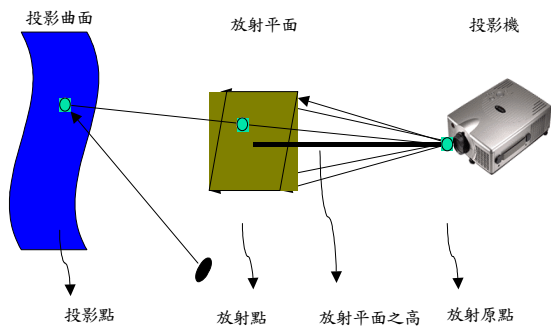
這裡此部分由於需要周邊軟、硬體的適當搭配，故除了基於之前 VR 場景開發的功力，積極開發出適當的跑步場景外，更在另一個子計畫目標裡，我們將拍攝真實外面影像，以融入 VR 畫面中，而增加真實感。而在與周邊系統的搭配上，除了在平台上適當搭建跑步機，我們更結合運動強度的設定、虛擬路況參觀、運動生理資料等，以增加使用者的興趣及提高此系統各方面的功能(如圖十一所示)。



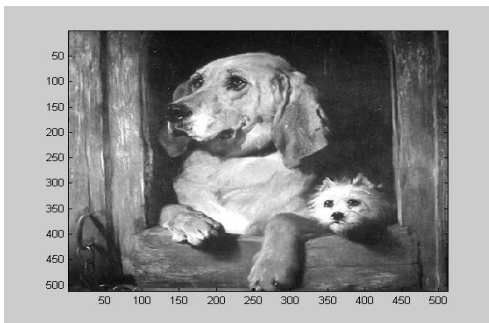
圖十一：VR 跑步場景整合系統示意圖 D. VR 投影技術之探討

這個部分技術的探討主要是為了在

影像投射的呈現上，不會受限於場地、投影機等。故我們首要目標是希望能夠依照場地布置方式改變投影機的所在位置、投射角度；此外更需依照所要投影至投影曲面的影像去決定投影平面所需呈現的畫面，故這裡投影曲面上的影像希望可以是任意形狀，投影曲面也可以隨意選擇(圖十二為投影概念示意圖)。我們於圖十三~十五展示在不同曲面上的放射面成像，圖十六與十七則展示轉動投影機方向的放射面成像。



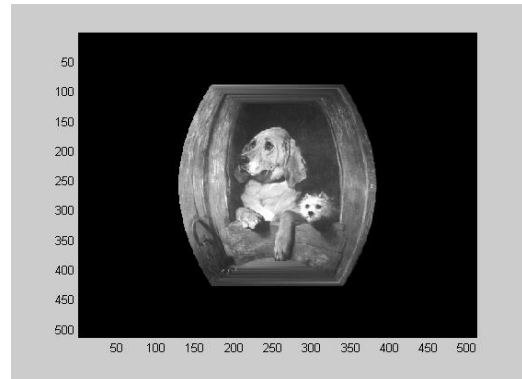
圖十二：投影概念示意圖



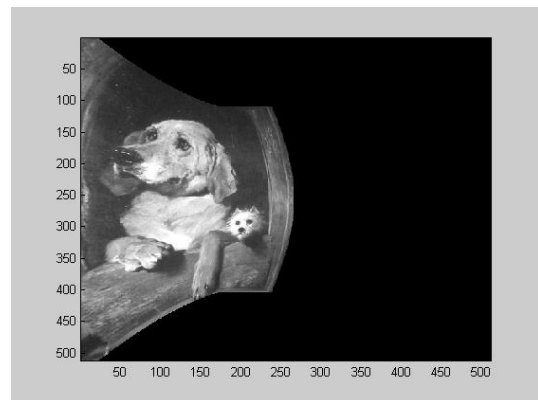
圖十三：平面投影上的放射面成像



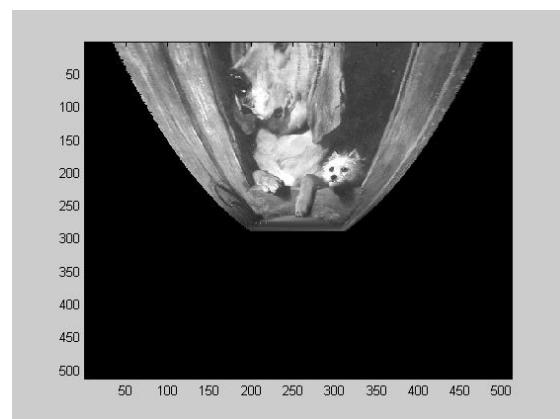
圖十四：圓柱投影上的放射面成像



圖十五：內凹球面投影上的放射面成像



圖十六：投影機向右偏轉時之球面投影



圖十七：投影機向下偏轉時之球面投影

子計畫三：動態 VR 運動復健輔助系統之互動式實景影像設計與研發

一、計畫概述

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。最後的部份則是對於下年度的展望。

二、本階段研究計畫之目的及成果

本研究報告為國科會整合性研究計畫「泛用型虛擬實境操控與運動復健輔助系統研發」中之子計畫三「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，本計畫已達成上列預定的 (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。圖十八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。



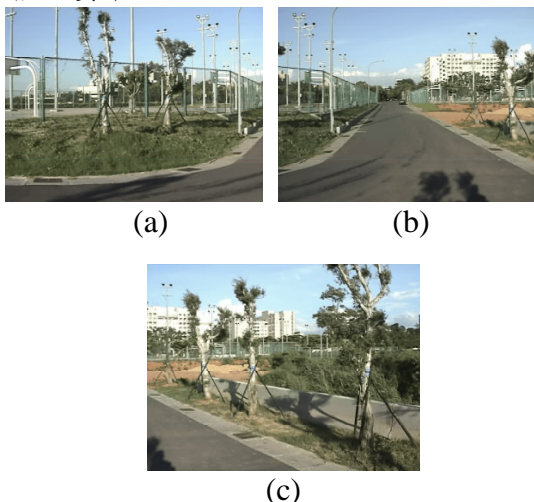
圖十八：影像擷取系統

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)。圖十九所示即是左、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，所以從圖十九可以看出來，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。圖二十是經過指定相關點，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 *AVIedit*，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。



圖十九：(a)左側攝影機影像、(b)中央攝影機影像、(c)右側攝影機影像



圖二十：接合並做亮度調整之廣角影像

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，為了方便快速編碼和解碼，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。

(一)*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*)。

(二) 彈性註冊方法 (*Elastic Registration*) [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。

(三) 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：

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

此處

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

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：

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

或

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

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

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

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

$$\begin{bmatrix} x_0^{(1)} x_1^{(1)} & y_0^{(1)} x_1^{(1)} & x_1^{(1)} & x_0^{(1)} y_1^{(1)} & y_0^{(1)} y_1^{(1)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_0^{(9)} x_1^{(9)} & y_0^{(9)} x_1^{(9)} & x_1^{(9)} & x_0^{(9)} y_1^{(9)} & y_0^{(9)} y_1^{(9)} \end{bmatrix} \begin{bmatrix} f_{11} \\ f_{21} \\ f_{31} \\ f_{12} \\ f_{22} \\ f_{32} \\ f_{13} \\ f_{23} \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} \text{----(5)}$$

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法(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法(Least Median Square)是所有方法中估計結果最正確的。

(四) 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。以圖二十一為例，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

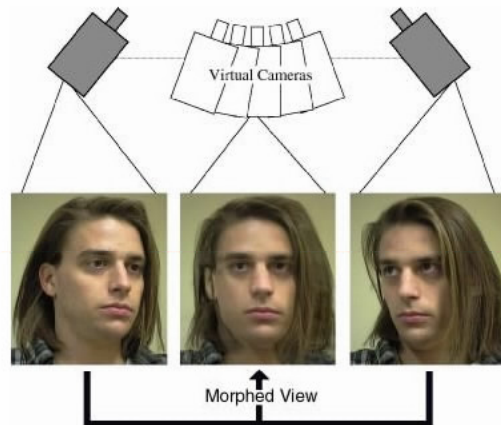
transformation)把既有的兩張影像(I_0 、 I_1)轉換成兩張光軸平行的影像(\hat{I}_0 、 \hat{I}_1)。

(b) Morph:

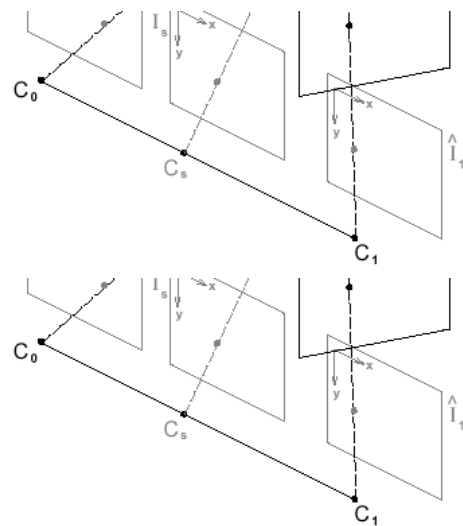
其次使用線性內差法產生所要的中間影像 \hat{I}_s 。

(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 。



圖二十一：View morphing



圖二十二：Warping

(五) 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。為了達到即時的目的，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，它的優點是可以用很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，而在最後一年去除了所有限制，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劃的最終目標。

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. Armangué and J. Page, “A Survey addressing the Fundamental Matrix Problem,” *Image Processing*, 2001. Proceedings. 2001 IEEE International Conference, Vol. 2, pp.209-212, 2001.
- [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.

子計畫四：動態 VR 運動復健輔助系統 之人機溝通介面及使用者情 緒瞭解

一、計畫概述

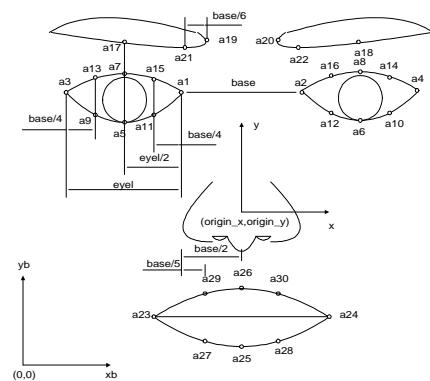
動態模擬系統之動態感覺，決定了虛擬實境之逼真程度，但如何表示動態模擬系統使用者之感覺概念，並傳遞給行為轉換與控制模組(子計畫一)及六軸運動平台、力回饋模組(子計畫二、三)做適當的參數調整與控制，是本計畫之重點。為了表示使用者的感覺概念，道，以將操控者的使用感受忠實的告知子計畫一與二進行模擬器的控制變數轉換及後續階層式控制法則之設計，並藉此以改善整個控制迴路及調整參數的方法，而提高了模擬器的真實性，本計畫可視為整個模擬器系統的高階回授單元。本計畫之兩個主要子題：計畫將發展出一套即時之人情緒狀態判定系統及能與機器系統做抽象式觀念感覺溝通與對映之方法，下面是我們在這一年計畫中，對以上兩方面所提出來實現方法的描述。

二、目前研究進度：

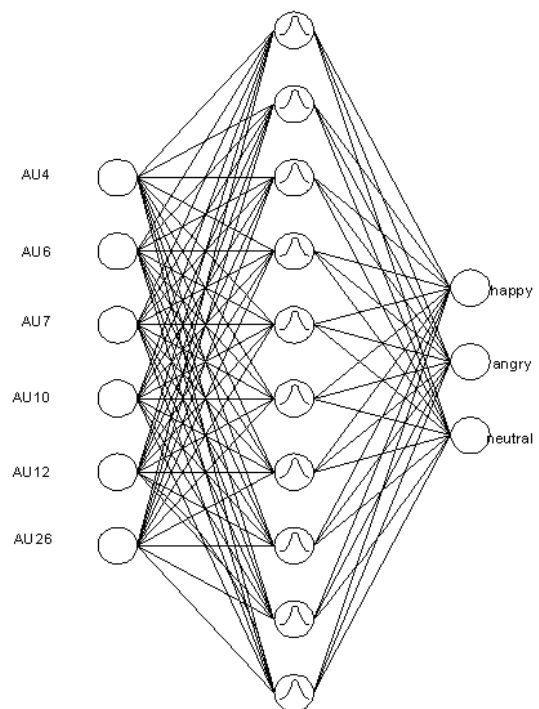
(1) 人類表情與情緒的辨識

人類表情與情緒的辨識此子系統第一部份為臉部特徵的選取。臉部特徵的選取對於辨識人臉和臉部表情是很重要的。人類臉部有些明顯的特徵如眼睛、嘴巴、眉毛、臉部輪廓等等。在本計畫中，我們擷取眉毛、眼睛及嘴巴作為主要的特徵。我們以前的研究顯示，我們首先使用粗略輪廓預測程序法(rough contour estimation routine)以及我們所提出的點輪廓偵測法(point contour detection method)這兩種影像處理的技術，來擷取眉毛、眼睛和嘴巴這三個特徵器官的正確輪廓。由於使用粗略輪廓預測程序法所得到的眉毛輪廓已經相當精準，足夠做為辨識之用，故不用再加強其輪廓。而使用粗略輪廓預測程序法

所得到的眼睛和嘴巴的輪廓不夠精準，所以我們再使用點輪廓偵測法現正進行臉部表情運動單元的分析。圖二十三顯示 30 個臉部特徵點；臉部特徵點可以用來代表臉部的特徵。因此當我們得到了這些特徵器官的精準輪廓之後，我們定義臉部特徵點來描述這三個臉部特徵的位置和形狀。接下來可由臉部特徵點組成運動單元。由於運動單元可以用來描述人類臉部基本的肌肉運動，所以臉部快樂、生氣、與中性情緒狀態表情可以藉由一放射狀函數基礎之神經網路辨識，它的效果經過驗證可達 92% 辨識率，如圖二十四所示。



圖二十三：臉部特徵定義圖



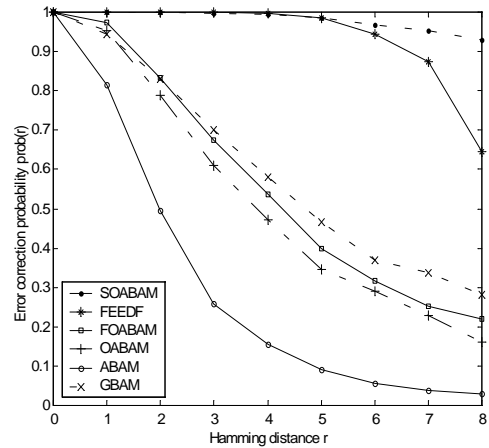
圖二十四：藉由放射狀函數基礎神經網

路辨識表情情緒狀態

(2) 人機系統溝通與對映法

人機系統溝通與對映之法 Takagi 等人係利用關聯網路 (Associate Memory; 簡稱 AM) 設計之赫比式 (Hebbian) 學習方法來建構之對映傳遞之關係, 但 AM 之兩對映是由(0,1)之串列對映至(0,1)之串列, 與上述之感覺對應, 為 0~1 間的任意實數對映, 我們初步的研究其對映精確度很差。本計畫主要是擴展並衍生觀念模糊集的概念, 以某一層的觀念節點的節點活性度值組表示一個感覺概念, 以機電整合受控體之物理量為輸入層, 感覺概念的溝通與學習可由函數大約化之對映方法來執行感覺概念之交流, 以使電腦有感覺概念意識與溝通的能力, 而對於觀念節點溝通所需之函數大約化, 我們擬設計出一套對映網路的裝置來進行, 亦即在受控機器中建立並學習某一感覺概念交流所須之對映網路。第一年我們以 Kosko 之模糊關聯網路 Fuzzy Associate Memory 設計, 所設計關聯網路(Associate Memory) 可以雙向聯想記憶 (Bidirectional Associative Memory), 簡稱 BAM, 實現。近來雙向聯想記憶已經在聯想記憶研究當中扮演一個很重要的角色。非對稱式雙向聯想記憶 (Asymmetric Bidirectional Associative Memory) 為雙向聯想記憶放寬鍵結權重必須要對稱的限制之結果, 且相較於常見的雙向聯想記憶結構有較好的記憶與回想效能。高階非對稱式雙向聯想記憶 (High-Order Asymmetric Bidirectional Associative Memory) 的記憶容量比一階的好很多, 然而新的高階聯想記憶設計法卻很少被提出來。在本篇論文裡, 我們所關心的是設計具有最大吸引範圍的二階非對稱式雙向聯想記憶 (Second-Order Asymmetric Bidirectional Associative Memory)。將它延伸到高階非對稱式雙向聯想記憶 (High-Order Asymmetric Bidirectional Associative Memory) 是有可能而且很簡單的。我們首先推導出對於二階非對稱式聯想記憶的鍵結權重矩陣能夠保證將所有標準圖樣

組回憶出來之充分條件。為了要遵守完全回憶定理, 接著闡述學習步伐大小也是適應性的局部訓練法則, 它將導致一個較快的設計時間。最後我們推導出下列定理: 在設計 SOABAM 時, 增大符合完全回憶定理的數值, 將會增加一個有雜訊的圖樣能夠正確地收斂到它的聯想圖樣之能力; 以這個定理當作基礎, 我們的演算法也予以修改, 能夠保證每一個訓練圖樣能儲存在具有越大的吸引範圍。針對 color graphics adapter (CGA) 字型的電腦模擬, 已經證明出我們所提出的局部訓練法則效果優於其他主要的 BAM 的設計, 其結果如圖二十五與圖二十六所示。



圖二十五：使用有 r 個雜訊 CGA 圖樣之各式聯想記憶辨識結果

	$r=0$	$r=1$	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$	$r=7$	$r=8$
SOABAM	1.000	1.000	0.999	0.997	0.993	0.984	0.968	0.952	0.929
FEEDF	1.000	1.000	1.000	0.999	0.997	0.986	0.945	0.873	0.646
FOABAM	1.000	0.973	0.833	0.674	0.536	0.398	0.317	0.251	0.220
GBAM	1.000	0.944	0.829	0.702	0.581	0.466	0.369	0.337	0.281
OABAM	1.000	0.954	0.788	0.610	0.471	0.347	0.291	0.228	0.162
ABAM	1.000	0.814	0.497	0.258	0.155	0.092	0.057	0.039	0.030

圖二十六：Average error correction probability prob(r) versus Hamming distance r of various BAM design

schemes.

三、結論與討論：

一理想的虛擬實境系統，應能隨時掌握操控者的情況與感受，並能讓操控者與系統模擬器做直覺性的溝通，而讓使用者有完全沈浸其中的感受，因此研究發展任何虛擬實境技術時，使用者與機器系統之間的互動與溝通為相當重要的一環。本子計畫藉電子攝影機隨時紀錄操控者臉部，藉影像處理技術偵測出操控者之情緒狀態，並建立操控者與模擬器間的人性化溝通管道，以將操控者的使用感受忠實的告知子計畫一與二進行模擬器的控制變數轉換及後續階層式控制法則之設計，並藉此以改善整個控制迴路及調整參數的方法，而提高了模擬器的真實性，本計畫可視為整個模擬器系統的高階回授單元。本子計畫已發展出一套即時之人情緒狀態判定系統藉線上拍攝操控者臉部影像序列，由影像處理技術偵測出操控者之臉部變化，進而推敲操控者快樂、生氣、與中性情緒狀態，它的效果經過驗證可達 92% 辨識率。在人機系統溝通與對映方面，我們以模糊關聯網路 Fuzzy Associate Memory 設計，實現感覺概念交流所須之對映網路，模擬證明我們所提出的局部訓練法則效果優於其他主要的 BAM 的設計。

四、參考文獻

- [1] T. Takagi, A. Imura, H. Ushida, and T. Yamaguchi, "Conceptual fuzzy sets as a meaning representation and their inductive construction," *International Journal of Intelligent Systems*, vol. 10, pp. 929-945, 1995.
- [2] F. L. Chung and T. Lee, "A new look at solving a system of fuzzy relational equations," *Fuzzy Sets and Systems*, vol. 88, pp.343-353, 1997.
- [3] Cootes, T. F., C. J. Taylor, D. H. Cooper, and J. Graham, "Active shape models their training and application," *Computer Vision and Image Understanding*, Vol. 61, No. 1, pp. 38-59, 1995.
- [4] P. Ekman, and W. V. Friesen, *The Facial Action Coding System*, Consulting Psychologist Press, San Francisco, CA, 1978.
- [5] Huang, C. L., and C. W. Chen, "Human facial feature extraction for face interpretation and recognition," *Pattern Recognition*, Vol. 25, No. 12, pp. 1435- 1444, 1992.
- [6] Kobayashi, H., and F. Hara, "Analysis of the neural network recognition characteristics of six basic facial expressions," in *Proceedings of 3rd IEEE Int. Workshop on Robot and Human Communication*, Nagoya, Japan, pp. 222-227, 1994.
- [7] Lanitis, A., C. J. Taylor, and T. F. Cootes, "Automatic interpretation and coding of face images using flexible models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 743-756, 1997.
- [8] Turk, M. A., and A. P. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuro-science*, Vol. 3, No. 1, pp. 71-86, 1991.
- [9] Yuille, A. L., P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," *International Journal of Computer Vision*, Vol. 8, No. 2, pp. 99-111, 1992.
- [10] W. Pedrycz, "Processing in relational structures: Fuzzy relational equations," *Fuzzy Sets and Systems*, vol. 40, pp. 77-106, 1991.
- [11] Y. F. Wang, J. B. Cruz, Jr., and J. H. Mulligan, Jr., "On multiple training pairs for bidirectional associative memory" *IEEE Trans. on Neural Networks*, vol. 1, pp. 275-276, Sept. 1991.
- [12] T. Wang, X. Zhuang, and X. Xing, "Designing bi-directional associative memories with optimal stability," *IEEE Trans. on Syst. Man, Cybern.*

- vol. 24, pp. 778-790, May. 1994.
- [13] H. Oh and S. C. Kothari, "Adaptation of the relaxation method for learning in bidirectional associative memory," *IEEE Trans. on Neural Networks*, vol. 5, pp. 576-583, July. 1994.
- [14] Y. Wu and D. A. Pados, "A feedforward bidirectional associative memory," *IEEE Trans. on Neural Networks*, vol. 11, pp. 859-866, July. 2000.
- [15] C.-S. Leung, L.-W. Chan, and E. Lai, "Stability and statistical properties of second-order bidirectional associative memory," *IEEE Trans. on Neural Networks*, vol. 8, pp. 267-277, Mar. 1997.
- [16] H. Shi, Y. Zhao, and X. Zhuang, "A general model for bidirectional associative memories," *IEEE Trans. on Syst. Man, Cybern. B*, vol. 28, pp. 511-519, Aug. 1998.
- [17] J. Park, C.-H. Kwon, and D. Park, "An optimization-based design procedure for asymmetric bidirectional associative memories," *IEEE Trans. on Neural Networks*, vol. 12, pp. 169-170, Jan. 2001.
- [18] F. L. Chung and T. Lee, "On Fuzzy Associative Memory with Multiple-Rule Storage Capacity," *IEEE Trans. Fuzzy Syst.*, vol. 4, no. 3, pp. 375-384, 1996.
- [19] H.-C. Chang, H.-C. Chen, and J.-H. Fang, "Lithology Determination from Well Logs with Fuzzy Associative Memory Neural Network," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, no. 3, pp. 773-780, 1997.

子計畫五：動態 VR 運動復健輔助系統 之人機溝通介面及使用者情 緒瞭解

一、中文摘要

本計畫提出一智慧型教導系統，以指導使用者操作技巧。方法為學習由專家或資深的訓練輔助員所提供有效率且正確的操作運動技巧。學習方法主要以類神經模糊網路來完成。目前我們已發展出一具監督式學習能力之新型遞迴式類神經模糊網路之架構設計及學習演算法推導。此網路的特點為具有線上之架構與參數學習之能力，並具高度學習能力。經由測試及與其他遞迴式網路之比較已驗證其優點。

關鍵詞：類神經網路、模糊網路、監督式學習、線上學習

Abstract

The project proposes an intelligent teaching system to guide user's operation skill. The system is built by learning an efficient and proper exercising skill provided by experts or experienced operator. Learning is performed mainly by a neural fuzzy network. We have developed structure of a new type of recurrent neural fuzzy network and its learning algorithm. The network is characterized with on-line structure and parameter learning and owns high learning ability. After testing and comparisons with other recurrent network, we have verified the advantages of the proposed network.

Keywords: Neural network, Fuzzy network, Supervised learning, On-line learning

二、Motivation And Objective

The project aims at setting up an intelligent teaching system to guide and evaluate user's operation skill. The role of the system can be regarded as an instructor. The development of the system relies on proper operation and exercising skill provided by experts. To learn expert's knowledge and skill, a learning network and mechanism should be developed. Since the

operation of a person usually depends on previous states, a recurrent network is more suitable than a feedforward network for the learning task. For nonlinear system learning, the most commonly used model is the neural or neural fuzzy network. The performance of a neural fuzzy network has shown to be superior than a neural network. For this reason, in this project, a recurrent fuzzy network is proposed.

Some recurrent neural fuzzy networks have been proposed in the literature. In [1], a dynamic fuzzy logic system is proposed with multiple external output. In [2], a recurrent neuro-fuzzy model is proposed. In [3], a dynamic fuzzy neural network (DFNN) is proposed. In DFNN, the premise and defuzzification parts are static while the consequent parts of the fuzzy rules are recurrent neural networks with internal feedback and time delay synapses. For these recurrent neural fuzzy networks, one common characteristic is that the recurrence is achieved by including external feedback. To apply these networks to dynamic system processing problems, we still need to know the order of both the system. To conquer this problem, a TSK-type Recurrent Fuzzy Network (TRFN) [4] is proposed. In contrast to the aforementioned networks, in TRFN, the recurrent property is achieved by including internal feedback. The internal variables, derived from fuzzy firing strengths, are fed back to both network input and output layers in TRFN, and no *a priori* knowledge of the plant order is required. In TRFN, there are no rules initially, they are constructed by on-line structure and parameter learning. TRFN has the following advantages over a neural network: faster learning speed, higher learning accuracy, and each node and weight has its corresponding meaning.

≡ 、 Structure and Learning of TRFN

A. Structure of TRFN

In this section, structure of the TRFN (as shown in Fig. 27) is introduced. A network with two external inputs and a single output is considered here for convenience. This six-layered network realizes a recurrent fuzzy network of the following form:

Rule 1: IF $x_1(t)$ is A_{11} and $x_2(t)$ is A_{12} and $h_1(t)$ is G

THEN $y(t+1)$ is $a_{10}+a_{11}x_1(t)+a_{12}x_2(t)+a_{13}h_1(t)$

and $h_1(t+1)$ is w_{11} and $h_2(t+1)$ is w_{21}

Rule 2 : IF $x_1(t)$ is A_{21} and $x_2(t)$ is A_{22} and $h_2(t)$ is G

THEN $y(t+1)$ is $a_{20}+a_{21}x_1(t)+a_{22}x_2(t)+a_{23}h_2(t)$

and $h_1(t+1)$ is w_{12} and $h_2(t+1)$ is w_{22}

where A and G are fuzzy sets, w and a are the consequent parameters for inference output h and

y , respectively. The consequent part for the external output y is of TSK-type and is a linear combination of the external input variables x and internal variables h , plus a constant.

To give a clear understanding of the mathematical function of each node, we will describe functions of TRNFN layer by layer. For notation convenience, the net input to the i th node in layer k is denoted by $u_i^{(k)}$ and the output value by $O_i^{(k)}$.

Layer 1: No function is performed in this layer. The node only transmits input values to layer 2.

Layer 2: Two types of membership functions are used in this layer. For external input x_j , the following Gaussian membership function is used,

$$O_i^{(2)} = \exp\left\{-\frac{(u_j^{(2)} - m_{ij})^2}{f_{ij}^2}\right\} \text{ and } u_j^{(2)} = O_j^{(1)} \quad (1)$$

where m_{ij} and f_{ij} are, respectively, the center and the width of the Gaussian membership function of the i th term of the j th input variable x_j . For internal variable h_i , the following sigmoid membership function is used,

$$O_i^{(2)} = \frac{1}{1 + \exp\{-u_i^{(2)}\}} \text{ and } u_i^{(2)} = O_i^{(5)} \quad (2)$$

Links in layer 2 are all set to unity.

Layer 3: The output of each node in this layer is determined by fuzzy AND operation. Here, the product operation is utilized to determine the firing strength of each rule. The function of each rule is

$$O_i^{(3)} = \prod_{j=1}^{n+1} O_j^{(2)} = \frac{1}{1 + \exp\{-O_i^{(5)}\}} * \exp\left\{-\left(\sum_{j=1}^n \frac{(O_j^{(2)} - m_{ij})^2}{f_{ij}^2}\right)\right\} \quad (3)$$

where n is the number of external inputs. The link weights are all set to unity.

Layer 4: Nodes in this layer perform a linear summation. The mathematical function of each node i is

$$O_i^{(4)} = \sum_{j=0}^{n+1} a_{ij} u_j^{(4)} = a_{i0} + \sum_{j=1}^n a_{ij} x_j + a_{i(n+1)} h_i \quad (4)$$

where n is the number of external input variables, and a_{ij} , $j=0, \dots, n+1$, are the parameters to be tuned. Links from this layer to layer 6 are all equal to unity.

Layer 5: The context node functions as a defuzzifier for the fuzzy rules with inference output

h . The link weights represent the singleton values in the consequent part of the internal rules. The simple weighted sum is calculated in each node,

$$h_i = O_i^{(5)} = \sum_{j=1}^r O_j^{(3)} w_{ij} \quad (5)$$

As in Fig. 27, the delayed value of h_i is fed back to layer 1 and acts as an input variable to the precondition part of a rule. Each rule has a corresponding internal variable h and is used to decide the influence degree of temporal history to the current rule.

Layer 6: The node in this layer computes the output signal y of the TRNFN. The output node together with links connected to it act as a defuzzifier. The mathematical function is

$$y = O^{(6)} = \frac{\sum_{j=1}^r O_j^{(3)} O_j^{(4)}}{\sum_{j=1}^r O_j^{(3)}} \quad (6)$$

B. Learning of TRFN

The task of constructing the TRFN is divided into two subtasks: structure learning and parameter learning. The objective of the structure learning is to decide the number of fuzzy rules, initial location of membership functions, and initial consequent parameters. On the contrary, the objective of parameter learning is to tune the free parameters of the constructed network to an optimal extent.

Since there are no rules initially in TRFN, the first task in structure learning is to decide when to generate a new rule. Clustering on the external input x , which represents the spatial information, is used as the criterion. Based on this concept, the spatial firing strength

$$F^i(x) = \prod_{k=1}^n O_k^{(2)} = \exp \left\{ - \left(\sum_{j=1}^n (x_j - m_{ij})^2 f_{ij}^2 \right) \right\} \in [0, 1] \quad (7)$$

is used as the criterion to decide if a new fuzzy rule should be generated. For each incoming data $x(t)$, find

$$I = \arg \max_{1 \leq i \leq r(t)} F^i(x) \quad (8)$$

where $r(t)$ is the number of existing rules at time t . If $F^I < F_m^I(t)$, then a new rule is generated, where $F_m^I(t) \in (0, 1)$ is a pre-specified threshold that decays during the learning process. Once a new rule is generated, the next step is to assign initial centers and widths of the corresponding membership functions. Here, they are decided by

$$m_{(r(t)+1)i} = x_i(t) \quad (9)$$

$$f_{(r(t)+1)i} = S \cdot \sum_{j=1}^n \frac{(x_j - m_{ij})^2}{f_{ij}^2}$$

for $i = 1 \cdots n$, according to the

first-nearest-neighbor heuristic, where $S \geq 0$ decides the overlap degree between two clusters. The number of fuzzy sets in each external input dimension is equal to the number of fuzzy rules. Once a new rule is newly generated during the presentation of $(x(t), y(t))$ data, generation of the corresponding consequent node in layer 4 and context node in layer 5 follows. The initial constant value a_{io} connected to layer 4 is set to $y(t)$, and the other a_{ij} parameters are assigned as small random signals in $[-0.05, 0.05]$ initially. For the newly generated context node, its fan-in comes from all the existing rule nodes in layer 3. The initial link weights w are set as random values in $[-1, 1]$ to make the initial values of internal variables h locate in the sensitive region of membership function G . This way, a quick parameter learning can be reached at the beginning. The output, h , of the new context node is fed back as input in the precondition part of the newly generated rule. With this setting, each rule has its own memory elements for memorizing the temporal firing strength history. Via repeating the above process for every incoming training data, a new recurrent rule is generated, one after another, and a whole TRFN is constructed finally.

As to the parameter learning, the real time recurrent learning is used. Detailed learning algorithm may refer to [4].

四、Results And Discussion

Suppose the control action sequence u provided by the expert is described by the following equation:

$$u(k+1) = 0.72 u(k) + 0.025 u(k-1) x(k-1) + 0.01 x^2(k-2) + 0.2 x(k-3)$$

where x denotes the state of the operated plant. To learn the mapping by TRFN, only the current state x and action u are fed as the network input. After 9000 time steps of training, there rules are generated and there are 33 network parameters in total. A test sequence is generated and the test result is shown in Fig. 28, and the root mean square (RMS) error is 0.0313. To show the effectiveness and efficiency of the recurrent part in TRFN, the performance of Elman's recurrent neural network (ERNN) [5] is compared. There are five hidden nodes in ERNN and the number of network parameters is 54. The test RMS error for the test sequence is 0.078. That is, the performance of TRFN is better than that of ERNN.

五、Conclusion

An intelligent teaching system based on the proposed TSK-type recurrent fuzzy network

(TRFN) is proposed. The network enables us to learn the operation data provided by an expert or experienced operator. Besides, the proposed TRFN has also been accepted and published in international journals[4].

六、References

- [1] G. C. Mouzouri and J. M. Mend el, Dynamic nonsingleton fuzzy logic systems for nonlinear modeling, *IEEE Trans. Fuzzy Syst.*, vol. 5, no. 2, 199-208, 1997.
- [2] J. Zhang and A. J. Morris, Recurrent neuro-fuzzy networks for nonlinear process modeling, *IEEE Trans. Neural Networks*, vol. 10, no. 2, pp. 313-326, 1999.
- [3] P. A. Mastorocostas and J. B. Theocharis, A recurrent fuzzy-neural model for dynamic system identification, *IEEE Trans. Syst., Man and Cyber., Part B: Cybernetics*, vol. 32, no. 2, pp. 176-190, 2002.
- [4] C. F. Juang, "A TSK-type recurrent fuzzy network for dynamic systems processing by neural network and genetic algorithm," *IEEE Trans. Fuzzy Systems*, vol. 10, no. 2, pp. 155-170, April, 2002.
- [5] J. L. Elman, "Finding structure in time," *Cognitive Science*, Vol. 14, pp. 179-211, 1990.

Figure 28: Outputs of the desired action sequence (solid curve) and model TRFN.

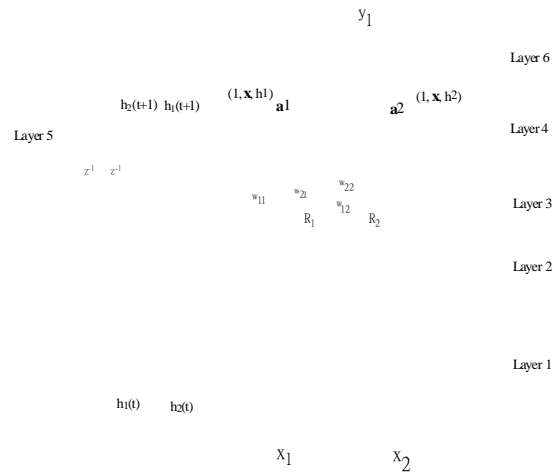


Figure 27: Structure of the proposed TRFN.

