

行政院國家科學委員會補助專題研究計畫 成果報告
 期中進度報告

先進家用機器人系統研發(2/3) —子計畫四
家用機器人之智慧型互動系統研發
Development of the Intelligent Interaction System
of an Advanced Home Robot

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC-91-2213-E-009-027

執行期間： 91年 08月 01日至 92年 07月 31日

計畫主持人：宋開泰教授

共同主持人：

計畫參與人員： 蔡奇謐、陳文俊、林嘉豪、王耀慶、陳志豪

成果報告類型(依經費核定清單規定繳交)： 精簡報告 完整報告

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

處理方式：除產學合作研究計畫、提升產業技術及人才培育研究計畫、列管計畫及下列情形者外，得立即公開查詢

涉及專利或其他智慧財產權， 一年 二年後可公開查詢

執行單位：國立交通大學電機與控制工程學系

中 華 民 國 92年 05月 29日

摘 要

本計畫之目的在發展家用機器人之智慧型互動系統(Intelligent Interaction System, IIS)。在計畫執行第二年期間，我們完成了家用機器人之網路遙控系統設計與測試，可經由個人電腦、筆記型電腦或 PDA 等上網裝置在世界任何地方操控家用機器人，是一種實用之互動介面。對於家庭中成員互動部分，我們已完成以機器人頭部之 web camera 取像之人臉辨識系統。經由此項技術，機器人可以在環境搜尋特定之對象(如家庭成員)，進而追蹤這個人之移動，達成一種互動之控制。基於第一年之語音辨識研究成果，我們完成家用機器人語音辨識操控設計，機器人可以接受特定人之語音命令進行設定之動作，在下一階段，我們將整合人臉辨識與語音操控及網路遙控功能，達成一種智慧型之互動系統。另外，本計畫也已完成一 3 自由度之機械手臂設計與製作，並安裝於家用機器人上，接下來我們將進行抓取物品之整合互動設計與實驗。

Abstract

The purpose of this project is to develop an intelligent interaction system (IIS) for the home robot. In the second year of the project, we have achieved a remote control of the robot using internet (through a PC, a notebook or a PDA). This capability provides a easy-to-access interaction between the host and his/her home robot from virtually any place in the world. We also completed a face recognition and searching system using the on-board web camera. This provides the robot a capacity to recognize the members of the family and interacts with the member by following him/her or waiting for commands. Based on the voice recognition achieved in the previous year, we developed a voice control system for the home robot. Therefore, in the near future, an integration of the face recognition with the voice control system will be carried out. Moreover, we also designed and constructed a 3-degree-of-freedom robot arm for the home robot. A more practical interaction system will be developed using these features in the next phase of project.

第一章 計劃緣由與目的

家用機器人之智慧型互動系統包括機器人與家庭成員之互動、機器人與環境中物體之互動、及機器人與其他機器人之間之互動。圖 1 顯示家庭機器人之互動系統示意圖。在實際作法上將牽涉到機器人本體之設計、各種智慧型感測器之介面設計、訊號處理、機器人決策行為之架構、機器學習、以及多機器人之間之合作機制等等技術開發。為了使機器人的使用更為「自然」與「方便」，並且賦予機器人與使用者之間能夠產生溝通的能力，需要使用到智慧型人機介面來達成這個需求。本計畫一方面發展語音辨識系統，讓機器人能夠藉由語音的方式，得到使用者目前希望給予機器人的訊息，並轉化為實質的控制訊號，作出適時適當的反應動作。另一方面亦可藉由影像辨識技術來達成，即發展人臉辨識系統提供機器人作家庭成員的辨識，讓機器能對不同的使用者提供不同的服務及互動。

本計畫第一年已完成家用機器人的本體機構及控制模組，並且研製出超音波環境感測模組與反應操控系統；第二年之工作重點是機器人與家庭成員間之智慧型機人介面(Robot-man interface)系統研發，這包括應用無線網路遙控機器人，使用 PDA 操控及監測，將人臉辨識與去年的影像追蹤技術結合，達成與家庭成員的互動，進一部將語音辨識功能應用到語音互動操控。

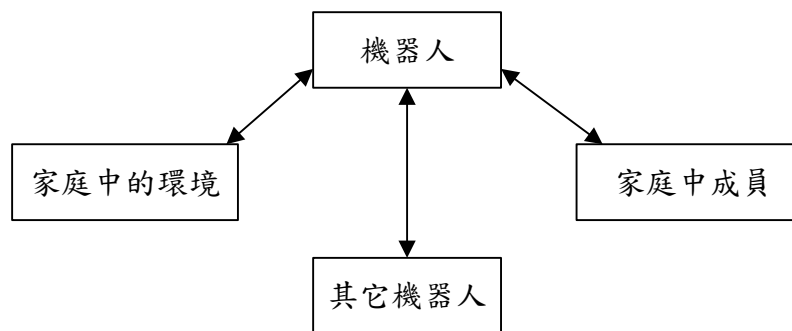


圖 1 家庭機器人之互動系統示意圖

第二章 研究方法與成果

本計劃第二年之研究重點在家用機器人與家庭成員互動系統的研製。為了測試與展示所發展的方法，我們將所發展的法則與軟體皆整合到所發展的家用機器人雛型，如圖 2 所示。在家用機器人與家庭成員之互動系統設計方面，我們的目標是整合家庭成員辨識系統及語音辨識系統，以提高家用機器人與人的互動性及使家用機器人能提供使用者更多的幫助。在設計上，我們將使家用機器人能夠達成以下的功能：

- (1) 使用者可藉由語音對家用機器人下達家庭成員搜尋的命令。
- (2) 藉由家庭成員辨識系統對使用者的語音控制命令進行權限的確認。以確保家用機器人的安全性。
- (3) 使家用機器人藉由簡單的語音辨識與人對話，並且藉由家庭成員辨識系統，使家用機器人能對不同的使用者進行不同的對話方式。



圖2 家用機器人H1實體圖

圖 3 為整合語音辨識系統及家庭成員辨識系統之互動系統架構圖。在此架構中，我們可以藉由網路或 PDA 對家用機器人下達命令，亦可藉由語音辨識進行語音命令的辨別。利用語音辨識系統進行使用者命令的辨認及與使用者對話。整合家庭成員辨識確認使用者所下達的命令是否超過使用者可行使的權限，或對家庭中成員進行搜尋與追蹤。在與人對話方面，配合上家庭成員辨識系統將使對話內容更豐富多變，而不在只是單一的對答方式。

本實驗室發展之網際網路遠端操控機器人系統，可以讓我們在世界任何一個角落，透過 internet 遙控家裡的機器人，不但能夠直接控制機器人的移動、安排它巡邏的路徑，監控機器人所在的位置狀況，還可以看到機器人所在的環境影像。使用者不需要作任何安裝的動作，只要輕鬆上網瀏覽即可。使用者可以隨時隨地透過網路操控機器人，更換巡邏路線，取代固定式的監視器的有監視死角的缺點，達到居家看護與安全監控的效用，如圖4所示。

在機器人端，我們架構一精簡之 Linux 系統，使用 CF 卡為儲存媒介增加系統之強韌性。透過無線網路可即時將 camera 抓取之家中影像壓縮並透過無線網路傳回 Server。另外也回傳機器人的狀態與位置提供使用者參考。

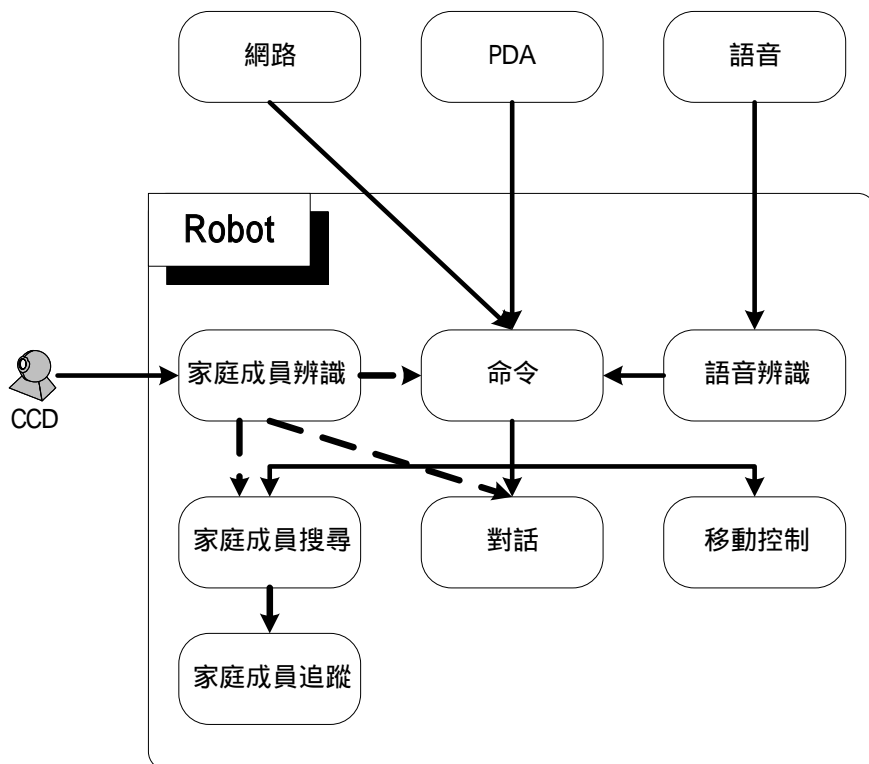


圖 3 家用機器人與家庭成員互動系統(整合語音辨識及家庭成員辨識系統)

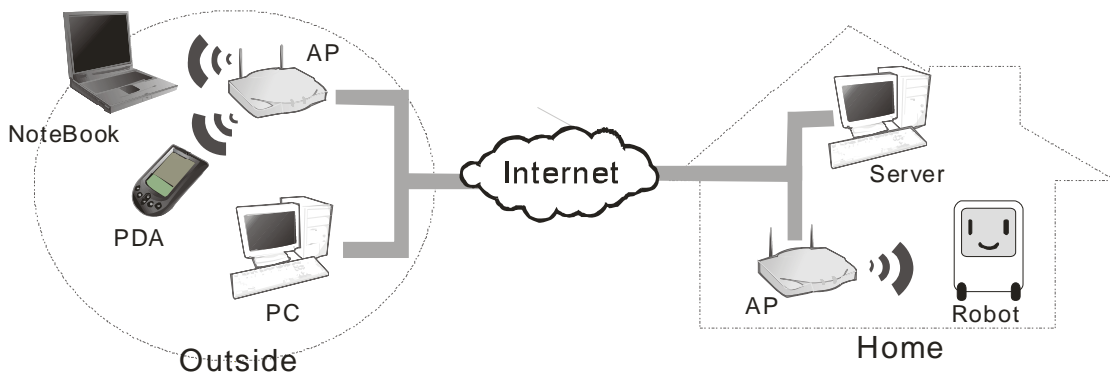


圖4 網際網路遠端操控個人機器人系統

在Server端，替使用者端以及機器人端交換資訊，包括使用者的命令、機械人的狀態以及機器人所抓取之影像。透過 Server 管理連線可減輕機器人的負擔，並且按照用戶端的狀況調整影像品質，維持最佳操縱狀態。在使用者端，利用 JAVA 跨平台之特性製作操作介面程式。使用者可透過 PDA 或者任意平台電腦上的網頁瀏覽器操控家中之機器人，檢視家中狀態，安排巡邏路徑或者自由行走皆可。操作介面以簡單為主，只要 click 即可輕鬆操作。未來再配合上醫療儀器或警報系統，將可使家用機器人擁有更多與人互動的應用，可達成家庭看護及家庭保全的功能。

2.1 人臉辨識系統

為了使家用機器能夠對不同的家庭成員提供適當的服務與互動，使家用機器人能更人性化，我們發展一種家庭成員的辨識方法，使家用機器人能夠辨識不同的家庭成員。人臉辨識屬於主動式辨識方法，不需高解析度的攝影機或近距離拍攝，只需要家庭成員距離攝影機某一範圍內就可進行辨識，再配合家用機器人的運動控制，家用機器人將可以主動的去進行人臉的搜尋與辨識，不需家庭成員自己走到機器人面前等待進行辨識。

要完成家庭成員辨識系統，除了要能夠辨別出不同的家庭成員之外，在進行辨識之前，我們還需要能夠從影像中找出人臉在影像位置及大小。圖 5 顯示家庭成員辨識系統及互動的主要架構。其主要分為三部分：(1)人臉偵測、(2)人臉辨識、(3)人臉追蹤。本系統主要藉由家用機器人上之 Web Camera 作為影像輸入設備，將其擷取到的影像傳到機器人上的主電腦進行人臉搜尋、臉部調整及人臉辨識。使用者可藉由網路要求機器人對特定家庭成員進行搜尋，並且整合人臉追蹤系統[1]，使其能對搜尋到的目標進行持續的追蹤，以等待使用者的下一步要求，例如，與特定家庭成員進行視訊交談。

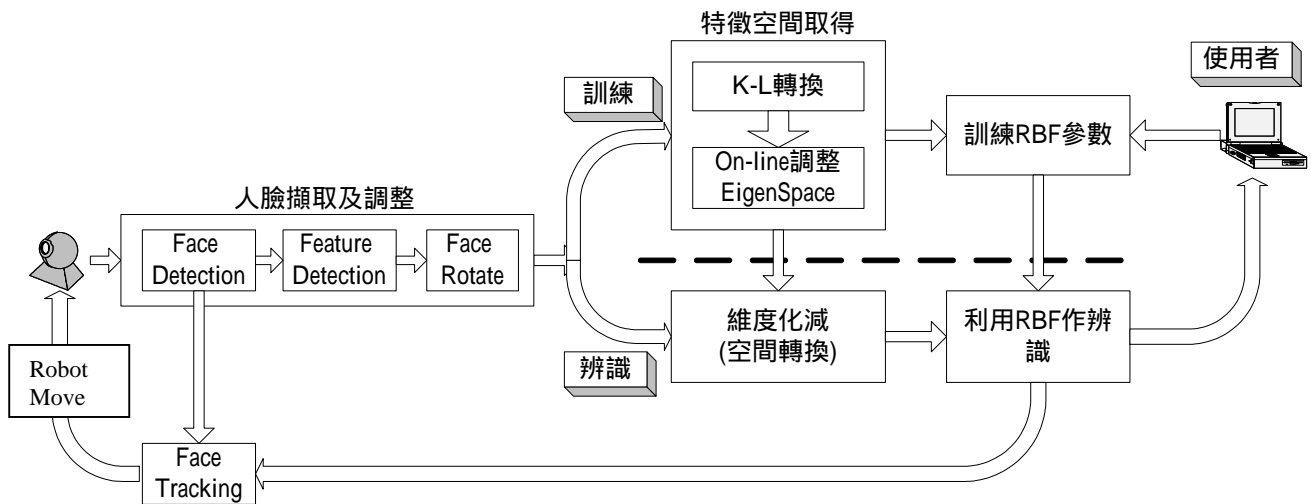


圖 5 家庭成員辨識系統及機器人與人互動

2.2 人臉搜尋

在家用機器人進行人臉辨識之前，我們必須了解經由 Web Camera 所擷取到的影像中的人臉位置與大小，所以我們需在影像中對人臉進行搜尋。在人臉搜尋方法，一般常利用人臉的某些特徵進行人臉與非人臉的判別。其中膚色是利用於識別人臉的主要特徵之一[1, 2, 3]。我們經由 Web Camera 所得到的影像資料為紅、綠、藍(RGB)三原色，RGB 三原色的色彩形成與光的照度、物體表面、攝影機的特性都有關。RGB 三原色所構成的色彩空間(color space)很容易受到光線的影響，會隨光線的變化有很大的變動，所以 RGB 三色並不適合直接使用於膚色區域分割，必須要做色彩空間的轉換以減少 RGB 三原色對光之靈敏度。我們使用 Normalized RGB 的色彩空間作膚色區域分割，因為 Normalized RGB 的轉換只需簡單的數學運算就可達成，且能夠降低光的照度對影像影響。式(1)為 RGB 三原色轉換到正規化 RGB 的公式，其中 r 、 g 、 b 為所求， Y 為影像之灰階值。

$$\begin{aligned}
 r &= \frac{R}{R+G+B} & g &= \frac{G}{R+G+B} & b &= \frac{B}{R+G+B} \\
 Y &= \frac{R+G+B}{3}
 \end{aligned}
 \tag{1}$$

為了找出所有的膚色區塊，我們利用 Web Camera 擷取多張人臉影像，並且將所有膚色區塊取出，將其從 RGB 色彩空間投影到 Normalized RGB 色彩空間，我們求出式(2)能夠將膚色與非膚色區隔出來。往後我們只需判斷影像的 Y , r , g 的值是否在式(2)所圍起來的區域內，就可分離影像的主體區(膚色部分)及背景區(非膚色部分)。

$$\left\{ \begin{array}{l} Y > 30 \\ Y < 190 \\ g > 0.00036 \times Y + 0.25 \\ g < -0.00052 \times Y + 0.44 \\ r > 0.00088 \times Y + 0.2 \\ r < 0.0008 \times Y + 0.58 \end{array} \right. \quad (2)$$

在圖 6 中，使用 Web Camera 實際擷取一張彩色影像，利用式(2)進行膚色區塊的分割。

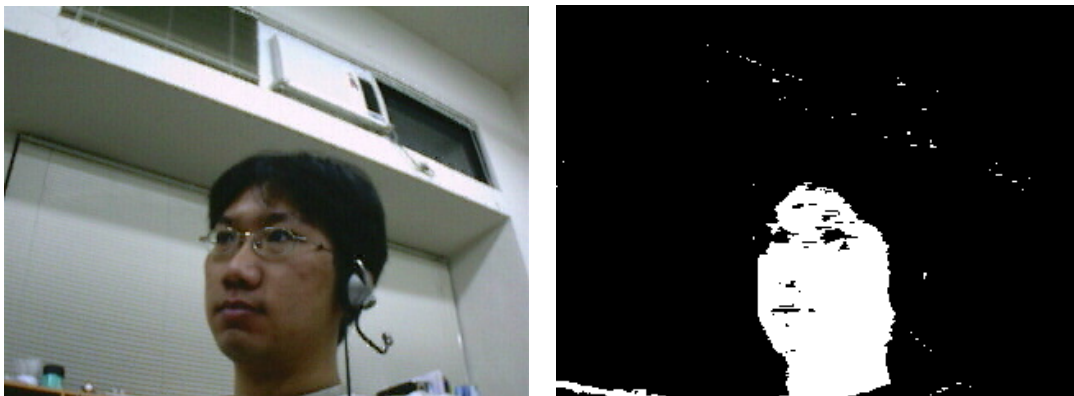
單用膚色作為人臉判別，容易受到其它與膚色相似的物體的影響而產生誤判，因此，除了膚色之外，我們亦使用另一人臉的特徵—輪廓。在文獻[4, 5]中提到，人臉的輪廓近似於橢圓形，且其橢圓形之長短軸半徑比約為 1.2:1。在輪廓判別部份，我們利用 Sobel 的方法對灰階影像作邊緣偵測，再以邊緣偵測的結果作橢圓形的輪廓搜尋。最後整合膚色分割與輪廓判別的結果，找出臉部位置。

由於 Web Camera 輸出的影像會有許多雜訊存在，使系統在作橢圓形判別時容易出錯，所以在利用 Sobel 作邊緣偵前，先使用低通濾波器來將雜訊去除使影像平滑。在此使用一個 3×3 的遮罩矩陣來與原影像做摺積(convolution)，如式(3)所示。

$$I_{lpf} = g * h = \sum \sum h(j-m, i-n) g(m, n)$$

$$h = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad (3)$$

I_{lpf} 為濾波後之影像。 g 為原始灰階影像。 h 為低通濾波之遮罩。



(a) 原始彩色影像 (b) 膚色區塊分割結果
圖 6 彩色影像膚色區塊分割(白色為膚色區塊，黑色為非膚色區塊)

Sobel 運算則是利用水平方向梯度運算子 G_x 及垂直方向梯度運算子 G_y 對原始灰階影像作摺積後，邊緣影像即為此兩個分量的絕對值和。最後再設定一邊緣梯度的門檻值區分出邊緣及非邊緣區塊。水平方向梯度運算子 G_x 及垂直方向梯度運算子 G_y 兩個運算子分別以下面兩個矩陣表示。

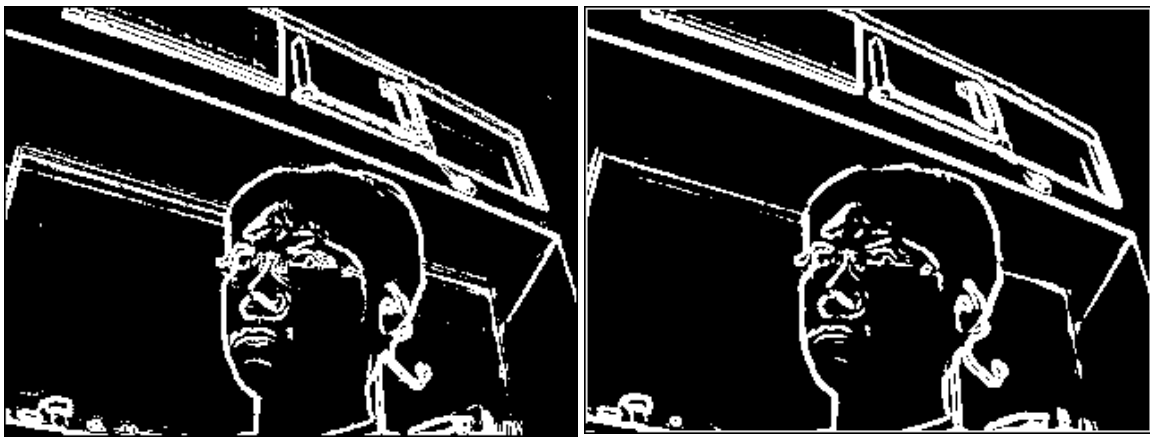
$$G_x = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \quad G_y = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

圖 7 顯示出未經過低通濾與經過低通濾波後的邊緣擷取結果。

在人臉搜尋系統中，我們整合膚色分割及橢圓搜尋方法，但在實際的實作過程中，我們發現在 320×240 大小的影像中執行橢圓形搜尋相當的浪費時間。為了加速人臉搜尋速度，以提供家用機器人一個即時人臉搜尋系統，所以在進行橢圓形搜尋之前，利用膚色分割後的資料，求其在 X 軸及 Y 軸上的投影量，以判斷出可以人臉的區塊，減少橢圓形搜尋的範圍。圖 8 為圖 6(b) 影像在 X 軸及 Y 軸上的投影。藉由膚色在 X 軸及 Y 軸上的投影量，我們分別設定一個門檻值，找出有可能的人臉範圍，最後在利用橢圓形搜尋得到圖 9 的結果，圖 9 中黑線所框出來的區域為可能人臉區塊，白線框出的為橢圓形搜尋後的結果。

2.3 人臉影像前置處理

為了要達到提高辨識率的目標，除了要有強健的辨識演算法之外，還需要有一適當的人臉影像的前置處理機制，以減少輸入影像資料間的差異性(諸如：不同照度下、臉部影像大小不同、臉部的旋轉角度不同及表情不同等情況)。



(a) 未經過低通濾波

(b) 經過低通濾波

圖 7 邊緣擷取

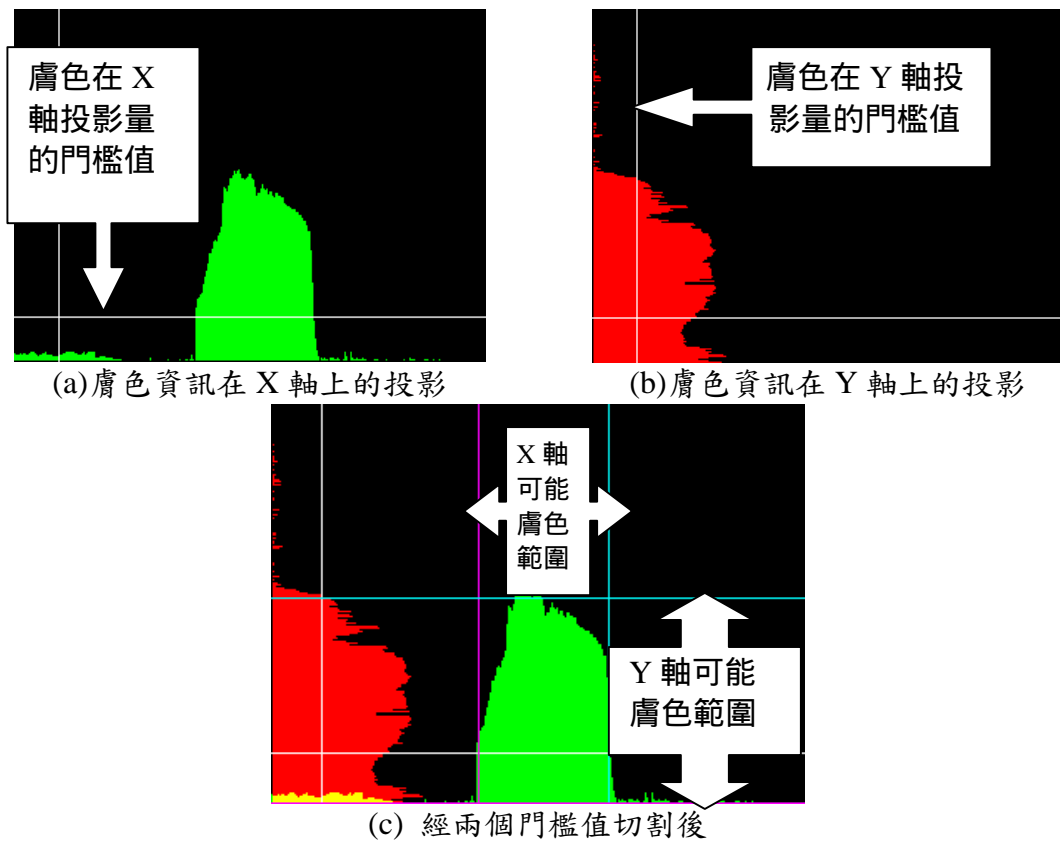


圖 8 膚色資訊在 X 軸及 Y 軸上的投影

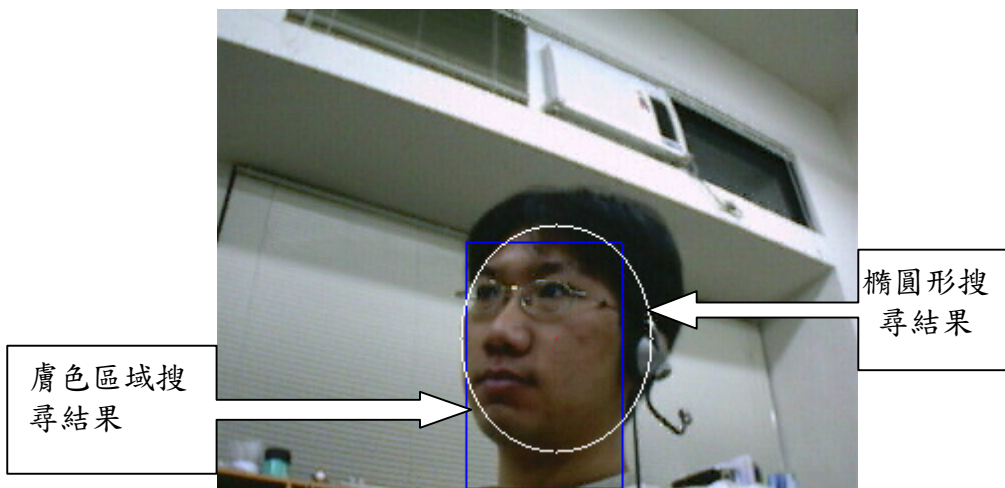


圖 9 膚色分割及橢圓形搜尋結果

在人臉影像的前置處理方面，我們將輸入的影像經過正規化的調整，以得到標準的影像形式。希望藉由減少輸入影像之間的差異性，提高辨識率。

本計畫提出的人臉影像正規化處理過程包括：

- (1) 特徵(眼睛)的擷取[2]
- (2) 臉部的調整
- (3) 臉部照度補償

在臉部照度補償部分，因不同的照度對影像所造成的變化是整體性的，也就是說，當光線變暗，整體的灰階值會偏暗，反之亦然。所以我們利用此特性，將每張人臉的灰階值進行調整，使其平均值為 128。

假設我們有一大小為 N 個像素的影像 I ， I_i 表示影像中的第 i 個像素的灰階值，此影像 I 之平均值 I_{ave} 可由式(4)求出。再利用式(5)對影像 I 中的所有像素進行調整，以得到照度補償的臉部影像。

$$I_{ave} = \frac{1}{N} \sum_{i=1}^N I_i \quad (4)$$

$$I_i = (128 - I_{ave}) + I_i \quad (5)$$

圖 10 為利用多張未經照度補償的臉部影像，經過為經照度補償後的結果。由其結果可看出 確實改善了光照度的問題。

2.4 辨識系統

本系統主要的目的是為了應用在家用機器人上，所以系統的即時性對我們來說非常的重要。我們希望儘可能的減少辨識時所需花費的時間。在此，我們使用主要元素分析法(Principal Component Analysis)的特徵臉演算法(Eigenfaces approach)[3]對臉部影像作維度的化減，再利用 Radial Basis Function Neural Network 作最後的辨識。特徵臉演算法是利用主要元素分析法(Principal Component Analysis)透過 K-L 轉換來達成影像上統計資訊的維度化減的方法，其可將輸入訊號的高維度資訊中所有的相關性消去，以化減資料的維度。在演算法的運算過程中，將二維矩陣的資訊，以一維的向量方式來表示，並且從輸入的所有影像中萃取出臉部特徵空間，臉部特徵空間的維度比原空間維度小，將輸入影像投影到臉部特徵空間，可以得到較少的維度的資訊，以達到減少維度的目的。

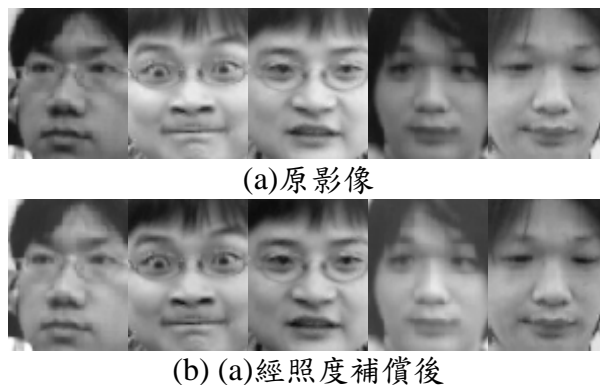


圖 10 臉部影像照度補償

在求特徵空間的方法中，一般利用 K-L 轉換求得，但因為其需要龐大的計算量。在此，我們利用文獻[4]中提到化減 K-L 轉換的方法，以減少求特徵空間時所需的計算量。但每次需新增資料庫成員時，需要將所有的資料庫成員取出重新計算，相當麻煩。在此，我們利用[5]所提到 On-Line 更新特徵空間的方法。利用此方法，我們便可直接利用新的影像資料對特徵空間作調整，快速的更新特徵空間。

在辨識方法，因為 Radial Basis Function 的類神經網路具有快速學習的好處，相當適合於即時(Real-Time)系統。在此，我們利用 Radial Basis Function 作人臉的訓練及辨識，並且利用 Normal Equation[16]方法來求得 RBF 的權重值。再利用最小均方誤差(Least Mean Square Error; LMSE)作參數細部的調整，以求得最佳的 RBF 參數。圖 11 為 RBF 類神經網路的架構。

2.5 人臉辨識測試結果

目前已實際藉由 Web Camera 抓取影像進行人臉搜尋、臉部調整及人臉辨識。以下就各個階段的實際展示其結果。

a) 人臉搜尋及臉部調整實際測試結果：

在此，我們在機器人主控電腦上實現人臉搜尋及臉部調整的演算法，該程式可直接藉由 Web Camera 擷取影像進行處理，表 1 為利用程式進行實際測式的結果。

b) 人臉辨識系統測試結果

目前利用實驗室中 4 位成員的臉部影像進行人臉辨識測試，每位成員取 10 張影像，且 4 位成員中取其中 3 位建立資料庫，另一位則用來測試非資料庫成員時的情況。表 2 為 3 位成員各取 8 張用以建立人臉資料庫。

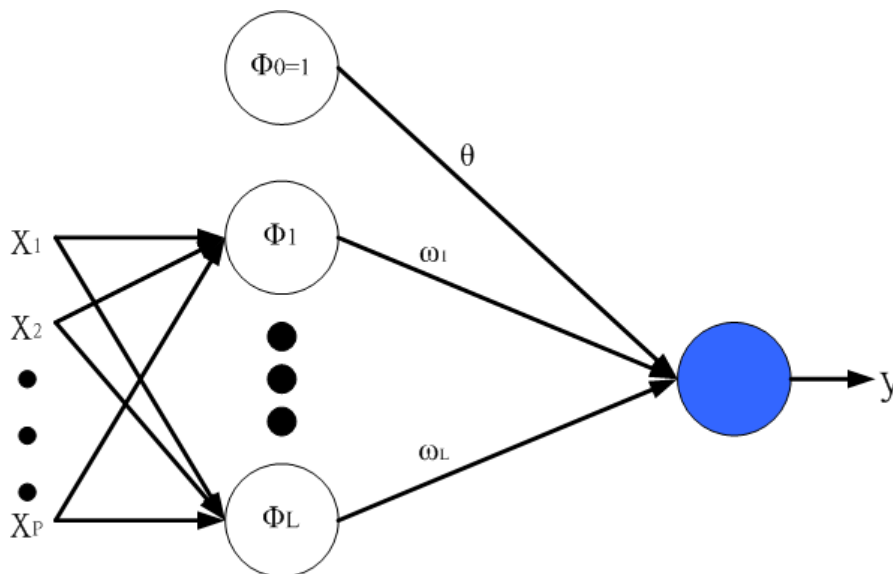





圖 11 RBF Neural Network 架構

表 1 人臉搜尋及臉部調整測試結果

由 Web Camera 擷取到的影像	臉部經正規化調整	經照度補償
		
		

表 2 資料庫成員臉部影像

文俊	
正憲	
忠富	

在人臉辨識部分，我們設計了兩個實驗，任取一張人臉影像(不限定資料庫中成員)，與資料庫中所有成員比對。下表 3 為測試結果。若該成員的 RBF 的輸出值接近 1，則表示此輸入影像為該成員。反之，若 RBF 的輸出值接近 0，則表示此輸入影像非為該成員。




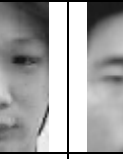
之所以辨識系統的輸出值能夠如此的接近 0，主要是因為在 RBF 的輸出值 y 再經過 $\frac{1}{1+e^{-y}}$ 計算輸出。

- (1) 取出資料庫中某位成員之參數，對環境中的 4 張人臉進行比對。下表 4 為測試結果。
- (2) 由前面兩實驗結果，我們可以看出本系統對於存在於資料庫中成員的辨識率很高，而且 RBF 的輸出值不是接近 0，就是接近 1。其幾乎不會出現在模糊地帶(0.5 左右)，而造成不容易如何區分輸出的結果。對於不存在資料庫中之成員，容易造成誤判為資料庫中的某位成員。

表 3 人臉辨識實驗結果一

環境中輸入辨識系統之影像		資料庫中影像		資料庫中影像		資料庫中影像		資料庫中影像	
									
辨識系統之辨識結果，輸出為成員名字		文俊	文俊	正憲	正憲	忠富	忠富	忠富	正憲
辨識系統輸出值	文俊	1	1	5.60E-65	3.13E-45	3.71E-38	9.73E-12	2.09E-45	
	正憲	2.32E-51	1.01E-56	1	1	3.53E-48	2.43E-14	1	
	忠富	3.18E-45	8.56E-45	2.88E-45	3.30E-43	1	2.54E-10	2.32E-31	

表 4 人臉辨識結果二

環境中人臉影像					
選取之資料庫成員	文俊	1	1.95E-64	2.12E-19	2.09E-45
	正憲	1.69E-66	1	7.43E-9	1
	忠富	4.76E-45	2.77E-45	1	2.32E-31

2.6 家用機器人之特定人搜尋

為了提供家用機器人對家庭成員的識別及其與人的互動，我們將家庭成員辨識系統及去年所完成之人臉追蹤系統整合到家用機器人上，製作一家用機器人對特定人之搜尋系統。

在流程設計上，整個系統主要可以分為三種狀態：等待、搜尋及追蹤。機器人一開始處於等待狀態，當使用者透過網路下達搜尋命令後，機器人會切換為搜尋狀態，在家庭中自由移動，並將在行進中擷取到的影像進行人臉偵測及與使用者所指定的家庭人員作比對。若順利的找到使用者所指定的家庭成員時，會進入追蹤狀態，對該家庭成員進行持續的追蹤，等待使用者下一步的命令，例如：要求與該家庭成員進行通話或拿東西給該家庭成員。若整個環境搜尋完皆找不到指定的成員，機器人將會進行等待的狀態。圖 12 為主要的流程圖，除了開始時為等待狀態及結束時為追蹤狀態，其它系統大部分都是在執行搜尋的動作。

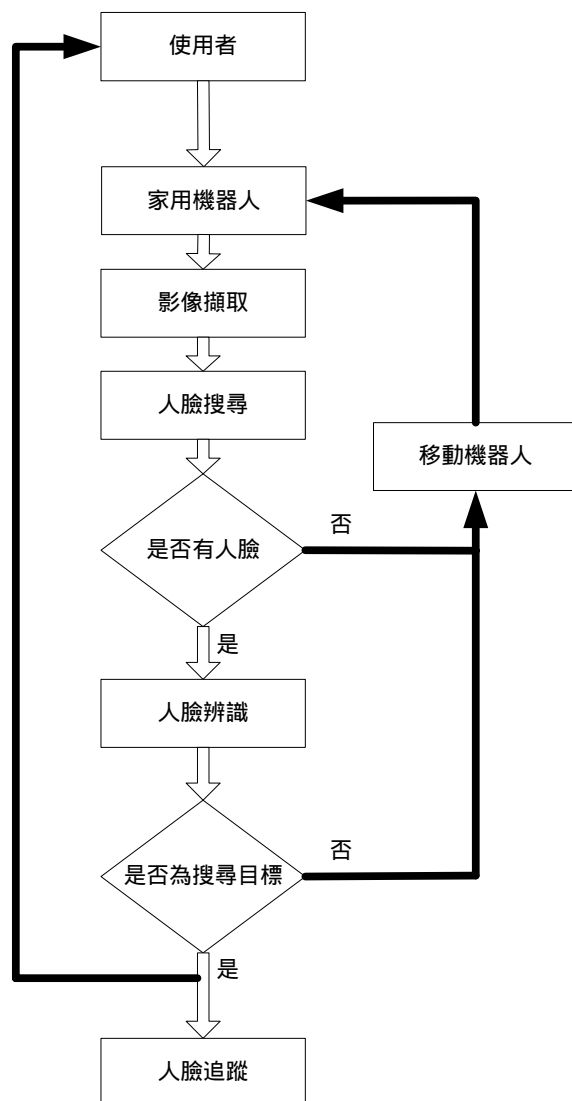


圖 12 家用機器人對特定人搜尋之流程

在影像輸入介面方面，我們採用 37 萬像素的 Web Camera，其使用 USB 介面與機器人之主電腦進行資料的傳遞。為了減少在進行特定人搜尋時的計算時間，但不希望因影像解析度太低，而造成影像太過模糊，所以在這我們影像採用 320×240 的大小進行影像的處理。在移動式追蹤方面，我使用第一年完成之 2 自由度頭部控制系統及本體操控系統，對搜尋到的目標進行移動式的持續追蹤。在人與機器人的溝通介面方面，我們透過無線網路進行人與機器人之間的資訊傳遞。目前即將進行實地的進行家用機器人對特定人搜尋，至於在 PC 上的測試與模擬結果，如 2.1.4 節所顯示。

2.7 家用機器人之語音操控

目前我們已完成語音操控系統，圖 13 為實際操控狀態。家用機器人是為了要給使用者帶來便利，為了方便使用者對家用機器人的操作，實驗室學長利用第一年所完成之語音辨識系統，製作一對於家用機器人之行為操控系統。

設計上，機器人除了需要持續目前的運動狀態，亦需隨時等待使用者的呼叫。只要呼叫的命令輸入，就可以介入機器人的控制狀態之中。並且讓機器人進入下一個狀態，即等待命令的部分。對於特定的任務模式，例如家電控制、資訊傳達等功用來說。機器人需要使用者給定所需要的資訊，利用這些資訊，讓家用機器人可以正確地執行任務。圖 14 動作(Action)包含了兩個概念，一是機器人運動控制的部份，二是在機器人的語音回應。在特定的時刻，機器人可以用語音的方式回應資訊給使用者，增加與使用者之間的互動性，回應之內容是以



圖 13 實際語音操控

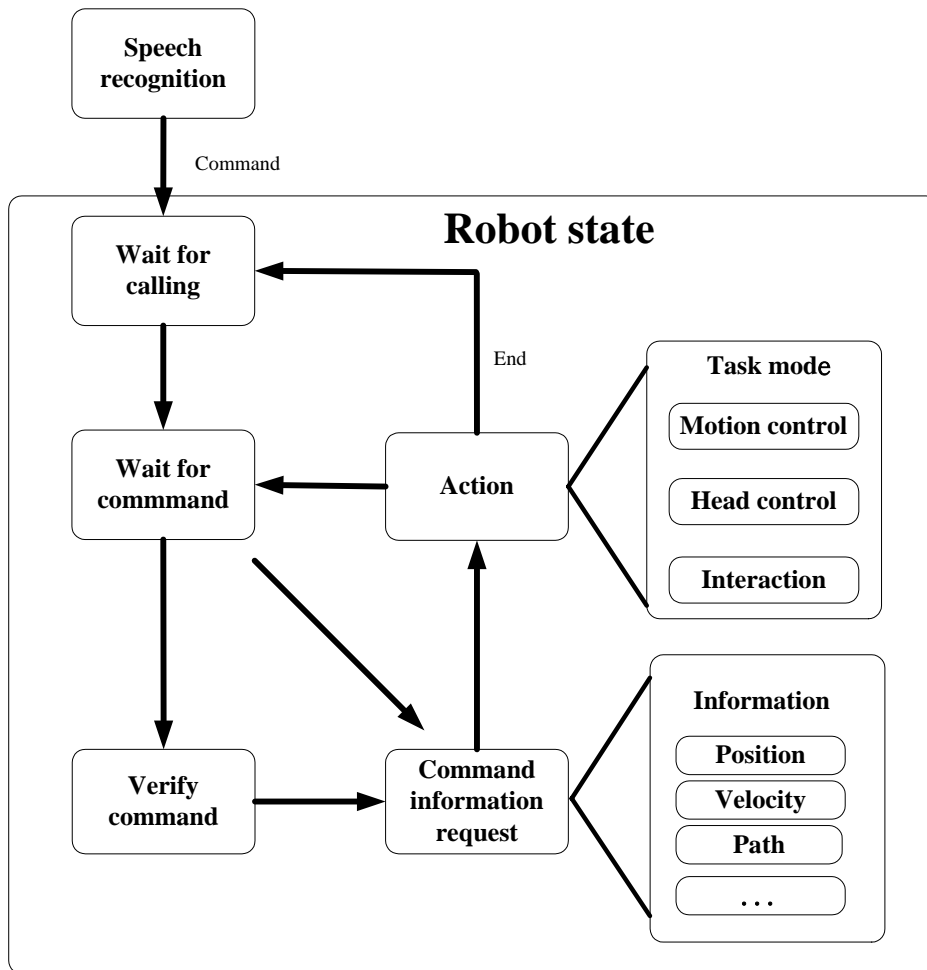


圖 14 機器人內部狀態流程圖

機器人自己目前的狀態或是對於資訊的要求。在機器人執行任務時，使用者仍隨時可以透過以語音的方式變更其目前之狀態，這樣的設計是具彈性的，可以使得使用者在任何時刻皆可以對機器人的狀況進行控制。

在語音的輸入介面，目前是透過無線麥克風的方式，將語音信號傳入 DSK Board，因此使用者可以在近端或遠端對家用機器人發出語音命令。目前使用之無線麥克風，有效距離可達 20 公尺，因此，在此有效距離之內，使用者可以利用無線麥克風直接的對機器人發出呼叫或命令。家用機器人語音資料庫設計為小字彙量的範疇，主要會可分為三個部分，運動控制命令、頭部運動控制命令及互動對話詞彙，並以中、英文混用的方式，中文以雙字詞為主、英文以單字為主。

在實際的語音操控上，藉由無線麥克風將信號傳入 DSK Board，將處理後將辨識結果經由 RS232 介面傳入家用機器人之主控電腦加以處理。

在實際對本實驗室多位成員進行短暫地語音訓練之後，進而測試統計在實際上對於一般使用者之適用性及通用性。語音辨識功能測試的實際測試數據如圖 15 所示。從多人測試結果

發現，在簡短的語音訓練過程之後，四位測試同學測試的語音辨識成功率最低的測試數據為平均值 76%，最高的測試數據為平均值 93%。從此一結果顯示，本辨識系統可以符合一般的適用性，即在簡短的訓練過程後就可以得到一定程度之語音辨識效果。而在進行語音辨識的測試上，五位實驗同學以自己的名字作為測試，讓系統辨別使用者是誰，在圖 16 中可以看出，辨識率分佈在 70%~90%之間。在訓練次數上的分析，增加次數在辨識率的表現上差異不大。實際將語音辨識系統結合於家用機器人之上，並實際地利用語音辨識的功能進行對家用機器人之操控。實際測試之語音辨識成功率與在系統發展時於電腦上測試的效果並無太大之差別。針對家用機器人之操控表現，設計了一測試實驗驗證本論文之語音辨識系統之實用性。

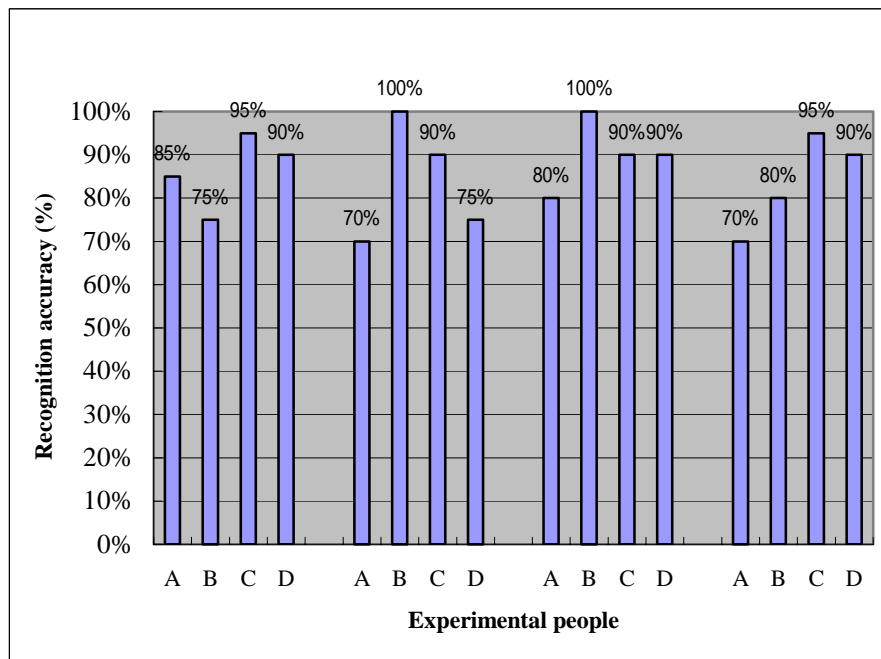


圖 15 多人測試語音辨識成功率統計

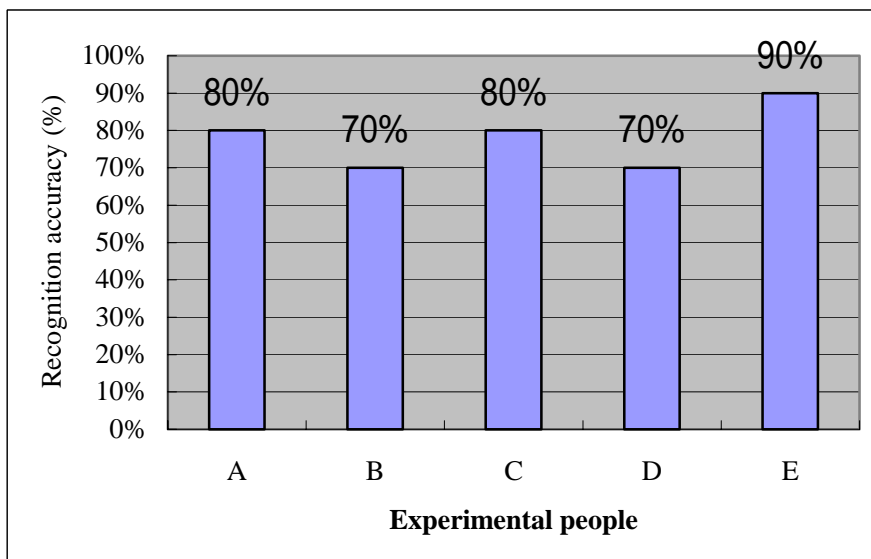


圖 16 分辨語者之測試結果

第三章 計畫成果自評與未來工作

本計畫第二年的執行期間我們完成了家用機器人經由網路操控系統，使用者可以透過 PDA 等上網裝置與機器人互動。藉由語音辨識及人臉辨識，我們發展出家庭成員之互動系統，即語音操控家用機器人及家用機器人對特定家庭成員搜尋。在影像辨識方面，我們提出了一個人臉辨識系統架構，並應用於家用機器人上，使其能夠與人互動，此系統架構主要包含了：(1)人臉搜尋、(2)特徵擷取、(3)臉部正規化、(4)人臉辨識、(5)人臉追蹤、(6)機器人頭部運動控制、(7)本體行為控制。使機器人能主動的在環境中搜尋人臉，並進行辨識。亦可利用頭部及本體的移動對特定的目標進行持續的追蹤或搜索。語音辨識系統的架構包含以 DSP 實現辨識演算法及在家用機器人主電腦之處理架構。已完成一平均辨識成功辨識率可達到 90% 的中英文混合語音辨識系統，並且在家用機器人上完成實驗測試，可以符合一般人說話的方式對家用機器人發出控制命令，讓家用機器人在使用者的指揮下，作指定的動作。

在人臉辨識部分，我們目前使用的是一般的 Web Camera，常因為焦距的問題，而照成影像模糊，使的辨識結果容易出錯。系統會有背光的問題，當物體後面有強光時，膚色就無法顯示出來，這是人臉偵測時最容易造成搜尋不到人臉的原因。在家用機器人上，背光的問題是有需要被解決的。因為家用機器人在移動時，有可能會行成背光的情況。

我們接下來將整合語音操控與影像辨識，使機器人有更完整的互動系統，增加其實用性。另外在下一階段，我們也將展開機械手臂在家用機器人的應用設計，並展開多機器人系統的設計與應用研究。

計畫成果自評：本計畫之研究內容與原計畫完全符合並已達到預期目標，所發展之機器人影像追蹤控制設計及家庭成員辨識設計等研究成果具有學術價值與實用性，已有四研究論文被接受發表於國際研討會 ICMT 2002、IEEE ICRA2003、IEEE CIRA 2003 及中華民國自動控制研討會，較完整的論文將投稿於學術期刊。

參考文獻

- [1] Jordao, L.; Perrone, M.; Costeira, J.P.; Santos-Victor, J., “Active face and feature tracking”, Image Analysis and Processing, 1999. Proceedings. International Conference on , 1999Page(s): 572 -576.
- [2] Han CC, Liao HY, Yu GJ, Chen LH. “Fast face detection via morphology-based pre-processing”, Pattern Recognition, 2000, 33(10):1707-1712
- [3] 黃士銓, “灰階人臉辨識之研究”, 碩士論文, 國立交通大學, 民國 91 年
- [4] 黃俊欽, “灰階人臉辨識之研究”, 碩士論文, 國立交通大學, 民國 85 年
- [5] P. Hall, D. Marshall, and R. Martin. “Incremental eigenanalysis for classification, In British Machine Vision Conference, volume 1, pages 286-295, September 1998.
- [6] J.O. Rawlings. “Applied Regression Analysis”, Wadsworth & Brooks/Cole, Pacific Grove, CA, 1988.