

# 致理技術學院

## 資訊網路技術系 實務專題報告

### 手勢辨識簡報播放系統

指導教師：高楊達 老師

學生：蔡孟純(19634111)

邱佩盈(19634118)

鄭惠馨(19634125)

許銀佳(19634127)

陳韋廷(19634136)

中華民國 99 年 12 月

# 致理技術學院

## 資訊網路技術系 實務專題報告

### 手勢辨識簡報播放系統

學生：蔡孟純(19634111)

邱佩盈(19634118)

鄭惠馨(19634125)

許銀佳(19634127)

陳韋廷(19634136)

本成果報告書經審查及口試合格特此證明。

指導老師：\_\_\_\_\_

中華民國 99 年 12 月

# 專題研究授權書

本授權書所授權之專題研究為\_\_\_\_\_

共\_\_\_\_\_人，在致理技術學院資訊網路技術系\_\_\_\_\_學年度第\_\_\_\_\_學期完成資網實務專題。

專題名稱：\_\_\_\_\_

同意      不同意

本組同學共\_\_\_\_\_人，皆同意著作財產權之論文全文資料，授予教育部指定送繳之圖書館及本人畢業學校圖書館，為學術研究之目的以各種方法重製，或為上述目的再授權他人以各種方法重製，

不限地域與時間，惟每人以一份為限。

上述授權內容均無須訂立讓與及授權契約書。依本授權之發行權為非專屬性發行權利。依本授權所為之收錄、重製、發行及學術研發利用均為無償。上述同意與不同意之欄位若未勾選，該組同學皆同意視同授權。

指導教師姓名：

專題學生簽名：

(親筆正楷)

學號：

(務必填寫)

中華民國      年      月      日

# 致 謝

大學四年裡，最重要的，莫過於三下到四上這最重要的專題研究，轉眼間，這半年即將邁入尾聲，憶起這段時間的點點滴滴，心中感慨不已，感謝讓我們有所成長的一切。

首先誠摯的感謝指導老師高楊達老師，老師教學嚴謹，為人務實認真，在老師的諄諄教誨下，使我們得以一探影像處理領域的驚奇之旅，在我們遇到瓶頸時，也適時地提供建議及思考方向，讓我們在求學的路上可以循著正確的方向邁進。

在此特別感謝本系系主任、系上師長及助理，感謝他們這些日子以來的陪伴，並給予每位專題組員適時的指導，對學生於專題研究室，甚至是各項實驗的儀器也能給予很大的協助，讓學生能無後顧之憂地使用專題研究室。

感謝本組組員們彼此之間的包容與體諒，且能將自身的知識與經驗，在遭遇困難時能以冷靜的態度與努力不懈的精神面對，在困難時彼此激勵與幫助，由於大家互助合作，努力耕耘及孜孜不倦的學習精神，才使得專題研究順利完成。

最後還要感謝家人、同學與學弟妹們，在這段期間內所給予的支持與鼓勵。在此，僅以此論文獻給所有幫助過，指導過我們的人。

# 摘要

隨著科技的發展，電腦為人們帶來極多便利的生活，然而「科技始終來自於人性」也一直是所有科技人員不斷研發新技術的動力。為此，發展出多種方便人機互動的介面，其中尤以「手勢」最為容易與機器互動。本論文提出一套以影像為基礎之手勢辨識系統，用於簡報時透過手勢的辨識即可自由的播放投影片，本系統處理了投影片播放時常用到的幾項功能：「下一頁」、「上一頁」、「重頭開始播放」以及「停止播放」等四項。論文的架構分為「手勢資料庫建立」及「即時手勢辨識」兩大模組，透過攝影機對手勢進行取像，並經由偵測影像中特定顏色定位出手部座標，加以使用型態影像學技巧處理影像破碎現象、最後將二維平面座標轉為極座標，並以極座標之「角度」及「距離」當作手勢的特徵。由於手勢與攝影機的距離與拍攝角度等因素，造成偵測到的手勢大小及方向略有不同，無法進行的有效辨識，因此，本論文利用爬山式序列表示法(Mountain Climbing Sequence(MCS))及大小正規化兩種技術，順利的解決了旋轉及放大及縮小等問題，最後，再使用歐式距離(Euclidean Distance)進行影像的比對動作。未來，本論文所實作之系統除了使用在簡報撥放功能外，亦可應用在其他的人機互動的介面上，以期減少人類與機器間的溝通問題。

關鍵字：手部偵測、型態影像學、極座標、爬山式序列法、歐式距離。

# 目 錄

專題研究授權書.....	i
致 謝.....	ii
目 錄.....	iv
圖 目 錄.....	vii
表 目 錄.....	xi
第壹章 緒論.....	1
1.1 研究背景 .....	1
1.2 研究動機 .....	3
1.3 研究目的 .....	5
1.4 研究重點 .....	7
第貳章 影像辨識之理論及技術探討.....	10
2.1 研究理論及應用 .....	10
2.1.1 手勢辨識方式 .....	11
2.1.2 手勢辨識設備 .....	15
2.2 技術 1 探討 手勢辨識系統[9] .....	18
2.3 技術 2 探討 電腦互動技術實驗課程設計與教學手冊[13]	23
2.4 技術 3 探討 虛擬滑鼠：以視覺為基礎之手勢辨識[14]	27
2.4.1 手勢辨識的規則 .....	27

2.4.2 影像分割的規則 .....	28
2.4.3 特徵擷取 .....	28
2.4.4 手勢分析與結果 .....	28
2.5 技術 4 探討 手勢辨識應用於遙控音樂播放系統[15] ...	32
第參章 手勢辨識簡報播放系統.....	39
3.1 取得影像[16] .....	40
3.2 影像前處理 .....	41
3.2.1 手部偵測 .....	41
3.2.2 影像二值化[18] .....	43
3.2.3 中間值濾波[Median Filter] .....	45
3.2.4 形態學處理 .....	48
3.2.5 相連區域法[Connected Component] .....	50
3.3 特徵抽取 .....	52
3.3.1 Sobel 邊緣偵測法 .....	52
3.3.2 特徵抽取---極座標[27] .....	53
3.3.3 特徵正規化 .....	56
3.4 手勢辨識 .....	59
3.5 播放簡報[30] .....	60
第肆章 系統流程.....	61

4.1 環境設定 .....	61
4.2 系統流程.....	62
第五章 結論.....	69
參考文獻.....	71

# 圖目錄

圖 1 專題研究架構圖 .....	9
圖 2 RGB 的混色結果 .....	12
圖 3 RGB 的座標系統 .....	13
圖 4 HSI 的座標系統 .....	13
圖 5 學習紅外線機制 .....	17
圖 6 膚色以色調與飽和度表示之分佈區域 .....	19
圖 7 之膚色偵測流程圖[9] .....	20
圖 8 運用最小平方線到手掌影像範例圖 .....	21
圖 9 手部外圍輪廓波形掃描示意圖 .....	21
圖 10 手勢辨識系統的七種手勢範例圖 .....	22
圖 11 經由平均結果所得之背景 .....	23
圖 12 物件進入畫面中 .....	24
圖 13 白色部分是系統偵測到的物件變化 .....	24
圖 14 Connected Component 演算法使用的 model .....	25
圖 15 偵測最大物件的範例 .....	26
圖 16 光線劇烈變化導致程式誤判的範例 .....	27
圖 17 攝取影像序列之取像軟體 .....	29
圖 18 取像環境 .....	30

圖 19 為校正之紅色十字記號 .....	30
圖 20 攝影機位置校正之情況 .....	31
圖 21 背景相減法 .....	33
圖 22 擴張運算 .....	34
圖 23 侵蝕運算 .....	35
圖 24 封閉算子 .....	35
圖 25 開放算子 .....	35
圖 26 小雜訊濾除 .....	36
圖 27 左手臂影像擷取： .....	36
圖 28 邊緣化 .....	37
圖 29 系統流程圖 .....	39
圖 30 本系統所使用的四種手勢 .....	40
圖 31 (a)擷取到的原圖、(b)系統保留綠色像素結果.....	43
圖 32 圖影像二值化函數轉換 .....	44
圖 33 二值化前的圖 .....	45
圖 34 二值化後的圖 .....	45
圖 35 以數值 33 為中心點的矩陣 .....	46
圖 36 使用中間值濾波後的範例圖 .....	47
圖 37 手勢影像的原圖 .....	48

圖 38 二值化後經中間值濾波之結果 .....	48
圖 39 8 鄰居範例圖 .....	49
圖 40 一個紅色框標示出一個區域 .....	51
圖 41 條件判斷後，紅框為我們要的手部區域。 .....	51
圖 42 Sobel 運算子.....	53
圖 43 Sobel 邊緣偵測的結果 .....	53
圖 44 極座標 .....	54
圖 45 已取得手勢各維度的極座標 .....	56
圖 46 數值的正規化之後的結果 .....	58
圖 47 為圖 46 移位後的結果.....	58
圖 48 系統介面.....	62
圖 49 系統使用的手勢圖。 .....	63
圖 50 開啟簡報後出現畫面 .....	64
圖 51 此為開啟簡報檔後出現的畫面 .....	64
圖 52 此為拳頭手勢 .....	65
圖 53 此為辨識完手勢所在的頁面 .....	65
圖 54 此為一級棒的手勢 .....	66
圖 55 此為辨識後成功之畫面 .....	66
圖 56 此為五的手勢 .....	67

圖 57 此為回到第一頁的畫面 .....	67
圖 58 此為結束手勢 .....	68
圖 59 此為結束後的原始畫面 .....	68

# 表 目 錄

表 1 系統開發環境表。 .....	61
--------------------	----

# 第壹章 緒論

## 1.1 研究背景

隨著科技的發展，電腦與機器為人們帶來極多便利的生活，然而「科技始終來自於人性」也一直是所有科技人員不斷研發新技術的驅動力。為此，甚至有研發出可以讓人機互動 (Human-Computer Interaction, HCI)[1]的語言，也就是使用程式語言如 Visual Basic、C#、C++、Java 等來與電腦溝通，再透過電腦來控制機器的動作。而現代的溝通方式擁有上百種的語言，但在人類尚未發展出語言文字之前，人與人之間的互動方式便是使用身體語言(Body Language)來進行溝通，這種通用的溝通方式如果應用在科技方面，這樣大多數的人都能夠接受其操作。

隨著多媒體資訊科技的進步，如數位相機、數位攝影機、掃描器、照相手機等科技產品以及網際網路的普及，使得數位影像隨手可得，也研發出更多相關於影像處理的技術，而影像辨識就是其中的應用。

手部正是人們用來進行肢體溝通上最重要、也是最容易清楚辨識的部分，因此手勢辨識系統因應而出，相關應用在於人機介面設計、醫療復健、虛擬實境、數位藝術創作與遊戲設計等方面都有所發展。

在現實中，許多年長的人不會使用電腦科技產物，那有什麼方式可

以幫助他們快速學習呢？用人類與生俱來的才能或許是最適當的，對於手勢辨識這方面的技術，我們認為值得去做深入的研究；但要如何達到「科技始終來自於人性」的願景，便是近代人機互動介面發展的重要關鍵。

近年來的科技發展不斷地朝人機介面的目標前進，而影像辨識成了重要的研究區域。我們所談到的影像處理系統其實就是一種以影像資料為基礎加以發展，其主要目的是要讓人的生活或工作更加便利，如：人臉辨識(Face Recognition)系統、保全監控系統、看護系統等。人臉辨識系統是使用影像擷取器，如數位相機等得到影像後再取出特徵向量，並使用人工智慧或圖形辨識讓系統進行表情分析，可以廣泛應用到簽到或者是人與電腦互動式的教學；而保全系統則是結合人體動作特徵及人臉偵測(Face Detection)，以便於如果有犯罪情形發生的時候，可以快速的從影像當中找到嫌疑犯並加以逮捕；另外看護系統是一種應用到醫療方面的影像處理系統，其可以藉由病人房內所拍攝到的影像判斷病人目前的狀況，如果當病人處於危險的情況之下，系統便可即時通知值班的醫護人員讓病人可以得到妥善的照顧。以往的影像系統的發展，通常都侷限於影像擷取器所得到的資訊不夠清晰導致系統的效能受到限制，然而現今隨著數位相機、數位攝影機效能的提升，這項缺點已經漸漸被

克服。

這些系統不僅能按照使用者的需求，精確地接收資訊並執行動作，這都是人機互動介面技術的一大躍進。在目前的人機互動系統中，人雖然被稱為使用者，但卻沒有主控系統的反應能力。因此在研究未來的系統時，人將會是主動的參與者，電腦將會針對人的各種動作而做出反應。人機介面的發展趨勢，終將會從以電腦主控為中心進而轉變成以人為中心。[2]

## 1.2 研究動機

在許多操作機器、家電、遊戲等介面，使用者除了傳統上使用按鍵或紅外線遙控來控制系統的方式之外，也另外提供了一種使人與機器更加自然的溝通方式，使得使用者在操作上更加具有便利性及真實感[3]。我們希望可以研究出讓使用者能直接以人體的手勢及姿態來控制機器介面，進而簡化操作流程。因此必須透過手勢辨識之影像處理來記錄手部動作，比對資料庫已有的手勢，再辨識出使用者所要求的指令，傳達到所連接的硬體並作回應，以達到人機介面互動的效果。

我們觀察到不論是在課堂上的教學或是正式的演講當中，主講人在台上演說時必須一邊操控滑鼠切換頁次，才能使整個演講流暢地進行，

但有時這樣的舉動會造成主講人不經意的分神，而無法達成完美的演講，因此也有些人會請旁人操控滑鼠來輔助演講，但其效果通常不彰；因為協助者與主講人若沒有十足的默契，便很難得知主講人的進度為何，時而過快、或時而過慢，所以切換頁次的時機便要拿捏得當方能順利進行，卻也要花些許練習時間以培養彼此的默契，這對主講人也是相當的不便，而無法暢所欲言。

但若能以手勢來取代點滑鼠以達到切換頁次的動作，且亦能透過手勢自然地擺動，除了能展現出演講時的自信，還能成功的吸引住聽眾的目光，這樣對整場演講而言會有加分的效果。

藉由影像辨識的輔助之下，我們希望能發展出一套「手勢辨識簡報播放系統」來針對在教學或演講時，以手勢取代滑鼠控制的方式來改善傳統控制方式的不便。

而手勢辨識(Hand Gesture Recognition)是目前熱門研究的主题，應用的範圍非常廣，包括機器人操控，遊戲，或是家電的控制等[4]。因此，除了能偵測到手勢的位置、型態進而加以辨識之外，要如何正確地讀取手部資訊，以及考量到時間及空間的變化，也是我們影像處理中要解決的重要議題。

在辨識之前有一個重要的步驟就是在於影像前處理的部分，傳統上

用來處理色彩長條圖的方式大多採用均等量化法或查表量化法，這些方法的缺點就是在於量化的誤差過大，會使得影像比對時容易產生錯誤。而且若採用 RGB 的色彩模型來處理顏色，則是易受到亮度的變化 (lighting intensity changes) 干擾而誤判。另外，若只考慮顏色上的比對，有可能會使顏色的分佈區域相近，但就會有影像內含的物件卻是不相同的情形發生。

還有一個要考慮的要點是在於如何將圖像中的物件或是其組成的特徵分割出來，就是指在影像切割的部分。如何能準確的、正確的切割出所需要的資訊，是在許多影像分析應用上一項非常重要的影像處理技術，因為切割出來的物件及特徵可以當成描述和辨識物件特徵的功能。但絕大多數的情況下，並非所有資訊都是我們所要處理分析的。利用分割的概念將影像中具有意義的資料取出並做有效的處理，這一動作決定了影像分析最後的結果。所以本專題研究希望除了能實現以手勢播放簡報的功能外，更能夠找出最合適影像辨識的處理技術。

### 1.3 研究目的

由於電腦與人類的日常生活愈來愈密不可分，能夠擁有更方便且更具人性化的新型人機介面，取代現有鍵盤及滑鼠等輸入裝置，是多數人

心中之期盼。希望這次的研究，使後續有愈來愈多研究者能投入新型人機介面的研發工作，提出更新穎的構想，讓大眾能早日享受到更直覺、更方便、以及更具人性化的新型人機介面裝置所帶來的便利。此項研究，是基於除了用手去控制滑鼠外，還可以用手勢去操作簡報播放的一種技術。

為了避免環境因素如光線、背景、相機等的影響造成膚色、手部偵測錯誤，或高計算量的需求限制，因此我們的主題研究需要一個平價且列為標準配備的攝影機，以及低計算量的人機介面。

在選擇人機介面的部分，我們認為簡報的環境較為單純，因其多半以螢幕輸出為主，在背景相減的部分較無與膚色相近的問題。而我們最終的目的，便是設計出一套以虛擬滑鼠為主的「非觸碰式操作系統」，讓使用者能夠自在地運用手勢操作，以改良教學或演講品質。

本專題目的就是要能設計出一整套人性化的人機控制介面，並不是僅限制於在使用者在操控簡報系統而已。只要舉起手或比個手勢，就可在不需拿取任何控制器來遠端遙控機器，舉凡家庭電器、娛樂設備、室內控制等，這可將此技術運用到極致。以單一攝影機達成擷取當前的手部動作，操作方式既簡單，在此將性質相近的操作又可以以同樣的動作來代表。不論是大人、小孩、各種身材各種年齡層的使用者皆可順利使

用。最終我們希望能讓系統有相當的強健性，屏除背景雜訊，而不受到背景的影響。

#### 1.4 研究重點

我們的研究重點是在於如何成功地將使用者與內建資料庫定義出的手勢進行辨識後，比對出相對應的指令，再透過電腦程式，利用游標的(X, Y)坐標，傳達到電腦系統負責控制滑鼠的位置，並模擬滑鼠的動作以操控電腦螢幕上的簡報畫面。例如：指定滑鼠至特定的位置、切換上/下頁、顯示右鍵的功能並選取等。並且，我們為了使用上的便利性，程式將在背景執行。

為了增加辨識的成功機率，在研究過程中有以下重點為需考量到的必要條件，因為人體的膚色在色調裡屬於暖色系，與背景色彩的差異性並不大，所以在一般的場合中較難辨識。因此我們必須以戴手套的方式來遮住手腕以上的部位。在試用過白色、橘色及草綠色的手套進行影像辨識之後，發現綠色的效果最好。因為屬於白色的物體太多，反而會在取特徵的過程中發生誤判，而橘色則是仍太接近人體的膚色。因此我們最後是採用較適合成為對比色的草綠色手套，以增強手勢影像在背景中的顯著性，使之差異性更大，更容易進行影像的辨識。

我們採用最基本的影像擷取方式，便是只需架設一台普通功能相機，透過影像傳輸至電腦立即作影像處理。

## 1.5 研究範圍與報告書架構

本次專題研究的範圍是以影像辨識為主，從影像的前處理到偵測手勢，經辨識結果後判讀出該使用者欲執行的指令並傳達至電腦中，藉此以手勢取代滑鼠的動作來操作簡報系統。

本篇論文主要分為五大章節貫穿全文，以下為各章節概述：

- A. 第 1 章為緒論，主要在於簡述影像辨識的背景及研究的目的與動機，並列出幾點條件為此研究當中的要點，以利作出較為精準的辨識結果。
- B. 第 2 章則為影像辨識的文獻回顧及相關研究技術的理論與探討，也包含了本研究當中會使用到的技術。
- C. 第 3 章為手形偵測的方法，包含著型態影像學各種原理與技術。
- D. 第 4 章為本專題技術成果，以 C++ Builder 6.0 開發操作介面來呈現。
- E. 第 5 章為結論，對本研究做個最後的總結，並期許未來能有更進一步的研究重點及方向。

圖一為本系統之流程圖。

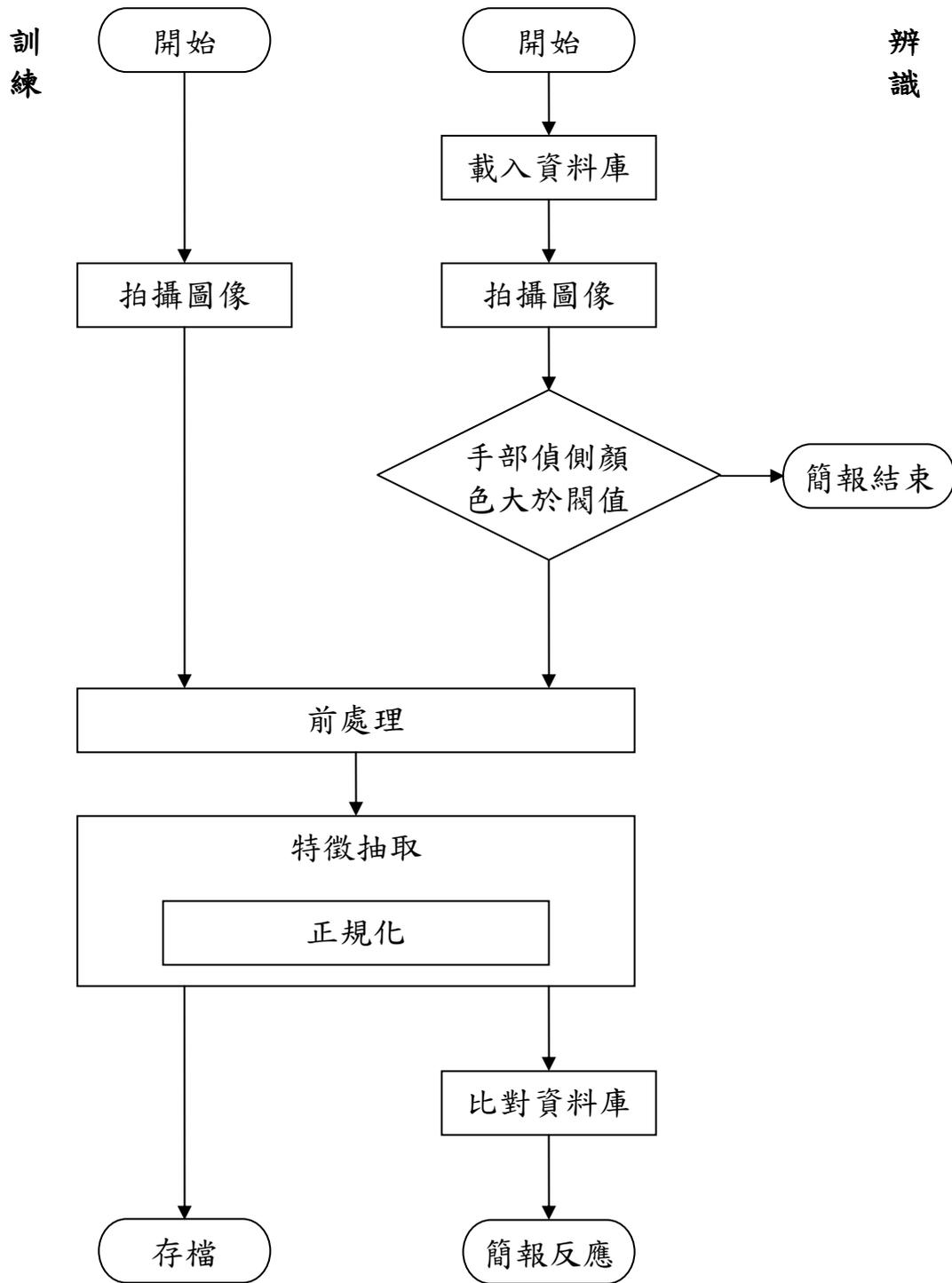


圖 1 專題研究架構圖

## 第貳章 影像辨識之理論及技術探討

### 2.1 研究理論及應用

近年來關於人體動作辨識的研究相當多，尤其以手勢辨識居多，此類研究可應用在機器人操作介面、家電操作介面及遙控系統等方面，未來能讓使用者隨時隨地輕鬆用手勢控制家電設備，是新一代數位住宅的構想。

而手勢辨識的技術在娛樂方面已經有很大的突破，透過手持裝置結合在虛擬實境的操作，最令人耳目一新的代表應用就是日本電玩業者於2006年所推出的家用遊樂器主機 Wii。

Wii 是以手持式遊戲搖桿 Remote 內建的 CMOS 晶片，在主機前方 5 公尺的活動區內感應到 Wii 主機上的 Sensor Bar 所發出的紅外線訊號，並將所感應到與主機的距離、Remote 的姿勢等訊息，藉由藍牙通訊傳回 Wii 的主機[[4]，使用者僅須揮動手勢便可模擬畫面中的動作。如此便利的遊戲設計，不只大受玩家的歡迎，就連不曾接觸電玩遊戲的人或小孩或女人，甚至是老人，都願意開始嘗試這種簡單的操作介面。

此辨識功能雖非以影像為主，但是以透過紅外線感應手部的動作及位置來操作電腦，仍屬手勢辨識的範疇。Wii 遊戲機改變了舊有的按鈕式或鍵盤式操作介面，直覺性的操作介面大幅增加了使用者的接受度，這

也更確立了使用手勢來控制介面是未來趨勢。除了在娛樂方面已出現影像辨識的應用，在治安、門禁或看護系統也有以人臉偵測之影像辨識為主。而在人機介面上也出現以手勢控制家電設備，甚至還可以遠端操作，而這些都是生活上俯拾皆是的相關應用。

手勢辨識方式可分成兩種，一種是以建立規則(Rule Based)為主，透過設計者事先定義好的特徵規則輸入至資料庫以供影像辨識，常應用在人工智慧的研究上。另一種則是以學習(Learning Based)為主的辨識方式，將蒐集到的資料樣本輸入至分類器(Classifier)，如 SVM，好讓分類器訓練參數並加以儲存為辨識之時可以用。

以下便來探討目前用來研究手勢辨識的相關技術：

### 2.1.1 手勢辨識方式

手勢辨識可以分成靜態手勢辨識(Static Hand Gesture Recognition)與動態手勢辨識(Gesture Recognition)兩種類型，靜態是以單一影像的手形來代表手勢，而動態則是以連續影像的手部位置移動來代表手勢。因靜態辨識屬於單一影像處理的關係，所以在研究技術及所需的設備較為簡單，以下列出較常使用的技術，並作簡單的介紹。

#### 2.1.1.1 HSI 色彩模型

RGB 為彩色顯示器所使用之色彩空間，由於相當易於瞭解，因此在繪

圖或呈現上多以此空間為基礎。如圖 2 為用來表示色彩組成的元素：

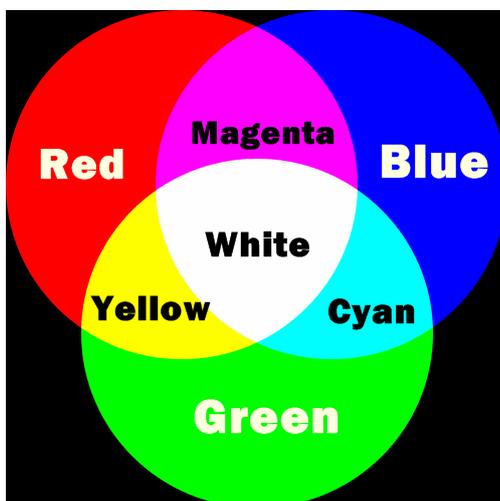


圖 2 RGB 的混色結果

然而 RGB 在表達現實世界中的光影變化並非為線性關係，亮度變化也會使同一種顏色在 RGB 色彩空間中呈現不同值，因此當環境的亮度改變時極易產生誤差。而且以 RGB 色彩空間做影像處理也較沒有效率[5]，這將會對影像造成辨識上的困難度。因此，在數位影像處理的過程中，像是在作影像辨識前置處理的部份，通常就會先將 RGB 的色彩模型轉換成 HSI 或者是 YCbCr 模型，再去作形態學的处理。

而 HSI 分別表示色相 (Hue)、飽和度 (Saturation)、強度 (Intensity)。色相是純色彩(純紅、純綠)的區分，是色彩由物體反射或透射出來的顏色，表示的方式為色彩在標準色環上的位置，以介於 0 到 360 的角度表示。飽和度則是表示顏色的強度或純度，也是灰色和色相的

相對比率量，表示的方式是從 0%到 100%。強度則是指顏色的相對亮度或明暗度，一樣是從 0%到 100%來表示[6]。

RGB 的座標系統的模型為正方體，如圖 3 所示。而圖 4 為 HSI 色彩空間的座標系統，是呈錐狀體：

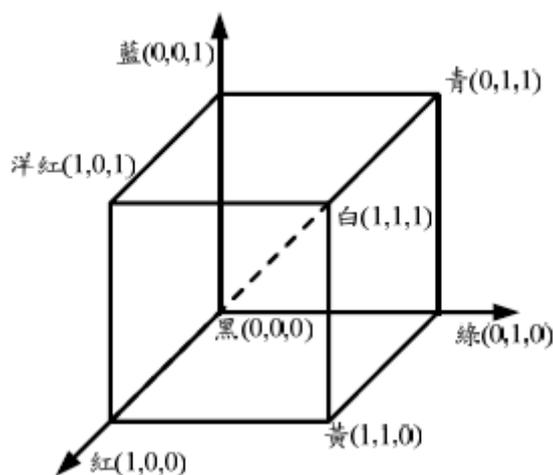


圖 3 RGB 的座標系統

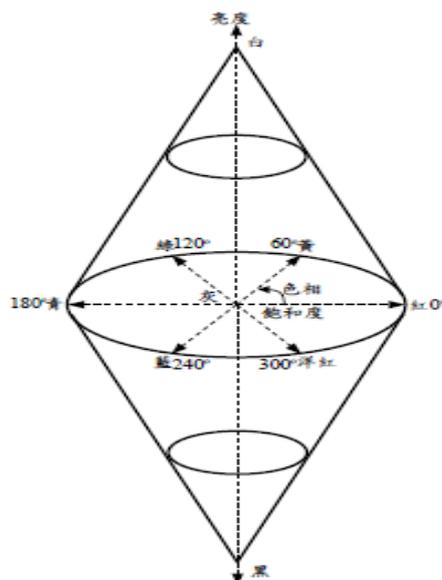


圖 4 HSI 的座標系統

此色彩轉換可以將亮度(luminance)和色度(chrominance)做良好的分離[7]。先將 RGB 轉換後，再來統計人各人種膚色的區間，統計出膚色會落在該色彩空間的那個區域內，以此區間來判斷出是否為膚色。而兩者轉換[8]的公式(1)~(4)：

$$H_1 = \text{COS}^{-1} \left\{ \frac{\frac{1}{2}((R-G)+(R-B))}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right\} \quad (1)$$

$$H = \begin{cases} H_1 & \text{if } B \leq G \\ 360^\circ - H_1 & \text{if } B > G \end{cases} \quad (2)$$

$$S = \frac{\text{Max}(R, G, B) - \text{Min}(R, G, B)}{\text{Max}(R, G, B)} \quad (3)$$

$$I = \frac{\text{Max}(R, G, B)}{255} \quad (4)$$

因為此模型較易於區別各種不同顏色的區域，並有利於我們之後取影像的邊緣特徵，所以本系統採用的是 HSI 色彩模型。

### 2.1.1.2 邊緣偵測

影像的邊緣乃是兩個不同區域之間的連接像素的集合，也是不同的幾何物件最明顯的區隔，因此邊緣就成為影像當中各幾何物件重要的特徵之一。邊緣偵測其實就是求某一個方向的一階導數，即影像  $f(x, y)$  在  $x, y$  方向的梯度變化量。如公式(5)所示：

$$\nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \quad (5)$$

在眾多方法中，Sobel 邊緣偵測法與 Canny 邊緣偵測法(Canny Edge Detection)最常被應用[9]。

先將原始影像先經過長條圖等化處理，然後以 Sobel 演算法或 Canny 來進行邊界偵測。經過長條圖等化處理後可以強化影像的對比，改善亮度不佳的問題。然後再透過邊緣偵測，並且只要適當調整偵測臨界值，便可得到理想的結果，即完成了前處理的作業流程。

Canny 為 Sobel 改良後的演算法，是 Canny 於 1986 年所發表的一種萃取目標物特徵的方法，對於一個對影像高頻部分具有很高敏感度，適合處理較精密的影像。主要是利用影像當中邊緣位置的梯度值通常較週遭來的大的特性來偵測[10]。Canny 的目標是找到一個最優的邊緣檢測演算法，故採 canny 邊緣檢測時是使用高斯模糊消除雜訊，並估計圖像的背景，並創建結構元素，再用該結構對原圖像進行數學型態的操作，估計背景亮度，並使用閾值操作將圖像轉換為二值化圖像。

此偵測法需用到較複雜的演算法，在計算的過程上也花費較多的計算時間，通常適用於處理較多細部雜訊的人臉偵測系統上。再加上 Sobel 本身就有不錯的偵測邊緣效果，故本系統便不使用 Sobel 改良後的 Canny。

### 2.1.2 手勢辨識設備

手勢辨識的系統依輸入裝置可分為以影像為基礎的辨識 (vision-based approach) 及以手套為基礎的辨識 (glove-based approach) 兩類。

以影像為基礎的辨識使用攝影機擷取手部影像，再將影像手勢特徵化再進行辨識，其優點在於使用者不需在身上穿戴任何的額外裝置，因此使用者接受度較高；而相較於以手套為基礎的辨識則是使用資料手套來取得手指彎曲度資訊，再由彎曲度資訊進行辨識。然而，以影像為基礎來辨識手勢其運算量較大，由於必須先進行影像的前處理以取得手部特徵，因此辨識步驟較為繁雜。

另外，使用單一攝影機拍攝手部時，可能會有影像死角，無法取得完整手部資訊，因此能處理的手勢較少。如果加設攝影機數量，可取得較完整資訊，但相對運算量會大得多，對一般系統而言負擔很大，成本也會大為增高。

如使用資料手套做為輸入元件，雖然使用者必須額外穿戴手套較為不便。然而，資料手套可取得的資訊卻較為完整，運算量也較低，適用於即時辨識系統，也可以處理較多種手勢，多半應用於動態辨識系統上。以下便是要介紹手勢辨識系統常會使用到的硬體設備：

### 2.1.3.1 紅外線發射器[11]

紅外線常被應用於控制家中硬體設備，傳送指令至特定紅外線發射端，再藉由無線網路經紅外線訊號控制盒操控特定家電。當加入新家電的新功能時，家電遙控器對準主控端的紅外線接收器發送訊號，系統會將訊號格式編碼複製。往後選定功能，系統發送一模一樣的紅外線訊號來控制家電，如圖 5。

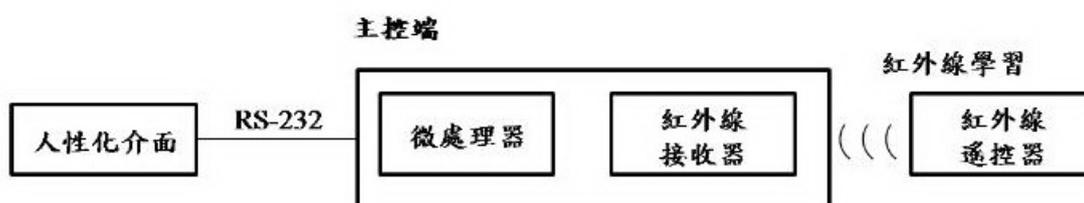


圖 5 學習紅外線機制[11]

另一個應用則是在於娛樂方面的家用遊樂器主機 Wii。在可偵測到的活動區內感應到紅外線的訊號，藉由藍牙通訊傳回 Wii 的主機，便可模擬畫面中的動作。

### 2.1.3.2 攝影機(Video Camera)[12]

利用攝影機擷取手部或配戴於手部的感測器在空間中影像，還可以連續拍攝影像方式提供電腦計算出平均值的影像，對使用者而言是較為直覺的方式，也是目前最常見且花費成本較低的一種設備，常應用於靜態辨識系統。然而若要獲得較完整的資訊，所需付出的代價為運算量的增加及辨識速度的降低，這些需要硬體的支援及良好的演算法配合。

### 2.1.3.3 資料手套

主要是為了偵測使用者手部方位（X 軸與 Z 軸兩自由度）及手指彎曲度的功能。資料手套上鑲有兩軸（X 軸與 Z 軸）方位感測器，其功能為感測手部方位的改變，測量角度在 $\pm 60$ 度時呈線性變化。在資料手套內則貼附著光纖並在每隻手指安裝一個感測器，測量手指的平均彎曲量，將測量值傳回電腦做運算並反應成相對的動作。因為能夠精準抓住手部動作，所以適用於動態的手勢辨識系統，但因考量到成本及使用不便的因素，故在本專題的硬體設備需求上並不適用。

## 2.2 技術 1 探討 手勢辨識系統[[9]]

技術中採用 RGB 色彩模型轉換成 HSI 的模型，原因為數位影像的膚色像素值投影在 HSI 色彩模型時，會在 HSI 座標軸上形成一個近乎橢圓的幾何形狀，如圖 6，利用這個橢圓可以明確且輕易地區隔出膚色部分與非膚色部份。

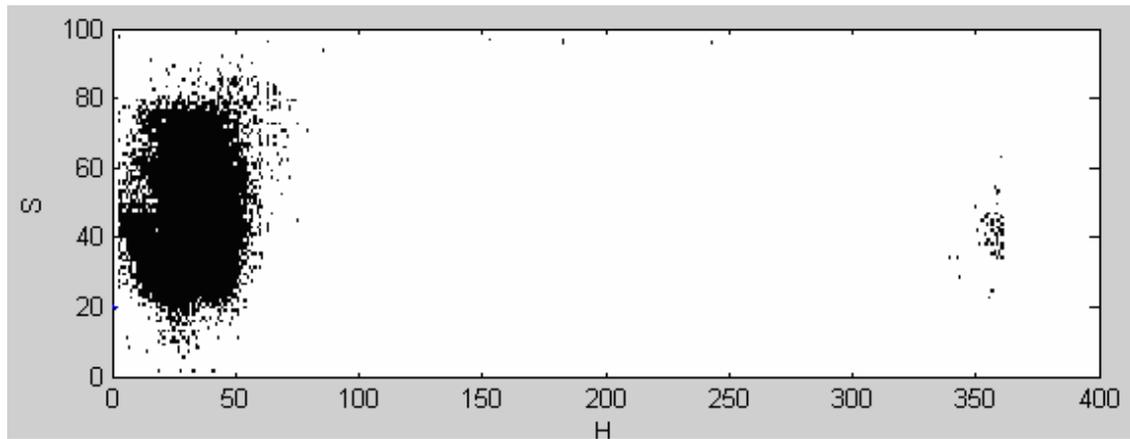


圖 6 膚色以色調與飽和度表示之分佈區域[9]

輸入之影像經過 HSI 轉換後，經過統計可發現膚色分部都集中在  $H\_Range[0, 56]$ 、 $S\_Range[15, 83]$  這組範圍內，可利用此一範圍多次分析測試後找出一組可靠門檻值，往後將以此門檻值為基礎作為偵測手勢之依據。

由於 RGB 轉換 HSI 公式將像素用 HSI 表示，初期所得到的膚色像素點會有因為膚色誤判而被保留或缺少，為了彌補這些像素點，研究中在膚色偵測之後再經由型態學當中的斷開和閉合技術，讓多餘的點消失，缺少點補回一些。圖 7 為膚色偵測流程圖。

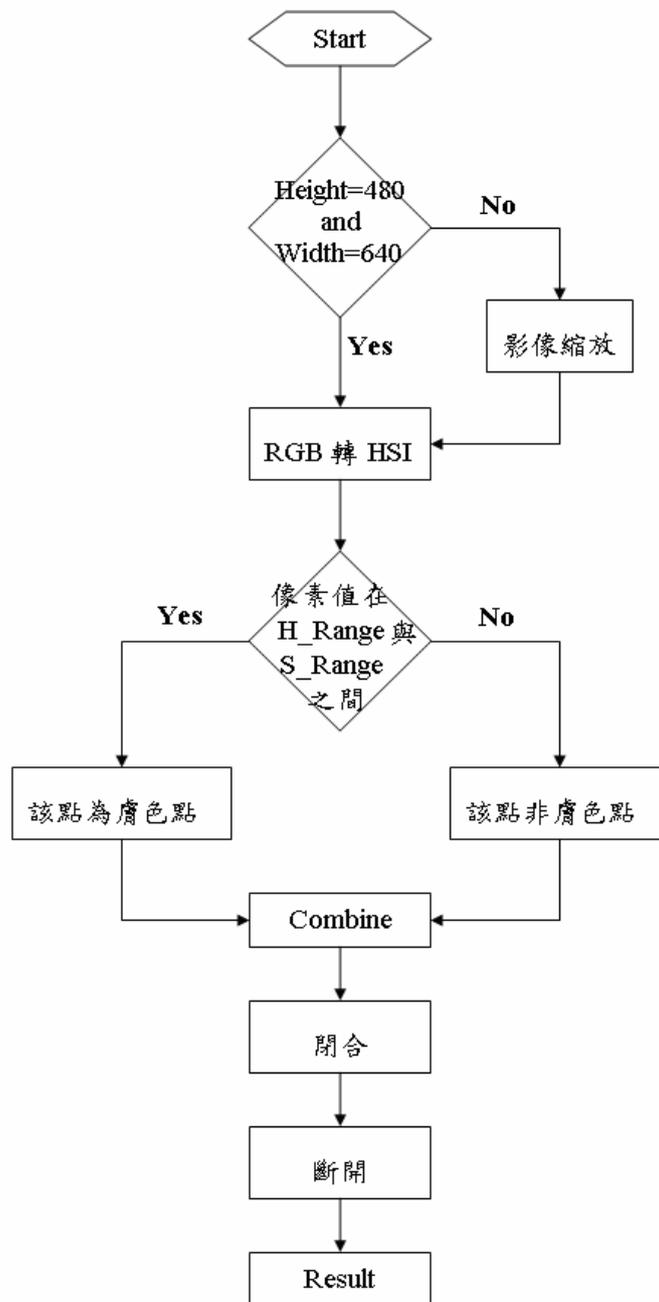


圖 7 之膚色偵測流程圖[[9]

在特徵擷取這方面，此技術是利用手部大小正規化來處理，為了要準確的尋找相對位置，必須先將手勢做轉正的處理，而為了轉正則需要用到最小平方線的技術，最小平方線是一種在數值方法中用來計算出座

標軸當中各數值最接近的直線的方法，在系統當中用來計算出手部物件的代表直線，接著利用這條直線所代表的角度算出物件的角度，再根據這個角度將影像加以旋轉以期手掌影像可以旋轉到接近垂直的位置，如圖 8 所示，將手勢轉正之後，利用手部外圍輪廓波形掃描手勢的輪廓。幾何物件的外圍輪廓波形掃描是利用幾何物件的中心點，掃描  $360^\circ$  求出各角度物件輪廓點與中心點距離所形成的波形構成，如圖 9 所示。



(a)原圖



(b)結果圖

圖 8 運用最小平方線到手掌影像範例圖[9]

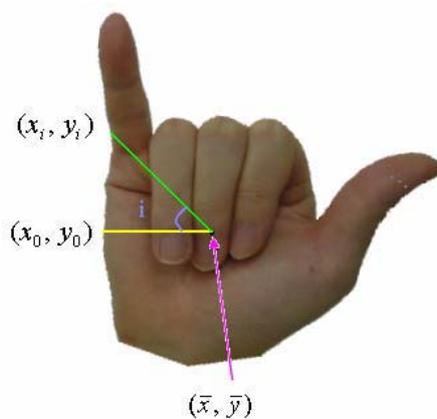


圖 9 手部外圍輪廓波形掃描示意圖[9]

此技術之缺點有三：

- A. 由於系統只定義七種手勢，故所取得的影像必須在七種手勢當中，如圖 10，如果測試影像是在已經經過訓練的資料群組當中，整體辨識正確率可以達到 95% ，而如果測試影像不在訓練資料群組當中整體辨識正確率為 87% 。

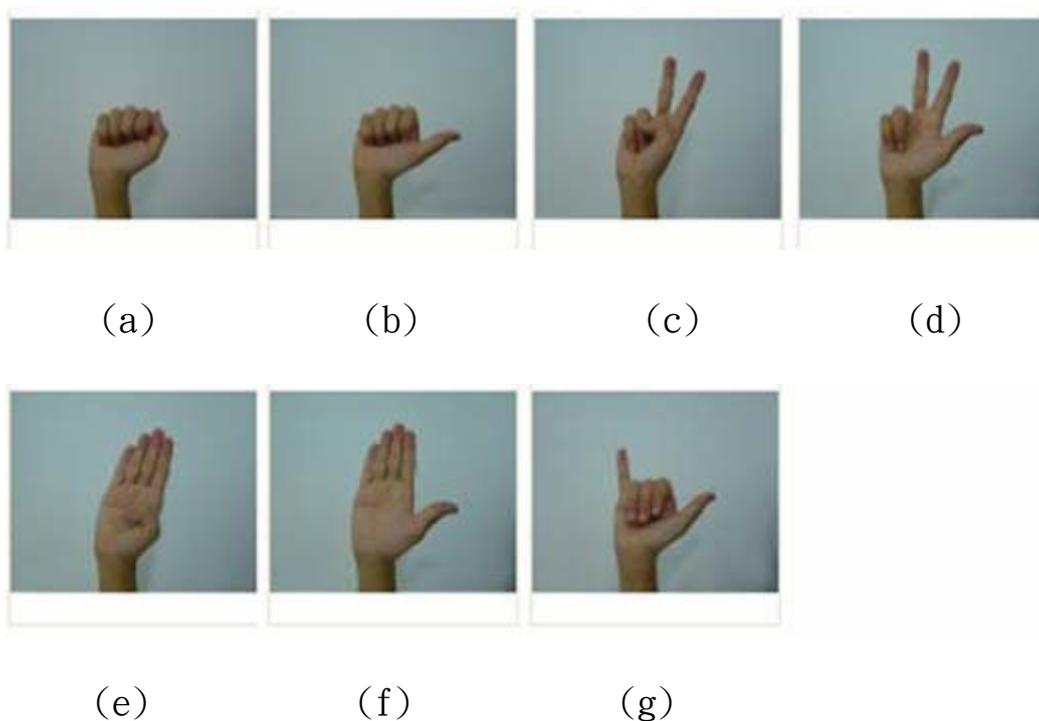


圖 10 手勢辨識系統的七種手勢範例圖[9]

- B. 輸入的影像其拍攝時的環境照明必須要充足。
- C. 手掌必須是右手影像，且拍攝的時候手部中心軸允許些微的傾斜。

## 2.3 技術 2 探討 電腦互動技術實驗課程設計與教學手冊[[13]

此技術主要是希望能藉由一個觀念簡單的實驗，以及一些簡單的影像處理，以協助各位了解電腦如何藉由攝影機與環境產生互動。技術中是以根據影像資料建立一組背景，再藉由比對擷取的影像與建立的背景之間的差異，來判斷影像中是否有物件正在進行移動。

在手勢偵測時，因為影像中包含許多複雜的訊息，所以將影像大致區分為手勢區域與背景區域。

此論文技術是利用實驗的內容，是在一秒內連取十張影像，並取這十張影像的像素值平均值作背景，如圖 11。

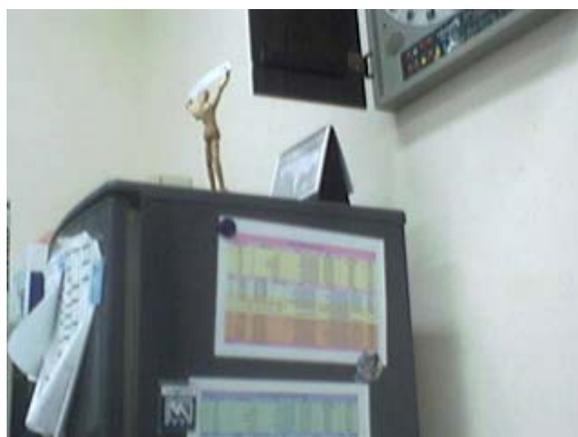


圖 11 經由平均結果所得之背景[13]

有了背景之後，便可持續的將目前擷取裝置擷取到的影像跟背景相減，當目前擷取到的影像如圖 12，不同於目前的背景時，則畫面中新進入的物件，其相對位置上的像素值與背景的像素值之間的差異便會相當的大，當此差異值大過一個門檻值後，我們便能認為此位置的像素與背

景不同，並且將此點設為白色，顯示在程式的右邊畫面中，如圖 13 所示。



圖 12 物件進入畫面中[13]



圖 13 白色部分是系統偵測到的物件變化[13]

雖然可以藉由上述的方法偵測到物件的移動，但是在通常的情況下，物件的擷取成果並不會如圖 12 所示的那麼漂亮，經常會有一些雜訊散落在各處。為了要能準確的判斷物件的所在位置，[[13]的方法將再計算圖 13 中的 Connected Components，並將較小的 Connected Components 視為雜訊濾除掉。

此演算法的步驟如下：

- A. 自左上至右下逐一判斷每個像素，如圖 14 所示，X 所在處代表目前位置，A、B、C、D 分別代表 X 之位置的左上，正上，右上，以及正左四個位置的像素值。
- B. 若 X 位置的像素值為 1（即圖 4 中的白色處），則判斷 A、B、C、D 四個點是否屬於某個 Connected Component，若均不屬於任何一個 Connected Component，則將此點標示為一個新的 Connected Component。若此 A、B、C、D 中只有一點屬於某個 Connected Component，則 X 亦屬於該 Connected Component。
- C. 若 A、B、C、D 中，A、C 分別屬於不同的 Connected Component，則 X、A、C 均應視為相同的 Connected Component。同理，若 C、D 分別屬於不同的 Connected Component，則 X、C、D 也均視為相同的 Connected Component。

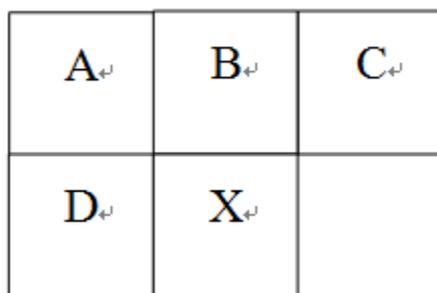


圖 14 Connected Component 演算法使用的 model[13]

完成了 Connected Components 的演算法之後，即可計算每一個

Connected component 的大小，除了前述所說，我們可以將較小的 Connected component 視為雜訊濾除掉之外，完成了 Connected Components 的演算法之後，也可以做許多其他的應用，例如，假設我們主要是偵測目前畫面中，體積最大的移動物件，我們擁有了每個 Connected Components 的大小資訊，便可以很容易的將目前最大的移動物件顯示出來，如圖 15 所示。



(a)物件進入畫面中

(b)白色部分是最大的移動物件

圖 15 偵測最大物件的範例[13]

但此技術的缺點是由於一開始只有建立一次背景，當環境產生某些變化（如光線的變化）之後，雖然景物本身並沒有改變，但由於光線的劇烈變化，就可能導致程式的誤判，如圖 16 所示。



(a) 正常的光線



(b) 全黑表示為正常的狀態



(c) 光線改變但景物不變



(d) 白色部分是系統偵測到的物件變化

圖 16 光線劇烈變化導致程式誤判的範例[13]

## 2.4 技術 3 探討 虛擬滑鼠：以視覺為基礎之手勢辨識[14]

這項技術是以視覺為基礎之手勢辨識，所採用的手勢較為簡單，因為眾所皆知，滑鼠，只有三個鍵，左鍵、右鍵以及中間的捲軸鍵。此虛擬滑鼠的構想是以左鍵及右鍵做為研究，要二種手勢取代這二個按鍵，且讓軟體能夠辨識出使用者的手勢並操作滑鼠。

### 2.4.1 手勢辨識的規則

手勢開始時先以一個較慢的速度從起始位置出發，繼續以一個較快的運動，最後結束回到手勢開始時的起始位置。當手在運動中，假設其

處於某個特定狀態。當手靜止時，個別手指的運動有他的意義在，當然手勢必需被限制在一個固定的活動範圍內。

#### 2.4.2 影像分割的規則

影像分割後的區域應具有單一性，每一區域應代表相同之性質，且所有的單一區域之聯集組成原影像。分割的區域結構應儘量簡單化，且要避免區域內部有太多小洞出現，分割後，相鄰之兩個區域彼此間應有明確之差異性性質存在，區域之邊界和邊緣應簡單清楚，可供辨識。

#### 2.4.3 特徵擷取

簡單來說，就是要從影像中辨識出手勢的動作，就必須先從影像中擷取及追蹤某些特徵資訊，如手掌中心位置、手掌中心位置之水平與垂直移動量、指尖位置、以及指尖位置之水平與垂直移動量等特徵資訊，大部份以特徵為基礎的影像比對技術，多半是用邊、點、區塊等特徵，與其周遭鄰近像素進行比對。

#### 2.4.4 手勢分析與結果

常用的手勢分析為馬可夫模型以及隱藏式馬可夫模型。在過去的研究中，隱藏式馬可夫模型已被成功應用於語音辨識上。近年來，它也逐漸被應用於影像辨識和電腦視覺之研究中，如手寫字辨識、人體動作辨識、手語辨識、以及表情辨別等方面。

以視覺為基礎之方法進行手勢辨識的研究，實作方法及操作步驟大致分為以下五個部份進行：

- A. 透過 CCD 取像系統擷取連續手勢影像。
- B. 將手勢從影像中分離出來。
- C. 擷取特徵資訊辨識手勢動作。
- D. 決定手勢動作所代表之滑鼠事件。
- E. 傳送訊息控制系統滑鼠動作。

有好的手勢影像，當然就要有好的取像環境，這裡是使用一張安裝於電腦上的數位影像擷取卡，連接一台彩色 CCD 攝影機，透過安裝於電腦上的取像軟體，如圖 17 所示，進行影像序列攝取工作，此系統每秒最高可擷取 12 張彩色影像。整套設備設置於一隔絕外界光源之暗房內，如圖 18 所示。

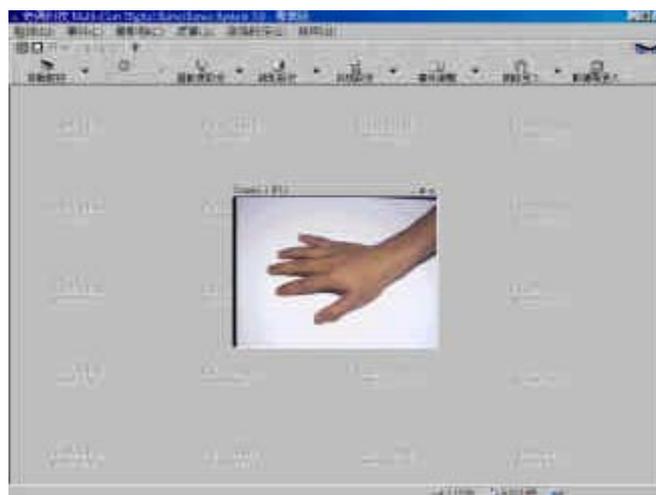


圖 17 攝取影像序列之取像軟體[14]

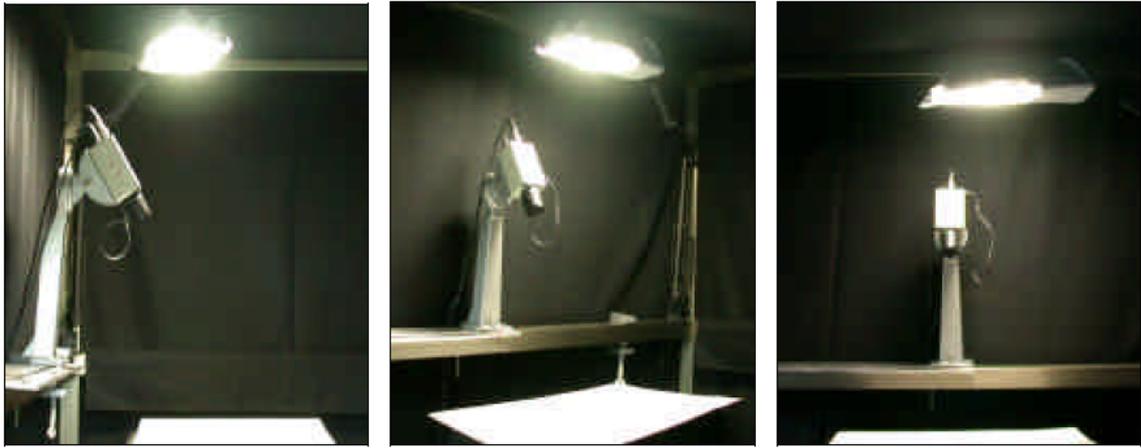


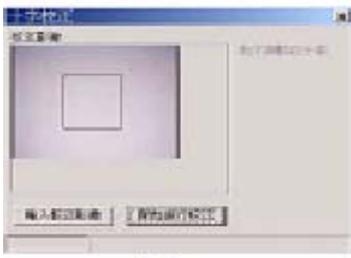
圖 18 取像環境[14]

有好的取像環境，如果沒有再三確認所拍攝的影像是否正確，那再好的環境及設備也是沒有用的，所以之後必需要針對拍攝出來的影像做校正的工作，方法是在一張白色硬紙板中央畫上紅色十字記號，並將這個標有十字記號的紙板平放於拍攝序列的桌面上，其中垂直交叉之水平線與垂直線之長度皆為 20 公分，如圖 19 所示。

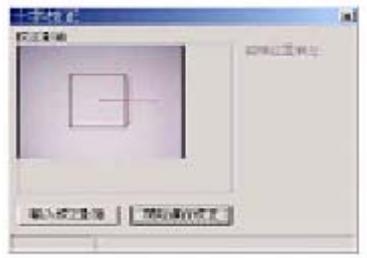


圖 19 為校正之紅色十字記號[14]

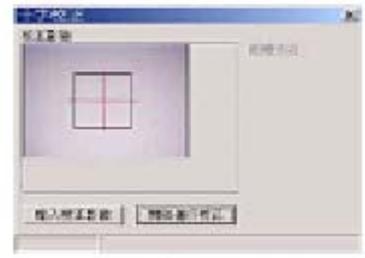
以下的六張圖，為圖 20 的測試情形。



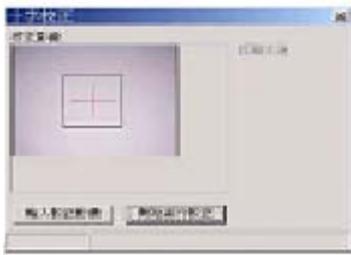
(a)無十字記號



(b)水平位置偏左



(c)垂直位置過近



(d)垂直位置過遠



(e)多種位置偏移情形



(f)設定正確

同時發生

圖 20 攝影機位置校正之情況[14]

在影像分割方面，由於使用者的桌面環境中可能包含鍵盤，而鍵盤也是電腦的輸入裝置，因此此項研究限定，當人手移動至鍵盤區域時所做的動作皆視為無效且不做處理。所以在進行影像分割時，還要辨識影像中是否存在鍵盤，若存在鍵盤則需將其分離出來，並計算其所涵蓋之範圍。

這裡採用的手勢為平常控制滑鼠的動作做為輸入手勢：部分研究為了能夠清楚和快速地辨識出手勢指令動作，因而定義外觀差異較大的手勢做為系統輸入的手勢。如定義剪刀、石頭、布三種手勢做為輸入手勢；

定義伸展不同手指隻數做為輸入手勢；定義數字 1 至 5 之手勢做為輸入手勢。此項研究考量到多數人已習慣於平時操控滑鼠之動作，因此採用大家熟悉的滑鼠控制動作做為輸入手勢，讓使用者不需花費時間學習和習慣一套新的輸入動作，以提高使用者於操作上的便利性和使用上的接受度。

優點：此項研究只使用單一 CCD 攝影機做為取像設備。

缺點：由於 CCD 架設之位置與拍攝之角度等因素影響，造成系統無法從攝取之影像中，完整獲得欲觀察物體之所有特徵資訊，此即為遮蔽問題。

## 2.5 技術 4 探討 手勢辨識應用於遙控音樂播放系統[15]

此技術主要的目標就是能夠透過手勢可以輕易控制家電音響設備的運作。目標在於能夠發展一套可藉由手勢遙控的音樂播放系統，此系統分為二部份：

### A. 影像輸入

在硬體方面使用一台個人電腦與一台數位攝影機讀取輸入影像，軟體方面使用 Borland C++ Builder 6.0 當做此研究的視窗介面研發軟體。

### B. 影像辨識

又分為影像前置處理及手勢辨識兩個單元，在影像前置處理這個單元裡，我們首先要做的是背景相減法。

### 1. 背景相減法

主要是將背景影像與輸入影像在相同的影像平面空間中相減來取得差異影像。當系統一開始操作，我們就取第一張影像做為我們的背景影像，與隨後影像相減後，影像相同部分會被消除掉，差異部份會被保留。

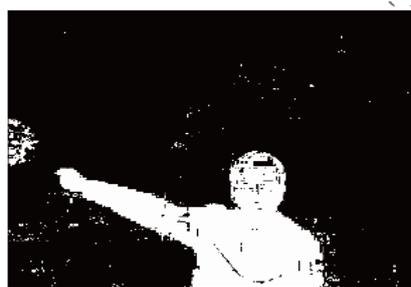
圖 21 為背景相減法的結果。



(a)背景影像



(b)輸入影像



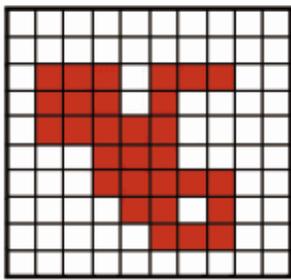
(c)背景相減後影像

圖 21 背景相減法[15]

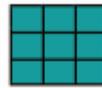
## 2. 小雜訊濾除與平滑化處理及區域填充

擴張算子能將區域鄰近的小雜訊合併在一起，遠端的雜訊仍會處於孤立狀態；侵蝕算子則區域旁鄰近雜訊仍會保留，但遠端的會被去除。

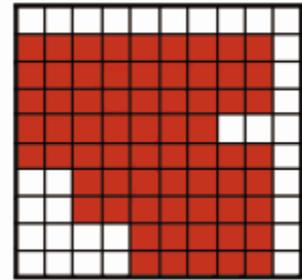
在此研究裡，做者先執行開放算子將目標物體外的細小雜訊去除，然後執行封閉算子修補目標物體中的微小碎塊、平滑整個影像。但擴張與侵蝕運算若疊代太多次，可能會造成原輸入影像的失真，所以若物體中存在較大的碎塊雜訊，則擴張與侵蝕運算將無能為力。圖 22 為擴張運算結果，圖 23 為侵蝕運算結果，圖 24 為封閉運算結果，圖 25 為開放運算結果，圖 26 為雜訊濾除過程。



(a)待測的影像

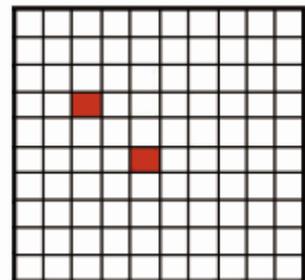
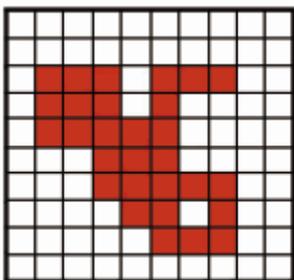


(b)3x3 的遮罩



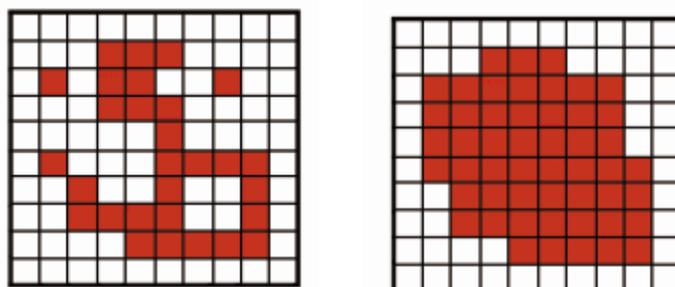
(c)被膨脹後的影像

圖 22 擴張運算[15]



(a)待測的影像                      (b)3×3 的遮罩                      (c)被侵蝕後的影像

圖 23 侵蝕運算[15]



(a)待測的影像                      (b)封閉算子執行後

圖 24 封閉算子[15]



(a)待測的影像                      (b)開放封閉算子執行後

圖 25 開放算子[15]



(a)待處理影像                      (b)開放算子處理後的結果



(c)封閉算子處理後的結果

圖 26 小雜訊濾除[15]

因上述提到物體中存在較大的碎塊雜訊問題，所以我們這裡用的是八鄰點的連通法來做處理。在此研究中將會擷取左手臂的影像，手臂寬度限定在 50 像素以下，長度在 60 像素以上。圖 27 為擷取左手臂影像的情形。



(a)待處理影像



(b)濾除身軀部位影像



(c)取出手臂部位影像



(d)留下左手臂部位影像

圖 27 左手臂影像擷取[15]

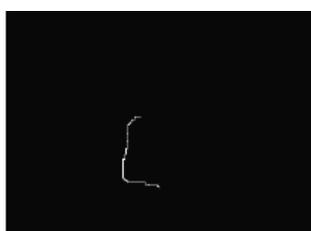
順利擷取左手臂影像後，我們要對擷取到的影像做邊緣化。圖 28 為邊緣化的情況。



(a)欲邊緣化影像



(b)側邊運算後



(c)取出的手臂特徵影像

圖 28 邊緣化[15]

接著，是影像辨識中的另一個單元，手勢辨識。

手臂的特徵影像可用兩條的直線來描述它，依據手臂的特徵影像呈現的位置與角度的相對關係來判定手臂姿態。在這裡我們使用的演算法是 Hough Transform。

Hough Transform 是一種直線偵測的演算法，主要是偵測影像中的直線線段，主要用途是在二值化影像中的輪廓形狀偵測。且越接近直線的位置與角度，亮點就越多，密度越大亮度也就越高。

當在此項研究中能找出一組可靠的動態門檻值的來判斷直線是否存在，往後將以此門檻值為基礎作為偵測直線的依據。在做完上述步驟後即可將處理過後的影像顯示在畫面中了。

## 第參章 手勢辨識簡報播放系統

本專題系統流程圖如下所示。其中共同的部份為：前處理、特徵抽取、正規化，以下章節將詳細介紹系統流程中的各個處理步驟。

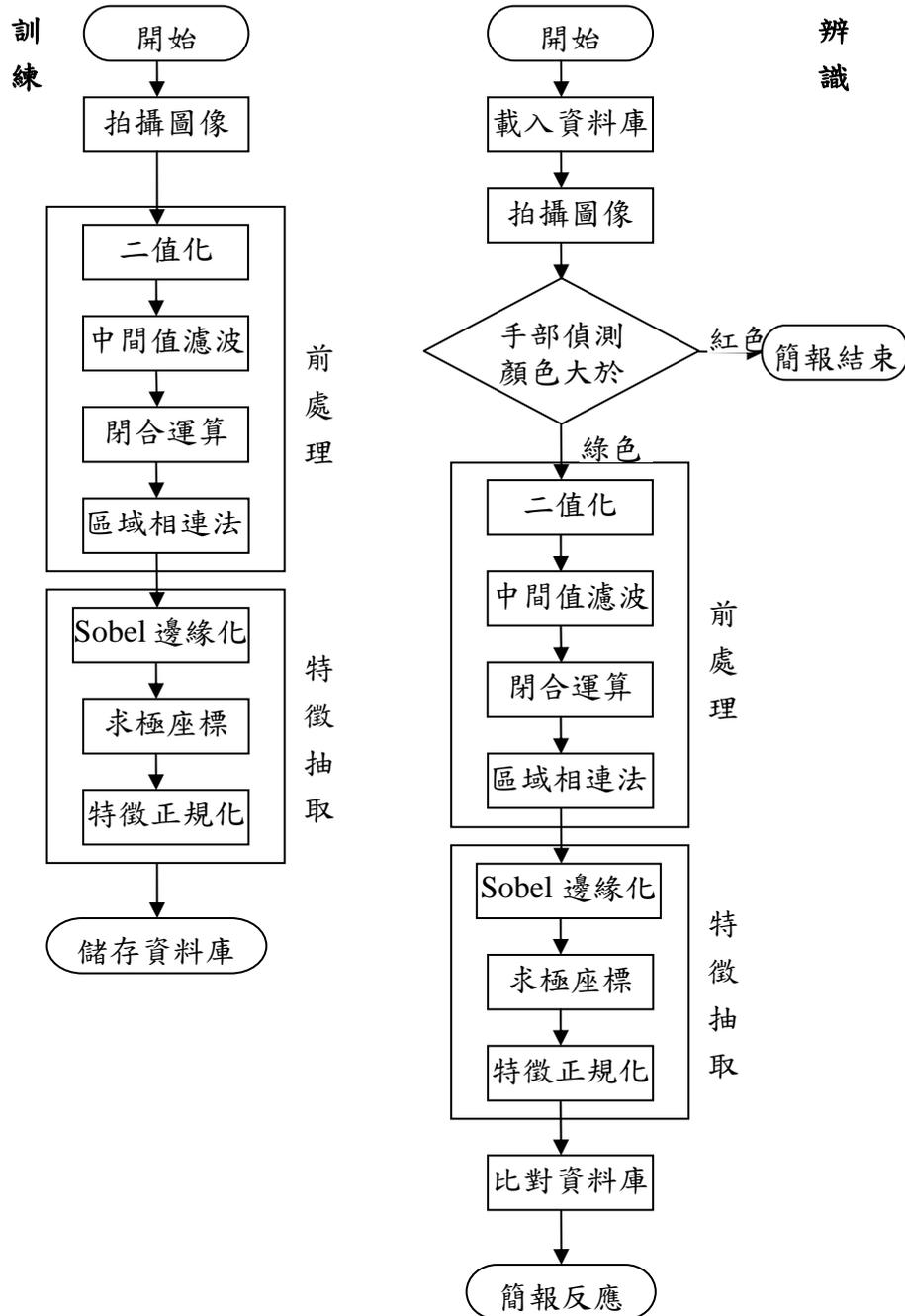


圖 29 系統流程圖

### 3.1 取得影像[16]

由於希望使用者播放簡報時，不必為了切換頁面，而影響演講的流暢度，所以用即時拍攝的方式，讓使用者只需比一個手勢就能達到切換頁面的目的。

想要取得系統所需影像，要先做以下兩個步驟：

- A. 取得訓練圖庫：我們的系統是利用特定的手勢變化去操作簡報 (PowerPoint)，致使達到簡報播放的成果，因此在使用這套系統之前，必須先設定四個手勢。因為每個人手的大小及擺放的距離都不太一樣，故先以三個人的手去做拍攝，一個綠色手勢拍攝三張圖，共九張圖留作資料庫建立，且這套系統需要戴上我們自製的紅綠手套。圖 30 為本系統所使用的四種手勢。



圖 30 本系統所使用的四種手勢

首先要先對影像做拍攝，戴上手套，三個手勢三個人共拍攝九張圖，以供手勢建檔，在按下開始拍攝後，系統會自動間隔 0.5 秒後拍攝一張圖，然後給予影像名稱，而所拍攝的這些影像，則用做辨識的

部分。

- B. 資料庫建立：這個部分，是後端的一個操作流程，主要是針對第一個步驟所取得的九張圖去做一個建檔的動作。我們會先對這些圖做特徵抽取，把所拍攝到的三個手勢九張圖的特徵值算出每個手勢的特徵值的平均儲存於記事本內[17]，建完檔後的記事本內會有三筆手勢的特徵值資料，這樣就完成了手勢建檔的部分，而這三筆資料就是辨識影像的關鍵要素。

## 3.2 影像前處理

在這個部分，我們要先將影像做一些基本的處理，使影像在辨識的時候能夠快速且正確的取得我們所要的部分。而在這一節當中，我們會講到影像的二值化是如何去做處理的，還有中間值濾波的部分，以及在影像辨識中幾乎都會用到的型態學，而我們是如何將影像做膨脹、侵蝕的動作，都將會在以下小節中呈現。

### 3.2.1 手部偵測

在進行手部偵測時，由於要偵測的目標的顏色是紅色與綠色，所以在攝影的畫面中，限制不能出現綠色與紅色，否則會影響系統的判斷。

在拍攝的畫面中，我們經由手套顏色的偵測，來判斷是哪一種手勢，

若是紅色的像素點大於綠色且紅色的像數大於閾值，則結束放映簡報，本統設定閾值為 1500 個像素點；若是綠色的像素點大於紅色且超過閾值，則進行下一階段的處理。

現今數位彩色影像絕大多數是利用 RGB 色彩模型描述影像中像素的色彩像素值，在偵測手套顏色時，若使用 RGB 色彩空間來表示的話，由於在拍攝圖像時，光源受到放投影片時而有環境上的限制，使得光源穩定度極差，在 RGB 色彩空間便會出現明顯的差異，因此 RGB 並無法有效的分離出不同顏色區域，所以本專題將利用色彩空間轉換至其他色彩模式下以期達到較好的顏色分離效果。色彩空間的模式有很多種，其中 HSI 為色相 H(Hue)、飽和度 S(Saturation)、亮度 I(Intensity)是根據人類對顏色的認知所建立的顏色模型，由於人的視覺對亮度的敏感程度遠強於對顏色濃淡的敏感程度，此外 HSI 有另一個優點，可以降低影像中色彩灰階度訊息的相關性。為了便於色彩處理和識別，因此我們選擇以 HSI 來進行色彩轉換。

以下先定義出手部偵測所用到的四個物理量：綠色色調範圍值(H\_GRange)、綠色飽和度範圍值(S\_GRange)、紅色色調範圍值(H\_RRange)以及紅色飽和度範圍值(S\_RRange)。

$$\begin{aligned} 320 \leq H\_RRange < 360 \text{ or } H\_RRange < 10 \\ 90 < S\_RRange \end{aligned} \quad (6)$$

$$\begin{aligned}
 50 < H\_GRange < 95 \\
 100 < S\_GRange
 \end{aligned}
 \tag{7}$$

影像經過 2.1 研究理論及應用中 RGB 轉換 HSI 公式將影像像素值以 HSI 表示，並由上列的公式色調與飽和度的範圍得到影像當中的手部像素點，若是綠色的像素點大於紅色且超過閾值，則將手部保留下來，其餘非手部的部分將以白色表示；但是此時所得到的手部像素點大多都會有因為誤判而缺少，如圖 31(a)為取得的原影像，31(b)是經由顏色保留後的結果，或是被系統判斷成手部像素點而被保留下來，我們可在下面章節的處理步驟進行修正。



圖 31 (a)擷取到的原圖、(b)系統保留綠色像素結果

### 3.2.2 影像二值化[18]

二值化又稱為灰度分劃(Threshold)，此動作將一般影像的灰度分劃成只有兩種灰度值，凡是影像本身灰度大於某一閾值便令其為『亮點』(白

色)；而灰度值低於閾值的，便令其為『暗點』(黑色)，如此可得到一個二元影像，將影像分為黑與白，使影像資料簡化以利往後系統運作。

單一臨界值法是目前最常用來處理影像二值化的轉換方式，假設一個二維灰階值影像的函數為  $f(x, y)$ ，是以 0(黑色)到 255(白色)為其值， $x$  與  $y$  則是影像座標，經過二值化後影像的函數為  $g(x, y)$ ，其函數轉換表示如圖 32 所示，則二值化的單一臨界值法的轉換公式(8)如下所示。

$$g(x, y) = \begin{cases} 255 & \text{if } f(x, y) \geq T \\ 0 & \text{if } f(x, y) < T \end{cases} \quad (8)$$

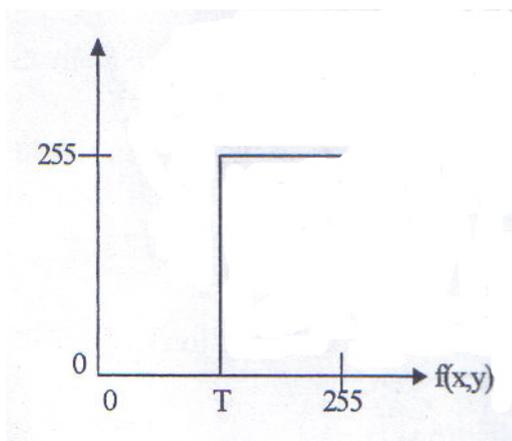


圖 32 圖影像二值化函數轉換

其中  $T$  便是代表臨界值，尋找臨界值的方法便是計算出整個欲轉換區域的灰階值後再取平均值，此平均值便當作是臨界值，與  $N$  分別代表影像的寬度與長度。灰階值為 0~255 共 256 組色階，0~ $T$  畫分為 0， $T$ ~255 劃分為 255。本論文的影像經上一節手部偵測後，非綠色區域值已劃分為 255(白色)，保留下來的綠色區域值不為 255(白色)即  $T$  劃分為 0(黑色)，

影像二值化結果如圖 33 和圖 34。



圖 33 二值化前的圖

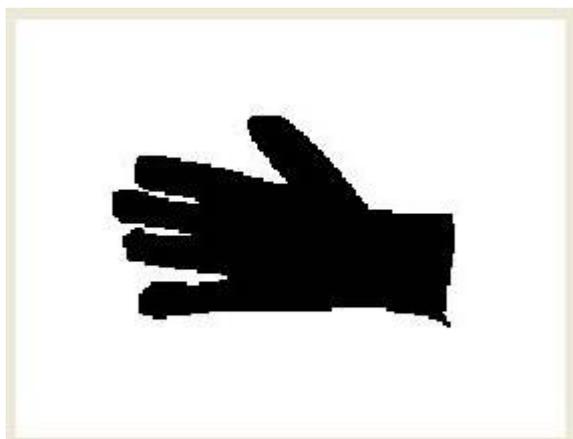


圖 34 二值化後的圖

### 3.2.3 中間值濾波[Median Filter]

在經過二值化處理之後，我們可得到一個二值化後的影像。影像是由像素點較多、面積較大的區域所構成，而雜訊往往是以孤立的點的形式出現，雖然這些點對應的像點數很少，但在我們訓練及辨識的過程中會被視為影像的有效像素點進而影響到辨識的準確率及資料庫建檔的正確性。因此，在數位影像處理中，常會藉由濾波器來去除雜訊，其中間

值濾波技術可以很平均的把影像重點部份表現出來[19]。

中間值濾波是以局部處理方式用來去除影像中的孤立點，並能保持原影像的銳利度[20]。中間值濾波能抑制影像中的雜訊，是一種很好用的工具。主要是用來除去影像中的雜點，雜點是以亮點的方式雜亂地分佈在影像中，因其灰度與周圍像素的灰度比通常是排在最亮的位置，所以取中間值可降低雜點的亮度。而它對影像最大的損害是使影像喪失原有的高頻特性，產生失真的情形。但對佈滿雜點的影像卻是值得的，因為它可獲得沒有雜點的影像。

其原理是在每個像素上放置一個  $m * m$  大小的框架，並利用該像素鄰域上灰階的中間值來代替，而不是以它的平均值來代替[21]。以下為處理流程及舉例說明。首先我們將以中心點劃分為  $3 \times 3$  的矩陣，其區域內的像素值為(21, 23, 22, 32, 33, 11, 110, 31, 22)，並假設其中有一個雜訊，像素值為 110。如圖 35 所示：

21	33	22
32	33	11
110	31	22

圖 35 以數值 33 為中心點的矩陣

再經由小而大排序後為 11，21，22，22，**31**，33，33，110，然後用  
 順位第 5 的灰階值，來取代中心點的像素值。依此類推，原始影像的每  
 一點均進行中通濾波處理，可得一新影像，其中可發現雜訊已被消除。

如圖 36 所示：

0	21	12
31	<b>31</b>	24
31	31	26

圖 36 使用中間值濾波後的範例圖

中值濾波器在影像處理上有其統計意義，當所取出的 9 個點有較多  
 偏暗的點時，也就是灰階值偏低時，則排序後的中間值亦會是偏暗的點；  
 反之，若 9 個點有較多偏亮的點，也就是灰階值偏高時，則排序後的輸  
 出結果也會是較為偏亮的點，因此中值濾波器可以強迫影像中各個點之  
 灰階值與其鄰點更為相近，達到抑制雜訊使影像平滑的效果[22]。因此，  
 中值濾波器的功能是強迫有不同灰階度的點更接近其鄰近點。消除雜訊  
 後的結果可得到較完整的手部區域，如圖 37 和圖 38 所示



圖 37 手勢影像的原圖



圖 38 二值化後經中間值濾波之結果

#### 3.2.4 形態學處理

在型態影像學中，所有的應用都是以「膨脹」、「侵蝕」兩種運算混合使用，如此便可達到欲達成之理想效果。

在偵測手勢時，有時會遇到待測區域之像素點不完整，可能是區域中像素太多，或是區域中像素破碎凌亂不堪，此時為了讓我們可以更準確的偵測出手勢，我們可利用膨脹(dilation)與侵蝕(erosion)來增減區域之像素點。

膨脹運算會以一點為原點，往它的八鄰居向外擴張，讓原本的像素

點往外擴張長出一層邊，將外型均勻的加粗，稱為「膨脹」；相反的，侵蝕運算會讓原本的像素點往內慢慢減少，將外型均勻的縮小變細，稱為「侵蝕」。如圖 39 所示，假設 E 為中心點，A~I 為其 8 鄰居，膨脹與侵蝕的定義互為相反，膨脹是只要 E 之周圍有一鄰居是 1，則輸出為 1，而侵蝕是只要 E 之周圍有一鄰居是 0，則刪除此點[23]。其膨脹與侵蝕公式(9)(10)分別如下所示，I 代表 AND 運算，Y 代表 OR 運算。

A	B	C
D	E	F
G	H	I

圖 39 8 鄰居範例圖

$$g(I, J) = E \cup (A \cup B \cup C \cup D \cup F \cup G \cup H \cup I) \quad (9)$$

$$g(I, J) = E \cap (A \cap B \cap C \cap D \cap F \cap G \cap H \cap I) \quad (10)$$

影像之開閉運算即是使用膨脹與侵蝕兩種方法的運算組合，開運算(opening)是將影像先侵蝕運算再經過膨脹運算來處理，可清除影像之雜訊，使物體輪廓平滑；閉運算(closing)則是先使用膨脹運算再經由侵蝕運算來處理，可用來填補影像內之輪廓上的破洞或將一些狹窄的中斷部分和長細缺口連接起來。本系統之型態影像使用方式，因之前步驟做過

影像相減後再經由手部偵測的處理，造成了影像較為殘缺不堪，為了配合後續之條件判斷，我們將先做一次閉運算來處理。

### 3.2.5 相連區域法[Connected Component]

在前面的處理步驟已去除大部分雜訊，但仍無法留下主要手部位置，所以本節將以相連區域法取出更正確手部位置[24]，此方法主要是利用 8 鄰居鄰接的方式，利用堆疊的原理將二值化影像中的相同區域連接起來，便可以清楚的區分出每個像素分別在哪個區域，並可獲得各區域的長、寬大小，以利於後續的條件判斷[25]。在影像中白點為背景，黑點為有效的影像，下面將介紹相連區域法之步驟[26]。

- A. 以影像左上角為起點，從左至右、由上至下掃描影像，找出第一個黑點，並將此點標示為已完成，以此點為中心，搜尋8鄰居，將8鄰居中未標示為已完成的黑點位置，存入堆疊中。
- B. 從堆疊中取出另一個黑點的位置，並標示為已完成，同樣以此點為中心，搜尋 8 鄰居，將 8 鄰居中未標示為已完成的黑點位置，存入堆疊中，如此一直循環。
- C. 直到堆疊清空，即可完成一次的相連區域法。
- D. 重覆的執行 1~3 步驟，直到影像中所有黑點都標示為已完成，即可找出所有的相連區域，如圖 40 所示。

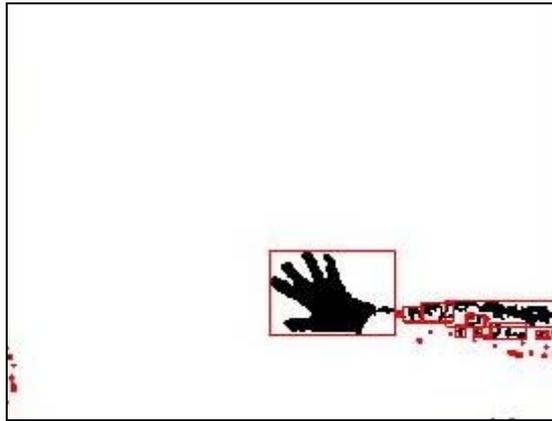


圖 40 一個紅色框標示出一個區域

在找出所有的相連區域後，針對每一個區域執行條件判斷，條件如下：

- A. 第一個黑點座標在(5, 5)以上。雜訊通常會出現在整張圖四邊上，可以藉由此條件過濾掉一部分的雜訊。
- B. 符合長 X 寬大於 5500 像素。手部是一個連續區域，範圍大小會有一個最小的門檻值。

符合上述的條件，便是我們要的手部位置，如圖 41 所示。

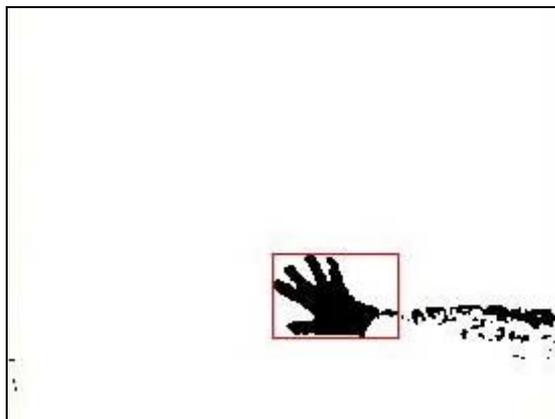


圖 41 條件判斷後，紅框為我們要的手部區域。

在確認手部位置之後，即可進一步找出手部位置的重心座標，此重心為 3.3.2 特徵抽取-極座標時所需要。計算重心如公式(11)：

$$G(x_p, y_p) = G\left(\frac{\sum_{i=1}^{tp} x_i}{tp}, \frac{\sum_{i=1}^{tp} y_i}{tp}\right) \quad (11)$$

$G(x_p, y_p)$  為重心座標， $x_i$  為紅框內所有黑點的 x 座標， $y_i$  為紅框內所有黑點的 y 座標，tp 為紅框內黑點的總數。

### 3.3 特徵抽取

想要從影像中辨識出手勢的動作，就必須先從影像中擷取及追蹤某些特徵資訊，如手部的中心位置。下面小節將說明在手勢辨識中所用到的特徵擷取理論及技術。

#### 3.3.1 Sobel 邊緣偵測法

在本小節，我們使用 Sobel 邊緣偵測法，主要是擷取手勢影像中最外緣的邊緣點。Sobel 邊緣偵測法是使用微分的概念設計出一組 Sobel 運算子，使用邊緣化主要的原因是因為手勢外緣部份的影像擷取比較簡易、輪廓較明顯且較細線化法更能清楚的表達整個手勢的姿態，而取得邊緣影像的做法就是由左至右由上至下掃描影像，以此方法找出邊緣線。

圖 42，輸入的影像在經過與這組運算子計算之後會得到一組水平分

量  $G_x$ 、與垂直分量  $G_y$ ，再經由  $\nabla f \approx |G_x| + |G_y|$  求得梯度，圖 43 為經過 Sobel 邊緣偵測法之後的結果圖。

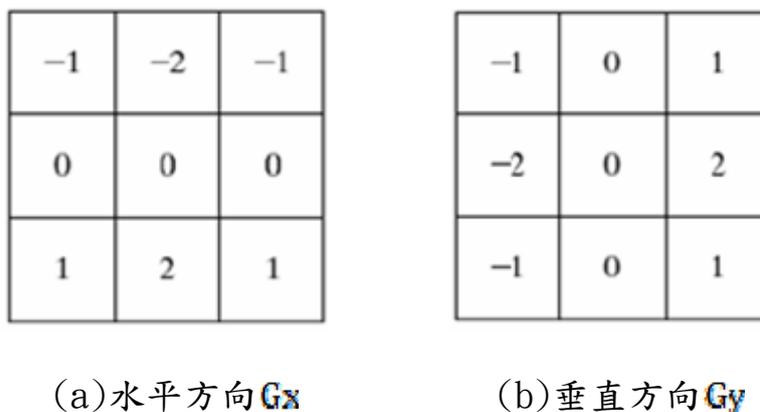


圖 42 Sobel 運算子



(a)為原始影像                      (b)為 Sobel 後的結果

圖 43 Sobel 邊緣偵測的結果

### 3.3.2 特徵抽取---極座標[27]

極座標(Polar Coordinate)是以向右的水平線為基準，該點與原點所成的直線距離  $r$ ，及與該線與基準線的角度  $\theta$  來表示該點的座標。為求得影像的極座標，就必須先求出手部的外圍輪廓及重心座標。

由 Sobel 邊緣偵測法對影像的每一個像素點作處理後，可得到此影像的邊緣圖，即為手勢的邊緣特徵。在 3.2.5 節所提到相連區域法主要

的功能在於取得較為完整的手部區域，並且在掃描的同時將每個相同標記值的點的坐標累加起來，再除以面積即可得到重心。我們以重心坐標為極點 P，再取一條由極點 P 出發之半線(Half-Line)稱為極軸(Polar axis)。通常極軸是由極點 P 指向右方的一條水平線，恰巧對應於直角座標中之正 X 軸(positive X-axis)。並使影像邊緣的極點 O 與極點 P 連成一線，令 r 為 P 至極點 O 之距離，若  $P=0$  則用  $(0, \theta)$  來表示 P。 $\theta$  則為起始於極軸夾角於直線之角度， $(r, \theta)$  便稱為點 O 之極座標(Polar Coordinates)，而  $\theta$  可為任何值。如圖 44 所示：

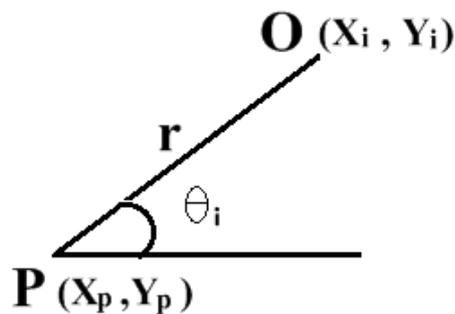


圖 44 極座標[27]

在經過 Sobel 偵測後會得到影像的邊緣輪廓，而輪廓的點有很多個，包含手指之間的縫隙也會被留下，但整個手部的範圍不會超過 360 度，因此偵測到的每個點還會先求出極座標的角度，然後同角度只會留下最遠的距離。這樣就可得到外圍輪廓的 360 個極座標，即影像中的特徵點(feature points)。

我們先以迴圈來掃描，從左上角的座標(0,0)開始，由左至右，從上到下的方式，一系列地去掃過圖檔，當判斷有黑點的時候，就可以求得 $(x_i, y_i)$ 的座標，然後再跟重心座標 $(x_p, y_p)$ 作距離運算。

而我們是採用歐氏距離演算法(Euclidean Distance)來計算各手部輪廓各點與重心點距離，如下列公式所示：

$$R(P, O) = \sqrt{(x_p - x_i)^2 + (y_p - y_i)^2} \quad (12)$$

R：重心座標與輪廓點座標之直線距離。

$P(x_p, y_p)$ ：即重心座標。

$O(x_i, y_i)$ ：即手勢所取得之邊緣點。

計算手的重心點到每個手部輪廓點的距離之後，我們再去以下列公式來求出每個線段與水平的角度。

$$\theta_i = \tan^{-1} \left( \frac{O_i(y) - Y_p}{O_i(x) - X_p} \right) \quad (13)$$

其中的 $O_{ix}$ 為下一個輪廓點座標， $\tan^{-1}$ 為反正切值，是以 $(y/x)$ 求出極座標的角度。在程式部分，是利用`atan2 (double x, double y)`[28]這個函式去取得二個變數 $(x_i, y_i)$ 的反正切值，這類似去計算 $x/y$ 的反正切值，但此函式能使用參數 $x, y$ 的正負號來決定其結果所在的象限。因此，此函式傳回的結果 $(\theta_i)$ 會落在正負的圓周率 $(\pi)$ 之間。在三角函數中，

$\tan\theta$ 有四種特殊的角度分別是  $0^\circ$ 、 $90^\circ$ 、 $180^\circ$ 、 $270^\circ$ ，其值在  $0^\circ$ 、 $180^\circ$  時為 0， $90^\circ$ 、 $270^\circ$  則為無意義。而  $\text{atan2}(0,0)$  在 BCB 裡算出來的不會是 0 度，而是會跳出一個 DOMAIN Error 的例外，所以把可以把  $\text{atan2}(0,0)$  當成判斷 0 度的條件。至於其它的特殊角度的計算方式，我們則是利用前一個角度  $\theta_{i-1}$  跟後一個角度  $\theta_{i+1}$  相加後再除以二來當作特殊角度  $\theta_i$  的值。

算出角度之後，根據角度與距離的關係，保留住最遠的距離，這樣便可得到 360 個極點即為手部邊緣的輪廓點。我們亦可以得到一個直角座標的曲線圖，橫軸代表角度，由左到右為 0 到 359 度，縱軸代表手部重心到手部輪廓點的距離，其最高的峰值代表著我們所取得的極座標。

如圖 45 所示：

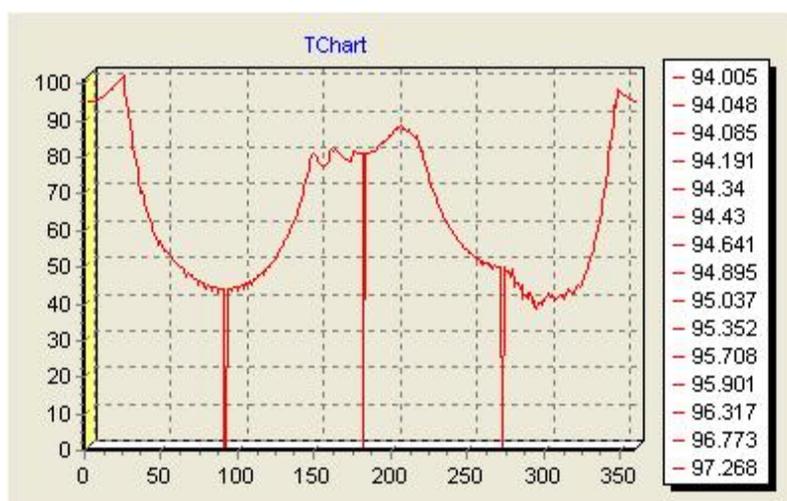


圖 45 已取得手勢各維度的極座標

### 3.3.3 特徵正規化

在訓練資料庫或辨識前，我們必須先取得特徵，而正確的取得特徵

能降低在辨識比對時產生的誤判。影像在經由前處理將手部部分的資訊留下，去除雜訊及不需要的資訊，並修補手部的像素點使我們取得較完整的手部形狀，接著才能從中取得有效的特徵。

經由 3.3.2 特徵抽取---極座標的處理取得了從 0 到 359 度的極座標，將此存放在一個長度為 360 維度的一維陣列中，以角度作為其索引值，重心座標與輪廓點座標之直線距離為元素內容，以便做接下來的處理步驟。

由於為了讓使用者在操作本系統時能容易上手，在拍攝影像時，並未定義手部在畫面中應佔的比例大小及手勢的角度；而每個人手部的形狀、大小都不一致，在 3.3.2 特徵抽取---極座標所求得的重心點到外圍輪廓點的最遠距離也不一定是相同的，比對時所用的特徵必須公平地在同樣的基礎、條件上，為此則需解決圖形放大或縮小、旋轉等問題，所以首先對一維陣列所存的距離值作數值的正規化，使其範圍落在 0~100 間，公式如下：

$$his[i] = \frac{his[i] \times 100}{MaxD} \quad (14)$$

$his[i]$ ：存放特徵的一維陣列。

$MaxD$ ：重心座標與輪廓點座標之最遠距離。

接著再對一維陣列做平滑化的處理，以彌補數值正規化後所產生的

些微誤差，到目前已解決了多數影像中放大或縮小的問題，最後角度旋轉的問題，我們使用爬山式序列表示法(Mountain Climbing Sequence (MCS))[29]。此法主要是算出此 360 維中，在給予不同權重後，所得之最小累計數，並將該序列平移至原點，此種表示法能輕易克服旋轉的問題。

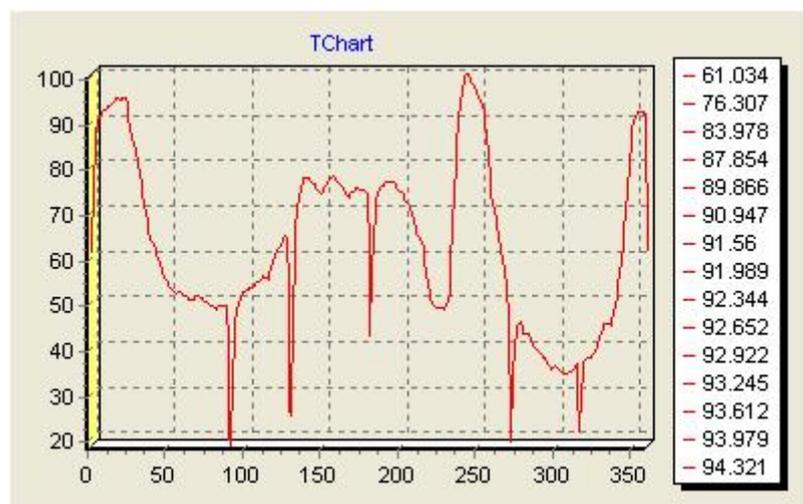


圖 46 數值的正規化之後的結果

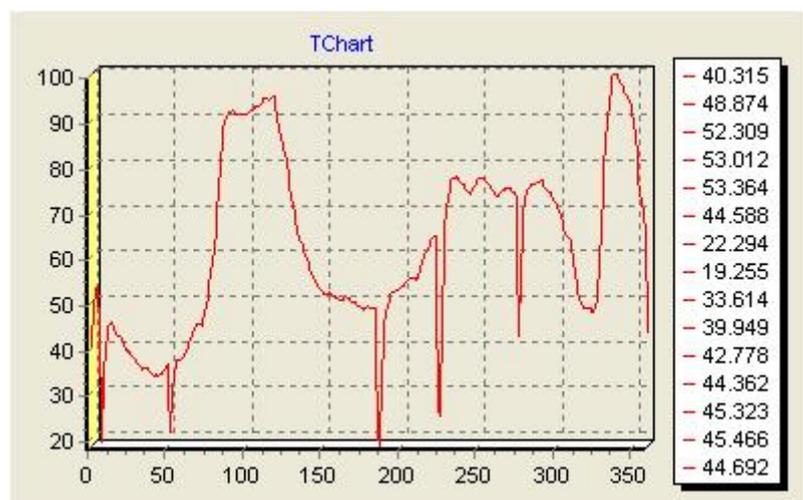


圖 47 為圖 46 移位後的結果

### 3.4 手勢辨識

於上一章節作完特徵值的正規化處理之後，其取得的影像特徵值與資料庫所建好的已是 1:1 的大小比例了。我們會針對本系統所建立的三種手勢圖 30 中的綠色手勢，去作一維陣列的特徵運算。本節我們是採用距離度量來計算原影像及使用者所比的手勢之間的差異性。

距離度量(Distance measure)乃是所有辨識系統上最基本的一環，在眾多的距離度量公式中，最傳統的度量法稱之為歐幾里得距離(Euclidean distance)，就是所謂的歐氏距離，亦稱為個體差距平方和。而本系統的辨識方式便是利用歐氏距離去作運算，計算重心距離公式如下。

$$D = \sqrt{\sum_{i=0}^{359} (d_i - d_i')^2} \quad (15)$$

$D$ ：原影像與欲辨識之影像的差異值。

$d_i$ ：欲辨識之影像極座標的特徵值。

$d_i'$ ：原影像極座標的特徵值。

我們將上一節所得到的特徵值，與資料庫裡的手勢影像的特徵平均值去做相減的動作，並分別對三個手勢平方開根號的方式所得到的值會是個正數值。而數值  $D$  越小，就代表此影像越符合資料庫其中的一個手勢圖，然後本系統便會針對所辨識的結果來執行該手勢所代表的動作。

### 3.5 播放簡報[30]

系統中放置一個 PowerPointPresentation 元件[31]，就可以透過 C++Builder 來操控 PowerPoint。之後透過程式碼的處理便可以使 PowerPoint 做換頁的動作。

當使用者按下播放鍵時，此時 PowerPoint 簡報就會藉由攝影機所辨識到的手勢，也就是辨識手勢的形狀，判斷系統該送出什麼動作指令給 PowerPoint 簡報，然後簡報依指令動作，如此一直重覆，直到簡報播放結束。

## 第肆章 系統流程

### 4.1 環境設定

本論文之硬體規格採用羅技 QuickCan Pro for Notebooks之WebCam進行拍攝，電腦執行硬體規格為CPU U7300@1.30GHz，記憶體4G DDR3 RAM，測試影像為96\*96的解析度640\*480像素的bmp影像並製作一組紅綠雙面手套。

在軟體部分作業系統的軟體為 Microsoft Windows XP，所使用的程式語言是以 C++ Builder 6.0 來開發操作。表 1 為系統的開發硬體環境與系統的參數。訓練用的手勢影像共有 3 組資料，每一組資料當中均含有 3 種手勢，測試影像共有 3 小組 9 張影像，分別是用自製的綠色及紅色手套來拍攝。

表 1 系統開發環境表。

CPU	U7300@1.30GHz
RAM	DDR3 4.00GB
攝影機規格	羅技 QuickCan Pro for Notebooks
系統實作軟體	Microsoft Windows XP

## 4.2 系統流程

此節將針對本系統的流程做一個簡單的介紹，將會說明本系統的介面以及如何操作此系統，所需要的圖像及畫面，都會在本節一一呈現。

A. 系統介面：此為本系統的主要介面，也是圖像呈現的地方，以下我們分為 A 和 B 兩部分做說明。如圖 48。

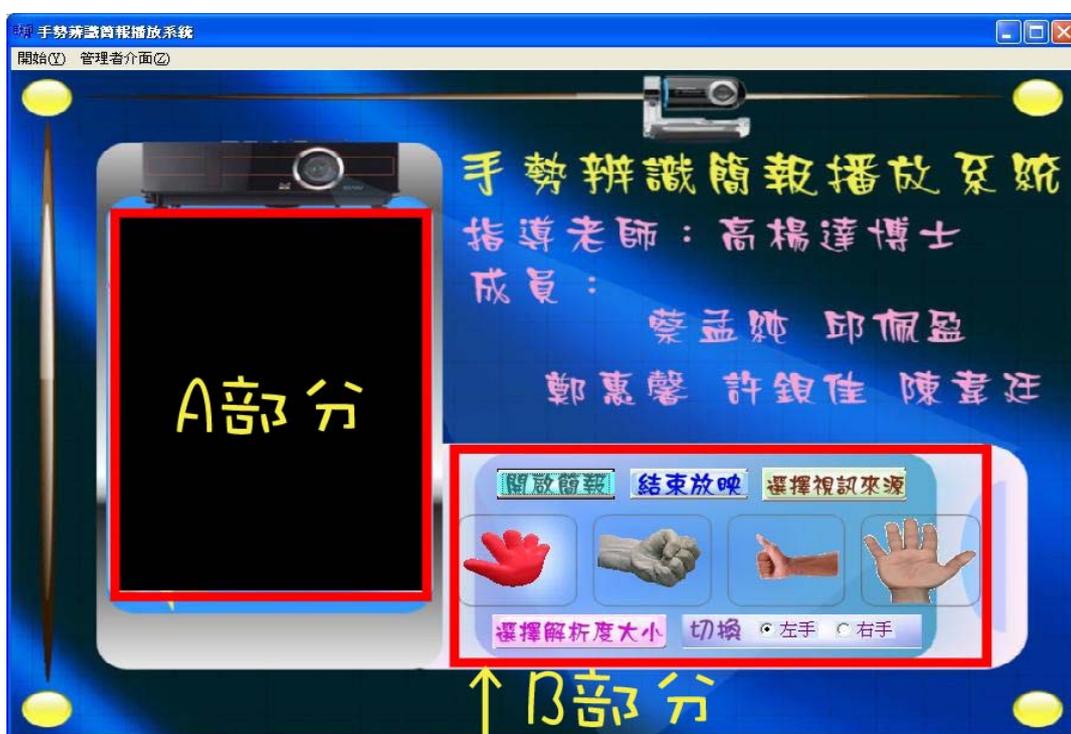


圖 48 系統介面

1. A 部分：為攝影機所接收的畫面，也就是視訊畫面的部分，WebCam 拍攝時應距離手套約 200 公分，電腦會將使用者所比的手勢以每 0.5 秒一張的速度去做拍攝的動作，並去讀取其手勢的圖像，而決定要下一頁、結束、上一頁或是回到第一頁。

2. B 部分：這個框裡面的四個圖像，正是本系統所需用到的手勢影像，以及我們會用到的按鈕，每個按鈕上都有說明其用途，以方便操作。
- 在此有四個手勢，如圖 49。由於本系統在顏色偵測方面是取綠色及紅色，所以要比手勢時，必需先戴上我們自製的手套，由此手套去比出所需要的手勢，讓系統來做辨識的動作，另外要特別注意的是在拍攝畫面中除了手套之外不能出現紅色或綠色。



圖 49 系統使用的手勢圖。

- B. 選擇簡報檔 PPT：在此，我們會點選 B 部份的「開啟簡報」按鈕，並選取需要的簡報檔。如圖 50。



圖 50 開啟簡報後出現畫面

- C. 播放的畫面：在這個部分我們將會開始撥放使用者所選取的簡報檔，並停留在簡報的第一頁，如圖 51。



圖 51 此為開啟簡報檔後出現的畫面

- D. 手勢辨識播放簡報過程：在這裡，當使用者要往下一頁進行講解時，就會比拳頭的手勢，如圖 49 的①，代表下一頁，此時簡報就會跳到

第二頁，如圖 52 及圖 53。



圖 52 此為拳頭手勢



圖 53 此為辨識完手勢所在的頁面

假設簡報播放至第十頁，因授課內容必須回到上一頁時，這個時候就比一級棒的手勢，如圖 49 的②，待系統辨識完後，就會將頁面回到上一頁。如圖 54 及圖 55。



圖 54 此為一級棒的手勢



圖 55 此為辨識後成功之畫面

接著上一個假設，如果畫面停在第十二頁，而老師的內容卻在第一頁時，是不是要一直比著手勢呢？此時就對著鏡頭比五的手勢，如圖 49 的③，畫面將跳回第一頁。如圖 56 及圖 57。



圖 56 此為五的手勢



圖 57 此為回到第一頁的畫面

E. 回到程式介面：當簡報結束，也就是讀到紅色手套的時候，如圖 49 的④，就會回到程式的原始畫面，可再做下一個選取簡報的動作。如圖 58 及圖 59。如此就完成了一次完整的簡報檔播放動作。



圖 58 此為結束手勢



圖 59 此為結束後的原始畫面

## 第五章 結論

本專題主要是透過手勢辨識的方式來操作簡報的上下頁切換、停止、回到第一頁流程，以省去使用者移動滑鼠或需走動的麻煩。一般的影像辨識主要是以膚色偵測為判斷手部區域，而本系統則採用綠紅雙色的手套來區隔手勢及影像背景，以提高影像的辨識率。

在影像前處理的部份，雖然已將大部分的雜訊除去，但仍無法判斷主要的區域，所以利用相連區域法(Connected Component)的方式，透過8 鄰居連接的偵測，得到一個完整的手部區域後，接著我們再取極座標的角度與距離作為手識辨識的特徵。由於每次拍攝到的影像距離及角度會導致偵測到的手勢大小及方向略有不同，而無法進行的有效辨識，因此，本系統利用爬山式序列表示法(Mountain Climbing Sequence(MCS))及大小正規化兩種技術，以解決旋轉及放大縮小等問題。

在辨識的部分，我們使用歐式距離(Euclidean Distance)將已經正規化後的特徵值與系統中內建好的資料庫作比對的動作，當所得到的值若愈小，便代表所辨識的影像越符合資料庫中的某一個手勢圖，系統會依其手勢的指令然後執行該相對的簡報反應。

此專題困難的地方在於「顏色偵測」和「系統處理時間」，這兩個地方將會影響到辨識的結果。在「顏色辨識」時，若是無法將有效的影像

留下並取出手部區域，而後的系統處理將會得到錯誤的辨識結果，而造成電腦資源的浪費，與增加系統處理時間，在「系統處理時間」方面，為了能即時反應在簡報播放上，必須將系統處理時間縮短至下一個拍攝時間點前，但為了能正確的辨識，每個處理步驟都是不可或缺的，所以選用計算時間花費較短的演算法，是一項重要的課題。因此我們捨棄效果較好的 Canny 而選擇速度較快的 Sobel 來做影像處理，都是為了將系統處理時間降到合理的範圍，以達到即時性。

本專題主要的應用在於即時手勢辨識方面，未來除了可以播放簡報外，更可以進一步在簡報上劃記重點，也適用於如機器人遙控、即時操控家電的開關、病患緊急求救系統，更進階的應用在於即時手語辨識系統。

## 參考文獻

- [1] Ping Zhang, Dennis Galletta, Na Li, Heshan Sun, " HUMAN-COMPUTER INTERACTION" , in Introduction, pp.1。
- [2] 唐國豪，" 人機互動，科技發展第 368 期" ，逢甲大學工業工程系，pp.18，民國 92 年 8 月。
- [3] 王立洋，蔡維哲，" 基於膚色的任意角度手勢辨識" ，國立南台科技大學電子工程學系論文，pp.1，民國 97 年。
- [4]" Wii 介紹-楊國棟" ，  
<http://www2.nuk.edu.tw/lib/e-news/20071101/3-3.htm>，國立高雄大學圖書資訊館通訊第 14 期。
- [5] 鍾國亮，" 影像處理與電腦視覺" ，台灣台北東華書局，民國 95 年 3 月三版。
- [6] 徐健峰，賴皇志，" IP CAM 即時人流計數" ，大同大學資訊工程學系專題報告，pp.6，民國 96 年 12 月。
- [7] 黃建邦，" 彩色視覺技術於多人人臉辨識上的應用" ，崑山科技大

學電機工程系碩士論文，pp. 1，民國 94 年 7 月。

[8] ” HSI to RGB 的轉換公式” ，

[http://delphi.ktop.com.tw/board.php?cid=168&fid=921&tid=5](http://delphi.ktop.com.tw/board.php?cid=168&fid=921&tid=57705)

7705，Delphi K. Top 討論區。

[9] 陳佳源，” 結合平移距離與傅立葉特徵之手勢辨識系統” ，中原大

學資訊工程學系碩士學位論文，pp. 18，民國 96 年 7 月。

[10] 林景騰，” 應用時間序列雷達影像於水線萃取之研究” ，國立中央

大學太空科學研究所碩士論文，pp. 40，民國 96 年 6 月。

[11] 王瑋群，” 智慧型人性化家電人機介面” ，國立中央大學資訊工程

學系碩士論文，pp. 56，民國 98 年 6 月。

[12] 洪兆欣，” 以軌跡辨識為基礎之手勢辨識系統” ，國立中央大學資

訊工程研究所碩士論文，pp. 17，民國 95 年 7 月。

[13] ” 電腦互動技術實驗課程設計與教學手冊” ，銘傳大學資訊工程學

系-教育部 95 年度教學卓越計畫，民國 96 年 4 月。

[14] 陳治宇，” 虛擬滑鼠：以視覺為基礎之手勢辨識” ，國立中山大學

資訊工程學系碩士論文，民國 92 年 6 月。

[15] 吳怡明，” 手勢辨識應用於遙控音樂播放系統” ，國立台灣科技大學電機工程系碩士學位論文，民國 98 年 1 月。

[16] ” Webcam 影像擷取” ，  
<http://delphi.ktop.com.tw/board.php?cid=168&fid=921&tid=62758>，Delphi K. Top 討論區。

[17] ” BCB-開啟文字檔” ，  
<http://steven1103.blogspot.com/2009/03/bcb.html> 。

[18] 施博銘、呂柏緯、林鼎軒、張寄園，” 室內環境人臉偵測” ，致理技術學院資訊網路技術系實務專題報告，pp. 11，民國 96 年 12 月。

[19] 蔡博智，” 影像追蹤方法應用在監控系統之研究” ，中原大學機械工程學系碩士論文，民國 91 年 7 月。

[20] 廖士頤，” 軟式切換濾波器於影像雜訊消除之研究” ，成功大學電機工程學系碩士論文，pp. 23，民國 95 年 7 月。

[21] ” 關於中值濾波器” ，

<http://www.programmer-club.com.tw/ShowSameTitleN/cb/16665.html>，程式設計俱樂部。

[22] 沈珀楷，”探討圖像在空間域和頻域之技術”，南台灣科技大學資訊工程系碩士作業，民國 98 年 12 月。

[23] ”型態學(侵蝕與膨脹)詳細的方程式介紹”，  
<http://delphi.ktop.com.tw/board.php?cid=168&fid=921&tid=44322>，Delphi K. Top 討論區。

[24] 陳繼崇，”結合膚色區域分割與主要成份分析於多人臉部辨識”，國立臺灣海洋大學機械與機電工程學系碩士學位論文，pp. 18，民國 95 年 7 月。

[25] 張家榮，”使用等高集合法對目標移動物追蹤之研究”，元智大學電機工程研究所碩士論文，pp. 30-32，民國 90 年 7 月。

[26] 林展緯，”以全對焦影像合成技術修正失焦顯微鏡影像”，中原大學電機工程學系碩士學位論文，pp. 3-6 - 3-7，民國 94 年 6 月。

[27] ”向量與極座標—極座標”，

<http://www.math.isu.edu.tw/calculus/cllang/Lecture/極座標.htm>，義守大學應用數學系。

[28] “math 數學函式庫” ，

<http://www.php5.idv.tw/modules.php?mod=books&act=show&shid=367>，PHP5 網管實驗室。

[29] 高揚達，”以內容為基礎之影像查詢研究”，淡江大學資訊工程學系博士班學位論文，民國 95 年。

[30] ”BCB 可自動播放投影片的函式 TPowerPointPresentation 範例” ，

<http://delphi.ktop.com.tw/board.php?cid=168&fid=921&tid=61763>，Delphi K. Top 討論區。

[31] ”BCB 的 TpowerPointPresentation 控制元件” ，

[http://www.delphi2007.net/delphiblog/html/delphi\\_200822515573015334.html](http://www.delphi2007.net/delphiblog/html/delphi_200822515573015334.html)，delphi2007.net 教程網。



