

逢甲大學

自動控制工程學系 90 年畢業專題報告

結合腳控與聲控輔助之光學式瞳位追蹤器

指導教授：林宸生

學 生：葉勝凱

第一章 前言

1.1 研究背景及動機

近年來由於資訊科技的發達與網際網路的普及，電腦已成為日常生活中另一項必需品，不管何時何地都離不開它。對一般人而言，利用電腦打打字、聽音樂、漫遊在網際網路的世界，是一項既簡單而又容易上手的事，但對於長期躺在床上或是肢體上有缺陷的病友而言，操作電腦會是非常困難且又不容易上手，如何解決這方面的問題，將是現今研究的重點與趨勢。因此就有許多的學術單位、研究單位、私人企業...等，便開使著手研發各式各樣的儀器設備，加上電腦處理運算速度和效能大幅的提昇，新穎實用性高的人機介面（Human - machine Interface）應運而生，其最終目的均是想讓操作者與電腦之間的控制和操作能夠更加靈活方便、更加符合人性。

由本系主導研發的『光學式瞳位追蹤器』，其最初目的就是架構在四肢癱瘓或半身不遂的殘障病友身上，使用者只需靠著移動眼球，就能控制滑鼠在螢幕上移動的軌跡，藉由追瞳系統與電腦溝通並下達指令，對於操作不便者是一大福音。但有鑒於使用追瞳器必須注視目標物零點五秒以上，系統才會做觸發的動作，系統本身又無法做連續兩下左鍵和右鍵的動作，且輸入觸發的方式稍嫌不足。而以上追瞳器的不足與缺失，便是主導本專題研究的最主要動機，利用足控與聲控來增加追瞳系統的功能，有其必要性與迫切性，另外就是增強其功能，讓大家都可以輕易地上手。

經過討論和評估可行性，另外再加入了兩項觸發方式－足控與聲

控。足控主要設定在手受傷或上肢有缺陷，不易使用手來做任何觸發的動作，這時就可利用腳踏板來作命令點選的動作；除此之外，聲音訊號也可當作一種輸入的方式，聲音透過麥克風，經過音效卡，將訊號傳給電腦，再經由電腦一連串的擷取與分析後，控制器便可取代滑鼠執行命令。以上的種種方法都是使操作者有更多選擇的輸入觸發方式，讓追瞳器在瀏覽網頁或操作微軟視窗方面能有更好的效率。雖然市面上有許多改良式的滑鼠產品，但受測者使用後的結果，均是不甚理想，唯有使用追瞳器與控制系統，才是一條正確的路徑。未來還必須突破現今已有的技術，讓主架構追瞳系統靈敏度更高，視線的瀏覽能更加的平滑流暢，未來可應用在週遭的日常生活上，如電玩、多媒體系統、虛擬時境...等等，對生活將有莫大的助益。

1.2 瞳位追蹤資料文獻

關於瞳位追蹤資料，最早的記載是在十九世紀的時候，當時就有許多相關眼球運動紀錄的實驗，研究者是利用麻醉藥麻醉眼球角膜後，在角膜上放置一小塊玻璃或金屬板，當眼球再轉動時，眼球將牽動角膜上的玻璃或金屬板，利用玻璃或金屬板牽動，把傳輸轉動信號至記錄裝置中，如同眼球有一個小鏡子般，以此來記錄眼球的運動，但是這種實驗對受測者的眼睛會有極大的傷害。因此之後的研究者開始使用攝影的技術來減輕實驗者的負擔。二十世紀後，漸漸的有許多的學者專家開始利用眼睛的各種特性來記錄眼球運動的方式，如電的特性、眼睛形狀、光的特性或直接接觸眼球的方式等，最主要的方法大約可分成下列五種方式：

1. 眼電圖法（ Electro-Oculography ）簡稱 EOG [1]。

2. 搜尋線圈法 (Search Coil) 簡稱 SC [1]。
3. 紅外線眼動圖法 (Infra-red Oculography) 簡稱 IROG [2]。
4. 紅外線視訊系統法 (Infra-red Video System) 簡稱 IRVS [1][3]。
5. Purkinje 影像追蹤法 (Dual-Purkinje-image) 簡稱 DPI [1]。

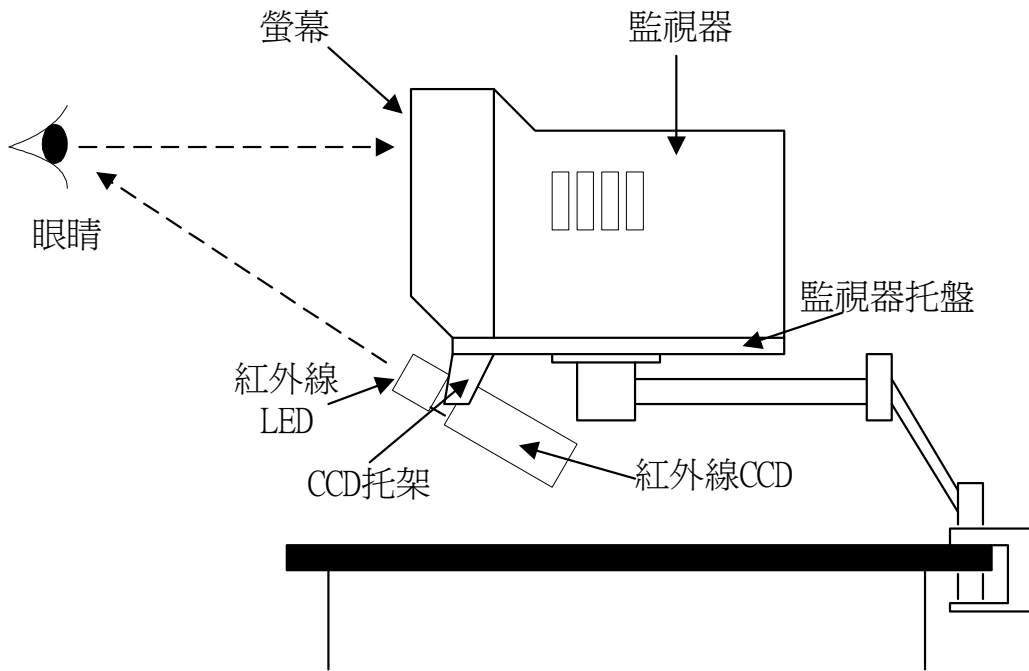


圖 1.1 紅外線視訊系統法 [3]

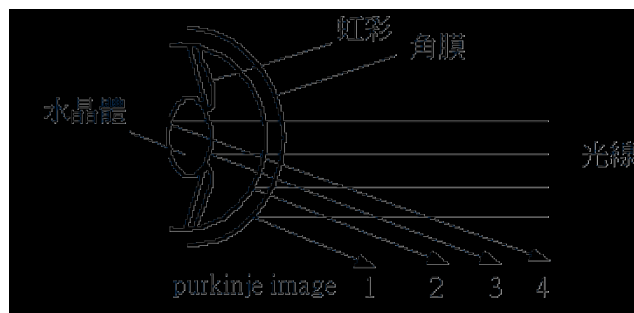


圖 1.2 Purkinje-images 影像追蹤法[1]

但無論所採用的技術為何，一個理想的眼球追蹤系統應包含如下的特點：

- 系統必須能夠提供使用者無障礙的視野。
- 系統不與受測者接觸（非接觸式量測）。
- 系統必須具有高準確度，準確度受限於非線性、扭曲、雜訊與延遲等誤差。
- 系統必須具有高解析度以量測微小的眼球運動，解析度受限於設備的雜訊。
- 系統必須能夠計算三維的眼球轉動，而對眼睛的平移不靈敏。
- 系統必須能夠適用於不同的受測者與環境

1.3 研究方法

1.3.1 專題研究方向

- 1.足控輔助裝置的架設
- 2.聲控系統的架設與運用
- 3.光學式追瞳器輔助系統之整合與功能

在整個研究的過程當中，首先我們先針對追瞳器輔助系統作處理，依照原先預定的成果結論與執行功能，設計出適合系統的軟體需求與硬體架構，經過測試無誤後，再將輔助系統結合到已研發完成的『光學式瞳位追蹤器』上，並搭配頭配顯示器，讓操作者直接從頭配顯示器內的螢幕，作移動滑鼠與選取目標物的動作，用輔助系統來作指令確定的觸發與選取。

第一項足控方面，使用了 8255 晶片作為與電腦溝通的主要媒介，先以腳踏板當作命令輸入的開關，透過 RS232 連接埠，將觸發按鍵的訊號傳入電腦中，再用軟體系統去判斷有無輸入訊號，如果有訊號觸發則系

統執行滑鼠按鍵的動作。

第二項聲控方面，以麥克風與音效卡作為溝通的橋樑，利用聲音當作輸入的訊號，再以軟體取得聲音波形的訊號，經過一連串的信號擷取與分析後，控制器便可取代滑鼠執行命令。

最後再將其輔助設備加入主系統追瞳器中，當使用者透過追瞳器移動滑鼠游標，欲點選某個目標物時，只要在目標物的範圍之內，踏下腳踏板開關或輸入一段語音給麥克風，都可以讓輔助系統達到觸發滑鼠左右鍵的命令功能。如何讓系統模擬出更接近滑鼠，將會是整個研究的主要重點。

1.3.2 預期結果

- 1.足控裝置與追瞳器結合
- 2.聲控系統的架設與運用
- 3.完成系統軟體操作介面
- 4.完成系統整合及測試
- 5.評估整套系統的可行性及改良
- 6.完成論文報告

1.3.3 步驟與時間分配

表 1.1 詳細步驟與時間分配表

工作項目 \ 月次	第一月	第二月	第三月	第四月	第五月	第六月	第七月	第八月	第九月	第十月
專題相關書籍研讀	■	■								
實驗硬體採購與佈置		■	■							
實驗整合之設置			■	■	■					
影像擷取之測試與分析		■	■	■	■	■				
瞳位樣板判斷法則與評估				■	■	■				
足控系統設計與測試				■	■	■	■			
聲控系統設計與測試					■	■	■	■		
系統整合與實際測試				■	■	■	■	■	■	
軟硬體之整體評估與改良							■	■	■	
測試資料與數據之整理分析								■	■	■
書面報告撰寫									■	■
預定進度累積百分比	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%

第二章 理論基礎

2.1 CCD 簡介

一般應用於機械視覺的取像裝置有光導攝像管、V8 攝影機、CCD 攝影機…等，這些器材均為換能器裝置，可將一影像中各像點的亮度轉換成電壓訊號而輸出成像。

在實驗中，我們採用針孔式紅外線 CCD 攝影機（如圖 2.1 所示），其詳細規格如表 2.1 所示。由於針孔式攝影機體積小、重量輕、價格便宜、不佔空間，可直接與頭配顯示器相結合，當我們注視頭配顯示器內的螢幕時，CCD 隨時可攝取瞳孔的影像，送回電腦做影像處理，並與我們眼睛的距離保持固定，可減少不必要的誤差。

表 2.1 本實驗使用之 CCD 規格

成像裝置大小	SONY 1/3 英吋 INTERLINE TRANSFER CCD
信號系統	510(H)*492(V)NTSC ; 500(H)*582(V) PAL
解析度	380 TV line
最小要求照度	2 LUX / F3.5
掃瞄系統	2 : 1 交錯式
自動電子快門速度	1 到 1/100,000 秒
透鏡設備	針孔式透鏡 5.5mm F3.5
信號雜訊比	> 48 d B
電源	12VDC

CCD 攝影機相較於其他傳統取像裝置相比，的確有其優異之處，以下簡述其優點：

1. 因為沒有真空管，所以耐衝擊及抗震性強。
2. 體積小，重量輕。
3. 產品壽命長、不易老化、耗電低、穩定性佳。
4. 因屬 IC 製造技術，其幾何精密度好，影像不會扭曲變形。
5. 因為沒有偏向磁場，所以不受電磁干擾。
6. 每一次取像終了，電荷一定移出，故沒有殘像的問題。

如此的考量，在未來將整套追瞳系統模組化時的效能評估及誤差率的影響都是相當的重要的。



圖 2.1 CCD 攝影機之實體圖

2.2 輔助光源之影響

為了使 CCD 攝影機能清楚的抓取眼睛影像，良好的光源選擇非常的重要。

要成為良好的光源需符合下列幾項特點：

1. 獲得明顯的對比以凸顯其輪廓。
2. 顯露出影像的細節。
3. 強化物體的特徵。

4. 增加訊噪比，以確保影像資料的真實性。

我們所使用的光源若符合以上的特點越多，則得到的效果越好。而在選擇光源的另一重點，則是需瞭解光源的特性，針對不同環境採用不同的光源，以得到較佳性能及降低對使用者的傷害。以下列舉幾項光源的相關特性：

1. 光的亮度。
2. 亮度的穩定性。
3. 光的頻譜。
4. 是否有螢光。
5. 是否有紫外線。
6. 是否有紅外線。
7. 使用壽命。
8. 是否易於安裝與固定。
9. 色溫。
10. 演色性。

光源體接受電能後，除轉換一部份可見光之外，其餘電能則轉變為紅外線、紫外線、螢光及熱能等等，這些因素不僅會影響到光源本身的壽命，並對與光源有直接接觸的使用者造成若干程度的傷害。以下列舉數種常見的人工光源，列表匯總與比較，如表 2.2 所示。

由表可看出各類光源的特性，本實驗除了機械視覺的適用性需為佳，也就是可使得 CCD 能更清楚的攝取影像之外，尚需考慮對使用者眼睛的傷害程度，譬如氙氣燈的發光頻率包含紫外線的範圍，直視過量會造成眼睛的傷害；而有些螢光燈（日光燈）的頻譜分佈並不符合 CCD 攝影機取像頻率的要求，所以也不適合採用。而一般常見的普通燈泡其

輻射紅外光強，亮度穩定性高，不會有閃爍現象，且具有連續光譜、價格便宜、容易取得、不受環境溫度影響等優點，故我們採用普通燈泡為本實驗的人工光源，經實驗證實，確實有不錯的效果。

表 2.2 各類人工光源匯總與比較

	普通燈	鹵素燈	螢光燈	氬氣燈
紅外線	極多	極多	無	多
紫外線	無	無	許多	少許
螢光	無	無	少許	少許
發光效率(lm/W)	22	27	50~104	30
平均壽命(小時)	1000	2000	5000	-
機器視覺適用性	佳	佳	佳	佳

2.3 二值化討論

當我們在進行影像處理時，若被處理的物體和不需要的背景能加以分離的話，將使我們方便許多，並更快處理速度。一張影像是由像素點所構成的二圍矩陣，每一個像素的光強度函數（ Light-Intensity Function ）以 $f(x, y)$ 表示之，俗稱為灰度值（ Gray Level ）或明亮度（ Brightness ），其中 (x, y) 為影像中的座標。一般灰度值的範圍如下：

$$0 \leq f(x, y) \leq 255 \quad (2.1)$$

其中 0 表示最暗或黑色，而 255 表示最亮或白色。如圖 2.2 所示的灰度分佈統計圖，為影像中的像素對應於 $f(x, y)$ 所產生的圖形，整幅

影像就由亮的背景和暗的物體所組成，在兩個模式間存在一個臨界點 T (閾值)，若我們想從背景中提出一個物體，即可以 T 為判斷值，如果 $f(x, y) < T$ 則稱為物體點，否則為背景點。在本次專題實驗中，我們將臨界的 T 值設為可調，以調整出最佳的分離狀況。

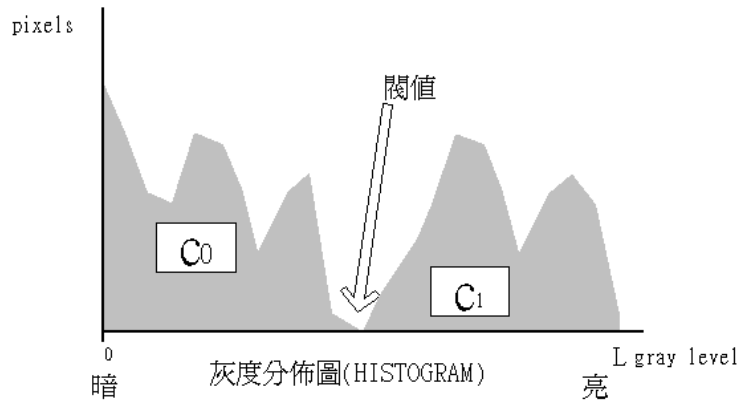


圖 2.2 灰度分佈統計圖

一個二值化的影像 $g(x, y)$ 定義為

$$g(x, y) = \begin{cases} 1 & f(x, y) > T \\ 0 & f(x, y) \leq T \end{cases} \quad (2.2)$$

影像中的像素若被視為 1 時即被視為背景，標記為 0 則被視為物體，因此在我們的實驗中，瞳孔的影像經二值化處理後就會清楚的與背景分離，以便演算法則判斷。其次，如圖 2.3 所示為光源對眼球影像的影響，圖 2.3(a)未加光源時，圖 2.3(b)則是加入光源後的影響。

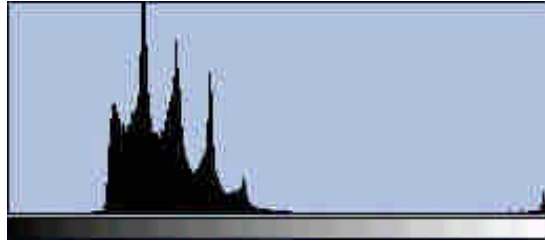


圖 2.3(a) 打光前之眼球影像的灰階亮度統計圖

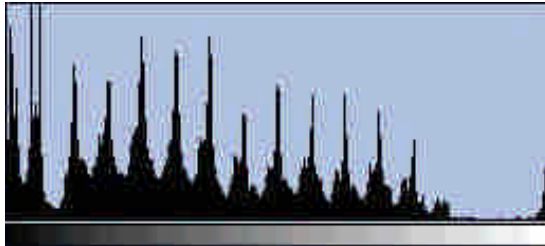


圖 2.3(b) 打光前後之眼球影像的灰階亮度統計圖

2.4 視覺基本概要

眼睛是人體中一個非常重要的感官器官。人們要知道外界環境的改變，或要適應環境，必須依靠感覺器官傳遞信息來做出反應。在外界信息中，百分之八十是通過視覺系統得到的。

正常眼睛的視線：

正常人的眼睛在觀看一物件時，兩眼球是需要互相協調，而使得兩眼的注視焦點在同一點上。當我們眼睛在看遠物時，兩眼的視線是近似平行；當兩眼看近物時，兩眼的視線會向內側對準焦點，使得映照在兩眼的影像感覺是一個物體。在生活中，常常會聽到一句話『那個人在看我耶！！』，其原因是因為我們人類有敏感的視覺焦點判斷能力，因此我們能很輕易的察覺到對方眼睛的注視點。當兩眼睛注視同平面之一點時，注視不同位置，兩眼將協調而產生不同的移動位置及角度。當兩眼

睛注視不同平面時，可能發生兩眼睛中，其中一眼睛的移動位置及角度相同。因此，我們也可能會發生誤判對方注視焦點的窘境。在本研究光學式瞳位追蹤器，也是根據眼球觀看同一平面物件時，眼球將產生不同的移動位置及角度，所以我們可以分析眼睛的焦距，來判斷眼睛注視的焦點。

2.5 瞳位演算法

對於瞳孔中心的追蹤步驟上，我們使用了最佳追蹤理論的觀念，消除偏移的誤差並使系統能正確的判斷出瞳孔注視的位置。圖 2.4 中，眼睛注視上方校正點此時系統在初值中經過取像得知標準的定位點，這時能計算出瞳孔在注視時的相對座標值，經注視的校正後的座標做出相對關係而相對比例比例設為 $\beta = 2$ 。

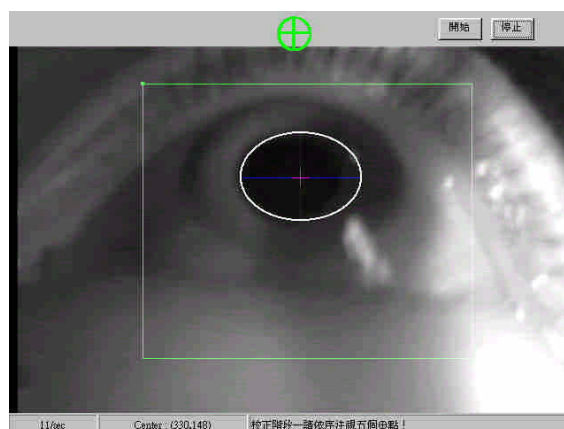


圖 2.4 眼睛注視上方校正點之圖形

首先在瞳孔座標的處理中，我們先採取估測內徑重心 (G_x, G_y) 將全部 (V_{xn}, V_{yn}) $n=0 \sim n_{max}$ 採樣之，其半徑加標準差外圓圓週外之雜點消去，並重新計算以新點瞳孔注視畫面中心點，這時映對座標求得為 (M_x, M_y) ，系

統定位固定注視點座標設為 (Sx, Sy) ，此時則可算出下點即將預測產生之座標 (Sxx, Syy) 。估計時採樣的注視點共取樣 N 次，並得平均值為 (Mx', My') ，並同時取樣 100 次內的標準差 σ (如無 100 次則取樣 100 次內)。在系統計算第十次後 σ 如出現不合理的數據偏差，我們將此取樣點取消，並記錄合理之 σ ，當其由大變小後又漸轉為大時或絕對誤差在 b 以內時，我們將第 n 次取樣瞳孔位置求得之 (Mx, My) 為此次的標準校正點。而在平面上 Sxx, Syy 也將與 (Sx, Sy) 產生相對的關係。我們所使用的公式如下：

$$(Gx, Gy) = \frac{\sum_{n=0}^{n_{\max}} (Vxn, Vyn)}{n} \quad (2.3)$$

$$(Mx, My) = \frac{\sum_{n=0}^{new_max} (Vnew_xn, Vnew_yn)}{new_max} \quad (2.4)$$

new_max is 慮雜訊後之正式保留像素之總點數， $(Vnew_xn, Vnew_yn)$ 為率雜訊後正式保留像素之點座標

$$(Mx', My') = \frac{\sum_{n=0}^n (Mxn, Myn)}{n}, n \leq 100 \quad (2.5)$$

$$\sigma = \sqrt{\frac{1}{m-1} \sum_{i=1}^m (M_{(xm,ym)} - M_{(x,y)'})^2} \quad (2.6)$$

$$\beta_{(x,y)} = \frac{S_{(xn,yn)} - (Sx, Sy)}{(G_{(x,y)} - S_{(cx,cy)})} \quad (2.7)$$

$n = 1, 2, 3, 4$

瞳位校正第二階段則實施 $\beta_{(x,y)}$ 距離調整的動作，我們以上式的相同方式求出第二點 Gx, y 以最先求得之 (Mx, My) 為標準點校正點(此點我們設為 (Scx, Scy))以初始設定 (Sxn, Syn) $n=1, 2, 3, 4$ 之座標比對，將取樣之點做 β 微調即可得到一個正確之 β 係數，並以 $(Sx, Sy)(Sxn, Syn)$ 之間距

離來限定相對位置。接著我們將分配四象現來校正 $\beta_{1,2,3,4}$ (即右上左下) 在各象現之值。當系統計算出標準校正點中心 (Scx, Scy) 時會自動切割四個象現來計算相對位置，因而類推，我們可以使用四個校正點及原點之定位點來做定位並求出 $\beta_1, \beta_2, \beta_3, \beta_4$ 。因取樣之 β 必須計算兩次才能得到一正確的相對座標。

表 2.3 以 β 雙採樣來進行點的計算

象限	使用 β 係數
一	β_1, β_2
二	β_2, β_3
三	β_3, β_4
四	β_4, β_1

根據表 2.3 的讀入計算判斷，我們可針對，眼睛瞳孔計算 (Mx, My) 之所在不同位置映射至螢幕座標 (Rx, Ry) 這種分段多層次的判斷方式，可精確的計算出瞳孔真正的座標值，如發現系統計算出格的狀況發生機率 > 0.3 則使用者的配戴位置可能有許偏差此時如 (Mx, My) 持續出格，則可以加以最大出格位置偏移量 (Ax, Ay) 來計算出新的瞳孔位置。

$$\begin{aligned}
 Ax &= | Mx_outside - Mx | \\
 Ay &= | My_outside - My | \\
 (Mx_outside, My_outside) & \text{ 為電腦計算出出格之點。}
 \end{aligned}
 \tag{2.8}$$

則參考點 Sx, Sy 將自動依 Ax, Ay 相減修正。

即系統求出之相對於螢幕座標為 (Rx, Ry) ，其於作圖表示上如 圖 2.5 所示。

$$R_{(x,y)} = (Sx, Sy) + \beta_{(1,2)} (G_{(x,y)} - (S_{(cx,cy)} \pm A_{(x,y)}))
 \tag{2.9}$$

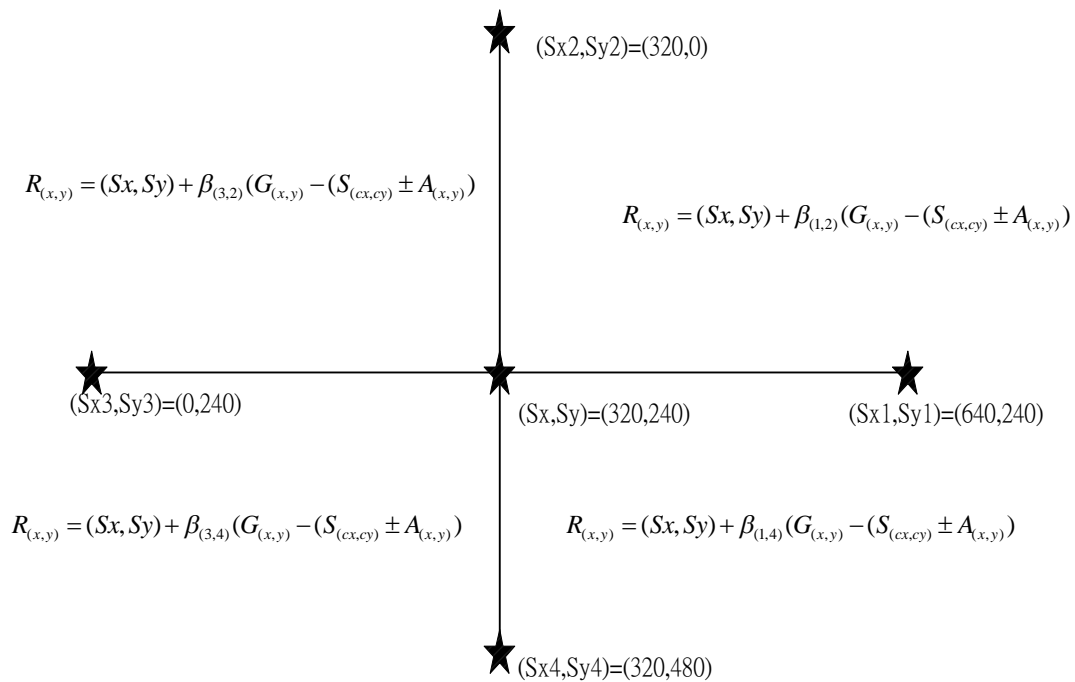


圖 2.5 圖示座標關係

2.6 8255 介面卡

8255 為一可程式規劃的週邊介面 (PROGRAMMABLE PERIPHERAL INTERFACE) 晶片，簡稱 PPI，可連接多種不同的微處理器。由於其價格便宜且容易使用，因此為微處理器對外界傳送資料最重要的晶片。

2.6.1 8255 方塊圖

8255 為一可程式化的週邊介面元件，其內部方塊如圖 2.6 所示，由圖可看出 8255 可區分為以下四部分：

1. 資料匯流排緩衝器 (Data Bus Buffer)

此匯流排為雙向的 8 位元緩衝器，用來做為 CPU 與 8255 間資料傳輸的介面，當 CPU 執行 IN、OUT 指令，資料將經由此暫存器輸入或輸出。

2. 讀寫控制邏輯(READ/WRITE CONTROL LOGIC)

在讀寫控制邏輯中含有 RD、WR、A1、A0、RESET 與 CS 之控制腳，其主要之功能是用來控制資料的傳輸方向與寫入控制暫存器的控制字組狀態。表 3.1 為讀寫控制邏輯的基本操作。

3. A、B 組控制單元(GROUP A & GROUP B CONTROL)

此單元用來控制 A、B、C 三埠的動作，其動作是由 RD、WR 與控制暫存器決定。A 組控制 PA7~PA0、PC7~PC4；B 組控制 PB7~PB0、PC3~PC0。

4. A 埠、B 埠、C 埠：

此三個埠均為 8 位元的資料輸入/輸出門鎖器或緩衝器，共分為兩組來加以規劃，其中 A 埠及 C 埠上半部(PC7~PC4)屬於 A 組控制、B 埠及 C 埠下半部(PC3~PC0)屬於 B 組控制。C 埠除了可分成兩個四位元的 I/O 組合外，亦可用於產生 A 埠及 B 埠所需的交握性號線。

表 3.1 讀/寫控制邏輯的基本操作

/CS	/RD	/WR	A1	A2	動作說明
0	0	1	0	0	讀取：A 埠→資料匯流排→CPU
0	0	1	0	1	讀取：B 埠→資料匯流排→CPU
0	0	1	1	0	讀取：C 埠→資料匯流排→CPU
0	0	1	1	1	禁能：條件不符
0	1	0	0	0	寫入：CPU→資料匯流排→A 埠
0	1	0	0	1	寫入：CPU→資料匯流排→B 埠
0	1	0	1	0	寫入：CPU→資料匯流排→C 埠
0	1	0	1	1	寫入：CPU→資料匯流排→控制暫存器
1	X	X	X	X	資料匯流排為三態
0	1	1	X	X	資料匯流排為三態

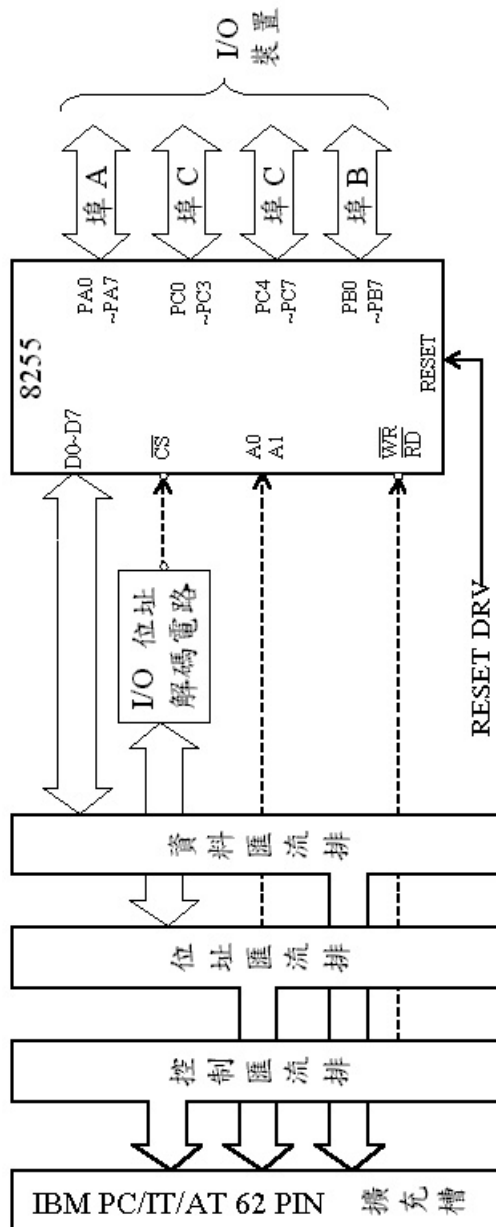


圖 2.6 8255 內部方塊

2.6.2 8255 的控制字組

前面曾經提過 8255 為可程式周邊介面元件，透過控制暫存器的控制字語，可將埠 A、埠 B 及埠 C 規劃成三種工作模式。8255 對 24 條 I/O 線化分成 A 群及 B 群來進行工作模式之規劃，所謂 A 群是指埠 A 及埠 C 上半部；而 B 群是指埠 B 及埠 C 下半部。

控制字語為寫入控制暫存器的 8 位元資料，其最高位元 D7 之值具有兩種意義：當 D7=1，用來設定 8255 三個 I/O 埠之工作模式；反之，當 D7=0，用來設定清除埠 C 的各位元值。圖 2.7 所示為 8255 的模式定義控制字組 (Mode Definition Word) 之格式。

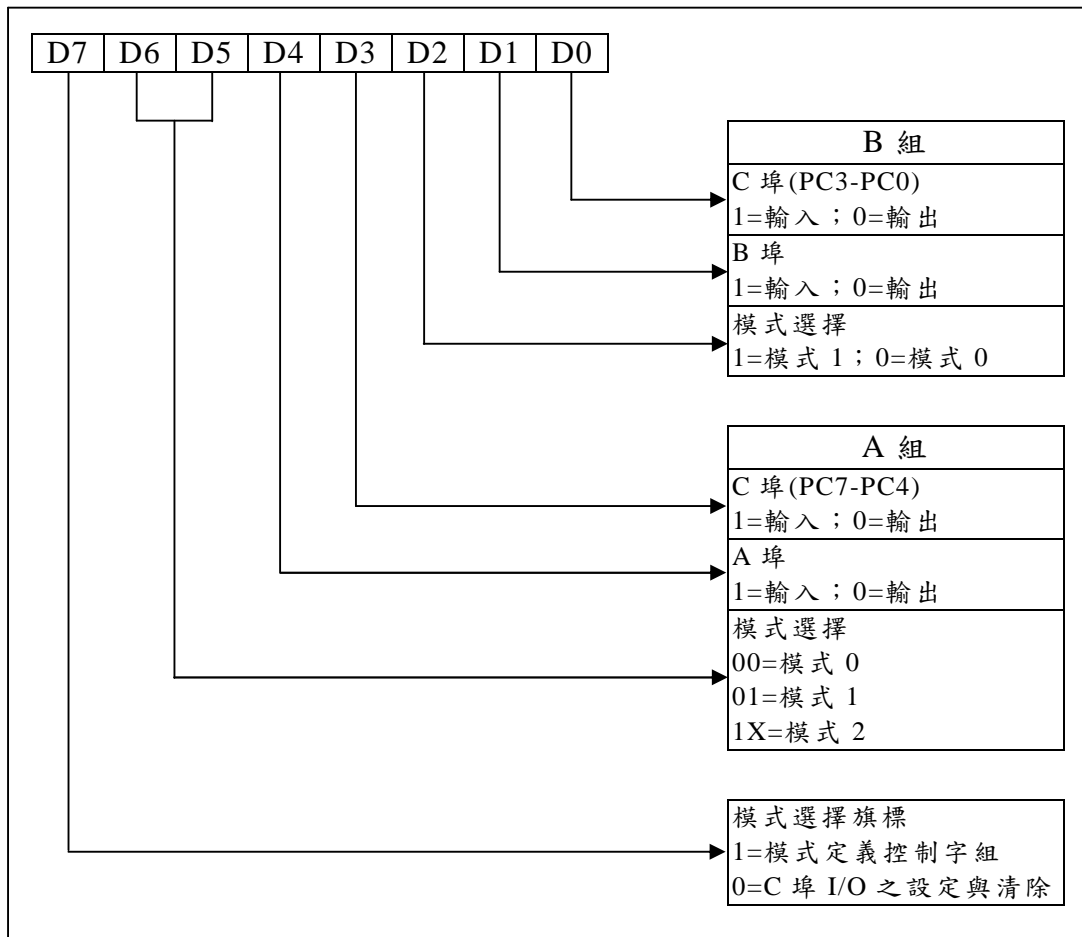


圖 2.7 8255 的模式定義控制字組

2.7 PC 聲霸卡 (Sound Blaster)

想要在個人電腦聽到美妙的歌曲與動人的音樂，那麼圖 2.8 聲霸卡將是依個不可或缺的產品，沒有它的話，充其量電腦只是一台計算機而

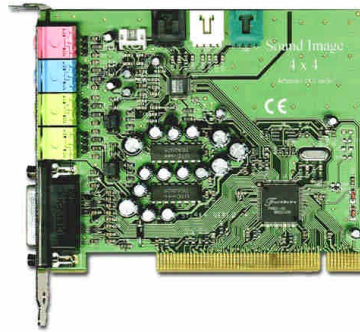


圖 2.8 聲霸卡

已，根本不適用於一般家用的個人電腦。但是有了它的存在，才能稱為一台多媒體電腦。同時它對發出音色的好壞與錄製聲音的品質，也同樣扮演著非常重要的角色，所以市面上才会有許許多多價格差異很大的聲霸卡，當然輸出的音響配備與輸入聲音的裝置好壞，在音質與抗雜音上，也佔有一席之地。

2.7.1 認識語音信號

任何信號都可以用波的形式表示，語音信號也不例外。談到波形，不免要談到振幅、週期等等。而像語音訊號的波形，我們可以在不同的時間區段上找到不同的週期，造成這種現象的主要原因，乃是因為語音訊號本身是由很多具有不同週期的訊號所組成。這種週期隨時間變化的訊號，我們稱為非固定式（nonstationary）的訊號，如圖 2.9(a)所示，利用 Windows 的錄音程式錄製出的聲音波形，任何時刻週期都不一樣。而固定式如 \sin 、 \cos ...

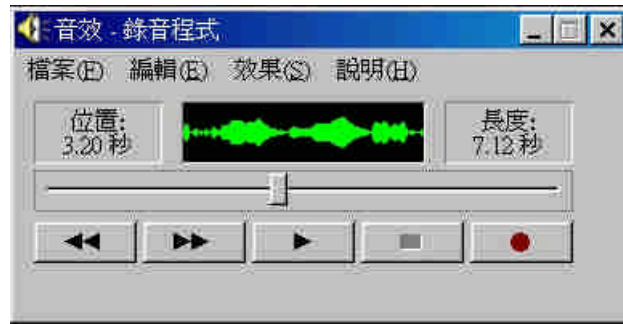


圖 2.9(a) 錄製出的聲音波形

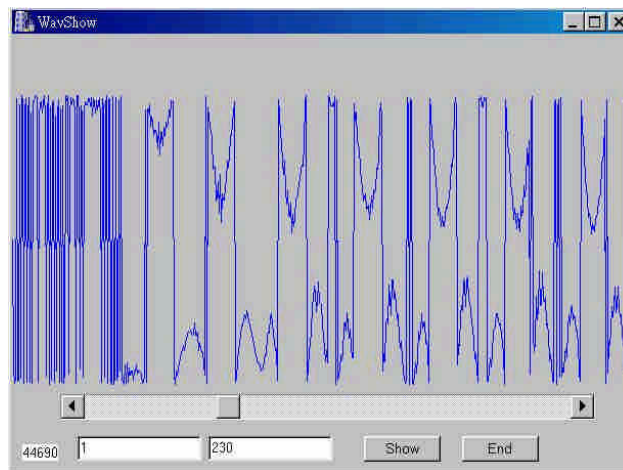


圖 2.9(b) 聲音波形的繪製

然而對非固定式的訊號來說，要處理它們並不容易，以語音訊號來說，由於在不同的時間區段有不同的週期與振幅大小，因此一段非固定式的訊號所具有的週期與振幅等參數數目，都比固定式的訊號來得多。雖然語音是非固定式的訊號。這樣一來，我們就可以用處理固定式的訊號的方式來對不固定式的訊號來作處理。再語音處理上，每一時間區段是一個短時距（short time）或被稱為一個音框（frame），如圖 2.10 所示音框取法。圖 2.9(b)代表的意義為圖 2.9(a)聲音波形片段重新經過放大而繪製成的。

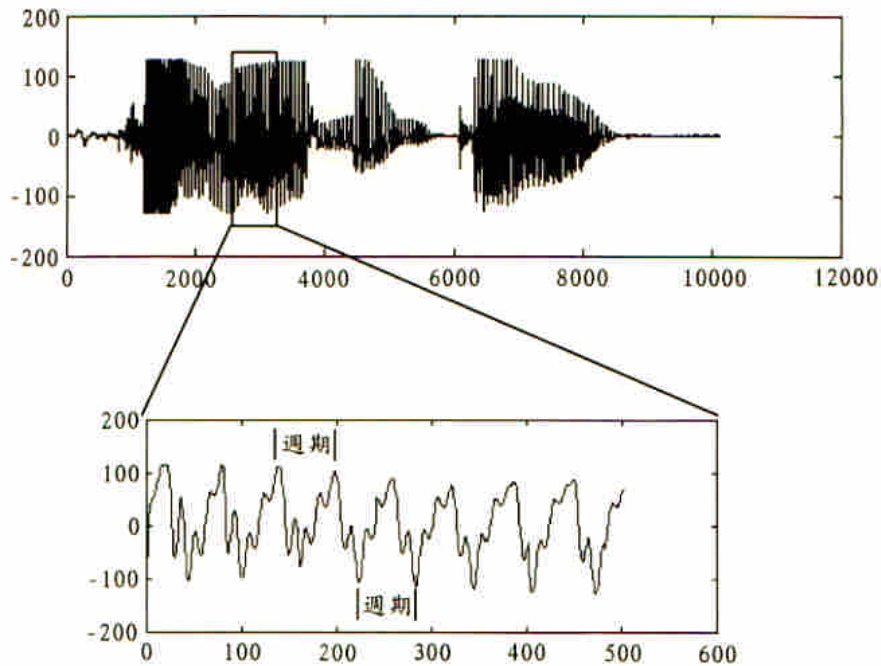


圖 2.10 音框取法

語音訊號還有一項重要的特性：在不同時間，雖然說的是同一句話或一個音，但其波形卻不盡相同，也可以說語音是一種隨時間而變的動態性訊號。

2.7.2 語音訊號數位化

語音在空氣中是以波的形式來傳遞，我們將以這種波的訊號稱為類比訊號 (Analog signal)。電腦本身並無法直接處理這些類比訊號，因為在電腦中，所有的資料都是以 0 或 1 表示。訊號以數值大小表示時稱為數位訊號 (Digital Signal)。如果要用電腦來處理類比訊號，就必須將類比訊號轉為數位訊號才行。

語音訊號在進入電腦前是一種連續性的訊號，所謂的連續性號是指

時間上的連續，透過麥克風我們把聲音傳到電腦裡面，這中間的過程稱為數位化。

原來連續性的訊號，經過數位化的處理後，變成一種不連續的訊號，這個訊號只在某些固定的時間刻度上有值，這些刻度稱為取樣點，取樣點上的振幅大小稱之為取樣值，兩個取樣點之間的時間間隔稱之為取樣週期，取樣週期的倒數稱為取樣頻率。取樣頻率的另一種說法是，在一秒中內要對原始訊號做幾次取樣，取樣頻率的大小可以由使用者自定，但是一般來說要符合取樣定理 (Sampling Theorem)。就是取樣頻率需大於兩倍取樣頻率，否則會造成取樣失真 (Aliasing)，一旦取樣失真發生，所得到的聲音取樣值便不能代表原來的訊號。因此在對語音訊號作取樣時，由於語音的最高頻率不會超過 4000Hz(圖 2.11(a))，所以取樣頻率定在 8000Hz(圖 2.11(b))以上，就保證取樣失真不會發生。

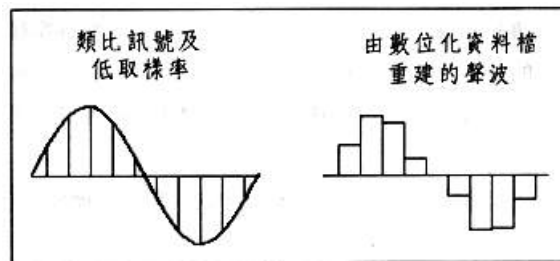


圖 2.11(a) 低頻率取樣

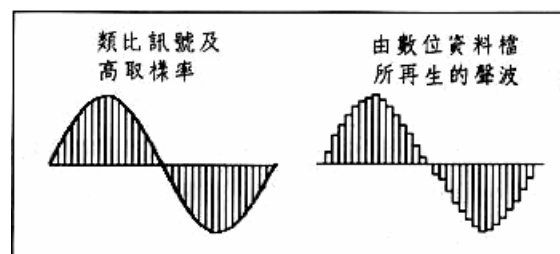


圖 2.11(b) 高頻率取樣

在每個取樣點上，訊號的振幅大小也被數位化，在麥克風上的訊號是一種細微交流電壓的型態，譬如說：原始訊號的振幅在-50mv ~50mv 之間變動，數位化的過程就是把電壓轉換成數字大小，如以-128 代表 -50mv，127 代表 50mv，-128~127 中間的數值代表電壓大小，則以線性（Linear）或是以非線性（Non-Linear）的方式遞增。

第三章 實驗方法

本專題研究共分三個方向，『足控輔助裝置的架設』、『聲控系統的架設與運用』與『光學式追瞳器輔助系統之整合與功能』，前二項先個別獨立出來設計，最後再一起整合到第三項中，經故測試無誤後，再請操作者測試其性能。

3.1 足控輔助裝置的架設

3.1.1 實驗設備：

1. PC 個人電腦：

Pentium II-350 電腦，64MB RAM

2. 使用的軟體：

Borland C++ Builder 4.0

3. 8255 晶片：

笛卡兒 8255 介面卡

4. 輸入裝置：

腳踏板

3.1.2 實驗步驟

1. 首先架構出足控的硬體配備(如下圖 3.1)。



圖 3.1 足控系統硬體配置圖

2. 8255 介面卡之設定與規劃

在使用 8255 介面卡時，我們必須先確認需求的工作模式，模式 0 其工作功能為單純的輸入、輸出，沒有檢查用的信號線(握手線)；模式 1 其工作功能較模式 0 多出了兩條檢查用的信號線，目的為檢查輸出端的資料是否被接收出去，此功能可減少程式設計者寫程式的一些麻煩。由於本專題為基本的輸入輸出，故在此使用工作模式 0。

設定完工作模式後，接下來的步驟就是要選定輸入輸出埠，根據實際狀況的要求，選擇適合的腳位來使用。為了模擬出滑鼠的左右鍵，我們需要兩個輸入的信號接腳，所以選用埠 C 下半部(PC7-PC4)，來當作觸發信號的輸入腳位。

整合上述的種種設定，規劃出如底下程式的寫法，其中 `outportb` 為電腦對外界溝通的指令；`0x` 在程式的語法中代表後面的數值為十六進位的寫法；`303` 為 8255 晶片的控制字組位址；最後面的十六進位值 `88`，其二進位寫法為 `1000 1000`，參照圖 2.7 而言，意思為使用模式 0、使用埠 C 下半部作輸入信號端。

```
outportb(0x303,0x88);
```

3. 接收並判斷輸入的信號

由於『光學式瞳位追蹤器』於抓取影像時，每秒鐘抓取 10 到 15 張影像，所以在程式的寫法上，當每抓一張影像，系統將會執行底下的程式碼，其中 `inportb` 指電腦接收外界資料的指令；PCD 其值為 0x302 意義為埠 C 資料的暫存器位址，搭配前面的 `inportb` 指令，而成為接收埠 C 位址的資料；最後 `& 0x80` 這一項，則是比對觸發訊號是鍵入左踏板，還是右踏板。

```
#define PCD 0x302
```

```
Foot_Left = inportb(PCD) & 0x80;
```

```
Foot_Right = inportb(PCD) & 0x20;
```

3.1.3 實驗討論

由於 8255 卡是跟電腦作連結，若要從外界輸入訊號到電腦裡面，直接給一 V_{high} 的電壓可能會對電腦造成某些問題，為了避免這個問題，在輸入訊號與 8255 卡之間加上了保護裝置，以防突然產生高電流進入電腦，造成電腦的損壞。

3.2 聲控系統的架設與運用

3.2.1 實驗設備：

1. PC 個人電腦：

Pentium II-350 電腦，64MB RAM

2. 使用的軟體：

Borland C++ Builder 4.0

3. 音效卡：

16bit 天音卡

4. 輸入裝置：

麥克風

5. 輔助配備：

耳機

3.2.2 實驗步驟

1. 首先架構出聲控的硬體配備(如下圖 3.2)。

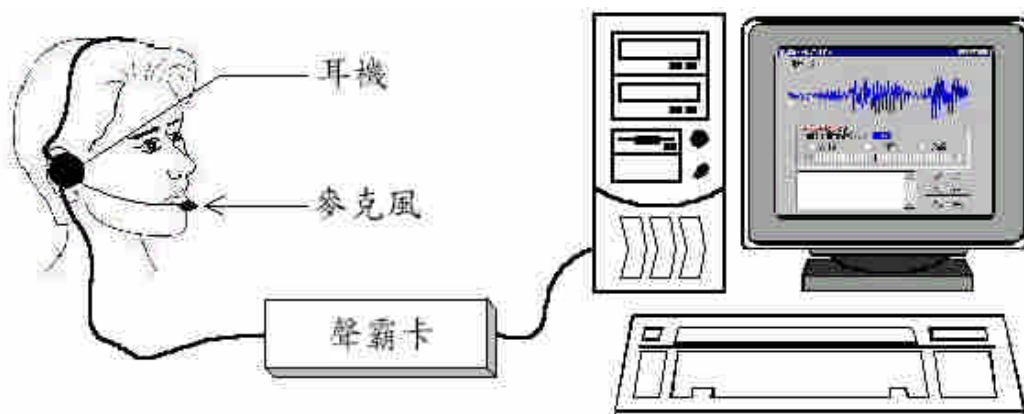


圖 3.2 聲控系統硬體配置圖

2. 錄音程式的撰寫

要抓取聲音的訊號，最快最容易的方法就是使用錄製聲音的軟體，透過麥克風可將受測的聲音波形錄製並儲存在電腦裡面，在經由其他軟

體對待測聲音進行分析。然而這次研究需用到的錄音程式，必須長時間處於錄音狀態，利用現成的 Windows 錄音程式，只能錄製一分鐘左右的時間，根本不符合這次的實驗，所以必須另外設計新的錄音程式，又能



圖 3.3 錄放音程式

無限制的抓取，如圖 3.3 為一可錄製和播放聲音的程式，可無限制的錄製，直到操作者關閉整個系統為止。同時又能錄製到檔案裡面，非常符合實驗要求。圖中，當按下 Start Record 的按鍵，在狀態欄後面的文字，Recording 會有顏色上的變化，用來告知操作者目前正在錄音當中。撥放聲音時也同樣有這種情形。

3. 錄音程式的改寫

既然可以存取聲音的波形，接下來的步驟就是分析判斷操作者是否輸入語音至麥克風中，並且判定操作者想要的出發按鍵，是按下滑鼠左鍵、左鍵兩下或是滑鼠右鍵，完成這些使用目的，才算是聲控輔助系統

最終的目的。所以在程式的寫法上，利用先前的錄音程式，去除用不到的程式碼，增加系統需求的功能及目的。完成如圖 3.4 的程式，工作說明如下：

系統錄製聲音的時間長短大約每 0.9 秒一次，接著關閉錄音程式。此時系統會取執行分析聲音波形的程式，並繪製出其聲音波形，同一時間內也做波形的判斷與滑鼠鍵的辨別及觸發執行，全部的步驟總共花費大約一秒鐘左右，因此操作者使用本系統時，輸入語音的時間須高於一秒鐘左右，使用本系統這樣才能更順暢。

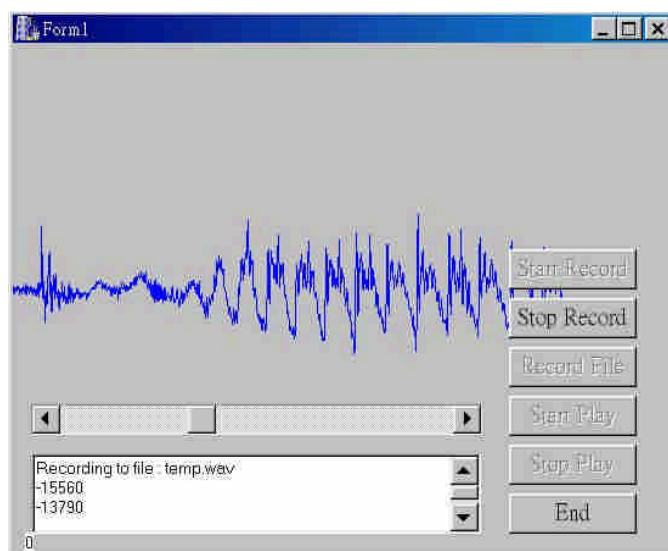


圖 3.4 判斷聲音輸入的程式

3.2.3 實驗討論

因為麥克風非常地靠近嘴邊，所以當有訊號輸入時，其聲音的波形振幅變化非常的明顯，即時有大聲的外界雜音進入麥克風，影響的振幅大小並不嚴重。但是麥克風品質的優劣、端子在音效卡的接觸情形與信號線的好壞，則影響振幅大小非常深，突如其來的雜訊，可能會影響波

形，導致程式在判斷程式上的錯誤與缺失。

在圖 3.4 中，下方的左右移動捲軸，功用為調整波形振幅判別的臨界值，假如操作者處於很吵雜的環境下或是硬體本身品質不良，產生雜訊介入時，調整捲軸可改善降低影響程度。但是，相對的使用者需提高輸入語音的音量，因為音量經過麥克風和音效卡進到電腦裡面，會改以電壓的型態來表示每取樣時刻的聲音強度，在根據這些電壓大小，轉換成資料型態，這其中當然也包含了雜音與雜訊的資料，所以提高臨界值的方法，可修正程式誤判的情形。

3.3 光學式追瞳器輔助系統之整合與功能

3.3.1 實驗設備：

1. PC 個人電腦：

Pentium II-350 電腦，64MB RAM

2. 使用的軟體：

Borland C++ Builder 4.0

3. 音效卡：

16bit 天音卡

4. 輸入裝置：

麥克風

5. 輔助配備：

耳機

6. 8255 晶片：

笛卡兒 8255 介面卡

7. 輸入裝置：

腳踏板

3.3.2 實驗步驟

1. 整體系統架構圖(如下圖 3.5)

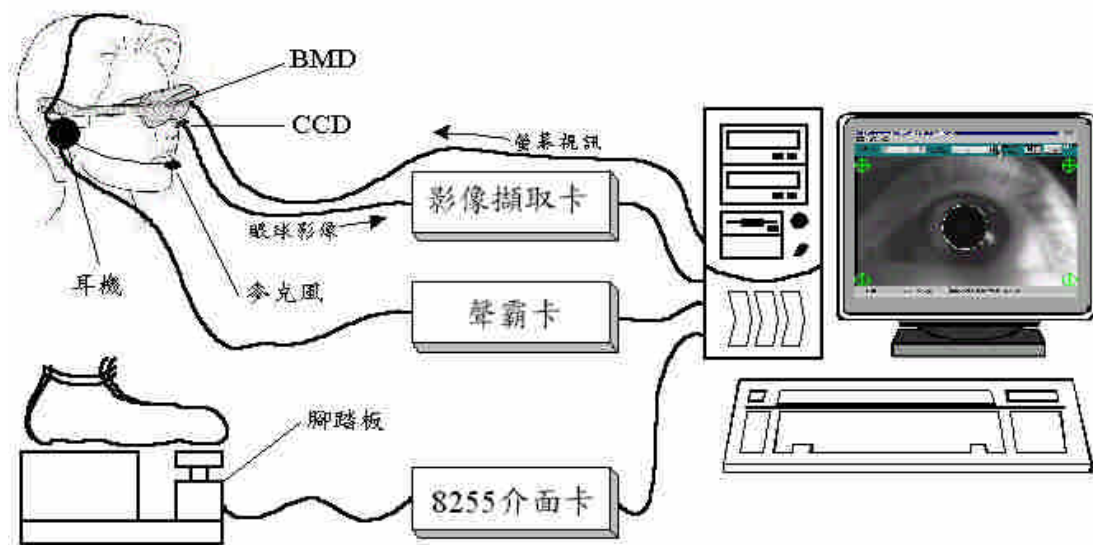


圖 3.5 整體系統架構圖

2. 整合輔助系統與追瞳系統

主系統追瞳器與聲控系統分別屬於兩個獨立系統，它的判斷法則不能使用足控的方式，在擷取影像後聲音的判別，因為抓取影像時，每秒處理 10 張影像，而聲音的錄製與處理的時間也近似一秒鐘，若強制整合在一起，必定會產生嚴重的問題，程式會起衝突。基於以上種種，把程式設計成為主系統與副系統，主系統為追瞳系統，副系統為聲控系統。當主程式第一次執行時，於程式執行某個階段後，主系統將會驅動執行

副系統，接著程式將同時執行兩套系統。主系統負責移動滑鼠的游標，副系統則專注於觸發滑鼠按鍵，這樣整個研究架構才算完成。

3. 相關操作畫面

架設完成整個硬體及軟體架構後，將頭配顯示器與耳機戴在頭上，接著開啟追瞳主程式，如圖 3.6 瞳位搜尋畫面。畫面上是由針孔式紅外線 CCD 攝影機所拍攝到的眼球瞳孔影像，為做二值化處理後的影像，必須適時的調整 CCD 的位置，使瞳孔落在畫面的中心處。綠色方框為搜尋邊界，可依照不同的眼睛大小，調整其搜尋邊界。白色圓圈為瞳孔的範圍，紅色十字代表白色圓圈的圓心，也就是眼睛瞳孔的中心，調整畫面上方的臨界值與對比值，使瞳孔大小能落在白色的圓圈內。

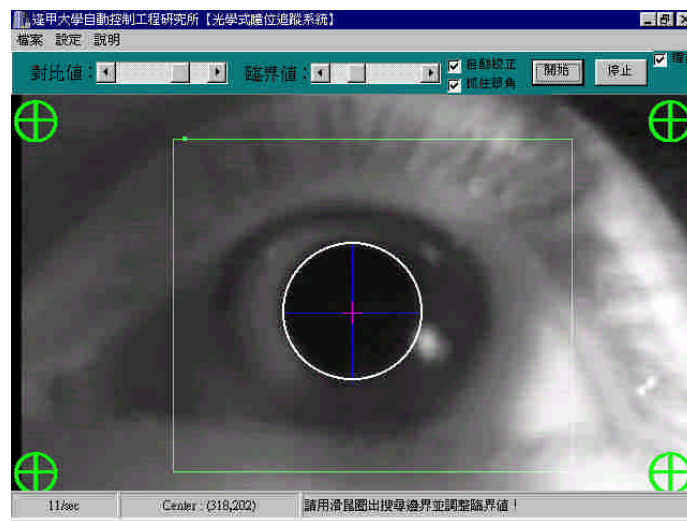


圖 3.6 瞳位搜尋畫面

若要使用輔助裝置，則需點選上方的設定鍵，選取輔助裝置參數設定，出現圖 3.7(a)的畫面，這是選擇輔助器具的設定選單，和滑鼠按鍵的一些設定，使用聲控環境狀況的設定。點選啟動聲控，變成圖 3.7(b)畫面。

圖 3.7(c)為凝視設定畫面，可選擇已有的凝視功能或是取消它。圖 3.7(d)設定滑鼠按鍵畫面，發長音時，設定其動作型態。

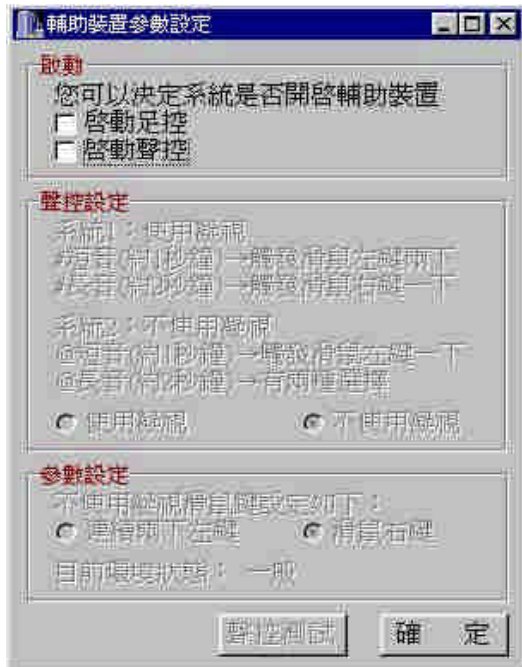


圖 3.7(a) 輔助裝置參數設定

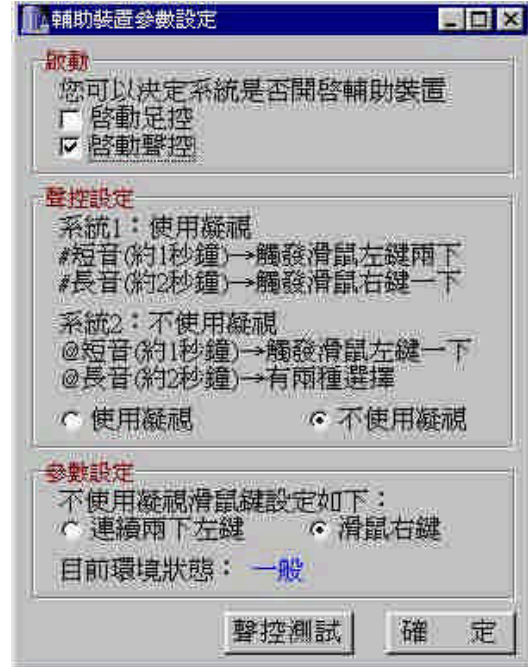


圖 3.7(b) 啟動聲控畫面

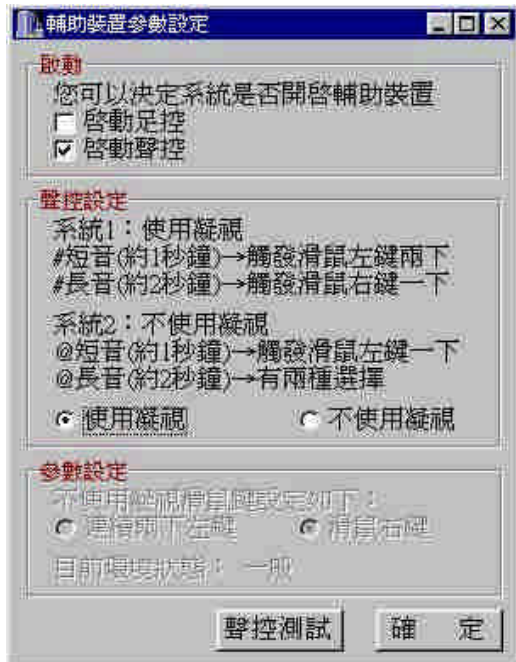


圖 3.7(c) 凝視設定畫面

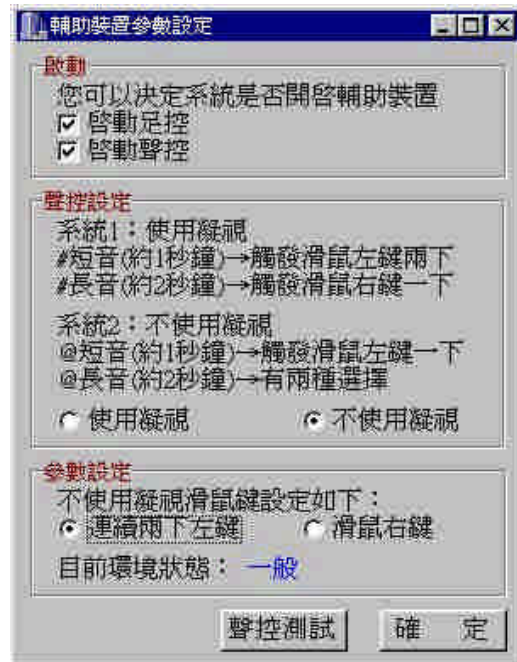


圖 3.7(d) 設定滑鼠按鍵畫面

在上圖中，若點選「啟動聲控」，就能使用底下的聲控測試，點選「聲控測試」，出現如圖 3.8(a)，初始環境狀態為一般，但可依照當時環



圖 3.8(a) 聲控測試起始狀態畫面

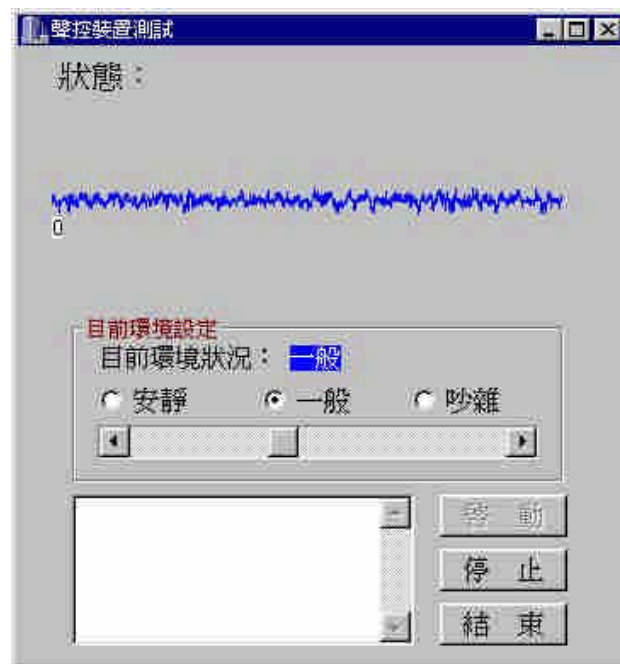


圖 3.8(b) 使用者未輸入語音前畫面

境的狀況，來選擇適當的環境參數。接著按下「啟動」鍵，開啟錄音程式。

若使用者未曾輸入語音，則會出現圖 3.8(b)的畫面，表示的意義是讓使用者得知是否麥克接線接對、有無故障，若是任何波形都沒有，則可能是接線接錯或則是麥克風有問題。若突然有波形振幅落差非常的明顯，則有可能是麥克風品質不良或是外界環境很吵雜，此時可調整臨界值來改善這問題。

輸入語音後就會出現如圖 3.8(c)的畫面，系統大約每一秒鐘變換一次輸入波形，目的就是要讓操作者得知自己輸入音量大小夠不夠，輸入語音的時間長短是否足夠，更重要的是要了解本系統的工作方式，經過多次的訓練，了解系統動作方式後，這樣在往後進入追瞳主系統中，才能依照多次的經驗，讓操作者更快速更容易的使用此系統。

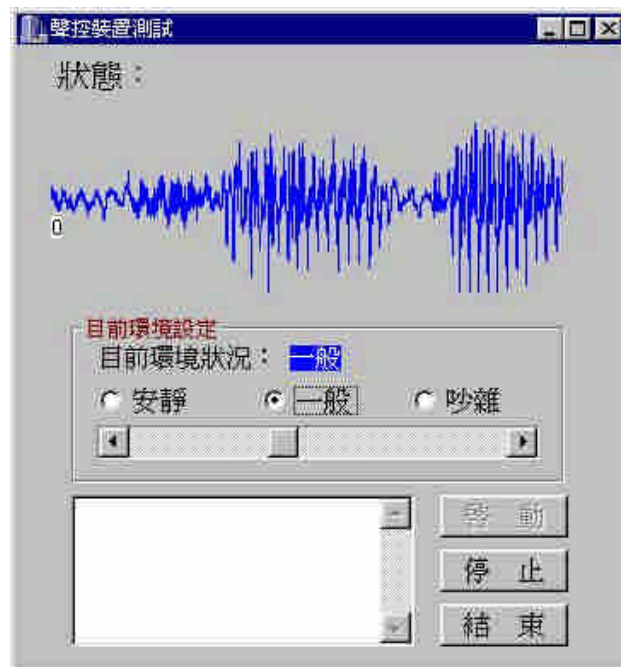


圖 3.8(c) 使用者輸入語音後畫面

等上述的動作完成後，點選圖 3.6 右上方中的確定鍵，開始作校正的步驟，追瞳系統的判斷法則會開始對瞳孔影像分析，在螢幕上會依序在上方、左方、中間、右方及下方出現標記符號(如圖 3.9(a)~圖 3.9(e)所示)，使用者只需配合注視頭配顯示器裡螢幕上的標記符號位置，直到 5 次的搜尋動作結束後，則整個瞳位追蹤的判斷完成。整個校正流程圖如圖 3.10 所示。

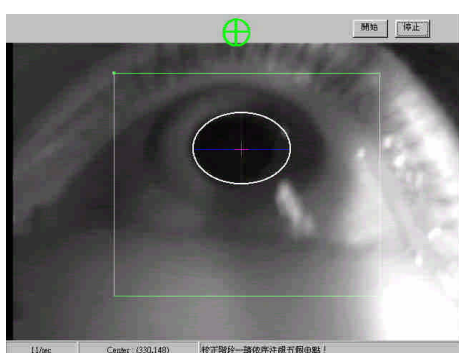


圖 3.9(a) 注視螢幕上方

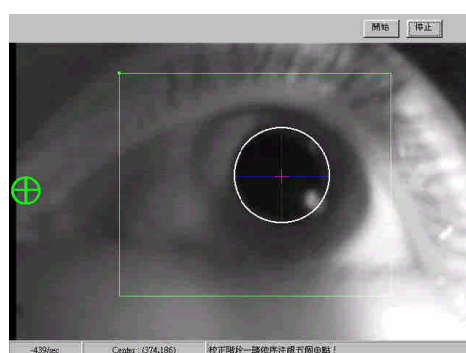


圖 3.9(b) 注視螢幕左方

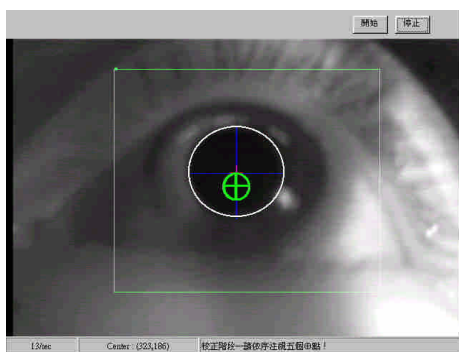


圖 3.9(b) 注視螢幕中間

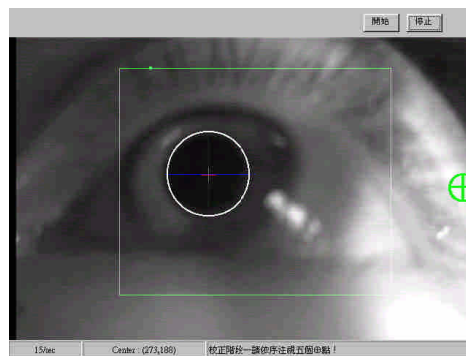


圖 3.9(a) 注視螢幕右方

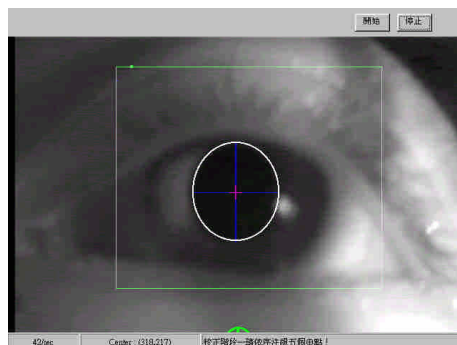


圖 3.9(a) 注視螢幕下方

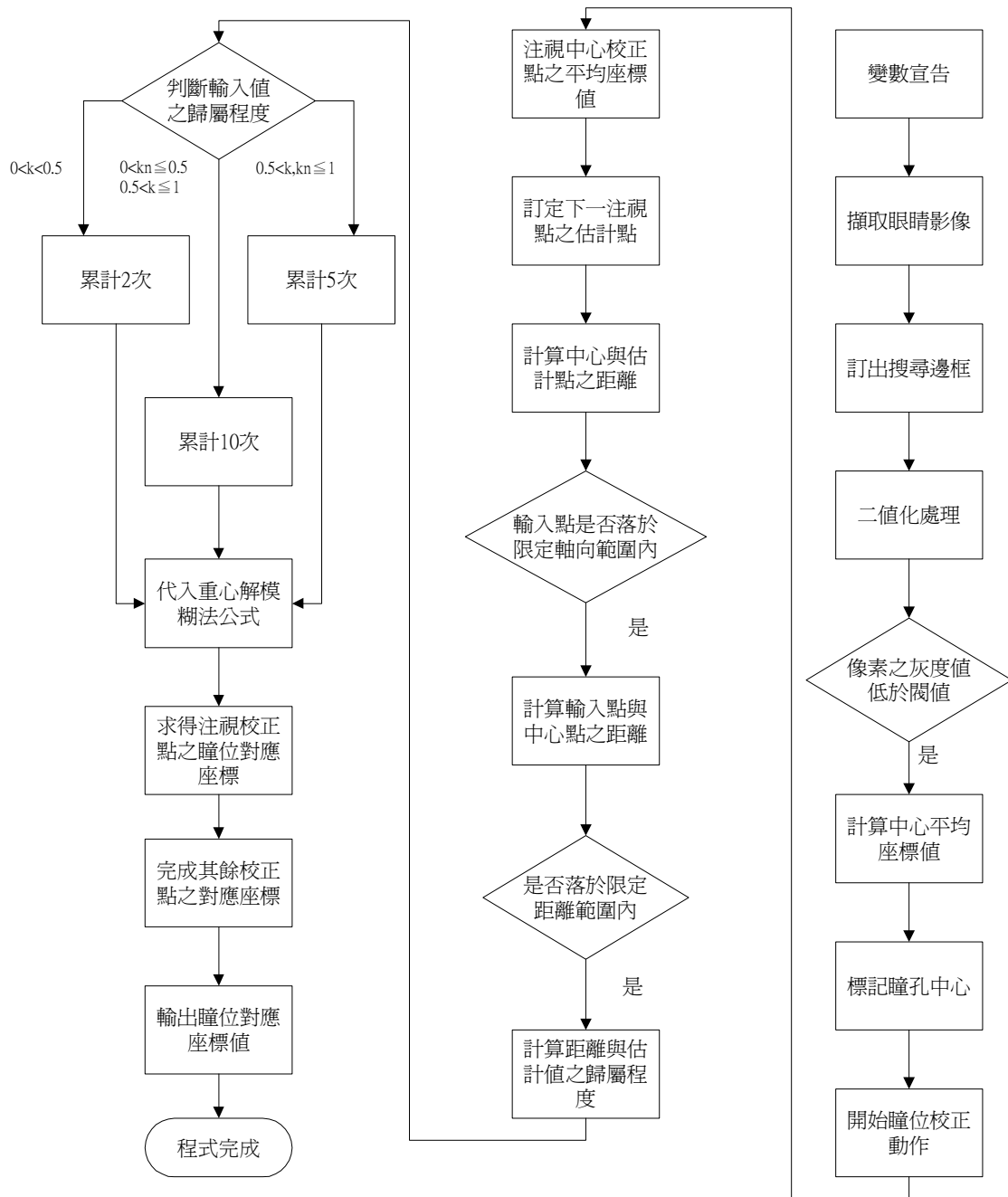


圖 3.10 瞳位追蹤校正程式之流程圖

做完校正步驟後，會出現如圖 3.11 確認畫面，紅色十字線條代表剛剛校正後，眼球移動得軌跡，此軌跡的形狀會影響以後用眼睛移動滑鼠游標的移動軌跡，造成移動上的誤差，相對的在點選目標物時，也會變的不怎麼靈敏，所以校正步驟是要進追瞳系統，非常重要的一道手續。按下「確定」鍵後，就開始使用追瞳系統，而且輔助裝置也能發揮它的功效了。同時按下「確定」鍵這步驟也是決定是否開啟聲控的功能，若在之前有勾選「啟動聲控」的話，那在進入應用程式後，便會驅動聲控的功能；倘若沒有勾選「啟動聲控」，則系統不會讓聲控作動。



圖 3.11 確認畫面

最後進入系統後開始使用輔助系統，其相關流程圖如下(圖 3.12(a)與圖 3.12(b))：

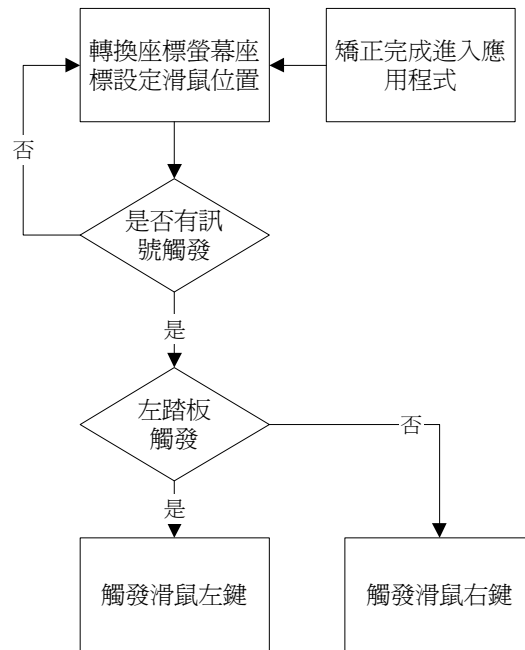


圖 3.12(a) 足控觸發之流程圖

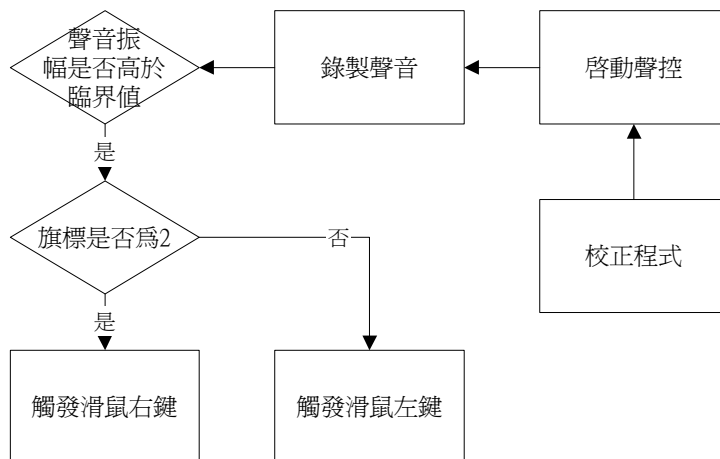


圖 3.12(b) 聲控觸發之流程圖

第四章 結果與討論

本專題最大收穫的地方在於參加『千禧年第二屆輔具大展』，那時有許許多多的參觀者，試用了本系統最初的成果，由衷的希望這套系統能幫助他們，解決操作電腦上的不便。然而沒有任何系統是完美的，許多使用者操作過後，都會給一些使用上的建議，為的就是讓使用者更容易上手，幫助更多的人。

系統完成後，我們比較各種觸發輸入的方式，統計出如下表六所示。在凝視這項中，由於使用者眼球容易飄動，並不容易停留在某個範圍作觸發，所需時間可能很短也可能很長。

表 4.1 各種輸入方式比較

類別	滑鼠	凝視	眨眼	足控	聲控
所需時間	最快	大於 0.5 秒 或更長	約 0.5 秒	次之	約 1 秒

另一個問題就是追瞳系統很容易產生誤差，理想狀況下，當我們注視到某一目標物時，滑鼠游標的游標應該停在目標物上。但事實上，根本做不到的，總是有那麼一點點的誤差。而且使用一段時間後，滑鼠游標也開始不聽話了，大多的原因是頭配顯示器與瞳孔間的距離，發生變化所導致的結果，此時又必須重新校正，才能順暢的使用。希望未來廠商能開發出自動對焦裝置，只需作一次的校正，儘管已移動頭配顯示器的位置，也能根據之前設定的焦距進行調整，增加追瞳器的可行性。

第五章 結論

在林宸生教授的帶領下，瞳位追蹤器的應用上已有很基礎的函式，接下去要突破的就是硬體介面的功能，古云：「工欲善其事，必先利其器」一點也沒錯，由於頭配顯示器的螢幕解析度不佳，所以會導致使用者操作幾分鐘後，眼睛就會產生疲累，這對使用者而言是一大傷害。其次就是 CCD 解析度不佳，使得程式在做判斷與轉換座標時，無法很正確的判斷出來，而使得滑鼠游標不聽使喚。另外在未來我們希望將聲控徹徹底底變成語音辨識控制，提高追瞳的實用性。將追瞳器推廣至家家戶戶讓大家在操作電腦上多一個新方式，在使用上多一分新感受。

參考文獻

- [1] Ramesh Jain, "Machine Vision", McGraw -Hill Corp, 1995
- [2] 胡錦標、林宸生、謝宏榮等, "精密光電技術", 高立書局, 民 79.12
- [3] 林宸生, "光電精密量測", 全欣資訊書局, 民 85.7
- [4] 楊忠煌、黃博俊、李文昌, "單晶片微電腦原理與實習－8051",
全華科技圖書, 民國 85 年 8 月
- [5] 黃肇祥, "C++ Buider 實務攻略", 松岡電腦圖書, 1998
- [6] 余明興、吳明哲, "C++ Buider 5 學習範本", 民國 89 年 7 月
- [7] 鄭鈞文、王敏成、邵守都 譯, "聲霸卡的聲音世界", 松崗電腦圖書,
民國 83 年 3 月初版
- [8] 許志興, "聲霸卡之應用與語音辨識", 旗標書局, 民國 83 年初版
- [9] 林宸生, "數位信號：影像與語音處理", 全華圖書, 1997
- [10] 楊武智, "數位信號處理入門", 全華書局, 民國 82 年初版
- [11] 宗孔德、胡廣書, "數位訊號處理", 儒林書局, 民國 81 年初版