

逢 甲 大 學
自 動 控 制 工 程 學 系 學 士 論 文

吹 控
Blowing Input System

學 生 : 劉 家 明
王 鵬 舜
指 導 教 授 : 林 宸 生

中 華 民 國 九 十 一 年 四 月

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另外，電腦網路的弗遠弗界，更縮短了人與人之間的距離，也使得網路資源的應用與取得更加容易。目前電腦在視窗的作業環境下，使用越來越容易，大部份的操作都可以藉由滑鼠來完成。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，操作電腦將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。對於有心想只用電腦的殘障人士(如手指不靈活或手殘等)而言並不適合。基於此，為了這些殘障人士提供無障礙的電腦使用環境，讓他們也能遨遊於電腦的世界裡，可以讓他們打打字，利用吹氣控制讓他們表示出他們想說的話，也可

以作為一個呼叫器來使用。針對於此，我們特別開發了這一套只要吹氣就可以控制滑鼠的系統，以其對這些不善表達的病有些許的幫助。

1.2 文獻回顧

十八世紀出，法國數學家 Jean Baptiste Fourier 探討出如何將自然的聲音分解成一個個相同的基本元素，由他所發展的這套數學方法可以將一個聲音分解成有限個不同的正弦波。這個步驟就是著名的傅利葉分析。

較複雜的聲控，可以將典型語音處理技術分成下列五種：1. 語音編碼 2. 語音合成 3. 語音辨認 4. 語音辨識 5. 語音了解。通常，一般的語音處理將之著重在語音辨識、語音編碼、和語音合成。

在這科技發達的時代，語音辨識的技術已經可以做到有：證券業語音辨識股票下單系統、自動總機系統、電話銀行服務系統、航空或鐵路訂票系統、查號台服務系統...等。

而就以吹控輔具所需，較重要的是基本的語音辨認，並不需要其他較深的語音技術。我們只需利用聲霸卡抓取信號，然後再將訊號轉為數位化，加入判斷後我們就可以代替手動滑鼠在電腦上移動與點選的行為了。

1.3 研究方法

1. 聲控系統的架設與運用

在整個研究過程當中，我們首先針對吹控輔助系統的聲霸卡以麥克風作為溝通的橋樑，利用對麥克風吹氣之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，依得到的訊號結論，再設計出吹

氣控制器可取代手動滑鼠的基本的執行命令。

2. 聲控系統判斷方式簡述

我們區分計算聲音輸入後所計算之次數，其方式主要可分為兩種：

(1) 觸發一信號後於一單位時間內如果沒有第二個觸發信號，就判定為一次觸發，次數為1；連續觸發兩信號後，在一單位時間內沒有第三觸發信號，就判定為二次觸發，次數為2，三次、四次.... 以此類推。

(2) 規劃一單位時間後計算所有時間內觸發數之總數，為本系統採用的計算方式。

3. 靈敏度的調整

我們設計一可調整的 scroll bar 來去除雜訊。

第二章 理論基礎

2.1 信號概說

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，我們把它稱為數位訊號 (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。

吹氣訊號在進入電腦之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，中間沒有遺漏。而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，如下圖 2.1 所示，利用 c++ Builder 軟體所顯示的波形，任何時刻週期都不一樣。

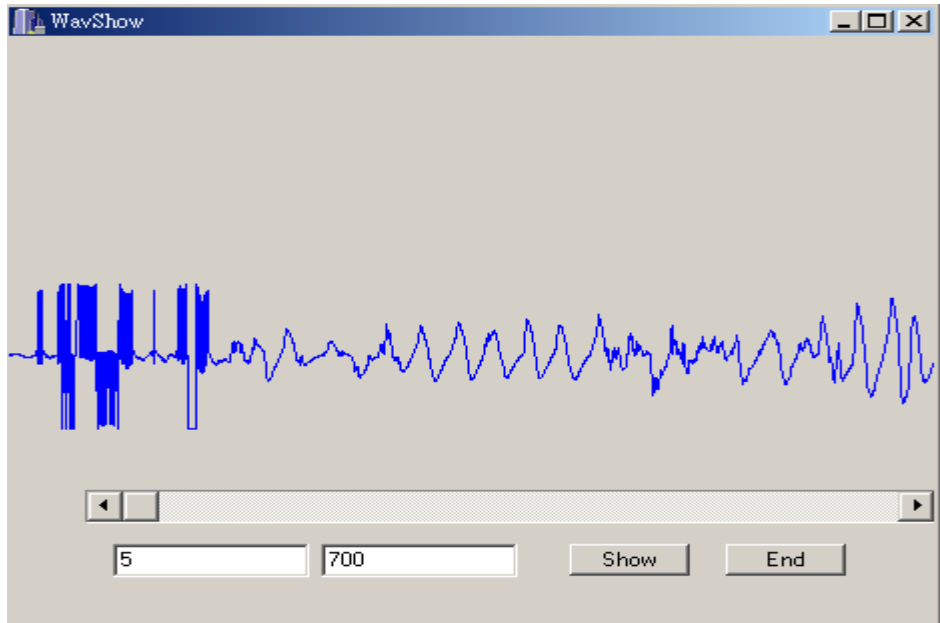


圖 2.1 波形圖

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。因此依照 Shannon Theorem[1] 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 2.2(a))，所以取樣頻率設定在 8000Hz(圖 2.2(b))以上，就能保證取樣失真不會發生。

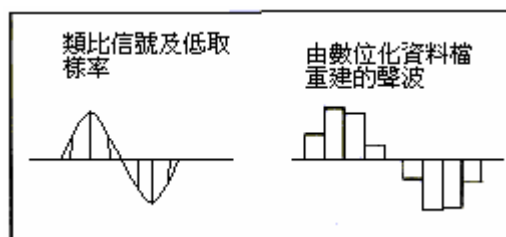


圖 2.2(a) 低頻率取樣類比信號圖

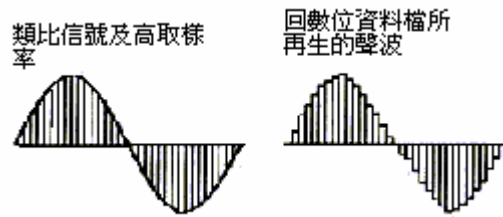


圖 2.2(b)高頻率取樣類比信號圖

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。

2.2 快速傅利葉(FFT)

早期的傅利葉轉換大都利用手算來完成，因此傅利葉轉換標本個數都很小，連續時間下的傅利葉轉換其定義為： $F(w) = \int_{-\infty}^{\infty} f(x)e^{-iwx} dx$ [2]

在連續時間下的系統輸入之信號多為類比的方式，但是在許多情形下的某些系統輸入函數卻呈現著離散的型態，例如每隔一段時間間隔才量取的實驗數據，每隔一段距離間隔才量化的影像資料，這時候我們就必須將原來的傅利葉轉換作視當的修正，從連續時間下的傅利葉轉換轉到離散時間傅利葉轉換

(Discrete-Time Fourier Transform; DTFT)時，其定義為：

$$X(e^{jw}) = \sum_{n=-\infty}^{\infty} x[n]e^{-jwn} \quad [2]$$

我們留意到離散時間之序列之 n 值為正負無窮大，而其傅利葉逆轉換仍為一連續型態，因此使用積分來表示，其本身仍為一連續變數之函數。但是在實際的情形我們處理的數據資料其長度應該是有限的，亦即序列之 n 值為有限，因此這時候我們必須再將連續時間傅利葉轉換作修正，這便轉換到了所謂的離散傅利葉轉換 (Discrete Fourier Transform; DFT)，其定義為；

$$X[k] = \sum_{n=0}^{N-1} x[n] W_N^{kn} \text{ 其中 } k = 0, 1, \dots, N-1; n = 0, 1, \dots, N-1 \quad [2]$$

由於離散傅利葉轉換在計算上可以有取巧的地方，就變成所謂的快速傅利葉轉換。需注意的是 FFT 之峰值會受信號頻率，與取樣時間影響。

2.3 PC 聲霸卡

圖 2.3 聲霸卡將是一個不可或缺的產品，沒有它的話，充其量電腦只是一台計算機而已，根本不適用從事語音處理。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。

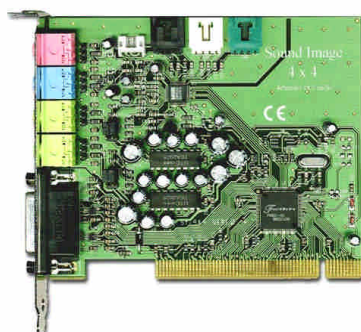


圖 2.3 聲霸卡

2.3.1 8-bit 單音

當了解的聲霸卡的聲音取樣格式之後，就可以利用程式語言對信號作出轉換與讀取的動作，進而再加上預期的功能。一般來說，音效卡的規格基本都具有 8-bit 及 16-bit 單音及立體聲錄放音，使用者定義取樣頻率從 5kHz ~ 48 kHz，但較好的可以達到 24bit 之錄放音。在一個 8bit 的取樣標本裡頭，所有的 byte 都接連的存放，當信號經由數位化之後其值就介於 0~255 中間，再藉著濾波器的使用就可以消除不必要的雜訊。以下是其軟體程式的宣告：

```
__fastcall TAudioLineParameters::TAudioLineParameters() : TPersistent()  
{FChannels=1;FFrequency=11025;FBitsPerSample=8;}[3]
```

2.3.2 8-bit 立體

在立體取樣中，我們以兩個 channel 分別來代表左右聲道，在一個立體的取樣中，資料是依交錯格式安排的，也就是說在一個取樣中是由 2byte 所組成，一個是左聲道另一個是右聲道，永遠一個交替的順序下去。

2.3.3 16-bit 取樣

對於一個 16-bit 的取樣，其值介於 -32768~+32767。在這情況下，0 這個值就代表中間的數值。

第三章 實驗方法

本次專題報告的研究為一個部份，『吹氣控制滑鼠』以市面一般的麥克風作改良，加上音效卡，以個人電腦為將得到的因緣作處理。

3.1 吹控輔具裝置

3.1.1 實驗設備：

1. PC 個人電腦：

AMD-450 電腦，128MB 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(如下圖 3.1 所示)

5. 輔助配備：

耳機



圖 3.1 耳機附麥克風

3.1.2 實驗步驟

1. 我們特地將所買的麥克風上面的氣孔予以鑽孔加大，再架構出如下圖 3.2 的聲控硬體配備。

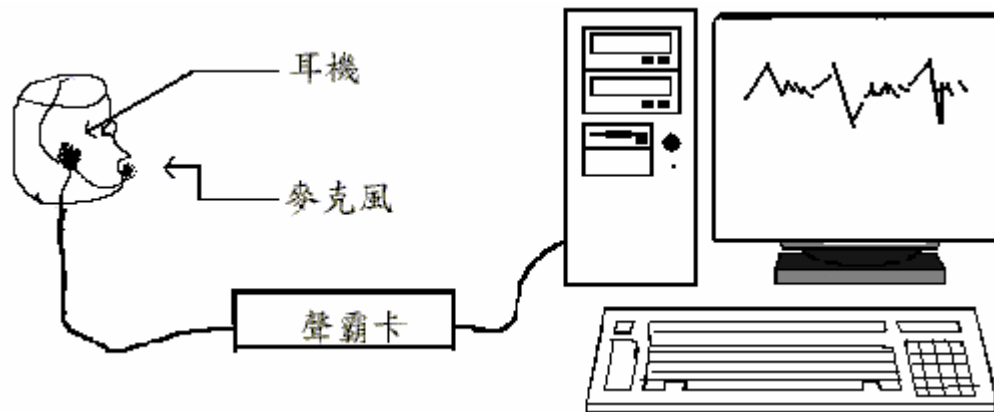


圖 3.2 聲控系統硬體配置圖

2. 再來由使用者吹氣測試訊號是否已由軟體抓取，如下兩圖。



圖 3.3 (a) 無吹氣訊號圖



圖 3.3 (b) 有吹氣訊號圖

3. 按下啟動聲控的按鈕之後，依照吹一下向上、吹兩下向下、吹三下向左、吹四下向右之方式開始移動。於移動中，再吹氣一次即可使滑鼠停止不動，停止

時滑鼠不點選。如果欲使滑鼠點選一次就停留原地不吹氣五秒，如欲使滑鼠雙擊則停留原地兩次週期(也就是 5+5=10 秒)。而倒數時間的快慢可以經過使用者自己的調整。

4. 可開始上網閱讀文章，也可使用 word 書寫文章和使用一些基本操作功能。

3.1.3 滑鼠移動與點選順序

下圖為滑鼠移動與點選順序流程圖 3.4

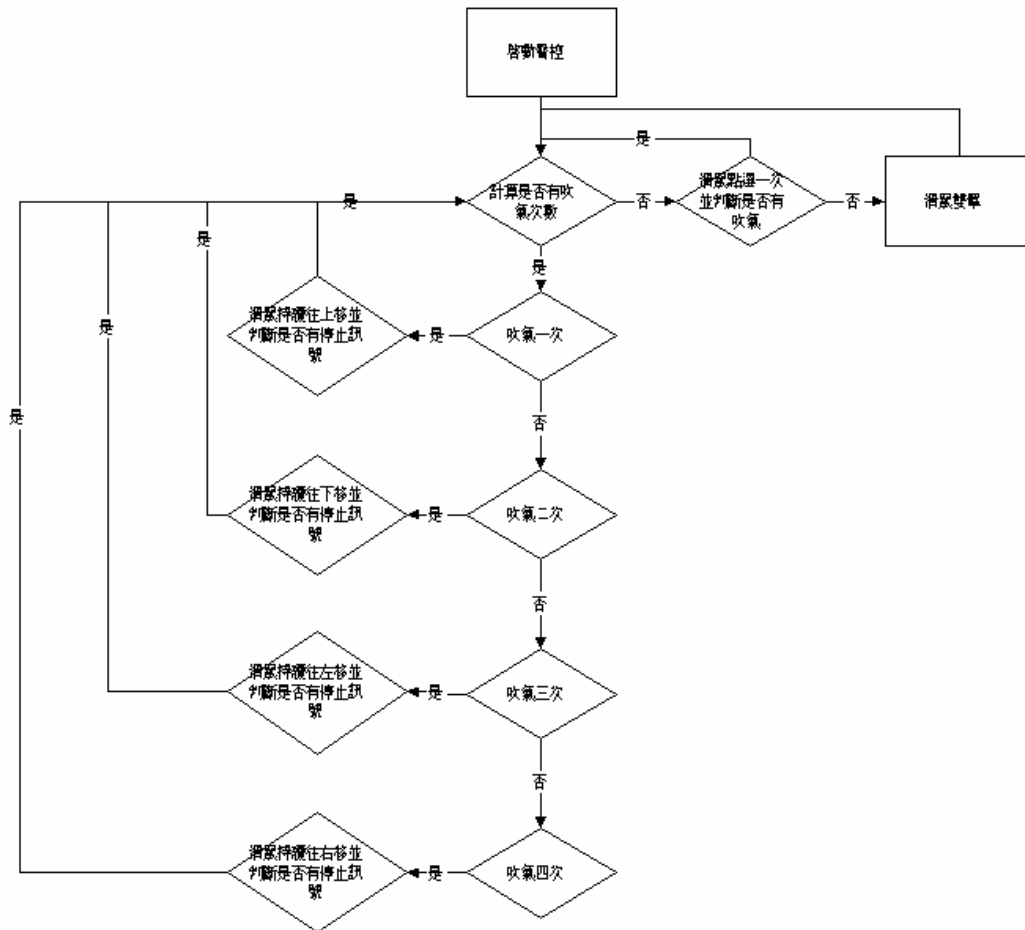


圖 3.4 滑鼠移動與點選控制流程圖

3.1.4 實驗討論

經過自己與患者的測試之後，一般來說，對於輸入訊號抓取的不正確、失誤大可分為硬體與軟體的問題。

在硬體方面，因為我們的吹控輔具是設計給行動不方便之人使用，而這些患者可能因長年缺乏運動，所以其呼氣之力道也較微弱。因此，他們所使用的麥克風需較一般人靠近嘴邊，而且系統的靈敏度要求也比一般人高，在吹氣之麥克風的位置要求也更嚴格，這樣才可符合他們的需求。也因為麥克風離臉部較近，所以當患者移動頸部時便可能碰撞到麥克風，在呼吸時的氣流也可能形成觸發信號，而此我們在麥克風前加一層薄海綿以消除這些雜訊，否則這些雜訊有可能進而影響吹氣次數的計算，這也是需要使用者注意的地方。此外，經過我們測試麥克風品質的優劣，發現端子在音效卡的接觸情形與信號線的好壞，則影響信號大小非常深，突如其來的雜訊，可能會被當成輸入，進而導致程式在判斷程式上的錯誤與缺失。

在軟體方面，因為在程式的部分我們只是做比較基本的控制，例如：我們並不擔心使用 FFT 轉換時，因不同的輸入信號（吹氣力道之大小）導致不同的峰值，因其轉換出來之峰值遠大於我們所設定觸發之門檻，反而比較重要的是吹氣區域是否落於有效取樣點上，信號之取樣如下圖。而取樣時間間隔的設定在經過我們的病友—瑞澤的實際測試之後，發現於每 0.2 秒判斷信號一次，可以是最舒服、輕鬆的。

虛線區域為吹氣之長短時間，如果虛線區沒有落於0.2、0.4、0.6之取樣點上，則吹入之信號無效，也就是不計算次數。

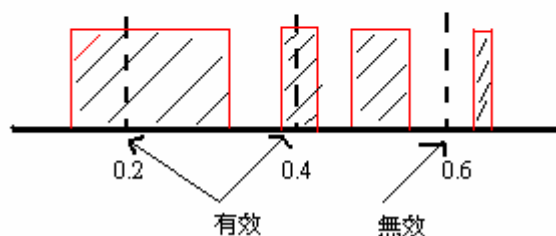


圖 3.5 有效訊號圖

第四章 結果與討論

這一套自行製作的陽春吹控裝置還有許多的缺點與功能上需要加以努力改進的。首先我們比較了病友—瑞澤自身的吹控裝置，他使用的吹控裝置是從國外引進的，其功能如下：滑鼠的移動是由嘴中的搖桿控制，滑鼠的點選功能以一個感測器判斷吸氣跟吐氣，然後作拖曳與點選的作用。所以基本上，就以單純的控制而言，其敏捷度與準確度並不輸給手控滑鼠，其控制效率當然是在麥克風控制系統之上。但是就如果是用於舌頭的移動也不是很靈活的病友而言，相信只用麥克風控制滑鼠的系統將可以為他們帶來一點便利與實用性。

再來，我們和瑞澤實際操作本系統時發現的一些缺點如下圖。



4.1 病友實際測試圖

再經過與病友的實際討論之後，發現本系統麥克風是否準確的接收吹氣之訊號，對病友的體力上有很大的影響。

下表 4.1 是本系統在 Windows 作業系統下測試的時間表：

功能	秒數
開啟瀏覽器	15
輸入一中文字	<60
移動速度	50pixel/s
關機	40

表 4.1 測試時間表

第五章 結論

本系統在多方改善與檢討之後，發現在信號計算次數方面可以在更快一點，我們將改採第二章的計算聲音輸入後，計算觸發次數方式中的第一種方式重新規劃本系統。如此，只要使用者可以適應在 2.5 秒內輸入想要之信號次數，理論上就可以省下將近一半的時間，像是輸入一個中文字就可以控制在 30 秒之內。

再來就是規劃新的硬體，如下圖，以 8255 卡作為訊號的轉換，再加上一具有活塞之光感測器來偵測吸與吹之動作，在以設計出類似搖桿的硬體來控制方向。如果這個概念可以實現，我們就可以把吹控的功能大幅度的向上提昇，也能為行動不方便之病友帶來更有用，方便的滑鼠替代工具。[4]

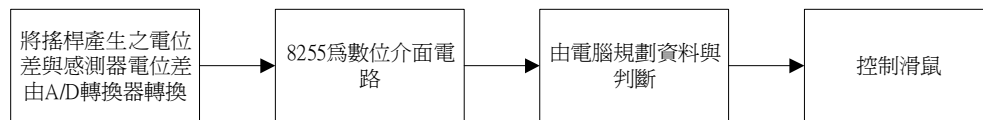


圖 5.1 吹控硬體概念圖

參考文獻

- [1] 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
- [2] 林宸生，"數位信號 - 影像與語音處理"，全華圖書，民國 86 年 12 月
- [3] <http://www.wischik.com/lu>
- [4] 陳龍三，"聲控電腦製作與應用入門"，松崗電腦圖書，1997 年 11 月初版