

手語辨識與翻譯系統之研究

A Study of the recognition and translation system of sign-language

專題組員：蔡明勳、范元駿、陳維沼、蔡育庭

專題編號：PROJ2015-BIOINFO-10103

指導老師：董其樺老師

摘要

本專題研究旨在第一人稱利用 Leap motion 感應裝置與 Leap Trainer 進行手勢偵測，應用於受測者比畫各種手語意義的準確率之探討。此外，本專題亦使用 Processing 語言結合 Leap Motion，編寫出能夠辨識出左右手數字 1~10 之程式，探究了解偵測手部動作的原理和規則。經測試，本專題所實作的裝置，所進行的手語辨識度平均為 49.75%，而最高辨識度可達 85%。我們期望此研究未來能為全球約三億六千萬名以上的聾啞人士和一般大眾立即、自在地進行溝通交流與對話。

關鍵詞：手語、影像辨識、Leap motion

1. 簡介

在現今資訊傳遞發達，溝通便利的時代中，根據統計，聽力損失的全球患病率估計值約 3.6 億人，超過世界人口的 5%[1]，其中有超過 12 萬人在台灣[2]。我們日常生活中覺得再簡單不過的閒聊，在聾啞人士眼裡是件多麼奢求的事。除此之外，搭捷運時曾看到站務人員與聾啞人士溝通，需利用電腦設備搭配已有的網路視訊功能來向遠端的手語翻譯人員溝通。由

於此系統需要使用到電腦、攝影機、網路，但不是隨處都有良好的網路訊號，且電腦設備並非隨時能夠攜帶。即使現在已有方便於攜帶的智慧型行動裝置也不是最簡單、最即時的溝通管道，在諸多不便利之下，時間上可能導致有所拖延，無法讓聾啞人士在第一時間或很緊急的時刻向一般民眾提出溝通的需求。因此，若能開發相關系統或產品，讓聾啞人士和一般大眾能更沒有距離、更自在地進行溝通交流與對話，造福國內(甚至是全世界)的聽障或語言障礙的同胞，使其感受到這個社會無私的關懷與無差別的尊重，如此一來，這個世界將能更加和諧。

未來達成這個願景，現今在美國新創公司 MotionSavvy 已開發出一個叫做 Uni，能翻譯手語的平板電腦，可以捕捉手部動作並將其翻譯為文字[3]。此 Uni 裝置是以第一人稱方式擷取聽障者所比劃出的手語動作，加以翻譯。但聽障者一隻手要拿著平板，一隻手比劃手語，可能會讓使用者能比的手語種類有所侷限，也無法流暢地與平常人溝通。

為了能夠偵測手勢動作，目前有許多研究利用影像辨識相關之演算法。例如：「在複雜背景下做即時的手

部偵測與追蹤」中提到的物件偵測演算法[4]；「手勢影像辨識率增強方法之研究」中所使用的演算法，則偵測輸入影像的邊界框比例及質心的不同，並可去除上手臂的資訊[5]；「基於手形輪廓之智慧型手勢辨識系統設計」論文中採用類神經網路的演算法，其研究結果顯示其對於手勢 1 至 5 的辨識率可以達到 100%，而對數字 0 至 9 的辨識率經由增加的輸入層神經元個數至 90 個時，也可達到 100%[6]。除了手勢辨識演算法外，也有許多研究是辨認手勢後應用至硬體控制上，例如「即時手勢辨識系統應用於機上盒控制」、「猜拳機」、「手勢之影像辨識及應用於高爾夫球撿球車」等，都是擷取手勢影像，達成即時辨識目的來判斷手勢指令，再將指令轉而控制電子儀器[7-9]。

本專題研究欲改善先前研究裡以第二人稱進行翻譯手語會產生的問題，率先嘗試以第一人稱裝置欲幫助聽障者與平常人更直接自在的溝通，與以往其他研究所採用的視角有所不同。我們所設計的系統，初步是透過手勢辨識裝置 Leap motion 配戴在語障者胸前，用此偵測使用者手部動作，翻譯其比劃的手語。在此專題裡，我們以不同測試者和不同手語動作會有不同手勢比劃的角度，做系統性的準確率計算。

2. 專題進行方式

Leap motion：是一台內含動作感應裝置的小型機盒。動作感應裝置會根據內置的兩個攝像頭從不同角度捕捉的畫面，重建出手掌在真實世界三維空間的運動軌跡。檢測範圍大約在

動作感應器上方 2.5 公分到 60 公分之間。此裝置連接電腦後，我們可以透過觀察器偵測手指及手掌，在一定視野範圍內，能獲得手指(長、寬、方向、指間位置及速度)的動作，加以識別。



圖一、Leap motion 手勢擷取裝置。此裝置內含兩個攝影鏡頭，可偵測立體空間內，手掌及手臂的動作。

Processing：一種電子藝術和視覺互動設計開發的程式語言。此語言建立在 Java 語言的基礎之上，但使用簡化的語法和圖形編程模型，目的是輕易寫出可視覺化的互動式應用程式。此語言可搭配 Leap motion，以此做為外部控制裝置。

Leap Trainer：手勢辨識之網頁程式，結合 Leap motion 後，可辨識手勢且能學習手語並捕捉運動期間的姿態。該網頁利用類神經網路等機器學習演算法，讀取 Leap motion 所輸入之手勢訊號。轉換成量化數值後，再經類神經網路的分析，即可將所學習記憶的手勢動作儲存下來。

Processing 結合 Leap motion 及 Leap Trainer 結合 Leap motion，共同之處就是能以第一人稱大量辨識手

語；不同之處，則是 Processing 偏向顯示圖文並非學習手勢，但 Leap Trainer 能學習手勢並記憶。

圖二顯示我們透過 Processing 程式，可偵測使用者比的手語數字 1 至 10。此 Processing 程式不但能正確偵測數字手語，也能區分左右手。



圖二、Leap motion 與 Processing 程式語言結合。圖中顯示當右手比 5、左手比 6 時，可分別偵測出並顯示其數字。



圖三、Leap motion 與 Leap Trainer 整合。圖中使用者比出手語「人」，螢幕正確辨識出人，並顯示符合程度 78%。

此手語辨識系統模型之外觀，其設計的原始理念是做成項圈的形

式戴在使用者胸前，但是研究後發現此裝置胸前面積需要足夠大，以穩定整個 Leap motion 裝置，增加手勢辨識的穩定性。因此我們現在是利用厚紙板製作成能讓受測者自由穿脫手語辨識模型，模型上 Leap motion 的位置也和當初的理念項圈上位置一樣，不影響偵測手勢的結果。

Leap Trainer 其原理是將 Leap motion 所擷取的手勢影像，根據其向量、角度、位置、速度等數值，經由類神經網路等機器學習演算法，輸出轉換成編碼。我們可將其編碼，輸入至 Jscript 程式「leaptrainer-ui.js」裡的 `trainer.fromJSON('')`；這段程式碼的單引號中。如此即是將學習記憶後的手勢動作紀錄至網頁裡。當重新開啟 Trainer 網頁，所學得的手勢便能載入並顯示於 Trainer 網頁，留待辨識使用者的手語手勢(如圖三)。

3. 主要成果

圖四表示此裝置之雛形 (prototype)，以厚紙板做成外殼，並可讓 Leap motion 附掛其上。將此裝置配戴在比畫手語的使用者胸前，實際使用情況則如圖五所示。

表一、手語辨識次數及其辨識率

手語	運動	你	車禍	敲打	走開
成功辨識次數	17	19	33	23	11
失敗辨識次數	18	17	5	5	1
無法辨識次數	5	4	2	12	28
辨識率	42.50%	47.50%	82.50%	57.50%	27.50%
平均符合程度	67%	71.47%	70.76%	69.22%	66.82%

手語	如何	人	趕快	不客氣	擔心
成功辨識次數	25	34	5	26	6
失敗辨識次數	15	4	5	0	0
無法辨識次數	0	2	30	14	34
辨識率	62.50%	85%	12.50%	65%	15%
平均符合程度	70.68%	73.00%	66.20%	71%	67.17%



圖四、Leap motion 項圈模擬模型。

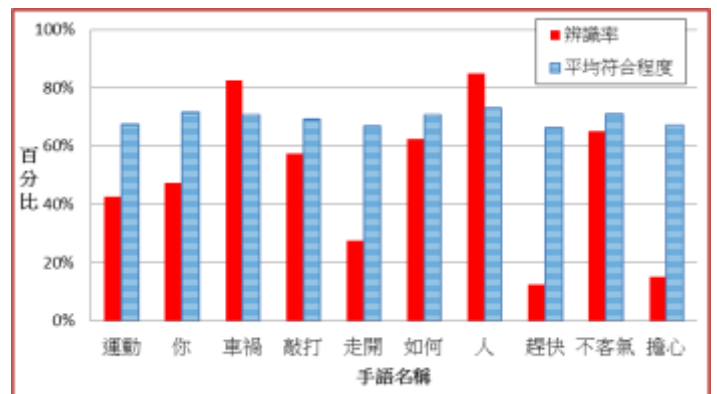
最後統整結果，計算每種手勢的辨識率和平符合度，公式如下。辨識率為總辨識次數中辨識成功次數的百分比，如公式(1)；平均符合度是所有成功辨識的次數中，各辨識成功百分率總和之平均值，如公式(2)。這兩筆百分率值同時越高，表示該手勢具有越高的辨識穩定性。

$$\text{辨識率} = \frac{\text{辨識成功次數}}{\text{總辨識次數}} \times 100\% \dots\dots\dots (1)$$

$$\text{平均符合程度} = \frac{\text{各辨識成功百分率總和}}{\text{辨識成功次數}} \times 100\% \dots\dots\dots (2)$$



圖五、手語辨譯項圈模擬模型實戴。



圖六、手語辨識率及符合程度長條圖

為了測試手語辨識的準確度，我們所設計的實驗方法，是四名受測者，每人各測試 10 種手勢，每種手勢反覆測試 10 次。每次皆記錄正確辨識、不正確辨識及無法偵測的結果。

在表一與圖六裡的辨識率和平符合度兩者數值越高且越相近，表

示該手語的平均辨識率越良好。由圖六可知，在所有測試的手勢當中，以「人」的平均辨識率是最高的。推測是因為此動作變動範圍較小且與其他測試手語較不類似及衝突，所以辨識成功的次數較高。此外，辨識率次高的手語是「車禍」。此手勢因動作較於簡單、手指與手指和手掌間不交疊且動作分明、手勢的變動範圍不大，又剛好位於裝置鏡頭的正前方，利於鏡頭擷取完整的手勢，導致平均辨識率較高。

相反地，十套測試手語中，「趕快」和「擔心」這兩種手語辨識成功的機率較差。我們推測，因為此兩手勢是一直重複不停的動作，在學習時動作軌跡較漂浮不定，而且手指相互重疊的部分較多。在這兩種不利辨識的狀況下，在 Leap Trainer 分析擷取比對時，較難與當初學習手勢的吻合，因而導致平均辨識率較差。而且，若手臂揮動的範圍幅度太大，超出裝置鏡頭的範圍，無法擷取完整的手勢。

4. 評估與展望

本專題進行至此的進展，已初步實現以第一人稱進行手語辨譯的想法。然而現在這個作品，在現階段仍屬於半成品的模樣，因為穩定性與材質的關係，造成辨識率並不是十分的高，況且必須接著電腦才能夠使用，導致無法獨立帶著出門辨識，與當初的想法有點落差，且各國的手勢都不太一樣，收入資料庫方面目前有一定難度，只能先作出簡單的手勢，並由其他網頁記錄手勢位置，

在國內一篇「數字手語之辨識於語音合成轉換系統之研究」的碩士論

文[10]，提出了一套可翻譯數字手語並發出語音的系統。該研究將手語翻譯成文字後，再透過文句轉語音之軟體，使其能夠產生語音的輸出。未來，我們也期望可以結合文字輸出語音的軟體，讓我們裝置所翻譯的語意，可以撥放出聲音。

近期有一個研究「手語翻譯手套」[11]研發出手語手套，使用者能夠戴在手套比手語，並在手機 app 上進行翻譯發聲。他們的穿戴方式是以手套的形式穿上，而我們的穿戴方式是以項鍊的方式戴著。以便利性而言，我們的概念較能使聽障人士容易使用和操作。

我們未來理想的手語辨譯項圈外觀，呈現的是服貼在使用者的頸部，如圖七所示。本裝置會像配戴飾品一樣的自在、便利、美觀，又能以第一人稱的方式幫助聽障者辨譯手語與正常人直接無礙的溝通。



圖七、手語辨譯項圈模型呈現

5. 結語

本專題實作出一個可學習辨識位於胸前簡單手勢的項圈裝置模型，並且經系統性測試，整理出各辨識率統計圖表。結果顯示，本系統可分別辨識兩手比出數字的手語，而日常常見

的十種手語平均辨識率為 49.75%，而其中比「人」的手語辨識度最高，可達 85%。本專題目前只能辨識出部分的手勢，未來希望以第一人稱的方式配戴項圈，能辨識出更多手勢，並可翻譯其比劃的手語，最後更期望能發出其手語意思的語音，造福聾啞人士們，使其感受到這個社會無私的關懷與尊重，讓這個世界將能更加美好。

6. 銘謝

感謝董其樺老師對於本專題的細心指導，在一年多以來的詩間，本團隊在老師的引領下，接觸了完全不熟悉的手勢感應裝置和程式語言。剛開始接觸 Leap motion 及 Processing 時，所有人都充滿了挑戰與期待，老師與我們皆是初次摸索該裝置，在老師按部就班的帶領下，一起研究、嘗試、探討，培養出彼此間齊心協力的默契，在這領域上有所了解與成長，感謝卓越的董其樺老師奉獻大量的心血與時間在本專題上，對於本專題增添了不少的精采。謝謝老師。

7. 參考文獻

- [1] World Health Organization (2013) *Multi-country assessment of national capacity to provide hearing care*. Switzerland: WHO Document Production Services
- [2] 統計處 (2012 年 8 月 25 日)。一〇一年第三十四週內政統計通報(101 年上半年身心障礙者福利統計)。內政部。
- [3] 李英嘉(2015 年 2 月 24 日)。由聽障者親自打造的無障礙溝通平板，讓你「聽見」手語的聲音。社企流。2015 年 12 月 10 日，取自 <http://www.seinsights.asia/taxonomy/term/3232>
- [4] 毛綱增、吳怡樂(2009)。在複雜背景下做即時的手部偵測與追蹤，國立台灣科技大學資訊工程學系碩士學位論文
- [5] 周士強、陳佑冠(2010)。手勢影像辨識率增強方法之研究，華梵大學電子工程學系碩士學位論文
- [6] 余玉田、陳永平(2010)。基於手形輪廓之智慧型手勢辨識系統設計，國立交通大學電機與控制學程碩士學位論文
- [7] 李經寧、蘇木春(2009)。即時手勢辨識系統應用於機上盒控制，國立中央大學資訊工程學系碩士學位論文
- [8] 曹文潔、王文俊(2007)。猜拳機，國立中央大學電機工程學系碩士學位論文
- [9] 吳佳展、王明賢(2007)。手勢之影像辨識及應用於高爾夫球撿球車，南台科技大學電機工程學系碩士學位論文
- [10] 方如玉、陳世中(2009)。數字手語之辨識於語音合成轉換系統之研究，南台科技大學電子工程學系碩士學位論文
- [11] 吳柏軒(2015 年 10 月 12 日)。手語即時翻譯手套 動手也能說 Hello。自由時報，取自 <http://news.ltn.com.tw/news/life/paper/922837>