

# 資工系實務專題研究計畫成果報告

## 數字語音辨識

專題編號：098CSIE-S001

執行期限：97 年 1 學期至 98 年 1 學期

指導教授：杭學鳴老師

專題計劃參與人員：許文瀚 許湘婷

### 一、中文摘要

本專題目的是辨識單音的數字語音。我們以前人的語音辨識經驗與演算法為起點，採用許多語音技巧，進行辨識比對，此實務專題將語音辨識的理論與實作兩者合一，學習到研讀資料、擷取數據、實作程式、分析結果等專題研究的技巧與經驗。專題主要分四大部分：第一部份使用 Gold wave 錄音程式錄製待測語音，第二部份將待測語音輸入到 Matlab 後，進行語音的前置處理，第三部份是計算語音辨識特定的一些演算公式，例如：自相關係數、線性預估、倒頻譜、轉移倒頻譜，第四部份是訓練語音模型，利用平均切割與不平均切割來求出最小變異數，加總當作特徵值，完成辨識的流程。

**關鍵詞：**自相關係數、線性預估、倒頻譜、轉移倒頻譜、語音辨識

### 二、緣由與目的

科技的日新月異，以機器化代替人力的技術，確實帶給世界上的人更便利且更有效率的生活型態。在語音辨識這方面技術與應用已經趨於成熟的階段，像是運用在手機上，利用人說話輸入電話號碼以及姓名。或是運用在電腦的輸入，可以不用鍵盤，透過麥克風，說的話即可立即顯示在 word 編輯文件裡。由於好奇心理，開啟了我對於語音辨識領域的初步了解。

語音的技術有很多種，我們主要做語音辨識(speech recognition)，讓機器聽懂人說的話，即為自動語音辨識。我們專題的主要目的，在於學習現有的語音辨識理論與技巧。經由實作經驗，結合所學到的知識，一步一步學以致用。

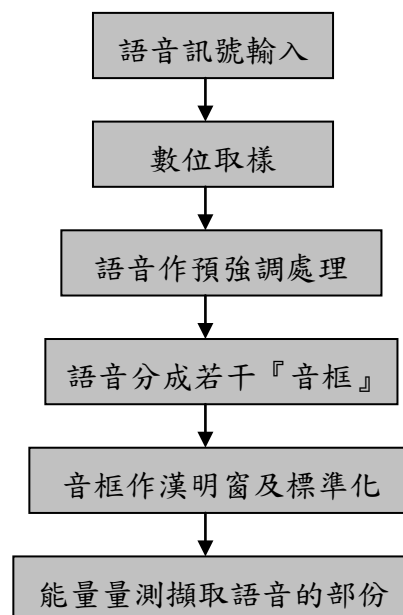
### 三、研究報告應含的內容

#### 一、開發環境與工具：

Goldwave 軟體、MATLAB、麥克風

#### 二、系統架構：

##### (1)前置處理

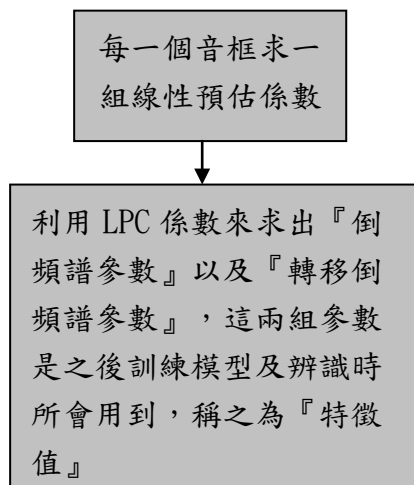


(圖一)為前置處理動作

當語音從 Goldwave 輸入到 MATLAB 後，將輸入的類比信號數位化，取樣頻率取 8KHz。為了補償衰減掉的高頻部份，先將語音作預強調處理。接著，執行“短時

段”的處理方法，將音框切割成 240 點段落，乘上漢明窗，可聚集中間部分的特徵並壓抑兩端信號對語音參數的影響。透過標準化將原始語音信號作等比例的放大或縮小，避免端點偵測時導致比對上的錯誤，再利用能量量測方法切割語音存在的起點與終點。(圖一)

## (2) 辨識演算法



(圖二)為辨識演算法流程

得到具高能量語音存在的範圍後，對每個音框先作自相關係數。通常線性模型的階層 $p$ 為12。亦即對本身的語音位移從0至 $p$ 位置後，與本身的語音作相乘。相乘後值越小，表示語音彼此之間相關性小；反之，彼此語音之間的相關性大。計算完音框的自相關係數之後，利用一種求出反矩陣的杜賓演算法(Durbin algorithm)，得到一組線性預估係數。將線性預估模式依演算法方式，進行倒頻譜及轉移倒頻譜之計算。之後，可得到訓練模型所用到的特徵值。倒頻譜主要的目的，在於表示語音訊號中頻譜的波峰及細部變化的特徵，進而掌握語音頻譜特性。語音大多會受到外界雜訊的影響，因此將倒頻譜參數作偏微分後，會得到轉移倒頻譜參數的特徵值。轉移倒頻譜具有使特徵值具有抗雜訊的能力。(圖二)

經由演算法處理所得到之特徵值，使用不平衡切割方法，先訓練推求語音模

型。利用語音模型，計算待測語音和個別語音模型的特徵值所得到的差值，協助辨識比對工作。以下為最後實驗結果的一例。

六個狀態的資料庫比對數字為=>(1 2 5)  
 相似度的值最小為=>4.199921 比較偏向數字=>1  
 九個狀態的資料庫比對數字為=>(0 3 6 7 8 9)  
 相似度的值最小為=>8.548793 比較偏向數字=>0  
 三個狀態的資料庫比對數字為=>(4) 不採用  
 比對結果得到:  
 待測者與資料庫比對得最接近數字=>1  
 (表一)待測語音與資料庫比對結果圖

## 四、結論

經由實作，檢驗辨識的結果，確實可以達成語音數字辨識的能力。為獲得比較好的效果，在某些情況下，對於數字音框多少的改變及依據差值條件的判斷式，以及利用實驗當中重複執行參數值的設定。經由本實務專題，我學習到語音各方面技術，跟剛踏入語音專題前的知識背景了解相比，受益良多。也符合實作與理論結合之應用目的。

## 五、參考文獻

- [1]聲音辨識原理  
<http://andrew.csie.ncyu.edu.tw/DOC3/>
- [2]人類發音器官  
<http://www.uijin.idv.tw/new/phonology.handout/>
- [3]立德管理學院 語音 辨識專題  
<http://74.6.146.127/search>
- [4]楊鎮光，” Visual Basic 與語音辨識-讓電腦聽話”，文魁資訊, 2002
- [5]王小川，”語音訊號處理”，全華, 民93

