

# 語音會議記錄

靜宜大學資訊工程學系

專題成果報告書

指導老師:羅峻旗

專題學生:資工四 B 410817487 林育昇

資工四 B 410803292 余兆堤

資工四 B 410817649 李承翰

實驗室名稱:主顧樓 517 實驗室

## 目錄

●前言.....	3
●系統功能.....	3
●系統特色.....	3
●使用說明.....	4
●系統介面.....	9
●使用對象.....	9
●開發工具.....	10
●成本分析.....	11
●結論與未來發展.....	11

## 前言

會議進行時，參與者之間溝通頻繁，記錄員常常跟不上會議內容，記錄員當下也會來不及反應內容是哪位參與者所講的，在會議後用錄音，檔案的後續處理很耗費時間，我們想要做出一個可以解決以上問題的系統。

## 系統功能

透過使用者的聲音進行特徵擷取，再經由訓練得到多位使用者各自聲音所擁有的特徵，可由訓練完的特徵辨識是哪位使用者，最後再利用 GOOGLE 語音轉文字的功能，將使用者當下所說的轉為文字檔。



## 系統特色

1. 可以及時將語音轉為文字顯示出來
2. 可由訓練後辨識出正在說話的使用者

# 使用說明

## 1. 開樣本檔



## 2. 錄製訓練樣本



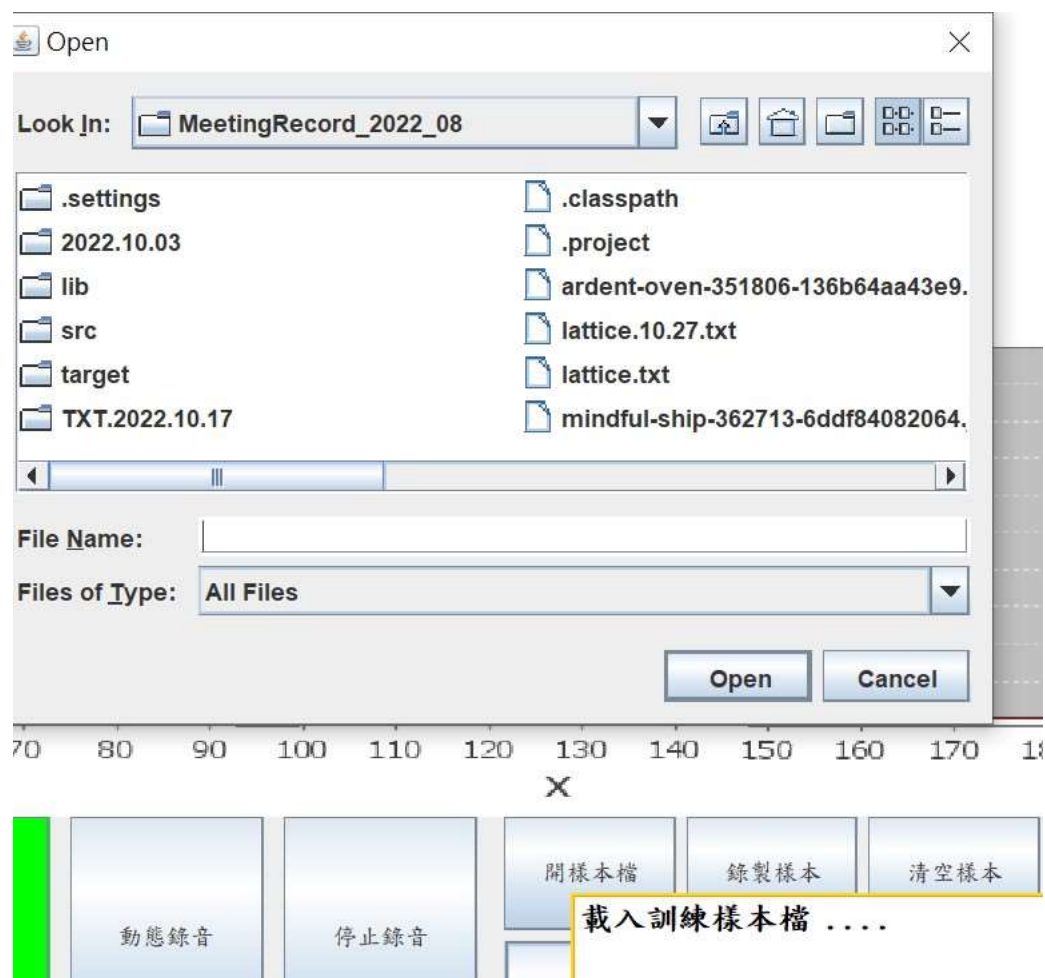
## 3. 停止錄製



#### 4. 儲存樣本



#### 5. 載入訓練樣本檔



## 6. 特徵擷取

```
"name": "林育昇", "chain": { "name": "林育昇", "count": 0, "nodes":  
  { "common": false, "cnt": 38, "prob": 0.003143093465674111, "peaks": [90, 114, 120, 107, 94, 99]},  
  { "common": false, "cnt": 36, "prob": 0.002977667493796526, "peaks": [120, 91, 114, 97, 103, 109]},  
  { "common": true, "cnt": 169, "prob": 0.013978494623655914, "peaks": [89, 10, 14, 109, 96, 101]},  
  { "common": false, "cnt": 83, "prob": 0.006865177832919769, "peaks": [59, 91, 82, 5, 13, 86]},  
  { "common": true, "cnt": 212, "prob": 0.01753515301902399, "peaks": [97, 5, 51, 9, 14, 87]},  
  { "common": false, "cnt": 77, "prob": 0.0063688999172870145, "peaks": [4, 113, 9, 58, 92, 88]},  
  { "common": false, "cnt": 43, "prob": 0.003556658395368073, "peaks": [13, 9, 4, 58, 53, 88]},  
  { "common": true, "cnt": 61, "prob": 0.005045492142266336, "peaks": [87, 26, 91, 22, 18, 56]},  
  { "common": false, "cnt": 169, "prob": 0.013978494623655914, "peaks": [47, 22, 17, 51, 26, 56]},  
  { "common": true, "cnt": 81, "prob": 0.006699751861042184, "peaks": [121, 98, 103, 109, 115, 6]},  
  { "common": true, "cnt": 72, "prob": 0.005955334987593052, "peaks": [87, 5, 112, 108, 103, 10]},  
  { "common": false, "cnt": 66, "prob": 0.005459057071960298, "peaks": [61, 87, 116, 107, 10, 112]},
```

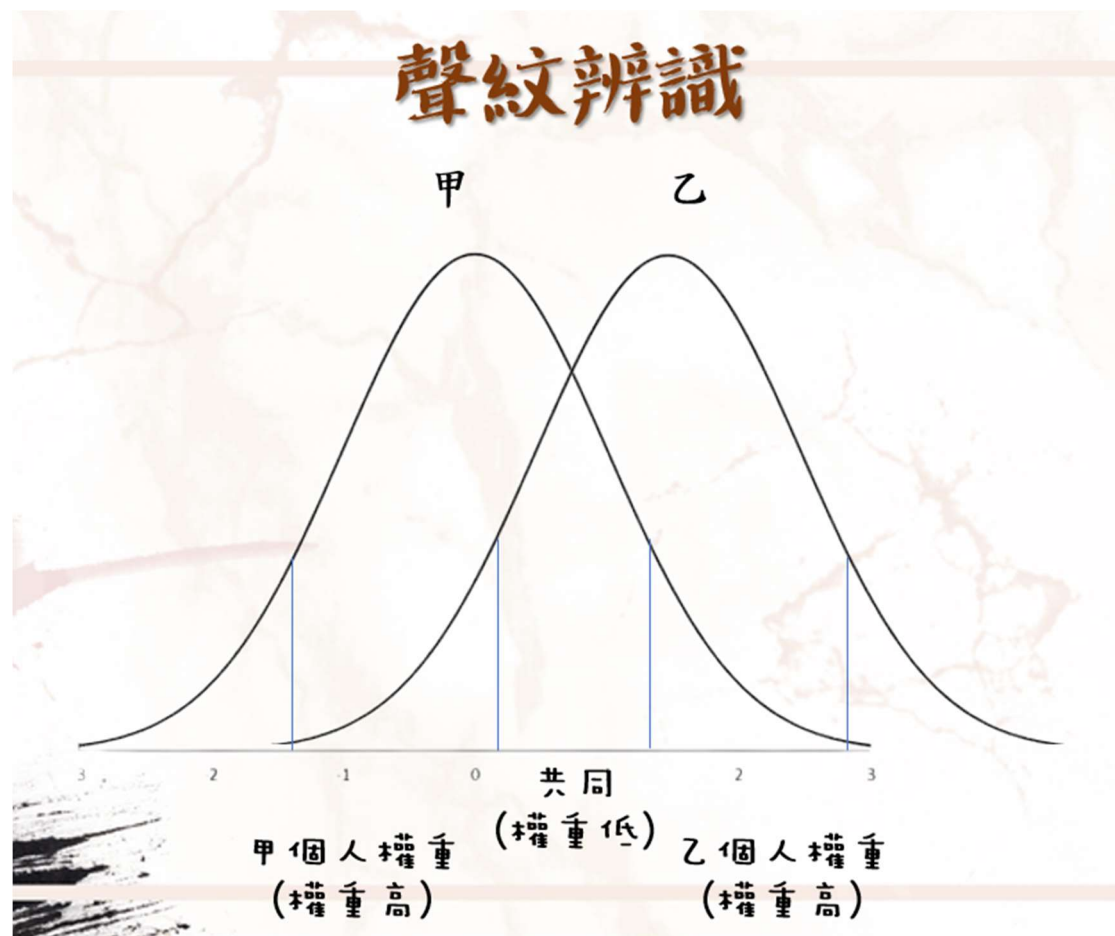
我們取出每秒聲音的六個波峰來進行特徵擷取，common  
true 就是這段特徵跟別人有重疊，cnt 就是此特徵出現的  
次數，prob 是我們根據出現的次數算的機率，最後的  
peaks 是這段聲音採用的六個波峰值。

### 6-1. 進行訓練



```
{"name": "Lattice", "cats": [{"name": "  
--- Lattice loaded-----  
Lattice: Lattice  
Category: 余兆堤|  
Chain Size: 143  
Category: 李承翰  
Chain Size: 134  
Category: 林育昇  
Chain Size: 118
```

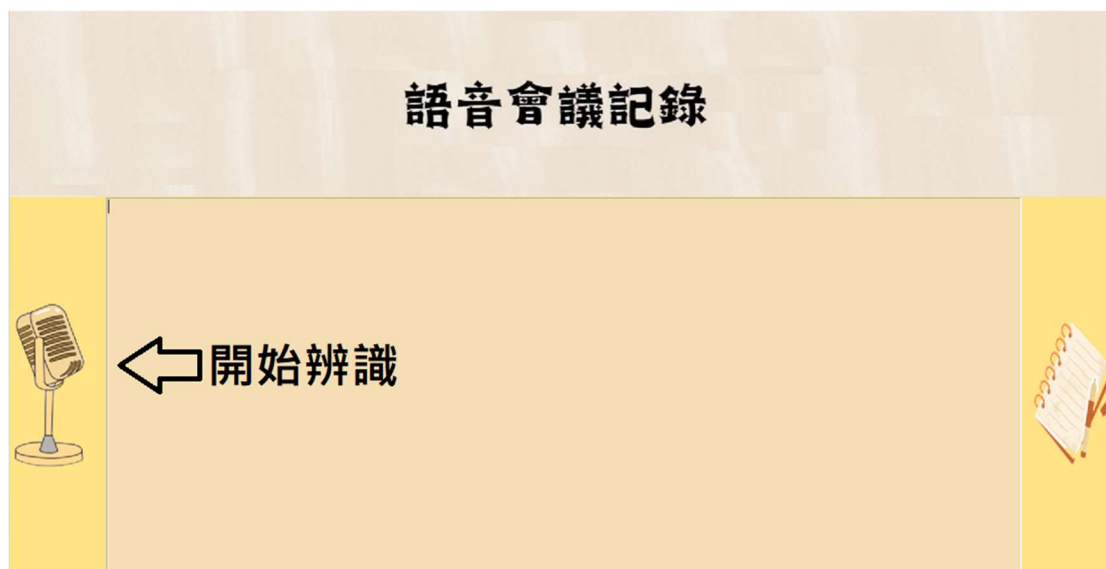
## 6-2. 聲紋辨別



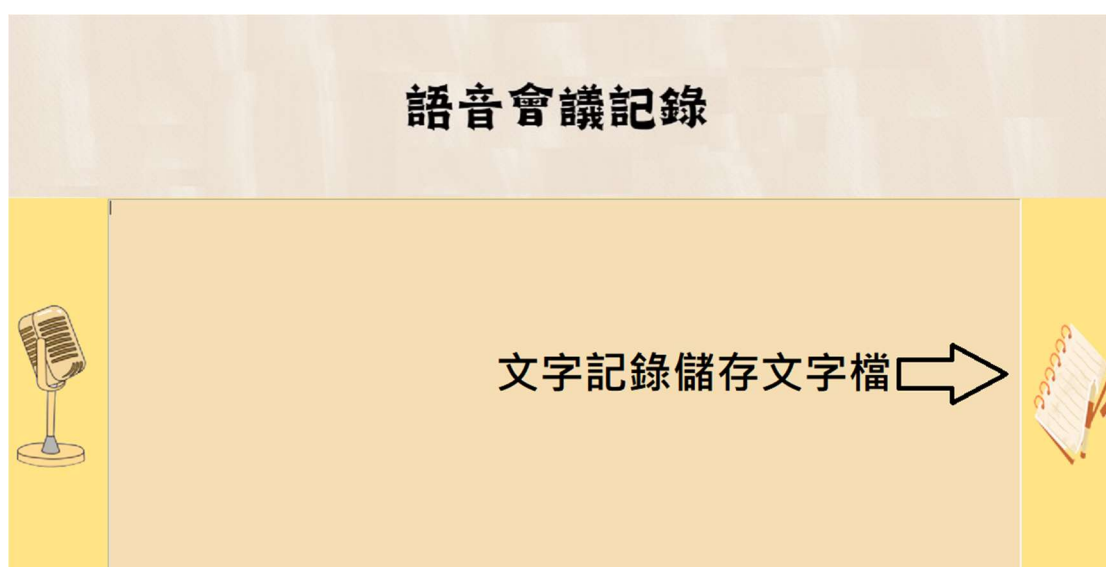
擷取完的特徵會形成一個類似常態分布的圖，我們先把常態分佈機率比較低的部份去掉。

當有多個人的常態分佈出現之後，特徵若有重疊表示是共有的特徵，就會將重疊部分的權重降低。

## 7. 開始辨識

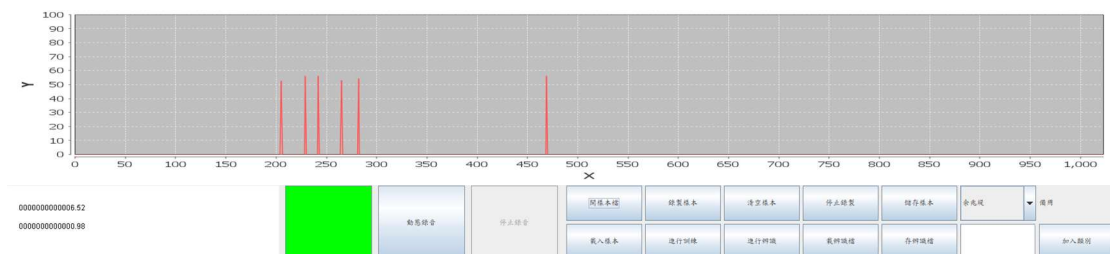


## 8. 儲存成文字檔





## 系統介面

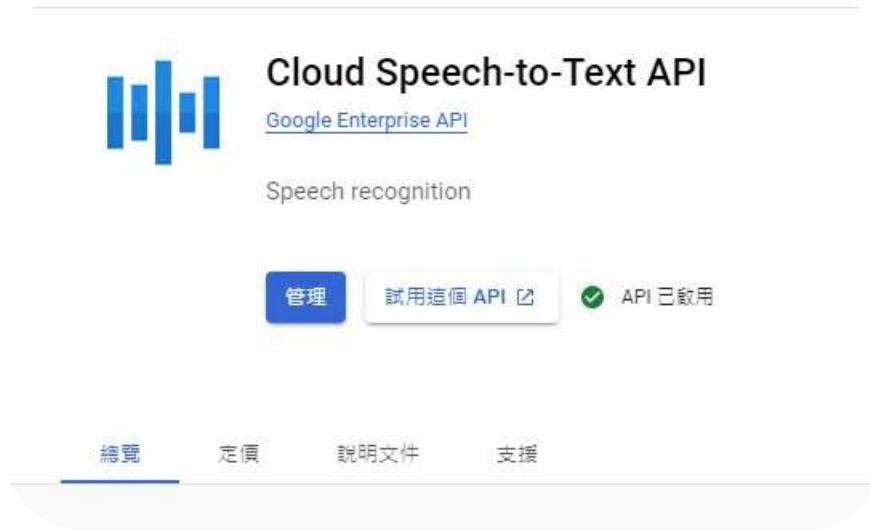


## 使用對象

1. 會議記錄員
2. 語音辨識工具

## 開發工具

### 語音轉文字



### 開發環境



### 開發語言



## 成本分析

項目名稱	說明	單位	數量	單價	小計	備註
電腦	專案之 進行	部	1	30000	30000	實驗室 提供
麥克風	專案之 進行	部	1	2000	2000	實驗室 提供
雜支費	印刷、 文具	批	1		500	自行負 擔
共計					32500	

## 結論與未來發展

### 結論

雖然在吵雜的環境之下較難執行，語音轉文字的輸出時間也可能會受到網路影響，但利用特徵值的方式去辨識，目前辨識結果可以接受。

### 未來發展

透過這套系統不只可以應用在會議記錄也可以應用在辯論會上，因為發現市長辯論會的模式是固定發言時間，輪

流發言卻沒有及時轉錄的字幕。甚至可以應用在現今智慧家居聲音安全性辨識上面。

現今攜帶式設備普及，如果未來能使用在手機平台上，就可以更便利的應用。