

目錄

第一章 緒論.....	1
1.1 研究動機.....	1
1.2 相關研究簡介.....	1
1.3 改進方法簡介.....	2
1.4 系統簡介.....	2
1.5 系統應用.....	4
1.6 章節概要.....	5
第二章 基礎理論與技術.....	6
2.1 語音辨識概論.....	6
2.1.1 動態時間扭曲法 (DTW, Dynamic Time Warping).....	6
2.2.2 隱藏式馬可夫模型 (HMM, Hidden Markov Model).....	6
2.2 語音特徵參數擷取.....	7
2.2.1 音框化 (Frame Blocking).....	9
2.2.2 計算能量 (Energy).....	9
2.2.3 預強調 (Pre-Emphasis).....	9
2.2.4 漢明窗 (Hamming Window).....	9
2.2.5 快速傅立葉轉換 (FFT, Fast Fourier Transform).....	10
2.2.6 三角帶通濾波器 (Triangular band-pass filter).....	11
2.2.7 離散餘弦轉換 (DCT, Discrete Cosine Transformatiion).....	11
2.2.8 差量倒頻譜參數 (Delta Cepstrum Coefficients).....	11
2.3 聲學模型訓練.....	12
2.4 語音訊號與聲學模型之比對.....	13
第三章 改進方法.....	14

3.1 計算能量 (Energy)	15
3.2 預強調 (Pre-Emphasis).....	17
3.3 漢明窗 (Hamming Window)	17
3.4 快速傅立葉轉換 (FFT).....	19
3.5 三角帶通濾波器.....	22
3.6 離散餘弦轉換.....	24
3.7 計算能量對數.....	25
3.8 開平方根表格建立.....	27
第四章 實驗數據與分析.....	28
4.1 精準度.....	28
4.1.1 Cos and Sin Table 之誤差比較：	28
4.1.2 Logarithm and Square Root Table 之誤差比較	29
4.1.3 MFCC 各個步驟誤差的比較	29
4.1.4 辨識率比較.....	33
4.1.5 執行時間比較.....	34
4.2 錯誤分析.....	35
第五章 未來研究方向.....	36
參考文獻.....	37

圖表目錄

圖 1-1 語料與參數轉換關係圖	3
圖 1-2 不同語料之轉換參數設定抽取特徵及 HTK 訓練語料圖	3
圖 1-3 根據不同語料辨識結果圖	4
圖 2-1 語音頻率與人類感知頻率關係圖	8
圖 2-2 MFCC 語音特徵擷取流程圖	8
圖 2-3 原始訊號與乘上 HAMMING WINDOW 過後的訊號比較	10
圖 2-4 狀態轉移示意圖	12
圖 2-5 VITERBI 計算流程圖	13
圖 3-1 TCC300 語料於計算能量與調整參數值資料溢位的機率	16
圖 3-2 TIMIT 語料於計算能量與調整參數值資料溢位的機率	16
圖 3-3 不同倍率放大之整數漢明窗值的分佈圖	18
圖 3-4 放大倍率 2^1 與 2^{14} 的圖形	18
圖 3-5 DCT ALGORITHM IN MFCC PROCEDURE	19
圖 3-6 WEIGHT CEPSTRUM ALGORITHM IN MFCC PROCEDURE	20
圖 3-7 SUB-ROUTINE OF FFT AND FFT FUNCTION PARTS OF ALGORITHM IN MFCC PROCEDURE	20
圖 3-8 SIN TABLE 結構圖	22
圖 3-9 TIMIT 語料中 $ X_a[k] ^2$ 與不同倍率 FILTER TABLE 產生溢位之機率圖 ..	23

圖 3-10 TCC300 語料中 $ X_a[k] ^2$ 與不同倍率 FILTER TABLE 產生溢位之機率圖	24
圖 3-11 $\ln(x)$ X 軸與 Y 軸的相對應關係	26
圖 3-12 整數 $\ln(x)$ TABLE X 軸與 Y 軸的相對應關係	26
圖 3-13 整數 $Sqrt(x)$ TABLE X 軸與 Y 軸的相對應關係	27
圖 4-1 AVERAGE RELATIVE ERROR COMPARISON OF TCC300 CORPUS	32
圖 4-2 AVERAGE RELATIVE ERROR COMPARISON OF TIMIT CORPUS	33
圖 4-3 [2][3]與我們的方法所需時間比較	35



表格目錄

表格 3-1 唐詩語料資訊	14
表格 3-2 TCC300 及 TIMIT 的語料資訊	14
表格 3-3 INPUT VALUE FOR CALCULATING SIN IN FFT SUB-ROUTINE.....	21
表格 3-4 INPUT VALUE FOR CALCULATING SIN IN FFT FUNCTION	21
表格 3-5 [2][3]與我們的建表方式比較.....	27
表格 4-1 COS TABLE 的誤差	28
表格 4-2 SIN TABLE 的誤差	29
表格 4-3 COS 與 SIN TABLE 佔記憶體大小比較	29
表格 4-4 LOGARITHM TABLE 的誤差.....	29
表格 4-5 SQUARE ROOT TABLE 的誤差.....	29
表格 4-6 [2][3]的方法針對 TCC300 語料所產生的誤差	30
表格 4-7 我們的方法針對 TCC300 語料所產生的誤差	30
表格 4-8 [2][3]的方法針對 TIMIT 語料所產生的誤差.....	31
表格 4-9 我們的方法針對 TIMIT 語料所產生的誤差	31
表格 4-10 TANG POEM 資訊	33
表格 4-11 TCC300 辨識率比較.....	34
表格 4-12 TIMIT 辨識率比較	34
表格 4-13 [2][3]及我們的方法所需計算時間的比較.....	34