

# 利用DSP晶片實作語音控制之研究

蔡學承、李立民

E-mail: 386773@mail.dyu.edu.tw

## 摘要

本研究中，我們建立一個語音命令辨識系統，將此系統實現在德州儀器數位訊號處理器 TMS320C6713 DSK Board上，並實作機器人語音控制之實驗。我們取12維向量的線性預估倒頻譜參數及語音能量強度作為語音特徵參數，並取上述特徵之一次與二次微分。語音辨識系統中，使用隱藏式馬可夫模型來做模型訓練。系統接收來自麥克風的訊號，經過DSP板做語音辨認之後將命令信號傳送到一個紅外線遙控器，發送訊號以控制機器人之動作。語音辨認系統採用Matlab與Simulink做高階系統設計，利用Embedded Coder將Matlab程式碼轉成C語言，再利用DSP晶片之編譯工具轉成DSP板的執行檔後下載到DSP板執行。系統包含端點偵測、特徵參數計算、資料緩衝與儲存及語音辨識核心程式。實際用五個命令操作時，在比較沒有吵雜的環境中，語音辨識率可達到將近100%，在比較吵雜的環境中也有80%以上的語音辨識率。

關鍵詞：數位訊號處理器、語音辨識、隱藏式馬可夫模型、聲控機器人

## 目錄

封面內頁 簽名頁 中文摘要 . . . . .	iii	英文摘要 . . . . .	iii
. . . . . iv 誌謝 . . . . .	iv	目錄 . . . . .	v
. . . . . vi 圖目錄 . . . . .	vi	表目錄 . . . . .	viii
. . . . . x 第一章 緒論 . . . . .	1	1.1 研究緒論 . . . . .	1
. . . . . 1.1.1 研究緒論 . . . . .	1	1.1.2 研究背景 . . . . .	2
. . . . . 1.1.2 研究背景 . . . . .	2	1.2 研究目的 . . . . .	3
第二章 語音特徵參數擷取 . . . . .	4	2.1 語音特徵參數擷取 . . . . .	4
2.1 語音特徵參數擷取 . . . . .	4	2.2 梅爾頻率倒頻譜參數 . . . . .	14
2.2 梅爾頻率倒頻譜參數 . . . . .	14	第三章 隱藏式馬可夫模型 . . . . .	17
第三章 隱藏式馬可夫模型 . . . . .	17	3.1 以距離及動態規劃為觀點的語音辨識 . . . . .	17
3.1 以距離及動態規劃為觀點的語音辨識 . . . . .	17	3.2 隱藏式馬可夫模型的組成 . . . . .	20
3.2 隱藏式馬可夫模型的組成 . . . . .	20	3.3 使用Viterbi方法做語音模型訓練與語音辨認 . . . . .	22
3.3 使用Viterbi方法做語音模型訓練與語音辨認 . . . . .	22	3.4 使用EM演算法做模型訓練 . . . . .	24
3.4 使用EM演算法做模型訓練 . . . . .	24	3.4.1 正算程序 . . . . .	25
3.4.1 正算程序 . . . . .	25	3.4.2 逆算程序 . . . . .	27
3.4.2 逆算程序 . . . . .	27	3.4.3 模型參數重估 . . . . .	28
3.4.3 模型參數重估 . . . . .	28	第四章 實驗過程 . . . . .	32
第四章 實驗過程 . . . . .	32	4.1 語音資料庫格式 . . . . .	32
4.1 語音資料庫格式 . . . . .	32	4.2 狀態數與維度的差異 . . . . .	33
4.2 狀態數與維度的差異 . . . . .	33	4.2.1 維度比較 . . . . .	35
4.2.1 維度比較 . . . . .	35	4.3 實際測試 . . . . .	40
4.3 實際測試 . . . . .	40	4.4 紅外線遙控器原理 . . . . .	43
4.4 紅外線遙控器原理 . . . . .	43	第五章 結論 . . . . .	46
第五章 結論 . . . . .	46	5.1 結論 . . . . .	46
5.1 結論 . . . . .	46	5.2 未來研究方向 . . . . .	47
5.2 未來研究方向 . . . . .	47	參考文獻 . . . . .	48
參考文獻 . . . . .	48		

## 參考文獻

- [1] J. Makhoul, "Linear prediction: A tutorial review," Proc. of IEEE, vol. 63, no. 4, pp. 561-580, Apr. 1975.
- [2] 王小川, "語音訊號處理", 全華圖書, 台北縣土城市, 2008.
- [3] 維基百科, [http://en.wikipedia.org/wiki/Hidden\\_Markov\\_model](http://en.wikipedia.org/wiki/Hidden_Markov_model).
- [4] K.-F. Lee, "Automatic Speech Recognition: The Development of the SPHINX System," Kluwer Academic Publishers, Boston, 1989.
- [5] B. Bogert, M. Healy and J. Tukey, "The quefrency analysis of time series for echoes," Proc. Symp. On Time series Analysis, New York, J. Wiley, 1963
- [6] S. Davis, P. Mermelstein, "Comparing of Parametric Representations for Monosyllable Word Recognition in Continuously Spoken Sentence," IEEE Trans. on Acoustic, Speech and Signal Processing, 1980.
- [7] 蔡博宇, "以FPGA為平台之嵌入式類神經網路語音辨識系統實現", 碩士論文, 電機工程學系, 高雄大學, 2009
- [8] 李霄瑜, "結合基因演算法及經驗模態分解進行強健性語音辨識與FPGA晶片實現", 碩士論文, 資訊工程學系, 高雄大學, 2009
- [9] 楊俊哲, "隱藏式馬可夫模型之語音辨識在電視控制系統之應用", 碩士論文, 工程科學及海洋工程學研究所, 台灣大學, 2012
- [10] 蔡宜亨, "具有強健性語音辨識的無線語音控制系統研製", 碩士論文, 資訊工程學系, 高雄大學, 2011
- [11] TEXAS INSTRUMENTS Technical Reference, [http://c6000.spectrumdigital.com/dsk6713/V2/docs/dsk6713\\_TechRef.pdf](http://c6000.spectrumdigital.com/dsk6713/V2/docs/dsk6713_TechRef.pdf)
- [12] HOLTAK data sheet, [http://www.holtek.com.tw/pdf/consumer/6221\\_2v170.pdf](http://www.holtek.com.tw/pdf/consumer/6221_2v170.pdf)