

# Incorporating Bi-Word Units to Improve Continuous Mandarin Digit Recognition System

張育豪、李立民

E-mail: 9315086@mail.dyu.edu.tw

## ABSTRACT

In this thesis, we analyze the errors in the continuous Mandarin digit recognition system and try to reduce the major errors. Initial experimental results show that the deletion errors are much more serious than the substitution errors and the insertion errors. Experiments also show that increasing the number of Gaussian mixture of the HMM state could not effectively reduce the number of deletion errors. Error statistics show that the deletion of digits "1" and "5" are the major causes of errors and these deletion are most likely to happen in sentences containing bi-word strings "11", "71", and "55". To reduce deletion errors of these sub-string, we incorporate these bi-word units into the original word loop network of single digit. Experimental results show that this method can successfully reduce the deletion errors of digits "5" and "1". An unwanted side effect of this method is the increase of insertion errors of digits "5" and "1". Nevertheless, when the TOP-N error is considered, the number of increased insertion errors is smaller than the number of reduced deletion error, and the overall system performance is improved.

Keywords : bi-word units ; continuous Mandarin digit recognition

## Table of Contents

目錄 封面內頁 簽名頁 授權書 . . . . .	iii	中文摘要 . . . . .	
. . . . . iv 英文摘要 . . . . .	iv	v 誌謝 . . . . .	
. . . . . vi 目錄 . . . . .	vi	vii 圖目錄 . . . . .	
. . . . . ix 表目錄 . . . . .	ix		
. xi 第一章 前言 . . . . .		1 1.1 研究主題與動機 . . . . .	
. 1 1.2 研究背景 . . . . .		2 1.3 研究方向 . . . . .	
節概要 . . . . .	4	2 1.4 章節概要 . . . . .	
音辨認簡介 . . . . .	4	2 2.1 語音特徵參數擷取 . . . . .	
參考模型的建構 . . . . .	5	2 2.2 擷取特徵參數 . . . . .	
. . . . . 12	3.1	國語數字辨認單元 . . . . .	
. . . . . 13	3.2	隱藏式馬可夫模型的建立 . . . . .	
. . . . . 20	3.3	15 3.3.1 正算程序 . . . . .	
. . . . . 22	3.3.2	逆算程序 . . . . .	
. . . . . 22	3.5	模型比對-維特比演算法 . . . . .	
. . . . . 26	4.1	語音資料庫簡介 . . . . .	
. . . . . 26	4.2	基本系統的建立 . . . . .	
. . . . . 26	4.3	辨認效果評估 . . . . .	
. . . . . 30	4.4	基本系統的實驗結果 . . . . .	
. . . . . 35	4.5	錯誤分析 . . . . .	
. . . . . 38	5.1	雙字模型的建立 . . . . .	
. . . . . 38	5.2	修正型系統的實驗結果 . . . . .	
. . . . . 42	5.3	錯誤分析 . . . . .	
. . . . . 43	第六章	結論及未來研究方向 . . . . .	
. . . . . 43	圖目錄	圖2.1 語音辨認系統基本流程圖 . . . . .	
. . . . . 5	圖2.2	數字"9"語音波形 . . . . .	
. . . . . 6	圖2.3	數字"9"聲譜圖 . . . . .	
. . . . . 7	圖2.4	特徵向量擷取流程圖 . . . . .	
. . . . . 7	圖2.5	音框示意圖 . . . . .	
. . . . . 9	圖2.6	三角帶通濾波器組之頻率響應 . . . . .	
. . . . . 11	圖3.1	基本決策樹 . . . . .	
. . . . . 14	圖3.2	隱藏式馬可夫模型狀態轉換圖 . . . . .	
. . . . . 15	圖3.3	高斯混合模型示意圖 . . . . .	
. . . . . 18	圖3.4	正算程序示意圖 . . . . .	
. . . . . 21	圖3.5	逆算程序示意圖 . . . . .	
. . . . . 22	圖3.6	示意圖 . . . . .	
. . . . . 23	圖4.1	靜音模型與短停頓模型之關係 . . . . .	
. . . . . 28	圖4.2	錯誤類型示意圖 . . . . .	
. . . . . 29	圖4.3	基本系統句子辨認率 . . . . .	
. . . . . 30	圖4.4	基本系統單字辨認率 . . . . .	
. . . . . 30	圖5.1	連續數字"55"波形 . . . . .	
. . . . . 36	圖5.2	連續數字"11"波形 . . . . .	
. . . . . 36	圖5.3	連續數字"171"波形 . . . . .	
. . . . . 37	圖5.4	連續數字"171"聲譜圖 . . . . .	
. . . . . 37	圖5.5	修正型系統句子辨認率 . . . . .	
. . . . . 38	圖5.6	修正型系統單字辨認率 . . . . .	
. . . . . 38	圖5.7	基本系統與修正型系統辨認率比較(混合數=16) . . . . .	
. . . . . 39	表目錄	表2.1 本文所使用之特徵向量擷取規格表 . . . . .	
. . . . . 8	表3.1	國語數字聲母、韻母 . . . . .	

.....	12	表3.2 以前後音影響做模型的細分	.....	13	表4.1 國語數字資料庫錄音格式
.....	26	表4.2 單音模型組成	.....	27	表4.3 聲韻母模型狀態分配表
.....	27	表4.4 刪除型錯誤個數統計表	.....	31	表4.5 插入型錯誤個數統計表
.....	31	表4.6 替代型錯誤個數統計表	.....	32	表4.7 基本系統數字"5"及數字"1"刪除型錯誤個數
.....	33	表4.8 基本系統數字"5"及數字"1"插入型錯誤個數	.....	34	表5.1 刪除數字"5"的錯誤統計
.....	35	表5.2 刪除數字"1"的錯誤統計	.....	35	表5.3 雙字模型字典
.....	38	表5.4 修正型系統數字"5"及數字"1"刪除型錯誤個數	.....	40	表5.5 修正型系統數字"5"及數字"1"插入型錯誤個數
.....	41		.....	41	

## REFERENCES

- [1] 陳信宏, 謝寶華, "使用前後文相關HMM模型之國語連續語音辨認", 碩士論文, 國立交通大學電信工程系, 新竹, 民國90年6月。
- [2] 陳信宏, 呂儲仰, "國語連續音節辨認系統之改進與分析", 碩士論文, 國立交通大學電信工程系, 新竹, 民國91年6月。
- [3] Zang Bo; Juan, L.; Gang, P.; Wang, W.S.-Y. "A High Performance Mandarin Digit Recognizer ", Signal Processing and Its Applications, 1999. ISSPA '99. Proceedings of the Fifth International Symposium on, Volume: 2, 22-25 Aug. 1999, pp629 - 632 vol.2.
- [4] Yuan-yuan Shi; Jia Liu; Run-sheng Liu; "Mandarin Digit String Speech Recognition using Linear Discriminant Analysis and Tone Discrimination Model", TENCON '02. Proceedings. 2002 IEEE Region 10 Conference on Computers, Communications, Control and Power Engineering, Volume: 1, 28-31 Oct. 2002, pp461 - 464 vol.1.
- [5] 林進燈, 黃進民, "以修改後的廣義共同向量為基礎之不特定語者中文數字語音辨識", 碩士論文, 國立交通大學電資學院電機與控制學程碩士班, 新竹, 民國92年6月。
- [6] 陳遵立, 陳松琳, "以類神經網路為架構之語音辨識系統", 碩士論文, 國立中山大學電機工程學系研究所, 高雄, 民國91年6月。
- [7] 謝秀琴, "數位語音訊號基本原理", 全華科技, 台北市, 1996。
- [8] 謝依蘭, "語音訊號數位處理", 松崗, 台北市, 1992。
- [9] 陳明熒, "PC電腦語音辨認實作", 旗標, 台北市, 1994。
- [10] K. Beulen, H. Ney, "Automatic Question Generation for Decision Tree Based State Tying". In Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing, 1998, pp805-808 Vol.2.
- [11] Thomas F. Quatieri, "Discrete-Time speech signal processing principles and practice" Prentice-Hall PTR, 2000.
- [12] X.D. Xuang, Y. Ariki, M.A. Jack, "Hidden Markov Models for Speech Recognition". Edinburgh University Press, 1990, pp187-205, Chap7.
- [13] Steve Young, The HTK Book Version 3.0, 2000.