

# 語音驗證系統之研究

黃威勳、李立民

E-mail: 9707896@mail.dyu.edu.tw

## 摘要

語者驗證主要目的為依據語音訊號所包含的相關訊息來決定語者的身份，但是對電腦而言，要掌握這些資訊的差異需要經過很多步驟。本文使用梅爾倒頻譜參數做為語音特徵參數，以符合人類發音與聽覺的特徵。高斯混合模型在文本不固定(Text-Independent)語者驗證的領域裡受到廣泛的研究，然而說話人的語音特徵，除了因為口腔形狀與聲帶的不同會造成聲音的差異之外，說話的速度也可以作為語者驗證的依據。由於高斯混合模型無法考慮說話速度的差異，本文以高階統計式高斯模型作文本不固定的語者驗證系統，將兩種模型依照相同的條件進行實驗，結果顯示高階統計式高斯模型確實可以改善高斯混合模型的缺失，相等錯誤率(Equal Error Rate)降低3.8個百分比。

關鍵詞：高斯混合模型；高階統計式高斯模型；語者驗證

## 目錄

目錄 封面內頁 簽名頁 授權書 . . . . .	iii
中文摘要 . . . . .	iii
iv 英文摘要 . . . . .	iv
v 誌謝 . . . . .	v
vi 目錄 . . . . .	vi
vii 圖目錄 . . . . .	vii
ix 表目錄 . . . . .	ix
. xi 第一章 緒論 . . . . .	xi
1.1 研究動機 . . . . .	1.1.1
1.2 語者驗證系統概論 . . . . .	1.2
2 研究背景 . . . . .	2
3 1.4 研究目的 . . . . .	3.1.4
4 1.5 章節概要 . . . . .	4.1.5
5 第二章 語音特徵參數擷取 . . . . .	5
6 2.1 語者驗證系統概要 . . . . .	6.2.1
6.2.2 量化誤差 . . . . .	6.2.2
7 2.3 端點偵測 . . . . .	7.2.3
7.2.4 梅爾倒頻譜參數 . . . . .	7.2.4
10 第三章 向量量化 . . . . .	10
18 3.1 簡介 . . . . .	18.3.1
18 3.2 K-means分群法 . . . . .	18.3.2
18 3.2 分裂分群法 . . . . .	18.3.2
21 第四章 建立語者模型 . . . . .	21
23 4.1 適用於語者驗證的模型 . . . . .	23.4.1
23 4.2 高斯混合模型 . . . . .	23.4.2
24 4.3 目標函式最佳化 . . . . .	24.4.3
26 4.4 高階統計式高斯模型 . . . . .	26.4.4
27 第五章 實驗結果 . . . . .	27
33 5.1 語音資料庫簡介 . . . . .	33.5.1
33 5.2 實驗流程設定 . . . . .	33.5.2
34 5.3 系統評估方式 . . . . .	34.5.3
34 5.2 實驗結果 . . . . .	34.5.2
36 第五章 結論及未來研究方向 . . . . .	36
49 6.1 結論 . . . . .	49.6.1
49 6.2 未來研究方向 . . . . .	49.6.2
50 參考文獻 . . . . .	50
. . . . .	51

## 參考文獻

- [1] D. A. Reynolds, " Robust Text-Independent Speaker Identification Using Gaussian Mixture Speaker Models ", IEEE Trans. Speech and Audio Processing, Vol. 3, No. 1, pp. 72-83, 1995.
- [2] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, " Speaker Verification Using Adapted Gaussian Mixture Models ", Digital Signal Processing, Vol. 10, pp. 19-41, 2000.
- [3] S. Davis, and P. Mermelstein, " Comparing of Parametric Representations for Monosyllable Word Recognition in Continuously Spoken Sentence ", IEEE Trans. On Acoustic, Speech and Signal Processing, pp. 357-366, 1980.
- [4] L. R. Rabiner, and M. R. Sambur, " An algorithm for determining the endpoints of isolated utterances ", The Bell System Technique Journal, Vol. 54, pp. 297 - 315, 1975.
- [5] T. K. Moon, " The Expectation-Maximization Algorithm ", IEEE, Signal Processing Magazine, Nov. 1996.
- [6] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki, " The DET curve in assessment of detection task performance ", In Proceedings of the European Conference on Speech Communication and Technology, pp. 1895-1898, 1997.
- [7] S. Young, The HTK book, Version 3.4, Cambridge University Engineering Department, 2002.

[8] 鍾偉仁, “語者辨認與驗證之初步研究”, 國立台灣大學碩士論文, 民國90年6月。

[9] 吳金池, “語者辨識系統之研究”, 國立中央大學碩士論文, 民國90年6月。

[10] 王小川, “語音訊號處理”, 全華科技圖書有限公司, 2005。

[11] 李佳蓀, “以高階隱藏式馬可夫模型作語音辨認之研究”, 大葉大學碩士論文, 民國95年6月。