

A Study of High-Order Hidden Markov Models for Speech Recognition

李佳蓓、李立民

E-mail: 9510173@mail.dyu.edu.tw

ABSTRACT

In order to improve the flaw of traditional first order Hidden Markov Models, we propose high-order hidden Markov models (HO-HMM) to improve speech recognition systems. In the proposed model, both the state transition and output probability depend not only on the current state but also several previous states. Therefore, the HO-HMM can precisely capture the state duration and speech dynamic trajectory. We developed an extended Viterbi algorithm for HO-HMM to train the model and recognize input utterance. We conducted experiments on speaker independent Mandarin digits recognition to investigate the performance of HO-HMM. Experimental results show that the performance of HO-HMM system with both high-order state transition and output observation dependencies is superior to that of first order HMM. We also found that as the order of HO-HMM increases, the error rate of recognition reduced. Experimental results also show that the HO-HMM system is more robust against environmental noise.

Keywords : speech recognition ; high-order hidden Markov models ; dynamic trajectory ; extended Viterbi algorithm

Table of Contents

封面內頁 簽名頁 授權書	iii	中文摘要	
.	iv	英文摘要	
. v 誌謝		vi 目錄	
.	vii	圖目錄	x
目錄	xii	第一章 緒論	
.	1	1.1 研究目的	1
.	1	1.2 研究背景	1
.	2	1.3 研究方法	2
.	3	1.4 章節概要	3
.	3	第二章 語音特徵參數擷取	4
辨識系統概要	4	2.1 語音	4
梅爾頻率倒頻譜參數擷取	5	2.2 語音特徵參數	5
. 12 3.1 隱藏式馬可夫模型	12	2.3 梅爾頻率倒頻譜參數擷取	12
. . 15 3.2.1 初始狀態分佈	15	3 第三章 隱藏式馬可夫模型	12
. 16 3.2.3 觀測符號機率分佈	16	. 12 3.1 隱藏式馬可夫模型	12
. 19 3.4 參數重估	19	. 12 3.2 建立隱藏式馬可夫模型	12
.	23	. 15 3.2.1 初始狀態分佈	15
.	23	. 15 3.2.2 狀態轉移分佈	15
.	25	. 16 3.2.3 觀測符號機率分佈	16
.	30	. 17 3.3.4 Viterbi演算法	17
.	30	21 第四章 高階隱藏式馬可夫模型	21
獨立音實驗結果	33	. 23 4.1 高階隱藏式馬可夫模型	23
雜訊之實驗結果	35	. 23 4.2 擴充型 Viterbi演算法	23
.	37	. 25 4.3 模型訓練	25
.	39	30 5.1 語音資料庫	30
.	42	. 31 5.3 獨立音實驗	31
.	47	. 32 5.3.1 乾淨語音之	32
.	48	. 35 5.3.2 雜訊語音之獨立音實驗	35
.	48	. 35 5.3.2.1 語音摻雜 20dB 訊雜比白	35
.	48	. 37 5.4 混淆音實驗	37
.	48	. 38 5.4.1 未加雜訊之混淆音實	38
.	48	. 40 5.4.2 雜訊語音之混淆音實驗結果	40
.	48	. 40 5.4.2.1 混淆音摻雜 20dB 訊雜比白雜訊之	40
.	48	. 41 5.4.2.2 混淆音摻雜 10dB 訊雜比白雜訊之實驗結果	41
.	48	44 第六章 結論與未來研究方向	44
.	48	. 47 6.1 結論	47
.	48	. 47 6.2 未來研究方向	47
.	48	48 參考文獻	48

REFERENCES

[1] S. Furui, "Speaker independent isolated word recognition using dynamic features of speech spectrum," IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP-34, pp. 52 – 59, 1986.

- [2] J.-F. Mari, J.-P. Haton, and A. Kriouile, "Automatic word recognition based on second-order hidden Markov models." *IEEE Transactions on Speech and Audio Processing*, vol. 5 no. 1, pp. 22 – 25, 1997.
- [3] Y. He, "Extended Viterbi algorithm for second-order hidden Markov process." In *Proceedings of the IEEE 9th International Conference on Pattern Recognition*, pp.718 – 720, 1988.
- [4] J.A. du Preez, "Algorithms for high order hidden Markov modeling," *Proceedings of the IEEE South African Symposium on Communications and Signal Processing*, 9-10 Sept. pp. 101 -106, 1997.
- [5] L. Deng, M. Aksmanovic, D. Sun, and C. F. J. Wu. "Speech recognition using hidden Markov models with polynomial regression functions as nonstationary states," *IEEE Transactions on Speech and Audio Processing*, Vol. 2, No. 4, pp. 507-520, October, 1994.
- [6] 王小川, 語音訊號處理, 全華科技, 台北市, 2005.
- [7] 鄭智寬, "語音特徵抽取方法對連續音辨認影響之研究", 大葉大學碩士論文, 彰化, 民國93年6月。
- [8] 廖子傑, "國語連續數字辨認之研究", 大葉大學碩士論文, 彰化, 民國93年6月。
- [9] T. F. Quatieri, *Discrete-Time speech signal processing principles and practice*, Prentice Hall PTR, 2002.
- [10] S. E. Levinson, L. R. Rabiner, M. M. Sondhi, "An Introduction to the Application of the Theory of Probabilistic Function of a Markov Process to Automatic Speech Recognition," *The Bell System Technical Journal*, Vol.62, No.4, April 1983.
- [11] X.D. Xuang, Y. Ariki, M.A. Jack, *Hidden Markov Models for Speech Recognition*. Edinburgh University Press, pp. 187-205, 1990.
- [12] S. Young, *The HTK Book, Version 3.2*, Cambridge University Engineering Department, 2002.
- [13] S. Davis, P. Mermelstein, "Comparing of Parametric Representation for Monisyllable Word Recognition in Continuously Spoken Sentence," *IEEE Trans. On Acoustic, Speech and Signal Processing*, pp.357-366, 1980.
- [14] Lee L. M. and Lee J. C. "A Study on High-Order Hidden Markov Models and Applications to Speech Recognition," *IEA/AIE 2006, Springer Lecture Notes in Artificial Intelligence*, vol. 4031, pp. 682 – 690, Jun. 2006.