

# 雜訊環境下語音辨認補償方法研究

林耕弘、李立民

E-mail: 9223672@mail.dyu.edu.tw

## 摘要

在安靜環境中所發展出來的語音辨識系統移植到雜訊環境時，其辨認率將會急遽下降，故雜訊效應的補償方法將為語音辨認之重要課題。在本文中，我們對語音之逆頻譜特徵係數在外加雜訊下的行為進行探討。我們發現逆頻譜向量的值會隨著外加雜訊的增加而漸縮小，當環境為白色雜訊時，其變動的方向會與乾淨語音的逆頻譜向量的方向相反，且會縮小到原點，但是在非白色雜訊下，變動的方向則往純雜訊與乾淨語音間的特徵向量差值方向行進，基於這個行為，我們提出特徵偏移向量，將其引入參考模型中，我們利用少量的混雜語音與參考模型中每個狀態之逆頻譜均值向量間的差值，求得特徵偏移向量，但考慮到求出特徵偏移向量所使用的混雜語音，可能與實際使用時的環境並不匹配，所以我們再利用一最佳純量係數，乘上特徵偏移向量作為最佳線性調整量，加至原始的語音參考模型，使模型能夠更適應的在雜訊環境下使用，實驗結果顯示這個方法能夠有效的使用在白色雜訊與其它非白色雜訊環境下。

關鍵詞：語音辨識；環境適應；外加雜訊；特徵偏移向量

## 目錄

封面內頁 簽名頁 授權書 . . . . .	iii	中文摘要 . . . . .	iii
. . . . .	iv	英文摘要 . . . . .	v
. . . . .	vi	誌謝 . . . . .	v
. . . . .	vi	目錄 . . . . .	vii
. . . . .	ix	圖目錄 . . . . .	x
. . . . .	ix	表目錄 . . . . .	xi
. . . . .	1	第一章 緒論 . . . . .	1
. . . . .	1	1.1 研究主題與動機 . . . . .	1
. . . . .	1	1.2 研究背景 . . . . .	1
. . . . .	1	1.3 研究方向 . . . . .	2
. . . . .	3	1.1 章節概要 . . . . .	3
. . . . .	3	第二章 語音特徵參數擷取與辨認系統架構 . . . . .	4
. . . . .	4	2.1 語音分析簡介 . . . . .	4
. . . . .	4	2.2 語音特徵參數擷取 . . . . .	5
. . . . .	14	2.3 參考模型架構 . . . . .	14
. . . . .	14	2.4 辨認效果評估 . . . . .	14
. . . . .	21	2.5 本章結論 . . . . .	22
. . . . .	23	第三章 雜訊效應之補償方法 . . . . .	23
. . . . .	23	3.1 語音逆頻譜向量在雜訊環境下之行為 . . . . .	23
. . . . .	31	3.2 逆頻譜向量在雜訊下之表示 . . . . .	31
. . . . .	31	3.3 逆頻譜均值向量之線性調整法 . . . . .	33
. . . . .	35	3.4 特徵偏移向量之線性調整法 . . . . .	35
. . . . .	35	3.5 本章結論 . . . . .	37
. . . . .	38	第四章 各種補償方法在不同雜訊環境下的實驗 . . . . .	38
. . . . .	38	4.1 語音資料庫簡介 . . . . .	38
. . . . .	39	4.2 實驗環境設定 . . . . .	39
. . . . .	39	4.3 基本系統在不同雜訊環境下的實驗結果 . . . . .	40
. . . . .	41	4.4 逆頻譜線性調整法的實驗結果 . . . . .	41
. . . . .	41	4.5 特徵偏移向量之線性調整法的實驗結果 . . . . .	45
. . . . .	49	4.6 本章結論 . . . . .	49
. . . . .	49	第五章 結論 . . . . .	50
. . . . .	52	參考文獻 . . . . .	50
. . . . .	52	附錄A . . . . .	52
. . . . .	55		55

## 參考文獻

- [1]李立民，雜訊環境下語音辨認之研究，國立清華大學電機工程研究所博士論文，1995。
- [2]謝秀琴，數位語音訊號基本原理，全華科技，台北，1996。
- [3]Guan, C., Chen, Y. and Wu, B., " Direct modification on LPC coefficients with application to speech enhancement and improving the performance of speech recognition in noise," IEEE int. conf. Acoust., Speech Signal Processing, Vol. , pp.107-110, 1993.
- [4]Gales, M.J.F. and Young, S., " An improved approach to the hidden Markov model decomposition of speech and noise," IEEE Int. Conf. Acoust., Speech, Signal Processing., pp.1-233-236, 1992.
- [5]Gales, M.J.F. and Young, S.J., " Cepstral parameter compensation for HMM recognition in noise," Speech Communication, Vol.12, No.3, pp. 231-239, 1993.
- [6]Hwang, T.-H., Lee, L.-M. and Wang, H.-C., " Feature adaptation using deviation vector for robust speech recognition in noisy environment," Proceedings of IEEE international conference on Acoust., Speech Signal Processing, pp.1227-1230, 1997.

- [7]Hwang, T.-H., Lee, L.-M. and Wang, H.-C., " Cepstral behaviors due to additive noise and a compensation scheme for noisy speech recognition, " IEE Proc.-Vis. Image Signal Process., Vol.145, No.5, pp.316-321, 1998.
- [8]Juang, B.H., Levinson, S.E. and Sondhi, M.M., " Maximum likelihood estimation for multivariate mixture observations of Markov chains, " IEEE Trans. Inform., vol. IT-32, No.2, pp. 307-309, 1986.
- [9]Lee, L.-M., Chen, J.-K. and Wang, H.-C., " Nonlinear cepstral equalization method for noisy speech recognition, " IEE Proc.-Vis. Image Signal Process., Vol.141, No.6, pp.397-402, 1994.
- [10]Lee, L.-M. and Wang, H.-C., " An extended Levinson-Durbin algorithm for the analysis of noisy autoregressive process, " IEEE Signal Process. Letters, Vol.3, No.1, pp.13-15, 1996.
- [11]Lee, C.-H. and Huo Q., " On adaptive Decision Rules and Decision Parameter Adaptation for Automatic Speech Recognition, " Proceedings of IEEE, Vol.88, No.8, pp.1241-1269, August, 2000.
- [12]Mansour, D. and Juang, B.H., " A family of distortion measures based upon projection operation for robust speech recognition, " IEEE Trans. Acoust. Speech Signal Process., No. 37, pp. 1659-1671, 1989.
- [13]Moreno, P.J., Raj, B., and Stern , R.M., " A vector Taylor series approach for environment-independent speech recognition, " Proceedings of IEEE international conference on Acoust, Speech Signal Processing , pp.733-736, 1996.
- [14]Nadas, A. , Nahamoo, D. and Picheny, M.A., " Speech recognition using noise adaptive prototypes, " IEEE Trans. Acoust., Speech, Signal Processing, Vol-37, No.10, pp.1495-1503, 1989.
- [15]Quatieri, T.F., Discrete-Time Speech Signal Processing, Prentice Hall, New Jersey, 2002.
- [16]Rabiner, L.R., " A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, " Proceedings of IEEE , Vol.77, No.2, pp.257-286, 1989.
- [17]Rabiner, L.R., Lee, C.H., Juang, B.H., Wilpon, J.G., " HMM clustering for connected word recognition, " IEEE Int. Conf. Acoust., Speech Signal Processing, Vol.1, pp.405-408, 1989.
- [18]Rabiner, L.R. and Juang, B.H., Fundamentals of Speech Recognition, Prentice Hall, New Jersey, 1993.
- [19]Varga, A.P. and Moore, R.K. , " Hidden Markov model decomposition of speech and noise, " IEEE Int. Conf. Acoust., Speech, Signal Processing, pp.845-848, 1991.