Endpoint Detection For Speech Under Noisy Environments

鄭凱元、李立民

E-mail: 9607414@mail.dyu.edu.tw

ABSTRACT

In this paper, we investigate endpoint detection algorithms for noisy speech. We examine and analyze the performances of three endpoint detection methods for speech under five kinds of noisy environments and four SNR levels. In our detecting process, the first several frames of signal were supposed to be background noise and used to determine the threshold values for detection algorithms. The low energy threshold values are adjusted by constraining the durations between starting point and end point to the detected body of speech. This can increase the detection accuracy. In addition, we use smoothing method to smooth out the irregular fluctuation of the estimated feature curve so that detection can be easier. Then we try to adjust the coefficient for threshold values to improve the detection accuracy. Experimental results show that the energy-ZCR based method have the best result under high SNR environment. The KL distance based method has the most consistent performance under all conditions. The entropy based method has the least performance degradation under high level of noise. Finally, we use non-linear spectral subtraction (NSS) method to remove noise and obtain robust estimate of speech energy. Experimental results show that NSS based detection method can largely improve the accuracy for low SNR environments. Key Words : endpoint detection, energy-ZCR, voice activity detection

Keywords : endpoint detection ; energy-ZCR ; voice activity detection

Table of Contents

REFERENCES

[1] Xuedong Huang, Alex Acero, Hsiao-Wuen Hon,“ Spoken Language Processing ”, 1st Edition, Prentice Hall, 2005.
[2] 王小川， 語音訊號處理， 全華科技，台北市， 2005。
[3] L. Lamel, L. Rabiner, A. Rosenberg, and J. Wilpon,“ An improved endpoint detector for isolated word recognition ”, IEEE ASSP Mag, vol.29, pp. 777-785, 1981.
[4] Jean-Claude Junqua, Brian Mak and Ben Reaves,“ A Robust Algorithm for Word Boundary Detection in the Presence of Noise ”, IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, Vol. 2, No. 3, July 1997, pp. 406-412.
[5] G. S. Ying, C. D. Mitchell, and L. H. Jamieson,“ Endpoint Detection of Isolated Utterances Based on a Modified Teager Energy measurement ”, Proceedings of the 1993 IEEE International Conference on Acoustics, Speech, and Signal Processing, Minneapolis, Minnesota, April 1993, pp. II.732-II.735.
[6] Lingyun Gu and Stephen A. Zahorian,“ A New Robust Algorithm for Isolated Word Endpoint Detection ”, IV-4161 International Conference on Acoustics, Speech, and Signal Processing, Orlando, FL, May 13-17, 2002.
[7] Beena Ahmed and W. Harvey Holmes,“ A Voice Activity Detector Using The Chi-Square ”, ICASSP, 2004.
[8] Javier Ramirez, Jose C. Segura, Carmen Benitez, Angel de la Torre, Antonio Rubio,“ Improved Voice Activity Detection Combining Noise Reduction and Subband Divergence Measures ”, Proc. INTERSPEECH 2004, ICSLP, Jeju Island, Korea, October 2004.
[9] 賴辰瑋，“ 強健性語音辨認之研究:語音前端端點偵測與語音強化法 ”， 國立暨南國際大學碩士論文，南投，民國94年6月。
[10] Jialin Shen, Jeihweih Hung, Linshan Lee,“ Robust Entropy-based Endpoint Detection For Speech Recognition In Noisy Environments ”, International Conference on Spoken Language Processing, Sydney, 1998.
[11] JIA Chuan, Bo XU,“ An Improved Entropy-based Endpoint Detection Algorithm ”, ISCSLP, 2002, paper 96.
[12] S. Young, The HTK Book, Version 3.3, Cambridge University Engineering Department, 2005.
[13] 蘇盈安，“ 以動態音框長度調整作語者驗證之研究 ”， 大葉大學碩士論文， 彰化， 民國94年6月。
[14] 郭正雄，“ 強健性語音辨識之研究:改良式頻譜消去法 ”， 國立暨南國際大學碩士論文，民國93年6月。
[15] 韋曉東，胡光銳，任曉林，“ 應用倒譜特徵的帶噪語音端點檢測方法 ”， 第34卷第2期2000年2月上海交通大學學報。
[16] Liang-sheng Huang, Chung-go Yang,“ A Novel Approach to Robust Speech Endpoint Eetection In Car Environments ”, International Conference On Acoustics, Speech And Signal Processing, 2000.