Journal of Engineering Technology
Volume. 3, Jan. 2015, Pages 77-80
ORIGINAL ARTICLE

On a Transition Detection Method by Flattering the Spectrum of Speech Signal

Seong-Geon Bae and Myung-Jin Bae

Department Electronic Communication, Daelim University College 29, Imgok-rd, Dongan-gu, Anyang-Si, Gyeonggi-Do, 431-715, Korea
Information and Telecommunication Department, Soongsil University, 1-1 Sangdo 5 dong DongJak-Ku,Seoul, 156-743, Republic of Korea

Abstract. In speech signal processing, there are various methods to enhanced the signals. But due to the characteristics of the deterioration of the Speech, the restoration of the clipped signals are focused on the voiced sound, making the researchers focus on the quality enhancement of the low frequency. Therefore we have focused on emphasizing the ratio of the low frequency and the differentiation of the voiced and the unvoiced sounds in enhancing the quality and restoration of the speech. In this research we propose a method which enhances the voice signals by dividing the down mix and the low frequency. We differentiate the signals and use the fact that signals have little effect on perception to make voice part, and isolated the high frequency part through low pass filter.

Keywords: Non-uniform sampling, Low-pass Filter, Formant

1 INTRODUCTION

In speech signal processing, the main point is to consider especially the data transfer and compressibility, processing speed in the information the signals are transferring. The remaining components that exists in the speech signal is through to be because of the high correlation between samples. Therefore, the improvement of the clipped signals, since it is mostly made in a voiced signal which has relatively high amplitude, is done by we must properly use even sampling methods and low frequency wave filter to remove the remains in a sample, the sample that has high correlation between the samples and have low effect on perception. According to this, when looking at the characteristics of the remains in the angle of voice perception, the peak and valley in a speech sample act as a very important element. Especially, since the remaining information are just peaks and valleys when it undergoes differentiating of the sound signals and clipping, it is clear that the samples between the peak and the valley is a unessential in perception level.[4].

In this research we have proposed a method that detects low frequency division and peak & valley and compares it with detected reference spectrum. We used the fact that the signal that was made when the original signal was differentiated and clipped had little effect on the perspective to make a voice sound section and isolated high frequency through low frequency[1][2].

2 EXISTING METHODS

In voice signal processing, the main point is to consider especially the data transfer and compressibility, processing speed in the information the sound signals are transferring. The remaining components that exists in the sound signal is through to be because of the high correlation between samples. Therefore, the improvement of the clipped sounds, since it is mostly made in a voiced sound which has relatively high amplitude, is done by we must properly use even sampling methods and low frequency wave filter to remove the remains in a sample, the sample that has high correlation between the samples and have low effect on perception. According to this, when looking at the characteristics of the remains in the angle of voice perception, the peak and valley in a sound sample act as a very important element. Especially, since the remaining information are just peaks and valleys when it undergoes differentiating of the sound signals and clipping, it is clear that the samples between the peak and the valley is a unessential in perception level.[4].

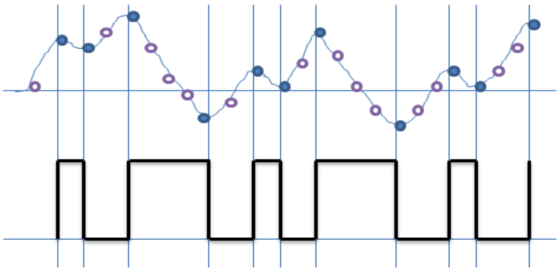


Figure. 1. Quantization influence signal.

In voice signals the transition section is a key element in playing a continued sound. Therefore the detection of the transition section it very important in saving the sound's nature. The comparison method, on the frequency side, will have difficulties in detecting at the precise section or noise section since it detects by making a comparison. Also many methods were proposed to overcome such error, but realistic misconception rate is still high[1][5].

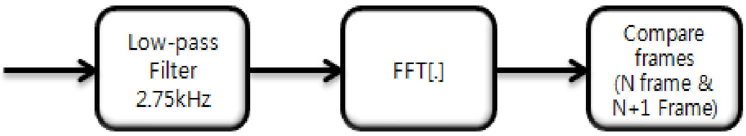


Figure. 2. General Signal Processing in Speech Signal.

As it can be seen in the picture when the deviation between the n frame and the n+1 spectrum is low, we consider it as the same sound and when it is high we consider it as a transition section. Therefore we get to reconstruct signal through down-mix, nonuniform sampling, and low frequency filtering[5][6].

3 PROPOSED METHOD

In the thesis a detection method that great in a transition section that connects two sounds is proposed. The inputted signals goes through the low frequency filter and eliminates irregular signals. In the previous method, we would done FTT and detect the basic signal and use the

Journal of Engineering Technology
Volume. 3, Jan. 2015, Pages 77-80
ORIGINAL ARTICLE

characteristics of differentiation and integration signals to detect the peak and the valley. Afterwards we reconstruct the signal and isolate the unnecessary component. Since the recomposed signals are made into cosine signal, the signal is reconstructed into a very clean state. Therefore we detect the transition section by comparing the signals in its original and recomposed state and we run it through a filter[1][7][8].

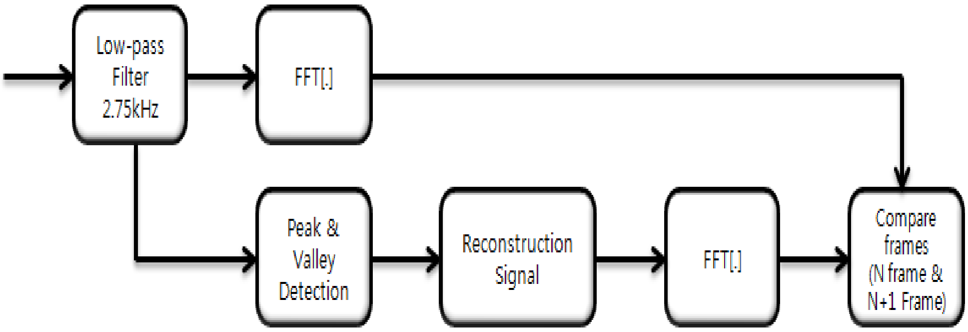
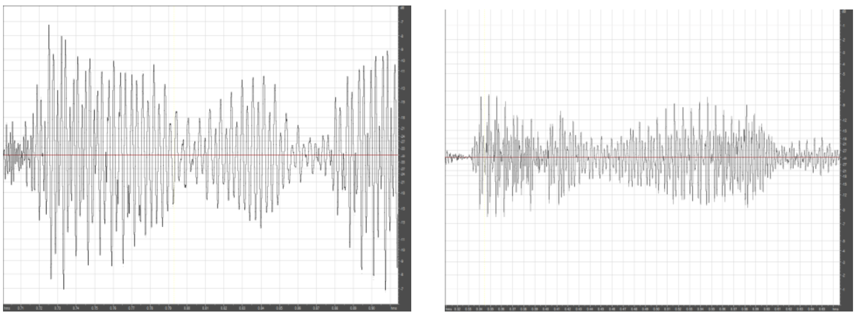
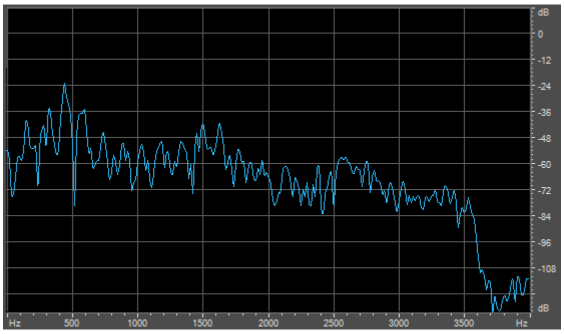


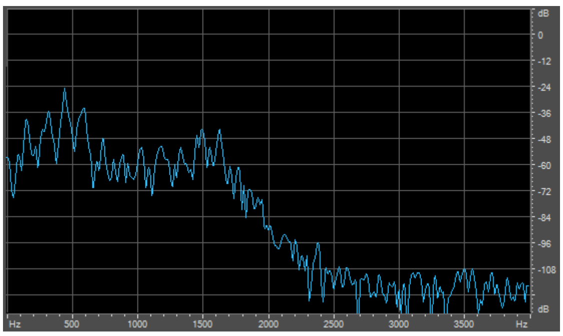
Figure. 3. Proposed method



(a) Original Signal (b) Low-pass Filtered Signal



(a) FFT of Original signal



(b) FFT of Low-pass Filtered Signal

(c) Table of Detection rate
Fig. 4. Results of proposed method

Table 1. Results of Proposed method

Title1	Title2		Title3	
	Errors	Detection	Errors	Detection
Test1	3%	89%	2%	93%
Test2	5%	89%	3%	96%
Test3	7%	89%	3%	85%
Total	5%	88.3%	3.3%	91.3%

4 CONCLUSION

In this thesis we have proposed a method that detects transition section to vitalize the nature that is important in a continuous sound signal. In sound signals, since the transition section has the highest change of sound, a precise detection is hard. In this thesis we have used spectrum comparison to this characteristics and used peak & valley method to used reference spectrum. Although we could use various techniques to recompose the signals, since its process emphasizes perception, it has a suitable characteristics to make and use reference spectrum.

REFERENCES

1. S. Gazor and W. Zhang, “A soft voice activity detector based on a Laplacian-Gaussian model," IEEE Trans. Speech Audio Processing, Vol. 11, No. 5, pp. 498-505, Sep. 2003
2. T. Agarwal and P. Kabal, “Pre-processing of noisy speech for voice coders," in Proc. IEEE Workshop on Speech Coding, Tsukaba, Japan, pp.169-171, Oct. 2002
3. J. Thyssen, Y. Gao, A. Benyassine, E. Shlomot, C. C. Murgia, H. Y. Su, K. Mano, Y. Hiwasaki, H. Ehara, K. Yasunaga, C. Lamblin, B. Kovesi, J. Stegmann, and H. G. Gang, “A candidate for the ITU-T 4 kbit/s speech coding standard," in Proc. 26th IEEE Int. Conf. Acoust. Speech Signal Process., ICASSP-01, Salt Lake City, UT, Vol. 2, pp.681-684, May. 2001
4. J.Sohn, N. S. Kim and W. Sung, “A statistical model-based voice activity detector,” IEEE Signal Processing Lett., Vol. 6, No 1, pp. 1-3, Jan. 1999.
5. A. J. Accardi and R. V. Cox, “A modular approach to speech enhancement with an application to speech coding,” J.Acout. Soc. Am, Vol. 10, No. 3, p.1245, Sep. 2001.
6. Jeong chan-joong, Bae myeong-jin, “Analysis regarding vocalizations of teuroteu singers”, The institute of electronics engineers of Korea, Summer scholarship contest of 2009 p. 1090~1091, 2009.
7. M. J. Bae and S. H. Lee, “ Digital Voice Signal Analysis”, Books Publishing Dong Young, 1998, ch. 3,6
8. J. Sohn, N. S. Kim, and W. Sung, “A statistical model-based voice activity detector," IEEE Signal Processing Lett., Vol. 6, No. 1, pp. 1-3, Jan. 1999