Pitch Error Improved with SNR Compensation

Hyung-Woo Park, Seong-Geon Bae and Myung-Jin Bae[1]

Information and Telecommunication Department,Soongsil University,
1-1 Sangdo 5 dong DongJak-Ku,Seoul,
156-743, Republic of Korea
park.hyungwoo@gmail.com, sgbae123@empal.com, mjbae@ssu.ac.kr

Abstract

On speech signal processing, it is very important to find the fundamental frequency of voice. The reason is why it is used in variable places, such as speech-enhancement-system, speech-recognition system, speaker-classification-system, and handicapped assisting-system. However the pitch detection is difficult when the original signal is corrupted by noise, or put in transition section of voice. In this paper, we make proposal of the method that enhance accuracy of pitch detection system, through SNR compensation using time-domain SNR estimator with continuous voice signal. And we proved the performance of the detector, in drawing pitch contour of variable SNR signals.

Keywords: *Pitch detection, Pitch Gross error, SNR compensation, continuous speech SNR.*

1. Introduction

On speech signal processing, we applied different method of processing to the voices as per the type. We distinguished the types of sound – voiced/unvoiced and silence section, and as per reason of sound creation. In 1950's, Fant proposed a linear-combination model of speech which is consisted with voiced and unvoiced section in time domain. That is operating with impulse- train and white-noise source by separated time slot [1, 2].

In the voice, Pitch is the basic vibration of vocal cords. And it means fundamental frequency of speaker. And Fundamental frequency has characteristics of speaker. When speech recognition system is operating, we normally remove the pitch information, In order to reduce the property of speaker. And if we can detect the pitch more accurate, we can apply variable place to use, such as in speaker classification, in speech enhance system, in speech synthesis system or speech change system [1-4].

In this paper, we make proposal of the method that enhance accuracy of pitch detection system, through SNR compensation using continuous speech signal. When the pitch detection is difficult, the original signal is corrupted by noise, or put in transition section of voiced/unvoiced. Because those section, the value of SNR is low. We propose the enhanced pitch detection system using SNR estimation method for continuous speech signal. And using that enhances the speech signal before pitch calculation. Before this paper, we have proposed a SNR estimation system for continuous speech signal in [3, 5]. And this system can make a segment SNR in continuous speech signal.

In Chapter 2 introduce the SNR estimation. In Chapter 3 investigated for pitch detection method. In Chapter 4 the proposed method are discussed. Chapter 5 is Experimental methods and results. Chapter 6 is conclusion.

[1] Corresponding author : Myung Jin Bae(mjbae@ssu.ac.kr)

2. SNR estimation

In speech signal processing, usually noise estimation algorithm need flexibility for variable environment and it can only apply on silence region to avoid effects of speech signal. So we have to preprocess finding voiced region before noise estimation. However, if received signal has not silence region, we cannot apply that method.

We use the new method to get SNR for continuous speech signal without silence region detection. The speech signal can be separated with voiced and unvoiced signal. And we use that property of speech signal. We estimate the SNR in separate section of voiced and unvoiced. In voiced region, we used the cross-correlation method to decide noise level factor of the signal. And in unvoiced region, we use log-spectrum distance method for calculate the noise level factor from received signal to estimated LPC parameter. Figure 1 is the block diagram of SNR estimation system which is used in this paper [3, 4].

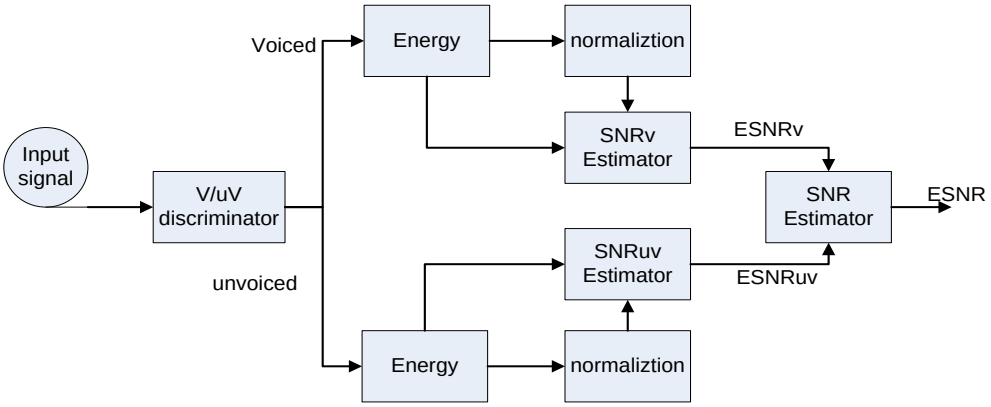

Figure 1. Block Diagram of SNR Estimator [3]

3. Detection Method of the Pitch and Harmonics

The Pitch detection method has already been proposed, various methods have been studied [1][2]. Proposed and studied, but how the characteristics of each one, and presented a clear pitch detection method does not. Existing pitch detection method, there are several ways to check. For example peak comparison method using parallel processing, using the autocorrelation peak highlighted, clipping using the comparative method, using the value of the difference between a voice signal average magnitude difference function (AMDF) law or in the frequency domain to detect cepstral method [1-6].

4. Proposed Method of Pitch Detection with SNR Enhancement Processing

Speech signal is time variant signal which includes pitch and formant. However, in short time analysis, the signal is quasi-periodic that component is pitch and also formant is group energy envelope. In noise corrupted voice signal, start and end point of speech, transition interval of the voice signal, hard to find each component and it is easily confused about some element of the signal. In this short time section of the signal-to-noise ratio analysis, the result of SNR is low [3, 4]. And Low SNR section, we can enhance the corrupted signal with SNR estimator and speech enhancer. After the SNR compensator, we can get more accurate pitch

of voiced signal. And also, we can apply logarithm compensation with small signal and low SNR signal, after that operation. In Figure 2 represent the proposed method by block diagram.

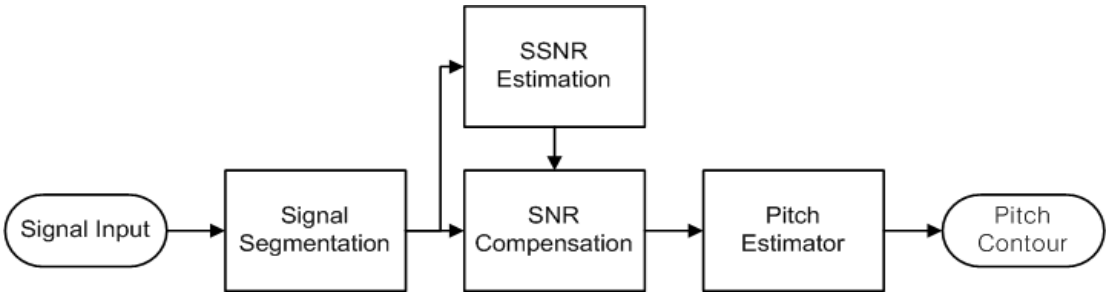

Figure 2. Block Diagram of Proposed Method

We separate the input signal to short time domain. And we analyze the segmented-SNR of each section of signal. When we need pitch information in SNR estimator, we use pitch search algorithm. After get the SSNR of speech, we are using decision logic of the threshold SNR compensation. Finally, we consider the propriety of calculated pitch.

5. Experiment and Result

5.1 Experiment Method

To verify the performance of the proposed method in a lab environment to prepare recorded audio signal, SSNR estimate of the sample. Based on the estimated SSNR, SNR reinforces lower SNR part of the signal before the pitch detection and compare of the results with processed value and non-processed value. The data recorded by each five men and women repeatedly, 'National Charter of Education' was used. Experimental data is sampled at 8 ㎑ and the number of bits per sample, 16 bits / sample was used.

We experiment in the same order as in Figure 2. We processed lower SSNR section of the voice signal to strengthen the periodicity emphasis and emphasized harmonics. And we use the autocorrelation method for pitch detection which is '3 level central cutting'. 80 ~ 500 ㎐ pitch inspection at intervals, and the other is detected in the range deemed to be zero value for incorrect value.

5.2 Experiment Result

The Figure 3 and Figure 4 are one of the experiment results. We experiment the data of the SNR is the signal by varying. We make the data that the SNR changes from 30dB to 0dB. And we enhance the signals by estimated SSNR. We use white Gaussian noise for mixing 0~30dB SNR signals. In every Figure, a) is the input speech signal, b) is ordinary pitch detection result and c) is suggested enhanced signal pitch contour. Each result plot shows in b), error is getting more in low SNR signal, and in figure c), errors are corrected.

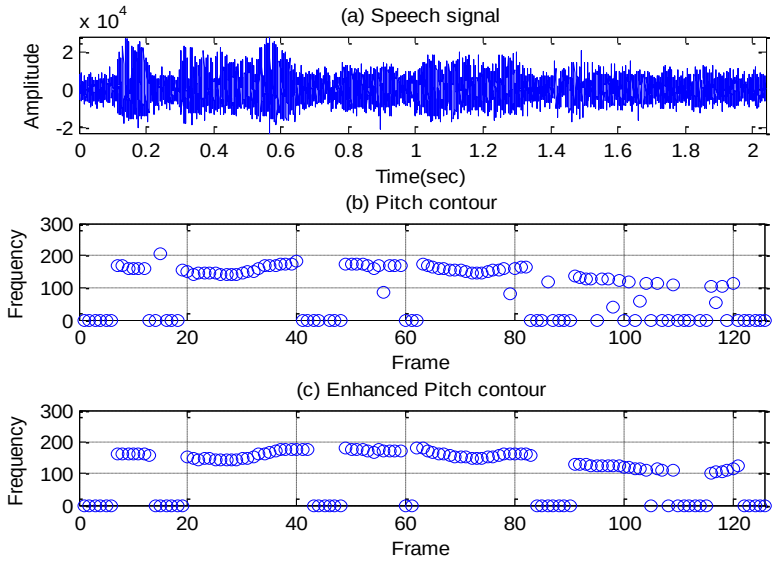

Figure 3. Experiment Result of 0dB SNR Signal

Figure 4 is zoomed data of 0 dB SNR signal that shows 80~130 frames. Especially in this interval, the performance of the proposed pitch detector that I will appear. Figure 4. b) shows more then 8 errors but enhanced result has less than 2 errors even though only in low SNR signal, but we still find advantages.

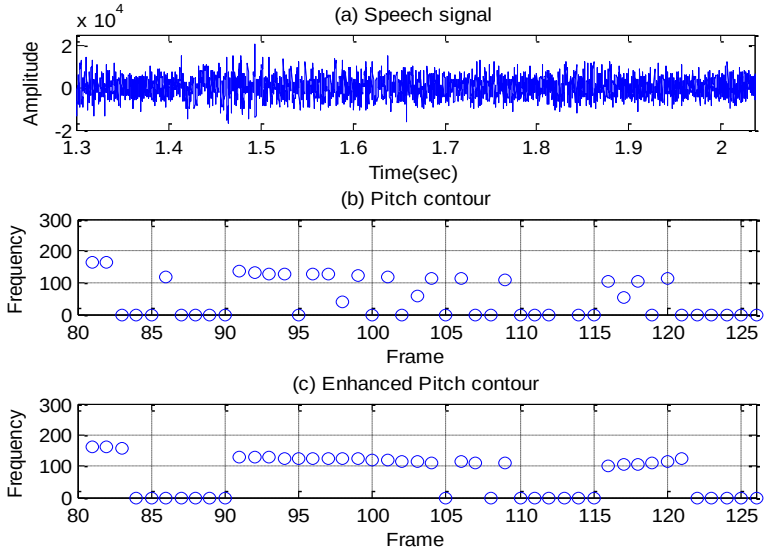

Figure 4. Zoomed Result of 0dB SNR Signal

6. Conclusion

This Accurate detection of pitch in speech processing is important. Using a linear model proposed by Fant analysis and processing errors occurred during the subsequent processing of pitch detection error accumulation leads to persistent. In this paper, in order to improve the accuracy of pitch detection SNR, SSNR estimation of speech signals and to compensate for pitch detection through preprocessing to smoothly and accurately proposed and experimental results are compared to leading pitch.

The pitch of the voice is used in a variety of voice processing. The pitch height of the naturalness of a speech synthesizer or speech recognizer is used to increase the recognition rate, and determines the characteristics of speech signals is an important value. In order to accurately detect these pitches in this paper for speech signals, we present the lower section of the voice signal that occurs in SSNR, propose ways to reduce the pitch gross error. By estimating the SNR of the continuous speech, an error can occur with its value in the interval by compensating the signal processing method is proposed. Experiments results by the proposed method, Speech, and the transition interval between phoneme and phoneme pitch in the lower part of the energy has been confirmed that improved detection capability.

References

[1] M. J. Bae and S. H. Lee, Editor, Digital Voice Signal Analysis, Books Publishing Dong Young, Korea **(1998)**.
[2] L. R. Rabiner and R. W. Schafer, Editor, Introduction to Digital Speech Processing, The essence of knowledge, USA, **(2007)**.
[3] H. W. Park and M. J. Bae, "A study on SNR Estimation of Continuous Speech Signal with Additive Colored Noise", GESTS Society, GESTS Int'l Trans. Computer Science and Engineering, Vol. 64, No. 1, pp. 137-147 **(2011)**.
[4] Y. H. Song, J. H. Ahn and M. J. Bae, "On the noise detection from correlation of near pitch waveforms", GESTS Society, GESTS Int'l Trans. Computer Science and Engineering, Vol. 44, No. 1, pp. 45-54, **(2008)** January.
[5] Ji-Soo O, Jeong-Jin Kang, Myung-Jae Lim, Ki-Young Lee, "Design and Implementation of an Emotion Recognition System using Physiological Signal", IWIT, Vol. 10, No. 1, pp. 57-62 **(2010)**.
[6] Yeon-Soo Lee and Young B. Park, "An acoustic study of feeling information extracting method", IWIT, Vol. 10, No. 1, pp. 51-55 **(2010)**.
[7] W.Y. Yang and Y. S. Jo, Editor, Digital Signal Processing and MATLAB, BRAINKOREA, Korea, **(2001)**.
[8] Y.W Oh, Editor, Voice Information Processing, Hong-Reung Press, Korea, **(1998)**.

Authors

Hyung-woo Park

Education : He received the MS. degree from Soongsil University in 2007. He is currently the Ph.D student at Soongsil University.

E-mail: park.hyungwoo@gmail.com

Seong-Geon Bae

He received the M.S. degree in Electronics Engineering from Konkuk University in 1995. He is currently the under Ph.D. degree at Soongsil University.

E-mail: sgbae123@empal.com

Myung-Jin Bae

Education & Work experience: He received the Ph.D. degree in Electronic Engineering from Seoul National University in 1987. He is currently the Professor of the Dept. of Information & Tele communication at Soongsil University.

E-mail: mjbae@ssu.ac.kr