THE JOURNAL OF THE ACOUSTICAL SOCIETY OF KOREA Vol.22, No.3E 2003. 9. pp. 110~112

Filtering of a Dissonant Frequency for Speech Enhancement

Sang-Ki Kang*, Seong-Joon Baek**, Ki-Yong Lee***, Koeng-Mo Sung****
*Modem Algorithm Lab. Telecommunication R&D Center Telecommunication Network Samsung Electronics Co., Ltd
**School of Electronics & Computer Engineering, Chonnam National University
***School of Electronic Engineering, Soongsil University
****School of Electric & Computer Engineering, Seoul National University
(Received November 4 2002; revised August 7 2003; accepted August 19 2003)

Abstract

There have been numerous studies on the enhancement of the noisy speech signal. In this paper, we propose a completely new speech enhancement scheme, that is, a filtering of a dissonant frequency (especially F# in each octave of the tempered scale) based on the fundamental frequency which is developed in frequency domain. In order to evaluate the performance of the proposed enhancement scheme, subjective tests (MOS tests) were conducted. The subjective test results indicate that the proposed method provides a significant gain in audible improvement especially for speech contaminated by colored noise and speaking in a husky voice. Therefore when the filter is employed as a pre-filter for speech enhancement, the output speech quality and intelligibility is greatly enhanced.

Keywords: *Dissonant frequency, Fundamental frequency, Parametric cubic convolution*

I. Introduction

Degradation of the quality or intelligibility of speech caused by the acoustic background noise is common in most practical situations. In general, the addition of noise reduces intelligibility and degrades the performance of digital voice processors used for applications such as speech compression and recognition. Therefore, the problem of removing the uncorrelated noise component from the noisy speech signal, i.e., speech enhancement, has received considerable attention. There have been numerous studies on the enhancement of the noisy speech signal. Many different types of speech enhancement algorithms have been proposed and tested[1-7]. They can be grouped into three categories. The first category contains speech enhancement algorithms based on the short-time spectral estimation such as the spectrum subtraction[1] and Wiener filtering[2,3] techniques. The algorithms in the second category are comb filtering and adaptive noise canceling techniques which exploit the quasi-periodic nature of the speech signal[4-6]. The third category contains algorithms that are based on the statistical model of the speech signal and use hidden Markov model (HMM) or expectation and maximization (EM) for speech enhancement[7].

In this paper, we propose a completely new speech enhancement scheme, that is, a filtering of a dissonant frequency (especially F# in each octave of the tempered

Corresponding author: Sang-Ki Kang (sangki.kang@samsung.com)
Modem Algorithm Lab. Telecommunication R&D Center Telecommunication Network Samsung Electronics Co., Ltd, Dong Suwon P.O.BOX 105, 416, Maetan-3dong, Paldal-gu, Suwon-si, Gyeonggi-do, 442-600, Korea

scale) based on an improved fundamental frequency estimation using Parametric Cubic Convolution after peak sharpening and modified Sub-Harmonic Summation (SHS) [8].

II. A Dissonant Frequency Filtering

The standard musicological definition is that a musical interval is consonant if it sounds pleasant or restful. Dissonance, on the other hand, is the degree to which an interval sounds unpleasant or rough. Dissonant intervals generally feel tense and unresolved. When reference frequency is C, C# and B and F# are such examples. Among them, F# is known as "the Devil's interval" in music[9]. So we eliminates F# as a typical example to make speech less annoying and pleasant. In a dissonant frequency filtering of input speech signal to enhance perceptual speech intelligibility, it is important accurately to estimate the fundamental frequency that is closely related to a dissonant frequency filtering. A filtering of a dissonant frequency based on the fundamental frequency estimation which is developed in frequency domain is as follows. First, frequency transform of the original input signal using Fast Fourier Transform (FFT) is carried out and then an improved fundamental frequency estimation is performed using Parametric Cubic Convolution[8]. Next, filtering of the dissonant frequency based on the obtained fundamental frequency (pitch) is performed. The dissonant frequency F_d corresponding to F# relative to fundamental frequency is defined as follows[10].

$$F\_d = F\_0 \times 2^{(n+1/2)}, \qquad n = 0,1,\ldots,7 \tag{1}$$

Here is F_d dissonant frequency, F_0 is fundamental frequency obtained by an improved Parametric Cubic Convolution and n varies from 0 to 7 on condition that F_d is less than the half of sampling frequency. The filtering of dissonant frequency F_d is performed if the following condition is satisfied.

$$F\_0 \times 2^{(n+11/24)} < F\_d < F\_0 \times 2^{(n+13/24)}, \quad n = 0,1,\ldots,7 \tag{2}$$

where the range corresponds to a half tone or a semitone.

Very simple filtering algorithm is employed in this paper. If speech shows a peak in the region defined above, that peak is smeared out, that is, the magnitude in that region is lowered to the neighboring magnitude while the phase is kept. Finally, the inverse Fourier Transform of filtered signal is carried out. Finally, the inverse Fourier Transform of filtered signal is carried out. This is illustrated in Fig. 1.

III. Experimental Results

The database for performance evaluation consists of 14 speech files collected from 5 males and 2 females, each one delivering 2 Korean sentences. Also we obtained a husky voice data. All utterances were sampled at 8 kHz. Also 512-point FFT padded with 256 zeros is used. Noise types considered in our experiments include computer fan, babble noise and car noise recorded inside a car moving approximately at a speed of 90 km/h. We obtain the noisy speech by adding noise to clean speech with the noise power being adjusted to achieve SNR=5 and 10 dB. In order to evaluate the performance of the proposed enhancement scheme, subjective tests (MOS tests) were conducted. Ten listeners participated in the test. The listeners were asked to give a score ranging from 1 (very bad) to 5 (excellent) according to the perceived quality. The results are illustrated in Table 1.

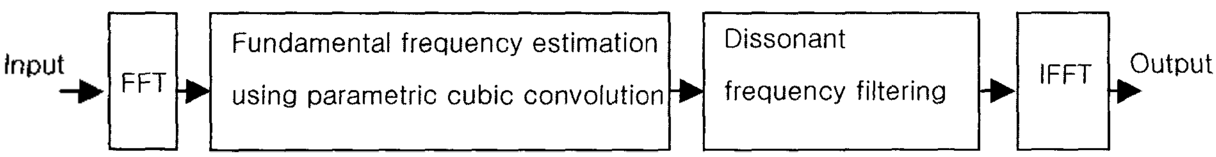

Figure 1. Speech quality improvement scheme.

Table 1. Results of subjective measures of enhanced speech.

Noise Types	SNR	MOS (unprocessed)	MOS (processed)
Computer	5 dB	2.83	3.04
Fan noise	10 dB	2.85	3.14
Babble	5 dB	2.81	3.17
Noise	10 dB	2.86	3.38
Car	5 dB	2.79	3.05
Noise	10 dB	2.83	3.29
Husky Voice		2.98	3.27

The subjective test results indicate that the proposed method provides a significant gain in audible improvement especially for speech contaminated by especially colored noise and a husky voice signal.

IV. Conclusions

We proposed a new speech enhancement scheme, that is, a filtering of a dissonant frequency (especially F#). The subjective test results indicated that the proposed method delivered improvement in terms of both speech intelligibility and perceived quality when compared with the unprocessed case and in a husky voice. Therefore when the filter is employed as a pre-filter for speech enhancement, the output speech quality and intelligibility is enhanced.

References

1. S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-**29**, 113-120, April 1979.
2. J. S. Lim and A. V. Oppenheim, "All-pole modeling of degraded speech," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-**26**, 197-210, June 1978.
3. J. Hansen and M. Clements, "Constrained iterative speech enhancement with application to speech recognition," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-**39** (4), 21-27, April 1991.
4. M. R. Samber, "Adaptive noise canceling for speech signals," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-**26**, 419-423, Oct. 1978.
5. J. S. Lim, A. V. Oppenheim, and L. D. Braida, "Evaluation of an adaptive comb filtering method for enhancing speech degraded by white noise addition," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-**26** (4), 354-358, April 1991.
6. S. F. Boll and D. C. Pulsipher, "Suppression of acoustic noise in speech using two microphone adaptive noise cancelation," *IEEE Trans. Acoust., Speech, Signal Processing*, ASSP-**28** (6), 752-753, Dec. 1980.
7. Y. Ephraim, "Statistical-model-based speech enhancement systems," *Proc. of IEEE*, **80** (10), 1526-1555, Oct. 1992.
8. H.-S. Pang, S.-J. Baek, and K.-M. Sung, "Improved fundamental frequency estimation using parametric cubic convolution," *IEICE Trans. Fundamentals*, E83-A (12), 2747-2750, 2000.
9. J. Tennenbaum, "The foundations of scientific Musical Tunning," *FIDELIO Magazine*, 1 (1), Winter 1991-92.
10. M. David Howard and J. Angus, *Acoustics and Psychoacoustics*, Focal Press, 1996.

[Profile]

• Sang-Ki Kang
The Journal of the Acoustical Society of Korea, Vol. 21, No. 4.

• Seong-Joon Baek
The Journal of the Acoustical Society of Korea, Vol. 16, No. 2.

• Ki-Yong Lee
The Journal of the Acoustical Society of Korea, Vol. 15, No. 3E.

• Koeng-Mo Sung
The Journal of the Acoustical Society of Korea, Vol. 14, No. 2.