International Journal on Recent and Innovation Trends in Computing and Communication
Volume: 2 Issue: 11
ISSN: 2321-8169
3344 - 3345

Echo Path Transfer Function Estimation for Spectral Subtraction-based Acoustic Echo Suppression

Jong Han Joo[1], Jung Hoon Lee[2], Young Sun Kim[2], Jae Young Kang[2], and Seung Ho Choi[2]
[1]Dept. of Electronic Engineering
[2]Dept. of Electronic and IT Media Engineering
Seoul National University of Science and Technology
E-mail address: shchoi@seoultech.ac.kr

***Abstract*—**In this study, we propose a novel technique for spectral subtraction (SS)-based acoustic echo suppression (AES). Conventional AES methods based on SS apply fixed weights to the estimated echo path transfer function (EPTF) at the current signal segment and to the EPTF estimated until the previous time interval. We propose a new EPTF estimation approach that adaptively updates the weight parameters in response to abrupt changes in the acoustic environment. From the experiments, we conclude that the developed techniques can be successfully used for the SS-based AES systems.

Keywords*—Acoustic echo suppression, echo path transfer function, spectral subtraction*

*****

I. INTRODUCTION

Recent studies have proved that there are various advantages in adopting spectral subtraction (SS) methods in a short-time Fourier transform (STFT) domain for acoustic echo suppression (AES) [1, 2]. In conventional SS-based AES methods, for updating the echo path transfer function (EPTF), fixed weights are applied to the estimated EPTF at the current signal segment and to the EPTF updated until the previous time interval [2]. We propose an adaptive algorithm that updates the weight parameters to consider abrupt changes in the acoustic environment due to background noises or double-talk (DT)

II. EPTF ESTIMATION FOR AES

A. Spectral Subtraction-Based AES

In the echo suppression system as shown in Fig. 1, $x(n)$ is a message prompt signal and $y(n)$ is an echo signal that is a portion of $d(n)$ transmitted from the near-end microphone. The $h(n)$ represents an echo path impulse response. In a barge-in situation, the residual echo $e(n)=d(n)-\hat{y}(n)$ is the estimate of user's voice $v(n)$.

In the spectral subtraction-based AES method, $X(i,k)$, $D(i,k)$, $\hat{H}(i,k)$ denote the spectra of the far-end speech signal, microphone input signal , estimated EPTF with frequency index i and frame index k, respectively. $\hat{H}(i,k)$ is iteratively computed as follows:

$$\hat{H}(i,k)=\frac{H_{num}(i,k)}{H_{den}(i,k)} \tag{1}$$

where

$$H_{num}=\lambda H_{num}(i,k-1)+(1-\lambda)X^{*}(i,k)D(i,k) \tag{2}$$

$$H_{den}=\lambda H_{den}(i,k-1)+(1-\lambda)X^{*}(i,k)X(i,k) \tag{3}$$

and λ is the weight parameter [3]. Then, the estimated echo magnitude spectrum $\hat{Y}(i,k)$ is given by

$$\hat{Y}(i,k)=\hat{H}(i,k)\,|X(i,k)| \tag{4}$$

and $|\hat{E}(i,k)|$, denoting the estimated short-time magnitude spectrum of residual echo signal, is given by

$$|\hat{E}(i,k)|=(|D(i,k)|^{\alpha}-\beta|\hat{Y}(i,k)|^{\alpha})^{\frac{1}{\alpha}}. \tag{5}$$

Subsequently, short-time phase of the microphone input, $\angle \mathrm{D(i,k)}$, is used as the phase of $\hat{\mathrm{E}}(\mathrm{i,k})$, i.e.

$$\hat{E}(i,k)=|E(i,k)|\,e^{j\angle \mathrm{D(i,k)}}. \tag{6}$$

B. Proposed EPTF Estimation

Conventional AES methods apply fixed-weights to the estimated EPTF at the current signal segment and to the EPTF estimated until the previous time interval, as in Eqs. (2) and (3). We replace the λ with a time-varying $\lambda(k)$ that is adaptively updated in response to an abrupt change in acoustic environment.

The parameter $\lambda(k)$ is controlled by the cross-correlation coefficient $\rho(k)$ between $|\hat{Y}(k)|$ and $|D(k)|$,

$$\rho(k)=\frac{\frac{1}{N}\sum_{i=0}^{N-1}|D(i,k)|\,|\hat{Y}(i,k)|}{||D(k)||\,||\hat{Y}(k)||}. \tag{7}$$

Then, we let $\lambda(k)$ decrease linearly with $\rho(k)$:

$$\lambda(k)=a\rho(k)+b,\ a<0. \tag{8}$$

ISSN: 2321-8169
3344 - 3345

III. PERFORMANCE EVALUATION

To evaluate the performance of the proposed AES, we used a woman's voice as the message signal $x(n)$, and four voices of two men and two women as the input signal $v(n)$. The audio files were sampled at 16 kHz. $x(n)$ was convoluted with the acoustic echo path impulse response before being mixed. For each frame of the Hamming-windowed signal, $x(n)$ and $d(n)$ were transformed into their spectra through 256-point DFT after zero padding. To obtain an objective comparison, we evaluated the performance of echo return loss enhancement (ERLE) [4], log-spectral distance (LSD) which are defined by

$$ERLE(dB) = 10\log_{10}[\frac{E[d^2(n)]}{E[e(n)^2]}] \quad (9)$$

$$LSD = \sqrt{\frac{1}{N}\sum_{k=0}^{N-1}(10\log_{10}\frac{|X(k)|^2}{|D(k)|^2})^2} \quad (10)$$

where n is the sample index and $E[\cdot]$ indicates the expected value. We used the LSD during the DT period, and the ERLE is used when DT is not included. Table 1 compares the values of ERLE and LSD obtained.

Conventional technique of AES usually obtained the best values of LSD and ERLE it can when λ was 0.9 — 0.95; however the proposed AES method obtained better ERLE and LSD values. It is evident that the performance of the proposed AES is superior to that of conventional AES.

TABLE I. COMPARISON OF ERLE AND LSD RESULTS

λ	Average LSD	Average ERLE
0.85	1.37	14.2
0.9	1.36	14.4
0.95	1.35	14.9
Proposed	1.30	18.2

IV. IV. CONCLUSION

We described a novel method for echo path transfer function estimation in the spectral subtraction-based acoustic echo suppression system. The proposed method adaptively updates weight parameters to consider the abrupt changes of acoustic environment. The experimental results showed that the developed techniques can be used successfully in both single-talk and double-talk conditions.

V. ACKNOWLEDGEMENTS

This study was supported by the Research Program funded by the Seoul National University of Science and Technology.

VI. REFERENCES

[1] C. Avendano, "Acoustic echo suppression in the STFT domain," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, Oct. 2001.

[2] C. Faller and J. Chen, "Suppressing acoustic echo in a spectral envelope space," *IEEE Trans. on Speech and Audio Processing*, vol. 13, no. 5, pp. 1048-1062, Sep. 2005.

[3] C. Faller and C. Tournery, "Robust echo control using a simple echo path model," *in Proc. IEEE Int. Conf. Acous., Speech Signal Processing*, vol. 5, pp. 281-284, 2006.

[4] T. Aboulnasr and K. Mayyas, "A robust variable step-size LMS-type algorithms: analysis and simulations," *IEEE Trans. on signal Processing*, vol. 45, no. 3, pp. 631-639, Mar. 1997.

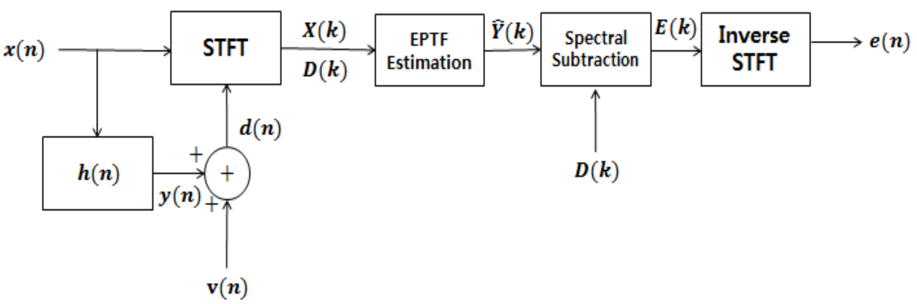

Figure 1. Block diagram of the conventional method

Figure 2. Block diagram of the proposed method