

RESEARCH

Open Access

Integrated acoustic echo and background noise suppression technique based on soft decision

Yun-Sik Park¹ and Joon-Hyuk Chang^{2*}

Abstract

In this paper, we propose an efficient integrated acoustic echo and noise suppression algorithm using the combined power of acoustic echo and background noise within a soft decision framework. The combined power of the acoustic echo and noise is adopted to the integrated suppression algorithm based on soft decision to address the artifacts such as the nonlinear distortion and the disturbed noise introduced from the conventional methods. Specifically, in the unified frequency domain architecture, the acoustic echo and noise signal are efficiently able to be suppressed through the acoustic echo suppression algorithm based on soft decision without the help of the additional noise reduction technique.

1 Introduction

Recently, hands-free systems are widely used for safety and convenience in the mobile communication. However, such an equipment introduces specific technical difficulties due to the background noise and the echoes by acoustic coupling between a loudspeaker and a microphone of this equipment [1,2]. Thus, for hands-free mobile equipment, the serial combination of the acoustic echo cancellation (AEC) and noise reduction (NR) algorithm has been predominantly considered to achieve the improved performance and sufficient quality of the transmitted speech signal [3,4]. Indeed, the performance of the conventional integrated system is significantly affected by the combined structure of the AEC and NR algorithm. Generally, in the conventional unified structure where the NR module exists after the AEC algorithm, noise estimation can be disturbed by the AEC processing. Also, in the unified structure where the NR algorithm is placed before the AEC algorithm, it also introduces non-linear distortions on the echo signal which can disturb the identification operation [5]. Therefore, much work has been dedicated to the problem of improving the performance of the combined structure depending on AEC and NR algorithm. In [6], Gustaffson et al. used a single perceptually motivated weighted rule to suppress both noise and residual echo in a frequency domain. However, this method needs the

adaptive echo canceller to identify the echo path impulse response for eliminating the undesired echo effect, which also affects the performance of the NR algorithm. In [7], Habets et al. presented the joint suppression technique of stationary (e.g., background noise) and non-stationary interference (e.g., echo) using a soft decision approach. But, an estimate of the variance of the echo signal was assumed to be known *a priori*, which inherently requires the AEC before the NR module. Other closely related technique by same authors is an approach of combined suppression of residual echo, reverberation, and background noise in a fashion of the post-filter following the traditional AEC [8]. But, the cancellation is performed directly on the waveform as in [7,8]. The algorithm is sensitive to the misalignment in the echo path response estimate. Also, it is hard to efficiently model the impulse responses lasting above milliseconds long with hundreds of coefficients. From this viewpoint, it is noted that a low complexity acoustic echo suppression (AES) algorithm by Faller [9] uses a spectral modification technique by incorporating the echo path response filter characterizing the actual echo path in a frequency domain. Recently, our previous approach in [10] presented the novel acoustic echo suppression (AES) algorithm based on soft decision without the help of the AEC and an additional residual echo suppression (RES), which conventional methods substantially need [10]. However, this technique has a problem in that the background noise is not taken into

* Correspondence: jchang@hanyang.ac.kr

²School of Electronic Engineering, Hanyang University, Seoul 133-791, Korea
Full list of author information is available at the end of the article

consideration for suppression, which can not be considered realistic.

In this paper, we propose a novel approach to the integrated suppression algorithm where the combined power of acoustic echo and background noise is incorporated based on soft decision as in [10] to directly suppress both strong acoustic echo and noise signal in a frequency domain. The proposed method efficiently estimates the echo and noise power separately and summates them to provide the unified framework in determining and modifying the suppression gain based on soft decision. This is clearly different from the conventional integrated strategies requiring the AEC and NR independently. For this, our approach directly estimates the spectral envelope of the echo signal instead of identifying the echo path impulse response in a time domain. Also, the background noise is estimated during near-end speech and echo-absent periods. In particular, the acoustic echo and noise signal are able to be reduced at a time through a single gain based on soft decision using the estimated combined power. Based on this, the proposed method can efficiently suppress the acoustic echo and noise without the help of an additional residual signal suppressor. Accordingly, the proposed unified structure addresses the problems associated with the residual echo and noise produced by the conventional unified structure where the NR operation is placed after the AEC algorithm or vice versa. The performance of the proposed algorithm is evaluated by both the subjective and objective quality tests and is demonstrated to be better than that of the conventional methods.

2 Proposed integrated suppression algorithm based on soft decision

In the previous section, we note that the previous AES technique in [10] needs the additional NR before/after the AES architecture for suppressing noise. However, this procedure could have a drawback such as the non-linear distortion on echo or the disturbed noise power estimate as happened in the conventional integrated system [5]. Considering the case that the NR operation is placed after the AES algorithm, the noise power estimation can be disturbed by the AES processing. On the contrary, in the unified structure where the NR algorithm is simply placed before AES, it also introduces non-linear distortions on echo signal, which can disturb the identification operation. In order to reduce the problem resulting from serially combined structure, we propose a novel approach as the integrated suppression system based on the combined power of acoustic echo and background noise as in Figure 1 showing the block diagram of the proposed system based on soft decision. From the figure, it can be seen in advance that the proposed method can suppress the acoustic echo and the

noise signal with a single gain based on soft decision. For this, the noise and echo spectral are separately and efficiently estimated and combined by a single power in the soft decision framework. Since we take the frequency domain AES algorithm in [10] as a baseline, we should reassume that two hypotheses to incorporating the discrete Fourier transform (DFT) spectrum of the noise signal $D(i,k)$, H_0 and H_1 , indicate near-end speech absence and presence as follows:

$$\begin{aligned} H_0 : \text{near - end speech absent} : Y(i, k) &= D(i, k) + E(i, k) \\ H_1 : \text{near - end speech present} : Y(i, k) &= D(i, k) + E(i, k) + S(i, k) \end{aligned} \quad (1)$$

where $E(i, k)$, $S(i, k)$, and $Y(i, k)$ represent the DFT spectra of the echo signal, the near-end speech, and the input signal picked up by the microphone with a time index i and frequency index k .

Under the assumption that $D(i, k)$, $E(i, k)$, and $S(i, k)$ are characterized by separate zero-mean complex Gaussian distributions, the following are obtained [10].

$$p(Y(i, k)|H_0) = \frac{1}{\pi \{\lambda_e(i, k) + \lambda_d(i, k)\}} \exp \left[-\frac{|Y(i, k)|^2}{\{\lambda_e(i, k) + \lambda_d(i, k)\}} \right] \quad (2)$$

$$p(Y(i, k)|H_1) = \frac{1}{\pi \{\lambda_s(i, k) + \lambda_e(i, k) + \lambda_d(i, k)\}} \exp \left[-\frac{|Y(i, k)|^2}{\{\lambda_s(i, k) + \lambda_e(i, k) + \lambda_d(i, k)\}} \right] \quad (3)$$

where $\lambda_e(i, k)$, $\lambda_d(i, k)$, and $\lambda_s(i, k)$ are the variance of the echo, noise, and near-end speech, respectively. The near-end speech absence probability (NSAP) $p(H_0|Y(i, k))$ for each frequency band is derived from Bayes' rule such that [10]:

$$\begin{aligned} p(H_0|Y(i, k)) &= \frac{p(Y(i, k)|H_0)p(H_0)}{p(Y(i, k)|H_0)p(H_0) + p(Y(i, k)|H_1)p(H_1)} \\ &= \frac{1}{1 + q\Lambda(Y(i, k))} \end{aligned} \quad (4)$$

where $q = p(H_1)/p(H_0)$ and $p(H_0) (= 1 - p(H_1))$ represent the *a priori* probability of near-end speech absence. Substituting (2) and (3) into (4), the likelihood ratio $\Lambda(Y(i, k))$ can be computed as follows:

$$\begin{aligned} \Lambda(Y(i, k)) &= \frac{p(Y(i, k)|H_1)}{p(Y(i, k)|H_0)} \\ &= \frac{1}{1 + \xi(i, k)} \exp \left[\frac{\gamma(i, k)\xi(i, k)}{1 + \xi(i, k)} \right] \end{aligned} \quad (5)$$

For (5), we define the *a posteriori* signal-to-combined power ratio (SCR) $\gamma(i, k)$ and the *a priori* SCR $\xi(i, k)$ by

$$\gamma(i, k) = \frac{|Y(i, k)|^2}{\lambda_{cd}(i, k)}, \quad \xi(i, k) \equiv \frac{\lambda_s(i, k)}{\lambda_{cd}(i, k)}. \quad (6)$$

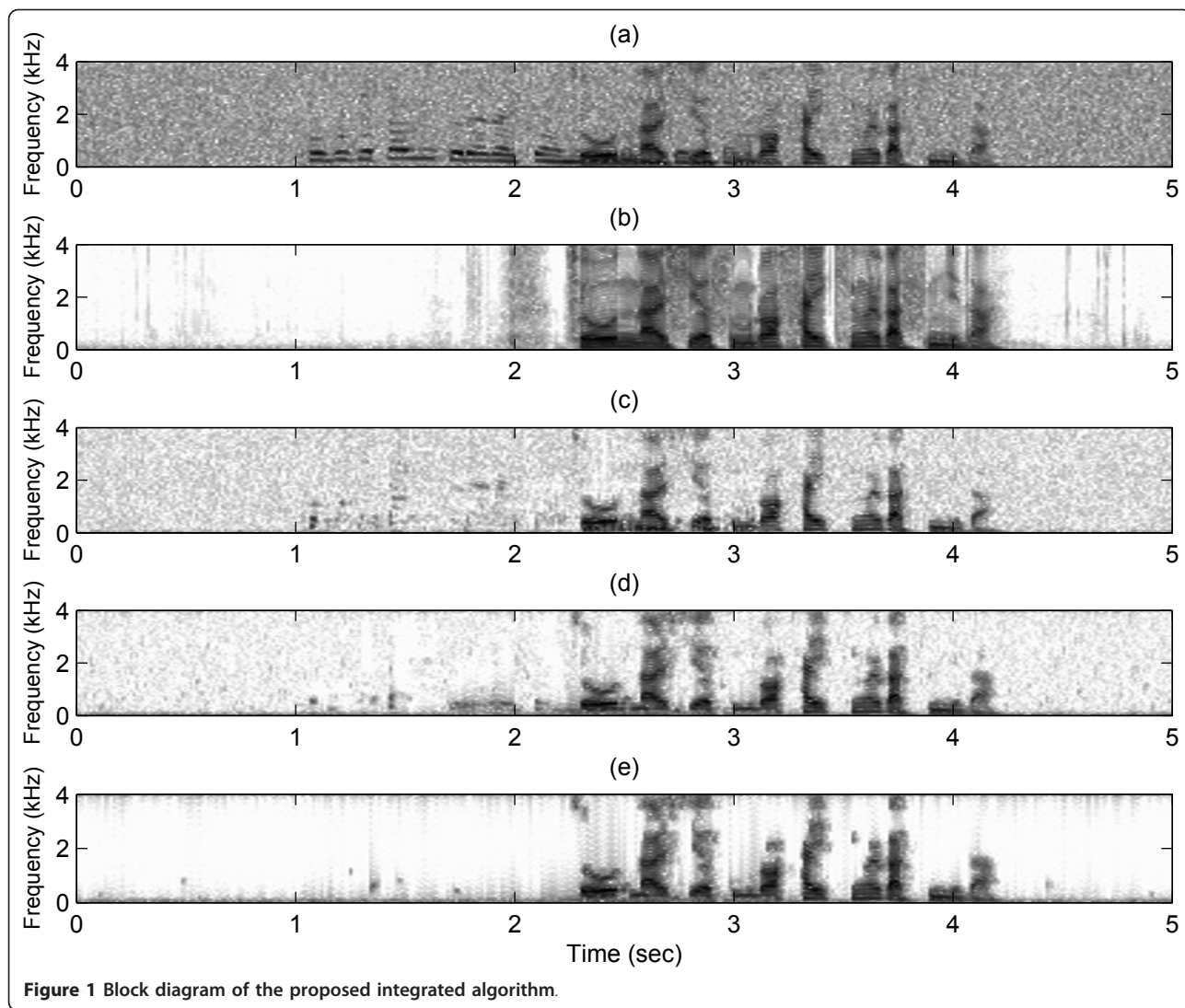


Figure 1 Block diagram of the proposed integrated algorithm.

where $\lambda_{cb}(i, k)$ denotes the combined power of the echo and noise to simultaneously suppress, which should be estimated carefully. Also, $\zeta(i, k)$ is estimated with the help of the well-known decision-directed (DD) approach [10]. Then

$$\hat{\xi}(i, k) = \alpha_{DD} \frac{|\hat{S}(i-1, k)|^2}{\hat{\lambda}_{cd}(i-1, k)} + (1 - \alpha_{DD})P[\gamma(i, k) - 1] \quad (7)$$

where α_{DD} is a weight and $P[z] = z$ if $z \geq 0$, and $P[z] = 0$ otherwise. Also, $\hat{S}(i-1, k)$ is a k th frequency estimate of the near-end speech at the previous frame, and $\hat{\lambda}_{cd}(i, k)$ is the estimate for $\lambda_{cb}(i, k)$.

For $\hat{\lambda}_{cd}(i, k)$, we first estimate the power of the echo signal when the near-end speech signal is not present in the observation (single-talk), as given by

$$\hat{\lambda}_e(i, k) = \alpha_{\lambda_e} \hat{\lambda}_e(i-1, k) + (1 - \alpha_{\lambda_e})|\hat{E}(i, k)|^2 \quad (8)$$

where α_{λ_e} is a smoothing parameter. Note that noise is not taken into account in this update scheme, since it is assumed that the echo is not correlated with the noise and the power of the echo signal is more dominant than the noise power. The estimated magnitude spectrum of echo $|\hat{E}(i, k)|$ is given by

$$|\hat{E}(i, k)| = H(i, k)|X_d(i, k)| \quad (9)$$

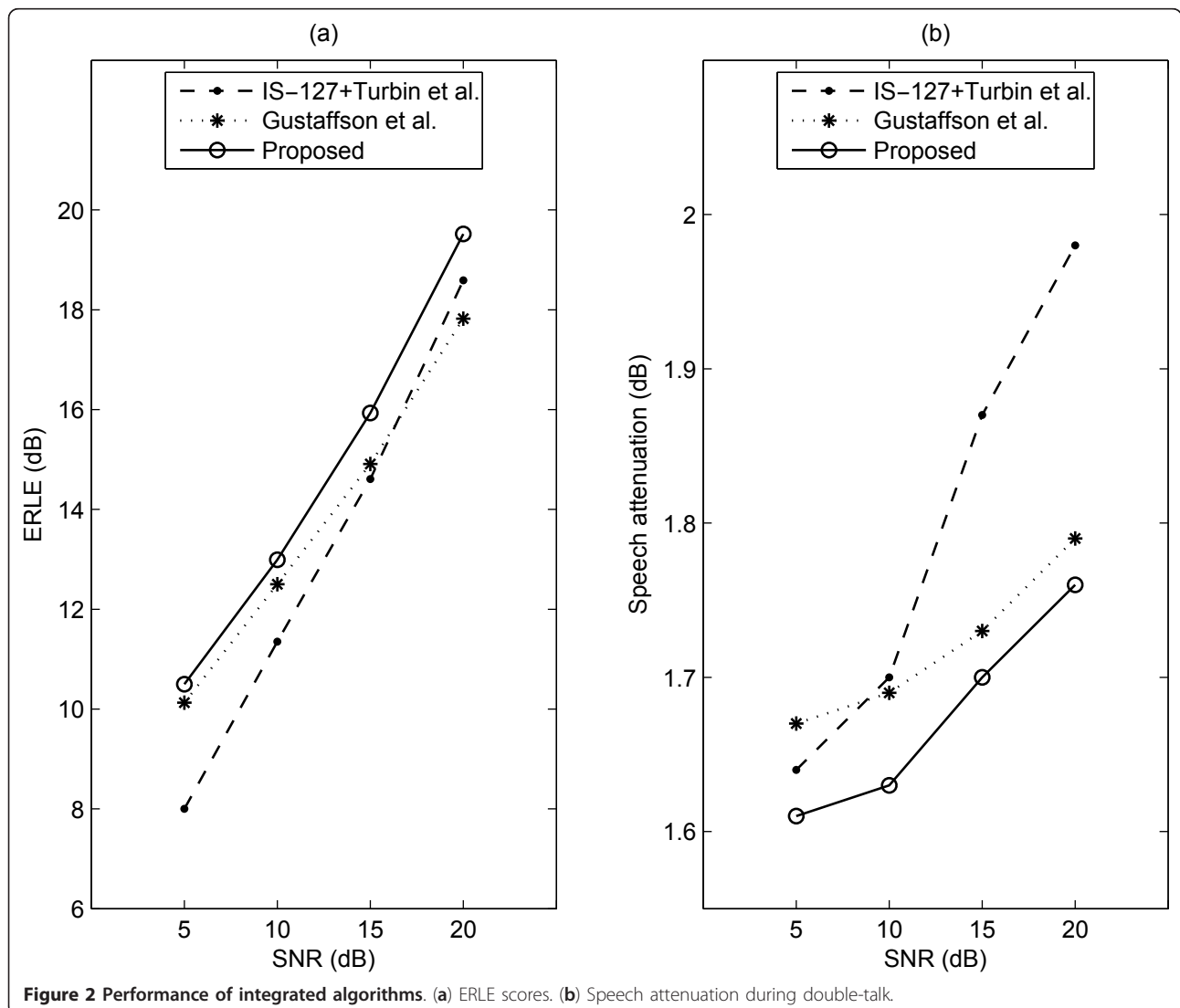
with the far-end speech signal $X_d(i, k)$ and the gain filter $H(i, k)$ characterizing the response of the echo path that is achieved by the magnitude of the least squares estimator [9]

$$H(i, k) = \left| \frac{E[X_d^*(i, k)Y(i, k)]}{E[X_d^*(i, k)X_d(i, k)]} \right| \quad (10)$$

where $*$ denotes the complex conjugate and d indicates d samples delay. Since the echo path is time varying, $H(i, k)$ is estimated iteratively as in [10]. Note that, since $Y(i, k)$ is not affected by the NR algorithm, the estimate of the echo path response does not suffer from the non-linear distortion by the NR operation. And the update of the estimate $H(i, k)$ should be frozen during the double-talk periods to prevent the divergence of $H(i, k)$. To detect a double-talk period, the cross-correlation coefficients-based double-talk detection method proposed by [4] in the frequency domain is implemented. More specifically, (1) the cross-correlation coefficient between the microphone input and the estimate echo, and (2) the cross-correlation coefficient between microphone input and the residual error of the suppressor are computed and used to detect double-talk periods on each frame.

Based on the estimated echo power, we propose the combined power incorporating both the echo power and the background noise power. This is clearly different from the previous approach in [10] in that the method of [10] does not substantially estimate and include the background noise power because of the difficulty in estimating the noise power after the AES algorithm as explained in the first paragraph of Section 2. Specifically, the combined power $\lambda_{cb}(i, k)$ is estimated by assuming that the acoustic echo and noise are uncorrelated and then combining the estimated echo and noise power based on the long-term smoothing scheme with a parameter $\alpha_{\lambda_{cb}}$ such that

$$\hat{\lambda}_{cd}(i, k) = \alpha_{\lambda_{cd}} \hat{\lambda}_{cd}(i - 1, k) + (1 - \alpha_{\lambda_{cd}}) \{\hat{\lambda}_e(i, k) + E[|D(i, k)|^2 | Y(i, k)]\} \quad (11)$$



where $\hat{\lambda}_e(i, k)$ is derived as in (8).

Actually, notice that if $E[|D(i, k)|^2|Y(i, k)] \cong 0$, (11) becomes the original AES algorithm as in [10], while (11) results in the conventional NR algorithm in case that $\hat{\lambda}_e(i, k)$ is nearly zero. Actually, the noise power estimate $E[|D(i, k)|^2|Y(i, k)]$ is obtained during noise-only periods, which is achieved by the voice activity detection (VAD) algorithm that is a similar method as in IS-127 noise reduction algorithm known to give robust performance under various noise conditions [11]. For this reason, we can avoid the disturbed estimate of the noise power incurred by the AES algorithm. Note that since both $e(t)$ and $s(t)$ have a role as a dominant speech, the additional VAD to detect the noise signal periods is needed at the near-end. In addition, the proposed integrated algorithm is further

improved in that distinct values of q 's in (4) are estimated for different frames and frequency bins such as $q(i, k)$ that can be tracked in time [12]. Therefore, the proposed algorithm employs a decision rule to decide whether the near-end speech signal is present in the k th bin, as given by

$$q(i, k) = \alpha_q q(i - 1, k) + (1 - \alpha_q) I(i, k) \quad (12)$$

in which the smoothing parameter α_q is set as 0.3 and $I(i, k)$ denotes an indicator function for the result in (6), that is, $I(i, k) = 1$ if $\eta(i, k) > \eta_{th}$ and $I(i, k) = 0$ otherwise. The value of $q(i, k)$ can be easily updated using the $\eta(i, k)$ as $\eta(i, k) \underset{\hat{H}_0}{\overset{\hat{H}_1}{\geq}} \eta_{th}$ where the threshold η_{th} is set to 5.0 considering the desired significance level.

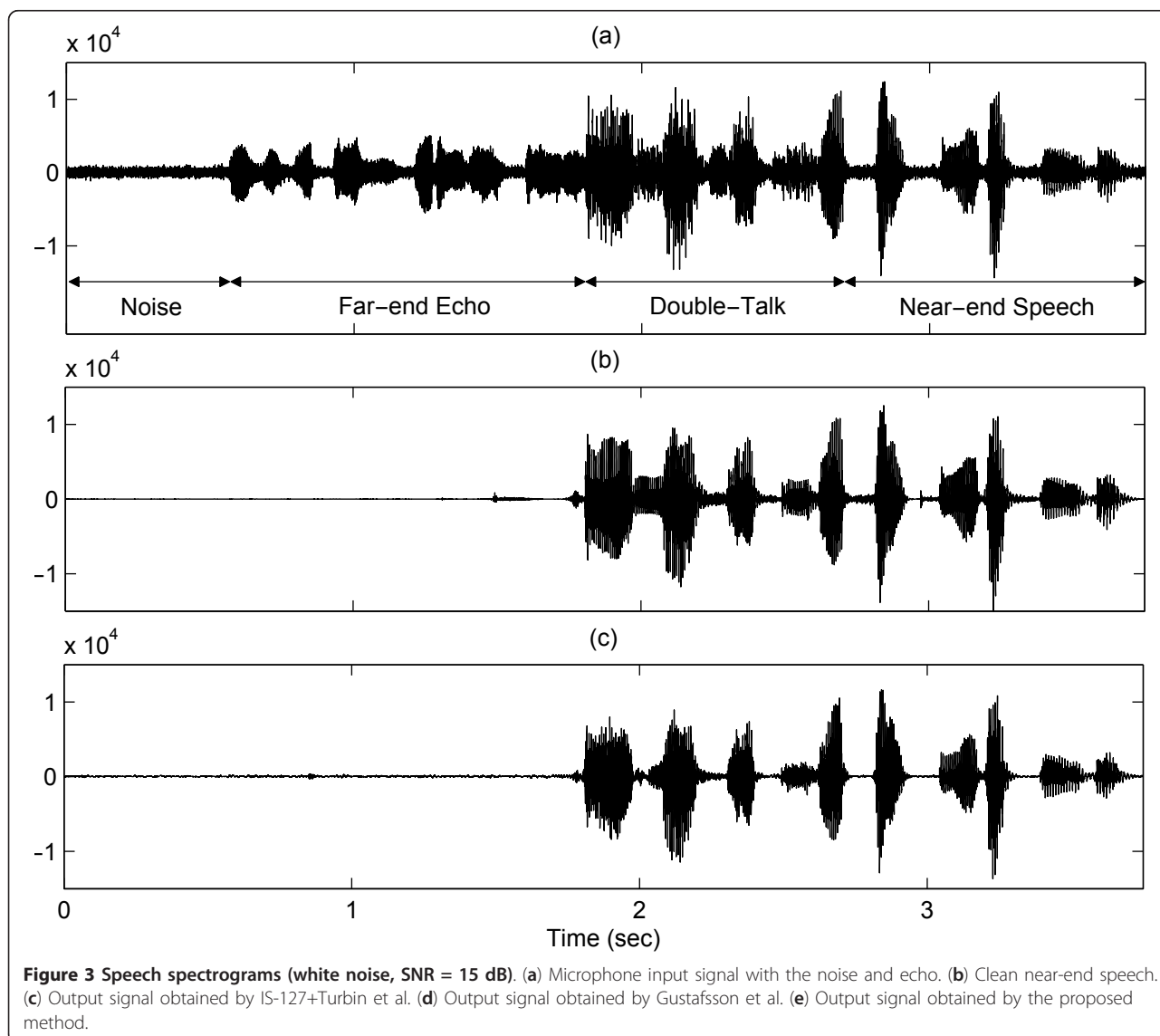


Figure 3 Speech spectrograms (white noise, SNR = 15 dB). (a) Microphone input signal with the noise and echo. (b) Clean near-end speech. (c) Output signal obtained by IS-127+Turbin et al. (d) Output signal obtained by Gustafsson et al. (e) Output signal obtained by the proposed method.

Finally, the estimated near-end speech $\hat{S}(i, k)$ for the echo and noise to be suppressed can be expressed as

$$\hat{S}(i, k) = (1 - p(H_0|Y(i, k))) G(i, k) Y(i, k) = \tilde{G}(i, k) Y(i, k) \quad (13)$$

where $p(H_0|Y(i, k))$, $G(i, k)$ and $\tilde{G}(i, k)$ are the NSAP in (4), suppression gain and overall suppression gain for the integrated system, respectively. Here, $G(i, k)$ for each frequency band is derived from the Wiener filter such that

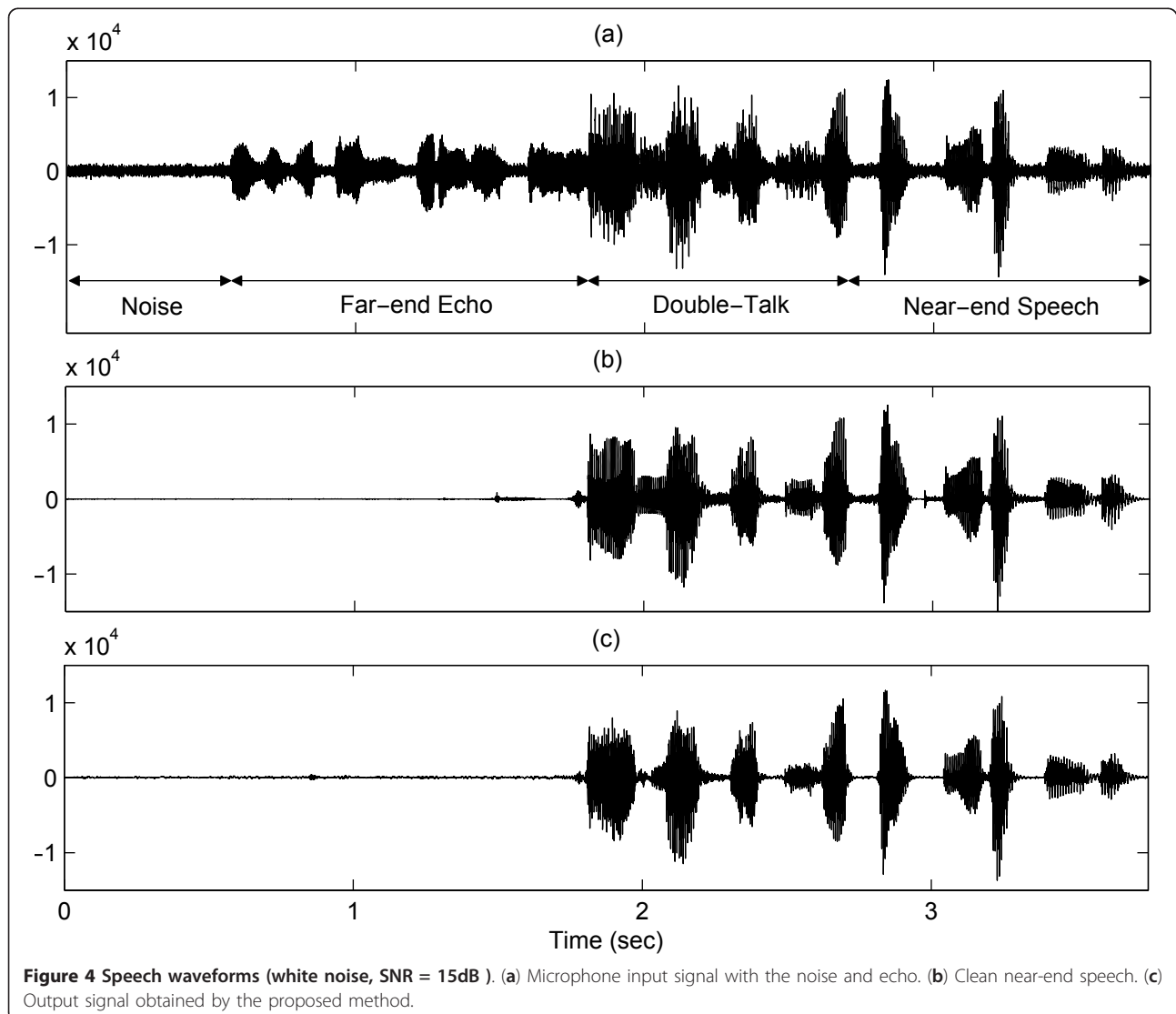
$$G(i, k) = \frac{\hat{\xi}(i, k)}{1 + \hat{\xi}(i, k)}. \quad (14)$$

Notice that a better echo and noise suppression rule through $\tilde{G}(i, k)$ is formulated to apply higher attenuation using $(1 - p(H_0|Y(i, k)))$ consisting of echo or noise

(or both) alone while preserving the quality of the near-end speech.

3 Experiments and results

In order to compare the performance of the proposed integrated algorithm compared with the conventional methods, we conducted a quantitative comparison and subjective quality test under various noise conditions. Twenty test phrases, spoken by seven speakers and sampled at 8 kHz, were used as the experimental data. For assessing the performance of the proposed method, we artificially created 20 data files, where each file was obtained by mixing the far-end signal with the near-end signal. Each frame of the windowed signal was transformed into its corresponding spectrum through 128-point DFT after zero padding. We then achieved 16 frequency sub-bands to entirely cover full frequency



ranges (~4 kHz) of the narrow band speech signal, which is analogous to that of the IS-127 noise suppression algorithm [11]. The far-end speech signal was convolved with a filter simulating the acoustic echo path before being mixed [13,14]. The simulation environment was designed to fit a small office room having a size of $5 \times 4 \times 3 \text{ m}^3$. The length of the simulated acoustic impulse response corresponds to 1,400 tap with the reverberation time $T_{60} = 0.14 \text{ s}$. The echo level measured at the input microphone was 3.5 dB lower than that of the input near-end speech on average. In order to create noisy conditions, white, babble, and vehicular noises from the NOISEX-92 database were added to clean near-end speech signals at signal-to-noise ratios (SNRs) of 5, 10, 15, and 20dB. For the purpose of an objective comparison, we evaluated the performance of the proposed scheme and that of the conventional integrated algorithm. The performance of the approach was measured in terms of echo return loss enhancement (ERLE) and speech attenuation (SA), which are defined in [13].

To see the performance of the conventional integrated algorithm for comparison, we also evaluated the performance of the conventional acoustic echo and noise suppression algorithm by Gustafsson et al. [3],^a which is a serial algorithm on the basis of a time-domain AEC and an additional noise and residual echo reduction filter. Also, we included the other integrated system in which the NR algorithm, that is, IS-127 noise suppression [11] is followed by the AEC with the post-filter as in [15]. For the AEC, a normalized least mean square (NLMS) adaptive filter with the number of filter taps, $L = 128$, was used, because we consider the used DFT size (i.e., 128) in our AES approach in terms of the computational complexity. Given noise environments, overall results for the aforementioned 20 data files are shown in Figure 2. ERLE and SAs

scores were averaged to yield final mean score results for the case of three types of noise sources. From Figure 2a, it is evident that in most noisy conditions, the proposed integrated algorithm based on soft decision yielded a higher ERLE compared to the conventional techniques. This means that the proposed method effectively suppresses both the acoustic echo and noise signal. The SAs of the proposed method during double-talk periods are shown in Figure 2b, where we can observe that the SAs of the proposed scheme were better than that of the methods by Gustafsson et al. and Turbin et al. in all the tested conditions. This phenomenon indicates that the proposed algorithm preserves the near-end talk signal well during the double-talk periods. Also, the speech spectrograms are presented in Figure 3. From Figure 3e yielded by the proposed method, the residual echo and background noise are further reduced compared to the conventional techniques (Figure 3c and 3d) during the active far-end speech and noise period while preserving the near-end speech quite well. In addition, Figure 4 illustrates the speech segments that are results of the proposed algorithm. When we see the double-talk periods carefully, it can be easily seen that the enhanced output signal is successfully obtained even during the double-talk periods.

Finally, in order to evaluate the subjective quality of the proposed algorithm in terms of the distortion of the near-end speech and the residual echo, we carried out a set of informal listening tests. Opinion scores were, respectively, recorded by eleven listeners, and all the scores from the listeners were then averaged to yield final mean opinion score (MOS) results. Eleven listeners (6 men and 5 women) whose ages ranged from 20 to 35 participated in the experiment. Eight of them were students specialized in signal processing, while the others were not specialist. Ten test phrases,

Table 1 Comparison of MOS results (with 95% confidence interval)

Environments		MOS		
Noise	SNR (dB)	IS-127+Turbin et al.	Gustafsson et al.	Proposed
White	5	1.10 ± 0.14	1.35 ± 0.23	1.50 ± 0.36
	10	1.45 ± 0.24	1.90 ± 0.40	2.40 ± 0.47
	15	1.95 ± 0.39	2.70 ± 0.38	2.75 ± 0.43
	20	1.85 ± 0.38	2.80 ± 0.39	3.10 ± 0.50
Babble	5	1.20 ± 0.24	1.15 ± 0.17	1.35 ± 0.27
	10	1.40 ± 0.24	1.45 ± 0.28	1.50 ± 0.24
	15	1.55 ± 0.24	2.10 ± 0.30	2.10 ± 0.30
	20	2.25 ± 0.30	2.40 ± 0.28	2.45 ± 0.24
Vehicle	5	2.15 ± 0.31	3.10 ± 0.40	3.25 ± 0.48
	10	2.25 ± 0.21	3.20 ± 0.24	3.40 ± 0.35
	15	2.35 ± 0.27	3.20 ± 0.24	3.25 ± 0.30
	20	2.45 ± 0.36	3.40 ± 0.38	3.50 ± 0.39

Table 2 Comparison of noise rating scale results (with 95% confidence interval)

Environments		Noise rating scale		
Noise	SNR (dB)	IS-127+Turbin et al.	Gustafsson et al.	Proposed
White	5	1.40 ± 0.24	1.65 ± 0.35	2.55 ± 0.36
	10	1.45 ± 0.24	2.20 ± 0.62	2.60 ± 0.53
	15	1.85 ± 0.51	2.75 ± 0.30	2.80 ± 0.54
	20	2.35 ± 0.51	3.20 ± 0.45	3.20 ± 0.39
Babble	5	1.20 ± 0.24	1.20 ± 0.19	1.30 ± 0.22
	10	1.60 ± 0.38	1.65 ± 0.41	1.75 ± 0.37
	15	1.95 ± 0.39	1.90 ± 0.37	2.15 ± 0.46
	20	2.30 ± 0.40	2.45 ± 0.32	2.45 ± 0.32
Vehicle	5	1.95 ± 0.32	3.20 ± 0.29	3.60 ± 0.53
	10	2.10 ± 0.26	3.35 ± 0.27	3.55 ± 0.28
	15	2.00 ± 0.30	3.40 ± 0.28	3.45 ± 0.28
	20	2.10 ± 0.40	3.40 ± 0.32	3.60 ± 0.35

where five were spoken by a male speaker and the other were spoken by a female speaker, were used as the experimental data. Each phrase consisted of the two different meaningful sentences and lasted 8s as suggested in [16]

Table 1 illustrates that the proposed approach outperformed or at least was comparable to the conventional methods in terms of overall subjective quality under the given noise conditions. In addition, we separately checked the performance of noise reduction which is one of the major goals in this work, which was achieved by the ITU-T P.835 [16], that is, the subjective quality test in terms of the background noise rating scale (5: not noticeable, 4: slightly noticeable, 3: noticeable but not intrusive, 2: somewhat intrusive, 1: very intrusive) in a similar manner as in the previous MOS test. As Table 2 shows, the performance improvement was found for all cases at all SNRs. These results confirm that the proposed integrated system is effective in suppressing the background noise.

4 Conclusions

In this paper, we have proposed a novel integrated suppression algorithm based on soft decision using the combined power of the estimated echo and noise power. The principal contribution of this study is that the proposed method can efficiently suppress the acoustic echo and noise signal through the suppression gain based on soft decision without the help of an additional residual echo and noise suppressor. The performance of the proposed algorithm has been found to be superior to that of the conventional technique. Future study areas may include the other superior statistical models characterizing the input signals such as the Laplacian and gamma as in [17], even though the Gaussian model can lead to more tractable mathematics.

Endnotes

^aFor [3], we set T_n to 0.05 where T_n denotes a minimum threshold.

Acknowledgements

This work was supported by the IT R&D program of MKE/KEIT. [2009-S-036-01, Development of New Virtual Machine Specification and Technology], by National Research Foundation of Korea(NRF) grant funded by the Korean Government(MEST) (NRF-2011-0009182), and by the research fund of Hanyang University (HY-2011-201110000000210). Note: Please send all correspondence related with this manuscript to Prof. J.-H. Chang at the address below.

Author details

¹School of Electronic Engineering, Inha University, Incheon 402-751, Korea
²School of Electronic Engineering, Hanyang University, Seoul 133-791, Korea

Competing interests

The authors declare that they have no competing interests.

Received: 19 May 2011 Accepted: 17 January 2012

Published: 17 January 2012

References

1. H Puder, P Dreiseitel, Implementation of a hands-free car phone with echo cancellation and noise-dependent loss control. *Proc IEEE Int Conf Acoust Speech Signal Process.* **6**, 3622–3625 (2000)
2. P Dreiseitel, E Hänslér, H Puder, Acoustic echo and noise control—a long lasting challenge. *Proc EUSIPCO 945–952* (Sep. 1998)
3. S Gustafsson, R Martin, P Vary, Combined acoustic echo control and noise reduction for hands-free telephony. *Signal Process.* **64**(1), 21–32 (1998). doi:10.1016/S0165-1684(97)00173-4
4. SJ Park, CG Cho, C Lee, DH Youn, Integrated echo and noise canceler for hands-free applications. *IEEE Trans Circuits Syst II.* **49**(3), 186–195 (2002)
5. Y Guelou, A Benamar, P Scalart, Analysis of two structures for combined acoustic echo cancellation and noise reduction, in *Proc IEEE Int Conf Acoust Speech Signal Process.* **2**, 637–640 (1996)
6. S Gustafsson, R Martin, P Jax, P Vary, A psychoacoustic approach to combined acoustic echo cancellation and noise reduction. *IEEE Trans Speech Audio Process.* **10**(5), 245–256 (2002). doi:10.1109/TSA.2002.800553
7. E Habets, I Cohen, S Gannot, MMSE log-spectral amplitude estimator for multiple interferences, in *Proc Int Workshop Acoust Echo Noise Control, IWAENC'06*, (Paris, France, Sept. 2006)
8. E Habets, S Gannot, I Cohen, P Sommen, Joint dereverberation and residual echo suppression of speech signals in noisy environments. *IEEE Trans Audio Speech Lang Process.* **16**(8), 1433–1451 (2008)

9. C Faller, C Tournery, Estimating the delay and coloration effect of the acoustic echo path for low complexity echo suppression. in *Proc Intl Works on Acoust Echo and Noise Control (IWAENC)* 53–56 (Oct. 2005)
10. YS Park, JH Chang, Frequency domain acoustic echo suppression based on soft decision. *IEEE Signal Process Lett.* **161**, 53–56 (2009)
11. TIA/EIA/IS-127, Enhanced variable rate codec, speech service option 3 for wideband spread spectrum digital systems (1996)
12. D Malah, R Cox, A Accardi, Tracking speech-presence uncertainty to improve speech enhancement in non-stationary noise environments. in *Proc IEEE Int Conf Acoust Speech Signal Process* 789–792 (1999)
13. SY Lee, NS Kim, A statistical model based residual echo suppression. *IEEE Signal Process Lett.* **14**(10), 758–761 (2007)
14. S McGovern, A Model for Room Acoustics, 2003. <http://sgm-audio.com/research/rir/rir.html>
15. V Turbin, A Gilloire, P Scalart, Comparison of three post-filtering algorithms for residual acoustic echo reduction. in *Proc IEEE Int Conf Acoust Speech Signal Process* 307–310 (1997)
16. ITU-T Recommendation P.835, Subjective test methodology for evaluating speech communication systems that include noise suppression algorithm (Nov. 2003)
17. JH Chang, S Gazor, NS Kim, SK Mitra, Voice activity detection based on multiple statistical models. *IEEE Trans Signal Process.* **54**(6), 1965–1976 (2006)

doi:10.1186/1687-6180-2012-11

Cite this article as: Park and Chang: **Integrated acoustic echo and background noise suppression technique based on soft decision.** *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:11.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
