

논문 2009-46SP-5-19

주파수 영역에서 구간조건을 이용한 음향학적 반향 제거

(Frequency Domain Acoustic Echo Suppression Based on Boundary Condition)

이규호*, 장준혁**

(Kyu-Ho Lee and Joon-Hyuk Chang)

요약

본 논문은 적응 반향 제거기인 parametric wiener filter (PWF)의 반향신호 제거 파라미터를 구간 조건에 따라 차등적으로 적용시키는 새로운 음향학적 반향제거 (Acoustic Echo Suppression, AES) 알고리즘을 제안한다. PWF는 반향 신호 추정의 부정확성을 보상하고자 반향신호 제거 파라미터를 사용한다. 기존의 방법은 단일통화 구간과 동시통화 구간에 관계없이 동일한 반향신호 제거 파라미터 값을 적용하여 동시통화 환경에서 근단 신호의 왜곡을 발생 시킨다. 본 논문에서는 상관계수 (Correlation Coefficient)를 이용한 동시통화(Double-Talk) 검출 알고리즘과 음성구간검출기 (Voice Activity Detector, VAD)의 결과로부터, 원단 화자만 있는 구간과 동시통화 구간을 구분하고 각각의 경우에 반향제거 파라미터를 달리 적용하여 동일한 반향신호 제거 파라미터를 사용함에 따른 동시통화 구간내의 근단 화자신호의 음질 저하 없이 반향신호를 보다 효과적으로 제거하는 방법을 제안한다. 제안된 방법의 결과는 객관적인 실험을 통해 비교 평가한 결과 우수한 성능을 보였다.

Abstract

In this paper, we propose a novel approach of an acoustic echo cancellation (AEC) algorithm which is differently adopted in the relevant period condition by the suppression parameter of a parametric wiener filter (PWF). The PWF uses the suppression parameter to compensate uncertainty of acoustic echo signal estimation. The existing PWF method using the fixed suppression parameter derives the distortion of the near-end signal at the double-talk. To solve this problem, the boundary condition is devised using decision of the double-talk detection (DTD) algorithm and voice activity detector (VAD). The boundary condition makes it possible to treat differently depending on the case of the single-talk and double-talk. According to the experimental results, the proposed approach is found to be effective for acoustic echo cancellation using the boundary condition.

Keywords: Acoustic Echo Suppression(AES), Parametric Wiener Filter (PWF), Double-Talk Detection(DTD), Voice Activity Detector (VAD).

I. 서론

정보화 사회의 발전과 더불어 빠르게 진행되고 있는

통신수단의 발전은 이동 환경에 좀 더 편리한 통화환경 조성을 위해 송수기 (Hand-Set)를 사용하지 않는 핸드프리 (Hands-Free)통신방식을 지향하고 있다. 즉, 이동환경의 편리성을 위한 단말기의 소형화가 급속히 이루어지고 있으며, 이에 따라 스피커로 출력된 음성이 벽이나 물건, 또는 사람의 피부 등에 반사되어 마이크의 입력으로 들어가게 되는 음향 반향 신호에 대한 문제가 중요하게 대두되고 있다. 전통적인 음향 반향 신호 제거 방법은 적응 반향제거기 (Acoustic Echo Canceler, AEC)를 이용하는 것이다^[1~2]. 반향제거기는 반향 신호를 추정하고, 입력신호에서 추정된 반향신호

* 학생회원, ** 정회원, 인하대학교 전자공학부
(Department of Electronics Engineering,
Inha University)

※ 본 연구는 지식경제부와 한국산업기술재단의 전력 기술인력양성사업으로 수행된 연구결과이며 또한 본 연구는 지식경제부 출연금으로 ETRI, SoC산업진흥센터에서 수행한 IT SoC 핵심설계인력양성사업의 연구결과입니다.

접수일자: 2009년2월16일, 수정완료일: 2009년8월27일

를 빼줌으로서 반향신호를 제거한다. 일반적으로 AEC를 이용한 반향제거에는 상시적인 혹은 급작스런 반향경로의 변화에 따른 잔여 반향신호를 제거하기 위한 반향 억제기 (Acoustic Echo Suppression, AES)가 함께 사용된다. 그러나 최근에는 AEC를 적용하지 않고 AES를 이용하여 모든 반향 신호를 제거하는 방식이 제안되어 왔다^[3~4]. 이와 같은 방식은 [5]에서 parametric wiener filter (PWF)를 AES로 사용한 방식으로 보고되었다. PWF는 추정된 반향 신호의 오차를 보상할 수 있는 반향신호 제거 파라미터를 갖는다. 이때, 원단화자만 존재하는 단일통화 구간과 원단화자와 근단 화자가 동시에 존재하는 동시통화 구간에 대한 구분 없이 동일한 반향신호 제거 파라미터를 적용하여 반향 신호 제거 효과를 높여야 할 경우, 원단 화자만 있는 구간의 반향 신호 제거에는 효과적 이지만 동시통화 구간내의 근단 화자 신호에 대한 음질 저하를 유발하게 되는 문제점이 있다. 본 논문은 [6]에서 제안된 통계적 모델 기반의 우도비 (likelihood ratio test, LRT) 시험으로 부터의 음성구간검출기 (Voice Activity Detector, VAD)와 [7]에서 제안된 마이크 입력신호와 추정된 반향신호, 적응 반향 제거기 사이의 각각 두 가지 상호 상관계수를 적절한 문턱값과 비교하여 동시통화 구간을 검출하는 동시통화 (Double-Talk) 검출 알고리즘의 결과로부터, 원단 화자만 존재하는 단일통화 구간과 동시통화 구간을 구분하고 각각의 경우에 반향신호 제거 파라미터를 달리 적용하여 모든 구간에 동일한 반향신호 제거 파라미터를 적용함에 따른 문제점을 해결하는 방법을 제안한다. 제안된 방법은 ERLE (Echo Return Loss Enhancement)를 측정하여 평가하였으며, 고정된 반향신호 제거 파라미터를 사용한 경우보다 우수한 결과를 나타내었다. 본 논문의 구성은 다음과 같다. II장에서 PWF에 고정된 반향신호 제거 파라미터를 적용할 때 발생하는 문제점에 대해 언급하고 III장에서는 이에 대한 해결 방안을 제시한다. IV장장에서는 실험을 통해 제안한 알고리즘의 성능을 평가하고 마지막으로 V장에서 결론을 맺는다.

II. 고정 suppression parameter를 적용한 경우의 PWF의 문제점

마이크로폰으로 입력되는 음향학적 반향신호는 다양한 반사 경로를 거쳐 입력되기 때문에 원단 신호로부터

반사 경로를 고려한 임펄스응답에 대한 추정 과정이 필요하다. 음향학적 반향 제거기에서 반향신호 $d(t)$, 배경 잡음 $w(t)$, 근단 화자신호 $s(t)$, 원단신호와 마이크 입력신호를 각각 $x(t)$, $y(t)$ 라 하고 $Y(i,k)$ 와 $S(i,k)$ 은 각각 $y(t)$ 와 $s(t)$ 의 i 번째 프레임의 k 번째 주파수 성분을 나타낸다. 이때, 반향 경로의 임펄스응답을 고려하여 d 샘플만큼 지연된 원단신호의 주파수 성분 $X_d(i,k)$ 부터 추정된 반향신호 $\hat{Y}(i,k)$ 는 다음과 같이 나타낼 수 있다^[5].

$$|\hat{Y}(i,k)| = G_V(i,k)|X_d(i,k)| \tag{1}$$

여기서 원단신호로부터 음향학적 반향신호를 추정하기 위한 최소 자승 오차 추정 기반의 이득 필터 $G_V(i,k)$ 는 다음의 식과 같다^[5].

$$G_V(i,k) = \left| \frac{E\{X_d^*(i,k)Y(i,k)\}}{E\{X_d^*(i,k)X_d(i,k)\}} \right| \tag{2}$$

여기서 *는 complex conjugate를 의미한다. 반향신호는 다양한 반향 경로를 거쳐 마이크로폰으로 입력되기 때문에 수시로 변화하는 반향 경로에 대한 영향을 줄이기 위해 다음식과 같이 long-term smoothing을 적용한다.

$$G_V(i,k) = \frac{a_{12}(i,k)}{a_{22}(i,k)} \tag{3}$$

$$a_{12}(i,k) = \epsilon a_{12}(i,k-1) + (1-\epsilon)|X_d^*(i,k)Y(i,k)| \tag{4}$$

$$a_{22}(i,k) = \epsilon a_{22}(i,k-1) + (1-\epsilon)X_d^*(i,k)X_d(i,k) \tag{5}$$

여기서 ϵ 는 $0 \leq \epsilon \leq 1$ 인 가중치 파라미터로 0.998을 사용한다. 식 (1)에서 추정된 반향신호는 AES인 PWF의 wiener 이득 필터에 적용되며 다음과 같이 표현된다^[5].

$$G(i,k) = \left[\frac{\max(|Y(i,k)|^\alpha - \beta|\hat{Y}(i,k)|^\alpha, 0)}{|Y(i,k)|^\alpha} \right]^{\frac{1}{\alpha}} \tag{6}$$

여기서 α 와 β 는 반향 제거를 제어하기 위한 설계 parameter들이다. 마지막으로, 식 (6)에 의하여 마이크 입력신호 $y(t)$ 에서 추정된 음향학적 반향신호가 제거된 출력신호를 $e(t)$ 라 하면 다음과 같이 나타낼 수 있다^[5].

$$e(t) = F^{-1}[G(i,k)Y(i,k)] \tag{7}$$

여기서 $F^{-1}[\cdot]$ 은 역 (Inverse) 이산 푸리에 변환 (discrete Fourier transform, DFT) 연산자를 의미한다.

식 (6)과 (7)에 의해 표현된 PWF에 의한 반향 제거 과정에서, 추정된 반향신호의 오차를 보상하는 반향신호 제거 파라미터 (β)값에 따라 wiener 이득 필터가 결정되고, 이는 다시 출력신호 $e(t)$ 에 영향을 미치게 됨을 볼 수 있다. 이때, 구간조건 없이 반향 신호에 대한 보다 효과적인 제거를 위해 반향신호 제거 파라미터 값을 증가 시킬 경우 원단 화자만 존재하는 단일통화 구간에서는 추정된 반향 신호에 대해 효과적인 제거를 기대 할 수 있지만 동시통화 구간에서의 반향 신호에 대한 잘못된 추정의 결과를 증폭하여 근단 화자 신호의 음질이 떨어지게 될 수 있는 위험성이 있다.

III. 제안된 구간 조건에 따른 차등적 suppression parameter 적용

지금까지 우리는 PWF를 이용한 반향제거에서 원단 화자 신호만 존재하는 단일통화 구간과 동시통화 구간에 대한 구분 없이 동일한 반향신호 제거 파라미터 값을 사용하여 반향 제거 효과를 높이려 할 경우의 문제점에 대해서 알아보았다. 이러한 문제를 해결하기 위해 차등적 반향신호 제거 파라미터를 적용하기 위한 구간 조건이 요구된다. 구간조건은 원단화자 신호에 대해, 동시통화 구간을 발생시키는 근단 화자 신호가 존재하지 않을 경우와 존재할 경우 각각에 대한 가정 H_0 , H_1 으로 다음과 같이 표현할 수 있다.

H_0 : Near-end speech absent: $Y(i, k) = \hat{Y}(i, k)$

H_1 : Near-end speech present:

$$Y(i, k) = \hat{Y}(i, k) + S(i, k) \tag{8}$$

여기서 추정된 반향신호 $\hat{Y}(i, k)$ 는 근단 화자 신호인 $S(i, k)$ 와 통계적으로 독립이라고 가정한다. H_1 은 [7]의 상관계수를 이용한 동시통화검출 알고리즘을 이용하여 결정되어지며, H_0 는 음성구간검출기로 부터의 원단화자만 존재하는 구간에 대한 결과에서 H_1 으로 결정된 구간을 제외시킴으로써 구분되어진다. 본 논문에서는 이와 같은 구간 조건을 이용하여 근단 화자 신호의 왜곡에 대한 고려가 필요 없는, 반향신호만 존재하는 구간에 대해서는 반향 제거를 위한 보다 강력한 파라미터를 적용하고 이와 반대로 근단 화자 신호의 왜곡에 대한 고려가 요구되는 동시통화 구간에 대해서는 실험적으로 결정된 적절한 파라미터를 차등적으로 적용한다. 구간 조건으로부터 차

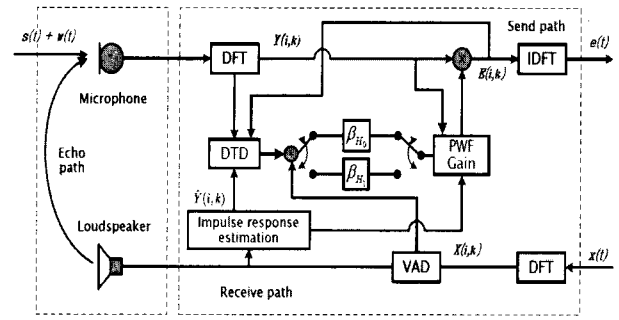


그림 1. 제안된 음향학적 반향 제거기의 블록도
Fig. 1. Block diagram of proposed AES with differential suppression parameter.

등 적용한 PWF의 wiener 이득 필터 식은 다음과 같이 $G_{H_0}(i, k)$, $G_{H_1}(i, k)$ 으로 나누어 표현되어진다.

$$G_{H_0}(i, k) = \left[\frac{\max(|Y(i, k)|^\alpha - (\beta_{H_0})|\hat{Y}(i, k)|^\alpha, 0)}{|Y(i, k)|^\alpha} \right]^{\frac{1}{\alpha}}$$

$$G_{H_1}(i, k) = \left[\frac{\max(|Y(i, k)|^\alpha - (\beta_{H_1})|\hat{Y}(i, k)|^\alpha, 0)}{|Y(i, k)|^\alpha} \right]^{\frac{1}{\alpha}} \tag{10}$$

β_{H_0} 과 β_{H_1} 은 각각 H_0 와 H_1 으로 결정 되었을 때, PWF의 반향신호 제거 파라미터를 의미하며 2.0과 1.0의 값을 갖는다. 이로써 원단화자만 존재하는 구간 (H_0)에서 추정된 반향신호에 대해서는 두 배의 제거 효과가 반영된 wiener 이득 필터가 ($G_{H_0}(i, k)$) 적용되고 동시통화 구간 (H_1)에서는 추정된 반향신호의 결과를 그대로 적용한 wiener 이득 필터가 ($G_{H_1}(i, k)$) 적용되어 동시통화 구간내의 근단 화자 신호에 대한 왜곡 없이 원단화자만 있는 구간에서의 반향신호를 보다 효과적으로 제거할 수 있게 된다. 이와 같이 제안된 차등적 반향신호 제거 파라미터를 적용한 알고리즘은 그림 1의 순서도와 같다.

IV. 실험 결과

본 논문에서는 제안된 알고리즘의 성능 평가를 위해 다양한 잡음 환경에서 객관적인 실험을 수행하였다. 성능 평가는 원단 화자 신호만 있는 구간에서 반향제거 후 여전히 남아 있는 잔여 반향의 제거 정도를 측정하는 식 (11)과 같이 표현되는 ERLE 수치를 비교함으로써 이루어졌다. t 는 샘플 인덱스를 나타낸다.

$$ERLE(t) = 10 \log_{10} \left[\frac{E\{y^2(t)\}}{E\{e^2(t)\}} \right] \tag{11}$$

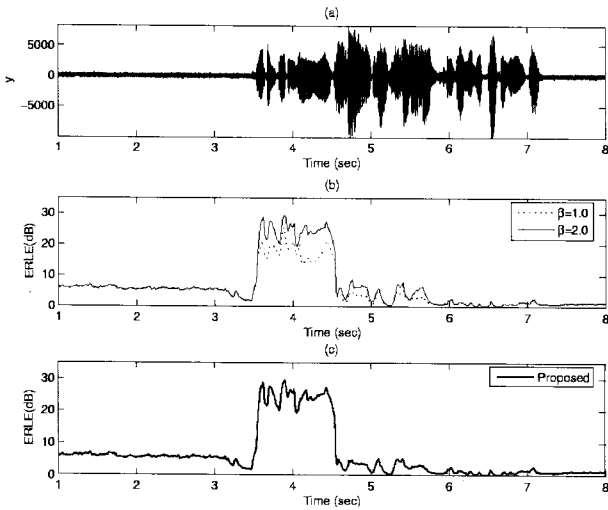


그림 2. (a) white 잡음이 부가된 근단 입력 신호 $y(t)$
 (b) 동일한 suppression parameter를 사용한 경우 ERLE 비교 (c) 제안된 경우의 ERLE

Fig. 2. (a) Near-end speech signal $y(t)$ under the white noise (SNR = 20 dB). (b) ERLE of fixed suppression parameter 2.0 and 1.0. (c) ERLE of proposed algorithm.

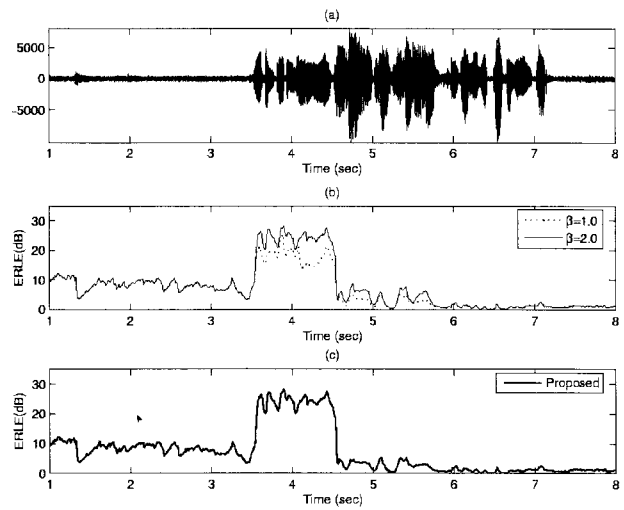


그림 4. (a) office잡음이 부가된 근단 입력 신호 $y(t)$
 (b) 동일한 suppression parameter를 사용한 경우 ERLE 비교 (c) 제안된 경우의 ERLE

Fig. 4. (a) Near-end speech signal $y(t)$ under the office noise (SNR = 20 dB). (b) ERLE of fixed suppression parameter 2.0 and 1.0. (c) ERLE of proposed algorithm.

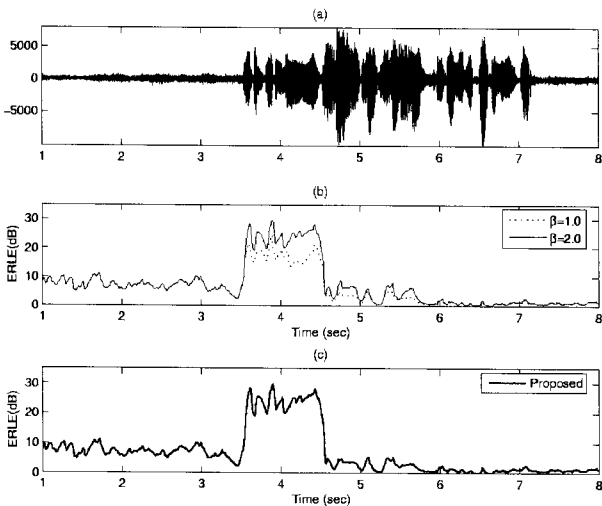


그림 3. (a) babble 잡음이 부가된 근단 입력 신호 $y(t)$
 (b) 동일한 suppression parameter를 사용한 경우 ERLE 비교 (c) 제안된 경우의 ERLE

Fig. 3. (a) Near-end speech signal $y(t)$ under the babble noise (SNR = 20 dB). (b) ERLE of fixed suppression parameter 2.0 and 1.0. (c) ERLE of proposed algorithm.

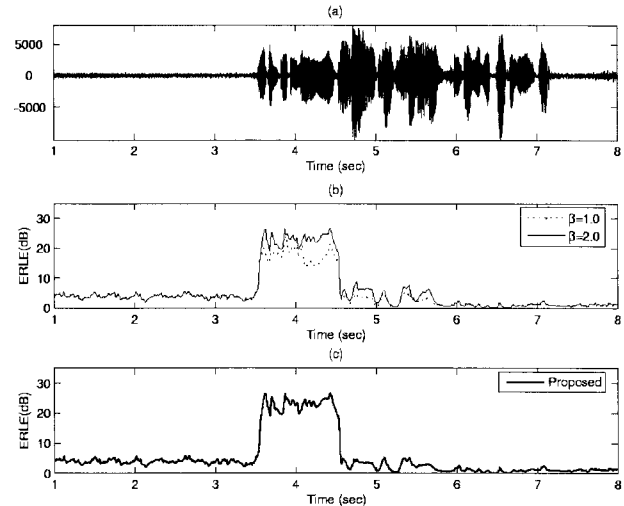


그림 5. (a) street잡음이 부가된 근단 입력 신호 $y(t)$ (b) 동일한 suppression parameter를 사용한 경우 ERLE 비교 (c) 제안된 경우의 ERLE

Fig. 5. (a) Near-end speech signal $y(t)$ under the street noise (SNR = 20 dB). (b) ERLE of fixed suppression parameter 2.0 and 1.0. (c) ERLE of proposed algorithm.

테스트는 10명의 화자로부터 8kHz로 샘플링 된 문장을 수집하여 수행 되었다. 수집된 10명의 화자 중 5명의 화자는 원단 화자 신호로 사용하고 나머지 5명은 근단 화자 신호로 사용하였으며 분류된 두 그룹의 신호를 무작위로 섞어서 동시통화 구간이 약 1초 발생하는 문장을 생성하였다. 원단 화자 신호로 분류된 5명의 화자

에 대해서는 근단 화자 신호와 섞기 전에 실제 반사 환경을 모델링하기 위해 근단 화자 신호보다 3.5 dB 작게 함과 동시에 임펄스 응답 필터를 통과 시켰다^[8-9]. 모델링 환경의 장소는 $5 \times 4 \times 3 m^3$ 크기로 설정 하였다.

그림 2부터 5는 각각 SNR 20 dB로 white, babble, office, street 잡음이 부가된 입력신호에 대하여 반향신

호 제거 파라미터에 따른 ERLE 변화를 보여주고 있다. 적용된 모든 잡음환경에서 반향신호 제거 파라미터가 구간에 관계없이 1.0으로 고정된 값을 사용한 경우에 비해 2.0으로 고정된 값을 사용하였을 경우, 원단화자만 존재하는 구간에서 향상된 ERLE 수치를 볼 수 있지만, 동시 통화 구간에서도 ERLE 수치가 함께 높아지면서 동시통화 구간내의 원단화자에 대한 왜곡이 발생되고 있음을 볼 수 있다. 그러나 본 논문이 제안하는 구간조건에 따라 반향신호 제거 파라미터를 달리 적용한 경우에는 이러한 동시 통화 구간에서의 문제점 없이 원단화자의 반향신호만을 효과적으로 제거시킬 수 있다.

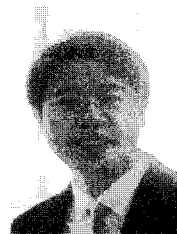
VI. 결 론

본 논문에서는 적응 반향 제거기인 PWF의 반향신호 제거 파라미터를 구간조건에 따라 다른 값을 적용하여 동시통화 구간내의 근단 화자신호의 음질 저하 없이 반향신호를 제거하는 알고리즘을 제안하였다. 구간조건은 상호 상관계수를 이용한 동시통화 검출 알고리즘과 음성구간검출기의 결과를 이용하여 결정 하였다. 제안된 방법은 구간 조건에 대한 고려 없이 동일한 반향신호 제거 파라미터를 사용한 경우와 비교한 결과 보다 향상된 반향제거 효과를 보였다.

참 고 문 헌

- [1] P. S. R. Diniz, Adaptive Filtering: Algorithm and Practical Implementation. Norwell, MA: Kluwer, 1997.
- [2] C. Avendano, "Acoustic echo suppression in the STFT domain," in *Proc. IEEE Workshop on Appl. of Sig. Proc. to Audio and Acoust.*, Oct. 2001.
- [3] C. Faller and J. Chen, "Suppression acoustic echo in a sampled auditory envelope space," *IEEE Trans. on Speech and Audio Proc.*, Aug. 2003.
- [4] C. Faller, "Perceptually motivated low complexity acoustic echo control," in *Preprint 114th Conv. Aud. Eng. Soc.*, Mar. 2003.
- [5] C. Faller and C. Tournery, "Estimating the delay and coloration effect of the acoustic echo path for low complexity echo suppression," in *Proc. IEEE Works. on Acoust. Echo and Noise Control (IWAENC)*, Sept. 2005.
- [6] J. Sohn, N. S. Kim and W. Sung, "Statistical Model-Based Voice Activity Detection," *IEEE Signal Processing Letters*, vol. 6, no. 1, pp. 1-3, Jan. 1999.
- [7] S. J. Park, C. G. Cho, C. Lee, and D. H. Youn, "Integrated echo and noise canceler for hands-free applications," *IEEE Trans. on Circuits and Systems II*, vol. 49, no. 3, pp. 186-195, Mar. 2002.
- [8] S. McGovern, A Model for Room Acoustics, 2003 [Online]. Available: <http://2pi.us/rir.html>
- [9] S. Y. Lee and N. S. Kim, "A statistical model based residual echo suppression," *IEEE Sig. Proc. Letter*, vol. 14, no. 10, pp. 758-761, Oct. 2007.

— 저 자 소 개 —



이 규 호(학생회원)
2008년 인하대학교 전자전기
공학부 학사
2008년~현재 인하대학교
전자공학과 석사과정
<주관심분야 : 디지털신호처리>



장 준 혁(정회원)
1998년 경북대학교 전자공학과
학사.
2000년 서울대학교 전기공학부
석사.
2004년 서울대학교 전기컴퓨터공
학부 박사.
2000년~2005년 (주)넷더스 연구소장
2004년~2005년 캘리포니아 주립대학, 산타바바
라(UCSB) 박사후연구원
2005년 한국과학기술연구원(KIST) 연구원
2005년~현재 인하대학교 전자공학과 조교수
<주관심분야 : 음성 신호처리, 오디오 신호처리,
통신 신호처리, 휴먼/컴퓨터 인터페이스>