

새로운 잡음전력 추정 기법을 적용한 음향학적 반향 및 배경잡음 제거 통합시스템

A New Unified System of Acoustic Echo and Noise Suppression Incorporating a Novel Noise Power Estimation

박 윤 식*, 장 준 혁*

(Yun-Sik Park*, Joon-Hyuk Chang*)

*인하대학교 전자공학부

(접수일자: 2009년 8월 19일; 채택일자: 2009년 10월 6일)

본 논문에서는 주파수 영역에서 음향학적 반향 및 잡음 제거의 통합 시스템을 위한 효과적인 잡음전력 추정 기법을 제안한다. 제안된 방법은 잡음 제거 (NS, noise suppression)가 음향학적 반향 억제 (AES, acoustic echo suppression)의 후처리단으로 결합하여 사용되는 구조에서 발생하는 잡음전력 추정오차를 줄이기 위해 마이크로폰 입력신호의 음성부재확률 (SAP, speech absence probability)을 잡음전력 갱신을 위한 스무딩 (smoothing) 파라미터로 적용한다. 따라서 제안된 기법에서는 반향 억제 후 신호에서 잡음전력 갱신을 위한 SAP를 추출하는 대신 입력신호에 대한 SAP를 NS 알고리즘에 적용함으로써 잡음 제거가 반향 억제 후 왜곡된 잡음 스펙트럼 구간에서는 잡음전력을 갱신하지 않도록 한다. 제안된 알고리즘은 기존의 방법과 객관적인 실험을 통해 비교 평가한 결과 다양한 배경잡음 환경에서 우수한 성능을 보였다.

핵심용어: 음향학적 반향 억제, 잡음 제거, 음성부재확률

투고분야: 음성 처리 분야 (2,3)

In this paper, we propose a efficient noise power estimation technique for an integrated acoustic echo and noise suppression system in a frequency domain. The proposed method uses speech absence probability (SAP) derived from the microphone input signal as the smoothing parameter updating noise power to reduce the noise power estimation error resulted from the distortions in the unified structure where the noise suppression (NS) operation is placed after the acoustic echo suppression (AES) algorithm. Therefore, in the proposed approach, the smoothing parameter based on SAP derived from the input signal instead of echo-suppressed signal should stop updating noise power estimates during the distorted noise spectrum periods. The performance of the proposed algorithm is evaluated by the objective test under various environments, and yields better results compared with the conventional scheme.

Keywords: Acoustic echo suppression, Noise suppression, Speech absence probability

ASK subject classification: Speech Signal Processing (2,3)

1. 서론

차량용 핸드프리나 화상통화가 가능한 휴대용 단말기의 보급이 확산되면서 관련 음성 통신 방식에서 발생할 수 있는 음향학적 반향 (acoustic echo)이나 배경잡음 제거에 대한 연구가 주목받고 있다 [1][2]. 최근에는 이러한 음향학적 반향 및 잡음 제거기의 통합된 구조가 일반화 되면서 각 알고리즘에 대한 상호 특성을 고려해 전체 알

고리즘의 성능을 향상시키기 위한 연구가 활발히 진행되고 있다 [3][4]. 일반적인 통합구조는 두 알고리즘이 서로 독립이라고 가정하고 병렬로 설계되어 직렬로 나열하는 방식이지만 실제로는 서로간의 위치에 따라 성능의 변화가 큰 것으로 알려져 있다. 구체적으로, 잡음제거기가 반향제거기의 앞단에 위치하게 되면, 잡음제거기의 비선형적인 연산으로 인해 반향제거기의 성능이 저하되고, 반대로 반향제거기가 잡음제거기의 앞단에 위치하게 되면, 잡음제거기가 추정해야할 잡음의 스펙트럼이 반향제거기에 의해 왜곡이 생겨 잡음추정이 심각히 저하되는 구조적 문제점이 97년 IEEE ICASSP에서 France Telecom의 연

책임저자: 장 준 혁 (changjh@inha.ac.kr)
인천시 남구 용현동 253 인하대학교 전자공학부
(전화: 032-860-7423; 팩스: 032-868-3654)

구조에 의해 분석된 바 있다 [5]. 따라서 본 논문에서는 이러한 통합 환경에서의 구조적 문제점, 특히 선(pre)반향-후(post)잡음 제거 통합구조의 잡음 제거 (NS, noise suppression) 알고리즘에서 음향학적 반향 억제 (AES, acoustic echo suppression)로 인해 발생하는 잡음전력 추정오차를 줄이기 위해 마이크로폰으로 유입되는 입력 신호의 음성부재확률 (SAP, speech absence probability)을 잡음전력 갱신을 위한 스무딩 (smoothing) 파라미터로 적용하는 잡음전력 추정 기법을 제안한다 [6]-[8]. 구체적으로 반향신호가 제기된 신호에서 잡음전력 갱신을 위한 SAP를 추출하는 대신 입력신호에 대한 SAP를 NS 알고리즘에 적용함으로써 잡음 제거기가 반향 억제 후 잔여 반향 (residual echo) 신호 등이 여전히 존재하여 잡음전력 추정오차를 증가시킬 수 있는 구간에서는 잡음전력을 갱신하지 않도록 한다. 제안된 방법의 성능 비교를 위해 실제 잡음신호에 대한 상대추정오차 (relative estimation error), segmental SNR (Signal to Noise Ratio) 및 ITU-T P.826 perceptual evaluation of speech quality (PESQ) 테스트를 실시하였고 제안된 기법은 다양한 잡음 환경에서 기존의 방법보다 우수한 성능을 보였다.

II. 음향학적 반향 및 잡음 제거의 통합 시스템

이 장에서는 제안된 잡음전력 추정 기법을 위해 음향학적 반향 및 잡음 제거의 통합시스템을 구성하고 있는 주파수 영역 기반 AES와 NS 알고리즘에 대해 간략하게 설명한다.

2.1 음향학적 반향 억제

그림 1은 일반적인 AES와 NS 통합구조에 대한 블록도를 보여주고 있다. 통합 알고리즘에서 반향신호 $d(t)$, 배경잡음 $w(t)$, 근단 (near-end) 화자신호 $s(t)$, 원단 (far-end) 신호와 마이크 입력신호를 각각 $x(t)$, $y(t)$ 라 하고 $Y(i,k)$ 를 $y(t)$ 의 i 번째 프레임의 k 번째 주파수 성분이라 하면 입력신호 $Y(i,k)$ 로부터 반향신호가 제거된 신호 $\hat{Y}(i,k)$ 는 다음과 같이 표현된다 [9].

$$\hat{E}(i,k) = G_{AES}(i,k) Y(i,k) \tag{1}$$

여기서 $G_{AES}(i,k)$ 는 AES 이득 (gain)으로서 다음과 같

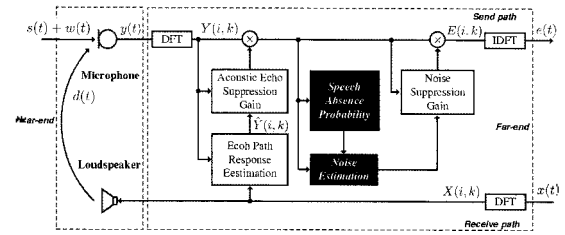


그림 1. 음향학적 반향제거와 잡음 제거 통합프레임의 직렬 블록도
Fig. 1. Block diagram of the structure for unified AES and NS.

은 Wiener 필터를 사용한다.

$$G_{AES}(i,k) = \frac{\max(|Y(i,k)| - |\hat{Y}(i,k)|, 0)}{|Y(i,k)|} \tag{2}$$

(2)식에서 $\hat{Y}(i,k)$ 는 원단신호 $X(i,k)$ 로부터 추정된 음향학적 반향신호이다. 일반적으로 입력 마이크로폰으로 전달되는 음향학적 반향신호는 다양한 반사 경로를 거쳐 입력되기 때문에 원단신호로부터 반사 경로를 고려한 임펄스응답에 대한 추정 과정이 필요하다. 따라서 반향 경로의 임펄스응답을 고려하여 원단신호의 주파수 성분 $X(i,k)$ 부터 추정된 반향신호 $\hat{Y}(i,k)$ 는 least squares 추정에 기반하여 다음과 같이 나타낼 수 있다 [9].

$$\hat{Y}(i,k) = \frac{E\{X_d^*(i,k) Y(i,k)\}}{E\{X_d^*(i,k) X_d(i,k)\}} |X(i,k)| \tag{3}$$

여기서 *는 complex conjugate를 의미하고 $X_d(i,k)$ 는 d 샘플 delay된 원단신호의 주파수 성분을 나타낸다. 일반적으로 실제 환경에서는 실제 마이크로폰으로 입력된 반향신호와 추정된 반향신호와의 불일치로 잔여 반향신호가 존재하게 된다. 이러한 잔여 반향신호는 잡음신호와 비슷한 성질을 보임으로서 이후 잡음제거기 및 추가적으로 적용될 수 있는 다양한 어플리케이션의 성능에 문제점을 일으킬 수 있다. 따라서 잔여 반향신호를 감소시키기 위한 다양한 잔여 반향제거 알고리즘이 추가적으로 적용될 수 있다 [4], [11].

2.2 잡음 제거

이 장에서는 본 논문의 통합구조에서의 잡음제거 알고리즘을 설명하기 위해 직렬 통합구조이거나 단독적으로 잡음제거기가 사용되는 경우에 음성의 통계적 모델에 기반한 기존의 잡음제거 방법에 대하여 소개한다. 이러한 잡음제거 기법에서는 음성 통계모델에 기반한 SAP를 잡

음성 전력 갱신을 위한 스무딩 파라미터로 적용한다 [8]. 구체적으로 음성의 통계모델에 기반한 soft decision 추정을 위해 반향신호가 제거된 신호 $\hat{E}(i, k)$ 에서 근단화자 신호 $S(i, k)$ 존재 할 때와 부재일 때의 각각의 가설 H_0 , H_1 에 대하여 다음과 같이 표현할 수 있다.

$$H_0: \text{near-end speech absent} : \hat{E}(i, k) = W(i, k) \quad (4)$$

$$H_1: \text{near-end speech present} :$$

$$\hat{E}(i, k) = W(i, k) + S(i, k) \quad (5)$$

여기서 잡음신호 $W(i, k)$ 는 음성신호인 $S(i, k)$ 와 통계적으로 독립이라고 가정한다. 음성과 잡음신호가 복소가우시안 분포를 따른다는 가정에서 H_0 와 H_1 의 확률밀도 함수는 다음과 같다 [6].

$$p(\hat{E}(i, k)|H_0) = \frac{1}{\pi \lambda_w(i, k)} \exp\left[-\frac{|\hat{E}(i, k)|^2}{\lambda_w(i, k)}\right] \quad (6)$$

$$p(\hat{E}(i, k)|H_1) = \frac{1}{\pi(\lambda_s(i, k) + \lambda_w(i, k))} \cdot \exp\left[-\frac{|\hat{E}(i, k)|^2}{\lambda_s(i, k) + \lambda_w(i, k)}\right] \quad (7)$$

여기서 $\lambda_s(i, k)$, $\lambda_w(i, k)$ 는 각각 음성과 잡음신호의 전력을 나타내며 Bayes' rule에 의하여 각 주파수 채널별 SAP를 구하면 아래와 같이 표현된다 [6].

$$p(H_0|\hat{E}(i, k)) = \frac{p(\hat{E}(i, k)|H_0)p(H_0)}{p(\hat{E}(i, k)|H_0)p(H_0) + p(\hat{E}(i, k)|H_1)p(H_1)} = \frac{1}{1 + qA(\hat{E}(i, k))} \quad (8)$$

여기서 $q = p(H_1)/p(H_0)$ 이고 위의 (6)식과 (7)식을 (8)식에 대입하면 likelihood ratio $A(\hat{E}(i, k))$ 는 다음과 같다.

$$A(\hat{E}(i, k)) = \frac{p(\hat{E}(i, k)|H_1)}{p(\hat{E}(i, k)|H_0)} = \frac{1}{1 + \xi(i, k)} \exp\left[\frac{\gamma(i, k)\xi(i, k)}{1 + \xi(i, k)}\right] \quad (9)$$

여기서 $\gamma(i, k)$, $\xi(i, k)$ 는 각각 *a posteriori* SNR (Signal to Noise Ratio)와 *a priori* SNR로 아래와 같이 정의 된다 [10].

$$\gamma(i, k) = \frac{|\hat{E}(i, k)|^2}{\lambda_w(i, k)} \quad (10)$$

$$\xi(i, k) = \frac{\lambda_s(i, k)}{\lambda_w(i, k)} \quad (11)$$

스무딩 파라미터에 의한 갱신으로 잡음전력을 추정하는 soft decision 기반의 잡음전력 추정은 long-term 스무딩된 전력 스펙트럼 $\hat{\lambda}_w(i, k)$ 로 다음과 같다 [8].

$$\hat{\lambda}_w(i, k) = \zeta_w \hat{\lambda}_w(i-1, k) + (1 - \zeta_w) E[|W(i, k)|^2 \hat{E}(i, k)] \quad (12)$$

$$\hat{\lambda}_s(i, k) = \zeta_s \hat{\lambda}_s(i-1, k) + (1 - \zeta_s) E[|S(i, k)|^2 \hat{E}(i, k)] \quad (13)$$

여기서 ζ_w ($0 < \zeta_w < 1$)와 ζ_s ($0 < \zeta_s < 1$)는 스무딩 파라미터이고 음성부재에 따른 잡음전력 추정을 구하기 위해 주파수 채널별 SAP가 적용되면 다음과 같이 표현된다 [8].

$$E[|W(i, k)|^2 \hat{E}(i, k)] = E[|W(i, k)|^2 \hat{E}(i, k) | H_0] p(H_0 | \hat{E}(i, k)) + E[|W(i, k)|^2 \hat{E}(i, k) | H_1] p(H_1 | \hat{E}(i, k)) \quad (14)$$

$$E[|S(i, k)|^2 \hat{E}(i, k)] = E[|S(i, k)|^2 \hat{E}(i, k) | H_0] p(H_0 | \hat{E}(i, k)) + E[|S(i, k)|^2 \hat{E}(i, k) | H_1] p(H_1 | \hat{E}(i, k)) \quad (15)$$

여기서 $E[\cdot]$ 는 기대값 연산자이고 잡음 제거 이득은 MMSE (minimum mean square error) 추정에 기반한 잡음 제거 필터를 적용하였다 [10].

III. 제안된 통합구조에서의 잡음전력 추정 기법

2장에서는 주파수 영역에서 AES와 NS 통합구조에 적용된 각각의 알고리즘에 대하여 간략히 설명하였다. 3장에서는 주파수 기반의 선반향-후잡음 통합구조에 적용된 잡음 제거 알고리즘에서 잡음전력 추정 오차를 줄이기 위한 새로운 잡음전력 추정을 기법을 제안한다. 일반적으로 AES와 NS 알고리즘이 독립적으로 설계되어 직렬로 연결되는 선반향-후잡음 통합구조에서는 NS 알고리즘은 순차적으로 음향학적 반향이 제거된 신호 $\hat{E}(i, k)$ 를 기반으로 하여 잡음 제거가 이루어진다. 하지만 97년 IEEE ICASSP에서 France Telecom의 연구진에 의해 이

러한 통합구조에서 반향제거기에 의해 선처리된 신호는 잡음 스펙트럼이 반향 제거기에 의해 왜곡이 생겨 잡음 제거 알고리즘에서 정확한 잡음전력 추정이 어려운 것으로 알려져 있다 [5]. 특히 반향신호가 존재했던 구간은 완벽한 AES가 어렵기 때문에 잔여 반향신호 [11]가 여전히 존재하게 되고 이러한 babble 잡음의 특성을 보이는 잡음 스펙트럼이 잡음전력 갱신에 포함되면 원래 잡음전력보다 크게 잡음전력이 추정되고 이후 근단화자신호에 대한 SNR을 감소시킬 수 있다. 따라서 본 논문에서는 잡음전력 추정오차를 감소시키기 위해 잡음전력 갱신에 스무딩 파라미터로 적용되는 SAP를 반향이 제거된 신호 $\hat{E}(i,k)$ 에 대한 SAP $p(H_0|\hat{E}(i,k))$ 대신 마이크로폰 입력 신호 $Y(i,k)$ 로부터 도출된 SAP $p(H_0|Y(i,k))$ 를 잡음전력 갱신에 적용한다. 그림 2와 3은 각각 제안된 잡음전력 추정 기법의 블록도와 입력신호 및 반향이 제거된 신호에 대한 SAP 차이를 보여 주고 있다. 그림 3의 (c)는 (a)와 (b) 신호에 대하여 각 구간의 SAP 변화를 보여주고 있다. (c)로부터 마이크로폰 입력신호에 대한 SAP $p(H_0|Y(i,k))$ 는 반향신호가 존재하는 구간에서 0에 가까운 값을 나타내지만 $p(H_0|\hat{E}(i,k))$ 는 1에 가까운 값을 가짐으로써 (14)식에 의해 잡음전력이 갱신되는 것을 알 수 있다. 이러한 잔여 반향신호가 남아 있는 구간에서의 잡음전력 갱신으로 인해 실제 잡음전력에 대한 추정오차가 증가될 수 있다. 따라서 $p(H_0|Y(i,k))$ 를 잡음전력 갱신을 위한 스무딩 파라미터로 적용하여 잔여 반향 신호구간에서는 잡음전력 갱신이 되지 않도록 하여 잡음전력 추정오차를 감소시킬 수 있다. 입력신호에 대한 SAP는 다음과 같다.

$$p(H_0|Y(i,k)) = \frac{p(Y(i,k)|H_0)p(H_0)}{p(Y(i,k)|H_0)p(H_0) + p(Y(i,k)|H_1)p(H_1)} = \frac{1}{1 + qA(Y(i,k))} \quad (16)$$

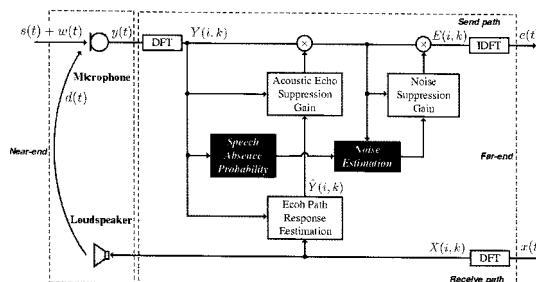


그림 2. 제안된 잡음전력추정 방법을 기반으로한 NS와 AES의 통합 블록도.

Fig. 2. Block diagram of the proposed noise power estimation method for the unified frame.

최종적으로 입력신호의 SAP가 적용된 잡음전력은 (12) 식과 (13)식을 이용하여 다음과 같이 표현된다.

$$\hat{\lambda}_w(i,k) = \zeta_w \hat{\lambda}_w(i-1,k) + (1 - \zeta_w) E[|W(i,k)|^2 | Y(i,k)] \quad (17)$$

$$\hat{\lambda}_s(i,k) = \zeta_s \hat{\lambda}_s(i-1,k) + (1 - \zeta_s) E[|S(i,k)|^2 | Y(i,k)] \quad (18)$$

여기서 $E[|W(i,k)|^2 | Y(i,k)]$ 와 $E[|S(i,k)|^2 | Y(i,k)]$ 는 다음과 같다.

$$E[|W(i,k)|^2 | Y(i,k)] = E[|W(i,k)|^2 | Y(i,k), H_0] p(H_0 | Y(i,k)) + E[|W(i,k)|^2 | Y(i,k), H_1] p(H_1 | Y(i,k)) \quad (19)$$

$$E[|S(i,k)|^2 | Y(i,k)] = E[|S(i,k)|^2 | Y(i,k), H_0] p(H_0 | Y(i,k)) + E[|S(i,k)|^2 | Y(i,k), H_1] p(H_1 | Y(i,k)) \quad (20)$$

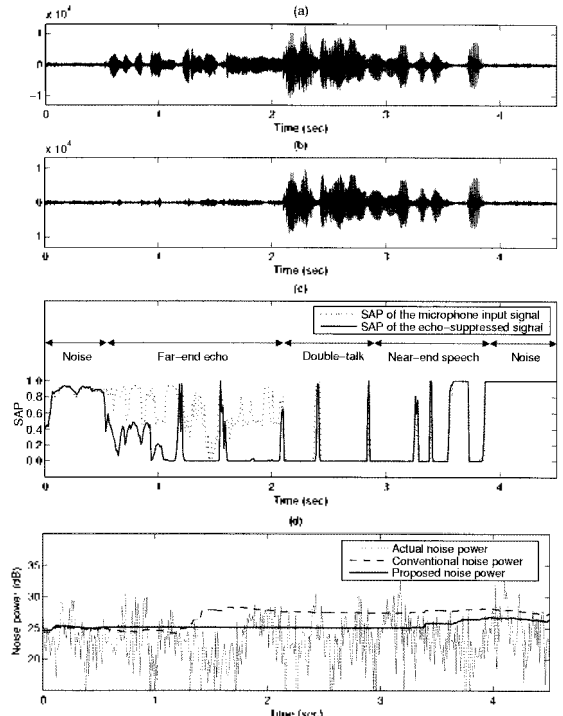


그림 3. SAP 비교 (babble noise SNR=20 dB): (a) 마이크로폰 입력신호에 대한 음성파형 (b) 반향신호가 제거된 신호에 대한 음성 파형 (c) 각 신호에 대한 SAP 비교 (d) 제안된 알고리즘에 의한 추정된 잡음전력 비교.

Fig. 3. SAP (babble noise SNR=20 dB): (a) the microphone input signal waveform (b) the echo-suppressed signal waveform (c) comparison of SAP for the microphone input and echo-suppressed signals (d) Comparison of Noise power estimation by the proposed algorithm.

IV. 실험 및 결과고찰

본 논문에서는 제안된 알고리즘의 성능 평가를 위해 다양한 잡음 환경에서 객관적인 실험을 수행하였다. 성능 평가는 실제 잡음전력과 추정된 잡음전력을 비교하기 위해 상대추정오차 (relative estimation error)와 segmental SNR [12] 그리고 객관적 음질평가를 위해 ITU-T P.826 PESQ 테스트를 실시하였으며 상대추정오차 ϵ_n 은 다음과 같다 [13].

$$\epsilon_n = \frac{1}{N} \sum_{n=0}^{N-1} \frac{\sum_k [\hat{\lambda}_w(i,k) - \lambda_w(i,k)]^2}{\sum_k \lambda_w^2(i,k)} \quad (21)$$

여기서 N 은 분석된 신호의 프레임 수이고, $\lambda_w(i,k)$ 는 실제 잡음신호부터 구해진 잡음전력을 의미하며 $\hat{\lambda}_w(i,k)$ 는 테스트 방법으로 부터 추정된 잡음전력을 나타낸다.

테스트 샘플을 위해 7명의 화자로 부터 얻은 8kHz로 샘플링 된 20개의 문장을 수집하고 각 문장을 원단화자와 근단화자신호로 분류하여 합성하였다. 원단화자신호로 분류된 음성은 섞기 전에 반사 경로를 고려한 실제 환경의 시뮬레이션을 위해 [14]에서 제안하는 room impulse response 모델링 기법에 기반한 임펄스응답 필터를 통과시키고 [15] 입력 마이크로폰으로 들어가는 반향신호는 근단화자신호 보다 3.5 dB 작게 설정하였다. 모델링 환경의 장소는 $5 \times 4 \times 3 \text{ m}^3$ 크기로 설정 하였고 잡음 환경을 위해서 babble과 vehicular 잡음을 다양한 SNR로 부가하였다. 그림 3의 (d)는 (c)의 SAP를 추출위한 테스트 샘플에 대하여 기존의 잡음전력 추정 기법에 기반한 잡음전력 추정치와 제안된 방법에 의한 잡음전력 추정치의 비교를 보여주고 있다. 기존의 방법은 원단 반향신호 구간 즉, 반향신호가 제거된 후 잔여 반향신호가 존재할 수 있는 구간에서 SAP에 의해 잡음전력이 갱신이 되어 실제 잡음전력보다 증가된 잡음전력 추정치를 나타내고 있다. 하지만 제안된 방법에서는 왜곡된 잡음 스펙트럼이 존재할 수 있는 구간에서 잡음전력 갱신을 범추게 되어 기존 방법에 의한 잡음전력 추정치보다 실제 잡음전력 의 평균값에 보다 가까운 잡음전력 추정치의 변화를 보이는 것을 알 수 있다. 표 1은 다양한 잡음 환경에서의 각각의 알고리즘에 의한 상대추정오차를 보여주고 있다. 상대추정오차는 (21)식에 의해 잡음 환경 및 SNR에 따른 각각의 테스트 샘플들에 대한 평균값으로 도출하였으며 표 1로부터 제안된 방법에 의한 잡음전력 추정치가 테스트의 모든

잡음 환경에서 기존의 방법보다 작은 상대추정오차를 보임으로서 향상된 성능을 나타내고 있는 것을 볼 수 있다. 또한 잡음제거 후 결과신호에 대하여 SNR 개선 정도를 평가하기 위해 segmental SNR 테스트를 실시하였다. 표 2로부터 대부분의 잡음환경에 대하여 제안된 기법이 향상된 SNR을 보이는 것을 알 수 있다. 최종적으로 통합프

표 1. 기존과 제안된 방법에 의한 상대추정오차 비교
Table 1. Relative estimation error obtained from conventional and proposed methods.

Environments		Relative Estimation Error	
Noise	SNR (dB)	Conventional	Proposed
Babble	5	0.818	0.685
	10	0.816	0.678
	15	0.807	0.679
	20	0.811	0.689
Vehicle	5	0.495	0.476
	10	0.502	0.477
	15	0.600	0.481
	20	0.434	0.430

표 2. 기존과 제안된 방법에 의한 segmental SNR 비교
Table 2. Segmental SNR obtained from conventional and proposed methods.

Environments		Segmental SNR	
Noise	SNR (dB)	Conventional	Proposed
Babble	5	3.247	3.479
	10	5.539	5.925
	15	4.068	4.387
	20	6.789	6.926
Vehicle	5	5.749	5.801
	10	3.754	3.838
	15	6.946	7.057
	20	7.297	7.299

표 3. 기존과 제안된 방법에 의한 PESQ 비교 (with 95% confidence interval)

Table 3. PESQ obtained from conventional and proposed methods (with 95% confidence interval).

Environments		PESQ	
Noise	SNR (dB)	Conventional	Proposed
Babble	5	1.727 ± 0.190	1.750 ± 0.190
	10	2.058 ± 0.160	2.074 ± 0.166
	15	2.318 ± 0.131	2.330 ± 0.139
	20	2.529 ± 0.094	2.542 ± 0.088
Vehicle	5	2.803 ± 0.092	2.815 ± 0.093
	10	2.908 ± 0.091	2.921 ± 0.091
	15	2.987 ± 0.091	2.998 ± 0.091
	20	3.049 ± 0.100	3.058 ± 0.100

래임의 전반적인 성능 비교를 위하여 최종 결과신호에 대한 객관적 음질평가를 위해 PESQ 테스트를 하였다. 표 3의 PESQ 수치로부터 제안된 알고리즘의 PESQ 수치가 기존의 방법보다 모든 잡음환경에 대하여 개선된 결과를 나타내는 것을 알 수 있다.

V. 결론

본 논문에서는 주파수 영역에서 음향학적 방향 및 잡음 제거의 통합 시스템을 위한 효과적인 잡음전력 추정 기법을 제안하였다. AES 후처리단으로 결합되는 NS 알고리즘에서 왜곡된 잡음 스펙트럼에 의한 잡음전력 추정오차를 줄이기 위해 반향신호가 제거된 신호 대신 마이크로폰 입력신호의 SAP를 잡음전력 갱신을 위한 스무딩 파라미터로 적용하였다. 객관적 테스트 결과로부터 제안된 방법이 기존의 방법보다 개선된 결과를 나타내었다.

감사의 글

본 연구는 지식경제부와 한국산업기술재단의 전략기술인력양성사업으로 수행된 연구결과이며, 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (ITTA-2009-C1090-0902-0010).

참고 문헌

1. H. Puder and P. Dreiseitel, "Implementation of a hands-free car phone with echo cancellation and noise-dependent loss control," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3622-3625, 2000.
2. P. Dreiseitel, E. Hansler, and H. Puder, "Acoustic echo and noise control—a long lasting challenge," in *Proc. EUSIPCO*, pp. 945-952, Sep. 1998.
3. S. Gustafsson, R. Martin, and P. Vary, "Combined acoustic echo control and noise reduction for hands-free telephony," *Signal Processing*, vol. 64, no. 1, pp. 21-32, 1998.
4. 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, issue 3, pp. 186-195, 2002.
5. Y. Guelou, A. Benamar, and P. Scalart, "Analysis of two structures for combined acoustic echo cancellation and noise reduction," *IEEE ICASSP*, vol. 2, pp.637-640, May, 1996.

6. R. J. McAulay and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-28, pp. 137-145, 1980.
7. J. Sohn and W. Sung, "A voice activity detector employing soft decision based noise spectrum adaptation," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. 1, pp. 365-368, 1998.
8. N. S. Kim and J.-H. Chang, "Spectral enhancement based on global soft decision," *IEEE Signal Processing Letters*, vol. 7, no. 5, pp. 108-110, 2000.
9. C. Faller and 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)*, pp. 53-56, Oct. 2005.
10. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-32, no. 6, pp. 1109-1121, 1984.
11. J.-H. Chang, H. G. Kim, and S. Kim, "Residual echo reduction based on MMSE estimator in acoustic echo canceller," *IEICE Electronic Express*, vol. 4, no. 24, pp. 762-767, 2007.
12. N. Virag, "Single channel speech enhancement based on masking properties of the human auditory system," *IEEE Trans. Speech Audio Process.*, vol. 7, no. 2, pp. 126-137, 1999.
13. I. Cohen and B. Berdugo, "Noise estimation by minima controlled recursive averaging for robust speech enhancement," *IEEE Signal Processing Letters*, vol. 9, no. 1, pp. 12-15, 2002.
14. *A model for room acoustics*, <http://l2pi.us/rir.html>, 2003.
15. S. Y. Lee and N. S. Kim, "A statistical model based residual echo suppression," *IEEE Signal Processing Letters*, vol. 14, no. 10, pp. 758-761, 2007.

저자 약력

•박 윤 식(Yun-Sik Park)



2006년 2월 : 인하대학교 전자공학과 학사
 2008년 2월 : 인하대학교 전자공학부 석사
 2008년 3월~현재 : 인하대학교 전자공학부 박사 과정

•장 준 희(Joon-Hyuk Chang)



1998년 2월 : 경북대학교 전자공학과 학사
 2000년 2월 : 서울대학교 전기공학부 석사
 2004년 2월 : 서울대학교 전기컴퓨터공학부 박사
 2000년 3월~2005년 4월 : 휴넷데스 연구소장
 2004년 5월~2005년 4월 : 캘리포니아 수립대학 산타바바라 (UCSB) 박사후연구원
 2005년 5월~2005년 8월 : 한국과학기술연구원 (KIST) 연구원
 2005년 9월~현재 : 인하대학교 전자공학부 조교수