한국음향학회지 제30권 제3호 pp. 123~128 (2011)

환경인식 기반의 향상된 Minimum Statistics 잡음전력 추정기법

Improved Minimum Statistics Based on Environment-Awareness for Noise Power Estimation

손 영 호*, 최 재 훈*, 장 준 혁**
(Young-Ho Son*, Jae-Hun Choi*, Joon-Hyuk Chang**)

*인하대학교 전자공학부, **한양대학교 융합전자공학부
(접수일자: 2011년 1월 20일; 채택일자: 2011년 2월 24일)

본 논문에서는 다양한 잡음 환경에서 음성향상을 위한 Minimum Statistics (MS) 잡음전력 추정 기법을 제시한다. 기존의 방법에서는 최소값 추적을 위해서 유한한 서치 (search)윈도우를 사용하여 최적으로 신호의 파워 스펙트럼을 수무딩하고 최소 확률을 적용하는 것을 기본으로 한다. 본 논문에서 제안된 알고리즘은 기존의 최소값 서치 윈도우가 다양한 잡음 환경에 상관없이 고정된 사이즈를 사용하는 것에 환경인식 정보를 적용하여 서치 윈도우 사이즈가 Gaussian mixture model (GMM)기반의 잡음 분류 알고리즘을 이용한 결과 값의 비교로 잡음 환경에 따라 변화 하도록 한다. 제안된 음성 향상 기법은 ITU-T P.862 perceptual evaluation of speech quality (PESQ)를 이용하여 평가하였고 기존의 MS방법보다 향상된 결과를 보였다.
핵심용어: Minimum Statistics (MS), Gaussian mixture model (GMM), 환경인식
투고분야: 음성처리 분야 (2)

In this paper, we propose the improved noise power estimation in speech enhancement under various noise environments. The previous MS algorithm tracking the minimum value of finite search window uses the optimal power spectrum of signal for smoothing and adopts minimum probability. From the investigation of the previous MS-based methods it can be seen that a fixed size of the minimum search window is assumed regardless of the various environment. To achieve the different search window size, we use the noise classification algorithm based on the Gaussian mixture model (GMM). Performance of the proposed enhancement algorithm is evaluated by ITU-T P.862 perceptual evaluation of speech quality (PESQ) under various noise environments. Based on this, we show that the proposed algorithm yields better result compared to the conventional MS method.
Keywords: Minimum Statistics (MS), Gaussian mixture model (GMM), Environment-awareness
ASK subject classification: Speech Signal Processing (2)

I. 서 론

최근 스마트폰 사용자가 증가하면서 이동환경에서의 음성 통신의 중요성이 점차 증가하고 있다. 실제적인 음성향상 시스템에서 잡음을 정확하게 추정하는 것이 가장 중요하며, 특히 비정상잡음 신호를 처리 할 수 있어야 한다. 잡음 신호의 추정은 음성 향상 시스템에 미치는 영향이 크기 때문에 추정된 잡음신호가 너무 작을 경우 자연스럽지 못한 잔류 잡음이 생기며, 너무 클 경우 음성 신호가 둔탁하게 들려 명료도가 떨어진다. 이러한 음성향상 기술의 성능에 주요 영향을 미치는 세부 기술로는 음성/잡음신호에 대한 통계적 모델 추정 [1-2], 스펙트럼 이득 수정 [3-6] 및 잡음신호의 추정 [7-9] 등으로 정리를 할 수 있는데, 대표적인 잡음전력 추정 방법으로 Minimum Statistics (MS) 기반의 잡음전력 추정법이 우수한 성능을 보인다고 알려져 있다 [7]. MS 잡음전력 추정은 음성신호와 잡음신호가 통계적으로 독립이라는 것과 잡음 섞인 음성 신호의 파워가 빈번히 잡음 신호의 파워 레벨까

책임저자: 장 준 혁 (jchang@hanyang.ac.kr)
서울시 성동구 행당동 17 공업센터별관 502-1호 한양대학교 융합전자공학부
(전화: 02-2220-0355)

지 감소하는 점에서 착안되었다. 그러므로 잡음 섞인 음성 신호의 파워 스펙트럼의 최소값을 추적하여 정확한 잡음 신호의 파워 스펙트럼을 추정하는 것이 가능하다는 것이다. 그러나 일반적으로 최소는 평균보다 작으므로 최소 추적 방법은 바이어스를 보상 함으로써 잡음전력을 추정한다 [10-11]. 하지만 이와 같은 잡음전력 추적 방법은 다양한 잡음 환경에 상관없이 고정된 서치 윈도우 사이즈를 사용하기 때문에 정상상태 (stationary) 또는 비정상상태 (non-stationary)의 다양한 잡음에 동일한 사이즈가 적용되므로 잡음전력 추적을 함에 있어서 적합하지 않을 수 있다.

본 논문에서는 다양한 잡음 환경에서 동일한 서치 윈도우 사이즈를 사용하는 기존의 MS방법대신에 환경인식 정보를 잡음 종류에 따라 서치 윈도우 사이즈가 가변 할 수 있도록 적용한 새로운 알고리즘을 제안하며 잡음의 종류를 구별하기 위해서 Gaussian mixture model (GMM)을 사용한다 [12-13]. 그 결과 잡음 전력 추정을 향상시켜 우수한 음성 향상 기법을 도출하였으며, 객관적 음질 평가 방법인 perceptual evaluation of speech quality (PESQ) 테스트 결과 기존의 MS보다 향상된 결과를 나타내었다.

II. MS (Minimum Statistics)의 고찰 및 최적의 서치 윈도우 사이즈 D 결정

2.1. Minimum Statistics (MS)의 고찰

주파수 대역이 제한되어 있고 샘플링 되어 있는 잡음 섞인 음성 신호를 $y(i)$라 하고, $y(i)=s(i)+n(i)$라고 한다. 여기서 $s(i)$는 잡음이 없는 깨끗한 음성신호이고 $n(i)$는 잡음 신호이다. 여기서 i는 시간 인덱스이며 퓨리에 변환을 통해 주파수 축에서 다음과 같이 표현 된다.

$$Y(t,k)=S(t,k)+N(t,k) \tag{1}$$

여기서 $Y(t,k)$, $S(t,k)$, $N(t,k)$는 각각 $y(i)$, $s(i)$, $n(i)$의 t번째 프레임에서의 k번째 주파수 성분이 된다. MS 추정은 음성 구간에서의 정확한 잡음전력 추정을 위해 음성의 짧은 휴지기간에서 갖는 최소 음성신호 전력 값을 이용하며 이러한 최소 전력 값은 순간적으로 잡음전력의 레벨까지 감소하기 때문에 최소 전력 값을 바탕으로 잡음 전력의 추정이 가능하다 [11]. 반복적인 스무딩된 스펙트럼 군으로 간단한 추적 알고리즘을 다음과 같이 구현할 수 있다.

$$P(t,k)=\alpha P(t-1,k)+(1-\alpha)|Y(t,k)|^2 \tag{2}$$

여기서 $P(t,k)$는 스무딩으로 구해진 전력스펙트럼 성분이며 고정 스무딩 파라미터 $\alpha=0.85$이다. 음성 신호의 존재 유무에 상관없이 잡음 신호의 파워 스펙트럼 추정 $\hat{\sigma}_N^2(t.k)$는 연속된 96개의 $P(t,k)$에서 최소값을 선택하여 구하며 고정 스무딩 파라미터를 적용하여 구해진 전력 스펙트럼의 최소값은 잡음전력의 평균보다 작은 값을 갖는 바이어스 상태이기 때문에 최종적인 잡음전력 추정을 위해 바이어스 보상을 다음과 같이 표현 된다.

$$\hat{\sigma}_N^2(t,k)=B_{\min}(t,k)P_{\min}(t,k) \tag{3}$$

여기서, $\hat{\sigma}_N^2(t,k)$는 바이어스 보상이 적용된 추정된 잡음 신호이며 $P_{\min}(t,k)$는 최소 전력추정의 확률 밀도이며, $B_{\min}(t,k)$는 바이어스 보상 함수이다. 최소 전력추정을 위해 윈도우 사이즈 D를 기반으로 하는 $P_{\min}(t,k)$은 다음과 같이 표현 된다 [11].

$$P_{\min}(t,k)=\min\{P(t,k),P(t-1,k),\cdots,P(t-D+1,k)\} \tag{4}$$

바이어스 보상은 오염된 음성신호의 분산에 기반하며 바이어스 보상 함수 $B_{\min}(t,k)$는 근사식으로 다음과 같이 표현 된다 [11].

$$B_{\min}(t,k)\approx 1+(D-1)\frac{2}{Q_{\text{eq}}(t,k)} \tag{5}$$

식 (5)에서 $Q_{\text{eq}}(t,k)$는 다음과 같이 표현된다.

$$\frac{1}{Q_{\text{eq}}(t,k)}\approx\frac{\widehat{\text{var}}P(t,k)}{2\hat{\sigma}_N^4(t-1,k)} \tag{6}$$

여기서 $\widehat{\text{var}}P(t,k)$는 스무딩된 전력 스펙트럼 $P(t,k)$의 분산으로 다음과 같이 표현된다.

$$\widehat{\text{var}}P(t,k)=\overline{P^2}(t,k)-\overline{P}^2(t,k) \tag{7}$$

여기서, $\overline{P}(t,k)$와 $\overline{P^2}(t,k)$는 다음과 같이 표현된다.

$$\overline{P}(t,k)=\beta(t,k)\overline{P}(t-1,k)+(1-\beta(t,k))P(t,k) \tag{8}$$

$$\overline{P^2}(t,k)=\beta(t,k)\overline{P^2}(t-1,k)+(1-\beta(t,k))P^2(t,k) \tag{9}$$

여기서 스무딩 파라미터인 $\beta(t,k)=\alpha^2(t,k)$로 최대값 0.8을 갖는다.

표 1. 다양한 잡음 환경에서 PESQ 수치 비교를 통한 최적화된 윈도우 사이즈 D (frame)
Table 1. Optimal window size of D (frame) is selected by comparing PESQ score.

Noise type	D
Babble noise	55
White noise	135
Vehicle noise	140
Office noise	65
F16 nosie	140

2.2. 최적의 서치 윈도우 사이즈 D 결정 방법

최소값 서치 윈도우사이즈 D를 다양한 잡음환경별의 최적화된 값을 찾기 위해서 40에서 140까지 1단위로 변화시켜가며 테스트를 진행하였다. 각 잡음별로 가장 음질이 우수한 D를 결정하기 위해 객관적인 음질평가로 공인된 ITU-T 862 perceptual evaluation of speech quality (PESQ)평가 방법을 사용하였으며, 가장 PESQ점수가 높은 D를 해당 잡음의 최적화된 D로 설정하였다. 이때 사용한 음성데이터를 한 프레임의 크기를 10 ms에서 8 kHz로 샘플링 하여 8초길이로, 남성, 여성 모두 96개의 문장이 발음된 음성에 표 1에 언급된 NOISEX-92의 총 5개의 잡음 (babble, white, vehicle, office, F16)을 5, 10, 15 dB SNR을 인가하여 테스트를 진행하였다.

III. GMM 결합기반의 MS 잡음전력추정

3.1. Gaussian Mixture Model (GMM)

제안된 잡음 분류에 사용되어지는 GMM은 화자 인식과 음악 인식에서 뛰어난 성능을 보이는 패턴 인식기이다 [12-13]. GMM은 주어진 표본 데이터 집합의 분포 밀도를 단 하나의 확률 밀도 함수로 모델링하는 방법을 개선한 밀도 추정 방법으로 복수 개의 가우시안 확률 밀도 함수로 데이터의 분포를 모델링 하는 방법이다. 따라서 GMM으로 데이터의 분포를 모델링할 경우에 혼합 성분 개수가 충분히 주어지고, 적절한 파라미터 값들만 주어진다면, 이론적으로는 어떠한 연속적인 분포도 거의 완벽하게 추정하여 모델링 할 수 있다. GMM패턴 인식기는 훈련부와 인식부로 구성되어 있으며, 훈련부에서는 각 잡음의 모델을 만들고 인식부에서는 이 모델을 이용하여 잡음을 인식한다. 먼저 분류 시스템에서 사용되는 GMM은 가우시안 혼합성분 밀도의 가중치 합인 함수로서 다음과 같이 표현된다.

$$P(\vec{x}\,|\,\lambda)=\sum_{i=1}^{M}\alpha_i P_i(\vec{x}) \tag{10}$$

$$P_i(\vec{x})=\frac{1}{(2\pi)^{\frac{L}{2}}\left|\sum_i\right|^{\frac{1}{2}}}exp\left\{-\frac{1}{2}(\vec{x}-\vec{\mu_i})^T(\sum_i)^{-1}(\vec{x}-\vec{\mu_i})\right\} \tag{11}$$

여기서 α_i는 혼합 성분의 가중치를 나타내고 $\vec{\mu_i}$는 평균 벡터, Σ_i는 공분산 행렬을 나타낸다.

$$\lambda=\left\{p_i,\vec{\mu_i},\Sigma_i\right\},\quad i=1,...,M \tag{12}$$

여기서, 훈련부는 식 (12)와 같은 파라미터를 가지고 Expectation Maximization (EM)알고리즘 기반의 학습을 통하여 잡음에 대한 혼합 가우시안 모델 λ를 추정하고 인식부는 훈련부에 만들어진 λ를 이용해서 입력된 음성신호에 대한 사후 확률을 구하여 가장 큰 확률을 갖는 모델을 찾는다. 실제로 구성된 모델에 실시간으로 입력받는 데이터의 특징벡터를 입력을 받으며 사용된 특징벡터로는 자기 상관함수 (Autocorrelation Function)와 반사계수 (Reflection Coefficients)를 사용한 Levinson-Durbin 알고리즘을 사용하여 구한 Linear Prediction Coding (LPC) 계수 10차, LPC 분석에서의 에러 성분에 대한 잔류 에너지 1차, 잔류에너지의 이동평균1차, 최소값 10을 가진 프레임의 에너지 1차, 프레임 에너지의 이동평균 1차를 사용 총14차의 특징벡터를 사용하였으며 최종적으로 다음과 같이 추정된 구간별 최적모델에 따른 우도 (likelihood)값을 비교하여 가장 큰 우도를 갖는 것으로 분류한다.

$$\hat{G}=\underset{1\le g\le S}{\arg max}\sum_{t=1}^{N}\log p(\vec{x}\,|\,\lambda), \tag{13}$$
$$g=5\,(1:babble,\,2:car,\,3:f16,\,4:office,\,5:white)$$

3.2. 제안된 잡음전력추정

기존의 MS 잡음전력 추정기법에서는 최소 잡음전력 추정을 위하여 $P_{\min}(t,k)$을 식 (4)와 같이 표현 하였다. 하지만 $P_{\min}(t,k)$을 구하는데 있어서 고정된 최소값 서치 윈도우 사이즈 D는 잡음전력추정 구간을 지정하는데 있어서 잡음환경을 고려하지 않고 고정된 값을 사용한

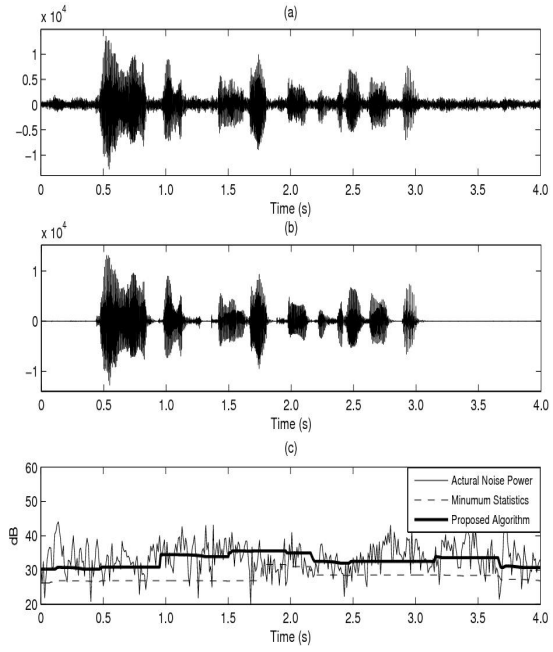


그림 1. babble 잡음 (SNR = 10 dB) 에서의 잡음전력 추정 비교 (a) 잡음 섞인 음성 파형, (b) 깨끗한 음성 파형 (c) 실제 잡음전력 (실선), MS기반 추정된 잡음전력 (일점선), 제안된 방법 기반 추정된 잡음전력 (굵은선)

Fig. 1. Comparison of noise power estimation (k=3) under babble noise (SNR = 10 dB). (a) Noisy speech, (b) Clean speech, (c) Actual noise power (solid line), estimated noise power based on MS(dashed line), estimated noise power based on proposed method.

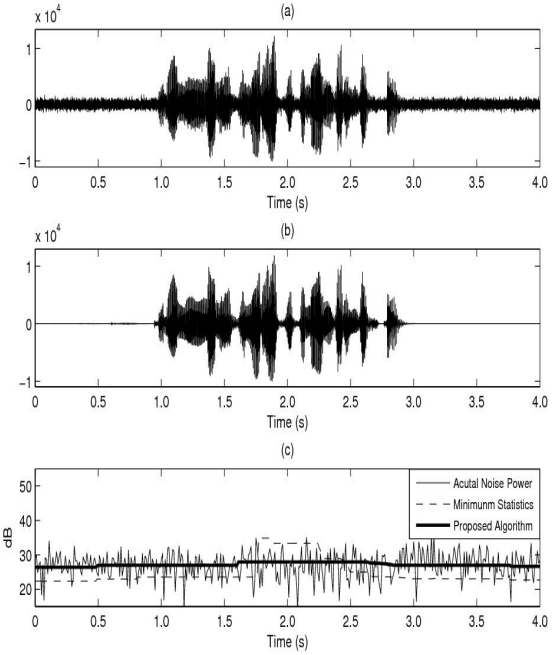


그림 2. white 잡음 (SNR = 10 dB) 에서의 잡음 전력 추정 비교 (a) 잡음 섞인 음성 파형, (b) 깨끗한 음성 파형, (c) 실제 잡음전력 (실선), MS기반 추정된 잡음전력 (일점선), 제안된 방법 기반 추정된 잡음전력 (굵은선)

Fig. 2. Comparison of noise power estimation (k=3) under white noise (SNR = 10 dB). (a) Noisy speech, (b) Clean speech, (c) Actual noise power (solid line), estimated noise power based on MS(dashed line), estimated noise power based on proposed method.

것이기 때문에 다양한 잡음 환경에서 적용할 경우 음성의 왜곡이 발생 할 수 있으며, 잡음전력 추정에 있어서 지연이 발생 할 수도 있다. 이러한 기존의 MS알고리즘을 보완하기 위하여 다양한 잡음 환경에서 최소값 서치 윈도우 사이즈 D를 최적화하기 위한 새로운 알고리즘을 다음과 같이 제안한다.

$$\hat{\sigma}_N^2(t,k) = \hat{B}_{\min}(t,k)\hat{P}_{\min}(t,k)) \tag{14}$$

여기서, $\hat{\sigma}_N^2(t,k)$ 는 잡음신호의 파워 스펙트럼 밀도이며, $\hat{B}_{\min}(t,k)$ 은 바이어스 보상 함수이며, $\hat{P}_{\min}(t,k)$은 최소 전력추정의 확률 밀도 이며 다음과 같이 표현된다.

$$\hat{B}_{\min}(t,k) \approx 1 + (\widehat{D}_g - 1)\frac{2}{Q_{eq}(t,k)} \tag{15}$$

$$\hat{P}_{\min}(t,k) = \min\left\{P(t,k), P(t-1,k), \cdots, P(t-\widehat{D}_g+1,k)\right\} \tag{16}$$

여기서 사용되는 최소값 서치 윈도우 사이즈 $\widehat{D}_g$는 식 (13)을 통하여 구해진다. g의 값이 1로 분류될 경우에는 babble 잡음으로 인식하여 서치 윈도우 사이즈 $\hat{D}_1$의 값을 55, 2로 분류될 경우 car 잡음으로 인식하여 $\hat{D}_2$의 값을 140, 3으로 분류될 경우 f16 잡음으로 인식하여 $\hat{D}_3$의 값을 140, 4로 분류될 경우 office 잡음으로 인식하여 $\hat{D}_4$의 값을 65, 5로 분류될 경우 $\hat{D}_5$의 값을 135로 인식을 한다. 여기서 $\widehat{D}_g(t)$는 서치 윈도우 사이즈 $\widehat{D}_g$가 잡음환경에 따라 갑작스럽게 바뀌는 것을 방지하기 위하여 Long-term 스무딩을 해주며 수식적으로 다음과 같이 표현된다.

$$\widehat{D}_g(t) = \gamma\widehat{D}_g(t-1) + (1-\gamma)\widehat{D}_g(t) \tag{17}$$

여기서, γ는 스무딩 파라미터로 0.99의 값을 갖는다. 또한 GMM패턴 인식기를 사용하는데 있어서 음성이 섞인 구간에서의 잡음정보 분류를 할 경우 다른 잡음으로 인식하기 때문에 이 같은 오류를 막기 위하여 speech absence probability (SAP)를 이용하여 잡음으로 인식된 구간에

서만 GMM패턴 인식기를 사용하였다 [14]. 그림 1, 2는 환경인식 정보를 통해 얻어진 가변 하는 서치 윈도우 사이즈와 기존의 고정된 서치 윈도우 사이즈를 적용 하였을 때 얻은 잡음전력 추정 값을 보여 주고 있다. 기존의 MS에서 고정된 서치 윈도우 사이즈는 다양한 잡음 환경을 고려하지 않았기 때문에 잡음 추정 구간이 일정하여 잡음전력 추정에 지연이 발생하는 것을 그림 1 (c), 2 (c)에서 볼 수 있다. 제안된 가변 하는 서치 윈도우 사이즈는 환경인식 결과를 통해 비정상상태 (non-stationary)잡음의 경우 잡음전력 추정 구간이 짧게 정상상태 (stationary) 잡음의 경우 잡음전력 추정 기간이 길게 선택되어 잡음전력을 추정하며 기존의 고정 값을 사용하던 MS보다 향상된 것을 볼 수 있다.

IV. 실험 결과

본 논문에서 제안된 음성향상 알고리즘의 성능을 평가하기 위해 널리 적용되고 있는 ITU-T P.862 PESQ 테스트를 통하여 객관적인 성능 평가를 하였다. 표 2의 PESQ 테스트를 위해 남성, 여성화자 각각이 48개의 문장을 발음하도록 한 총 96개의 음성 데이터를 한 프레임의 크기를 10 ms에서 8 kHz로 샘플링 하여 세 가지 형태의 잡음이 부가된 오염된 음성을 사용하였고 잡음은 NOISEX-92 데이터베이스의 babble noise, vehicle noise, f16 noise, office noise, white gaussian noise (WGN)를 사용 하였으며 SNR을 5, 10, 15 dB 세 가지로 나누어 테스트 하였다. PESQ값은 이들 샘플에 대한 평균 수치로 나타냈다. 기존의 MS 알고리즘의 서치 윈도우 사이즈 D를 96으로 설정 하여 실험하였으며 제안된 알고리즘의 경우 환경인식 인식 결과 정보에 따른 서치 윈도우 사이즈 $\widehat{D}_g$를 잡음 환경 에 따라 선택 하도록 하였다. 환경인식 정보기반의 잡음전력 추정방법이 기존의 MS 잡음전력 추정방법보다 PESQ수치로 babble, white, vehicle, f16, office noise에서 각각 평균 0.01, 0.05, 0.03, 0.01, 0.03 정도 향상된 수치를 보여다. 기타 다른 잡음에서는 원래의 MS기반 알고리즘과 유사한 성능을 보였다. 이는 제안된 알고리즘에서 잡음전력 추정을 위해 잡음 환경에 따라 서치 윈도우 사이즈가 환경인식 결과를 적용하여 가변하도록 하는 것이 고정된 서치 윈도우 사이즈를 사용하는 것보다 잡음전력 추정에 있어서 성능의 향상이 있음을 PESQ수치로 보여주고 있다.

표 2. 다양한 잡음 환경에서 Martin이 제안한 기존 방법과 제안된 알고리즘의 PESQ 수치 비교
Table 2. PESQ score of the original method made by Martin and proposed algorithm.

Noise type	Method	SNR (dB)		
		5	10	15
Babble noise	MS	2.30	2.63	2.93
	Proposed	2.31	2.65	2.94
White noise	MS	2.10	2.43	2.77
	Proposed	2.13	2.46	2.80
Vehicle noise	MS	3.48	3.69	3.87
	Proposed	3.52	3.73	3.92
Office noise	MS	2.34	2.63	2.94
	Proposed	2.35	2.64	2.95
F16 nosie	MS	2.87	3.11	3.37
	Proposed	2.89	3.14	3.41

V. 결 론

본 논문에서는 기존의 MS에 환경인식 정보를 적용하여 최소 잡음전력을 더욱 정확하게 추정하는 새로운 알고리즘을 제안하였다. 기존의 방법은 최소 잡음전력을 추정하기 위하여 고정된 서치 윈도우 사이즈를 사용하기 때문에 다양하게 변화하는 잡음 환경에서 정확도가 떨어진다는 단점을 가지고 있었다. 이러한 단점을 보완하기 위하여 환경인식 정보를 이용하여 잡음 환경에 따라 서치 윈도우 사이즈를 선택하도록 새로운 알고리즘을 제안 하였다. 그 결과, 음성 향상에 있어서 실험에 사용된 모든 잡음 환경과 신호 대 잡음 비 환경에서 기존의 MS방법보다 우수한 성능을 보였다.

감사의 글

이 논문은 2009년 정부 (교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (2009-0085162) 또한 본 연구는 지식경제부와 한국산업기술재단의 전략기술인력양성사업으로 수행된 연구결과임.

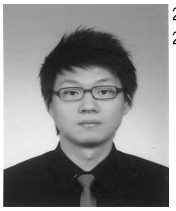
참 고 문 헌

1. J. W. Shin, J.-H. Chang, and N. S. Kim, "Statistical modeling of speech signals based on generalized gamma distribution," *IEEE Signal Processing Letters*, vol. 12, no. 3, pp. 258-261, Mar. 2005.
2. J.-H. Chang and N. S. Kim, "Voice activity detection based on complex Laplacian. model," *IEE Electronics Letters*, vol.

39, no. 7, pp. 632–634, Apr. 2003.
3. S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP–27, no. 2, pp. 113–120, Apr. 1979.
4. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean–square error log–spectral amplitude estimator," *IEEE Trans. ASSP*, vol. 33, no. 2, pp 443–445, Apr. 1985.
5. B. L. Sim, Y. C. Tong, J. S. Chang, and C. T. Tan, "A parametric formulation of the generalized spectral subtraction method," *IEEE Trans. on Speech and Audio Processing*, vol. 6, no. 4, pp. 328–336, July 1998.
6. R. J. McAualy 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, Apr. 1980.
7. R. Martin, "Spectral subtraction based on minimum statistics," *in Proc. Eur. Signal Processing Conf.*, pp. 1182–1185, 1994.
8. J. Sohn, W. Sung, "A voice activity detector employing soft decision based noise spectrum adaptation," *in Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, pp. 365–368, 1998.
9. 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, Jan. 2002.
10. G. Doblinger, "Computationally efficient speech enhancement by spectral minima tracking in subbands," *in Proc. EUROSPEECH*, vol. 2, pp. 1513–1516, 1995.
11. R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics," *IEEE. Trans. on Speech and Audio Processing*, vol. 9, no. 5, pp. 504–512, Jul. 2001.
12. G. Xuan, W.Zhang, and P. Chai, "EM algorithm of gaussian mixture model and hidden Markov model," in *Proc. IEEE International Conference on Image Processing*, vol. 1, pp. 145–148, Oct. 2001.
13. D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models," *Digital Signal Processing*, vol. 10, pp. 19–41, Jan. 2000.
14. 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, May 2000.

저자 약력

•손 영 호 (Young–Ho Son)



2010년 2월: 수원대학교 전자공학과 학사
2010년 3월 ~ 현재: 인하대학교 전자공학부 석사 과정

•최 재 훈 (Jae–Hun Choi)

2007년 2월: 인하대학교 전자공학과 학사
2007년 1월 ~ 2008년 2월: 삼성전자 정보통신 총괄 연구원
2008년 3월 ~ 2010년 2월: 인하대학교 전자공학과 석사
2010년 3월 ~ 현재: 인하대학교 전자공학과 박사 과정

•장 준 혁 (Joon–Hyuk Chang)

1998년 2월: 경북대학교 전자공학과 학사
2000년 2월: 서울대학교 전기공학부 석사
2004년 2월: 서울대학교 전기컴퓨터공학부 박사
2000년 3월 ~ 2005년 4월: ㈜넷더스 연구소장
2004년 5월 ~ 2005년 4월: 캘리포니아 주립대학, 산타바바라 (UCSB) 박사후연구원
2005년 5월 ~ 2005년 8월: 한국과학기술연구원 (KIST) 연구원
2005년 9월 ~ 2011년 2월: 인하대학교 전자공학부 조교수
2011년 3월 ~ 현재: 한양대학교 융합전자공학부 부교수