

# 강인한 음성향상을 위한 Minimum Statistics와 Soft Decision의 확률적 결합의 새로운 잡음전력 추정기법

## A Probabilistic Combination Method of Minimum Statistics and Soft Decision for Robust Noise Power Estimation in Speech Enhancement

박 윤 식\*, 장 준 혁\*

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

\*인하대학교 전자전기공학부

(접수일자: 2007년 2월 28일, 수정일자: 2007년 4월 24일, 채택일자: 2007년 5월 16일)

본 논문에서는 비정상적인 잡음 환경에서 음성향상을 위한 새로운 잡음 추정 기법을 제시한다. 제안된 방법은 잡음 전력 추정을 위해 주파수 채널별 음성부재확률 (SAP, Speech Absence Probability)을 선택적 가중 파라미터로 적용하여 음성 구간에서는 기존의 Minimum Statistics (MS)에 의한 잡음전력 추정치에 비중을 두고 비음성 구간에서는 Soft Decision (SD)에 기반한 잡음전력 추정치를 선택하도록 기존의 알고리즘을 결합한다. 제안된 알고리즘의 성능은 다양한 잡음 환경에서 음성향상기법에 적용하여 주관적인 음질평가 결과에 의해 평가하여 기존의 MS 또는 SD에 기반한 방법보다 향상된 결과를 나타내었다.

핵심용어: 잡음전력추정, 최소 확률, Soft decision, 음성향상, 음성부재확률

투고분야: 음성 처리 분야 (2,3)

This paper presents a new approach to noise estimation to improve speech enhancement in non-stationary noisy environments. The proposed method combines the two separate noise power estimates provided by the minimum statistics (MS) for speech presence and soft decision (SD) for speech absence in accordance with SAP (Speech Absence Probability) on a separate frequency bin. The performance of the proposed algorithm is evaluated by the subjective test under various noise environments and yields better results compared with the conventional MS or SD-based schemes.

*Key words:* Noise power estimation, Minimum statistics, Soft decision, Speech enhancement, SAP

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

## 1. 서론

음성향상기술은 다양한 음성신호처리시스템의 전처리기 (preprocessor)로 사용되어 시스템의 성능을 향상시키며 특히, 음성통신을 기반한 음성부호화기의 전처리기로써 이동환경에서의 배경잡음에 대한 직접적인 음질 향상 및 음성 부호화기의 성능 향상에 중요한 요소로 작용한다. 이러한 음성향상기술의 성능에 주요 영향을 미치는

세부 기술로는 음성/잡음신호에 대한 통계적 모델 추정 [1], [2], 스펙트럼 이득 추정 [3]-[6] 및 잡음신호의 추정 [7]-[9] 등으로 정리 할 수 있으며, 최근 이동통신 기술의 발달로 이동 중에 발생할 수 있는 비정상적 (non-stationary)인 배경잡음과 음성 구간에서 배경잡음 추정의 어려움 때문에 정확한 잡음 추정을 위한 연구가 주목 받고 있다.

현재 대표적인 잡음전력 추정 방법으로 Minimum Statistics (MS)와 Soft Decision (SD)에 기반한 잡음전력 추정법이 우수한 성능을 보인다고 알려져 있다 [7], [8]. MS 잡음전력 추정은 음성신호와 잡음신호가 통계적으로 독립이라는 가정 하에 오염된 음성 구간의 휴지구간

에서 음성전력이 잡음신호의 전력 레벨까지 감소하는 점에서 착안되었다. 따라서, 이러한 음성 구간의 레벨 사이를 연결해줄 충분히 큰 윈도우를 사용하여 최적화된 전력 스펙트럼 스무딩 (smoothing)을 거친 최소 잡음전력을 추정한 후 각 스펙트럼 성분에 해당하는 바이어스 (bias) 보상을 함으로써 잡음전력을 추정한다 [10], [11].

이에 반하여 SD에 기반한 잡음전력 추정은 현재 잡음 신호 추정을 위하여 확률에 기반한 프레임간 스무딩을 이용해 이전의 추정신호를 갱신하여 현재의 잡음전력을 추정하는 방법이다 [8], [12]. 다시 말해 잡음신호로 확실하게 추정된 기존의 잡음신호를 기반으로 통계적 스무딩을 위해 가중치 파라미터로 사용되는 음성부재확률 (SAP, Speech Absence Probability)에 따라 스펙트럼 성분을 갱신함으로써 잡음전력을 추정한다.

본 논문에서는 MD와 SD에 기반한 잡음전력 추정법을 분석하여 이를 토대로 새로운 잡음추정기법을 제안한다. 구체적으로, SD 추정에서 사용된 주파수 채널별 SAP를 가중치로 사용하여 [14] 음성 구간에서는 MS로 추정된 잡음전력에 더 큰 가중치를 두고, 비음성 구간에서는 SD로 추정된 잡음전력에 큰 가중치가 적용되도록 하는 결합 잡음전력추정 CNPE (Combined Noise Power Estimation) 알고리즘을 제안하였다. 제시된 알고리즘의 성능은 다양한 잡음 환경에서 Mean Opinion Score (MOS)에 의해 평가 했고 기존의 MS 또는 SD보다 향상된 결과를 나타내었다.

## II. MS 와 SD

이 장에서는 제안된 잡음전력 추정을 위해 MS와 SD를 이용한 잡음전력 추정 기법에 대해 간략하게 설명한다.

### 2.1. MS 기반 잡음전력 추정

잡음신호  $n$ 이 음성신호  $x$ 에 인가되면 오염된 음성신호  $y$ 를 만들어 내게 되고 푸리에 변환을 통해 주파수 축에서 다음과 같이 표현된다.

$$Y(t, k) = X(t, k) + N(t, k) \quad (1)$$

여기서  $Y(t, k)$ ,  $X(t, k)$ ,  $N(t, k)$ 는 각각  $y$ ,  $x$ ,  $n$ 의  $t$ 번째 프레임에서의  $k$ 번째 주파수 성분이 된다. MS 추정은 음성 구간에서의 정확한 잡음전력 추정을 위해 음성의 짧

은 휴지기간에서 갖는 최소 음성신호 전력값을 이용한다. 이러한 최소 전력값은 순간적으로 잡음전력의 레벨까지 감소하기 때문에 최소 전력값을 바탕으로 잡음전력의 추정이 가능하다 [11]. 최소 잡음전력을 얻기 위해 스무딩이 적용된 전력 스펙트럼 성분을 다음과 같이 구한다.

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

여기서  $P(t, k)$ 는 스무딩으로 구해진 전력스펙트럼 성분이며 고정 스무딩 파라미터  $\alpha=0.85$  이다. 하지만 1에 가까운 스무딩 파라미터에 의해 추정된  $P(t, k)$ 는 넓은 동적 범위의 음성 구간을 가지게 되며 이는 부정확한 잡음 전력 추정을 유도할 수 있다. 이러한 문제를 해결하기 위해 최적화된 스무딩을 위한 가변 스무딩 파라미터  $\alpha_{opt}(t, k)$ 는 아래와 같이 제시된다.

$$\alpha_{opt}(t, k) = \frac{1}{1 + (P(t-1, k)/\lambda_n(t, k) - 1)^2} \quad (3)$$

여기서  $\lambda_n(t, k) = E\{|N(t, k)|^2\}$ 는 잡음신호의 전력 스펙트럼이며  $\alpha_{opt}(t, k)$ 는  $P(t-1, k)/\lambda_n(t, k)$ 가 1일 때 발생하는 교착상태를 고려해 최대값을 0.96으로 제한한다.

가변 스무딩 파라미터를 적용하여 구해진 전력 스펙트럼의 최소값은 잡음전력의 평균보다 작은 값을 갖는 바이어스 상태이기 때문에 최종적인 잡음전력 추정을 위해 바이어스 보상을 다음과 같이 한다 [11].

$$\hat{\lambda}_n^{MS}(t, k) = B_{min}(t, k) \lambda_{min, n}(t, k) \quad (4)$$

여기서,  $\hat{\lambda}_n^{MS}(t, k)$ 는 바이어스 보상이 적용된 추정된 잡음신호이며  $B_{min}(t, k)$ 는 바이어스 보상 함수,  $\lambda_{min, n}(t, k)$ 는 최소 잡음전력이다. 바이어스 보상은 오염된 음성신호의 분산에 기반하며 바이어스 보상 함수  $B_{min}(t, k)$ 는 아래와 같이 근사식으로 표현할 수 있다.

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

(5)식에서  $D$ 는 한 프레임당 스펙트럼 개수로 128이며,  $Q_{eq}(t, k)$ 는 다음과 같이 표현된다.

$$\frac{1}{Q_{eq}(t, k)} \approx \frac{\widehat{\text{var}} P(t, k)}{2\hat{\lambda}_n^2(t-1, k)} \quad (6)$$

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

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

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

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

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

### 2.2. SD 기반 잡음전력 추정

음성의 통계모델에 기반한 SD 추정을 위해 음성이 존재 할 때와 비음성일 때의 각각의 가정  $H_0, H_1$ 에 대하여 오염된 음성은 (1)식과 비교하여 다음과 같이 표현할 수 있다.

$$H_0 : \text{speech absent} : Y(t, k) = N(t, k) \tag{10}$$

$$H_1 : \text{speech present} : Y(t, k) = X(t, k) + N(t, k)$$

여기서 잡음신호  $N(t, k)$ 는 음성신호인  $X(t, k)$ 와 통계적으로 독립이라고 가정한다. 음성과 잡음신호가 복소가우시안 분포를 따른다는 가정에서  $H_0$ 와  $H_1$ 의 확률밀도함수는 다음과 같다 [4].

$$p(Y(t, k)|H_0) = \frac{1}{\pi\lambda_n(t, k)} \exp\left[-\frac{|Y(t, k)|^2}{\lambda_n(t, k)}\right] \tag{11}$$

$$p(Y(t, k)|H_1) = \frac{1}{\pi(\lambda_s(t, k) + \lambda_n(t, k))} \cdot \exp\left[-\frac{|Y(t, k)|^2}{\lambda_s(t, k) + \lambda_n(t, k)}\right] \tag{12}$$

여기서  $\lambda_s(t, k), \lambda_n(t, k)$ 는 각각 음성과 잡음신호의 전력을 나타내며 Bayes' rule에 의하여 각 주파수 채널별 SAP를 구하면 아래와 같이 표현된다 [6].

$$p(H_0|Y(t, k)) = \frac{p(Y(t, k)|H_0)p(H_0)}{p(Y(t, k)|H_0)p(H_0) + p(Y(t, k)|H_1)p(H_1)} \tag{13}$$

$$= \frac{1}{1 + qA(Y(t, k))}$$

여기서  $q = p(H_1)/p(H_0)$ 이고 위의 (11)식과 (12)식을 (13)식에 대입하면 likelihood ratio  $A(Y(t, k))$ 는 다음과 같다 [13].

$$A(Y(t, k)) = \frac{p(Y(t, k)|H_1)}{p(Y(t, k)|H_0)} = \frac{1}{1 + \xi(t, k)} \exp\left[\frac{\gamma(t, k)\xi(t, k)}{1 + \xi(t, k)}\right] \tag{14}$$

여기서  $\gamma(t, k), \xi(t, k)$ 는 각각 *a posteriori* SNR (Signal to Noise Ratio)와 *a priori* SNR로 아래와 같이 정의 된다 [4].

$$\gamma(t, k) \equiv \frac{|Y(t, k)|^2}{\lambda_n(t, k)}, \quad \xi(t, k) \equiv \frac{\lambda_s(t, k)}{\lambda_n(t, k)} \tag{15}$$

스무딩 파라미터에 의한 갱신으로 잡음전력을 추정하는 SD 기반의 잡음전력 추정은 Long-Term 스무딩된 전력 스펙트럼  $\hat{\lambda}_n(t, k)$ 로 아래와 같다.

$$\hat{\lambda}_n^{SD}(t+1, k) = \zeta_n \hat{\lambda}_n^{SD}(t, k) + (1 - \zeta_n) E[|N(t, k)|^2 | Y(t, k)] \tag{16}$$

여기서  $\zeta_n$ 은  $0 < \zeta_n < 1$ 의 범위를 갖는 스무딩 파라미터로  $\zeta_n = 0.99$ 이며 음성부재에 따른  $E[|N(t, k)|^2 | Y(t, k)]$ 를 구하기 위해 주파수 채널별 SAP가 적용되면 아래와 같이 표현 된다 [12].

$$E[|N(t, k)|^2 | Y(t, k)] = E[|N(t, k)|^2 | Y(t, k), H_0] p(H_0 | Y(t, k)) + E[|N(t, k)|^2 | Y(t, k), H_1] p(H_1 | Y(t, k)) \tag{17}$$

여기서

$$E[|N(t, k)|^2 | Y(t, k), H_0] = |Y(t, k)|^2 \tag{18}$$

$$E[|N(t, k)|^2 | Y(t, k), H_1] = \left(\frac{\xi(t, k)}{1 + \xi(t, k)}\right) \hat{\lambda}_n^{SD}(t, k) + \left(\frac{1}{1 + \xi(t, k)}\right)^2 |Y(t, k)|^2 \tag{19}$$

### III. 확률적 결합기반 잡음전력추정

MS 추정은 음성신호의 짧은 휴지기간의 최소값을 이용하여 잡음전력을 추정하기 때문에 스무딩에 의한 SD추정 보다 음성구간에서 정확한 잡음전력 추정이 가능하다. 구체적으로 SD는 음성부재의 확률에 근거한 스무딩을 이용하므로 음성이 지속되는 경우, 전력이 급격히 감소하는 휴지구간에서는 이전 전력 스펙트럼 성분에 영향을 받는 스무딩에 의해서 정확한 잡음전력을 추정하기가 어렵지만 MS추정은 이러한 휴지기간의 최소 잡음전력을 추정된 상태에서 바이어스 보상이 되기 때문에 휴지구간에서도 효과적으로 잡음전력을 추정할 수 있다. Babble noise가

SNR 5 dB로 인가된 잡음 음성을 각 추정 방법에 의해 잡음이 제거된 음성신호와 깨끗한 음성신호의 에너지를 음성구간에서 비교한 그림 1에서 볼 수 있듯이 MS 추정의 잡음전력을 이용한 음성향상 알고리즘에 의해 잡음이 제거된 음성신호가 음성 구간의 휴지부분에서 SD보다 본래의 깨끗한 음성신호에 더 가까운 것을 볼 수 있다.

반면에, 잡음 구간에서는 음성 구간에서와는 다르게 MS 기법은 주어진 음성신호의 최소 전력값으로부터 잡음전력을 지속적으로 추정하는 과정상의 에러가 상대적으로 크지만, SD기법은 SAP에 의해 확실히 잡음구간이라고 판별이 되면 식 (16)에 기반하여 그림 2에서 보듯이 Long-Term 스무딩된 형태로 실제 잡음전력의 궤적을 그대로 따라가는 구조이므로 MS기법에 비해 보다 정확히 잡음전력을 추정할 수 있는 장점이 있다. 결론적으로, SNR 10 dB의 white gaussian noise (WGN)에 대해 잡음 구간에서의 각각의 추정 방법에 의해 추정된 잡음전력을 실제 잡음전력과 비교하여 도시한 그림 2에서 보듯이 잡음 구간에서 SD 추정에 의한 잡음전력이 MS보다 실제 잡음신호를 정확하게 추정하는 것을 확인할 수 있다.

따라서, 본 논문에서는 효과적인 잡음전력 추정을 위해 SD 추정에서 구해진 주파수 채널별 SAP를 가중치 파라미터로 적용하여 [14] 음성 구간에서는 MS에 의한 잡음전력 추정치  $\hat{\lambda}_n^{MS}$ 에 비중을 두고 비음성 구간에서는 SD에 의한 잡음전력 추정치  $\hat{\lambda}_n^{SD}$ 에 비중을 두도록 기존의 알고리즘을 확률적으로 결합하여 새로운 잡음전력  $\hat{\lambda}_n^{CNPE}$ 을 구하는 결합잡음전력추정 (Combined Noise Power Estimation, CNPE) 알고리즘을 아래와 같이 제안한다.

$$\hat{\lambda}_n^{CNPE} = (1 - p(H_0|Y(t, k)))\hat{\lambda}_n^{MS} + p(H_0|Y(t, k))\hat{\lambda}_n^{SD} \quad (20)$$

식(20)에서 CNPE에 의하여 구해진 잡음전력  $\hat{\lambda}_n^{CNPE}$ 은 일반적으로 오염된 음성신호와 잡음제거이득의 곱으로

$$\hat{X}(t, k) = G(\xi(t, k), \gamma(t, k)) Y(t, k) \quad (21)$$

식(21)과 같이 표현되는 잡음제거 방법에서 아래식과 같이 MMSE (Minimum Mean Square Error)에 기반한 잡음제거이득  $G(\cdot, \cdot)$ 의 주요파라미터인 *a priori* SNR  $\xi(t, k)$ 과 *a posteriori* SNR  $\gamma(t, k)$ 에 적용 된다 [4].

$$G(\xi(t, k), \gamma(t, k)) = \frac{\sqrt{\pi v(t, k)}}{2\gamma(t, k)} \exp\left\{-\frac{v(t, k)}{2}\right\} \left[ (1 + v(t, k))I_0\left(\frac{v(t, k)}{2}\right) + v(t, k)I_1\left(\frac{v(t, k)}{2}\right) \right] \quad (22)$$

여기서  $v(t, k)$ 는 아래와 같이 주어진다.

$$v(t, k) = \frac{\xi(t, k)}{1 + \xi(t, k)} \gamma(t, k) \quad (23)$$

잡음제거이득 파라미터 *a priori* SNR을 구하기 위해 깨끗한 음성신호 전력은 다음과 같이 추정된다 [12].

$$\hat{\lambda}_s(t+1, k) = \zeta_s \hat{\lambda}_s(t, k) + (1 - \zeta_s) E[|X(t, k)|^2 | Y(t, k)] \quad (24)$$

여기서  $\zeta_s$ 은  $0 < \zeta_s < 1$ 의 범위를 갖는 스무딩 파라미터로  $\zeta_s = 0.98$ 이고 음성부재를 고려한  $E[|X(t, k)|^2 | Y(t, k)]$ 를 구하기 위해 주파수 채널별 SAP이 적용되면 다음과 같이 표현 된다.

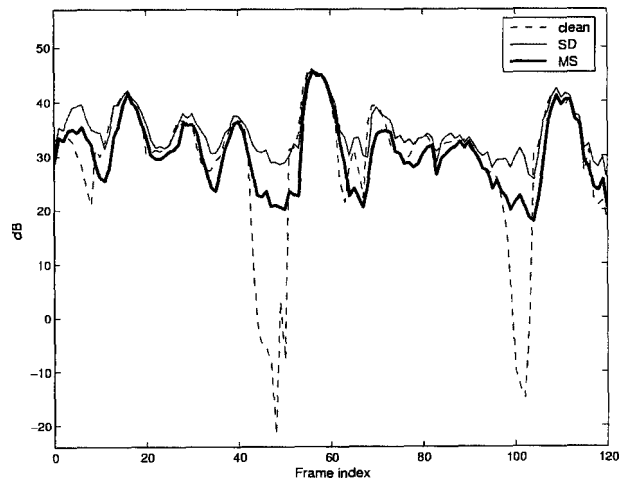


그림 1. 음성 구간에서의 깨끗한 음성신호와 MS와 SD에 의해 잡음이 제거된 음성신호의 전력 비교

Fig 1. Power contours of the clean and enhanced speech signal in speech periods.

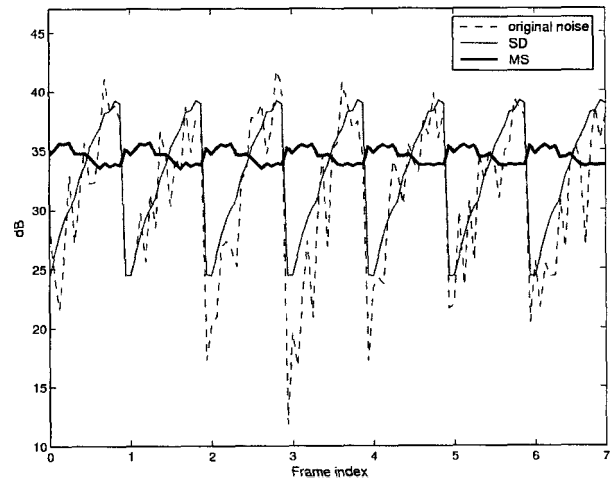


그림 2. 잡음 구간에서의 실제 잡음신호와 MS와 SD에 의해 추정된 잡음신호의 전력 비교

Fig 2. Original and estimated noise power in noise periods.

$$E[|X(t,k)|^2|Y(t,k)] = E[|X(t,k)|^2|Y(t,k), H_0]P(H_0|Y(t,k)) + E[|X(t,k)|^2|Y(t,k), H_1]P(H_1|Y(t,k)) \quad (25)$$

여기서

$$E[|X(t,k)|^2|Y(t,k), H_0] = 0 \quad (26)$$

$$E[|X(t,k)|^2|Y(t,k), H_1] = \left( \frac{1}{1 + \hat{\xi}(t,k)} \right) \hat{\lambda}_s(t,k) + \left( \frac{\hat{\xi}(t,k)}{1 + \hat{\xi}(t,k)} \right)^2 |Y(t,k)|^2 \quad (27)$$

이고, 따라서 새로운 잡음전력 추정치  $\hat{\lambda}_n^{CNPE}(t,k)$ 가 적용된 *a priori* SNR  $\hat{\xi}(t,k)$ 는 식(20)과 (24)을 이용하여 아래와 같이 다시 표현된다.

$$\hat{\xi}(t,k) = \frac{\hat{\lambda}_s(t,k)}{\hat{\lambda}_n^{CNPE}(t,k)} \quad (28)$$

두 번째 주요파라미터인 *a posteriori* SNR도 아래와 같이 다시 주어진다.

$$\hat{\gamma}(t,k) = \frac{|Y(t,k)|^2}{\hat{\lambda}_n^{CNPE}(t,k)} \quad (29)$$

#### IV. 실험 및 결과고찰

본 논문에서는 제안된 음성향상 알고리즘의 음질 평가를 위해 MOS 테스트를 수행하였다. MOS 테스트를 위해 남성, 여성화자 각각이 5개의 문장을 발음하도록 한 총 10개의 음성 데이터를 한 프레임의 크기를 10 ms에서 8 kHz로 샘플링 하여 세가지형태의 잡음이 부가된 오염된 음성을 사용하였고 잡음은 NOISEX-92 데이터베이스의 white gaussian noise (WGN), babble noise, vehicle

표 1. 다양한 노이즈 환경에서 MOS 비교  
Table 1. Comparison of the MOS under the various noise type.

Noise type	Method	SNR (dB)		
		5	10	15
WGN	MS	1.57	2.20	2.69
	SD	1.78	2.34	2.91
	CNPE	1.86	2.51	2.91
Babble noise	MS	2.40	2.98	3.39
	SD	2.27	2.90	3.42
	CNPE	2.45	3.02	3.70
Vehicle noise	MS	2.76	3.40	3.93
	SD	3.06	3.69	4.31
	CNPE	3.20	3.87	4.43

noise를 사용 하였으며 SNR을 5, 10, 15 dB로 달리하였다.

테스트에 사용된 음성 데이터에서 제안된 기법에 의해 추정된 잡음 전력 추정치는 MMSE 기반의 잡음제거 알고리즘에 적용되어 잡음이 제거된 음성 데이터를 대상으로 10명의 청취자에 의해 MOS를 결정하였다. 표 1은 다양한 잡음 환경과 SNR에 대한 각 알고리즘의 MOS 결과를 보여주고 있다. 결과를 고찰해보면, 대부분의 잡음 환경에 대해서 제안된 CNPE 추정이 기존의 MS 또는 SD에 기반한 추정 방법보다 음성향상에 우수한 성능을 가지고 있음을 알 수 있다.

#### V. 결론

본 논문에서는 잡음전력 추정을 위해 MS 기법이나 SD 기법에 기반한 잡음전력 추정치를 각 주파수 채널별 SAP에 따라 결합하는 새로운 잡음전력 추정 알고리즘을 제안하였다. 주파수 채널별 SAP를 기준으로 적용하여 음성구간에서는 MS에 의해 얻어진 잡음전력 추정치에 비중을 두고 비음성 구간에서는 SD에 기반한 잡음전력 추정치를 선택하도록 기존의 알고리즘을 결합하였다. 제안된 음성향상기술의 성능을 평가하기 위해 MOS 평가를 했으며 다양한 잡음 환경에서 제시된 CNPE 추정이 기존의 MS나 SD보다 향상된 결과를 나타내었다.

#### 참고 문헌

1. J. W. Shin, J.-H. Chang, N. S. Kim, "Statistical modeling of speech signals based on generalized gamma distribution," IEEE Signal Processing Letters, 12 (3) 258-261, Mar. 2005.
2. J.-H. Chang and N. S. Kim, "Voice activity detection based on complex Laplacian model," IEE Electronics Letters, 39 (7) 632-634, April 2003.
3. S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," IEEE Trans. Acoust., Speech, Signal Process., ASSP-27 (2) 113-120, Apr. 1979.
4. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," IEEE Trans. Acoust., Speech, Signal Process., ASSP-32, (6) 1109-1121, Dec. 1984.
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, 6 (4) 328-336, July 1998.
6. R. J. McAulay and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter," IEEE Trans. Acoust.,

Speech, Signal Processing, ASSP-28, 137-145, Apr. 1980.

7. R. Martin, "Spectral subtraction based on minimum statistics," in Proc. Eur. Signal Processing Conf., 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, 365-368, 1998.
9. I. Cohen, B. Berdugo, "Noise estimation by minima controlled recursive averaging for robust speech enhancement," IEEE Signal Processing Letters, 9 (1) 12-15, Jan. 2002.
10. G. Doblinger, "Computationally efficient speech enhancement by spectral minima tracking in subbands," in Proc. EUROSPEECH, 2, 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, 9 (5) 504-512, July 2001.
12. N. S. Kim, J.-H. Chang, "Spectral enhancement based on global soft decision," IEEE Signal Processing Letters, 7 (5) 108-110, May 2000.
13. J. Sohn, N. S. Kim, W. Sung, "A statistical model-based voice activity detection," IEEE Signal Processing Letters, 6 (1) 1-3, Jan. 1999.
14. N. S. Kim, Y. J. Kim, and H. W. Kim, "Feature compensation based on soft decision," IEEE Signal Processing Letters, 11 (3) 378-381, Mar. 2004.

---

**저자 약력**

---

• **박 윤 식 (Yun-Sik Park)**



2006년 2월: 인하대학교 전자공학과 학사  
 2006년 3월~현재: 인하대학교 전자공학과 석사과정

• **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: 넷넷더스 연구소장  
 2004년 5월~2005년 4월: 캘리포니아 주립대학, 산타바바라 (UCSB) 박사후연구원  
 2005년 5월~2005년 8월: 한국과학기술연구원 (KIST) 연구원

2005년 9월~현재 : 인하대학교 전자전기공학부 조교수