

잡음 패턴의 지능적 추정을 통한 음질 개선 알고리즘

Tone Quality Improvement Algorithm using Intelligent Estimation of Noise Pattern

서정국 · 차형태

Joung-Kook Seo and Hyung-Tai Cha

숭실대학교 전자공학과

요 약

본 논문에서는 잡음에 열화된 대역에서 지능적으로 잡음의 패턴을 추정하여 지각 필터의 성능을 개선시켜 오디오 신호의 음질을 개선하는 알고리즘을 제안한다. 제안된 알고리즘은 묵음 구간에서 추정된 고정된 추정 잡음을 사용하는 기존의 방식을 사용하지 않고, 신호 구간마다 달라지는 신호의 세기와 잡음의 영향 정도를 고려하여 신호가 열세하고 잡음이 대부분을 차지하는 대역, 즉 잡음에 열화된 대역에서 잡음 패턴을 추정하여 신호를 열화시키는 잡음을 효과적으로 제어하는 방식으로 지각적으로 개선된 음질의 신호를 얻을 수 있다. 기존 방식과의 비교를 위해 다양한 신호 대 잡음 비(signal-to-noise ratio, SNR)에서 열화된 오디오 신호를 입력으로 사용하였다. 입력 SNR이 5dB, 10dB, 15dB와 20dB의 각각의 경우에 대하여 객관적인 평가인 세그멘탈 신호 대 잡음비(Segmental SNR, SSNR)와 잡음 대 마스크 비(Noise-to-mask ratio, NMR), 또한 주관적인 평가인 청감 테스트(Mean opinion score, MOS test)를 비교하였다. 그 결과, SSNR 측면에서는 각각의 경우에 대해 7.4dB, 6.8dB, 5.7dB, 5.1dB 그리고, NMR 측면에서는 15.7dB, 15.5dB, 15.2dB, 14.8dB 의 개선을 확인할 수 있었다. 또한 주관적인 측정 결과인 테스트의 개선 확인 및 음질 개선에 의한 성능을 확인할 수 있었다.

Abstract

In this paper, we propose an algorithm that improves a tone quality of a noisy audio signal in order to enhance a performance of perceptual filter using intelligent estimation of noise pattern from a band degraded by additive noise. The proposed method doesn't use the estimated noise which is obtained from silent range. Instead new estimated noise according to the power of signal and effect of noise variation is considered for each frame. So the noisy audio signal is enhanced by the method which controls a estimation of noise pattern effectively in a noise corruption band. To show the performance of the proposed algorithm, various input signals which had a different signal-to-noise ratio(SNR) such as 5dB, 10dB, 15dB and 20dB were used to test the proposed algorithm. we carry out SSNR and NMR of objective measurement and MOS test of subjective measurement. An approximate improvement of 7.4dB, 6.8dB, 5.7dB, 5.1dB in SSNR and 15.7dB, 15.5dB, 15.2dB, 14.8dB in NMR is achieved with the input signals, respectively. And we confirm the enhancement of tone quality in terms of mean opinion score(MOS) test which is result of subjective measurement.

Key Words : 잡음 패턴, 열화된 대역, 지각 필터, 잡음 추정, 잡음 이득 제어.

1. 서 론

최근 시스템 환경의 발전과 더불어 다양한 고품질의 음성 및 오디오 신호의 디지털 표현 기술의 발전으로 인해 열화된 신호 개선에 있어 정확한 의미 전달의 관점에서의 신호 명료도 개선만으로는 음질 측면에서 늘어나는 사용자의 요구 조건을 충족시키기에는 부족하게 되었다. 그 결과 최근 잡음에 오염된 신호의 개선을 위한 알고리즘에 있어 높은 SNR 개선과 더불어 신호에 대한 지각적인 측면에서도 무결성을 보

장하기 위해 다양한 방법이 제안되어 왔다[1-4]. 또한 인간의 청각 특성을 모델링하여 시간-주파수 영역의 특성을 이용한 방법[5-6], 청각 모델 기반의 마스크 특성을 이용한 방법[7-10] 등이 있다. 하지만, 위의 방법들은 잡음을 제거하기 위해서 잡음에 대한 통계적인 정보가 필요하고, 만일 그 정보를 알 수 없을 때는 보통 묵음 구간을 통해 잡음을 추정한다. 하지만, 정확한 묵음 구간 검출은 매우 어려울 뿐더러 잡음 추정 오차는 에너지 차감 방식의 음질 개선 성능에 많은 영향을 주며, 개선 결과 잔여 잡음 및 뮤지컬 톤 잡음 등의 부가적인 왜곡을 발생시키므로 이러한 추정 오차의 영향에 대한 보상이 요구된다.

본 논문에서는 인간의 청각 시스템에서 주파수 변별력을 나타내는 임계 대역에서의 청각 자극 에너지를 이용한다. 이를 이용한 에너지 차감 형식의 필터에 제안한 지능적인 잡음 패턴의 추정을 통해 얻은 추정 잡음을 적용하여, 열화된 오

접수일자 : 2004년 11월 2일

완료일자 : 2005년 3월 14일

감사의 글 : 본 연구는 한국과학재단 목적기초연구 R01-2002-000-00073-0(2004) 지원으로 수행되었습니다.

디오 신호를 잡음 에너지 제어 측면에서 보상해 줌으로써 신호의 음질을 개선하는 알고리즘을 제안한다.

제안하는 알고리즘에서는 백색 가우시안 잡음처럼 각 대역에 일정하게 영향을 주는 잡음인 경우 처리되는 매 구간마다 잡음에 열화된 대역에서 잡음이 신호에 비해 우세하다는 것을 고려[9]하여 임계 대역에서 잡음을 처리함으로써 지각 필터의 성능에 효과적인 개선뿐만 아니라 낮은 입력 신호 대 잡음비에서의 개선 효과도 확인할 수 있었다.

2. 청각 자극 에너지 차감 필터

일반적인 주파수 차감 형식의 음성 개선 알고리즘들은 다양한 환경에 대응할 수 있도록 자유 파라미터를 가진다. 이것들은 잡음 에너지 감소와 잔여 잡음 에너지의 존재 및 제어 과정에서 부수적으로 발생하게 되는 음성 신호의 원치 않는 왜곡 등의 사이에 존재하는 Tradeoff에 대한 Variation을 제공한다. 최근 인간의 청각적 지각 특성을 이용하여 잡음 에너지 차감 파라미터를 마스킹 특성에 근거해서 적용시키거나[7], 가청 잡음의 Psycho-acoustical noise shaping을 통해 잡음 에너지 감소[6]와 음성의 명료도(Intelligibility) 증가를 통해 많은 개선을 가져왔다.

임의의 시간 프레임 인덱스 i 에서의 잡음에 의해 열화된 신호와 원 신호의 전력 스펙트럼을 각각 $Y(\omega, i)$, $X(\omega, i)$ 라고 하면, 각각의 신호에 대한 청각 시스템에서의 지각 패턴에 대한 신호의 표현인 기저 막에서의 자극 에너지(Excitation energy)는 임의의 임계 대역(Critical band) 또는 바르(Bark) 인덱스, z 에 대해 각각 시간 주파수 지각 에너지 $Y_{cl}(z, i)$, $X_{cl}(z, i)$ 로 표현할 수 있다.

이때, 만약 임의의 음질 개선 처리를 통해서, 이러한 $Y_{cl}(z, i)$ 로부터 가장 가능한 잡음 에너지의 영향을 제거했다면, 개선된 신호의 전력 스펙트럼, $X(\omega, i)$ 에 의한 기저 막에서의 자극 에너지는 순수 신호의 지각 에너지와 일치하게 될 것이다.

$$X_{cl}(z, i) = X_{cl}(z, i) \quad (1)$$

이러한 신호의 지각적인 자극 에너지를 이용하여, 추정 잡음의 자극 에너지를 통한 음질 개선 필터를 이용한 방법이 제안되었고[6], 신호의 지각적인 응답 특성을 이용한 음질 개선 지각 필터, $H_c(z, i)$ 가 식 1을 만족하는 $X(\omega, i)$ 를 생성하고, 임의의 주파수 ω 에 대한 임출력 신호의 전력 스펙트럼을 임의의 임계대역 z 의 자극 에너지로 표현함으로써 추정 잡음을 이용한 신호의 자극 에너지 차감 형식의 필터는 다음과 같이 나타낼 수 있다.

$$H_c(z, i) = 1 - \frac{D_{cl}(z, i)}{Y_{cl}(z, i)}, \text{ for } 0 \leq z \leq Z-1 \quad (2)$$

여기서, Z 는 전체 임계 대역의 수를 나타낸다. 이 때, 임계 대역에서의 지각적으로 느끼는 잡음 에너지의 제어 필터는 시간 영역에서의 에너지 확산 영향이 없다고 가정함으로써[6], 다음과 같이 표현 할 수 있다.

$$H_c(z, i) = 1 - \frac{D_c(z, i)}{Y_c(z, i)} \text{ for } 0 \leq z \leq Z-1 \quad (3)$$

이때, 이러한 신호의 지각적 응답 특성을 이용한 음질 개

선 필터, $H_c(z, i)$ 를 통해 개선된 신호의 자극 에너지, $X(z, i)$ 는 다음과 같이 얻을 수 있다.

$$X(z, i) = H_c(z, i) Y(z, i) \quad (4)$$

식 4의 방법으로 개선된 신호는 추정 잡음의 자극 에너지, $D_c(z, i)$ 에 의해 개선 효과가 달라짐을 확인할 수 있다. 즉, 추정 잡음의 자극 에너지가 추정 오차가 없는 입력 SNR 20dB 환경에서는 잡음에 의해 열화된 신호보다 3dB정도의 개선된 신호를 얻는 반면, -3dB의 정도의 추정 오차가 있는 경우는 성능 개선이 약 SNR 1dB 정도로 나타났다

3. 지능적인 잡음 패턴 추정 알고리즘

기존 지각 필터에 적용되는 잡음은 식 5와 같이 잡음에 의해 열화된 신호의 세 프레임의 묶음 구간 평균을 이용하여 추정 잡음으로 사용하였다.

$$D(\omega, i) = \text{mean}\{Y(\omega, i)\}, \text{ for } 1 \leq i \leq 3 \quad (5)$$

여기서 구한 추정 잡음, $D(\omega, i)$ 를 자극 에너지 $D_c(z, i)$ 로 변환시켜 지각 필터에 적용 하지만, 이는 필터에 적용되는 구간마다 동일한 추정 잡음 에너지를 사용하게 되며, 실제로 첨가된 잡음 에너지와의 추정 오차가 -3dB 정도를 갖게 됨으로 지각 필터의 성능 저하를 가져오게 되는 주요 원인이 된다.

제안된 방법은 이러한 고정적인 추정 잡음을 사용하지 않고, 처리되는 프레임 단위로 잡음에 열화된 대역에서 지능적으로 잡음을 추정한다. 처리되는 매 프레임마다 가변적인 추정 잡음을 지각 필터에 적용하는 방법이다.

원 신호를 $x(n)$ 이라고 하면, 첨가 잡음, $d(n)$ 에 의해 열화된 신호는 다음과 같이 얻을 수 있다.

$$y(n) = x(n) + d(n) \text{ for } 0 \leq n \leq N-1 \quad (6)$$

여기서, 잡음에 의해 열화된 신호, $y(n)$ 에서 원 신호가 대부분을 차지하고, 잡음 신호가 열세한 대역을 $NL_{low}^d(\omega)$ 이라면 다음과 같은 관계를 얻을 수 있다[9].

$$\begin{cases} NL_{low}^d(\omega, i) = D(\omega, i), \\ NL_{low}^y(\omega, i) = Y(\omega, i) \neq NL_{low}^d(\omega, i) \end{cases} \quad (7)$$

(where, $0 \leq \omega < 5\text{kHz}$)

또한, 원 신호가 열세하고, 잡음이 대부분을 차지하는 대역 즉, 잡음에 열화된 대역을 $NL_{high}^y(\omega)$ 라고 하면, 다음과 같은 관계도 얻을 수 있다. 단, fs 는 44.1kHz의 샘플링율을 나타낸다.

$$\begin{cases} NL_{high}^d(\omega, i) = D(\omega, i) \\ NL_{high}^y(\omega, i) = Y(\omega, i) \approx NL_{high}^d(\omega, i) \end{cases} \quad (8)$$

(where, $10 \leq \omega < fs/2 \text{ kHz}$)

이는 잡음에 열화된 대역에서는 잡음에 의해 열화된 진폭 스펙트럼과 실제 잡음 신호의 진폭 스펙트럼은 그림 1과 같

이 거의 유사함을 알 수 있다.

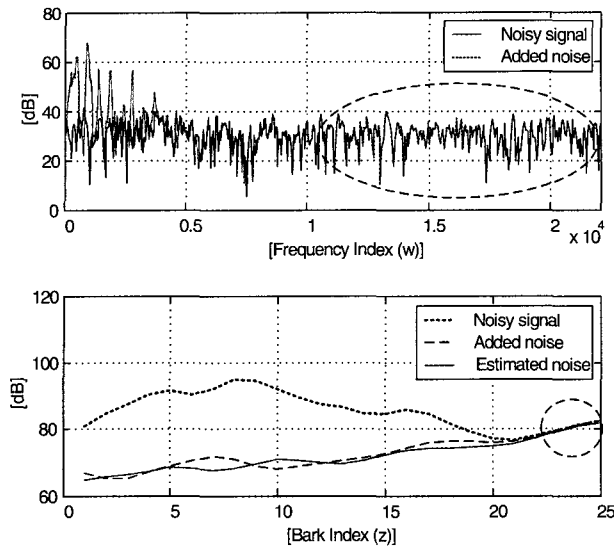


그림 1. 잡음에 열화된 대역
Fig. 1. Noise corruption band

여기서, \bigcirc 표시 되어 있는 부분은 주파수 영역과 바크 (Bark) 영역에서의 유사함을 보여주는 부분이다. 주파수 영역에서의 10~22.05kHz 의 범위는 바크 영역에서는 23~25Bark 로 측정된다. 이러한 특성을 이용하여 잡음에 열화된 대역에서 잡음을 추정하여 지각 필터에 적용되는 추정 잡음을 생성시키는 방법은 다음 식 9와 같다.

$$\begin{aligned} \overline{D}_{low}(\omega, i) &= \begin{cases} Y(\omega + 10, i), & 0 \leq \omega \leq 10 \text{ [kHz]} \\ 0, & \text{otherwise} \end{cases} \\ \overline{D}_{high}(\omega, i) &= \begin{cases} Y(\omega, i), & 10 \leq \omega \leq fs/2 \text{ [kHz]} \\ 0, & \text{otherwise} \end{cases} \\ \mathcal{D}(\omega, i) &= \overline{D}_{low}(\omega, i) + \overline{D}_{high}(\omega, i) \end{aligned} \quad (9)$$

여기서, $\overline{D}_{high}(\omega, i)$ 는 잡음에 열화된 대역에서 추출된 신호이며, 추정되어 사용될 잡음의 나머지 대역의 신호는 $\overline{D}_{low}(\omega, i)$ 이다. 결과적으로 1차로 추정된 잡음은 $\mathcal{D}(\omega, i)$ 이 된다[9].

하지만, 추정된 잡음은 열화된 대역에서 얻어진 잡음의 패턴을 이용하여 재구성된 것이므로, 저 대역 부분의 신호 추가로 인해 발생하는 마스킹 효과와 임계 대역 에너지 처리 과정으로 인해 고 대역 부분에서 과 추정(over estimation)이 발생하게 된다. 이러한 과 추정으로 인해 개선되는 신호의 왜곡을 초래하게 된다. 이러한 문제를 해결하기 위해 과 추정 부분의 이득을 제어함으로써 과 추정되는 부분을 보상해 주는 알고리즘을 추가한다. 그림 2는 제안된 잡음 패턴 추정 알고리즘을 보여준다.

그림 2에서 잡음 에너지 제어를 통한 잡음 추정 방법은 고 대역부분에 발생되는 과 추정 잡음을 제거하기 위한 방법으로 1차로 추정된 잡음의 자극 에너지, $\mathcal{D}(z, i)$ 가 잡음에 열화된 에너지보다 큰 구간이 4개 이상이 존재할 때 적용이 된다. 잡음 에너지를 제어시켜 주는 알고리즘은 다음 식과 같다.

$$\mathcal{D}(\omega, i) = \overline{D}_{low}(\omega, i) + \alpha \times \overline{D}_{high}(\omega, i) \quad (10)$$

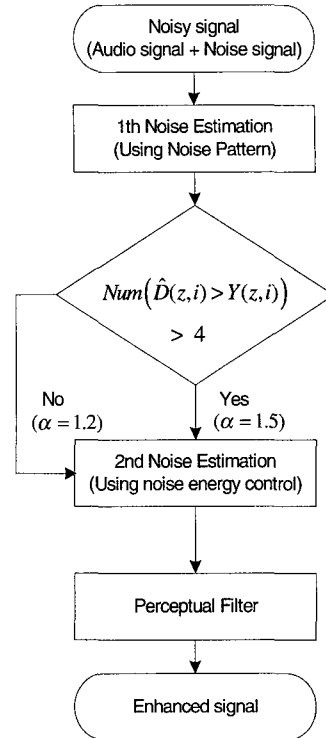


그림 2. 잡음 패턴 추정 알고리즘
Fig. 2. Estimation algorithm of noise pattern

여기서, α 는 잡음 에너지 제어 파라미터로 과 추정된 구간의 개수가 4개 이상인 경우는 $\alpha=1.5$, 이외에는 $\alpha=1.2$ 이다. 이 값은 실험으로 얻어낸 값으로, 잡음에 열화된 신호의 음압 레벨이 큰 경우에는 마스킹 효과에 의해 잡음의 영향이 작기 때문에, 고주파 부분의 임계 대역 내에서 과 추정이 발생하는 빈도가 낮아진다. 음압레벨이 작은 경우, 즉 구간이 4개인 경우를 8가지 다른 신호를 이용하여 실험한 결과를 구간 별 평균을 내어 찾아낸 실험치이다. 이렇게 얻어진 추정된 잡음의 전력 스펙트럼, $\mathcal{D}(\omega, i)$ 에 의한 기저 막에서의 자극 에너지는 $\mathcal{D}_e(z, i)$ 으로 표현할 수 있다. 제안된 방법으로 구해진 추정 잡음, $\mathcal{D}_e(z, i)$ 을 이용하여 지각 필터에 적용한다.

잡음에 열화된 신호의 자극에너지, $Y_e(z, i)$ 와 추정된 잡음의 자극 에너지, $\mathcal{D}_e(z, i)$ 를 식 3의 지각 필터에 적용함으로써, 필터 응답 및 개선된 신호의 자극 에너지를 구할 수 있다.

그림 3(상)의 필터 응답을 보면, 기존의 고정된 묵음 구간을 사용하는 방법을 이용한 지각 필터(conventional)는 첨가된 잡음을 그대로 사용한 최적화 필터(optimum)에 비해 고 대역의 잔여 잡음이 적게 제거되는 모습을 볼 수 있다. 이로 인해 기존의 자극 에너지 결과에서도 잡음이 각 대역에서 개선되지 않았음을 볼 수 있다.

하지만, 제안된 추정 잡음 방법을 이용한 지각 필터(proposed)를 적용한 것은 최적화 필터 응답과 거의 유사하게 잡음이 제거 되었으며, 이렇게 개선된 신호의 자극 에너지도 원 신호(clean)에 가까운 결과를 볼 수 있다.

그림 3는 입력 잡음 SNR이 20dB인 경우 측정된 결과 그림이다. 실험 결과 입력 잡음 SNR이 작아질수록 원 신호와의 차이가 나는 이유는 잡음의 양이 많아짐으로써, 대역 별

로 영향을 주는 잔여 잡음의 양이 많아 필터 적용 시 필터링의 양이 증가하기 때문이다. 이로 인해 음질에는 약간의 왜곡이 생기지만, 잔여 잡음의 제거로 인해 명료성의 저하나 귀의 피로감을 줄일 수 있다.

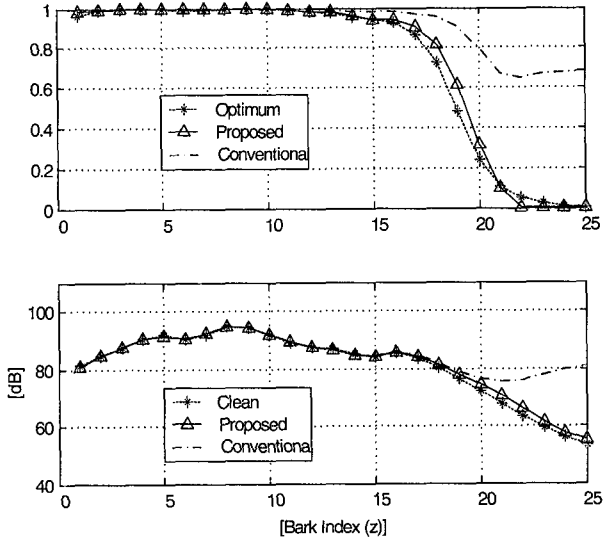


그림 3. 개선 신호의 필터 응답(상) 및 자극 에너지(하)
Fig. 3. Filter response(top) and excitation energy(bottom) of enhanced signal

이렇게 개선된 신호의 자극 에너지를 주파수 영역에서의 전력 스펙트럼 결과로 원 신호의 전력 스펙트럼과 유사함을 그림 4에서 다시 확인할 수 있다.

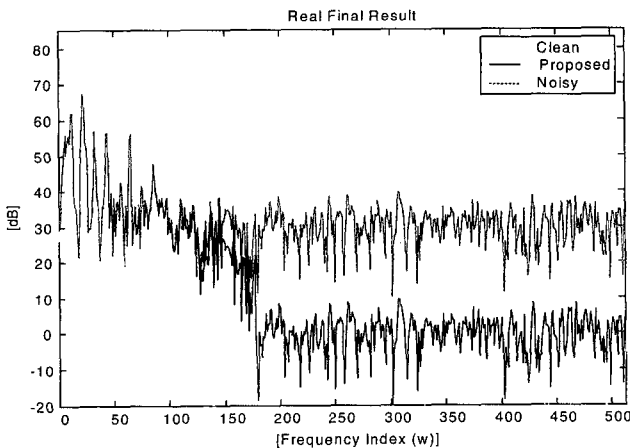


그림 4. 전력 스펙트럼 결과
Fig. 4. Result of power spectrum

4. 실험 결과

실험에 사용된 신호들은 일반 오디오 CD로부터 획득한 44.1kHz의 샘플링 율과 16bit/sample의 정밀도를 갖는 오디오 신호로 Classic, Jazz, Rock, Vocal등의 4가지 장르의 오디오 신호를 사용하였다.

신호를 열화 시키는 잡음으로는 백색 가우시안 잡음을 첨가시켜 입력 신호 대 잡음 비(Signal-to-Noise Ratio, SNR)

20dB, 15dB, 10dB, 5dB등의 입력 신호로 이용하였고, 해닝 윈도우 (Hanning Window) 함수와 오버랩 애드 (Overlap Add) 방식을 주파수축 변환에 사용하였다.

입력되는 열화된 신호의 음질 개선 결과를 보기 위해 다음의 4가지 필터를 적용 후 결과를 비교 하였다. 첫 번째 방법은 주파수 차감 방식인 잡음 에너지 차감 방식의 필터 (SPS, Simple Power Subtract)[1]를 사용 하였으며, 두 번째로는 기존의 복음 구간에서 잡음을 추정 한 후 지각 필터 (PF, Conventional method)[6]를 적용하였다. 세 번째 방법으로는 제안된 방법으로 잡음에 열화 된 대역에서 잡음을 추정한 후 지각 필터(PPF1, 1th proposed method)[9]에 적용 하였다. 최종적으로 제안한 방법은 잡음 패턴의 추정을 통해 지각 필터(PPF2, 2nd proposed method)에 적용하였다.

4.1 객관적 평가

각 방법으로 얻은 개선 결과에 대한 비교로는 객관적 측정(Objective Measurement) 결과로 사용되는 프레임 단위 신호 대 잡음 비(Segmental Signal-to-Noise Ratio, SSNR)와 잡음 대 마스크 비(Noise-to-mask ratio, NMR)를 비교한 결과는 표 1와 같다. 이는 입력 SNR 별로 측정된 결과이며, 입력된 원 신호는 클래식을 사용한 경우이다.

표 1. 객관적 측정 결과(클래식에 적용한 경우)

(a) 세그멘탈 신호 대 잡음 비 (dB)

(b) 잡음 대 마스크 비 (dB)

Table 1. Objective measurement result (in case of classic)

(a) Segmental SNR (dB)

(b) NMR (dB)

Input SNR	Method (SSNR)			
	SPS	PF	PPF1	PPF2
20 dB	19.23	19.27	23.07	23.30
15 dB	14.31	14.38	18.88	19.06
10 dB	9.35	9.47	14.41	14.87
5 dB	4.40	4.55	10.61	10.80

(a)

Input SNR	Method (NMR)			
	SPS	PF	PPF1	PPF2
20 dB	2.31	1.96	-9.78	-11.36
15 dB	7.30	6.95	-5.04	-6.75
10 dB	12.30	11.95	-0.16	-2.07
5 dB	17.30	16.94	4.77	2.79

(b)

표 1의 (a)에서 볼 수 있듯이 SSNR의 결과는 제안된 방식(PPF2)으로 개선된 신호는 기존의 PF인 경우 보다 대략 3dB이상이며, 입력 SNR이 작아질수록 PF와 PPF2의 SSNR 값의 차가 더 커짐을 알 수 있다. 이는 기존의 SPS 방법 및 PF 방법에 비해 제안된 방법을 적용한 결과의 성능이 개선되었음을 보여 준다. 또한 1차 잡음 추정만 사용한 PPF1보다는 SNR이 대략 0.3dB 정도 개선되었음을 확인할 수 있었다.

그리고, 표 1의 (b)는 잡음 에너지를 마스크 임계 값에 비교하여 나타내고 것으로 가청 가능한 잡음의 양을 나타내고 있다. 그러므로 NMR의 결과는 SSNR의 결과 값과는 반대로 작은 값이 나올수록 가청 가능한 잡음의 양이 적다는

의미이다. 입력 SNR이 작아질수록 PF와 PPF2의 차가 커질 수 있다. 이는 입력 SNR이 작을수록 첨가되는 잡음의 양이 많기 때문에, 잡음에 오염된 신호에서 제거된 잡음의 양이 기존의 방법보다 많이 필터링 되었음을 보여준다. 즉, 제안된 방법으로 개선된 신호에 존재하는 가청 가능한 잡음의 많은 양이 감소함 되었음을 알 수 있다. 또한 1차 잡음 추정만 사용한 PPF1보다는 NMR이 대략 2dB 정도 개선되었음을 확인할 수 있었다.

다음의 표 2는 입력되는 원 신호를 장르 별로 SSNR 과 NMR 을 측정한 결과로써, 개선된 양을 보기 위해 잡음에 열화된 신호에서 개선된 신호의 차로 개선량을 나타낸 표이다.

표 2. 장르 별 SSNR 과 SMR 개선 결과 비교(dB)
Table 2. Improvement SSNR and NMR results with various inputs (dB)

Input Wave		SSNR		NMR	
		Improvement	Improvement	Improvement	Improvement
		PF	PPF2	PF	PPF2
Classic	15 dB	1.07	5.74	1.56	15.27
	5 dB	1.24	7.49	1.58	15.72
Rock	15 dB	0.50	3.22	1.30	8.34
	5 dB	0.89	5.92	1.46	11.79
Jazz	15 dB	1.03	5.58	1.20	10.32
	5 dB	1.25	7.67	1.60	10.64
Vocal	15 dB	1.15	6.66	1.81	9.50
	5 dB	1.36	9.07	1.81	9.52

4.2 주관적 평가

위 표 2와 3에서의 객관적인 측정 결과에서의 개선 효과를 지각적으로 느낄 때의 개선 효과를 보기 위해 주관적 측정 방법(Subjective Measurement)인 청감 테스트 (Mean Opinion Score, MOS)를 실시하여 비교한 결과는 표 4와 같다. 입력된 신호로는 클래식 장르를 사용 하였으며, MOS score 와 각 방법들의 SNR 개선 량도 표기하였다.

-테스트는 20세 이상 오디오 전문분야에 종사하지 않는 남녀 30명을 대상으로 강의실 환경에서 10초의 분량의 샘플을 이용하여 수행하였다. 잡음에 의해 열화된 신호를 1이라고 하고 원음의 음질을 5라고 할 때 원음에 가까운 음질 정도를 기준으로 1과 5사이의 값으로 표현하도록 하여 수행하였으며, 주변 잡음과의 차음 확보를 위해 Etymotic Reaserch사의 Ear-Canal형 헤드폰 ER4S와 Beyerdynamic사의 레퍼런스 헤드폰 DT-880 및 HeadRoom사의 Headphone Amp를 이용하였다.

표 3. 주관적 측정 결과
Table 3. Objective measurement result

Method	in case of input SNR 15dB		in case of input SNR 5dB	
	MOS score	SNR improvement	MOS score	SNR improvement
	SPS	2.7	1.00 dB	1.4
PF(-3dB)	3.0	1.07 dB	1.9	1.24 dB
PF(0dB)	3.7	5.26 dB	3.1	6.84 dB
PPF2	4.3	5.74 dB	3.5	7.49 dB

표 3의 측정 결과에서 볼 수 있듯이 주파수 차감법을 이용

(SPS)한 결과와 목음 구간에서 잡음을 추정하여 사용한 PF(-3dB)의 경우 성능이 저하 되는 반면, 제안된 방법으로 잡음 에너지를 제어한 후 지각 필터에 추정한 잡음을 적용시킨 경우인 PPF2는 MOS score 뿐만 아니라 SNR 개선률에서도 우수한 성능을 보이고 있음을 알 수 있다. 또한 입력 SNR이 5dB인 경우처럼 낮은 입력 SNR 인 경우에도 기존의 방법보다 성능이 우수하였다.

5. 결 론

백색 가우시안 잡음을 적용하여 잡음에 의해 열화된 오디오 신호의 청각 자극 에너지를 이용하여 지각 필터에 적용함으로써 잡음 에너지를 줄이고, 음질 개선을 위해 잡음을 기존의 방법보다 정확하게 추정함으로써 지각 필터의 성능을 개선하여 음질을 개선하는 알고리즘을 제안하였다.

제안된 방식에 의해 지각 필터에 적용되는 추정 잡음을 유사하게 추정함으로써 효과적으로 잔여 잡음을 제거하여, 지각적으로 원 신호에 가까운 개선 신호를 얻을 수 있었다. 또한 잡음 에너지 차감 필터나 기존의 지각 필터와 비교할 때, SSNR과 NMR 비교, 그리고 청감 테스트 결과에서 추정 오차가 없는 경우의 잡음을 적용시킨 경우와 같이 좋은 성능의 음질 개선을 확인할 수 있었다.

하지만, 제안된 알고리즘으로는 저 대역에 많은 영향을 주는 잡음인 경우, 잡음 추정에 있어서 오차가 생김을 확인할 수 있었다. 이러한 추정 잡음의 오차로 인해 필터 적용 후에도 잔여 잡음이 존재하게 되어 음질의 명료도를 감소시키는 것을 확인할 수 있었다.

그러므로 향후 진행될 연구 과제는 잡음에 열화된 대역이 신호 성분이 많은 대역에 존재하게 하는 잡음을 추정 할 수 있는 방법에 대한 연구가 진행될 예정이다.

참 고 문 헌

- [1] S. F. Boll, "Suppression of acoustic noise speech using spectral subtraction", *IEEE Trans. Acoustic, Speech, and Signal Processing*, Vol ASSP-27, 1997.
- [2] R. J. MacAuly, "Speech enhancement using a soft-decision noise suppression filter", *IEEE Trans. Acoustic, Speech, and Signal Processing*, Vol 28, 1980.
- [3] Y. Ephraim and D. Malah, "Speech enhancement using a MMS error short-time spectral amplitude estimator", *IEEE Trans. Acoustic, Speech, and Signal Processing*, Vol 32, 1984
- [4] Y. Ephraim, "A Bayesian estimation approach for Speech enhancement using HMM", *IEEE Trans. Signal Processing*, Vol 40, 1992.
- [5] D. Tsoukalas, J. Mourjopoulos, and M. Paraskevas, "Speech enhancement using Psychoacoustic Criteria", *Journal of Audio Eng. Soc.*, vol. 45, no. 1/2, pp. 22-35, IEEE ICASSP, 1994
- [6] D. Tsoukalas, J. Mourjopoulos, and G. Kokkinakis, "Perceptual filters for audio signal enhancement," *Journal of Audio Eng. Soc.*, vol. 45, no. 1/2, pp.

22-35, Jan/Feb. 1997

[7] N. Virag, "Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System," *IEEE Trans. Speech and Audio Processing*, vol. 7, no. 2, pp. 126-137, Mar. 1999.

[8] 김현중, 차형태, "잡음 신호의 지각 패턴 제어를 통한 음질 개선 알고리즘 개발에 관한 연구", *한국음향학회 하계 학술대회 논문집* Vol 21, No. 1(s), pp. 199-202, July. 2002

[9] 서정국, 차혁근, 차형태, "잡음에 열화된 대역에서의 잡음 추정을 이용한 오디오 음질 개선". *한국음향학회 추계 학술대회 논문집* Vol 22, No. 2, pp. 293-296, Oct.~Nov. 2003

[10] 차형태, "가변 임계값을 이용한 지각 필터의 적응적인 음질 개선 알고리즘", *한국음향학회 논문지* Vol 23, No. 6, pp. 446-453, Aug. 2004



차형태 (Hyung-Tai Cha)

1993년 : The University of Pittsburgh (공학박사)

1993~96 : 삼성전자 신호처리 연구소 선임 연구원

1996~현재 : 숭실대학교 정보통신전자공학부 부교수

2001~현재 : 신호 및 시스템학회 이사

2001~현재 : 한국 방송공학회지 편집위원장, 편집위원

관심분야 : Multimedia Systems and Applications Audio and Video Signal Processing, Communication System, ASIC Implementation of Digital System,

Phone : 02) 820-0711

Fax : 02) 820-0711

E-mail : hcha@ssu.ac.kr

저 자 소 개



서정국 (Joung-Kook Seo)

2003년 : 숭실대 정보통신 전자공학부 졸업.

2003년~현재 : 동 대학원 전자공학과 석사과정

관심분야 : 오디오 및 음성 신호처리, 통신 신호처리

Phone : 02) 826 9063

Fax : 02) 826 9063

E mail : seojk@mms.ssu.ac.kr