

잡음 에너지 제어를 통한 지각 필터 성능 개선

Performance Improvement of Perceptual Filter Using Noise Energy Control

서정국*, 차형태*
(Joung-Kook Seo*, Hyung-Tai Cha*)

*송실대학교 전자공학과

(접수일자: 2004년 10월 15일; 수정일자: 2005년 1월 10일; 채택일자: 2005년 1월 17일)

본 논문에서는 잡음 에너지 제어를 통한 지각 필터의 성능을 향상시킴으로써 잡음에 의해 열화 된 오디오 신호의 음질을 개선하는 알고리즘을 제안한다. 기존의 방식에서는 묵음 구간에서 획득한 잡음 에너지를 사용하여 필터를 구성하여 사용하지만, 신호 구간마다 달라지는 신호의 세기 및 잡음의 환경 정도에 많은 영향을 받아 잡음의 에너지가 급격하게 변화한다면 음질의 개선률이 감소함을 알 수 있다. 그러나 제안하는 방식에서는 묵음 구간에서 추정된 잡음의 에너지 제어를 통해 초기 추정 잡음보다 가까운 추정 잡음을 얻음으로써 잡음 에너지가 급격하게 변화하여도 음질 개선률에는 변화가 적음을 알 수 있었다. 또한 저 대역에 영향을 미치는 잡음의 경우에도 다른 방법들과는 달리 음질의 개선이 뚜렷하였다. 기존 방식과의 비교를 위해 다양한 신호 대 잡음 비 (signal-to-noise ratio, SNR)에서 열화 된 오디오 신호를 입력으로 사용하였다. 입력 SNR이 5dB, 10dB, 15dB와 20dB의 각각의 경우에 대하여 SSNR (Segmental SNR)과 잡음 대 마스크 비 (Noise-to-mask ratio, NMR), 음질 테스트를 수행한 결과, 청감 테스트 (Mean Opinion Score, MOS Test) 결과의 향상과 음질의 개선을 확인할 수 있었다.

핵심용어: 자극 에너지, 지각 필터, 오디오 음질 개선, 잡음 에너지 제어

투고분야: 음향 신호처리 분야 (1.4), 음악음향 및 음향심리 분야 (8.7)

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 noise energy control. Most of the algorithms which were proposed by the other researchers usually applied a filter using the noise energy acquired from a silent range. In this case, the improvement rate of tone quality decreases if the noise energy is changed by the magnitude or environment variation in a signal frame. But the proposed method provides the means to find a good estimated noise through energy control of the estimated noise which is obtained from a silent range. Also we can get the enhancement of tone quality in low frequency band unlike other methods. 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. With the proposed algorithm, we could confirm the enhancement of tone quality in terms of segmental SNR (SSNR), noise-to-mask ration (NMR) and mean opinion score (MOS) test.

Keywords: Excitation energy, Perceptual filter, Audio signal enhancement, Noise energy control

ASK subject classification: Acoustic Signal Processing (1.4), Musical Acoustics and Psychoacoustics (8.7)

I. 서론

기술의 발전에 따라 고음질을 요구하는 시스템의 개발이 활발히 이루어지고 있으며, 이를 위해 음질 개선을

위한 고성능 음성 잡음처리 알고리즘들이 많이 연구되었다. 일반적으로 잡음은 명료성을 떨어뜨리고, 청자에게 불쾌감이나 귀의 피로감을 안겨주며 이러한 문제로 인하여 잡음을 줄이기 위한 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[7-9] 등이 있다. 하지만, 위

책임저자: 차형태 (hcha@ssu.ac.kr)
서울시 동작구 상도동 1-1 송실대학교 전자공학과
멀티미디어시스템 연구실
(전화: 02-820-0711; 팩스: 02-820-0711)

의 방법들은 잡음을 제거하기 위해서 잡음에 대한 통계적인 정보가 필요하고, 만일 그 정보를 알 수 없을 때는 보통 묵음 구간에서 잡음을 추정하여 사용한다. 하지만, 정확한 묵음 구간 검출은 매우 어려울 뿐만 아니라 잡음 추정 오차는 에너지 차감 방식의 음질 개선 성능에 많은 영향을 주게 되어 개선 결과 잔여 잡음 및 뮤지컬 톤 잡음 등의 부가적인 왜곡을 발생시키므로 이러한 추정 오차의 영향에 대한 보상이 필요하다. 또한 백색 가우시안 잡음 (White Gaussian Noise)에서는 성능이 우수하지만, 저 대역에 상대적으로 영향을 많이 미치는 잡음인 경우는 음질 개선이 저하됨을 볼 수 있다.

본 논문에서는 청각 자극 에너지를 이용한 에너지 차감 형식의 필터를 정적인 통계적 특성 (stationary static characteristic)을 갖는 잡음 환경에서 잡음에 열화 된 오디오 신호를 잡음 에너지 제어 측면에서 보상해 줌으로써 신호의 음질을 개선하는 방법을 제안하고 있다.

제안하는 알고리즘은 묵음 구간에서의 잡음 검출에 따른 잡음의 청각 자극 에너지의 이득 오차로 인해 발생하는 지각 필터의 성능 저하를 줄이기 위해 잡음 에너지의 제어를 통해 임계대역 에너지 영역에서 잡음을 처리하는 새로운 알고리즘을 적용함으로써 지각 필터의 성능을 크게 개선하였다.

II. 지각 패턴을 이용한 자극 에너지 차감 필터

일반적인 주파수 차감 형식의 음성 개선 알고리즘들은 다양한 환경에 대응할 수 있도록 자유 파라미터를 통해, 잡음 에너지 감소와 잔여 잡음 에너지의 존재 및 제어 과정에서 부수적으로 발생하게 되는 음성 신호의 원치 않는 왜곡 등의 사이에 존재하는 Tradeoff에 대한 Variation을 제공하게 되는데, 최근 인간의 청각적 지각 특성을 이용하여 잡음 에너지 차감 파라미터를 마스킹 특성에 근거해서 적응시키거나[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 열화 된 신호와 원 신호의 전력 스펙트럼을 각각 $Y(\omega, i)$, $X(\omega, i)$ 라고 하면, 각각의 신호에 대한 청각 시스템에서의 지각 패턴에 대한 신호의 표현인 기저막에서의 자

극 에너지 (Excitation energy)는 임의의 임계 대역 (Critical band) 또는 바크 (Bark) 인덱스, z 에 대해 각각 시간-주파수 지각 에너지 $Y_e(z, i)$, $X_e(z, i)$ 로 표현할 수 있다.

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

$$\hat{X}_e(z, i) = X_e(z, i) \tag{1}$$

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

$$H_e(z, i) = 1 - \frac{\hat{D}_e(z, i)}{Y_e(z, i)}, \text{ for } 0 \leq z \leq Z-1 \tag{2}$$

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 표현할 수 있다.

$$H_e(z, i) = 1 - \frac{\hat{D}_e(z, i)}{Y_e(z, i)} \text{ for } 0 \leq z \leq Z-1 \tag{3}$$

이때, 이러한 신호의 지각적 응답 특성을 이용한 음질 개선 필터, $H_e(z, i)$ 를 통해 개선된 신호의 자극 에너지, $\hat{X}(z, i)$ 는 다음과 같이 얻을 수 있다.

$$\hat{X}_e(z, i) = H_e(z, i) Y_e(z, i) \tag{4}$$

식 4의 방법으로 개선된 신호는 추정 잡음의 자극 에너지, $\hat{D}_e(z, i)$ 에 의해 개선 효과가 틀려짐을 확인할 수 있다. 즉, 추정 잡음의 자극 에너지가 추정 오차가 없

는 입력 SNR 20dB 환경에서는 잡음에 의해 열화 된 신호보다 3dB정도의 개선된 신호를 얻는 반면, -3dB의 정도의 추정 오차가 있는 경우는 성능 개선이 약 SNR 1dB 정도로 나타났다.

III. 잡음 에너지 제어 알고리즘

3.1. 잡음 신호의 분석

원 신호는 44.1kHz의 샘플링 율과 16Bit/sample의 해상도를 가진 오디오 신호이며, 잡음은 정적인 통계적 특성을 갖는 백색 가우시안 잡음과 헬리콥터 잡음을 각각 신호에 더해 줌으로써 신호의 SNR 이 20dB, 15dB, 10dB 와 5dB 되도록 설정하였다. 그림 1은 첨가 잡음으로 사용된 신호의 스펙트럼을 보여준다.

그림 1의 두 잡음은 잡음의 변화 (variance)가 없는 정적인 통계적 특성 (stationary static characteristic)을 볼 수 있다. 하지만, 두 잡음의 자극 에너지를 살펴보면 백색 가우시안 잡음인 경우는 각 대역에 일정하게 영향을 주는 반면에 일정한 속도로 비행중인 경우에 기내에 영향을 미치는 프로펠러의 소리와 같은 헬리콥터 잡음은 저 대역에 많은 영향을 주는 것을 알 수 있다. 이는 기존의 자각 필터가 백색 가우시안 잡음에서의 성능은 우수하지만, 저 대역에 많은 영향을 끼치는 잡음인 경우 자각필터가 잘 작동하지 않아서 신호 개선의 효과가 감소하는 원인이 된다. 그림 2는 두 잡음의 전력 스펙트럼과 자극 에너지를 보여준다.

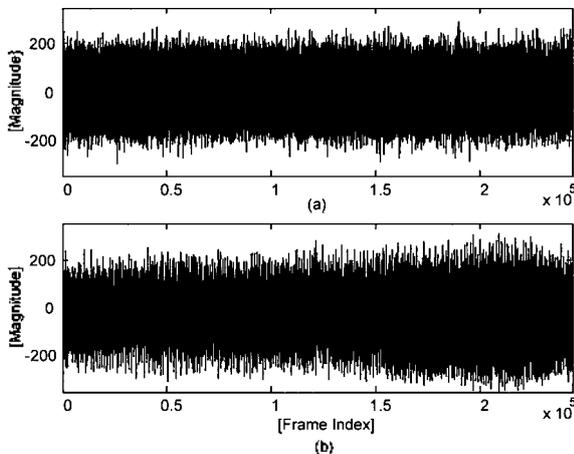


그림 1. 첨가 잡음 신호의 스펙트럼
(a) 백색 가우시안 잡음, (b) 헬리콥터 잡음
Fig. 1. Spectrum of added Noise.
(a) White gaussian noise. (b) Helicopter noise.

그림 2의 (a)에서는 이 두 잡음의 전력 스펙트럼 상의 차이를 볼 수 있다. 백색 가우시안 잡음은 전 대역에 걸쳐 영향을 주지만, 헬리콥터 잡음인 경우 저 대역에 많은 영향을 주는 것을 볼 수 있다. 이러한 잡음의 분석을 위해 임계 대역 (Critical Band) 에너지에 의한 지각적인 확산 영향을 나타내는 자극 에너지를 이용한다. 이는 첨가 잡음에 대해 더욱 정확한 추정 환경을 제공해 주게 되어 주파수 영역에서의 에너지 차감 방식에 비해 높은 개선 결과를 나타내기 때문이다[8]. 그림 2의 (b)는 이러한 두 첨가 잡음의 자극 에너지를 보여주고 있다.

3.2. 잡음 에너지 제어 기법

일반적으로 기존 자각 필터에 적용되는 잡음은 식 5와 같이 잡음에 의해 열화 된 신호에서 세 프레임의 목음 구간 평균을 이용하여 추정 잡음으로 사용하였다.

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

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

제안된 방법은 이러한 고정적인 추정 잡음을 사용하지 않고, 입력되는 열화 된 신호의 전력 (power)과 잡음의 variance를 이용하여 첨가된 잡음의 신호 대 잡음비

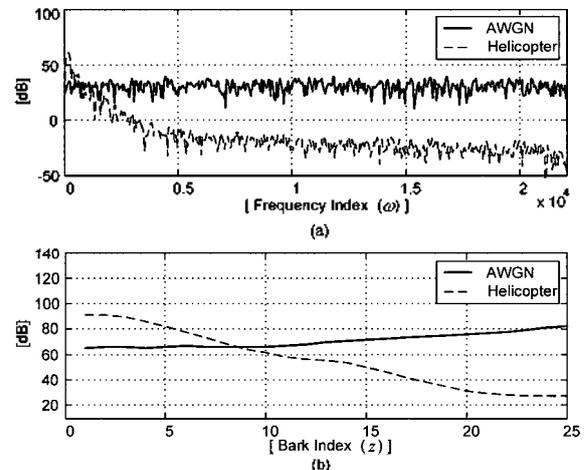


그림 2. 첨가 잡음 신호의 전력 스펙트럼과 자극 에너지
(a) 전력 스펙트럼, (b) 자극 에너지
Fig. 2. Power spectrum and excitation energy of noise.
(a) Power spectrum, (b) Excitation energy.

(signal-to-noise, SNR)를 추정한다. 이는 첨가된 잡음의 에너지가 급격하게 변화하였을 때 지각 필터 성능 저하의 원인이 되므로, 추정 입력 SNR 을 이용하여 첨가되는 잡음 에너지를 추정하기 위한 작업이다.

추정된 SNR 에 따른 임계치 (threshold)를 구하여, 대역 별로 control 상수를 적용함으로써 일차로 추정된 잡음의 에너지를 제어하게 됨으로써 적용된 추정 잡음을 생성하게 된다. 제한한 잡음 에너지 제어 알고리즘은 그림 3과 같다.

원 신호를 $x(n)$ 이라고 하면, 첨가 잡음, $d(n)$ 에 의해 열화 된 신호는 다음 식 6과 같이 얻을 수 있다.

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

여기서, 잡음에 열화 된 신호, $Y(\omega, i)$ 의 묵음 구간에서 추정된 잡음, $\bar{D}(\omega, i)$ 를 이용한다. 이때, 첨가되는 잡음에 대한 정보를 알 수가 없으므로 첨가되는 잡음의 SNR이 작을수록 신호를 열화 시키는 정도가 커짐으로써 지각 필터를 통해 처리되더라도 잔여 하는 잡음의 양이 많아서 개선의 효과를 볼 수 없다. 이러한 환경에서 열화 된 신호, $Y(\omega, i)$ 와 추정된 잡음, $\bar{D}(\omega, i)$ 을 이용하여 잡음의 SNR 을 추정함으로써 첨가되는 잡음 에너지의 변화량을 미리 예측함으로써 잔여 잡음을 제거하여 신호를 개선시키는데 필요한 정보로 사용된다. 첨가되는 잡음의 SNR을 추정하는 방법은 계산에 사용되는 한 프레임의 샘플수가 N이라고 할 때 식 7과 같다.

$$NS = \frac{1}{N} \cdot \sum_{i=1}^N (Y)^2$$

$$NO = \frac{1}{N} \cdot \sum_{i=1}^N \left(\frac{\bar{D} \cdot \text{std}(\bar{D})}{D(1:N)} \right)^2 \quad (7)$$

$$ES = 10 \cdot \log_{10} \left(\frac{NS}{NO} \right)$$

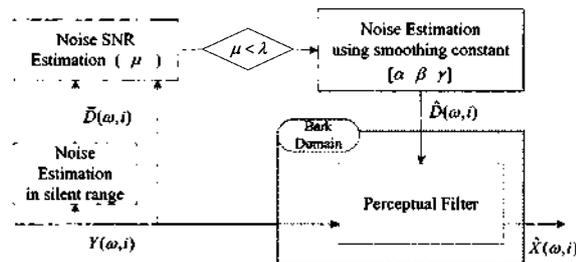


그림 3. 잡음 에너지 제어 알고리즘
Fig. 3. Noise energy control algorithm.

표 1. 임계치와 control 상수

Table 1. Threshold and control constants.

Estimate SNR (λ)	White Gaussian Noise			Helicopter Noise		
	α	β	γ	α	β	γ
10	1.8	1.9	1.9	2.0	1.0	0.5
15	1.7	1.8	1.9	1.8	1.0	0.3
20	1.6	1.7	1.9	1.5	0.5	0.2
25	1.5	1.6	1.9	1.5	0.3	0.1

여기서, NS 는 잡음에 열화 된 신호의 평균 전력 값이며, NO 는 묵음 구간에서 추정된 잡음의 변화량 값이고, std는 표준편차를 나타낸다. ES 은 추정된 SNR 값으로 그림 3에서의 μ 로 정의된다. 이런 μ 값에 따라 첨가된 잡음의 에너지 변화량을 20dB, 15dB, 10dB, 5dB로 추정하게 된다. 이러한 과정은 입력 잡음의 SNR 에 따라 필터 응답에 많은 영향을 주는 원인이 되기 때문에 잡음 환경에 맞추어 적당한 control 상수 값을 부여해 주기 위함이다.

위의 μ 값을 이용하여 대역 별로 control 상수 값인 $[\alpha \beta \gamma]$ 로 잡음 에너지 제어 값을 부여하여 새롭게 추정된 잡음, $\hat{D}(\omega)$ 를 생성 시키는 과정은 식 8과 같다.

$$\begin{cases} \xi_L(\omega) = \alpha \cdot \bar{D}(\omega), & (1 \leq \omega \leq LF) \\ \xi_M(\omega) = \beta \cdot \bar{D}(\omega), & (LF < \omega \leq MF) \\ \xi_H(\omega) = \gamma \cdot \bar{D}(\omega), & (MF < \omega) \end{cases} \quad (8)$$

$$\hat{D}(\omega) = \xi_L(\omega) + \xi_M(\omega) + \xi_H(\omega)$$

여기서, 그림 3에 사용되는 λ 는 추정된 잡음의 SNR 의 임계치이며, α, β, γ 는 대역 별로 적용되는 control 상수 값이다. 잡음 에너지 제어를 정확히 하기 위해 대역별로 나누어 적용하게 된다. 이 때, LF와 MF 는 대역을 나누어 주는 임계치 값으로 백색 가우시안 잡음인 경우 LF=1000Hz (10Bark), MF=2500Hz (15Bark) 로 사용하였으며, 헬리콥터 잡음인 경우는 LF=2500Hz (15Bark), MF=7000Hz (20Bark)로 사용하였다. 이는 잡음의 통계적 특성을 알기 위한 실험을 통해 적용된 값으로, 각 잡음이 대역별로 미치는 영향이 다르므로 대역 별로 적용되는 잡음 제어 값을 가변적으로 적용하기 위함이다. 다음 표 1은 제한한 알고리즘에 사용된 임계치 값과 control 상수 값이다.

이러한 제안된 잡음 제어 알고리즘으로 추정된 잡음 신호의 자극 에너지는 아래 그림 4와 같다.

위의 그림은 입력 SNR이 15dB로 추정 되었을 때의 결과 그림이다. 실제로 예측되는 SNR이 작아질수록 열화시키는 잡음의 양이 많아짐으로써, 에너지의 크기가 증가하게 된다. 이는 저 대역에 영향을 주는 잡음 환경에서는 잡음 추정이 목음 구간을 이용하는 것만으로 잔여 잡음을 제거 할 수가 없음을 보여준다.

그림 4의 (b)의 경우 기존의 목음 구간을 이용한 에너지는 고 대역에서도 차이를 보일 뿐만 아니라, 저 대역에서도 실제로 첨가된 잡음과 비교해 보아도 차이가 날 수 있다. 하지만, 제안된 알고리즘으로 이용하여

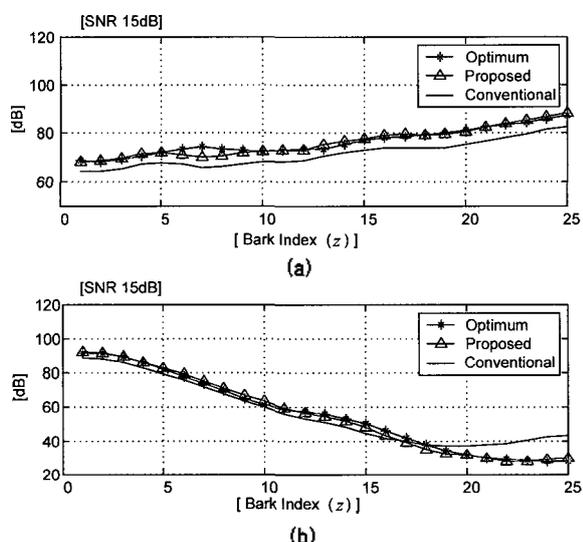


그림 4. 추정 잡음의 자극 에너지 (입력 SNR=15dB)
 (a) 백색 가우시안 잡음, (b) 헬리콥터 잡음
 Fig. 4. Excitation energy of estimated noise (SNR 15dB).
 (a) White gaussian noise, (b) Helicopter noise.

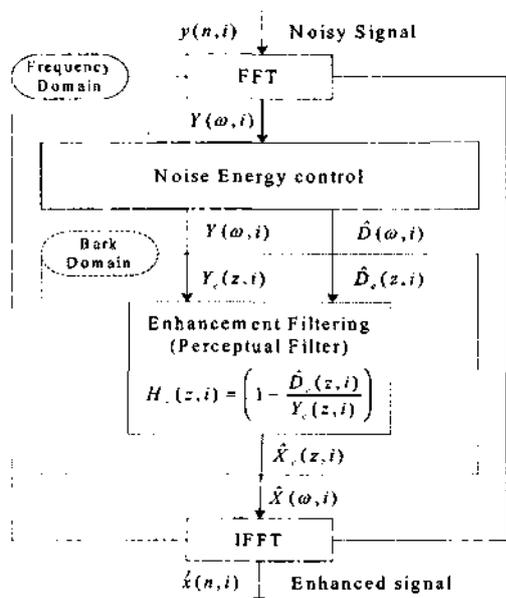


그림 5. 지각 필터의 적용
 Fig. 5. Application of perceptual filter.

추정 된 잡음은 잡음의 입력 SNR이 낮은 경우에도 적용적으로 실제 첨가된 잡음을 비슷하게 추정함을 볼 수 있다. 이렇게 추정된 잡음은 지각 필터에 적용 시 잔여 잡음을 기존의 방법보다 더 제거할 수 있도록 만들어 주는 계기가 된다.

3.3. 지각 필터의 적용 방법

제안된 방법으로 구해진 추정 잡음, $\hat{D}(\omega, i)$ 를 이용하여 지각 필터에 적용하는 개념도는 그림 5와 같다.

지각 필터에 적용하기 이전에 추정 잡음, $\hat{D}(\omega, i)$ 과 잡음에 의해 열화 된 신호, $Y(\omega, i)$ 의 자극 에너지를 구하면, $\hat{D}_c(z, i)$, $Y_c(z, i)$ 와 같이 구할 수 있다. 이런 자극 에너지를 식 4에서 주어진 지각 필터에 적용함으로써, 그림 6, 7과 같은 필터 응답 및 개선된 신호의 자극 에너지를 구할 수 있다.

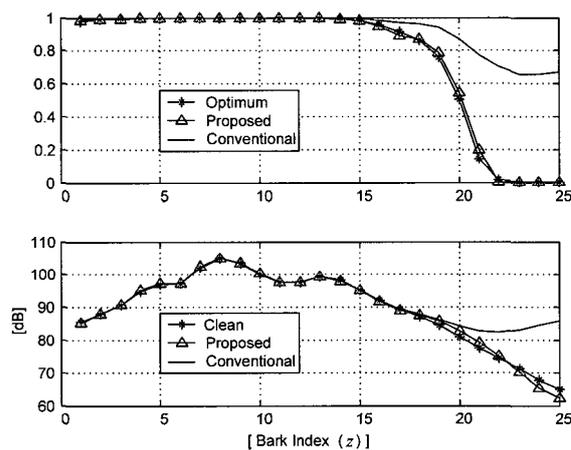


그림 6. 백색 가우시안 잡음 하에서 개선 신호의 필터 응답 (상)과 자극 에너지(하)
 Fig. 6. Filter response (top) and excitation energy (bottom) of enhanced signal with white gaussian noise.

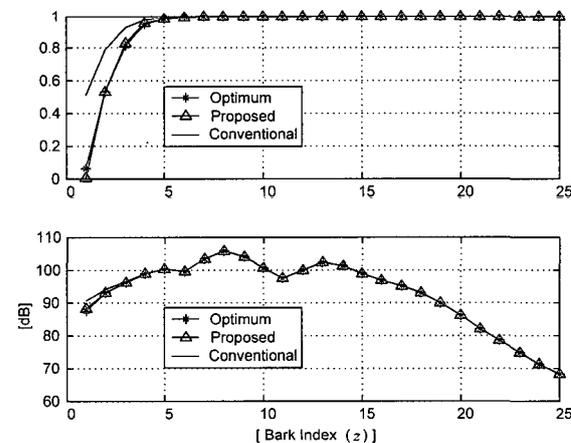


그림 7. 헬리콥터 잡음 하에서 개선 신호의 필터 응답(상)과 자극 에너지(하)
 Fig. 7. Filter response (top) and excitation energy (bottom) of enhanced signal with helicopter noise.

표 2. 객관적 측정 결과 (클래식에 적용한 경우)

- (a) 세그멘탈 신호 대 잡음비 (dB)
- (b) 잡음 대 마스킹 비 (dB)

Table 2. Objective measurement result (in case of Classic).

- (a) Segmental SNR (dB)
- (b) NMR (dB)

Input SNR	White Gaussian Noise				Helicopter Noise			
	SPS	PF	APF	PPF	SPS	PF	APF	PPF
20 dB	19.23	19.27	22.02	23.22	20.13	20.49	20.55	21.84
15 dB	14.31	14.38	17.25	19.02	16.90	16.63	16.72	18.74
10 dB	9.35	9.47	11.12	14.84	12.54	12.89	12.90	15.73
5 dB	4.40	4.55	7.89	10.78	8.87	8.99	9.12	12.42

(a)

Input SNR	White Gaussian Noise				Helicopter Noise			
	SPS	PF	APF	PPF	SPS	PF	APF	PPF
20 dB	2.31	1.96	-11.01	-12.04	-15.53	-14.91	-14.84	-18.55
15 dB	7.30	6.95	-5.20	-7.57	-11.10	-11.20	-11.82	-16.55
10 dB	12.30	11.95	1.35	-2.88	-6.32	-7.01	-7.56	-14.19
5 dB	17.30	16.94	4.05	1.90	-1.42	-2.39	-3.02	-12.64

(b)

그림 6은 백색 가우시안 잡음을 적용 하였을 때의 결과 그림이다. 필터 응답을 보면, 기존의 지각 필터는 첨가된 잡음을 그대로 사용한 최적화 필터에 비해 고 대역의 잔여 잡음이 적게 제거되는 모습을 볼 수 있다.

이와 반대로 그림 7의 경우는 저 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 필터를 적용한 것은 최적화 필터 응답과 거의 유사하게 잡음이 제거 되었으며, 이렇게 개선된 신호의 자극 에너지도 원 신호에 가까운 결과를 볼 수 있다.

그림 6과 그림 7은 입력 잡음 SNR이 15dB인 경우 측정된 결과 그림이다. 실험 결과 입력 잡음 SNR이 작아질수록 원 신호와의 차이가 나는 이유는 잡음의 양이 많아짐으로써, 대역 별로 영향을 주는 잔여 잡음의 양이 많아 필터 적용 시 필터링의 양이 증가되었기 때문이다. 이로 인해 음질에는 약간의 왜곡이 생기지만, 잔여 잡음의 제거로 인해 명료성의 저하나 귀의 피로감을 줄일 수 있다.

IV. 실험 결과

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

지 장르의 오디오 신호를 사용하였다.

신호를 열화 시키는 잡음으로는 정적인 통계적 특성을 갖는 백색 가우시안 잡음과 헬리콥터 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 열화 된 신호의 음질 개선 결과를 보기 위해 다음의 3가지 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 묵음 구간에서 잡음을 추정한 후 지각 필터 (PF, Conventional method)[6]를 적용하였다. 세 번째 방법으로는 제안된 방법으로 잡음에너지를 제어한 후 지각 필터에 적용 (PPF, proposed method)하였다.

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 마스킹 비 (Noise-to-mask ratio, NMR)를 비교한 결과는 표 2와 같다. 이는 입력 SNR 별로 측정된 결과이며, 입력된 원 신호는 클래식을 사용한 경우이다.

표 2의 (a)에서 볼 수 있듯이 SSNR의 결과는 제안된 방식 (PPF)으로 개선된 신호는 기존의 PF인 경우 보다 대략 3dB이상이며, 입력 SNR이 작아질수록 PF와 PPF의 SSNR 값의 차가 더 커짐을 알 수 있다. 이는 기존의

SPS 방법 및 PF 방법, APF 방법에 비해 제안된 방법을 적용한 결과의 성능이 많이 개선되었음을 보여 준다. 특히 헬리콥터 잡음인 경우 기존의 방법들의 결과값들은 거의 유사하지만, 제안한 방법으로 처리한 결과는 입력 잡음 SNR이 20dB, 15dB, 10dB, 5dB인 경우에 각각 1.29dB, 2.02dB, 2.83dB, 3.3dB의 개선 효과를 얻을 수 있다. 이는 잡음에 열화 된 정도가 강할수록 개선의 효과가 큼을 확인할 수 있다.

그리고, 표 2의 (b)는 잡음 에너지를 마스킹 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 PF, APF와 PPF의 차가 커짐을 알 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 감소되었음을 알 수 있다. 백색 가우시안 잡음인 경우 개선 효과를 볼 수 있었지만, 헬리콥터 잡음인 경우 APF의 방법과 비교해보면, 입력 잡음의 SNR이 작아질수록 개선률에서 각각 3.71dB, 4.43dB, 6.63dB, 9.62dB의 큰 개선 효과를 확인할 수 있다.

다음의 표 3은 입력되는 원 신호를 장르별로 SSNR 과 NMR 을 측정한 결과로써, 개선된 양을 보기 위해 잡음에 열화 된 신호에서 개선된 신호의 차로 개선량을 나타낸 표이다.

위 표 3에서는 입력 잡음의 열화 된 정도가 클수록 개선 효과를 보이기 위해 입력 잡음 SNR이 15dB와 5dB인 경우에 측정된 결과를 보여준다. SSNR 개선량 중 Rock 장르는 헬리콥터 잡음을 이용한 결과 전체적으로 낮은 수치의 결과를 얻었다. 이는 다른 장르에 비해 잡음의 주파수 특징이 유사하고 타악기등의 세기가 센 악기에 의해 추정 잡음 오차에 따른 성능 차이나 필터간의 성능 차이가 거의 없어 전체적으로 낮은 수치를 보임을 확인할 수 있다. 음원이 잡음과 유사한 특징을 지닌다면 필터의 성능을 비교하는데 부적합 함을 알 수 있다. 하지만, NMR 개선량 측면에서는 모든 장르별로 개선된 신호에 존재하는 가청 가능한 잡음의 많은 양이 감소한 결과를 확인할 수 있다.

위 표 2와 3에서의 객관적인 측정 결과에서의 개선 효과를 자각적으로 느낄 때의 개선 효과를 보기위해 주관

표 3. 장르 별 개선 결과 비교
(a) 세그멘탈 신호 대 잡음비
(b) 잡음 대 마스킹 비

Table 3. Improvement results with various inputs.

(a) Segmental SNR

(b) NMR

Input Wave		White Gaussian Noise		Helicopter Noise	
		PF	PPF	PF	PPF
Classic	15 dB	1.07	5.71	2.46	4.56
	5 dB	1.24	7.47	4.02	8.25
Rock	15 dB	0.5	2.89	0.15	0.26
	5 dB	0.89	5.61	0.16	0.46
Jazz	15 dB	1.03	5.51	0.42	0.68
	5 dB	1.25	7.66	2.43	2.65
Vocal	15 dB	1.15	7.39	0.26	0.54
	5 dB	1.36	9.02	1.50	3.07

(a)

Input Wave		White Gaussian Noise		Helicopter Noise	
		PF	PPF	PF	PPF
Classic	15 dB	1.56	16.01	6.89	12.24
	5 dB	1.58	16.58	8.06	18.31
Rock	15 dB	1.30	8.12	0.63	1.12
	5 dB	1.46	11.12	2.37	4.82
Jazz	15 dB	1.20	10.95	2.65	6.20
	5 dB	1.60	13.18	3.07	10.29
Vocal	15 dB	1.81	11.21	3.47	8.85
	5 dB	1.81	9.81	3.55	13.13

(b)

적 측정 방법 (Subjective Measurement)인 청감 테스트 (Mean Opinion Score, MOS)를 실시하여 비교한 결과는 표 4와 같다. 입력된 신호로는 클래식 장르를 사용하였으며, MOS score 와 각 방법들의 SNR 개선량도 표기하였다.

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

표 4의 측정 결과에서 볼 수 있듯이 주파수 차감법을 이용 (SPS)한 결과와 묵음 구간에서 잡음을 추정하여 사용한 PF (-3dB)의 경우 성능이 저하되는 반면, 제안된 방법으로 잡음 에너지를 제어한 후 지각 필터에 추정된 잡음을 적용시킨 경우인 PPF는 MOS score 뿐만 아니라 SNR 개선률에서도 우수한 성능을 보이고 있음을 알 수

표 4. 주관적 측정 결과

(a) 입력 신호 대 잡음 비 15dB 인 경우
 (b) 입력 신호 대 잡음 비 5dB 인 경우

Table 4. Objective measurement result.

(a) In case of input SNR 15dB
 (b) In case of input SNR 5dB

Enhancement method	White Gaussian Noise		Helicopter Noise	
	MOS Score	SNR improvement	MOS Score	SNR improvement
SPS	2.7	1.00 dB	2.6	3.73 dB
PF (-3dB)	3.0	1.07 dB	3.2	2.53 dB
PF (0dB)	3.7	5.26 dB	3.8	3.61 dB
PPF	4.1	5.71 dB	4.0	4.64 dB

(a)

Enhancement method	White Gaussian Noise		Helicopter Noise	
	MOS Score	SNR improvement	MOS Score	SNR improvement
SPS	1.4	1.09 dB	1.3	4.70 dB
PF (-3dB)	1.9	1.24 dB	1.6	4.82 dB
PF (0dB)	3.1	6.84 dB	3.3	7.46 dB
PPF	3.3	7.47 dB	3.4	8.25 dB

(b)

있다. 또한 입력 SNR이 5dB인 경우처럼 낮은 입력 SNR인 경우에도 기존의 방법보다 성능이 우수하였다.

V. 결론

통계적 특성을 갖는 백색 가우시안 잡음과 실험에 사용된 헬리콥터 잡음을 적용하여 잡음에 의해 열화 된 오디오 신호의 청각 자극 에너지를 이용하여 지각 필터에 적용함으로써 잡음 에너지를 줄이고, 음질 개선을 위해 잡음을 기존의 방법보다 정확하게 추정함으로써 지각 필터의 성능을 개선하는 방법을 제안하였다.

제안된 방식에 의해 지각 필터에 적용되는 추정 잡음을 유사하게 추정함으로써 효과적으로 잔여 잡음을 제거하여, 지각적으로 원 신호에 가까운 개선 신호를 얻을 수 있었다. 특히, 저 대역에 상대적으로 영향을 주는 헬리콥터 잡음을 적용한 결과 SSNR과 NMR 개선량 측면에서는 기존의 방법들보다 입력 잡음에 열화 된 정도가 강할수록 커짐을 확인할 수 있었으며, 평균 개선량은 SSNR인 경우 2.7dB, NMR인 경우 6.1dB의 성능 개선을 확인하였다.

또한 전체적으로 잡음 에너지 차감 필터나 기존의 지각 필터, 적용 지각 필터와 비교할 때, SSNR과 NMR 비교, 그리고 청감 테스트 결과에서 추정 오차가 없는 경

우의 잡음을 적용시킨 경우와 같이 좋은 성능의 음질 개선을 확인 할 수 있었다.

하지만, 제안된 알고리즘으로는 비정적 특성을 지닌 잡음에 대해서는 잡음의 변화 (variance)가 시간에 따라 비정적이므로 열화 시키는 잡음의 양이 시간에 따라 변화하므로, 잡음 추정에 있어서 오차가 생김을 확인할 수 있었다. 이러한 추정 잡음의 오차로 인해 필터 적용 후에도 잔여 잡음이 존재하게 된다.

그러므로 향후 진행될 연구 과제는 정적인 통계적 특성을 갖지 않는 비정적 신호 (Non-stationary signal) 잡음을 추정 할 수 있는 방법에 대한 연구가 진행될 예정이다.

감사의 글

본 연구는 숭실대학교 교내연구비 지원으로 이루어졌습니다.

참고 문헌

1. S. F. Boll, "Suppression of acoustic noise speech using spectral subtraction", IEEE Trans. ASSP 27, 1997.
2. R. J. MacAuly, "Speech enhancement using a soft-decision noise suppression filter", IEEE Trans, 28, 1980.
3. Y. Ephraim and D. Malah, "Speech enhancement using a MMS error short-time spectral amplitude estimator", IEEE Trans, 32, 1984.
4. Y. Ephraim, "A Bayesian estimation approach for Speech enhancement using HMM", IEEE Trans, 40, 1992.
5. D. Tsoukalas, J. Mourjopoulos, and M. Paraskevas, "Speech enhancement using Psychoacoustic Criteria", Journal of Audio Eng. Soc, 45 (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, 45 (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, 7 (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 21 (1(s)), pp. 199-202, July. 2002
9. 엄혜영, 한현수, 홍민철, 차형태, "적응 지각 필터를 이용한 오디오 음질 개선 알고리즘", 한국음향학회 논문지 22 (8), pp. 687-693, Nov. 2003

저자 약력

● 서정국 (Joung-Kook Seo)



2003년 2월: 송실대학교 정보통신전자공학부 (공학사)
2003년 3월~현재: 송실대학교 전자공학과 석사 과정 재학
※ 주관심분야: 오디오 및 음성 신호처리, 통신 신호처리

● 차형태 (Hyung-Tai Cha)



1993년: The University of Pittsburgh (공학박사)
1993~96: 삼성전자 신호처리 연구소 선임연구원
1996~현재: 송실대학교 정보통신전자공학부 부교수
2001~현재: 한국 신호 및 시스템학회 이사
2001~현재: 한국 방송공학회지 편집위원장, 편집위원
※ 주관심분야: Multimedia Systems and
Applications Audio and Video
Signal Processing, Communication
System, ASIC Implementation of
Digital System