

# 지각 필터 응답 보상을 통한 음질 개선

채 병 국, 차 혁 근, 차 형 태  
송실 대학교 전자공학과

## Improvement of Sound Quality using Compensation of Perceptual Filter Response

Byoung-Koog Chae, Hyuk-Geun Cha, Hyung-Tai Cha

Dept. Of Electronics Engineering, Soongsil Univ

E-main : ekstns@mms.ssu.ac.kr, hgcha@mms.ssu.ac.kr , hcha@ssu.ac.kr

### 요약

본 논문에서는 잡음에 오염된 신호의 지각관계를 해석하여 지각 필터 응답 제어를 통한 음성 신호 개선 알고리즘을 제안한다. 음성 신호 개선 기법은 단일 채널환경에서 사람의 청각시스템에서의 주파수 변별력을 나타내는 각각의 임계대역에 대한 전체 에너지를 나타내는 임계대역 에너지의 지각적인 확산의 영향 즉, 마스킹 확산의 영향을 나타내는 자극에너지를 이용하여 신호와 잡음 에너지에 의해 변화하는 잡음에 의한 신호의 마스킹 구간을 검출하여 목음 구간 추출 잡음 필터응답과 추정 잡음 오차를 보상 시킨 필터응답을 통한 지각 필터 응답을 보상하여 신호를 개선 하는 방법이다. 실험 결과 제안한 방법을 통해 SNR에 개선과 음질 개선 효과를 얻을 수 있음을 테스트를 통해 확인 하였다.

### 1. 서론

사람의 음성은 통신 수단으로 사용되어 왔기 때문에 신호처리 분야에서 흥미로운 분야의 하나로 자리 잡고 있다. 최근에는 시스템 환경의 발전과 다양한 고품질의 음성 신호의 디지털 표현 기술의 발전으로 인해 열화 된 신호 개선에 있어 정확한 의미 전달의 관점에서 신호의 명료도 개선만으로는 사용자의 요구 조건을 충족시키기

에는 부족하다.[1] 그 결과 최근 잡음에 오염된 신호의 개선을 위한 알고리즘에 있어 높은 SNR 개선과 신호에 대한 지각적인 측면에서도 무결성을 보장하기 위해 인간의 청각적 지각 특성을 이용하여 잡음에 의한 영향을 최소화하기 위해 마크 영역에서의 신호처리를 통한 신호 개선 기법[2] 등을 적용 하거나 사람의 청각 시스템에서의 신호에 대한 지각에 직접적인 접근 방식으로 청각적 자극에너지를 통한 잡음 에너지의 감소와 음질 개선[3]에 발전을 가져 왔다.

본 논문에서는 지각 필터 응답 보상을 통한 음성 신호 개선 방법을 제안 한다.

### 2. 마스킹 특성 및 자극 에너지에 대한 해석

사람의 청각 시스템에서의 주파수 응답 특성을 반영하는 각각의 임계대역에서의 전체 신호 에너지에 대한 마스킹 특성 및 지각적인 에너지 확산현상을 나타내는 자극 에너지에 대한 해석은 다음과 같다.[3][4][5][6][7] 길이  $N$ 의 윈도우 사이즈로 구분되는 입력 신호  $x(n;i)$ 에 대해 푸리에 변환을 통해 전력 스펙트럼을 계산한다

$$X(\omega; i) = \left[ \frac{1}{N} \sum_{n=0}^{N-1} \omega(n) \cdot x(n; i) e^{-j \frac{2\pi n \omega}{N}} \right]^2$$

where,  $0 \leq n \leq N-1$ ,  $0 \leq \omega \leq \frac{N}{2}$

$\omega(n)$ : hanwindow (1)

$i$ : 윈도우를 통해 구분되어지는 시간 영역의 프레임 인덱스

사람의 청각 시스템에서의 주파수 변별력을 나타내는 마크 인덱스  $z$  에서 신호의 전력 스펙트럼  $X(\omega; i)$  에 대한 임계대역 에너지를 계산한다.

$$X_z(z; i) = \sum_{\omega=\omega_z}^{\omega_{z+1}} X(\omega; i)$$

for,  $0 \leq z \leq Z-1$

where,  $\omega_z, \omega_{z+1}$ : 임의의 임계대역에 대한 저주파 경계와 고주파 경계

$Z$ : 전체 임계대역의 수 (2)

임계대역 에너지를 계산하면 다음과 같다.

$$X_z(z; i) = a_0(z) X_i(z; i) \quad (\text{for, } 0 \leq z \leq Z-1)$$

$a_0(z)$ : 의어부터 중이까지의 감쇠특성 (3)

$X_z(z; i)$  는 기저 막에서의 신호의 에너지의 확산현상을 나타내는 지각적인 에너지 확산함수를 통해 컨벌루션을 취함으로써 기저 막에서의 신호 에너지에 대한 응답을 나타내는 자극에너지를 다음과 같이 표현 할 수 있다.

$$X_z(z; i) = \sum_{v=0}^{z-1} sf_2(v, z-v) X_o(v; i) + \sum_{v=z}^{Z-1} sf_1(v, v-z) X_o(v; i)$$

for,  $0 \leq z \leq Z-1$  (4)

where,  $sf_1(v, v-1)$ : Backward Spreading  
 $sf_2(v, z-v)$ : Forward Spreading

이때  $sf_1(v, v-z)$  과  $sf_2(v, z-v)$  을 다음과 같이 단일함수를 이용하여 표현하면

$$SF(v, z) = \begin{cases} sf_2(v, z-v), & v < z \\ sf_1(v, v-z), & v \geq z \end{cases} \quad (5)$$

임계대역의 신호 에너지에 대한 지각 패턴을 나타내는 자극에너지는 다음과 같이 표현 할 수 있다.

$$X_z(z; i) = \sum_{v=0}^{z-1} [SF(v, z) X_o(v; i)]$$

for,  $0 \leq z \leq Z-1$  (6)

확산함수  $SF(v, z)$  는 다음과 같이 표현 되어질 수 있다.

$$SF(\Delta z) = 15.81 + 7.5(\Delta z + 0.474) - 17.5 \sqrt{1 + (\Delta z + 0.474)^2} \quad (\text{dB SPL}) \quad (7)$$

$\Delta z = z - v$ : 마크인덱스로 표현된 임계대역간 차이

기저 막의 신호 에너지에 대한 지각적인 주파수 응답을 나타내는 청각적 자극 에너지는 다음과 같이 표현할 수 있다.

$$X_z(z; i) = H_z^p(z; i) \cdot Y_z(z; i)$$

where,  $0 \leq H_z^p(z; i) \leq 1$  (8)

임의의 첨가 잡음에 열화 된 신호는 신호와 잡음이 서로 상관관계가 없다고 가정하면 잡음에 열화 된 신호는 다음과 같이 표현할 수 있다.

$$Y(\omega; i) = X(\omega; i) + D(\omega; i) \quad (9)$$

이때  $Y(\omega; i)$  와  $D(\omega; i)$  는 각각 잡음에 열화 된 신호와 첨가잡음의 전력 스펙트럼을 나타낸다.

일반적으로 단일 채널 음질 개선 알고리즘들의 경우 현재 프레임에 첨가된 잡음을 알 수 없으므로, 묵음 구간 동안 추정된 잡음의 전력 스펙트럼  $\hat{D}(\omega; i)$  를 이용한 임의의 에너지 차감 방식의 필터  $H(\omega; i)$  를 통해 다음과 같이 나타낼 수 있다.

$$\hat{X}(\omega; i) = H(\omega; i) \cdot Y(\omega; i) = \left( 1 - \frac{\hat{D}(\omega; i)}{Y(\omega; i)} \right) Y(\omega; i) \quad (10)$$

with,  $0 \leq H(\omega; i) \leq 1$

이때 식 (8)에서와 같은 신호에 대한 청각 시스템에서의 지각적 응답을 나타내는 기저 막에서의 자극 에너지를 통해 식 (9)과 같은 관계식으로 표현할 수 있으며, 이러한 신호의 지각적인 자극 에너지를 이용하여 식 (10)에서와 같이 묵음 구간추정 잡음의 청각적 자극 에너지  $\hat{D}_z(z; i)$  의 차감을 통한 음질 개선 필터는 다음과 같이 표현된다.

$$\hat{H}_z^p(z; i) = \left( 1 - \frac{\hat{D}_z(z; i)}{Y_z(z; i)} \right)$$

for,  $1 \leq z \leq Z-1$  (11)

에너지 차감 형식의 필터,  $H_z^p(z; i)$  에 대한 잡음에 열화 된 신호에서 잡음 에너지의 이득을 조절하는 임의의 이득 조정 함수와 같이 해석을 할 수 있다.

### 3. 잡음 추정 오차와 필터 응답 오차 보상을 통한 음질 개선

잡음 추정 오차의 영향은 신호에 첨가된 잡음과 추정 잡음 사이의 오차에 의한 잡음 이득 제어 오차로 반영된다. 이러한 에너지 차감 방식의 신호 개선 방식은 추정 오차의 정도에 따라 잡음의 이득 제어 정도가 급격하게 감소하여 추정 오차가 신호 개선에 큰 영향을 미치게 된다. 잡음 추정 오차,  $D_i^c(z; i)$ 의 영향은 에너지 차감을 통해 다음과 같이 된 신호,  $X_i(z; i)$ 와 추정 잡음,  $\hat{D}_i(z; i)$ 를 통해 개선된 신호,  $\hat{X}_e(z; i)$ 을 통해 다음과 같이 나타낼 수 있다.

$$X_i(z; i) = \hat{X}_i(z; i) \pm D_i^c(z; i) \quad (12)$$

이때 이러한  $D_i^c(z; i)$ 의 영향은 개선된 신호 내에 존재하는 잔여 잡음과 뮤지컬 톤 잡음으로 반영되며 부가적인 음질 열화의 원인이 된다. 추정 오차를 보정한 잡음의 자극 에너지,  $\hat{D}_i^c(z; i)$ 는 이러한 임계 대역 에너지,  $Y_i(z; i)$ 와  $\hat{D}_i(z; i)$ 를 통해 보정된 잡음의 임계 대역 에너지  $\hat{D}_i^c(z; i)$ 를 통해 확산 함수를 적용하여 계산된다. 이러한 보정된 잡음의 임계 대역 에너지,  $\hat{D}_i^c(z; i)$ 는 다음과 같이 표현 된다.

$$\hat{D}_i^c(z; i) = \begin{cases} \hat{D}_i(z; i) \pm \text{avg} D_i^c & \text{for } \Delta z_i^c \\ \hat{D}_i(z; i) \pm D_i^c(z; i) & \text{for } \Delta \hat{z}_i^c \end{cases}$$

$$\text{avg} D_i^c = \frac{1}{N_i} \sum_{z \in \Delta \hat{z}_i^c} 10 * \log_{10} D_i^c(z; i) \quad (13)$$

where,  $N_i$  = the number of  $z$  in  $\Delta \hat{z}_i^c$

여기서  $\text{avg} D_i^c$ 는 잡음 지각 구간,  $\Delta \hat{z}_i^c$ 에서의 추정 오차 에너지 레벨 평균을 나타낸다. 이때 추정 오차가 보상된 추정 잡음,  $\hat{D}_e^c(z; i)$ 을 통한 필터 응답,  $H_e^c(z; i)$ 는 뮤지컬 톤 잡음이 발생한 임의의 임계 대역,  $z$ 에 대해 인접 임계 대역으로부터의 자극 에너지 확산을 통한 보상을 위해 다음과 같이 구할 수 있다.

$$\hat{X}_i^c(z; i) = H_e^c(z; i) Y_i(z; i) \quad (14)$$

$$\hat{X}_e^c(z; i) = \sum_{v=0}^{z-1} SF(v; z) \hat{X}_i^c(z; i) \quad (15)$$

where,  $0 \leq z \leq Z-1$

$$\frac{X_i^c(z; i)}{Y_i(z; i)} = H_e^c(z; i) \quad (16)$$

잡음 지각 구간,  $\Delta \hat{z}_i^c$ 과 지각하지 못하는 임계 대역  $\Delta z_i^c$ 에서 목음 구간에서 추정된 잡음 필터 응답,  $H_e^p(z; i)$ 와 추정 오차가 보상된 추정 잡음 필터 응답,  $H_e^c(z; i)$ 를 통해 개선된 필터 응답이 보상된다. 보상된 지각 필터 응답,  $H_e^m(z; i)$ 은 다음과 같이 표현이 된다.

$$H_e^m(z; i) = \begin{cases} \{H_e^p(z; i) + H_e^c(z; i)\} / 2 & \text{for } \Delta z_i^c \\ H_e^p(z; i) / 2 & \text{for } \Delta \hat{z}_i^c \end{cases} \quad (17)$$

개선된 신호의 진폭 스펙트럼은 다음과 같이 구해지게 된다.

$$|\hat{x}(w; i)| = |y(w; i)| \sqrt{H_e^m(w; i)} \quad (18)$$

이때  $H_e^m(w; i)$ 는  $H_e^m(z; i)$ 를 주파수 대역으로 확산 시킨 후 임계 대역간 필터 응답의 평탄화에 대해 선형 보간법을 통해 스무딩 처리한 응답이다.

### 4. 실험결과

8kHz, 16bit/sample의 일반 음성 신호를 원 신호로 이용하였고, 잡음에 의해 열화 된 신호는 가우시안 잡음 신호 대 잡음 비 5dB로 원 신호를 열화 시켜 얻은 신호를 사용하였다.

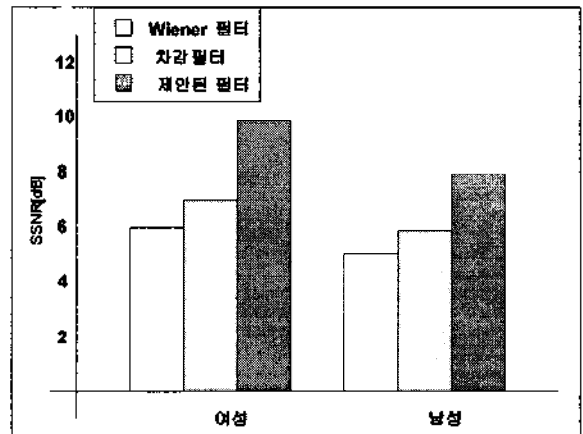


그림1. 신호 대 잡음비(SSNR)

잡음에 의해 열화 된 신호를 필터응답 보상을 통하여 얻은 개선 신호와 지각적·자극 에너지를 이용한 차감 필터를 통하여 얻은 개선 신호를 SSNR(Segmental signal-to-noise ratio)과 MOS(Mean Opinion Score)를 이용하여 객관적인 방법으로 성능을 비교하였다.

그림 1에서 SSNR의 비교를 나타내었다. 여성 신호는 위너 필터 신호보다 4dB, 차감 필터보다 3dB가 개선이 되고 남성 신호는 위너 필터 신호보다 3dB, 차감 필터보다 2dB 개선이 되었다.

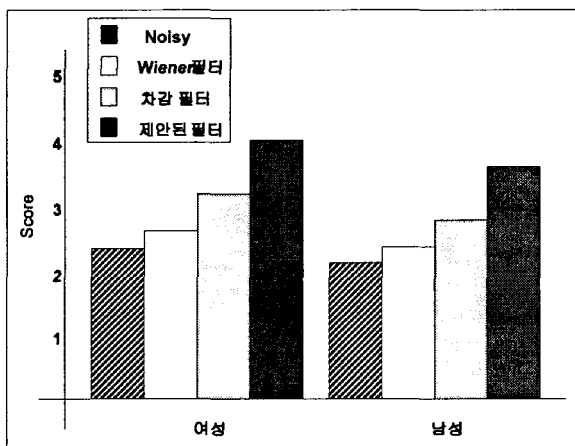


그림 2. 청각 테스트(MOS) 결과

그림 2는 20세 이상 청각적 장애가 없는 일반인 남녀 30명을 대상으로 동일한 오디오 샘플을 이용하여 1점에서 5점까지의 점수로 MOS을 평가한 결과이다. 필터 응답 보상을 통한 개선 방법이 그림 1에서 SSNR결과를 비교한 결과와 같이 성능이 개선되었음을 평가되었다.

## 5. 결과

본 논문에서는 필터 응답 보상을 통한 지각적 특성을 이용하여 개선하는 알고리즘을 제안하였다. 제안한 알고리즘의 성능 비교를 위해 SSNR 비교를 통한 객관적 비교를 수행하였다. 실험 결과 지각적 음질 개선 효과를 평가하는 청각 테스트를 통해 음질 개선 효과를 확인할 수 있었다. 향후 현재까지 검토한 결과를 바탕으로 여러 가지 환경에서 음성 신호와 다양한 잡음에 대한 성능의

변화 및 개선에 관한 연구가 필요 하겠다.

## 참고문헌

1. 한진수, "음성신호처리", 오성 미디어, 2000
2. R. Vetter, "Single Channel Speech Enhancement Using MDL-Based Subspace Approach in Bark Domain," Proc. IEEE, Acoustics, Speech, and Signal Processing, Vol. 1, pp. 641-644, Salt Lake City, UT, USA, May, 2001.
3. D. E. Tsoukalas, J. Mourjopoulos, G. Kokkinakis, "Perceptual Filter for Audio Signal Enhancement", J. Audio Eng. Soc, vol. 45, no. 1/2, Jan. 1997
4. E. Zwicker, H. Fastl, Psychoacoustics : Facts and Models, Springer 2<sup>nd</sup> Edition, 1999.
5. C. J. Moore, Hearing, Academic Press, 1955
6. B. Lincoln, "An Experimental High Fidelity Perceptual Audio Coder Project in Mus420 Win 97",
7. 김현중, "지각적인 간섭현상 해석을 통한 음질 개선에 관한 연구", 숭실 대학교 대학원 전자공학과 논문집, 2003