

가변 임계값을 이용한 지각 필터의 적응적인 음질 개선 알고리즘

Adaptive Enhancement Algorithm of Perceptual Filter Using Variable Threshold

차 형 태*
(Hyung-Tai Cha*)

*송실대학교 전자공학과

(접수일자: 2004년 2월 24일; 수정일자: 2004년 7월 5일; 채택일자: 2004년 8월 2일)

본 논문에서는 잡음에 의해 열화된 오디오 신호를 가변 임계값을 이용한 적응 지각 필터를 사용하여 음질을 개선하는 알고리즘을 제안한다. 제안된 적응 지각 필터는 신호 구간마다 달라지는 신호의 세기와 잡음의 영향 정도를 고려하여 임계값을 가변적으로 조정함으로써 잔여 잡음을 효과적으로 제어하는 방식으로 지각적으로 개선된 음질의 신호를 얻을 수 있다. 제안한 방식은 잡음에 의해 열화된 오디오 신호를 주파수 영역으로 변환한 후 임계 대역 기반의 임계 대역 에너지(Critical intensity energy)와 마스킹 영향이 고려된 청각 자극 에너지(Excitation energy)를 계산한 다음, 지각 필터를 기반으로 한 적응 지각 필터 알고리즘으로 각 단계별 지각 필터 응답을 임계값으로 이용하여 가변 임계값이 재조정되는 단계를 결정하게 된다. 신호의 구간별 에너지 크기에 의한 잡음에 의해 열화된 정도의 차이를 가변 임계값을 이용하여 고려함으로써 잔여 잡음의 효과적인 제어가 가능하게 된다. 제안한 방법은 다양한 신호대 잡음비에서 열화된 오디오 신호를 입력으로 사용하였다. 입력 신호대 잡음비가 15dB, 20dB, 25dB와 30dB의 각각의 경우에 대하여 잡음대 마스킹비(Noise-to-mask ratio, NMR)와 청감 테스트(Mean opinion score, MOS Test)를 시행하였다. 그 결과, 잡음대 마스킹비의 개선 측면에서 각각의 경우에 대해 17.4dB, 15.3dB, 12.8dB, 9.8dB의 개선을 확인할 수 있었고, 청감 테스트의 개선 측면에서는 각각 2.9, 2.5, 2.3, 1.7의 개선된 음질을 확인할 수 있었다.

핵심용어: 지각 필터, 가변 임계값, 오디오 음질 개선, 적응 신호 처리, 임계 대역 처리

투고분야: 음악음향 및 심리음향 분야 (8.7)

In this paper, a new adaptive perceptual filter using variable threshold to enhance audio signals degraded by additively nonstationary noise is proposed. The adaptive perceptual filter updates variable threshold each time according to the power of signal and the effect of noise variation. So the noisy audio signal is enhanced by the method which controls a residual noise effectively. The proposed algorithm uses the perceptual filter which transforms a time domain signal into frequency domain and calculates an intensity energy and an excitation energy in bark domain. In this method, the stage updated the response of filter is decided by threshold. The proposed algorithm using variable threshold effectively controls a residual noise using the energy difference of audio signals degraded by the additive nonstationary noise. The proposed method is tested with the noisy audio signals degraded by nonstationary noise at various signal-to-noise ratios (SNR). We carry out NMR and MOS test when the input SNR is 15dB, 20dB, 25dB and 30dB. An approximate improvement of 17.4dB, 15.3dB, 12.8dB, 9.8dB in NMR and enhancement of 2.9, 2.5, 2.3, 1.7 in MOS test is achieved with the input signals, respectively.

Keywords: Perceptual filter, Variable threshold, Audio enhancement, Adaptive signal processing, Bark domain processing

ASK subject classification: Musical acoustics and psychoacoustics (8.7)

I. 서론

원하지 않는 잡음의 노출에 의해 오디오 신호에는 음질의 열화가 발생하게 된다. 이때, 잡음에 의해 열화된 오디오 신호로부터 원 신호에 가까운 음질로 개선하고자 지금까지 많은 연구가 진행되었으며, 연구 방법에는 청각 모델 기반의 마스킹 특성을 이용한 방법[1,2], 인간의 청각 특성을 모델링하여 시간-주파수 영역의 특성을 이용한 방법[3], 주파수 차감을 이용한 방법[4], 소프트 디시전 필터링을 이용한 방법[5], 최소 평균 오차 (MMSE, Minimum mean square error)를 이용한 방법[6] 등이 있다.

오디오 신호에 대해 분석이 이루어지는 대역폭은 음에 따라 다르게 되는데, 음에 따라 반응하는 대역으로 나누어 놓은 주파수 대역을 임계 대역 (Critical band)이라고 한다. 본 논문에서는 오디오 신호의 주파수 대역에 대하여 인간의 청각 특성인 주파수 변별력을 나타내는 임계 대역으로 잡음에 의해 열화된 오디오 신호를 변환하고, 마스킹 특성이 고려된 청각 자극 에너지를 이용한 지각 필터를 적용적인 방법으로 적용함으로써 잡음에 의해 열화된 오디오 신호의 음질을 개선하고 잔여 잡음을 제어하는 알고리즘을 제안한다. 제안하는 알고리즘은 신호 구간마다 신호의 세기와 잡음의 영향이 변화한다는 것을 고려하기 위해 가변 임계값을 이용한다. 적용 지각 필터의 초기 응답을 이용하여 가변 임계값을 설정하고, 묵음 구간에서 추정된 잡음의 양과 제거되는 누적 잡음의 비를 임계값으로 이용한다.

적용 지각 필터에 의해 개선된 음질을 얻는 과정에서 가변 임계값을 이용하여 적용 지각 필터를 적용한 결과는 신호의 세기와 잡음의 영향이 변화함에 따라 임계값을 가변적으로 조정하고, 필터 응답에 대한 분석을 통하여 응답을 적용시키기 위한 복잡한 계산량을 줄이며 지각적으로 보다 나은 음질 개선을 확인할 수 있었다.

II. 적용 지각 필터

적용 지각 필터는 오디오 신호에 대해서 각각의 주파수 성분에 대한 처리를 하지 않고 임계 대역의 대표 에너지를 이용하여 음질을 개선한다. 푸리에 변환 (Fourier Transform)을 이용하여 주파수 변환된 오디오 신호는

주파수 변별력을 나타내는 임계 대역에서 하나의 대표 에너지로 인지하게 된다. 임의의 i 번째 프레임의 임의의 임계 대역, z 에서의 임계 대역 에너지를 $X(z, i)$ 라고 할 때, 각각의 임계 대역 에너지 (Critical intensity energy)는 인접 대역 및 다른 대역에 마스킹 영향을 끼치게 된다. 마스킹 확산 함수를 이용하여 임계 대역 에너지의 마스킹 현상이 고려된 청각 자극 에너지 (Excitation energy)는 다음과 같이 나타낼 수 있다 [1,2,3,7].

$$X_e(z, i) = \sum_{v=1}^Z [SF(v, z)X(v, i)] \quad (1)$$

where, $1 \leq z \leq Z$ [Bark]

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

이때, z, i 는 바크 인덱스와 프레임 윈도우의 인덱스를, $SF(v, z)$ 는 임계 대역 에너지의 마스킹 현상을 고려한 하기 위해 마스킹 확산 함수를 나타낸다. z 는 신호의 샘플링 주파수에 의해 결정되어지는 전체 임계대역 수를 나타낸다. 44.1kHz의 샘플링 주파수를 갖는 오디오 신호의 경우 z 는 25 Bark를 갖게 된다[8,9].

신호와 잡음이 상관관계가 없다고 가정하면, 잡음에 의해 열화된 신호는 다음과 같이 나타낼 수 있다.

$$Y(z, i) = X(z, i) + D(z, i) \quad (2)$$

where, $1 \leq z \leq Z$

이때, $X(z, i)$ 와 $D(z, i)$ 는 원 신호와 잡음 신호의 임계 대역 에너지를 나타낸 것이다. 식 (2)의 임계 대역 에너지에 의해 표현된 잡음에 의해 열화된 신호는 마스킹 현상을 고려하여 청각 자극 에너지로 나타내면 다음과 같다.

$$Y_e(z, i) = X_e(z, i) + D_e(z, i) \quad (3)$$

where, $1 \leq z \leq Z$

이때, $X_e(z, i)$ 와 $D_e(z, i)$ 는 원 신호와 잡음 신호의 청각 자극 에너지를 나타낸 것이다.

잡음에 의해 열화된 신호로부터 원 신호에 가까운 개선된 신호를 얻기 위한 지각 필터 (Perceptual filter)를 다음과 같이 나타낼 수 있다[3].

$$H_c(z, i) = 1 - \frac{\widehat{D}_c(z, i)}{Y_c(z, i)} \quad (4)$$

where, $1 \leq z \leq Z, 0 < H_c(z, i) \leq 1$

이때, $Y_c(z, i)$ 는 잡음에 의해 열화된 신호의 청각 자극 에너지를 나타내고, $\widehat{D}_c(z, i)$ 는 잡음에 의해 열화된 신호의 묵음 구간으로부터 추정된 잡음 신호의 청각 자극 에너지로, 추정된 잡음이 잡음에 의해 열화된 신호보다 작은 경우만 고려하기로 한다.

지각 필터에 의해 개선된 신호는 묵음 구간의 추정 잡음에 따른 오차가 존재하는 신호 구간에서 잔여 잡음이 존재하게 되어 음질의 열화 원인이 되기때문에 음질 개선 측면에서 잔여 잡음에 대한 제어가 요구된다. 본 논문에서는 잔여 잡음을 제어하기 위하여 지각 필터를 적응적으로 적용함으로써 신호의 왜곡을 최소화하면서 잔여 잡음을 효과적으로 제어하고자 한다. 이때, 잔여 잡음은 처음 묵음 구간에서 추정된 잡음의 양과 동일하지 않으므로 적응적으로 적용되는 각 시점에서 얻어지는 결과 신호에 의해, 식 (4)의 지각 필터를 구성하는 묵음 구간 추정 잡음과 잡음에 의해 열화된 신호의 청각 자극 에너지의 관계를 재조정하게 된다.

즉, 잡음에 의해 열화된 신호는 지각 필터에 의해 잡음 에너지가 감소된 개선 신호를 얻게 된다. 이때, 개선 신호에 남아 음질을 열화시키는 잔여 잡음이 존재할 경우, 지각 필터의 인자인 묵음 구간 추정 잡음과 잡음에 의해 열화된 신호의 청각 자극 에너지를 적응적으로 재조정함으로써 지각 필터의 응답을 재조정하고자 한다.

묵음 구간 추정 잡음은 잡음에 의해 열화된 신호 중에서 원 신호가 아주 미비하고 잡음이 대부분을 차지하는 신호로, 다음과 같이 나타낼 수 있다.

$$\widehat{D}(z, i) \approx Y_c^D(z, i) = X_c^D(z, i) + D_c^D(z, i) \quad (5)$$

where, $X_c^D(z, i) \approx 0$

지각적으로 잡음 에너지의 영향이 없는 개선 신호를 얻는데 적응 지각 필터가 m 번 적용된다고 할 때, 잡음에 의해 열화된 신호와 잡음이 대부분을 차지하는 신호인 묵음 구간 추정 잡음을 적응적으로 재조정하는 필터 적용 과정은 다음과 같이 나타낼 수 있다[1].

$$\widehat{X}_c(z, i)^{(m)} = H_c^A(z, i)^{(m-1)} Y_c(z, i)^{(m-1)} \quad (6)$$

$$= \sqrt{1 - \frac{\widehat{D}_c(z, i)^{(m-1)}}{Y_c(z, i)^{(m-1)}}} Y_c(z, i)^{(m-1)}$$

where, $Y_c(z, i)^{(m-1)} = \widehat{X}_c(z, i)^{(m-1)}$

$m \geq 1, 0 < H_c^A(z, i) \leq 1$

$$\widehat{X}_c^D(z, i)^{(m)} = H_c^A(z, i)^{(m-1)} Y_c^D(z, i)^{(m-1)} \quad (7)$$

$$= \sqrt{1 - \frac{\widehat{D}_c(z, i)^{(m-1)}}{Y_c(z, i)^{(m-1)}}} Y_c^D(z, i)^{(m-1)}$$

where, $\widehat{D}_c(z, i)^{(m-1)} \approx Y_c^D(z, i)^{(m-1)} = \widehat{X}_c^D(z, i)^{(m-1)}$

$m \geq 1, 0 < H_c^A(z, i) \leq 1$

이때, 적응적으로 재조정하는 필터 적용 과정은 식 (6)과 식 (7)에 의해서 얻어진 결과, $\widehat{X}_c(z, i)^{(m)}$, $\widehat{X}_c^D(z, i)^{(m)}$ 이 식 (1)에 의해 재조정된 청각 자극 에너지, $Y_c(z, i)^{(m)}$ 과 $\widehat{D}_c(z, i)^{(m)}$ 을 각각 얻을 수 있고, 이 재조정된 에너지는 식 (8)에서 식 (11)까지의 전개를 통하여 점차 감소하게 됨을 알 수 있다.

$$Y_c(z, i)^{(m-1)} = \widehat{X}_c(z, i)^{(m-1)} \quad (8)$$

식 (6)은 식 (8)의 관계에 의해 식 (9)과 같이 표현할 수 있다.

$$\begin{aligned} \widehat{X}_c(z, i)^{(m)} &= H_c^A(z, i)^{(m-1)} Y_c(z, i)^{(m-1)} \\ &= H_c^A(z, i)^{(m-1)} \widehat{X}_c(z, i)^{(m-1)} \end{aligned} \quad (9)$$

또한 적응 지각 필터 응답은 0과 1사이의 범위에서 존재하기 때문에 식 (9)의 m 번 재조정된 결과와 이전 결과는 식 (10)과 같은 관계를 얻을 수 있다.

$$\widehat{X}_c(z, i)^{(m)} \leq \widehat{X}_c(z, i)^{(m-1)} \quad (10)$$

$$\therefore \widehat{X}_c(z, i)^{(m-1)} - \widehat{X}_c(z, i)^{(m)} \geq 0$$

식 (10)의 임계 대역 에너지에 의한 결과로부터 마스킹 확산 현상을 고려하기 위하여 확산 함수와의 컨벌루션 연산을 통한 청각 자극 에너지를 다음과 같이 계산할 수 있다.

$$\sum_{v=0}^{z-1} [SF(v, z) \{ \widehat{X}_c(z, i)^{(m-1)} - \widehat{X}_c(z, i)^{(m)} \}] \quad (11)$$

$$\begin{aligned}
 &= \sum_{i=0}^{z-1} [SF(v, z) \hat{X}_i(z, i)^{(m-1)}] - \sum_{i=0}^{z-1} [SF(v, z) \hat{X}_i(z, i)^{(m)}] \\
 &= \hat{X}_e(z, i)^{(m-1)} - \hat{X}_e(z, i)^{(m)} \geq 0 \\
 \therefore \hat{X}_e(z, i)^{(m-1)} &\geq \hat{X}_e(z, i)^{(m)} \quad \text{where, } m \geq 1
 \end{aligned}$$

식 (11)로부터 재조정 단계에 의해 얻어지는 청각 자극 에너지는 점차 감소하게 됨을 확인할 수 있다.

잡음에 의해 열화된 신호와 묵음 구간 추정 잡음의 청각 자극 에너지를 재조정하여, 식 (4)의 분자, 분모에 해당하는 성분이 재조정되는 두 인자로 구성된 적응 지각 필터 (Adaptive perceptual filter)[1]는 다음과 같이 일반화시킬 수 있다.

$$H_e^A(z, i)^{(m)} = 1 - \frac{\hat{D}_e(z, i)^{(m)}}{Y_e(z, i)^{(m)}} \quad (12)$$

where, $1 \leq z \leq Z, m \geq 1, 0 < H_e^A(z, i)^{(m)} \leq 1$

이때, $\hat{D}_e(z, i)^{(m)}$ 와 $Y_e(z, i)^{(m)}$ 는 재조정된 묵음 구간 추정 잡음과 잡음 신호의 청각 자극 에너지를 나타내고, m 은 적응 지각 필터가 적용되는 횟수를 나타낸다.

III. 가변 임계값 적용

식 (7)의 추정 잡음을 재조정하는 식의 차감되는 잡음 에너지 즉, 제거된 잡음을 $D_e^R(z, i)$ 라고 하면, m 번 적용된 경우에 대하여 다음과 같이 표현할 수 있다.

$$\begin{aligned}
 D_e^R(z, i)^{(m)} &= \hat{X}_i^D(z, i)^{(m-1)} - \hat{X}_i^D(z, i)^{(m)} \\
 &= \hat{D}_i(z, i)^{(m-1)} - \hat{D}_i(z, i)^{(m)} \quad (13)
 \end{aligned}$$

where, $1 \leq z \leq Z, 0 < H_e^A(z, i) \leq 1$

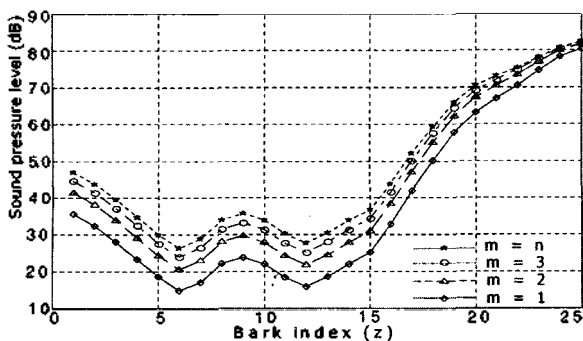


그림 1. 제거된 누적 잡음
Fig 1. Removed cumulative noise.

적용 지각 필터를 적용할 때마다 식 (13)에서 구해지는 제거된 잡음 에너지는 식 (10)과 식 (11)로부터 양수임을 알 수 있고, 식 (1)에 의해서 청각 자극 에너지, $D_e^R(z, i)^{(m)}$ 를 얻을 수 있다.

적용 지각 필터를 잡음에 의해 열화된 신호에 적용하는 단계마다 제거된 누적 잡음의 청각 자극 에너지는 다음과 같다.

$$\sum_{k=1}^m D_e^R(z, i)^{(k)} = D_e^R(z, i)^{(1)} + D_e^R(z, i)^{(2)} + \dots + D_e^R(z, i)^{(m)}$$

where, $m \geq 1, 1 \leq z \leq Z$ (14)

식 (14)에 의해 적응적인 적용 단계에서 얻어지는 누적 잡음의 청각 자극 에너지는 그림 1과 같이 변화함을 확인할 수 있다.

그림 1은 단계마다 제거된 누적 잡음 에너지를 나타낸 것이다. 필터의 적용 단계마다 동일한 양의 잡음이 제거되는 것이 아니라 초기의 잡음량보다 적응 지각 필터의 적용 단계를 거칠수록 잔여 잡음의 영향이 줄어들기 때문에 점차 적은 양의 잡음이 제거됨을 알 수 있다. 또한, 25바크 대역의 제거된 누적 잡음 에너지가 다른 임계 대역에 비해 많은 양의 잡음을 제거함을 확인할 수 있다. 이로부터 고주파에 해당하는 임계 대역에서 대부분의 잔여 잡음을 제거할 수 있다.

묵음 구간에서 추정한 잡음 에너지와 식 (14)의 제거된 누적 잡음 에너지의 비를 r 라고 하면, r 는 다음과 같이 나타낼 수 있다.

$$\begin{aligned}
 r &= \frac{\sum_{k=1}^m D_e^R(z, i)^{(k)}}{\hat{D}_e(z, i)^{(0)}} \\
 &= \frac{D_e^R(z, i)^{(1)} + D_e^R(z, i)^{(2)} + \dots + D_e^R(z, i)^{(m)}}{\hat{D}_e(z, i)^{(0)}}
 \end{aligned}$$

where, $m \geq 1, 1 \leq z \leq Z, 0 \leq r \leq 1$ (15)

이때, r 는 잡음에 의해 열화된 신호에 적응 지각 필터가 적용될 적용 횟수를 결정하기 위한 임계값 역할을 하게 된다. 그림 1에서 확인한 대부분의 잡음이 제거되는 대역을 이용하기 위하여 묵음 구간 추정 잡음과 제거된 누적 잡음의 비를 다음과 같이 다시 표현할 수 있다.

$$r = \frac{\sum_{k=1}^m D_e^k(Z, i)^{(k)}}{\widehat{D}_e(Z, i)^{(0)}} \quad (16)$$

where, $m \geq 1$

이때, Z 는 샘플링 주파수에 의해 결정되는 전체 임계 대역의 수를 나타낸과 동시에 고주파에 해당하는 마지막 임계 대역을 나타낸다. 식 (1)에서와 같이 44.1kHz의 경우 Z 는 25Bark를 나타낸다. 즉, 임계값은 목음 구간 추정 잡음과 제거된 누적 잡음의 청각 자극 에너지의 마지막 임계 대역에서의 비를 나타낸다. 이때, 적응 지각 필터가 임의의 신호 구간에 적용된 횟수를 n 이라고 할 때, 식 (16)에 대해서 n 번 적용된 $r(n)$ 와 $n+1$ 번 적용된 $r(n+1)$ 의 사이에는 다음과 같은 관계가 있음을 알 수 있다.

$$r(n) = \frac{\sum_{k=1}^n D_e^k(Z, i)^{(k)}}{\widehat{D}_e(Z, i)^{(0)}}$$

$$r(n+1) = \frac{\sum_{k=1}^{n+1} D_e^k(Z, i)^{(k)}}{\widehat{D}_e(Z, i)^{(0)}}$$

$$= \frac{\sum_{k=1}^n D_e^k(Z, i)^{(k)} + D_e^{n+1}(Z, i)^{(n+1)}}{\widehat{D}_e(Z, i)^{(0)}}$$

$$= r(n) + \frac{D_e^{n+1}(Z, i)^{(n+1)}}{\widehat{D}_e(Z, i)^{(0)}}$$

$\therefore r(n) < r(n+1)$ (17)

where, $\widehat{D}_e(Z, i)^{(0)} \geq 0, D_e^k(Z, i)^{(n+1)} \geq 0$

식 (17)로부터 적응 지각 필터가 잡음에 의해 열화된 신호에 적용되는 횟수가 증가함에 따라 r 의 값이 증가함을 알 수 있다.

잡음의 영향이 신호의 세기에 의해 변화하는 상황을

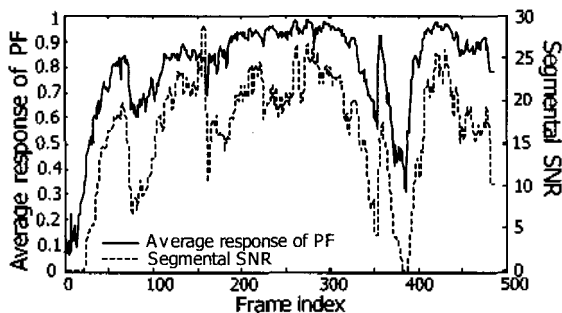


그림 2. 신호대 잡음비와 평균 지각 필터 응답 비교
Fig 2. Signal-to-noise ratio vs Average perceptual filter response.

고려하기 위하여 신호의 프레임 단위의 특징으로 잡음의 영향에 따라 변화하는 식 (4)에서의 지각 필터 응답을 이용할 수가 있다.

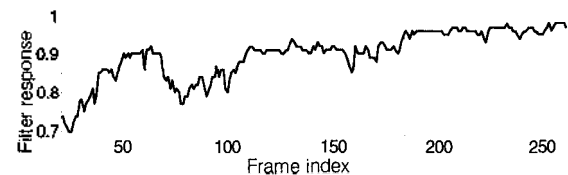
그림 2는 프레임단위의 신호대 잡음비의 변화와 평균 지각 필터 응답을 나타낸 것이다. 평균 지각 필터 응답 (Average PF response)은 좌축 (左軸)에, 신호대 잡음비 (Segmental SNR)는 우축 (右軸)에 단위를 서로 다르게 나타내어 프레임단위에 대한 신호대 잡음비와 평균 지각 필터 응답을 비교할 수 있다. 그림 2로부터 신호대 잡음비와 평균 지각 필터 응답의 변화에서 유사성을 찾을 수 있다. 평균 지각 필터 응답으로부터 신호의 세기를 추정하고, 크게 추정된 구간에서는 임계값을 적게, 적게 추정된 구간에서는 임계값을 크게하여 이용함으로써 잡음의 영향을 고려하여 적응 지각 필터를 적용하였다. 가변적인 임계값을 r 로 나타낼 때 다음과 같은 임계값의 범위를 얻을 수 있다.

$$\begin{cases} r_v = 1 - (H_e(z, i)_{mean} - r_{v_{min}}) & \text{if, } H_e(z, i)_{mean} > r_{v_{min}} \\ r_v = r_{v_{min}} & \text{, otherwise} \end{cases} \quad (18)$$

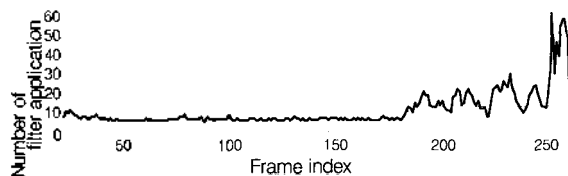
where, $H_e(z, i)_{mean} = \frac{1}{Z} \sum_{i=1}^Z H_e(z, i)$

지각 필터의 평균 응답을 이용한 가변 임계값을 적용함에 있어 지각 필터 응답에 따른 적응 지각 필터 응답의 적용을 위한 계산량을 비교하면 다음과 같다.

그림 3은 지각 필터의 평균 응답과 응답의 적응 횟수 (계산량)를 비교한 것으로 그림 (b)의 지각 필터 응답이 0.95 이상인 경우에서 (a)의 적응 횟수가 급격히 증가하



(a) 지각 필터 평균 응답
(a) Average response of perceptual filter



(b) 복잡도
(b) Complexity

그림 3. 지각 필터의 평균 응답과 복잡도 비교
Fig 3. Average response of perceptual filter vs Complexity.

는 것을 확인할 수 있다. 가변 임계값의 최대값을 0.95라고 할 때 최대 가변 임계값보다 큰 구간에서 적응 횟수가 급격히 증가하게 된다. 이 구간은 잡음의 영향이 미세한 구간으로 신호의 마스킹 영향력으로 지각되는 잡음의 양이 거의 없다. 이러한 구간에 대한 많은 적응 횟수는 신호의 음질 개선에 불필요한 계산량이 되기 때문에 적응 지각 필터 응답의 작용을 조정할 필요가 있다. 식 (18)의 가변 임계값의 범위를 다음과 같이 다시 표현할 수 있다.

$$\begin{cases} r_{vt} = r_{vt\max} & , \text{ if } H_{e\text{mean}} < r_{vt\min} \\ r_{vt} = 1 - H_{e\text{mean}} & , \text{ if } H_{e\text{mean}} > r_{vt\max} \\ r_{vt} = 1 - [H_{e\text{mean}} - r_{vt\min}] & , \text{ otherwise} \end{cases} \quad (19)$$

where, $r_{vt\min} \leq r_{vt} \leq r_{vt\max}$

식 (19)에 의해 지각 필터의 평균 응답이 최대 가변 임계값보다 큰 구간의 적응 횟수를 제한한 후 적응 횟수는 다음과 같이 얻을 수 있다.

그림 4는 지각 필터의 평균 응답이 최대 가변 임계값의 범위를 넘을 때 식 (19)를 이용하여 적응 횟수를 제한

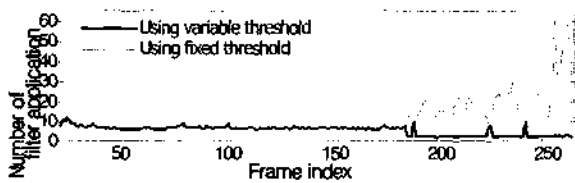


그림 4. 계산량의 개선
Fig 4. Improvement of complexity.

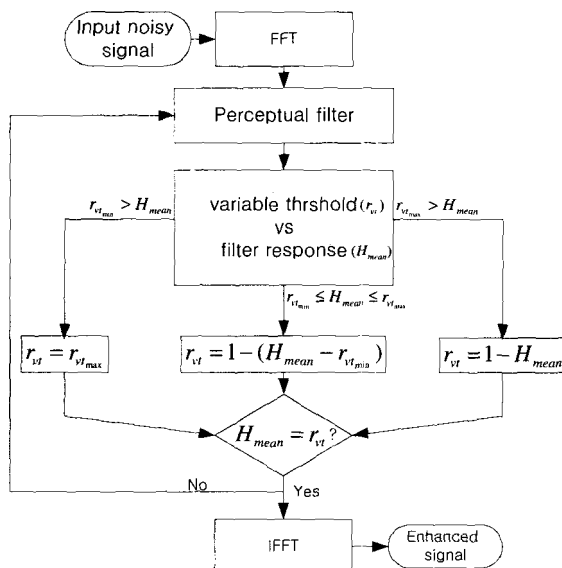


그림 5. 제안된 시스템
Fig 5. Proposed system.

표 1. 프레임단위 신호대 잡음비
Table 1. Segmental Signal-to-noise ratio.

입력 신호대 잡음비(dB)	15.0	20.0	25.0	30.0	
개선 결과 (dB)	APF	21.4	25.4	29.3	33.4
	PF	18.6	23.2	27.7	32.1
	SPS	16.8	21.7	26.5	31.3

함으로써 가변 임계값을 지각 필터의 적용한 결과로 고정 임계값을 이용하여 지각 필터를 적용한 경우보다 계산량을 줄일 수 있음을 보여준다. 특히 지각 필터의 응답이 0.95보다 크게 나온 잡음의 영향이 미세한 구간에서의 불필요한 연산을 줄일 수 있다.

VI. 실험 결과

다음 그림은 제안된 음질 개선 시스템의 전체적인 블록도를 나타낸다.

실험에 이용한 오디오 신호는 Classic, Jazz, Rock, Vocal등의 4가지 장르의 오디오 신호를 백색 가우시안 잡음 (White Gaussian Noise)에 의해 신호대 잡음비 (Signal-to-noise ratio, SNR) 15dB, 20dB, 25dB 그리고 30dB로 열화시킨 신호를 이용하여 실험을 하였다.

다음 그림은 입력 신호로 사용한 잡음에 열화된 오디오 신호를 나타낸 것이다.

그림 6는 본 논문에서 잡음 신호를 제거하기 위해 사용한 오디오 신호를 나타낸 그림으로 각각의 프레임단위로 실험을 적용하였다.

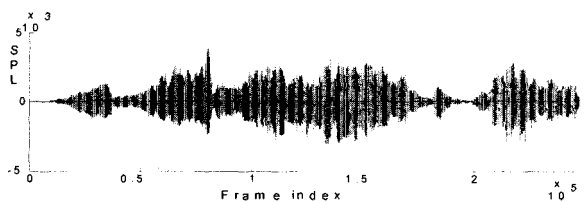


그림 6. 입력 신호
Fig 6. Input signal.

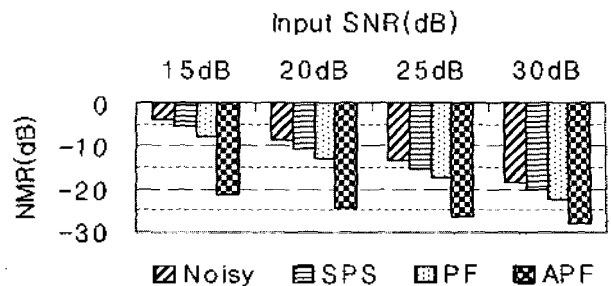


그림 7. 잡음대 마스킹비
Fig 7. Noise-to-mask ratio.

잡음 에너지 차감 방식의 필터 (Simple power subtract, SPS)와 지각 필터 (Perceptual filter, PF), 그리고 본 논문에서 제안하는 방식의 가변 임계값의 범위로 적용한 적응 지각 필터 (Adaptive perceptual filter, APF) 각각의 방식을 15dB, 20dB, 25dB 그리고 30dB의 신호대 잡음비로 열화된 신호 (Noisy)에 적용한 결과 개선된 신호들의 프레임단위 신호대 잡음비 (SSNR, Segmental signal-to-noise ratio)를 비교한 결과는 다음 표와 같다.

표 1에서 확인할 수 있듯이 제안한 방법에 의해 다양한 입력 신호대 잡음비의 신호에 대해 6.4dB, 5.4dB, 4.3dB, 3.4dB의 개선을 얻을 수 있었다. 각각의 경우에 대한 잡음대 마스킹비 (NMR, Noise-to-mask ratio)를 비교한 결과는 다음 그림 7과 같다.

그림 7은 잡음 에너지를 마스킹 임계값에 비교하여 나타내는 것으로 가청 가능한 잡음의 양을 나타내고 있다. 이 결과로부터 신호대 잡음비 30dB의 경우 NMR의 감소에 의해 개선 신호에 존재하는 가청 가능한 잡음의 많은 양이 감소함을 알 수 있다.

앞에서 확인한 신호대 잡음비와 잡음대 마스킹비의 수치적 성능 비교를 통하여 확인한 적응 지각 필터의 성능을 주관적인 음질의 개선을 비교하기 위하여 청감 테스트를 수행하였다. 청감 테스트는 실험에서 이용한 4가지 장르의 신호에 대해서 20세 이상 성인 남녀 30명을 대상으로 SNR 15dB의 잡음에 의해 열화된 신호를 1이라고 하고 원음의 음질을 5라고 할 때 원음에 가까운 음질 정도를 기준으로 1과 5사이의 값으로 음질을 표현하도록 하여 수행하였으며, 주변 잡음과의 차음성확보를 위해 Etymotic Research사의 Ear-Canal형 헤드폰 ER4S와 Beyerdynamic사의 레퍼런스 헤드폰 DT-880 및 HeadRoom사의 Headphone Amp를 이용하였다. 청감 테스트를 수행한 결과는 다음 그림 8과 같다.

입력 신호대 잡음비 15dB, 20dB, 25dB 그리고 30dB

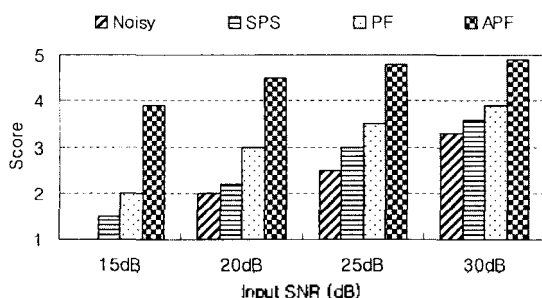


그림 8. 청감 테스트
Fig 8. Mean opinion score Test.

의 신호에 대하여 잡음 에너지 차감 필터와 지각 필터 그리고 본 논문에서 제안한 방식의 적응 지각 필터의 개선 신호에 대하여 청감 테스트를 진행하였다. 청감 테스트 결과, 적응 지각 필터를 적용함으로써 신호대 잡음비 25dB와 30dB의 경우에는 원음과 거의 흡사한 음질을 얻을 수 있었고, 지각 필터의 결과 신호에 비하여 음질의 개선을 확인할 수 있다.

V. 결론

적응 지각 필터를 적용하여 잡음에 의해 열화된 오디오 신호의 음질을 개선하는 알고리즘을 제안하였다. 이때, 적응 지각 필터의 적용 단계마다 제거된 누적 잡음과 묵음 구간 추정 잡음의 비율을 임계값으로 이용하여 신호의 에너지 변화에 따라 가변적으로 적용하였다. 또한, 필터 응답을 분석함으로써 필터 응답의 적용에 대한 그림 4와 같이 계산량을 개선할 수 있었다.

잡음 에너지 차감 필터나 지각 필터와 비교할 때 제안한 방식의 적응 지각 필터의 개선 신호에서 SSNR, NMR 개선을 확인할 수 있었고, 청감 테스트의 경우 신호대 잡음비 25dB와 30dB에서는 원음을 구분할 수 없을 정도의 원음에 가까운 음질을 얻을 수 있었다.

향후 과제로 적응 지각 필터 방법은 필터 응답을 적용시키기 위하여 사용되는 많은 계산량을 줄이기 위한 방안의 연구와 다양한 잡음 환경에 응용할 수 있는 연구가 진행될 예정이다.

감사의 글

본 연구는 숭실대학교 교내연구비 지원으로 이루어졌습니다.

참고 문헌

1. 임해영, 한현수, 홍민철, 차형테, "적응 지각 필터를 이용한 오디오 음질 개선 알고리즘", 한국음향학회지, 22(8), 667-693, Nov. 2003.
2. N. Virag, "Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System", IEEE Trans. Speech and Audio Processing, 7(2), 126-137, Mar. 1999.

3. Tsoukalas, D. E., Mourjopoulos, J. N., and Kokkinakis, G., "Perceptual filters for audio signal enhancement", J of Audio Eng. Soc. 45(1/2), 22-35, Jan/Feb. 1997.
4. P. Lockwood and J. Boudy, "Experiments with a nonlinear spectral subtractor (NSS), hidden Markov models and projection, for robust recognition in cars", Speech Commu., 11, 215-228, Jun. 1992.
5. R. J. McAulay and M. L. Malpass, "Speech enhancement using a soft-decision noise suppression filter", IEEE Trans. Acoustic., Speech, Signal Processing, 28, 137-145, Apr. 1980.
6. Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator", IEEE Trans. Acoustic., Speech, Signal Processing, 32, 1109-1121, Dec. 1984.
7. Andrzej Czyewski, Rafa Krlkowski, "Noise Reduction in Audio Signals Based on the Perceptual Coding Approach", Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 17-20.10 New Paltz, NY, USA, 147-150.
8. J. Johnston "Transform coding of audio signals using perceptual noise criteria", IEEE Journal of Selected Areas in Communications, 6(2), Feb. 1988.
9. E. Zwicker, and H. Fastl, *Psychoacoustics : Facts and Models*, Springer 2nd Edition, 1999.

저자 약력

◆ 차 형 태 (Hyung-Tai Cha)



1993년: The University of Pittsburgh (공학박사)
 1993~96: 삼성전자 신호처리 연구소 선임연구원
 1996~현재: 숭실대학교 정보통신전자공학부 부교수
 2001~현재: 신호및시스템학회 이사
 2001~현재: 한국 방송공학회지 편집위원장, 편집위원

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