

잡음에 열화된 오디오 신호의 분석을 통한 효율적인 적응지각필터 반복 수행 계수의 예측

류일현, 차형태, 구교식, 서보국
숭실대학교 정보통신전자공학부

Prediction of Efficient Adaptive Perceptual Filter Iterate Coefficient through Analysis of Noisy Signal

Il-hyun Ryu, Hyung-tai Cha, Kyo-sik Koo, Bo-kook Seo
School of Electronic Engineering, Soong-sil University

요약

디지털 미디어 기술의 발전은 코딩 분야를 비롯하여 다양하게 발전하고 있다. 특히 오디오 신호 처리 분야에서는 디지털 오디오 신호의 생성, 압축, 복원의 단계가 다양한 형태로 개발되고 있다. 오디오 신호 처리에서 인간의 청각 기관을 모델링한 심리음향 기법은 이용하여 압축뿐만 아니라 잡음 신호의 개선에서도 효과적으로 이용되고 있다. 이러한 심리음향모델을 기반으로 하여 구성된 적응지각필터는 지각필터를 이용하여 적응적으로 잡음에 열화된 신호를 개선한다. 이때, 적응지각필터 반복 수행 계수의 효과적인 결정은 오디오 신호의 청각적 손실을 줄이는 동시에 정확한 잡음 제거를 수행한다. 성능을 확인하기 위해서 SNR 및 NMR 비교를 수행하였다.

I. 서론

오디오 신호는 전송, 압축 과정에서 예기치 않은 잡음에 의해 열화되거나, 청취 환경의 환경 잡음에 의해 음질의 열화를 일으키게 된다. 이러한 여러 요인에 의한 잡음에 열화된 신호의 음질을 원 신호에 가깝게 개선하고자 인간의 청각 특성을 이용하였다. 오디오 신호의 음질 개선을 위해 청각 특성 모델에서의 마스킹특성을 이용한 연구를 비롯하여 청각 특성 모델을 음성 신호에 이용하여 음질 개선한 연구, 네트워크에서의 전송을 위한 코덱에 이용되어 압축효율을 높이기 위한 연구 등이 진행되어져 왔다.

이러한 연구들은 잡음이 존재하는 환경에서 오디오 신호의 효율적인 음질 개선을 위해서 주파수 차감법, 지각필터 등의 단계에 이르게 되었다.[1,2,3,4]

본 논문에서는 청각 특성 분석을 통한 지각필터를 이용하여 적응적인 잡음 제어를 함에 있어서 입력 신호의 분석을 통하여 과도한 반복 연산을 피하고 효율적으로 잡음을 제어할 수 있도록 적응지각필터의 반복 수행 계수를 결정 기법을 제안한다.

II. 본론

1. 적응 지각 필터

청각 특성의 분석을 통하여 임계 대역에서의 주파수 변별력을 고려하여 각 대역별로 존재하는 신호를 에너지로 표현한 임계 대역 에너지와 마스킹 특성을 고려한 청각 자극 에너지를 계산할 수 있다. 잡음에 의해 열화된 신호와 묵음 구간으로부터 추정된 잡음의 청각 자극 에너지에 의해 지각 필터를 표현하게 된다. [1,4,5]

원래의 오디오 신호의 청각 자극 에너지를 $X_e(z)$ 라고 하면 잡음에 의해 열화된 신호로부터 지각적으로 원신호에 가까운 개선된 신호 에너지를 얻을 수 있는 지각필터를 다음과 같이 나타낼 수 있다.

$$H_e(z) = \frac{X_e(z)}{Y_e(z)} \approx 1 - \frac{\widehat{D}_e(z)}{Y_e(z)} \quad (1)$$

where, $1 \leq z \leq Z$

이때, $\widehat{D}_e(z)$ 는 잡음에 의해 열화된 신호로부터 추정된 잡음의 청각 자극 에너지를 나타내고 추정된 잡음이 잡음에 의해 열화된 신호보다 크기 많은 경우에만

고려하기로 한다. [6]

잡음에 의해 열화 된 신호로부터 식 (1)의 지각 필터에 의한 잡음 에너지의 차감 작용을 이용하여 지각 필터를 구성하는 묵음 구간의 추정 잡음, $\widehat{D}_e(z)$ 와 잡음에 의해 열화 된 신호, $Y_e(z)$ 을 재조정하면, 잡음에 의해 열화 된 신호로부터 잔여 잡음을 적응적으로 제어하여 개선된 신호를 얻을 수 있다[1]. 초기의 지각 필터 입력인 잡음에 의해 열화 된 신호에 영향을 주는 잡음의 에너지와 지각 필터에 의해 얻어진 결과 신호에 영향을 주는 잔여 잡음의 에너지는 동일하지 않다.

지각 필터에 의해 얻은 결과 신호에서 잡음에 의해 열화 된 신호보다 잡음 에너지의 영향이 적지만 잔여 잡음이 존재함을 확인할 수 있다.

결과 신호에 영향을 미치는 잔여 잡음을 제어하기 위해 잔여 잡음 에너지를 검출한 시점에서 잡음에 의해 열화 된 신호와 잡음의 청각 자극 에너지의 관계에 재조정을 필요로 한다.[2,3]

잡음에 의해 열화 된 신호를 재조정하는 과정은 다음 두 식을 통해 나타낼 수 있다.

$$\widehat{X}_i^{(m)}(z) = H_e^{(m-1)} Y_i^{(m-1)} = \left(1 - \frac{\widehat{D}_e^{(m-1)}(z)}{Y_e^{(m-1)}(z)}\right) Y_i^{(m-1)} \quad (2)$$

$$\text{where, } Y_i^{(m-1)}(z) = \widehat{X}_i^{(m-1)}(z), 1 \leq z \leq Z$$

이때, m은 재조정 과정을 통해 신호에 적용된 횟수를 나타낸다. 식 (2)에서 얻어진 재조정된 임계 대역 에너지는 마스킹 특성을 고려하기 위해 식 (3)의 확산 함수를 이용하여 재조정된 청각 자극 에너지 식 (4)를 얻는다.

$$Y_e(z) = \sum_{v=1}^Z [F(v, z) Y_i(v)] \quad \text{where, } 1 \leq z \leq Z \quad (3)$$

$$Y_e^{(m)}(z) = \sum_{v=1}^Z [SF(v, z) Y_i^{(m)}(v)] \quad (4)$$

$$\text{where, } Y_i^{(m)}(z) = \widehat{X}_i^{(m)}(z), 1 \leq z \leq Z$$

묵음 구간의 추정 잡음은 신호 성분이 거의 없이 잡음이 대부분을 차지하는 신호로 다음과 같이 표현할 수 있다.

$$\widehat{D}_e(z) \approx Y_e^D(z) = X_e^D(z) + D_e^D(z) \quad (5)$$

$$\text{where, } X_e^D(z) \approx 0, 1 \leq z \leq Z$$

지각 필터의 두 번째 구성 요소인 추정 잡음을 재조정하는 과정은 다음과 같다.

$$\widehat{X}_i^{D(m)}(z) = H_e^{(m-1)} Y_i^{D(m-1)} = \left(1 - \frac{\widehat{D}_e^{(m-1)}(z)}{Y_e^{(m-1)}(z)}\right) Y_i^{D(m-1)} \quad (6)$$

$$\text{where, } Y_i^{D(m-1)}(z) = \widehat{X}_i^{D(m-1)}(z)$$

$$\widehat{D}_e^{(m-1)}(z) = Y_e^{D(m-1)}(z), 1 \leq z \leq Z$$

식 (6)에 의해 재조정된 묵음 구간의 추정 잡음의 임계 대역 에너지는 식 (5)와 같은 과정에 의해 마스킹 특성이 고려된 재조정된 추정 잡음의 청각 자극 에너지, $\widehat{D}_e^{(m)}(z)$ 을 얻을 수 있게 된다.

잡음에 의해 열화 된 신호와 묵음 구간 추정 잡음의 재조정 과정을 거치면서 잔여 잡음의 제어에 의해 지각적으로 원신호에 가까운 개선된 신호를 얻을 수 있게 된다.

이때, 적응적으로 재조정하는 시점에서 얻어지는 제거된 잡음의 누적 에너지와 묵음 구간의 추정 잡음과의 비를 임계값으로 이용하여 효율적인 잔여 잡음 제어를 통한 개선 신호를 얻을 수 있다.

2. 전처리 잡음 추정

일반적으로 지각필터에서 사용되는 추정 잡음은 묵음 구간의 특성을 분석하여 이용한다. 이렇게 추정된 잡음은 묵음 구간에서만 이뤄지기 때문에 잡음에 의해 열화 된 신호에 적용할 때 적지 않은 오차를 포함하게 된다. 지각 필터에서 잡음의 오차는 잡음 제어의 목적을 크게 훼손하는 요인이 된다.

$$\widehat{D}(w, i) = \text{mean} Y(w, i), \quad \text{for } 1 \leq i \leq 3 \quad (7)$$

식 (7)은 기존의 잡음 추정 방법을 나타낸 식이다. 이는 묵음 구간 3 프레임의 평균을 취해 얻어진다.

본 논문에서 사용한 방법은 기존의 고정적인 잡음 추정에서 벗어나 처리되는 프레임 단위로 잡음을 추정하고 원 잡음 크기에 가까운 잡음을 추정함으로써 필터의 성능을 개선한다.

원신호를 $x(n)$ 이라고 하면, 첨가잡음 $d(n)$ 에 의해 열화 된 신호 $y(n)$ 은 다음과 같이 얻을 수 있다.

$$y(n) = x(n) + d(n) \quad \text{for } 0 \leq n \leq N-1 \quad (8)$$

여기서, $y(n)$ 에서 원신호가 대부분을 차지하고 잡음의 크기가 열세한 구간을 NL_{y-low} 라고 하면, 다음과 같은 관계를 얻을 수 있다.

$$NL_{d-low}(w) = [D(w)]_0^{5kHz}$$

$$NL_{y-low}(w) = [Y(w)]_0^{5kHz} \neq NL_{d-low}(w) \quad (9)$$

이와 반대로 잡음이 대부분을 차지하는 대역을 $NL_{y-low}(w)$ 로 표현하면 아래의 관계도 얻을 수 있다.

$$\begin{aligned}
 NL_{d-high}(w) &= [D(w)]_{5kHz}^{f_s/2} \\
 NL_{y-high}(w) &= [Y(w)]_{5kHz}^{f_s/2} \approx NL_{d-high}(w) \quad (10) \\
 \text{where, } f_s &= 44.1kHz
 \end{aligned}$$

이 관계식의 의미는 고대역에서의 잡음에 의해 열화된 진폭 스펙트럼이 실제 잡음 신호의 진폭 스펙트럼과 매우 유사함을 나타낸다.

백색잡음의 전 대역에 걸쳐 일정한 크기의 스펙트럼을 갖는 특성을 이용하여 지각 필터에 적용하는 추정 잡음을 다음과 같이 결정한다.

$$\mathcal{D}_{high} = \begin{cases} \widehat{D}(\omega, i) = Y(\omega) & 15 \leq \omega \leq f_s/2 [kHz] \\ \widehat{D}(\omega, i) = 0, & \text{otherwise} \end{cases} \quad (11)$$

$$\mathcal{D}_{low} = \begin{cases} \widehat{D}(\omega, i) = Y(\omega + 15) & 0 \leq \omega \leq 5 [kHz] \\ \widehat{D}(\omega, i) = Y(\omega + 10) & 5 \leq \omega \leq 10 [kHz] \\ \widehat{D}(\omega, i) = Y(\omega + 5) & 10 \leq \omega \leq 15 [kHz] \\ \widehat{D}(\omega, i) = 0, & \text{otherwise} \end{cases} \quad (12)$$

$$\mathcal{D}_e = \mathcal{D}_{low} + \mathcal{D}_{high} \quad (13)$$

III. 제안한 알고리즘

본 논문에서는 적응지각필터의 적용 횟수를 추정하기 위하여 입력 신호와 추정 잡음의 크기를 이용한다. 입력 신호와 적응지각필터 동작 특성의 상관관계를 적용 횟수에 따른 SNR 과 NMR을 이용하여 측정한다.

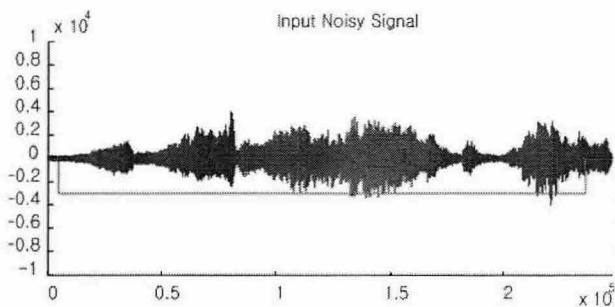


그림 1. 입력 Noisy 신호

그림 1은 적응 지각 필터의 입력 신호이다. 20dB의 White Gaussian Noise를 오디오 신호에 첨가하여 얻은 잡음에 열화된 신호이다. 그림 2,3,4는 각각 적응 지각 필터에 의하여 각각 1회, 3회, 5회의 필터링을 거쳤을 때, 잡음 개선을 나타내는 결과이다.

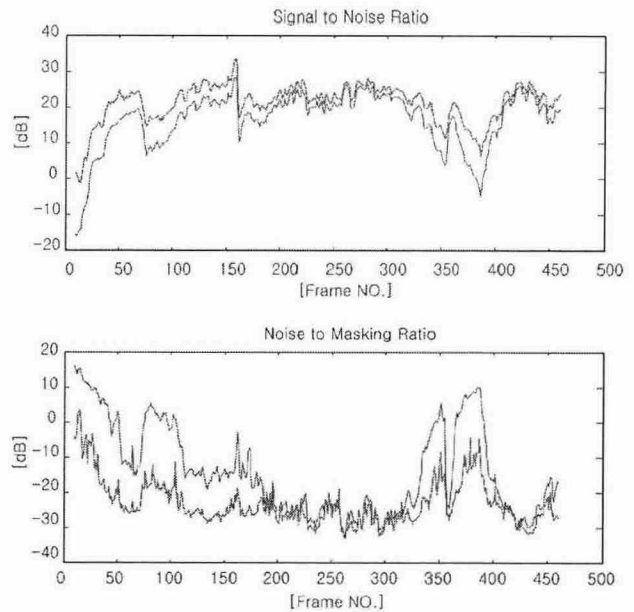


그림 2. 1회 적용 SNR(상)과 NMR(하)

그림 2에서 필터링의 결과로 SNR이 전체적으로 증가한 것을 알 수 있다. 특히 잡음이 대부분을 차지하는 1~20 프레임 구간에서는 결과 SNR이 그림 4에서 가장 잘 개선된 것을 볼 수 있다. 반면에 입력 신호가 큰 250 프레임 전, 후의 구간에서는 필터의 결과가 1회나 5회가 유사함을 알 수 있다.

따라서 이러한 입력 오디오 신호의 에너지의 크기에 따른 잡음 제어 특성을 이용하여, 적응 지각 필터의 적용 횟수를 미리 예측하여, 내부의 비교 연산을 수행하지 않으면서, 적응지각필터를 구현할 수 있다.

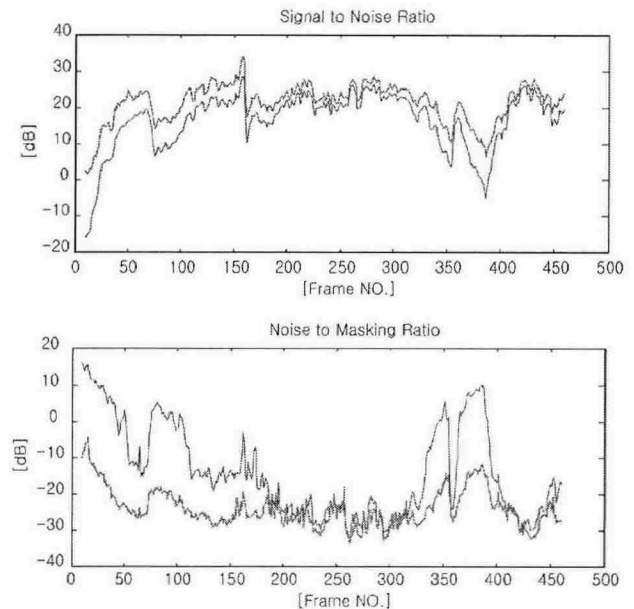


그림 3. 3회 적용 SNR(상) 및 NMR(하)

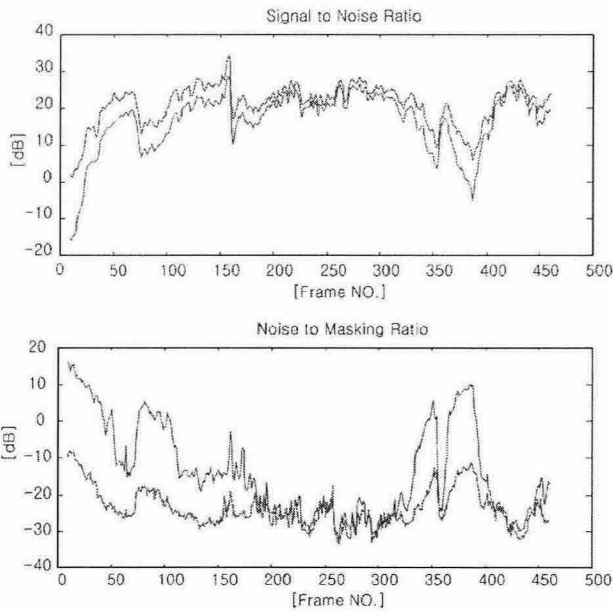


그림 4. 5회 적용 SNR(상) 및 NMR(하)

위에서 나타난 신호의 특성을 다양한 장르의 오디오 신호를 통하여 추출하였다. 아래의 표 1은 잡음 대 신호 비에 따른 반복 횟수 표이다.

평균 잡음 대 신호 비	반복 횟수
10 dB 이상	2
-3 dB 이상	3
-7 dB 이상	4
-7 dB 이하	5

표 1. 적용된 반복 횟수 표

IV. 실험 및 결과

제안하는 알고리즘의 실험을 위해 잡음에 열화된 오디오 신호를 적응지각필터를 통하여 각각 1~5회까지 반복 적용한 결과와 제안하는 알고리즘을 통하여 반복 수행된 결과의 SNR 및 NMR을 비교하였다. 표 2는 20dB 잡음을 첨가하였을 때, 장르에 따라 성능을 평가한 것이다.

적용횟수	Classic	Vocal	Jazz	Rock
1	23.4845	26.7354	24.8537	22.3467
2	23.7430	26.8188	24.9688	22.2582
3	23.7431	26.6193	24.8294	21.8653
4	23.6393	26.3923	24.6096	21.3906
5	23.4854	26.1763	24.3603	20.9054
Proposed	23.7195	26.7676	24.9172	22.0926

표 2. 실험 결과(SNR)

적용횟수	Classic	Vocal	Jazz	Rock
1	-14.3706	-3.8922	-18.4322	-21.5807
2	-18.3651	-8.9836	-20.4384	-21.5027
3	-19.6183	-12.2937	-20.9800	-21.0535
4	-19.9182	-14.2943	-21.0457	-20.5762
5	-19.8825	-15.3348	-20.9513	-20.1332
Proposed	-20.0443	-14.1549	-21.1486	-21.1636

표 3. 실험 결과(NMR)

V. 결론

실험을 통하여 추정된 잡음의 크기와 입력 신호의 크기를 통하여 적응지각필터의 반복 계수를 예측할 수 있음을 볼 수 있었다. 부분적으로 측정하면, 필요 이상의 잡음 제거가 이루어진 부분이 존재하여 NMR이 감소한 부분이 존재한다. 이러한 오차에 대하여는 잡음과 신호의 에너지 비율 이외의 다양한 기법을 통하여 보완해야 할 것이다. 또한 음악 장르의 차이에 의해 제안된 알고리즘에 의해 수행된 성능이 떨어지는 것을 볼 수 있다. 다양한 방법의 접근을 통하여 보완해야 한다. 이러한 적응지각필터의 반복 계수 예측은 단순한 연산량 감소뿐만 아니라 반복 횟수 예측을 통하여 추정 잡음 에너지의 보상 기법으로 사용될 수 있다.

연속된 연구로는 적응지각필터의 반복 횟수 예측을 통한 잡음 에너지 보정 기법을 진행 중에 있다.

참고 문헌

- [1] D. Tsoukalas, . Mourjopoulos, and G. Lollinakis, "Perceptual Filter for Audio Signal Enhancement", Journal of Audio Eng. Soc, vol. 45, no. 1/2, 1997
- [2] 엄혜영, 서정국, 차형태 "잔여 잡음의 적응적 제어를 통한 오디오 음질 개선", 한국음향학회, 추계. 2003
- [3] 유일현, 서정국, 차형태 "잡음 추정을 이용한 적응 지각필터 속도 향상을 위한 전처리 알고리즘", 한국 음향학회, 추계. 2004
- [4] N. Virag, "Single Channel Speech Enhancement Based on Masking Properties of Human Auditory System", IEEE Trans. Speech and audio processing, vol. 7, no. 2, 1999
- [5] E. Zwicker, H. Fastl, Ψ Psychoacoustics : fact and models, Springer, 2nd Edition, 1999
- [6] 서정국 차혁근, 차형태, "잡음에 열화 된 대역에서의 잡음 추정을 이용한 음질개선", 한국 음향학회, 추계, 2003