

정적 통계적 특성 잡음의 추정을 통한 지각 필터 성능 개선

서정국, 류일현, 차형태
승실대학교 전자공학과

Perceptual Filter Performance Improvement through Estimation of Stationary Static Characteristic Noise

Joungkook Seo, Ilhyun Ryu, Hyungtai Cha
Dept. Of Electronics Engineering, Soongsil Univ.

Email : seojk@mms.ssu.sc.kr, fbdlfgus@mms.ssu.ac.kr, hcha@ssu.sc.kr

요약

본 논문에서는 잡음의 변화(variance)가 없는 정적인 통계적 특성(Stationary Static Characteristic)을 갖는 환경에서 잡음 추정을 통해 지각 필터의 성능을 개선하는 알고리즘을 제안한다. 제안된 잡음 추정 알고리즘은 입력되는 잡음에 열화 된 신호의 묵음 구간에서 추정된 잡음을 이용하여 입력되는 잡음의 SNR을 추정 후, 대역 별로 smoothing 상수 값으로 잡음 에너지를 제어하여 첨가된 잡음을 추정함으로써 초기 추정 잡음 보다 가까운 추정 잡음을 얻을 수 있게 된다. 이는 신호를 열화 시킨 잡음을 보다 정확한 추정을 제공함으로써, 지각 필터의 응답을 개선할 수 있고 더불어 잡음에 의해 열화 된 신호의 음질을 개선할 수 있다. 또한 저 대역에 영향을 미치는 잡음인 경우 다른 방법들과는 달리 음질의 개선이 뚜렷하다. 기존의 방식과 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 음질 개선의 개선을 확인할 수 있었다

1. 서론

음질 개선을 위한 각종 음성 처리 시스템이 발전하고 있으며, 그에 따라 고성능 음성 잡음 처리 시스템의 필요는 점점 증가하고 있다. 일반적으로 잡음은 명료성을 떨어뜨리고, 청자에게 불쾌감이나 귀의 피로감을 안겨 준다. 이러한 문제로 잡음을 줄이기 위한 다양한 방법이 제안되어 왔다[1][2][3][4]. 또한 인간의 청각 특성을 모델링하여 시간-주파수 영역의 특성을 이용한 방법 [5][6], 청각 모델 기반의 마스크 특성을 이용한 방법 [7][8][9] 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 뿐만 아니라 잘못된 묵음 구간 검출은 오히려 음성을 더욱 왜곡시키는 결과를 초래한다. 또한 백색 가우시안 잡음(White Gaussian Noise)에서는 성능이 우수하지만, 저 대역에 상대적으로 영향을 미치는 잡음인 경우는 음질 개선이 저하됨을 볼 수 있다.

본 논문에서는 청각 자극 에너지를 이용한 에너지 차감 형식의 필터를 정적인 통계적 특성을 갖는 잡음 환경에서 잡음에 열화 된 오디오 신호를 잡음 에너지 제어 측면에서 보상해 줌으로써 신호의 음질을 개선하는

방법을 제안하고 있다. 또한 추정 오차로 인한 지각 필터의 성능 저하를 줄이기 위해 임계 대역 에너지 영역에서 잡음을 처리하는 알고리즘을 적용함으로써 지각 필터의 성능을 크게 개선하였다.

2. 제안된 잡음 추정 알고리즘

2.1 잡음 추정 알고리즘

일반적으로 기존 지각 필터에 적용되는 잡음은 식 1과 같이 잡음에 의해 열화 된 신호의 세 프레임의 묵음 구간 평균을 이용하여 추정 잡음으로 사용하였다.

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

여기서 구한 추정 잡음, $\hat{D}(\omega, i)$ 를 자극 에너지 $\hat{D}_e(z, i)$ 로 변환시켜 지각 필터에 적용하지만, 1이는 필터에 적용되는 구간마다 동일한 추정 잡음 에너지를 사용하게 됨으로써, 지각 필터의 성능 저하를 가져오게 되는 원인이 된다.

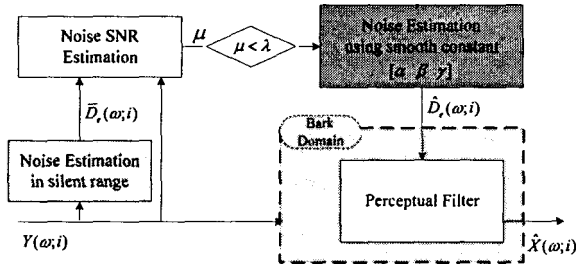


그림 2. 제안한 잡음 추정 알고리즘

그림 2는 제안한 잡음 추정 알고리즘이다. 제안된 방법은 고정적인 추정 잡음을 사용하지 않고, 입력되는 열화 된 신호의 전력과 잡음의 variance를 이용하여 첨가된 잡음의 신호 대 잡음 비(signal-to-noise, SNR)를 추정한다. 그 후, 추정된 SNR에 따른 임계치(threshold)를 구하여 대역 별로 smoothing 상수를 적용함으로써, 일차적으로 추정된 잡음 에너지를 제어하게 됨으로써 적용된 추정 잡음을 생성하게 된다.

원 신호를 $x(n)$ 이라고 하면, 첨가 잡음, $d(n)$ 에 의해 열화 된 신호는 다음 식 2와 같이 얻을 수 있다.

$$y(n) = x(n) + d(n) \quad \text{for}, 0 \leq n \leq N-1 \quad (2)$$

여기서, 잡음에 열화 된 신호, $Y(\omega, i)$ 에서 묵음 구간에서 추정된 잡음, $\bar{D}(\omega, i)$ 을 이용한다. 이때, 첨가되는 잡음에 대한 정보를 알 수 없으므로 첨가되는 잡음의 SNR이 작을수록 신호를 열화 시키는 정도가 커짐으로써 지각 필터를 통해 처리되더라도 잔여 잡음이 많아서 개선의 효과를 볼 수 없다. 이러한 환경에서 열화 된 신호, $Y(\omega, i)$ 와 추정된 잡음, $\bar{D}(\omega, i)$ 을 이용하여 잡음의 SNR을 추정하여 첨가되는 잡음 에너지의 변화량을 미리 예측함으로써, 잔여 잡음을 제거하여 신호를 개선 시키는데 필요한 정보로 사용된다. 첨가되는 잡음의 SNR을 추정하는 방법은 한 프레임의 샘플수가 N이라고 할 때, 식 3과 같다.

$$\begin{aligned} NS_{power} &= \frac{1}{N} \cdot \sum_{i=1}^N (Y)^2 \\ NE_{variance} &= \frac{1}{N} \cdot \sum_{i=1}^N \left(\frac{\bar{D} \cdot \text{std}(\bar{D})}{\bar{D}(1:N)} \right)^2 \quad (3) \\ ES_{SNR} &= 10 * \log_{10} \left(\frac{NS_{power}}{NE_{variance}} \right) \end{aligned}$$

여기서, ES_{SNR} 은 그림 2에서의 μ 로 표현된다. μ 값에 따라 첨가된 잡음의 SNR을 20dB, 15dB, 10dB, 5dB로 추정하게 된다. 이러한 과정은 잡음의 SNR에 따라 필터 응답에 많은 영향을 주는 원인이 되기 때문에 잡음 환경에 맞추어 적당한 smoothing 상수 값을 부여해 주기 위함이다. 위의 μ 값을 이용하여 대역 별로 smoothing 상수 값인 α, β, γ 를 부여하여 새롭게 추정된 잡음, $\hat{D}(\omega, i)$ 를 생성 시킨다.

λ 는 추정된 잡음의 SNR의 임계치이며, LF와 MF는 대역을 나누어 주는 임계치 값으로 LF=2500Hz (15Bark), MF=7000Hz (20Bark)로 사용하였다. 다음 표는 제안한 알고리즘에 사용된 임계치 값들이다.

λ	α	β	γ
10	2.0	1.0	0.5
15	1.8	1.0	0.3
20	1.5	0.5	0.2
25	1.5	0.3	0.1

표 1. 임계치(threshold) 값

2.2 지각 필터 적용

지각 필터에 적용하기 이전의 추정 잡음, $\hat{D}(\omega, i)$ 과 잡음에 의해 열화 된 신호, $Y(\omega, i)$ 의 자극 에너지를 구하면, $\hat{D}(z, i)$, $Y_e(z, i)$ 과 같이 구할 수 있다. 이런 자극 에너지를 지각 필터에 적용함으로써, 그림 3과 같은 개선된 신호의 필터 응답 및 자극 에너지를 구할 수 있다.

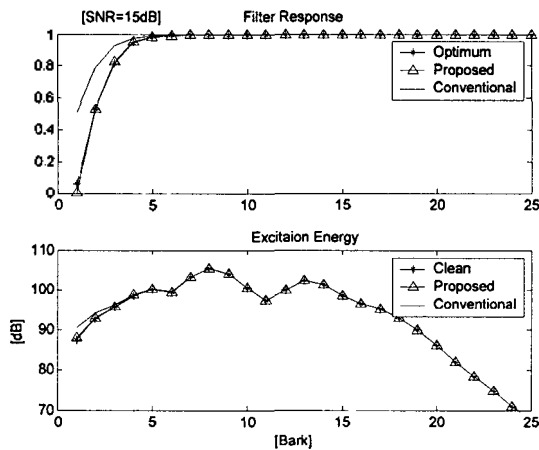


그림 3. 개선된 신호의 필터 응답(상) 및 자극 에너지(하)

그림 3의 필터 응답을 보면, 기존의 지각 필터(—)는 첨가된 잡음을 그대로 사용한 최적화 필터(→)에 비해 저 대역의 잔여 잡음이 적게 제거되는 모습을 볼 수 있다. 이로 인해 기존의 자극 에너지 결과에서도 잡음이 저 대역에서 개선되지 않았음을 볼 수 있다.

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

잡음 SNR이 작아 질수록 원 신호와의 차이가 나는 이유는 잡음의 양이 많아짐으로써, 저 대역에 영향을 주는 잔여 잡음의 양이 많아 필터 적용 시 저 대역에서 필터링의 양이 증가되었기 때문이다. 이로 인해 음질에는 약간의 왜곡이 생기지만, 잔여 잡음의 제거로 인해 명료성의 저하나 귀의 피로감을 줄일 수 있다.

3. 실험 결과

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

신호를 열화 시키는 잡음으로는 정적인 통계적 특성을 갖는 헬리콥터 잡음을 첨가시켜 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 비교하였다. 첫 번째 방법은 주파수 차감 방식인 잡음 에너지 차감 방식의 필터(SPS, Simple power subtract)]를 사용하였으며 두 번째로는 기존의 묵음 구간에서 잡음을 추정 한 후 지각 필터(PF, Conventional method)를 적용하였다. 세 번째 방법으로는 제안된 방법으로 지각 필터에 적용(PPF, Proposed method)하였다.

Input SNR	SSNR			NMR		
	SPS	PF	PPF	SPS	PF	PPF
20dB	20.13	20.49	21.84	-15.53	-14.91	-18.55
15dB	16.90	16.63	18.74	-11.10	-11.20	-16.55
10dB	12.54	12.89	15.73	-6.32	-7.01	-14.19
5dB	8.87	8.99	12.42	-1.42	-2.39	-12.64

표 1. SSNR과 NMR 측정 결과

표 1에서 볼 수 있듯이 제안된 방식으로 개선된 신호는 기존의 PF(-3dB)인 경우 보다 대략 3dB이상이며, 잡음의 SNR이 작아질수록 SSNR의 값이 더 커짐에 따라 성능이 개선되었음을 확인할 수 있다. 그리고, NMR은 잡음 에너지를 마스킹 임계 값에 비교하여 나타내고 것으로 가청 가능한 잡음의 양을 나타내고 있다. 이 결과로부터 SNR이 20dB인 경우 NMR의 감소에 의해 개선 신호에 존재하는 가청 가능한 잡음의 많은 양이 감소함을 알 수 있다.

위의 객관적인 평가인 SSNR과 NMR에서의 개선

효과를 지각적으로 느낄 때의 개선 효과를 보기 위해 청각 테스트 (Mean Opinion Score, MOS) Test를 실시하여 비교한 그림은 다음과 같다.

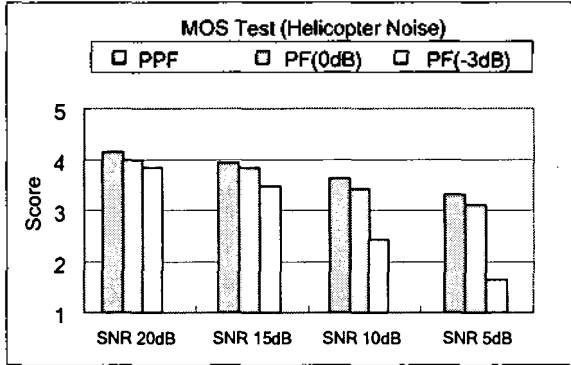


그림 4. MOS 테스트 결과

그림 4의 테스트는 20세 이상 오디오 전문분야에 종사하지 않는 남녀 30명을 대상으로 SNR 15dB의 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 차 음성 확보를 위해 Etymotic Research사의 Ear-Canal형 헤드폰 ER4S와 Beyerdynamic사의 레퍼런스 헤드폰 DT-880 및 HeadRoom사의 Headphone Amp를 이용하였다.

결과에서 볼 수 있듯이 묵음 구간에서 잡음을 추정하여 사용한 PF(-3dB)의 경우 성능이 저하 되는 반면, 제안된 방법으로 잡음을 추정 후 지각 필터에 적용시킨 경우, PPF는 PF보다 우수한 성능을 보이고 있음을 알 수 있다.

4. 결론

통계적 특성을 잡음을 적용하여 잡음에 의해 열화 된 오디오 신호의 청각적 자극 에너지를 이용하여 지각 필터에 적용함으로써 잡음 에너지를 줄이고, 음질 개선을 위해 잡음을 기존의 방법보다 정확하게 추정함으로써 지각 필터를 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. 또한 SSNR과 NMR비교, 그리고 청감 테스트 결과에서 추정 오차가 없는 경우의 잡음을 적용시킨 경우와 같이 좋은 성능의 음질 개선을 확인할 수 있었다.

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

참고문헌

1. S.F. Boll, "Suppression of acoustic noise speech using spectral subtraction", IEEE Trans, vol ASSP-27, 1979.
2. R.J. MacAuly "Speech enhancement using a soft-decision noise suppression filter", IEEE Trans, vol 28, 1980.
3. Y. Ephraim and D. Malah, "Speech enhancement using a MMS error short-time spectral amplitude estimator" IEEE Trans, vol 32, 1984.
4. Y. Ephraim, "A Bayesian Estimation Approach for Speech Enhancement using HMM", IEEE Trans, vol 40, 1992.
5. D. Tsoukalas, M. Paraskevas and G. Kokkinakis, "Perceptual Filter for Audio Signal Enhancement," Journal of Audio Eng. Soc, vol. 45, no. 1/2, pp. 23-35, Jan/Feb. 1997.
6. D. Tsoukalas, M. Paraskevas and J. Mourjopoulos, "Speech enhancement using Psychoacoustic Criteria," proc. IEEE ICASSP, 1994
7. N. Virag, "Single Channel Speech Enhancement based on Masking property 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. 차형태, "가변 임계값을 이용한 지각 필터의 적응적인 음질 개선 알고리즘", 한국음향학회 논문지 Vol 23, No6, pp 446-453, Aug. 2004