

웨이브렛 계수의 표준편차를 이용한 음성신호의 적응 잡음 제거

Adaptive Noise Reduction using Standard Deviation of Wavelet Coefficients in Speech Signal

황항자*† · 정광일** · 이상태*** · 김종교****

Hyang-Ja Hwang*† · Kwang-Il Jeong** · Sang-Tae Lee*** · Chong-Kyo Kim****

한국원자력안전기술원*

Korea Institute of Nuclear Safety

한국원자력연구소**

Korea Atomic Energy Research Institute

한국표준과학연구원***

Korea Research Institute of Standards and Science

전북대학교 전자공학과****

Department of Electronic Engineering, Chonbuk National University

Abstract : This paper proposed a new time adapted threshold using the standard deviations of Wavelet coefficients after Wavelet transform by frame scale. The time adapted threshold is set up using the sum of standard deviations of Wavelet coefficient in cA3 and weighted cD1. cA3 coefficients represent the voiced sound with low frequency and cD1 coefficients represent the unvoiced sound with high frequency.

From simulation results, it is demonstrated that the proposed algorithm improves SNR and MSE performance more than Wavelet transform and Wavelet packet transform does. Moreover, the reconstructed signals by the proposed algorithm resemble the original signal in terms of plosive sound, fricative sound and affricate sound but Wavelet transform and Wavelet packet transform reduce those sounds seriously.

Key words : Wavelet, Wavelet Packet, Noise Reduction, Speech Signal

요약 : 일상생활의 대화중에 포함되는 잡음, 특히 모든 주파수 대역에 포함되는 백색잡음에 의해 오염된 음성신호는 청각적으로 심한 불쾌감과 거부감을 주며 대화의 명료성을 저해시키는 요인으로 작용할 수 있다. 본 논문은 이러한 잡음환경 하에서 음성인식을 위하여 음성에 부가된 잡음을 제거하는 방법으로 프레임 단위로 웨이브렛

† 교신저자 : 황항자(한국원자력안전기술원)

E-mail : hhjbluesky@naver.com

TEL : 016-545-1964, 042-868-0656,

FAX : 042-868-0415

변환을 하여 웨이브렛 계수의 표준편차를 이용하여 시간 적응 임계값을 정하는 새로운 방법을 제안한다. 음성의 특성을 고려하기 위하여 고주파 성분을 많이 가지는 무성음의 경우는 cD1 신호에서, 저주파 성분을 많이 가지는 유성음의 경우는 cA3 신호의 표준편차를 이용하여 시간 적응 임계값을 설정하였다. 실험을 통해 제안한 방법이 일반적인 웨이브렛 변환과 웨이브렛 패킷 변환을 이용한 방법보다 SNR과 MSE 측면에서 향상됨을 확인할 수 있었다. 또한 웨이브렛 변환과 웨이브렛 패킷 변환에서는 과열음, 마찰음 및 과찰음 성분이 많이 제거되는 반면 제안한 방법은 본래 신호와 유사하게 복원됨을 실험 결과 확인할 수 있었다.

주제어 : 웨이브렛, 웨이브렛 패킷, 잡음 제거, 음성신호

1. 서론

음성인식 시스템의 실용화가 늘어남에 따라 최근에는 주변 잡음에 대한 인식 시스템의 성능저하가 문제제시되고 있다. 이러한 이유는 잡음이 없거나 비교적 조용한 실험실에서는 우수한 성능을 나타내는 음성인식 시스템의 성능이 잡음환경의 영향을 고려하지 않은 음성인식 시스템의 실제 환경에서는 성능이 급격하게 감소하기 때문이다. 따라서 잡음에 의해 오염된 음성신호, 특히 백색잡음에 의해 오염된 음성신호는 청각적으로 심한 거부감을 주게 되므로, 이러한 오염된 음성신호에서 잡음을 제거하는 기술은 매우 중요하다고 볼 수 있다.

최근 들어 활발히 연구되고 있는 웨이브렛 변환(wavelet transform)은 해석하고자 하는 주파수 성분에 따라 가변할 수 있는 모함수(mother function)를 제공한다. 가변하는 모함수에 의한 웨이브렛 변환의 다중해상도 특성과 시간-주파수 국부성은 통계적 특성을 모르거나 시간적으로 예측하기 힘든 비정상적(non-stationary)인 신호해석에 매우 유용한 것으로 밝혀졌다[8, 9].

본 논문에서는 웨이브렛 변환을 이용한 잡음 제거 방법에서 가장 중요한 부분인 임계값을 적응적으로 구하는 새로운 방법을 제안한다. 웨이브렛 변환을 수행한 후 유성음 구간에 해당하는 저주파 성분과 무성음 구간에 해당하는 고주파 성분을 이용하여 시간 적응 임계값을 제안하였다. 성능 평가를 위해서 SNR(Signal to Noise Ratio)과 MSE(Mean

Squared Error)을 계산하여 기존의 웨이브렛 변환 및 웨이브렛 패킷 변환과 비교하였다.

2. 웨이브렛을 이용한 잡음제거 방법

웨이브렛을 이용한 thresholding 방법의 기본 원리는 백색 가우시안 잡음(white gaussian noise)에 의해 오염된 신호를 웨이브렛 변환했을 때 각 스케일에 포함된 잡음 성분은 신호 성분의 크기보다 상대적으로 작은 값을 가지므로 적절한 임계값 이하의 값을 제거한 후 다시 합성함으로써 효과적으로 잡음을 제거할 수 있음을 이용한다. 즉, 기저 웨이브렛을 선택한 후 잡음에 손상된 음성으로부터 웨이브렛 변환을 취하여 계수를 구한 후 각각의 레벨별로 임계값 λ 를 정한다. soft thresholding을 하여 웨이브렛 계수의 임계값 이하의 성분을 제거한 후 역 웨이브렛 변환을 취함으로써 잡음이 제거된 음성을 구할 수 있다[3, 6, 7].

임계값 λ 는 웨이브렛 변환일 경우 식 (2)를 사용하며, 웨이브렛 패킷 변환일 경우 식 (3)을 사용한다[6].

$$T_{soft}(X) = \begin{cases} \text{sgn}(X)(|X| - \lambda) & , |X| \geq \lambda \\ 0 & , |X| < \lambda \end{cases} \quad (1)$$

$$\lambda = \sigma \sqrt{2 \log(N)} \quad (2)$$

$$\lambda = \sigma \sqrt{2 \log(N \log_2 N)} \quad (3)$$

여기에서 $T_{soft}(X)$ 는 잡음이 제거된 신호이며, X 는 잡음이 첨가된 음성신호, N 은 신호의 샘플

수이고, σ 는 선택되어진 웨이브렛 계수의 표준편차이다. 표준편차는 잡음이 포함된 웨이브렛 계수의 중간값을 이용하여 식 (4)와 같이 계산한다.

$$\sigma = MAD/0.6745 \quad (4)$$

여기에서 MAD는 median absolute deviation을 의미한다[5, 7].

3. 시간 적응 임계값

웨이브렛 변환으로 얻어지는 웨이브렛 계수가 음성 신호 구간과 잡음 신호 구간에서 다른 특성을 가진다는 것을 이용하였다[2, 3]. 웨이브렛 계수의 표준편차는 잡음 신호 구간에서 작은 값을 가지며 음성이 포함된 구간에서는 비교적 큰 값을 가진다. 웨이브렛 변환에서 스케일이 작을수록 빠르게 변화하는 부분에서 정확하게 검출할 수 있어서 고주파 구간에서 웨이브렛 계수의 표준편차는 작아지게 된다. 잡음이 섞인 음성신호의 초기부분과 끝부분은 잡음 신호만이 존재한다고 가정한다[1~4].

잡음이 첨가된 음성신호를 프레임 단위로 나누어 각각의 프레임에 대하여 웨이브렛 변환한 후 cD1과 cA3를 이용하면 잡음 환경 하에서도 음성 구간을 검출할 수 있음을 이용하여 시간 적응 임계값을 생성한다. 이는 잡음이 섞인 음성신호에 존재하는 파열음, 마찰음 및 파찰음의 경우 신호의 에너지는 유성음구간에 비해 상대적으로 작지만 주파수 영역에서 고주파 부분에 많은 에너지를 가지게 되며, 유성음 구간의 경우 저주파 부분에 많은 에너지를 가지게 됨을 이용한 것이다. cD1의 프레임 단위 표준편차는 C_{D1}^k 로 표시하고 cA3의 표준편차는 C_{A3}^k 로 표시하며, k는 프레임의 수를 나타낸다.

본 논문은 “소프트웨어” 음성파형을 보여주는 그림 1에서 C_{A3}^k 의 출력 파형이 음성 파형과 비슷함을 이용하여 시간 적응 임계값으로 사용할 수 있음을 알게 되었다.

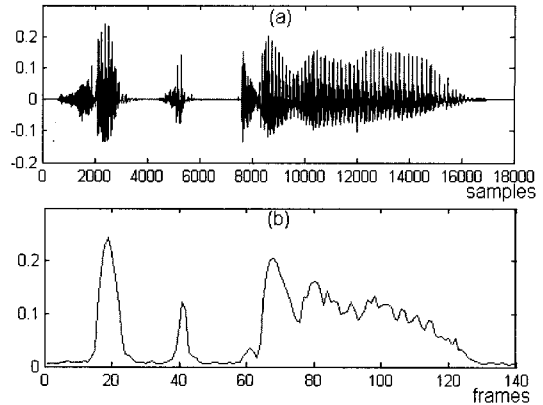


그림 1. (a) 음성 파형 (b) C_{A3}^k 의 파형

그러나 음성파형에 존재하는 고주파 성분의 과찰음을 저주파 성분을 표시하는 C_{A3}^k 에서는 제대로 나타내지 못한다. 따라서 고주파 성분인 과찰음의 정보를 잘 나타내는 C_{D1}^k 를 C_{A3}^k 에 적절하게 더함으로써 음성파형 전부를 나타낼 수 있다.

이러한 성질을 이용하여 웨이브렛 영역에서 음성 검출을 위한 파라미터를 식 (5)와 같이 정의하여 사용할 수 있다[2].

$$T^k = C_{A3}^k + \alpha C_{D1}^k \quad k=1, \dots, N. \quad (5)$$

k는 프레임의 수이며, 가중치 α 값은 [3]에서는 ‘6’의 값을 사용하여 음성 구간을 검출하였다.

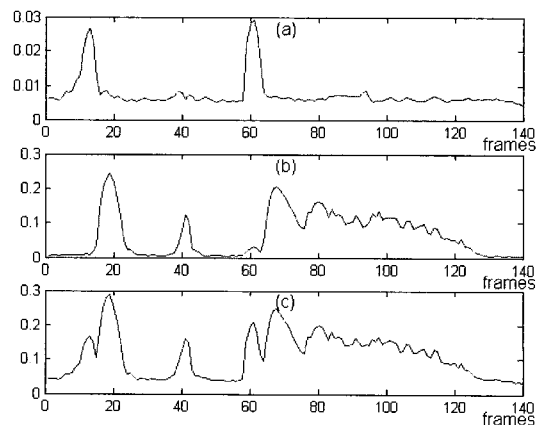


그림 2. (a) C_{D1}^k (b) C_{A3}^k (c) T^k

α 값을 ‘6’으로 설정한 T^k 의 그래프를 살펴보면 그림 2와 같다. C_{D1}^k 의 값에 ‘6’을 곱하여 C_{A3}^k 에 더함으로써 T^k 의 기저선이 원래 C_{A3}^k 의 기저선 0.013 정도에서 0.06 정도로 상승했음을 볼 수 있다. 즉, C_{D1}^k 의 기저선의 6배가 더해져서 T^k 의 기저선이 상승하였다. 이러한 기저선 상승의 결과는 [2]의 논문에서와 같이 끝점 검출만을 목적으로 사용할 수는 있다. 그러나 본 논문에서는 T^k 값을 시간 적응 임계값으로 활용할 수 있도록 다음과 같은 과정을 거친다.

- 1) 식 (5)에서 α 의 값을 C_{D1}^k 과 C_{A3}^k 의 최대값의 비로 구한다.

$$\alpha = \frac{\max(C_{A3}^k)}{\max(C_{D1}^k)} \quad (6)$$

- 2) 위에서 구한 α 값을 식 5에 적용하여 T^k 값을 구한 후 기저선 제거를 위해서 T^k 의 평균값을 구한 후, C_{D1}^k 에서 이 평균값만큼 빼준다 (그림 3(a)).

$$T^{k'} = T^k - \beta \times \text{mean}(T^k) \quad (7)$$

여기에서 β 는 조정 파라미터이다.

- 3) 파라미터 값의 최대값이 ‘1’이 되도록 정규화한다(그림 3(b)).

$$T^{k''} = T^{k'} \times \frac{1}{\max(T^{k'})} \quad (8)$$

- 4) 재샘플링을 통하여 시간 적응 임계값을 구한다(그림 3(c)). 즉, $T^{k''}$ 는 원 신호의 샘플수가 아닌 프레임의 수이므로 원 신호의 샘플수로 재샘플링을 통하여 시간 적응 임계값을 결정할 수 있다.

- 5) 파라미터 $T^{k''}$ 를 이용하여 시간 적응 임계값을 식 2에서 구한 λ 값을 곱하여 사용한다(그림 3(d)).

$$\lambda^{k'} = \delta \lambda (1 - T^{k''}) \quad (9)$$

여기에서 δ 는 조정 파라미터이다.

- 6) 마지막으로 잡음음성신호와 시간 적응 임계값을 비교하면서 soft thresholding을 이용하여 잡음을 제거한다.

4. 실험방법 및 결과

4.1 실험방법

실험에 사용된 음성은 16kHz 샘플링되고 16bit로 양자화된 “소프트웨어”와 “1월 15일까지 할 수 있는데요”를 발음한 음성신호를 사용하였다.

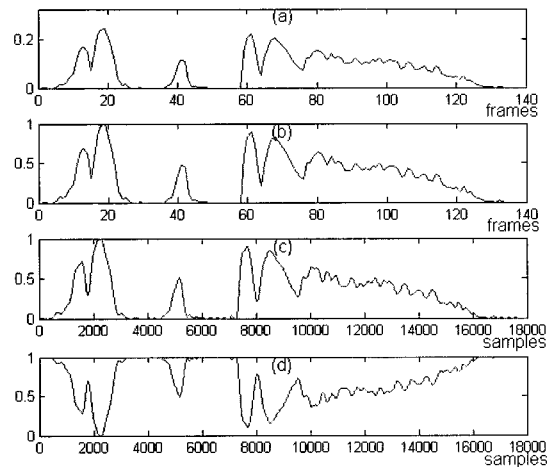


그림 3. (a) $T^{k'}$ (b) $T^{k''}$ (c) 재샘플링 (d) $\lambda^{k'}$

잡음음성을 위해서는 원음성에 white gaussian noise를 첨가하여 0dB에서 20dB까지의 잡음 음성신호를 만들어 실험을 수행하였다. 또한 잡음 제거의 객관적인 성능 평가를 위하여 SNR과 MSE를 계산한다[3]. 이를 수식으로 나타내면 다음과 같다.

$$SNR(\delta) = 10 \log \frac{\sum_i \hat{s}_i^2}{\sum_i n_i^2} \quad (10)$$

$$MSE(\delta) = \frac{\sum_{i=0}^N (s_i - \hat{s}_i)^2}{N} \quad (11)$$

여기에서 s , n , \hat{s} 는 각각 원 신호, 잡음, 원 신호의 추정값을 나타내며, N 은 샘플수이다.

SNR은 신호 대 잡음비로서 값이 클수록 잡음이 많이 제거되었음을 나타내며, MSE값은 작을수록 원음성과 유사하게 복원함을 나타낸다.

4.2 실험결과

음성특징 벡터의 추출에 필요한 계산량에 영향을 줄 뿐만 아니라 음성특징 벡터의 성질에도 영향을 주는 프레임 크기와 주파수 대역 분할을 수행하는 웨이브렛 모함수를 결정한다.

프레임 크기에 따른 성능을 평가하기 위해서 “소프트웨어” 음성샘플에 white gaussian noise 10dB를 첨가하여 잡음신호를 만들어 실험을 수행하였다. 웨이브렛 함수는 daubechies 함수[8] 8tap을 사용하였으며 분석 프레임의 1/2을 overlap-and-add 하였다[1].

표 1. 프레임 크기에 따른 SNR과 MSE의 평가

프레임 크기	SNR	MSE
128	12.783	0.0597
150	12.578	0.0626
200	12.8184	0.0592
256	12.880	0.0583
300	12.872	0.0580

표 1의 결과를 보면 256 샘플의 경우가 다른 프레임 크기의 경우보다 약간의 성능 차이를 보이고 있어서 256 샘플의 경우가 제안한 방법에 적합함을 알 수 있다.

또한 위의 실험 결과와 daubechies 함수, coiflet 함수[9], symlets 함수[9] 등의 여러 모함수를 변경하

며 실험한 결과, 256 샘플의 분석 프레임과 symlets 3tap 모함수가 제안한 방법에 적합한 것으로 평가되었다[4]. 위의 결과를 이용하여 white gaussian noise를 5dB 단위로 0dB에서 20dB까지 섞인 음성 샘플을 생성하여 실험하였다. 다음은 각각의 음성 에 대한 실험 결과이다.

4.2.1 “소프트웨어” 음성의 실험결과

그림 4와 같이 5dB white gaussian noise을 첨가하였을 경우 잡음 제거 결과는 WT(Wavelet Transform)과 WPT(Wavelet Packet Transform)을 비교하였을 때 SNR과 MSE 모두 약간의 성능 향상이 있었다.

그림 4,5,6,7에서 ‘original’은 잡음이 없는 원음성 신호, ‘noise signal’은 백색잡음이 혼합된 신호, ‘WT_denoise’는 WT를 이용한 잡음 제거 후 신호, ‘WPT_denoise’는 WPT를 이용한 잡음 제거 후 신호, 그리고 ‘proposed’는 제안된 방법으로 잡음 제거된 신호를 나타낸다.

WT와 WPT의 경우 1100~1800 샘플 구간의 파열음 성분과 잡음 성분을 구분하지 못하였으나, 제안한 알고리즘은 파열음 성분과 잡음 성분을 구분하여 잡음 제거 후에도 파열음 성분을 볼 수 있었다. MSE 값으로 비교해보면, WT는 0.155, WPT는

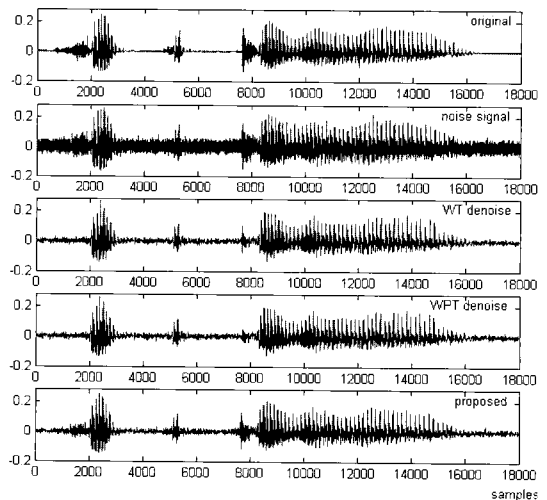


그림 4. 5dB white gaussian noise를 첨가한 “소프트웨어” 음성의 잡음 제거

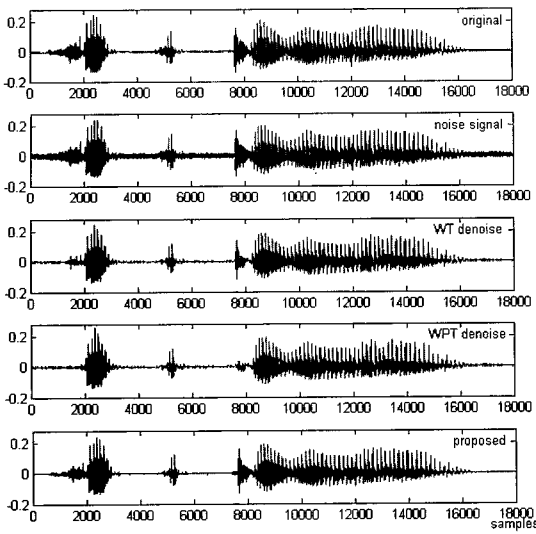


그림 5. 15dB white gaussian noise를 첨가한 “소프트웨어” 음성의 잡음 제거

0.212, 제안한 알고리즘은 0.099으로서 WT에 비해 36%, WPT에 비해 53% 정도가 향상되었으며, 원음성과 유사하게 복원함을 알 수 있다. 또한 잡음 dB가 증가할수록 제안한 알고리즘의 경우 1100~1800 샘플 구간의 파열음 성분을 원음성과 유사하게 복원하고 있음을 그림 4와 그림 5의 비교를 통해서 알 수 있다.

표 2. “소프트웨어” 음성의 SNR 비교

잡음 크기	WT	WPT	Proposed
0dB	6,8495	6,7832	6,258
5dB	9,3704	9,0316	9,825
10dB	11,3744	10,1672	12,880
15dB	13,110	10,571	15,832
20dB	15,011	10,701	18,440

표 3. “소프트웨어” 음성의 MSE 비교 (단위 : 1.0e-003)

잡음 크기	WT	WPT	Proposed
0dB	2,331	2,367	2,672
5dB	1,305	1,411	1,177
10dB	0,822	1,086	0,583
15dB	0,551	0,989	0,296
20dB	0,355	0,960	0,162

표 2와 표 3의 SNR 값과 MSE 값의 비교를 통해서 제안한 알고리즘이 잡음 dB가 증가할수록 SNR은 증가하고, MSE는 감소하게 됨을 알 수 있으며, 이는 기존의 WT와 WPT보다 잡음 제거 능력이 우수함을 보여주고 있다.

4.2.2 “1월 15일까지 할 수 있는데요” 음성의 실험 결과

그림 6과 같이 10dB white gaussian noise를 첨가하였을 경우 WT와 WPT의 경우 5000~7000 샘플 구간의 “1(십)”의 음성의 마찰음을 잡음으로 간주하고 많이 제거된 것을 볼 수 있으며, MSE 값으로 비교해보면 WT는 0.874, WPT는 0.975의 값을 나타내고 있다.

그러나 제안한 알고리즘의 경우 MSE 값이 0.475의 작은 값을 나타내므로 원음성과 유사하게 복원하고 있음을 알 수 있으며, WT에 비해서 45%, WPT에 비해서 51% 정도의 성능 향상이 있었다. 또한 17000~19000 샘플 구간의 “수”에 해당하는 마찰음 성분도 위와 비슷한 결과를 나타낸다.

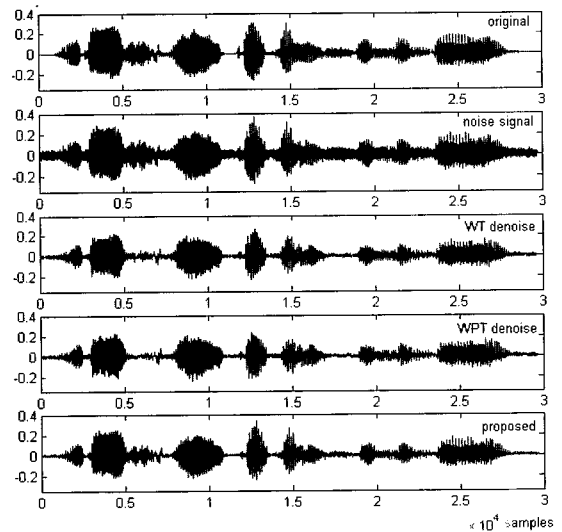


그림 6. 10dB white gaussian noise를 첨가한 “1월 15일까지 할 수 있는데요” 음성의 잡음 제거

표 4. “1월 15일까지 할 수 있는데요” 음성의 SNR

잡음 크기	WT	WPT	Proposed
0dB	5,197	5,071	5,728
5dB	7,053	6,486	8,635
10dB	8,565	7,044	11,161
15dB	10,091	7,220	13,444
20dB	11,517	7,281	15,503

표 5. “1월 15일까지 할 수 있는데요” 음성의 MSE
(단위 : 1.0e-003)

잡음 크기	WT	WPT	Proposed
0dB	0,715	0,736	0,633
5dB	0,466	0,531	0,324
10dB	0,329	0,467	0,181
15dB	0,231	0,449	0,107
20dB	0,167	0,442	0,066

표 4와 표 5는 음성 “1월 15일까지 할 수 있는데요”에 대한 잡음 크기에 따른 SNR과 MSE 값을 보여준다. 0dB의 경우 WT 및 WPT와 약간의 차이가 있었으나 잡음 dB가 증가할수록 제안한 알고리즘의 성능이 우수함을 보여준다.

4.2.3 “청와대” 음성의 실험결과

마지막으로 파찰음 성분이 있는 샘플 구간에서의 잡음 제거 성능을 평가하기 위해서 “청와대”라는 음성을 사용하여 실험을 실시하였다. 그림 7과 같이 10dB white gaussian noise를 첨가하였을 경우 WT와 WPT의 경우 1000~2000 샘플 구간의 “ㄷ”의 파찰음을 잡음으로 간주하고 많이 제거된 것을 볼 수 있으며, MSE 값으로 비교해보면 WT는 0.794, WPT는 1.156의 값을 나타내고 있다. 그러나 제안한 알고리즘의 경우 MSE 값이 0.473의 작은 값을 나타내므로 원음성과 유사하게 복원하고 있음을 알 수 있으며, WT에 비해서 40%, WPT에 비해서 59% 정도의 성능 향상이 있었다.

표 6과 표 7은 음성 “청와대”에 대한 잡음 크기에 따른 SNR과 MSE 값을 보여준다.

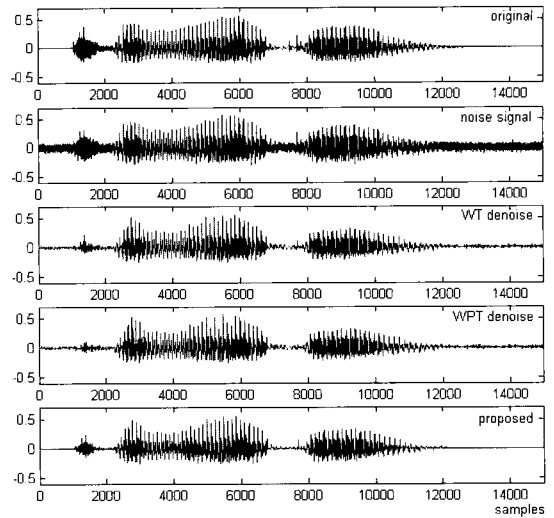


그림 7. 10dB white gaussian noise를 첨가한 “청와대” 음성의 잡음 제거

표 6. “청와대” 음성의 SNR 비교

잡음 크기	WT	WPT	Proposed
0dB	5,885	5,819	6,912
5dB	8,177	7,640	9,671
10dB	9,987	8,358	12,240
15dB	11,643	8,592	14,517
20dB	13,079	8,682	16,308

표 7. “청와대” 음성의 MSE 비교 (단위 : 1.0e-003)

잡음 크기	WT	WPT	Proposed
0dB	2,043	2,074	1,613
5dB	1,206	1,364	0,854
10dB	0,794	1,156	0,473
15dB	0,543	1,095	0,280
20dB	0,390	1,073	0,185

5. 결 론

본 논문은 잡음 환경의 음성인식을 위하여 잡음에 오염된 음성신호를 프레임 단위로 웨이브렛 변환을 수행한 후 웨이브렛 계수의 표준편차를 이용하여 시간 적응 임계값을 정하고 잡음을 제거하는 방법을 제안하였다. 음성의 특성을 고려하여 고주

과 성분을 많이 가지는 무성음의 경우는 첫 번째 스케일의 detail 신호에서, 저주파 성분을 많이 가지는 유성음의 경우는 세 번째 스케일의 approximation 신호의 표준 편차를 구하여 시간 적응 임계값을 설정하였다.

원음성에 white gaussian noise를 5dB단위로 0dB에서 20dB까지 첨가하여 제안한 시간 적응 임계값을 사용하여 잡음을 제거하였다. 단어와 문장의 음성샘플 잡음제거 실험을 통하여 웨이브렛이나 웨이브렛 패킷 변환이 파열음, 마찰음 성분 및 파찰음 성분을 많이 제거하는 반면, 제안한 알고리즘은 원 신호와 유사하게 복원하였다. 또한 일반적인 웨이브렛 변환과 웨이브렛 패킷 변환을 이용한 방법보다 SNR과 MSE가 향상됨으로써 시간 적응 임계값 방법이 효과적으로 잡음을 제거함을 알 수 있었으며, 잡음이 제거된 음성신호를 직접 들어보았을 때 불쾌감 및 거부감이 없었으며, 원래의 음성내용을 정확하게 인식할 수 있었다.

이러한 결과는 고정임계값을 쓰는 기존 웨이브렛 변환 및 웨이브렛 패킷변환의 잡음제거와 시간 적응 임계값을 사용하는 제안한 잡음제거의 차이로 볼 수 있다. 즉, 웨이브렛 변환과 웨이브렛 패킷변환의 경우, 음성샘플 전체구간에 대하여 고정된 임계값을 사용하여 잡음을 제거함으로써, 잡음이 아닌 구간에서도 잡음으로 인식하고 제거하는 반면, 제안한 시간 적응 임계값을 사용하면, 프레임단위(실제로는 overlap-and-add를 수행하므로 128 샘플구간)마다 표준편차를 이용하여 임계값을 다르게 계산하여, 잡음을 제거함으로써 제안된 방법이 기존 방법보다 향상된 결과를 나타낸다.

향후 연구과제로는 잡음을 제거하기 전과 후의 인식을 실험을 인간의 청각과 컴퓨터 프로그램을 통하여 수행함으로써 제안한 방법의 성능을 재평가하는 것이다.

참고문헌

- [1] 김현기, 이상운, 홍재근 (2001). 이산 웨이브렛 변환영역에서의 스펙트럼 차감법을 이용한 잡음 제거, 멀티미디어학회 논문지, 4(4), 306-315.
- [2] 석종원, 배건성 (1999). 웨이브렛 변환을 이용한 음성신호의 끝점검출, 한국음향학회 논문지, 18(6), 57-64.
- [3] 석종원 (1999). Wavelet Transform-Based Speech Signal Processing: Speech Enhancement and Endpoint Detection, 박사학위논문, 경북대학교.
- [4] 황향자 (2004). 음성신호에서 웨이브렛 변환을 이용한 적응 잡음 제거, 석사학위논문, 전북대학교.
- [5] Bahoura, M., Fouat, J. (2001). Wavelet speech enhancement based on the teager energy operator, IEEE Signal Processing Letters, 8(1), 10-12.
- [6] C. S. Burrus, R. A. Gopinath, H. Guo (1998). Introduction to Wavelets and Wavelet Transforms, Prentice-Hall.
- [7] Chang, S., Kwon, Y., Yang S., Kim I. (2002). Speech Enhancement for Non-stationary Noise Environment by adaptive Wavelet Packet, IEEE Transaction on Acoustics, Speech and Signal Processing, 1, 1-561-1-564.
- [8] Daubechies (1990). The Wavelet transform, time-frequency localization and signal analysis, IEEE Transactions on Information Theory, 36(5), 961-1005.
- [9] Michael Misiti, Yves Misiti, Georges Oppenheim, Jean Michael Poggi (2001). Wavelet Toolbox for Use with MATLAB, The Math Work Inc.