

# 위너필터에 의한 음성 중의 잡음제거 알고리즘

최재승\*

Noise Reduction Algorithm in Speech by Wiener Filter

Jae-Seung Choi\*

요 약

본 논문에서는 음성신호를 개선할 목적으로 잡음으로 오염된 음성신호로부터 잡음성분을 제거하기 위한 위너 필터를 사용한 잡음제거 알고리즘을 제안한다. 제안한 알고리즘은 먼저 잡음 복원 및 제거 방법에 기초하여 잡음으로 오염된 신호로부터 각 프레임에서 백색잡음의 잡음 스펙트럼을 제거한다. 또한 본 알고리즘은 선형예측 분석 방법에 기초한 위너 필터를 사용하여 음성신호를 강조한다. 본 실험에서는 일본 남성화자에 의한 음성과 잡음데이터를 사용하여 본 알고리즘의 실험 결과를 나타낸다. 백색잡음에 의하여 오염된 음성신호에 대하여 스펙트럼 왜곡률 척도를 사용하여 본 알고리즘이 유효하다는 것을 확인한다. 실험으로부터 백색잡음에 대하여 이전의 위너 필터와 비교하여 최대 4.94 dB의 출력 스펙트럼 왜곡률이 개선된 것을 확인할 수 있었다.

ABSTRACT

This paper proposes a noise reduction algorithm using Wiener filter to remove the noise components from the noisy speech in order to improve the speech signal. The proposed algorithm first removes the noise spectrums of white noise from the noisy signal based on the noise reshaping and reduction method at each frame. And this algorithm enhances the speech signal using Wiener filter based on linear predictive coding analysis. In this experiment, experimental results of the proposed algorithm demonstrate using the speech and noise data by Japanese male speaker. Based on measuring the spectral distortion (SD) measure, experiments confirm that the proposed algorithm is effective for the speech by contaminated white noise. From the experiments, the maximum improvement in the output SD values was 4.94 dB better for white noise compared with former Wiener filter.

키워드

Wiener Filter, LPC Analysis, Spectral Distortion, Noise Reduction.  
위너필터, LPC 분석, 스펙트럼 왜곡, 잡음제거

## 1. 서론

최근에 잡음으로 오염된 음성신호로부터 잡음을 제거하고 음성신호의 품질과 음성신호의 명료도를 향상시키는 문제는 현재의 실생활에 존재하는 여러 경우에 필요한 연구 과제 중의 하나이다. 지금까지 이러한 문

제의 연구 과제를 중심으로 진행된 연구 논문 등이 최근 발표되고 있다[1, 2, 3, 4, 5, 6, 7]. 또한 일반적으로 잡음에 오염된 음성신호를 추출하여 잡음을 제거하는 필터로서는, 과거에는 매치드 필터(Matched filter)[8, 9], 칼만 필터(Kalman filter)[10, 11], 위너 필터(Wiener filter)[12, 13] 등이 널리 알려져 있다. 그러나

\* 교신저자(corresponding author) : 신라대학교 전자공학과(jschoi@silla.ac.kr)  
접수일자 : 2013. 06. 25

심사(수정)일자 : 2013. 07. 23

게재확정일자 : 2013. 08. 23

음성신호에 따른 여러 조건과 그 목적의 차이로부터 이러한 것들을 그대로 잡음 중의 음성 중에서 음성만을 추출하고 잡음을 제거하는 것은 쉽지 않다. 잡음의 제거는 먼저 잡음을 성질을 충분히 파악할 필요가 있다. 잡음에는 여러 가지가 존재하지만, 본 연구에서는 정상 가우스 백색잡음(White noise)을 사용하여 실험을 진행한다.

음향적으로 잡음제거를 고려할 때에 가장 중요한 요소 중의 하나는 인간의 청각이 위상정보에 둔감하고, 주로 진폭 스펙트럼 정보, 즉 진폭정보를 중요한 실마리로 인식하여 음성을 지각하고 있다는 사실을 들 수 있다[14]. 따라서 본 연구에서는 이러한 관점으로부터 인간의 청각 기강의 성질을 충분히 고려하여 음성의 진폭 스펙트럼 성분의 추출을 첫 번째의 목표로 하는 위너 필터의 특성을 제안한다. 따라서 본 논문에서 제안하는 알고리즘은 선형 예측 분석(Linear Predictive Coding, LPC)[15, 16]에 의한 위너 필터를 사용하여 잡음성분을 제거하는 비선형적인 잡음 제거 알고리즘이다. 또한 실제의 위너 필터의 실험을 위해서, LPC를 실행하여 주파수 영역에서 위너 필터를 구성하여 잡음이 중첩된 음성 중에서 잡음을 제거하는 위너 필터 알고리즘을 구성한다.

본 실험에서는 위너 필터 알고리즘을 사용하여 각 프레임에서 잡음을 제거하는 실험결과를 나타낸다. 잡음에 의하여 오염된 음성에 대하여 스펙트럼 왜곡률(Spectral Distortion, SD)[17]를 사용하여 본 알고리즘이 유효하다는 것을 확인한다. 백색잡음에 대하여 4.94 dB 이상이 개선된 것을 확인할 수 있었다.

본 논문에서의 구성은 다음과 같다. 2장에서는 기존의 잡음 처리 방법인 위너 필터와 스펙트럼 차감 방법에 대하여 설명하고, 3장에서는 본 논문에서 제안하는 위너 필터 알고리즘의 개요 및 방법을 기술한다. 4장에서는 실험에 사용한 음성데이터 및 잡음데이터를 기술하며, 또한 본 논문에서 제안한 알고리즘의 실험 결과를 나타낸다. 마지막으로 5장에서는 결론을 기술한다.

## II. 기존의 잡음 처리 방법

본 장에서는 잡음 환경 하에서 음성인식 및 음성강

조 분야에 응용되는 잡음 처리 방법의 하나인 기존의 위너 필터 방법과 스펙트럼 차감 방법(Spectral subtraction, SS)에 대하여 기술한다.

그림 1은 기존의 위너 필터에 의한 방법의 블록도를 나타낸다. 그림 1에서 위너 필터는 원 신호와 필터링된 신호와의 차를 최소화하도록 하며, 이 필터는 잡음으로 오염된 입력 음성신호를 처리하여 신호대 잡음비(Signal-to-noise ratio, SNR)를 최대한으로 향상시키는 것을 목적으로 한다. 이것은 음성인식의 성능을 개선하기 위하여 널리 이용하는 방법이다. 본 블록에서 잡음성분을 이산 여현 변환(Discrete Cosine Transform, DCT)을 한 후에 잡음성분을 추정한다. 추정된 잡음성분을 위너 필터에 의하여 원 신호에 가까운 신호로 처리한 후에 역이산 여현 변환(Inverse Discrete Cosine Transform, IDCT)을 하여 출력음성을 만든다.

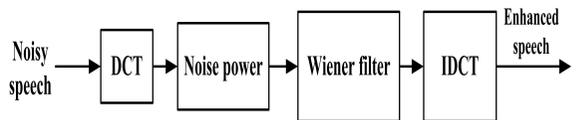


그림 1. 기존의 위너 필터법의 블록도  
Fig. 1 Block diagram of conventional Wiener filter method

스펙트럼 차감법[4]은 원하는 신호의 짧은 프레임 구간에서 진폭 스펙트럼을 추정하는 방법의 하나이며, 잡음이 중첩된 입력 음성신호의 스펙트럼으로부터 잡음신호의 스펙트럼을 차감하는 것으로부터 원하는 신호의 스펙트럼을 추정할 수 있다.

음성신호를  $s(t)$ , 잡음을  $n(t)$ 이라 하면, 관측된 신호  $y(t)$ 은 다음식과 같이 나타낼 수 있다.

$$y(t) = s(t) + n(t) \quad (1)$$

이 신호를 시간축 상에서 짧은 구간으로 분할하여, 시간축에 통과시켜 단구간 푸리에 변환(Fourier transform)을 실시함으로써 입력신호의 단구간 전력스펙트럼  $|Y(k)|^2$ 를 구한다. 여기에서, 음성신호  $s(t)$ 와 잡음  $n(t)$ 는 서로 상관이 없다고 가정하며, 또한 잡음이 정상적이라고 가정한다. 스펙트럼 차감 방법에서는 다음 식에 의하여 목적신호의 단구간 스펙트럼의 추정값  $|\hat{S}(k)|^2$ 를 구한다.

$$|\hat{S}(k)|^2 = |Y(k)|^2 - E[N(k)]^2 \quad (2)$$

식 (2)로부터 단구간 진폭 스펙트럼의 추정값  $|\hat{S}(k)|$ 를 구하여, 이것에 입력신호의 위상을 부가하여 역 푸리에변환을 수행함으로써 구하고자 하는 목적신호를 회복한다. 기존의 스펙트럼 차감법의 블록도를 그림 2에 나타낸다. 스펙트럼 차감 방법에서는 목적신호의 단구간 진폭 스펙트럼을 추정할 때에 잡음을 정상으로 가정하여, 잡음의 평균 진폭 스펙트럼을 사용한다. 그러나 가령 정상적이어도 잡음의 단구간 진폭 스펙트럼은 평균값의 주위에 분산하고 있어서, 이 분산 때문에 목적음의 단구간 진폭 스펙트럼의 추정 정확도가 떨어지기도 한다.

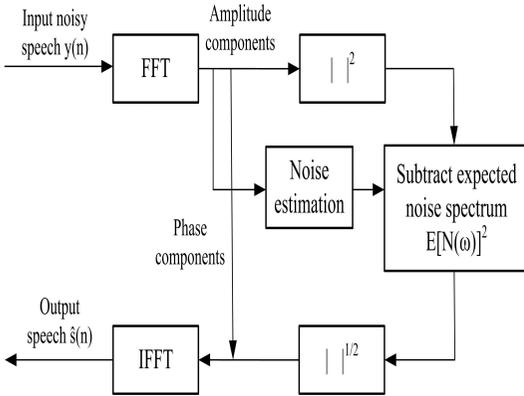


그림 2. 기존의 스펙트럼 차감법의 블록도  
Fig. 2 Block diagram of conventional spectrum subtraction

### III. 제안한 위너 필터 알고리즘

그림 3은 본 논문에서 제안하는 에러 가중치 필터 방법에 의한 알고리즘을 나타낸다. 본 알고리즘은 잡음으로 오염된 음성신호를 전처리 과정을 거친다. 전처리 과정은 전체 음성구간을 한 프레임이 128 샘플인 구간으로 분리한 후에 해밍창 (Hamming window)을 거친다. 이 후에 처음의 6 프레임을 잡음 구간으로 판단하여 잡음 성분으로 추정하여, 음성신호 중에서 잡음 성분을 어느 정도 제거한다. 그리고 본 논문에서는 LPC 분석에 의하여 10차의 LPC 켈프스트럼 계수를 이용하였다. LPC 켈프스트럼 계수는 자기상관 함수에 의

하여 잡음 신호로부터 구한 자기상관 계수를 이득으로 나눈 값을 합하여 10차의 LPC 켈프스트럼 계수를 구한다. 마지막으로 추정된 깨끗한 음성 성분과 추정된 잡음 성분을 이용하여 최종적인 위너 필터를 설계한다. 따라서 구해진 위너 필터로 현재 프레임을 처리하여 깨끗한 음성 신호를 만드는 과정을 반복한다. 최종적으로 위너 필터에 의한 음성신호를 재 구성함으로써 강조된 음성신호를 만들 수 있다.

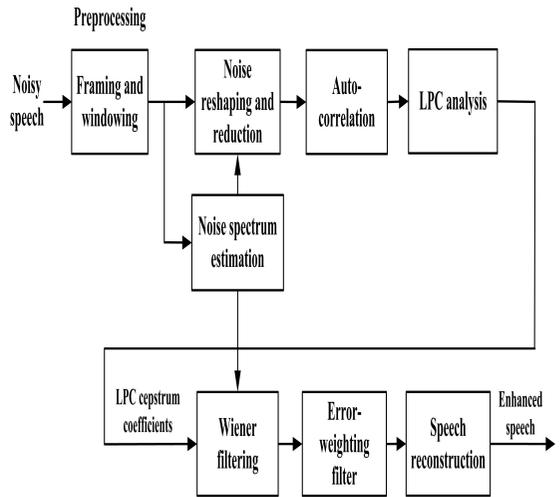


그림 3. 제안하는 위너 필터 알고리즘  
Fig. 3 Proposed wiener filter algorithm

## IV. 실험 및 결과

### 4.1 음성데이터

본 실험에서 사용한 음성데이터는 일본인 남성화자에 의한 단어 “bizen(JM1)”, “shimonoseki(JM2)”를 사용하였으며, 샘플링 주파수는 8 kHz이며 양자화 비트수는 12비트이다. 그리고 백색잡음은 컴퓨터에 의해서 작성한 가우스 잡음을 사용하였으며 샘플링 주파수는 8 kHz이다.

본 실험에 있어서의 평가방법으로는 스펙트럼 왜곡율(Spectral Distortion, SD)[17]을 사용한다. SD는 입력 신호의 대수 스펙트럼의 차를 구한 것으로 식 (3)과 같이 정의한다.

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^N \frac{1}{W} \int_0^w S_x(w) - S_y(w)^2 dw} \quad (3)$$

단,  $S_x(w)$ ,  $S_y(w)$ 는 입출력 신호의 대수 스펙트럼(dB)이며,  $N$ 은 측정구간의 프레임 수,  $W$ 는 신호의 대역폭이다.

#### 4.2 특징벡터 추출

음성 실험에 사용하는 특징벡터는 10차의 선형예측 코스트럼 계수(LPC)를 사용한다. 입력되는 음성신호는 각 프레임에서 128샘플 단위로 해밍창을 한다. 본 실험에 사용한 음성신호는 한 프레임의 분석구간의 길이가 16 ms이며, 분석구간의 테이터에 대하여 식 (4)의 해밍창을 통과시켰다. 여기에서  $N$ 은 분석구간의 음성테이터의 샘플수이다.

$$W(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) (n = 0, \dots, N-1) \quad (4)$$

#### 4.3 실험 결과

본 논문에서는 위너 필터에 의한 알고리즘 구현을 위하여, 제안한 알고리즘을 이용하여 실험을 한다. 실험환경은 Intel(R) Core(TM)2 Quad 2.4GHz CPU와 3.25GB RAM이 장착된 IBM 호환 컴퓨터에서 Windows XP OS를 사용하여 Microsoft VC++ 6.0.0으로 구현하였다.

본 논문은 입력되는 음성신호에 포함되는 잡음을 제거하기 위하여, SD를 사용하여 음성 데이터에 대한 잡음제거의 실험결과에 대해서 기술한다.

표 1과 표 2는 제안한 알고리즘과 기존의 위너필터 [18]에 의한 방법을 비교한 실험결과이다. 표 1은 백색잡음으로 오염된 음성신호의 입력 SD값이 각각 21.10 dB과 18.01 dB일 경우에, 제안한 알고리즘의 출력 SD값은 각각 10.15 dB과 9.85 dB[표 1의 Prop.]이 구해졌으며, 입력 SD값과 비교하여 각각 9.95 dB과 8.16 dB [표 1의 Impr1]이 개선된 것을 알 수 있다. 또한 제안한 알고리즘의 출력 SD값은 이전의 위너 필터 방법 [18]보다 각각 4.12 dB과 4.94 dB[표 1의 Impr2]이 개선되었다. 표 2는 백색잡음으로 오염된 음성신호의 입력 SD값이 각각 20.77 dB과 18.76 dB일 경우에, 제안한 알고리즘의 출력 SD값은 각각 11.46 dB과 11.14

dB[표 2의 Prop.]이 구해졌으며, 입력 SD값과 비교하여 각각 9.31 dB과 7.62 dB[표 2의 Impr1]이 개선된 것을 알 수 있다. 또한 제안한 알고리즘의 출력 SD값은 이전의 위너 필터 방법보다 각각 4.61 dB과 3.83 dB[표 2의 Impr2]이 개선되었다.

표 1과 표 2의 결과로부터 본 논문에서 제안한 위너 필터 알고리즘에 의하여 출력 SD값이 최대 4.94 dB[표 1의 경우]이 개선된 것을 알 수 있다. 따라서 본 논문에서 제안한 알고리즘은 백색잡음의 중첩된 양에 따라서 음성 중의 잡음 성분을 제거하는 효과가 서로 다르다는 것을 알 수 있다.

SD 및 SNR의 평가방법이 서로 다르기 때문에 단순히 비교하기는 쉽지 않지만, PCM(Pulse Code Modulation) 방식의 평가에 의하면 SD의 약 3.5 dB의 개선량은 SNR의 20 dB에 해당한다고 알려져 있기 때문에(문헌 [19]), 원 음성으로부터의 SD의 개선이 4.94 dB로 상당한 효과가 있다고 판단된다.

표 1. 백색잡음을 사용한 경우의 기존 방법과의 SD 성능비교 (Input SD = 20.10 dB, 18.01 dB)

Table 1. The SD comparison for the conventional method using white noise (Input SD = 20.10 dB, 18.01 dB).

Data	Input SD[dB]	Output SD [dB]		Impr. SD[dB]	
		Pre-Wiener	Prop.	Impr1	Impr2
JM1	20.10	15.98	10.15	9.95	4.12
JM2	18.01	14.79	9.85	8.16	4.94

표 2. 백색잡음을 사용한 경우의 기존 방법과의 SD 성능비교 (Input SD = 20.77 dB, 18.76 dB)

Table 2. The SD comparison for the conventional method using white noise (Input SD = 20.77 dB, 18.76 dB).

Data	Input SD[dB]	Output SD [dB]		Impr. SD[dB]	
		Pre-Wiener	Prop.	Impr1	Impr2
JM1	20.77	16.07	11.46	9.31	4.61
JM2	18.76	14.97	11.14	7.62	3.83

## V. 결론

본 논문에서는 인간의 청각에서 중요하게 고려하는 음성의 진폭 스펙트럼 성분을 강조하는 것을 목적으로

하여 위너 필터의 특성에 기초한 알고리즘을 제안하였다. 본 논문에서 제안한 알고리즘은 먼저 잡음 복원 및 제거 방법에 기초하여 잡음으로 오염된 신호로부터 각 프레임에서 백색잡음의 잡음 스펙트럼을 제거한다. 또한 본 알고리즘은 선형예측 분석 방법에 기초한 위너 필터를 사용하여 음성신호를 강조한다. 특히 본 알고리즘은 선형 예측 분석에 의한 위너 필터를 사용하여 잡음성분을 제거하는 비선형적인 잡음 제거 알고리즘이다. 또한 실제의 위너 필터의 실현을 위해서, LPC를 실행하여 주파수 영역에서 위너 필터를 구성하여 잡음이 중첩된 음성 중에서 잡음을 제거하는 위너 필터 알고리즘을 구성하였다.

본 실험에서는 일본 남성화자에 의한 음성신호와 잡음신호를 사용하여 본 알고리즘의 실험 결과를 나타내었다. 백색잡음에 의하여 오염된 음성신호에 대하여 스펙트럼 왜곡률 척도를 사용하여 본 알고리즘이 유효하다는 것을 확인하였다. 실험으로부터 백색잡음에 대하여 이전의 위너 필터와 비교하여 최대 4.94 dB의 출력 스펙트럼 왜곡률이 개선된 것을 확인할 수 있었다. 향후의 연구과제로는 다양한 배경잡음 및 음성데이터를 사용하여 본 알고리즘을 개선할 필요가 있다.

## 참고 문헌

- [1] K. Daqrouq, I. N. Abu-Isbeih, M. Alfauri, "Speech signal enhancement using neural network and wavelet transform", 2009 6th International Multi-Conference on Systems, Signals and Devices, pp. 1-6, 2009.
- [2] W. G. Knecht, M. E. Schenkel, G. S. Moschytz, "Neural network filters for speech enhancement", IEEE Trans. Speech and Audio Processing, Vol. 3, No. 6, pp. 433-438, 1995.
- [3] J. Chen, K. K. Paliwal and S. Nakamura, "Sub-band based additive noise removal for robust speech recognition", in Proc. Eurospeech, pp. 571-574, 2001.
- [4] S.F. Boll, "Suppression of acoustic noise in speech using spectral subtraction", IEEE Trans. Acoust., Speech, Signal Processing. Vol. 27, No. 2, pp. 113-120, 1979.
- [5] Y.P. Kim and H. Y. Lee, "A Study on Improved Method of Voice Recognition Rate", The Journal of The Korea Institute of Electronic Communication Sciences, Vol. 8, No. 1, pp. 77-83, 2013.
- [6] C.K. Lee and D.I. Kim, "Adaptive Noise Reduction of Speech Using Wavelet Transform", The Journal of The Korea Institute of Electronic Communication Sciences, Vol. 4, No. 3, pp. 190-196, 2009.
- [7] C.Y. Lee, "A Study on the Removal of Unusual Feature Vectors in Speech Recognition", The Journal of The Korea Institute of Electronic Communication Sciences, Vol. 8, No. 4, pp. 561-567, 2013.
- [8] J. Freudenberger, S. Stenzel, "Blind Matched Filtering for Speech Recording in Uncorrelated Noise", International Workshop on Acoustic Signal Enhancement, pp. 1-4, September 2012.
- [9] J.S. Hu, C.H. Yang, "Speech enhancement using transfer function ratio beamformer and matched filter array", International Conference on Information and Automation, pp. 1161-1166, June 2009.
- [10] Y. Wang, J. An, V. Sethu, E. Ambikairajah, "Perceptually motivated pre-filter for speech enhancement using Kalman filtering", 2007 6th International Conference on Information, Communications & Signal Processing, pp. 1-4, December 2007.
- [11] Y. Shao, C.H. Chang, "A Kalman filter based on wavelet filter-bank and psychoacoustic modeling for speech enhancement", 2006 IEEE International Symposium on Circuits and Systems, ISCAS 2006. Proceedings, May, 2006.
- [12] L. Lin, W. H. Holmes and E. Ambikairajah, "Subband noise estimation for speech enhancement using a perceptual Wiener filter", IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 1, pp. I-80 - I-80-3, 2003.
- [13] X. Dang, T. Nakai, "Noise reduction using modified phase spectra and Wiener Filter", 2011 IEEE International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1-5, September, 2011.
- [14] J. H. L. Hansen and S. Nandkumar, "Robust estimation of speech in noisy backgrounds based on aspects of the auditory process", The Journal of the Acoustical Society of

- America, Vol. 97, No. 6, pp. 3833-3849, 1995.
- [15] X. Zhang, Y. Guo, Xuemei Hou, " A speech Recognition Method of Isolated Words Based on Modified LPC Cepstrum", IEEE International Conference on Granular Computing, pp. 481-484, 2007.
- [16] P.B. Patil: Multilayered network for LPC based speech recognition. IEEE Transactions on Consumer Electronics, Vol. 44, No. 2, pp. 435 - 438, 1998.
- [17] S.H. Jo, C.D. Yoo, "Psychoacoustically Constrained and Distortion Minimized Speech Enhancement," IEEE Transactions on Audio, Speech, and Language Processing, Vol. 18, No. 8, pp. 2099-2110, November 2010.
- [18] J. S. Choi, "Formant Enhancement Algorithm of Speech Using Auditory Filter", Journal of Korean Institute of Information Technology, Vol. 11, No. 7, June 2013.
- [19] K. Itoh, N. Kitawaki, and K. Kakehi, "A Study of Objective Quality Measures for Digital Speech Waveform Coding Systems", The Journal of Institute of Electronics, Information and Communication Engineers, Vol. J66-A, No. 3, pp. 274-281, 1983.

## 저자 소개



### 최재승(Jae-Seung Choi)

1989년 조선대학교 전자공학과 공학사

1995년 일본 오사카시립대학 전자정보공학부 공학석사

1999년 일본 오사카시립대학 전자정보공학부 공학박사

2000년~2001년 일본 마쯔시타 전기산업주식회사 (현, 파나소닉 주식회사) AVC사 연구원

2002년~2007 경북대 디지털기술연구소 책임연구원

2007년~현재 신라대학교 전자공학과 교수

※ 관심분야 : 음성신호처리, 신경회로망, 적응필터와 잡음제거, 디지털 TV 및 멀티미디어 등