

신경망 구조의 적응 Wiener 필터를 이용한 비선형 잡음감쇠기

이행우*

Nonlinear Noise Attenuator by Adaptive Wiener Filter with Neural Network

Haeng-Woo Lee*

요약

본 논문은 음향잡음감쇠기에서 신경망 구조의 Wiener 필터를 이용하여 비선형 잡음을 감쇠시키는 방법에 대하여 연구하였다. 이 시스템은 기존의 적응필터를 이용하는 대신 신경망 위너필터를 이용한 심층학습 알고리즘으로 비선형 잡음감쇠 성능을 개선한다. 128-neuron, 8-neuron 은닉층과 오차 역전파(back propagation) 알고리즘을 이용하여 비선형 잡음이 포함된 단일입력 음성신호로부터 음성을 추정한다. 본 연구에서 비선형 잡음에 대한 감쇠 성능을 검증하기 위하여 Keras 라이브러리를 사용한 시뮬레이션 프로그램을 작성하고 모의 실험을 수행하였다. 모의실험 결과, 본 시스템은 비선형 잡음이 포함되어 있는 경우에도 위너필터 대신 FNN 필터를 사용하면 잡음감쇠 성능이 상당히 개선되는 것을 볼 수 있다. 이는 FNN 필터의 복잡한 구조가 어떤 형태의 비선형 특성도 잘 표현하기 때문이다.

ABSTRACT

This paper studied a method of attenuating nonlinear noise using a Wiener filter of a neural network structure in an acoustic noise attenuator. This system improves nonlinear noise attenuation performance with a deep learning algorithm using a neural network Wiener filter instead of using a conventional adaptive filter. A voice is estimated from a single input voice signal containing nonlinear noise using a 128-neuron, 8-neuron hidden layer and an error back propagation algorithm. In this study, a simulation program using the Keras library was written and a simulation was performed to verify the attenuation performance for nonlinear noise. As a result of the simulation, it can be seen that the noise attenuation performance of this system is significantly improved when the FNN filter is used instead of the Wiener filter even when nonlinear noise is included. This is because the complex structure of the FNN filter expresses any type of nonlinear characteristics well.

키워드

Deep Learning, Neural Network, Noise Attenuation, Nonlinear, Wiener Filter

심층 학습, 신경망, 잡음 감쇠, 비선형, 위너 필터

* 남서울대학교 지능정보통신공학과
• 접수일 : 2022. 11. 24
• 수정완료일 : 2023. 01. 03
• 게재확정일 : 2023. 02. 17

• Received : Nov. 24, 2022, Revised : Jan. 03, 2023, Accepted : Feb. 17, 2023
• Corresponding Author : Haeng-Woo Lee
Dept. of Information Communication Engineering, Namseoul University,
Email : hwlee@nsu.ac.kr

I. 서 론

음성신호에 포함된 잡음을 감쇠시키는 음성개선기술에 대해 지금까지 많은 연구가 이루어지고 있다. 적응 잡음감쇠기는 음향센서의 수에 따라 단일입력과 다중입력 시스템으로 구분되는데 단일입력시스템[1]은 하나의 마이크를 통해서 음성신호가 입력된다. 음성신호의 유성음 구간이 나타내는 준주기적 특성을 이용하면 잡음이 포함된 마이크 입력신호로부터 음성신호를 추정할 수 있다.

선형적응방식은 비선형성이 있는 경우 부경로를 정확히 추정하지 못한다. 증폭기나 스피커와 같은 전자장치의 제한된 품질로 인해 능동잡음제어(ANC) 분야에서 잡음방지에 비선형 왜곡이 불가피하게 유입된다. 결과적으로 부정확하게 추정된 부경로는 전체 노이즈 감쇠 성능을 저하시킨다. 부경로에 비선형 요소가 포함된 경우 작은 비선형성이 적응필터 동작에 상당한 영향을 미치게 된다. 비선형 왜곡을 해결하기 위해 많은 연구에서 적응형 비선형 ANC 알고리즘이 제안되었다. Volterra 확장은 부드럽거나 약한 비선형성을 모델링하는데 효과적인 것으로 나타났으며[2], 비선형 왜곡이 있는 경우 순방향 능동잡음제어를 위해 잘린 2차 Volterra 기반 FxLMS 알고리즘이 제안되었다[3-4]. Napoli와 Piroddi[5]는 보다 효율적이고 신뢰할 수 있는 비선형 ANC를 위해 외인성 변수가 있는 다항식 비선형 Autoregressive 모델을 사용하였다. 비선형 FxLMS와 접선 쌍곡선 함수 기반 FxLMS는 부경로를 포화형 비선형성으로 모델링하여 ANC 시스템의 비선형성을 처리하기 위해 도입되었다[6]. Bilinear FxLMS[7], Filtered-s LMS[8], Leaky FxLMS[9]와 같은 다른 알고리즘도 비선형성을 해결하기 위해 연구되었다. 그러나 강한 비선형성이 있는 경우 성능이 제한된다.

비선형 관계를 처리하는 능력을 이용하여 신경망 기술이 비선형 ANC를 처리하기 위해 도입되었다[10]. 다층 퍼셉트론(MLP) 신경망은 진동의 능동적 제어를 위해 도입되었으며[11], 여기서 신경망의 가중치는 적응 filtered-x 역전파(back propagation)를 사용하여 업데이트된다. 또한 비선형 능동제어구조를 기반으로 수렴속도를 높이고 훈련의 연산량 부하를 줄이기 위해 개선된 훈련 알고리즘을 사용하였다[12-13]. 그리

고 ANC에서 비선형 효과를 해결하기 위해 기능적 연결 신경망을 기반으로 하는 효율적인 ANC 구조를 제안하였다[14-15].

딥러닝은 신경망을 기반으로 많은 수의 은닉층을 사용하는 복잡한 머신러닝 모델이다. 다층 신경망을 학습시키는 오차 역전파 알고리즘을 사용함으로써 많은 층으로 구성된 심층 신경망(DNN)도 학습이 가능하게 되었다[16-19]. 딥러닝은 복잡한 비선형 관계를 모델링할 수 있으며 잠재적으로 비선형 ANC 문제를 해결하는 데 중요한 역할을 할 수 있다. 실제 애플리케이션에 적합하려면 ANC 시스템이 다양한 소음을 감쇠하고 변화하는 음향환경에 대처할 수 있어야 한다. 딥러닝 기반 지도 ANC의 경우 훈련 중 ANC를 다양한 노이즈와 변동에 노출시키려면 대규모 다중 조건 훈련이 필요하다. 본 논문에서는 특히 비선형 ANC 문제를 해결하기 위해 딥러닝 기술을 제안한다. 능동잡음제어기의 적응필터 대신에 FNN 신경망 필터의 심층학습(deep learning) 알고리즘을 이용하여 비선형 잡음을 감쇠시키는 방법이다. 우리의 접근방식은 다양한 노이즈 소스에 해당하는 최적의 제어 매개변수를 인코딩하도록 훈련된 딥러닝 모델을 사용한다.

논문의 내용은 II절에서 Wiener 기반 비선형 잡음 감쇠기에 대해 알아보고, III절에서는 FNN 신경망 필터의 구조를 설명하였다. 그리고 IV절에서 선형 및 비선형 시스템에 대한 시뮬레이션 및 그 결과에 대하여 기술하였고, 끝으로 V절에서 결론을 도출하였다.

II. Wiener 필터 기반 비선형 잡음감쇠기

그림 1은 위너(Wiener) 필터를 기반으로 한 적응 잡음감쇠기를 보여주고 있다. 음성신호의 준주기적 특성을 이용하여 지연된 신호들로부터 현재 음성을 추정한다. 또한 이것은 단일 입력을 사용하는 능동 잡음 감쇠기이다. 지연된 음성 샘플들 간에는 높은 상관관계를 갖지만 잡음과는 어떤 상관관계가 존재하지 않는다. 그러므로 음성신호는 잡음과 독립된 상태에서 음성 추정오차가 최소가 되도록 수렴해나간다.

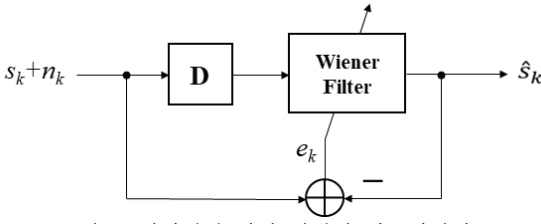


그림 1. 위너필터 기반 비선형 잡음감쇠기
Fig. 1 Wiener filter based nonlinear noise attenuator

필터 계수는 평균 제곱 오차가 최소화되도록 갱신된다. 이 오차는 입력신호의 원래 음성성분에 대한 추정 음성의 차이를 말한다. 정상상태에서 예측된 신호가 원래 음성신호와 동일한 경우 필터가 완전히 수렴된 것이다.

음성신호와 잡음은 서로 상관관계가 없다고 가정하면 음성신호의 추정값은 식 (1)과 같이 구해진다.

$$\hat{s}_k = \sum_{n=0}^{N-1} w_k(n) x_k(k-n) \quad (1)$$

여기서 N은 필터의 차수이며 위너필터의 출력은 입력신호에 포함된 음성신호를 추정한다. 이 신호가 입력신호 $x_k (= s_k + n_k)$ 로부터 차감되면 추정 오차가 되며 잡음신호를 포함하게 된다. 추정 오차는 식 (2)로 표현된다.

$$e_k = x_k - \hat{s}_k \quad (2)$$

잡음추정신호는 필터의 계수 조정을 위한 오차로 사용되며 식 (3)과 같이 오차신호의 평균전력이 최소가 되도록 업데이트해 나간다.

$$E\{e_k^2\} = 0 \quad (3)$$

여기서 $E\{\cdot\}$ 는 평균값이며 최소값일 때 조건은

$$\partial(E\{e_k^2\})/\partial w_k = 0 \quad (4)$$

즉 풀어서 다시 쓰면

$$\begin{aligned} \partial(E\{e_k^2\})/\partial w_k &= 2e_k E\{\partial e_k/\partial w_k\} \\ &= -2e_k E\{\partial \hat{s}_k/\partial w_k\} \end{aligned} \quad (5)$$

식 (1)로부터 벡터 형식으로 표현하면

$$\partial(E\{e_k^2\})/\partial w_k = -2e_k \mathbf{x}_k \quad (6)$$

이것이 필터 계수가 조정되어야 할 값이며 계수의 갱신은 식 (7)과 같이 이루어진다.

$$\mathbf{w}_{k+1} = \mathbf{w}_k - 2\delta e_k \mathbf{x}_k \quad (7)$$

여기서 δ 는 조정스텝의 크기를 나타내며 $E\{e_k^2\}$ 를 최소화시키는 것은 음성신호의 추정오차 $E\{(s_k - \hat{s}_k)^2\}$ 를 최소화시키는 것이다. 이때 필터의 출력인 음성신호의 추정치 \hat{s}_k 는 음성신호를 가장 잘 추정하게 된다.

본 논문에서는 Wiener 필터를 FNN 신경망 필터로 대체함으로써 비선형 성분까지도 추정하고자 한다. FNN 신경망 구조는 모든 뉴런에 격자형으로 연결되어 있어 선형뿐만 아니라 샘플들 간의 비선형 성분도 나타낼 수 있다.

III. FNN 신경망 필터의 구조

FNN 신경망은 한 계층의 모든 뉴런이 다른 계층의 모든 뉴런에 완전히 연결되어 있는 구조이다. 이러한 신경망은 샘플들 간의 연결 형태, 즉 선형 및 비선형 관계에 구애받지 않는다. 그러므로 입력 데이터에 대해 특별한 가정이 필요하지 않다.

FNN 기반 신경망 필터는 모든 뉴런과 연결되기 때문에 연산량이 많지만 은닉층의 수에 따라 복잡한 비선형 특성을 표현할 수 있다. 본 논문에서는 2-Layer 은닉층과 각 뉴런의 수를 128개 및 8개로 구성하였다. 그림 2에서 회로 구조를 보여주고 있으며 모든 뉴런의 입력 데이터들이 중첩되어 연결된 것을 볼 수 있다.

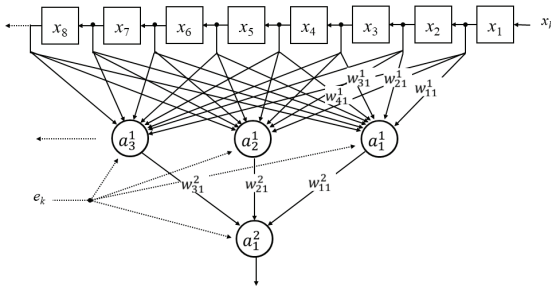


그림 2. FNN 신경망 구조
Fig. 2 Structure of FNN

바이어스(bias)는 제로라고 가정하며 활성화 함수로 ReLU(Rectified Linear Unit)를 이용하면 각 계층(layer)의 출력은 식 (8), (9)와 같이 구해진다.

$$a_j^1 = \phi_{ReLU}(u_j^h) = \begin{cases} \sum_{i=1}^l w_{ij}^1 x_i & \text{for } u_j^h > 0 \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

$$a_k^2 = \phi_{ReLU}(u_k^o) = \begin{cases} \sum_{j=1}^m w_{jk}^2 a_j^1 & \text{for } u_k^o > 0 \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

목표값 대비 출력뉴런의 오차값은 $e_k = x_k - a_k^2$ 이고 활성화함수의 미분값은 $\phi'(u) = 1$ ($for u > 0$) 이다. NLMS(Normalized Least Mean Square) 알고리즘을 이용하면 각 레이어의 가중치는 $u > 0$ 인 경우에 식 (10), (11)과 같이 업데이트된다.

$$w_{jk}^2(t+1) = w_{jk}^2(t) + \eta e_k \frac{a_j^1}{E\{a_m^1\}} \quad (10)$$

$$w_{ij}^1(t+1) = w_{ij}^1(t) + \eta \sum_{k=1}^m w_{jk}^2 e_k \frac{x_i}{E\{x_i\}} \quad (11)$$

IV. 모의실험 결과

본 논문에서 제안한 비선형 잡음감쇠기의 성능을 검증하기 위해 Keras 라이브러리를 이용하여 Spyder 툴로 시뮬레이션 프로그램을 작성하였다. 입력신호는 음성과 백색잡음이 혼합되어 8kHz로 샘플링되어 16-bit로 표현되며, 100,000 샘플(12.5sec)을 준비하였

다. 이 시스템은 지도학습에 해당되므로 입력데이터는 내부적으로 $128 \times 99,980$ 샘플의 입력배열과 99,980 샘플의 목표값으로 구성된다. 잡음감쇠 성능은 다음과 같이 정의되는 NMSE(Normalized Mean Square Error)를 사용한다.

$$NMSE[dB] = 10 \log \left(\frac{E\{(e_k)^2\}}{E\{x_i\}^2} \right) \quad (12)$$

또한 음성 예측값과 잔류 오차의 비를 데시벨 단위로 나타내는 SNR을 사용한다.

그림 3에서 선형 위너필터의 NMSE 특성곡선을 보여주고 있다. 여러 군데에서 비선형 잡음성분이 감소되지 않고 많이 남아있는 것을 볼 수 있다. 그리고 그림 4에서는 신호대잡음비를 데시벨 단위로 보여준다. 이 그림으로부터 비선형 잡음성분 때문에 잡음감쇠 성능이 상당한 정도로 영향을 받고 있음을 알 수 있다.

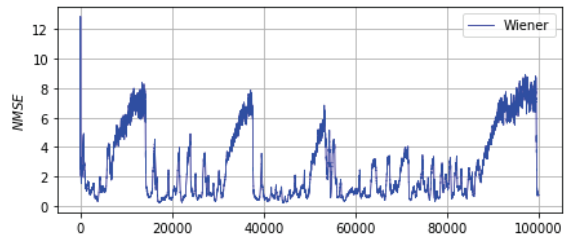


그림 3. 선형 위너필터의 NMSE 특성곡선
Fig. 3 NMSE curve of Wiener filter

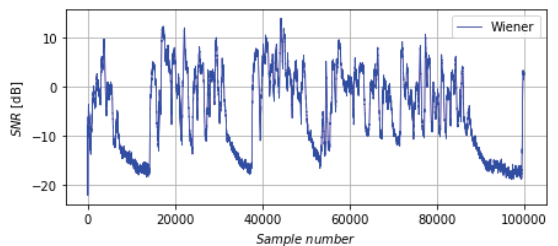


그림 4. 선형 위너필터의 SNR 특성곡선
Fig. 4 SNR curve of Wiener filter

그림 5는 비선형 FNN 필터의 NMSE 특성곡선을 나타낸다. 곡선으로부터 선형 위너필터에 비해 FNN

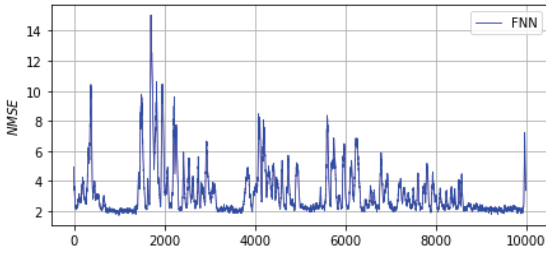


그림 5. 비선형 FNN 필터의 NMSE 특성곡선
Fig. 5 NMSE curve of FNN filter

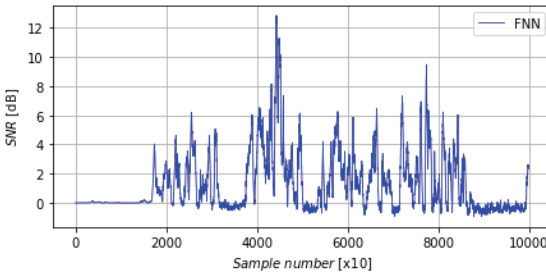


그림 6. 비선형 FNN 필터의 SNR 특성곡선
Fig. 6 SNR curve of FNN filter

필터에서는 비선형 잡음성분이 많이 감소하였음을 알 수 있다. 또한 그림 6에서도 비선형 FNN 필터의 SNR 특성곡선을 보여주고 있다. 이 그림으로부터 비선형 잡음에 대해서도 양호한 SNR 성능을 유지한다는 것을 알 수 있다.

V. 결 론

본 논문에서 FNN 딥러닝 기술을 이용한 잡음감쇠기에서 비선형 잡음이 시스템 성능에 미치는 영향을 살펴보았다. 잡음감쇠기는 128-neuron, 8-neuron 은닉층과 오차 역전과 알고리즘을 이용하여 구현하였다.

Keras 라이브러리를 사용하여 모델을 코딩하였고 필터 구성에 따라 NMSE 및 SNR 값이 어떻게 변화하는지 관찰하였다. 모의실험 결과, 본 시스템은 비선형 잡음이 포함된 경우에도 FNN 필터를 사용하면 위너필터를 사용하는 것에 비해 잡음감쇠 성능이 개선되는 것을 보여준다. 이것은 FNN 필터의 복잡한 연결구조 때문으로 어떤 형태의 비선형성도 표현이 가능하다.

앞으로 좀더 다양한 잡음환경에서 여러 형태의 비선형 잡음 데이터를 사용하여 성능을 검증하기 위한 후속 연구가 이루어져야 할 것이다.

감사의 글

이 논문은 2022년도 남서울대학교 학술연구비 지원에 의해 연구되었음.

References

- [1] M. R. Sambur, "Adaptive noise canceling for speech signals," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-26, Oct. 1978, pp. 419-423.
- [2] K. Lashkari, "A novel Volterra-Wiener model for equalization of loudspeaker distortions," In *2006 IEEE international conference on acoustics speech and signal processing proceedings*, Toulouse, France, 2006, pp. V.
- [3] X. Guo, Y. Li, J. Jiang, C. Dong, S. Du, and L. Tan, "Sparse modeling of nonlinear secondary path for nonlinear active noise control," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 3, 2018, pp. 482-496.
- [4] L. Tan and J. Jiang, "Adaptive Volterra filters for active control of nonlinear noise processes," *IEEE Trans. Signal Processing*, vol. 49, no. 8, Aug. 2001, pp. 1667-1676.
- [5] R. Napoli and L. Piroddi, "Nonlinear active noise control with NARX models," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 18, no. 2, Feb. 2010, pp. 286-295.
- [6] S. Ghasemi, R. Kamil, and M. H. Marhaban, "Nonlinear THF-FXLMS algorithm for active noise control with loudspeaker nonlinearity," *Asian Journal of Control*, vol. 18, no. 3, 2016, pp. 502-513.
- [7] S. M. Kuo and H.-T. Wu, "Nonlinear adaptive bilinear filters for active noise control systems,"

- IEEE Transactions on Circuits and Systems*, vol. 52, no. 3, 2005, pp. 617-624.
- [8] D. P. Das and G. Panda, "Active mitigation of nonlinear noise processes using a novel filtered-s LMS algorithm," *IEEE Transactions on Speech and Audio Processing*, vol. 12, 2004, pp. 313-322.
- [9] O. J. Tobias and R. Seara, "Leaky-FXLMS algorithm: stochastic analysis for Gaussian data and secondary path modeling error," *IEEE Transactions on Speech and Audio Processing*, vol. 13, no. 6, 2005, pp. 1217-1230.
- [10] N. V. George and G. Panda, "Advances in active noise control: A survey, with emphasis on recent nonlinear techniques," *Signal Processing*, vol. 93, 2013, pp. 363-377.
- [11] S. D. Snyder and N. Tanaka, "Active control of vibration using a neural network," *IEEE Transactions on Neural Networks*, vol. 6, no. 4, 1995, pp. 819-828.
- [12] M. Bouchard, B. Paillard, and C. T. Le Dinh, "Improved training of neural networks for the nonlinear active control of sound and vibration," *IEEE Transactions on Neural Networks*, vol. 10, no. 2, Mar. 1999, pp. 391-401.
- [13] C. Y. Chang and F. B. Luoh, "Enhancement of active noise control using neural-based filtered-x algorithm," *Journal of Sound and Vibration*, vol. 305, no. 1-2, 2007, pp. 348-356.
- [14] T. Krukowicz, "Active noise control algorithm based on a neural network and nonlinear input-output system identification model," *Archives of Acoustics*, vol. 35, no. 2, 2010, pp. 191-202.
- [15] G. Panda and D. P. Das, "Functional link artificial neural network for active control of nonlinear noise processes," In *2003 international workshop on acoustic echo and noise control*, Kyoto, Japan, 2003, pp. 163-166.
- [16] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, 2015, pp. 85-117.
- [17] J. S. Choi, "Speech and Noise Recognition System by Neural Network," *Journal of the Korea Institute of Electronic Communication Sciences*, vol. 5, Aug. 2010, pp. 357-362.
- [18] H. W. Lee, "Optimization of the kernel size in CNN noise attenuator," *Journal of the Korea Institute of Electronic Communication Sciences*, vol. 15, Dec. 2020, pp. 987-994.
- [19] H. W. Lee, "Optimization of the number of filter in CNN noise attenuator," *Journal of the Korea Institute of Electronic Communication Sciences*, vol. 16, Aug. 2021, pp. 625-632.

저자 소개

이행우(Haeng-Woo Lee)

1985년 광운대학교 전자공학과 (공학사)

1987년 서강대학교 대학원 전자공학과 (공학석사)



2001년 전북대학교 대학원 전자공학과 (공학박사)

1987년~1998년 한국전자통신연구원 선임연구원

2001년~현재 남서울대학교 지능정보통신공학과 교수

※ 관심분야 : VLSI 설계, 딥러닝, 음향잡음감쇠