

음향반향제거기에서 기하학적 개념의 BSS를 이용한 동시통화 제어

이행우*

Double-talk Control using Blind Signal Separation based on Geometric Concept in Acoustic Echo Canceller

Haeng-Woo Lee*

요 약

본 논문은 기하학적 개념에 기반한 암묵신호분리를 이용하여 동시통화문제를 제어하는 음향반향제거기에 관한 것이다. 음향반향제거기는 동시통화 구간에서 성능이 저하되거나 발산하게 된다. 따라서 혼합된 마이크 입력신호로부터 근단화자신호를 분리해서 동시통화상태를 검출하기 위하여 암묵신호분리기술을 이용한다. 암묵신호분리는 미지의 입력신호들로부터 기하학적 개념에 기반하여 변형과 회전의 두 단계를 거쳐 근단화자신호를 추정해낸다. 컴퓨터 시뮬레이션을 통하여 이 음향반향제거기의 성능을 검증하였다. 동시통화 구간에서는 반향제거필터의 계수가 발산하는 것을 방지하기 위하여 계수 갱신작업을 중지하도록 하였다. 시뮬레이션 결과, 이 방법을 사용한 음향반향제거기는 암묵신호분리의 빠른 수렴속도로 인해 동시통화의 유무에 상관없이 안전하게 동작함을 확인하였다.

ABSTRACT

This paper describes an acoustic echo canceller with double-talk using BSS(Blind Signal Separation) based on the geometric concept. The acoustic echo canceller may be deteriorated or diverged during the double-talk period. So we use the blind signal separation to detect the double talking by separating the near-end speech signal from the mixed microphone signal. In the closed reverberation environment, the blind signal separation extracts the near-end signal from unknown signals with the transformation and rotation based on the geometric concept. By this method, the acoustic echo canceller operates irrespective of double-talking. We verified performances of the proposed acoustic echo canceller by computer simulations. The results show that the acoustic echo canceller with this algorithm detects the double-talk periods thoroughly, and operates stably in the normal state without diverging of coefficients after ending the double-talking.

키워드

Blind Signal Separation, Acoustic Echo Canceller, Double Talk, Geometric Concept, Hands-Free
암묵 신호 분리, 음향 반향 제거기, 동시 통화, 기하학적 개념, 핸드즈 프리

1. 서 론

음향반향신호는 원단화자로부터의 수신신호가 스피

커를 통해 방사된 후 마이크에서 근단화자의 송신신호와 결합됨으로서 원단에서 근단화자의 음성청취를 방해하게 된다. 반향신호를 제거하기 위하여 적응필터

* 남서울대학교 정보통신공학과(hwlee@nsu.ac.kr)
• 접수일 : 2016. 07. 01
• 수정완료일 : 2017. 05. 31
• 게재확정일 : 2017. 06. 16

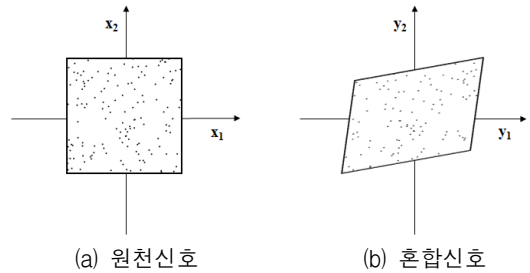
• Received : Jul. 01, 2016, Revised : May. 31, 2017, Accepted : Jun. 16, 2017
• Corresponding Author : Haeng-Woo Lee
Dept. of Information Communication, Namseoul University,
Email : hwlee@nsu.ac.kr

를 이용하게 되는데 이 회로는 원단화자신호만 존재하는 일방통화 시에만 정상적으로 동작하고 근단화자신호도 함께 존재하는 동시통화 구간에서는 반향신호를 제거하는 능력이 갑자기 저하되는 현상이 발생한다. 따라서 이 동시통화문제를 처리할 수 있는 기능이 포함되어야 한다[1-2].

일반적으로 동시통화상태를 검출하여 반향제거의 계수 갱신을 중지시킴으로서 계수의 발산을 방지하는 방법[3]이 있다. 그러나 종래 이 방법은 검출시간이 비교적 길어 검출이 되기 전에 반향제거의 계수가 발산하는 경우가 있다. 따라서 본 논문에서는 신속하게 동시통화상태를 검출하기 위하여 새로운 음원분리 기술을 사용한다. 음원분리기술은 관찰된 혼합음으로부터 각 음원신호를 분리해낼 수 있다. 특히 입력신호의 통계적인 특성 등 사전 정보를 알지 못하는 상태에서 신호들을 분리하는 블라인드 음원분리(BSS) 방법이 있다. 이것은 음성간의 통계적 특성이 독립이라는 가정하에 여러 음원을 분리하는 방법이다[4]. Jutten과 Herault가 제안한 BSS 알고리즘은 선형 혼합된 신호를 간단히 분리할 수 있다고 발표하였으며 이후 여러 후속 연구가 이루어졌다[5]. Jutten과 Herault는 시간영역에서 잡음과 신호를 분리하는 BSS 알고리즘을 제안하였다. 그러나 이 알고리즘은 수천 개의 많은 템플 사용하므로 큰 메모리가 필요하고 계산량도 많아진다. Bell과 Sejnowski는 독립성분분석(ICA)을 통하여 지연을 추정하는 방법을 제안하였다[6-8]. 그러나 음성신호는 주기적인 특성 때문에 지연을 추정하는데 있어 초기값에 따라서는 최적해로 수렴하지 못하는 경우도 있다. 이를 해결하기 위해 Torkko는 정보량 최대화 알고리즘을 이용하여 지연을 추정하였다[9]. 그러나 계산량이 많고 주변잡음이 있으면 정확하게 지연을 추정하지 못한다. 따라서 최근에는 주파수영역에서 블라인드 음원분리를 하는 연구가 많이 이루어지고 있다[10]. 또한 폐쇄된 공간에서 반향신호가 존재하는 다채널 환경에서 신호를 분리하는 문제는 다차원의 많은 계산량을 요구한다[11-12].

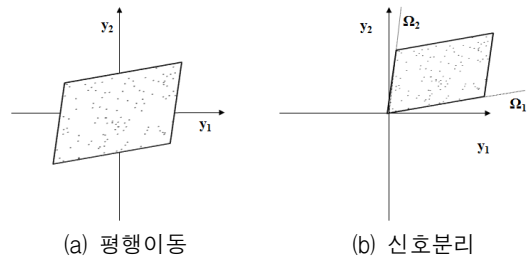
많은 연구결과에서 보는 바와 같이 순시혼합모델은 이해는 쉬우나 구현이 용이하지 않고 연산량이 많다. 그래서 일부 연구자들은 특정 이산신호를 처리할 수 있는 간단한 기하학적 접근방법[13]을 제안하였다. 그

리고 후에는 아날로그신호를 분리하도록 개선하였다[14]. 기본 개념은 기하학적 관점의 독립성을 이용하는 것이다. 실제로 2개의 독립된 신호 $x_1(n)$ 과 $x_2(n)$ 의 산포도(scatter plot)는 직사각형을 이룬다. 이 신호들의 혼합은 직사각형을 평행사변형 모양으로 기하학적 변형시키는 것이다.



(a) 원천신호 (b) 혼합신호
그림 1. 신호 혼합의 기하학적 개념
Fig. 1 Geometric concept of signals mixture

기하학적 개념에 기반한 암묵신호분리는 그림 2와 같이 두 단계의 과정을 거친다. 첫째 평행사변형을 제 1사분면으로 평행 이동시킨다. 둘째 평행사변형의 모서리 Ω_1 과 Ω_2 의 기울기를 구한다. 여기서 신호의 확률밀도함수는 uniform하고 입력 원천신호는 2개만 존재한다고 가정한다.



(a) 평행이동 (b) 신호분리
그림 2. 신호 분리의 기하학적 개념
Fig. 2 Geometric concept of signals separation

한편 음성신호 또는 Gamma 확률밀도함수를 갖는 신호들은 평행사변형의 모서리를 찾기가 어렵다. 따라서 연구자들은 대신 평행사변형의 대칭 기축 Ω_1 과 Ω_2 의 기울기를 구하는 알고리즘을 제안하였다. 이것은 혼합평면을 여러 클러스터로 분할한 후 대칭축을 더 세력이 큰 클러스터의 방향에 대응시킨다. 성능은 축 Ω_1 과 Ω_2 의 기울기의 산출에 좌우된다. 기울기의

산출은 클러스터의 수와 우수한 클러스터를 선정하기 위한 판정의 수에 따라 달라진다. 이러한 기하학적 원리를 이용하여 신호를 분리하고 동시통화상태를 검출하는 방법에 대하여 연구하였다.

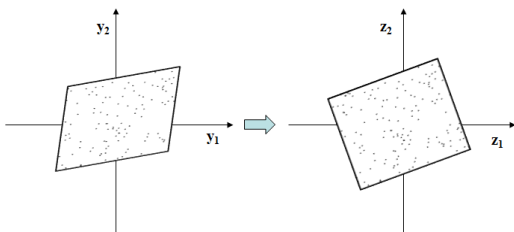
본 논문은 다음과 같이 구성되었다. II장에서 압묵 신호분리 알고리즘의 개념을 논하였고, III장에서는 음향반향제거기에서 압묵신호분리를 통한 동시통화 검출방법에 대하여 살펴보았다. 그리고 모의실험과 그 결과에 대한 검토를 IV장에서 기술하였고, V장에서 결론을 도출하였다.

II. 기하학적 개념에 기반한 압묵신호분리

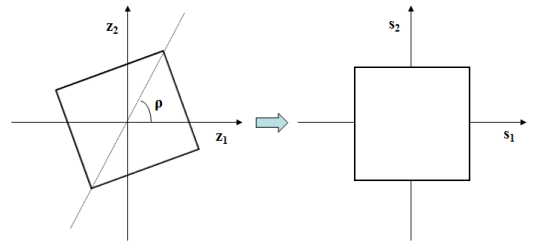
원천신호의 수는 혼합신호의 공분산 행렬에 있어서 크기가 큰 고유치의 수로 정해질 수 있다. 여기서 우리는 소스와 센서의 수가 동일하고 순시혼합인 경우라고 가정한다. 이러한 순시혼합의 분리는 조합과 스케일 인자를 통해 이루어질 수 있다. 즉 산출된 원천신호는 조합과 스케일 인자로 표현된 원천신호와 같다. 따라서 \mathbf{W} 가 분리행렬이고 \mathbf{M} 이 혼합행렬일 때 전역행렬 $\mathbf{G}(=\mathbf{WM})$ 은 다음 관계를 만족해야 한다 [13].

$$\mathbf{G} = \mathbf{PD} \quad (1)$$

여기서 \mathbf{P} 는 어떤 조합, \mathbf{D} 는 모든 계수를 포함하는 대각행렬을 나타낸다. 순시혼합은 관찰공간에서의 기하학적 변형을 가져온다. 신호분리는 변형(transformation)과 회전(rotation)의 2단계 과정을 통하여 이루어질 수 있다. 먼저 두 개의 신호가 혼합된다고 가정한다.



(a) 변형



(b) 회전

그림 3. 신호 분리의 2단계 과정
Fig. 3 Two steps for signals separation

Cholesky 분해를 이용하여 혼합신호 \mathbf{Y} 로부터 직교신호 \mathbf{Z} 들을 생성할 수 있다. 즉 혼합 평행사변형을 정사각형으로 변형한다. \mathbf{R}_Y 가 $\mathbf{Y}(t)$ 의 공분산 행렬이라 하면

$$\mathbf{R}_Y = \mathbf{E} \mathbf{Y} \mathbf{Y}^T \quad (2)$$

Cholesky 분해를 이용하여 \mathbf{R}_Y 의 제곱근 \mathbf{L} 을 구할 수 있다.

$$\mathbf{R}_Y = \mathbf{L} \mathbf{L}^T \quad (3)$$

원천신호들이 통계적으로 독립이고 원천신호의 수가 센서의 수와 같다면 혼합 공분산 행렬 \mathbf{R}_Y 는 그 제곱근과 마찬가지로 모든 계수를 포함하는 행렬이 된다. 따라서 직교 혼합신호 \mathbf{Z} 는 다음과 같이 구할 수 있다.

$$\mathbf{Z}(t) = \mathbf{L}^{-1} \mathbf{Y}(t) \quad (4)$$

그리고 공분산 행렬 \mathbf{R}_Z 는 고유행렬과 같게 된다.

$$\mathbf{R}_Z = \mathbf{E} \mathbf{Z} \mathbf{Z}^T = \mathbf{L}^{-1} \mathbf{E} \mathbf{Y} \mathbf{Y}^T \mathbf{L}^{-T} = \mathbf{L}^{-1} \mathbf{R}_Y \mathbf{L}^{-T} = \mathbf{I}_p \quad (5)$$

또한 산출신호 $\mathbf{S}(t)$ 는 각 θ 의 회전을 통하여 $\mathbf{Z}(t)$ 로부터 얻어질 수 있다.

$$\mathbf{S}(t) = \mathbf{R}(\theta) \mathbf{Z}(t) \quad (6)$$

여기서 $\mathbf{R}(\theta)$ 는 삼각함수 행렬이다. 1차 대각선과 수평축 사이의 각을 ρ 라고 하자. 이 각은 원점으로부터 가장 먼 거리에 있는 점을 이용하여 결정되는데 unique하게 정해지지 않으며 이것은 주로 조합의 불확정성에 기인한다. 모든 물리적 신호는 한정적이고 스케일 인자의 불확정성을 이용한다면 소스는 최대 크기가 같거나 소스의 산포도가 정방형이라고 생각할 수 있다. 따라서 신호들을 분리하기 위해서는 다음과 같은 각도로 신호들을 회전시켜야 한다.

$$\theta_V = \frac{\pi}{4} - \rho \tag{7}$$

이 각은 원천신호의 전력밀도함수(pdf)가 균일할 때 계산될 수 있다. 이 신호들이 대칭 감마 전력밀도 함수나 음성신호와 같이 대칭 감마에 근접한 전력밀도 함수처럼 unimodal 전력밀도 함수를 갖는 경우, 이 신호들의 산포도는 그림 4와 같이 직사각형 보다는 십자가 형태가 될 것이다.

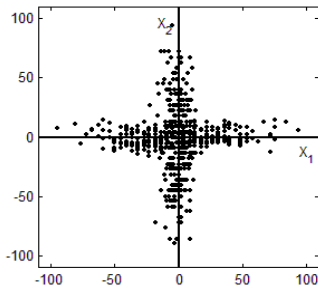


그림 4. 분리된 음성신호의 산포도
Fig. 4 Scatter plot of separated speech signals

이때 가장 먼 거리에 있는 점은 두 직교축 중 하나에 대응된다. 따라서 각은 다음과 같이 구해진다.

$$\theta_G = -\rho \tag{8}$$

III. 동시통화 음향반향제거기

폐쇄된 공간에서 핸드프리 통화는 음향반향을 발생시키며, 이 원하지 않는 반향신호를 제거하기 위하여

적용필터를 사용한다. 그런데 원단화자와 근단화자가 동시에 말을 하는 동시통화의 경우에는 마이크로 유입되는 근단화자신호로 인하여 반향제거기의 계수가 발산할 수 있다. 이 문제를 해결하기 위하여 기하학적 개념에 기반한 BSS 방법을 이용한다. 마이크로 유입되는 반향신호와 근단화자신호는 암묵신호분리를 통하여 근단화자신호를 얻을 수 있다. 그래서 BSS에서 동시통화구간을 검출하고 이 구간에는 필터계수의 갱신을 보류하도록 한다. 그림 5에서 BSS를 이용한 음향반향제거기의 구조를 보여주고 있다.

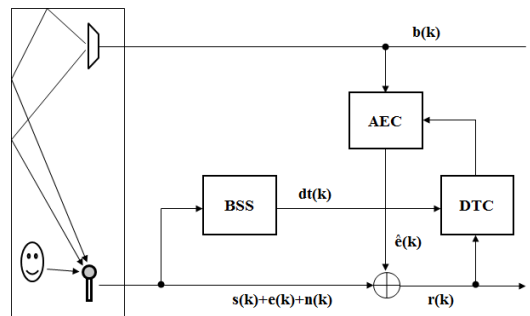


그림 5. BSS를 이용한 음향반향제거기
Fig. 5 Acoustic echo canceller with BSS

동시통화구간에서 음향반향제거기의 계수는 잔차신호에 포함된 근단신호로 인하여 수렴상태에서 이탈하여 발산하고 수렴상태로 복귀하지 못할 수도 있다. 따라서 근단화자 및 반향신호 등이 포함된 마이크 입력신호에 대하여 BSS를 사용하여 근단화자신호와 반향신호를 분리하고, 검출된 근단화자신호의 유무에 따라 AEC의 잔차신호를 이용하여 적응필터의 계수를 갱신할지 여부를 정한다. 즉 동시통화상태에서는 필터의 계수가 발산하는 것을 방지하기 위하여 계수의 갱신 작업을 일시 중지하도록 한다. 이때 암묵신호분리작업이 늦어지면 AEC 계수들이 발산할 수 있으므로 근단화자신호가 나타날 때 신속하게 BSS가 수렴해야 한다.

스피커에서 발생한 음향반향신호는 다채널의 반향 경로를 거쳐 마이크로 유입되는데 이 마이크 신호는 BSS의 입력신호로도 사용된다. 반향신호를 제거하기 위하여 NLMS 알고리즘을 이용하며 이와 관련된 수식은 다음과 같다[3].

$$\hat{e}(k) = \sum_{n=0}^{N-1} c_n(k)b(k-n) \quad (9)$$

$$r(k) = s(k) + e(k) - \hat{e}(k) \quad (10)$$

$$c_n(k+1) = c_n(k) + 2\mu \frac{r(k)}{\sigma_e^2} b(k-n) \quad (11)$$

여기서 $c(k)$ 는 반향제거기의 계수이며, μ 는 알고리즘의 수렴속도를 결정하는 상수이다.

IV. 모의실험 결과

기하학적 개념에 기반한 음원분리 및 반향제거기의 성능을 상위수준에서 검증하기 위해 C 언어를 사용하여 시스템의 동작을 표현하는 시뮬레이션 프로그램을 작성하였다. 앞에서 설명한 알고리즘을 적용하여 반복된 심볼 수에 대한 ERLE의 특성곡선을 관찰하였다.

입력은 근단화자의 음성 녹음신호와 원단의 음성신호 등 두 음원으로부터 마이크에 유입된 혼합신호를 사용하였다. 원단으로부터의 수신신호는 400Hz~4kHz의 음성주파수 사이에서 무색 성분을 갖는 신호이다. 마이크 입력신호는 8kHz로 샘플링되며, 사용된 반향제거기는 256-tap FIR 구조이다. 적응 알고리즘으로 NLMS 알고리즘을 사용하였으며, 적응스텝의 크기는 $\mu=0.005$ 로 설정하였다. 반향경로의 응답특성은 시간이 지남에 따라 크기가 지수함수적으로 감소하고 끝부분에서는 상당히 긴 꼬리가 지속된다. 이 실험에서 반향경로는 다음과 같은 수학적 모델을 사용하였다.

$$h(i) = 0.0914 \text{rand}(i) * (-1.0346)^{-i} \quad (12)$$

여기서 $\text{rand}(i)$ 는 정수에 따라 불규칙하게 (+1) 또는 (-1)을 갖는 함수이다.

본 기술은 두 음향신호가 혼합되어 유입될 때 둘 중 어느 하나의 신호를 분리해낸다. 먼저 근단화자의 음성신호와 원단의 음향반향신호가 혼합되어 유입될 때 각 신호의 분리 실험을 수행하였다. 사용된 근단화자의 음성신호는 「실제 윈도우즈 프로그램 내에서」 라는 여성의 녹음된 음성이다. 그림 6은 두 음원신호의 파형을 보여주고 있다.

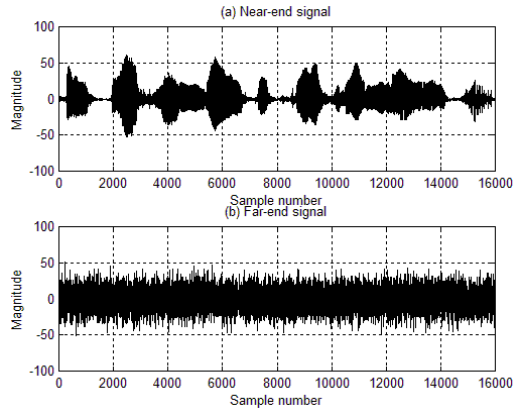


그림 6. 근단신호와 원단신호의 음원파형
Fig. 6 Waveforms of near-end and far-end signal

그리고 그림 7은 음향 반향된 원단신호에 근단 음성신호가 혼합된 마이크 입력신호의 파형을 보여주고 있다. 혼합된 음성은 근단화자신호 중에서 5,000 - 11,500 번째 샘플이 삽입된 형태이다.

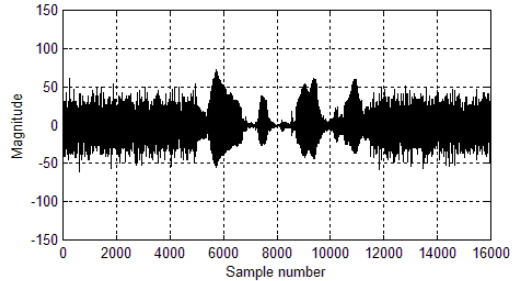


그림 7. 마이크 입력신호
Fig. 7 Microphone input signal

이어서 두 신호가 혼합되어 BSS에 유입될 때 각 신호의 분리 결과에 대하여 살펴본다. 두 음원인 음향 반향신호와 음성신호를 분리한 경우의 각 출력신호 파형을 그림 8에서 보여주고 있다. 그림 (a)에서 초기 계수가 수렴하지 않았을 때는 이상 파형이 나타났으나 곧 수렴하여 잘 분리된 음성신호를 보여준다.

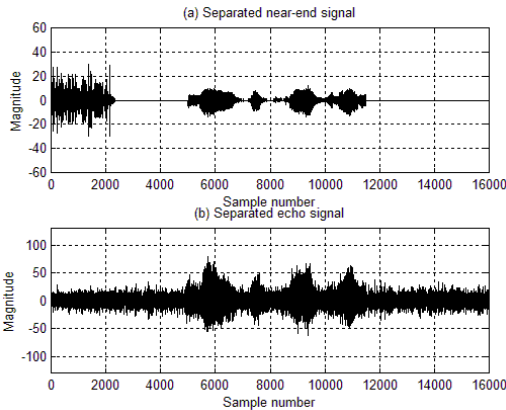


그림 8. 분리된 근단신호 및 원단신호
Fig. 8 Separated near-end and far-end signal

그리고 그림 9는 마이크 신호에서 반향추정신호를 제거한 잔차신호와 암묵신호분리의 출력인 동시통화 구간을 보여준다. 이 그림으로부터 잔차신호는 동시통화 구간에서 근단화자신호를 포함하고 있는 것을 볼 수 있다. 동시통화는 high 값일 때 발생하며 이 구간에서는 AEC 계수의 갱신을 중지하도록 한다.

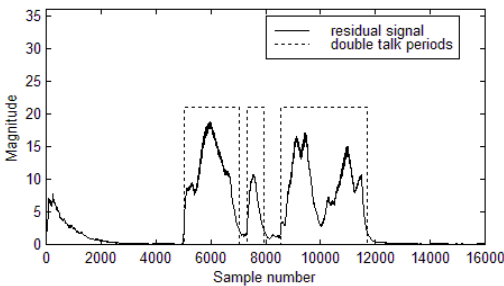


그림 9. 잔차신호와 동시통화 구간
Fig. 9 Residual signal and double talk period

끝으로 그림 10은 음향반향제거기의 성능을 나타내는 ERLE 곡선을 보여준다. 이 그림으로부터 동시통화 구간에서는 근단화자신호 때문에 크기가 작아짐을 볼 수 있고, 동시통화가 종료된 직후에도 ERLE 값이 높은 크기를 유지하고 있음을 알 수 있다.

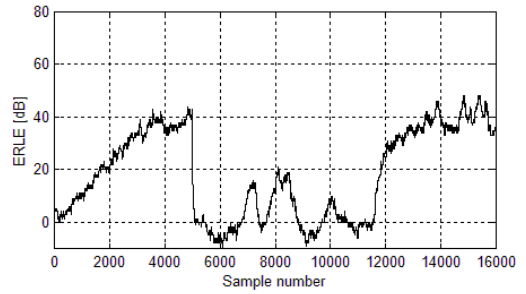


그림 10. 제안한 음향반향제거기의 ERLE 곡선
Fig. 10 ERLE curves of the proposed AEC

V. 결 론

본 논문은 음향반향제거기에서 기하학적 개념에 기반한 신호분리기술을 사용하여 혼합된 신호로부터 반향신호와 근단화자신호를 분리하고, 분리된 근단화자신호를 이용하여 동시통화 구간에는 반향제거기의 계수 갱신을 일시 중지함으로써 동시통화상태에서도 반향제거기가 정상적으로 동작하는 방법에 관한 것이다. 폐쇄된 반향환경에서 암묵신호분리의 혼합모델은 다 채널이지만 근단화자신호는 단일채널로 볼 수 있기 때문에 BSS 알고리즘을 이용하여 두 입력신호 성분을 분리할 수 있다. 모의실험 결과, 제안된 알고리즘은 암묵신호분리과정을 통해 근단화자신호와 반향신호를 분리하고, 동시통화 시에도 발산하지 않고 안정적으로 동작함을 보였다. 향후 실제 반향경로에서도 동일한 성능을 유지할 수 있도록 후속연구를 수행할 예정이다.

감사의 글

이 논문은 2017년도 남서울대학교 학술연구비 지원에 의해 연구되었음.

References

- [1] S. Minami and T. Kawasaki, "A Double Talk Detection Method for an Echo Canceller," *ICC(International Conference on Communications)* '85, Chicago, USA, Jun. 1985, pp.1492-1497.

- [2] H. Ye and B. Wu, "A New Double-Talk Detection Algorithm Based on the Orthogonality Theorem," *IEEE Trans. on Comm.*, vol.39, no.11, Nov. 1991, pp.1542-1545.
- [3] W. Hsu, F. Chui, and D. Hodges, "An Acoustic Echo Canceler," *IEEE J. of Solid-state Circuits*, vol.24, no.6, Dec. 1989, pp.1639-1646.
- [4] A. Bell and T. Sejnowski, "An Information Maximization Approach to Blind Separation and Blind Deconvolution," *Neural Computation*, vol.7, no.6, 1995, pp.1129-1159.
- [5] K. Torkkola, "Blind Separation of Convolved Sources Based on Information Maximization," *Neural Networks for Signal Processing, IEEE Signal Processing Society Workshop*, Corfu, Greece, Jun. 1996.
- [6] H. Saruwatari, T. Kawamura, and K. Shikano, "Blind Source Separation for Speech Based on Fast Convergence Algorithm with ICA and Beamforming," *Proceedings of the Eurospeech 2001*, Aalborg, Denmark, Sep. 2001, pp.2603-2606.
- [7] P. Comon, "Independent component analysis, A new concept," *Signal Processing*, vol.36, no.3, 1994, pp.287-314.
- [8] M. Kawamoto, K. Matsuoka, and N. Ohnishi, "A method of blind separation for convolved non-stationary signals," *Neuro Computing*, vol.22, no.1-3, 1998, pp.157-171.
- [9] L. Parra and C. Spence, "Convolutive blind separation of non-stationary sources," *IEEE Trans. Speech Audio Process.*, vol.8, no.3, 2000, pp.320-327.
- [10] D. Schobben and P. Sommen, "A frequency domain blind signal separation method based on decorrelation," *IEEE Trans. Signal Process.*, vol.50, no.8, 2002, pp.1855-1865.
- [11] E. Weinstein, M. Feder, and A. Oppenheim, "Multi-channel signal separation by decorrelation," *IEEE Trans. Speech Audio Process.*, vol.1, no.4, 1993, pp.405-413.
- [12] D. Yellin and E. Weinstein, "Multichannel signal separation: methods and analysis," *IEEE Trans. Signal Process.*, vol.44, no.1, 1996, pp.106-118.
- [13] C. Puntinet, A. Prieto, C. Jutten, M. Alvarez, and J. Ortega, "Separation of sources : A geometry-based procedure for reconstruction of n-valued signals," *Signal Processing*, vol.46, no.3, 1995, pp.267-284.
- [14] C. Guntonet and A. Prieto, "Neural net approach for blind separation of sources based on geometric properties," *Neuro Computing*, vol.18, no.3, 1998, pp.141- 164.

저자 소개

이행우(Haeng-Woo Lee)

1985년 광운대학교 전자공학과 졸업(공학사)

1987년 서강대학교 대학원 전자공학과 졸업(공학석사)

2001년 전북대학교 대학원 전자공학과 졸업(공학박사)

2001년 남서울대학교 정보통신공학과 교수

1987년~1998년 한국전자통신연구원 선임연구원

※관심분야 : VLSI 설계, 음향신호처리, 적응필터

