

IVA 기반의 2채널 암묵적신호분리에서 주파수빈 뒤섞임 문제 해결을 위한 후처리 과정

Post-Processing of IVA-Based 2-Channel Blind Source Separation for Solving the Frequency Bin Permutation Problem

추쯔하오¹⁾ · 배 건 성²⁾

Chu, Zhihao · Bae, Keunsung

ABSTRACT

The IVA(Independent Vector Analysis) is a well-known FD-ICA method used to solve the frequency permutation problem. It generally works quite well for blind source separation problems, but still needs some improvements in the frequency bin permutation problem. This paper proposes a post-processing method which can improve the source separation performance with the IVA by fixing the remaining frequency permutation problem. The proposed method makes use of the correlation coefficient of power ratio between frequency bins for separated signals with the IVA-based 2-channel source separation. Experimental results verified that the proposed method could fix the remaining frequency permutation problem in the IVA and improve the speech quality of the separated signals.

Keywords: blind source separation, independent vector analysis, frequency permutation

1. 서론

ICA(Independent Component Analysis)는 확률통계모델에 기반한 암묵적신호분리 기법으로, 다채널 혼합신호로부터 독립적인 원 신호를 찾는 방법이다[1]. 곱의 형태로 표현되는 ICA의 순시혼합모델(instantaneous mixture model)에 컨볼루션 형태로 모델링 되는 혼합신호를 적용하기 위해서는 일반적으로 컨볼루션이 곱의 형태로 표현되는 주파수영역에서의 ICA, 즉, FD-ICA(Frequency Domain ICA)를 사용하게 된다. FD-ICA[2]는 DFT(Discrete Fourier Transform)를 이용해서 얻은 각 주파수 빈(bin)에 대해 ICA의 기본모델을 적용하게 되는데, 이때

각 주파수 빈에서 분리된 출력신호의 순서가 일정하지 않은 ICA의 모호성(ambiguity)으로 인해 분리된 신호에 다른 신호의 주파수 빈이 뒤섞이는 문제가 발생하여 원래의 신호를 제대로 분리해 낼 수 없게 된다.

IVA(Independent Vector Analysis)[3] 기법은 주파수 빈의 뒤섞임 문제를 해결하기 위해 제안된 FD-ICA 방법 중의 하나로, FD-ICA를 수행할 때 각 주파수 빈 별로 독립적으로 ICA를 수행하는 대신에 각 신호의 주파수 빈 전체를 하나의 벡터로 간주하여 주파수 빈 상호간의 의존성을 고려하면서 ICA를 수행함으로써 FD-ICA에서 필연적으로 수반되는 주파수 빈의 뒤섞임 문제를 해결하고자 한 방법이다. IVA는 KLD(Kullback Leibler divergence)를 기반으로 주파수 빈 사이의 의존성을 고려하면서 ICA 모델의 분리행렬(unmixing matrix)을 구함으로써 컨볼루션 형태의 혼합신호에 대해 대체적으로 좋은 분리 성능을 보이지만, 때때로 임의의 주파수 대역에서 또는 음성 신호의 에너지가 상대적으로 적은 고주파의 넓은 대역에서 신호가 뒤섞여서 분리된 신호의 음질이 저하되는 현상이 관찰되었다[4]. 즉, IVA를 이용한 신호 분리의 경우에도 이러한 주파

1) 경북대학교, chuzhihao1016@yeah.net

2) 경북대학교, ksbae@ee.knu.ac.kr, 교신저자

이 논문은 2012[2013,2014]학년도 경북대학교 학술연구비에 의하여 연구되었음

접수일자: 2013년 11월 2일

수정일자: 2013년 11월 30일

게재결정: 2013년 12월 15일

수 빈의 뒤섞임 문제에 대한 여지가 남아 있다.

FD-ICA에서 주파수 빈의 뒤섞임 문제를 해결하기 위해 다양한 방법들이 연구되어 왔는데, 최근에 주파수 영역에서 ICA로 분리된 신호의 각 주파수 빈에 대한 전력비의 상관계수를 이용하여 동일 채널 신호의 주파수 빈 의존성을 판정하는 방법을 제안하고 실험을 통해 주파수 빈의 뒤섞임 문제를 해결할 수 있음을 보였다[5]. 그러나 위 연구에서는 복소수 형태로 주어지는 스펙트럼 값에 어떤 ICA 기법을 적용했는지가 분명하지 않으며, 가장 대표적인 fastICA를 이용할 경우에도 주파수영역에서 주어지는 복소수 변수의 적절한 처리 방법에 대한 연구가 필요하다.

본 연구에서는 대체적으로 우수한 신호 분리 성능을 보이는 FD-ICA 기법중의 하나인 IVA에서 때때로 발생하는 주파수 빈의 뒤섞임 문제를 해결하여 신호 분리 성능을 향상시킬 수 있는 방법을 제안하였다. 제안한 방법은 IVA로 분리된 신호의 각 주파수 빈에 대한 전력비의 상관계수를 이용한 후처리 과정으로, 분리가 잘못된 주파수 빈을 검출하고 뒤섞임 문제를 해결할 수 있는 알고리즘을 제안하였으며 실험을 통해 제안한 방법의 타당성을 검증하였다. 본 논문의 구성은 다음과 같다. 2장에서는 주파수 빈의 전력비의 상관도를 이용하여 본 연구에서 제안한 IVA 기법의 후처리 과정을 설명하고, 3장에서는 컴퓨터 모의실험 조건 및 과정, 실험결과를 제시하며, 마지막으로 4장에서 결론을 맺는다.

2. 전력비의 상관도를 이용한 IVA 후처리 과정

FD-ICA의 하나인 IVA의 기본 모델을 행렬식의 형태로 나타내면 식 (1) 및 (2)와 같다.

$$X(n, f) = A(f)S(n, f) \quad (1)$$

$$Y(n, f) = W(f)X(n, f) \quad (2)$$

여기서 $S(n, f)$ 는 원 신호의 스펙트럼, $X(n, f)$ 는 혼합신호의 스펙트럼, $Y(n, f)$ 는 분리된 신호의 스펙트럼, 즉, $S(n, f)$ 의 추정된 값이며, $A(f)$ 는 주파수영역에서 복소수 형태로 표현되는 혼합행렬, $W(f)$ 는 분리행렬, 즉, 추정된 $A(f)$ 의 역행렬이며, n 은 시간 축에서의 프레임 인덱스, f 는 DFT로 얻어지는 각 주파수 빈을 나타낸다. IVA를 수행하여 분리된 i 번째 채널 신호의 각 주파수 빈에 대한 전력은 $\|Y_i(n, f)\|^2$ 로 표시할 수 있으며, 분리된 채널의 모든 신호 전력에 대한 각 주파수 빈의 전력비는 식 (3)과 같이 정의된다.

$$P_i(n, f) = \frac{\|Y_i(n, f)\|^2}{\sum_{j=1}^N \|Y_j(n, f)\|^2} \quad (3)$$

음원이 2개이고 마이크가 2개인 경우에, 각 주파수 빈에 대한 전력비는 식 (4)와 같이 표현된다.

$$P_1(n, f) = \frac{\|Y_1(n, f)\|^2}{\|Y_1(n, f)\|^2 + \|Y_2(n, f)\|^2} \quad (4.a)$$

$$P_2(n, f) = \frac{\|Y_2(n, f)\|^2}{\|Y_1(n, f)\|^2 + \|Y_2(n, f)\|^2} \quad (4.b)$$

분리된 i 번째 채널 신호에서, 시간 축에서의 모든 프레임에 대한 각 주파수 빈 사이의 상관도를 구하기 위해 식 (5)와 같은 행렬을 구하고, $C_i(f_m, f_k)$ 의 공분산 행렬을 이용하여 각 주파수 빈 사이의 상관계수(correlation coefficient) 값을 식 (6)으로 구한다.

$$C_i(f_m, f_k) = P_i(n, f_m)^T P_i(n, f_k) \quad (5)$$

$$\rho_{m,k} = \frac{\text{cov}[C_i(f_m, f_k)]}{\sqrt{C_i(f_m, f_m)} \sqrt{C_i(f_k, f_k)}} \quad (6)$$

그림 1은 잡음음성을 512-point DFT의 IVA를 이용하여 분리된 신호의 스펙트로그램을 보인 것이다. 그림에서 보면 분리된 음성신호와 잡음신호가 3.3 kHz 근처부터 고주파대역이 서로 뒤바뀌었음을 알 수 있다. 그림 2는 분리된 음성신호에 대해 식 (6)에서 구한 상관계수 행렬 값을 나타낸 것인데, 상관계수는 -1에서 1 사이의 값을 가지므로 어두운 부분은 상관계수 값이 음의 값을 갖는 영역에 해당된다. 그림 2에서, 주파수 빈 110 근처를 경계로 하여 주파수 빈 사이의 상관도가 저주파영역과 고주파영역에서 확연히 다른 모습을 볼 수 있다. 다양한 실험을 통해 주파수 빈의 뒤섞임이 발생하는 경우에는 그림 3(a)에 예시한 것과 같이 상관계수 행렬의 대각선을 중심으로 인접한 주파수 빈 사이에서 부호의 변화가 생기는 현상을 발견하였다. 이것은 동일한 음원의 신호일 경우 전체 신호의 상관도가 높고 주파수 빈 사이에 같은 방향의 상관성을 갖는데 비해 다른 독립적인 특성을 갖는 음원과는 반대의 상관성을 보이는 것이라 생각된다. 따라서 이웃한 주파수 빈 f_m 과 f_{m+1} 의 상관계수 값이 양(+)이면 두 주파수 빈은 같은 음원의 신호이고, 만약 음(-)의 값이면 이웃한 두 주파수 빈은 서로 다른 음원의 신호에 해당되어 뒤섞임 문제가 발생했음을 의미하게 된다. 따라서 그림 3(a)와 같이 이웃한 주파수 빈 사이의 상관계수 값의 부호를 조사하여 그림 3(b)에서와 같이 스펙트럼이 뒤바뀐 주파수 빈을 찾아내고, 그에 해당하는 IVA 분리행렬의 행을 바꾸어줌으로써 주파수 빈의 뒤섞임 문제를 해결할 수 있다.

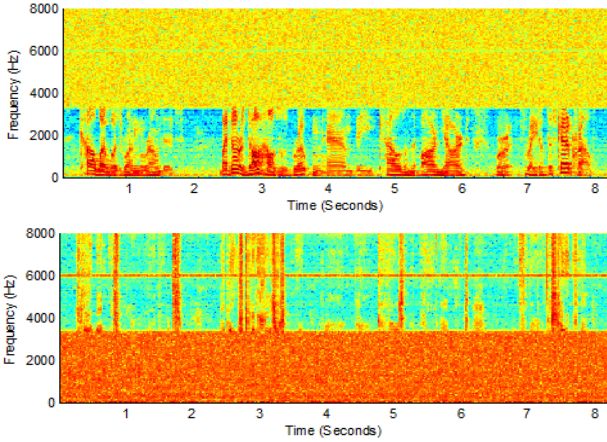


그림 1. IVA를 이용하여 분리한 신호의 스펙트로그램
Figure 1. Spectrograms of separated signals by IVA

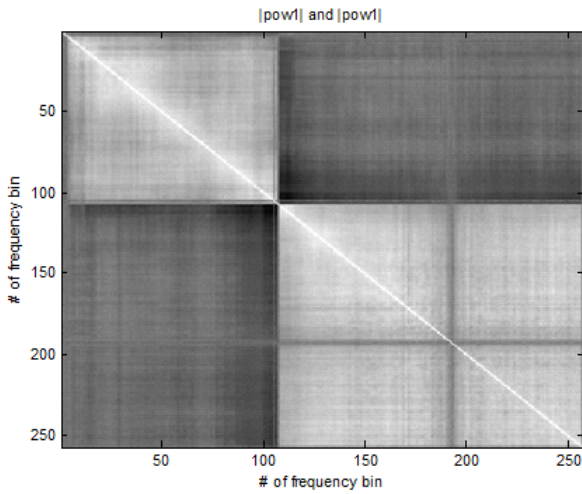
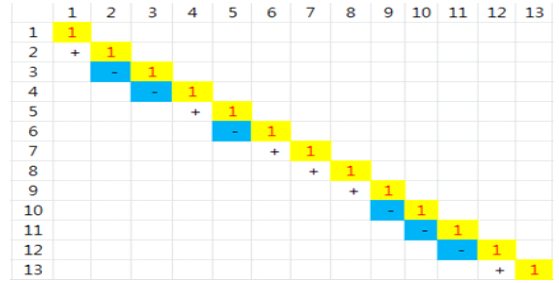


그림 2. 분리된 음성신호의 상관계수
Figure 2. Correlation coefficients of separated signals

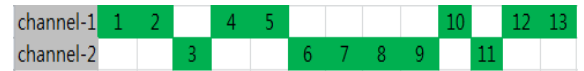
IVA를 이용한 2채널 신호분리에서, 주파수 빈의 뒤섞임 문제를 해결하기 위해 본 연구에서 제안한 방법을 요약하면 다음과 같다.

- 1) IVA를 이용하여 분리된 2채널의 신호 $Y_1(n, f)$ 및 $Y_2(n, f)$ 를 구한다.
- 2) 식 (4) ~ (6)을 이용하여 분리된 임의의 신호에서 주파수 빈의 전력비를 이용한 상관계수 $\rho_{m,k}$ 로 구성되는 행렬을 구한다.
- 3) $k=1$ 부터 마지막 주파수 빈 $k=M-1$ 까지 이웃한 주파수 빈의 상관계수 $\rho_{k+1,k}$ 의 부호를 조사하면서 아래 과정을 수행한다.
 - i) $k=1$.
 - ii) If $\rho_{k+1,k} < 0$, then $k_e = k+1$, and go to step iii).
Otherwise $k = k+1$, and go to step ii).
 - iii) $k = k+1$.

iv) If $\rho_{k+1,k} < 0$, then $k_e = k$, and $k = k+1$.



(a)



(b)

그림 3. 전력비의 상관계수 행렬과 주파수 빈 뒤섞임 관계
Figure 3. Relationship between correlation coefficient matrix of power ratio and frequency bin permutation

Change the spectral bins from $k = k_b$ to $k = k_e$ between two separated signals.
Otherwise $k = k+1$, and go to step iv).
v) if $k < M$, then go to step ii).

3. 실험 및 검토

제안한 방법의 신호분리 및 음질개선 성능을 평가하기 위하여 <그림 4>와 같은 공간을 가정하고 image method[6]를 이용하여 다양한 위치에서의 2채널 혼합신호를 생성하였다. 음원신호는 샘플링주파수 16 kHz인 남자 음성신호 m1, m2, m3와 여자 음성신호 f1, f2, f3, 총 6개 음원 파일을 <그림 5>와 같이 임의로 2개를 선택하여 9개의 조합을 만들고, 음원의 위치는 두 화자가 동일 방향에 있는 경우도 포함하여 <그림 4>의 AL, BD, DF, DE, GL, GH, BC에 해당하는 7가지 조합으로 배치하여 총 63개 경우의 2채널 혼합신호를 생성하였다. <표 1>에 실험 조건을 정리하였다.

표 1. 실험 조건

Table 1. Experimental conditions

Source signals	3 male speech and 3 female speech
Sampling rate	16KHz
Microphone distance	6cm
Microphone and source height	1.5m
Source location	AI, BD, DF, DE, GI, GH, BC
Simulation room size	Length 6m, width 4m, height 3.5m

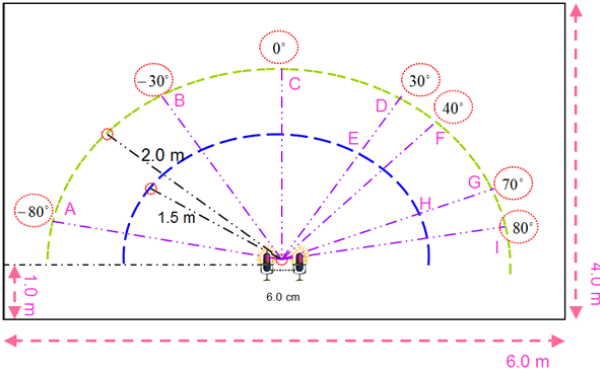


그림 4. 혼합신호 생성을 위한 시뮬레이션 환경
Figure 4. Simulation environment for making mixed signals

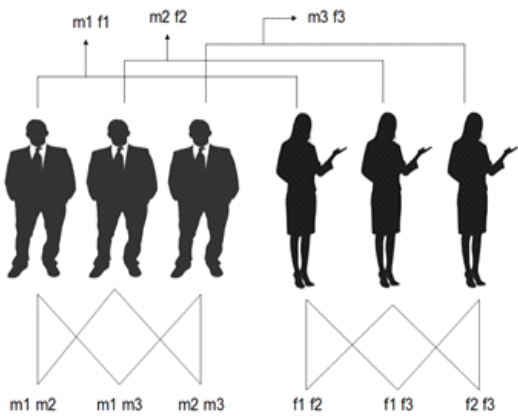


그림 5. 혼합신호 생성을 위한 2채널 음원 조합
Figure 5. 2-channel source combinations for making mixed signals

<그림 6>은 m1, f1의 두 음원신호와, 이를 DF 위치에 배치했을 때의 혼합신호, IVA로 분리된 신호, 제안한 방법으로 후처리를 한 신호의 스펙트로그램을 보인 것이다. <그림 6(c)>의 IVA 결과를 보면 2700Hz 부근부터 고주파 대역까지 스펙트로그램 6(a)의 원 신호와 비교했을 때 서로 뒤바뀌어 있음을 볼 수 있다. 그러나 제안한 방법으로 후처리 과정을 수행할 경우 <그림 6(d)>에서 보여주는 바와 같이 스펙트럼 변화가 원 신호와 비슷함을 알 수 있다. 청취 테스트에서도 6(c)의 경우에는 분리된 신호에서 다른 원 신호가 많이 남아있는 반면에 6(d)에서는 깨끗하게 분리되었음을 확인할 수 있었다.

본 연구에서는 혼합신호의 신호에 주안점을 두고 있으므로 음질개선 성능평가를 위한 척도로 객관적인 음질평가 척도인 PESQ를 사용하여 분리한 음성신호의 음질을 평가하였다[7].

PESQ는 깨끗한 원 음성을 기준으로 한 MOS(Mean Opinion Score) 값을 <표 2>에 주어진 것과 같이 표현하는데, ITU-T P.862 권고안으로 표준화된 객관적인 음질평가 방법으로 주관적인 평가결과와 상관도가 아주 높은 것으로 알려져 있으며 공개된 자동음질 평가 프로그램을 이용할 수 있다.

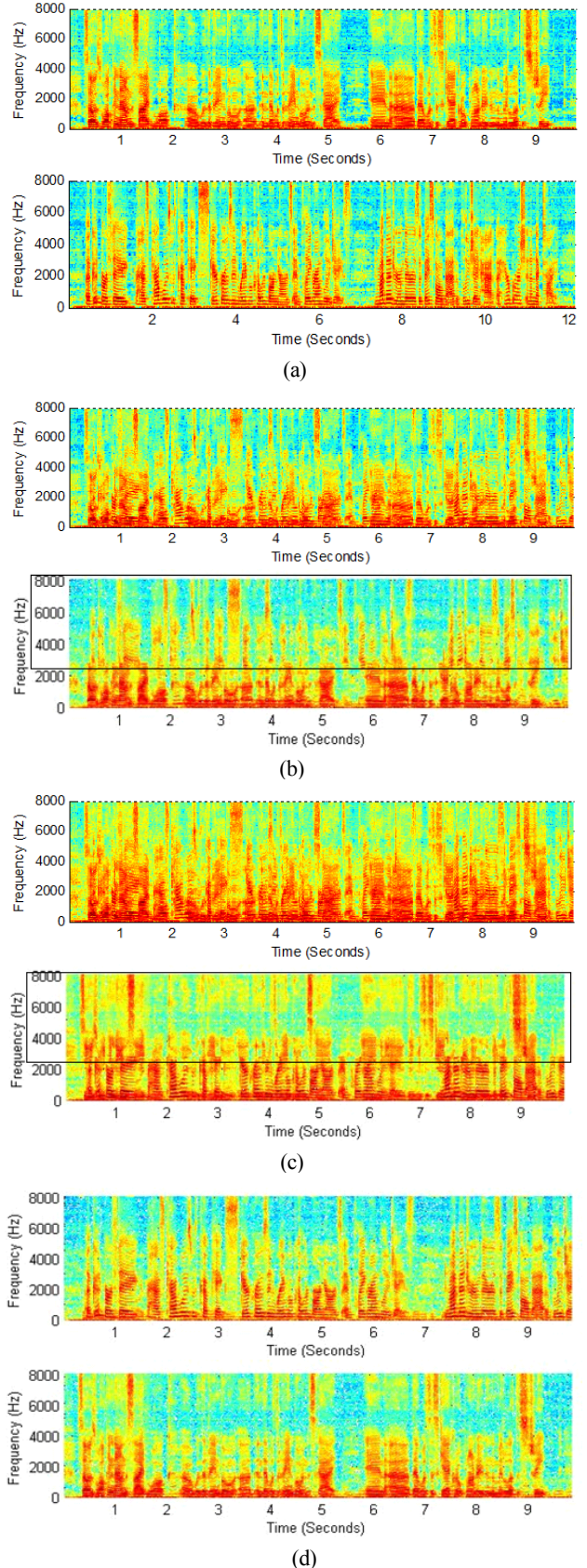


그림 6. 실험결과 얻어진 신호의 스펙트로그램 예
(a) 원 음원신호 (b) 혼합신호 (c) IVA 출력신호
(d) 제안한 방법을 적용하여 얻은 신호

Figure 6. Sample spectrogram obtained from experimental results. (a) Original source signal (b) mixed signal (c) IVA output signal (d) acquired signal by the proposed method

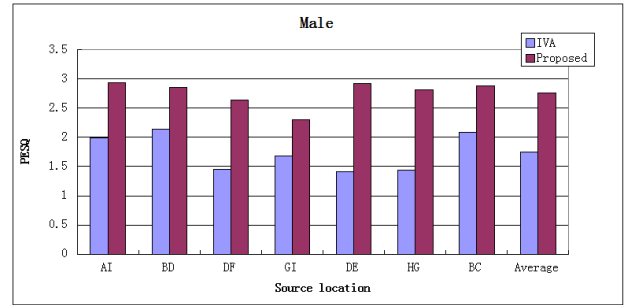
<그림 7>은 IVA를 이용한 신호분리 결과와 본 연구에서 제안한 방법을 적용하여 얻은 신호에 대해 얻은 PESQ 결과를 보인 것이다. MOS 값이 남녀 화자 구분 없이 1정도 향상된 것을 볼 수 있으며, 음원 위치에 상관없이 제안한 방법을 적용하였을 때 음질이 뚜렷하게 많이 개선되었음을 알 수 있다. 이것은 암묵적신호분리에서 좋은 성능을 보이는 기존의 IVA 기법에서도 여전히 주파수 뒤섞임 문제가 남아있을 수 있음을 의미하며, 실제로 본 실험에 사용한 63개 경우의 2채널 혼합 신호에서 IVA를 이용하여 신호 분리를 수행한 결과 그 중 19개 경우에서 주파수 빈 뒤섞임 문제가 발생하였음을 확인하였다.

표 2. PESQ로 표현되는 MOS 값
Table 2. MOS values in PESQ

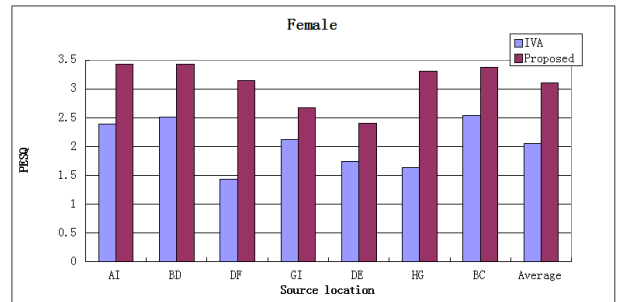
MOS	Quality	Impairment
5	Excellent	Imperceptible
4	Good	Perceptible but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

4. 결론

IVA는 주파수영역에서 ICA를 적용할 때 필연적으로 발생하는 주파수 빈의 뒤섞임 문제를 해결한 FD-ICA 기법중의 하나이지만 때때로 넓은 주파수대역에서 신호가 뒤섞여 신호분리 성능이 저하되는 현상이 발생하고는 한다. 본 논문에서는 IVA로 분리된 신호에서 발생할 수 있는 주파수 빈 뒤섞임 문제를 전력비의 상관행렬을 이용한 적절한 후처리 과정을 통해서 해결할 수 있는 방법을 제안하고, 모의실험을 통해 제안한 방법의 타당성을 검증하였다. IVA의 후처리 과정으로 제안한 방법을 적용한 경우 주파수 빈의 뒤섞임 문제를 완전하게 해결할 수 있었으며, 이로 인해 다양한 경우의 혼합음성에 대한 실험에서 분리된 음성의 음질이 PESQ 값으로 1정도 향상함을 볼 수 있었다. 향후 제안한 방법을 발전시켜 일반적인 FD-ICA에 적용할 수 있는 방법에 대해 연구할 계획이다.



(a)



(b)

그림 7. 혼합신호의 음원 분리 실험 결과
(a) 분리된 남성화자 신호에 대한 PESQ 값
(b) 분리된 여성화자 신호에 대한 PESQ 값

Figure 7. Results of source separation experiments on mixed signals (a) PESQ values on separated male speech signals (b) PESQ values on separated female speech signals

참고문헌

- [1] P. Comon. (1994). Independent component analysis, a new concept?, *Signal Processing*, 36, 287-314.
- [2] E. Bingham and A. HyvÄärinen. (2000). A fast fixed-point algorithm for independent component analysis of complex valued signals, *International Journal of Neural Systems*, 10, 1-8.
- [3] T. Kim, H. T. Attias, S.-Y. Lee, and T.-W. Lee. (2007). Blind source separation exploiting higher-order frequency dependencies, *IEEE Transactions on, Audio, Speech, and Language Processing*, 15(1), 70-79.
- [4] 왕성양, 전성일, 배건성 (2011). 독립벡터 분석기법을 이용한 마이크로폰 배열 기반의 음성개선, 제 28 회 음성통신 및 신호처리 학술대회, 200-203.
- [5] H. Sawada, S. Araki, and S. Makino. (2007). Measuring dependence of bin-wise separated signals for permutation alignment in frequency-domain BSS, *IEEE International Symposium on Circuits and Systems (ISCAS)*, 3247-3250.

- [6] K. D. Donohue. Systems Array Processing Toolbox.
<http://www.engr.uky.edu/~donohue/>
- [7] PESQ, <http://www.itu.int/rec/T-REC-P.862/en>
- [8] J. Benesty, et al. (2008), *Microphone array signal processing*.
Berlin: Springer Verlag.
- [9] NoiseX-92, <http://www.speech.cs.cmu.edu/comp.speech/Section1/Data/noisex.html>

• **추쑤하오 (Chu, Zhihao)**

Email: chuzhihao1016@yeah.net
경북대학교 대학원 전자공학부
관심분야: 음성신호처리, ICA 및 IVA

• **배건성 (Bae, Keunsung) 교신저자**

경북대학교 IT대학 전자공학부
대구 북구 산격동 1370번지
Tel: 053-950-5527
Email: ksbae@ee.knu.ac.kr
관심분야: 음성신호처리, 디지털신호처리, 수중음향신호처리
1979~현재 전자공학부 교수