

분리행렬의 가중 내적 제한조건을 이용한 FDICA 알고리즘의 수렴속도 향상

Improvement of convergence speed in FDICA algorithm with weighted inner product constraint of unmixing matrix

전 성 일¹⁾ · 배 건 성²⁾

Quan, Xingri · Bae, Keunsung

ABSTRACT

For blind source separation of convolutive mixtures, FDICA(Frequency Domain Independent Component Analysis) algorithms are generally used. Since FDICA algorithm such as Sawada FDICA, IVA(Independent Vector Analysis) works on the frequency bin basis with a natural gradient descent method, it takes much time to converge. In this paper, we propose a new method to improve convergence speed in FDICA algorithm. The proposed method reduces the number of iteration drastically in the process of natural gradient descent method by applying a weighted inner product constraint of unmixing matrix. Experimental results have shown that the proposed method achieved large improvement of convergence speed without degrading the separation performance of the baseline algorithms.

Keywords: Hadamard product form, FDICA

1. 서론

음원과 전달 채널에 대한 사전정보 없이 여러 음원이 혼합된 신호에서 원래의 음원을 추출하는 기법, 즉, 암목적 신호분리(BSS: Blind Source Separation)에 대한 연구는 오디오 및 음성 신호처리 분야에서 활발히 연구되고 있다. ICA(Independent Component Analysis)는 BSS의 한 가지 접근방식으로서, 음원이 확률 통계적으로 서로 독립이라는 가정 하에 확률모델을 이용하여 혼합된 신호로부터 음원을 분리한다. 1994년 Comon이 상호정보량(mutual information)을 최소화 하는 방법으로 암목적 신호분리가 가능함을 보이면서[1] ICA 알고리즘에 대한 대폭적인 연구가 시작되었다[2,3]. 2000년대 초반까지는 혼합신호가 원 음원의 선형 조합 형태로 모델링되는 순시혼합모델(instantaneous mixing model)을 이용한 기본 ICA 알고리즘들이

정립되면서[4-6], 이를 전달 채널이 컨볼루션 형태로 주어지는 혼합모델(convolutive mixing model)에 적용하기 위한 주파수영역 ICA(FDICA: Frequency Domain ICA)에 관한 많은 연구가 이루어지고 있다[7-9].

시간영역에서의 컨볼루션 형태는 주파수영역에서 곱의 형태로 주어지므로 FDICA에서는 각 주파수 빈(frequency bin)에 대해 순시혼합모델이 되며, 각각의 주파수 빈 별로 순시혼합모델의 ICA 알고리즘을 수행할 수 있다. 이 경우에도 ICA의 고유문제인 스케일 모호성(scale ambiguity)과 치환 모호성(permutation ambiguity) 문제가 발생하는데, 분리된 신호의 이득이 달라지는 스케일 모호성 문제는 최소왜곡정리(minimal distortion principle)[10]로 일정하게 보상 가능하다. 그러나 분리된 신호의 순서가 바뀌는 치환 모호성은 FDICA에서 분리된 음원의 주파수 빈이 다른 음원의 주파수 빈과 뒤섞인 형태로 나타나므로 시간영역에서 바르게 분리된 신호를 얻기 위해서는 주파수 빈 뒤섞임 문제의 해결이 필수 조건이며, 이에 대해서는 여전히 많은 연구가 수행되고 있다[17,8,9,11,12].

주파수 영역 ICA를 구현하기 위하여서는 STFT(Short Time Fourier Transform)를 하여야 하는데, 일반적으로 잔향이 있는

1) 경북대학교, comet263@hotmail.com

2) 경북대학교, ksbae@ee.knu.ac.kr, 교신저자

접수일자: 2015년 11월 12일

수정일자: 2015년 12월 14일

게재결정: 2015년 12월 14일

환경에서는 프레임 길이를 크게 해주어야하기에 계산량이 많아지게 된다. 또한 일반화된 공분산행렬을 구하여야 하는데 이때 일정한 프레임 수를 확보하여야 하므로 계산량이 증가하게 되면서 FDICA 알고리즘이 각 주파수 빈 별로 많은 차수의 iteration이 소요되므로 수렴속도가 느려진다. [14]에서는 수렴속도를 개선하기 위하여 FDICA에서 일반적으로 사용하는 natural gradient descent 알고리즘 대신에 수렴속도가 빠르다고 알려져 있는 Newton method로 접근하였지만 여전히 100차 정도의 iteration을 수행하여야 한다.

본 논문에서는 Sawada FDICA와 IVA 같은 FDICA 알고리즘의 분석을 통해 이 기법들의 분리행렬을 구하는 주요 과정이 아다마르 곱(Hadamard product) 형식으로 표현됨을 보이고, 아다마르 곱 형식으로 표현된 FDICA 알고리즘의 iteration 과정에서 분리행렬에 가중 내적 제한조건을 추가하고 해를 구하는 방법을 제안함으로써 분리 성능의 저하없이 FDICA 알고리즘 수렴 과정에서의 iteration 차수를 줄이고 수렴속도를 현저하게 개선할 수 있음을 보였다.

본 논문의 구성은 다음과 같다. 2장에서 먼저 FDICA 혼합 모델 및 Sawada FDICA와 IVA 알고리즘을 소개하고, 3장에서 이 두 가지 FDICA 알고리즘이 모두 아다마르 곱 형식으로 표현됨을 보이며, 제안한 분리행렬에 가중내적 제한조건을 추가하고 해를 구하는 FDICA 알고리즘에 대해 설명한다. 4장에서는 제안한 방법의 타당성을 검증하기 위한 실험 및 결과에 대해 설명하고, 5장에서 결론을 맺는다.

2. FDICA 혼합모델 및 분리 알고리즘

2.1 FDICA 혼합 모델

FDICA 혼합 및 분리 모델은 <그림 1>과 같이 표현할 수 있는데, 여기서 $s_i(t)$, $x_m(t)$, $y_k(t)$ ($i, m, k = 1, 2, \dots, N$)는 각각 시간영역에서의 독립 음원, 마이크로폰 혼합 신호 및 분리된 신호이며, $h_{mi}(t)$ 는 i 번 음원으로부터 m 번 마이크로폰까지의 임펄스 응답, $w_{km}(t)$ 는 FDICA 알고리즘의 분리행렬 인자로 m 번 혼합신호로부터 k 번 분리 신호까지의 분리 필터이다.

ICA는 $s_i(t)$, $h_{mi}(t)$ 에 대해 모르는 상황에서, 혼합 신호 $x_m(t)$ 만 이용하여 분리 필터 $w_{km}(t)$ 를 추정하여 분리 신호

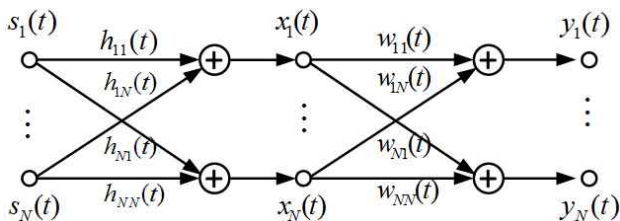


그림 1. FDICA 혼합 및 분리 모델

Figure 1. FDICA mixing and separation model

$y_k(t)$ 를 얻는 것이다. <그림 1>의 혼합모델을 수식으로 표현하면 식 (1)과 같이 주어지고, 주파수 영역으로 변환하면 식 (2)와 같이 되며, 간단하게 행렬 형태로 정리하면 식 (3)과 같다

$$x_m(t) = \sum_{i=1}^N h_{mi}(t) * s_i(t) \quad (1-1)$$

$$y_k(t) = \sum_{j=1}^N w_{kj}(t) * x_j(t) \quad (1-2)$$

$$X_m(\omega, \tau) = \sum_{i=1}^N H_{mi}(\omega) S_i(\omega, \tau) \quad (2-1)$$

$$Y_k(\omega, \tau) = \sum_{j=1}^N W_{kj}(\omega) X_j(\omega, \tau) \quad (2-2)$$

$$\mathbf{X}(\omega, \tau) = \mathbf{H}(\omega) \mathbf{S}(\omega, \tau) \quad (3-1)$$

$$\mathbf{Y}(\omega, \tau) = \mathbf{W}(\omega) \mathbf{X}(\omega, \tau) \quad (3-2)$$

식 (3)에서 $\mathbf{S}(\omega, \tau)$, $\mathbf{X}(\omega, \tau)$, $\mathbf{Y}(\omega, \tau)$ 는 각각 음원, 혼합 신호 및 분리 신호의 STFT이고, τ 는 프레임 번호, ω 는 주파수 빈을 나타내며, $\mathbf{S}(\omega, \tau) = [S_1(\omega, \tau), S_2(\omega, \tau), \dots, S_N(\omega, \tau)]^T$ 이다. $\mathbf{H}(\omega)$, $\mathbf{W}(\omega)$ 는 각 주파수 빈에서의 혼합행렬과 분리행렬을 의미하며, 음원 수와 혼합신호의 수가 모두 N 이라고 가정하면 $N \times N$ 행렬이 되는데 $\mathbf{H}(\omega)$ 행렬의 (m, i) 번째 원소는 $H_{mi}(\omega)$ 로 표현된다.

2.2 Sawada FDICA 알고리즘

논문 [7]에서 Sawada는 복소수 영역에서의 $\tanh(\cdot)$ 함수를 극좌표계 형태로 사용하면 음원 분리 성능이 향상됨을 보였다. Sawada FDICA 알고리즘의 업데이트 수식은 식 (4)와 같다.

$$\Delta \mathbf{W}(\omega) = [\mathbf{I} - E[\Phi(\mathbf{Y}(\omega, \tau)) \mathbf{Y}^H(\omega, \tau)]] \mathbf{W}(\omega) \quad (4-1)$$

$$\mathbf{W}(\omega) \leftarrow \mathbf{W}(\omega) + \eta \Delta \mathbf{W}(\omega) \quad (4-2)$$

$$\mathbf{Y}(\omega, \tau) = \mathbf{W}(\omega) \mathbf{X}(\omega, \tau) \quad (4-3)$$

식 (4)에서 \mathbf{I} 는 단위행렬, $E[\cdot]$ 는 τ 에 관한 기대치, η 는 학습률(learning rate), 활성화 함수(activation function) Φ 는 식 (5)와 같이 정의된다.

$$\Phi[Y_k(\omega, \tau)] = \tanh[\gamma |Y_k(\omega, \tau)|] \exp[j \text{Arg}(Y_k(\omega, \tau))] \quad (5)$$

여기서 γ 는 shape 파라미터, $|\cdot|$ 는 복소수 magnitude 연산, $\text{Arg}(\cdot)$ 는 복소수 argument 연산이며, $j = \sqrt{-1}$ 이다. Φ 는 벡터 함수인데 아래 모든 벡터 함수는 각 원소별 비선형 함수를 적용한 벡터 함수로 표기한다. 즉, $\boldsymbol{\zeta} = [\zeta_1 \dots \zeta_N]^T$ $\Phi(\boldsymbol{\zeta}) = [\Phi(\zeta_1), \dots, \Phi(\zeta_N)]^T$ 이다. Sawada FDICA의 활성화 함수

$\Phi[Y_k(\omega, \tau)]$ 는 $|Y_k(\omega, \tau)|$ 에 $\tanh(\cdot)$ 함수를 사용하고, 위상은 $Y_k(\omega, \tau)$ 와 같다. $\mathbf{C}(\omega) = E[\Phi(\mathbf{Y}(\omega, \tau))\mathbf{Y}^H(\omega, \tau)]$ 를 일반화된 공분산 행렬이라고 명명하면, FDICA는 이 행렬이 단위행렬에 가까워지도록 식 (4)의 과정을 반복하게 된다. Sawada FDICA는 각각의 주파수 빈에 대해 순시혼합모델 ICA를 수행함으로써 컨볼루션으로 혼합된 신호를 분리할 수 있으며, ICA의 치환 모호성 문제는 채널 특성과 신호 특성에 관한 정보를 추가하는 방식으로 접근하였다[11,12].

2.3 IVA 알고리즘

Sawada FDICA 알고리즘은 각각의 주파수 빈에 대해 단독으로 처리하기에 치환 모호성 문제가 심각하며 후처리 과정이 필요한 반면, IVA 알고리즘은 전체 주파수 빈을 하나의 벡터로 간주하고 벡터의 확률 분포를 모델링 하는 방법으로 주파수 빈 사이의 상관성을 고려한다.

k 번째 채널의 STFT $Y_k(\omega, \tau)$ 에서 전체 주파수 빈을 하나의 랜덤 벡터로 보면 (관측 샘플은 프레임 수만큼 있음), 랜덤 변수 $\mathbf{Y}_k = [Y_k(1), \dots, Y_k(\omega), \dots]^T$ 의 확률 분포 $p_{\mathbf{Y}_k}(\mathbf{Y}_k)$ 를 식 (6)과 같이 평균이 0이고 공분산이 단위행렬인 형태로 모델링한다[8].

$$p_{\mathbf{Y}_k}(\mathbf{Y}_k) = \exp[-\sqrt{\mathbf{Y}_k^H \mathbf{Y}_k}] = \exp[-\sqrt{\sum_{\omega} |Y_k(\omega)|^2}] \quad (6)$$

IVA 알고리즘의 비용함수는 식 (7)과 같으며, 이를 최소화하기 위한 활성화 함수는 $\Phi[\mathbf{Y}_k] = \frac{\partial}{\partial Y_k(\omega)} \log[p_{\mathbf{Y}_k}(\mathbf{Y}_k)]$ 로 정의되므로[15], IVA의 활성화 함수는 식 (8)과 같다.

$$\rho(\mathbf{W}) = -\log(|\det(\mathbf{W})|) - \sum_{k=1}^N \log[p_{\mathbf{Y}_k}(\mathbf{Y}_k)] \quad (7)$$

$$\Phi[Y_i(\omega, \tau)] = \frac{Y_i(\omega, \tau)}{\sqrt{\sum_{\omega} |Y_i(\omega, \tau)|^2}} \quad (8)$$

식 (8)에서 분모는 전체 주파수 빈 파워의 합으로 사용됨으로써 IVA에서는 주파수 빈 사이의 상관성이 고려되게 된다. IVA 분리 알고리즘은 활성화 함수에 대한 정의만 다를 뿐 분리행렬 업데이트는 식 (4)의 Sawada FDICA와 똑같다. 식 (8)에서 보면 IVA의 활성화 함수 $\Phi[Y_k(\omega, \tau)]$ 는, magnitude는 $|Y_i(\omega, \tau)| / \sqrt{\sum_{\omega} |Y_i(\omega, \tau)|^2}$ 이고, 위상은 $Y_k(\omega, \tau)$ 와 같다.

3. FDICA 알고리즘의 아다마르 곱 형식 및 분리행렬의 가중 내적 제한조건 적용

3.1 FDICA 알고리즘의 아다마르 곱 형식

Sawada FDICA에서 사용하는 활성화 함수 식 (5)를 일반화된 공분산 행렬에 대입하면 식 (9)가 얻어진다.

$$\begin{aligned} \mathbf{C}(\omega) &= E[\Phi(\mathbf{Y}(\omega, \tau))\mathbf{Y}^H(\omega, \tau)] \\ &= E[\tanh(\gamma|\mathbf{Y}(\omega, \tau)|) \exp(j \text{Arg}(\mathbf{Y}(\omega, \tau))) \mathbf{Y}^H(\omega, \tau)] \\ &= E\left[\left(\frac{\tanh(\gamma|\mathbf{Y}(\omega, \tau)|)}{|\mathbf{Y}(\omega, \tau)|} \circ \mathbf{Y}(\omega, \tau)\right) \mathbf{Y}^H(\omega, \tau)\right] \end{aligned} \quad (9)$$

여기서 \circ 은 아다마르 곱 연산자이다. IVA 알고리즘의 활성화 함수 식 (8)을 일반화된 공분산 행렬에 대입하면 식 (10)과 같다.

$$\mathbf{C}(\omega) = E\left[\left(\frac{1}{\sqrt{\sum_{\omega} |\mathbf{Y}(\omega, \tau)|^2}} \circ \mathbf{Y}(\omega, \tau)\right) \mathbf{Y}^H(\omega, \tau)\right] \quad (10)$$

식 (9)와 (10)은 모두 아다마르 곱 형식 이므로 식 (11)과 같은 형식으로 일반화 할 수 있다.

$$\mathbf{C}(\omega) = E[(\mathbf{G}(\omega, \tau) \circ \mathbf{Y}(\omega, \tau)) \mathbf{Y}^H(\omega, \tau)] \quad (11)$$

식 (11)에서 $\mathbf{G}(\omega, \tau) = [g_1(\omega, \tau), \dots, g_N(\omega, \tau)]^T$ 는 $\mathbf{Y}(\omega, \tau)$ 에 관련된 비선형 이득 벡터 함수로 볼 수 있으며, Sawada FDICA[7], IVA[8], 서브밴드 IVA[13] 등 다양한 FDICA 알고리즘들을 모두 이와 같은 아다마르 곱 형식으로 변환 시킬 수 있다.

$\mathbf{Y}(\omega, \tau) = \mathbf{W}(\omega) \mathbf{X}(\omega, \tau)$ 를 식 (11)에 대입하면 식 (12), (13)과 같이 정리할 수 있다(부록 1).

$$\mathbf{R}_i(\omega) = E[g_i(\omega, \tau) \mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau)] \quad (12)$$

$$\mathbf{C}_{ik}(\omega) = \mathbf{W}_i(\omega) \mathbf{R}_i(\omega) \mathbf{W}_k^H(\omega) \quad (13)$$

$\mathbf{C}_{ik}(\omega)$ 는 $\mathbf{C}(\omega)$ 행렬의 (i, k) 번째 원소이고, $\mathbf{W}_i(\omega)$ 는 분리행렬 $\mathbf{W}(\omega)$ 의 i 번째 행벡터이다. 여기서 $\mathbf{R}_i(\omega)$ 는 Hermitian positive-semidefinite 행렬이다(부록 1). 혼합신호 입력과 분리된 출력 신호가 각각 2개인 2×2 경우 식 (12), (13)은 아래와 같이 주어진다,

$$\mathbf{R}_1(\omega) = E[g_1(\omega, \tau) \mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau)] \quad (14)$$

$$\mathbf{R}_2(\omega) = E[g_2(\omega, \tau) \mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau)] \quad (15)$$

$$\mathbf{C}(\omega) = \begin{bmatrix} \mathbf{W}_1(\omega) \mathbf{R}_1(\omega) \mathbf{W}_1^H(\omega) & \mathbf{W}_1(\omega) \mathbf{R}_1(\omega) \mathbf{W}_2^H(\omega) \\ \mathbf{W}_2(\omega) \mathbf{R}_2(\omega) \mathbf{W}_1^H(\omega) & \mathbf{W}_2(\omega) \mathbf{R}_2(\omega) \mathbf{W}_2^H(\omega) \end{bmatrix} \quad (16)$$

3.2 가중 내적 제한조건을 이용한 FDICA

FDICA는 일반화된 공분산 행렬 $\mathbf{C}(\omega)$ 가 단위행렬이 될 때까지 식 (4)를 반복하여 수행한다. $\mathbf{C}(\omega) = \mathbf{I}$ 는 분리행렬의 i

번째 행 $W_i(\omega)$ 와 k 번째 행 $W_k(\omega)$ ($k \neq i$)의 가장 내적이 영이 됨을 의미한다. 본 논문에서는 2×2 경우에 $C(\omega) = I$ 로 되는 분리 행렬의 각 행 $W_i(\omega)$ 를 명시적으로 찾을 수 있음을 보였다(부록 2). 본 논문에서는 이 명시적인 해를 이용하여 가장 내적 제한조건을 추가하는 FDICA 알고리즘을 제안하였다. 제안한 FDICA 알고리즘의 의사코드는 <그림 2>와 같다. 먼저 $Y(\omega, \tau)$ 를 초기화한 다음, Sawada FDICA는 식 (9), IVA는 식 (10)과 같이 $Y(\omega, \tau)$ 를 이용하여 $G(\omega, \tau)$ 를 계산한다. 그런 다음, 식 (12)와 같이 모든 i 에 대하여 $R_i(\omega)$ 를 계산하고, $C(\omega) = I$ 로 되는 분리 행렬의 각 행 $W_i(\omega)$ 를 명시적으로 구한다. 마지막으로 $Y(\omega, \tau) = W(\omega)X(\omega, \tau)$ 로 분리 신호를 업데이트 하고 다음 iteration 과정으로 들어간다.

```

 $Y(\omega, \tau)$  초기화,  $Y(\omega, \tau) = X(\omega, \tau)$ 
for iteration = 1 : max_iteration
    for  $\omega = 1 : \Omega$ 
         $Y(\omega, \tau)$ 를 이용하여  $G(\omega, \tau)$  계산,  $\forall \omega, \tau$ 
         $R_i(\omega) = E[g_i(\omega, \tau)X(\omega, \tau)X^H(\omega, \tau)]$ ,  $\forall i$ 
         $R_i(\omega) \rightarrow W_i(\omega)$ ,  $\forall i$ 
         $Y(\omega, \tau) = W(\omega)X(\omega, \tau)$ 
    end
수렴 조건 판단
end
    
```

그림 2. 가장 내적 제한조건을 이용하는 아다마르 곱 형식의 FDICA 알고리즘 의사코드
 Figure 2. Pseudo-code of Hadamard product form FDICA algorithm with weighted inner product constraint

4. 실험 및 검토

FDICA 음원분리 과정에서 제안한 알고리즘의 수렴속도 개선 성능을 분석하기 위하여 image method toolbox[16]로 룸 임펄스 응답을 생성하고 음성신호와 컨볼루션 하여 마이크로폰 혼합신호를 생성하였다. 룸 환경은 <그림 3>과 같고, 룸 크기는 $7m \times 5m \times 2.75m$, 마이크로폰은 6cm 간격을 두고 두 개 설치하며[8], 잔향 파라미터는 $T_{60} = 0.15$ 초로 설정하였다. <그림 4>는 생성한 룸 임펄스 응답 예를 보여준다.

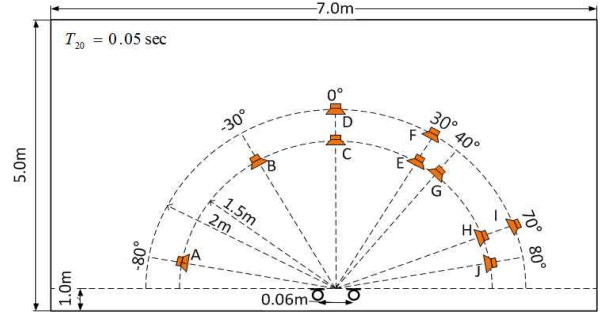


그림 3. 실험 환경 설정
 Figure 3. Experimental environment setting

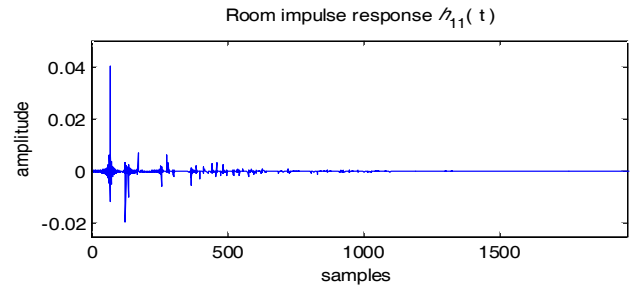


그림 4. 룸 임펄스 응답 예
 Figure 4. An example of room impulse response

음원은 인터넷[17]에서 내려 받은 남자 화자 3명과 여자 화자 3명의 영어 문장을 3초 길이로 잘라서 둘 씩 조합하여 사용하였으며, 샘플링주파수는 16kHz이다. 음원의 위치는 멀리 떨어져 있는 경우, 가까이 있는 경우, 동일 방위각에 있는 경우를 고려하여 AH, BG, EG, HJ, CD, EF, HI 일곱 가지로 조합하여 실험하였고[8], 혼합신호는 $f1+f2$, $m1+m2$, $f3+m3$ 의 세 가지 조합으로 총 21가지 경우에 대해 실험하였다. 표 1에 음원 및 위치의 조합, 즉, setting의 번호를 나타내었다.

STFT는 프레임 사이즈 2048 샘플, 75% 오버랩 시키며, 해밍 윈도우를 사용하고 2048 포인트 FFT를 수행하며 minimal distortion principle[10]을 적용한다. 수렴 여부는 IVA 알고리즘에서와 똑같이 식 (7)을 사용하며, 값의 변화율이 10^{-6} 이하가 되면 iteration 과정을 마치도록 하였다. Sawada FDICA는 주파수 빈 뒤섞임 정정을 위한 후처리 과정이 필요한데 전력비를 이용하는 방법을 사용하였다[18].

본 연구에서 제안한 방법을 적용하여 수렴속도 성능 향상을 언더라도 음원 분리 성능이 저하하지 않음을 보이기 위해 식 (17)과 같이 정의되는 SIR(Signal to Interference Ratio)[12]을 사용하였다.

$$SIR_i = 10 \log \frac{\sum_t |A_{ii}(t) * s_i(t)|^2}{\sum_t \sum_{k \neq i} |A_{ik}(t) * s_k(t)|^2} \quad (17)$$

식 (17)에서 $A(\omega) = W(\omega)H(\omega)$, $A_{ik}(\omega)$ 는 $A(\omega)$ 의 i 번째

표 1. 음원 및 위치 조합을 나타내는 실험 setting 번호

Table 1. Experimental setting number representing mixed sound sources and their positions

음원 \ 위치	AH	BG	EG	HJ	CD	EF	HI
f1+f2	1	4	7	10	13	16	19
m1+m2	2	5	8	11	14	17	20
f3+m3	3	6	9	12	15	18	21

행 k 번째 열이며, $A_{i,k}(t)$ 는 $A_{i,k}(\omega)$ 의 IFFT이다.

$$Y(\omega) = W(\omega)X(\omega) = W(\omega)H(\omega)S(\omega) = A(\omega)S(\omega)$$

이므로 분리 신호의 순서와 음원의 순서가 같다고 가정하면,

$$Y_i(\omega) = A_{ii}(\omega)S_i(\omega) + \sum_{k \neq i} A_{ik}(\omega)S_k(\omega) \quad (18)$$

로 표현되어 $A_{ii}(\omega)S_i(\omega)$ 는 원하는 신호, $\sum_{k \neq i} A_{ik}(\omega)S_k(\omega)$ 는 간섭 신호가 된다. 위 분석으로부터 식 (17)은 원하는 신호의 에너지 대 간섭 신호 에너지의 비를 뜻함을 알 수 있다.

4.1 학습률 파라미터 선택

Sawada FDICA와 IVA 모두 natural gradient descent 알고리즘을 사용하는데, 학습률 η 가 크면 발산할 수 있고, 작으면 수렴속도가 느려지게 된다. 따라서 실험적 방법으로 적절한 학습률 파라미터를 선택할 필요가 있다.

IVA 알고리즘에서 $\eta = 0.1, 0.2, 0.3$ 에 대해 실험해보면 SIR은 <그림 5(a)>와 같이 학습률의 영향을 거의 받지 않는다. $\eta = 0.4$ 부터 발산하는 현상이 발생했는데, 수렴속도는 <그림 5(b)>에 η 에 따른 수렴곡선 변화(100차까지만 보여줌)를 보인 것처럼 예상대로 $\eta = 0.3$ 일 때 가장 빠르게 나타난다. 학습률이 크면 불안정한 경우가 나타날 수 있으므로, 좀 더 안정적으로 동작할 것을 고려하여 $\eta = 0.2$ 를 선택하였다.

Sawada FDICA의 분리 성능은 <그림 6>에 보여주는 바와 같다. 실험에 사용한 데이터에서 학습률 $\eta = 0.2$ 일 때 발산하는 현상이 발생하였으므로 $\eta = 0.1$ 로 선택한다. IVA는 SIR이 약 7dB 정도 되는데 반하여, Sawada FDICA는 1~3번(음원 위치 AH)과 5번(음원 위치 BG) 실험에서 비교적 높은 분리 성능을 보이고 기타 실험환경에서는 음원 분리가 잘 안됨을 볼 수 있다. 이 결과는 IVA 논문 [8]에서 관찰한 바와 비슷한데, 음원의 위치가 가까이 있는 경우 린 임펄스 응답이 매우 비슷하고, 이때 Sawada FDICA 알고리즘에서는 본 논문에서 사용한 후처리 과정[18]이 잔향환경에서 성능이 좋지 않아 음원 분리 성능이 저하한 것으로 생각된다.

4.2 분리 성능 및 수렴속도 분석

제한한 가중 내적 제한 조건을 이용한 Fast IVA와 기존의 IVA 알고리즘의 분리 성능 및 수렴속도를 <그림 7>에 보여주고 있다(50차 반복까지만 나타내었음). SIR 분리 성능은 기존의 방법과 거의 비슷하고 수렴속도는 가중 내적 제한조건을 이용하는 방법이 기존 알고리즘보다 훨씬 더 빠름을 볼 수 있다. <표 2>에서 21개 실험 환경에 대한 반복 차수와 수렴 소요 시간을 보여주는데, 기존의 IVA는 평균 240차 정도의 반복 차수가 필요하고 평균 16초 정도의 시간이 소요되는데 반해, 제한한 가중 내적 제한조건을 이용하는 IVA는 평균 15차 반복 차수에 3초 정도의 알고리즘 수렴 시간이 소요되었음을 볼 수 있다. 이때 실험에 사용한 PC는 CPU가 Intel Core i7-2.93GHz이며, MATLAB을 이용하였다.

제한한 가중 내적 제한 조건을 이용한 Fast Sawada FDICA와 기존의 Sawada FDICA 알고리즘의 분리 성능 및 수렴속도를 <그림 8>에 나타내었다(200차 반복까지만 나타내었음). SIR 분리 성능은 대체적으로 비슷한 양상을 보이고 있고, 수렴속도는 가중 내적 제한조건을 이용하는 방법이 기존 알고리즘보다 현저하게 빨리 수렴함을 볼 수 있다. <표 3>에서 21개 실험 환경에 대한 반복 차수와 수렴 소요 시간을 보여주는데, 제한한 가중 내적 제한조건을 이용함으로써 평균 245차 반복차수를 평균 41차로 낮출 수 있으며, 알고리즘 실행 시간도 33초에서 10초 이내로 줄일 수 있음을 볼 수 있다.

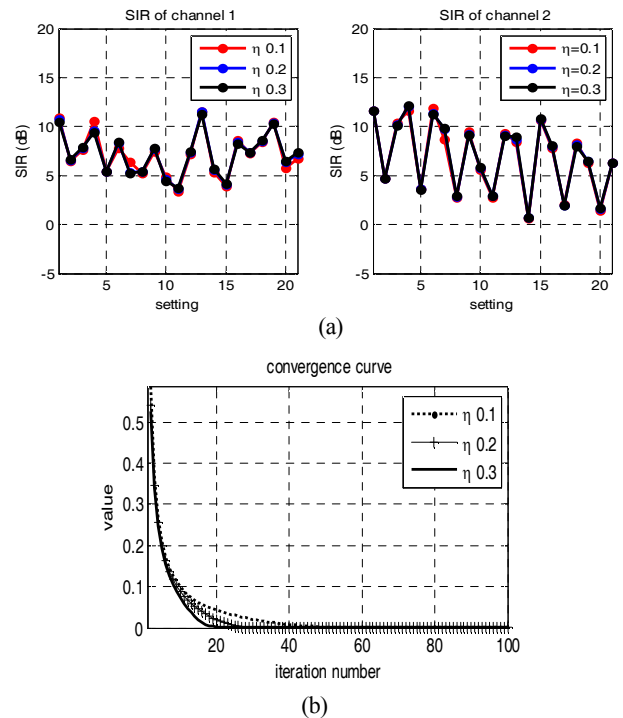


그림 5. 학습률에 따른 IVA 알고리즘 분리 성능 (a) SIR, (b) 수렴 곡선 예 (Setting 1)

Figure 5. IVA algorithm performance depending on the learning rate (a) SIR, (b) an example of convergence curve

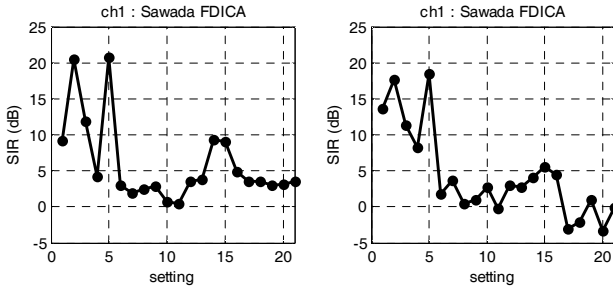
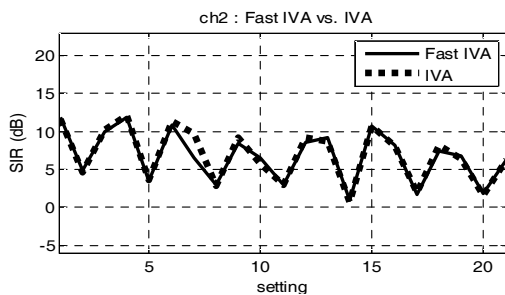
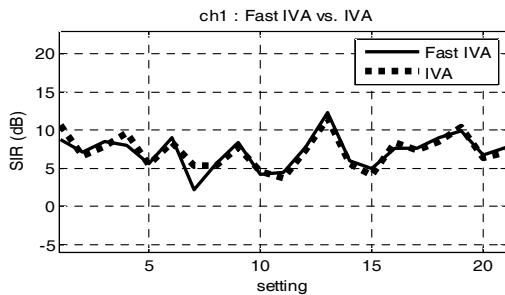
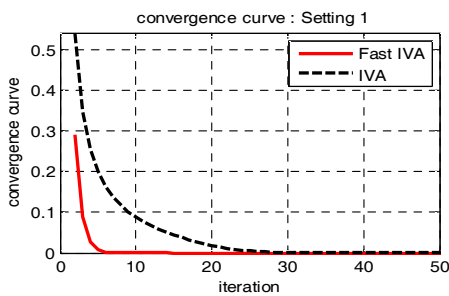


그림 6. Sawada FDICA 알고리즘 분리 성능
Figure 6. SIR of Sawada FDICA



(a)



(b)

그림 7. 가중 내적 제한조건을 이용한 IVA와 기존의 IVA 알고리즘의 분리 성능 및 수렴속도 비교
(a) SIR, (b) 수렴속도

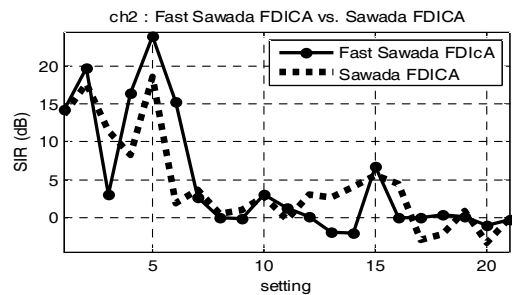
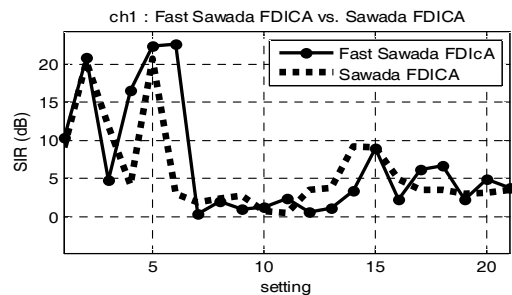
Figure 7. Performance comparison of SIR and convergence speed between IVA with weighted inner product constraint and original IVA algorithm (a) SIR, (b) convergence speed

<표 4>에 가중 내적 제한조건을 이용하는 FDICA와 기존 FDICA 알고리즘의 21차례 실험에 대한 평균 SIR, 평균 반복 차수와 평균 소요시간을 정리하였다. 제안한 방법은 기존의 알고리즘에 비해 평균 SIR이 비슷하면서, 매 차 반복에서 계산량

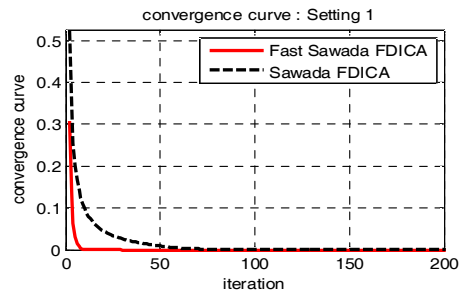
이 조금 증가하지만 반복 차수가 현저히 줄어들어 수렴에 소요되는 전체 시간이 현저하게 짧음을 볼 수 있다.

4. 결론

컨볼루션 형태의 혼합 신호를 암묵적으로 분리하기 위하여 다양한 형태의 FDICA 기법들이 제안되고 있는데, 일반적으로 알고리즘의 수렴에 필요한 iteration 차수가 많으며, 따라서 수렴 소요 시간이 긴 문제점이 있다. 본 논문에서는 FDICA 알고리즘을 아다마르 곱 형식으로 표현하고, 2×2 경우의 분리행렬에 가중 내적 제한조건을 적용하고 명시적인 해를 구하는 방법을 제안하였고, 실험을 통해 제안한 방법은 FDICA 알고리즘의 신호분리 성능은 저하시키지 않으면서 iteration 차수를 현저히 줄임으로써 수렴속도를 개선할 수 있음을 보였다.



(a)



(b)

그림 8. 가중 내적 제한조건을 이용한 Sawada FDICA와 기존의 Sawada FDICA 알고리즘의 분리 성능 및 수렴속도 비교 (a) SIR, (b) 수렴속도

Figure 8. Performance comparison of SIR and convergence speed between Sawada FDICA with weighted inner product constraint and original Sawada FDICA (a) SIR, (b) convergence speed

표 2. 수렴속도 비교 : IVA
Table 2. Comparison of convergence speed : IVA

실험 번호	IVA		가중 내적 제한조건을 이용한 IVA	
	반복 차수	수렴 시간(초)	반복 차수	수렴 시간(초)
1	238	15.9	14	2.6
2	250	16.4	14	3.0
3	230	15.4	14	3.1
4	164	11.1	14	3.0
5	168	11.0	14	3.1
6	326	22.4	14	2.9
7	436	28.5	15	3.2
8	258	16.8	14	3.0
9	193	16.5	14	2.9
10	212	14.2	15	3.2
11	238	15.3	14	3.0
12	460	30.7	14	2.9
13	273	18.0	26	5.4
14	185	12.1	15	3.2
15	161	10.9	14	2.9
16	164	10.9	15	3.1
17	237	15.9	14	2.9
18	199	13.3	14	2.9
19	185	12.3	14	3.0
20	211	14.2	14	2.9
21	232	15.5	14	3.0
평균	239.05	16.06	14.76	3.10

표 3. 수렴속도 비교 : Sawada FDICA

Table3. Comparison of convergence speed : Sawada FDICA

실험 번호	Sawada FDICA		가중 내적 제한조건을 이용한 Sawada FDICA	
	반복 차수	수렴 시간(초)	반복 차수	수렴 시간(초)
1	223	38.2	29	7.6
2	191	28.6	27	6.3
3	256	37.6	42	10.7
4	203	30.7	32	6.7
5	200	30.4	30	9.0
6	219	32.2	30	7.2
7	247	35.7	48	9.7
8	248	35.0	47	11.5
9	274	38.7	47	11.7
10	236	34.7	41	8.3
11	234	33.7	45	9.0
12	252	35.5	36	7.3
13	268	31.3	44	8.9
14	286	33.7	37	7.5
15	321	37.8	50	10.1
16	250	29.4	45	9.0
17	216	25.4	42	8.4
18	287	33.9	51	10.2
19	259	30.5	45	9.1
20	229	27.8	38	7.6
21	256	33.9	54	10.7
평균	245.48	33.07	40.95	8.89

표 4. 기존 알고리즘과 제안한 방법의 성능 분석
Table 4. Comparison of performance between original and
proposed FDICA algorithms

	IVA	제안한 기법을 적용한 IVA	Sawada FDICA	제안한 기법을 적용한 Sawada FDICA
평균 SIR (dB)	7.09	6.98	5.15	5.87
평균 반복차수	239	15	245	41
평균 수렴 시간 (초)	16.06	3.10	33.07	8.89

이 논문은 방위사업청과 국방과학연구소의 지원으로
수행되었습니다.(계약번호:UD130007DD)

참고문헌

[1] P. Comon, (1994). Independent component analysis, a new concept?, *Signal Processing*, vol. 36, pp. 287-314.

[2] A. Hyvärinen, J. Karhunen, and E. Oja, (2001). *Independent component analysis*. John Wiley & Sons.

[3] A. Cichocki and S. Amari, (2003). *Adaptive Blind Signal and Image Processing (Learning Algorithms and Applications)*. New York: John Wiley.

[4] A. J. Bell and T. J. Sejnowski, (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Computation*, vol. 7, pp. 1129-1159.

[5] S.-i. Amari, A. Cichocki, and H. H. Yang, (1996). A new learning algorithm for blind signal separation. *Advances in neural information processing systems*, pp. 757-763.

[6] A. Hyvärinen and E. Oja, (1997). A fast fixed-point algorithm for independent component analysis. *Neural Computation*, vol. 9, pp. 1483-1492.

[7] H. Sawada, R. Mukai, S. Araki, and S. Makino, (2001). A polar-coordinate based activation function for frequency domain blind source separation. *in Proc. Int. Conf. ICA and BSS*, pp. 663-668.

[8] Kim, T., H. T. Attias, Lee, S.-Y., and Lee, T.-W., (2007). Blind source separation exploiting higher-order frequency dependencies. *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 15, pp. 70-79.

[9] F. Nesta, P. Svaizer, and M. Omologo, (2011). Convolutional BSS of short mixtures by ICA recursively regularized across frequencies. *Audio, Speech, and Language Processing, IEEE*

Transactions on, vol. 19, pp. 624-639.

- [10] K. Matsuoka, (2002). Minimal distortion principle for blind source separation. *Proceedings of the 41st SICE Annual Conference*, vol.4, pp. 2138-2143.
- [11] H. Sawada, S. Araki, R. Mukai, and S. Makino, (2007). Grouping separated frequency components by estimating propagation model parameters in frequency-domain blind source separation. *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 15, pp. 1592-1604.
- [12] H. Sawada, S. Araki, and S. Makino, (2007) Measuring dependence of bin-wise separated signals for permutation alignment in frequency-domain BSS. in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, pp. 3247-3250.
- [13] Lee, I. and Jamg, G. J., (2012). Independent vector analysis based on overlapped cliques of variable width for frequency-domain blind signal separation. *EURASIP Journal on Advances in Signal Processing*, vol. 2012, p. 113.
- [14] Lee, I., Kim, T. and Lee, T.-W., (2007). Fast fixed-point independent vector analysis algorithms for convolutive blind source separation. *Signal Processing*, vol. 87, pp. 1859-1871.
- [15] S. Haykin, (2009). *Neural networks and learning machines vol. 3*. Prentice Hall.
- [16] E. A. Lehmann. (2012). Image-source method: Matlab code implementation. www.eric-lehmann.com
- [17] K. d. Donohue. (2009). Audio Systems Lab Experimental Data. <http://www.engr.uky.edu/~donohue/audio/Data/audioexpdata.htm>
- [18] Quan, X. and Bae, K. (2014). Frequency bin alignment using covariance of power ratio of separated signals in multi-channel FD-ICA. *Phonetics and Speech Sciences*, Vol6, pp. 149-153. (전성일, 배건성, (2014). 다채널 주파수영역 독립성분분석에서 분리된 신호 전력비의 공분산을 이용한 주파수 빈 정렬. 말소리와 음성과학, vol. 6, pp. 149-153.)

부록 1

식 (11)에서 일반화된 공분산 행렬은 아다마르 곱 형식으로 표현됨을 보였다. 아래, $\mathbf{Y}(\omega, \tau) = \mathbf{W}(\omega) \mathbf{X}(\omega, \tau)$ 를 식 (11)에 대입해 식 (12), (13)을 유도하도록 한다.

$$\mathbf{C}(\omega) = E[(\mathbf{G}(\omega, \tau) \circ \mathbf{Y}(\omega, \tau)) \mathbf{Y}^H(\omega, \tau)]$$

여기서, $\mathbf{G}(\omega, \tau) = [g_1(\omega, \tau), \dots, g_N(\omega, \tau)]^T$,

$$\mathbf{Y}(\omega, \tau) = [Y_1(\omega, \tau), Y_2(\omega, \tau), \dots, Y_N(\omega, \tau)]^T,$$

$$C_{ik}(\omega) = E[g_i(\omega, \tau) Y_i(\omega, \tau) Y_k^H(\omega, \tau)]$$

$$Y_i(\omega, \tau) = \mathbf{W}_i(\omega) \mathbf{X}(\omega, \tau) \text{이며,}$$

여기서, $\mathbf{W}_i(\omega)$ 는 분리 행렬 $\mathbf{W}(\omega)$ 의 i 번째 행벡터,

$$\mathbf{X}(\omega, \tau) = [X_1(\omega, \tau), X_2(\omega, \tau), \dots, X_N(\omega, \tau)]^T$$

$$C_{ik}(\omega) = E[g_i(\omega, \tau) \mathbf{W}_i(\omega) \mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau) \mathbf{W}_k^H(\omega)]$$

$$= \mathbf{W}_i(\omega) E[g_i(\omega, \tau) \mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau)] \mathbf{W}_k^H(\omega) \text{ 이다.}$$

$\mathbf{R}_i(\omega) = E[g_i(\omega, \tau) \mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau)]$ 라고 정의하면,

$$C_{ik}(\omega) = \mathbf{W}_i(\omega) \mathbf{R}_i(\omega) \mathbf{W}_k^H(\omega) \text{로 표현된다.}$$

관찰: $\mathbf{R}_i(\omega)$ 는 Hermitian Positive-semidefinite 행렬이다.

증명: $\mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau)$ 행렬은 Hermitian positive-semidefinite, $g_i(\omega, \tau) > 0$ 이라 하면, 식 (9)의 Sawada FDICA와 식 (10)의 IVA에서 $g_i(\omega, \tau)$ 계산 시 모두 magnitude만 사용하므로 $g_i(\omega, \tau)$ 는 실수이고, 따라서 $g_i(\omega, \tau) \mathbf{X}(\omega, \tau) \mathbf{X}^H(\omega, \tau)$ 도 Hermitian positive-semidefinite로 된다. Hermitian 행렬의 합은 여전히 Hermitian행렬이고, positive-definite 행렬에 정수 가중치를 추가한 행렬의 합도 여전히 positive-definite 행렬로 되므로, $\mathbf{R}_i(\omega)$ 는 Hermitian positive-semidefinite 행렬이 된다. $\mathbf{X}(\omega, \tau) \neq 0, \forall \tau$ 이면 $\mathbf{R}_i(\omega)$ 는 full rank이며 positive definite로 된다.

부록 2

부록 2에서는 2×2 경우를 예로 $\mathbf{C}(\omega) = \mathbf{I}$ 를 만족시키는 분리행렬의 각 행 $\mathbf{W}_i(\omega)$ 를 명시적으로 구할 수 있음을 보인다.

statement: $\mathbf{R}_1, \mathbf{R}_2$ 는 Hermitian positive-definite(invertible) 2×2 정방행렬이고, $\mathbf{W}_1, \mathbf{W}_2$ 는 1×2 행벡터이다.

$$\begin{bmatrix} \mathbf{W}_1 \mathbf{R}_1 \mathbf{W}_1^H & \mathbf{W}_1 \mathbf{R}_2 \mathbf{W}_2^H \\ \mathbf{W}_2 \mathbf{R}_2 \mathbf{W}_2^H & \mathbf{W}_2 \mathbf{R}_1 \mathbf{W}_1^H \end{bmatrix} = \mathbf{I}$$

를 만족시키는 $\mathbf{W}_1, \mathbf{W}_2$ 를 찾는 문제가 된다.

$\mathbf{R}_1, \mathbf{R}_2$ 는 Hermitian positive-definite 행렬이므로, 고유치 분해를 하면, $\mathbf{R}_1 = \mathbf{U}_1 \mathbf{\Lambda}_1 \mathbf{U}_1^H$, $\mathbf{\Lambda}_1 > 0$, $\mathbf{U}_1^H \mathbf{U}_1 = \mathbf{U}_1 \mathbf{U}_1^H = \mathbf{I}$ 로 표

현된다. 행벡터 $\xi_i = W_i U_1 \Lambda_1^{1/2}$ 이라 하면,

$$\begin{bmatrix} \xi_1 \xi_1^H & \xi_1 \xi_2^H \\ \xi_2 M \xi_1^H & \xi_2 M \xi_2^H \end{bmatrix} = I \text{ 가 되며,}$$

여기서 $M = \Lambda_1^{-1/2} U_1^H R_2 U_1 \Lambda_1^{-1/2}$, $M = M^H$, M 행렬도 Hermitian positive-definite로 된다. ξ_1 은 단위 벡터이고, ξ_2 는 ξ_1 에 직교임과 동시에, ξ_1 과 M -orthogonal 하므로 M 행렬의 고유벡터 방향으로 선정할 수 있다. M 행렬에 대해 고유치 분해를 하면, $M = U_M \Lambda_M U_M^H$, $\Lambda_M > 0$, $U_M^H U_M = U_M U_M^H = I$ 이 되는데, $\xi_1 = U_{M2}^H$, $\xi_2 = \lambda_{M1}^{-1/2} U_{M1}^H$ 로 취하면 위 관계식을 모두 만족시킬 수 있다. 여기서 U_{M1} , U_{M2} 는 U_M 의 두 고유벡터(열벡터)이고, λ_{M1} 는 U_{M1} 에 대응되는 고유치이다. 고유벡터에 임의의 스칼라를 곱하여도 여전히 고유벡터가 되므로, 복소수 행벡터 ξ_1, ξ_2 에 임의의 $\exp[j\theta]$ 을 곱하여도 위 관계는 여전히 성립된다.

W_1, W_2 를 찾는 과정을 정리하면 아래와 같다.

$$\begin{aligned} R_1 &= U_1 \Lambda_1 U_1^H \\ M &= \Lambda_1^{-1/2} U_1^H R_2 U_1 \Lambda_1^{-1/2} \\ M &= U_M \Lambda_M U_M^H \\ W_1 &= U_{M2}^H \Lambda_1^{-1/2} U_1^H \\ W_2 &= \lambda_{M1}^{-1/2} U_{M1}^H \Lambda_1^{-1/2} U_1^H \end{aligned}$$

- **전성일 (Quan, Xingri)**
 경북대학교 IT 대학 전자공학부
 대구 북구 산격3동 1370번지
 Tel: 053-940-8627
 Email: comet263@hotmail.com
 관심 분야: 마이크로폰 배열 음성신호처리, adaptive filter
 2015년 8월 경북대학교 IT대학 박사학위 취득
- **배건성 (Bae, Keunsung), 교신저자**
 경북대학교 IT 대학 전자공학부
 대구 북구 산격3동 1370번지
 Tel: 053-940-8627
 Email: ksbae@mir.knu.ac.kr
 관심 분야: 음성신호처리, 디지털신호처리
 현재 경북대학교 IT 대학 전자공학부 교수