

## PCA 기반 오디오 신호 분리 알고리즘 구현

\*전재현<sup>1</sup> \*조두리<sup>1</sup> \*정제창<sup>‡</sup>

\*한양대학교 융합전자공학부

\*jaehyun12@hanyang.ac.kr \*serenitist@hanyang.ac.kr \*jjeong@hanyang.ac.kr

## Audio signal separation Algorithm Implementation based PCA

\*Jae-Hyeon Jeon \*Du-ri Jo \*Je-chang Jeong

\*Dept. of Electronic Engineering, Hanyang University, Korea

## 요약

다수의 음원이 특정한 공간에 산재하고 있을 때, 그 중 특정 음원에 주목하면 다른 음원과 분리되어 특정 음원만 들리는 현상을 카테일파티 현상이라고 한다. 심리적인 이 현상에 영감을 받아 음원을 분리하는 알고리즘이 만들어졌다. 이런 음원 분리방법을 Blind Source Separation(BSS) 이라고 하는데, 여러 신호가 섞이는 과정을 모르는 상태에서 음원을 분리한다는 뜻에서 Blind Source Separation 이라고 한다. BSS에 사용되는 알고리즘으로 주로 PCA, ICA이 있다. PCA는 2차원의 경우를, ICA는 그 이상의 고차원의 통계적 특성을 이용한다.

이에 본 논문은 PCA를 이용하여 두 음원을 분리하는 알고리즘을 구현하는데 역점을 두었다. PCA는 주로 음원보다는 이미 지 신호 처리에 초점이 맞추어져 있지만, 음원 분리에 있어서도 충분한 성능을 보여주므로, ICA를 이용한 음원 분리 알고리즘과의 비교를 통하여 장, 단점을 알아보고 추후 PCA의 응용 가능성을 알아보았다.

## 1. 서론

선형조합으로 섞이게 된 알 수 없는 기존 음원을 분리해 내는 과정은 Blind Source Separation(BSS)라고 한다. 이때 주로 사용되는 방법은 PCA(Principle Component Analysis), ICA(Independent Component Analysis), LDA(Linear Discriminant Analysis) 등이 있다. ICA는 3차 이상의 통계 특성에서 독립 성분을 분리하는데 초점이 맞추어져 있는데 반해, PCA는 2차 통계 특성분리에 특화되어있다. PCA는 두 개의 변수 집합의 공분산을 이용하여 주로 어떠한 분포를 나타내는지 알아내는 것이 목적이다.

본 논문에서는 PCA를 이용하여 음원을 분리하는 알고리즘을 구현하는데 목적이 있다. 구현한 알고리즘을 토대로 여러 종류의 오디오 음원 샘플을 적용시킨 후 각 샘플의 통계적 특징에 따라 분리 정도를 비교하였다. 기존 음원과 분리된 음원의 차이를 노이즈라 생각하고 Signal to Noise Ratio(SNR)을 구하여 음원의 특징에 따라 분리되는 정도를 비교하였다.

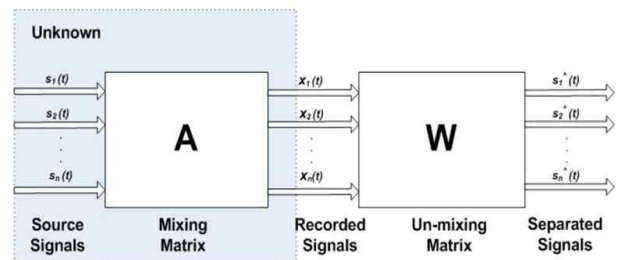
## 2. 문제 설계

관측한 오디오 신호를 기존 오디오 신호가 임의의 Mixing matrix를 통하여 선형 조합된 신호로 본다. 이때, 관측된 오디오 신호를  $X(t) = [x_1(t) \ x_2(t) \ \dots \ x_n(t)]^T$  라 하면 다음과 같은 식을 쓸 수 있다.

$$X(t) = A S(t) \quad (1)$$

여기서  $A$ 는 Mixing matrix,  $S$ 는  $n$ 차 기존 오디오 신호 벡터이

다. 본 논문에서는 관측된 오디오 신호인  $X(t)$ 를 가지고 Mixing matrix인  $A$ 의 역행렬을 PCA를 통해 추정하고 이를 통해  $S(t) = A^{-1}X(t)$ 를 이용하여 기존의 섞이기 전 오디오 신호  $S(t)$ 를 분리하는데 목적이 있다.



<그림 1. 기존 오디오 신호와 Mixing matrix, PCA algorithm을 이용한 신호 분리 >

이때 필요한 가정은 다음과 같다.

1. Mixing matrix  $A$ 는 역행렬을 가져야 한다.
2. 각각의 오디오 신호는 공간적으로 상관관계가 없고, 시간에 대해 상관관계를 갖는 신호라 가정한다.

즉,  $E\{s_i(t) s_j(t - \tau)\} = r_{ij}(\tau)$ 으로 나타낼 수 있다. 이때, 공간적으로 상관관계가 없다는 것은  $i \neq j$ 일 때,  $r_{ij}(\tau) = 0$ 이 됨을 의미한다. 시간적으로 상관관계를 갖기 때문에  $i = j$ 일 때,  $r_{ij}(\tau)$ 은 0이 아닌 값을 가진다. 본 논문에서는 모든  $i, j$ 에 대하여  $E\{s_i(t) s_j(t - \tau)\}$ 을 벡터 행렬  $R_s(\tau)$ 로 표현하였다.  $R_s(\tau)$ 는

위의 성질에 따라 Diagonal matrix가 된다.

### 3. PCA 기반 신호 분리 알고리즘 구현

이 부분에서는 PCA 기반으로 신호를 분리하는 알고리즘을 소개한다. PCA는 선형 특징을 추출하는데 유용한 데이터 분석 방법이다. 본 논문에서는 2차원 오디오 신호를 PCA를 이용하여 분리하는 알고리즘을 소개한다.

#### 3.1 Whitening

Whitening이란 기존 상관관계가 있는 벡터들을 상관관계가 없는 새로운 벡터로 만드는 과정이다. 관측한 오디오 신호인  $X(t) = [x_1(t) \ x_2(t)]^T$ 의 공분산 행렬을 선형 변환을 통하여 Whitening된 오디오 신호 벡터  $V$ 를 만든다. 이때, 오디오 신호  $X$ 의 공분산 행렬의 고유값과 고유벡터를 이용한다. 여기서는 Eigenvalue Decomposition(EVD)를 이용하여 접근을 하였지만 보다 일반적인 알고리즘은 Singular Value Decomposition(SVD)를 이용하여 구현할 수 있다. 벡터  $X$ 의 공분산 행렬을 EVD한 결과는 다음과 같다.

$$R_X(0) = E\{X(t)X(t)^T\} = UDU^T \quad (2)$$

$U$ 는  $E\{X(t)X(t)^T\}$ 의 고유벡터 행렬(Orthogonal matrix)이고,  $D$ 는 고유값의 대각행렬이다. ( $D = \text{diag}\{\lambda_1, \lambda_2 \dots \lambda_n\}$ ) 관측된 오디오 신호 벡터  $X$ 를 다음과 같은 선형 변환  $W$ 을 거쳐 Whitening된 신호 벡터  $V$ 를 만든다.

$$W = D^{-1/2}U^T \quad (3)$$

$V = WX = D^{-1/2}U^T X$  이므로, 벡터  $V(t)$ 의 공분산 행렬은 다음과 같다.

$$\begin{aligned} R_V(0) &= E\{V(t)V(t)^T\} \\ &= E\{D^{-1/2}U^T X(t)X(t)^T U D^{-1/2}\} \\ &= D^{-1/2}U^T E\{X(t)X(t)^T\} U D^{-1/2} \\ &= D^{-1/2}U^T (UDU^T) U D^{-1/2} \\ &= I \end{aligned} \quad (4)$$

Whitening 작업을 통하여 공분산 행렬이 단위행렬이 됨을 알 수 있다. 즉, 두 신호 벡터의 상관관계를 없애는 것이 whitening의 목적이다.

#### 3.2 PCA Algorithm

Whitening을 한 후에 구해진 벡터  $V(t)$ 는

$$V(t) = WX(t) = WAS(t) = BS(t) \quad (5)$$

라고 쓸 수 있다. 이 때  $B = WA$ 는 orthogonal mixing matrix가 된다.

새로운 신호  $Y(t)$ 를  $V(t)$ 와  $V(t)$ 의 1 sample delay 된 신호를 더한 신호로 정의한다. 즉,  $Y(t) = V(t) + V(t-1)$  가 된다. 벡터  $Y(t)$ 의 공분산 행렬을  $R_Y(0)$ 라 하면

$$\begin{aligned} R_Y(0) &= E\{[V(t) + V(t-1)][V(t) + V(t-1)]^T\} \\ &= 2I + R_V(1) + R_V(1)^T \end{aligned} \quad (6)$$

이때  $R_Y(0)$ 은 symmetric하므로 이를 EVD하여 나타내면

$$R_Y(0) = U_2 D_2 U_2^T \quad (7)$$

$2I + R_S(1) + R_S(1)^T$ 가 diagonal matrix고  $B = WA$ 가 orthogonal matrix이기 때문에 (5), (6)에 의하여 다음이 성립한다.

$$R_Y(0) = B[2I + R_S(1) + R_S(1)^T]B^T \quad (8)$$

식 (7), (8)을 비교함으로써  $U_2 = B$ 가 성립함을 알 수 있다. 따라서 다음 식이 성립한다.

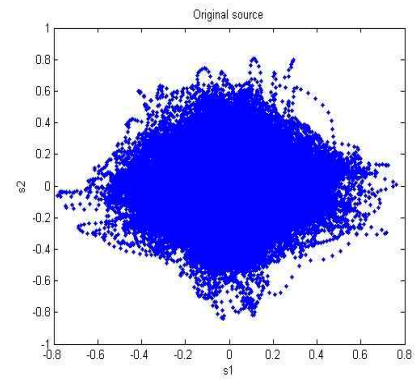
$$B = WA = U_2$$

$$\therefore A^{-1} = U_2^T W \quad (9)$$

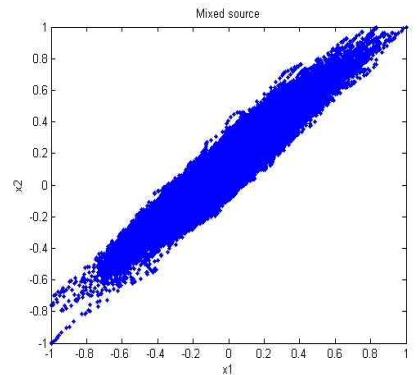
원 신호와 mixing된 관측 신호사이의 관계인  $X = AS$ 에서 가정1에 의해  $A^{-1}$ 이 존재하므로,  $S = A^{-1}X = U_2 D^{-1/2} U X$ 을 이용하여 원래의 오디오 신호를 복원할 수 있다.

### 4. MATLAB 구현 및 실험

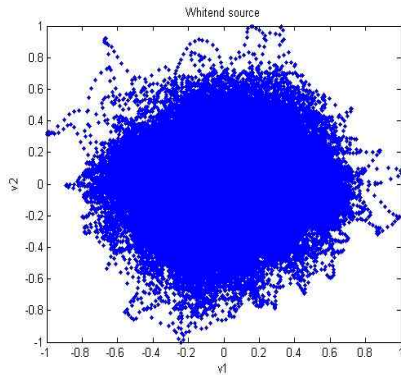
몇 개의 샘플 오디오 신호를 받아들이고 임의의 Mixing matrix를 이용하여 신호를 합성시킨다. 합성된 신호의 공분산 행렬을 EVD하여 Whitening을 한다. 다음은 기존 신호, 선형 조합된 관측 신호, whitening된 신호의 분포를 보여준다.



<그림 2. Original source>



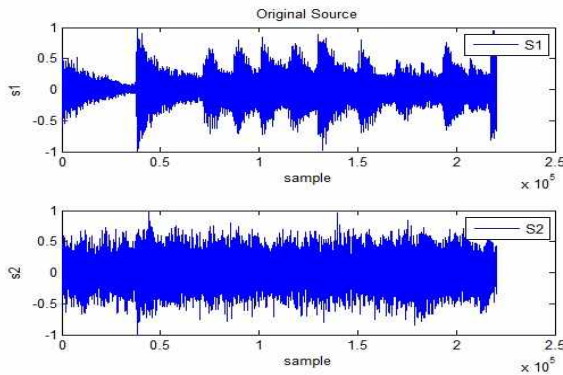
<그림 3. Mixed source>



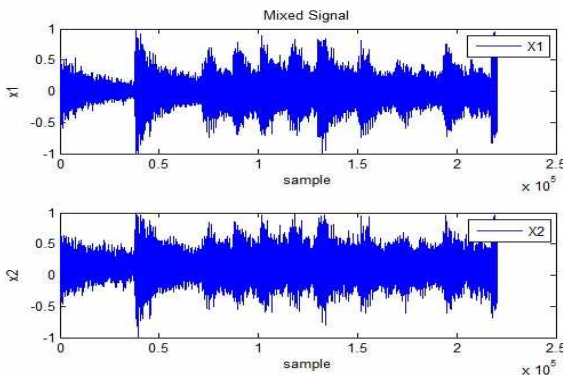
<그림 4. Whitened source>

그림에서 보듯이 임의의 Mixing matrix를 거쳐 선형 조합된 오디오 신호는 어느 정도의 상관관계를 보인다. 이를 선형변환  $W = D^{-1/2} E^T$ 를 통해 whitening을 하여  $V$  벡터를 구한 결과 Mixing 전의 기존 오디오 신호와 같이 상관관계가 없는 분포를 지니는 것을 알 수 있다.

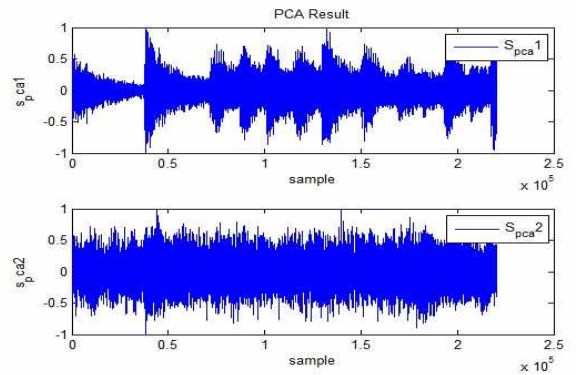
이후 PCA 알고리즘을 통하여  $A^{-1} = U_2^T W$  벡터를 구하고 관측된 오디오 신호인  $X(t)$ 와 곱하여  $S'(t)$ 를 추정한다.



<그림 5. Mixing matrix를 통과하지 않은 원본 오디오 신호>



<그림 6. Mixing matrix를 통과한 관측된 오디오 신호>



<그림 7. PCA 기반 신호 분리 알고리즘을 적용하여 분리된 오디오 신호>

그림 5,6,7은 원본 오디오 신호, 임의로 조합된 오디오 신호, PCA 기반 신호 분리 알고리즘을 적용하여 분리한 신호이다. 파형에서 보듯이 PCA 기반 신호 분리 알고리즘을 적용하여  $S(t)$ 와 유사한  $S'(t)$ 를 추정할 수 있었다.

기존 신호  $S(t)$ 와  $S'(t)$ 의 차이를 노이즈라 가정하고 Signal to Noise Ratio를 구하여 보았다.

$$Signal(t) = S(t)$$

$$Noise(t) = S(t) - S'(t) \quad (10)$$

Signal to Noise Ratio는 원본 오디오 신호  $S(t)$ 의 Power와 노이즈인  $S(t) - S'(t)$ 의 Power를 비교하여 구할 수 있다.

$$SNR = 10 \log \left( \frac{P_{signal}}{P_{noise}} \right) \quad (11)$$

본 실험에서는 임의의 음원 4개를 대상으로 그 중 두 신호에 대해 SNR을 각각 구하였다.

	$SNR_1$	$SNR_2$
$X_1, X_2$	17.9707	15.3016
$X_1, X_3$	18.2927	18.8468
$X_1, X_4$	7.2000	7.6444
$X_2, X_3$	25.6156	27.0254
$X_2, X_4$	30.5314	29.0495
$X_3, X_4$	7.1871	4.7459

<표 1. 각각의 신호와 분리된 신호와의 SNR>

평균적으로 SNR은 17.4509으로 약 원래 신호의 Power의 1/55정도의 노이즈 값을 가지는 것을 알 수 있다.

## 5. 결과

본 논문에 제시된 PCA 알고리즘 외에도 Mutual Information을 최소화하고, Non-Gaussianity를 최대화하는 방법인 ICA, 데이터가 효율적으로 분포되는 축을 찾는 LDA와 같은 알고리즘이 사용될 수 있다. 하지만 Independency에 대한 정의가 명확하고 구현상 간단함을 이점으로 PCA 알고리즘을 적용했다. PCA 기반으로 오디오 신호 분리 알고리즘을 구현한 결과, 통계적 특성을 이용하여 임의로 믹싱된 오디오

오 신호를 분리해 낼 수 있었다. 기존의 음원과 비교하여 SNR을 측정 하였을 때, 상당히 낮은 노이즈 수준을 보였다. 본 논문에서 다루진 않았지만 향후 ICA와 LDA등으로 SNR을 구한 후 이 중 가장 최적의 성능을 갖는 방법으로 data를 분리해 내는 방안역시 가능할 것으로 기대 된다.

## 6. 감사의 글

"본 연구는 미래창조과학부 및 정보통신산업진흥원의 대학 IT연구센터 육성지원 사업의 연구결과로 수행되었음"  
(NIPA-2013-H0301-13-1011)

## 7. 참조

- [1] Jean-Francois Cardoso *et al* "Blind signal separation: statistical principles" *Proc. IEEE*, vol. 9, no.10, pp. 2009-2025, Oct, 1998.
- [2] Tao Li and Nicholas D. Sidiropoulos "Blind Digital Signal Separation Using Successive Interference Cancellation Iterative Least Squares" *IEEE Transactions on Signal Processing*, vol.48, no.11, November, 2000.
- [3] Ganesh R. Naik "Introduction: Independent Component Analysis" in *Independent Component Analysis for Audio and Biosignal Applications*, InTech, October, 2012, pp.1-18.
- [4] Anil Lal and Wenwu Wang "Monaural Audio Separation Using Spectral Template and Isolated Note Information" in *Independent Component Analysis for Audio and Biosignal Applications*, InTech, October, 2012, pp.67-69.
- [5] Fernando I. Mato-Mendez and Manuel A. Sobreira-Seoane "Blind Implicit Source Separation - A New Concept in BSS Theory." in *Independent Component Analysis for Audio and Biosignal Applications*, InTech, October, 2012, pp.321-339.
- [6] A. Hyvarinen, E. Oja. "Independent Component Analysis: Algorithms and Applications," *Neural Networks*, 13(4-5):411-430, 2000.
- [7] 김민제 *et al.* "블라인드 방식의 리듬 음원 분리," *한국음향학회지* vol.28, no.8, 2009, pp.697-705.
- [8] 이해경 *et al* "주성분 분석을 이용한 블라인드 신호 분리" *한국정보과학회 가을학술발표논문집* vol.28, no.2, 2001.