

특집논문 (Special Paper)

방송공학회논문지 제17권 제6호, 2012년 11월 (JBE Vol. 17, No. 6, November 2012)

http://dx.doi.org/10.5909/JBE.2012.17.6.924

ISSN 1226-7953(Print)

## 오디오 업믹스를 위한 효율적인 주성분-주변성분 분리 알고리즘

백 용 현<sup>a)‡</sup>, 전 세 운<sup>b)</sup>, 이 석 필<sup>c)</sup>, 박 영 철<sup>a)</sup>

### Efficient Primary-Ambient Decomposition Algorithm for Audio Upmix

Yong-Hyun Baek<sup>a)‡</sup>, Se-Woon Jeon<sup>b)</sup>, Seok-pil Lee<sup>c)</sup>, and Young-Cheol Park<sup>a)</sup>

#### 요 약

스테레오 업믹스(Upmix)에서 음원을 주성분(Primary)과 주변성분(Ambient)으로 분리하는 것은 주된 전처리 과정이며 주성분 분석법(Principal Component Analysis - PCA)을 이용한 연구가 진행되고 있다. 그러나 주성분 분석법은 분리 성능이 스테레오 음원이 가지는 주성분과 주변성분의 파워비(Primary Ambient Power Ratio - PAR Ratio) 및 주성분의 페닝 각도에 영향을 받는다는 단점이 있다. 이전 연구에 따르면 PAR에 따른 단점을 극복하기 위한 변형된 주성분 분석법(Modified PCA) 방법이 제안되었으나 여전히 페닝 각도에 대한 단점은 극복하지 못하였다. 본 논문에서는 PAR 및 페닝 각도에 영향을 받지 않는 새로운 주성분 분석법 기반의 알고리즘을 제안하였다. 제안된 알고리즘은 스테레오 음원의 파워를 보존하는 기준을 두고 고유치의 비를 이용한 적절한 스케일 값을 통해 문제를 해결하였다. 제안된 알고리즘은 실험결과 PAR 및 주성분의 페닝 각도에 영향을 받지 않고 정확한 분리 성능을 보여줌을 확인하였다.

#### Abstract

Decomposition of a stereo signal into the primary and ambient components is a key step to the stereo upmix and it is often based on the principal component analysis (PCA). However, major shortcoming of the PCA-based method is that accuracy of the decomposed components is dependent on both the primary-to-ambient power ratio (PAR) and the panning angle. Previously, a modified PCA was suggested to solve the PAR-dependent problem. However, its performance is still dependent on the panning angle of the primary signal. In this paper, we proposed a new PCA-based primary-ambient decomposition algorithm whose performance is not affected by the PAR as well as the panning angle. The proposed algorithm finds scale factors based on a criterion that is set to preserve the powers of the mixed components, so that the original primary and ambient powers are correctly retrieved. Simulation results are presented to show the effectiveness of the proposed algorithm.

Keyword : PCA, Panning Angle, Correlation, Eigen Value Ratio

a) 연세대학교 전산학과(Computer Science, Yonsei Univ.)  
b) 연세대학교 전기전자공학부(Electrical & Electronic Eng., Yonsei Univ.)  
c) 상명대학교(Sangmyung Univ.)  
‡ Corresponding Author : 백용현 (Yong-Hyun Baek)  
E-mail: hyen0330@yonsei.ac.kr  
Tel: +82-33-760-2756

Manuscript received September 3, 2012 Revised October 17, 2012  
Accepted October 17, 2012

## 1. 서론

현재 3차원 영상과 더불어 오디오 또한 실감 음원 청취에 대한 사람들의 욕구가 증가하면서 각 가정에 돌비 서라운드(Dolby Surround)<sup>[1]</sup>와 같은 다채널 재생 시스템이 널리

보급 되고 있다. 그러나 현재 대부분의 오디오 콘텐츠들은 여전히 스테레오로 서비스 되는 것이 대부분이며 이것을 다채널 환경에서 재생하기 위하여 채널 포맷 변환 및 분배 기술인 업믹스 기술을 사용한다. 업믹스 기술은 크게 두 단계 과정을 거치는데 첫 번째 단계로 스테레오에서 주성분과 주변성분을 분리하고 두 번째 단계로 분리된 음원을 다채널 스피커 포맷에 맞게 분배하는 과정이다. 주성분은 스테레오 음원에서 패닝 기법<sup>[2]</sup>으로 믹싱되거나 목소리 등의 방향성이 있는 패닝 음원을 말하고 주변성분은 그 외 패닝 음원이 벽에 반사되어 청취자의 귀에 전달되는 잔향이나 관객의 합성소리 같은 배경음을 말한다. 본 논문에서는 업믹스 기술의 첫 번째 단계인 주성분-주변성분을 분리하는 효과적인 방법을 제안한다.

기존의 주성분-주변성분 분리 방법으로 다양한 방법들이 제안되었다<sup>[3-5]</sup>. 이들 중 본 논문에서는 주성분 분석법을 이용한 분리 방법에 대하여 고찰하고 주성분 분석법이 가지는 문제점을 완화한 향상된 주성분 분석법을 제안한다. 주성분 분석법의 문제점으로는 크게 두 가지로 첫 번째로 분리해낸 주성분에 잔여 주변성분이 남아 있다는 점, 두 번째로 입력 스테레오 채널에 주성분이 되는 주성분이 존재하지 않는 경우 에너지가 큰 주변성분이 주성분으로 분리된다는 점이다. 이러한 문제를 해결하기 위하여 Modified PCA 알고리즘이 제안되었다<sup>[6]</sup>. 이는 주성분 같은 음원이 단지 크기의 차이만 두고 스테레오 두 채널에 존재하여 상관관계(correlation)가 높고 반대로 주변성분은 여러 경로를 거쳐 반사된 음원이 두 채널에 독립적으로 존재하여 상관관계가 매우 낮은 특성을 이용하여 주성분이 존재하지 않는 경우를 정규화 된 상관 계수(Normalized Correlation Coefficient)를 이용하여 판단한다. 그러나 이러한 상관 계수가 두 채널에 주변성분만 있을 경우 뿐 만 아니라 한쪽 채널로 극단적으로 패닝 된 음원의 경우에도 상관 계수가 작아지기 때문에 주성분과 주변성분을 분리하지 못하게 되는 문제가 있다.

이를 해결하기 위하여 본 논문에서는 상관 계수가 아닌 두 채널간의 고유치값의 비를 이용하여 주성분과 주변성분을 분리한다. 고유치값의 비는 패닝 각도에 영향을 받지 않기 때문에 상호 상관도를 이용할 때의 문제점을 극복할 수

있다. 또한 주변성분을 각 채널별로 상관성이 없는 성분을 두 채널 분리할 수 있으며 주성분에 섞인 잔여 주변성분 역시 제거하였다.

본 논문의 구성은 다음과 같다. 2장에서는 스테레오 신호 모델 및 기존의 주성분 분석법을 분리 알고리즘들에 대해 설명하고 3장에서 제안된 알고리즘을, 4장에서는 모의실험 결과를 보이고 5장에서 본 논문의 결론을 맺는다.

## II. PCA 기반의 음원 분리 알고리즘

### 1. 스테레오 신호 모델

업믹스 기술에서 입력 스테레오 신호를 분석할 때 단시간 푸리에 변환(short-time Fourier transform, STFT)을 이용하여 분석한다. 이렇게 주파수 도메인으로 변환된 계수들을 사람의 청각 특성에 따라 각각 서브밴드별로 묶어 나누어 처리한다<sup>[7]</sup>. 이후의 수식들은 주파수 영역에서의 표현들이며 수식의 간편함을 위해 시간 프레임 인덱스와 주파수 대역 인덱스는 생략한다. 입력 스테레오 신호는 음원이 패닝 기법에 의하여 스테레오 스피커 포맷에 따라  $-30^\circ$ 에서  $30^\circ$  사이에 레벨 패닝 되어 정위 되었다고 가정하고 여기에 각각의 상관도가 없는 주변성분이 더해진 형태로 모델링할 수 있다. 모델링 된 입력 두 채널의 신호는 수식 (1)과 같이 표현 할 수 있다.

$$\begin{aligned} X_L &= a_L S + N_L \\ X_R &= a_R S + N_R \end{aligned} \quad (1)$$

여기서  $a_L$ 과  $a_R$ 은 패닝 계수,  $S$ 는 패닝 성분이 되고  $N_L$ 과  $N_R$ 은 각 채널에 더해진 주변성분이다. 레벨 패닝 기법에 의해 패닝 이득은 수식 (2)와 같은 관계를 가지며 각  $S, N_L, N_R$ 은 수식 (3)과 같은 관계를 갖는다고 가정한다.

$$a_L^2 + a_R^2 = 1 \quad (2)$$

$$E\{S^H N_L\} = E\{S^H N_R\} = E\{N_L^H N_R\} = 0 \quad (3)$$

여기서  $E\{\cdot\}$ 는 평균을 나타내는 오퍼레이터이고 구현 시 1차의 IIR 필터를 이용하였다. 또한  $H$ 는 허미션 연산을 나타낸다. 따라서 수식 (3)과 같이 입력 패닝 성분과 주변성분, 또한 각 채널 주변성분간에는 상관성이 없고, 수식 (4)에서 처럼 두 채널에 더해지는 주변성분들의 에너지는 같다고 가정한다.

$$E\{N_L^H N_L\} = E\{N_R^H N_R\} \tag{4}$$

### 2. 주성분 분석법을 이용한 음원 분리 방법

주성분 분석법을 이용하여 음원을 분리 할 때 먼저 패닝 성분에 곱해진 패닝 이득을 구해야 한다. 패닝 이득은 입력 스테레오 채널의 2x2 공분산 행렬로부터 추정 할 수 있고 공분산 행렬은 아래 식과 같이 계산된다.

$$R = \begin{bmatrix} r_{LL} & r_{LR} \\ r_{RL} & r_{RR} \end{bmatrix}, r_{ij} = E\{X_i^H X_j\} \tag{5}$$

공분산 행렬  $R$ 을 고유치 분해하여 다시 표현 하면 다음과 같다.

$$R = [V_1 \ V_2] \begin{bmatrix} \lambda_1 & \\ & \lambda_2 \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix}, \lambda_1 \geq \lambda_2 \tag{6}$$

여기서  $\lambda_{1,2}$ 는 고유치 값이고  $V_{1,2}$ 는 각 고유치 값에 대응하는 고유치 벡터가 된다. 이때 두 고유치 값 중 큰 값을 가지는 고유치 값은 주성분에 해당하는 성분의 에너지가 되며 이에 대응하는 고유치 벡터의 방향이 패닝 각도가 되므로 따라서 추정된 패닝 이득은 수식 (7)과 같다<sup>[8]</sup>.

$$\hat{a}_L = v_{1,1}, \hat{a}_R = v_{1,2} \tag{7}$$

추정된 패닝 계수으로부터 주성분과 주변성분은 다음과 같은 식으로 구할 수 있다.

$$\hat{S} = \hat{a}_L X_L + \hat{a}_R X_R \tag{8}$$

$$\begin{aligned} \hat{N}_L &= X_L - \hat{a}_L \hat{S}, \\ \hat{N}_R &= X_R - \hat{a}_R \hat{S} \end{aligned} \tag{9}$$

### 3. 변형된 주성분 분석법을 이용한 음원 분리 방법

3.1 에서 설명한 기존의 주성분 분석법 방법으로 음원을 분리할 경우 서론에서 언급했듯이 몇 가지 문제점이 발생한다. 먼저 수식 (8)을 본 논문에서 가정한 입력 신호 모델링 수식 (1)을 적용하여 다시 쓰면 다음과 같다.

$$\begin{aligned} \hat{S} &= \hat{a}_L X_L + \hat{a}_R X_R \\ &= \hat{a}_L (a_L S + N_L) + \hat{a}_R (a_R S + N_R) \\ &= S + \hat{a}_L N_L + \hat{a}_R N_R \end{aligned} \tag{10}$$

따라서 주성분에 잔여 주변성분이 여전히 남아 있는 채로 분리되게 된다. 또한 주성분 분석법에서 가장 큰 고유치 값을 가지는 방향을 주성분으로 가정하는데 만약 입력 스테레오 채널에서 이러한 패닝 된 주성분이 존재하지 않는 경우 에너지가 큰 주변성분을 주성분으로 분리하는 문제가 발생한다. 이 경우 정확하게 분리 하자면 두 입력 채널이 그대로 주변성분으로 분리 되어야 된다.

이러한 문제를 해결하기 위하여 변형된 주성분 분석법 알고리즘이 제안되었다. 변형된 주성분 분석법 알고리즘은 다음과 같은 식을 이용하여 주성분과 주변성분을 분리한다.

$$\begin{aligned} \tilde{N}_L &= |\Phi| \hat{N}_L + (1 - |\Phi|) X_L \\ \tilde{N}_R &= |\Phi| \hat{N}_R + (1 - |\Phi|) X_R \end{aligned} \tag{11}$$

여기서  $|\Phi|$  은 스테레오 두 채널간의 정규화 된 상관 계수를 나타낸다. 따라서 패닝 성분이 존재 할 경우 상관 계수는 1의 가까운 값을 가지게 되어 기존의 주성분 분석법을 그대로 사용하게 되고 패닝 성분이 존재 하지 않는 경우 상관 계수가 0에 가까운 값을 가지게 되어 입력 신호를 그대로 주변성분으로 분리하게 된다.

패닝 된 주성분 또한 상관 계수를 아래 식을 이용하여 분리한다.

$$\tilde{S} = |\Phi| \hat{S} \tag{12}$$

### III. 제안된 음원 분리 알고리즘

#### 1. 주변성분 신호 분리

앞장에서 설명한 변형된 주성분 분석법에서 사용한 상관 계수는 페닝 신호가 극단적으로 페닝 될 경우 페닝 신호가 존재함에도 불구하고 역시 0에 가까워지게 된다. 따라서 이 경우 주성분과 주변성분 신호를 분리 하지 못하고 입력 신호가 그대로 주변성분으로 빠져 나오게 된다. 그 결과 대부분의 페닝 신호가 주변성분으로 가게 되어 업믹스시 페닝 신호의 음상이 후방에 위치하게 되는 문제가 발생하게 된다.

제안된 알고리즘은 앞서 3-1에서 추정된 주성분과 주변성분으로부터 각각의 파워가 입력 스테레오의 실제 파워가 유지되도록 하는 스케일 값을 이용하여 분리한다. 먼저 주변성분 신호의 분리는 다음 수식 (13)을 이용하여 분리한다.

$$\begin{aligned} \overline{N}_L &= X_L - \Gamma_L \hat{a}_L \hat{S}, \\ \overline{N}_R &= X_R - \Gamma_R \hat{a}_R \hat{S} \end{aligned} \quad (13)$$

여기서  $\Gamma_L$ 과  $\Gamma_R$ 은 각각 왼쪽과 오른쪽의 스케일 값이다. 스케일 값을 구하기 위하여 추정된 주변성분의 파워가 실제 주변성분의 파워와 같도록 조건(criterion)은 다음과 같다.

$$\sigma_{\overline{N}_L}^2 = \sigma_{\overline{N}_R}^2 = \sigma_N^2 \quad (14)$$

수식 (13)의 왼쪽 채널의 대한 주변성분에 대한 식을 수식 (1)을 이용하여 주성분과 주변성분의 항으로 정리하면 수식 (15)와 같고 이 때 추정된 페닝 계수는 근사화 되기에 충분히 정확하게 구해졌다고 가정한다.

$$\overline{N}_L = (1 - \Gamma_L) a_L S + (1 - \Gamma_L a_L^2) N_L - \Gamma_L a_L a_R N_R \quad (15)$$

수식 (15)로부터 주변성분의 파워를 구하면 수식 (16)과 같다.

$$\begin{aligned} \sigma_{\overline{N}_L}^2 &= (1 - \Gamma_L)^2 a_L^2 \sigma_S^2 + (1 - \Gamma_L a_L^2)^2 \sigma_{N_L}^2 + \Gamma_L^2 a_L^2 a_R^2 \sigma_{N_R}^2 \\ &= (1 - \Gamma_L)^2 a_L^2 \sigma_S^2 + (1 - 2\Gamma_L a_L^2 + \Gamma_L^2 a_L^2) \sigma_{N_L}^2 \end{aligned} \quad (16)$$

여기서 조건을 이용하여 스케일 값에 대한 식으로 정리하면 다음과 같다.

$$(1 - \Gamma_L)^2 \sigma_S^2 + (\Gamma_L^2 - 2\Gamma_L) \sigma_{N_L}^2 = 0 \quad (17)$$

따라서 스케일 값  $\Gamma_L$ 은 다음과 같다.

$$\Gamma_L = 1 - \sqrt{\frac{\sigma_{N_L}^2}{\sigma_S^2 + \sigma_{N_L}^2}} \quad (18)$$

오른쪽 채널의 주변 성분에 대해서도 동일한 방법으로 구하면 두 채널의 주변 성분의 파워가 같다고 가정하였기 때문에 수식 (18)과 동일하다. 또한 여기서 수식 (18)의 분모와 분자는 수식 (6)의 두 고유치 값에 대응된다<sup>[8]</sup>. 따라서 수식 (13)을 제안된 알고리즘의 스케일 값을 이용하여 다시 정리하면 다음과 같은 식을 얻을 수 있다.

$$\begin{aligned} \overline{N}_L &= X_L - \left(1 - \sqrt{\frac{\lambda_2}{\lambda_1}}\right) \hat{a}_L \hat{S}, \\ \overline{N}_R &= X_R - \left(1 - \sqrt{\frac{\lambda_2}{\lambda_1}}\right) \hat{a}_R \hat{S} \end{aligned} \quad (19)$$

수식 (19)에서 보듯 스케일 값은 고유치 값의 비로 나타남을 알 수 있다. 만약 입력 스테레오 신호에 페닝 신호가 존재 하지 않는 경우 고유치 값의 비는 1에 가까운 값을 가지게 되므로 입력 신호를 그대로 통과 시키게 되며 페닝 신호가 존재 할 경우는 주성분과 주변 성분의 에너지의 비로써 두 신호를 분리 하게 된다. 즉, 페닝 각도가 변한다 하더라도 입력 스테레오 채널에 존재하는 주성분과 주변성분의 에너지는 변하지 않기 때문에 고유치 비는 일정하게 유지된다. 따라서 제안된 알고리즘은 주성분과 주변 성분의 변하지 않는 에너지의 비로써 분리하게 되기 때문에 페닝 각도에 강인한 분리 성능을 가진다.

다음 그림 1은 페닝 각도에 따른 상관 계수와 고유치 비

를 나타낸 것이다. 상관 계수는 패닝 각도가 극단적으로 갈수록 값이 떨어지며 변화하나 고유치 비는 일정하게 유지됨을 볼 수 있다.

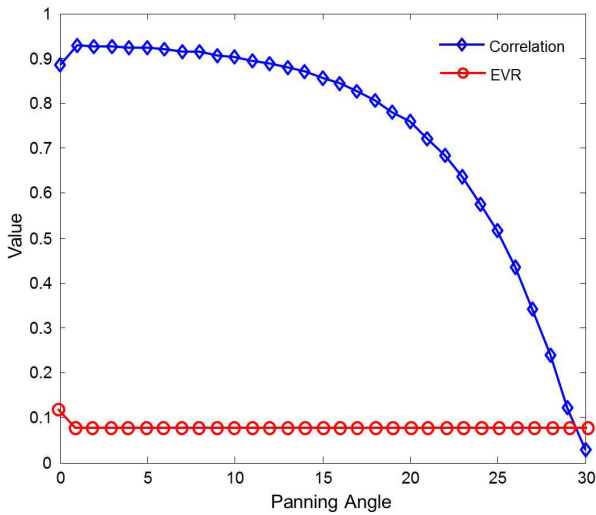


그림 1. 패닝 각도에 따른 상관 계수와 고유치의 비  
Fig 1. Normalized correlation coefficient and eigenvalue ratio according to the panning angle

## 2. 주성분 분리

주성분에 대한 스케일 값 역시 주변성분을 분리할 때와 동일하게 실제 주성분의 파워가 유지되도록 하는 조건을 이용하여 분리한다. 주성분에 분리하기 위한 조건은 다음과 같이 주어진다.

$$\sigma_s^2 = \sigma_s^2 \tag{20}$$

제안된 알고리즘에서 추정된 주성분 식은 다음과 같다.

$$\bar{S} = \Gamma_S \hat{S} \tag{21}$$

수식 (21)로부터 주성분의 파워를 구하면,

$$\sigma_s^2 = \Gamma_S^2 (\sigma_s^2 + \sigma_N^2) \tag{22}$$

수식 (20)을 이용하여 (22)를 스케일 값에 대해 정리하면 다음과 같다.

$$\Gamma_S = \sqrt{\frac{\sigma_s^2}{\sigma_s^2 + \sigma_N^2}} = \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1}} \tag{23}$$

스케일 값을 이용하여 다시 수식 (21)을 정리하면 최종적으로 주성분을 분리하는 식을 얻을 수 있다.

$$\bar{S} = \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1}} \hat{S} \tag{24}$$

따라서 잔여 주변성분이 섞여 있는 추정된 패닝 신호로부터 실제 패닝 신호 에너지에 해당하는 성분만을 추출함으로써 실제 패닝 신호를 더 정확하게 분리하여 추정할 수 있다.

## IV. 모의 실험

모의 실험은 기존의 주성분 분석법, 변형된 주성분 분석법 그리고 제안된 알고리즘의 분리 성능을 비교하였다. 신호는 사인(sine) 원도우와 오버랩(overlap add)을 적용한 단 구간 푸리에 변환을 이용하여 프레임 단위로 분석하며 여기서 샘플링 주파수 44.1Khz에 프레임 크기는 1024를 사용하였다. FFT 이후 주파수 영역에서 사람의 청각적 특성을 반영한 임계 대역 (critical band)을 이용하여 각 대역별로 주성분과 주변성분을 분리한 뒤 합성한다. 입력은 패닝 된 음악신호에 이상적인 실험을 위해 상관관계가 없는 백색잡음을 주변성분으로 더한 입력을 사용하였다.

그림 2와 3은 PAR 10dB에서 제안된 알고리즘이 실제 입력 신호의 에너지를 그대로 유지하는지 알아보기 위해 패닝 각도를 달리 하여 각 알고리즘으로부터 추정된 주변성분파워를 실제 파워와 비교한 그림이다. 그림 2에서 보듯 0도에서는 세 알고리즘이 실제 파워와 유사하나 제안된 알고리즘이 더 정확하게 주변성분의 파워를 추정함을 볼 수 있다.

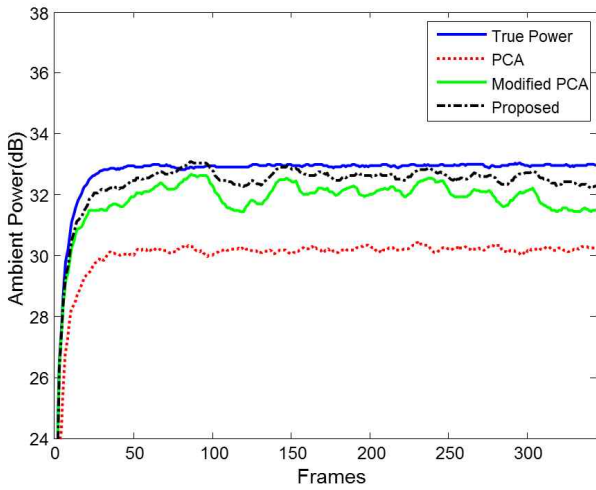


그림 2. 추정된 주변성분 파워 비교 (PAR=10dB,  $\theta=0^\circ$ )  
 Fig 2. Comparison of the estimated ambient powers (PAR=10dB,  $\theta=0^\circ$ )

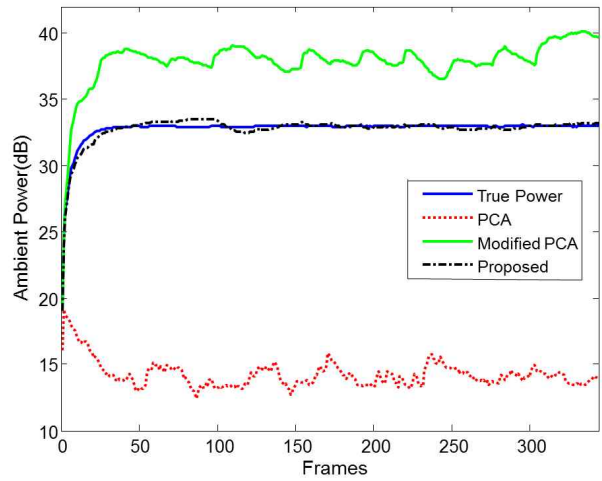


그림 3. 추정된 주변성분 파워 비교 (PAR=10dB,  $\theta=20^\circ$ )  
 Fig 3. Comparison of the estimated ambient powers (PAR=10dB,  $\theta=20^\circ$ )

그림 3과 같이 신호의 파워가 극단적으로 패닝 된 경우, 기존의 주성분 분석법은 패닝 된 방향으로의 패닝 이득이 1에 가까워져 수식 (8)과 (9)에 의해 추정된 패닝 신호가 입력 신호와 같아져서 급격히 작아지고 변형된 주성분 분석법은 극단적으로 패닝 된 경우 입력 패닝 신호를 그대로

주변성분로 보내기 때문에 더 커짐을 알 수 있다. 반면 제안된 알고리즘은 극단적으로 패닝 된 경우에도 실제 주변 성분의 파워를 유지하여 분리함을 알 수 있다.

그림 4와 5는 패닝 각도  $30^\circ$ 에서 변형된 주성분 분석법과 제안된 알고리즘으로 추정한 주변성분을 나타낸다. 여기서

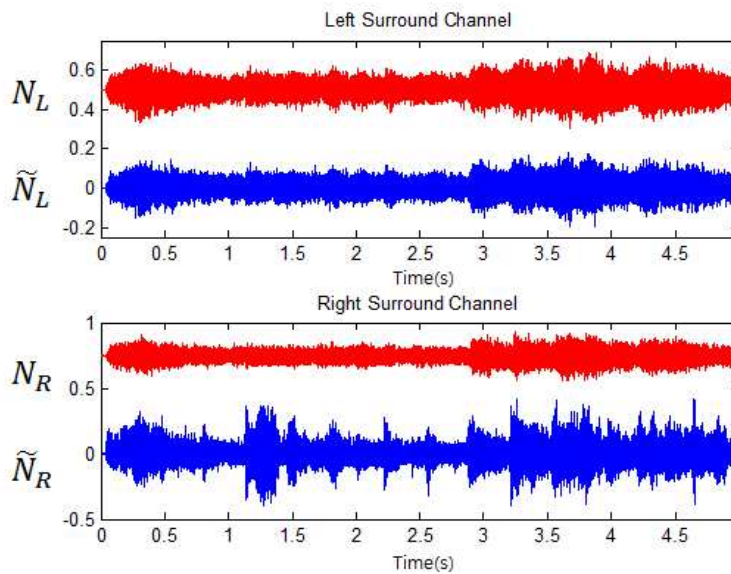


그림 4 추정된 주변성분 (변형된 주성분 분석법,  $\theta=30^\circ$ )  
 Fig 4 Estimated Ambient (변형된 주성분 분석법,  $\theta=30^\circ$ )

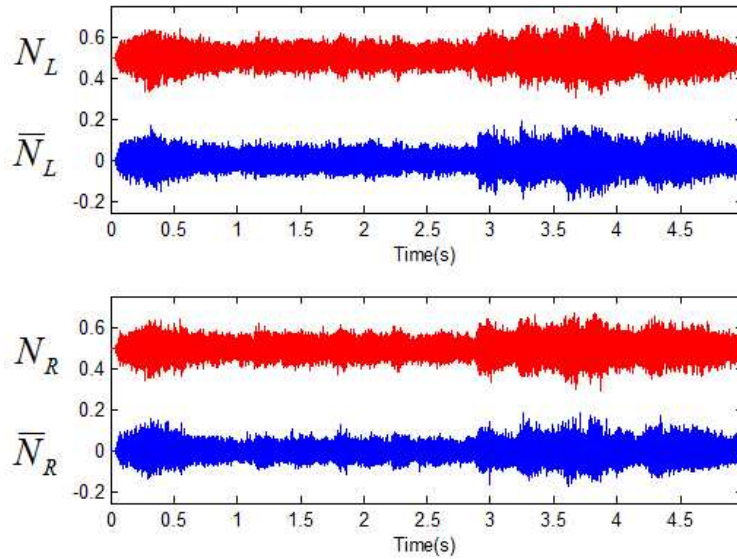


그림 5. 추정된 주변성분 (제안된 알고리즘  $\Theta=30^\circ$ )  
 Fig 5. Estimated Ambient (Proposed Algorithm,  $\Theta=30^\circ$ )

입력은 하나의 모노 음원을 각도를 바꾸어 가면서 패닝 시키고, 모델링된 룸 임펄스 응답을 이용하여 주변성분을 만들어 각 스테레오 왼쪽과 오른쪽 채널에 더해준 스테레오를 사용하였다. 그림 4에서 보듯 변형된 주성분 분석법은 상관 계수가 극단적으로 패닝 될수록 0에 가까워져 수식

(11)에 따라 입력 신호 그 자체가 주변성분으로 나오게 되기 때문에 제대로 분리 하지 못하는 반면 그림 5의 제안된 알고리즘은 각도가 극단적으로 패닝 되었다 하더라도 실제 주변성분을 정확하게 추정하여 분리 하고 있음을 보여 준다. 이는 각 채널에서 고유치 비를 이용하여 주성분과 주변

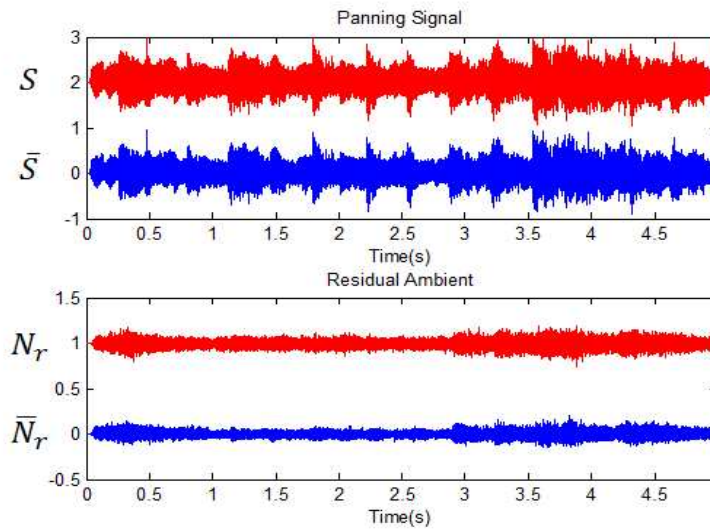


그림 6. 추정된 패닝 신호와 잔여 주변성분(제안된 알고리즘)  
 Fig 6 Estimated panning signal and residual ambient signal (Proposed Algorithm)

성분의 에너지의 비를 구할 수 있고 에너지는 패닝 각도에 영향을 받지 않고 일정 하므로 주변성분으로는 스테레오 신호에서 주변성분이 가지는 에너지만큼 분리되어 나오게 된다.

그림 6은 제안된 알고리즘의 패닝 신호를 추정된 결과이다. 앞서 언급한 대로 패닝 신호에 주변성분이 함께 존재하기 때문에 스케일 값은 이러한 패닝 신호의 섞인 주변성분을 제거하게 된다. 그림 6에 상단은 실제 패닝 된 주성분과 추정된 주성분을 비교한 것으로 정확하게 패닝 된 주성분이 분리되어 나옴을 볼 수 있다. 하단은 실제 주변성분을 이용하여 잔여 성분을 구한 것과 제안된 알고리즘에서 추정된 주변성분으로 잔여 주변성분을 구한 것이다. 즉, 여기서  $\bar{N}r = \hat{a}_L \tilde{N}_L + \hat{a}_R \tilde{N}_R$  을 의미한다.

그림 6에서 추정된 주변 성분으로부터 패닝 된 주성분에 섞인 잔여 주변성분을 제거하여 더 정확하게 주성분을 분리하고 있음을 보여준다.

## V. 결론

본 논문에서는 주성분 분석법을 이용한 향상된 주성분-주변성분 분리 알고리즘을 제안하였다. 기존 변형된 주성분 분석 알고리즘이 상관 계수를 사용하여 패닝 각도에 따라 음원 분리 성능이 저하되나 제안한 알고리즘은 두 입력 채널의 주성분과 주변성분의 에너지에 기반 한 공분산 행렬의 고유치 비를 이용함으로써 패닝 성분이 존재하지 않고 주변성분만 있을 경우는 물론 패닝각도에도 상관없이 음원 분리 성능이 강건하게 유지된다. 또한 패닝 신호에서

잔여 주변성분을 제거하여 더 정확한 실제 패닝 신호를 분리해 낼 수 있다. 향후 제안 된 알고리즘을 다채널 환경에 맞추어 재분배 할 경우 입력 스테레오 신호의 에너지를 그대로 유지 할 뿐만 아니라 패닝 신호의 음상이 후방으로 빠지는 경우도 방지 할 수 있을 것으로 보인다. 또한 분리해 낸 주변성분간의 상관도 역시 실제 주변성분의 상관도와 유사하게 낮은 상관도를 가지는 신호로 분리할 수 있기 때문에 청취자의 공간감(Envelopment)을 높일 수 있을 것으로 보인다.

## 참 고 문 헌

- [1] R. Dressler, "Dolby Surround Prologic II de-coder principles of operation," Tech. Rep. Dolby Laboratories, 2000
- [2] Pulkki, Ville. "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", AES convention. vol 45, Issue 6, pp 456-466, 1997
- [3] J. Usher and J. Benesty, "Enhancement of spatial sound quality: A new reverberation-extraction audio upmixer," IEEE Trans. Audio, Speech, and Language Processing, vol. 15, no. 7, pp. 2141-2150, 2007
- [4] C. Faller, "Multiple-loudspeaker playback of stereo signals," J. AES, vol. 54, no.11, pp. 1051-1064, 2006.
- [5] S.W. Jeon, Y.C. Park, S.P. Lee, D.H. Yoon, "Robust Representation of Spatial Sound in Stereo-to Multichannel Upmix", AES convention, 2010
- [6] Goodwin. M.M., "Geometric Signal Decompositions for Spatial Audio Enhancement", IEEE ICASSP conf, pp 409-412, 2008
- [7] E. Zwicker, E. Terhardt, "Analytical Expressions for Critical-Band rate and Critical Bandwidth as a Function of Frequency", ASOA, Vol 68, pp 1523-1525, 1980
- [8] Briand, Manuel, Martin, Nadine, Virette, David, "Parametric Representation of Multichannel Audio Based on Principal Component Analysis", AES convention, 2006



---

저 자 소 개

---



백 용 현

- 2009년 : 연세대학교 정보처리기술학부 학사 졸업
- 2009년 ~ 현재 : 연세대학교 전산학과 석박사 통합과정
- 주관심분야 : 음성/오디오 신호처리, 3D 오디오, 디지털 신호처리



전 세 운

- 2007년 : 연세대학교 전기전자공학사 졸업
- 2007년 ~ 현재 : 연세대학교 전기전자공학과 석박사 통합과정
- 주관심분야 : 오디오 신호처리, 3D 오디오, 디지털 신호처리



이 석 필

- 1990년 : 연세대학교 전기공학과 학사
- 1992년 : 연세대학교 전기공학과 석사
- 1997년 : 연세대학교 전기전자공학과 박사
- 2002년 ~ 현재 : KETI 디지털미디어 연구센터 센터장
- 2012년 ~ 현재 : 상명대학교 교수
- 주관심분야 : 오디오 신호처리, 디지털 신호처리, 방송통신 시스템, 멀티미디어 검색



박 영 철

- 1986년 : 연세대학교 전자공학과 학사 졸업
- 1988년 : 연세대학교 전자공학과 석사 졸업
- 1993년 : 연세대학교 전자공학과 박사 졸업
- 2002년 ~ 현재 : 연세대학교 컴퓨터 정보통신공학부 교수
- 주관심분야 : 음성/오디오 신호처리, 3D 오디오, 디지털 신호처리, 적응 신호처리