

반위상 주요성분에 기반을 둔 개선된 음수 채널간 상관도 파라미터 합성 기법

Improved Synthesis Method of Negative Inter-channel Correlation Parameter Based on Anti-phase Primary Component

현동일 · 이석필* · 박영철** · 윤대희

(Dong-il Hyun, Seok-Pil Lee*, Young-cheol Park**, and Dae Hee Youn)

연세대학교 전기전자공학과, *전자부품연구원, **연세대학교 컴퓨터정보통신공학부
(접수일자: 2012년 2월 29일; 수정일자: 2012년 5월 9일; 채택일자: 2012년 7월 11일)

초 록: 파라메트릭 스테레오와 MPEG 서라운드 는 대표적인 공간 오디오 부호화 기법이다. 본 논문에서는 기존의 공간 오디오 부호화 기법에서 채널간 상관도 합성의 문제점을 분석하였다. 기존의 방법에서는 출력신호에 합성되는 주요성분들은 서로 동위상 관계에 있는 반면 잔향성분들은 서로 반위상 관계에 있다고 가정하고 있다. 음수 채널간 상관도에 대해서 이러한 가정은 잔향성분을 과도하게 포함하도록 하는 문제점을 야기할 수 있다. 이에 대한 해결책으로 음수 채널간 상관도를 가지는 경우에 대해서는 주요성분들이 반위상 관계에 있다는 가정을 기반으로 새로운 채널간 상관도 합성 기법을 제안하였다. 제안된 기법은 실제 부호화 및 복호화기에서 동작하는 근사화 과정에도 적용되었다. 컴퓨터 시뮬레이션을 통하여 제안된 기법의 성능을 평가하였으며 주관적 청취 실험을 통하여 제안된 방법이 헤드폰 환경뿐만 아니라 스피커 환경에서도 효과적으로 동작함을 검증하였다.

핵심용어: 파라메트릭 스테레오, MPEG 서라운드, 공간 오디오 부호화, 채널간 상관도, 반위상

투고분야: 음향 신호처리 분야(1)

ABSTRACT: Parametric stereo(PS) and MPEG surround(MPS) are major spatial audio coding(SAC) tools. In this paper, the problem of the inter-channel correlation(ICC) synthesis in the conventional SAC is analyzed. Conventional methods assume that ambient components mixed to two output channels are anti-phased, while the primary components are assumed to be in-phased. This assumption can cause excessive ambient mixing for a negative-valued ICC. As a remedy to this problem, we propose a new ICC synthesis method based on an assumption that the primary components are anti-phased each other for a negative ICC. The proposed method is also applied to the approximation which works in practice. The performance of the proposed method was evaluated by computer simulations and the subjective listening tests verified that the proposed method is effective in not only headphones but also loudspeakers playback.

Key words: Parametric Stereo, MPEG Surround, Spatial Audio Coding, Inter-Channel Correlation, Anti-Phase

ASK subject classification: Acoustic Signal Processing (1)

1. 서 론

공간 오디오 부호화(SAC: spatial audio coding)에 대한 폭넓은 연구가 지난 십여 년간 진행되어 왔다. SAC에서는 멀티채널 오디오를 분석하여 모노 또는

스테레오의 다운믹스 신호와 공간 파라미터를 얻고 이러한 공간 파라미터들을 다운믹스 신호에 적용하여 인지적으로 원본과 유사한 멀티채널 오디오를 합성한다.^[1,2] 이 다운믹스 신호는 후방호환성을 가지며 따라서 기존의 복호화기에서도 동작한다.

양이단서부호화(BCC: binaural cue coding)는 SAC 기술을 구성하는 기본 기술이며 채널간 레벨차(ICLD:

*Corresponding author: 현동일(fireloaf@dsp.yonsei.ac.kr)
120-749 서울특별시 서대문구 신촌동 134 연세대학교 제2공학관 B601 디지털 신호처리 연구실
(전화: 02-2123-4534; 팩스: 02-364-4870)

inter-channel level difference), 채널간 시간차(ICTD: inter-channel time difference), 채널간 상관도(ICC: inter-channel correlation)의 세 가지의 공간 파라미터를 이용한다. ICLD와 ICTD는 음원의 방향과 연관된 파라미터이며 ICC는 음원의 폭과 연관된 파라미터이다.^[3] SAC 기술을 응용한 표준화된 코덱으로는 파라메트릭 스테레오(PS: parametric stereo)와 MPEG 서라운드(MPS: MPEG Surround)가 잘 알려져 있다. PS의 부호화 및 복호화 모듈은 MPS의 부호화 및 복호화 모듈인 TTO(two-to-one)과 OTT(one-to-two) 모듈에 대응된다.^[2] PS는 낮은 비트율로 스테레오를 표현하기 위한 다양한 코딩 기술에 적용되었다. PS는 2004년에 3GPP와 MPEG에 의해 각각 표준화된 enhanced aacPlus^[4]와 HE-AAC v2^[5]에 포함되었으며 가장 최근에 표준화된 코덱인 음성/음악 통합 압축기술(USAC: unified speech and audio codec)^[6]에도 MPS에 PS의 위상 정보가 추가된 스테레오 수정버전인 MPS 2-1-2 모드로서 포함되었다.

PS는 채널간 강도차(IID: interchannel intensity difference), 채널간 일관성(IC: interchannel coherence), 채널간 위상차(IPD: interchannel phase difference), 전체 위상차(OPD: overall phase difference)의 네 개의 공간 파라미터를 이용한다. PS의 성능에 있어 가장 중요한 것은 이러한 공간 파라미터들을 정확하게 분석하고 효과적으로 합성하는 것이다. 이러한 공간 파라미터 분석 및 합성에 관련된 연구들이 계속적으로 진행되었다.^[7-11]

MPS는 스피커 재생 시스템에 적용되며 해당 환경에서는 채널간 위상정보가 인지적인 성능 개선을 가져다주지 못한다고 알려져 있다.^[3,12] 따라서 MPS는 각각 PS의 IID와 ICC에 해당하는 채널 레벨차(CLD: channel level difference)와 채널간 상관도(ICC: inter-channel correlation)만을 이용한다. 따라서 채널간 위상 정보는 MPS에 채용되지 않는다.

공간 파라미터는 시간축, 주파수축, 양자화 해상도 등의 다양한 측면에서 확장성(scalability)을 제공하기 위해 선택적으로 사용될 수 있다. 이를 위한 한 방법으로 PS는 위상 파라미터를 포함하지 않는 모드를 제공한다. 이 경우 누락된 위상 파라미터를 보상해 주기 위하여 ICC 분석을 변형하여 사용한다. 원래

의 ICC는 0과 1사이의 값으로 제한되는 반면에 변형된 ICC 분석은 음수의 ICC를 가지기도 한다.^[1]

기존의 복호화기는 양수 및 음수 ICC에 대해 동일하게 동작하는 업믹스 행렬을 이용하여 합성한다. 업믹스 행렬은 다운믹스 신호와 다운믹스 신호를 디코릴레이션하여 얻은 디코릴레이션된 신호들을 각 채널에 분배함으로써 업믹스된 출력 신호의 CLD와 ICC가 전송된 공간 파라미터와 같아지도록 하는 방법을 통해 설계되었다. 업믹스 행렬의 입력으로 사용되는 다운믹스와 디코릴레이션된 신호들은 각각 주요성분분석(PCA: principal component analysis)을 통해 얻어지는 주요성분과 잔향성분에 해당된다.^[1] 업믹스 행렬 설계에는 위상에 대한 두 가지 가정이 전제된다. 하나는 출력채널 내의 주요성분들은 동일한 위상을 가진다는 것이며 다른 하나는 잔향성분들은 서로 반대의 위상을 가진다는 것이다. 이러한 가정 하에서 ICC가 양수일 경우, 출력 신호 내에 존재하는 다운믹스(주요성분) 신호의 비율은 항상 디코릴레이션된(잔향성분) 신호의 비율보다 높게 된다. 그러나 ICC가 음수인 경우에는 잔향성분이 다운믹스 신호보다 큰 에너지를 가지고 출력 신호에 분배될 수 있게 된다. 업믹스 행렬은 PCA를 기반으로 얻어지는데 이러한 경우는 PCA의 기본 가정인 주요성분이 항상 잔향성분보다 크다는 것을 위반하게 되며 잔향성분의 비율이 증가할수록 합성된 오디오 신호의 자연성이 결여되는 경향을 가진다.

본 논문에서는 음수 ICC 합성으로 인한 문제를 해결하기 위해 새로운 업믹스 행렬을 제안하였다. 복호화 단에서 음수 ICC를 접하게 될 경우, 두 개의 출력 채널 내의 주요성분들은 서로 반위상 관계에 있으며 잔향성분들은 동일위상을 가진다고 가정한다. 음수 ICC는 주요성분간의 IPD가 보다 큰 경우에 나타나기 때문에 이러한 가정이 가능하며 이를 기반으로 새로운 업믹스 행렬을 제안한다. 제안된 업믹스 행렬은 출력 채널간에 대하여 기존과 같이 동일한 ICC를 얻을 수 있게 하며 주요성분이 항상 잔향성분보다 크게 해준다. 제안된 업믹스 행렬의 성능을 검증하기 위하여 컴퓨터 시뮬레이션과 주관적 음질 평가를 수행하였다.

본 논문은 다음과 같이 구성되었다. 2장에서는 기

존의 ICC 합성 방법을 검토하고 ICC 업믹스 행렬에 대한 가정과 음수 ICC 합성으로 인한 문제를 검토하였다. 제안된 방법에 대한 상세한 설명은 3장에 기술하였다. 4장에서는 주관적 음질평가를 통하여 제안된 방법의 성능을 검증하였으며 최종적으로 5장에서 결론을 맺는다.

II. 배경지식

이 장에서는 기존의 ICC 분석 방법을 검토하고 음수 ICC 합성 시에 발생하는 문제를 살펴보았다.

2.1 공간 파라미터 분석

PS는 스테레오 오디오를 위한 SAC의 특정한 버전이며 두 가지 형태의 부호화 및 복호화 모드를 가진다. 첫 번째 모드는 CLD와 ICC와 같은 레벨 파라미터뿐만 아니라 IPD와 OPD와 같은 위상 파라미터도 사용하며 두 번째 구조는 공간 파라미터 양자화를 위한 비트율을 줄이기 위하여 CLD와 ICC만을 사용한다.^[1] 이러한 구조는 MPS의 TTO와 OTT 모듈과 같은 형태를 가진다.^[2] PS는 다음과 같이 공간 파라미터를 계산한다.^[1]

$$\begin{aligned} \text{CLD}[b] &= \frac{\sum_m \sum_{k=k_0}^{k_{0+1}-1} L_m[k] L_m^*[k]}{\sum_m \sum_{k=k_0}^{k_{0+1}-1} R_m[k] R_m^*[k]} \\ \text{ICC}_1[b] &= \frac{\left| \sum_m \sum_{k=k_0}^{k_{0+1}-1} L_m[k] R_m^*[k] \right|}{\sqrt{\sum_m \sum_{k=k_0}^{k_{0+1}-1} L_m[k] L_m^*[k]} \sqrt{\sum_m \sum_{k=k_0}^{k_{0+1}-1} R_m[k] R_m^*[k]}} \quad (1) \\ \text{IPD}[b] &= \angle \left(\sum_m \sum_{k=k_0}^{k_{0+1}-1} L_m[k] R_m^*[k] \right) \\ \text{OPD}[b] &= \angle \left(\sum_m \sum_{k=k_0}^{k_{0+1}-1} L_m[k] S_m^*[k] \right) \end{aligned}$$

b 는 파라미터 밴드의 인덱스이며 k_0 는 파라미터 밴드의 시작 인덱스이다. CLD와 ICC는 다음과 같이 양자화된다^[1]. CLD는 청각특성을 반영하기 위하여 dB 단위로 양자화된다.

$$\begin{aligned} \text{CLD} &= [-50, -45, -40, -35, -30, -25, -22, -19, -16, -13, -10, -8, \\ &\quad -6, -4, -2, 0, 2, 4, 6, 8, 10, 13, 16, 19, 22, 25, 30, 35, 40, 45, 50], \\ \text{ICC} &= [1, 0.937, 0.84118, 0.60092, 0.36764, 0, -0.589, -0.99] \\ \text{ICC와 IPD} &\text{는 각각 정규화된 상호상관도의 크기와} \end{aligned}$$

위상이다. 따라서 ICC는 0부터 1사이의 값으로 제한된다. 위상 파라미터를 사용하지 않는 모드에서 ICC는 다음 식과 같이 나타내어진다.^[1]

$$\text{ICC}_2[b] = \frac{\text{Re} \left\{ \sum_m \sum_{k=k_0}^{k_{0+1}-1} L_m[k] R_m^*[k] \right\}}{\sqrt{\sum_m \sum_{k=k_0}^{k_{0+1}-1} L_m[k] L_m^*[k]} \sqrt{\sum_m \sum_{k=k_0}^{k_{0+1}-1} R_m[k] R_m^*[k]}} \quad (2)$$

정규화된 상호상관도는 복소수가 될 수 있다. 이 경우 ICC₂는 식(2)과 같이 정규화된 상호상관도의 실수부만을 취한 값으로 얻어진다. 만약 두 입력 채널이 동일한 위상을 가진다면 ICC₁과 ICC₂는 동일한 값을 가진다. 그러나 IPD가 증가함에 따라 ICC₁과 ICC₂간의 차이도 증가하게 되며 특히 IPD가 $\pi/2$ 보다 커지게 되면 ICC₂는 음수가 된다. 실제로 ICC₂는 -1부터 1까지의 범위를 가진다.

2.2 위상 파라미터를 사용하는 기존의 업믹스 행렬

업믹스 행렬은 전송된 공간 파라미터에 따라 다른 형태의 구조를 가진다. 출력 신호는 다운믹스 신호와 디코릴레이션된 신호들을 믹싱해 줌으로써 합성할 수 있다. 이 경우 업믹스 행렬은 전송된 공간 파라미터들이 출력 신호에 대해서도 유지되도록 설계되어 있다. PS의 업믹스 행렬은 다음과 같이 정의된다.^[6]

$$\begin{bmatrix} \hat{L}[k] \\ \hat{R}[k] \end{bmatrix} = \begin{bmatrix} \lambda_1 \cos(\beta + \alpha) & \lambda_1 \sin(\beta + \alpha) \\ \lambda_2 \cos(\beta - \alpha) & \lambda_2 \sin(\beta - \alpha) \end{bmatrix} \begin{bmatrix} \hat{S}[k] \\ D(\hat{S}[k]) \end{bmatrix} \quad (3)$$

이 때 각각의 계수들은 식(4)과 같이 정의된다.

$$\begin{aligned} \lambda_1 &= \sqrt{\frac{10^{\text{CLD}[b]/10}}{1+10^{\text{CLD}[b]/10}}}, \quad \lambda_2 = \sqrt{\frac{1}{1+10^{\text{CLD}[b]/10}}} \\ \alpha &= \frac{1}{2} \arccos(\text{ICC}[b]), \quad \beta = \arctan \left[\frac{\lambda_2 - \lambda_1 \tan(\alpha)}{\lambda_2 + \lambda_1} \right] \end{aligned} \quad (4)$$

합성된 출력 신호 내의 주요성분과 잔향성분은 그림 1에 나타난 것처럼 벡터로 표현될 수 있다.^[13] 벡터 공간은 서로 같은 크기를 가지며 직교하는 다운믹스(주요성분) $\hat{S}[k]$ 과 디코릴레이션된(잔향성분) $D(\hat{S}[k])$ 에 의해 확장되며 CLD와 ICC는 신호 간에 이루는 크기와 각도에 의해 통제된다.

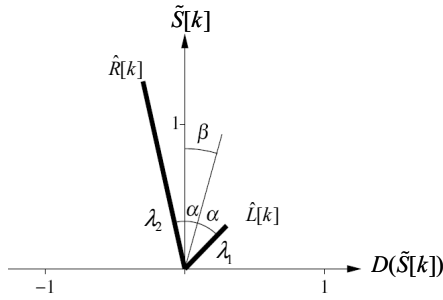


그림 1. 동위상 주요성분 가정하에서의 주요-잔향 영역에서의 업믹싱 표현
 Fig. 1. Representation of upmixing in primary-ambient domain under the assumption of in-phase primary components.

각 채널의 이득은 CLD에 의해 결정되며 회전각도 α 는 ICC에 의해 결정된다. 행렬의 원소들의 제곱의 합은 항상 1이 된다. 상수 β 는 출력 신호 내의 다운믹스 신호의 비율이 최대가 되도록 하는 값으로 결정되며 출력의 크기가 일정하므로 다르게 해석하면 출력 신호 내의 디코릴레이션된 신호의 비율이 최소가 되도록 하는 것이다. 이와 같이 출력 신호 내의 디코릴레이션된 신호의 비율이 최소가 되는 것은 각 채널에 대한 이득이 부호는 반대이며 크기가 같은 경우에 얻어진다. 따라서 $\lambda_1 \sin(\beta + \alpha)$ 는 $-\lambda_2 \sin(\beta - \alpha)$ 와 같아지게 되며 이는 식(5)과 같이 PS에서 가정한 신호 모델에 부합한다.^[1]

$$\begin{aligned} L[k] &= a_1 S[k] + D[k] \\ R[k] &= a_2 S[k] - D[k] \end{aligned} \quad (5)$$

β 를 사용함으로써 다운믹스 신호의 에너지는 같거나 증가하고 디코릴레이션된 신호의 에너지는 같거나 감소하게 되며 이는 식(6)과 같이 표현된다.

$$\begin{aligned} \lambda_1^2 \cos^2(\beta + \alpha) + \lambda_2^2 \cos^2(\beta - \alpha) &\geq \lambda_1^2 \cos^2(\alpha) + \lambda_2^2 \cos^2(-\alpha) \\ \lambda_1^2 \sin^2(\beta + \alpha) + \lambda_2^2 \sin^2(\beta - \alpha) &\leq \lambda_1^2 \sin^2(\alpha) + \lambda_2^2 \sin^2(-\alpha) \end{aligned} \quad (6)$$

ICC가 양수인 경우 a 의 크기는 항상 $\pi/4$ 보다 작아지며 이 경우 a 는 다음의 부등식을 항상 만족시킨다.

$$\begin{aligned} \lambda_1^2 \cos^2(\alpha) + \lambda_2^2 \cos^2(-\alpha) &> \lambda_1^2 \sin^2(\alpha) + \lambda_2^2 \sin^2(-\alpha), \\ (\lambda_1^2 + \lambda_2^2) \cos^2 \alpha &> (\lambda_1^2 + \lambda_2^2) \sin^2 \alpha, \\ \cos^2 \alpha &> \sin^2 \alpha \end{aligned} \quad (7)$$

식(6)과 (7)로부터 ICC가 양수일 경우, 출력 신호 내의 다운믹스 신호의 에너지는 항상 디코릴레이션된 신호보다 큰 것을 확인할 수 있다.

2.3 음수 ICC합성에 대한 문제점

식(4)로부터 ICC가 음수인 경우, a 는 $\pi/4$ 부터 $\pi/2$ 사이의 값으로 제한된다. 또한 CLD가 0dB에 접근함에 따라 β 도 0으로 접근하고 출력 신호 내의 잔향성분의 비가 주요성분의 비보다 커지게 된다. 이러한 경우는 PS의 부호화 및 복호화의 핵심적인 원리인 PCA의 가정을 위반하게 된다. 식(3)과 같은 기존의 업믹스 행렬에서는 각 채널에 대한 주요성분과 잔향성분이 각각 동위상과 반위상인 것으로 가정한다. ICC가 -1이고 CLD가 0dB일 경우 기존의 업믹스 행

$$\begin{bmatrix} 0 & \sqrt{2}/2 \\ 0 & -\sqrt{2}/2 \end{bmatrix} \text{로 얻어진다.}$$

이러한 극단적 경우에 한 채널은 다른 채널에 대해 완벽하게 반위상이 되며 결과적으로, 출력 신호 내에 주요성분은 없고 잔향성분만이 존재하게 된다. 또한 디코릴레이션된 잔향성분은 인공적으로 만들어졌기 때문에 합성된 출력 신호는 자연성이 결여되며 전체적인 오디오 품질 악화를 야기할 수 있다.

그림 2는 역양자화된 CLD와 ICC에 따른 주요성분 및 잔향성분 대 출력 신호의 에너지 비율을 나타낸다. $D(\tilde{S}[k])$ 의 에너지는 $\tilde{S}[k]$ 의 에너지와 같아지도록 정규화되었으며 모든 이득의 제곱의 합은 1이기 때문에 각 에너지 비율은 다음과 같이 계산되었다.

$$\begin{aligned} \text{primary: } &\lambda_1^2 \cos^2(\beta + \alpha) + \lambda_2^2 \cos^2(\beta - \alpha) \\ \text{ambient: } &\lambda_1^2 \sin^2(\beta + \alpha) + \lambda_2^2 \sin^2(\beta - \alpha) \end{aligned} \quad (8)$$

가로축은 ICC의 양자화된 인덱스이며 CLD는 0, 2, 4, 6 dB의 값들이 선택되었다. ICC 인덱스 중 6과 7은 각각 -0.589와 -0.99에 대응되며 이 두 경우 ICC는 음수값을 가지게 된다. ICC가 음수이고 CLD가 0dB에 근접하는 경우 잔향성분이 주요성분보다 커지는 것을 볼 수 있다. CLD가 6dB보다 커지면 ICC와 무관하게 주요성분은 잔향성분보다 항상 커지게 된다.

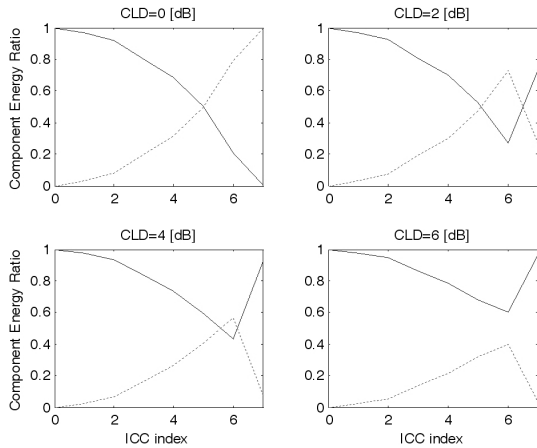


그림 2. 기존의 업믹스 행렬을 이용하여 얻은 합성된 출력 신호 내의 주요성분과 잔향성분의 에너지 비율(식 (3), (4))(실선: 다운믹스 신호, 파선: 디코릴레이션된 신호)

Fig 2. The energy ratios of the primary and ambient components in the synthesized output signal obtained using the conventional upmix matrix (Eqs. (3) and (4)) (solid: downmix, dashed: decorrelated).

III. 반위상 주요성분 가정에 기반을 둔 업믹스 행렬

3.1 제안된 업믹스 행렬

앞서 언급된 것처럼 음수 ICC의 경우, 동위상 주요 성분 가정은 출력 신호 내의 잔향의 비율이 증가하는 문제를 야기할 수 있다. 식(2)로부터 주요성분이 동위상인 경우 IPD는 0이 되며 주요성분이 반위상인 경우 IPD가 π 가 되는 점을 알 수 있다. 따라서 음수 ICC는 IPD가 $\pi/2$ 보다 커지는 것을 의미하며 각 채널에 대한 주요성분들이 동위상이기 보다는 반위상에 가깝게 된다. 결과적으로 ICC가 음수이면 각 채널의 주요성분들이 반위상이라고 가정하는 것이 보다 적절하다. 이러한 가정 하에서 주요-잔향 영역에서의 업믹싱 과정은 그림 3과 같은 구조를 기반으로 처리될 수 있다. 주요성분과 잔향성분에 대한 가정이 서로 뒤바뀌었기 때문에 업믹싱 과정은 각 성분들의 축을 바꾸는 것으로 표현될 수 있다.

그림 3을 통하여 음수 ICC에 대해 반위상 주요 성분 가정을 만족하는 새로운 업믹스 행렬을 얻을 수 있다. 새로운 업믹스 행렬은 식(3)의 기존 방법과 유사한 구조를 가진다.

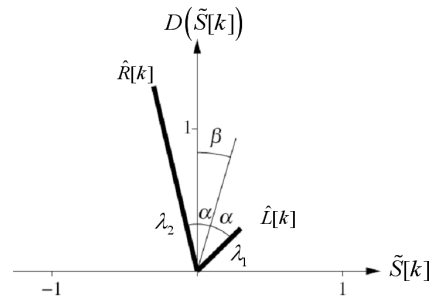


그림 3. 반위상 주요성분 가정하에서의 주요-잔향 영역에서의 업믹싱 표현

Fig 3. Representation of upmixing in primary-ambient domain under the assumption of anti-phase primary components.

$$\begin{bmatrix} \hat{I}[k] \\ \hat{R}[k] \end{bmatrix} = \begin{bmatrix} \lambda_1 \sin(\beta' + \alpha) & \lambda_1 \cos(\beta' + \alpha) \\ \lambda_2 \sin(\beta' - \alpha) & \lambda_2 \cos(\beta' - \alpha) \end{bmatrix} \begin{bmatrix} \hat{S}[k] \\ D(\hat{S}[k]) \end{bmatrix} \quad (9)$$

$\lambda_1, \lambda_2, \alpha$ 는 식(4)을 이용하여 계산되며 β' 는 주요성분에 대한 새로운 가정을 만족시키기 위하여 다른 방법으로 계산된다. 식(3), (4)와 같은 기존의 업믹스 행렬은 반대 부호의 잔향 이득을 가지는 반면에 새로운 업믹스 행렬에서의 잔향 이득은 동일한 값을 가지도록 제한되며 이러한 제한을 만족하는 β' 는 다음과 같이 결정된다.

$$\beta' = \pm \arctan \left[\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \cdot \frac{1}{\tan(\alpha)} \right] \quad (10)$$

제안된 업믹스 행렬은 주요성분의 에너지가 잔향성분의 에너지보다 커지도록 유지시켜 준다. 단 제안된 업믹스 행렬은 한 채널의 주요성분의 위상을 반전시키기 때문에 조심스런 고려가 필요하다. 두 개의 해를 갖는 것은 이러한 이유 때문이다. 위상 반전으로 인해 발생할 수 있는 문제를 방지하기 위하여 위상 반전은 작은 이득을 가지는 채널에 대해 적용된다. 이를 만족하는 β' 은 다음과 같이 항상 음수 값을 가지게 된다.

$$\begin{aligned} \text{if } \lambda_1 > \lambda_2, \text{ then } \beta' &= \arctan \left[\frac{\lambda_2 - \lambda_1}{\lambda_2 + \lambda_1} \cdot \frac{1}{\tan(\alpha)} \right] \\ \text{if } \lambda_1 < \lambda_2, \text{ then } \beta' &= \arctan \left[\frac{\lambda_1 - \lambda_2}{\lambda_2 + \lambda_1} \cdot \frac{1}{\tan(\alpha)} \right] \end{aligned} \quad (11)$$

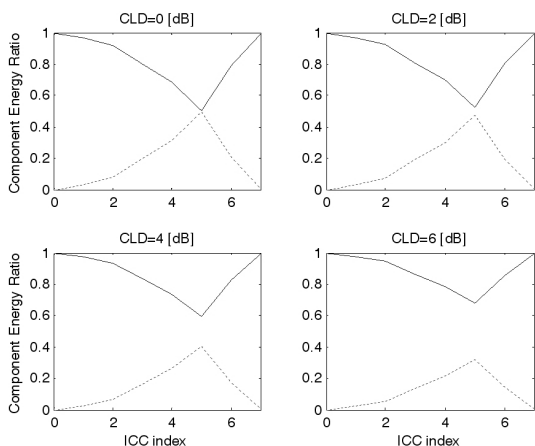


그림 4. 새로운 업믹스 행렬을 이용하여 얻은 합성된 출력 신호 내의 주요성분과 잔향성분의 에너지 비율(식 (9), (4)) (실선: 다운믹스 신호, 파선: 디코릴레이션된 신호)
 Fig 4. The energy ratios of the primary and ambient components in the synthesized output signal obtained using the new upmix matrix (Eqs. (9) and (4)) (solid: downmix, dashed: decorrelated).

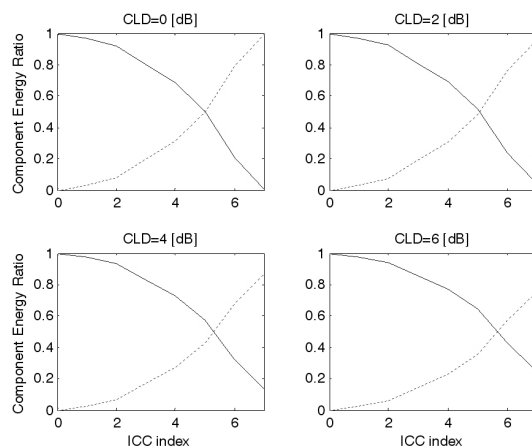


그림 5. 식(12)을 이용하여 얻은 합성된 출력 신호 내의 주요성분과 잔향성분의 에너지 비율(실선: 다운믹스 신호, 파선: 디코릴레이션된 신호)
 Fig 5. The energy ratios of the primary and ambient components in the synthesized output signal obtained using Eq. (12) (solid: downmix, dashed: decorrelated).

인접 프레임 간에 부호 바뀔이 일어나게 되면 주파수 영역의 값들을 시간 영역의 값으로 변환하는 과정에서 에너지 상쇄가 발생하게 된다. 이는 업믹스 행렬을 프레임 간에 보간시킴으로써 완화시킬 수 있다. 채널 이득이 동일한 경우에는 현재 프레임에서의 부호가 이전 프레임에서의 부호와 같도록 설정하였다. 이는 인접한 프레임간의 빈번한 부호 바뀔로 인해 발생할 수 있는 왜곡을 막기 위해서이다.

그림 4는 새로운 업믹스 행렬을 이용하여 얻어지는 출력 신호 내의 주요성분과 잔향성분의 에너지 비율을 보여준다. 그림에서 나타난 것과 같이 모든 CLD와 ICC에 대해 주요성분이 잔향성분보다 더 큰 것을 볼 수 있다.

3.2 근사화된 업믹스 행렬

실제 구현 과정에서 식(4)의 β 를 계산하기 위해서는 역탄젠트와 나눗셈 연산이 필요하며 이는 상당한 복잡도를 요구한다. 따라서 실제 복호화기에서는 계산 복잡도를 낮추기 위해 다음과 같이 근사화된 값으로 계산된다.^[13]

$$\beta = \alpha \frac{(\lambda_2 - \lambda_1)}{\sqrt{2}} \quad (12)$$

그러나 위 식의 근사화 과정은 잔향성분의 에너지가 우세하게 되는 문제점을 더욱 악화시킨다. 그림 5는 식(12)을 이용하여 얻은 출력 신호 내의 주요성분과 잔향성분의 에너지 비율을 보여준다. 그림 2에서의 정확한 값을 사용한 경우와 비교하면 음수 ICC의 경우 특히 잔향성분의 에너지가 우세하게 되는 문제점이 심해지는 것을 볼 수 있다.

식(12)과 유사한 방법으로 새롭게 정의된 파라미터 β' 는 다음과 같이 근사화할 수 있다.

$$\begin{aligned} \text{if } \lambda_1 > \lambda_2, \text{ then } \beta' &= \frac{1}{\alpha} \cdot \frac{(\lambda_2 - \lambda_1)}{\sqrt{2}} \\ \text{if } \lambda_1 < \lambda_2, \text{ then } \beta' &= \frac{1}{\alpha} \cdot \frac{(\lambda_1 - \lambda_2)}{\sqrt{2}} \end{aligned} \quad (13)$$

식(12)의 근사화와는 달리 식(13)에서의 근사화는 매우 효과적이다. 식(13)을 이용하여 얻은 에너지 비율을 그림 6에 표시하였다. 제안된 업믹스 행렬에 대한 근사화는 성분들의 에너지 비율에 뚜렷한 영향을 주지 않았다.

IV. 성능 평가

제안된 업믹스 행렬의 성능을 검증하기 위하여 주관적 청취 실험을 수행하였다. 청취 실험에 대한 경험

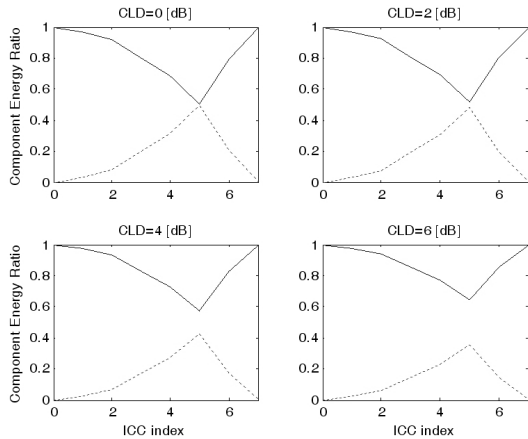


그림 6. 식(13)의 근사화된 파라미터를 이용하여 얻은 합성된 출력 신호 내의 주요성분과 잔향성분의 에너지 비율(실선: 다운믹스 신호, 파선: 디코릴레이션된 신호)

Fig 6. The energy ratios of the primary and ambient components in the synthesized output signal obtained using the approximated parameter in Eq. (13)(solid: downmix, dashed: decorrelated).

이 있는 8명의 피실험자가 참여하였으며 이중-맹검법을 사용하는 MUSHRA 실험 방법을 이용하였다.^[4] 평가는 표 1에 표기된 실험 샘플에 대해 Sennheiser HD600 헤드폰을 이용한 헤드폰 환경과 Egosys nEar04 모니터링 스피커를 이용한 스피커 환경의 두 경우에 대해 반복하여 수행되었다. 실험에는 48 kHz의 샘플링 주파수를 가지며 8초의 길이를 가지는 네 개의 샘플이 사용되었다. 위상 파라미터들은 1.5kHz 미만에서 현저한 역할을 하며 음수 ICC 합성의 누락된 위상 파라미터들에 대한 보상을 주된 목적으로 하기 때문에, 상대적으로 저대역에 에너지가 집중된 음성 신호뿐만 아니라, 일반적인 음악 신호에 대해서도 제안된 알고리즘의 성능을 검증하기 위하여 음악 신호도 실험 샘플로서 선택하였다. 앵커(anchor) 신호는 실험 신호들을 3.5kHz 차단 주파수를 가지는 저역통과필터를 이용하여 생성한다.

각각의 샘플들은 기존의 PS 부호화기에 의해 다운믹스 신호와 공간 파라미터로 분석되었다.^[4] PS에서 다운믹스 신호는 다양한 오디오 부호화기를 통해 부호화될 수 있다. 실험에서는 양자화로 인한 왜곡을 배제하기 위하여 양자화되지 않은 다운믹스 신호를 복호화기에 그대로 사용하였다. 공간 파라미터인 CLD와 ICC는 기존 PS에서 사용하는 양자화기를 사

표 1. 실험 샘플

Table 1. Test samples.

샘플명	설명
arirang	음성(한국인 남성)
wedding	음성(한국인 남성)
Music_1	음악(헤비메탈)
Music_3	음악(악기연주)

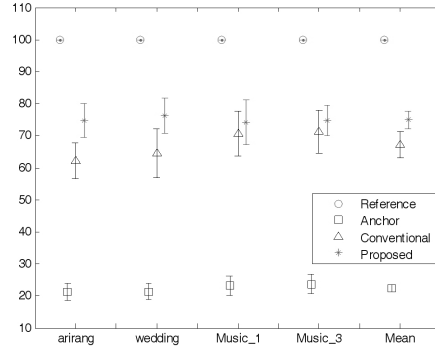


그림 7. MUSHRA 청취 실험 결과(헤드폰 환경)

Fig 7. MUSHRA listening test results (headphones playback).

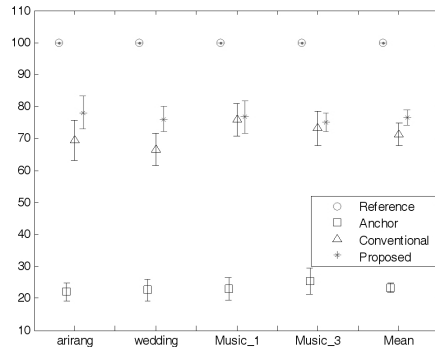


그림 8. MUSHRA 청취 실험 결과(스피커 환경)

Fig 8. MUSHRA listening test results (loudspeakers playback).

용하여 양자화하였다.^[4] 다운믹스와 공간 파라미터들은 각각 기존 PS의 업믹스 행렬^[1]과 제안된 업믹스 행렬에 의해 복호화되었다. 이 때 계수 β 는 근사화를 통해 계산하였다.

헤드폰 환경과 스피커 환경에 대한 실험결과를 각각 그림 7과 8에 나타내었다. 실험 결과에 대한 전체 평균과 95% 신뢰구간을 나타내었다. 실험 결과는 제안된 알고리즘이 기존에 비해 모든 경우에 향상된 음질을 보임을 나타낸다. 평균적으로 10점 정도의

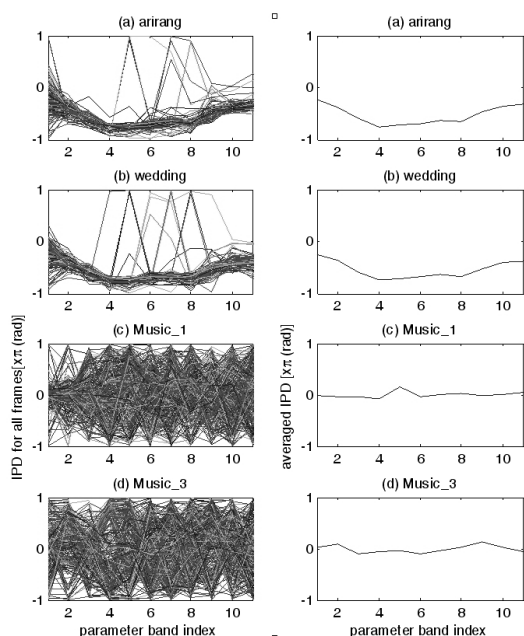


그림 9. 각 샘플의 IPD(좌: 모든 프레임에 대한 IPD들, 우: IPD의 평균값)

Fig. 9. IPD of each test sample (left: IPDs for all frames, right: averaged IPD).

성능개선을 확인할 수 있다.

스피커 환경의 경우 전체적으로 음질이 향상되는 결과를 보여주었는데 이는 헤드폰 환경의 경우 과도한 잔향성분으로 인한 왜곡이 귀에 직접적으로 전해지는 반면에 스피커 환경의 경우 공간의 영향을 받기 때문에 상대적으로 왜곡이 덜 민감하게 인지되기 때문이다. 스피커 환경에서도 제안된 업믹스 행렬을 적용한 경우에 상대적인 성능향상은 감소했지만 헤드폰 환경과 유사한 수준까지의 성능 향상을 보이는 것을 확인할 수 있었다.

각 샘플의 IPD를 그림 9에 나타내었다. 좌측은 모든 프레임에 대한 IPD를 나타낸 것이며 우측은 IPD의 평균값을 나타낸 것이다. 샘플 'arirang'과 'wedding'은 큰 IPD값을 가지는 반면에 샘플 'Music_1'과 'Music_3'는 평균적으로 0에 가까운 IPD값을 가지며 시간변화에 따라 다양하게 변화하는 IPD를 보여주었다. 제안된 알고리즘은 양수 ICC가 분석되는 경우에는 기존 PS와 동일하게 동작하며 반위상에 가까운 신호로부터 음수 ICC가 얻어지는 경우에 성능향상이 기대된다. 그림 7과 8에 나타난 실험 결과는 이를 잘 나타내준다. 상대적으로 음수 ICC의 빈도가 낮은

음악 신호의 경우 음성 신호에 비해 성능 개선이 크지 않은 것을 보여준다. 그러나 이와 같은 음악 신호에 대해서도 기존 방법과 비교할 때 거의 유사하거나 약간의 성능 개선을 보임을 확인할 수 있다.

V. 결 론

본 논문에서는 주요성분에 대한 동위상 가정에 의해 야기될 수 있는 문제점을 분석하였다. 동위상 가정은 음수 ICC를 접할 경우 잔향성분의 비율이 과도하게 커지게 하는 것을 수식을 이용한 분석을 통하여 나타내었다. 이러한 문제점을 완화하기 위하여 음수 ICC를 접했을 때 주요성분이 반위상이라는 가정을 기반으로 하는 새로운 업믹스 행렬을 제안하였다. 제안된 업믹스 행렬은 의도된 CLD와 ICC가 기존과 동일하게 획득될 수 있도록 하면서도 출력 신호 내의 주요성분이 잔향성분보다 항상 우세하도록 유지시켜 준다. 주관적 음질 평가를 통하여 제안된 업믹스 행렬의 성능을 검증하였으며 실험 결과로부터 제안된 알고리즘이 기존에 비해 항상 뛰어난 것을 확인할 수 있었다. 제안된 알고리즘은 헤드폰뿐만 아니라 스피커 재생환경에서 평가되었으며 따라서 스테레오 헤드폰을 대상으로 하는 PS뿐만 아니라 멀티채널 스피커를 사용하는 MPS에도 적용될 수 있다.

감사의 글

이 논문은 2011년도 정부(지식경제부)의 재원으로 산업원천기반구축개발사업의 지원을 받아 수행된 연구임(No. 10037244).

참고문헌

1. J. Breebaart, S. van de Par, A. Kohlrausch, and E. Schuijers, "Parametric coding of stereo audio," *EURASIP J. Appl. Signal Process.*, vol. 9, pp. 1305-1322, 2004.
2. J. Breebaart, G. Hotho, J. Koppens, E. Schuijers, W. Oomen, and S. van de Par, "Background, concept, and architecture for the recent MPEG surround standard on multichannel audio compression," *J. Audio Eng. Soc.* vol. 55, no. 5, pp. 331-351, 2007.

3. C. Faller and F. Baumgarte, "Binaural cue coding - part II: schemes and applications," *IEEE Trans. on Speech and Audio Proc.*, vol. 11, no. 6, pp. 520-531, Nov. 2003.
4. 3GPP TS 26 series Rel. 9, *Enhanced aacPlus general audio codec: Encoder specification parametric stereo part*, 2009.
5. M. Wolters, K. Kjörling, D. Himm, and H. Purnhagen, "A closer look into MPEG-4 high efficiency AAC," *AES 115th Convention*, paper no. 5871, Oct. 2003.
6. Study on ISO/IEC 23003-3:201x/DIS, *Information technology - MPEG audio technologies - Part 3: Unified speech and audio coding*, N12013, 2011.
7. J. Kim, E. Oh, and J. Robilliard, "Enhanced stereo coding with phase parameters for MPEG unified speech and audio coding," *AES 127th Convention*, paper no. 7875, Oct. 2009.
8. D. Hyun, J. Seo, Y. Park, and D. Youn, "Robust interchannel correlation (ICC) estimation using constant interchannel time difference (ICTD) compensation," *AES 127th Convention*, paper no. 7934, Oct. 2009.
9. M. Kim, E. Oh, and H. Shim, "Stereo audio coding improved by phase parameters," *AES 129th Convention*, paper no. 8289, Nov. 2010.
10. E. Oh and M. Kim, "Enhanced stereo algorithms in the unified speech and audio coding," *AES 43rd Int. Conf.*, Sep. 2011.
11. D. Hyun, Y. Park, S. Lee, and D. Youn, "Enhanced interchannel correlation (ICC) synthesis for spatial audio coding," *AES 43rd Int. Conf.*, Sep. 2011.
12. F. Baumgarte and C. Faller, "Why binaural cue coding is better than intensity stereo coding," *AES 112th Convention*, paper no. 5575, May 2002.
13. H. Purnhagen, "Low complexity parametric stereo coding in MPEG-4," *7th Int. Conf. on Audio Effects (DAFX-04)*, Oct. 2004.
14. ITU-R BS.1534-1, *Method for the subjective assessment of intermediate sound quality (MUSHRA)*, International Telecommunications Union, Geneva, Switzerland, 2001.

저자 약력

▶ 현 동 일(Dong-il Hyun)

2005년: 연세대학교 기계전자공학부(학사)
 2007년: 연세대학교 전기전자공학과(석사)
 2007년 ~ 현재: 연세대학교 전기전자공학과(박사과정)
 <관심 분야> 오디오 신호처리, 3D 오디오, 오디오 부호화

▶ 이 석 필(Seok-Pil Lee)

1990년: 연세대학교 전기공학과(학사)
 1992년: 연세대학교 전기공학과(석사)
 1997년: 연세대학교 전기공학과(박사)
 2009년 ~ 현재: 전자부품연구원(KETI) 디지털미디어연구센터 센터장
 <관심 분야> 디지털방송통신융합시스템, 멀티미디어 신호처리

▶ 박 영 철(Young-cheol Park)

1986년: 연세대학교 전자공학과(학사)
 1988년: 연세대학교 전자공학과(석사)
 1993년: 연세대학교 전자공학과(박사)
 2002년 ~ 현재: 연세대학교 컴퓨터정보통신공학부 교수
 <관심 분야> 음성 신호처리, 적응 신호처리, 오디오 신호처리

▶ 윤 대 희(Dae Hee Youn)

1977년: 연세대학교 전자공학과(학사)
 1979년: Kansas State University(석사)
 1982년: Kansas State University(박사)
 1985년 ~ 현재: 연세대학교 전기전자공학과 교수
 <관심 분야> 적응 신호처리, 음성 신호처리, 오디오 신호처리