

논문 2014-51-6-26

토널리티 정보를 기반으로 한 파라메트릭 스테레오 부호화의 개선된 합성 기법

(An Improved Synthesis Method of Parametric Stereo Coding Based on
Tonality Information)

이 동 금*, 박 영 철**, 윤 대 희***

(Tung chin Lee[©], Young-cheol Park, and Dae Hee Youn)

요 약

본 논문에서는 PS의 복호화과정에서 톤 성분에 영향을 주는 잔향 성분을 효과적으로 억제할 수 있는 합성 방법을 제안하였다. PS에서 잔향 성분은 비상관 필터를 이용하여 구할 수 있으며, 부호화단에서 분석된 IC 파라미터를 통해서 합성되는 잔향의 비중이 결정된다. 하지만 파라미터들은 서브밴드 도메인에서 분석되기 때문에, 톤 성분이 존재하는 대역에서도 낮은 IC값이 분석될 수 있고, 이는 출력 신호의 음질 열화를 야기시킨다. 본 논문에서는 이러한 문제를 보완하기 위해 복호화단으로 입력되는 다운믹스 신호의 토널리티를 측정하였고, 이 측정된 값을 통해 합성되는 잔향 성분의 비중을 조절해주었다. 실험은 시뮬레이션 결과를 통해 성능을 검증한 후에 주관적 음질 평가를 수행하였고, 전체적으로 음질 향상이 있음을 확인하였다.

Abstract

In this paper, we propose a synthesis method that can effectively suppress the ambience which affects tonal components in the PS decoder. Ambience component was obtained by using decorrelation filter and the weighting of the ambience in the decoder was determined through IC parameter. However, since the parameters are extracted in the sub-band domain, a low IC value could be analyzed even if the tonal component is dominant. The quality of the output signal may be degraded. To prevent this problem, the tonality was measured in the downmixed signal and the weighting of the ambience components were adjusted appropriately according to the measured tonality index. The performance of the proposed method was evaluated by simulations. Furthermore, the subjective test was performed and the results confirmed that the proposed method offers improved quality.

Keywords : Parametric stereo, MPEG surround, tonality, spatial audio coding

I. 서 론

멀티미디어 산업의 발전으로 인터넷, 원격 및 무선 방송 통신 등에 대한 관심이 꾸준히 증가하고 있다. 이에 따라서 비디오 및 오디오 서비스에 대해서 보다 나은 서비스를 제공하기 위해 많은 연구가 진행되었다. 오디오의 경우, 제한된 채널 환경에서 최상의 음질을 재생할 수 있도록 오디오 부호화 기술에 대한 연구

* 학생회원, *** 평생회원, 연세대학교 전기전자공학과
(School of Electrical and Electronic Engineering,
Yonsei University)

** 평생회원, 연세대학교 컴퓨터 정보통신공학부
(Computer and Telecommunications Engineering
Division, Yonsei University)

© Corresponding Author(E-mail: riglord@dsp.yonsei.ac.kr)

접수일자: 2014년02월08일, 수정일자: 2014년05월21일
수정완료: 2014년05월25일

가 20년 넘게 진행되어 오고 있는데, 초창기에는 심리 음향(psychoacoustic) 기술^[1]을 적용하여 사람이 인지하지 못하는 오디오 신호 내의 정보들을 제거하여 압축하는 기술 (MPEG Layer-III(MP3)^[2], Advanced Audio Coding(AAC)^[3])이 추가 되었으나, 2000년 이후에는 오디오 신호의 정보들을 파라미터(parameter)로 표현하여 부호화하는 기술 (High Efficiency Advanced Audio Coding (HE-AAC)^[4], MPEG Surround^[5])이 개발 되었다.

파라메트릭 스테레오 (Parametric Stereo, PS) 부호화기^[6]는 파라미터 부호화기의 한 종류로서 스테레오 (stereo) 신호를 모노(mono) 신호로 다운믹스(downmix)하고, 입력 신호의 공간 정보를 파라미터로 추출하여 부호화하는 방법이다. MPEG HE-AAC의 한 기술로 표준화된 PS는 낮은 비트율로도 고품질의 스테레오 부호화가 가능하다. 이 후, MPEG에서는 PS 부호화 방법을 멀티채널 환경에 적합하도록 확장하여 MPEG 서라운드(MPEG Surround)라는 이름으로도 표준화를 완료하였다^[5].

PS나 MPEG Surround는 모두 부호화 과정에서 채널 수가 줄어들고, 복호화 과정에서는 오직 파라미터 정보만을 이용하여 합성하기 때문에 입력 신호의 파형(waveform)을 완벽하게 복원하기 어렵다^[7]. 하지만 입력 신호의 공간감을 출력 신호에서도 복원하기 위해 복호화단에서 비상관(decorrelation) 필터^[6]를 이용하여 잔향신호를 생성하고, 전송된 파라미터 정보를 이용하여 그 비중을 조절한다. 하지만 새 잔향 신호는 입력 신호의 특성이 잘 복원되도록 생성된 신호가 아니며, 복호화 과정이 서브밴드(sub-band) 단위로 파라미터들이 적용되기 때문에 복원된 신호는 잔향 성분에 의해 음질 열화가 발생한다. 특히 톤(tone) 특성이 강하게 존재하는 주파수 대역에서도 불필요한 잔향(ambience) 신호가 함께 합성되는 경우가 발생하여 출력 신호에 영향을 끼친다. 본 논문에서는 이러한 문제를 해결하기 위해 복호화단으로 입력된 다운믹스 신호의 토널리티(tonality)^[8]를 추정하였고, 추정된 토널리티 인덱스(index)에 따라서 합성된 신호내의 잔향 성분 비중을 조절해주었다. 제안된 방법은 복호화단에서만 이루어지기 때문에 부호화단에서 어떤 추가 정보를 필요로 하지 않는 장점이 있다.

본 논문의 구성은 다음과 같다. II장에서는 PS의 기

본 구조를 간단히 소개하고 합성 과정에서의 문제점에 대해서 설명하였다. III장에서는 본 논문에서 제안하는 토널리티 기반의 합성 방법을 기술하였다. IV장에서는 제안된 방법의 성능을 평가하기 위해 주관적 실험 결과를 나타내었고 마지막 V장에서 결론을 맺는다.

II. PS의 기본적인 구조

기본적인 PS 부호화의 부호화기와 복호화기의 구조는 그림 1과 같다. 입력된 스테리오는 하이브리드 QMF 필터뱅크(Hybrid QMF filterbank)^[5]를 통해 시간-주파수 신호 $X_L(n,k)$, $X_R(n,k)$ 로 변환된다. 다음 파라미터 밴드 단위로 아래의 4가지 공간 파라미터 Interchannel Intensity Difference(IID), Interchannel Coherence(IC), Interchannel Phase Difference(IPD), Overall Phase Difference(OPD)를 분석한다^[6].

$$IID(b) = 10 \log_{10} \frac{\sum_n \sum_{k=k_0}^{k_{0+1}-1} X_L(n,k) X_L^*(n,k)}{\sum_n \sum_{k=k_0}^{k_{0+1}-1} X_R(n,k) X_R^*(n,k)}$$

$$IC(b) = \frac{\left| \sum_n \sum_{k=k_0}^{k_{0+1}-1} X_L(n,k) X_R^*(n,k) \right|}{\sqrt{\left(\sum_n \sum_{k=k_0}^{k_{0+1}-1} X_L(n,k) X_L^*(n,k) \right) \left(\sum_n \sum_{k=k_0}^{k_{0+1}-1} X_R(n,k) X_R^*(n,k) \right)}} \quad (1)$$

$$IPD(b) = \angle \left(\sum_n \sum_{k=k_0}^{k_{0+1}-1} X_L(n,k) X_R^*(n,k) \right)$$

$$OPD(b) = \angle \left(\sum_n \sum_{k=k_0}^{k_{0+1}-1} X_L(n,k) X_S^*(n,k) \right)$$

PS에서는 입력 스테레오 신호를 다음과 같은 신호

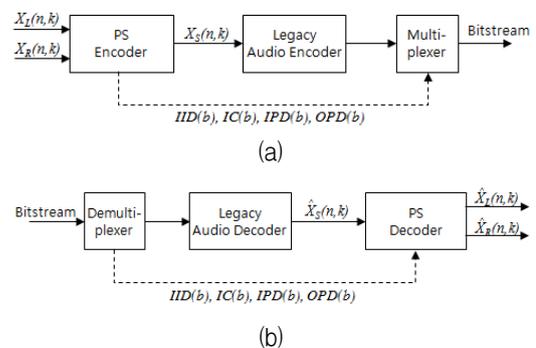


그림 1. 파라메트릭 스테레오 부호화. (a) 부호화기 (b) 복호화기.
Fig. 1. Parametric stereo coding. (a) Encoder (b) Decoder.

모델로 가정하고 있다.

$$\begin{aligned} X_L(n,k) &= g_L X_S(n,k) + D(n,k) \\ X_R(n,k) &= g_R X_S(n,k) - D(n,k), \end{aligned} \quad (2)$$

입력 좌우 채널에 주요 성분(primary component)인 $X_S(n,k)$ 가 공통적으로 포함되고, 잔향 성분 신호 $D(n,k)$ 가 반대 위상(anti-phase)의 관계를 가지며 양쪽 채널에 더해진다. 다운 믹스 신호는 아래 식으로 계산된다.

$$X_S(n,k) = \frac{X_L(n,k) + X_R(n,k)}{(g_L + g_R)} \quad (3)$$

식에서 $(g_L + g_R)$ 은 다운믹스 신호의 에너지를 보존하기 위한 게인(gain)값으로, 식 (4)와 같이 좌우 채널의 파워의 합이 다운믹스 신호 파워의 합과 같도록 정해진다.

$$E[|X_S(n,k)|^2] = E[|X_L(n,k)|^2] + E[|X_R(n,k)|^2] \quad (4)$$

다음, 다운 믹스된 모노 신호는 현존하는 오디오 부호화기 (MP3, AAC 등)를 이용하여 부호화되어, 분석된 파라미터와 함께 비트스트림으로 전송된다.

그림 1 (b)의 PS 복호화기는 비트스트림으로부터 전송된 복호화된 다운믹스 신호와 공간 부호화 파라미터를 이용하여 아래 식을 이용하여 스테레오 신호를 합성한다.

$$\begin{aligned} \begin{bmatrix} \hat{X}_L \\ \hat{X}_R \end{bmatrix} &= \begin{bmatrix} e^{jOPD(b)} & 0 \\ 0 & e^{-j(OPD(b) - IPD(b))} \end{bmatrix} \times \\ &\begin{bmatrix} \lambda_L & 0 \\ 0 & \lambda_R \end{bmatrix} \begin{bmatrix} \cos(\alpha + \beta) & \sin(\alpha + \beta) \\ \cos(-\alpha + \beta) & \sin(-\alpha + \beta) \end{bmatrix} \begin{bmatrix} \hat{X}_S(n,k) \\ D(\hat{X}_S(n,k)) \end{bmatrix} \\ &= \begin{bmatrix} \lambda_L e^{jOPD(b)} \cos(\alpha + \beta) & \lambda_L e^{-j(OPD(b) - IPD(b))} \sin(\alpha + \beta) \\ \lambda_R e^{jOPD(b)} \cos(-\alpha + \beta) & \lambda_R e^{-j(OPD(b) - IPD(b))} \sin(-\alpha + \beta) \end{bmatrix} \times \\ &\begin{bmatrix} \hat{X}_S(n,k) \\ D(\hat{X}_S(n,k)) \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} \lambda_L &= \sqrt{\frac{10^{IID(b)/10}}{1 + 10^{IID(b)/10}}}, \lambda_R = \sqrt{\frac{1}{1 + 10^{IID(b)/10}}}, \\ \alpha &= 0.5 \arccos(IC(b)), \beta = \arctan\left(\frac{\lambda_L - \lambda_R}{\lambda_L + \lambda_R} \tan(\alpha)\right). \end{aligned} \quad (5)$$

식에서 $D(\hat{X}_S(n,k))$ 는 비상관 필터를 이용하여 만든 인공 잔향 신호이다. 이는 다운믹스 과정에서 잔향 성분이 제거되었다고 가정하고 있기 때문에, 출력 신호에서도 입력 신호와 똑같은 공간감을 나타내기 위한 목적으

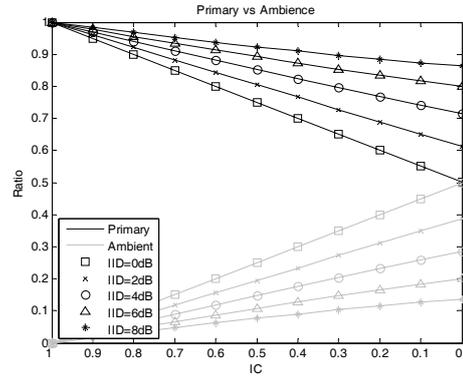


그림 2. IID 와 IC 값에 따른 주요 성분과 잔향 성분의 상대적 비중.

Fig. 2. The relative weights of primary and ambience component according to IID and IC value.

로 이용된다. 이러한 이유로, 전송된 파라미터 정보는 출력신호에서도 똑같이 유지되며, 입력 신호의 공간 정보를 복원해낼 수 있다.

부호화단에서 전송된 파라미터중 $IID(b)$ 는 좌우 채널의 크기 레벨 λ_L 과 λ_R 로, $IPD(b)$ 와 $OPD(b)$ 는 그대로 좌우 채널의 위상차 정보로 사용된다. $IC(b)$ 는 식 (5)에서 회전 행렬 요소(element)의 α 와 β 를 계산하기 위해 사용되는데, 이는 합성 과정에서 주요 성분과 잔향 성분의 비중을 결정한다^[9]. 그림 2는 $IID(b)$ 와 $IC(b)$ 의 변화에 따른 주요 성분과 잔향 성분의 비중을 그림으로 비교하였다.

그림 2에서 검은선이 주요성분, 회색선이 잔향성분을 의미하며, IID별로 다른 마커로 표시하였다. 그림에서 볼 수 있듯이, 두 신호의 에너지가 비슷하고, IC가 0에 가까울수록 잔향 성분에 대한 비중이 커진다. 하지만 합성 과정에서 파라미터 정보들은 서브 밴드 단위로 적용되기 때문에 임의의 밴드 내에서 잔향이 불필요한 성분에도 똑같이 적용되어 음질 열화가 발생한다. 본 논문에서는 이러한 문제점을 분석하고, 이를 보완할 수 있는 합성 행렬식을 새로 제안하였다.

III. 제안된 PS의 합성 방법

1. 합성 행렬식의 문제점

일반적으로 잔향 성분은 시간 및 공간적으로 변화가 크지 않은 특징이 있지만, $D(\hat{X}_S(n,k))$ 는 다운 믹스 신호를 기반으로 부호화단에서 생성된 신호이므로 $D(n,k)$ 의 특

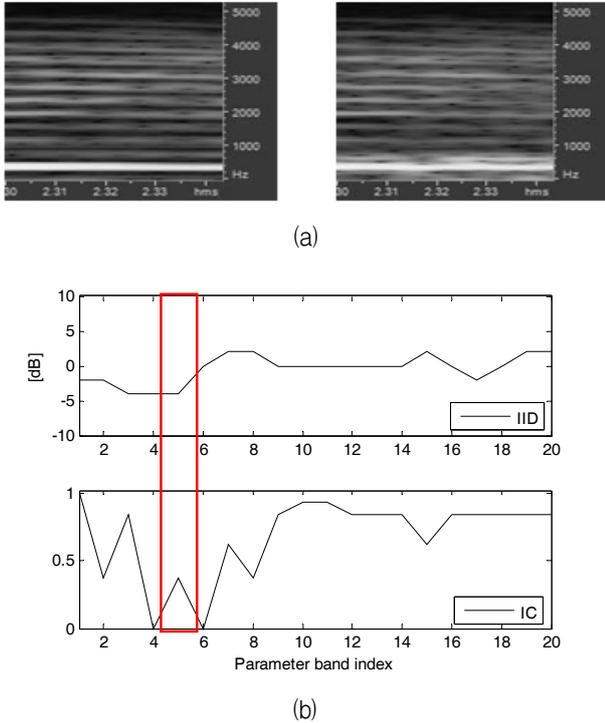


그림 3. (a) 왼쪽 채널의 입력과 출력 신호의 스펙트럼 그래프. (b) 임의의 프레임에서 분석된 IID와 IC.
 Fig. 3. (a) Spectrogram of the input and output signal of left channel. (b) Analyzed IID and IC at an arbitrary frame.

성과는 다르다. 특히 다운 믹스 신호의 특정 대역에 톤 성분이 존재하면, 이러한 특성이 새롭게 생성된 $D(\hat{X}_S(n,k))$ 에도 그대로 반영되어 출력 신호에 영향을 주게 되고, 이는 곧 음질 열화를 초래한다. 그림 3은 임의의 프레임에서 분석된 파라미터와 스펙트로그램을 나타낸 것이다.

그림 3 (a)의 스펙트로그램은 약 5kHz까지만 나타내었다. 그림 3 (a)를 보면 입력 신호에서 약 400Hz 대역에 존재하는 톤 성분 (빨간색 다각형)이 출력 신호에서는 많이 뭉개져 있음을 확인할 수 있다. 이러한 이유는 그림 3 (b)에 분석된 파라미터 값을 통해서 확인할 수 있는데, 해당 주파수 밴드 (5번째 밴드)의 $IC(b)$ 가 낮게 분석되어서 합성 과정에서 잔향 신호가 과도하게 합성되었기 때문이다.

톤은 잔향의 특성과 상반되기 때문에 오직 톤만 존재하는 주파수 대역에서는 $IC(b)$ 값이 1로 분석되어야 한다. 하지만 일반적으로 소리가 녹음될 때에는 배경 잔향 성분도 함께 녹음된다. 식 (2)에 정의된 신호 모델에서도 주요성분과 잔향 성분을 함께 고려하고 있다. 이러한

이유로, $IC(b)$ 는 임의의 주파수 밴드 내에 톤이 존재함에도 불구하고 1보다 낮은 값으로 분석될 수 있다.

2. 토널리티 기반의 합성행렬

본 논문에서는 합성 과정에서 톤 정보의 손실을 최소화하여 음질을 향상시키는 방법을 제안하였다. 제안된 기본적인 블록도는 그림 4와 같다.

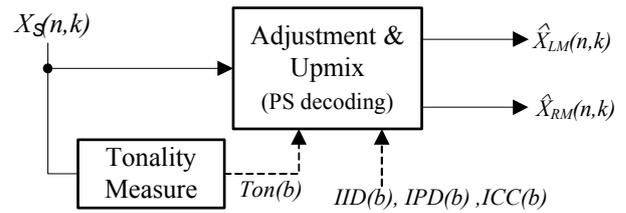


그림 4. 제안된 방법의 블록도.
 Fig. 4. Block diagram of the proposed method

먼저 입력 신호의 톤 정보를 확인하기 위해 다운믹스 신호의 토널리티를 측정하였다. 토널리티는 아래 식과 같이 spectral flatness measure(SFM) 방법^[8]을 이용하여 있으며 각 파라미터 밴드에 대해서 수행하였다.

$$SFM_{dB} = 10 \log_{10} \frac{\left[\prod_{k=k_b}^{k_{b+1}} |X_S(n,k)|^2 \right]^{\frac{1}{k_{b+1}-k_b+1}}}{\frac{1}{k_{b+1}-k_b+1} \sum_{k=k_b}^{k_{b+1}} |X_S(n,k)|^2}, \quad (6)$$

$$Ton(b) = \min\left(\frac{SFM_{dB}}{SFM_{dBmax}}, 1\right), \quad SFM_{dBmax} = 60dB.$$

다음, 그림 3에 보인 문제의 프레임에 대해서 토널리티를 측정하여 다운 믹스 신호의 파워 스펙트럼과 함께 그림 5에 나타내었다. 그림 5 (b)의 IC와 그림 5 (c)의 측정된 토널리티 인덱스를 입력 스테레오 신호의 Power Spectral Density(PSD)와 비교하기 위해 가로축 파라미터 밴드 인덱스를 해당 대역의 주파수 대역폭을 고려하여 주파수로 변환하여 나타내었다.

그림을 보면 톤이 존재하는 주파수 대역에서는 상대적으로 높은 토널리티가 측정되는 것을 확인할 수 있다. 하지만 앞장에서 언급했듯이, 합성 과정에서 사용되는 $D(\hat{X}_S(n,k))$ 는 IC가 낮으면 톤 성분이 있는 대역에서도 잔향 성분의 영향이 커진다. 그러므로 이러한 문제를 보완하기 위해 본 논문에서는 다음과 같은 합성

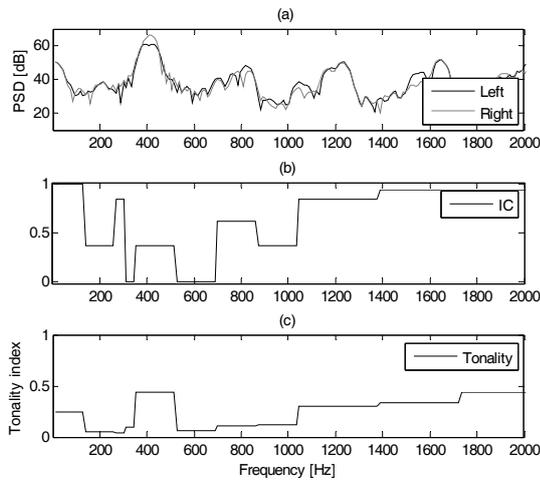


그림 5. (a) 입력 스테레오의 파워 스펙트럼, (b) 분석된 IC 값과 (c) 측정된 토널리티.

Fig. 5. (a) Power spectrum of the input stereo signal, (b) analyzed IC value and (c) measured tonality.

행렬식을 이용하였다.

$$\begin{bmatrix} \hat{X}_{LM} \\ \hat{X}_{RM} \end{bmatrix} = \begin{bmatrix} e^{jOPD(b)} & 0 \\ 0 & e^{-j(OPD(b)-IPD(b))} \end{bmatrix} \begin{bmatrix} \lambda_L & 0 \\ 0 & \lambda_R \end{bmatrix} \\ \times \begin{bmatrix} G_L(b)\cos(\alpha+\beta) & Ton(b)\sin(\alpha+\beta) \\ G_R(b)\cos(-\alpha+\beta) & Ton(b)\sin(-\alpha+\beta) \end{bmatrix} \begin{bmatrix} \hat{X}_S(n,k) \\ D(\hat{X}_S(n,k)) \end{bmatrix}. \quad (7)$$

식에서 $Ton(b)$ 는 임의의 밴드에서의 토널리티를 의미한다. 즉, 다운 믹스 신호로부터 측정된 토널리티에 따라서 잔향의 비중을 결정하였으며, 토널리티 값이 1에 가까울수록 잔향의 비중은 줄어들도록 설계하였다. 하지만 만약 $Ton(b)$ 을 통해 잔향 성분만 조절되면 전체적인 출력 신호의 에너지의 크기가 줄어든다. 그러므로 위 행렬식과 같이 신호의 에너지가 보존될 수 있도록 두 채널의 주요 성분에도 계인 $G_L(b)$ 과 $G_R(b)$ 을 곱해주었다. $G_L(b)$ 은 식 (8)과 같이 계산될 수 있으며, 같은 방법으로 $G_R(b)$ 를 구할 수 있다.

$$G_L^2(b)\cos^2(\alpha+\beta) + Ton^2(b)\sin^2(\alpha+\beta) = 1, \quad (8)$$

$$G_L(b) = \sqrt{\frac{1 - Ton^2(b)\sin^2(\alpha+\beta)}{\cos^2(\alpha+\beta)}}.$$

그림 6에서는 제안된 방법을 이용하여 구한 출력 신호의 스펙트로그램을 기존 방법을 이용한 결과와 함께 비교하였다. 그림 6 (b)에서 400Hz 대역에 톤 성분 주변

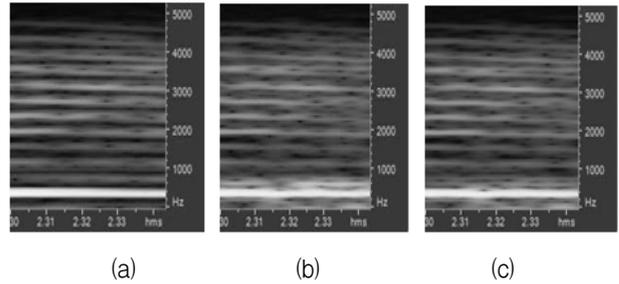


그림 6. 스펙트로그램 (a) 입력 신호, (b) 출력 신호 (PS), (c) 출력 신호 (제안된 방법)

Fig. 6. Spectrograms. (a) Input signal, (b) output signal (PS), (c) output signal (proposed method).

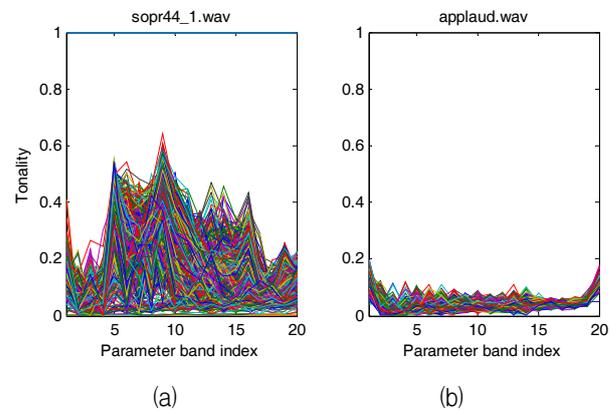


그림 7. 토널리티 측정. (a) 톤 성분이 많은 신호. (b) 잔향 성분이 많은 신호.

Fig. 7. Tonality measure. (a) Tone dominant signal. (b) Ambience dominant signal.

으로 존재하던 잔향 성분이 그림 6 (c)에서는 많이 완화되었음을 확인할 수 있다.

그림 7에서는 톤 성분이 많은 신호(sopr44_1.wav)와 톤 성분은 거의 없지만 잔향 성분이 풍부하여 공간감이 느껴지는 applaud 신호(applaud.wav)에 대해서 토널리티를 측정하였다. 그림 7 (a)의 경우, 특정 저주파 대역에서 높은 토널리티가 측정되었지만, 그림 7 (b) 대해서는 전 대역에 대해서 매우 낮은 토널리티가 측정되었다. 이는 제안된 합성 행렬식을 사용하더라도 성능 변화가 없음을 의미한다. 즉, applaud와 같이 잔향 특성이 절대적인 신호에 대해서는 잔향의 비중이 큰 영향을 주지 않기 때문에 성능 열화가 발생하지 않는다.

PS와 MPEG 서라운드 부호화기는 적용환경이 다르지만, 서론에서 간단히 언급했듯이 두 부호화기는 같은 패러다임을 갖고 개발되었기 때문에, 본 논문에서 제안된 방법은 두 부호화기에 모두 적용시킬 수 있다.

이외로, 합성 과정에서 본 논문에서 제안된 방법 대

신 개선된 잔향성분을 이용하더라도 부호화기의 성능을 향상시킬 수 있다. MPEG Surround 부호화기에서는 부호화단에서 잔여 신호(residual signal)를 추가적으로 생성하여 복호화단에서 해당 신호를 인공 잔향 성분 $D(\hat{X}_s(n,k))$ 대신 사용함으로써 발생 가능한 음질 열화를 보완해주고 있다^[5]. 하지만 이러한 과정은 추가적인 정보가 필요할 뿐만 아니라 전체 연산량에도 적지 않은 영향을 끼친다. 제안된 방법을 이용하면 추가 정보 없이 적은 연산으로도 음질 향상을 기대할 수 있다.

IV. 실험

제안된 방법의 성능을 평가하기 위해 주관적 음질 평가를 수행하였다. 총 5명의 숙련된 실험자가 참여하였으며, 모두 헤드폰(Sennheiser HD600)을 착용하고 실험을 진행하였다. 실험에는 톤 성분으로만 구성된 샘플 “glock”, Unified Speech Audio Coding(USAC)^[10] 표준화 과정에서 사용된 테스트 샘플(“te15”, “phi7”, sampling rate : 48KkHz)과 Lame’s 테스트 샘플(“pipes”, “sopr441”, sampling rate : 44.1kHz)들 중 톤 특성이 뚜렷한 네 개의 신호를 선별하여 총 5개의 샘플과 잔향 특성이 많은 “applaud” 샘플을 이용하였다.

실험은 MUSHRA^[11]방법으로 평가되었으며, 전체 평균과 95%의 신뢰구간을 계산하여 그림 8에 나타내었다. 실험 결과를 보면 모든 테스트 샘플에서 신뢰구간이 겹친다. 이는 제안된 방법이 기존 방법에 비해 두드러진 성능 향상을 보이지 않음을 의미한다. 하지만 그림에서 평균값은 전체적으로 기존 방법에 비해 높게 평

가된 것으로 보아, 기존 방법에 비해서 좀 더 향상된 성능을 기대할 수 있다.

실험 결과에서 tonal 성분이 적은 “applaud”는 큰 성능 향상을 보이지 않았다. 하지만, 샘플 “sopr44_1”의 경우 다른 실험 샘플들의 결과와 비교해서 상대적으로 가장 큰 성능차를 보이는데, 이는 톤 성분이 존재하는 주파수 대역에 낮은 IC가 분석되어 과도한 잔향 성분의 합성으로 음질의 열화가 발생하였지만, 제안된 방법을 통해 이러한 문제점을 효과적으로 해결하였기 때문이다. 다른 네 가지 실험 샘플에 대해서도 비슷한 결과를 나타내었다.

V. 결론

본 논문에서는 PS 오디오 부호화에서 잔향 합성에 대한 문제점을 분석하였고 이를 보완할 수 있는 방법을 제안하였다. PS는 복호화단에서 다운믹스 신호 기반의 잔향신호를 생성하여 합성하기 때문에, 톤 성분이 존재하는 대역에 IC가 낮게 측정되면, 출력 신호에 잔향 성분이 과도하게 합성되어 음질 열화가 발생한다. 이러한 문제를 보완하기 위해, 다운믹스의 토널리티를 측정하여 잔향 성분의 비중을 조절해주었다. 주관적 음질 평가 결과 제안된 방법을 이용하여 출력 신호의 음질을 향상시킬 수 있음을 확인하였다. 이러한 개선 과정에서 제안된 방법은 부호화단에서 어떤 추가 정보를 필요로 하지 않는 장점이 있다.

REFERENCES

- [1] T. Painter, A. Spanias, “Perceptual coding of digital audio,” Proc. of IEEE, vol. 88, no. 4, pp. 451-515, Apr., 2000.
- [2] ISO/IEC 11172-3, Coding of Moving Pictures and Associated Audio for Digital Storage Media up to about 1.5 Mb/s (part 3: MPEG-Audio), August 1993.
- [3] ISO/IEC 13818-7 Information Technology - Generic Coding of Moving Pictures and Associated Audio, Part 7: Advanced Audio Coding, 1997.
- [4] M. Wolters, K. Kjörling, D. Homm, and H. Purnhagen, “A closer look into MPEG-4 high efficiency AAC,” in Proc. 115th AES Convention,

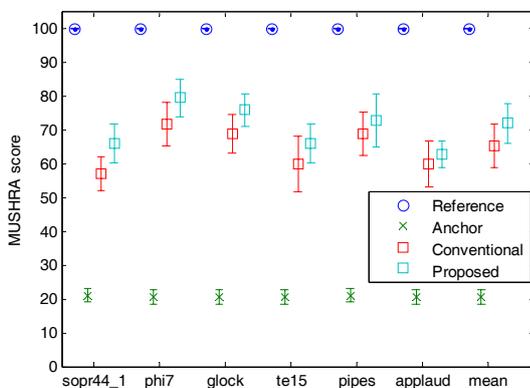
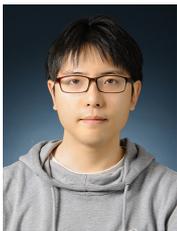


그림 8. 주관적 음질 평가 결과.

Fig. 8. Subjective evaluation results.

- New York, USA, October 2003.
- [5] 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, pp. 331-351, 2007.
- [6] 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.
- [7] G. Hotho, L. Villemoes, J. Breebaart, "A Backward-Compatible Multichannel Audio Codec," IEEE Trans. on Audio, Signal and Lang. Proc., Vol. 16, no. 1, pp. 83-93, Jan. 2008.
- [8] J.D. Johnston, "Transform coding of audio signal using perceptual noise criteria," IEEE J. on Sel. Areas in Comm., Vol. 6, no. 2, pp. 314-323, Feb., 1988.
- [9] H. Purnhagen: "Low Complexity Parametric Stereo Coding in MPEG-4," 7th Inter. Conf. on Audio Effects (DAFX-04), pp. 163-168, Naples, Italy, Oct., 2004.
- [10] M. Neuendorf, et al., "Unified Speech and Audio Coding Scheme for High Quality at Low Bitrates," Proc. ICASSP, pp. 1-4, Taipei, Taiwan, Apr., 2009.
- [11] Method for the Subjective Assessment of Intermediate Quality Level of Coding Systems 2003, ITU-R BS.1534-1.

— 저 자 소 개 —



이 동 금(학생회원)
2005년 연세대학교 전기전자
공학부 학사 졸업
2007년 연세대학교 전기전자
공학과 석사 졸업
2008년~현재 연세대학교 전기
전자공학과 박사과정

<주관심분야 : 오디오 신호처리, 오디오 부호화,
음성 부호화>



박 영 철(평생회원)
1986년 연세대학교 전자공학과
학사 졸업
1988년 연세대학교 전자공학과
석사 졸업
1993년 연세대학교 전자공학과
박사 졸업

현재 연세대학교 컴퓨터정보통신공학부 교수
<주관심분야 : 오디오 신호처리, 음성 신호처리,
적용 신호처리>



윤 대 희(평생회원)
1977년 연세대학교 전자공학과
학사 졸업
1979년 Kansas State University
석사 졸업
1982년 Kansas State University
박사 졸업

1985년~현재 연세대학교 전기전자공학과 교수
<주관심분야 : 오디오 신호처리, 음성 신호처리,
적용 신호처리>