

멀티채널 오디오 압축 코덱 음질의 객관적인 측정방법

최인용*, 전상배**, 성평모***

서울대학교 공과대학 전기·컴퓨터공학부

Objective measurement of spatial auditory quality for multi channel audio codecs

In Yong Choi*, Sang Bae Chon**, Koeng-Mo Sung***

School of Electrical Engineering and Computer Science

Seoul National University

E-mail : *ciy@acoustics.snu.ac.kr, **strlen@acoustics.snu.ac.kr, ***kmsung@acoustics.snu.ac.kr

Abstract

본 논문은 멀티채널 오디오 압축 코덱의 음질을 객관적으로 평가할 수 있는 시스템 및 파라미터에 관한 것으로, 멀티채널 오디오 신호로부터 양이입력신호(ear input signals)를 만들어내는 전처리 과정과 이 과정을 통해 출력되는 양이입력신호로부터 양이레벨차이왜곡(inter-aural level difference distortion)을 구하는 과정 및 양이레벨차이왜곡이 청취평가 결과와 일관적인 상관관계를 보임을 서술한다.

본 연구에 의하면 멀티채널 오디오 압축 코덱의 음질을 선별된 청취자에 의한 주관적인 평가와 통계처리 없이 객관적인 측정만을 통해 평가하는 것이 가능하며, 이를 사용하면 멀티채널 오디오 압축 코덱 개발자들이 시간, 경제적 부담 없이 자신이 개발한 압축 코덱의 음질을 간단하게 평가해볼 수 있다.

I. 서론

본 연구는 멀티채널 오디오 압축 코덱에 의해 부호

화 되었다가 복호화된 오디오 신호가 압축 전의 본래 신호에 대비하여 청각적 음질 열화가 발생한 정도를 청취자에 의한 주관적인 평가 없이 객관적인 측정을 통해 알아내기 위한 것으로, 연구의 결과로서 멀티 채널 오디오 신호로부터 청취 환경에서 양이에 입력되는 소리 신호와 유사한 두 채널의 양이입력신호를 산출하는 전처리 과정 및 이 양이입력신호로부터 주관 평가 결과와 높은 상관 관계를 보이는 출력 변수를 추출해내는 과정을 제안한다. 이는 단순히 압축 후 복구된 신호의 주파수별 에너지가 본래 신호에 대비해 왜곡된 정도를 수치화한 것과는 다르며, 멀티채널 오디오 재생 환경에서 선별된 청취자가 느끼는 청각적인 평가를 적절한 방법으로 통계 처리한 것과 유사한 결과를 얻기 위한 것이다.

단일 채널 오디오 신호 압축 코덱에 대해서는 객관적 음질 평가 방법에 관한 연구가 오랜 기간 진행된 바 있으며 ITU-R 의 권고안이 채택된 바 있다.[1] (이하 이를 Perceptual Evaluation of Audio Quality 의 약자인 PEAQ 로 표기) 그러나 이는 공간감 인자를 측정하지 않아 특히 공간감 면의 음질이 중요한 멀티 채널 오디오 압축 코덱에는 사용할 수 없다는 한계를 지니고 있다.

한편 평가의 대상이 될 멀티 채널 압축 코덱은 MPEG 표준화 그룹(ISO/IEC JTC1/SC29/WG11)에서 개발

본 연구는 한국전자통신연구원의 지원하에 이루어졌음.

논의가 활발히 진행 중이며, 여러 기관에서 개발한 코덱이 발표되었는데 이 코덱들의 음질 평가는 역시 ITU-R 권고안인 MUSHRA 에 기반한 청취 주관 평가 기법 [2]을 통해 이루어지고 있다.

이 방법을 이용해 수행한 다수의 코덱들에 대한 청취 평가 결과가 발표되었는데 [4][5] 본 발명에서 제안하는 객관적 음질 평가 방법에 의해 추출한 양이레벨차이 왜곡은 MPEG 에 기고된 해당 코덱들의 주관 음질 평가 결과와 높은 상관관계를 보인다.

II. 전처리 과정 및 양이레벨차이 왜곡의 산출

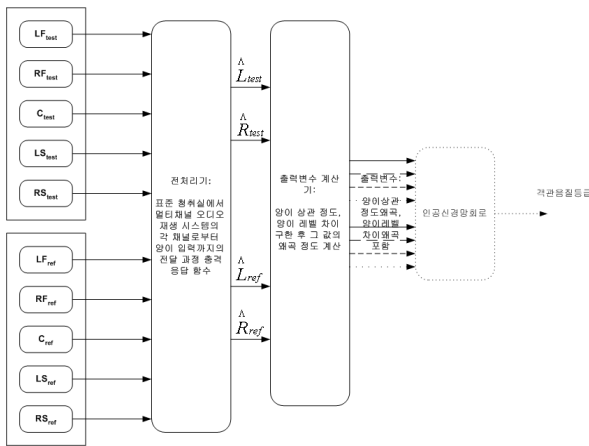


그림 1. 멀티채널 오디오 압축 코덱 음질 평가 시스템 전 과정의 블록도

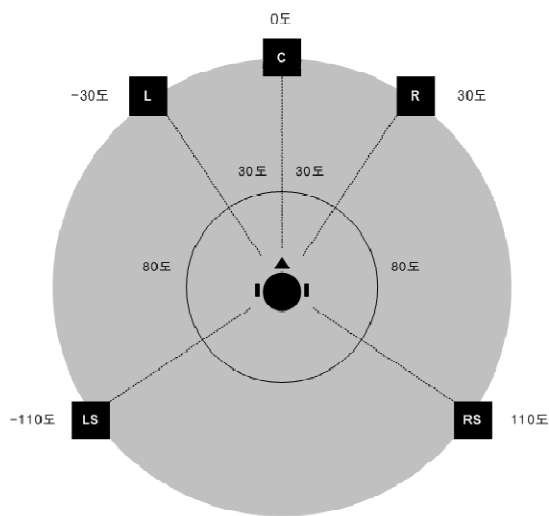


그림 2. ITU-R의 멀티채널 오디오 재생 시스템 배치 권고안

특정 멀티채널 오디오 신호가 있고, 이 오디오 신호를 우리가 평가하려는 압축 코덱을 이용해 부호화 및 복호화한 신호가 있다. 현재 일반적으로 멀티채널 오디오 오는 좌측전방(이하 LF 로 표기), 우측전방(이하 RF 로 표기), 중앙(이하 C 로 표기), 저역효과음(이하 LEF 로 표기), 좌측 후방(이하 LS 로 표기), 우측 후방(이하 RS 로 표기)의 6 채널(혹은 5.1 채널)을 가지는데, 이 중 LEF 채널은 실제 사용하지 않는 경우가 많으므로 여기서는 LF, RF, C, LS, RS 의 다섯 채널만을 사용하기로 한다. 본래 음원의 다섯 채널은 LF_{ref} , RF_{ref} , C_{ref} , LS_{ref} , RS_{ref} 로 칭하며 평가할 코덱에 의해 부호화 및 복호화된 음원의 다섯 채널은 LF_{test} , RF_{test} , C_{test} , LS_{test} , RS_{test} 라 칭한다.

그림 1 과 같은 음질 평가 시스템 구조에서 최초 입력 부분을 보면 LF_{ref} , RF_{ref} , C_{ref} , LS_{ref} , RS_{ref} , LF_{test} , RF_{test} , C_{test} , LS_{test} , RS_{test} 등 총 10 개의 신호는 전처리로 입력된다. 이 전처리는 ITU-R 에서 권고하는 표준 멀티채널 오디오 재생 환경에서 바람직한 재생 시스템과 인체(두부 및 상반신)를 모사한 양이 마이크로폰을 이용해 측정된 소리 전달 경로의 충격 응답이 전달함수화 되어 포함되어 있다. ITU-R 에서 권고하는 멀티채널 오디오 재생 시스템 배치는 그림 2 와 같으며, 이 때 총 소리 전달 경로는 10 개인데, 이 10 개의 소리 전달 경로의 충격응답 그래프는 그림 3 과 같다.

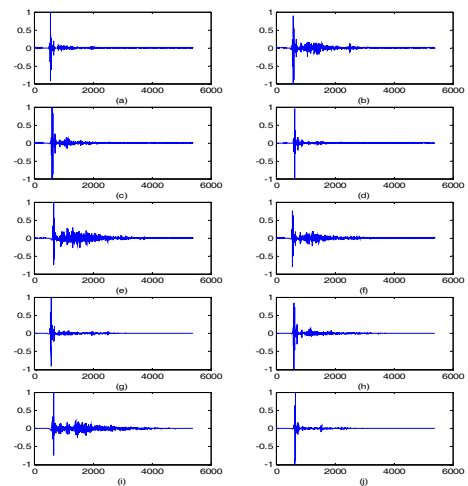


그림 3. 멀티채널 오디오 시스템 표준 청취실에서 측정된 10 가지 청각 전달 경로의 충격응답. (a) LF 채널에서 왼쪽 귀 (b) RF 채널에서 왼쪽 귀 (c) C 채널에서 왼쪽 귀 (d) LS 채널에서 왼쪽 귀 (e) RS 채널에서 왼쪽 귀 (f) LF 채널에서 오른쪽 귀

(g) RF 채널에서 오른쪽 귀 (h) C 채널에서 오른쪽 귀 (i) LS 채널에서 오른쪽 귀 (j) RS 채널에서 오른쪽 귀

이 소리 전달 경로 이후 양이입력신호를 구해내는 과정은 그림 4 와 같다. 이런 과정을 통해 두 가지의 양이입력신호(L_{ref} , R_{ref} 및 L_{test} , R_{test})가 출력된다.

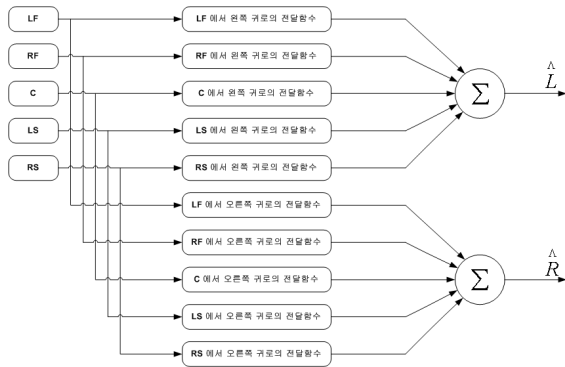


그림 4. 멀티채널 음원으로부터 양이입력신호를 구해내는 과정의 블록도

전처리에서 출력된 두 가지 양이 입력 신호는 출력변수계산기로 입력된다. 이 출력변수계산기는 양이레벨차이왜곡을 계산해 출력하며, 여기에서 잡음강도 (Noise Loudness: NLA), 오류신호조화구조(Harmonic Structure of Error: EHS) 등 PEAQ 에서 사용하는 여타 다른 출력 변수가 계산될 수도 있다.

양이레벨차이는 양쪽 귀로 입력되는 신호의 에너지 비율인데, 흔히 ILD(Interaural Level Difference)라 칭하며, 이하 서술에서 이 표기를 따른다. 이 ILD 는 방향감 인지 및 이를 바탕으로한 공간감 인지에서 중요한 역할을 하는 것으로 알려져 있다.[3] 여기서 압축하지 않은 본래 오디오 신호의 ILD 를 ILD_{ref} 라 하고 압축 코덱에 의해 부호화 및 복호화된 오디오 신호의 ILD 를 ILD_{test} 라 하기로 한다. ILD 는 44100Hz 의 주파수로 샘플링되었을 때 2048 개의 샘플을 50%씩 중첩하며 진행하는 시간 프레임에서 24 개의 청각 임계대역에 대해 각각 계산하였다. 이후 동일 시간 프레임에서 ILD_{ref} 와 ILD_{test} 의 차이를 계산하였으며 이를 양이레벨차이왜곡(Interaural Level Difference Distortion, 이하 ILDDist)이라 하기로 한다. 한 개의 시간 프레임에서 전 청각 대역의 ILDDist 를 구하기 위해 24 개의 청각 임계대역에 대한 ILDDist 를 가중 평균한다. 이를 전 시간 프레임에 대해 평균함으로써

해당 멀티채널 압축 코덱의 ILDDist 를 계산할 수 있다.

III. 양이레벨차이왜곡과 청취평가결과 사이의 상관관계

그림 5 는 72 차 MPEG 표준화 회의에서 발표된 4 가지 멀티채널 오디오 압축 코덱 음질의 주관평가결과 (Mean Opinion Score: MOS)[4][5]와 앞에서 서술한 방법으로 측정된 ILDDist 간의 상관관계를 나타낸 그래프이다. 5 가지의 표준 음원(Chostakovitch, fountain music, indie2, jackson1, pops)에 대해 측정하였으며 각각의 음원은 오케스트라의 연주, 클래식 소품의 연주, 영화음악, 대중음악 등 여러가지 장르의 특징적인 부분을 약 20 초 길이의 표본으로 만들어 MPEG 에서 평가 용도로 사용하고 있는 것이다. 비교 결과 주관평가에서 높은 점수를 받은 음원일수록 ILDDist 가 작은 것으로 나타났으며, 이는 ILDDist 가 객관적인 음질 측정을 위한 하나의 변수가 될 수 있음을 의미한다.

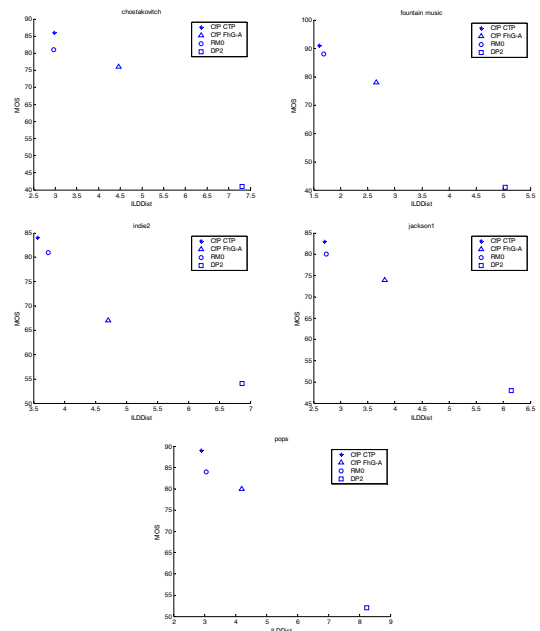


그림 5. 다섯가지의 음원에 대해 CfP_CTP, CfP_FHG-A, RM0, Dolby Pro-logic II 등 네 가지의 멀티채널 압축 코덱의 주관 평가 결과 [4][5]와 각각의 ILDDist 값과의 비교. 그래프는 Chostakovitch, fountain music, indie2, jackson1, pops 의 순서로 나타내었으며 주관평가결과는 MOS 로 표시하였는데 이는 100 점 만점의 범위에서 점수로 나타내며 ILDDist 는 dB 단위로 표시하였다.

한편 양이입력신호에 대해 PEAQ 에서 사용하는 다섯가지 음질 측정 변수(Rms Modulation Difference, Noise Loudness, Averaged Linear Distortion, Noise-to-Mask Ratio, and Harmonic Structure of Error)를 측정해본 결과 ILDDist 와 마찬가지로 모두 MOS 값에 대해 음의 상관관계를 보였다. (그림 6) 이 값들과 함께 ILDDist 를 인공지능망 회로에 입력하면 공간감을 포함한 멀티채널 오디오 음질의 객관음질등급(Objective Difference Grade)을 출력할 수 있을 것으로 보인다.

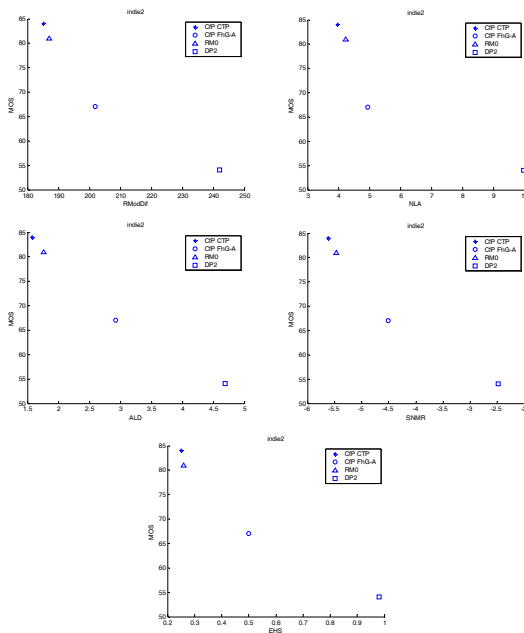


그림 6. CfP_CTP, CfP_FHG-A, RM0, Dolby Pro-logic II 등 네 가지의 멀티채널 압축 코덱의 주관평가 결과[4][5]와 PEAQ의 다섯가지 출력변수 측정값과의 비교. Rms Modulation Difference, Noise Loudness, Averaged Linear Distortion, Noise-to-Mask Ratio, and Harmonic Structure of Error의 순서로 나타내었고 음원은 indie2를 사용하였다.

IV. 결론

서론에서 언급하였듯이 단일 채널 오디오 신호 압축 코덱에 대해서는 객관적 음질 평가 방법에 관한 연구가 오랜 기간 진행되어 이미 PEAQ가 개발되었고 이것이 ITU-R의 권고안으로 채택되었다.[1] 그러나 이는 멀티 채널 오디오 압축 코덱에는 사용할 수 없으며 PEAQ에서 측정하는 변수들은 주파수 분포의 청각적인 밸런스 및 잡음 유무 등을 판단하는 것에 한정되어 있

어 멀티채널 오디오에서 특히 중요하다고 할 수 있는 공간감을 측정하지 못한다. 본 논문을 통하여 제안되는 ILDDist는 공간감 형성에 중요한 파라미터로 알려진 ILD가 왜곡된 정도를 측정하는 값으로, 주관평가 결과와 일관된 음의 상관관계를 가진다는 것이 입증되었으며 멀티채널 오디오 압축 코덱의 객관적 음질 평가를 위한 주요 측정 변수로 사용될 수 있다. 한편 본 논문에서는 멀티채널 오디오 신호로부터 양이입력신호를 구해내는 전처리 과정 역시 제안하였는데 이를 통해 ILD 및 ILDDist를 측정할 수 있으며 PEAQ에서 사용중인 음질 변수 역시 제안된 전처리 과정을 거친 양이입력신호를 사용해 측정하는 것이 보다 합리적이다.

참고문헌

- [1] ITU-R Recommendation BS. 1387-1, "Method for objective measurements of perceived audio quality", International Telecommunication Union, Geneva, Switzerland, 1998
- [2] ITU-R Recommendation BS. 1534-1, "Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA)", International Telecommunication Union, Geneva, Switzerland, 2001
- [3] J. Blauert, "Spatial Hearing: The Psychophysics of Human Sound Localization," MIT press, 1983
- [4] ISO/IEC JTC1/SC29/WG11 (MPEG), N7138, "Report on MPEG Spatial Audio Coding RM0 Listening Tests"
- [5] ISO/IEC JTC1/SC29/WG11 (MPEG), N7139, "Spatial Audio Coding RM0 listening test data"