

압축을 변화에 따른 멀티채널 오디오의 품질 및 Predistortion의 영향 평가

Quality Assessment and Predistortion Evaluation of the Multi-channel Audio Codec according to the bitrate changing

차 경 환*, 장 대 영**, 김 성 한**, 김 천 덕***

(Kyung-Hwan Cha*, Dae-Young Jang**, Sung-Han Kim**, Chun-Duck Kim***)

요 약

본 논문에서는 멀티채널 오디오의 전송 및 재생에 있어서 압축을 변화에 따른 음질의 주관 평가와 matrixing과 dematrixing 과정에서 발생하는 가청잡음을 보상해 주는 predistortion의 영향을 평가하였다. 시뮬레이션은 지각 부호화를 이용하는 MPEG-2 오디오 계층2의 알고리즘을 사용하였으며, 압출율은 384, 320, 256, 128kbps로 변화시키면서 predistortion의 유/무에 따른 음질의 영향을 평가하였다. double blind 법에 의한 주관 평가 결과 3/2채널에서 5점 열화 척도가 320kbps까지는 -1이하로 원음과 차이가 없거나 거슬리지 않는 것으로 평가되었으며, predistortion의 영향은 128kbps에서 최고1정도 품질이 향상되었으며, 특히 음성시료가 음악시료보다 더욱 향상된 결과를 얻었다.

ABSTRACT

This paper describes the subjective assessment of the multi-channel audio quality according to the bitrate changing and evaluates the predistortion effect to avoid the unmasked noise after matrixing/dematrixing process in transmission and regeneration of the multi-channel audio. The simulation is processed by the perceptual coding that is MPEG-2 Audio layer II algorithm. We evaluate the quality improvement about predistortion using or not by 384, 320, 256, 128 kbps. As the result of the double blind subjective assessment, 5 Grade-Impairment Scale is scored under minus one to 320kbps and so audio quality is evaluated to be perceptible, but not annoying in 3/2 channel. The effect of the predistortion is improved one level in 128kbps and especially speech test material is better improved than music test materials.

I. 서 론

최근의 ISDN 및 VLSI 기술의 발전에 의해서 광대역 음향신호의 부호화는 디지털 오디오, 디지털 음향방송, 고선명 텔레비전 및 멀티미디어 등에서 그 중요성이 증대되고 있다. 고품질의 멀티채널 오디오의 도입과 더불어 효율적이고, 경제적인 방법으로 낮은 비트율로 멀티채널의 입체음향 오디오를 제공하기 위한 다양한 신호 압축 기술들이 활발히 연구되고 있으며, 특히 인간의 음향심리모형을 이용한 압축 부호화 기술의 연구가 활발히 진행되고 있다[1].

MPEG-2 오디오 표준은 고품질의 디지털 서라운드음과 다중언어 오디오 정보를 부호화 할 수 있도록 심리음

향 현상을 이용하여 저비트율의 멀티채널 오디오에 대한 압축 부호화 알고리즘을 기술한 것이다. 이 표준의 기본 특징은 MPEG-1 오디오를 표준으로 부호화된 모노, 스테레오 또는 듀얼채널 오디오 프로그램에 대한 역방향 호환성(backward compatibility), 멀티채널 오디오의 전송 및 재생, 저주파수 효과 채널 등을 제공하는 것이다. 멀티채널 오디오의 압축 부호화에서는 모든 채널의 정보를 효율적으로 수용하기 위한 공통비트 운용(common bit pool), 결합 스테레오 부호화, 공통 마스크 한계(common masking threshold), 서라운드 신호의 동적누화(dynamic crosstalk), matrixing과 dematrixing 과정에서 발생하는 코딩왜곡을 보상해 주는 predistortion 등의 기법이 추가됨에 따라서 고품질을 유지하면서 효율적으로 압축이 가능하다[2].

본 논문에서는 멀티채널 오디오의 전송 및 재생을 위한 소프트웨어를 MPEG-2 오디오 계층2의 알고리즘으

* 동서대학교 정보통신공학과
** 한국전자통신연구소 음향통신연구실
*** 부산수산대학교 전기공학과
접수일자: 1996년 2월 21일

로 구현하고, 3/2채널에서 어느 압축율까지 원음과 동일한 품질로 어느 압축율까지 평가되는지를 double blind 법에 의해 주관 평가하였다. 또한 제한된 압축율상에서 동일하지 않은 채널신호에 의한 matrixing과 dematrixing 이후 각 채널의 양자화 오차의 조합에 의해 발생하는 가청잡음을 마스크하는데 소요되는 비트들에 의한 음질저하가 어느 압축율까지 영향을 받지 않는지, 그리고 가청잡음을 감소시키는 predistortion 기법을 사용하였을 경우에 압축율의 변화에 따라서 어느 정도 멀티채널의 음질에 영향을 주는지를 청취실에서 10명의 청취자를 대상으로 평가하였다.

II. Predistortion을 이용한 멀티채널 오디오 부호화

2.1 Predistortion에 의한 가청잡음 보상기법

멀티채널 오디오의 인코더와 디코더에서 matrixing/dematrixing 과정을 통해서 복원된 신호는 명료한 음질 (transparent quality)을 유지한다고 가정한다. 그러나 이러한 가정은 matrixing/dematrixing 과정에서 사용되는 신호가 동일한 신호이어야만 성립한다. 그러나 matrixing/dematrixing 사이의 양자화과정에 의해 발생하는 각 채널의 코딩왜곡의 조합에 의해 dematrixing된 이후의 신호는 matrixing이전의 신호와는 차이가 생기게 된다. 이러한 가청잡음은 matrixing과정 이전에 미리 양자화와 역양자화를 거친 신호들을 사용함으로써 감소시킬 수 있는데 이러한 과정을 predistortion이라고 한다.

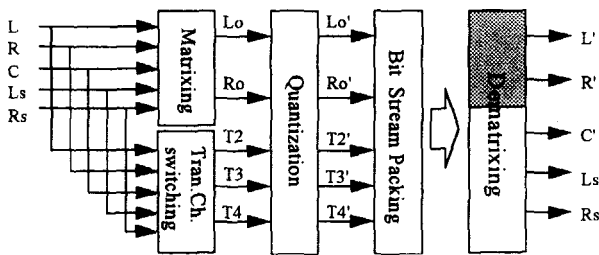


그림 1. L, R, C 채널의 전송 블록도

예를들어 그림 1.에서와 같이 L, R, C 채널이 전송채널 할당에 의해 확장채널(T₂, T₃, T₄)에 할당되었다고 가정하면 은 식(1)에서와 같이 matrixing에 의해 T₀, T₁에 할당된다. 그러면 L₀, R₀, L, R, C의 5채널은 식(2)에서와 같이 L₀', R₀', C', L', R'의 형태로 양자화 되어진후 디코더에 비트스트림으로 전달된다. 여기서 e()는 양자화 오차를 나타낸다.

$$\begin{aligned}
 T_0: L_0 &= L + \alpha C + \beta L_s \\
 T_1: R_0 &= R + \alpha C + \beta R_s \\
 \text{attenuation factor } \alpha &= 1/\sqrt{2}, \beta = 1/\sqrt{2} \quad (1)
 \end{aligned}$$

$$\begin{aligned}
 T_0: L_0' &= L_0 + e(L_0) \\
 T_1: R_0' &= R_0 + e(R_0) \\
 T_2: C' &= C + e(C) \\
 T_3: L' &= L + e(L) \\
 T_4: R' &= R + e(R) \quad (2)
 \end{aligned}$$

전송채널할당에 의해 전송되지 않은 2개의 채널 L_s, R_s은 디코더의 dematrixing과정에 의해서 식(3)과 같이 복원된다.

$$\begin{aligned}
 L_s' &= (L_0 - \alpha C - L')/\beta \\
 R_s' &= (R_0 - \alpha C - R')/\beta \quad (3)
 \end{aligned}$$

식(3)에 식(2)을 대입하여 인코더에서 전송되지 않은 2개의 채널 L_s, R_s을 복원시킨 식(4)에는 다른 채널의 양자화 오차가 포함되어져서 왜곡되어져 있는 것을 알수있다. 따라서 이러한 왜곡성분들의 조합에 의해 마스크되지 않는 가청잡음이 들리게 된다.

$$\begin{aligned}
 L_s' &= L_s + e(L_0)/\beta - e(L)/\beta - \alpha e(C)/\beta \\
 R_s' &= R_s + e(R_0)/\beta - e(R)/\beta - \alpha e(C)/\beta \quad (4)
 \end{aligned}$$

그림 2.는 이러한 가청잡음을 감소시키기 위한 predistortion과정을 나타내고 있다. 전송채널할당에 의해 선택된 L, R, C 채널의 신호를 양자화/역양자화 시킨 후 matrixing을 수행하면 식(1)이 식(5)와 같이 양자화/역양자화에 의한 양자화 오차가 포함된 형태로 양자화되고, 전송채널할당에 의한 L, R, C 채널은 양자화된 T₂', T₃', T₄'의 형태로 디코더에 비트스트림으로 전송된다.

$$\begin{aligned}
 L_0 &= L + e(L) + \alpha C + e(\alpha C) + \beta L_s \\
 R_0 &= R + e(R) + \alpha C + e(\alpha C) + \beta R_s \quad (5)
 \end{aligned}$$

따라서 디코더의 dematrixing과정을 거치면서 전송되지 않은 2채널 은 식(3)에 의해 복원시키면 식(6)으로 표현되어진다.

$$\begin{aligned}
 L_s' &= (L_0 + e(L_0) - \alpha[C + e(C)] - [L + e(L)])/\beta \\
 &= \{L + e(L) - \alpha(C) + e(\alpha C) + \beta L_s + e(L_0) \\
 &\quad - \alpha C - e(\alpha C) - L - e(L)\}/\beta \\
 R_s' &= (R_0 + e(R_0) - \alpha[C + e(C)] - [R + e(R)])/\beta \\
 &= \{R + e(R) - \alpha(C) + e(\alpha C) + \beta R_s + e(R_0) \\
 &\quad - \alpha C - e(\alpha C) - R - e(R)\}/\beta \quad (6)
 \end{aligned}$$

식(6)을 간단히 하면 predistortion과정을 수행한 후, 복원된 L_s, R_s은 식(7)로 나타낼 수 있다. predistortion과정을 고려하지 않은 식(4)와 비교해 보면 다른 채널의 양자화 오차가 제거된 형태로 코딩왜곡이 감소되는 것을

알 수 있으며, 이러한 왜곡성분들의 조합에 의한 가청잡음이 저감된 것을 확인할 수 있다.

$$L'_s = L_s + e(L_0)/\beta$$

$$R'_s = R_s + e(R_0)/\beta \quad (7)$$

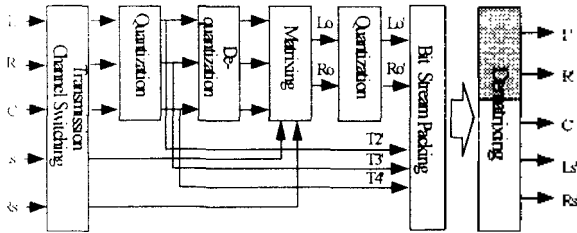


그림 2. Predistortion과정을 고려한 L, R, C채널의 전송 블록도

2.2 Predistortion기법을 이용한 부호화

멀티채널 오디오의 인코딩 및 디코딩은 인간의 청각적 특성을 이용하여 감도가 낮은 부분의 정보를 생략해서 부호량을 감소시키는 지각 부호화(perceptual coding) 방식을 이용하는 MPEG-2 오디오 계층2의 알고리즘으로 구현하였다. 인코더 및 디코더의 처리 블록 구성도는 그림 3과 그림 4와 같다.

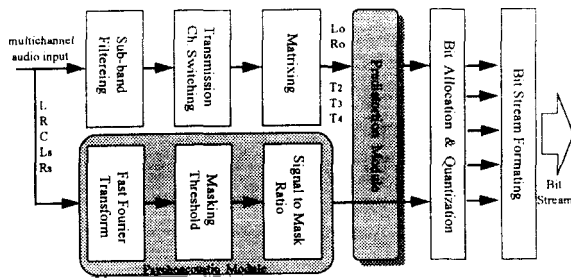


그림 3. Predistortion 기법을 사용한 멀티채널 오디오의 인코더 흐름도

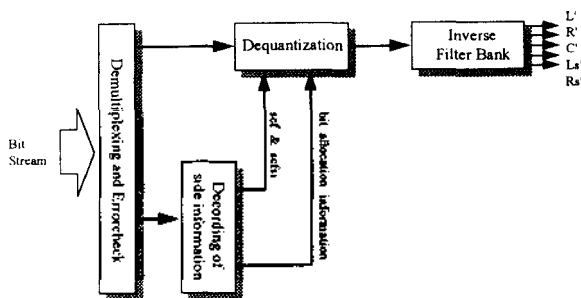


그림 4. 멀티채널 오디오의 디코더 흐름도

멀티채널 오디오의 부호화는 기본적으로 모노 또는 스테레오 오디오의 부호화와 차이가 없으며, 특히 멀티채널의 효율적 부호화를 위한 처리가 추가된다. 먼저 심리음향모델에 의한 SMR(Signal-to-Masking Ratio)값을 계산하는데, 1024샘플로 구성되는 프레임에 대해 FFT 분석을 수행하여 PDS(Power Density Spectrum)를 측정정한 후 신호의 주기성(tonal) 성분과 비주기성(non-tonal) 성분을 구하고, 개별 마스킹 역치와 전체마스킹 역치를 계산하고, 각 서브밴드의 최대 신호레벨과 최소 마스킹 역치에 의해 32개의 서브밴드에 대한 SMR값을 계산한다. 다위상 필터뱅크를 사용하여 각 채널의 1152개의 입력 샘플을 받아서 서브밴드당 36개의 샘플을 갖는 32개의 서브밴드를 출력한다. 각 서브밴드에 대한 척도계수를 구하여 각 블록의 샘플이 효율적으로 양자화될 수 있도록 정규화하는데 사용한다. 5개의 멀티채널(L, R, C, L_s, R_s) 신호로 부터 음향효과를 증가시키는 전송채널할당(transmission channel switching)은 T₀는 T₀로, R₀는 T₁으로 할당하고, 나머지는 구성방법에 따라서 T₂~T₄에 할당한다. Matrixing과정에서는 스테레오 신호를 계산할때 과부하를 줄이기 위해 신호를 감쇄시키고, 역방향 호환성을 위해 downmixing하여 L₀, R₀를 계산한다. Predistortion 기법을 이용하지 않을 경우에는 T₀~T₄의 멀티 전송채널에 사용가능한 전체 비트수에 NMR(Noise-to-Masking Ratio)이 작은 순서대로 각 서브밴드의 채널별로 한 비트씩 할당하는 과정을 반복한다. 만약 predistortion 기법을 이용하는 경우에는 앞에서 설명한 그림 2의 절차를 수행한후 비트 할당과 비트 스트림을 생성하게 된다.

멀티채널 오디오 복호화는 그림 4에서와 같이 해더와 비트할당 정보, CRC, scf, scfsi, 멀티채널 오디오 데이터 등으로 전송된 비트스트림을 분해하고, dematrixing 과정을 통해서 인코더에서 전송되지 않은 2개의 채널을 복원시킨 후 역필터뱅크에 의해 멀티채널 오디오 신호(L', R', C', L_s', R_s')를 재생하게 된다.

III. 시뮬레이션

압축률 변화에 따른 멀티채널 오디오의 품질과 predistortion의 유/무에 따른 음질향상 정도를 평가하기 위한 시뮬레이션의 환경은 그림 5와 같다. 5채널의 소스를 멀티채널 오디오 기록 및 재생 시스템(MAPS: Multi-channel Audio Processing System)으로 디지털 오디오 데이터로 변환시킨 후 개인용 컴퓨터의 보조기억장치에 저장한다. 이 데이터들은 LAN에 접속된 워크스테이션으로 전송되어져서 개발한 압축 부호화 프로그램에 의해 384, 320, 258, 128 kbps의 압축률 변화와 각 압축율에서 predistortion의 유/무에 따른 시뮬레이션을 수행하였다. 워크스테이션에서 인코딩된 비트스트림은 디코딩 처리된 후 LAN을 통해서 개인용 컴퓨터에 저장되고, 멀티채

년 오디오 기록 및 재생 시스템에 의해 청취실에서 3/2채널의 스피커 시스템으로 멀티채널 오디오 데이터를 재생하였다.

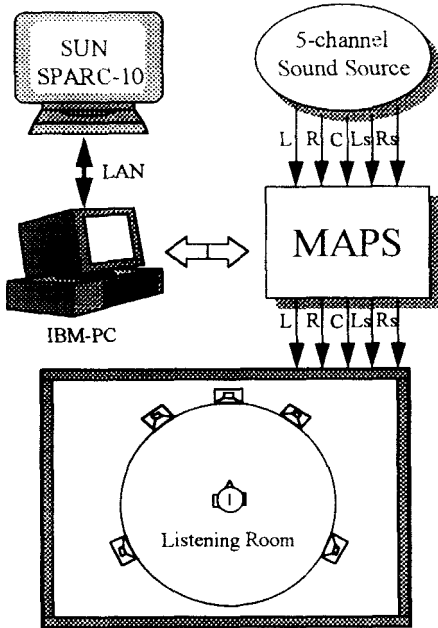


그림 5. 멀티채널 오디오 품질 평가를 위한 시뮬레이션 환경

탐의 주관적인 평가 결과들을 상호 비교 검증할 수 있도록 권고하고 있는 내용을 고려하여 그림 6과 같은 청취조건을 설정하였다. 또한 실험에 사용된 청취실은 국제적인 단간이 되고있는 ITU-R의 TG 10-3(subjective assessment)과 EBU 전문가 그룹 G1/1IST(subjective assessment and listening condition)의 규격에 맞게 설계된 청취실(간향시간: 0.2초, 실내용적: 4.7m X 6m X 3m, 초기반사음: -14dB, 암소음: 40dBA)에서 행하여졌다.

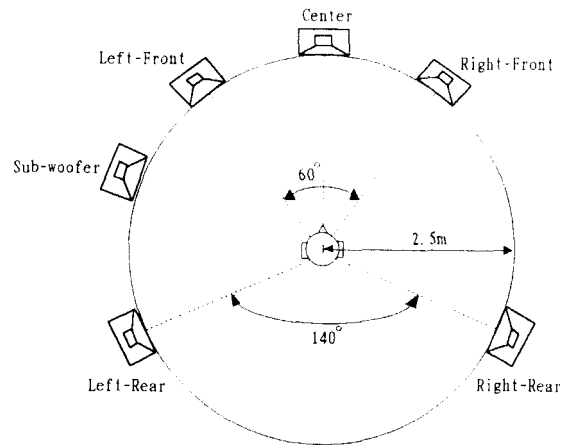


그림 6. 멀티채널 오디오의 주관평가를 위한 스피커 배치와 청취위치

IV. 멀티채널 오디오의 주관 품질 평가

4.1 시험재료의 선정 및 청취조건

멀티채널 오디오의 주관 평가에 사용되는 멀티채널 오디오 소스는 독일방송연구소(IRT)에서 제작한 멀티채널 오디오 품질 평가용 음악중 5곡을 선정하였으며, 표 1은 주관 평가에 사용된 시험음악을 나타내고 있다.

표 1. 멀티채널 오디오의 주관 평가에 사용된 시험음악

No.	내용	시간(secs)
1	영화 음악(Indiana Johnse)	13
2	영화 효과 음악(Night Bird)	12
3	영화 효과 음악(Ski Jump)	13
4	테니스게임 중계 방송	15
5	영어 남자 음성	8

주관 평가에 있어서 기준 청취위치에 청취자가 위치하고, 기준 스피커 배열에 의한 기준 음장이 만들어 질 때, 청취자에 영향을 미치는 기준 음장에 대한 복잡한 음향적인 요구조건인 기준 청취조건(reference listening conditions)은 그 평가결과를 상호 비교하기 위해서도 상당히 중요하다. 본 실험에서는 ITU-R에서 사운드 시스

4.2 주관 품질 평가방법 및 척도구성

주관 평가 실험에서는 “double blind triple stimulus with hidden reference” 방법에 의해서 평가 하고자 한다. 여기서 double blind 기법이란, 통상 피험자만 자극의 제시순서나 내용을 모르게 하는 single blind 기법과는 달리, 피험자는 물론 실험을 관리하는 실험자도 어떤 시험음악이 제시되는가를 사전에 알 수 없게 하여, 실험 결과에 피험자나 실험자의 의도가 배제되도록 하는 기법을 말한다[6].

이 방법에서 실험자는 사전에 5개의 시험음악 모두에 대해 각각 2개의 원음과 1개의 부호화된 음악을 준비한다. 항상 원음(이하 A로 표시)을 먼저 제시하고, 다음에 사전에 불류칙하게 할당된 부호화된 음악(이하 B 또는 C)이나 원음(이하 B 또는 C)을 각각 제시하여, B 또는 C가 A에 비해 얼마나 열화되었는지 또는 품질이 상대적으로 어느 정도인지를 청취자로 하여금 표 2에 나타난 연속적인 5등급-열화 척도(Continuous 5 Grade-Impairment Scale)에 따라 반응지에 평가하게 한다. 즉, 5등급-열화 척도인 경우에는 식(8)과 같이 차이점수를 계산하여, 그 결과를 그래프 형태로 나타낸다:

$$\text{차이점수(Diff-Grade)} = \text{숨겨진 원음의 점수} - \text{부호화된 음의 점수} \quad (8)$$

표 2. ITU-R의 5점-열화척도와 대응되는 차이점수

손상 정도	점수	대응되는 차이점수
느껴지지 않는다	5.0	0
느껴지지만, 거슬리지 않는다	4.0	-1
약간 거슬린다	3.0	-2
거슬린다	2.0	-3
매우 거슬린다	1.0	-4

4.3 실험결과 및 고찰

그림 7은 double blind 기법에 의한 원음과 압축율 변화에 따른 멀티채널 오디오의 평균차이점수를 나타내고 있다. 시험음악 5곡에 대한 평균 차이점수가 압축율 320kbps 까지는 모두 1 이하로 멀티채널 오디오의 재생 음질이 원음과 차이를 느끼지 못하거나 느끼지만 거슬리지 않는 정도로 평가되었다. 그러나 압축율 128kbps에서는 평균 차이점수가 3수준으로 귀에 거슬리는 것으로 평가되었다. 따라서 양자화 잡음이 마스킹 효과에 의한 전체적인 마스킹 곡선 위쪽으로 올라와서 마스킹 되지 못하고 가청잡음으로 들릴 수 있는 일체 압축율이 256kbps 인 것으로 평가되었다.

각 압축율에서 predistortion의 유/무에 따른 음질향상의 주관 평가 결과는 압축율이 256kbps까지는 평균차이점수의 향상이 없는 것으로 평가되었고, 그림 8에서와 같이 압축율이 128kbps에서는 평균차이점수가 0.4~1.5 정도 향상된 것으로 평가되었다. 이러한 결과는 평균차이

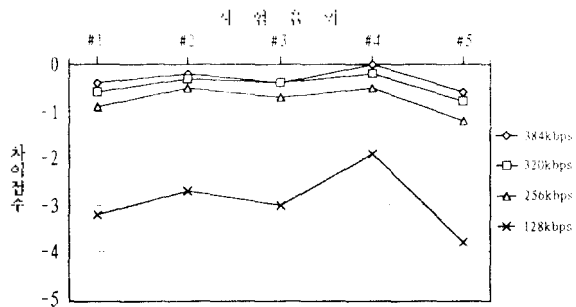


그림 7. 압축율 변화에 따른 멀티채널 오디오의 평균차이점수

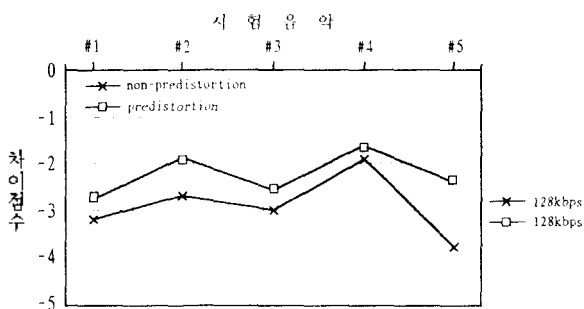


그림 8. 128kbps에서 predistortion 기법에 의한 품질향상

점수가 1이하의 귀에 거슬리지 않는 품질에서는 predistortion에 의해 가청잡음이 감소된 품질차이를 고도의 전문가가 아니면 구별할 수 없는 것으로 판단된다. 또한 128kbps에서 남자 음성(시험음악 5)이 다른 시험음악에 비해 높게 향상된 것으로 평가된 것은 복잡한 주파수 성분으로 구성된 음악에 비해 남자 음성은 수평선 상부에 단순하고 비교적 사람의 귀에 익숙해 있기 때문에 predistortion에 의한 품질향상을 구별하기가 용이하였다고 평가된다.

V. 결 론

본 논문에서는 디지털 텔레비전이나 고품위 텔레비전의 개발에 따른 실감있는 멀티채널 오디오의 효율적 전송 및 재생에 요구되는 멀티채널 오디오 부호화 기술을 MPEG-2 오디오 계층2의 알고리즘으로 구현하고, 한정된 저장매체나 제한된 전송채널의 조건하에서 압축율 변화에 따른 멀티채널 오디오의 주관 평가를 double blind 기법에 의해 10명의 청취자를 대상으로 청취실에서 평가하였다. 또한 matrixing/dematrixing 과정에서 각 채널의 양자화 오차의 조합에 의해 발생하는 가청잡음을 감소시켜 주는 predistortion의 유/무에 따른 품질향상을 각 압축율마다 평가하였다.

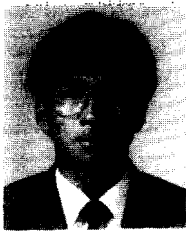
3/2채널의 스피커 시스템에서 압축율 변화에 따른 멀티채널 오디오의 품질이 256kbps를 경계로 5점 열화척도에 의한 평균차이점수가 급격히 열화되는 것으로 평가되었다. 따라서 Double blind 기법에 의한 주관 평가에서는 귀에 거슬리지 않는 -1수준의 품질을 얻기 위해서는 압축율이 256kbps 보다 높은 압축율로 멀티채널 오디오를 부호화 해야 된다고 평가되었다.

predistortion 기법에 의한 품질향상은 평균차이점수가 1이하인 256kbps 이상에서는 코딩왜곡의 감소로 인한 음질향상을 구별할 수 없고, -1이상인 128kbps에서는 척도 1정도 향상되었다. 특히 음악보다는 평소 귀에 익숙한 음성시료에서 predistortion에 의한 가청잡음의 감소를 잘 구별하였다.

참 고 문 헌

1. ISO/IEC 11172-3, "Coding of moving pictures and associated audio for digital storage media at up about 5Mbit/s", 1993.
2. ISO/IEC 13818-3, "Generic Coding of Moving Pictures and Associated Audio", 1994.
3. B. Grill, J. Herre, K. H. Brandenburg, E. Eberlein, J. Koller, J. Muller, "Improved MPEG-2 Audio Multi-channel Encoding", '96th AES Convention, Amsterdam 1994, preprint 3865.
4. Gerhard Stoll, Soren Nielsen, Andreas Silzle, "Extension of ISO/MPEG-Audio Layer II to Multi-Channel

- Coding", 94th AES Convension, Berlin 1993, preprint 3550(W4-3).
5. Gunther Theile, "The New Sound Format 3/2-Stereo", 94th Convention, Berlin 1993, prepint 3550
 6. ITU-R, "Methods for the Subjective Assessment of Small Impairments in Audio Systems including Multi-channel Sound Systems", ITU-R Draft New Recommendation, Document 10/65, November 1993.
 7. ITU-R draft New Recommendation: "Subjective Assessment of Sound Systems with Small Impairments including Multichannel Sound Systems", ITU-R Document TG10-/TEMP/10(Rev. 2), March 1993.
 8. 홍진우, 장대영, 김성환, 강성훈 "HDTV용 멀티 채널 오디오 소스 코덱 개발", JCC'95 논문집, pp. 470~473, 1995. 4
 9. 강성훈, 강경욱 "멀티채널 오디오 음질평가를 위한 기준 청취실 규격 및 설계", 한국음향학회 전기음향 워크샵 발표 논문집, pp. 153~159, 1995. 10



▲차 경 환(Kyung-Hwan Cha)정회원

1963년 4월 16일생

1985년 2월 : 부산수산대학교 전자통신공학과(공학사)

1990년 2월 : 부산수산대학교 대학원 전자통신공학과(공학석사)

1996년 2월 : 부산수산대학교 대학원 전자공학과(박사과정 수료)

1990년 1월~1995년 2월 : LG전자(주) 생활시스템연구소 주임연구원

1995년 4월~1996년 3월 : 한국전자통신연구소 위촉연구원

1995년 3월~현재 : 동서대학교 정보통신공학과 전임강사

▲장 대 영(Dae-Young Jang) 정회원

1936년 2월 24일생

현재 : 한국전자통신연구소 음향통신연구실 선임연구원 (1992년 제11권 1호 참조)

▲김 성 한(Sung-Han Kim) 정회원

현재 : 한국전자통신연구소 음향통신연구실 연구원

▲김 천 덕(Chun-Duck Kim) 정회원

1946년 5월 23일생

현재 : 부산수산대학교 전기공학과 교수

1988년 제7권 2호 참조)