

論文96-33B-4-17

지각 부호화를 이용한 스테레오 오디오 코덱의 구현 및 음질 평가

(Implementation and Evaluation of Stereo Audio Codec using Perceptual Coding)

車 炯 煥 *, 張 大 永 **, 洪 鎮 祐 **, 金 千 德 ***

(Kyung Hwan Cha, Dae Young Jang, Jin Woo Hong, and Chun Duck Kim)

요 약

본 논문에서는 저비트율의 고품질 오디오를 제공하기 위하여 TMS320C40 DSP(Digital Signal Processing)칩을 이용한 실시간 스테레오 오디오 코덱의 구현과 음질평가에 대하여 기술하였다. 과대한 반복계산과 처리과정의 복잡성으로 인하여 발생하는 오디오 압축 알고리즘의 실시간 처리 문제를 해결 하도록 하드웨어와 소프트웨어를 구현하였다. 또한, 지각 부호화를 이용하는 오디오 코덱에서 나타날 수 있는 5 가지 유형의 왜곡 특성을 분석하고, 이런 왜곡을 나타낼 수 있는 시험음악을 EBU(European Broadcast Union)에서 제작한 SQAM(Sound Quality Assessment Material) 422-2-4-2중에서 8곡을 선정하여 평가하였다. 주관평가는 double blind 기법에 의해 청취실에서 11명의 피험자를 대상으로 코덱 구현 결과 및 품질을 실험하였다. 그 결과 5등급-열화척도(5 Grade-Impairment Scale)가 -1 이하로 원음에 대해 차이를 느끼지 못하거나 거슬리지 않는 정도로 평가 되었다.

Abstract

In this paper, we described the implementation and the sound quality assessment of a real-time stereo audio codec using TMS320C40 DSP(Digital Signal Processing) chip for low bitrate and high quality audio. We implemented hardware and software in order to overcome a real-time processing problem of audio compression algorithm that can be produced by largely recursive computing and complexity of the process. We have studied five types of distortion that can be produced by perceptual coding and the codec was evaluated by eight test musics that are selected in SQAM(Sound Quality Assessment Material) 422-2-4-2 produced by EBU(European Broadcast Union). The subjective listening tests were carried out on the codec quality and performance by double blind method in a listening room with eleven listeners. As a result, 5 Grade-Impairment Scale was scored under minus one and the codec quality was evaluated to be perceptible, but not annoying.

I. 서 론

고음질 오디오 부호화 기술은 디지털 오디오, 디지털

음향방송, 고선명 텔레비전 및 멀티미디어등에서 그 중요성이 증대되고 있다. 오디오 신호를 고능률로 압축시켜 부호화하는 방법은 크게 두종류가 있다. 하나는 신호 자체의 통계적인 특성을 이용하여 정보를 잃지 않고 효율 있게 부호화하는 가변장 부호화(entropy coding)방법이며, 다른 하나는 인간의 청각 특성을 이용하여 감도가 낮은 미세한 정보는 생략하여 부호량을 감소시키는 지각 부호화(perceptual coding)방법이 있다. 지각 부호화는 가변장 부호화와는 달리 원신호와 재생신호사이에 정보의 객관적 측정량의 차이(정보 손실량)는 있으나 주관적 측정량의 차이(지각 평가량)가

* 正會員, 東西大學教 情報通信工學科

(Dept. of Information and Communication Eng. Dongseo Univ.)

** 正會員, 韓國電子通信研究所 音響通信研究室

(ETRI Acoustic Communication Section)

*** 正會員, 釜山水産大學教 電氣工學科

(Dept. of Elec. Eng. Nat'l Fisheries Univ. of Pusan)

接受日字: 1996年1月15日, 수정완료일: 1996年3月16日

없는 특징이 있다^{[1]-[4]}. 그러나 지각 부호화 방법은 가변장 부호화 방법에 비하여 계산량이 많고 처리과정이 복잡하기 때문에 처리시간이 많이 요구된다. 따라서 실시간 처리를 위한 많은 기법들이 필요하게 되는데 본 논문에서는 이러한 문제점을 해결하여 품질이 저하되지 않으면서 실시간으로 동작하도록 오디오 코덱을 구현하였다.

인간의 심리음향 특성(psychoacoustic property)을 고려하여 만든 오디오 코덱의 품질은 객관적 또는 주관적으로 평가할 수 있으나, 실제로 사람이 느끼는 품질은 주관평가를 통해서 가장 잘 예측될 수 있다. 주관평가는 비용과 시간이 많이 소모된다는 단점이 있으나, 지각 부호화를 이용한 오디오 코덱의 품질에 대한 주관적인 평가를 정확히 산출해 낼 수 없는 객관적 평가 알고리즘의 한계로 인하여 현재까지는 효율적으로 사용되고 있다. 그러나, 단순히 원음과 개발된 코덱과의 지각(perception) 및 인지(cognition) 수준에서 음질을 평가하는 MOS(Mean Opinion Score)만으로는 정확한 상관관계의 분석과 평가가 되지 못한다. 따라서, 오디오 코덱의 품질에 대한 주관적인 평가를 주어진 여건하에서 가장 경제적이고 신뢰성 있게 그리고 정확히 산출하기 위해서는, 연구하고자 하는 목적에 적합한 시험재료의 선정 및 반응척도의 구성과 실험조건에 따른 실험설계가 반드시 필요하다.

본 논문에서는 DSP 구조를 이용한 실시간 스테레오 오디오 코덱을 구현하고, 지각 부호화과정에서 나타날 수 있는 5가지 유형의 왜곡에 대한 특성을 나타내는 음악을 선정하여 구현한 코덱의 품질을 평가하였다. 부호화 알고리즘은 MPEG-1 오디오 계층 2의 알고리즘^[1]을 사용하였으며, 주관평가 방법은 double blind 기법에 의한 5 등급-열화척도(5 Grade Impairment Scale)을 사용하여 오디오에 익숙한 11명의 피험자를 대상으로 하여 40dBA의 암소음을 갖는 청취실에서 평가하였다.

II. 지각 부호화를 이용한 오디오 코덱 구현

1. 하드웨어 구현

인코더의 하드웨어는 그림 1에서와 같이 TI사의 TMS320C40(40MHz) 1개, 48kHz/16bit의 분해능을 갖는 2채널 ADC(Analog to Digital Converter) 2

개, 2채널 디지털 오디오의 인터페이스를 위한 DAIR (Digital Audio Interface Receiver) 2개, ROM (32KBytes * 32bits) 2개와 RAM (32KBytes * 16bits) 4개, 관련 기능 및 프로세서 기능을 위한 FPGA(Altera사의 EPF8820) 1개, 그리고 이들 수행에 필요한 일반 TTL 부품 및 소자들을 이용하여 설계하였다. 디코더의 하드웨어는 그림 2에서와 같이 인코더와 동일한 구조로 설계되었기 때문에, 2채널 디지털 오디오의 인터페이스를 위한 DAIR 대신에 DAIT (Digital Audio Interface Transmitter) 2개와 관련 기능 및 프로세서 기능을 위한 FPGA(Altera사의 EPF8452) 1개가 다르게 구성 되어있다.

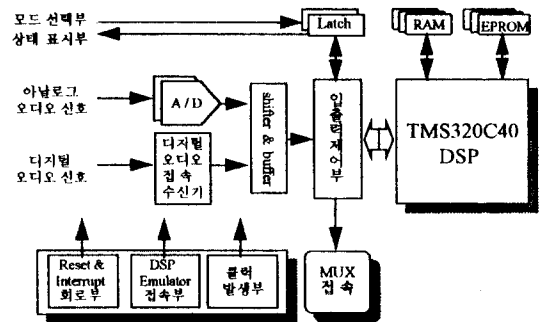


그림 1. 오디오 인코더의 블럭 구성도
Fig. 1. The block diagram of audio encoder.

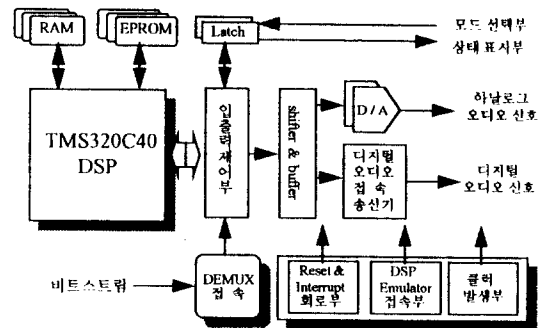


그림 2. 오디오 디코더의 블럭 구성도
Fig. 2. The block diagram of audio decoder.

지각 부호화를 이용하는 스테레오 오디오 코덱은 개발의 효율성을 위해 인코더와 디코더가 그림 1과 그림 2에서와 같이 동일한 하드웨어 구조를 가지고, 확장성이 용이한 구조로 설계되었다. 또한 최대 2채널의 디지털 오디오 신호를 AES/EBU, IEC958등의 규격으로

입력할 수 있으며 아날로그 또는 디지털 오디오 입력으로 부터 입력되어 저장된 데이터에 인코딩 알고리즘을 적용하여 생성된 부호화 비트 스트림을 비디오 신호와 다중화하도록 다중화부로 전달하는 기능을 가지고 있으며, 사용자가 필요에 따라 3개의 on/off 스위치를 이용하여 인코더의 인코딩 채널을 선택하거나, 선택된 결과를 LED에 표시함으로써 사용자가 확인할 수 있도록 하는 기능을 가지고 있다.

2. 지각 부호화를 이용한 소프트웨어 구현

고음질 부호화기에서 가장 중요하게 사용되는 청각 특성이 바로 마스킹 현상(masking effect)이다^[5]. 마스킹 현상은 한 음계 대역에서 음압(sound pressure level)이 큰 음이 존재할 때 주위 대역에서 양자화 잡음과 같은 작은 신호들은 들리지 않게 되는 것을 말한다. 따라서 절대 가청 한계와 입력 신호에 따른 마스킹 효과를 이용하여 전체적인 마스킹 곡선을 얻어내고, 이를 신호의 스펙트럼과 비교하여 양자화 잡음이 곡선 이하에 존재하도록 부호화하면, 잡음이 존재하더라도 크게 인식되지 않으므로, 지각적으로 동일한 복원음을 얻으면서 압축율을 향상시킬 수 있다.

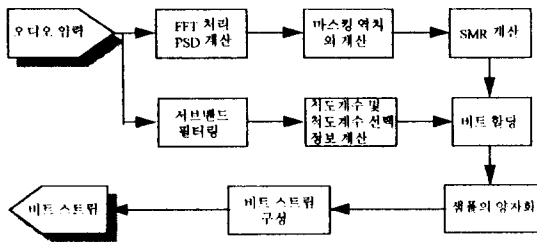


그림 3. 스테레오 오디오 부호화 절차
Fig. 3. Stereo audio encoding process.

그림 3은 오디오 부호화 알고리즘의 흐름도를 나타낸다. 부호화 절차는 먼저 심리음향모델에 의한 SMR (Signal to Mask Ratio)을 계산하는데, 1024샘플로 구성되는 프레임에 대해 FFT분석을 수행하여 PDS (Power Density Spectrum)를 추정한 후, 신호의 주기성(tonal)성분과 비주기성(non-tonal)성분을 구하고 개별 마스킹 역치와 전체 마스킹 역치를 계산하고, 각 서브밴드의 최대 신호레벨에서 최소 마스킹 역치를 빼주면 32개의 서브밴드에 대한 SMR이 얻어진다. 결국 SMR이 작다면 신호의 음압레벨이 작거나 마스킹이 많이 되는 것이므로 SMR이 작을수록 적은 비트를 가

지고 효과적으로 양자화를 할 수 있다.

다위상 필터뱅크를 사용하여 각 채널의 1,152개의 입력 샘플을 받아서 서브밴드당 36개의 샘플을 갖는 32개의 서브밴드를 출력한다. 각 서브밴드에 대한 척도계수(scf:scale factor)를 구하여 각 블럭의 샘플이 효율적으로 양자화될 수 있도록 정규화하는데 사용한다. 비트 할당은 그림 4의 비트할당 흐름도와 같이 그 프레임에 할당된 비트의 수가 초과하지 않을 때까지 반복적으로 증가시키면서 비트할당을 한다.

서브밴드 샘플의 양자화를 위해서 서브밴드 샘플을 척도계수로 나누어 정규화하고 그 결과를 비트할당 블럭의 비트수에 따라 양자화한다. 낮은 주파수 영역에 있는 서브밴드 샘플들은 15개의 양자화군으로, 중간 주파수 영역에 있는 서브밴드 샘플들은 7개의 양자화군으로, 높은주파수 영역에 있는 서브밴드 샘플들은 3개의 다른 양자화군으로 양자화된다.

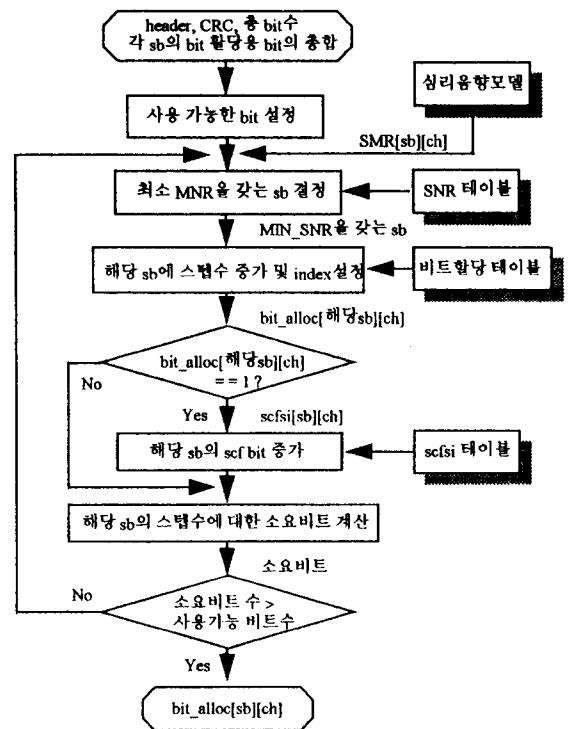


그림 4. 비트 할당 흐름도
Fig. 4. Bit allocation flow.

고음질 오디오의 복호화는 심리음향 모델이 없으므로 부호화보다는 간단하며, 부호화 절차는 비트스트림을 분석, 역양자화 및 정규화, 서브밴드필터 합성, 시간영역 PCM 데이터 복원의 순으로 처리된다.

3. 실시간 처리를 위한 하드웨어의 및 소프트웨어 개선 사항

1) 하드웨어의 개선 사항

인코더는 디코더 보다 처리 과정이 복잡하여 2~3배 정도의 처리 시간이 더 걸리기 때문에 인코더의 실시간 처리 방법에 대해서 설명하고, 이 방법은 디코더에도 동일하게 적용되었다. 인코더에서 A/D입력과 비트 스트림의 출력 방식이 1개의 워드(word) 혹은 1개의 바이트(byte) 단위로 인터럽트에 의해 수행되는 관계로 1초에 4,800개씩 9,600개의 인터럽트를 처리해야 하는 DSP의 부하를 줄이기 위하여 입력과 출력 제어부에 각각 FIFO(1Kbyte * 16bits)를 두었으며, 이들 FIFO와 DSP간에는 DMA에 의해 처리하도록 하였다. 또한 소프트웨어의 빠른 처리를 위하여 실행 프로그램은 ROM으로 부터 TMS320C40의 bootstrap기능을 이용하여 static RAM(access time 15nsec)으로 load시킨 후 RAM에서 프로그램 명령어가 실행되도록 처리하였다. 그리고 TMS320C40내에 설정되어 있는 내부 메모리(2Kbytes * 32bits)를 프로그램 RAM과 데이터 RAM으로 구성한 후 데이터를 처리하였다.

2) 소프트웨어 개선 사항

처음에 C언어를 이용하여 프로그램을 작성하고 디버깅 하였으며, 각 함수의 실행시간을 측정한 후 시간이 많이 걸리는 함수를 중심으로 최적화를 수행하였다. 이 과정에서 만족되지 않는 부분은 TMS320C40 어셈블리 언어로 프로그램을 구성하였다. 실행시간이 많이 걸리는 부분으로는 SMR계산을 위한 심리 음향 모델링, 서브 밴드 필터링, 비트 할당 부분이었다. 입출력 드라이브, 심리 음향 모델링부의 FFT계산, 서브 밴드 필터링부의 주파수 분석, 비트 할당부의 비트 처리 등이 어셈블리 언어로 구성되었다. TMS320C40 어셈블리 언어의 구성에 있어서 실행 cycle이 큰 RPTS명령, BR 명령 등의 사용을 억제하고, delayed 명령어와 parallel 명령어를 이용한 TMS320C40의 pipeline structure를 최대한 활용하였다. 또한 내부 메모리의 일부를 stack으로 사용하여 push와 pop시간을 단축하였고, 나머지를 heap으로 처리하여 cycle buffer로 구성한 후 모든 데이터의 처리를 이 내부 메모리에서 수행함으로써 데이터의 access time을 최대한 줄이는 방법을 이용하였다. 그리고 look-up table방식을 활용함으로써 산출식으로 처리되는 알고리즘과 logic계산에 필요한 데이터들의 처리시간을 감축하였다.

48 kHz 표본화 및 16비트 부호화에 의해 입력되는 오디오 샘플의 한 프레임(1152샘플)을 처리하기 위한 시간은 24msec로 제한된다. 따라서 TMS320C40 40MHz를 사용할 경우 실행 cycle이 50nsec가 되기 때문에 24msec동안 처리할 수 있는 전체 실행 cycle은 480,000이 된다. 본 논문에서 구현한 코덱의 경우 상기의 실시간 처리를 위한 하드웨어 기법 및 소프트웨어 기법을 이용한 결과 최대 440,000 실행 cycle이 내에서 처리됨을 확인하였고 이것은 TMS320C40 40MHz의 처리능력을 최대로 사용하는 경우 약 92% 사용한 결과가 된다.

III. 오디오 코덱의 주관품질 평가

1. 실험방법

그림 5는 구현된 실시간 스테레오 오디오 코덱의 품질을 평가하기 위한 절차를 나타내고 있다. 콤팩트디스크(CD)의 아날로그 스테레오 오디오 출력의 원음은 구현한 인코더의 A/D변환기에 의해 48KHz 샘플링과, 16비트 부호화에 의해 디지털 데이터로 변환되고(스테레오의 경우 1,536kbps), 앞에서 기술한 압축 알고리즘에 의해 256kbps의 비트스트림으로(약 6:1 압축) 디코더에 전송된다. 디코더에서는 전송 받은 비트스트림을 복호화하여 아날로그 스테레오 오디오를 재생하고, 이 신호를 DAT(Digital Audio Tape)에 녹음시켰다.

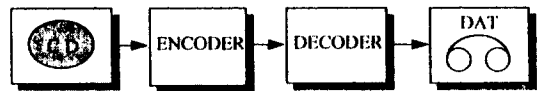


그림 5. 실시간 스테레오 오디오 코덱의 녹음과정
Fig. 5. Real-time stereo audio codec recording process.

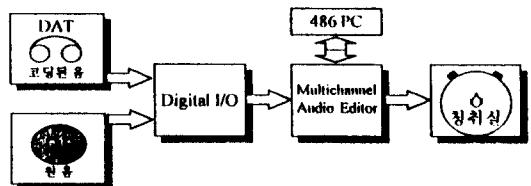


그림 6. 주관품질 평가 실험을 위한 편집 과정
Fig. 6. Editing process for the subjective listening test.

그림 6은 원음과 코덱에 의한 재생음의 품질평가를

위한 편집과정을 나타내고 있다. CD의 원음과 DAT의 코딩된 음은, 멀티채널 오디오 재생기에서 double blind 기법에 의한 주관평가 실험이 용이하도록, 원음과 코딩된 음의 재생 순서를 편집한 후 PC의 하드디스크에 저장한다. 청취실에서 주관평가 실험을 할 때, 청취자의 요구에 따라 원음과 코딩된 음은 적절하게 들려줄 수 있도록 준비를 해둔다.

2. 시험음악의 선정

인간의 심리음향 특성을 고려하여 만든 오디오 코덱 들은 오디오 신호를 주파수영역으로 변환시키고, 어떤 성분을 들리게 할지를 결정하기 위해, 마스킹 곡선들을 사용한다. 그리고 마스킹 역치 이하의 부분을 제외한 가청 주파수 성분만을 코덱에서 전송하게 된다. 만일 이러한 가청성분들이 낮은 해상도로 양자화 되었을 때, 양자화 과정에 의해 왜곡성분들이 발생할 것이다. 이러한 지각 부호화 과정에서 나타날 수 있는 왜곡은 5가지의 유형으로 분류할 수 있다¹⁶⁾.

(1) 블럭율 왜곡(block rate distortion): 한 블럭에서 다음 블럭으로 신호처리가 바뀔때, 블럭사이의 아주 짧은 지연속성에 의해 발생하는 왜곡으로 지속적인 순음(pure tone)에 의해 평가 할 수 있다. 이러한 왜곡은 순음의 성분이 지속적으로 발생하는 목금(marimba)과 철금(glockenspiel) 등에서 나타난다.

(2) 프리에코(pre-echo): 계속되는 블럭처리에서 갑작스러운 입력신호의 레벨변화에 대해, 출력신호가 입력신호의 변화이전에 크기가 변동하는 왜곡이다. 이러한 왜곡은 하나의 주기를 가지는 단속적인 정현파에 의해 평가할 수 있다. 프리에코는 주로 타악기의 사운드에서 나타난다.

(3) 잡음침투(noise intrusion): 지각 부호화 알고리즘이 처리하기에는 부하가 많은 복잡한 신호가 입력되었을 경우, 레벨이 마스킹 역치(masking threshold) 이하가 되어야될 여러신호들이 가청범위로 침투되는 것을 말한다. 이러한 왜곡은 다중음(multitone signals)에 의한 출력 스펙트럼으로 평가할 수 있다.

(4) 대역간 누화(interband crosstalk): 입력신호를 서브밴드로 분할할 때, 디지털 필터의 설계 및 구현의 잘못에 의해 생기는 것으로, 특정 대역의 신호가 다른 대역에서 반응을 나타내는 왜곡으로, 블럭율 왜곡과 거의 동시에 일어난다. 디지털 필터의 aliasing 효과에 의해 문제가 있는 특정 주파수가 지속적으로 발생하는

음악에서 들을 수 있다.

(5) 스펙트럼 형상(spectrum shaping): 비선형계에서만 발생하는 특별한 유형의 왜곡으로, 부호화 알고리즘이 여러 주파수로 구성된 신호중 특정 주파수만을 증폭시키거나, 중요한 주파수 성분을 빠뜨리고 부호화하는 경우에 발생된다. 이러한 왜곡은 광대역의 주파수 특성을 가지는 다중음에 의한 주파수 응답으로 평가할 수 있다.

상기와 같은 왜곡을 나타낼 수 있는 음악을 EBU에서 제작한 SQAM 422-2-4-2중에서 8곡을 선정하였다. 표 1은 주관평가에 사용된 총 8곡의 시험음악들을 나타내고 있다.

표 1. 주관평가에 사용된 시험음악
Table 1. Test music used to evaluate the subjective performance.

| No. | 왜곡 평가 특성 | 시험 음악 | 시간 (secs) |
|-----|----------|-----------------------|-----------|
| 1 | 블럭율 왜곡 | Double-bass | 6 |
| 2 | 블록율 왜곡 | Clarinet | 7 |
| 3 | 프리에코 | Side-drum with snares | 7 |
| 4 | 잡음침투 | Piano(Schubert) | 14 |
| 5 | 잡음침투 | Soprano(Mozart) | 18 |
| 6 | 대역간 누화 | Flute | 6 |
| 7 | 대역간 누화 | Trumpet | 8 |
| 8 | 스펙트럼 형상 | Vibraphone | 5 |

3. 주관평가 실험

주관평가 실험은 "duble blind triple stimulus with hidden reference" 방법¹³⁾으로 평가 하고자 한다. 여기서, double blind 기법이란 통상 자극의 제시 순서나 내용을 피험자만 모르게 하는 single blind 기법과는 달리, 피험자는 물론 실험을 관리하는 실험자도 어떤 시험음악이 제시되는가를 사전에 알 수 없게 하여, 실험결과에 피험자나 실험자의 의도가 배제되도록 하는 기법을 말한다.

이 방법에서 실험자는 사전에 8개의 시험음악 모두에 대해 각각 2개의 원음과 1개의 부호화된 음악을 준비한다. 항상 원음(이하 A로 표시)을 먼저 제시하고, 다음에 미리 불규칙하게 배열되어 있는 부호화된 음악(이하 B 또는 C)이나 원음(이하 B 또는 C)을 각각 제

시하여, B 또는 C가 A에 비해 얼마나 열화되었는지 또는 품질이 상대적으로 어느 정도인지를 피험자로 하여금 표 2에 나타난 연속적인 5 등급-열화 척도(continuous 5 grade-Impairment scale)에 따라 반응지에 평가 한다. 평가 대상인 부호화된 음악과 숨겨진 원음의 평가 점수간 차이가 통계분석의 입력치로 사용된다. 즉, 5 등급-열화 척도인 경우에는 식(1)과 같이 차이점수를 계산하여, 그 결과를 그래프 형태로 나타낸다. ITU-R의 5 등급-열화 척도는 평가에서 좀 더 미세한 크기까지를 도출해 내기 위하여 소숫점의 사용도 허용한다.

차이점수(diff-grade) =

$$\text{숨겨진 원음의 점수} - \text{부호화된 음의 점수} \quad (1)$$

표 2. ITU-R의 5 등급-열화척도와 대응되는 차이점수

Table 2. Five grade-impairment of ITU-R and diff-grade.

| 손상 정도 | 점수 | 대응되는 차이점수 |
|----------------|-----|-----------|
| 느낄수 없다 | 5.0 | 0 |
| 느끼지만, 거슬리지 않는다 | 4.0 | -1 |
| 약간 거슬린다 | 3.0 | -2 |
| 거슬린다 | 2.0 | -3 |
| 매우 거슬린다 | 1.0 | -4 |

4. 청취조건

주관 평가에 있어서 기준 청취 위치에 청취자가 위치하고, 기준 스피커 배열에 의한 기준 음장이 만들어질 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 시스템의 주관적인 평가 결과들을 상호비교 검증할 수 있도록 권고하고 있는 내용을 고려하여 그림 7과 같은 청취조건을 설정하였다. 또한 실험에 사용된 청취실은 국제적인 근간이 되고있는 ITU-R의 TG 10 3(subjective assessment)과 EBU 전문가 그룹 G1/LIST(subjective assessment and listening condition)의 규격에 맞게 설계된 청취실(잔향시간: 0.2 초, 실내용적: 4.7m x 6m x 3m, 초기반사음: -14dB, 흡소음: 40dBA)에서 행하여졌다.

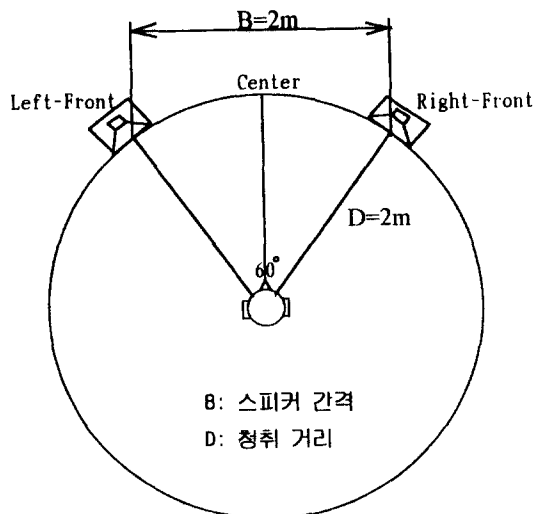


그림 7. 주관평가를 위한 스피커 배치와 청취위치
Fig. 7. Speaker position and listening point for subjective listening test.

5. 실험결과 및 고찰

그림 8은 double blind 기법에 의한 원음과 실시간 스테레오 오디오 코덱의 평균 차이점수를 나타내고 있다. 시험음악 8곡에 대한 구현된 코덱의 품질은 평균 차이점수가 모두 -1이하로 표 2에서 제시한 평가척도에 의하면 원음과 차이를 느끼지 못하거나, 느끼지만 거슬리지 않는 정도로 평가되었다.

부호화 과정에서 나타날 수 있는 5가지 유형의 왜곡 특성을 분석하기 위해서 선정된 시험음악을 통해서 코덱을 분석해 보면, 특정 시험음악에 대해서 주관평가의 차이점수가 특별히 높게 평가된 것은 없으므로, 부호화 과정에서 블럭처리의 불연속성, 임계대역에서 복잡한 신호에 의한 신호처리의 과부하로 마스킹 역치로의 침투로 인한 왜곡성분들은 발생하지 않은 것으로 평가된다.

그러나 trumpet(시험음악 7)과 같은 특정주파수가 지속적으로 발생하는 시험음악에서 상대적으로 차이점수가 높게 평가된 것은, 특정주파수가 서브밴드 분할필터인 디지털 필터의 aliasing이 발생하는 서브밴드 경계부근에서 자주 일어나기 때문이라고 분석된다. 따라서, 일시적 마스킹 효과를 이용하여 서브밴드내의 오디오 블럭을 최적으로 부호화 시키는데 사용되는 서브밴드 분할필터의 주파수 분해능을 향상시킬 수 있는 디지털 필터의 성능 향상이 대역간 혼입의 왜곡을 감소시킬 수 있다고 분석된다.

한편 side-drum with snares(시험음악 3)는 프리에코에 의한 왜곡이 발생하는지를 평가하기 위하여 선정되었는데, 원음과 차이를 느끼지 않을 정도로 평가되었다. 타악기 음과 같이 신호레벨의 갑작스러운 변화에 의한 프리에코 성분의 왜곡은 2ms ~ 4ms 이내의 아주 짧은 구간에서 발생하므로 고도의 훈련을 받은 전문가가 아니면 왜곡을 느끼지 못하기 때문에 펄스성분의 시험신호에 의한 스펙트럼 분석으로 평가 되어져야 할 것이다.

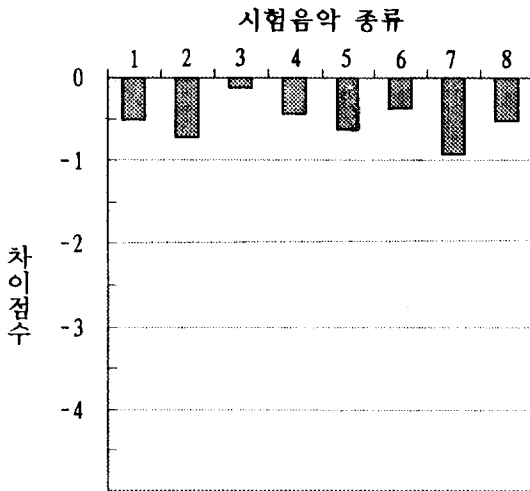


그림 8. 실시간 스테레오 오디오 코덱의 평균 차이점수

Fig. 8. Mean diffgrades of real-time stereo audio codec.

IV. 결 론

본 논문에서는 한정된 저장매체나 제한된 전송채널의 조건하에서 저비트율의 고음질을 요구하는 실시간 처리용 스테레오 오디오 코덱의 구현과 품질 평가에 대하여 기술하였다. 코덱은 인코더와 디코더가 유사한 하드웨어 구조를 가지고, 확장성이 용이한 구조로 설계되었으며, TMS320C40 DSP가 갖고 있는 장점을 충분히 활용할 수 있도록 하드웨어 및 소프트웨어를 구현하여 TMS320C40 40MHz의 처리능력의 약 92%를 사용한 결과로써 실시간 처리됨을 확인하였다.

코덱의 주관품질 평가는 지각 부호화 과정에서 나타날 수 있는 5가지 유형의 왜곡 특성을 분석하고, 그 왜곡을 나타낼 수 있는 시험음악을 EBU의 SQAM중에서

8곡을 선정하여, ITU-R 과 EBU에서 권고하는 규격에 맞게 설계된 청취실에서 오디오에 익숙한 연구원 11명을 대상으로 하여 double blind 기법으로 평가하였다. 실험결과 5 등급-열화척도에 의한 차이점수가 -1이하로 산출되었으며 원음과 차이를 느끼지 못하거나 느끼지만 거슬리지 않는 것으로 평가되어 본 논문에서 구현한 스테레오 오디오 코덱의 우수성을 입증하였다.

왜곡특징 분석을 위해 선정된 8곡의 시험음악에 의한 왜곡평가가 크게 차이나지 않아서 구현된 코덱에서는 두드러지게 나타난 왜곡성분들은 없었으며, 상대적으로 프리에코왜곡이 가장 낮게 평가되었고, 대역간 혼입왜곡이 가장 높게 평가되었다.

앞으로, 스테레오 오디오 코덱의 객관적 품질평가 방법에 대한 연구가 지속되어야 할 것이며, 멀티 채널 오디오 코덱의 구현 및 그 평가 방법에 대한 연구가 있어야 할 것이다.

참 고 문 헌

- [1] ISO/IEC 11172-3, Coding of moving pictures and associated audio for digital storage media at up about 1.5Mbit/s, 1993.
- [2] ITU-R, "Methods for the Subjective Assessment of Small Impairments in Audio Systems including Multichannel Sound Systems", ITU-R Draft New Recommendation, Document 10/65, November 1993.
- [3] J. G. Beerends and J. A. Stemerdink, "A Perceptual Audio Quality Measure Based on a Psychoacoustic Sound Representation", *Journal of Audio Engineering Society*, Vol. 40, No. 12, pp. 936~978, December 1992.
- [4] D. Wiese and G. Stoll, "Bitrate Reduction of High Quality Audio Signals by Modelling the Ears Masking Thresholds," The 89th AES Convention, LA, Preprint No. 2970, Sep. 21 ~ 25, 1990.
- [5] E. Zwicker, *psychoacoustics*. Springer-Verlag, new York, 1982.
- [6] A. Milne, "New Test Methods for Digital Audio Data Compression Algorithms",

Proceedings of the Audio Engineering Society 11th International Conference, pp. 210 ~215, Portland, Oregon, USA, May, 1992.

[7] C.Grewin and T.Ryden, Subjective Assessment on Low Bit-Rate Audio Codecs, Audio Engineering Society 10th International Conference, pp. 91 ~102, 1991.

[8] 홍진우, 장대영, 김성환, 강성훈 HDTV용 멀티 채널 오디오 소스 코덱 개발, JCCI95 논문집, pp. 470 ~473, 1995.4

[9] 강성훈, 강경옥 멀티채널 오디오 음질평가를 위한 기준 청취실 규격 및 설계, 한국음향학회 전기음향 워크샵 발표 논문집, pp. 153 ~159, 1995. 10

— 저 자 소 개 —



車 垌 煥(正會員)

1995년 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 ~ 1995년 2월 한국전자통신연구소 위촉연구원. 1995년 3월 ~ 현재 동서대학교 정보통신공학과 전임강사.



張 大 永(正會員)

1966년 2월 24일생. 1991년 2월 부산수산대학교 전자공학과(학사). 1991년 1월 ~ 현재 한국전자통신연구소 선임연구원.

주관심 분야: 오디오 부호화, 실감음향, 가상현실.

洪 鎮 祐(正會員) 第 30卷 B編 第 10號 參照

金 天 德(正會員)

1971년 2월 부산수산대학 어업학과(수산학사). 1981년 2월 일본동북대학 전기 및 통신공학과(공학석사). 1984년 2월 일본동북대학 전기 및 통신공학과(공학박사). 1984년 4월 ~ 1988년 9월 부산수산대학 전자통신공학과 조교수. 1986년 8월 ~ 1987년 8월 펜실바니아주립대학 객원연구원. 1988년 10월 ~ 1993년 10월 부산수산대학교 전자공학과 부교수. 1993년 10월 ~ 현재 부산수산대학교 전기공학과 교수.