

MPEG 오디오 부호화 기술 분석

홍진우, 강경옥

한국전자통신연구원 무선·방송기술연구소

Analysis of MPEG Audio Coding Technology

Jin-Woo Hong and Kyeong-Ok Kang

ETRI-Radio & Broadcasting Technology Laboratory

jwhong@etri.re.kr

요약

MPEG 오디오 그룹에서는 오디오 부호화 기술의 국제 표준으로 MPEG-1 오디오, MPEG-2 오디오 BC, MPEG-2 AAC의 규격 제정을 완료하였고, 현재 MPEG-4 오디오 및 MPEG-7 오디오의 국제 표준을 제정하고 있다. 본 논문에서는 이들 표준에 대한 요구 기능 및 기술 특징을 분석하고, 각각의 표준에 대한 응용분야와 향후의 계획에 대하여 기술한다.

1. 서론

방송, 영상전화, 대화형 멀티미디어 서비스와 같은 시각통신은 기본적으로 오디오를 필요로 하며, 미래의 문자, 그래픽, 팩스, 영상, 전자메일 등도 음성이나 오디오를 하나의 수단으로 제공하는 서비스로 발전하게 될 것이다. 따라서, 음성이나 오디오 신호를 처리하는 기술은 매우 중요한 연구 분야가 된다.

한정된 저장 매체 및 제한된 전송 용량에 많은 양의 오디오 정보를 제공하기 위한 디지털 오디오 신호의 압축 기술이 꾸준히 발전하고 있으며, 다양한 응용분야에 적용되고 있다. 특히, ISO/IEC SC29 산하의 MPEG(Moving Picture Experts Group) 오디오 그룹에서는 MPEG-1 오디오와 MPEG-2 오디오의 국제표준을 제정하였으며, MPEG-4 오디오, MPEG-7 오디오 등의 국제표준을 제정하기 위하여 노력하고 있다[1]-[5].

MPEG-1 오디오는 스테레오 오디오 부호화의 국제 표준이고, MPEG-2 오디오는 멀티채널(5.1 채널) 오디오 부호화의 국제표준으로 2가지 방식이 제정되었다. 하나는 MPEG-1 오디오와 호환성을 갖는 MPEG-2 오디오 BC(Backward Compatibility)이고, 다른 하나는 MPEG-1 오디오와 호환성을 갖지 않는 MPEG-2 AAC(Advanced Audio Coding)이다.

MPEG-4 오디오는 '99년 1월에 국제표준(ISO/IEC 14496-3) 제정을 목표로 하고 있으며, 채널당 64kb/s 이하의 데이터율을 갖는 오디오 및 음성 부호화 표준이다. MPEG-7 오디오는 다양한 오디오 정보의 효율적 검색을 위한 오디오 내용 표현 인터페이스를 의미한다.

본 논문에서는 고품질을 유지하면서 다양한 응용분야에 적용할 수 있는 MPEG 오디오 기술의 특징, 요구사항, 적용분야 등에 대해 분석하고, 기술하고자 한다.

2. MPEG-1 오디오(ISO/IEC 11172-3)

1992년에 국제 표준으로 제정된 MPEG-1 오디오는 스테레오(2 채널) 오디오에 대한 부호화 방식으로 고품질, 고압축율을 실현하였으며, 압축 부호화 기술을 적용하기 위한 분야가 다양하다는 점을 감안하여 품질

(quality), 처리지연(delay), 복잡도(complexity), 에러율(error-sensitivity), 비트율(bit rate) 등의 기술적 파라미터를 고려한 3가지 계층(계층 I, 계층 II, 계층 III)의 부호화 기술을 규정하였다[6][7].

계층 I은 매우 낮은 비트율을 꼭 필요로 하지 않는 응용분야에 적합하다. 즉, 가정용 디지털 테이프 레코더(DAT), 윈체스터(Winchester) 디스크 또는 자광 디스크(magneto-optical disk) 등에 이용될 수 있다. 계층 II는 계층 I에 대한 중복성(redundancy)과 무관성(irrelevance)을 제거함으로써 계층 I보다 더 많은 압축을 수행하고, 오디오 방송, 텔레비전, 레코딩, 전기통신, 그리고 멀티미디어 등과 같은 분야의 일반용 및 전문가용의 오디오 시스템에 적용될 수 있다. 계층 III은 혼합형 필터뱅크(hybrid filterbank)를 사용함으로써 주파수의 분해도를 높이고, 비선형 양자화 및 가변장 부호화(entropy coding)를 적용하여 부호화의 효율을 향상시킨 기술로서, 매우 낮은 비트율을 필요로 하는 전기통신 분야, 특히 협대역 ISDN과 음악 재생기, 전문가용의 시스템에 적용될 수 있다.

MPEG-1 오디오 계층 I, II의 인코더 및 디코더의 블럭 구성도는 그림 1과 같다. MPEG-1 오디오 표준에서 오디오와 비디오를 포함한 경우의 데이터 압축 정보량은 약 1.5Mbit/s 이고, 이 표준은 비디오 CD, CD-ROM, VOD(Video-on-Demand), 위성방송 등에 응용될 수 있다. 이중 오디오는 32kHz, 44.1kHz, 48kHz의 표본화 주파수와 16비트 부호화를 갖는 PCM 오디오 신호를 모노의 경우 32 ~ 192kbit/s, 스테레오의 경우 64 ~ 384kbit/s의 비트율로 압축하며, 단일 채널, 이중 채널(bilingual), 스테레오, 결합 스테레오 등의 동작 모드를 수용하며, 서브밴드 부호화 방식을 이용한다.

각 서브밴드에 대한 scf(scale factor)의 계산은 12개의 샘플로 구성된 하나의 블럭에서 수행된다. 12개 샘플들의 절대값의 최대치가 결정된 후 전체 120dB의 동적 범위를 수용하는 6비트의 워드길이를 양자화된다. scf의 비트율을 줄이기 위해 귀의 일시적 마스크 현상을 이용하는 부호화 전략을 사용하여 한 프레임내의 각 서브밴드에서 연속된 3개의 scf를 어떤 크기의 형태로 구분하고, 그 형태에 따라 서브밴드당 2비트로 구성된 scfsi(scale factor information)를 가지는 1 ~ 3개의 scf를 전송한다.

비트 할당과 비트할당 정보의 부호화에서 샘플을 부호화하는데 유용한 비트의 수(전송 비트율)를 결정하여야만 한다. 이것은 scf, scfsi, 비트할당 정보, 그리고 보조 데이터(ancillary data)에 필요한 비트의 수에 따라 결정된다. 비트할당은 각 서브밴드에 있는 마스크 한계

의 최대 레벨과 최소 레벨의 차로부터 유도된다. 즉, 비트할당 처리는 모든 서브밴드와 전체 프레임에 대한 NMR(noise-to-mask ratio)의 합을 최소화 시킨 후 서브밴드의 양자화 레벨의 수를 그 프레임에 유용한 비트의 수를 초과하지 않을 때까지 반복적으로 증가시키면서 수행한다.

서브밴드 샘플의 양자화를 위해서 먼저 한 블럭의 12개 서브밴드 샘플의 각각이 scf로 나누어져 정규화되고, 그 결과가 비트할당 블럭의 비트수에 따라 양자화된다. 낮은 주파수 영역에 있는 서브밴드의 샘플들은 15개의 양자화군으로, 중간 영역의 주파수 영역에 있는 서브밴드의 샘플들은 7개의 양자화군으로, 높은 주파수 영역에 있는 서브밴드 샘플들은 3개의 다른 양자화군으로 양자화된다. 이 양자화군은 3, 5, 7, 9, 15,, 65535 양자화 레벨을 포함한다. 이때 3, 5, 9의 양자화 레벨은 단지 2, 3, 또는 4비트로 구성된 코드워드들 효율적으로 사용할 수 없기 때문에 3개의 연속된 서브밴드 샘플들을 모아서 덩어리(granule)로 집단화하고, 이 덩어리를 하나의 코드워드로 부호화 한다. 집단화를 사용한 경우의 부호화 이득은 37.5%까지 얻을 수 있고, 특히, 고주파수 영역에 있는 서브밴드 샘플들이 3, 5, 7, 9 양자화 레벨로 양자화 되기 때문에 이 영역에서 집단화의 효율이 높다.

MPEG-1 오디오 비트 스트림은 384개의 PCM 샘플(48kHz 표본화율인 경우 계층 I에서 8ms) 또는 1152개의 PCM 샘플(48kHz 표본화율인 경우 계층 II에서 24ms)에 일치하는 짧은 프레임으로 구성되며, 데이터 구조는 그림 2와 같다.

3. MPEG-2 오디오 BC(ISO/IEC 13818-3)

ISO/IEC SC-29 내의 MPEG 오디오 그룹에서 1994년 11월에 국제 표준으로 제정한 멀티채널 오디오 부호화 기술을 의미한다[8][9][10]. MPEG-2 오디오 BC는 MPEG-1 오디오와 순방향 호환성 및 역방향 호환성을 유지하도록 개발된 부호화 방식("MPEG-2 오디오 BC"라고 한다)이다. MPEG-2 오디오 인코더의 처리 블럭 구성도는 그림 3과 같으며, 다음과 같은 특징을 갖는다.

1) 멀티채널 부호화

전통적인 2채널 사운드 포맷에서 확장된 개념인 3/2+1의 구성을 갖는 멀티채널 시스템이다. MPEG-2 오디오 BC 부호화의 표준(ISO 13818-3)은 역방향 호환성을 갖는 저비트율 멀티채널 인코딩 시스템의 기술들을 포함한다. MPEG-1의 좌측(L)과 우측(R) 채널은 매트릭스된 신호 Lc와 Rc로 대체된다. 여기서

$$L_c = L + a \cdot C + b \cdot L_s, \quad R_c = R + a \cdot C + b \cdot R_s$$

이고, MPEG-1 부호화기처럼 부호화된다. 그러므로, MPEG-1 복호화기는 전체 5채널 정보를 가지고 다운믹스한 것을 재생할 수 있어야 한다.

2) 다중언어 부호화(multilingual coding)

역방향 호환성 멀티채널 부호화와 더불어 다중의 사운드 채널을 전송하기 위한 추가 채널이 제공된다. 이들 추가 채널에는 다중 언어 해설 사운드가 제공될 수 있다.

3) 낮은 표본화 주파수에서의 부호화

MPEG-2 오디오에서의 또 다른 기능은 32kHz 이하의

낮은 표본화 주파수를 사용하는 모드의 추가이다. 이들 모드는 광대역 음성과 채널당 16kbps ~ 64kbps 사이의 비트율에 중간급의 음질을 필요로 하는 분야의 전송에 유용하기 때문에, 해설뿐만 아니라 원격회의 시스템, 그리고 비트율이 제한되는 곳의 응용분야에 적합하다. 낮은 표본화 주파수를 추가한 기본 개념은 고주파수 분해능 필터뱅크에 대한 부호화 이득을 증가시키기 위한 것이다. 낮은 표본화 주파수의 또 다른 장점은 주정보와 부정보(특히, 헤더)의 개선율이다.

4. MPEG-2 AAC(ISO/IEC 13818-7)

MPEG 오디오 그룹에서 1997년 4월에 제정한 국제 표준으로 MPEG-1 오디오와의 역방향 호환성을 유지하지 않는 멀티채널 오디오 부호화 방식을 의미한다[4].

MPEG-2 오디오가 MPEG-1 오디오와의 역방향 호환성을 유지하여야 한다는 요구 사항 때문에 명료한 음질을 재생하기 위하여 알고리즘이 복잡해지고, 처음에 목표 설정하였던 비트율을 만족할 수 없었다. 이 때문에 역방향 호환성의 요구 사항을 배제함으로써 MPEG-2 오디오보다 통계적으로 성능이 좋은 부호화기를 개발하여야 한다는 인식이 MPEG 오디오 그룹내에서 이루어졌으며, 이것이 MPEG-2 AAC를 개발하게 된 동기가 된다.

MPEG-2 AAC의 기본적인 요구 사항은 역방향 호환성이 필요 없다는 것 외에는 MPEG-2 오디오 BC의 요구 사항과 유사하다. 가장 중요한 요구 사항들을 열거하면 다음과 같다.

- 1) 표본화 주파수는 32kHz, 44.1kHz, 그리고 48kHz를 수용한다.
- 2) 입력 채널의 구성은 1/0(모노) 및 2/0(스테레오)로부터 최대 3/2+1(전방의 좌측/우측 중앙, 후방의 좌측/우측, 저주파수 효과 채널)까지의 다른 구성이 가능하다.
- 3) 증가된 채널 수뿐만 아니라 감소된 채널 수까지 재생할 수 있는 방법이 알고리즘에 포함된다.
- 4) 오디오 품질은 ITU-R의 방송 품질에 부합하도록 3/2+1의 채널 구성에 384kbps의 비트율 이내로 한다.
- 5) 편집 가능한 최소 형태(granularity)의 오디오 포맷을 갖기 위해 미리 정의된 오디오의 접근 단위를 설정하고, 이를 반영한다.
- 6) 오차의 복구성을 높이기 위해 비트 오차 발생시 비트 스트림 동기를 유지하기 위한 구조와 오차 방지를 위한 구조를 제공한다.

이 방식은 명료한 음질을 확보하기 위하여 비트율이 크지 않으며, 사용된 알고리즘의 일부가 멀티미디어용 부호화 방식인 MPEG-4 오디오의 일부 기술로 활용된다는 장점이 있으나 기존의 MPEG-1 오디오 부호화 방식을 수용하지 못한다는 단점이 있다.

MPEG-2 AAC 인코더 및 디코더의 처리 블럭 구성도는 그림 4와 같다.

5. MPEG-4 오디오(ISO/IEC 14496-3)

MPEG-4 오디오는 저작자, 서비스 제공자, 사용자의 욕구를 충족시키기 위한 부호화 기술로서 오디오 정보

의 표현 단위인 오디오 오브젝트(object)를 대상으로 하며, 음성 및 음악과 같은 자연음(natural audio)과 구조적 표현에 근거한 합성음(synthetic audio) 모두를 재현하기 위한 기술(tool)을 제공한다. 이러한 신호의 재현은 압축과 다른 기능(scalability 또는 다른 속도의 재생 등)을 제공하고, 특히, 합성음의 재현은 문자 또는 악기 표현에 의해서, 그리고 잔향과 공간음과 같은 효과를 제공하는 특징을 부호화함으로써 만들어진다[11].

1) 자연음 부호화

MPEG-4 오디오 표준은 자연음에 대해 2kb/s부터 64kb/s까지의 범위를 갖는 비트율로 부호화한다. MPEG-4 기술의 일부인 MPEG-2 AAC 표준을 이용하면 이 비트율보다 더 높은 비트율의 오디오 부호화를 수행할 수 있다. MPEG-4 오디오에서는 전체 영역의 비트율 내에서 고품질의 오디오는 물론 특별한 기능들을 제공하기 위해서 크게 3가지 유형의 부호화 기술이 정의된다. 첫번째는 8kHz 표본화 주파수를 갖는 음성을 2~4kb/s로, 그리고 8kHz 또는 16kHz 표본화 주파수를 갖는 오디오를 4~16kb/s로 압축하는 저비트율 부호화 방식인 Parametric coding 기술, 두번째는 8kHz 표본화 주파수를 갖는 협대역 음성과 16kHz 표본화 주파수를 갖는 광대역 음성에 대해 6~24kb/s로 압축하는 중간 비트율 부호화 방식인 CELP coding 기술, 그리고 세번째는 8kHz 이상의 표본화 주파수를 갖는 오디오 신호에 대해 16~64kb/s 또는 그 이상의 비트율로 압축하는 고품질 부호화 방식인 T/F coding 기술 등이다.

오디오 신호 대역, 비트율, MPEG-4 오디오 부호화 기술과의 관계를 도시하면 그림 5와 같다.

2) 합성음 부호화

MPEG-4 오디오에서는 구조화된 입력을 기반으로 한 음을 생성한다. 즉, 문자 입력을 합성음으로 변환(TTS : Text-to-Speech)할 수 있으며, 음악을 포함한 일반 음이 합성될 수 있다. 합성음악은 정확한 음을 표현하면서 매우 낮은 비트율로 처리된다.

TTS의 경우 일반 문자 및 운율 파라미터(pitch contour, phoneme duration 등)를 가진 문자들이 이해성이 높은 합성음을 만들기 위한 입력으로 사용되며, 다음의 특별한 기능을 갖는다.

- 1) 원음의 운율을 사용한 음성합성
- 2) 운율 정보를 갖는 얼굴 애니메이션 제어
- 3) 트릭 모드 기능 : 휴지, 제시작, jump forward/backward
- 4) 문자를 위한 국제 언어 제공
- 5) 음소를 위한 국제 부호 제공
- 6) 나이, 성별, 언어, 화자의 방언을 규정하는 방법 제공

악보 합성(SDS : Score Driven Synthesis)의 경우 SAOL(Structured Audio Orchestra Language)라 불리는 특별한 합성 언어에 의해서 처리된다. 이 언어는 오케스트라가 악기를 모으도록 정의하는 제어 데이터를 만들고, 처리하는데 사용된다. MPEG-4 오디오는 합성방법을 표준화하는 것이 아니라 합성을 표현하는 방법을 표준화하는 것이다. 향후의 합성방법은 파형표(wavetable), FM, 부가어(additive), 물리적 모델링, 분리합성뿐만 아

니라 이들 방식의 비파라미터성 혼합형을 포함하는 SAOL로 표현될 수 있다.

합성의 제어는 비트스트림에 있는 악보 또는 각본을 다운로드 받음에 의해서 수행된다. 악보는 특정한 시간에 다양한 악기가 전체 음악 연주 또는 음향 효과의 발생을 하도록 하는 순차적 명령어의 집합이다. SASL(Structured Audio Score Language)이라 불리는 언어로 다운로드된 악보 표현은 새로운 음을 만드는데 사용될 수 있다. 이것은 최종 합성음상에서 더 세련된 제어를 작곡가에게 부여한다. 이렇게 세련된 제어를 필요로 하지 않는 합성 처리에 대해 설정된 MIDI 프로파일로 오케스트라를 제어하기 위하여 사용될 수 있다.

MPEG-4 오디오에서는 이해성있는 음성으로부터 고품질 다채널 오디오까지 그리고, 자연음으로부터 합성음까지 광범위한 응용분야에 이용할 수 있는 많은 기능들을 제공하고, 다음과 같은 구성을 갖는 오디오 오브젝트의 매우 효율적인 표현기술을 제공한다.

- 1) 음성신호 : 음성 부호화 도구를 사용하여 2~24kb/s의 비트율을 갖는 음성 부호화
- 2) 합성음 : 혼합형 scalable TTS는 문자 또는 운율 파라미터를 갖는 문자들을 이해성있는 합성음을 만들기 위한 입력으로 사용한다.
- 3) 낮은 대역폭 오디오 신호 : 인터넷 오디오 서비스와 같이 대역이 엄격히 제한된 채널상에서 음악과 음성을 표현하기 위한 기능
- 4) 높은 대역폭 오디오 신호 : 고품질 오디오를 제공하기 위한 기능으로 모노에서 다채널까지의 방송품질 오디오 표현이 가능한 기능
- 5) 합성 오디오 : 악보 제어 정보를 특별한 언어로 표현된 음악 악기에 적용하는 구조화된 오디오를 제공하는 기능
- 6) 정제와 복잡도를 갖는 합성 오디오 : 표준화된 파형표의 처리를 제공하는 구조화된 오디오 구현 기능.

MPEG-4 오디오 부호화는 다음과 같은 새로운 기능들을 포함한다.

가변등급(scalability) : 가변등급이란 비트 스트림의 어떤 부분만으로 복호화가 가능하며, 더 낮은 품질과 대역 또는 선택된 내용으로 의미있는 오디오 신호를 만드는 것을 의미한다. 가변등급은 MPEG-4의 주요 특징 중 하나로서 MPEG-4의 입장에서 보면 몇가지 중요한 기술들을 필요로 한다. MPEG-4의 중요한 기술적 어려움(부호화 효율을 향상시키는 것을 제외하면)은 부호화 효율의 영향을 줄이면서 가변단위를 실현하는 것이 된다.

제한된 시간의 오디오 스트림(limited time bit stream) : 몇개의 오디오 채널은 선택적으로 혹은 하나의 오디오 출력 채널로 혼합되어 재생될 수 있다. 이러한 시간 제한은 사용자가 어떤 영상물에 속해 있거나 한 부분으로 되어 있는 오디오 데이터의 재생을 제어할 수 있게 한다. 효율적인 부호화에 의해 현재의 멀티 채널 부호화에 소요되는 비트율보다 낮게 부호화할 수 있다.

피치 조절/시간축 조절(pitch change/time scale change) : 이 기능에 의해 일정 시간 구간의 신호를 뒤집어하는 것에 의해 여러가지 오디오 신호의 동기를 맞추는 것

이 가능하다. 현재 이 기능은 후처리에 의해 수행되지만 고려되어야 하는 사항은 효율적으로 피치 조절/시간축 조절을 구현하는 것이다.

· 편집성(editability) : 압축된 오디오 데이터의 조작이 필요한 관건이다. 비트 스트림의 임의의 내용에 대한 자유롭게 접근이 필요하다.

· 지연(delay) : 응용분야에 따라 긴 지연이 허용될 수도 있고(단방향의 비대화형 서비스의 경우), 허용되지 않을 수도 있다(양방향 통신분야의 경우).

6. MPEG-7 오디오

MPEG-7 오디오는 다양한 형태의 오디오 정보의 표현 방법을 표준화하는 것으로서 Multimedia(audio) Content Description Interface 라고 한다. 표현(description)은 오디오의 내용 그 자체와 연관되어 있으며, 사용자가 원하는 자료를 편리하고, 빠르게, 그리고 효율적으로 검색할 수 있는 인터페이스를 제공하는 것이다. 예를 들면, 휘파람 소리로 특정 오디오 신호를 발생하면 해당되는 운율과 음에 맞는 정보를 검색하여 제공하여 주는 표준 인터페이스의 표현 방법을 기술하는 것이다.

MPEG-7 오디오 표현은 오디오 정보의 압축/복원에는 상관없지만 오디오 데이터를 시공간에서 특정한 연관성을 갖는 객체로 부호화하는 수단을 제공한다. MPEG-7은 현재 요구사항이 정의되어 기술 제안서(Call for proposals)를 모집하고 있는 상태이며, 세부 일정은 다음과 같다[5].

- Call for proposals : 1998. 10.
- Working Draft : 1999. 12.
- Committee Draft : 2000. 10.
- Draft International Standard : 2001. 7.
- International Standard : 2001. 9.

MPEG-7 오디오의 요구사항은 다음과 같다.

- 1) Query classes : MPEG-7은 주파수 윤곽선, 음색, 화음, 주파수 프로파일, 크기 포락선, 운율, 자체 내용, 음파 추정, 음원, 공간음 구조 등과 같은 query class 들을 허용하는 오디오 표현을 제공하여야 한다.
- 2) 청각 데이터 유형 : MPEG-7은 다음과 같은 유형의 청각 데이터 검색을 제공하여야 한다.
 - MPEG-1 오디오, CD 와 같은 디지털 오디오
 - 아날로그 오디오
 - 가라오케를 포함하는 MIDI
 - Model-based audio & Production audio
- 3) 청각 sub-class : MPEG-7은 다음 sub-class 의 청각 데이터에 적용할 수 있는 메타 정보를 제공해야 한다.
 - 사운드 트랙 & 사운드 효과
 - 음악 & 음성
 - 기호화된 오디오 표현 & mixing 정보

7. 결론

디지털 오디오 부호화 방식의 국제 표준 기술 특징 및 응용 분야에 대하여 기술하였다. MPEG 오디오 국제 표준은 사용 영역과 응용 분야에 따라 다양한 적용이 가능하도록 제정되었기 때문에 일반 사용자뿐만 아니라 전문가들도 오디오 품질, 비트율, 지연시간, 처리 복잡

도, 에러 처리 능력 등의 요인을 고려한다면 합리적인 방식을 선택하여 사용할 수 있다.

MPEG-1 오디오 계층 II는 현재 CD-I, CD-Video 등에 적용되고 있기 때문에 호환성을 갖는 한 계속해서 이러한 시스템에 채택될 것이고, MPEG-1 과 MPEG-2 오디오 계층 III은 위성 라디오, 고품질 오디오의 인터넷 전송, ISDN 전송 등에 사용될 것이다.

MPEG-2 AAC는 가장 최근에 제정된 국제 표준이기 때문에 멀티채널 오디오의 새로운 기술의 부호화 표준이 될 것이며, 소비자용 및 전문가용 모두를 위한 새로운 고품질 사운드의 저장매체와 전송의 응용분야에 사용될 것이다.

MPEG-4 오디오는 미래의 멀티미디어 표준이 될 것이며, 다소 복잡하지만 오디오 또는 오디오-비주얼 데이터의 전송과 저장을 위해 모든 것을 할 수 있는 시스템이 될 것이다. MPEG-4 오디오가 확장가능하고, 개방화된 표준으로 제정되고 있기 때문에 미래에 오디오 부호화에 대한 새로운 방법이 제시되어도 적용할 수 있다.

감사의 글 : 본 연구는 정보통신부 국제공동연구 지원 사업중 "차세대 멀티미디어용 고품질 오디오 처리 기술 개발" 과제의 일환으로 수행되었습니다.

참고문헌

- [1] K. Brandenburg, M. Bosi, "Overview of MPEG-Audio: Current and Future Standards for Low Bit-Rate Audio Coding," The 99th AES Convention, New York, preprint No.4130, Oct. 1995.
- [2] K. Brandenburg and G. Stoll, "The ISO/MEPG-Audio CODEC: A Generic Standard for Coding of High Quality Digital Audio," The 92th AES Convention, Vienna, Preprint No. 2970, Mar. 24 ~ 27, 1992.
- [3] B. Grill, J. Herre, K.H. Brandenburg, E. Eberlein, J. Koller, and J. Muller, "Improved MPEG-2 Audio Multi-Channel Encoding," The 96th AES Convention, Amsterdam, Preprint No. 3865, Feb. 26 ~ Mar. 1, 1994.
- [4] ISO/IEC JTC1/SC29/WG11 MPEG94/304, "Refined Requirements for MPEG-2 Audio NBC Coding Mode Extension," 1994.
- [5] ISO/IEC JTC1/SC29/WG11 WG11/N1734, "Second draft of MPEG-7 requirement," July, 1997.
- [6] ISO/IEC 11172-3, 1993.
- [7] 最新 MPEG, マルチメディア通信研究會, 1994. 7.
- [8] 홍진우, "멀티채널 오디오 부호화(MPEG-2) 기술," ETRI 전자통신 기술동향, No.4, 1995.10.
- [9] R.N.J. Veldhuis, M. Breeuwer, and R.G. Van Der Waal, "Subband Coding of Digital Audio Signals, Philips Journal of Research, Vol.44, No. 2, pp.329 ~ 343, 1989.
- [10] G. Theile, "The New Sound Format 3/2-Stereo," The 94th AES Convention, Berlin, preprint 3550, Mar. 1993.
- [11] ISO/IEC JTC1/SC29/WG11 WG11/N0973, "MPEG-4 Proposal Package Description(PPD) - Revision 2," 1995.

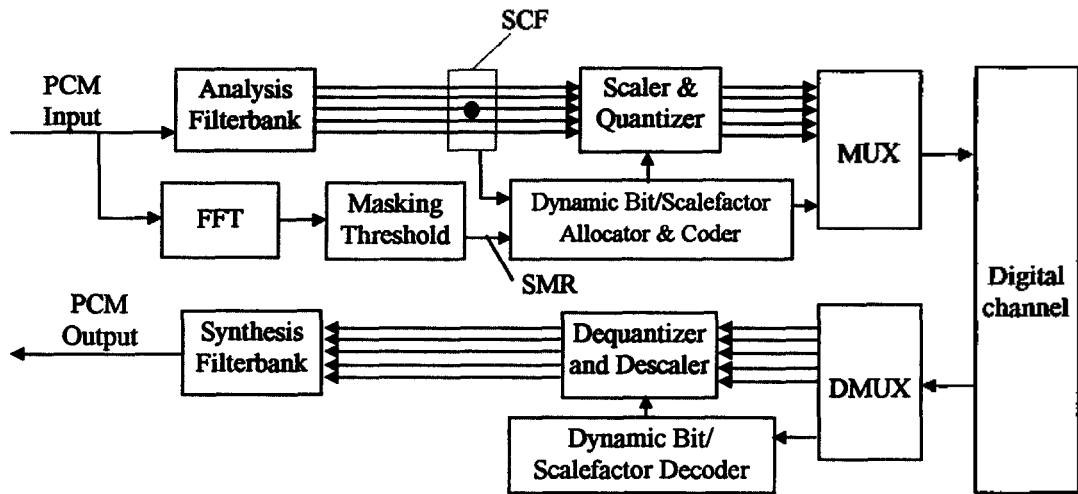


그림 1. MPEG-1 오디오 인코더 및 디코더의 구조(계층 I, II)

계층 I 프레임 구조(8ms)
(48kHz 표본화에 대한 384 PCM 오디오 입력 샘플)

헤더 (Header)	CRC	비트할당 (bit allocation)	scale factor	서브밴드 샘플	보조 데이터
12비트 sync. 20비트 system info.	16 비트	4비트 (linear)	6비트 (linear)	32PCM 오디오 입력 샘플에 일치하는 1 서브밴드 샘플 (12 서브밴드 = 32 샘플)	이점의 길이

계층 II 프레임 구조(24ms)
(48kHz 표본화에 대한 1152 PCM 오디오 입력 샘플)

헤더 (Header)	CRC	비트할당 (bit allocation)	SCPSI	SCF (scale factor)	서브밴드 샘플	보조 데이터
12비트 sync. 20비트 system info.	16 비트	지역 sb (4비트) 중역 sb (3비트) 고역 sb (4비트)	2비트 00 01 10 11	6비트 (linear)	GR 1 2 3 4 5 6 7 8 9 10 11 12 96PCM 오디오 입력 샘플에 일치하는 3 서브밴드 샘플 각 서브밴드 샘플의 3샘플로 구성되는 12granule (96)	이점의 길이

그림 2. MPEG-1 스테레오 오디오 비트스트림의 구조

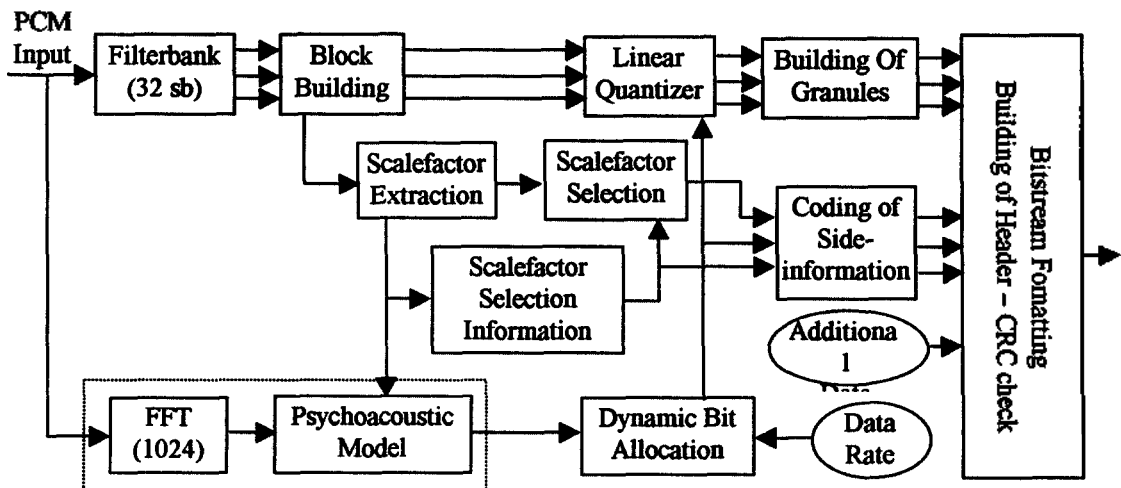
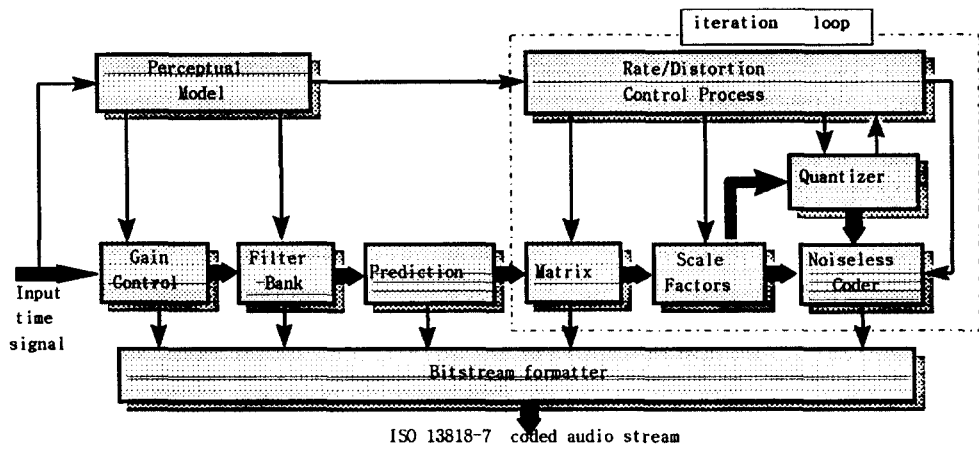
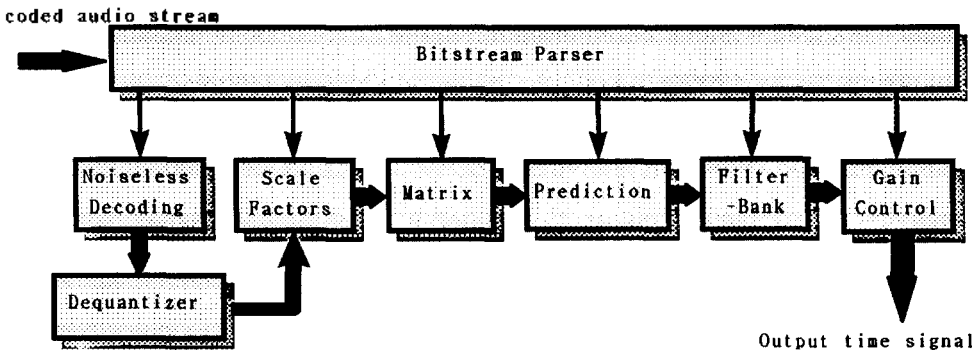


그림 3. MPEG-2 오디오 인코더 블록 구성도



(a) Encoder



(b) Decoder

그림 4. 지각 부호화 방법의 기본 구조

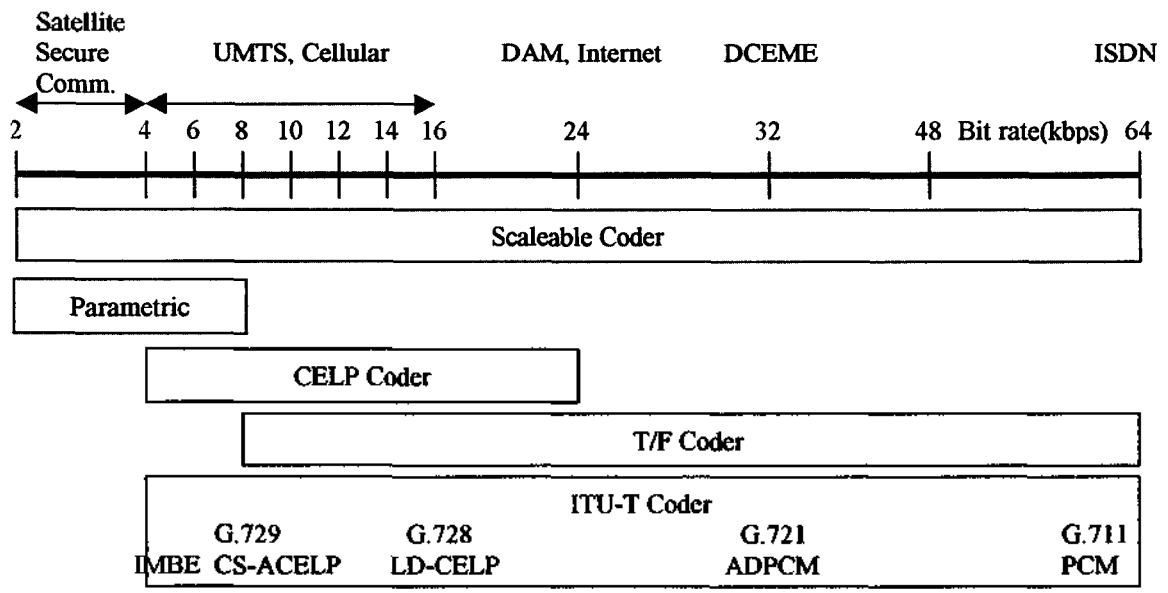


그림 5. MPEG-4 오디오의 부호화 영역