

다채널 오디오 코딩을 위한 MPEG Surround-2부: 다양한 모드 및 툴들

MPEG Surround for Multi-Channel Audio Coding-Part 2: Various Modes and Tools

방 희 석*
(Hee-Suk Pang*)

*세종대학교 전자공학과
(접수일자: 2009년 8월 31일; 채택일자: 2009년 9월 17일)

본 논문에서는 MPEG Surround의 다양한 모드들과 툴들에 대해서 소개한다. MPEG Surround 바이노럴 모드의 경우 HRTFs (Head-Related Transfer Functions)를 이용하여 가상의 5.1 채널을 지원하기 때문에 휴대용 오디오 기기에서 헤드폰이나 이어폰을 이용하여 재생할 수 있다. 또한, MPEG Surround는 부가 정보 없이 스테레오 채널을 5.1 채널을 지원하는 Enhanced Matrix 모드, 기존의 3D 코딩 신호에 적용할 수 있는 3D 스테레오 모드, 복호화 과정에서 연산량을 크게 줄일 수 있는 저전력 (Low Power) 버전 등을 지원한다. 이 외에도, MPEG Surround는 아티스틱 다운믹스 신호에 대해 적용하는 Arbitrary Downmix Gain (ADG) 툴, 기존의 행렬 근거 방법들에 의한 다운믹스 신호에 적용되는 행렬 호환성 (Matrix Compatibility) 툴, 고비트율을 사용할 수 있을 때 적용되는 잔차 신호 코딩 (Residual Coding) 툴, 박수 소리 등 특정 신호에 대해 사용될 수 있는 GES (Guided Envelope Shaping) 툴 등을 지원한다. 중요한 모드 및 툴들에 대해 여러 기관에 의해 수행된 청취 평가의 결과도 함께 보이고 있다.

핵심용어: MPEG Surround, 바이노럴 모드, HRTF, Enhanced Matrix 모드, 저전력 버전

투고분야: 뉴미디어 분야 (13,3)

An overview of various modes and tools of MPEG Surround is provided. Because the binaural mode of MPEG Surround supports the virtual 5.1-channel playback based on HRTFs, it can be played via headphones and earphones for portable audio devices. MPEG Surround also supports the enhanced matrix mode which converts stereo signals to 5.1-channel signals without side information, the 3D stereo mode which deals with 3D-coded signals, the low power version which greatly reduces the computational load in the decoding process. Besides, MPEG Surround provides the arbitrary downmix gains (ADGs) tool which is applied to artistic downmix signals, the matrix compatibility tool which is applied to downmix signals by conventional matrix-based methods, the residual coding tool which can be used at high bit rates, and the GES tool which is applied to specific sound such as applause. The listening test results by various companies and organizations are also presented for important modes and tools.

Keywords: MPEG Surround, Binaural Mode, HRTF, Enhanced Matrix Mode, Low Power Version

ASK subject classification: New Media (13,3)

I. 서론

본 논문의 2부에서는 MPEG Surround의 다양한 모드들과 툴들에 대해 소개한다. 이들 중 일부는 RM0에서부

터 표준에 포함된 것들도 존재하며 [1], 또한 많은 기술들이 표준화 과정에서 제안되어 최종적인 표준에 포함되었다 [2].

우선 MPEG Surround의 모드에는 가상 5.1 채널을 지원하여 기존의 오디오 재생 기기에서 헤드폰이나 이어폰을 이용한 재생이 가능하게 하는 바이노럴 모드 (Binaural Mode), 이미 3D 정보가 포함되어 있는 신호에 대한 다채널 합성 방법을 지원하는 3D 스테레오 모드 (3D Stereo

책임저자: 방 희 석 (hspang@sejong.ac.kr)
134-747 서울시 광진구 군자동 세종대학교
전자정보공학대학 전자공학과
(전화: 02-3408-3727; 팩스: 02-3408-4329)

Mode), 부가 정보 없이 스테레오 채널로부터 5.1 채널을 합성하는 Enhanced Matrix 모드, 복호화 과정에서 연산량을 대폭 줄이면서도 음질은 거의 유사하게 유지하는 저전력 버전 등을 들 수 있다. 또한, MPEG Surround의 툴들은 아티스틱 다운믹스 신호에 적용되는 Arbitrary Downmix Gains (ADGs) 툴, MPEG Surround가 아닌 기술에 의해 생성된 다운믹스 신호에 적용되는 행렬 호환성 (Matrix Compatibility) 툴, 고비트율을 사용하여 고음질을 지원할 수 있는 잔차 신호 코딩 (Residual Coding) 툴, 박수 소리 등 특수한 소리에 대한 음질을 향상시키는 GES (Guided Envelope Shaping) 툴 등을 들 수 있다. 이들 중 바이노럴 모드와 3D 스테레오 모드는 2장에서 소개하고, 나머지 모드들과 툴들은 3장에서 소개한다. 또한, 중요한 모드들과 툴들에 대한 성능 평가는 MPEG에서의 성능 평가서 [3][4]를 근거로 하여 4장에서 소개한다.

II. 바이노럴 모드와 3D 스테레오 모드

MPEG Surround의 표준화 과정 중 여러 개의 스피커를 이용한 다채널 재생 뿐 아니라 헤드폰이나 이어폰을 이용한 가상 5.1 채널 재생을 지원하자는 의견 [5]과, 이미 HRTF들로 가상의 5.1 채널을 지원하도록 처리가 된 스테레오 다운믹스 신호에 대해 MPEG Surround를 적용하여 다채널 재생을 지원할 수 있게 하자는 의견 [6]이 제기되었다. 이에 대해 MPEG 오디오 회의에서는 다양한 토론을 통해 전자의 기술을 바이노럴 모드, 후자의 기술을 3D 스테레오 모드로 명명하고 CE (Core Experiment)를 진행하였다 [7]. 그 후 이에 응한 다양한 기술들에 대해 성능 평가가 이루어졌고 [8], 각 기술에 대해 최적의 안을 선택하여 표준안을 선택하였다. 이 장에서는 이 둘에 대해서 소개한다.

2.1. 바이노럴 모드

바이노럴 모드에서 입력은 다운믹스 신호와 MPEG Surround의 비트열이며, 특히 이들은 5-1-5 구조나 5-2-5 구조를 가져야 한다. 이는 바이노럴 모드를 위한 다운믹스 신호와 비트열이 따로 존재하는 것이 아니고, 다채널로 재생을 할 것인지 아니면 헤드폰이나 이어폰을 통해 가상의 5.1 채널을 스테레오로 재생할 것인지를 복호화부에서 선택할 수 있음을 의미한다.

바이노럴 모드에서의 복호화 과정은 그림 1과 같다 [9]. 다운믹스 신호는 다채널 복호화 과정과 같이 우선 QMF 및 Nyquist 필터 분석을 통해서 Hybrid 서브밴드 영역으로 변환된다. 한편 MPEG Surround 비트열은 공간 파라미터 정보로 변환되는데, 이 때 HRTF의 3D 정보를 포함하여 변환된다. 즉, HRTF가 서브밴드 밴드 영역에서 적용되는 것으로 볼 수 있다. 다운믹스 신호와 3D 정보를 포함한 공간 파라미터를 이용해 바이노럴 신호를 생성한 후, Nyquist 필터 및 QMF 합성을 수행하면 시간 영역의 신호가 합성된다. 이 과정은 5.1 채널 신호를 모두 합성한 후 각 채널별로 HRTF를 적용하는 것을 간략화한 구조로 볼 수 있으며, 따라서 연산량 (Computational Load)나 메모리 사용량 (Memory Requirement)을 줄일 수 있다. MPEG Surround에서는 특정 HRTF의 사용을 정의하고 있지 않기 때문에, 오디오 기기에 대해서 사용자에게 특화된 HRTF를 사용하여 개인별로 최적화된 3D 오디오를 재생하는 것이 가능하다. MPEG Surround에서 HRTF는 서브밴드 영역에서 적용하는 것과 파라미터 형태로 적용하는 것 등 두 가지가 가능한데, 이에 대한 자세한 정보는 표준의 Annex B와 C에 설명되어 있다 [2].

그림 2에서는 5.1 채널 신호 중 FL, FR, C, SL, SR 채널에 대한 바이노럴 모드의 개념도를 표현하였다. 이는 각 채널의 신호에 대해 해당 각도의 HRTF들을 적용한 후 왼쪽 헤드폰을 통해 FL, SL, C 채널이, 오른쪽 헤드폰을

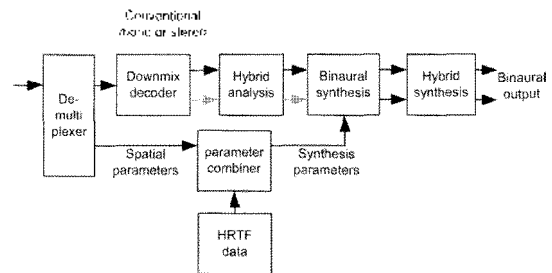


그림 1. MPEG Surround 바이노럴 디코더
Fig. 1. MPEG Surround binaural decoder.

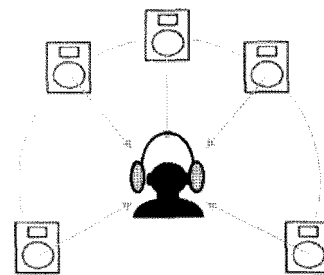


그림 2. 바이노럴 모드의 개념도
Fig. 2. Concept of the binaural mode.

통해 FR, SR, C 채널이 재생되는 것으로 볼 수 있다. 좀 더 현실적인 공간감을 느끼기 위해서 잔향기 (Reverb-berator)가 사용될 수 있는데, 바이노럴 모드에서는 시간 영역 혹은 서브밴드 영역에서 잔향기를 추가할 수 있다. 이에 대한 간단한 구조가 MPEG Surround 표준 중 Annex D에 소개되어 있으며 [2], 잔향기는 사용자가 선호하는 것으로 구현할 수 있다.

MPEG Surround에서는 사용자가 다운믹스 신호와 다 채널 신호를 선택하여 재생할 수 있기 때문에, 바이노럴 모드의 경우에도 헤드폰이나 이어폰을 통해 바로 모노 혹은 스테레오의 다운믹스 신호를 재생하는 것이 가능하다.

2.2. 3D 스테레오 모드

바이노럴 모드에서는 복호화 과정 중 HRTF를 적용하는 반면에, 3D 스테레오 모드의 경우는 MPEG Surround 부호화부 등에서 미리 HRTF를 적용하여 3D 처리가 적용된 다운믹스 신호를 생성한다. 이는 휴대용 오디오 기기 등을 통해 3D 사운드를 재생하는 것을 목적으로 하고 있으며, 현재 일부에서 상용화가 되고 있는 서비스이기도 하다. 이 경우, 다운믹스 신호는 HRTF가 적용되어 있기 때문에 바로 MPEG Surround 복호화를 적용하면 출력 다채널 신호의 음질에 문제가 생기며, 이를 위해 추가적인 처리 과정을 요구하게 된다.

그림 3에서 3D 스테레오 모드의 복호화기를 보이고 있다 [2]. 이는 기본적인 MPEG Surround의 다채널 및 바이노럴 모드와 유사한 구조를 가지는데, MPEG Surround 부호화부에서 사용한 HRTF 정보가 비트열에 포함된다. 즉, 다운믹스 신호만 재생할 때는 이미 HRTF 정보가 포함되어 있는 다운믹스 신호가 3D 오디오의 형태로 재생되고, 이를 다채널 신호로 합성할 때는 비트열의 HRTF 정보를 이용해 다운믹스 신호에서 HRTF

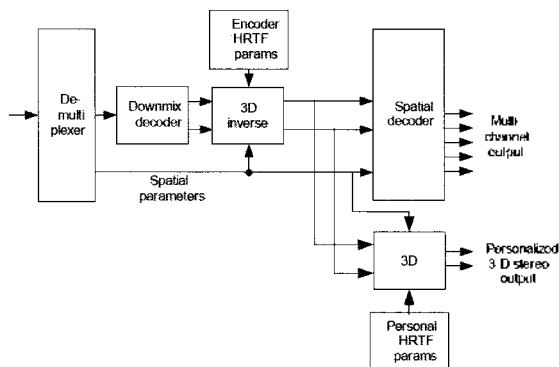


그림 3. MPEG Surround의 3D 스테레오 복호화기
Fig. 3. MPEG Surround 3D stereo decoder.

정보를 제거하는 보정 작업을 한 후 MPEG Surround 복호화 과정을 계속 수행한다. 한편, 그림의 아래쪽에서 보이는 바이노럴 복호화 모드는 다운믹스 신호에 포함된 HRTF 정보를 제거한 후 다시 복호화부에서 HRTF를 적용하는 과정이며, 이는 MPEG Surround 부호화부에서 일괄적으로 사용된 HRTF를 사용자별 HRTF로 변경하는 것을 의미한다.

III. 다양한 모드 및 톨들

MPEG Surround는 RMO 및 표준화 과정을 통해 많은 모드 및 톨들이 포함되어 있다. 이 장에서는 이들에 대해서 간단히 소개한다.

3.1. Enhanced Matrix 모드

특정 환경에서는 MPEG Surround의 부가 정보를 보내는 것이 부적절하거나 아예 불가능할 수 있다. 예를 들어, 비트열에 신호 이외의 추가 정보를 더하는 기능이 없는 다운믹스 코덱이 있을 수도 있고 아날로그 시스템 또한 부가 정보를 같이 보내는 것이 어렵다. 따라서, MPEG Surround에서는 사용 범위를 넓히기 위해서 부가 정보 없이 다채널 코딩을 지원하는 모드를 정의하고 있으며 이를 Enhanced Matrix 모드로 정의한다.

그림 4에서 Enhanced Matrix 모드에 대한 구조를 보이고 있다 [9]. 복호화부에서 스테레오 다운믹스 신호를 Hybrid 서브밴드 영역으로 변환하고 좌우 신호를 분석하여 미리 정의된 테이블을 이용해 파라미터 밴드별로 CLD, ICC 등의 공간 파라미터를 결정한다. 일단 공간 파라미터의 값들이 정의되면 기존의 MPEG Surround 복호화 과정에 의해서 다채널 및 바이노럴 복호화 모드가 모두 지원 가능하다. 이 모드는 여러 주파수 밴드별 및 시간별로 값이 변하는 공간 파라미터를 사용하기 때문에, 기존의 행렬 기반 방법들에 비해 향상된 성능을 보인다 [4].

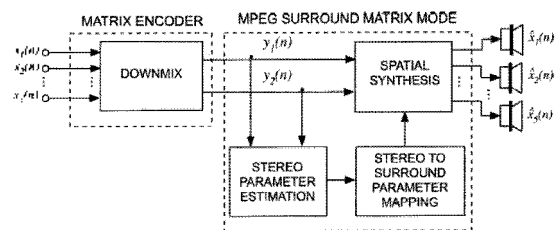


그림 4. Enhanced matrix 모드
Fig. 4. Enhanced matrix mode.

3.2. Arbitrary Downmix Gain (ADG) 툴

DVD 등의 오디오 매체에서는 5.1 채널 등의 다채널 신호 뿐 아니라 스테레오 다운믹스 신호까지 같이 제공되는 경우가 많다. 특히 다운믹스 신호들이 사운드 엔지니어들에 의해 믹싱 과정을 통해 생성되는 경우, 각각의 음원들에 대해 다른 믹싱 파라미터 및 다양한 이펙트들이 추가되게 된다. 이는 MPEG Surround 부호화기에 의해 생성되는 다운믹스와 사운드 엔지니어에 의해 생성되는 다운믹스 신호가 매우 다를 수 있음을 의미한다. 전자의 경우를 자동 다운믹스 (Automatic Downmix), 후자의 경우를 아티스틱 다운믹스 (Artistic Downmix)라고 정의한다.

MPEG Surround에서는 자동 다운믹스 뿐 아니라 아티스틱 다운믹스도 사용할 수 있다. 한편, 아티스틱 다운믹스 신호의 경우는 다운믹스 신호를 바로 재생할 때는 문제가 되지 않지만, MPEG Surround 복호화 과정을 통해서 다채널 신호를 재생할 때 문제가 될 수 있는데 이는 MPEG Surround의 복호화 과정이 기본적으로 자동 다운믹스에 기반을 두고 있기 때문이다. 아티스틱 다운믹스는 기본적으로 주파수 영역별 및 채널별 게인이 자동 다운믹스와 다르며, 특히 다채널 신호에서 모노나 스테레오 신호로의 믹싱 과정에서 원신호의 모든 채널의 소리가 포함되는 경우와 일부 채널의 소리 자체가 제외되는 경우로 나눌 수 있다. 이 절에서는 전자에 대한 처리 방법을 설명하며, 후자의 경우는 다음 절에서 설명되는 잔차 신호 코딩을 이용하여야 한다.

그림 5에서 아티스틱 다운믹스 신호에 대한 처리 과정을 보이고 있다. 이는 ADG라는 일종의 다운믹스 신호에 대한 파라미터 밴드별 게인 조정 정보를 사용하여 처리가 되며, MPEG Surround에서의 다채널 합성 과정이 진행되기 이전에 다운믹스 신호에 대해 채널별로 적용된다. ADG는 CLD와 같은 양자화 및 허프만 테이블을 이용하는 등 동일한 포맷을 가진다. 또한, ADG는 부호화기에서 다운믹스 과정을 수행할 때 클리핑 (Clipping)을 방지하기 위해 동적 범위 제어 (Dynamic Range Control)를 적용할 수 있으며, 이에 대한 내용은 표준의 Annex F에서 소개하고 있다 [2].

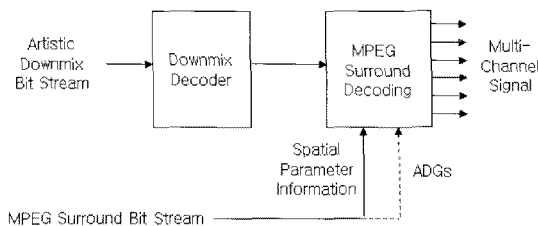


그림 5. 아티스틱 다운믹스 신호의 처리 과정
Fig. 5. Processing of an artistic downmix signal.

3.3. 잔차 신호 코딩 (Residual Coding) 툴

파라미터에 근거한 다채널 오디오 코딩 방법은 저비트율에서 좋은 음질의 다채널 신호를 생성할 수는 있지만 고비트율에서 원음과 동일한 수준의 음질은 재생할 수 없으며, 이는 각 채널별로 따로 코딩을 하는 방법으로만 가능하다.

그림 6에서는 MPEG Surround의 잔차 신호 코딩에 의한 합성음의 구조를 보이고 있다 [9]. 이는 기본적으로 저주파 영역에 대해서는 잔차 신호로, 고주파 부분에 대해서는 기존의 방법대로 파라미터를 기반으로 한 합성 신호로 구성하는 일종의 하이브리드 형태를 가진다. 이와 같은 구조를 가지는 이유는 사람이 저주파 부분에 대해 더욱 민감하기 때문이다. 잔차 신호는 MPEG2 AAC LC (Low Complexity) 프로파일을 사용해서 부호화 및 복호화하여야 한다 [2]. 또한, 이 과정에서는 MDCT-QMF 변환을 이용해야 한다고 규정하고 있으므로, 자세한 내용은 표준을 참고하기 바란다 [2].

잔차 신호 코딩 툴을 이용하면 MPEG Surround의 비트율에 따라서 잔차 신호의 대역폭 및 AAC 부호화에 사용되는 비트율을 조절할 수 있으며, 따라서 비트율에 따른 음질의 조절이 가능하다. 또한, AAC 복호화가 불가능한 경우 등에 대해 전체 비트열 중 잔차 신호를 복호화하지 않고 파라미터 기반의 복호화 과정만 수행하는 것도 가능하다.

3.4. Guided Envelope Shaping (GES) 툴

MPEG Surround의 표준화 과정 중에 박수 소리와 같은 특정 음에 대해서 음질 저하가 보고되었고, 이의 원인인 시간 영역에서의 포락선의 왜곡을 해결하기 위해 GES 툴이 표준에 포함되었다. GES에서는 그림 7 [9]과 같이 광대역의 세밀한 시간 포락선 정보를 각 채널에 대해 부가

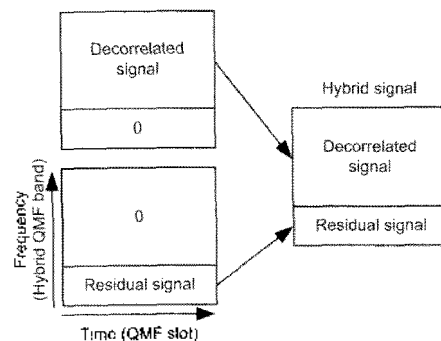


그림 6. 잔차 신호 코딩에 의한 합성 신호의 구조
Fig. 6. Structure of synthesized signals based on residual coding.

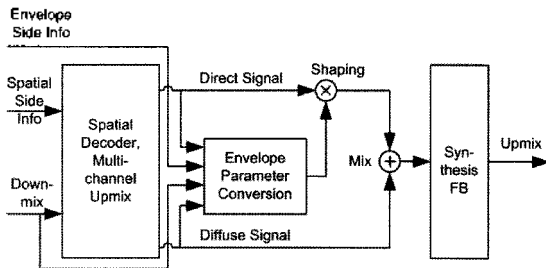


그림 7. GES 복호화 과정
Fig. 7. GES decoding.

정보로 사용하며, 따라서 다채널 신호에서의 정확한 시간-공간 광형을 복원할 수 있게 한다. MPEG Surround의 복호화부에서는 공간 파라미터 중 ICC를 처리하는 과정에서 직접 신호(Direct Signal)와 디퓨즈 신호(Diffuse Signal)를 섞는 과정을 거치는데, GES는 직접 신호 부분만 수정하며 특히 박수 소리와 같은 신호에 대한 음질 향상을 기대할 수 있다.

3.5. 행렬 호환 (Matrix Compatibility) 틀

MPEG Surround 부호화부에서는 자동 다운믹스나 아티스틱 다운믹스 이외에도 기존의 행렬 기반(Matrix-Surround, MTX) 방법과 호환성을 가지는 스테레오 다운믹스 신호를 생성할 수 있다. 이 기능은 스테레오 신호에 대해 행렬 기반의 5.1 채널 신호는 복호화할 수 있으나 공간 파라미터는 복호화하지 못하는 기기에 사용될 때 유용하다. 이 틀은 서브밴드 영역에서 작동하며, MPEG Surround 부호화부에서는 자동 다운믹스 신호가 생성된 후 행렬 처리를 거쳐서 MTX 다운믹스로 변형되며, 복호화부에서는 MTX 다운믹스 신호가 역행렬 처리되어 자동 다운믹스 신호로 변형된 뒤 MPEG Surround 복호화 과정을 계속하게 된다.

3.6. Parameter Smoothing 틀

비트율이 매우 낮을 때는 양자화 간격을 넓혀서 세밀하지 않게 부호화하는 것이 필요할 수 있다. 이와 같은 방법은 특정 신호에 대해 음질 저하를 초래할 수 있는데, 예를 들어 공간 파라미터가 두 양자화 구간의 중간에 있으면 최종 양자화 값이 둘 사이를 계속 이동할 수 있다. 따라서, 이 틀이 작동되면 공간 파라미터의 복호화 시 값이 서서히 변화하도록 해서 급격한 변화를 막는 역할을 수행한다.

3.7. 저전력 버전

MPEG Surround의 복호화 버전은 기본 버전 이외에도

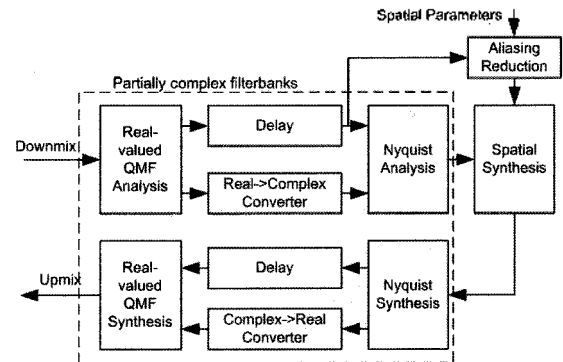


그림 8. 저전력 모드의 복호화부
Fig. 8. Decoding process of the LP mode.

연산량을 대폭 낮춘 저전력 (Low Power, LP) 버전을 지원한다. 기본 버전은 저전력 버전과 구분하기 위해서 고음질 (High Quality, HQ) 버전으로 표시한다. LP 버전은 HQ 버전 중 QMF 필터뱅크들과 역상관기 (Decorrelator)를 더 간단한 구조로 바꾼 것이다. 두 버전 모두 동일한 MPEG Surround 비트열에 대해 동작하며, 복호화 과정 중에서 둘 중 하나를 선택할 수 있다.

그림 8에서 LP 버전의 구조를 보이고 있다 [9]. HQ 버전에서는 신호의 분석 및 합성을 위해서 복소수 QMF를 사용하며, 이는 MPEG Surround 부호화기의 전체 연산량 중 많은 부분을 차지한다. LP 버전에서는 저주파 부분에 대해서 복소수 QMF 변환을 실수 QMF 변환으로 대체하는데, 즉 복소수 QMF 분석은 실수 QMF 분석 및 실수-복소수 변환으로, 복소수 QMF 합성은 복소수-실수 변환 및 실수 QMF 합성으로 대체한다. 이는 실수 QMF 필터뱅크가 임계 샘플링 (Critical Sampling)에 의해 파라미터 밴드들의 경계에서 앨리어싱 (Aliasing)이 발생할 수 있으며, 이는 저주파 부분에서 그 효과가 더 크게 나타나기 때문이다. 고주파 부분에 대해서는 상대적으로 이 효과가 적기 때문에 저주파 부분과 시간 지연만 맞추어 준다. 한편, 고주파 부분에 대해 파라미터 밴드 경계에서 정현파에 가까운 신호 (Tonal Signal)가 있을 경우 Aliasing Reduction 부분에서 공간 파라미터들의 정보가 파라미터 밴드별로 급격한 변화를 가지지 않도록 수정된다. HQ 버전에서 명확한 신호 분리를 위해 고음질의 격자 IIR 역상관기 (Lattice IIR Decorrelator)를 사용함에 반해, LP 버전에서는 실수 역상관기와 함께 Parametric Stereo [10]에서 사용하는 역상관기를 섞어서 사용한다. 다운믹스 신호가 모노인 경우 등 특정한 구조에서는 좀 더 간단한 역상관기의 구조가 사용될 수 있다 [9]. 표 1에서는 HQ 버전과 LP 버전의 차이를 이상에서 기술한 것들 및

표 1. 고음질 (HQ) 및 저전력 (LP) 버전의 비교
Table 1. Comparison of HQ and LP versions.

기능	HQ 버전	LP 버전
필터뱅크	복소수 QMF	실수 QMF 기반
역상관계	고음질	저복잡도
앨리어싱 제거	없음	있음
적용 HRTF	서브밴드 및 파라미터	파라미터
잔차신호 코딩	모든 영역 가능	저주파 영역

바이노럴 복호화 모드에서 사용가능한 HRTF, 잔차 신호 코딩 적용 범위 등에 대해 보이고 있다.

III. 성능 평가

MPEG Surround의 성능은 본 논문의 1부에서 소개한 것처럼 MPEG에서 다양한 청취 평가를 수행하여 공식 문서로 공표한 바 있다. 본 논문의 2부에서는 13음 기반으로 잔차 신호 코딩 틀 및 ADG 틀, [4]를 기반으로 Enhanced Matrix 모드, 바이노럴 모드, 저전력 버전 등 5 가지 모드 및 틀에 대한 성능을 보인다. 본 논문의 1부의 표 1에서 [3]의 테스트 아이템을, 표 2에서 [4]의 테스트 아이템을 보이고 있다. 청취 평가 방법인 MUSURA 방법 [11]을 포함하여, 청취 평가와 관련된 일반적인 내용은 본 논문의 1부 내용 중 성능 평가 부분을 참고하기 바란다. 특별한 언급이 없을 경우 MPEG Surround는 고음질 버전을 사용하여 5.1 채널을 재생하며, 전체 비트율 (다운믹스 코덱+ 부가 정보)은 160 kbps를 사용한다.

우선 그림 9에서는 13명의 청취자에 대한 잔차 신호 코딩 틀의 성능 평가 결과를 보이고 있다 [3]. 참고로 [3]은 MPEG 여러 기고문들에서의 결과들 중 유효한 것들을 추출한 것으로서, 그림 9와 10에서는 전체 평균과 함께 각 아이템들의 결과도 보이고 있다. 그림 9에서는 은닉 원음 (Hidden Reference), 앵커음 (3.5 kHz BW limited), AAC LC 프루파일로 다채널 코딩 (AAC-LC)을 192 kbps와 320 kbps의 비트율로 한 것, AAC와 MPEG Surround를 사용한 것 (SAC)에 대해 160 kbps, 192 kbps, 320 kbps의 비트율을 사용한 것 등의 7개 신호에 대해서 순서대로 보이고 있다. 여기에서 SAC (Spatial Audio Coding)는 MPEG Surround의 개명 전의 이름이며, SAC 160 kbps는 잔차 신호 코딩 틀을 사용하지 않은 것이고, SAC 192 kbps와 SAC 320 kbps는 잔차 신호 코딩 틀을 사용한 것을 의미한다. 각 아이템 및 평균에서 마지막 3개의 결과들 보면 SAC 160 kbps, SAC 192 kbps, SAC 320 kbps에

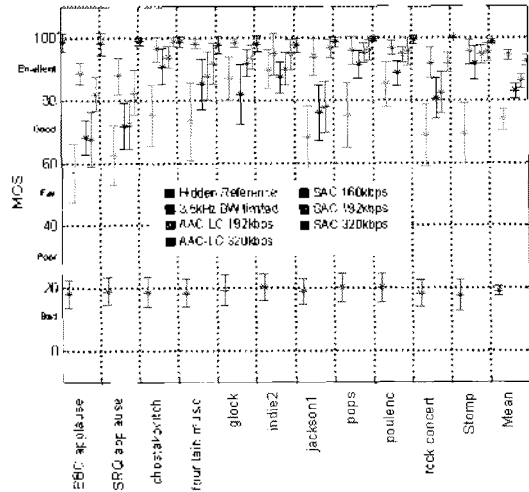


그림 9. 잔차 신호 코딩에 의한 청취 평가 결과
Fig. 9. Listening test results of residual coding.

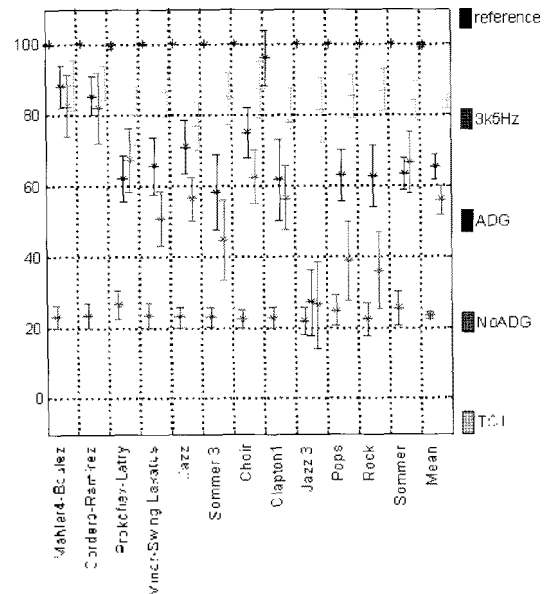


그림 10. 아티스틱 다운믹스 신호에 대한 청취 평가 결과
Fig. 10. Listening test results of artistic downmix signals.

서 점차 음질이 향상됨을 알 수 있으며, 이는 잔차 신호 코딩의 효용성을 입증한다. 특히 4번째 결과인 AAC-LC 320 kbps와 7번째 결과인 SAC 320 kbps는 그 성능이 거의 유사함을 알 수 있다.

그림 10에서는 10명의 청취자에 대해 ADG 틀의 성능 평가를 한 것을 보이고 있다 [3]. 그림에서는 은닉 원음 (reference), 앵커음 (3.5 kHz), 아티스틱 다운믹스 신호와 함께 ADG 틀을 사용한 것 (ADG), 아티스틱 다운믹스 신호만 사용하고 ADG 틀을 사용하지 않은 것 (NoADG), 자동 다운믹스 신호를 사용한 일반적인 MPEG Surround (TCI)의 순서로 결과를 보이고 있다. 이 청취평가에서는 아티스틱 다운믹스 생성을 위해 MPEG Surround의 표준

화 시 사용한 테스트 아이템들과는 다른 아이템들을 사용하였으며, 이들에 대한 자세한 설명은 [3]을 참고하기 바란다. 결과에서 아티스틱 다운믹스 신호에 ADG 툴을 사용하는 것이 사용하지 않는 것에 비해서 성능이 좋음을 알 수 있다. 아티스틱 다운믹스 신호와 ADG 툴을 사용하는 경우가 TCI에 비해서 성능이 떨어지는 이유 중 하나는 ADG를 사용하더라도 양자화나 노이즈 증폭과 같은 문제로 인해서 아티스틱 다운믹스 신호를 자동 다운믹스 신호로 완전하게 복원하기 힘들기 때문이다.

다음으로 MPEG Surround 성능 평가 보고서를 근거로 하여 Enhanced Matrix 모드, 바이노럴 복호화 모드, 저전력 버전 등의 결과에 대해 보인다 [4]. 우선 그림 11에서는 41명의 청취자에 대해 AAC와 MPEG Surround를 사용한 것 (AAC_MPS), AAC와 MPEG Surround를 Enhanced

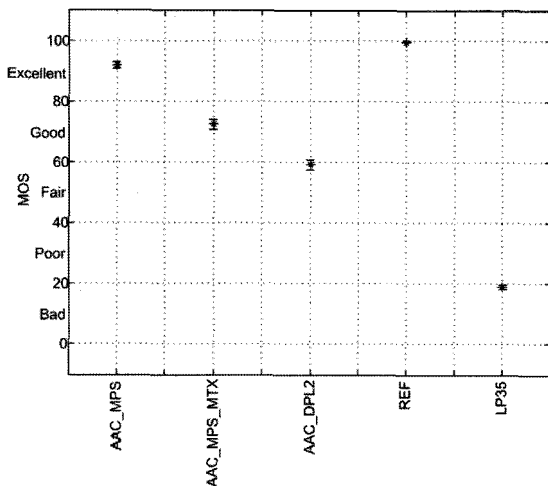


그림 11. Enhanced Matrix 모드 대한 청취 평가 결과
Fig. 11. Listening test results of the enhanced matrix mode.

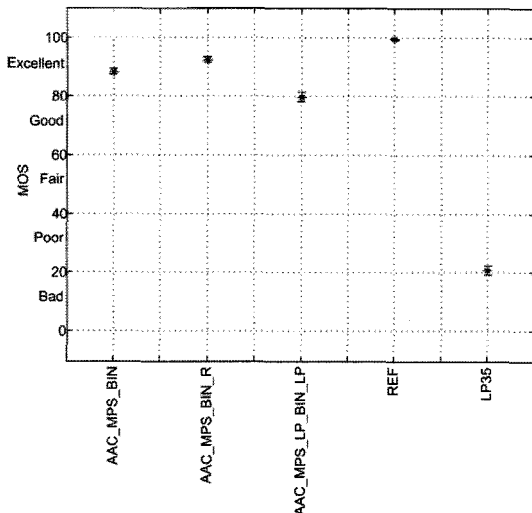


그림 12. 바이노럴 모드에 대한 청취 평가 결과
Fig. 12. Listening test results of the binaural mode.

Matrix 모드로 사용한 것 (AAC_MPS_MTX), AAC와 Dolby Prologic II를 사용한 것 (AAC_DPL2)의 결과를 은닉 원음 (REF) 및 앵커음 (LP35)의 결과와 함께 보이고 있다 [4]. 이들 결과로부터 MPEG Surround의 Enhanced Matrix 모드와 Dolby Prologic II는 모두 부가정보를 사용하지 않지만 전자가 후자보다 월등한 성능을 보임을 알 수 있다. 또한, MPEG Surround에서 부가 정보를 사용하는 경우와 사용하지 않는 경우에도 음질 차이가 나는 것을 확인할 수 있다.

그림 12에서는 30명의 청취자 결과를 바탕으로 고음질 버전의 MPEG Surround 바이노럴 모드 (AAC_MPS_BIN), 저전력 버전의 MPEG Surround 바이노럴 모드 (AAC_MPS_LP_BIN_LP), MPEG Surround를 5.1 채널로 출력한 후 각각의 채널에 HRTF들을 적용한 경우 (AAC_MPS_BIN_R)의 결과를 은닉 원음 (REF) 및 앵커음 (LP35)의 결과와 함께 보이고 있다 [4]. 결과로부터 MPEG Surround로 5.1 채널 출력 후 HRTF들을 적용한 것, 고음질 버전의 MPEG Surround 바이노럴 모드, 저전력 버전의 MPEG Surround 바이노럴 모드 순으로 음질이 좋음을 알 수 있다. 그러나, MPEG Surround의 바이노럴 모드들은 연산량 및 메모리 사용량을 상대적으로 크게 줄이면서도 80~90점 수준의 좋은 음질을 지원한다는 장점이 있다.

마지막으로 18명의 청취자에 대해 HE-AAC와 MPEG Surround를 함께 사용한 경우에 대한 청취 평가 결과를 보인다. 그림 13에서는 이 경우의 조합에 대해 160 kbps의 전체 비트율을 사용하고 MPEG Surround가 고음질 버전인 것 (HE-AAC_MPS_A)과 저전력 버전인 것 (HE-AAC_MPS_A_LP), 64 kbps의 전체 비트율을 사용하고

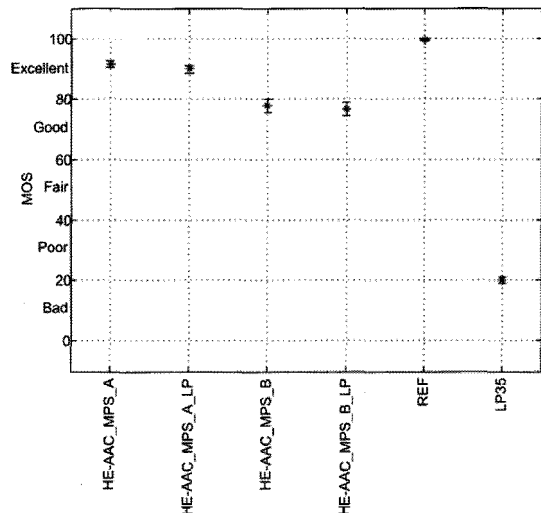


그림 13. 저전력 버전에 대한 청취 평가 결과
Fig. 13. Listening test results of the low power mode.

MPEG Surround가 고음질 버전인 것 (HE-AAC_MPS_B) 과 저전력 버전인 것 (HE-AAC_MPS_B_LP)의 결과를 은닉 원음 (REF) 및 앵커음 (LP35)의 결과와 함께 보이고 있다. 결과로부터, 5.1 채널 재생에 대해 비트율에 의한 음질 차이는 존재하지만 고음질 버전과 저전력 버전의 음질 차이는 거의 없음을 알 수 있다. 이는 그림 12에서 바이노럴 모드에 대해 고음질 버전과 저전력 버전의 음질 차이가 일부 존재하는 것과는 차이가 있다.

IV. 결론

본 논문에서는 1부의 MPEG Surround 기본 구조 소개에 이어서, MPEG Surround의 다양한 모드 및 톨들에 대해서 소개하고 음질 평가의 결과를 이용해 성능을 보였다. 이들 중 특히 바이노럴 모드의 경우 휴대용 오디오 기기에 사용될 수 있으므로, DMB (Digital Multimedia Broadcasting)이나 DAB (Digital Audio Broadcasting) 등의 방송 시스템에서 다채널 오디오 서비스가 상용화될 경우 활용도가 높을 것으로 예상된다. 또한, Enhanced Matrix 모드는 부가 정보 없이 기존의 스테레오 채널을 5.1 채널로 확장시킬 수 있다는 점에서 홈씨어터나 차량용 음향 시스템에서 유용하게 사용될 수 있을 것으로 판단된다.

감사의 글

본 논문은 지식경제부 IT 산업원천 사업의 지원을 받았습니다. MPEG Surround 표준화 과정에 같이 참여하였던 LG 전자 DM 연구소의 김동수, 임재현, 윤성용, 이현국 연구원과 DTV 연구소의 오현오, 정양원 연구원께 감사의 말씀을 드립니다.

참고 문헌

1. ISO/IEC JTC 1/SC 29/WG 11, *Text of Working Draft for Spatial Audio Coding*, N7136, Apr. 2005.
2. ISO/IEC 23003-1, *Information technology - MPEG audio technologies - Part 1: MPEG Surround*, 2007.
3. ISO/IEC JTC 1/SC 29/WG 11, *Performance of MPEG Surround Technology*, N7950, Jan. 2006.
4. ISO/IEC JTC 1/SC 29/WG 11, *Report on MPEG Surround Verification Test*, N8851, Jan. 2007.
5. J. Jakka et al., *New use cases for spatial audio coding*, ISO/IEC JTC 1/SC 29/WG 11, M12913, Jan. 2006.
6. K. S. Chong et al., *Core Experiment on adding 3D stereo support to MPEG Surround*, M12845, Jan. 2006.
7. ISO/IEC JTC 1/SC 29/WG 11, *Work plan for core experiment on binaural stereo*, N7949, Jan. 2006.
8. ISO/IEC JTC 1/SC 29/WG 11, *MPEG Surround Binaural/3D-Stereo Test Report*, N8178, Apr. 2006.
9. J. Herre, K. Kjørting, J. Breebaart, C. Faller, S. Disch, H. Purnhagen, J. Koppens, J. Hilpert, J. Roden, W. Oomen, K. Linzmeier, and K. S. Chong, "MPEG Surround - the ISO/MPEG standard for efficient and compatible multichannel audio coding," *J. Audio Eng. Soc.*, vol. 56, no. 11, pp. 932-955, 2008.
10. ISO/IEC 14496-3: 2005, *Information technology - Coding of audio-visual objects - Part 3: Audio*, 2005.
11. ITU-R BS 1534-1, *Method for the Subjective Assessment of Intermediate Quality Levels of Coding Systems*, 2003.

저자 약력

• 방희석 (Hee-Suk Pang)



1994년: 서울대학교 전자공학과 (학사)
 1996년: 서울대학교 전자공학과 (석사)
 2001년: 서울대학교 전기컴퓨터공학부 (박사)
 2001년~2008년: LG전자 DM연구소 책임연구원
 2008년~현재: 세종대학교 전자공학과 조교수
 *주관심 분야: 오디오/음성 코딩, 오디오 신호처리, 음악 신호 분석