

주 제

# MPEG Surround 멀티채널 오디오 압축기술의 개요와 적용방안

한국전자통신연구원 서정일, 백승권, 장인선, 장대영

차례

I. 서론

II. Spatial Audio Coding

III. MPEG Surround

IV. MPEG Surround 성능

V. MPEG Surround 적용방안

VI. 결론

## 요약

공간 오디오 부호화(spatial audio coding)는 기존의 스테레오 오디오 시스템과 호환성을 유지하면서 멀티채널 오디오 신호를 효과적으로 압축하기 위한 기술이다. MPEG에서는 AAC 및 MP3와 같은 기존의 스테레오 오디오 압축 표준과 호환성을 제공하면서 멀티채널 오디오 신호의 음장감을 유지하고 저 비트율로 압축하기 위한 기술을 MPEG Surround란 이름으로 표준화가 진행 중이다.

본 논문에서는 MPEG Surround 기술의 개요와 표준화 진행상황을 소개하고, 향후 적용방안에 대해서 살펴보고자 한다.

## I. 서론

멀티채널 환경에서의 오디오 부호화 기술에 대한 연구는 1990년대 중반부터 MPEG을 중심으로 이루어졌다. DVD 등의 멀티채널 소스의 폭발적인 증가와, 사용자들의 멀티채널에 대한 욕구증가와, 통신환경을 통한 서비스의 필요성 증가가 멀티채널 오디오 기술에 대한 연구를 더욱 부추기게 하였다. 이러한 이유로 MPEG-2 BC(backward compatible), MPEG-2/4 AAC(advanced audio coding) 등과 같은 멀티채널 오디오 부호화기술이 표준화되어 상업적으로 큰 실효를 거두었으며, MPEG-4 및 MPEG-7에 대한 표준화 작업이 Microsoft 등의 대기업과 Fraunhofer 등의 연구소를 중심으로 이루어지고 있다. 이들을 축으로 한 연구 그룹에서 발표한 연구 결과들이 세계적인 표준으로 자리 잡음으로써,

학계뿐만 아니라 오디오 산업에 매우 큰 영향을 미치고 있는 것이 현실이다. 아울러 최근에는 BCC (binaural cue coding)과 같은 혁신적인 공간오디오 부호화(spatial audio coding) 기법에 대한 연구가 활발히 진행되고 있는데[1], 이를 통해서 보다 적은 정보량을 이용하여 현실감 있는 멀티채널 오디오 데이터의 전달을 목적으로 하고 있다.

MPEG에서는 2002년도부터 MPEG Surround란 이름으로 공간 오디오 부호화 기술에 대한 표준화가 진행 중이다[2]. MPEG Surround는 M개의 오디오 신호를 N개의(M>N) 오디오 신호와 사람이 음원의 위치를 판단하는 공간단서 파라미터들로 구성되는 부가정보(side information)를 이용하여 나타내는 파라메트릭 멀티채널 오디오 압축 기술이다. MPEG Surround 부호화기는 멀티채널 오디오 신호를 모노 또는 스테레오 채널로 다운믹스(downmix)한 후 기존의 MPEG-4 오디오 도구(MPEG-4 AAC, MPEG-4 HE-AAC 등)로 압축하고, 멀티채널 오디오 신호로부터 공간단서(spatial cues)를 추출하여 다운믹스 오디오 신호와 다중화한다. MPEG Surround 복호화기는 역다중화기를 이용하여 다운믹스 오디오 신호와 공간단서를 분리하고, 다운믹스 오디오 신호에 공간단서를 적용하여 멀티채널 오디오 신호를 합성한다. 따라서 MPEG Surround는 다음과 같은 특징을 갖는다.

첫째, MPEG Surround는 기존의 모노 혹은 스테레오 오디오 부호화 시스템과의 역호환성(backward compatibility)을 제공한다. MPEG Surround 부호화기에서 생성된 다운믹스 신호는 독립적으로 서비스가 가능하므로 기존의 모노 또는 스테레오 오디오 시스템과의 호환이 가능하다.

둘째, MPEG Surround는 멀티채널 오디오 신호를 위한 효과적인 파라메트릭 코딩 방식을 지원한다. 멀티채널 오디오 신호를 표현하기 위한 부가정보는

매우 작은 정보량(채널당 4kbps 이내)이므로 모노 또는 스테레오 오디오 서비스만 제공할 수 있는 대역폭에서도 고품질의 멀티채널 오디오 서비스를 제공할 수 있다.

셋째, MPEG Surround는 핸드폰과 같이 멀티채널 재생환경을 구성하기 어려운 환경에서도 가상적인 멀티채널 오디오 신호를 제공할 수 있는 바이노럴 복호화 기능을 제공한다. 바이노럴 복호화기는 멀티채널로 복원하지 않고 공간단서와 머리전달함수(head related transfer function, HRTF)를 이용하여 가상적인 멀티채널 신호를 적은 연산량으로 효과적으로 생성할 수 있다.

본 논문의 구성은 다음과 같다. II장과 III장에서는 공간 오디오 부호화 기술과 MPEG Surround의 기본 개념과 표준화 동향을 소개하며, IV장에서는 현재까지 표준화가 완료된 MPEG Surround RM0 (reference model 0) 코덱의 성능을 살펴본다. V장에서는 앞의 내용들을 기반으로 MPEG Surround의 적용 방안을 제시함으로써 향후 멀티채널 오디오 서비스의 발전 방향을 분석한 후, VI장에서 결론을 맺는다.

## II. Spatial Audio Coding

3차원 오디오 신호와 같이 다채널 및 다객체 환경에서는 전송에 필요한 대역폭이 채널의 개수나 오디오 객체의 개수에 비례하여 증가한다. 그러나 다채널 및 다객체 오디오 신호를 모노 또는 스테레오 채널로 다운믹스하여 압축하고 객체들과 채널들에 대한 정보는 별도의 부가정보로 표현하여 전송하면 압축효율을 극대화할 수 있다. Faller 등에 의해 제안된 BCC(binaural cue coding) [1] [3]란 이름으로 제안된 공간 오디오 부호화(spatial audio coding; SAC)

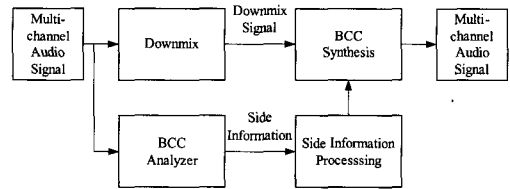
는 (그림 1)과 같이 멀티채널 오디오 신호를 최소한의 채널로 다운믹스한 오디오 신호와 멀티채널 오디오 신호의 음향 공간상의 단서(spatial cue)로 이루어진 부가정보로 분할하여 압축하는 파라미터 멀티채널 오디오 부호화 방식이다[3].

자유공간(free field) 상에서 단일 음원에서 발생된 음파는 (그림 2)와 같이 청취자의 두 귀에 크기차와 도달 시간차를 가지며 도착하게 된다. 청취자는 이를 통하여 음원의 위치를 파악하므로 이를 양이효과(binaural effect)라고 부른다. 이때 두 귀 사이의 음향신호의 크기 차이를 ILD (interaural level difference)라 부르고, 두 귀에 도착하는 신호 도달 시간의 차이를 ITD(interaural time difference)라 부르며 이들을 양이단서(binaural cue)라고 통칭한다. 실험을 통하여 1.5kHz 이상의 높은 주파수 대역에 대해서는 ILD가, 1.5kHz 이하의 낮은 주파수 대역에 대해서는 ITD가 음원의 위치를 구별하는 데 중요한 영향을 준다고 알려져 있다[4].

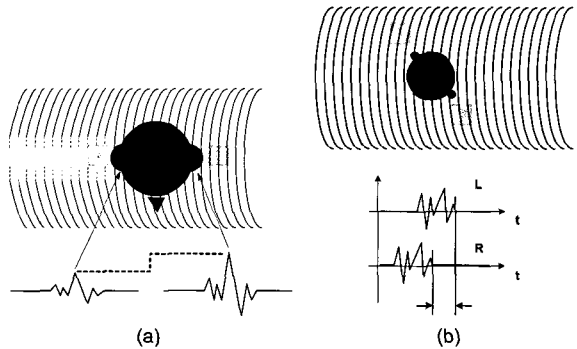
하나의 가상음원으로 발생하는 모든 음향 신호 성분들은 청취자의 귀에서 일정한 ILD와 ITD 값을 가지므로, 각 음향 신호 성분들은 자신만의 영역을 차지하며 이들의 크기는 전력(power)으로 표현할 수 있다. 물론, 특정 음원의 시간-주파수 평면에서 차지하는 영역이 다른 음원과 겹치는 경우도 발생한다. 그러나 사람의 귀는 특정 시간-주파수 대역에서 두 개 이상의 음향 신호 성분이 존재할 때 전력이 큰 신호로 인지하는 차폐 특성(masking effect)이 있으므로 겹치는 영역에 대해서는 전력이 큰 신호의 영역으로 가정할 수 있다.

만약 (그림 3)과 같이 세 개의 음원들이 시간-주파수 평면상에서 겹치지 않는 영역에서 존재한다고 가정하면, 이들이 생성하는 음향학적 이미지는 모든 음원들의 합신호를 이용하여 재현할 수 있다. 왜냐하면, 세 개의 음원들에 대한 시간-주파수 평면상의 임

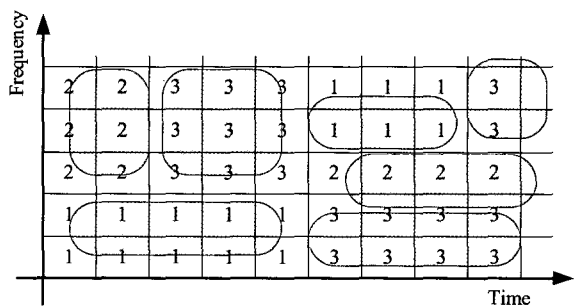
의 영역에서 합신호에 양이단서 파라미터를 적절히 적용하면 원래의 음향학적 이미지를 재현할 수 있으며, 시간-주파수 평면이 각 음원들에 대한 영역으로 겹치지 않게 구분할 수 있다면 이론적으로 완벽하게 원음의 음향학적 이미지를 복원할 수 있기 때문이다.



(그림 1) Binaural cue coding 시스템 개념도



(그림 2) 두 귀 사이의 신호 크기차(a)와 신호 도달 시간차(b)



(그림 3) 모노 합신호에 포함되어 있는 음원들과 이들의 영역들

### III. MPEG Surround

MPEG Surround는 최소의 채널(모노 혹은 스테레오)로 합쳐진 다운믹스 오디오 신호와 멀티채널 오디오 신호에 대한 인간의 지각 특성을 나타내는 공간 파라미터(spatial parameter)들로 구성된 부가정보를 이용하여 멀티채널 오디오 신호를 부호화하는 기술이다[5].

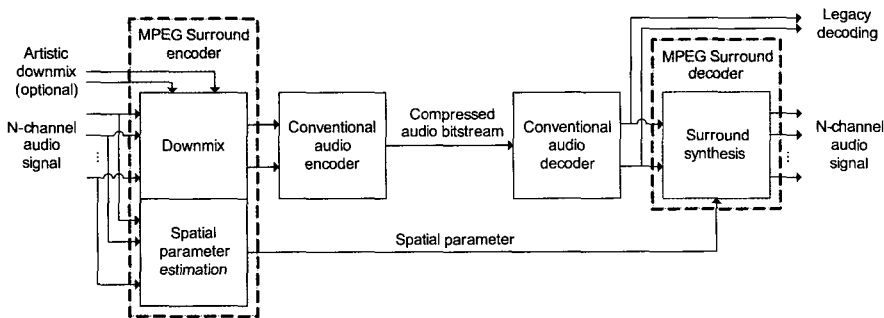
(그림 4)는 MPEG Surround의 기본 개념도이다. 그림에서 보는 바와 같이, MPEG Surround 부호화기는 N개의 멀티채널 오디오 신호를 입력받아 부가정보로써 양이효과(binaural effect)에 바탕을 둔 두 귀 사이의 소리 크기 차이와 채널 사이의 상관도 등으로 표현되는 공간 파라미터를 추출한다. 추출된 공간 파라미터는 매우 작은 정보량(채널당 4kbps 이내)이므로 모노 또는 스테레오 오디오 서비스만 제공할 수 있는 대역폭에서도 고품질의 멀티채널 오디오 서비스를 가능케 한다. MPEG Surround 부호화기는 입력받은 멀티채널 입력신호로부터 다운믹스 신호를 생성하며, 이는 음향엔지니어 등이 제작한 다운믹스(artistic/hand-mixed downmix) 신호로 대체할 수 있다. 생성된 다운믹스 신호는 기존의 고성능 오디오 압축 기술인 MPEG-4 AAC, MPEG-4 HE-AAC 및 MPEG-4 BSAC 등으로 부호화되어 공간

파라미터와 함께 전송된다.

MPEG Surround 복호화기는 전송된 오디오 비트 스트림과 공간 파라미터를 합성하여 N개의 멀티채널 오디오 신호를 복원한다. 멀티채널 오디오의 복호화가 불가능한 경우에는 기존의 오디오 복호화기를 이용하여 다운믹스 신호만을 복호화 할 수 있으므로 독립적인 서비스가 가능하다. 따라서 기존의 모노 또는 스테레오 오디오 부호화 시스템과 호환성을 제공할 수 있다.

#### 1. MPEG Surround 부호화기

(그림 5)는 MPEG Surround 부호화기의 블록도이다. 부호화기에 입력된 N개의 멀티채널 신호는 해석 필터뱅크(analysis filter bank)에 의해 주파수 밴드로 분해된다. 주파수 영역의 서브 밴드로 분할하는 방법으로써 MPEG Surround에서는 낮은 복잡도로 이를 수행하기 위하여 QMF (quadrature mirror filter)를 사용한다. QMF는 SBR (Spectral Band Replication)과 같은 틀과의 호환성도 제공함으로써 보다 효율적인 부호화를 유도 할 수 있다. QMF를 거친 각 서브밴드는 Nyquist 필터뱅크를 이용하여 균등분할구조인 서브밴드로 나누고 이를 사람의 청각 시스템의 주파수 분해능과 유사하게 재구성하며, 이



(그림 4) MPEG Surround의 개념도

와 같은 QMF와 Nyquist 필터뱅크 전체 구조를 통칭하여 Hybrid QMF라 부른다.

그 후 서브밴드 신호들로부터 공간 지각과 관련 있는 공간 특성들을 해석하며, 채널 간 크기 차이(CLD; Channel Level Difference), 채널 간 유사도(ICC; InterChannel Correlation), 채널 예측 계수(CPC; Channel Prediction Coefficients) 및 예측 오차(Prediction errors or residual signal) 4가지의 공간 파라미터들을 부호화 동작 모드에 따라 선택적으로 추출한다.

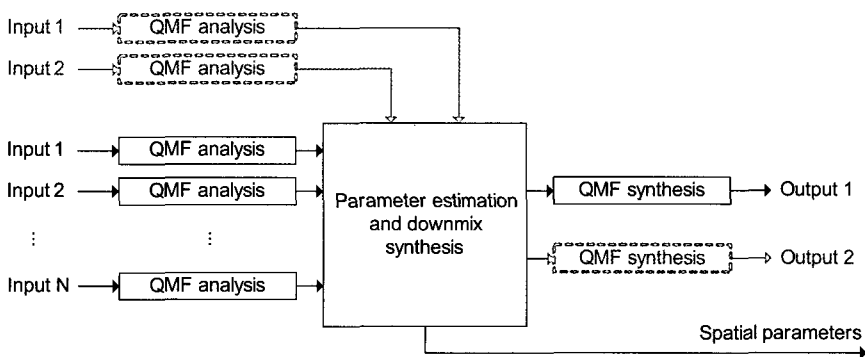
또한, 입력신호들은 다운믹싱 처리 후 QMF 합성뱅크를 통과하여 시간영역의 다운믹스 신호로 변환된다. 다운믹스 신호는 부호화기에서 자동적으로 생성되며, 이는 모노/스테레오 재생 혹은 행렬 서라운드 복호화기(예로, Dolby Prologic 등)에 따른 재생을 위해 최적화된 형태를 띠게 된다. 또한, 무선 전송을 위한 후처리 결과로 혹은 스튜디오 엔지니어에 의해서 생성된(artistic downmix) 다운믹스 신호가 MPEG Surround 부호화기의 다운믹스 신호로 제공되는 경우, 부호화기에서는 제공된 다운믹스 신호에 기반하여 공간 파라미터를 조정한 후 전송함으로써 복호화기에서의 멀티채널 복원을 최적화한다.

## 2. MPEG Surround 복호화기

(그림 6)은 MPEG Surround 복호화기의 블록도이다. 시간영역에서의 다운믹스 신호는 QMF 해석 필터뱅크 통과되어 시간-주파수 영역으로 변환된다. 그 후, 공간 합성 모듈에서 전송된 공간 파라미터와 결합되며, QMF 합성 필터뱅크를 통과하여 시간영역의 멀티채널 신호로 변환된다.

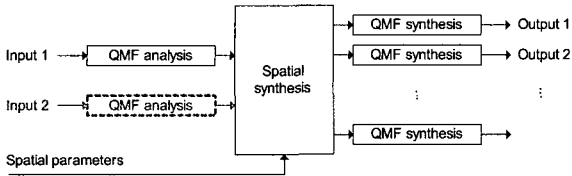
공간 합성 단계는 입력된 다운믹스 신호와 전송된 공간 파라미터를 결합하여 2개 채널 신호를 6개 채널 신호로 업믹스(upmix) 함으로써 멀티채널 신호를 생성하는 역할을 하며 행렬화(matrixing)와 디코리레이션(decorrelation)으로 구성된다. (그림 7)은 스테레오 다운믹스 신호에서 5.1 채널 신호 생성을 위한 공간 합성 모듈의 구조도이다. 시간-주파수 영역의 입력신호들은 전믹싱 행렬(pre-mixing matrix) M1에 의해 처리된 후, 후믹싱 행렬(post-mixing matrix)에 바로 입력되거나 디코리레이션 회로 D1~D3를 통과한 후 입력되어 5.1채널 형태의 출력 신호로 변환된다.

참고로, 잔차 신호(residual signal) res는 선택적으로 제공되며, MPEG Surround는 잔차 신호를 이용하여 음질(transparent quality)에 대한 확장성

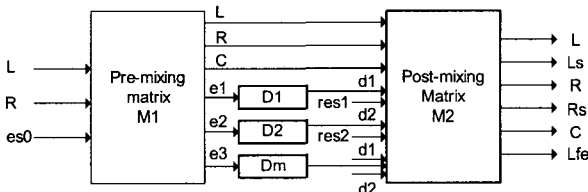


(그림 5) MPEG Surround 부호화기

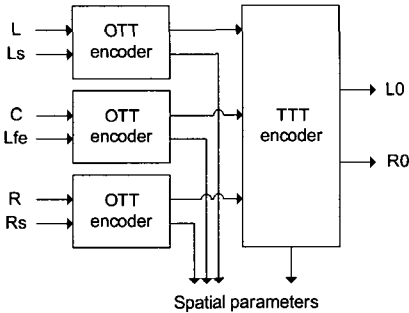
(scalability)을 제공할 수 있다.



(그림 6) MPEG Surround 복호화기



(그림 7) 공간 합성의 구조도



(그림 8) 공간 파라미터 추출 예

### 3. Spatial Parameter

MPEG Surround에서는 OTT (One-to-Two) 모듈과 TTT (Two-to-Three) 모듈을 조합한 임의의 매핑을 이용하여 다운믹스 신호 및 멀티채널 신호를 생성한다. 한 예로, (그림 8)은 TTT 모듈 한 개와 OTT 모듈 3개를 조합하여 5.1채널 신호로부터 스테레오 다운믹스 신호와 공간 파라미터를 얻는 MPEG Surround 부호화기이다.

Parametric stereo[6]와 BCC에서 발전된 개념

인 OTT 모듈은, 위 그림에서 보듯이 스테레오 입력 신호로부터 모노 다운믹스 신호를 생성하며, 또한 아래의 공간 파라미터들을 추출한다.

- 채널 간 크기 차이(CLD): CLD 파라미터는 시간-주파수에 따른 두 채널간의 크기 차이를 나타내며, 로그스케일로 비선형 양자화 한다.
- 채널 간 유사도(ICC): ICC 파라미터는 시간-주파수에 따른 두 채널간의 유사도이며, 비선형 스케일로 양자화 한다.
- 잔차 신호(residual signal): 원신호와 공간 파라미터들을 이용하여 형성된 파형간의 오차 신호로써, MPEG Surround 복호화기에서는 이에 대한 정보를 이용하여 원 신호와 가깝게 복원한다.

TTT 모듈은 부호화 단계에서, (그림 8)에서와 같이, 입력신호로부터 스테레오 다운믹스 신호를 생성하고 아래와 같은 공간 파라미터들을 추출한다.

- 채널 예측 계수(CPC): CPC는 시간입력 채널 혹은 입력 채널간의 결합으로부터 출력 채널 혹은 출력 채널간의 결합에 대한 예측 계수를 나타내며, 비선형 스케일로 양자화 한다.
- 채널 간 유사도(ICC): ICC 파라미터는 시간-주파수에 따른 두 입력 채널간의 유사도이며, OTT 모듈의 ICC 파라미터와 유사하다.
- 예측 오차(prediction error): 원신호와 채널 예측 계수를 이용하여 모델링된 파형간의 오차 신호로써, MPEG Surround 복호화기에서는 이에 대한 정보를 이용하여 원신호와 가깝게 복원할 수 있다.

또한, 부호화 단계에서 TTT 모듈은 스테레오 다

운믹스 신호로부터 세 번째 채널 즉, 중앙(center) 채널을 모델링하여 생성한다. 이는 다운믹스 신호  $L0, R0$ 를 원 신호  $L, R, C$ 의 선형 조합으로 표현된다는 가정을 바탕으로 하며, 식(1)과 같은 선형식으로 모델링할 수 있다.

$$\begin{bmatrix} L0 \\ R0 \end{bmatrix} = H_{Trr} \begin{bmatrix} L \\ C \\ R \end{bmatrix} \quad (1)$$

따라서 부호화 단계에서의 다운믹스 행렬(HTTT)이 공간 파라미터(CPC)로써 전송되면, 복호화기에서는 이를 이용하여 다운믹스 신호로부터 세 개 채널의 신호를 생성할 수 있다.

#### 4. Decorrelation

(그림 7)에서 보듯이, 공간 합성 단계는 행렬화와 디코리레이션으로 구성된다. 디코리레이션 모듈은 아래와 같은 입출력 신호의 관계에 대한 요구사항을 만족시키면서, 전송된 ICC 파라미터에 따른 유사도를 갖는 출력 신호를 합성하는 기능을 담당한다.

- 입출력 신호의 유사도는 0에 가까워야 한다.
- 입출력 신호의 주파수 영역 및 시간 영역에서의 포락선은 거의 유사해야 한다.
- 디코리레이터 출력신호들은 입출력 신호의 제한을 만족하며, 서로 유사하지 않아야 한다.

격자 전역 필터(lattice allpass filter)와 그 밖의 툴들을 사용하여 위의 요구사항을 만족하는 디코리레이터를 구현하였으며, 일반적인 구조는 (그림 9)와 같다. 입력 신호가 동적으로 변하는 경우에는 전역 필터만으로는 두 번째 요구사항을 만족시킬 수 없기 때문에, 주파수 영역에서의 포락선은 에너지 비교(energy comparison) 단계에 의해, 시간 영역에서

의 포락선은 TES (Temporal Envelope Shaping) 혹은 TP (Temporal Processing) 등에 의해 조절된다.

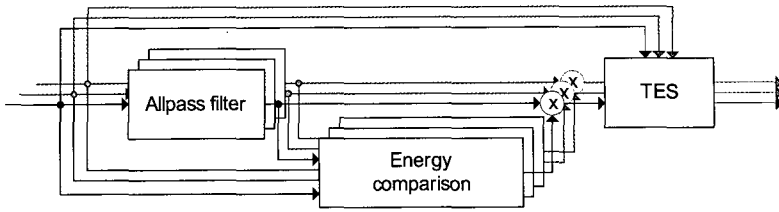
### 5. MPEG Surround Binaural Decoder

핸드폰 등과 같이 스테레오 오디오 출력만을 가진 단말에서 멀티채널 오디오 신호를 머리전달함수(head related transfer function: HRTF)를 이용하여 가상적인 멀티채널 신호로 변환할 수 있다[7]. MPEG Surround로 압축된 멀티채널 오디오 스트림에서 부가정보 신호를 이용하면 멀티채널로 복원하지 않고 직접 가상 멀티채널 신호로 복원할 수 있다[8]. 즉, (그림 10)의 상단부에 표현된 바이노럴 디코더는 스테레오 다운믹스 신호를 공간 파라미터와 머리전달함수를 이용하여 가상 멀티채널 신호(3D stereo)로 구성함으로써 적은 연산량으로 MPEG Surround 디코딩 과정과 가상 멀티채널 처리 과정을 동시에 처리하는 것이다.

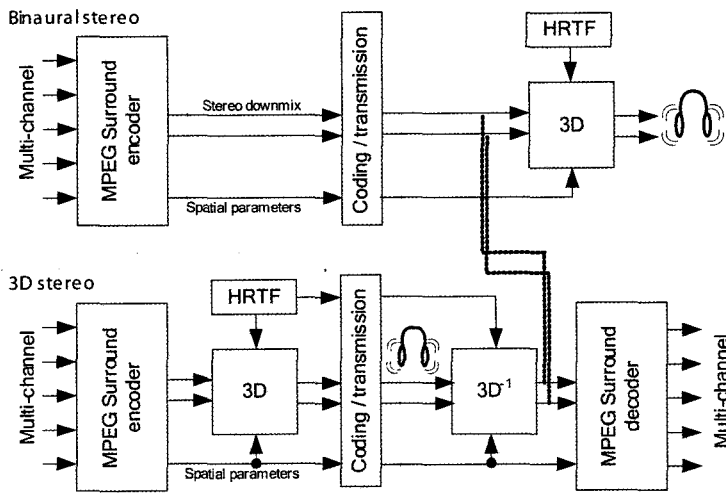
특히, MPEG Surround 바이노럴 디코더는 5채널에 대한 머리전달함수를 아래 수식과 같이 2X2 매트릭스(H)로 변환함으로써 연산량을 최소화하면서도 멀티채널 음장감을 제공한다. 아래 수식에서  $L0, R0$ 는 입력된 스테레오 다운믹스 신호를,  $Lb, Rb$ 는 출력 바이노럴 스테레오 신호를 나타낸다.

$$\begin{bmatrix} L_b[k, m] \\ R_b[k, m] \end{bmatrix} = \sum_{i=0}^{L_0-1} \begin{bmatrix} h_{11}[i] & h_{12}[i] \\ h_{21}[i] & h_{22}[i] \end{bmatrix}_k \begin{bmatrix} L_0[k, m-i] \\ R_0[k, m-i] \end{bmatrix} \quad (2)$$

(그림 10)의 하단부에 표현된 3D 스테레오는 가상 멀티채널 처리 과정을 부호화기에서 처리하도록 하여 디코더에서 아무런 처리과정 없이 가상 멀티채널 신호를 재생할 수 있도록 하는 기법도 적용할 수



(그림 9) QMF 영역에서의 디코리레이터 구조도



(그림 10) MPEG Surround 바이노럴 디코더와 3D 스테레오 코덱 구조

있다. 물론, 멀티채널 복원을 위해서는 부호화 과정중에 적용된 가상 멀티채널 효과를 역변환 과정을 통해 상쇄시킴으로써 원음에 가까운 멀티채널 신호를 합성한다.

#### IV. MPEG Surround 성능

MPEG 오디오 서브 그룹에서는 MPEG Surround의 성능 평가를 위하여 ITU-R에서 권고하는 MUSHRA (Multi-Stimulus test with Hidden Reference and Anchor) [9] 방법을 선정

하였다.[10] 이는 중간 정도의 음질을 가지는 오디오 시스템의 성능을 평가하기 위하여 고안된 평가 방법으로 숨겨진 기준 신호(hidden reference)와 저역 필터로 음질을 열화 시킨 앵커(anchor) 신호들을 포함한 실험 시스템을 구성하여 주관적으로 음질을 평가하는 방법이다.

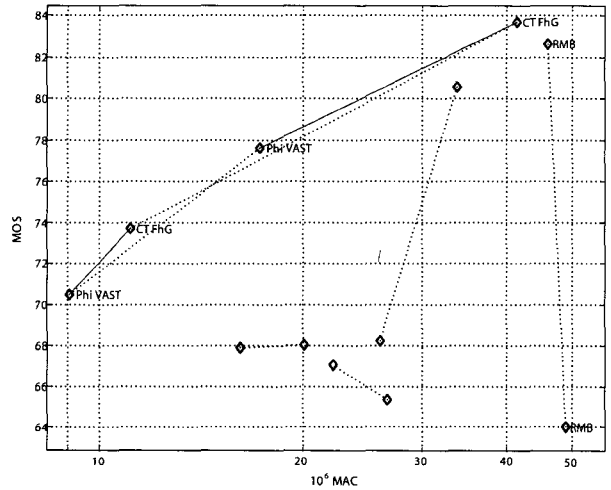
(그림 11)은 MPEG Surround의 전반적인 음질 평가 결과이다[11]. 그림에서 보듯이, 행렬화 기반의 멀티채널 재생방법인 Dolby Prologic II와 Logic7과의 비교 및 다양한 비트율의 AAC 코덱과의 비교를 통하여 MPEG Surround의 음질성능을 살펴볼 수 있다. 여기서 MPEG Surround의 다운믹스 신호에



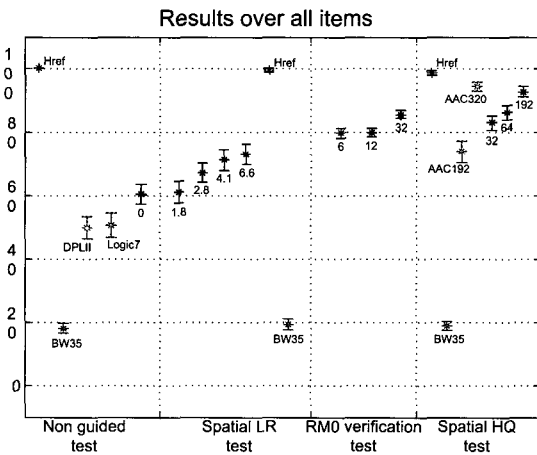
대한 비트율은 128 kbps이며 그래프 하단에 있는 숫자는 부가정보의 비트율을 나타낸다. 기본적으로 MPEG Surround는 기존 행렬화 기반의 코덱보다 고음질을 제공할 수 있음을 확인할 수 있다. 또한 MPEG Surround의 부가 정보의 비트율에 따라 저비트율(LR; Low-rate) 모드, 기본 모드(RM0 verification), 고효율(HQ; High-quality) 모드 등으로 나뉘어 질 수 있으며, 평가 결과, 부가정보의 비트율이 증가함에 따라 점진적인 음질향상이 관측되었다. 즉, MPEG Surround는 기존 저비트율 코덱의 성능으로부터 끊김 없는(seamless) 점진적 음질향상을 기대할 수 있으며 비트율이 증가할수록 원음과 가까운 손실 없는(lossless) 음질을 제공할 수 있으며 이를 기반으로 확장성(scalability)을 제공할 수 있다.

(그림 12)는 MPEG Surround 바이노럴 디코더 RM0(reference model 0)로 선정된 CT/FhG 모델과 Phi/VAST 모델의 음질(quality)로 표현된 음질과 복잡도(complexity)를 나타낸 그림이다[12]. 그림에서와 같이 CT/FhG 모델이 가장 우수한 음질을

제공하는 반면 Phi/VAST 모델은 낮은 복잡도를 제공하면서도 적절한 음질을 제공하고 있다. 그림에서 같은 모델이면서 우수한 음질을 제공하는 것은 스테레오 다운믹스 신호일 경우이며, RMB는 MPEG Surround RM0를 이용했을 경우를 나타낸다.



(그림 12) MPEG Surround 바이노럴 디코더 RM0로 선정된 CT/FhG와 Phi/VAST 모델의 성능(quality)과 복잡도(complexity) 비교



(그림 11) 부가정보 비트율에 따른 MPEG Surround 코덱의 성능

## V. MPEG Surround 적용방안

오늘날 대부분의 오디오 시스템이 스테레오를 기반으로 하고 있으므로, 멀티채널 오디오 서비스를 위해서는 기본 오디오 전송 채널의 수용 능력(capability)과 기존 오디오 재생 시스템과의 호환성이 제공되어야 한다.

MPEG Surround 기술은 스테레오 (혹은 모노) 신호와 적은 양의 부가 정보로써 멀티채널 신호를 압축하여 전송함으로써, 80% 이상의 압축 효율을 얻을 수 있기 때문에 기존에 압축된 스테레오 신호의 전송 채널을 이용하여 전송이 가능하며, 또한 전송된 스테

레오 (혹은 모노) 다운믹스 신호는 기존의 재생 시스템에서 독립적인 서비스가 가능하다. 따라서 MPEG Surround는 멀티채널 신호에 대한 압축 효율과 기존 오디오 시스템과의 역 호환성이 요구되는 어떠한 서비스에도 적용이 가능하다.

MPEG Surround 기술에 대한 잠재력이 큰 응용 분야로는 스테레오 방송을 기반으로 하는 DAB (Digital Audio Broadcasting) 및 DMB(Digital Multimedia Broadcasting)를 들 수 있다. 기존 대다수의 방송 시스템은 전송채널 수용 능력(channel capacity)의 한계로 인해 멀티채널 오디오 방송을 지원하지 못하지만, 앞서 설명한 MPEG Surround의 압축효율 및 역호환성(backward compatibility)을 활용하여 제한된 대역폭 상에서 기존의 DAB 및 DMB 수신기뿐만 아니라 멀티채널 오디오 시스템에서도 재생 가능한 오디오 스트림을 전송함으로써 향후 멀티채널 DAB/DMB 서비스가 가능할 것으로 전망된다. 또한, 멀티채널 DAB/DMB 서비스는 멀티채널 재생환경 특성상 차량에 가장 빠르게 보편화 될 것으로 보인다. 즉, 차량 내부에는 5개 이상의 라우드 스피커와 우퍼 등이 장착되어 있으므로 멀티채널 재생이 가능하며 또한, 가정이나 사무실과 같은 환경에서는 청취자가 라우드 스피커의 위치에 따른 최적 청취영역(sweet spot)에 위치해야만 멀티채널 오디오에 대한 적절한 음장을 제공받을 수 있음에 반해, 차량 내부에서는 청취자의 위치가 고정되어 있으므로 상대적으로 적절한 음장의 제공하는 멀티채널 오디오 청취가 용이하기 때문에, 멀티채널 DAB/DMB 서비스의 파급효과가 가장 클 것으로 예상된다. 또한, 핸드폰과 같은 스테레오 출력장치만을 구비한 소형 단말에서는 MPEG Surround 바이노럴 디코더나 3D 스테레오 기능을 이용하여 가상 멀티채널 서비스를 제공해줌으로써 시청자에게 보다 실감있는 음향을 제공할 수 있다.

그 밖의 MPEG Surround에 대한 응용 분야로는 음악 다운로드 서비스, 음악 스트리밍 서비스, 인터넷 라디오, 멀티채널 원격회의 서비스 및 게임을 위한 실감 오디오 서비스 등을 들 수 있다.

## VI. 결 론

MPEG에서 표준화가 진행중인 MPEG Surround는 기존의 스테레오 오디오 코덱과 호환성을 제공하면서도 멀티채널 오디오 신호를 기존의 스테레오 오디오 대역폭으로 압축할 수 있는 파라메트릭 오디오 압축 도구이다. MPEG Surround가 제공하는 호환성과 고압축율은 DAB나 DMB와 같은 스테레오 오디오 방송 시스템을 멀티채널로의 손쉬운 확장이 가능할 뿐만 아니라, MPEG Surround의 바이노럴 디코더 모드를 적용하면 스테레오 오디오 출력 장치에서도 가상적인 멀티채널 음향을 제공할 수 있다. DVD와 지상파 DTV를 필두로 보급화가 급속도로 진행되고 있는 가정용 멀티채널 오디오 기기와 이미 멀티채널 재생환경이 구비되어 있는 차량에서 MPEG Surround를 이용한 멀티채널 오디오 방송 및 콘텐츠 서비스가 머지않아 적용될 것으로 예상된다.

### [참 고 문 헌]

- [1] C. Faller and F. Baumgarte, "Binaural Cue Coding: A Novel and Efficient Representation of Spatial Audio," in Proc. ICASSP 2002, Orlando, FL, May 2002.
- [2] ISO/IEC JTC1/SC29/WG11 (MPEG),

- Document N6455, "Call for Proposals on Spatial Audio Coding," Munich, Mar. 2004.F.
- [3] Baumgarte and C. Faller, "Estimation of Auditory Spatial Cues for Binaural Cue Coding (BCC)," in Proc. ICASSP 2002, Orlando, FL, May 2002.
- [4] F. L. Wightman and D. J. Kistler, "The dominant role of low-frequency interaural time differences in sound localization," Journal Acoustical Society of America, vol. 91, no. 3, pp. 1648-1661, Mar. 1992.
- [5] ISO/IEC JTC1/SC29/WG11 (MPEG), Document N7947, "ISO/IEC 23003-1: 2006/FCD, MPEG Surround," Bangkok, Jan. 2006.
- [6] H. Purnhagen, "Low complexity parametric stereo coding in MPEG-4," in Proc. 7th DAFX04, Naples, Italy, Oct. 2004.
- [7] 강성훈, 강경욱, 임채음향, 기전연구소, 1997.
- [8] ISO/IEC JTC1/SC29/WG11 (MPEG), Document N8177, "Study on Text of ISO/IEC 23003-1, MPEG Surround," Montreux, April. 2006.
- [9] ITU-R Recommendation BS. 1534-1, "Method for the subjective assessment of intermediate sound quality (MUSHRA)," International Telecommunications Union, Geneva, Switzerland, 2001.
- [10] ISO/IEC JTC1/SC29/WG11 (MPEG), Document N6691, "Procedures for the Evaluation of Spatial Audio Coding Systems," Redmond, July 2004.
- [11] ISO/IEC JTC1/SC29/WG11 (MPEG), Document N7701, "Performance of MPEG Surround Technology," Nice, 2005.
- [12] ISO/IEC JTC1/SC29/WG11 (MPEG), Document N8178, "MPEG Surround Binaural/3D-Stereo test Report," Montreux, April, 2006.



**서정일**

1994년 경북대학교 전자공학과 (공학사)  
1996년 경북대학교 대학원 전자공학과 (공학석사)  
2005년 경북대학교 대학원 전자공학과 (공학박사)  
1998년 ~ 1999년 LG 반도체 근무  
1999년 ~ 2000년 현대 전자 근무  
2000년 ~ 현재 한국전자통신연구원 전파방송연구

단 방송미디어연구그룹 선임연구원

관심분야 : 음향신호처리, 음성신호처리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등



**백승권**

1999년 한국항공대학교 항공전자공학과 (공학사)  
2001년 한국정보통신대학교 대학원 전자공학과 (공학석사)  
2005년 한국정보통신대학교 대학원 전자공학과 (공학박사)  
2005년 ~ 현재 한국전자통신연구원 전파방송연구

단 방송미디어연구그룹 선임연구원

관심분야 : 오디오 신호처리, 음원분리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등



**장인선**

2001년 충북대학교 전기전자공학부 (공학사)  
2004년 포항공과대학교 대학원 컴퓨터공학과 (공학석사)  
2004년 ~ 현재 한국전자통신연구원 전파방송연구  
단 방송미디어연구그룹 연구원

관심분야 : 오디오 신호처리, 음원분리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등



**장대영**

1991년 부경대학교 전자공학과 (공학사)  
2000년 배재대학교 대학원 컴퓨터공학과 (공학석사)  
1991년 ~ 현재 한국전자통신연구원 전파방송연구  
단 방송미디어연구그룹 책임연구원

관심분야 : 오디오 신호처리, 음원분리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등