



## A Study on Ultra Multi-channel Audio Coding for Realistic Audio Services

Kwangki Kim\*

*School of IT Convergence, Korea Nazarene University*

### ABSTRACT

Realistic audio services based on ultra multi-channel audio signals are indispensable for realistic high quality video such as 3D movies and UHD broadcasting. Generally, since the ultra multi-channel audio signals to be more than 10.2 channels has very high data rate, an efficient ultra multi-channel audio coding method is required to enable the realistic audio services in a wired/wireless network or communication environment. In this paper, we propose the ultra multi-channel audio coding based on space division of more than 10.2 channels playback layout for efficiently handling the ultra multi-channel audio signals with low bit rate and complexity. The proposed method is originated by considering the conventional 5.1 channel playback system and the usage of the multi-channel audio coder such as MPEG Surround or SSLCC. Input ultra multi-channel audio signals are divided into five spaces such as Center, Left Front, Left Back, Right Front, and Right Back and allocated signals to each space are independently encoded. One down-mix and power ratio parameters for each space are extracted through the encoding process and five down-mix signals from each space are coded by the conventional multi-channel audio coding. We implemented a simple ultra multi-channel audio coder adopting the down-mixing and the power ratio of the audio signal to check the feasibility of the proposed method and we confirmed that the proposed ultra multi-channel audio coding has low bit-rate to be less than 200 kbps for 10.2 channel audio signals. Since the proposed method focused on low bit-rate and complexity and it may be suffered from the degradation of sound quality, researches on the improvement of sound quality using a variable subband and a residual signal processing remain as future works.

© 2018 KKITS All rights reserved

**KEYWORDS:** Realistic audio services, Ultra multi-channel audio signals, Ultra multi-channel audio coding, Down-mixing, Power ratio

**ARTICLE INFO:** Received 3 January 2018, Revised 31 January 2018, Accepted 8 February 2018.

\*Corresponding author is with School of IT Convergence,  
Korea Nazarene University, 48 Wolbong-ro Seobuk-gu

Cheonan-si Chungcheongnam-do, 31172, S.KOREA.  
E-mail address: k2kim@kornu.ac.kr

## 1. 서론

최근, 3D 영화 및 UHD (Ultra High Definition) 방송과 같은 초 고화질 비디오의 출현으로 기존의 재생 시스템을 통해 현실감 넘치는 사운드를 재생할 수 있는 실감음향 기술에 대한 관심이 커지고 있으며, 많은 연구가 진행되고 있다. 실감음향 기술은 10.2 채널에서 22.2 채널 이상의 초다채널 오디오 재생시스템, 전송된 오디오 신호의 채널 수와 사용자의 재생 환경의 채널 수가 불일치 할 때 이를 보상하기 위한 채널 매핑 기술, 스테레오 헤드폰으로 3D 사운드를 즐길 수 있도록 하는 바이노럴 렌더링 기술, 초다채널 오디오 신호의 높은 데이터양을 줄이기 위한 초다채널 오디오 부호화 기술 등으로 나눌 수 있다 [1-4]. 이 중 초다채널 오디오 신호의 압축/복원과 관련된 부호화 기술에 대한 연구는 부족한 상황이며, 현재 연구가 진행되고 있는 초다채널 오디오 부호화기는 기존의 오디오 부호화기를 재활용하는 것에 머무르고 있어 초다채널 신호의 채널 수에 비례해서 데이터율과 복잡도가 크게 증가하여 유/무선 통신 네트워크나 통신환경에서 실감음향 서비스를 제공하기에는 어렵다는 문제를 지니고 있다. 따라서, 본 논문에서는 유/무선 네트워크나 통신환경에서 실감음향 서비스가 가능하도록 하기 위해 필수적인 기술인 초다채널 오디오 부호 기술을 다루도록 한다.

초다채널 오디오 신호의 데이터 전송률은 채널 수에 비례해서 크게 증가하므로 초다채널 신호를 부호화 과정 없이 네트워크나 통신환경을 통해 전송하여 실감음향 서비스를 제공하는 것은 불가능하다. 이러한 문제점을 해결하기 위해 모노 / 스테레오 신호 또는 다채널 오디오 신호의 부호화에 사용되었던 오디오 부호화기를 재활용하여 초다채널 오디오 신호를 효율적으로 압축/복원하기 위한 연구가 수행되고 있다. 대표적으로 MPEG 오디오

그룹은 10.2 채널에서 22.2 채널 이상의 초다채널 신호를 효율적으로 부호화/복호화 하는 MPEG-H 3D Audio 표준을 개발하였다 [5,6]. 모노/스테레오 음성/오디오 통합 부호화기인 MPEG-D USAC (Unified Speech and Audio Coding)[7,8]에 기반한 MPEG-H 3D Audio는 초다채널 오디오 신호를 2개의 채널씩 묶은 후, 각각의 스테레오 신호를 MPEG-D USAC으로 부호화를 수행한다. MPEG-H 3D Audio는 입력 초다채널 신호를 2개의 신호씩 묶어서 부호화하기 때문에 22.2 채널의 경우 11개의 USAC 부호화기가 필요하며 이를 실행하기 위해서는 높은 비트율과 고사양의 시스템이 필요하다는 제약을 지니고 있다. 따라서 본 논문에서는 MPEG Surround[9,10]나 SSLCC (Sound Source Location Coefficient Coding)[11-13]와 같은 기존에 개발되어있는 다채널 오디오 부호화기를 활용한 초다채널 오디오 부호화기 구조를 제안한다. 제안된 부호화기 구조에서는 초다채널 신호를 5개의 공간으로 분할하여 부호화함으로써 하나 또는 두 개의 기존의 부호화기만으로 초다채널 오디오 신호를 부호화하기 때문에 높은 압축율과 낮은 복잡도를 얻을 수 있어 실감음향 서비스가 유/무선 네트워크 및 통신환경과 모바일 환경에서도 가능하게 된다.

본 논문은 다음과 같이 구성된다. 먼저 2장에서 제안된 초다채널 오디오 부호화기의 개요를 설명하고, 3장에서는 공간 분할 기반의 초다채널 오디오 신호의 압축과 복원 방법을 자세히 설명한다. 4장에서는 제안된 방법의 실제 구현을 통한 간단한 실험 결과와 추가적인 고려사항을 제시하고, 5장에서 결론을 맺도록 한다.

## 2. 제안된 초다채널 오디오 부호화기의 개요

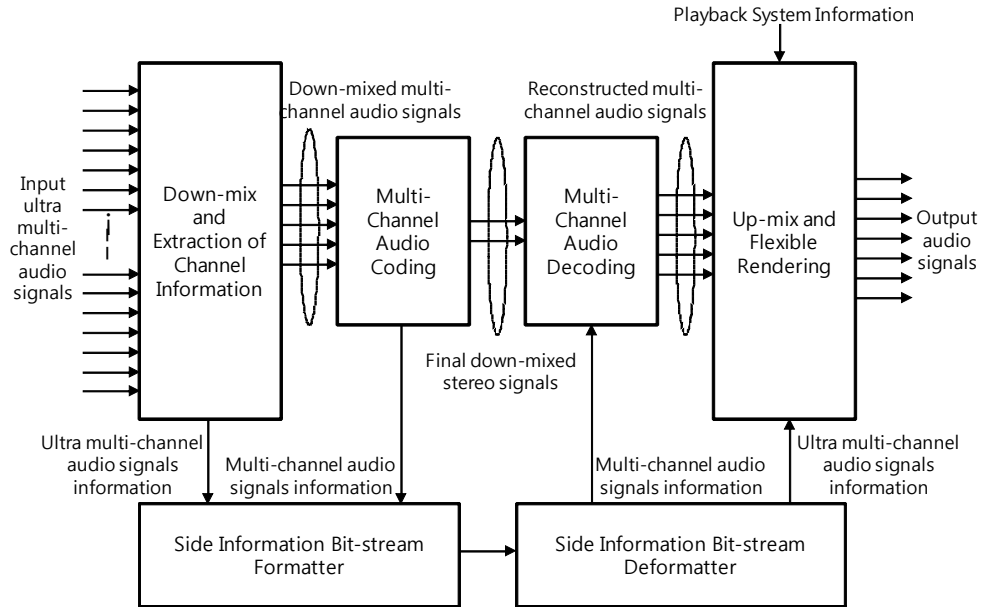


그림 1. 제안된 초다채널 오디오 부호화기 구조  
 Figure 1. Proposed ultra multi-channel audio coding structure

〈그림 1〉은 기존의 다채널 오디오 부호화기를 재사용하여 제안된 초다채널 오디오 부호화기의 구조를 보여준다. 5.1 채널 재생시스템이 널리 상용화 되어 있어 이를 최대한 재활용하고, 이미 개발되어 있는 다채널 오디오 부호화기가 5.1채널 신호를 효율적으로 처리할 수 있기 때문에 MPEG Surround 나 SSLCC와 같은 다채널 오디오 부호화기에 기반한 초다채널 오디오 부호화기를 제안한 것이다. 입력된 초다채널 오디오 신호는 5.1 채널 오디오 신호로 다운믹싱이 되며, 이 과정에서 초다채널 신호로의 복원을 위해 각각의 채널 정보(채널의 파워, 상관도 등)를 추출하여 복호화 단으로 전송한다. 또한, 다운믹싱을 통해 생성된 5.1 채널 오디오 신호는 기존의 다채널 오디오 부호화기를 이용하여 부호화하며, 다채널 오디오 부호화기를 이용할 경우 부호화 과정을 통해 최종적인 스테레오 다운믹스 신호와 5.1 채널 신호로의 복원을 위한 각각의 채널 정보(채널의 파워, 상관도 등)를

추출하여 복호화 단으로 전송한다. 정리하면, 전체 초다채널 부호화 과정을 통해 하나의 스테레오 다운믹스 신호, 스테레오 다운믹스 신호를 5.1 채널로 복원하기 위한 5.1 채널 파라미터, 그리고 5.1채널 신호를 초다채널 신호로 복원하기 위한 초다채널 파라미터가 생성되어 복호화 단으로 전송되는 것이다.

복호화 과정은 부호화 과정의 역순으로 진행되며, 먼저 전송된 스테레오 다운믹스 신호와 5.1 채널 파라미터를 이용하여 5.1 채널 신호를 복원한다. 다음으로 복원된 5.1 채널 신호와 초다채널 파라미터를 이용하여 초다채널 신호를 복원하며, 사용자의 재생환경 정보에 따라서 초다채널 신호의 렌더링을 통해 최종 출력신호를 생성하게 된다. 최종 출력신호를 생성하는 렌더링은 스피커 재생환경을 위해서는 매트릭스 기반의 채널 매핑 기법을 적용하며 헤드폰 재생환경을 위해서는 바이노럴 렌더링 기법을 적용한다.

### 3. 제안된 초다채널 오디오 부호화기의 상세

초다채널 오디오 신호를 부호화하기 위해서 초다채널 재생시스템을 <그림 2>에서 보는 바와 같이 5개의 공간 (Center, Left Front, Left Back, Right Front, Right Back)으로 분할한다. 각 공간 별로 분할된 공간에 해당되는 채널 신호를 이용하여 하나의 다운믹스 신호와 각 공간에 해당되는 채널 신호의 복원을 위한 파라미터를 추출한다. 이는 일반 가정에서 널리 쓰이는 5.1 채널 재생시스템과 이미 개발되어 있는 다채널 오디오 부호화기의 재활용을 고려한 것으로 10.2 채널 이상의 초다채널 오디오 신호를 5.1 채널로 다운믹스하며, 각각의 공간 별로 파라미터를 추출하여 원래의 10.2 채널 이상의 초다채널 오디오 신호를 복원하도록 한다. 한편, 5.1 채널 다운믹스 신호는 필요에 따라 다채널 오디오 부호화기에 의해서 스테레오 다운믹스 신호와 파라미터로 압축/복원될 수 있다.

그림 3은 Left Front 공간의 부호화 예를 나타낸다. 그림과 같이 Left Front 공간상에  $L_1, L_2, L_3, L_4$  총 4개의 채널신호가 존재할 때, 하나의 다운믹스 신호는 식 (1)과 같이 네 개의 신호를 모두 더해서 생성하며 생성된 다운믹스 신호는 5.1 채널 신호의 Lf 신호로 간주한다.

$$Lf = L_1 + L_2 + L_3 + L_4 \quad (1)$$

생성된 Lf를 이용하여 Left Front 공간상의 4개의 원신호를 복원하기 위한 공간 파라미터는 각각의 채널의 파워비를 이용하며, 아래 식 (2)를 이용하여 계산한다.

$$PR_i = \frac{P_i}{P_{max}} \text{ for } 1 \leq i \leq 4 \quad (2)$$

여기에서  $PR_i$ 와  $P_i$ 는  $i$ 번째 채널의 파워비와 파워를 각각 나타내며,  $P_{max}$ 는 Left Front에 있는 전체 채널 신호 중 가장 큰 파워를 나타낸다. 추출된 파워비는 전송을 위하여 4비트 양자화 과정을 거치며,  $P_{max}$ 를 이용해 파워비를 0과 1사이로 제한함으로써 효율적인 양자화를 할 수 있다. 이러한 다운믹싱과 파라미터 추출은 전체 재생공간으로 확대하여 나머지 4개의 공간에 대해서도 수행되며, 최종적으로 5.1 채널 다운믹스 신호와 5개의 공간 각각의 신호 복원을 위한 파라미터를 추출하게 된다.

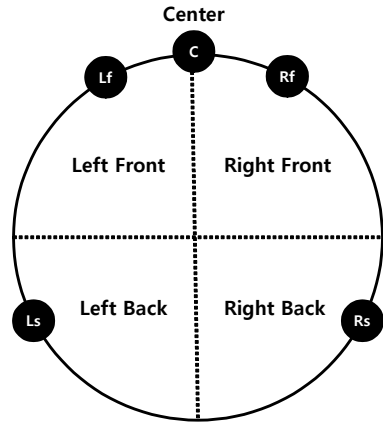


그림 2. 초다채널 오디오 부호화를 위한 재생공간분할  
Figure 2. Space segmentation of the ultra multi-channel playback system

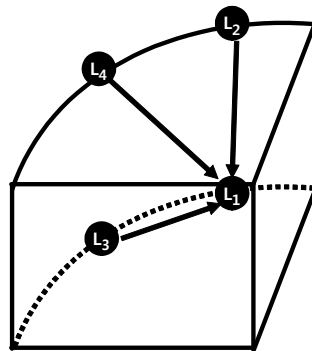


그림 3. Left front 공간의 부호화 예  
Figure 3. An example of coding in the left front space.

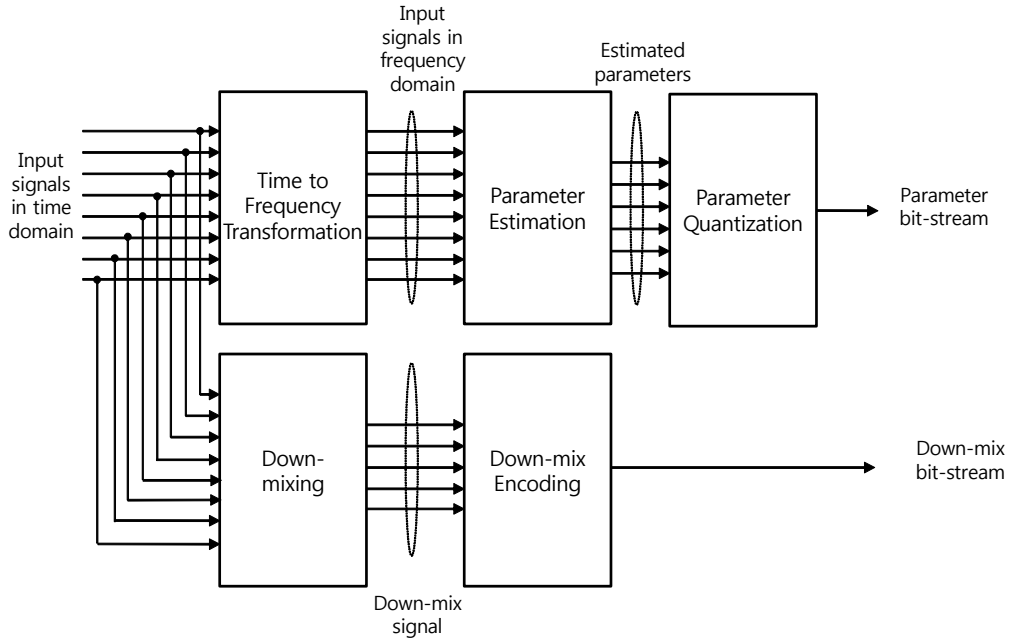


그림4. 초다채널 오디오 부호화 상세도

Figure 4. Detailed process of the ultra multi-channel audio encoding

다운믹스 신호  $L_f$ 와 파라미터를 이용한 Left Front 공간상의 원신호의 복원은 식 (3)을 이용하여 간단히 수행된다.

$$\hat{L}_i = L_f \times \frac{PR_i}{\sum_{j=1}^4 PR_j} \text{ for } 1 \leq i \leq 4 \quad (3)$$

여기에서  $\hat{L}_i$  는  $i$ 번째 채널의 복원 신호를 나타낸다. 식 (3)은 Left Front 공간상의 4개의 신호의 합이 다운믹스 신호이므로 다운믹스 신호에 4개의 신호 각각이 차지하는 파워의 비율을 곱함으로써 쉽게 원신호를 계산할 수 있음을 보여준다. 다운믹스 신호와 파라미터를 이용한 공간상의 원신호의 복원은 나머지 4개의 공간에 대해서도 동일하게 수행되어 최종적으로 초다채널 오디오 신호를 복원할 수 있다.

<그림 4>는 제안된 초다채널 오디오 부호화의 상세도를 나타낸다. 초다채널 오디오 부호화는 다운믹싱과 원신호 복원을 위한 파라미터 추출 두 단계로 나뉜다. 다운믹싱은 식 (1)을 이용하여 시간영역에서 수행되고, 파라미터 추출은 식 (2)를 이용하여 주파수영역에서 수행된다. 특히, 파라미터는 인간의 청각 특성을 반영하여 주파수 영역을 28개로 나눈 서브밴드에서 추출된다 [14,15]. 즉, 인간의 청각이 민감한 저주파 영역은 서브밴드의 대역을 좁게 설정하고, 상대적으로 둔감한 고주파 대역은 서브밴드의 대역을 넓게 설정하여 파라미터를 추출함으로써 복원된 신호의 음질의 열화를 최소화 하도록 한다. 생성된 다운믹스 신호는 다채널 오디오 부호화기에 의한 다운믹스 부호화 과정을 거쳐 다운믹스 비트스트림을 생성하며, 추출된 파라미터는 4비트 16레벨 양자화 과정을 거쳐 파라미터 비트스트림을 생성한다.

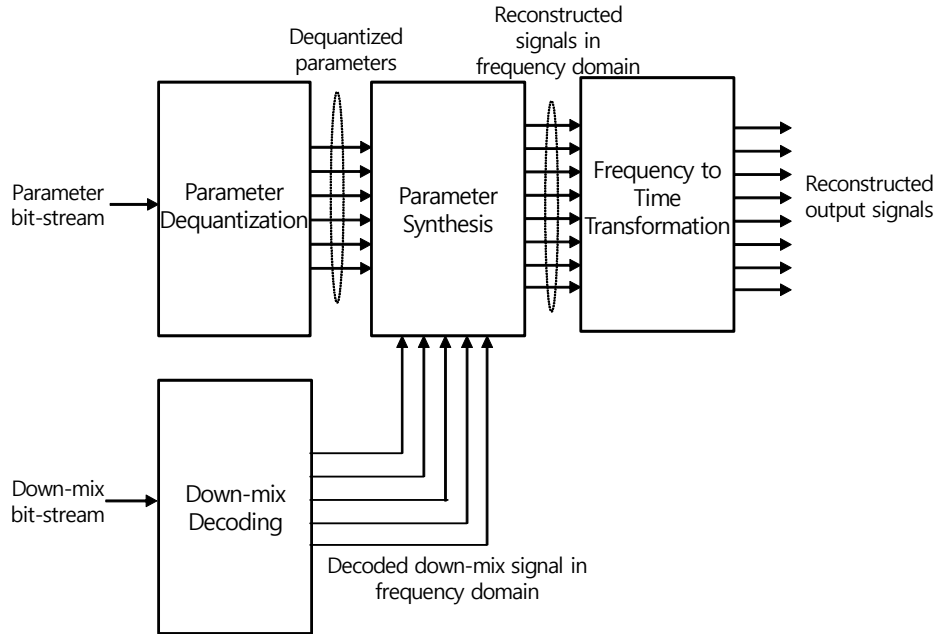


그림5. 초다채널 오디오 복호화 상세도

Figure 5. Detailed process of the ultra multi-channel audio decoding

<그림 5>는 제안된 초다채널 오디오 복호화의 상세도를 나타낸다. 그림에서 보는 바와 같이 초다채널 오디오 복호화는 부호화의 역 과정으로 수행되며, 다운믹스 신호와 파라미터를 이용한 초다채널 신호의 복원은 주파수 영역에서 수행된다. 복호화를 통해 복원된 다운믹스 신호와 역양자화를 거친 파라미터를 식 (3)에 대입하여 주파수 영역에서의 초다채널 오디오 신호를 생성하며, 주파수-시간 영역 변환을 통해 시간영역에서의 초다채널 신호를 생성한다.

#### 4. 실험 및 추가적인 고려사항

앞에서 설명한 다운믹싱 방법과 파워비를 파라미터로 갖는 간단한 초다채널 오디오 부호화기를 구현하여 비트율을 확인함으로써 제안된 초다채널 부호화 방법의 타당성을 확인하였다. 이를 위해

10.2 채널 오디오 신호에서 저음 강화를 위한 LFE 채널을 제외한 10 개의 채널 신호를 부호화하여 비트율을 확인하였다.

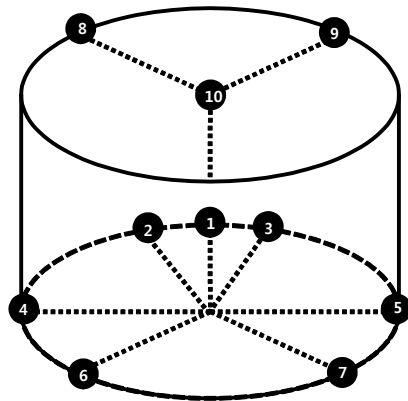


그림 6. 10.2 채널 레이아웃

Figure 6. 10.2 channel configuration

표 1. 10.2 채널 재생공간 분할

Table 1. Space segmentation of 10.2 channel configuration

분할 공간	할당 채널
Center	1, 10
Left front	2, 8
Left back	4, 6
Right front	3, 9
Right back	5, 7

<그림 6>의 10.2 채널 재생 레이아웃을 <표 1>과 같이 재생 공간을 분할하여 분할된 공간 별로 다운믹싱과 파라미터를 추출한 후, 각 공간의 다운믹스 신호를 기존의 다채널 부호화기를 활용하여 부호화 하도록 하였다. 여기에서 기존의 다채널 부호화기는 SSLCC를 활용하였다. 실험을 위하여 3개의 10.2 채널 오디오 콘텐츠를 활용하였으며, 오디오 콘텐츠는 20초 이내의 지속시간과 44.1 kHz 샘플링 주파수, 16 비트 양자화를 갖는다.

실험 결과 다운믹스 신호의 비트율을 제외한 각 재생공간에서의 파워비 파라미터만의 전체 비트율은 약 50 kbps이며, 기존의 다채널 오디오 부호화기를 활용하여 다운믹스 신호를 부호화 하였을 경우 소모되는 비트율은 대략적으로 140 kbps로 측정되었다. 또한, 현재 파라미터 비트율은 파라미터를 일괄적으로 4 비트 양자화를 적용하여 측정하였기 때문에 추후에 양자화 인덱스의 확률분포를 고려한 허프만 코딩을 적용하였을 경우 비트율을 더욱 줄일 수 있을 것으로 예상된다.

본 실험은 초다채널 오디오 부호화기의 간단한 구현을 통해 비트율만 조사한 것으로 복원된 신호의 음질확인을 위한 청취평가가 필요하다. 청취평가가 수행되지 않았지만, 본 논문에서 제안한 간단한 초다채널 오디오 부호화 방법은 음질의 열화를 피할 수 없다. 이는 다운믹스된 신호와 주파수 영역을 28개의 영역으로 나눈 서브밴드에서 추출한 파라미터만을 이용하여 초다채널 신호를 복원하고자 하는 부호화 기술 자체가 지니는 원천적인 문

제로 이를 해결하기 위하여 다음과 같은 부호화 모드가 필요하다.

표 2. 초다채널 오디오 부호화기 모드 제안

Table 2. Suggestion of ultra multi-channel audio coding mode

모드	설명	비트율 (10.2채널 기준)
일반 모드	28개 서브밴드에서 추출한 파워비 파라미터만을 이용하여 압축/복원	약 190 kbps (다운믹스 140 kbps+채널당 5kbps)
가변 서브밴드 모드	인간의 청각특성과 신호의 주파수 특성을 반영하여 서브밴드의 수를 28, 40, 55, 70 개 등으로 가변하여 파라미터를 추출함으로써 주파수 해상도를 높여 음질 개선	약 290 kbps (다운믹스 140 kbps +채널당 15kbps)
잔차 신호 모드	파워비 파라미터와 다운믹스 신호만으로 복원된 신호와 원신호 사이의 차이를 나타내는 잔차신호를 추출하여 이를 추가적인 파라미터로 이용함으로써 복원된 신호를 최대한 원신호와 유사하게 만들어 음질 개선	약 440 kbps (다운믹스 140 kbps+채널당 30kbps)

먼저, 가변 서브밴드 모드는 일반모드와 달리 고정된 서브밴드 수가 아닌 신호의 주파수 특성에 따라서 서브밴드의 수를 달리 함으로써 주파수 해상도를 높여 음질 열화를 최소화하기 위함이다. 특정 채널의 신호의 주파수 특성이 크게 변화하지 않고 일정할 경우에는 기존의 28개의 서브밴드만으로 주파수 특성을 표현할 수 있지만, 주파수 특성의 변화가 심할 때는 서브밴드의 수를 늘려 주파수 해상도를 높여야 복원된 신호의 음질 열화를 피할 수 있기 때문이다. 다음으로 잔차신호 모드는 다운믹스 신호와 서브밴드에서 추출하여 양자화를 거친 파워비 파라미터를 이용하여 원래의 신호를

복원하였을 때, 원 신호와 복원된 신호 사이의 차이가 존재하므로 이를 잔차신호로 추출하여 이용함으로써 복원된 신호가 최대한 원신호와 가깝도록 만들어 음질을 개선하기 위해 제안된 것이다. 추후에 이 두 개의 모드를 구현함으로써 현재의 초다채널 오디오 부호화기의 음질 측면에서의 성능을 크게 향상시킬 수 있을 것으로 기대된다.

## 5. 결론

본 논문에서는 10.2 채널에서 22.2 채널 이상의 초다채널 오디오 신호를 효율적으로 처리하기 위한 재생 공간 분할 기반의 초다채널 오디오 부호화 방법을 제안하였다. 제안된 방법은 5.1 채널 재생환경을 고려하여 입력된 초다채널 오디오 신호를 Center, Left Front, Left Back, Right Front, Right Back으로 나눠서 각 공간별로 다운믹스와 파라미터 추출을 수행한 후, 각 공간별 다운믹스 신호를 기존에 개발되어 있는 MPEG Surround나 SSLCC와 같은 다채널 오디오 부호화기로 부호화하도록 한다. 제안된 초다채널 오디오 부호화기는 낮은 비트율과 적은 복잡도를 지니기 때문에 초다채널 오디오 신호에 의한 실감음향 서비스가 유/무선 네트워크 및 통신 환경에서도 가능하도록 한다. 검증을 위해 초다채널 오디오 부호화기를 구현하고, 이를 활용하여 10.2 채널 오디오 신호를 부호화 한 결과 비트율은 약 190 kbps를 보여 제안된 초다채널 오디오 부호화기의 타당성을 확인할 수 있었다.

제안된 초다채널 오디오 부호화기는 비트율과 복잡도의 최소화에 초점을 맞춰서 구현된 상태로 압축/복원된 오디오 신호의 음질 평가를 통한 검증이 추가로 진행되어야 한다. 일반적으로 다운믹싱과 서브밴드 처리에 의한 음질의 열화는 피할 수 없으므로 가변 서브밴드 처리에 의한 주파수

해상도 개선과 원신호와 복원된 신호의 차이를 최소화 하기 위한 잔차신호 적용을 통한 음질의 향상은 향후 과제로 남아있다. 추가적으로 전송된 초다채널 오디오 신호와 일치하는 재생환경을 가지고 있지 않은 사용자를 고려하여 초다채널 신호를 사용자의 재생환경에 맞도록 매핑하는 방법과 스테레오 헤드폰 재생환경을 고려한 바이노럴 사운드 생성 방법에 대한 연구도 필요하다.

## References

- [1] T. Holman, *The history and future of DSPs in consumer audio equipment-Part II: emerging areas and the future*, ICCE2008.
- [2] K. Hamasaki, S. Komiyama, K. Hiyama, and H. Okubo, *5.1 and 22.2 multichannel sound productions using an integrated surround sound panning system*, 2005 NAB BEC Proceedings, Apr. 2005.
- [3] Y. Lee, S. Kim, H. Jo, Y. Park, and Y. J. Kim, *Virtual height speaker rendering for Samsung 10.2-channel vertical surround system*, 131st AES Convention, Oct. 2011.
- [4] D. Jang, J. Seo, T. Lee, and K. Kang, *Present and future of UHD sound technology*, Korea Society Broadcast Engineers Magazine Vol. 17, No. 5, pp. 47-59, 2012.
- [5] Herre, J., Hilpert, J., Kuntz, A., & Plogsties, J., *MPEG-H 3D audio-the new standard for coding of immersive spatial audio*, IEEE Journal of selected topics in signal processing, Vol. 9, No. 5, pp. 770-779, 2015.
- [6] ISO/IEC 23008-3:2015, *MPEG-H (High efficiency coding and media deliver in heterogeneous environments), Part 3: 3D Audio*, 2015.



[7] M. Neuendorf, M. Multrus, N. Rettelbach, G. Fuchs, J. Robilliard, J. Lecomte, and R. Lefebvre, *The ISO/MPEG unified speech and audio coding standard-consistent high quality for all content types and at all bit rates*, Journal of the Audio Engineering Society, Vol. 61, No. 12, pp. 956-977, 2013.

[8] ISO/IEC 23003-3:2012, *Information Technology-MPEG Audio Technologies-Part 3: Unified Speech and Audio Coding*, 2012

[9] J. Herre, H. Purnhagen, and J. Breebard, *The reference model architecture for MPEG spatial audio coding*, in Proceedings of the 118th AES Convention, Barcelona, Spain, 2005.

[10] ISO/IEC 23003-1:2007, *Information technology-MPEG audio technologies-part 1: MPEG surround*, 2007.

[11] H. Moon, J. Seo, S. Beack and K. Sung, *A multi-channel audio compression method with virtual source location information for MPEG-4 SAC*, IEEE Transactions on Consumer Electronics, Vol. 51, No. 4, pp. 1253-1259, 2005.

[12] S. Beack, J. Seo, H. Moon, K. Kang, M. Hahn, *Angle-based virtual source location representation for spatial audio coding*. ETRI journal, Vol. 28, No. 2, pp. 219-222, 2006.

[13] K. Kim, S. Beack, J. Seo, D. Jang, and M. Hahn, *Improved channel level difference quantization for spatial audio coding*, ETRI Journal, Vol. 29, No. 1, pp. 99-102, 2007.

[14] F. Baumgarte and C. Faller, *Binaural cue coding—part I: psychoacoustic fundamentals and design principles*, IEEE Transactions on Speech and Audio Processing, Vol. 11, No. 6, pp. 509-519, 2003.

[15] Zwicker, Eberhard, and Hugo Fastl. *Psychoacoustics: Facts and models*, Vol. 2. Berlin: Springer, 1999.

---

## 실감음향 서비스를 위한 초다채널 오디오 부호화에 관한 연구

김광기

나사렛대학교 IT융합학부

---

### 요 약

3D 영화나 UHD 방송과 같은 실감나는 고화질 비디오의 출현과 함께 이를 뒷받침하는 현실감 넘치는 사운드를 재생하기 위한 초다채널 오디오 신호 기반의 실감 음향 서비스는 필수적이다. 일반적으로 초다채널 오디오 신호는 10.2에서 22.2 이상의 채널 수를 가지므로 처리해야 될 데이터 양이 매우 크기 때문에 유/무선 네트워크나 통신환경에서 실감음향 서비스가 가능하도록 만들기 위해서는 초다채널 오디오 신호를 효율적으로 부호화하는 방법이 필요하다. 본 논문에서는 10.2 채널 이상의 초다채널 오디오 신호를 낮은 비트율과 복잡도를 가지고 처리하기 위한 재생 공간 분할 기반의 초다채널 오디오 부호화 방법을 제안한다. 제안된 방법은 5.1 채널 재생 환경과 MPEG Surround나 SSLCC 같은 다채널 오디오 부호화기의 사용을 고려하여 고안되었으며, 입력된 초다채널 신호를 Center, Left Front, Left Back, Right Front, Right Back의 5개 공간으로 나눠 각 공간별로 부호화를 수행한다. 공간별로 하나의 다운믹스 신호와 원신호 복원을 위한 파워비 파라미터를 추출하고, 5개의 다운믹스 신호는 다시 기존의 다채널 오디오 부호화기에 의해서 부호화함으로써 초다채널 오디오 신호를 효율적으로 부호화 할 수 있다. 검증을 위해 제안된 초다채널 오디오 부호화기를 구현하였으며, 200 kbps 미만의 낮은 비트율로 10.2 채널 오디오 신호를 부호화할 수 있음을 확인하였다. 제안된 초다채널 오디오 부호화기는 비트율과 복잡도의 최소화에 초점을 맞춰서 구현된 것으로 복원된 신호의 음질의 열화를 피할 수 없기 때문에 가변 서브밴드 처리나 잔차신호 적용 등과 같은 음질 향상 방안에 관한 연구는 향후 과제로 남아있다.

---

## Acknowledgments

본 논문은 2017년도 나사렛대학교 교비학술연구조성비 지원에 의해 연구되었으며, 2017년도 정부(교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(No. 2017R1D1A3B03034951).



**Kwangki Kim** received the B.S. degree in electronic engineering from Korea Aviation University, Koyang, South Korea, in 2002, the M.S. degree and the Ph.D. degree in department of Information and Communications Engineering at Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2004 and 2011, respectively. He worked as a researcher in Samsung DMC R&D center in 2012. Currently, He is an assistant professor in Korea Nazarene University, Chon-an, Korea. His research interests include multi-channel/-object audio coding, spatial audio, 3D sound, and their applications. He is a life member of the KKITS.

*E-mail address:* k2kim@kornu.ac.kr