

파라메트릭 멀티채널 오디오 코덱의 다운믹스 방식에 대한 연구

문한길, 이철우
삼성전자 주식회사 DM연구소

Study on the downmix method of parametric multichannel audio codec

Han-gil Moon, Chu-Iwoo Lee
DM R&D Center, Samsung Electronics

Abstract - DVD/BD 및 HDTV의 보급으로 인해 다수의 오디오 콘텐츠들이 멀티채널(5.1채널 이상) 형식으로 제작되고 있다. 오디오 정보를 담고 있는 물리적인 채널의 수가 증가하면, 이에 따라 정보량도 선형적으로 증가하게 된다. 선형적으로 증가된 정보를 기존의 오디오 코덱을 이용해서 압축할 경우, 압축에 필요한 비트레이트의 선형적인 증가를 피할 수 없다. 최근 채널 수 증가로 야기되는 비트레이트의 증가를 최소화하고 효율적으로 멀티채널 오디오 신호를 압축할 수 있는 방법으로 MPEG surround와 같은 파라메트릭 멀티채널 오디오 코딩 방식이 제안되었다. 파라메트릭 멀티채널 오디오 코딩 방식의 경우, 멀티채널 오디오 신호를 채널 수가 감소된 다운믹스 신호와 다운믹스 신호로부터 다시 멀티채널 오디오 업믹스 하는데 필요한 파라미터로 표현하는 방식이다. 따라서 다운믹스 방식 및 업믹스에 필요한 파라미터에 따라 업믹스된 멀티채널 오디오 신호의 품질이 달라진다. 본 논문에서는 MPEG surround에서 사용하고 있는 기존의 ITU-R 다운믹스 방식의 문제점을 실제 멀티채널 오디오 신호의 사례를 통해 제시하고 이 문제점을 해결하기 위한 새로운 다운믹스 방식과 파라미터를 제안하고자 한다.

서 4배까지 증가되었다. 기존의 오디오 부호화 방식을 이용해서 늘어난 오디오 채널을 부호화 할 경우, 비트레이트 역시 3배에서 4배 증가한다. 그러나 파라메트릭 멀티채널 오디오 부호화 방식을 이용할 경우, 5.1채널 혹은 7.1채널의 입력 신호를 부호화 하더라도 그에 소요되는 비트레이트는 스테레오 신호를 부호화 하는데 필요한 비트레이트와 크게 차이가 나지 않는다. 그 이유는 그림 1에서의 같이 다채널 오디오 신호를 스테레오 혹은 모노 신호로 다운믹스하고, 다운믹스된 신호로부터 원래의 다채널 오디오 신호를 복원하는데 필요한 추가정보를 매우 효율적으로 부호화하기 때문이다. 일반적으로 파라메트릭 멀티채널 오디오 코덱에서 사용하는 다운믹스 방식은 ITU에서 정의한 ITU 다운믹스 방식을 사용한다.

1. 서 론

파라메트릭 멀티채널 오디오 코덱(Parametric Multichannel Audio Codec)은 다채널 오디오 신호를 보다 효율적으로 압축하기 위하여 다채널 오디오 신호(n>2)를 스테레오 혹은 모노신호와 이 스테레오 또는 모노 신호로부터 다채널 오디오 신호를 복원하기 위한 추가정보로 표현하는 방법이다. 파라메트릭 멀티채널 오디오 코덱을 이용하여 효율적으로 멀티채널 신호를 압축하기 위해서는 압축 효율면에서 이득이 많은 우수한 추가정보를 정의하여 사용하는 것이 중요하다. 또한 압축된 스테레오 또는 모노 신호로부터 복원한 멀티채널 오디오 신호의 품질을 유지하기 위해서는 멀티채널 오디오 신호를 어떤 방식으로 다운믹스 했는지도 중요하다. 기존의 파라메트릭 멀티채널 오디오 코덱 가운데, 표준화된 기술인 MPEG surround의 경우, ICLD, OPD, IPD, ICC등을 추가정보로 사용하고 ITU 다운믹스를 사용한다[1]. 본 논문에서는 기존 방법보다 우수한 멀티채널 오디오 음질을 확보하기 위해서 새로운 다운믹스 방식을 제안하고, 그 성능을 기존 방법과 비교하고자 한다. 성능 평가를 위해서는 현재 MPEG 오디오 그룹에서 멀티채널 오디오의 성능 평가를 위해 사용 중인 음원을 포함한 다양한 장르의 음원을 사용하였고, 평가 방법으로는 MUSHRA test[3]를 이용하였다.

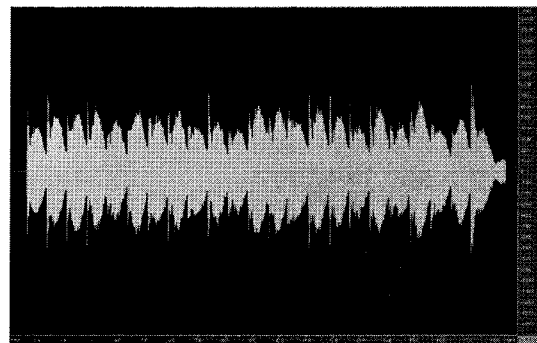
2.2 ITU 다운믹스

다운믹스는 다채널 오디오 신호를 스테레오 신호로 만들기 위한 일종의 방정식이다. 예를 들어 5.1채널의 채널수를 갖는 다채널 오디오 신호를 스테레오 신호로 만들 수 있는 방법은 이론적으로 무한하지만, 일반적인 경우 ITU(International Telecommunication Union)에서 정한 표준 다운믹스[4] 방식을 이용해서 5.1채널을 스테레오로 다운믹스 하는데 ITU 다운믹스는 다음과 같은 수식(1), 수식(2)로 표현된다.

$$L_t = 1.00L + 0.71C + 0.71L_s \quad \text{수식(1)}$$

$$R_t = 1.00R + 0.71C + 0.71R_s \quad \text{수식(2)}$$

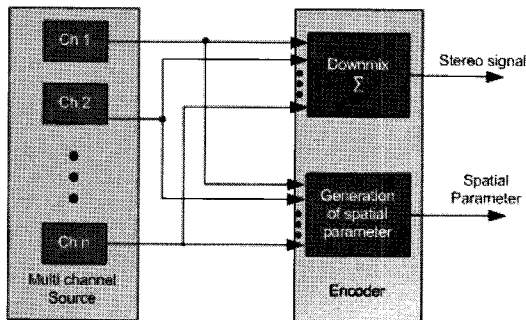
2.2.1 ITU 다운믹스의 문제점



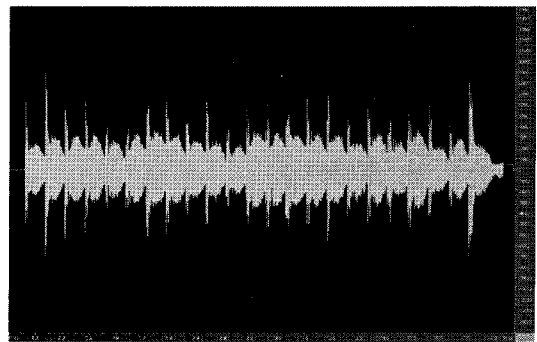
<그림 2> Glock의 우전방 신호 (Input)

2. 본 론

2.1 파라메트릭 멀티채널



<그림 1> 파라메트릭 멀티채널 오디오 인코더 구조



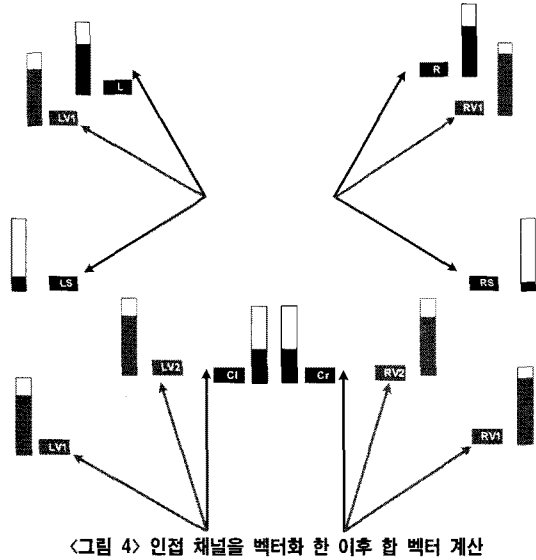
<그림 3> Glock의 우전방 신호 (Conventional output)

다채널 오디오 신호를 기존의 채널별 오디오 코딩 방식을 이용해서 부호화 하면, 채널 수 증가에 따른 비트레이트의 선형적인 증가를 피할 수 없다. 최근 HDTV, DVD 및 BD등에 채용되고 있는 오디오 규격은 5.1채널 혹은 7.1채널 오디오 신호로 기존의 스테레오 신호에 비해 채널 수는 3배에

위의 그림 2는 멀티채널 테스트 신호 가운데 하나인 glock의 우전방(R) 신호의 파형을 나타내고, 그림 3은 파라메트릭 멀티채널 오디오 코덱 중 하나인 MPEG surround를 이용하여 멀티채널 신호(glock)를 부호화 및 복호화 하여 얻은 신호의 우전방(R) 신호의 파형이다. 두 그림을 비교해보면

MPEG surround를 이용하여 부호화 및 복호화 한 신호의 시간 파형(그림 3)이 원본(그림 2)와는 차이가 있음을 확인할 수 있다. 부호화중에 거치는 다운믹스를 통해서 여러 채널의 오디오 신호가 하나의 신호로 합쳐지게 되는데, 이 과정에서 phase 성분 차이로 인해 사라지는 오디오 신호가 발생하게 되고, 이렇게 사라지는 오디오 신호는 다운믹스된 신호로부터 다채널 오디오 신호를 복원하기 위해 별도로 부호화하는 부가정보가 아무리 정확하더라도 다시 복원될 수 없다. 본 논문에서는 ITU 다운믹스 방식의 이러한 문제점을 해결하기 위한 새로운 방식을 제안하고자 한다.

2.3 벡터 기반의 다운믹스



〈그림 4〉 인접 채널을 벡터화 한 이후 합 벡터 계산

위의 그림 4는 5.1채널을 다운믹스하는 방식을 도식화 한 그림이다. 스칼라 계산을 통해서 다운믹스 신호를 만들어내는 ITU 다운믹스 방식과는 달리 본 논문을 통해서 제안하는 방식은 다운믹스가 일어나는 인접 채널의 오디오 신호를 서로 직교하는 두 개의 위치 벡터에 매핑 한 이후 합 벡터를 구하고, 구해진 합 벡터와 다른 인접 채널의 오디오 신호 역시 서로 직교 좌표를 갖도록 매핑 한 후 합 벡터를 구하는 방식으로 다운믹스를 순차적으로 계속한다. 다운믹스를 수행하면서 추출하는 부가정보는 다음과 수식과 같은 방식으로 추출하고 복원한다.

$$\sin^2 \theta + \cos^2 \theta = 1$$

$$\theta_m = \frac{\theta_{pan} - \theta_1}{\theta_2 - \theta_1} \times 90 \quad \theta_2 \geq \theta_{pan} \geq \theta_1$$

$$g_1 = \cos \theta_m \quad g_2 = \sin \theta_m \quad \text{수식(3)}$$

$$\text{수식(4)}$$

$$\text{수식(5)}$$

$$\sin^2 \theta + \cos^2 \theta = 1$$

$$\theta_n = \frac{\theta_{pan} - \theta_1}{\theta_2 - \theta_1} \times 90 \quad \theta_2 \geq \theta_{pan} \geq \theta_1$$

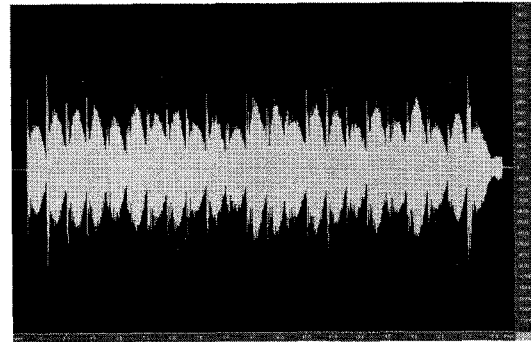
$$g_1 = \cos \theta_n \quad g_2 = \sin \theta_n \quad \text{수식(6)}$$

$$\text{수식(7)}$$

$$\text{수식(8)}$$

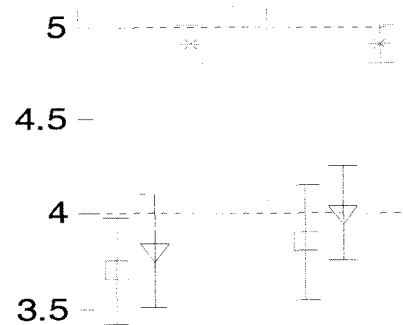
2.3.1 벡터 기반 다운믹스의 결과

다음 그림 5는 본 논문을 통해 제안한 방식으로 다운믹스 하여 부호화하고 복호화한 신호의 우전방 신호 파형이다. 원본 신호(그림 2) 및 기존 ITU 다운믹스 방식으로 만든 신호(그림 3)와 비교해보면, 제안한 방식으로 얻은 신호의 파형이 보다 원본 신호(그림 2)에 가까운 모양을 갖고 있음을 확인할 수 있다.



〈그림 5〉 Glock의 우전방 신호 (Proposed output)

2.4 MUSHRA 테스트를 이용한 청취평가



〈그림 6〉 MUSHRA 테스트

제안한 다운믹스 방식의 실제적인 성능 측정을 위해서 MUSHRA 테스트를 수행하였다. 테스트를 위해 총 7개의 5.1채널 음원을 사용하였고 MUSHRA 테스트에 참여한 인원은 총 8명이다. 7개의 음원 가운데 4개의 음원에서는 평가 결과 제안된 방식이 더 우수한 결과를 나타냈고, 2개의 음원에 대해서는 이전과 같은 점수를 나타냈고, 하나의 음원에 대해서는 낮은 점수를 나타냈으나 전체적으로는 음질이 향상되는 것을 확인할 수 있었다.

3. 결 론

본 논문에서는 MPEG surround와 같은 파라메트릭 멀티채널 오디오 코덱에서 사용하고 있는 다운믹스 방식을 제안하였다. 기존의 ITU 다운믹스 방식의 경우, 스칼라 합을 통해서 다운믹스 신호를 만들고 이 과정에서 phase차로 인해 사라지는 오디오 정보가 존재했다. 이 문제점을 해결하기 위해 본 논문에서는 벡터 기반의 다운믹스 방식을 제안하였고, 제안한 방식으로 구현한 코덱을 통해 얻은 파형과 MUSHRA 테스트를 통해 복호화된 오디오 신호의 품질이 향상됨을 확인할 수 있었다.

〔참 고 문 헌〕

- [1] J.Breebaart, J.Herre, C.Faller, J.Roden., F.Myburg, S.Disch, H.Purnhagen, G.Hotho, M.Neusinger, K.Kjorling, W.Oomen, "MPEG Spatial Audio Coding / MPGE Surround: Overview and Current Status", 119th Conv. Audio Eng. Soc., Oct. 2005
- [2] Han-gil Moon; Jeong-il Seo; Seungkwon Baek; Koeng-Mo Sung; "A multi-channel audio compression method with virtual source location information for MPEG-4 SAC" Consumer Electronics, IEEE Transactions on Volume 51, Issue 4, Nov. 2005 Page(s):1253-1259
- [3] ITU-R Recommendation BS.1534-1, "Method for the Subjective Assessment of Intermediate Sound Quality (MUSHRA)", ITU, Geneva, Switzerland, 2001
- [4] ITU-R Recommendation BS.775, "Multichannel stereophonic sound system with and without accompanying picture", ITU, Geneva, Switzerland, 1992