



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2018-0009337
 (43) 공개일자 2018년01월26일

- (51) 국제특허분류(Int. Cl.)
G10L 19/16 (2013.01) *G10L 19/00* (2006.01)
G10L 19/002 (2013.01) *G10L 19/008* (2014.01)
- (52) CPC특허분류
G10L 19/167 (2013.01)
G10L 19/0017 (2013.01)
- (21) 출원번호 10-2017-7033556
- (22) 출원일자(국제) 2016년06월17일
 심사청구일자 없음
- (85) 번역문제출일자 2017년11월20일
- (86) 국제출원번호 PCT/KR2016/006495
- (87) 국제공개번호 WO 2016/204581
 국제공개일자 2016년12월22일
- (30) 우선권주장
 62/181,096 2015년06월17일 미국(US)
 (뒷면에 계속)

- (71) 출원인
삼성전자주식회사
 경기도 수원시 영통구 삼성로 129 (매탄동)
- (72) 발명자
김선민
 경기도 용인시 수지구 신봉1로 110 507동 2002호
전상배
 경기도 수원시 영통구 매영로 110 103동 504호
- (74) 대리인
리엔텍특허법인

전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **저연산 포맷 변환을 위한 인터널 채널 처리 방법 및 장치**

(57) 요약

본 발명의 일 실시예에 따른 오디오 신호를 처리하는 방법은, MPS212(MPEG Surround 212)를 이용하여 인코딩 된 오디오 비트스트림을 수신하는 단계; 수신된 오디오 비트스트림 및 포맷 변환기에 정의된 MPS212 출력 채널들에 대한 렌더링 파라미터 중 EQ(Equalization) 값들 및 게인 값들에 기초하여, 하나의 CPE(Channel Pair Element)에 대한 인터널 채널 신호를 생성하는 단계; 및 생성된 인터널 채널 신호에 기초하여 스테레오 출력 신호들을 생성하는 단계;를 포함한다.

(52) CPC특허분류

G10L 19/002 (2013.01)

G10L 19/008 (2013.01)

(30) 우선권주장

62/241,082 2015년10월13일 미국(US)

62/241,098 2015년10월13일 미국(US)

62/245,191 2015년10월22일 미국(US)

명세서

청구범위

청구항 1

MPS212(MPEG Surround 212)를 이용하여 인코딩된 오디오 비트스트림을 수신하는 단계;

상기 수신된 오디오 비트스트림 및 포맷 변환기에 정의된 MPS212 출력 채널들에 대한 EQ(Equalization) 값들 및 게인 값들에 기초하여, 하나의 CPE(Channel Pair Element)에 대한 인터널 채널 신호를 생성하는 단계; 및
상기 생성된 인터널 채널 신호에 기초하여 스테레오 출력 신호들을 생성하는 단계;를 포함하는,
오디오 신호를 처리하는 방법.

청구항 2

제 1 항에 있어서,

상기 인터널 채널 신호를 생성하는 단계는,

상기 수신된 오디오 비트스트림을, MPS212 페이로드에 포함된 CLD(Channel Level Difference)에 기초하여, 상기 하나의 CPE에 포함된 채널 쌍에 대한 신호로 업믹싱하는 단계;

상기 업믹싱된 비트스트림을, 상기 EQ(Equalization) 값들 및 게인 값들에 기초하여, 스케일링하는 단계; 및

상기 스케일링된 비트스트림을 믹싱하는 단계;를 포함하는,

오디오 신호를 처리하는 방법.

청구항 3

제 1 항에 있어서,

상기 인터널 채널 신호를 생성하는 단계는,

상기 하나의 CPE에 대한 인터널 채널 신호 생성 여부를 결정하는 단계를 더 포함하는,

오디오 신호를 처리하는 방법.

청구항 4

제 3 항에 있어서,

상기 인터널 채널 신호 생성 여부는,

상기 하나의 CPE에 포함되는 채널 쌍이 같은 인터널 채널 그룹에 해당하는지 여부에 기초하여 결정되는,

오디오 신호를 처리하는 방법.

청구항 5

제 4 항에 있어서,

상기 하나의 CPE에 포함되는 채널 쌍이 모두 좌측 인터널 채널 그룹에 포함되는 경우, 상기 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널로만 출력되고,

상기 하나의 CPE에 포함되는 채널 쌍이 모두 우측 인터널 채널 그룹에 포함되는 경우, 상기 인터널 채널 신호는 스테레오 출력 채널 중 우측 출력 채널로만 출력되는,

오디오 신호를 처리하는 방법.

청구항 6

제 4 항에 있어서,

상기 하나의 CPE에 포함되는 채널 쌍이 모두 센터 인터널 채널 그룹에 포함되거나, 모두 LFE(Low Frequency Effect) 인터널 채널 그룹에 포함되는 경우, 상기 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널과 우측 출력 채널로 균등하게 출력되는,

오디오 신호를 처리하는 방법.

청구항 7

제 1 항에 있어서,

상기 인터널 채널 신호를 생성하는 단계는,

인터널 채널 게인을 계산하는 단계; 및

상기 인터널 채널 게인을 적용하는 단계;를 더 포함하는,

오디오 신호를 처리하는 방법.

청구항 8

MPS212(MPEG Surround 212)를 이용하여 인코딩 된 오디오 비트스트림을 수신하는 수신부;

상기 수신된 오디오 비트스트림 및 포맷 변환기에 정의된 MPS212 출력 채널들에 대한 EQ(Equalization) 값들 및 게인 값들에 기초하여, 하나의 CPE(Channel Pair Element)에 대한 인터널 채널 신호를 생성하는 인터널 채널 신호 생성부; 및

상기 생성된 인터널 채널 신호에 기초하여 스테레오 출력 신호들을 생성하는 스테레오 출력 신호 생성부;를 포함하는,

오디오 신호를 처리하는 장치.

청구항 9

제 8 항에 있어서,

상기 인터널 채널 신호 생성부는,

상기 수신된 오디오 비트스트림을, MPS212 페이로드에 포함된 CLD(Channel Level Difference)에 기초하여, 상기 하나의 CPE에 포함된 채널 쌍에 대한 신호로 업믹싱하고,

상기 업믹싱된 비트스트림을, 상기 EQ(Equalization) 값들 및 게인 값들들에 기초하여, 스케일링하고,

상기 스케일링된 비트스트림을 믹싱하는,

오디오 신호를 처리하는 장치.

청구항 10

제 8 항에 있어서,

상기 인터널 채널 신호 생성부는,

상기 하나의 CPE에 대한 인터널 채널 신호 생성 여부를 결정하는,

오디오 신호를 처리하는 방법.

청구항 11

제 10 항에 있어서,

상기 인터널 채널 신호 생성 여부는 상기 하나의 CPE에 포함되는 채널 쌍이 같은 인터널 채널 그룹에 해당하는 지 여부에 기초하여 결정되는,

오디오 신호를 처리하는 장치.

청구항 12

제 11 항에 있어서,

상기 하나의 CPE에 포함되는 채널 쌍이 모두 좌측 인터널 채널 그룹에 포함되는 경우, 상기 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널로만 출력되고,

상기 하나의 CPE에 포함되는 채널 쌍이 모두 우측 인터널 채널 그룹에 포함되는 경우, 상기 인터널 채널 신호는 스테레오 출력 채널 중 우측 출력 채널로만 출력되는,

오디오 신호를 처리하는 장치.

청구항 13

제 11 항에 있어서,

상기 하나의 CPE에 포함되는 채널 쌍이 모두 센터 인터널 채널 그룹에 포함되거나, 모두 LFE(Low Frequency Effect) 인터널 채널 그룹에 포함되는 경우, 상기 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널과 우측 출력 채널로 균등하게 출력되는,

오디오 신호를 처리하는 장치.

청구항 14

제 8 항에 있어서,

상기 인터널 채널 신호 생성부는,

인터널 채널 계인을 계산하고, 상기 인터널 채널 계인을 적용하는,

오디오 신호를 처리하는 장치.

청구항 15

제 1 항에 따른 방법을 실행하기 위한 컴퓨터 프로그램을 기록하는 컴퓨터 판독 가능한 기록 매체.

발명의 설명

기술 분야

[0001] 본 발명은 저연산 포맷 변환을 위한 인터널 채널 처리 방법 및 장치에 대한 것으로, 보다 상세하게는, 스테레오 출력 레이아웃 환경에서 입력 채널들에 대한 인터널 채널 처리를 수행함으로써 포맷 변환기의 입력 채널 개수를 감소시켜 포맷 변환기에서 수행되는 공분산 연산 횟수를 감소시키는 방법 및 장치에 대한 발명이다.

배경 기술

[0002] MPEG-H 3D 오디오는 다양한 종류의 신호를 처리할 수 있으며, 입출력 형태의 제어가 용이하여 차세대 오디오 신호 처리를 위한 해결책으로 기능한다. 또한, 기기의 소형화 경향 및 시대의 흐름에 따라 오디오 재생 환경은 스테레오 재생 환경의 모바일 기기를 통해 재생되는 비율이 높아지고 있다.

[0003] 22.2 채널 등 다채널로 구현되는 실감 오디오(immersive audio) 신호가 스테레오 재생 시스템으로 전달되는 경우, 모든 입력 채널이 디코딩되어야 하며 실감 오디오 신호를 다운믹스하여 스테레오 포맷으로 변환해야 한다.

[0004] 입력 채널의 개수가 증가할수록, 또한 출력 채널의 개수가 감소할수록 이와 같은 과정에서 공분산 분석 및 위상 정합을 위해 필요한 디코더의 복잡도는 증가한다. 이와 같은 복잡도의 증가는 모바일 기기에서 연산 속도뿐 아니라 배터리 소모에도 큰 영향을 미치게 된다.

발명의 내용

해결하려는 과제

[0005] 상술한 바와 같이, 실감 음향을 제공하기 위하여 입력 채널의 개수는 증가하는 반면 휴대성을 위하여 출력 채널

의 개수는 감소하는 환경에서, 디코딩시 포맷 변환을 위한 복잡도가 문제된다.

[0006] 본 발명은 전술한 종래 기술의 문제점을 해결하며, 디코더에서 포맷 변환의 복잡도를 감소시키는 것을 그 목적으로 한다.

과제의 해결 수단

[0007] 상기 목적을 달성하기 위한 본 발명의 대표적인 구성은 다음과 같다.

[0008] 상기 기술적 과제를 해결하기 위한 본 발명의 일 실시예에 따른 오디오 신호를 처리하는 방법은, MPS212(MPEG Surround 212)를 이용하여 인코딩 된 오디오 비트스트림을 수신하는 단계; 수신된 오디오 비트스트림 및 포맷 변환기에 정의된 MPS212 출력 채널들에 대한 렌더링 파라미터 중 EQ(Equalization) 값들 및 게인 값들에 기초하여, 하나의 CPE(Channel Pair Element)에 대한 인터널 채널 신호를 생성하는 단계; 및 생성된 인터널 채널 신호에 기초하여 스테레오 출력 신호들을 생성하는 단계;를 포함한다.

[0009] 본 발명의 또 다른 실시예에 따르면, 인터널 채널 신호를 생성하는 단계는, 수신된 오디오 비트스트림을, MPS212 페이지드에 포함된 CLD(Channel Level Difference)에 기초하여, 하나의 CPE에 포함된 채널 쌍에 대한 신호로 업믹싱하는 단계; 및 업믹싱된 비트스트림을, 렌더링 파라미터들에 기초하여, 스케일링하는 단계; 및 스케일링된 비트스트림을 믹싱하는 단계;를 포함한다.

[0010] 본 발명의 또 다른 실시예에 따르면, 인터널 채널 신호를 생성하는 단계는, 하나의 CPE에 대한 인터널 채널 신호 생성 여부를 결정하는 단계를 더 포함한다.

[0011] 본 발명의 또 다른 실시예에 따르면, 인터널 채널 신호 생성 여부는, 하나의 CPE에 포함되는 채널 쌍이 같은 인터널 채널 그룹에 해당하는지 여부에 기초하여 결정된다.

[0012] 본 발명의 또 다른 실시예에 따르면, 하나의 CPE에 포함되는 채널 쌍이 모두 좌측 인터널 채널 그룹에 포함되는 경우, 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널로만 출력되고, 하나의 CPE에 포함되는 채널 쌍이 모두 우측 인터널 채널 그룹에 포함되는 경우, 인터널 채널 신호는 스테레오 출력 채널 중 우측 출력 채널로만 출력된다.

[0013] 본 발명의 또 다른 실시예에 따르면, 하나의 CPE에 포함되는 채널 쌍이 모두 센터 인터널 채널 그룹에 포함되거나, 모두 LFE(Low Frequency Effect)채널 그룹에 포함되는 경우, 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널과 우측 출력 채널로 균등하게 출력된다.

[0014] 본 발명의 또 다른 실시예에 따르면, 오디오 신호는 실감 오디오(immersive audio) 신호이다.

[0015] 본 발명의 또 다른 실시예에 따르면, 인터널 채널 신호를 생성하는 단계는, 인터널 채널 게인을 계산하는 단계; 및 인터널 채널 게인을 적용하는 단계;를 더 포함한다.

[0016] 상기 기술적 과제를 해결하기 위한 오디오 신호를 처리하는 장치는, MPS212(MPEG Surround 212)를 이용하여 인코딩 된 오디오 비트스트림을 수신하는 수신부; 수신된 오디오 비트스트림 및 포맷 변환기에 정의된 MPS212 출력 채널들에 대한 렌더링 파라미터 중 EQ(Equalization) 값들 및 게인 값들에 기초하여, 하나의 CPE(Channel Pair Element)에 대한 인터널 채널 신호를 생성하는 인터널 채널 생성부; 및 생성된 인터널 채널 신호에 기초하여 스테레오 출력 신호들을 생성하는 스테레오 출력 신호 생성부;를 포함한다.

[0017] 본 발명의 또 다른 실시예에 따르면, 인터널 채널 신호 생성부는, 수신된 오디오 비트스트림을, MPS212 페이지드에 포함된 CLD(Channel Level Difference)에 기초하여, 하나의 CPE에 포함된 채널 쌍에 대한 신호로 업믹싱하고, 업믹싱된 비트스트림을, 렌더링 파라미터들에 기초하여, 스케일링하고, 스케일링된 비트스트림을 믹싱한다.

[0018] 본 발명의 또 다른 실시예에 따르면, 인터널 채널 생성부는, 하나의 CPE에 대한 인터널 채널 신호 생성 여부를 결정한다.

[0019] 본 발명의 또 다른 실시예에 따르면, 하나의 CPE에 포함되는 채널 쌍이 같은 인터널 채널 그룹에 해당하는지 여부에 기초하여 결정된다.

[0020] 본 발명의 또 다른 실시예에 따르면, 하나의 CPE에 포함되는 채널 쌍이 모두 좌측 인터널 채널 그룹에 포함되는 경우, 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널로만 출력되고, 하나의 CPE에 포함되는 채널 쌍이 모두 우측 인터널 채널 그룹에 포함되는 경우, 인터널 채널 신호는 스테레오 출력 채널 중 우측 출력 채널

로만 출력된다.

- [0021] 본 발명의 또 다른 실시예에 따르면, 하나의 CPE에 포함되는 채널 쌍이 모두 센터 인터널 채널 그룹에 포함되거나, 모두 LFE(Low Frequency Effect)채널 그룹에 포함되는 경우, 인터널 채널 신호는 스테레오 출력 채널 중 좌측 출력 채널과 우측 출력 채널로 균등하게 출력된다.
- [0022] 본 발명의 또 다른 실시예에 따르면, 오디오 신호는 실감 오디오(immersive audio) 신호이다.
- [0023] 본 발명의 또 다른 실시예에 따르면, 인터널 채널 신호 생성부는, 인터널 채널 게인을 계산하고, 인터널 채널 게인을 적용한다.
- [0024] 한편, 본 발명의 일 실시예에 따르면, 전술한 방법을 실행하기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.
- [0025] 이 외에도, 본 발명을 구현하기 위한 다른 방법, 다른 시스템 및 상기 방법을 실행하기 위한 컴퓨터 프로그램을 기록하는 컴퓨터 판독 가능한 기록 매체가 더 제공된다.

발명의 효과

- [0026] 본 발명에 의하면, 인터널 채널을 이용함으로써 포맷 변환기(format converter)에 입력되는 채널의 숫자를 감소시켜, 포맷 변환기의 복잡도를 감소시킬 수 있다. 보다 구체적으로, 포맷 변환기에 입력되는 채널의 숫자가 감소됨으로써 포맷 변환기에서 수행되는 공분산 분석이 간소화되어 복잡도가 감소되는 효과가 있다.

도면의 간단한 설명

- [0027] 도 1 은 24개의 입력 채널을 스테레오 출력 채널로 포맷 변환하는 디코딩 구조에 대한 일 실시예를 나타낸다.
- 도 2 는 22.2 채널 실감 오디오(immersive audio) 신호를 13 개의 인터널 채널을 이용하여 스테레오 출력 채널로 포맷 변환하는 디코딩 구조에 대한 일 실시예를 나타낸다.
- 도 3 은 하나의 CPE로부터 하나의 인터널 채널을 생성하는 일 실시예를 나타낸다.
- 도 4 는 본 발명의 일 실시예에 따른, 디코더에서 인터널 채널 신호에 인터널 채널 게인 적용부의 세부 블록도이다.
- 도 5 는 본 발명의 일 실시예에 따른, 인코더에서 인터널 채널 게인이 전처리(pre-processed)되는 경우의 디코딩 블록도이다.
- 도 6 은 본 발명의 일 실시예에 따른, CPE가 스테레오 재생 레이아웃으로 출력되는 경우, 모노 SBR 디코딩 후 MPS 디코딩하는 구조에서의 인터널 채널 처리 방법의 순서도이다.
- 도 7 은 본 발명의 일 실시예에 따른, CPE가 스테레오 재생 레이아웃으로 출력되는 경우, MPS디코딩 후 스테레오 SBR 디코딩하는 구조에서의 인터널 채널 처리 방법의 순서도이다.
- 도 8 은 본 발명의 일 실시예에 따른, QCE가 스테레오 재생 레이아웃으로 출력되는 경우, 스테레오 SBR을 이용하는 구조에서의 인터널 채널 처리 방법의 블록도이다.
- 도 9 는 본 발명의 또 다른 실시예에 따른, QCE가 스테레오 재생 레이아웃으로 출력되는 경우, 스테레오 SBR을 이용하는 구조에서의 인터널 채널 처리 방법의 블록도이다.
- 도 10a 는 첫번째 포락선의 시작 경계들이 서로 같고, 마지막 포락선의 종료 경계들이 서로 같은 경우의 시간 포락선 그리드를 결정하는 일 실시예를 나타낸다.
- 도 10b 는 첫번째 포락선의 시작 경계들은 서로 다르고, 마지막 포락선의 종료 경계들은 같은 경우의 시간 포락선 그리드를 결정하는 일 실시예를 나타낸다.
- 도 10c 는 첫번째 포락선의 시작 경계들은 서로 같고, 마지막 포락선의 종료 경계들은 서로 다른 경우의 시간 포락선 그리드를 결정하는 일 실시예를 나타낸다.
- 도 10d 는 첫번째 포락선의 시작 경계들이 서로 다르고, 마지막 포락선의 종료 경계들이 서로 다른 경우의 시간 포락선 그리드를 결정하는 일 실시예를 나타낸다.
- 표 1 은 22.2 채널 실감 오디오(immersive audio) 신호를 스테레오 신호로 렌더링하는 포맷 변환기의 믹싱 매트

릭스의 일 실시예를 나타낸다.

표 2 는 22.2 채널 실감 오디오(immersive audio) 신호를 인터널 채널을 스테레오 신호로 렌더링하는 포맷 변환기의 믹싱 매트릭스의 일 실시예를 나타낸다.

표 3 은 본 발명의 일 실시예에 따른, 22.2 채널을 인터널 채널로 구성하기 위한 CPE 구조를 나타낸다.

표 4 는 본 발명의 일 실시예에 따른, 디코더 입력 채널에 대응되는 인터널 채널들의 타입을 나타낸다.

표 5 는 본 발명의 일 실시예에 따른, 인터널 채널 타입에 따라 추가적으로 정의되는 채널의 위치를 나타낸다.

표 6 은 본 발명의 일 실시예에 따른, 인터널 채널 타입에 대응되는 포맷 변환기 출력 채널 및 각 출력 채널에 적용될 게인과 EQ 인덱스를 나타낸다.

표 7 은 본 발명의 일 실시예에 따른, ICGConfig의 선택스를 나타낸다.

표 8 은 본 발명의 일 실시예에 따른, mpeg3daExtElementConfig()의 선택스를 나타낸다.

표 9 는 본 발명의 일 실시예에 따른, usacExtElementType를 나타낸다.

표 10 은 본 발명의 일 실시예에 따른, speakerLayoutType을 나타낸다.

표 11 은 본 발명의 일 실시예에 따른, SpeakerConfig3d()의 선택스를 나타낸다.

표 12 는 본 발명의 일 실시예에 따른, immersiveDownmixFlag를 나타낸다.

표 13 은 본 발명의 일 실시예에 따른, SAOC3DgetNumChannels()의 선택스를 나타낸다.

표 14 는 본 발명의 일 실시예에 따른, 채널 할당 순서를 나타낸다.

표 15는 본 발명의 일 실시예에 따른, mpeg3daChannelPairElementConfig()의 선택스를 나타낸다.

표 16 은 본 발명의 일 실시예에 따른, 채널 구성 요소 및 재생 레이아웃에 기초하여 결정되는 MPS와 SBR의 디코딩 시나리오를 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0028]

발명의 실시를 위한 최선의 형태

[0029]

상기 목적을 달성하기 위한 본 발명의 대표적인 구성은 다음과 같다.

[0030]

상기 기술적 과제를 해결하기 위한 본 발명의 일 실시예에 따른 오디오 신호를 처리하는 방법은, MPS212(MPEG Surround 212)를 이용하여 인코딩 된 오디오 비트스트림을 수신하는 단계; 수신된 오디오 비트스트림 및 포맷 변환기에 정의된 MPS212 출력 채널들에 대한 렌더링 파라미터 중 EQ(Equalization) 값들 및 게인 값들에 기초하여, 하나의 CPE(Channel Pair Element)에 대한 인터널 채널 신호를 생성하는 단계; 및 생성된 인터널 채널 신호에 기초하여 스테레오 출력 신호들을 생성하는 단계;를 포함한다.

[0031]

발명의 실시를 위한 형태

[0032]

후술하는 본 발명에 대한 상세한 설명은, 본 발명이 실시될 수 있는 특정 실시예를 예시로서 도시하는 첨부 도면을 참조한다. 이러한 실시예는 당업자가 본 발명을 실시할 수 있기에 충분하도록 상세히 설명된다. 본 발명의 다양한 실시예는 서로 다르지만 상호 배타적일 필요는 없음이 이해되어야 한다.

[0033]

예를 들어, 본 명세서에 기재되어 있는 특정 형상, 구조 및 특성은 본 발명의 정신과 범위를 벗어나지 않으면서 일 실시예로부터 다른 실시예로 변경되어 구현될 수 있다. 또한, 각각의 실시예 내의 개별 구성요소의 위치 또는 배치도 본 발명의 정신과 범위를 벗어나지 않으면서 변경될 수 있음이 이해되어야 한다. 따라서, 후술하는 상세한 설명은 한정적인 의미로서 행하여지는 것이 아니며, 본 발명의 범위는 특허청구범위의 청구항들이 청구하는 범위 및 그와 균등한 모든 범위를 포괄하는 것으로 받아들여져야 한다.

[0034]

도면에서 유사한 참조부호는 여러 측면에 걸쳐서 동일하거나 유사한 구성요소를 나타낸다. 그리고 도면에서 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 유사한 부분에 대해서는 유사한 도면 부호를 붙였다.

[0035]

이하에서는, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 용이하게 실시할 수 있도록 하

기 위하여, 본 발명의 여러 실시예에 관하여 첨부된 도면을 참조하여 상세히 설명하기로 한다. 그러나 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예에 한정되지 않는다.

- [0036] 명세서 전체에서, 어떤 부분이 다른 부분과 "연결"되어 있다고 할 때, 이는 "직접적으로 연결"되어 있는 경우뿐 아니라, 그 중간에 다른 소자를 사이에 두고 "전기적으로 연결"되어 있는 경우도 포함한다. 또한 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0037] 본 명세서 사용되는 용어의 정의는 다음과 같다.
- [0038] 인터널 채널(IC, Internal Channel)은, MPS212(MPEG Surround stereo) 업믹싱과 포맷 변환기(FC, Format Converter) 다운믹싱에서 발생하는 불필요한 연산을 제거하기 위해, 포맷 변환 과정에서 사용되는 가상의 중간 단계(intermediate) 채널로, 스테레오 출력을 고려한다.
- [0039] 인터널 채널 신호(internal channel signal)는, 스테레오 신호를 제공하기 위하여 포맷 변환기에서 믹싱되는 모노신호로, 인터널 채널 게인을 이용하여 생성된다.
- [0040] 인터널 채널 처리(internal channel processing)는, MPS212 디코딩 블록에 기초하여 인터널 채널 신호를 생성하는 처리를 의미하며, 인터널 채널 처리 블록에서 수행된다.
- [0041] 인터널 채널 게인(ICG, Internal Channel Gain)은, CLD(Channel Level Difference) 값과 포맷 변환 파라미터들로부터 계산되는, 인터널 채널 신호에 적용되는 게인을 의미한다.
- [0042] 인터널 채널 그룹(internal channel group)은, 코어 코덱(core codec) 출력 채널 위치에 기초하여 결정되는 인터널 채널의 타입을 의미하며 코어 코덱 출력 채널 위치 및 인터널 채널 그룹은 표 4에 정의되어 있다(후술함).
- [0043] 이하 첨부된 도면을 참고하여 본 발명을 상세히 설명하기로 한다.
- [0044] 도 1 은 24개의 입력 채널을 스테레오 출력 채널로 포맷 변환하는 디코딩 구조에 대한 일 실시예를 나타낸다.
- [0045] 멀티 채널 입력의 비트스트림이 디코더로 전달되면, 디코더에서는 입력 채널 레이아웃이 재생 시스템의 출력 채널 레이아웃에 맞게 다운믹싱된다. 예를 들어, 도 1 과 같이 MPEG 표준을 따르는 22.2 채널 입력 신호가 스테레오 채널 출력 시스템으로 재생될 때, 디코더에 포함되는 포맷 변환기(130)는 포맷 변환기 내부에 규정된 포맷 변환기 규칙에 따라 24 개의 입력 채널 레이아웃을 2 개의 출력 채널 레이아웃으로 다운믹싱한다.
- [0046] 이 때, 디코더에 입력되는 22.2 채널 입력 신호는 하나의 CPE(Channel Pair Element)에 포함되는 두개의 채널에 대한 신호들이 다운믹스된 CPE 비트스트림(110)들을 포함한다. CPE 비트스트림은 MPS212(MPEG Surround based stereo)를 이용해 인코딩되어 있으므로, 수신된 CPE 비트스트림은 MPS212(120)를 이용해 디코딩된다. 이 때, LFE채널, 즉 우퍼채널은 CPE로 구성되지 않는다. 따라서, 22.2 채널 입력이라면 디코더 입력 신호는11개의 CPE 에 대한 비트스트림과 두개의 우퍼채널에 대한 비트스트림으로 구성된다.
- [0047] 22.2 채널 입력 신호를 구성하는 CPE 비트스트림들에 대한 MPS212 디코딩이 수행되면, 각각의 CPE에 대한 두개의 MPS212 출력 채널(121, 122)이 생성되며, MPS212를 이용해 디코딩 된 출력 채널(121, 122)들은 포맷 변환기의 입력 채널이 된다. 도 1 과 같은 경우 포맷 변환기의 입력 채널 개수 N_{in} 은 우퍼채널을 포함하여 24가 된다. 따라서, 포맷 변환기에서는 24×2 다운믹싱이 수행되어야 한다.
- [0048] 포맷 변환기에서는, 멀티 채널 신호들 사이의 위상 차이에 의한 음색 왜곡(timbral distortion)을 방지하기 위해, 공분산 분석(covariance analysis)에 따른 위상 정합(phase alignment)이 수행된다. 이 때, 공분산 매트릭스는 $N_{in} \times N_{in}$ 차원을 가지므로, 공분산 매트릭스를 분석하기 위해서는 이론적으로 $(N_{in} \times (N_{in}-1)/2 + N_{in}) \times 71\text{band} \times 2 \times 16 \times X (48000/2048)$ 번의 복소수 곱셈이 수행되어야 한다.
- [0049] 입력 채널의 개수 N_{in} 이 24인 경우, 한번의 복소수 곱셈을 위해서는 네번의 연산이 수행되어야 하며 대략 64 MOPS(Million Operations Per Second)의 성능이 요구된다.
- [0050] 표 1 은 22.2 채널 실감 오디오(immersive audio) 신호를 스테레오 신호로 렌더링하는 포맷 변환기의 믹싱 매트릭스의 일 실시예를 나타낸다.
- [0051] 표 1 의 믹싱 매트릭스에서 가로축(140)과 세로축(150)은 24개의 입력 채널에 대해 번호를 매긴 것으로 공분산 분석에서 그 순서는 큰 의미를 가지지 않는다. 표 1 에 개시된 실시예에서 믹싱 매트릭스의 각 요소(element)가 1의 값을 갖는 경우(160)는 공분산 분석이 필요하지만, 0의 값을 갖는 경우(170)는 공분산 분석이 생략될 수

있다.

[0052] 예를 들어, CM_M_L030 및 CH_M_R030 채널과 같이 스테레오 출력 레이아웃으로의 포맷 변환 과정에서 서로 믹싱되지 않은 입력 채널들의 경우, 믹싱 매트릭스에서 해당 요소의 값이 0이 되며 서로 믹싱되지 않는 CM_M_L030 및 CH_M_R030채널 상호간의 공분산 분석 과정이 생략될 수 있다.

[0053] 따라서 24*24번의 공분산 분석 중 서로 믹싱되지 않은 입력 채널들에 대한 128번의 공분산 분석을 제외할 수 있다.

[0054] 또한, 믹싱 매트릭스는 입력 채널에 따라 대칭으로 구성되므로, 표 1 에서

표 1

0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0
2	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1
3	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1
4	1	0	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
5	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1
6	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0	1	1	0
7	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1
8	1	1	1	1	0	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1
9	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0	1
10	1	1	1	1	1	0	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1	0
11	0	1	1	1	1	1	0	1	1	1	1	0	1	1	1	0	1	1	1	0	1	1	1
12	1	0	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0
13	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0
14	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	0	1	1	1
15	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	0	1	1
16	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1	0
17	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1	1
18	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1	1
19	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0
21	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0
22	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0
23	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0

[0055]

[0056] 대각선을 기준으로 하단(190)과 상단(180)을 나누어, 하단에 해당하는 영역은 공분산 분석을 생략할 수 있다. 또한 대각선의 상단에 해당하는 영역 중 굵은 글씨로 기재된 부분에 대해서만 공분산 분석이 수행되므로, 최종적으로 236번의 공분산 분석이 수행된다.

[0057] 이와 같이 믹싱 매트릭스의 값이 0인 경우(서로 믹싱되지 않는 채널들) 및 믹싱 매트릭스의 대칭성을 이용해 불필요한 공분산 분석 과정을 제거하면 공분산 분석은 236 X 71band X 2 X 16 X (48000/2048)번의 복소수 곱셈이 수행되어야 한다.

[0058] 따라서, 이와 같은 경우 50 MOPS가 요구되므로 모든 믹싱 매트릭스에 대하여 공분산 분석을 수행하는 경우에 비해 공분산 분석에 의한 시스템 부하가 개선되는 효과가 있다.

[0059] 도 2 는 22.2 채널 실감 오디오(immersive audio) 신호를 13개의 인터널 채널을 이용하여 스테레오 출력 채널로

포맷 변환하는 디코딩 구조에 대한 일 실시예를 나타낸다.

- [0060] 한편 MPEG-H 3D 오디오는 제한된 전송 환경에서 멀티채널 오디오 신호를 보다 효율적으로 전달하기 위해 CPE를 이용한다. 하나의 채널 쌍에 해당하는 두 개의 채널들이 스테레오 레이아웃으로 믹싱되는 경우, 채널간 상관도 (ICC, Inter Channel Correlation) ICC=1로 설정되어 역상관기(decorrelator)가 적용되지 않으므로 두 개의 채널은 서로 같은 위상 정보를 갖는다.
- [0061] 즉, 스테레오 출력을 고려하여 각각의 CPE에 포함되는 채널쌍을 결정하면, 업믹스된 채널쌍들은 서로 같은 패닝 계수를 갖게 된다(후술함).
- [0062] 하나의 인터널 채널은 하나의 CPE에 포함되는 두 개의 동위상(in-phase) 채널이 믹싱되어 생성된다. 하나의 인터널 채널 신호는 인터널 채널에 포함되는 두개의 입력 채널이, 스테레오 출력 채널로 변환되는 경우의 포맷 변환기 변환 규칙에 따른 믹싱 계인과 EQ(Equalization) 값에 기초하여 다운믹싱된다. 이 때 하나의 CPE에 포함되는 채널쌍은 서로 동위상 채널이므로 다운믹싱 후 채널간 위상을 정합하는 과정이 필요하지 않다.
- [0063] MPS212 업믹서의 스테레오 출력 신호들은 위상차이가 없지만, 도 1 에 개시된 실시예에서는 이를 고려하지 않으므로 복잡도가 불필요하게 증가하게 된다. 재생 레이아웃이 스테레오인 경우, 포맷 변환기의 입력으로 업믹스된 CPE 채널 쌍 대신 하나의 인터널 채널을 이용함으로써, 포맷 변환기의 입력 채널 개수를 줄일 수 있다.
- [0064] 도 2 에 개시된 실시예에서는, CPE 비트스트림(210)을 MPS212 업믹싱하여 두개의 채널을 생성하는 과정 대신, CPE 비트스트림을 인터널 채널 처리(220)하여 하나의 인터널 채널(221)을 생성한다. 이 때, 우퍼 채널은 CPE로 구성되지 않으므로 각각의 우퍼채널 신호가 인터널 채널 신호가 된다.
- [0065] 도 2에 개시된 실시예에서 22.2 채널인 경우를 가정하면 이론적으로 22개의 일반 채널에 대한 11개의 CPE에 대한 인터널 채널과 2개의 우퍼 채널에 대한 인터널 채널을 포함하여 $N_{in}=13$ 개의 인터널 채널이 포맷 변환기의 입력 채널이 된다. 따라서, 포맷 변환기에서는 $13*2$ 다운믹싱이 수행된다.
- [0066] 이와 같이 스테레오 재생 레이아웃인 경우, 인터널 채널을 이용함으로써 MPS212를 통해 업믹싱하고, 포맷 변환을 통해 다시 다운믹싱하는 과정에서 발생하는 불필요한 과정을 추가적으로 제거함으로써 디코더의 복잡도를 보다 더 감소시킬 수 있다.
- [0067] 하나의 CPE에 대한 두 개의 출력 채널 i, j 에 대한 믹싱 매트릭스 $M_{mix}(i,j)$ 값이 1인 경우 채널간 상관도 (ICC, Inter Channel Correlation) $ICC^{i,m}=1$ 로 설정되며, 역상관기(decorrelation) 및 잔여(residual) 처리 단계는 생략될 수 있다.
- [0068] 인터널 채널은 포맷 변환기의 입력에 해당하는 가상의 중간단계 채널로 정의된다. 도 2 에 도시된 것과 같이 각각의 인터널 채널 처리 블록(220)은 CLD와 같은 MPS212 페이로드 및 EQ, 게인값과 같은 렌더링 파라미터들을 이용하여 인터널 채널 신호를 생성한다. 이 때, EQ 및 게인값들은 포맷 변환기의 변환 규칙 테이블에 정의된, MPS212 블록의 출력 채널에 대한 렌더링 파라미터를 의미한다.
- [0069] 표 2 는 22.2 채널 실감 오디오(immersive audio) 신호를 인터널 채널을 스테레오 신호로 렌더링하는 포맷 변환기의 믹싱 매트릭스의 일 실시예를 나타낸다.

표 2

	A	B	C	D	E	F	G	H	I	J	K	L	M
A	1	1	1	1	1	1	1	1	1	1	1	1	1
B	1	1	1	1	1	1	1	1	1	1	1	1	1
C	1	1	1	1	1	1	1	1	1	1	1	1	1
D	1	1	1	1	1	1	1	1	1	1	1	1	1
E	1	1	1	1	1	1	1	1	1	1	1	1	1
F	1	1	1	1	1	1	1	1	1	0	0	0	0
G	1	1	1	1	1	1	1	1	1	0	0	0	0
H	1	1	1	1	1	1	1	1	1	0	0	0	0
I	1	1	1	1	1	1	1	1	1	0	0	0	0
J	1	1	1	1	1	0	0	0	0	1	1	1	1
K	1	1	1	1	1	0	0	0	0	1	1	1	1
L	1	1	1	1	1	0	0	0	0	1	1	1	1
M	1	1	1	1	1	0	0	0	0	1	1	1	1

[0070]

[0071] 표 1 과 마찬가지로, 표 2 의 믹싱 매트릭스에서 가로축과 세로축은 입력 채널의 인덱스를 의미하며 공분산 분석에서 그 순서는 큰 의미를 가지지 않는다.

[0072] 상술한 바와 같이, 믹싱 매트릭스는 대각선을 기준으로 대칭 성질을 가지므로, 표 2 에 개시된 믹싱 매트릭스 역시 대각선을 기준으로 상단 혹은 하단의 구성을 선택함으로써 일부에 대한 공분산 분석을 생략할 수 있다. 또한, 스테레오 출력 채널 레이아웃으로 포맷 변환 과정에서 믹싱되지 않는 입력 채널들에 대해서 역시 공분산 분석을 생략할 수 있다.

[0073] 그러나 표 1 에 개시된 실시예와 달리 표 2 에 개시된 실시예에서는 22개의 일반 채널로 구성되는 11개의 인터널 채널 및 2개의 우퍼 채널을 포함하는 13개의 채널이 스테레오 출력 채널로 다운믹싱되며, 포맷 변환기의 입력 채널 개수 N_{in} 은 13이 된다.

[0074] 그 결과, 표 2 와 같이 인터널 채널을 이용하는 경우의 실시예에서는, 75번의 공분산 분석이 수행되며 이론적으로 19MOPS가 요구되므로 인터널 채널을 이용하지 않는 경우와 비교해 공분산 분석에 의한 포맷 변환기의 부하를 크게 감소시킬 수 있다.

[0075] 포맷 변환기는, 다운믹싱을 위한 다운믹스 매트릭스 M_{Dmx} 가 정의되어 있으며, 믹싱 매트릭스 M_{Mix} 는 다음과 같이 M_{Dmx} 를 이용하여 계산된다.

```

MMix = zero Nin × Nin Matrix
for i = 1 to Nout
  for j = 1 to Nin
    set_i = 0
    if MDmx (i, j) > 0.0
      set_i = 1
    end
    for k = 1 to Nin
      set_k = 0
      if MDmx (i, j) > 0.0
        set_k = 1
      end
      if set_i == 1 and set_k == 1
        MMix (j, k) = 1
      end
    end
  end
end
end
end

```

[0076]

[0077]

[0078]

end

각 OTT 디코딩 블록은 채널 번호 i 및 j에 해당하는 두 개의 채널을 출력하며, 믹싱 매트릭스 $M_{Mix}(i,j)$ 가 1인 경우는 $ICC_{i,m}=1$ 로 설정되어 업믹스 매트릭스 $R_2^{i,m}$ 의 $H11_{OTT}^{i,m}$ 와 $H21_{OTT}^{i,m}$ 가 계산되므로 역상관기를 사용하지 않는다.

표 3 은 본 발명의 일 실시예에 따른, 22.2 채널을 인터널 채널로 구성하기 위한 CPE 구조를 나타낸다.

표 3

Input Channel	Element	Mixing Gain to L	Mixing Gain to R	Internal Channel
CH_M_000	CPE	0.707	0.707	ICH_A
CH_L_000				
CH_U_000	CPE	0.707	0.707	ICH_B
CH_T_000				
CH_M_180	CPE	0.707	0.707	ICH_C
CH_U_180				
CH_LFE2	LFE	0.707	0.707	ICH_D
CH_LFE3	LFE	0.707	0.707	ICH_E
CH_M_L135	CPE	1	0	ICH_F
CH_U_L135				
CH_M_L030	CPE	1	0	ICH_G
CH_L_L045				
CH_M_L090	CPE	1	0	ICH_H
CH_U_L090				
CH_M_L060	CPE	1	0	ICH_I
CH_U_L045				
CH_M_R135	CPE	0	1	ICH_J
CH_U_R135				
CH_M_R030	CPE	0	1	ICH_K
CH_L_R045				
CH_M_R090	CPE	0	1	ICH_L
CH_U_R090				
CH_M_R060	CPE	0	1	ICH_M
CH_U_R045				

[0079]

[0080]

22.2 채널 비트스트림이 표 3 과 같은 구조를 가지면, 13개의 인터널 채널들은 ICH_A 부터 ICH_M까지 정의될 수

있으며, 13개의 인터널 채널들에 대한 믹싱 매트릭스는 표 2와 같이 결정될 수 있다.

- [0081] 표 3의 첫번째 열은 입력 채널에 대한 인덱스를 나타내며, 첫번째 행은 입력 채널이 CPE를 구성하는지 여부, 스테레오 채널로의 믹싱 게인 및 인터널 채널 인덱스를 나타낸다.
- [0082] 예를 들어 CM_M_000과 CM_L_000이 하나의 CPE로 구성되는 ICH_A 인터널 채널의 경우, 이 CPE를 스테레오 출력 채널로 업믹싱하기 위해 좌 출력 채널에 적용되는 믹싱 게인과 우 출력 채널에 적용되는 믹싱 게인은 모두 0.707의 값을 갖는다. 즉, 좌 출력 채널과 우 출력 채널로 업믹싱된 신호는 서로 동일한 크기로 재생된다.
- [0083] 또는, CH_M_L135와 CH_U_L135가 하나의 CPE로 구성되는 ICH_F 인터널 채널의 경우, 이 CPE를 스테레오 출력 채널로 업믹싱하기 위해 좌 출력 채널에 적용되는 믹싱 게인은 1, 우 출력 채널에 적용되는 믹싱 게인은 0의 값을 갖는다. 즉, 모든 신호는 좌 출력 채널로만 재생되며 우 출력 채널로는 재생되지 않는다.
- [0084] 반대로, CH_M_R135와 CH_U_R135가 하나의 CPE로 구성되는 ICH_J 인터널 채널의 경우, 이 CPE를 스테레오 출력 채널로 업믹싱하기 위해 좌 출력 채널에 적용되는 믹싱 게인은 0, 우 출력 채널에 적용되는 믹싱 게인은 1의 값을 갖는다. 즉, 모든 신호는 좌 출력 채널로는 재생되지 않고 우 출력 채널로만 재생된다.
- [0085] 도 3은 하나의 CPE로부터 하나의 인터널 채널을 생성하는 장치의 일 실시예를 나타낸다.
- [0086] 하나의 CPE에 대한 인터널 채널은 다운믹스된 모노신호에 CLD와 게인 및 EQ와 같은 QMF 도메인의 포맷 변환 파라미터들을 적용함으로써 유도될 수 있다.
- [0087] 도 3에 개시된 인터널 채널을 생성하는 장치는, 업믹서(310), 스케일러(320) 및 믹서(330)를 포함한다.
- [0088] CH_M_000과 CH_L_000의 채널쌍에 대한 신호가 다운믹스된 CPE(340)가 입력되는 경우를 가정하면, 업믹서(310)는 CLD 파라미터를 이용하여 CPE 신호를 업믹싱한다. 업믹서(310)를 통과한 CPE 신호는 CH_M_000에 대한 신호(351)와 CH_L_000에 대한 신호(352)로 업믹스되며, 업믹스된 신호들의 위상 역시 같게 유지되며 포맷 변환기에서 함께 믹스될 수 있다.
- [0089] 업믹스된 CH_M_000 채널 신호 및 CH_L_000 채널 신호 각각은 포맷 변환기에 정의되어 있는 변환 규칙에 해당하는 게인 및 EQ에 의해 각 서브밴드별로 스케일링(320, 321)된다.
- [0090] CH_M_000과 CH_L_000의 채널 쌍에 대해 각각 스케일링된 신호들(361, 362)이 생성되면, 믹서(330)는 스케일링된 신호들(361, 362)을 믹싱하고, 믹싱된 신호를 파워 정규화함으로써 포맷 변환을 위한 중간단계 채널 신호(intermediate channel signal)인 인터널 채널 신호 ICH_A(370)를 생성한다.
- [0091] 이 때, CLD를 이용하여 업믹스되지 않는, SCE(Single Channel Element) 및 우퍼채널 등의 경우, 인터널 채널은 원래의 입력 채널과 동일하다.
- [0092] 인터널 채널을 이용하는 코어 코덱 출력은 하이브리드 QMF(Quadrature Mirror Filter) 도메인에서 수행되므로, ISO IEC23308-3 10.3.5.2의 과정은 처리되지 않는다. 코어 코더의 각 채널을 할당하기 위하여, 표 4 내지 표 6과 같은 추가적인 채널 할당 규칙 및 다운믹스 규칙이 정의된다.
- [0093] 표 4는 본 발명의 일 실시예에 따른, 디코더 입력 채널에 대응되는 인터널 채널들의 타입을 나타낸다.

표 4

Type	Channels	Panning (L,R)
CH-I-LFE	CH_LFE1, CH_LFE2, CH_LFE3	(0.707, 0.707)
CH-I-CNTR	CH_M_000, CH_L_000, CH_U_000, CH_T_000, CH_M_180, CH_U_180	(0.707, 0.707)
CH-I-LEFT	CH_M_L022, CH_M_L030, CH_M_L045, CH_M_L060, CH_M_L090, CH_M_L110, CH_M_L135, CH_M_L150, CH_L_L045, CH_U_L045, CH_U_L030, CH_U_L045, CH_U_L090, CH_U_L110, CH_U_L135, CH_M_LSCR, CH_M_LSCH	(1, 0)
CH-R-RIGHT	CH_M_R022, CH_M_R030, CH_M_R045, CH_M_R060, CH_M_R090, CH_M_R110, CH_M_R135, CH_M_R150, CH_L_R045, CH_U_R045, CH_U_R030, CH_U_R045, CH_U_R090, CH_U_R110, CH_U_R135, CH_M_RSCR, CH_M_RSCH	(0,1)

[0094]

[0095]

인터널 채널은 코어 코더와 포맷 변환기의 입력채널 사이의 중간단계 채널에 해당하며 우퍼채널, 센터 채널, 좌측 채널 및 우측 채널의 네가지 타입이 있다.

[0096]

CPE로 표현된 각 타입의 채널 쌍들이 동일한 인터널 채널 타입일 경우, 포맷 변환기에서 동일한 패닝 계수와 믹싱 매트릭스를 가지므로 인터널 채널을 이용할 수 있다. 즉, CPE에 포함된 채널 쌍이 같은 인터널 채널 타입을 갖는 경우 인터널 채널 처리가 가능하며 따라서 CPE를 구성할 때 같은 인터널 채널 타입을 갖는 채널들로 CPE를 구성하도록 할 필요가 있다.

[0097]

디코더 입력 채널이 우퍼 채널, 즉 CH_LFE1, CH_LFE2 또는 CH_LFE3에 해당하면, 인터널 채널 타입은 우퍼 채널인 CH_I_LFE로 결정된다.

[0098]

디코더 입력 채널이 센터 채널, 즉 CH_M_000, CH_L_000, CH_U_000, CH_T_000, CH_M_180 또는 CH_U_180에 해당하면, 인터널 채널 타입은 센터 채널인 CH_I_CNTR로 결정된다.

[0099]

디코더 입력 채널이 좌측 채널, 즉 CH_M_L022, CH_M_L030, CH_M_L045, CH_M_L060, CH_M_L090, CH_M_L110,

CH_M_L135, CH_M_L150, CH_L_L045, CH_U_L045, CH_U_L030, CH_U_L045, CH_U_L090, CH_U_L110, CH_U_L135, CH_M_LSCR 또는 CH_M_LSCH에 해당하면, 인터널 채널 타입은 좌측 채널인 CH_I_LEFT로 결정된다.

[0100] 디코더 입력 채널이 우측 채널, 즉 CH_M_R022, CH_M_R030, CH_M_R045, CH_M_R060, CH_M_R090, CH_M_R110, CH_M_R135, CH_M_R150, CH_L_R045, CH_U_R045, CH_U_R030, CH_U_R045, CH_U_R090, CH_U_R110, CH_U_R135, CH_M_RSCR 또는 CH_M_RSCH에 해당하면, 인터널 채널 타입은 우측 채널인 CH_I_RIGHT로 결정된다.

[0101] 표 5 는 본 발명의 일 실시예에 따른, 인터널 채널 타입에 따라 추가적으로 정의되는 채널의 위치를 나타낸다.

표 5

LoudspeakerGeometry as defined in ISO /IEC 23001-8)	Channel	Azimuth [deg]	Elevation [deg]	Azimuth start angle of sector [deg]	Azimuth end angle of sector [deg]	Elevation start angle of sector [deg]	Elevation end angle of sector [deg]	Ch. is LFE	Position is relative
43	CH_I_CNTR	0	0	0	0	0	0	0	0
44	CH_I_LFE	0	n/a	n/a	n/a	n/a	n/a	1	0
45	CH_I_LEFT	30	0	30	30	0	0	0	0
46	CH_I_RIGHT	-30	0	-30	-30	0	0	0	0

[0102]

[0103] CH_I_LFE는 우측채널로 고도각 0도에 위치하고 CH_I_CNTR 은 고도각 및 방위각이 모두 0도에 위치하는 채널에 해당한다. CH_I_LFE는 고도각은 0도이고 방위각은 좌측 30도~60도 사이의 섹터에 위치하는 채널에 해당하며, CH_I_RIGHT 는 고도각은 0도이고 방위각은 우측 30도~60도 사이의 섹터에 위치하는 채널에 해당한다.

[0104] 이 때 새로 정의된 인터널 채널들의 위치는 채널간 상대적 위치가 아닌 기준점에 대한 절대적인 위치이다.

[0105] CPE 쌍으로 구성된 QCE(Quadruple Channel Element)의 경우에도 인터널 채널이 적용될 수 있다(후술함).

- [0106] 인터널 채널을 생성하는 구체적인 방법은 두가지로 구현될 수 있다.
- [0107] 첫번째는 MPEG-H 3D 오디오 인코더에서 전처리(pre-processing) 하는 방법이고, 두번째는 MPEG-H 3D 오디오 디코더에서 후처리(post-processing) 하는 방법이다.
- [0108] 인터널 채널이 MPEG에서 이용되는 경우, 표 5 는 ISO/IEC 23008-3 표 90에 새로운 행으로 추가될 수 있다.
- [0109] 표 6 은 본 발명의 일 실시예에 따른, 인터널 채널 타입에 대응되는 포맷 변환기 출력 채널 및 각 출력 채널에 적용될 게인과 EQ 인덱스를 나타낸다.

표 6

Source	Destination	Gain	EQ_index
CH_I_CNTR	CH_M_L030, CH_M_R030	1.0	0 (off)
CH_I_LFE	CH_M_L030, CH_M_R030	1.0	0 (off)
CH_I_LEFT	CH_M_L030	1.0	0 (off)
CH_I_RIGHT	CH_M_L030	1.0	0 (off)

- [0110]
- [0111] 인터널 채널을 이용하기 위하여, 포맷 변환기에는 표 6 과 같은 추가적인 규칙이 추가되어야 한다.
- [0112] 인터널 채널 신호는 포맷 변환기의 게인 및 EQ 값들을 고려하여 생성된다. 따라서, 표 6에 나타난 것과 같이 게인 값은 1이고, EQ 인덱스는 0인, 추가적인 변환 규칙을 이용하여 인터널 채널 신호를 생성할 수 있다.
- [0113] 인터널 채널 타입이 센터 채널에 해당하는 CH_I_CNTR 채널이거나, 우퍼 채널에 해당하는 CH_I_LFE이면, 출력 채널은 CH_M_L030 및 CH_M_R030이 된다. 이 때, 게인값은 1로, EQ 인덱스는 0으로 결정되며 두개의 스테레오 출력 채널을 모두 이용하므로 출력 신호의 파워를 유지하기 위해 각 출력 채널 신호에 $1/\sqrt{2}$ 를 곱해야 한다.
- [0114] 인터널 채널 타입이 좌측 채널에 해당하는 CH_I_LEFT이면, 출력 채널은 CH_M_L030이 된다. 이 때, 게인값은 1로, EQ 인덱스는 0으로 결정되며 좌측의 출력 채널만을 이용하므로, CH_M_L030에는 게인 1이 적용되고 CH_M_R030에는 게인 0이 적용된다.
- [0115] 인터널 채널 타입이 우측 채널에 해당하는 CH_I_RIGHT이면, 출력 채널은 CH_M_R030이 된다. 이 때, 게인값은 1로, EQ 인덱스는 0으로 결정되며 우측의 출력 채널만을 이용하므로 CH_M_R030에는 게인 1이 적용되고 CH_M_L030에는 게인 0이 적용된다.
- [0116] 이 때, 인터널 채널과 입력 채널이 동일한 SCE 채널등의 경우, 일반적인 포맷 변환 규칙이 적용된다.
- [0117] 인터널 채널이 MPEG에서 이용되는 경우, 표 6 은 ISO/IEC 23008-3 표 96에 새로운 행으로 추가될 수 있다.
- [0118] 표 7 내지 표 15 는 MPEG에서 인터널 채널을 사용하기 위해 기존 규격이 변경되어야 하는 부분을 나타낸다.
- [0119] 표 7 은 본 발명의 일 실시예에 따른, ICGConfig의 선택스를 나타낸다.

표 7

Syntax	No. of bits	Mnemonic
<pre> ICGConfig () { if (ICGDisabledPresent) { for (elemIdx=0; elemCPE=0; elemIdx<numElements; ++elemIdx) { if (usacElementType[elemIdx] == ID_USAC_CPE) { ICGDisabledCPE[elemCPE]; elemCPE++; } } } if (ICGPreAppliedPresent) { for (elemIdx=0; elemCPE=0; elemIdx<numElements; ++elemIdx) { if (usacElementType[elemIdx] == ID_USAC_CPE) { ICGPreAppliedCPE[elemCPE]; elemCPE++; } } } } </pre>	1	Uimbsf

[0120]

[0121]

[0122]

[0123]

[0124]

[0125]

[0126]

[0127]

표 7 에 도시된 ICGconfig는, 인터널 채널 처리 블록에서 처리되어야 하는 프로세스의 타입들을 정의한다.

ICGDisabledPresent는 CPE들에 대한 적어도 하나의 인터널 채널 처리가 채널 할당을 이유로 사용되지 않는지 (disable) 여부를 나타낸다. 즉, 적어도 하나의 ICGDisabledCPE가 1의 값을 갖는지 여부를 나타내는 인디케이터이다.

ICGDisabledCPE는 CPE들에 대한 각각의 인터널 채널 처리가 채널 할당을 이유로 사용되지 않는지 여부를 나타낸다. 즉, 각각의 CPE가 인터널 채널을 사용하는지 여부를 나타내는 인디케이터이다.

ICGPreAppliedPresent는 적어도 하나의 CPE가 인터널 채널 게인을 고려하여 인코딩 되었는지 여부를 나타낸다.

ICGPreAppliedCPE는 각각의 CPE가 인터널 채널 게인을 고려하여 인코딩 되었는지 여부, 즉 인터널 채널 게인이 인코더에서 전처리되었는지 여부를 나타내는 인디케이터이다.

각각의 CPE에 대하여, ICGAppliedPresent가 1로 설정되어 있으면, ICGPreAppliedCPE의 1비트 플래그인 ICGPreAppliedCPE를 읽는다. 즉, 각각의 CPE에 인터널 채널 게인(ICG, Internal Channel Gain)이 적용되어야 하는지 여부를 확인하고, 적용되어야 한다면 전처리 된 것인지 여부를 확인하여 인코더에서 전처리 되었다면 디코더에서 인터널 채널 게인을 적용하지 않고, 인코더에서 전처리되지 않았다면 디코더에서 인터널 채널 게인을 적용하는 것이다.

실감 오디오 입력 신호가 CPE 또는 QCE를 이용해 MPS212 인코딩 되고 출력 레이아웃이 스트레오이면, 포맷 변환기 입력 채널 개수를 감소시키기 위해 코어 코덱 디코더 내부에서 인터널 채널 신호가 생성된다. 이 때, ICGDisabledCPE 가 1로 설정된 CPE에 대해서는 인터널 채널 신호 생성이 생략된다. 인터널 채널 처리는 디코딩 된 모노 신호에 인터널 채널 게인을 곱하는 과정에 해당하며, 인터널 채널 게인은 CLD와 포맷 변환 파라미터로부터 계산된다.

- [0128] ICGDisabledCPE[n] 은 n번째 CPE가 인터널 채널 처리가 가능한지 여부를 나타낸다. n번째 CPE 채널 쌍에 포함되는 두 개의 채널이 표 4에 정의된 같은 채널 그룹에 속하는 경우, 해당 CPE는 인터널 채널 처리가 가능하며 ICGDisabledCPE[n]은 0으로 설정된다.
- [0129] 예를 들어, 입력 채널 중 CH_M_L060 과 CH_T_L045가 하나의 CPE로 구성된 경우, 두 채널은 같은 채널 그룹에 속하므로 ICGDisabledCPE[n]는 0으로 설정되며 CH_I_LEFT의 인터널 채널이 생성될 수 있다. 반면, 입력 채널 중 CH_M_L060 and CH_M_000이 하나의 CPE로 구성된 경우, 두 채널은 서로 다른 채널 그룹에 속하므로 ICGDisabledCPE[n]는 1로 설정되며, 인터널 채널 처리가 수행되지 않는다.
- [0130] CPE 쌍으로 구성된 QCE(Quadruple Channel Element)의 경우, (1) QCE가 하나의 그룹에 속하는 네개의 채널로 구성된 경우 또는 (2) QCE가 하나의 그룹에 속하는 두개의 채널과 또 다른 그룹에 속하는 두개의 채널로 구성된 경우 인터널 채널 처리가 가능하며 ICGDisableCPE[n] 와 ICGDisableCPE[n+1]이 모두 0으로 설정된다.
- [0131] (1)의 예를 들면, QCE가 CH_M_000, CH_L_000, CH_U_000 및 CH_T_000 네개의 채널로 구성된 경우, 인터널 채널 처리가 가능하며 인터널 채널 타입은 CH_I_CNTR이 된다. (2)의 예를 들면, QCE가 CH_M_L060, CH_U_L045, CH_M_R060 및 CH_U_R045 네 개의 채널로 구성된 경우, 인터널 처리가 가능하며 인터널 채널 타입은 CH_I_LEFT 및 CH_I_RIGHT이 된다.
- [0132] (1) 또는 (2) 경우를 제외하면, 해당 QCE를 구성하는 CPE 쌍에 대한 ICGDisableCPE[n] 와 ICGDisableCPE[n+1]이 모두 1로 설정되어야 한다.
- [0133] 인터널 채널 게인이 인코더에서 적용되면 인터널 채널 게인이 디코더에서 적용되는 경우와 비교해 디코더에 요구되는 복잡도를 감소시킬 수 있다.
- [0134] ICGConfig의 ICGPreAppliedCPE[n]는 n번째 CPE가 인코더에서 인터널 채널 게인이 적용되었는지 여부를 나타낸다. 만일 ICGPreAppliedCPE[n]가 참이라면, 디코더의 인터널 채널 처리 블록은 n번째 CPE의 스테레오 재생을 위한 다운믹스 신호를 바이패스(bypass)한다. 반면, ICGPreAppliedCPE[n]가 거짓이라면, 다운믹스 신호는 디코더의 인터널 채널 처리 블록은 다운믹스 신호에 인터널 채널 게인을 적용한다.
- [0135] ICGDisableCPE[n]가 1이면 해당 QCE 또는 CPE를 위한 인터널 채널 게인의 계산이 불가능하므로 ICGPreApplied[n]은 0으로 설정된다. CPE 쌍으로 구성되는 QCE에 대해서는, QCE에 포함되는 각각의 CPE에 대한 인덱스 ICGPreApplied[n] 와 ICGPreApplied[n+1]가 같은 값을 가져야 한다.
- [0136] 이하에서는 표 8 내지 표 16 을 이용해 인터널 채널 처리를 위해 변경 또는 추가되어야 하는 비트스트림 구성 및 선택스를 설명한다.
- [0137] 표 8 은 본 발명의 일 실시예에 따른, mpeg3daExtElementConfig()의 선택스를 나타낸다.

표 8

Syntax	No. of bits	Mnemonic
<pre> mpegh3daExtElementConfig() { usacExtElementType = escapedValue(4, 8, 16); usacExtElementConfigLength = escapedValue(4, 8, 16); if (usacExtElementDefaultLengthPresent) { usacExtElementDefaultLength = escapedValue(8, 16, 0) + 1; } else { usacExtElementDefaultLength = 0; } usacExtElementPayloadFrag; switch (usacExtElementType) { case ID_EXT_ELE_FILL: /* No configuration element */ break; case ID_EXT_ELE_MPEGS: SpatialSpecificConfig(); break; case ID_EXT_ELE_SAOC: SAOCSpecificConfig(); break; case ID_EXT_ELE_AUDIOFREROLL: /* No configuration element */ break; case ID_EXT_ELE_UNI_DRC: mpegH3daUniDrcConfig(); break; case ID_EXT_ELE_OBJ_METADATA: ObjectMetadataConfig(); break; case ID_EXT_ELE_SAOC_3D: SAOC3DSpecificConfig(); break; case ID_EXT_ELE_HOA: HOACConfig(); break; case ID_EXT_ELE_FMT_CNRVTR: /* No configuration element */ break; case ID_EXT_ELE_ICG: ICGConfig(); break; default: while (usacExtElementConfigLength-- > 0) { tmp; } break; } } </pre>	<p>1</p> <p>1</p> <p>NOTE</p> <p>8</p>	<p>uimsbf</p> <p>uimsbf</p> <p>uimsbf</p>
<p>NOTE: The default entry for the usacExtElementType is used for unknown extElementTypes so that legacy decoders can cope with future extensions.</p>		

[0138]

[0139]

표 8 의 mpegH3daExtElementConfig()에서의 예시와 같이 Configuration 과정에서 ICGConfig()를 호출하여 각 표 7과 같은 Internal Channel 사용 여부 및 ICG 적용 여부를 획득할 수 있다.

[0140]

표 9 는 본 발명의 일 실시예에 따른, usacExtElementType를 나타낸다.

표 9

usacExtElementType	Value
ID_EXT_ELE_FILL	0
ID_EXT_ELE_MPEGS	1
ID_EXT_ELE_SAOC	2
ID_EXT_ELE_AUDIOPREROLL	3
ID_EXT_ELE_UNI_DRC	4
ID_EXT_ELE_OBJ_METADATA	5
ID_EXT_ELE_SAOC_3D	6
ID_EXT_ELE_HOA	7
ID_EXT_ELE_FMT_CNVTRTR	8
ID_EXT_ELE_ICG	9
/*reserved for ISO use*/	10-127
/*reserved for use outside of ISO scope*/	128 and higher

NOTE : Application-specific usacExtElementType values are mandated to be in the space reserved for use outside of ISO scope. These are skipped by a decoder as a minimum of structure is required by the decoder to skip these extensions.

[0141]

[0142]

표 9 에 나타난 바와같이, usacExtElementType에서는 인터널 채널 처리를 위해 ID_EXT_ELE_ICG가 추가되고, 그 값은 9가 될 수 있다.

[0143]

표 10 은 본 발명의 일 실시예에 따른, speakerLayoutType을 나타낸다.

표 10

Value	Meaning
0	Loudspeaker layout is signaled by means of ChannelConfiguration index as defined in ISO/IEC 23001-8.
1	Loudspeaker layout is signaled by means of a list of LoudspeakerGeometry indices as defined in ISO/IEC 23001-8
2	Loudspeaker layout is signaled by means of a list of explicit geometric position information.
3	Loudspeaker layout is signaled by means of LCChannelConfiguration index. Note that the LCChannelConfiguration has same layout with ChannelConfiguration but different channel orders to enable the optimal internal channel structure using CPE.

[0144]

[0145]

인터널 채널 처리를 위해서는 인터널 채널을 위한 스피커 레이아웃 타입 speakerLayoutType 이 정의되어야 한다. 표 10은 의 speakerLayoutType각 값에 대한 의미를 나타낸다.

[0146]

speakerLayoutType==3인 경우, 라우드 스피커 레이아웃은 LCChannelConfiguration 인덱스의 의미에 의해 시그널링된다. LCChannelConfiguration는 ChannelConfiguration과 동일한 레이아웃을 가지지만, CPE를 이용하는 최적의 인터널 채널 구조를 가능하기 위한 채널 할당 순서(order)를 가진다.

[0147]

표 11 은 본 발명의 일 실시예에 따른, SpeakerConfig3d()의 선택스를 나타낸다.

표 11

Syntax	No. of bits	Mnemonic
SpeakerConfig3d()		
{		
speakerLayoutType;	2	uimsbf
if (speakerLayoutType == 0 speakerLayoutType == 3){		
CICPSpeakerLayoutIdx;	6	uimsbf
}		
else {		
numSpeakers = escapedValue(5, 8, 16) + 1;		
if (speakerLayoutType == 1){		
for (i = 0; i < numSpeakers; i++){		
CICPSpeakerIdx;	7	uimsbf
}		
}		
if (speakerLayoutType == 2){		
mpeg3daFlexibleSpeakerConfig(numSpeakers);		
}		
}		
}		

[0148]

[0149] 상술한 바와 같이 speakerLayoutType==3인 경우, CICPSpeakerLayoutIdx와 동일한 레이아웃을 이용하지만 인터널 채널을 위한 최적화된 채널 할당 순서(ordering)에서 차이가 있다.

[0150] speakerLayoutType==3이고 출력 레이아웃이 스테레오인 경우, 입력 채널 번호 Nin 는 코어 코덱 이후 인터널 채널의 번호로 변경된다.

[0151] 표 12 는 본 발명의 일 실시예에 따른, immersiveDownmixFlag를 나타낸다.

표 12

immersiveDownmixFlag	Meaning
0	Generic format converter shall be applied as defined in clause 10.
1	If the local loudspeaker setup, signaled by LoudspeakerRendering(), is signaled as (speakerLayoutType==0 or 3,CICPSpeakerLayoutIdx==5) or as (speakerLayoutType==0 or 3,CICPSpeakerLayoutIdx==6), independently of potentially signaled loudspeaker displacement angles, then immersive rendering format converter shall be applied as defined in clause 11. In all other case the generic format converter shall be applied as defined in clause 10.

[0152]

[0153] 인터널 채널을 위한 스피커 레이아웃 타입을 새로 정의함으로써 immersiveDownmixFlag 역시 수정되어야 한다. immersiveDownmixFlag가 1인 경우, 표 12와 같이 speakerLayoutType==3인 경우의 처리를 위한 구문이 추가되어야 한다.

[0154] 오브젝트 확산(object spreading)은 다음과 같은 요건이 만족되어야 한다.

[0155] - 로컬 라우드스피커 설정은 LoudspeakerRendering()에 의해 시그널링되며,

[0156] - the speakerLayoutType은 0 또는 3이어야 하며,

[0157] - CICPSpeakerLayoutIdx는 4, 5, 6, 7, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 중 하나의 값을 가진다.

[0158] 표 13 은 본 발명의 일 실시예에 따른, SAOC3DgetNumChannels()의 신택스를 나타낸다.

[0159] SAOC3DgetNumChannels 은 표 13 에 나타난 것처럼 speakerLayoutType==3인 경우를 포함하도록 수정되어야 한다.

표 13

Syntax	No. of bits	Mnemonic
SAOC3DgetNumChannels(Layout)		Note 1
{		
numChannels = numSpeakers;		Note 2
for (i = 0; i < numSpeakers; i++){		
if (Layout.isLFE[i] == 1){		
numChannels = numChannels - 1;		
}		
}		
return numChannels;		
}		
Note 1: The function SAOC3DgetNumChannels() returns the number of available non-LFE channels numChannels.		
Note 2: numSpeakers is defined in Syntax of SpeakerConfig3d(). If speakerLayoutType == 0 or speakerLayoutType == 3 numSpeakers represents the number of loudspeakers corresponding to the ChannelConfiguration value, CIPsSpeakerLayoutIdx, as defined in ISO/IEC 23001-8.		

[0160]

[0161] 표 14 는 본 발명의 일 실시예에 따른, 채널 할당 순서를 나타낸다.

[0162] 표 14 는 인터널 채널을 위해 새로 정의되는 채널 할당 순서로, 라우드스피커 레이아웃 또는 LCChannelConfiguration에 따른 채널 개수, 순서 및 가능한 인터널 채널 타입을 나타낸다.

표 14

Loudspeaker Layout Index or LCChannelConfiguration	Number of Channels	Channels (with ordering)	Possible Internal Channel Type
1	1	CH_M_000	Center
2	2	CH_M_L030, CH_M_R030	Left Right
3	3	CH_M_000, CH_M_L030, CH_M_R030	Center Left Right
4	4	CH_M_000, CH_M180, CH_M_L030, CH_M_R030	Center Left Right
5	5	CH_M_000, CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	Center Left Right
6	6	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	Center Left Left Right
7	8	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_L060, CH_M_R030, CH_M_R110, CH_M_R060	Center Left Left Right
8		n.a.	
9	3	CH_M_180, CH_M_L030, CH_M_R030	Center Left Right
10	4	CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	Left Right
11	7	CH_M_000, CH_M_180, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_R030, CH_M_R110	Center Left Left Right
12	8	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_L135, CH_M_R030, CH_M_R110, CH_M_R135	Center Left Left Right
13	24	CH_M_000, CH_L_000, CH_U_000, CH_T_000, CH_M_180, CH_T_180, CH_LFE2, CH_LFE3, CH_M_L135, CH_U_L135, CH_M_L030, CH_L_L045, CH_M_L090, CH_U_L090, CH_M_L060, CH_U_L045, CH_M_R135, CH_U_R135, CH_M_R030, CH_L_R045, CH_M_R090, CH_U_R090, CH_M_R060, CH_U_R045	Center Left Left Right

[0163]

14	8	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_U_L030, CH_M_R030, CH_M_R110, CH_U_R030	Center Left Left Right
15	12	CH_M_000, CH_U_180, CH_LFE2, CH_LFE3, CH_M_L030, CH_M_L135, CH_M_L090, CH_U_L045, CH_M_R030, CH_M_R135, CH_M_R090, CH_U_R045	Center Left Left Right
16	10	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_U_L030, CH_U_L110, CH_M_R030, CH_M_R110, CH_U_R030, CH_U_R110	Center Left Left Right
17	12	CH_M_000, CH_U_000, CH_T_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_U_L030, CH_U_L110, CH_M_R030, CH_M_R110, CH_U_R030, CH_U_R110	Center Left Left Right
18	14	CH_M_000, CH_U_000, CH_T_000, CH_LFE1, CH_M_L030, CH_M_L110, CH_M_L150, CH_U_L030, CH_U_L110, CH_M_R030, CH_M_R110, CH_M_R150, CH_U_R030, CH_U_R110	Center Left Left Right
19	12	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L135, CH_M_L090, CH_U_L030, CH_U_L135, CH_M_R030, CH_M_R135, CH_M_R090, CH_U_R030, CH_U_R135	Center Left Left Right
20	14	CH_M_000, CH_LFE1, CH_M_L030, CH_M_L135, CH_M_L090, CH_U_L045, CH_U_L135, CH_M_LSCR, CH_M_R030, CH_M_R135, CH_M_R090, CH_U_R045, CH_U_R135, CH_M_RSCR	Center Left Left Right

[0164]

[0165]

표 15는 본 발명의 일 실시예에 따른, mpeg3daChannelPairElementConfig()의 선택스를 나타낸다.

[0166]

인터널 채널 처리를 위해, 표 15에 개시된 것과 같이 stereoConfigIndex가 0보다 큰 경우 Mps212Config()처리를 한 후 isInternal Channel Processed()가 처리되도록 mpeg3daChannelPairElementConfig ()가 수정되어야 한다.

표 15

Syntax	No. of bits	Mnemonic
<pre> mpeg3daChannelPairElementConfig(sbrRatioIndex) { mpeg3daCoreConfig(); if (enhancedNoiseFilling) { igflIndependentTiling; } if (sbrRatioIndex > 0) { SbrConfig(); stereoConfigIndex; } else { stereoConfigIndex = 0; } if (stereoConfigIndex > 0) { Mps212Config(stereoConfigIndex); } } </pre>		
isInternalChannelProcessed	1	uimsbf
qcelIndex;	2	uimsbf
if(qcelIndex > 0) {		
shiftIndex0;	1	uimsbf
if(shiftIndex0 > 0) {		
shiftChannel0;	nBits¹⁾	
}		
shiftIndex1;	1	uimsbf
if(shiftIndex1 > 0) {		
shiftChannel1;	nBits¹⁾	
}		
}		
¹⁾ nBits = floor(log2(numAudioChannels + numAudioObjects + numHOATransportChannels + numSAOCTransportChannels - 1)) + 1		

[0167]

[0168]

도 4 는 본 발명의 일 실시예에 따른, 디코더에서 인터널 채널 신호에 인터널 채널 게인 적용부의 세부 블록도이다.

[0169]

speakerLayout==3이고, isInternalProcessed가 0이고, 재생 레이아웃이 스테레오인 조건이 만족되어 Internal Channel Gain을 Decoder에서 적용시키는 경우, 도 4 와 같은 인터널 채널 처리 과정이 수행된다.

[0170]

도 4 에 개시된 인터널 채널 게인 적용부는 인터널 채널 게인 획득부(410) 및 곱셈기(420)를 포함한다.

[0171]

입력되는 CPE가 CH_M_000과 CH_L_000의 채널쌍으로 구성되는 경우를 가정하면, 해당 CPE에 대한 모노 QMF 서브밴드 샘플들(430)이 입력되면 인터널 채널 게인 획득부(410)는 CLD를 이용하여 인터널 채널 게인을 획득한다. 곱셈기(420)는 획득된 인터널 채널 게인을 수신된 모노 QMF 서브밴드 샘플에 곱해줌으로써 인터널 채널 신호 ICH_A(440)를 획득한다.

[0172]

인터널 채널 신호는 CPE에 대한 모노 QMF 서브밴드 샘플들에 인터널 채널 게인 $G_{ICH}^{l,m}$ 을 곱함으로써 간단히 재구성될 수 있다. 이 때, l은 시간 인덱스 m은 주파수 인덱스를 나타낸다.

[0173]

인터널 채널 게인 $G_{ICH}^{l,m}$ 은 [수학식 1]과 같이 정의된다.

수학식 1

$$G_{ICH}^{l,m} = \sqrt{\frac{(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m)^2 + (c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m)^2}{(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m + c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m)^2}} \times c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m + c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m$$

[0174]

이 때, $c_{left}^{l,m}$ 및 $c_{right}^{l,m}$ 는 CLD의 패닝 계수를, G_{left} 및 G_{right} 는 포맷 변환 규칙에 정의된 게인을, $G_{EQ,left}^m$ 및 $G_{EQ,right}^m$ 는 포맷 변환 규칙에 정의된 EQ의 m번째 밴드의 게인을 의미한다.

[0176]

도 5 는 본 발명의 일 실시예에 따른, 인코더에서 인터널 채널 게인이 전처리(pre-processed)되는 경우의 디코딩 블록도이다.

- [0177] speakerLayout==3이고, isInternalProcessed가 1이고, 재생 레이아웃이 스테레오인 조건이 만족되어 Internal Channel Gain이 Encoder에서 적용되어 전달되는 경우, 도 5 와 같은 인터널 채널 처리 과정이 수행된다.
- [0178] 출력 레이아웃이 스테레오인 경우라면, MPEG-H 3D 오디오 인코더에서 CPE에 해당하는 인터널 채널 게인을 미리 처리함으로써 디코더에서 MPS212가 바이패스될 수 있으므로 디코더 복잡도를 더 감소시킬 수 있다.
- [0179] 그러나 출력 레이아웃이 스테레오가 아닌 경우라면 인터널 채널 처리를 수행하지 않으므로, 디코더에서 도 5와 같이 인터널 채널 게인의 역수 $\frac{1}{G_{ICH}^{i,m}}$ 를 곱하고 MPS212 처리하여 원복하는 과정이 필요하다.
- [0180] 도 3 및 도 4와 마찬가지로 입력되는 CPE가 CH_M_000과 CH_L_000의 채널쌍으로 구성되는 경우를 가정한다. 인코더에서 인터널 채널 게인이 전처리된 모노 QMF 서브밴드 샘플들(540)이 입력되면, 디코더는 출력 레이아웃이 스테레오인지 여부를 판단(510)한다.
- [0181] 출력 레이아웃이 스테레오라면, 인터널 채널을 사용하는 경우이므로 인터널 채널 ICH_A(550)에 대한 인터널 채널 신호로 수신된 모노 QMF 서브밴드 샘플들(540)를 출력한다. 반면 출력 레이아웃이 스테레오가 아니라면, 인터널 채널 처리가 인터널 채널을 사용하지 않으므로, 역 인터널 채널 게인(inverse internal channel gain)처리(520)를 수행하여 인터널 채널 처리된 신호를 원복(560)하고, 원복된 신호를 MPS212 업믹싱(530)하여 CH_M_000(571) 및 CH_L_000(572) 각각에 대한 신호를 출력한다.
- [0182] 포맷 변환기의 공분산 분석에 의한 부하가 문제되는 경우는 입력 채널의 개수에 비해 많고 출력 채널의 개수가 적은 경우로, MPEG-H 오디오에서 출력 레이아웃이 스테레오인 경우 가장 큰 디코딩 복잡도를 갖는다.
- [0183] 반면 스테레오가 아닌 다른 출력 레이아웃인 경우 인터널 채널 게인의 역수를 곱하기 위해 추가되는 연산량은, 프레임당 두 세트의 CLD인 경우를 가정하면 (곱셈 5번, 덧셈 2번, 나눗셈 1번, 제곱근 1번 ≈ 55 연산) X(71밴드) X (2 파라미터 세트) X(48000/2048) X(13 인터널 채널)로, 대략 2.4 MOPS가 되며 시스템에 큰 부하로 작용하지 않는다.
- [0184] 인터널 채널이 생성된 후, 인터널 채널의 QMF 서브밴드 샘플들, 인터널 채널들의 개수 및 각 인터널 채널들의 타입은 포맷 변환기로 전달되며, 인터널 채널들의 개수는 포맷 변환기에서 공분산 매트릭스의 크기를 결정한다.
- [0185] 표 16 은 본 발명의 일 실시예에 따른, 채널 구성 요소 및 재생 레이아웃에 기초하여 결정되는 MPS와 SBR의 디코딩 시나리오를 나타낸다.

표 16

Reproduction Layout	Element	Order of MPS and SBR
Stereo	CPE	An MPS after mono SBR
Stereo	CPE	An MPS before stereo SBR
Stereo	QCE	Two MPS before two stereo SBR
Non-stereo	CPE/QCE	Independent of the order

- [0186]
- [0187] MPS는 최소의 채널(모노 또는 스테레오)로 합쳐진 다운믹스와 멀티채널 오디오 신호에 대한 인간의 지각 특성을 나타내는 공간 단서(spatial cue)파라미터들로 구성된 부가 정보(ancillary data)를 이용하여 멀티 채널 오디오 신호를 인코딩하는 기술이다.
- [0188] MPS 인코더는 N개의 멀티채널 오디오 신호를 입력받아 부가정보로서 양이 효과에 바탕을 둔 두 귀 사이의 소리 크기 차이 및 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 오디오를 제공할 수 있다.
- [0189] 또한 MPS 인코더는 입력된 멀티 채널 입력 신호로부터 다운믹스 신호를 생성하며, 생성된 다운믹스 신호는 오디오 입축 기술인 MPEG USAC 등으로 부호화되어 공간 파라미터와 함께 전송된다.

- [0190] 이 때 부호화기에 입력된 N개의 멀티 채널 신호는 분석 필터 뱅크(analysis filter bank)에 의해 주파수 밴드로 분해된다. 주파수 영역을 서브밴드로 분할하는 대표적인 방법은 DFT(Discrete Fourier Transform)이나 QMF(Quadrature Mirror Filter)를 이용하는 것인데, MPEG 서라운드에서는 보다 낮은 복잡도로 이를 수행하기 위하여 QMF 필터를 사용한다. QMF 필터를 사용하는 경우, SBR(Spectral Band Replication)과의 호환성이 보장되어 보다 효율적인 부호화를 수행할 수 있다.
- [0191] SBR은 사람이 상대적으로 감지하기 힘든 대역인 고주파 대역에 저주파 대역을 복사하여 붙여넣고 고주파 대역 신호에 대한 정보는 파라미터화 하여 전송하는 기술로, 낮은 비트레이트로 넓은 대역폭을 구현할 수 있다. 압축율이 높고 비트레이트는 낮은 코덱에서 주로 사용되며 고주파 대역의 일부 정보가 소실되므로 고조파(harmonics)를 표현하기 어렵지만, 가청 주파수 내에서는 높은 복원율을 갖는다.
- [0192] 인터널 채널 처리에 적용되는 SBR은 처리되는 도메인에 대한 차이를 제외하고 ISO/IEC 23003-3:2012와 동일하다. ISO/IEC 23003-3:2012의 SBR은 QMF 도메인에서 정의되지만 인터널 채널은 하이브리드 QMF 도메인에서 처리된다. 따라서, QMF 도메인의 인덱스 개수가 k라면 인터널 채널에 대한 전체 SBR 처리를 위한 주파수 인덱스 개수는 k+7이 된다.
- [0193] CPE가 스테레오 재생 레이아웃으로 출력되는 경우로, 모노 SBR 디코딩 후 MPS 디코딩하는 디코딩 시나리오에 대한 실시예는 도 6 에서 개시된다.
- [0194] CPE가 스테레오 재생 레이아웃으로 출력되는 경우로, MPS 디코딩 후 스테레오 SBR 디코딩하는 디코딩 시나리오에 대한 실시예는 도 7 에서 개시된다.
- [0195] QCE가 스테레오 재생 레이아웃으로 출력되는 경우로, CPE 쌍에 대하여 각각 MPS 디코딩 후, 디코딩 된 신호들에 대하여 각각 스테레오 SBR 디코딩하는 시나리오에 대한 실시예는 도 8 및 도 9에서 개시된다.
- [0196] CPE 또는 QCE가 출력되는 재생 레이아웃이 스테레오가 아닌 경우는 MPS 디코딩과 SBR 디코딩 순서는 상관없다.
- [0197] 디코더에서 처리되는, MPS212로 인코딩 된 CPE 신호들에 대한 정의는 다음과 같다.
- [0198] `cplx_out_dmx[]` 복소수 예측(complex prediction) 스테레오 디코딩 된 CPE 다운믹스 신호
- [0199] `cplx_out_dmx_preICG[]` 하이브리드 QMF 도메인에서 복소수 예측 스테레오 디코딩 및 하이브리드 QMF 분석 필터 뱅크 디코딩을 하여, ICG가 인코더에서 이미 적용된 모노 신호
- [0200] `cplx_out_dmx_postICG[]` 하이브리드 QMF 도메인에서 복소수 예측 스테레오 디코딩 및 인터널 채널 처리를 한, ICG가 디코더에서 적용되는 모노 신호
- [0201] `cplx_out_dmx_ICG[]` 하이브리드 QMF 도메인의 전대역(fullband) 인터널 채널 신호
- [0202] 디코더에서 처리되는, MPS212로 인코딩 된 QCE 신호들에 대한 정의는 다음과 같다.
- [0203] `cplx_out_dmx_L[]` 복소수 예측 스테레오 디코딩 된 제 1 CPE의 제 1 채널 신호
- [0204] `cplx_out_dmx_R[]` 복소수 예측 스테레오 디코딩 된 제 1 CPE의 제 2 채널 신호
- [0205] `cplx_out_dmx_L_preICG[]` 하이브리드 QMF 도메인의 제 1 ICG 선적용(pre-applied) 인터널 채널 신호
- [0206] `cplx_out_dmx_R_preICG[]` 하이브리드 QMF 도메인의 제 2 ICG 선적용 인터널 채널 신호
- [0207] `cplx_out_dmx_L_postICG[]` 하이브리드 QMF 도메인의 제 1 ICG 후적용(post-applied) 인터널 채널 신호
- [0208] `cplx_out_dmx_R_postICG[]` 하이브리드 QMF 도메인의 제 2 ICG 후적용 인터널 채널 신호
- [0209] `cplx_out_dmx_L_ICG_SBR 22.2-to-2` 포맷 변환을 위한 다운믹스된 파라미터들과 SBR에 의해 생성된 고주파 요소(component)를 포함하는, 제 1 전대역 디코딩된 인터널 채널 신호
- [0210] `cplx_out_dmx_R_ICG_SBR 22.2-to-2` 포맷 변환을 위한 다운믹스된 파라미터들과 SBR에 의해 생성된 고주파 요소(component)를 포함하는, 제 2 전대역 디코딩된 인터널 채널 신호
- [0211] 도 6 은 본 발명의 일 실시예에 따른, CPE가 스테레오 재생 레이아웃으로 출력되는 경우, 모노 SBR 디코딩 후 MPS 디코딩하는 구조에서의 인터널 채널 처리 방법의 순서도이다.
- [0212] CPE 비트스트림이 수신되면, 제일 먼저 `ICGDisabledCPE[n]` 플래그를 통해 CPE 사용 여부를 판단(610)한다.

- [0213] ICGDisabledCPE[n]이 TRUE라면, CPE 비트스트림은 ISO/IEC 23008-3에 정의된대로 디코딩(620)되며, ICGDisabledCPE[n]이 FALSE라면, SBR이 필요한경우 CPE 비트스트림에 모노 SBR을 수행하고, 스테레오 디코딩하여 다운믹스 시그널 cplx_out_dmx를 생성(630) 한다.
- [0214] 인터널 채널 게인이 인코더단에서 이미 적용되었는지 여부를 판단하기 위해 ICGPreAppliedCPE를 확인(640)한다.
- [0215] ICGPreAppliedCPE[n]이 FALSE이면, 다운믹스 시그널 cplx_out_dmx를 하이브리드 QMF 도메인 인터널 채널 처리(650)하여 ICG 후적용 다운믹스 신호 cplx_out_dmx_postICG가 생성된다. 인터널 채널 처리(650)에서, 인터널 채널 게인을 계산하기 위해 MPS 파라미터가 사용된다. CPE에 대한 역양자화된 선형 CLD 값은 ISO/IEC 23008-3에 의해 계산되며, 인터널 채널 게인은 수학적 식 2에 따라 계산된다.
- [0216] 다운믹스 신호 cplx_out_dmx에 수학적 식 2에 의한 인터널 채널 게인을 곱하여 cplx_out_dmx_postICG가 생성된다.

수학적 식 2

[0217]
$$G_{ICH}^{l,m} = \sqrt{(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m)^2 + (c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m)^2}$$

[0218] 이 때, $c_{left}^{l,m}$ 와 $c_{right}^{l,m}$ 는 CPE 신호 에 대한 1번째 타임슬롯과 m번째 하이브리드 QMF 밴드의 역양자화된 선형 CLD 값을, G_{left} 와 G_{right} 는 ISO/IEC 23008-3 테이블 96, 즉 포맷 변환 규칙 테이블에 정의된 출력 채널에 대한 게인 열의 값을 나타내며, $G_{EQ,left}^m$ 와 $G_{EQ,right}^m$ 는 포맷 변환 규칙 테이블에 정의된 출력 채널에 대한 EQ의 m번째 밴드의 게인을 나타낸다.

[0219] ICGPreAppliedCPE[n]이 TRUE이면, 다운믹스 시그널 cplx_out_dmx를 분석(660)하여 ICG 선적용 다운믹스 신호 cplx_out_dmx_preICG를 획득한다.

[0220] ICGPreApplied CPE[n]의 설정에 따라 cplx_out_dmx_preICG 또는 cplx_out_dmx_postICG가 최종 인터널 채널 처리 출력 신호 cplx_out_dmx_ICG가 된다.

[0221] 도 7 은 본 발명의 일 실시예에 따른, CPE가 스테레오 재생 레이아웃으로 출력되는 경우, MPS디코딩 후 스테레오 SBR 디코딩하는 인터널 채널 처리 방법의 순서도이다.

[0222] 도 7 에 개시된 실시예에서는, 도 6 에 개시된 실시예와 달리, MPS 디코딩 된 후 SBR 디코딩을 수행하므로 인터널 채널을 이용하지 않는 경우라면 스테레오 SBR 디코딩이 수행되지만, 인터널 채널을 이용하는 경우라면 모노 SBR이 수행되며 이를 위해 스테레오 SBR을 위한 파라미터들이 다운믹스된다.

[0223] 따라서, 도 6 과 비교하여 도 7 에서는, 2채널에 대한 SBR 파라미터들을 다운믹스하여 1채널에 대한 SBR 파라미터를 생성(780)하고, 생성된 SBR 파라미터를 이용하여 모노 SBR을 수행(770)하는 단계가 추가되며, 모노 SBR이 수행된 cplx_out_dmx_ICG가 최종 인터널 채널 처리 출력 신호 cplx_out_dmx_ICG가 된다.

[0224] 도 7 과 같은 동작 순서 배치에서, 인터널 채널 처리 후 SBR가 수행되어 고주파 성분이 확장되므로 cplx_out_dmx_preICG 신호 또는 cplx_out_dmx_postICG 신호는 대역제한(bandlimited) 신호에 해당한다. 업믹스된 스테레오 신호를 위한 SBR 파라미터 쌍은 대역제한된 인터널 채널 신호 cplx_out_dmx_preICG 또는 cplx_out_dmx_postICG의 대역폭 확장을 위해 파라미터 도메인에서 다운믹스 되어야 한다.

[0225] SBR 파라미터 다운믹서는 SBR에 의해 확장된 고주파 밴드들을 포맷 변환기의 EQ, 게인 파라미터와 곱하는 과정을 포함해야 한다. SBR 파라미터를 다운믹싱하는 구체적인 방법은 후술한다.

[0226] 도 8 은 본 발명의 일 실시예에 따른, QCE가 스테레오 재생 레이아웃으로 출력되는 경우, 스테레오 SBR을 이용하는 구조에서의 인터널 채널 처리 방법의 블록도이다.

[0227] 도 8 에 개시된 실시예는, ICGPreApplied[n]과 ICGPreApplied[n+1]이 모두 0인 경우, 즉 디코더에서 인터널 채널 게인을 적용하는 방법에 대한 실시예이다.

[0228] 도 8 에 도시된 블록도에서 전체적인 디코딩 구조는, 비트스트림 디코딩(810), 스테레오 디코딩(820), 하이브리드 QMF 분석(830), 인터널 채널 처리(840) 및 스테레오 SBR(850)의 순서로 진행된다.

- [0229] QCE를 구성하는 CPE 쌍 각각에 대한 비트스트림이 디코딩(811, 812)되면, 디코딩 된 신호로부터 SBR 페이로드, MPS212 페이로드 및 CplxPred 페이로드가 추출된다.
- [0230] 디코딩된 CplxPred 페이로드를 이용하여 스테레오 디코딩(821)이 수행되며, 스테레오 디코딩 된 신호 cplx_dmx_L와 cplx_dmx_R가 각각 하이브리드 QMF 분석(831, 832)과정을 거쳐 인터널 채널 처리부(841, 842)의 입력 신호로 전달된다.
- [0231] 이 때, 생성되는 인터널 채널 신호 cplx_dmx_L_PostICG 및 cplx_dmx_R_PostICG는 대역제한 신호이다. 따라서, 두 인터널 채널 신호는 각 CPE의 비트스트림에서 추출한 SBR 페이로드를 다운믹스한 다운믹스 SBR 파라미터들을 이용하여 스테레오 SBR(851)된다. 대역 제한 인터널 채널 신호는 스테레오 SBR을 통해 고주파 확장되어, 전대역 인터널 채널 처리 출력 신호인 cplx_dmx_L_ICG 및 cplx_dmx_R_ICG가 생성된다.
- [0232] 다운믹스 SBR 파라미터들은 대역제한된 인터널 채널 신호를 대역 확장하여 전대역 인터널 채널 신호를 생성하기 위해 이용된다.
- [0233] 이와 같이, QCE에 대한 인터널 채널을 이용하는 경우, 스테레오 디코딩 블록 및 스테레오 SBR 블록은 두개가 아닌 하나만 이용하므로 스테레오 디코딩 블록 822 및 스테레오 SBR 블록 852는 생략될 수 있다. 즉 QCE를 이용함으로써 각각의 CPE를 각각 처리하는 경우보다 보다 간단한 디코딩 구조의 구현이 가능하다.
- [0234] 도 9 는 본 발명의 또 다른 실시예에 따른, QCE가 스테레오 재생 레이아웃으로 출력되는 경우, 스테레오 SBR을 이용하는 구조에서의 인터널 채널 처리 방법의 블록도이다.
- [0235] 도 9 에 개시된 실시예는, ICGPreApplied[n]과 ICGPreApplied[n+1]이 모두 1인 경우, 즉 인코더에서 인터널 채널 계인을 적용하는 방법에 대한 실시예이다.
- [0236] 도 9 에 도시된 블록도에서 전체적인 디코딩 구조는, 비트스트림 디코딩(910), 스테레오 디코딩(920), 하이브리드 QMF 분석(930) 및 스테레오 SBR(950)의 순서로 진행된다.
- [0237] 인코더에서 인터널 채널 계인을 적용한 경우, 디코더에서는 별도의 인터널 채널 처리를 수행하지 않으므로, 도 9 에서는 도 8과 비교해 인터널 채널 처리 블록(841, 842)이 생략되어 있다. 그 외의 과정은 도 8과 유사하므로 중복되는 설명은 생략한다.
- [0238] 스테레오 디코딩 된 신호 cplx_dmx_L와 cplx_dmx_R는 각각 하이브리드 QMF 분석(931, 932)과정을 거쳐 바로 스테레오 SBR 블록(951)의 입력 신호로 전달된다. 스테레오 SBR 블록을 통과하면, 전대역 인터널 채널 처리 출력 신호인 cplx_dmx_L_ICG 및 cplx_dmx_R_ICG가 생성된다.
- [0239] 한편, 출력 채널이 스테레오 채널이 아닌 경우 인터널 채널을 이용하는 것을 적절하지 않을 수 있다. 따라서, 인코더에서 인터널 채널 계인이 적용된 경우, 출력 채널이 스테레오 출력 채널이 아니라면 디코더는 역 인터널 채널 계인(IG, inverse ICG)을 적용하여야 한다.
- [0240] 이 때, 표 8에 도시된 것처럼 MPS와 SBR의 디코딩 순서는 상관없이 있지만, 설명의 편의를 위해 모노 SBR 디코딩 후 MPS212 디코딩 하는 시나리오에 대해 설명한다.
- [0241] 역 인터널 채널 계인 IG는 MPS 파라미터들과 포맷 변환 파라미터들을 이용하여, 수학적 식 3과 같이 계산된다.

수학적 식 3

[0242]
$$IG_{CH}^{l,m} = \frac{1}{\sqrt{(c_{left}^{l,m} \times G_{left} \times G_{EQ,left}^m)^2 + (c_{right}^{l,m} \times G_{right} \times G_{EQ,right}^m)^2}}$$

[0243] 이 때, $c_{left}^{l,m}$ 와 $c_{right}^{l,m}$ 는 CPE 신호 에 대한 1번째 타임슬롯과 m번째 하이브리드 QMF 밴드의 역양자화된 선형 CLD 값을, G_{left} 와 G_{right} 는 ISO/IEC 23008-3 테이블 96, 즉 포맷 변환 규칙 테이블에 정의된 출력 채널에 대한 계인 열의 값을 나타내며, $G_{EQ,left}^m$ 와 $G_{EQ,right}^m$ 는 포맷 변환 규칙 테이블에 정의된 출력 채널에 대한 EQ의 m번째 밴드의 계인을 나타낸다.

[0244] 만일 ICGPreAppliedCPE[n]가 참이라면, n 번째 cplx_dmx는 MPS 블록 이전에 역 인터널채널 계인이 곱해져야 하

며 나머지 디코딩 과정은 ISO/IEC 23008-3를 따라야 한다.

[0245] 디코더에서 인터널 채널 처리 블록이 이용되거나 인코더에서 인터널 채널 게인이 전처리되고, 출력 레이아웃이 스테레오인 경우, SBR 블록 이전단에서 CPE/QCE에 대한 MPS 업믹스된 스테레오/쿼드 채널 신호 대신 대역제한 인터널 채널 신호가 생성된다.

[0246] SBR 페이로드는 MPS 업믹스된 스테레오/쿼드 채널 신호를 위해 스테레오 SBR 인코딩 되어 있으므로, 인터널 채널 처리를 위해서 스테레오 SBR 페이로드는 파라미터 도메인에서 포맷 변환기의 게인 및 EQ를 곱함으로써 다운믹스되어야 한다.

[0247] 이하에서는 스테레오 SBR을 파라미터 다운믹싱하는 구체적인 방법에 대해 설명한다.

[0248] (1) 역필터링(inverse filtering)

[0249] 각 노이즈 플로어 밴드에서 스테레오 SBR 파라미터들이 최대값을 가지도록 함으로써 역필터링 모드가 선택된다.

[0250] 이를 구현하기 위한 구체적인 식은 [수학식 4]와 같다.

수학식 4

[0251]
$$\text{for } (i = 0; i < N_Q; i++)$$

$$\text{bs_invf_mode}_{\text{Downmixed}}(i) = \text{MAX}(\text{bs_invf_mode}_{\text{ch1}}(i), \text{bs_invf_mode}_{\text{ch2}}(i))$$

$$\begin{pmatrix} \text{ch1} \\ \text{ch2} \end{pmatrix} = \begin{cases} \begin{pmatrix} \text{Left of CPE1} \\ \text{Left of CPE2} \end{pmatrix} & \text{in case of Cplx_out_dmx_L} \\ \begin{pmatrix} \text{Right of CPE1} \\ \text{Right of CPE2} \end{pmatrix} & \text{in case of Cplx_out_dmx_R} \end{cases}$$

[0252] (2) 추가 고조파(additional harmonics)

[0253] 기본주파수(f) 및 기본주파수의 홀수 고조파(3f, 5f, 7f, ...)로 구성된 음파는, 반파대칭이 된다. 그러나 기본 주파수의 짝수 고조파(0f, 2f, 3, ...)로 구성된 음파는 대칭성을 가지지 않는다. 반대로, 단순 스케일링 또는 이동 이외의 음원 파형 변화를 야기하는 비선형 시스템은 추가 고조파를 발생시켜 조화 왜곡(harmonic distortion)이 일어난다.

[0254] 추가 고조파는 추가적인 사인파들의 조합으로, 수학식 5와 같이 표현될 수 있다.

수학식 5

[0255]
$$\text{for } (i = 0; i < N_{\text{High}}; i++)$$

$$\text{bs_add_harmonic}_{\text{Downmixed}}(i) = \text{OR}(\text{bs_add_harmonic}_{\text{ch1}}(i), \text{bs_add_harmonic}_{\text{ch2}}(i))$$

[0256] (3) 포락선 시간 경계(envelope time borders)

[0257] 도 10 은 본 발명의 일 실시예에 따른, SBR 파라미터인 시간 경계를 결정하는 방법을 나타내는 도면이다.

[0258] 도 10a 는 첫번째 포락선의 시작 경계들이 서로 같고, 마지막 포락선의 종료 경계들이 서로 같은 경우의 시간 포락선 그리드를 나타낸다.

[0259] 도 10b 는 첫번째 포락선의 시작 경계들은 서로 다르고, 마지막 포락선의 종료 경계들은 같은 경우의 시간 포락선 그리드를 나타낸다.

[0260] 도 10c 는 첫번째 포락선의 시작 경계들은 서로 같고, 마지막 포락선의 종료 경계들은 서로 다른 경우의 시간 포락선 그리드를 나타낸다.

[0261] 도 10d 는 첫번째 포락선의 시작 경계들이 서로 다르고, 마지막 포락선의 종료 경계들이 서로 다른 경우의 시간 포락선 그리드를 나타낸다.

[0262] 인터널 채널 SBR에 대한 시간 포락선 그리드 f_{B_Merged} 는 스테레오 SBR 시간 그리드를 가장 높은 해상도를 갖는 가장 작은 조각들로 분할함으로써 생성된다.

- [0263] t_{E_Merged} 의 시작 경계값은 스테레오 채널에 대한 가장 큰 시작 경계값으로 설정된다. 시간 그리드 0과 시작 경계 사이의 포락선은 이전 프레임에서 이미 처리되어 있다. 두 채널들의 마지막 포락선들의 종료 경계들 중에서 가장 큰 값이 마지막 포락선의 종료 경계로 선택된다.
- [0264] 도 10 에 도시된 것과 같이, 두 채널들의 시간 경계들의 교차점을 취함으로써 첫번째 포락선과 마지막 포락선의 시작/종료 경계들은 가장 세분화된 해상도를 갖도록 결정된다. 만일 5개 이상의 포락선이 있다면, t_{E_Merged} 의 종료점에서부터 시작하여 역으로 t_{E_Merged} 의 시작점까지를 검색하여 4 보다 작은 개수의 포락선을 찾아 그 포락선의 시작 경계를 제거함으로써 포락선의 개수를 감소시켜야 한다. 이와 같은 과정은 5개의 포락선이 남을때까지 계속된다.
- [0265] (4) 잡음 시간 경계(noise time borders)
- [0266] 다운믹스된 잡음 시간 경계 L_{Q_Merged} 의 개수는 두개의 채널들의 잡음 시간 경계들 사이에 큰 값을 취함으로써 결정된다. 첫번째 그리드와 병합 잡음 시간 경계 t_{Q_Merged} 는 포락선 시간 경계 t_{E_Merged} 의 첫번째 그리드와 마지막 그리드를 취함으로써 결정된다.
- [0267] 만일 잡음 시간 경계 L_{Q_Merged} 가 1보다 크다면, $t_{Q_Merged}^{(1)}$ 이 잡음 시간 경계 L_Q 가 1보다 큰 채널의 $t_Q^{(1)}$ 로 선택된다. 만일 두 채널 모두 1보다 큰 L_Q 를 갖는다면, $t_Q^{(1)}$ 의 최소값이 $t_{Q_Merged}^{(1)}$ 로 선택된다.
- [0268] (5) 포락선 데이터(envelope data)
- [0269] 도 11 은 본 발명의 일 실시예에 따른, SBR 파라미터인 주파수 해상도를 병합하는 방법을 설명하기 위한 도면이다.
- [0270] 병합 포락선 시간 경계의 주파수 해상도 r_{Merged} 가 선택된다. r_{Merged} 의 각 섹션에 대한 주파수 해상도 r_{ch1} , r_{ch2} 사이이 최대값이 도 11과 같이 r_{Merged} 로 선택된다.
- [0271] 모든 포락선에 대한 포락선 데이터 E_{Orig_Merged} 는 포락선 데이터 E_{Orig} 로부터 포맷 변환 파라미터들을 고려하여, 수학식 6과 같이 계산된다.

수학식 6

$$E_{OrigMerged}(k, l) = E_{ch1Orig}(g_{ch1}(k), h_{ch1}(l)) \times (EQ_{ch1}(k, h_{ch1}(l)))^2 + E_{ch2Orig}(g_{ch2}(k), h_{ch2}(l)) \times (EQ_{ch2}(k, h_{ch2}(l)))^2$$

[0272] 이 때,

$$EQ_{ch1}(k, l) = \frac{\sum_m (G_{ch1}^m \times G_{EQ, ch1}^m)}{F(k+1, r_{Merged}(l)) - F(k, r_{Merged}(l))}, F(k, r_{Merged}(l)) \leq m < F(k+1, r_{Merged}(l))$$

$$EQ_{ch2}(k, l) = \frac{\sum_m (G_{ch2}^m \times G_{EQ, ch2}^m)}{F(k+1, r_{Merged}(l)) - F(k, r_{Merged}(l))}, F(k, r_{Merged}(l)) \leq m < F(k+1, r_{Merged}(l))$$

$$0 \leq k < n(r_{Merged}(l)), 0 \leq l < L_{E_Merged} \text{ 이며,}$$

$$t_{E_ch1}(h_{ch1}(l)) \leq t_{E_Merged}(l) < t_{E_ch1}(h_{ch1}(l) + 1)$$

$$t_{E_ch2}(h_{ch2}(l)) \leq t_{E_Merged}(l) < t_{E_ch2}(h_{ch2}(l) + 1) \text{ 이 고,}$$

$$g_{ch1}(k) \leq F(g_{ch1}(k), r_{ch1}(h_{ch1}(l))) \leq F(k, r_{Merged}(l)) < F(g_{ch1}(k) + 1, r_{ch1}(h_{ch1}(l)))$$

[0280] $g_{ch2}(k)$ 는 $F(g_{ch2}(k), r_{ch2}(h_{ch2}(l))) \leq F(k, r_{Merged}(l)) < F(g_{ch2}(k) + 1, r_{ch2}(h_{ch2}(l)))$ 와 같이 정의된다.

[0281] (6) 잡음 플로어 데이터(noise floor data)

[0282] 병합된 잡음 플로어 데이터 $Q_{OrigMerged}$ 는 두 채널 데이터의 합으로 수학식 7에 따라 결정된다.

수학식 7

[0283] $Q_{OrigMerged}(k, l) = Q_{Origch1}(k, h_{ch1}(l)) + Q_{Origch2}(k, h_{ch2}(l)), 0 \leq k < N_Q, 0 \leq l < L_{Q,Merged}$

[0284] 이 때,

[0285] $h_{ch1}(l)$ 는 $t_{Q,ch1}(h_{ch1}(l)) \leq t_{Q,Merged}(l) < t_{Q,ch1}(h_{ch1}(l) + 1)$,

[0286] $h_{ch2}(l)$ 는 $t_{Q,ch2}(h_{ch2}(l)) \leq t_{Q,Merged}(l) < t_{Q,ch2}(h_{ch2}(l) + 1)$ 와 같이 정의된다.

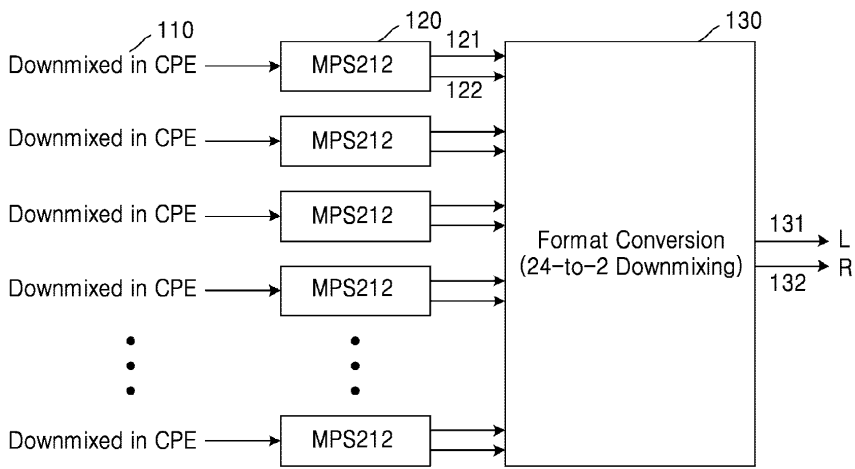
[0287] 이상 설명된 본 발명에 따른 실시예는 다양한 컴퓨터 구성요소를 통하여 실행될 수 있는 프로그램 명령어의 형태로 구현되어 컴퓨터 판독 가능한 기록 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능한 기록 매체는 프로그램 명령어, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 컴퓨터 판독 가능한 기록 매체에 기록되는 프로그램 명령어는 본 발명을 위하여 특별히 설계되고 구성된 것이거나 컴퓨터 소프트웨어 분야의 당업자에게 공지되어 사용 가능한 것일 수 있다. 컴퓨터 판독 가능한 기록 매체의 예에는, 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체, CD-ROM 및 DVD와 같은 광기록 매체, 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical medium), 및 ROM, RAM, 플래시 메모리 등과 같은, 프로그램 명령어를 저장하고 실행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령어의 예에는, 컴파일러에 의하여 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용하여 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드도 포함된다. 하드웨어 장치는 본 발명에 따른 처리를 수행하기 위하여 하나 이상의 소프트웨어 모듈로 변경될 수 있으며, 그 역도 마찬가지이다.

[0288] 이상에서 본 발명이 구체적인 구성요소 등과 같은 특정 사항과 한정된 실시예 및 도면에 의하여 설명되었으나, 이는 본 발명의 보다 전반적인 이해를 돕기 위하여 제공된 것일 뿐, 본 발명이 상기 실시예에 한정되는 것은 아니며, 본 발명이 속하는 기술분야에서 통상적인 지식을 가진 자라면 이러한 기재로부터 다양한 수정과 변경을 꾀할 수 있다.

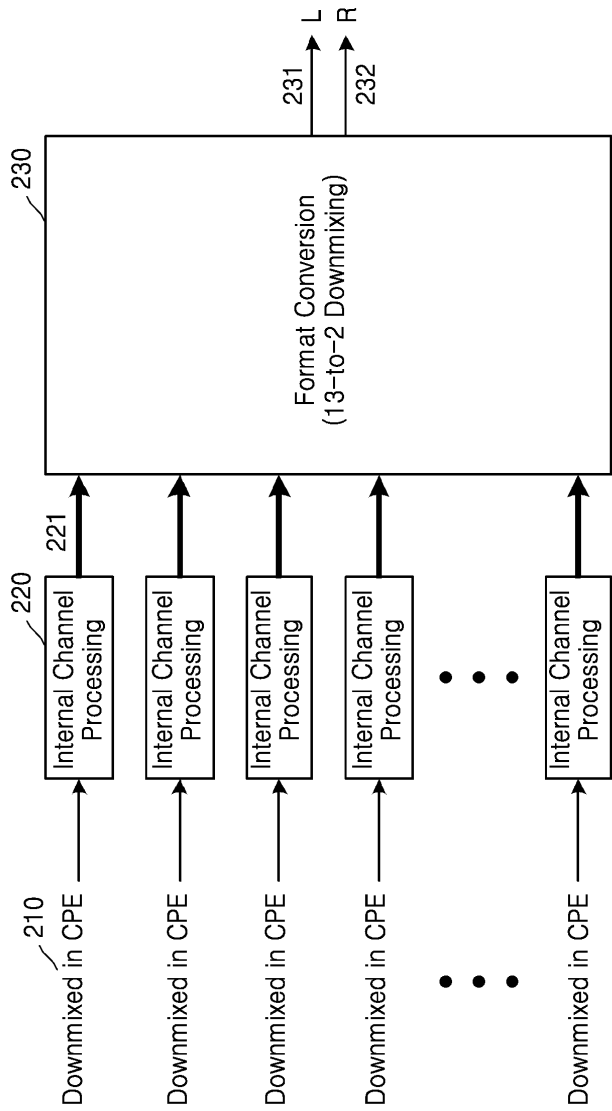
[0289] 따라서, 본 발명의 사상은 상기 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐만 아니라 이 특허청구범위와 균등한 또는 이로부터 등가적으로 변경된 모든 범위는 본 발명의 사상의 범주에 속한다고 할 것이다.

도면

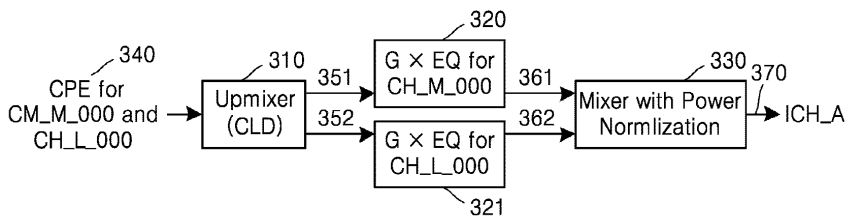
도면1



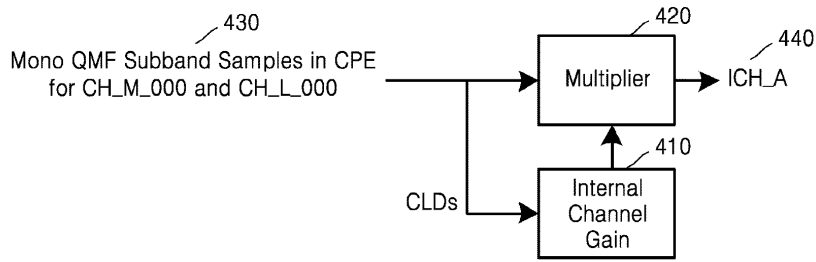
도면2



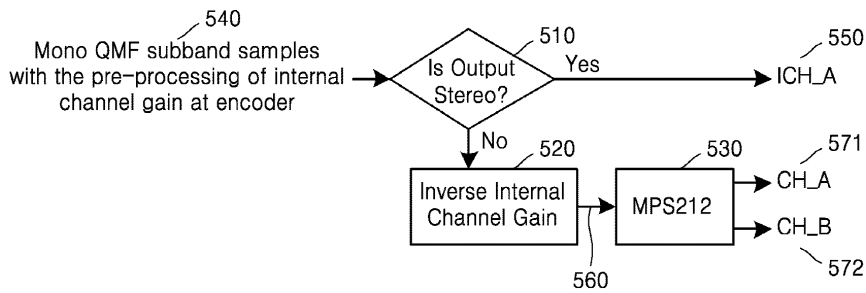
도면3



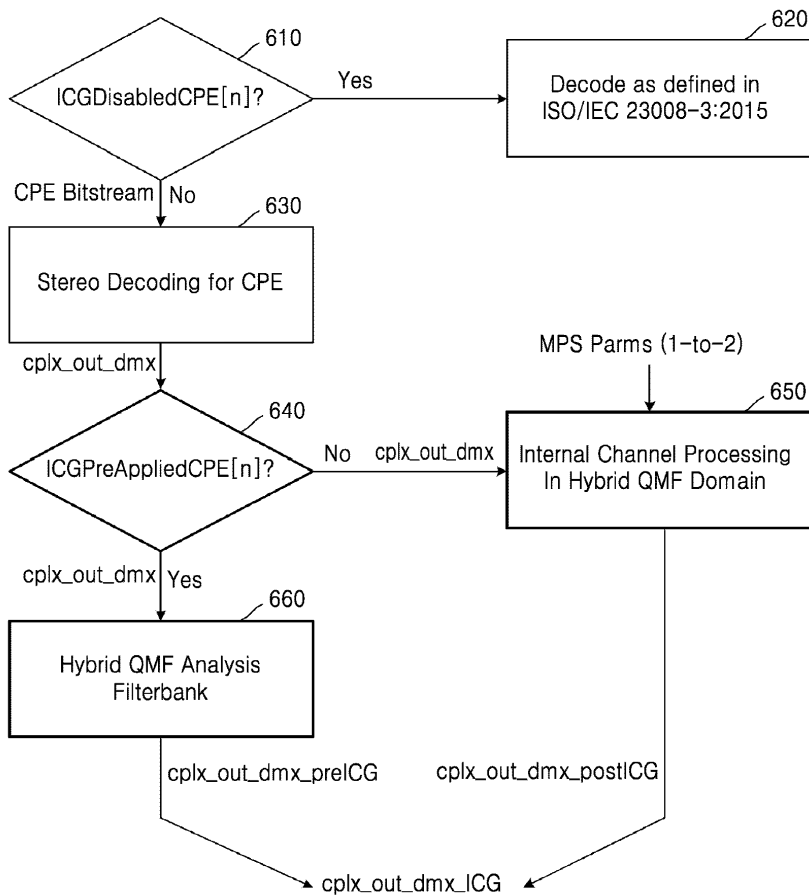
도면4



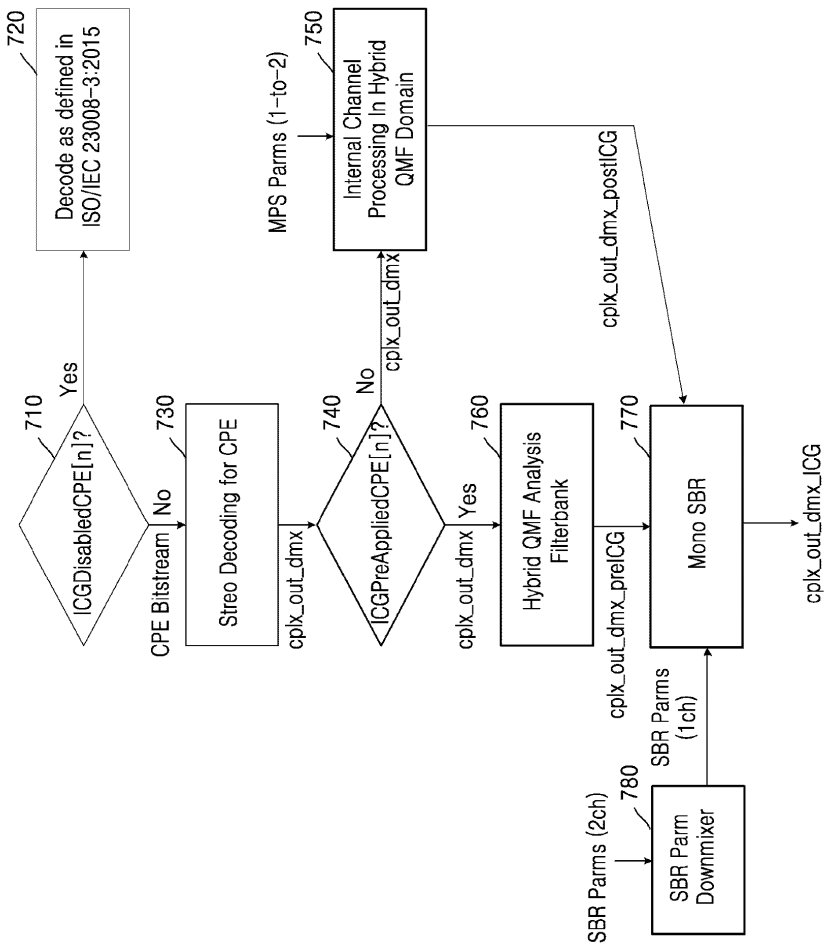
도면5



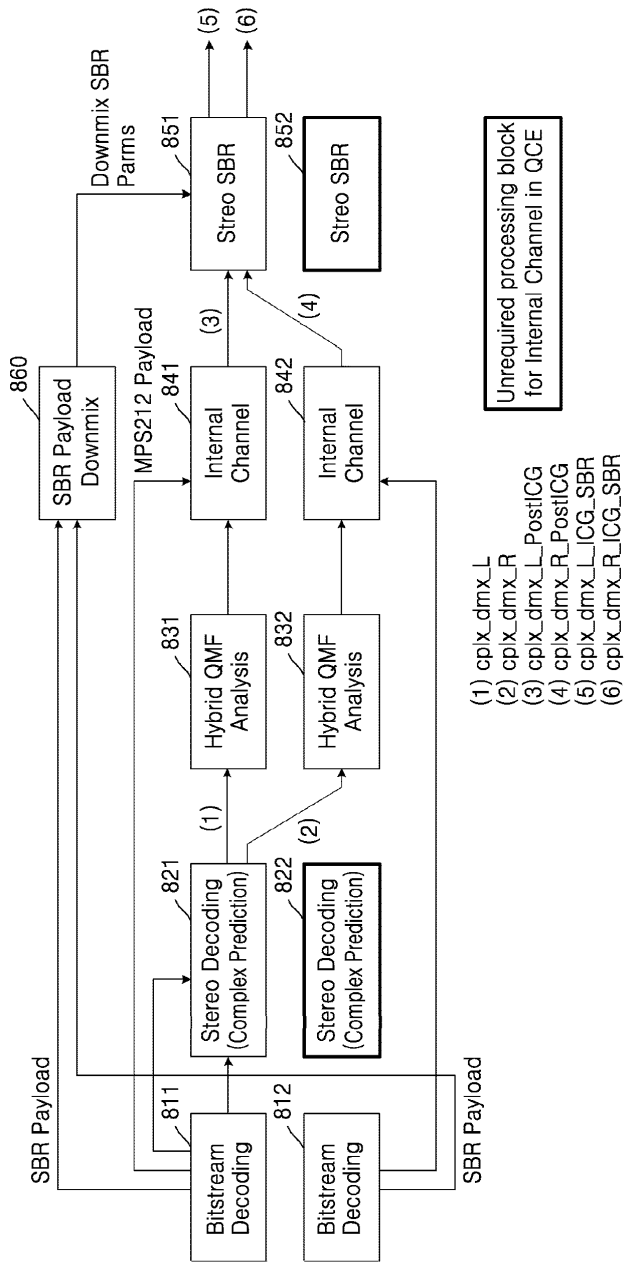
도면6



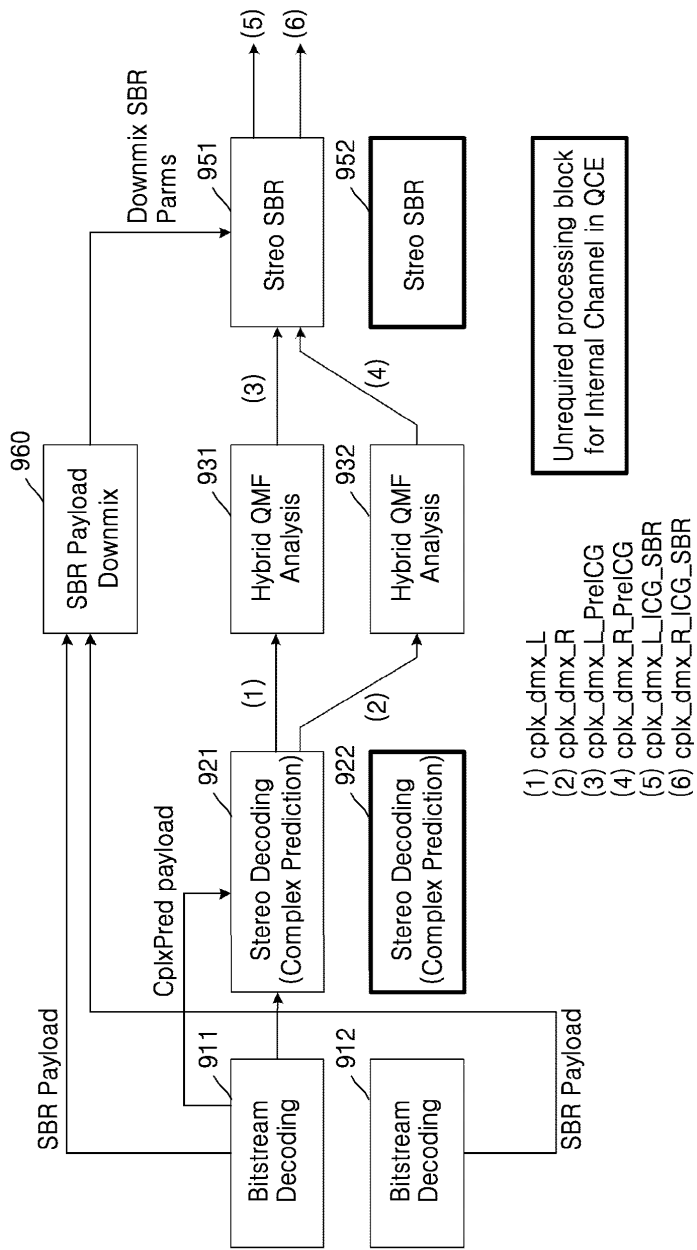
도면7



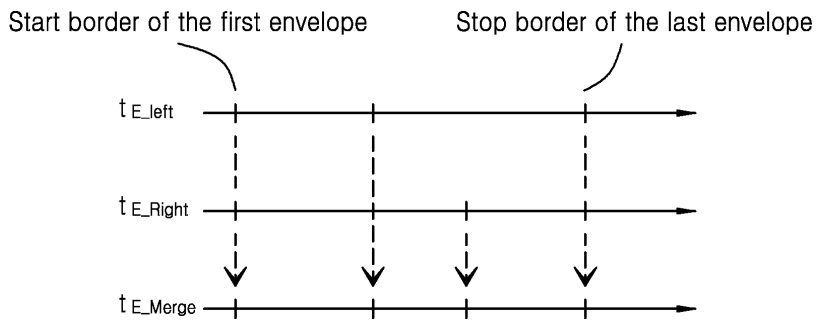
도면8



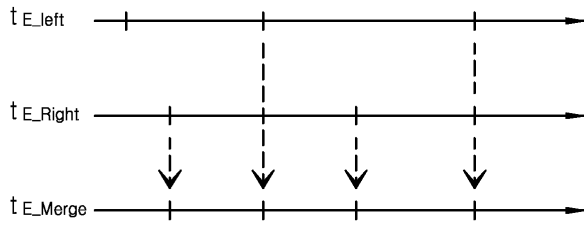
도면9



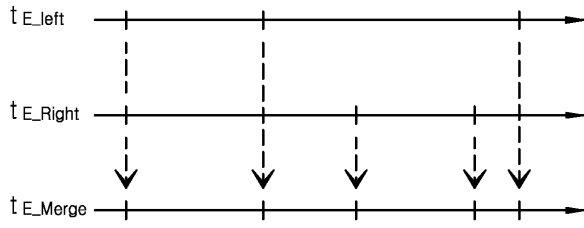
도면10a



도면10b



도면10c



도면10d

