



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년09월28일
 (11) 등록번호 10-1782916
 (24) 등록일자 2017년09월22일

- | | |
|--|--|
| (51) 국제특허분류(Int. Cl.)
HO4S 3/00 (2006.01)
(52) CPC특허분류
HO4S 3/00 (2013.01)
HO4S 3/008 (2013.01)
(21) 출원번호 10-2016-7006857
(22) 출원일자(국제) 2014년09월17일
심사청구일자 2016년03월15일
(85) 번역문제출일자 2016년03월15일
(65) 공개번호 10-2016-0048832
(43) 공개일자 2016년05월04일
(86) 국제출원번호 PCT/KR2014/008677
(87) 국제공개번호 WO 2015/041476
국제공개일자 2015년03월26일
(30) 우선권주장
1020130125936 2013년10월22일 대한민국(KR)
(뒷면에 계속)
(56) 선행기술조사문헌
KR100754220 B1
KR1020090020813 A | (73) 특허권자
주식회사 윌러스표준기술연구소
서울특별시 서초구 마방로 48, 2층(양재동, 경
원빌딩)
(72) 발명자
오현오
경기도 과천시 별양로 12, 322-1102 (래미안슈르
아파트)
이태규
경기도 성남시 분당구 장미로 139,210-304(매화마
을주공2단지아파트)
(74) 대리인
임국일 |
|--|--|

전체 청구항 수 : 총 14 항

심사관 : 권영학

(54) 발명의 명칭 **오디오 신호 처리 방법 및 장치**

(57) 요약

본 발명은 오디오 신호를 효과적으로 재생하기 위한 신호 처리 방법 및 장치에 관한 것으로서, 더욱 상세하게는 멀티채널 혹은 멀티오브젝트 오디오 신호를 스테레오로 재생하기 위한 바이노럴 렌더링을 낮은 연산량으로 구현하기 위한 오디오 신호 처리 방법 및 장치에 관한 것이다.

이를 위해 본 발명은, 멀티채널 또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호를 수신하는 단계; 상기 멀티 오디오 신호의 필터링을 위한 절단된 서브밴드 필터 계수들을 수신하는 단계, 상기 절단된 서브밴드 필터 계수는 상기 멀티 오디오 신호의 바이노럴 필터링을 위한 BRIR(Binaural Room Impulse Response) 필터 계수로부터 획득된 서브밴드 필터 계수의 적어도 일 부분이며, 상기 절단된 서브밴드 필터 계수의 길이는 해당 서브밴드 필터 계수에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 오더 정보에 기초하여 결정됨; 및 상기 멀티 오디오 신호의 각 서브밴드 신호에 대응하는 상기 절단된 서브밴드 필터 계수를 이용하여 상기 서브밴드 신호를 필터링 하는 단계; 를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법 및 이를 이용한 오디오 신호 처리 장치를 제공한다.

(52) CPC특허분류

H04S 2400/01 (2013.01)

H04S 2420/03 (2013.01)

(30) 우선권주장

61/878,638 2013년09월17일 미국(US)

61/894,442 2013년10월23일 미국(US)

명세서

청구범위

청구항 1

입력 오디오 신호를 수신하는 단계, 상기 입력 오디오 신호는 복수의 서브밴드 신호들을 포함함;

상기 입력 오디오 신호의 필터링을 위한 절단된 서브밴드 필터 계수들을 수신하는 단계, 상기 절단된 서브밴드 필터 계수들은 상기 입력 오디오 신호의 바이노럴 필터링을 위한 BRIR(Binaural Room Impulse Response) 필터 계수들로부터 획득된 서브밴드 필터 계수들의 적어도 일 부분이며, 상기 절단된 서브밴드 필터 계수들의 길이는 해당 서브밴드 필터 계수들에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 결정되고, 적어도 하나의 서브밴드의 절단된 서브밴드 필터 계수들의 길이는 다른 서브밴드의 절단된 서브밴드 필터 계수들의 길이와 다름; 및

상기 입력 오디오 신호의 각 서브밴드 신호에 대응하는 절단된 서브밴드 필터 계수들을 이용하여 상기 서브밴드 신호를 필터링 하는 단계;

를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

청구항 2

제1 항에 있어서,

상기 특성 정보는 해당 서브밴드 필터 계수들의 제1 잔향 시간 정보를 포함하며, 상기 필터 차수 정보는 각 서브밴드 별로 하나의 값을 갖는 것을 특징으로 하는 오디오 신호 처리 방법.

청구항 3

제1 항에 있어서,

적어도 하나의 서브밴드의 절단된 서브밴드 필터 계수들의 길이는 2의 거듭 제곱의 배수 값을 갖는 것을 특징으로 하는 오디오 신호 처리 방법.

청구항 4

제1 항에 있어서,

상기 복수의 서브밴드 신호들은 각각 기 설정된 주파수 밴드를 기준으로 한 저주파수의 제1 서브밴드 그룹과 고주파수의 제2 서브밴드 그룹을 포함하며,

상기 필터링 하는 단계는 상기 제1 서브밴드 그룹의 절단된 서브밴드 필터 계수들 및 서브밴드 신호들에 대하여 수행되는 것을 특징으로 하는 오디오 신호 처리 방법.

청구항 5

제1 항에 있어서,

상기 필터링 하는 단계는 해당 서브밴드 필터 계수들의 제1 잔향 시간 정보에 적어도 부분적으로 기초하여 절단된 프론트 서브밴드 필터 계수들을 이용하여 수행되며,

상기 서브밴드 필터 계수들 중 상기 프론트 서브밴드 필터 계수들 이후의 구간에 대응하는 상기 서브밴드 신호의 잔향 처리 단계를 더 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

청구항 6

제5 항에 있어서,

상기 잔향 처리 단계는,

각 서브밴드에 대한 다운믹스 서브밴드 필터 계수들을 수신하는 단계, 상기 다운믹스 서브밴드 필터 계수들은 해당 서브밴드의 각 리어 서브밴드 필터 계수들을 조합하여 생성되며, 상기 리어 서브밴드 필터 계수들은 해당

서브밴드 필터 계수들 중 상기 프론트 서브밴드 필터 계수들 이후의 구간으로부터 획득됨;

상기 각 서브밴드에 대한 다운믹스 서브밴드 신호를 생성하는 단계, 상기 다운믹스 서브밴드 신호는 해당 서브밴드의 각 서브밴드 신호들을 다운믹스 하여 생성됨; 및

상기 다운믹스 서브밴드 신호와 이에 대응하는 상기 다운믹스 서브밴드 필터 계수들을 이용하여 2채널의 좌, 우 서브밴드 잔향 신호를 생성하는 단계;

를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

청구항 7

제6 항에 있어서,

상기 다운믹스 서브밴드 신호는 모노 서브밴드 신호이고,

상기 다운믹스 서브밴드 필터 계수들은 상기 다운믹스 서브밴드 신호에 대한 잔향부의 에너지 감소 특성을 반영하며,

상기 모노 서브밴드 신호에 대한 디코릴레이션(decorrelation) 신호를 생성하는 단계; 및

상기 모노 서브밴드 신호와 상기 디코릴레이션 신호 간의 가중치 합산을 수행하여 2채널의 좌, 우 신호를 생성하는 단계를 더 포함하는 것을 특징으로 하는 오디오 신호 처리 방법.

청구항 8

입력 오디오 신호에 대한 바이노럴 렌더링을 수행하기 위한 오디오 신호 처리 장치로서, 상기 입력 오디오 신호는 복수의 서브밴드 신호들을 포함하며,

상기 각 서브밴드 신호에 대한 직접음 및 초기 반사음 파트의 렌더링을 수행하기 위한 고속 콘볼루션부; 및

상기 각 서브밴드 신호에 대한 후기잔향 파트의 렌더링을 수행하기 위한 후기잔향 생성부를 포함하되,

상기 고속 콘볼루션부는,

상기 입력 오디오 신호의 필터링을 위한 절단된 서브밴드 필터 계수들을 수신하되, 상기 절단된 서브밴드 필터 계수들은 상기 입력 오디오 신호의 바이노럴 필터링을 위한 BRIR(Binaural Room Impulse Response) 필터 계수들로부터 획득된 서브밴드 필터 계수들의 적어도 일 부분이며, 상기 절단된 서브밴드 필터 계수들의 길이는 해당 서브밴드 필터 계수들에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 결정되고, 적어도 하나의 서브밴드의 절단된 서브밴드 필터 계수들의 길이는 다른 서브밴드의 절단된 서브밴드 필터 계수들의 길이와 다름,

상기 입력 오디오 신호의 각 서브밴드 신호에 대응하는 절단된 서브밴드 필터 계수들을 이용하여 상기 서브밴드 신호를 필터링 하는 것을 특징으로 하는 오디오 신호 처리 장치.

청구항 9

제8 항에 있어서,

상기 특성 정보는 해당 서브밴드 필터 계수들의 제1 잔향 시간 정보를 포함하며, 상기 필터 차수 정보는 각 서브밴드 별로 하나의 값을 갖는 것을 특징으로 하는 오디오 신호 처리 장치.

청구항 10

제8 항에 있어서,

적어도 하나의 서브밴드의 절단된 서브밴드 필터 계수들의 길이는 2의 거듭 제곱의 배수 값을 갖는 것을 특징으로 하는 오디오 신호 처리 장치.

청구항 11

제8 항에 있어서,

상기 복수의 서브밴드 신호들은 각각 기 설정된 주파수 밴드를 기준으로 한 저주파수의 제1 서브밴드 그룹과 고

주파수의 제2 서브밴드 그룹을 포함하며,

상기 고속 콘볼루션부는, 상기 제1 서브밴드 그룹의 절단된 서브밴드 필터 계수들을 이용하여 상기 제1 서브밴드 그룹의 서브밴드 신호들에 대하여 필터링을 수행하는 것을 특징으로 하는 오디오 신호 처리 장치.

청구항 12

제8 항에 있어서,

상기 고속 콘볼루션부는, 해당 서브밴드 필터 계수들의 제1 잔향 시간 정보에 적어도 부분적으로 기초하여 절단된 프론트 서브밴드 필터 계수들을 이용하여 상기 필터링을 수행하며,

상기 오디오 신호 처리 장치는,

상기 서브밴드 필터 계수들 중 상기 프론트 서브밴드 필터 계수들 이후의 구간에 대응하는 상기 서브밴드 신호의 잔향 처리를 수행하는 후기잔향 생성부를 더 포함하는 것을 특징으로 하는 오디오 신호 처리 장치.

청구항 13

제12 항에 있어서,

상기 후기잔향 생성부는,

각 서브밴드에 대한 다운믹스 서브밴드 필터 계수들을 수신하되, 상기 다운믹스 서브밴드 필터 계수들은 해당 서브밴드의 각 리어 서브밴드 필터 계수들을 조합하여 생성되며, 상기 리어 서브밴드 필터 계수들은 해당 서브밴드 필터 계수들 중 상기 프론트 서브밴드 필터 계수들 이후의 구간으로부터 획득되고,

상기 각 서브밴드에 대한 다운믹스 서브밴드 신호를 생성하되, 상기 다운믹스 서브밴드 신호는 해당 서브밴드의 각 서브밴드 신호들을 다운믹스 하여 생성되고,

상기 다운믹스 서브밴드 신호와 이에 대응하는 상기 다운믹스 서브밴드 필터 계수들을 이용하여 2채널의 좌, 우 서브밴드 잔향 신호를 생성하는 것을 특징으로 하는 오디오 신호 처리 장치.

청구항 14

제13 항에 있어서,

상기 다운믹스 서브밴드 신호는 모노 서브밴드 신호이고,

상기 다운믹스 서브밴드 필터 계수들은 상기 다운믹스 서브밴드 신호에 대한 잔향부의 에너지 감소 특성을 반영하며,

상기 후기잔향 생성부는,

상기 모노 서브밴드 신호에 대한 디코릴레이션(decorrelation) 신호를 생성하는 디코릴레이터; 및

상기 모노 서브밴드 신호와 상기 디코릴레이션 신호 간의 가중치 합산을 수행하여 2채널의 좌, 우 신호를 생성하는 IC 매칭부;

를 더 포함하는 것을 특징으로 하는 오디오 신호 처리 장치.

발명의 설명

기술 분야

[0001] 본 발명은 오디오 신호를 효과적으로 재생하기 위한 신호 처리 방법 및 장치에 관한 것으로서, 더욱 상세하게는 멀티채널 혹은 멀티오브젝트 오디오 신호를 스테레오로 재생하기 위한 바이노럴 렌더링을 낮은 연산량으로 구현하기 위한 오디오 신호 처리 방법 및 장치에 관한 것이다.

배경 기술

[0002] 멀티채널 신호를 스테레오로 청취하기 위한 바이노럴 렌더링(binaural rendering)은 타겟 필터의 길이가 증가할수록 많은 연산량을 요구하는 문제가 있다. 특히, 녹음실의 특성을 반영한 BRIR(Binaural Room Impulse

Response) 필터를 이용하는 경우 그 길이는 48,000 ~ 96,000 샘플에 이르기도 한다. 여기에 22.2 채널 포맷과 같이 입력 채널 수가 증가하게 되면 그 연산량은 막대하다.

[0003] i 번째 채널의 입력 신호를 $x_i(n)$, 해당 채널의 좌, 우 BRIR 필터를 각각 $b_i^L(n)$, $b_i^R(n)$, 출력 신호를 $y^L(n)$, $y^R(n)$ 이라고 하면, 바이노럴 필터링(binaural filtering)은 다음과 같은 식으로 표현할 수 있다.

수학식 1

$$y^m(n) = \sum_i x_i(n) * b_i^m(n)$$

[0004]

[0005] 여기에서 $m \in \{L, R\}$ 이며, *는 컨볼루션(convolution)을 의미한다. 위의 시간-도메인 컨볼루션은 일반적으로 고속 푸리에 변환(Fast Fourier Transform, FFT)에 기반한 고속 컨볼루션(fast convolution)을 이용하여 수행된다. 고속 컨볼루션을 이용하여 바이노럴 렌더링을 수행하는 경우, 입력 채널수에 해당하는 횟수의 FFT와 출력 채널수에 해당하는 횟수의 역 고속 푸리에 변환(Inverse FFT)을 수행해야 한다. 게다가 멀티채널 오디오 코덱과 같이 실시간 재생 환경에서의 경우 딜레이를 고려해야 하기 때문에 블록 단위(block-wise)의 고속 컨볼루션을 수행해야 하며, 이는 전체 길이에 대하여 단순히 고속 컨볼루션을 수행했을 때보다 더 많은 연산량을 소모할 수 있다.

[0006] 그러나 대부분의 코딩 방식(coding scheme)들은 주파수 도메인에서 이루어지며, 일부 코딩 방식(이들테면, HE-AAC, USAC 등)의 경우 복호화 과정의 마지막 단계가 QMF 도메인에서 수행된다. 따라서 위의 수학식 1과 같이 바이노럴 필터링이 시간 도메인에서 수행될 경우 채널 수만큼의 QMF 합성(QMF synthesis)을 위한 연산이 추가적으로 필요하므로 매우 비효율적이다. 따라서 바이노럴 렌더링을 QMF 도메인에서 직접 수행할 경우 이득이 있다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 멀티채널 혹은 멀티오브젝트 신호를 스테레오로 재생함에 있어서, 원신호와 같은 입체감을 보존하기 위한 바이노럴 렌더링에서 많은 연산량을 필요로 하는 필터링 과정을 음질 손실을 최소화하면서도 매우 낮은 연산량으로 구현하기 위한 목적을 가지고 있다.

[0008] 또한, 본 발명은 입력 신호 자체에 왜곡이 있는 경우 고품질 필터를 통해 왜곡의 확산이 발생하는 것을 최소화하고자 하는 목적을 가지고 있다.

[0009] 또한, 본 발명은 매우 긴 길이를 갖는 FIR(Finite Impulse Response) 필터를 더 작은 길이의 필터로 구현하고자 하는 목적을 가지고 있다.

[0010] 또한, 본 발명은 축약된 FIR 필터를 이용한 필터링의 수행시, 누락된 필터 계수에 의해 손상된 부분의 왜곡을 최소화하고자 하는 목적을 가지고 있다.

과제의 해결 수단

[0011] 상기와 같은 과제를 해결하기 위해, 본 발명은 다음과 같은 오디오 신호 처리 방법 및 오디오 신호 처리 장치를 제공한다.

[0012] 먼저 본 발명은, 멀티채널 또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호를 수신하는 단계; 상기 멀티 오디오 신호의 필터링을 위한 절단된 서브밴드 필터 계수들을 수신하는 단계, 상기 절단된 서브밴드 필터 계수는 상기 멀티 오디오 신호의 바이노럴 필터링을 위한 BRIR(Binaural Room Impulse Response) 필터 계수로부터 획득된 서브밴드 필터 계수의 적어도 일 부분이며, 상기 절단된 서브밴드 필터 계수의 길이는 해당 서브밴드 필터 계수에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 결정되고, 적

어도 하나의 상기 절단된 서브밴드 필터 계수의 길이는 다른 서브밴드의 절단된 서브밴드 필터 계수의 길이와 다름; 및 상기 멀티 오디오 신호의 각 서브밴드 신호에 대응하는 상기 절단된 서브밴드 필터 계수를 이용하여 상기 서브밴드 신호를 필터링 하는 단계; 를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법을 제공한다.

[0013] 또한, 멀티채널 또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호에 대한 바이노럴 렌더링을 수행하기 위한 오디오 신호 처리 장치로서, 상기 멀티 오디오 신호는 각각 복수의 서브밴드 신호들을 포함하며, 상기 각 서브밴드 신호에 대한 직접음 및 초기 반사음 파트의 렌더링을 수행하기 위한 고속 콘볼루션부; 및 상기 각 서브밴드 신호에 대한 후기잔향 파트의 렌더링을 수행하기 위한 후기잔향 생성부를 포함하되, 상기 고속 콘볼루션부는, 상기 멀티 오디오 신호의 필터링을 위한 절단된 서브밴드 필터 계수들을 수신하되, 상기 절단된 서브밴드 필터 계수는 상기 멀티 오디오 신호의 바이노럴 필터링을 위한 BRIR(Binaural Room Impulse Response) 필터 계수로부터 획득된 서브밴드 필터 계수의 적어도 일 부분이며, 상기 절단된 서브밴드 필터 계수의 길이는 해당 서브밴드 필터 계수에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 결정되고, 적어도 하나의 상기 절단된 서브밴드 필터 계수의 길이는 다른 서브밴드의 절단된 서브밴드 필터 계수의 길이와 다름, 상기 멀티 오디오 신호의 각 서브밴드 신호에 대응하는 상기 절단된 서브밴드 필터 계수를 이용하여 상기 서브밴드 신호를 필터링 하는 것을 특징으로 하는 오디오 신호 처리 장치를 제공한다.

[0014] 이때, 상기 특성 정보는 해당 서브밴드 필터 계수의 제1 잔향 시간 정보를 포함하며, 상기 필터 차수 정보는 각 서브밴드 별로 하나의 값을 갖는 것을 특징으로 한다.

[0015] 또한, 상기 절단된 서브밴드 필터의 길이는 2의 거듭 제곱의 배수 값을 갖는 것을 특징으로 한다.

[0016] 본 발명의 실시예에 따르면, 상기 복수의 서브밴드 필터 계수들 및 복수의 서브밴드 신호들은 각각 기 설정된 주파수 밴드를 기준으로 한 저주파수의 제1 서브밴드 그룹과 고주파수의 제2 서브밴드 그룹을 포함하며, 상기 필터링 하는 단계는 상기 제1 서브밴드 그룹의 절단된 서브밴드 필터 계수들 및 서브밴드 신호들에 대하여 수행되는 것을 특징으로 한다.

[0017] 또한 본 발명의 실시예에 따르면, 상기 필터링 하는 단계는 해당 서브밴드 필터 계수의 제1 잔향 시간 정보에 적어도 부분적으로 기초하여 절단된 프론트 서브밴드 필터 계수를 이용하여 수행되며, 상기 서브밴드 필터 계수 중 상기 프론트 서브밴드 필터 계수 이후의 구간에 대응하는 상기 서브밴드 신호의 잔향 처리 단계를 더 포함하는 것을 특징으로 한다.

[0018] 이때, 상기 잔향 처리 단계는, 각 서브밴드에 대한 다운믹스 서브밴드 필터 계수를 수신하는 단계, 상기 다운믹스 서브밴드 필터 계수는 해당 서브밴드의 각 채널 별 또는 각 오브젝트 별 리어 서브밴드 필터 계수들을 조합하여 생성되며, 상기 리어 서브밴드 필터 계수는 해당 서브밴드 필터 계수 중 상기 프론트 서브밴드 필터 계수 이후의 구간으로부터 획득됨; 상기 각 서브밴드에 대한 다운믹스 서브밴드 신호를 생성하는 단계, 상기 다운믹스 서브밴드 신호는 해당 서브밴드의 각 채널 별 또는 각 오브젝트 별 서브밴드 신호들을 다운믹스 하여 생성됨; 및 상기 다운믹스 서브밴드 신호와 이에 대응하는 상기 다운믹스 서브밴드 필터 계수를 이용하여 2채널의 좌, 우 서브밴드 잔향 신호를 생성하는 단계를 포함하는 것을 특징으로 한다.

[0019] 본 발명의 일 실시예에 따르면, 상기 다운믹스 서브밴드 신호는 모노 서브밴드 신호이고, 상기 다운믹스 서브밴드 필터 계수는 해당 서브밴드 신호에 대한 잔향부의 에너지 감소 특성을 반영하며, 상기 필터링 된 모노 서브밴드 신호에 대한 디코릴레이션(decorrelation) 신호를 생성하는 단계; 및 상기 필터링 된 모노 서브밴드 신호와 상기 디코릴레이션 신호 간의 가중치 합산을 수행하여 2채널의 좌, 우 신호를 생성하는 단계를 더 포함하는 것을 특징으로 한다.

[0020] 본 발명의 다른 실시예에 따르면, 멀티채널 또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호를 수신하는 단계, 상기 멀티 오디오 신호는 각각 복수의 서브밴드 신호들을 포함하며, 상기 복수의 서브밴드 신호들은 기 설정된 주파수 밴드를 기준으로 한 저주파수의 제1 서브밴드 그룹의 신호와 고주파수의 제2 서브밴드 그룹의 신호를 포함함; 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 적어도 하나의 파라미터를 수신하는 단계, 상기 적어도 하나의 파라미터는 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 BRIR(Binaural Room Impulse Response) 서브밴드 필터 계수로부터 추출됨; 상기 수신된 파라미터를 이용하여 상기 제2 서브밴드 그룹의 서브밴드 신호에 대하여 탭-딜레이 라인 필터링을 수행하는 단계; 를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법을 제공한다.

[0021] 또한, 멀티채널 또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호에 대한 바이노럴 렌더링을 수행하기 위한

오디오 신호 처리 장치로서, 상기 멀티 오디오 신호는 각각 복수의 서브밴드 신호들을 포함하며, 상기 복수의 서브밴드 신호들은 기 설정된 주파수 밴드를 기준으로 한 저주파수의 제1 서브밴드 그룹의 신호와 고주파수의 제2 서브밴드 그룹의 신호를 포함하고, 상기 제1 서브밴드 그룹의 각 서브밴드 신호에 대한 렌더링을 수행하기 위한 고속 콘볼루션부; 및 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대한 렌더링을 수행하기 위한 탭-딜레이 라인 프로세싱부를 포함하되, 상기 탭-딜레이 라인 프로세싱부는, 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 적어도 하나의 파라미터를 수신하되, 상기 적어도 하나의 파라미터는 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 BRIR(Binaural Room Impulse Response) 서브밴드 필터 계수로부터 추출되고, 상기 수신된 파라미터를 이용하여 상기 제2 서브밴드 그룹의 서브밴드 신호에 대하여 탭-딜레이 라인 필터링을 수행하는 것을 특징으로 하는 오디오 신호 처리 장치를 제공한다.

[0022] 이때, 상기 파라미터는 해당 BRIR 서브밴드 필터 계수에 대한 하나의 딜레이 정보 및 상기 딜레이 정보에 대응하는 하나의 게인 정보를 포함하는 것을 특징으로 한다.

[0023] 본 발명의 실시예에 따르면, 상기 탭-딜레이 라인 필터링은 상기 파라미터를 이용한 원-탭-딜레이 라인 필터링인 것을 특징으로 한다.

[0024] 또한, 상기 딜레이 정보는 상기 BRIR 서브밴드 필터 계수 중 최대 피크에 대한 위치 정보를 나타내는 것으로 한다.

[0025] 또한, 상기 딜레이 정보는 QMF 도메인에서 샘플 단위의 정수 값을 갖는 것을 특징으로 한다.

[0026] 또한, 상기 게인 정보는 복소수 값을 갖는 것을 특징으로 한다.

[0027] 본 발명의 일 실시예에 따르면, 상기 필터링 된 멀티 오디오 신호를 각 서브밴드 별로 2채널의 좌, 우 서브밴드 신호로 합산하는 단계; 상기 합산된 좌, 우 서브밴드 신호를 상기 제1 서브밴드 그룹의 멀티 오디오 신호로부터 생성된 좌, 우 서브밴드 신호와 결합하는 단계; 및 상기 결합된 좌, 우 서브밴드 신호를 각각 QMF 합성하는 단계; 를 더 포함하는 것을 특징으로 한다.

[0028] 본 발명의 또 다른 실시예에 따르면, 복수의 서브밴드를 갖는 멀티미디어 신호를 수신하는 단계; 상기 멀티미디어 신호의 각 서브밴드 신호의 필터링을 위한 적어도 하나의 원형(proto-type) 렌더링 필터 계수를 수신하는 단계; 상기 원형 렌더링 필터 계수를 복수의 서브밴드 필터 계수들로 변환하는 단계; 상기 각 서브밴드 필터 계수를 해당 서브밴드 필터 계수에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 절단하는 단계, 적어도 하나의 상기 절단된 서브밴드 필터 계수의 길이는 다른 서브밴드의 절단된 서브밴드 필터 계수의 길이와 다름; 및 상기 각 서브밴드 신호에 대응하는 상기 절단된 서브밴드 필터 계수를 이용하여 상기 멀티미디어 신호를 필터링하는 단계; 를 포함하는 것을 특징으로 하는 멀티미디어 신호 처리 방법을 제공한다.

[0029] 또한, 복수의 서브밴드를 갖는 멀티미디어 신호 처리 장치로서, 상기 멀티미디어 신호의 각 서브밴드 신호의 필터링을 위한 적어도 하나의 원형(proto-type) 렌더링 필터 계수를 수신하고, 상기 원형 렌더링 필터 계수를 복수의 서브밴드 필터 계수들로 변환하고, 상기 각 서브밴드 필터 계수를 해당 서브밴드 필터 계수에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 절단하되, 적어도 하나의 상기 절단된 서브밴드 필터 계수의 길이는 다른 서브밴드의 절단된 서브밴드 필터 계수의 길이와 다름, 파라미터화부; 및 상기 멀티미디어 신호를 수신하고, 상기 각 서브밴드 신호에 대응하는 상기 절단된 서브밴드 필터 계수를 이용하여 상기 멀티미디어 신호를 필터링하는 렌더링 유닛;을 포함하는 것을 특징으로 하는 멀티미디어 신호 처리 장치를 제공한다.

[0030] 이때, 상기 멀티미디어 신호는 멀티채널 또는 멀티오브젝트 신호를 포함하며, 상기 원형 렌더링 필터 계수는 시간 도메인의 BRIR 필터 계수인 것을 특징으로 한다.

[0031] 또한, 상기 특성 정보는 해당 서브밴드 필터 계수의 에너지 감쇄 시간 정보를 포함하며, 상기 필터 차수 정보는 각 서브밴드 별로 하나의 값을 갖는 것을 특징으로 한다.

[0032] 본 발명의 또 다른 실시예에 따르면, 멀티채널 또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호를 수신하는 단계, 상기 멀티 오디오 신호는 각각 복수의 서브밴드 신호를 포함하며, 상기 복수의 서브밴드 신호는 기 설정된 주파수 밴드를 기준으로 한 저주파수의 제1 서브밴드 그룹의 신호와 고주파수의 제2 서브밴드 그룹의 신호를 포함함; 상기 제1 서브밴드 그룹의 상기 멀티 오디오 신호의 필터링을 위한 절단된 서브밴드 필터 계수들을 수신하는 단계, 상기 절단된 서브밴드 필터 계수는 상기 멀티 오디오 신호의 바이노럴 필터링을 위한

BRIR(Binaural Room Impulse Response) 필터 계수로부터 획득된 제1 서브밴드 그룹의 서브밴드 필터 계수의 적어도 일 부분이며, 상기 절단된 서브밴드 필터 계수의 길이는 해당 서브밴드 필터 계수에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 결정됨; 및 상기 절단된 서브밴드 필터 계수를 이용하여 상기 제1 서브밴드 그룹의 서브밴드 신호를 필터링 하는 단계; 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 적어도 하나의 파라미터를 수신하는 단계, 상기 적어도 하나의 파라미터는 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 서브밴드 필터 계수로부터 추출됨; 및 상기 수신된 파라미터를 이용하여 상기 제2 서브밴드 그룹의 서브밴드 신호에 대하여 탭-딜레이 라인 필터링을 수행하는 단계; 를 포함하는 것을 특징으로 하는 오디오 신호 처리 방법을 제공한다.

[0033] 또한, 멀티채널 또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호에 대한 바이노럴 렌더링을 수행하기 위한 오디오 신호 처리 장치로서, 상기 멀티 오디오 신호는 각각 복수의 서브밴드 신호들을 포함하며, 상기 복수의 서브밴드 신호들은 기 설정된 주파수 밴드를 기준으로 한 저주파수의 제1 서브밴드 그룹의 신호와 고주파수의 제2 서브밴드 그룹의 신호를 포함하고, 상기 제1 서브밴드 그룹의 각 서브밴드 신호에 대한 렌더링을 수행하기 위한 고속 콘볼루션부; 및 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대한 렌더링을 수행하기 위한 탭-딜레이 라인 프로세싱부를 포함하되, 상기 고속 콘볼루션부는, 상기 제1 서브밴드 그룹의 상기 멀티 오디오 신호의 필터링을 위한 절단된 서브밴드 필터 계수들을 수신하되, 상기 절단된 서브밴드 필터 계수는 상기 멀티 오디오 신호의 바이노럴 필터링을 위한 BRIR(Binaural Room Impulse Response) 필터 계수로부터 획득된 제1 서브밴드 그룹의 서브밴드 필터 계수의 적어도 일 부분이며, 상기 절단된 서브밴드 필터 계수의 길이는 해당 서브밴드 필터 계수에서 추출된 특성 정보를 적어도 부분적으로 이용하여 획득된 필터 차수 정보에 기초하여 결정되고, 상기 절단된 서브밴드 필터 계수를 이용하여 상기 제1 서브밴드 그룹의 서브밴드 신호를 필터링 하며, 상기 탭-딜레이 라인 프로세싱부는, 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 적어도 하나의 파라미터를 수신하되, 상기 적어도 하나의 파라미터는 상기 제2 서브밴드 그룹의 각 서브밴드 신호에 대응하는 BRIR(Binaural Room Impulse Response) 서브밴드 필터 계수로부터 추출되고, 상기 수신된 파라미터를 이용하여 상기 제2 서브밴드 그룹의 서브밴드 신호에 대하여 탭-딜레이 라인 필터링을 수행하는 것을 특징으로 하는 오디오 신호 처리 장치를 제공한다.

[0034] 이때, 상기 제1 서브밴드 그룹의 서브밴드 신호를 필터링하여 생성된 2채널의 좌, 우 서브밴드 신호와 상기 제2 서브밴드 그룹의 서브밴드 신호를 탭-딜레이 라인 필터링하여 생성된 2채널의 좌, 우 서브밴드 신호를 결합하는 단계; 및 상기 결합된 좌, 우 서브밴드 신호를 각각 QMF 합성하는 단계; 를 더 포함하는 것을 특징으로 한다.

발명의 효과

[0035] 본 발명의 실시예에 따르면, 멀티채널 혹은 멀티오브젝트 신호에 대한 바이노럴 렌더링의 수행시 음질 손실을 최소화 하면서 연산량을 획기적으로 낮출 수 있다.

[0036] 본 발명의 실시예에 따르면, 기존에 저전력 장치에서 실시간 처리가 불가능했던 멀티채널 혹은 멀티오브젝트 오디오 신호에 대한 고음질의 바이노럴 렌더링이 가능하도록 한다.

도면의 간단한 설명

- [0037] 도 1은 본 발명의 실시예에 따른 오디오 신호 디코더를 나타낸 블록도.
- 도 2는 본 발명의 일 실시예에 따른 바이노럴 렌더러의 각 구성을 나타낸 블록도.
- 도 3 내지 도 7은 본 발명에 따른 오디오 신호 처리 장치의 다양한 실시예들을 나타낸 도면.
- 도 8 내지 도 10은 본 발명의 실시예에 따른 바이노럴 렌더링을 위한 FIR 필터 생성 방법을 나타낸 도면.
- 도 11 내지 도 14는 본 발명의 P-파트 렌더링부의 다양한 실시예를 나타낸 도면.
- 도 15 및 도 16은 본 발명의 QTDL 프로세싱의 다양한 실시예를 나타낸 도면.

발명을 실시하기 위한 구체적인 내용

- [0038] **발명의 실시를 위한 최선의 형태**
- [0039] 본 명세서에서 사용되는 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어를 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도, 관례 또는 새로운 기술의 출현 등에 따라 달라질 수 있

다. 또한 특정 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 그 의미를 기재할 것이다. 따라서 본 명세서에서 사용되는 용어는, 단순한 용어의 명칭이 아닌 그 용어가 가진 실질적인 의미와 본 명세서의 전반에 걸친 내용을 토대로 해석되어야 함을 밝혀두고자 한다.

- [0040] 도 1은 본 발명의 실시예에 따른 오디오 신호 디코더를 나타낸 블록도이다. 본 발명의 오디오 신호 디코더는 코어 디코더(10), 렌더링 유닛(20), 믹서(30), 및 포스트 프로세싱 유닛(40)을 포함한다.
- [0041] 먼저, 코어 디코더(10)는 스피커(loudspeaker) 채널 신호, 개별(discrete) 오브젝트 신호, 오브젝트 다운믹스 신호 및 사전-렌더링된(pre-rendered) 신호 등을 복호화 한다. 일 실시예에 따르면 상기 코어 디코더(10)에서는 USAC(Unified Speech and Audio Coding) 기반의 코덱이 사용될 수 있다. 코어 디코더(10)는 수신된 비트스트림을 복호화 하여 렌더링 유닛(20)으로 전달한다.
- [0042] 렌더링 유닛(20)은 코어 디코더(10)에 의해 복호화 된 신호를 재생 레이아웃(reproduction layout) 정보를 이용하여 렌더링 한다. 렌더링 유닛(20)은 포맷 컨버터(22), 오브젝트 렌더러(24), OAM 디코더(25), SAOC 디코더(26) 및 HOA 디코더(28)를 포함할 수 있다. 렌더링 유닛(20)은 복호화 된 신호의 타입에 따라 상기 구성 중 어느 하나를 이용하여 렌더링을 수행한다.
- [0043] 포맷 컨버터(22)는 전송된 채널 신호를 출력 스피커 채널 신호로 변환한다. 즉, 포맷 컨버터(22)는 전송된 채널 구성(configuration)과 재생될 스피커 채널 구성 간의 변환을 수행한다. 만약, 출력 스피커 채널의 개수(이들테면, 5.1 채널)가 전송된 채널의 개수(이들테면, 22.2 채널)보다 적거나 전송된 채널 구성과 재생될 채널 구성이 다를 경우, 포맷 컨버터(22)는 전송된 채널 신호에 대한 다운믹스를 수행한다. 본 발명의 오디오 신호 디코더는 입력 채널 신호와 출력 스피커 채널 신호간의 조합을 이용하여 최적의 다운믹스 매트릭스를 생성하고, 상기 매트릭스를 이용하여 다운믹스를 수행할 수 있다. 본 발명의 실시예에 따르면, 포맷 컨버터(22)가 처리하는 채널 신호에는 사전-렌더링된 오브젝트 신호가 포함될 수 있다. 일 실시예에 따르면, 오디오 신호의 부호화 전에 적어도 하나의 오브젝트 신호가 사전-렌더링되어 채널 신호에 믹스(mix)될 수 있다. 이와 같이 믹스 된 오브젝트 신호는 채널 신호와 함께 포맷 컨버터(22)에 의해 출력 스피커 채널 신호로 변환될 수 있다.
- [0044] 오브젝트 렌더러(24) 및 SAOC 디코더(26)는 오브젝트 기반의 오디오 신호에 대한 렌더링을 수행한다. 오브젝트 기반의 오디오 신호에는 개별 오브젝트 웨이브폼과 파라메트릭 오브젝트 웨이브폼이 포함될 수 있다. 개별 오브젝트 웨이브폼의 경우, 각 오브젝트 신호들은 모노포닉(monophonic) 웨이브폼으로 인코더에 제공되며, 인코더는 단일 채널 엘리먼트들(Single Channel Elements, SCEs)을 이용하여 각 오브젝트 신호들을 전송한다. 파라메트릭 오브젝트 웨이브폼의 경우, 복수의 오브젝트 신호들이 적어도 하나의 채널 신호로 다운믹스 되며, 각 오브젝트의 특징과 이들 간의 관계가 SAOC(Spatial Audio Object Coding) 파라미터로 표현된다. 오브젝트 신호들은 다운믹스 되어 코어 코덱으로 부호화되며, 이때 생성되는 파라메트릭 정보가 함께 디코더로 전송된다.
- [0045] 한편, 개별 오브젝트 웨이브폼 또는 파라메트릭 오브젝트 웨이브폼이 오디오 신호 디코더로 전송될 때, 이에 대응하는 압축된 오브젝트 메타데이터가 함께 전송될 수 있다. 오브젝트 메타데이터는 오브젝트 속성을 시간과 공간 단위로 양자화하여 3차원 공간에서의 각 오브젝트의 위치 및 이득값을 지정한다. 렌더링 유닛(20)의 OAM 디코더(25)는 압축된 오브젝트 메타데이터를 수신하고, 이를 복호화하여 오브젝트 렌더러(24) 및/또는 SAOC 디코더(26)로 전달한다.
- [0046] 오브젝트 렌더러(24)는 오브젝트 메타데이터를 이용하여 각 오브젝트 신호를 주어진 재생 포맷에 따라 렌더링한다. 이때, 각 오브젝트 신호는 오브젝트 메타데이터에 기초하여 특정 출력 채널들로 렌더링될 수 있다. SAOC 디코더(26)는 복호화된 SAOC 전송 채널들과 파라메트릭 정보로부터 오브젝트/채널 신호를 복원한다. 상기 SAOC 디코더(26)는 재생 레이아웃 정보와 오브젝트 메타데이터에 기초하여 출력 오디오 신호를 생성할 수 있다. 이와 같이 오브젝트 렌더러(24) 및 SAOC 디코더(26)는 오브젝트 신호를 채널 신호로 렌더링할 수 있다.
- [0047] HOA 디코더(28)는 HOA(Higher Order Ambisonics) 신호 및 HOA 부가 정보를 수신하고, 이를 복호화한다. HOA 디코더(28)는 채널 신호나 오브젝트 신호를 별도의 수학적 식으로 모델링하여 사운드씬(sound scene)을 생성한다. 생성된 사운드씬에서 스피커가 있는 공간상의 위치를 선택하면, 스피커 채널 신호로 렌더링이 수행될 수 있다.
- [0048] 한편, 도 1에는 도시되지 않았지만, 렌더링 유닛(20)의 각 구성요소로 오디오 신호가 전달될 때, 전처리 과정으로서 동적 범위 제어(Dynamic Range Control, DRC)가 수행될 수 있다. DRC는 재생되는 오디오 신호의 동적 범위를 일정 레벨로 제한하는 것으로, 기 설정된 쓰레쉬(threshold) 보다 작은 소리는 더 크게, 기 설정된 쓰레쉬 보다 큰 소리는 더 작게 조정 한다.
- [0049] 렌더링 유닛(20)에 의해 처리된 채널 기반의 오디오 신호 및 오브젝트 기반의 오디오 신호는 믹서(30)로 전달된

다. 믹서(30)는 채널 기반의 웨이브폼과 렌더링된 오브젝트 웨이브폼의 딜레이(delay)를 조정하고, 이를 샘플 단위로 합산한다. 믹서(30)에 의해 합산된 오디오 신호는 포스트 프로세싱 유닛(40)으로 전달된다.

[0050] 포스트 프로세싱 유닛(40)은 스피커 렌더러(100)와 바이노럴 렌더러(200)를 포함한다. 스피커 렌더러(100)는 믹서(30)로부터 전달된 멀티채널 및/또는 멀티오브젝트 오디오 신호를 출력하기 위한 포스트 프로세싱을 수행한다. 이러한 포스트 프로세싱에는 동적 범위 제어(DRC), 음량 정규화(Loudness Normalization, LN) 및 피크 제한(Peak Limiter, PL) 등이 포함될 수 있다.

[0051] 바이노럴 렌더러(200)는 멀티채널 및/또는 멀티오브젝트 오디오 신호의 바이노럴 다운믹스 신호를 생성한다. 바이노럴 다운믹스 신호는 각 입력 채널/오브젝트 신호가 3차원상에 위치한 가상의 음원에 의해 표현되도록 하는 2채널의 오디오 신호이다. 바이노럴 렌더러(200)는 스피커 렌더러(100)에 공급되는 오디오 신호를 입력 신호로서 수신할 수 있다. 바이노럴 렌더링은 BRIR(Binaural Room Impulse Response) 필터를 기초로 수행되며, 시간 도메인 또는 QMF 도메인 상에서 수행될 수 있다. 실시예에 따르면, 바이노럴 렌더링의 후처리 과정으로서 전술한 동적 범위 제어(DRC), 음량 정규화(LN) 및 피크 제한(PL) 등이 추가로 수행될 수 있다.

[0052] 도 2는 본 발명의 일 실시예에 따른 바이노럴 렌더러의 각 구성을 나타낸 블록도이다. 도시된 바와 같이, 본 발명의 실시예에 따른 바이노럴 렌더러(200)는 BRIR 파라미터화부(210), 고속 콘볼루션부(230), 후기잔향 생성부(240), QTDL 프로세싱부(250), 믹서&콤바이너(260)를 포함할 수 있다.

[0053] 바이노럴 렌더러(200)는 다양한 타입의 입력 신호에 대한 바이노럴 렌더링을 수행하여 3D 오디오 헤드폰 신호(즉, 3D 오디오 2채널 신호)를 생성한다. 이때, 입력 신호는 채널 신호(즉, 스피커 채널 신호), 오브젝트 신호 및 HOA 신호 중 적어도 하나를 포함하는 오디오 신호가 될 수 있다. 본 발명의 다른 실시예에 따르면, 바이노럴 렌더러(200)가 별도의 디코더를 포함할 경우, 상기 입력 신호는 전술한 오디오 신호의 부호화된 비트스트림이 될 수 있다. 바이노럴 렌더링은 복호화된 입력 신호를 바이노럴 다운믹스 신호로 변환하여, 헤드폰으로 청취시 서라운드 음향을 체험할 수 있도록 한다.

[0054] 본 발명의 실시예에 따르면, 바이노럴 렌더러(200)는 입력 신호에 대한 바이노럴 렌더링을 QMF 도메인 상에서 수행할 수 있다. 이를테면, 바이노럴 렌더러(200)는 QMF 도메인의 멀티채널(N channels) 신호를 수신하고, QMF 도메인의 BRIR 서브밴드 필터를 이용하여 상기 멀티채널 신호에 대한 바이노럴 렌더링을 수행할 수 있다. QMF

분석 필터뱅크를 통과한 i 번째 채널의 k 번째 서브밴드(subband) 신호를 $x_{k,i}(l)$, 서브밴드 도메인에서의 시간 인덱스를 l 이라고 하면, QMF 도메인에서의 바이노럴 렌더링은 다음과 같은 식으로 표현할 수 있다.

수학식 2

$$y_k^m(l) = \sum_i x_{k,i}(l) * b_{k,i}^m(l)$$

[0055]

여기서, $m \in \{L, R\}$ 이며, $b_{k,i}^m(l)$ 은 시간 도메인 BRIR 필터를 QMF 도메인의 서브밴드 필터로 변환한 것이다.

[0056]

즉, 바이노럴 렌더링은 QMF 도메인의 채널 신호 또는 오브젝트 신호를 복수의 서브밴드 신호로 나누고, 각 서브밴드 신호를 이에 대응하는 BRIR 서브밴드 필터와 콘볼루션 한 후 합산하는 방법으로 수행될 수 있다.

[0057]

[0058] BRIR 파라미터화부(210)는 QMF 도메인에서의 바이노럴 렌더링을 위해 BRIR 필터 계수를 변환 및 편집하고 각종 파라미터를 생성한다. 먼저, BRIR 파라미터화부(210)는 멀티채널 또는 멀티오브젝트에 대한 시간 도메인 BRIR 필터 계수를 수신하고, 이를 QMF 도메인 BRIR 필터 계수로 변환한다. 이때, QMF 도메인 BRIR 필터 계수는 복수의 주파수 밴드에 각각 대응하는 복수의 서브밴드 필터 계수들을 포함한다. 본 발명에서 서브밴드 필터 계수는 QMF 변환된 서브밴드 도메인의 각 BRIR 필터 계수를 가리킨다. 본 명세서에서 서브밴드 필터 계수는 BRIR 서브밴드 필터 계수로도 지칭될 수 있다. BRIR 파라미터화부(210)는 QMF 도메인의 복수의 BRIR 서브밴드 필터 계수를 각각 편집하고, 편집된 서브밴드 필터 계수를 고속 콘볼루션부(230) 등에 전달할 수 있다. 본 발명의 실시예에 따르면, BRIR 파라미터화부(210)는 바이노럴 렌더러(200)의 일 구성요소로 포함될 수도 있으며, 별도의 장치로 구비될 수도 있다. 일 실시예에 따르면, BRIR 파라미터화부(210)를 제외한 고속 콘볼루션부(230), 후기잔향

생성부(240), QTDL 프로세싱부(250), 믹서&컴바이너(260)를 포함하는 구성이 바이노럴 렌더링 유닛(220)으로 분류될 수 있다.

[0059] 일 실시예에 따르면, BRIR 파라미터화부(210)는 가상 재생 공간의 적어도 하나의 위치에 대응되는 BRIR 필터 계수를 입력으로 수신할 수 있다. 상기 가상 재생 공간의 각 위치는 멀티채널 시스템의 각 스피커 위치에 대응될 수 있다. 일 실시예에 따르면, BRIR 파라미터화부(210)가 수신한 각 BRIR 필터 계수는 바이노럴 렌더러(200)의 입력 신호의 각 채널 또는 각 오브젝트에 직접 매칭될 수 있다. 반면에, 본 발명의 다른 실시예에 따르면 상기 수신된 각 BRIR 필터 계수는 바이노럴 렌더러(200)의 입력 신호에 독립적인 구성(configuration)을 가질 수 있다. 즉, BRIR 파라미터화부(210)가 수신한 BRIR 필터 계수 중 적어도 일부는 바이노럴 렌더러(200)의 입력 신호에 직접 매칭되지 않을 수 있으며, 수신된 BRIR 필터 계수의 개수는 입력 신호의 채널 및/또는 오브젝트의 총 개수보다 작거나 클 수도 있다.

[0060] 본 발명의 실시예에 따르면, BRIR 파라미터화부(210)는 바이노럴 렌더러(200)의 입력 신호의 각 채널 또는 각 오브젝트에 대응하는 BRIR 필터 계수를 변환 및 편집하여 바이노럴 렌더링 유닛(220)으로 전달할 수 있다. 상기 대응하는 BRIR 필터 계수는 각 채널 또는 각 오브젝트에 대한 매칭 BRIR 또는 폴백(fallback) BRIR이 될 수 있다. BRIR 매칭은 가상 재생 공간상에서 각 채널 또는 각 오브젝트의 위치를 타겟으로 하는 BRIR 필터 계수가 존재하는지 여부에 따라 결정될 수 있다. 만약, 입력 신호의 각 채널 또는 각 오브젝트의 위치 중 적어도 하나를 타겟으로 하는 BRIR 필터 계수가 존재할 경우, 해당 BRIR 필터 계수는 입력 신호의 매칭 BRIR이 될 수 있다. 그러나 특정 채널 또는 오브젝트의 위치를 타겟으로 하는 BRIR 필터 계수가 존재하지 않을 경우, 바이노럴 렌더링 유닛(220)은 해당 채널 또는 오브젝트와 가장 유사한 위치를 타겟으로 하는 BRIR 필터 계수를 해당 채널 또는 오브젝트에 대한 폴백 BRIR로 제공할 수 있다.

[0061] 한편 본 발명의 다른 실시예에 따르면, BRIR 파라미터화부(210)는 수신된 BRIR 필터 계수 전체를 변환 및 편집하여 바이노럴 렌더링 유닛(220)으로 전달할 수 있다. 이때, 입력 신호의 각 채널 또는 각 오브젝트에 대응하는 BRIR 필터 계수(또는, 편집된 BRIR 필터 계수)의 선별 작업은 바이노럴 렌더링 유닛(220)에서 수행될 수 있다.

[0062] 바이노럴 렌더링 유닛(220)은 고속 콘볼루션부(230), 후기잔향 생성부(240) 및 QTDL 프로세싱부(250)를 포함하며, 멀티채널 및/또는 멀티오브젝트 신호를 포함하는 멀티 오디오 신호를 수신한다. 본 명세서에서는 멀티채널 및/또는 멀티오브젝트 신호를 포함하는 입력 신호를 멀티 오디오 신호로 지칭하기로 한다. 도 2에서는 일 실시예에 따라 바이노럴 렌더링 유닛(220)이 QMF 도메인의 멀티채널 신호를 수신하는 것으로 도시되어 있으나, 바이노럴 렌더링 유닛(220)의 입력 신호에는 시간 도메인 멀티채널 신호 및 멀티오브젝트 신호 등이 포함될 수 있다. 또한, 바이노럴 렌더링 유닛(220)이 별도의 디코더를 추가적으로 포함할 경우, 상기 입력 신호는 상기 멀티 오디오 신호의 부호화된 비트스트림이 될 수 있다. 이에 더하여, 본 명세서에서는 멀티 오디오 신호에 대한 BRIR 렌더링을 수행하는 케이스를 기준으로 본 발명을 설명하지만, 본 발명은 이에 한정되지 않는다. 즉, 본 발명에서 제공하는 특징들은 BRIR이 아닌 다른 종류의 렌더링 필터에도 적용될 수 있으며, 멀티 오디오 신호가 아닌 단일 채널 또는 단일 오브젝트의 오디오 신호에 대해서도 적용될 수 있다.

[0063] 고속 콘볼루션부(230)는 입력 신호와 BRIR 필터간의 고속 콘볼루션을 수행하여 입력 신호에 대한 직접음(direct sound)과 초기 반사음(early reflection)을 처리한다. 이를 위해, 고속 콘볼루션부(230)는 절단된(truncated) BRIR을 사용하여 고속 콘볼루션을 수행할 수 있다. 절단된 BRIR은 각 서브밴드 주파수에 종속적으로 절단된 복수의 서브밴드 필터 계수를 포함하며, BRIR 파라미터화부(210)에서 생성된다. 이때, 각 절단된 서브밴드 필터 계수의 길이는 해당 서브밴드의 주파수에 종속적으로 결정된다. 고속 콘볼루션부(230)는 서브밴드에 따라 서로 다른 길이를 갖는 절단된 서브밴드 필터 계수를 이용함으로써 주파수 도메인에서의 가변차수(variable order) 필터링을 수행할 수 있다. 즉, 각 주파수 밴드 별로 QMF 도메인 서브밴드 오디오 신호와 이에 대응하는 QMF 도메인의 절단된 서브밴드 필터들 간의 고속 콘볼루션이 수행될 수 있다. 본 명세서에서 직접음 및 초기 반사음(Direct sound & Early reflection, D&E) 파트는 F(front)-파트로 지칭될 수 있다.

[0064] 후기잔향 생성부(240)는 입력 신호에 대한 후기잔향(late reverberation) 신호를 생성한다. 후기잔향 신호는 고속 콘볼루션부(230)에서 생성된 직접음 및 초기 반사음 이후의 출력 신호를 나타낸다. 후기잔향 생성부(240)는 BRIR 파라미터화부(210)로부터 전달된 각 서브밴드 필터 계수로부터 결정된 잔향 시간 정보에 기초하여 입력 신호를 처리할 수 있다. 본 발명의 실시예에 따르면, 후기잔향 생성부(240)는 입력 오디오 신호에 대한 모노 또는 스테레오 다운믹스 신호를 생성하고, 생성된 다운믹스 신호에 대한 후기잔향 처리를 수행할 수 있다. 본 명세서에서 후기잔향(Late Reverberation, LR) 파트는 P(parametric)-파트로 지칭될 수 있다.

[0065] QTDL(QMF domain Tapped Delay Line) 프로세싱부(250)는 입력 오디오 신호 중 고 주파수 밴드의 신호를 처리한

다. QTDL 프로세싱부(250)는 고 주파수 밴드의 각 서브밴드 신호에 대응하는 적어도 하나의 파라미터를 BRIR 파라미터화부(210)로부터 수신하고, 수신된 파라미터를 이용하여 QMF 도메인에서 탭-딜레이 라인 필터링을 수행한다. 본 발명의 실시예에 따르면, 바이노럴 렌더러(200)는 기 설정된 상수 또는 기 설정된 주파수 밴드를 기초로 입력 오디오 신호를 저 주파수 밴드 신호와 고 주파수 밴드 신호로 분리하고, 저 주파수 밴드 신호는 고속 콘볼루션부(230) 및 후기잔향 생성부(240)에서, 고 주파수 밴드 신호는 QTDL 프로세싱부(250)에서 각각 처리할 수 있다.

[0066] 고속 콘볼루션부(230), 후기잔향 생성부(240) 및 QTDL 프로세싱부(250)는 각각 2채널의 QMF 도메인 서브밴드 신호를 출력한다. 믹서&콤바이너(260)는 고속 콘볼루션부(230)의 출력 신호, 후기잔향 생성부(240)의 출력 신호 및 QTDL 프로세싱부(250)의 출력 신호를 결합하여 믹싱을 수행한다. 이때, 출력 신호의 결합은 2채널의 좌, 우 출력 신호에 대해 각각 별도로 수행된다. 바이노럴 렌더러(200)는 결합된 출력 신호를 QMF 합성하여 시간 도메인의 최종 출력 오디오 신호를 생성한다.

[0067] 이하, 각 도면을 참조로 하여 도 2의 고속 콘볼루션부(230), 후기잔향 생성부(240), QTDL 프로세싱부(250) 및 이들의 조합에 대한 다양한 실시예들을 구체적으로 설명하도록 한다.

[0068] 도 3 내지 도 7은 본 발명에 따른 오디오 신호 처리 장치의 다양한 실시예들을 나타내고 있다. 본 발명에서 오디오 신호 처리 장치는 협의의 의미로는 도 2에 도시된 바이노럴 렌더러(200) 또는 바이노럴 렌더링 유닛(220)을 가리킬 수 있다. 그러나 본 발명에서 오디오 신호 처리 장치는 광의의 의미로는 바이노럴 렌더러를 포함하는 도 1의 오디오 신호 디코더를 가리킬 수 있다. 도 3 내지 도 7에 도시된 각 바이노럴 렌더러는 설명의 편의를 위해 도 2에 도시된 바이노럴 렌더러(200)의 일부 구성만을 나타낼 수 있다. 또한, 이하 본 명세서에서는 멀티채널 입력 신호에 대한 실시예를 주로 기술할 수 있으나, 별도의 언급이 없을 경우 채널, 멀티채널 및 멀티채널 입력 신호는 각각 오브젝트, 멀티오브젝트 및 멀티오브젝트 입력 신호를 포함하는 개념으로 사용될 수 있다. 뿐만 아니라, 멀티채널 입력 신호는 HOA 디코딩 및 렌더링된 신호를 포함하는 개념으로도 사용될 수 있다.

[0069] 도 3은 본 발명의 일 실시예에 따른 바이노럴 렌더러(200A)를 나타내고 있다. BRIR을 이용한 바이노럴 렌더링을 일반화하면 M개의 채널을 갖는 멀티채널의 입력 신호에 대해 0개의 출력신호를 얻기 위한 M-to-0 프로세싱이다. 바이노럴 필터링은 이 과정에서 각각의 입력 채널과 출력 채널에 대응되는 필터 계수를 이용한 필터링으로 볼 수 있다. 도 3에서 원본 필터 셋 H는 각 채널 신호의 스피커 위치에서부터 좌, 우 귀의 위치까지의 전달함수들을 의미한다. 이러한 전달함수 중 일반적인 청음공간, 즉 잔향이 있는 공간에서 측정된 것을 Binaural Room Impulse Response(BRIR)라 부른다. 반면 재생 공간의 영향이 없도록 무향실에서 측정된 것을 Head Related Impulse Response(HRIR)이라고 하며, 이에 대한 전달함수를 Head Related Transfer Function(HRTF)라 부른다. 따라서, BRIR은 HRTF와는 다르게 방향 정보뿐만 아니라 재생 공간의 정보를 함께 담고 있다. 일 실시예에 따르면, HRTF와 인공 잔향기(artificial reverberator)를 이용하여 BRIR을 대체할 수도 있다. 본 명세서에서는 BRIR을 이용한 바이노럴 렌더링에 대하여 설명하지만, 본 발명은 이에 한정되지 않으며 다양한 형태의 FIR 필터를 이용한 바이노럴 렌더링에도 동일하게 적용 가능하다. 한편, BRIR은 전술한 바와 같이 96K개의 샘플 길이를 가질 수 있으며, 멀티 채널 바이노럴 렌더링은 M*0개의 서로 다른 필터를 이용하여 수행되므로 고 연산량의 처리 과정이 요구된다.

[0070] 본 발명의 실시예에 따르면, BRIR 파라미터화부(210)는 연산량 최적화를 위해 원본 필터 셋 H로부터 변형된 필터 계수들을 생성할 수 있다. BRIR 파라미터화부(210)는 원본 필터 계수를 F(front)-파트 계수와 P(parametric)-파트 계수로 분리한다. 여기서, F-파트는 직접음 및 초기 반사음(D&E) 파트를 나타내고, P-파트는 후기잔향(LR) 파트를 나타낸다. 예를 들어, 96K 샘플 길이를 갖는 원본 필터 계수는 앞의 4K 샘플까지만을 절단한 F-파트와, 나머지 92K 샘플에 대응되는 부분인 P-파트로 각각 분리될 수 있다.

[0071] 바이노럴 렌더링 유닛(220)은 BRIR 파라미터화부(210)로부터 F-파트 계수 및 P-파트 계수를 각각 수신하고, 이를 이용하여 멀티채널 입력 신호를 렌더링 한다. 본 발명의 실시예에 따르면, 도 2에 도시된 고속 콘볼루션부(230)는 BRIR 파라미터화부(210)로부터 수신된 F-파트 계수를 이용하여 멀티 오디오 신호를 렌더링 하고, 후기 잔향 생성부(240)는 BRIR 파라미터화부(210)로부터 수신된 P-파트 계수를 이용하여 멀티 오디오 신호를 렌더링 할 수 있다. 즉, 고속 콘볼루션부(230)와 후기잔향 생성부(240)는 각각 본 발명의 F-파트 렌더링부와 P-파트 렌더링부에 대응될 수 있다. 일 실시예에 따르면, F-파트 렌더링(F-파트 계수를 이용한 바이노럴 렌더링)은 통상적인 FIR(Finite Impulse Response) 필터로 구현되고, P-파트 렌더링(P-파트 계수를 이용한 바이노럴 렌더링)은 파라메트릭한 방법으로 구현될 수 있다. 한편, 유저 또는 제어 시스템에 의해 제공되는 복잡도-품질 제어 입력은 F-파트 및/또는 P-파트로 생성되는 정보를 결정하는데 사용될 수 있다.

- [0072] 도 4는 본 발명의 다른 실시예에 따른 바이노럴 렌더러(200B)로서, F-파트 렌더링을 구현하는 보다 상세한 방법을 도시하고 있다. 설명의 편의를 위해 도 4에서 P-파트 렌더링부는 생략되었다. 또한, 도 4에서는 QMF 도메인에서 구현된 필터를 나타내고 있으나, 본 발명은 이에 한정하지 않으며 다른 도메인의 서브밴드 프로세싱에 모두 적용 가능하다.
- [0073] 도 4를 참조하면, F-파트 렌더링은 QMF 도메인 상에서 고속 콘볼루션부(230)에 의해 수행될 수 있다. QMF 도메인 상에서의 렌더링을 위해 QMF 분석부(222)는 시간 도메인 입력 신호 x_0, x_1, \dots, x_{M-1} 을 QMF 도메인 신호 X_0, X_1, \dots, X_{M-1} 으로 변환한다. 이때, 입력신호 x_0, x_1, \dots, x_{M-1} 은 멀티채널 오디오 신호, 이를테면 22.2 채널 스피커에 대응되는 채널 신호일 수 있다. QMF 도메인은 총 64개의 서브밴드를 사용할 수 있으나, 본 발명은 이에 한정하지 않는다. 한편, 본 발명의 일 실시예에 따르면 QMF 분석부(222)는 바이노럴 렌더러(200B)에서 생략될 수 있다. SBR(Spectral Band Replication)을 사용하는 HE-AAC나 USAC의 경우 QMF 도메인에서 프로세싱을 수행하므로, 바이노럴 렌더러(200B)는 QMF 분석 없이 바로 QMF 도메인 신호 X_0, X_1, \dots, X_{M-1} 을 입력으로 수신할 수 있다. 따라서, 이와 같이 QMF 도메인 신호를 직접 입력으로 수신하는 경우, 본 발명에 따른 바이노럴 렌더러에서 이용하는 QMF는 이전 처리부(이를테면, SBR)에서 사용하는 QMF와 동일한 것을 특징으로 한다. QMF 합성부(244)는 바이노럴 렌더링이 수행된 2채널의 좌, 우 신호 Y_L, Y_R 을 QMF 합성하여 시간 도메인의 2채널 출력 오디오 신호 y_L, y_R 을 생성한다.
- [0074] 도 5 내지 도 7은 각각 F-파트 렌더링과 P-파트 렌더링을 함께 수행하는 바이노럴 렌더러(200C, 200D, 200E)의 실시예를 나타내고 있다. 도 5 내지 도 7의 실시예에서 F-파트 렌더링은 QMF 도메인 상에서 고속 콘볼루션부(230)에 의해 수행되며, P-파트 렌더링은 QMF 도메인 또는 시간 도메인 상에서 후기잔향 생성부(240)에 의해 수행된다. 도 5 내지 도 7의 실시예에서 이전 도면의 실시예와 중복되는 부분은 구체적인 설명을 생략하도록 한다.
- [0075] 도 5를 참조하면, 바이노럴 렌더러(200C)는 F-파트 렌더링 및 P-파트 렌더링을 모두 QMF 도메인에서 수행할 수 있다. 즉, 바이노럴 렌더러(200C)의 QMF 분석부(222)는 시간 도메인 입력 신호 x_0, x_1, \dots, x_{M-1} 을 QMF 도메인 신호 X_0, X_1, \dots, X_{M-1} 으로 변환하여 각각 고속 콘볼루션부(230) 및 후기잔향 생성부(240)로 전달한다. 고속 콘볼루션부(230) 및 후기잔향 생성부(240)는 QMF 도메인 신호 X_0, X_1, \dots, X_{M-1} 을 렌더링하여 각각 2채널의 출력 신호 Y_L, Y_R 및 Y_{Lp}, Y_{Rp} 를 생성한다. 이때, 고속 콘볼루션부(230) 및 후기잔향 생성부(240)는 BRIR 파라미터화부(210)에서 수신한 F-파트 필터 계수 및 P-파트 필터 계수를 각각 이용하여 렌더링을 수행할 수 있다. F-파트 렌더링의 출력 신호 Y_L, Y_R 과 P-파트 렌더링의 출력 신호 Y_{Lp}, Y_{Rp} 는 믹서&컴바이너(260)에서 좌, 우 채널 별로 결합되어 QMF 합성부(224)로 전달된다. QMF 합성부(224)는 입력된 2채널의 좌, 우 신호를 QMF 합성하여 시간 도메인의 2채널 출력 오디오 신호 y_L, y_R 을 생성한다.
- [0076] 도 6을 참조하면, 바이노럴 렌더러(200D)는 QMF 도메인에서 F-파트 렌더링을, 시간 도메인에서 P-파트 렌더링을 각각 수행할 수 있다. 바이노럴 렌더러(200D)의 QMF 분석부(222)는 시간 도메인 입력 신호를 QMF 변환하여 고속 콘볼루션부(230)로 전달한다. 고속 콘볼루션부(230)는 QMF 도메인 신호를 F-파트 렌더링하여 2채널의 출력 신호 Y_L, Y_R 을 생성한다. QMF 합성부(224)는 F-파트 렌더링의 출력 신호를 시간 도메인 출력 신호로 변환하여 믹서&컴바이너(260)로 전달한다. 한편, 후기잔향 생성부(240)는 시간 도메인 입력 신호를 직접 수신하여 P-파트 렌더링을 수행한다. P-파트 렌더링의 출력 신호 y_{Lp}, y_{Rp} 는 믹서&컴바이너(260)로 전달된다. 믹서&컴바이너(260)는 시간 도메인 상에서 F-파트 렌더링 출력 신호 및 P-파트 렌더링 출력 신호를 각각 결합하여 시간 도메인의 2채널 출력 오디오 신호 y_L, y_R 을 생성한다.
- [0077] 도 5와 도 6의 실시예에서 F-파트 렌더링 및 P-파트 렌더링이 각각 병렬적(parallel)으로 수행된 반면, 도 7의 실시예에 따르면 바이노럴 렌더러(200E)는 F-파트 렌더링과 P-파트 렌더링을 각각 순차적(sequential)으로 수행할 수 있다. 즉, 고속 콘볼루션부(230)는 QMF 변환된 입력 신호를 F-파트 렌더링하고, F-파트 렌더링된 2채널 신호 Y_L, Y_R 은 QMF 합성부(224)에서 시간 도메인 신호로 변환된 뒤 후기잔향 생성부(240)로 전달될 수 있다. 후기잔향 생성부(240)는 입력된 2채널 신호에 대한 P-파트 렌더링을 수행하여 시간 도메인의 2채널 출력 오디오 신호 y_L, y_R 을 생성한다.
- [0078] 도 5 내지 도 7은 각각 F-파트 렌더링과 P-파트 렌더링을 수행하는 일 실시예를 나타낸 것이며, 각 도면의 실시예들을 조합 또는 변형하여 바이노럴 렌더링을 수행할 수 있다. 이를테면, 각 실시예에서 바이노럴 렌더러는 입력된 멀티 오디오 신호 각각에 대해 개별적으로 P-파트 렌더링을 수행할 수도 있지만, 입력 신호를 2채널의 좌, 우 신호 또는 모노 신호로 다운믹스 한 후 다운믹스 신호에 대한 P-파트 렌더링을 수행할 수도 있다.

- [0079] <주파수 도메인 가변차수 필터링(Variable Order Filtering in Frequency-domain, VOFF)>
- [0080] 도 8 내지 도 10은 본 발명의 실시예에 따른 바이노럴 렌더링을 위한 FIR 필터 생성 방법을 나타내고 있다. 본 발명의 실시예에 따르면, QMF 도메인에서의 바이노럴 렌더링을 위해, QMF 도메인의 복수의 서브밴드 필터로 변환된 FIR 필터가 사용될 수 있다. 이때, F-파트 렌더링에는 각 서브밴드 주파수에 종속적으로 절단된 서브밴드 필터들이 사용될 수 있다. 즉, 바이노럴 렌더러의 고속 콘볼루션부는 서브밴드에 따라 서로 다른 길이를 갖는 절단된 서브밴드 필터를 이용함으로써 QMF 도메인에서의 가변차수 필터링을 수행할 수 있다. 이하, 설명되는 도 8 내지 도 10의 필터 생성 실시예는 도 2의 BRIR 파라미터화부(210)에 의해 수행될 수 있다.
- [0081] 도 8은 바이노럴 렌더링에 사용되는 QMF 도메인 필터의 각 QMF 밴드에 따른 길이의 일 실시예를 나타내고 있다. 도 8의 실시예에서 FIR 필터는 I개의 QMF 서브밴드 필터로 변환되며, F_i 는 QMF 서브밴드 i 의 절단된 서브밴드 필터를 나타낸다. QMF 도메인은 총 64개의 서브밴드를 사용할 수 있으나, 본 발명은 이에 한정하지 않는다. 또한, N 은 원본 서브밴드 필터의 길이(탭 수)를 나타내며, 절단된 서브밴드 필터의 길이는 각각 N_1 , N_2 , N_3 로 표현되었다. 이때, 길이 N , N_1 , N_2 및 N_3 는 다운 샘플된 QMF 도메인에서의 탭 수를 나타낸다.
- [0082] 본 발명의 실시예에 따르면, 각 서브밴드에 따라 서로 다른 길이(N_1 , N_2 , N_3)를 갖는 절단된 서브밴드 필터가 F-파트 렌더링에 사용될 수 있다. 이때, 절단된 서브밴드 필터는 원본 서브밴드 필터에서 절단된 앞부분(front)의 필터이며, 프론트 서브밴드 필터로도 지칭될 수 있다. 또한, 원본 서브밴드 필터의 절단 이후의 뒷부분(rear)은 리어 서브밴드 필터로도 지칭될 수 있으며, P-파트 렌더링에 이용될 수 있다.
- [0083] BRIR 필터를 이용한 렌더링의 경우, 각 서브밴드 별 필터 차수(즉, 필터 길이)는 원본 BRIR 필터로부터 추출된 파라미터들 이를 테면, 각 서브밴드 필터 별 잔향 시간(Reverberation Time, RT) 정보, EDC(Energy Decay Curve) 값, 에너지 감쇄 시간 정보 등에 기초하여 결정될 수 있다. 각 주파수 별로 공기 중에서의 감쇄, 벽 및 천장의 재질에 따른 흡음 정도가 다른 음향적 특성으로 인해, 잔향 시간은 주파수에 따라 서로 달라질 수 있다. 일반적으로는 낮은 주파수의 신호일수록 잔향 시간이 긴 특성을 갖는다. 잔향 시간이 길면 FIR 필터의 뒷부분에 많은 정보가 남아 있음을 의미하므로, 해당 필터를 길게 절단하여 사용하는 것이 잔향 정보를 제대로 전달하는데 바람직하다. 따라서, 본 발명의 각 절단된 서브밴드 필터의 길이는 해당 서브밴드 필터에서 추출된 특성 정보(이를테면, 잔향 시간 정보)에 적어도 부분적으로 기초하여 결정된다.
- [0084] 절단된 서브밴드 필터의 길이는 다양한 실시예에 따라 결정될 수 있다. 먼저 일 실시예에 따르면, 각 서브밴드는 복수의 그룹으로 분류되고, 분류된 그룹에 따라 각 절단된 서브밴드 필터의 길이가 결정될 수 있다. 도 8의 예시에 따르면, 각 서브밴드는 3개의 구역(Zone 1, Zone 2, Zone 3)으로 분류될 수 있는데, 저 주파수에 해당하는 Zone 1의 절단된 서브밴드 필터들은 고 주파수에 해당하는 Zone 2 및 Zone 3의 절단된 서브밴드 필터들보다 더 긴 필터 차수(즉, 필터 길이)를 가질 수 있다. 또한, 고 주파수의 구역으로 감에 따라 해당 구역의 절단된 서브밴드 필터의 필터 차수는 점점 줄어들 수 있다.
- [0085] 본 발명의 다른 실시예에 따르면, 각 절단된 서브밴드 필터의 길이는 원본 서브밴드 필터의 특성 정보에 따라 각 서브밴드 별로 독립적 및 가변적으로 결정될 수 있다. 각 절단된 서브밴드 필터의 길이는 해당 서브밴드에서 결정된 절단 길이에 기초하여 결정되며, 이웃하는 또는 다른 서브밴드의 절단된 서브밴드 필터의 길이에 영향을 받지 않는다. 이를테면, Zone 2의 일부 혹은 전부의 절단된 서브밴드 필터의 길이가 Zone 1의 적어도 하나의 절단된 서브밴드 필터의 길이보다 길 수 있다.
- [0086] 본 발명의 또 다른 실시예에 따르면, 복수의 그룹으로 분류된 서브밴드 중 일부에 대해서만 주파수 도메인 가변차수 필터링이 수행될 수 있다. 즉, 분류된 적어도 2개의 그룹 중 일부의 그룹에 속한 서브밴드에 대해서만 서로 다른 길이를 갖는 절단된 서브밴드 필터가 생성될 수 있다. 일 실시예에 따르면, 절단된 서브밴드 필터가 생성되는 그룹은 기 설정된 상수 또는 기 설정된 주파수 밴드를 기초로 저 주파수 밴드로 분류된 서브밴드 그룹(이를테면, Zone 1)이 될 수 있다.
- [0087] 절단된 필터의 길이는 오디오 신호 처리 장치가 획득한 추가적인 정보 이를 테면, 디코더의 복잡도(complexity), 복잡도 레벨(프로파일), 또는 요구되는 퀄리티 정보에 기초하여 결정될 수 있다. 복잡도는 오디오 신호 처리 장치의 하드웨어 리소스(resource)에 따라 결정되거나 사용자가 직접 입력한 값에 따라 결정될 수 있다. 퀄리티는 유저의 요구에 따라 결정되거나, 비트스트림을 통해 전송된 값 또는 비트스트림에 포함된 다른 정보를 참조하여 결정될 수 있다. 또한, 퀄리티는 전송되는 오디오 신호의 품질을 추정된 값에 따라 결정될 수도 있는데, 이를테면 비트 레이트가 높을수록 더 높은 퀄리티로 간주할 수 있다. 이때, 각 절단된 서브밴드 필터의 길이는 복잡도 및 퀄리티에 따라 비례적으로 증가할 수도 있고, 각 밴드별로 서로 다른 비율로 변화할 수

도 있다. 또한, 각 절단된 서브밴드 필터의 길이는 후술하는 FFT 등의 고속 프로세싱에 의한 추가적인 이득을 얻기 위해 이에 대응되는 크기 단위 이를 테면, 2의 거듭제곱의 배수로 결정될 수 있다. 반면, 절단된 서브밴드 필터의 결정된 길이가 실제 서브밴드 필터의 총 길이보다 길 경우, 절단된 서브밴드 필터의 길이는 실제 서브밴드 필터의 길이로 조정될 수 있다.

[0088] BRIR 파라미터화부는 전술한 실시예에 따라 결정된 각 절단된 서브밴드 필터에 대응하는 절단된 서브밴드 필터 계수(F-파트 계수)들을 생성하고, 이를 고속 콘볼루션부로 전달한다. 고속 콘볼루션부는 절단된 서브밴드 필터 계수를 이용하여 멀티 오디오 신호의 각 서브밴드 신호에 대한 주파수 도메인 가변차수 필터링을 수행한다.

[0089] 도 9는 바이노럴 렌더링에 사용되는 QMF 도메인 필터의 각 QMF 밴드 별 길이의 다른 실시예를 나타내고 있다. 도 9의 실시예에서 도 8의 실시예와 동일하거나 상응하는 부분은 중복적인 설명을 생략하도록 한다.

[0090] 도 9의 실시예에서 F_{iL} , F_{iR} 은 각각 QMF 서브밴드 i 의 F-파트 렌더링에 사용되는 절단된 서브밴드 필터(프론트 서브밴드 필터)를 나타내며, P_i 는 QMF 서브밴드 i 의 P-파트 렌더링에 사용되는 리어 서브밴드 필터를 나타낸다. N 은 원본 서브밴드 필터의 길이(탭 수)를 나타내며, N_{iF} 및 N_{iP} 는 각각 서브밴드 i 의 프론트 서브밴드 필터 및 리어 서브밴드 필터의 길이를 나타낸다. 전술한 바와 같이, N_{iF} 및 N_{iP} 는 다운 샘플된 QMF 도메인에서의 탭 수를 나타낸다.

[0091] 도 9의 실시예에 따르면, 프론트 서브밴드 필터뿐만 아니라 리어 서브밴드 필터의 길이도 원본 서브밴드 필터에서 추출된 파라미터에 기초하여 결정될 수 있다. 즉, 각 서브밴드의 프론트 서브밴드 필터 및 리어 서브밴드 필터의 길이는 해당 서브밴드 필터에서 추출된 특성 정보에 적어도 부분적으로 기초하여 결정된다. 예를 들어, 프론트 서브밴드 필터의 길이는 해당 서브밴드 필터의 제1 잔향 시간 정보에 기초하여, 리어 서브밴드 필터의 길이는 제2 잔향 시간 정보에 기초하여 결정될 수 있다. 즉, 프론트 서브밴드 필터는 원본 서브밴드 필터에서 제1 잔향 시간 정보에 기초하여 절단된 앞부분의 필터이며, 리어 서브밴드 필터는 프론트 서브밴드 필터 이후의 구간으로서 제1 잔향 시간과 제2 잔향 시간 사이의 구간에 대응하는 뒷부분의 필터가 될 수 있다. 일 실시예에 따르면 제1 잔향 시간 정보는 RT_{20} , 제2 잔향 시간 정보는 RT_{60} 이 될 수 있으나, 본 발명은 이에 한정하지 않는다.

[0092] 제2 잔향 시간 이내에는 초기 반사음 파트에서 후기잔향 파트로 전환되는 부분이 존재한다. 즉, 결정성(deterministic characteristic)을 갖는 구간에서 추계적 특성(stochastic characteristic)을 갖는 구간으로 전환 되는 지점이 존재하며, 전체 밴드의 BRIR의 관점에서 이 지점을 믹싱 타임이라고 부른다. 믹싱 타임 이전 구간의 경우 각 위치 별로 방향성을 제공하는 정보가 주로 존재하며, 이는 채널 별로 고유하다. 반면에 후기잔향 파트의 경우 채널 별로 공통된 특징을 지니기 때문에 복수개의 채널을 한꺼번에 처리하는 것이 효율적일 수 있다. 따라서 서브밴드 별 믹싱 타임을 추정하여 믹싱 타임 이전에 대해서는 F-파트 렌더링을 통해 고속 콘볼루션을 수행하고, 믹싱 타임 이후에 대해서는 P-파트 렌더링을 통해 각 채널 별 공통된 특성이 반영된 프로세싱을 수행할 수 있다.

[0093] 그러나 믹싱 타임을 추정하는 것은 지각적(perceptual) 관점에서 편향(bias)에 의한 에러가 발생할 수 있다. 따라서, 정확한 믹싱 타임을 추정하여 해당 경계를 기준으로 F-파트와 P-파트로 나누어 처리하는 것 보다는, F-파트의 길이를 최대한 길게 하여 고속 콘볼루션을 수행하는 것이 퀄리티 관점에서는 우수하다. 따라서, F-파트의 길이 즉, 프론트 서브밴드 필터의 길이는 복잡도-퀄리티 제어에 따라 믹싱 타임에 해당하는 길이보다 길거나 짧아질 수 있다.

[0094] 이에 더하여, 각 서브밴드 필터의 길이를 줄이기 위해 전술한 바와 같이 절단하는 방법 이외에도, 특정 서브밴드의 주파수 응답이 단조로울(monotonic) 경우 해당 서브밴드의 필터를 낮은 차수로 감소시키는 모델링이 가능하다. 대표적인 방법으로는, 주파수 샘플링을 이용한 FIR 필터 모델링이 있으며, 최소 자승 관점에서 최소화되는 필터를 디자인할 수 있다.

[0095] 본 발명의 실시예에 따르면, 각 서브밴드 별 프론트 서브밴드 필터 및/또는 리어 서브밴드 필터의 길이는 해당 서브밴드의 각 채널에 대해 동일한 값을 가질 수 있다. BRIR에는 측정 상의 에러가 존재할 수 있으며, 잔향 시간을 추정하는데 있어서도 편향 등의 오차 요소가 존재한다. 따라서, 이러한 영향을 줄이기 위해 채널간 또는 서브밴드간 상호 관계에 기초하여 필터의 길이가 결정될 수 있다. 일 실시예에 따르면, BRIR 파라미터화부는 동일한 서브밴드의 각 채널에 대응하는 서브밴드 필터로부터 각각 제1 특성 정보(이를 테면, 제1 잔향 시간 정보)를 추출하고, 추출된 제1 특성 정보를 조합하여 해당 서브밴드에 대한 하나의 필터 차수 정보(또는, 제1 절단 지점 정보)를 획득할 수 있다. 해당 서브밴드의 각 채널 별 프론트 서브밴드 필터는 상기 획득된 필터 차수

정보(또는, 제1 절단 지점 정보)에 기초하여 동일한 길이를 갖도록 결정될 수 있다. 마찬가지로, BRIR 파라미터 화부는 동일한 서브밴드의 각 채널에 대응하는 서브밴드 필터로부터 각각 제2 특성 정보(이를 테면, 제2 잔향 시간 정보)를 추출하고, 추출된 제2 특성 정보를 조합하여 해당 서브밴드의 각 채널에 대응하는 리어 서브밴드 필터에 공통으로 적용될 제2 절단 지점 정보를 획득할 수 있다. 여기서, 프론트 서브밴드 필터는 원본 서브밴드 필터에서 제1 절단 지점 정보에 기초하여 절단된 앞부분의 필터이며, 리어 서브밴드 필터는 프론트 서브밴드 필터 이후의 구간으로서 제1 절단 지점과 제2 절단 지점 사이의 구간에 대응하는 뒷부분의 필터가 될 수 있다

[0096] 한편 본 발명의 다른 실시예에 따르면, 특정 서브밴드 그룹의 서브밴드에 대해서는 F-파트 프로세싱만 수행될 수 있다. 이때, 해당 서브밴드에 대해서 제1 절단 지점까지의 필터만 이용하여 프로세싱이 수행되면, 전체 서브밴드 필터를 이용하여 프로세싱이 수행될 때와 비교하여 처리되는 필터의 에너지 차이에 의하여 사용자가 지각할 수 있는 수준의 왜곡이 발생할 수 있다. 이러한 왜곡을 방지하기 위해, 해당 서브밴드 필터에서 프로세싱에 사용되지 않는 영역, 즉 제1 절단 지점 이후의 영역에 대한 에너지 보상이 이루어 질 수 있다. 상기 에너지 보상은 F-파트 계수(프론트 서브밴드 필터 계수)에 해당 서브밴드 필터의 제1 절단 지점까지의 필터 파워를 나누고, 원하는 영역의 에너지, 이를테면 해당 서브밴드 필터의 전체 파워를 곱함으로써 수행 가능하다. 따라서, F-파트 계수의 에너지가 전체 서브밴드 필터의 에너지와 같아지도록 조정될 수 있다. 또한, BRIR 파라미터화부에서 P-파트 계수가 전송되었음에도 불구하고, 바이노럴 렌더링 유닛에서는 복잡도-퀄리티 제어에 기초하여 P-파트 프로세싱을 수행하지 않을 수 있다. 이 경우, 바이노럴 렌더링 유닛은 P-파트 계수를 이용하여 F-파트 계수에 대한 상기 에너지 보상을 수행할 수 있다.

[0097] 전술한 방법들에 의한 F-파트 프로세싱에 있어서, 각 서브밴드 별로 다른 길이를 갖는 절단된 서브밴드 필터의 필터 계수는 하나의 시간 영역 필터(즉, proto-type 필터)로부터 획득된다. 즉, 하나의 시간 영역 필터를 복수의 QMF 서브밴드 필터로 변환하고, 각 서브밴드에 대응되는 필터들의 길이를 가변 시킨 것이므로, 각 절단된 서브밴드 필터는 하나의 원형필터로부터 획득된 것이다.

[0098] BRIR 파라미터화부는 전술한 실시예에 따라 결정된 각 프론트 서브밴드 필터에 대응하는 프론트 서브밴드 필터 계수(F-파트 계수)를 생성하고, 이를 고속 콘볼루션부로 전달한다. 고속 콘볼루션부는 수신된 프론트 서브밴드 필터 계수를 이용하여 멀티 오디오 신호의 각 서브밴드 신호에 대한 주파수 도메인 가변차수 필터링을 수행한다. 또한, BRIR 파라미터화부는 전술한 실시예에 따라 결정된 각 리어 서브밴드 필터에 대응하는 리어 서브밴드 필터 계수(P-파트 계수)를 생성하고, 이를 후기잔향 생성부로 전달할 수 있다. 후기잔향 생성부는 수신된 리어 서브밴드 필터 계수를 이용하여 각 서브밴드 신호에 대한 잔향 처리를 수행할 수 있다. 본 발명의 실시예에 따르면, BRIR 파라미터화부는 각 채널 별 리어 서브밴드 필터 계수들을 조합하여 다운믹스 서브밴드 필터 계수(다운믹스 P-파트 계수)를 생성하고, 이를 후기잔향 생성부로 전달할 수 있다. 후술하는 바와 같이, 후기잔향 생성부는 수신된 다운믹스 서브밴드 필터 계수를 이용하여 2채널의 좌, 우 서브밴드 잔향 신호를 생성할 수 있다.

[0099] 도 10은 바이노럴 렌더링에 사용되는 FIR 필터 생성 방법의 또 다른 실시예를 나타내고 있다. 도 10의 실시예에서 도 8 및 도 9의 실시예와 동일하거나 상응하는 부분은 중복적인 설명을 생략하도록 한다.

[0100] 도 10을 참조하면, QMF 변환된 복수의 서브밴드 필터들은 복수의 그룹으로 분류되고, 분류된 각 그룹별로 서로 다른 프로세싱이 적용될 수 있다. 예를 들어, 복수의 서브밴드는 기 설정된 주파수 밴드(QMF 밴드 i)를 기준으로 한 저 주파수의 제1 서브밴드 그룹(Zone 1)과, 고 주파수의 제2 서브밴드 그룹(Zone 2)으로 분류될 수 있다. 이때, 제1 서브밴드 그룹의 입력 서브밴드 신호들에 대해서는 F-파트 렌더링이, 제2 서브밴드 그룹의 입력 서브밴드 신호들에 대해서는 후술하는 QTDL 프로세싱이 수행될 수 있다.

[0101] 따라서, BRIR 파라미터화부는 제1 서브밴드 그룹의 각 서브밴드 별로 프론트 서브밴드 필터 계수를 생성하고, 이를 고속 콘볼루션부에 전달한다. 고속 콘볼루션부는 수신된 프론트 서브밴드 필터 계수를 이용하여 제1 서브밴드 그룹의 서브밴드 신호에 대한 F-파트 렌더링을 수행한다. 실시예에 따라서, 제1 서브밴드 그룹의 서브밴드 신호에 대한 P-파트 렌더링이 후기잔향 생성부에 의해 추가적으로 수행될 수도 있다. 또한, BRIR 파라미터화부는 제2 서브밴드 그룹의 각 서브밴드 필터 계수로부터 적어도 하나의 파라미터를 획득하고 이를 QTDL 프로세싱부로 전달한다. QTDL 프로세싱부는 획득된 파라미터를 이용하여 후술하는 바와 같이 제2 서브밴드 그룹의 각 서브밴드 신호에 대한 탭-딜레이 라인 필터링을 수행한다. 본 발명의 실시예에 따르면, 제1 서브밴드 그룹과 제2 서브밴드 그룹을 구분하는 기 설정된 주파수(QMF 밴드 i)는 사전에 정해진 상수 값에 기초하여 결정될 수도 있고, 전송된 오디오 입력 신호의 비트열 특성에 따라 결정될 수도 있다. 이를 테면, SBR을 사용하는 오디오 신호의 경우, 제2 서브밴드 그룹이 SBR 밴드에 대응하도록 설정될 수 있다.

[0102] 본 발명의 다른 실시예에 따르면, 복수의 서브밴드는 기 설정된 제1 주파수 밴드(QMF 밴드 i) 및 제 2 주파수 밴드(QMF 밴드 j)를 기초로 3개의 서브밴드 그룹으로 분류될 수도 있다. 즉, 복수의 서브밴드는 제1 주파수 밴드보다 작거나 같은 저 주파수 구역인 제1 서브밴드 그룹(Zone 1), 제1 주파수 밴드 보다 크고 제2 주파수 밴드 보다 작거나 같은 중간 주파수 구역인 제2 서브밴드 그룹(Zone 2), 및 제2 주파수 밴드 보다 큰 고 주파수 구역인 제3 서브밴드 그룹(Zone 3)으로 분류될 수 있다. 이때, 제1 서브밴드 그룹의 서브밴드 신호 및 제2 서브밴드 그룹의 서브밴드 신호들에 대해서는 전술한 바와 같이 각각 F-파트 렌더링 및 QTDL 프로세싱이 수행될 수 있으며, 제3 서브밴드 그룹의 서브밴드 신호들에 대해서는 렌더링을 수행하지 않을 수 있다.

[0103] <후기잔향 렌더링>

[0104] 다음으로 도 11 내지 도 14를 참조로 본 발명의 P-파트 렌더링의 다양한 실시예에 대해 설명하도록 한다. 즉, QMF 도메인에서 P-파트 렌더링을 수행하는 도 2의 후기잔향 생성부(240)의 다양한 실시예가 도 11 내지 도 14를 참조로 설명된다. 도 11 내지 도 14의 실시예에서 멀티채널 입력 신호는 QMF 도메인의 서브밴드 신호로 수신된다고 가정한다. 따라서, 도 11 내지 도 14의 각 구성 즉, 디코릴레이터(241), 서브밴드 필터링부(242), IC 매칭부(243), 다운 믹스부(244) 및 에너지 감쇄 매칭부(246)의 프로세싱은 각 QMF 서브밴드 별로 수행될 수 있다. 도 11 내지 도 14의 실시예에서 이전 도면의 실시예와 중복되는 부분은 구체적인 설명을 생략하도록 한다.

[0105] 전술한 도 8 내지 도 10의 실시예에서 P-파트에 대응되는 $P_i(P_1, P_2, P_3, \dots)$ 는 주파수 가변 절단에 따라 제거된 각 서브밴드 필터의 뒷부분에 해당하며, 통상적으로 후기잔향에 대한 정보를 포함하고 있다. 복잡도-퀄리티 제어에 따라 P-파트의 길이는 각 서브밴드 필터의 절단된 지점 이후의 전체 필터로 정의될 수도 있고, 해당 서브밴드 필터의 제2 잔향 시간 정보를 참조하여 보다 작은 길이로 정의될 수도 있다.

[0106] P-파트 렌더링은 각 채널 별로 독립적으로 수행될 수도 있고, 다운믹스 된 채널에 대해 수행될 수도 있다. 또한, P-파트 렌더링은 기 설정된 서브밴드 그룹 별로 또는 각 서브밴드 별로 서로 다른 프로세싱을 통해 적용될 수도 있으며, 전체 서브밴드에 대하여 동일한 프로세싱으로 적용될 수도 있다. 이때, P-파트에 적용 가능한 프로세싱으로는 입력 신호에 대한 에너지 감소 보상, 탭-딜레이 라인 필터링, IIR(Infinite Impulse Response) 필터를 이용한 프로세싱, 인공 잔향기(artificial reverberator)를 이용한 프로세싱, FIIC(Frequency-independent Interaural Coherence) 보상, FDIC(Frequency-dependent Interaural Coherence) 보상 등이 포함될 수 있다.

[0107] 한편, P-파트에 대한 파라메트릭(parametric) 프로세싱을 위해서는 크게 두 가지 특징 즉, EDR(Energy Decay Relief)과 FDIC(Frequency-dependent Interaural Coherence)의 특징을 보존하는 것이 중요하다. 먼저, P-파트를 에너지 관점에서 관찰하면, 각 채널 별로 EDR이 동일 또는 유사함을 알 수 있다. 각 채널이 공통된 EDR을 가지고 있기 때문에, 모든 채널을 한 개 또는 두 개의 채널로 다운 믹스한 후, 다운 믹스 된 채널에 대한 P-파트 렌더링을 수행하는 것은 에너지 관점에서 타당하다. 이때, M개의 채널에 대하여 M회의 콘볼루션을 수행해야 하는 P-파트 렌더링의 연산을, M-to-0 다운믹스와 1회(또는, 2회)의 콘볼루션으로 줄임으로 상당한 연산량의 이득을 제공할 수 있다.

[0108] 다음으로, P-파트 렌더링에서 FDIC를 보상하는 과정이 필요하다. FDIC를 추정하는 방법에는 여러 가지가 있으나, 다음과 같은 수식이 사용될 수 있다.

수학식 3

$$IC(i) = \frac{\Re \left[\sum_{k=0}^K H_L(i,k) H_R(i,k)^* \right]}{\sqrt{\sum_{k=0}^K |H_L(i,k)|^2 \sum_{k=0}^K |H_R(i,k)|^2}}$$

[0109]

[0110] 여기서, $H_m(i,k)$ 는 임펄스 응답 $h_m(n)$ 의 STFT(Short Time Fourier Transform) 계수, n은 시간 인덱스, i는 주파수 인덱스, k는 프레임 인덱스, m은 출력 채널 인덱스(L, R)를 나타낸다. 또한, 분자의 함수 $\Re(x)$

는 입력 x 의 실수 값을 출력하고, x^* 는 x 의 복소 켈레(complex conjugate) 값을 나타낸다. 상기 수식에서 분자 부분은 실수 값 대신 절댓값을 취하는 함수로 치환될 수도 있다.

[0111] 한편, 본 발명에서 바이노럴 렌더링은 QMF 도메인에서 수행되기 때문에, FDIC는 다음과 같은 수식으로 정의될 수도 있다.

수학식 4

$$IC(i) = \frac{\Re \left[\sum_{k=0}^K h_L(i,k) h_R(i,k)^* \right]}{\sqrt{\sum_{k=0}^K |h_L(i,k)|^2 \sum_{k=0}^K |h_R(i,k)|^2}}$$

[0112]

[0113] 여기서, i 는 서브밴드 인덱스, k 는 서브밴드에서의 시간 인덱스, $h_m(i,k)$ 는 BRIR의 서브밴드 필터를 나타낸다.

[0114] 후기잔향 파트의 FDIC는 BRIR이 레코딩 될 때 두 마이크로폰의 위치에 주로 영향을 받는 파라미터로서, 스피커의 위치, 즉 방향과 거리에는 영향을 받지 않는다. 청자의 머리를 구(sphere)라고 가정했을 때, BRIR의 이론적 인 FDIC(IC_{ideal})는 다음과 같은 수식을 만족시킬 수 있다.

수학식 5

$$IC_{ideal}(k) = \frac{\sin(kr)}{kr}$$

[0115]

[0116] 여기서, r 은 청자의 양 귀 사이의 거리, 즉 두 마이크로폰 사이의 거리이며, k 는 주파수 인덱스이다.

[0117] 복수 채널의 BRIR을 이용한 FDIC를 분석하면, F-파트에 주로 포함된 초기 반사음은 각 채널 별로 매우 다름을 알 수 있다. 즉, F-파트의 FDIC는 채널 별로 매우 다르게 변화한다. 한편, 고 주파수 대역의 경우 FDIC가 매우 크게 변화하지만, 이는 에너지가 빠르게 감쇄하는 고 주파수 대역 신호의 특성상 큰 측정 오차가 발생하기 때문이고, 채널 별 평균을 취하게 될 경우 FDIC는 거의 0에 수렴한다. 반면, P-파트의 경우에도 측정 오차로 인해 채널 별 FDIC의 차이가 발생하지만, 평균적으로 수학식 5와 같은 싱크 함수에 수렴하는 것을 확인할 수 있다. 본 발명의 실시예에 따르면, 전술한 바와 같은 특성에 기초하여 P-파트 렌더링을 위한 후기잔향 생성부를 구현할 수 있다.

[0118] 도 11은 본 발명의 일 실시예에 따른 후기잔향 생성부(240A)를 나타내고 있다. 도 11의 실시예에 따르면, 후기잔향 생성부(240A)는 서브밴드 필터링부(242) 및 다운믹스부(244a, 244b)를 포함할 수 있다.

[0119] 서브밴드 필터링부(242)는 P-파트 계수를 이용하여 멀티채널 입력 신호 X_0, X_1, \dots, X_{M-1} 을 각 서브밴드 별로 필터링 한다. P-파트 계수는 전술한 바와 같이 BRIR 파라미터화부(미도시)로부터 수신되며, 서브밴드 별로 서로 다른 길이를 갖는 리어 서브밴드 필터의 계수를 포함할 수 있다. 서브밴드 필터링부(242)는 각 주파수 별로 QMF 도메인 서브밴드 신호와 이에 대응하는 QMF 도메인의 리어 서브밴드 필터 간의 고속 콘볼루션을 수행한다. 이때, 리어 서브밴드 필터의 길이는 전술한 바와 같이 RT60에 기초하여 결정될 수 있으나, 복잡도-퀄리티 제어에 따라 RT60보다 크거나 작은 값으로 설정될 수도 있다.

[0120] 멀티채널 입력 신호는 서브밴드 필터링부(242)에 의해 각각 좌 채널 신호인 $X_{L0}, X_{L1}, \dots, X_{L_{M-1}}$ 과 우 채널 신호인 $X_{R0}, X_{R1}, \dots, X_{R_{M-1}}$ 로 렌더링 된다. 다운믹스부(244a, 244b)는 렌더링 된 복수의 좌 채널 신호와 복수의 우 채널 신호를 각각 좌, 우 채널 별로 다운믹스 하여 2채널의 좌, 우 출력 신호 Y_{Lp}, Y_{Rp} 를 생성한다.

[0121] 도 12는 본 발명의 다른 실시예에 따른 후기잔향 생성부(240B)를 나타내고 있다. 도 12의 실시예에 따르면, 후기잔향 생성부(240B)는 디코릴레이터(241), IC 매칭부(243), 다운믹스부(244a, 244b) 및 에너지 감쇄 매칭부(246a, 246b)를 포함할 수 있다. 또한, 후기잔향 생성부(240B)의 프로세싱을 위해, BRIR 파라미터화부(미도시)는 IC 추정부(213) 및 다운믹스 서브밴드 필터 생성부(216)를 포함할 수 있다.

[0122] 도 12의 실시예에 따르면, 후기잔향 생성부(240B)는 후기잔향 파트의 채널 별 에너지 감쇄(energy decay) 특성이 동일함을 이용하여 연산량을 감소시킬 수 있다. 즉, 후기잔향 생성부(240B)는 각 멀티채널 신호에 대한 디코릴레이션 및 IC(Interaural Coherence) 조정을 수행하고, 조정된 각 채널 별 입력 신호 및 디코릴레이션 신호를 좌, 우 채널 신호로 다운믹스 한 뒤, 다운믹스 된 신호의 에너지 감쇄를 보상하여 2채널의 좌, 우 출력 신호를 생성한다. 더욱 구체적으로, 디코릴레이터(241)는 각 멀티채널 입력 신호 X_0, X_1, \dots, X_{M-1} 에 대한 디코릴레이션 신호 D_0, D_1, \dots, D_{M-1} 를 생성한다. 디코릴레이터(241)는 양 귀 간의 코히어런스(coherence)를 조정하기 위한 일종의 전처리기로서, 위상 랜덤화기(phase randomizer)가 사용될 수 있으며, 연산량의 효율을 위해 90도 단위로 입력 신호의 위상을 바꾸어 줄 수도 있다.

[0123] 한편, BRIR 파라미터화부(미도시)의 IC 추정부(213)는 IC 값을 추정하여 바이노럴 렌더링 유닛(미도시)으로 전달한다. 바이노럴 렌더링 유닛은 수신된 IC 값을 메모리(255)에 저장할 수 있으며, IC 매칭부(243)로 전달한다. IC 매칭부(243)는 BRIR 파라미터화부로부터 IC 값을 직접 수신할 수도 있으며, 메모리(255)에 기 저장된 IC 값을 획득할 수도 있다. 각 채널 별 입력 신호 및 디코릴레이션 신호는 IC 매칭부(243)에서 좌 채널 신호인 $X_{L0}, X_{L1}, \dots, X_{L_{M-1}}$ 과 우 채널 신호인 $X_{R0}, X_{R1}, \dots, X_{R_{M-1}}$ 로 렌더링 된다. IC 매칭부(243)는 IC 값을 참조하여 각 채널 별로 디코릴레이션 신호와 원래의 입력 신호 간의 가중치 합산을 수행하며, 이를 통해 두 채널 신호 간의 코히어런스를 조정한다. 이때, 각 채널 별 입력 신호는 서브밴드 도메인의 신호이므로, 전술한 FDIC의

매칭이 가능하다. 원래의 채널 신호를 X , 디코릴레이션 채널 신호를 D , 해당 서브밴드의 IC를 ϕ 라고 했을 때, IC 매칭이 수행된 좌, 우 채널 신호 X_L, X_R 은 다음 수식과 같이 표현될 수 있다.

수학식 6

$$X_L = \sqrt{(1+\phi)/2}X \pm \sqrt{(1-\phi)/2}D$$

$$X_R = \sqrt{(1+\phi)/2}X \mp \sqrt{(1-\phi)/2}D$$

[0124]

(복부호 동순)

[0125]

[0126] 다운믹스부(244a, 244b)는 IC 매칭을 통해 렌더링 된 복수의 좌 채널 신호와 복수의 우 채널 신호를 각각 좌, 우 채널 별로 다운믹스 하여 2채널의 좌, 우 렌더링 신호를 생성한다. 다음으로, 에너지 감쇄 매칭부(246a, 246b)는 2채널의 좌, 우 렌더링 신호의 에너지 감쇄를 각각 반영하여 2채널의 좌, 우 출력 신호 Y_{Lp}, Y_{Rp} 를 생성한다. 에너지 감쇄 매칭부(246a, 246b)는 다운믹스 서브밴드 필터 생성부(216)로부터 획득된 다운믹스 서브밴드 필터 계수를 이용하여 에너지 감쇄 매칭을 수행한다. 다운믹스 서브밴드 필터 계수는 해당 서브밴드의 각 채널 별 리어 서브밴드 필터 계수의 조합에 의해 생성된다. 이를테면, 다운믹스 서브밴드 필터 계수는 해당 서브밴드에 대하여 각 채널 별 리어 서브밴드 필터 계수의 제곱 진폭 응답의 평균에 루트를 취한 서브밴드 필터 계수를 포함할 수 있다. 따라서 다운믹스 서브밴드 필터 계수는 해당 서브밴드 신호에 대한 후기잔향 파트의 에너지 감소 특성을 반영한다. 다운믹스 서브밴드 필터 계수는 실시예에 따라 모노 또는 스테레오로 다운믹스 된 다운믹스 서브밴드 필터 계수를 포함할 수 있으며, FDIC와 마찬가지로 BRIR 파라미터화부로부터 직접 수신되거나, 메모리(225)에 기 저장된 값으로부터 획득될 수 있다. M개의 채널 중 k번째 채널에서의 F-파트가 절단된

BRIR을 $BRIR_k$, k번째 채널에서의 N번째 샘플까지의 절단된 BRIR을 $BRIR_{T,k}$, N번째 샘플 이후의 절단된 부분의 에너지를 보상한 다운믹스 서브밴드 필터 계수를 $BRIR_E$ 라고 한다면, $BRIR_E$ 는 다음과 같은 수식을 이용하여 구할 수 있다.

수학식 7

$$BRIR_E(m) = \frac{\sqrt{\sum_{k=0}^{M-1} \sum_{m'=0}^{\infty} (BRIR_k(m'))^2}}{\sqrt{\sum_{k=0}^{M-1} \sum_{m'=0}^{N-1} (BRIR_{T,k}(m'))^2}} \sqrt{\frac{\sum_{k=0}^{M-1} (BRIR_{T,k}(m))^2}{M}}$$

[0127]

$$BRIR_{T,k}(m) = \begin{cases} BRIR_k(m) & m < N \\ 0 & otherwise \end{cases}$$

[0128]

도 13은 본 발명의 또 다른 실시예에 따른 후기잔향 생성부(240C)를 나타내고 있다. 도 13의 후기잔향 생성부(240C)의 각 구성은 도 12의 실시예에서 설명된 후기잔향 생성부(240B)의 각 구성과 동일할 수 있으며, 각 구성 간의 데이터 처리 순서가 일부 다를 수 있다.

[0129]

도 13의 실시예에 따르면, 후기잔향 생성부(240C)는 후기잔향 파트의 각 채널 별 FDIC가 동일함을 이용하여 연산량을 더욱 감소시킬 수 있다. 즉, 후기잔향 생성부(240C)는 각 멀티채널 신호를 좌, 우 채널 신호로 다운믹스 하고, 다운믹스 된 좌, 우 채널 신호의 IC를 조정된 뒤, 조정된 좌, 우 채널 신호에 대한 에너지 감쇄를 보상하여 2채널의 좌, 우 출력 신호를 생성할 수 있다.

[0130]

더욱 구체적으로, 디코릴레이터(241)는 각 멀티채널 입력 신호 X₀, X₁, ..., X_{M-1}에 대한 디코릴레이션 신호 D₀, D₁, ..., D_{M-1}를 생성한다. 다음으로, 다운믹스부(244a, 244b)는 멀티채널 입력 신호 및 디코릴레이션 신호를 각각 다운믹스하여 2채널 다운믹스 신호 X_{DMX}, D_{DMX}를 생성한다. IC 매칭부(243)는 IC 값을 참조하여 2채널 다운믹스 신호를 가중치 합산하며, 이를 통해 두 채널 신호 간의 코히어런스를 조정한다. 에너지 감쇄 보상부(246a, 246b)는 IC 매칭부(243)에 의해 IC 매칭이 수행된 좌, 우 채널 신호 X_L, X_R 각각에 대한 에너지 보상을 수행하여 2채널의 좌, 우 출력 신호 X_{Lp}, Y_{Rp}를 생성한다. 이때, 에너지 보상에 사용되는 에너지 보상 정보에는 각 서브밴드에 대한 다운믹스 서브밴드 필터 계수가 포함될 수 있다.

[0131]

도 14는 본 발명의 또 다른 실시예에 따른 후기잔향 생성부(240D)를 나타내고 있다. 도 14의 후기잔향 생성부(240D)의 각 구성은 도 12 및 도 13의 실시예에서 설명된 후기잔향 생성부(240B, 240C)의 각 구성과 동일할 수 있으나, 더욱 간략화 된 특징을 갖는다.

[0132]

먼저, 다운 믹스부(244)는 멀티채널 입력 신호 X₀, X₁, ..., X_{M-1}를 각 서브밴드 별로 다운믹스 하여 모노 다운믹스 신호(즉, 모노 서브밴드 신호) X_{DMX}를 생성한다. 에너지 감쇄 매칭부(246)는 생성된 모노 다운믹스 신호에 대한 에너지 감쇄를 반영한다. 이때, 에너지 감쇄를 반영하기 위해 각 서브밴드에 대한 다운믹스 서브밴드 필터 계수가 사용될 수 있다. 다음으로, 디코릴레이터(241)는 에너지 감쇄가 반영된 모노 다운믹스 신호의 디코릴레이션 신호 D_{DMX}를 생성한다. IC 매칭부(243)는 에너지 감쇄가 반영된 모노 다운믹스 신호와 디코릴레이션 신호를 FDIC 값을 참조하여 가중치 합산하며, 이를 통해 2채널의 좌, 우 출력 신호 Y_{Lp}, Y_{Rp}를 생성한다. 도 14의 실시예에 따르면, 에너지 감쇄 매칭을 모노 다운믹스 신호(X_{DMX})에 대해 1회만 수행하게 되므로 연산량을 더욱 절약할 수 있다.

[0133]

<고 주파수 밴드의 QTDL 프로세싱>

[0134]

다음으로 도 15 및 도 16을 참조로 본 발명의 QTDL 프로세싱의 다양한 실시예에 대해 설명하도록 한다. 즉, QMF 도메인에서 QTDL 프로세싱을 수행하는 도 2의 QTDL 프로세싱부(250)의 다양한 실시예가 도 15 및 도 16을 참조로 설명된다. 도 15 및 도 16의 실시예에서 멀티채널 입력 신호는 QMF 도메인의 서브밴드 신호로 수신된다고 가정한다. 따라서, 도 15 및 도 16의 실시예에서 탭-딜레이 라인 필터 및 원-탭-딜레이 라인 필터는 각 QMF 서브밴드 별로 프로세싱을 수행할 수 있다. 또한, QTDL 프로세싱은 전술한 바와 같이 기 설정된 상수 또는 기 설정된 주파수 밴드를 기초로 분류된 고 주파수 밴드의 입력 신호에 대해서만 수행될 수 있다. 만약, 입력 오디오 신호에 SBR(Spectral Band Replication)이 적용된 경우, 상기 고 주파수 밴드는 SBR 밴드에 대응될 수 있다. 도 15 및 도 16의 실시예에서 이전 도면의 실시예와 중복되는 부분은 구체적인 설명을 생략하도록 한다.

[0135]

고 주파수 대역의 효율적인 부호화를 위해 사용되는 SBR(Spectral Band Replication)은 저 비트율 부호화 시 고 주파수 대역의 신호를 버림으로 인해 좁아진 밴드 폭을 다시 확장하여, 원 신호만큼의 밴드 폭을 확보하기 위한

도구이다. 이때, 고 주파수 대역은 부호화되어 전송되는 저 주파수 대역의 정보와 인코더에서 전송한 고 주파수 대역 신호의 부가 정보를 활용하여 생성된다. 그러나 SBR을 이용하여 생성된 고 주파수 성분은 부정확한 고조파(harmonic)의 생성으로 인하여 왜곡이 발생할 수 있다. 또한, SBR 밴드는 고 주파수 대역이며, 전술한 바와 같이 해당 주파수 대역의 잔향 시간은 매우 짧다. 즉, SBR 밴드의 BRIR 서브밴드 필터는 유효 정보가 적으며, 빠른 감쇄율을 갖는다. 따라서, SBR 대역에 준하는 고 주파수 대역에 대한 BRIR 렌더링은 콘볼루션을 수행하는 것보다는 유효한 소수의 탭을 이용하여 렌더링을 수행하는 것이 음질의 퀄리티 대비 연산량 측면에서 매우 효과적일 수 있다.

[0137] 도 15는 본 발명의 일 실시예에 따른 QTDL 프로세싱부(250A)를 나타내고 있다. 도 15의 실시예에 따르면, QTDL 프로세싱부(250A)는 탭-딜레이 라인 필터를 이용하여 멀티채널 입력 신호 X_0, X_1, \dots, X_{M-1} 에 대한 서브밴드 별 필터링을 수행한다. 탭-딜레이 라인 필터는 각 채널 신호에 대하여 기 설정된 소수의 탭만 콘볼루션을 수행한다. 이때 사용되는 소수의 탭은 해당 서브밴드 신호에 대응하는 BRIR 서브밴드 필터 계수로부터 직접 추출된 파라미터에 기초하여 결정될 수 있다. 상기 파라미터는 탭-딜레이 라인 필터에 사용될 각 탭에 대한 딜레이(delay) 정보 및 이에 대응하는 게인(gain) 정보를 포함한다.

[0138] 탭-딜레이 라인 필터에 사용되는 탭의 수는 복잡도-퀄리티 제어에 의해 결정될 수 있다. QTDL 프로세싱부(250A)는 기 결정된 탭 수에 기초하여, 각 채널 및 서브밴드 별로 해당 탭 수에 대응하는 파라미터(게인 정보, 딜레이 정보) 셋을 BRIR 파라미터화부로부터 수신한다. 이때, 수신되는 파라미터 셋은 해당 서브밴드 신호에 대응하는 BRIR 서브밴드 필터 계수로부터 추출되며, 다양한 실시예에 따라 결정될 수 있다. 이를테면, 해당 BRIR 서브밴드 필터 계수의 복수의 피크들 중에서 절대 값 크기 순, 실수 값 크기 순, 또는 허수 값 크기 순으로 상기 기 결정된 탭 수 만큼 추출된 피크들 각각에 대한 파라미터의 셋이 수신될 수 있다. 이때, 각 파라미터의 딜레이 정보는 해당 피크의 위치 정보를 나타내며, QMF 도메인에서 샘플 단위의 정수 값을 갖는다. 또한, 게인 정보는 해당 딜레이 정보에 대응하는 피크의 크기에 기초하여 결정된다. 이때, 게인 정보는 서브밴드 필터 계수에서의 해당 피크값 자체가 사용될 수도 있지만, 전체 서브밴드 필터 계수에 대한 에너지 보상이 수행된 이후의 해당 피크의 가중치 값이 사용될 수도 있다. 상기 게인 정보는 해당 피크에 대한 실수 가중치 및 허수 가중치를 함께 이용하여 획득되며, 따라서 복소수 값을 갖는다.

[0139] 탭-딜레이 라인 필터에 의해 필터링 된 복수의 채널 신호는 각 서브밴드 별로 2채널의 좌, 우 출력 신호 Y_L, Y_R 로 합산된다. 한편, QTDL 프로세싱부(250A)의 각 탭-딜레이 라인 필터에서 사용되는 파라미터는 바이노럴 렌더링의 초기화 과정에서 메모리에 저장될 수 있으며, 파라미터 추출을 위한 추가적인 연산 없이 QTDL 프로세싱이 수행될 수 있다.

[0140] 도 16은 본 발명의 다른 실시예에 따른 QTDL 프로세싱부(250B)를 나타내고 있다. 도 16의 실시예에 따르면, QTDL 프로세싱부(250B)는 원-탭-딜레이 라인 필터를 이용하여 멀티채널 입력 신호 X_0, X_1, \dots, X_{M-1} 에 대한 서브밴드 별 필터링을 수행한다. 원-탭-딜레이 라인 필터는 각 채널 신호에 대하여 한 개의 탭에서만 콘볼루션을 수행하는 것으로 이해할 수 있다. 이때 사용되는 탭은 해당 서브밴드 신호에 대응하는 BRIR 서브밴드 필터 계수로부터 직접 추출된 파라미터에 기초하여 결정될 수 있다. 상기 파라미터는 전술한 바와 같이 BRIR 서브밴드 필터 계수로부터 추출된 딜레이 정보 및 이에 대응하는 게인 정보를 포함한다.

[0141] 도 16에서 L_0, L_1, \dots, L_{M-1} 은 각각 M개의 채널에서 왼쪽 귀로의 BRIR에 대한 딜레이를 나타내고, R_0, R_1, \dots, R_{M-1} 은 각각 M개의 채널에서 오른쪽 귀로의 BRIR에 대한 딜레이를 나타낸다. 이때, 딜레이 정보는 해당 BRIR 서브밴드 필터 계수 중 절대 값 크기 순, 실수 값 크기 순, 또는 허수 값 크기 순으로 최대 피크에 대한 위치 정보를 나타낸다. 또한, 도 16에서 $G_{L_0}, G_{L_1}, \dots, G_{L_{M-1}}$ 은 좌 채널의 각 딜레이 정보에 대응하는 게인을 나타내고, $G_{R_0}, G_{R_1}, \dots, G_{R_{M-1}}$ 은 각각 우 채널의 각 딜레이 정보에 대응하는 게인을 나타낸다. 전술한 바와 같이, 각 게인 정보는 해당 딜레이 정보에 대응하는 피크의 크기에 기초하여 결정된다. 이때, 게인 정보는 서브밴드 필터 계수에서의 해당 피크값 자체가 사용될 수도 있지만, 전체 서브밴드 필터 계수에 대한 에너지 보상이 수행된 이후의 해당 피크의 가중치 값이 사용될 수도 있다. 상기 게인 정보는 해당 피크에 대한 실수 가중치 및 허수 가중치를 함께 이용하여 획득되며, 따라서 복소수 값을 갖는다.

[0142] 도 15의 실시예와 같이, 원-탭-딜레이 라인 필터에 의해 필터링 된 복수의 채널 신호는 각 서브밴드 별로 2채널의 좌, 우 출력 신호 Y_L, Y_R 로 합산된다. 또한, QTDL 프로세싱부(250B)의 각 원-탭-딜레이 라인 필터에서 사용되는 파라미터는 바이노럴 렌더링의 초기화 과정에서 메모리에 저장될 수 있으며, 파라미터 추출을 위한 추가적인 연산 없이 QTDL 프로세싱이 수행될 수 있다.

[0143] 이상에서는 본 발명을 구체적인 실시예를 통하여 설명하였으나, 당업자라면 본 발명의 취지 및 범위를 벗어나지

않고 수정, 변경을 할 수 있다. 즉, 본 발명은 멀티 오디오 신호에 대한 바이노럴 렌더링의 실시예에 대하여 설명하였지만, 본 발명은 오디오 신호 뿐만 아니라 비디오 신호를 포함하는 다양한 멀티미디어 신호에도 동일하게 적용 및 확장 가능하다. 따라서 본 발명의 상세한 설명 및 실시예로부터 본 발명이 속하는 기술분야에 속한 사람이 용이하게 유추할 수 있는 것은 본 발명의 권리범위에 속하는 것으로 해석된다.

[0144] **발명의 실시를 위한 형태**

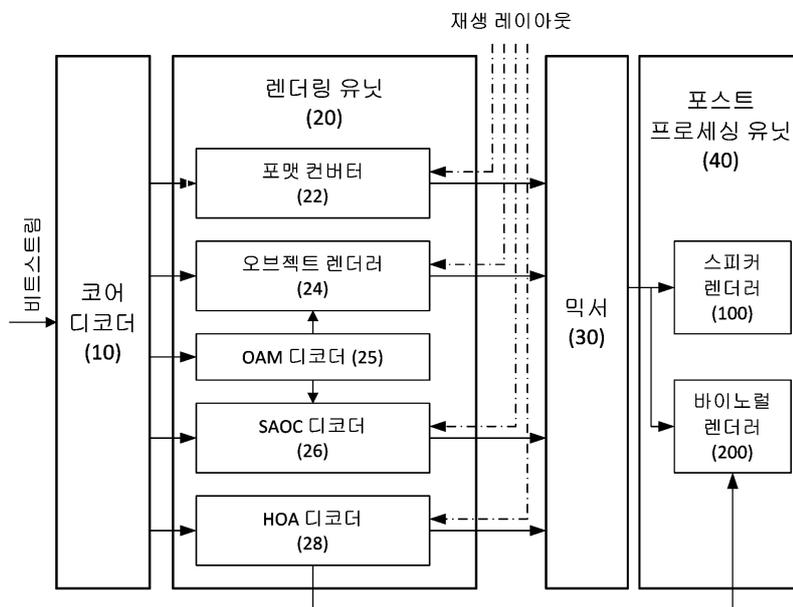
[0145] 전술한 바와 같이, 발명의 실시를 위한 최선의 형태에서 관련된 사항을 서술하였다.

산업상 이용가능성

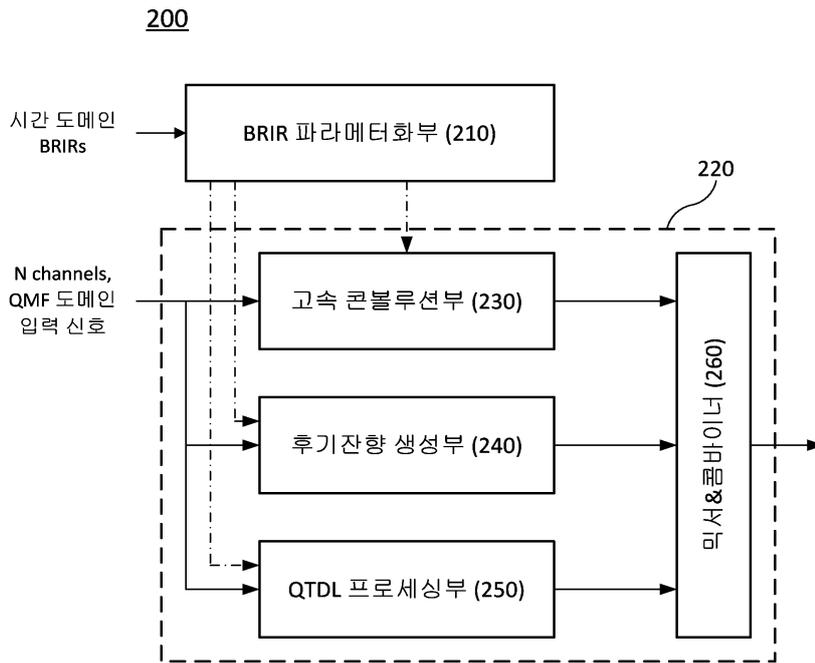
[0146] 본 발명은 다양한 형태의 오디오 신호 처리 장치 및 비디오 신호 처리 장치 등을 포함하는 멀티미디어 신호 처리 장치에 적용될 수 있다.

도면

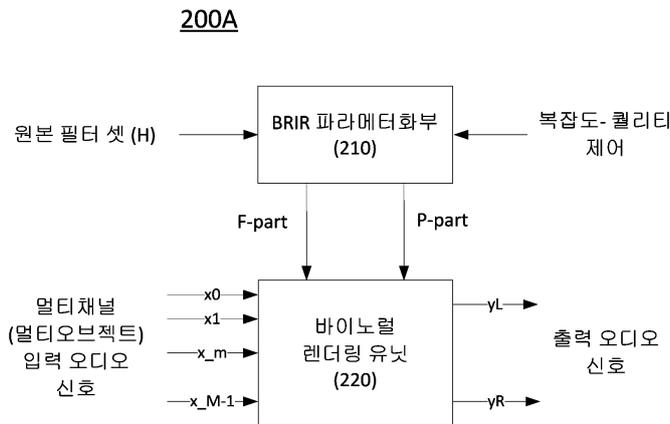
도면1



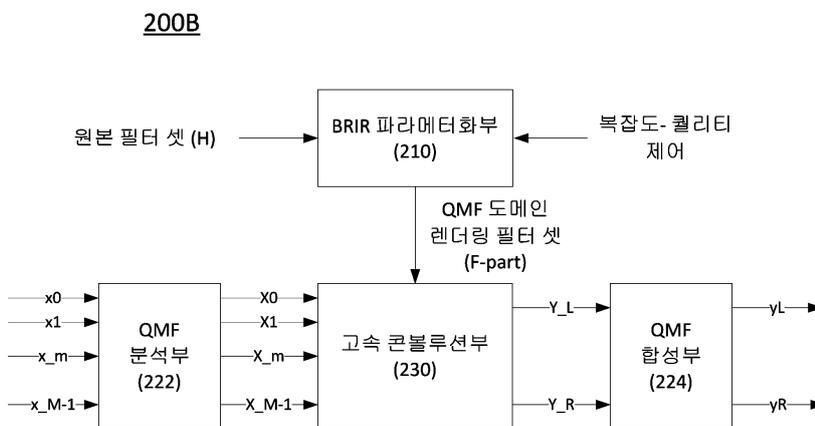
도면2



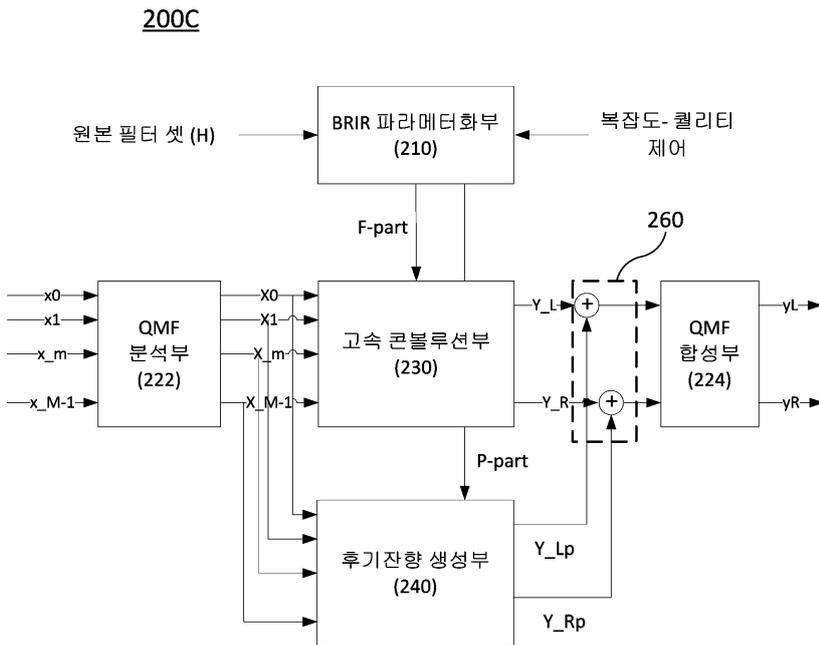
도면3



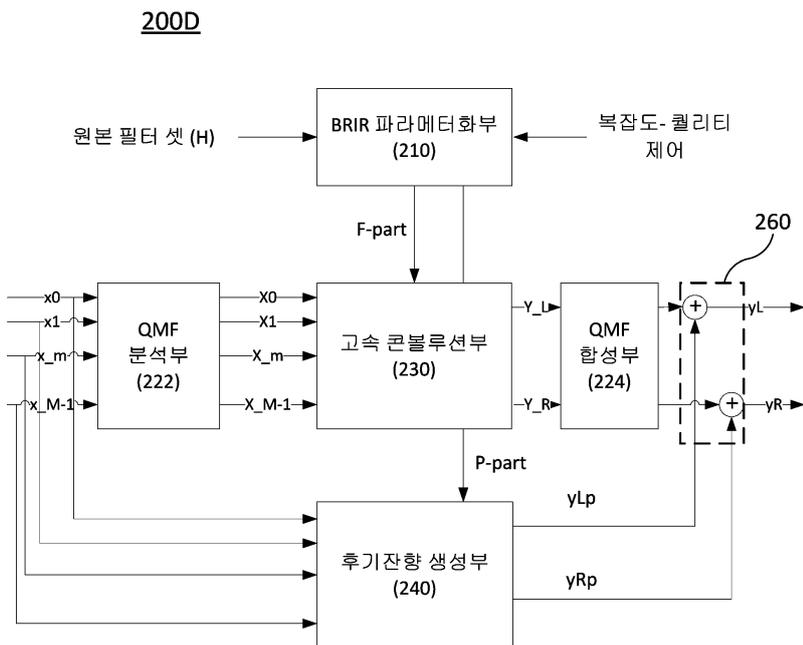
도면4



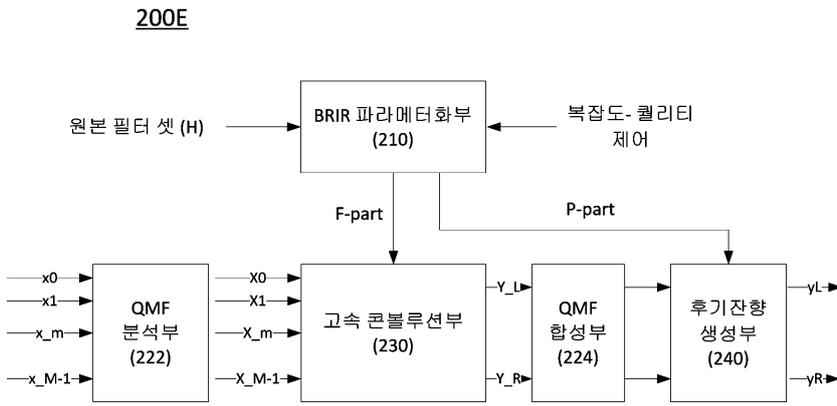
도면5



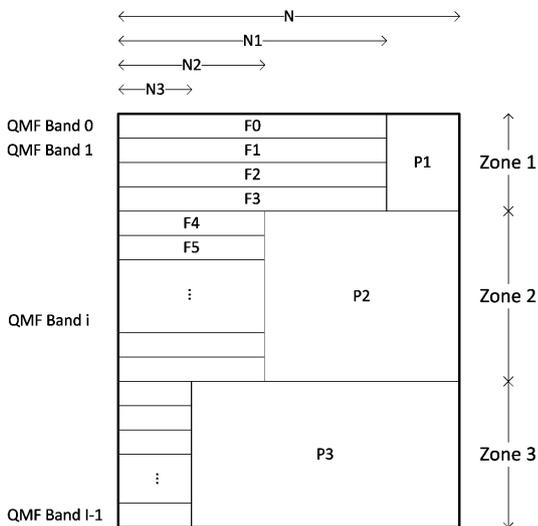
도면6



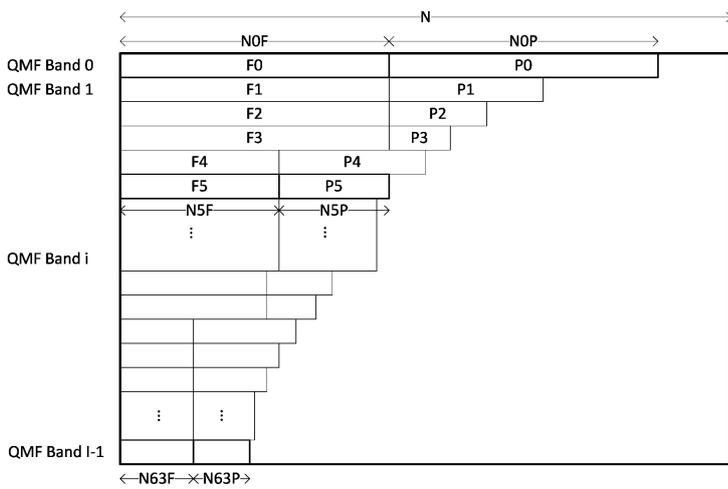
도면7



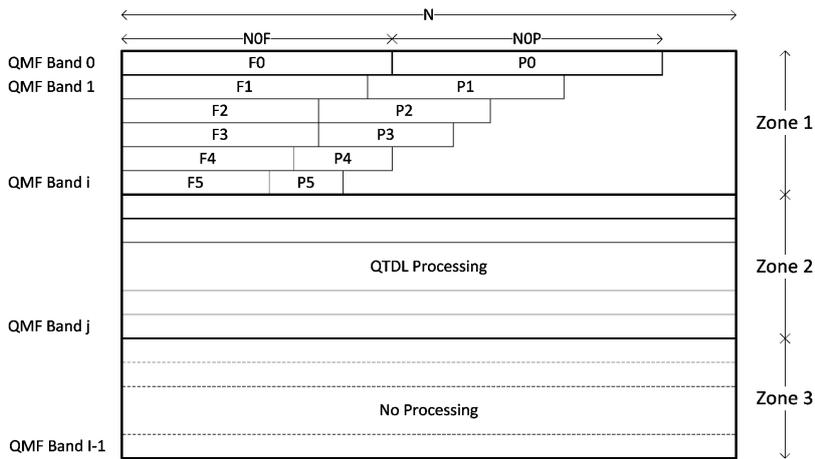
도면8



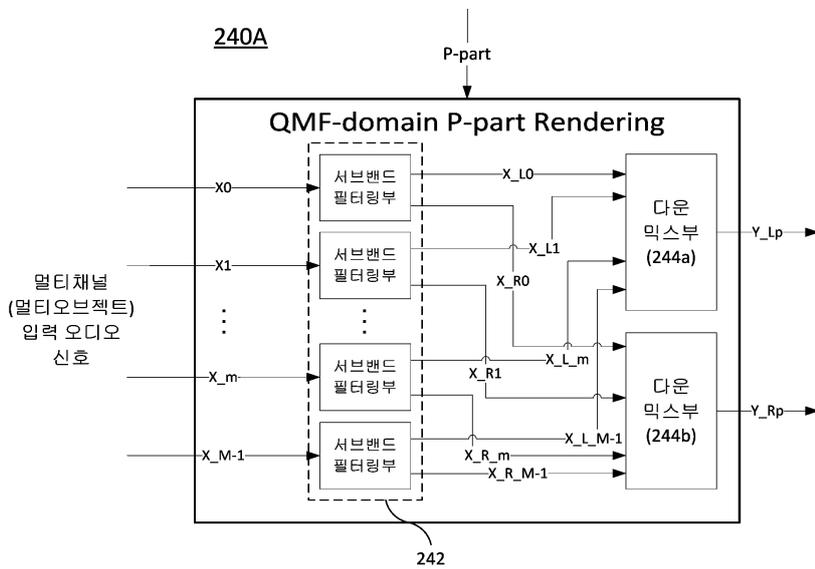
도면9



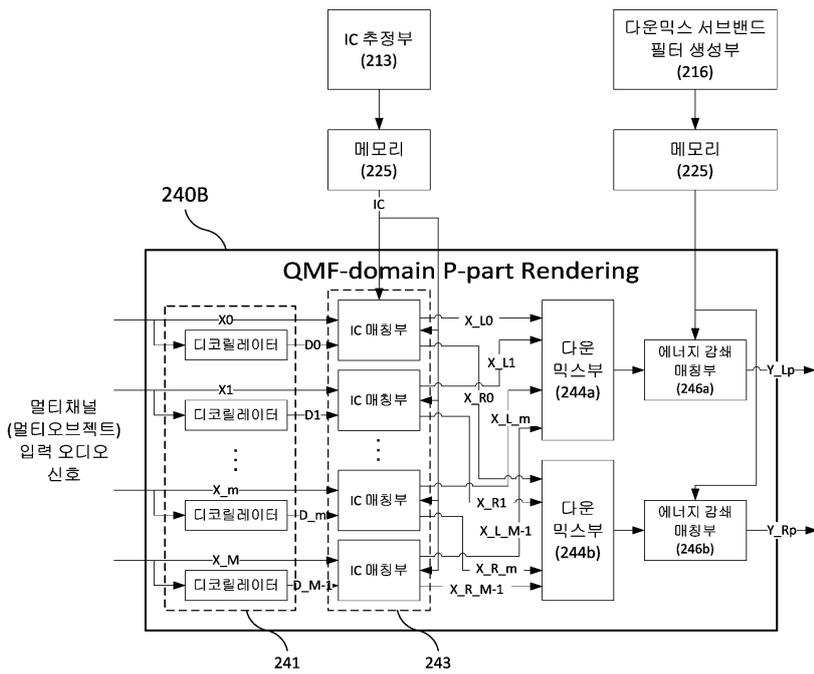
도면10



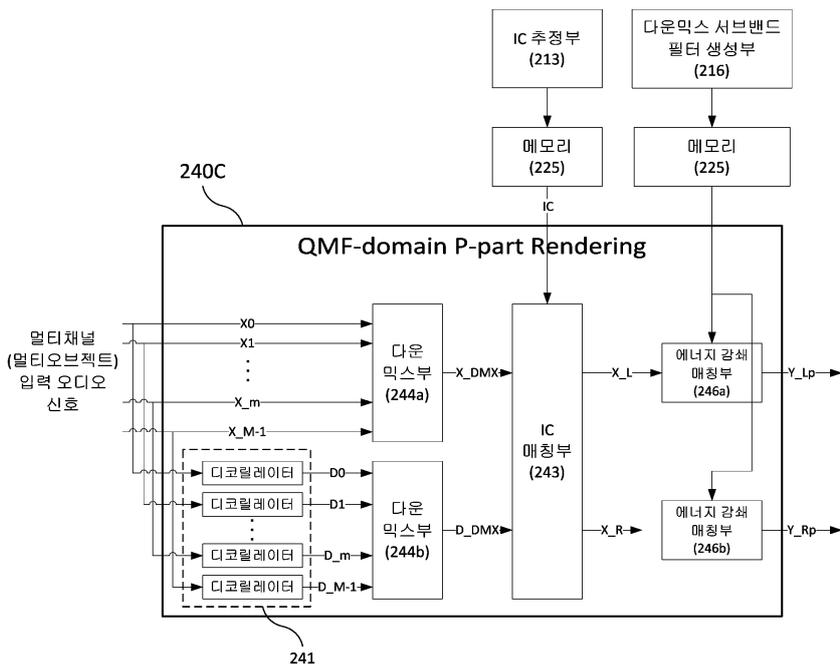
도면11



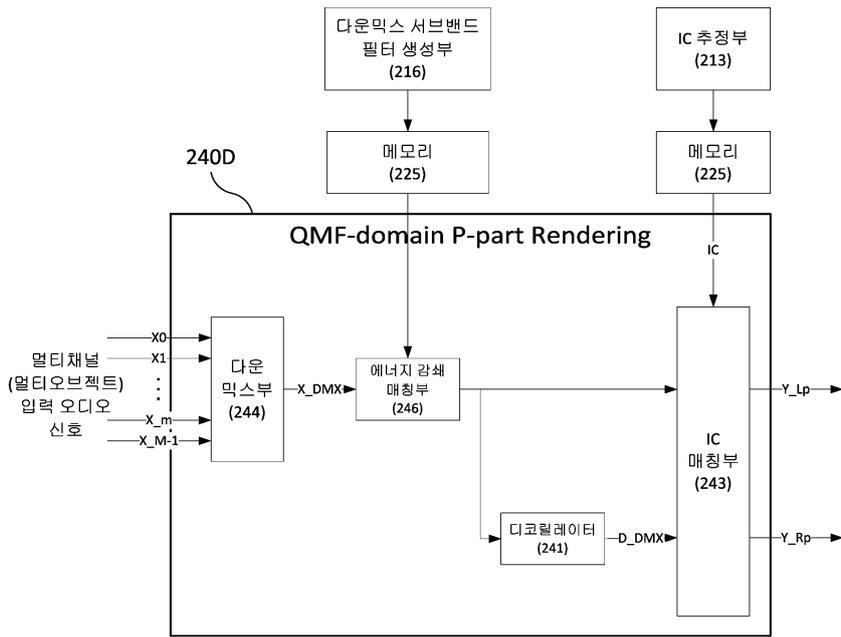
도면12



도면13

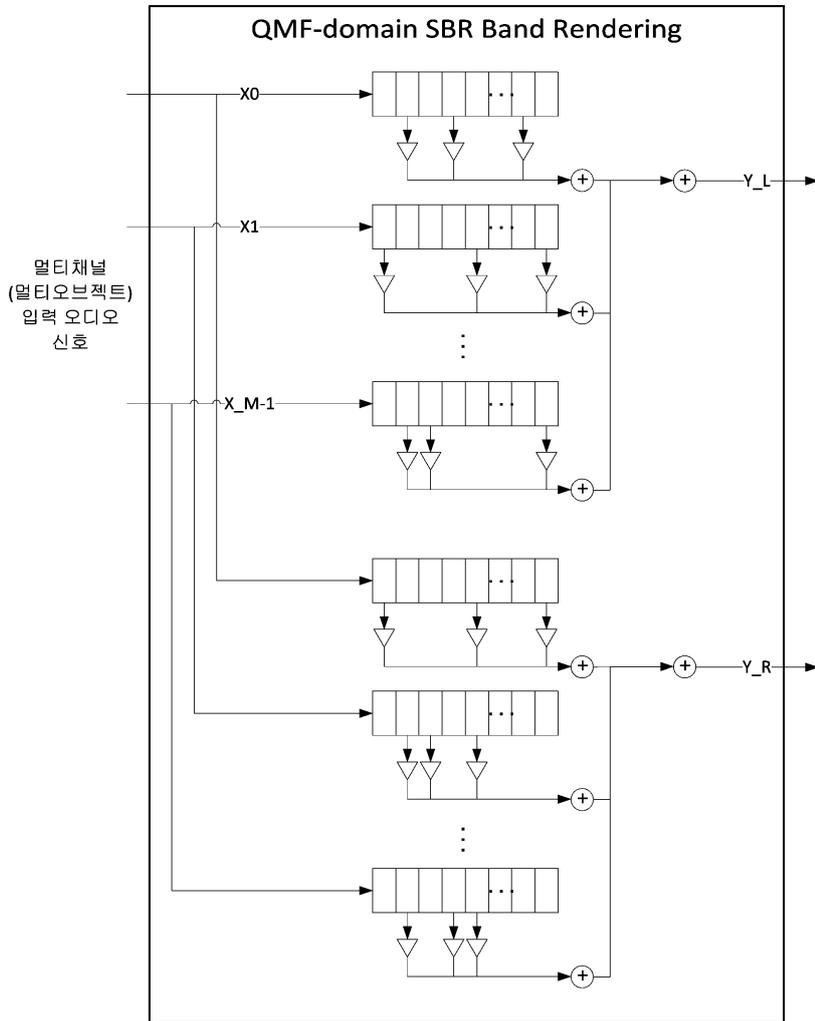


도면14



도면15

250A



도면16

250B

