



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2015년10월22일  
(11) 등록번호 10-1562379  
(24) 등록일자 2015년10월15일

- (51) 국제특허분류(Int. Cl.)  
G10L 19/008 (2014.01) H03M 7/30 (2006.01)  
H04S 3/00 (2006.01)
- (21) 출원번호 10-2014-7036845(분할)
- (22) 출원일자(국제) 2006년08월31일  
심사청구일자 2014년12월29일
- (85) 번역문제출일자 2014년12월29일
- (65) 공개번호 10-2015-0008932
- (43) 공개일자 2015년01월23일
- (62) 원출원 특허 10-2008-7008708  
원출원일자(국제) 2006년08월31일  
심사청구일자 2011년08월29일
- (86) 국제출원번호 PCT/IB2006/053040
- (87) 국제공개번호 WO 2007/031896  
국제공개일자 2007년03월22일
- (30) 우선권주장  
05108405.1 2005년09월13일  
유럽특허청(EPO)(EP)  
06110231.5 2006년02월21일  
유럽특허청(EPO)(EP)
- (56) 선행기술조사문헌  
WO2004028204 A2  
US20050180579 A1  
HERRE, J. et al. THE REFERENCE MODEL  
ARCHITECTURE FOR MPEG SPATIAL AUDIO CODING.  
Proceeding of Audio Engineering Society 118th  
Convention. 2005.05.28. (제1면 내지 제13면)

- (73) 특허권자  
코닌클리케 필립스 엔.브이.  
네덜란드, 아인트호벤 5656 에이이, 하이 테크 캠퍼스 5
- (72) 발명자  
브레바르트, 더크, 제이  
네덜란드, 아인트호벤 아아 엔엘-5656, 홀스트란 6 내
- (74) 대리인  
장훈

전체 청구항 수 : 총 2 항

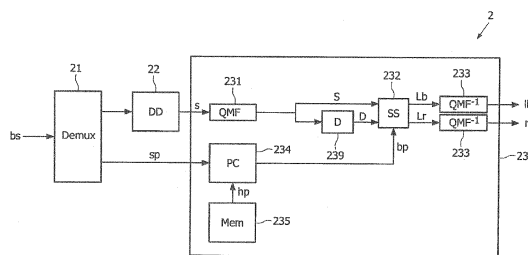
심사관 : 김주식

(54) 발명의 명칭 공간 디코더 유닛 및 한 쌍의 바이노럴 출력 채널들을 생성하기 위한 방법

(57) 요약

공간 디코더 유닛(23)은 하나 이상의 오디오 채널들(s; l; r)을 한 쌍의 바이노럴 출력 채널들(lb, rb)로 변환하기 위해 구성된다. 장치는 공간 파라미터들(sp)을 바이노럴 정보를 포함한 바이노럴 파라미터들(bp)로 변환하기 위한 파라미터 변환 유닛(234)을 포함한다. 장치는 바이노럴 파라미터들(bp)을 이용하면서 부가적으로 오디오 채널(뒷면에 계속)

대표도



널들(L, R)을 한 쌍의 바이노럴 신호들(Lb, Rb)로 변환하기 위한 공간 합성 유닛(232)을 포함한다. 공간 합성 유닛(232)은 QMF 도메인과 같은 변환 도메인에서 동작하는 것이 바람직하다.

---

명세서

청구범위

청구항 1

공간 파라미터들(sp) 및 2개의 오디오 입력 채널(l, r)을 이용하여 한 쌍의 바이노럴(binaural) 출력 채널들(lb, rb)을 생성하기 위한 공간 디코더 유닛(23)에 있어서,

- 파라미터화된 지각 전달 함수들(perceptual transfer functions)(hp)을 이용하여 상기 공간 파라미터들(sp)을 바이노럴 파라미터들(bp)로 변환하기 위한 파라미터 변환 유닛(234)으로서, 상기 바이노럴 파라미터들은 상기 공간 파라미터들 및 상기 파라미터화된 지각 전달 함수들에 의존하는, 상기 파라미터 변환 유닛(234);

- 상기 2개의 오디오 입력 채널을 QMF 도메인의 변환된 오디오 입력 채널로 변환하기 위한 2개의 변환 유닛들:

$$\begin{bmatrix} Lb[k,m] \\ Rb[k,m] \end{bmatrix} = H_k \begin{bmatrix} L[k,m] \\ R[k,m] \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}_k \begin{bmatrix} L[k,m] \\ R[k,m] \end{bmatrix}$$

로서 한 쌍의 변환된 바이노럴 채널들(Lb, Rb)을 합성하기 위한 공간 합성 유닛(232) - 여기서, 인덱스 k는 QMF 하이브리드 주파수 대역 인덱스를 나타내고 인덱스 m은 QMF 슬롯 시간 인덱스를 나타내고, 행렬(H<sub>k</sub>)의 파라미터들(h<sub>ij</sub>)은 상기 바이노럴 파라미터들에 의해 결정되고 L[k,m] 및 R[k,m]은 QMF 변환된 오디오 입력 신호들임 -,

- 상기 한 쌍의 변환된 바이노럴 채널들(Lb, Rb)을 상기 한 쌍의 바이노럴 출력 채널들로 역변환하기 위한 한 쌍의 역변환 유닛들 - 여기서 상기 매트릭스(H<sub>k</sub>)의 파라미터들은

$$h_{11} = m_{11}H_l(l) + m_{21}H_l(r) + m_{31}H_l(c)$$

$$h_{12} = m_{12}H_l(l) + m_{22}H_l(r) + m_{32}H_l(c)$$

$$h_{21} = m_{11}H_r(l) + m_{21}H_r(r) + m_{31}H_r(c)$$

$$h_{22} = m_{12}H_r(l) + m_{22}H_r(r) + m_{32}H_r(c)$$

에 의해 주어지고, 여기서

$$H_l(c) = P_l(c)e^{+j\phi(c)/2}$$

$$H_r(c) = P_r(c)e^{-j\phi(c)/2}$$

$$H_l(l) = \sqrt{w_l^2 P_l^2(lf) + w_r^2 P_l^2(lr)}$$

$$H_r(l) = e^{-j(w_l^2\phi(lf) + w_r^2\phi(lr))} \sqrt{w_l^2 P_l^2(lf) + w_r^2 P_l^2(lr)}$$

$$H_l(r) = e^{+j(w_l^2\phi(rf) + w_r^2\phi(rr))} \sqrt{w_l^2 P_l^2(rf) + w_r^2 P_l^2(rs)} ; \text{ 및}$$

$$H_r(r) = \sqrt{w_l^2 P_r^2(rf) + w_r^2 P_r^2(rr)}$$

이고

$$w_{lf}^2 = \frac{10^{CLD_l/10}}{1 + 10^{CLD_l/10}}$$

$$w_{lr}^2 = \frac{1}{1 + 10^{CLD_l/10}}$$

$$w_{rf}^2 = \frac{10^{CLD_r/10}}{1 + 10^{CLD_r/10}}$$

$$w_{rr}^2 = \frac{1}{1 + 10^{CLD_r/10}}$$

이며,

여기서,

$CLD_l$ 은 좌측 전방 채널 대 좌측 후방 채널에 대한 레벨 차이를 규정하는 공간 파라미터들의 파라미터이고;

$CLD_r$ 는 우측 전방 채널 대 우측 후방 채널에 대한 레벨 차이를 규정하는 공간 파라미터들의 파라미터이고; 그리고

$P_l(c)$ 는 중앙 채널로부터 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 전력 레벨이고;

$P_r(c)$ 는 상기 중앙 채널로부터 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 전력 레벨이고;

$P_l(lf)$ 는 상기 좌측 전방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$P_l(rf)$ 는 상기 우측 전방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$P_l(lr)$ 는 상기 좌측 후방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$P_l(rr)$ 는 상기 우측 후방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$P_r(lf)$ 는 상기 좌측 전방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$P_r(rf)$ 는 상기 우측 전방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$P_r(lr)$ 는 상기 좌측 전방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$P_r(rr)$ 는 상기 우측 후방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$\phi(c)$ 는 상기 중앙 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(lf)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 좌측 전방 채널로부터의 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(rf)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 우측 전방 채널로부터의 상기

우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(lr)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 좌측 후방 채널로부터의 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(rr)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 우측 후방 채널로부터의 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$m_{11}$ ,  $m_{21}$ ,  $m_{31}$ ,  $m_{12}$ ,  $m_{22}$ , 및  $m_{32}$ 는 5.1 MPEG 사라운드 디코더의 2 대 3 업믹스 유닛(two-to-three upmix unit)에 서와 같은 방식으로 상기 공간 파라미터들에 의존하는 것을 특징으로 하는, 공간 디코더 유닛.

**청구항 2**

공간 파라미터들(sp) 및 2개의 오디오 입력 채널(l, r)을 이용하여 한 쌍의 바이노럴(binaural) 출력 채널들 (lb, rb)을 생성하기 위한 방법에 있어서,

- 파라미터화된 지각 전달 함수들(perceptual transfer functions)(hp)을 이용하여 상기 공간 파라미터들(sp) 을 바이노럴 파라미터들(bp)로 변환하는 단계로서, 상기 바이노럴 파라미터들은 상기 공간 파라미터들 및 상기 파라미터화된 지각 전달 함수들에 의존하는, 상기 공간 파라미터들(sp)을 바이노럴 파라미터들(bp)로 변환하는 단계;

- 상기 2개의 오디오 입력 채널을 QMF 도메인의 변환된 오디오 입력 채널로 변환하는 단계:

$$\begin{bmatrix} Lb[k, m] \\ Rb[k, m] \end{bmatrix} = H_k \begin{bmatrix} L[k, m] \\ R[k, m] \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}_k \begin{bmatrix} L[k, m] \\ R[k, m] \end{bmatrix}$$

로서 한 쌍의 변환된 바이노럴 채널들(Lb, Rb)을 합성하는 단계 - 여기서, 인덱스 k는 QMF 하이브리드 주파수 대역 인덱스를 나타내고 인덱스 m은 QMF 슬롯 시간 인덱스를 나타내고, 행렬( $H_k$ )의 파라미터들( $h_{ij}$ )은 상기 바이노럴 파라미터들에 의해 결정되고 L[k,m] 및 R[k,m]은 QMF 변환된 오디오 입력 신호들임 -,

- 상기 한 쌍의 변환된 바이노럴 채널들(Lb, Rb)을 상기 한 쌍의 바이노럴 출력 채널들로 역변환하는 단계 - 여기서 상기 매트릭스( $H_k$ )의 파라미터들은

$$h_{11} = m_{11}H_l(l) + m_{21}H_l(r) + m_{31}H_l(c)$$

$$h_{12} = m_{12}H_l(l) + m_{22}H_l(r) + m_{32}H_l(c)$$

$$h_{21} = m_{11}H_r(l) + m_{21}H_r(r) + m_{31}H_r(c)$$

$$h_{22} = m_{12}H_r(l) + m_{22}H_r(r) + m_{32}H_r(c)$$

에 의해 주어지고, 여기서

$$H_l(c) = P_l(c)e^{+j\phi(c)/2}$$

$$H_r(c) = P_r(c)e^{-j\phi(c)/2}$$

$$H_l(l) = \sqrt{w_l^2 P_l^2(lf) + w_r^2 P_l^2(lr)}$$

$$H_r(l) = e^{-j(w_l^2\phi(lf) + w_r^2\phi(lr))} \sqrt{w_l^2 P_r^2(lf) + w_r^2 P_r^2(lr)}$$

$$H_l(r) = e^{+j(w_{rf}^2\phi(rf)+w_{rr}^2\phi(rr))} \sqrt{w_{rf}^2 P_l^2(rf) + w_{rr}^2 P_l^2(rr)} ; \text{ 및}$$

$$H_r(r) = \sqrt{w_{rf}^2 P_r^2(rf) + w_{rr}^2 P_r^2(rr)}$$

이고

$$w_{lr}^2 = \frac{10^{CLD_l/10}}{1 + 10^{CLD_l/10}}$$

$$w_{lr}^2 = \frac{1}{1 + 10^{CLD_l/10}}$$

$$w_{rf}^2 = \frac{10^{CLD_r/10}}{1 + 10^{CLD_r/10}}$$

$$w_{rr}^2 = \frac{1}{1 + 10^{CLD_r/10}}$$

이며,

여기서,

CLD<sub>l</sub>은 좌측 전방 채널 대 좌측 후방 채널에 대한 레벨 차이를 규정하는 공간 파라미터들의 파라미터이고;

CLD<sub>r</sub>는 우측 전방 채널 대 우측 후방 채널에 대한 레벨 차이를 규정하는 공간 파라미터들의 파라미터이고; 그리고

P<sub>l</sub>(c)는 중앙 채널로부터 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 전력 레벨이고;

P<sub>r</sub>(c)는 상기 중앙 채널로부터 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 전력 레벨이고;

P<sub>l</sub>(lf)는 상기 좌측 전방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

P<sub>l</sub>(rf)는 상기 우측 전방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

P<sub>l</sub>(lr)는 상기 좌측 후방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

P<sub>l</sub>(rr)는 상기 우측 후방 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

P<sub>r</sub>(lf)는 상기 좌측 전방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

P<sub>r</sub>(rf)는 상기 우측 전방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

P<sub>r</sub>(lr)는 상기 좌측 전방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

P<sub>r</sub>(rr)는 상기 우측 후방 채널로부터 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수의 주파수 대역 당 평균 레벨이고;

$\phi(c)$ 는 상기 중앙 채널로부터 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(lf)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 좌측 전방 채널로부터의 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(rf)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 우측 전방 채널로부터의 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(lr)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 좌측 후방 채널로부터의 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$\phi(rr)$ 는 상기 좌측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수와 상기 우측 후방 채널로부터의 상기 우측 바이노럴 채널을 위한 파라미터화된 지각 전달 함수 간의 주파수 대역 당 평균 위상 차이이고;

$m_{11}$ ,  $m_{21}$ ,  $m_{31}$ ,  $m_{12}$ ,  $m_{22}$ , 및  $m_{32}$ 는 5.1 MPEG 사라운드 디코더의 2 대 3 업믹스 유닛(two-to-three upmix unit)에 서와 같은 방식으로 상기 공간 파라미터들에 의존하는 것을 특징으로 하는, 한 쌍의 바이노럴 출력 채널들을 생성하기 위한 방법.

**발명의 설명**

**기술 분야**

[0001]

본 발명은 오디오 코딩에 관한 것이다. 보다 특히, 본 발명은 오디오 입력 신호를 바이노럴 출력 신호로 변환하기 위한 장치 및 방법에 관한 것이며, 여기서 입력 신호는 적어도 하나의 오디오 채널 및 부가 채널들을 나타내는 파라미터들을 포함한다.

**배경 기술**

[0002]

바이노럴(binaural) 오디오 신호들, 즉 인간의 귀가 민감하게 반응하는 특정 방향성 정보를 포함하는 오디오 신호들을 기록 및 재생하는 것이 잘 알려져 있다. 바이노럴 기록들은 일반적으로 인형 사람 머리에 장착된 두 개의 마이크로폰들을 사용하여 이루어지며, 따라서 기록된 사운드는 인간 귀에 의해 캡처된 사운드에 대응하고 머리 및 귀의 형태로 인한 임의의 영향들을 포함한다. 바이노럴 기록들은 바이노럴 기록의 재생이 헤드셋을 요구한다는 점에서 스테레오(즉, 입체음향) 기록들과 상이하지만, 반면 스테레오 기록은 라우드스피커들에 의한 재생을 위해 이루어진다. 바이노럴 기록이 단지 두 개의 채널들만을 이용하여 모든 공간 정보의 재생을 허용하는 반면, 스테레오 기록은 동일한 공간 지각을 제공하지 않을 것이다.

[0003]

정규의 이중 채널(입체 음향) 또는 다중 채널(예로서, 5.1) 기록들은 지각 전달 함수들의 세트로 각각의 정규 신호를 컨볼빙(convolving)함으로써 바이노럴 기록들로 변환될 수 있다. 그러한 지각 전달 함수들은 인간 머리, 및 가능하게는 다른 대상들의 영향을 신호상에 모델링한다. 잘 알려진 유형의 지각 전달 함수는 소위 머리 전달 함수(Head-Related Transfer Function; HRTF)이다. 또한 벽들, 천장, 및 방 바닥에 의해 야기되는 반사들을 고려하는, 대안적 유형의 지각 전달 함수는 바이노럴 룸 임펄스 응답(Binaural Room Impulse Response; BRIR)이다.

[0004]

다중 채널 신호들의 경우에, 지각 함수들의 세트로 신호들을 바이노럴 기록 신호들로 변환하는 것은 모든 채널들의 신호들을 갖는 지각 함수들의 컨볼루션(convolution)을 수반한다. 통상적인 컨볼루션이 계산적으로 요구하기 때문에, 신호들 및 HRTF는 통상적으로 컨볼루션이 계산적으로 훨씬 적게 요구하는 곱셈으로 대체되는 주파수(푸리에) 도메인으로 변환된다.

[0005]

원 채널들 간의 관계들을 나타내는 보다 작은 수의 채널들 및 파라미터들에 의해 채널들의 원래 수를 나타냄으로써 전송되거나 또는 저장될 오디오 채널들의 수를 감소시키는 것이 더 잘 알려져 있다. 따라서, 스테레오 신호들의 세트는 단일 (모노) 채널 더하기 다수의 연관된 공간 파라미터들에 의해 나타내어질 수 있으며, 반면 5.1 신호들의 세트는 두 개의 채널들 및 연관된 공간 채널들의 세트, 또는 심지어 단일 채널 더하기 연관된 공간 파라미터들에 의해 표현될 수 있다. 공간 인코더들에서 이러한 다중 오디오 채널들의 "다운믹싱(downmixing)", 및 공간 디코더들에서 오디오 신호들의 대응하는 "업믹싱(upmixing)"은 통상적으로 변환 도메인

또는 서브밴드 도메인, 예를 들면 QMF(Quadrature Mirror Filter) 도메인에서 실행된다.

- [0006] 특허 상호 조약 특허 출원번호 제W02004/028204호는 머리 전달 함수들을 이용하여 바이노럴 신호들을 생성하는 시스템을 개시한다. 2005년 5월 28일, 뉴욕 음향 엔지니어링 협회 학술 논문의 Herre 등의, 문서 "The Reference Model Architecture for MPEG Spatial Audio Coding"; XP009059973은 MPEG 참조 모델 구조를 개시한다; 2004년 5월 8일 AES 학술에서 제공된 논문들의 견본 인쇄 페이지 1 내지 12, Endegard 등의 문서 "Synthetic Ambiance in Parametric Stereo Coding", XP008048096은 파라메트릭 스테레오 코딩의 예를 개시하며, 2004년 5월 8일, 음향 엔지니어링 협회, 학술 견본 인쇄, Herre 등의 문서 "MP3 Surround: Efficient and Compatible Coding of Multi-Channel Audio"; XP002338414는 MP3 서라운드 인코딩의 예를 개시한다.
- [0007] 다운믹싱된 입력 채널들이 바이노럴 출력 채널들로 변환되는 경우, 종래의 방식은 먼저 업믹싱된 중간 채널들을 생성하기 위해 공간 디코더를 이용하여 입력 채널들을 업믹싱하고, 그 후 이들 업믹싱된 중간 채널들을 바이노럴 채널들로 변환하는 것이다. 이러한 절차는 통상적으로 5 또는 6개의 중간 채널들을 생성하며, 이것은 그 후 두 개의 바이노럴 채널들로 감소되어야 한다. 먼저 채널들의 수를 확장하고 그 후 감소시키는 것은 확실히 효율적이지 않으며 계산 복잡도를 증가시킨다. 또한, 다중 채널 라우드스피커 재생이 되는 5 또는 6개의 중간 채널들을 단지 바이노럴 재생이 되는 2개의 채널들로 감소시키는 것은 필연적으로 아티팩트들을 도입하고 따라서 사운드 품질을 감소시킨다.
- [0008] 상기 참조된 QMF 도메인은 동일하지는 않지만 주파수 (푸리에 변환) 도메인과 유사하다. 공간 디코더가 바이노럴 출력 신호들을 생성한다면, 다운믹싱된 오디오 신호들은 먼저 업믹싱을 위해 QMF 도메인으로 변환되어야 하며, 그 후 시간 도메인 중간 신호들을 생성하기 위해 역으로 QMF 변환되며, 후속하여 (푸리에 변환된) HRTF와 곱셈을 위해 주파수 도메인으로 변환되며, 마지막으로 시간 도메인 출력 신호들을 생성하기 위해 역으로 변환된다. 이러한 절차는 여러 변환들이 연속하여 수행되어야 하므로, 효율적이지 않음이 명백하다.
- [0009] 이러한 종래 기술 방식에 수반된 계산들의 수는 다운믹싱된 오디오 신호들로부터 바이노럴 출력 신호들을 생성할 수 있는, 이동 MP3 플레이어와 같은 핸드헬드 소비자 장치를 설계하는 것을 매우 어렵게 만들 것이다. 비록 그러한 장치가 구현될 수 있을지라도, 그것의 배터리 수명은 요구되는 계산 부하로 인하여 매우 짧을 것이다.

**발명의 내용**

- [0010] 본 발명의 목적은 종래 기술의 이들 및 다른 문제점들을 해결하고, 공간 파라미터들의 연관된 세트 및 하나 이상의 오디오 입력 채널들로 표현되는 다운믹싱된 오디오 채널들의 세트로부터 한 쌍의 바이노럴 출력 채널들을 생성할 수 있는 공간 디코더 유닛을 제공하며, 여기서 디코더는 증가된 효율성을 갖는다.
- [0011] 따라서, 본 발명은 공간 파라미터들 및 하나 이상의 오디오 입력 채널들을 이용하여 한 쌍의 바이노럴 출력 채널들을 생성하기 위한 공간 디코더 유닛을 제공하고, 상기 장치는 파라미터화된 지각 전달 함수들을 이용하여 공간 파라미터들을 바이노럴 파라미터들로 변환하기 위한 파라미터 변환 유닛으로서, 상기 바이노럴 파라미터들은 공간 파라미터들 및 파라미터화된 지각 전달 함수들 모두에 종속하는, 상기 파라미터 변환 유닛; 상기 단일 오디오 입력 채널을 변환된 오디오 채널로 변환하기 위한 단일 변환 유닛; 변환된 역상관된 신호를 생성하기 위해 변환된 오디오 채널을 역상관하는 역상관 유닛; 및 바이노럴 파라미터들을 변환된 오디오 채널(S) 및 변환된 역상관된 신호에 인가함으로써 한 쌍의 변환된 바이노럴 채널들을 합성하기 위한 공간 합성 유닛; 및 변환된 바이노럴 채널들을 한 쌍의 바이노럴 출력 채널들로 역변환하기 위한 한 쌍의 역변환 유닛들을 포함한다.
- [0012] 공간 파라미터들을 바이노럴 파라미터들로 변환함으로써, 공간 합성 유닛은 부가적인 바이노럴 합성 유닛을 요구하지 않고, 한 쌍의 바이노럴 채널들을 직접 합성할 수 있다. 불필요한 중간 신호들이 생성되지 않기 때문에, 아티팩트들의 도입이 실제로 제거되면서 계산 요구사항들은 감소된다.
- [0013] 본 발명의 공간 디코더 유닛에서, 바이노럴 채널들의 합성은 주파수 도메인에서의 변환 및 다음의 시간 도메인에서의 역변환의 부가 단계들을 요구하지 않고, 변환 도메인, 예를 들면, QMF 도메인에서 실행될 수 있다. 두 개의 변환 단계들이 생략될 수 있기 때문에, 계산들의 수 및 메모리 요건들 모두가 상당히 감소된다. 그러므로, 본 발명의 공간 디코더 유닛은 이동 소비자 장치에서 비교적 쉽게 구현될 수 있다.
- [0014] 더욱이, 본 발명의 공간 디코더 유닛에서, 바이노럴 채널들은 다운믹싱된 채널들로부터 직접 생성되며, 각각의 바이노럴 채널은 헤드셋 또는 유사 장치를 이용하여 바이노럴 재생을 위한 바이노럴 신호들을 포함한다. 파라미터 변환 유닛은 공간(즉, 업믹싱) 파라미터들로부터 바이노럴 채널들을 생성하기 위해 사용된 바이노럴 파라미



터들을 유도한다. 이러한 바이노럴 파라미터들의 유도는 HRTF들(Head-Related Transfer Functions) 및/또는 바이노럴 룸 임펄스 응답들(BRIRs)과 같은 파라미터화된 지각 전달 함수들을 수반한다. 그러므로, 본 발명에 따르면, 지각 전달 함수들의 처리는 파라미터 도메인에서 수행되는 반면, 종래 기술에서 이러한 처리는 시간 도메인 또는 주파수 도메인에서 실행되었다. 이것은 파라미터 도메인에서의 해결책이 시간 도메인 또는 주파수 도메인에서의 해결책보다 통상적으로 적기 때문에 계산 복잡도의 추가적인 감소를 가져올 수 있다.

[0015] 파라미터 변환 유닛은 바이노럴 파라미터들을 결정하기 위해, 파라미터 도메인에서, 입력 (다운 믹싱) 오디오 채널들이 바이노럴 채널들로 만들어지는 모든 지각 전달 함수 기여들을 결합하기 위해 구성되는 것이 바람직하다. 달리 말하면, 공간 파라미터들 및 파라미터화된 지각 전달 함수들은 결합된 파라미터들이 업믹싱된 중간 신호들을 수반하는 종래 기술의 방법에서 획득된 것과 유사한 통계적 특성들을 갖는 바이노럴 출력 신호를 야기하는 방식으로 결합된다.

[0016] 바람직한 실시예에서, 본 발명의 공간 디코더 유닛은 오디오 입력 채널들을 변환된 오디오 입력 채널들로 변환하기 위한 하나 이상의 변환 유닛들, 및 한 쌍의 변환된 바이노럴 채널들을 한 쌍의 바이노럴 출력 채널들로 역변환하기 위한 한 쌍의 역변환 유닛들을 더 포함하며, 여기서 공간 합성 유닛은 변환 도메인, 또는 서브 밴드 도메인, 바람직하게는 QMF 도메인에서 동작하도록 구성된다.

[0017] 본 발명의 공간 디코더 유닛은 두 개의 변환 유닛들을 포함할 수 있으며, 파라미터 변환 유닛은 단지 세 개의 채널들을 포함한 지각 전달 함수 파라미터들을 이용하도록 구성되고, 이들 세 개의 채널들 중 두 개의 채널은 합성의 전방 및 후방 채널들의 기여들을 포함한다. 그러한 실시예에서, 파라미터 변환 유닛은 채널 레벨(예로서, CLD), 채널 일관성(예로서, ICC), 채널 예측(예로서, CPC), 및/또는 위상(예로서, IPD) 파라미터들을 처리하기 위해 구성될 수 있다.

[0018] 본 발명의 공간 디코더 유닛은 단일 변환 유닛만을 포함할 수 있으며, 단일 변환 유닛에 의해 변환된 단일 채널 출력을 역상관시키기 위해 역상관 유닛을 더 포함한다. 그러한 실시예에서, 파라미터 변환 유닛은 채널 레벨(예로서, CLD), 채널 일관성(예로서, ICC), 및/또는 위상(예로서, IPD) 파라미터들을 처리하기 위해 구성될 수 있다.

[0019] 본 발명의 공간 디코더 유닛은 스테레오 반향 유닛(stereo reverberation unit)을 부가적으로 포함할 수 있다. 그러한 스테레오 반향 유닛은 변환 도메인 또는 서브 밴드(예로서, QMF) 도메인에서 또는 시간 도메인에서 동작하도록 구성될 수 있다.

[0020] 본 발명은 또한 입력 비트스트림으로부터 한 쌍의 바이노럴 출력 채널들을 생성하기 위해 공간 디코더 장치를 제공하며, 장치는 입력 비트스트림을 단일 다운믹스 채널 및 신호 파라미터들로 역다중화하기 위한 역다중화기 유닛, 단일 다운믹스 채널을 디코딩하기 위한 다운믹스 디코더 유닛, 및 공간 파라미터들 및 단일 다운믹스 채널을 이용하여 한 쌍의 바이노럴 출력 채널들을 생성하기 위한 공간 디코더 유닛을 포함한다.

[0021] 또한, 본 발명은 상기 정의된 공간 디코더 유닛 및/또는 공간 디코더 장치를 포함하는 소비자 장치 및 오디오 시스템을 제공한다. 본 발명은 또한 공간 파라미터들 및 단일 오디오 입력 채널들을 이용하여 한 쌍의 바이노럴 출력 채널들을 생성하는 방법을 제공하며, 상기 방법은 파라미터화된 지각 전달 함수들을 이용하여 공간 파라미터들을 바이노럴 파라미터들로 변환하는 단계로서, 상기 바이노럴 파라미터들은 공간 파라미터들 및 파라미터화된 지각 전달 함수들 모두에 종속하는, 상기 변환 단계; 단일 오디오 입력 채널을 변환된 오디오 채널로 변환하는 단계; 변환된 역상관된 신호를 생성하기 위해 변환된 오디오 채널을 역상관시키는 단계; 미 바이노럴 파라미터들을 변환된 오디오 채널(S) 및 변환된 역상관된 신호로 인가함으로써 한 쌍의 변환된 바이노럴 채널들을 합성하는 단계 및 변환된 바이노럴 채널들을 한 쌍의 바이노럴 출력 채널들로 역변환하는 단계를 포함한다. 본 발명에 따른 방법의 다른 양상들이 이하의 상세한 설명으로부터 명백해질 것이다.

[0022] 또한, 본 발명은 상기 정의된 공간 디코더 유닛 및/또는 공간 디코더 장치를 포함하는 소비자 장치 및 오디오 시스템을 제공한다. 본 발명은 또한 공간 파라미터들 및 하나 이상의 오디오 입력 채널들을 이용하여 한 쌍의 바이노럴 출력 채널들을 생성하는 방법을 제공하며, 상기 방법은 파라미터화된 지각 전달 함수들을 이용하여 공간 파라미터들을 바이노럴 파라미터들로 변환하는 단계, 바이노럴 파라미터들 및 오디오 채널들을 이용하여 한 쌍의 바이노럴 채널들을 합성하는 단계를 포함한다. 본 발명에 따른 방법의 다른 양상들이 이하의 상세한 설명으로부터 명백해질 것이다.

[0023] 본 발명은 상기 정의된 방법을 실행하기 위한 컴퓨터 프로그램 제품을 부가적으로 제공한다. 컴퓨터 프로그램 제품은 CD 또는 DVD와 같은 데이터 캐리어에 저장된 컴퓨터 실행가능한 명령들의 세트를 포함할 수 있다. 프로

그래밍가능한 컴퓨터가 상기 정의된 방법을 실행하도록 허용하는, 컴퓨터 실행가능한 명령들의 세트는 또한 예로서 인터넷을 통해 원격 서버로부터 다운로드하기 위해 이용할 수 있다.

[0024] 본 발명은 첨부된 도면들에 예시된 대표적인 실시예들을 참조하여 이하에 추가로 설명될 것이다.

**도면의 간단한 설명**

[0025] 도 1은 종래 기술에 따른 머리 전달 함수들의 애플리케이션을 개략적으로 도시한 도면.

도 2는 종래 기술에 따른 공간 오디오 인코더 장치를 개략적으로 도시한 도면.

도 3은 바이노럴 합성 장치에 결합된 종래 기술에 따른 공간 오디오 디코더 장치를 개략적으로 도시한 도면.

도 4는 종래 기술에 따른 공간 오디오 디코더 유닛을 개략적으로 도시한 도면.

도 5는 공간 오디오 디코더 유닛의 예를 개략적으로 도시한 도면.

도 6은 본 발명에 따른 공간 오디오 디코더 장치를 개략적으로 도시한 도면.

도 7은 변환 도메인 반향 유닛이 제공된, 도 5의 공간 오디오 디코더 유닛을 개략적으로 도시한 도면.

도 8은 시간 도메인 반향 유닛이 제공된, 도 5의 공간 오디오 디코더 유닛을 개략적으로 도시한 도면.

도 9는 본 발명에 따른 공간 오디오 디코더 장치가 제공된 소비자 장치를 개략적으로 도시한 도면.

[실시예]

종래 기술에 따른, 머리 전달 함수들(HRTF들)과 같은, 지각 전달 함수들의 애플리케이션이 도 1에 개략적으로 도시된다. 바이노럴 합성 장치(3)는 6개의 HRTF 유닛들(31)을 포함하도록 도시되며, 각각은 입력 채널 및 출력 채널의 특정 조합을 위해 전달 함수를 포함한다. 도시된 예에서, 채널들 l(좌), c(중앙), 및 r(우)에 대응할 수 있는 세 개의 오디오 입력 채널들(ch1, ch2, 및 ch3)이 존재한다. 제 1 채널(ch1)은 각각 HRTF(1,L) 및 HRTF(1,R)를 포함하는 두 개의 HRTF 유닛들(31)에 공급된다. 이러한 예에서, HRTF(1,L)은 좌측 바이노럴 신호에 대한 제 1 채널의 기여를 결정하는 머리 전달 함수이다.

이 기술 분야의 숙련자들은 정규 (스테레오) 기록들 및 바이노럴 기록들 모두를 만들고, 정규 기록에 대한 바이노럴 기록의 형상을 나타내는 전달 함수를 유도함으로써 결정될 수 있다. 바이노럴 기록들은 인형의 사람 머리에 장착된 두 개의 마이크로폰들을 이용하여 만들어지며, 따라서 기록된 사운드는 사람 귀에 의해 캡처된 사운드에 대응하고 머리 및 귀의 형태, 심지어 머리카락 및 어깨들의 존재로 인한 임의의 영향들을 포함한다.

HRTF 처리가 시간 도메인에서 일어난다면, HRTF들은 채널들의 (시간 도메인) 오디오 신호들로 컨벌빙된다. 그러나, 일반적으로 HRTF들은 주파수 도메인으로 변환되며, 그 결과적인 전달 함수들 및 오디오 신호들의 주파수 스펙트럼들은 그 후 곱하여진다(푸리에 변환 유닛들 및 역 푸리에 변환 유닛들은 도 1에 도시되지 않음). 시간 프레임들을 오버랩시키는 것을 수반하는 적합한 오버랩 및 부가(Overlap-and Add; OLA) 기술이 빠른 푸리에 변환(FFT) 프레임들보다 긴 길이를 갖는 HRTF들을 수용하기 위해 사용될 수 있다.

적절한 HRTF 유닛(31)에 의한 HRTF 처리 후, 그 결과적인 좌측 및 우측 신호들은 (시간 도메인) 좌측 바이노럴 신호(lb) 및 우측 바이노럴 신호(rb)를 산출하기 위해 각각의 가산기(32)에 의해 가산된다.

도 1의 대표적인 종래 기술의 바이노럴 합성 장치(3)는 세 개의 입력 채널들을 갖는다. 현대의 오디오 시스템들은 종종 소위 5.1 시스템들에서의 경우와 같이, 5 또는 6개의 채널들을 갖는다. 그러나, 전달 및/또는 저장될 데이터의 양을 감소시키기 위해, 다수의 오디오 채널들은 일반적으로 하나 또는 두 개의 채널들로 감소된다("다운믹싱된다"). 원 채널들의 특성들 및 상호 관계들을 나타내는 다수의 신호 파라미터들은 채널들의 원래 수로 하나 또는 두 개 채널들의 확장("업믹싱")을 허용한다. 종래 기술에 따른 대표적인 공간 인코더 장치(1)가 도 2에 개략적으로 예시된다.

공간 인코더 장치(1)는 공간 인코딩(SE) 유닛(11), 다운믹스 인코딩(DE) 유닛(12), 및 다중화기(Mux)(13)를 포함한다. 공간 인코딩 유닛(11)은 다섯 개의 오디오 입력 채널들(lf(좌측 전방), lr(좌측 후방), rf(우측 전방), rr(우측 후방), 및 c(중앙))을 수신한다. 공간 인코딩 유닛(11)은 신호 파라미터들(sp)(공간 인코딩 유닛(11)은 두 개의 채널들(l 및 r) 대신에 단일 채널을 생성할 수 있다)뿐만 아니라, 두 개의 채널들 (l(좌측) 및 r(우

측))을 생성하기 위해 5개의 입력 채널들을 다운믹싱한다. 5개의 채널들이 두 개의 채널들(소위 5-2-5 구성)로 다운믹싱되는, 도시된 실시예에서, 단일 파라미터들(sp)이 예로서 다음을 구성할 수 있다.

파라미터	설명
CPC <sub>1</sub>	2 to 3 변환을 위한 예측/에너지 파라미터
CPC <sub>2</sub>	2 to 3 변환을 위한 예측/에너지 파라미터
CLD <sub>l</sub>	좌측 전방 대 좌측 후방 레벨 차이
CLD <sub>r</sub>	우측 전방 대 우측 후방 레벨 차이
ICC <sub>l</sub>	좌측 전방 대 좌측 후방 상관
ICC <sub>r</sub>	우측 전방 대 우측 후방 상관
ICC <sub>f</sub>	2 to 3 변환을 위한 상관 파라미터
CLD <sub>lfe</sub>	중심 대 lfe 레벨 차이(적용가능하다면)

"lfe"는 선택적인 저 주파수(서브-우퍼) 채널이며, "후방(rear)" 채널들은 또한 "서라운드" 채널들로 알려져 있음을 주의한다.

공간 인코딩 유닛(11)에 의해 생성된 두 개의 다운믹스 채널들(l 및 r)은 다운믹스 인코딩(DE) 유닛(12)으로 공급되며, 이것은 일반적으로 데이터의 양을 감소시킬 목적의 코딩의 유형을 이용한다. 따라서, 인코딩된 다운믹스 채널들(l 및 r), 및 신호 파라미터들(sp)은 출력 비트스트림(bs)을 생성하기 위해 다중화기 유닛(13)에 의해 다중화된다.

대안적 실시예(미도시)에서, 5(또는 6)개의 채널들은 단일(모노) 채널(소위 5-1-5 구성)로 다운믹싱되며, 신호 파라미터들(sp)은 예를 들면 다음을 포함할 수 있다.

파라미터	설명
CLD <sub>fs</sub>	전방 대 후방 레벨 차이
CLD <sub>fc</sub>	전방 대 중앙 레벨 차이
CLD <sub>f</sub>	전방 좌측 대 전방 우측 레벨 차이
CLD <sub>s</sub>	좌측 후방 대 우측 후방 레벨 차이
ICC <sub>fs</sub>	전방 대 후방 상관
ICC <sub>fc</sub>	전방 대 중앙 상관
ICC <sub>f</sub>	전방 좌측 대 전방 우측 상관
ICC <sub>s</sub>	좌측 후방 대 우측 후방 상관
CLD <sub>lfe</sub>	중앙 대 lfe 레벨 차이(적용가능하다면)

이러한 대안적 실시예에서, 인코딩된 다운 믹스 채널(s), 뿐만 아니라 신호 파라미터(sp)는 또한 출력 비트스트림(bs)을 생성하기 위해 다중화기 유닛(13)에 의해 다중화된다.

이러한 비트스트림(bs)이 한 쌍의 바이노럴 채널들을 생성하기 위해 사용되는 것이라면, 종래 기술의 방식은 5 또는 6개의 원 채널들을 생성하기 위해 두 개의 다운 믹스 채널들(l 및 r)(또는 대안적으로, 단일 다운믹스 채널)을 우선 업믹싱할 것이고, 그 후 이들 5 또는 6개의 채널들을 두 개의 바이노럴 채널들로 변환한다. 이러한 종래 기술의 방식의 예가 도 3에 도시된다.

종래 기술에 따른 공간 디코더 장치(2')는 역다중화기(Demux) 유닛(21'), 다운믹스 디코딩 유닛(22'), 및 공간 디코더 유닛(23')을 포함한다. 바이노럴 합성 장치(3)는 공간 디코더 장치(2')의 공간 디코더 유닛(23')에 결합된다.

역다중화기 유닛(21')은 도 2의 비트스트림(bs)과 동일할 수 있는, 비트스트림(bs)을 수신하고, 신호 파라미터들(sp) 및 두 개의 인코딩된 다운믹스 채널들을 출력한다. 신호 파라미터들(sp)이 공간 디코더 유닛(23')에 전송되는 반면, 인코딩된 다운믹스 채널들은 먼저 디코딩된 다운믹스 채널들(l 및 r)을 생성하기 위해 다운믹스 디코딩 유닛(22')에 의해 디코딩된다. 공간 디코더 유닛(23')은 본질적으로 도 2에서의 공간 인코딩 유닛(11)의 역동작들을 실행하고, 5개의 오디오 채널들을 출력한다. 이들 5개의 오디오 채널들은 바이노럴 장치(3)에 공급되며, 이것은 도 3의 장치와 유사한 구조를 가질 수 있지만, 모든 5개의 채널들을 수용하기 위해 부가적인 HRTF 유닛들(31)을 가진다. 도 1에 도시된 바와 같이, 바이노럴 합성 장치(3)는 두 개의 바이노럴 채널들(lb(좌측 바이노럴) 및 rb(우측 바이노럴))을 출력한다.

종래 기술의 공간 디코더 유닛(23')의 대표적인 구조는 도 4에 도시된다. 도 4의 유닛(23')은 2-대-3 업믹스 유닛(230'), 3개의 공간 합성(SS) 유닛들(232'), 및 3개의 역상관(D) 유닛들(239')을 포함한다. 2-대-3 업믹스 유닛(230')은 다운믹스 채널들(l&r) 및 신호 파라미터들(sp)을 수신하고 3개의 채널들(l, r, 및 ce)을 생성한다. 이들 채널들의 각각의 각각의 채널의 역상관된 버전을 생성하는 역상관기 유닛(239')에 공급된다. 각각의 채널

(1, r, 및 ce), 그 각각의 역상관된 버전, 및 연관된 신호 파라미터들(sp)은 각각의 공간 합성(또는 업믹스) 유닛(232')에 공급된다. 채널(1)을 수신하는 공간 합성 유닛(232')은 예를 들면 출력 채널들(1f(좌측 전방) 및 1r(좌측 후방))을 출력한다. 공간 합성 유닛들(232')은 일반적으로 매트릭스 행렬의 곱을 수행하며, 행렬의 파라미터들은 신호 파라미터들(sp)에 의해 결정된다.

도 4의 예에서, 6개의 출력 채널들이 생성됨을 주의한다. 몇몇 실시예들에 있어서, 제 3 역상관 유닛(239') 및 제 3 공간 합성 유닛(232')은 생략될 수 있으며, 따라서 단지 5개의 출력 채널들만을 생성할 수 있다. 그러나, 모든 실시예들에서, 종래 기술의 공간 합성 유닛(23')은 2개 이상의 출력 채널들을 생성할 것이다. 임의의(QMF) 변환 유닛들은 예시를 분명하게 하기 위해 도 4의 단순히 예시적인 예로부터 생략될 수 있음을 또한 주의한다. 실제 예들에서, 공간 디코딩은 QMF 도메인과 같은 변환 도메인에서 수행될 것이다.

도 3의 구성은 효율적이지 않다. 공간 디코더 장치(2')는 2개의 다운믹스 채널들(1 및 r)을 5개의 업믹싱된(중간) 채널들로 변환하지만, 바이노럴 합성 장치(3)는 그 후 5개의 업믹싱된 채널들을 2개의 바이노럴 채널들로 감소시킨다. 또한, 공간 디코더 유닛(2')에서 업믹싱은 일반적으로 QMF(Quadrature Mirror Filter) 도메인과 같은 서브밴드 도메인에서 실행된다. 그러나, 바이노럴 합성 장치(3)는 일반적으로 주파수(즉, 푸리에 변환) 도메인에서 신호들을 처리한다. 이들 두 개의 도메인들이 동일하지 않기 때문에, 공간 디코더 장치(2')는 먼저 다운믹스 채널들의 신호들을 QMF 도메인으로 변환하고, 변환된 신호들을 처리한 후, 업믹싱된 신호들을 시간 도메인으로 되돌려 변환한다. 이어서, 바이노럴 합성 장치(3)는 모든(현재 예에서 5개) 이들 업믹싱된 신호들을 주파수 도메인으로 변환하고, 변환된 신호들을 처리한 후, 바이노럴 신호들을 시간 도메인으로 되돌려 변환한다. 수반된 계산 노력이 상당하며, 보다 효율적인 신호 처리, 특히 이러한 처리가 핸드헬드 장치에서 실행될 경우 요구된다는 것이 명백할 것이다.

본 발명은 공간 디코더 장치에 바이노럴 합성 장치를 통합하고 파라미터에서 바이노럴 처리를 효과적으로 실행함으로써 훨씬 더 효율적인 처리를 제공한다. 공간 디코더 유닛의 예가 도 5에 개략적으로 예시되지만, 본 발명에 따른 조합된 공간 및 바이노럴 디코더 장치(간결함을 위해 공간 디코더 장치로 불리우는)가 도 6에 예시된다.

도 5에 도시된 공간 디코더 유닛(23)은 변환 유닛들(231), 공간 합성(SS) 유닛(232), 역변환 유닛들(233), 파라미터 변환(PC) 유닛(234) 및 메모리(Mem) 유닛(235)을 포함한다. 도 5의 대표적인 예에서, 공간 디코더 유닛(23)은 2개의 변환 유닛들(231)을 포함하지만, 다른 예들에서, 단지 단일 변환 유닛(231)(도 6에서와 같이), 또는 2 이상의 변환 유닛들(231)이 다운믹스 채널들의 수에 종속하여 존재할 수 있다.

변환 유닛들(231) 각각은 다운 믹스 채널 1 및 r을 각각 수신한다(도 3에 보여지는 바와 같이). 각각의 변환 유닛(231)은 각각의 채널(의 신호)을 적합한 변환 또는 서브밴드 도메인으로, 본 예에서는 QMF 도메인으로 변환하도록 구성된다. QMF 변환된 채널들(L 및 R)은 바람직하게는 변환 도메인 바이노럴 채널들(Lb 및 Rb)을 생성하기 위해 채널들(L 및 R)에 매트릭스 동작을 실행하는 공간 합성 유닛(232)에 공급된다. 역변환 유닛들(233)은 본 발명의 예의 역 QMF 변환에서, 바이노럴 시간 도메인 채널들(lb 및 rb)를 생성하기 위해 역변환을 실행한다.

공간 합성 유닛(232)은 도 4의 종래 기술의 공간 합성 유닛(232')과 유사하거나 또는 동일할 수 있다. 그러나, 이러한 유닛에 의해 사용된 파라미터들은 종래 기술에 사용된 것과 상이하다. 보다 특히, 파라미터 변환 유닛(234)은 메모리 유닛(235)에 저장된 HRTF 파라미터들(hp)을 이용하여 종래의 공간 파라미터들(sp)을 바이노럴 파라미터들(bp)로 변환한다. 이들 HRTF 파라미터들(hp)은 다음을 포함할 수 있다.

- 방위각(수평 평면의 각), 고도(수직 평면의 각), 및 거리의 함수로서 좌측 전달 함수를 위한 주파수 대역 당 평균 레벨,
- 방위각, 고도, 및 거리의 함수로서 우측 전달 함수를 위한 주파수 대역 당 평균 레벨, 및
- 방위각, 고도, 및 거리의 함수로서 주파수 대역 당 평균 위상 또는 시간 차.

또한, 다음 파라미터들이 포함될 수 있다.

- 방위각, 고도, 및 거리의 함수로서 HRTF 주파수 대역 당 좌측 및 우측 전달 함수들의 코히어런스 측정치(coherence measure), 및/또는
- 방위각, 고도, 및 거리의 함수로서 좌측 및 우측 전달 함수들에 대한 절대적 위상 및/또는 시간 파라미터들.

실제 사용된 HRTF 파라미터들은 특정 실시예들에 종속할 수 있다.

공간 합성 유닛(232)은 다음의 공식을 이용하여 바이노럴 채널들(Lb 및 Rb)을 결정할 수 있다.

$$\begin{bmatrix} Lb[k,m] \\ Rb[k,m] \end{bmatrix} = H_k \begin{bmatrix} L[k,m] \\ R[k,m] \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}_k \begin{bmatrix} L[k,m] \\ R[k,m] \end{bmatrix} \quad (1)$$

여기서, 인덱스 k는 QMF 하이브리드 (주파수) 대역 인덱스를 나타내고, 인덱스 m은 QMF 슬롯 (시간) 인덱스를 나타낸다. 행렬  $H_k$ 의 파라미터들( $h_{ij}$ )은 바이노럴 파라미터들(도 5에서 bp)에 의해 결정된다. 인덱스 k에 의해 나타내어지는 바와 같이, 행렬  $H_k$ 는 QMF 하이브리드 대역에 종속할 수 있다. 일 예에서, 파라미터 변환 유닛(도 5에서 234)은 공간 합성 유닛(232)에 의해 그 후 행렬 파라미터들( $h_{ij}$ )로 변환되는 바이노럴 파라미터들을 생성한다. 또 다른 예에서, 행렬 파라미터들( $h_{ij}$ )은 파라미터 변환 유닛(도 5에서 234)에 의해 생성된 바이노럴 파라미터들과 동일하며 변환되지 않고 공간 합성 유닛(232)에 의해 직접 사용될 수 있다.

행렬  $H_k$ 의 파라미터들( $h_{ij}$ )은 2개의 다운믹스 채널들(5-2-5 구성)의 경우에서 다음의 방식으로 결정될 수 있다. 도 4의 종래 기술의 공간 디코더 유닛에서, 2-대-3 디코더 유닛(230')은 2개의 (입력) 다운믹스 채널들(l 및 r)을 3개의 (출력) 채널들(l, r, 및 c)로 변환한다(입력 채널들이 다음 논의에서  $l_0$  및  $r_0$ 로 라벨링되는 이유로, 출력 채널들(l 및 r)은 일반적으로 입력 채널들(l 및 r)과 동일하지 않다).

예에 따르면, 파라미터 변환 유닛(도 5 및 도 6에서 234)은 단지 3개의 채널들만(예로서, l, r, 및 c)의 기여가 고려되는 지각 전달 함수 파라미터들을 활용하기 위해 구성되며, 이들 3개의 채널들 중 2개의 채널들(예로서, l 및 r)은 합성의 각 전방(lf, rf) 및 후방(lr, rr) 채널들을 포함한다. 즉, 각각의 전방 및 후방 채널들은 효율성을 향상시키기 위해 함께 그룹핑된다.

2-대-3 업믹스 유닛(230')의 동작은 다음의 행렬 동작에 의해 설명될 수 있다.

$$\begin{bmatrix} l \\ r \\ c \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \\ m_{31} & m_{32} \end{bmatrix} \begin{bmatrix} l_0 \\ r_0 \end{bmatrix} \quad (2)$$

여기서, 행렬 엔트리들( $m_{ij}$ )은 공간 파라미터들에 종속한다. 공간 파라미터들 및 행렬 엔트리들의 관계는 5.1 MPEG 서라운드 디코더의 것과 동일하다. 세 개의 결과로서 생긴 신호들 l, r, 및 c의 각각에 대해서, 그 효과는 이들 사운드 소스들의 원하는 (인지된) 위치에 대응하는 지각 전달 함수(현재 예에서: HRTF) 파라미터들로 결정된다. 중심 채널(c)에 대해서, 사운드 소스 위치의 공간 파라미터들은 직접 이용될 수 있으며, 중심을 위한 두 개의 출력 신호들( $l_B(c)$  및  $r_B(c)$ )을 야기한다.

$$\begin{bmatrix} l_B(c) \\ r_B(c) \end{bmatrix} = \begin{bmatrix} H_l(c) \\ H_r(c) \end{bmatrix} c = \begin{bmatrix} P_l(c)e^{+j\phi(c)/2} \\ P_r(c)e^{-j\phi(c)/2} \end{bmatrix} c \quad (3)$$

식(3)으로부터 관찰될 수 있는 바와 같이, HRTF 파라미터 처리는 중심 채널의 사운드 소스 위치에 대응하는, 평균 전력 레벨들( $P_l$  및  $P_r$ )과 산호의 곱으로 이루어지는 반면, 위상 차이는 대칭적으로 분포된다. 이러한 처리는 한편으로는 HRTF 파라미터들에서 QMF 필터 뱅크로의 매핑, 및 다른 한편으로는 공간 파라미터들에서 QMF 대역으로의 매핑을 이용하여, 각각의 QMF 대역을 위해 독립적으로 수행된다.

좌측(l) 채널에 대해서, 좌측 전방 및 좌측 후방 채널들로부터의 HRTF 파라미터들은 가중치들( $w_{lf}$  및  $w_{rf}$ )을 이용하여 단일 기여로 결합된다. 그 결과 합성 파라미터들이 통계적으로 전방 및 후방 채널들 모두의 효과를 시뮬레이션한다. 다음 식들은 좌측 채널에 대한 바이노럴 출력쌍( $l_b$ ,  $r_b$ )을 생성하기 위해 사용된다.

$$\begin{bmatrix} l_b(l) \\ r_b(l) \end{bmatrix} = \begin{bmatrix} H_l(l) \\ H_r(l) \end{bmatrix} l \quad (4)$$

여기서,

$$H_l(l) = \sqrt{w_f^2 P_l^2(lf) + w_r^2 P_l^2(lr)} \quad (5)$$

및

$$H_r(l) = e^{-j(w_f^2 \phi(lf) + w_r^2 \phi(lr))} \sqrt{w_f^2 P_r^2(lf) + w_r^2 P_r^2(lr)} \quad (6)$$

가중들( $w_{lf}$  및  $w_{lr}$ )은  $lf$  및  $lr$ 을 위한 1-대-2 유닛의 CLD 파라미터에 의존한다.

$$w_f^2 = \frac{10^{CLD_r/10}}{1+10^{CLD_r/10}} \quad (7)$$

$$w_r^2 = \frac{1}{1+10^{CLD_r/10}} \quad (8)$$

유사한 방식으로, 우측 채널에 대한 바이노럴 출력은 다음에 따라 획득된다.

$$\begin{bmatrix} L_b(r) \\ R_b(r) \end{bmatrix} = \begin{bmatrix} H_l(r) \\ H_r(r) \end{bmatrix} r \quad (9)$$

여기서,

$$H_l(r) = e^{+j(w_f^2 \phi(rf) + w_r^2 \phi(rr))} \sqrt{w_f^2 P_l^2(rf) + w_r^2 P_l^2(rs)} \quad (10)$$

$$H_r(r) = \sqrt{w_f^2 P_r^2(rf) + w_r^2 P_r^2(rr)} \quad (11)$$

$$w_f^2 = \frac{10^{CLD_r/10}}{1+10^{CLD_r/10}} \quad (12)$$

$$w_r^2 = \frac{1}{1+10^{CLD_r/10}} \quad (13)$$

위상 변경 기간은 양쪽 경우들 모두에서 대측성 귀(contra-lateral ear)에 적용됨을 주의한다. 더욱이, 인간의 청각 시스템은 근사치 2 kHz 위 주파수들에 대한 바이노럴 위상에 크게 영향을 받지 않기 때문에, 위상 변경 기간은 단지 낮은 주파수 영역에서 적용되도록 요구된다. 따라서, 주파수 범위의 나머지에 대해서, 실수치 처리가 충분하다(실수치  $m_{ij}$ 를 가정하면).

그것은 또한 상기 식들이  $lf$  및  $lr$ 의 (HRTF) 필터링된 신호들의 비간섭성 부가를 가정함을 주의한다. 하나의 가능한 확장은 전방/후방 상관에 대한 고려뿐만 아니라 식들에서  $lf$  및  $lr$ (및  $lf$  및  $rr$ )의 전송된 채널 내 코히어런스(Inter-Channel Coherence; ICC)을 포함하는 것이다.

상술된 모든 처리 단계들은 단일, 신호-도메인 2x2 행렬을 야기하도록 파라미터 도메인에서 조합될 수 있다.

$$\begin{bmatrix} l_b \\ r_b \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} l_0 \\ r_0 \end{bmatrix} \quad (14)$$

$$h_{11} = m_{11}H_l(l) + m_{21}H_l(r) + m_{31}H_l(c) \quad (15a)$$

$$h_{12} = m_{12}H_l(l) + m_{22}H_l(r) + m_{32}H_l(c) \quad (15b)$$

$$h_{21} = m_{11}H_r(l) + m_{21}H_r(r) + m_{31}H_r(c) \quad (15c)$$

$$h_{22} = m_{12}H_r(l) + m_{22}H_r(r) + m_{32}H_r(c) \quad (15d)$$

본 발명은 본질적으로 종래 기술에서와 같은 주파수 또는 시간 도메인 대신에, 파라미터 도메인에서 바이노럴 (즉, HRTF) 정보를 처리한다. 이러한 방식으로, 상당한 계산 절약이 달성될 수 있다.

도 6에서 단지 비제한적 예로 도시된 본 발명에 따른 공간 디코더 장치(2)는 역다중화기(Demux) 유닛(21), 다운믹스 디코딩 유닛(22), 및 공간/바이노럴 디코더 유닛(23)을 포함한다. 역다중화기 유닛(21) 및 다운믹스 디코딩 유닛(22)은 도 3에 도시된 다운믹스 디코딩 유닛(22') 및 종래 기술의 역다중화기 유닛(21')과 유사할 것이다. 도 6의 공간 디코더 유닛(23)은 다운믹스 채널들 및 연관된 변환 유닛들의 수를 제외하고, 도 5의 공간 디코더 유닛(23)과 동일하다. 도 6의 공간 디코더 장치가 단일 다운믹스 채널(s)을 위해 구성되기 때문에, 단지 단일 변환 유닛(231)은 역상관(D) 유닛(239)이 (변환 도메인) 다운믹스 신호(S)의 역상관된 버전(D)을 생성하기 위해 부가되는 동안 제공된다. 단일의 다운믹스 채널(s)에 연관된 단일 파라미터들(sp)이 두 개의 다운믹스 채널들과 연관된 것과 상이하기 때문에, 파라미터 변환 유닛(234)에 의해 생성된 바이노럴 파라미터들(bp)은 일반적으로 도 5의 실시예의 것과 상이하다.

도 6의 구성에서, 바이노럴 디코더의 입력은 공간 파라미터들(sp)에 의해 수반된 모노 입력 신호(s)를 포함한다. 바이노럴 합성 유닛은 원래의 5.1 유닛의 HRTF 처리로부터 기인하는 그들 특성들에 근사하는 통계적 특성들로 스테레오 출력 신호를 생성하며, 이것은 다음에 의해 설명될 수 있다.

$$l_b = H_l(lf) \otimes lf + H_l(rf) \otimes rf + H_l(lr) \otimes lr + H_l(Rr) \otimes Rr + H_l(c) \otimes c \quad (16)$$

$$r_b = H_r(lf) \otimes lf + H_r(rf) \otimes rf + H_r(lr) \otimes lr + H_r(rr) \otimes rr + H_r(c) \otimes c \quad (17)$$

통계적 특성들, 채널들(lf, rf, lr, rr, 및 c)의 상호관계, 및 HRTF 임펄스 응답들의 파라미터를 설명하는 공간 파라미터들이 주어지면, 바이노럴 출력 쌍( $l_b, r_b$ )의 통계적 특성들(즉, 바이노럴 파라미터들의 근사치)을 추정하는 것이 또한 가능하다. 보다 특히, 평균 에너지(각 채널에 대해), 평균 위상차, 및 코히어런스가 추정되며, 이어서 모노 입력 신호의 역상관 및 매트릭싱(matrixing)에 의해 원상태로 된다.

바이노럴 파라미터들은 두 개의 바이노럴 출력 채널들의 각각에 대한 (상대적) 레벨 변화(및 이후 채널 레벨 차 파라미터를 정의함), (평균) 위상차, 및 코히어런스 측정(변환 도메인 시간/주파수 타일 당)을 포함한다.

제 1 단계로서, 5(또는 6) 채널(5.1) 신호의 상대적 파워들(모노 입력 신호의 파워에 대한)은 전송된 CLD 파라미터들을 이용하여 계산된다. 좌측 전방 채널의 상대적 파워는 다음에 의해 제공된다.

$$\sigma_{lf}^2 = r_1(\text{CLD}_{fs})r_1(\text{CLD}_{fc})r_1(\text{CLD}_f) \quad (18)$$

여기서,

$$r_1(\text{CLD}) = \frac{10^{\text{CLD}/10}}{1 + 10^{\text{CLD}/10}} \quad (19)$$

및

$$r_2(\text{CLD}) = \frac{1}{1 + 10^{\text{CLD}/10}} \quad (20)$$

유사하게, 다른 채널들의 상대적 파워들이 다음에 의해 제공된다.

$$\sigma_{rf}^2 = r_1(\text{CLD}_{fs})r_1(\text{CLD}_{fc})r_2(\text{CLD}_f) \quad (21a)$$

$$\sigma_c^2 = r_1(\text{CLD}_{fs})r_2(\text{CLD}_{fc}) \quad (21b)$$

$$\sigma_b^2 = r_2(\text{CLD}_b) r_1(\text{CLD}_s) \quad (21c)$$

$$\sigma_r^2 = r_2(\text{CLD}_b) r_2(\text{CLD}_s) \quad (21d)$$

좌측 바이노럴 출력 채널의 상대적 파워의 기대치( $\sigma_L^2$ ) (모노 입력 채널에 대한), 우측 바이노럴 출력 채널의 상대적 파워의 기대치( $\sigma_R^2$ ), 및 외적(cross product)의 기대치( $L_B R_B^*$ )가 그 후 계산될 수 있다. 바이노럴 출력의 코히어런스(ICC<sub>B</sub>)가 그 후 다음에 의해 제공된다.

$$\text{ICC}_B = \frac{\langle L_B R_B^* \rangle}{\sigma_L \sigma_R} \quad (22)$$

및 평균 위상 각(IPD<sub>B</sub>)가 다음에 의해 제공된다.

$$\text{IPD}_B = \arg\langle L_B R_B^* \rangle \quad (23)$$

바이노럴 출력의 채널 레벨차(CLD<sub>B</sub>)가 다음에 의해 제공된다.

$$\text{CLD}_B = 10 \log_{10} \left( \frac{\sigma_L^2}{\sigma_R^2} \right) \quad (24)$$

마지막으로, 모노 입력에 비교되는 바이노럴 출력의 전체 (선형) 이득( $g_B$ )은 다음에 의해 제공된다.

$$g_B = \sqrt{\sigma_L^2 + \sigma_R^2} \quad (25)$$

바이노럴 매트릭스에서 IPD<sub>B</sub>, CLD<sub>B</sub>, ICC<sub>B</sub>, 및  $g_B$  파라미터들을 원상태로 하기 위해 요구되는 매트릭스 계수들은 전체 이득들( $g_B$ )로 확장된, 종래의 파라메트릭 스테레오 디코더로부터 간단히 획득된다.

$$h_{11} = g_B c_L \cos(\alpha + \beta) \exp(j\text{IPD}_B / 2) \quad (26a)$$

$$h_{12} = g_B c_L \sin(\alpha + \beta) \exp(j\text{IPD}_B / 2) \quad (26b)$$

$$h_{21} = g_B c_R \cos(-\alpha + \beta) \exp(-j\text{IPD}_B / 2) \quad (26c)$$

$$h_{22} = g_B c_R \sin(-\alpha + \beta) \exp(-j\text{IPD}_B / 2) \quad (26d)$$

여기서,

$$\alpha = 0.5 \arccos(\text{ICC}_B) \quad (27)$$

$$\beta = \arctan \left( \frac{c_R - c_L}{c_R + c_L} \tan(\alpha) \right) \quad (28)$$

$$c_L = \sqrt{\frac{10^{\text{CLD}_B / 10}}{1 + 10^{\text{CLD}_B / 10}}} \quad (29)$$



$$c_r = \sqrt{\frac{1}{1+10^{CLD_p/10}}} \quad (30)$$

본 발명의 공간 디코더 유닛의 다른 실시예들은 반향 유닛을 포함할 수 있다. 반향을 추가하는 것은 바이노럴 사운드가 생성될 때 감지된 거리를 향상시킨다는 것이 알려져 있다. 이러한 이유로, 도 7의 공간 디코더 유닛(23)에는 공간 합성 유닛(232)과 평행하여 접속된 스테레오 반향 유닛(237)이 제공된다. 도 7의 스테레오 반향 유닛(237)은 QMF 변환 도메인 단일 다운믹스 신호(S)를 수신하고, 부가 유닛들(238)에 의해 변환 도메인 바이노럴 신호들(도 6에서 채널들(Lb 및 Lr))에 부가되는, 두 개의 반향 신호들을 출력한다. 조합된 신호들은 그 후 출력하기 전에 역변환 유닛들(233)에 의해 역변환된다.

도 8의 실시예에서, 스테레오 반향 유닛(237)은 시간 도메인에서 반향을 생성하기 위해 구성되고 시간 도메인 단일 다운믹스 신호(s)를 수신한다. 스테레오 반향 유닛(237)은 시간 도메인 반향 신호들을 출력하며, 이것은 부가 유닛들(238)에 의해 바이노럴 채널들(lb 및 rb)의 시간 도메인 신호들에 부가된다. 이들 중 하나의 실시예가 적절한 반향을 제공한다.

본 발명은 핸드 헬드 소비자 장치, 및 상술된 바와 같은 공간 디코더 유닛 또는 공간 디코더 장치를 포함하는 오디오 시스템과 같은 소비자 장치를 제공한다. 핸드헬드 소비자 장치는 MP3 플레이어 또는 유사한 장치에 의해 구성될 수 있다. 소비자 장치는 도 9에 개략적으로 예시된다. 소비자 장치(50)는 본 발명에 따른 공간 디코더 장치(2)를 포함하도록 도시된다(도 6 참조).

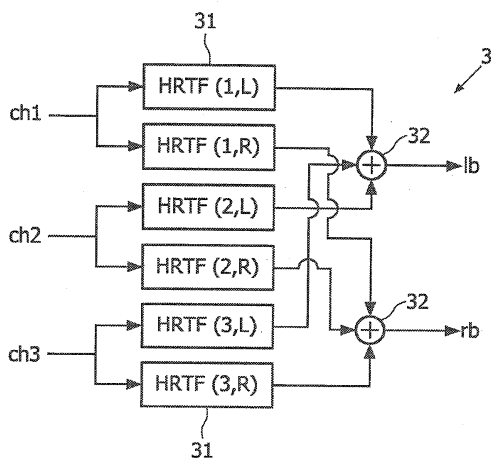
본 발명은 조합된 공간 디코더 장치 및 바이노럴 합성 장치의 계산 복잡도가 바이노럴 정보에 따라 공간 파라미터들을 변경함으로써 상당히 감소될 수 있다는 통찰에 기초한다. 이것은 공간 디코더 장치로 하여금 임의의 아티팩트들의 도입을 방지하면서, 동일한 신호 처리 동작에서 공간 디코딩 및 지각 전달 함수 처리를 효과적으로 실행하도록 허용한다.

본 명세서에 사용된 임의의 용어는 본 발명의 영역을 제한하기 위해 해석되어서는 안된다는 것을 주의한다. 특히, 단어들 "포함하다" 및 "포함하는"은 명확하게 서술되지 않은 임의의 요소들을 배제하는 것을 의미하지 않는다. 단일 (회로) 소자들은 다수의 (회로) 소자들 또는 그 등가물들로 대체될 수 있다.

이 기술분야의 숙련자들은 본 발명이 상기 예시된 실시예에 제한되지 않으며, 많은 변경들 및 부가들이 첨부한 청구항들에서 정의된 바와 같이 본 발명의 범위로부터 벗어나지 않고 이루어질 수 있음을 이해할 것이다.

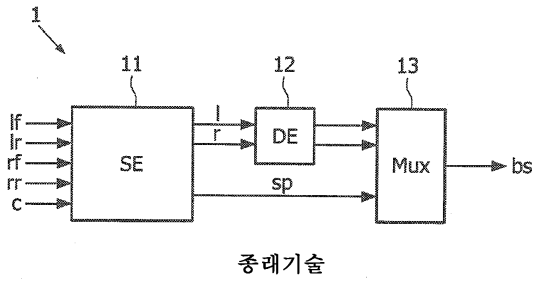
**도면**

**도면1**

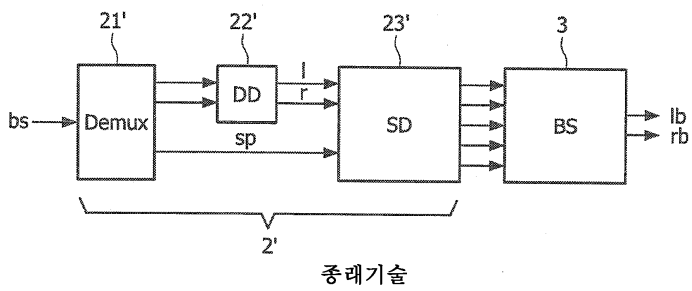


종래기술

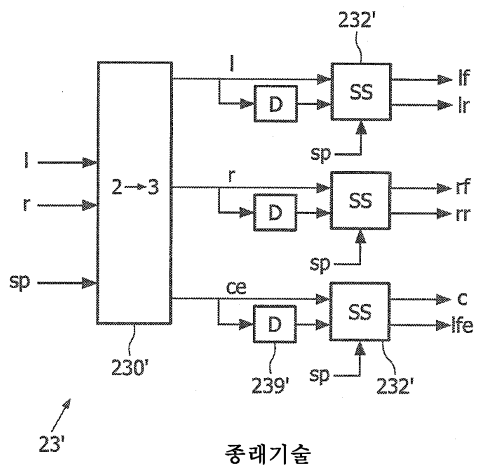
도면2



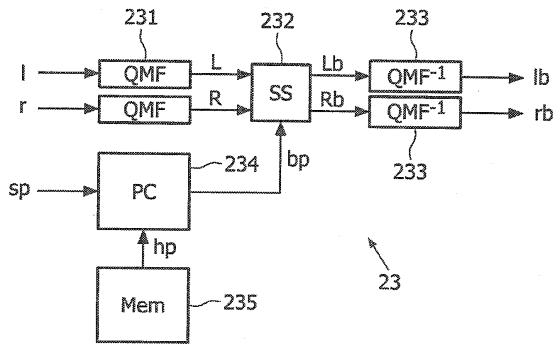
도면3



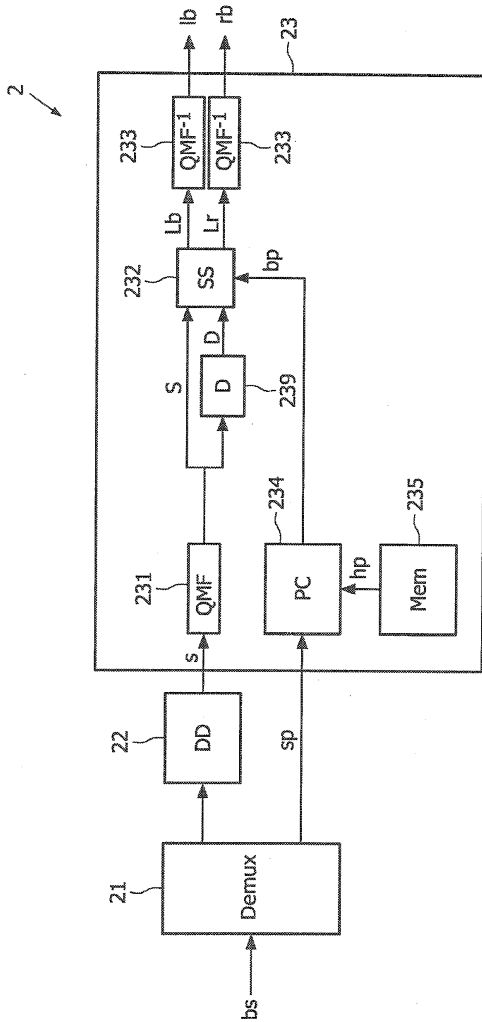
도면4



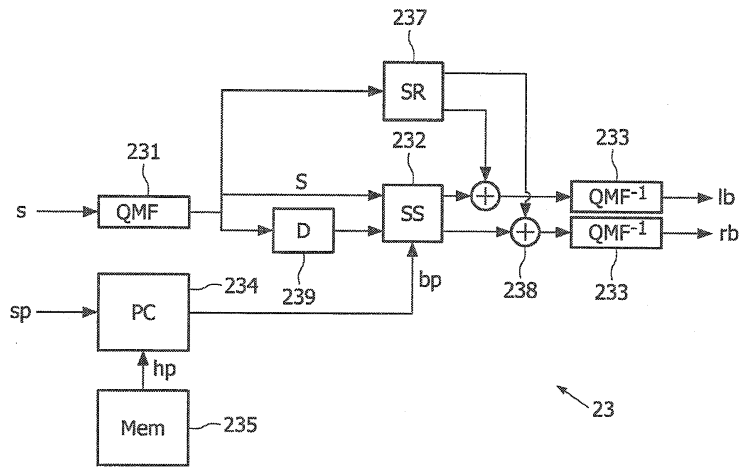
도면5



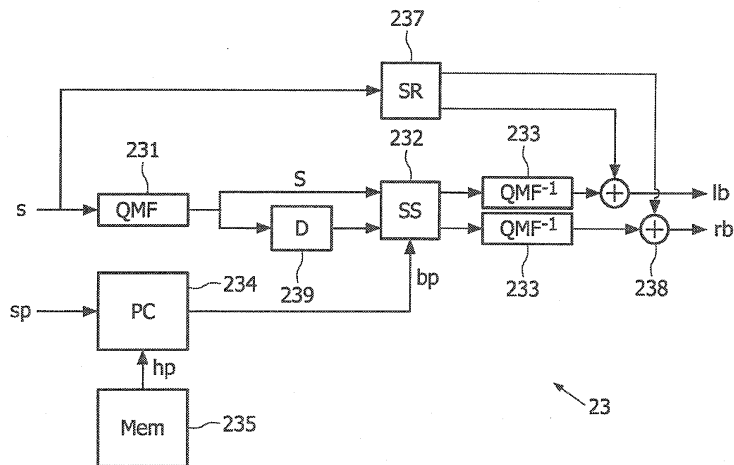
도면6



도면7



도면8



도면9

