



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년04월12일
(11) 등록번호 10-1726205
(24) 등록일자 2017년04월06일

- (51) 국제특허분류(Int. Cl.)
G10L 19/032 (2013.01) G10L 19/008 (2014.01)
G10L 19/16 (2013.01)
- (52) CPC특허분류
G10L 19/032 (2013.01)
G10L 19/008 (2013.01)
- (21) 출원번호 10-2015-7011796
- (22) 출원일자(국제) 2013년11월04일
심사청구일자 2015년05월06일
- (85) 번역문제출일자 2015년05월06일
- (65) 공개번호 10-2015-0066565
- (43) 공개일자 2015년06월16일
- (86) 국제출원번호 PCT/EP2013/072961
- (87) 국제공개번호 WO 2014/072260
국제공개일자 2014년05월15일
- (30) 우선권주장
61/723,687 2012년11월07일 미국(US)
- (56) 선행기술조사문헌
JP2000347679 A*
KR1020120057384 A*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
돌비 인터네셔널 에이비
네덜란드 1101 씨엔 암스트레담 주이두스트 헤리
커베르그백 1-35 3이 아폴로 빌딩
돌비 레버러토리즈 라이쎬싱 코오포레이션
미합중국, 캘리포니아 94103, 샌프란시스코, 마켓
스트리트 1275
- (72) 발명자
슈그, 미하엘
독일 90429 뉘른베르크 도이치헤른스트라쎄 15-19
돌비 저머니 게엠베하 내
윌리엄스, 필립
미국 캘리포니아 94103-4938 샌프란시스코 브래넨
스트리트 999 돌비 레버러토리즈 인코포레이티드
내
- (74) 대리인
장훈

전체 청구항 수 : 총 14 항

심사관 : 옥윤철

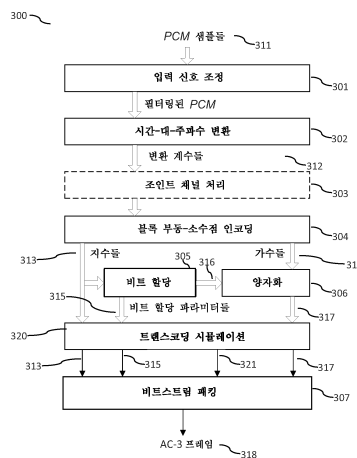
(54) 발명의 명칭 감소된 복잡성 변환기 SNR 계산

(57) 요약

본 문서는 오디오 인코딩/디코딩에 관한 것이다. 특히 본 문서는 오디오 인코딩/디코딩의 컨텍스트에서 이용되는 비트 할당 처리의 복잡성을 감소시키기 위한 방법 및 시스템에 관한 것이다. 제 1 오디오 코덱 시스템에 따라 오디오 신호를 인코딩하도록 구성된 오디오 인코더(300)가 기술된다. 오디오 인코더(300)는 오디오 신호에 기초하

(뒷면에 계속)

대표도 - 도3



여 스펙트럼 계수들(312)의 세트를 결정하도록 구성된 변환 유닛(302)을 포함한다. 또한, 인코더(300)는 스펙트럼 계수들(312)의 세트에 기초하여, 스케일 팩터들의 세트 및 스케일링된 값들(314)의 세트를 결정하고; 스케일 팩터들의 세트를 인코딩하여 인코딩된 스케일 팩터들(313)의 세트를 산출하도록 구성된 부동-소수점 인코딩 유닛(304)을 포함한다. 또한, 인코더(300)는, 제 1 타겟 데이터-레이트에 기초하고 인코딩된 스케일 팩터들(313)의 세트를 위해 이용된 비트들의 수에 기초하여, 스케일링된 값들(314)의 세트를 양자화하기 위해 이용가능한 비트들의 총수를 결정하고; 스케일링된 값들(314)의 세트의 스케일링된 값들을 양자화하기 위해 이용가능한 비트들의 총수의 할당을 표시하는 제 1 제어 파라미터(315)를 결정하고; 제 1 제어 파라미터(315)에 따라 스케일링된 값들(314)의 세트를 양자화하여 양자화된 스케일링된 값들(317)의 세트를 산출하도록 구성된 비트 할당 및 양자화 유닛(305, 306)을 포함한다. 또한, 인코더(300)는 제 1 제어 파라미터(315)에 기초하여 제 2 제어 파라미터(321)를 결정하도록 구성된 트랜스코딩 시뮬레이션 유닛(320)을 포함하고; 제 2 제어 파라미터(321)는 트랜스코더가 제 1 비트스트림을 제 2 타겟 데이터-레이트의 제 2 비트스트림으로 변환할 수 있게 하고; 제 2 비트스트림은 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따르고; 제 1 비트스트림은 제 2 제어 파라미터를 포함한다.

(52) CPC특허분류

G10L 19/173 (2013.01)

명세서

청구범위

청구항 1

삭제

청구항 2

제 1 오디오 코덱 시스템에 따라 오디오 신호의 프레임을 인코딩하고, 그에 의해 제 1 타겟 데이터-레이트의 제 1 비트스트림을 산출하도록 구성된 오디오 인코더(300)에 있어서,

- 상기 오디오 신호의 상기 프레임에 기초하여 스펙트럼 계수들(312)의 세트를 결정하도록 구성된 변환 유닛(302);

- 부동-소수점 인코딩 유닛(304)으로서,

- 상기 스펙트럼 계수들(312)의 세트에 기초하여, 스케일 팩터들의 세트 및 스케일링된 값들(314)의 세트를 결정하고;

- 상기 스케일 팩터들의 세트를 인코딩하여 인코딩된 스케일 팩터들(313)의 세트를 산출하도록 구성된, 상기 부동-소수점 인코딩 유닛(304);

- 비트 할당 및 양자화 유닛(305, 306)으로서,

- 상기 제 1 타겟 데이터-레이트에 기초하고 상기 인코딩된 스케일 팩터들(313)의 세트를 위해 이용된 비트들의 수에 기초하여, 상기 스케일링된 값들(314)의 세트를 양자화하기 위해 이용가능한 비트들의 총수를 결정하고;

- 상기 스케일링된 값들(314)의 세트의 상기 스케일링된 값들을 양자화하기 위해 상기 이용가능한 비트들의 총수의 할당을 표시하는 제 1 제어 파라미터(315)를 결정하고;

- 상기 제 1 제어 파라미터(315)에 따라 상기 스케일링된 값들(314)의 세트를 양자화하여 양자화된 스케일링된 값들(317)의 세트를 산출하도록 구성된, 상기 비트 할당 및 양자화 유닛(305, 306);

- 트랜스코더가 상기 제 1 비트스트림을 제 2 타겟 데이터-레이트의 제 2 비트스트림으로 변환할 수 있게 하기 위한 제 2 제어 파라미터(321)를 도출하도록 구성된 트랜스코딩 시뮬레이션 유닛(320)으로서, 상기 제 2 비트스트림은 상기 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따르고; 상기 트랜스코딩 시뮬레이션 유닛(320)은 상기 제 1 제어 파라미터(315)로부터 상기 제 2 제어 파라미터(321)를 도출하도록 구성되는, 상기 트랜스코딩 시뮬레이션 유닛(320); 및

- 상기 양자화된 스케일링된 값들(317)의 세트, 상기 인코딩된 스케일 팩터들(313)의 세트, 상기 제 1 제어 파라미터(315) 및 상기 제 2 제어 파라미터(321)를 포함하는 상기 제 1 비트스트림을 생성하도록 구성된 비트스트림 패킹 유닛(307)을 포함하고,

상기 트랜스코딩 시뮬레이션 유닛(320)은 상기 제 1 제어 파라미터(315) 단독으로부터 상기 제 2 제어 파라미터(321)를 도출하도록 구성되는, 오디오 인코더.

청구항 3

제 2 항에 있어서,

상기 트랜스코딩 시뮬레이션 유닛(320)은 상기 제 2 제어 파라미터(321)의 값을 상기 제 1 제어 파라미터(315)의 값과 동일하게 설정하도록 구성되는, 오디오 인코더.

청구항 4

제 2 항에 있어서,

상기 트랜스코딩 시뮬레이션 유닛(320)은 상기 제 2 오디오 코덱 시스템에 따라 비트 할당 처리를 수행하지 않고 상기 제 2 제어 파라미터(321)를 도출하도록 구성되는, 오디오 인코더.

청구항 5

제 2 항에 있어서,

- 상기 제 1 제어 파라미터(315)는 거친 구성요소(coarse component) 및 미세 구성요소(fine component)를 포함하고;
- 상기 트랜스코딩 시뮬레이션 유닛(320)은 상기 거친 및 미세 구성요소들을 조합하여 상기 제 2 제어 파라미터(321)를 산출하도록 구성되는, 오디오 인코더.

청구항 6

제 2 항에 있어서,

- 상기 제 1 비트스트림은 제 1 포맷에 따르고;
- 상기 제 2 비트스트림은 제 2 포맷에 따르고;
- 상기 트랜스코딩 시뮬레이션 유닛(320)은 상기 제 2 포맷에 의해 요구된 초과 비트들의 수를 결정하여 상기 양자화된 스케일링된 값들(317)의 세트 및 상기 인코딩된 스케일 팩터들(313)의 세트를 표현하도록 구성되고;
- 상기 비트 할당 및 양자화 유닛(305, 306)은 상기 초과 비트들의 수에도 또한 기초하여 상기 이용가능한 비트들의 총수를 결정하도록 구성되는, 오디오 인코더.

청구항 7

청구항 7은(는) 설정등록료 납부시 포기되었습니다.

제 6 항에 있어서,

상기 비트 할당 및 양자화 유닛(305, 306)은 상기 이용가능한 비트들의 총수를 상기 초과 비트들의 수만큼 감소시키도록 구성되는, 오디오 인코더.

청구항 8

청구항 8은(는) 설정등록료 납부시 포기되었습니다.

제 6 항에 있어서,

상기 초과 비트들의 수는,

- 상기 오디오 신호의 프레임에 대해 구체적으로 결정되거나;
- 미리-결정된 값인, 오디오 인코더.

청구항 9

청구항 9은(는) 설정등록료 납부시 포기되었습니다.

제 5 항에 있어서,

상기 제 1 타겟 데이터-레이트는 상기 제 2 타겟 데이터-레이트와 동일한, 오디오 인코더.

청구항 10

청구항 10은(는) 설정등록료 납부시 포기되었습니다.

제 2 항에 있어서,

상기 트랜스코딩 시뮬레이션 유닛(320)은,

- 상기 제 1 제어 파라미터에 기초하여 디폴트 제 2 제어 파라미터를 결정하고;

- 상기 디폴트 제 2 제어 파라미터에 기초하여 트랜스코딩되는 디폴트 제 2 비트스트림이 상기 제 2 타겟 데이터-레이트를 초과하는지의 여부를 결정하고;
- 상기 디폴트 제 2 비트스트림이 상기 제 2 타겟 데이터-레이트를 초과하지 않는 경우, 상기 디폴트 제 2 제어 파라미터에 기초하여 상기 제 2 제어 파라미터를 결정하도록 구성되는, 오디오 인코더.

청구항 11

청구항 11은(는) 설정등록료 납부시 포기되었습니다.

제 10 항에 있어서,

상기 트랜스코딩 시뮬레이션 유닛(320)은,

- 상기 제 1 제어 파라미터(315)를 이용하여 상기 양자화된 스케일링된 값들(317)의 세트를 역-양자화하여 역-양자화된 스케일링된 값들의 세트를 산출하고;
- 상기 디폴트 제 2 제어 파라미터(321)를 이용하여 상기 역-양자화된 스케일링된 값들의 세트를 재-양자화하여 재-양자화된 스케일링된 값들의 세트를 산출하도록 구성되는, 오디오 인코더.

청구항 12

청구항 12은(는) 설정등록료 납부시 포기되었습니다.

제 11 항에 있어서,

상기 디폴트 제 2 비트스트림이 상기 제 2 타겟 데이터-레이트를 초과한다고 결정되는 경우, 상기 트랜스코딩 시뮬레이션 유닛(320)은, 상기 제 2 제어 파라미터(321)에 기초하여 트랜스코딩되는 상기 제 2 비트스트림이 상기 제 2 타겟 데이터-레이트를 초과하지 않도록, 상기 제 2 오디오 코덱 시스템에 따라 비트 할당 및 양자화를 수행하여 상기 제 2 제어 파라미터를 결정하도록 구성되는, 오디오 인코더.

청구항 13

청구항 13은(는) 설정등록료 납부시 포기되었습니다.

제 12 항에 있어서,

상기 제 2 오디오 코덱 시스템에 따른 비트 할당 및 양자화는,

- 상기 제 2 타겟 데이터-레이트에 기초하고 상기 제 2 오디오 코덱 시스템에 따라 상기 인코딩된 스케일 팩터들(313)의 세트를 재-인코딩하는데 이용된 비트들의 수에 기초하여, 상기 역-양자화된 스케일링된 값들의 세트를 양자화하기 위해 이용가능한 비트들의 제 2 총수를 결정하는 것; 및
- 상기 역-양자화된 스케일링된 값들의 세트의 스케일링된 값들을 양자화하기 위해 상기 이용가능한 비트들의 제 2 총수의 할당을 표시하는 제 2 제어 파라미터(321)를 결정하는 것을 포함하는, 오디오 인코더.

청구항 14

청구항 14은(는) 설정등록료 납부시 포기되었습니다.

제 13 항에 있어서,

상기 제 2 오디오 코덱 시스템에 따른 비트 할당 및 양자화는,

- 상기 인코딩된 스케일 팩터들(313)의 세트에 기초하여 PSD 분포(410)로 칭해지는 전력 스펙트럼 밀도를 결정하는 것;
- 상기 인코딩된 스케일 팩터들(313)의 세트에 기초하여 마스킹 곡선(441)을 결정하는 것;
- 중간 제 2 제어 파라미터를 이용하여 상기 마스킹 곡선(441)을 오프셋함으로써 오프셋 마스킹 곡선(441)을 결정하는 것;
- 상기 PSD 분포(410)의 및 상기 오프셋 마스킹 곡선(441)의 비교에 기초하여, 상기 역-양자화된 스케일링된 값들의 세트의 역-양자화된 스케일링된 값들을 양자화하기 위해 요구된 비트들의 수를 결정하는 것; 및

- 상기 요구된 비트들의 수와 상기 이용가능한 비트들의 제 2 총수 사이의 차가 감소되고 상기 요구된 비트들의 수가 상기 이용가능한 비트들의 제 2 총수를 초과하지 않도록, 상기 중간 제 2 제어 파라미터를 반복 처리로 조정하고, 그에 의해 상기 제 2 제어 파라미터(321)를 산출하는 것을 추가로 포함하는, 오디오 인코더.

청구항 15

청구항 15은(는) 설정등록료 납부시 포기되었습니다.

제 14 항에 있어서,

상기 트랜스코딩 시뮬레이션 유닛(320)은,

- 상기 중간 제 2 제어 파라미터를 상기 제 1 제어 파라미터로 초기화하고/하거나;
- 상기 PSD 분포(410)의 및 상기 오프셋 마스킹 곡선(441)의 상기 비교에 기초하여 결정된 양자화 잡음이 미리-결정된 잡음 임계값 아래에 있는 경우 상기 반복 절차를 중단하도록 구성되는, 오디오 인코더(300).

청구항 16

청구항 16은(는) 설정등록료 납부시 포기되었습니다.

제 11 항에 있어서,

상기 디폴트 제 2 비트스트림이 상기 제 2 타겟 데이터-레이트를 초과한다고 결정되는 경우, 상기 트랜스코딩 시뮬레이션 유닛(320)은 상기 디폴트 제 2 제어 파라미터를 미리-결정된 제어 파라미터 오프셋 값만큼 오프셋함으로써 상기 제 2 제어 파라미터(321)를 결정하도록 구성되는, 오디오 인코더.

청구항 17

청구항 17은(는) 설정등록료 납부시 포기되었습니다.

제 2 항에 있어서,

상기 변환 유닛(302)은 상기 오디오 신호의 프레임으로부터 도출된 하나 이상의 블록들에 대해 변형 이산 코사인 변환(Modified Discrete Cosine Transform)을 수행하도록 구성되는, 오디오 인코더.

청구항 18

청구항 18은(는) 설정등록료 납부시 포기되었습니다.

제 2 항에 있어서,

- 상기 스케일 팩터들은 지수들 e에 대응하고;
- 상기 스케일링된 값들은 가수들 m에 대응하고;
- 상기 부동-소수점 인코딩 유닛(304)은 공식 $X = m \cdot 2^e$ 를 이용하여 변환 계수 X에 대한 지수 e 및 가수 m을 결정하도록 구성되는, 오디오 인코더.

청구항 19

청구항 19은(는) 설정등록료 납부시 포기되었습니다.

제 2 항에 있어서,

상기 비트 할당 및 양자화 유닛(305, 306)은,

- 상기 인코딩된 스케일 팩터들(313)의 세트에 기초하여 PSD 분포(410)로 칭해진, 전력 스펙트럼 밀도를 결정하고;
- 상기 인코딩된 스케일 팩터들(313)의 세트에 기초하여 마스킹 곡선(441)을 결정하고;
- 중간 제 1 제어 파라미터를 이용하여 상기 마스킹 곡선(441)을 오프셋함으로써 오프셋 마스킹 곡선(441)을 결정하고;

- 상기 PSD 분포(410)의 및 상기 오프셋 마스크링 곡선(441)의 비교에 기초하여, 상기 스케일링된 값들(314)의 세트의 스케일링된 값들을 양자화하기 위해 요구된 비트들의 수를 결정하고;
- 상기 요구된 비트들의 수와 상기 이용가능한 비트들의 총수 사이의 차가 감소되고 상기 요구된 비트들의 수가 상기 이용가능한 비트들의 총수를 초과하지 않도록, 상기 중간 제 1 제어 파라미터를 조정하고, 그에 의해 상기 제 1 제어 파라미터를 산출함으로써, 상기 제 1 제어 파라미터(315)를 결정하도록 구성되는, 오디오 인코더.

청구항 20

청구항 20은(는) 설정등록료 납부시 포기되었습니다.

제 2 항에 있어서,

상기 비트스트림 패킹 유닛(307)은 상기 제 1 비트스트림이 상기 제 1 타겟 데이터-레이트에 따르도록 하나 이상의 채움 비트(fill bit)들을 상기 제 1 비트스트림에 삽입하도록 구성되는, 오디오 인코더.

청구항 21

청구항 21은(는) 설정등록료 납부시 포기되었습니다.

제 2 항에 있어서,

상기 오디오 신호는 다중 채널 오디오 신호인, 오디오 인코더.

청구항 22

청구항 22은(는) 설정등록료 납부시 포기되었습니다.

제 2 항에 있어서,

상기 프레임은 상기 오디오 신호의 미리 결정된 수의 샘플들을 포함하는, 오디오 인코더.

청구항 23

제 2 항에 있어서,

- 상기 제 1 오디오 코덱 시스템은 돌비 디지털 플러스 코덱 시스템에 따르고/따르거나;
- 상기 제 1 제어 파라미터는 돌비 디지털 플러스 SNR 오프셋 값을 포함하고/하거나;
- 상기 제 2 오디오 코덱 시스템은 돌비 디지털 코덱 시스템에 따르고/따르거나;
- 상기 제 2 제어 파라미터는 돌비 디지털 SNR 오프셋 값을 포함하는, 오디오 인코더.

청구항 24

제 2 항에 있어서,

- 상기 제 1 타겟 데이터-레이트는: 384kbps, 448kbps, 640kbps 중 하나이고/이거나;
- 상기 제 2 타겟 데이터-레이트는 640kbps인, 오디오 인코더.

청구항 25

오디오 트랜스코더에 있어서,

- 제 1 데이터-레이트의 제 1 비트스트림을 수신하는 것으로서;
 - 상기 제 1 비트스트림은 제 1 오디오 코덱 시스템에 따라 인코딩된 오디오 신호의 프레임을 표시하고;
 - 상기 제 1 비트스트림은 양자화된 스케일링된 값들(317)의 세트, 인코딩된 스케일 팩터들(313)의 세트, 제 1 제어 파라미터(315) 및 제 2 제어 파라미터(321)를 포함하고;
 - 상기 양자화된 스케일링된 값들(317)의 세트 및 상기 인코딩된 스케일 팩터들(313)의 세트는 상기 오

디오 신호의 프레임의 스펙트럼 구성요소들을 표시하고;

- 상기 제 1 제어 파라미터(315)는 상기 양자화된 스케일링된 값들(317)의 세트를 양자화하는데 이용된 양자화기의 해상도를 표시하고;

- 상기 제 2 제어 파라미터(321)는 제 2 타겟 데이터-레이트의 제 2 비트스트림에 대한 상기 양자화된 스케일링된 값들(317)의 세트를 재-양자화하기 위해 상기 트랜스코더에 의해 이용될 양자화기를 표시하고;

- 상기 제 2 비트스트림은 상기 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따르는, 상기 제 1 비트스트림을 수신하고;

- 상기 제 1 데이터-레이트가 상기 제 2 타겟 데이터-레이트와 동일한지의 여부를 결정하고;

- 상기 제 1 제어 파라미터가 상기 제 2 제어 파라미터에 대응하는지의 여부를 결정하고;

- 상기 제 1 데이터-레이트가 상기 제 2 타겟 데이터-레이트와 동일한 경우 및 상기 제 1 제어 파라미터가 상기 제 2 제어 파라미터에 대응하는 경우, 상기 양자화된 스케일링된 값들(317)의 세트, 상기 인코딩된 스케일 팩터들(313)의 세트, 및 상기 제 2 제어 파라미터(321)를 상기 제 2 비트스트림에 복사함으로써 상기 제 2 비트스트림을 결정하도록 구성되는, 오디오 트랜스코더.

청구항 26

제 25 항에 있어서,

상기 제 1 데이터-레이트가 상기 제 2 타겟 데이터-레이트보다 작은 경우 및 상기 제 1 제어 파라미터가 상기 제 2 제어 파라미터에 대응하는 경우,

- 상기 제 1 비트스트림이 결합 채널 및/또는 풀 채널을 포함하는지의 여부를 결정하고;

- 상기 양자화된 스케일링된 값들(317)의 세트의 양자화된 스케일링된 값들 및 상기 풀 채널과 연관되는 상기 인코딩된 스케일 팩터들(313)의 세트의 인코딩된 스케일 팩터들을 상기 제 2 비트스트림에 복사하도록 추가로 구성되는, 오디오 트랜스코더.

청구항 27

제 26 항에 있어서,

- 상기 양자화된 스케일링된 값들(317)의 세트의 양자화된 스케일링된 값들과 상기 결합 채널과 연관되는 상기 인코딩된 스케일 팩터들(313)의 세트의 인코딩된 스케일 팩터들을 분리하여, 그에 의해 양자화된 스케일링된 값들의 제 1 세트 및 인코딩된 스케일 팩터들의 제 1 세트를 산출하고;

- 상기 제 1 제어 파라미터를 이용하여 상기 양자화된 스케일링된 값들의 제 1 세트를 역-양자화하여 역-양자화된 스케일링된 값들의 제 1 세트를 산출하고;

- 상기 제 2 제어 파라미터를 이용하여 상기 역-양자화된 스케일링된 값들의 제 1 세트를 재-양자화하여, 그에 의해 재-양자화된 스케일링된 값들의 제 1 세트를 산출하고;

- 상기 재-양자화된 스케일링된 값들의 제 1 세트를 상기 제 2 비트스트림에 삽입하도록 추가로 구성되는, 오디오 트랜스코더.

청구항 28

삭제

청구항 29

제 1 오디오 코덱 시스템에 따라 인코딩된 오디오 신호를 표시하는 제 1 비트스트림을 상기 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따른 제 2 비트스트림으로 트랜스코딩하기 위한 방법에 있어서,

- 제 1 데이터-레이트의 상기 제 1 비트스트림을 수신하는 단계로서,

- 상기 제 1 비트스트림은 양자화된 스케일링된 값들(317)의 세트, 인코딩된 스케일 팩터들(313)의 세트, 제 1 제어 파라미터(315) 및 제 2 제어 파라미터(321)를 포함하고;

- 상기 양자화된 스케일링된 값들(317)의 세트 및 상기 인코딩된 스케일 팩터들(313)의 세트는 상기 오디오 신호의 스펙트럼 구성요소들을 표시하고;

- 상기 제 1 제어 파라미터(315)는 상기 양자화된 스케일링된 값들(317)의 세트를 양자화하는데 이용된 양자화기를 표시하고;

- 상기 제 2 제어 파라미터(321)는 제 2 타겟 데이터-레이트의 제 2 비트스트림에 대한 상기 양자화된 스케일링된 값들(317)의 세트를 재-양자화하기 위해 트랜스코더에 의해 이용될 양자화기를 표시하는, 상기 제 1 비트스트림을 수신하는 단계;

- 상기 제 1 데이터-레이트가 상기 제 2 타겟 데이터-레이트와 동일한지의 여부를 결정하는 단계;

- 상기 제 1 제어 파라미터가 상기 제 2 제어 파라미터에 대응하는지의 여부를 결정하는 단계; 및

- 상기 제 1 데이터-레이트가 상기 제 2 타겟 데이터-레이트와 동일한 경우 및 상기 제 1 제어 파라미터가 상기 제 2 제어 파라미터에 대응하는 경우, 상기 양자화된 스케일링된 값들(317)의 세트, 상기 인코딩된 스케일 팩터들(313)의 세트, 및 상기 제 2 제어 파라미터(321)를 상기 제 2 비트스트림에 복사함으로써 상기 제 2 비트스트림을 결정하는 단계를 포함하는, 제 1 비트스트림을 트랜스코딩하기 위한 방법.

청구항 30

돌비 디지털 플러스 코덱 시스템에 따라 오디오 신호를 인코딩하고, 그에 의해 제 1 타겟 데이터-레이트의 제 1 비트스트림을 산출하도록 구성된 오디오 인코더(300)에 있어서,

- 상기 돌비 디지털 플러스 코덱 시스템에 따라 상기 제 1 타겟 데이터-레이트에 대한 snroffset 파라미터(315)를 결정하고;

- 트랜스코더가 상기 제 1 비트스트림을 제 2 타겟 데이터-레이트의 제 2 비트스트림으로 변환할 수 있게 하기 위한 convsnroffset 파라미터(321)를 상기 snroffset 파라미터(315)로부터 도출하도록 구성되고; 상기 제 2 비트스트림은 돌비 디지털 코덱 시스템에 따르고; 상기 제 1 비트스트림은 상기 snroffset 파라미터(315) 및 상기 convsnroffset 파라미터(321)를 포함하는, 오디오 인코더.

청구항 31

제 1 포맷에 대응하는 제 1 비트스트림의 제 2 포맷에 대응하는 제 2 비트스트림으로의 변환을 가능하게 하는 방법으로서, 상기 제 1 및 제 2 비트스트림들은 인코딩된 오디오 신호의 적어도 하나 및 동일 프레임에 관련되고, 상기 제 1 비트스트림은 상기 제 1 비트스트림과 연관된 제 1 비트 할당 처리를 표시하는 제 1 제어 파라미터를 포함하고, 상기 제 1 제어 파라미터는 거친 구성요소 및 미세 구성요소를 포함하고, 상기 제 2 비트스트림은 상기 제 2 비트스트림과 연관된 제 2 비트 할당 처리를 표시하는 제 2 제어 파라미터를 포함하고, 상기 제 2 비트스트림은 상기 제 2 제어 파라미터를 이용하여 상기 제 1 비트스트림으로부터 생성되는, 상기 방법에 있어서,

- 상기 거친 및 미세 구성요소들의 조합에만 기초하여 상기 제 2 제어 파라미터를 결정하는 단계; 및

- 상기 제 2 제어 파라미터를 상기 제 1 비트스트림에 삽입하는 단계를 포함하는, 제 1 비트스트림의 제 2 비트스트림으로의 변환을 가능하게 하는 방법.

청구항 32

오디오 트랜스코더에 있어서,

- 제 1 데이터-레이트의 제 1 비트스트림을 수신하는 것으로서;

- 상기 제 1 비트스트림은 돌비 디지털 플러스 코덱 시스템에 따라 인코딩된 오디오 신호를 표시하고;

- 상기 제 1 비트스트림은 양자화된 스케일링된 값들(317)의 세트, snroffset 파라미터(315) 및 convsnroffset 파라미터(321)를 포함하고;

- 상기 convsnroffset 파라미터(321)는 제 2 타겟 데이터-레이트의 제 2 비트스트림을 생성하기 위해 상기 트랜스코더에 의해 이용될 양자화기를 표시하고;

- 상기 제 2 비트스트림은 돌비 디지털 오디오 코덱 시스템에 따르는, 상기 제 1 비트스트림을 수신하고;
- 상기 제 1 데이터-레이트가 상기 제 2 타겟 데이터-레이트와 동일한지의 여부를 결정하고;
- 상기 snroffset 파라미터가 상기 convsnroffset 파라미터에 대응하는지의 여부를 결정하고;
- 상기 제 1 데이터-레이트가 상기 제 2 타겟 데이터-레이트와 동일한 경우 및 상기 snroffset 파라미터가 상기 convsnroffset 파라미터에 대응하는 경우, 상기 양자화된 스케일링된 값들(317)의 세트 및 상기 convsnroffset 파라미터(321)를 상기 제 2 비트스트림에 복사함으로써 상기 제 2 비트스트림을 결정하도록 구성되는, 오디오 트랜스코더.

발명의 설명

기술 분야

[0001] 관련 출원들에 대한 교차-참조

[0002] 본 출원은, 2012년 11월 7일에 출원된 미국 가 특허 출원 제 61/723,687 호에 대한 우선권을 주장하며, 이것은 전체로서 본 명세서에 참조로 포함된다.

[0003] 본 발명은 오디오 인코딩/디코딩에 관한 것이다. 특히, 본 발명은 오디오 인코딩/디코딩의 컨텍스트에서 이용되는 비트 할당 처리의 복잡성을 감소시키기 위한 방법 및 시스템에 관한 것이다.

배경 기술

[0004] 다양한 단일-채널 및/또는 5.1, 7.1 또는 9.1 다중-채널 오디오 렌더링 시스템들과 같은 다중-채널 오디오 렌더링 시스템들이 현재 이용중이다. 오디오 렌더링 시스템들은 예를 들면, 5+1, 7+1 또는 9+1 스피커 위치들에서 각각 발신하는 서라운드 사운드의 생성을 허용한다. 대응하는 단일-채널 또는 다중-채널 오디오 신호들의 효율적인 송신을 위해 또는 효율적인 저장을 위해, 돌비 디지털(Dolby Digital; DD) 또는 돌비 디지털 플러스(DD+)와 같은 오디오 코덱(인코더/디코더) 시스템들이 이용되고 있다.

[0005] 특정 오디오 코덱 시스템(예로서, 돌비 디지털)을 이용하여 인코딩된 오디오 신호들을 디코딩하도록 구성된 오디오 렌더링 디바이스들의 중요한 설치 기반이 존재할 수 있다. 특정 오디오 코덱 시스템은 예를 들면, 제 2 오디오 코덱으로 칭해질 수 있다. 반면에, 오디오 코덱 시스템들의 진화는 업데이트된 오디오 코덱 시스템(예로서, 돌비 디지털 플러스)으로 이어질 수 있고, 이것은 예를 들면, 제 1 오디오 코덱 시스템으로 칭해질 수 있다. 업데이트된 오디오 코덱 시스템은 부가 특징들(예로서, 증가된 수의 채널들) 및/또는 개선된 코딩 품질을 제공할 수 있다. 이와 같이, 콘텐츠 제공자들은 업데이트된 오디오 코덱 시스템에 따라 그들의 콘텐츠를 제공하려는 경향이 있을 수 있다.

[0006] 그러나, 제 2 오디오 코덱 시스템의 디코더를 구비한 오디오 렌더링 디바이스를 가진 사용자는 제 1 오디오 코덱 시스템에 따라 인코딩된 오디오 콘텐츠를 여전히 렌더링할 수 있어야 한다. 이것은 제 1 오디오 코덱 시스템에 따라 인코딩된 오디오 콘텐츠를 제 2 오디오 코덱 시스템에 따라 인코딩되는 수정된 오디오 콘텐츠로 변환하도록 구성된 소위 트랜스코더 또는 변환기에 의해 달성될 수 있다. 이러한 트랜스코더들/변환기들(예로서, 셋톱박스들 내에서 구현되는)의 비용을 절감하기 위해서는, 변환의 계산 복잡성이 비교적 낮아야 한다. 이를 위해, 제 1 오디오 코덱 시스템에 따라 동작하는 인코더는 하나 이상의 제어 파라미터들을 인코딩된 오디오 콘텐츠를 포함하는 비트스트림에 삽입하도록 구성될 수 있다. 하나 이상의 제어 파라미터들은 감소된 계산 복잡성으로 변환을 수행하기 위해 트랜스코더에 의해 이용될 수 있다. 반면에, 하나 이상의 제어 파라미터들의 생성은 통상적으로 인코더의 계산 복잡성을 증가시킨다.

발명의 내용

해결하려는 과제

[0007] 본 문서에서는, 감소된 계산 복잡성으로 제 1 포맷(제 1 오디오 코덱 시스템에 따라)으로부터 제 2 포맷(제 2 오디오 코덱 시스템에 따라)으로의 오디오 콘텐츠의 변환을 가능하게 하는 방법들 및 시스템들이 기술된다. 본 문서에 기술된 방법들 및 시스템들은 인코더에서 및/또는 트랜스코더에서 계산 복잡성을 감소시키는데 이용될

수 있다.

과제의 해결 수단

- [0008] 일 양태에 따라 제 1 오디오 코덱 시스템에 따라 오디오 신호의 프레임을 인코딩하도록 구성된 오디오 인코더가 기술된다. 오디오 신호는 다중 채널 오디오 신호, 예를 들면, 5.1, 7.1 또는 9.1 다중-채널 오디오 신호를 포함할 수 있다. 오디오 신호는 프레임들의 시퀀스로 분할될 수 있고, 프레임들은 오디오 신호의 미리-결정된 수의 샘플들, 예를 들면, 1536개의 샘플들을 포함할 수 있다. 제 1 오디오 코덱 시스템은 돌비 디지털 플러스 코덱 시스템, 예를 들면, 저 복잡성 돌비 디지털 플러스 시스템을 포함할 수 있거나 이에 따를 수 있다. 오디오 인코더는 오디오 신호를 제 1 타겟 데이터-레이트의 제 1 비트스트림으로 인코딩하도록 구성될 수 있다. 제 1 타겟 데이터-레이트(또는 제 1 데이터-레이트)에 대한 예들은 384kbps, 448kbps 또는 640kbps(특히 5.1 다중-채널 오디오 신호의 경우에)이다. 특히 다른 타입들의 다중-채널 오디오 신호들에 대해 다른 제 1 타겟 데이터-레이트들이 가능함을 유념해야 한다.
- [0009] 오디오 인코더는 오디오 신호의 프레임에 기초하여 스펙트럼 계수들의 세트를 결정하도록 구성된 변환 유닛을 포함할 수 있다. 즉, 변환 유닛은 오디오 신호의 하나 이상의 스펙트럼 구성요소들을 결정하도록 구성될 수 있다. 변환 유닛은 오디오 신호의 프레임으로부터 복수의 블록들을 결정하도록 구성될 수 있다. 또한, 변환 유닛은 샘플들의 블록들을 시간-도메인에서 주파수-도메인으로 변환하도록 구성될 수 있다. 예로서, 변환 유닛은 오디오 신호의 프레임으로부터 도출된 하나 이상의 블록들에 대해 변형 이산 코사인 변환(Modified Discrete Cosine Transform; MDCT)을 수행하도록 구성될 수 있다.
- [0010] 인코더는 스펙트럼 계수들의 세트에 기초하여, 스케일 팩터들의 세트 및 스케일링된 값들의 세트를 결정하도록 구성된 부동-소수점 인코딩 유닛을 포함할 수 있다. 스케일 팩터들은 지수들 e에 대응할 수 있고 스케일링된 값들은 가수들 m에 대응할 수 있다. 부동-소수점 인코딩 유닛은 공식 $X = m \cdot 2^e$ 를 이용하여 변환 계수 X에 대한 지수 e 및 가수 m을 결정하도록 구성될 수 있다. 이렇게 함으로써, 스펙트럼 계수들의 세트로부터의 모든 스펙트럼 계수들에 대해, 스케일 팩터들의 세트 및 스케일링된 값들의 세트가 결정될 수 있다.
- [0011] 또한, 부동-소수점 인코딩 유닛은 스케일 팩터들의 세트를 인코딩하여 인코딩된 스케일 팩터들의 세트를 산출하도록 구성될 수 있다. 스케일 팩터들의 세트의 인코딩은 예를 들면, 오디오 신호의 프레임의 모든 블록들에 대한 스케일 팩터들에 기초할 수 있다. 인코딩은 인코딩된 스케일 팩터들이 스케일 팩터들의 값들과 상이한 값들을 표현하도록 스케일 팩터의 수정을 유발할 수 있다.
- [0012] 인코더는, 제 1 타겟 데이터-레이트에 기초하고 인코딩된 스케일 팩터들의 세트를 위해 이용된 비트들의 수에 기초하여, 스케일링된 값들의 세트를 양자화하기 위해 이용가능한 비트들의 총수를 결정하도록 구성된 비트 할당 및 양자화 유닛을 포함할 수 있다. 이를 위해, 제 1 타겟 데이터-레이트는 프레임 당 비트들의 총수로 변환될 수 있고 인코딩된 스케일 팩터들의 세트를 위해 이용된 비트들(뿐만 아니라, 다른 목적들을 위해 예약될 수 있거나 이용되었을 수 있는 비트들)의 수가 비트들의 총수로부터 차감될 수 있고, 그에 의해 스케일링된 값들의 세트를 양자화하기 위해 이용가능한 비트들의 총수를 산출한다.
- [0013] 비트 할당 및 양자화 유닛은 스케일링된 값들을 양자화하기 위한 양자화의 해상도를 결정하기 위해 반복 비트 할당 처리를 수행하도록 구성될 수 있다. 양자화의 해상도는 스케일링된 값들의 세트를 양자화하기 위해 이용가능한 비트들의 총수가 초과되지 않고 지각적인 양자화 잡음이 최소화(또는 감소)되도록 결정되어야 한다. 이 요건을 충족하는 양자화는 제 1 제어 파라미터를 이용하여 식별될 수 있다. 즉, 비트 할당 및 양자화 유닛은 스케일링된 값들의 세트의 스케일링된 값들을 양자화하기 위해 이용가능한 비트들의 총수의 할당을 표시하는, 즉 스케일링된 값들의 세트의 스케일링된 값들을 양자화하기 위한 양자화를 표시하는 제 1 제어 파라미터를 결정하도록 구성될 수 있다. 제 1 제어 파라미터는 예를 들면, 돌비 디지털 플러스 snroffset(또는 SNR 오프셋) 값일 수 있거나 이를 포함할 수 있다.
- [0014] 예로서, 비트 할당 및 양자화 유닛은 인코딩된 스케일 팩터들의 세트에 기초하여 변환 계수들의 세트의 전력 스펙트럼 밀도(PSD) 분포를 결정함으로써 제 1 제어 파라미터를 결정하도록 구성될 수 있다. 인코딩된 스케일 팩터들의 세트는 통상적으로 제 1 비트스트림에 삽입되고 따라서 대응하는 디코더(또는 트랜스코더)에 알려져 있다. 이와 같이, PSD 분포는 또한 대응하는 디코더(또는 트랜스코더)에서 결정될 수 있다. 또한, 비트 할당 및 양자화 유닛은 인코딩된 스케일 팩터들의 세트에 기초하여 마스킹 곡선을 결정하도록 구성될 수 있다. 따라서, 마스킹 곡선은 통상적으로 대응하는 디코더(또는 트랜스코더)에서도 또한 도출가능하다. 마스킹 곡선은 이웃하는 스펙트럼 구성요소들(즉, 인접한 주파수들에서의 스펙트럼 구성요소들) 또는 오디오 신호의 변환 계수들 사

이의 마스킹을 표시할 수 있다. 또한, 비트 할당 및 양자화 유닛은 중간 제 1 제어 파라미터를 이용하여 마스킹 곡선을 오프셋함으로써 오프셋 마스킹 곡선을 결정하도록 구성될 수 있다. 특히, 중간 제 1 제어 파라미터는 오프셋 마스킹 곡선을 상향/하향으로 이동시키는데 이용될 수 있고, 그에 의해 마스킹되는 더 적은/ 더 많은 스펙트럼 구성요소들을 산출하고, 즉 그에 의해 양자화되어야 하는 더 적은/ 더 많은 스펙트럼 구성요소들을 산출한다. 비트 할당 및 양자화 유닛은 또한 PSD 분포의 및 오프셋 마스킹 곡선의 비교에 기초하여, 스케일링된 값들의 세트의 스케일링된 값들을 양자화하기 위해 요구된 비트들의 수를 결정하도록 구성될 수 있다. 중간 제 1 제어 파라미터는 요구된 비트들의 수와 이용가능한 비트들의 총수 사이의 차가 감소(예로서, 최소화)되도록 조정(반복 방식으로)될 수 있고, 그에 의해 제 1 제어 파라미터를 차를 감소(예로서, 최소화)시키는 중간 제 1 제어 파라미터로서 산출한다. 통상적으로, 차는 요구된 비트들의 수가 이용가능한 비트들의 총수를 초과하지 않도록 되어야 한다.

[0015] 상기에 언급된 반복 비트 할당 처리의 결과로서, 스케일링된 값들의 세트를 양자화하기 위한 양자화기를 규정하는 제 1 제어 파라미터가 획득된다. 비트 할당 및 양자화 유닛은 제 1 제어 파라미터에 따라 스케일링된 값들의 세트를 양자화하여 양자화된 스케일링된 값들의 세트를 산출하도록 구성될 수 있다.

[0016] 인코더는, 트랜스코더가 제 1 비트스트림을 제 2 타겟 데이터-레이트의 제 2 비트스트림으로 변환할 수 있게 하기 위한 제 2 제어 파라미터를 도출하도록 구성된 트랜스코딩 시뮬레이션 유닛을 추가로 포함할 수 있다. 제 2 비트스트림은 통상적으로 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따른다. 예로서, 제 2 코덱 시스템은 돌비 디지털 코덱 시스템에 따를 수 있고 제 2 제어 파라미터는 돌비 디지털 SNR 오프셋 값에 대응할 수 있거나 이를 포함할 수 있다. 제 2 타겟 데이터-레이트는 예를 들면, 640kpbs(특히, 5.1 다중-채널 오디오 신호의 경우에)일 수 있다. 제 2 타겟 데이터-레이트는 제 1 타겟 데이터-레이트 이상일 수 있다. 특히 다른 타입들의 다중-채널 오디오 신호들에 대해 다른 제 2 타겟 데이터-레이트들이 가능함을 유념해야 한다.

[0017] 트랜스코딩 시뮬레이션 유닛은 제 1 제어 파라미터로부터 제 2 제어 파라미터를 도출하도록 구성될 수 있다. 특히, 트랜스코딩 시뮬레이션 유닛은 제 1 제어 파라미터 단독으로부터 제 2 제어 파라미터를 도출하도록 구성될 수 있다. 일 실시예에서, 트랜스코딩 시뮬레이션 유닛은 제 2 오디오 코덱 시스템에 따라 비트 할당 처리를 수행하지 않고 제 2 제어 파라미터를 도출하도록 구성된다. 특정 실시예에서, 트랜스코딩 시뮬레이션 유닛은 제 2 제어 파라미터의 값을 제 1 제어 파라미터의 값과 동일하게 설정하도록 구성될 수 있다. 이와 같이, 인코더는 감소된 계산 복잡성으로 제 2 제어 파라미터를 결정하도록 구성될 수 있다. 제 1 제어 파라미터는 거친 구성요소와 미세 구성요소를 포함할 수 있다. 예로서(DD / DD+ 오디오 코덱 시스템의 경우에, csnroffset 및 fsnroffset 파라미터). 트랜스코딩 시뮬레이션 유닛은 거친 및 미세 구성요소들을 조합하여 제 2 제어 파라미터(예로서, convsnroffset 파라미터)를 산출하도록 구성될 수 있다.

[0018] 또한, 인코더는 양자화된 스케일링된 값들의 세트, 인코딩된 스케일 팩터들의 세트, 제 1 제어 파라미터 및/또는 제 2 제어 파라미터를 포함하는 제 1 비트스트림을 생성하도록 구성된 비트스트림 패키징 유닛을 포함할 수 있다. 제 1 비트스트림은 대응하는 디코더에 제공될 수 있다. 대안적으로 또는 부가적으로, 제 1 비트스트림은 제 1 비트스트림을 제 2 비트스트림으로 변환하도록 구성된 트랜스코더에 제공될 수 있다. 비트스트림 패키징 유닛은 제 1 비트스트림이 제 1 타겟 데이터-레이트에 따르도록 하나 이상의 스킵 비트들(또한 폐기 비트(waste bit)들 또는 미사용 비트들 또는 채움 비트(fill bit)들로 칭해질 수 있음)을 제 1 비트스트림에 삽입하도록 구성될 수 있다.

[0019] 제 1 비트스트림은 제 1 포맷에 따를 수 있고 제 2 비트스트림은 제 2 포맷에 따를 수 있다. 트랜스코딩 시뮬레이션 유닛은 양자화된 스케일링된 값들의 세트 및 인코딩된 스케일 팩터들의 세트를 표현하기 위해 제 2 포맷에 의해 요구된 초과 비트들의 수를 결정하도록 구성될 수 있다. 즉, 트랜스코딩 시뮬레이션 유닛은 제 1 포맷에 따른 표현에 비해 제 2 포맷에 따라 오디오 신호를 표현하는데 요구되는 부가 비트들의 수로서 초과 비트들의 수를 결정하도록 구성될 수 있다. 초과 비트들의 수는 오디오 신호의 특히 프레임에 대해 결정될 수 있거나 초과 비트들의 수는 미리-결정된 값, 예를 들면, 최악의 경우의 값일 수 있다. 인코더의 비트 할당 및 양자화 유닛은 초과 비트들의 수에도 또한 기초하여 이용가능한 비트들의 총수를 결정하도록 구성될 수 있다. 특히, 비트 할당 및 양자화 유닛은 이용가능한 비트들의 총수를 초과 비트들의 수만큼 감소시키도록 구성될 수 있다. 이렇게 함으로써, 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과하지 않는 것(특히 제 1 타겟 데이터-레이트가 제 2 타겟 데이터-레이트에 대응하거나 동일한 경우에)이 보장될 수 있다.

[0020] 트랜스코딩 시뮬레이션 유닛은 제 1 제어 파라미터에 기초하여 디폴트 제 2 제어 파라미터를 예를 들면, 제 1 제어 파라미터에 대응하거나 동일한 디폴트 제 2 제어 파라미터를 결정하도록 구성될 수 있다. 또한, 트랜스코

딩 시뮬레이션 유닛은, 디폴트 제 2 제어 파라미터에 기초하여 트랜스코딩되는 디폴트 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과하는지의 여부를 결정하도록 구성될 수 있다. 즉, 트랜스코딩 시뮬레이션 유닛은 디폴트 제 2 제어 파라미터를 이용하여 제 1 비트스트림을 제 2 비트스트림으로 변환하는 트랜스코더를 시뮬레이션하도록 구성될 수 있다. 이를 위해, 트랜스코딩 시뮬레이션 유닛은 제 1 제어 파라미터를 이용하여 양자화된 스케일링된 값들의 세트를 역-양자화하여 역-양자화된 스케일링된 값들의 세트를 산출하고, 디폴트 제 2 제어 파라미터를 이용하여 역-양자화된 스케일링된 값들의 세트를 재-양자화하여 재-양자화된 스케일링된 값들의 세트를 산출하도록 구성될 수 있다.

[0021] 디폴트 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과하지 않는 경우, 트랜스코딩 시뮬레이션 유닛은 디폴트 제 2 제어 파라미터에 기초하여 제 2 제어 파라미터를 결정하도록 구성될 수 있다. 예로서, 제 2 제어 파라미터는 디폴트 제 2 제어 파라미터와 동일하게 설정될 수 있다. 이와 같이, 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과하지 않는 것이 - 제 2 오디오 코덱 시스템에 따라 명시적 및/또는 반복 비트 할당 처리를 수행할 필요없이 - 보장된다.

[0022] 반면에, 디폴트 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과한다고 결정되는 경우, 트랜스코딩 시뮬레이션 유닛은 제 2 제어 파라미터에 기초하여 트랜스코딩되는 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과하지 않도록 제 2 제어 파라미터를 결정하기 위해 제 2 오디오 코덱 시스템에 따라 비트 할당 및 양자화를 수행하도록 구성될 수 있다. 즉, 디폴트 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과한다고 결정되는 경우에만, 제 2 오디오 코덱 시스템에 따라 비트 할당 및 양자화 처리를 수행할 필요가 있을 수 있다.

[0023] 제 2 오디오 코덱 시스템에 따른 비트 할당 및 양자화 처리는 제 2 타겟 데이터-레이트에 기초하고 제 2 오디오 코덱 시스템에 따라 인코딩된 스케일 팩터들의 세트를 재-인코딩하는데 이용된 비트들의 수에 기초하여, 역-양자화된 스케일링된 값들의 세트를 양자화하기 위해 이용가능한 비트들의 제 2 총수를 결정하는 단계를 포함할 수 있다. 또한, 비트 할당 및 양자화 처리는 역-양자화된 스케일링된 값들의 세트의 스케일링된 값들을 양자화하기 위해 이용가능한 비트들의 제 2 총수의 할당을 표시하는 제 2 제어 파라미터를 결정하는 단계를 포함할 수 있다.

[0024] 제 2 제어 파라미터의 결정은 반복 비트 할당 처리와 함께 수행될 수 있다. 이 반복 비트 할당 처리는 인코딩된 스케일 팩터들의 세트에 기초하여(예로서, 제 2 오디오 코덱 시스템에 따라 인코딩되는 인코딩된 스케일 팩터들의 세트에 기초하여) 전력 스펙트럼 밀도(PSD) 분포를 결정하는 단계를 포함할 수 있다. 또한, 반복 비트 할당 처리는 인코딩된 스케일 팩터들의 세트에 기초하여 마스킹 곡선을 결정하는 단계를 포함할 수 있다. 오프셋 마스킹 곡선은 중간 제 2 제어 파라미터를 이용하여 마스킹 곡선을 오프셋함으로써 결정될 수 있다. 또한, 역-양자화된 스케일링된 값들의 세트의 역-양자화된 스케일링된 값들을 양자화하기 위해 요구된 비트들의 수는 PSD 분포의 및 오프셋 마스킹 곡선의 비교에 기초하여 결정될 수 있다. 중간 제 2 제어 파라미터는 요구된 비트들의 수와 이용가능한 비트들의 제 2 총수 사이의 차가 감소(예로서, 최소화)되도록 반복 처리에서 조정될 수 있고, 그에 의해 제 2 제어 파라미터를 산출한다. 즉, 트랜스코딩 시뮬레이션 유닛은 제 2 오디오 코덱 시스템에 따라 반복 비트 할당 처리를 수행하도록 구성될 수 있고, 이것은 제 1 오디오 코덱 시스템에 따른 비트 할당 처리와 유사하다(예로서, 동일하다).

[0025] 트랜스코딩 시뮬레이션 유닛은 중간 제 2 제어 파라미터를 제 1 제어 파라미터로 초기화하도록 구성될 수 있고, 그에 의해 제 2 타겟 데이터-레이트에 및/또는 양자화 잡음에 관한 요건들을 충족하는 제 2 제어 파라미터를 결정하는데 요구되는 반복 횟수를 잠재적으로 감소시킨다. 대안적으로 또는 부가적으로, 트랜스코딩 시뮬레이션 유닛은 PSD 분포의 및 오프셋 마스킹 곡선의 비교에 기초하여 결정된 양자화 잡음이 미리-결정된 잡음 임계값 아래에 있는 경우 반복 절차를 중단하도록 구성될 수 있고, 그에 의해 요구되는 반복들의 수를 잠재적으로 감소시킨다.

[0026] 대안적으로 또는 부가적으로, 디폴트 제 2 비트스트림이 제 2 타겟 데이터-레이트를 초과한다고 결정되는 경우, 트랜스코딩 시뮬레이션 유닛은 디폴트 제 2 제어 파라미터를 미리 결정된 제어 파라미터 오프셋 값만큼 오프셋함으로써 제 2 제어 파라미터를 결정하도록 구성될 수 있다. 미리-결정된 제어 파라미터 오프셋 값은 예를 들면, 제 1 오디오 코덱 시스템에 따라 수행되는 비트 할당 및 양자화 처리에 기초하여 결정될 수 있다. 비트 할당 및 양자화 유닛에 의해 수행되는 이러한 비트 할당 및 양자화 처리는 제 2 제어 파라미터가 얼마만큼 오프셋되어야 하는지에 관한 표시를 제공할 수 있어서, 제 2 비트스트림은 제 2 타겟 데이터-레이트를 충족한다(예로서, 제 2 타겟 데이터-레이트를 초과하지 않는다).

[0027] 또 다른 양태에 따라, 제 1 데이터-레이트(예로서, 제 1 타겟 데이터-레이트)에서의 제 1 비트스트림을 수신하

도록 구성된 오디오 트랜스코더(또한 오디오 변환기로 칭해짐)가 기술된다. 상기에 개요된 바와 같이, 제 1 비트스트림은 제 1 오디오 코덱 시스템에 따라 인코딩된 오디오 신호의 프레임을 표시할 수 있다. 제 1 비트스트림은 양자화된 스케일링된 값들의 세트, 인코딩된 스케일 팩터들의 세트, 제 1 제어 파라미터 및 제 2 제어 파라미터를 포함할 수 있다. 양자화된 스케일링된 값들의 세트 및 인코딩된 스케일 팩터들의 세트는 오디오 신호의 프레임의 스펙트럼 구성요소들을 표시할 수 있고, 제 1 제어 파라미터는 양자화된 스케일링된 값들의 세트를 양자화하는데 이용된 양자화기의 해상도를 표시할 수 있다. 제 2 제어 파라미터는 제 2 타겟 데이터-레이트의 제 2 비트스트림에 대한 양자화된 스케일링된 값들의 세트를 재-양자화하기 위해 트랜스코더에 의해 이용될 양자화기를 표시할 수 있고, 제 2 비트스트림은 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따른다.

[0028] 트랜스코더는 제 1 데이터-레이트가 제 2 타겟 데이터-레이트와 동일한지의 여부를 결정하고 제 1 제어 파라미터가 제 2 제어 파라미터에 대응하는지의 여부를 결정하도록 구성될 수 있다. 제 1 데이터-레이트가 제 2 타겟 데이터-레이트와 동일한 경우 및 제 1 제어 파라미터가 제 2 제어 파라미터에 대응하는 경우, 트랜스코더는 양자화된 스케일링된 값들의 세트, 인코딩된 스케일 팩터들의 세트, 및 제 2 제어 파라미터를 제 2 비트스트림에 복사함으로써 제 2 비트스트림을 결정하도록 구성될 수 있다. 이와 같이, 트랜스코더는 양자화된 스케일링된 값들의 세트를 역-양자화할 필요없이(제 1 제어 파라미터를 이용하여), 및 역-양자화된 스케일링된 값들을 재-양자화할 필요없이(제 2 제어 파라미터를 이용하여), 제 2 비트스트림을 생성하도록 구성될 수 있다. 결과적으로, 트랜스코더의 계산 복잡성이 감소될 수 있다.

[0029] 제 1 데이터-레이트가 제 2 타겟 데이터-레이트보다 작은 경우 및 제 1 제어 파라미터가 제 2 제어 파라미터에 대응하는 경우, 트랜스코더는 제 1 비트스트림이 결합 채널 및/또는 풀 채널을 포함하는지의 여부(예로서, 다중-채널 오디오 신호들의 경우에)를 결정하도록 구성될 수 있다. 트랜스코더는 양자화된 스케일링된 값들의 세트의 양자화된 스케일링된 값들 및 인코딩된 스케일 팩터들의 세트의 인코딩된 스케일 팩터들을 복사하도록 구성될 수 있으며, 이들은 제 2 비트스트림에 대한 풀 채널과 연관된다. 이와 같이, 풀 채널들에 대해, 트랜스코더는 양자화된 스케일링된 값들의 세트(풀 채널과 연관됨)를 역-양자화할 필요 없고, 역-양자화된 스케일링된 값들(풀 채널과 연관됨)을 재-양자화할 필요가 없고, 그에 의해 트랜스코더의 계산 복잡성을 감소시킨다.

[0030] 또한, 오디오 트랜스코더는 결합 채널과 연관되는 양자화된 스케일링된 값들의 세트의 양자화된 스케일링된 값들 및 인코딩된 스케일 팩터들의 세트의 인코딩된 스케일 팩터들을 분리(de-couple)하고, 그에 의해 양자화된 스케일링된 값들의 제 1 세트 및 인코딩된 스케일 팩터들의 제 1 세트를 산출하도록 구성될 수 있다. 또한, 트랜스코더는 제 1 제어 파라미터를 이용하여 양자화된 스케일링된 값들의 제 1 세트를 역-양자화하여 역-양자화된 스케일링된 값들의 제 1 세트를 산출하고, 제 2 제어 파라미터를 이용하여 역-양자화된 스케일링된 값들의 제 1 세트를 재-양자화하고, 그에 의해 재-양자화된 스케일링된 값들의 제 1 세트를 산출하도록 구성될 수 있다. 재-양자화된 스케일링된 값들의 제 1 세트는 제 2 비트스트림에 삽입될 수 있다. 이와 같이, 제 2 오디오 코덱 시스템의 디코더에는 결합 채널들을 포함하지 않는, 즉 풀 채널들만 포함하는 제 2 비트스트림이 제공될 수 있다.

[0031] 또 다른 양태에 따라, 제 1 오디오 코덱 시스템에 따라 오디오 신호를 제 1 비트스트림으로 인코딩하기 위한 방법(및 대응하는 인코더)이 기술된다. 이 방법은 오디오 신호의 스펙트럼 구성요소들에 기초하여(예로서, 변환 계수들의 세트에 기초하여), 스케일 팩터들의 세트 및 스케일링된 값들의 세트를 결정하는 단계를 포함한다. 이 방법은 제 1 오디오 코덱 시스템에 따라 반복 비트 할당 처리를 이용하여 스케일링된 값들의 세트를 양자화하기 위한 양자화기의 해상도를 표시하는 제 1 제어 파라미터를 결정하는 단계를 진행한다. 양자화기의 해상도는 제 1 비트스트림의 제 1 타겟 데이터-레이트에 의존할 수 있다. 또한, 이 방법은 제 1 비트스트림의 제 2 타겟 데이터-레이트의 제 2 비트스트림으로의 변환을 가능하게 하기 위한 제 2 제어 파라미터를 결정하는 단계를 포함할 수 있다. 상기에 개요된 바와 같이, 제 2 비트스트림은 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따를 수 있다. 제 2 제어 파라미터 결정 단계는 제 1 제어 파라미터에 기초하여, 예를 들면, 제 2 오디오 코덱 시스템에 따라 반복 비트 할당 처리를 수행하지 않고 제 2 제어 파라미터를 결정하는 단계를 포함할 수 있다. 상기에 개요된 바와 같이, 제 1 제어 파라미터에 기초한 제 2 제어 파라미터의 결정은 하나 이상의 조건들(예로서, 제 2 타겟 데이터-레이트를 충족하는 제 2 비트스트림에 대해)에 의거될 수 있다. 제 1 비트스트림은 제 1 및 제 2 제어 파라미터들을 표시할 수 있다.

[0032] 또 다른 양태에 따라, 제 1 오디오 코덱 시스템에 따라 인코딩된 오디오 신호를 표시하는 제 1 비트스트림을 제 1 오디오 코덱 시스템과 상이한 제 2 오디오 코덱 시스템에 따른 제 2 비트스트림으로 트랜스코딩하기 위한 방법(및 대응하는 트랜스코더)이 기술된다. 이 방법은 제 1 데이터-레이트의 제 1 비트스트림을 수신하는 단계를

포함한다. 제 1 비트스트림은 양자화된 스케일링된 값들의 세트, 인코딩된 스케일 팩터들의 세트, 제 1 제어 파라미터 및 제 2 제어 파라미터를 포함할 수 있다. 양자화된 스케일링된 값들의 세트 및 인코딩된 스케일 팩터들의 세트는 오디오 신호의 스펙트럼 구성요소들을 표시할 수 있고, 제 1 제어 파라미터는 양자화된 스케일링된 값들의 세트를 양자화하는데 이용된 양자화기를 표시할 수 있다. 제 2 제어 파라미터는 제 2 타겟 데이터-레이트의 제 2 비트스트림에 대한 양자화된 스케일링된 값들의 세트를 제-양자화하기 위해 트랜스코더에 의해 이용될 양자화기를 표시할 수 있다. 이 방법은 제 1 데이터-레이트가 제 2 타겟 데이터-레이트와 동일한지의 여부를 결정하는 단계, 및 제 1 제어 파라미터가 제 2 제어 파라미터에 대응하는지의 여부를 결정하는 단계를 추가로 포함할 수 있다. 제 1 데이터-레이트가 제 2 타겟 데이터-레이트와 동일한 경우 및 제 1 제어 파라미터가 제 2 제어 파라미터에 대응하는 경우(예로서, 제 2 제어 파라미터와 값이 동일한 경우), 이 방법은 양자화된 스케일링된 값들의 세트, 인코딩된 스케일 팩터들의 세트, 및 제 2 제어 파라미터를 제 2 비트스트림에 복사함으로써 제 2 비트스트림을 결정하는 단계로 진행할 수 있다.

[0033] 또 다른 양태에 따라, 돌비 디지털 플러스 코덱 시스템에 따라 오디오 신호를 인코딩하도록 구성되고, 그에 의해 제 1 타겟 데이터-레이트의 제 1 비트스트림을 산출하는 오디오 인코더(및 대응하는 방법)가 기술된다. 오디오 인코더는 돌비 디지털 플러스 코덱 시스템에 따라 제 1 타겟 데이터-레이트에 대한 snroffset 파라미터를 결정하도록 구성될 수 있다. 또한, 인코더는 트랜스코더가 제 1 비트스트림을 제 2 타겟 데이터-레이트의 제 2 비트스트림으로 변환할 수 있게 하기 위한 convsnroffset 파라미터를 snroffset 파라미터로부터 도출하도록 구성될 수 있다. 제 2 비트스트림은 돌비 디지털 코덱 시스템에 따를 수 있고, 제 1 비트스트림은 snroffset 파라미터 및 convsnroffset 파라미터를 포함할 수 있다.

[0034] 또 다른 양태에 따라, 제 1 포맷에 대응하는 제 1 비트스트림의 제 2 포맷에 대응하는 제 2 비트스트림으로의 변환을 가능하게 하는 방법이 기술된다. 또한, 변환을 가능하게 하는 방법을 수행하도록 구성된 대응하는 장치(특히, 대응하는 오디오 인코더)가 기술된다. 제 1 비트스트림의 제 2 비트스트림으로의 실제 변환은 상이한 엔티티에 의해(예로서, 트랜스코더에 의해) 수행될 수 있다.

[0035] 제 1 및 제 2 포맷들은 본 문서에서 기술된 제 1 및 제 2 오디오 코덱 시스템들의 포맷들에 대응할 수 있다. 제 1 및 제 2 비트스트림들은 통상적으로 인코딩된 오디오 신호의 적어도 하나 및 동일 프레임에 관련된다. 즉, 제 1 및 제 2 비트스트림들은 통상적으로 오디오 신호의 대응하는 하나 이상의 프레임들을 기술한다. 제 1 비트스트림은 제 1 비트스트림과 연관된 제 1 비트 할당 처리를 표시하는 제 1 제어 파라미터를 포함한다. 제 1 비트 할당 처리는 제 1 오디오 코덱 시스템에 따라 수행될 수 있다. 본 문서에 개요된 바와 같이, 제 1 제어 파라미터는 거친 구성요소 및 미세 구성요소를 포함할 수 있다.

[0036] 제 2 비트스트림은 제 2 비트스트림과 연관된 제 2 비트 할당 처리를 표시하는 제 2 제어 파라미터를 포함할 수 있다. 제 2 비트 할당 처리는 제 2 오디오 코덱 시스템에 따라 수행될 수 있다. 또한, 제 2 비트스트림은 제 2 제어 파라미터를 이용하여 제 1 비트스트림으로부터 생성될 수 있다. 특히, 제 2 제어 파라미터는 제 1 비트스트림을 제 2 비트스트림으로 변환하기 위해 트랜스코더(인코더에 대해 원격일 수 있음)에 의해 이용될 수 있다.

[0037] 이 방법은 제 1 제어 파라미터에만 기초하여 제 2 제어 파라미터를 결정하는 단계를 포함할 수 있다. 특히, 제 2 제어 파라미터는 제 1 제어 파라미터의 거친 및 미세 구성요소들의 조합에만 기초하여 결정될 수 있다. 또한, 이 방법은 제 2 제어 파라미터를 제 1 비트스트림에 삽입하는 단계를 포함할 수 있다. 이와 같이, 제 1 비트스트림(제 1 및 제 2 제어 파라미터들을 포함하는)은 트랜스코더로 송신될 수 있고, 그에 의해 트랜스코더가 감소된 계산 복잡성으로(및 제 2 비트스트림을 송신할 필요없이) 제 1 비트스트림으로부터 제 2 비트스트림을 결정할 수 있게 한다.

[0038] 또 다른 양태에 따라, 오디오 트랜스코더(및 대응하는 트랜스코딩 방법)가 기술된다. 오디오 트랜스코더는 제 1 데이터-레이트의 제 1 비트스트림을 수신하도록 구성된다. 제 1 비트스트림은 돌비 디지털 플러스 코덱 시스템에 따라 인코딩된 오디오 신호를 표시할 수 있다. 제 1 비트스트림은 양자화된 스케일링된 값들의 세트, snroffset 파라미터 및 convsnroffset 파라미터를 포함할 수 있다. convsnroffset 파라미터는 제 2 타겟 데이터-레이트의 제 2 비트스트림을 생성하기 위해 트랜스코더에 의해 이용될 양자화기를 표시할 수 있고, 제 2 비트스트림은 돌비 디지털 오디오 코덱 시스템에 따른다. 트랜스코더는 제 1 데이터-레이트가 제 2 타겟 데이터-레이트와 동일한지의 여부를 결정하고 snroffset 파라미터가 convsnroffset 파라미터에 대응하는지의 여부를 결정하도록 구성될 수 있다. 제 1 데이터-레이트가 제 2 타겟 데이터-레이트와 동일한 경우 및 snroffset 파라미터가 convsnroffset 파라미터에 대응하는 경우, 트랜스코더는 양자화된 스케일링된 값들의 세트 및 convsnroffset 파라미터를 제 2 비트스트림에 복사함으로써 제 2 비트스트림을 결정하도록 구성될 수 있다.

- [0039] 또 다른 양태에 따라, 소프트웨어 프로그램이 기술된다. 소프트웨어 프로그램은 처리기 상에서의 실행을 위해 적응될 수 있고 처리기 상에서 실행될 때 본 문서에 개요된 방법 단계들을 수행하기 위해 적응될 수 있다.
- [0040] 또 다른 양태에 따라, 저장 매체가 기술된다. 저장 매체는 처리기 상에서의 실행을 위해 적응되고 처리기 상에서 실행될 때 본 문서에 개요된 방법 단계들을 수행하기 위해 적응되는 소프트웨어 프로그램을 포함할 수 있다.
- [0041] 또 다른 양태에 따라, 컴퓨터 프로그램 제품이 기술된다. 컴퓨터 프로그램은 컴퓨터 상에서 실행될 때 본 문서에 개요된 방법 단계들을 수행하기 위한 실행가능한 지시들을 포함할 수 있다.
- [0042] 본 특허 출원에 개요된 양호한 실시예들을 포함하는 방법들 및 시스템들은 독립형으로 이용될 수 있거나 이 문서에 개시된 다른 방법들 및 시스템들과 조합하여 이용될 수 있음을 유념해야 한다. 또한, 본 특허 출원에 개요된 방법들 및 시스템들의 모든 양태들은 임의로 조합될 수 있다. 특히, 청구항들의 특징들은 임의 방식으로 서로 조합될 수 있다.
- [0043] 본 발명은 첨부 도면들을 참조하여 일 예시적인 방식으로 하기에 설명된다.

도면의 간단한 설명

- [0044] 도 1a는 일 예시적인 다중-채널 오디오 인코더의 고 레벨 블록도.
- 도 1b는 인코딩된 프레임들의 일 예시적인 시퀀스를 도시한 도면.
- 도 2a는 예시적인 다중-채널 오디오 디코더들의 고 레벨 블록도.
- 도 2b는 7.1 다중-채널 오디오 신호에 대한 일 예시적인 스피커 배열을 도시한 도면.
- 도 3은 다중-채널 오디오 인코더의 예시적인 구성요소들의 블록도.
- 도 4a 내지 도 4e는 일 예시적인 다중-채널 오디오 인코더의 특정 양태들을 도시한 도면들.
- 도 5는 복수의 예시적인 프레임들에 대한 DD+ 비트스트림 포맷에 및 DD 비트스트림 포맷에 이용되는 고정된 비트들의 수를 도시한 도면.
- 도 6은 청취 테스트들의 예시적인 실험 결과들을 도시한 도면.

발명을 실시하기 위한 구체적인 내용

- [0045] 특정 다중-채널 오디오 디코더에 의해 디코딩되는 채널들의 수에 관해 하위 호환가능한 비트스트림들을 생성하는 다중-채널 오디오 코덱 시스템들을 제공하는 것이 바람직하다. 특히, N.1 다중-채널 오디오 디코더에 의해 디코딩될 수 있도록 M.1 다중 채널 오디오 신호를 인코딩하는 것이 바람직하며, $N < M$ 이다. 예로서, 5.1 오디오 디코더에 의해 디코딩될 수 있도록 7.1 오디오 신호를 인코딩하는 것이 바람직하다. 하위 호환성을 허용하기 위해, 다중-채널 오디오 코덱 시스템들은 통상적으로 M.1 다중-채널 오디오 신호를 감소된 수의 채널들(예로서, N.1 채널들)을 포함하는 독립형 (서브)스트림("IS")으로, 및 전체 M.1 오디오 신호를 디코딩 및 렌더링하기 위해 대체 및/또는 확장 채널들을 포함하는 하나 이상의 종속형 (서브)스트림들("DS")로 인코딩한다.
- [0046] 또한, 이전 버전의 오디오 디코더가 업데이트된 버전의 오디오 인코더에 의해 생성된 비트스트림을 디코딩할 수 있게 하는 비트스트림을 제공하는 것이 바람직하다. 즉, 비트스트림의 디코딩에 관해 하위 호환성(동일한 수 N.1의 채널들을 표현하는 비트스트림들도)을 허용하는 것이 바람직하다. 이것은 업데이트된 버전의 오디오 인코더를 이용하여 인코딩된 비트스트림을 이전 버전의 오디오 디코더에 의해 디코딩될 수 있는 비트스트림으로 변환하는 소위 트랜스코더 또는 변환기의 이용으로 달성될 수 있다. 이러한 트랜스코더는, 예를 들면, 비트스트림(업데이트된 버전의 오디오 인코더를 이용하여 인코딩된)을 수신하도록 구성되고 이전 버전의 오디오 디코더에 의해 디코딩될 수 있는 수정된 비트스트림을 제공하도록 구성될 수 있는 셋톱 박스에서 제공된다. 예로서, 트랜스코더는 돌비 디지털 플러스(DD+) 비트스트림을 수신하고 수신된 비트스트림을 돌비 디지털 오디오 디코더에 의해 디코딩될 수 있는 돌비 디지털(DD) 비트스트림으로 트랜스코딩하도록 구성될 수 있다. 이와 같이, 오디오 디코더들의(예로서, 텔레비전 세트들 내의 돌비 디지털 오디오 디코더들의) 설치 기반이 보호될 수 있는 동시에, 개선된 오디오 인코딩/디코딩 시스템들(돌비 디지털 플러스 코덱 시스템과 같은)에 대한 진화를 차단하지 않는다.
- [0047] 이러한 컨텍스트에서, 비트스트림의 인코딩에 연계된 및/또는 비트스트림의 트랜스코딩에 연계된 계산 복잡성을 감소시키는 것이 바람직하다. 본 문서에서, 감소된 계산 복잡성으로 비트스트림의 생성을 가능하게 하는 방법들

및 시스템들이 기술된다. 이 방법들 및 시스템들은 돌비 디지털 플러스(DD+) 코덱 시스템(또는 인헨스드 AC-3으로 칭해짐)에 기초하여 기술된다. DD+ 코덱 시스템은 차세대 텔레비전 시스템 위원회(Advanced Television Systems Committee; ATSC)의 2010년 11월 22일자의 문서 A/52 : 2010, "디지털 오디오 압축 표준(Digital Audio Compression Standard)(AC-3, E-AC-3)"에 명시되며, 그 콘텐츠는 참조로 포함된다. 그러나, 본 문서에 기술된 방법들 및 시스템들은 일반적으로 적용가능하고, 오디오 신호들을 인코딩하고 비트스트림을 트랜스코더에 제공하는 다른 오디오 코덱 시스템에 적용될 수 있어서, 비트스트림은 비트스트림의 저 복잡성 트랜스코딩을 가능하게 함을 유념해야 한다.

[0048] 빈번히 이용되는 다중-채널 구성들(다중-채널 오디오 신호들)은 7.1 구성 및 5.1 구성이다. 5.1 다중-채널 구성은 통상적으로 L(좌측 프론트), C(중앙 프론트), R(우측 프론트), Ls(좌측 서라운드), Rs(우측 서라운드), 및 LFE(저주파 효과들) 채널을 포함한다. 7.1 다중-채널 구성은 Lb(좌측 서라운드 백) 및 Rb(우측 서라운드 백) 채널을 추가로 포함한다. 일 예시적인 7.1 다중-채널 구성은 도 2b에 도시된다. DD+에서 7.1 채널들을 송신하기 위해서는, 2개의 서브스트림들이 이용된다. 제 1 서브스트림(독립형 서브스트림, "IS"으로 칭해짐)은 5.1 채널 믹스를 포함하고, 제 2 서브스트림(종속형 서브스트림, "DS"으로 칭해짐)은 확장 채널들 및 대체 채널들을 포함한다. 예를 들면, 서라운드 백 채널들 Lb 및 Rb를 이용하여 7.1 다중-채널 오디오 신호를 인코딩하여 송신하기 위해, 독립형 서브스트림은 채널들 L(좌측 프론트), C(중앙 정면), R(우측 프론트), Lst(좌측 서라운드 다운믹스), Rst(우측 서라운드 다운믹스), LFE(저주파 효과들)를 운반하고, 및 종속형 채널은 확장형 채널들 Lb(좌측 서라운드 백), Rb(우측 서라운드 백) 및 대체 채널들 Ls(좌측 서라운드), Rs(우측 서라운드)를 운반한다. 풀 7.1 신호 디코드가 수행될 때, 종속형 서브스트림으로부터의 Ls 및 Rs 채널들은 독립형 서브스트림으로부터의 Lst 및 Rst 채널들을 대체한다.

[0049] 도 1a는 5.1 및 7.1 채널들 사이의 관계를 도시하는 일 예시적인 DD+ 7.1 다중-채널 오디오 인코더(100)의 고레벨 블록도를 도시한다. 다중-채널 오디오 신호의 칠(7) 플러스 일(1) 오디오 채널들(101)(L, C, R, Ls, Lb, Rs 및 Rb 플러스 LFE)은 두 그룹들의 오디오 채널들로 나누어진다. 기본 그룹(121)의 채널들은 오디오 채널들 L, C, R 및 LFE 뿐만 아니라, 다운믹스드 서라운드 채널들 Lst(102) 및 Rst(103)를 포함하며, 이들은 통상적으로 7.1 서라운드 채널들 Ls, Rs 및 7.1 백 채널들 Lb, Rb로부터 도출된다. 예로서, 다운믹스드 서라운드 채널들(102, 103)은 다운믹스 유닛(109)에서 Lb 및 Rb 채널들의 일부 또는 전부와 7.1 서라운드 채널들 Ls, Rs를 더하여 도출된다. 다운믹스드 서라운드 채널들 Lst(102) 및 Rst(103)는 다른 방식들로 결정될 수 있음을 유념해야 한다. 예로서, 다운믹스드 서라운드 채널들 Lst(102) 및 Rst(103)는 2개의 7.1 채널들, 예를 들면, 7.1 서라운드 채널들 Ls, Rs로부터 직접 결정될 수 있다.

[0050] 기본 그룹(121)의 채널들은 DD+ 5.1 오디오 인코더(105)에서 인코딩되고, 그에 의해 DD+ 코어 프레임(151)으로 송신되는 독립형 서브스트림("IS")(110)을 산출한다(도 1b 참조). 코어 프레임(151)은 또한 IS 프레임으로 칭해진다. 제 2 그룹(122)의 오디오 채널들은 7.1 서라운드 채널들 Ls, Rs 및 7.1 서라운드 백 채널들 Lb, Rb를 포함한다. 제 2 그룹(122)의 채널들은 DD+ 4.0 오디오 인코더(106)에서 인코딩되고, 하나 이상의 DD+ 확장 프레임(152, 153)으로 송신되는 종속형 서브스트림("DS")(120)을 산출한다(도 1b 참조). 제 2 그룹(122)의 채널들은 본 명세서에서 확장 그룹(122)의 채널들로 칭해지고 확장 프레임들(152, 153)은 DS 프레임들(152, 153)로 칭해진다.

[0051] 도 1b는 인코딩된 오디오 프레임들(151, 152, 153, 161, 162)의 일 예시적인 시퀀스(150)를 도시한다. 도시된 예는 IS 프레임들(151 및 161)을 각각 포함하는 2개의 독립형 서브스트림들 IS0 및 IS1을 포함한다. 다수의 IS(및 각각의 DS)는 다수의 연관된 오디오 신호들(예로서, 상이한 언어들의 영화에 대해 또는 상이한 프로그램들에 대해)을 제공하는데 이용될 수 있다. 독립형 서브스트림들 각각은 하나 이상의 종속형 서브스트림들 DS0, DS1을 각각 포함한다. 종속형 서브스트림들 각각은 각각의 DS 프레임들(152, 153 및 162)을 포함한다. 또한, 도 1b는 다중-채널 오디오 신호의 완전한 오디오 프레임의 시간 길이(170)를 표시한다. 오디오 프레임의 시간 길이(170)는 32ms(예로서, 샘플링 레이트 fs=48kHz에서)일 수 있다. 즉, 도 1b는 하나 이상의 IS 프레임들(151, 161) 및 각각의 DS 프레임들(152, 153, 162)로 인코딩되는 오디오 프레임의 시간의 길이(170)를 표시한다.

[0052] 인코더(100)는 서브스트림들의 상이한 코딩 포맷으로의 효율적인 트랜스코딩을 허용하는 서브스트림들에 데이터를 포함하도록 구성될 수 있다. 예로서, 서브스트림들은 DD+ 독립형 서브스트림 IS0을 DD 비트스트림으로 트랜스코딩하도록 허용하는 데이터를 포함할 수 있다. 좀더 일반적인 관점에서, 인코더(100)는 제 1 오디오 코덱(예로서, DD+)에 호환가능한 제 1 비트스트림을 생성하도록 구성될 수 있다. 제 1 비트스트림은 트랜스코더가 감소된 복잡성으로 제 2 오디오 코덱(예로서, DD)과 호환가능한 제 2 비트스트림을 생성하도록 허용하는 데이터를 포함할 수 있다. 이를 위해, 인코더(100)는 제 2 오디오 코덱(예로서, DD)에 따라 오디오 채널들(101)의 일부

또는 전부를 인코딩하고 하나 이상의 제어 파라미터들을 결정하도록 구성될 수 있고, 이들 제어 파라미터들은 트랜스코더가 제 1 비트스트림으로부터 제 2 비트스트림을 효율적인 방식으로 생성할 수 있게 한다. 대역폭 효율성을 고려하여, 제 1 비트스트림은 제 2 오디오 코덱에 따라 인코딩되는 오디오 데이터가 아닌, 제 1 오디오 코덱에 따라 인코딩되는 오디오 데이터만을 포함해야 함을 유념해야 한다. 즉, 하나 이상의 제어 파라미터들은 오디오 데이터의 트랜스코딩에만 관한 것이어야 한다.

[0053] 도 2a는 예시적인 다중-채널 디코더 시스템들(200, 210)의 고 레벨 블록도들을 도시한다. 특히, 도 2a는, 인코딩된 기본 그룹(121)의 채널들을 포함하는 인코딩된 IS(201)를 수신하는 일 예시적인 5.1 다중-채널 디코더 시스템(200)을 도시한다. 인코딩된 IS(201)은 수신된 비트스트림의 IS 프레임들(151)로부터 취해진다(예로서, 도시되지 않은 디멀티플렉서를 이용하여). IS 프레임들(151)은 인코딩된 기본 그룹(121)의 채널들을 포함하고 5.1 다중-채널 디코더(205)를 이용하여 디코딩되고, 그에 의해 디코딩된 기본 그룹(221)의 채널을 포함하는 디코딩된 5.1 다중-채널 오디오 신호를 산출한다. 또한, 도 2a는, 인코딩된 기본 그룹(121)의 채널들을 포함하는 인코딩된 IS(121) 및 인코딩된 확장 그룹(122)의 채널들을 포함하는 인코딩된 DS(202)를 수신하는 일 예시적인 7.1 다중-채널 디코더 시스템(210)을 도시한다. 상기에 개요된 바와 같이, 인코딩된 IS(201)은 IS 프레임들(151)로부터 취해질 수 있고 인코딩된 DS(202)는 수신된 비트스트림의 DS 프레임들(152, 153)로부터 취해질 수 있다(예로서, 도시되지 않은 디멀티플렉서를 이용하여). 디코딩 후에, 디코딩된 기본 그룹(221)의 채널들 및 디코딩된 확장 그룹(222)의 채널들을 포함하는 디코딩된 7.1 다중-채널 오디오 신호가 획득된다. 7.1 다중-채널 디코더(215)가 디코딩된 확장 그룹(222)의 채널들을 대신 이용할 때, 다운믹스드 서라운드 채널들 Lst, Rst(211)는 드롭될 수 있음을 유념해야 한다. 7.1 다중-채널 오디오 신호의 통상적인 렌더링 위치들(232)은 도 2b의 다중-채널 구성(230)에 도시되고, 다중-채널 구성(230)은 또한 청취자의 일 예시적인 위치(231) 및 비디오 렌더링을 위한 스크린의 일 예시적인 위치(233)를 도시한다.

[0054] 현재, DD+에서의 7.1 채널 오디오 신호들의 인코딩은 제 1 코어 5.1 채널 DD+인코더(105) 및 제 2 DD+ 인코더(106)에 의해 수행된다. 제 1 DD+ 인코더(105)는 기본 그룹(121)의 5.1 채널들을 인코딩하고(따라서 5.1 채널 인코더로 칭해질 수 있고) 제 2 DD+ 인코더(106)는 확장 그룹(122)의 4.0 채널들을 인코딩한다(따라서 4.0 채널 인코더로 칭해질 수 있다). 기본 그룹(121) 및 확장 그룹(122)의 채널들에 대한 인코더들(105, 106)은 통상적으로 서로 아무런 지식을 가지지 않는다. 두 인코더들(105, 106) 각각에는 총 이용가능한 데이터-레이트의 고정된 부분에 대응하는 데이터-레이트가 제공된다. 즉, IS에 대한 인코더(105) 및 DS에 대한 인코더(106)에는 총 이용가능한 데이터-레이트의 고정된 분수(fixed fraction)(예로서, IS 인코더(105)에 대한 총 이용가능한 데이터-레이트의 Z%("IS 데이터-레이트"로 칭해짐) 및 DS 인코더(106)에 대한 총 이용가능한 데이터-레이트의 100%-Z%("DS 데이터-레이트"로 칭해짐), 예를 들면, Z=50)가 제공된다. 각각의 할당된 데이터-레이트들(즉, IS 데이터-레이트 및 DS 데이터-레이트)을 이용하여, IS 인코더(105) 및 DS 인코더(106)는 기본 그룹(121)의 채널들 및 확장 그룹(122)의 채널들의 독립된 인코딩을 각각 수행한다.

[0055] 다음에서는, IS 인코더(105) 및 DS 인코더(106)의 구성요소들에 관한 다른 세부사항들이 일 예시적인 DD+ 다중-채널 인코더(300)의 블록도를 도시한 도 3의 컨텍스트에 기술된다. IS 인코더(105) 및/또는 DS 인코더(106)는 도 3의 DD+ 다중-채널 인코더(300)에 의해 구현될 수 있다. 인코더(300)의 구성요소들을 기술하는 것에 후속하여, 다중-채널 인코더(300)가 어떻게 제 1 비트스트림(제 1 오디오 코덱 시스템을 이용하여 인코딩된)으로부터 제 2 비트스트림(제 2 오디오 코덱 시스템을 이용하여 인코딩된)으로의 효율적인 트랜스코딩을 가능하도록 적용될 수 있는지가 기술된다.

[0056] 다중-채널 인코더(300)는 다중-채널 입력 신호(예로서, 5.1 입력 신호)의 상이한 채널들에 대응하는 PCM 샘플들의 스트림들(311)을 수신한다. PCM 샘플들의 스트림들(311)은 PCM 샘플들의 프레임들로 정렬될 수 있다. 프레임들 각각은 다중-채널 오디오 신호의 특정 채널들의 미리 결정된 수의 PCM 샘플들(예로서, 1536개의 샘플들)을 포함할 수 있다. 이와 같이, 다중 채널 오디오 신호의 각각의 시간 세그먼트에 대해, 다중-채널 오디오 신호의 상이한 채널들 각각에 상이한 오디오 프레임들이 제공된다. 다중-채널 오디오 인코더(300)는 다음에서 다중-채널 오디오 신호의 특정 채널에 대해 기술된다. 그러나, 결과로서 생긴 AC-3 프레임(318)은 통상적으로 다중-채널 오디오 신호의 모든 채널들의 인코딩된 데이터를 포함함을 유념해야 한다.

[0057] PCM 샘플들(311)을 포함하는 오디오 프레임은 입력 신호 조정 유닛(301)에서 필터링될 수 있다. 후속적으로, (필터링된) 샘플들(311)은 시간-대-주파수 변환 유닛(302)에서 시간-도메인으로부터 주파수-도메인으로 변환될 수 있다. 이를 위해, 오디오 프레임은 샘플들의 복수의 블록들로 세분될 수 있다. 블록들은 미리 결정된 길이 L(예로서, 블록 당 256개의 샘플들)을 가질 수 있다. 또한, 인접한 블록들은 오디오 프레임으로부터 샘플들의 특정 정도의 오버랩(예로서, 50% 오버랩)을 가질 수 있다. 오디오 프레임 당 블록들의 수는 오디오 프레임의 특

징(예로서, 과도 존재)에 의존할 수 있다. 통상적으로, 시간-대-주파수 변환 유닛(320)은 오디오 프레임으로부터 도출되는 PCM 샘플들 각각의 블록에 시간-대-주파수 변환(예로서, MDCT(Modified Discrete Cosine Transform) 변환)을 적용한다. 이와 같이, 샘플들의 블록의 각각에 대해, 시간-대-주파수 변환 유닛(302)의 출력에서 변환 계수들(312)의 블록이 획득된다.

[0058] 다중-채널 입력 신호의 각각의 채널은 별도로 처리될 수 있고, 그에 의해 다중-채널 입력 신호의 상이한 채널들에 대해 변환 계수들(312)의 블록들의 개별 시퀀스들을 제공한다. 다중-채널 입력 신호의 채널들의 일부 사이의 상관들(예로서, 서라운드 신호들 Ls 및 Rs 사이의 상관들)을 고려하여, 조인트 채널 처리가 조인트 채널 처리 유닛(303)에서 수행될 수 있다. 일 예시적인 실시예에서, 조인트 채널 처리 유닛(303)은 채널 결합을 수행하고, 그에 의해 결합된 채널들의 그룹을 단일 합성 채널로 변환할 뿐만 아니라, 단일 합성 채널로부터 개별 채널들을 재구성하기 위해 대응하는 디코더 시스템(200, 210)에 의해 이용될 수 있는 보조 정보(side information)를 결합한다. 예로서, 5.1 오디오 신호의 Ls 및 Rs 채널들이 결합될 수 있거나, L, C, R, Ls, 및 Rs 채널들이 결합될 수 있다. 결합이 유닛(303)에서 이용되는 경우, 단일 합성 채널만이 도 3에 도시된 다른 처리 유닛들에 제공된다. 그렇지 않으면, 개별 채널들(즉, 변환 계수들(312)의 블록의 개별 시퀀스들)이 인코더(300)의 다른 처리 유닛들에 넘겨진다.

[0059] 다음에서는, 변환 계수들(312)의 블록들의 일 예시적인 시퀀스에 대한 인코더의 다른 처리 유닛들이 기술된다. 이 기술은 인코딩되는 채널들 각각에(예로서, 다중-채널 입력 신호의 개별 채널들에 또는 채널 결합으로부터 유발되는 하나 이상의 합성 채널들에) 적용가능하다.

[0060] 블록 부동-소수점 인코딩 유닛(304)은 채널(전체 대역폭 채널들(예로서, L, C 및 R 채널들), LFE(저주파 효과들) 채널, 및 결합 채널을 포함한 모든 채널들에 적용가능한)의 변환 계수들(312)을 지수/가수 포맷으로 변환하도록 구성된다. 변환 계수들(312)을 지수/가수 포맷으로 변환함으로써, 변환 계수들(312)의 양자화로부터 유발되는 양자화 잡음은 절대 입력 신호 레벨과 무관해질 수 있다.

[0061] 통상적으로, 유닛(304)에서 수행되는 블록 부동-소수점 인코딩은 변환 계수들(312) 각각을 지수 및 가수로 변환할 수 있다. 지수들은 인코딩된 지수들(313)을 송신하는데 요구되는 데이터-레이트 오버헤드를 감소시키기 위해 가능한 효율적으로 인코딩될 것이다. 동시에, 지수들은 변환 계수들(312)의 스펙트럼 해상도를 잃는 것을 회피하기 위해 가능한 정확하게 인코딩되어야 한다. 다음에는, 상기에 언급된 목적들을 달성하기 위해 DD+에서(및 DD에서) 이용되는 일 예시적인 블록 부동-소수점 인코딩 방식이 간략하게 기술된다. DD+ 인코딩 방식(및 특히, DD+에 의해 이용되는 블록 부동-소수점 인코딩 방식)에 관한 다른 상세들에 대해, 2004년 10월 28일-31일, AEC 컨벤션에서, 문서 Fielder, L.D. 등에 의한 "돌비 디지털 플러스에 대한 도입, 및 돌비 디지털 코딩 시스템에 대한 증진(Introduction to Dolby Digital Plus, and Enhancement to the Dolby Digital Coding System)"에 대한 참조가 이루어지며, 그 콘텐츠는 참조로 포함된다.

[0062] 블록 부동-소수점 인코딩의 제 1 단계에서, 미가공 지수들(raw exponents)이 변환 계수들(312)의 블록을 위해 결정될 수 있다. 이것은, 변환 계수들(402)의 일 예시적인 블록에 대한 미가공 지수들(401)의 블록이 도 4a에 도시된다. 변환 계수들(402)은 값 X를 가진다고 가정되고, 변환 계수들(402)은 X가 1 이하가 되도록 정규화될 수 있다. 값 X는 가수/지수 포맷 $X = m \cdot 2^e$ 로 표현될 수 있고, m은 가수($m \leq 1$)(또한 스케일링된 값으로 칭해짐)이고 e는 지수(또한 스케일 팩터로 칭해짐)이다. 일 실시예에서, 미가공 지수(401)는 0과 24 사이의 값들을 취할 수 있고, 그에 의해 144dB(즉, 2(-0) 내지 2(-24))를 넘는 다이내믹 레인지를 커버한다.

[0063] (미가공) 지수들(401)을 인코딩하기 위해 요구된 비트들의 수를 더욱 감소시키기 위해, 완전한 오디오 프레임의 변환 계수(312)의 블록들(통상적으로 오디오 프레임 당 6개의 블록들)에 걸친 지수들의 시간 공유(time sharing)와 같은 다양한 방식들이 적용될 수 있다. 또한, 지수들은 주파수들에 걸쳐(즉, 변환/주파수-도메인에서 인접한 주파수 빈들에 걸쳐) 공유될 수 있다. 예로서, 지수는 두 개 또는 네 개의 주파수 빈들에 걸쳐 공유될 수 있다. 또한, 변환 계수들(312)의 블록의 지수들은 인접한 지수들 사이의 차가 미리 결정된 최대값, 예를 들면, +/-2를 초과하지 않는 것을 보장하기 위해 텐팅(tenting)될 수 있다. 이것은 변환 계수들(312)의 블록의 지수들의 효율적인 차동 인코딩을 허용한다(예로서, 5개의 차동들을 이용하여). 지수들을 인코딩하기 위해 요구되는 데이터-레이트를 감소시키기 위한 상기에 언급된 방식들(예로서, 시간 공유, 주파수 공유, 텐팅 및 차동 인코딩)은 지수들을 인코딩하기 위해 이용되는 상이한 데이터-레이트들을 유발하는 상이한 지수 코딩 모드들을 규정하기 위해 상이한 방식들로 조합될 수 있다. 상기에 언급된 지수 코딩의 결과로서, 오디오 프레임의 변환 계수들(312)의 블록들(예로서, 오디오 프레임 당 6개의 블록들)에 대해 인코딩된 지수들(313)의 시퀀스가 획득된다.

- [0064] 유닛(304)에서 수행되는 블록 부동-소수점 인코딩 방식의 다른 단계로서, 오리지널 변환 계수들(402)의 가수들 m' 는 대응하는 결과로서 생긴 인코딩된 지수 e' 에 의해 정규화된다. 결과로서 생긴 인코딩된 지수 e' 는 상기에 언급된 미가공 지수 e 와 상이할 수 있다(시간 공유, 주파수 공유 및/또는 텐팅 단계들로 인해). 도 4a의 각각의 변환 계수들(402)에 대해, 정규화된 가수 m' 는 $X = m' \cdot 2^{-e'}$ 로서 결정될 수 있고, 여기서 X 는 오리지널 변환 계수들(402)의 값이다. 오디오 프레임의 블록들에 대한 정규화된 가수들 m' (314)은 가수들(314)의 양자화를 위해 양자화 유닛(306)에 넘겨진다. 가수들(314)의 양자화, 즉 양자화된 가수들(317)의 정확도는 가수 양자화에 이용 가능한 데이터-레이트에 의존한다. 이용가능한 데이터-레이트는 비트 할당 유닛(305)에서 결정된다.
- [0065] 유닛(305)에서 수행되는 비트 할당 처리는, 음향 심리학 원리들에 따라 정규화된 가수들(314)의 각각에 할당될 수 있는 비트들의 수를 결정한다. 비트 할당 처리는 오디오 프레임의 정규화된 가수들을 양자화하기 위해 이용 가능한 비트 카운트를 결정하는 단계를 포함한다. 또한, 비트 할당 처리는 각각의 채널에 대해 전력 스펙트럼 밀도(PSD) 분포 및 주파수-도메인 마스킹 곡선(음향 심리학 모델에 기초하여)을 결정한다. PSD 분포 및 주파수-도메인 마스킹 곡선은 오디오 프레임의 상이한 정규화된 가수들(314)에 대한 이용가능한 비트들의 실질적으로 최적인 분포를 결정하는데 이용된다.
- [0066] 비트 할당 처리에서의 제 1 단계는 얼마나 많은 가수 비트들이 정규화된 가수들(314)을 인코딩하기 위해 이용 가능한지를 결정하는 것이다. 타겟 데이터-레이트는 현재 오디오 프레임을 인코딩하기 위해 이용가능한 비트들의 총수로 변환된다. 특히, 타겟-데이터-레이트는 인코딩된 다중-채널 오디오 신호에 대한 수 k bits/s를 명시한다. T 초의 프레임 길이를 고려하면, 비트들의 총수는 $T*k$ 로서 결정될 수 있다. 가수 비트들의 이용가능한 수는, 메타 데이터, 블록 스위치 플래그들(검출된 과도들 및 선택된 블록 길이들을 시그널링하기 위한), 결합 스케일 팩터들, 지수들 등과 같이, 오디오 프레임을 인코딩하기 위해 이미 이용된 비트들을 차감함으로써 비트들의 총수로부터 결정될 수 있다. 메타데이터는 예를 들면, 트랜스코딩 목적들을 위해 이용될 수 있는 정보를 포함할 수 있다. 비트 할당 처리는 또한, 비트 할당 파라미터들(315)과 같이 다른 양태들에 여전히 할당될 필요가 있는 비트들을 차감할 수 있다(하기 참조). 결과적으로, 이용가능한 가수 비트들의 총수가 결정될 수 있다. 이용가능한 가수 비트들의 총수는 그 후에 오디오 프레임의 모든(예로서, 하나, 둘, 셋 또는 여섯) 블록들에 걸쳐 모든 채널들(예로서, 주요 채널, LFE 채널, 및 결합 채널) 사이에 분배될 수 있다.
- [0067] 또 다른 단계로서, 변환 계수들(312)의 블록의 전력 스펙트럼 밀도("PSD") 분포가 결정될 수 있다. PSD는 입력 신호의 각각의 변환 계수 주파수 bin에서 신호 에너지의 척도이다. PSD는 인코딩된 지수들(313)에 기초하여 결정될 수 있고, 그에 의해 대응하는 다중-채널 오디오 디코더 시스템(200, 210)이 다중-채널 오디오 인코더(300)와 동일한 방식으로 PSD를 결정할 수 있게 한다. 도 4b는 인코딩된 지수들(313)로부터 도출된 변환 계수들(312)의 블록의 PSD 분포(410)를 도시한다. PSD 분포(410)는 변환 계수들(312)의 블록에 대한 주파수-도메인 마스킹 곡선(431)을 계산하는데 이용될 수 있다(도 4d 참조). 주파수-도메인 마스킹 곡선(431)은, 마스크 주파수가 마스크 주파수의 바로 근처의 주파수들을 마스킹하고, 그에 의해 이들 에너지가 특정 마스킹 임계값 아래에 있는 경우 마스크 주파수의 바로 근처의 주파수들을 들리지 않게 하는 현상을 기술하는 음향 심리학 마스킹 효과들을 고려한다. 도 4c는 마스크 주파수(421) 및 이웃 주파수들에 대한 마스킹 임계 곡선(422)을 도시한다. 실제 마스킹 임계 곡선(422)은 DD+ 인코더에서 이용되는 (두 개-세그먼트)(구분적 선형) 마스킹 템플릿(423)에 의해 모델링될 수 있다.
- [0068] 마스킹 임계 곡선(422)의 형상(및 결과에 의해 또한 마스킹 템플릿(423))은, 예를 들면, Zwicker에 의해, 규정된 바와 같은 임계 대역 스케일 상에서(또는 대수 스케일 상에서) 상이한 마스크 주파수들에 대해 실질적으로 변하지 않은 상태로 남아있는 것이 관찰되었다. 이 관찰에 기초하여, DD+ 인코더는 마스킹 템플릿(423)을 밴드(banded) PSD 분포 상에 적용한다(밴드 PSD 분포는 대역들이 임계 대역들의 대략 절반 넓이인 임계 대역 스케일 상에서의 PSD 분포에 대응한다). 밴드 PSD 분포의 경우, 단일 PSD 값은 임계 대역 스케일 상에서(또는 대수 스케일 상에서)의 복수의 대역들 각각에 대해 결정된다. 도 4d는 도 4b의 선형-이격된 PSD 분포(410)에 대한 일 예시적인 밴드 PSD 분포(430)를 도시한다. 밴드 PSD 분포(430)는, 임계 대역 스케일 상에서(또는 대수 스케일 상에서)의 동일 대역 내에 있는 선형-이격된 PSD 분포(410)로부터 PSD 값들을 조합함으로써(예로서, 로그-덧셈 연산을 이용하여) 선형-이격된 PSD 분포(410)로부터 결정될 수 있다. 마스킹 템플릿(423)은 밴드 PSD 분포(430)의 각각의 PSD 값에 적용될 수 있고, 그에 의해 임계 대역 스케일 상에서(또는 대수 스케일 상에서)의 변환 계수들(402)의 블록에 대한 전체 주파수-대역 마스킹 곡선(431)을 산출한다(도 4d 참조).
- [0069] 도 4d의 전체 주파수-도메인 마스킹 곡선(431)은 선형 주파수 해상도로 다시 확장될 수 있고 도 4b에 도시된 변환 계수들(402)의 블록의 선형 PSD 분포(410)와 비교될 수 있다. 이것은, 선형 해상도에 대한 주파수-도메인 마

스킹 곡선(441) 뿐만 아니라, 선형 해상도에 대한 PSD 분포(410)를 도시하는 도 4e에 도시된다. 주파수-도메인 마스킹 곡선(441)은 또한 청각 곡선(hearing curve)의 절대 임계값을 고려해야 할 수 있음을 유념해야 한다.

[0070] 특정 주파수 bin의 변환 계수들(402)의 가수를 인코딩하기 위한 비트들의 수는 PSD 분포(410)에 기초하여 및 마스킹 곡선(441)에 기초하여 결정될 수 있다. 특히, 마스킹 곡선(441) 아래에 있는 PSD 분포(410)의 PSD 값들은 지각적으로 상관없는 가수들에 대응한다(그러한 주파수 bin들에서 오디오 신호의 주파수 성분이 그 부근에서 마스커 주파수에 의해 마스킹되기 때문에). 결과적으로, 이러한 변환 계수들(402)의 가수들에는 어떠한 비트들도 전혀 할당될 필요가 없다. 반면, 마스킹 곡선(441) 위에 있는 PSD 분포(410)의 PSD 값들은 이들 주파수 bin들에서 변환 계수들(402)의 가수들에 인코딩을 위해 비트들이 할당되어야함을 나타낸다. 이러한 가수들에 할당되는 비트들의 수는 PSD 분포(410)의 PSD 값과 마스킹 곡선(441)의 값 사이의 증가하는 차에 따라 증가되어야 한다. 상기에 언급된 비트 할당 처리는 도 4e에 도시된 상이한 변환 계수들(402)에 대한 비트들의 할당(442)을 유발한다.

[0071] 상기에 언급된 비트 할당 처리는 모든 채널들(예로서, 다이렉트 채널들, LFE 채널 및 결합 채널)에 대해 및 오디오 프레임의 모든 블록들에 대해 수행되고, 그에 의해 할당된 비트들의 전체(예비) 수를 산출한다. 할당된 비트들의 이러한 전체 예비 수는 이용가능한 가수 비트들의 총수와 매칭할(즉, 동일할) 가능성이 없다. 일부 경우들에서(예로서, 복잡한 오디오 신호들에 대해), 할당된 비트들의 전체 예비 수는 이용가능한 가수 비트들의 수를 초과할 수 있다(비트 스타베이션(bit starvation)). 다른 경우들에서(예로서, 단순한 오디오 신호들의 경우 예), 할당된 비트들의 전체 예비 수는 이용가능한 가수 비트들의 수보다 아래에 있을 수 있다(비트 서플러스(bit surplus)). 인코더(300)는 통상적으로 할당된 비트들의 전체(최종) 수를 이용가능한 가수 비트들의 수에 가능한 가깝게 매칭하려고 한다. 이를 위해, 인코더(300)는 소위 SNR 오프셋 파라미터를 이용할 수 있다. SNR 오프셋은, 마스킹 곡선(441)을 PSD 분포(410)에 대해 상향 또는 하향으로 이동시킴으로써, 마스킹 곡선(441)의 조정을 허용한다. 마스킹 곡선(441)을 상향 또는 하향으로 이동시킴으로써, 할당된 비트들의(예비) 수는 각각 감소되거나 증가될 수 있다. 이와 같이, SNR 오프셋은 종단 기준(예로서, 할당된 비트들의 예비 수가 이용가능한 비트들의 수(보다 아래에 있지만)에 가능한 근접한 기준; 또는 미리 결정된 최대 수의 반복들이 수행된 기준)이 충족될 때까지 반복 방식으로 조정될 수 있다.

[0072] 상기에 나타난 바와 같이, 할당된 비트들의 최종 수와 이용가능한 비트들의 수 사이의 최상을 매치를 허용하는 SNR 오프셋에 대한 반복 탐색은 이진 탐색을 이용할 수 있다. 이러한 반복에서, 할당된 비트들의 예비 수가 이용가능한 비트들의 수를 초과하는지의 여부가 결정된다. 이러한 결정 단계에 기초하여, SNR 오프셋이 수정되고 다른 반복이 수행된다. 이진 탐색은 $(\log_2(K)+1)$ 반복들을 이용하여 최상의 매치(및 대응하는 SNR 오프셋)를 결정하도록 구성되고, 여기서 K는 가능한 SNR 오프셋들의 수이다. 반복 탐색의 종단 후, 할당된 비트들의 최종 수가 획득된다(이것은 통상적으로, 할당된 비트들의 이전에 결정된 예비 수들 중 하나에 대응한다). 할당된 비트들의 최종 수는 이용가능한 비트들의 수보다(약간) 낮음을 유념해야 한다. 이러한 경우들에서, 스킵 비트들 또는 채움 비트들은 할당된 비트들의 최종 수를 이용가능한 비트들의 수에 완전히 정렬시키는데 이용될 수 있다.

[0073] SNR 오프셋은, 제로의 SNR 오프셋이 오리지널 오디오 신호와 인코딩된 신호 사이의 "바로-인식가능한 차(just-noticeable difference)"로서 알려진 인코딩 조건을 유발하는 인코딩된 가수들을 유발하도록 규정될 수 있다. 즉, 제로의 SNR 오프셋에서, 인코더(300)는 지각 모델에 따라 동작한다. SNR 오프셋의 양의 값은 마스킹 곡선(441)을 하향으로 이동시킬 수 있고, 그에 의해 할당된 비트들의 수를 증가시킨다(통상적으로 아무런 인식가능한 품질 개선 없이). SNR 오프셋의 음의 값은 마스킹 곡선(441)을 상향으로 이동시킬 수 있고, 그에 의해 할당된 비트들의 수를 감소시킨다(및 그에 의해 통상적으로 가청 양자화 잡음을 증가시킨다). SNR 오프셋은 예를 들면, -48에서 +144dB까지의 유효 범위를 가진 10-비트 파라미터일 수 있다. 최적의 SNR 오프셋 값을 찾기 위해, 인코더(300)는 반복 이진 탐색을 수행할 수 있다. 반복 이진 탐색은 그 후에 PSD 분포(410)/마스킹 곡선(441) 비교들의 최대 11회 반복들(10-비트 파라미터의 경우에)을 요구할 수 있다. 실제로 이용되는 SNR 오프셋 값은 비트 할당 파라미터(315)로서 대응하는 디코더로 송신될 수 있다. 또한, 가수들은(최종) 할당된 비트들에 따라 인코딩되고, 그에 의해 양자화된 가수들(317)의 세트를 산출한다.

[0074] DD 및 DD+ 오디오 코덱 시스템의 경우에, 각각의 블록에 대해 csnroffset으로 칭해지는 6 비트 거친(coarse) SNR 오프셋이 존재할 수 있고 각각의 채널에 대해 fsnroffset로 칭해지는 4 비트 미세한(fine) SNR 오프셋 값이 존재할 수 있다. csnroffset 값은 프레임의 모든 블록들에 대해 동일하게 될 수 있고 fsnroffset 값은 프레임의 모든 블록들 및 채널들에 대해 동일하게 될 수 있다. DD+ 오디오 코덱 시스템에서, 파라미터들 csnroffset 및 fsnroffset을 6 비트 frmcsnroffset 및 4 비트 frmfsnroffset 파라미터로서 프레임 당 1회만 송신하도록 선택

될 수 있다.

- [0075] 본 문서에 개요된 바와 같이, DD+ 오디오 코덱 시스템에서, convsnroffset 파라미터가 제공될 수 있다. convsnroffset 파라미터는 통상적으로 두 부분들로 나누어지는 것이 아니라, convsnroffset는 DD+ 비트스트림 내의 각각의 오디오 블록에 대해 통상적으로 10 비트 값이다. 따라서, convsnroffset 파라미터가 csnroffset 및 fsnroffset 파라미터들에 기초하여 결정되는 경우(본 문서에 기재된 바와 같이), convsnroffset 파라미터는 6 비트 csnroffset 및 4 비트 fsnroffset을 단일 값으로 조합하여 결정될 수 있다.
- [0076] 이와 같이, SNR(Signal-to-Noise-Ratio) 오프셋 파라미터는 인코딩된 다중-채널 오디오 신호의 코딩 품질의 표시자로서 이용될 수 있다. 상기에 언급된 SNR 오프셋의 관습에 따라, 제로의 SNR 오프셋은 오리지널 다중-채널 오디오 신호에 대한 "바로-인식가능한 차"를 가진 인코딩된 다중-채널 오디오 신호를 표시한다. 양의 SNR 오프셋은 오리지널 다중-채널 오디오 신호에 대한 적어도 "바로-인식가능한 차"의 품질을 가진 인코딩된 다중-채널 오디오 신호를 표시한다. 음의 SNR 오프셋은 오리지널 다중-채널 오디오 신호에 대한 "바로-인식가능한 차"보다 낮은 품질을 가진 인코딩된 다중-채널 오디오 신호를 표시한다. SNR 오프셋 파라미터의 다른 관습들이 가능함(예로서, 역 관습)을 유념해야 한다.
- [0077] 인코더(300)는 인코딩된 지수들(313), 양자화된 가수들(317), 비트 할당 파라미터들(315) 뿐만 아니라, 다른 인코딩 데이터(예로서, 블록 스위치 플래그들, 메타데이터, 결합 스케일 팩터들 등)를 미리 결정된 프레임 구조(예로서, AC-3 프레임 구조)로 배열하도록 구성된 비트스트림 패킹 유닛(307)을 추가로 포함하고, 그에 의해 다중-채널 오디오 신호의 오디오 프레임에 대한 인코딩된 프레임(318)을 산출한다.
- [0078] 상기에 나타난 바와 같이, 인코더(100, 300)는, 트랜스코더가 제 1 오디오 코덱 시스템(예로서, DD+)에 따라 인코딩되는 인코딩된 프레임(318)을 제 2 오디오 코덱 시스템(예로서, DD)의 디코더에 의해 디코딩될 수 있는 수정된 프레임으로 트랜스코딩할 수 있게 하는 하나 이상의 제어 파라미터들을 결정하도록 구성될 수 있다. 이를 위해, 인코더(100, 300)는, 제 2 오디오 코덱 시스템에 따라 동작하는 오디오 인코더를 시뮬레이션하고 그에 의해 제어 파라미터들을 결정하도록 구성될 수 있다.
- [0079] 이것은 트랜스코딩 시뮬레이션 유닛(320)을 포함하는 도 3의 인코더(300)에 도시된다. 트랜스코딩 시뮬레이션 유닛(320)은 인코더(300)에 의해 이용되는 인코딩된 지수들(313), 양자화된 가수들(317) 및 하나 이상의 비트 할당 파라미터들(315)을 수신하여 제 1 오디오 코덱 시스템에 따라 오디오 신호의 프레임을 인코딩할 수 있다. 또한, 트랜스코딩 시뮬레이션 유닛(320)은 트랜스코더의 기능들을 시뮬레이션(예로서, 제 2 오디오 코덱 시스템에 따라 양자화된 가수들(317)을 역-양자화하고 가수들(317)을 양자화)하도록 구성될 수 있다. 특히, 트랜스코딩 시뮬레이션 유닛(320)은, 트랜스 코딩의 계산 복잡성을 감소시키기 위해 트랜스코더로 송신될 수 있는 제 2 제어 파라미터들(321)(예로서, 하나 이상의 제 2 비트 할당 파라미터들)을 결정하도록 구성될 수 있다.
- [0080] 예로서, DD+ 인코더는 통상적으로, 트랜스코더가 DD+ 비트스트림(복수의 인코딩된 프레임들(318)을 포함)을 640kbps DD 비트스트림으로 변환할 수 있게 하는 소위 convsnroffset 파라미터(즉, 제어 파라미터)를 결정하도록 구성된다. convsnroffset 파라미터는 또한 변환 SNR 오프셋 파라미터로 칭해질 수 있거나, 더욱 일반적으로 제어 파라미터로 칭해질 수 있다. convsnroffset 파라미터의 계산은, 트랜스코더(또한 디코더 변환기 또는 변환기로 칭해질)에서의 DD 포맷으로의 변환의 복잡성을 감소시키도록 돕기 위해 DD+ 인코딩 처리의 컨텍스트에서 수행될 수 있다. convsnroffset 파라미터의 계산은 통상적으로 DD+ 비트스트림의 부분 디코딩 및 인코더(100, 300)에 의한 640kbps DD 인코딩의 시뮬레이션을 필요로 한다. 이것은 인코더(100, 300)가 DD+ 인코더에 대해서 뿐만 아니라, DD 인코더에 대해서도 도 3 및 도 4a 내지 도 4e의 컨텍스트에서 기술된 인코딩 처리를 수행해야 하기 때문에 상당한 계산 복잡성을 유발한다. convsnroffset 파라미터는 통상적으로 640kb/s의 타겟 비트 레이트에서 동작하는 DD 인코더에 대해 도출되는 상기에 언급된 SNR 오프셋에 대응한다. 본 문서에서는 convsnroffset 파라미터를 결정하기 위한 계산 복잡성을 감소시키도록 허용하는 방법들 및 시스템들이 기술된다. 또한, 기술된 방법들 및 시스템들은 DD+ 비트스트림으로부터 DD 비트스트림으로의 트랜스코딩을 수행하는 계산 복잡성을 감소시키는 것을 허용할 수 있다.
- [0081] DD+ 인코더(300)는 인코딩된 오디오 신호의 비트 레이트(주어진 품질에서)를 감소시키거나 인코딩된 오디오 신호의 품질(주어진 비트 레이트에서)을 증가시키기 위한 하나 이상의 코딩 툴들을 이용할 수 있다. 이러한 코딩 툴들은 예를 들면, AHT(Adaptive Hybrid Transform)의 이용, ECPL(Enhanced Coupling)의 이용, SPX(Spectral Extension)의 이용 및/또는 TPNP(Temporal Pre-Noise Processing)의 이용이다. 저 복잡성 DD+ 인코더로서 알려진 변형(예로서, 모바일 디바이스들과 같은 계산 복잡성이 제한된 컴퓨팅 디바이스들과 함께 이용되는)은 통상적으로 상기에 언급된 DD+코딩 툴들을 이용하지 않는다. 이와 같이, DD+ LC 인코더는, 인코딩된 지수들, 양자화

된 가수들, 비트 할당 파라미터 등을 통상적으로 DD 비트스트림 포맷과 상이한 DD+ 비트스트림 포맷으로 인코딩하는 DD 인코더와 유사하거나 이에 대응한다. 이와 같이, (저 복잡성) DD+ 인코더와 DD 인코더 사이에 상당한 오버랩이 존재하는 것이 관찰되었다. 이 오버랩 또는 유사성은 convsnroffset 파라미터를 결정하기 위한 계산 복잡성을 감소시키는데 이용될 수 있다.

[0082] 상기에 나타난 바와 같이, 통상적인 DD+ 인코더(300)는, 트랜스코더에서 DD+ 비트스트림의 640kbps DD 비트스트림으로의 효율적인 변환을 가능하게 하기 위해, convsnroffset 파라미터를 결정한다. convsnroffset 파라미터를 DD+ 비트스트림에 삽입함으로써, 트랜스코더는, convsnroffset 파라미터에 의해 주어진 해상도를 가진 양자화기를 이용하여 가수들을 직접 재-양자화할 수 있으므로, 상기에 언급된 반복 비트 할당 처리(예로서, 11회 반복들을 포함)를 수행할 필요가 없다. 이와 같이, DD 비트스트림에 대한 복잡한 SNR 오프셋 계산은 변환기/트랜스코더에서 인코더로 이동되고 결과는 DD+ 비트스트림 내에서 convsnroffset 파라미터로서 송신된다. 인코더(300)에서의 convsnroffset 파라미터의 계산(소위 스테퍼(stuffer) 내에서 수행되는)은 총 DD+ 인코더 복잡성의 약 25-40%를 필요로 한다. 따라서, convsnroffset 파라미터를 계산하기 위한 복잡성을 감소시키는 것이 바람직하다.

[0083] 본 문서에서, convsnroffset 파라미터를 감소된 복잡성으로 결정하는 것을 허용하는 간소화된 스테퍼가 기술된다. 상기에 개요된 바와 같이, 통상적으로 DD+ 인코더와 DD 인코더 사이에 큰 오버랩이 존재한다. 특히, 도 3 및 도 4a 내지 도 4e의 컨텍스트에서 기술된 부동-소수점 인코딩에 관해 큰 오버랩이 존재한다. 이것은 특히, DD 디코더와 LC DD+ 인코더 사이의 차만이 비트스트림 포맷이 될 수 있는 경우에 저 복잡성 (LC) DD+ 인코더에 대해 참이다. 지수들 및 가수들을 결정하기 위한 방식, 및 지수들을 인코딩하고 가수들을 양자화하기 위한 방식들은 통상적으로 동일하다. 따라서, 스테퍼에 대한 DD+ SNR 오프셋을 재이용하고 동일한 SNR 오프셋 파라미터를 이용하여 DD+ 비트스트림을 DD 비트스트림으로 변환하는 것이 가능할 수 있다. 즉, SNR 오프셋 파라미터(DD+ 코덱의 컨텍스트에서 이용되는)를 convsnroffset 파라미터로서 재이용하는 것이 가능할 수 있고, 그에 의해 명시적 convsnroffset 파라미터 계산을 필요 없게 하고, 그에 의해 (LC) DD+ 인코더의 계산 복잡성을 상당히 감소시킨다.

[0084] 또한, convsnroffset 파라미터로서 SNR 오프셋 파라미터의 재이용은 트랜스코딩된 DD 인코딩된 오디오 신호의 오디오 품질에 관해 이로울 수 있다. 특히, 트랜스코더는 오리지널 DD+ 표현이 유지되기 때문에 오디오 품질에 영향을 주지 않는다. 특히, DD+ 타겟 비트 레이트가 DD 타겟 비트 레이트에 대응하는 경우들에서, 즉 DD+ 비트스트림의 및 DD 비트스트림의 타겟 비트 레이트들이 동일한(예로서, 640kbps) 경우들에서, 트랜스코더는 DD+ 비트스트림으로부터 DD 비트스트림을 생성하기 위한 지수들 및/또는 양자화된 가수들을 재이용하도록 구성될 수 있다. 결과적으로, DD+ 비트스트림 내에 포함된 오디오 신호의 오디오 품질 및 DD 비트스트림 내에 포함된 오디오 신호의 오디오 품질은 동일하다. 또한, 트랜스코더가 DD 비트스트림을 생성할 때 가수들을 역-양자화 및 재-양자화할 필요가 없으므로, 트랜스코더의 복잡성이 감소된다.

[0085] 상기에 나타난 바와 같이, LC DD+ 인코더는 인코딩된 지수들, 양자화된 가수들 등을 DD+ 비트스트림 포맷으로 인코딩하는 DD 인코더로서 뷰잉될 수 있다. DD+ 비트스트림 포맷은 통상적으로 DD 비트스트림 포맷과 상이하다. 특히, DD 비트스트림 포맷에 대한 고정된 비트들의 양(동기화 정보(si); 비트스트림 정보(bsi); 오디오프레임(audfrm); 보조 데이터(auxdata); 에러체크; 지수들; 등에 대한)은 통상적으로 DD+ 비트스트림 포맷에 비해 더 크다. 이것은, DD+ 비트스트림 포맷 및 DD 비트스트림 포맷에 이용되는 고정된 비트들의 수 사이의 차(500)가 복수의 프레임들에 대해 도시되는 도 5에서 알 수 있다. DD 비트스트림 포맷은 DD+ 비트스트림 포맷보다 많은 평균적으로 대략 80 내지 100개의 고정된 비트들을 필요로 한다는 것을 알 수 있다. 결과적으로, DD 비트스트림을 생성하기 위해 DD+ SNR 오프셋을 이용하는 경우에는 640kbps 프레임 크기(640kbps = 20480bits/frame)에서 이용가능한 것보다 많은 비트들을 필요로 하는 비트스트림을 산출하게 된다. 즉 convsnroffset 파라미터로서 DD+에 대해 결정된 SNR 오프셋 파라미터를 이용할 때, 이것은 640kbit/s의 타겟 비트 레이트를 약간 초과하는 DD 비트스트림을 유발하게 된다. 그러나, 트랜스코더가 통상적으로 20480bits/frame의 고정된 프레임 크기, 즉 타겟 비트 레이트에 대응하는 고정된 프레임 크기를 제공하므로, 이것은 일반적으로 수용가능하지 않다.

[0086] 이 문제를 극복하기 위한 상이한 방식들이 이용될 수 있으며, 방식들은 DD+ 타겟 비트 레이트에 의존한다. 640kbits/s의 DD+ 타겟 비트 레이트의 경우에, 즉 DD 타겟 비트 레이트에 대응하는 DD+ 타겟 비트 레이트의 경우에, 상기에 언급된 문제는 DD+ 인코더(300)의 비트 할당 처리의 컨텍스트에서 DD / DD+ 고정된 비트들 차를 고려함으로써 극복될 수 있다. 상기에 개요된 바와 같이, 반복 비트 할당 처리는 이용가능한 가수 비트들의 총수, 즉 가수들의 양자화에 할당될 수 있는 비트들의 총수를 결정하는 것으로 시작한다. 본 문서에는 이용가능한 가수 비트들의 DD+ 특정 총수로부터 DD / DD+ 고정된 비트들 차를 차감하고, 그에 의해 DD로의 가능한 트랜스코딩을 고려하는 이용가능한 가수 비트들의 감소된 총수를 산출하는 것이 제안된다. 차감되는 DD / DD+ 고정된 비

트들 차는 프레임 특정 방식으로 결정될 수 있거나 그것은 평균 또는 최악의 경우의 값에 대응할 수 있다. DD+ SNR 오프셋 계산은 그 후에 이용가능한 가수 비트들의 감소된 총수를 이용하여 수행될 수 있다.

[0087] 결과적으로, DD+인코딩된 오디오 신호의 품질은 약간 감소된다. 그러나, 관찰된 최악의 경우의 페널티가 3kbps의 비트 레이트 또는 총 DD+ 타겟 비트 레이트의 0.5%에 대응하는 프레임 당 DD / DD+ 고정된 비트들 차의 102 비트들의 범위에 있다는 사실로 인해, 오디오 품질에 대한 영향은 낮다. 상기에 나타난 바와 같이, 이용가능한 가수 비트들의 감소된 총 수로 인해 DD+ 비트스트림 내에서 이용되지 않은 비트들은 스킵 비트들 또는 채움 비트들로 채워질 수 있고, 그에 의해 640kbps/s의 DD+ 타겟 비트 레이트에서 DD+ 호환가능한 프레임들을 산출한다.

[0088] 다른 결과로서, DD+ 인코딩 처리의 컨텍스트에서 계산된 SNR 오프셋은 이제 convsnroffset 파라미터로서 이용될 수 있다. 이제 트랜스코딩된 DD 비트스트림은 640kbps의 DD 타겟 비트 레이트를 충족한다는 것이 보장된다.

[0089] 다른 이점으로서, 트랜스코더(또는 변환기)의 복잡성이 감소될 수 있음을 유념해야 한다. 트랜스코더는 부분적 DD+ 디코딩 및 DD 재-인코딩을 수행할 필요없이, DD+ 인코딩된 지수들 및 DD+ 양자화된 가수들을 DD 비트스트림으로 복사할 수 있다.

[0090] 다른 방식은 DD+ 타겟 비트 레이트가 DD 타겟 비트 레이트보다 작은 상황에서 취해질 수 있다. 예로서, DD+ 타겟 비트 레이트는 448kbps 또는 384kbps일 수 있다. 변환기는 통상적으로 감소된 DD+ 타겟 비트 레이트들이 이용가능하지 않도록 단 하나의 DD 타겟 비트 레이트(예로서, 640kbps)에 제한된다. 그럼에도 불구하고, DD+ 인코딩의 컨텍스트에서 결정되는 SNR 오프셋은 convsnroffset 파라미터로서 재이용될 수 있다. 이것은 임의의 경우에 DD+ 인코딩된 오디오 신호의 품질이 DD+ 타겟 비트 레이트에 의해 제한된다는 사실로 인해 가능하다. DD 타겟 비트 레이트보다 낮은 DD+ 타겟 비트 레이트에서 인코딩된 DD+ 인코딩된 오디오 신호의 트랜스코딩은 DD+ 인코딩된 오디오 신호보다 높은 오디오 품질을 가진 DD 인코딩된 오디오 신호를 제공할 수 없다.

[0091] 그러나, 비교적 낮은 DD+ 타겟 비트 레이트에서 동작하는 DD+ 인코더는 DD 인코더에 의해 이용되지 않은 코딩 톨들을 이용할 수 있다. 이와 같이, 이들 코딩 톨들의 영향은 고려되어야 한다. DD+ 인코더가 풀 채널들의 인코딩된 지수들 및 양자화된 가수들을 제공하는 경우, 이들 풀 채널들(즉, 인코딩된 지수들 및 양자화된 가수들)은 DD 비트스트림으로 복사될 수 있고, 그에 의해 DD+ 디코딩 및 DD 재-인코딩의 단계들이 불필요해지기 때문에 통상적인 트랜스코더들에 비해 오디오 품질(즉, 신호대 잡음비)을 개선한다.

[0092] DD+ 인코더가 하나 이상의 결합 채널들을 제공하는 경우(통상적으로, DD 및 DD+ 인코더가 단일 결합 채널만을 제공하는 경우), DD 타겟 비트 레이트(640kbps)에서의 DD 인코더가 통상적으로 결합을 이용하지 않기 때문에, 결합 채널들은 통상적으로 DD 비트스트림 내의 풀 채널들로서 개별적으로 디인코딩 및 재-인코딩되어야 한다. 이 트랜스코딩은 DD+ 인코딩된 오디오 신호에 비해 DD 인코딩된 오디오 신호의 품질 손상을 유발할 수 있다(DD+ 디코딩 및 DD 재-인코딩 동작들로 인해). 또한, 복수의 풀 채널들의 DD 인코딩은 통상적으로 감소된 수의 결합 채널들의 DD+ 인코딩에 비해 증가된 양의 비트들을 필요로 한다. 예로서, 5.1 다중-채널 오디오 신호의 모든 5개의 채널들이 결합되었을 수 있고, 이것은 단일 오리지널 결합 채널이 DD 인코더에 의해 5회 인코딩되어야 하는 상황을 유발한다. 오리지널 결합 채널을 다수회(예로서, 5회) 인코딩하기 위해 필요한 부가 비트들은 풀 채널들에 대한 더 작은 수요(결합 채널들에 대한 비트 수요에 비해)에 의해 보상될 수 있다.

[0093] 도 6은, 복수의 상이한 오디오 신호들의 오디오 품질이 분석되는 예시적인 MUSHRA(MUltiple Stimuli with Hidden Reference and Anchor) 테스트들을 도시한다. 특히, 명시적으로 계산된 convsnroffset 파라미터를 이용하여 트랜스코딩된 트랜스코딩된 신호의 오디오 품질(601)은 DD+ 인코딩된 오디오 신호의 SNR 오프셋에 대응하는 convsnroffset 파라미터를 이용하여 트랜스코딩된 트랜스코딩 신호의 오디오 품질(602)과 비교된다. 도시된 예에서, DD+ 타겟 비트 레이트는 384kbps이고 DD 타겟 비트 레이트는 640kbps이다. 도시된 예에서, DD+ 인코더(300)는 결합(약 10kHz에서 결합 시작 주파수와)을 이용한다. 예시된 복수의 상이한 오디오 신호들에 대해, 상당한 품질 저하가 관찰될 수 없음이 관찰될 수 있다. 반면, 인코더(300)에서의 계산 복잡성 및 트랜스코더에서의 가능한 계산 복잡성이 상당히 감소되었다.

[0094] 변환된(즉, 트랜스코딩된) 비트스트림의 비트 레이트는 DD 타겟 비트 레이트(예로서, 640kbps)를 초과할 수 있음을 유념해야 한다. 이것은, 최악-경우의 DD+/DD 고정된 비트 차가 정확하게 결정되지 않는 경우(즉, 너무 낮다고 가정되는 경우), 640kbps DD+ 경우에 대해(즉, DD+ 타겟 비트 레이트가 DD 타겟 비트 레이트에 대응하는 경우에 대해) 발생할 수 있다. 대안적으로 또는 부가적으로, 이것은 하나 이상의 확장된 결합 채널들이 변환시 이용가능한 것보다 많은 비트들을 필요로 하는 경우, 더 낮은 데이터 레이트들에 대해(즉, DD+ 타겟 비트 레이트

트가 DD 타겟 비트 레이트보다 낮은 경우에 대해) 발생할 수 있다.

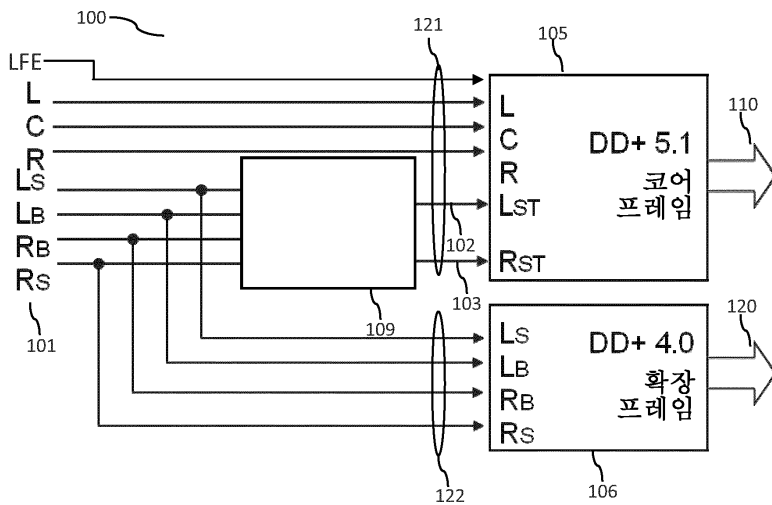
- [0095] 인코더(300)는, DD+ SNR 오프셋이 convsnroffset 파라미터로서 이용되는 경우, 변환된 DD 비트스트림이 DD 타겟 비트 레이트를 초과하는 상기에 언급된 상황을 검출하도록 구성될 수 있다. 특히, DD+ 인코더(300)는 단일 비트 할당 반복(convsnroffset 파라미터의 명시적 결정을 위해 필요한 11회 반복들에 비해)으로 변환된 DD 비트 스트림에 대한 DD+ SNR 오프셋을 유효화하도록 구성될 수 있다. 이것은 프레임마다 기초하여 확인될 수 있다.
- [0096] (특정 프레임에 대해) DD+ SNR 오프셋을 convsnroffset 파라미터로서 이용하는 것이 DD 타겟 비트 레이트를 초과하는 비트들의 수를 유발한다고 결정되는 경우, 인코더(300)는 하나 이상의 복구 전략들을 적용할 수 있다: 예로서, 인코더(300)는 명시적 convsnroffset 계산을 폴백(fallback)으로 수행하도록 구성될 수 있다. DD+ SNR 오프셋은 개선된 시작점으로 이용될 수 있고, 그에 의해 요구된 반복들의 횟수를 잠재적으로 감소시킨다. 대안적으로 또는 부가적으로, DD+ SNR 오프셋에 기초하여 초기 SNR 오프셋을 결정하는데 실증적 분석이 이용될 수 있고, 초기 SNR 오프셋은 비트 할당 반복들의 수를 감소(예로서, 최소화)시킨다. 대안적으로 또는 부가적으로, 명시적인 convsnroffset 계산이 이용될 수 있지만, 반복 처리는 충분하다고 간주되는(예로서, 마스킹 임계값 아래의 6dB인 양자화 잡음을 유발하는) 중간 결과가 얻어질 때 중단될 수 있다.
- [0097] 본 문서에서는 SNR 오프셋 값을 트랜스코딩/변환기에서 DD 인코딩에 이용되는 convsnroffset 값에 복사하는 것이 제안되었다. 이 방식은 특히 640kbps에서 동작하는 LC DD+ 인코더에 관련되며, 이것은 LC DD+ 인코더가 이 타겟 비트 레이트를 위한 결합 또는 DD+ 툴들 중 어느 것도 이용하지 않기 때문이다. 더 낮은 비트레이트들에 대해, LC DD+ 인코더는 통상적으로 결합을 이용한다. 그러나, DD+ SNR 오프셋 값은 오디오 품질의 적은 잠재적인 저하만으로 convsnroffset 값에 대해 이용될 수 있다.
- [0098] 상기에 개요된 바와 같이, 640kbps DD 포맷은 통상적으로, 보조 정보를 저장하는데 640kbps DD+ 포맷보다 더 많은 비트들을 필요로 한다. 본 문서에는 DD+ 인코딩 처리 동안에 비트 차를 고려하는 것이 제안된다. DD+에 대한 손상된 비트 레이트의 최대량은 3kbps 또는 총 비트 레이트의 0.5%로 측정되었고, 이것은 DD+ 비트스트림의 가청 저하를 유발하지 않는다. 그러나, DD+ 인코딩 동안 비트 차를 고려함으로써, DD+대 DD 트랜스코딩에 대해서 뿐만 아니라, DD+ 인코딩에 대해 동일한 SNR 오프셋을 이용하는 것이 가능하다. DD+ 비트스트림의 및 트랜스코딩된 DD 비트스트림의 결과로서 생긴 디코더 출력은 통상적으로 DD+ 디코더에 의해 및 DD 디코더에 의해 적용된 상이한 디더링(different dithering)을 제외하면 동일하다.
- [0099] LC DD+ 인코더의 더 낮은 비트 레이트들에 대해(예로서, 448kbps 및 384kbps), 결합은 통상적으로 LC DD+ 인코더에 의해 이용된다. 변환기는 통상적으로 DD+ 비트스트림을 결합없이 640kbps DD 비트스트림으로 변환한다. 청취 테스트는, 변환기에 대한 DD+ SNR 오프셋을 이용(즉, convsnroffset을 DD+ SNR 오프셋과 동일하게 설정)하는 것이 명시적으로 계산된 convsnroffset 파라미터를 이용하여 변환기에 의해 도출된 트랜스코딩된 신호의 오디오 품질에 비교할 수 있는 트랜스코딩된 신호의 오디오 품질을 산출한다는 것을 보여준다. 실험 결과들은 또한 폴 채널들이 통상적으로 DD 타겟 비트 레이트(예로서, 640 kbps의)에 의해 설정된 제한을 초과하지 않을 때 결합 채널들의 인코딩에 의해 유발되는 비트들이 증가하는 것을 보여주었다.
- [0100] DD+ 인코더는 DD+ SNR 오프셋이 변환된 DD 비트스트림에 대해 무효한지의 여부(즉, DD 비트스트림을 생성하기 위해 변환기 내에서 DD+ SNR 오프셋을 이용할 때 초과하는 비트들의 수가 존재하는지의 여부)를 결정하도록 구성될 수 있다. 이것이 그 경우이면, 명시적인 변환기 snroffset(즉, convsnroffset) 파라미터 계산을 이러한 비트 오버플로가 발생하는 특정 프레임에 대한 폴백으로서 이용하는 것이 가능하다. 그럼에도 불구하고, DD+ snroffset 값을 convsnroffset 파라미터 계산에 대한 더 양호한 시작점으로서 이용함으로써 및/또는 예를 들면, 중간 결과가 이미 미리-결정된 품질 기준을 충족할 때, 최적의 결과를 찾기 전에 반복을 중단함으로써, 계산 복잡성을 감소시키는 것이 가능할 수 있다.
- [0101] 본 문서에 기술된 방법들 및 시스템들은 소프트웨어, 펌웨어 및/또는 하드웨어로서 구현될 수 있다. 특정 구성 요소들은 예를 들면, 디지털 신호 처리기 또는 마이크로프로세서 상에서 구동하는 소프트웨어로서 구현될 수 있다. 다른 구성요소들은 예를 들면, 하드웨어로서 및/또는 주문형 반도체들로서 구현될 수 있다. 기술된 방법들 및 시스템들에서 접하게 되는 신호들은 랜덤 액세스 메모리 또는 광 저장 매체들과 같은 매체들 상에 저장될 수 있다. 이들은 라디오 네트워크들, 위성 네트워크들, 무선 네트워크들 또는 유선 네트워크들, 예를 들면, 인터넷과 같은 네트워크들을 통해 이전될 수 있다. 본 문서에 기술된 방법들 및 시스템들을 이용하는 통상적인 디바이스들은 오디오 신호들을 저장 및/또는 렌더링하는데 이용되는 휴대용 전자 디바이스들 또는 다른 소비자 기기이다.

부호의 설명

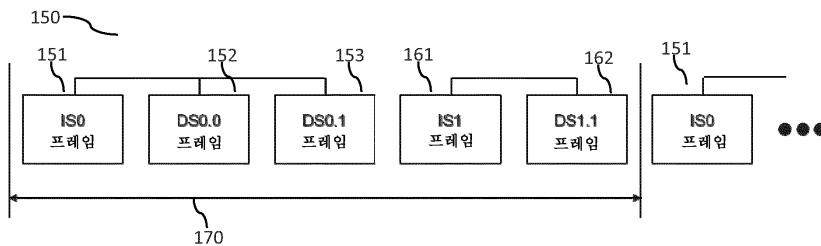
- [0102] 100: DD+ 7.1 다중-채널 오디오 인코더
- 101: 칠(7) 플러스 일(1) 오디오 채널들
- 102, 103: 다운믹스드 서라운드 채널들
- 105: IS 인코더
- 106: DS 인코더
- 200, 210: 다중-채널 디코더 시스템
- 205: 5.1 다중-채널 디코더
- 215: 7.1 다중-채널 디코더
- 300: DD+ 다중-채널 인코더
- 302: 시간-대-주파수 변환 유닛
- 304: 블록 부동-소수점 인코딩 유닛
- 307: 비트스트림 패킹 유닛
- 320: 시간-대-주파수 변환 유닛

도면

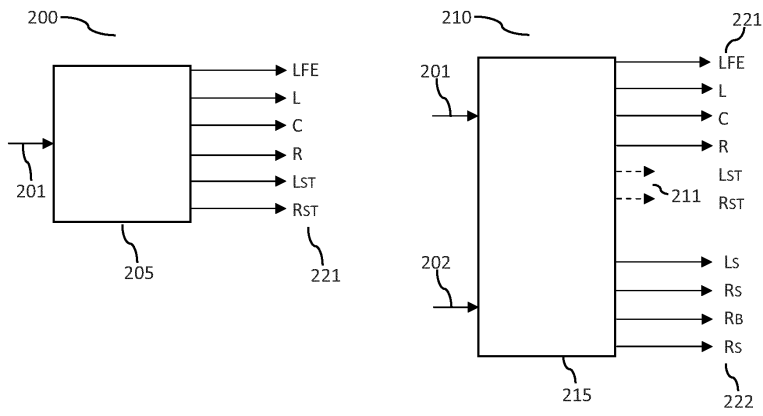
도면1a



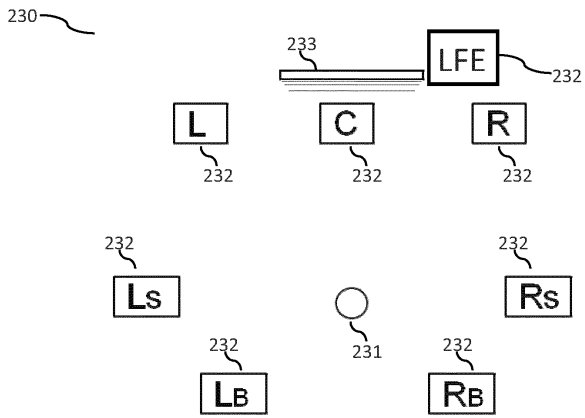
도면1b



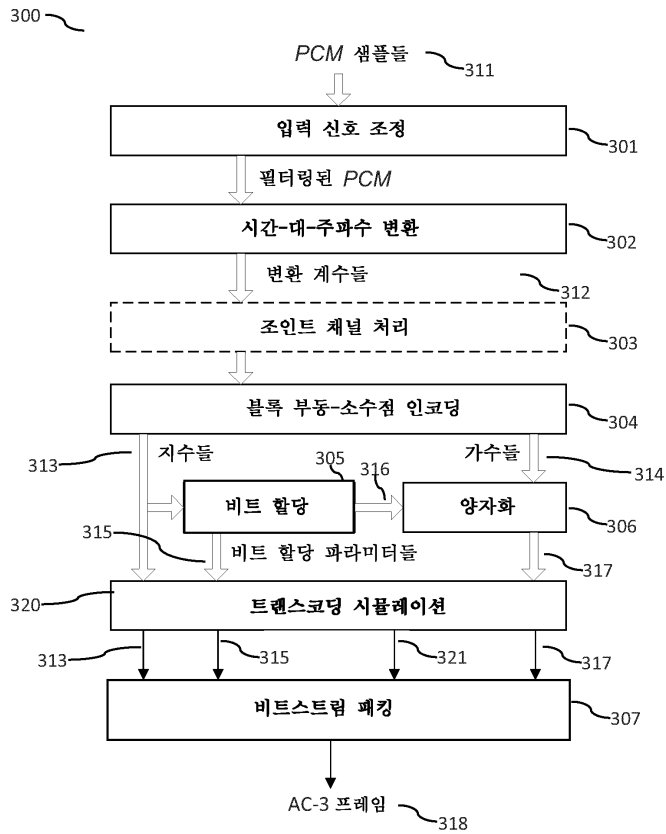
도면2a



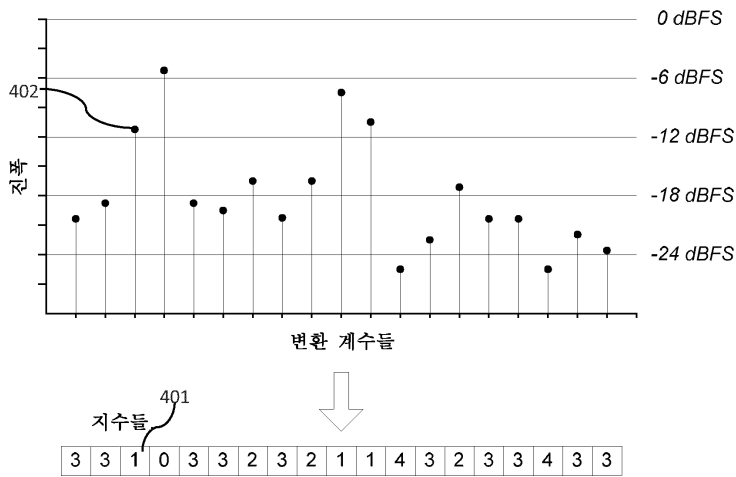
도면2b



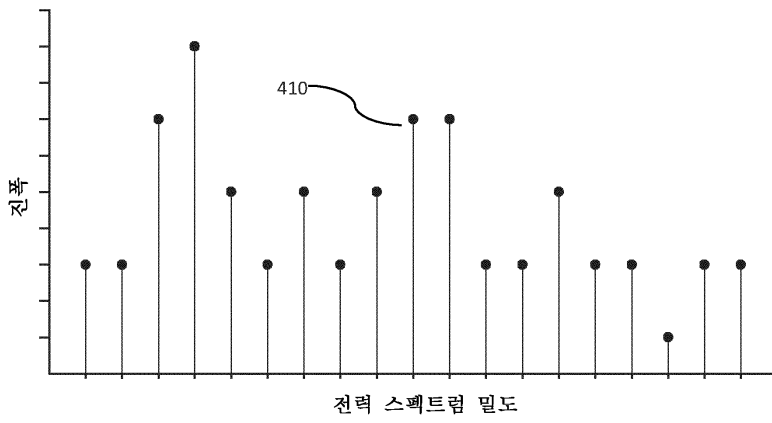
도면3



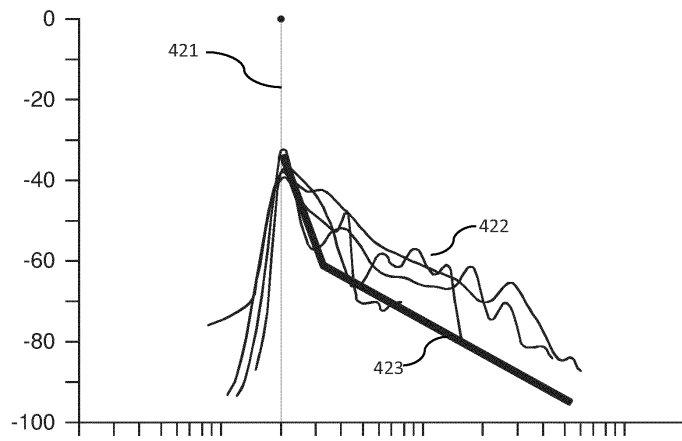
도면4a



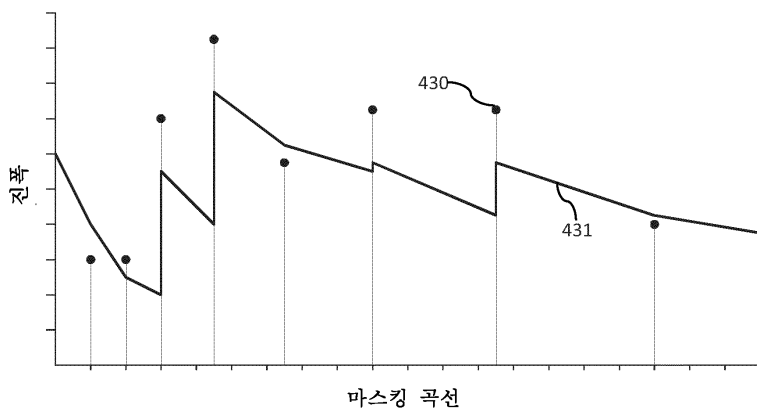
도면4b



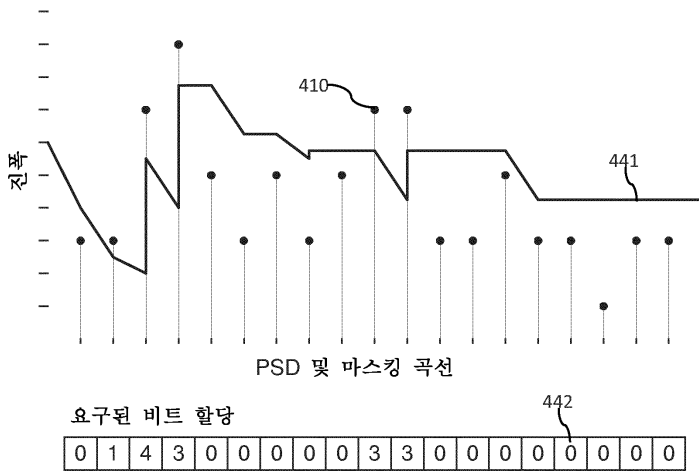
도면4c



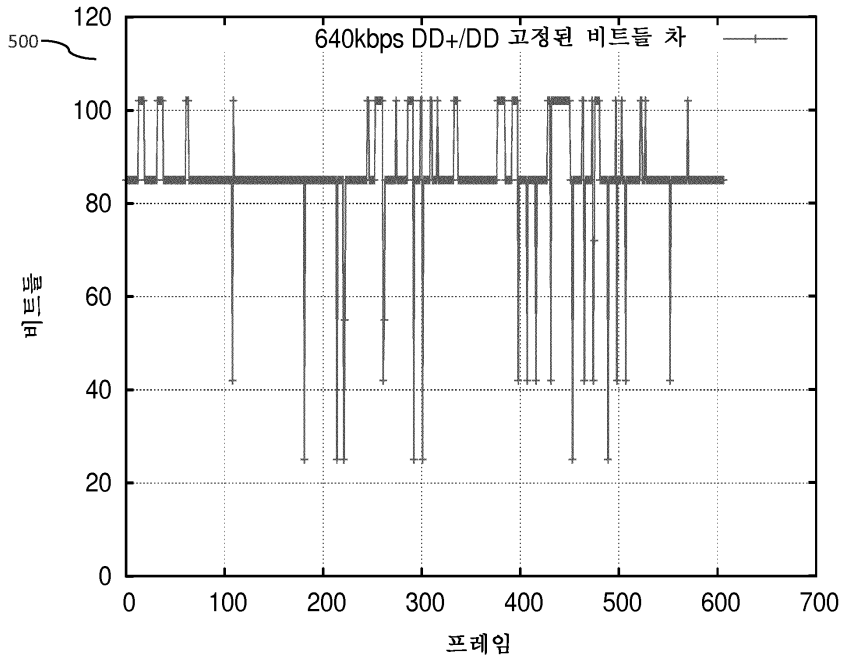
도면4d



도면4e



도면5



도면6

