



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0128126
(43) 공개일자 2019년11월15일

(51) 국제특허분류(Int. Cl.)
G10L 19/02 (2006.01) G10L 19/00 (2006.01)
G10L 19/002 (2013.01) G10L 19/008 (2014.01)
G10L 19/16 (2013.01)
(52) CPC특허분류
G10L 19/02 (2013.01)
G10L 19/0017 (2013.01)
(21) 출원번호 10-2019-0140945(분할)
(22) 출원일자 2019년11월06일
심사청구일자 2019년11월06일
(62) 원출원 특허 10-2012-0059434
원출원일자 2012년06월01일
심사청구일자 2017년05월17일
(30) 우선권주장
2011121982 2011년06월01일 러시아(RU)

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
(72) 발명자
주기현
경기도 수원시 영통구 삼성로 129(매탄동)
포로브 안톤
러시아 연방 196240 상트페테르부르크 크라스노푸
틸로브스카야 스트리트 99-35
오시포브 콘스탄틴
러시아연방 198205 상트페테르부르크 파르티자나
제르마나 스트리트 28-2-133
(74) 대리인
리엔목특허법인

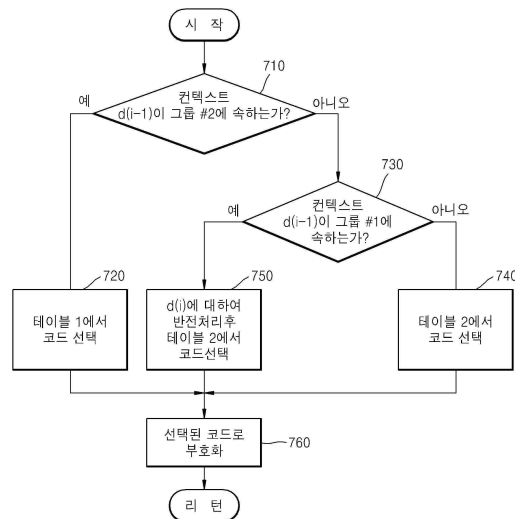
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 오디오 부호화방법 및 장치, 오디오 복호화방법 및 장치, 그 기록매체 및 이를 채용하는 멀티미디어 기기

(57) 요약

오디오 부호화방법은 오디오 스펙트럼에 대하여, 소정의 서브밴드 단위로 엔벨로프를 획득하는 단계; 상기 서브밴드 단위로, 상기 엔벨로프에 대하여 양자화하는 단계; 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 부호화를 수행하는 단계를 포함한다. 이에 따르면, 한정된 비트 범위에서 오디오 스펙트럼의 엔벨로프 정보를 부호화하는데 소요되는 비트수를 감소시킴으로써 실제 스펙트럼 성분을 부호화하는데 소요되는 비트수를 증가시킬 수 있다.

대표도 - 도7



(52) CPC특허분류

G10L 19/002 (2013.01)

G10L 19/008 (2013.01)

G10L 19/167 (2013.01)

명세서

청구범위

청구항 1

주파수 도메인에서 구분된 복수의 서브밴드로 이루어지는 오디오 스펙트럼에 대하여 엔벨로프를 획득하는 단계;
이전 서브밴드의 양자화 인덱스와 현재 서브밴드의 양자화 인덱스를 포함하는 양자화 인덱스를 얻기 위하여 상기 엔벨로프에 대해 양자화하는 단계;

상기 이전 서브밴드의 양자화 인덱스와 상기 현재 서브밴드의 양자화 인덱스로부터 상기 현재 서브밴드의 차분 양자화 인덱스를 획득하는 단계;

상기 이전 서브밴드의 차분 양자화 인덱스를 이용하여 상기 현재 서브밴드의 컨텍스트를 획득하는 단계; 및

상기 현재 서브밴드의 컨텍스트에 따라 상기 현재 서브밴드의 차분 양자화 인덱스에 대해 무손실 부호화하는 단계를 포함하되,

상기 현재 서브밴드의 이전 서브밴드가 첫 번째 서브밴드인 경우, 상기 현재 서브밴드의 컨텍스트는 상기 첫 번째 서브밴드의 양자화 인덱스를 이용하여 획득되는 오디오 부호화 방법.

청구항 2

제1 항에 있어서, 상기 엔벨로프는 대응하는 서브밴드의 평균 에너지, 평균 진폭, 파워 및 norm 값 중 하나인 오디오 부호화 방법.

청구항 3

제1 항에 있어서, 상기 무손실 부호화하는 단계는 상기 차분 양자화 인덱스가 소정 범위를 갖도록 조정된 다음 허프만 부호화하는 것을 특징으로 하는 오디오 부호화 방법.

청구항 4

제1 항에 있어서, 상기 무손실 부호화하는 단계는,

상기 컨텍스트에 대응하는 이전 서브밴드의 차분 양자화 인덱스를 복수의 그룹 중 하나로 그룹핑하는 단계와,

각 그룹에 대하여 정의된 허프만 테이블을 이용하여 상기 현재 서브밴드의 차분 양자화 인덱스에 대해 허프만 부호화하는 단계를 포함하는 것을 특징으로 하는 오디오 부호화 방법.

청구항 5

제1 항에 있어서, 상기 무손실 부호화하는 단계는,

상기 컨텍스트에 대응하는 이전 서브밴드의 차분 양자화 인덱스를 제1 내지 제3 그룹 중 하나로 그룹핑하는 단계;

상기 제2 그룹을 위한 제1 허프만 테이블과 상기 제1 및 제3 그룹이 공유하는 제2 허프만 테이블을 포함하는 2개의 허프만 테이블을 할당하여 상기 현재 서브밴드의 차분 양자화 인덱스에 대해 허프만 부호화하는 단계를 포함하는 것을 특징으로 하는 오디오 부호화 방법.

청구항 6

제1 항 내지 제5 항 중 어느 한 항에 기재된 오디오 부호화 방법을 컴퓨터에 의해 실행할 수 있는 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기술 분야

[0001] 본 발명은 오디오 부호화/복호화에 관한 것으로서, 보다 구체적으로는, 한정된 비트 범위에서 오디오 스펙트럼의 엔벨로프 정보를 부호화하는데 소요되는 비트수를 감소시킴으로써 실제 스펙트럼 성분을 부호화하는데 소요되는 비트수를 증가시킬 수 있는 오디오 부호화방법 및 장치, 오디오 복호화방법 및 장치, 그 기록매체 및 이를 채용하는 멀티미디어 기기에 관한 것이다.

배경 기술

[0002] 오디오 신호의 부호화시 실제의 스펙트럼 성분 이외에 엔벨로프와 같은 부가정보가 비트스트림에 포함될 수 있다. 이때, 손실을 최소화하면서 부가정보의 부호화에 할당되는 비트수를 감소시킴으로써, 실제의 스펙트럼 성분의 부호화에 할당되는 비트수를 증가시킬 수 있다.

[0003] 즉, 오디오 신호를 부호화하거나 복호화하는 경우, 특히 낮은 비트율에서 한정된 비트를 효율적으로 이용함으로써, 해당 비트 범위에서 최상의 음질을 갖는 오디오 신호를 복원하는 것이 요구된다.

발명의 내용

해결하려는 과제

[0004] 본 발명이 해결하고자 하는 과제는 한정된 비트 범위에서 오디오 스펙트럼의 엔벨로프 정보를 부호화하는데 소요되는 비트수를 감소시키는 반면 실제 스펙트럼 성분을 부호화하는데 소요되는 비트수를 증가시킬 수 있는 오디오 부호화방법 및 장치, 오디오 복호화방법 및 장치, 그 기록매체와 이를 채용하는 멀티미디어 기기를 제공하는데 있다.

과제의 해결 수단

[0005] 상기 과제를 달성하기 위한 본 발명의 일실시예에 따른 오디오 부호화방법은, 오디오 스펙트럼에 대하여, 소정의 서브밴드 단위로 엔벨로프를 획득하는 단계; 상기 서브밴드 단위로, 상기 엔벨로프에 대하여 양자화하는 단계; 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 부호화를 수행하는 단계를 포함할 수 있다.

[0006] 상기 과제를 달성하기 위한 본 발명의 일실시예에 따른 오디오 부호화장치는, 오디오 스펙트럼에 대하여, 소정의 서브밴드 단위로 엔벨로프를 획득하는 엔벨로프 획득부; 상기 서브밴드 단위로, 상기 엔벨로프에 대하여 양자화하는 엔벨로프 양자화부; 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 부호화를 수행하는 엔벨로프 부호화부; 상기 오디오 스펙트럼에 대하여 양자화 및 무손실 부호화를 수행하는 스펙트럼 부호화부를 포함할 수 있다.

[0007] 상기 과제를 달성하기 위한 본 발명의 일실시예에 따른 오디오 복호화방법은, 비트스트림으로부터 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 복호화를 수행하는 단계; 및 상기 무손실 복호화결과 복원된 현재 서브밴드의 차이값으로부터 서브밴드 단위로 상기 양자화된 엔벨로프를 구하여 역양자화를 수행하는 단계를 포함할 수 있다.

[0008] 상기 과제를 달성하기 위한 본 발명의 일실시예에 따른 오디오 복호화장치는, 비트스트림으로부터 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 복호화를 수행하는 엔벨로프 복호화부; 상기 무손실 복호화결과 복원된 현재 서브밴드의 차이값으로부터 서브밴드 단위로 상기 양자화된 엔벨로프를 구하여 역양자화를 수행하는 엔벨로프 역양자화부; 및 상기 비트스트림에 포함된 스펙트럼 성분에 대하여 무손실 복호화 및 역양자화를 수행하는 스펙트럼 복호화부를 포함할 수 있다.

[0009] 상기 과제를 달성하기 위한 본 발명의 일실시예에 따른 멀티미디어 기기는, 오디오 스펙트럼에 대하여, 소정의 서브밴드 단위로 엔벨로프를 획득하고, 상기 서브밴드 단위로, 상기 엔벨로프에 대하여 양자화하고, 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 부호화를 수행하는 부호화모듈을 포함할 수 있다.

[0010] 상기 멀티미디어 기기는 비트스트림으로부터 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 복호화를 수

행하고, 상기 무손실 복호화결과 복원된 현재 서브밴드의 차이값으로부터 서브밴드 단위로 상기 양자화된 엔벨로프를 구하여 역양자화를 수행하는 복호화모듈을 더 포함할 수 있다.

발명의 효과

[0011] 복잡도의 증가 및 복원된 음질의 열화없이, 한정된 비트 범위에서 오디오 스펙트럼의 엔벨로프 정보를 부호화하는데 소요되는 비트수를 감소시킴으로써 실제 스펙트럼 성분을 부호화하는데 소요되는 비트수를 증가시킬 수 있다.

도면의 간단한 설명

[0012] 도 1은 본 발명의 일실시예에 따른 디지털신호 처리장치의 구성을 나타낸 블록도이다.

도 2는 본 발명의 다른 실시예에 따른 디지털신호 처리장치의 구성을 나타낸 블록도이다.

도 3a 및 도 3b는 양자화 해상도가 0.5이고 양자화 스텝사이즈가 3.01인 경우, 최적화되지 않은 로그 스케일과 최적화된 로그 스케일을 비교한 도면이다.

도 4a 및 도 4b는 양자화 해상도가 1이고 양자화 스텝사이즈가 6.02인 경우, 최적화되지 않은 로그 스케일과 최적화된 로그 스케일을 비교한 도면이다.

도 5는 최적화되지 않은 로그 스케일의 양자화 결과와 최적화된 로그 스케일의 양자화 결과를 비교한 도면이다.

도 6은 이전 서브밴드의 양자화 델타값을 컨텍스트로 사용하는 경우, 선택되는 3개 그룹의 확률 분포를 나타낸 도면이다.

도 7은 도 1의 엔벨로프 부호화부에서의 컨텍스트 기반 부호화 동작을 설명하는 도면이다.

도 8은 도 2의 엔벨로프 복호화부에서의 컨텍스트 기반 복호화 동작을 설명하는 도면이다.

도 9는 본 발명의 일실시예에 따른 부호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.

도 10은 본 발명의 일실시예에 따른 복호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.

도 11은 본 발명의 일실시예에 따른 부호화모듈과 복호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0013] 본 발명은 다양한 변환을 가할 수 있고 여러가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 구체적으로 설명하고자 한다. 그러나 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 기술적 사상 및 기술 범위에 포함되는 모든 변환, 균등물 내지 대체물을 포함하는 것으로 이해될 수 있다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0014] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들이 용어들에 의해 한정되는 것은 아니다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.

[0015] 본 발명에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 본 발명에서 사용한 용어는 본 발명에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나 이는 당 분야에 종사하는 기술자의 의도, 관례, 또는 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 발명에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 발명의 전반에 걸친 내용을 토대로 정의되어야 한다.

[0016] 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 발명에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0017] 이하, 본 발명의 실시예들을 첨부 도면을 참조하여 상세히 설명하기로 하며, 첨부 도면을 참조하여 설명함에 있

어, 동일하거나 대응하는 구성요소는 동일한 도면번호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다.

[0018] 도 1은 본 발명의 일실시예에 따른 디지털신호 처리장치의 구성을 나타낸 블록도이다.

[0019] 도 1에 도시된 디지털신호 처리장치(100)는 변환부(110), 엔벨로프 획득부(120), 엔벨로프 양자화부(130), 엔벨로프 부호화부(140), 스펙트럼 정규화부(150) 및 스펙트럼 부호화부(160)를 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나의 이상의 프로세서(미도시)로 구현될 수 있다. 여기서, 디지털신호는 비디오, 이미지, 오디오 혹은 음성, 혹은 오디오와 음성의 혼합신호를 나타내는 사운드 등의 미디어 신호를 의미할 수 있으나, 이하에서는 설명의 편의를 위하여 오디오 신호를 지칭하기로 한다.

[0020] 도 1을 참조하면, 변환부(130)는 시간 도메인의 오디오 신호를 주파수 도메인으로 변환하여 오디오 스펙트럼을 생성할 수 있다. 이때, 시간/주파수 도메인 변환은 MDCT(Modified Discrete Cosine Transform) 등과 같은 공지된 다양한 방법을 사용하여 수행할 수 있다. 일예를 들어, 시간 도메인의 오디오신호에 대한 MDCT는 하기 수학적 식 1에서와 같이 수행될 수 있다.

수학적 식 1

$$x_i = \sum_{j=0}^{2N-1} h_j s_j \cos[\pi(j+(N+1)/2)(i+1/2)/N], \quad i = 0, \dots, N-1$$

[0022] 여기서, N은 한 프레임에 포함된 샘플의 개수 즉 프레임 사이즈, h_j 는 적용된 윈도우, s_j 는 시간 도메인의 오디오 신호, x_i 는 MDCT 변환계수를 나타낸다. 한편, 수학적 식 1의 코사인 윈도우 대신 사인 윈도우 예를 들면,

$$h_j = \sin[\pi(j+1/2)/2N]$$

가 사용될 수도 있다.

[0023] 변환부(110)로부터 얻어지는 오디오 스펙트럼의 변환계수들, 예를 들면 MDCT 계수 x_i 는 엔벨로프 획득부(120)로 제공된다.

[0024] 엔벨로프 획득부(120)는 변환부(110)로부터 제공되는 변환계수들로부터, 소정의 서브밴드 단위로 엔벨로프 값을 획득할 수 있다. 서브밴드는 오디오 스펙트럼의 샘플들을 그룹핑한 단위로서, 임계대역을 반영하여 균일 혹은 비균일 길이를 가질 수 있다. 비균일한 경우, 한 프레임에 대하여 시작 샘플에서부터 마지막 샘플에 이르기까지 서브밴드에 포함되는 샘플의 개수가 점점 증가하도록 서브밴드를 설정할 수 있다. 또한 다중 비트율을 지원하는 경우, 서로 다른 비트율에서 대응하는 각 서브밴드에 포함되는 샘플의 개수가 동일해지도록 설정할 수 있다. 한 프레임에 포함되는 서브밴드의 개수 혹은 서브밴드에 포함되는 샘플의 개수는 미리 결정될 수 있다. 엔벨로프 값은 서브밴드에 포함되는 변환계수들의 평균 진폭, 평균 에너지, 파워 혹은 norm 값 등을 의미할 수 있다.

[0025] 각 서브밴드의 엔벨로프 값은 하기 수학적 식 2에 근거하여 산출할 수 있으나, 이에 한정되는 것은 아니다.

수학적 식 2

$$n = \sqrt{\frac{1}{w} \sum_{i=1}^w x_i^2}$$

[0027] 여기서, w는 서브밴드에 포함되는 변환계수의 개수 즉 서브밴드 사이즈, x_i 는 변환계수, n은 서브밴드의 엔벨로프 값을 나타낸다.

[0028] 엔벨로프 양자화부(130)는 각 서브밴드의 엔벨로프 값(n)에 대하여 최적화된 로그 스케일(logarithmic scale)로 양자화를 수행할 수 있다. 엔벨로프 양자화부(130)로부터 얻어지는 각 서브밴드에 대한 엔벨로프 값의 양자화 인덱스(n_q)는 예를 들면 하기 수학적 식 3에 의해 얻어질 수 있다.

수학식 3

$$n_q = \left\lfloor \frac{1}{r} \log_c n + \frac{b}{r} \right\rfloor$$

[0029]

[0030]

여기서, b는 라운딩 계수로서 최적화되기 이전의 초기값은 r/2이다. c는 로그 스케일의 베이스, r은 양자화 해상도를 각각 나타낸다.

[0031]

실시예에 따르면, 엔벨로프 양자화부(130)에서는 각 양자화 인덱스에 대응하는 양자화 영역내에서의 전체 양자화 오차가 최소가 되도록, 각 양자화 인덱스에 대응하는 양자화 영역의 좌측 및 우측 경계를 가변시킬 수 있다. 이를 위하여, 각 양자화 인덱스에 대응하는 양자화 영역의 좌측 및 우측 경계와 양자화 인덱스사이에서 각각 얻어지는 좌측 및 우측 양자화 오차가 동일해지도록 라운딩 계수(b)를 조정한다. 엔벨로프 양자화부(130)의 세부적인 동작에 대해서는 후술하기로 한다.

[0032]

한편, 각 서브밴드에 대한 엔벨로프 값의 양자화 인덱스(n_q)의 역양자화는 하기 수학식 4에 의해 수행될 수 있다.

수학식 4

$$\tilde{n} = c^{m_q}$$

[0033]

여기서, $\tilde{n}$ 는 각 서브밴드에 대하여 역양자화된 엔벨로프 값, r은 양자화 해상도, c는 로그 스케일의 베이스를 각각 나타낸다.

[0035]

엔벨로프 양자화부(130)에서 얻어지는 각 서브밴드에 대한 엔벨로프 값의 양자화 인덱스(n_q)는 엔벨로프 부호화부(140)로, 각 서브밴드에 대한 역양자화된 엔벨로프 값($\tilde{n}$)은 스펙트럼 정규화부(150)로 제공될 수 있다.

[0036]

한편, 도시되지 않았으나, 각 서브밴드 단위로 구해지는 엔벨로프 값은 정규화된 스펙트럼 즉, 정규화된 변환계수를 부호화하는데 필요로 하는 비트 할당에 사용될 수 있다. 이 경우, 각 서브밴드 단위로 양자화 및 무손실 부호화된 엔벨로프 값은 비트스트림에 포함되어 복호화장치로 제공될 수 있다. 각 서브밴드의 엔벨로프 값을 이용한 비트할당과 관련하여, 부호화장치와 복호화장치에서 동일한 프로세스를 이용할 수 있도록 역양자화된 엔벨로프 값을 사용할 수 있다.

[0037]

엔벨로프 값으로서 norm 값을 예로 들 경우, 각 서브밴드 단위로 norm 값을 이용하여 마스킹 임계치를 계산하고, 마스킹 임계치를 이용하여 지각적으로 필요한 비트수를 예측할 수 있다. 즉, 마스킹 임계치는 JND(Just Noticeable Distortion)에 해당하는 값으로서, 양자화 잡음이 마스킹 임계치보다 작을 경우 지각적인 잡음을 느낄 수 없다. 따라서, 지각적인 잡음을 느낄 수 없도록 하는데 필요한 최소 비트수를 마스킹 임계치를 이용하여 계산할 수 있다. 일실시예로, 각 서브밴드 단위로, norm 값과 마스킹 임계치와의 비를 이용하여 SMR(Signal-to-Mask Ratio)를 계산하고, SMR에 대하여 6.025 dB $\approx$ 1 비트의 관계를 이용하여 마스킹 임계치를 만족하는 비트수를 예측할 수 있다. 여기서, 예측된 비트수는 지각적인 잡음을 느낄 수 없도록 하는데 필요한 최소 비트수이지만, 압축 측면에서 보면 예측된 비트수 이상으로 사용할 필요가 없으므로 서브밴드 단위로 허용되는 최대 비트수(이하, 허용 비트수라 약함)로 간주될 수 있다. 이때, 각 서브밴드의 허용 비트수는 소수점 단위로 표현될 수 있으나, 이에 한정되는 것은 아니다.

[0038]

한편, 각 서브밴드 단위의 비트 할당은 norm 값을 이용하여 소수점 단위로 수행할 수 있으나, 이에 한정되는 것은 아니다. 이때, norm 값이 큰 서브밴드에서부터 순차적으로 비트를 할당하는데, 각 서브밴드의 norm 값에 대하여 각 서브밴드의 지각적 중요도에 따라서 가중치를 부여함으로써 지각적으로 중요한 서브밴드에 더 많은 비트가 할당되도록 조정할 수 있다. 지각적 중요도는 일례로 ITU-T G.719 에서와 같은 심리음향 가중을 통하여

결정할 수 있다.

[0039] 다시 도 1로 돌아가서, 엔벨로프 부호화부(140)는 엔벨로프 양자화부(130)로부터 제공되는 각 서브밴드에 대한 엔벨로프 값의 양자화 인덱스(n_q)에 대하여 양자화 델타값을 구하고, 양자화 델타값에 대하여 컨텍스트에 기반한 무손실 부호화를 수행하고, 그 결과를 비트스트림에 포함시켜 전송 및 저장에 사용할 수 있다. 여기서, 컨텍스트는 이전 서브밴드의 양자화 델타값을 사용할 수 있다. 엔벨로프 부호화부(140)의 세부적인 동작에 대해서는 후술하기로 한다.

[0040] 스펙트럼 정규화부(150)는 각 서브밴드의 역양자화된 엔벨로프 값 $\tilde{n} = C^{m_q}$ 을 이용하여, $y_i = x_i / \tilde{n}$ 에서와 같이 변환계수에 대하여 정규화를 수행함으로써, 각 서브밴드의 스펙트럼 평균 에너지가 1이 되도록 한다.

[0041] 스펙트럼 부호화부(160)는 정규화된 변환계수에 대하여 양자화 및 무손실 부호화를 수행하고, 그 결과를 비트스트림에 포함시켜 전송 및 저장에 사용할 수 있다. 이때, 스펙트럼 부호화부(160)는 각 서브밴드 단위로 엔벨로프 값에 근거하여 최종적으로 결정된 할당 비트수를 이용하여, 정규화된 변환계수를 양자화 및 무손실 부호화할 수 있다.

[0042] 정규화된 변환계수에 대한 무손실 부호화는 예를 들면 팩토리얼 펄스 코딩(Factorial Pulse Coding, 이하 FPC라 약함)을 사용할 수 있다. FPC는 단위 크기 펄스들(unit magnitude pulses)을 사용하여 정보 신호를 효율적으로 부호화하는 방법이다. FPC에 따르면, 정보 콘텐츠는 4가지 성분 즉, 난-제로 펄스 위치의 개수, 난-제로 펄스들의 위치, 난-제로 펄스들의 크기, 및 난-제로 펄스들의 부호로 나타낼 수 있다. 구체적으로, FPC는

$$m = \sum_{i=0}^{k-1} |\tilde{y}_i|$$
, (여기서 m 은 단위 크기 펄스들의 전체 개수)을 만족하면서 서브밴드의 원래의 벡터 y 와 FPC 벡터 $\tilde{y}$ 의 차이가 최소가 되는 MSE(mean square error) 기준에 근거하여 $\tilde{y} = \{\tilde{y}_1, \tilde{y}_2, \tilde{y}_3, \dots, \tilde{y}_{k-1}\}$ 에 대한 최적 해(solution)을 결정할 수 있다.

[0043] 최적 해는 하기 수학식 5에서와 같이, Lagrangian 함수를 이용하여 조건부 극값(conditional extreme value)을 찾음으로써 얻을 수 있다.

수학식 5

$$L = \sum (y_i - \tilde{y}_i)^2 + \lambda (\sum \tilde{y}_i - m)$$

$$\begin{cases} \frac{\partial L}{\partial \tilde{y}_i} = 2\tilde{y}_i - 2y_i + \lambda \tilde{y}_i = 0 \\ \frac{\partial L}{\partial \lambda} = \sum \tilde{y}_i - m = 0 \end{cases}$$

[0044]

$$\tilde{y}_i = \text{Round}\left(\frac{y_i m}{\sum y_i}\right)$$

[0045]

[0046] 여기서, L 은 Lagrangian 함수, m 은 서브밴드에 있는 단위 크기 펄스의 전체 갯수, λ 는 최적화계수인 Lagrange

multiplier로서, 주어진 함수의 최소값을 찾기 위한 컨트롤 파라미터, y_i 는 정규화된 변환계수, $\tilde{y}_i$ 는 위치 i 에서 요구되는 펄스의 최적 개수를 나타낸다.

[0047]

FPC를 이용하여 무손실 부호화를 수행하면, 각 서브밴드별로 얻어진 전체 세트의 $\tilde{y}_i$ 가 비트스트림에 포함되어 전송될 수 있다. 또한, 각 서브밴드에서 양자화 오차를 최소화시키고 평균 에너지의 얼라인먼트(alignment)를 수행하기 위한 최적 승수(optimum multiplier)도 비트스트림에 포함되어 전송될 수 있다. 최적 승수는 하기의 수학식 6에서와 같이 구할 수 있다.

수학식 6

$$D = \frac{\sum (y_i - G \tilde{y}_i)^2}{\sum y_i^2} \rightarrow 0$$

$$\frac{\partial D}{\partial G} = 0$$

$$G = \frac{\sum y_i \tilde{y}_i}{\sum \tilde{y}_i^2}$$

[0048]

여기서, D 는 양자화 오차, G 는 최적 승수를 나타낸다.

[0049]

도 2는 본 발명의 일실시예에 따른 디지털 신호 복호화장치의 구성을 나타낸 블록도이다.

[0050]

도 2에 도시된 디지털 신호 복호화장치(200)는 엔벨로프 복호화부(210), 엔벨로프 역양자화부(220), 스펙트럼 복호화부(230), 스펙트럼 역정규화부(240), 역변환부(250)를 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다. 여기서, 디지털 신호는 비디오, 이미지, 오디오 혹은 음성, 혹은 오디오와 음성의 혼합신호를 나타내는 사운드 등의 미디어 신호를 의미할 수 있으나, 이하에서는 도 1의 부호화장치에 대응되도록 오디오신호를 지칭하기로 한다.

[0051]

도 2를 참조하면, 엔벨로프 복호화부(210)는 통신 채널 혹은 네트워크를 통하여 비트스트림을 수신하고, 비트스트림에 포함된 각 서브밴드의 양자화 델타값을 무손실 복호화하여 각 서브밴드에 대한 엔벨로프 값의 양자화 인덱스(n_q)를 복원할 수 있다.

[0052]

엔벨로프 역양자화부(220)는 각 서브밴드에 대하여 복호화된 엔벨로프 값의 양자화 인덱스(n_q)에 대하여 역양자

화를 수행하여, 역양자화된 엔벨로프 값 $\tilde{n} = c^{m_q}$ 을 얻을 수 있다.

[0053]

스펙트럼 복호화부(230)는 수신된 비트스트림에 대하여 무손실 복호화 및 역양자화를 수행하여 정규화된 변환계수를 복원할 수 있다. 예를 들어, 부호화장치에서 FPC 를 사용한 경우 각 서브밴드에 대하여 전체 세트의

$\tilde{y}_i$ 을 무손실 복호화 및 역양자화할 수 있다. 이때, 각 서브밴드의 평균 에너지 얼라인먼트는 최적 승수(G)를 이용하여 하기 수학식 7에 의해 수행될 수 있다.

수학식 7

$$\tilde{y}_i = \tilde{y}_i G$$

[0055]

[0056]

스펙트럼 복호화부(230)는 도 1의 스펙트럼 부호화부(160)에서와 마찬가지로 각 서브밴드 단위로 엔벨로프 값에 근거하여 최종적으로 결정된 할당 비트수를 이용하여 무손실 복호화 및 역양자화를 수행할 수 있다.

[0057]

스펙트럼 역정규화부(240)는 엔벨로프 역양자화부(220)로부터 제공되는 역양자화된 엔벨로프 값을 이용하여, 스펙트럼 복호화부(210)로부터 제공되는 정규화된 변환계수에 대하여 역정규화(denormalization)를 수행할 수 있

다. 예를 들어, 부호화장치에서 FPC 를 사용한 경우 에너지 얼라인먼트가 수행된 $\tilde{y}_i$ 에 대하여 역양자화된

엔벨로프 값 $\tilde{n}$ 를 이용하여 $\tilde{x}_i = \tilde{y}_i \tilde{n}$ 에서와 같이 역정규화를 수행한다. 역정규화를 수행함으로써, 각 서브밴드에 대하여 원래의 스펙트럼 평균 에너지가 복원된다.

[0058]

역변환부(250)는 스펙트럼 역정규화부(240)로부터 제공되는 변환계수에 대하여 역변환을 수행하여 시간 도메인의 오디오 신호를 복원할 수 있다. 예를 들어, 상기 수학식 1에 대응되는 하기 수학식 8을 이용하여 스펙트럼

성분 $\tilde{x}_i$ 에 대하여 역변환을 수행하여 시간영역의 오디오신호 s_j 를 구할 수 있다.

수학식 8

$$s_j = \frac{1}{N} h_j \sum_{i=0}^{N-1} x_i \cos[\pi(j+(N+1)/2)(i+1/2)/N], \quad j = 0, \dots, 2N-1$$

[0059]

[0060]

이하에서는, 도 1에 도시된 엔벨로프 양자화부(130)의 동작에 대하여 좀 더 구체적으로 설명하기로 한다.

[0061]

엔벨로프 양자화부(130)에서 각 서브밴드의 엔벨로프 값에 대하여 베이스가 c인 로그 스케일로 양자화를 수행할

경우, 양자화 인덱스에 대응하는 양자화 영역의 경계(B_i)는 $B_i = c^{(S_i+S_{i+1})/2}$, 근사화 포인트

(approximating points, A_i) 즉, 양자화 인덱스는 $A_i = c^{S_i}$, 양자화 해상도(r)는 $r = S_i - S_{i-1}$, 양

자화 스텝사이즈는 $20 \lg A_i - 20 \lg A_{i-1} = 20r \lg c$ 와 같이 나타낼 수 있다. 이때, 각 서브밴드에 대한 엔벨로프 값(n)의 양자화 인덱스(n_q)는 상기 수학식 3에서와 같이 구해질 수 있다.

[0062]

그런데, 최적화되지 않은 선형 스케일의 경우, 양자화 인덱스(n_q)에 대응하는 양자화 영역의 좌측 및 우측 경계는 근사화 포인트로부터 서로 다른 거리만큼 떨어져서 존재한다. 이와 같은 차이로 인하여, 도 3a 및 도 4a에 도시된 바와 같이 양자화에 대한 SNR(signal-to-ratio) 척도 즉 양자화 오차가 근사화 포인트로부터 좌측 경계와 우측 경계에 대하여 서로 다른 값을 갖게 된다. 여기서, 도 3a는 양자화 해상도가 0.5, 양자화 스텝사이즈가 3.01 dB인 최적화되지 않은 로그 스케일(베이스는 2)의 양자화를 도시한 것이다. 양자화 영역의 좌측과 우측 경계에서 근사화 포인트로부터의 양자화 오차 SNR_L 및 SNR_R 은 14.46 dB와 15.96 dB로 서로 다를 수 있다. 도 4a은 양자화 해상도가 1, 양자화 스텝사이즈가 6.02 dB인 최적화되지 않은 로그 스케일(베이스는 2)의 양자화를 도시한 것이다. 양자화 영역의 좌측과 우측 경계에서 근사화 포인트로부터의 양자화 오차 SNR_L 및

SNR_R 은 7.65 dB와 10.66 dB 로 서로 다를 수 있다.

[0063] 일실시예에 따르면, 양자화 인덱스에 대응하는 양자화 영역의 경계를 가변시킴으로써, 각 양자화 인덱스에 대응하는 양자화 영역내의 전체 양자화 오차가 최소가 되도록 할 수 있다. 양자화 영역내 전체 양자화 오차는, 근사화 포인트로부터 양자화 영역의 좌측 및 우측 경계에서 얻어지는 양자화 오차가 동일할 경우 최소가 될 수 있다. 양자화 영역의 경계 쉬프트는 라운딩 계수 b 를 가변시킴으로써 얻어질 수 있다.

[0064] 양자화 인덱스에 대응하는 양자화 영역의 좌측 및 우측 경계에서 근사화 포인트에 대한 양자화 오차 SNR_L , SNR_R 은 각각 다음 수학적 식 9와 같이 나타낼 수 있다.

수학적 식 9

$$SNR_L = -20 \lg((c^{S_i} - c^{(S_i+S_{i-1})/2})/c^{(S_i+S_{i-1})/2})$$

$$SNR_R = -20 \lg((c^{(S_i+S_{i+1})/2} - c^{S_i})/c^{(S_i+S_{i+1})/2})$$

[0065]

[0066] 여기서, c 는 로그 스케일의 베이스, S_i 는 양자화 인덱스(i)에 대응하는 양자화 영역의 경계에 대한 지수(exponent)를 나타낸다.

[0067] 양자화 인덱스에 대응하는 양자화 영역의 좌측 및 우측 경계에 대한 지수 쉬프트는 파라미터 b_L 및 b_R 을 통하여 하기 수학적 식 10과 같이 나타낼 수 있다.

수학적 식 10

$$b_L = S_i - (S_i + S_{i-1})/2$$

$$b_R = (S_i + S_{i+1})/2 - S_i$$

[0068]

[0069] 여기서, S_i 는 양자화 인덱스(i)에 대응하는 양자화 영역의 경계에 대한 지수, b_L 및 b_R 는 양자화 영역의 좌측 및 우측 경계에서 근사화 포인트에 대한 지수 쉬프트를 각각 나타낸다.

[0070] 양자화 영역의 좌측 및 우측 경계에서 근사화 포인트에 대한 지수 쉬프트의 합은 양자화 해상도와 동일하며, 따라서 하기의 수학적 식 11과 같이 나타낼 수 있다.

수학적 식 11

$$b_L + b_R = r$$

[0071]

[0072] 한편, 양자화의 일반적인 특성에 근거하여, 라운딩 계수는 양자화 인덱스에 대응하는 양자화 영역의 좌측 경계에서 근사화 포인트에 대한 지수 쉬프트와 동일하다. 따라서, 상기 수학적 식 9는 다음 수학적 식 12와 같이 나타낼 수 있다.

수학식 12

$$SNR_L = -20 \lg((c^{S_i} - c^{S_i+b_L})/c^{S_i+b_L}) = -20 \lg(c^{b_L} - 1)$$

$$SNR_R = -20 \lg((c^{S_i+b_R} - c^{S_i})/c^{S_i+b_R}) = -20 \lg(1 - c^{-r+b_L})$$

[0073]

[0074]

양자화 인덱스에 대응하는 양자화 영역의 좌측 및 우측 경계에서 근사화 포인트에 대한 SNR을 동일하게 함으로써, 하기 수학식 13에서와 같이 파라미터 b_L 을 결정할 수 있다.

수학식 13

$$-20 \lg(c^{b_L} - 1) = -20 \lg(1 - c^{-r+b_L})$$

$$c = c^{b_L} + c^{-r+b_L} = c^{b_L}(1 + c^{-r})$$

[0075]

[0076]

따라서, 라운딩 계수(b_L)는 하기 수학식 14와 같이 나타낼 수 있다.

수학식 14

$$b_L = 1 - \log_c(1 + c^{-r})$$

[0077]

[0078]

도 3b는 양자화간격이 3.01 dB이고, 양자화 해상도가 0.5인 최적화된 로그 스케일(베이스는 2)의 양자화를 도시한 것이다. 양자화 영역의 좌측과 우측 경계에서 근사화 포인트로부터의 양자화 오차 SNR_L 및 SNR_R 은 15.31 dB로 동일함을 알 수 있다. 도 4b는 양자화간격이 6.02 dB이고, 양자화 해상도가 1.0인 최적화된 로그 스케일(베이스는 2)의 양자화를 도시한 것이다. 양자화 영역의 좌측과 우측 경계에서 근사화 포인트로부터의 양자화 오차 SNR_L 및 SNR_R 은 9.54 dB로 동일함을 알 수 있다.

[0079]

라운딩 계수 $b=b_L$ 은 양자화 인덱스에 대응하는 양자화 영역의 좌측 및 우측 경계에서부터 근사화 포인트까지의 지수에 대한 거리를 결정한다. 따라서, 일실시에 따른 양자화는 하기 수학식 15에서와 같이 수행될 수 있다.

수학식 15

$$n_q = \lfloor \frac{1}{r} \log_c n + \frac{b_L}{r} \rfloor$$

[0080]

[0081]

베이스 2인 로그 스케일에 의하여 양자화를 수행한 실험결과는 도 5a 및 도 5b에 도시되어 있다. 정보 이론에 따르면, 비트율-왜곡 함수 $H(D)$ 은 다양한 양자화 방법을 비교 분석할 수 있는 기준으로 사용될 수 있다. 양자화 인덱스 세트의 엔트로피는 비트율로 간주할 수 있고, 차원 b/s 를 가지며, dB 스케일의 SNR은 왜곡 척도로 간주할 수 있다.

[0082]

도 5a는 정상 분포에 대하여 양자화를 수행한 비교 그래프로서, 실선은 최적화되지 않은 로그 스케일의 양자화에 대한 비트율-왜곡 함수를, 점선은 최적화된 로그 스케일의 양자화에 대한 비트율-왜곡 함수를 나타낸다. 도 5b는 균일 분포에 대하여 양자화를 수행한 비교 그래프로서, 실선은 최적화되지 않은 로그 스케일의 양자화에 대한 비트율-왜곡 함수를, 점선은 최적화된 로그 스케일의 양자화에 대한 비트율-왜곡 함수를 나타낸다. 정상

및 균일 분포의 샘플들은 대응하는 분포법칙, 제로 기대값 및 단일 분산에 따라서 랜덤 갯수의 센서를 이용하여 생성된다. 비트율-왜곡 함수 $H(D)$ 는 여러가지 양자화 해상도에 대하여 산출될 수 있다. 도 5a 및 도 5b에 도시된 바와 같이, 점선은 실선 아래에 위치하며, 이는 최적화된 로그 스케일의 양자화가 최적화되지 않은 로그 스케일의 양자화에 비해 그 성능이 우수함을 의미한다.

[0083] 즉, 최적화된 로그 스케일의 양자화에 따르면, 동일한 비트율에 대하여 더 적은 양자화 오차로 양자화를 수행할 수 있거나, 동일한 비트율에 대하여 동일한 양자화 오차로 더 적은 비트수를 사용하여 양자화를 수행할 수 있다. 그 실험 결과는 다음 표 1 및 표 2에 도시되어 있으며, 표 1은 최적화되지 않은 로그 스케일의 양자화를, 표 2는 최적화된 로그 스케일의 양자화를 각각 나타낸다.

표 1

[0084] 양자화 해상도 (r)	2.0	1.0	0.5
라운드 계수 (b/r)	0.5	0.5	0.5
정상 분포			
비트레이트 (H), b/s	1.6179	2.5440	3.5059
양자화 오차 (D), dB	6.6442	13.8439	19.9534
균일 분포			
비트레이트 (H), b/s	1.6080	2.3227	3.0830
양자화 오차 (D), dB	6.6470	12.5018	19.3640

표 2

[0085] 양자화 해상도 (r)	2.0	1.0	0.5
라운드 계수 (b/r)	0.3390	0.4150	0.4569
정상 분포			
비트레이트 (H), b/s	1.6069	2.5446	3.5059
양자화 오차 (D), dB	8.2404	14.2284	20.0495
균일 분포			
비트레이트 (H), b/s	1.6345	2.3016	3.0449
양자화 오차 (D), dB	7.9208	12.8954	19.4922

[0086] 표 1 및 표 2에 따르면, 특성값 SNR은 양자화 해상도 0.5에서는 0.1 dB 개선되었고, 양자화 해상도 1.0에서는 0.45 dB 개선되었고, 양자화 해상도 2.0에서는 1.5 dB 개선되었음을 알 수 있다.

[0087] 일실시에에 따른 양자화 방법은 양자화 인덱스의 탐색 테이블만 라운드 계수에 따라서 갱신시키면 되기 때문에 복잡도를 증가시키지 않는다.

[0088] 다음, 도 1에 도시된 엔벨로프 복호화부(140)의 동작에 대하여 좀 더 구체적으로 설명하기로 한다.

[0089] 엔벨로프 값의 컨텍스트 기반 부호화는 델타 부호화(delta-coding)을 사용한다. 현재 서브밴드와 이전 서브밴드간의 엔벨로프 값에 대한 양자화 델타값은 하기 수학적 식 16에서와 같이 나타낼 수 있다.

수학적 식 16

$$d(i) = n_q(i+1) - n_q(i)$$

[0090]

[0091] 여기서, $d(i)$ 는 서브밴드(i+1)에 대한 양자화 델타값, $n_q(i)$ 는 서브밴드(i)에 대한 엔벨로프 값의 양자화 인덱스, $n_q(i+1)$ 는 서브밴드(i+1)에 대한 엔벨로프 값의 양자화 인덱스를 나타낸다.

[0092] 각 서브밴드에 대한 양자화 델타값 $d(i)$ 는 범위 [-15, 16]로 제한되며, 하기에서와 같이 우선 음수인 양자화 델타값을 조정한 다음, 양수인 양자화 델타값을 조정한다.

[0093] 먼저, 상기 수학적 식 16을 이용하여 양자화 델타값 $d(i)$ 을 고주파수 서브밴드에서부터 저주파수 서브밴드의 순서

로 구한다. 이때, $d(i) < -15$ 이면, $n_q(i)=n_q(i+1) + 15$ (여기서 $i=42, \dots, 0$)로 조정한다.

[0094] 다음, 상기 수학식 16을 이용하여 양자화 델타값 $d(i)$ 를 저주파수 서브밴드에서부터 고주파수 서브밴드의 순서로 구한다. 이때, $d(i) > 16$ 이면, $d(i) = 16$, $n_q(i+1)=n_q(i) + 16$ (여기서 $i=0, \dots, 42$)로 조정한다.

[0095] 이후, 구해진 모든 양자화 델타값 $d(i)$ 에 옵셋 15를 더하여, 최종적으로 범위 $[0, 31]$ 의 양자화 델타값을 생성한다.

[0096] 상기 수학식 16에 따르면, 한 프레임에 대하여 N 개의 서브밴드가 존재하는 경우, $n_q(0)$, $d(0)$, $d(1)$, $d(2), \dots, d(N-2)$ 가 구해진다. 현재 서브밴드의 양자화 델타값은 컨텍스트 모델(context model)을 사용하여 부호화되는데, 일실시에에 따르면 이전 서브밴드에 대한 양자화 델타값을 컨텍스트로 사용할 수 있다. 첫번째 서브밴드에 대한 $n_q(0)$ 은 $[0, 31]$ 의 범위에 존재하므로 5비트를 사용하여 그대로 무손실 부호화한다. 한편, 첫번째 서브밴드에 대한 $n_q(0)$ 이 $d(0)$ 의 컨텍스트로 사용될 경우에는, $n_q(0)$ 로부터 소정의 기준값을 이용하여 얻어지는 값을 사용할 수 있다. 즉, $d(i)$ 에 대한 허프만 부호화시에는 $d(i-1)$ 을 컨텍스트로 사용하고, $d(0)$ 에 대한 허프만 부호화시에는 $n_q(0)$ -기준값을 컨텍스트로 사용할 수 있다. 여기서, 소정의 기준값의 예로는 소정의 상수를 사용할 수 있으며, 미리 시뮬레이션을 통하여 혹은 실험적으로 최적 값으로 설정될 수 있다. 기준값은 비트 스트림에 포함되어 전송되거나, 부호화장치와 복호화장치에 미리 제공될 수 있다.

[0097] 일실시에에 따르면, 엔벨로프 부호화부(140)는 컨텍스트로 사용되는 이전 서브밴드의 양자화 델타값의 범위를 복수개의 그룹으로 나누고, 각 그룹별로 미리 정해진 허프만 테이블을 기준으로 하여 현재 서브밴드의 양자화 델타값에 대한 허프만 부호화를 수행할 수 있다. 여기서, 허프만 테이블은 예를 들어 대형 데이터베이스를 이용한 트레이닝 프로세스를 통하여 생성할 수 있으며, 소정의 기준에 의거하여 데이터를 수집하고, 수집된 데이터에 근거하여 허프만 테이블을 생성할 수 있다. 일실시에에 따르면, 이전 서브밴드의 양자화 델타값의 범위에 의거하여 현재 서브밴드의 양자화 델타값의 빈도수에 대한 데이터를 수집하여 각 그룹별로 허프만 테이블을 생성할 수 있다.

[0098] 이전 서브밴드의 양자화 델타값을 컨텍스트로 사용하여 얻어진 현재 서브밴드의 양자화 델타값의 확률 분포에 대한 분석결과를 이용하여, 다양한 분포 모델을 선택할 수 있고, 따라서 유사한 분포모델을 갖는 양자화 레벨의 그룹핑이 수행될 수 있다. 각 그룹의 파라미터는 다음 표 3에 도시되어 있다.

표 3

그룹 번호	차이값의 하한	차이값의 상한
#1	0	12
#2	13	17
#3	18	31

[0100] 한편, 3개 그룹에서의 확률 분포는 도 6에 도시되어 있다. 그룹 #1 및 그룹 #3의 확률 분포가 유사하며, x축에 의해 실질적으로 반전(혹은 플립)됨을 알 수 있다. 이는 부호화 효율에 대한 손실없이, 두개의 그룹 #1 및 #3에 대해서는 동일한 확률 모델을 사용해도 무방함을 의미한다. 즉, 그룹 #1은 그룹 #3와 동일한 허프만 테이블을 사용할 수 있다. 이에 따르면, 그룹 #2에 대한 허프만 테이블 1과, 그룹 #1 및 그룹 #3가 공유하는 허프만 테이블 2가 사용될 수 있다. 이때, 그룹 #1에 대한 코드의 인덱스는 그룹 #3에 대하여 반대로 표현하면 된다. 즉, 컨텍스트인 이전 서브밴드의 양자화 델타값에 의하여 현재 서브밴드의 양자화 델타값에 대한 허프만 테이블이 그룹 #1으로 결정된 경우, 부호화단에서 현재 서브밴드의 양자화 델타값($d(i)$)는 반전처리 과정 즉, $d'(i)=A-d(i)$ 의 값으로 변경되어 그룹 #3의 허프만 테이블을 참조하여 허프만 부호화를 수행할 수 있다. 한편, 복호화단에서는 그룹 #3의 허프만 테이블을 참조하여 허프만 복호화를 수행한 다음, $d'(i)$ 는 $d(i)=A-d'(i)$ 의 변환과정을 거쳐 최종 $d(i)$ 값을 추출하게 된다. 여기서, A 값은 그룹 #1과 그룹 #3의 확률 분포가 대칭되도록 만드는 값으로 설정될 수 있다. A 값은 부호화 및 복호화 과정에서 추출되는 것이 아니고, 미리 사전에 최적 값으로 설정될 있다. 한편, 그룹 #3의 허프만 테이블 대신에 그룹 #1의 허프만 테이블을 활용하고, 그룹 #3에서 양자화 델타값을 변경시켜서 수행할 수도 있다. 일실시에에 따르면, $d(i)$ 가 범위 $[0, 31]$ 의 값을 가지는 경우, A 값은 31을 사용할 수 있다.

[0101] 도 7은 도 1의 엔벨로프 부호화부(140)에서의 컨텍스트 기반 허프만 부호화 동작을 설명하는 도면으로서, 3개

그룹의 양자화 델타값의 확률분포에 의해 결정된 2가지 허프만 테이블을 이용한다. 여기서, 현재 서브밴드의 양자화 델타값($d(i)$)을 허프만 부호화함에 있어서, 이전 서브밴드의 양자화 델타값($d(i-1)$)을 컨텍스트로 활용하며, 그룹 #2에 대한 허프만 테이블 1과 그룹 #3에 대한 허프만 테이블 2가 사용되는 것을 예로 들기로 한다.

- [0102] 도 7을 참조하면, 710 단계에서는 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #2에 속하는지를 판단한다.
- [0103] 720 단계에서는, 710 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #2에 속하는 경우, 허프만 테이블 1로부터 현재 서브밴드의 양자화 델타값($d(i)$)에 대한 코드를 선택한다.
- [0104] 730 단계에서는, 710 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #2에 속하지 않는 경우, 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #1에 속하는지를 판단한다.
- [0105] 740 단계에서는, 730 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #1에 속하지 않는 경우 즉, 그룹 #3에 속하는 경우, 허프만 테이블 2로부터 현재 서브밴드의 양자화 델타값($d(i)$)에 대한 코드를 선택한다.
- [0106] 750 단계에서는, 730 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #1에 속하는 경우, 현재 서브밴드의 양자화 델타값($d(i)$)을 반전처리하고, 허프만 테이블 2로부터 반전처리된 현재 서브밴드의 양자화 델타값($d'(i)$)에 대한 코드를 선택한다.
- [0107] 760 단계에서는 720, 740 혹은 750 단계에서 선택된 코드를 이용하여 현재 서브밴드의 양자화 델타값($d(i)$)에 대하여 허프만 부호화를 수행한다.
- [0108] 도 8은 도 2의 엔벨로프 복호화부(210)에서의 컨텍스트 기반 허프만 복호화 동작을 설명하는 도면으로서, 도 7에서와 마찬가지로 3개 그룹의 양자화 델타값의 확률분포에 의해 결정된 2가지 허프만 테이블을 이용한다. 여기서, 현재 서브밴드의 양자화 델타값($d(i)$)을 허프만 복호화함에 있어서, 이전 서브밴드의 양자화 델타값($d(i-1)$)을 컨텍스트로 활용하며, 그룹 #2에 대한 허프만 테이블 1과 그룹 #3에 대한 허프만 테이블 2가 사용되는 것을 예로 들기로 한다.
- [0109] 도 8을 참조하면, 810 단계에서는 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #2에 속하는지를 판단한다.
- [0110] 820 단계에서는, 810 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #2에 속하는 경우, 허프만 테이블 1로부터 현재 서브밴드의 양자화 델타값($d(i)$)에 대한 코드를 선택한다.
- [0111] 830 단계에서는, 810 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #2에 속하지 않는 경우, 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #1에 속하는지를 판단한다.
- [0112] 840 단계에서는, 830 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #1에 속하지 않는 경우 즉, 그룹 #3에 속하는 경우, 허프만 테이블 2로부터 현재 서브밴드의 양자화 델타값($d(i)$)에 대한 코드를 선택한다.
- [0113] 850 단계에서는, 830 단계에서의 판단결과 이전 서브밴드의 양자화 델타값($d(i-1)$)이 그룹 #1에 속하는 경우, 현재 서브밴드의 양자화 델타값($d(i)$)을 반전처리하고, 허프만 테이블 2로부터 반전처리된 현재 서브밴드의 양자화 델타값($d'(i)$)에 대한 코드를 선택한다.
- [0114] 860 단계에서는 820, 840 혹은 850 단계에서 선택된 코드를 이용하여 현재 서브밴드의 양자화 델타값($d(i)$)에 대하여 허프만 복호화를 수행한다.
- [0115] 프레임별로 비트 코스트의 차이 분석은 다음 표 4에 도시되어 있다. 이에 따르면 상기 실시예에 따른 부호화 효율은 원래의 허프만 부호화 알고리즘에 비하여 평균 9% 증가하였음을 알 수 있다.

표 4

[0116]	알고리즘	비트율, kbps	이득, %
	허프만 엔코딩	6.25	-
	컨텍스트+허프만 엔코딩	5.7	9

- [0117] 도 9는 본 발명의 일실시예에 따른 부호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다. 도 9에

도시된 멀티미디어 기기(900)는 통신부(910)와 부호화모듈(930)을 포함할 수 있다. 또한, 부호화 결과 얻어지는 오디오 비트스트림의 용도에 따라서, 오디오 비트스트림을 저장하는 저장부(950)을 더 포함할 수 있다. 또한, 멀티미디어 기기(900)는 마이크로폰(970)을 더 포함할 수 있다. 즉, 저장부(950)와 마이크로폰(970)은 옵션으로 구비될 수 있다. 한편, 도 9에 도시된 멀티미디어 기기(900)는 임의의 복호화모듈(미도시), 예를 들면 일반적인 복호화 기능을 수행하는 복호화모듈 혹은 본 발명의 일실시예에 따른 복호화모듈을 더 포함할 수 있다. 여기서, 부호화모듈(930)은 멀티미디어 기기(900)에 구비되는 다른 구성요소(미도시)와 함께 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.

[0118] 도 9를 참조하면, 통신부(910)는 외부로부터 제공되는 오디오와 부호화된비트스트림 중 적어도 하나를 수신하거나, 복원된 오디오와 부호화모듈(930)의 부호화결과 얻어지는 오디오 비트스트림 중 적어도 하나를 송신할 수 있다.

[0119] 통신부(910)는 무선 인터넷, 무선 인트라넷, 무선 전화망, 무선 랜(LAN), 와이파이(Wi-Fi), 와이파이 다이렉트(WFD, Wi-Fi Direct), 3G(Generation), 4G(4 Generation), 블루투스(Bluetooth), 적외선 통신(IrDA, Infrared Data Association), RFID(Radio Frequency Identification), UWB(Ultra WideBand), 지그비(Zigbee), NFC(Near Field Communication)와 같은 무선 네트워크 또는 유선 전화망, 유선 인터넷과 같은 유선 네트워크를 통해 외부의 멀티미디어 기기와 데이터를 송수신할 수 있도록 구성된다.

[0120] 부호화모듈(930)은 일실시예에 따르면, 통신부(910) 혹은 마이크로폰(970)을 통하여 제공되는 시간 도메인의 오디오 신호를 주파수 도메인의 오디오 스펙트럼으로 변환하고, 오디오 스펙트럼에 대하여, 소정의 서브밴드 단위로 엔벨로프를 획득하고, 서브밴드 단위로, 상기 엔벨로프에 대하여 양자화를 수행하고, 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 부호화를 수행하여 비트스트림을 생성할 수 있다.

[0121] 부호화모듈(930)은 다른 실시예에 따르면, 엔벨로프의 양자화시, 소정의 양자화 인덱스에 대응하는 양자화 영역에서의 전체 양자화오차가 최소가 되도록 상기 양자화 영역의 경계를 조정하고, 이로부터 갱신되는 양자화 테이블을 이용하여 양자화를 수행할 수 있다.

[0122] 저장부(950)는 부호화 모듈(930)에서 생성되는 부호화된 비트스트림을 저장할 수 있다. 한편, 저장부(950)는 멀티미디어 기기(900)의 운용에 필요한 다양한 프로그램을 저장할 수 있다.

[0123] 마이크로폰(970)은 사용자 혹은 외부의 오디오신호를 부호화모듈(930)로 제공할 수 있다.

[0124] 도 10은 본 발명의 일실시예에 따른 복호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.

[0125] 도 10에 도시된 멀티미디어 기기(1000)는 통신부(1010)와 복호화모듈(1030)을 포함할 수 있다. 또한, 복호화 결과 얻어지는 복원된 오디오신호의 용도에 따라서, 복원된 오디오신호를 저장하는 저장부(1050)을 더 포함할 수 있다. 또한, 멀티미디어 기기(1000)는 스피커(1070)를 더 포함할 수 있다. 즉, 저장부(1050)와 스피커(1070)는 옵션으로 구비될 수 있다. 한편, 도 10에 도시된 멀티미디어 기기(1000)는 임의의 부호화모듈(미도시), 예를 들면 일반적인 부호화 기능을 수행하는 부호화모듈 혹은 본 발명의 일실시예에 따른 부호화모듈을 더 포함할 수 있다. 여기서, 복호화모듈(1030)은 멀티미디어 기기(1000)에 구비되는 다른 구성요소(미도시)와 함께 일체화되어 적어도 하나의 이상의 프로세서(미도시)로 구현될 수 있다.

[0126] 도 10을 참조하면, 통신부(1010)는 외부로부터 제공되는 부호화된 비트스트림과 오디오 신호 중 적어도 하나를 수신하거나 복호화 모듈(1030)의 복호화결과 얻어지는 복원된 오디오 신호와 부호화결과 얻어지는 오디오 비트스트림 중 적어도 하나를 송신할 수 있다. 한편, 통신부(1010)는 도 9의 통신부(910)와 실질적으로 유사하게 구현될 수 있다.

[0127] 복호화 모듈(1030)은 일실시예에 따르면, 통신부(1010)를 통하여 제공되는 비트스트림을 수신하고, 비트스트림으로부터 인접한 서브밴드에 대하여 양자화된 엔벨로프간의 차이값을 구하고, 이전 서브밴드의 차이값을 컨텍스트로 사용하여 현재 서브밴드의 차이값에 대하여 무손실 복호화를 수행하고, 무손실 복호화결과 복원된 현재 서브밴드의 차이값으로부터 서브밴드 단위로 상기 양자화된 엔벨로프를 구하여 역양자화를 수행할 수 있다.

[0128] 저장부(1050)는 복호화 모듈(1030)에서 생성되는 복원된 오디오신호를 저장할 수 있다. 한편, 저장부(1050)는 멀티미디어 기기(1000)의 운용에 필요한 다양한 프로그램을 저장할 수 있다.

[0129] 스피커(1070)는 복호화 모듈(1030)에서 생성되는 복원된 오디오신호를 외부로 출력할 수 있다.

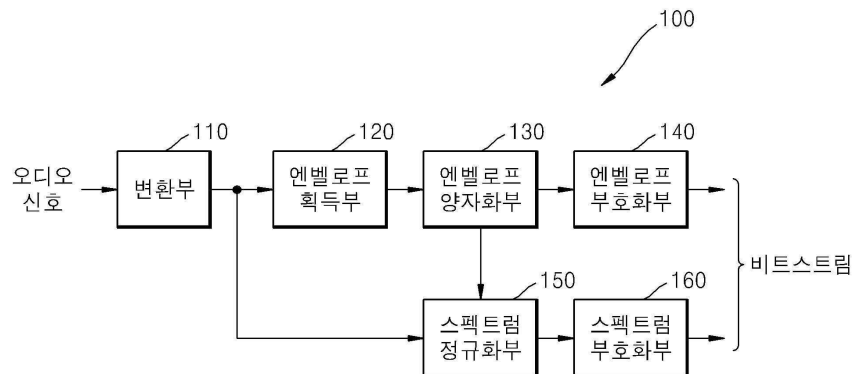
- [0130] 도 11은 본 발명의 일실시예에 따른 부호화모듈과 복호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.
- [0131] 도 11에 도시된 멀티미디어 기기(1100)는 통신부(1110), 부호화모듈(1120)과 복호화모듈(1130)을 포함할 수 있다. 또한, 부호화 결과 얻어지는 오디오 비트스트림 혹은 복호화 결과 얻어지는 복원된 오디오신호의 용도에 따라서, 오디오 비트스트림 혹은 복원된 오디오신호를 저장하는 저장부(1140)을 더 포함할 수 있다. 또한, 멀티미디어 기기(1100)는 마이크로폰(1150) 혹은 스피커(1160)를 더 포함할 수 있다. 여기서, 부호화모듈(1120)과 복호화모듈(1130)은 멀티미디어 기기(1100)에 구비되는 다른 구성요소(미도시)와 함께 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.
- [0132] 도 11에 도시된 각 구성요소는 도 9에 도시된 멀티미디어 기기(900)의 구성요소 혹은 도 10에 도시된 멀티미디어 기기(1000)의 구성요소와 중복되므로, 그 상세한 설명은 생략하기로 한다.
- [0133] 도 9 내지 도 11에 도시된 멀티미디어 기기(900, 1000, 1100)에는, 전화, 모바일 폰 등을 포함하는 음성통신 전용단말, TV, MP3 플레이어 등을 포함하는 방송 혹은 음악 전용장치, 혹은 음성통신 전용단말과 방송 혹은 음악 전용장치의 융합 단말장치가 포함될 수 있으나, 이에 한정되는 것은 아니다. 또한, 멀티미디어 기기(900, 1000, 1100)는 클라이언트, 서버 혹은 클라이언트와 서버 사이에 배치되는 변환기로서 사용될 수 있다.
- [0134] 한편, 멀티미디어 기기(900, 1000, 1100)가 예를 들어 모바일 폰인 경우, 도시되지 않았지만 키패드 등과 같은 유저 입력부, 유저 인터페이스 혹은 모바일 폰에서 처리되는 정보를 디스플레이하는 디스플레이부, 모바일 폰의 전반적인 기능을 제어하는 프로세서를 더 포함할 수 있다. 또한, 모바일 폰은 촬상 기능을 갖는 카메라부와 모바일 폰에서 필요로 하는 기능을 수행하는 적어도 하나 이상의 구성요소를 더 포함할 수 있다.
- [0135] 한편, 멀티미디어 기기(900, 1000, 1100)가 예를 들어 TV인 경우, 도시되지 않았지만 키패드 등과 같은 유저 입력부, 수신된 방송정보를 디스플레이하는 디스플레이부, TV의 전반적인 기능을 제어하는 프로세서를 더 포함할 수 있다. 또한, TV는 TV에서 필요로 하는 기능을 수행하는 적어도 하나 이상의 구성요소를 더 포함할 수 있다.
- [0136] 상기 실시예들에 따른 방법은 컴퓨터에서 실행될 수 있는 프로그램으로 작성가능하고, 컴퓨터로 읽을 수 있는 기록매체를 이용하여 상기 프로그램을 동작시키는 범용 디지털 컴퓨터에서 구현될 수 있다. 또한, 상술한 본 발명의 실시예들에서 사용될 수 있는 데이터 구조, 프로그램 명령, 혹은 데이터 파일은 컴퓨터로 읽을 수 있는 기록매체에 다양한 수단을 통하여 기록될 수 있다. 컴퓨터로 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 저장 장치를 포함할 수 있다. 컴퓨터로 읽을 수 있는 기록매체의 예로는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함될 수 있다. 또한, 컴퓨터로 읽을 수 있는 기록매체는 프로그램 명령, 데이터 구조 등을 지정하는 신호를 전송하는 전송 매체일 수도 있다. 프로그램 명령의 예로는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함할 수 있다.
- [0137] 이상과 같이 본 발명의 일실시예는 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명의 일실시예는 상기 설명된 실시예에 한정되는 것은 아니며, 이는 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 스코프는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 이의 균등 또는 등가적 변형 모두는 본 발명 기술적 사상의 범주에 속한다고 할 것이다.

부호의 설명

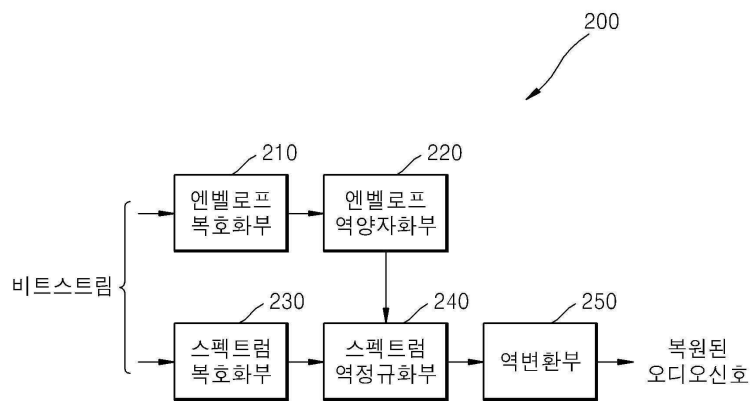
- [0138] 110: 변환부 120: 엔벨로프 획득부
- 130: 엔벨로프 양자화부 140: 엔벨로프 부호화부
- 150: 스펙트럼 정규화부 160: 스펙트럼 부호화부
- 210: 엔벨로프 복호화부 220: 엔벨로프 역양자화부
- 230: 스펙트럼 복호화부 240: 스펙트럼 역정규화부
- 250: 역변환부

도면

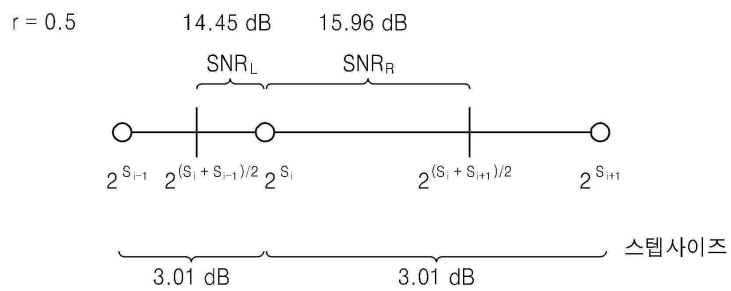
도면1



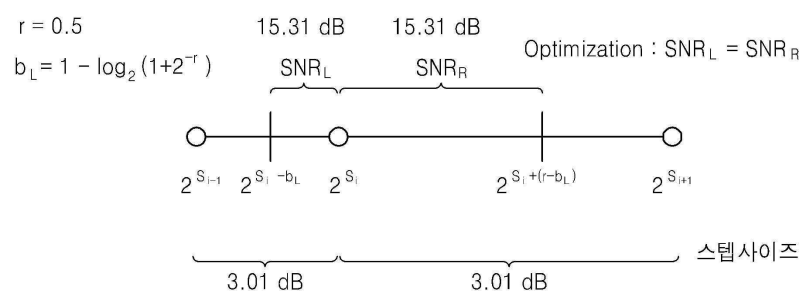
도면2



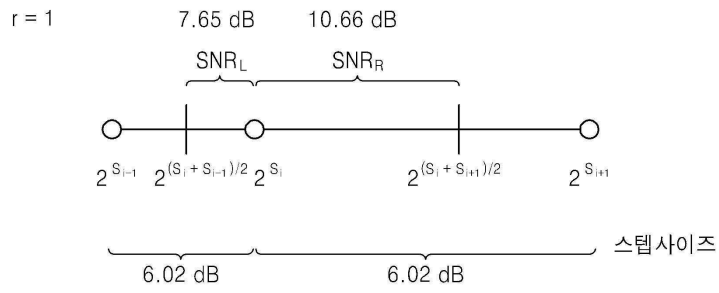
도면3a



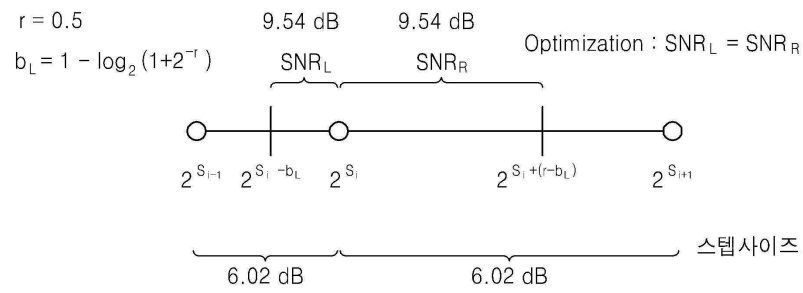
도면3b



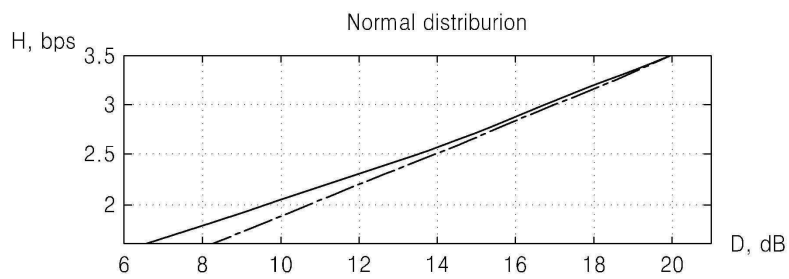
도면4a



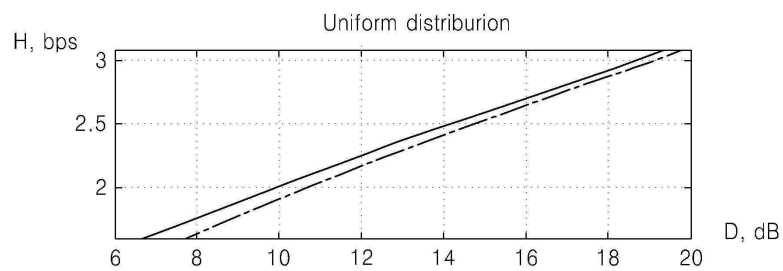
도면4b



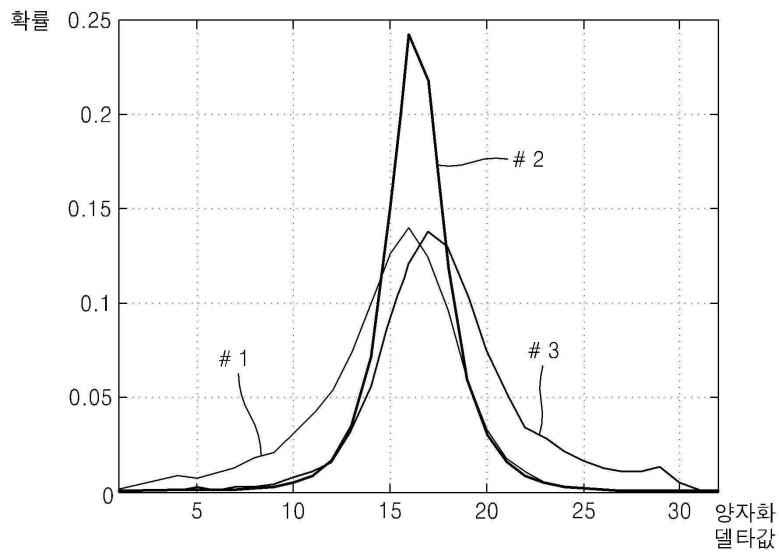
도면5a



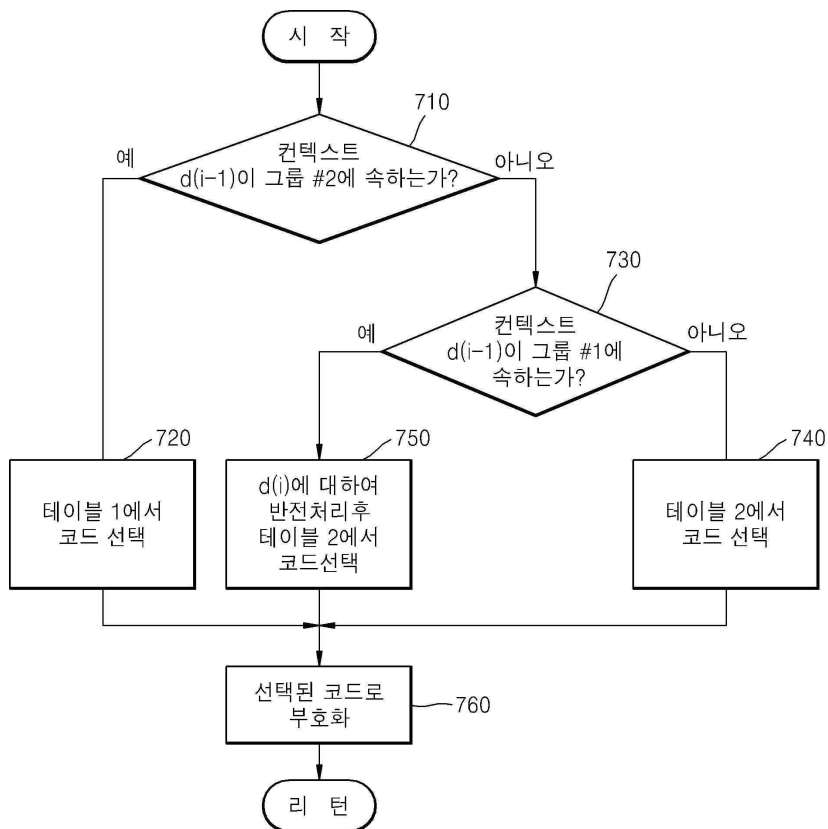
도면5b



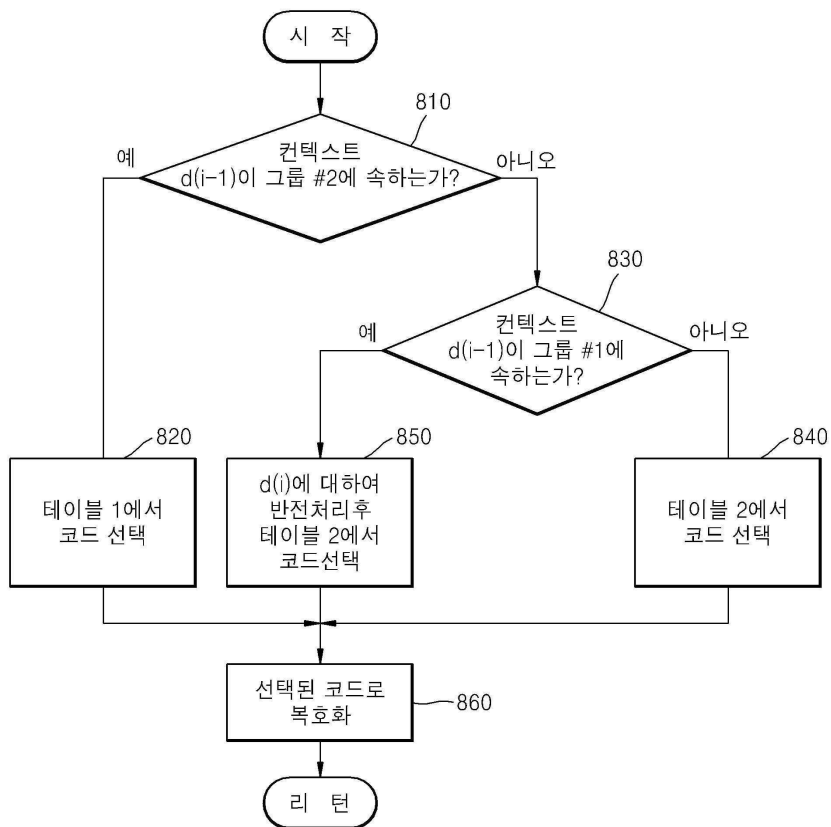
도면6



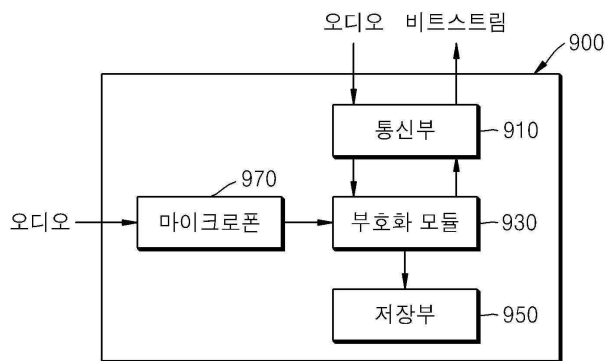
도면7



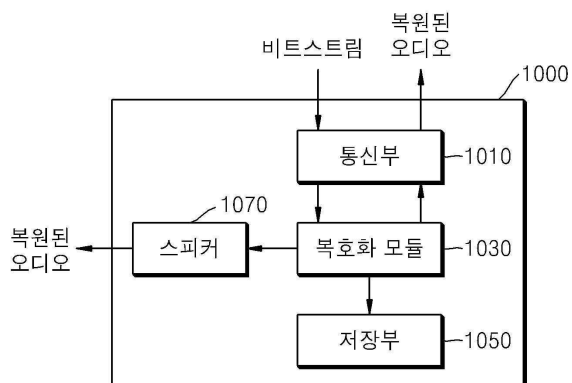
도면8



도면9



도면10



도면11

