(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2020년06월16일
(11) 등록번호 10-2123770
(24) 등록일자 2020년06월10일

(51) 국제특허분류(Int. Cl.)
*G10L 19/028* (2013.01) *G10L 19/038* (2013.01)
(52) CPC특허분류
*G10L 19/028* (2013.01)
*G10L 19/038* (2013.01)
(21) 출원번호 10-2019-7019105(분할)
(22) 출원일자(국제) 2012년10월30일
심사청구일자 2019년07월02일
(85) 번역문제출일자 2019년07월02일
(65) 공개번호 10-2019-0084131
(43) 공개일자 2019년07월15일
(62) 원출원 특허 10-2014-7030223
원출원일자(국제) 2012년10월30일
심사청구일자 2017년04월10일
(86) 국제출원번호 PCT/SE2012/051177
(87) 국제공개번호 WO 2013/147666
국제공개일자 2013년10월03일
(30) 우선권주장
61/617,216 2012년03월29일 미국(US)
(56) 선행기술조사문헌
US20070238415 A1
US20120029923 A1*
US20120046955 A1*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
**텔레폰악티에볼라겟엘엠에릭슨(펍)**
스웨덴왕국 스톡홀름 에스이-164 83
(72) 발명자
**그란카로프 볼로디야**
스웨덴 에스-171 67 솔라 앙크담스가탄 29
**토프트고르드 토마스**
스웨덴 에스-754 35 웁살라 아펠가탄 20에이
*(뒷면에 계속)*
(74) 대리인
**서장찬, 박병석**

전체 청구항 수 : 총 16 항 심사관 : 이남숙

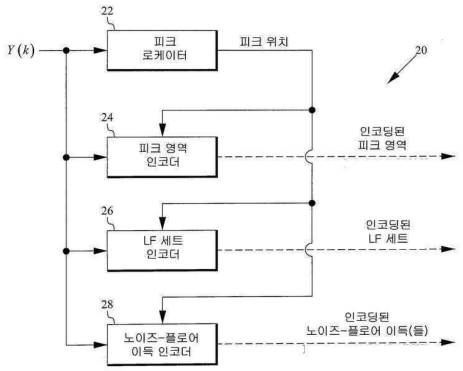
(54) 발명의 명칭 **하모닉 오디오 신호의 변환 인코딩/디코딩**

## (57) 요 약

하모닉 오디오 신호의 주파수 변환 계수((Y(k))를 인코딩하는 인코더(20)는 다음 구성요소를 포함한다. 미리 결정된 주파수 종속 임계값을 초과하는 크기를 갖는 스펙트럼 피크를 로케이팅하도록 구성되는 피크 로케이터(22). 상기 로케이팅된 피크를 포함하고 둘러싸는 피크 영역을 인코딩하도록 구성되는 피크 영역 인코더(24). 상기 피

(뒷면에 계속)

*대 표 도* - 도7



크 영역 밖에서, 상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 계수들의 적어도 하나의 저주파 세트를 인코딩하도록 구성되는 저주파 세트 인코더(26). 피크 영역 밖에서의 아직 인코딩되지 않은 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩하도록 구성되는 노이즈-플로어 이득 인코더(28).

(72) 발명자

**나슬룬트 세바스티안**

스웨덴 에스-16933 솔라 보크바겐 13

**포블로스 해럴드**

스웨덴 에스-187 54 태비 미트스포렛 29

# 명 세 서

## 청구범위

### 청구항 1

주파수 변환된 하모닉 오디오 신호를 인코딩하는 방법으로서,

주파수 변환된 하모닉 오디오 신호를 수신하는 단계와,

주파수 변환된 하모닉 오디오 신호에 대응하는 인코딩된 주파수 변환된 하모닉 오디오 신호를 생성하는 단계로서,

미리 결정된 주파수 종속 임계값을 초과하는 크기를 갖는 주파수 변환된 하모닉 오디오 신호 내에 스펙트럼 피크를 로케이팅하는 단계와,

상기 로케이팅된 스펙트럼 피크를 포함하고 둘러싸는 피크 영역을 인코딩하는 단계와;

상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 수정 이산 코사인 변환(MDCT: Modified Discrete Cosine Transform) 계수의 제1저주파 세트를, 상기 피크 영역 밖에서, 다수의 예약된 비트를 사용하여, 인코딩하고,

또한, 상기 피크 영역을 인코딩한 후 이용 가능한 예약되지 않은 비트가 있으면, 크로스오버 주파수 아래의 MDCT 계수들의 추가의 하나 이상의 저주파 세트를, 상기 피크 영역 밖에서, 인코딩하는 단계와;

피크 영역 밖에서 아직 인코딩되지 않은 MDCT 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩하는 단계에 기초하는,

상기 생성하는 단계와;

인코딩된 주파수 변환된 하모닉 오디오 신호를 출력하는 단계를 포함하는, 인코딩 방법.

### 청구항 2

청구항 1에 있어서,

피크 영역은,

피크의 스펙트럼 위치와 부호를 인코딩하고;

피크 이득을 양자화하며;

상기 양자화된 피크 이득을 인코딩하고;

상기 양자화된 피크 이득의 역에 의해 피크를 둘러싸는 미리 결정된 주파수 빈을 스케일링하며;

상기 스케일링된 주파수 빈을 형상 인코딩함으로써, 인코딩되는, 인코딩 방법.

### 청구항 3

청구항 1에 있어서,

MDCT 계수들의 저주파 세트를 인코딩하는 단계는 이득 형상 인코딩 방식에 기초하여 저주파 세트를 인코딩하는 단계를 포함하는, 인코딩 방법.

### 청구항 4

청구항 3에 있어서,

상기 이득 형상 인코딩 방식은 스칼라 이득 양자화 및 팩토리얼 펄스 형상 인코딩에 기초하는, 인코딩 방법.

### 청구항 5

청구항 1에 있어서,

2개의 고주파 세트들 각각에 대해 노이즈-플로어 이득을 인코딩하는 단계를 더 포함하는, 인코딩 방법.

**청구항 6**

제1항의 인코딩 방법에 따라 인코딩된 오디오 신호를 복원하는 방법으로서,

인코딩된 주파수 변환된 하모닉 오디오 신호를 수신하는 단계와,

인코딩된 주파수 변환된 하모닉 오디오 신호를 디코딩하고, 이에 의해서 복원된 주파수 변환된 하모닉 오디오 신호를 획득하는 단계로서,

인코딩된 주파수 변환된 하모닉 오디오 신호의 스펙트럼 피크 영역을 디코딩하는 단계로서, 상기 스펙트럼 피크 영역이 미리 결정된 주파수 종속 임계값을 초과하는 크기를 갖는 스펙트럼 피크를 포함하는, 디코딩하는 단계와,

인코딩된 주파수 변환된 하모닉 오디오 신호의 수정 이산 코사인 변환(MDCT: Modified Discrete Cosine Transform) 계수의 제1저주파 세트를 디코딩하는 단계와,

상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 각각의 저주파 세트의 MDCT 계수를, 상기 스펙트럼 피크 영역 밖에서, 분배하는 단계와,

스펙트럼 피크 영역 밖에 있는 인코딩된 주파수 변환된 하모닉 오디오 신호의 MDCT 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 디코딩하는 단계와,

대응하는 디코딩된 노이즈-플로어 이득을 갖는 노이즈로 MDCT 계수들의 각각의 고주파 세트를 충전하는 단계에 기초하는,

상기 획득하는 단계와,

복원된 주파수 변환 하모닉 오디오 신호를 출력하는 단계를 포함하는, 복원 방법.

**청구항 7**

청구항 6에 있어서,

피크 영역은,

피크의 스펙트럼 위치 및 부호를 디코딩하고,

피크 이득을 디코딩하고,

상기 피크를 둘러싸는 미리 결정된 주파수 빈의 형상을 디코딩하며,

디코딩된 피크 이득에 의해 디코딩된 형상을 스케일링함으로써, 디코딩되는, 복원 방법.

**청구항 8**

청구항 6에 있어서,

저주파 세트를 디코딩하는 단계는 이득-형상 디코딩 방식에 기초해서 저주파 세트를 디코딩하는 단계를 포함하는, 복원 방법.

**청구항 9**

청구항 8에 있어서,

이득-형상 디코딩 방식은 스칼라 이득 디코딩 및 팩토리얼 펄스 형상 디코딩에 기초하는, 복원 방법.

**청구항 10**

청구항 6에 있어서,

2개의 고주파 세트들 각각에 대해 노이즈-플로어 이득을 디코딩하는 단계를 포함하는, 복원 방법.

**청구항 11**

주파수 변환된 하모닉 오디오 신호를 인코딩하기 위한 인코더로서, 상기 인코더는 주파수 변환된 하모닉 오디오 신호를 획득하도록 구성되고, 프로세서 회로를 포함하며, 프로세서 회로는,

주파수 변환된 하모닉 오디오 신호에 대응하는 인코딩된 주파수 변환된 하모닉 오디오 신호를 생성하도록 구성되고,

미리 결정된 주파수 종속 임계값을 초과하는 크기를 갖는 주파수 변환된 하모닉 오디오 신호 내에 스펙트럼 피크를 로케이팅하며,

상기 로케이팅된 스펙트럼 피크를 포함하고 둘러싸는 피크 영역을 인코딩하고,

상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 수정 이산 코사인 변환(MDCT: Modified Discrete Cosine Transform) 계수의 제1저주파 세트를, 상기 피크 영역 밖에서, 다수의 예약된 비트를 사용하여, 인코딩하고,

또한, 상기 피크 영역을 인코딩한 후 이용 가능한 예약되지 않은 비트가 있으면, 크로스오버 주파수 아래의 MDCT 계수들의 추가의 하나 이상의 저주파 세트를, 상기 피크 영역 밖에서, 인코딩하며,

피크 영역 밖에서 아직 인코딩되지 않은 MDCT 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩하도록 구성되는 것에 기초하며,

인코딩된 주파수 변환된 하모닉 오디오 신호를 출력하도록 구성되는, 인코더.

**청구항 12**

청구항 11에 있어서,

상기 프로세서 회로는,

피크의 스펙트럼 위치와 부호를 인코딩하고,

피크 이득을 양자화하고,

양자화된 피크 이득을 인코딩하며,

상기 양자화된 피크 이득의 역에 의해 피크를 둘러싸는 미리 결정된 주파수 빈을 스케일링하고,

상기 스케일링된 주파수 빈을 형상 인코딩하도록 구성되는, 인코더.

**청구항 13**

청구항 11의 인코더를 포함하는 사용자 단말(UE)로서,

상기 인코더는, 원격 수신기에 전송하기 위해, UE의 무선 회로에 인코딩된 주파수 변환된 하모닉 오디오 신호를 출력하도록 구성되는, 사용자 단말.

**청구항 14**

제1항의 인코딩 방법에 따라 인코딩된 오디오 신호를 복원하도록 구성된 디코더로서, 상기 디코더는 인코딩된 주파수 변환된 하모닉 오디오 신호를 수신하도록 구성되고, 프로세서 회로를 포함하며, 프로세서 회로는,

인코딩된 주파수 변환된 하모닉 오디오 신호를 디코딩하고, 이에 의해서 복원된 주파수 변환된 하모닉 오디오 신호를 획득하도록 구성되고,

인코딩된 주파수 변환된 하모닉 오디오 신호의 스펙트럼 피크 영역을 디코딩하고, 상기 스펙트럼 피크 영역이 미리 결정된 주파수 종속 임계값을 초과하는 크기를 갖는 스펙트럼 피크를 포함하며,

수정 이산 코사인 변환(MDCT: Modified Discrete Cosine Transform) 계수의 제1저주파 세트를 디코딩하고,

상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 각각의 저주파 세트의 MDCT 계수를, 상기 스펙트럼 피크 영역 밖에서, 분배하며,

스펙트럼 피크 영역 밖에서 MDCT 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 디코딩하고,

대응하는 노이즈-플로어 이득을 갖는 노이즈로 MDCT 계수들의 각각의 고주파 세트를 충전하도록 구성되는 것에 기초하며,

복원된 주파수 변환 하모닉 오디오 신호를 출력하도록 구성되는, 디코더.

**청구항 15**

청구항 14에 있어서,

프로세서 회로는,

피크의 스펙트럼 위치 및 부호를 디코딩하고,

피크 이득을 디코딩하며,

상기 피크를 둘러싸는 미리 결정된 주파수 빈의 형상을 디코딩하고,

상기 디코딩된 피크 이득에 의해서 디코딩된 형상을 스케일링하도록 구성되는, 디코더.

**청구항 16**

청구항 14의 디코더를 포함하는 사용자 단말(UE)로서,

상기 디코더는, 대응하는 오디오 신호를 생성하기 위해, UE의 또 다른 오디오 신호 프로세서 회로에 복원된 변환된 하모닉 신호를 출력하도록 구성되는, 사용자 단말.

## *발명의 설명*

### *기 술 분 야*

[0001] 제안된 기술은 오디오 신호, 특히 하모닉(harmonic) 오디오 신호의 변환 인코딩/디코딩에 관한 것이다.

### *배 경 기 술*

[0002] 변환 인코딩(transform encoding)은 오디오 신호를 압축해서 송신하기 위해 사용되는 기본 기술이다. 변환 인코딩의 개념은 먼저 주파수 도메인(frequency domain)으로 신호를 변환하고 양자화하여 변환 계수를 송신하는 것이다. 디코더(decoder)는 수신된 변환 계수를 사용하여, 역 주파수 변환(inverse frequency transform)을 적용하여 신호 파형을 복원(reconstruct)한다(도 1 참조). 도 1에서, 오디오 신호 X(n)는 주파수 변환기(10)로 전송된다. 그 결과로 생긴 주파수 변환 Y(k)는 변환 인코더(12)로 전송되고, 인코딩된 변환은 디코더로 송신되며, 여기서 변환 디코더(14)에 의해 디코딩된다. 디코딩된 변환 $\hat{Y}(k)$ 는, 디코딩된 오디오 신호 $\hat{X}(n)$으로 변환시키는 역 주파수 변환기(16)로 전송된다. 이러한 방식 뒤에 있는 동기는 주파수 도메인 계수가 다음의 이유로서 더 효율적으로 양자화될 수 있다는 것이다.

[0003] 1) 변환 계수(도 1에서 Y(k))는 입력 신호 샘플(도 1에서 X(n))보다 더 비상관적(uncorrelate)이다.

[0004] 2) 주파수 변환은 에너지 컴팩션(compaction)을 제공한다(계수 Y(k)가 영(zero)에 더 근접하고, 무시될 수 있다).

[0005] 3) 변환 뒤에 있는 주관적 동기는 인간의 청각 시스템이 변환 도메인에서 동작하고, 그 도메인에서 지각적으로 중요한 신호 성분을 선택하는 것이 더 용이하다는 것이다.

[0006] 전형적인 변환 코덱(transform codec)에서, 신호 파형은 수정 이산 코사인 변환(MDCT: Modified Discrete Cosine Transform)을 사용하여, 블록 단위(50% 중첩)로 변환된다. MDCT형 변환 코덱에서, 블록 신호 파형 X(n)은 MDCT 벡터 Y(k)로 변환된다. 파형 블록의 길이는 20~40 ms 오디오 세그먼트(segments)에 대응한다. 길이를 2L로 표시하면, MDCT 변환은 다음과 같이 정의될 수 있다.

[0007] $$Y(k)=\sqrt{\frac{2}{L}}\sum_{n=0}^{2L-1}\sin\left[\left(n+\frac{1}{2}\right)\frac{\pi}{L}\right]\cos\left[\left(n+\frac{1}{2}+\frac{L}{2}\right)\left(k+\frac{1}{2}\right)\frac{\pi}{L}\right]X(n) \quad (1)$$

[0008] 여기서 k=0,...,L-1 이다. 이때 MDCT 벡터 Y(k)는 다중 대역(서브 벡터들)으로 분할되고, 각 대역에서의 에너지(또는 이득) G(j)는 다음과 같이 계산된다.

[0009] $$G(j)=\sqrt{\frac{1}{N_j}\sum_{k=m_j}^{m_j+N_j-1}Y^2(k)} \quad (2)$$

[0010] 여기서 $m_j$는 대역 j에서의 제1 계수이고, $N_j$는 해당 대역에서의 MDCT 계수의 개수를 의미한다(전형적인 범위는 8-32 계수를 포함). 균일한 대역 구조의 일례로서, 모든 j에 대해 $N_j$=8로 하면, G(0)는 첫 번째 8 계수의 에너지로 되고, G(1)은 그 다음 8 계수의 에너지 등으로 된다.

[0011] 이러한 에너지값 또는 이득은 양자화되는 스펙트럼 포락선(spectrum envelope)의 근사값을 제공하고, 양자화 인덱스는 디코더로 송신된다. 잔여 서브-벡터(sub-vector) 또는 형상(shape)은 해당 포락선 이득으로 MDCT 서브-벡터를 스케일링(scaling)함으로써 얻어지는데, 예를 들어 각 대역에서의 잔여분은 실효값(RMS) 단위를 갖도록 스케일링 된다. 이어서 잔여 서브-벡터 또는 형상은 해당 포락선 이득에 기초하하는 상이한 수의 비트로 양자화된다. 마지막으로, 디코더에서, MDCT 벡터는 해당 포락선 이득으로 잔여 서브-벡터 또는 형상을 스케일링함으로써 복원되고, 역 MDCT는 시간 도메인 오디오 프레임을 복원하는 데 사용된다.

[0012] 종래의 변환 인코딩 개념은 매우 하모닉인 오디오 신호, 예를 들어 단일 기기(single instruments)로 잘 작동하지 않는다. 이러한 하모닉 스펙트럼의 일례를 도 2에 도시하였다(비교를 위해 과도한 하모닉 없는 일반적인 오디오 스펙트럼을 도 3에 나타내었다). 그 이유는 스펙트럼 포락선을 갖는 정규화(normalization)가 "평면" 잔여 벡터(flat residual vector)로 이어지지 않고, 잔여 인코딩 방식이 적정 품질의 오디오 신호를 생성할 수 없기 때문이다. 신호 및 인코딩 모델 간의 이러한 불일치는 매우 높은 비트율(bitrates)에서만 해결될 수 있고, 대부분의 경우에는 이 방법이 적합하지 않다.

## 발명의 내용

### 해결하려는 과제

[0013] 제안된 기술의 목적은 하모닉 오디오 신호에 더 적합한 변환 인코딩/디코딩 방식이다.

### 과제의 해결 수단

[0014] 제안된 기술은 하모닉 오디오 신호의 주파수 변환 계수를 인코딩하는 방법을 포함한다. 상기 방법은

[0015] 미리 결정된 주파수 종속 임계값을 초과하는 크기를 갖는 스펙트럼 피크를 로케이팅(locating)하는 단계;

[0016] 상기 로케이팅된 피크를 포함하고 둘러싸는 피크 영역을 인코딩하는 단계;

[0017] 상기 피크 영역 밖에서, 상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 계수들의 적어도 하나의 저주파 세트를 인코딩하는 단계;

[0018] 피크 영역 밖에서 아직 인코딩되지 않은 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩하는 단계를 포함한다.

[0019] 제안된 기술은 또한 하모닉 오디오 신호의 주파수 변환 계수를 인코딩하는 인코더를 포함한다. 상기 인코더는

[0020] 미리 결정된 주파수 종속 임계값을 초과하는 크기를 갖는 스펙트럼 피크를 로케이팅하도록 구성되는 피크 로케이터;

[0021] 상기 로케이팅된 피크를 포함하고 둘러싸는 피크 영역을 인코딩하도록 구성되는 피크 영역 인코더;

[0022] 상기 피크 영역 밖에서, 상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 계수들의 적어도 하나의 저주파 세트를 인코딩하도록 구성되는 저주파 세트 인코더;

[0023] 피크 영역 밖에서 아직 인코딩되지 않은 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩하

도록 구성되는 노이즈-플로어 이득 인코더를 포함한다.

[0024] 제안된 기술은 또한 이러한 인코더를 포함하는 사용자 단말(UE)을 포함한다.

[0025] 제안된 기술은 또한 인코딩된 주파수 변환 하모닉 오디오 신호의 주파수 변환 계수를 복원하는 방법을 포함한다. 상기 방법은

[0026] 인코딩 주파수 변환 하모닉 오디오 신호의 스펙트럼 피크 영역을 디코딩하는 단계;

[0027] 계수들 중 적어도 하나의 저주파 세트를 디코딩하는 단계;

[0028] 피크 영역 밖에서 각각의 저주파 세트의 계수를 분배하는 단계;

[0029] 피크 영역 밖에서 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 디코딩하는 단계;

[0030] 대응하는 노이즈-플로어 이득을 갖는 노이즈로 각 고주파 세트를 충전하는 단계를 포함한다.

[0031] 제안된 기술은 또한 인코딩된 주파수 변환 하모닉 오디오 신호의 주파수 변환 계수를 복원하는 디코더를 포함한다. 상기 디코더는

[0032] 인코딩 주파수 변환 하모닉 오디오 신호의 스펙트럼 피크 영역을 디코딩하도록 구성되는 피크 영역 디코더;

[0033] 계수들 중 적어도 하나의 저주파 세트를 디코딩하도록 구성되는 저주파 세트 디코더;

[0034] 피크 영역 밖에서 각각의 저주파 세트의 계수를 분배하도록 구성되는 계수 분배기;

[0035] 피크 영역 밖에서 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 디코딩하도록 구성되는 노이즈-플로어 이득 디코더;

[0036] 대응하는 노이즈-플로어 이득을 갖는 노이즈로 각 고주파 세트를 충전하도록 구성되는 노이즈 필러를 포함한다.

[0037] 제안된 기술은 이러한 디코더를 포함하는 사용자 단말(UE)을 포함한다.

## *발명의 효과*

[0038] 제안된 하모닉 오디오 코딩 인코딩/디코딩 방식은 하모닉 오디오 신호의 큰 클래스에 대한 종래의 코딩 방식보다 더 좋은 지각적인 품질을 제공한다.

## *도면의 간단한 설명*

[0039] 본 기술은 그 추가적인 목적과 장점과 더불어, 첨부한 도면과 함께 다음의 설명을 참조함으로써 가장 잘 이해될 수 있을 것이다.

도 1은 주파수 변환 코딩 개념을 도시한다.

도 2는 하모닉 오디오 신호의 전형적인 스펙트럼을 도시한다.

도 3은 비하모닉(non-harmonic) 오디오 신호의 전형적인 스펙트럼을 도시한다.

도 4는 피크 영역을 도시한다.

도 5는 제안된 인코딩 방법을 도시하는 흐름도이다.

도 6A-D는 제안된 인코딩 방법의 실시예를 도시한다.

도 7은 제안된 인코더의 실시예의 블록도이다.

도 8은 제안된 디코딩 방법을 도시하는 흐름도이다.

도 9A-C는 제안된 디코딩 방법의 실시예를 도시한다.

도 10은 제안된 디코더의 실시예의 블록도이다.

도 11은 제안된 인코더의 실시예의 블록도이다.

도 12는 제안된 디코더의 실시예의 블록도이다.

도 13은 제안된 인코더를 포함하는 UE의 실시예의 블록도이다.

도 14는 제안된 디코더를 포함하는 UE의 실시예의 블록도이다.

도 15는 제안된 인코딩 방법의 일부에 대한 실시예의 흐름도이다.

도 16은 제안된 인코더에서 피크 영역 인코더의 실시예의 블록도이다.

도 17은 제안된 디코딩 방법의 일부에 대한 실시예의 흐름도이다.

도 18은 제안된 디코더에서 피크 영역 디코더의 실시예의 블록도이다.

## *발명을 실시하기 위한 구체적인 내용*

[0040] 도 2는 하모닉 오디오 신호의 전형적인 스펙트럼을 도시하고, 도 3은 비하모닉 오디오 신호의 전형적인 스펙트럼을 도시한다. 하모닉 오디오 신호의 스펙트럼은 더 약한 주파수 대역에 의해 분리되는 강한 스펙트럼 피크에 의해 형성되는 반면, 비하모닉 오디오 신호의 스펙트럼은 더 부드럽다.

[0041] 제안된 기술은 하모닉 오디오 신호를 잘 처리하는 다른 오디오 인코딩 모델을 제공한다. 주요 개념은 주파수 변환 벡터, 예를 들어 MDCT 벡터가 포락선과 잔여 부분으로 분할되지 않고, 대신에 스펙트럼 피크가 이웃하는 MDCT 빈(bins)과 함께 직접 추출되고 양자화된다는 것이다. 고주파수에서, 피크 이웃 외측의 저에너지 계수는 인코딩되지 않고, 디코더에서 노이즈가 채워진다. 여기에 기존의 인코딩에 사용되는 신호 모델{스펙트럼 포락선 + 잔류분}은 새로운 모델{스펙트럼 피크 + 노이즈-플로어(noise-floor)}로 대체된다. 저주파수에서, 피크 이웃 외측의 계수는 여전히 코딩되는데, 중요한 지각적 역할(perceptual role)이 있기 때문이다.

[0043] **인코더**

[0044] 인코더 측의 주요 단계는 다음과 같다.

[0045] ● 스펙트럼 피크 지역의 로케이팅 및 코딩

[0046] ● 저주파(LF) 스펙트럼 계수의 코딩. 코딩 영역의 크기는 피크 영역 코딩 후 남아있는 비트의 수에 의존한다.

[0047] ● 피크 영역 밖에서 스펙트럼 계수에 대한 노이즈-플로어 이득 코딩

[0048] 먼저 노이즈-플로어가 추정되고, 이어서 스펙트럼 피크가 피크 피킹(peak picking) 알고리즘에 의해 추출된다(해당 알고리즘은 부록 I-II에서 더 상세히 설명된다). 각 피크와 그 주변의 4 이웃들도 피크 위치에서의 단위 에너지로 정규화된다(도 4 참조). 다시 말해, 전체 영역은 피크가 진폭 1을 갖도록 스케일링된다. 피크 위치, 이득(피크 진폭, 크기를 표시) 및 부호가 양자화된다. 벡터 양자화(VQ: Vector Quantizer)는 피크를 둘러싸는 MDCT 빈에 적용되고 최상의 매치(match)를 제공하는 코드북 벡터(codebook vector)의 인덱스 $I_{shape}$를 검색한다. 주변의 형상 벡터뿐만 아니라 피크 위치, 이득 및 부호는 양자와 되고, 양자화 인덱스 $\{I_{positon}\ I_{gain}\ I_{sign}\ I_{shape}\}$는 디코더로 전송된다. 이들 지수에 추가하여, 디코더는 총 피크 수도 받는다.

[0049] 상기 예에서 각 피크 영역은 대칭적으로 피크를 둘러싸고 있는 4 이웃을 포함한다. 그러나 대칭 또는 비대칭 방식으로 피크를 둘러싸는 더 적거나 더 많은 이웃을 갖는 것도 가능하다.

[0050] 피크 영역이 양자화된 후, 가능한 모든 잔여 비트(노이즈-플로어 코딩에 대한 예약된 비트를 제외, 아래 참조)는 저주파 MDCT 계수를 양자화하기 위해 사용된다. 이것은 잔여 비양자화 MDCT 계수를, 예를 들면, 제1 빈에서 시작하는 24차 대역(dimensional bands)으로 그룹화하여 이루어진다. 따라서, 이 주파수 대역은 가장 낮은 주파수를 특정 크로스오버(crossover) 주파수까지 커버할 것이다. 피크 코딩에서 이미 양자화된 계수는 포함되지 않으므로, 상기 대역은 반드시 24 연속 계수로 구성될 필요는 없다. 이러한 이유로 대역은 아래의 "세트(sets)"로도 지칭될 것이다.

[0051] LF 대역 또는 세트의 총 수는 사용 가능한 비트의 수에 의존하지만, 적어도 하나의 세트를 생성하기 위해 예약된 충분한 비트가 항상 존재한다. 더 많은 비트를 사용할 수 있게 되면 세트당 비트의 최대 수에 대한 임계값에 도달할 때까지 제1 세트는 더 많은 비트를 할당받는다. 더 많은 비트를 사용할 수 있다면 다른 세트가 생성되고, 임계값에 도달할 때까지 이 세트에 비트가 할당된다. 가능한 모든 비트가 소비될 때까지 이 절차를 반복한다. 이것은 피크의 수가 프레임마다 변하기 때문에 이 프로세스가 정지되었을 때의 크로스오버 주파수가 프레임에 의존한다는 것을 의미한다. 피크 영역이 인코딩되었을 때 LF 인코딩에 이용 가능한 비트의 수에 의해 크로스오버 주파수가 결정된다.

[0052] LF 세트들의 양자화는 임의의 적절한 벡터 양자화 방식으로 수행될 수 있지만, 전형적으로 이득 형상 인코딩의

일부 형태가 사용된다. 예를 들어, 팩토리얼 펄스 코딩(factorial pulse coding)이 형상 벡터에 대해 이용될 수 있고, 스칼라 양자화기가 이득에 대해 사용될 수 있다.

[0053] 특정 수의 비트는 피크 영역 밖이고 LF 대역의 상한 주파수 위에서 적어도 하나의 고주파 대역의 계수들의 노이즈-플로어 이득을 인코딩하기 위해 항상 예약된다. 바람직하게는 2개의 이득이 이 목적을 위해 사용된다. 이러한 이득은 부록 I에 기재된 노이즈-플로어 알고리즘으로부터 얻을 수 있다. 팩토리얼 펄스 코딩이 저주파 대역을 인코딩하기 위해 사용되면 일부 LF 계수는 인코딩되지 않을 수도 있다. 이러한 이득은 대신에 고주파 대역 인코딩에 포함될 수 있다. LF 대역의 경우에서와 같이, HF 대역은 반드시 연속된 계수로 구성될 필요는 없다. 이러한 이유로 상기 대역은 아래의 "세트"로서도 지칭될 수 있다.

[0054] 가능하다면, 대역폭 확장(BWE) 영역에 대한 스펙트럼 포락선도 인코딩되고 전송된다. 대역의 개수(및 BWE 시작 시의 전이 주파수(transition frequency))는 비트율에 종속되며, 예를 들어 24 kbps에서 5.6 kHz 및 32 kbps에서 6.4 kHz이다.

[0055] 도 5는 일반적인 관점에서의 상기 제안된 인코딩 방법을 나타내는 흐름도이다. 단계 S1은 소정의 주파수 의존 임계값을 초과하는 크기를 갖는 스펙트럼 피크를 로케이팅한다. 단계 S2는 로케이팅된 피크를 포함하고 둘러싸는 피크 지역을 인코딩한다. 단계 S3은 피크 영역 밖에서, 피크 영역을 인코딩하는 데 사용되는 비트의 수에 따른 크로스오버 주파수 아래의, 계수들의 적어도 하나의 저주파 세트를 인코딩한다. 단계 S4는 피크 영역 밖에서 아직 인코딩되지 않은 (비인코딩 또는 잔여) 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩한다.

[0056] 도 6A-D는 제안된 인코딩 방법의 실시예를 도시한다. 도 6A는 인코딩할 신호 프레임의 MDCT 변환을 도시한다. 도면에서는 실제 신호보다 적은 수의 계수가 있다. 그러나 도면의 목적은 인코딩 과정만을 설명하기 위한 것이라는 점에 유념해야 한다. 도 6B는 이득-형상 인코딩을 위해 준비된 4 식별 피크 영역(identified peak regions)을 나타낸다. 부록 II에 기재한 방법이 이것들을 찾기 위해 사용될 수 있다. 다음으로 피크 영역 밖에서 LF 계수들이 도 6C에서 수집된다. 이들은 이득 형상 인코딩된 블록으로 연관된다. 도 6A에서의 원래 신호의 잔여 계수들은 도 6D에 도시된 고주파 계수들이다. 이들은 2 세트로 분할되고 각 세트에 대해 노이즈-플로어 이득에 의해 (연쇄 블록으로) 인코딩된다. 이러한 노이즈-플로어 이득은 각 세트의 에너지로부터, 또는 부록 I에 기재된 노이즈-플로어 추정 알고리즘에 의한 추정에 의해 얻어질 수 있다.

[0057] 도 7은 제안된 인코더(20)의 실시예에 대한 블록도이다. 피크 로케이터(peak locator)(22)는 소정의 주파수 종속 임계값을 초과하는 크기를 갖는 스펙트럼 피크를 로케이팅하도록 구성된다. 피크 영역 인코더(24)는 추출된 피크를 포함하고 둘러싸는 피크 영역을 인코딩하도록 구성된다. 저주파 세트 인코더(26)는 피크 영역 밖에서, 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의, 계수들의 적어도 하나의 저주파 세트를 인코딩하도록 구성된다. 노이즈-플로어 이득 인코더(28)는 피크 영역 밖에서, 아직 인코딩되지 않은 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩하도록 구성된다. 이 실시예에서, 인코더(24, 26, 28)는 어느 계수가 각각의 인코딩에 포함되는지를 결정하기 위해 상기 검출된 피크 위치를 사용한다.

[0059] **디코더**

[0060] 디코더의 주요 단계는 다음과 같다.

[0061] ● 스펙트럼 피크 영역을 복원(reconstruct)

[0062] ● LF 스펙트럼 계수를 복원

[0063] ● 수신된 노이즈-플로어 이득으로 스케일링된, 노이즈로 비코딩 영역(non-coded regions)을 노이즈 충전(noise-fill)

[0064] 오디오 디코더는 비트 스트림으로부터, 코딩된 피크 영역을 복원하기 위해 피크 영역의 수 및 양자화 인덱스 $\{I_{positon}\ I_{gain}\ I_{sign}\ I_{shape}\}$를 추출한다. 이러한 양자화 인덱스는 피크 이웃에 대해 최상의 매치를 제공하는 코드북 벡터에 대한 인덱스뿐만 아니라 스펙트럼 피크 위치, 이득 및 피크의 부호에 대한 정보를 포함한다.

[0065] 피크 영역 밖에서 MDCT 저주파 계수는 인코딩된 LF 계수로부터 복원된다.

[0066] 피크 영역 밖에서 MDCT 고주파 계수는 디코더에서 노이즈 충전된다. 노이즈-플로어 레벨은 가급적 2개의 코딩된 노이즈-플로어 이득의 형태(하나는 하부 및 하나는 상부 절반 또는 벡터의 일부)로 디코더에 의해 수신된다.

[0067] 가능하다면, 오디오 디코더는 HF MDCT 계수들에 대한 수신된 포락선 이득으로 미리 정의된 전이 주파수로부터

BWE를 실행한다.

[0068] 도 8은 일반적인 관점에서 상기 제안된 디코딩 방법을 나타내는 흐름도이다. 단계 S11은 인코딩된 주파수 변환 하모닉 오디오 신호의 스펙트럼의 피크 영역을 디코딩한다. 단계 S12는 계수들의 적어도 하나의 저주파 세트를 디코딩한다. 단계 S13은 피크 영역 밖에서 각각의 저주파 세트의 계수를 분배한다. 단계 S14는 피크 영역 밖에서 적어도 하나의 고주파 세트의 계수들의 노이즈-플로어 이득을 디코딩한다. 단계 S15는 대응하는 노이즈-플로어를 갖는 노이즈로 각 고주파 세트를 충전한다.

[0069] 일 실시예에서 저주파 세트의 디코딩은 이득-형상 디코딩 방식에 기초하한다.

[0070] 일 실시예에서 이득-형상 디코딩 방식은 스칼라 이득 디코딩 및 팩토리얼 펄스 디코딩에 기초하한다.

[0071] 일 실시예는 2개의 고주파 세트 각각에 대한 노이즈-플로어 이득을 디코딩하는 단계를 포함한다.

[0072] 도 9A-C는 상기 제안된 디코딩 방법의 실시예를 도시한다. 도 9A에 도시한 바와 같이, 주파수의 변환의 복원은 스펙트럼 피크 영역 및 그 위치를 이득-형상 디코딩함으로써 시작된다. 도 9B에서 LF 세트(들)은 이득-형상 디코딩되고 디코딩된 변환 계수는 피크 영역 밖의 블록에서 분배된다. 도 9C에서 노이즈-플로어 이득은 디코딩되고 잔여 변환 계수는 해당 노이즈-플로어 이득을 갖는 노이즈로 충전된다. 이러한 방식으로 도 6A의 변환은 대략적으로 복원된다. 도9C와 도 6A 및 6D의 비교는 노이즈 충전 영역이 상이한 개별 계수들을 갖지만, 예상한 대로, 동일한 에너지를 가짐을 나타낸다.

[0073] 도 10은 제안된 디코더(40)의 실시예의 블록도이다. 피크 영역 디코더(42)는 인코딩된 주파수 변환 하모닉 오디오 신호의 스펙트럼 피크 영역을 디코딩하도록 구성된다. 저주파 세트 디코더(44)는 계수들의 적어도 하나의 저주파 세트를 디코딩하도록 구성된다. 계수 분배기(46)는 피크 영역 밖에서 각각의 저주파 세트의 계수를 분배하도록 구성된다. 노이즈-플로어 이득 디코더(48)는 피크 영역 밖에서 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어를 디코딩하도록 구성된다. 노이즈 필러(noise filler)(50)는 해당 노이즈-플로어를 갖는 노이즈로 각 고주파 세트를 충전하도록 구성된다. 본 실시예에서 피크 위치는 피크 영역의 중복 기입을 방지하기 위해, 계수 분배기(46) 및 노이즈 필러(50)에 전송된다.

[0074] 여기에 기재된 단계, 기능, 절차 및/또는 블록들은, 범용 전자 회로 및 응용 전용 회로를 모두 포함하는, 이산 회로 또는 집적 회로 기술과 같은, 종래 기술을 사용하는 하드웨어로 구현할 수 있다.

[0075] 이와는 달리, 여기에 기재된 단계, 기능, 절차 및/또는 블록 중 적어도 일부는 적절한 처리 장비에 의한 실행을 위한 소프트웨어로 구현될 수 있다. 이러한 장비는, 예를 들어 하나 이상의 마이크로프로세서, 하나 이상의 디지털 신호 처리기(DSP), 하나 이상의 주문형 집적 회로(ASIC), 비디오 가속 하드웨어 또는 필드 프로그램어블 게이트 어레이(FPGA)와 같은 하나 이상의 적절한 프로그램어블 논리 소자(programmable logic devices)를 포함할 수 있다. 이러한 처리 요소들의 조합도 가능하다.

[0076] 이것은 또한 인코더/디코더에 이미 존재하는 일반적인 프로세싱 기능을 재사용하는 것도 가능할 수 있다는 것을 이해해야 한다. 이것은, 예를 들어, 기존 소프트웨어의 재프로그래밍 또는 새로운 소프트웨어 구성 요소의 추가에 의해 수행될 수 있다.

[0077] 도 11은 제안된 인코더(20)의 실시예에 대한 블록도이다. 이러한 실시예는 예를 들어, 마이크로프로세서와 같은 프로세서(110)에 기초하하는데, 이것은 피크를 로케이팅하는 소프트웨어(120), 피크 영역을 인코딩하는 소프트웨어(130), 적어도 하나의 저주파 세트를 인코딩하는 소프트웨어(140), 및 적어도 하나의 노이즈-플로어 이득을 인코딩하는 소프트웨어(150)를 실행한다. 소프트웨어는 메모리(160)에 저장된다. 프로세서(110)는 시스템 버스를 통해 상기 메모리와 통신한다. 입력되는 주파수 변환은, I/O(입출력) 버스를 제어하는 I/O 제어기(170)에 의해 수신되며, 이것은 프로세서(110) 및 메모리(160)에 연결된다. 소프트웨어(150)로부터 얻어지는 인코딩된 주파수는 변환은 I/O 버스를 통한 I/O 제어기(170)에 의해 메모리(160)로부터 출력된다.

[0078] 도 12는 제안된 디코더(40)의 실시예에 대한 블록도이다. 이러한 실시예는 예를 들어, 마이크로프로세서와 같은 프로세서(110)에 기초하하는데, 이것은 피크 영역을 디코딩하는 소프트웨어(220), 적어도 하나의 저주파 세트를 디코딩하는 소프트웨어(230), LF 계수를 분배하는 소프트웨어(240), 적어도 하나의 노이즈-플로어 이득을 디코딩하는 소프트웨어(250) 및 노이즈 충전을 하는 소프트웨어(260)를 실행한다. 소프트웨어는 메모리(270)에 저장된다. 프로세서(210)는 시스템 버스를 통해 상기 메모리와 통신한다. 입력되는 주파수 변환은, I/O 버스를 제어하는 I/O 제어기(280)에 의해 수신되며, 이것은 프로세서(210) 및 메모리(280)에 연결된다. 소프트웨어(260)로부터 얻어지는 복원된 주파수는 변환은 I/O 버스를 통한 I/O 제어기(280)에 의해 메모리(270)로부터 출력된다.

[0079] 상술한 기술은 이동 장치(예를 들어, 휴대 전화, 랩톱) 또는 개인용 컴퓨터와 같은 고정 장치에 사용될 수 있는 오디오 인코더/디코더에서 사용하고자 하는 것이다. 여기서 사용자 단말(UE: User Equipment)은 이러한 장치에 대한 일반적인 이름으로서 사용된다.

[0080] 도 13은 제안된 인코더를 포함하는 UE의 실시예에 대한 블록도이다. 마이크(70)로부터의 오디오 신호는 A/D 변환기(72)로 전송되고, 그 출력은 오디오 인코더(74)로 전송된다. 오디오 인코더(74)는 디지털 오디오 샘플을 주파수 도메인으로 변환하는 주파수 변환기(76)를 포함한다. 하모닉 신호 검출기(78)는 변환이 하모닉인지 또는 비하모닉 오디오인지를 결정한다. 비하모닉 오디오인 경우, 종래의 인코딩 모드(도시하지 않음)에서 인코딩된다. 하모닉 오디오인 경우, 제안된 기술에 따라 주파수 변환 인코더(20)로 전송된다. 인코딩된 신호는 수신기에 전송하기 위해 무선 유닛(80)에 전송된다.

[0081] 하모닉 신호 검출기(78)의 결정은 부록 I 및 II에서의 노이즈-플로어 에너지 $\overline{E}_{nf}$와 피크 에너지 $\overline{E}_{p}$에 기초한다. 논리는 다음과 같다. $\overline{E}_{p} / \overline{E}_{nf}$ 가 임계값보다 크고 검출된 피크의 수가 소정 범위에 있으면, 신호는 하모닉으로 분류된다. 그렇지 않은 경우 신호는 비하모닉으로 분류된다. 따라서 분류 및 인코딩 모드는 명시적으로 디코더에 시그널링된다.

[0082] 도 14는 제안된 디코더를 포함하는 UE의 실시예에 대한 블록도이다. 무선 유닛(82)에 의해 수신된 무선 신호를 기저 대역으로 변환되고, 채널은 디코딩되어 오디오 디코더(84)로 전송된다. 오디오 디코더는 디코딩 모드 선택기(86)를 포함하는데, 이것은 하모닉으로 분류된 경우에 상기 제안된 기술에 따라 신호를 주파수 변환 디코더(40)로 전송한다. 비하모닉 오디오로 분류된 경우, 종래의 디코더(도시하지 않음)에서 디코딩된다. 주파수 변환 디코더(40)는 상술한 바와 같이 주파수 변환을 복원한다. 복원된 주파수 변환은 역 주파수 변환기(88)에서 시간 도메인으로 변환된다. 결과적인 오디오 샘플들은 최종 오디오 신호를 스피커로 전송하는 D/A 변환 및 증폭 유닛(92)으로 전송된다.

[0083] 도 15는 제안된 인코딩 방법의 일부에 대한 실시예의 흐름도이다. 본 실시예에서, 도 5에서의 피크 영역 인코딩 단계 S2가 하위 단계 S2-A 내지 S2-E로 나누어진다. 단계 S2-A는 피크의 스펙트럼 위치 및 부호를 인코딩한다. 단계 S2-B는 피크 이득을 양자화한다. 단계 S2-C는 양자화 피크 이득을 인코딩한다. 단계 S2-D는 양자화 피크 이득의 역에 의해 피크를 둘러싸는 소정의 주파수 빈을 스케일링한다. 단계 S2-E는 스케일링된 주파수 빈을 형상 인코딩한다.

[0084] 도 16은 제안된 인코더에서 피크 영역 인코더에 대한 실시예의 블록도이다. 본 실시예에서는 피크 영역 인코더(24)가 구성요소 24-A 내지 24-D를 포함한다. 위치 및 부호 인코더 24-A는 피크의 스펙트럼 위치와 부호를 인코딩하도록 구성된다. 피크 이득 인코더 24-B는 피크 이득을 양자화하고 양자화된 피크 이득을 인코딩하도록 구성된다. 스케일링 유닛 24-C는 양자화된 피크 이득의 역에 의해 피크를 둘러싸는 소정의 주파수 빈들을 스케일링하도록 구성된다. 형상 인코더 24-D는 스케일링된 주파수 빈을 형상 인코딩하도록 구성된다.

[0085] 도 17은 제안된 디코딩 방법의 일부에 대한 실시예의 흐름도이다. 본 실시예에서는 도 8에서의 피크 영역 디코딩 단계 S11이 하위 단계 S11-A 내지 S11-D로 나누어진다. 단계 S11-A는 피크의 스펙트럼 위치와 부호를 디코딩한다. 단계 S11-B는 피크 이득을 디코딩한다. 단계 S11-C는 피크를 둘러싸는 소정의 주파수 빈의 형상을 디코딩한다. 단계 S11-D는 디코딩된 피크 이득에 의해 상기 디코딩된 형상을 스케일링한다.

[0086] 도 18은 제안된 디코더에서 피크 영역 디코더에 대한 실시예의 블록도이다. 본 실시예에서는 피크 영역 디코더(42)가 구성요소 42-A 내지 42-D를 포함한다. 위치 및 부호 디코더 42-A는 피크의 스펙트럼 위치와 부호를 디코딩하도록 구성된다. 피크 이득 디코더 42-B는 피크 이득을 디코딩하도록 구성된다. 형상 디코더 42-C 는 피크를 둘러싸는 소정의 주파수 빈의 형상을 디코딩하도록 구성된다. 스케일링 유닛 42-D는 디코딩된 피크 이득에 의해 상기 디코딩된 형상을 스케일링하도록 구성된다.

[0087] 24 kbps 모드에 대한 구체적인 세부 구현은 다음과 같다.

[0088] ● 코덱은 24 kbps의 비트율에서 프레임당 480 비트를 제공하는 20 ms 프레임으로 동작한다.

[0089] ● 처리된 오디오 신호는 32 kHz로 샘플링되고, 16 kHz의 오디오 대역폭을 갖는다.

[0090] ● 전이 주파수는 5.6 kHz(5.6 kHz보다 큰 모든 주파수 성분은 대역폭 확장임)로 설정된다.

[0091] ● 시그날링을 위한 예약된 비트 및 전이 주파수보다 큰 주파수의 대역폭 확장: ~30-40

[0092] ● 2개의 노이즈-플로어 이득을 코딩하기 위한 비트: 10.

[0093] ● 인코딩 스펙트럼 피크 영역의 수는 7-17이다. 피크 영역당 사용된 비트의 수는 ~20-22인데, 이것은 모든 피크 위치, 이득, 부호 및 형상을 인코딩하기 위한 총 개수 ~ 140-340을 제공한다.

[0094] ● 저주파 대역을 코딩하기 위한 비트: ~ 100-300

[0095] ● 코딩된 저주파 대역: 1-4 (각 대역은 8 MDCT 빈을 포함). 각각의 MDCT 빈은 25 Hz에 대응하기 때문에, 코딩된 저주파 영역은 200-800 Hz에 대응한다.

[0096] ● 대역폭 확장에 사용되는 이득과 피크 이득은 허프만(Huffman) 코딩되어 이에 의해 사용되는 비트 수는 심지어 일정한 수의 피크에 대한 프레임들 사이에서 변할 수 있다.

[0097] ● 피크 위치 및 부호 코딩은 피크 수가 증가함에 따라 더 효율적으로 되는 최적화를 사용한다. 7 피크에 대해, 위치와 부호는 피크당 약 6.9 비트를 필요로 하며, 17 피크에 대해 수는 피크당 약 5.7 비트이다.

[0098] ● 저주파 대역 코딩이 마지막에 오고 단지 비트들이 남아 있는 한 무엇이든 사용하기 때문에, 얼마나 많은 비트가 코딩의 상이한 단계에서 사용되었는지의 변동성은 문제가 되지 않는다. 그러나 하나의 저주파 대역을 인코딩하기 위한 충분한 비트가 남아 있도록 시스템이 설계된다.

[0099] 아래 표는 ITU-R BS.1534-1 MUSHRA(Multiple Stimuli with Hidden Reference and Anchor)에 기재된 절차에 따라 수행된 청취 시험의 결과를 보여준다. MUSHRA 시험에서 스케일은 0 내지 100인데, 여기서 낮은 값은 낮은 지각 품질에 대응하고, 높은 값은 높은 지각 품질에 대응한다. 두 코덱은 모두 24 kbps에서 동작한다. 시험 결과는 24 음악 항목에 대해 평균화되고 8 청취자로부터 선택된다.

[0100]

| 시험 중인 시스템 | MUSHRA 스코어 |
|---|---|
| 저주파 대역 앵커 신호(대역폭 7 kHz) | 48.89 |
| 종래의 코딩 방식 | 49.94 |
| 제안된 하모닉 코딩 방식 | 55.87 |
| 기준 신호(대역폭 16 kHz) | 1000.00 |

[0101] 다양한 수정 및 변형이 첨부된 청구범위에 의해 정의되는 본 발명의 범위로부터 벗어나지 않으면서 제안된 기술로 구현될 수 있음을 당업자는 이해할 수 있을 것이다.

[0103] **부록 I**

[0104] 노이즈-플로어 추정 알고리즘은 변환 계수 |Y(k)|의 절대값에서 동작한다. 순시 노이즈-플로어 에너지 $E_{nf}(k)$는 반복(recursion)에 따라 추정된다.

[0105] $$E_{nf}(k) = \alpha E_{nf}(k) + (1-\alpha)|Y(k)| \quad (3)$$

[0106] 여기서

[0107] $$\alpha = \begin{cases} 0.9578 & if \quad |Y(k)| > E_{nf}(k-1) \\ 0.6472 & if \quad |Y(k)| \le E_{nf}(k-1) \end{cases} \quad (4)$$

[0108] 가중 팩터 α의 특정한 형태는 고에너지 변환 계수의 영향을 최소화하고 저에너지 계수의 기여를 강조한다. 마지막으로 노이즈-플로어 레벨 $\overline{E}_{nf}$ 은 순시 에너지 $E_{nf}(k)$를 단순히 평균함으로써 추정된다.

[0109] **부록 II**

[0110] 피크-픽킹(peak-picking) 알고리즘은 노이즈-플로어 레벨과 스펙트럼 피크의 평균 레벨의 지식이 필요하다. 피크 에너지 추정 알고리즘은 노이즈-플로어 추정 알고리즘과 유사하지만, 저에너지 대신에 높은 스펙트럼 에너지를 추적한다.

[0111] $$E_p(k) = \beta E_p(k) + (1-\beta)|Y(k)| \quad (5)$$

[0112] 여기서

[0113] $$\beta = \begin{cases} 0.4223 & if \quad |Y(k)| > E_p(k-1) \\ 0.8029 & if \quad |Y(k)| \le E_p(k-1) \end{cases} \qquad (6)$$

[0114] 이 경우에, 가중 팩터는 저에너지 변환 계수의 영향을 최소화하고, 고에너지 계수의 기여를 강조한다. 전체적인 피크 에너지 $\overline{E}_p$ 는 순시 에너지를 단순히 평균함으로써 추정된다.

[0115] 피크와 노이즈-플로어 레벨이 산출될 때, 임계 레벨 θ은 $\gamma = 0.88579$ 로서 다음과 같이 형성된다.

[0116] $$\theta = \left( \frac{\overline{E}_p}{\overline{E}_{nf}} \right)^{\gamma} \overline{E}_{nf} \qquad (7)$$

[0117] 변환 계수는 임계값과 비교되고, 그보다 큰 진폭을 갖는 것들은 피크 후보의 벡터를 형성한다. 자연적인 소스가 예를 들어 80 Hz와 같이 매우 근접한 피크를 일반적으로 생성하지 않기 때문에, 피크 후보를 갖는 벡터는 더 정제된 다. 벡터 성분은 내림차순(decreasing order)으로 추출되고, 각 성분의 이웃은 영(0)으로 설정된다. 이러한 방식으로 특정 스펙트럼 영역에서 가장 큰 성분만이 남고, 이러한 성분들의 세트는 현재 프레임에 대한 스펙트럼 피크를 형성한다.

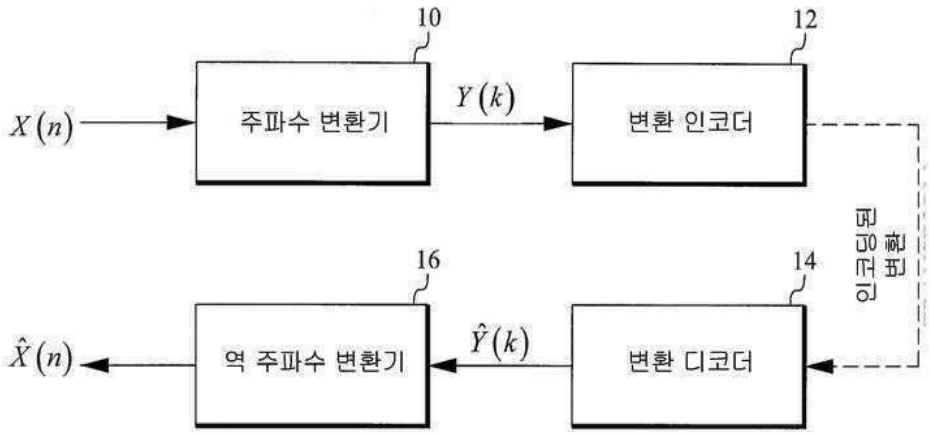
## 부호의 설명

[0118]

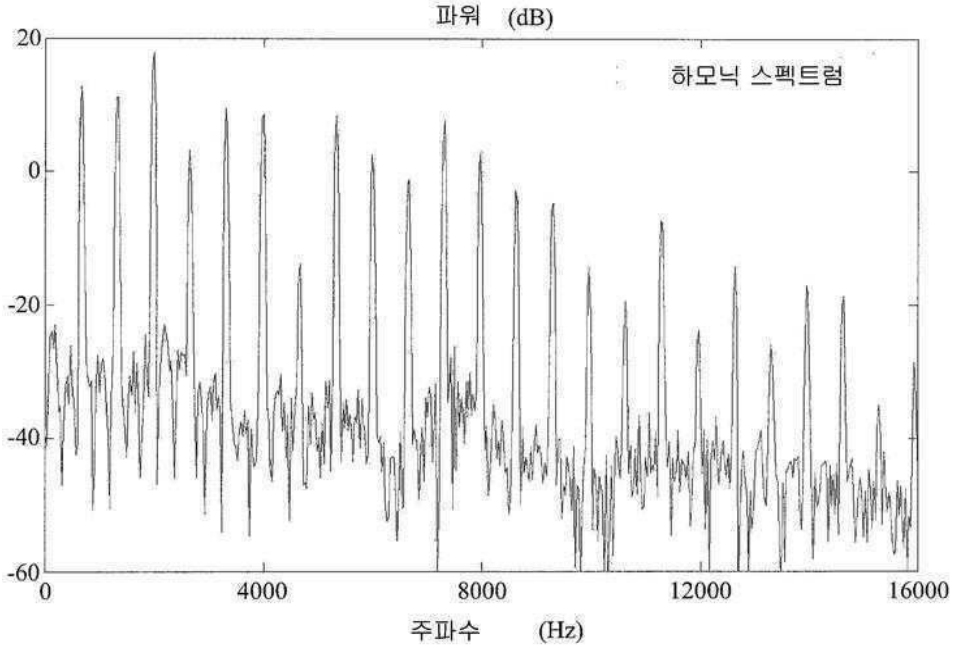
| | |
|---|---|
| ASIC | Application Specific Integrated Circuit |
| BWE | BandWidth Extension |
| DSP | Digital Signal Processors |
| FPGA | Field Programmable Gate Arrays |
| HF | High-Frequency |
| LF | Low-Frequency |
| MDCT | Modified Discrete Cosine Transform |
| RMS | Root Mean Square |
| VQ | Vector Quantizer |

## 도면

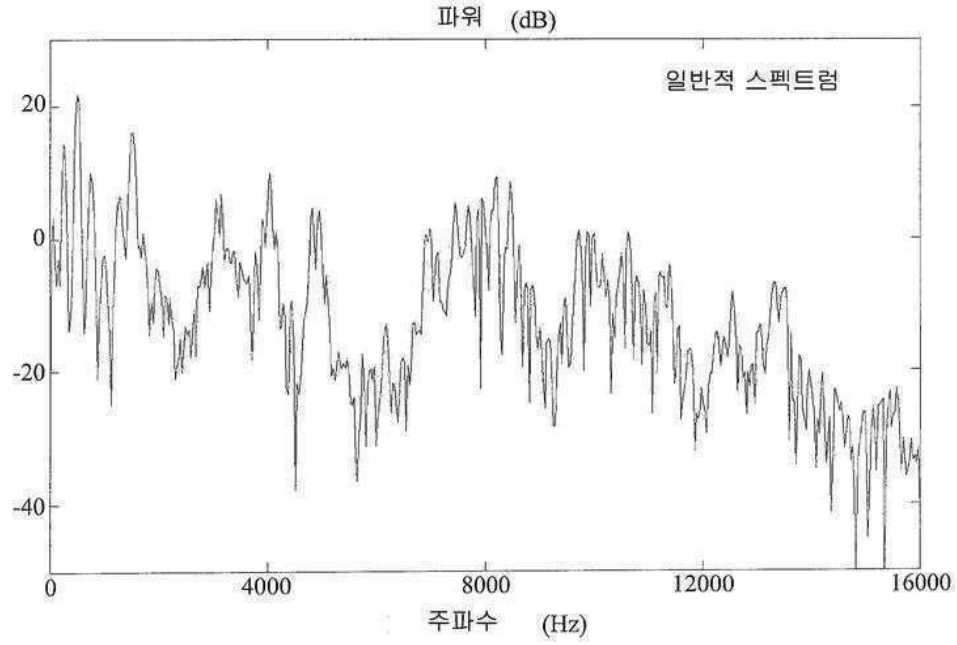
### 도면1



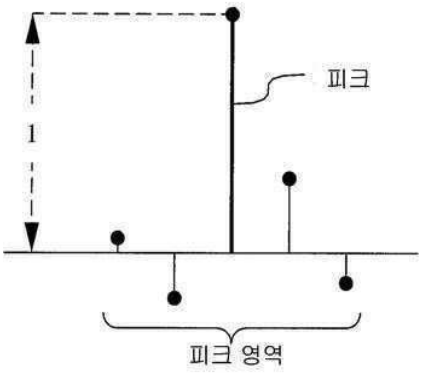
도면2


파워 (dB)
하모닉 스펙트럼
20
0
-20
-40
-60
0
4000
8000
12000
16000
주파수 (Hz)

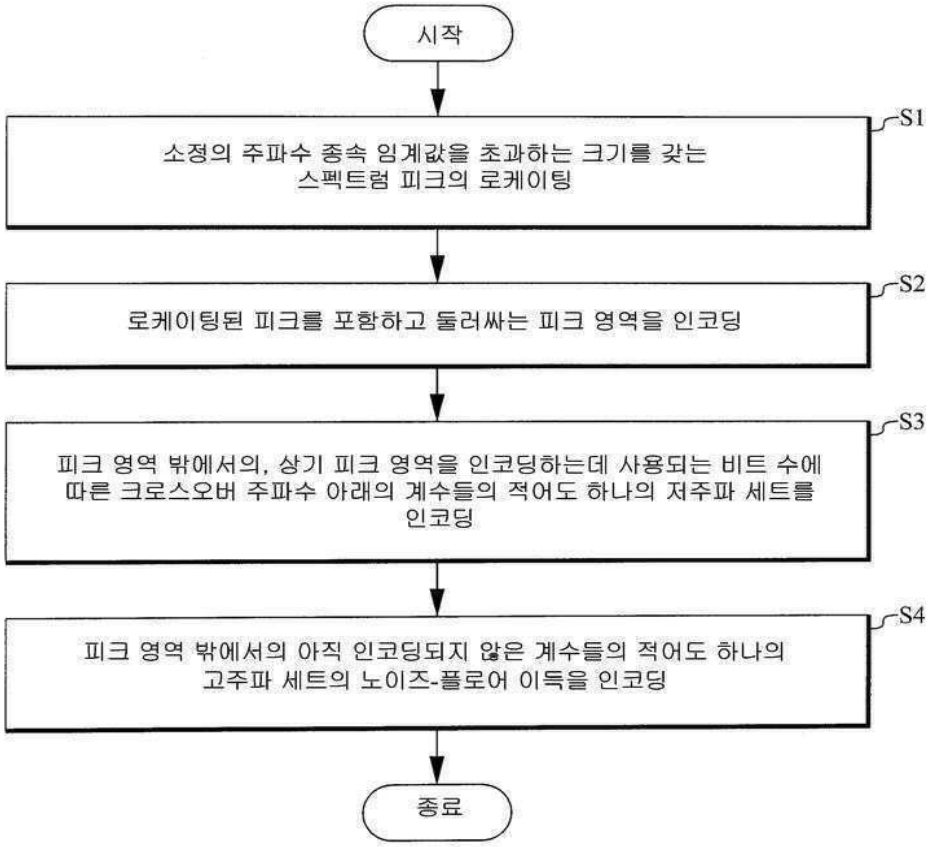
도면3


파워 (dB)
일반적 스펙트럼
20
0
-20
-40
0
4000
8000
12000
16000
주파수 (Hz)

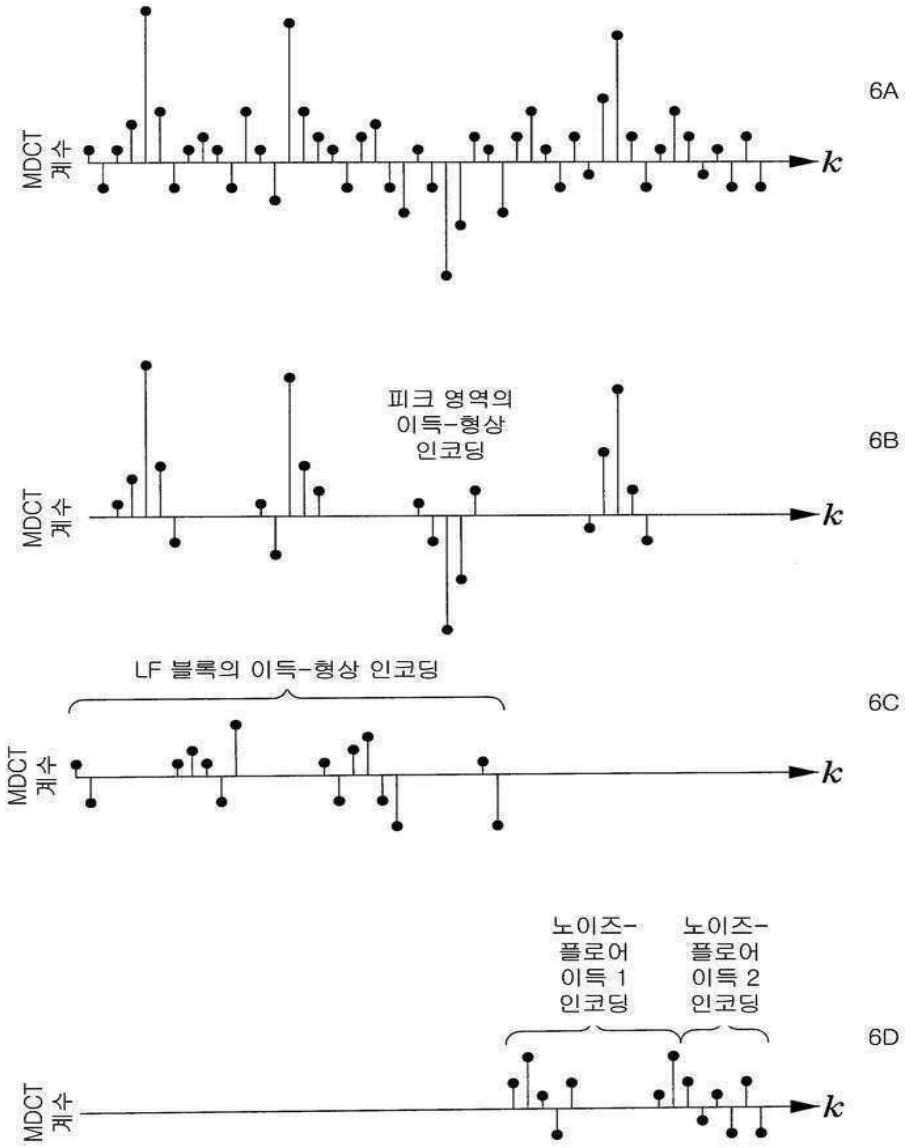
도면4


피크
1
피크 영역

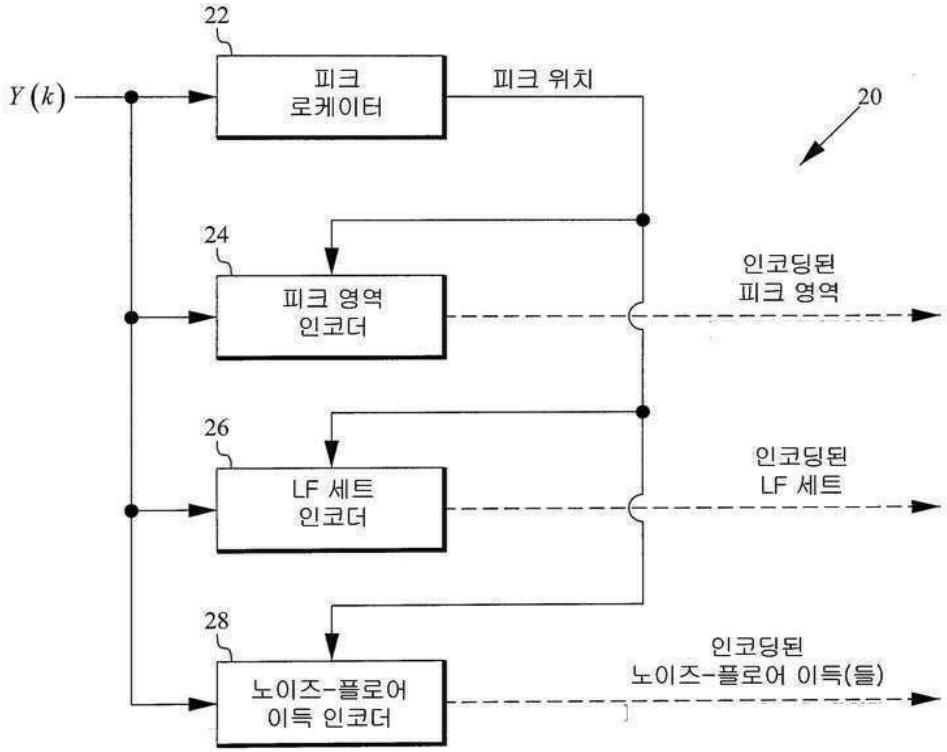
도면5


시작
소정의 주파수 종속 임계값을 초과하는 크기를 갖는 스펙트럼 피크의 로케이팅
S1
로케이팅된 피크를 포함하고 둘러싸는 피크 영역을 인코딩
S2
피크 영역 밖에서의, 상기 피크 영역을 인코딩하는데 사용되는 비트 수에 따른 크로스오버 주파수 아래의 계수들의 적어도 하나의 저주파 세트를 인코딩
S3
피크 영역 밖에서의 아직 인코딩되지 않은 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 인코딩
S4
종료

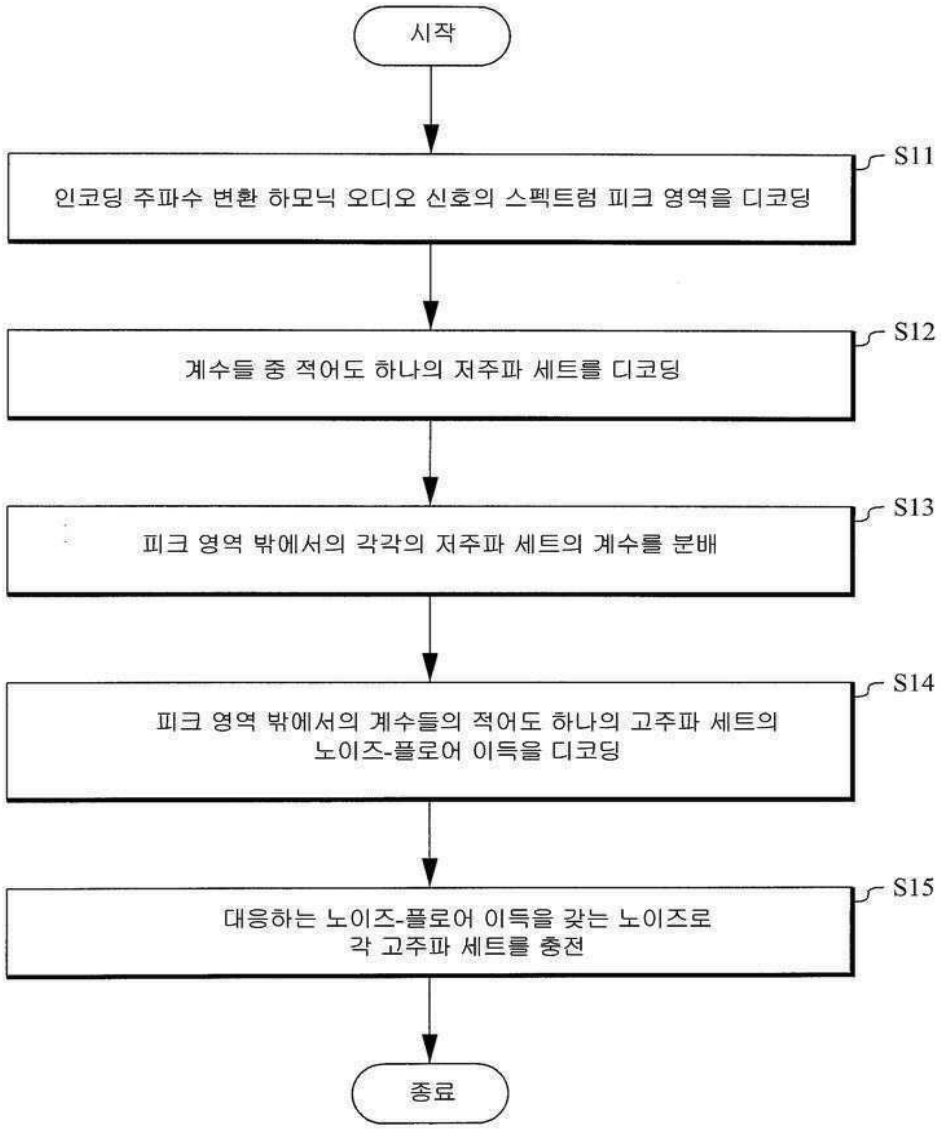
도면6


6A
MDCT 계수
k
6B
피크 영역의 이득-형상 인코딩
MDCT 계수
k
LF 블록의 이득-형상 인코딩
6C
MDCT 계수
k
노이즈-플로어 이득 1 인코딩
노이즈-플로어 이득 2 인코딩
6D
MDCT 계수
k

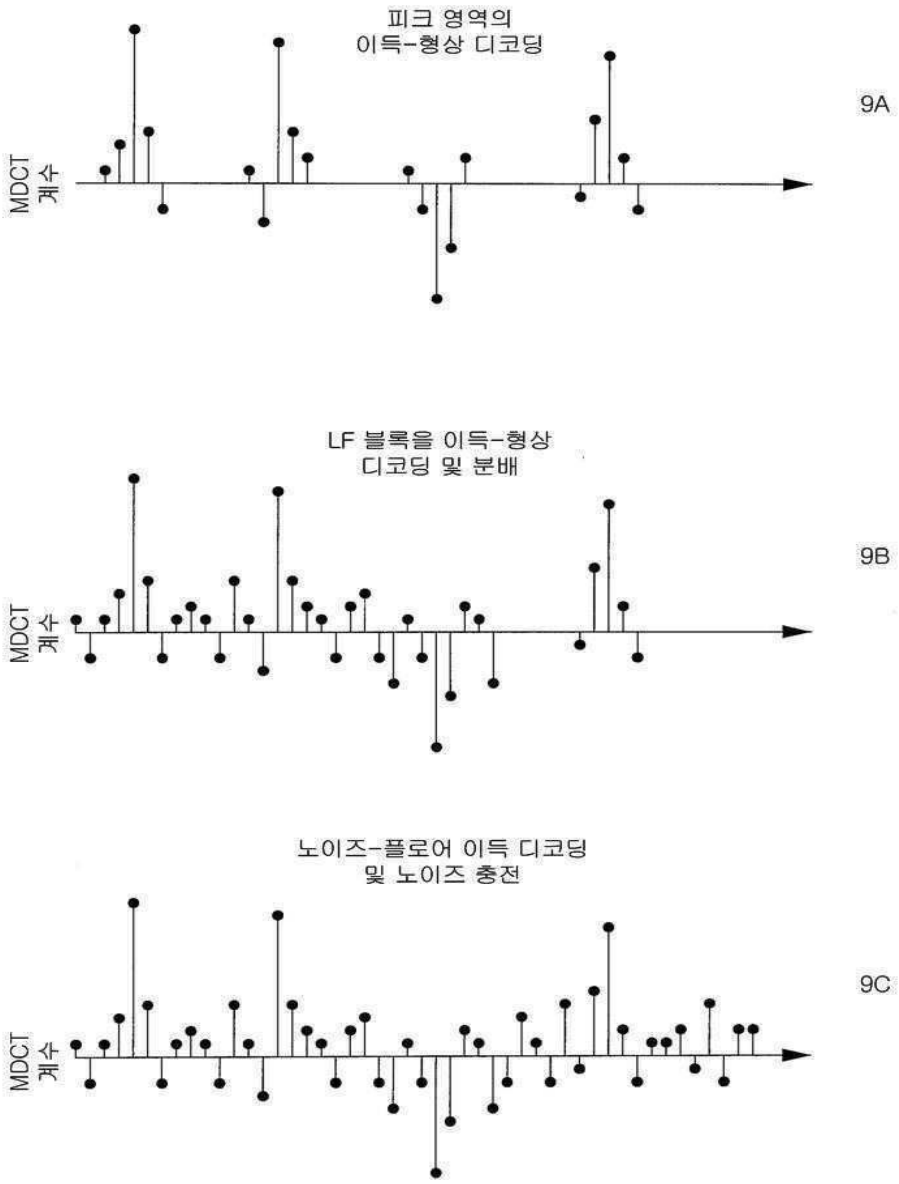
도면7


22
Y(k)
피크
로케이터
피크 위치
20
24
피크 영역
인코더
인코딩된
피크 영역
26
LF 세트
인코더
인코딩된
LF 세트
28
노이즈-플로어
이득 인코더
인코딩된
노이즈-플로어 이득(들)

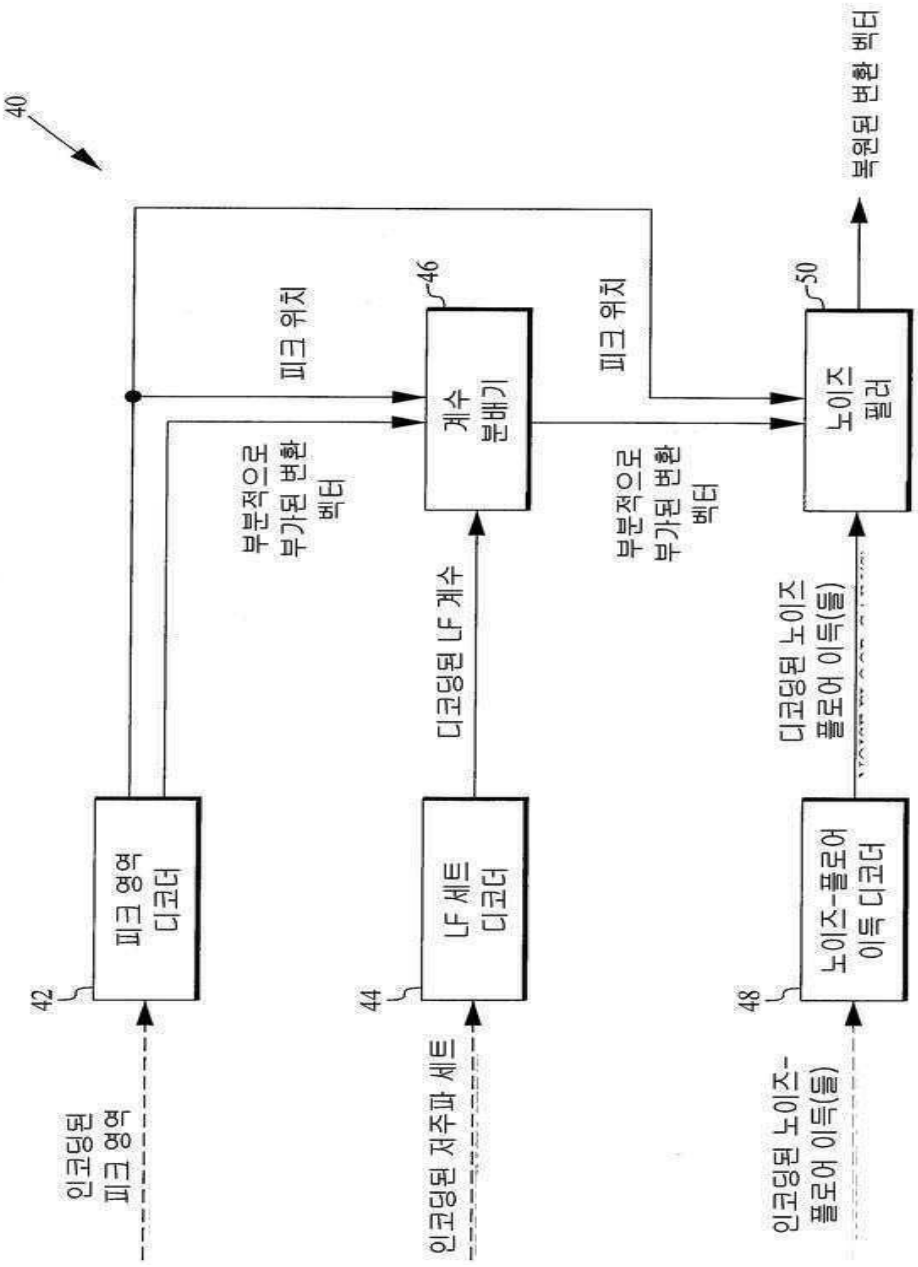
도면8


시작
인코딩 주파수 변환 하모닉 오디오 신호의 스펙트럼 피크 영역을 디코딩
S11
계수들 중 적어도 하나의 저주파 세트를 디코딩
S12
피크 영역 밖에서의 각각의 저주파 세트의 계수를 분배
S13
피크 영역 밖에서의 계수들의 적어도 하나의 고주파 세트의 노이즈-플로어 이득을 디코딩
S14
대응하는 노이즈-플로어 이득을 갖는 노이즈로 각 고주파 세트를 충전
S15
종료

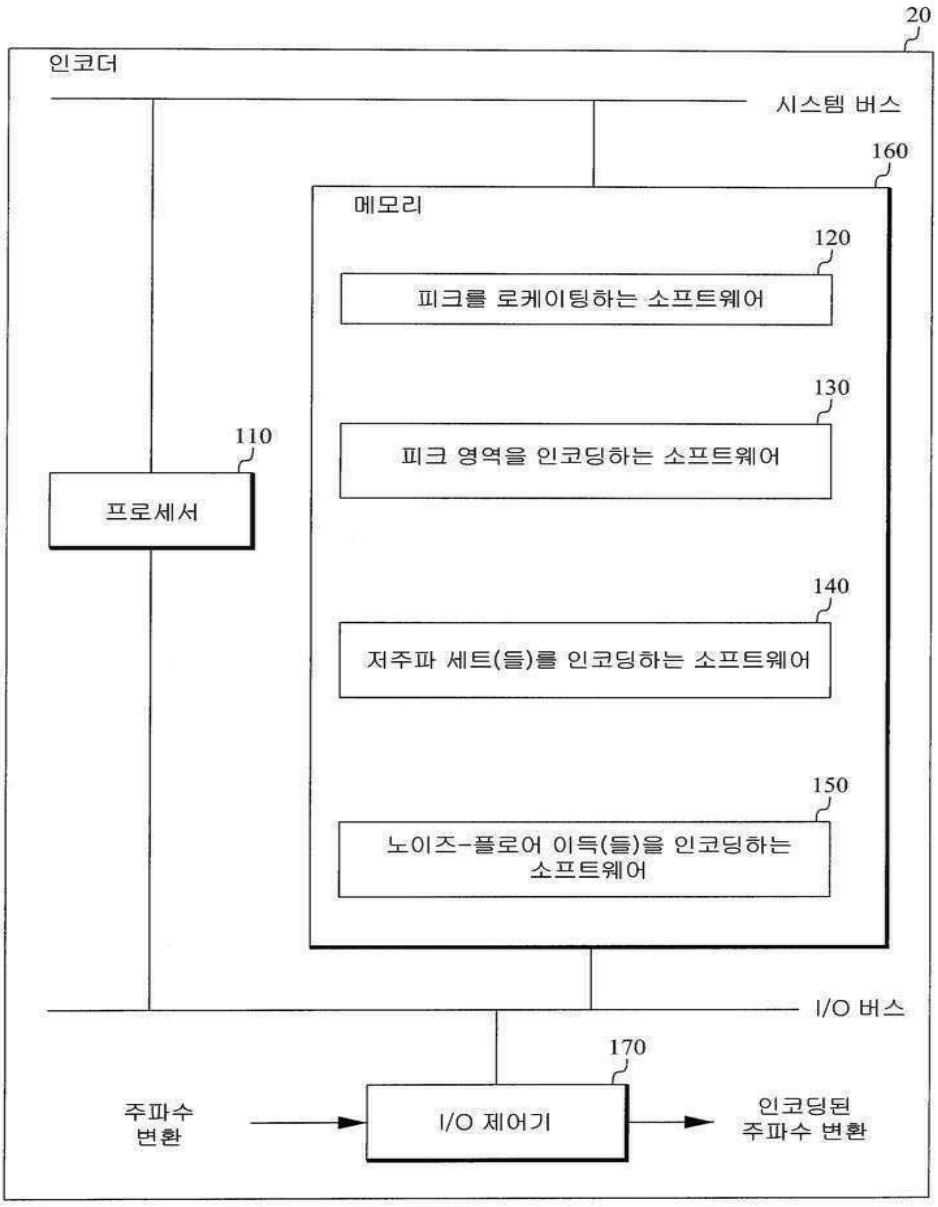
도면9


피크 영역의
이득-형상 디코딩
MDCT
계수
9A
LF 블록을 이득-형상
디코딩 및 분배
MDCT
계수
9B
노이즈-플로어 이득 디코딩
및 노이즈 충전
MDCT
계수
9C

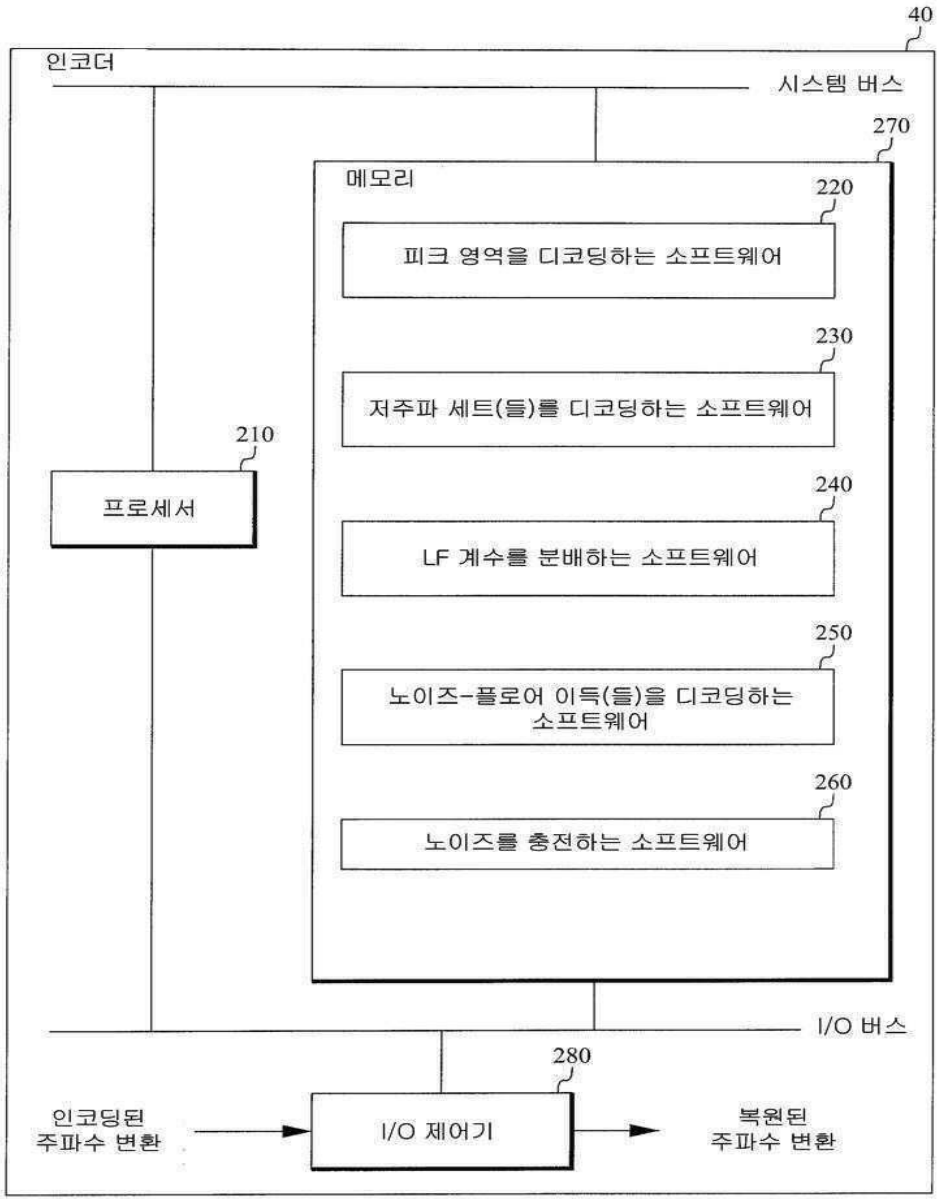
도면10


40
복원된 변환 벡터
피크 위치
46
계수 분배기
피크 위치
50
노이즈 필터
부분적으로 부가된 변환 벡터
부분적으로 부가된 변환 벡터
디코딩된 LF 계수
디코딩된 노이즈 플로어 이득(들)
피크 영역 디코더
LF 세트 디코더
노이즈-플로어 이득 디코더
42
44
48
인코딩된 피크 영역
인코딩된 저주파 세트
인코딩된 노이즈-플로어 이득(들)

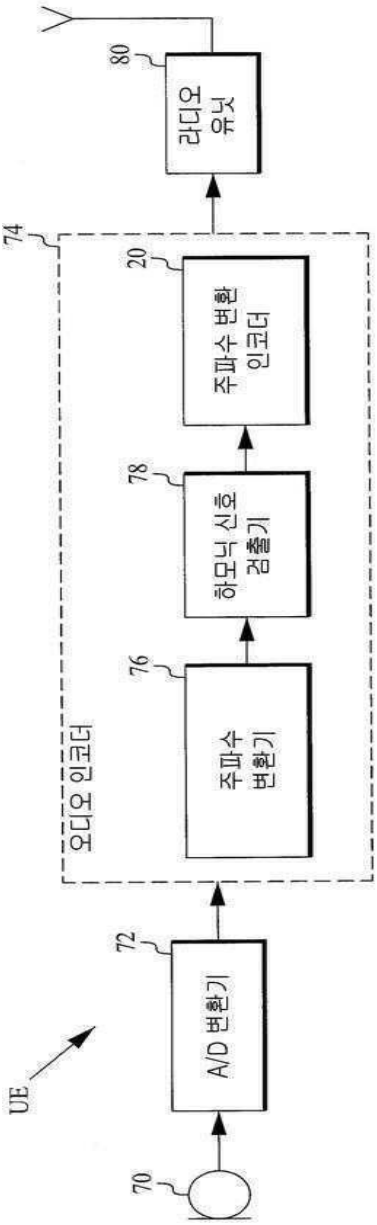
도면11


20
인코더
시스템 버스
160
메모리
120
피크를 로케이팅하는 소프트웨어
130
피크 영역을 인코딩하는 소프트웨어
110
프로세서
140
저주파 세트(들)를 인코딩하는 소프트웨어
150
노이즈-플로어 이득(들)을 인코딩하는 소프트웨어
I/O 버스
170
주파수 변환
I/O 제어기
인코딩된 주파수 변환

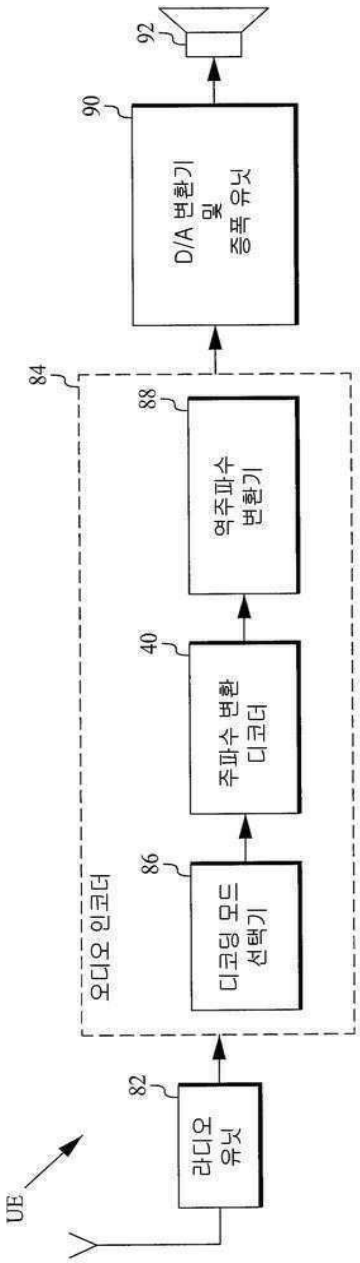
도면12


40
인코더
시스템 버스
270
메모리
220
피크 영역을 디코딩하는 소프트웨어
230
저주파 세트(들)를 디코딩하는 소프트웨어
210
프로세서
240
LF 계수를 분배하는 소프트웨어
250
노이즈-플로어 이득(들)을 디코딩하는 소프트웨어
260
노이즈를 충전하는 소프트웨어
I/O 버스
280
인코딩된 주파수 변환
I/O 제어기
복원된 주파수 변환

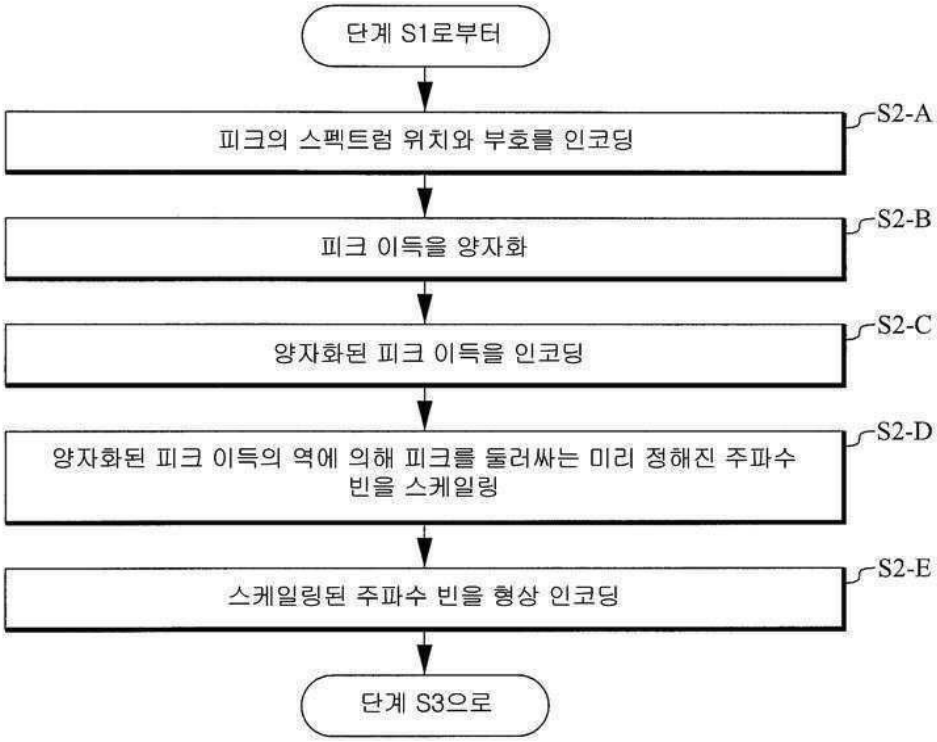
도면13


UE
70
A/D 변환기
72
오디오 인코더
74
주파수 변환기
76
하모닉 신호 검출기
78
주파수 변환 인코더
20
라디오 유닛
80

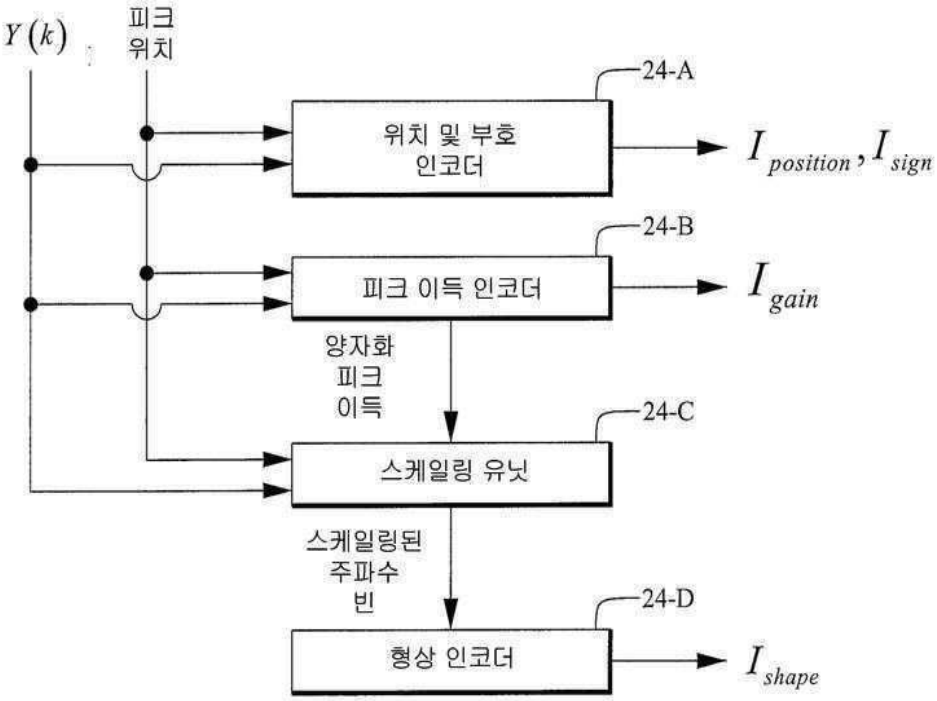
도면14


92
90
D/A 변환기
및
증폭 유닛
84
88
역주파수
변환기
40
주파수 변환
디코더
86
디코딩 모드
선택기
오디오 인코더
82
라디오
유닛
UE

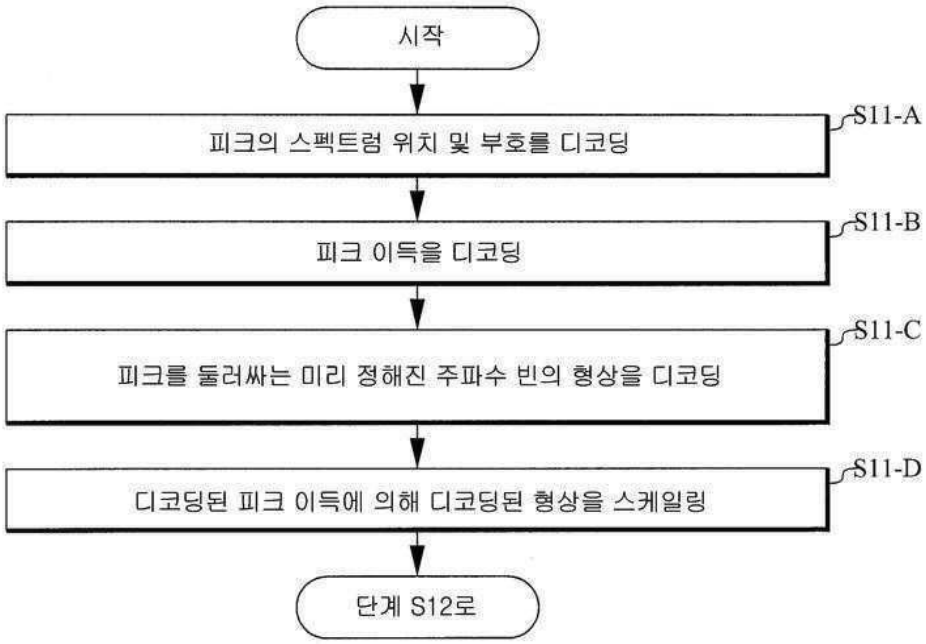
도면15


단계 S1로부터
피크의 스펙트럼 위치와 부호를 인코딩
S2-A
피크 이득을 양자화
S2-B
양자화된 피크 이득을 인코딩
S2-C
양자화된 피크 이득의 역에 의해 피크를 둘러싸는 미리 정해진 주파수 빈을 스케일링
S2-D
스케일링된 주파수 빈을 형상 인코딩
S2-E
단계 S3으로

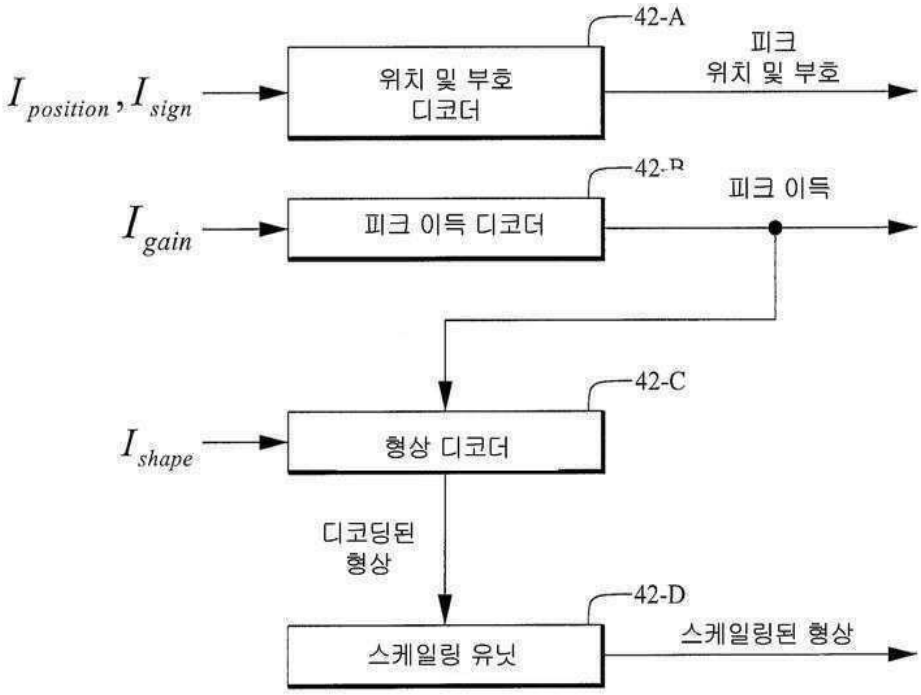
도면16


Y(k)
피크 위치
24-A
위치 및 부호 인코더
I position, I sign
24-B
피크 이득 인코더
I gain
양자화 피크 이득
24-C
스케일링 유닛
스케일링된 주파수 빈
24-D
형상 인코더
I shape

도면17


시작
피크의 스펙트럼 위치 및 부호를 디코딩
S11-A
피크 이득을 디코딩
S11-B
피크를 둘러싸는 미리 정해진 주파수 빈의 형상을 디코딩
S11-C
디코딩된 피크 이득에 의해 디코딩된 형상을 스케일링
S11-D
단계 S12로

도면18


$I_{position}, I_{sign}$
위치 및 부호 디코더
42-A
피크 위치 및 부호
$I_{gain}$
피크 이득 디코더
피크 이득
$I_{shape}$
형상 디코더
42-C
디코딩된 형상
스케일링 유닛
42-D
스케일링된 형상