



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2023-0066137  
(43) 공개일자 2023년05월12일

- |   |  |
|---|--|
| (51) 국제특허분류(Int. Cl.)<br>G10L 19/032 (2013.01) G10L 19/00 (2006.01)<br>G10L 19/16 (2013.01)<br>(52) CPC특허분류<br>G10L 19/032 (2013.01)<br>G10L 19/0017 (2013.01)<br>(21) 출원번호 10-2023-7015080(분할)<br>(22) 출원일자(국제) 2015년07월28일<br>심사청구일자 없음<br>(62) 원출원 특허 10-2017-7002772<br>원출원일자(국제) 2015년07월28일<br>심사청구일자 2020년07월21일<br>(85) 번역문제출일자 2023년05월02일<br>(86) 국제출원번호 PCT/KR2015/007901<br>(87) 국제공개번호 WO 2016/018058<br>국제공개일자 2016년02월04일<br>(30) 우선권주장<br>62/029,736 2014년07월28일 미국(US) | (71) 출원인<br>삼성전자주식회사<br>경기도 수원시 영통구 삼성로 129 (매탄동)<br>(72) 발명자<br>성호상<br>경기도 수원시 영통구 삼성로 129 (매탄동)<br>주기현<br>경기도 수원시 영통구 삼성로 129 (매탄동)<br>오은미<br>경기도 수원시 영통구 삼성로 129 (매탄동)<br>(74) 대리인<br>리엔목특허법인 |
|---|--|

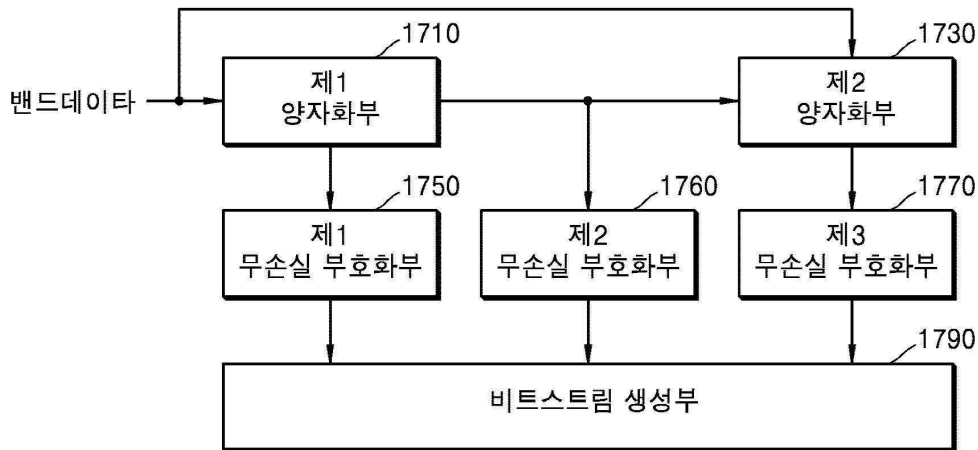
전체 청구항 수 : 총 8 항

(54) 발명의 명칭 **신호 부호화방법 및 장치와 신호 복호화방법 및 장치**

**(57) 요약**

스펙트럼 부호화방법은 현재 밴드의 스펙트럼 데이터에 대하여 제1 양자화방식에 따라서 양자화하는 단계, 스펙트럼 데이터 및 양자화된 스펙트럼 데이터를 이용하여 현재 밴드의 하위비트를 생성하는 단계, 현재 밴드의 하위비트가 포함된 하위비트 시퀀스에 대하여 제2 양자화방식에 따라서 양자화하는 단계, 및 양자화된 스펙트럼 데이터 중 N 비트(여기서 N는 1 이상)를 제외한 상위비트 및 양자화된 하위비트 시퀀스에 근거하여 비트스트림을 생성하는 단계를 포함할 수 있다.

**대표도** - 도17



(52) CPC특허분류  
*G10L 19/167* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

논제로(non-zero) 밴드의 스펙트럼 데이터를 USQ(Uniform Scalr Quantization)를 사용하여 양자화하는 단계;

상기 양자화된 스펙트럼 데이터의 하위 비트를 추출하는 단계;

모든 논제로 밴드의 상기 양자화된 스펙트럼 데이터의 하위 비트들을 수집하여 하위 비트 시퀀스를 생성하는 단계; 및

상기 하위 비트 시퀀스를 TCQ(Trellis Coded Quantization)을 사용하여 양자화하는 단계;

상기 양자화된 스펙트럼 데이터 중 상기 하위 비트를 제외한 상위 비트들 및 상기 양자화된 하위 비트 시퀀스에 기초하여 비트스트림을 생성하는 단계를 포함하고,

상기 TCQ에 사용되는 비트수는 각 논제로 밴드의 양자화를 위해 할당된 비트수에서 균등하게 추출되고, 상기 논제로 밴드에 남은 비트수는 USQ에 사용되는, 스펙트럼 부호화 방법.

#### 청구항 2

제1 항에 있어서, 상기 스펙트럼 데이터의 대역폭은 초광대역 혹은 전대역인 스펙트럼 부호화 방법.

#### 청구항 3

제1 항에 있어서, 상기 비트스트림을 생성하는 단계는,

상기 논제로 밴드의 중요 스펙트럼 성분의 개수, 위치 및 부호에 대하여 제1 무손실 부호화를 수행하는 단계;

상기 양자화된 스펙트럼 데이터 중 상기 하위 비트를 제외하여 구성되는 크기 정보에 대하여 제2 무손실 부호화를 수행하는 단계;

상기 양자화된 하위 비트 시퀀스에 대하여 제3 무손실 부호화를 수행하는 단계; 및

상기 제1 무손실 부호화, 제2 무손실 부호화 및 제3 무손실 부호화로부터 제공되는 데이터를 이용하여 상기 비트스트림을 생성하는 단계를 포함하는, 스펙트럼 부호화방법.

#### 청구항 4

제1 항에 있어서, 상기 스펙트럼 데이터는 입력 오디오 신호에 대하여 MDCT(Modified Discrete Cosine Transform) 및 정규화를 수행하여 얻어지는 스펙트럼 부호화방법.

#### 청구항 5

스펙트럼 부호화 장치에 있어서,

적어도 하나의 프로세서를 포함하고, 상기 적어도 하나의 프로세서는,

논제로(non-zero) 밴드의 스펙트럼 데이터를 USQ(Uniform Scalr Quantization)를 사용하여 양자화하고,

상기 양자화된 스펙트럼 데이터의 하위 비트를 추출하고,

모든 논제로 밴드의 상기 양자화된 스펙트럼 데이터의 하위 비트들을 수집하여 하위 비트 시퀀스를 생성하고,

상기 하위 비트 시퀀스를 TCQ(Trellis Coded Quantization)을 사용하여 양자화하고,

상기 양자화된 스펙트럼 데이터 중 상기 하위 비트를 제외한 상위 비트들 및 상기 양자화된 하위 비트 시퀀스에 기초하여 비트스트림을 생성하도록 설정되고,

상기 TCQ에 사용되는 비트수는 각 논제로 밴드의 양자화를 위해 할당된 비트수에서 균등하게 추출되고, 상기 논제로 밴드에 남은 비트수는 USQ에 사용되는, 스펙트럼 부호화 장치.

### 청구항 6

제5 항에 있어서, 상기 스펙트럼 데이터의 대역폭은 초광대역 혹은 전대역인 스펙트럼 부호화 장치.

### 청구항 7

제5 항에 있어서, 상기 적어도 하나의 프로세서는,

상기 논제로 밴드의 중요 스펙트럼 성분의 개수, 위치 및 부호에 대하여 제1 무손실 부호화를 수행하고,

상기 양자화된 스펙트럼 데이터 중 상기 하위 비트를 제외하여 구성되는 크기 정보에 대하여 제2 무손실 부호화를 수행하고,

상기 양자화된 하위 비트 시퀀스에 대하여 제3 무손실 부호화를 수행하고,

상기 제1 무손실 부호화, 제2 무손실 부호화 및 제3 무손실 부호화로부터 제공되는 데이터를 이용하여 상기 비트스트림을 생성하도록 설정되는, 스펙트럼 부호화 장치.

### 청구항 8

제5 항에 있어서, 상기 스펙트럼 데이터는 입력 오디오 신호에 대하여 MDCT(Modified Discrete Cosine Transform) 및 정규화를 수행하여 얻어지는 스펙트럼 부호화 장치.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 오디오 혹은 스피치 신호 부호화 및 복호화에 관한 것으로서, 보다 구체적으로는 주파수 도메인에서 스펙트럼 계수를 부호화 혹은 복호화하는 방법 및 장치에 관한 것이다.

### 배경 기술

[0002] 주파수 도메인에서 스펙트럼 계수들의 효율적인 부호화를 위해 다양한 방식의 양자화기가 제안되어 있다. 예를 들면, TCQ(Trellis Coded Quantization), USQ(Uniform Scalar Quantization), FPC(Factorial Pulse Coding), AVQ(Algebraic VQ), PVQ(Pyramid VQ) 등이 있으며, 각각의 양자화기에 최적화된 무손실 부호화기가 함께 구현될 수 있다.

### 발명의 내용

#### 해결하려는 과제

[0003] 본 발명이 해결하고자 하는 과제는 주파수 도메인에서 다양한 비트율 혹은 다양한 서브밴드의 크기에 적응적으로 스펙트럼 계수를 부호화 혹은 복호화하는 방법 및 장치를 제공하는데 있다.

[0004] 본 발명이 해결하고자 하는 다른 과제는 신호 부호화방법 혹은 복호화방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공하는데 있다.

[0005] 본 발명이 해결하고자 하는 다른 과제는 신호 부호화장치 혹은 복호화장치를 채용하는 멀티미디어 기기를 제공

하는데 있다.

**과제의 해결 수단**

- [0006] 상기 과제를 달성하기 위한 일측면에 따른 스펙트럼 부호화방법은 현재 밴드의 스펙트럼 데이터에 대하여 제1 양자화방식에 따라서 양자화하는 단계; 상기 스펙트럼 데이터 및 상기 양자화된 스펙트럼 데이터를 이용하여 상기 현재 밴드의 하위비트를 생성하는 단계; 상기 현재 밴드의 하위비트가 포함된 하위비트 시퀀스에 대하여 제2 양자화방식에 따라서 양자화하는 단계; 및 상기 양자화된 스펙트럼 데이터 중 N 비트(여기서 N는 1 이상)를 제외한 상위비트 및 상기 양자화된 하위비트 시퀀스를 포함하는 비트스트림을 생성하는 단계를 포함할 수 있다.
- [0007] 상기 과제를 달성하기 위한 일측면에 따른 스펙트럼 부호화장치는 현재 밴드의 스펙트럼 데이터에 대하여 제1 양자화방식에 따라서 양자화하고, 상기 스펙트럼 데이터 및 상기 양자화된 스펙트럼 데이터를 이용하여 상기 현재 밴드의 하위비트를 생성하고, 상기 현재 밴드의 하위비트가 포함된 하위비트 시퀀스에 대하여 제2 양자화방식에 따라서 양자화하고, 상기 양자화된 스펙트럼 데이터 중 N 비트(여기서 N는 1 이상)를 제외한 상위비트 및 상기 양자화된 하위비트 시퀀스를 포함하는 비트스트림을 생성하는 프로세서를 포함할 수 있다.
- [0008] 상기 과제를 달성하기 위한 일측면에 따른 스펙트럼 복호화방법은 비트스트림으로부터 ISC 정보를 추출하여 ISC의 갯수, 위치 및 부호를 디코딩하는 단계; 상기 비트스트림으로부터 하위비트를 제외한 나머지 비트를 추출하여 디코딩하는 단계; 상기 비트스트림으로부터 TCQ 경로 정보를 추출하여 하위비트 시퀀스를 디코딩하는 단계; 및 디코딩된 하위비트를 제외한 나머지 비트와 디코딩된 하위비트 시퀀스에 근거하여 스펙트럼 성분을 복원하는 단계를 포함할 수 있다.
- [0009] 상기 과제를 달성하기 위한 일측면에 따른 스펙트럼 복호화장치는 비트스트림으로부터 ISC 정보를 추출하여 ISC의 갯수, 위치 및 부호를 디코딩하고, 상기 비트스트림으로부터 하위비트를 제외한 나머지 비트를 추출하여 디코딩하고, 상기 비트스트림으로부터 TCQ 경로 정보를 추출하여 하위비트 시퀀스를 디코딩하고, 디코딩된 하위비트를 제외한 나머지 비트와 디코딩된 하위비트 시퀀스에 근거하여 스펙트럼 성분을 복원하는 프로세서를 포함할 수 있다.

**발명의 효과**

- [0010] 다양한 비트율과 다양한 서브밴드의 크기에 적응적인 스펙트럼 계수의 부호화 및 복호화가 가능하다. 또한, 멀티레이트를 지원하는 코덱에서 설계된 비트레이트 제어모듈을 이용하여 스펙트럼 계수를 USQ와 TCQ를 결합한 방식에 근거하여 부호화 혹은 복호화함으로써, 각 양자화방식의 이점을 극대화시킬 수 있다.

**도면의 간단한 설명**

- [0011] 도 1a 및 도 1b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 일예에 따른 구성을 각각 나타낸 블록도이다.
- 도 2a 및 도 2b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 다른 예에 따른 구성을 각각 나타낸 블록도이다.
- 도 3a 및 도 3b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 다른 예에 따른 구성을 각각 나타낸 블록도이다.
- 도 4a 및 도 4b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 다른 예에 따른 구성을 각각 나타낸 블록도이다.
- 도 5는 본 발명이 적용될 수 있는 주파수 도메인 오디오 부호화장치의 구성을 나타낸 블록도이다.
- 도 6은 본 발명이 적용될 수 있는 주파수 도메인 오디오 복호화장치의 구성을 나타낸 블록도이다.
- 도 7은 일실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블록도이다.
- 도 8은 서브밴드 분할의 예를 나타내는 도면이다.
- 도 9는 일실시예에 따른 스펙트럼 양자화장치의 구성을 나타내는 블록도이다.
- 도 10은 일실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블록도이다.

- 도 11은 일실시예에 따른 ISC 부호화장치의 구성을 나타내는 블록도이다.
- 도 12는 일실시예에 따른 ISC 정보 부호화장치의 구성을 나타내는 블록도이다.
- 도 13는 다른 실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블록도이다.
- 도 14는 다른 실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블록도이다.
- 도 15는 일실시예에 따른 ISC 수집 및 부호화과정의 개념을 나타내는 도면이다.
- 도 16은 USQ와 TCQ를 결합한 제2 조인트방식을 설명하는 도면이다.
- 도 17은 다른 실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블록도이다.
- 도 18은 도 17에 도시된 제2 양자화부의 구성을 나타내는 블록도이다.
- 도 19는 레지듀얼 데이터를 생성하는 방법을 설명하는 도면이다.
- 도 20은 TCQ의 일예를 나타내는 도면이다.
- 도 21은 일실시예에 따른 주파수 도메인 오디오 복호화장치의 구성을 나타낸 블록도이다.
- 도 22는 일실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다.
- 도 23은 일실시예에 따른 스펙트럼 역양자화장치의 구성을 나타내는 블록도이다.
- 도 24는 일실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다.
- 도 25는 일실시예에 따른 ISC 복호화장치의 구성을 나타내는 블록도이다.
- 도 26은 일실시예에 따른 ISC 정보 복호화장치의 구성을 나타내는 블록도이다.
- 도 27은 다른 실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다.
- 도 28은 다른 실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다.
- 도 29는 다른 실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다.
- 도 30은 도 29에 도시된 제3 디코딩부의 구성을 나타내는 블록도이다.
- 도 31은 일실시예에 따른 멀티미디어 기기의 구성을 나타낸 블록도이다.
- 도 32는 다른 실시예에 따른 멀티미디어 기기의 구성을 나타낸 블록도이다.
- 도 33은 다른 실시예에 따른 멀티미디어 기기의 구성을 나타낸 블록도이다.
- 도 34는 일실시예에 따른 스펙트럼 부호화방법을 설명하는 흐름도이다.
- 도 35는 일실시예에 따른 스펙트럼 복호화방법을 설명하는 흐름도이다.
- 도 36은 일실시예에 따른 비트할당장치의 구성을 나타낸 블록도이다.
- 도 37은 일실시예에 따른 부호화 모드 결정장치의 구성을 나타낸 블록도이다.
- 도 38은 도 37에 도시된 정정부에서 사용되는 상태 머신을 나타내는 도면이다.

**발명을 실시하기 위한 구체적인 내용**

- [0012] 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.
- [0013] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 구성요소들이 용어들에 의해 한정되는 것은 아니다. 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.
- [0014] 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.

[0015] 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.

[0016] 이하, 본 발명의 실시예들을 첨부 도면을 참조하여 상세히 설명하기로 한다.

[0017] 도 1a 및 도 1b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 일예에 따른 구성을 각각 나타낸 블록도이다.

[0018] 도 1a에 도시된 오디오 부호화장치(110)는 전처리부(112), 주파수도메인 부호화부(114), 및 파라미터 부호화부(116)를 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.

[0019] 도 1a에 있어서, 전처리부(112)는 입력신호에 대하여 필터링 혹은 다운샘플링 등을 수행할 수 있으나, 이에 한정되는 것은 아니다. 입력신호는 오디오, 뮤직 혹은 스피치, 혹은 이들의 혼합신호를 나타내는 사운드 등의 미디어 신호를 의미할 수 있으나, 이하에서는 설명의 편의를 위하여 오디오 신호로서 지칭하기로 한다.

[0020] 주파수도메인 부호화부(114)는 전처리부(112)로부터 제공되는 오디오 신호에 대하여 시간-주파수 변환을 수행하고, 오디오 신호의 채널 수, 부호화대역 및 비트율에 대응하여 부호화 톨을 선택하고, 선택된 부호화 톨을 이용하여 오디오 신호에 대한 부호화를 수행할 수 있다. 시간-주파수 변환은 MDCT(Modified Discrete Cosine Transform), MLT(Modulated Lapped Transform) 혹은 FFT(Fast Fourier Transform)를 사용하나, 이에 한정되는 것은 아니다. 여기서, 주어진 비트수가 충분한 경우 전체 대역에 대하여 일반적인 변환 부호화방식을 적용하며, 주어진 비트수가 충분하지 않은 경우 일부 대역에 대해서는 대역확장방식을 적용할 수 있다. 한편, 오디오 신호가 스테레오 혹은 멀티채널인 경우, 주어진 비트수가 충분하면 각 채널별로 부호화하고, 충분하지 않으면 다운믹싱방식을 적용할 수 있다. 주파수도메인 부호화부(114)로부터는 부호화된 스펙트럼 계수가 생성된다.

[0021] 파라미터 부호화부(116)는 주파수도메인 부호화부(114)로부터 제공되는 부호화된 스펙트럼 계수로부터 파라미터를 추출하고, 추출된 파라미터를 부호화할 수 있다. 파라미터는 예를 들어 서브밴드 혹은 밴드 별로 추출될 수 있으며, 이하에서는 설명의 간소화를 위하여 서브밴드라고 지칭하기로 한다. 각 서브밴드는 스펙트럼 계수들을 그루핑한 단위로써, 임계대역을 반영하여 균일 혹은 비균일 길이를 가질 수 있다. 비균일 길이를 가지는 경우, 저주파수 대역에 존재하는 서브밴드의 경우 고주파수 대역에서와 비교하여 상대적으로 적은 길이를 가질 수 있다. 한 프레임에 포함되는 서브밴드의 개수 및 길이는 코덱 알고리즘에 따라서 달라지며 부호화 성능에 영향을 미칠 수 있다. 한편, 파라미터는 서브밴드의 스케일팩터, 파워, 평균 에너지, 혹은 norm을 예로 들 수 있으나, 이에 한정되는 것은 아니다. 부호화결과 얻어지는 스펙트럼 계수와 파라미터는 비트스트림을 형성하며, 저장매체에 저장되거나 채널을 통하여 예를 들어 패킷 형태로 전송될 수 있다.

[0022] 도 1b에 도시된 오디오 복호화장치(130)는 파라미터 복호화부(132), 주파수도메인 복호화부(134), 및 후처리부(136)를 포함할 수 있다. 여기서, 주파수도메인 복호화부(134)는 프레임 소거 은닉(FEC: frame erasure concealment) 알고리즘 혹은 패킷 손실 은닉(PLC: packet loss concealment) 알고리즘을 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.

[0023] 도 1b에 있어서, 파라미터 복호화부(132)는 수신된 비트스트림으로부터 부호화된 파라미터를 복호화하고, 복호화된 파라미터로부터 프레임 단위로 소거 혹은 손실과 같은 에러가 발생하였는지를 체크할 수 있다. 에러 체크는 공지된 다양한 방법을 사용할 수 있으며, 현재 프레임이 정상 프레임인지 소거 혹은 손실 프레임인지에 대한 정보를 주파수도메인 복호화부(134)로 제공한다. 이하에서는 설명의 간소화를 위하여 소거 혹은 손실 프레임을 에러프레임으로 지칭하기로 한다.

[0024] 주파수도메인 복호화부(134)는 현재 프레임이 정상 프레임인 경우 일반적인 변환 복호화과정을 통하여 복호화를 수행하여 합성된 스펙트럼 계수를 생성할 수 있다. 한편, 주파수도메인 복호화부(134)는 현재 프레임이 에러프

레이민 경우 FEC 혹은 PLC 알고리즘을 통하여 이전 정상 프레임의 스펙트럼 계수를 에러프레임에 반복하여 사용하거나 회귀분석을 통하여 스케일링하여 반복함으로써, 합성된 스펙트럼 계수를 생성할 수 있다. 주파수도메인 복호화부(134)는 합성된 스펙트럼 계수에 대하여 주파수-시간 변환을 수행하여 시간도메인 신호를 생성할 수 있다.

- [0025] 후처리부(136)는 주파수도메인 복호화부(134)로부터 제공되는 시간도메인 신호에 대하여 음질 향상을 위한 필터링 혹은 업샘플링 등을 수행할 수 있으나, 이에 한정되는 것은 아니다. 후처리부(136)는 출력신호로서 복원된 오디오 신호를 제공한다.
- [0026] 도 2a 및 도 2b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 다른 예에 따른 구성을 각각 나타낸 블록도로서, 스위칭 구조를 가진다.
- [0027] 도 2a에 도시된 오디오 부호화장치(210)는 전처리부(212), 모드결정부(213), 주파수도메인 부호화부(214), 시간도메인 부호화부(215) 및 파라미터 부호화부(216)을 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.
- [0028] 도 2a에 있어서, 전처리부(212)는 도 1a의 전처리부(112)와 실질적으로 동일하므로 설명을 생략하기로 한다.
- [0029] 모드결정부(213)는 입력신호의 특성을 참조하여 부호화 모드를 결정할 수 있다. 입력신호의 특성에 따라서 현재 프레임에 적합한 부호화 모드가 음성모드인지 또는 음악모드인지 여부를 결정할 수 있고, 또한 현재 프레임에 효율적인 부호화 모드가 시간도메인 모드인지 아니면 주파수도메인 모드인지를 결정할 수 있다. 여기서, 프레임의 단구간 특성 혹은 복수의 프레임들에 대한 장구간 특성 등을 이용하여 입력신호의 특성을 파악할 수 있으나, 이에 한정되는 것은 아니다. 예를 들면, 입력신호가 음성신호에 해당하면 음성모드 혹은 시간도메인 모드로 결정하고, 입력신호가 음성신호 이외의 신호 즉, 음악신호 혹은 혼합신호에 해당하면 음악모드 혹은 주파수도메인 모드로 결정할 수 있다. 모드결정부(213)는 입력신호의 특성이 음악모드 혹은 주파수도메인 모드에 해당하는 경우에는 전처리부(212)의 출력신호를 주파수도메인 부호화부(214)로, 입력신호의 특성이 음성모드 혹은 시간도메인 모드에 해당하는 경우 시간도메인 부호화부(215)로 제공할 수 있다.
- [0030] 주파수도메인 부호화부(214)는 도 1a의 주파수도메인 부호화부(114)와 실질적으로 동일하므로 설명을 생략하기로 한다.
- [0031] 시간도메인 부호화부(215)는 전처리부(212)로부터 제공되는 오디오 신호에 대하여 CELP(Code Excited Linear Prediction) 부호화를 수행할 수 있다. 구체적으로, ACELP(Algebraic CELP)를 사용할 수 있으나, 이에 한정되는 것은 아니다.
- [0032] 파라미터 부호화부(216)는 주파수도메인 부호화부(214) 혹은 시간도메인 부호화부(215)로부터 제공되는 부호화된 스펙트럼 계수로부터 파라미터를 추출하고, 추출된 파라미터를 부호화한다. 파라미터 부호화부(216)는 도 1a의 파라미터 부호화부(116)와 실질적으로 동일하므로 설명을 생략하기로 한다. 부호화결과 얻어지는 스펙트럼 계수와 파라미터는 부호화 모드 정보와 함께 비트스트림을 형성하며, 채널을 통하여 패킷 형태로 전송되거나 저장매체에 저장될 수 있다.
- [0033] 도 2b에 도시된 오디오 복호화장치(230)는 파라미터 복호화부(232), 모드결정부(233), 주파수도메인 복호화부(234), 시간도메인 복호화부(235) 및 후처리부(236)를 포함할 수 있다. 여기서, 주파수도메인 복호화부(234)와 시간도메인 복호화부(235)는 각각 해당 도메인에서의 FEC 혹은 PLC 알고리즘을 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.
- [0034] 도 2b에 있어서, 파라미터 복호화부(232)는 패킷 형태로 전송되는 비트스트림으로부터 파라미터를 복호화하고, 복호화된 파라미터로부터 프레임 단위로 에러가 발생하였는지를 체크할 수 있다. 에러 체크는 공지된 다양한 방법을 사용할 수 있으며, 현재 프레임이 정상 프레임인지 에러프레임인지에 대한 정보를 주파수도메인 복호화부(234) 혹은 시간도메인 복호화부(235)로 제공한다.
- [0035] 모드결정부(233)는 비트스트림에 포함된 부호화 모드 정보를 체크하여 현재 프레임을 주파수도메인 복호화부(234) 혹은 시간도메인 복호화부(235)로 제공한다.
- [0036] 주파수도메인 복호화부(234)는 부호화 모드가 음악모드 혹은 주파수도메인 모드인 경우 동작하며, 현재 프레임이 정상 프레임인 경우 일반적인 변환 복호화과정을 통하여 복호화를 수행하여 합성된 스펙트럼 계수를 생성한다. 한편, 현재 프레임이 에러프레임이고, 이전 프레임의 부호화 모드가 음악모드 혹은 주파수도메인 모드인 경우 주파수 도메인에서의 FEC 혹은 PLC 알고리즘을 통하여 이전 정상 프레임의 스펙트럼 계수를 에러프레임에



반복하여 사용하거나 회귀분석을 통하여 스케일링하여 반복함으로써, 합성된 스펙트럼 계수를 생성할 수 있다. 주파수도메인 복호화부(234)는 합성된 스펙트럼 계수에 대하여 주파수-시간 변환을 수행하여 시간도메인 신호를 생성할 수 있다.

- [0037] 시간도메인 복호화부(235)는 부호화 모드가 음성모드 혹은 시간도메인 모드인 경우 동작하며, 현재 프레임이 정상 프레임인 경우 일반적인 CELP 복호화과정을 통하여 복호화를 수행하여 시간도메인 신호를 생성한다. 한편, 현재 프레임이 에러프레임이고, 이전 프레임의 부호화 모드가 음성모드 혹은 시간도메인 모드인 경우 시간 도메인에서의 FEC 혹은 PLC 알고리즘을 수행할 수 있다.
- [0038] 후처리부(236)는 주파수도메인 복호화부(234) 혹은 시간도메인 복호화부(235)로부터 제공되는 시간도메인 신호에 대하여 필터링 혹은 업샘플링 등을 수행할 수 있으나, 이에 한정되는 것은 아니다. 후처리부(236)는 출력신호로서 복원된 오디오신호를 제공한다.
- [0039] 도 3a 및 도 3b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 다른 예에 따른 구성을 각각 나타낸 블록도로서, 스위칭 구조를 가진다.
- [0040] 도 3a에 도시된 오디오 부호화장치(310)는 전처리부(312), LP(Linear Prediction) 분석부(313), 모드결정부(314), 주파수도메인 여기부호화부(315), 시간도메인 여기부호화부(316) 및 파라미터 부호화부(317)을 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.
- [0041] 도 3a에 있어서, 전처리부(312)는 도 1a의 전처리부(112)와 실질적으로 동일하므로 설명을 생략하기로 한다.
- [0042] LP 분석부(313)는 입력신호에 대하여 LP 분석을 수행하여 LP 계수를 추출하고, 추출된 LP 계수로부터 여기신호를 생성한다. 여기신호는 부호화 모드에 따라서 주파수도메인 여기부호화부(315)와 시간도메인 여기부호화부(316) 중 하나로 제공될 수 있다.
- [0043] 모드결정부(314)는 도 2b의 모드결정부(213)와 실질적으로 동일하므로 설명을 생략하기로 한다.
- [0044] 주파수도메인 여기부호화부(315)는 부호화 모드가 음악모드 혹은 주파수도메인 모드인 경우 동작하며, 입력신호가 여기신호인 것을 제외하고는 도 1a의 주파수도메인 부호화부(114)와 실질적으로 동일하므로 설명을 생략하기로 한다.
- [0045] 시간도메인 여기부호화부(316)는 부호화 모드가 음성모드 혹은 시간도메인 모드인 경우 동작하며, 입력신호가 여기신호인 것을 제외하고는 도 2a의 시간도메인 부호화부(215)와 실질적으로 동일하므로 설명을 생략하기로 한다.
- [0046] 파라미터 부호화부(317)는 주파수도메인 여기부호화부(315) 혹은 시간도메인 여기부호화부(316)로부터 제공되는 부호화된 스펙트럼 계수로부터 파라미터를 추출하고, 추출된 파라미터를 부호화한다. 파라미터 부호화부(317)는 도 1a의 파라미터 부호화부(116)와 실질적으로 동일하므로 설명을 생략하기로 한다. 부호화결과 얻어지는 스펙트럼 계수와 파라미터는 부호화 모드 정보와 함께 비트스트림을 형성하며, 채널을 통하여 패킷 형태로 전송되거나 저장매체에 저장될 수 있다.
- [0047] 도 3b에 도시된 오디오 복호화장치(330)는 파라미터 복호화부(332), 모드결정부(333), 주파수도메인 여기복호화부(334), 시간도메인 여기복호화부(335), LP 합성부(336) 및 후처리부(337)을 포함할 수 있다. 여기서, 주파수도메인 여기복호화부(334)와 시간도메인 여기복호화부(335)는 각각 해당 도메인에서의 FEC 혹은 PLC 알고리즘을 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.
- [0048] 도 3b에 있어서, 파라미터 복호화부(332)는 패킷 형태로 전송되는 비트스트림으로부터 파라미터를 복호화하고, 복호화된 파라미터로부터 프레임 단위로 에러가 발생하였는지를 체크할 수 있다. 에러 체크는 공지된 다양한 방법을 사용할 수 있으며, 현재 프레임이 정상 프레임인지 에러프레임인지에 대한 정보를 주파수도메인 여기복호화부(334) 혹은 시간도메인 여기복호화부(335)로 제공한다.
- [0049] 모드결정부(333)는 비트스트림에 포함된 부호화 모드 정보를 체크하여 현재 프레임을 주파수도메인 여기복호화부(334) 혹은 시간도메인 여기복호화부(335)로 제공한다.
- [0050] 주파수도메인 여기복호화부(334)는 부호화 모드가 음악모드 혹은 주파수도메인 모드인 경우 동작하며, 현재 프레임이 정상 프레임인 경우 일반적인 변환 복호화과정을 통하여 복호화를 수행하여 합성된 스펙트럼 계수를 생

성한다. 한편, 현재 프레임이 에러프레임이고, 이전 프레임의 부호화 모드가 음악모드 혹은 주파수도메인 모드인 경우 주파수 도메인에서의 FEC 혹은 PLC 알고리즘을 통하여 이전 정상 프레임의 스펙트럼 계수를 에러프레임에 반복하여 사용하거나 회귀분석을 통하여 스케일링하여 반복함으로써, 합성된 스펙트럼 계수를 생성할 수 있다. 주파수도메인 여기복호화부(334)는 합성된 스펙트럼 계수에 대하여 주파수-시간 변환을 수행하여 시간도메인 신호인 여기신호를 생성할 수 있다.

- [0051] 시간도메인 여기복호화부(335)는 부호화 모드가 음성모드 혹은 시간도메인 모드인 경우 동작하며, 현재 프레임이 정상 프레임인 경우 일반적인 CELP 복호화과정을 통하여 복호화를 수행하여 시간도메인 신호인 여기신호를 생성한다. 한편, 현재 프레임이 에러프레임이고, 이전 프레임의 부호화 모드가 음성모드 혹은 시간도메인 모드인 경우 시간 도메인에서의 FEC 혹은 PLC 알고리즘을 수행할 수 있다.
- [0052] LP 합성부(336)는 주파수도메인 여기복호화부(334) 혹은 시간도메인 여기복호화부(335)로부터 제공되는 여기신호에 대하여 LP 합성을 수행하여 시간도메인 신호를 생성한다.
- [0053] 후처리부(337)는 LP 합성부(336)로부터 제공되는 시간도메인 신호에 대하여 필터링 혹은 업샘플링 등을 수행할 수 있으나, 이에 한정되는 것은 아니다. 후처리부(337)는 출력신호로서 복원된 오디오신호를 제공한다.
- [0054] 도 4a 및 도 4b는 본 발명이 적용될 수 있는 오디오 부호화장치 및 복호화장치의 다른 예에 따른 구성을 각각 나타낸 블록도로서, 스위칭 구조를 가진다.
- [0055] 도 4a에 도시된 오디오 부호화장치(410)는 전처리부(412), 모드결정부(413), 주파수도메인 부호화부(414), LP 분석부(415), 주파수도메인 여기부호화부(416), 시간도메인 여기부호화부(417) 및 파라미터 부호화부(418)를 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다. 도 4a에 도시된 오디오 부호화장치(410)는 도 2a의 오디오 부호화장치(210)와 도 3a의 오디오 부호화장치(310)를 결합한 것으로 볼 수 있으므로, 공통되는 부분의 동작 설명은 생략하는 한편, 모드결정부(413)의 동작에 대하여 설명하기로 한다.
- [0056] 모드결정부(413)는 입력신호의 특성 및 비트율을 참조하여 입력신호의 부호화모드를 결정할 수 있다. 모드결정부(413)는 입력신호의 특성에 따라서 현재 프레임이 음성모드인지 또는 음악모드인지에 따라서, 또한 현재 프레임에 효율적인 부호화 모드가 시간도메인 모드인지 아니면 주파수도메인 모드인지에 따라서 CELP 모드와 그외의 모드로 결정할 수 있다. 만약, 입력신호의 특성이 음성모드인 경우에는 CELP 모드로 결정하고, 음악모드이면서 고비트율인 경우 FD 모드로 결정하고, 음악모드이면서 저비트율인 경우 오디오 모드로 결정할 수 있다. 모드결정부(413)는 FD 모드인 경우 입력신호를 주파수도메인 부호화부(414)로, 오디오 모드인 경우 LP 분석부(415)를 통하여 주파수도메인 여기부호화부(416)로, CELP 모드인 경우 LP 분석부(415)를 통하여 시간도메인 여기부호화부(417)로 제공할 수 있다.
- [0057] 주파수도메인 부호화부(414)는 도 1a의 오디오 부호화장치(110)의 주파수도메인 부호화부(114) 혹은 도 2a의 오디오 부호화장치(210)의 주파수도메인 부호화부(214)에, 주파수도메인 여기부호화부(416) 혹은 시간도메인 여기부호화부(417)는 도 3a의 오디오 부호화장치(310)의 주파수도메인 여기부호화부(315) 혹은 시간도메인 여기부호화부(316)에 대응될 수 있다.
- [0058] 도 4b에 도시된 오디오 복호화장치(430)는 파라미터 복호화부(432), 모드결정부(433), 주파수도메인 복호화부(434), 주파수도메인 여기복호화부(435), 시간도메인 여기복호화부(436), LP 합성부(437) 및 후처리부(438)를 포함할 수 있다. 여기서, 주파수도메인 복호화부(434), 주파수도메인 여기복호화부(435)와 시간도메인 여기복호화부(436)는 각각 해당 도메인에서의 FEC 혹은 PLC 알고리즘을 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다. 도 4b에 도시된 오디오 복호화장치(430)는 도 2b의 오디오 복호화장치(230)와 도 3b의 오디오 복호화장치(330)를 결합한 것으로 볼 수 있으므로, 공통되는 부분의 동작 설명은 생략하는 한편, 모드결정부(433)의 동작에 대하여 설명하기로 한다.
- [0059] 모드결정부(433)는 비트스트림에 포함된 부호화 모드 정보를 체크하여 현재 프레임을 주파수도메인 복호화부(434), 주파수도메인 여기복호화부(435) 혹은 시간도메인 여기복호화부(436)로 제공한다.
- [0060] 주파수도메인 복호화부(434)는 도 1b의 오디오 부호화장치(130)의 주파수도메인 복호화부(134) 혹은 도 2b의 오디오 복호화장치(230)의 주파수도메인 복호화부(234)에, 주파수도메인 여기복호화부(435) 혹은 시간도메인 여기복호화부(436)는 도 3b의 오디오 복호화장치(330)의 주파수도메인 여기복호화부(334) 혹은 시간도메인 여기복호화부(335)에 대응될 수 있다.

- [0061] 도 5는 본 발명이 적용되는 주파수 도메인 오디오 부호화장치의 구성을 나타낸 블록도이다.
- [0062] 도 5에 도시된 주파수 도메인 오디오 부호화장치(510)는 트랜지언트 검출부(511), 변환부(512), 신호분류부(513), 에너지 부호화부(514), 스펙트럼 정규화부(515), 비트할당부(516), 스펙트럼 부호화부(517) 및 다중화부(518)를 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다. 여기서, 주파수 도메인 오디오 부호화장치(510)는 도 2에 도시된 주파수 도메인 부호화부(214)의 모든 기능과 파라미터 부호화부(216)의 일부 기능을 수행할 수 있다. 한편, 주파수 도메인 오디오 부호화장치(510)는 신호분류부(513)를 제외하고는 ITU-T G.719 표준에 개시된 엔코더의 구성으로 대체될 수 있으며, 이때 변환부(512)는 50%의 오버랩 구간을 갖는 변환윈도우를 사용할 수 있다. 또한, 주파수 도메인 오디오 부호화장치(510)는 트랜지언트 검출부(511)와 신호분류부(513)를 제외하고는 ITU-T G.719 표준에 개시된 엔코더의 구성으로 대체될 수 있다. 각 경우에 있어서, 도시되지는 않았으나, ITU-T G.719 표준에서와 같이 스펙트럼 부호화부(517)의 후단에 노이즈 레벨 추정부를 더 구비하여, 비트할당 과정에서 제로 비트가 할당된 스펙트럼 계수를 위한 노이즈 레벨을 추정하여 비트스트림에 포함시킬 수 있다.
- [0063] 도 5를 참조하면, 트랜지언트 검출부(511)는 입력 신호를 분석하여 트랜지언트 특성을 나타내는 구간을 검출하고, 검출 결과에 대응하여 각 프레임에 대한 트랜지언트 시그널링 정보를 생성할 수 있다. 이때, 트랜지언트 구간의 검출에는 공지된 다양한 방법을 사용할 수 있다. 일실시예에 따르면, 트랜지언트 검출부(511)는 먼저 현재 프레임이 트랜지언트 프레임인지를 1차적으로 판단하고, 트랜지언트 프레임으로 판단된 현재 프레임에 대하여 2차적으로 검증을 수행할 수 있다. 트랜지언트 시그널링 정보는 다중화부(518)를 통하여 비트스트림에 포함되는 한편, 변환부(512)로 제공될 수 있다.
- [0064] 변환부(512)는 트랜지언트 구간의 검출 결과에 따라서, 변환에 사용되는 윈도우 사이즈를 결정하고, 결정된 윈도우 사이즈에 근거하여 시간-주파수 변환을 수행할 수 있다. 일례로서, 트랜지언트 구간이 검출된 서브밴드의 경우 단구간 윈도우(short window)를, 검출되지 않은 서브밴드의 경우 장구간 윈도우(long window)를 적용할 수 있다. 다른 예로서, 트랜지언트 구간을 포함하는 프레임에 대하여 단구간 윈도우를 적용할 수 있다.
- [0065] 신호분류부(513)는 변환부(512)로부터 제공되는 스펙트럼을 프레임 단위로 분석하여 각 프레임이 하모닉 프레임에 해당하는지 여부를 판단할 수 있다. 이때, 하모닉 프레임의 판단에는 공지된 다양한 방법을 사용할 수 있다. 일실시예에 따르면, 신호분류부(513)는 변환부(512)로부터 제공되는 스펙트럼을 복수의 서브밴드로 나누고, 각 서브밴드에 대하여 에너지의 피크치와 평균치를 구할 수 있다. 다음, 각 프레임에 대하여 에너지의 피크치가 평균치보다 소정 비율 이상 큰 서브밴드의 갯수를 구하고, 구해진 서브밴드의 갯수가 소정값 이상인 프레임을 하모닉 프레임으로 결정할 수 있다. 여기서, 소정 비율 및 소정값은 실험 혹은 시뮬레이션을 통하여 미리 결정될 수 있다. 하모닉 시그널링 정보는 다중화부(518)를 통하여 비트스트림에 포함될 수 있다.
- [0066] 에너지 부호화부(514)는 각 서브밴드 단위로 에너지를 구하여 양자화 및 무손실 부호화할 수 있다. 일실시예에 따르면, 에너지로서, 각 서브밴드의 평균 스펙트럼 에너지에 해당하는 Norm 값을 사용할 수 있으며, 스케일 팩터 혹은 파워를 대신 사용할 수 있으나 이에 한정되는 것은 아니다. 여기서, 각 서브밴드의 Norm 값은 스펙트럼 정규화부(515) 및 비트할당부(516)로 제공되는 한편, 다중화부(518)를 통하여 비트스트림에 포함될 수 있다.
- [0067] 스펙트럼 정규화부(515)는 각 서브밴드 단위로 구해진 Norm 값을 이용하여 스펙트럼을 정규화할 수 있다.
- [0068] 비트할당부(516)는 각 서브밴드 단위로 구해진 Norm 값을 이용하여 정수 단위 혹은 소수점 단위로 비트 할당을 수행할 수 있다. 또한, 비트할당부(516)는 각 서브밴드 단위로 구해진 Norm 값을 이용하여 마스킹 임계치를 계산하고, 마스킹 임계치를 이용하여 지각적으로 필요한 비트수 즉, 허용비트수를 추정할 수 있다. 다음, 비트할당부(516)는 각 서브밴드에 대하여 할당 비트수가 허용 비트수를 초과하지 않도록 제한할 수 있다. 한편, 비트할당부(516)는 Norm 값이 큰 서브밴드에서부터 순차적으로 비트를 할당하며, 각 서브밴드의 Norm 값에 대하여 각 서브밴드의 지각적 중요도에 따라서 가중치를 부여함으로써 지각적으로 중요한 서브밴드에 더 많은 비트가 할당되도록 조정할 수 있다. 이때, Norm 부호화부(514)로부터 비트할당부(516)로 제공되는 양자화된 Norm 값은 ITU-T G.719 에서와 마찬가지로 심리음향가중(psycho-acoustical weighting) 및 마스킹 효과를 고려하기 위하여 미리 조정된 다음 비트할당에 사용될 수 있다.
- [0069] 스펙트럼 부호화부(517)는 정규화된 스펙트럼에 대하여 각 서브밴드의 할당 비트수를 이용하여 양자화를 수행하고, 양자화된 결과에 대하여 무손실 부호화할 수 있다. 일례로서, 스펙트럼 부호화에 TCQ(Trellis Coded Quantizer), USQ(Uniform Scalar Quantizer), FPC(Factorial Pulse Coder), AVQ(Analog Vector Quantizer), PVQ(Predictive Vector Quantizer) 혹은 이들의 조합과, 각 양자화기에 대응되는 무손실 부호화기를 사용할 수

있다. 또한, 해당 코덱이 탑재되는 환경 혹은 사용자의 필요에 따라서 다양한 스펙트럼 부호화기법을 적용할 수 있다. 스펙트럼 부호화부(517)에서 부호화된 스펙트럼에 대한 정보는 다중화부(518)를 통하여 비트스트림에 포함될 수 있다.

- [0070] 도 6은 본 발명이 적용되는 주파수 도메인 오디오 부호화장치의 구성을 나타낸 블록도이다.
- [0071] 도 6에 도시된 오디오 부호화장치(600)는 전처리부(610), 주파수도메인 부호화부(630), 시간도메인 부호화부(650) 및 다중화부(670)을 포함할 수 있다. 주파수도메인 부호화부(630)는 트랜지언트 검출부(631), 변환부(633) 및 스펙트럼 부호화부(6355)를 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.
- [0072] 도 6에 있어서, 전처리부(610)는 입력신호에 대하여 필터링 혹은 다운샘플링 등을 수행할 수 있으나, 이에 한정되는 것은 아니다. 전처리부(610)는 신호 특성에 근거하여 부호화 모드를 결정할 수 있다. 신호 특성에 따라서 현재 프레임에 적합한 부호화 모드가 음성모드인지 또는 음악모드인지 여부를 결정할 수 있고, 또한 현재 프레임에 효율적인 부호화 모드가 시간도메인 모드인지 아니면 주파수도메인 모드인지를 결정할 수 있다. 여기서, 프레임의 단구간 특성 혹은 복수의 프레임들에 대한 장구간 특성 등을 이용하여 신호 특성을 파악할 수 있으나, 이에 한정되는 것은 아니다. 예를 들면, 입력신호가 음성신호에 해당하면 음성모드 혹은 시간도메인 모드로 결정하고, 입력신호가 음성신호 이외의 신호 즉, 음악신호 혹은 혼합신호에 해당하면 음악모드 혹은 주파수도메인 모드로 결정할 수 있다. 전처리부(610)는 신호 특성이 음악모드 혹은 주파수도메인 모드에 해당하는 경우에는 입력신호를 주파수도메인 부호화부(630)로, 신호 특성이 음성모드 혹은 시간도메인 모드에 해당하는 경우 입력신호를 시간도메인 부호화부(650)로 제공할 수 있다.
- [0073] 주파수도메인 부호화부(630)는 전처리부(610)로부터 제공되는 오디오 신호를 변환부호화에 근거하여 처리할 수 있다. 구체적으로, 트랜지언트 검출부(631)는 오디오신호로부터 트랜지언트 성분을 검출하여 현재 프레임이 트랜지언트 프레임인지를 판단할 수 있다. 변환부(633)는 트랜지언트 검출부(631)로부터 제공되는 프레임 타입 즉, 트랜지언트 정보에 근거하여 변환 윈도우의 길이 혹은 형태를 결정하고, 결정된 변환 윈도우에 근거하여 오디오 신호를 주파수 도메인으로 변환할 수 있다. 변환기법으로는 MDCT, FFT, 혹은 MLT를 적용할 수 있다. 일반적으로, 트랜지언트 성분을 갖는 프레임에 대해서는 짧은 길이의 변환 윈도우를 적용할 수 있다. 스펙트럼 부호화부(635)는 주파수 도메인으로 변환된 오디오 스펙트럼에 대하여 부호화를 수행할 수 있다. 스펙트럼 부호화부(635)에 대해서는 도 7 및 도 9를 참조하여 좀 더 구체적으로 설명하기로 한다.
- [0074] 시간도메인 부호화부(650)는 전처리부(610)로부터 제공되는 오디오 신호에 대하여 CELP(Code Excited Linear Prediction) 부호화를 수행할 수 있다. 구체적으로, ACELP(Algebraic CELP)를 사용할 수 있으나, 이에 한정되는 것은 아니다.
- [0075] 다중화부(670)는 주파수도메인 부호화부(630) 혹은 시간도메인 부호화부(650)에서 부호화결과 생성되는 스펙트럼 성분 혹은 신호 성분과 다양한 인덱스를 다중화하여 비트스트림을 생성하며, 비트스트림은 채널을 통하여 패킷 형태로 전송되거나 저장매체에 저장될 수 있다.
- [0076] 도 7은 일실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블록도이다. 도 7에 도시된 장치는 도 6의 스펙트럼 부호화부(635)에 대응되거나, 다른 주파수도메인 부호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0077] 도 7에 도시된 스펙트럼 부호화장치(700)는 에너지 추정부(710), 에너지 양자화 및 부호화부(720), 비트할당부(730), 스펙트럼 정규화부(740), 스펙트럼 양자화 및 부호화부(750) 및 노이즈 필링부(760)를 포함할 수 있다.
- [0078] 도 7을 참조하면, 에너지 추정부(710)는 원래의 스펙트럼 계수들에 대하여 서브밴드로 분리하고, 각 서브밴드별 에너지, 예를 들면 Norm값을 추정할 수 있다. 여기서, 하나의 프레임에서 각 서브밴드는 동일한 크기를 갖거나, 저대역에서부터 고대역으로 갈수록 각 서브밴드에 포함되는 스펙트럼 계수들의 갯수를 증가시킬 수 있다.
- [0079] 에너지 양자화 및 부호화부(720)는 각 서브밴드에 대하여 추정된 Norm값을 양자화 및 부호화할 수 있다. 이때, Norm값은 벡터 양자화, 스칼라 양자화, TCQ, LVQ(Lattice vector quantization) 등 다양한 방식으로 양자화될 수 있다. 에너지 양자화 및 부호화부(720)는 추가적인 부호화 효율을 향상시키기 위해 무손실 부호화를 추가적으로 수행할 수 있다.
- [0080] 비트할당부(730)는 서브밴드별로 양자화된 Norm값을 이용하여 프레임당 허용비트를 고려하면서 부호화에 필요한 비트를 할당할 수 있다.

- [0081] 스펙트럼 정규화부(740)는 서브밴드별로 양자화된 Norm값을 이용하여 스펙트럼에 대한 정규화를 수행할 수 있다.
- [0082] 스펙트럼 양자화 및 부호화부(750)는 정규화된 스펙트럼에 대하여 서브밴드별로 할당된 비트에 근거하여 양자화 및 부호화를 수행할 수 있다.
- [0083] 노이즈 필링부(760)는 스펙트럼 양자화 및 부호화부(750)에서 허용비트의 제약으로 0으로 양자화된 부분에 적절한 노이즈를 추가할 수 있다.
- [0084] 도 8은 서브밴드 분할의 예를 나타내는 도면이다.
- [0085] 도 8을 참조하면, 입력신호가 48kHz의 샘플링 주파수를 사용하며, 20ms의 프레임 크기를 갖는 경우, 매 프레임 당 처리할 샘플의 개수는 960개가 된다. 즉, 입력 신호를 MDCT를 이용하여 50%의 오버래핑을 적용하여 변환하면 960개의 스펙트럼 계수가 얻어진다. 여기서, 오버래핑의 비율은 부호화 방식에 따라 다양하게 설정될 수 있다. 주파수 도메인에서는 이론적으로 24kHz까지 처리가능하지만 인간의 가청대역을 고려하여 20kHz까지의 대역을 표현하기로 한다. 저대역인 0~3.2kHz까지는 8개의 스펙트럼 계수를 하나의 서브밴드로 묶어서 사용하며, 3.2~6.4kHz의 대역에서는 16개의 스펙트럼 계수를 하나의 서브밴드로 묶어서 사용한다. 6.4~13.6kHz의 대역에서는 24개의 스펙트럼 계수를 하나의 서브밴드로 묶어서 사용하며, 13.6~20kHz의 대역에서는 32개의 스펙트럼 계수를 하나의 서브밴드로 묶어서 사용한다. 실제 Norm값을 구해서 부호화를 할 경우, 부호화기에서 정해진 대역까지 Norm을 구하여 부호화할 수 있다. 정해진 대역 이후의 특정 고대역에서는 대역확장과 같은 다양한 방식에 근거한 부호화가 가능하다.
- [0086] 도 9는 일실시예에 따른 스펙트럼 양자화장치의 구성을 나타내는 블록도이다.
- [0087] 도 9에 도시된 장치는 양자화기 선택부(910), USQ(930) 및 TCQ(950)을 포함할 수 있다.
- [0088] 도 9에 있어서, 양자화기 선택부(910)는 입력신호 즉, 양자화될 신호의 특성에 따라서 다양한 양자화기 중에서 가장 효율적인 양자화기를 선택할 수 있다. 입력신호의 특성으로서는 밴드별 비트 할당정보, 밴드의 크기 정보 등이 사용 가능하다. 선택결과에 따라서, 양자화될 신호를 USQ(830) 및 TCQ(850) 중 하나로 제공되어 대응되는 양자화를 수행할 수 있다. 입력신호는 정규화된 MDCT(Modified Discrete Cosine Transform) 스펙트럼일 수 있다. 입력신호의 대역폭은 협대역(narrow band) 혹은 광대역(wide band)일 수 있다. 입력신호의 부호화모드는 정상 모드(normal mode)일 수 있다.
- [0089] 도 10은 일실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블록도이다. 도 10에 도시된 장치는 도 7의 스펙트럼 양자화 및 부호화부(750)에 대응되거나, 다른 주파수도메인 부호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0090] 도 10에 도시된 장치는 부호화방식 선택부(1010), 제로 부호화부(1020), 스케일링부(1030), ISC 부호화부(1040), 양자화 성분 복원부(1050) 및 역스케일링부(1060)를 포함할 수 있다. 여기서, 여기서, 양자화성분 복원부(1050) 및 역스케일링부(1060)는 옵션으로 구비될 수 있다.
- [0091] 도 10에 있어서, 부호화방식 선택부(1010)는 입력신호 특성을 고려하여 부호화방식을 선택할 수 있다. 입력신호 특성은 대역폭과 밴드별 할당된 비트 중 적어도 하나를 포함할 수 있다. 정규화된 스펙트럼은 밴드별로 선택된 부호화방식에 근거하여 제로 부호화부(1020) 혹은 스케일링부(1030)로 제공될 수 있다. 실시예에 따르면, 대역폭이 협대역 혹은 광대역인 경우, 밴드의 각 샘플에 할당된 평균 비트수가 소정값, 예를 들면 0.75 이상인 경우 해당 밴드는 매우 중요하다고 판단되어 USQ가 사용되는 한편, 모든 다른 밴드는 TCQ가 사용될 수 있다. 여기서, 평균 비트수는 밴드 길이 혹은 밴드 크기를 고려하여 결정될 수 있다. 선택된 부호화방식은 1 비트의 플래그를 이용하여 설정될 수 있다. 다른 실시예에 따르면, 대역폭이 초광대역(super wide band) 혹은 전대역(full band)인 경우, USQ와 TCQ를 결합한 방식을 사용할 수 있다.
- [0092] 제로 부호화부(1020)는 할당된 비트가 0인 밴드에 대하여 모든 샘플들을 0으로 부호화할 수 있다.
- [0093] 스케일링부(1030)는 밴드에 할당된 비트에 근거하여 스펙트럼에 대한 스케일링을 수행함으로써 비트율을 조절할 수 있다. 이때, 정규화된 스펙트럼이 사용될 수 있다. 스케일링부(1030)는 밴드에 포함된 각 샘플, 즉 스펙트럼 계수에 할당된 평균 비트수를 고려하여 스케일링을 수행할 수 있다. 예를 들면, 평균 비트수가 많을수록 더 큰 스케일링이 행해질 수 있다.
- [0094] 실시예에 따르면, 스케일링부(1030)는 밴드별로 비트 할당에 따라 적절한 스케일링값을 결정할 수 있다.

[0095] 구체적으로, 먼저 밴드 길이(band length)와 비트 할당 정보를 이용하여 현재 밴드를 위한 펄스 개수를 추정할 수 있다. 여기서, 펄스는 단위 펄스를 의미할 수 있다. 먼저 하기의 수학적 식 1에 근거하여 현재 밴드에서 실제로 필요한 비트(b)를 산출할 수 있다.

**수학적 식 1**

$$b = \log_2 \left( \sum_{i=1}^{\min(m,n)} 2^i \frac{n!}{(n-i)!i!} \frac{(m-1)!}{(i-1)!(m-i)!} \right)$$

[0096]

[0097] 여기서 n은 밴드의 길이를 나타내며, m은 펄스 개수(number of pulses)를 의미하고, i는 ISC(the Important Spectral Component)를 갖는 번째로 위치의 갯수를 의미한다.

[0098] 한편, 번째로 위치의 개수는 예를 들면, 하기 수학적 식 2에서와 같이 확률에 근거하여 얻어질 수 있다.

**수학적 식 2**

$$pN_{ZP}(i) = 2^{i-b} C_n^i C_{m-1}^{i-1}, \quad i \in \{1, \dots, \min(m, n)\}$$

[0099]

[0100] 그리고, 번째로 위치들을 위하여 필요한 비트수는 하기 수학적 식 3에서와 같이 추정될 수 있다.

**수학적 식 3**

$$b_{nzp} = \log_2(pN_{ZP}(i))$$

[0101]

[0102] 최종적으로, 최종적으로 펄스의 개수는 각 밴드에 할당된 비트에 가장 가까운 값을 갖는 b 값에 의해 선택될 수 있다.

[0103] 다음, 밴드별로 구해진 펄스 개수 추정치와 입력신호의 절대값을 이용하여 초기 스케일링 팩터를 결정할 수 있다. 입력신호는 초기 스케일링 팩터에 의해 스케일링될 수 있다. 만약, 스케일링된 원신호 즉, 양자화된 신호에 대한 펄스 개수의 합이 펄스 개수 추정치와 같지 않은 경우에는, 업데이트된 스케일링 팩터를 이용하여 펄스 재분배(redistribution) 처리를 수행할 수 있다. 펄스 재분배 처리는 현재 밴드에 대하여 선택된 펄스 개수가 밴드별로 구해진 펄스 개수 추정치보다 적은 경우에는 스케일링 팩터를 감소시켜 펄스 개수를 증가시키며, 반대로 많은 경우에는 스케일링 팩터를 증가시켜 펄스 개수를 감소시킨다. 이때, 원신호와의 왜곡을 최소화 하는 위치를 선택하여 미리 정해진 값만큼 증가시키거나 감소시킬 수 있다.

[0104] TSQ를 위한 왜곡 함수는 정확한 거리보다는 상대적인 크기를 필요로 하기 때문에 하기의 수학적 식 4에서와 같이 각 밴드에서 각 양자화 및 역양자화된 값의 자승 거리의 합으로 얻어질 수 있다.

**수학적 식 4**

$$d^2 = \sum_{i=1}^n (p_i - q_i)^2$$

[0105]

[0106] 여기서, p<sub>i</sub>는 실제 값이고, q<sub>i</sub>는 양자화된 값을 나타낸다.

[0107] 한편, USQ를 위한 왜곡 함수는, 최선의 양자화된 값을 결정하기 위하여 유클리디안 거리를 사용할 수 있다. 이때, 복잡도를 최소화하기 위하여 스케일링 팩터를 포함하는 수정된 수학적 식을 사용하며, 왜곡 함수는 하기 수학

식 5에 의해 산출될 수 있다.

**수학식 5**

$$d_1 = \sqrt{\sum_{i=1}^n (p_i - g_1 q_i)^2}$$

[0108]

[0109]

만약, 밴드당 펄스개수가 요구되는 값과 매칭되지 않는 경우, 최소 메트릭을 유지하면서 소정갯수의 펄스를 가감할 필요가 있다. 이는 하나의 펄스를 가감하는 과정을 펄스개수가 요구되는 값에 이를때까지 반복하는 방법에 의해 수행될 수 있다.

[0110]

하나의 펄스를 가감하기 위하여, 가장 최적의 왜곡값을 구하기 위한 n개의 왜곡값을 구할 필요가 있다. 예를 들어, 왜곡값 j는 하기 수학식 6에서와 같이 밴드에서 j 번째 위치에 펄스를 추가하는 것에 해당될 수 있다.

**수학식 6**

$$d_2^j = \sqrt{\sum_{i=1}^n (p_i - g_2 \hat{q}_i)^2}, j=1...n$$

[0111]

[0112]

상기 수학식 6을 n 번 수행하는 것을 피하기 위하여, 하기 수학식 7에서와 같은 편차(deviation)를 사용할 수 있다.

**수학식 7**

$$d_2^j = \sqrt{\sum_{i=1}^n (p_i - g_2 \hat{q}_i)^2} = \sqrt{\sum_{i=1}^n p_i^2 - 2g_2 \sum_{i=1}^n p_i \sum_{i=1}^n \hat{q}_i + g_2^2 \sum_{i=1}^n \hat{q}_i^2} = \left\{ \sum_{i=1}^n \hat{q}_i = \sum_{i=1}^n q_i + 1 \right\}$$

$$\left\{ \sum_{i=1}^n \hat{q}_i^2 = \sum_{i \in \{1...n\}, i \neq j} q_i^2 + (q_j + 1)^2 = \sum_{i=1}^n q_i^2 + 2q_j + 1 \right\} =$$

$$= \sum_{i=1}^n p_i^2 - 2g_2 \left( \sum_{i=1}^n q_i p_i + p_j \right) + g_2^2 \left( \sum_{i=1}^n q_i^2 + 2q_j + 1 \right), j=1...n$$

[0113]

[0114]

상기 수학식 7에서  $\sum_{i=1}^n q_i^2, \sum_{i=1}^n q_i p_i, \sum_{i=1}^n p_i^2$  는 한번만 계산하면 된다. 한편, n은 밴드 길이 즉, 밴드에 있는 계수 갯수를 나타내며, p는 원신호 즉, 양자화기의 입력신호, q는 양자화된 신호, g는 스케일링 팩터를 나타낸다. 최종적으로, 왜곡 d를 최소화하는 위치 j가 선택되고, qj 가 업데이트될 수 있다.

[0115]

한편, 비트율을 제어하기 위하여, 스케일링된 스펙트럼 계수를 사용하여 적절한 ISC를 선택하여 부호화할 수 있다. 구체적으로, 양자화하기 위한 스펙트럼 성분은 각 밴드의 비트 할당을 사용하여 선택될 수 있다. 이때, 스펙트럼 성분들의 분포 및 분산에 따른 다양한 조합에 근거하여 스펙트럼 성분을 선택할 수 있다. 다음, 실제 난제로위치들을 산출할 수 있다. 난제로위치는 스케일링 양과 재분배 동작을 분석하여 얻을 수 있으며, 이와 같이 선택된 난제로위치가 다른 말로 ISC라할 수 있다. 요약하면, 스케일링과 재분배 과정을 거친 신호의 크기를 분석하여 최적 스케일링 팩터와 ISC에 해당하는 난제로위치 정보를 구할 수 있다. 여기서 난제로위치 정보는 난제로위치의 개수와 위치를 의미한다. 만일 스케일링과 재분배 과정을 통해서 펄스 개수가 조절되지 않을 경우, 선택된 펄스를 실제 TCQ과정을 통해 양자화하고 그 결과를 이용하여 잉여 비트를 조정할 수 있다. 이 과

정은 다음과 같은 예가 가능하다.

- [0116] 

년제로위치 갯수와 밴드별로 구해진 펄스개수 추정치가 같지 않고 년제로위치의 개수가 소정값, 예를 들면 1보다 크고 구해진 양자화기 선택정보가 TCQ를 나타내는 조건의 경우, 실제 TCQ 양자화를 통해서 잉여 비트를 조정할 수 있다. 구체적으로 상기 조건에 해당하는 경우, 잉여 비트를 조정하기 위하여 먼저 TCQ양자화 과정을 거친다. 사전에 밴드별로 구해진 펄스개수 추정치에 비해 실제 TCQ 양자화를 통해 구해진 현재 밴드의 펄스개수가 더 적은 경우에는 그전에 정해진 스케일링 팩터에 1보다 큰 값, 예를 들어 1.1을 곱해서 스케일링 팩터를 증가시키고, 반대의 경우에는 1보다 적은 값, 예를 들어 0.9를 곱해서 스케일링 팩터를 감소시킨다. 이와 같은 과정을 반복하여 밴드별로 구해진 펄스 개수 추정치와 TCQ 양자화를 통해 구해진 현재 밴드의 펄스개수가 같아지는 경우에 실제 TCQ양자화 과정에서 사용된 비트를 계산하여 잉여 비트를 업데이트한다. 이와 같이 구해진 년제로위치가 ISC에 해당할 수 있다.
- [0117] 

ISC 부호화부(1040)에서는 최종적으로 선택된 ISC의 개수 정보 및 년제로위치 정보를 부호화할 수 있다. 이 과정에서 부호화 효율을 높이기 위해 무손실 부호화를 적용할 수도 있다. ISC 부호화부(1040)는 할당된 비트가 0이 아닌 년제로밴드에 대하여 선택된 양자화기를 이용하여 부호화를 수행할 수 있다. 구체적으로, ISC 부호화부(1040)는 정규화된 스펙트럼에 대하여 각 밴드별로 ISC를 선택하고, 각 밴드별로 선택된 ISC의 정보를 갯수, 위치, 크기 및 부호에 근거하여 부호화할 수 있다. 이때, ISC의 크기는 갯수, 위치 및 부호와는 다른 방식으로 부호화할 수 있다. 일례를 들면 ISC의 크기는 USQ와 TCQ 중 하나를 이용하여 양자화하고 산술부호화하는 한편, ISC의 갯수, 위치 및 부호에 대해서는 산술부호화를 수행할 수 있다. 일실시예에 따르면, 신호특성에 근거하여 TCQ와 USQ 중 하나를 선택하되, 각 밴드에 대한 원래의 비트할당 정보 뿐 아니라, 이전 부호화된 밴드로부터의 잉여 비트에 대한 2차 비트할당 처리를 추가적으로 이용하여 양자화기 선택이 수행되는 제1 조인트 방식이 사용될 수 있다. 제1 조인트 방식에서 2차 비트할당 처리는 이전 부호화된 밴드로부터의 잉여 비트들을 분배하는 것으로서, 두개 밴드를 선택할 수 있다. 여기서, 신호특성은 각 밴드에 할당된 비트 혹은 밴드의 길이를 포함할 수 있다. 예를 들어 특정 밴드가 중요한 정보를 포함하고 있다고 판단되는 경우 USQ를 사용하고, 그렇지 않은 경우 TCQ를 사용할 수 있다. 만약, 밴드에 포함된 각 샘플에 할당된 평균 비트수가 임계치, 예를 들어 0.75 이상인 경우 해당 밴드는 매우 중요한 정보를 포함하고 있는 것으로 판단할 수 있으므로 USQ가 사용될 수 있다. 한편, 밴드의 길이가 짧은 저대역의 경우에도 필요에 따라서 USQ가 사용될 수 있다. 제1 조인트 방식이 사용되는 입력신호의 대역폭은 협대역 혹은 광대역일 수 있다. 다른 실시예에 따르면, 모든 대역을 USQ를 사용하여 부호화하되, LSB(Least Significant Bit)에 대해서는 TCQ를 사용하는 제2 조인트 방식이 사용될 수 있다. 제2 조인트 방식이 사용되는 입력신호의 대역폭은 초광대역 혹은 전대역일 수 있다.
- [0118] 

양자화 성분 복원부(1050)는 양자화된 성분에 ISC의 위치, 크기 및 부호 정보를 부가하여 실제의 양자화된 성분을 복원할 수 있다. 여기서, 제로 위치 즉, 제로로 부호화된 스펙트럼 계수에는 0이 할당될 수 있다.
- [0119] 

역스케일링부(1060)는 복원된 양자화 성분에 대하여 역스케일링을 수행하여, 정규화된 입력 스펙트럼과 동일한 레벨의 양자화된 스펙트럼 계수를 출력할 수 있다. 스케일링부(1030)와 역스케일링부(1060)에서는 동일한 스케일링 팩터를 사용할 수 있다.
- [0120] 

도 11은 일실시예에 따른 ISC 부호화장치의 구성을 나타내는 블록도이다.
- [0121] 

도 11에 도시된 장치는 ISC 선택부(1110) 및 ISC 정보 부호화부(1130)을 포함할 수 있다. 도 11의 장치는 도 10의 ISC 부호화부(1040)에 대응되거나 독립적인 장치로 구현될 수 있다.
- [0122] 

도 11에 있어서, ISC 선택부(1110)는 비트율을 조절하기 위하여 스케일링된 스펙트럼으로부터 소정 기준에 의거하여 ISC를 선택할 수 있다. ISC 선택부(1110)는 스케일링된 스펙트럼으로부터 스케일링된 정도를 분석하여, 실제의 년제로위치를 구할 수 있다. 여기서, ISC는 스케일링 이전의 실제의 년제로 스펙트럼 계수에 해당할 수 있다. ISC 선택부(1110)는 밴드별 할당된 비트에 근거하여 스펙트럼 계수의 분포 및 분산을 고려하여 부호화할 스펙트럼 계수 즉 년제로 위치를 선택할 수 있다. ISC 선택을 위하여 TCQ를 사용할 수 있다.
- [0123] 

ISC 정보 부호화부(1130)는 선택된 ISC에 근거하여, ISC 정보 즉 ISC 개수정보, 위치정보, 크기정보 및 부호를 부호화할 수 있다.
- [0124] 

도 12는 일실시예에 따른 ISC 정보 부호화장치의 구성을 나타내는 블록도이다.
- [0125] 

도 12에 도시된 장치는 위치 부호화부(1210), 크기 부호화부(1230) 및 부호 부호화부(1250)를 포함할 수 있다.
- [0126] 

도 12에 있어서, 위치 부호화부(1210)는 ISC 선택부(도 11의 1110)에서 선택된 ISC의 위치정보 즉, 년-제로 스



펙트럼 계수의 위치정보를 부호화할 수 있다. 위치정보는 선택된 ISC의 갯수와 위치를 포함할 수 있다. 위치정보의 부호화에는 산술 부호화(arithmetic coding)가 사용될 수 있다. 한편, 선택된 ISC를 모아서 새로운 버퍼를 구성할 수 있다. ISC 수집을 위하여 제로 밴드와 선택되지 못한 스펙트럼은 제외될 수 있다.

- [0127] 크기 부호화부(1230)는 새롭게 구성된 ISC의 크기정보에 대하여 부호화를 수행할 수 있다. 이때, TCQ와 USQ 중 하나를 선택하여 양자화를 수행하고, 이어서 산술 부호화를 추가적으로 수행할 수 있다. 산술 부호화의 효율을 높이기 위하여 년-제로 위치정보와 ISC의 갯수가 사용될 수 있다.
- [0128] 부호 부호화부(1250)는 선택된 ISC의 부호정보에 대하여 부호화를 수행할 수 있다. 부호정보의 부호화에는 산술 부호화가 사용될 수 있다.
- [0129] 도 13은 다른 실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블럭도이다. 도 13에 도시된 장치는 도 7의 스펙트럼 양자화 및 부호화부(750)에 대응되거나, 다른 주파수도메인 부호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0130] 도 13에 도시된 장치는, 스케일링부(1330), ISC 부호화부(1340), 양자화 성분 복원부(1350) 및 역스케일링부(1360)를 포함할 수 있다. 도 10과 비교할 때, 제로 부호화부(1020)와 부호화방식 선택부(1010)가 생략되며, ISC 부호화부(1340)는 TCQ를 사용할 수 있다는 것을 제외하고 각 구성요소의 동작을 동일하다.
- [0131] 도 14는 다른 실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블럭도이다. 도 14에 도시된 장치는 도 7의 스펙트럼 양자화 및 부호화부(750)에 대응되거나, 다른 주파수도메인 부호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0132] 도 14에 도시된 장치는, 부호화방식 선택부(1410), 스케일링부(1430), ISC 부호화부(1440), 양자화 성분 복원부(1450) 및 역스케일링부(1460)를 포함할 수 있다. 도 10과 비교할 때, 제로 부호화부(1020)가 생략되어 있다는 것을 제외하고 각 구성요소의 동작을 동일하다.
- [0133] 도 15는 일실시예에 따른 ISC 수집 및 부호화과정ISC 수집과정의 개념을 나타내는 도면으로서, 먼저 제로 밴드 즉, 0으로 양자화될 밴드는 제외한다. 다음, 년-제로 밴드에 존재하는 스펙트럼 성분들 중에서 선택된 ISC를 이용하여 새로운 버퍼를 구성할 수 있다. 새로 구성된 ISC에 대하여 밴드 단위로 USC와 TCQ를 결합한 제1 혹은 제2 조인트방식에 의해 양자화를 수행하고, 대응하는 무손실 부호화를 수행할 수 있다.
- [0134] 도 16은 USQ와 TCQ를 결합한 제2 조인트방식을 설명하는 도면이다.
- [0135] 도 16을 참조하면, 먼저 밴드 단위로 스펙트럼 데이터에 대하여 USQ를 이용하여 양자화할 수 있다. 양자화된 스펙트럼 데이터가 1보다 큰 값을 갖는 경우, 제로 혹은 1에 해당하는 LSB를 포함할 수 있다. 각 밴드에 대하여 LSB 시퀀스를 얻고 LSB 시퀀스를 TCQ를 이용하여 양자화함으로써, LSB 시퀀스와 가능한 트렐리스 경로간의 베스트 매치를 찾을 수 있다. SNR 기준으로 양자화된 시퀀스에 일부 에러가 발생할 수 있으나, 대신 시퀀스의 길이를 줄일 수 있다.
- [0136] 제2 조인트방식에 따르면, 하나의 스킵에서 두가지 양자화기 즉, USQ와 TCQ의 이점을 활용할 수 있고, TCQ로 버퍼의 경로 제한을 배제시킬 수 있다.
- [0137] 도 17은 다른 실시예에 따른 스펙트럼 부호화장치의 구성을 나타내는 블럭도이다. 도 17에 도시된 장치는 도 10의 ISC 부호화부(1040)에 대응되거나, 독립적으로 구현될 수 있다.
- [0138] 도 17에 도시된 장치는 제1 양자화부(1710), 제2 양자화부(1730), 제1 무손실 부호화부(1750), 제2 무손실 부호화부(1760), 제3 무손실 부호화부(1770) 및 비트스트림 생성부(1790)을 포함할 수 있다. 장치에 포함된 구성요소들은 적어도 하나의 프로세서로 일체화될 수 있다.
- [0139] 도 17을 참조하면, 제1 양자화부(1710)는 밴드 즉 년제로 밴드의 스펙트럼 데이터에 대하여 USQ를 이용하여 양자화할 수 있다. 각 밴드의 양자화를 위하여 할당되는 비트수는 미리 결정될 수 있다. 이때, 제2 양자화부(1730)에서의 TCQ에 사용될 비트수를 각 년제로 밴드로부터 균등하게 확보하고, 년제로 밴드의 나머지 비트수를 이용하여 USQ를 수행할 수 있다. 여기서, 스펙트럼 데이터는 norms에 해당할 수 있고, 정규화된 스펙트럼 데이터일 수 있다.
- [0140] 제2 양자화부(1730)는 제1 양자화부(1710)에서 양자화된 스펙트럼 데이터 중 하위비트에 대하여 TCQ를 이용하여 양자화할 수 있다. 하위비트는 LSB에 해당할 수 있다. 이때, 모든 밴드로부터 하위비트 즉, 레지듀얼 데이터를 수집하여 TCQ를 수행할 수 있다. 레지듀얼 데이터는 양자화이후 년제로 데이터를 갖는 모든 밴드에 대하여

수집되며, 양자화된 스펙트럼 데이터와 양자화되지 않은 스펙트럼 데이터 즉, 원래의 스펙트럼 데이터간의 차이를 나타낼 수 있다. 난제로 밴드에서 일부 주파수가 제로로 양자화된 경우, 이들은 레지듀얼 데이터에 포함되지 않을 수 있다. 레지듀얼 데이터는 하나의 어레이를 구성할 수 있다.

- [0141] 제1 무손실 부호화부(1750)는 밴드에 포함된 ISC 정보, 예를 들면 ISC의 갯수, 위치 및 부호에 대하여 무손실 부호화를 수행할 수 있다. 이때, 산술부호화가 사용될 수 있다.
- [0142] 제2 무손실 부호화부(1760)는 제1 양자화부(1710)에서 양자화된 스펙트럼 데이터 중 하위비트를 제외한 나머지 비트로 이루어지는 크기 정보에 대하여 무손실 부호화를 수행할 수 있다. 이때, 산술부호화가 사용될 수 있다.
- [0143] 제3 무손실 부호화부(1770)는 제2 양자화부(1730)에서 양자화 결과 얻어지는 TCQ 정보 즉, 트렐리스 경로 데이터에 대하여 무손실 부호화를 수행할 수 있다. 이때, 산술부호화가 사용될 수 있다. 트렐리스 경로 데이터는 equi-probable 심볼로 부호화될 수 있다. 즉, 트렐리스 경로 데이터는 이진 시퀀스로서, 균일한 확률 모델을 갖는 산술부호화기를 사용하여 부호화될 수 있다.
- [0144] 비트스트림 생성부(1790)는 제1 내지 제2 무손실 부호화부(1750, 1760, 1770)로부터 제공되는 데이터를 이용하여 비트스트림을 생성할 수 있다.
- [0145] 도 18은 도 17에 도시된 제2 양자화부의 구성을 나타내는 블록도이다.
- [0146] 도 18에 도시된 장치는 하위비트 추출부(1810), 레지듀얼 데이터 생성부(1830) 및 TCQ(1850)을 포함할 수 있다. 장치에 포함된 구성요소들은 적어도 하나의 프로세서로 일체화될 수 있다.
- [0147] 도 18을 참조하면, 하위비트 추출부(1810)는 제1 양자화부(1710)로부터 제공되는 양자화된 난제로 스펙트럼 데이터와 원래의 난제로 스펙트럼 데이터간의 차이에 근거한 레지듀얼 데이터를 추출할 수 있다. 이때, 레지듀얼 데이터는 양자화된 난제로 스펙트럼 데이터의 하위비트 예를 들면 LSB에 해당할 수 있다.
- [0148] 레지듀얼 데이터 생성부(1830)는 모든 난제로 밴드에 존재하는 양자화된 난제로 스펙트럼 데이터와 원래의 난제로 스펙트럼 데이터간 차이를 수집하여 레지듀얼 어레이를 구성할 수 있다. 도 19는 레지듀얼 데이터를 생성하는 방법을 설명하고 있다.
- [0149] TCQ(1850)는 레지듀얼 데이터 생성부(1830)로부터 제공되는 레지듀얼 어레이에 대하여 TCQ를 수행할 수 있다. 이때, (7,5)<sub>8</sub> 코드라 알려진 코드레이트 1/2을 갖는 TCQ를 사용할 수 있다. 도 20은 4개의 스테이트를 갖는 TCQ의 일예를 나타내고 있다. 실시예에 따르면, TCQ를 이용한 양자화는 처음 2XTCQ\_AMP 개의 크기에 대하여 수행될 수 있으며, 이때 상수 TCQ\_AMP 가 10으로 정의되는 경우 부호화되는 프레임당 20개의 크기까지 허용될 수 있음을 의미한다. 양자화이후 경로 메트릭이 체크되어 베스트 트렐리스 경로가 선택되고, 트랙백 절차가 수행되면서 무손실 부호화를 위하여 베스트 트렐리스 경로에 대한 데이터가 별도의 어레이로 저장될 수 있다.
- [0150] 도 21은 일실시예에 따른 주파수 도메인 오디오 복호화장치의 구성을 나타낸 블록도이다.
- [0151] 도 21에 도시된 주파수 도메인 오디오 복호화장치(2100)는 프레임 에러 검출부(2110), 주파수도메인 복호화부(2130), 시간도메인 복호화부(2150) 및 후처리부(2170)을 포함할 수 있다. 주파수도메인 복호화부(2130)는 스펙트럼 복호화부(2131), 메모리 갱신부(2133), 역변환부(2135) 및 OLA(overlap and add)부(2137)을 포함할 수 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나의 프로세서(미도시)로 구현될 수 있다.
- [0152] 도 21을 참조하면, 프레임 에러 검출부(2110)는 수신된 비트스트림으로부터 프레임에러가 발생하였는지를 검출할 수 있다.
- [0153] 주파수도메인 복호화부(2130)는 부호화 모드가 음악모드 혹은 주파수도메인 모드인 경우 동작하며, 프레임에러가 발생한 경우 FEC 혹은 PLC 알고리즘을 동작시키며, 프레임에러가 발생하지 않은 경우 일반적인 변환 복호화 과정을 통하여 시간 도메인 신호를 생성한다. 구체적으로, 스펙트럼 복호화부(2131)는 복호화된 파라미터를 이용하여 스펙트럼 복호화를 수행하여 스펙트럼 계수를 합성할 수 있다. 스펙트럼 복호화부(2131)에 대해서는 도 22 및 도 23을 참조하여 좀 더 구체적으로 설명하기로 한다.
- [0154] 메모리 갱신부(2133)는 정상 프레임인 현재 프레임에 대하여 합성된 스펙트럼 계수, 복호화된 파라미터를 이용하여 얻어진 정보, 현재까지 연속된 에러프레임의 개수, 각 프레임의 신호특성 혹은 프레임 타입 정보 등을 다음 프레임을 위하여 갱신할 수 있다. 여기서, 신호특성은 트랜지언트 특성, 스테이셔너리 특성을 포함할 수 있다.

고, 프레임 타입은 트랜지언트 프레임, 스테이셔널 프레임 혹은 하모닉 프레임을 포함할 수 있다.

- [0155] 역변환부(2135)는 합성된 스펙트럼 계수에 대하여 시간-주파수 역변환을 수행하여 시간 도메인 신호를 생성할 수 있다.
- [0156] OLA부(2137)는 이전 프레임의 시간 도메인 신호를 이용하여 OLA 처리를 수행하고, 그 결과 현재 프레임에 대한 최종 시간 도메인 신호를 생성하여 후처리부(2170)으로 제공할 수 있다.
- [0157] 시간도메인 복호화부(2150)는 부호화 모드가 음성모드 혹은 시간도메인 모드인 경우 동작하며, 프레임에러가 발생한 경우 FEC 혹은 PLC 알고리즘을 동작시키며, 프레임에러가 발생하지 않은 경우 일반적인 CELP 복호화과정을 통하여 시간도메인 신호를 생성한다.
- [0158] 후처리부(2170)는 주파수도메인 복호화부(2130) 혹은 시간도메인 복호화부(2150)로부터 제공되는 시간도메인 신호에 대하여 필터링 혹은 업샘플링 등을 수행할 수 있으나, 이에 한정되는 것은 아니다. 후처리부(2170)는 출력신호로서 복원된 오디오신호를 제공한다.
- [0159] 도 22는 일실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다. 도 22에 도시된 장치는 도 21의 스펙트럼 복호화부(2131)에 대응되거나, 다른 주파수도메인 복호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0160] 도 22에 도시된 스펙트럼 복호화장치(2200)는 에너지 복호화 및 역양자화부(2110), 비트할당부(2230), 스펙트럼 복호화 및 역양자화부(2250), 노이즈 필링부(2270) 및 스펙트럼 셰이핑부(2290)을 포함할 수 있다. 여기서, 노이즈 필링부(2270)는 스펙트럼 셰이핑부(2290)의 후단에 위치할 수도 있다. 각 구성요소는 적어도 하나 이상의 모듈로 일체화되어 적어도 하나의 프로세서(미도시)로 구현될 수 있다.
- [0161] 도 22를 참조하면, 에너지 복호화 및 역양자화부(2210)는 부호화과정에서 무손실 부호화가 수행된 파라미터, 예를 들어 Norm값과 같은 에너지에 대하여 무손실 복호화를 수행하고, 복호화된 Norm값에 대하여 역양자화를 수행할 수 있다. 부호화과정에서 Norm값의 양자화된 방식에 대응되는 방식을 사용하여 역양자화를 수행할 수 있다.
- [0162] 비트할당부(2230)는 양자화된 Norm값 혹은 역양자화된 Norm값에 근거하여 서브밴드별로 필요로 하는 비트수를 할당할 수 있다. 이 경우, 서브밴드 단위로 할당된 비트수는 부호화과정에서 할당된 비트수와 동일할 수 있다.
- [0163] 스펙트럼 복호화 및 역양자화부(2250)는 부호화된 스펙트럼 계수에 대하여 서브밴드별로 할당된 비트수를 사용하여 무손실 복호화를 수행하고, 복호화된 스펙트럼 계수에 대하여 역양자화 과정을 수행하여 정규화된 스펙트럼 계수를 생성할 수 있다.
- [0164] 노이즈 필링부(2270)는 정규화된 스펙트럼 계수 중, 서브밴드별로 노이즈 필링을 필요로 하는 부분에 대하여 노이즈를 채울 수 있다.
- [0165] 스펙트럼 셰이핑부(2290)는 역양자화된 Norm값을 이용하여 정규화된 스펙트럼 계수를 셰이핑할 수 있다. 스펙트럼 셰이핑 과정을 통하여 최종적으로 복호화된 스펙트럼 계수가 얻어질 수 있다.
- [0166] 도 23은 일실시예에 따른 스펙트럼 역양자화장치의 구성을 나타내는 블록도이다.
- [0167] 도 23에 도시된 장치는 역양자화기 선택부(2310), USQ(2330) 및 TCQ(2350)을 포함할 수 있다.
- [0168] 도 23에 있어서, 역양자화기 선택부(2310)는 입력신호 즉, 역양자화될 신호의 특성에 따라서 다양한 역양자화기 중에서 가장 효율적인 역양자화기를 선택할 수 있다. 입력신호의 특성으로서는 밴드별 비트 할당정보, 밴드의 크기 정보 등이 사용 가능하다. 선택결과에 따라서, 역양자화될 신호를 USQ(2330) 및 TCQ(2350) 중 하나로 제공되어 대응하는 역양자화를 수행할 수 있다. 도 23은 제1 조인트방식에 대응할 수 있다.
- [0169] 도 24는 일실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다. 도 24에 도시된 장치는 도 22의 스펙트럼 복호화 및 역양자화부(2250)에 대응되거나, 다른 주파수도메인 복호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0170] 도 24에 도시된 장치는 복호화방식 선택부(2410), 제로 복호화부(2430), ISC 복호화부(2450), 양자화성분 복원부(2470) 및 역스케일링부(2490)을 포함할 수 있다. 여기서, 양자화성분 복원부(2470) 및 역스케일링부(2490)는 옵션으로 구비될 수 있다.
- [0171] 도 24에 있어서, 복호화방식 선택부(2410)는 밴드별 할당된 비트에 근거하여 복호화방식을 선택할 수 있다. 정규화된 스펙트럼은 밴드별로 선택된 복호화방식에 근거하여 제로 복호화부(2430) 혹은 ISC 복호화부(2450)로

제공될 수 있다.

- [0172] 제로 복호화부(2430)는 할당된 비트가 0인 밴드에 대하여 모든 샘플들을 0으로 복호화할 수 있다.
- [0173] ISC 복호화부(2450)는 할당된 비트가 0이 아닌 밴드에 대하여 선택된 역양자화기를 이용하여 복호화를 수행할 수 있다. ISC 복호화부(2450)는 부호화된 스펙트럼의 각 밴드별로 중요주파수성분의 정보를 얻고, 각 밴드별로 얻어진 중요주파수성분의 정보를 갯수, 위치, 크기 및 부호에 근거하여 복호화할 수 있다. 중요주파수성분의 크기는 갯수, 위치 및 부호와 다른 방식으로 복호화할 수 있다. 일 예를 들면 중요주파수성분의 크기는 산술 복호화하고 USQ와 TCQ 중 하나를 이용하여 역양자화하는 한편, 중요주파수성분의 갯수, 위치 및 부호에 대하여 산술복호화를 수행할 수 있다. 역양자화기 선택은 도 10에 도시된 ISC 부호화부(1040)에서와 동일한 결과를 이용하여 수행할 수 있다. ISC 복호화부(2450)는 할당된 비트가 0이 아닌 밴드에 대하여 제1 혹은 제2 조인트 방식에 대응되하여 역양자화를 수행할 수 있다.
- [0174] 양자화 성분 복원부(2470)는 복원된 ISC의 위치, 크기 및 부호 정보에 근거하여 실제의 양자화 성분을 복원할 수 있다. 여기서, 제로 위치 즉, 제로로 복호화된 스펙트럼 계수인 양자화되지 않은 부분에는 0이 할당될 수 있다.
- [0175] 추가적으로 역스케일링부(미도시)를 포함하여 복원된 양자화 성분에 대하여 역스케일링을 수행하여, 정규화된 스펙트럼과 동일한 레벨의 양자화된 스펙트럼 계수를 출력할 수 있다.
- [0176] 도 25는 일 실시예에 따른 ISC 복호화장치의 구성을 나타내는 블록도이다.
- [0177] 도 25의 장치는 펄스수 추정부(2510) 및 ISC 정보 복호화부(2530)를 포함할 수 있다. 도 25의 장치는 도 24의 ISC 복호화부(2450)에 대응되거나 독립적인 장치로 구현될 수 있다.
- [0178] 도 25에 있어서, 펄스수 추정부(2510)는 밴드 크기와 비트 할당 정보를 이용하여 현재 밴드에서 필요한 펄스 개수 추정치를 결정할 수 있다. 즉, 현재 프레임의 비트할당 정보가 인코더와 동일하므로 동일한 비트 할당 정보를 이용하여 동일한 펄스 개수 추정치를 도출하여 복호화를 진행한다.
- [0179] ISC 정보 복호화부(2530)는 추정된 펄스수에 근거하여, ISC 정보 즉 ISC 개수정보, 위치정보, 크기정보 및 부호를 복호화할 수 있다.
- [0180] 도 26은 일 실시예에 따른 ISC 정보 복호화장치의 구성을 나타내는 블록도이다.
- [0181] 도 26에 도시된 장치는 위치 복호화부(2610), 크기 복호화부(2630) 및 부호 복호화부(2650)를 포함할 수 있다.
- [0182] 도 26에 있어서, 위치 복호화부(2610)는 비트스트림에 포함된 위치정보와 관련된 인덱스를 복호화하여 ISC의 갯수 및 위치를 복원할 수 있다. 위치정보의 복호화에는 산술 복호화가 사용될 수 있다. 크기 복호화부(2630)는 비트스트림에 포함된 크기정보와 관련된 인덱스에 대하여 산술복호화를 수행하고, 복호화된 인덱스에 대하여 제 1 혹은 제2 조인트방식 중 하나에 대응하여 역양자화를 수행할 수 있다. 산술 복호화의 효율을 높이기 위하여 난-제로 위치정보와 ISC의 갯수가 사용될 수 있다. 부호 복호화부(2550)는 비트스트림에 포함된 부호정보와 관련된 인덱스를 복호화하여 ISC의 부호를 복원할 수 있다. 부호정보의 복호화에는 산술 복호화가 사용될 수 있다. 일 실시예에 따르면, 난-제로 밴드가 필요로 하는 펄스 갯수를 추정하여 위치정보, 크기정보 혹은 부호정보 복호화에 사용할 수 있다.
- [0183] 도 27는 다른 실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다. 도 27에 도시된 장치는 도 22의 스펙트럼 복호화 및 역양자화부(2250)에 대응되거나, 다른 주파수도메인 복호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0184] 도 27에 도시된 장치는 ISC 복호화부(2750), 양자화성분 복원부(2770) 및 역스케일링부(2790)를 포함할 수 있다. 도 24와 비교할 때, 복호화방식 선택부(2410)와 제로 복호화부(2430)가 생략되어 있고, ISC 복호화부(2450)가 TCQ를 사용한다는 것을 제외하고 각 구성요소의 동작을 동일하다.
- [0185] 도 28은 다른 실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다. 도 28에 도시된 장치는 도 22의 스펙트럼 복호화 및 역양자화부(2250)에 대응되거나, 다른 주파수도메인 복호화장치에 포함되거나, 독립적으로 구현될 수 있다.
- [0186] 도 28에 도시된 장치는 복호화방식 선택부(2810), ISC 복호화부(2850), 양자화성분 복원부(2870) 및 역스케일링부(2890)를 포함할 수 있다. 도 24와 비교할 때, 제로 복호화부(2430)가 생략되어 있다는 것을 제외하고 각 구

성요소의 동작을 동일하다.

- [0187] 도 29는 다른 실시예에 따른 스펙트럼 복호화장치의 구성을 나타내는 블록도이다. 도 29에 도시된 장치는 도 24의 ISC 복호화부(2450)에 대응되거나, 독립적으로 구현될 수 있다.
- [0188] 도 29에 도시된 장치는 제1 내지 제3 디코딩부(2910, 2930, 2950)와 스펙트럼 성분 복원부(2970)를 포함할 수 있다.
- [0189] 도 29에 있어서, 제1 디코딩부(2910)는 비트스트림으로부터 밴드의 ISC 정보를 추출하여 ISC의 갯수, 위치 및 부호를 디코딩할 수 있다. 디코딩된 ISC 정보는 스펙트럼 성분 복원부(2970)로 제공되고, ISC의 위치정보는 제 2 디코딩부(2930)로 제공될 수 있다.
- [0190] 제2 디코딩부(2930)는 제1 디코딩부(2910)로부터 제공되는 디코딩된 ISC의 위치정보 및 각 밴드에 할당된 비트 수에 근거하여, 스펙트럼 데이터의 하위비트를 제외한 나머지 비트를 밴드별로 디코딩할 수 있다. 이때, 각 밴드에 할당된 비트수와 실제 사용된 비트수간의 차이에 해당하는 잉여 비트는 누적되어 다음 밴드에 사용될 수 있다.
- [0191] 제3 디코딩부(2950)는 비트스트림으로부터 TCQ 경로정보를 추출하여 디코딩함으로써 하위비트 시퀀스에 해당하는 TCQ 레지듀얼 어레이를 복원할 수 있다.
- [0192] 스펙트럼 성분 복원부(2970)는 제1 내지 제3 디코딩부(2910, 2930, 2950)로부터 제공되는 데이터에 근거하여 스펙트럼 성분을 복원할 수 있다.
- [0193] 제1 내지 제3 디코딩부(2910, 2930, 2950)에서는 무손실 복호화방식으로 산술 복호화를 사용할 수 있다.
- [0194] 도 30은 도 29에 도시된 제3 디코딩부의 구성을 나타내는 블록도이다.
- [0195] 도 30에 도시된 제3 디코딩부는 TCQ 경로 복호화부(3010)와 TCQ 레지듀얼 데이터 복원부(3030)를 포함할 수 있다.
- [0196] 도 30에 있어서, TCQ 경로 복호화부(3010)는 비트스트림으로부터 얻어지는 TCQ 경로정보를 복호화할 수 있다.
- [0197] TCQ 레지듀얼 데이터 복원부(3030)는 복호화된 TCQ 경로정보에 근거하여 TCQ 레지듀얼 데이터를 복원할 수 있다. 구체적으로, 복호화된 트렐리스 스테이트에 따라서 레지듀얼 데이터 즉, 레지듀얼 어레이를 복원할 수 있다. 각 경로 비트로부터, 레지듀얼 어레이에 있는 2개의 LSB 비트들이 생성될 수 있다. 이를 pseudo code로 나타내면 다음과 같다.
- [0198] `for( state = 0, i = 0; i < bcount; i++)`
- [0199] `{`
- [0200] `residualbuffer[2*i] = dec_LSB[state][dpath[i]] & 0x1;`
- [0201] `residualbuffer [2*i + 1] = dec_LSB[state][dpath[i]] & 0x2;`
- [0202] `state = trellis_nextstate[state][dpath[i]];`
- [0203] `}`
- [0204] 즉, 스테이트 0에서부터 시작하여, 복호화된 dpath 비트를 사용하여 트렐리스를 따라 움직이고 현재 트렐리스 어레이에 해당하는 2개의 비트를 추출할 수 있다.
- [0205] 도 29 및 도 30에 도시된 구성은 도 17 및 도 18의 구성과 서로 가역적인 관계에 있다.
- [0206] 도 31은 본 발명의 일실시예에 따른 부호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.
- [0207] 도 31에 도시된 멀티미디어 기기(3100)는 통신부(3110)와 부호화모듈(3130)을 포함할 수 있다. 또한, 부호화 결과 얻어지는 오디오 비트스트림의 용도에 따라서, 오디오 비트스트림을 저장하는 저장부(3150)을 더 포함할 수 있다. 또한, 멀티미디어 기기(3100)는 마이크로폰(3170)을 더 포함할 수 있다. 즉, 저장부(3150)와 마이크로폰(3170)은 옵션으로 구비될 수 있다. 한편, 도 28에 도시된 멀티미디어 기기(3100)는 임의의 복호화모듈(미도시), 예를 들면 일반적인 복호화 기능을 수행하는 복호화모듈 혹은 본 발명의 일실시예에 따른 복호화모듈을 더 포함할 수 있다. 여기서, 부호화모듈(3130)은 멀티미디어 기기(3100)에 구비되는 다른 구성요소(미도시)와 함께 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.

- [0208] 도 31을 참조하면, 통신부(3110)는 외부로부터 제공되는 오디오와 부호화된 비트스트림 중 적어도 하나를 수신하거나, 복원된 오디오와 부호화모듈(3130)의 부호화결과 얻어지는 오디오 비트스트림 중 적어도 하나를 송신할 수 있다.
- [0209] 통신부(3110)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 기기 혹은 서버와 데이터를 송수신할 수 있도록 구성된다.
- [0210] 부호화모듈(3130)은 일실시예에 따르면, 현재 밴드의 스펙트럼 데이터에 대하여 제1 양자화방식에 따라서 양자화하고, 스펙트럼 데이터 및 양자화된 스펙트럼 데이터를 이용하여 현재 밴드의 하위비트를 생성하고, 현재 밴드의 하위비트가 포함된 하위비트 시퀀스에 대하여 제2 양자화방식에 따라서 양자화하고, 양자화된 스펙트럼 데이터 중 N 비트(여기서 N는 1 이상)를 제외한 상위비트 및 양자화된 하위비트 시퀀스에 근거하여 비트스트림을 생성할 수 있다.
- [0211] 저장부(3150)는 멀티미디어 기기(3100)의 운용에 필요한 다양한 프로그램을 저장할 수 있다.
- [0212] 마이크론(3170)은 사용자 혹은 외부의 오디오신호를 부호화모듈(2830)로 제공할 수 있다.
- [0213] 도 32는 본 발명의 일실시예에 따른 복호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.
- [0214] 도 32에 도시된 멀티미디어 기기(3200)는 통신부(3210)와 복호화모듈(3230)을 포함할 수 있다. 또한, 복호화결과 얻어지는 복원된 오디오신호의 용도에 따라서, 복원된 오디오신호를 저장하는 저장부(3250)을 더 포함할 수 있다. 또한, 멀티미디어 기기(3200)는 스피커(3270)를 더 포함할 수 있다. 즉, 저장부(2950)와 스피커(3270)는 옵션으로 구비될 수 있다. 한편, 도 32에 도시된 멀티미디어 기기(3200)는 임의의 부호화모듈(미도시), 예를 들면 일반적인 부호화 기능을 수행하는 부호화모듈 혹은 본 발명의 일실시예에 따른 부호화모듈을 더 포함할 수 있다. 여기서, 복호화모듈(3230)은 멀티미디어 기기(3200)에 구비되는 다른 구성요소(미도시)와 함께 일체화되어 적어도 하나의 이상의 프로세서(미도시)로 구현될 수 있다.
- [0215] 도 32를 참조하면, 통신부(3210)는 외부로부터 제공되는 부호화된 비트스트림과 오디오 신호 중 적어도 하나를 수신하거나 복호화 모듈(3230)의 복호화결과 얻어지는 복원된 오디오 신호와 부호화결과 얻어지는 오디오 비트스트림 중 적어도 하나를 송신할 수 있다. 한편, 통신부(3210)는 도 31의 통신부(3110)와 실질적으로 유사하게 구현될 수 있다.
- [0216] 복호화 모듈(3230)은 일실시예에 따르면, 통신부(3210)를 통하여 제공되는 비트스트림을 수신하고, TCQ 경로 정보를 추출하여 하위비트 시퀀스를 디코딩하고, ISC 정보를 추출하여 ISC의 갯수, 위치 및 부호를 디코딩하는 한편, 하위비트를 제외한 나머지 비트를 추출하여 디코딩하고, 디코딩된 하위비트 시퀀스와 디코딩된 하위비트를 제외한 나머지 비트에 근거하여 스펙트럼 성분을 복원할 수 있다.
- [0217] 저장부(3250)는 복호화 모듈(3230)에서 생성되는 복원된 오디오신호를 저장할 수 있다. 한편, 저장부(3250)는 멀티미디어 기기(3200)의 운용에 필요한 다양한 프로그램을 저장할 수 있다.
- [0218] 스피커(3270)는 복호화 모듈(3230)에서 생성되는 복원된 오디오신호를 외부로 출력할 수 있다.
- [0219] 도 33은 본 발명의 일실시예에 따른 부호화모듈과 복호화모듈을 포함하는 멀티미디어 기기의 구성을 나타낸 블록도이다.
- [0220] 도 33에 도시된 멀티미디어 기기(3300)는 통신부(3310), 부호화모듈(3320)과 복호화모듈(3330)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(3340)을 더 포함할 수 있다. 또한, 멀티미디어 기기(3300)는 마이크론(3350) 혹은 스피커(3360)를 더 포함할 수 있다. 여기서, 부호화모듈(3320)과 복호화모듈(3330)은 멀티미디어 기기(3300)에 구비되는 다른 구성요소(미도시)와 함께 일체화되어 적어도 하나 이상의 프로세서(미도시)로 구현될 수 있다.
- [0221] 도 33에 도시된 각 구성요소는 도 31에 도시된 멀티미디어 기기(3100)의 구성요소 혹은 도 32에 도시된 멀티미디어 기기(3200)의 구성요소와 중복되므로, 그 상세한 설명은 생략하기로 한다.
- [0222] 도 31 내지 도 33에 도시된 멀티미디어 기기(3100, 3200, 3300)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 단말장치, 텔레컨퍼런싱 혹은 인터랙션 시스템의 사용자 단말이 포함될 수 있으나, 이에 한정되는 것은 아니다. 또한, 멀티미디어 기기(3100, 3200, 3300)는 클라이언트, 서버 혹은 클라이언트와 서버 사이에 배치되는 변환기로서 사용될 수 있다.

- [0223] 한편, 멀티미디어 기기(3100, 3200, 3300)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.
- [0224] 한편, 멀티미디어 기기(3100, 3200, 3300)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.
- [0225] 도 34는 일실시예에 따른 스펙트럼 부호화방법의 동작을 나타낸 플로우차트이다.
- [0226] 도 34를 참조하면, 3410 단계에서는 현재 밴드의 스펙트럼 데이터에 대하여 제1 양자화방식에 따라서 양자화할 수 있다. 여기서, 제1 양자화방식은 스칼라 양자화일 수 있으며, 균일한 양자화 간격을 사용하는 USQ를 사용할 수 있다.
- [0227] 3430 단계에서는 스펙트럼 데이터 및 양자화된 스펙트럼 데이터를 이용하여 현재 밴드의 하위비트를 생성할 수 있다. 여기서, 하위비트는 스펙트럼 데이터와 양자화된 스펙트럼 데이터간의 차이에 근거하여 얻어질 수 있다. 제2 양자화방식은 트렐리스 부호 양자화(TCQ)에 해당할 수 있다.
- [0228] 3450 단계에서는 현재 밴드의 하위비트가 포함된 하위비트 시퀀스에 대하여 제2 양자화방식에 따라서 양자화할 수 있다.
- [0229] 3470 단계에서는 양자화된 스펙트럼 데이터 중 N 비트(여기서 N는 1 이상)를 제외한 상위비트 및 양자화된 하위비트 시퀀스에 근거하여 비트스트림을 생성할 수 있다.
- [0230] 한편, 도 34에 도시된 스펙트럼 부호화방법과 관련된 스펙트럼 데이터의 대역폭은 초광대역 혹은 전대역일 수 있다. 또한, 스펙트럼 데이터는 입력 오디오 신호에 대하여 MDCT(Modified Discrete Cosine Transform) 및 정규화를 수행하여 얻어질 수 있으며, 정상 모드로 부호화될 수 있다.
- [0231] 도 34의 각 단계에 대해서는 필요에 따라서 전술한 부호화장치의 각 구성요소의 동작이 더 부가될 수 있다.
- [0232] 도 35는 일실시예에 따른 스펙트럼 복호화방법의 동작을 나타낸 플로우차트이다.
- [0233] 도 35를 참조하면, 3510 단계에서는 비트스트림으로부터 ISC 정보를 추출하여 ISC의 갯수, 위치 및 부호를 디코딩하는 한편, 하위비트를 제외한 나머지 비트를 추출하여 디코딩할 수 있다.
- [0234] 3530 단계에서는 비트스트림으로부터 TCQ 경로 정보를 추출하여 하위비트 시퀀스를 디코딩할 수 있다.
- [0235] 3550 단계에서는 3510 단계에서 디코딩된 하위비트를 제외한 나머지 비트와 3530 단계에서 디코딩된 하위비트 시퀀스에 근거하여 스펙트럼 성분을 복원할 수 있다.
- [0236] 도 35의 각 단계에 대해서는 필요에 따라서 전술한 복호화장치의 각 구성요소의 동작이 더 부가될 수 있다.
- [0237] 도 36은 일실시예에 따른 비트할당장치의 구성을 나타내는 블럭도이다. 도 36에 도시된 장치는 도 5의 비트할당부(516), 도 7의 비트할당부(730) 혹은 도 22의 비트할당부(2230)에 대응되거나, 독립적으로 구현될 수 있다.
- [0238] 도 36에 도시된 장치는 비트 추정부(3610), 재분배부(3630) 및 조정부(3650)을 포함할 수 있으며, 적어도 하나의 프로세서로 일체화될 수 있다. 스펙트럼 양자화를 위한 비트할당을 위하여 프랙셔널 비트할당(fractional bit allocation)이 사용될 수 있다. 이에 따르면, 3비트의 프랙셔널 부분을 갖는 비트할당이 가능하게 되므로 좀 더 정밀한 비트할당이 가능해진다. 여기서, 부호화 모드는 일반 모드(generic mode)에 해당할 수 있다.
- [0239] 도 36에 있어서, 할당비트 추정부(3610)는 밴드의 평균 에너지, 예를 들면 norms에 근거하여 할당비트를 추정할 수 있다.
- [0240] 밴드의 초기비트  $R_0(p,0)$ 는 하기 수학적 식 8에 의해 추정될 수 있다.

수학식 8

$$R_0(p,0) = \max \left\{ 0, \frac{L_M(p)}{3} * \left( \hat{I}_M(i) - \frac{\sum_{i=0}^{N_{bands}-1} L_M(i) * \hat{I}_M(i) - 3 * TB}{\sum_{i=0}^{N_{bands}-1} L_M(i)} \right) \right\} \quad \text{for } p = 0, 1, \dots, N_{bands} - 1$$

[0241]

[0242]

여기서,  $L_M(p)$ 는 밴드  $p$ 에 있는 샘플당 1 비트로부터 얻어지는 비트수로서, 10개의 샘플이 존재할 경우 10 비트를 나타낸다.  $TB$ 는 허용가능한 전체 비트수(total bit budget)를 나타낸다.  $\hat{I}_M(i)$ 은 밴드  $i$ 의 양자화된 norms을 나타낸다.

[0243]

재분배부(3630)는 밴드의 초기비트를 소정 기준에 따라서 재분배할 수 있다.

[0244]

1 단계 반복처리를 통하여, 프레임에 포함된 모든 밴드의 초기비트의 합이 전체 비트수  $TB$ 와 같아질 때까지, 넉넉히 비트를 갖는 밴드에 할당비트를 재분배할 수 있다. 이는 하기 수학식 9에서와 같이 수행될 수 있다.

수학식 9

$$R_0(p,k) = \max \left\{ 0, R_0(p,k-1) - L_M(p) * \frac{\sum_{i=0}^{N_{bands}-1} R_0(p,k-1) - TB}{NSL_0(k-1)} \right\}$$

[0245]

[0246]

여기서,  $NSL_0(k-1)$ 은  $k$ 번째 반복처리후 할당 비트를 갖는 밴드들에 존재하는 스펙트럼 라인들의 수를 나타낸다.

[0247]

한편, 너무 적은 비트가 할당된 경우, SNR(signal-to-noise ratio)의 감소로 인하여 품질 저하를 초래할 수 있다. 이를 해결하기 위하여, 밴드의 할당 비트에 최소비트 제한을 적용할 수 있다. 제1 최소 비트수는 밴드 인덱스와 비트레이트에 따라서 상수값으로 이루어질 수 있다. 일례를 들면, 밴드  $p$ 의 제1 최소 비트수  $LNB(p)$ 는 밴드  $p$ 가 0 내지 15인 경우 3, 밴드  $p$ 가 16 내지 23인 경우 4, 밴드  $p$ 가 24 내지  $N_{bands}-1$ 인 경우 5로 결정할 수 있다.

[0248]

다음, 2단계 반복처리를 통하여, 할당비트의 재분배가 수행되어  $L_M(p)$  보다 많은 할당비트를 갖는 밴드에 비트를 할당할 수 있다. 여기서,  $L_M(p)$ 는 각 밴드의 제2 최소 비트수에 해당할 수 있다.

[0249]

먼저, 할당비트가 1단계 반복처리 결과와 제1 및 제2 최소 비트수에 근거하여 산출될 수 있다. 일례를 들면, 할당비트  $R_1(p,0)$ 는 하기 수학식 10에서와 같이 구해질 수 있다.

수학식 10

$$R_1(p,0) = \begin{cases} 0 & \text{if } R(p) < bs + LNB(p) \\ L_M(p) & \text{if } bs + LNB(p) \leq R(p) \leq L_M(p) \end{cases} \quad \text{for } p = 0, \dots, N_{bands} - 1$$

[0250]

[0251]

여기서,  $R(p)$ 는 1단계 반복처리후 밴드  $p$ 의 할당비트를 나타내고,  $bs$ 는 24 kbps의 경우 2, 32 kbps의 경우 3을



나타내나, 이에 한정되는 것은 아니다.

[0252] 한편, 허용가능한 전체 비트수 TB는  $L_M(p)$  비트를 갖는 밴드들의 비트수를 감산함으로써 갱신되고, 밴드 인덱스  $p$ 는  $L_M(p)$  비트보다 많은 비트를 갖는 밴드 인덱스를 나타내는  $p'$ 으로 갱신될 수 있다.  $N_{bands}$ 는  $p'$ 에 대한 밴드 수인  $N'_{bands}$ 로 갱신될 수 있다.

[0253] \*이어, 갱신된 TB 즉, TB'이  $L_M(p')$  비트보다 많은 비트수를 갖는 밴드들의 비트수와 같아질 때까지 2단계 반복 처리가 수행될 수 있다. 일례를 들면, 하기 수학식 11에서와 같이 수행될 수 있다.

**수학식 11**

$$R_1(p', k) = \max \left\{ L_M(p'), R_1(p', k-1) - L_M(p') * \frac{\sum_{i=0}^{N_{bands}-1} R_1(p', k-1) - TB'}{NSL_1(k-1)} \right\}$$

[0254]

[0255] 여기서,  $NSL_1(k-1)$ 은  $k$ 번째 반복처리후  $L_M(p')$  비트보다 많은 비트수를 갖는 밴드들에 존재하는 스펙트럼 라인들의 수를 나타낸다.

[0256] 제2 반복처리 동안,  $L_M(p')$  비트보다 많은 비트수를 갖는 밴드가 없는 경우, TB'이 0이 될때까지 넉제로 할당비트를 갖는 밴드의 할당비트가 최상위 밴드에서부터 0으로 설정될 수 있다.

[0257] 다음, 초과하여 할당된 비트들 및 적게 할당된 비트들에 대한 최종 재분배 처리가 수행될 수 있다. 이때, 소정 기준값에 근거하여 최종 재분배 처리가 수행수행될 수 있다.

[0258] 조정부(3650)는 할당비트의 소수점 부분(fractional part)이 소정 비트가 되도록 조정할 수 있다. 일례를 들면, 소수점 이하 3 비트로 조정할 수 있으며, 이는 하기 수학식 12를 통하여 수행될 수 있다.

**수학식 12**

$$R(p) = \lfloor R(p) * 8 \rfloor / 8 \quad \text{for } p = 0, \dots, N_{bands} - 1$$

[0259]

[0260] 도 37은 일실시예에 따른 모드 결정장치의 구성을 나타낸 블록도이다.

[0261] 도 37에 도시된 장치는 음성/음악 분류부(3710)와 정정부(3730)을 포함할 수 있다. 도 37에 도시된 장치는 도 2a의 모드결정부(213), 도 3a의 모드결정부(314) 혹은 도 4a의 모드결정부(413)에 포함되거나, 도 2a의 시간도메인 부호화부(215), 도 3a의 시간도메인 여기부호화부(316) 혹은 도 4a의 시간도메인 여기부호화부(417)에 추가적으로 포함되거나, 독립적으로 구현될 수 있다. 여기서, 각 구성요소는 별도의 하드웨어로 구현되어야 할 필요가 있는 경우를 제외하고는, 하나의 프로세서로 일체화될 수 있다. 한편, 오디오신호는 음악신호 혹은 음성신호, 혹은 음악과 음성의 혼합신호를 의미할 수 있다.

[0262] 도 37을 참조하면, 음성/음악 분류부(110)은 다양한 초기 분류 파라미터에 근거하여 오디오신호가 음악신호에 해당하는지 음성신호에 해당하는지 분류할 수 있다. 오디오 신호 분류과정은 적어도 하나 이상의 단계를 포함할 수 있다.

[0263] 일실시예에 따르면, 현재 프레임과 복수개의 이전 프레임의 신호 특성에 근거하여 오디오 신호를 음성 신호 혹은 음악 신호로 분류할 수 있다. 신호 특성은 단구간 특성과 장구간 특성 중 적어도 하나를 포함할 수 있다. 또한, 신호 특성은 시간 도메인 특성과 주파수 도메인 특성 중 적어도 하나를 포함할 수 있다. 여기서, 음성 신호로 분류되면, CELP(Code Excited Linear Prediction) 타입 코더를 이용하여 부호화될 수 있다. 한편, 음악 신호로 분류되면, 트랜스폼 코더를 이용하여 부호화될 수 있다. 여기서, 트랜스폼 코더의 일례로는

MDCT(Modified Discrete Cosine Transform) 코더를 들 수 있으나 이에 한정되는 것은 아니다.

[0264] 다른 실시예에 따르면, 오디오 신호 분류과정은 오디오 신호가 음성 특성을 갖는지 여부에 따라서 오디오 신호를 음성 신호와 일반적인 오디오 신호(generic audio signal) 즉, 음악 신호로 분류하는 제1 단계와 일반 오디오 신호가 GSC(Generic Signal audio Coder)에 적합한지를 판단하기 위한 제2 단계를 포함할 수 있다. 제1 단계의 분류결과와 제2 단계의 분류결과를 조합하여 오디오 신호가 음성신호로 분류될 수 있는지 음악신호로 분류될 수 있는지를 결정할 수 있다. 음성 신호로 분류되면 CELP 타입 코더로 부호화될 수 있다. CELP 타입 코더는 비트율 혹은 신호특성에 따라서 무성음 부호화(Unvoiced Coding; UC) 모드, 유성음 부호화(Voiced Coding; VC) 모드, 트랜지언트 부호화(Transition Coding; TC) 모드, 일반 부호화(Generic Coding; GC) 모드 중 복수개를 포함할 수 있다. 한편, GSC(Generic Signal audio Coding) 모드는 별도의 코더로 구현되거나 CELP 타입 코더의 하나의 모드로 포함될 수 있다. 음악 신호로 분류되면 트랜스폼 코더 혹은 CELP/트랜스폼 하이브리드 코더 중 하나를 이용하여 부호화될 수 있다. 세부적으로 트랜스폼 코더는 음악신호에, CELP/트랜스폼 하이브리드 코더는 음성신호가 아닌 비음악(non-music) 신호 혹은 음악과 음성이 혼합된 신호(mixed signal)에 적용될 수 있다. 일실시예에 따르면, 대역폭에 따라서 CELP 타입 코더, CELP/트랜스폼 하이브리드 코더와 트랜스폼 코더 모두가 사용되거나, CELP 타입 코더와 트랜스폼 코더가 사용될 수 있다. 예를 들어, 협대역(NB)인 경우 CELP 타입 코더와 트랜스폼 코더가 사용되고, 광대역(WB), 초광대역(SWB), 전대역(FB)의 경우 CELP 타입 코더, CELP/트랜스폼 하이브리드 코더와 트랜스폼 코더가 사용될 수 있다. CELP/트랜스폼 하이브리드 코더는 시간 도메인에서 동작하는 LP 기반 코더와 트랜스폼 도메인 코더를 결합한 것으로서, GSC(Generic Signal audio Coder)라고도 한다.

[0265] 제1 단계의 신호 분류는 GMM(Gaussian Mixture Model)에 근거할 수 있다. GMM을 위하여 다양한 신호 특성이 사용될 수 있다. 신호 특성의 예로는, 오픈 루프 피치, 정규화된 상관도, 스펙트럼 엔벨로프, 토널 안정도, 신호의 년-스테이셔너리티, LP 레지들 에러, 스펙트럼 차이값, 스펙트럼 스테이셔너리티 등과 같은 특성을 들 수 있으나, 이에 한정되는 것은 아니다. 제2 단계의 신호 분류를 위하여 사용되는 신호 특성의 예로는 스펙트럼 에너지 변동 특성, LP 분석 레지들 에너지의 틸트 특성, 고대역 스펙트럼 피키니스 특성, 상관도 특성, 보이싱 특성, 토널 특성 등을 들 수 있으나, 이에 한정되는 것은 아니다. 제1 단계에서 사용되는 특성은 CELP 타입 코더로 부호화하는 것이 적합한지를 판단하기 위하여 음성 특성인지 비음성 특성인지를 판단하기 위한 것이고, 제2 단계에서 사용되는 특성은 GSC로 부호화하는 것이 적합한지를 판단하기 위하여 음악 특성인지 비음악 특성인지를 판단하기 위한 것일 수 있다. 예를 들어, 제1 단계에서 음악 신호로 분류된 한 세트의 프레임들은 제2 단계에서 음성 신호로 전환되어 CELP 모드들 중 하나로 부호화될 수 있다. 즉, 큰 피치 주기 및 높은 안정도를 가지면서 상관도가 큰 신호 혹은 어택신호인 경우 제2 단계에서 음악 신호로부터 음성 신호로 전환될 수 있다. 이와 같은 신호 분류 결과에 따라서 부호화모드가 변경될 수 있다.

[0266] 정정부(3730)는 음성/음악 분류부(3710)의 분류 결과를 적어도 하나의 수정 파라미터에 근거하여 수정할 수 있다. 정정부(3730)는 컨텍스트에 기반하여 음성/음악 분류부(3710)의 분류 결과를 수정할 수 있다. 예를 들어, 현재 프레임이 음성 신호로 분류된 경우 음악 신호로 수정되거나 음성 신호로 유지될 수 있고, 현재 프레임이 음악 신호로 분류된 경우 음성 신호로 수정되거나 음악 신호로 유지될 수 있다. 현재 프레임의 분류 결과에 예러가 존재하는지를 판단하기 위하여 현재 프레임을 포함하는 복수개의 프레임들의 특성이 사용될 수 있다. 예를 들면, 8개의 프레임이 사용될 수 있으나, 이에 한정되는 것은 아니다.

[0267] 수정 파라미터의 예로는 토널리티, 선형예측에러, 보이싱, 상관도 등과 같은 특성 중 적어도 하나를 조합하여 사용될 수 있다. 여기서, 토널리티는 1~2 kHz 영역의 토널리티(ton<sub>2</sub>)과 2~4 kHz 영역의 토널리티(ton<sub>3</sub>)을 포함할 수 있으며, 각각 하기 수학적 식 13 및 14로 정의될 수 있다.

**수학적 식 13**

$$ton_2 = 0.2 * \log_{10} \left[ \sqrt{\frac{1}{8} \sum_{i=0}^7 \{tonality 2^{-i}\}^2} \right]$$

[0268]

**수학식 14**

$$ton_3 = 0.2 * \log_{10} \left[ \sqrt{\frac{1}{8} \sum_{i=0}^7 \{tonality3^{[-i]}\}^2} \right]$$

[0269]

[0270] 여기서, 위첨자(superscript) [-i]는 이전 프레임을 나타낸다. 예를 들어, tonality2<sup>[-1]</sup>는 1 프레임 이전 프레임의 1~2 kHz 영역의 토널리티를 나타낸다.

[0271] 한편, 저대역의 장구간 토널리티 ton<sub>LT</sub>는 ton<sub>LT</sub> = 0.2 \* log<sub>10</sub>[lt\_tonality]와 같이 정의될 수 있다. 여기서, lt\_tonality은 전대역의 장구간 토널리티를 나타낼 수 있다.

[0272] 한편, n 프레임에서 1~2 kHz 영역의 토널리티(ton<sub>2</sub>)과 2~4 kHz 영역의 토널리티(ton<sub>3</sub>)간의 차이 d<sub>t</sub>는 d<sub>t</sub> = 0.2 \* {log<sub>10</sub>(tonality2(n))-log<sub>10</sub>(tonality3(n))}와 같이 정의될 수 있다.

[0273] 다음, 선형예측에러 LP<sub>err</sub>은 하기 수학식 15로 정의될 수 있다.

**수학식 15**

$$LP_{err} = \sqrt{\frac{1}{8} \sum_{i=0}^7 [FV_s^{[-i]}(9)]^2}$$

[0274]

[0275] 여기서, FV<sub>s</sub>(9)는 FV<sub>s</sub>(i) = sfa<sub>i</sub>FV<sub>i</sub> + sfb<sub>i</sub> (여기서, i = 0, ..., 11)에 의해 정의되며, 음성/음악 분류부(3710)에서 사용되는 특징 파라미터 중 하기 수학식 16으로 정의되는 LP 레지들 로그-에너지 비율 특징 파라미터를 스케일링한 값에 해당하는 것이다. 여기서, sfa<sub>i</sub>, sfb<sub>i</sub>는 특징 파라미터의 종류 및 대역폭에 따라서 달라질 수 있으며, 각 특징 파라미터를 [0;1] 범위로 근사화하기 위하여 사용된다.

**수학식 16**

$$FV_9 = \log\left(\frac{E(13)}{E(1)}\right) + \log\left(\frac{E^{[-1]}(13)}{E^{[-1]}(1)}\right)$$

[0276]

[0277] 여기서, E(1)는 첫번째 LP 계수의 에너지, E(13)은 13번째 LP 계수의 에너지를 나타낸다.

[0278] 다음, 음성/음악 분류부(3710)에서 사용되는 특징 파라미터 중 하기 수학식 17에 의해 정의되는 정규화된 상관도 특징 혹은 보이싱 특징 FV<sub>1</sub>을 FV<sub>s</sub>(i) = sfa<sub>i</sub>FV<sub>i</sub> + sfb<sub>i</sub> (여기서, i = 0, ..., 11)에 근거하여 스케일링한 값 FV<sub>s</sub>(1)과 하기 수학식 18로 정의되는 상관도 맵 특징 FV(7)을 FV<sub>s</sub>(i) = sfa<sub>i</sub>FV<sub>i</sub> + sfb<sub>i</sub> (여기서, i = 0, ..., 11)에 근거하여 스케일링한 값 FV<sub>s</sub>(7) 간의 차이 d<sub>vcor</sub>는 d<sub>vcor</sub> = max(FV<sub>s</sub>(1)-FV<sub>s</sub>(7),0)로 정의될 수 있다.

**수학식 17**

$$FV_1 = C_{norm}^{[1]}$$

[0279]

[0280] 여기서, C<sub>norm</sub><sup>[1]</sup>은 첫번째 혹은 두번째 하프 프레임에서의 정규화된 상관도를 나타낸다.

수학식 18

$$FV_7 = \sum_{j=0}^{127} M_{cor}(j) + \sum_{j=0}^{127} M_{cor}^{[-1]}(j)$$

[0281]

[0282]

여기서,  $M_{cor}$ 은 프레임의 상관도 맵을 나타낸다.

[0283]

상기 복수개의 특징 파라미터를 조합하거나 단일한 특징 파라미터를 이용하여 다음 조건 1 내지 조건 4 중 적어도 하나 이상을 포함하는 수정 파라미터를 생성할 수 있다. 여기서, 조건 1과 조건 2은 음성 분류와 관련된 상태 머신에서 사용되는 조건을 의미하며, 조건 3과 조건 4는 음악 분류와 관련된 상태 머신에서 사용되는 조건을 의미할 수 있다. 구체적으로, 조건 1은 음성 상태(SPEECH\_STATE)를 0에서 1로 변경할 수 있고, 조건 2는 음성 상태(SPEECH\_STATE)를 1에서 0으로 변경할 수 있다. 한편, 조건 3은 음악 상태(MUSIC\_STATE)를 0에서 1로 변경할 수 있고, 조건 4는 음악 상태(MUSIC\_STATE)를 1에서 0으로 변경할 수 있다. 음성 상태(SPEECH\_STATE)가 1이면 음성일 확률이 높음 즉, CELP 타입 코딩이 적합함을 의미하고 0이면 음성이 아닐 확률이 높음을 의미할 수 있다. 일예를 들면, 음악 상태(MUSIC\_STATE)가 1이면 트랜스폼 코딩에 적합함을 의미하고, 0이면 CELP/트랜스폼 하이브리드 코딩 즉, GSC에 적합함을 의미할 수 있다. 다른 예를 들면, 음악 상태(MUSIC\_STATE)가 1이면 트랜스폼 코딩에 적합함을 의미하고, 0이면 CELP 타입 코딩에 적합함을 의미할 수 있다.

[0284]

조건 1( $cond_A$ )은 예를 들어 다음과 같이 정의될 수 있다. 즉,  $d_{vcor} > 0.4$  AND  $d_{ft} < 0.1$  AND  $FV_s(1) > (2 * FV_s(7) + 0.12)$  AND  $ton_2 < d_{vcor}$  AND  $ton_3 < d_{vcor}$  AND  $ton_{LT} < d_{vcor}$  AND  $FV_s(7) < d_{vcor}$  AND  $FV_s(1) > d_{vcor}$  AND  $FV_s(1) > 0.76$ 이면,  $cond_A$ 는 1로 설정될 수 있다.

[0285]

조건 2( $cond_B$ )은 예를 들어 다음과 같이 정의될 수 있다. 즉,  $d_{vcor} < 0.4$ 이면,  $cond_B$ 는 1로 설정될 수 있다.

[0286]

조건 3( $cond_C$ )은 예를 들어 다음과 같이 정의될 수 있다. 즉,  $0.26 < ton_2 < 0.54$  AND  $ton_3 > 0.22$  AND  $0.26 < ton_{LT} < 0.54$  AND  $LP_{err} > 0.5$ 이면,  $cond_C$ 는 1로 설정될 수 있다.

[0287]

조건 4( $cond_D$ )은 예를 들어 다음과 같이 정의될 수 있다. 즉,  $ton_2 < 0.34$  AND  $ton_3 < 0.26$  AND  $0.26 < ton_{LT} < 0.45$ 이면,  $cond_D$ 는 1로 설정될 수 있다.

[0288]

각 조건을 생성하기 위하여 사용된 특징 혹은 특징들의 조합은 이에 한정되지 않는다. 또한, 각 상수값은 예시적인 것에 불과하며 구현 방식에 따라서 최적값으로 설정될 수 있다.

[0289]

실시예에 따르면, 정정부(3730)는 두개의 독립적인 상태 머신, 예를 들면 음성 상태 머신과 음악 상태 머신을 이용하여 초기 분류 결과에 존재하는 에러를 정정할 수 있다. 각 상태 머신은 두개의 상태를 가지며, 각 상태에서 행오버가 사용되어 빈번한 트랜지션을 방지할 수 있다. 행오버는 예를 들어 6개 프레임으로 구성될 수 있다. 음성 상태 머신에서 행오버 변수를  $hang_{sp}$ 로, 음악 상태 머신에서 행오버 변수를  $hang_{mus}$ 로 나타내는 경우, 주어진 상태에서 분류 결과에 변화가 있는 경우, 각각 6으로 초기화되고, 이후 행오버가 각 다음 프레임에 대하여 1씩 감소하게 된다. 상태 변화는 행오버가 제로로 감소되는 경우에 발생할 수 있다. 각 상태 머신에는 오디오 신호로부터 추출되는 적어도 하나 이상의 특징이 조합되어 생성되는 수정 파라미터가 사용될 수 있다.

[0290]

도 38은 도 37에 도시된 정정부(3730)에서 사용되는 상태 머신을 나타내는 도면이다.

[0291]

도 38을 참조하면, 좌측은 일실시예에 따른 CELP 코어에 적합한 상태 머신 즉, 음성상태에서 컨텍스트 기반 정정을 위한 상태 머신을 나타내는 것이다. 정정부(3730)에서는 음악 상태 머신에서 결정되는 음악 상태와 음성 상태 머신에서 결정되는 음성상태에 따라서 분류 결과에 대한 수정(corection)이 적용될 수 있다. 예를 들어, 초기 분류결과가 음악 신호로 설정된 경우, 수정 파라미터에 근거하여 음성신호로 변경할 수 있다. 구체적으로, 초기 분류 결과 중 제1 단계의 분류결과가 음악 신호이고 음성 상태가 1이 된 경우, 제1 단계의 분류결과와 제2 단계의 분류결과 모두를 음성 신호로 변경할 수 있다. 이러한 경우, 초기 분류 결과에 에러가 존재하는 것으로 판단되어 분류 결과에 대한 수정이 이루어질 수 있다.

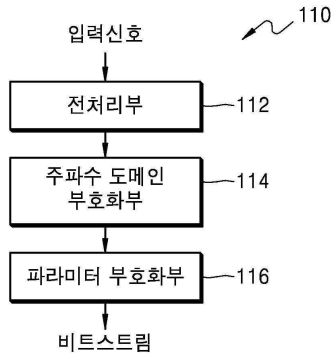
- [0292] 이에 대하여 좀 더 구체적으로 설명하면 다음과 같다.
- [0293] 먼저, 수정 파라미터, 예를 들면 조건 1 및 조건 2를 수신할 수 있다. 또한, 음성 상태 머신의 행오버 정보를 수신할 수 있다. 또한, 초기 분류 결과를 수신할 수 있다. 초기 분류 결과는 음성/음악 분류부(3710)로부터 제공될 수 있다.
- [0294] 초기 분류 결과 즉, 음성 상태가 0이면서 조건 1(cond<sub>A</sub>)이 1이고 음성 상태 머신의 행오버(hang<sub>sp</sub>)가 0인지를 판단할 수 있다. 음성 상태가 0이면서 조건 1이 1이고 음성 상태 머신의 행오버(hang<sub>sp</sub>)가 0인 것으로 판단된 경우, 430 단계에서 음성 상태를 1로 변경하고 행오버(hang<sub>sp</sub>)을 6으로 초기화시킬 수 있다.
- [0295] 한편, 초기 분류 결과 즉, 음성 상태가 1이면서 조건 2(cond<sub>B</sub>)가 1이고 음성 상태 머신의 행오버(hang<sub>sp</sub>)가 0인지를 판단할 수 있다. 음성 상태가 1이면서 조건 2가 1이고 음성 상태 머신의 행오버(hang<sub>sp</sub>)가 0인 것으로 판단된 경우, 음성 상태를 0으로 변경하고 행오버(hang<sub>sp</sub>)을 6으로 초기화시킬 수 있다. 한편, 음성 상태가 1이 아니거나 조건 2가 1이 아니거나 음성 상태 머신의 행오버(hang<sub>sp</sub>)가 0이 아닌 경우 행오버를 1만큼 감소시키는 행오버 업데이트를 수행할 수 있다.
- [0296] 도 38을 참조하면, 우측은 일실시예에 따른 HQ(High Quality) 코어에 적합한 상태 즉, 음악상태에서 컨텍스트 기반 정정을 위한 상태 머신을 나타내는 것이다. 정정부(37330)에서는 음악 상태 머신에서 결정되는 음악 상태와 음성 상태 머신에서 결정되는 음성 상태에 따라서 분류 결과에 대한 수정(corection)이 적용될 수 있다. 예를 들어, 초기 분류 결과가 음성 신호로 설정된 경우, 수정 파라미터에 근거하여 음악 신호로 변경할 수 있다. 구체적으로, 초기 분류 결과 중 제1 단계의 분류결과가 음성 신호이고 음악 상태가 1이 된 경우, 제1 단계의 분류결과와 제2 단계의 분류결과 모두를 음악 신호로 변경할 수 있다. 한편, 초기 분류 결과가 음악신호로 설정된 경우, 수정 파라미터에 근거하여 음성신호로 변경할 수 있다. 이러한 경우, 초기 분류 결과에 에러가 존재하는 것으로 판단되어 분류 결과에 대한 수정이 이루어질 수 있다.
- [0297] 이에 대하여 좀 더 구체적으로 설명하면 다음과 같다.
- [0298] 먼저, 수정 파라미터, 예를 들면 조건 3 및 조건 4를 수신할 수 있다. 또한, 음악 상태 머신의 행오버 정보를 수신할 수 있다. 또한, 초기 분류 결과를 수신할 수 있다. 초기 분류 결과는 음성/음악 분류부(3710)로부터 제공될 수 있다.
- [0299] 초기 분류 결과 즉, 음악 상태가 0이면서 조건 3(cond<sub>C</sub>)이 1이고 음악 상태 머신의 행오버(hang<sub>mus</sub>)가 0인지를 판단할 수 있다. 음악 상태가 0이면서 조건 3이 1이고 음악 상태 머신의 행오버(hang<sub>mus</sub>)가 0인 것으로 판단된 경우, 음악 상태를 1로 변경하고 행오버(hang<sub>mus</sub>)을 6으로 초기화시킬 수 있다.
- [0300] 한편, 초기 분류 결과 즉, 음악 상태가 1이면서 조건 4(cond<sub>D</sub>)가 1이고 음악 상태 머신의 행오버(hang<sub>mus</sub>)가 0인지를 판단할 수 있다. 음악 상태가 1이면서 조건 4가 1이고 음악 상태 머신의 행오버(hang<sub>mus</sub>)가 0인 것으로 판단된 경우, 음악 상태를 0으로 변경하고 행오버(hang<sub>mus</sub>)을 6으로 초기화시킬 수 있다. 한편, 음악 상태가 1이 아니거나 조건 4가 1이 아니거나 음악 상태 머신의 행오버(hang<sub>mus</sub>)가 0이 아닌 경우 행오버를 1만큼 감소시키는 행오버 업데이트를 수행할 수 있다.
- [0301] 상기 실시예들은 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.

[0302]

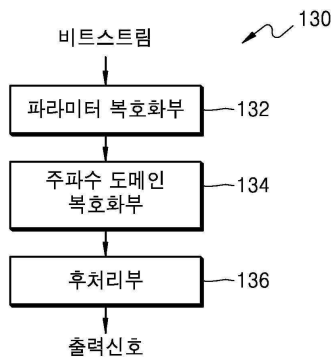
이상과 같이 본 발명의 일실시예는 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명의 일실시예는 상기 설명된 실시예에 한정되는 것은 아니며, 이는 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다. 따라서, 본 발명의 스코프는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 이의 균등 또는 등가적 변형 모두는 본 발명 기술적 사상의 범주에 속한다고 할 것이다.

**도면**

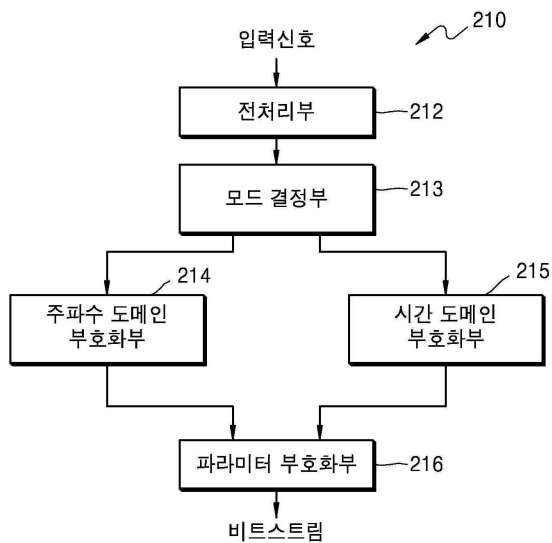
**도면1a**



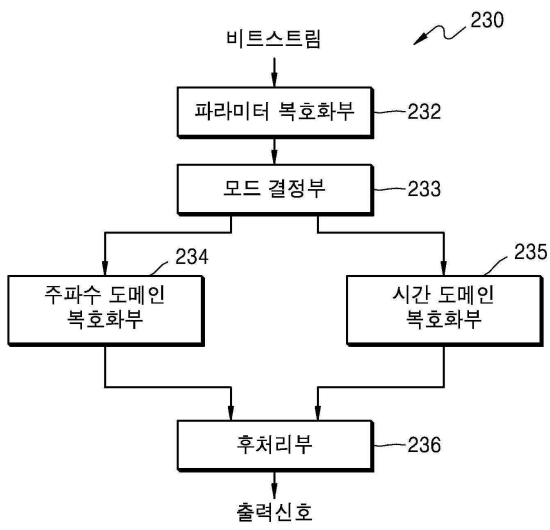
**도면1b**



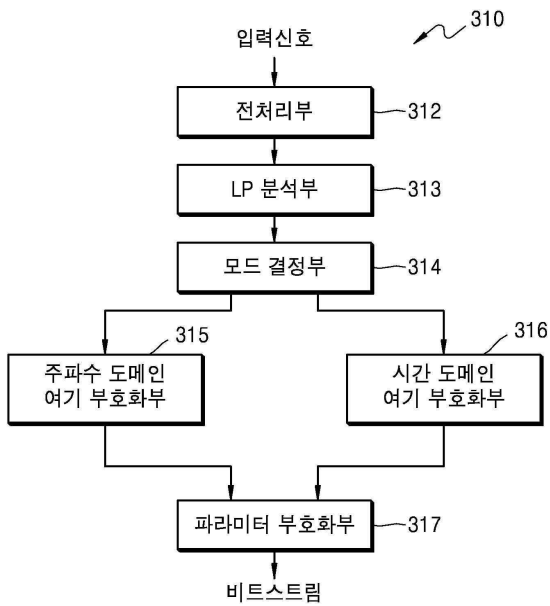
**도면2a**



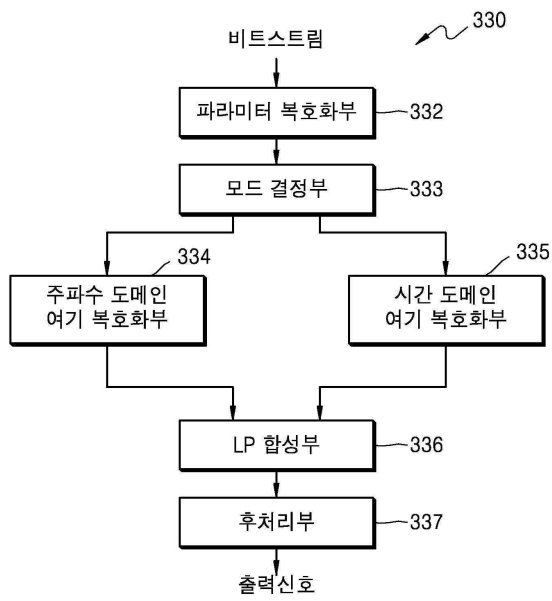
도면2b



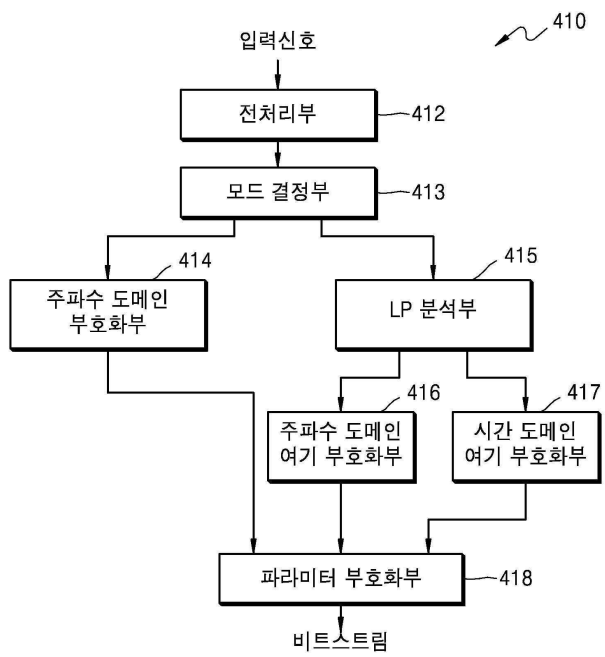
도면3a



도면3b

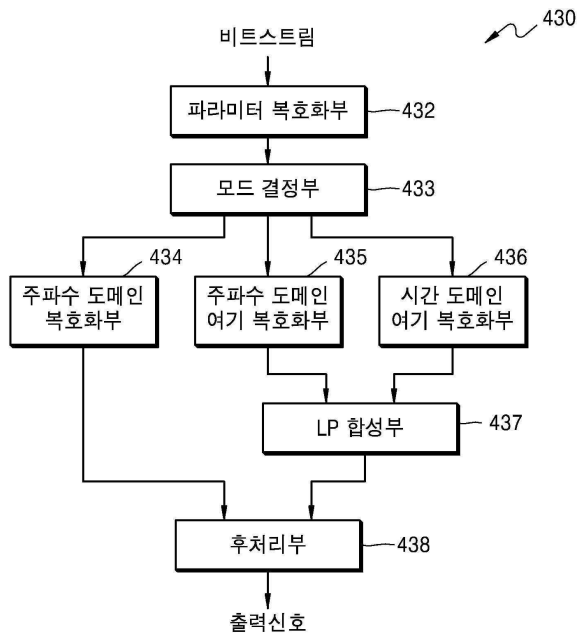


도면4a

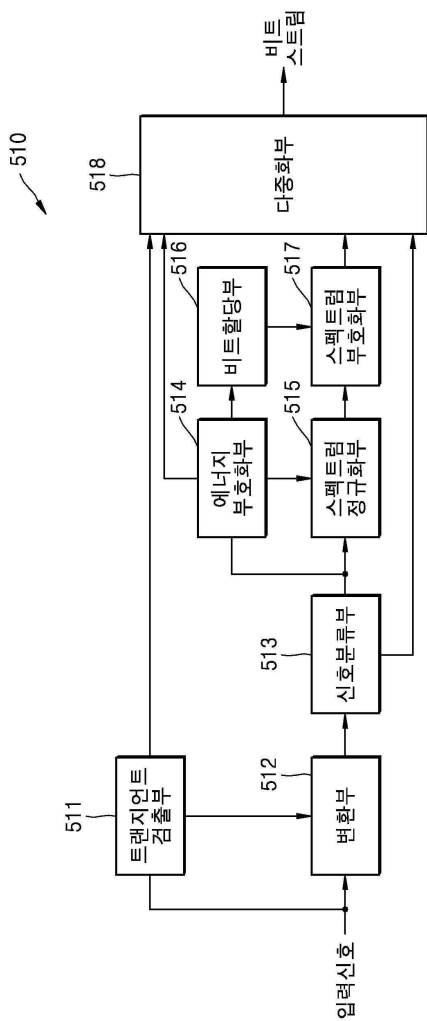




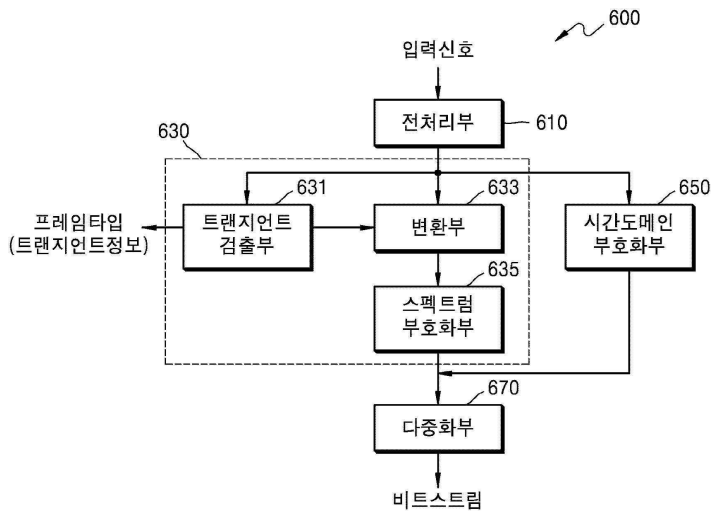
도면4b



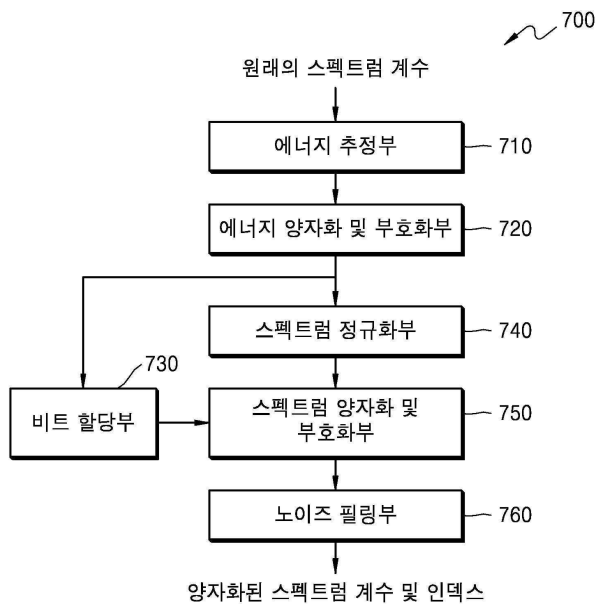
도면5



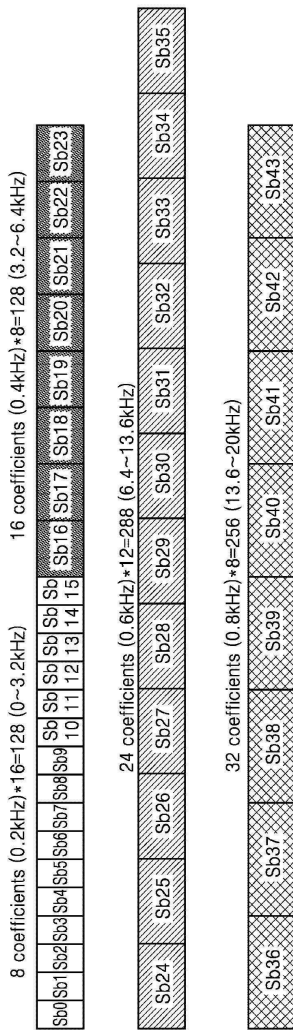
도면6



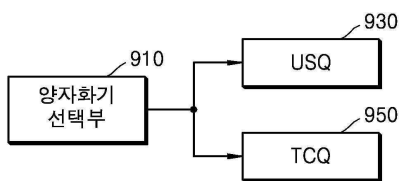
도면7



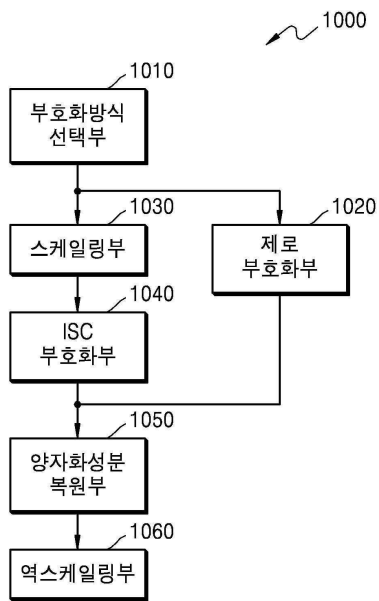
도면8



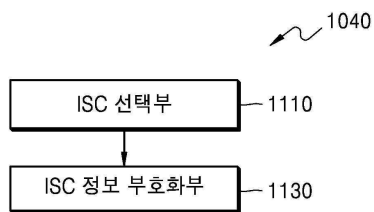
도면9



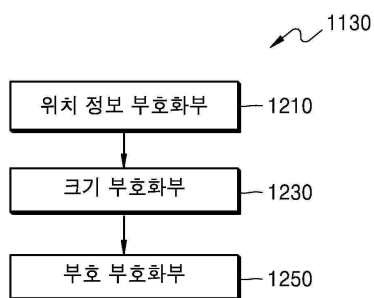
도면10



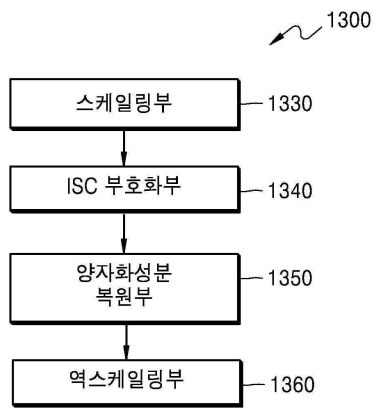
도면11



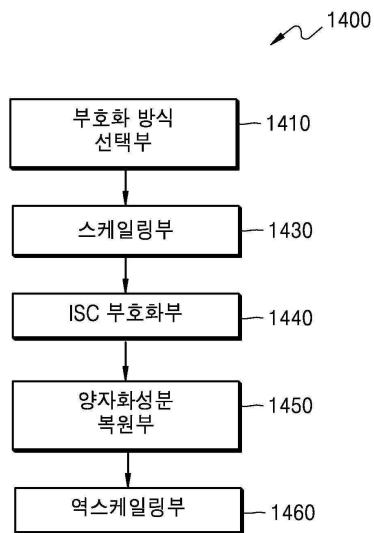
도면12



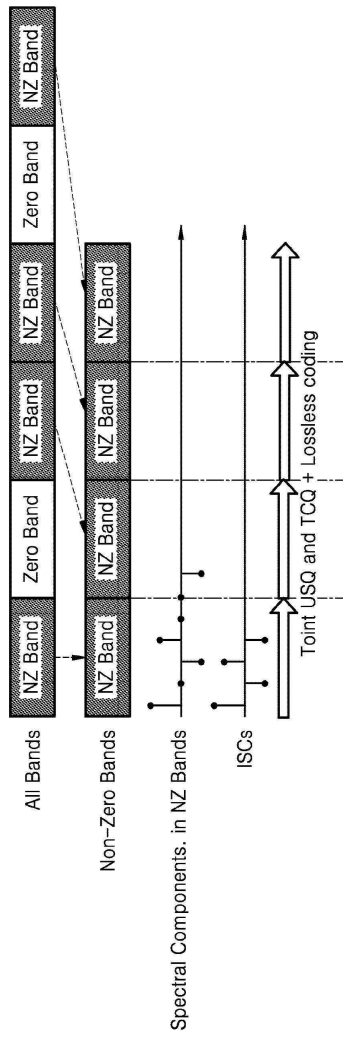
도면13



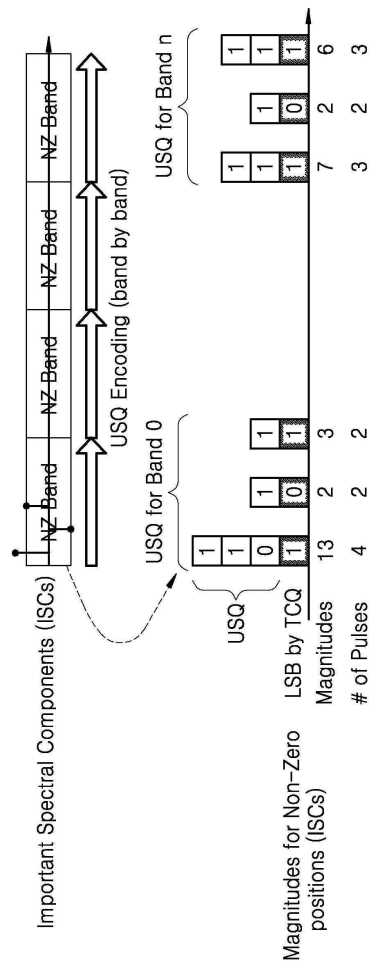
도면14



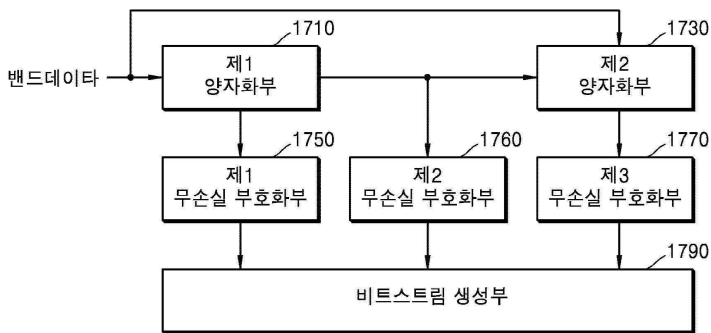
도면15



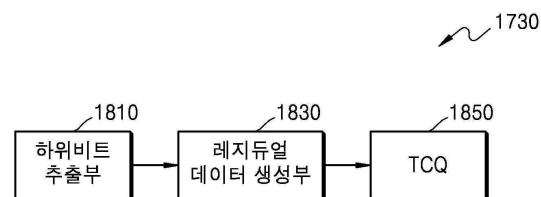
도면16



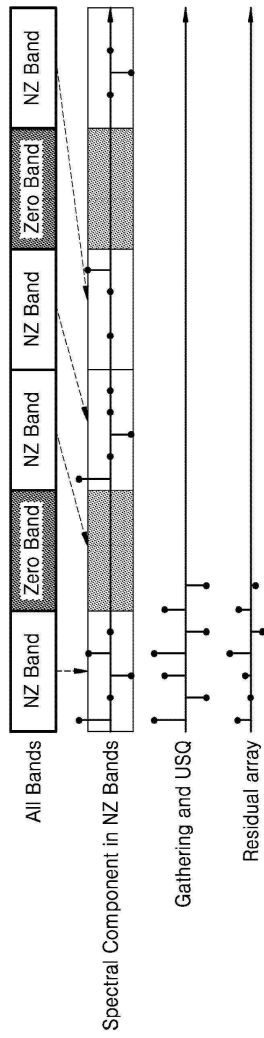
도면17



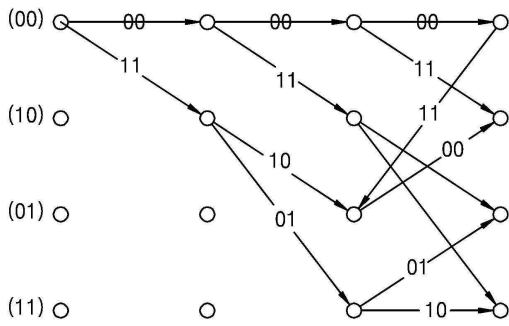
도면18



도면19

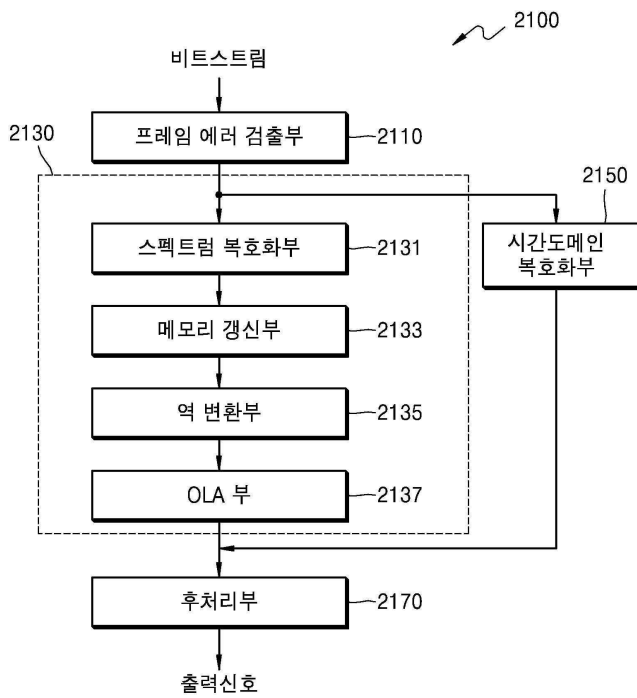


도면20

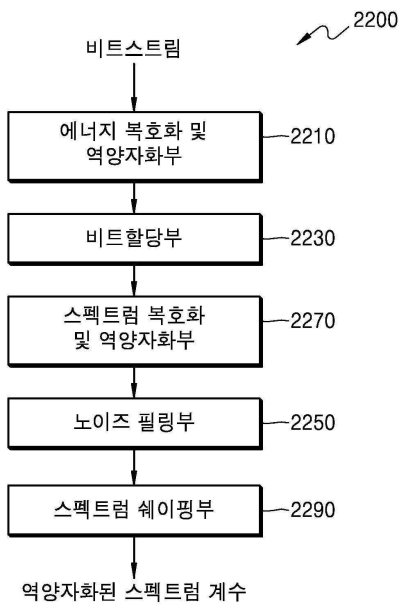




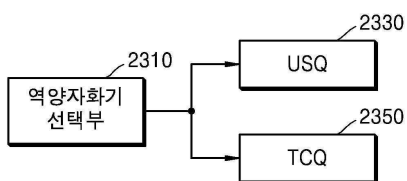
도면21



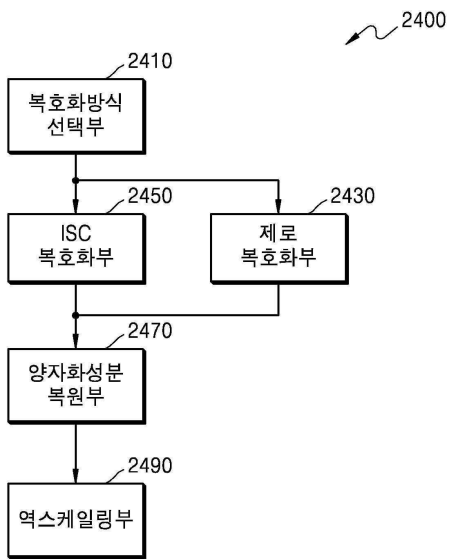
도면22



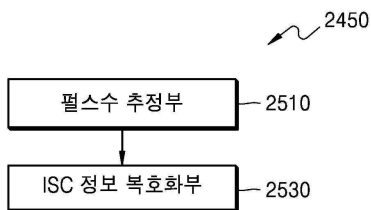
도면23



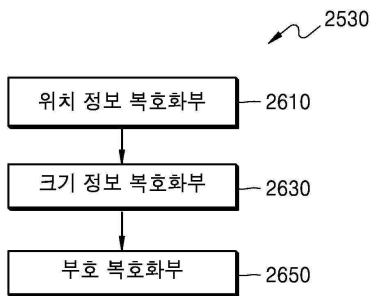
도면24



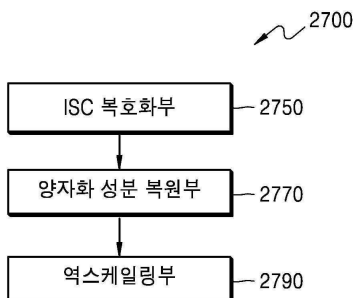
도면25



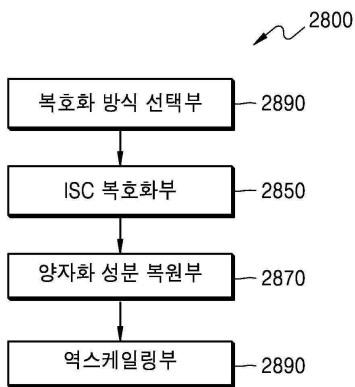
도면26



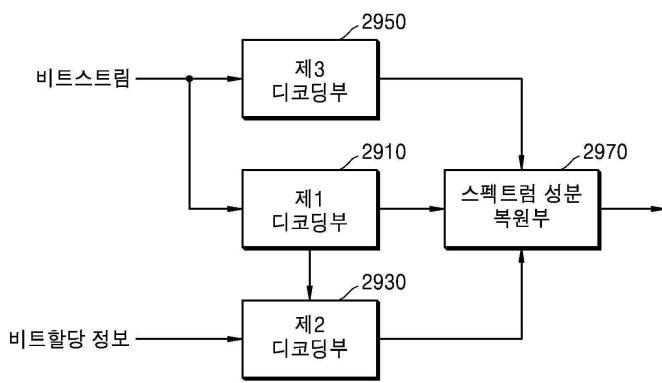
도면27



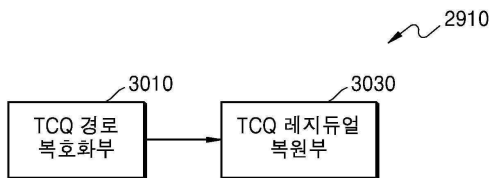
도면28



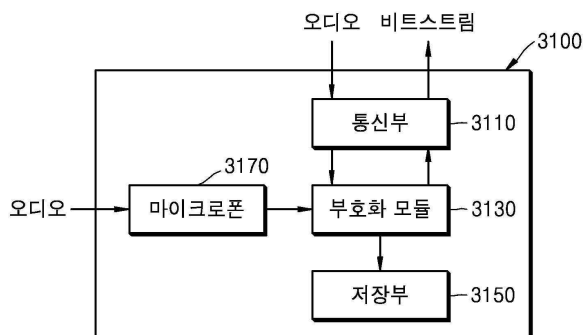
도면29



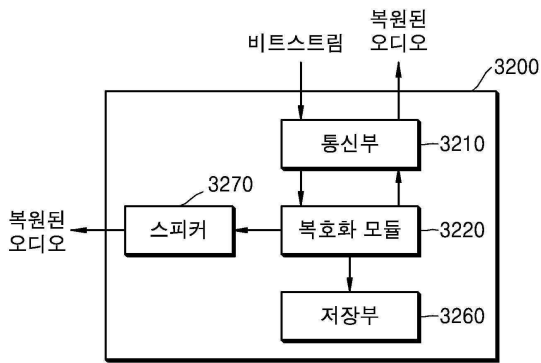
도면30



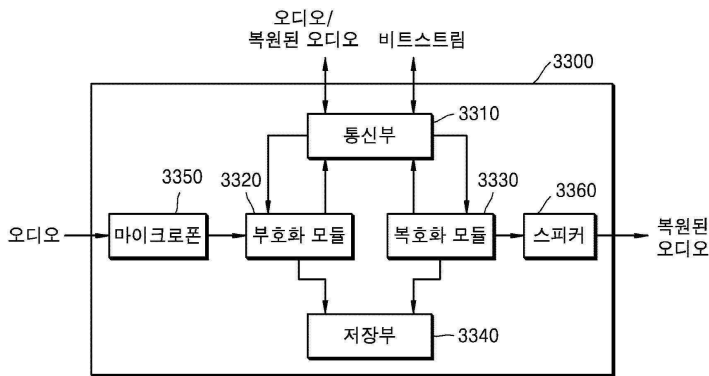
도면31



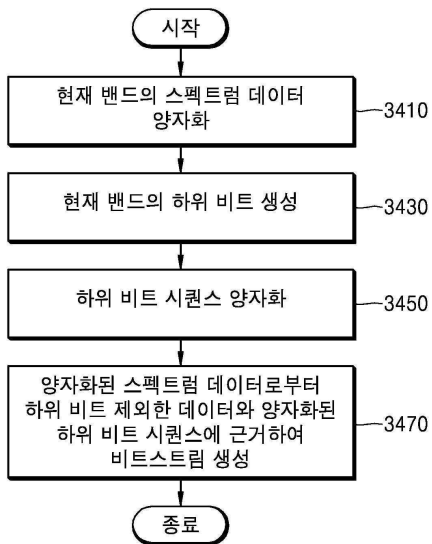
도면32



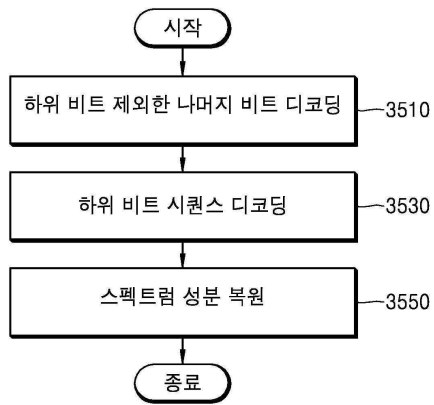
도면33



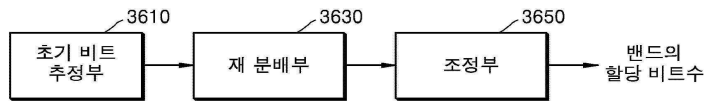
도면34



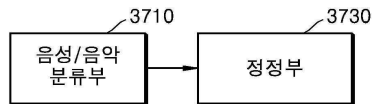
도면35



도면36



도면37



도면38

