



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2014년09월17일
(11) 등록번호 10-1441474
(24) 등록일자 2014년09월11일

(51) 국제특허분류(Int. Cl.)

H04B 7/26 (2006.01)

(21) 출원번호 10-2010-0013940

(22) 출원일자 2010년02월16일

심사청구일자 2013년05월06일

(65) 공개번호 10-2010-0093504

(43) 공개일자 2010년08월25일

(30) 우선권주장

1020090012356 2009년02월16일 대한민국(KR)

1020090092717 2009년09월29일 대한민국(KR)

(56) 선행기술조사문헌

KR1019980701162 A*

KR1020050078524 A*

*는 심사관에 의하여 인용된 문헌

기술이전 희망 : 기술양도, 실시권허여, 기술지도

(73) 특허권자

한국전자통신연구원

대전광역시 유성구 가정로 218 (가정동)

(72) 발명자

이미숙

대전광역시 서구 대덕대로 415, 109동 1103호 (만년동, 상아아파트)

배현주

대전광역시 유성구 가정로 63, 109동 401호 (신성동, 럭키하나아파트)

이병선

대전광역시 유성구 어은로 57, 107동 1104호 (어은동, 한빛아파트)

(74) 대리인

특허법인 신성

전체 청구항 수 : 총 12 항

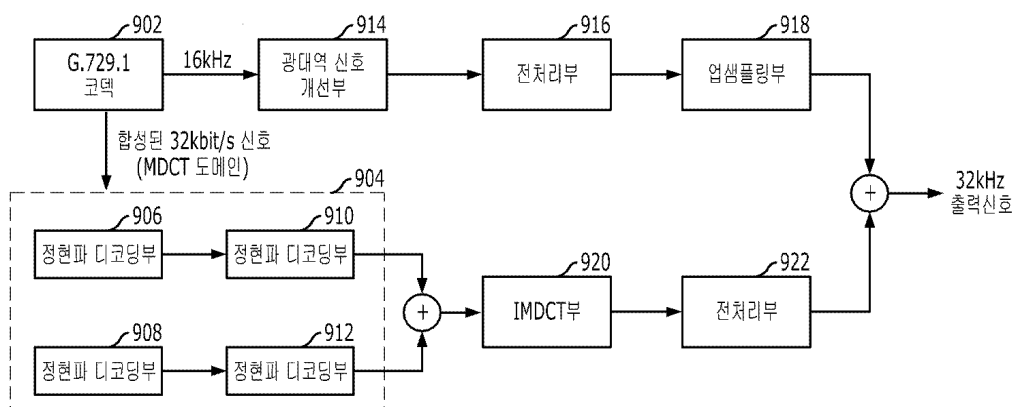
심사관 : 문해진

(54) 발명의 명칭 적응적 정현파 펄스 코딩을 이용한 오디오 신호의 인코딩 및 디코딩 방법 및 장치

(57) 요약

본 발명은 적응적 정현파 펄스 코딩을 이용한 오디오 신호의 인코딩 및 디코딩 방법 및 장치에 관한 것이다. 본 발명의 일 실시예에 의한 오디오 신호의 인코딩 방법은, 합성된 오디오 신호를 복수 개의 서브 대역들로 나누는 단계, 서브 대역들의 에너지를 각각 연산하는 단계, 서브 대역들 중 상대적으로 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택하는 단계 및 선택된 서브 대역들에 대하여 정현파 펄스 코딩을 수행하는 단계를 포함한다. 본 발명에 의하면 합성 신호의 각 서브 대역의 에너지 크기를 고려하여 정현파 펄스 코딩을 적용함으로써 보다 효율적으로 합성 신호의 품질을 향상시킬 수 있는 효과가 있다.

대표도



이 발명을 지원한 국가연구개발사업

과제고유번호 2008-S-011-01

부처명 지식경제부

연구사업명 IT성장동력기술개발

연구과제명 FMC 어커스틱 융합코덱 및 제어기술 연구 (표준화연계)

주관기관 한국전자통신연구원

연구기간 2008. 03. 01 ~ 2012. 02. 29

특허청구의 범위

청구항 1

오디오 신호를 부호화하는 방법에 있어서,

변환된 오디오 신호를 입력받는 단계;

상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하는 단계;

상기 서브 대역들의 에너지를 각각 연산하는 단계;

상기 서브 대역들의 에너지 크기가 큰 순서대로, 상기 서브 대역들에서 기 결정된 개수의 서브 대역들을 선택하는 단계; 및

상기 선택된 서브 대역들에 대해 정현파 코딩(Sinusoidal Coding)을 수행하는 단계;를 포함하며;

상기 정현파 코딩을 수행하는 단계는, 상기 서브 대역들의 에너지 크기를 바탕으로, 상기 선택된 서브 대역들을 트랙들로 선택한 후, 상기 트랙들에 대해 상기 정현파 코딩을 수행하는 것을 특징으로 하는 부호화 방법.

청구항 2

삭제

청구항 3

제1항에 있어서,

상기 정현파 코딩을 수행하는 단계는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 트랙으로 선택하는 것을 특징으로 하는 부호화 방법.

청구항 4

제1항에 있어서,

상기 정현파 코딩을 수행하는 단계는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 서브 대역으로 병합한 후, 상기 병합된 서브 대역에 대해 상기 정현파 코딩을 수행하는 것을 특징으로 하는 부호화 방법.

청구항 5

오디오 신호를 부호화하는 장치에 있어서,

변환된 오디오 신호를 입력받는 입력부;

상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하고, 상기 서브 대역들의 에너지를 각각 연산하며, 상기 서브 대역들의 에너지 크기가 큰 순서대로, 상기 서브 대역들에서 기 결정된 개수의 서브 대역들을 선택하는 연산부; 및

상기 선택된 서브 대역들에 대해 정현파 코딩(Sinusoidal Coding)을 수행하는 코딩부;를 포함하며;

상기 코딩부는, 상기 서브 대역들의 에너지 크기를 바탕으로, 상기 선택된 서브 대역들을 트랙들로 선택한 후, 상기 트랙들에 대해 상기 정현파 코딩을 수행하는 것을 특징으로 하는 부호화 장치.

청구항 6

삭제

청구항 7

제5항에 있어서,

상기 코딩부는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 트랙으로 선택하는 것을 특징으로 하는 부호화 장치.

청구항 8

제5항에 있어서,

상기 코딩부는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 서브 대역으로 병합한 후, 상기 병합된 서브 대역에 대해 상기 정현파 코딩을 수행하는 것을 특징으로 하는 부호화 장치.

청구항 9

오디오 신호를 복호화하는 방법에 있어서,

변환된 오디오 신호를 입력받는 단계;

상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하는 단계;

상기 서브 대역들의 에너지를 각각 연산하는 단계;

상기 서브 대역들의 에너지 크기가 큰 순서대로, 상기 서브 대역들에서 기 결정된 개수의 서브 대역들을 선택하는 단계; 및

상기 선택된 서브 대역들에 대해 정현파 디코딩(Sinusoidal Decoding)을 수행하는 단계;를 포함하며;

상기 정현파 디코딩을 수행하는 단계는, 상기 서브 대역들의 에너지를 바탕으로, 상기 선택된 서브 대역들을 트랙들로 선택한 후, 상기 트랙들에 대해 상기 정현파 디코딩을 수행하는 것을 특징으로 하는 복호화 방법.

청구항 10

삭제

청구항 11

제9항에 있어서,

상기 정현파 디코딩을 수행하는 단계는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 트랙으로 선택하는 것을 특징으로 하는 복호화 방법.

청구항 12

제9항에 있어서,

상기 정현파 디코딩을 수행하는 단계는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 서브 대역으로 병합한 후, 상기 병합된 서브 대역에 대해 상기 정현파 디코딩을 수행하는 것을 특징으로 하는 복호화 방법.

청구항 13

오디오 신호를 복호화하는 장치에 있어서,

변환된 오디오 신호를 입력받는 입력부;

상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하고, 상기 서브 대역들의 에너지를 각각 연산하며, 상기 서브 대역들의 에너지 크기가 큰 순서대로, 상기 서브 대역들에서 기 결정된 개수의 서브 대역들을 선택하는 연산부; 및

상기 선택된 서브 대역들에 대해 정현파 디코딩(Sinusoidal Decoding)을 수행하는 디코딩부;를 포함하며;

상기 디코딩부는, 상기 서브 대역들의 에너지 크기를 바탕으로, 상기 선택된 서브 대역들을 트랙들로 선택한 후, 상기 트랙들에 대해 상기 정현파 디코딩을 수행하는 것을 특징으로 하는 복호화 장치.

청구항 14

삭제

청구항 15

제13항에 있어서,

상기 디코딩부는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 트랙으로 선택하는 것을 특징으로 하는 복호화 장치.

청구항 16

제13항에 있어서,

상기 디코딩부는, 상기 선택된 서브 대역들 중 서로 인접한 서브 대역들을 하나의 서브 대역으로 병합한 후, 상기 병합된 서브 대역에 대해 상기 정현파 디코딩을 수행하는 것을 특징으로 하는 복호화 장치.

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

명세서

기술분야

- [0001] 본 발명은 오디오 신호의 인코딩 및 디코딩 방법 및 장치에 관한 것으로, 보다 상세하게는 적응적 정현과 펄스 코딩을 이용한 오디오 신호의 인코딩 및 디코딩 방법 및 장치에 관한 것이다.
- [0002] 본 발명은 지식경제부의 IT성장동력기술개발사업의 일환으로 수행한 연구로부터 도출된 것이다[과제관리번호: 2008-S-011-01, 과제명: FMC 어커스틱 융합코덱 및 제어기술 연구 (표준화연계)].

배경 기술

- [0003] 통신 기술의 발달과 함께 데이터 전송을 위한 대역폭이 증가하면서, 다채널 음성 및 오디오를 이용한 고품질 서비스에 대한 사용자의 요구가 점차 증가하고 있다. 고품질의 음성 및 오디오 서비스 제공을 위해서는 무엇보다도 스테레오 음성 및 오디오 신호를 효과적으로 압축하고 복원할 수 있는 코딩 기술이 필요하다.
- [0004] 이에 따라 협대역(Narrow Band: NB, 300~3,400 Hz), 광대역(Wide Band: WB, 50~7,000 Hz) 및 초광대역(Super Wide Band: SWB, 50~14,000 Hz) 신호를 코딩하는 코덱에 대한 연구가 활발히 진행되고 있다. 예를 들어, ITU-T G.729.1은 대표적인 확장코덱으로서, 협대역 코덱인 G.729를 기반으로 하는 광대역 확장코덱이다. 이 코덱은 8 kbit/s에서 G.729와 비트스트림 레벨의 호환성을 제공하고, 12 kbit/s에서는 보다 향상된 품질의 협대역 신호를 제공한다. 그리고 14 kbit/s부터 32 kbit/s에서는 2 kbit/s의 비트율 확장성을 가지고 광대역 신호를 코딩할 수 있으며, 비트율의 증가에 따라 출력신호의 품질도 좋아지는 특성을 가진다.
- [0005] 최근에는 G.729.1을 기반으로 초광대역 신호를 제공할 수 있는 확장코덱이 개발 중이다. 이 확장코덱은 협대역, 광대역, 그리고 초광대역 신호를 인코딩 및 디코딩할 수 있다.
- [0006] 이와 같은 확장코덱에서는 합성된 신호의 품질 향상을 위해 정현과 펄스 코딩을 이용하기도 한다. 그런데 이러한 정현과 펄스 코딩을 이용할 때, 코딩 효율을 높이기 위해서 입력 신호의 에너지를 고려할 필요가 있다. 특히 정현과 펄스 코딩에 사용할 수 있는 비트 수가 충분하지 않은 경우, 합성된 신호의 품질에 보다 많은 영향을 미치는 대역, 즉 상대적으로 에너지를 큰 대역을 우선적으로 코딩하는 것이 효과적이다.

발명의 내용

해결하려는 과제

- [0007] 본 발명은 합성 신호의 품질을 정현과 펄스 코딩을 이용하여 높일 수 있는 오디오 신호의 인코딩 및 디코딩 방법 및 장치를 제공하는 것을 일 목적으로 한다.
- [0008] 또한 본 발명은 합성 신호의 각 서브 대역의 에너지 크기를 고려하여 정현과 펄스 코딩을 적용함으로써 보다 효율적으로 합성 신호의 품질을 향상시킬 수 있는 오디오 신호의 인코딩 및 디코딩 방법 및 장치를 제공하는 것을 다른 목적으로 한다.
- [0009] 본 발명의 목적들은 이상에서 언급한 목적으로 제한되지 않으며, 언급되지 않은 본 발명의 다른 목적 및 장점들은 하기의 설명에 의해서 이해될 수 있고, 본 발명의 실시예에 의해 보다 분명하게 이해될 것이다. 또한, 본 발명의 목적 및 장점들은 특허 청구 범위에 나타낸 수단 및 그 조합에 의해 실현될 수 있음을 쉽게 알 수 있을 것이다.

과제의 해결 수단

- [0010] 이러한 목적을 달성하기 위한 본 발명은 오디오 신호의 인코딩 방법에 있어서, 변환된 오디오 신호를 복수 개의 서브 대역들로 나누는 단계, 서브 대역들의 에너지를 각각 연산하는 단계, 서브 대역들 중 상대적으로 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택하는 단계 및 선택된 서브 대역들에 대하여 정현과 펄스 코딩을 수행하는 단계를 포함하는 것을 일 특징으로 한다.
- [0011] 또한 본 발명은 오디오 신호의 인코딩 장치에 있어서, 변환된 오디오 신호를 입력받는 입력부, 합성된 오디오 신호를 복수 개의 서브 대역들로 나누고, 서브 대역들의 에너지를 각각 연산하며, 서브 대역들 중 상대적으로 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택하는 연산부 및 선택된 서브 대역들에 대하여 정현과

펄스 코딩을 수행하는 코딩부를 포함하는 것을 다른 특징으로 한다.

- [0012] 또한 본 발명은 오디오 신호의 디코딩 방법에 있어서, 변환된 오디오 신호를 입력받는 단계, 인코딩된 오디오 신호를 복수 개의 서브 대역들로 나누는 단계, 서브 대역들의 에너지를 각각 연산하는 단계, 서브 대역들 중 상대적으로 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택하는 단계 및 선택된 서브 대역들에 대하여 정현과 펄스 디코딩을 수행하는 단계를 포함하는 것을 또 다른 특징으로 한다.
- [0013] 또한 본 발명은 오디오 신호의 디코딩 장치에 있어서, 변환된 오디오 신호를 입력받는 입력부, 인코딩된 오디오 신호를 복수 개의 서브 대역들로 나누고, 서브 대역들의 에너지를 각각 연산하며, 서브 대역들 중 상대적으로 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택하는 연산부 및 선택된 서브 대역들에 대하여 정현과 펄스 디코딩을 수행하는 디코딩부를 포함하는 것을 또 다른 특징으로 한다.
- [0014] 또한 본 발명은 오디오 신호의 인코딩 방법에 있어서, 오디오 신호를 입력받는 단계, 오디오 신호에 대하여 MDCT(Modified Discrete Cosine Transform)를 수행하여 MDCT 계수를 출력하는 단계, MDCT 계수를 이용하여 고 주파수 오디오 신호를 합성하는 단계 및 고 주파수 오디오 신호에 대하여 정현과 펄스 코딩을 수행하는 단계를 포함하는 것을 또 다른 특징으로 한다.
- [0015] 또한 본 발명은 오디오 신호의 인코딩 장치에 있어서, 오디오 신호를 입력받는 입력부, 오디오 신호에 대하여 MDCT를 수행하여 MDCT 계수를 출력하는 MDCT부, MDCT 계수를 이용하여 고 주파수 오디오 신호를 합성하는 합성부 및 고 주파수 오디오 신호에 대하여 정현과 펄스 코딩을 수행하는 정현과 펄스 코딩부를 포함하는 것을 또 다른 특징으로 한다.
- [0016] 또한 본 발명은 오디오 신호의 디코딩 방법에 있어서, 오디오 신호를 입력받는 단계, 오디오 신호에 대하여 MDCT를 수행하여 MDCT 계수를 출력하는 단계, MDCT 계수를 이용하여 고 주파수 오디오 신호를 합성하는 단계 및 고 주파수 오디오 신호에 대하여 정현과 펄스 디코딩을 수행하는 단계를 포함하는 것을 또 다른 특징으로 한다.
- [0017] 또한 본 발명은 오디오 신호의 디코딩 장치에 있어서, 오디오 신호를 입력받는 입력부, 오디오 신호에 대하여 MDCT를 수행하여 MDCT 계수를 출력하는 MDCT부, MDCT 계수를 이용하여 고 주파수 오디오 신호를 합성하는 합성부 및 고 주파수 오디오 신호에 대하여 정현과 펄스 디코딩을 수행하는 정현과 펄스 디코딩부를 포함하는 것을 또 다른 특징으로 한다.

발명의 효과

- [0018] 전술한 바와 같은 본 발명에 의하면, 합성 신호의 품질을 정현과 펄스 코딩을 이용하여 높일 수 있는 장점이 있다.
- [0019] 또한 합성 신호의 각 서브 대역의 에너지 크기를 고려하여 정현과 펄스 코딩을 적용함으로써 보다 효율적으로 합성 신호의 품질을 향상시킬 수 있는 장점이 있다.

도면의 간단한 설명

- [0020] 도 1은 협대역 코덱과의 호환성을 제공하는 초광대역 확장코덱의 구조.
- 도 2는 본 발명의 일 실시예에 의한 오디오 신호 인코딩 장치의 구성도.
- 도 3은 본 발명의 일 실시예에 의한 오디오 신호 디코딩 장치의 구성도.
- 도 4는 본 발명의 일 실시예에 의한 오디오 신호 인코딩 방법의 흐름도.
- 도 5는 본 발명의 일 실시예에 의한 정현과 펄스 코딩 수행 단계(도 4의 s410)을 나타내는 흐름도.
- 도 6은 본 발명의 일 실시예에 의한 오디오 신호 디코딩 방법의 흐름도.
- 도 7은 기존 정현과 펄스 코딩 및 본 발명에 의한 적응형 정현과 펄스 코딩의 결과를 비교하기 위한 도면.
- 도 8은 본 발명의 다른 실시예에 의한 오디오 신호 인코딩 장치의 구성도.
- 도 9는 본 발명의 다른 실시예에 의한 오디오 신호 디코딩 장치의 구성도.

발명을 실시하기 위한 구체적인 내용

- [0021] 기술한 목적, 특징 및 장점은 첨부된 도면을 참조하여 상세하게 후술되며, 이에 따라 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명의 기술적 사상을 용이하게 실시할 수 있을 것이다. 본 발명을 설명함에 있어서 본 발명과 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우에는 상세한 설명을 생략한다. 이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 실시예를 상세히 설명하기로 한다. 도면에서 동일한 참조부호는 동일 또는 유사한 구성요소를 가리키는 것으로 사용된다.
- [0022] 도 1은 협대역 코덱과의 호환성을 제공하는 초광대역 확장코덱의 구조를 나타낸다.
- [0023] 일반적으로 확장코덱은 입력 신호를 여러 개의 주파수 대역으로 나눈 후에 각 주파수 대역의 신호를 인코딩 또는 디코딩하는 구조를 갖는다. 도 1을 참조하면, 입력된 신호는 1차 저대역 통과 필터(102) 및 1차 고대역 통과 필터(104)에 입력된다. 1차 저대역 통과 필터(102)는 필터링 및 다운 샘플링을 수행하여 입력 신호 중 저대역 신호 A(0-8kHz)를 출력한다. 그리고 1차 고대역 통과 필터(104)는 필터링 및 다운 샘플링을 수행하여 입력 신호 중 고대역 신호 B(8-16kHz)를 출력한다.
- [0024] 1차 저대역 통과 필터(102)에서 출력된 저대역 신호 A는 2차 저대역 통과 필터(106) 및 2차 고대역 통과 필터(108)에 입력된다. 2차 저대역 통과 필터(106)는 필터링 및 다운 샘플링을 수행하여 저-저대역 신호 A1(0-4kHz)를 출력하고, 2차 고대역 통과 필터(108)는 필터링 및 다운 샘플링을 수행하여 저-고대역 신호 A2(4-8kHz)를 출력한다.
- [0025] 결국 저-저대역 신호 A1은 협대역 코딩 모듈(110)로, 저-고대역 신호 A2는 광대역 확장코딩 모듈(112)로, 고대역 신호 B는 초광대역 확장코딩 모듈(114)로 각각 입력된다. 만약 협대역 코딩 모듈(110)만 동작할 경우에는 협대역 신호만이 재생되고, 협대역 코딩 모듈(110)과 광대역 확장코딩 모듈(112)이 동작할 경우에는 광대역 신호가 재생된다. 그리고 협대역 코딩 모듈(110), 광대역 확장코딩 모듈(112) 및 초광대역 확장코딩 모듈(114)이 동작하면 초광대역 신호가 재생된다.
- [0026] 도 1에 나타난 확장코덱의 대표적인 예로 ITU-T G.729.1을 들 수 있다. ITU-T G.729.1은 협대역 코덱인 G.729를 기반으로 하는 광대역 확장코덱이다. 이 코덱은 8 kbit/s에서 G.729와 비트스트림 레벨 호환성을 제공하고 12 kbit/s에서는 보다 향상된 품질의 협대역 신호를 제공한다. 그리고 14 kbit/s부터 32 kbit/s에서는 2 kbit/s의 비트율 확장성을 가지고 광대역 신호를 재생하는데, 비트율의 증가에 따라 출력신호의 품질도 좋아진다.
- [0027] 최근에는 G.729.1을 기반으로 초광대역 품질을 제공할 수 있는 확장코덱이 개발 중이다. 이 확장코덱은 협대역, 광대역, 그리고 초광대역 신호를 인코딩 및 디코딩할 수 있다.
- [0028] 이와 같은 확장코덱에서는 도 1과 같이 주파수 대역별로 다른 코딩방식을 적용할 수 있다. 예를 들어, G.729.1과 G.711.1코덱은 협대역 신호를 기존 협대역 코덱인 G.729와 G.711로 코딩하고, 나머지 신호에 대해서는 MDCT(Modified Discrete Cosine Transform)를 수행하여, 출력된 MDCT 계수를 코딩하는 방식을 사용한다.
- [0029] MDCT 영역 코딩에서는 MDCT 계수를 복수 개의 서브 대역으로 나누어서 각 서브 대역의 게인(gain)과 셰이프(shape)를 코딩하고, ACELP(Algebraic Code-Excited Linear Prediction) 혹은 정현파(sinusoidal) 펄스를 이용하여 MDCT 계수를 코딩한다. 확장코덱은 일반적으로 대역폭 확장을 위한 정보를 먼저 코딩한 후에 품질향상을 위한 정보를 코딩하는 구조를 가진다. 예를 들어, 각 서브 대역의 게인과 셰이프를 이용하여 7-14 kHz 대역의 신호를 합성한 후, ACELP 또는 정현파 펄스 코딩을 이용하여 합성된 신호의 품질을 향상시키는 구조가 그것이다.
- [0030] 즉, 초광대역 품질을 제공하는 첫 번째 계층에서는 게인과 셰이프 등의 정보를 이용하여 7-14 kHz 대역에 해당하는 신호를 합성한다. 그리고 추가적인 비트를 사용하여 합성된 신호의 품질향상을 위한 정현파 펄스 코딩 등을 적용한다. 이러한 구조를 통해, 비트율의 증가에 따라 합성된 신호의 품질을 개선시킬 수 있다.
- [0031] 일반적으로 정현파 펄스 코딩에서는 정해진 구간에서 크기가 가장 큰 펄스, 즉 품질에 가장 큰 영향을 미칠 수 있는 펄스의 위치, 크기, 그리고 부호 정보가 코딩된다. 이러한 펄스를 검색할 구간이 넓을수록 계산량은 증가한다. 따라서 전체 프레임(시간 영역의 경우) 또는 전체 주파수 대역에 대하여 정현파 펄스 코딩을 적용하는 것보다는 서브 프레임 또는 서브 대역별로 정현파 펄스 코딩을 적용하는 것이 바람직하다. 정현파 펄스 코딩은 하

나의 펄스를 전송하는데 상대적으로 많은 비트가 필요하지만 신호의 품질에 영향을 주는 신호를 정확히 표현할 수 있다는 장점을 가진다.

- [0032] 코덱의 입력 신호는 주파수에 따라 에너지 분포가 다양하게 나타난다. 특히 음악 신호의 경우에는 주파수에 따른 에너지의 변화가 음성 신호에 비해 큰 편이다. 에너지가 큰 서브 대역의 신호는 합성 신호의 품질에 보다 큰 영향을 미친다. 만일 전체 서브 대역을 코딩할 만큼 비트가 충분하다면 문제가 없지만, 그렇지 않은 경우에는 합성 신호의 품질에 영향을 많이 주는, 즉 에너지가 큰 서브 대역의 신호를 우선하여 코딩하는 것이 효과적이다.
- [0033] 본 발명은 도 1과 같은 확장코덱에서 한정된 비트 수를 고려하여 보다 효율적인 정현과 펄스 코딩을 수행함으로써 합성된 신호의 품질을 향상시킬 수 있는 오디오 신호의 인코딩 및 디코딩에 관한 것이다. 이하에서는 음성 및 오디오 신호를 오디오 신호로 지칭하여 본 발명을 설명한다.
- [0034] 도 2는 본 발명의 일 실시예에 의한 오디오 신호 인코딩 장치의 구성도이다.
- [0035] 도 2에 나타난 바와 같이, 오디오 신호 인코딩 장치(202)는 입력부(204), 연산부(206), 코딩부(208)를 포함한다. 입력부(204)는 변환된 오디오 신호, 예를 들면 오디오 신호가 MDCT에 의해 변환된 결과물인 MDCT 계수를 입력받는다.
- [0036] 연산부(206)는 입력부(204)를 통해 입력된 변환된 오디오 신호를 복수 개의 서브 대역들로 나누고, 이 서브 대역들의 에너지를 각각 연산한다. 그리고 연산부(206)는 서브 대역들 중 상대적으로 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택한다. 이 때, 하나의 서브 대역에서 코딩할 정현과 펄스 수 및 하나의 정현과 펄스를 코딩하는 데 필요한 비트 수에 의해 미리 정해진 수가 결정된다.
- [0037] 코딩부(208)는 연산부(206)에 의해 선택된 서브 대역들에 대하여 정현과 펄스 코딩을 수행한다. 코딩부(208)는 미리 정해진 수의 상대적으로 큰 에너지를 갖는 서브 대역들에 대하여 에너지가 큰 순서대로 정현과 펄스 코딩을 수행할 수 있다. 본 발명의 다른 실시예에서, 코딩부(208)는 미리 정해진 수의 상대적으로 큰 에너지를 갖는 서브 대역들에 대하여 에너지가 큰 순서가 아닌 다른 순서, 예를 들면 대역폭 또는 인덱스 순서로 정현과 펄스 코딩을 수행할 수 있다.
- [0038] 한편, 연산부(206)는 선택된 서브 대역들 중 서로 인접한 서브 대역들이 있는지 여부를 확인하고, 인접한 서브 대역들을 하나의 서브 대역으로 병합할 수 있다. 그리고 코딩부(208)는 이렇게 병합된 서브 대역에 대하여 정현과 펄스 코딩을 수행할 수 있다.
- [0039] 도 3은 본 발명의 일 실시예에 의한 오디오 신호 디코딩 장치의 구성도이다.
- [0040] 도 3에 나타난 바와 같이, 오디오 신호 디코딩 장치(302)는 입력부(304), 연산부(306), 디코딩부(308)를 포함한다. 입력부(304)는 변환된 오디오 신호, 예를 들면 MDCT 계수를 입력받는다.
- [0041] 연산부(306)는 입력부(304)를 통해 입력된 변환된 오디오 신호를 복수 개의 서브 대역들로 나누고, 이 서브 대역들의 에너지를 각각 연산한다. 그리고 연산부(306)는 서브 대역들 중 상대적으로 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택한다. 이 때, 하나의 서브 대역에서 코딩할 정현과 펄스 수 및 하나의 정현과 펄스를 코딩하는 데 필요한 비트 수에 의해 미리 정해진 수가 결정된다.
- [0042] 디코딩부(308)는 연산부(306)에 의해 선택된 서브 대역들에 대하여 정현과 펄스 디코딩을 수행한다. 디코딩부(308)는 미리 정해진 수의 상대적으로 큰 에너지를 갖는 서브 대역들에 대하여 에너지가 큰 순서대로 정현과 펄스 코딩을 수행할 수 있다. 본 발명의 다른 실시예에서, 디코딩부(308)는 미리 정해진 수의 상대적으로 큰 에너지를 갖는 서브 대역들에 대하여 에너지가 큰 순서가 아닌 다른 순서, 예를 들면 대역폭 또는 인덱스 순서로 정현과 펄스 코딩을 수행할 수 있다.
- [0043] 도 2 및 도 3에 나타난 오디오 신호 인코딩 장치(202) 및 오디오 신호 디코딩 장치(302)는 도 1의 협대역 코딩 모듈(110), 광대역 확장코딩 모듈(112) 또는 초광대역 확장코딩 모듈(114)에 포함될 수 있다.
- [0044] 이하에서는 도 4 내지 도 6 및 도 1의 초광대역 확장코딩 모듈(114)에서 오디오 신호를 인코딩 또는 디코딩하는 실시예를 통해 본 발명에 의한 오디오 신호 인코딩 및 디코딩 방법을 설명한다.
- [0045] 초광대역 확장코딩 모듈(114)은 7-14kHz에 해당하는 MDCT 계수를 여러 개의 서브 대역으로 나누고, 각 서브 대

역의 계인과 세이프를 코딩 또는 디코딩하여 오차 신호를 구한다. 그리고 나서 초광대역 확장코딩 모듈(114)은 오차 신호에 대하여 정현과 펄스 코딩 또는 디코딩을 수행한다. 만약 정현과 펄스 코딩에 사용되는 비트 수가 충분하다면 모든 서브 대역에 정현과 펄스 코딩이 적용될 수 있겠지만, 실제로 그러한 비트 수가 충분하지 않기 때문에 일부 서브 대역에 대해서만 정현과 펄스 코딩이 적용된다. 따라서 합성 신호의 품질에 더 많은 영향을 미치는 서브 대역에 정현과 펄스 코딩을 적용함으로써 같은 비트율이 주어졌을 때 보다 나은 신호 품질을 얻을 수 있다.

[0046] 도 4는 본 발명의 일 실시예에 의한 오디오 신호 인코딩 방법의 흐름도이다.

[0047] 도 4를 참조하면, 초광대역 확장코딩 모듈(114)에 포함된 오디오 신호 인코딩 장치는 변환된 오디오 신호, 예를 들면 7-14kHz에 해당하는 MDCT 계수를 입력받는다(s402). 그리고 입력받은 변환된 오디오 신호를 복수 개의 서브 대역으로 나누고(s404), 나누어진 복수 개의 서브 대역들의 에너지를 각각 연산한다(s406). 도 7은 9개의 서브 대역으로 나누어진 MDCT 계수와, 각 서브 대역의 상대적 에너지 크기를 나타낸다. 도 7을 참조하면, 서브 대역 1, 4, 5, 6, 7의 에너지가 다른 서브 대역의 에너지에 비해 상대적으로 큰 것을 알 수 있다.

[0048] 표 1은 8개의 서브 대역으로 나누어진 MDCT 계수의 인덱스 및 에너지를 나타낸다.

표 1

인덱스	1	2	3	4	5	6	7	8
에너지	350	278	657	245	1500	780	200	190

[0049]

[0050] 오디오 신호 인코딩 장치는 서브 대역들 중 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택한다(s408). 예를 들어, 표 1의 MDCT 계수는 표 2와 같이 에너지 순으로 정렬되고, 그 중 에너지가 가장 큰 5개의 서브 대역(음영 표시)이 선택된다.

표 2

인덱스	5	6	3	1	2	4	7	8
에너지	1500	780	657	350	278	245	200	190

[0051]

[0052] 본 발명에서는 표 2와 같이 미리 정해진 수, 예를 들면 5개의 서브 대역을 선택하게 된다. 이 미리 정해진 수는 하나의 서브 대역에서 코딩할 정현과 펄스 수 및 하나의 정현과 펄스를 코딩하는 데 필요한 비트 수에 의해 결정된다.

[0053] 하나의 정현과 펄스를 코딩하는 데 필요한 비트 수는 다음과 같다. 먼저 하나의 정현과 펄스의 부호(+, -)를 코딩하는 데 1비트가 필요하다. 그리고 정현과 펄스의 위치를 코딩하는 데 필요한 비트 수는 정현과 펄스 검색 구간의 크기, 예를 들면 서브 대역 하나의 크기에 의해 결정된다. 만약 서브 대역의 크기가 32라면 정현과 펄스의 위치를 코딩하는 데에는 5비트가 필요하다. ($2^5=32$) 그리고 정현과 펄스의 크기(이득)를 코딩하는 데 필요한 비트 수는 양자화기의 구조 및 코드 북의 크기에 의해 결정된다. 결국 하나의 정현과 펄스를 코딩하는 데 필요한 비트 수는 정현과 펄스의 부호, 위치, 크기를 코딩하는 데 필요한 비트들의 합이다.

[0054] 정현과 펄스 코딩을 위해 주어진 비트 및 하나의 정현과 펄스를 코딩하는 데 필요한 비트 수에 의해 10개의 정현과 펄스를 전송할 수 있을 경우, 하나의 서브 대역 당 2개의 정현과 펄스를 코딩한다고 하면 총 5개의 서브 대역에 정현과 펄스 코딩을 적용할 수 있다. 따라서 오디오 신호 코딩 장치는 표 2와 같이 에너지가 가장 큰 5개의 서브 대역을 선택하고, 선택된 서브 대역들(5, 6, 3, 1, 2)에 대하여 정현과 펄스 코딩을 수행한다(s410).

[0055] 도 5는 본 발명의 일 실시예에 의한 정현과 펄스 코딩 수행 단계(도 4의 s410)를 나타내는 흐름도이다.

[0056] 본 발명의 다른 실시예에서는 도 4의 단계 s408에서 선택된 서브 대역들 중 서로 인접한 서브 대역들이 있는지 여부를 확인하고(s502), 인접한 서브 대역들을 하나의 서브 대역으로 병합하며(s504), 병합된 서브 대역에 대하여 정현파 펄스 코딩을 수행한다(s506).

[0057] 예를 들어, 표 2에서 선택된 다섯 개의 서브 대역(5, 6, 3, 1, 2) 중 서브 대역 5와 인접한 서브 대역, 즉 4 또는 6이 있는지 여부를 확인한다. 다섯 개의 서브 대역 중에서 서브 대역 5와 인접한 서브 대역 6이 존재하므로, 오디오 신호 인코딩 장치는 서브 대역 5와 서브 대역 6 각각에 2개의 정현파 펄스를 코딩하는 대신, 두 서브 대역을 하나의 서브 대역으로 병합하고 이 하나의 서브 대역에 대하여 4개의 정현파 펄스를 코딩한다. 예를 들어 서브 대역 5가 서브 대역 6보다 큰 에너지를 갖고 있다면, 병합된 서브 대역에서는 4개의 정현파 펄스가 모두 서브 대역 5에만 위치할 수 있다. 이처럼 인접한 서브 대역을 병합하고 병합된 서브 대역에 정현파 펄스 코딩을 적용함으로써 보다 효율적인 정현파 펄스 코딩이 이루어진다.

[0058] 한편, 코덱의 특성에 따라 인코더와 디코더에서 합성된 7-14kHz 대역의 신호가 서로 일치하지 않을 수 있다. 인코더와 디코더에서 각각 연산된 서브 대역들의 에너지의 차이로 인한 오차를 줄이기 위하여, 오디오 신호 인코딩 장치는 표 3과 같이 서브 대역을 재정렬하여 정현파 펄스 코딩을 수행할 수도 있다.

표 3

인덱스	1	2	3	5	6	4	7	8
에너지	350	278	657	1500	780	245	200	190

[0059]

[0060] 즉, 오디오 신호 인코딩 장치는 5개의 서브 대역에 대하여 에너지 크기 순서대로 정현파 펄스 코딩을 수행하지 않고, 대역 폭 순서 또는 인덱스 순서에 따라 정현파 펄스 코딩을 수행할 수 있다. 이처럼 선택된 서브 대역들의 에너지 크기 순서를 고려하지 않음으로써, 인코더와 디코더에서 발생할 수 있는 상위대역 합성 신호의 차이로 인한 오차를 줄일 수 있다.

[0061] 도 6은 본 발명의 일 실시예에 의한 오디오 신호 디코딩 방법의 흐름도이다.

[0062] 먼저 변환된 오디오 신호를 입력받는다(s602). 그리고 변환된 오디오 신호를 복수 개의 서브 대역들로 나누고(s604), 서브 대역들의 에너지를 각각 연산한다(s606).

[0063] 그 다음, 서브 대역들 중 큰 에너지를 갖는 서브 대역들을 미리 정해진 수만큼 선택하고(s608), 선택된 서브 대역들에 대하여 정현파 펄스 디코딩을 수행한다(s610). 도 6의 단계 s602 내지 단계 s610은 앞서 설명한 본 발명의 일 실시예에 의한 오디오 신호 인코딩 방법의 각 단계와 유사하므로, 자세한 설명은 생략한다.

[0064] 도 7은 기존 정현파 펄스 코딩 및 본 발명에 의한 적응형 정현파 펄스 코딩의 결과를 비교하기 위한 도면이다.

[0065] 도 7의 (a)는 기존의 정현파 펄스 코딩에 의한 결과를 나타낸다. 도 7에 나타난 각 서브 대역의 상대적 에너지 크기를 보면, 서브 대역 1, 4, 5, 6, 7의 에너지가 다른 서브 대역에 비해 상대적으로 크다. 하지만 기존의 정현파 펄스 코딩은 서브 대역의 에너지 크기와는 상관 없이 대역 순 또는 인덱스 순으로 펄스 코딩을 적용하므로, (a)와 같이 서브 대역 1, 2, 3, 4, 5에 정현파 펄스가 코딩된다.

[0066] 도 7의 (b)는 본 발명에 의한 적응적 정현파 펄스 코딩에 의한 결과를 나타낸다. 본 발명에 의하면 (b)와 같이 상대적으로 에너지가 큰 서브 대역, 즉 서브 대역 1, 4, 5, 6, 7에 정현파 펄스 코딩이 적용된다.

[0067] 앞서 언급한 바와 같이, 본 발명은 음성을 포함하는 오디오 신호에 적용된다. 음성 신호의 에너지 분포를 보면, 유성음의 에너지는 대부분 낮은 주파수 대역에 위치하고, 무성음과 과열음의 에너지는 상대적으로 높은 주파수 대역에 위치한다. 반면에, 음악 신호는 주파수에 따라 에너지의 변화가 다양하게 나타난다. 따라서 음악 신호는 음성 신호와 같이 주파수 대역에 따른 에너지 분포의 특징을 정의하기가 어렵다. 합성된 신호의 품질은 에너지가 큰 주파수 대역의 신호에 의해 더 많은 영향을 받는다. 따라서 정현파 펄스 코딩을 적용할 서브 대역을 고정시키는 것보다는 본 발명과 같이 입력 신호의 특성에 따라 서브 대역을 선택하여 정현파 펄스 코딩을 적용함으로써 똑같은 비트율로 합성된 신호의 품질을 보다 향상시킬 수 있다.

- [0068] 이하에서는 도 8 및 도 9를 통해 본 발명의 다른 실시예에 의한 오디오 신호 인코딩 및 디코딩 방법 및 장치에 대해 설명한다.
- [0069] 도 8은 본 발명의 다른 실시예에 의한 오디오 신호 인코딩 장치의 구성도이다.
- [0070] 도 8에 나타난 오디오 신호 인코딩 장치는 32kHz의 입력 신호를 입력받고, 광대역 신호 및 초광대역 신호를 합성하여 출력한다. 이 오디오 신호 인코딩 장치는 광대역 확장코딩 모듈(802, 808, 822)과 초광대역 확장코딩 모듈(804, 806, 810, 812)로 구성된다. 광대역 확장코딩 모듈, 즉 G.729.1 코어 코덱(core codec)은 16kHz 신호를 이용하여 동작하는 반면에, 초광대역 확장코딩 모듈은 32kHz 신호를 이용한다. 초광대역 확장코딩은 MDCT 도메인에서 수행된다. 두 개의 모드, 즉 제네릭 모드(814)와 정현파 모드(816)가 초광대역 확장코딩 모듈의 첫 번째 계층을 코딩하기 위하여 이용된다. 제네릭 모드(814) 또는 정현파 모드(816) 중 어떤 것을 이용할지 여부는 입력 신호의 측정된 톤리티(Tonality)에 기반하여 결정된다. 보다 상위의 초광대역 계층들은 고 주파수 콘텐츠(content)의 품질을 개선하는 정현파 코딩부(818, 820), 또는 광대역 콘텐츠의 인지 품질(perceptual quality)을 개하는데 이용되는 광대역 신호 개선부(822)에 의하여 코딩된다.
- [0071] 32kHz의 입력 신호는 먼저 다운 샘플링부(802)에 입력되고, 16kHz로 다운 샘플링된다. 그리고 다운 샘플링된 16kHz 신호는 G.729.1 코덱(808)에 입력된다. G.729.1 코덱(808)은 입력된 16kHz 신호에 대하여 광대역 코딩을 수행한다. G.729.1 코덱(808)에서 출력된 합성된 32kbit/s 신호는 광대역 신호 개선부(822)로 입력되고, 광대역 신호 개선부(822)는 입력된 신호의 품질을 개선한다.
- [0072] 한편, 32kHz 입력 신호는 MDCT부(806)에 입력되고 MDCT 도메인으로 변환된다. MDCT 도메인으로 변환된 입력 신호는 톤리티 측정부(804)에 입력되고 입력 신호의 톤(tonal) 여부가 결정된다(810). 다시 말해, 첫 번째 초광대역 계층의 코딩 모드는 MDCT 도메인에서 입력 신호의 현재 프레임 및 이전 프레임의 로그 도메인 에너지(logarithmic domain energies)를 비교함으로써 수행되는 톤리티 측정에 기반하여 정의된다. 톤리티 측정은 입력 신호의 현재 프레임과 과거 프레임의 스펙트럴 피크(spectral peaks) 간의 상관관계 분석(correlation analysis)에 기반한다.
- [0073] 그 다음, 톤리티 측정부(804)에 의해 출력된 톤리티 정보에 의해 입력 신호가 톤인지 아닌지 여부가 결정된다(810). 예를 들어, 톤리티 정보가 특정 임계값(threshold)보다 크면 입력 신호는 톤인 것으로, 그렇지 않으면 입력 신호는 톤이 아닌 것으로 판단된다. 톤리티 정보는 또한 디코더로 전달되는 비트스트림에도 포함된다. 만약 입력 신호가 톤이면 정현파 모드(816)가, 그렇지 않으면 제네릭 모드(814)가 이용된다.
- [0074] 제네릭 모드(814)는 입력 신호의 프레임이 톤이 아닐 때(tonal=0) 이용된다. 제네릭 모드(814)는 고 주파수들을 코딩하기 위하여 G.729.1 광대역 코덱(808)의 코딩된 MDCT 도메인 표현을 활용한다. 고 주파수 대역(7-14kHz)은 4개의 서브 대역으로 나누어지고, 코딩되고 인벨로프 표준화된(envelope normalized) 광대역 콘텐츠로부터 각각의 서브 밴드에 대한 선택된 유사성 기준(similarity criteria)이 탐색된다. 가장 유사한 매치(match)는 합성된 고 주파수 콘텐츠를 획득하기 위하여 두 개의 스케일링 요소들, 즉 리니어(linear) 도메인의 첫 번째 스케일링 요소 및 로그 도메인의 두 번째 스케일링 요소에 의해 스케일링 된다. 이 콘텐츠는 또한 제네릭 모드(814) 및 정현파 코딩부(818) 내의 추가적인 정현파들에 의해 개선된다.
- [0075] 제네릭 모드(814)에서는 본 발명에 의한 오디오 인코딩 방법에 의하여, 코딩된 신호의 품질 개선이 이루어질 수 있다. 예를 들어, 비트 버짓(bit budget)은 첫 4kbit/s의 초광대역 계층에 두 개의 정현파를 추가하도록 허용한다. 추가할 정현파의 위치를 탐색할 트랙의 시작 위치는 합성된 고 주파수 신호의 서브 대역 에너지에 기반하여 선택된다. 합성된 서브 대역들의 에너지는 다음 수학적 식 1과 같이 연산될 수 있다.

수학적 식 1

$$SbE(k) = \sum_{n=0}^{n=31} \ddot{M}_{32}(k \times 32 + n)^2 \quad k = 0, \dots, 7$$

[0076]

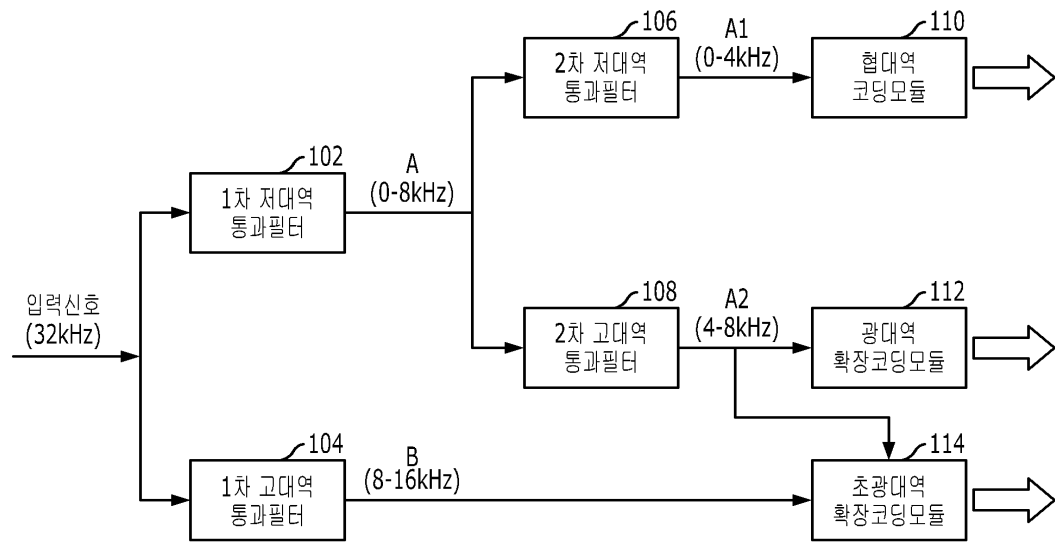
- [0077] 여기서, k 는 서브 대역 인덱스를 나타내고, $SbE(k)$ 는 k 번째 서브 대역의 에너지를 나타낸다. 또한 $\ddot{M}_{32}(k)$ 는 합성된 고 주파수 신호를 나타낸다. 각각의 서브 대역은 32개의 MDCT 계수들로 이루어진다. 상대적으로 큰 에너지를 갖는 서브 대역이 정현파 코딩의 탐색 트랙으로서 선택된다. 예를 들어, 탐색 트랙은 1의 단위 크기를 갖는 32개의 위치를 포함할 수 있다. 이러한 경우, 탐색 트랙은 서브 대역과 일치한다.
- [0078] 두 개의 정현파의 크기(amplitude)는 각기 4-bit, 1차원 코드북에 의하여 양자화된다.
- [0079] 정현파 모드(816)는 입력 신호가 토널일 때 이용된다. 정현파 모드(816)에서, 고 주파수 신호는 고 주파수 스펙트럼에 유한 개의 정현파 컴포넌트 세트를 추가함으로써 생성된다. 예를 들어, 추가되는 정현파의 총 개수는 10개인데, 4개는 7000-8600Hz 주파수 범위에, 4개는 8600-10200Hz 주파수 범위에, 1개는 10200-11800Hz 주파수 범위에, 1개는 11800-12600Hz 주파수 범위에 위치할 수 있다.
- [0080] 정현파 코딩부(818, 820)는 제네릭 모드(814) 또는 정현파 모드(816)에 의해 출력된 신호의 품질을 개선한다. 정현파 코딩부(818, 820)에 의해 추가되는 정현파의 수(N_{sin})는 비트 버킷에 따라 달라진다. 정현파 코딩부(818, 820)의 정현파 코딩을 위한 트랙들은 합성된 고 주파수 콘텐츠의 서브 대역 에너지에 기반하여 선택된다.
- [0081] 예를 들어, 7000-13400Hz 주파수 범위의 합성된 고 주파수 콘텐츠는 8개의 서브 대역으로 나누어진다. 각각의 서브 대역들은 32개의 MDCT 계수들로 구성되고, 서브 대역 에너지들은 각각 수학식 1과 같이 연산될 수 있다.
- [0082] 정현파 코딩을 위한 트랙들은 상대적으로 큰 에너지를 갖는 N_{sin}/N_{sin_track} 개의 서브 대역을 찾음으로써 선택된다. 여기서 N_{sin_track} 은 트랙 당 정현파의 개수이며 2로 설정된다. 선택된 N_{sin}/N_{sin_track} 개의 서브 대역들은 각각 정현파 코딩에 이용되는 트랙에 대응한다. 예를 들어, N_{sin} 이 4라면 처음 2개의 정현파가 가장 큰 서브 대역 에너지를 갖는 서브 대역에 위치하고, 남은 2개의 정현파는 두 번째로 큰 에너지를 갖는 서브 대역에 위치한다. 정현파 코딩을 위한 트랙 위치들은 이용 가능한 비트 버킷 및 고 주파수 신호 에너지 특성들에 따라 프레임 마다(frame by frame) 달라진다.
- [0083] 도 9는 본 발명의 다른 실시예에 의한 오디오 신호 디코딩 장치의 구성도이다.
- [0084] 도 9에 나타난 오디오 신호 디코딩 장치는 인코딩 장치에 의해 인코딩된 광대역 신호 및 초광대역 신호를 입력 받고, 이를 32kHz 신호로 출력한다. 이 오디오 신호 디코딩 장치는 광대역 확장디코딩 모듈(902, 914, 916, 918)과 초광대역 확장디코딩 모듈(904, 920, 922)로 구성된다. 광대역 확장디코딩 모듈은 입력된 16kHz 신호를 디코딩하며, 초광대역 확장디코딩 모듈은 32kHz 출력을 제공하기 위하여 고 주파수들을 디코딩한다. 초광대역 확장 디코딩은 대부분 MDCT 도메인에서 수행된다. 두 개의 모드, 즉 제네릭 모드(906) 및 정현파 모드(908)가 확장의 첫 번째 계층을 디코딩하기 위하여 이용되는데, 이는 처음으로 디코딩되는 토널리티 지시자(indicator)에 의존한다. 두 번째 계층은 광대역 신호 개선 및 추가적인 정현파 사이에 비트들을 분산시키기 위하여 인코더와 똑같은 비트 할당을 이용한다. 세 번째 초광대역 계층은 정현파 디코딩부(910, 912)로 구성되는데, 이는 고 주파수 콘텐츠의 품질을 개선한다. 네 번째 및 다섯 번째 확장 계층들은 광대역 신호 개선을 제공한다. 합성된 초광대역 콘텐츠를 개선하기 위하여 시간 도메인에서 전처리(post-processing)가 이용된다.
- [0085] 인코딩 장치에 의해 인코딩된 신호는 G.729.1 코덱(902)에 입력된다. G.729.1 코덱(902)는 16kHz의 합성 신호를 출력하고, 이는 광대역 신호 개선부(914)로 입력된다. 광대역 신호 개선부(914)는 입력된 신호의 품질을 개선한다. 광대역 신호 개선부(914)에서 출력된 신호는 전처리부(916)에 의한 전처리, 업 샘플링부(918)에 의한 업 샘플링을 거친다.
- [0086] 한편, 고 주파수 디코딩을 시작하기 이전에, 광대역 신호가 합성될 필요가 있다. 이러한 합성은 G.729.1 코덱(902)에 의하여 수행된다. 고 주파수 신호 디코딩에서는 일반적인 전처리 함수를 적용하기 이전에 32kbit/s 광대역 합성이 이용된다.
- [0087] 고 주파수 신호의 디코딩은 G.729.1 광대역 디코딩으로부터 합성된 MDCT 도메인 표현을 획득함으로써 시작된다. MDCT 도메인 광대역 콘텐츠는 제네릭 코딩 프레임의 고 주파수 신호를 디코딩하기 위하여 요구되는데, 여기서 고 주파수 신호는 광대역 주파수 범위로부터의 코딩된 서브 대역의 적응적 응답(adaptive replication)을 통해 구성된다.
- [0088] 제네릭 모드(906)는 적응적 서브 대역 응답에 의해 고 주파수 신호를 구성한다. 또한, 두 개의 정현파 컴포넌트

들이 첫 번째 4kbit/s 초광대역 확장 계층의 스펙트럼에 추가된다. 제네릭 모드(906)와 정현파 모드(808)는 정현파 모드 디코딩 기술에 기반한 유사한 향상 계층(enhancement layers)을 활용한다.

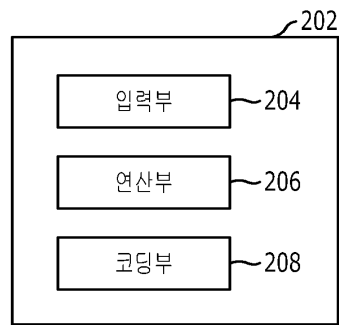
- [0089] 제네릭 모드(906)에서는 본 발명에 의한 오디오 디코딩 방법에 의하여, 디코딩된 신호의 품질 개선이 이루어질 수 있다. 제네릭 모드(906)는 두 개의 정현파 컴포넌트들을 재구성된 전체 고 주파수 스펙트럼에 추가한다. 이 정현파들은 위치, 부호 및 크기로 표현된다. 이 때, 정현파들을 추가하기 위한 트랙의 시작 위치는 위에서 언급한 바와 같이 상대적으로 큰 에너지를 갖는 서브 대역의 인덱스로부터 획득된다.
- [0090] 정현파 모드(908)에서, 고 주파수 신호는 유한 개의 정현파 컴포넌트 세트에 의해 생성된다. 예를 들어, 추가되는 정현파의 총 개수는 10개인데, 4개는 7000-8600Hz 주파수 범위에, 4개는 8600-10200Hz 주파수 범위에, 1개는 10200-11800Hz 주파수 범위에, 1개는 11800-12600Hz 주파수 범위에 위치할 수 있다.
- [0091] 정현파 디코딩부(910, 912)는 제네릭 모드(906) 또는 정현파 모드(908)에 의해 출력된 신호의 품질을 개선한다. 첫 번째 초광대역 향상 계층은 10개의 정현파 컴포넌트들을 정현파 모드 프레임의 고 주파수 신호 스펙트럼에 더 추가한다. 제네릭 모드 프레임에서, 추가되는 정현파 컴포넌트들의 수는 저 주파수 및 고 주파수 개선 사이의 적응적 비트 할당에 따라 설정된다.
- [0092] 정현파 디코딩부(910, 912)의 디코딩 과정은 다음과 같다. 먼저, 비트스트림으로부터 정현파의 위치가 획득된다. 그리고 나서 비트스트림은 전송된 부호 인덱스들 및 크기 코드북 인덱스들을 구하기 위해 디코딩된다.
- [0093] 정현파 디코딩을 위한 트랙들은 상대적으로 큰 에너지를 갖는 N_{sin}/N_{sin_track} 개의 서브 대역을 찾음으로써 선택된다. 여기서 N_{sin_track} 은 트랙 당 정현파의 개수이며 2로 설정된다. 선택된 N_{sin}/N_{sin_track} 개의 서브 대역들은 각각 정현파 디코딩에 이용되는 트랙에 대응한다.
- [0094] 각각의 대응하는 트랙들과 관계있는 10개의 정현파들의 위치 인덱스들은 비트스트림으로부터 처음으로 구해진다. 그리고 나서 10개의 정현파들의 부호들이 디코딩된다. 마지막으로, 정현파들의 크기(3개의 8비트 코드북 인덱스들)가 디코딩된다.
- [0095] 이렇게 정현파 디코딩부(910, 912)에 의해 품질이 개선된 신호들은 IMDCT(920)에 의한 역 MDCT, 전처리부(922)에 의한 전처리를 거친다. 업 샘플링부(918)의 출력 신호 및 전처리부(922)의 출력 신호는 더해져서 32kHz 출력 신호로 출력된다.
- [0096] 전술한 본 발명은, 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에게 있어 본 발명의 기술적 사상을 벗어나지 않는 범위 내에서 여러 가지 치환, 변형 및 변경이 가능하므로 전술한 실시예 및 첨부된 도면에 의해 한정되는 것이 아니다.

도면

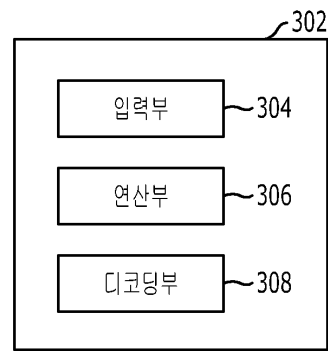
도면1



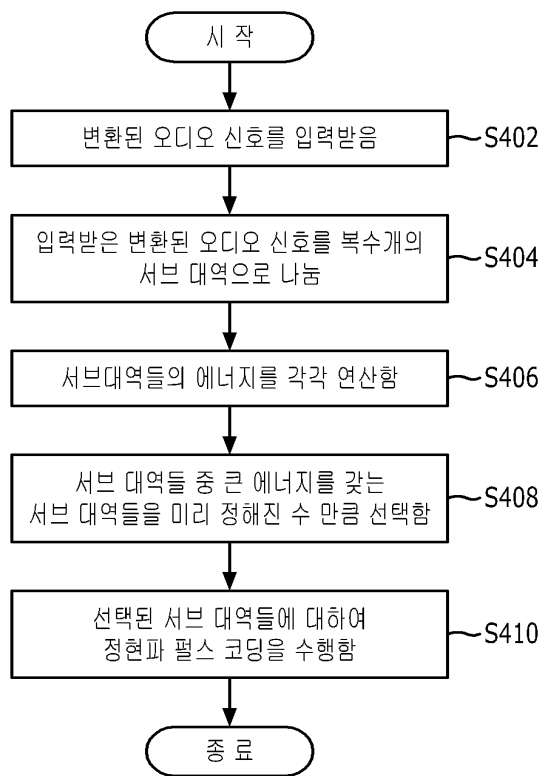
도면2



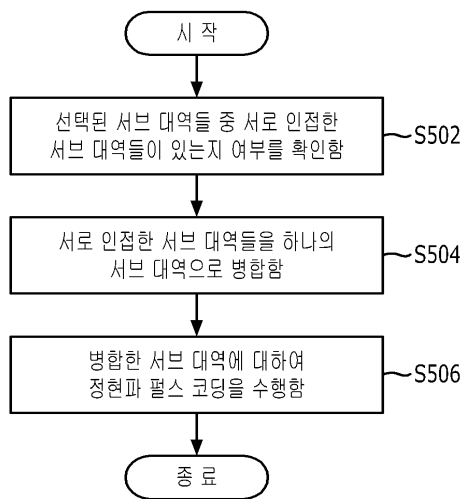
도면3



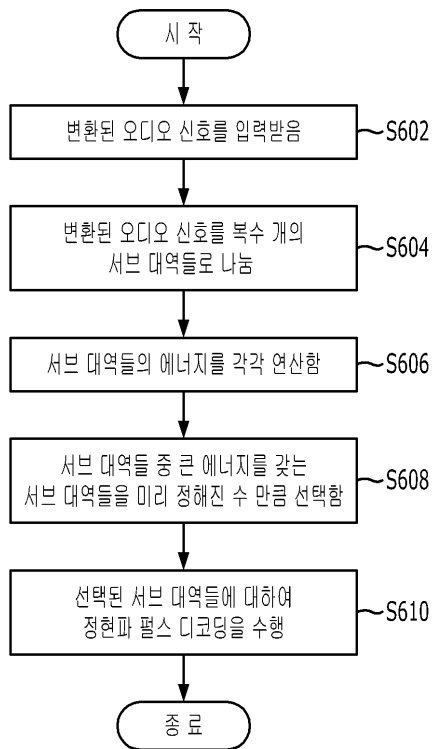
도면4



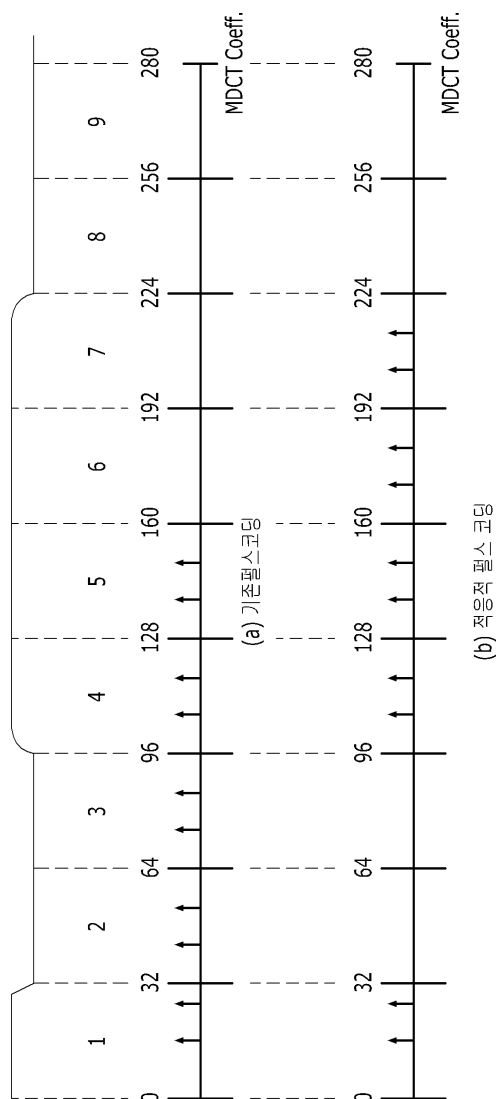
도면5



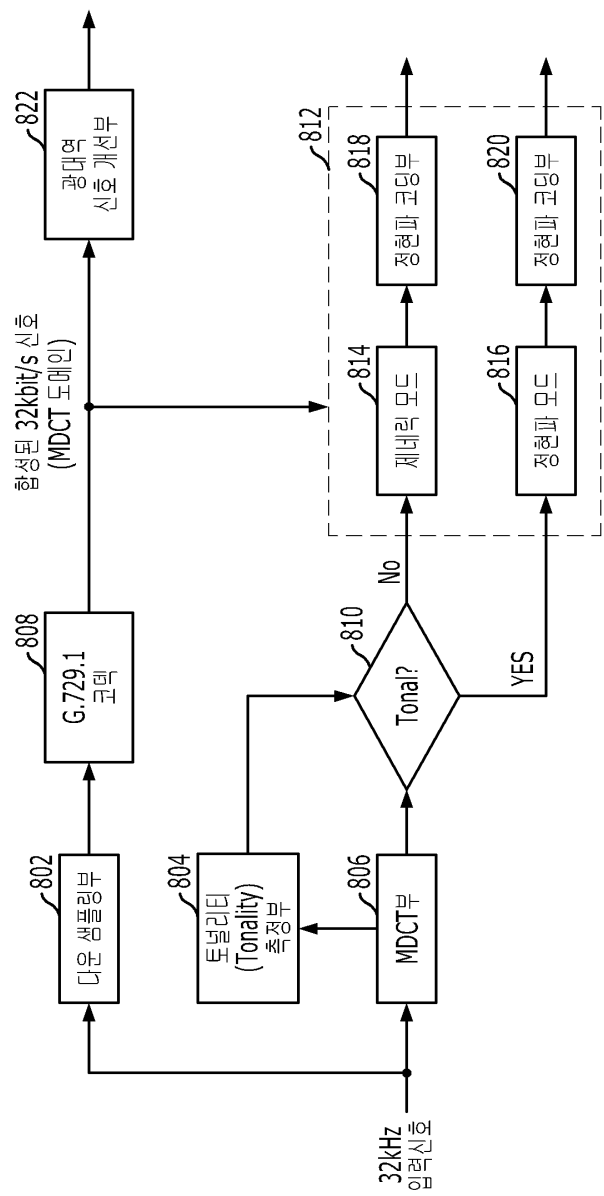
도면6



도면7



도면8



도면9

