



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년11월30일
(11) 등록번호 10-1924192
(24) 등록일자 2018년11월26일

(51) 국제특허분류(Int. Cl.)
G10L 19/00 (2006.01) G10L 19/02 (2006.01)
G10L 19/24 (2013.01) H03M 7/30 (2006.01)
(21) 출원번호 10-2010-0046928
(22) 출원일자 2010년05월19일
심사청구일자 2015년05월18일
(65) 공개번호 10-2010-0124678
(43) 공개일자 2010년11월29일
(30) 우선권주장
1020090043475 2009년05월19일 대한민국(KR)
1020090092701 2009년09월29일 대한민국(KR)
(56) 선행기술조사문헌
KR1020080086762 A*
KR1020090008611 A*
*는 심사관에 의하여 인용된 문헌
기술이전 회망 : 기술양도, 실시권허여, 기술지도

(73) 특허권자
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
이미숙
대전광역시 서구 대덕대로 415, 109동 1103호 (만년동, 상아아파트)
양희식
경상남도 김해시 구산로10번길 1 (구산동)
(74) 대리인
(뒷면에 계속)
특허법인 무한

전체 청구항 수 : 총 8 항

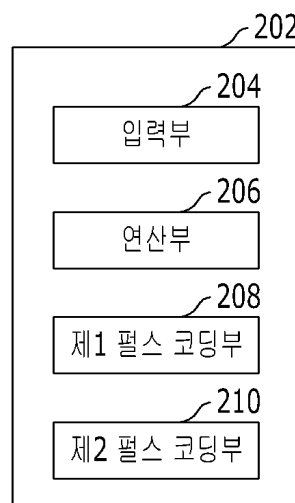
심사관 : 옥윤철

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

(57) 요약

본 발명은 오디오 신호의 인코딩 및 디코딩 방법 및 장치에 관한 것이다. 본 발명의 일 실시예에 의한 오디오 신호의 인코딩 방법은, 변환된 오디오 신호를 입력받는 단계, 변환된 오디오 신호를 복수 개의 서브 대역으로 나누는 단계, 복수 개의 서브 대역에 대하여 제1 정현파 코딩을 수행하는 단계, 제1 정현파 코딩의 코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 코딩의 수행 영역을 결정하는 단계 및 수행 영역에 대하여 제2 정현파 코딩을 수행하는 단계를 포함하고, 제1 정현파 코딩 수행 단계는 상기 코딩 정보에 따라 가변적으로 수행되는 것을 특징으로 한다. 본 발명에 의하면 계층형 정현파 코딩을 이용하여 상위 계층에서 오디오 신호를 인코딩 또는 디코딩할 때, 하위 계층의 정현파 코딩을 고려함으로써 합성 신호의 품질을 더욱 향상시킬 수 있는 효과가 있다.

대표도 - 도2



(72) 발명자

김현우

대전광역시 서구 둔산북로 175, 햇님아파트 1동
901호 (둔산동)

성종모

대전광역시 유성구 지족로 317, 반석마을1단지
108-702 (지족동)

배현주

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

이병선

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

이 발명을 지원한 국가연구개발사업

과제고유번호 2008-S-011-02

부처명 지식경제부

연구관리전문기관

연구사업명 IT성장동력기술개발

연구과제명 FMC 어커스틱 융합코덱 및 제어기술 연구(표준화연계)

주관기관 한국전자통신연구원

연구기간 2008.03.01 ~ 2012.02.29

명세서

청구범위

청구항 1

오디오 신호를 부호화하는 방법에 있어서,
 변환된 오디오 신호를 입력받는 단계;
 상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하는 단계;
 제1 계층에서 상기 서브 대역들에 대해 제1정현파 코딩(Sinusoidal Coding)을 수행하는 단계;
 상기 제1정현파 코딩의 코딩 정보를 바탕으로, 상기 서브 대역들 중 제2정현파 코딩을 위한 트랙들(tracks)을 결정하는 단계; 및
 상기 제1 계층과 다른 제2 계층에서 상기 트랙들에 대한 상기 제2정현파 코딩을 수행하는 단계;를 포함하며;
 상기 트랙들의 시작 지점들(starting positions)은, 상기 코딩 정보를 바탕으로 결정되고,
 상기 제1 정현파 코딩의 코딩 정보는, 제1 정현파 코딩에 할당된 비트 수 정보, 펄스 개수 정보, 펄스의 위치, 크기, 부호 정보 중 적어도 하나를 포함하는 부호화 방법.

청구항 2

삭제

청구항 3

제1항에 있어서,
 상기 트랙들의 시작 지점들은,
 상기 코딩 정보가 특정 값보다 작은 경우, 상기 복수 개의 서브 대역의 하위 대역에 위치하고;
 상기 코딩 정보가 특정 값보다 크거나 같은 경우, 상기 복수 개의 서브 대역의 상위 대역에 위치하는 것을 특징으로 하는 부호화 방법.

청구항 4

오디오 신호를 부호화하는 장치에 있어서,
 변환된 오디오 신호를 입력받는 입력부;
 상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하는 연산부;
 제1 계층에서 상기 서브 대역들에 대해 제1정현파 코딩(Sinusoidal Coding)을 수행하는 제1정현파 코딩부; 및
 상기 제1정현파 코딩의 코딩 정보를 바탕으로, 상기 서브 대역들 중 제2정현파 코딩을 위한 트랙들(tracks)을 결정하고, 상기 제1 계층과 다른 제2 계층에서 상기 트랙들에 대한 상기 제2정현파 코딩을 수행하는 제2정현파 코딩부;를 포함하며;
 상기 트랙들의 시작 지점들(starting positions)은, 상기 코딩 정보를 바탕으로 결정되고,
 상기 제1 정현파 코딩의 코딩 정보는, 제1 정현파 코딩에 할당된 비트 수 정보, 펄스 개수 정보, 펄스의 위치, 크기, 부호 정보 중 적어도 하나를 포함하는 부호화 장치.

청구항 5

삭제

청구항 6

제4항에 있어서,

상기 트랙들의 시작 지점들은,

상기 코딩 정보가 특정 값보다 작은 경우, 상기 복수 개의 서브 대역의 하위 대역에 위치하고;

상기 코딩 정보가 특정 값보다 크거나 같은 경우, 상기 복수 개의 서브 대역의 상위 대역에 위치하는 것을 특징으로 하는 부호화 장치.

청구항 7

오디오 신호를 복호화하는 방법에 있어서,

변환된 오디오 신호를 입력받는 단계;

상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하는 단계;

제1 계층에서 상기 서브 대역들에 대해 제1정현파 디코딩(Sinusoidal Decoding)을 수행하는 단계;

상기 제1정현파 디코딩의 디코딩 정보를 바탕으로, 상기 서브 대역들 중 제2정현파 디코딩을 위한 트랙들(tracks)을 결정하는 단계; 및

상기 제1 계층과 다른 제2 계층에서 상기 트랙들에 대한 상기 제2정현파 디코딩을 수행하는 단계;를 포함하며;

상기 트랙들의 시작 지점들(starting positions)은, 상기 디코딩 정보를 바탕으로 결정되고,

상기 디코딩 정보는,

상기 제1 정현파 디코딩에 할당된 비트 수 정보 또는 상기 제1정현파 디코딩에 할당된 펄스 개수 정보를 포함하는 것을 특징으로 하는 복호화 방법.

청구항 8

삭제

청구항 9

제7항에 있어서,

상기 트랙들의 시작 지점들은,

상기 디코딩 정보가 특정 값보다 작은 경우, 상기 복수 개의 서브 대역의 하위 대역에 위치하고;

상기 디코딩 정보가 특정 값보다 크거나 같은 경우, 상기 복수 개의 서브 대역의 상위 대역에 위치하는 것을 특징으로 하는 복호화 방법.

청구항 10

오디오 신호를 복호화하는 장치에 있어서,

변환된 오디오 신호를 입력받는 입력부;

상기 변환된 오디오 신호를 복수의 서브 대역들로 분할하는 연산부;

제1 계층에서 상기 서브 대역들에 대해 제1정현파 디코딩(Sinusoidal Decoding)을 수행하는 제1정현파

디코딩부; 및

상기 제1정현과 디코딩의 디코딩 정보를 바탕으로, 상기 서브 대역들 중 제2정현과 디코딩을 위한 트랙들(tracks)을 결정하고, 상기 제1 계층과 다른 제2 계층에서 상기 트랙들에 대한 상기 제2정현과 디코딩을 수행하는 제2정현과 디코딩부;를 포함하며;

상기 트랙들의 시작 지점들(starting positions)은, 상기 디코딩 정보를 바탕으로 결정되고,

상기 디코딩 정보는,

상기 제1 정현과 디코딩에 할당된 비트 수 정보 또는 상기 제1정현과 디코딩에 할당된 펄스 개수 정보를 포함하는 것을 특징으로 하는 복호화 장치.

청구항 11

삭제

청구항 12

제10항에 있어서,

상기 트랙들의 시작 지점들은,

상기 디코딩 정보가 특정 값보다 작은 경우, 상기 복수 개의 서브 대역의 하위 대역에 위치하고;

상기 디코딩 정보가 특정 값보다 크거나 같은 경우, 상기 복수 개의 서브 대역의 상위 대역에 위치하는 것을 특징으로 하는 복호화 장치.

발명의 설명

기술 분야

[0001] 본 발명은 오디오 신호의 인코딩 및 디코딩 방법 및 장치에 관한 것으로, 보다 상세하게는 계층형 정현과 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 코딩을 이용하기도 한다. 정현과 코딩은 여러 계층에 걸쳐 이루어질 수 있다. 만약 하위 계층에서 정현과 코딩에 할당되는 비트 또는 펄스 수가 프레임 단

위로 가변적인 경우, 상위 계층에서의 정현파 코딩에서 합성 신호의 품질을 높일 수 있는 방법이 요구된다.

발명의 내용

해결하려는 과제

[0007] 본 발명은 계층형 정현파 코딩을 이용하여 상위 계층에서 오디오 신호를 인코딩 또는 디코딩할 때, 하위 계층의 정현파 코딩을 고려함으로써 합성 신호의 품질을 더욱 향상시킬 수 있는 오디오 신호의 인코딩 및 디코딩 방법 및 장치를 제공하는 것을 목적으로 한다.

[0008] 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

[0009] 이러한 목적을 달성하기 위한 본 발명은 오디오 신호의 인코딩 방법에 있어서, 변환된 오디오 신호를 입력받는 단계, 변환된 오디오 신호를 복수 개의 서브 대역으로 나누는 단계, 복수 개의 서브 대역에 대하여 제1 정현파 코딩을 수행하는 단계, 제1 정현파 코딩의 코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 코딩의 수행 영역을 결정하는 단계 및 수행 영역에 대하여 제2 정현파 코딩을 수행하는 단계를 포함하고, 제1 정현파 코딩 수행 단계는 상기 코딩 정보에 따라 가변적으로 수행되는 것을 일 특징으로 한다.

[0010] 또한 본 발명은 오디오 신호의 인코딩 장치에 있어서, 변환된 오디오 신호를 입력받는 입력부, 변환된 오디오 신호를 복수 개의 서브 대역으로 나누는 연산부, 복수 개의 서브 대역에 대하여 제1 정현파 코딩을 수행하는 제1 정현파 코딩부 및 제1 정현파 코딩의 코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 코딩의 수행 영역을 결정하고, 수행 영역에 대하여 제2 정현파 코딩을 수행하는 제2 정현파 코딩부를 포함하고, 제1 정현파 코딩부는 코딩 정보에 따라 가변적으로 제1 정현파 코딩을 수행하는 것을 다른 특징으로 한다.

[0011] 또한 본 발명은 오디오 신호의 디코딩 방법에 있어서, 변환된 오디오 신호를 입력받는 단계, 변환된 오디오 신호를 복수 개의 서브 대역으로 나누는 단계, 복수 개의 서브 대역에 대하여 제1 정현파 디코딩을 수행하는 단계, 제1 정현파 디코딩의 코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 디코딩의 수행 영역을 결정하는 단계 및 수행 영역에 대하여 제2 정현파 디코딩을 수행하는 단계를 포함하고, 제1 정현파 디코딩 수행 단계는 디코딩 정보에 따라 가변적으로 수행되는 것을 또 다른 특징으로 한다.

[0012] 또한 본 발명은 오디오 신호의 디코딩 장치에 있어서, 변환된 오디오 신호를 입력받는 입력부, 변환된 오디오 신호를 복수 개의 서브 대역으로 나누는 연산부, 복수 개의 서브 대역에 대하여 제1 정현파 디코딩을 수행하는 제1 정현파 디코딩부 및 제1 정현파 디코딩의 디코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 디코딩의 수행 영역을 결정하고, 수행 영역에 대하여 제2 정현파 디코딩을 수행하는 제2 정현파 디코딩부를 포함하고, 제1 정현파 디코딩부는 디코딩 정보에 따라 가변적으로 제1 정현파 디코딩을 수행하는 것을 또 다른 특징으로 한다.

발명의 효과

[0013] 진술한 바와 같은 본 발명에 의하면, 계층형 정현파 코딩을 이용하여 상위 계층에서 오디오 신호를 인코딩 또는 디코딩할 때, 하위 계층의 정현파 코딩을 고려함으로써 합성 신호의 품질을 더욱 향상시킬 수 있는 장점이 있다.

도면의 간단한 설명

[0014] 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 두 개의 계층을 통해 7-14kHz에 해당하는 211개의 MDCT 계수에 정현파 코딩을 적용한 결과.

도 5는 본 발명의 일 실시예에 의한 계층형 정현파 코딩의 결과.

도 6은 본 발명의 다른 실시예에 의한 계층형 정현파 코딩의 결과.

도 7은 본 발명의 또 다른 실시예에 의한 계층형 정현파 코딩의 결과.

도 8은 기존의 정현파 코딩 방법과 본 발명에 의한 정현파 코딩 방법에 의해 합성된 MDCT 계수를 각각 나타내는 그래프.

도 9는 본 발명의 일 실시예에 의한 오디오 신호의 인코딩 방법을 설명하기 위한 흐름도.

도 10은 본 발명의 일 실시예에 의한 오디오 신호의 디코딩 방법을 설명하기 위한 흐름도.

도 11은 본 발명의 다른 실시예에 의한 오디오 신호 인코딩 장치의 구성도.

도 12는 본 발명의 다른 실시예에 의한 오디오 신호 디코딩 장치의 구성도.

발명을 실시하기 위한 구체적인 내용

[0015] 전술한 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.

[0016] 도 1은 협대역 코덱과의 호환성을 제공하는 초광대역 확장코덱의 구조를 나타낸다.

[0017] 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.

[0018] 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.

[0019] 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.

[0020] 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.

[0021] 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.

[0022] 이와 같은 확장코덱에서는 도 1과 같이 주파수 대역별로 다른 코딩방식을 적용할 수 있다. 예를 들어, G.729.1

과 G.711.1코덱은 협대역 신호를 기존 협대역 코덱인 G. 729 와 G. 711로 코딩하고, 나머지 신호에 대해서는 MDCT(Modified Discrete Cosine Transform)를 수행하여, 출력된 MDCT 계수를 코딩하는 방식을 사용한다.

- [0023] 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 코딩을 이용하여 합성된 신호의 품질을 향상시키는 구조가 그것이다.
- [0024] 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 코딩 등을 적용한다. 이러한 구조를 통해, 비트율의 증가에 따라 합성된 신호의 품질을 개선시킬 수 있다.
- [0025] 일반적으로 정현파 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 코딩을 적용하는 것보다는 서브 프레임 또는 서브 대역별로 정현파 코딩을 적용하는 것이 바람직하다. 정현파 코딩은 하나의 펄스를 전송하는데 상대적으로 많은 비트가 필요하지만 신호의 품질에 영향을 주는 신호를 정확히 표현할 수 있다는 장점을 가진다.
- [0026] 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.
- [0027] 서브 대역별로 정현파 코딩을 적용할 때, 계층형 정현파 코딩이 이용될 수 있다. 계층형 정현파 코딩이란 여러 계층에 걸쳐 정현파 코딩을 수행하는 것을 의미한다. 예를 들어, 첫 번째 계층에서는 전체 서브 대역 중 제1 영역에 대하여 정현파 코딩을 수행하고, 두 번째 계층에서는 전체 서브 대역 중 제2 영역에 대하여 정현파 코딩을 수행하는 것이다. 이러한 계층형 정현파 코딩을 수행함에 있어서, 앞서 언급한 바와 같은 신호의 주파수 대역 또는 에너지 등을 고려함으로써 오디오 신호의 품질을 보다 향상시키는 것이 가능하다.
- [0028] 본 발명은 도 1과 같은 확장코덱에서 계층형 정현파 코딩을 수행할 때, 이전 계층의 코딩 정보를 이용하여 다음 계층의 정현파 코딩을 수행함으로써, 합성된 신호의 품질을 보다 향상시킬 수 있는 오디오 신호의 인코딩 및 디코딩에 관한 것이다. 이하에서는 음성 및 오디오 신호를 오디오 신호로 지칭하여 본 발명을 설명한다.
- [0029] 도 2는 본 발명의 일 실시예에 의한 오디오 신호 인코딩 장치의 구성도이다.
- [0030] 도 2에 나타난 바와 같이, 오디오 신호 인코딩 장치(202)는 입력부(204), 연산부(206), 제1 정현파 코딩부(208), 제2 정현파 코딩부(210)를 포함한다.
- [0031] 입력부(204)는 변환된 오디오 신호, 예를 들면 오디오 신호가 MDCT에 의해 변환된 결과물인 MDCT 계수를 입력받는다.
- [0032] 연산부(206)는 입력부(204)를 통해 입력된 변환된 오디오 신호를 복수 개의 서브 대역으로 나눈다.
- [0033] 제1 정현파 코딩부(208)는 연산부(206)에 의해 나누어진 복수 개의 서브 대역에 대하여 제1 정현파 코딩을 수행한다. 제1 정현파 코딩부(208)는 코딩 정보에 따라 가변적으로 제1 정현파 코딩을 수행한다. 여기서 코딩 정보는 제1 정현파 코딩에 할당되는 비트 수 정보 또는 제1 정현파 코딩에 할당되는 펄스의 개수 정보일 수 있다. 또한, 제1 정현파 코딩을 '가변적'으로 수행한다는 것은 코딩 정보에 따라 비트 수 또는 펄스 개수를 달리하여 코딩하는 것, 또는 주파수 대역 순서가 아닌 각 서브 대역의 에너지 순으로 제1 정현파 코딩을 수행하는 것을 의미한다.
- [0034] 제2 정현파 코딩부(210)는 제1 정현파 코딩의 코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 코딩을 수행할 영역을 결정한다. 본 발명의 일 실시예에서, 제2 정현파 코딩부(210)는 코딩 정보가 특정 값보다 작은 경우, 복수 개의 서브 대역의 하위 대역을 수행 영역으로 결정하고, 코딩 정보가 특정 값보다 크거나 같은 경우, 복수 개의 서브 대역의 상위 대역을 수행 영역으로 결정할 수 있다. 본 발명의 다른 실시예에서, 제2 정현파 코딩부(210)는 제1 정현파 코딩이 적용되지 않은 가장 낮은 주파수 대역부터 제2 정현파 코딩을 적용할 수

있다. 그리고 제2 정현과 코딩부(210)는 결정된 수행 영역에 대하여 제2 정현과 코딩을 수행한다.

[0035] 도 3은 본 발명의 일 실시예에 의한 오디오 신호 디코딩 장치의 구성도이다.

[0036] 도 2에 나타난 바와 같이, 오디오 신호 디코딩 장치(302)는 입력부(304), 연산부(306), 제1 정현과 디코딩부(308), 제2 정현과 디코딩부(310)를 포함한다.

[0037] 입력부(304)는 변환된 오디오 신호, 예를 들면 오디오 신호가 MDCT에 의해 변환된 결과물인 MDCT 계수를 입력받는다.

[0038] 연산부(306)는 입력부(304)를 통해 입력된 변환된 오디오 신호를 복수 개의 서브 대역으로 나눈다.

[0039] 제1 정현과 디코딩부(308)는 연산부(306)에 의해 나누어진 복수 개의 서브 대역에 대하여 제1 정현과 디코딩을 수행한다. 제1 정현과 디코딩부(308)는 디코딩 정보에 따라 가변적으로 제1 정현과 코딩을 수행한다. 여기서 디코딩 정보는 제1 정현과 디코딩에 할당되는 비트 수 정보 또는 제1 정현과 디코딩에 할당되는 펄스의 개수 정보일 수 있다. 또한, 제1 정현과 디코딩을 '가변적'으로 수행한다는 것은 디코딩 정보에 따라 비트 수 또는 펄스 개수를 달리하여 디코딩하는 것, 또는 주파수 대역 순서가 아닌 각 서브 대역의 에너지 순으로 제1 정현과 디코딩을 수행하는 것을 의미한다.

[0040] 제2 정현과 디코딩부(310)는 제1 정현과 디코딩의 디코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현과 디코딩을 수행할 영역을 결정한다. 본 발명의 일 실시예에서, 제2 정현과 디코딩부(310)는 디코딩 정보가 특정 값보다 작은 경우, 복수 개의 서브 대역의 하위 대역을 수행 영역으로 결정하고, 코딩 정보가 특정 값보다 크거나 같은 경우, 복수 개의 서브 대역의 상위 대역을 수행 영역으로 결정할 수 있다. 본 발명의 다른 실시예에서, 제2 정현과 디코딩부(310)는 제1 정현과 디코딩이 적용되지 않은 가장 낮은 주파수 대역부터 제2 정현과 디코딩을 적용할 수 있다. 그리고 제2 정현과 디코딩부(310)는 결정된 수행 영역에 대하여 제2 정현과 디코딩을 수행한다.

[0041] 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.

[0042] 이하에서는 도 1 내지 도 8을 통해 본 발명에 의한 오디오 신호 인코딩 및 디코딩 방법의 일 실시예를 설명한다.

[0043] 초광대역 확장코딩 모듈(114)은 7-14kHz에 해당하는 MDCT 계수를 여러 개의 서브 대역으로 나누고, 각 서브 대역의 계인과 세이프를 코딩 또는 디코딩하여 오차 신호를 구한다. 그리고 나서 초광대역 확장코딩 모듈(114)은 오차 신호에 대하여 정현과 코딩 또는 디코딩을 수행한다. 이 때 정현과 코딩은 4kbit/s 또는 8kbit/s 단위로 비트율 조정이 가능한 계층형 구조라고 가정한다.

[0044] 초광대역 확장코딩 모듈(114)은 고대역(7-14kHz) 신호를 MDCT 영역으로 변환하고, 계층형 정현과 코딩을 통해 MDCT 계수를 코딩한다. 즉 고대역의 MDCT 계수를 복수 개의 서브 대역으로 나누고, 하나의 서브 대역 당 2개의 펄스를 코딩한다. 이 때 첫 번째 계층에서는 프레임에 따라 최대 10개의 펄스를 코딩할 수 있고, 두 번째 계층에서는 고정적으로 10개의 펄스를 코딩할 수 있다고 가정한다. 다시 말해, 첫 번째 계층에서는 펄스의 개수가 프레임에 따라 0에서 10까지 가변적이다. 하나의 서브 대역의 넓이는 0.8kHz(=32샘플)이며, 서브 대역의 시작점이 정해지면 그로부터 32개의 샘플이 하나의 서브 대역이 된다.

[0045] 도 4는 두 개의 계층을 통해 7-14kHz에 해당하는 211개의 MDCT 계수에 정현과 코딩을 적용한 결과를 나타낸다.

[0046] 도 4에서 N은 첫 번째 계층에서 정현과 코딩을 수행할 때 이용된 펄스의 개수를 나타낸다. 도 4를 참조하면, 첫 번째 계층에서는 정현과 코딩이 수행되지 않거나(N=0), 최대 10개의 펄스를 이용하여(N=10) 정현과 코딩이 수행될 수 있다. 하나의 서브 대역 당 2개의 펄스가 할당되므로, 이용되는 펄스의 수, 즉 N에 따라 정현과 코딩을 적용할 수 있는 서브 대역의 개수가 달라진다. 만일 N=2이면, 하나의 서브 대역에 대해서만 정현과 코딩이 적용되며, N=10인 경우 도 4와 같이 5개 서브 대역에 대하여 정현과 코딩이 적용된다.

[0047] 도 4에서, 두 번째 계층에서는 첫 번째 계층과는 독립적으로 항상 같은 서브 대역의 범위에 정현과 코딩이 적용된다. 즉, 첫 번째 계층의 정현과 코딩과는 무관하게 두 번째 계층에서는 항상 9.4kHz(=96번째 샘플)에서 정현과 코딩이 시작된다.

- [0048] 도 4와 같이 정현과 코딩을 수행할 때, 만약 첫 번째 계층에서 $N=6$ 일 경우, 두 번째 계층의 정현과 코딩을 수행하고 나면 7-13.4kHz의 대역에 빠짐 없이 정현과 코딩이 적용된다. 하지만 첫 번째 계층에서 $N=2$ 인 경우, 두 번째 계층의 정현과 코딩을 수행하고 나면 7.8-9.4kHz 대역에는 정현과 코딩이 적용되지 못하게 되며, 이는 합성된 신호의 품질 저하로 이어진다.
- [0049] 오디오 신호, 특히 음성 신호의 에너지 분포를 보면, 유성음의 에너지는 상대적으로 낮은 주파수 대역에 위치하며, 무성음과 파열음의 에너지는 상대적으로 높은 주파수 대역에 위치한다. 신호의 특성에 따라 다를 수 있으나, 대부분의 오디오 신호는 10kHz 이하에 많은 에너지를 갖는다. 즉, 도 4에 나타난 바와 같이 첫 번째 계층의 정현과 코딩과는 무관하게 두 번째 계층의 정현과 코딩이 수행될 경우, 일부 대역, 특히 음성 품질에 영향을 미치는 대역에 정현과 코딩이 적용되지 않는 경우가 발생하며, 이는 합성 신호의 품질 저하로 이어진다.
- [0050] 본 발명은 이와 같은 문제점을 극복하기 위해 첫 번째 계층의 정현과 코딩의 코딩 정보를 이용하여 두 번째 계층의 정현과 코딩을 수행함으로써 합성 신호의 품질을 향상시키는 오디오 신호의 인코딩 및 디코딩 방법을 제공한다.
- [0051] 도 5는 본 발명의 일 실시예에 의한 계층형 정현과 코딩의 결과를 나타낸다.
- [0052] 먼저 도 2의 입력부(204)는 MDCT 계수를 입력받는다. 그리고 연산부(206)는 입력받은 MDCT 계수를 도 5와 같이 복수 개의 서브 대역으로 나눈다. 이 때 하나의 서브 대역은 32개의 샘플을 갖는다.
- [0053] 제1 정현과 코딩부(208)는 첫 번째 계층의 정현과 코딩을 수행한다. 이 때 제1 정현과 코딩부(208)는 코딩 정보를 이용하여 가변적 정현과 코딩을 수행한다. 코딩 정보는 제1 정현과 코딩에 할당되는 비트 수 정보 또는 펄스의 개수 정보일 수 있다. 만약 제1 정현과 코딩을 위하여 4개의 정현과(또는 그에 대응하는 비트)가 할당되었다면, 제1 정현과 코딩부(208)는 이러한 정보를 이용하여 2개의 서브 대역에 대하여 제1 정현과 코딩을 수행한다. ($N=4$)
- [0054] 한편, 제2 정현과 코딩부(210)는 앞서 언급한 코딩 정보를 이용하여, 복수 개의 서브 대역 중 정현과 코딩을 수행할 영역을 결정한다. 제2 정현과 코딩부(210)는 제1 정현과 코딩부(208)로부터 제1 정현과 코딩에 할당된 비트 수 정보, 펄스 개수 정보, 펄스의 위치, 크기, 부호 정보 등이 포함된 코딩 정보를 전달받을 수 있다. 도 5를 참조하면, N 이 8보다 작은 경우 제2 정현과 코딩부(210)는 하위 대역(7-11kHz)에 대하여 제2 정현과 코딩을 수행하고, N 이 8보다 크거나 같은 경우 상위 대역(9.75-13.75kHz)에 대하여 제2 정현과 코딩을 수행한다.
- [0055] 이와 같은 계층형 정현과 코딩을 수행하면, 앞서 언급했던 기존 코딩의 문제점을 보완할 수 있다. 예를 들어 첫 번째 계층에서 $N=6$ 인 경우, 도 5에 의하면 두 번째 계층에서 하위 대역에 대하여 정현과 코딩을 수행하게 되므로 10kHz 이하에 대부분의 에너지를 갖고 있는 오디오 신호의 품질을 높일 수 있다.
- [0056] 도 6은 본 발명의 다른 실시예에 의한 계층형 정현과 코딩의 결과를 나타낸다.
- [0057] 본 실시예의 제2 정현과 코딩부(210)는 도 5를 통해 기술되었던 제2 정현과 코딩부(210)와 동일하게 제2 정현과 코딩을 수행한다. 다만, 본 실시예에서 제1 정현과 코딩부(208)는 주파수 대역 순서가 아닌 에너지가 많은 서브 대역 순으로 정현과 코딩을 '가변적으로' 수행한다.
- [0058] 도 7은 본 발명의 또 다른 실시예에 의한 계층형 정현과 코딩의 결과를 나타낸다.
- [0059] 본 실시예에서 제1 정현과 코딩부(208)는 도 4의 실시예에서와 마찬가지로 제1 정현과 코딩을 수행한다. 한편, 제2 정현과 코딩부(210)는 첫 번째 계층에서 제1 정현과 디코딩이 적용되지 않은 가장 낮은 주파수 대역에 대한 정보를 포함하는 코딩 정보를 이용하여, 제2 정현과 코딩을 수행한다. 예를 들어 도 7과 같이 $N=4$ 인 경우, 제2 정현과 코딩부(210)는 64번째 샘플에 해당하는 서브 대역부터 제2 정현과 코딩을 시작한다.
- [0060] 지금까지 설명한 본 발명의 일 실시예는 인코딩 뿐만 아니라 디코딩에도 유사하게 적용될 수 있다.
- [0061] 도 8은 기존의 정현과 코딩 방법과 본 발명에 의한 정현과 코딩 방법에 의해 합성된 MDCT 계수를 각각 나타내는 그래프이다.
- [0062] 도 8에서 파란색 선은 원래의 MDCT 계수를, 빨간색 선은 기존의 방법으로 인코딩 및 디코딩된 MDCT 계수를 나타낸다. 그리고 노란색 선은 본 발명에 의한 방법으로 인코딩 및 디코딩된 MDCT 계수를 나타낸다. 여기서 첫 번째 계층에서 $N=0$ 이고 두 번째 계층에서는 10개의 펄스가 코딩되었다. 따라서 본 발명에 의한 인코딩 및 디코딩에서 두 번째 계층에서는 7kHz부터 정현과 코딩 또는 디코딩이 시작된다. 도 8에 나타난 바와 같이 본 발명에 의한 인코딩 및 디코딩에서는 기존 방법과 비교할 때, 오디오 신호의 품질에 많은 영향을 미칠 수 있는 상대적으로

낮은 주파수 대역에서 큰 에너지를 갖는 신호를 잘 표현한다.

- [0063] 도 9는 본 발명의 일 실시예에 의한 오디오 신호의 인코딩 방법을 설명하기 위한 흐름도이다.
- [0064] 먼저 변환된 오디오 신호, 예를 들면 MDCT 계수를 입력받는다(902). 그리고 변환된 오디오 신호를 복수 개의 서브 대역으로 나눈다(904).
- [0065] 그 다음, 나뉜 복수 개의 서브 대역에 대하여 제1 정현파 코딩을 수행한다(906). 이 때 제1 정현파 코딩은 코딩 정보에 따라 가변적으로 제1 정현파 코딩을 수행한다. 여기서 코딩 정보는 제1 정현파 코딩에 할당되는 비트 수 정보 또는 제1 정현파 코딩에 할당되는 펄스의 개수 정보일 수 있다. 또한, 제1 정현파 코딩을 '가변적'으로 수행한다는 것은 코딩 정보에 따라 비트 수 또는 펄스 개수를 달리하여 코딩하는 것, 또는 주파수 대역 순서가 아닌 각 서브 대역의 에너지 순으로 제1 정현파 코딩을 수행하는 것을 의미한다.
- [0066] 그 다음, 제1 정현파 코딩의 코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 코딩을 수행할 영역을 결정한다(908). 이 때 코딩 정보가 특정 값보다 작은 경우, 복수 개의 서브 대역의 하위 대역을 수행 영역으로 결정하고, 코딩 정보가 특정 값보다 크거나 같은 경우, 복수 개의 서브 대역의 상위 대역을 수행 영역으로 결정할 수 있다. 또한 제1 정현파 코딩이 적용되지 않은 가장 낮은 주파수 대역부터 제2 정현파 코딩을 적용할 수도 있다. 그리고 나서, 결정된 수행 영역에 대하여 제2 정현파 코딩을 수행한다(910).
- [0067] 도 10은 본 발명의 일 실시예에 의한 오디오 신호의 디코딩 방법을 설명하기 위한 흐름도이다.
- [0068] 먼저 변환된 오디오 신호, 예를 들면 MDCT 계수를 입력받는다(1002). 그리고 변환된 오디오 신호를 복수 개의 서브 대역으로 나눈다(1004).
- [0069] 그 다음, 나뉜 복수 개의 서브 대역에 대하여 제1 정현파 디코딩을 수행한다(1006). 이 때 제1 정현파 디코딩은 디코딩 정보에 따라 가변적으로 제1 정현파 디코딩을 수행한다. 여기서 디코딩 정보는 제1 정현파 디코딩에 할당되는 비트 수 정보 또는 제1 정현파 디코딩에 할당되는 펄스의 개수 정보일 수 있다. 또한, 제1 정현파 디코딩을 '가변적'으로 수행한다는 것은 디코딩 정보에 따라 비트 수 또는 펄스 개수를 달리하여 디코딩하는 것, 또는 주파수 대역 순서가 아닌 각 서브 대역의 에너지 순으로 제1 정현파 디코딩을 수행하는 것을 의미한다.
- [0070] 그 다음, 제1 정현파 디코딩의 디코딩 정보를 이용하여, 복수 개의 서브 대역 중 제2 정현파 디코딩을 수행할 영역을 결정한다(1008). 이 때 디코딩 정보가 특정 값보다 작은 경우, 복수 개의 서브 대역의 하위 대역을 수행 영역으로 결정하고, 디코딩 정보가 특정 값보다 크거나 같은 경우, 복수 개의 서브 대역의 상위 대역을 수행 영역으로 결정할 수 있다. 또한 제1 정현파 디코딩이 적용되지 않은 가장 낮은 주파수 대역부터 제2 정현파 디코딩을 적용할 수도 있다. 그리고 나서, 결정된 수행 영역에 대하여 제2 정현파 디코딩을 수행한다(1010).
- [0071] 이하에서는 도 11 및 도 12를 통해 본 발명의 다른 실시예에 의한 오디오 신호 인코딩 및 디코딩 방법 및 장치에 대해 설명한다.
- [0072] 도 11은 본 발명의 다른 실시예에 의한 오디오 신호 인코딩 장치의 구성도이다.
- [0073] 도 11에 나타난 오디오 신호 인코딩 장치는 32kHz의 입력 신호를 입력받고, 광대역 신호 및 초광대역 신호를 합성하여 출력한다. 이 오디오 신호 인코딩 장치는 광대역 확장코딩 모듈(1102, 1108, 1122)과 초광대역 확장코딩 모듈(1104, 1106, 1110, 1112)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 체네틱 모드(1114)와 정현파 모드(1116)가 초광대역 확장코딩 모듈의 첫 번째 계층을 코딩하기 위하여 이용된다. 체네틱 모드(1114) 또는 정현파 모드(1116) 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(1118, 1120), 또는 광대역 콘텐츠의 인지 품질(perceptual quality)을 개하는데 이용되는 광대역 신호 개선부(11202)에 의하여 코딩된다.
- [0074] 32kHz의 입력 신호는 먼저 다운 샘플링부(1102)에 입력되고, 16kHz로 다운 샘플링된다. 그리고 다운 샘플링된 16kHz 신호는 G.729.1 코덱(1108)에 입력된다. G.729.1 코덱(1108)은 입력된 16kHz 신호에 대하여 광대역 코딩을 수행한다. G.729.1 코덱(1108)에서 출력된 합성된 32kbit/s 신호는 광대역 신호 개선부(1122)로 입력되고, 광대역 신호 개선부(1122)는 입력된 신호의 품질을 개선한다.
- [0075] 한편, 32kHz 입력 신호는 MDCT부(1106)에 입력되고 MDCT 도메인으로 변환된다. MDCT 도메인으로 변환된 입력 신

호는 토널리티 측정부(1104)에 입력되고 입력 신호의 토널(tonal) 여부가 결정된다(1110). 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 토널리티 측정에 기반하여 정의된다. 토널리티 측정은 입력 신호의 현재 프레임과 과거 프레임의 스펙트럴 피크(spectral peaks) 간의 상관관계 분석(correlation analysis)에 기반한다.

[0076] 그 다음, 토널리티 측정부(1104)에 의해 출력된 토널리티 정보에 의해 입력 신호가 토널인지 아닌지 여부가 결정된다(1110). 예를 들어, 토널리티 정보가 특정 임계값(threshold)보다 크면 입력 신호는 토널인 것으로, 그렇지 않으면 입력 신호는 토널이 아닌 것으로 판단된다. 토널리티 정보는 또한 디코더로 전달되는 비트스트림에도 포함된다. 만약 입력 신호가 토널이면 정현파 모드(1116)가, 그렇지 않으면 제네릭 모드(1114)가 이용된다.

[0077] 제네릭 모드(1114)는 입력 신호의 프레임이 토널이 아닐 때(tonal=0) 이용된다. 제네릭 모드(1114)는 고 주파수들을 코딩하기 위하여 G.729.1 광대역 코덱(1108)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(1114) 및 정현파 코딩부(1118) 내의 추가적인 펄스들에 의해 개선된다.

[0078] 제네릭 모드(1114)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 펄스를 추가하도록 허용한다. 추가할 펄스의 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

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

[0079]

[0080] 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.

[0081] 두 개의 펄스의 크기(amplitude)는 각기 4-bit, 1차원 코드북에 의하여 양자화된다.

[0082] 정현파 모드(1116)는 입력 신호가 토널일 때 이용된다. 정현파 모드(1116)에서, 고 주파수 신호는 예를 들어, 추가되는 펄스의 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.

[0083] 정현파 코딩부(1118, 1120)는 제네릭 모드(1114) 또는 정현파 모드(1116)에 의해 출력된 신호의 품질을 개선한다. 정현파 코딩부(1118, 1120)에 의해 추가되는 펄스의 수(N_{sin})는 비트 버짓에 따라 달라진다. 정현파 코딩부(1118, 1120)의 정현파 코딩을 위한 트랙들은 합성된 고 주파수 콘텐츠의 서브 대역 에너지에 기반하여 선택된다.

[0084] 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 수학적 식 1과 같이 연산될 수 있다.

[0085] 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 펄스의 개수이며 2로 설정된다. 선택된 N_{sin}/N_{sin_track} 개의 서브 대역들

은 각각 정현파 코딩에 이용되는 트랙에 대응한다. 예를 들어, N_{sin} 이 4라면 처음 2개의 펄스가 가장 큰 서브 대역 에너지를 갖는 서브 대역에 위치하고, 남은 2개의 펄스는 두 번째로 큰 에너지를 갖는 서브 대역에 위치한다. 정현파 코딩을 위한 트랙 위치들은 이용 가능한 비트 버짓 및 고 주파수 신호 에너지 특성들에 따라 프레임마다(frame by frame) 달라진다.

[0086] 한편, 또 다른 20개의 펄스가 고 주파수 신호에 두 단계로 추가된다. 이 때 추가되는 펄스의 트랙 구조는 제네릭 모드와 정현파 모드 프레임 사이에서 상이하다.

[0087] 제네릭 모드 프레임에서, 정현파 코딩을 위한 트랙들의 시작 위치는 N_{sin} 에 의존한다. 만약 N_{sin} 이 특정 임계값(threshold)보다 낮으면, 펄스들은 고 주파수 신호의 주파수 영역의 하위 부분에 위치한다. 만약 N_{sin} 이 임계값보다 크거나 같으면, 대부분의 펄스들은 고 주파수 신호의 주파수 영역의 상위 부분에 위치한다. 본 실시예에서 임계값은 8로 정의된다.

[0088] 첫 번째 단계에서, 10개의 펄스가 고 주파수 스펙트럼에 다음과 같이 추가된다. 먼저, 6개의 펄스는 각각 2개의 펄스를 가지며 7000-9400Hz 또는 9750-12150Hz의 주파수 대역에 위치하는 3개의 트랙으로 그룹화된다. 다음 4개의 펄스는 각각 2개의 펄스를 가지며 9400-11000Hz 또는 12150-13750Hz의 주파수 대역에 위치하는 2개의 트랙으로 그룹화된다.

[0089] 두 번째 단계에서, 남은 10개의 펄스는 다음과 같이 추가된다. 먼저, 6개의 펄스는 각각 2개의 펄스를 가지며 7800-10200Hz, 9400-11800Hz 또는 8600-11000Hz의 주파수 대역에 위치하는 3개의 트랙으로 그룹화된다. 마지막 4개의 펄스는 각각 2개의 펄스를 가지며 10200-11800Hz, 11800-13400Hz 또는 11000-12600Hz의 주파수 대역에 위치하는 2개의 트랙으로 그룹화된다.

[0090] 표 1은 위에서 설명한 제네릭 모드에서의 정현파 트랙의 구조, 즉 정현파 트랙의 시작 위치, 구간 크기(step size), 트랙 길이를 나타낸다.

표 1

N_{sin}	첫 번째 시작 위치	두 번째 시작 위치	구간 크기	길이
0, 2	280	312	3	32
	376	408	2	32
4, 6	280	376	3	32
	376	472	2	32
8, 10	390	344	3	32
	486	440	2	32

[0092] 정현파 모드에서는, 첫 10개 펄스가 다음과 같이 추가된다. 먼저, 6개의 펄스는 각각 2개의 펄스를 가지며 7000Hz와 9400Hz 사이의 주파수 대역에 위치하는 3개의 트랙으로 그룹화된다. 다음 4개의 펄스는 각각 2개의 펄스를 가지며 11000Hz와 12600Hz 사이의 주파수 대역에 위치하는 2개의 트랙으로 그룹화된다.

[0093] 두 번째 10개의 펄스는 다음과 같이 추가된다. 먼저, 4개의 펄스는 각각 2개의 펄스를 가지며 9400Hz와 11000Hz 사이의 주파수 대역에 위치하는 2개의 트랙으로 그룹화된다. 다음 6개의 펄스는 각각 2개의 펄스를 가지며 11000Hz와 13400Hz 사이의 주파수 대역에 위치하는 3개의 트랙으로 그룹화된다.

[0094] 표 2는 위에서 설명한 정현파 모드에서의 첫 번째 10개의 펄스의 정현파 트랙의 구조, 즉 정현파 트랙의 시작 위치, 구간 크기, 트랙 길이를 나타낸다. 그리고 표 3은 위에서 설명한 정현파 모드에서의 두 번째 10개의 펄스의 정현파 트랙의 구조, 즉 정현파 트랙의 시작 위치, 구간 크기, 트랙 길이를 나타낸다.

표 2

트랙	펄스 개수	시작 위치	구간 크기	길이
0	2	280	3	32
1	2	281	3	32
2	2	282	3	32

3	2	440	2	32
4	2	441	2	32

표 3

트랙	펄스 개수	시작 위치	구간 크기	길이
0	2	376	2	32
1	2	377	2	32
2	2	440	3	32
3	2	441	3	32
4	2	442	3	32

도 12는 본 발명의 다른 실시예에 의한 오디오 신호 디코딩 장치의 구성도이다.

도 12에 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(1202, 1214, 1216, 1218)과 초광대역 확장디코딩 모듈(1204, 1220, 1222)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(1206) 및 정현파 모드(1208)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(1210, 1212)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 후처리(post-processing)가 이용된다.

인코딩 장치에 의해 인코딩된 신호는 G.729.1 코덱(1202)에 입력된다. G.729.1 코덱(1202)는 16kHz의 합성 신호를 출력하고, 이는 광대역 신호 개선부(1214)로 입력된다. 광대역 신호 개선부(1214)는 입력된 신호의 품질을 개선한다. 광대역 신호 개선부(1214)에서 출력된 신호는 후처리부(1216)에 의한 후처리, 업 샘플링부(1218)에 의한 업 샘플링을 거친다.

한편, 고 주파수 디코딩을 시작하기 이전에, 광대역 신호가 합성될 필요가 있다. 이러한 합성은 G.729.1 코덱(1202)에 의하여 수행된다. 고 주파수 신호 디코딩에서는 일반적인 후처리 함수를 적용하기 이전에 32kbit/s 광대역 합성이 이용된다.

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.

제네릭 모드(1206)는 적응적 서브 대역 응답에 의해 고 주파수 신호를 구성한다. 또한, 두 개의 정현파 컴포넌트들이 첫 번째 4kbit/s 초광대역 확장 계층의 스펙트럼에 추가된다. 제네릭 모드(1206)와 정현파 모드(1208)는 정현파 모드 디코딩 기술에 기반한 유사한 향상 계층(enhancement layers)을 활용한다.

제네릭 모드(1206)에서는 본 발명에 의한 오디오 디코딩 방법에 의하여, 디코딩된 신호의 품질 개선이 이루어질 수 있다. 제네릭 모드(1206)는 두 개의 정현파 컴포넌트들을 재구성된 전체 고 주파수 스펙트럼에 추가한다. 이 펄스들은 위치, 부호 및 크기로 표현된다. 이 때, 펄스들을 추가하기 위한 트랙의 시작 위치는 위에서 언급한 바와 같이 상대적으로 큰 에너지를 갖는 서브 대역의 인덱스로부터 획득된다.

정현파 모드(1208)에서, 고 주파수 신호는 유한 개의 정현파 컴포넌트 세트에 의해 생성된다. 예를 들어, 추가되는 펄스의 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.

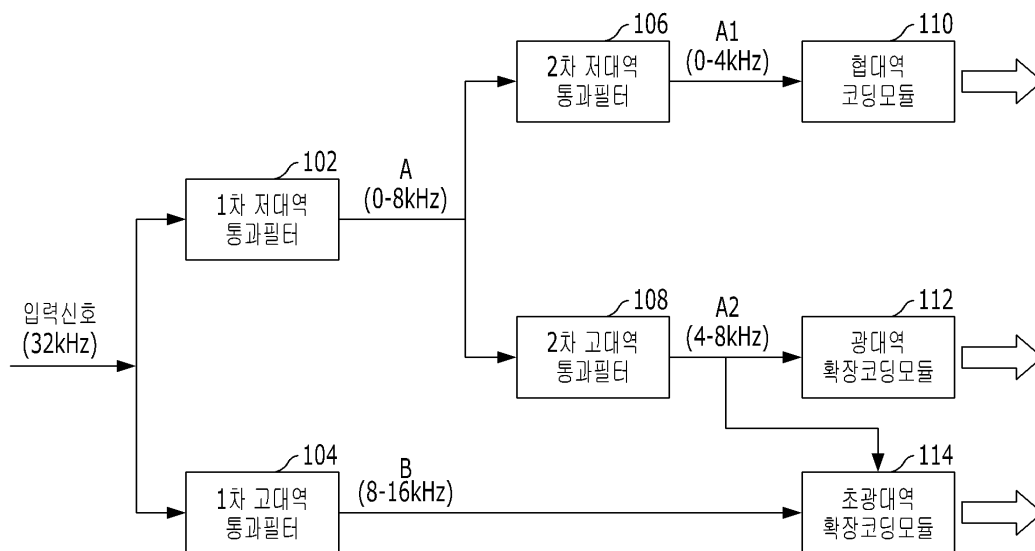
정현파 디코딩부(1210, 1212)는 제네릭 모드(1206) 또는 정현파 모드(1208)에 의해 출력된 신호의 품질을 개선한다. 첫 번째 초광대역 향상 계층은 10개의 정현파 컴포넌트들을 정현파 모드 프레임의 고 주파수 신호 스펙트럼에 더 추가한다. 제네릭 모드 프레임에서, 추가되는 정현파 컴포넌트들의 수는 저 주파수 및 고 파수 개선 사

이의 적응적 비트 할당에 따라 설정된다.

- [0106] 정현과 디코딩부(1210, 1212)의 디코딩 과정은 다음과 같다. 먼저, 비트스트림으로부터 펄스의 위치가 획득된다. 그리고 나서 비트스트림은 전송된 부호 인덱스들 및 크기 코드북 인덱스들을 구하기 위해 디코딩된다.
- [0107] 정현과 디코딩을 위한 트랙들은 상대적으로 큰 에너지를 갖는 N_{sin}/N_{sin_track} 개의 서브 대역을 찾음으로써 선택된다. 여기서 N_{sin_track} 은 트랙 당 펄스의 개수이며 2로 설정된다. 선택된 N_{sin}/N_{sin_track} 개의 서브 대역들은 각각 정현과 디코딩에 이용되는 트랙에 대응한다.
- [0108] 각각의 대응하는 트랙들과 관계있는 10개의 펄스들의 위치 인덱스들은 비트스트림으로부터 처음으로 구해진다. 그리고 나서 10개의 펄스들의 부호들이 디코딩된다. 마지막으로, 펄스들의 크기(3개의 8비트 코드북 인덱스들)가 디코딩된다.
- [0109] 한편, 디코딩 시에도 신호의 품질 향상을 위해 또 다른 20개의 펄스가 고 주파수 신호에 추가된다. 이 20개의 펄스 추가에 대해서는 위에서 자세히 설명한 바 있으므로 여기서는 그 설명을 생략한다.
- [0110] 이렇게 정현과 디코딩부(1210, 1212)에 의해 품질이 개선된 신호들은 IMDCT(1220)에 의한 역 MDCT, 후처리부(1222)에 의한 후처리를 거친다. 업 샘플링부(1218)의 출력 신호 및 후처리부(1222)의 출력 신호는 더해져서 32kHz 출력 신호로 출력된다.
- [0111] 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

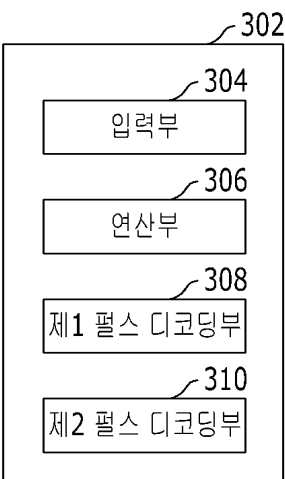
도면1



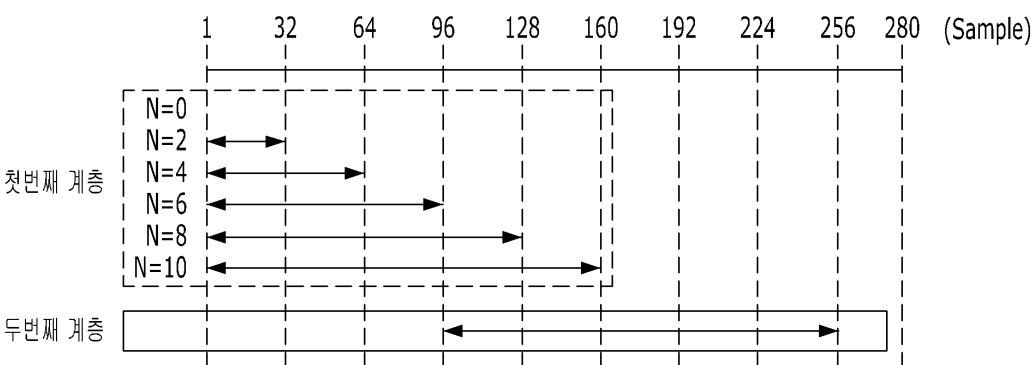
도면2



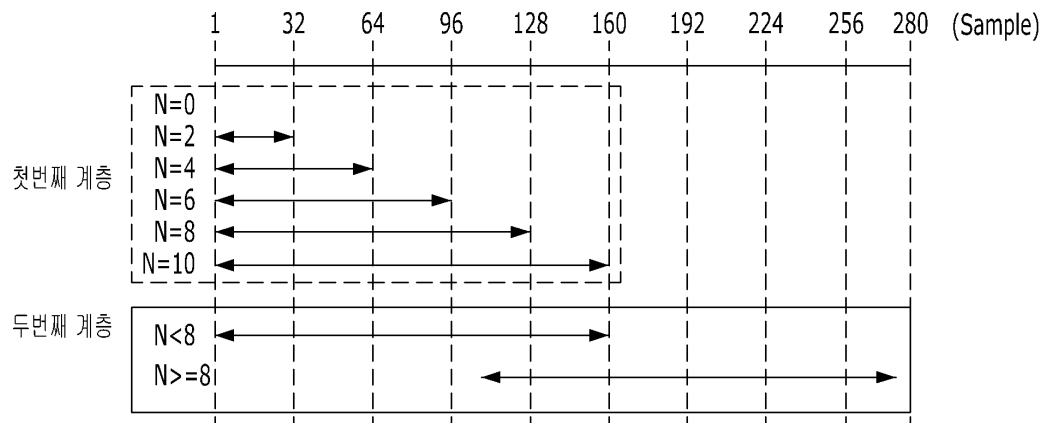
도면3



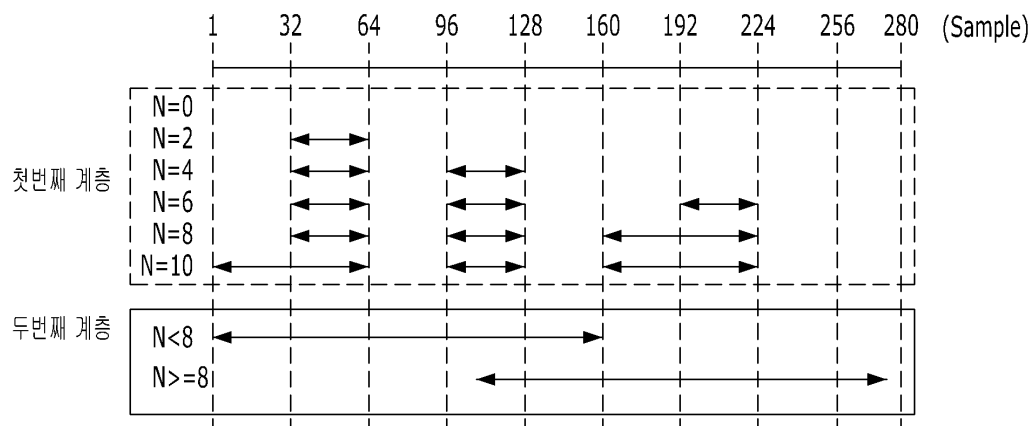
도면4



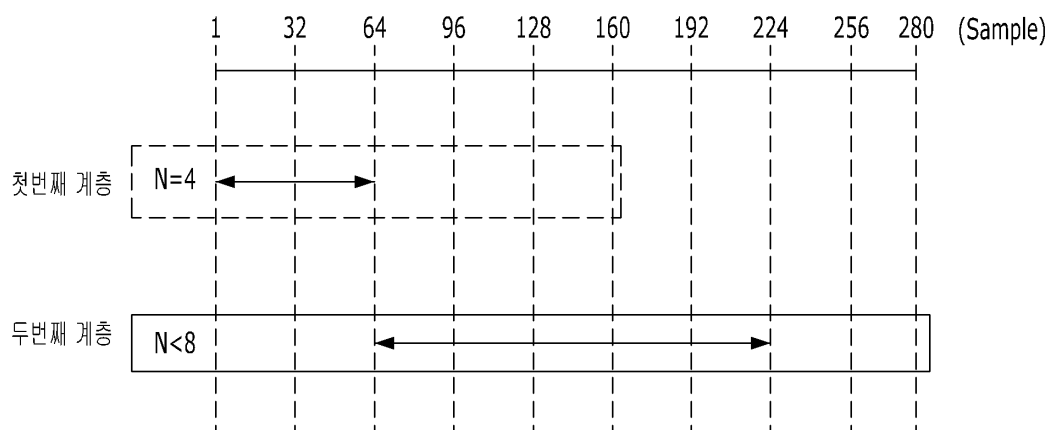
도면5



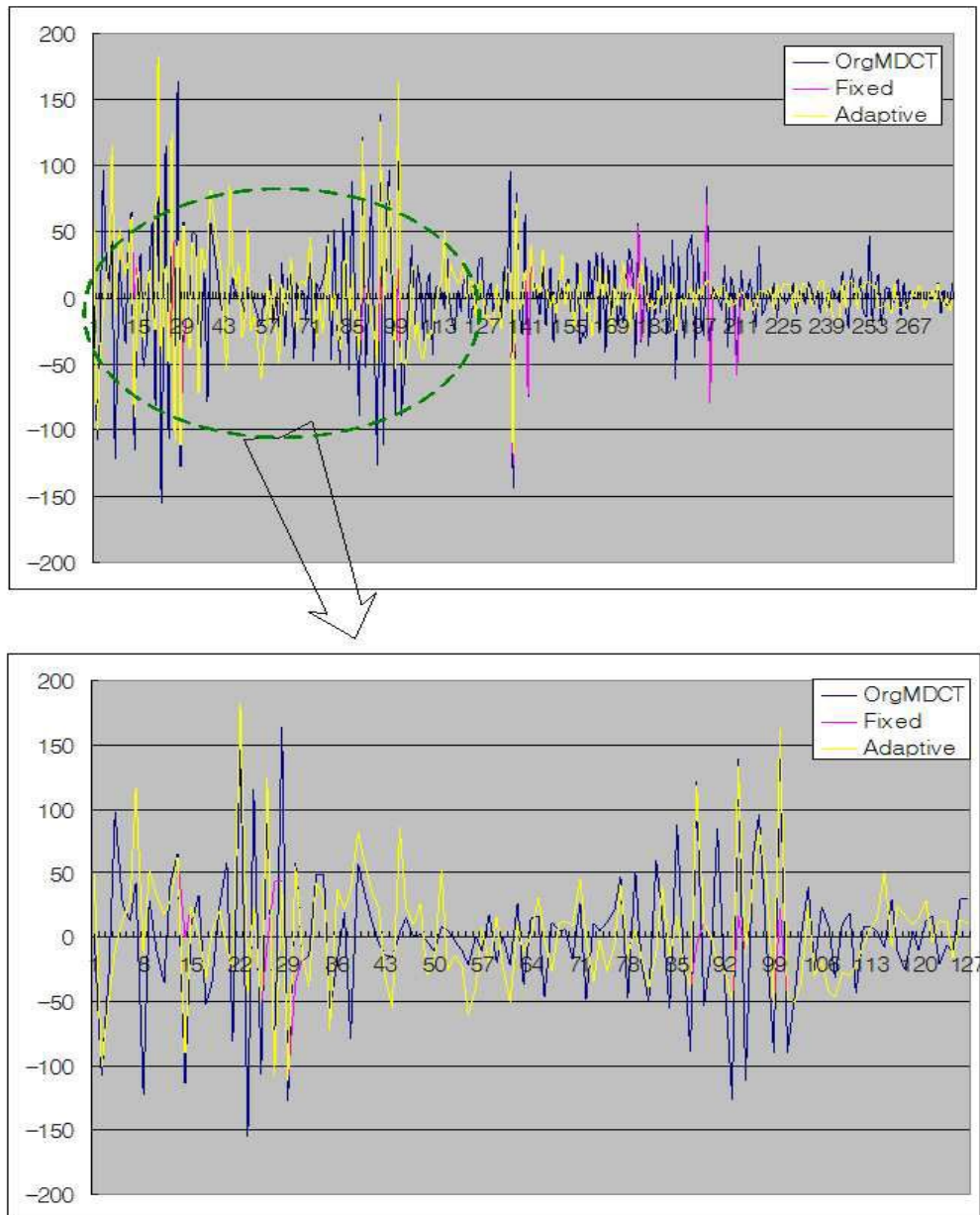
도면6



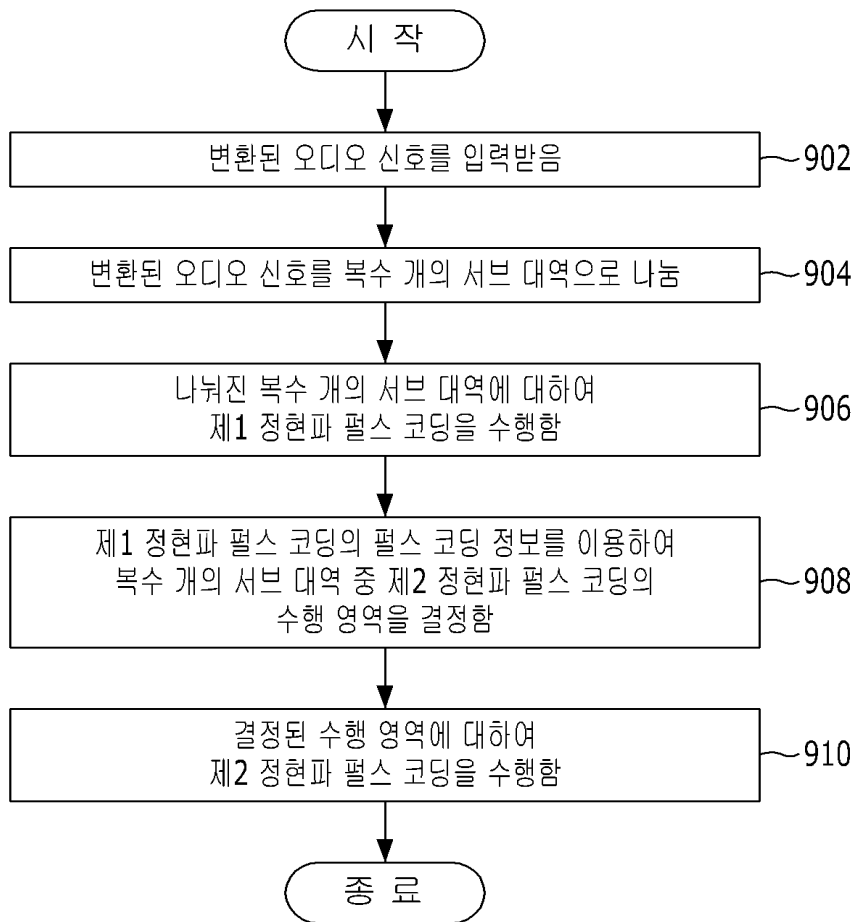
도면7



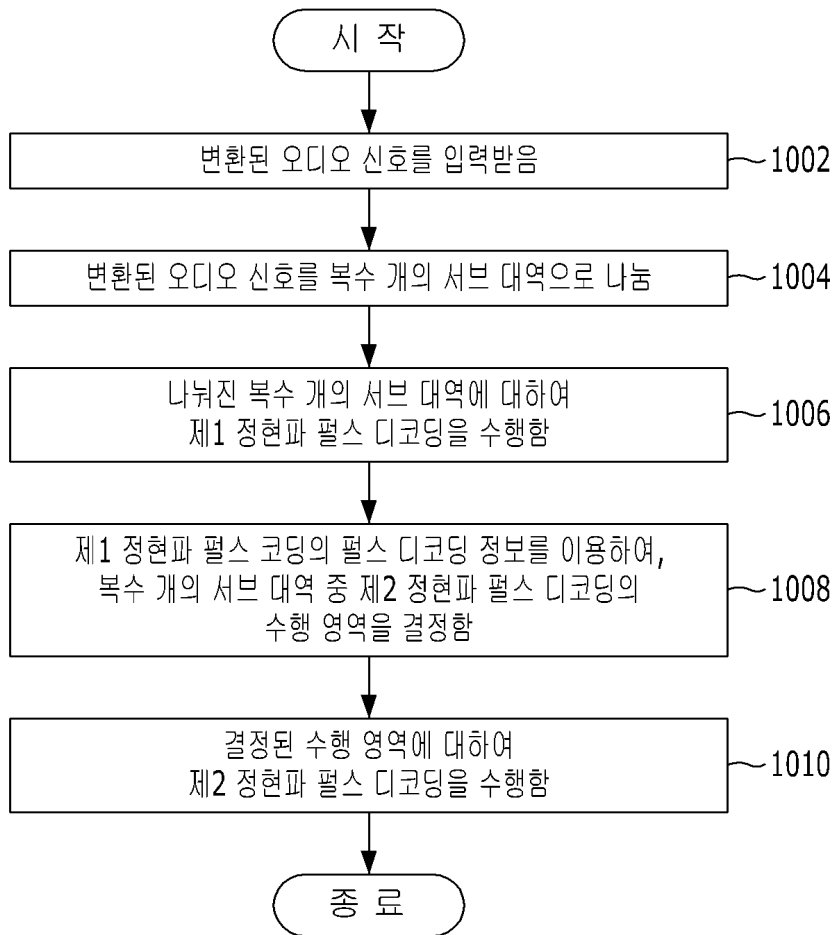
도면8



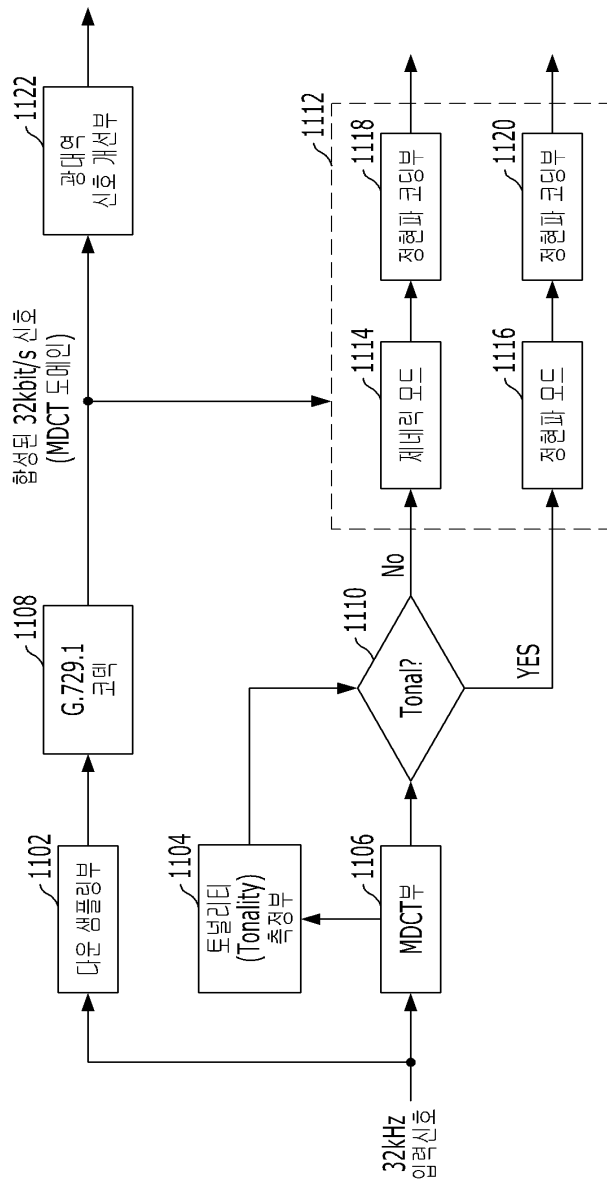
도면9



도면10



도면11



도면12

