



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년11월07일
(11) 등록번호 10-1794149
(24) 등록일자 2017년10월31일

(51) 국제특허분류(Int. Cl.)
G10L 19/028 (2013.01) G10L 19/12 (2006.01)
(52) CPC특허분류
G10L 19/028 (2013.01)
G10L 19/12 (2013.01)
(21) 출원번호 10-2015-7022400
(22) 출원일자(국제) 2014년01월28일
심사청구일자 2015년08월19일
(85) 번역문제출일자 2015년08월19일
(65) 공개번호 10-2015-0114966
(43) 공개일자 2015년10월13일
(86) 국제출원번호 PCT/EP2014/051649
(87) 국제공개번호 WO 2014/118192
국제공개일자 2014년08월07일
(30) 우선권주장
61/758,189 2013년01월29일 미국(US)
(56) 선행기술조사문헌
US06691085 B1*
Max Neuendorf, et al. MPEG unified speech and
audio coding-the ISO/MPEG standard for
high-efficiency audio coding of all content
types. Audio Engineering Society Convention
132. 2012.04.29.*
Universal Mobile Telecommunications System
(UMTS), Mandatory Speech Codec speech
processing functions AMR Wideband speech
codec, Transcoding functions. 3GPP TS 26.190
version 5.1.0 Release 5. 2001.12.*
ITU-T Rec. G.718. Frame error robust
narrow-band and wideband embedded variable
bit-rate coding of speech and audio from 8-32
kbit/s. ITU-T, 2008.06.*
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
프라운호퍼 게젤샤프트 쾰르 뢰르데룽 데어 안겐
반텐 포르슈 에. 베.
독일 80686 뮌헨 한자슈트라쎄 27 체
(72) 발명자
푸흐스, 구일라우메
독일 91058 에를랑겐 뢰르트 슈트라쎄 17
헨리히, 크리스티안
독일 91054 에를랑겐 하우프트슈트라쎄 68
(뒷면에 계속)
(74) 대리인
윤의섭, 김수진

전체 청구항 수 : 총 17 항

심사관 : 이남숙

(54) 발명의 명칭 CELP 같은 코더들을 위한 부가정보 없는 잡음 충전 장치 및 방법

(57) 요약

본 발명은 선형 예측 계수들(LPC)을 포함하는 인코딩된 오디오 정보에 기반한 디코딩 된 오디오 정보를 제공하는 오디오 디코더와 관련된다. 각각의 방법은 이러한 방법을 수행하는 컴퓨터 프로그램, 오디오 신호를 저장하는 저장 매체에 대한 오디오 신호, 이러한 오디오 신호를 취급하는 방법을 포함한다. 오디오 디코더는 기울기 정보를 (뒷면에 계속)

대표도



획득하기 위해 현재 프레임의 선형 예측 계수들을 이용하는 잡음의 기울기를 조정하도록 구성되는 기울기 조정기 그리고 기울기 연산기에 의해 획득된 기울기 정보에 의존하는 현재 프레임으로 잡음을 부가하도록 구성된 잡음 삽입기를 포함한다. 본 발명에 따른 다른 오디오 디코더는 잡음 레벨 정보를 획득하기 위한 적어도 하나의 이전 프레임의 선형 예측 계수를 이용하는 현재 프레임에 대한 잡음 레벨을 추정하도록 구성되는 잡음 레벨 추정기를 포함하고 잡음 레벨 추정기에 의해 제공된 잡음 레벨에 의존하는 현재 프레임으로 잡음을 부가하도록 구성되는 잡음 삽입기를 포함한다. 따라서, 비트 스트림에서 배경 잡음에 대한 부가 정보는 생략 될 수 있다.

(72) 발명자

젠더, 마뉴엘

독일 91052 에를랑겐 라이비그슈트라쎄 2

슈베르트, 벤자민

독일 90429 뉘른베르크 치크슈트라쎄 6

요코타니, 요시카즈

독일 91052 에를랑겐 루이트폴트슈트라쎄 85

명세서

청구범위

청구항 1

선형 예측 계수들(LPC, linear prediction coefficients)을 포함하는 인코딩된 오디오 정보에 기반하여 디코딩된 오디오 정보를 제공하는 오디오 디코더에 있어서,

배경 잡음의 기울기를 조정하도록 구성된 기울기 조정기 - 상기 기울기 조정기는 기울기 정보를 획득하기 위해 현재 프레임의 선형 예측 계수를 이용함 -;

잡음 레벨 추정기;

디코딩된 코어 코더 출력 신호를 획득하기 위해 현재 프레임의 선형 예측 계수들을 이용하여 현재 프레임의 오디오 정보를 디코딩하도록 구성되는 디코더 코어; 및

잡음 충전을 수행하기 위해, 현재 프레임으로 조정된 배경 잡음을 부가하도록 구성되는 잡음 삽입기;를 포함하며,

상기 기울기 조정기는 기울기 정보를 획득하기 위해 현재 프레임의 선형 예측 계수들의 제1차 분석 결과를 사용하도록 구성되며,

상기 기울기 조정기는 제1차 분석에 따라 현재 프레임의 선형 예측 계수들의 이득 g 의 연산을 이용하여 기울기 정보를 획득하도록 구성되며,

여기서 g 는 수식 $g = \sum(a_k \cdot a_{k+1}) / \sum(a_k \cdot a_k)$ 에 의해 주어지고,

상기 a_k 는 LPC 지수(index) k 에 위치한 현재 프레임의 선형 예측 계수인 것을 특징으로 하는 오디오 디코더.

청구항 2

제 1항에 있어서, 상기 오디오 디코더는 현재 프레임의 프레임 유형을 결정하는 프레임 유형 결정기를 포함하며, 상기 프레임 유형 결정기는 현재 프레임의 프레임 유형이 음성 유형으로 감지 될 때 배경 잡음의 기울기를 조정하는 기울기 조정기를 활성화 하도록 구성되는 것을 특징으로 하는 오디오 디코더.

청구항 3

삭제

청구항 4

삭제

청구항 5

제 1항에 있어서, 상기 오디오 디코더는 :

잡음 레벨 정보를 획득하기 위한 적어도 하나의 이전 프레임의 복수의 선형 예측 계수를 이용하여 현재 프레임에 대한 잡음 레벨을 추정하도록 구성된 잡음 레벨 추정기 - 상기 잡음 삽입기는 잡음 레벨 추정기에 의해 제공되는 잡음 레벨 정보에 의존하여 현재 프레임으로 배경 잡음을 부가하도록 구성됨 -;를 더 포함하고,

상기 오디오 디코더는 현재 프레임의 자극 신호를 디코딩하고 현재 프레임의 자극 신호의 제공근 e_{rms} 를 연산하도록 구성되며,

상기 오디오 디코더는 현재 프레임의 LPC 필터의 전달함수 최대 레벨(peak level) p 를 연산하도록 구성되고,

상기 오디오 디코더는 잡음 레벨 정보를 획득하기 위해 최대 레벨 p 및 제공근 e_{rms} 의 지수를 연산하는 것에 의해 현재 오디오 프레임의 스펙트럼 최소 m_f 를 연산하도록 구성되고,

상기 잡음 레벨 추정기는 둘 이상의 서로 다른 오디오 프레임에 기반한 잡음 레벨을 추정하도록 구성되는 것을 특징으로 하는 오디오 디코더.

청구항 6

선형 예측 계수(LPC)를 포함하는 인코딩된 오디오 정보에 기반한 디코딩된 오디오 정보를 제공하는 오디오 디코더에 있어서,

잡음 레벨 정보를 획득하기 위해 적어도 하나의 이전 프레임의 복수개의 선형 예측 계수를 이용하여 현재 프레임에 대한 잡음 레벨을 추정하도록 구성된 잡음 레벨 추정기; 및

상기 잡음 레벨 추정기에 의해 제공되는 잡음 레벨 정보에 의존하여 현재 프레임으로 잡음을 부가하도록 구성되는 잡음 삽입기;를 포함하고,

상기 오디오 디코더는 현재 프레임의 자극 신호를 디코딩하고 현재 프레임의 자극 신호의 제공근 e_{rms} 를 연산하도록 구성되며,

상기 오디오 디코더는 현재 프레임의 LPC 필터 전달 함수의 최대 레벨(peak level)을 연산하도록 구성되고,

상기 오디오 디코더는 잡음 레벨 정보를 획득하기 위해 최대 레벨 p 및 제공근 e_{rms} 의 지수를 연산함으로써 현재 오디오 프레임의 스펙트럼 최소 m_f 를 연산하도록 구성되며,

상기 잡음 레벨 추정기는 서로 다른 오디오 프레임의 둘 이상의 지수에 기반하여 잡음 레벨을 추정하도록 구성되고,

상기 오디오 디코더는 디코딩된 코어 코더 출력을 획득하기 위해 현재 프레임의 선형 예측 계수를 이용하여 현재 프레임의 오디오 정보를 디코딩하도록 구성된 디코더 코어를 포함하며, 상기 잡음 삽입기는 현재 프레임의 오디오 정보 디코딩에서 사용되고 하나 이상의 이전 프레임들의 오디오 정보 디코딩에서 사용되는 선형 예측 계수들에 의존하여 잡음을 부가하는 것을 특징으로 하는 오디오 디코더.

청구항 7

제 6항에 있어서, 상기 오디오 디코더는 현재 프레임의 프레임 유형을 결정하는 프레임 유형 결정기를 포함하고, 프레임 유형 결정기는 현재 프레임의 프레임 유형이 음성 또는 일반적인 오디오인 경우를 식별하기 위해 구성되고, 이에 따라 잡음 레벨 추정기는 현재 프레임의 프레임 유형에 의존하도록 수행될 수 있는 것을 특징으로 하는 오디오 디코더.

청구항 8

제 6항에 있어서, 상기 오디오 디코더는 현재 프레임이 음성 유형인 경우에서 잡음 레벨 정보를 획득하기 위해 현재 프레임의 시간 영역 표현으로부터 현재 프레임의 제공근 e_{rms} 를 연산하도록 구성되는 것을 특징으로 하는 오디오 디코더.

청구항 9

제 6항에 있어서, 상기 오디오 디코더는 만일 현재 프레임이 일반적인 오디오 유형인 경우 잡음 레벨 정보를 획득하기 위해 현재 프레임의 미형성 MDCT-자극을 디코딩하고 현재 프레임의 스펙트럼 영역 표현으로부터 제공근 e_{rms} 를 연산하도록 구성되는 것을 특징으로 하는 오디오 디코더.

청구항 10

제 6항에 있어서, 상기 오디오 디코더는 프레임 유형에 관계없이 잡음 레벨 추정기의 오디오 프레임으로부터 획득된 지수를 인큐잉(enqueue) 하도록 구성되며, 상기 잡음 레벨 추정기는 서로 다른 오디오 프레임의 둘 이상의 지수들에 대한 잡음 레벨 저장소를 포함하는 것을 특징으로 하는 오디오 디코더.

청구항 11

제 6항에 있어서, 상기 잡음 레벨 추정기는 서로 다른 오디오 프레임의 둘 이상의 지수의 통계적 분석에 기반하여 잡음 레벨을 추정하도록 구성되는 것을 특징으로 하는 오디오 디코더.

청구항 12

제 1항 또는 제6항 중 어느 한 항에 있어서, 상기 오디오 디코더는 현재 프레임을 강조 약화(de-emphasize)하기 위한 강조 약화 필터(de-emphasis filter)를 포함하고, 오디오 디코더는 잡음 삽입기가 현재 프레임으로 잡음을 부가한 이후 현재 프레임 상의 강조 약화 필터를 적용하도록 구성되는 것을 특징으로 하는 오디오 디코더.

청구항 13

제 1항 또는 제6항 중 어느 한 항에 있어서, 상기 오디오 디코더는 잡음 삽입기에 의해 현재 프레임으로 부가되는 잡음을 생성하도록 구성되는 잡음 생성기를 포함하는 것을 특징으로 하는 오디오 디코더.

청구항 14

제 1항 또는 제6항 중 어느 한 항에 있어서, 상기 오디오 디코더는 임의의 백색 잡음을 생성하도록 구성되는 잡음 생성기를 포함하는 것을 특징으로 하는 오디오 디코더.

청구항 15

제 1항 또는 제6항 중 어느 한 항에 있어서, 상기 오디오 디코더는

인코딩된 오디오 정보를 디코딩 하기 위해 하나 이상의 AMR-WB, G.718 또는 LD-USAC 또는 EVS 디코더에 기반한 디코더를 이용하도록 구성되는 것을 특징으로 하는 오디오 디코더.

청구항 16

선형 예측 계수들(LPC, linear prediction coefficients)을 포함하는 인코딩 된 오디오 정보에 기반한 디코딩 된 오디오 정보를 제공하는 방법에 있어서,

잡음 레벨을 추정하는 단계;

배경 잡음의 기울기를 조정하는 단계 - 현재 프레임의 선형 예측 계수들은 기울기 정보를 획득하기 위해 사용됨 -;

디코딩된 코어 코더 출력 신호를 획득하기 위해 현재 프레임의 선형 예측 계수를 이용하여 현재 프레임의 오디오 정보를 디코딩하는 단계; 및

잡음 충전을 수행하기 위해, 조정된 배경 잡음을 현재 프레임으로 부가하는 단계;를 포함하며,

기울기 정보를 획득하기 위해 현재 프레임의 선형 예측 계수들의 제1차 분석 결과가 사용되며,
제1차 분석에 따라 현재 프레임의 선형 예측 계수들의 이득 g 의 연산을 이용하여 기울기 정보가 획득되며,
여기서 g 는 수식 $g = \sum(a_k \cdot a_{k+1}) / \sum(a_k \cdot a_k)$ 에 의해 주어지고,
상기 a_k 는 LPC 지수(index) k 에 위치한 현재 프레임의 선형 예측 계수인 것을 특징으로 하는 방법.

청구항 17

컴퓨터 프로그램이 컴퓨터 상에서 구동될 때, 제 16항에 따른 방법을 수행하기 위한 컴퓨터 프로그램을 저장한 컴퓨터 판독가능 저장 매체.

청구항 18

선형 예측 계수들(LPC, linear prediction coefficients)을 포함하는 인코딩된 오디오 정보에 기반하여 디코딩된 오디오 정보를 제공하는 방법에 있어서,

잡음 레벨 정보를 획득하기 위해 적어도 하나의 이전 프레임의 복수의 선형 예측 계수들을 이용하여 현재 프레임에 대한 잡음 레벨을 추정하는 단계; 및

잡음 레벨 추정에 의해 제공되는 잡음 레벨 정보에 의존하여 현재 프레임으로 잡음을 부가하는 단계;를 포함하고,

디코딩된 코어 코더 출력 신호를 획득하기 위해 현재 프레임의 선형 예측 계수들을 이용하여 현재 프레임의 오디오 정보를 디코딩하는 단계; 및

현재 프레임의 오디오 정보 디코딩에서 사용되고 하나 이상의 이전 프레임의 오디오 정보 디코딩에서 사용되는 선형 예측 계수들에 의존하여 잡음을 부가하는 단계;를 포함하며,

상기 현재 프레임의 자극 신호가 디코딩되고 상기 현재 프레임의 자극 신호의 제곱근 e_{rms} 가 연산되며,

상기 현재 프레임의 LPC 필터의 전달 함수의 최대 레벨(peak level) p 가 연산되고,

현재 오디오 프레임의 스펙트럼 최소 m_l 는 잡음 레벨 정보를 획득하기 위해 최대 레벨 p 및 제곱근 e_{rms} 의 지수를 연산함으로써 계산되고,

상기 잡음 레벨은 서로 다른 오디오 프레임의 둘 이상의 지수들을 기반으로 추정되는 것을 특징으로 하는 방법.

청구항 19

컴퓨터 프로그램이 컴퓨터 상에서 구동될 때, 제 18항에 따른 방법을 수행하기 위한 컴퓨터 프로그램을 저장한 컴퓨터 판독가능 저장 매체.

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

발명의 설명

기술 분야

[0001] 본 발명의 실시 예는 선형 예측 계수(LPC, linear prediction coefficients)를 포함하는 인코딩 된 오디오 정보에 기초하여 디코딩된 오디오 정보를 제공하기 위한 오디오 디코더를 참조하고, 선형 예측 계수를 포함하는 인코딩 된 오디오 정보에 기초하여 디코딩된 오디오 정보를 제공하기 위한 방법을 참조한다. 이러한 방법을 수행하기 위한 컴퓨터 프로그램을 참조하고 상기 컴퓨터 프로그램은 컴퓨터에서 구동한다. 아울러 오디오 신호 또는 오디오 신호와 같은 신호를 저장하는 저장매체, 이러한 방법을 다루는 오디오 신호 등도 컴퓨터에서 구동한다.

배경 기술

[0002] 로우-비트-레이트 디지털 스피치 코더들(Low-bit-rate digital speech coders)은 코드 활성화 선형 예측(CELP, code-excited linear prediction)에 기반하고 코딩 원리는 일반적으로 비트-레이트가 샘플당 0.5에서 1로 떨어질 때 신호 회소성 인공물로부터 어려움을 겪고, 이는 다소 인공적인, 금속 소리를 유도한다. 특히 입력 음성이 배경에서의 환경 잡음을 가질 때, 로우-레이트 인공물들(low-rate artifacts)은 명확하게 들린다 : 구체적으로 배경 잡음은 활성 음성 섹션 동안 감쇄될 것이다. 본 발명은 AMR-WB [1] 및 G.718 [4, 7]과 같은 (A)CELP 코더들의 잡음 삽입 계획을 설명한다. 이는 잡음 충전 기술과 유사하게 xHE-AAC [5, 6]와 같은 코더들에 기반한 변형에서 사용되고, 주변 잡음을 재건하기 위한 음성 신호를 생성하는 임의의 노이즈 생성기(random noise generator)의 출력을 부가한다.

[0003] 국제 공개 WO 2012/110476 A1는 스펙트럼 도메인 잡음 형성에 기반 및 사용되는 선형 예측인 인코딩 개념을 보여준다. 스펙트럼 시퀀스를 포함하는 분광스펙트럼(spectrogram)으로의 오디오 입력 신호 스펙트럼 분해는 선형 예측 계수 연산 뿐만 아니라 선형예측 계수에 기반한 주파수-영역 형성을 위한 입력에 사용된다. 인용된 문서에 따라 오디오 인코더는 이로부터 선형 예측 계수들을 유도하기 때문에 입력 오디오를 분석하기 위한 선형 예측 분석기를 포함한다. 오디오 인코더의 주파수-영역 형성기는 선형 예측 분석기에 의해 제공되는 선형 예측 계수를 기초로 분광 스펙트럼의 스펙트럼 시퀀스의 현재 스펙트럼을 스펙트럼적으로 형성하도록 구성된다. 양자화된 그리고 스펙트럼적으로 형성된 스펙트럼은 스펙트럼 형성에서 사용되는 선형 예측 계수들 상의 정보와 함께 데이터 스트림에 삽입된다. 이에 따라 디코딩 측면에서, 탈-형성 및 역-양자화가 수행될 수 있다. 일시적인 잡음 형성 모듈은 또한 일시적인 잡음 형성을 위해 존재할 수 있다.

[0004] 종래 기술의 관점에서 방법 및 개선된 오디오 신호 또는 이러한 방법으로 처리한 오디오 신호 등을 저장하는 저장 매체를 수행하기 위한 개선된 오디오 디코더, 개선된 방법, 개선된 컴퓨터 프로그램이 요구된다. 보다 구체적으로, 인코딩된 비트스트림에서 전송된 오디오 정보의 소리 품질을 향상시키는 해결책을 찾는 것이 바람직하다.

발명의 내용

해결하려는 과제

[0005] 오디오 신호 등을 저장하는 저장 매체를 수행하기 위한 개선된 오디오 디코더, 개선된 방법, 개선된 컴퓨터 프로그램이 요구된다. 보다 구체적으로, 인코딩된 비트스트림에서 전송된 오디오 정보의 소리 품질을 향상시키는 해결책을 제공한다.

과제의 해결 수단

[0006] 청구 범위와 발명의 실시 예들의 상세한 설명에서 참조 부호는 단지 가독성을 향상시키기 위한 것일 뿐 한정을 의미하는 것은 아니다.

[0007] 본 발명의 목적은 선형 예측 계수를 포함하는 디코딩 된 오디오 정보를 기초로 디코딩 된 오디오 정보를 제공하는 오디오 디코더를 통해 해결되고, 오디오 디코더는 기울기 정보를 획득하기 위한 현재 프레임의 기울기 예측 계수들을 사용하는 잡음의 기울기를 조정하도록 구성되고, 잡음 삽입기는 기울기 연산기에 의해 획득되는 기울기 정보에 의존하는 현재 프레임으로 잡음을 부가하도록 구성된다. 또한, 본 발명의 목적은 선형 예측 계수(LPC, linear prediction coefficients)를 포함하는 인코딩된 오디오 정보에 기초한 디코딩 된 오디오 정보를 제공하는 방법에 의해 해결되고, 기울기 정보를 획득하기 위해 현재 프레임의 선형 예측 계수를 이용하는 잡음을 조정하는 단계 및 획득된 기울기 정보에 의존하는 현재 프레임으로 잡음을 부가하는 단계를 포함한다.

[0008] 두 번째 해결책은, 본 발명에서 선형 예측 계수(LPC)를 포함하는 인코딩된 오디오 정보에 기반한 디코딩 된 오디오 정보를 제공하는 오디오 디코더를 제안하고, 오디오 디코더는 잡음 레벨 정보를 획득하기 위한 적어도 하나의 이전 프레임의 선형 예측 계수를 이용하는 현재 프레임에 대한 잡음 레벨을 추적하도록 구성된 잡음 레벨 추정기를 포함하고 잡음 삽입기는 잡음 레벨 추정기에 의해 잡음 레벨 정보에 의존하는 현재 프레임으로 잡음을 부가하도록 구성된다. 또한, 본 발명의 목적은 선형 예측 계수(LPC)를 포함하는 인코딩된 오디오 정보에 기초한 디코딩 된 오디오 정보를 제공하는 방법에 의해 해결되고, 잡음 레벨 정보를 획득하기 위한 적어도 하나의 이전 프레임의 선형 예측 계수를 이용하는 현재 프레임에 대한 잡음 레벨을 추적하는 단계와 잡음 레벨 추정에 의해 제공되는 잡음 레벨 정보에 의존하는 현재 프레임으로 잡음을 부가하는 단계를 포함한다. 또한, 본 발명의 목적은 이러한 방법을 수행하는 컴퓨터 프로그램에 의해 해결된다. 상기 컴퓨터 프로그램은 컴퓨터에서 구동되고, 오디오 신호 또는 저장매체는 이러한 방법으로 취급되어온 오디오 신호를 저장한다.

[0009] 제안된 해결책은 잡음 충전 과정 측면에서 디코더 상에 제공되는 잡음을 조정하기 위한 CELP 비트스트림에서 부가 정보를 제공하는 것을 피한다. 이는 삽입된 잡음의 품질이 현재 또는 이전 디코딩 된 프레임들의 선형 예측 계수들에 단지 기반하여 증가 될 수 있는 동안 비트스트림으로 전달된 데이터의 양은 감소 될 수 있다는 것을 의미한다. 다시 말해 비트스트림으로 전달된 데이터의 양이 증가하는 잡음에 연관하는 부가정보가 생략될 수 있다. 본 발명은 로우-비트-레이트 디지털 코더(low-bit-rate digital coder)를 제공하도록 허용하고 비트 스트림에 관해 더 적은 대역폭을 소비하는 방법을 제공하고, 종래 기술에서의 해결책과 비교하여 주변 잡음의 개선된 품질을 제공한다.

[0010] 바람직하게 오디오 디코더는 현재 프레임의 프레임 유형을 결정하는 프레임 유형 결정기를 포함하고, 프레임 유

형 결정기는 현재 프레임의 프레임 유형이 음성 유형으로 감지될 때 잡음의 기울기를 조정하는 기울기 조정기를 활성화 하도록 구성된다.

[0011] 몇몇 실시예에 있어서, 프레임 유형은 프레임이 ACELP 또는 CELP 코딩 될 때 음성 유형 프레임으로서 프레임이 감지되도록 구성된다. 현재 프레임의 기울기에 따른 잡음 형성은 보다 자연스러운 배경 잡음을 제공할 수 있고 비트 스트림에서 코딩 된 원하는 신호의 배경 잡음에 대한 오디오 압축의 원하지 않는 효과를 줄일 수 있다. 이러한 원하지 않는 압축의 효과 및 인공물들은 음성 정보의 배경 잡음에 대해 두드러 질 수 있도록 하고, 현재 프레임으로 잡음을 추가하는 단계 전에 잡음의 기울기를 조정하는 단계에 의한 음성 유형 프레임 등을 추가하는 잡음의 품질을 향상시키는 이점이 될 수 있다. 이에 따라, 잡음 삽입기는 현재 프레임이 음성 프레임일 경우에 만 오직 현재 프레임으로 잡음을 추가 하도록 구성할 수 있고, 이 때문에 음성 프레임이 오직 잡음 충전에 의해 취급되는 경우에만 디코더 측면의 부하를 감소시킬 수 있다. 기울기 정보를 획득하기 위해 현재 프레임의 선형 예측 계수의 제1차 분석 결과를 사용하도록 구성된다. 선형 예측 계수들의 제1차 분석을 이용함으로써 비트스트림에서 노이즈 특성에 대한 추가 정보를 생략 가능하게 된다. 또한, 추가된 잡음의 조정은 현재 프레임의 선형 예측 계수들에 기반할 수 있고, 이는 현재 프레임의 오디오 정보 디코딩을 허용하기 위해 비트스트림으로 어쨌든 전송되어야 한다. 이는 현재 프레임의 선형 예측 계수들이 잡음의 기울기를 조정하는 과정에서 바람직하게 재 사용 된다는 것을 의미한다. 또한, 제1차 분석은 합리적으로 단순하기 때문에 이로써 오디오 디코더의 연산 복잡성이 크게 증가하지 않는다.

[0012] 본 발명의 몇몇 실시예에 있어서, 기울기 조정기는 제1차 분석으로서 현재 프레임의 선형 예측 계수들의 이득 g 의 연산을 이용하는 기울기 정보를 획득하도록 구성된다.

[0013] 보다 바람직하게, 이득 g 는 수식 $g = \sum[a_k \cdot a_{k+1}] / \sum[a_k \cdot a_k]$ 에 의해 주어지고, 상기 a_k 는 현재 프레임의 LPC 계수들이다. 몇몇 실시예에서, 둘 이상의 LPC 계수들 a_k 는 연산에서 사용된다. 바람직하게 $k = 0 \dots 15$ 이기 때문에 16개의 LPC 계수들의 총 개수가 사용된다. 발명의 실시예에서, 비트 스트림은 16개의 LPC 계수들보다 많거나 적게 코딩 될 수 있다. 현재 프레임의 선형 예측 계수들로서 비트스트림에서 쉽게 존재할 수 있기 때문에 기울기 정보는 추가 정보의 생성 없이 획득 될 수 있고, 따라서 비트스트림에서 전달된 데이터의 양을 감소 시킨다. 추가된 잡음은 단지 선형 예측 계수들을 이용함으로써 조정될 수 있고, 이는 인코딩된 오디오 정보의 디코딩에 필요하다.

[0014] 바람직하게, 기울기 조정기는 현재 프레임에 대한 직접적인 형태의 필터 $x(n)-g \cdot x(n-1)$ 의 전달 함수 연산을 이용한 기울기 정보를 획득 하도록 구성된다. 연산 유형은 합리적으로 용이하고, 디코더 측면에서 고 연산 전력을 요구하지 않는다. 도시된 바와 같이 이득 g 는 현재 프레임의 LPC 계수들로부터 쉽게 연산 될 수 있다. 이는 인코딩된 오디오 정보의 디코딩을 위해 순수 비트스트림 필수 데이터를 이용하는 동안 로우-비트-레이트 디지털 코더들에 대한 잡음 품질 향상을 허용한다.

[0015] 바람직한 실시예에 있어서, 잡음 삽입기는 현재 프레임으로 잡음을 추가하는 단계 전에 잡음의 기울기를 조정하기 위한 잡음으로 현재 프레임의 기울기 정보를 적용하도록 구성된다. 만일 잡음 삽입기가 이에 따라 구성되는 경우, 단순화된 오디오 디코더가 제공된다. 그리고 현재 프레임으로 조정된 잡음을 추가하는 단계는 기울기 정보의 첫 번째 적용에 의해 오디오 디코더의 단순하고 효과적인 방법으로 제공될 수 있다.

[0016] 발명의 실시예에 있어서, 오디오 디코더는 또한 잡음 레벨 정보를 획득하기 위한 적어도 하나의 이점 프레임의 선형 예측 계수를 이용하는 현재 프레임에 대한 잡음 레벨을 측정하도록 구성된 잡음 레벨 추정기를 포함하고 잡음 삽입기는 잡음 레벨 추정에 의해 제공되는 잡음 레벨 정보 상에 의존하는 현재 프레임으로의 잡음을 추가 하도록 구성된다. 이로써, 배경잡음의 품질과 그리고 전체 오디오 전달의 품질은 현재 프레임으로 추가된 잡음이 현재 프레임에서 아마도 존재하는 잡음 레벨에 따라 조정될 수 있기 때문에 향상된다. 예컨대 이전 프레임으로부터 측정 되었던 높은 잡음 레벨 때문에 높은 잡음 레벨이 현재 프레임에서 예상되는 경우, 잡음 삽입기는 현재 프레임으로 그것을 추가하는 단계 전에 현재 프레임으로 추가되는 노이즈의 레벨을 증가시키도록 구성될

수 있다. 따라서 추가되는 잡음은 현재 프레임에서 예상되는 잡음 레벨과 비교하여 너무 고요하거나 또는 시끄럽지 않도록 조정될 수 있다. 이러한 조정은 다시, 이전 프레임에서 잡음 레벨에 대한 정보를 제공하는 적어도 하나의 이전 프레임의 선형 예측 계수의 경우에서 비트스트림에서의 전달된 부가 정보에 기반하지 않는 반면 단지 비트스트림에서 전달된 필요한 데이터의 정보를 이용한다. 따라서, 바람직하게 현재 프레임으로 부가된 잡음은 유도된 그리고 노이즈 레벨 추정의 관점에서 스케일링된 기울기 g 를 사용하여 형성된다. 보다 바람직하게, 현재 프레임으로 부가되는 기울기 및 잡음 레벨은 현재 프레임이 음성 유형일 때 조정된다. 몇몇 실시예에 있어서, 기울기 및/또는 현재 프레임으로 부가된 잡음 레벨은 또한 예컨대 현재 프레임이 TCX 또는 DTX 유형과 같은 일반적인 오디오 유형일 때 조정된다.

[0017] 바람직하게, 오디오 디코더는 현재 프레임의 프레임 유형을 결정하는 유형 결정기를 포함하고 프레임 유형 결정기는 현재 프레임의 프레임 유형이 음성 또는 일반적인 오디오인 경우를 식별하도록 구성된다. 이에 따라 잡음 레벨 추정기는 현재 프레임의 프레임 유형에 의존하여 수행될 수 있다. 예컨대, 프레임 유형 결정기는 현재 프레임이 CELP 또는 ACELP인 경우를 감지하도록 구성되고, 이는 음성 프레임의 유형 또는 일반적인 오디오 프레임들인 TCX/MDCT 또는 DTX이다. 이러한 코딩 포맷들은 다른 원리를 따르기 때문에, 적절한 연산들이 프레임 유형에 의존하여 선택될 수 있도록 바람직하게 잡음 레벨 추정을 수행하는 단계 전에 프레임 유형을 결정하도록 한다.

[0018] 몇몇 실시예에 따른 본 발명의 오디오 디코더는 현재 프레임의 스펙트럼적으로 미 형성된 자극을 나타내는 제1 정보를 연산하도록 구성되고 오디오 디코더는 잡음 레벨 정보를 획득하도록 제1정보 및 제2정보의 지수를 연산하는 현재 프레임의 스펙트럼 스케일링에 대한 제2정보를 연산하도록 구성된다. 이로써 잡음 레벨 정보는 임의의 부가 정보의 사용 없이 획득될 수 있다. 따라서 코드의 비트 레이트는 낮게 유지된다.

[0019] 바람직하게, 오디오 디코더는 현재 프레임의 자극 신호를 디코딩 하도록 구성되고 현재 프레임이 음성 유형인 조건에서 잡음 레벨 정보를 획득하기 위해 제1정보로서 현재 프레임의 시간 영역 표현으로부터 그것의 제공된 e_{rms} 를 연산하도록 구성된다. 바람직하게 실시예에 있어서 오디오 디코더는 만일 현재 프레임이 CELP 또는 ACELP 유형인 경우에 따라 수행하도록 구성된다. 스펙트럼적으로 평탄화 된 자극신호(인지 영역에서)는 비트스트림으로부터 디코딩 되고 잡음 레벨 추정을 업데이트 하도록 사용된다. 현재 프레임에 대한 자극 신호의 제공된 e_{rms} 는 비트스트림이 읽혀진 이후 연산된다. 이러한 연산 유형은 소 소비 전력을 전혀 요구하지 않고 따라서 심지어 낮은 연산 전력으로의 오디오 디코더들에 의해 수행될 수 있다.

[0020] 바람직한 실시예에서 오디오 디코더는 제2정보로서 현재 프레임의 LPC 필터 전달함수의 최대 레벨 p 를 연산하도록 구성되고, 따라서 현재 프레임이 음성 유형인 조건 상에서 잡음 레벨 정보를 획득하기 위한 선형 예측 계수를 사용한다. 다시, 바람직하게 현재 프레임은 CELP 또는 ACELP 유형이다. 그리고 현재 프레임의 선형 예측 계수들을 재-사용 하는 것에 의해 최대 레벨 p 를 연산하는 것은 오히려 저렴하고, 이는 프레임에서 제한된 오디오 정보를 디코딩 하기 위해 또한 사용되고, 부가 정보는 생략될 수 있고 여전히 배경 잡음은 비트스트림의 데이터 레이트 증가 없이 향상될 수 있다.

[0021] 발명에 따른 바람직한 실시예에서, 오디오 디코더는 현재 프레임이 음성 유형인 경우 잡음 레벨 정보를 획득하기 위해 피크 레벨 p 및 제공된 e_{rms} 의 지수(quotient)를 연산함으로써 현재 오디오 프레임의 스펙트럼 최소(spectral minimum)를 연산하도록 구성된다.

[0022] 이 계산은 비교적 간단하고 다수의 오디오 프레임의 범위에서 잡음 레벨을 추정하는 데 유용한 수치를 제공할 수 있다. 따라서 현재 오디오 프레임들의 일련의 스펙트럼 최소 m_i 는 일련의 오디오 프레임에 의해 커버되는 시간 주기 동안 잡음 레벨을 추정하도록 사용될 수 있다. 이는 복잡도가 합리적으로 낮게 유지되는 동안 현재 프레임의 양호한 잡음 레벨 추정을 획득하도록 허용될 수 있다. 바람직하게 최대 레벨(peak level) p 는 공식 $p = \sum |a_k|$ 를 이용하여 연산될 수 있고, 상기 a_k 는 바람직하게 $k = 0 \dots 15$ 에서 선형 예측 계수들이다. 따라서 만일 프레임이 16개의 선형 예측 계수들을 포함하는 경우, p 는 몇몇 실시예에서 바람직하게 16 a_k 의 진폭을 통해 합산함으로써 연산된다.

- [0023] 바람직하게 오디오 디코더는 현재 프레임의 미 형성 MDCT-자극을 디코딩 하기 위해 구성되고 현재 프레임이 일반적인 오디오 유형인 경우 제1정보로서 잡음 레벨 정보를 획득하기 위해 현재 프레임의 스펙트럼 영역 표현으로부터 그것의 제공근 e_{rms} 을 연산 할 수 있도록 구성된다. 이는 현재 프레임이 음성 프레임이 아닌 일반적인 오디오 프레임 일 때마다 발명의 바람직한 실시 예가 된다. MDCT 또는 DTX에서의 스펙트럼 영역 표현은 예컨대, CELP 또는 (A)CELP 프레임들과 같은 음성 프레임에서의 시간영역 표현으로 대부분 동일하다. MDCT에서 다른 점은 Parseval의 이론을 고려하지 않는다는 것이다. 따라서 바람직하게 일반적인 오디오 프레임에 대한 제공근 e_{rms} 는 음성 프레임에 대한 제공근 e_{rms} 와 유사한 방법으로 연산된다. 그리고 바람직하게 WO 2012/110476 A1제시된 일반적인 오디오 프레임의 상응하는 LPC계수들을 연산한다. 선택적인 실시예에서, MDCT 전력 스펙트럼 주파수 대역은 일정한 주파수 대역을 가지고 이에 따라 스펙트럼의 스케일은 선형 스케일에 대응한다. 연산된LPC계수들에 상응하는 선형 스케일과 같이 예컨대 ACELP 또는 CELP 에서 연산된 것과 같이 동일한 프레임의 시간 영역 표현에서 LPC 계수들과 유사하다. 또한 바람직하게 현재 프레임이 일반적인 오디오 유형인 경우, WO 2012/110476 A1 에 개시된 MDCT 프레임으로부터 연산된 현재 프레임의 LPC 필터 전달함수의 최대 레벨 p 가 제2정보로서 연산되고, 따라서 현재 프레임이 일반적인 오디오 유형인 경우에서의 소음 레벨 정보를 획득하기 위한 선형 예측 계수를 사용한다. 그리고 만일 현재 프레임이 일반적인 오디오 유형인 경우, 현재 프레임이 일반적인 오디오 유형의 경우의 잡음 레벨 정보를 획득하기 위해 바람직하게 최대 레벨(peak level) p 및 제공근 e_{rms} 의 지수를 연산함으로써 현재 오디오 프레임의 스펙트럼 최소를 연산한다. 따라서 현재 오디오 프레임의 스펙트럼 최소 m_f 를 나타내는 지수는 만일 현재 프레임이 일반적인 유형 또는 음성 유형인 경우에 관계없이 획득될 수 있다.
- [0024] 바람직한 실시예에서, 오디오 디코더는 프레임의 유형에 관계없이 소음 레벨 추정기에서 현재 오디오 프레임으로부터 획득되는 지수를 인큐잉하도록 구성되고, 잡음 레벨 추정기는 다른 오디오 프레임들로부터 획득되는 둘 이상의 지수를 저장하는 소음 레벨을 포함한다. 이는 예컨대 낮은-지연 통합된 음성 및 오디오 디코딩(LD-USAC, EVS)을 적용할 때 음성 프레임과 일반 오디오 프레임의 디코딩 사이를 전환 하도록 구성된다. 이에 따라 여러 프레임들을 통한, 프레임 유형을 무시하는 평균 잡음 레벨이 획득될 수 있다. 바람직하게 잡음 레벨 저장은 열 개 또는 이상의 이전 오디오 프레임들로부터 획득한 열 개 또는 이상의 지수들을 유지할 수 있다. 예컨대, 잡음 레벨 저장은 30프레임들의 지수에 대한 공간을 포함할 수 있다. 따라서, 잡음 레벨은 확장된 이전시간에 대한 현재 프레임을 연산할 수 있다. 실시예에 있어서, 지수는 현재 프레임이 음성 유형으로 감지될 때 잡음 레벨 추정기에서 오직 인큐딩 된다. 다른 실시예에서, 지수는 현재 프레임이 일반적인 오디오 유형으로 감지될 때 잡음 레벨 추정기에서 오직 인큐딩 된다.
- [0025] 바람직하게 잡음 레벨 추정기는 다른 오디오 프레임의 둘 이상 지수의 통계적인 분석을 기초로 추정되도록 구성된다. 본 발명의 실시예에서, 오디오 디코더는 통계적으로 지수들을 분석하기 위해 잡음 전력 스펙트럼 밀도에 기초한 최소 제공 오차를 이용하도록 구성된다. 이러한 추적은 Hendriks, Heusdens and Jensen [2]의 공개에서 설명된다. 만일 [2]에 대한 방법이 적용되는 경우, 오디오 디코더는 통계적인 분석에서 추적 값의 제공근을 이용하도록 구성되고, 이 때문에 현재의 경우 진폭 스펙트럼이 직접적으로 검색된다. 본 발명의 다른 실시예에서, [3]으로부터 알려진 최소 통계는 다른 오디오 프레임의 둘 이상의 지수를 분석하도록 사용된다.
- [0026] 바람직한 실시예에 있어서, 디코딩 된 코어 코더 출력을 획득하기 위해 현재 프레임의 선형 예측 계수를 이용한 현재 프레임의 디코딩 된 오디오 정보로 구성되는 디코더를 포함하고 선형 예측 계수에 의존하는 잡음을 부가하는 잡음 삽입기를 포함하고, 이는 현재 프레임의 오디오 정보 디코딩 및/또는 하나 이상의 이전 프레임의 오디오 정보를 디코딩 할 때 사용된다. 따라서, 잡음 삽입기는 현재 프레임의 오디오 정보 디코딩에 사용되는 동일한 선형 예측 계수들의 사용을 만든다. 부가정보는 잡음 삽입기 재건을 위해 생략 될 수 있다.
- [0027] 바람직하게, 오디오 디코더는 현재 프레임을 강조 하지 않기(de-emphasize)위한 강조 약화 필터(de-emphasize filter)를 포함하고, 오디오 디코더는 현재 프레임으로 잡음 삽입기가 잡음을 부가한 이후 현재 프레임 상의 강조 약화 필터를 적용하도록 구성된다. 강조 약화(de-emphasis)는 저 주파수 부스팅 첫 번째 순서이기 때문에, 이는 낮은-복잡도, 낮은 주파수에서의 청각 잡음 인공물을 회피하는 부가된 잡음의 가파른 IIR 하이-패스 필터링을 허용한다.

[0028] 바람직하게, 오디오 디코더는 잡음 생성기를 포함하고, 잡음 생성기는 잡음 삽입기에 의해 현재 프레임으로 부가된 잡음을 생성하도록 구성된다. 오디오 디코더에 포함된 잡음 생성기를 구비하는 것은 더 이상 외부 잡음 생성기가 필요하지 않기 때문에 보다 편리한 오디오 디코더를 제공할 수 있다. 선택적으로, 잡음은 외부 잡음 생성기에 의해 공급될 수 있고, 이는 인터페이스를 통해 오디오 디코더와 연결될 수 있다. 예를 들어, 잡음 생성기의 특별한 유형은 현재 프레임에서 강화된 배경 잡음에 의존하여, 적용될 수 있다.

[0029] 바람직하게, 잡음 생성기는 임의의 백색 잡음을 생성하도록 구성된다. 이러한 잡음은 충분히 일반적인 배경 잡음과 유사하고 이러한 잡음생성기는 용이하게 제공 될 수 있다.

[0030] 본 발명의 바람직한 실시예에 있어서, 잡음 생성기는 1비트 당 샘플 보다 작은 인코딩된 오디오 정보의 비트 레이트 조건 상에서의 현재 프레임으로 잡음을 부가 하도록 구성된다. 바람직하게 인코딩된 오디오 정보의 비트 레이트는 0.8비트 당 샘플보다 작다. 보다 바람직하게 잡음 생성기는 0.5 비트당 샘플보다 작은 인코딩된 오디오 정보의 비트 레이트 조건 상의 현재 프레임으로 잡음을 부가하도록 구성된다.

[0031] 바람직한 실시예에서, 오디오 디코더는 코딩 된 오디오 정보를 디코딩 하기 위한 하나 이상의 코더들 AMR-WB, G.718또는 LD-USAC (EVS)에 기반한 코더를 사용하도록 구성된다. 이들은 잘 알려져 있고, 넓은 확산 (A) CELP 코더(wide spread (A)CELP coders)는 그러한 잡음 작성 방법의 추가 용도에 있어서 매우 유리할 수 있다.

발명의 효과

[0032] 비트 스트림에서 배경 잡음에 대한 부가 정보를 생략 할 수 있다.

도면의 간단한 설명

[0033] 본 발명의 실시 예들은 도면과 관련하여 이하에서 설명된다.

도 1은 본 발명의 실시예에 따른 오디오 디코더의 제 1 실시 예를 나타낸다.

도 2는 도 1에 따른 오디오 디코더에 의해 수행될 수 있는 본 발명에 따른 오디오 디코딩을 수행하기 위한 첫 번째 방법을 도시한다.

도 3은 본 발명에 따른 오디오 디코더의 두 번째 실시예를 도시한다;

도 4는 도 1에 따른 오디오 디코더에 의해 수행될 수 있는 본 발명에 따른 오디오 디코딩을 수행하기 위한 두 번째 방법을 나타낸 것이다.

도 5는 본 발명에 따른 오디오 디코더의 세 번째 실시 예를 도시 한 도면이다.

도 6은 도 5에 따른 오디오 디코더에 의해 수행 될 수 있는 본 발명에 따른 오디오 디코딩을 수행하기 위한 세 번째 방법을 도시한다.

도 7은 노이즈 레벨 추정을위한 스펙트럼 최소값의 m_f 를 연산하기 위한 방법의 예를 도시한다;

도 8은 LPC 계수들로부터 유도 된 기울기를 도시하는 도면을 나타낸다.

도 9는 MDCT 파워 스펙트럼으로부터 LPC 필터 등가물이 결정되는 방법을 예시하는 다이어그램을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0034] 본 발명은 도 1 내지 9와 관련하여 상세하게 설명된다. 본 발명은 도시된 그리고 설명된 실시예에 결코 한정되지 않는다.

[0035] 도 1은 본 발명의 실시예에 따른 오디오 디코더의 첫 번째 실시 예를 보여준다. 오디오 디코더는 인코딩된 오디오 정보에 기반한 디코딩 된 오디오 정보를 제공하도록 구성된다. 오디오 디코더는 인코딩된 오디오 정보를 디코딩 하기 위해 AMR-WB, G.718 및 LD-USAC (EVS)를 기반으로 할 수 있는 코더를 사용하도록 구성될 수 있다. 인코딩된 오디오 정보는 선형 예측 계수들(LPC, linear prediction coefficients)을 포함하고, 이는 계수 a_k 로서 개별적으로 지정될 수 있다. 오디오 디코더는 기울기 정보를 획득하기 위해 현재 프레임의 선형 예측 계수들을 사용하는 잡음의 기울기를 조정하도록 구성되는 기울기 조정기 및 기울기 연산기에 의해 획득된 기울기 정보에 의존하는 현재 프레임으로 잡음을 부가하도록 구성되는 잡음 삽입기를 포함한다. 잡음 삽입기는 1비트 당 샘플 보다 작은 인코딩된 오디오 정보의 비트레이트 조건 상의 현재 프레임으로 잡음을 부가하도록 구성된다. 또한, 잡음 삽입기는 현재 프레임이 음성 프레임인 조건 상에서 현재 프레임으로 잡음을 부가하도록 구성될 수 있다. 따라서, 잡음은 디코딩 된 오디오 정보의 전반적인 소리 품질을 향상시키기 위해 현재 프레임으로 부가될 수 있고, 이는 특히 음성 정보의 배경잡음에 관한 코딩 인공물들(artifacts) 때문에 손상 될 수 있다. 잡음의 기울기가 현재 오디오 프레임의 기울기 관점에서 조정될 때, 전반적인 소리 품질은 비트 스트림에서 부가 정보에 의존하는 것 없이 개선 될 수 있다. 따라서, 비트-스트림으로 전달되는 데이터의 양이 줄어들 수 있다.

[0036] 도 2는 도1에 따른 오디오 디코더에 의해 수행될 수 있는 본 발명에 따른 오디오 디코딩 수행에 대한 첫 번째 방법을 보여준다. 도 1에 도시된 오디오 디코더의 기술적인 세부사항은 방법적인 특징들과 함께 설명된다. 오디오 디코더는 인코딩된 오디오 정보의 비트스트림을 읽도록 구성된다. 오디오 디코더는 현재 프레임의 프레임 유형을 결정하는 프레임 유형 결정기를 포함하고, 프레임 유형 결정기는 현재 프레임의 프레임 유형이 음성 유형으로 감지 될 때 잡음의 기울기를 조정하는 기울기 조정기를 활성화 하도록 구성된다. 따라서, 오디오 디코더는 프레임 유형 결정기 적용에 의한 현재 오디오 프레임의 프레임 유형을 결정한다. 만일 현재 프레임이 ACELP프레임인 경우, 프레임 유형 결정기는 기울기 조정기를 활성화 한다. 기울기 조정기는 기울기 정보를 획득하기 위한 현재 프레임의 선형 예측 계수의 첫 번째 순서 분석 결과를 이용하도록 구성된다. 보다 구체적으로, 기울기 조

정기는 첫 번째 순서 분석으로서 공식
$$g = \sum [a_k \cdot a_{k+1}] / \sum [a_k \cdot a_k]$$
를 이용하여 이득 g 를 연산하고, 상기 a_k 는 현재 프레임의 LPC 계수들이다. 도 8은 LPC 계수들로부터 유도된 기울기를 나타내는 다이어그램이다. 도8은 “참조”의 두 프레임을 보여준다. “s”는 많은 양의 고 주파수를 구비하고, 기울기가 올라간다. “ee”는 많은 양의 저 주파를 구비하는 구비하고, 기울기가 내려간다. 도 8에서의 기울기 스펙트럼은 직접적인 형성 필터 $x(n) - g \cdot x(n-1)$ 의 전달함수를 도시하고, g 는 주어진 대로 정의된다. 따라서, 기울기 조정기는 비트 스트림에서 제공되는 LPC 계수들의 사용을 생성하고, 인코딩된 오디오 정보를 디코딩 하도록 한다. 부가 정보가 생략됨에 따라 비트 스트림으로 전달된 데이터의 양이 감소될 수 있다. 또한, 기울기 조정기는 직접적인 형성 필터 (direct form filter) $x(n) - g \cdot x(n-1)$ 의 전달 함수의 연산을 이용하는 기울기 정보를 획득하도록 구성된다. 이에 따라, 기울기 조정기는 이전에 연산된 이득 g 를 이용하는 직접적인 형성 필터 $x(n) - g \cdot x(n-1)$ 로 부터의 전달 함수를 연산함으로써 현재 프레임에서의 오디오 정보의 기울기를 연산한다. 기울기 정보가 획득된 후에, 기울기 조정기는 현재 프레임의 기울기 정보에 의존하는 현재 프레임으로 부가되는 잡음의 기울기를 조정한다. 이후에, 조정된 잡음은 현재 프레임으로 부가된다. 또한, 도 2에 도시되지 않은, 오디오 디코더는 현재 프레임을 약화시키는 강조 약화 필터(de-emphasis)를 포함하고, 오디오 디코더는 현재 프레임으로 잡음을 부가하는 노이즈 삽입기 이후 현재 프레임의 강조 약화 필터를 적용하도록 구성될 수 있다. 또한 저 복잡도 역할 프레임을 강조 약화한 후에, 낮은 복잡도, 부가된 잡음의 가파른 IIR 고 패스 필터링(high-pass filtering)로서 역할을 하고, 오디오 디코더는 디코딩 된 오디오 정보를 제공한다. 따라서, 도 2에 따른 방법은 배경 잡음의 품질을 개선하기 위해 현재 프레임으로 부가된 잡음의 기울기를 조정함으로써 오디오 정보의 소리 품질 향상을 허용한다.

[0037] 도 3은 본 발명의 실시예에 따라 오디오 디코더의 두 번째 실시 예를 보여준다. 오디오 디코더는 인코딩 된 오디오 정보를 기초로 디코딩 된 오디오 정보를 제공 하도록 다시 구성된다. 오디오 디코더는 인코딩된 오디오 정보를 디코딩 하기 위해 AMR-WB, G.718 and LD-USAC (EVS)에 기반할 수 있는 코더를 사용하도록 다시 구성된다. 인코딩된 오디오 정보는 a_k 계수들로서 개별적으로 지정 될 수 있는 선형 예측 계수들(LPC, linear prediction coefficients)를 다시 포함한다. 두 번째 실시예에 따른 오디오 디코더는 잡음 레벨 추정기에 의해 제공된 잡음 레벨 정보에 의존하는 현재 프레임으로 잡음을 부가하도록 구성되는 잡음 레벨 정보와 잡음 삽입기를 획득하기

위한 적어도 하나의 프레임의 선형 예측 계수를 이용하는 현재 프레임에 대한 잡음 레벨을 추정하도록 구성된다. 잡음 삽입기는 0.5 비트 당 샘플보다 작은 인코딩된 오디오 정보의 비트레이트 조건에서 현재 프레임으로 잡음을 부가하도록 구성된다. 또한, 잡음 삽입기는 현재 프레임이 음성 프레임인 경우 현재 프레임으로 잡음을 부가하도록 구성된다. 또한, 다시 잡음은 인공물 코딩 때문에 손상된 디코딩 된 오디오 정보의 전반적인 음향 품질을 개선하기 위해 현재 프레임으로, 특히 음성 정보의 배경 잡음에 관해 부가될 수 있다. 잡음의 잡음 레벨이 적어도 하나의 이전 오디오 프레임 관점에서 조정될 때, 전반적인 소리 품질은 비트 스트림에서 부가 정보 의존 없이 개선될 수 있다. 따라서, 비트 스트림으로 전송된 데이터의 양이 줄어들 수 있다.

[0038]

도 4는 도 3의 오디오 디코더에 의해 수행 될 수 있는 본 발명에 따른 오디오 디코딩을 수행하는 두 번째 방법을 보여준다. 도 3에 묘사된 기술적인 오디오 디코더의 기술적인 세부사항은 방법적인 특징들과 함께 설명된다. 도 4에 따라, 오디오 디코더는 현재 프레임의 프레임 유형을 결정하기 위해 비트스트림을 읽도록 구성된다. 또한, 오디오 디코더는 현재 프레임의 프레임 유형을 결정하는 프레임 유형 결정기를 포함하고, 프레임 유형 결정기는 현재 프레임의 프레임 유형이 음성 또는 일반적인 오디오 유형인 경우를 식별하도록 구성되고, 이로써 노이즈 레벨 추정은 현재 프레임의 프레임 유형에 의존하여 수행 될 수 있다. 일반적으로, 오디오 디코더는 현재 프레임의 스펙트럼적으로 미 형성된 자극을 나타내는 첫 번째 정보를 연산하도록 구성되고, 잡음 레벨 정보를 획득하기 위한 첫 번째 정보와 두 번째 정보의 지수를 연산하는 현재 프레임의 스펙트럼 스케일링을 고려하는 두 번째 정보를 연산하도록 구성된다. 예컨대, 만약 프레임 유형이 ACELP 인 경우, 이는 음성 프레임 유형이고, 오디오 디코더는 현재 프레임의 자극 신호를 디코딩 한다. 그리고 자극 신호의 시간 영역 표현으로부터 현재 프레임 f에 대한 그것의 제공근 e_{rms} 을 연산한다. 이는 오디오 디코더가 현재 프레임의 자극 신호를 디코딩 하도록 구성되고, 현재 프레임이 음성 유형인 조건 상에서 잡음 레벨 정보를 획득하기 위한 첫 번째 정보로서의 현재 프레임의 시간 영역 표현으로부터 그것의 제공근 e_{rms} 을 연산하는 것을 의미한다. 다른 경우에서, 만약 프레임 유형이 MDCT 또는 DTX인 경우, 이는 일반적으로 오디오 프레임 유형이고, 오디오 디코더는 현재 프레임의 자극 신호(excitation signal)를 디코딩하고 자극 신호의 시간 영역 표현 등가물 (time domain representation equivalent)로부터 현재 프레임 f에 대한 그것의 제공근을 연산한다. 이는, 오디오 디코더가 현재 프레임의 미 형성 MDCT-자극을 디코딩 하도록 구성되고, 현재 프레임이 일반적인 오디오 유형인 상태에서 잡음 레벨 정보를 획득하기 위해 첫 번째 정보로서 현재 프레임의 스펙트럼 영역 표현으로부터 그것의 제공근 e_{rms} 을 연산하도록 구성된다. WO 2012/110476 A1에서 구체적인 수행방법이 설명된다. 또한, 도 9는 LPC 필터 등가물이 어떻게 MDCT 파워-스펙트럼을 결정하는지 나타내는 다이어그램이다. 묘사된 스케일이 바크 스케일(Bark scale)인 경우, LPC 계수 등가물은 선형 스케일로부터 또한 획득될 수 있다. 특히, 그들은 선형 스케일로부터 획득될 때, 연산된 LPC 계수 등가물들이 예컨대 CELP에서 코딩될 때 동일한 프레임의 시간 영역 표현으로부터 연산된 것들과 매우 유사하다.

[0039]

또한, 도 3에 따른 오디오 디코더는, 도 4의 방법 차트에 의해 나타내어 진 것과 같이 두 번째 정보로서 현재 프레임의 LPC 필터 전달함수의 최대 레벨 p를 연산하도록 구성되고, 따라서 현재 프레임이 음성 유형인 조건 상에서의 잡음 레벨 정보를 획득하기 위한 선형 예측 계수를 이용한다. 이는, 오디오 디코더가 수식 $p = \sum |a_k|$ 에 따른 현재 프레임 f의 LPC 분석 필터의 전달 함수의 최대 레벨 p를 연산하는 것을 의미한다. 상기 a_k 는 $k = 0 \dots 15$ 로의 선형 예측 계수이다. 만약 프레임이 일반적인 오디오 프레임인 경우, LPC 계수 등가물(equivalent s)은 WO 2012/110476 A1와 앞서 기술한 바와 같이 도 9에서 보여진 대로, 현재 프레임의 스펙트럼 영역 표현으로부터 획득된다. 도 4에 도시된 바와 같이, 최대 레벨 p 연산 이후, 현재 프레임 f의 스펙트럼 최소 m_f 는 e_{rms} 를 p로 분할하는 것에 의해 연산된다. 따라서, 오디오 디코더는 이 실시예에서 e_{rms} 인, 현재 프레임의 스펙트럼적으로 미형성(unshaped)된 자극을 나타내는 첫 번째 정보와 이 실시예에서 최대 레벨 p인, 현재 프레임의 스펙트럼 스케일링에 대한 두 번째 정보를 연산하도록 구성되고, 잡음 레벨 정보를 획득하기 위해 첫 번째 정보 및 두 번째 정보의 지수를 연산하도록 구성된다. 그리고 현재 프레임의 스펙트럼 최소는 잡음 레벨 추정기에서 인큐잉(enqueued)되며, 오디오 디코더는 프레임 유형에 관계없이 오디오 레벨 추정기에서 현재 오디오 프레임으로부터 획득되는 지수를 인큐잉하도록 구성되고, 잡음 레벨 추정기는 다른 오디오 프레임으로부터 획득된, 이 경우에는 스펙트럼 최소 m_f 인, 둘 이상의 지수에 대한 잡음 레벨 저장매체를 포함한다. 보다 구체적으로, 잡음 레벨 저장매체는 잡음 레벨 추정을 위해 50프레임들로부터 지수들을 저장할 수 있다. 또한, 잡음 레벨 추정기는 다른 오디오 프레임의 둘 이상의 지수들의 통계적인 분석을 기반으로 잡음 레벨을 추정 하도록 구성되고, 따라

서 스펙트럼 최소 m_f 를 추정하도록 구성된다. 지수 m_f 연산에 대한 단계들은 도 7에서 세부적으로 묘사되고, 필요한 연산 단계들을 나타낸다. 두 번째 실시예에서, 잡음 레벨 추정기는 [3]으로부터 알려진 최소 통계에 기반하여 동작한다. 소음은 최소 통계에 기반한 현재 프레임의 추정된 잡음 레벨에 따라 스케일링 되고, 이후 현재 프레임이 음성 프레임 형태인 경우, 현재 프레임으로 부가된다. 마지막으로, 현재 프레임은 강조 약화(de-emphasized)된다(도 4 미 도시). 따라서, 두 번째 실시예 또한 비트 스트림으로 전달된 데이터의 양 감소를 허용하기 위해 잡음 충전에 대한 부가정보를 생략을 허용한다. 따라서, 오디오 정보의 소리 품질은 데이터 레이트 증가 없이 디코딩 단계 동안 배경 잡음 강화에 의해 개선될 수 있다. 시간/주파수 변형이 필요하지 않기 때문에, 그리고 잡음 레벨 추정기는 오직 프레임 당 한번이기 때문이라는 점에 유의하고(다중 서브-밴드 상이 아님) 표현된 잡음 충전은 잡음적인 음성의 낮은-비트-레이트 코딩을 개선할 수 있는 반면 매우 낮은 복잡도를 나타낸다.

[0040] 도 5는 본 발명에 따른 오디오 디코더의 세 번째 실시예를 나타낸다. 오디오 디코더는 인코딩된 오디오 정보에 기반한 디코딩된 오디오 정보를 제공하도록 구성된다. 오디오 디코더는 인코딩된 오디오 정보를 디코딩하기 위해 LD-USAC에 기반한 코더를 이용하도록 구성된다. 인코딩된 오디오 정보는 선형 예측 계수들(LPC, linear prediction coefficients)을 포함하고, 이는 계수들 a_k 로서 개별적으로 지정된다. 오디오 디코더는 기울기 정보를 획득하기 위한 현재 프레임의 선형 예측 계수들을 사용하는 잡음의 기울기를 조정하도록 구성되는 기울기 조정기를 포함하고 잡음 레벨 추정기는 잡음 레벨 정보를 획득하기 위한 적어도 하나의 이전 프레임의 선형 예측 계수를 이용하는 현재 프레임에 대한 잡음 레벨을 추정하도록 구성된다. 또한, 오디오 디코더는 기울기 연산기에 의해 획득된 기울기 정보에 의존하는 현재 프레임으로 잡음이 부가되고 잡음 레벨 추정기에 의해 제공되는 잡음 레벨 정보에 의존하도록 구성된 잡음 삽입기를 포함한다. 따라서, 잡음은 디코딩된 오디오 정보의 전반적인 소리 품질을 향상시키기 위해 현재 프레임으로 부가되고, 이는 특별히 음성 정보의 배경 소음에 대한, 코딩 인공물 때문에 손상될 수 있고, 기울기 연산기에 의해 획득된 기울기 정보에 의존하고, 잡음 레벨 추정기에 의해 제공된 잡음 레벨 정보에 의존한다. 이러한 실시예에서, 임의의 소음 생성기(도면 미도시)는 스펙트럼 적인 백색 잡음을 생성하는 오디오 디코더에 의해 포함된다. 이는 그리고 잡음 레벨 정보에 따라 스케일링 되고, 앞서 전술한 바와 같이, g -유도된 기울기를 이용하여 미형성된다.

[0041] 도 6은 도5에 따른 오디오 디코더에 의해 수행될 수 있는 본 발명의 실시예에 따른 오디오 디코딩 수행에 대한 세 번째 방법을 보여준다. 비트 스트림은 읽혀지고, 프레임 유형 결정기라고 불리는 프레임 유형 결정기는 현재 프레임이 음성 프레임(ACELP) 또는 일반적인 오디오 프레임(TCX/MDCT)인 경우를 결정한다. 프레임 유형에 관계없이, 프레임 헤더는 디코딩 되고, 스펙트럼 적으로 평평하고, 인지 영역에서 미형성 자극은 디코딩 된다. 음성 프레임의 경우, 자극 신호는 앞서 전술한 바와 같이 시간-영역 자극이다. 만일 프레임이 일반적인 오디오 프레임인 경우, MDCT-영역 잔류는 디코딩 된다(도메인 스펙트럼). 시간 영역 표현과 스펙트럼 영역 표현은 각각 도 7에서 도시된 대로 잡음 레벨을 추정하고, 앞서 전술한 바와 같이 LPC계수들을 또한 이용하는 것은 임의의 부가정보 또는 부가 LPC 계수들을 이용하는 대신 비트스트림을 디코딩 한다. 프레임의 두 가지 유형의 잡음 정보는 기울기를 조정하도록 인코딩되고 잡음의 잡음 레벨은 현재 프레임이 음성 프레임인 조건 상에서 현재 프레임으로 부가된다. ACELP 음성 프레임으로의 잡음을 부가한 이후, ACELP 음성 프레임은 IIR 및 음성 프레임들에 의해 강조가 약화 되고(de-emphasized) 일반적인 오디오 프레임은 시간 신호에서 결합되고, 디코딩된 오디오 정보를 나타낸다. 부가된 잡음의 약화된 스펙트럼상의 가파른 하이-패스 영향은 작게 삽입된 도 6의 도 I, II, 및 III에 의해 표시된다. 다시 말해, 도 6에 따라, 앞서 전술된 ACELP 잡음 충전 시스템은 LD-USAC (EVS), xHE-AAC [6]의 낮은 지연 변형 디코더에서 구현되었고, 이는 프레임 당 기준으로 코딩 ACELP (음성)와 MDCT (음악 / 소음) 사이를 전환할 수 있다. 도 6에 따른 삽입 과정은 아래와 같이 요약된다.

[0042] 1.비트 스트림이 읽혀지면, 그리고 현재 프레임이ACELP 또는 MDCT 또는 DTX 프레임인 경우를 결정한다. 프레임 유형에 관계없이, 스펙트럼적으로 평평한 자극 신호(인지 영역에서)는 디코딩 되고 그리고 세부적으로 이하에 설명된 바와 같이 잡음 레벨 추정을 업데이트 하기 위해 디코딩 된다. 그리고 신호는 강조 약화(de-emphasis)를 위해 완전히 재건 되고, 이는 마지막 단계이다.

[0043] 2. 만일 프레임이 ACELP로 코딩된 경우, 잡음 삽입에 대한 기울기(전반적인 스펙트럼 모양)는 LPC 필터 계수들의 첫 번째 순서 LPC 분석에 의해 연산된다. 기울기는 16개의 LPC 계수들의 이득 g 로부터 유도되고, 이는 $g = \sum [a_k \cdot a_{k+1}] / \sum [a_k \cdot a_k]$ 에 의해 주어진다.

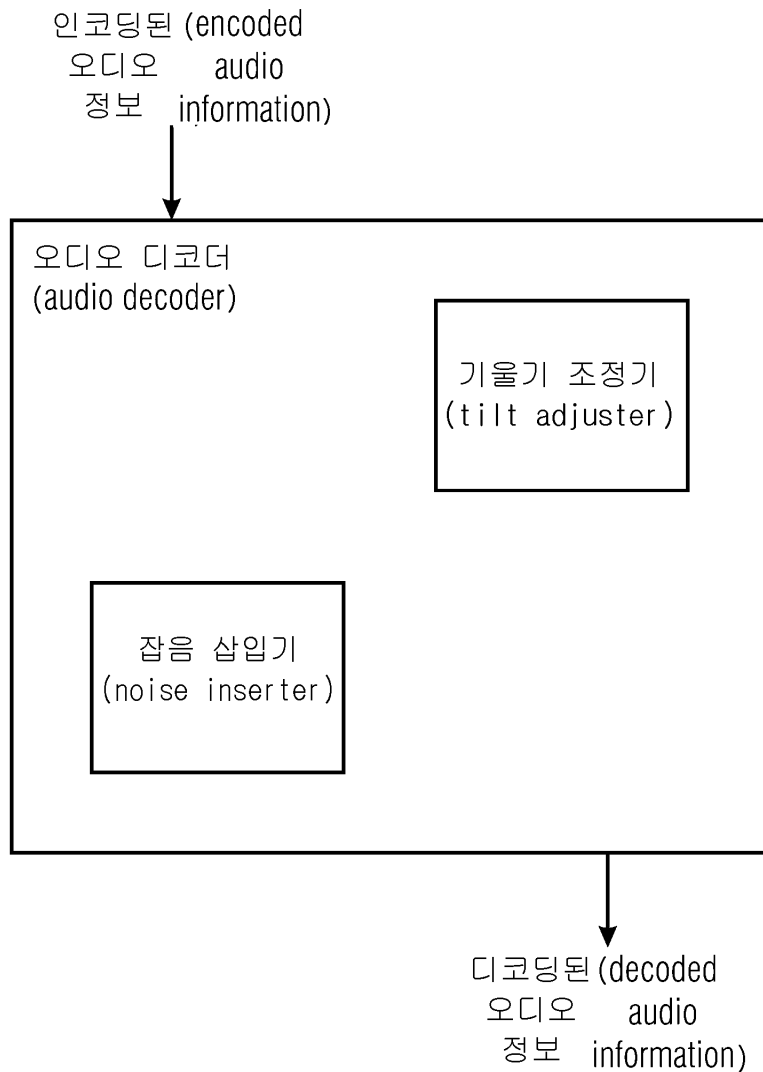
- [0044] 3. 만일 프레임이 ACELP 코딩 되면, 잡음 형성 레벨과 기울기는 디코딩 된 프레임으로의 잡음 부가를 수행하기 위해 이용된다: 임의의 잡음 생성기는 스펙트럼적인 백색 잡음 신호를 생성하고, 그리고 이는 g - 유도된 기울기를 이용하여 스케일링 및 형성된다.
- [0045] 4. ACELP 프레임에 대해 형성된 그리고 레벨 링 된 잡음 신호는 마지막 강조 약화 필터링 단계 바로 전에 디코딩 된 신호로 부가 된다. 강조 약화(de-emphasis)는 낮은 주파수 첫 번째 IIR 증폭(boosting)때문에, 도 6에 도시된 바와 같이 낮은 주파수에서 청각적인 잡음 인공물을 피하는 낮은 복잡성을 허용하고, 부가된 잡음의 가파른 IIR 하이 패스 필터링을 허용한다.
- [0046] 1단계에서 잡음 레벨 추정은 현재 프레임(또는 MDCT-영역이 시간 영역과 등가인 경우, e_{rms} 가 만일 ACELP 프레임인 경우에 대해 연산되는 것을 의미함)에 대한 자극 신호의 제공된 e_{rms} 연산에 의해 수행되고 그리고 LPC 분석 필터 전달 함수의 최대 레벨 p 에 의한 분할에 의해 수행된다. 이는 도 7에서의 프레임 f 의 스펙트럼 최소의 레벨 m_f 를 산출하고, m_f 는 최종적으로 예컨대, 최소 통계[3]에 기반하여 연산하는 잡음 레벨 추정기에서 인큐잉된다. 시간/주파수 변형이 필요하지 않다는 것을 유의하고, 레벨 추정기는 프레임 당 오직 한번 동작하기 때문에 (다수 서브-밴드들 상 아님) 전술한 CELP 잡음 충전 시스템은 잡음 음성의 낮은 비트 레이트 코딩을 개선할 수 있는 반면 낮은 복잡도를 나타낸다.
- [0047] 몇몇 양태는 오디오 디코더의 문맥으로 기술되었지만, 이러한 양상은 또한 블록 또는 장치가 방법 단계 또는 방법 단계의 기능에 대응하는 대응하는 방법의 설명을 나타내는 것이 분명하다.
- [0048] Analogously, aspects described in the context of a method step also represent a description of a corresponding block or item or feature of a corresponding audio decoder. Some or all of the method steps may be executed by (or using) a hardware apparatus, like for example, a microprocessor, a programmable computer or an electronic circuit. In some embodiments, some one or more of the most important method steps may be executed by such an apparatus.
- [0049] 유사하게, 방법 단계들의 문맥으로 설명 양태는 해당 오디오 디코더의 대응하는 블록 또는 항목 또는 기능에 대한 설명을 나타낸다. 방법 단계들의 일부 또는 전부는, 예를 들어 같은 하드웨어 장치 (또는 사용)에 의해 실행될 수 있는, 마이크로 프로세서, 프로그래머블 컴퓨터 또는 전자 회로. 일부 실시예에서 가장 중요한 방법 단계 중 일부는 하나 이상의 이러한 장치에 의해 실행될 수 있다.
- [0050] 본 발명의 인코딩 된 오디오 신호는 디지털 저장 매체에 저장 될 수도 있고 이러한 무선 전송 매체 나 인터넷 등의 유선의 전송 매체로 전송 매체를 통해 전송 될 수 있다.
- [0051] 특정 구현 요구 사항에 따라, 본 발명의 실시 예들은 하드웨어 또는 소프트웨어로 구현 될 수 있다. 구현은, 예를 들면, 디지털 저장 매체를 이용하여 플로피 디스크, DVD를 행할 수 전자적으로 판독 가능한 제어 신호를 갖는 블루 레이, CD, ROM, PROM, EPROM, EEPROM 또는 FLASH 메모리, 저장 각각의 방법을 수행되도록 프로그램 가능한 컴퓨터 시스템과 협력 (또는 협력 할 수 있는)하여 구현된다. 따라서, 디지털 저장 매체는 판독가능한 컴퓨터 일 수 있다.
- [0052] 본 발명에 따른 일부 실시 예는 본원에 기재된 방법 중 하나가 수행되도록 프로그램 가능한 컴퓨터 시스템과 협력 할 수 있는 전자적으로 판독 가능한 제어 신호를 갖는 데이터 캐리어를 포함한다.
- [0053] 일반적으로, 본 발명의 실시 예는 프로그램 코드를 가진 컴퓨터 프로그램 제품, 컴퓨터 프로그램 제품이 컴퓨터 상에서 실행될 때 방법 중 하나를 수행하기 위한 프로그램 코드 인 동작으로 구현될 수 있다. 프로그램 코드는, 예를 들면 머신 판독 가능 매체 상에 저장 될 수 있다.

- [0054] 다른 실시 예는 기계 판독 가능 캐리어 상에 저장된 본 명세서에 기술 된 방법 중 하나를 수행하기 위한 컴퓨터 프로그램을 포함한다.
- [0055] 즉, 본 발명의 방법의 실시 예는, 따라서, 컴퓨터 프로그램이 컴퓨터상에서 실행될 때, 여기에 설명 된 방법 중 하나를 수행하기 위한 프로그램 코드를 갖는 컴퓨터 프로그램을 포함한다.
- [0056] 본 발명의 방법의 또 다른 실시 예에 따라서, 데이터 캐리어 (또는 디지털 저장 매체, 또는 컴퓨터 판독 가능 매체)를 포함하는 본원에 기재된 방법 중 하나를 수행하기 위한, 컴퓨터 프로그램을 기록한 것이다. 데이터 캐리어는, 디지털 저장 매체 또는 기록 매체는 일반적으로 가시적인 및/또는 비트랜지셔널(transitionary)이다.
- [0057] 본 발명의 방법의 또 다른 실시 예에 따라서, 인, 데이터 스트림 또는 본원에 기재된 방법 중 하나를 수행하기 위한 컴퓨터 프로그램을 나타내는 신호들의 시퀀스, 데이터 스트림 또는 신호 서열은 예를 들어 인터넷을 통해, 예를 들어, 데이터 통신 접속을 통해 전송하도록 구성 될 수 있다.
- [0058] 또 다른 실시 예는, 예를 들면, 컴퓨터, 또는 프로그래머블 로직 디바이스로 구성되거나 본원에 기술된 방법 중 하나를 수행하도록 구성된 처리 수단을 포함한다.
- [0059] 또 다른 실시 예는 컴퓨터가 여기에 설명된 방법 중 하나를 수행하기 위한 컴퓨터 프로그램 설치를 포함한다.
- [0060] 본 발명에 따른 또 다른 실시 예는 장치 또는 수신기에 본원에 기재된 방법 중 하나를 수행하기 위한 (전자적 또는 광학적으로, 예를 들어) 컴퓨터 프로그램을 전송하도록 구성되는 시스템을 포함한다. 수신기는, 예를 들면, 컴퓨터, 모바일 장치, 메모리 장치 등일 수 있다. 장치 또는 시스템, 예를 들어, 수신기에 컴퓨터 프로그램을 전송하는 파일 서버를 포함할 수 있다.
- [0061] 일부 실시 예에서, (예컨대 필드 프로그래머블 게이트 어레이) 프로그래머블 논리 디바이스는 여기에 설명된 방법의 기능의 일부 또는 전부를 수행하기 위해 사용될 수 있다. 일부 실시 예에서, 필드 프로그래머블 게이트 어레이는 본원에 기재된 방법 중 하나를 수행하기 위해 마이크로 프로세서와 협력 할 수 있다. 일반적으로, 방법은 바람직하게는 임의의 하드웨어 장치에 의해 수행된다.
- [0062] 본원에 기술 된 장치는 하드웨어 장치를 이용하여, 또는 컴퓨터를 사용하여, 또는 하드웨어 장치 및 컴퓨터의 조합을 사용하여 구현 될 수 있다.
- [0063] 여기에 설명된 방법은 하드웨어 장치를 이용하여, 또는 컴퓨터를 사용하여, 또는 하드웨어 장치 및 컴퓨터의 조합을 이용하여 수행 될 수 있다.
- [0064] 상술 한 실시 예들은 본 발명의 원리에 대한 예시에 불과하다. 이는 본원에 기재된 변형 및 배치의 변형 및 상세 다른 당업자에게 명백 할 것이라는 점을 알 수있다. 따라서, 단지 임박한 특허 청구 범위에 의해서만 아니라 본원의 실시 예에 대한 설명 및 설명에 의해 제시된 특정 세부 사항에 의해 제한되는 의도이다.
- [0065] 비특허문헌(List of cited non-patent literature)
- [0066] [1] B. Bessette et al., "The Adaptive Multi-rate Wideband Speech Codec (AMR-WB)," IEEE Trans. On Speech and Audio Processing, Vol. 10, No. 8, Nov. 2002.
- [0067] [2] R. C. Hendriks, R. Heusdens and J. Jensen, "MMSE based noise PSD tracking with low complexity," in IEEE Int. Conf. Acoust., Speech, Signal Processing, pp. 4266 - 4269, March 2010.

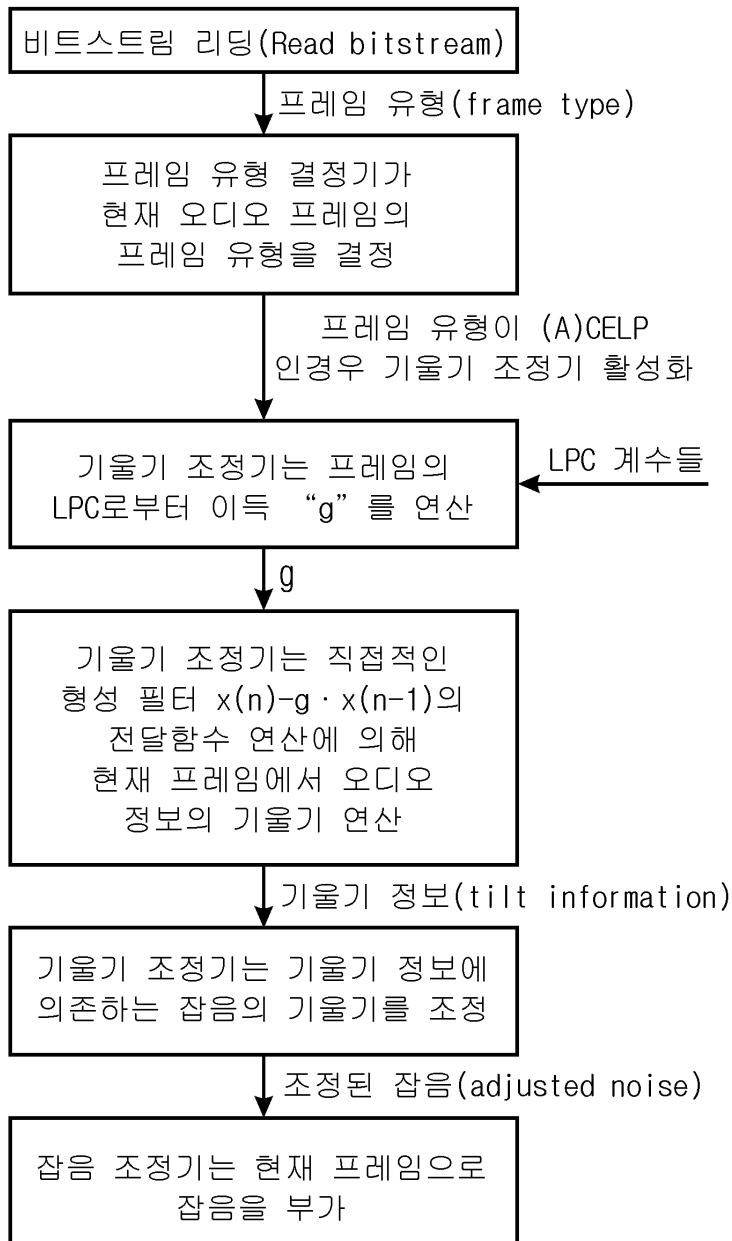
- [0068] [3] R. Martin, "Noise Power Spectral Density Estimation Based on Optimal Smoothing and Minimum Statistics," IEEE Trans. On Speech and Audio Processing, Vol. 9, No. 5, Jul. 2001.
- [0069] [4] M. Jelinek and R. Salami, "Wideband Speech Coding Advances in VMR-WB Standard," IEEE Trans. On Audio, Speech, and Language Processing, Vol. 15, No. 4, May 2007.
- [0070] [5] J. Makinen et al., "AMR-WB+: A New Audio Coding Standard for 3rd Generation Mobile Audio Services," in Proc. ICASSP 2005, Philadelphia, USA, Mar. 2005.
- [0071] [6] M. Neuendorf et al., "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types," in Proc. 132nd AES Convention, Budapest, Hungary, Apr. 2012. Also appears in the Journal of the AES, 2013.
- [0072] [7] T. Vaillancourt et al., "ITU-T EV-VBR: A Robust 8 -32 kbit/s Scalable Coder for Error Prone Telecommunications Channels," in Proc. EUSIPCO 2008, Lausanne, Switzerland, Aug. 2008.
- [0073]

도면

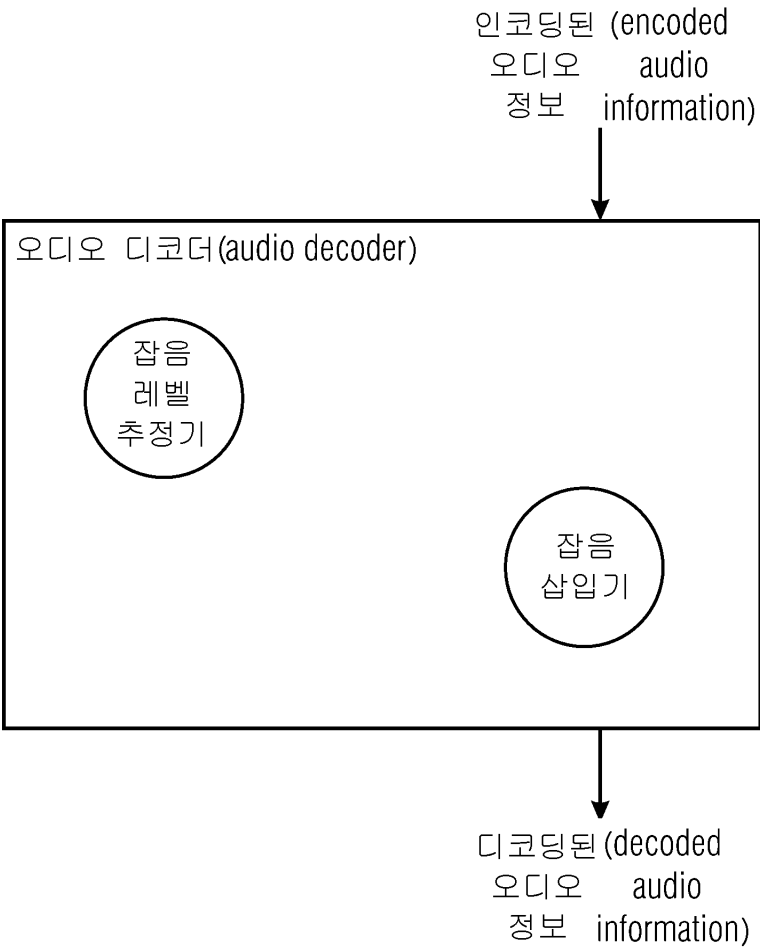
도면1



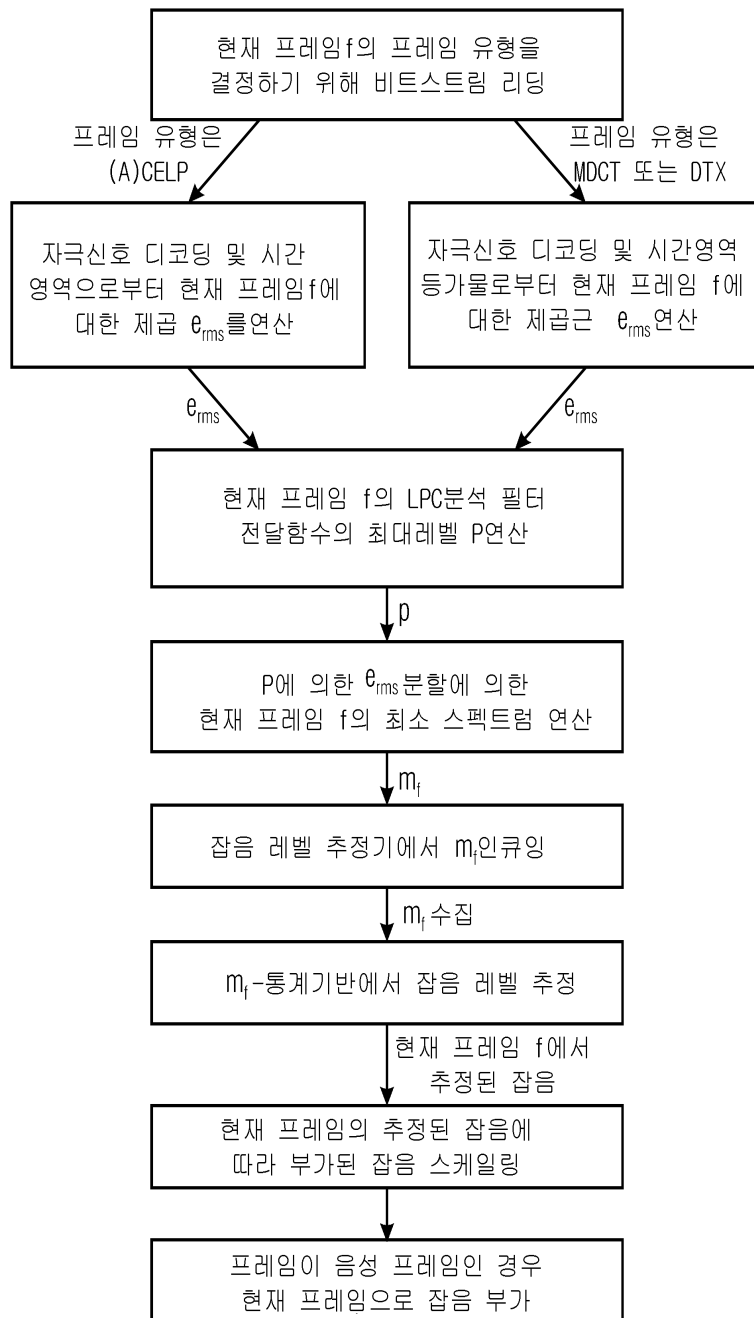
도면2



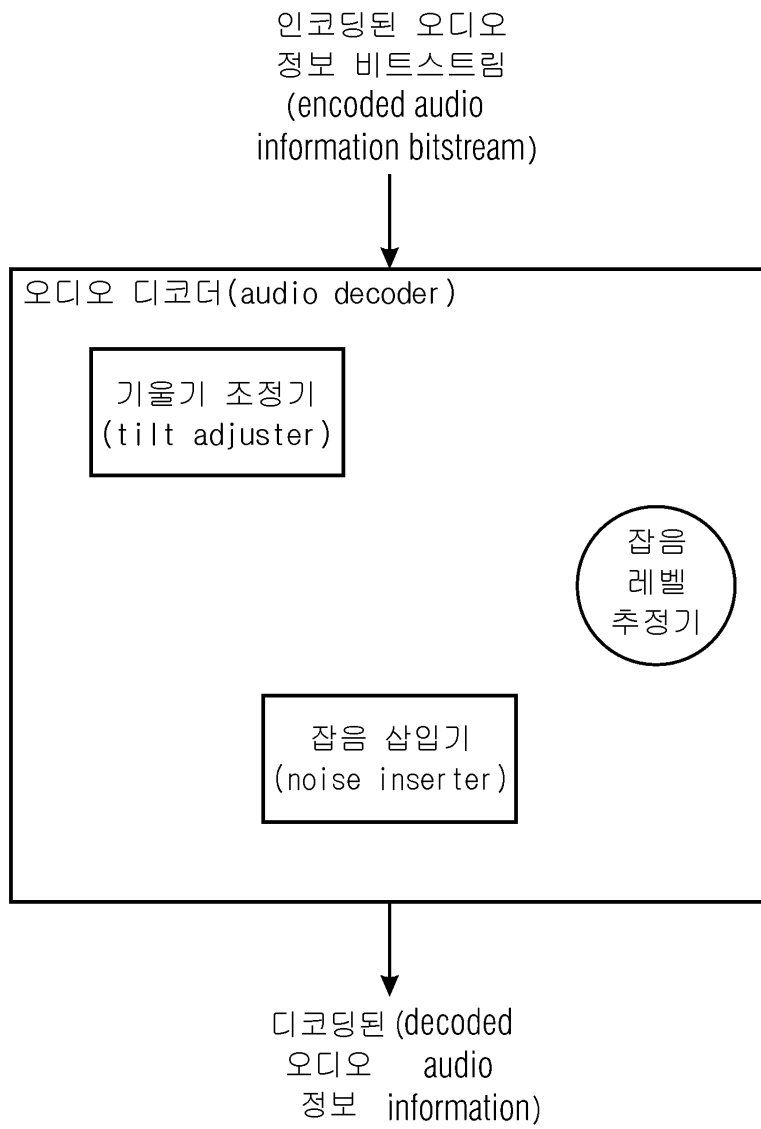
도면3



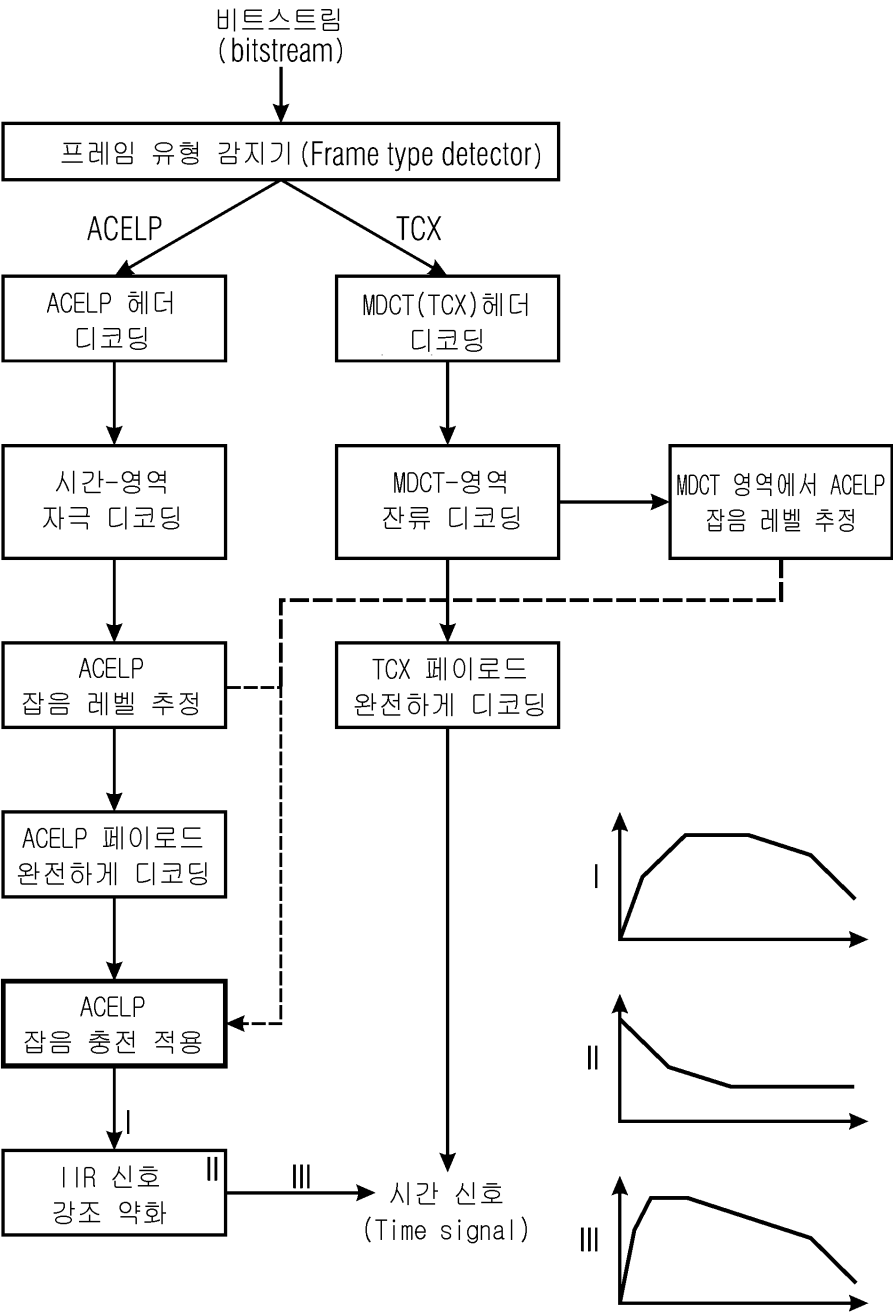
도면4



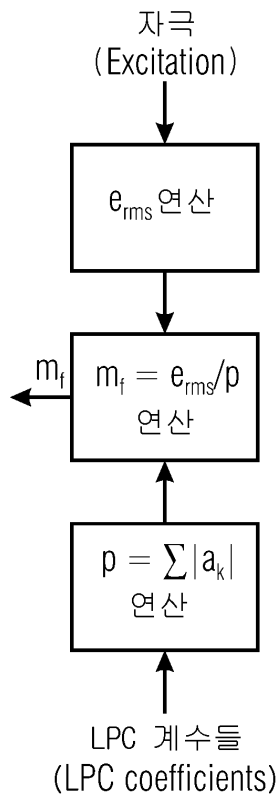
도면5



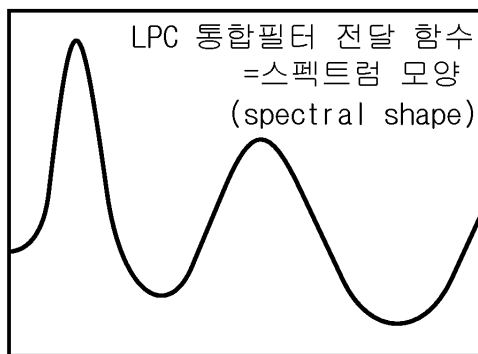
도면6



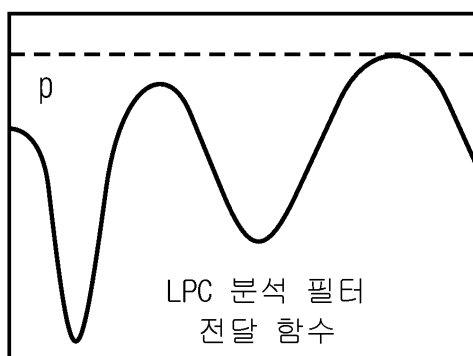
도면7a



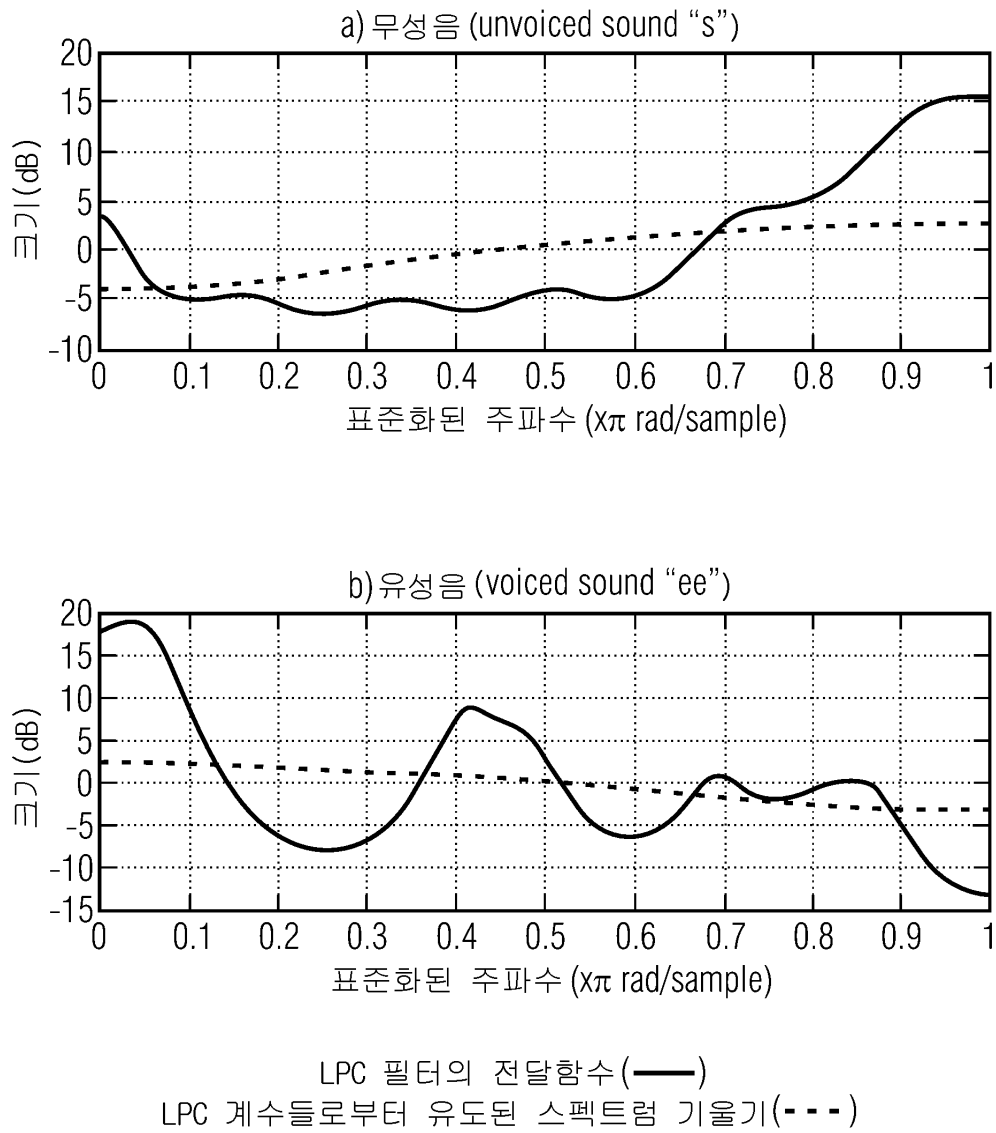
도면7b



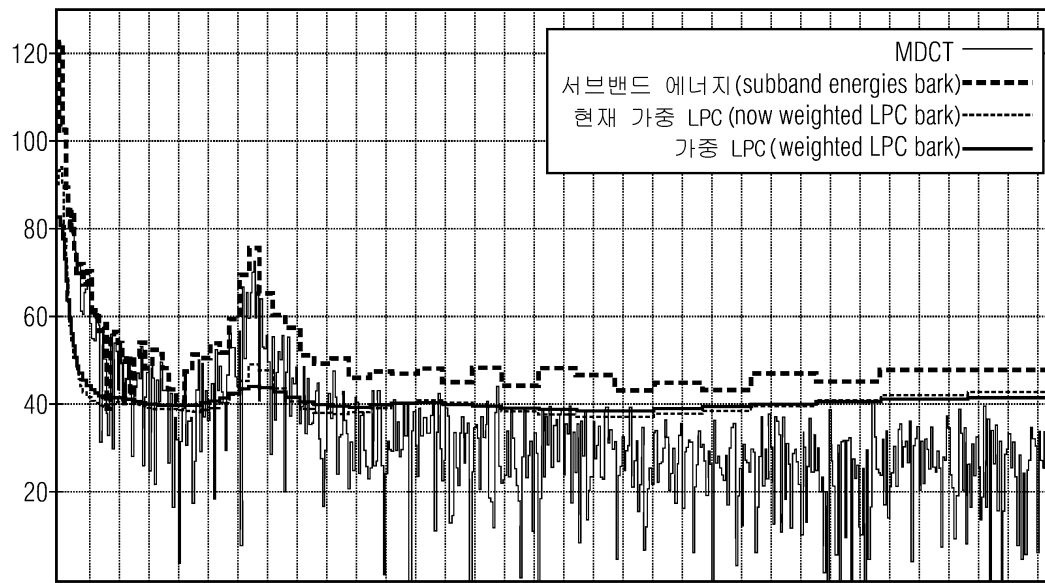
도면7c



도면8



도면9



MDCT 파워스펙트럼으로부터의 LPC(MDCT 값의 제공)