



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2014년06월13일
(11) 등록번호 10-1407120
(24) 등록일자 2014년06월05일

- (51) 국제특허분류(Int. Cl.)
G10L 19/02 (2006.01) *G10L 21/04* (2006.01)
- (21) 출원번호 10-2013-7010454
- (22) 출원일자(국제) 2011년10월04일
심사청구일자 2013년04월24일
- (85) 번역문제출일자 2013년04월24일
- (65) 공개번호 10-2013-0069821
- (43) 공개일자 2013년06월26일
- (86) 국제출원번호 PCT/EP2011/067318
- (87) 국제공개번호 WO 2012/045744
국제공개일자 2012년04월12일
- (30) 우선권주장
61/390,267 2010년10월06일 미국(US)
- (56) 선행기술조사문현
US20100153122 A1

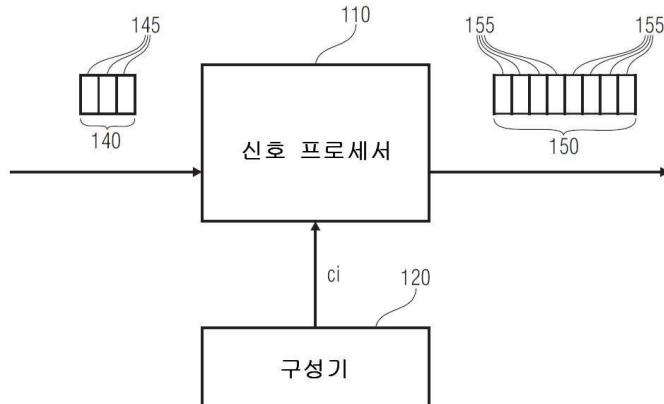
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 오디오 신호를 처리하고 결합된 통합형 음성 및 오디오 코덱 (U S A C) 을 위한 보다 높은 시간적 입도를 제공하기 위한 장치 및 방법

(57) 요 약

오디오 신호를 처리하는 장치가 제공된다. 장치는 신호 프로세서(110; 205; 405) 및 구성기(120; 208; 408)를 포함한다. 신호 프로세서(110; 205; 405)는 오디오 신호의 제1 구성가능한 갯수의 샘플들을 갖는 제1 오디오 신호 프레임을 수신하도록 적응된다. 게다가, 신호 프로세서(110; 205; 405)는 오디오 신호를 구성가능한 업샘플링 계수만큼 업샘플링하여, 처리된 오디오 신호를 획득하도록 적응된다. 더 나아가, 신호 프로세서(110; 205; 405)는 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들을 갖는 제2 오디오 신호 프레임을 출력하도록 적응된다. 구성기(120; 208; 408)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 구성가능한 업샘플링 계수가 제1 업샘플링 값과 동일하도록 하는 구성 정보에 기초하여 신호 프로세서(110; 205; 405)를 구성하도록 적응된다. 또한, 구성기(120; 208; 408)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 업샘플링 계수가 상이한 제2 업샘플링 값과 동일하도록, 신호 프로세서(110; 205; 405)를 구성하도록 적응된다. 제1 또는 제2 비율값은 정수값이 아니다.

대 표 도 - 도1



(72) 발명자

레텔바흐 니콜라우스

독일 뉘른베르그 90427 스페사르트스트라쎄 38

푸쉬 켈라움

독일 에를랑겐 91058 퀴르터 스트라쎄 17

노이엔도르프 막스

독일 뉘른베르크 90459 파라디에스트라쎄 20

베세테 브루노

캐나다 퀘벡 셔브룩 제이1엔 4지5 루 무리요 1600

르페브르 로흐

캐나다 퀘벡 마곡 제이1엑스 0엘6 루 드 라 부르가
드 259

구르네 필립

캐나다 퀘벡 셔브룩 제이1엘 0에이2 루 드 쏘비뇽
3012

빌데 스텔판

독일 뉘른베르크 90419 잔드라르트스트라쎄 4

특허청구의 범위

청구항 1

오디오 신호를 처리하기 위한 장치에 있어서,

오디오 신호의 제1 구성가능한 갯수의 샘플들을 갖는 제1 오디오 신호 프레임을 수신하도록 적응되고, 상기 오디오 신호를 구성가능한 업샘플링 계수만큼 업샘플링하여 처리된 오디오 신호를 획득하도록 적응되며, 상기 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들을 갖는 제2 오디오 신호 프레임을 출력하도록 적응된, 신호 프로세서(110; 205; 405); 및

상기 신호 프로세서(110; 205; 405)를 구성하도록 적응된 구성기(120; 208; 408)

를 포함하고,

상기 구성기(120; 208; 408)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 제1 업샘플링값과 동일하도록 하는 구성 정보에 기초하여 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응되고,

상기 구성기(120; 208; 408)는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 상이한 제2 업샘플링값과 동일하도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응되며,

상기 제1 비율값 또는 상기 제2 비율값은 정수값이 아닌 것인, 오디오 신호 처리 장치.

청구항 2

제1항에 있어서, 상기 구성기(120; 208; 408)는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상기 제1 비율이 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상기 제1 비율보다 클 때, 상기 상이한 제2 업샘플링값이 상기 제1 업샘플링값보다 크도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 3

제1항에 있어서,

상기 구성기(120; 208; 408)는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상기 제1 비율이 상기 제1 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 상기 제1 비율값과 동일하도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응되며,

상기 구성기(120; 208; 408)는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상기 제2 비율이 상기 상이한 제2 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 상기 상이한 제2 비율값과 동일하도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 4

제1항에 있어서,

상기 구성기(120; 208; 408)는, 상기 제1 비율이 상기 제1 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 2와 동일하도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응되며,

상기 구성기(120; 208; 408)는, 상기 제2 비율이 상기 상이한 제2 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 8/3과 동일하도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 5

제1항에 있어서,

상기 구성기(120; 208; 408)는, 상기 제1 비율이 상기 제1 비율값을 가질 때, 상기 제1 구성가능한 샘플들의 갯

수가 1024와 동일하고 상기 제2 구성가능한 샘플들의 갯수가 2048과 동일하도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응되며,

상기 구성기(120; 208; 408)는, 상기 제2 비율이 상기 상이한 제2 비율값을 가질 때, 상기 제1 구성가능한 샘플들의 갯수가 768과 동일하고 상기 제2 구성가능한 샘플들의 갯수가 2048과 동일하도록, 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 6

제1항에 있어서, 상기 신호 프로세서(110; 205; 405)는,

상기 오디오 신호를 디코딩하여 사전처리된 오디오 신호를 획득하기 위한 코어 디코더 모듈(210),

제1 사전처리된 오디오 신호를 시간 도메인에서 주파수 도메인으로 변환시켜서 복수의 서브대역 신호들을 포함한 사전처리된 주파수 도메인 오디오 신호를 획득하기 위한 복수의 분석 필터 뱅크 채널들을 갖는 분석 필터 뱅크(220),

상기 사전처리된 주파수 도메인 오디오 신호를 위한 추가적인 서브대역 신호들을 생성하고 추가하기 위한 서브대역 발생기(230), 및

상기 제1 사전처리된 오디오 신호를 주파수 도메인에서 시간 도메인으로 변환시켜서 상기 처리된 오디오 신호를 획득하기 위한 복수의 합성 필터 뱅크 채널들을 갖는 합성 필터 뱅크(240)

를 포함하며,

상기 구성기(120; 208; 408)는 상기 구성가능한 업샘플링 계수가 상기 분석 필터 뱅크 채널들의 갯수에 대한 상기 합성 필터 뱅크 채널들의 갯수의 제3 비율과 동일하도록, 상기 합성 필터 뱅크 채널들의 갯수 또는 상기 분석 필터 뱅크 채널들의 갯수를 구성시킴으로써 상기 신호 프로세서(110; 205; 405)를 구성시키도록 적응된 것인, 오디오 신호 처리 장치.

청구항 7

제6항에 있어서, 상기 서브대역 발생기(230)는 상기 사전처리된 주파수 도메인 오디오 신호를 위한 상기 추가적인 서브대역 신호들을 생성하기 위해 사전처리된 오디오 신호 발생기의 서브대역 신호들을 복제하도록 적응된 스펙트럼 대역 복제기인 것인, 오디오 신호 처리 장치.

청구항 8

제6항에 있어서, 상기 신호 프로세서(110; 205; 405)는 상기 사전처리된 오디오 신호를 디코딩하여 스테레오 또는 서라운드 채널들을 포함한 사전처리된 오디오 신호들을 획득하기 위한 MPEG 서라운드 디코더(410)를 더 포함하며,

상기 서브대역 발생기(230)는, 상기 사전처리된 주파수 도메인 오디오 신호를 위한 상기 추가적인 서브대역 신호들이 생성되어 상기 사전처리된 주파수 도메인 오디오 신호에 추가된 후 상기 사전처리된 주파수 도메인 오디오 신호를 상기 MPEG 서라운드 디코더(410)에 제공하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 9

제6항에 있어서, 상기 코어 디코더 모듈(210)은 제1 코어 디코더(510)와 제2 코어 디코더(520)를 포함하고, 상기 제1 코어 디코더(510)는 시간 도메인에서 동작하도록 적응되고, 상기 제2 코어 디코더(520)는 주파수 도메인에서 동작하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 10

제9항에 있어서, 상기 제1 코어 디코더(510)는 ACELP 디코더이고, 상기 제2 코어 디코더(520)는 FD 변환 디코더 또는 TCX 변환 디코더인 것인, 오디오 신호 처리 장치.

청구항 11

제10항에 있어서, 상기 ACELP 디코더(510)는 상기 제1 오디오 신호 프레임을 처리하도록 적응되고, 상기 제1 오디오 신호 프레임은 4개의 ACELP 프레임들을 가지며, 상기 제1 오디오 신호 프레임의 상기 제1 구성가능한 샘플

들의 갯수가 768과 동일할 때, 상기 ACELP 프레임들 각각은 192개의 오디오 신호 샘플들을 갖는 것인, 오디오 신호 처리 장치.

청구항 12

제10항에 있어서, 상기 ACELP 디코더(510)는 상기 제1 오디오 신호 프레임을 처리하도록 적응되고, 상기 제1 오디오 신호 프레임은 3개의 ACELP 프레임들을 가지며, 상기 제1 오디오 신호 프레임의 상기 제1 구성가능한 샘플들의 갯수가 768과 동일할 때, 상기 ACELP 프레임들 각각은 256개의 오디오 신호 샘플들을 갖는 것인, 오디오 신호 처리 장치.

청구항 13

제1항에 있어서, 상기 구성기(120; 208; 408)는, 상기 오디오 신호의 상기 제1 구성가능한 샘플들의 갯수 또는 상기 처리된 오디오 신호의 상기 제2 구성가능한 샘플들의 갯수 중 적어도 하나를 표시하는 구성 정보에 기초하여 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 14

제1항에 있어서, 상기 구성기(120; 208; 408)는 상기 구성 정보에 기초하여 상기 신호 프로세서(110; 205; 405)를 구성하도록 적응되고, 상기 구성 정보는 상기 오디오 신호의 상기 제1 구성가능한 샘플들의 갯수 및 상기 처리된 오디오 신호의 상기 제2 구성가능한 샘플들의 갯수를 표시하며, 상기 구성 정보는 구성 인덱스인 것인, 오디오 신호 처리 장치.

청구항 15

오디오 신호를 처리하기 위한 방법에 있어서,

구성가능한 업샘플링 계수를 구성하는 단계;

오디오 신호의 제1 구성가능한 갯수의 샘플들을 갖는 제1 오디오 신호 프레임을 수신하는 단계; 및

상기 오디오 신호를 상기 구성가능한 업샘플링 계수만큼 업샘플링하여 처리된 오디오 신호를 획득하고, 상기 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들을 갖는 제2 오디오 프레임을 출력하도록 적응되는 단계를 포함하며,

상기 구성가능한 업샘플링 계수는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 제1 업샘플링값과 동일하도록 하는 구성 정보에 기초하여 구성되고,

상기 구성가능한 업샘플링 계수는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 상기 구성가능한 업샘플링 계수가 상이한 제2 업샘플링값과 동일하도록 구성되며,

상기 제1 비율값 또는 상기 제2 비율값은 정수값이 아닌 것인, 오디오 신호 처리 방법.

청구항 16

오디오 신호를 처리하기 위한 장치에 있어서,

오디오 신호의 제1 구성가능한 갯수의 샘플들을 갖는 제1 오디오 신호 프레임을 수신하도록 적응되고, 상기 오디오 신호를 구성가능한 다운샘플링 계수만큼 다운샘플링하여 처리된 오디오 신호를 획득하도록 적응되며, 상기 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들을 갖는 제2 오디오 프레임을 출력하도록 적응된 신호 프로세서(910); 및

상기 신호 프로세서를 구성하도록 적응된 구성기(920)

를 포함하며,

상기 구성기(920)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 상기 구성가능한 다운샘플링 계수가 제1 다운샘플링값과 동일하도록 하는 구성 정보에 기초하

여 상기 신호 프로세서(910)를 구성하도록 적응되고,

상기 구성기(920)는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 상기 구성가능한 다운샘플링 계수가 상이한 제2 다운샘플링 값과 동일하도록, 상기 신호 프로세서(910)를 구성하도록 적응되며,

상기 제1 비율값 또는 상기 제2 비율값은 정수값이 아닌 것인, 오디오 신호 처리 장치.

청구항 17

제16항에 있어서, 상기 구성기는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상기 제1 비율이 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상기 제2 비율보다 작을 때, 상기 제1 다운샘플링 값이 상기 상이한 제2 다운샘플링 값보다 작도록, 상기 신호 프로세서(910)를 구성하도록 적응된 것인, 오디오 신호 처리 장치.

청구항 18

오디오 신호를 처리하기 위한 방법에 있어서,

구성가능한 다운샘플링 계수를 구성하는 단계;

오디오 신호의 제1 구성가능한 갯수의 샘플들을 갖는 제1 오디오 신호 프레임을 수신하는 단계; 및

상기 오디오 신호를 상기 구성가능한 다운샘플링 계수만큼 다운샘플링하여 처리된 오디오 신호를 획득하고, 상기 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들을 갖는 제2 오디오 프레임을 출력하도록 적응되는 단계를 포함하며,

상기 구성가능한 다운샘플링 계수는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 상기 구성가능한 다운샘플링 계수가 제1 다운샘플링 값과 동일하도록 하는 구성정보에 기초하여 구성되고,

상기 구성가능한 다운샘플링 계수는, 상기 제1 구성가능한 샘플들의 갯수에 대한 상기 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 상기 구성가능한 다운샘플링 계수가 상이한 제2 다운샘플링 값과 동일하도록 구성되며,

상기 제1 비율값 또는 상기 제2 비율값은 정수값이 아닌 것인, 오디오 신호 처리 방법.

청구항 19

컴퓨터 프로그램이 컴퓨터 또는 프로세서에 의해 실행될 시에 제15항 또는 제18항의 방법을 수행하기 위한 컴퓨터 프로그램이 저장된 컴퓨터로 읽을 수 있는 매체.

명세서

기술 분야

[0001]

본 발명은 오디오 처리에 관한 것이며, 특히 오디오 신호를 처리하고 결합된 통합형 음성 및 오디오 코덱(USAC)을 위한 보다 높은 시간적 입도(temporal granularity)를 제공하기 위한 장치 및 방법에 관한 것이다.

배경 기술

[0002]

다른 오디오 코덱들과 같이, USAC는 고정된 프레임 크기(USAC: 2048개 샘플들/프레임)를 나타낸다. 하나의 프레임 내에서 보다 짧은 변환 크기들의 제한된 세트로 스위칭할 가능성이 존재하지만, 프레임 크기는 완전 시스템의 시간적 해상도를 여전히 제한시킨다. 완전 시스템의 시간적 입도를 증가시키기 위해, 통상적인 오디오 코덱들의 경우 샘플링 레이트는 증가되고, 이것은 시간적으로 하나의 프레임의 보다 짧은 지속기간(예컨대, 밀리초)을 야기시킨다. 하지만, 이것은 USAC 코덱에 대해서는 쉽게 가능하지가 않다.

[0003]

USAC 코덱은 ACELP(ACELP = Algebraic Code Excited Linear Prediction)와 같은, 통상적인 음성 코더들로부터의 틀들에 더하여, AAC(Advanced Audio Coding) 변환 코더, SBR(Spectral Band Replication) 및 MPEG(MPEG = Moving Picture Experts Group) 서라운드와 같은, 통상적인 일반적 오디오 코덱들로부터의 틀들의 조합을 포함

한다. ACELP와 변환 코더 모두는 보통 동일한 환경(즉, 프레임 크기, 샘플링 레이트) 내에서 동시에 구동되고, 손쉽게 스위칭될 수 있는데, 보통, 클린 음성 신호들의 경우에서는, ACELP 툴이 이용되고, 음악, 혼합 신호들의 경우에서는, 변환 코더가 이용된다.

발명의 내용

해결하려는 과제

[0004] 동시에 ACELP 툴은 상대적으로 낮은 샘플링 레이트들에서만 동작하도록 제한된다. 24kbit/s의 경우, 단지 17075Hz의 샘플링 레이트만이 이용된다. 보다 높은 샘플링 레이트들의 경우에서, ACELP 툴은 성능이 상당히 떨어지기 시작한다. 하지만 SBR 및 MPEG 서라운드뿐만이 아니라 변환 코더는 예컨대, 변환 코더에 대해서는 22050Hz, SBR 및 MPEG 서라운드에 대해서는 44100Hz와 같이, 훨씬 더 높은 샘플링 레이트로부터 이로울 것이다. 하지만, 지금까지, ACELP 툴은 완전 시스템의 샘플링 레이트를 제한시켰으며, 이것은 특히 음악 신호들에 대한 차선적인 시스템을 야기시킨다.

과제의 해결 수단

[0005] 본 발명의 목적은 오디오 신호를 처리하기 위한 방법 및 장치에 대한 개선된 개념들을 제공하는 것이다. 본 발명의 목적은 청구항 제1항에 따른 장치, 청구항 제15항에 따른 방법, 청구항 제16항에 따른 장치, 청구항 제18항에 따른 방법 및 청구항 제19항에 따른 컴퓨터 프로그램에 의해 해결된다.

[0006] 현재의 USAC RM은 8kbit/s와 같은 매우 낮은 비트레이트들로부터 128kbit/s 이상의 비트레이트들에서의 명료한 퀄리티까지 이르는 범위에서, 방대한 수의 동작점들에 대해 높은 코딩 성능을 제공한다. 이러한 광범위한 비트레이트들에 대한 이러한 높은 퀄리티에 도달하기 위해, MPEG 서라운드, SBR, ACELP 및 통상적인 변환 코더들과 같은, 툴들의 조합이 이용된다. 물론 이러한 툴들의 조합은 툴 상호연동 및 이러한 툴들이 배치된 공통 환경의 연합 최적화 처리를 필요로 한다.

[0007] 이러한 연합 최적화 처리에서 몇몇의 툴들은 중간 비트레이트 범위(24kbit/s ~ 32kbit/s)에서의 높은 시간적 구조를 노출시키는 신호 재생 결합들을 갖는다는 것을 발견하였다. 특히 MPEG 서라운드, SBR 및 FD 변환 코더들(FD, TCX)(FD = Frequency Domain; TCX = Transform Coded Excitation) 툴들, 즉 주파수 도메인에서 동작하는 모든 툴들은 시간 도메인에서 보다 짧은 프레임 크기와 동일한, 보다 높은 시간적 입도를 갖고 동작될 때 보다 우수하게 수행할 수 있다.

[0008] 최신 기술의 HE-AACv2 인코더(High-Efficiency AAC v2 인코더)와 비교하여, 현재의 USAC 참조 퀄리티 인코더는 동일한 프레임 크기(샘플들)를 이용하면서, 보다 상당히 낮은 샘플링 레이트에서 24 kbit/s 및 32 kbit/s와 같은 비트레이트들로 동작한다는 것을 발견하였다. 이것은 수밀리초의 프레임들의 지속기간이 상당히 길다는 것을 의미한다. 이러한 결합들을 보상하기 위해, 시간적 입도는 증가될 필요가 있다. 이것은 (예컨대 고정 프레임 크기를 이용한 시스템들의) 샘플링 주파수를 증가시키거나 또는 프레임 크기들을 단축시킴으로써 달성될 수 있다.

[0009] 시간적 동적 신호들에 대한 성능을 증가시키기 위해 샘플링 주파수를 증가시키는 것이 SBR 및 MPEG 서라운드에 대한 합리적인 방법이지만, 이것은 모든 코어 코더 툴들에 대해 유용되는 것은 아닐 것이다: 보다 높은 샘플링 주파수는 변환 코더에 유리할 것이지만, 동시에 ACELP 툴의 성능을 급격하게 감소시킨다는 것은 잘 알려져 있다.

[0010] 오디오 신호를 처리하는 장치가 제공된다. 장치는 신호 프로세서 및 구성기를 포함한다. 신호 프로세서는 오디오 신호의 제1 구성가능한 갯수의 샘플들을 갖는 제1 오디오 신호 프레임을 수신하도록 적응된다. 게다가, 신호 프로세서는 오디오 신호를 구성가능한 업샘플링 계수만큼 업샘플링하여, 처리된 오디오 신호를 획득하도록 적응된다. 더 나아가, 신호 프로세서는 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들을 갖는 제2 오디오 신호 프레임을 출력하도록 적응된다.

[0011] 구성기는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 구성가능한 업샘플링 계수가 제1 업샘플링값과 동일하도록 하는 구성 정보에 기초하여 신호 프로세서를 구성하도록 적응된다. 게다가, 구성기는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 업샘플링 계수가 상이한 제2 업샘플링값과 동일하도록, 신호 프로세서를 구성하도록 적응된다. 제1 또는 제2 비율값은 정수값이 아니다.

[0012] 앞서 설명한 실시예에 따르면, 신호 프로세서는 오디오 신호를 업샘플링하여 업샘플링 처리된 오디오 신호를 획

득한다. 위 실시예에서, 업샘플링 계수는 구성가능하며, 이것은 비정수값일 수 있다. 업샘플링 계수가 비정수값일 수 있다라는 사실 및 그 구성가능성은 장치의 유연성을 증가시킨다. 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 갖는 경우, 구성가능한 업샘플링 계수는 상이한 제2 업샘플링값을 갖는다. 따라서, 장치는 제1 및 제2 오디오 신호 프레임의 프레임 길이(즉, 샘플들의 갯수)의 비율과 업샘플링 계수 사이의 관계를 고려하도록 적응된다.

[0013] 실시예에서, 구성기는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제2 비율이 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율보다 클 때, 상이한 제2 업샘플링값이 제1 업샘플링값보다 크도록, 신호 프로세서를 구성하도록 적응된다.

[0014] 실시예에 따르면, USAC 코덱에 대한 새로운 동작 모드(이하에서는 "추가 세팅"이라고 칭한다)가 제안되며, 이 새로운 동작 모드는 24 kbit/s 및 32 kbit/s와 같은 중간 데이터 레이트들에 대한 시스템 성능을 증대시킨다. 이러한 동작점들에 대해, 현재의 USAC 참조 코덱의 시간적 해상도는 너무 낮다는 것이 발견되었다. 그러므로, a) 코더 코더에 대한 샘플링 레이트를 증가시키지 않고서 코어 코더 프레임 크기들을 단축시킴으로써 이러한 시간적 해상도를 증가시키는 것과, 추가로 b) 이러한 툴들에 대한 프레임 크기를 변경시키지 않고서 SBR 및 MPEG 서라운드에 대한 샘플링 레이트를 증가시키는 것이 제안된다.

[0015] 제안된 추가 세팅은 ACELP 툴을 포함한 시스템이 44.1 및 48 kHz와 같은 보다 높은 샘플링 레이트들에서 동작되도록 해주기 때문에, 시스템의 유연성을 많이 향상시킨다. 이러한 샘플링 레이트들은 일반적으로 시장에서 요청되는 것들이기 때문에, 이것은 USAC 코덱의 수용에 도움을 줄 것이라고 예상된다.

[0016] 현재의 MPEG 통합형 음성 및 오디오 코딩(USAC) 작업 아이템에 대한 새로운 동작 모드는 완전한 오디오 코덱의 시간적 입도를 증가시킴으로써 전체 코덱의 시간적 유연성을 증가시킨다. (샘플들의 제2 갯수가 동일하게 남아 있는 것을 가정하여) 제2 비율이 제1 비율보다 크다면, 제1 구성가능한 샘플들의 갯수는 감소되었고, 즉 제1 오디오 신호 프레임의 프레임 크기는 단축되었다. 이것은 보다 높은 시간적 입도를 야기시키며, 주파수 도메인에서 동작하고 제1 오디오 신호 프레임을 처리하는 모든 툴들은 보다 우수하게 수행될 수 있다. 하지만, 이러한 고효율적인 동작 모드에서는, 업샘플링된 오디오 신호를 포함한 제2 오디오 신호 프레임을 처리하는 툴들의 성능을 증가시키는 것이 또한 바람직하다. 이와 같이 이러한 툴들의 성능 증가는 업샘플링된 오디오 신호의 보다 높은 샘플링 레이트에 의해, 즉 이러한 동작 모드에 대한 업샘플링 계수를 증가시킴으로써 실현될 수 있다. 게다가, 주파수 도메인에서 동작하지 않으며, 제1 오디오 신호 프레임을 처리하고 (원래의) 오디오 신호의 샘플링 레이트가 비교적 낮을 때 최상으로 동작하는, USAC에서의 ACELP 디코더와 같은 툴들이 존재한다. 이러한 툴들은 높은 업샘플링 계수로부터 이익을 얻는데, 그 이유는 이것은 업샘플링된 오디오 신호의 샘플링 레이트에 비해 (원래의) 오디오 신호의 샘플링 레이트가 비교적 낮다는 것을 의미하기 때문이다. 상술한 실시예는 이러한 환경에 대한 효율적인 동작 모드를 위한 구성 모드를 제공하도록 적응된 장치를 제공한다.

[0017] 새로운 동작 모드는 완전한 오디오 코덱의 시간적 입도를 증가시킴으로써 전체 코덱의 시간적 유연성을 증가시킨다.

[0018] 실시예에서, 구성기는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 구성가능한 업샘플링 계수가 제1 비율값과 동일하도록, 신호 프로세서를 구성하도록 적응되며, 구성기는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 업샘플링 계수가 상이한 제2 비율값과 동일하도록, 신호 프로세서를 구성하도록 적응된다.

[0019] 실시예에서, 구성기는, 제1 비율이 제1 비율값을 가질 때, 구성가능한 업샘플링 계수가 2와 동일하도록, 신호 프로세서를 구성하도록 적응되며, 구성기는, 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 업샘플링 계수가 8/3과 동일하도록, 신호 프로세서를 구성하도록 적응된다.

[0020] 추가적인 실시예에 따르면, 구성기는, 제1 비율이 제1 비율값을 가질 때, 제1 구성가능한 샘플들의 갯수가 1024와 동일하고 제2 구성가능한 샘플들의 갯수가 2048과 동일하도록, 신호 프로세서를 구성하도록 적응되며, 구성기는, 제2 비율이 상이한 제2 비율값을 가질 때, 제1 구성가능한 샘플들의 갯수가 768과 동일하고 제2 구성가능한 샘플들의 갯수가 2048과 동일하도록, 신호 프로세서를 구성하도록 적응된다.

[0021] 실시예에서, 코어 코더가 보다 짧은 프레임 크기(1024개 대신에 768개 샘플들)에서 동작하는 추가적인 USAC 코더 세팅을 도입시키는 것이 제안된다. 더 나아가, SBR 및 MPEG 서라운드가 보다 높은 샘플링 레이트에서 동작되도록 하기 위해 본 환경에서 SBR 디코더 내부의 리샘플링을 2:1로부터 8:3으로 수정하는 것이 제안된다.

- [0022] 더 나아가, 실시예에 따르면, 코어 코더의 시간적 입도는 코어 코더 프레임 크기를 1024개의 샘플들로부터 768 개의 샘플들로 감축시킴으로써 증가된다. 이 단계에 의해, 코어 코더의 시간적 입도는 샘플링 레이트를 일정하게 유지하면서 4/3배 증가된다. 이것은 ACELP가 적절한 샘플링 주파수(Fs)에서 구동되도록 해준다.
- [0023] 게다가, SBR 툴에서, 8/3 비율(지금까지는 2 비율)의 리샘플링이 적용되어, 3/8 Fs에서의 768 크기의 코어 코더 프레임을 Fs에서의 2048 크기의 출력 프레임으로 변환시킨다. 이것은 SBR 툴과 MPEG 서라운드 툴이 통상적으로 높은 샘플링 레이트(예컨대, 44100 Hz)에서 구동되도록 해준다. 따라서, 모든 툴들이 각자의 최적의 동작점에서 구동되기 때문에, 우수한 웰리티의 음성 및 음악 신호들이 제공된다.
- [0024] 실시예에서, 신호 프로세서는, 오디오 신호를 디코딩하여 사전처리된 오디오 신호를 획득하기 위한 코어 디코더 모듈, 제1 사전처리된 오디오 신호를 시간 도메인에서 주파수 도메인으로 변환시켜서 복수의 서브대역 신호들을 포함한 사전처리된 주파수 도메인 오디오 신호를 획득하기 위한 복수의 분석 필터 뱅크 채널들을 갖는 분석 필터 뱅크, 사전처리된 주파수 도메인 오디오 신호를 위한 추가적인 서브대역 신호들을 생성하고 추가하기 위한 서브대역 발생기, 및 제1 사전처리된 오디오 신호를 주파수 도메인에서 시간 도메인으로 변환시켜서 처리된 오디오 신호를 획득하기 위한 복수의 합성 필터 뱅크 채널들을 갖는 합성 필터 뱅크를 포함한다. 구성기는 구성 가능한 업샘플링 계수가 분석 필터 뱅크 채널들의 갯수에 대한 합성 필터 뱅크 채널들의 갯수의 제3 비율과 동일하도록, 합성 필터 뱅크 채널들의 갯수 또는 분석 필터 뱅크 채널들의 갯수를 구성시킴으로써 신호 프로세서를 구성시키도록 적응될 수 있다. 서브대역 발생기는 사전처리된 주파수 도메인 오디오 신호를 위한 추가적인 서브대역 신호들을 생성하기 위해 사전처리된 오디오 신호 발생기의 서브대역 신호들을 복제하도록 적응된 스펙트럼 대역 복제기일 수 있다. 신호 프로세서는 사전처리된 오디오 신호를 디코딩하여 스테레오 또는 서라운드 채널들을 포함한 사전처리된 오디오 신호를 획득하기 위한 MPEG 서라운드 디코더를 더 포함할 수 있다. 게다가, 서브대역 발생기는 사전처리된 주파수 도메인 오디오 신호를 위한 추가적인 서브대역 신호들이 생성되어 사전처리된 주파수 도메인 오디오 신호에 추가된 후 사전처리된 주파수 도메인 오디오 신호를 MPEG 서라운드 디코더에 제공하도록 적응될 수 있다.
- [0025] 코어 디코더 모듈은 제1 코어 디코더와 제2 코어 디코더를 포함할 수 있으며, 제1 코어 디코더는 시간 도메인에서 동작하도록 적응될 수 있고, 제2 코어 디코더는 주파수 도메인에서 동작하도록 적응될 수 있다. 제1 코어 디코더는 ACELP 디코더일 수 있고 제2 코어 디코더는 FD 변환 디코더 또는 TCX 변환 디코더일 수 있다.
- [0026] 실시예에서, ACELP 코덱을 위한 수퍼 프레임 크기는 1024개 샘플들에서 768개 샘플들로 감소된다. 이것은 192 크기(64 크기의 3개 서브 프레임들)의 4개의 ACELP 프레임들을 768 크기의 하나의 코어 코더 프레임으로 결합시킴으로써 행해질 수 있다(이전에는, 256 크기의 4개의 ACELP 프레임들이 1024 크기의 코어 코더 프레임으로 결합되었다). 768개 샘플들의 코어 코더 프레임 크기에 도달하기 위한 또 다른 솔루션은 예컨대 256 크기(64 크기의 4개의 서브 프레임들)의 3개의 ACELP 프레임들을 결합하는 것일 것이다.
- [0027] 추가적인 실시예에 따르면, 구성기는, 오디오 신호의 제1 구성가능한 샘플들의 갯수 또는 처리된 오디오 신호의 제2 구성가능한 샘플들의 갯수 중 적어도 하나를 표시하는 구성 정보에 기초하여 신호 프로세서를 구성하도록 적응된다.
- [0028] 또 다른 실시예에서, 구성기는 구성 정보에 기초하여 신호 프로세서를 구성하도록 적응되며, 구성 정보는 오디오 신호의 제1 구성가능한 샘플들의 갯수 및 처리된 오디오 신호의 제2 구성가능한 샘플들의 갯수를 표시하며, 구성 정보는 구성 인덱스이다.
- [0029] 게다가, 오디오 신호를 처리하기 위한 장치가 제공된다. 장치는 신호 프로세서 및 구성기를 포함한다. 신호 프로세서는 오디오 신호의 제1 구성가능한 갯수의 샘플들을 갖는 제1 오디오 신호 프레임을 수신하도록 적응된다. 게다가, 신호 프로세서는 오디오 신호를 구성가능한 다운샘플링 계수만큼 다운샘플링하여, 처리된 오디오 신호를 획득하도록 적응된다. 더 나아가, 신호 프로세서는 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들을 갖는 제2 오디오 신호 프레임을 출력하도록 적응된다.
- [0030] 구성기는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 구성가능한 다운샘플링 계수가 제1 다운샘플링값과 동일하도록 하는 구성 정보에 기초하여 신호 프로세서를 구성하도록 적응될 수 있다. 게다가, 구성기는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 다운샘플링 계수가 상이한 제2 다운샘플링값과 동일하도록, 신호 프로세서를 구성하도록 적응된다. 제1 또는 제2 비율값은 정수값이 아니다.

도면의 간단한 설명

[0031]

본 발명의 바람직한 실시예들을 첨부 도면들과 관련지어서 후속하여 논의한다.

도 1은 실시예에 따른 오디오 신호를 처리하기 위한 장치를 도시한다.

도 2는 또다른 실시예에 따른 오디오 신호를 처리하기 위한 장치를 도시한다.

도 3은 실시예에 따른 장치에 의해 수행되는 업샘플링 처리를 도시한다.

도 4는 추가적인 실시예에 따른 오디오 신호를 처리하기 위한 장치를 도시한다.

도 5a는 실시예에 따른 코어 디코더 모듈을 도시한다.

도 5b는 도 5a에 따른 코어 디코더 모듈과 함께 도 4의 실시예에 따른 오디오 신호를 처리하기 위한 장치를 도시한다.

도 6a는 4개의 ACELP 프레임들을 포함한 ACELP 수퍼 프레임을 도시한다.

도 6b는 3개의 ACELP 프레임들을 포함한 ACELP 수퍼 프레임을 도시한다.

도 7a는 USAC의 디폴트 세팅을 도시한다.

도 7b는 실시예에 따른 USAC에 대한 추가 세팅을 도시한다.

도 8a와 도 8b는 MUSHRA 방법론에 따른 리스닝 테스트의 결과들을 도시한다.

도 9는 대안적인 실시예에 따른 오디오 신호를 처리하기 위한 장치를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0032]

도 1은 실시예에 따른 오디오 신호를 처리하기 위한 장치를 도시한다. 장치는 신호 프로세서(110) 및 구성기(120)를 포함한다. 신호 프로세서(110)는 오디오 신호의 제1 구성가능한 갯수의 샘플들(145)을 갖는 제1 오디오 신호 프레임(140)을 수신하도록 적응된다. 게다가, 신호 프로세서(110)는 오디오 신호를 구성가능한 업샘플링 계수만큼 업샘플링하여, 처리된 오디오 신호를 획득하도록 적응된다. 더 나아가, 신호 프로세서는 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들(155)을 갖는 제2 오디오 신호 프레임(150)을 출력하도록 적응된다.

[0033]

구성기(120)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 구성가능한 업샘플링 계수가 제1 업샘플링 값과 동일하도록 하는 구성 정보(ci)에 기초하여 신호 프로세서(110)를 구성하도록 적응된다. 게다가, 구성기(120)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 업샘플링 계수가 상이한 제2 업샘플링 값과 동일하도록, 신호 프로세서(110)를 구성하도록 적응된다. 제1 또는 제2 비율값은 정수값이 아니다.

[0034]

도 1에 따른 장치는 예컨대 디코딩 처리에서 활용될 수 있다.

[0035]

실시예에 따르면, 구성기(120)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제2 비율이 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율보다 클 때, 상이한 제2 업샘플링 값이 상이한 제1 업샘플링 값보다 크도록, 신호 프로세서(110)를 구성하도록 적응될 수 있다. 추가적인 실시예에서, 구성기(120)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 구성가능한 업샘플링 계수가 제1 비율값과 동일하도록, 신호 프로세서(110)를 구성하도록 적응되며, 구성기(120)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 업샘플링 계수가 상이한 제2 비율값과 동일하도록, 신호 프로세서(110)를 구성하도록 적응된다.

[0036]

또다른 실시예에서, 구성기(120)는, 제1 비율이 제1 비율값을 가질 때, 구성가능한 업샘플링 계수가 2와 동일하도록, 신호 프로세서(110)를 구성하도록 적응되며, 구성기(120)는, 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 업샘플링 계수가 8/3과 동일하도록, 신호 프로세서(110)를 구성하도록 적응된다. 추가적인 실시예에 따르면, 구성기(120)는, 제1 비율이 제1 비율값을 가질 때, 제1 구성가능한 샘플들의 갯수가 1024와 동일하고 제2 구성가능한 샘플들의 갯수가 2048과 동일하도록, 신호 프로세서(110)를 구성하도록 적응되며, 구성기(120)는, 제2 비율이 상이한 제2 비율값을 가질 때, 제1 구성가능한 샘플들의 갯수가 768과 동일하고 제2 구성가능한

샘플들의 갯수가 2048과 동일하도록, 신호 프로세서(110)를 구성하도록 적응된다.

[0037] 실시예에서, 구성기(120)는 구성 정보(ci)에 기초하여 신호 프로세서(110)를 구성하도록 적응되며, 구성 정보(ci)는 업샘플링 계수, 오디오 신호의 제1 구성가능한 샘플들의 갯수 및 처리된 오디오 신호의 제2 구성가능한 샘플들의 갯수를 표시하며, 구성 정보는 구성 인덱스이다.

[0038] 아래의 [표]는 구성 정보로서의 구성 인덱스에 대한 예시를 도시한다:

인덱스	코어코더프레임길이	sbr비율	출력프레임길이
2	768	8:3	2048
3	1024	2:1	2048

[0039] 여기서 "인덱스"는 구성 인덱스를 표시하고, "코어코더프레임길이"는 오디오 신호의 제1 구성가능한 샘플들의 갯수를 표시하고, "sbr비율"은 업샘플링 계수를 표시하며, "출력프레임길이"는 처리된 오디오 신호의 제2 구성가능한 샘플들의 갯수를 표시한다.

[0040] 도 2는 또 다른 실시예에 따른 장치를 도시한다. 장치는 신호 프로세서(205) 및 구성기(208)를 포함한다. 신호 프로세서(205)는 코어 디코더 모듈(210), 분석 필터 뱅크(220), 서브대역 발생기(230), 및 합성 필터 뱅크(240)를 포함한다.

[0041] 코어 디코더 모듈(210)은 오디오 신호(as1)를 수신하도록 적응된다. 코어 디코더 모듈(210)은, 오디오 신호(as1)를 수신한 후, 오디오 신호를 디코딩하여 사전처리된 오디오 신호(as2)를 획득한다. 그런 후, 코어 디코더 모듈(210)은 시간 도메인으로 표현되는 사전처리된 오디오 신호(as2)를 분석 필터 뱅크(220)에 제공한다.

[0042] 분석 필터 뱅크(220)는 사전처리된 오디오 신호(as2)를 시간 도메인에서 주파수 도메인으로 변환시켜서 복수의 서브대역 신호들을 포함한 사전처리된 주파수 도메인 오디오 신호(as3)를 획득하도록 적응된다. 분석 필터 뱅크(220)는 구성가능한 갯수의 분석 필터 뱅크 채널들(분석 필터 뱅크 대역들)을 갖는다. 분석 필터 뱅크 채널들의 갯수는 사전처리된 시간 도메인 오디오 신호(as2)로부터 생성된 서브대역 신호들의 갯수를 결정한다. 실시예에서, 분석 필터 뱅크 채널들의 갯수는 구성가능한 파라미터(c1)의 값을 세팅함으로써 세팅될 수 있다. 예를 들어, 분석 필터 뱅크(220)는 32개 또는 24개의 분석 필터 뱅크 채널들을 갖도록 구성될 수 있다. 도 2의 실시예에서, 분석 필터 뱅크 채널들의 갯수는 구성기(208)의 구성 정보(ci)에 따라 세팅될 수 있다. 분석 필터 뱅크(220)는, 사전처리된 오디오 신호(as2)를 주파수 도메인으로 변환시킨 후, 사전처리된 주파수 도메인 오디오 신호(as3)를 서브대역 발생기(230)에 제공한다.

[0043] 서브대역 발생기(230)는 주파수 도메인 오디오 신호(as3)에 대한 추가적인 서브대역 신호들을 생성하도록 적응된다. 게다가, 서브대역 발생기(230)는 사전처리된 주파수 도메인 오디오 신호(as3)를 수정하여, 서브대역 발생기(230)에 의해 생성된 추가적인 서브대역 신호들 및 사전처리된 주파수 도메인 오디오 신호(as3)의 서브대역 신호들을 포함하는 수정된 주파수 도메인 오디오 신호(as4)를 획득하도록 적응된다. 서브대역 발생기(230)에 의해 생성된 추가적인 서브대역 신호들의 갯수는 구성가능하다. 실시예에서, 서브대역 발생기는 스펙트럼 대역 복제기(Spectral Band Replicator; SBR)이다. 그런 후 서브대역 발생기(230)는 수정된 사전처리 주파수 도메인 오디오 신호(as4)를 합성 필터 뱅크에 제공한다.

[0044] 합성 필터 뱅크(240)는 수정된 사전처리 주파수 도메인 오디오 신호(as4)를 주파수 도메인에서 시간 도메인으로 변환시켜서 처리된 시간 도메인 오디오 신호(as5)를 획득하도록 적응된다. 합성 필터 뱅크(240)는 구성가능한 갯수의 합성 필터 뱅크 채널들(합성 필터 뱅크 대역들)을 갖는다. 합성 필터 뱅크 채널들의 갯수는 구성가능하다. 실시예에서, 합성 필터 뱅크 채널들의 갯수는 구성가능한 파라미터(c2)의 값을 세팅함으로써 세팅될 수 있다. 예를 들어, 합성 필터 뱅크(240)는 64개의 합성 필터 뱅크 채널들을 갖도록 구성될 수 있다. 도 2의 실시예에서, 구성기(208)의 구성 정보(ci)는 분석 필터 뱅크 채널들의 갯수를 세팅할 수 있다. 수정된 사전처리 주파수 도메인 오디오 신호(as4)를 시간 도메인으로 변환시킴으로써, 처리된 오디오 신호(as5)가 획득된다.

[0045] 실시예에서, 수정된 사전처리 주파수 도메인 오디오 신호(as4)의 서브대역 채널들의 갯수는 합성 필터 뱅크 채널들의 갯수와 동일하다. 이러한 실시예에서, 구성기(208)는 서브대역 발생기(230)에 의해 생성된 추가적인 서브대역 채널들의 갯수를 구성하도록 적응된다. 구성기(208)는 구성기(208)에 의해 구성된 합성 필터 뱅크 채널들의 갯수(c2)가 사전처리된 주파수 도메인 오디오 신호(as3)의 서브대역 채널들의 갯수 더하기 서브대역 발생기(230)에 의해 생성된 추가적인 서브대역 신호들의 갯수와 동일하도록, 서브대역 발생기(230)에 의해 생성된 추가적인 서브대역 채널들의 갯수를 구성하도록 적응될 수 있다. 이에 의해, 합성 필터 뱅크 채널들의 갯수는

수정된 사전처리 주파수 도메인 오디오 신호(as4)의 서브대역 신호들의 갯수와 동일하다.

[0047] 오디오 신호(as1)가 샘플링 레이트(sr1)를 갖는다는 것과, 분석 필터 뱅크(220)가 c1개의 분석 필터 뱅크 채널들을 갖고 합성 필터 뱅크(240)가 c2개의 합성 필터 뱅크 채널들을 갖는다는 것을 가정하면, 처리된 오디오 신호(as5)는 샘플링 레이트(sr5)를 갖는다:

$$sr5 = (c2/c1) \cdot sr1$$

[0049] $c2/c1$ 은 업샘플링 계수(u)를 결정한다:

$$u = c2/c1$$

[0051] 도 2의 실시예에서, 업샘플링 계수(u)는 정수값이 아닌 숫자로 세팅될 수 있다. 예를 들어, 업샘플링 계수(u)는 분석 필터 뱅크 채널들의 갯수를 $c1 = 24$ 로 세팅하고, 합성 필터 뱅크 채널들의 갯수를 $c2 = 64$ 로 세팅함으로써 $8/3$ 값으로 세팅될 수 있다:

$$u = 8/3 = 64/24$$

[0053] 서브대역 발생기(230)가 스펙트럼 대역 복제기인 것을 가정하면, 실시예에 따른 스펙트럼 대역 복제기는 원래 서브대역들로부터 임의적인 갯수의 추가적인 서브대역들을 생성할 수 있으며, 이미 이용가능한 서브대역들의 갯수에 대한 생성된 추가적인 서브대역들의 갯수의 비율은 정수일 필요는 없다. 예를 들어, 실시예에 따른 스펙트럼 대역 복제기는 아래의 단계들을 수행할 수 있다:

[0054] 첫번째 단계에서, 스펙트럼 대역 복제기는 복수의 추가적인 서브대역들을 생성시킴으로써 해당 갯수의 서브대역 신호들을 복제하고, 생성된 추가적인 서브대역들의 갯수는 이미 이용가능한 서브대역들의 갯수의 정수배일 수 있다. 예를 들어, 24개(또는, 예컨대 48개)의 추가적인 서브대역 신호들은 오디오 신호의 24개의 원래의 서브대역 신호들로부터 생성될 수 있다(예컨대, 서브대역 신호들의 총 갯수는 두 배 또는 세 배일 수 있다).

[0055] 두번째 단계에서, 서브대역 신호들의 희망하는 갯수가 $c12$ 이고 실제로 이용가능한 서브대역 신호들의 갯수가 $c11$ 이라고 가정하면, 세 개의 상이한 상황들이 구별될 수 있다:

[0056] $c11$ 이 $c12$ 와 동일하면, 이용가능한 서브대역 신호들의 갯수($c11$)는 필요한 서브대역 신호들의 갯수($c12$)와 동일하다. 어떠한 서브대역 조정도 필요하지 않다.

[0057] $c12$ 가 $c11$ 보다 작다면, 이용가능한 서브대역 신호들의 갯수($c11$)는 필요한 서브대역 신호들의 갯수($c12$)보다 크다. 실시예에 따르면, 최고 주파수 서브대역 신호들이 삭제될 수 있다. 예를 들어, 만약 64개의 서브대역 신호들이 이용가능하고 단지 61개의 서브대역 신호들만이 필요하다면, 최고 주파수를 갖는 세 개의 서브대역 신호들은 폐기될 수 있다.

[0058] $c12$ 가 $c11$ 보다 크다면, 이용가능한 서브대역 신호들의 갯수($c11$)는 필요한 서브대역 신호들의 갯수($c12$)보다 작다.

[0059] 실시예에 따르면, 추가적인 서브대역 신호들로서 제로 신호들, 즉 각각의 서브대역 샘플의 진폭값들이 제로와 동일한 신호들을 추가함으로써 추가적인 서브대역 신호들이 생성될 수 있다. 또 다른 실시예에 따르면, 의사난수 서브대역 신호들, 즉 각각의 서브대역 샘플의 값들이 의사난수 데이터를 포함하는 서브대역 신호들을 추가적인 서브대역 신호들로서 추가함으로써 추가적인 서브대역 신호들이 생성될 수 있다. 또 다른 실시예에서, 추가적인 서브대역 신호들은 최고 서브대역 신호 또는 최고 서브대역 신호들의 샘플값들을 복사하고, 이것들을 추가적인 서브대역 신호들(복사된 서브대역 신호들)의 샘플값들로서 이용함으로써 생성될 수 있다.

[0060] 실시예에 따른 스펙트럼 대역 복제기에서, 모든 서브대역들이 채워지도록, 이용가능한 기저대역 서브대역들은 복사되어 최고 서브대역들로서 활용될 수 있다. 누락된 모든 서브대역들이 값들로 채워질 수 있도록, 동일한 기저대역 서브대역은 두 번 또는 여러 번 복사될 수 있다.

[0061] 도 3은 실시예에 따른 장치에 의해 수행되는 업샘플링 처리를 도시한다. 시간 도메인 오디오 신호(310) 및 오디오 신호(310)의 몇몇의 샘플들(315)이 도시된다. 세 개의 서브대역 신호들(330)을 포함한 주파수 도메인 오디오 신호(320)를 획득하기 위해 오디오 신호는 주파수 도메인으로 변환되는데, 예컨대 시간 도메인에서 주파수 도메인으로 변환된다. (이 단순예에서는, 분석 필터 뱅크가 3개의 채널들을 포함하는 것으로 가정한다.) 그런 후 세 개의 추가적인 서브대역 신호들(335)을 획득하기 위해 주파수 도메인 오디오 신호(330)의 서브대역 신호들은 복제될 수 있으며 이로써 주파수 도메인 오디오 신호(320)는 원래의 세 개의 서브대역 신호들(330)과 세 개의 생

성된 추가적인 서브대역 신호들(335)을 포함한다. 그런 후, 두 개의 또다른 추가적인 서브대역 신호들(338), 예컨대 제로 신호들, 의사난수 서브대역 신호들 또는 복사된 서브대역 신호들이 생성된다. 그런 후 주파수 도메인 오디오 신호는 시간 도메인으로 역변환되어 그 결과 원래의 시간 도메인 오디오 신호(310)의 8/3배 샘플링 레이트인 샘플링 레이트를 갖는 시간 도메인 오디오 신호(350)를 초래시킨다.

[0062] 도 4는 추가적인 실시예에 따른 장치를 도시한다. 장치는 신호 프로세서(405) 및 구성기(408)를 포함한다. 신호 프로세서(405)는 도 2의 실시예에서의 각각의 유닛들에 대응하는, 코어 디코더 모듈(210), 분석 필터 맹크(220), 서브대역 발생기(230), 및 합성 필터 맹크(240)를 포함한다. 신호 프로세서(405)는 사전처리된 오디오 신호를 디코딩하여 스테레오 또는 서라운드 채널들을 갖는 사전처리된 오디오 신호를 획득하기 위한 MPEG 서라운드 디코더(410)(MPS 디코더)를 더 포함한다. 서브대역 발생기(230)는 사전처리된 주파수 도메인 오디오 신호를 위한 추가적인 서브대역 신호들이 생성되어 사전처리된 주파수 도메인 오디오 신호에 추가된 후 사전처리된 주파수 도메인 오디오 신호를 MPEG 서라운드 디코더(410)에 제공하도록 적응된다.

[0063] 도 5a는 실시예에 따른 코어 디코더 모듈을 도시한다. 코어 디코더 모듈은 제1 코어 디코더(510)와 제2 코어 디코더(520)를 포함한다. 제1 코어 디코더(510)는 시간 도메인에서 동작하도록 적응되고 제2 코어 디코더(520)는 주파수 도메인에서 동작하도록 적응된다. 도 5a에서, 제1 코어 디코더(510)는 ACELP 디코더이고 제2 코어 디코더(520)는 FD 변환 디코더, 예컨대 AAC 변환 디코더이다. 대안적인 실시예에서, 제2 코어 디코더(520)는 TCX 변환 디코더이다. 도착하는 오디오 신호 부분(asp)이 음성 데이터 또는 다른 오디오 데이터를 포함하는지 여부에 의존하여, 도착하는 오디오 신호 부분(asp)은 ACELP 디코더(510)에 의해 또는 FD 변환 디코더(520)에 의해 처리된다. 코어 디코더 모듈의 출력은 오디오 신호의 사전처리된 부분(pp-asp)이다.

[0064] 도 5b는 도 5a에 따른 코어 디코더 모듈과 함께 도 4의 실시예에 따른 오디오 신호를 처리하기 위한 장치를 도시한다.

[0065] 실시예에서, ACELP 코덱을 위한 수퍼 프레임 크기는 1024개 샘플들에서 768개 샘플들로 감소된다. 이것은 192 크기(64 크기의 3개 서브 프레임들)의 4개의 ACELP 프레임들을 768 크기의 하나의 코어 코더 프레임으로 결합시킴으로써 행해질 수 있다(이전에는, 256 크기의 4개의 ACELP 프레임들이 1024 크기의 코어 코더 프레임으로 결합되었다). 도 6a는 4개의 ACELP 프레임들(610)을 포함한 ACELP 수퍼 프레임(605)을 도시한다. ACELP 프레임들(610) 각각은 3개의 서브프레임들(615)을 포함한다.

[0066] 768개 샘플들의 코어 코더 프레임 크기에 도달하기 위한 또다른 솔루션은 예컨대 256 크기(64 크기의 4개의 서브 프레임들)의 3개의 ACELP 프레임들을 결합하는 것일 것이다. 도 6b는 3개의 ACELP 프레임들(630)을 포함한 ACELP 수퍼 프레임(625)을 도시한다. ACELP 프레임들(630) 각각은 4개의 서브프레임들(635)을 포함한다.

[0067] 도 7b는 디코더 관점으로부터의 제안된 추가적인 세팅을 약술하며 이것을 통상적인 USAC 세팅과 비교한다. 도 7a과 도 7b는 24 kbit/s 또는 32 kbit/s로서 동작점들에서 일반적으로 이용되는 디코더 구조를 약술한다.

[0068] USAC RM9(USAC 참조 모델 9) 디폴트 세팅을 도시하는 도 7a에서, 오디오 신호 프레임은 QMF 분석 필터 맹크(710)에 입력된다. QMF 분석 필터 맹크(710)는 32개의 채널들을 갖는다. QMF 분석 필터 맹크(710)는 시간 도메인 오디오 신호를 주파수 도메인으로 변환시키도록 적응되고, 주파수 도메인 오디오 신호는 32개의 서브대역들을 포함한다. 그런 후 주파수 도메인 오디오 신호는 업샘플러(720)에 입력된다. 업샘플러(720)는 주파수 도메인 오디오 신호를 업샘플링 계수(2)만큼 업샘플링하도록 적응된다. 따라서, 64개 서브대역들을 포함한 주파수 도메인 업샘플러 출력 신호가 업샘플러에 의해 생성된다. 업샘플러(720)는 스펙트럼 대역 복제(Spectral Band Replication; SBR) 업샘플러이다. 이미 언급한 바와 같이, 스펙트럼 대역 복제는 스펙트럼 대역 복제기 내로 입력되는 보다 낮은 주파수 서브대역들로부터 보다 높은 주파수 서브대역들을 생성하기 위해 활용된다.

[0069] 그런 후 업샘플링된 주파수 도메인 오디오 신호는 MPEG 서라운드(MPEG Surround; MPS) 디코더(730)에 제공된다. MPS 디코더(730)는 다운믹싱된 서라운드 신호를 디코딩하여 서라운드 신호의 주파수 도메인 채널들을 유도해내도록 적응된다. 예를 들어, MPS 디코더(730)는 주파수 도메인 서라운드 신호의 2개의 업믹싱된 주파수 도메인 서라운드 채널들을 생성하도록 적응될 수 있다. 또다른 실시예에서, MPS 디코더(730)는 주파수 도메인 서라운드 신호의 5개의 업믹싱된 주파수 도메인 서라운드 채널들을 생성하도록 적응될 수 있다. 그런 후 주파수 도메인 서라운드 신호의 채널들은 QMF 합성 필터 맹크(740)에 제공된다. QMF 합성 필터 맹크(740)는 주파수 도메인 서라운드 신호의 채널들을 시간 도메인으로 변환시켜서 서라운드 신호의 시간 도메인 채널들을 획득하도록 적응된다.

[0070] 살펴볼 수 있는 바와 같이, USAC 디코더는 2:1 시스템으로서 자신의 디폴트 세팅으로 동작한다. 코어 코덱은 출

력 샘플링 레이트(f_{out})의 절반의 1024개 샘플들/프레임의 입도에서 동작한다. 32개 대역 분석 QMF 필터 뱅크를 동일 레이트로 구동하는 64개 대역 합성 QMF 뱅크와 결합시킴으로써, 2배의 업샘플링이 SBR 툴 내부에서 암시적으로 수행된다. SBR 툴은 f_{out} 에서 2048 크기의 프레임들을 출력한다.

[0071] 도 7b는 USAC에 대한 제안된 추가 세팅을 도시한다. QMF 분석 필터 뱅크(750), 업샘플러(760), MPS 디코더(770) 및 합성 필터 뱅크(780)가 도시된다.

[0072] 디폴트 세팅과는 대조적으로, USAC 코덱은 8/3 시스템으로서 제안된 추가 세팅에서 동작한다. 코어 코더는 출력 샘플링 레이트(f_{out})의 3/8배 샘플링 레이트로 구동한다. 동일한 환경에서, 코어 코더 프레임 크기는 $\frac{3}{4}$ 배로 스케일링 다운되었다. SBR 툴 내부에서의 24개 대역 분석 QMF 필터 뱅크와 64개 대역 합성 필터 뱅크의 조합에 의해, 2048개 샘플들의 프레임 길이에서의 f_{out} 의 출력 샘플링 레이트가 달성될 수 있다.

[0073] 이러한 세팅은 코어 코더 및 추가적인 툴들 모두에 대한 시간적 입도를 훨씬 많이 증가시키도록 해준다: SBR 및 MPEG 서라운드와 같은 툴들은 보다 높은 샘플링 레이트에서 동작될 수 있는 반면에, 코어 코더 샘플링 레이트는 감소되고 대신에 프레임 길이는 단축된다. 이러한 방식에 의해, 모든 컴포넌트들은 자신의 최적의 환경에서 동작할 수 있다.

[0074] 실시예에서, AAC 코더가 출력 샘플링 레이트(f_{out})의 3/8배 샘플링 레이트로 동작한다 하더라도, 코어 코더로서 활용된 AAC 코더는 $\frac{1}{2} f_{out}$ 샘플링 레이트에 기초하여 여전히 스케일계수들을 결정할 수 있다.

[0075] 아래의 [표]는 USAC 참조 웰리티 인코더에서 이용되는 USAC에 대한 샘플링 레이트들 및 프레임 지속기간에 대한 상세한 수치들을 제공한다. 살펴볼 수 있는 바와 같이, 제안된 새로운 세팅에서의 프레임 지속기간은 거의 25% 만큼 감소될 수 있는데, 코딩 노이즈의 확산이 또한 이와 마찬가지 비율만큼 감소될 수 있으므로, 이것은 모든 비정지형(non-stationary signal) 신호들에 대해 긍정적인 영향들을 야기시킨다. 이러한 감소는 ACELP 툴을 최적화된 동작 범위 밖으로 이탈시킬 코어 코더 샘플링 주파수 증가 없이 달성될 수 있다.

	샘플링 레이트 코어 코더	샘플링 레이트 SBR	프레임당 지속기간
USAC 디폴트	17075 Hz	34150 Hz	60 ms
제안된 새로운 세팅	16537.5 Hz	44100 Hz	46 ms

[0076]

[표]는 24 kbit/s에서의 참조 웰리티 인코더에서 이용되는 디폴트 세팅 및 제안된 새로운 세팅에 대한 샘플링 레이트들 및 프레임 지속기간을 도시한다.

[0078] 아래에서는, 제안된 새로운 세팅을 구현하기 위해, USAC 디코더에 대한 필요한 수정들을 보다 상세하게 설명한다.

[0079] 변환 코더와 관련하여, 변환 및 윈도우 크기들을 $\frac{3}{4}$ 배로 스케일링함으로써 보다 짧은 프레임 크기들이 손쉽게 달성될 수 있다. FD 코더는 표준 모드에서 1024 및 128의 변환 크기들을 갖고 동작하는 반면에, 새로운 세팅에 의해 768 및 96 크기의 추가적인 변환들이 도입된다. TCX의 경우, 768, 384 및 192의 크기의 추가적인 변환들이 필요하다. 윈도우 계수들에 따른 새로운 변환 크기들을 규정하는 것 이외에, 변환 코더는 변동없이 유지될 수 있다.

[0080] ACELP 툴과 관련하여, 전체 프레임 크기는 768개 샘플들로 적응될 필요가 있다. 이러한 목적을 달성하기 위한 한가지 방법은 192개 샘플들의 4개의 ACELP 프레임들을 768개 샘플들의 각 프레임 내에 끼워 맞추되 프레임의 전체적인 구조는 변경하지 않은 채로 남겨두는 것이다. 감소된 프레임 크기로의 적응은 프레임당 서브프레임들의 갯수를 4개에서 3개로 감소시킴으로써 달성된다. ACELP 서브프레임 길이는 변경되지 않고 64개 샘플들로 유지된다. 서브프레임들의 갯수가 감소되도록 하기 위해, 약간 상이한 기법을 이용하여 피치 정보가 인코딩되며, 세 개의 피치값들은 표준 모델에서 9, 6, 9 및 6 비트들을 이용하는 절대-상대-절대-상대 방식 대신에 9, 6 및 6 비트들을 각각 이용하는 절대-상대-상대 방식을 이용하여 인코딩된다. 하지만, 피치 정보를 코딩하는 다른 방법들도 가능하다. 다양한 양자화기들(LPC 필터들, 이득들 등)뿐만이 아니라 ACELP 코드북들과 같은, ACELP 코덱의 다른 엘리먼트들은 변동없다.

[0081] 768개 샘플들의 전체 프레임 크기를 달성하기 위한 또 다른 방법은 768 크기의 하나의 코어 코더 프레임에 대한

256 크기의 세 개의 ACELP 프레임들을 결합하는 것일 것이다.

[0082] SBR 툴의 기능성은 변동없다. 하지만, 8/3배의 업샘플링을 가능하게 하기 위해, 32개 대역 분석 QMF, 24개 대역 분석 QMF에 대한 추가가 필요하다.

[0083] 아래에서는, 계산 복잡도에 대한 제안된 추가 동작점의 영향을 설명한다. 이것은 제일먼저 코덱 툴 별로 이뤄지고 마지막에는 요약을 한다. 복잡도는 이러한 동작점들에 대한 대응하는 HE-AACv2 세팅에 필적하는 보다 높은 비트레이트들에서 USAC 참조 퀄리티 인코더에 의해 이용되는 보다 높은 샘플링 레이트 모드와 디폴트의 낮은 샘플링 레이트 모드에 대해 비교되어 진다.

[0084] 변환 코더와 관련하여, 변환 코더 부분들의 복잡도는 샘플링 레이트와 변환 길이로 스케일링된다. 제안된 코어 코더 샘플링 레이트들은 대체로 동일하게 잔존한다. 변환 크기들은 $\frac{1}{4}$ 배만큼 감소된다. 이에 의해, 기저 FFT들에 대한 혼합 기수법을 가정하면, 계산 복잡도는 이와 거의 동일한 배수만큼 감소된다. 전반적으로, 변환 기반 디코더의 복잡도는 현재의 USAC 동작점에 비해 약간 감소되고 높은 샘플링 동작 모드에 비해 $\frac{1}{4}$ 배만큼 감소될 것으로 예상된다.

[0085] ACELP와 관련하여, ACELP 툴들의 복잡도는 주로 다음의 동작들로 집결된다:

[0086] 디코딩 여기(Decoding of the excitation): 상기 동작의 복잡도는 초당 서브프레임들의 갯수에 비례하며, 이는 결국 코어 코더 샘플링 주파수에 정비례한다(서브프레임 크기는 64개 샘플들로 변동없다). 그러므로, 이것은 새로운 세팅과 거의 동일하다.

[0087] 베이스-포스트필터(bass-postfilter)를 비롯한 LPC 필터링 및 다른 합성 동작들: 이 동작의 복잡도는 코어 코더 샘플링 주파수에 정비례하며 그러므로 거의 동일하다.

[0088] 전반적으로, ACELP 디코더의 예상된 복잡도는 현재의 USAC 동작점에 비해 변동없고 높은 샘플링 동작 모드에 비해 $\frac{1}{4}$ 배만큼 감소될 것으로 예상된다.

[0089] SBR과 관련하여, SBR 복잡도에 대한 주요 기여자들은 QMF 필터뱅크들이다. 여기서 복잡도는 샘플링 레이트와 변환 크기로 스케일링된다. 특히, 분석 필터뱅크의 복잡도는 대체로 $\frac{1}{4}$ 배만큼 감소된다.

[0090] MPEG 서라운드와 관련하여, MPEG 서라운드 부분의 복잡도는 샘플링 레이트로 스케일링된다. 제안된 추가 동작 모드는 MPEG 서라운드 툴의 복잡도에 어떠한 직접적인 영향도 미치지 않는다.

[0091] 전체적으로, 제안된 새로운 동작 모드의 복잡도는 낮은 샘플링 레이트 모드에 비해 약간 더 복잡하지만, 보다 높은 샘플링 레이트 모드에서 구동될 경우에는 USAC 디코더의 복잡도 미만이라는 것이 발견되었다(USAC RM9, 하이 SR: 13.4 MOPS, 제안된 새로운 동작점: 12.8 MOPS).

[0092] 테스트된 동작점에 대해, 복잡도는 다음과 같이 평가된다:

[0093] 34.15kHz에서 동작되는 USAC RM9: 대략 4.6 WMOPS;

[0094] 44.1kHz에서 동작되는 USAC RM9: 대략 5.6 WMOPS;

[0095] 제안된 새로운 동작점: 대략 5.0 WMOPS.

[0096] USAC 디코더는 디폴트 구성에서 48 kHz까지의 샘플링 레이트들을 처리가능할 필요가 있을 것으로 예상되기 때문에, 이 제안된 새로운 동작점에 의해서는 어떠한 결점도 예상되지 않는다.

[0097] 메모리 수요량과 관련하여, 제안된 추가 동작 모드는 추가적인 MDCT 윈도우 프로토타입들의 저장소를 필요로 하는데, 이것은 총 합산하여 900 워드(32비트) 미만의 추가적인 ROM 수요량에 불과하다. 대체로 25k워드인 전체 디코더 ROM 수요량에 비추어, 이것은 무시가능한 것으로 보인다.

[0098] 리스닝 테스트 결과들은 음성 항목들에 대한 퀄리티의 저하없이, 음악 및 혼합 테스트 항목들에 대한 상당한 개선을 보여준다. 이러한 추가 세팅은 USAC 코덱의 추가적인 동작 모드로서 의도된다.

[0099] 24 kbit/s 모노의 제안된 새로운 세팅의 성능을 평가하기 위해 MUSHRA 방법론에 따른 리스닝 테스트가 수행되었다. 아래의 조건들이 테스트에서 포함되었다: 숨은 참조; 3.5 kHz 로우패스 앵커; USAC WD7 참조 퀄리티 (WD7@34.15kHz); 높은 샘플링 레이트로 동작되는 USAC WD7(WD7@44.1kHz); 및 USAC WD7 참조 퀄리티, 제안된 새로운 세팅(WD7_CE@44.1kHz).

[0100] 테스트는 USAC 테스트 세트로부터의 12개의 테스트 항목들, 및 다음의 추가적인 항목들: si02: 카스타넷; 벨벳:

전자 음악; 및 실로폰: 음악 박스를 다뤘다.

[0101] 도 8a 및 도 8b는 테스트의 결과들을 도시한다. 22개의 주체들이 리스닝 테스트에 관여되었다. 평가를 위해 스튜던트-t 확률 분포가 이용되었다.

[0102] 평균 점수들의 평가(95% 유의수준)에서, 44.1 kHz의 보다 높은 샘플링 레이트에서 동작하는 WD7은 두 개의 항목들(es01, 해리포터)에 대해서 WD7보다 상당히 불량하게 수행한다는 것이 관찰될 수 있다. WD7과 본 기술을 피처링한 WD7 사이에서는, 어떠한 중요한 차이도 관찰될 수 없다.

[0103] 편차 점수들의 평가에서는, 44.1 kHz에서 동작하는 WD7은 6개 항목들(es01, 루이스_라킨, te1, 웨딩음성, 해리포터, 음악 내 음성_4)에 대해서 WD7보다 불량하게 수행한다는 것과 모든 항목들에 대해 평균화되었다는 것이 관찰될 수 있다. 불량하게 수행된 항목들은 완전히 순수한 음성 항목들 및 혼합 음성/음악 항목들 중 두 개를 포함한다. 더 나아가 44.1 kHz에서 동작하는 WD7은 네 개의 항목들(트윈클, 샐베이션, si02, 벨벳)에 대해서 WD7보다 상당히 우수하게 수행한다는 것이 관찰될 수 있다. 이러한 항목들 모두는 상당한 음악 신호 부분들을 포함하거나 또는 음악으로서 분류된 것들이다.

[0104] 테스트하에 있는 기술에 대해서 다섯 개의 항목들(트윈클, 샐베이션, te15, si02, 벨벳)에 대해서 WD7보다 우수하게 수행하고, 추가적으로 모든 항목들에 대해 평균화되었다는 것이 관찰될 수 있다. 우수하게 수행된 항목들 모두는 상당한 음악 신호 부분들을 포함하거나 또는 음악으로서 분류된 것들이다. 어떠한 저하도 관찰될 수 없었다.

[0105] 상술한 실시예들에 의해, 중간 USAC 비트레이트들에 대한 새로운 세팅이 제공된다. 이러한 새로운 세팅은 USAC 코덱으로 하여금, ACELP 틀의 퀄리티를 희생시키지 않고서, 변환 코더들, SBR 및 MPEG 서라운드와 같은, 모든 관련 틀들에 대한 시간적 입도를 증가시키도록 한다. 이에 의해, 중간 비트레이트 범위, 특히 높은 시간적 구조를 나타내는 음악 및 혼합 신호들에 대한 퀄리티는 향상될 수 있다. 더 나아가, ACELP 틀을 포함한 USAC 코덱은 이제 44.1 kHz와 같은, 보다 폭넓은 범위의 샘플링 레이트들에서 이용될 수 있기 때문에, USAC 시스템들은 유연성을 얻는다.

[0106] 도 9는 오디오 신호를 처리하기 위한 장치를 도시한다. 장치는 신호 프로세서(910) 및 구성기(920)를 포함한다. 신호 프로세서(910)는 오디오 신호의 제1 구성가능한 갯수의 샘플들(945)을 갖는 제1 오디오 신호 프레임(940)을 수신하도록 적응된다. 게다가, 신호 프로세서(910)는 오디오 신호를 구성가능한 다운샘플링 계수만큼 다운샘플링하여, 처리된 오디오 신호를 획득하도록 적응된다. 더 나아가, 신호 프로세서는 처리된 오디오 신호의 제2 구성가능한 갯수의 샘플들(955)을 갖는 제2 오디오 신호 프레임(950)을 출력하도록 적응된다.

[0107] 구성기(920)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 제1 비율이 제1 비율값을 가질 때, 구성가능한 다운샘플링 계수가 제1 다운샘플링 값과 동일하도록 하는 구성 정보(ci2)에 기초하여 신호 프로세서(910)를 구성하도록 적응된다. 게다가, 구성기(920)는, 제1 구성가능한 샘플들의 갯수에 대한 제2 구성가능한 샘플들의 갯수의 상이한 제2 비율이 상이한 제2 비율값을 가질 때, 구성가능한 다운샘플링 계수가 상이한 제2 다운샘플링 값과 동일하도록, 신호 프로세서(910)를 구성하도록 적응된다. 제1 또는 제2 비율값은 정수값이 아니다.

[0108] 도 9에 따른 장치는 예컨대 인코딩 처리에서 활용될 수 있다.

[0109] 비록 몇몇 양태들은 장치의 관점에서 설명되었지만, 이러한 양태들은 또한 대응 방법의 설명을 나타낸다는 것이 명백하며, 여기서 블록 또는 디바이스는 방법 단계 또는 방법 단계의 특징에 대응한다. 마찬가지로, 방법 단계의 관점에서 설명된 양태들은 또한 대응하는 장치의 대응하는 블록 또는 아이템 또는 특징의 설명을 나타낸다.

[0110] 본 발명의 분해된 신호는 디지털 저장 매체상에 저장될 수 있거나 또는 인터넷과 같은 무선 전송 매체 또는 유선 전송 매체와 같은 전송 매체를 통해 전송될 수 있다.

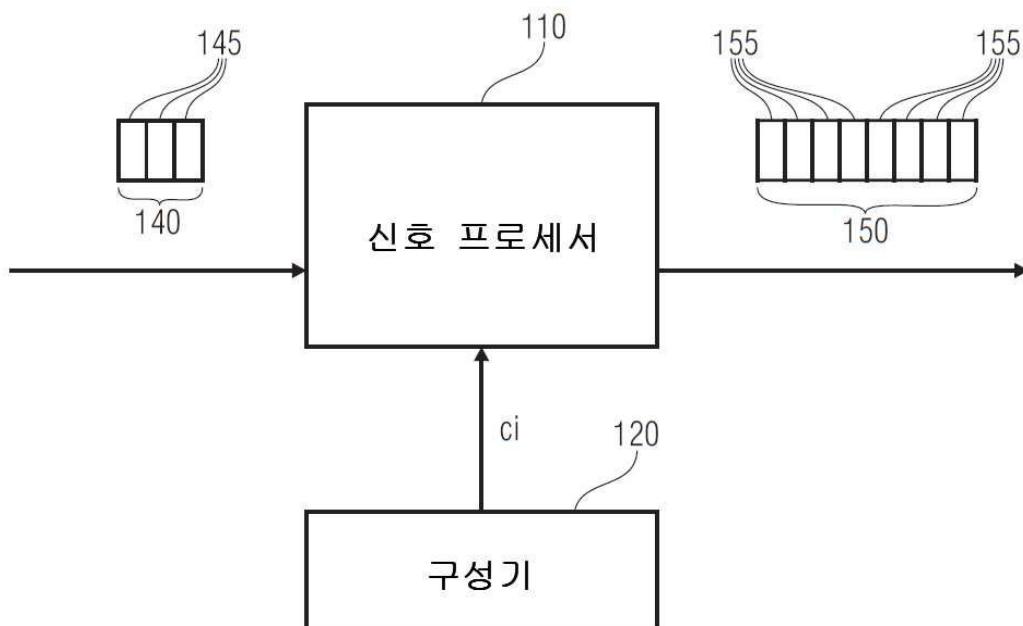
[0111] 일정한 구현 요구에 따라, 본 발명의 실시예들은 하드웨어나 소프트웨어로 구현될 수 있다. 이러한 구현은 전자적으로 판독가능한 제어 신호들이 저장되어 있으며, 각각의 방법이 수행되도록 프로그램가능한 컴퓨터 시스템과 협동하는(또는 이와 협동가능한) 디지털 저장 매체, 예컨대 플로피 디스크, DVD, CD, ROM, PROM, EPROM, EEPROM 또는 FLASH 메모리를 이용하여 수행될 수 있다.

[0112] 본 발명에 따른 몇몇의 실시예들은 여기서 설명된 방법들 중 하나의 방법이 수행되도록, 프로그램가능한 컴퓨터 시스템과 협동할 수 있는 전자적으로 판독가능한 제어 신호들을 갖는 비일시적 데이터 캐리어를 포함한다.

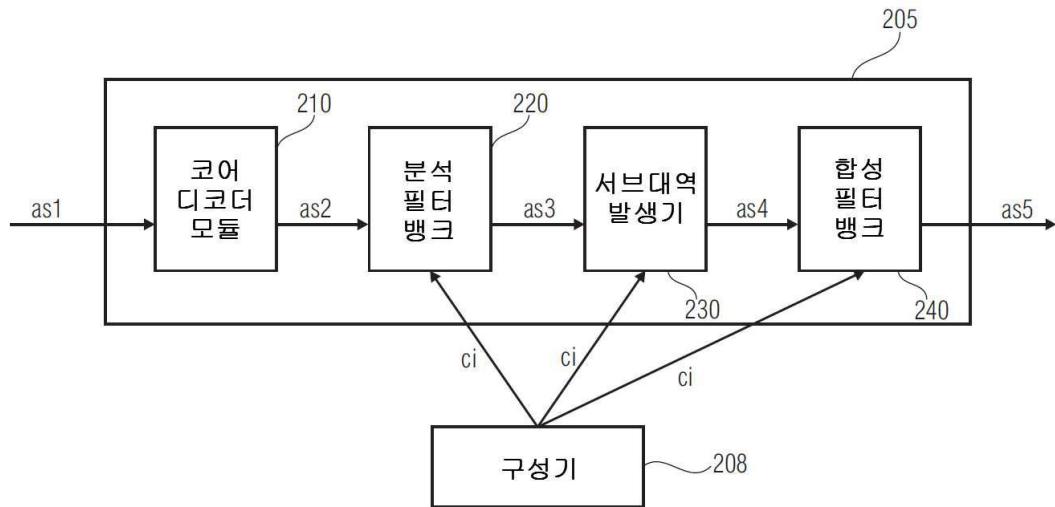
- [0113] 일반적으로, 본 발명의 실시예들은 컴퓨터 프로그램 제품이 컴퓨터 상에서 구동될 때 본 방법들 중 하나의 방법을 수행하기 위해 동작되는 프로그램 코드를 갖는 컴퓨터 프로그램 제품으로서 구현될 수 있다. 프로그램 코드는 예컨대 머신 판독가능한 캐리어 상에 저장될 수 있다.
- [0114] 다른 실시예들은 머신 판독가능한 캐리어 상에서 저장되는, 여기서 설명된 방법들 중 하나의 방법을 수행하기 위한 컴퓨터 프로그램을 포함한다.
- [0115] 다시 말하면, 본 발명의 방법의 실시예는, 따라서, 컴퓨터 상에서 컴퓨터 프로그램이 구동될 때, 여기서 설명된 방법들 중 하나의 방법을 수행하기 위한 프로그램 코드를 갖는 컴퓨터 프로그램이다.
- [0116] 본 발명의 방법들의 추가적인 실시예는, 이에 따라 여기서 설명된 방법들 중 하나의 방법을 수행하기 위한 컴퓨터 프로그램이 기록되어 있는 데이터 캐리어(또는 디지털 저장 매체, 또는 컴퓨터 판독가능한 매체)이다.
- [0117] 본 발명의 방법의 추가적인 실시예는, 이에 따라 여기서 설명된 방법들 중 하나의 방법을 수행하기 위한 컴퓨터 프로그램을 표현한 신호들의 시퀀스 또는 데이터 스트림이다. 신호들의 시퀀스 또는 데이터 스트림은 데이터 통신 접속, 예컨대 인터넷을 통해 전송되도록 구성될 수 있다.
- [0118] 추가적인 실시예는 여기서 설명된 방법들 중 하나의 방법을 수행하도록 구성되거나 적응된 처리수단, 예컨대 컴퓨터, 또는 프로그램가능 논리 디바이스를 포함한다.
- [0119] 추가적인 실시예는 여기서 설명된 방법들 중 하나의 방법을 수행하기 위한 컴퓨터 프로그램이 설치된 컴퓨터를 포함한다.
- [0120] 몇몇의 실시예들에서, 프로그램가능한 논리 디바이스(예컨대 필드 프로그램가능한 게이트 어레이)는 여기서 설명된 방법들의 기능들 모두 또는 그 일부를 수행하기 위해 이용될 수 있다. 몇몇의 실시예들에서, 여기서 설명된 방법들 중 하나의 방법을 수행하기 위해 필드 프로그램가능한 게이트 어레이가 마이크로프로세서와 협동할 수 있다. 일반적으로, 본 방법들은 바람직하게는 임의의 하드웨어 장치에 의해 수행된다.
- [0121] 상술한 실시예들은 본 발명의 원리들에 대한 일례에 불과하다. 여기서 설명된 구성 및 상세사항의 수정 및 변형은 본 발명분야의 당업자에게 자명할 것으로 이해된다. 그러므로, 본 발명은 계류중인 본 특허 청구항들의 범위에 의해서만 제한이 되며 여기서의 실시예들의 설명 및 해설을 통해 제시된 특정한 세부사항들에 의해서는 제한되지 않는다는 것이 본 취지이다.

도면

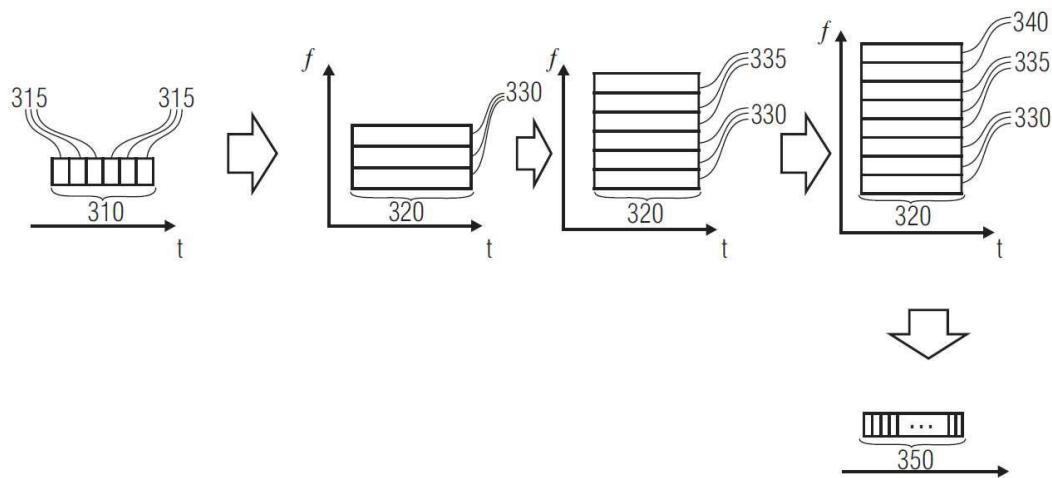
도면1



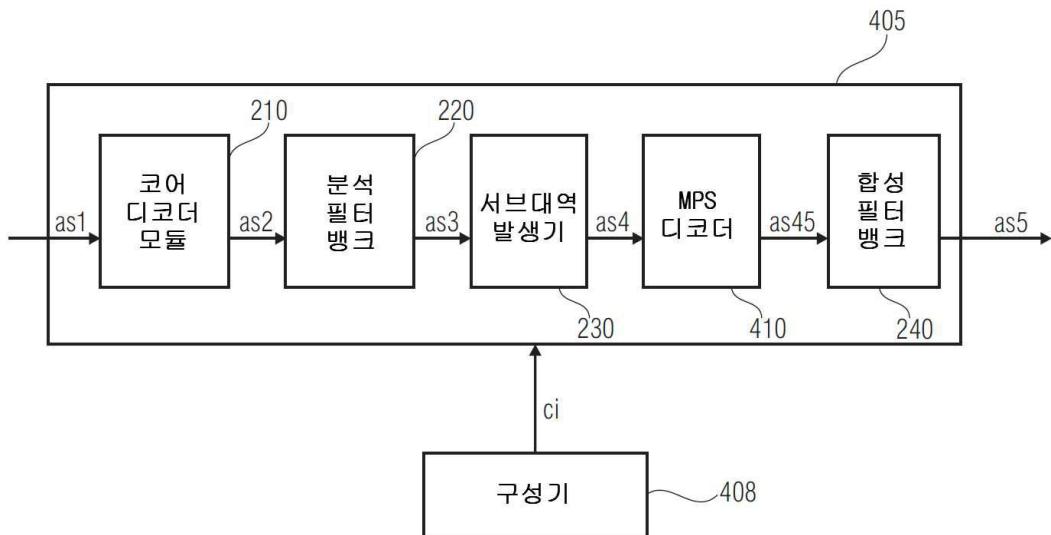
도면2



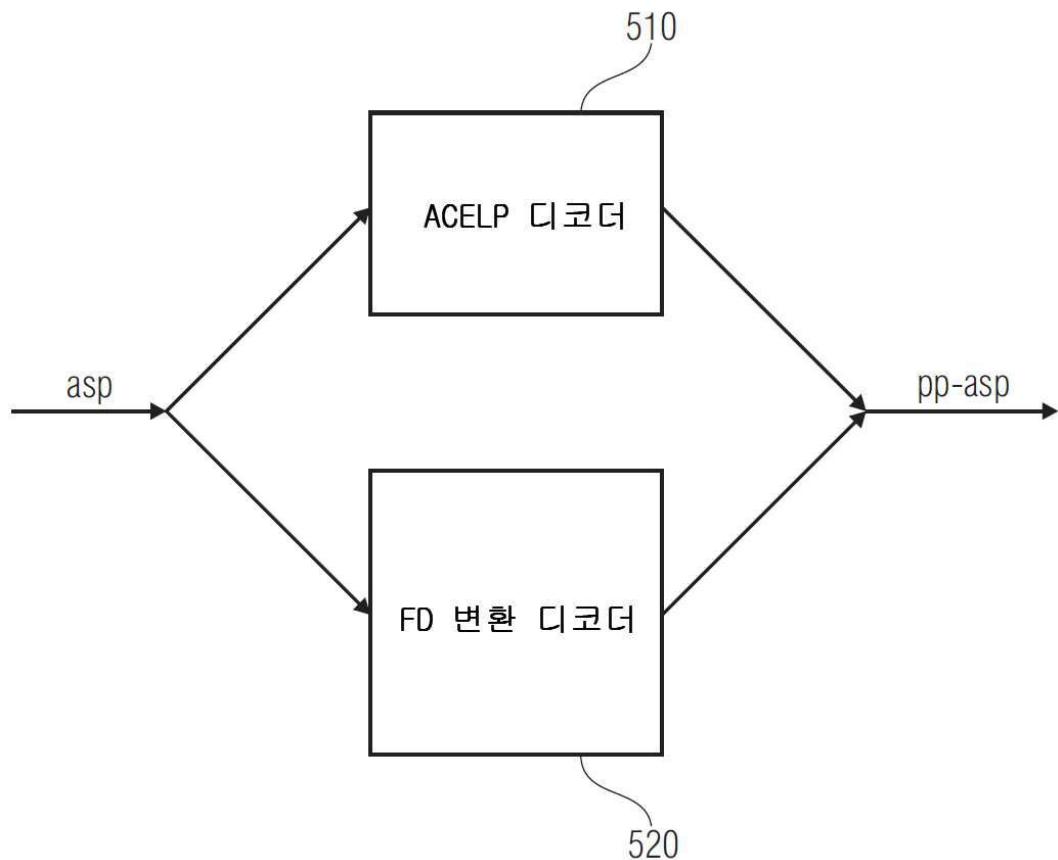
도면3



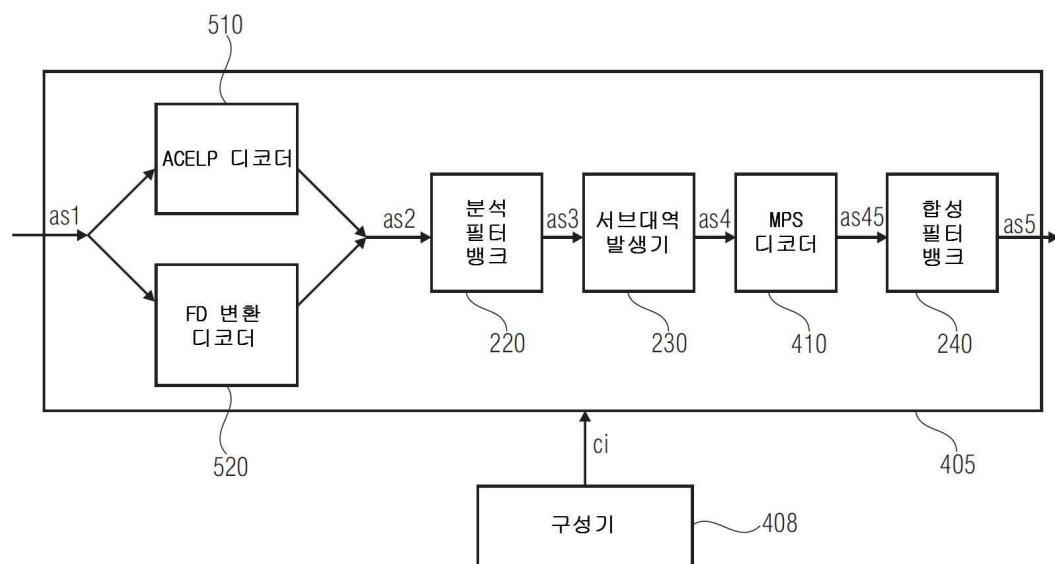
도면4



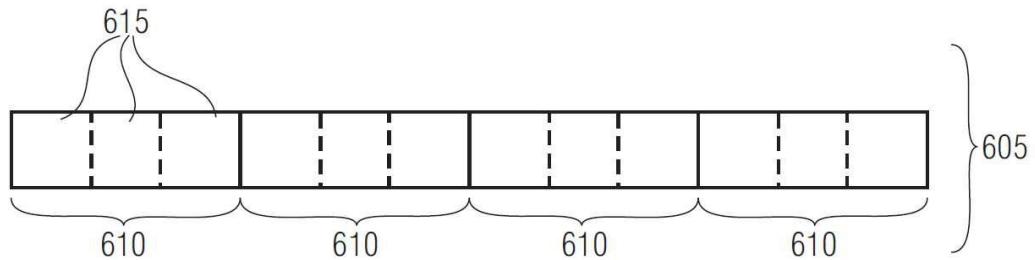
도면5a



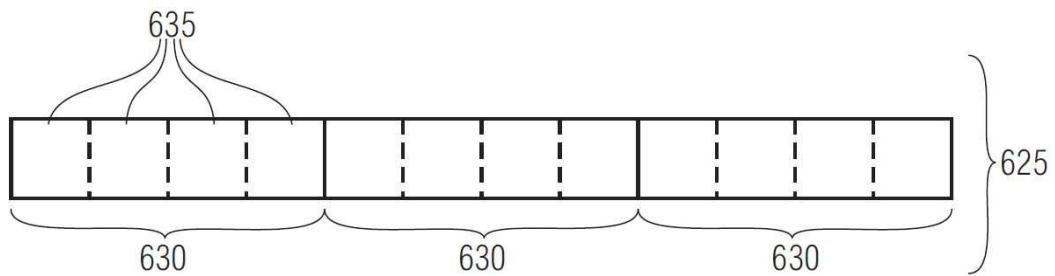
도면5b



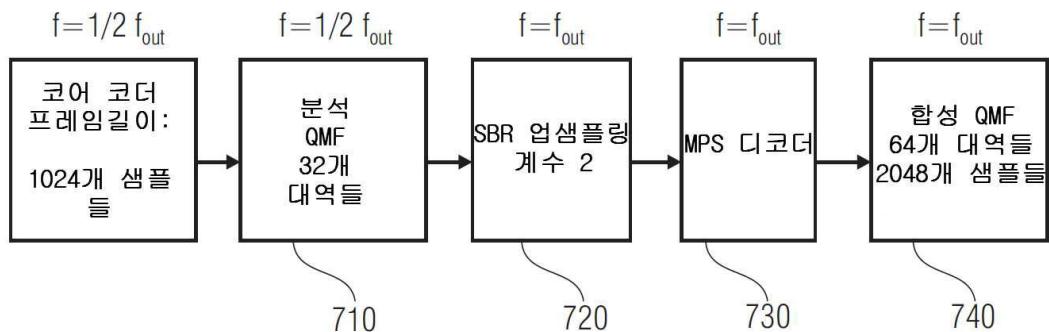
도면6a



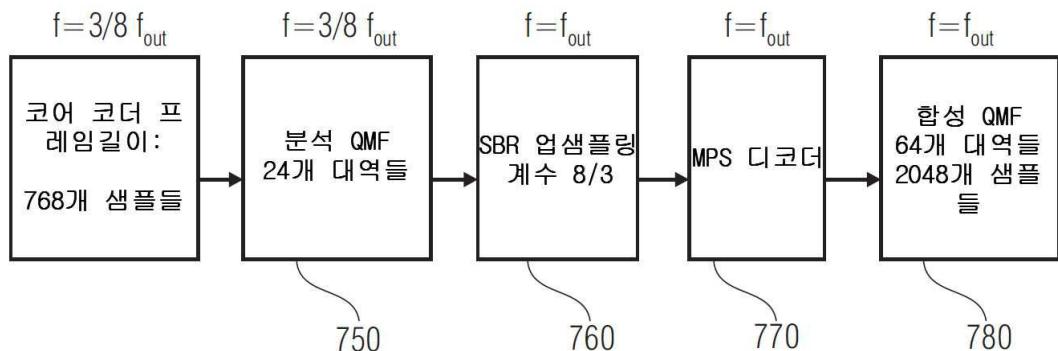
도면6b



도면7a

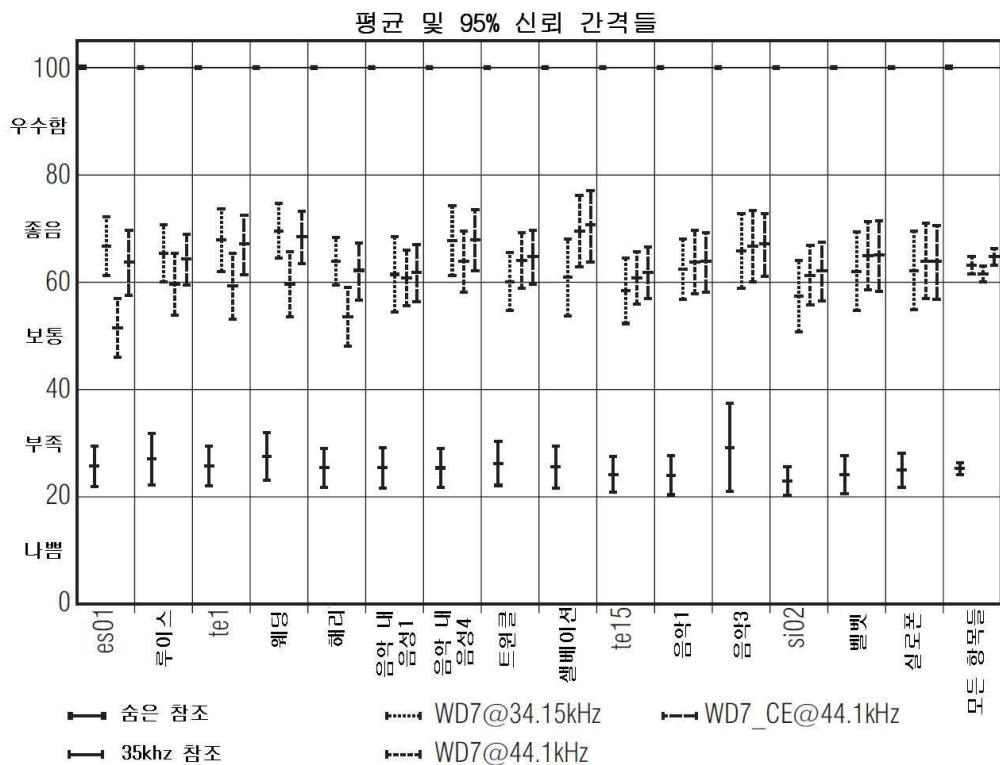


도면7b



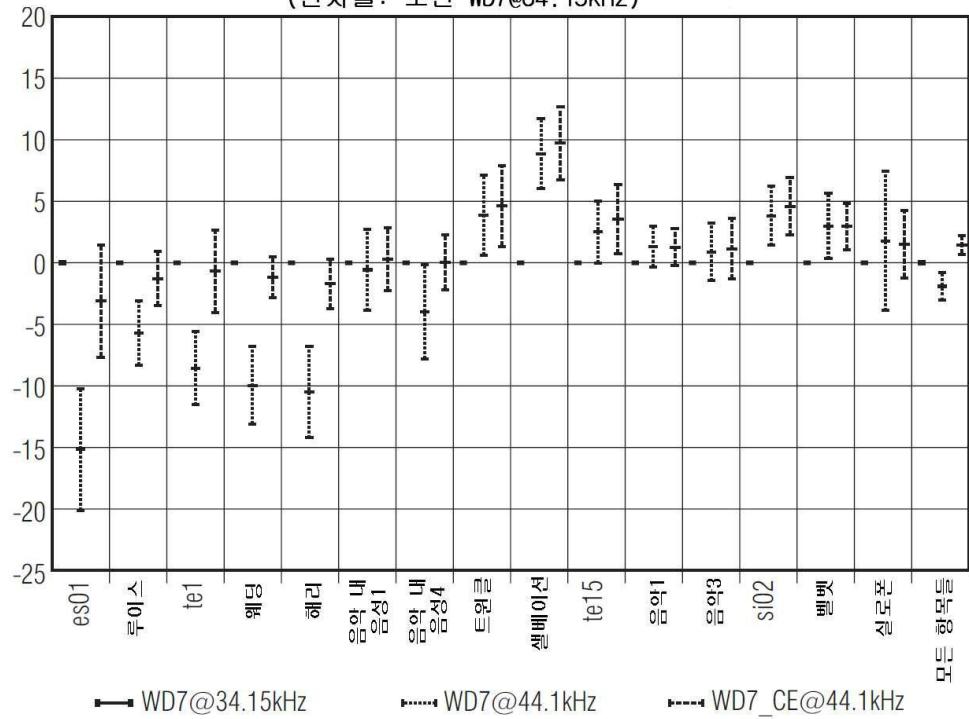
도면8a

세션 1AQ_테스트24m, 22개 주제들, 스튜던트-t 분포



도면8b

세션 1AQ_테스트24m, 22개 주제들, 스튜던트-t 분포

평균 및 95% 신뢰 간격들
(편차들: 조건-WD7@34.15kHz)

도면9

