



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2016-0024924
(43) 공개일자 2016년03월07일

- (51) 국제특허분류(Int. C1.)
G10L 19/02 (2006.01) *G10L 19/025* (2013.01)
G10L 19/03 (2013.01) *G10L 21/0388* (2013.01)
- (52) CPC특허분류
G10L 19/02 (2013.01)
G10L 19/025 (2013.01)
- (21) 출원번호 10-2016-7001383
- (22) 출원일자(국제) 2014년07월15일
 심사청구일자 2016년01월18일
- (85) 번역문제출일자 2016년01월18일
- (86) 국제출원번호 PCT/EP2014/065118
- (87) 국제공개번호 WO 2015/010953
 국제공개일자 2015년01월29일
- (30) 우선권주장
 13177346.7 2013년07월22일
 유럽특허청(EPO)(EP)
 (뒷면에 계속)
- (71) 출원인
 프라운호퍼 게젤샤프트 주르 폴트데룽 테어 안젠
 반텐 포르슝 에. 베.
 독일 80686 뮌헨 한자슈트라쎄 27 체
- (72) 발명자
 디쉬, 사샤
 독일 90766 뮌헨 빌헬름슈트라쎄 70
 가이거, 랄프
 독일 91052 에를랑겐 야콥-헤르츠-베그 36
 (뒷면에 계속)
- (74) 대리인
 윤의섭, 김수진

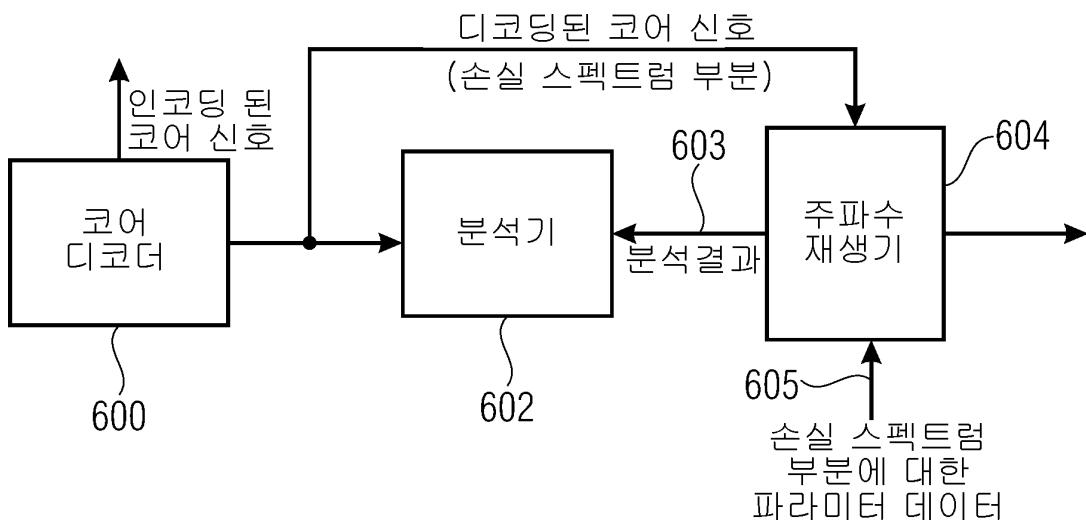
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 인코딩된 오디오 신호를 디코딩하기 위한 장치, 방법 및 컴퓨터 프로그램

(57) 요 약

디코딩된 코어 신호를 얻기 위해 상기 인코딩된 코어 신호를 디코딩하는 코어 디코더(600); 분석 결과(603)를 제공하기 위해 주파수 재생 작업 전 또는 후 상기 디코딩된 코어 신호를 분석하는 분석기(602); 및 상기 분석 결과(603), 파라미터 데이터(605) 및 상기 디코딩된 코어 신호의 스펙트럼 부분을 이용하여 상기 디코딩된 코어 신호에 포함되지 않는 스펙트럼 부분들을 재생하는 주파수 재생기(604);를 포함하는, 인코딩된 코어 신호 및 파라미터 데이터를 포함하는 인코딩된 오디오 신호를 디코딩하기 위한 장치.

대 표 도 - 도6a



(52) CPC특허분류

G10L 19/03 (2013.01)

G10L 21/0388 (2013.01)

(72) 발명자

헵리히, 크리스티안

독일 91054 에를랑겐 하우프트슈트라쎄 68

나겔, 프레드리크

독일 90425 뉘른베르크 빌헬름스하베너 슈트라쎄
72

노이кам, 크리스티안

독일 90562 칼호로이트 바이스가쎄 24

슈미트, 콘스탄틴

독일 90489 뉘른베르크 히르바겐슈트라쎄 21

피셔, 미하엘

독일 91054 에를랑겐 하아크슈트라쎄 24

(30) 우선권주장

13177350.9 2013년07월22일

유럽특허청(EPO)(EP)

13177353.3 2013년07월22일

유럽특허청(EPO)(EP)

13177348.3 2013년07월22일

유럽특허청(EPO)(EP)

13189382.8 2013년10월18일

유럽특허청(EPO)(EP)

명세서

청구범위

청구항 1

인코딩된 코어 신호 및 파라미터 데이터를 포함하는 인코딩된 오디오 신호를 디코딩하기 위한 장치에 있어서, 디코딩된 코어 신호를 얻기 위해 상기 인코딩된 코어 신호를 디코딩하는 코어 디코더(600); 분석 결과(603)를 제공하기 위해 주파수 재생 작업 전 또는 후 상기 디코딩된 코어 신호를 분석하는 분석기(602); 및 상기 분석 결과(603), 파라미터 데이터(605) 및 상기 디코딩된 코어 신호의 스펙트럼 부분을 이용하여 상기 디코딩된 코어 신호에 포함되지 않는 스펙트럼 부분들을 재생하는 주파수 재생기(604);를 포함하는, 인코딩된 코어 신호 및 파라미터 데이터를 포함하는 인코딩된 오디오 신호를 디코딩하기 위한 장치.

청구항 2

제1항에 있어서,

상기 분석기(614)는 주파수 재생 전 또는 후에 상기 디코딩된 코어 신호의 하나 이상의 지역적 스펙트럼 최소값들을 위치시키기 위해 주파수 재생 전 또는 후 상기 디코딩된 코어 신호를 분석하도록 구성되며, 상기 분석기(614)는 상기 로컬 스펙트럼 최소값들을 식별하는 상기 분석 결과(603)를 제공하도록 구성되며, 상기 주파수 재생기(604, 616)는 상기 스펙트럼 부분을 재생하도록 구성되며, 상기 재생된 스펙트럼 부분의 또는 상기 디코딩된 신호의 상기 스펙트럼 부분의 주파수 타일 경계들이 하나 이상의 스펙트럼 최소값들에서 설정되는 것을 특징으로 하는 장치.

청구항 3

제1항 또는 제2항에 있어서,

상기 주파수 재생기(604)는 예비 재생 신호(703)를 발생시키도록 구성되고, 상기 분석기(602)는 아티팩트-생성 신호 부분들을 검출하기 위해 상기 예비 재생 신호를 분석(704)하도록 구성되며, 상기 주파수 재생기(604)는 상기 예비 재생 신호를 조작하는 또는 상기 재생된 신호의 아티팩트-생성 신호를 감소 또는 제거하기 위해 상기 예비 재생에 대한 파라미터들과 다른 파라미터들로 추가 재생을 수행하는 조작기(722)를 더 포함하는 것을 특징으로 하는 장치.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서,

상기 주파수 재생기(604)는 재생된 스펙트럼 부분들을 얻기 위해 상기 디코딩된 코어 신호의 스펙트럼 부분을 이용하여 상기 디코딩된 코어 신호에 포함되지 않는 스펙트럼 부분들을 갖는 예비 재생 신호(703)를 재생하도록 구성되며,

상기 주파수 재생기(604)는 상기 디코딩된 코어 신호 및 재생된 스펙트럼 부분 사이의 주파수 경계 근처에 또는 상기 디코딩된 코어 신호의 동일 또는 상이한 스펙트럼 부분들을 이용하여 발생된 두개의 재생된 스펙트럼 부분들 사이의 주파수 경계 근처에 아티팩트 생성 신호 부분들을 검출(704)하도록 구성되고,

상기 주파수 재생기(604)는 상기 예비 재생 신호를 조작하는 또는 상기 예비 재생 신호를 발생시키기 위해 이용

되는 상기 제어 데이터와 다른 상기 조작된 제어 데이터를 이용하여 재생된 신호를 새로 발생시키기 위해 제어 데이터를 조작하는 조작기(722)를 더 포함하는 것을 특징으로 하는 장치.

청구항 5

제4항에 있어서,

상기 주파수 재생기(604)는 미가공(raw) 스펙트럼 부분들을 얻기 위해 상기 디코딩된 신호의 하나 이상의 스펙트럼 부분들을 이용하여 상기 스펙트럼 부분들을 유도하도록 구성되는 타일 발생기(820)를 포함하며,

상기 조작기(824)는 조작된 스펙트럼 부분들을 얻기 위해 상기 주파수 타일 발생기(820) 또는 상기 미가공(raw) 스펙트럼 부분들을 조작하도록 구성되며,

상기 주파수 재생기(604)는 상기 파라미터 데이터(605)를 이용하여 상기 조작된 스펙트럼 부분들의 엔벨로프 조정을 수행하도록 구성되는 스펙트럼 엔벨로프 조정기(826)를 더 포함하는 것을 특징으로 하는 장치.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서,

상기 분석기(602)는 주파수 검출 범위에 위치된 음조 신호 부분들을 검출하도록 구성되며, 상기 주파수 검출 범위는 1 바크(Bark)인 미리 결정된 검출 대역폭 또는 복원 주파수 범위 또는 소스 주파수 범위의 대역폭의 20%보다 작은 미리 결정된 검출 대역폭만큼 복원 범위 내의 인접 주파수 타일들 사이에서 또는 복원 범위의 주파수 경계로부터 확장하는 것을 특징으로 하는 장치.

청구항 7

제6항에 있어서,

상기 조작기(824)는 상기 미리 결정된 검출 대역폭에서 재생된 신호의 음조 부분들을 포함하는 스펙트럼 부분들을 감쇠시키거나 제거(708)하도록 구성되는 것을 특징으로 하는 장치.

청구항 8

제7항에 있어서,

상기 조작기(722, 824)는 주파수로 상기 음조 부분(802)의 시작에 위치된 주파수 시작 스펙트럼 부분 및 주파수로 상기 음조 부분(802)의 끝 주파수에 위치된 끝 스펙트럼 부분을 결정하고, 보간된 신호 부분을 얻기 위해 상기 시작 주파수 및 상기 끝 주파수 사이를 보간(804)하고, 상기 보간된 신호 부분(806)에 의해 시작 주파수 및 끝 주파수 사이의 음조 부분을 교체하도록 구성되는, 장치.

청구항 9

제7항에 있어서,

상기 조작기(822)는 상기 재생된 스펙트럼 부분들(810)의 비-음조 신호 부분 또는 상기 디코딩된 코어 신호의 비-음조 신호 부분에 의해 결정되는 에너지를 갖는 스펙트럼 라인들(808)을 랜덤으로 또는 비-랜덤으로 발생시키도록 구성되는 것을 특징으로 하는 장치.

청구항 10

제4항 내지 제9항 중 어느 한 항에 있어서,

상기 분석기는 특정 주파수에서 상기 아티팩트-생성 신호 부분들을 검출하도록 구성되고, 상기 조작기(722, 824)는 타일 발생기를 제어하도록 구성되어 상기 타일 발생기가 상기 재생된 스펙트럼 부분의 주파수 경계 또는 상기 디코딩된 코어 신호의 스펙트럼 부분의 주파수 경계를 변경하도록 구성되고 상기 아티팩트-생성 신호 부분이 더 적게 아티팩트-생성하거나 아티팩트-생성을 하지 않도록 하는 것을 특징으로 하는 장치.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서,

상기 분석기(602)는 재생된 신호의 최대 주파수 경계에서 또는 상기 디코딩된 코어 신호의 동일 또는 상이한 스펙트럼 부분들을 이용하여 재생된 두개의 재생된 스펙트럼 부분들 사이의 주파수 경계에서 또는 상기 디코딩된 코어 신호의 주파수 경계에서 재생된 신호의 또는 상기 디코딩된 코어 신호의 스펙트럼 부분의 피크 부분의 중간-분할(midway-splitting)을 검출하도록 구성되며,

상기 주파수 재생기는 상기 분할이 감소되거나 제거되도록 상기 디코딩된 신호의 동일 또는 상이한 스펙트럼 부분들을 이용하여 발생된 두개의 재생된 스펙트럼 부분들 사이의 주파수 경계 또는 상기 재생된 신호 및 상기 디코딩된 코어 신호 사이의 주파수 경계를 변경하거나 상기 최대 주파수를 변경하도록 구성되는 것을 특징으로 하는 장치.

청구항 12

상기 제1항 내지 제11항 중 어느 한 항에 있어서,

상기 주파수 재생기(604)는 타일 발생기(820)를 포함하며, 상기 타일 발생기(820)는 상기 디코딩된 코어 신호의 동일 또는 상이한 스펙트럼 부분들을 이용하여 제1스펙트럼 부분에 대한 제1주파수 타일 및 제2스펙트럼 부분에 대한 제2주파수 타일을 발생시키도록 구성되며,

상기 제2주파수 타일의 상기 하부 주파수 경계는 상기 제1주파수 타일의 상부 주파수 경계와 일치하고,

상기 분석기(602)는 상기 피크 스펙트럼 부분이 제2주파수 타일의 하부 주파수 경계 또는 제1주파수 타일의 상부 주파수 경계 또는 제1주파수 타일의 하부 주파수 경계 및 상기 디코딩된 코어 신호의 미리 결정된 캡 필링 시작 주파수(309)에 의해 클리핑(clipped)되는지 여부를 검출하도록 구성되며,

상기 조작기(824)는 상기 타일 발생기(820)가 상기 클리핑이 감소되거나 제거되도록 변경된 수정된 시작 또는 정지 주파수 경계들을 갖는 수정된 주파수 타일들을 발생시키도록 상기 타일 발생기(820)를 제어하도록 구성되는 것을 특징으로 하는 장치.

청구항 13

제1항 내지 제12항 중 어느 한 항에 있어서,

상기 코어 디코더는 제로(zero) 표현과 다른 스펙트럼 값들에 의해 표현되는 제1스펙트럼 부분들의 제1세트를 포함하는 주파수 영역 디코딩된 스펙트럼 부분들을 얻도록 구성되고 제2스펙트럼 부분들의 제2세트는 스펙트럼 값들에 대해 제로(zero) 표현에 의해 표현되며, 상기 파라미터 정보는 상기 제2스펙트럼 부분들의 제2세트에 대해 제공되며, 상기 주파수 재생기(604)는 제1스펙트럼 부분들의 제1세트에 포함되지 않는 복원 대역 내의 스펙트럼 부분들을 재생하기 위해 제1스펙트럼 부분들의 제1세트로부터 디코딩된 스펙트럼 부분들을 이용하도록 구성되며, 상기 장치는 상기 재생된 스펙트럼 부분들 및 상기 디코딩된 코어 신호의 스펙트럼 부분들을 시간 표현으로 변환하는 주파수-시간 변환기(828)를 더 포함하는 것을 특징으로 하는 장치.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서,

상기 코어 디코더(600)는 변형 이산 코사인 변환(MDCT) 스펙트럼 값들을 출력하도록 구성되고,

상기 주파수-시간 변환기(828)는 이후 얻어지는 MDCT 프레임들에 오버랩-애드(overlap-add) 처리를 적용하는 역 MDCT 변환(512, 514, 516)을 수행하기 위한 프로세서를 포함하는 것을 특징으로 하는 장치.

청구항 15

제1항 내지 제14항 중 어느 한 항에 있어서,

상기 주파수 재생기(604)는 예비 재생 신호를 발생(702)시킬도록 구성되며,

상기 주파수 재생기(604)는 상기 예비 재생 신호의 음조 구성요소들을 검출(704)하도록 구성되며,

상기 주파수 재생기는 재생된 신호를 발생시키기 위해 검출(704)의 결과에 기반하여 상기 복원 범위의 인접 주파수 타일들 사이의 또는 복원 범위 및 소스 범위 사이의 전이 주파수들을 조정하도록 구성되며,

상기 재생기는 상기 전이 주파수들 주변의 검출 범위에 위치되는 음조 구성요소들(708)을 제거하도록 더 구성되며,

상기 주파수 재생기는 상기 전이 주파수들 주변의 크로스-오버 범위에서 제거된 음조 구성요소들을 갖는 신호를 크로스-오버 필터링하기 위한 크로스-오버 필터(710)를 더 포함하며,

상기 주파수 재생기는 상기 파라미터 데이터(605)를 이용하여 상기 크로스-필터의 결과를 스펙트럼 엔벨로프 성형하기 위한 스펙트럼 엔벨로프 성형기(712)를 더 포함하는 것을 특징으로 하는 장치.

청구항 16

인코딩된 코어 신호 및 파라미터 데이터를 포함하는 인코딩된 오디오 신호를 디코딩하기 위한 방법에 있어서,

디코딩된 코어 신호를 얻기 위해 상기 인코딩된 코어 신호를 디코딩(600)하는 단계;

분석 결과(603)를 제공하기 위해 주파수 재생 작업을 수행하기 전 또는 후에 상기 디코딩된 코어 신호를 분석(602)하는 단계; 및

상기 분석 결과(603), 상기 파라미터 데이터(605), 및 상기 디코딩된 코어 신호의 스펙트럼 부분을 이용하여 상기 디코딩된 코어 신호에 포함되지 않는 스펙트럼 부분들을 재생(604)하는 단계;를 포함하는,

인코딩된 코어 신호 및 파라미터 데이터를 포함하는 인코딩된 오디오 신호를 디코딩하기 위한 방법.

청구항 17

컴퓨터 또는 프로세서 상에서 수행될 때, 제16항의 방법을 수행하기 위한 컴퓨터 프로그램.

발명의 설명

기술 분야

[0001] 본 발명은 오디오 코딩/디코딩에 관한 것으로서, 특히 지능형 갭 필링(Intelligent Gap Filling, IGF)을 사용하는 오디오 코딩에 관한 것이다.

배경 기술

[0002] 오디오 코딩은 음향심리학적 지식을 사용하여 오디오 신호들 내의 중복과 무관성(irrelevancy)의 이용을 처리하는 신호 압축의 영역이다. 오늘날 오디오 코덱들은 일반적으로 거의 모든 종류의 오디오 신호의 지각적 투명 코

팅(perceptually transparent coding)을 위하여 약 60 kbps/채널이 필요하다. 새로운 코덱들은 대역폭 확장(BWE)과 같은 기술들을 사용하여 신호 내의 스펙트럼 유사성들을 이용함으로써 코딩 비트레이트를 감소시키는 것을 목표로 한다. 대역폭 확장 전략은 오디오 신호의 고주파수 성분들을 표현하기 위하여 낮은 비트레이트 파라미터 설정을 사용한다. 고주파수 스펙트럼은 저주파수 영역들과 스펙트럼 형상, 텔드 및 원래 신호의 음색(timbre)과 색(color)을 유지하도록 조정된 시간 지속성으로부터 스펙트럼 콘텐츠로 채워진다. 그러한 대역폭 확장 방법들은 오디오 코덱들이 약 24 kbps/채널의 낮은 비트레이트에서도 뛰어난 품질을 유지하는 것을 가능하게 한다.

[0003] 발명의 오디오 코딩 시스템은 광범위한 비트레이트들에서 임의 오디오 신호들을 효율적으로 코딩한다. 반면에, 높은 비트레이트를 위하여, 본 발명의 시스템은 투명성(transparency)에 집중되는데, 그 이유는 낮은 비트레이트 지각적 어노이언스(annoyance)가 최소화되기 때문이다. 따라서, 이용가능한 비트레이트의 주된 공유는 인코더 내의 신호의 바로 지각적으로 가장 관련 있는 구조를 과형 코딩하도록 사용되고, 결과로서 생기는 스펙트럼 캡들은 원래 스펙트럼과 대략 근사치인 신호 콘텐츠를 갖는 디코더 내에 채워진다. 인코더로부터 디코더로 전송되는 전용 부가 정보에 의해 파라미터 구동되는 이른바 스펙트럼 지능형 캡 필링을 제어하기 위하여 매우 제한된 비트 예산(bit budget)이 소모된다.

[0004] 오디오 신호들의 저장 또는 전송은 종종 엄격한 비트레이트 제약(bitrate constraint)의 대상의 된다. 과거에, 코더들은 매우 낮은 비트레이트만이 이용가능할 때 전송된 오디오 대역폭을 극적으로 감소시키도록 강요당했다.

[0005] 현대 오디오 코덱들은 오늘날 대역폭 확장 방법들을 사용함으로써 광대역 신호들을 코딩할 수 있다([1]). 이러한 알고리즘들은 고주파수 콘텐츠의 파라미터 표현에 의존하는데, 이는 고주파수 스펙트럼 영역("패칭(patching)") 내로의 전위(transposition) 및 파라미터 구동 후처리의 적용에 의해 디코딩된 신호의 과형 코딩된 저주파수로부터 발생된다. 대역폭 확장 전략들에서, 주어진 이른바 크로스오버 주파수(cross-over frequency) 위의 고주파수 스펙트럼 영역의 재구성은 종종 스펙트럼 패칭을 기초로 한다. 일반적으로, 고주파수 영역은 다수의 인접한 패치로 분해되고 이러한 패치들 각각은 주어진 크로스오버 주파수 아래의 저주파수 스펙트럼의 대역-통과(band-pass, BP) 영역들로부터 근원된다. 최신 시스템들은 소스(source)로부터 표적 영역으로 인접한 부대역 계수들을 복사함으로써 필터뱅크 표현, 예를 들면 직각 대칭 필터(Quadrature Mirror Filter, QMF) 내에서 패칭을 실행한다.

[0006] 압축 효율을 증가시키고 이에 의해 낮은 비트레이트에서 확장된 오디오 대역폭을 가능하게 하는 오늘날의 오디오 코덱들에서 발견되는 또 다른 기술은 오디오 스펙트럼의 적합한 부분들의 파라미터 구동 합성 대체(parasmetric driven synthetic replacement)이다. 예를 들면, 원래 오디오 신호의 잡음(noise) 같은 신호 부분들은 디코더 내에서 발생되고 부가 정보 파라미터들에 의해 스케일링되는 인공 잡음에 의한 주관적 품질의 실질적 손실 없이 대체될 수 있다. 일례가 MPEG-4 고급 오디오 코딩(AAC)에 포함되는 지각적 잡음 대체(perceptual noise substitution)이다([5]).

[0007] 또한 확장된 오디오 대역폭을 가능하게 하는 또 다른 제공은 MPED-D 통합 음성 오디오 코딩(Unified Speech and Audio Coding, USAC) 내에 포함된 잡음 필링(noise filling) 기술이다([7]). 너무 거친(coarse) 양자화에 기인하여 양자화기(quantizer)의 사각지대(dead-zone)에 의해 추론되는 스펙트럼 캡들(제로들)은 그 뒤에 디코더 내의 인공 잡음으로 채워지고 파라미터-구동 후-처리에 의해 스케일링된다.

[0008] 또 다른 최신 시스템은 정확한 스펙트럼 대체(Accurate Spectral Replacement, ASR)이다([2-4]). 과형 코덱에 더하여, 정확한 스펙트럼 대체는 디코더에서 신호의 지각적으로 중요한 정현파(sinusoidal) 부분들을 저장하는 전용 신호 합성 단계를 이용한다. 또한, [5]에서 설명되는 시스템은 낮은 비트레이트에서 적절한 지각 품질을 갖는 확장된 오디오 대역폭을 가능하게 하기 위하여 과형 코더의 고주파수 영역 내의 정현파 모델링에 의존한다. 이러한 모든 방법은 변형 이산 코사인 변환(MDCT)을 제외하고 제 2 도메인 내로의 데이터의 변환을 포함하며 또한 고주파수 정현파 성분들의 보존을 위한 상당히 복잡한 분석/합성 단계들을 포함한다.

[0009] 도 13a는 예를 들면, 고효율 고급 오디오 코딩에서 사용되는 것과 같은, 대역폭 확장 기술을 위한 오디오 인코더의 개략적인 다이어그램을 도시한다. 라인(1300)에서의 오디오 신호는 저역 통과(low pass, (1302) 및 고역 통과(1304)를 포함하는 필터 시스템 내로 입력된다. 고역 통과 필터(1304)에 의해 출력되는 신호는 파라미터 추출기/코더(1306) 내로 입력된다. 파라미터 추출기/코더(1306)는 예를 들면, 스펙트럼 엔벨로프 파라미터, 잡음 첨가 파라미터, 손실 고조파(missing harmonics) 파라미터, 또는 역 필터링 파라미터와 같은 파라미터들을 계산하고 코딩하도록 구성된다. 이러한 추출된 파라미터들은 비트 스트림 멀티플렉서(1308) 내로 입력된다. 저역 통과 출력 신호는 일반적으로 다운 샘플러(down sampler, 1310) 및 코어 코더(1312)의 기능성을 포함하는 프로세서

내로 입력된다. 저역 통과(1302)는 대역폭이 라인(1300) 상의 원래 입력 오디오 신호에서 발생하는 것보다 상당히 작은 대역폭으로 인코딩되도록 제한한다. 이는 코어 코더에서 발생하는 전체 기능이 감소된 대역폭을 갖는 신호 상에서 운용하여야만 한다는 사실에 기인하여 중요한 코딩 이득을 제공한다. 예를 들면 라인(1300) 상의 오디오 신호의 대역폭이 20 kHz이고 저역 통과 필터(1302)가 바람직하게는 4 kHz의 대역폭을 가질 때, 샘플링 정리(sampling theorem)을 충족시키기 위하여, 다운 샘플러 뒤의 신호는 적어도 40 kHz이어야만 하는 오디오 신호(1300)를 위하여 필요한 샘플링 레이트에 대한 실질적인 감소인, 8 kHz의 샘플링 주파수를 갖는 것이 이론적으로 충분하다.

[0010] 도 13b는 상응하는 대역폭 확장 디코더의 개략적인 다이어그램을 도시한다. 디코더는 비트스트림 멀티플렉서(1320)를 포함한다. 비트스트림 디멀티플렉서(1320)는 코더 디코더(1322)를 위한 입력 신호 및 파라미터 디코더(1324)를 위한 입력 신호를 포함한다. 코어 디코더 출력 신호는 위의 예에서, 8 kHz의 샘플링 레이트, 따라서 4 kHz의 대역폭을 가지며 반면에 완전한 대역폭 재구성을 위하여, 고주파수 재구성기(1330)의 출력 신호는 적어도 40 kHz의 샘플링 레이트를 필요로 하는 20 kHz에 존재하여야만 한다. 이를 가능하게 하기 위하여, 업샘플러(upsampler, 1325)와 필터뱅크(1326)의 기능성을 갖는 디코더 프로세서가 필요하다. 고주파수 재구성기(1330)는 그리고 나서 필터뱅크(1326)에 의해 주파수 분석된 저주파수 신호 출력을 수신하고 고주파수 대역의 파라미터 표현을 사용하여 도 13a의 고역 통과 필터(1304)에 의해 정의되는 주파수 범위를 재구성한다. 고주파수 재구성기(1330)는 저주파수 범위 내의 소스 범위를 사용하여 상부 주파수 범위의 재생, 스펙트럼 엔벨로프 조정, 잡음 첨가 기능 및 상부 주파수 범위 내에 순실 고조파를 도입하기 위한 기능과 같은 몇몇 기능을 가지며 만일 도 13a의 인코더 내에 적용되고 계산되면, 고주파수 범위가 일반적으로 낮은 주파수 범위만큼 음조(tonal)가 아니라는 것을 설명하기 위하여 역 필터링의 작동을 갖는다. 고효율 고급 오디오 코딩에서, 순실 고조파는 디코더면 상에서 재합성되고 정확하게 재구성 대역의 중간에 위치된다. 따라서, 특정 재구성 대역 내에서 결정된 모든 순실 고조파 라인은 원래 신호 내에 위치된 주파수 값들에 위치되지 않는다. 대신에, 그러한 순실 고조파 라인들은 특정 대역의 중앙 내의 주파수들에 위치된다. 따라서, 원래 신호 내의 순실 고조파 라인이 원래 신호의 재구성 대역 경계와 매우 가깝게 위치되었을 때, 이러한 순실 고조파를 대역의 중앙에서 재구성된 신호 내에 위치시킴으로써 도입되는 주파수 내의 오류는 파라미터들이 발생되고 전송된, 개별 재구성 대역의 50%에 가깝다.

[0011] 게다가, 일반적인 오디오 코어 코더들이 스펙트럼 도메인 내에서 운용하더라도, 코어 디코더는 그럼에도 불구하고 그때 다시 필터 뱅크(1326) 기능성에 의해 스펙트럼 도인으로 전환되는, 시간 도메인 신호를 발생시킨다. 이는 부가적인 처리 지연을 도입하고, 처음에 스펙트럼 도메인으로부터 주파수 도메인으로 변환하고 다시 일반적으로 서로 다른 주파수 도메인으로 변환하는 이중 처리(tandem processing)에 기인하는 아티팩트(artifact)들을 도입할 수 있으며, 물론 이는 상당한 양의 계산 복잡도 및 이에 의한 대역폭 확장이 휴대폰, 태블릿 또는 랩톱 컴퓨터 등과 같은 모바일 장치에 적용될 때 문제가 되는, 전력을 필요로 한다.

[0012] 현재 오디오 코덱들은 코딩 전략의 구성 요소로서 대역폭 확장을 사용하여 낮은 비트레이트 오디오 코딩을 실행한다. 대역폭 확장 기술들은 고주파수 콘텐츠만을 대체하도록 한정된다. 게다가, 주어진 크로스오버 주파수 위의 지각적으로 중요한 콘텐츠가 과형 코딩되도록 허용하지 않는다. 따라서, 현재의 오디오 코덱들은 대역폭 확장이 구현될 때 고주파 상세부분 또는 음색을 손실하는데, 그 이유는 신호의 음조 고조파의 정확한 정렬이 대부분의 시스템에서 고려되지 않기 때문이다.

[0013] 현재 기술의 대역폭 확장 시스템들의 또 다른 단점은 오디오 신호의 대역폭 확장의 구현을 위한 새로운 도메인으로의 변환(예를 들면, 변형 이산 코사인 변환으로부터 직각 대칭 필터 도메인으로의 변환)을 위한 필요성이다. 이는 동기화의 복잡성, 부가적인 계산 복잡도 및 증가되는 메모리 요구량에 이르게 한다.

[0014] 오디오 신호들의 저장 또는 전송은 때때로 엄격한 비트레이트 제약의 대상이 된다. 과거에, 코더들은 매우 낮은 비트레이트만이 이용가능할 때 전송된 오디오 대역폭을 대폭 감소시키도록 강요당했다. 현대 오디오 코더들은 이제 대역폭 확장 방법들을 이용함으로써 광대역 신호들을 코딩할 수 있다([1-2]). 이러한 알고리즘들은 고주파수 콘텐츠(HF)의 파라미터 표현에 의존하며, 고주파수 스펙트럼 영역으로의 치환 ("패칭(patching)") 및 파라미터 구동 후처리의 적용에 의해 디코딩된 신호의 과형 코딩된 저주파수 부분(LF)으로부터 발생된다.

[0015] 대역폭 확장 전략들에서, 주어지는 이른바 크로스오버 주파수 위의 고주파수 스펙트럼 영역의 재구성은 종종 스펙트럼 패칭을 기초로 한다. 예를 들어, 인텔리전트 캡 필팅(지능형 캡 필팅(IGF)) 같은, 스펙트럼 캡을 채우기 위해 기능하는 다른 설계들(schemes)은 오디오 신호 고주파수 스펙트럼의 부분들을 재생(regenerate)하기 위해 인접한 소위 스펙트럼 타일들을 이용한다. 일반적으로, 고주파수 영역은 다수의 인접한 패치로 구성되고 이러한 패치들 각각은 주어진 크로스오버 주파수 아래의 저주파수 스펙트럼의 대역-통과 영역들로부터 기원한다. 최신

시스템들은 스스로부터 표적 영역으로 인접한 부대역 계수들의 세트를 복사함으로써 필터뱅크 표현 내에서 패칭을 효율적으로 실행한다. 여전히, 몇몇 신호 컨텐츠에 대해, 고주파수 대역 내의 인접 패치들 및 저주파수 대역으로부터 재생된(복원된, reconstructed) 신호들의 집합(seemblage)은 비팅(beating), 부조화음(dissonance) 및 청각적 거칠음(auditory roughness)를 야기할 수 있다.

[0016] 그래서, [19]에서, 부조화음 가드-대역(보호-대역, guard-band) 필터링이 필터뱅크-기반 BEW 시스템의 컨텍스트에서 제시된다. 스펙트럼 컨텐츠를 제로들(zeros) 또는 노이즈로 교체하고 부조화음의 가능성은 피하기 위해 LF 및 BWE(대역폭 확장)-재생된 고주파수(HF) 사이의 크로스-오버(cross-over) 주파수에서 극사적으로 1 바크(Bark) 대역폭의 노치 필터를 효과적으로 적용하는 것이 제안된다.

[0017] 그러나, [19]에서 제안된 솔루션들은 몇몇 결점들을 갖는다. 첫째로, 스펙트럼 컨텐츠를 제로들(zeros) 또는 노이즈 중 어느 하나로의 엄격한 교체는 신호의 지각적 품질을 악화시킬 수 있다. 게다가, 제안된 처리(프로세싱)은 신호 적응적이 아니며 그래서 몇몇 경우들에서 지각적 품질을 손상시킬 수 있다. 예를 들어, 만약 신호가 트랜지언트들(과도상태, transients)을 포함하는 경우, 이는 전- 및 후-에코들을 야기할 수 있다.

[0018] 두번째로, 부조화가 연속 HF 패치들 사이의 전이들(트랜지션, transitions)에서 일어날 수도 있다. [19]에서 제안된 솔루션은 LF 및 BWE-재생 HF 주파수 사이의 크로스-오버 주파수에서 일어나는 부조화를 처리하는데에만 기능한다.

[0019] 마지막으로, [19]에서 제안된 것처럼 필터 뱅크 기반 시스템들에 반대로, BWE 시스템들(대역폭 확장 시스템들)은 예를 들어, 변형 이산 코사인 변환(MDCT) 같은 변환 기반 실시예들에서 실현될 수도 있다. MDCT 같은 변환들은 스펙트럼 계수들의 밴드패스(대역 통과, band pass) 영역들이 복제되거나 스펙트럼 계수들이 [19]에서 제안된 것처럼 제로(zero)로 설정되는 경우 발생하는 소위 위블링(warbling)[20] 또는 링잉(ringing) 아티팩트들을 야기하기 쉽다.

[0020] 특히, 미국 특히 8,412,365는 필터뱅크 기반 변환 또는 폴딩(folding)에서, 제로(zero)로 설정되는 하나 또는 몇몇 부대역 채널들로 만들어지고 삽입되는 소위 가드-대역들(보호-대역들)을 이용하는 것을 공개한다. 필터뱅크 채널들의 숫자는 보호-대역들에 따라 이용되고, 보호-대역의 대역폭은 0,5 바크(Bark)이다. 이러한 부조화 보호-대역들은 랜덤 백색 노이즈 신호들을 이용하여 부분적으로 복원되고, 즉, 부대역들은 제로(zero, 0) 대신에 백색 노이즈와 함께 입력된다. 보호 대역들은 처리될 현재 신호에 관계없이 삽입된다.

발명의 내용

해결하려는 과제

[0021] 본 발명의 목적은 인코딩된 오디오 신호를 디코딩하는 향상된 개념을 제공하는 것이다.

과제의 해결 수단

[0022] 본 발명의 목적은 청구항 1항의 인코딩된 오디오 신호를 디코딩하기 위한 장치, 청구항 16항의 인코딩된 오디오 신호를 디코딩하기 위한 방법, 또는 17항의 컴퓨터 프로그램에 의해 달성된다.

[0023] 이와 같이, 고정된 디코더-설정과 대조적으로, 패칭 또는 주파수 타일링(tiling)이 고정된 방식으로 수행되는 곳에서, 즉 특정 소스 범위가 코어 신호로부터 취해지고 특정 고정 주파수 경계들이 소스 범위 및 복원 범위 사이의 주파수 또는 복원 범위 내의 타일들 및 두개의 인접 주파수 패치들 사이의 주파수 경계를 설정하기 위해 적용되는 곳에서, 신호-의존 패칭 또는 타일링(tiling)이 수행되고, 여기서 예를 들어, 코어 신호는 코어 신호의 지역적 최소값(local minima)를 찾도록 분석될 수 있고, 이후, 코어 범위는 코어 범위의 주파수 경계들이 코어 신호 스펙트럼에서 지역적 최소값과 일치하도록 선택된다.

[0024] 대안적으로 또는 추가적으로, 신호 분석은 예비 재생 신호 또는 예비 주파수-패칭 또는 타일링 신호에서 수행될 수 있고, 여기서, 예비 주파수 재생 절차 후에, 코어 범위 및 복원 범위 사이의 경계는 그것들이 복원시 비팅(beatting) 아티팩트를 발생시키기 위해 서로 꽈 개까이 있다는 점에서 문제가 되는 음조(tonal) 부분 같이 어떤 한 아티팩트-생성 신호 부분들을 검출하기 위해 분석된다. 대안적으로 또는 추가적으로, 경계들은 음조 부분의 중간-클리핑(halfway-clipping)이 검출되고 음조 부분의 이 클리핑이 그대로 복원시 아티팩트를 생성하는 방식으로 검사될 수도 있다. 이러한 절차들을 피하기 위해, 복원 범위 및/또는 소스 범위의 및/또는 복원 범위의 두 개의 개별 주파수 타일들 또는 패치들의 주파수 경계는 새로 설정된 경계들로 복원을 다시 수행하기 위해 신호

조작기(signal manipulator)에 의해 수정(변형, modified)될 수 있다.

[0025] 추가적으로, 또는 대안적으로, 주파수 재생(regeneration)은 주파수 경계들이 그것들 그대로 남겨진다는 점에서 분석 결과에 기반한 재생이며, 복원 범위 내의 두개의 개별 주파수 타일들 또는 폐치들 사이의 또는 복원 범위 및 소스 범위 사이의 주파수 경계들에 가까운 문제가 있는 음조 부분들의 제거 또는 적어도 감쇄가 수행된다. 그러한 음조 부분들은 비팅 아티팩트를 도출하는 또는 중간-클리핑된(halfway-clipped) 음조 부분들일 수 있는 근접한 톤들(tones)일 수 있다.

[0026] 특히, MDCT 같은 비-에너지 보존 변환의 경우, 단일 톤(단일 음조, single tone)은 단일 스펙트럼 라인으로 직접 맵핑되지 않는다. 대신에, 단일 톤은 톤의 위상에 의존하여 특정 진폭들을 갖는 스펙트럼 라인들의 그룹에 맵핑될 것이다. 폐칭 작업이 이러한 음조 부분을 클립핑(clip)할 때, 이는 MDCT 복원기에서 완벽한 복원이 적용되더라도 복원 후에 아티팩트를 도출하게 된다. 이는 MDCT 복원기가 이러한 톤을 정확히 최종적으로 복원하기 위해 톤에 대한 완전한 음조 패턴을 요구한다는 사실 때문이다. 클리핑이 이전에 일어났다는 사실 때문에, 이는 더 이상 가능하지 않으며, 그래서, 시간 변화 위블링 아티팩트가 생성될 것이다. 본 발명에 따른 분석에 기반하여, 주파수 재생기는, 아티팩트를 생성하는 완전한 음조 부분을 감쇠시켜, 또는 이전에 논의된 것처럼, 대응 경계 주파수들을 변경하여, 또는 양쪽 방법들을 적용하여 또는 그러한 음조 패턴들에서 어떠한 미리 알려진 지식에 기반하여 클리핑된 부분을 복원하여, 이러한 상황을 피할 것이다.

[0027] 추가적으로 또는 대안적으로, 크로스-오버 필터링은 제1주파수 타일 및 제2주파수 타일을 스펙트럼적으로 크로스-오버 필터링하기 위해 또는 캡 필팅 주파수로부터 제1타일 정지 주파수까지 확장하는 주파수들을 갖는 제1주파수 타일 및 디코딩된 코어 신호를 스펙트럼적으로 크로스-오버 필터링하기 위해 적용될 수 있다.

[0028] 이 크로스-오버 필터링은 소위 필터 링잉(filter ringing)을 감소시키는데 유용하다.

[0029] 발전된 접근은 MDCT 같은 변환에 기반하여 대역폭 확장(BWE) 내에 적용되는 것이 주로 의도된다. 그럼에도 불구하고 본 발명의 기술들은 일반적으로, 예를 들어, 쿼드라쳐 미리 필터 뱅크(Quadrature Mirror Filter bank)(QMF) 기반 시스템 내에서 유사하게, 특히 예를 들어 실수값(real-valued) QMF 표현 같이, 상기 시스템이 결정적으로(임계적으로, critically) 샘플링되는 경우, 적용될 수 있다.

[0030] 진보적인 접근은 (크로스-오버 주파수 또는 폐치 경계들 같은) 전이 지점들(transition points)에 가까운 스펙트럼 영역들에 신호 컨텐츠가 매우 음조적인(very tonal) 경우 청각적 거칠음, 비팅들 및 부조화들이 일어날 수 있다는 관찰에 기반한다. 그래서, 이러한 구성요소들의 차후 감쇄 또는 제거 및 전이 영역들에서 음조 구성요소들의 신호 적응적 검출로 구성되는 상기 결점들에 대해 제안된 솔루션이 최신 기술에서 발견되었다. 이러한 구성요소들의 감쇄 또는 제거는 제로 또는 노이즈 삽입에 의해 대안적으로, 또는 그러한 구성요소의 풋(foot)에서 풋(foot)까지의 스펙트럼적 보간에 의해 바람직하게 달성될 수 있다. 대안적으로, 전이들의 스펙트럼적 위치는 전이 아티팩트들이 최소화되도록 적응적으로 선택될 신호일 수 있다.

[0031] 게다가, 이러한 기술은 필터 링잉을 감소시키거나 심지어 회피하기 위해 이용될 수도 있다. 특히 트랜지언트-유사 신호들에 대해, 링잉은 청취가능하고 불쾌한 아티팩트이다. 필터 링잉 아티팩트들은 전이 대역(transition band)에서 필터의 소위 브릭-월(brick-wall) 특성에 의해 야기된다(차단 주파수에서 통과 대역으로부터 정지 대역까지의 가파른 전이(transition)). 그러한 필터들은 시간-주파수 변환의 주파수 영역에서 제로(zero)로 하나의 계수 또는 계수들의 그룹을 설정하여 효율적으로 실행될 수 있다. 그래서, 대역폭 확장의 경우에, 우리는 링잉 효과를 감소시키기 위한 제1폐치 및 코어 대역 사이의 또는 폐치들 사이의 각 전이 주파수(transition frequency)에서 크로스-오버 필터를 적용하는 것을 제안한다. 크로스-오버 필터는 적합한 이득 함수(gain functions)들을 이용하여 변환 영역에서 스펙트럼 가중에 의해 실행될 수 있다.

[0032] 본 발명의 추가 관점에 따라, 인코딩된 오디오 신호를 디코딩하기 위한 장치는 코어 디코더, 디코딩된 코어 신호의 스펙트럼 부분을 이용하여 디코딩된 코어 신호에 포함되지 않는 주파수들을 갖는 하나 이상의 스펙트럼 타일들을 발생시키기 위한 타일 발생기 및 디코딩된 코어 신호를 스펙트럼적으로 크로스-오버 필터링하기 위한 크로스-오버 필터 및 타일 및 추가 주파수 타일을 스펙트럼적으로 크로스-오버 필터링하기 위한 또는 캡 필팅 주파수로부터 제1타일 정지 주파수까지 확장하는 주파수들을 갖는 제1주파수 타일, 주파수 타일의 상부 경계 주파수(upper border frequency)에 주파수-인접한 하부 경계 주파수(lower border frequency)를 갖는 추가 주파수 타일(further frequency tile)을 포함한다.

[0033] 바람직하게, 이러한 절차는 MDCT 같은 변환에 기반한 대역폭 확장 내에서 적용되는 것이 의도된다. 그러나, 본 발명은 일반적으로 적용가능하며, 특히 쿼드라쳐 미리 필터 뱅크(QMF)에 의존하는 대역폭 확장 시나리오에서, 특

히 주파수-시간 변환에 따라 또는 시간-주파수 변환에 따라 실수값 QMF 표현이 존재할 때의 예처럼, 시스템이 임계적으로 샘플링될 때 특히 적용가능하다.

[0034] 실시예는 트랜지언트-유사 신호들(transient-like signals)에 대해 특히 유용하며, 이는 그러한 트랜지언트-유사 신호들에 대해, 링잉(ringing)은 청취가능하고 불쾌한 아티팩트이기 때문이다. 필터 링잉 아티팩트들은 전이 대역(transition band)에서 필터의 소위 브릭-월(brick-wall) 특성, 즉 차단 주파수에서 통과 대역으로부터 정지 대역까지의 가파른 전이에 의해 야기된다. 그러한 필터들은 시간-주파수 변환의 주파수 영역에서 하나의 계수 또는 계수들의 그룹을 제로(zero)로 설정하여 효율적으로 실행될 수 있다. 그래서, 본 발명은 이러한 링잉 아티팩트들을 감소시키기 위해 폐치들/타일들 사이의 또는 코어 대역 및 제1폐치/타일 사이의 각 전이 주파수에서 크로스-오버 필터에 의존한다. 크로스-오버 필터는 적합한 이득 함수(gain functions)들을 이용하여 변환 영역에서 스펙트럼 가중에 의해 실행될 수 있다.

[0035] 바람직하게, 크로스-오버 필터는 신호-적응적이며 두개의 필터들, 낮은 스펙트럼 영역에 적용되는, 페이드-아웃 필터(fade-out filter), 높은 스펙트럼 영역에 적용되는 페이드-인 필터를 포함한다. 필터들은 특정 실시예에 의존하여 대칭적 또는 비대칭적일 수 있다.

[0036] 추가 실시예에서, 주파수 타일 또는 주파수 폐치는 크로스-오버 필터링의 유일한 대상이 되는 것은 아니지만, 타일 발생기는 바람직하게는, 크로스-오버 필터링을 수행하기 전에, 전이 주파수들 주변의 전이 범위들에 잔존하는 음조 부분들의 제거 또는 감쇠 및 스펙트럼 최소값들에서 주파수 경계들의 설정을 포함하는 폐치 적응을 수행한다.

발명의 효과

[0037] 오디오 코딩 시스템은 광범위한 비트레이트들에서 임의 오디오 신호들을 효율적으로 코딩한다. 반면에, 높은 비트레이트를 위하여, 본 발명의 시스템은 투명성(transparency)에 집중되는데, 그 이유는 낮은 비트레이트 지각적 어노이언스(annoyance)가 최소화되기 때문이다. 따라서, 이용가능한 비트레이트의 주된 공유는 인코더 내의 신호의 바로 지각적으로 가장 관련 있는 구조를 파형 코딩하도록 사용되고, 결과로서 생기는 스펙트럼 캡들은 원래 스펙트럼과 대략 근사치인 신호 콘텐츠를 갖는 디코더 내에 채워진다. 인코더로부터 디코더로 전송되는 전용 부가 정보에 의해 파라미터 구동되는 이른바 스펙트럼 지능형 캡 필팅을 제어하기 위하여 매우 제한된 비트 예산(bit budget)이 소모된다.

도면의 간단한 설명

[0038] 바람직한 실시 예들이 첨부된 도면들을 참조하여 이후에 논의된다.

도 1a는 오디오 신호를 인코딩하기 위한 장치를 도시한다.

도 1b는 도 1a의 인코더와 매칭하는 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한다.

도 2a는 디코더의 바람직한 일 구현을 도시한다.

도 2b는 인코더의 바람직한 일 구현을 도시한다.

도 3a는 도 1b의 스펙트럼 도메인 디코더에 의해 발생된 것과 같은 스펙트럼의 개략적 표현을 도시한다.

도 3b는 스케일 인자 대역들을 위한 스케일 인자들과 재구성 대역들을 위한 에너지들 및 잡음 필팅 대역을 위한 잡음 필팅 정보 사이의 관계를 나타내는 테이블을 도시한다.

도 4a는 스펙트럼 부분들의 제 1 및 제 2 세트 내로 스펙트럼 부분의 선택을 적용하기 위한 스펙트럼 도메인 인코더의 기능을 도시한다.

도 4b는 도 4a의 기능의 일 구현을 도시한다.

도 5a는 변형 이산 코사인 변환 인코더의 기능을 도시한다.

도 5b는 변형 이산 코사인 변환 기술을 갖는 디코더의 기능을 도시한다.

도 5c는 주파수 재생기의 일 구현을 도시한다.

도 6a는 하나의 실시예에 따라 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한다.

도 6b는 인코딩된 오디오 신호를 디코딩하기 위한 장치의 추가 실시예이다.

도 7a는 도 6a 또는 6b의 주파수 재생기의 바람직한 실시예를 도시한다.

도 7b는 주파수 재생기 및 분석기 사이의 협력의 추가 실시예를 도시한다.

도 8는 주파수 재생기의 추가 실시예를 도시한다.

도 8b는 본 발명의 추가 실시예를 도시한다.

도 9a는 재생 주파수 범위를 위한 에너지 값들을 사용하는 주파수 재생 기술을 갖는 디코더를 도시한다.

도 9b는 도 9a의 주파수 재생의 더 상세한 구현을 도시한다.

도 9c는 도 9b의 기능을 개략적으로 도시한다.

도 9d는 도 9a의 디코더의 또 다른 구현을 도시한다.

도 10a는 도 9a의 디코더와 매칭하는 인코더의 블록 다이어그램을 도시한다.

도 10b는 도 10a의 파라미터 계산기의 또 다른 기능을 나타내기 위한 블록 다이어그램을 도시한다.

도 10c는 도 10a의 파라미터 계산기의 또 다른 기능을 나타내는 블록 다이어그램을 도시한다.

도 10d는 도 10a의 파라미터 계산기의 또 다른 기능을 나타내는 블록 다이어그램을 도시한다.

도 11a는 트랜지언트 주변의 필터 링잉(filter ringing)의 스펙트럼을 도시한다.

도 11b는 대역폭 확장 적용 후 트랜지언트의 스펙트로그램을 도시한다.

도 11c는 필터 링잉 감소를 갖는 대역폭 확장을 적용한 후 트랜지언트의 스펙트럼을 도시한다.

도 12a는 인코딩된 오디오 신호를 디코딩하기 위한 장치의 블록 다이어그램을 도시한다.

도 12b는 음조 신호의 (양식화된) 크기 스펙트럼, 패치/타일 적응없는 카피-업, 아티팩트-생성 음조 부분들의 추가적 제거 및 변화된 주파수 경계들을 갖는 카피-업(copy-up)을 도시한다.

도 12c는 예시 크로스-페이드(cross-face) 함수를 도시한다.

도 13a는 대역폭 확장을 갖는 종래 기술의 인코더를 도시한다.

도 13b는 대역폭 확장을 갖는 종래 기술의 디코더를 도시한다.

도 14a는 크로스-오버 필터를 이용하여 인코딩된 오디오 신호를 디코딩하기 위한 추가 장치를 도시한다.

도 14b는 예시적 크로스-오버 필터의 더 세부적인 설명을 도시한다.

발명을 실시하기 위한 구체적인 내용

[0039] 도 6a는 인코딩된 코어 신호 및 파라미터 데이터를 포함하는 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한다. 상기 장치는 디코딩된 코어 신호를 얻기 위해 인코딩된 코어 신호를 디코딩하는 코어 디코더(600) 및 주파수 재생 작업을 수행하기 전 또는 후에 디코딩된 코어 신호를 분석하기 위한 분석기(602)를 포함한다. 상기 분석기(602)는 분석 결과(603)를 제공하도록 구성된다. 주파수 재생기(604)는 분석 결과(603) 및 손실 스펙트럼 부분들(missing spectral portions)에 대한 엔벨로프 데이터(605), 디코딩된 코어 신호의 스펙트럼 부분을 이용하여 상기 디코딩된 코어 신호에 포함되지 않는 스펙트럼 부분들을 재생하도록 구성된다. 이와 같이, 초기 실시 예들과 대조적으로, 주파수 재생은 디코더-측에서 신호-독립적으로 수행되지 않고, 신호-의존적으로 수행된다. 이는 문제가 존재하지 않을 때, 주파수 재생은 그대로 수행된다는 장점을 가지지만, 문제적 신호 부분들이 존재 할 때, 이는 분석 결과(603)에 의해 검출되고 주파수 재생기(604)는 이후, 예를 들어, 복원 대역 내의 두개의 개별 타일들/패치들 사이의 주파수 경계의 변화 또는 복원 대역 및 코어 영역 사이의 초기 주파수 경계의 변화가 될 수 있는 주파수 재생의 적응된 방식을 수행할 수 있다. 보호-대역(guard-bands)의 실시예에 대조적으로, 이는 특정 절차들이 요구될 때만 아니면, 보호-대역 실시에서처럼, 어떠한 신호-의존성 없이 언제나 수행되는 이점을 가진다.

[0040] 바람직하게, 코어 디코더(600)는 도 6b에서 도시되는 것처럼 엔트로피 (예를 들어, 허프만(Huffman) 또는 산술 디코더) 디코딩 및 역양자화 단계(612)로 실행된다. 코어 디코더(600)은 이후 코어 신호 스펙트럼을 출력하고

상기 스펙트럼은, 도 6a에서 도시되는 것처럼, 시간 영역 신호를 분석할 수도 있는 어떠한 임의의 분석기보다 스펙트럼 분석기로서 실행되는, 도 6a에서의 분석기(602)에 꽤 유사한, 스펙트럼 분석기(614)에 의해 분석된다. 도 6b의 실시예에서, 스펙트럼 분석기는 스펙트럼 신호를 분석하도록 구성되며 소스 대역에서 및/또는 타겟 대역에서, 즉 주파수 폐치들 또는 주파수 타일들에서, 지역적(로컬, local) 최소값들(minima)이 결정된다. 이후, 주파수 재생기(604)는, (616)에서 도시되는 것처럼, 폐치 경계들이 소스 대역 및/또는 타겟 대역에서 최소값들에 위치되는 주파수 재생을 수행한다.

[0041] 이후, 도 7a는 도 6a의 주파수 재생기(604)의 바람직한 실시예를 설명하기 위해 논의된다. 예비 신호 재생기(preliminary signal regenerator, 702)는, 입력으로, 소스 대역으로부터 소스 데이터를, 그리고 추가적으로, 예비 경계 주파수들 같은 예비 폐치 정보(preliminary patch information)를 수신한다. 이후, 예비 재생 신호(preliminary regenerated signal)(703)가 발생되며, 이는 예비 복원 신호(preliminary reconstructed signal)(703) 내의 음조 구성요소들(tonal components)을 검출하는 검출기(detector, 704)에 의해 검출된다. 대안적으로 또는 추가적으로, 소스 데이터(705)는 도 6a의 분석기(602)에 대응하는 검출기에 의해 분석될 수도 있다. 이후, 예비 신호 표현 단계는 필요하지 않을 것이다. 소스 데이터로부터 복원 데이터로의 잘 정의된(well-defined) 맵핑(mapping)이 있을 때, 도 12b의 관점에서 나중에 논의될 것처럼, 두개의 개별적으로 발생된 주파수 타일들 사이의 주파수 경계에서 또는 코어 범위의 상부 경계(upper border)에 근접한(가까운) 음조 부분들이 있는지 여부가, 이후 최소값들 또는 음조 부분들은 소스 데이터만 고려하여 검출될 수도 있다.

[0042] 문제가 있는 음조 구성요소들이 주파수 경계들 부근에서 발견된 경우, 전이 주파수 조정기(transition frequency adjuster, 706)는 복원 대역에서 동일 소스 데이터 및 하나에 의해 발생된 개별 주파수 부분들 사이의 또는 복원 대역 및 코어 대역 사이의 갭 필링 시작 주파수(gap filling start frequency) 또는 크로스-오버 주파수 또는 전이 주파수 같은 전이 주파수의 조정을 수행한다. 블록(706)의 출력 신호는 경계들에서 음조 구성요소들의 제거기(remover, 708)로 포워딩된다. 제거기는 블록(706)에 의해 전이 주파수 조정 이후에 여전히 거기에 있는 잔존 음조 구성요소들을 제거하도록 구성된다. 제거기(708)의 결과는 이후 필터 링잉 문제(filter ringing problem)을 다루기 위해 크로스-오버 필터(710)에 포워딩되고 크로스-오버 필터(cross-over filter, 710)의 결과는 이후 복원 대역의 스펙트럼 엔벨로프 성형(spectral envelope shaping)을 수행하는 스펙트럼 엔벨로프 성형 블록(712)로 입력된다.

[0043] 도 7a의 맥락에서 논의된 것처럼, 블록(704)의 음조 구성요소들의 검출은 예비 복원 신호(703) 또는 소스 데이터(705) 양쪽에서 수행될 수 있다. 이 실시예는 도 7b에서 도시되며, 여기서 예비 재생 신호는 블록(718)에서 보여지듯이 생성된다. 도 7a의 신호(703)에 대응하는 신호는 이후 아티팩트-생성 구성요소들을 검출하는 검출기(detector, 720)로 포워딩된다. 비록 검출기(720)가 도 7a에서 (704)로 도시되는 것처럼 주파수 경계들에서 음조 구성요소들을 검출하는 검출기(detector)가 되도록 구성되지만, 상기 검출기는 다른 아티팩트-생성 구성요소들을 검출하도록 실행될 수도 있다. 그러한 스펙트럼 구성요소들은 음조 구성요소들보다는 다른 구성요소들일 수 있고 아티팩트들이 생성되었는지 여부에 대한 검출은 어떠한 것이 아티팩트-생성 구성요소들을 제공하였는지를 찾아내기 위해 상이한 재생들을 시도하고 상이한 재생 결과들을 비교하여 수행될 수 있다.

[0044] 검출기(720)는 이제 상기 신호, 즉 예비 재생 신호,를 조작하는 조작기(manipulator, 722)를 제어한다. 이러한 조작은 라인(line, 723)에 의해 예비 재생 신호를 실제로 처리하여 또는, 지금 예를 들어, 라인(724)에 의해 도시되는 것처럼 수정된 전이 주파수들과 함께, 재생을 새로이 수행하여 처리될 수 있다.

[0045] 조작 절차의 하나의 실시예는 도 7a에서 도시되는 것처럼 전이 주파수가 조정되는 것이다. 추가 실시예가 도 8a에서 도시되는 것처럼, 이는 도 7a의 블록(706)과 함께 또는 블록(706) 대신에 수행될 수 있다. 검출기(802)는 문제가 있는 음조 부분의 시작 및 끝 주파수들을 검출하도록 제공된다. 이후, 보간기(interpolator, 804)는 보간 그리고, 바람직하게는 스펙트럼 범위 내의 음조 부분의 시작 및 끝 사이를 복합 보간(complex interpolating)하도록 구성된다. 이후, 블록(806)에 의해 도 8a에서 도시되는 것처럼, 음조 부분은 보간 결과에 의해 교체된다.

[0046] 대안적 실시예는 블록(808, 810)에 의해 도 8a에서 도시된다. 보간을 수행하는 대신에, 스펙트럼 라인들(808)의 임의 발생이 음조 부분의 시작 및 끝 사이에서 수행된다. 이후, 임의로 발생된 스펙트럼 라인들의 에너지 조정은 (810)에서 도시되는 것처럼 수행되고, 에너지가 인접 비-음조 스펙트럼 부분들에 유사하도록 랜덤으로 발생된 스펙트럼 라인들이 설정된다. 이후, 음조 부분은 엔벨로프-조정되어 랜덤으로 발생된 스펙트럼 라인들에 의해 교체된다. 스펙트럼 라인들은 가능한한, 아티팩트가 없는 교체 신호를 제공하기 위해 랜덤으로(임의로, randomly) 발생되거나 또는 슈도 랜덤(pseudo randomly)으로 발생된다.

[0047]

추가 실시예는 도 8b에서 도시된다. 도 6a의 주파수 재생기(604) 내에 위치되는 주파수 타일 발생기는 블록(820)에서 도시된다. 주파수 타일 발생기는 미리 결정된 주파수 경계들을 이용한다. 이후, 분석기는 주파수 타일 발생기에 의해 발생된 신호를 분석하며, 주파수 타일 발생기(frequency tile generator, 820)는 바람직하게 복수의 주파수 타일들을 발생시키기 위해 복수의 타일링 작업들(tiling operations)을 수행하도록 구성된다. 이후, 도 8b의 조작기(824)는 분석기(822)에 의해 출력되는 분석 결과에 따라 주파수 타일 발생기의 결과를 조작한다. 조작은 주파수 경계들의 변화 또는 개별 부분들의 감쇠일 수 있다. 이후, 스펙트럼 엔벨로프 조정기(spectral envelope adjuster, 826)는 도 6a의 맥락에서 이미 논의된 것처럼 파라미터 정보(605)를 이용하여 스펙트럼 엔벨로프 조정을 수행한다.

[0048]

블록(826)에 의해 출력된 스펙트럼적으로 조정된 신호는, 제1스펙트럼 부분들, 즉, 코어 디코더(600)의 출력 신호의 스펙트럼 표현을, 추가적으로, 수신하는 주파수-시간 변환기로 입력된다. 주파수-시간 변환기(828)의 출력은 이후 오디오 렌더링(audio rendering)의 확성기에 전송되도록 또는 저장되도록 이용될 수 있다.

[0049]

본 발명은 도 13a, 13b에서 도시되는 것처럼 알려진 주파수 재생 절차들에도 적용될 수 있고 또는 도 1a 내지 5b 및 9a 내지 10d에 관해 이후 설명되는, 지능형 캡 필팅 맥락 내에서 바람직하게 적용될 수 있다.

[0050]

도 1a는 오디오 신호(99)를 인코딩하기 위한 장치를 도시한다. 오디오 신호(99)는 샘플링 레이트를 갖는 오디오 신호를 시간 스펙트럼 변환기(time spectrum converter)에 의해 출력되는 스펙트럼 표현(101)으로 변환하기 위한 시간 스펙트럼 변환기(100) 내로 입력된다. 스펙트럼(101)은 스펙트럼 표현(101)을 분석하기 위하여 스펙트럼 분석기(102) 내로 입력된다. 스펙트럼 분석기(102)는 제 1 스펙트럼 해상도로 인코딩되려는 제 1 스펙트럼 부분들(103)의 제 1 세트 및 제 2 스펙트럼 해상도로 인코딩되려는 제 2 스펙트럼 부분들(105)의 다른 제 2 세트를 결정하도록 구성된다. 제 2 스펙트럼 해상도는 제 1 스펙트럼 해상도보다 작다. 제 2 스펙트럼 부분들(105)의 제 2 세트는 제 2 스펙트럼 해상도를 갖는 스펙트럼 엔벨로프 정보를 계산하기 위한 파라미터 계산기 또는 파라미터 코더(104) 내로 입력된다. 게다가, 제 1 스펙트럼 해상도를 갖는 제 1 스펙트럼 부분들의 제 1 인코딩된 표현을 발생시키기 위하여 스펙트럼 도메인 오디오 코더(105)가 제공된다. 게다가, 파라미터 계산기/파라미터 코더(106)가 제 2 스펙트럼 부분들의 제 2 세트의 제 2 인코딩된 표현을 발생시키도록 구성된다. 제 1 인코딩된 표현(107)과 제 2 인코딩된 표현(109)은 비트 스트림 멀티플렉서 또는 비트 스트림 포머(108) 내로 입력되고 블록(108)이 최종적으로 저장 장치 상으로의 전송 또는 저장을 위하여 인코딩된 오디오 신호를 출력한다.

[0051]

일반적으로, 도 3a의 306과 같은 제 1 스펙트럼 부분은 307a, 307b와 같은 두 개의 제 2 스펙트럼 부분에 의해 둘러싸일 것이다. 이는 코어 코더 주파수 범위가 대역 제한적인 고효율 고급 오디오 코딩에서는 그렇지 않다.

[0052]

도 1b는 도 1a의 인코더와의 디코더 매칭을 도시한다. 제 1 인코딩된 표현(107)은 제 1 스펙트럼 부분들이 제 1 세트의 제 1 디코딩된 표현을 발생시키기 위하여 스펙트럼 도메인 오디오 디코더(112) 내로 입력되며, 디코딩된 표현은 제 1 스펙트럼 해상도를 갖는다. 게다가, 제 2 인코딩된 표현(109)은 제 1 스펙트럼 해상도보다 낮은 제 2 스펙트럼 해상도를 갖는 제 2 스펙트럼 부분들의 제 2 세트의 제 2 디코딩된 표현을 발생시키기 위하여 파라미터 디코더(114) 내로 입력된다.

[0053]

디코더는 제 1 스펙트럼 부분을 사용하여 제 1 스펙트럼 해상도를 갖는 재구성된 제 2 스펙트럼 부분을 발생시키기 위한 주파수 재생기(116)를 더 포함한다. 주파수 재생기(116)는 타일 필팅 운용을 실행하는데, 즉 타일 또는 제 1 스펙트럼 부분들이 제 1 세트를 사용하고 이러한 제 1 스펙트럼 부분들이 제 1 세트를 제 2 스펙트럼 부분을 갖는 재구성 범위 또는 재구성 대역 내로 복사하며 일반적으로 파라미터 디코더(114)에 의해, 즉 제 2 스펙트럼 부분들의 제 2 세트에 대한 정보를 사용함으로써 출력된 디코딩된 제 2 표현에 의해 나타낸 것과 같이 스펙트럼 엔벨로프 형상화 또는 다른 운용을 실행한다. 라인(117) 상기 주파수 재생기(116)의 출력에서 나타내는 것과 같이 제 1 스펙트럼 부분들의 디코딩된 제 1 세트 및 스펙트럼 부분들의 재구성된 제 2 세트는 제 1 디코딩된 표현과 재구성된 제 2 스펙트럼 부분을 시간 표현(119)으로 변환하도록 구성되는 스펙트럼-시간 변환기(118) 내로 입력되며, 시간 표현은 특정한 높은 샘플링 레이트를 갖는다.

[0054]

도 2b는 도 1a의 인코더의 일 구현을 도시한다. 오디오 입력 신호(99)는 도 1a의 시간 스펙트럼 변환기(100)와 상응하는 분석 필터뱅크(220) 내로 입력된다. 그리고 나서, 시간적 잡음 형상화 블록(222) 내에서 시간적 잡음 형상화 운용이 실행된다. 따라서, 도 2b의 블록 음조 마스크(226)와 상응하는 도 1a의 스펙트럼 분석기(102) 내로의 입력은 시간적 잡음 형상화/시간적 타일 형상화 운용이 적용되지 않을 때 완전한 스펙트럼 값들일 수 있거나 또는 도 2b에 도시된 것과 같이 시간적 잡음 형상화 운용 블록(222)이 적용될 때, 스펙트럼 잔류 값들일 수 있다. 2-채널 신호들 또는 다중 채널 신호들을 위하여, 조인트 채널 코딩(228)이 부가적으로 실행될 수 있으며,

따라서 도 1a의 스펙트럼 도메인 인코더(106)는 조인트 채널 코딩 블록(228)을 포함할 수 있다. 게다가, 또한 도 1a의 스펙트럼 도메인 인코더(106)의 일부분인, 무손실 데이터 압축을 실행하기 위한 엔트로피 코더(232)가 제공된다.

[0055] 스펙트럼 분석기/음조 마스크(226)는 시간적 잡음 형상화 블록(222)의 출력을 제 1 스펙트럼 부분들(103)의 제 1 세트와 상응하는 코어 대역과 음조 성분들 및 도 1a의 제 2 스펙트럼 부분들(105)의 제 2 세트와 상응하는 잔류 성분들로 분리한다. 지능형 캡 필팅 파라미터 추출 인코딩으로서 표시되는 블록(224)은 도 1a의 파라미터 코어(104)와 상응하고 비트스트림 멀티플렉서(230)는 도 1a의 비트스트림 멀티플렉서(108)와 상응한다.

[0056] 바람직하게는, 분석 필터뱅크(222)는 변형 이산 코사인 변환 필터뱅크로서 구현되고 변형 이산 코사인 변환 필터뱅크는 주파수 분석 공구로서 작용하는 변형 이산 코사인 변환으로 신호(99)를 시간-주파수 도메인으로 변환시키도록 사용된다.

[0057] 스펙트럼 분석기(226)는 바람직하게는 음조성 마스크(tonality mask)를 적용한다. 음조성 마스크 추정 단계는 신호 내의 잡음 유사 성분들로부터 음조 성분들을 분리하도록 사용된다. 이는 코어 코더(228)가 모든 음조 성분을 음향심리학적 모듈로 인코딩하도록 허용한다. 음조성 마스크 추정 단계는 다양한 서로 다른 방법들로 구현될 수 있고 바람직하게는 음성/오디오 코딩을 위한 사인 및 잡음-모델링([8, 9]) 또는 [10]에서 설명되는 HILN(Harmonic and Individual Line plus Noise) 모델 기반 오디오 코더에서 사용된다. 바람직하게는, 출생-사망 궤도를 dnb지할 필요 없이 구현하기에 쉬운 구현이 사용되나, 다른 음조성 또는 잡음 검출기가 또한 사용될 수 있다.

[0058] 지능형 캡 필팅 모듈은 소스 영역과 표적 영역 사이에 존재하는 유사성을 계산한다. 표적 영역은 소스 영역으로부터의 스펙트럼에 의해 표현될 것이다. 소스 및 표적 영역 사이의 유사성의 측정은 상호 상관관계 접근법을 사용하여 수행된다. 표적 영역은 $nTar$ 비-오버래핑 주파수 타일들로 분할된다. 표적 영역 내의 모든 타일을 위하여, 고정된 시작 주파수(start frequency)로부터 $nSrc$ 소스 타일들이 생성된다. 이러한 소스 타일들은 0과 1 사이의 인자에 의해 오버래핑되는데, 0은 0% 오버랩을 의미하고 1은 100% 오버랩을 의미한다. 이러한 소스 타일들 각각은 표적 타일과 최적으로 매칭하는 소스 타일을 찾기 위하여 다양한 랜덤에서 표적 타일과 상호 연관된다.

[0059] 최적으로 매칭하는 타일 수는 $tileNum[idx_tar]$ 내에 저장되고, 표적과 가장 상호 관련되는 랜덤은 $xcorr_lag[idx_tar][idx_src]$ 내에 저장되며 상관관계의 부호는 $xcorr_sign[idx_tar][idx_src]$ 내에 저장된다. 상관관계가 고도로 음성일 때, 소스 타일은 디코더에서의 타일 필팅 과정 전에 -1로 곱해질 필요가 있다. 지능형 캡 필팅 모듈은 또한 스펙트럼 내의 음조 성분들의 비-오버라이팅(overwriting)을 처리하는데, 그 이유는 음조 성분들이 음조성 마스크를 사용하여 보존되기 때문이다. 스펙트럼을 정확하게 재구성하는 것을 가능하게 하는 표적 영역을 에너지를 저장하기 위하여 대역 방법 에너지 파라미터가 사용된다.

[0060] 이러한 방법은 다중-톤(tone) 신호의 고조파 그리드(grid)가 코어 코더에 의해 보존되며 정현파들 사이의 갭들이 소스 영역으로부터 최적으로 매칭하는 "형상화된 잡음"에 의해 보존된다는 점에서 고전적인 스펙트럼 대역 복제([1])에 대하여 특정 장점을 갖는다. ASR (정밀 스펙트럼 교체, Accurate Spectral Replacement)[2-4]와 비교하여 이러한 시스템의 또 다른 장점은 디코더에서 신호의 중요한 부분을 생성하는 신호 합성 단계의 부재이다. 대신에, 이러한 작업은 스펙트럼의 중요한 성분들의 보존을 가능하게 하는, 코어 코더에 의해 발생된다. 제안된 시스템의 또 다른 장점은 특징들이 제공하는 지속적인 확장성(scalability)이다. $tileNum[idx_tar]$ 및 $xcorr_lag=0$ 만의 사용은 모든 타일을 위하여 총 입상도 매칭으로 불리고 변수 $xcorr_lag$ 를 사용하는 동안에 낮은 비트레이트를 위하여 사용될 수 있고 모든 타일을 위하여 표적 및 소스 스펙트럼을 더 잘 매칭하는 것을 가능하게 한다.

[0061] 게다가, 트릴링(trilling)과 음악 잡음과 같은 주파수 도메인 아티팩트들을 제거하는 타일 선택 안정화 기술이 제안된다.

[0062] 스테레오 채널 쌍들의 경우에 부가적인 조인트 스테레오 과정이 적용된다. 이는 특정 도착지 범위를 위하여 신호가 고도로 상관된 페닝된(panned) 음원을 할 수 있기 때문에 필요하다. 이러한 특정 영역을 위하여 선택된 소스 영역들이 상관관계를 잘 갖지 않는 경우에 있어서, 공간적 이미지는 비-상관된 소스 영역들 때문에 악화될 수 있다. 인코더는 일반적으로 스펙트럼 갭들의 상호 상관관계를 실행하는, 각각의 목적지 영역 에너지 대역을 분석하고 만일 특정 한계값이 초과되면, 이러한 에너지 대역을 위한 조인트 플래그를 설정한다. 디코더에서 왼쪽 및 오른쪽 채널 에너지 대역은 만일 이러한 조인트 스테레오 플래그가 설정되지 않으면, 개별적으로 처리된다. 조인트 스테레오 플래그가 설정된 경우에 있어서, 에너지들과 페칭 모두 조인트 스테레오 도메인 내에서 실

행된다. 지능형 갭 필링 영역들을 위한 조인트 스테레오 정보는 예측이 경우에 예측의 방향이 다운믹스로부터 잔류로 되거나 이와 반대인지를 나타내는 플래그를 포함하는, 코어 코딩을 위한 조인트 스테레오 정보와 유사하게 시그널링된다.

[0063] 에너지들은 왼쪽/오른쪽 도메인 내의 전송된 에너지들로부터 계산될 수 있다.

[0064] $midNrg[k] = leftNrg[k] + rightNrg[k];$

[0065] $sideNrg[k] = leftNrg[k] - rightNrg[k];$

[0066] 여기서 k 는 변환 도메인 내의 주파수 지수이다.

[0067] 도 다른 해결책은 조인트 스테레오가 활성인 대역들을 위하여 조인트 스테레오 도메인 내에서 직접적으로 에너지들을 계산하고 전송하는 것이며, 따라서 디코더 면에서 어떠한 부가적인 변환도 필요하지 않다.

[0068] 소스 타일들은 항상 중면(mid/side) 매트릭스에 따라 생성된다:

[0069] $midTile[k] = 0.5 \cdot (leftTile[k] + rightTile[k])$

[0070] $sideTile[k] = 0.5 \cdot (leftTile[k] - rightTile[k])$

[0071] 에너지 조정:

[0072] $midTile[k] = midTile[k] * midNrg[k];$

[0073] $sideTile[k] = sideTile[k] * sideNrg[k];$

[0074] 조인트 스테레오 \rightarrow LR 변환

[0075] 만일 어떠한 부가적인 예측 파라미터도 코딩되지 않으면:

[0076] $leftTile[k] = midTile[k] + sideTile[k]$

[0077] $rightTile[k] = midTile[k] - sideTile[k]$

[0078] 만일 부가적인 예측 파라미터가 코딩되고 시그널링된 방향이 중간으로부터 측면이면:

[0079] $sideTile[k] = sideTile[k] - predictionCoeff \cdot midTile[k]$

[0080] $leftTile[k] = midTile[k] + sideTile[k]$

[0081] $leftTile[k] = midTile[k] - sideTile[k]$

[0082] 만일 시그널링된 방향이 측면으로부터 중간이면:

[0083] $midTile[k] = midTile[k] - predictionCoeff \cdot sideTile[k]$

[0084] $leftTile[k] = midTile[k] - sideTile[k]$

[0085] $leftTile[k] = midTile[k] + sideTile[k]$

[0086] 이러한 과정은 고도로 상관관계인 목적지 영역들과 퍼닝된 목적지 영역들을 발생시키기 위하여 사용되는 타일들로부터, 결과로서 생기는 왼쪽과 오른쪽 채널들은 만일 소스 영역들이 상관되지 않더라도, 여전히 상관되고 퍼닝된 음원을 표현하는 것을 보장하며, 그러한 영역들을 위한 스테레오 이미지를 보존한다.

[0087] 바꾸어 말하면, 비트스트림 내에서, 일반적인 조인트 스테레오 코딩을 위한 예로서 왼쪽/오른쪽 또는 중/면이 사용되어야만 하는지를 나타내는 조인트 스테레오 플래그들이 전송된다. 디코더에서, 첫 번째로, 코어 신호가 코어 대역들을 위한 조인트 스테레오 플래그들에 의해 나타낸 것과 같이 디코딩된다. 두 번째로, 코어 신호는 오른쪽/왼쪽 및 중간/측면 모두에 저장된다. 지능형 갭 필링 타일 필링을 위하여, 지능형 갭 필링 대역들을 위한 조인트 스테레오 정보에 의해 나타낸 것과 같이 표적 타일 표현에 맞추기 위하여 소스 타일 표현이 선택된다.

[0088] 시간적 잡음 형상화는 표준 기술이고 고급 오디오 코딩의 일부분이다([11-13]). 시간적 잡음 형상화는 필터뱅크와 양자화 단계 사이의 선택적 초리 단계를 삽입하는, 지각적 코더의 기본 전략의 확장으로서 고려될 수 있다. 시간적 잡음 형상화 모듈의 주요 작업은 트랜지언트 유사 신호들의 시간적 마스킹 영역 내에 생산된 양자화 잡음을 감추는 것이며 따라서 이는 더 효율적인 코딩 전략에 이르게 한다. 우선, 시간적 잡음 형상화는 변환 도메

인내의 "순방향 예측", 예들 들면 변형 이산 코사인 변환을 사용하여 예측 계수들의 세트를 계산한다. 이러한 계수들은 그리고 나서 신호의 시간적 엔벨로프를 평탄화(flattening)하기 위하여 사용된다. 양자화가 시간적 잡음 형상화 필터링된 스펙트럼에 영향을 미치기 때문에, 양자화 잡음이 시간적으로 편평하다. 디코더 면 상에 역 시간적 잡음 형상화를 적용함으로써, 양자화 잡음이 시간적 잡음 형상화 필터의 시간적 엔벨로프에 따라 형상화되고 따라서 양자화 잡음이 트랜지언트에 의해 마스킹된다.

[0089] 지능형 캡 필팅은 변형 이산 코사인 변환 표현을 기초로 한다. 효율적인 코딩을 위하여, 바람직하게는 약 20 ms의 긴 블록들이 사용되어야만 한다. 만일 그러한 긴 블록 내의 신호가 트랜지언트들을 포함하면, 타일 필팅에 기인하여 지능형 캡 필팅 스펙트럼 대역들 내에 가청(audible) 전- 및 후-에코들이 발생한다. 도 7c는 지능형 캡 필팅에 기인하는 트랜지언트 개시 전에 일반적인 전-에코 효과를 도시한다. 원쪽 면 상에, 원래 신호의 스펙트로그램이 도시되고 오른쪽 면 상에 시간적 잡음 형상화 필터링 없는 대역폭 확장된 신호의 스펙트로그램이 도시된다.

[0090] 전-에코 효과는 지능형 캡 필팅 콘텍스트 내의 시간적 잡음 형상화를 사용하여 감소된다. 여기서, 시간적 잡음 형상화는 시간적 타일 형상화로서 사용되는데 그 이유는 디코더 내의 스펙트럼 재생이 시간적 잡음 형상화 잔류 신호 상에서 실행되기 때문이다. 일반적으로 인코더 면 상에서 완전한 스펙트럼을 사용하여 필요한 시간적 타일 형상화 예측 계수들이 계산되고 적용된다. 시간적 잡음 형상화/시간적 타일 형상화 시작 및 중단 주파수(stop frequency)들은 지능형 캡 필팅 공구의 지능형 캡 필팅 시작 주파수($f_{IGFstart}$)에 의해 영향을 받지 않는다. 레거시 시간적 잡음 형상화와 비교하여, 시간적 타일 형상화 중단 주파수는 지능형 캡 필팅 공구의 중단 주파수로 증가되며, 이는 지능형 캡 필팅 시작 주파수보다 높다. 디코더 면 상에서 시간적 잡음 형상화/시간적 타일 형상화 계수들이 완전한 스펙트럼 상에 다시 적용되는데, 즉 코어 스펙트럼 플러스 재생된 스펙트럼 플러스 음조성 맵으로부터의 음조 성분들이다(도 7e 참조). 시간적 타일 형상화의 적용은 원래 신호의 엔벨로프를 다시 매칭하기 위하여 재생된 스펙트럼의 시간적 엔벨로프를 형성하는데 필요하다. 따라서 도시된 전-에코들이 감소된다. 게다가, 시간적 잡음 형상화에서와 같이 지능형 캡 필팅 시작 주파수 아래의 신호 내에 여전히 양자화 잡음을 형상화한다.

[0091] 레거시 디코더들에서, 오디오 신호에 대한 스펙트럼 패칭은 패치 경계들에서 스펙트럼 상관관계에 오류를 발생시키고 이에 의해 분산(dispersion)을 도입함으로써 오디오 신호의 시간적 엔벨로프를 손상시킨다. 따라서, 잔류 신호 상의 지능형 캡 필팅 타일 필팅 실행의 또 다른 혜택은 형상화 필터의 적용 후에 타일 경계들이 균일하게 상관되고, 신호의 더 신뢰할 수 있는 시간적 재생을 야기한다는 것이다.

[0092] 본 발명의 인코더에서, 명기된 시간적 잡음 형상화/시간적 타일 형상화를 갖는 스펙트럼, 음조성 마스크 처리 및 지능형 캡 필팅 파라미터 추정은 음조 성분들을 제외하고 지능형 캡 필팅 시작 주파수 위의 어떠한 신호도 없다. 이러한 회소 스펙트럼은 이제 산술 코딩과 예측 코딩의 원리들을 사용하여 코어 코더에 의해 코딩된다. 시그널링 비트들과 함께 이러한 코딩된 성분들은 오디오의 비트스트림을 형성한다.

[0093] 도 2a는 상용하는 디코더 구현을 도시한다. 인코딩된 오디오 신호와 상용하는 도 2a의 비트스트림은 도 1b와 관련하여 블록들(112 및 114)에 연결될 수 있는 디멀티플렉서/디코더 내로 입력된다. 비트스트림 디멀티플렉서는 입력 오디오 신호를 도 1b의 제 1 인코딩된 표현(107) 및 도 1b의 제 2 인코딩된 표현(109)의 입력 오디오 신호로 분리한다. 제 1 스펙트럼 부분들의 제 1 세트를 갖는 제 1 인코딩된 표현은 도 1b의 스펙트럼 도메인 디코더와 상용하는 조인트 채널 디코딩 블록(204) 내로 입력된다. 제 2 인코딩된 표현은 도 2a에 도시되지 않은 파라미터 디코더(114) 내로 입력되고 그리고 나서 제 1b의 주파수 재생기(116)와 상용하는 지능형 캡 필팅 블록(202) 내로 입력된다. 주파수 재생을 위하여 필요한 제 1 스펙트럼 부분들의 제 1 세트는 라인(203)을 거쳐 지능형 캡 필팅 블록(202) 내로 입력된다. 게다가, 음조 마스크의 출력이 스펙트럼 도메인 디코더(112)의 출력과 상용하도록 조인트 채널 디코딩(204) 뒤에 특정 코어 디코딩이 음조 마스크 블록(206) 내에 적용된다. 그리고 나서, 결합기(208)에 의한 결합, 즉 결합기(208)의 출력이 이제 완전 범위 스펙트럼을 가지나, 여전히 시간적 잡음 형상화/시간적 타일 형상화 필터링된 도메인 내에 존재하는 프레임 빌딩이 실행된다. 그리고 나서, 블록(210)에서, 라인(109)을 통하여 제공되는 시간적 잡음 형상화/시간적 타일 형상화 정보를 사용하여 역 시간적 잡음 형상화/시간적 타일 형상화 운용이 실행되는데, 즉 시간적 타일 형상화 부가 정보는 바람직하게는 예를 들면 간단한 고급 오디오 코딩 또는 통합 음성 오디오 코딩일 수 있는, 스펙트럼 도메인 인코더(106)에 의해 발생되는 제 1 인코딩된 표현 내에 포함되거나, 또는 제 2 인코딩된 표현 내에 포함될 수 있다. 블록(210)의 출력에서, 최대 주파수까지 원래 입력 신호의 샘플링 레이트에 의해 정의되는 완전 범위 주파수인, 완전한 스펙트럼이 제공된다. 그리고 나서, 최종적으로 오디오 출력 신호를 획득하기 위하여 합성 필터뱅크(212) 내에서 스펙트럼/

시간 변환이 실행된다.

[0094] 도 3a는 스펙트럼의 개략적인 표현을 도시한다. 스펙트럼은 도 3a에 도시된 예에서 7개의 스케일 인자 대역(SCB1 내지 SCB7)이 존재하는 스케일 인자 대역(SCB)들로 세분된다. 스케일 인자 대역들은 고급 오디오 코딩 표준에서 정의되고 도 3에 개략적으로 도시된 것과 같이 상부 주파수들까지 증가하는 대역폭을 갖는 고급 오디오 코딩 스케일 인자일 수 있다. 스펙트럼의 바로 시작부터, 즉 낮은 주파수들에서 지능형 캡 필링을 실행하는 것이 아니고, 309에 도시된 지능형 캡 필링 시작 주파수에서 지능형 캡 필링 운용을 시작하는 것이 바람직하다. 따라서, 코어 주파수 대역은 스펙트럼 최저 주파수로부터 지능형 캡 필링 주파수로 확장한다. 지능형 캡 필링 시작 주파수 위에서, 제 2 스펙트럼 부분들의 제 2 세트에 의해 표현되는 낮은 해상도 성분들로부터 높은 해상도 스펙트럼 성분들(304, 305, 306, 307, 제 1 스펙트럼 부분들의 제 1 세트)을 분리하기 위하여 스펙트럼 분석이 적용된다. 도 3a는 바람직하게는 스펙트럼 도메인 인코더(106) 또는 조인트 채널 코더(228) 내로 입력되는 스펙트럼을 도시하는데, 즉 코어 인코더는 완전한 범위 내에서 운용하나, 상당한 양의 제로 스펙트럼 값들을 인코딩하는데, 즉 이러한 제로 스펙트럼 값들은 0으로 양자화되거나 또는 양자화 전에 양자화 뒤에 0으로 설정된다. 어쨌든, 코어 인코더는 완전한 범위 내에서 운용하는데, 즉 스펙트럼이 도시된 것과 같을 수 있으며, 즉 코어 디코더는 어떠한 지능형 캡 필링 또는 낮은 스펙트럼 해상도를 갖는 제 2 스펙트럼 부분들의 제 2 세트의 인코딩을 반드시 지각할 필요는 없다.

[0095] 바람직하게는, 높은 해상도는 변형 이산 코사인 변환 라인들과 같은 스펙트럼 라인들의 라인 방식 코딩에 의해 정의되며, 제 2 해상도 또는 낮은 해상도는 예를 들면, 스케일 인자 대역 당 하나의 단일 스펙트럼 값을 계산함으로써 정의되는데, 스케일 인자는 일부 주파수 라인들을 포함한다. 따라서, 제 2 낮은 해상도는 그것의 스펙트럼 해상도와 관련하여, 일반적으로 고급 오디오 코딩 또는 통합 음성 오디오 코딩 코어 인코더와 같은 코어 인코더에 의해 적용되는 라인 방식 코딩에 의해 정의되는 제 1 또는 높은 해상도보다 훨씬 낮다.

[0096] 스케일 인자 또는 에너지 계산과 관련하여, 상황이 도 3b에 도시된다. 인코더가 코어 인코더라는 사실 때문에, 그리고 각각의 대역 내의 스펙트럼 부분들의 제 1 세트의 성분들일 수는 있으나 반드시 그럴 필요는 없다는 사실 때문에, 코어 인코더는 지능형 캡 필링 시작 주파수(309) 아래뿐만 아니라, 샘플링 주파수의 반, 즉 $f_{s/2}$ 와 유사하거나 또는 동일한 최대 주파수($f_{IGF_{stop}}$)까지 지능형 캡 필링 시작 주파수 위의 코어 범위 내의 각각의 대역을 위한 스케일 인자를 계산한다. 따라서, 도 3a의 인코딩된 음조 부분들(302, 304, 305, 306, 307)은 본 실시 예에서는 스케일 인자들(SCB1 내지 SCB7)과 함께 높은 해상도 스펙트럼 데이터와 상응한다. 낮은 해상도 스펙트럼 데이터는 지능형 캡 필링 시작 주파수로부터 시작하여 계산되며 스케일 인자들(SCB1 내지 SCB7)과 함께 전송되는, 에너지 정보 값들(E_1, E_2, E_3, E_4)과 상응한다.

[0097] 특히, 코어 인코더가 낮은 비트레이트 상태 하일 때, 코어 대역, 즉 지능형 캡 필링 시작 주파수보다 주파수가 낮은, 즉 스케일 인자 대역들(SCB1 내지 SCB7) 내에서 부가적인 잡음-필링 운용이 적용될 수 있다. 디코더 면상에서, 0으로 양자화된 이러한 값들은 재합성되고 재합성된 스펙트럼 값들은 도 3b의 308에 도시된 NF_2 와 같은 잡음-필링 에너지를 사용하여 그것들의 진폭 내에서 조정된다. 특히 통합 음성 오디오 코딩에서와 같이 스케일 인자와 관련하여 절대 항들 또는 상대 항들 내에서 주어질 수 있는 잡음-필링 에너지는 0으로 양자화된 스펙트럼 값들의 세트의 에너지와 상응한다. 이러한 잡음-필링 스펙트럼 라인들은 또한 소스 범위와 에너지 정보(E_1, E_2, E_3, E_4)로부터 스펙트럼 값들을 사용하여 주파수 타일들을 재생하기 위하여 다른 주파수들로부터 주파수 타일들을 사용하는 주파수 재생에 의존하는 어떠한 지능형 캡 필링 운용 없이 간단한 잡음-필링 합성에 의해 재생되는 제 3 스펙트럼 부분들의 제 3 세트로서 고려될 수 있다.

[0098] 바람직하게는, 에너지 정보가 계산되는 대역들은 스케일 인자 대역들과 일치한다. 다른 실시 예들에서, 에너지 정보 값 그룹화(grouping)가 적용되며 따라서 예를 들면 스케일 인자 대역들(4 및 5)을 위하여, 단일 에너지 정보 값만이 전송되나, 본 실시 예에서도 그룹화된 재구성 대역들의 경계들은 스케일 인자 대역들의 경계들과 일치한다. 만일 서로 다른 대역 분리가 적용되면, 특정 재-계산들 또는 동기화 계산들이 적용될 수 있으며, 이는 특정 구현에 따라 이해할 수 있다.

[0099] 바람직하게는, 도 1a의 스펙트럼 도메인 인코더(106)는 도 4a에 도시된 것과 같이 음향심리학적으로 구동되는 인코더이다. 일반적으로, 예를 들면 MPEG2/4 고급 오디오 코딩 표준 또는 MPEG1/2 계층 3 표준에서 나타내는 것과 같이, 스펙트럼 범위(도 4a의 401)로 변환된 후에 인코딩되려는 오디오 신호는 스케일 인자 계산기(400)로 전달된다. 스케일 인자 계산기는 부가적으로 인코딩되려는 오디오 신호를 수신하거나 또는 MPEG1/2 계층 3 또는 MPEG 고급 오디오 코딩 표준에서와 같이 오디오 신호의 복소수 스펙트럼 신호를 수신하는 음향심리학적 모델에

의해 제어된다. 음향심리학적 모델은 각각의 스케일 인자 대역을 위하여, 음향 심리학적 한계를 표현하는 스케일 인자를 계산한다. 부가적으로, 시간 인자들은 그리고 나서, 잘 알려진 내부 또는 외부 반복 루프들에 이하거나 또는 어떠한 다른 적절한 인코딩 과정에 의해 특정 비트레이트 상태들이 충족되도록 조정된다. 그리고 나서 한편으로는 양자화되려는 스펙트럼 값들 및 다른 한편으로는 계산된 스케일 인자들이 양자화기 프로세서(404) 내로 입력된다. 간단한 오디오 인코더 운용에서, 양자화되려는 스펙트럼 값들은 스케일 인자들에 의해 가중되고, 가중된 스케일 스펙트럼 값들은 그리고 나서 일반적으로 상부 진폭 범위들에 대하여 압축 기능을 갖는 고정된 양자화기 내로 입력된다. 그리고 나서, 양자화기 프로세서의 출력에서 일반적으로 인접한 주파수 값들을 위한 제로-양자화 지수들의 세트, 또는 종래에 제로 값들의 런(run)"으로 불린 것과 같은, 특이적이고 매우 효율적인 코딩을 갖는 엔트로피 인코더 내로 전달되는 양자화 지수들이 존재한다.

[0100] 그러나, 도 1a의 오디오 인코더에서, 양자화기 프로세서는 일반적으로 스펙트럼 분석기로부터 제 2 스펙트럼 부분들에 대한 정보를 수신한다. 따라서, 양자화기 프로세서(404)는 양자화기 프로세서(404)의 출력에 있어서, 스펙트럼 분석기(102)에 의해 정의되는 것과 같이 제 2 스펙트럼 부분들이 0이거나 또는 특히 스펙트럼 내에 제로 값들의 "런"이 존재할 때, 매우 효율적으로 코딩될 수 있는 제로 표현으로서 인코더 또는 디코더에 의해 인정되는 표현을 갖도록 한다.

[0101] 도 4b는 양자화기 프로세서의 일 구현을 도시한다. 변형 이산 코사인 변환 스펙트럼 값들은 제로 블록(410)으로의 설정 세트 내로 입력될 수 있다. 그리고 나서, 제 2 스펙트럼 부분들은 블록(412) 내에서 스펙트럼 인자들에 의한 가중 이전에 이미 0으로 설정된다. 부가적인 구현에서, 블록(410)은 제공되지 않으나, 가중 블록(412) 뒤의 블록(418)에서 0으로의 설정 협력이 실행된다. 또 다른 구현에서, 양자화 블록(420) 내에서의 양자화 뒤에 0으로의 설정 블록(422) 내에서 0으로의 설정 운용이 또한 실행될 수 있다. 이러한 구현에서, 블록들(410 및 413)은 존재하지 않을 수 있다. 일반적으로, 특정 구현에 따라 블록들(410, 418, 422) 중 적어도 하나가 제공된다.

[0102] 그리고 나서, 블록(422)의 출력에서, 도 3a에 도시된 것과 상응하는 양자화된 스펙트럼이 획득된다. 양자화된 스펙트럼은 그리고 나서 허프만 코더(Huffman coder) 또는 예를 들면 통합 음성 오디오 코딩 표준에서 정의되는 것과 같은 산술 코더일 수 있는, 도 2b의 232와 같은 엔트로피 코더 내로 입력된다.

[0103] 대안으로서 서로 또는 병렬로 제공되는 0으로의 설정 블록들(410, 418, 422)은 스펙트럼 분석기(424)에 의해 제어된다. 스펙트럼 분석기는 바람직하게는 잘 알려진 음조성 검출기의 어떠한 구현을 포함하거나 또는 스펙트럼을 높은 해상도로 인코딩되려는 성분들과 낮은 해상도로 인코딩되려는 성분들로 분리하도록 운용되는 어떠한 다른 종류의 검출기를 포함한다. 스펙트럼 분석기에서 구현되는 그러한 다른 알고리즘들은 보이스 활성 검출기(voice activity detector), 잡음 검출기, 음성 검출기 또는 서로 다른 스펙트럼 부분들을 위한 해상도 요구사항에 대한 스펙트럼 정보 또는 관련 메타데이터에 따라 결정하는 어떠한 다른 검출기일 수 있다.

[0104] 도 5a는 예를 들면 고급 오디오 코딩 또는 통합 음성 오디오 코딩에서 구현되는 것과 같은, 도 1a의 시간 스펙트럼 변환기(100)의 바람직한 구현을 도시한다. 시간 스펙트럼 변환기(100)는 트랜지언트 검출기(504)에 의해 제어되는 윈도우어(windower, 502)를 포함한다. 트랜지언트 검출기(504)가 트랜지언트를 검출할 때, 긴 윈도우들로부터 짧은 윈도우들로의 전환(switchover)이 윈도우어에 시그널링된다. 윈도우어(502)는 그리고 나서 블록들의 오버래핑을 위하여, 윈도우잉된(windowed) 프레임들을 계산하는데, 각각의 윈도우잉된 프레임은 일반적으로 2048 값들과 같은 두 개의 N 값을 갖는다. 그리고 나서, 블록 변환기(506) 내의 변환이 실행되고, 이러한 블록 변환기는 일반적으로 부가적으로 대시메이션(decimation)을 제공하며, 따라서 변형 이산 코사인 변환 스펙트럼 값들과 같은 N 값들을 갖는 스펙트럼 프레임을 획득하기 위하여 결합된 대시메이션/변환이 실행된다. 따라서, 긴 윈도우 운용을 위하여, 블록(506)의 입력에서의 프레임은 2048 값들과 같은 두 개의 N 값을 포함하고 스펙트럼 프레임은 그때 1024 값들을 갖는다. 그러나, 그리고 나서 8개의 짧은 블록이 실행될 때 짧은 블록들로의 전환이 실행되며 각각의 짧은 블록은 긴 윈도우와 비교하여 1/8 윈도우잉된 시간 도메인을 갖고 각각의 스펙트럼 블록은 긴 블록과 비교하여 1/8 스펙트럼 값들을 갖는다. 따라서, 대시메이션이 윈도우어의 50% 오버랩 운용과 결합될 때, 스펙트럼은 시간 도메인 오디오 신호(99)의 중요하게 샘플링된 버전이다.

[0105] 그 뒤에, 더 1b의 주파수 재생기(118) 및 스펙트럼-시간 변환기(118), 또는 도 2a의 블록들(208, 212)의 운용의 특정 구현을 도시한 도 5b가 참조된다. 도 5b에서, 도 3a의 스케일 인자 대역(6)과 같은 특정 재구성 대역이 고려된다. 이러한 재구성 대역 내의 제 1 스펙트럼 부분, 즉 도 3a의 제 1 스펙트럼 부분(306)은 프레임 빌더(frame builder)/조정기 블록(510) 내로 입력된다. 게다가, 스케일 인자 대역(6)을 위한 재구성된 제 2 스펙트럼 부분이 또한 프레임 빌더/조정기(510) 내로 입력된다. 게다가, 스케일 인자 대역(6)을 위한 도 3b의 E₃과 같

은 에너지 정보가 또한 블록(510) 내로 입력된다. 재구성 대역 내의 재구성된 제 2 스펙트럼 부분은 이미 소스 영역을 사용하여 주파수 타일 필링에 의해 발생되었고 재구성 대역은 그리고 나서 표적 범위와 상응한다. 이제, 그리고 나서 예를 들면 도 2a의 결합기(208)의 출력에서 획득된 것과 같은 N 값을 갖는 최종적으로 완전한 재구성된 프레임을 획득하기 위하여 프레임의 에너지 조정이 실행된다. 그리고 나서 블록(512)에서, 예를 들면 블록(512)의 입력에서 124 스펙트럼 값들을 위한 248 시간 도메인 값들을 획득하기 위하여 역 블록 변환/보간(interpolation)이 실행된다. 그리고 나서, 인코딩된 오디오 신호 내의 부가 정보로서 전송된 긴 원도우/짧은 원도우 표시에 의해 다시 제어되는 합성 원도우인 운용이 실행된다. 그리고 나서 블록(516)에서, 이전 시간 프레임(previous time frame)을 갖는 오버랩/애드 운용이 실행된다. 바람직하게는, 변형 이산 코사인 변환은 50% 오버랩을 적용하며, 따라서 2N 값들의 각각의 새로운 시간 프레임을 위하여, N 시간 도메인 값들이 최종적으로 출력된다. 50% 오버랩은 블록(516) 내의 오버랩/애드 운용에 기인하여 하나의 프레임으로부터 그 다음 프레임으로 중요한 샘플링과 지속적인 크로스오버를 제공한다는 사실 때문에 매우 바람직하다.

[0106] 도 3a의 301에 도시된 것과 같이, 잡음-필링 운용은 부가적으로 도 3a의 스케일 인자 대역과 일치하는 고려된 재구성 대역을 위한 것과 같이 지능형 캡 필링 시작 주파수 아래에서뿐만 아니라 지능형 캡 필링 시작 주파수 위에서도 적용될 수 있다. 그리고 나서, 잡음-필링 스펙트럼 값들은 또한 프레임 빌더/조정기(510) 내로 입력될 수 있고 잡음-필링 스펙트럼 값들의 조정이 또한 이러한 블록 내에 적용될 수 있거나 또는 잡음-필링 스펙트럼 값들은 프레임 빌더/조정기(510) 내로 입력되기 전에 잡음-필링 에너지를 사용하여 이미 조정될 수 있다.

[0107] 바람직하게는, 지능형 캡 필링 운용, 즉 다른 부분들로부터 스펙트럼 값들을 사용하는 주파수 타일 필링 운용은 완전한 스펙트럼 내에서 적용될 수 있다. 따라서, 스펙트럼 타일 필링 운용은 지능형 캡 필링 시작 주파수 위의 고대역에서 뿐만 아니라 저대역에서도 적용될 수 있다. 게다가, 주파수 타일 필링이 없는 잡음-필링은 또한 지능형 캡 필링 시작 주파수 아래에서뿐만 아니라 지능형 캡 필링 시작 주파수 위에서도 적용될 수 있다. 그러나, 높은 품질 및 고효율 오디오 인코딩은 잡음-필링 운용이 지능형 캡 필링 시작 주파수 아래의 주파수 범위에 한정될 때, 그리고 주파수 타일 필링 운용이 도 3a에 도시된 것과 같이 지능형 캡 필링 시작 주파수 위의 주파수 범위에 제한될 때 획득될 수 있다는 것을 발견하였다.

[0108] 바람직하게는, 표적 터일(TT, 지능형 캡 필링 시작 주파수보다 큰 주파수들을 갖는)은 완전 비율 코더의 스케일 인자 대역 경계들로 향한다. 정보가 얻어지는 소스 타일들(ST)은, 즉 지능형 캡 필링 시작 주파수보다 낮은 주파수들을 위하여 스케일 인자 대역 경계들에 엮매이지 않는다. 소스 타일의 크기는 관련된 표적 터일의 크기와 상응하여야만 한다. 이는 다음의 예를 사용하여 설명된다. TT[0]은 10 변형 이산 코사인 변환 빈들의 크기를 갖는다. 이는 정확하게 두 개의 뒤따르는 스케일 인자 대역의 길이(4+6과 같이)와 상응한다. 그리고 나서 TT[0]과 상관되는 모든 가능한 소스 타일이 10 빈의 길이를 갖는다. TT[0]에 인접한 제 2 표적 터일(TT[1])은 15 빈의 길이를 갖는다(7+8의 길이를 갖는 스케일 인자 대역). 그리고 나서, 이를 위한 소스 타일은 TT[0]에 대하여 10 빈보다는 15 빈의 길이를 갖는다.

[0109] 표적 터일의 길이를 갖는 소스 타일을 위한 표적 터일을 발견할 수 없는 경우가 발생하면(예를 들면 표적 터일의 길이가 이용가능한 소스 범위보다 클 때), 상관관계는 계산되지 않으며 소스 범위는 표적 터일(TT)이 완전히 채워질 때까지 이러한 표적 터일 내로 여러 번 복사된다(복사는 두 번째 복사의 최저 주파수를 위한 주파수 라인이 첫 번째 복사의 최고 주파수를 위한 주파수 라인을 바로 뒤따르도록 번갈아 수행된다).

[0110] 그 뒤에, 도 1b의 주파수 재생기(116) 또는 도 2a의 지능형 캡 필링 블록(202)의 또 다른 바람직한 실시 예를 도시한 도 5c가 참조된다. 블록(522)은 표적 대역 식별뿐만 아니라 부가적으로 소스 대역 식별을 수신하는 주파수 타일 발생기이다. 바람직하게는, 인코더 면 상에서 도 3a의 스케일 인자 대역(3)이 스케일 인자 대역(7)을 재구성하기에 매우 적합하다는 것을 결정하였다. 따라서, 소스 대역 식별은 2일 수 있으며 표적 대역 식별은 7일 수 있다. 이러한 정보를 기초로 하여, 주파수 타일 발생기(522)는 스펙트럼 성분들(523)의 원시 제 2 부분들을 발생시키기 위하여 고조파 타일 필링 운용 또는 어떠한 다른 필링 운용까지 복사를 적용한다. 스펙트럼 성분들의 원시 제 2 부분들은 제 1 스펙트럼 부분들의 제 1 세트 내에 포함된 주파수 해상도와 동일한 주파수 해상도를 갖는다.

[0111] 그리고 나서, 도 3a의 307과 같은 재구성 대역의 제 1 스펙트럼 부분이 프레임 빌더(524) 내로 입력되고 원시 제 2 부분(523)이 또한 프레임 빌더(524) 내로 입력된다. 그리고 나서, 재구성된 프레임은 이득 인자 계산기(528)에 의해 계산되는 재구성 대역을 위한 이득 인자를 사용하여 조정기(526)에 의해 조정된다. 그러나 중요하게도, 프레임 내의 제 1 스펙트럼 부분은 조정기(526)에 의해 영향을 받지 않으나, 재구성 프레임을 위한 원시 제 2 부분만이 조정기(526)에 의해 영향을 받는다. 이를 위하여, 이득 인자 계산기(528)는 소스 대역 또는 원시

제 2 부분(523)을 분석하고 부가적으로 스케일 인자 대역(7)이 고려될 때, 조정기(526)에 의해 출력된 조정된 프레임의 에너지가 에너지(E_4)를 갖도록 최종적으로 정확한 이득 인자(527)를 찾기 위하여 재구성 대역 내의 제 1 스펙트럼 부분을 분석한다.

[0112] 이러한 맥락에서, 고효율 고급 오디오 코딩과 비교하여 본 발명의 고주파수 재구성 정확도를 평가하는 것은 매우 중요하다. 이는 도 3a의 스케일 인자 대역(7)과 관련하여 설명된다. 도 13a에 도시된 것과 같은 종래의 인코더는 "손실 고조파"로서 높은 해상도로 인코딩되려는 스펙트럼 부분(307)을 검출할 수 있는 것으로 추정된다. 그리고 나서, 이러한 스펙트럼 성분의 에너지는 스케일 인자 대역(7)과 같은 재구성 대역을 위한 스펙트럼 엔벨로프 정보와 함께 디코더로 전송될 수 있다. 그리고 나서, 디코더는 손실 고조파를 재생성할 수 있다. 그러나, 손실 고조파(307)가 도 13a의 종래의 디코더에 의해 재구성될 수 있는 스펙트럼 값은 재구성 주파수(390)에 의해 나타내는 주파수에서 대역(7)의 중간에 존재할 수 있다. 따라서, 본 발명은 도 13d의 종래의 디코더에 의해 도입될 수 있는 주파수 오류(391)를 방지한다.

[0113] 일 구현에서, 스펙트럼 분석기는 또한 제 1 스펙트럼 부분들과 제 2 스펙트럼 부분들 사이의 유사성을 계산하도록 구현되고 계산된 유사성을 기초로 하여, 재구성 범위 내의 제 2 스펙트럼 부분을 위하여 가능한 한 빨리 제 2 스펙트럼 부분과 매칭하는 제 1 스펙트럼 부분을 결정한다. 그리고 나서 이러한 가변 소스/목적지 범위 구현에서, 파라미터 코더가 부가적으로 제 2 인코딩된 표현 내로 도입될 것이며 매칭 정보는 각각의 목적지 범위를 위하여 매칭 소스 범위를 나타낸다. 디코더 면 상에서, 이러한 정보는 그리고 나서 소스 대역 식별과 표적 대역 식별을 기초로 하여 원시 제 2 부분(523)의 발생을 나타내는 도 5c의 주파수 타일 발생기(522)에 의해 사용될 수 있다.

[0114] 게다가, 도 3a에 도시된 것과 같이, 스펙트럼 분석기는 샘플링 주파수의 반 이하의 적은 양이고 바람직하게는 샘플링 주파수의 적어도 1/4 또는 일반적으로 높은 최대 분석 주파수까지 스펙트럼 표현을 분석하도록 구성된다.

[0115] 도시된 것과 같이, 인코더는 다운샘플링 없이 운용하고 디코더는 업샘플링 없이 운용한다. 바꾸어 말하면, 스펙트럼 도메인 오디오 코더는 원래 입력된 오디오 신호의 샘플링 레이트에 의해 정의되는 나이퀴스트 주파수 (Nyquist frequency)를 갖는 스펙트럼 표현을 발생시키도록 구성된다.

[0116] 게다가, 도 3a에 도시된 것과 같이, 스펙트럼 분석기는 캡 필링 시작 주파수로 시작하여 스펙트럼 표현 내에 포함된 최대 주파수에 이해 표현되는 최대 주파수로 끝나는 스펙트럼 표현을 분석하도록 구성되며, 최소 주파수로부터 캡 필링 시작 주파수까지 확장하는 스펙트럼 부분은 스펙트럼 부분들의 제 1 세트에 속하고 캡 필링 주파수 위의 주파수 값들을 갖는 304, 305, 306, 307과 같은 또 다른 스펙트럼 부분은 부가적으로 제 1 스펙트럼 부분들의 제 1 세트 내에 포함된다.

[0117] 설명된 것과 같이, 스펙트럼 도메인 오디오 디코더(112)는 제 1 디코딩된 표현 내의 스펙트럼 값에 의해 표현되는 최대 주파수가 샘플링 레이트를 갖는 시간 주파수 내에 포함되는 최대 주파수와 동일하도록 구성되며 제 1 스펙트럼 부분들의 제 1 세트 내의 최대 주파수를 위한 스펙트럼 값은 0이거나 또는 0과 다르다. 어쨌든, 스펙트럼 성분들의 제 1 세트 내의 최대 주파수를 위하여 스케일 인자 대역 내의 모든 스펙트럼 값이 도 3a와 3b의 맥락에서 논의된 것과 같이 0으로 설정되거나 또는 그렇지 않은지와 관계없이 발생되거나 또는 전송되는, 스케일 인자 대역을 위한 스케일 인자가 존재한다.

[0118] 따라서, 본 발명은 압축 효율을 증가시키기 위한 다른 파라미터 기술들, 예를 들면 잡음 대체 및 잡음 필링(이러한 기술들은 독점적으로 잡음 같은 국부적 신호 콘텐츠의 효율적인 표현을 위한 것이다)과 관련하여 본 발명은 음조 성분들의 정확한 주파수 재생을 허용한다는 점에서 바람직하다. 지금까지, 어떠한 최신 기술도 저대역과 고대역 내의 고정된 연역적 세분의 제한 없이 스펙트럼 캡 필링에 의한 임의적 신호 콘텐츠의 효율적인 파라미터 표현을 다루지 못한다.

[0119] 본 발명의 시스템의 실시 예들은 종래 접근법들을 향상시키고 이에 의해 낮은 비트레이트들에서도 높은 압축 효율, 없거나 매우 적은 지각적 어노이언스(annoyance), 및 완전한 오디오 대역폭을 제공한다.

[0120] *일반적인 시스템은 다음을 구성한다:

- 완전 대역 코어 코딩
- 지능형 캡 필링(타일 필링 또는 잡음 필링)

- 음조 마스크에 의해 선택되는 코어 내의 희소 음조 부분들
- 터일 필링을 포함하는, 완전 대역을 위한 조인트 스테레오 쌍 코딩
- 타일 상의 시간적 잡음 형상화
- 지능형 캡 필링 범위 내의 스펙트럼 백색화

[0127] 더 효율적인 시스템을 향한 제 1 단계는 스펙트럼 데이터를 코어 코더와 다른 제 2 변환 도메인으로 변환하기 위한 필요성을 제거하는 것이다. 예를 들면 고급오디오 코딩과 같은, 대부분의 오디오 코덱들이 기본 변환으로서 변형 이산 코사인 변환을 사용하기 때문에 변형 이산 코사인 변환 도메인 내의 대역폭 확장을 실행하는 것이 또한 유용하다. 대역폭 확장 시스템의 두 번째 요구사항은 고주파수 음조 성분들이 보존되고 코딩된 오디오의 품질이 따라서 현존하는 시스템들보다 뛰어난 음조 그리드를 보존하기 위한 필요성일 수 있다. 위에 언급된 대역폭 확장 전략을 위한 요구사항 모두를 처리하기 위하여, 지능형 캡 필링으로 불리는 새로운 시스템이 제안된다. 도 2b는 인코더 면 상의 제안된 시스템의 다이어그램을 도시하고 도 2a는 디코더 면 상의 시스템을 도시한다.

[0128] 도 9a는 제 1 스펙트럼 부분들의 제 1 세트의 인코딩된 표현 및 제 2 스펙트럼 부분들의 제 2 세트를 위한 스펙트럼 에너지들을 나타내는 파라미터 데이터의 인코딩된 표현을 포함하는 인코딩된 오디오 신호를 디코딩하기 위한 장치를 도시한다. 스펙트럼 부분들의 제 1 세트는 도 9a에서 901a에 표시되고, 파라미터 데이터의 인코딩된 표현은 도 9a에서 901b에 표시된다.

[0129] 오디오 디코더(900)는 디코딩된 제 1 스펙트럼 부분들(904)의 제 1 세트를 획득하도록 제 1 스펙트럼 부분들의 제 1 세트의 인코딩된 표현(901a)을 디코딩하기 위하여, 그리고 개별 재구성 대역들을 위한 개별 에너지들을 나타내는 제 2 스펙트럼 부분들의 제 2 세트를 위한 디코딩된 파라미터 데이터(902)를 획득하도록 파라미터 데이터의 인코딩된 표현을 디코딩하기 위하여 제공되며, 제 2 스펙트럼 부분들은 재구성 대역들 내에 위치된다. 게다가, 제 2 스펙트럼 부분을 포함하는 재구성 대역의 스펙트럼 값들을 재구성하기 위하여 주파수 재생기(906)가 제공된다. 주파수 재생기(906)는 제 1 스펙트럼 부분들의 제 1 세트의 하나의 제 1 스펙트럼 부분 및 재구성 대역을 위한 하나의 개별 에너지 정보를 사용하며, 재구성 대역은 제 1 스펙트럼 부분과 제 2 스펙트럼 부분을 포함한다.

[0130] 주파수 재생기(906)는 재구성 대역 내의 주파수들을 갖는 제 1 스펙트럼 부분의 축적된 에너지를 포함하는 생존 에너지 정보를 결정하기 위한 계산기(912)를 포함한다. 게다가, 주파수 재생기(906)는 재구성 대역의 또 다른 스펙트럼 부분들의 타일 에너지 정보를 결정하고 제 1 스펙트럼 부분과 다른 주파수 값들을 결정하기 위한 계산기(918)를 포함하며, 이러한 주파수 값들은 재구성 대역 내의 주파수들을 가지며, 또 다른 스펙트럼 부분들은 재구성 대역 내의 제 1 스펙트럼 부분과 다른 제 1 스펙트럼 부분을 사용하는 주파수 재생에 의해 발생된다.

[0131] 주파수 재생기(906)는 재구성 대역 내의 손실 에너지를 위한 계산기(914)를 더 포함하며, 계산기(914)는 재구성 대역을 위한 개별 에너지 및 블록(912)에 의해 발생되는 생존 에너지를 사용하여 운용한다. 게다가, 주파수 재생기(906)는 손실 에너지 정보 및 블록(918)에 의해 발생된 에너지 정보를 기초로 하여 재구성 대역 내의 또 다른 스펙트럼 부분들을 조정하기 위한 스펙트럼 엔벨로프 조정기(916)를 포함한다.

[0132] 특정 재구성 대역(920)을 도시한 도 9c가 참조된다. 재구성 대역은 921에 개략적으로 도시된 도 3a에서의 제 1 스펙트럼 부분(306)과 같은 재구성 대역 내의 제 1 스펙트럼 부분을 포함한다. 게다가, 재구성 대역 내의 스펙트럼 값들의 나머지는 예를 들면 도 3a의 지능형 캡 필링 시작 주파수(309) 아래의 스케일 인자 대역(1, 2, 3)으로부터, 소스 영역을 사용하여 발생된다. 주파수 재생기(906)는 제 2 스펙트럼 부분들(922 및 923)을 위한 원시 스펙트럼 값들을 발생시키도록 구성된다. 그리고 나서, 이제 동일한 스펙트럼 해상도, 즉 제 1 스펙트럼 부분(921)과 동일한 라인 거리를 갖는 재구성 대역(920) 내의 재구성되고 조정된 제 2 스펙트럼 부분들을 획득하기 위하여 최종적으로 주파수 대역들(922, 923) 내의 원시 스펙트럼 값들을 조정하도록 도 9c에 도시된 것과 같이 이득 인자(g)가 계산된다. 도 9c의 921에 도시된 재구성 대역 내의 제 1 스펙트럼 부분은 오디오 디코더(900)에 의해 디코딩되고 도 9b의 엔벨로프 조정 실행된 블록(916)에 의해 영향을 받지 않는다는 것을 이해하는 것이 중요하다. 대신에, 921에 표시된 재구성 대역 내의 제 1 스펙트럼 부분은 그대로 남는데, 그 이유는 이러한 스펙트럼 부분이 라인(904)을 통하여 완전 대역폭 또는 완전 비율 오디오 디코더(900)에 의해 출력되기 때문이다.

[0133] 그 뒤에, 실수들을 갖는 특정 예가 설명된다. 블록(912)에 의해 계산된 것과 같은 나머지 생존 에너지는 예를

들면, 5개의 에너지 유닛이며 이러한 에너지는 제 1 스펙트럼 부분(921) 내에 바람직하게 표시된 4개의 스펙트럼 라인의 에너지이다.

[0134] 게다가, 도 3b 또는 3a의 스케일 인자 대역(6)과 상응하는 재구성 대역을 위한 에너지 값(E3)은 10개의 유닛과 동일하다. 중요하게도, 에너지 값은 스펙트럼 부분들(922, 923)의 에너지뿐만 아니라,

[0135] 인코더-면 상에서, 즉 예를 들면 음조성 마스크를 사용하여 스펙트럼 분석을 실행하기 전에 계산된 것과 같은 재구성 대역(920)의 완전한 에너지를 포함한다. 따라서, 10개의 에너지 유닛은 재구성 대역 내의 제 1 및 제 2 스펙트럼 부분들을 커버한다. 그때, 블록들(922, 923) 또는 블록들(922, 923)을 위한 원시 표적 범위 데이터를 위한 소스 범위 데이터의 에너지는 8개의 에너지 유닛과 동일한 것으로 가정된다. 따라서, 5개의 유닛의 손실 에너지가 계산된다.

[0136] 타일 에너지(t_{Ek})에 의해 세분된 손실 에너지를 기초로 하여, 0.79의 이득 인자가 계산된다. 그리고 나서 제 2 스펙트럼 부분들(922, 923)을 위한 원시 스펙트럼 라인들은 계산된 이득 인자와 곱해진다. 따라서, 제 2 스펙트럼 부분들(922, 923)을 위한 스펙트럼 값들만이 조정되고 제 1 스펙트럼 부분들(921)을 위한 스펙트럼 라인들은 이러한 엔벨로프 조정에 의해 영향을 받지 않는다. 제 2 스펙트럼 부분들(922, 923)을 위한 원시 스펙트럼 값들을 곱한 후에, 재구성 대역 내의 제 1 스펙트럼 부분들로 구성되고, 재구성 대역(920) 내의 제 2 스펙트럼 부분들(922, 923) 내의 스펙트럼 라인들로 구성되는 완전한 재구성 대역이 계산되었다.

[0137] 바람직하게는, 대역들(922, 923) 내의 원시 스펙트럼 데이터를 발생시키기 위한 소스 범위는 주파수와 관련하여, 지능형 캡 필링 시작 주파수(309) 아래이고 재구성 대역(920)은 지능형 캡 필링 시작 주파수(309) 위이다.

[0138] 게다가, 재구성 대역 경계들이 스케일 인자 대역 경계들과 일치하는 것이 바람직하다. 따라서, 재구성 대역은 일 실시 예에서, 코어 오디오 디코더의 스케일 인자 대역들의 상응하는 크기를 갖거나 또는 에너지 쌍이 적용될 때, 재구성 대역을 위한 에너지 값이 스케일 인자 대역들의 2 또는 그 이상의 정수를 제공하도록 크기화된다. 따라서, 스케일 인자 대역(4), 스케일 인자 대역(5) 및 스케일 인자 대역(6)을 위하여 에너지 축적이 실행된다. 그리고 가정될 때, 재구성 대역(920)의 낮은 주파수 경계는 스케일 인자 대역(4)의 낮은 경계와 동일하고 재구성 대역(920)의 높은 주파수 경계는 스케일 인자 대역(6)의 높은 경계와 동일하다.

[0139] 그 뒤에, 도 9a의 디코더의 도 다른 기능을 나타내기 위하여 도 9d가 설명된다. 오디오 디코더(900)는 스펙트럼 부분들의 제 1 세트의 제 1 스펙트럼 부분들과 상응하는 탈양자화된 스펙트럼 값들을 수신하고, 부가적으로 도 3b에 도시된 것과 같은 스케일 인자 대역들을 위한 스케일 인자들이 역 스케일링 블록(940)에 제공된다. 역 스케일링 블록(940)은 도 3a의 지능형 캡 필링 시작 주파수(309) 아래의 제 1 스펙트럼 부분들의 모든 제 1 세트, 및 부가적으로 지능형 캡 필링 시작 주파수(309) 위의 제 1 스펙트럼 부분들, 즉 도 9d의 941에 도시된 것과 같은 재구성 대역 내에 모두 위치되는 도 3a의 제 1 스펙트럼 부분들(304, 305, 306, 307)을 제공한다. 게다가, 재구성 대역 내의 주파수 타일 필링을 위하여 사용되는 소스 대역 내의 제 1 스펙트럼 부분들이 엔벨로프 조정기/계산기(942)에 제공되고 이러한 블록은 부가적으로 도 9d의 943에 도시된 것과 같이 인코딩된 오디오 신호에 대한 파라미터 부가 정보로서 제공되는 재구성 대역을 위한 에너지 정보를 수신한다. 그리고 나서, 엔벨로프 조정기/계산기(942)는 도 9b와 9c의 기능들을 제공하고 최종적으로 재구성 대역 내의 제 2 스펙트럼 부분들을 위한 조정된 스펙트럼 값들을 출력한다. 재구성 대역 내의 제 2 스펙트럼 부분들의 이러한 조정된 스펙트럼 값들(922, 923) 및 도 9d의 라인(941)에 표시된 재구성 대역 내의 제 1 스펙트럼 부분들은 공동으로 재구성 대역 내의 완전한 스펙트럼 표현을 표현한다.

[0140] 그 뒤에, 인코딩된 오디오 신호를 제공하거나 또는 발생시키기 위하여 오디오 신호의 인코딩을 위한 오디오 인코더의 바람직한 실시 예들을 설명하기 위하여 도 10a 내지 10b가 참조된다. 인코더는 스펙트럼 분석기(1004)를 공급하는 시간/스펙트럼 변환기(1002)를 포함하고, 스펙트럼 분석기(1004)는 한편으로는 파라미터 계산기(1006)에, 다른 한편으로는 오디오 인코더(1008)에 연결된다. 오디오 인코더(1008)는 제 1 스펙트럼 부분들의 제 1 세트의 인코딩된 표현을 제공하고 제 2 스펙트럼 부분들의 제 2 세트를 포함하지는 않는다. 다른 한편으로는, 파라미터 계산기(1006)는 제 1 및 제 2 스펙트럼 부분들을 포함하는 재구성 대역을 위한 에너지 정보를 제공한다. 게다가, 오디오 인코더(1008)는 제 1 스펙트럼 해상도를 갖는 제 1 스펙트럼 부분들의 제 1 세트의 제 1 인코딩된 표현을 발생시키도록 구성되며, 오디오 인코더(1008)는 블록(1002)에 의해 발생되는 스펙트럼 표현의 모든 대역을 위한 스케일 인자들을 제공한다. 부가적으로, 도 3b에 도시된 것과 같이, 인코더는 주파수와 관련하여, 적어도 도 3a에 도시된 것과 같은 지능형 캡 필링 시작 주파수(309) 위에 위치되는 재구성 대역들을 위한 에너지 정보를 제공한다. 따라서, 바람직하게는 스케일 인자 대역들과 일치하거나 또는 스케일 인자 대역들의

그룹들과 일치하는 재구성 대역들을 위하여, 두 개의 값, 즉 오디오 인코더(1008)로부터의 상응하는 스케일 인자 및 부가적으로, 파라미터 계산기(1006)에 의해 출력되는 에너지 정보가 주어진다.

[0141] 오디오 인코더는 바람직하게는 서로 다른 주파수 대역폭들, 즉 서로 다른 수의 스펙트럼 값들을 갖는 스케일 인자 대역들을 갖는다. 따라서, 파라미터 계산기는 특정 재구성 대역의 대역폭과 관련하여 사로 다른 대역폭을 위한 에너지들을 정규화하기 위한 정규화기(normalizer, 1012)를 포함한다. 이를 위하여, 정규화기(1012)는 입력들로서, 대역 내의 에너지 및 대역 내의 스펙트럼 값들의 수를 수신하고 정규화기(1012)는 그리고 나서 재구성/스케일 인자 대역 당 정규화된 에너지를 출력한다.

[0142] 게다가, 도 10a의 파라미터 계산기(1006a)는 도 10a의 라인(1007)에 의해 도시된 것과 같이 코어 또는 오디오 인코더(1008)로부터 제어 정보를 수신하는 에너지 값 계산기를 포함한다. 이러한 제어 정보는 오디오 인코더에 의해 사용되는 긴/짧은 블록들에 대한 정보 및/또는 그룹화 정보를 포함할 수 있다. 따라서, 긴/짧은 블록들에 대한 정보 및 짧은 윈도우들에 대한 그룹화 정보는 "시간" 그룹화와 관련되며, 그룹화 정보는 부가적으로 스펙트럼 그룹화, 즉 두 스케일 인자 대역의 단일 재구성 대역으로의 그룹화를 언급할 수 있다. 따라서, 에너지 값 계산기(1006)는 스펙트럼 부분들만이 그룹화되었을 때 제 1 및 제 2 스펙트럼 부분을 포함하는 각각의 그룹화된 대역을 위한 단일 에너지 값을 출력한다.

[0143] 도 10d는 스펙트럼 그룹화를 구현하는 또 다른 실시 예를 도시한다. 이를 위하여, 블록(1016)은 두 개의 인접한 대역을 위한 에너지 값들을 계산하도록 구성된다. 그리고 나서, 블록(1018)에서, 인접한 대역들을 위한 에너지 값들이 비교되고, 에너지 값들이 예를 들면 한계(threshold)에 의해 정의되는 것과 크게 다르지 않거나 또는 약간 다를 때, 블록(1020)에 표시된 것과 같이 두 대역을 위한 단일(정규화된) 값이 발생된다. 라인(1024)에 의해 도시된 것과 같이, 블록(1018)은 우회될 수 있다. 게다가, 블록(1020)에 의해 실행되는 두 개 또는 그 이상의 대역을 위한 단일 값의 발생은 인코더 비트레이트 제어(1024)에 의해 제어될 수 있다. 따라서, 비트레이트가 감소될 때, 인코딩된 비트레이트 제어(1024)는 블록(1018) 내의 비교가 에너지 정보 값들을 그룹화하도록 허용되지 않았더라도 두 개 또는 그 이상의 대역을 위한 단일의 정규화된 값을 발생시키도록 블록(1020)을 제어한다.

[0144] 오디오 인코더가 두 개 또는 그 이상의 짧은 윈도우의 그룹화를 실행하는 경우에 있어서, 이러한 그룹화는 또한 에너지 정보를 위하여 적용된다. 코어 인코더가 두 개 또는 그 이상의 짧은 블록의 그룹화를 실행할 때, 이러한 두 개 또는 그 이상의 블록을 위하여, 스케일 인자들의 단일 세트만이 계산되고 전송된다. 디코더-면 상에서, 오디오 디코더는 그리고 나서 두 그룹화된 윈도우를 위하여 동일한 스케일 인자들의 세트를 적용한다.

[0145] 에너지 정보 계산과 관련하여, 재구성 대역 내의 스펙트럼 값들은 두 개 또는 그 이상의 짧은 윈도우에 대하여 축적된다. 바꾸어 말하면, 이는 짧은 블록 및 뒤따르는 짧은 블록을 위한 특정 재구성 대역 내의 스펙트럼 값들이 함께 축적되고 두 개의 짧은 블록을 포함하는 이러한 재구성 대역을 위하여 단일 에너지 정보 값만이 전송되는 것을 의미한다. 그리고 나서 디코더-면 상에서, 도 9a 내지 9d와 관련하여 설명된 엔벨로프 조정이 각각의 짧은 블록을 위하여 개별적으로 실행되지 않고 그룹화된 짧은 윈도우들의 세트를 위하여 함께 실행된다.

[0146] 상응하는 정규화는 그리고 나서 주파수 내의 어떠한 그룹화 또는 시간 내의 그룹화가 실행되었더라도 디코더-면 상에서 에너지 값 정보 계산을 위하여, 한편으로는 에너지 정보 값 및 재구성 대역 또는 그룹화된 재구성 대역들의 세트 내의 스펙트럼 라인들의 양이 알려져야만 하도록 쉽게 허용한다.

[0147] 게다가, 스펙트럼 에너지들에 대한 정보, 개별 에너지들에 대한 정보 또는 개별 에너지 정보, 생존 에너지(survive energy)에 대한 정보 또는 생존 에너지 정보, 타일 에너지에 대한 정보 또는 타일 에너지 정보, 또는 손실 에너지에 대한 정보 또는 손실 에너지 정보는 에너지 값 뿐만 아니라 최종 에너지 값이 유도될 수 있는, (예를 들어, 절대) 진폭 값, 레벨 값 또는 어떠한 다른 값을 포함할 수 있다. 이와 같이, 에너지에 대한 정보는, 예를 들어, 절대 진폭의 및/또는 진폭의 및/또는 레벨의 값 및/또는, 에너지 값 그 자체를 포함할 수 있다.

[0148] 도 12a는 디코딩을 위한 장치의 추가 실시 예를 도시한다. 비트스트림은, 예를 들어, AAC 디코더 일 수 있는 코어 디코더(1200)에 의해 수신된다. 상기 결과는 예를 들어 주파수 재생기(604)에 대응하는 대역폭 확장 패칭 또는 타일링(1202)을 수행하는 단계로 구성된다. 이후, 패치/타일 적용의 절차 및 후-처리가 수행되며, 패치 적용이 수행될 때, 주파수 재생기(1202)는 예를 들어, 조정된 주파수 경계들과 함께 추가 주파수 재생을 수행하도록 제어된다. 게다가, 음조 라인들의 제거 및 감쇠 같은 것에 의해 패치 처리가 수행될 때, 상기 결과는 예를 들어, 블록(712 또는 826)의 맥락에서 논의된 것처럼 파라미터-구동 대역폭 엔벨로프 성형을 수행하는 블록(1206)으로 포워딩된다(보내진다, forwarded). 상기 결과는 이후 예를 들어, 도 12a에서 도시되는 것처럼 PCM

출력 영역인 최종 출력 영역으로 변환을 수행하기 위해 합성 변환 블록(1208)으로 포워딩된다.

[0149] 본 발명 실시예의 주요 특징은 다음과 같다:

[0150] 바람직한 실시예는 음조 스펙트럼 영역들이 크로스-오버 주파수의 안좋은 선택 및/또는 폐치 마진들(patch margins)에 의해 축소되는 경우, 또는 음조 구성요소들은 폐치 경계들의 부근에 너무 가깝게 위치되는 경우, 상기 참조된 워블링(warbling) 아티팩트들을 나타내는 MDCT에 기반한다.

[0151] 도 12b는 어떻게 새로 제안된 기술이 최신 대역폭 확장 방법들에서 발견되는 아티팩트들을 감소시키는지를 보여준다. 도 12 폐널 (2)에서, 현대 대역폭 확장 방식의 출력의 양식화된 크기 스펙트럼이 보여진다. 이 예에서, 상기 신호는 두개의 인접 톤들에 의해 야기되는 비팅(beating)들에 의해, 그리고 톤(tone)의 분할에 의해서도 지각적으로 악화된다. 양쪽의 문제가 있는 스펙트럼 영역들은 각각 원(circle)으로 마킹된다.

[0152] 이러한 문제들을 극복하기 위해, 새로운 기술은 먼저 신호에 포함된 음조 구성요소들의 스펙트럼 위치를 검출한다. 이후, 본 발명의 하나의 관점에 따라, 음조 구성요소들의 분할 또는 비팅이 최소화되도록 (주어진 제한 내에서) 개별 시프트들(shifts)에 의해 LF(저주파수) 및 모든 폐치들 사이의 전이 주파수들을 조정하는 것이 시도된다. 이러한 목적을 위해, 전이 주파수는 바람직하게는 지역적 스펙트럼 최소값을 매칭시켜야 한다. 이 단계는

도 12b 폐널 (2) 및 폐널 (3)에서 보여지며, 여기서 전이 주파수 f_{x2} 는 고주파수들로 시프팅되며, f'_{x2} 를 도출한다.

[0153] 본 발명의 또 다른 관점에 따라, 전이 영역에서 문제가 있는 스펙트럼 컨텐츠가 잔존하는 경우, 적어도 하나의 잘못 위치된 음조 구성요소들이 워블링(warbling) 또는 전이 주파수들에서 비팅 아티팩트들을 감소시키기 위해 제거된다. 이는 도 2 폐널 (3)에서 보여지는 것처럼, 스펙트럼 외삽 또는 보간 / 필터링을 통해 처리된다. 음조 구성요소는 그래서 풋-포인트(foot-point)에서 풋-포인트까지, 즉 그 왼쪽 지역적 최소값에서 그것의 오른쪽 지역적 최소값까지, 제거된다. 본 기술의 응용 후에 상기 결과 스펙트럼 도 12b 폐널 (4)에서 보여진다.

[0154] 다른 말로, 도 12b는, 상부 왼쪽 코너에서, 즉 폐널 (1)에서, 원래 신호를 도시한다. 상부 오른쪽 코너에서, 즉 폐널 (2)에서, 타원들(1220 및 1221)에 의해 마킹된 문제가 되는 영역들을 갖는 비교 대역폭 확장 신호가 보여진다. 하부 왼쪽 코너에서, 즉 폐널 (3)에서, 두개의 바람직한 폐치 또는 주파수 타일 처리 특징들이 도시된다. 음조 부분들의 분할(splitting)은 주파수 경계 f'_{x2} 를 증가시키도록 처리되어 대응 음조 부분의 클리핑(clipping)이 거기에 더 이상 없게 된다. 게다가, 음조 부분(1031 및 1032)을 제거하기 위한 이득 함수(1030)가 적용되거나, 또는, 대안적으로, (1033)에 의해 도시되는 보간이 표시된다. 최종적으로, 도 12b의 하부 오른쪽 코너는, 즉 폐널(4)는 한쪽에서 타일/폐치 주파수의 조합으로부터 도출되는 향상된 신호 또는 문제가 있는 음조 부분들의 적어도 감쇠를 나타낸다.

[0155] 도 12b의 폐널 (1)은, 원래 스펙트럼 이전에 논의된 것처럼, 원래 스펙트럼은 크로스-오버 또는 캡 필링 시작 주파수 f_{x1} 까지의 코어 주파수 범위를 갖는다.

[0156] 이와 같이, 주파수 f_{x1} 는 나이퀴스트 주파수(Nyquist frequency)($f_{Nyquist}$)보다 작거나 그와 동일한 최대 주파수 및 경계 주파수(1250) 사이에서 확장하는 복원 범위(1254) 및 소스 범위(1252) 사이의 경계 주파수(1250)를 도시한다. 인코더-측면에서, 신호는 f_{x1} 에서 대역폭-제한되는 것으로 가정되고, 또는 지능형 캡 필링에 관한 기술이 적용될 때, f_{x1} 이 도 3a의 캡 필링 시작 주파수(309)에 대응하는 것으로 가정된다. 상기 기술에 기반하여, f_{x1} 위의 복원 범위는 (도 13a, 13b의 경우에) 비어있게(empty) 될 것이고 도 3a의 맥락에서 논의된 것처럼 고해상도로 인코딩될 특정 제1스펙트럼 부분들을 포함할 것이다.

[0157] 도 12b, 폐널(2)는 예를 들어, 문제가 있는 두개의 부분들을 갖는 도 7a의 블록(702)에 의해 발생되는, 예비 재생 신호(preliminary regenerated signal)를 도시한다. 문제가 있는 하나의 부분은 (1220)에서 도시된다. (1220a)에서 도시되는 코어 영역 내의 음조 부분 및 (1220b)에서 도시되는 주파수 타일의 시작에서의 음조 부분 사이의 주파수 거리는 너무 작아서 비팅 아티팩트(beating artifact)가 생성될 것이다. 추가 문제는 (1225)에서 도시된 주파수 타일링(tiling) 작업 또는 제1폐칭 작업에 의해 발생된 제1주파수 타일의 상부 경계에서의 중간-클리핑된(halfway-clipped) 또는 분할된 음조 부분(1226)이다. 이 음조 부분(1226)이 도 12b의 다른 음조 부분들과 비교될 때, 그 너비가 일반적인 음조 부분의 너비보다 작다는 것을 명확해 질 것이고 이는 이러한 음조 부분이 소스 범위(1252)의 잘못된 위치에서 제1주파수 타일(1225) 및 제2주파수 타일(1227) 사이의 주파수 경계를 설정하여 분할되었다는 것을 의미한다. 이러한 이슈를 다루기 위해, 경계 주파수 f_{x2} 는 도 12b의 폐널 (3)에서

도시되는 것처럼 조금 더 커지도록 수정되었고, 이 음조 부분의 클리핑은 일어나지 않는다.

[0158] 다른 한편으로, f'_{x2} 이 변화된 이러한 절차는 비팅 문제(beating problem)를 효과적으로 처리하지 않으며, 그래서 이는 도 7a의 블록(708)의 맥락에서 논의되는 것처럼 어떠한 다른 절차들 또는 보간 또는 필터링에 의해 음조 구성요소들의 제거에 의해 처리된다. 이와 같이, 도 12b는 (708)에서 도시되는 경계들에서 음조 부분들의 제거 및 전이 주파수 조정(706)의 순차적 적용을 도시한다.

[0159] 또 다른 옵션은 전이 경계 f_{x1} 가 약간 낮아지도록 그래서 음조 부분(1220a)이 코어 범위에 더 이상 있지 않도록 설정하는 것일 것이다. 이후, 음조 부분(1220a)은 낮은 값(lower value)으로 전이 주파수(transition frequency) f_{x1} 를 설정하여 제거되거나 감소되었다.

[0160] 이러한 절차는 문제가 있는 음조 구성요소(1032)를 갖는 이슈를 처리하도록 작동하였다. f'_{x2} 를 훨씬 더 높게 설정하여, 음조 부분(1032)가 위치되는 스펙트럼 부분은 제1폐칭 작업(1225) 내에서 재생될 수 있었고, 그래서 두개의 인접 또는 근처 음조 부분들은 발생되지 않았을 것이다.

[0161] 기본적으로, 비팅 문제는 인접 음조 부분들의 주파수에서 거리 및 진폭에 의존한다. 검출기(704, 720) 또는 언급된 더 일반적인, 분석기(602)는 바람직하게는 어떠한 음조 구성요소를 위치시키기 위해 f_{x1} , f_{x2} , f'_{x2} 같은 전이 주파수 아래의 주파수에 위치된 하부 스펙트럼 부분의 분석이 분석되는 방식으로 구성된다. 게다가, 전이 주파수위의 스펙트럼 범위가 음조 구성요소를 검출하기 위해서도 분석된다. 상기 검출(detection)가, 주파수에 관해 전이 주파수의 왼쪽에 하나 그리고 (상승하는 주파수에 관해) 오른쪽에 하나로, 두개의 음조 구성요소들을 도출할 때, 이후 도 7a에서 (708)로 도시된 경계들에서의 음조 구성요소들의 제거가 활성화된다(activated). 상기 음조 구성요소의 검출은, 대응 대역의 대역폭에 관해 양쪽 방향으로 적어도 20%인, 전이 주파수로부터, 확장하는 특정 검출 범위에서 수행되며, 바람직하게는 상기 대역폭에 관련된 전이 주파수의 오른쪽에 대해 위쪽으로 그리고 전이 주파수의 왼쪽에 대해 아래쪽으로 10%까지만 확장하는, 즉 한쪽으로는 소스 범위의 대역폭 그리고 다른 한쪽으로는 복원 범위에서이며, 또는, 전이 주파수가 두개의 주파수 타일들(1225, 1227) 사이의 전이 주파수일 때, 대응 주파수 타일의 대응하는 10% 양에서이다. 추가 실시예에서, 미리 결정된 검출 대역폭은 1 바크(one Bark)이다. 폐치 경계 주변의 1 바크의 범위 내의 음조 부분들을 제거하는 것이 가능해야만 하며, 그래서 완전 검출 범위는 2 바크이고, 즉 저대역의 1 바크 그리고 고대역의 1 바크이며, 여기서 저대역의 1 바크는 고대역의 1 바크에 바로 인접한다.

[0162] 본 발명의 또 다른 관점에 따라서, 필터 링잉 아티팩트를 감소시키기 위해, 주파수 영역의 크로스-오버 필터는 두개의 연속 스펙트럼 영역들에 적용되며, 즉 제1폐치와 코어 대역 사이 또는 두개의 폐치들 사이에서 적용된다. 바람직하게는, 크로스-오버 필터는 신호 적응적이다.

[0163] 크로스 오버 필터는, 낮은 스펙트럼 영역에 적용되는 페이드-아웃 필터 \mathbf{h}_{out} , 및 높은 스펙트럼 영역에 적용되는 페이드-인 필터 \mathbf{h}_{in} 인, 두개의 필터들로 구성된다.

[0164] 각 필터들은 N 의 길이를 갖는다.

[0165] 게다가, 양쪽 필터들의 기울기는 $0 \leq Xbias \leq N$ 와 함께, 크로스-오버 필터의 노치(notch) 특성을 결정하는 $Xbias$ 로 불리는 신호 적응 값에 의해 특징지어진다:

[0166] 만약 $Xbias = 0$ 인 경우, 양쪽 필터들의 합은 1과 동일하고, 즉 결과 필터에 노치 필터 특성이 없다는 것이다.

[0167] 만약 $Xbias = N$ 라면, 양쪽 필터들은 완전히 제로(zero)이다.

[0168] 크로스-오버 필터들의 기본 디자인은 다음 방정식들에 제약된다:

$$\mathbf{h}_{out}(k) = \mathbf{h}_{in}(N - 1 - k), \forall Xbias$$

$$\mathbf{h}_{out}(k) + \mathbf{h}_{in}(k) = 1, Xbias = 0$$

[0171] $k = 0, 1, \dots, N - 1$ 는 주파수 지수(frequency index)이다. 도 12c는 그러한 크로스-오버 필터의 예를 보여준다.

[0172] 이 예에서, 다음 방정식은 필터 h_{out} 를 생성하는데 이용된다:

$$h_{out}(k) = 0.5 + 0.5 \cdot \cos \left(\frac{k}{N-1-Xbias} \cdot \pi \right), k = 0, 1, \dots, N-1-Xbias$$

[0174] 다음 방정식은 필터들 h_{in} 및 h_{out} 이 이후 어떻게 적용되는지를 설명한다.

$$Y(k_t - (N-1) + k) = \mathbf{LF}(k_t - (N-1) + k) \cdot h_{out}(k) +$$

$$\mathbf{HF}(k_t - (N-1) + k) \cdot h_{in}(k), \quad k = 0, 1, \dots, N-1$$

[0177] Y 는 집합된 스펙트럼(assembled spectrum)을 나타내고, k_t 는 전이 주파수이며, \mathbf{LF} 는 저주파수 컨텐츠 그리고

\mathbf{HF} 는 고주파수 컨텐츠이다.

[0178] 다음, 이 기술의 이점의 증거가 표현된다. 다음 예들에서 원래 신호는 트랜지언트-유사 신호(transient-like signal)이며, 특히 22kHz의 차단 주파수(cut-off frequency)를 갖는, 그것의 저역 통과 필터링된 버전이다. 첫째로, 이 트랜지언트는 변환 영역에서 6 kHz로 제한된 대역이다. 이후, 저역 통과 필터링된 원래 신호의 대역폭은 24 kHz로 확장된다. 대역폭 확장은 변환 내에서 6 kHz 위에서 이용가능한 주파수 범위를 완전히 채우기 위해 LF(저주파수) 대역을 세번 복제하는 것을 통해 달성된다.

[0179] 도 11a는 (스피치 퍼크들 (1100)) 변환의 브릭-월(장벽, brick-wall) 특성 때문에 트랜지언트를 스펙트럼적으로 둘러싸는 필터 링잉 아티팩트의 일반적 스펙트럼으로 고려될 수 있는, 이러한 신호의 스펙트럼을 보여준다. 발명의 접근을 적용하여, 필터 링잉은 각 전이 주파수에서 근사적으로 20 dB만큼 감소된다 (감소된 스피치 퍼크들).

[0180] 동일 효과는, 상이한 도시에서, 도 11b, 11c에서 보여진다. 도 11b는 어떠한 필터 링잉 감소 없이 설명된 대역폭 확장 기술 위에 적용한 후 트랜지언트를 시간적으로 선행하고 후행하는 필터 링잉 아티팩트를 갖는 언급된 트랜지언트 유사 신호의 스펙트로그램을 보여준다. 수평 라인들 각각은 연속 패치들 사이의 전이 주파수에서 필터 링잉을 표현한다. 도 6은 대역폭 확장 내에서 발명의 접근을 적용한 후의 동일 신호를 보여준다. 링잉 감소의 적용을 통해, 필터 링잉은 이전 도면에서 표시된 신호와 비교하여 근사적으로 20 dB만큼 감소된다.

[0181] 이후, 도 14a, 14b는 분석기 특징을 갖는 맥락에서 이미 논의된 크로스-오버 필터 발명 관점을 더 도시하기 위해 논의된다. 그러나, 크로스-오버 필터(710)는 도 6a-7b의 맥락에서 논의된 발명에 독립적으로 실행될 수도 있다.

[0182] 도 14a는 파라미터 데이터에 대한 정보 및 인코딩된 코어 신호를 포함하는 인코딩된 오디오 신호를 디코딩하는 장치를 도시한다. 상기 장치는 디코딩된 코어 신호를 얻기 위해 인코딩된 코어 신호를 디코딩하는 코어 디코더(1400)를 포함한다. 코딩된 코어 신호는 도 13a의 맥락에서 대역폭이 제한될 수 있고, 도 13b 실시예 또는 코어 디코더는 도 1 내지 5c 또는 9a-10d의 맥락에서 풀 레이트 코더(full rate coder) 또는 총 주파수 범위(full frequency range)일 수 있다.

[0183] 게다가, 디코딩된 코어 신호에 포함되지 않는 주파수들을 갖는 하나 이상의 스펙트럼 타일들을 재생하기 위한 타일 발생기(1404)는 디코딩된 코어 신호의 스펙트럼 부분을 이용하여 발생된다. 상기 타일들은, 예를 들어, 도 3a의 맥락에서 도시된 것처럼, 복원 대역 내에서 복원된 제2스펙트럼 부분들일 수 있고 또는 이는 고해상도로 복원될 제1스펙트럼 부분들을 포함할 수 있지만, 대안적으로, 스펙트럼 타일들은 또한 도 13dptj 도시된 것처럼 인코더가 엄격한 대역 제한을 수행할 때 완전히 빈 주파수 대역들을 포함할 수 있다.

[0184] 게다가, 크로스-오버 필터(1406)는 갑 필팅 주파수(309)로부터 제1타일 정지 주파수까지 확장하는 주파수들을 갖는 제1주파수 타일 및 디코딩된 코어 신호를 크로스-오버 필터링하기 위해 또는 제2주파수 타일 및 제1주파수 타일(1225)을 스펙트럼적으로 크로스-오버 필터링하기 위해 제공되며, 제2주파수 타일은 제1주파수 타일(1225)의 상부 경계 주파수에 주파수-인접한 하부 경계 주파수를 갖는다.

[0185] 추가 실시예에서, 상기 크로스-오버 필터(1406) 출력 신호는 엔벨로프-조정되어 재생된 신호(envelope-adjusted

regenerated signal)를 최종적으로 얻기 위해 파라미터 부가 정보로서 인코딩된 오디오 신호에 포함된 파라미터 스펙트럼 엔벨로프 정보를 적용하는 엔벨로프 조정기(envelope adjuster, 1408)로 입력된다. 구성요소들(1404, 1406, 1408)은 예를 들어, 도 13b, 도 1b 또는 도 6a에서 도시된 것처럼, 주파수 재생기로서 실행될 수 있다.

[0186] 도 14b는 크로스-오버 필터(1406)의 추가 실시예를 도시한다. 크로스-오버 필터(1406)는 제1입력 신호 IN1을 수신하는 페이드-아웃 서브필터, 및 제2입력 IN2를 수신하는 두번쩨 페이드-인 서브필터(1422)를 포함하며 양 필터들(1420) 및 (1422)의 결과들 또는 출력들은 예를 들어, 가산기(애더, adder)인 결합기(combiner, 1424)에 제공된다. 가산기 또는 결합기(1424)는 주파수 빈들에 대해 스펙트럼 값들을 출력한다. 도 12c는 페이드-아웃 서브필터 특성(1420a) 및 페이드-인 서브필터 특성(1422a)을 포함하는 예시 크로스-페이드 함수를 도시한다. 양 쪽 필터들은 21과 동일한, 즉 $N=21$ 인, 도 12c의 예에서 특정 주파수 중첩(오버랩, overlap)을 갖는다. 이와 같이, 예를 들어, 소스 영역(1252)의 다른 주파수 값들은 영향받지 않는다. 오직 소스 범위(1252)의 가장 높은 21 주파수 빈들은 페이드-아웃 함수(1420a)에 의해 영향받는다.

[0187] 다른 한편으로, 오직 제1주파수 타일(1225)의 가장 낮은 21 주파수 라인들은 페이드-인 함수(1422a)에 의해 영향받는다.

[0188] 추가적으로, 9 및 13 사이의 주파수 라인들이 영향받는다는 것이 크로스-페이드 함수로부터 명확해지지만, 페이드-인 함수는 사실 1 및 9 사이의 주파수 라인들에 영향을 미치지 않으며 페이드-아웃 함수(1420a)는 13 및 21 사이의 주파수 함수들에 영향을 미치지 않는다. 이는 주파수 라인들 9 및 13 사이에서 오직 중첩이 필요하며, f_{x1} 같은 크로스-오버 주파수가 주파수 샘플 또는 주파수 빈(11)에 위치된다는 것을 의미한다. 이와 같이, 오직 소스 범위 및 제1주파수 타일 사이의 주파수 값들 또는 두개의 주파수 빈들의 중첩은 크로스-오버 또는 크로스-페이드 함수를 실행하기 위해 요구될 것이다.

[0189] 특정 실시예에 의존하여, 높은 또는 낮은 중첩이 적용될 수 있고, 추가적으로, 코사인 함수로부터 떨어진 다른 페이딩 함수들이 이용될 수 있다. 게다가, 도 12c에서 도시된 것처럼, 크로스-오버 범위에서 특정 노치(notch)를 적용하는 것이 바람직하다. 다르게 언급하자면, 경계 범위에서의 에너지는 노치가 없는(노치-프리, notch-free) 크로스-페이드 함수의 경우가 되는 것처럼 양쪽 필터 함수들이 결국 통합되지 않는다는 사실 때문에 감소될 것이다. 주파수 타일의 경계들에 대한 에너지 손실, 즉 제1주파수 타일은 상부측 경계 및 하부측 경계에서 감쇠될 것이고, 에너지들은 대역들의 중간에 더 집중된다. 그러나, 스펙트럼 엔벨로프 조정이 크로스-오버 필터에 의해 처리되는 것 이후에 일어난다는 사실 때문에, 전체 주파수는 터치되지 않고(not touched), 도 3a의 맥락에서 논의된 것처럼 대응 스케일 인수들같이 스펙트럼 엔벨로프 데이터에 의해 정의된다. 다른 말로, 도 9b의 계산기(918)는, 크로스-오버 필터의 출력인, "이미 발생된 원(raw) 타겟 범위"를 이후 계산한다. 게다가, 보간에 의한 음조 부분의 제거로 인한 에너지 손실은 또한 이러한 제거가 낮은 타일 에너지를 도출한다는 사실 때문에 보상될 것이고 완전 복원 대역에 대한 이득 인자는 더 높아질 것이다. 그러나, 다른 한편으로, 크로스-오버 주파수는 주파수 타일의 중간에 더 많은 에너지 집중을 도출하며, 이는 결국, 도 11a-11c의 맥락에서 논의된 것처럼 트랜지언트들에 의해 부분적으로 야기되는, 아티팩트들을 효과적으로 감소시킨다.

[0190] 도 14b는 상이한 입력 조합들을 도시한다. 소스 주파수 범위 및 주파수 타일 사이의 경계에서 필터링하기 위하여, 입력 1은 코어 범위의 상부 스펙트럼 부분이며, 입력 2는 제1주파수 타일의 또는, 오직 단일 주파수 타일만이 존재할 때, 단일 주파수 타일의 낮은 스펙트럼부분이다. 게다가, 상기 입력은 제1주파수 타일이 될 수 있고 전이 주파수는 제1타일의 상부 주파수 경계가 될 수 있으며 서브필터(1422)에 대한 입력은 제2주파수 타일의 낮은 부분이 될 것이다. 추가 제3주파수 타일이 존재할 때, 추가 전이 주파수는 제2주파수 타일 및 제3주파수 타일 사이의 주파수 경계가 될 것이고, 도 12 특성이 이용될 때, 페이드-아웃 서브필터(1421)에 대한 입력은 필터 파라미터에 의해 결정되는 것처럼 제2주파수 타일의 상부 스펙트럼 범위가 될 것이며, 페이드-인 서브필터(1422)에 대한 입력은 제3주파수 타일의 낮은 부분 및, 도 12c의 예에서처럼, 가장 낮은 21 스펙트럼 라인들이 될 것이다.

[0191] 도 12c에서 도시되는 것처럼, 페이드-인 서브필터 및 페이드-아웃 서브필터에 대해 파라미터 N 이 동일하도록 하는 것이 바람직하다. 그러나 이는 필수적인 것은 아니다. N 에대한 값은 변화할 수 있고 그 결과는 이후 필터 "노치"가 하부 및 상부 범위 사이에서 비대칭적이 될 것이다. 추가적으로, 페이드-인/페이드-아웃 함수들이 반드시 도 12c에서처럼 동일 특성이어야 할 필요는 없다. 대신에, 비대칭적 특성들이 이용될 수도 있다.

[0192] 게다가, 크로스-오버 필터 특성을 신호-적용적으로 만드는 것을 바람직하다. 그래서, 신호 분석에 기반하여, 필터 특성은 적용적이다. 크로스-오버 필터가 특히 트랜지언트 신호들에 대해 유용하다는 사실 때문에, 트랜지언트 신호들이 발생하는지 여부가 검출된다. 트랜지언트 신호들이 발생할 때, 도 12c에서 도시된 것 같은 필터 특

성이 이용될 수도 있다. 그러나, 비-트랜지언트 신호가 검출될 때, 크로스-오버 필터의 영향력을 감소시키기 위해 필터 특성을 변화시키는 것이 바람직하다. 예를 들어, 이는 N을 제로(zero, 0)로 설정하여 또는 X_{bias} 를 제로로 설정하여 얻어질 수 있고 이에 의해 양쪽 필터들의 합은 1과 동일해지며, 즉 결과 필터에 노치 필터 특성이 없다는 것이다. 대안적으로, 크로스-오버 필터(1406)은 비-트랜지언트 신호들의 경우에 간단히 우회될 수 있다. 그러나, 바람직하게는, 파라미터 N , X_{bias} 을 변화시켜 상대적으로 느리게 필터 특성을 변화시키는 것이 필터 특성들의 빠른 변화로 인해 얻어지는 아티팩트들을 피하기 위해 바람직하다. 게다가, 저역-통과 필터는 특정 트랜지언트/음조 검출기에 의해 검출될 때 비록 신호가 더 빠르게 변화하더라도 그러한 상대적으로 작은 필터 특성 변화만을 허용하는 것이 바람직하다. 검출기는 도 14a의 (1405)에서 도시된다. 그것은 타일 발생기(1404)의 출력 신호 또는 타일 발생기에 대한 입력 신호를 수신할 수 있고 또는 예를 들어, AAC 디코딩으로부터 짧은 블록 표시같은 트랜지언트/비-트랜지언트 정보를 얻기 위해 코어 디코더(1400)에 연결될 수도 있다. 자연스럽게, 도 12c에서 보여지는 것과 다른 어떠한 다른 크로스오버 필터 또한 이용될 수 있다.

[0193] 이후, 트랜지언트 검출에 기반하여, 또는 음조 검출에 기반하여 또는 어떠한 다른 신호 특성 검출에 기반하여, 크로스-오버 필터(1406) 특성은 논의된 것처럼 변경된다.

[0194] 인코딩 또는 디코딩을 위한 장치의 맥락에서 일부 양상들이 설명되었으나, 이러한 양상들은 또한 블록 또는 장치가 방법 단계 또는 방법 단계의 특징과 상응하는, 상응하는 방법의 설명을 나타낸다는 것은 자명하다. 유사하게, 방법 단계의 맥락에서 설명된 양상들은 또한 상응하는 블록 아이템 혹은 상응하는 장치의 특징을 나타낸다. 일부 또는 모든 방법 단계는 예를 들면 마이크로프로세서, 프로그램가능 컴퓨터 또는 전자 회로 같은 하드웨어 장치에 의해(또는 사용하여) 실행될 수 있다. 일부 실시 예들에서, 하나 또는 그 이상의 각종 중요한 방법 단계는 그러한 장치에 의해 실행될 수 있다.

[0195] 본 발명에 따른 일부 실시 예들은 여기에 설명된 방법들 중 어느 하나가 실행되는 것과 같이, 프로그램가능 컴퓨터 시스템과 협력할 수 있는, 전자적으로 판독가능한 제어 신호들을 갖는 데이터 캐리어를 포함한다.

[0196] 일반적으로, 본 발명의 실시 예들은 프로그램 코드를 갖는 컴퓨터 프로그램 제품으로서 구현될 수 있으며, 프로그램 코드는 컴퓨터 프로그램 제품이 컴퓨터 상에서 구동할 때 방법들 중 어느 하나를 실행하도록 운영될 수 있다. 프로그램 코드는 예를 들면, 기계 판독가능 캐리어 상에 저장될 수 있다.

[0197] 다른 실시 예들은 기계 판독가능 캐리어 상에 저장되는, 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 컴퓨터 프로그램을 포함한다.

[0198] 바꾸어 말하면, 본 발명의 방법의 일 실시 예는 따라서 컴퓨터 프로그램이 컴퓨터 상에 구동할 때, 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 프로그램 코드를 갖는 컴퓨터 프로그램이다.

[0199] 본 발명의 방법의 또 다른 실시 예는 따라서 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 컴퓨터 프로그램을 포함하는, 그 안에 기록되는 데이터 캐리어(또는 데이터 저장 매체, 또는 컴퓨터 판독가능 매체)이다. 데이터 캐리어, 디지털 저장 매체 또는 기록된 매체는 일반적으로 유형 및/또는 비-일시적이다.

[0200] 본 발명의 방법의 또 다른 실시 예는 따라서 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 컴퓨터 프로그램을 나타내는 데이터 스트림 또는 신호들의 시퀀스이다. 데이터 스트림 또는 신호들의 시퀀스는 예를 들면 데이터 통신 연결, 예를 들면 인터넷을 거쳐 전송되도록 구성될 수 있다.

[0201] 또 다른 실시 예는 여기에 설명된 방법들 중 어느 하나를 실행하도록 구성되거나 혹은 적용되는, 처리 수단, 예를 들면 컴퓨터, 또는 프로그램가능 논리 장치를 포함한다.

[0202] 또 다른 실시 예는 그 안에 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 컴퓨터 프로그램이 설치된 컴퓨터를 포함한다.

[0203] 본 발명에 따른 또 다른 실시 예는 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 컴퓨터를 수신기로 전달하도록(예를 들면, 전자적으로 또는 광학적으로) 구성되는 장치 또는 시스템을 포함한다. 수신기는 예를 들면, 컴퓨터 모바일 장치, 메모리 장치 등일 수 있다. 장치 또는 시스템은 예를 들면 컴퓨터 프로그램을 수신기로 전달하기 위한 파일 서버를 포함할 수 있다.

[0204] 일부 실시 예들에서, 여기에 설명된 방법들 중 일부 또는 모두를 실행하기 위하여 프로그램가능 논리 장치(예를 들면, 필드 프로그램가능 게이트 어레이)가 사용될 수 있다. 일부 실시 예들에서, 필드 프로그램가능 게이트 어레이는 여기에 설명된 방법들 중 어느 하나를 실행하기 위하여 마이크로프로세서와 협력할 수 있다.

일반적으로, 방법들은 바람직하게는 어떠한 하드웨어 장치에 의해 실행된다.

[0205] 위에 설명된 실시 예들은 단지 본 발명의 원리들을 위한 설명이다. 여기에 설명된 배치들과 상세내용들의 변형과 변경은 통상의 지식을 가진 자들에 자명할 것이라는 것을 이해할 것이다. 따라서, 본 발명은 여기에 설명된 실시 예들의 설명에 의해 표현된 특정 상세내용이 아닌 특허 청구항의 범위에 의해서만 한정되는 것으로 의도된다.

[0206] 인용문헌의 목록

[0207] [1] Dietz, L. Liljeryd, K. Kjrling and O. Kunz, Spectral Band Replication, a novel approach in audio coding, in 112th AES Convention, Munich, May 2002.

[0208] [2] Ferreira, D. Sinha, Accurate Spectral Replacement, Audio Engineering Society Convention, Barcelona, Spain 2005.

[0209] [3] D. Sinha, A. Ferreira1 and E. Harinarayanan, A Novel Integrated Audio Bandwidth Extension Toolkit (ABET), Audio Engineering Society Convention, Paris, France 2006.

[0210] [4] R. Annadana, E. Harinarayanan, A. Ferreira and D. Sinha, New Results in Low Bit Rate Speech Coding and Bandwidth Extension, Audio Engineering Society Convention, San Francisco, USA 2006.

[0211] [5] T. ernicki, M. Bartkowiak, Audio bandwidth extension by frequency scaling of sinusoidal partials, Audio Engineering Society Convention, San Francisco, USA 2008.

[0212] [6] J. Herre, D. Schulz, Extending the MPEG-4 AAC Codec by Perceptual Noise Substitution, 104th AES Convention, Amsterdam, 1998, Preprint 4720.

[0213] [7] M. Neuendorf, M. Multrus, N. Rettelbach, et al., MPEG Unified Speech and Audio Coding-The ISO/MPEG Standard for High-Efficiency Audio Coding of all Content Types, 132nd AES Convention, Budapest, Hungary, April, 2012.

[0214] [8] McAulay, Robert J., Quatieri, Thomas F. Speech Analysis/Synthesis Based on a Sinusoidal Representation. IEEE Transactions on Acoustics, Speech, And Signal Processing, Vol 34(4), August 1986.

[0215] [9] Smith, J.O., Serra, X. PARSHL: An analysis/synthesis program for non-harmonic sounds based on a sinusoidal representation, Proceedings of the International Computer Music Conference, 1987.

[0216] [10] Purnhagen, H.; Meine, Nikolaus, "HILN-the MPEG-4 parametric audio coding tools," Circuits and Systems, 2000. Proceedings. ISCAS 2000 Geneva. The 2000 IEEE International Symposium on , vol.3, no., pp.201,204 vol.3, 2000

[0217] [11] International Standard ISO/IEC 13818-3, Generic Coding of Moving Pictures and Associated Audio: Audio, Geneva, 1998.

[0218] [12] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, K. Akagiri, H. Fuchs, M. Dietz, J. Herre, G. Davidson, Oikawa: "MPEG-2 Advanced Audio Coding", 101st AES Convention, Los Angeles 1996

[0219] [13] J. Herre, Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction, 17th AES International Conference on High Quality Audio Coding, August 1999

[0220] [14] J. Herre, Temporal Noise Shaping, Quantization and Coding methods in Perceptual Audio Coding: A Tutorial introduction, 17th AES International Conference on High Quality Audio Coding, August 1999

[0221] [15] International Standard ISO/IEC 23001-3:2010, Unified speech and audio coding Audio, Geneva, 2010.

[0222] [16] International Standard ISO/IEC 14496-3:2005, Information technology - Coding of audio-visual objects - Part 3: Audio, Geneva, 2005.

[0223] [17] P. Ekstrand, Bandwidth Extension of Audio Signals by Spectral Band Replication, in Proceedings of 1st IEEE Benelux Workshop on MPCA, Leuven, November 2002

[0224] [18] F. Nagel, S. Disch, S. Wilde, A continuous modulated single sideband bandwidth extension, ICASSP

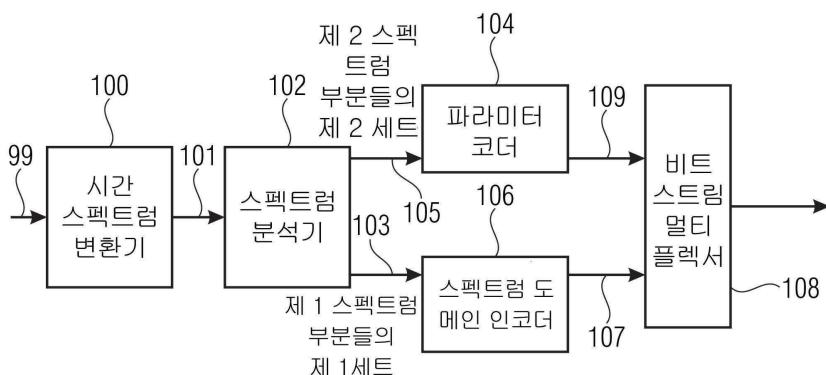
International Conference on Acoustics, Speech and Signal Processing, Dallas, Texas (USA), April 2010

[0225] [19] Liljeryd, Lars; Ekstrand, Per; Henn, Fredrik; Kjorling, Kristofer: Spectral translation/folding in the subband domain, United States Patent 8,412,365, April 2, 2013.

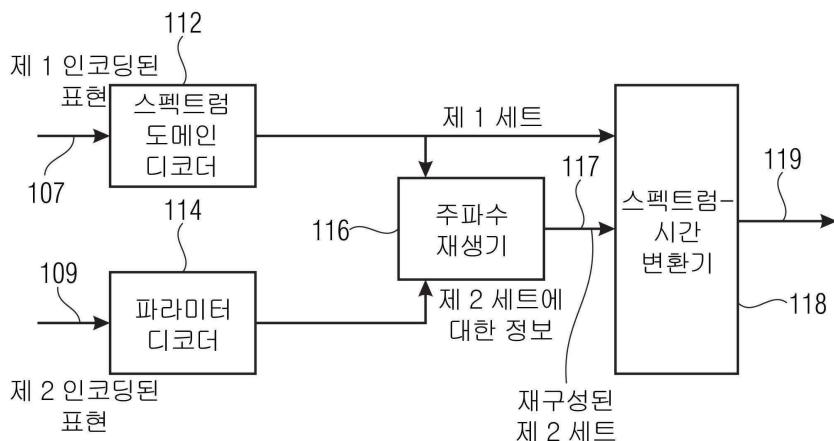
[0226] [20] Daudet, L.; Sandler, M.; "MDCT analysis of sinusoids: exact results and applications to coding artifacts reduction," Speech and Audio Processing, IEEE Transactions on, vol.12, no.3, pp. 302- 312, May 2004.

도면

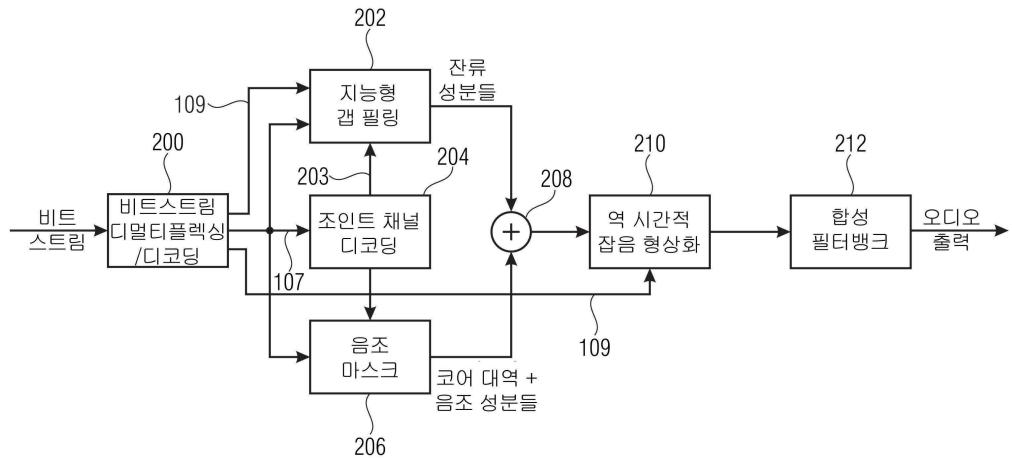
도면1a



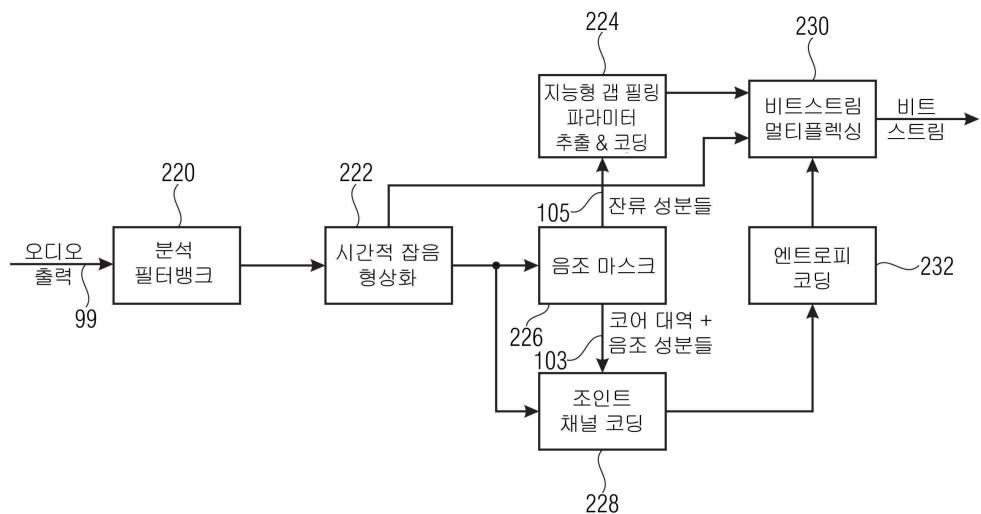
도면1b



도면2a

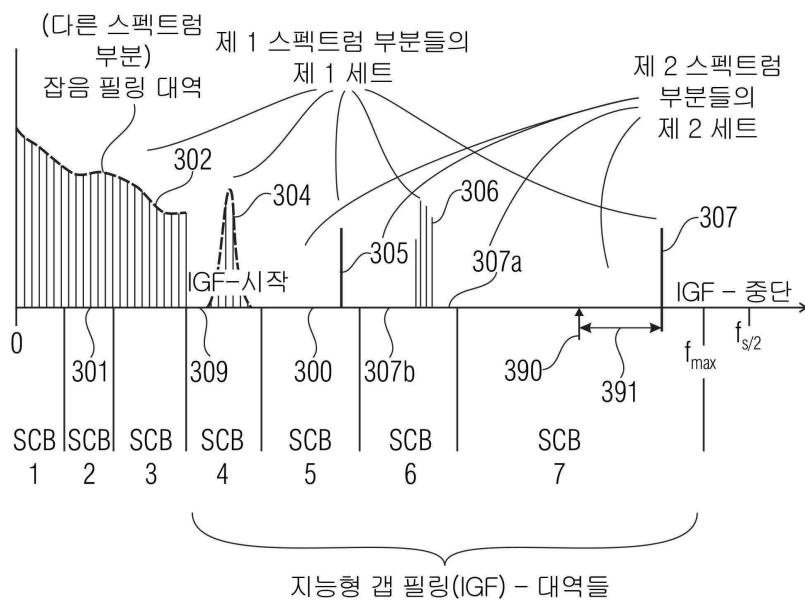


도면2b



도면3a

- 제 1 세트의 “엔벨로프:”를 위한
제 1 해상도(고해상도) (라인-방식코딩);
- 제 2세트의 “엔벨로프”를 위한
제 1 해상도(저해상도) (SCB 당 스케일 인자);

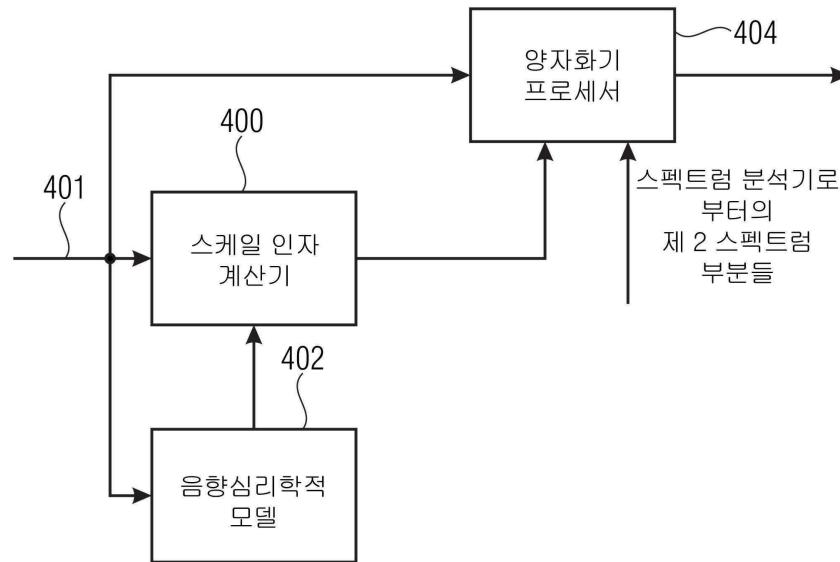


도면3b

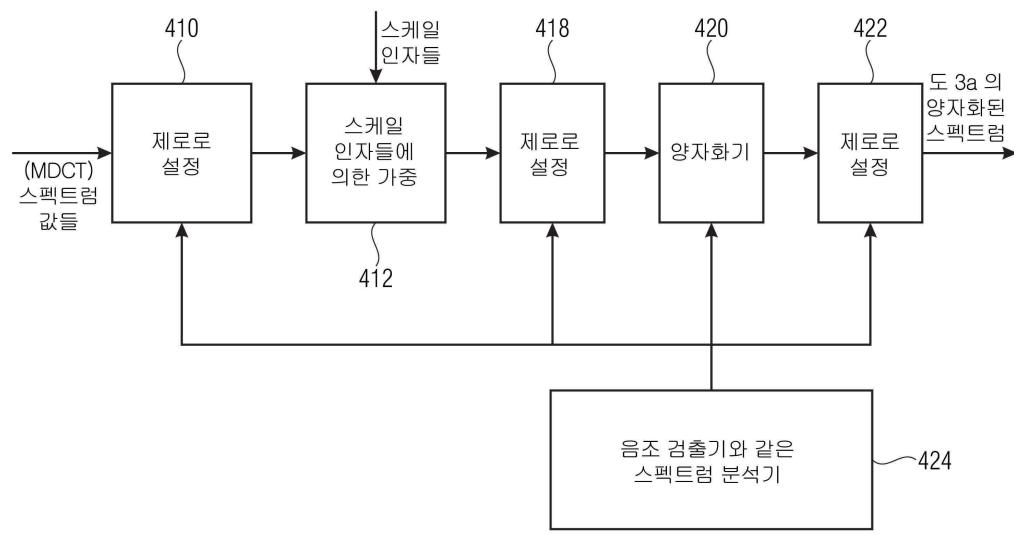
SCB1	SCB2	SCB3	SCB4	SCB5	SCB6	SCB7
SF1	SF2	SF3	SF4	SF5	SF6	SF7
			E_1	E_2	E_3	E_4
	NF_2					

Below the table, brackets indicate groupings: 308 (SCB1, SCB2, SCB3), 310 (SCB4, SCB5), and 312 (SCB6, SCB7).

도면4a

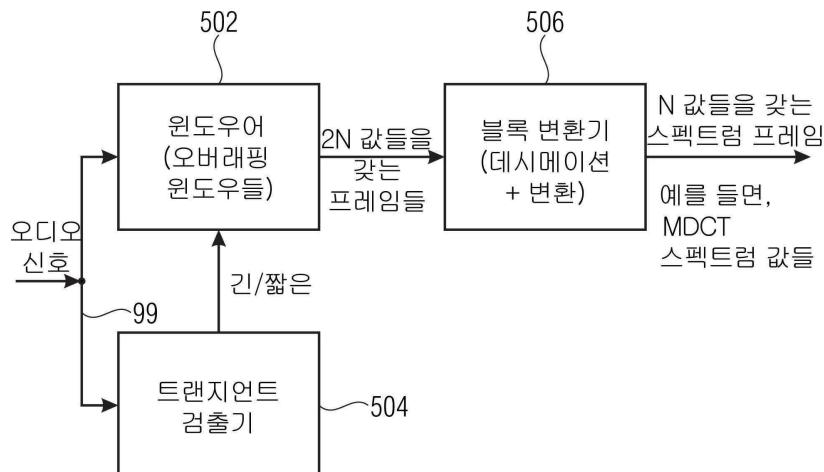


도면4b



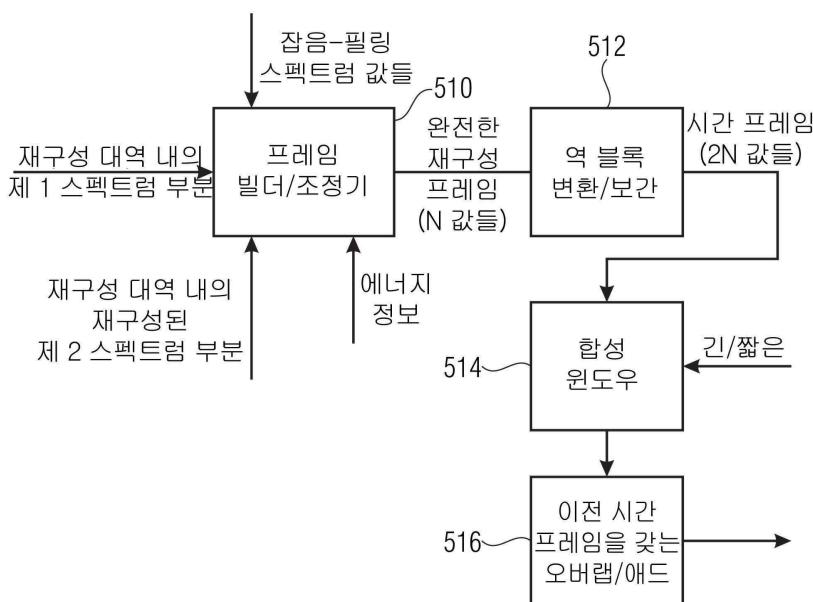
(양자화기 프로세서)

도면5a

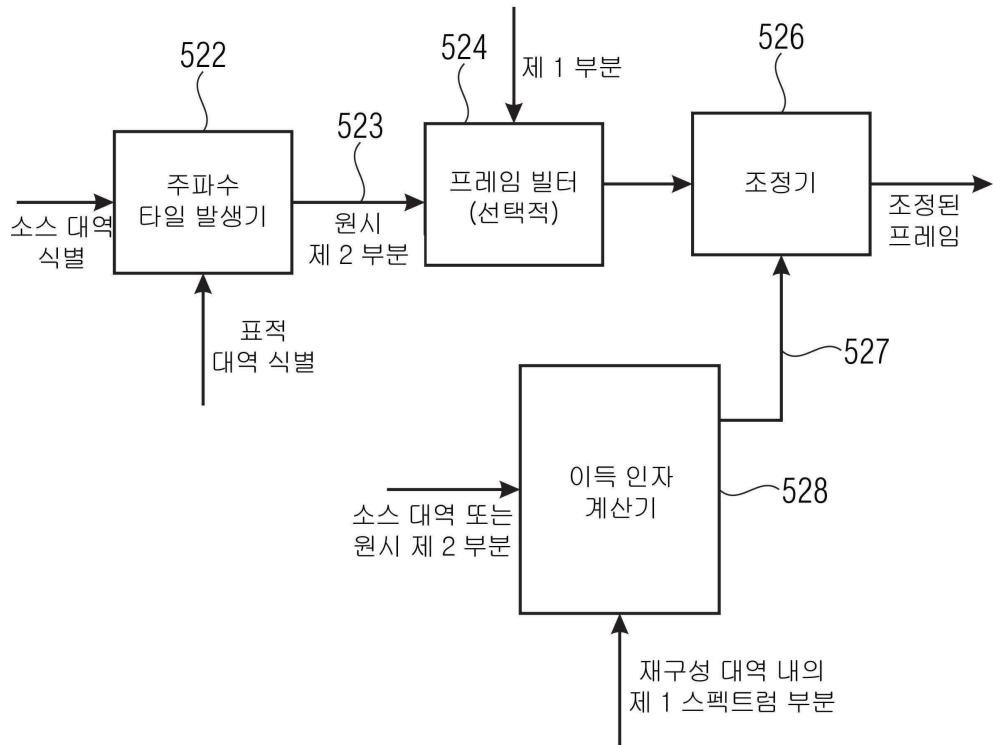


(다른 스펙트럼 부분들)

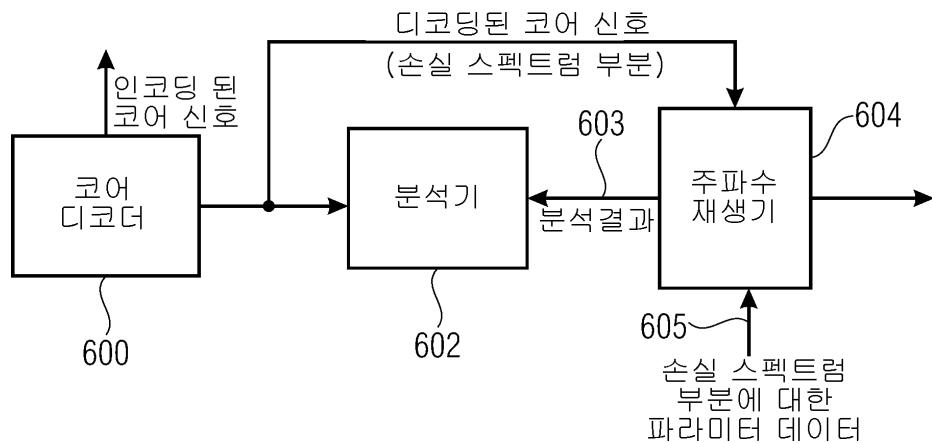
도면5b



도면5c



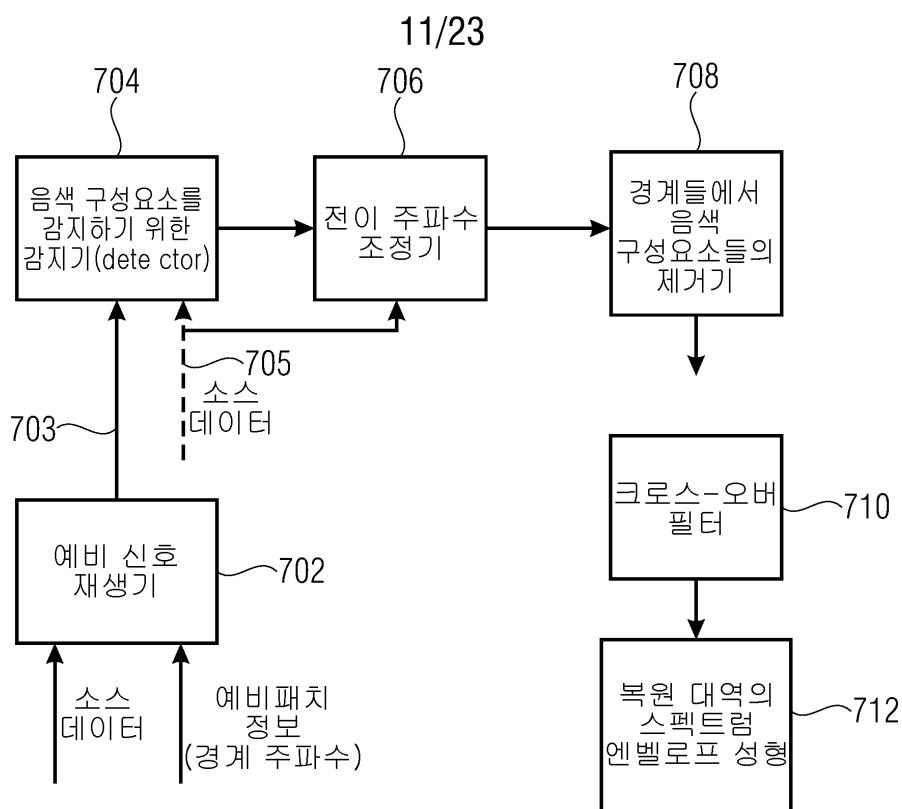
도면6a



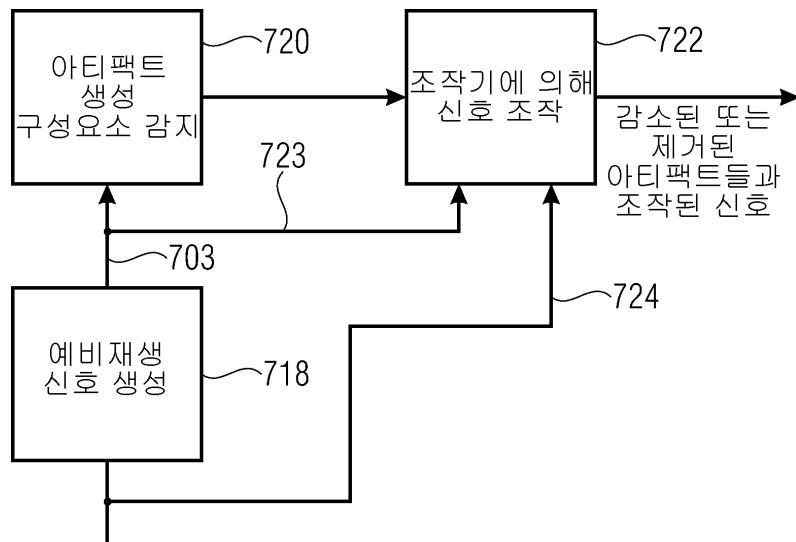
도면6b



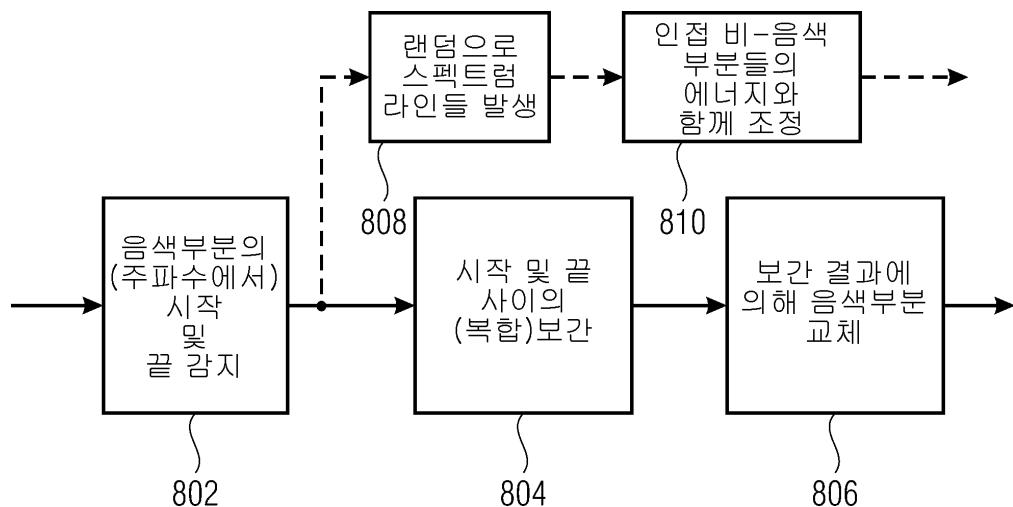
도면7a



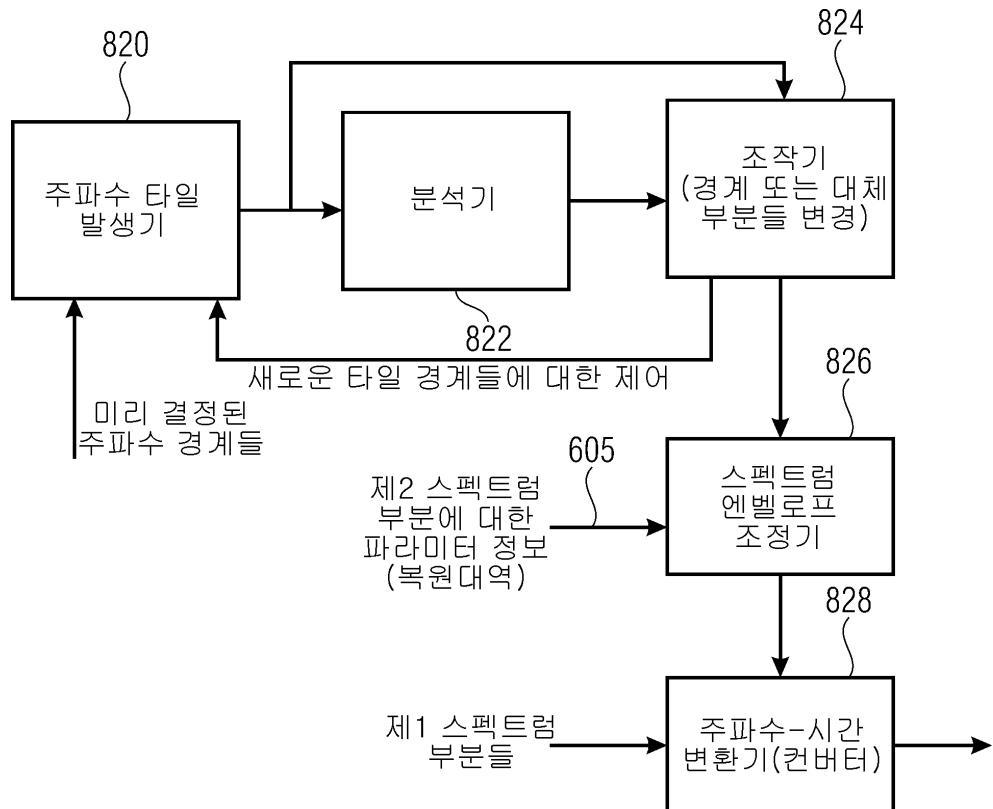
도면7b



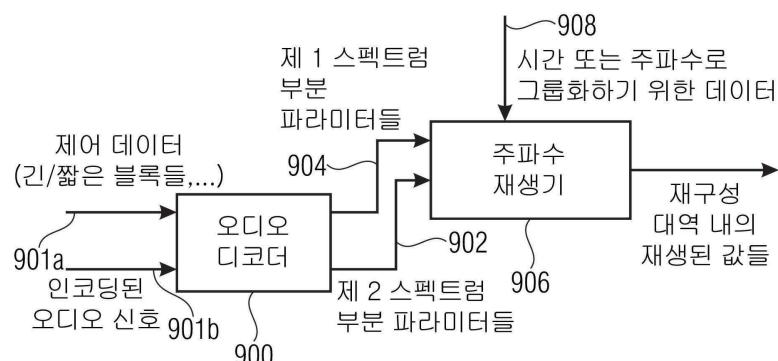
도면8a



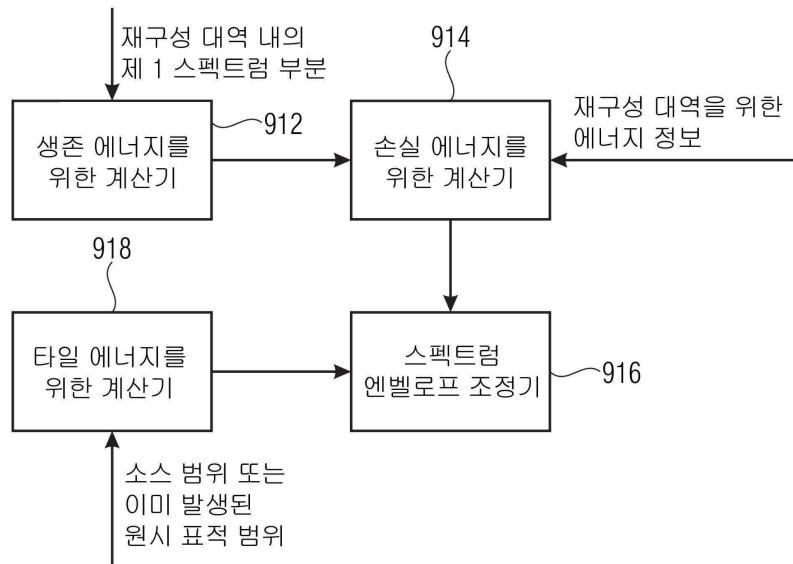
도면8b



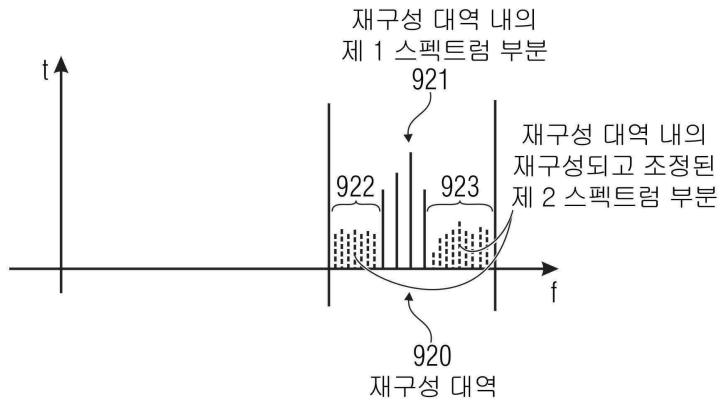
도면9a



도면9b



도면9c



e.g. • 생존 에너지

5 유닛

- 재구성 대역을 위한 에너지 값

10 유닛 (재구성 대역
내의 제 1 및
제 2 스펙트럼
대역들을 포함)

- 소스 범위 데이터 또는
원시 표적 범위 데이터의 에너지

8 유닛

- 손실 에너지

5 유닛

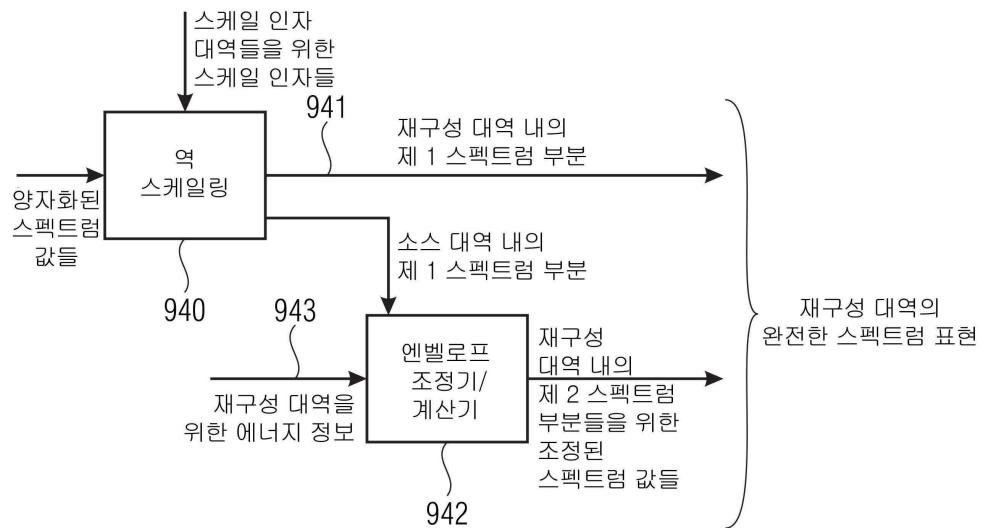
- 이득 인자

$$g := \sqrt{\frac{mE_k}{pE_k}} = 0.79$$

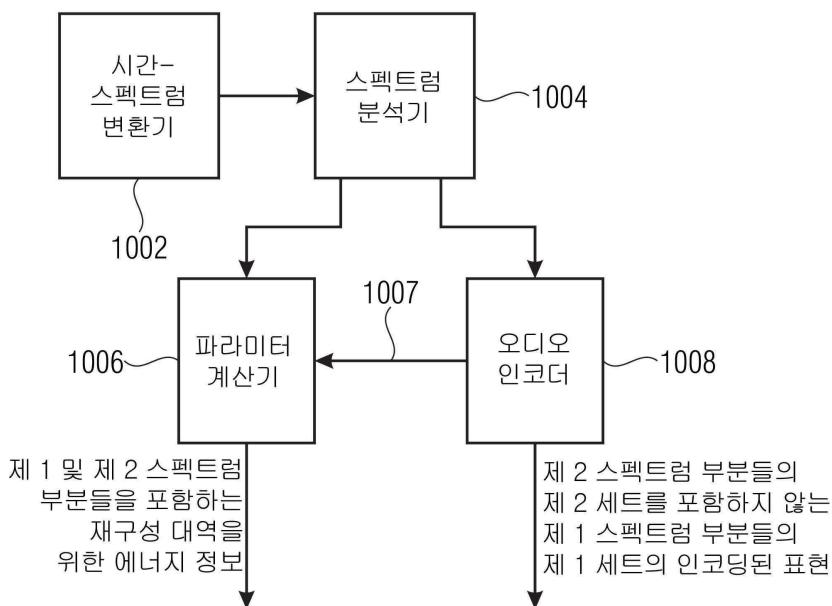
→ 제 2 스펙트럼 부분들을 위한 스펙트럼 값들만이 조정됨

→ 제 1 스펙트럼 부분들은 엔벨로프 조정에 의해 영향을 받지 않음

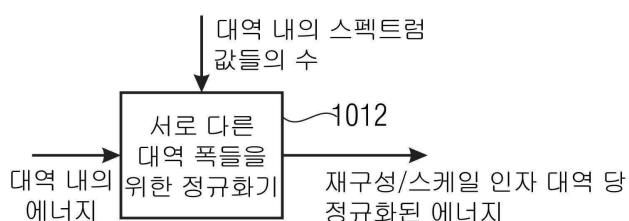
도면9d



도면10a

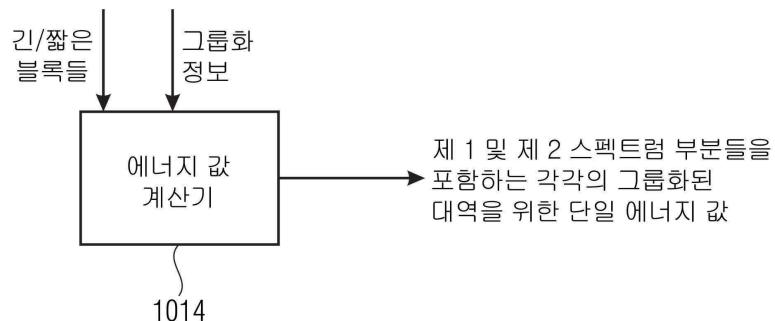


도면10b

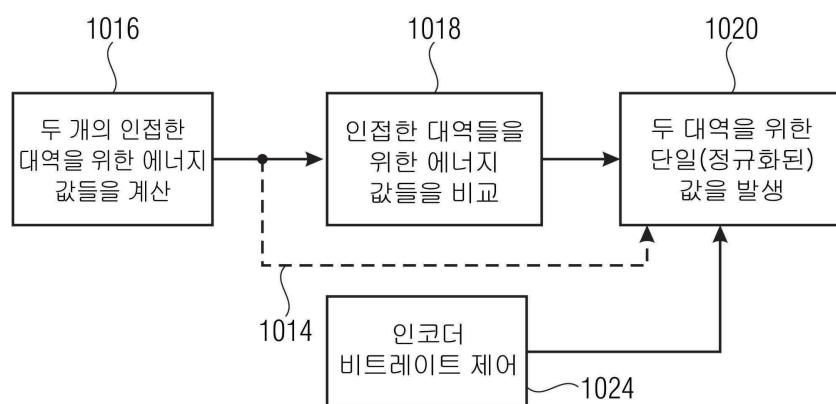


도면10c

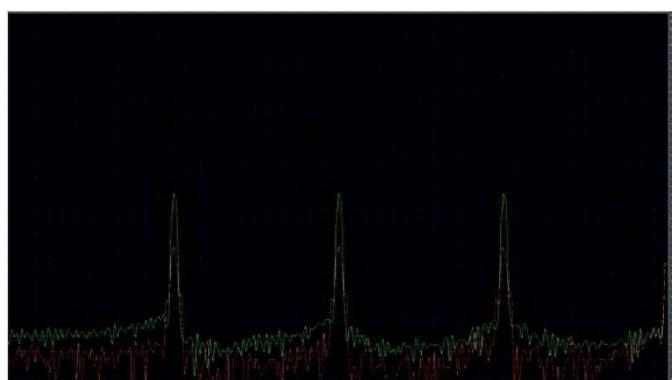
코더 인코더로부터의 제어 정보



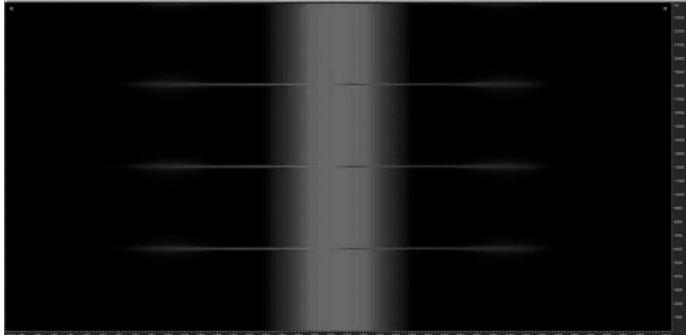
도면10d



도면11a

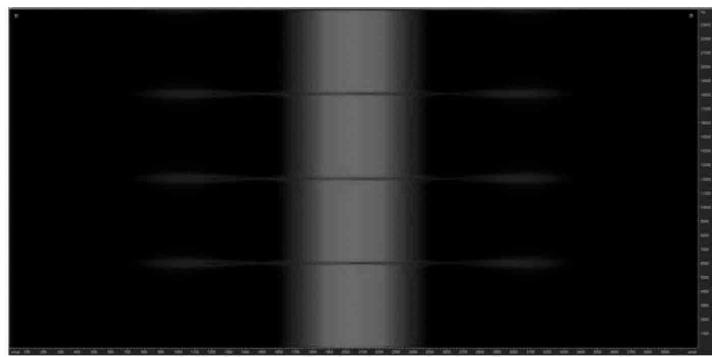


도면11b



BWE 적용 후 트랜지언트의 스펙트로 그램
X축은 시간, Y축은 주파수를 표현

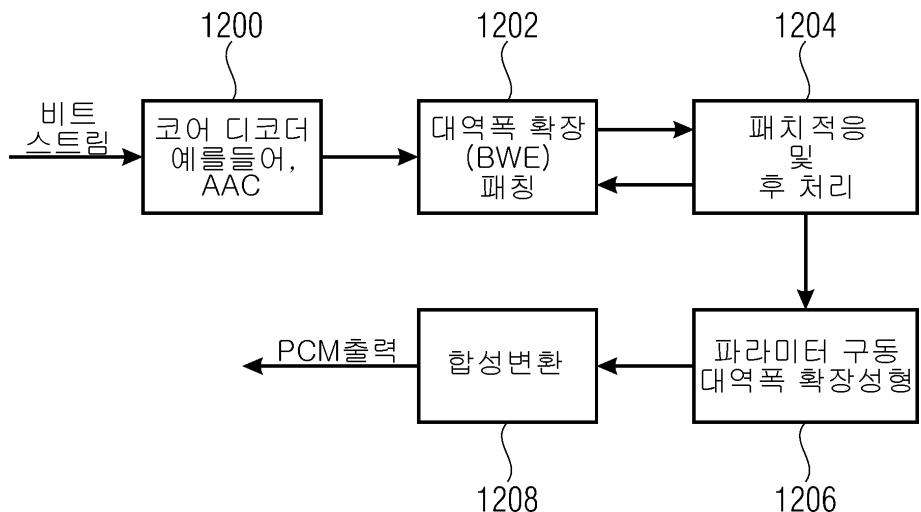
도면11c



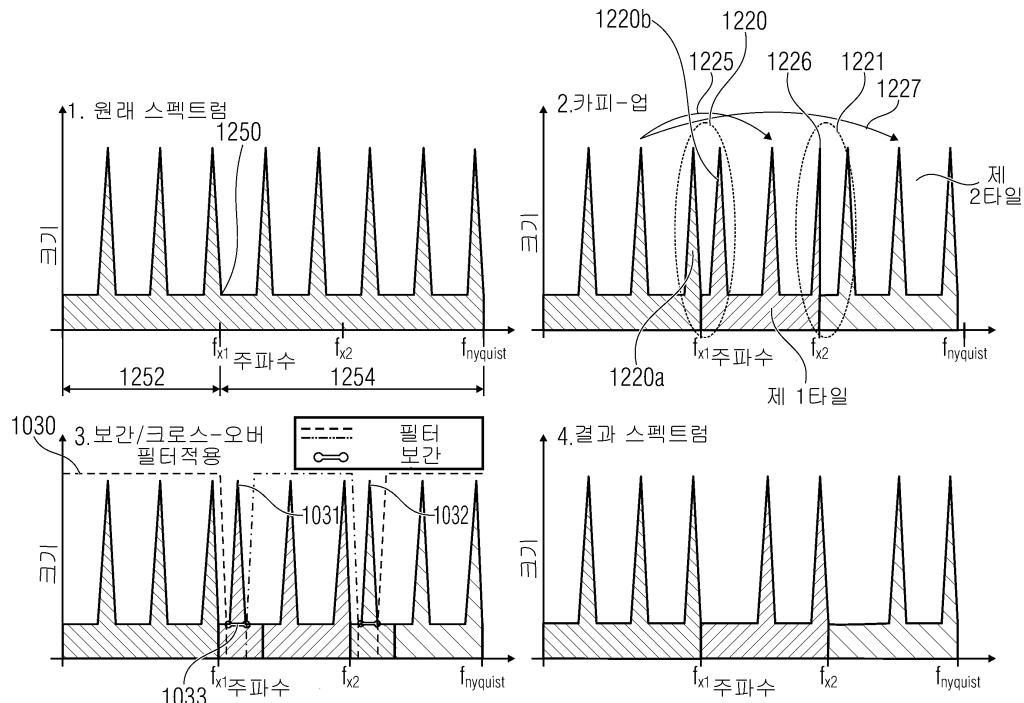
BWE 적용후 트랜지언트의 스펙트로 그램
X축은 시간, Y축은 주파수를 표현.

필터 링잉 감소의 적용을 통해 필터 링잉은 근사적으로 20dB가 감소됨

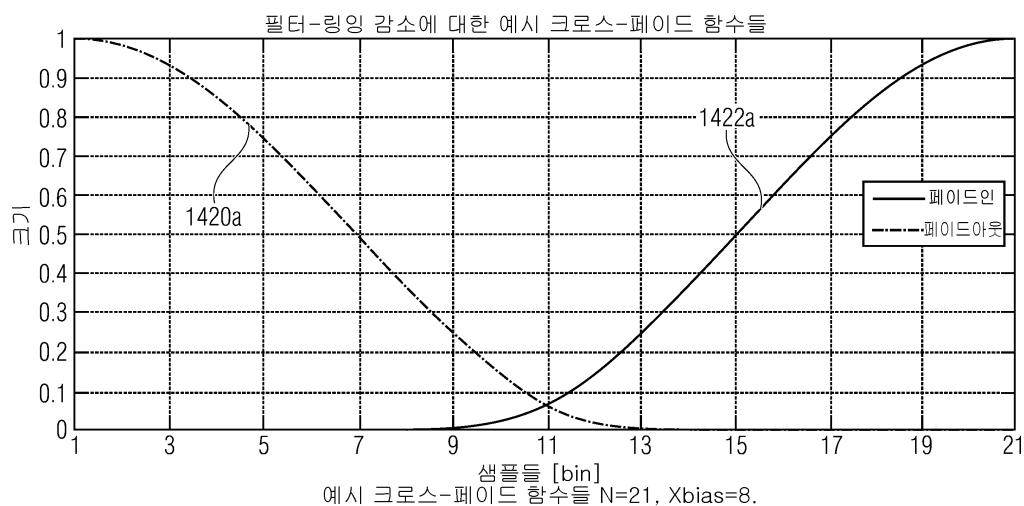
도면12a



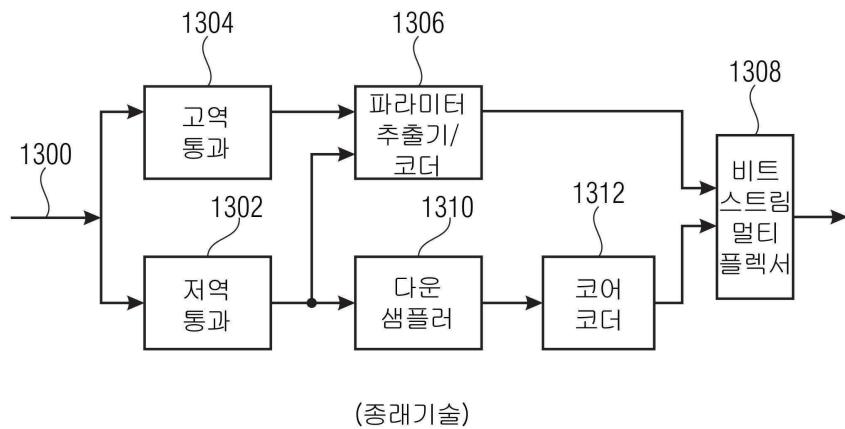
도면12b



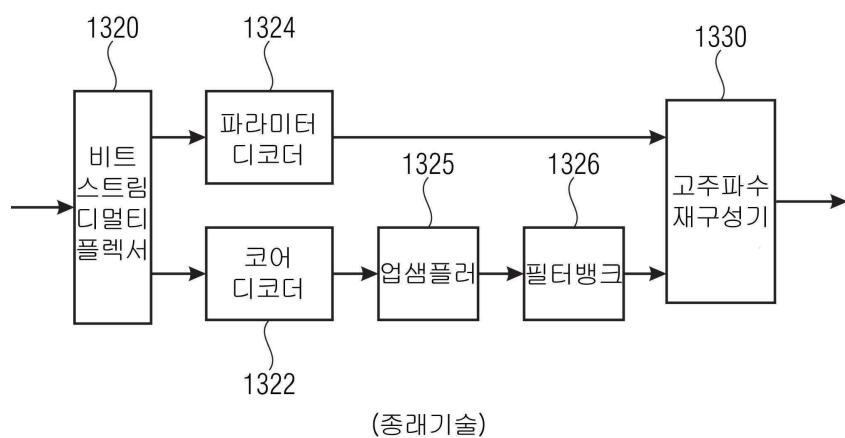
도면12c



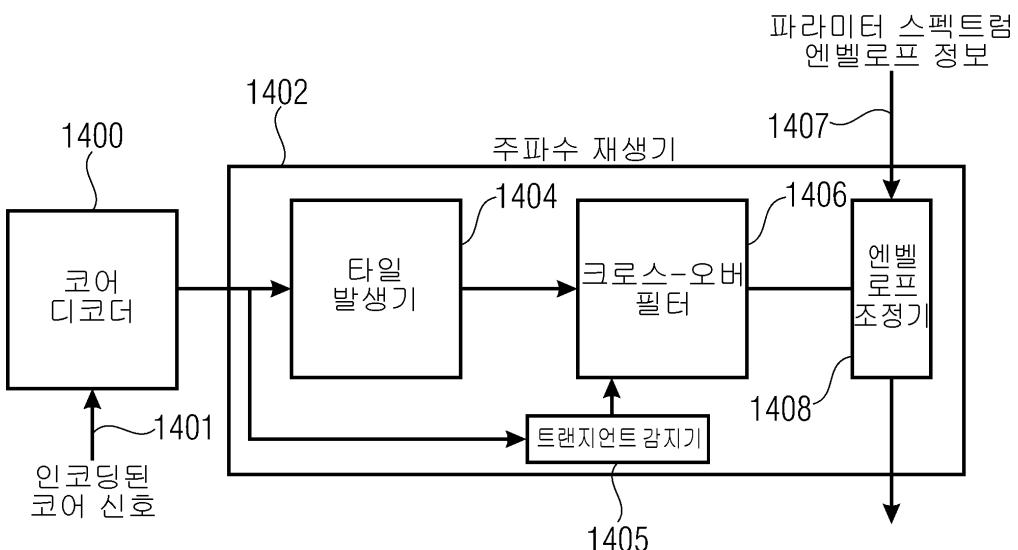
도면13a



도면13b



도면14a



도면14b

