



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년07월14일
 (11) 등록번호 10-1757341
 (24) 등록일자 2017년07월06일

- (51) 국제특허분류(Int. Cl.)
G10L 19/035 (2013.01) *G10L 19/02* (2006.01)
G10L 25/03 (2013.01)
- (52) CPC특허분류
G10L 19/035 (2013.01)
G10L 19/0212 (2013.01)
- (21) 출원번호 10-2015-7022139
- (22) 출원일자(국제) 2014년01월28일
 심사청구일자 2015년08월17일
- (85) 번역문제출일자 2015년08월17일
- (65) 공개번호 10-2015-0118954
- (43) 공개일자 2015년10월23일
- (86) 국제출원번호 PCT/EP2014/051624
- (87) 국제공개번호 WO 2014/118171
 국제공개일자 2014년08월07일
- (30) 우선권주장
 61/758,191 2013년01월29일 미국(US)
- (56) 선행기술조사문헌
 Laurent Daudet, et al. MDCT analysis of sinusoids: exact results and applications to coding artifacts reduction. IEEE Transactions on Speech and Audio Processing, 2004.
 Marie Oger, et al. Model-based deadzone optimization for stack-run audio coding with uniform scalar quantization. IEEE International Conference on Acoustics, Speech and Signal Processing. 2008.
 US20080240235 A1

- (73) 특허권자
 프라운호퍼 게젤샤프트 쥘르 뫼르데롱 데어 안겐
 반텐 포르슈 에.베.
 독일 80686 뮌헨 한자슈트라쎄 27 체
- (72) 발명자
 디에츠, 마틴
 독일 90429 뉘른베르크 도이치헤른슈트라쎄 37
 푸호스, 구일라우메
 독일 91058 에를랑겐 뫼르트 슈트라쎄 17
 (뒷면에 계속)
- (74) 대리인
 윤의섭, 김수진

전체 청구항 수 : 총 17 항

심사관 : 이남숙

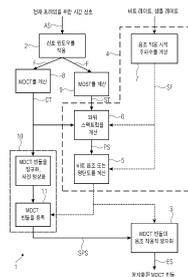
(54) 발명의 명칭 **저-복잡도 음조-적응 오디오 신호 양자화**

(57) 요약

본 발명은 그것으로부터 인코딩된 신호(ES)를 생산하기 위하여 오디오 신호(AS)를 인코딩하기 위한 오디오 인코더(1)를 제공하는데, 오디오 인코더(1)는: 오디오 신호(AS)로부터 프레임들(F)을 추출하도록 구성되는 프레임 장치(2); 오디오 신호(AS)의 프레임(F)으로부터 유도된 스펙트럼 신호(SPS)의 스펙트럼 라인들(SL₁₋₃₂, 도 2

(뒷면에 계속)

대표도



참조)을 양자화 지수들(I_0, I_1)에 매핑하도록 구성되는 양자화기(3)를 포함하되, 상기 양자화기(3)는 스펙트럼 라인들(SL_{1-32})이 양자화 지수 제로(I_0)에 매핑되는, 데드 존(DZ, 도 2 참조)을 가짐; 및 데드 존(DZ)을 변형하도록 구성되는 제어 장치(4);를 포함하고, 제어 장치(4)는 적어도 하나의 스펙트럼 라인(SL_{1-32}) 또는 적어도 하나의 스펙트럼 라인들(SL_{1-32})의 그룹을 위한 적어도 하나의 음조 표시 값(TI_{5-32})을 계산하도록 구성되는 음조 계산 장치(5)를 포함하고, 제어 장치(4)는 음조 표시 값(TI_{5-32})에 의존하여 적어도 하나의 스펙트럼 라인(SL_{1-32}) 또는 적어도 하나의 스펙트럼 라인들(SL_{1-32})의 그룹을 위하여 데드 존(DZ)을 변형하도록 구성된다.

(52) CPC특허분류

G10L 25/03 (2013.01)

(72) 발명자

험리히, 크리스티안

독일 91054 에를랑겐 하우프트슈트라쎄 68

마르코비치, 고란

독일 90425 뉘른베르크 아헤너 슈트라쎄 19

명세서

청구범위

청구항 1

인코딩된 신호(ES)를 생산하기 위하여 오디오 신호(AS)를 인코딩하기 위한 오디오 인코더(1)에 있어서,
 상기 오디오 신호(AS)로부터 프레임들(F)을 추출하도록 구성되는 프레임링 장치(2);
 상기 오디오 신호(AS)의 상기 프레임(F)으로부터 유도된 스펙트럼 신호(SPS)의 스펙트럼 라인들(SL₁₋₃₂)을 양자화 지수들(I₀, I₁)에 매핑하도록 구성되는 양자화기(3)를 포함하되, 상기 양자화기(3)는 상기 스펙트럼 라인들(SL₁₋₃₂)이 양자화 지수 제로(I₀)에 매핑되는, 데드 존(DZ)을 가짐; 및
 상기 데드 존(DZ)을 변형하도록 구성되는 제어 장치(4);를 포함하고,
 상기 제어 장치(4)는 적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위한 적어도 하나의 음조 표시 값(TI₅₋₃₂)을 계산하도록 구성되는 음조 계산 장치(5)를 포함하고,
 상기 제어 장치(4)는 각각의 음조 표시 값(TI₅₋₃₂)에 의존하여 상기 적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 상기 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위하여 상기 데드 존(DZ)을 변형하도록 구성되는 것을 특징으로 하는 오디오 인코더.

청구항 2

제 1항에 있어서, 상기 제어 장치(4)는 상기 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나에서의 상기 데드 존(DZ)이 더 큰 음조를 갖는 상기 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나에서의 상기 데드 존(DZ)보다 크도록 또는 상기 스펙트럼 라인들(SL₁₋₃₂)의 그룹들 중 어느 하나에서의 데드 존(DZ)이 더 큰 음조를 갖는 상기 스펙트럼 라인들(SL₁₋₃₂)의 그룹들 중 어느 하나에서의 데드 존(DZ)보다 크도록 상기 데드 존(DZ)을 변형하도록 구성되는 것을 특징으로 하는 오디오 인코더.

청구항 3

제 1항에 있어서, 상기 제어 장치(4)는 상기 오디오 신호(AS)의 상기 프레임(F)의 파워 스펙트럼(PS)을 계산하도록 구성되는 파워 스펙트럼 계산 장치(6)를 포함하고, 상기 파워 스펙트럼(PS)은 스펙트럼 라인들(SL₁₋₃₂) 또는 스펙트럼 라인들(SL₁₋₃₂)의 그룹들을 위한 파워 값들(PS₅₋₃₂)을 포함하며, 상기 음조 계산 장치(5)는 상기 파워 스펙트럼(PS)에 의존하여 상기 적어도 하나의 음조 표시 값(TI₅₋₃₂)을 계산하도록 구성되는 것을 특징으로 하는 오디오 인코더.

청구항 4

제 3항에 있어서, 상기 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나를 위한 상기 음조 표시 값(TI₅₋₃₂)은 각각의 스펙트럼 라인들(SL₁₋₃₂)을 위한 상기 파워 값(PS₅₋₃₂) 및 미리 정의된 수의 상기 파워 스펙트럼(PS)의 그 주위의 파워 값들(PS₅₋₃₂)의 합의 비교를 기초로 하거나, 또는 상기 스펙트럼 라인들의 그룹들(SL₁₋₃₂) 중 어느 하나를 위한 상기 음조 표시 값(TI₅₋₃₂)은 상기 각각의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위한 상기 파워 값(PS₅₋₃₂) 및 미리 정의된 수의 상기 파워 스펙트럼(PS)의 그 주위의 파워 값들(PS₅₋₃₂)의 합의 비교를 기초로 하는 것을 특징으로 하

는 오디오 인코더.

청구항 5

제 1항에 있어서, 상기 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나를 위한 상기 음조 표시 값(TI₅₋₃₂)은 상기 오디오 신호(AS)의 선행 프레임(F)의 상기 스펙트럼 라인(SL₁₋₃₂)의 상기 음조 표시 값(TI₅₋₃₂)을 기초로 하거나, 혹은 상기 스펙트럼 라인들(SL₁₋₃₂)의 그룹들 중 어느 하나를 위한 상기 음조 표시 값(TI₅₋₃₂)은 상기 오디오 신호(AS)의 선행 프레임(F)을 위한 상기 스펙트럼 라인들(SL₁₋₃₂)의 그룹의 상기 음조 표시 값(TI₅₋₃₂)을 기초로 하는 것을 특징으로 하는 오디오 인코더.

청구항 6

제 3항에 있어서, 상기 음조 표시 값(TI₅₋₃₂)은 아래의 공식에 의해 계산되는데:

$$T_{k,i} = f\left(\frac{P_{k-7,i} + \dots + P_{k-1,i} + P_{k+1,i} + \dots + P_{k+7,i}}{P_{k,i}}, \frac{P_{k-7,i-1} + \dots + P_{k-1,i-1} + P_{k+1,i-1} + \dots + P_{k+7,i-1}}{P_{k,i-1}}\right),$$

여기서 i 는 상기 오디오 신호(AS)의 특정 프레임(F)을 나타내는 지수이고, k 는 특정 스펙트럼 라인(SL₁₋₃₂)을 나타내는 지수이며, $T_{k,i}$ 는 i 번째 프레임(F)의 상기 k 번째 스펙트럼 라인(SL₁₋₃₂)의 음조 표시 값(TI₅₋₃₂)이며, $P_{k,i}$ 는 상기 i 번째 프레임(F)의 상기 k 번째 스펙트럼 라인(SL₁₋₃₂)의 파워 값(PS₅₋₃₂)이며, 혹은 상기 음조 표시 값(TI₅₋₃₂)은 아래의 공식에 의해 계산되며:

$$T_{m,i} = f\left(\frac{P_{m-7,i} + \dots + P_{m-1,i} + P_{m+1,i} + \dots + P_{m+7,i}}{P_{m,i}}, \frac{P_{m-7,i-1} + \dots + P_{m-1,i-1} + P_{m+1,i-1} + \dots + P_{m+7,i-1}}{P_{m,i-1}}\right),$$

여기서 i 는 상기 오디오 신호(AS)의 특정 프레임(F)을 나타내는 지수이고, m 은 특정 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 나타내는 지수이며, $P_{m,i}$ 는 상기 i 번째 프레임의 상기 m 번째 스펙트럼 라인들(SL₁₋₃₂)의 그룹의 파워 값인 것을 특징으로 하는 오디오 인코더.

청구항 7

제 1항에 있어서, 상기 오디오 인코더(1)는 상기 데드 존(DZ)을 변경하기 위하여 시작 주파수(SF)를 계산하도록 구성되는 시작 주파수 계산 장치(7)를 포함하고, 상기 데드 존(DZ)은 상기 시작 주파수(SF)보다 높거나 또는 동일한 주파수를 표현하는 스펙트럼 라인들(SL₁₋₃₂)만을 위하여 변형되는 것을 특징으로 하는 오디오 인코더.

청구항 8

제 7항에 있어서, 상기 시작 주파수 계산 장치(7)는 상기 오디오 신호(AS)의 샘플 레이트를 기초로 하고 그리고 /또는 상기 인코딩된 신호(ES)로부터 생산된 비트스트림을 위하여 예상되는 최대 비트-레이트를 기초로 하여 상기 시작 주파수(SF)를 계산하도록 구성되는 것을 특징으로 하는 오디오 인코더.

청구항 9

제 3항에 있어서, 상기 오디오 인코더(1)는 상기 오디오 신호(AS)의 상기 프레임(F)으로부터 변형 이산 코사인 변환(CT)을 계산하도록 구성되는 변형 이산 코사인 변환 계산 장치(8) 및 상기 오디오 신호(AS)의 상기 프레임(F)으로부터 변형 이산 사인 변환(ST)을 계산하도록 구성되는 변형 이산 사인 변환 계산 장치(9)를 포함하고, 상기 파워 스펙트럼 계산 장치(6)는 상기 변형 이산 코사인 변환(CT) 및 상기 변형 이산 사인 변환(ST)을 기초

로 하여 상기 파워 스펙트럼(PS)을 계산하도록 구성되는 것을 특징으로 하는 오디오 인코더.

청구항 10

제 9항에 있어서, 상기 파워 스펙트럼 계산 장치(6)는 공식 $P_{k,i} = (MDCT_{k,i})^2 + (MDST_{k,i})^2$ 에 따라 상기 파워 값들(PS₅₋₃₂)을 계산하도록 구성되는데, 여기서 i 는 상기 오디오 신호(AS)의 특정 프레임(F)을 나타내는 지수이고, k 는 특정 스펙트럼 라인을 나타내는 지수이며, $MDCT_{k,i}$ 는 i 번째 프레임의 상기 k 번째 스펙트럼 라인에서의 상기 변형 이산 코사인 변환(CT)의 값이며, $MDST_{k,i}$ 는 상기 i 번째 프레임의 상기 k 번째 스펙트럼 라인에서의 상기 변형 이산 사인 변환(ST)의 값이며, $P_{k,i}$ 는 상기 i 번째 프레임의 상기 k 번째 스펙트럼 라인의 상기 파워 값(PS₅₋₃₂)인 것을 특징으로 하는 오디오 인코더.

청구항 11

제 9항에 있어서, 상기 오디오 인코더(1)는 상기 스펙트럼 신호(SPS)를 생산하도록 구성되는 스펙트럼 신호 계산 장치(10)를 포함하고, 상기 스펙트럼 신호 계산 장치(10)는 상기 데드 존(DZ)의 변형에 기인한 에너지 손실이 보상되도록 상기 스펙트럼 신호(SPS)의 상기 스펙트럼 라인들(SL₁₋₃₂)의 진폭들을 설정하도록 구성되는 진폭 설정 장치(11)를 포함하는 것을 특징으로 하는 오디오 인코더.

청구항 12

제 11항에 있어서, 상기 진폭 설정 장치(11)는 상기 각각의 스펙트럼 라인(SL₁₋₃₂)에서 상기 데드 존(DZ)의 변형에 의존하여 상기 스펙트럼 신호(SPS)의 상기 진폭들을 설정하도록 구성되는 것을 특징으로 하는 오디오 인코더.

청구항 13

제 11항에 있어서, 상기 스펙트럼 신호 계산 장치(10)는 정규화 장치(12)를 포함하는 것을 특징으로 하는 오디오 인코더.

청구항 14

제 11항에 있어서, 상기 변형 이산 코사인 변환 계산 장치(8)에 의해 계산된 상기 오디오 신호(AS)의 상기 프레임(F)으로부터의 상기 변형 이산 코사인 변환(CT)은 상기 스펙트럼 신호 계산 장치(10)에 제공되는 것을 특징으로 하는 오디오 인코더.

청구항 15

인코더(1) 및 디코더를 포함하는 시스템에 있어서, 상기 인코더(1)는 제 1항 내지 14항 중 어느 한 항에 따라 구성되는 것을 특징으로 하는 시스템.

청구항 16

인코딩된 신호(ES)를 생산하기 위하여 오디오 신호(AS)를 인코딩하기 위한 방법에 있어서, 상기 오디오 신호(AS)로부터 프레임들(F)을 추출하는 단계;

상기 오디오 신호(AS)의 상기 프레임(F)으로부터 유도된 스펙트럼 신호(SPS)의 스펙트럼 라인들(SL₁₋₃₂)을 양자화 지수들(I₀, I₁)에 매핑하는 단계 - 여기서 입력 스펙트럼 라인들(SL₁₋₃₂)이 양자화 지수 제로(I₀)에 매핑되는 데드 존(DZ)이 사용됨 -; 및

상기 데드 존(DZ)을 변형하는 단계;를 포함하고,

적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위한 적어도 하나의 음조 표시 값(TI₅₋₃₂)이 계산되고,

상기 각각의 음조 표시 값(TI₅₋₃₂)에 의존하여 상기 적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 상기 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위하여 상기 데드 존(DZ)이 변형되는 것을 특징으로 하는 오디오 신호(AS)를 인코딩하기 위한 방법.

청구항 17

컴퓨터 또는 프로세서 상에서 구동할 때, 제 16항의 방법을 실행하기 위한 컴퓨터 프로그램을 저장한 컴퓨터 판독가능 저장 매체.

발명의 설명

기술 분야

[0001] 본 발명은 디지털 오디오 신호 처리에 관한 것이다. 더 구체적으로 본 발명은 오디오 신호 양자화에 관한 것이다.

배경 기술

[0002] 매우 낮은 비트레이트의 변환 코딩에 있어서, 프레임 당 비트들의 수는 일반적으로 디코딩된 신호 내의 아티팩트(artifact)들을 방지하는데 충분하지 않다. 음악 소음(musical noise)은 특히 하나의 프레임으로부터 그 다음 프레임으로의 특정 주파수에서 "켜지고 꺼지는", 즉 제로로 양자화되거나 또는 제로로 양자화되지 않은 변환 라인들(빈(bin)들)에 기인하여 정상 음악 또는 소음 스펙트럼에 나타날 수 있다. 그러한 코딩 접근법은 디코딩된 신호 영역에 원래 신호가 갖는 것보다 더 많은 음조 특성(따라서 용어 음악 소음)을 제공하지 못할 뿐만 아니라, 이는 또한 상기 스펙트럼 영역에 대한 중요한 장점을 전혀 생산하지 않고 대신에 확장 고효율-고급 오디오 코딩(xHE-AAC) [4]에서 사용되는 변환 코딩 여기(transform coded excitation, TCX) 또는 FD 코딩 시스템들에서의 소음 추가 알고리즘들 같은 빈-대체 기술을 적용한다. 실제로, 음속 코딩 소음의 경향이 있는 영역들의 분명하나 불충분한 코딩은 다른 스펙트럼 영역들, 특히 인간 청각 시스템이 민감한 낮은 주파수들에서 음속으로 더 잘 소비되는, 변환 코더의 엔트로피 코딩 단계 내의 비트들을 필요로 한다.

[0003] 낮은 비트 레이트 오디오 코딩에서 음악 소음의 발생을 감소시키는 한 가지 방법은 즉각적인 입력 신호 특성 및 양자화된 스펙트럼의 비트 소비에 적응하도록 하기 위하여 입력 스펙트럼 라인들을 양자화 지수들에 매핑하는 양자화기의 행동을 변형하는 것이다. 더 정확하게는, 양자화 동안에 사용된 데드 존(dead-zone)은 신호 적응적으로 변경된다. 일부 접근법들이 공개되었다[5, 6, 및 그 안의 참고문헌들]. [5]에서, 양자화 적응은 코딩되려는 전체 스펙트럼 상에서 실행된다. 따라서 적응된 양자화기는 주어진 프레임의 모든 스펙트럼 빈에 대하여 동일하게 행동한다. 게다가, 최적 데드 존(z_{opt})을 갖는 양자화의 경우에 있어서, 비트 레이트 및 백워드 호환성 패널티(backward compatibility penalty)를 표현하는, 2 비트의 부가 정보가 디코더에 전송되어야만 한다. [6]에서, 양자화기는 주파수 대역 당 기반 상에 적용되나, 두 가지 대역 당 두 가지 양자화 시도가 수행되며, 더 나은 시도(특정 결정에 따른)만이 전송을 위하여 사용된다. 이는 복잡하다.

발명의 내용

해결하려는 과제

[0004] 본 발명의 목적은 오디오 신호 처리를 위한 향상된 개념들을 제공하는 것이다. 더 구체적으로, 본 발명은 적응

오디오 신호 양자화를 위한 향상된 개념들을 제공하는 것이다. 본 발명의 목적은 청구항 1항에 따른 오디오 인코더, 청구항 15항에 따른 시스템, 청구항 16항에 따른 방법 및 청구항 17항에 따른 컴퓨터 프로그램에 의해 달성된다.

과제의 해결 수단

- [0005] 일 양상에서 본 발명은 그것으로부터 인코딩된 신호를 생산하기 위하여 오디오 신호의 인코딩을 위한 오디오 인코더를 제공하는데, 오디오 인코더는:
- [0006] 오디오 신호로부터 프레임들을 추출하도록 구성되는 프레임링 장치(framing device);
- [0007] 오디오 신호의 프레임으로부터 유도된 스펙트럼 라인들을 양자화 지수들에 매핑하도록 구성되는 양자화기를 포함하되, 상기 양자화기는 스펙트럼 라인들이 양자화 지수 제로에 매핑되는, 데드 존을 가짐; 및
- [0008] 데드 존을 변형하도록 구성되는 제어 장치;를 포함하고,
- [0009] 제어 장치는 적어도 하나의 스펙트럼 라인 또는 적어도 하나의 스펙트럼 라인들의 그룹을 위한 적어도 하나의 음조 표시 값을 계산하도록 구성되는 음조 계산 장치를 포함하고,
- [0010] 제어 장치는 음조 표시 값에 의존하여 적어도 하나의 스펙트럼 라인 또는 적어도 하나의 스펙트럼 라인들의 그룹을 위하여 데드 존을 변형하도록 구성된다.
- [0011] 프레임링 장치는 오디오 신호에 대한 윈도우 함수의 적용에 의해 오디오 신호로부터 프레임들을 추출하도록 구성될 수 있다. 신호 처리에서, 윈도우 함수(또한 아포디제이션 함수(apodization function) 또는 테이퍼링 함수(tapering function)로서 알려진)는 일부 선택된 간격 외부의 제로 값인 수학 함수이다. 신호에 대한 윈도우 함수의 적용에 의해, 신호는 일반적으로 프레임들로 불리는, 짧은 세그먼트들로 깨질 수 있다.
- [0012] 디지털 오디오 신호 처리에서, 양자화는 큰 입력 값들의 세트를 (계수 가능한) 작은 세트(일부 정밀 단위에 대한 반올림 값과 같은)로 매핑하는 과정이다. 양자화를 실행하는 장치 또는 알고리즘 함수는 양자화기로 불린다.
- [0013] 본 발명에 따르면, 스펙트럼 신호는 오디오 신호의 프레임들을 위하여 계산된다. 스펙트럼 신호는 시간-도메인 신호인, 오디오 신호의 각각의 프레임들의 스펙트럼을 포함할 수 있고, 각각의 스펙트럼은 주파수 도메인 내의 프레임들 중 어느 하나의 표현이다. 주파수 스펙트럼은 신호의 수학적 변환을 거쳐 발생될 수 있고, 결과로서 생기는 값들은 일반적으로 진폭 대 주파수로서 표현된다.
- [0014] 데드 존은 양자화 동안에 사용되는 구역이고, 스펙트럼 라인들(주파수 빈(frequency bin)들) 또는 스펙트럼 라인들의 그룹들(주파수 대역들)이 제로로 매핑된다. 데드 존은 일반적으로 제로의 진폭에 위치하는, 하부 한계, 및 상부 한계를 가지며, 이는 서로 다른 스펙트럼 라인들 또는 스펙트럼 라인들의 그룹들을 위하여 변경될 수 있다.
- [0015] 본 발명에 따르면, 데드 존은 제어 장치에 의해 변형될 수 있다. 제어 장치는 적어도 하나의 스펙트럼 라인 또는 적어도 하나의 스펙트럼 라인들의 그룹을 위한 적어도 하나의 음조 표시 값을 계산하도록 구성되는 음조 계산 장치를 포함한다.
- [0016] 용어 "음조"는 스펙트럼 신호의 음색 특성을 언급한다. 일반적으로 음조는 스펙트럼이 주기적 성분을 대부분 포함하는 경우에 높다고 말할 수 있는데, 이는 프레임의 스펙트럼이 두드러진 피크들을 포함하는 것을 의미한다. 음조 특성의 반대는 소음 특성이다. 후자의 경우에 프레임의 스펙트럼은 더 평평한(flat) 스펙트럼이다.
- [0017] 게다가, 제어 장치는 각각의 음조 표시 값에 의존하여 적어도 하나의 스펙트럼 라인 또는 적어도 하나의 스펙트럼 라인들의 그룹을 위하여 데드 존을 변형하도록 구성된다.
- [0018] 본 발명은 다음과 같은, 신호-적용 데드 존을 갖는 양자화 전략을 나타낸다:
- [0019] ● 현존하는 미디어 코덱들에서의 사용을 허용하는, 어떠한 부가 정보도 필요로 하지 않음,
- [0020] ● 양자화 이전에 복잡도를 절약하는, 어떠한 데드 존을 빈 또는 대역 당 사용하는지를 결정함,
- [0021] ● 대역 주파수 및/또는 신호 음조를 기초로 하여 빈 당 또는 대역 당 데드 존을 결정할 수 있음.
- [0022] 본 발명은 현존하는 코딩 인프라구조에 적용될 수 있는데 그 이유는 인코더 내의 신호 양자화기가 변경되고, 상응하는 디코더가 인코딩된 신호로부터 생산된 (변경되지 않은) 비트스트림을 판독하고 출력을 디코딩할 것이기

때문이다. [6] 및 그 안의 참조문헌들과 달리, 각각의 스펙트럼 라인들의 그룹 또는 각각의 스펙트럼 라인을 위한 데드 존은 양자화 이전에 선택되고, 따라서 그룹 또는 스펙트럼 라인 당 하나의 양자화 연산만이 필요하다. 최종적으로, 양자화기 결정은 두 개의 가능한 데드 존 값 사이를 선택하는 것에 한정되지 않고, 전체 범위의 값을 선택하는데 한정된다. 결정은 아래에 상세히 설명된다. 위에 설명된 음조 적응 양자화 전략은 저 지연-통합 음성 및 오디오 코딩(LD-USAC) 인코더의 변환 코딩 여기(transform coded excitation, TCX) 경로, 확장 고효율-고급 오디오 코딩의 저-지연 변형에서 구현될 수 있다[4].

[0023] 본 발명의 바람직한 일 실시 예에 따르면, 제어 장치는 스펙트럼 라인들 중 어느 하나의 데드 존이 더 큰 음조를 갖는 스펙트럼 라인들 중 어느 하나에서의 데드 존보다 크거나 또는 스펙트럼 라인들의 그룹들 중 어느 하나의 데드 존이 더 큰 음조를 갖는 스펙트럼 라인들의 그룹들 중 어느 하나에서의 데드 존보다 큰 것과 같은 방법으로 데드 존을 변형하도록 구성된다. 이러한 특징들에 의해, 비-음조 스펙트럼 영역들은 제로로 양자화되는 경향이 있을 것인데, 이는 데이터의 양이 감소될 수 있다는 것을 의미한다.

[0024] 본 발명의 바람직한 일 실시 예에 따르면, 제어 장치는 오디오 신호의 프레임의 파워 스펙트럼(power spectrum)을 계산하도록 구성되는 파워 스펙트럼 계산 장치를 포함하고, 파워 스펙트럼은 스펙트럼 라인들 또는 스펙트럼 라인들의 그룹들을 위한 파워 값들을 포함하며, 음조 계산 장치는 파워 스펙트럼을 기초로 하여 적어도 하나의 음조 표시 값을 계산하도록 구성된다. 파워 스펙트럼을 기초로 하여 적어도 하나의 음조 표시 값을 계산함으로써 계산 복잡도는 아주 낮게 유지된다.

[0025] 본 발명의 바람직한 일 실시 예에 따르면, 스펙트럼 라인들 중 어느 하나를 위한 음조 표시 값은 각각의 스펙트럼 라인을 위한 파워 값 및 미리 정의된 수의 파워 스펙트럼의 그 주위의 파워 값들의 합의 비교를 기초로 하거나, 또는 스펙트럼 라인들의 그룹들 중 어느 하나를 위한 음조 표시 값은 각각의 스펙트럼 라인들의 그룹을 위한 파워 값 및 미리 정의된 수의 파워 스펙트럼의 그 주위의 파워 값들의 합의 비교를 기초로 한다. 파워 값을 이웃의 파워 값들과 비교함으로써 파워 스펙트럼의 피크 영역 또는 평면 영역이 쉽게 식별될 수 있고 따라서 음조 표시 값이 쉬운 방법으로 계산될 수 있다.

[0026] 본 발명의 바람직한 일 실시 예에 따르면, 스펙트럼 라인들 중 어느 하나를 위한 음조 표시 값은 오디오 신호의 선행 프레임의 스펙트럼 라인의 음조 표시 값을 기초로 하거나, 또는 스펙트럼 라인들의 그룹들 중 어느 하나를 위한 음조 표시 값은 오디오 신호의 선행 프레임을 위한 스펙트럼 라인들 그룹의 음조 표시 값을 기초로 한다. 이러한 특징들에 의해 데드 존은 시간에 따라 원활한 방식으로 변형될 것이다.

[0027] 본 발명의 바람직한 일 실시 예에 따르면, 음조 표시 값은 다음의 공식에 계산되는데:

$$T_{k,i} = f\left(\frac{P_{k-7,i} + \dots + P_{k-1,i} + P_{k+1,i} + \dots + P_{k+7,i}}{P_{k,i}}, \frac{P_{k-7,i-1} + \dots + P_{k-1,i-1} + P_{k+1,i-1} + \dots + P_{k+7,i-1}}{P_{k,i-1}}\right),$$

[0028] 여기서 i 는 오디오 신호의 특정 프레임을 나타내는 지수이고, k 는 특정 스펙트럼 라인을 나타내는 지수이며, $P_{k,i}$ 는 i 번째 프레임의 k 번째 스펙트럼 라인의 파워 값이며, 혹은 음조 표시 값은 다음의 공식에 계산되며:

$$T_{m,i} = f\left(\frac{P_{m-7,i} + \dots + P_{m-1,i} + P_{m+1,i} + \dots + P_{m+7,i}}{P_{m,i}}, \frac{P_{m-7,i-1} + \dots + P_{m-1,i-1} + P_{m+1,i-1} + \dots + P_{m+7,i-1}}{P_{m,i-1}}\right),$$

[0030] 여기서 i 는 오디오 신호의 특정 프레임을 나타내는 지수이고, m 은 특정 스펙트럼 라인들의 그룹을 나타내는 지수이며, $P_{m,i}$ 는 i 번째 프레임의 m 번째 스펙트럼 라인들의 그룹의 파워 값이다. 공식으로부터 음조 표시 값은 현재 프레임인, i 번째 프레임의 파워 값, 및 선행 프레임인, $i-1$ 번째 프레임으로부터 계산된다는 것을 이해할 것이다. 공식은 $i-1$ 번째 프레임으로부터의 의존성을 생략함으로써 변경될 수 있다. 여기서 k 번째 파워 값의 왼쪽 7개 및 오른쪽 7개의 이웃하는 파워 값들의 합이 계산되고 각각의 파워 값으로 나뉜다. 이러한 공식을 사용하여 낮은 음조 표시 값은 높은 음조를 나타낸다.

[0031] 본 발명의 바람직한 일 실시 예에 따르면, 오디오 인코더는 데드 존을 변경하기 위하여 시작 주파수(start frequency)를 계산하도록 구성되는 시작 주파수 계산 장치를 포함하고, 데드 존은 시작 주파수보다 높거나 또는 동일한 주파수를 표현하는 스펙트럼 라인들만을 위하여 변형된다. 이는 데드 존이 저주파수들을 위하여 고정되고 고주파수들을 위하여 변동되는 것을 의미한다. 이러한 특징들은 더 나은 오디오 품질에 이르게 하는데 그 이유는 인간 청각 시스템이 저주파수에서 더 민감하기 때문이다.

[0032] 본 발명의 바람직한 일 실시 예에 따르면, 시작 주파수 계산 장치는 오디오 신호의 샘플 레이트를 기초로 하여

나 및/또는 인코딩된 신호로부터 생산된 비트스트림을 위하여 예상되는 최대 비트-레이트를 기초로 하여 시작 주파수를 계산하도록 구성된다, 이러한 특징에 의해 오디오 품질은 최적화될 수 있다.

[0034] 본 발명의 바람직한 일 실시 예에 따르면, 오디오 인코더는 오디오 신호의 프레임으로부터 변형 이산 코사인 변환을 계산하도록 구성되는 변형 이산 코사인 변환 계산 장치 및 오디오 신호의 프레임으로부터 변형 이산 사인 변환을 계산하도록 구성되는 변형 이산 사인 변환 계산 장치를 포함하고, 파워 스펙트럼 계산 장치는 변형 이산 코사인 변환 및 변형 이산 사인 변환을 기초로 하여 파워 스펙트럼을 계산하도록 구성된다. 변형 이산 코사인 변환은 어쨌든 오디오 신호를 인코딩하는 목적으로 계산되어야만 한다. 따라서, 변형 이산 사인 변환만이 음조 적응 양자화의 목적을 위하여 부가적으로 계산되어진다. 따라서, 복잡도가 감소될 수 있다. 그러나, 이산 푸리에 변환 또는 홀수 이산 푸리에 변환과 같은 다른 변환들이 사용될 수 있다.

[0035] 본 발명의 바람직한 일 실시 예에 따르면, 파워 스펙트럼 계산 장치는 공식 $R_{k,i} = (MDCT_{k,i})^2 + (MDST_{k,i})^2$ 에 따라 파워 값들을 계산하도록 구성되는데, 여기서 i 는 오디오 신호의 특정 프레임을 나타내는 지수이고, k 는 특정 스펙트럼 라인을 나타내는 지수이며, $MDCT_{k,i}$ 는 i 번째 프레임의 k 번째 스펙트럼 라인에서의 변형 이산 코사인 변환의 값이며, $MDST_{k,i}$ 는 i 번째 프레임의 k 번째 스펙트럼 라인에서의 변형 이산 사인 변환의 값이며, $R_{k,i}$ 는 i 번째 프레임의 k 번째 스펙트럼 라인의 파워 값이다. 위의 공식은 쉬운 방법으로 파워 값들을 계산하도록 허용한다.

[0036] 본 발명의 바람직한 일 실시 예에 따르면, 오디오 인코더는 스펙트럼 신호를 생산하도록 구성되는 스펙트럼 신호 계산 장치를 포함하고, 스펙트럼 신호 계산 장치는 데드 존의 변형에 의해 에너지 손실이 보상되는 것과 같은 방법으로 스펙트럼 신호의 스펙트럼 라인들의 진폭들을 설정하도록 구성되는 진폭 설정 장치(amplitude setting device)를 포함한다. 이러한 특징들에 의해 양자화는 에너지 보존 방식으로 수행될 수 있다.

[0037] 본 발명의 바람직한 일 실시 예에 따르면, 진폭 설정 장치는 각각의 스펙트럼 라인에서 데드 존의 변형에 의존하여 스펙트럼 신호의 진폭들을 설정하도록 구성된다. 예를 들면, 데드 존이 커진 스펙트럼 라인들은 이러한 목적을 위하여 약간 증폭될 수 있다.

[0038] 본 발명의 바람직한 일 실시 예에 따르면, 스펙트럼 신호 계산 장치는 정규화 장치(normalizing device)를 포함한다. 이러한 특징에 의해 뒤따르는 양자화 단계가 쉬운 방법으로 수행될 수 있다.

[0039] 본 발명의 바람직한 일 실시 예에 따르면, 변형 이산 코사인 변환 계산 장치에 의해 계산된 오디오 신호의 프레임으로부터 변형 이산 코사인 변환은 스펙트럼 신호 계산 장치에 제공된다. 이러한 특징에 의해 변형 이산 코사인 변환은 양자화 적응을 위한 목적 및 인코딩된 신호의 계산을 위한 목적으로 사용된다.

[0040] 일 양상에서, 본 발명은 인코더 및 디코더를 포함하는 시스템을 제공하고, 인코더는 본 발명에 따라 지정된다.

[0041] 일 양상에서, 본 발명은 그것으로부터 인코딩된 신호를 생산하기 위하여 오디오 신호를 인코딩하기 위한 방법을 제공하고 방법은:

[0042] 오디오 신호로부터 프레임들을 추출하는 단계;

[0043] 오디오 신호의 프레임으로부터 유도된 스펙트럼 라인들을 양자화 지수들에 매핑하는 단계를 포함하되, 상기 양자화기는 입력 스펙트럼 라인들이 양자화 지수 제로에 매핑되는, 데드 존이 사용됨; 및

[0044] 데드 존을 변형하는 단계;를 포함하고,

[0045] 적어도 하나의 스펙트럼 라인 또는 적어도 하나의 스펙트럼 라인들의 그룹을 위한 적어도 하나의 음조 표시 값이 계산되고,

[0046] 음조 표시 값에 의존하여 적어도 하나의 스펙트럼 라인 또는 적어도 하나의 스펙트럼 라인들의 그룹을 위하여 데드 존이 변형된다.

발명의 효과

[0047] 일 양상에서 본 발명은 컴퓨터 또는 프로세서 상에서 구동될 때, 본 발명에 따른 방법을 실행하기 위한 컴퓨터 프로그램을 제공한다.

도면의 간단한 설명

[0048] 본 발명의 바람직한 실시 예들이 첨부된 도면들과 관련하여 그 뒤에 설명될 것이다.
 도 1은 본 발명에 따른 인코더의 실시 예를 도시한다.
 도 2는 본 발명에 따른 인코더의 작업 원리를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0049] 도 1은 본 발명에 따라 그것으로부터 인코딩된 신호(ES)를 생산하기 위하여 오디오 신호(AS)를 인코딩하기 위한 오디오 인코더(1)를 도시한다. 오디오 인코더(1)는:

[0050] 오디오 신호(AS)로부터 프레임들(F)을 추출하도록 구성되는 프레이밍 장치(2);

[0051] 오디오 신호(AS)의 프레임(F)으로부터 유도된 스펙트럼 신호(SPS)의 스펙트럼 라인들(SL₁₋₃₂, 도 2 참조)을 양자화 지수들(I₀, I₁)에 매핑하도록 구성되는 양자화기(3)를 포함하되, 상기 양자화기(3)는 스펙트럼 라인들(SL₁₋₃₂)이 양자화 지수 제로(I₀)에 매핑되는, 데드 존(DZ, 도 2 참조)을 가짐; 및

[0052] 데드 존(DZ)을 변형하도록 구성되는 제어 장치(4);를 포함하고,

[0053] 제어 장치(4)는 적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위한 적어도 하나의 음조 표시 값(TI₅₋₃₂)을 계산하도록 구성되는 음조 계산 장치(5)를 포함하고,

[0054] 제어 장치(4)는 음조 표시 값(TI₅₋₃₂)에 의존하여 적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위하여 데드 존(DZ)을 변형하도록 구성된다.

[0055] 프레이밍 장치(2)는 오디오 신호(AS)에 대한 윈도우 함수의 적용에 의해 오디오 신호(AS)로부터 프레임들(F)을 추출하도록 구성될 수 있다. 신호 처리에서, 윈도우 함수(또한 아포디제이션 함수 또는 테이퍼링 함수로서 알려진)는 일부 선택된 간격의 외부의 제로 값인 수학 함수이다. 신호(AS)에 대한 윈도우 함수의 적용에 의해, 신호(AS)는 일반적으로 프레임들(F)로 불리는, 짧은 세그먼트들로 깨질 수 있다.

[0056] 디지털 오디오 신호 처리에서, 양자화는 큰 입력 값들의 세트를 (계수 가능한) 작은 세트(일부 정밀 단위에 대한 반올림 값과 같은)로 매핑하는 과정이다. 양자화를 실행하는 장치 또는 알고리즘 함수는 양자화기로 불린다.

[0057] 본 발명에 따르면, 스펙트럼 신호(SPS)는 오디오 신호(AS)의 프레임들(F)을 위하여 계산된다. 스펙트럼 신호(SPS)는 시간-도메인 신호인, 오디오 신호(AS)의 각각의 프레임들(F)의 스펙트럼을 포함할 수 있고, 각각의 스펙트럼은 주파수 도메인 내의 프레임들(F) 중 어느 하나의 표현이다. 주파수 스펙트럼은 신호(AS)의 수학적 변환을 거쳐 발생될 수 있고, 결과로서 생기는 값들은 일반적으로 진폭 대 주파수로서 표현된다.

[0058] 데드 존(DZ)은 양자화 동안에 사용되는 구역이고, 스펙트럼 라인들((SL₁₋₃₂, 주파수 빈들) 또는 스펙트럼 라인들(SL₁₋₃₂)의 그룹들(주파수 대역들)이 제로로 매핑된다. 데드 존(DZ)은 일반적으로 제로의 진폭에 위치하는, 하부 한계, 및 상부 한계를 가지며, 이는 서로 다른 스펙트럼 라인들(SL₁₋₃₂) 또는 스펙트럼 라인들(SL₁₋₃₂)의 그룹들을 위하여 변경될 수 있다.

[0059] 본 발명에 따르면, 데드 존(DZ)은 제어 장치(4)에 의해 변형될 수 있다. 제어 장치(4)는 적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위한 적어도 하나의 음조 표시 값(TI₅₋₃₂)을 계산하도록 구성되는 음조 계산 장치(5)를 포함한다.

[0060] 용어 "음조"는 스펙트럼 신호(SPS)의 음색 특성을 언급한다. 일반적으로 음조는 스펙트럼이 주기적 성분을 대부분 포함하는 경우에 높다고 말할 수 있는데, 이는 프레임(F)의 스펙트럼 또는 그 일부가 두드러진 피크들을 포함하는 것을 의미한다. 음색 특성의 반대는 소음 특성이다. 후자의 경우에 프레임(F)의 스펙트럼 또는 그것들의 부분은 더 평평하다.

[0061] 게다가, 제어 장치(4)는 각각의 음조 표시 값(TI₅₋₃₂)에 의존하여 적어도 하나의 스펙트럼 라인(SL₁₋₃₂) 또는 적어도 하나의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위하여 데드 존(DZ)을 변형하도록 구성된다.

[0062] 본 발명은 다음과 같은, 신호-적용 데드 존(DZ)을 갖는 양자화 전략을 나타낸다:

- [0063] ● 현존하는 미디어 코덱들에서의 사용을 허용하는, 어떠한 부가 정보도 필요로 하지 않음,
- [0064] ● 양자화 이전에 복잡도를 절약하는, 어떠한 데드 존(DZ)을 빈 또는 대역 당 사용하는지를 결정함,
- [0065] ● 대역 주파수 및/또는 신호 음조를 기초로 하여 빈 당 또는 대역 당 데드 존(DZ)을 결정할 수 있음.
- [0066] 본 발명은 현존하는 코딩 인프라구조에 적용될 수 있는데 그 이유는 인코더(1) 내의 신호 양자화기(3)가 변경되고, 상응하는 디코더가 인코딩된 신호로부터 생산된 (변경되지 않은) 비트스트림을 판독하고 출력을 디코딩할 것이기 때문이다. [6] 및 그 안의 참조문헌들과 달리, 각각의 스펙트럼 라인들(SL₁₋₃₂)의 그룹 또는 각각의 스펙트럼 라인(SL₁₋₃₂)을 위한 데드 존(DZ)은 양자화 이전에 선택되고, 따라서 그룹 또는 스펙트럼 라인(SL₁₋₃₂) 당 하나의 양자화 연산만이 필요하다. 최종적으로, 양자화기 결정은 두 개의 가능한 데드 존 값 사이를 선택하는 것에 한정되지 않고, 전체 범위의 값을 선택하는데 한정된다. 결정은 아래에 상세히 설명된다. 위에 설명된 음조 적용 양자화 전략은 저 지연-통합 음성 및 오디오 코딩(LD-USAC) 인코더의 변환 코딩 여기(TCX) 경로, 확장 고 효율-고급 오디오 코딩의 저-지연 변형에서 구현될 수 있다[4].
- [0067] 본 발명의 바람직한 일 실시 예에 따르면, 제어 장치(4)는 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나에서의 데드 존(DZ)이 더 큰 음조를 갖는 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나에서의 데드 존(DZ)보다 크거나 또는 스펙트럼 라인들(SL₁₋₃₂)의 그룹들 중 어느 하나에서의 데드 존(DZ)이 더 큰 음조를 갖는 스펙트럼 라인들(SL₁₋₃₂)의 그룹들 중 어느 하나에서의 데드 존(DZ)보다 큰 것과 같은 방법으로 데드 존(DZ)을 변형하도록 구성된다. 이러한 특징들에 의해, 비-음조 스펙트럼 영역들은 제로로 양자화되는 경향이 있을 것인데, 이는 데이터의 양이 감소될 수 있다는 것을 의미한다.
- [0068] 본 발명의 바람직한 일 실시 예에 따르면, 제어 장치(4)는 오디오 신호(AS)의 프레임(F)의 파워 스펙트럼(PS, 또한 도 2 참조)을 계산하도록 구성되는 파워 스펙트럼 계산 장치(6)를 포함하고, 파워 스펙트럼(PS)은 스펙트럼 라인들(SL₁₋₃₂) 또는 스펙트럼 라인들(SL₁₋₃₂)의 그룹들을 위한 파워 값들(PS₅₋₃₂)을 포함하며, 음조 계산 장치(5)는 파워 스펙트럼(PS)을 기초로 하여 적어도 하나의 음조 표시 값(PS₅₋₃₂)을 계산하도록 구성된다. 파워 스펙트럼(PS)을 기초로 하여 적어도 하나의 음조 표시 값(PS₅₋₃₂)을 계산함으로써 계산 복잡도는 아주 낮게 유지된다. 게다가 정확도가 향상될 수 있다.
- [0069] 본 발명의 바람직한 일 실시 예에 따르면, 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나를 위한 음조 표시 값(TI₅₋₃₂)은 각각의 스펙트럼 라인(SL₁₋₃₂)을 위한 파워 값(PS₅₋₃₂) 및 미리 정의된 수의 파워 스펙트럼(PS)의 그 주위의 파워 값들(PS₅₋₃₂)의 합의 비교를 기초로 하거나, 또는 스펙트럼 라인들의 그룹들(SL₁₋₃₂) 중 어느 하나를 위한 음조 표시 값(TI₅₋₃₂)은 각각의 스펙트럼 라인들(SL₁₋₃₂)의 그룹을 위한 파워 값(PS₅₋₃₂) 및 미리 정의된 수의 파워 스펙트럼(PS)의 그 주위의 파워 값들(PS₅₋₃₂)의 합의 비교를 기초로 한다. 파워 값(PS₅₋₃₂)을 이웃의 파워 값들(PS₅₋₃₂)과 비교함으로써 파워 스펙트럼(PS)의 피크 영역 또는 평면 영역이 쉽게 식별될 수 있고 따라서 음조 표시 값(TI₅₋₃₂)이 쉬운 방법으로 계산될 수 있다.
- [0070] 본 발명의 바람직한 일 실시 예에 따르면, 스펙트럼 라인들(SL₁₋₃₂) 중 어느 하나를 위한 음조 표시 값(TI₅₋₃₂)은 오디오 신호(AS)의 선행 프레임(F)의 스펙트럼 라인(SL₁₋₃₂)의 음조 표시 값(TI₅₋₃₂)을 기초로 하거나, 또는 스펙트럼 라인들(SL₁₋₃₂)의 그룹들 중 어느 하나를 위한 음조 표시 값(TI₅₋₃₂)은 오디오 신호(AS)의 선행 프레임(F)에 대한 스펙트럼 라인들(SL₁₋₃₂)의 그룹의 음조 표시 값(TI₅₋₃₂)을 기초로 한다. 이러한 특징들에 의해 데드 존(DZ)은 시간에 따라 원활한 방식으로 변형될 것이다.
- [0071] 본 발명의 바람직한 일 실시 예에 따르면, 음조 표시 값(TI₅₋₃₂)은 다음의 공식에 계산되는데:

[0072]
$$T_{k,i} = f\left(\frac{P_{k-7,i} + \dots + P_{k-1,i} + P_{k+1,i} + \dots + P_{k+7,i}}{P_{k,i}}, \frac{P_{k-7,i-1} + \dots + P_{k-1,i-1} + P_{k+1,i-1} + \dots + P_{k+7,i-1}}{P_{k,i-1}}\right),$$

[0073] 여기서 i 는 오디오 신호(AS)의 특정 프레임(F)을 나타내는 지수이고, k 는 특정 스펙트럼 라인(SL₁₋₃₂)을 나타내는 지수이며, $T_{k,i}$ 는 상기 i 번째 프레임(F)의 상기 k 번째 스펙트럼 라인(SL₁₋₃₂)의 음조 표시 값(TI₅₋₃₂)이며, $P_{k,i}$

는 i 번째 프레임의 k 번째 스펙트럼 라인(SL_{1-32})의 파워 값(PS_{5-32})이며, 혹은 음조 표시 값(TI_{5-32})은 다음의 공식에 계산되며:

$$T_{m,i} = f\left(\frac{P_{m-7,i} + \dots + P_{m-1,i} + P_{m+1,i} + \dots + P_{m+7,i}}{P_{m,i}}, \frac{P_{m-7,i-1} + \dots + P_{m-1,i-1} + P_{m+1,i-1} + \dots + P_{m+7,i-1}}{P_{m,i-1}}\right),$$

[0074]

[0075]

여기서 i 는 오디오 신호(AS)의 특정 프레임(F)을 나타내는 지수이고, m 은 특정 스펙트럼 라인들(SL_{1-32})의 그룹을 나타내는 지수이며, $P_{m,i}$ 는 i 번째 프레임의 m 번째 스펙트럼 라인들(SL_{1-32})의 그룹의 파워 값(PS_{5-32})이다. 공식으로부터 음조 표시 값(TI_{5-32})은 현재 프레임(F)인, i 번째 프레임의 파워 값(PS_{5-32}), 및 선행 프레임(F)인, $i-1$ 번째 프레임으로부터 계산된다는 것을 이해할 것이다. 공식은 $i-1$ 번째 프레임으로부터의 의존성을 생략함으로써 변경될 수 있다. 여기서 특정 스펙트럼 라인(SL_{1-32})의 k 번째 파워 값(PS_{5-32}) 또는 스펙트럼 라인들(SL_{1-32})의 그룹의 m 번째 파워 값(PS_{5-32})의 왼쪽 7개 및 오른쪽 7개의 이웃하는 파워 값들(PS_{5-32})의 합이 계산되고 각각의 파워 값(PS_{5-32})으로 나뉜다. 이러한 공식을 사용하여 낮은 음조 표시 값(TI_{5-32})은 높은 음조를 나타낸다.

[0076]

본 발명의 바람직한 일 실시 예에 따르면, 오디오 인코더(1)는 데드 존(DZ)을 변경하기 위하여 시작 주파수(SF)를 계산하도록 구성되는 시작 주파수 계산 장치(7)를 포함하고, 데드 존(DZ)은 시작 주파수(SF)보다 높거나 또는 동일한 주파수를 표현하는 스펙트럼 라인들(SL_{1-32})만을 위하여 변형된다. 이는 데드 존(DZ)이 저주파수들을 위하여 고정되고 고주파수들을 위하여 변동되는 것을 의미한다. 이러한 특징들은 더 나은 오디오 품질에 이르게 하는데 그 이유는 인간 청각 시스템이 저주파수에서 더 민감하기 때문이다.

[0077]

본 발명의 바람직한 일 실시 예에 따르면, 시작 주파수 계산 장치(7)는 오디오 신호(AS)의 샘플 레이트를 기초로 하거나 및/또는 인코딩된 신호(ES)로부터 생산된 비트스트림을 위하여 예상되는 최대 비트-레이트를 기초로 하여 시작 주파수(SF)를 계산하도록 구성된다, 이러한 특징에 의해 오디오 품질은 최적화될 수 있다.

[0078]

본 발명의 바람직한 일 실시 예에 따르면, 오디오 인코더(1)는 오디오 신호(AS)의 프레임(F)으로부터 변형 이산 코사인 변환(CT)을 계산하도록 구성되는 변형 이산 코사인 변환 계산 장치(8) 및 오디오 신호(AS)의 프레임(F)으로부터 변형 이산 사인 변환(ST)을 계산하도록 구성되는 변형 이산 사인 변환 계산 장치(9)를 포함하고, 파워 스펙트럼 계산 장치(6)는 변형 이산 코사인 변환(CT) 및 변형 이산 사인 변환(ST)을 기초로 하여 파워 스펙트럼(PS)을 계산하도록 구성된다. 변형 이산 코사인 변환(CT)은 어쨌든 오디오 신호(AS)를 인코딩하는 목적으로 계산되어야만 한다. 따라서, 변형 이산 사인 변환(ST)만이 음조 적응 양자화의 목적을 위하여 부가적으로 계산되어진다. 따라서, 복잡도가 감소될 수 있다. 그러나, 이산 푸리에 변환 또는 홀수 이산 푸리에 변환과 같은 다른 변환들이 사용될 수 있다.

[0079]

본 발명의 바람직한 일 실시 예에 따르면, 파워 스펙트럼 계산 장치(6)는 공식 $P_{k,i} = (MDCT_{k,i})^2 + (MDST_{k,i})^2$ 에 따라 파워 값들을 계산하도록 구성되는데, 여기서 i 는 오디오 신호(AS)의 특정 프레임(F)을 나타내는 지수이고, k 는 특정 스펙트럼 라인(SL_{1-32})을 나타내는 지수이며, $MDCT_{k,i}$ 는 i 번째 프레임의 k 번째 스펙트럼 라인에서의 변형 이산 코사인 변환(CT)의 값이며, $MDST_{k,i}$ 는 i 번째 프레임의 k 번째 스펙트럼 라인에서의 변형 이산 사인 변환(ST)의 값이며, $P_{k,i}$ 는 i 번째 프레임의 k 번째 스펙트럼 라인의 파워 값(PS_{5-32})이다. 위의 공식은 쉬운 방법으로 파워 값들(PS_{5-32})을 계산하도록 허용한다.

[0080]

본 발명의 바람직한 일 실시 예에 따르면, 오디오 인코더(1)는 스펙트럼 신호(SPS)를 생산하도록 구성되는 스펙트럼 신호 계산 장치(10)를 포함하고, 스펙트럼 신호 계산 장치(10)는 데드 존(DZ)의 변형에 기인한 에너지 손실이 보상되는 것과 같은 방법으로 스펙트럼 신호(SPS)의 스펙트럼 라인들(SL_{1-32})의 진폭들을 설정하도록 구성되는 진폭 설정 장치(11)를 포함한다. 이러한 특징들에 의해 양자화는 에너지 보존 방식으로 수행될 수 있다.

[0081]

본 발명의 바람직한 일 실시 예에 따르면, 진폭 설정 장치(11)는 각각의 스펙트럼 라인(SL_{1-32})에서 데드 존(DZ)의 변형에 의존하여 스펙트럼 신호(SPS)의 진폭들을 설정하도록 구성된다. 예를 들면, 데드 존(DZ)이 커진 스펙트럼 라인들(SL_{1-32})은 이러한 목적을 위하여 약간 증폭될 수 있다.

[0082]

본 발명의 바람직한 일 실시 예에 따르면, 스펙트럼 신호 계산 장치(10)는 정규화 장치(12)를 포함한다. 이러한 특징에 의해 뒤따르는 양자화 단계가 쉬운 방법으로 수행될 수 있다.

- [0083] 본 발명의 바람직한 일 실시 예에 따르면, 변형 이산 코사인 변환 계산 장치(8)에 의해 계산된 오디오 신호(AS)의 프레임(F)으로부터의 변형 이산 코사인 변환(CT)은 스펙트럼 신호 계산 장치(10)에 제공된다. 이러한 특징에 의해 변형 이산 코사인 변환(CT)은 양자화 적응을 위한 목적 및 인코딩된 신호(AS)의 계산을 위한 목적으로 사용된다.
- [0084] 도 1은 본 발명의 적응 인코더(1) 내의 데이터 및 제어 정보의 흐름을 도시한다. 특정 주파수(SF) 위의 비-음조 스펙트럼 영역들은 낮은 비트-레이트들에서 상당히 광범위하게 제로로 양자화되는 경향을 가질 것이라는 것을 이해하여야 한다. 그러나, 이는 디코더 내의 제로 빈들 상에 적용된 소음 삽입이 소음 유사 스펙트럼들을 충분히 재구성할 것이고, 양자화는 저주파수 빈들을 더 미세하게 양자화하도록 사용될 수 있는, 비트들을 절약할 것으로 의도된다.
- [0085] 도 2는 본 발명에 따른 인코더의 작업 원리를 도시한다. 여기서, 본 발명에 따른 오디오 인코더(1)의 데드 존(DZ), 오디오 신호(AS)의 프레임(F)의 파워 값들(PS_{5-32})을 갖는 파워 스펙트럼(PS), 값들(TL_{5-32})을 포함하는 음조 및 스펙트럼(SP)의 스펙트럼 라인들(SL_{1-32})이 공동 좌표 시스템 내에 도시되는데, 여기서 x-축은 주파수를 나타내고 y-축은 진폭을 나타낸다. 단순성을 위하여 도 2에서 1보다 큰 매핑 지수는 도시되지 않는다는 것을 이해하여야 한다.
- [0086] 시작 주파수 계산 장치(7)에 의해 계산될 수 있는, 시작 주파수(SF) 아래에, 데드 존은 고정된 크기이다. 실시 예에서 스펙트럼 라인(SL_1)은 데드 존(DZ) 외부에서 끝나고 따라서 이는 지수 0(I_0)으로 매핑될 수 있으나, 반면에 스펙트럼 라인(SL_7)은 데드 존(DZ) 내에서 끝나고 따라서 이는 지수 1(I_1)으로 매핑될 수 있을 것이다. 그러나, 시작 주파수(SF)로의 시작과 더 높은 주파수로의 이동에서, 데드 존(DZ)의 크기는 제어 장치(4)에 의해 변형될 수 있다. 그러한 목적을 위하여, 파워 값들(PS_{5-32})이 위에 설명된 것과 같이 계산된다. 게다가, 값들(TL_{5-32})을 나타내는 음조가 파워 값들(PS_{5-32})로부터 계산된다.
- [0087] $k=20$ 부터 $k=23$ 까지의 영역에서, 파워 스펙트럼(PS)은 값들(TI_{20-23})을 표시하는 낮은 음조를 야기하고, 높은 음조를 표시하는 피크를 갖는다. 위의 나머지 영역들에서 파워 스펙트럼(PS)을 위한 시작 주파수(SF)는 더 평평하고 따라서 그것들 각각의 영역들에서 낮은 음조를 나타내는, 값들(TI_{12-19} 및 TI_{24-32})을 나타내는 음조는 상대적으로 높다. 그 결과 데드 존(DZ)은 $k=12$ 부터 $k=19$ 까지의 영역 및 $k=24$ 부터 $k=32$ 까지의 영역에서 확대된다. 데드 존(DZ)의 이러한 확대는 예를 들면, 음조 적응 양자화 없이 지수 1로 매핑될 수 있었던, 스펙트럼 라인(SL_{12}) 및 스펙트럼 라인(SL_{25})이 이제 지수 0으로 매핑되는 것을 야기한다. 이러한 제로 양자화는 디코더로 전송되려는 데이터의 양을 감소시킨다.
- [0088] 본 발명의 바람직한 구현에서, 인코더 양자화는 다음과 같이 요약된다:
- [0089] 1. 시간-대-주파수 변환 단계 동안에, 변형 이산 코사인 변환(코사인 부분) 및 변형 이산 사인 변환(사인 부분)은 주어진 프레임을 위하여 윈도우잉된 입력 신호로부터 계산된다.
- [0090] 2. 입력 프레임의 변형 이산 코사인 변환은 양자화, 코딩, 및 변환을 위하여 사용된다. 변형 이산 사인 변환은 또한 빈 당 파워 스펙트럼($P_k = MDCT_k^2 + MDST_k^2$)을 계산하도록 사용된다.
- [0091] 3. 코딩 대역 당, 또는 바람직하게는 빈 당 P_k 로, 음조 또는 스펙트럼 평탄도 측정이 계산된다. 이를 달성하기 위한 일부 방법들이 문헌(1, 2, 3]에 설명된다. 바람직하게는, 빈 당 소수의 연산만을 갖는 낮은 복잡도 버전이 사용된다. 본 발명의 경우에 있어서, P_k 및 그 주위의 $P_{k-7} \dots P_{k+7}$ 의 합 사이의 비교가 만들어지고 [3]에 설명된 탄생/소멸 트래커(birth/death tracker)와 유사한 이력으로 향상된다. 게다가, 특정 비트 레이트 의존 주파수 아래의 빈들은 항상 음조로서 고려된다.
- [0092] 4. 선택적 단계로서, 큰 양자화기 데드 존에 기인한 에너지 손실을 보상하기 위하여 양자화 이전에 스펙트럼의 약간의 증폭을 실행하도록 음조 또는 평탄도 측정이 사용될 수 있다. 더 정확하게는, 큰 양자화기 데드 존이 적용되는 빈들이 약간 증폭되고, 정상 또는 정상 데드 존(즉, 에너지를 보존하는 경향이 있는)에 사용되는 빈들은 변형되지 않는다.
- [0093] 5. 단계 3의 음조 또는 평탄도 측정은 이제 각각의 주파수 빈들 양자화하도록 사용되는 데드 존의 선택을 제어한다. $P_{k-7} \dots P_{k+7}/P_k$ 의 낮은 값들을 측정하는, 높은 음조를 갖는 것으로 결정된 빈들은 디폴드(즉, 대충 에너지를 보존하는) 데드 존으로 양자화되고, 낮은 음조를 갖는 빈들은 새로운 확대된 데드 존으로 양자화된다. 낮은

음조 빈은 따라서 높은 음조 빈보다 더 자주 제로로 양자화되려는 경향이 있다. 선택적으로, 빈의 데드 존의 크기는 디폴트(가장 작은) 및 최대 데드 존 크기 사이의 범위를 갖는, 빈 음조의 연속적인 함수로서 정의될 수 있다.

- [0094] 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 가장 중요한 방법 단계는 그러한 장치에 의해 실행될 수 있다.
- [0095] 특정 구현 요구사항들에 따라, 본 발명의 실시 예들은 하드웨어 또는 소프트웨어에서 구현될 수 있다. 구현은 예를 들면, 각각의 방법이 실행될 것과 같이 프로그램가능 컴퓨터 시스템과 협력하는(또는 협력할 수 있는), 그 안에 저장되는 전자적으로 판독 가능한 제어 신호들을 갖는, 디지털 저장 매체, 예를 들면, 플로피 디스크, DVD, CD, ROM, PROM, EPROM, EEPROM 또는 플래시 메모리를 사용하여 실행될 수 있다.
- [0096] 본 발명에 따른 일부 실시 예들은 여기에 설명된 방법들 중 어느 하나가 실행되는 것과 같이, 프로그램가능 컴퓨터 시스템과 협력할 수 있는, 전자적으로 판독 가능한 제어 신호들을 갖는 데이터 캐리어를 포함한다.
- [0097] 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.
- [0098] 다른 실시 예들은 기계 판독가능 캐리어 상에 저장되는, 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 컴퓨터 프로그램을 포함한다.
- [0099] 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.
- [0100] 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.
- [0101] 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.
- [0102] 또 다른 실시 예는 여기에 설명된 방법들 중 어느 하나를 실행하도록 구성되거나 혹은 적용되는, 처리 수단, 예를 들면 컴퓨터, 또는 프로그램가능 논리 장치를 포함한다.
- [0103] 또 다른 실시 예는 그 안에 여기에 설명된 방법들 중 어느 하나를 실행하기 위한 컴퓨터 프로그램이 설치된 컴퓨터를 포함한다.
- [0104] 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.
- [0105] 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.
- [0106] 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.
- [0107] 참고문헌:

- [0108] [1] L. Daudet, "Sparse and Structured Decomposition of Signals with the Molecular Matching Pursuit," IEEE Trans. on Audio, Speech, and Lang. Processing, Vol. 14, No. 5, Sep. 2006.
- [0109] [2] F. Keiler, "Survey on Extraction of Sinusoids in Stationary Sounds," in Proc. DAFX, 2002.
- [0110] [3] R.J. McAulay and T.F. Quatieri, "Speech Analysis/Synthesis Based on a Sinusoidal Representation," IEEE Trans. Acoustics, Speech, and Sig. Processing, Vol. 34, No. 4, Aug. 1986.
- [0111] [4] M. Neuendorf et al., "MPEG Unified Speech and Audio Coding - The ISO/MPEG Standard for High-Efficiency Audio Coding of All Content Types," in Proc. 132nd Convention of the AES, Budapest, Hungary, Apr. 2012. Also to appear in the Journal of the AES, 2013.
- [0112] [5] M. Oger et al., "Model-Based Deadzone Optimization for Stack-Run Audio Coding with Uniform Scalar Quantization," in Proc. ICASSP 2008, Las Vegas, USA, Apr. 2008.
- [0113] [6] M. Schug, EP2122615, "Apparatus and method for encoding an information signal", 2007.

부호의 설명

- [0114] 1 : 오디오 인코더
- 2 : 프레임링 장치
- 3 : 양자화기
- 4 : 제어 장치
- 5 : 음조 계산 장치
- 6 : 파워 스펙트럼 계산 장치
- 7 : 시작 주파수 계산 장치
- 8 : 변형 이산 코사인 변환 계산 장치
- 9 : 변형 이산 사인 변환 계산 장치
- 10 : 스펙트럼 신호 계산 장치
- 11. 진폭 설정 장치
- 12 : 정규화 장치
- AS : 오디오 신호
- ES : 인코딩된 신호
- F : 프레임
- SL_{1-32} : 스펙트럼 라인
- SPS : 스펙트럼 신호
- I : 지수
- DZ : 데드 존
- TI_{5-32} : 음조 표시 값
- PS : 파워 스펙트럼
- PS_{5-32} : 파워 값
- SF : 시작 주파수
- CT : 변형 이산 코사인 변환

ST : 변형 이산 사인 변환

도면

도면1

