

특허청구의 범위

청구항 1

이진 출력 시퀀스(Φ)로 디지털 오디오 신호 프레임(S)을 코딩하는 방법-코딩 비트들의 최대 개수(N_{max})는 상기 신호 프레임에 따라 계산될 수 있는 파라미터들의 세트에 대하여 정의되고, 상기 파라미터들의 세트는 제1 및 제2 서브세트로 구성됨-으로서,

- 상기 제1 서브세트의 파라미터들을 계산하고, N_0 개($N_0 < N_{max}$)의 코딩 비트에 대해 상기 파라미터들을 코딩하는 단계;
- 상기 제2 서브세트의 파라미터들에 대한 $N_{max} - N_0$ 코딩 비트들의 할당을 판단하는 단계; 및
- 정해진 순서로 상기 제2 서브세트의 파라미터들에 할당된 상기 $N_{max} - N_0$ 코딩 비트들의 랭킹을 매기는 단계를 포함하고,

상기 $N_{max} - N_0$ 코딩 비트들의 할당 및/또는 랭킹 순서는 상기 제1 서브세트의 코딩된 파라미터들의 함수로서 판단되며,

상기 방법은, $N_0 < N \leq N_{max}$ 인 상기 파라미터들의 세트의 코딩에 사용가능한 상기 이진 출력 시퀀스의 N 개 비트의 표시에 응답하여,

- 상기 순서에서 첫번째 랭킹으로 매겨진 $N - N_0$ 코딩 비트들이 할당된 상기 제2 서브세트의 파라미터들을 선택하는 단계;
- 상기 제2 서브세트의 상기 선택된 파라미터들을 계산하고, 첫번째 랭킹으로 매겨진 상기 $N - N_0$ 코딩 비트들을 생성하도록 상기 파라미터들을 코딩하는 단계; 및
- 상기 제2 서브세트의 상기 선택된 파라미터들의 $N - N_0$ 코딩 비트들뿐만 아니라, 상기 제1 서브세트의 N_0 코딩 비트들을 상기 출력 시퀀스에 삽입하는 단계

를 더 포함하는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 2

제1항에 있어서, 상기 제2 서브세트의 상기 파라미터들에 할당된 상기 코딩 비트들의 랭킹 순서는 한 프레임부터 다른 프레임까지 변화가능한 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 3

제1항에 있어서, $N < N_{max}$ 인 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 4

제1항에 있어서, 상기 제2 서브세트의 상기 파라미터들에 할당된 상기 코딩 비트들의 랭킹 순서는 적어도 상기 제1 서브세트의 상기 코딩된 파라미터들의 함수로서 판단된 중요도가 감소하는 순서인 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 5

제4항에 있어서, 상기 제2 서브세트의 상기 파라미터들에 할당된 상기 코딩 비트들의 랭킹 순서는 적어도 하나의 심리음향적 기준의 도움으로 상기 제1 서브세트의 상기 코딩된 파라미터들의 함수로서 판단되는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 6

제5항에 있어서, 상기 제2 서브세트의 상기 파라미터들은 상기 신호의 스펙트럼 대역에 속하고, 상기 코딩된 신호의 스펙트럼 엔벨로프(spectral envelope)는 상기 제1 서브세트의 상기 코딩된 파라미터들을 기초로 하여 추정되며(estimated), 주파수 마스킹 커브는 청각적 인식 모델을 상기 추정된 스펙트럼 엔벨로프에 적용함으로써 계산되고, 상기 심리음향적 기준은 각 스펙트럼 대역의 상기 마스킹 커브와 관련된 상기 추정된 스펙트럼 엔벨

로프의 레벨을 참조하는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 7

제4항에 있어서, $N_{max} = N$ 인 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 8

제1항에 있어서, 상기 제1 서브세트의 N_0 코딩 비트들이 상기 제2 서브세트의 상기 선택된 파라미터들의 상기 $N - N_0$ 코딩 비트들에 우선하는 방식, 및 상기 제2 서브세트의 상기 선택된 파라미터들의 상기 각 코딩 비트들이 상기 코딩 비트들에 대하여 판단된 상기 순서로 나타나는 방식으로, 상기 코딩 비트들이 상기 출력 시퀀스에서 순서화되는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 9

제1항에 있어서, 상기 개수(N)는 하나의 프레임부터 또 다른 프레임까지 변화하는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 10

제1항에 있어서, 상기 제1 서브세트의 상기 파라미터들은 가변 비트 레이트로 코딩되어, 개수(N_0)가 하나의 프레임부터 또 다른 프레임까지 변화하는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 11

제1항에 있어서, 상기 제1 서브세트는 코더 커널(coder kernel: 1)에 의해 계산되는 파라미터들을 포함하는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 12

제11항에 있어서, 상기 코더 커널(1)은 코딩될 상기 신호의 대역폭보다 낮은 동작 주파수 대역을 갖고, 또한 상기 제1 서브세트는 상기 코더 커널의 상기 동작 대역보다 높은 주파수 대역과 관련된 상기 오디오 신호의 에너지 레벨을 포함하는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 13

제12항에 있어서, 상기 동작 대역보다 높은 주파수 대역과 관련된 상기 에너지 레벨의 상기 코딩 비트들이 상기 코더 커널에 의해 계산된 상기 파라미터들의 상기 코딩 비트들에 즉시 후속하는 방식으로, 상기 제1 서브세트의 상기 코딩 비트들이 상기 출력 시퀀스에서 순서화되는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 14

제11항에 있어서, 코딩될 상기 신호와 상기 코더 커널에 의해 생성된 상기 코딩된 파라미터들로부터 나오는 합성 신호와의 차분(difference) 신호가 추정되고, 또한 상기 제1 서브세트는 상기 코더 커널의 상기 동작 대역에 포함되는 주파수 대역과 관련된 상기 차분 신호의 에너지 레벨을 포함하는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 15

제12항에 있어서, 상기 주파수 대역과 관련된 상기 에너지 레벨의 상기 코딩 비트들이 상기 코더 커널(1)에 의해 계산된 상기 파라미터들의 상기 코딩 비트들에 후속하는 방식으로, 상기 제1 서브세트의 상기 코딩 비트들이 상기 출력 시퀀스에서 순서화되는 것을 특징으로 하는 디지털 오디오 신호 프레임 코딩 방법.

청구항 16

디지털 오디오 신호(\hat{S})를 합성하기 위해 이진 입력 시퀀스(Φ')를 디코딩하는 방법-코딩 비트들의 최대 개수(N_{max})는 신호 프레임을 기술하기 위한 파라미터들의 세트에 대하여 정의되고, 상기 파라미터들의 세트는 제1 및 제2 서브세트로 구성되며, 상기 입력 시퀀스는 신호 프레임에 대하여 상기 파라미터들의 세트에 대한 N' 개

($N' \leq N_{max}$)의 코딩 비트들을 포함함으로써,

- $N_0 < N'$ 인 경우에, 상기 입력 시퀀스의 상기 N' 비트들로부터 상기 제1 서브세트의 상기 파라미터들의 N_0 개의 코딩 비트들을 추출하는 단계;
- 추출된 상기 N_0 코딩 비트들을 기초로 하여 상기 제1 서브세트의 상기 파라미터들을 복원하는 단계;
- 상기 제2 서브세트의 상기 파라미터들에 대한 $N_{max} - N_0$ 코딩 비트들의 할당을 판단하는 단계; 및
- 정해진 순서로 상기 제2 서브세트의 상기 파라미터들에 할당된 $N_{max} - N_0$ 코딩 비트들의 랭킹을 매기는 단계를 포함하고,

상기 $N_{max} - N_0$ 코딩 비트들의 할당 및/또는 랭킹 순서는 상기 제1 서브세트의 상기 복원된 파라미터들의 함수로서 판단되며,

상기 방법은,

- 상기 순서에서 첫번째 랭킹으로 매겨진 상기 $N' - N_0$ 코딩 비트들이 할당되는 상기 제2 서브세트의 파라미터들을 선택하는 단계;
- 상기 입력 시퀀스의 상기 N' 비트들로부터, 상기 제2 서브세트의 상기 선택된 파라미터들의 $N' - N_0$ 코딩 비트들을 추출하는 단계;
- 추출된 상기 $N' - N_0$ 코딩 비트들에 기초하여 상기 제2 서브세트의 상기 선택된 파라미터들을 복원하는 단계; 및
- 상기 제1 및 제2 서브세트들의 상기 복원된 파라미터들을 이용하여 상기 신호 프레임을 합성하는 단계를 포함하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 17

제16항에 있어서, 상기 제2 서브세트의 상기 파라미터들에 할당된 상기 코딩 비트들의 랭킹 순서는 하나의 프레임 임부터 또 다른 프레임까지 변화가능한 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 18

제16항에 있어서, $N' < N_{max}$ 인 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 19

제16항에 있어서, 상기 제2 서브세트의 상기 파라미터들에 할당된 상기 코딩 비트들의 랭킹 순서는 적어도 상기 제1 서브세트의 상기 복원된 파라미터들의 함수로서 판단되는 중요도가 감소하는 순서인 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 20

제19항에 있어서, 상기 제2 서브세트의 상기 파라미터들에 할당된 상기 코딩 비트들의 랭킹 순서는 적어도 하나의 심리음향적 기준의 도움으로 상기 제1 서브세트의 상기 복원된 파라미터들의 함수로서 판단되는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 21

제20항에 있어서, 상기 제2 서브세트의 상기 파라미터들은 상기 신호의 스펙트럼 대역들에 속하고, 상기 신호의 스펙트럼 엔벨로프는 상기 제1 서브세트의 상기 복원된 파라미터들에 기초하여 추정되며, 주파수 마스킹 커브는 청각적 인식 모델을 상기 추정된 스펙트럼 엔벨로프에 적용함으로써 계산되고, 상기 심리음향적 기준은 각 스펙트럼 대역의 상기 마스킹 커브와 관련된 상기 추정된 스펙트럼 엔벨로프의 레벨을 참조하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 22

제16항에 있어서, 상기 제1 서브세트의 상기 파라미터들의 N_0 코딩 비트들은 상기 제2 서브세트의 상기 선택된 파라미터들의 상기 $N' - N_0$ 코딩 비트들이 추출된 위치에 우선하는 상기 시퀀스의 위치에서 수신된 N' 비트들로

부터 추출되는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 23

제16항에 있어서, 상기 신호 프레임을 합성하기 위해, 상기 제2 서브세트의 비선택된 파라미터들은 추출된 상기 N' - NO 코딩 비트들에 기초하여 복원된 적어도 선택된 파라미터들에 기초하여 내삽(interpolation)에 의해 추정되는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 24

제16항에 있어서, 상기 제1 서브세트는 디코더 커널(21)의 입력 파라미터들을 포함하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 25

제24항에 있어서, 상기 디코더 커널(21)은 합성될 상기 신호의 대역폭보다 낮은 동작 주파수 대역을 갖고, 또한 상기 제1 서브세트는 상기 디코더 커널의 상기 동작 대역보다 높은 주파수 대역과 관련된 상기 오디오 신호의 에너지 레벨을 포함하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 26

제25항에 있어서, 상기 동작 대역보다 높은 주파수 대역과 관련된 상기 에너지 레벨의 상기 코딩 비트들이 상기 디코더 커널(21)의 상기 입력 파라미터들의 상기 코딩 비트들에 즉시 후속하는 방식으로, 상기 입력 시퀀스 내의 상기 제1 서브세트의 상기 코딩 비트들이 상기 출력 시퀀스에서 순서회되는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 27

제26항에 있어서, 상기 입력 시퀀스(Φ')의 N' 비트들이 상기 디코더 커널(21)의 상기 입력 파라미터들의 상기 코딩 비트들 및 상기 동작 대역보다 높은 주파수 대역과 관련된 상기 에너지 레벨의 상기 코딩 비트들의 적어도 일부분으로 제한되는 경우에는,

- 상기 입력 시퀀스로부터 상기 디코더 커널의 상기 입력 파라미터들의 상기 코딩 비트들 및 상기 에너지 레벨의 상기 코딩 비트들의 상기 일부분을 추출하는 단계;
- 상기 디코더 커널에서 기저 신호(S')를 합성하고, 상기 추출된 코딩 비트들에 기초하여 상기 동작 대역보다 높은 주파수 대역과 관련된 에너지 레벨을 복원하는 단계;
- 상기 기저 신호의 스펙트럼을 계산하는 단계;
- 에너지 레벨을 상기 입력 시퀀스 내의 코딩되지 않은 에너지 레벨과 관련된 상기 동작 대역보다 높은 대역 각각에 할당하는 단계;
- 대응하는 에너지 레벨 및 상기 스펙트럼의 적어도 하나의 대역 내의 상기 기저 신호의 상기 스펙트럼에 기초하여 상기 동작 대역보다 높은 주파수 대역 각각에 대한 스펙트럼 성분들을 합성하는 단계;
- 기저 신호 보정(correction) 신호를 얻기 위해 상기 합성된 스펙트럼 성분들을 시간 영역으로 변환하는 단계; 및
- 상기 신호 프레임을 합성하기 위해 상기 기저 신호 및 상기 보정 신호를 모두 가산(add)하는 단계를 포함하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 28

제27항에 있어서, 상기 입력 시퀀스 내의 코딩되지 않은 에너지 레벨과 관련된 상기 동작 대역보다 높은 대역에 할당된 에너지 레벨은 상기 추출된 코딩 비트들에 기초하여 복원된 상기 에너지 레벨 및 상기 기저 신호의 상기 스펙트럼에 따라 계산되는 인식적 마스킹 레벨의 분수(fraction)인 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 29

제24항에 있어서, 기저 신호(S')는 상기 디코더 커널에서 합성되고, 또한 상기 제1 서브세트는 합성될 상기 신호와 상기 코더 커널의 상기 동작 대역에 포함되는 주파수 대역과 관련된 상기 기저 신호 사이의 차분 신호의 에너지 레벨을 포함하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 30

제29항에 있어서, $N_0 < N' < N_{max}$ 에 대하여, 주파수 대역 내의 스펙트럼 성분에 속하는 상기 제2 서브세트의 비선택된 파라미터들은 추출된 상기 $N' - N_0$ 코딩 비트들에 기초하여 복원된 선택된 파라미터들 및/또는 상기 기저 신호의 계산된 스펙트럼의 도움으로 추정되는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 31

제30항에 있어서, 주파수 대역 내의 상기 제2 서브세트의 상기 비선택된 파라미터들은 상기 대역의 인접 스펙트럼의 도움으로 추정되고, 인접 스펙트럼은 상기 입력 시퀀스의 상기 N' 코딩 비트들에 기초하여 판단되는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 32

제25항에 있어서, 상기 디코더 커널(21)의 상기 입력 파라미터들의 상기 코딩 비트들은 상기 주파수 대역과 관련된 상기 에너지 레벨의 상기 코딩 비트들이 추출된 위치에 우선하는 상기 시퀀스의 위치에서 수신되는 상기 N' 비트들로부터 추출되는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 33

제16항에 있어서, 상기 개수(N')는 하나의 프레임부터 또 다른 프레임까지 변화하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 34

제16항에 있어서, 상기 개수(N_0)는 하나의 프레임부터 또 다른 프레임까지 변화하는 것을 특징으로 하는 이진 입력 시퀀스의 디코딩 방법.

청구항 35

제1항 내지 제15항 중 어느 한 항의 코딩 방법을 구현하도록 고안된 디지털 신호 프로세싱 수단을 포함하는 오디오 코더.

청구항 36

제16항 내지 제34항 중 어느 한 항의 디코딩 방법을 구현하도록 고안된 디지털 신호 프로세싱 수단을 포함하는 오디오 디코더.

명세서

기술분야

[0001] 본 발명은 오디오 신호의 인코딩 및 디코딩을 위한 장치에 관한 것이고, 상세하게는 디지털화 및 압축된 오디오 신호(음성 및/또는 사운드)의 전송 및 저장 애플리케이션에 포함된다.

배경기술

[0002] 더 상세하게는, 본 발명은 멀티레이트 코딩 시스템으로 불리는, 가변 비트 레이트를 제공하는 능력을 갖는 오디오 코딩 시스템에 속한다. 이러한 시스템은 프로세싱 시에 코딩 비트 레이트가 변화가능하다는 점에서 고정 레이트 코더(coder)와 구별되는데, 이 시스템은 특히 혼성(heterogeneous) 액세스 네트워크 상의 전송에 적합하다. 혼성 액세스 네트워크는 고정 및 이동 액세스, 높은 비트 레이트(ADSL), 낮은 비트 레이트(RTC, GPRS 모뎀) 또는 가변 능력을 갖는 단말기(이동전화, PC 등)를 포함하는 IP 타입의 네트워크이다.

[0003] 본질적으로, 멀티레이트 코더의 두 개의 카테고리, 즉 "스위칭 가능한" 멀티레이트 코더와 "계층적 코더"는 구별된다.

- [0004] "스위칭 가능한" 멀티레이트 코더는 하나의 기술적 계열(예컨대 CELP, 사인과 같은 일시적 코딩 또는 주파수 코딩, 또는 변환에 의한 코딩)에 속하는 코딩 아키텍처에 의존하며, 이러한 아키텍처에서는 비트 레이트의 표시가 코더 및 디코더에 동시에 제공된다. 코더는 이 정보를 사용하여 선택된 비트 레이트에 적절한 테이블 및 알고리즘의 부분을 선택한다. 디코더는 대칭적 방식으로 동작한다. 오디오 코딩을 위해 다양한 스위칭 가능한 멀티레이트 코딩 구조가 제안되었다. 3GPP(3rd Generation Partnership Project) 기구, 전화 대역에서의 NB-AMR("Narrow Band Adaptive Multirate", 기술 규격 3GPP TS 26.090, 버전 5.0.0, 2002년 6월), 또는 광대역에서의 WB-AMR("Wide Band Adaptive Multirate", 기술 규격 3GPP TS 26.190, 버전 5.1.0, 2001년 12월)가 그 예이다. 이러한 코더는 상당히 큰 입도(granularity; NB-AMR에 대하여 8 비트 레이트, WB-AMR에 대하여 9 비트 레이트)를 가지며, 상당히 넓은 비트 레이트 범위(NB-AMR에 대하여 4.75부터 12.2 kbit/s, WB-AMR에 대하여 6.60부터 23.85 kbit/s)에서 동작한다. 그러나, 이 유연성(flexibility)에 대하여 지불되는 비용은 구조의 복잡성이다. 이러한 비트 레이트들 모두를 호스팅하기 위해, 이 코더는 다양하고 상이한 옵션들, 변화되는 양자화 테이블 등을 지원해야 한다. 동작 커브는 비트 레이트와 함께 점진적으로 증가하지만, 이러한 점진적 증가는 선형적이지 않고, 어떤 비트 레이트는 다른 것들 이상으로 최적화된다.
- [0005] "가변(scalable)"으로 불리는 소위 "계층적" 코딩 시스템에서, 코딩 동작으로부터 발생하는 이진(binary) 데이터는 연속적인 계층에 분배된다. "커널(kernel)"로 불리는 기저(base) 계층은 이진 트레인(train)의 디코딩에 절대적으로 필요한 이진 요소(element)를 형성하고, 디코딩의 최소 품질을 결정한다.
- [0006] 후속 레이어는 디코딩 동작으로부터 발생하는 신호의 품질을 점진적으로 향상시키고, 디코더에 의해 이용되는 새로운 정보를 가져오는 각각의 새로운 계층은 향상된 품질의 신호를 출력에 제공한다.
- [0007] 계층적 코딩의 특별한 특징 중 하나는 코더 또는 디코더에 특정 표시를 제공할 필요 없이 이진 트레인의 일부분을 삭제하기 위해 어떤 레벨에서 전송 또는 저장 체인 중 어떤 것에 개입할 수 있다는 점이다. 디코더는 이진 정보를 사용하여 대응하는 품질의 신호를 수신 및 생성한다.
- [0008] 계층적 코딩 구조의 분야는 많은 과제를 발생시켰다. 어떤 계층적 코딩 구조는 계층화된 코딩 정보를 전달하도록 설계된 하나의 타입의 코더를 기초로 하여 동작한다. 추가적인 계층이 대역폭을 변경하지 않고 출력 신호의 품질을 향상시키는 경우는 "임베디드 코더"라고 부른다(예컨대, R.D. Lacovo 외의 "Embedded CELP Coding for Variable Bit-Rate Between 6.4 and 9.6 kbit/s, Proc. ICASSP 1991, 681~686 페이지를 보시오). 그러나 이러한 타입의 코더는 제안되는 가장 낮은 비트 레이트와 가장 높은 비트 레이트 사이의 큰 갭(gap)을 허용하지 않는다.
- [0009] 신호의 대역폭을 증가시키기 위해 종종 계층(hierarchy)이 사용된다. 커널은 기저대역(baseband) 신호[예컨대 전화 신호(300~3400 Hz)]를 제공하고, 후속 레이어는 추가적인 주파수 대역(예컨대, 7 kHz까지의 광대역, 20 kHz까지의 HiFi 대역 또는 중간 대역 등)을 제공한다. 서브대역 코더, 즉 시간/주파수 변환을 사용하는 코더는 J.P. Princen외의 문헌 "Subband/transform coding using filter banks designs based on time domain aliasing cancellation"(Proc. IEEE ICASSP-87, 2161~2164 페이지) 및 Y. Mahieux 외의 "High Quality Audio Transform Coding at 64 kbit/s"(IEEE Trans. Commun. Vol.42, No.11, 1994년 11월, 3010~3019 페이지)에 기술된 바와 같다.
- [0010] 또한, 커널 및 추가적인 레이어들을 코딩하는 모듈 또는 모듈들에 대하여 상이한 코딩 기술이 종종 사용되며, 하나의 레이어는 각 단계가 서브코더로 구성된 다양한 코딩 단계들을 말한다. 주어진 레벨의 단계의 서브코더는 이전 단계에 의해 코딩되지 않은 신호의 일부를 코딩하거나, 또는 이전 단계 코딩의 나머지(residual)를 코딩할 수 있고, 나머지는 원신호(original signal)에서 디코딩된 신호를 감산함으로써 얻어진다.
- [0011] 이 구조의 장점은 높은 비트 레이트에서 양호한 품질을 제공하고, 충분한 품질을 유지하면서도 비교적 낮은 비트 레이트까지 내려갈 수 있다는 점이다. 특히, 낮은 비트 레이트에 사용되는 기술은 일반적으로 높은 비트 레이트에 유효하지 않으며, 그 반대도 마찬가지이다.
- [0012] 두 개의 상이한 기술(예컨대, CELP 및 시간/주파수 변환 등)을 사용할 수 있게 하는 이러한 구조는 특히 넓은 비트 레이트 범위를 스위핑(sweeping)하는데 특히 유효하다.
- [0013] 그러나, 종래 기술에서 제안된 계층적 코딩 구조는 각 중간 레이어에 할당된 비트 레이트를 정확하게 정의한다. 계층적 이진 트레인의 입도(granularity) 및 특정 파라미터들의 인코딩에 대응하는 각 레이어는 이 파라미터들에 할당된 비트 레이트에 의존한다(전형적으로 레이어는, 한 프레임당 수십 비트를 포함할 수 있고, 신호 프레임은 주어진 지속시간 동안 특정한 개수의 신호 샘플로 구성되며, 후술하는 예에서는 60 ms 신호에 대

응하는 960개의 샘플을 갖는 하나의 프레임을 고려한다).

[0014] 또한, 디코딩된 신호의 대역폭이 이진 요소의 레이어들의 레벨에 따라 변화될 수 있는 경우에, 라인 비트 레이트의 변화는 청취를 방해하는 아티팩트를 생성할 수 있다.

발명의 상세한 설명

[0015] 특히, 본 발명은 기존의 계층적 코딩 및 스위칭 가능한 코딩의 사용시 언급되는 결점을 줄이는 멀티레이트 코딩 솔루션을 제공하는 것을 목적으로 한다.

[0016] 따라서, 본 발명은 이진 출력 시퀀스로 디지털 오디오 신호를 코딩하는 방법을 제공하며, 코딩 비트의 최대 개수(N_{max})는 신호 프레임에 따라 계산될 수 있는 파라미터들의 세트에 대하여 정의되고, 파라미터들의 세트는 제 1 및 제2 서브세트로 구성된다. 제안되는 방법은 다음의 단계들을 포함한다.

[0017] - 제1 서브세트의 파라미터들을 계산하고, $N_0 < N_{max}$ 인 N_0 개의 코딩 비트로 이 파라미터들을 코딩하는 단계;

[0018] - 제2 서브세트의 파라미터들에 대한 $N_{max} - N_0$ 코딩 비트의 할당을 결정하는 단계; 및

[0019] - 결정된 순서에서 제2 서브세트의 파라미터들에 할당된 $N_{max} - N_0$ 코딩 비트의 랭킹(rank)을 매기는 단계.

[0020] $N_{max} - N_0$ 코딩 비트의 랭킹 순서 및/또는 할당은 제1 서브세트의 코딩된 파라미터들의 함수로 결정된다. 또한 코딩 방법은 상기 파라미터들의 세트를 코딩하는데 사용가능한 이진 출력 시퀀스의 비트수(N , 여기서 $N_0 < N \leq N_{max}$)의 표시에 응답하여 다음의 단계들을 포함한다.

[0021] - 상기 순서에서 최우선 랭킹으로 매겨진 $N - N_0$ 코딩 비트가 할당된 제2 서브세트의 파라미터들을 선택하는 단계;

[0022] - 제2 서브세트의 선택된 파라미터들을 계산하고, 최우선 랭킹으로 매겨진 상기 $N - N_0$ 코딩 비트를 생성하도록 이 파라미터들을 코딩하는 단계; 및

[0023] - 제1 서브세트의 N_0 코딩 비트뿐만 아니라 제2 서브세트의 선택된 파라미터들의 $N - N_0$ 코딩 비트를 출력 시퀀스에 삽입하는 단계.

[0024] 본 발명에 따른 방법은 각 프레임에 대하여 적어도 N_0 부터 N_{max} 까지 범위의 비트들에 대응하는 범위에서 동작하는 멀티레이트 코딩을 정의하는 것을 가능케 한다.

[0025] 따라서 기존의 계층적 코딩 및 스위칭가능한 코딩과 관련된 미리 설정된 비트 레이트의 개념은 "커서(cursor)"의 개념에 의해 대체되어, (N_0 미만의 비트수(N)에 대응가능한) 최소값과 (N_{max} 에 대응하는) 최대값 사이에서 비트 레이트를 자유롭게 변화시킬 수 있게 된다. 이 극값들은 잠재적으로 서로 떨어져 있다. 본 방법은 선택된 비트 레이트에 관계없이 코딩 효율에 있어서 양호한 성능을 제공한다.

[0026] 유리하게도, 이진 출력 시퀀스의 비트수(N)는 엄밀하게 N_{max} 미만이다. 코더에 대하여 주목할 점은 사용되는 비트 할당이 코더의 실제 출력 비트 레이트를 참조하지 않고, 디코더에 맞는 또 다른 개수(N_{max})를 참조한다는 것이다.

[0027] 그러나 전송 채널에 사용가능한 순간적인 비트 레이트의 함수에서 $N = N_{max}$ 로 고정하는 것이 가능하다. 이와 같이 스위칭 가능한 멀티레이트 코더의 출력 시퀀스는, 디코더가 N_{max} 를 알고 있어서 제2 서브세트의 코딩 비트들의 구조를 복원할 수 있는 한, 전체 시퀀스를 수신하지 않는 디코더에 의해 프로세싱될 수 있다.

[0028] $N = N_{max}$ 가 가능한 또 다른 경우는 최대 코딩 레이트로 오디오 데이터를 저장하는 경우이다. 더욱 낮은 비트 레이트로 저장된 이 콘텐츠의 N' 비트들을 판독하는 경우에, 디코더는 $N' \geq N_0$ 인 한 제2 서브세트의 코딩 비트 구조를 복원할 수 있다.

[0029] 제2 서브세트의 파라미터들에 할당된 코딩 비트들의 랭킹 순서는 미리 설정된 순서일 수 있다.

[0030] 바람직한 실시예에서, 제2 서브세트의 파라미터들에 할당된 코딩 비트의 랭킹 순서는 가변적이다. 특히, 그것은 제1 서브세트의 코딩된 파라미터들의 함수로 판단되는 중요도가 감소하는 순서일 수 있다. 따라서 그 프레임에 대하여, $N_0 \leq N' \leq N_{max}$ 인 N' 비트들의 이진 시퀀스를 수신하는 디코더는 제1 서브세트의 코딩을 위해 수신된 N_0 비트들로부터 이 순서를 추론할 수 있을 것이다.

[0031] 제2 서브세트의 파라미터들의 코딩을 위한 $N_{max} - N_0$ 비트들의 할당은 고정된 방식으로 수행될 수 있다(이 경우

에, 이 비트들의 랭킹 순서는 적어도 제1 서브세트의 코딩된 파라미터에 의존한다).

- [0032] 바람직한 실시예에서, 제2 서브세트의 파라미터들의 코딩에 대한 $N_{max} - N_0$ 비트들의 할당은 제1 서브세트의 코딩된 파라미터들의 함수이다.
- [0033] 유리하게도, 제2 서브세트의 파라미터들에 할당되는 코딩 비트들의 랭킹 순서는 제1 서브세트의 코딩된 파라미터들의 함수로서 적어도 하나의 심리음향적 기준의 도움으로 판단된다.
- [0034] 제2 서브세트의 파라미터들은 신호의 스펙트럼 대역에 속한다. 이 경우에, 본 방법은 제1 서브세트의 코딩된 파라미터들에 기초하여 코딩된 신호의 스펙트럼 엔벨로프를 추정(estimate)하는 단계, 청각적 인식 모델을 추정된 스펙트럼 엔벨로프에 적용함으로써 주파수 마스킹 커브를 계산하는 단계를 포함하고, 심리음향적 기준은 각 스펙트럼 대역 내의 마스킹 커브와 관련하여 추정된 스펙트럼 엔벨로프의 레벨을 참조한다.
- [0035] 실시예에서, 제1 서브세트의 N_0 코딩 비트들이 제2 서브세트의 선택된 파라미터들의 $N - N_0$ 코딩 비트들에 우선하는 방식, 및 제2 서브세트의 선택된 파라미터들의 각 코딩 비트들이 코딩 비트들에 대하여 판단된 순서로 나타나는 방식으로, 코딩 비트들이 출력 시퀀스에서 순서화된다. 이것은 이진 시퀀스가 잘리는(truncated) 경우에 가장 중요한 부분을 수신하는 것을 가능케 한다.
- [0036] 개수(N)는 하나의 프레임부터 또 다른 프레임까지 변화할 수 있는데, 특히 예컨대 전송 리소스의 사용가능한 용량의 함수로 변화할 수 있다.
- [0037] 본 발명에 따른 멀티레이트 오디오 코딩은 N_0 와 N_{max} 사이에서 자유롭게 선택되는 전송되는 비트수를 특정 시간에 선택할 수 있기 때문에, 매우 유연한 계층적 모드 또는 스위칭 가능한 모드에 사용될 수 있고, 이것을 프레임 바이 프레임(frame by frame)으로 부른다.
- [0038] 제1 서브세트 파라미터들은 가변 비트 레이트로 코딩되어, 개수(N_0)는 하나의 프레임부터 또 다른 프레임까지 변화한다. 이것은 코딩되는 프레임의 함수로 비트 배치를 최적으로 조정하는 것을 가능케 한다.
- [0039] 실시예에서, 제1 서브세트는 코더 커널에 의해 계산되는 파라미터들을 포함한다. 유리하게는, 코더 커널은 코딩되는 신호의 대역폭보다 더 낮은 동작 주파수 대역을 갖고, 또한 제1 서브세트는 코더 커널의 동작 대역보다 더 높은 주파수 대역과 관련된 오디오 신호의 에너지 레벨을 포함한다. 이러한 구조의 타입은 2 레벨의 계층적 코더의 타입이고, 이것은 코더 커널을 통해 충분한 품질로 코딩된 신호를 전달하고, 사용가능한 비트 레이트의 함수로서, 본 발명에 따른 코딩 방법으로부터 나오는 추가적인 정보를 이용하여 코더 커널에 의해 수행되는 코딩을 보충한다.
- [0040] 바람직하게는, 더 높은 주파수 대역과 관련된 에너지 레벨의 코딩 비트들이 코더 커널에 의해 계산된 파라미터들의 코딩 비트들의 바로 뒤를 후속하는 방식으로, 제1 서브세트의 코딩 비트들이 출력 시퀀스에서 순서화된다. 이것은 디코더가 코더 커널의 정보 및 더 높은 주파수 대역과 관련된 코딩된 에너지 레벨의 정보를 소유하는데 충분한 비트들을 수신하는 한, 성공적으로 코딩된 프레임에 대한 하나의 동일한 대역폭을 보장한다.
- [0041] 실시예에서, 코딩되는 신호와 코더 커널에 의해 생성된 코딩된 파라미터들로부터 나오는 합성 신호 사이의 차분 신호가 추정되고, 또한 제1 서브세트는 코더 커널의 동작 대역에 포함되는 주파수 대역과 관련되는 차분 신호의 에너지 레벨을 포함한다.
- [0042] 본 발명의 제2 태양은 본 발명의 코딩 방법에 따라 코딩된 프레임의 디코딩에 대응하는 디지털 오디오 신호를 합성하기 위해 이진 입력 시퀀스를 디코딩하는 방법이다. 이 방법에 따르면, 코딩 비트들의 최대 개수는 신호 프레임을 기술하기 위한 파라미터들의 세트에 대하여 정의되고, 파라미터들의 세트는 제1 및 제2 서브세트로 구성된다. 신호 프레임에 대한 입력 시퀀스는 파라미터들의 세트에 대한 N' 개($N' \leq N_{max}$)의 코딩 비트들을 포함한다. 본 발명에 따른 디코딩 방법은,
 - [0043] - $N_0 < N'$ 인 경우에, 입력 시퀀스의 N' 비트들로부터 제1 서브세트의 파라미터들의 N_0 개의 코딩 비트들을 추출하는 단계;
 - [0044] - 추출된 N_0 코딩 비트들을 기초로 하여 제1 서브세트의 파라미터들을 복원하는 단계;
 - [0045] - 제2 서브세트의 파라미터들에 대한 $N_{max} - N_0$ 코딩 비트들의 할당을 판단하는 단계; 및
 - [0046] - 판단된 순서로 제2 서브세트의 파라미터들에 할당된 $N_{max} - N_0$ 코딩 비트들의 랭킹을 매기는 단계를 포함하고,

- [0047] N_{max} - N_0 코딩 비트들의 할당 및/또는 랭킹 순서는 제1 서브세트의 복원된 파라미터들의 함수로 판단된다. 이 디코딩 방법은,
- [0048] - 상기 순서에서 첫번째로 랭크된 상기 N' - N_0 코딩 비트들이 할당되는 제2 서브세트의 파라미터들을 선택하는 단계;
- [0049] - 입력 시퀀스의 N' 비트들로부터, 제2 서브세트의 선택된 파라미터들의 N' - N_0 코딩 비트들을 추출하는 단계;
- [0050] - 추출된 N' - N_0 코딩 비트들에 기초하여 제2 서브세트의 선택된 파라미터들을 복원하는 단계; 및
- [0051] - 제1 및 제2 서브세트들의 복원된 파라미터들을 이용하여 신호를 합성하는 단계를 더 포함한다.
- [0052] 이 디코딩 방법은 코더에 의해 가상으로 또는 실제적으로 생성된 N_{max} 비트 시퀀스의 잘라버림(truncation)으로 인한 손실(missing) 파라미터들을 재생성하기 위한 절차와 관계가 있다.
- [0053] 본 발명의 제3 태양은 본 발명에 따른 코딩 방법을 구현하도록 고안된 디지털 신호 프로세싱 수단을 포함하는 오디오 코더이다.
- [0054] 본 발명의 또 다른 태양은 본 발명에 따른 디코딩 방법을 구현하도록 고안된 디지털 신호 프로세싱 수단을 포함하는 오디오 디코더이다.
- [0055] 본 발명의 특징들 및 장점들은 첨부된 도면을 참조하여 예시적인 실시예의 상세한 설명으로 명확해질 것이다.

실시예

- [0059] 도 1에 도시된 코더는 2 코딩 단계의 계층적 구조를 갖는다. 제1 코딩 단계(1)는 예컨대 CELP 타입의 전화 대역(300~3400 Hz)에서의 코더 커널로 구성된다. 이 코더는 예컨대 6.4 kbit/s로 고정된 모드에서 ITU-T(International Telecommunication Union)에 의해 표준화된 G.723.1 코더로 간주된다. 이 코더는 표준에 따라 G.723.1 파라미터를 계산하고, 30 ms의 프레임마다 192 코딩 비트(P_1)에 의해 그 파라미터들을 양자화한다.
- [0060] 대역을 광대역(50~7000 Hz)으로 증가시키게 하는 제2 코딩 단계(2)는 도 1의 감산기(2)에 의해 공급되는 제1 단계 코딩의 나머지(E) 상에서 동작한다. 신호 동기 모듈(4)은 코더 커널(10)의 프로세싱에 소요되는 시간만큼 오디오 신호 프레임(S)을 지연시킨다. 신호 동기 모듈(4)의 출력은 감산기(3)에 제공되고, 감산기(3)는 그 출력으로부터, 코더 커널의 출력 비트(P_1)에 의해 표현되는 양자화된 파라미터를 기반으로 동작하는 디코더 커널의 출력과 등가인 합성 신호(S')를 감산한다. 통상적으로, 코더(10)는 로컬 디코더 공급(S')을 합성한다.
- [0061] 코딩될 오디오 신호(S)는 예컨대 7 kHz의 대역폭을 갖고, 16 kHz로 샘플링된다. 하나의 프레임은 예컨대 960 샘플, 즉 60 ms의 신호 또는 코더 커널 G.723.1의 2개의 기본 프레임으로 구성된다. 후자는 8 kHz로 샘플링된 신호 상에서 동작하기 때문에, 신호(S)는 코더 커널(1)의 입력에서 팩터 2로 샘플링된다. 유사하게, 합성 신호(S')는 코더 커널(10)의 출력에서 16 kHz로 오버샘플링된다.
- [0062] 제1 단계(10)의 비트 레이트는 6.4 kbit/s($2 \times N_1 = 2 \times 192 = 384$ bit/frame)이다. 코더가 32 kbit/s($N_{max} = 1920$ bits/frame)의 최대 비트 레이트를 갖는다면, 제2 단계의 최대 비트 레이트는 25.6 kbit/s($1920 - 384 = 1536$ bits/frame)이다. 제2 단계는 예컨대 20 ms(16 kHz에서의 320 샘플)의 기본 프레임들 또는 서브프레임 상에서 동작한다.
- [0063] 제2 단계(2)는 예컨대 MDCT(Modified Discrete Cosine Transform) 타입의 시간/주파수 변환 모듈(5)을 포함하고, 감산기(3)에 의해 얻어진 나머지(E)가 시간/주파수 변환 모듈(5)에 입력된다. 실제적으로, 도 1의 모듈(3, 5)의 동작 방식은 각 20 ms 서브프레임 동안 다음의 동작을 수행함으로써 행해진다.
- [0064] - 모듈(4)에 의해 지연된 입력 신호(S)의 MDCT 변환(여기서 모듈(4)은 320 MDCT 계수를 제공함). 7225 Hz로 제한되는 스펙트럼에서, 단지 최초의 289 MDCT 계수만이 0이 아니다.
- [0065] - 합성 신호(S')의 MDCT 변환. 전화 대역 신호의 스펙트럼을 처리하기 때문에, 최초의 139 MDCT 계수만이 0이 아니다(3450 Hz까지).
- [0066] - 이전 스펙트럼들 사이의 차분(difference) 스펙트럼의 계산.
- [0067] 그 결과 나오는 스펙트럼은 모듈(6)에 의해 서로 상이한 폭의 다양한 대역으로 분배된다. 예컨대, G.723.1 코덱의 대역폭은 21 대역으로 분할될 수 있고, 더 높은 주파수는 추가적인 11 대역으로 분배된다. 이 추가적인

11 대역에서, 나머지(E)는 입력 신호(S)와 동일하다.

- [0068] 모듈(7)은 나머지(E)의 스펙트럼의 엔벨로프(envelope) 코딩을 수행한다. 이것은 차분 스펙트럼의 각 대역의 MDCT 계수 에너지를 계산함으로써 시작된다. 이하에서는 이 에너지를 "스케일 팩터"라고 한다. 32 스케일 팩터들은 차분(difference) 신호의 스펙트럼 엔벨로프를 구성한다. 그 후에 모듈(7)은 두 부분의 양자화를 진행한다. 제1 부분은 전화 대역(0부터 3450 kHz까지의 최초 21 대역)에 대응하고, 제2 부분은 높은 대역(3450부터 7225 Hz까지의 끝 11 대역)에 대응한다. 각 부분에서, 가변 비트 레이트를 갖는 종래의 허프만(Huffman) 코딩을 사용함으로써, 제1 스케일 팩터는 절대적 기준으로 양자화되고, 후속 팩터들은 차분(differential) 기반으로 양자화된다. 이러한 32 스케일 팩터들은 랭크(rank: $i, i=1, 2, 3$)의 각 서브프레임에 대하여 비트($P2$)의 가변 개수 $[N2(i)]$ 로 양자화된다.
- [0069] 양자화된 스케일 팩터들은 도 1의 FQ로 표시된다. 코더 커널(1)의 양자화된 파라미터들로 구성된 제1 서브세트의 양자화 비트($P1, P2$) 및 양자화된 스케일 팩터(FQ)는 숫자 $[N0 = (2 \times N1) + N2(1) + n2(2) + N2(3)]$ 에서 변화가능하다. 차분 $[Nmax - N0 = 1536 - N2(1) - N2(2) - N2(3)]$ 은 더 세밀하게 대역의 스펙트럼들을 양자화하는데 사용가능하다.
- [0070] 모듈(8)은 모듈(6)에 의해 대역들로 분배된 MDCT 계수들을 밴드들에 대해 각각 결정된 양자화된 스케일 팩터(FQ)로 나눔으로써 MDCT 계수들을 정규화한다. 정규화된 스펙트럼들은 공지된 타입의 벡터 양자화 방식을 사용하는 양자화 모듈(9)에 제공된다. 모듈(9)로부터 나오는 양자화 비트는 도 1에서 P3로 표시된다.
- [0071] 출력 멀티플렉서(10)는 모듈들(1, 7, 9)로부터 나오는 비트들($P1, P2, P3$)을 합성하여 코더의 이진 출력 시퀀스(Φ)를 형성한다.
- [0072] 본 발명에 따르면, 현재의 프레임을 표현하는 출력 시퀀스의 비트 총 개수(N)는 $Nmax$ 와 반드시 등가는 아니다. 그것은 $Nmax$ 보다 적을 수 있다. 그러나, 양자화 비트의 밴드 할당은 개수($Nmax$) 기반으로 수행된다.
- [0073] 도 1에서, 이러한 할당은 양자화된 스케일 팩터(FQ) 및 모듈(11)에 의해 계산되는 $Nmax - N0$ 개의 스펙트럼 마스킹(masking) 커브에 기초하여 모듈(12)에 의해 각 서브프레임에 대하여 수행된다.
- [0074] 모듈(11)의 동작 방식은 다음과 같다. 먼저, 모듈(11)은 모듈(7)에 의해 양자화된 차분 신호의 스펙트럼 엔벨로프에 기초하여 신호(S)의 원 스펙트럼 엔벨로프의 근사값을 결정하고, 코더 커널로부터 나오는 합성 신호(S')에 대하여 동일한 해상도(resolution)로 스펙트럼 엔벨로프의 근사값을 판단한다. 또한, 마지막 2 엔벨로프는 단지 상기 제1 서브세트의 파라미터들이 제공되는 디코더에 의해 판단가능하다. 따라서, 신호(S)의 추정된(estimated) 스펙트럼 엔벨로프는 또한 디코더에 사용가능할 것이다. 그 후에, 모듈(11)은 그 자체가 공지된 방식으로, 즉 원래의 추정된 스펙트럼 엔벨로프에 대역간 청각적 인식의 모델(model of band by band auditory perception)을 적용함으로써 스펙트럼 마스킹 커브를 계산한다. 이 커브(11)는 고려되는 각 대역에 대한 마스킹 레벨을 산출한다.
- [0075] 모듈(12)은 차분 신호의 3개의 MDCT 변환의 3×32 대역들 중 시퀀스(Φ)의 $Nmax - N0$ 나머지 비트의 동적 할당을 수행한다. 여기 개시된 본 발명의 구현에서, 각 밴드에서의 마스킹 커브와 관련하여 추정된 스펙트럼 엔벨로프의 레벨을 참조하는 중요한 심리음향적 인지 기준의 함수로써, 이 레벨에 비례하는 비트 레이트가 각 밴드에 할당된다. 다른 랭킹 기준을 사용할 수도 있다.
- [0076] 이러한 비트 할당 후에, 모듈(9)은 얼마나 많은 비트들이 각 서브프레임 내의 각 밴드의 양자화를 위해 고려되어야 하는지를 인식한다.
- [0077] 그럼에도 불구하고, $N < Nmax$ 인 경우에, 이렇게 할당된 비트들 모두가 필수적으로 사용되는 것은 아니다. 밴드들을 표현하는 비트들의 순서화는 중요한 인식 기준의 함수로써 모듈(13)에 의해 수행된다. 모듈(13)은 신호 대 마스크 비율(각 밴드에 있어서 추정된 스펙트럼 엔벨로프 및 마스킹 커브 사이의 비율)의 감소 순서인 감소 중요도의 순서로 3×32 대역의 랭킹을 매긴다. 이 순서는 본 발명에 따라 이진 시퀀스(Φ)를 구성하는데 사용된다.
- [0078] 시퀀스(Φ) 내의 현재 프레임의 코딩에 바람직한 비트 개수(N)의 함수로써, 모듈(9)에 의해 양자화되는 대역은 모듈(13)에 의해 첫번째 랭킹이 매겨진 대역을 선택함으로써, 또한 선택된 각 밴드에 대하여 비트의 개수를 모듈(12)에 의해 결정된 것과 동일하게 유지함으로써 결정된다.
- [0079] 선택된 각 대역의 MDCT 계수들은 총 비트수가 $N - N0$ 와 등가인 비트를 생성하기 위해, 예컨대 벡터 양자화기의

도움으로 할당된 비트수에 따라 모듈(9)에 의해 양자화된다.

- [0080] 출력 멀티플렉서(10)는 도 2에 표현된 후술하는 순서의 시퀀스의 최초 N 비트로 구성되는 이진 시퀀스(Φ)를 구성한다($N = N_{max}$ 인 경우).
- [0081] a) 2개의 G.723.1 프레임에 대응하는 최초의 이진 트레인(384 비트);
- [0082] b) 22번째 스펙트럼 대역(전화 대역을 넘는 최초의 대역)부터 32번째 대역(가변 레이트 허프만 코딩)까지, 3개의 서브프레임($i=1, 2, 3$)에 대한 스케일 팩터를 양자화하기 위한 다음 비트들($F_{22}^{(i)}, \dots, F_{32}^{(i)}$);
- [0083] c) 첫번째 스펙트럼 대역부터 21번째 대역까지(가변 레이트 허프만 코딩), 3개의 서브프레임들($i=1, 2, 3$)에 대한 스케일 팩터들을 양자화하기 위한 다음 비트들($F_1^{(i)}, \dots, F_{21}^{(i)}$);
- [0084] d) 마지막으로, 모듈(13)에 의해 결정된 순서에 따라, 가장 중요한 대역에서부터 가장 덜 중요한 대역까지, 인지적으로 중요한 순서로 96 대역들의 벡터 양자화 인덱스들($M_{c1}, M_{c2}, \dots, M_{c96}$).
- [0085] G.723.1 파라미터들 및 높은 대역의 스케일 팩터들(a 및 b)을 먼저 위치시킴으로써, 이 그룹들(a, b)의 수신에 대응하는 최소값을 넘는 실제 비트 레이트에 관계없이 디코더에 의해 복원 가능한 신호에 대하여 동일한 대역폭을 유지할 수 있다. G.723.1 코딩에 추가하여 높은 대역의 $3 \times 11 = 33$ 스케일 팩터들의 허프만 코딩에 충분한 최소값은 예컨대 8 kbit/s이다.
- [0086] 상기의 코딩 방법은 디코더가 $N_0 \leq N' \leq N_{max}$ 인 N' 비트들을 수신하는 경우에 프레임의 디코딩을 가능케 한다. 이 개수(N')는 일반적으로 하나의 프레임으로부터 또 다른 프레임으로 변화 가능할 것이다.
- [0087] 이러한 예에 대응하는 본 발명에 따른 디코더는 도 3에 예시된다. 디멀티플렉서(20)는 코딩 비트들(P1, P2)를 수신된 비트 시퀀스(Φ')로부터 추출하기 위해, 수신된 비트 시퀀스(Φ')를 분리한다. 384 비트(P1)는 G.723.1 타입의 디코더 커널(21)에 제공되어, 디코더 커널(21)이 전화 대역에서의 기저 신호(S')의 2 프레임들을 합성한다. 비트(P2)는 모듈(22)에 의해 허프만 알고리즘에 따라 디코딩되고, 3 서브프레임 각각에 대한 양자화된 스케일 팩터들(FQ)를 복원(recover)한다.
- [0088] 도 1의 코더의 모듈(11)과 동일한, 마스크 커브를 계산하는 모듈(23)은 기저 신호(S') 및 양자화된 스케일 팩터(FQ)를 수신하고, 96 대역의 각각에 대한 스펙트럼 마스크 레벨을 생성한다. 양자화된 스케일 팩터(FQ) 및 기저(既知) 개수(N_{max})의 이러한 마스크 레벨(또한 모듈(22)에 의해 비트(P2)의 허프만 디코딩으로부터 추론되는 개수(N_0)의 마스크 레벨)에 기초하여, 모듈(24)은 도 1의 모듈(12)과 동일한 방식으로 비트 할당을 결정한다. 또한, 모듈(25)은 도 1을 참조하여 기술되는 모듈(13)과 동일한 랭킹 기준에 따라 대역들을 순서화한다.
- [0089] 모듈(24, 25)에 의해 제공된 정보에 따라, 모듈(26)은 입력 시퀀스(Φ')의 비트(P3)를 추출하고, 시퀀스(Φ')에 표현된 대역들과 관련된 정규화된 MDCT 계수들을 합성한다. 적절한 경우에($N' < N_{max}$), 손실 대역과 관련된 표준화된 MDCT 계수들은 이하에 기술되는 바와 같이 내삽(interpolation) 또는 외삽(extrapolation)에 의해 합성될 수 있다[모듈(27)]. $N < N_{max}$ 인 손실 대역들은 코더에 의한 잘라버림(truncation)으로 제거되었거나, 전송 과정에서 제거되었을 것이다($N' < N$).
- [0090] 모듈(26) 및/또는 모듈(27)에 의해 합성된 표준화 MDCT 계수들은 코더의 모듈(5)에 의해 수행되는 MDCT 역변환인 주파수/시간 변환을 수행하는 모듈(29)로 제공되기 전에, 그것들 각각의 양자화된 스케일 팩터들로 곱해진다[멀티플렉서(28)]. 그 결과로부터 나오는 일시적 보정(correction) 신호가 디코더 커널(21)에 의해 전달되는 합성 신호(S')와 더해져서[가산기(30)], 디코더의 출력 오디오 신호(\hat{S})를 생성한다.
- [0091] 시퀀스의 최초 N_0 비트들을 수신하지 않는 경우에도, 디코더는 신호(\hat{S})를 합성할 수 있다는 것을 유념하라.
- [0092] 상기 목록의 일부분에 대응하는 $2 \times N_1$ 비트를 수신하는 것으로 충분하고, 디코딩은 "저하" 모드가 된다. 이러한 저하 모드는 디코딩된 신호를 얻기 위해 MDCT 합성을 사용하지 않는다. 이 모드와 다른 모드들 사이에서 중단없이 스위칭하는 것을 보장하기 위해, 디코더는 3 MDCT 분석을 수행하고, 그 후에 3 MDCT 합성을 수행하여, MDCT 변환의 메모리 갱신을 가능케 한다. 출력 신호는 전화 대역 품질의 신호를 포함한다. 최초 $2 \times N_1$ 비트가 수신되지 않은 경우라도, 디코더는 소거된 것에 대응하는 프레임을 고려하여, 소거된 프레임들을 알기 위한 공지된 알고리즘을 사용할 수 있다.
- [0093] 디코더가 부분(a) 및 부분(b)에 대응하는 $2 \times N_1$ 비트를 수신하는 경우, 디코더는 광 대역 신호의 합성을 개시할

수 있다. 이 방법은 특히 다음과 같은 절차로 수행될 수 있다.

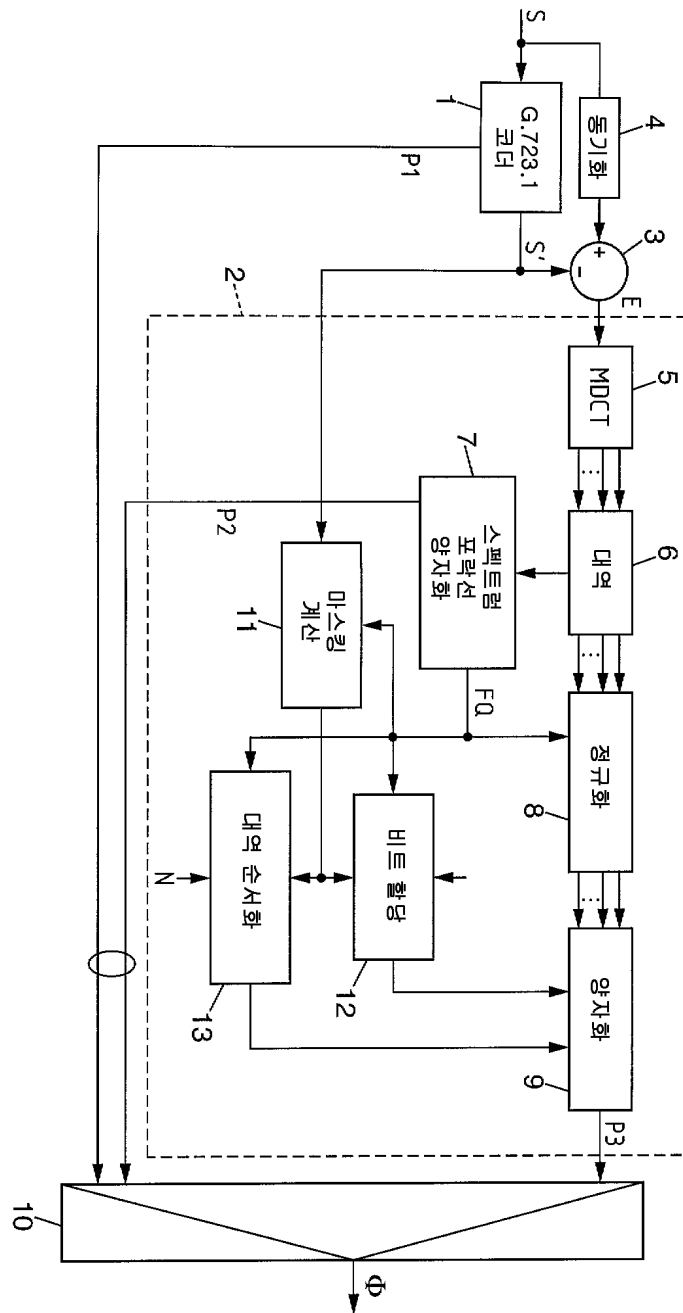
- [0094] 1) 모듈(22)는 수신된 3 스펙트럼 엔벨로프의 일부분을 복원한다.
- [0095] 2) 수신되지 않은 대역들은 일시적으로 0으로 설정된 그 대역들의 스케일 팩터들을 갖는다.
- [0096] 3) 스펙트럼 엔벨로프의 낮은 부분은 G.723.1 디코딩 후에 얻어진 신호 상에서 수행되는 MDCT 분석에 기초하여 계산되고, 모듈(23)은 얻어진 엔벨로프 상의 3 마스킹 커브를 계산한다.
- [0097] 4) 수신되지 않은 대역으로 인한 널(null)을 회피함에 의해 스펙트럼 엔벨로프를 조정하도록 스펙트럼 엔벨로프가 보정된다. 스펙트럼 엔벨로프(FQ)의 높은 부분의 제로값은 예컨대 이전에 계산된 마스킹 커브 값의 100분의 1로 대체되어, 그것들은 청취불가능하게 된다. 낮은 대역의 완성된 스펙트럼 및 높은 대역의 스펙트럼 엔벨로프를 이 시점에 알게 된다.
- [0098] 5) 그 후에 모듈(27)은 높은 스펙트럼을 생성한다. 이 대역의 미세한 구조는 스케일 팩터들로 가중(멀티플렉서: 28)되기 전에 그것에 인접하는 기지의 미세한 구조를 반영하여 생성된다. 비트들(P3) 중 어떤 것도 수신되지 않은 경우에, "기지의 인접 스펙트럼"은 G.723.1 디코더 커널에 의해 생성된 신호(S')의 스펙트럼에 대응한다. 기지의 인접 스펙트럼의 "반영"은 "기지의 인접 스펙트럼"으로부터 떨어진 거리에 비례하여 감소하는 변화량을 갖는 표준화된 MDCT 스펙트럼의 값을 복사하는 것으로 수행될 수 있다.
- [0099] 6) MDCT 역변환(29) 및 그 결과로 나오는 보정 신호를 디코더 커널의 출력 신호와 가산(30)한 후에, 광 대역 합성 신호가 얻어진다.
- [0100] 디코더가 차분 신호의 낮은 엔벨로프의 적어도 일부분을 수신하는 경우에(부분 c), 디코더는 단계 3의 스펙트럼 엔벨로프를 정제(refine)하는데 이 정보를 고려하거나 고려하지 않을 수 있다.
- [0101] 디코더(10)가 적어도 시퀀스의 부분(d)에서 첫번째 랭킹이 매겨진 가장 중요한 대역의 MDCT 계수들을 디코딩할 충분한 비트(P3)를 수신하는 경우에, 모듈(26)은 모듈(24, 25)에 의해 표시되는 할당 및 순서에 따라 정규화된 MDCT 계수들을 어느 정도까지 복원한다. 그러므로 이 MDCT 계수들은 상기 단계 5와 같이 내삽될 필요가 없다. 다른 대역에 대하여는, 단계 1 내지 단계 6의 프로세스가 이전과 동일한 방식으로 모듈(27)에 의해 적용가능하고, 특정 대역에 대하여 수신된 MDCT 계수의 인식은 단계 5에서 더욱 신뢰가능한 내삽을 가능케 한다.
- [0102] 수신되지 않은 대역들은 하나의 MDCT 서브프레임부터 다음 MDCT 서브프레임까지 변화할 수 있다. 손실 대역의 "기지의 인접 스펙트럼"은 빠지지 않은 또 다른 서브프레임 내의 동일한 대역 및/또는 동일한 서브프레임 코스의 주파수 영역에 가장 근접한 하나 이상의 대역들에 대응할 것이다. 또한, "기지의 인접 스펙트럼"의 여러 대역들/서브프레임들을 기반으로 추정된 공현도의 가중된 합을 계산함으로써, 서브프레임에 대한 밴드로부터 손실 MDCT 스펙트럼을 재생성하는 것도 가능하다.
- [0103] 프레임당 N' 비트의 실제 비트 레이트가 임의적으로 주어진 프레임의 최후 비트에 위치하는 한, 송신되는 최후 코딩 파라미터는 경우에 따라 완전하게 또는 부분적으로 송신될 수 있다. 다음과 같은 두 가지의 경우가 일어날 수 있다.
- [0104] - 채택된 코딩 구조가 수신된 부분 정보를 이용하는 것이 가능한 경우(스칼라 양자화기, 또는 분할된 사전(partitioned dictionary)들로 벡터 양자화하는 경우)
- [0105] - 또는, 코딩 구조는 수신된 부분 정보를 이용하는 것이 불가능하고, 완전히(fully) 수신되지 않은 파라미터가 수신되지 않은 다른 파라미터처럼 프로세싱된다. 후자의 경우에, 비트들의 순서가 각 프레임에 따라 변화한다면, 이로 인해 손실된 비트들의 개수는 가변적이고, N' 비트들의 선택은 더 적은 개수의 비트들로 얻어지는 것보다 더욱 나은 품질로, 디코딩된 프레임들의 전체 세트를 평균하여 생성될 것이다.

도면의 간단한 설명

- [0056] 도 1은 본 발명에 따른 예시적인 오디오 코더의 개략도이다.
- [0057] 도 2는 본 발명의 실시예에서 N 비트의 이진 출력 시퀀스를 나타낸다.
- [0058] 도 3은 본 발명에 따른 오디오 디코더의 개략도이다.

도면

도면1



도면3

