



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2008-0101872
(43) 공개일자 2008년11월21일

(51) Int. Cl.

G10L 19/00 (2006.01) H04N 7/24 (2006.01)

(21) 출원번호 10-2008-7016356

(22) 출원일자 2008년07월04일

심사청구일자 없음

번역문제출일자 2008년07월04일

(86) 국제출원번호 PCT/KR2007/000305

국제출원일자 2007년01월18일

(87) 국제공개번호 WO 2007/083934

국제공개일자 2007년07월26일

(30) 우선권주장

60/759,622 2006년01월18일 미국(US)

(뒷면에 계속)

(71) 출원인

연세대학교 산학협력단

서울 서대문구 신촌동 134 연세대학교

엘지전자 주식회사

서울특별시 영등포구 여의도동 20번지

(72) 발명자

정양원

서울 강남구 도곡동 역삼한신아파트 2동 803호

오현오

경기 고양시 일산서구 주엽1동 강선마을3단지한신아파트 306동403호

(뒷면에 계속)

(74) 대리인

김용인, 박영복

전체 청구항 수 : 총 30 항

(54) 부호화/복호화 장치 및 방법

(57) 요약

본 발명은 부호화/복호화 장치 및 방법에 관한 것이다. 그 복호화 방법은, 입력되는 비트스트림으로부터 부호화된 복수의 신호들과 분할 정보를 추출하는 단계; 부호화된 복수의 신호들 각각에 대해 복호화 방식을 결정하는 단계; 복수의 신호들을 결정된 복호화 방식에 따라 복호화하는 단계; 및 분할 정보를 이용하여 상기 복호화된 복수의 신호들을 합성하는 단계를 포함하는 것을 특징으로 한다.

본 발명에 따른 부호화/복호화 장치 및 방법에 의하면, 특성에 따라 신호를 분류하여 그에 맞는 부호화기를 사용해 신호를 부호화함으로써, 서로 다른 특성을 가지는 신호들을 최적의 비트율로 부호화할 수 있다. 그에 따라 오디오 신호, 음성 신호 등의 여러 신호들을 모두 효율적으로 부호화할 수 있다.

(72) 발명자

김효진

서울 마포구 성산1동 258-18번지 303호

최승중

서울특별시 서초구 반포4동 미도아파트 305-904

이동금

서울특별시 서대문구 신촌동 134 연세대학교 전기
전자공학부

강홍구

서울특별시 서대문구 신촌동 134 연세대학교 전기
전자공학부

이재성

서울특별시 서대문구 신촌동 134 연세대학교 전기
전자공학부

(30) 우선권주장

60/797,782 2006년05월03일 미국(US)

60/817,926 2006년06월29일 미국(US)

60/844,510 2006년09월13일 미국(US)

60/848,217 2006년09월29일 미국(US)

60/860,822 2006년11월24일 미국(US)

특허청구의 범위

청구항 1

입력되는 비트스트림으로부터 부호화된 복수의 신호들과 상기 복수의 신호들의 분할 정보를 추출하는 단계;

상기 부호화된 복수의 신호들 각각에 대해, 복수의 복호화 방식들 중 상기 신호를 복호화할 방식을 결정하는 단계;

상기 복수의 신호들을 상기 결정된 복호화 방식에 따라 복호화하는 단계; 및

상기 추출된 분할 정보를 이용하여, 상기 복호화된 복수의 신호들을 합성하는 단계를 포함하는 것을 특징으로 하는 복호화 방법.

청구항 2

제1항에 있어서, 상기 분할 정보는

상기 복수의 신호들의 개수 또는 주파수 대역에 대한 정보를 포함하는 것을 특징으로 하는 복호화 방법.

청구항 3

제1항에 있어서, 상기 부호화된 복수의 신호들은 복수의 주파수 대역들로 분할된 것을 특징으로 하는 복호화 방법.

청구항 4

제3항에 있어서,

상기 복수의 주파수 대역들은 가변적인 것을 특징으로 하는 복호화 방법.

청구항 5

제1항에 있어서, 상기 부호화된 복수의 신호들은 음성 복호화기를 이용해 효율적으로 복호화될 수 있는 신호와 오디오 부호화기를 이용해 효율적으로 복호화될 수 있는 신호로 분할된 것을 특징으로 하는 복호화 방법.

청구항 6

제1항에 있어서,

상기 부호화된 복수의 신호들 각각의 클래스 정보를 상기 비트스트림으로부터 추출하는 단계를 포함하고,

상기 복호화 방식 결정 단계는 상기 추출된 클래스 정보를 이용하여 상기 복수의 복호화 방식들 중 상기 신호의 복호화 방식을 결정하는 것을 특징으로 하는 복호화 방법.

청구항 7

제6항에 있어서, 상기 클래스 정보는

상기 신호의 부호화 또는 복호화 방식에 대한 정보와 상기 신호의 특성에 대한 정보 중 적어도 하나를 포함하는 것을 특징으로 하는 복호화 방법.

청구항 8

제6항에 있어서, 상기 클래스 정보는

음성 복호화 방식과 오디오 복호화 방식 중 상기 신호를 가장 효율적으로 복호화할 방식에 대한 정보인 것을 특징으로 하는 복호화 방법.

청구항 9

제6항에 있어서, 상기 클래스 정보는

상기 신호가 모델링이 잘되는 신호인지 여부에 대한 정보인 것을 특징으로 하는 복호화 방법.

청구항 10

제1항에 있어서, 상기 복호화 방식 결정 단계는

상기 신호가 모델링이 잘되는 신호인 경우 상기 신호를 음성 복호화 방식을 이용해 복호화하도록 결정하고, 상기 신호가 모델링이 잘되지 않는 신호인 경우 상기 신호를 오디오 복호화 방식을 이용해 복호화하도록 결정하는 것을 특징으로 하는 복호화 방법.

청구항 11

제8항 내지 제10항 중 어느 한 항에 있어서,

상기 음성 복호화 방식은 상기 신호를 시간 영역에서 복호화하며, 상기 오디오 복호화 방식은 상기 신호를 주파수 영역에서 복호화하는 것을 특징으로 하는 복호화 방법.

청구항 12

제1항에 있어서,

상기 신호의 복호화 방식은 상기 신호의 변화량 또는 토널 성분 포함 정도에 따라 결정되는 것을 특징으로 하는 복호화 방법.

청구항 13

제1항에 있어서, 상기 합성 단계는

상기 복호화된 복수의 신호들 중 적어도 하나를 복수의 신호들로 분할하는 단계; 및

상기 분할된 복수의 신호들 중 2 이상의 신호를 하나의 신호로 병합하는 단계를 포함하는 것을 특징으로 하는 복호화 방법.

청구항 14

제1항에 있어서, 상기 합성 단계는

상기 복호화된 복수의 신호들 중 2 이상의 신호를 합성하는 단계; 및

상기 복호화된 복수의 신호들 중 적어도 하나와 상기 합성된 신호를 합성하는 단계를 포함하는 것을 특징으로 하는 복호화 방법.

청구항 15

입력되는 비트스트림으로부터 부호화된 복수의 신호들과 상기 복수의 신호들의 분할 정보를 추출하는 비트언팩킹부;

상기 부호화된 복수의 신호들 각각에 대해, 복수의 복호화기들 중 상기 신호를 복호화할 복호화기를 결정하는 복호화기결정부;

상기 복수의 복호화기들을 포함하며, 상기 복수의 신호들을 상기 결정된 복호화기를 이용하여 복호화하는 복호화부; 및

상기 추출된 분할 정보를 이용하여, 상기 복호화된 복수의 신호들을 합성하는 합성부를 포함하는 것을 특징으로 하는 복호화 장치.

청구항 16

제15항에 있어서, 상기 분할 정보는

상기 복수의 신호들의 개수 또는 주파수 대역에 대한 정보를 포함하는 것을 특징으로 하는 복호화 장치.

청구항 17

제15항에 있어서, 상기 비트언팩킹부는

상기 복수의 신호들 각각의 복호화기 정보를 상기 비트스트림으로부터 추출하는 것을 특징으로 하는 복호화 장치.

청구항 18

제15항에 있어서, 상기 복호화부는

음성 복호화기 및 오디오 복호화기를 포함하고,

상기 복호화기결정부는

상기 신호가 모델링이 잘되는 신호인 경우 음성 복호화를 이용해 상기 신호를 복호화하도록 결정하며, 상기 신호가 모델링이 잘되지 않는 신호인 경우 오디오 복호화를 이용해 상기 신호를 복호화하도록 결정하는 것을 특징으로 하는 복호화 장치.

청구항 19

입력되는 신호를 복수의 신호들로 분할하는 단계;

상기 분할된 복수의 신호들 각각을 상기 신호의 특성에 기초하여 복수의 클래스들 중 어느 하나로 분류하는 단계;

상기 분류된 클래스에 대응되는 부호화 방식을 이용하여, 상기 복수의 신호들을 부호화하는 단계; 및

상기 부호화된 복수의 신호를 이용하여 비트스트림을 생성하는 단계를 포함하는 것을 특징으로 하는 부호화 방법.

청구항 20

제19항에 있어서, 상기 신호 분할 단계는

극(extrema)의 개수와 제로-크로싱(zero-crossing)의 개수의 차이가 1 이하이며, 최대값들에 의한 포락선과 최소값들에 의한 포락선의 평균이 실질적으로 0인 복수의 신호들로 분할하는 것을 특징으로 하는 부호화 방법.

청구항 21

제19항에 있어서, 상기 신호 분할 단계는

상기 입력되는 신호를 음성 부호화 방식을 이용해 효율적으로 부호화될 수 있는 신호와 오디오 부호화 방식을 이용해 효율적으로 부호화될 수 있는 신호로 분할하는 것을 특징으로 하는 부호화 방법.

청구항 22

제19항에 있어서, 상기 신호 분할 단계는

상기 입력되는 신호를 복수의 신호들로 분할하는 단계; 및

상기 분할된 복수의 신호들 중 2 이상의 신호들을 하나의 신호로 병합하는 단계를 포함하는 것을 특징으로 하는 부호화 방법.

청구항 23

제22항에 있어서, 상기 병합 단계는

상기 분할된 복수의 신호들 중 서로 이웃하지 않는 2 이상의 신호들을 하나의 신호로 병합하며, 상기 병합되는 2 이상의 신호들은 유사한 특성을 가지는 것을 특징으로 하는 부호화 방법.

청구항 24

제19항에 있어서, 상기 신호 분할 단계는

상기 입력되는 신호를 복수의 신호들로 분할하는 단계; 및

상기 분할된 복수의 신호들 중 적어도 하나를 2 이상의 신호들로 분할하는 단계를 포함하는 것을 특징으로 하는

부호화 방법.

청구항 25

제19항에 있어서, 상기 분류 단계는

음성 부호화 방식과 오디오 부호화 방식 중 상기 신호를 가장 효율적으로 부호화할 수 있는 방식을 판단하는 것을 특징으로 하는 부호화 방법.

청구항 26

입력되는 신호를 복수의 신호들로 분할하고, 상기 분할된 복수의 신호들 각각을 상기 신호의 특성에 기초하여 복수의 클래스들 중 어느 하나로 분류하는 분류부;

상기 복수의 클래스들에 대응되는 복수의 부호화기를 포함하며, 상기 복수의 신호들을 상기 분류된 클래스에 대응되는 부호화기를 이용하여 부호화하는 부호화부; 및

상기 부호화된 복수의 신호를 이용하여 비트스트림을 생성하는 비트패킹부를 포함하는 것을 특징으로 하는 부호화 장치.

청구항 27

제26항에 있어서, 상기 분류부는

상기 입력되는 신호를 복수의 신호들로 분할하는 분할부; 및

상기 분할된 복수의 신호들 중 2 이상의 신호들을 하나의 신호로 병합하는 병합부를 포함하는 것을 특징으로 하는 부호화 장치.

청구항 28

제26항에 있어서, 상기 분류부는

상기 입력되는 신호를 복수의 신호들로 분할하는 제1 분할부; 및

상기 분할된 복수의 신호들 중 적어도 하나를 2 이상의 신호들로 분할하는 제2 분할부를 포함하는 것을 특징으로 하는 부호화 장치.

청구항 29

제26항에 있어서, 상기 부호화부는

음성 부호화기와 오디오 부호화기를 포함하고,

상기 분류부는

상기 음성 부호화기와 오디오 부호화기 중 상기 신호를 가장 효율적으로 부호화할 수 있는 부호화기를 판단하는 것을 특징으로 하는 부호화 장치.

청구항 30

제1항 내지 제14항 및 제19항 내지 제25항 중 어느 한 항에 기재된 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

명세서

기술 분야

<1> 본 발명은 신호의 부호화/복호화 장치 및 방법에 관한 것으로, 더욱 상세하게는 신호의 특성에 따라 최적의 비트율로 부호화/복호화할 수 있도록 하는 효율적인 부호화/복호화 장치 및 방법에 관한 것이다.

배경 기술

<2> 종래의 오디오 부호화기는 48kbps 이상의 높은 비트율에서는 고음질의 오디오 신호를 제공하지만 음성 신호의

처리에는 비효율적이며, 종래의 음성 부호화기는 12kbps 이하의 낮은 비트율에서 음성 신호를 효과적으로 부호화할 수 있지만 다양한 오디오 신호를 부호화하기에 부족하다.

발명의 상세한 설명

<3> 기술적 과제

<4> 본 발명이 이루고자 하는 기술적 과제는, 음성 신호, 오디오 신호 등과 같이 서로 다른 특성을 가지는 신호들을 최적의 비트율로 부호화할 수 있도록 하는 부호화/복호화 장치 및 방법을 제공하는데 있다.

<5> 기술적 해결방법

<6> 상술한 기술적 과제를 해결하기 위한 본 발명에 따른 복호화 방법은, 입력되는 비트스트림으로부터 부호화된 복수의 신호들과 상기 복수의 신호들의 분할 정보를 추출하는 단계; 상기 부호화된 복수의 신호들 각각에 대해, 복수의 복호화 방식들 중 상기 신호를 복호화할 방식을 결정하는 단계; 상기 복수의 신호들을 상기 결정된 복호화 방식에 따라 복호화하는 단계; 및 상기 추출된 분할 정보를 이용하여, 상기 복호화된 복수의 신호들을 합성하는 단계를 포함하는 것을 특징으로 한다.

<7> 상술한 기술적 과제를 해결하기 위한 본 발명에 따른 복호화 장치는, 입력되는 비트스트림으로부터 부호화된 복수의 신호들과 상기 복수의 신호들의 분할 정보를 추출하는 비트언팩킹부; 상기 부호화된 복수의 신호들 각각에 대해, 복수의 복호화기들 중 상기 신호를 복호화할 복호화기를 결정하는 복호화기결정부; 상기 복수의 복호화기들을 포함하며, 상기 복수의 신호들을 상기 결정된 복호화기를 이용하여 복호화하는 복호화부; 및 상기 추출된 분할 정보를 이용하여, 상기 복호화된 복수의 신호들을 합성하는 합성부를 포함하는 것을 특징으로 한다.

<8> 상술한 기술적 과제를 해결하기 위한 본 발명에 따른 부호화 방법은, 입력되는 신호를 복수의 신호들로 분할하는 단계; 상기 분할된 복수의 신호들 각각을 상기 신호의 특성에 기초하여 복수의 클래스들 중 어느 하나로 분류하는 단계; 상기 분류된 클래스에 대응되는 부호화 방식을 이용하여, 상기 복수의 신호들을 부호화하는 단계; 및 상기 부호화된 복수의 신호를 이용하여 비트스트림을 생성하는 단계를 포함하는 것을 특징으로 한다.

<9> 상술한 기술적 과제를 해결하기 위한 본 발명에 따른 부호화 장치는, 입력되는 신호를 복수의 신호들로 분할하고, 상기 분할된 복수의 신호들 각각을 상기 신호의 특성에 기초하여 복수의 클래스들 중 어느 하나로 분류하는 분류부; 상기 복수의 클래스들에 대응되는 복수의 부호화기를 포함하며, 상기 복수의 신호들을 상기 분류된 클래스에 대응되는 부호화기를 이용하여 부호화하는 부호화부; 및 상기 부호화된 복수의 신호를 이용하여 비트스트림을 생성하는 비트팩킹부를 포함하는 것을 특징으로 한다.

<10> 유리한 효과

<11> 본 발명에 따른 부호화/복호화 장치 및 방법에 의하면, 특성에 따라 신호를 분류하여 그에 맞는 부호화기를 사용해 신호를 부호화함으로써, 서로 다른 특성을 가지는 신호들을 최적의 비트율로 부호화할 수 있다. 그에 따라 오디오 신호, 음성 신호 등의 여러 신호들을 모두 효율적으로 부호화할 수 있다.

산업상 이용 가능성

<224> 상기한 바와 같은 본 발명에 따른 부호화/복호화 방법 및 장치에 의하면, 특성에 따라 신호를 분류하여 그에 맞는 부호화기를 사용해 신호를 부호화함으로써, 서로 다른 특성을 가지는 신호들을 최적의 비트율로 부호화할 수 있다. 그에 따라 오디오 신호, 음성 신호 등의 여러 신호들을 모두 효율적으로 부호화할 수 있다.

도면의 간단한 설명

<12> 도 1은 본 발명에 따른 부호화 장치의 구성에 대한 제1 실시예를 나타내는 블록도이다.

<13> 도 2는 도 1에 도시된 분류부의 구성에 대한 제1 실시예를 나타내는 블록도이다.

<14> 도 3은 도 2에 도시된 전처리부의 구성에 대한 일실시예를 나타내는 블록도이다.

<15> 도 4는 입력되는 신호의 지각적 엔트로피(Perceptual Entrophy, PE)를 계산하는 장치의 구성에 대한 일실시예를 나타내는 블록도이다.

<16> 도 5는 도 1에 도시된 분류부의 구성에 대한 제2 실시예를 나타내는 블록도이다.

- <17> 도 6은 도 5에 도시된 신호분할부의 구성에 대한 제1 실시예를 나타내는 블록도이다.
- <18> 도 7 및 도 8은 복수의 신호들을 병합하는 방법에 대한 실시예들을 설명하기 위한 도면이다.
- <19> 도 9는 도 5에 도시된 신호분할부의 구성에 대한 제2 실시예를 나타내는 블록도이다.
- <20> 도 10은 입력 신호를 복수의 신호들로 분할하는 방법에 대한 일실시예를 나타내는 도면이다.
- <21> 도 11은 도 5에 도시된 결정부의 구성에 대한 일실시예를 나타내는 블록도이다.
- <22> 도 12는 도 1에 도시된 부호화부의 구성에 대한 제1 실시예를 나타내는 블록도이다.
- <23> 도 13은 도 1에 도시된 부호화부의 구성에 대한 제2 실시예를 나타내는 블록도이다.
- <24> 도 14는 본 발명에 따른 부호화 장치의 구성에 대한 제2 실시예를 나타내는 블록도이다.
- <25> 도 15는 본 발명에 따른 복호화 장치의 구성에 대한 일실시예를 나타내는 블록도이다.
- <26> 도 16은 도 15에 도시된 합성부의 구성에 대한 일실시예를 나타내는 블록도이다.
- <27> **발명의 실시를 위한 최선의 형태**
- <28> 이하, 본 발명의 바람직한 실시예를 첨부된 도면을 참조하여 본 발명에 따른 부호화/복호화 장치 및 방법에 관해 상세히 설명한다.
- <29> 도 1은 본 발명에 따른 부호화 장치의 구성에 대한 제1 실시예를 블록도로 도시한 것으로, 도시된 부호화 장치는 분류부(100), 부호화부(200) 및 비트패킹부(300)를 포함하여 이루어진다.
- <30> 도 1에 도시된 바와 같이, 본 발명에 따른 부호화 장치는 서로 다른 방식으로 부호화를 수행하는 복수의 부호화부들(210, 220)을 포함한다.
- <31> 분류부(100)는 입력되는 신호를 복수의 신호들로 분할한 후, 상기 분할된 복수의 신호들 각각을 복수의 부호화부들(210, 220) 중 어느 하나에 대응시킨다. 복수의 부호화부들(210, 220) 중 일부만이 상기 복수의 신호들에 대응될 수도 있으며, 복수의 부호화부들(210, 220) 중 어느 하나는 2 이상의 신호들에 대응될 수도 있다.
- <32> 분류부(100)는 상기 분할된 복수의 신호들 각각에 대해 부호화 비트수를 할당하거나, 부호화 순서를 결정할 수도 있다.
- <33> 부호화부(200)는 상기 분할된 복수의 신호들을 상기 분류부(100)에 의해 대응된 부호화부들을 이용해 부호화한다. 분류부(100)는 상기 복수의 신호들 각각의 특성을 분석하여, 상기 분석된 특성에 기반하여 복수의 부호화부들(210, 220) 중 상기 신호를 가장 효율적으로 부호화할 수 있는 부호화기를 선택한다.
- <34> 신호를 가장 효율적으로 부호화할 수 있는 부호화기는 상기 신호를 부호화했을 때 압축 효율이 가장 높은 부호화기를 의미할 수 있다.
- <35> 예를 들어, 상기 분할된 신호가 특정 계수와 잔차 신호로 모델링이 잘되는 신호인 경우 음성 부호화기를 이용하여 부호화하는 것이 효율적일 수 있으며, 상기 분할된 신호가 모델링이 잘되지 않는 신호인 경우에는 오디오 부호화기를 이용하여 부호화하는 것이 효율적일 수 있다.
- <36> 상기 분할된 신호를 특정 계수와 잔차 신호로 모델링했을 때, 상기 분할된 신호의 에너지 레벨에 대한 상기 잔차 신호의 에너지 레벨 비가 미리 설정된 기준값보다 작은 경우, 상기 분할된 신호가 모델링이 잘되는 신호라고 판단할 수 있다.
- <37> 상기 분할된 신호가 시간 축 상에서 높은 중복성을 가지는 경우, 상기 신호는 과거 신호로부터 현재 신호를 예측하는 선형 예측에 의해 모델링이 잘 될 수 있으므로, 선형 예측 부호화를 이용하는 음성 부호화기를 통해 부호화되는 것이 효율적이다.
- <38> 비트패킹부(300)는 상기 부호화된 신호 및 부호화와 관련된 부가 정보 등을 이용하여 전송할 비트스트림을 생성한다. 비트패킹부(300)는 비트-플레인(bit-plain) 방식이나 BSAC(Bit Sliced Arithmetic Coding) 방식 등을 사용하여 가변 비트율 특성을 가지는 비트스트림을 생성할 수 있다.
- <39> 비트율의 제한으로 부호화되지 않는 신호 또는 대역에 대해서는, 복호화 장치에서 복호화된 신호 또는 대역으로부터 interpolation, extrapolation, replication 등의 방법을 이용하여 재생될 수 있다. 또한, 상기 분할된

복수의 신호들 중 부호화되지 않은 신호에 대한 보상 정보는 전송되는 비트스트림에 포함될 수 있다.

- <40> 도 1을 참조하면, 분류부(100)는 복수의 분류부들(110, 120)을 포함하는 것이 바람직하다. 제1 분류부 내지 제n 분류부들(110, 120) 각각은 입력되는 신호를 복수의 신호들로 분할하거나, 입력되는 신호의 도메인을 변환하거나, 입력되는 신호의 특성을 추출하거나, 신호의 특성에 따라 입력되는 신호를 분류하거나, 입력되는 신호를 복수의 부호화부들(210, 220) 중 어느 하나에 대응시키는 등의 역할을 할 수 있다.
- <41> 제1 분류부 내지 제n 분류부들(110, 120) 중 어느 하나는 입력되는 신호에 전처리를 수행하여 부호화에 효율적인 신호로 변환하는 전처리부일 수 있다. 상기 전처리부는 입력되는 신호를 복수의 성분들, 예를 들어 계수 성분과 신호 성분으로 분할할 수 있으며, 다른 분류부들에 앞서 입력되는 신호에 대해 전처리를 수행하는 것이 바람직하다.
- <42> 상기 전처리는 입력되는 신호의 특성, 외부 환경 요인, 목표 비트율 등에 따라 선택적으로 사용되도록 할 수 있으며, 상기 분할된 복수의 신호들 중 일부의 신호에 대해 선택적으로 사용될 수도 있다.
- <43> 분류부(100)는 심리음향모델링부(400)로부터 입력받은 신호의 지각적 특성 정보, 예를 들어 마스킹 임계치, 신호 대 마스크 비(SMR, Signal-to-Mask Ratio), 지각적 엔트로피(Perceptual Entrophy) 등을 이용하여 입력되는 신호를 분류할 수 있다.
- <44> 즉, 분류부(100)는 심리음향모델링부(400)로부터 입력받은 신호의 지각적 특성 정보, 예를 들어 마스킹 임계치 및 신호 대 마스크 비(SMR, Signal-to-Mask Ratio) 등을 이용하여 입력되는 신호를 복수의 신호들로 분할하거나, 상기 분할된 복수의 신호들을 부호화부에 대응시킬 수 있다.
- <45> 또한, 분류부(100)는 신호의 톤 성분 포함 정도(tonality), ZCR(Zero Crossing Rate), 선형 예측 계수, 이전 프레임의 분류 정보 등을 입력받아, 상기 입력된 정보를 이용하여 입력되는 신호를 분류할 수 있다.
- <46> 도 1에 도시된 바와 같이, 부호화부(200)로부터 출력되는 부호화 결과에 대한 정보가 분류부(100)로 피드백될 수 있다.
- <47> 분류부(100)에 의해 입력되는 신호가 복수의 신호들로 분할되고, 상기 복수의 신호들 각각에 대한 부호화기, 부호화 비트수 또는 부호화 순서 등이 결정되면, 상기 결정된 부호화기 및 부호화 순서 등에 따라 상기 복수의 신호들에 대한 부호화가 진행된다. 부호화를 수행한 후, 상기 부호화에 실제 사용된 비트수는 분류부(100)가 할당한 부호화 비트수와 다를 수 있다.
- <48> 상기와 같은 할당된 부호화 비트수와 실제 사용된 비트수의 차이에 대한 정보가 분류부(100)로 피드백되어, 실제 사용된 비트수가 작은 경우 분류부(100)가 다른 신호에 대해 할당된 비트수를 증가시키도록 할 수 있으며, 실제 사용된 비트수가 많은 경우 분류부(100)가 다른 신호에 대해 할당된 비트수를 감소시키도록 할 수 있다.
- <49> 분류부(100)가 분류된 신호에 대해 대응시킨 부호화기와 실제 부호화를 수행한 부호화기가 다를 수 있으므로, 그러한 경우 부호화기의 변경 정보가 분류부(100)로 피드백되어 분할된 신호에 부호화기를 대응시키는 작업이 다시 수행되도록 할 수 있다.
- <50> 또한, 상기 피드백된 부호화 결과 정보를 이용하여 분류부(100)는 입력 신호를 복수의 신호들로 재분할할 수도 있으며, 그러한 경우 분류부(100)는 종전에 분할된 복수의 신호들과 다른 분할 구조를 가진 복수의 신호들로 상기 입력 신호를 분할할 수 있다.
- <51> 상기에서 설명한 것 이외에도, 분류부(100)에 의해 결정된 부호화 관련 정보가 실제 부호화 과정에서 수행된 것과 상이한 경우, 상기 상이한 내용을 분류부(100)로 피드백하여 상기 부호화 관련 정보가 재결정되도록 할 수 있다.
- <52> 도 2는 도 1에 도시된 분류부의 구성에 대한 제1 실시예를 블록도로 도시한 것으로, 도시된 바와 같이 제1 분류부(110)는 입력되는 신호를 부호화에 효율적인 신호의 형태로 변환하기 위한 전처리를 수행하는 전처리부일 수 있다.
- <53> 도 2를 참조하면, 제1 분류부(110)는 서로 다른 방식의 전처리를 수행하는 복수의 전처리부들(111, 112)을 포함할 수 있으며, 입력되는 신호의 특성, 외부 환경 요인, 목표 비트율 등에 따라 제1 전처리부 내지 제n 전처리부(111, 112) 중 어느 하나를 이용하여 입력되는 신호에 대해 전처리를 수행할 수 있다. 또한, 제1 분류부(110)는 복수의 전처리부를 이용하여 입력되는 신호에 대해 2 이상의 전처리를 수행할 수도 있다.

<54> 도 3은 도 2에 도시된 전처리부의 구성에 대한 일실시예를 블록도로 도시한 것으로, 도시된 전처리부는 계수 추출부(113) 및 잔차신호추출부(114)를 포함하여 이루어진다.

<55> 계수추출부(113)는 입력되는 신호를 분석을 수행하여 상기 신호의 특성을 나타내는 계수를 추출하고, 잔차신호추출부(114) 상기 추출된 계수를 이용하여 입력신호로부터 중복 성분이 제거된 잔차 신호를 추출한다.

<56> 상기 전처리부는 입력되는 신호에 대해 선형 예측 부호화를 수행할 수 있으며, 이 경우 계수추출부(113)는 입력되는 신호에 대해 선형 예측 분석을 수행하여 선형 예측 계수를 추출하고, 잔차신호추출부(114)는 상기 추출된 선형 예측 계수를 이용하여 상기 입력되는 신호로부터 잔차 신호(residual signal)를 추출한다. 상기 잔차 신호는 중복성이 제거되었기 때문에 백색 잡음과 같은 형태를 가질 수 있다.

<57> 이하에서는, 본 발명에 따른 선형 예측 분석 방법에 대해 상세히 설명하기로 한다.

<58> 선형 예측 분석에 의해 추정된 신호는 다음의 수학적 식 1과 같이 과거 입력 신호의 선형 조합으로 이루어질 수 있다.

<59> 수학적 식 1

$$\hat{x}(n) = \sum_{j=1}^p \alpha_j x(n-j)$$

<60>

<61> 상기 수학적 식 1에서 p 는 선형 예측 차수이고, α_1 내지 α_p 는 선형 예측 계수로서 입력 신호와 추정 신호의 mean square error(MSE)를 최소화하는 과정을 통하여 구해진다.

<62> 선형 예측 분석을 위한 전달함수는 다음의 수학적 식 2와 같이 표현될 수 있다.

<63> 수학적 식 2

$$P(z) = \sum_{k=1}^p \alpha_k z^{-k}$$

<64>

<65> 전처리부는 또 다른 선형 예측 분석 방법인 Warped linear prediction coding(이하, WLPC라고 한다)을 이용하여 입력되는 신호로부터 선형 예측 계수와 잔차 신호를 추출할 수 있다. 상기 WLPC는 유닛 딜레이(unit delay)인 z^{-1} 를 다음의 수학적 식 3과 같은 전달 함수를 가지는 올-패스 필터(all-pass filter)로 대체함으로써 구현될 수 있다.

<66> 수학적 식 3

$$A(z) = \left[\frac{z^{-1} - \lambda}{1 - \lambda z^{-1}} \right]$$

<67>

<68> 상기 수학적 식 3에서 λ 는 올-패스 계수이다. 상기 올-패스 계수 λ 를 변화시킴에 따라 분석하고자 하는 신호의 해상도를 변화시킬 수 있다. 따라서, 일부 주파수 대역에 신호 집중도가 높은 신호의 경우, 예를 들어 저주파 대역에 신호의 집중도가 높은 오디오 신호에 대해, 저주파 대역 신호의 해상도를 높이도록 올-패스 계수 λ 를 설정함으로써 신호를 효율적으로 부호화할 수 있다.

<69> WLPC는 고주파수 영역의 신호보다 저주파수 영역의 신호를 더 높은 해상도를 가지고 분석하여, 저주파수 영역 신호에 대해 높은 예측 성능을 보인다. 그에 따라 WLPC는 저주파수 영역 신호를 더 잘 모델할 수 있다.

<70> 입력되는 신호의 특성, 외부 환경 요인, 목표 비트율 등에 따라, 올-패스 계수 λ 를 시간 축 상에서 변화시킬 수 있다. 다만, 올-패스 계수 λ 가 시간에 따라 변화하는 경우, 복호화된 오디오 신호에서 큰 왜곡이 발생할 수 있다. 따라서 올-패스 계수 λ 가 변화하는 시점에서 스무딩(smoothing) 기법 등을 적용하여, 올-패스 계수 λ 가 연속적으로 변화하도록 하여 왜곡을 최소화할 수 있다. 바람직하게는, 시간 축 상에서 현재의 올-패스 계수 λ 로 사용할 수 있는 값의 범위를 이전의 올-패스 계수 λ 값에 의해 결정하도록 할 수 있다.

- <71> 선형 예측 계수 추정을 위한 입력으로 원본 신호가 아닌 심리음향모델의 마스킹 임계치를 사용할 수 있다. 즉, 마스킹 임계치를 시간 영역 신호로 변환한 후, 상기 변환된 신호를 입력으로 하는 WLPC를 수행할 수 있다. 또한, 잔차 신호를 입력으로 한 차례 더 선형 예측 계수를 추정할 수 있다. 즉, 선형 예측 분석을 복수회 수행함으로써, 보다 화이트닝(whitening)된 잔차 신호를 얻을 수 있다.
- <72> 도 2에 도시된 제1 분류부(110)는 상기 수학적식 1 및 수학적식 2를 참조하여 설명한 선형 예측 분석을 수행하는 제1 전처리부(111)와 상기 WLP를 수행하는 제2 전처리부(미도시)를 포함하고, 입력되는 신호의 특성, 외부 환경 요인, 목표 비트율 등에 따라 제1, 2 전처리부 중 하나를 선택하거나, 입력 신호에 대해 선형 예측 분석을 수행하지 않도록 결정할 수 있다.
- <73> 상기 WLP를 수행하는 제2 전처리부(미도시)에서 λ 가 0인 경우는 제1 전처리부(111)와 동일하므로, 제1 분류부(110)는 상기 제2 전처리부(미도시)만을 포함하고, 상기 λ 값을 이용하여 상기 두 선형 예측 부호화 방법들 중 하나를 선택할 수도 있다. 또한, 상기 선형 예측 분석의 수행 여부 결정 및 선형 예측 부호화 방식의 선택은 프레임 단위로 수행될 수 있다.
- <74> 상기 선형 예측 분석의 사용 여부 및 선택된 선형 예측 부호화기에 대한 정보는 전송되는 비트스트림에 포함될 수 있다.
- <75> 비트팩킹부(300)는 제1 분류부(110)로부터 선형 예측 계수, 선형 예측 부호화 사용 여부 및 사용된 선형 예측 부호화기 정보 등을 입력받아 전송할 비트스트림에 포함시킬 수 있다.
- <76> 입력되는 신호를 주파수 영역에서 지각적으로 구별되지 않는 음질로 부호화하기 위한 비트수는 상기 신호의 지각적 엔트로피(Perceptual Entropy, PE)를 계산함에 의해 구해질 수 있다.
- <77> 도 4는 상기 지각적 엔트로피(Perceptual Entropy, PE)를 계산하는 장치의 구성에 대한 일실시예를 블록도로 도시한 것으로, 도시된 장치는 필터뱅크(115), 선형예측부(1160), 심리음향모델링부(117), 제1 비트계산부(118) 및 제2 비트계산부(119)를 포함하여 이루어진다.
- <78> 상기 지각적 엔트로피(PE)는 다음의 수학적식 4를 이용하여 계산될 수 있다.
- <79> 수학적식 4
- $$PE = \frac{1}{2\pi} \int_0^{\pi} \max \left[0, \log_2 \frac{X(e^{jw})}{T(e^{jw})} \right] dw \quad (bit/sample)$$
- <80>
- <81> 상기 수학적식 4에서, $X(e^{jw})$ 는 원본 신호의 에너지를 의미하며, $T(e^{jw})$ 는 마스킹 임계치(masking threshold)를 의미한다.
- <82> 올-패스 필터를 이용하는 WLPC의 경우에 있어 지각적 엔트로피(WPE)는 잔차 신호의 에너지와 잔차 신호의 마스킹 임계치(masking threshold)의 비를 이용하여 계산될 수 있다. 따라서 WLPC를 이용하는 부호화 장치에 있어 지각적 엔트로피(WPE)는 다음의 수학적식 5를 이용하여 계산될 수 있다.
- <83> 수학적식 5
- $$PE = \frac{1}{2\pi} \int_0^{\pi} \max \left[0, \log_2 \frac{R(e^{jw})}{T'(e^{jw})} \right] dw \quad (bit/sample)$$
- <84>
- <85> 상기 수학적식 5에서, $R(e^{jw})$ 는 잔차 신호의 에너지를 의미하며, $T'(e^{jw})$ 는 잔차 신호의 마스킹 임계치(masking threshold)를 의미한다.
- <86> 또한, 상기 잔차 신호의 마스킹 임계치 $T'(e^{jw})$ 는 다음의 수학적식 6과 같이 표현될 수 있다.

<87> 수학식 6

$$T'(e^{jw}) = T(e^{jw}) / |H(e^{jw})|^2$$

<88>

<89> 상기 수학식 6에서, $T(e^{jw})$ 는 원본 신호의 마스크 임계치를 의미하며, $H(e^{jw})$ 는 WLPC의 전달 함수를 의미한다. 심리음향모델링부(320)가 스케일팩터 밴드(scalefactor band) 도메인에서 원본 신호의 마스크 임계치와 WLPC의 전달 함수를 이용하여 상기 잔차 신호의 마스크 임계치($T'(e^{jw})$)를 계산할 수 있다.

<90> 도 4를 참조하면, 제1 비트계산부(118)는 선형예측부(116)에서 출력되는 WLPC가 수행된 잔차 신호와 심리음향모델링부(117)로부터 출력되는 마스크 임계치를 입력받는다. 또한, 원본 신호는 필터뱅크(115)를 통해 주파수 변환된 후 심리음향모델링부(117) 및 제2 비트계산부(119)로 입력되는 것이 바람직하다. 필터뱅크(115)는 상기 원본 신호에 대해 푸리에 변환 등을 수행할 수 있다.

<91> 제1 비트계산부(118)는 원본 신호의 마스크 임계치를 WLPC 합성 필터의 전달함수 스펙트럼으로 나눈 값과 잔차 신호 에너지의 비를 이용하여 WPE를 계산한다.

<92> 제2 비트계산부(119)는 필터뱅크(115)로부터 입력되는 원본 신호의 에너지와 심리음향모델링부(117)로부터 입력되는 마스크 임계치의 비를 이용하여 PE를 계산한다.

<93> 60개 이상의 동일하지 않은 대역폭을 가지는 밴드들(non-uniform partition bands)로 분할된 신호에 대해서는, WLPC를 이용하는 경우의 지각적 엔트로피(Perceptual Entropy, WPE)는 다음의 수학식 7과 같이 계산될 수 있다.

<94> 수학식 7

$$WPE = - \sum_{b=1}^{b_{max}} (w_{high}(b) - w_{low}(b)) \cdot \log_{10} \left(\frac{nb_{res}(b)}{e_{res}(b)} \right)$$

$$e_{res}(b) = \sum_{w=w_{low}(b)}^{w_{high}(b)} res(w)^2 \quad nb_{res}(b) = \sum_{w=w_{low}(b)}^{w_{high}(b)} \frac{nb_{linear}(w)}{h(w)^2}$$

<95>

<96> 상기 수학식 7에서, b 는 심리음향모델에서 나뉘어진 파티션 밴드의 인덱스를 나타내며, $e_{res}(b)$ 는 파티션 밴드에서의 잔차 신호의 에너지 합을 나타내고, $w_{low}(b)$ 및 $w_{high}(b)$ 는 각각 파티션 밴드에서의 가장 낮은 주파수와 가장 높은 주파수를 의미한다. 또한, $nb_{linear}(w)$ 는 선형적으로 맵핑된 파티션 밴드에서의 마스크 임계치를 나타내며, $h(w)^2$ 는 해당 프레임에서의 LPC 에너지 스펙트럼(energy spectrum)을 나타낸다. $nb_{res}(w)$ 는 잔차 신호에 해당하는 선형적인 마스크 임계치를 나타낸다.

<97> 동일한 대역폭을 가지는 복수의 서브밴드(subband)로 분할된 신호에 대해서는 60개 이상의 동일하지 않은 대역폭을 가지는 밴드들(non-uniform partition bands)로 분할된 신호에 대해서는, WLPC를 이용하는 경우의 지각적 엔트로피(Perceptual Entropy, WPE)는 다음의 수학식 8과 같이 계산될 수 있다.

<98> 수학적식 8

$$nb_{sub}(s) = \min_{s_{low}(s) < w < s_{high}(s)} \left(\frac{nb_{linear}(w)}{h(w)^2} \right)$$

$$WPE_{sub} = - \sum_{s=1}^{smax} (s_{high}(s) - s_{low}(s)) \cdot \log_{10} \left(\frac{nb_{sub}(s)}{e_{sub}(s)} \right)$$

$$e_{sub}(s) = \sum_{w=s_{low}(s)}^{s_{high}(s)} res(w)^2$$

<99>

<100> 상기 수학적식 8에서, s는 선형적으로 분할된 서브 밴드의 인덱스를 나타내며, s_{low}(w)와 s_{high}(w)는 서브 밴드에서 가장 낮은 주파수와 가장 높은 주파수를 각각 나타낸다. nb_{sub}(s)는 서브 밴드에서의 마스킹 임계치(masking threshold)를 나타내며, 상기 수학적식 8과 같이 해당 서브 밴드에서의 마스킹 임계치들 중 가장 최소값을 취한다. e_{sub}(s)는 서브 밴드의 에너지를 나타내는 것으로, 가장 낮은 주파수부터 가장 높은 주파수까지의 값을 합한 것이다.

<101> 상기와 같은 동일한 대역폭을 가지는 밴드들에서, 임계치(threshold)가 입력 스펙트럼의 합보다 큰 밴드에 대해서는 지각적 엔트로피가 계산되지 않으므로, 저주파수 대역에서 해상도가 높은 수학적식 7의 경우보다 WPE가 낮게 계산될 수 있다.

<102> 동일하지 않은 대역폭을 가지는 스케일팩터 밴드에서, WLPC를 이용하는 경우의 지각적 엔트로피(Perceptual Entropy, WPE)는 다음의 수학적식 9와 같이 계산될 수 있다.

<103> 수학적식 9

$$nb_{sf}(s) = \min_{sf_{low}(s) < w < sf_{high}(s)} \left(\frac{nb_{linear}(w)}{h(w)^2} \right)$$

$$WPE_{sf} = - \sum_{f=1}^{fmax} (s_{high}(f) - s_{low}(f)) \cdot \log_{10} \left(\frac{nb_{sf}(f)}{e_{sf}(f)} \right)$$

$$e_{sf}(s) = \sum_{w=sf_{low}(s)}^{sf_{high}(s)} res(w)^2$$

<104> 상기 수학적식 9에서, f는 스케일 팩터(scale factor) 밴드의 인덱스를 나타내며, nb_{sf}(f)는 스케일 팩터 밴드에서 가장 작은 값의 마스킹 임계치를 나타냅니다. 또한, WPE_{sf}는 모든 스케일 팩터 밴드의 입력 신호와 마스킹 임계치의 비를 나타내며, e_{sf}(s)는 스케일 팩터의 가장 낮은 주파수부터 가장 높은 주파수의 값을 합하여 해당 서브 밴드의 에너지를 나타낸 것이다.

<106> 도 5는 도 1에 도시된 분류부의 구성에 대한 제2 실시예를 블록도로 도시한 것으로, 도시된 분류부는 신호분할부(121) 및 결정부(122)를 포함하여 이루어진다.

<107> 도 5를 참조하면, 신호분할부(121)는 입력되는 신호를 복수의 신호들로 분할한다. 예를 들어, 신호분할부(121)는 서브밴드 필터(subband filter)를 사용하여 입력되는 신호를 복수의 주파수 대역으로 분할할 수 있다. 상기 분할되는 주파수 대역들의 대역폭은 서로 동일하거나 상이할 수 있다. 상기한 바와 같이, 상기 분할된 복수의

신호들 각각은 상기 신호의 특성에 맞는 부호화부에 의해 독립적으로 부호화되는 것이 바람직하다.

- <108> 신호분할부(121)는 분할되는 복수의 신호들, 예를 들어 복수의 주파수 대역신호들 간의 간섭이 최소화될 수 있도록, 입력되는 신호를 분할하는 것이 바람직하다. 또한, 신호분할부(121)는 분할된 복수의 신호들을 한번 더 분할하는 이중 필터뱅크 구조를 가질 수도 있다.
- <109> 상기 분할된 복수의 신호들에 대한 정보, 예를 들어 분할된 신호의 개수, 분할된 주파수 대역 정보 등의 분할 정보는 전송되는 비트스트림에 포함될 수 있으며, 복호화 장치에서는 상기 분할 정보를 이용하여 독립적으로 복호화된 복수의 신호들을 합성하여 원 신호를 복원할 수 있다.
- <110> 상기 분할 정보는 하나의 테이블(table)로 구성될 수 있으며, 비트스트림에는 미리 설정된 복수의 분할 정보 테이블들 중 상기 신호 분할에 사용된 테이블에 대한 식별 정보가 포함될 수 있다.
- <111> 또한, 상기 분할된 복수의 신호들, 예를 들어 복수의 주파수 대역 신호들 각각에 대해 음질에 미치는 중요도 정하고, 상기 정해진 중요도에 따라 상기 신호들 각각의 비트율을 조절할 수 있다. 이때, 상기 복수의 신호들 각각의 중요도는 고정된 값이 사용되거나, 각 프레임별로 입력되는 신호의 특성에 따라 가변적으로 결정될 수 있다.
- <112> 또한, 입력되는 신호에 음성 신호와 오디오 신호가 섞여있는 경우, 신호분할부(121)는 음성 신호의 특성과 오디오 신호의 특성을 고려하여 상기 입력되는 신호를 음성 신호와 오디오 신호로 분할할 수도 있다.
- <113> 결정부(122)는 상기 분할된 복수의 신호들 각각에 대해, 상기 신호를 가장 효율적으로 부호화할 수 있는 부호화부를 복수의 부호화부들(210, 220) 중에서 결정한다. 예를 들어, 부호화부(200)에 m개의 부호화부가 포함된 경우, 결정부(122)는 상기 분할된 신호가 상기 m개의 부호화부들 중 어느 부호화부에 의해 가장 효율적으로 부호화될 수 있는지 여부를 결정할 수 있다.
- <114> 결정부(122)는 상기 분할된 신호들 각각을 그 특성에 따라 분류한다. 예를 들어, 결정부(122)는 상기 입력되는 신호를 미리 설정된 N개의 클래스 중 어느 하나로 분류하고, 상기 N개의 클래스들은 복수의 부호화부들(210, 220)과 일대일 대응될 수 있다. 결정부(122)는 상기 분할된 신호들 각각을 복수의 클래스들 중 어느 하나로 분류함에 따라, 상기 신호를 부호화할 부호화부를 결정할 수 있다.
- <115> 부호화부(200)에 m개의 부호화부가 포함된 경우, 결정부(122)는 상기 복수의 신호들 각각을 제1 부호화부에 의해 가장 효율적으로 부호화될 수 있는 제1 클래스, 제2 부호화부에 의해 가장 효율적으로 부호화될 수 있는 제2 클래스, ..., 제m 부호화부에 의해 가장 효율적으로 부호화될 수 있는 제m 클래스 중 어느 하나의 클래스로 분류할 수 있다.
- <116> 상기 m개의 부호화부들 각각이 가장 효율적으로 부호화할 수 있는 신호들의 특성을 미리 파악하여, 상기 파악된 특성들을 상기 제1 내지 제m 클래스 각각의 특성으로 설정할 수 있다. 결정부(122)는 상기 복수의 신호들 각각에 대해 특성을 추출한 후, 상기 m개의 클래스들 중 상기 추출한 특성을 가지는 클래스로 상기 신호를 분류할 수 있다.
- <117> 신호가 분류되는 복수의 클래스들에 대한 일실시예로, 음성 음성 클래스, 무성 음성 클래스, 백그라운드 잡음 클래스, 묵음 클래스, 톤(tonal)이 있는 오디오 클래스, 톤이 없는 오디오 클래스, 음성 음성과 오디오가 혼합된 클래스 등이 있을 수 있다.
- <118> 결정부(122)는 심리음향모델링부(400)로부터 신호의 지각적 특성, 예를 들어 마스킹 임계치, 신호 대 마스크 비 또는 지각적 엔트로피 등을 이용하여, 상기 분할된 복수의 신호들 각각을 부호화할 부호화부를 결정할 수 있다.
- <119> 결정부(122)는 상기 신호의 지각적 특성 정보를 이용하여 상기 복수의 신호들 각각에 할당되는 부호화 비트수 또는 상기 복수의 신호들의 부호화 순서 등을 결정할 수 있다.
- <120> 결정부(122)에 의해 결정된 정보들, 예를 들어 상기 분할된 신호들 각각에 대해 결정된 부호화기 정보, 부호화 비트수 정보, 부호화 순서 정보 등은 전송되는 비트스트림에 포함되는 것이 바람직하다.
- <121> 도 6은 도시된 신호분할부의 구성에 대한 제1 실시예를 블록도로 도시한 것으로, 도시된 신호분할부는 분할부(123) 및 병합부(124)를 포함하여 이루어진다.
- <122> 도 6을 참조하면, 분할부(123)는 입력되는 신호를 복수의 신호들로 분할하고, 병합부(124)는 상기 분할된 신호들 중 유사한 특성을 가지는 신호들을 하나의 신호로 병합할 수 있다. 이를 위해, 병합부(124)는 합성 필터뱅크

(synthesis filter bank)를 포함할 수 있다.

- <123> 예를 들어, 분할부(123)는 입력되는 신호를 256개의 주파수 대역으로 분할하고, 병합부(124)는 상기 분할된 주파수 대역들 중 유사한 특성을 가지는 대역들을 하나의 대역으로 병합할 수 있다.
- <124> 도 7에 도시된 바와 같이, 병합부(124)는 분할된 복수의 신호들 중 이웃한 신호들을 하나의 신호로 병합할 수 있다. 이 경우, 병합부(124)는 신호의 특성에 관계없이, 상기 복수의 신호들 중 이웃한 신호들을 정해진 규칙에 따라 병합할 수도 있다.
- <125> 또한, 도 8에 도시된 바와 같이, 병합부(124)는 분할된 복수의 신호들 중 이웃한 신호들 뿐 아니라, 유사한 특성을 가지는 서로 이웃하지 않는 복수의 신호들을 하나의 신호로 병합할 수도 있다. 이 때, 병합부(124)는 동일한 부호화기를 이용하여 효율적으로 부호화될 수 있는 복수의 신호들을 하나의 신호로 병합하는 것이 바람직하다.
- <126> 도 9는 도 5에 도시된 신호분할부의 구성에 대한 제2 실시예를 블록도로 도시한 것으로, 도시된 신호분할부는 제1 분할부(125), 제2 분할부(126) 및 제3 분할부(127)를 포함하여 이루어진다.
- <127> 도 9를 참조하면, 신호분할부(121)는 입력되는 신호를 계층적으로 분할할 수 있다. 예를 들어, 입력되는 신호는 제1 분할부(125)를 통해 2개의 신호로 분할되고, 상기 분할된 두 신호 중 하나는 제2 분할부(126)를 통해 3개의 신호로 분할되며, 상기 분할된 3개의 신호 중 하나는 제3 분할부(127)를 통해 3개의 신호로 분할되어, 상기 입력 신호는 6개의 신호로 분할될 수 있다. 입력 신호를 복수의 주파수 대역으로 분할하는 경우, 상기과 같은 계층적 분할에 의해 분할되는 주파수 대역들의 대역폭을 상이하게 할 수 있다.
- <128> 상기에서는 입력되는 신호를 3 계층으로 나누어 분할하는 것으로 계층 분할 방법을 설명하였으나, 3 계층 이외에 2 계층 또는 4 이상의 계층으로 나누어 상기 입력 신호를 복수의 신호들로 분할하는 것도 가능하다.
- <129> 또한, 신호분할부(121)에 포함된 복수의 분할부들 중 어느 하나는 입력되는 신호를 복수의 시간 대역 신호들로 분할할 수도 있다.
- <130> 도 10은 신호분할부(121)가 입력 신호를 복수의 신호들로 분할하는 방법에 대한 일실시예를 설명하기 위해 도시한 것이다.
- <131> 음성 또는 오디오 신호의 경우 짧은 프레임 길이 동안은 변하지 않는(stationary) 특성을 가지는 것이 일반적이지만, 변화(transition) 구간 등에서는 변화하는(non-stationary) 특성을 가질 수 있다.
- <132> 상기과 같이 변화하는(non-stationary) 특성을 가지는 신호를 효과적으로 분석하여 부호화 효율을 높이기 위해, 본 발명에 따른 부호화 장치는 wavelet 또는 EMD(Empirical Mode Decomposition) 등의 방법을 사용하여 입력되는 신호를 효과적으로 분석하는 것이 바람직하다. 즉, 본 발명에 따른 부호화 장치는 입력되는 신호를 고정되지 않은 변환 함수를 사용해 분할하여 특성을 분석할 수 있다. 예를 들어, 신호분할부(121)는 주파수 대역이 고정되지 않은 서브밴드 필터링(subband filtering) 방법을 사용하여 입력되는 신호를 가변적인 대역폭을 가지는 복수의 주파수 대역으로 분할할 수 있다.
- <133> 이하에서는, EMD를 이용하여 입력되는 신호를 복수의 신호들로 분할하는 방법에 대해 설명하기로 한다.
- <134> EMD는 입력되는 신호 $s(t)$ 를 IMF(Intrinsic Mode Function) $C_m(t)$ 으로 분할한다. 상기 IMF들 각각은 극값(extrema)의 개수와 zero-crossing의 개수는 반드시 같거나 최대 1개의 차이가 나야 하며, 최대값들로 정의되는 포락선과 최소값들로 정의되는 포락선의 평균값이 0이 되어야 하는 조건을 만족하여야 한다.
- <135> 상기과 같은 조건을 만족하는 IMF는 하모닉(harmonic) 함수처럼 단순하게 진동하는 형태를 가질 수 있으므로, EMD를 이용하여 입력되는 신호를 그 특성에 따라 효과적으로 분할할 수 있다.
- <136> 상기 IMF를 추출하기 위한 방법은 다음과 같다. 우선, 입력 신호의 최대값들로 정의되는 함수의 극값들을 cubic spline interpolation을 이용해 연결하여 상부 포락선을 구성하고, 입력 신호의 최소값들로 정의되는 함수의 극값들을 cubic spline interpolation을 이용해 연결하여 하부 포락선을 구성한다.
- <137> 입력되는 신호가 가질 수 있는 값들은 상기 구성된 상부 포락선과 하부 포락선 사이에 포함될 수 있다.
- <138> 다음으로, 상기 상부 포락선과 하부 포락선의 평균 $m(t)$ 를 구하고, 다음의 수학적 식 10과 같이 계산하여 상기 구해진 평균 $m(t)$ 를 입력 신호 $s(t)$ 로부터 제거시켜 첫 번째 성분 $h_1(t)$ 를 구한다.

<139> 수학식 10

$$s(t) - m_1(t) = h_1(t)$$

<140>

<141> 상기 수학식 10에 의해 구해진 $h_1(t)$ 가 상기에서 설명한 IMF 조건을 만족하지 못하면, $h_1(t)$ 가 입력 신호로 간주되어 상기에서 설명한 과정을 반복한다.

<142> 상기와 같은 반복을 통해 상기 IMF 조건을 만족하는 첫번째 신호 $C_1(t)$ 가 구해지면, 다음의 수학식 11과 같이 입력 신호 $s(t)$ 에서 $C_1(t)$ 를 제거하여 여분의 신호인 $r_1(t)$ 를 구한다.

<143> 수학식 11

$$s(t) - c_1(t) = r_1(t)$$

<144>

<145> 다음에는, 상기 여분 신호 $r_1(t)$ 가 새로운 입력 신호로 간주되어, 상기한 바와 같은 과정을 반복해 두 번째 IMF $c_2(t)$ 와 여분 신호 $r_1(t)$ 를 구한다.

<146> 상기와 같은 과정을 반복하여 IMF들과 여분 신호들을 구하는 중, 여분 신호 $r_n(t)$ 가 상수이거나, 단조 증가 함수 또는 하나의 주기를 가지는 함수로 극값이 한개 또는 존재하지 않게 되는 경우 IMF 추출 과정을 종료한다.

<147> 상기와 같은 과정을 통해 입력 신호 $s(t)$ 는 다음의 수학식 12와 같이 복수의 IMF들과 하나의 여분 신호 $r_m(t)$ 의 합으로 표현될 수 있다.

<148> 수학식 12

$$s(t) = \sum_{m=0}^M C_m(t) + r_m(t)$$

<149>

<150> 상기 수학식 12에서, M은 추출된 IMF들의 개수이며, $r_m(t)$ 는 최종으로 남은 여분 신호이다. 입력 신호의 전체적인 변화 특성은 상기 최종 여분 신호 $r_m(t)$ 에 나타날 수 있다.

<151> 도 10은 EMD를 이용하여 입력신호를 복수의 신호들로 분할한 결과에 대한 예를 도시한 것으로, 입력신호(Original Input Signal)로부터 추출된 11개의 IMF들과 최종 여분 신호(Residue)를 나타낸 것이다.

<152> 도 10을 참조하면, EMD를 이용해 분할된 복수의 IMF들의 주파수는 추출된 순서에 반비례하여 감소한다.

<153> 또한, 다음의 수학식 13과 같은 여분 신호의 표준 편차 SD(standard deviation)를 이용하여 IMF를 추출하는 과정을 단순화할 수 있다.

<154> 수학식 13

$$SD = \sum_{t=0}^T \left[\frac{|h_{l(k-1)}(t) - h_{lk}(t)|^2}{h_{l(k-1)}^2(t)} \right]$$

<155>

<156> 상기 수학식 13에서, $h_{l(k-1)}$ 는 이전 과정에서 구해진 여분 신호이며, h_{lk} 는 현재 과정에서 구해진 여분 신호를 의미한다. 이전 과정에서 구한 여분 신호와 현재 과정에서 구해진 여분 신호의 차이가 기준 값 이하인 경우, 예를 들어 상기 두 여분 신호의 표준 편차 SD가 03 이하인 경우, 상기 현재 과정에서 구해진 여분 신호 h_{lk} 를 IMF로 간주할 수 있다.

<157> 다음의 수학식 14와 같이 표현되는 힐버트 변환(Hilbert Transform)을 이용하면, 신호 $x(t)$ 를 분석이 용이한 신호(analytic signal)로 변환할 수 있다.

<158> 수학식 14

$$z(t) = x(t) + jH\{x(t)\} = a(t)e^{j\theta(t)}$$

<159>

<160> 상기 수학식 14에서, $a(t)$ 는 instantaneous amplitude를 의미하며, $\theta(t)$ 는 instantaneous phase를 의미하고, $H[\cdot]$ 는 힐버트 변환을 의미한다.

<161> 힐버트 변환(Hilbert transform)은 입력된 신호를 imaginary 신호로 변환하며, 상기 입력된 신호와 변환된 신호를 이용하여 분석이 용이한 신호(analytic signal)을 만들 수 있다.

<162> 평균이 0인 신호에 상기와 같은 힐버트 변환(Hilbert transform)을 적용하면, 시간 영역과 주파수 영역 모두에 대해 높은 해상도의 주파수 성분들을 얻을 수 있다.

<163> 이하에서는, 도 5에 도시된 결정부(122)가 상기 분할된 복수의 신호들 각각에 대해 상기 신호를 부호화할 부호화기를 결정하는 방법에 대한 실시예를 설명하기로 한다.

<164> 결정부(122)는 상기 복수의 신호들 각각에 대해 음성 부호화기와 오디오 부호화기 중 어느 부호화기를 이용하여 상기 신호를 더욱 효율적으로 부호화할 수 있는지 여부를 결정할 수 있다. 즉, 결정부(122)는 상기 분할된 신호가 음성 부호화기를 이용해 더욱 효율적으로 부호화될 수 있는 신호인 경우 상기 신호를 복수의 부호화부들(210, 220) 중 음성 부호화기를 이용해 부호화하는 것으로 결정하고, 상기 분할된 신호가 오디오 부호화기를 이용해 더욱 효율적으로 부호화될 수 있는 신호인 경우 복수의 부호화부들(210, 220) 중 오디오 부호화기를 이용해 상기 신호를 부호화하는 것으로 결정할 수 있다.

<165> 이하에서는, 결정부(122)가 음성 부호화기와 오디오 부호화기 중 어느 부호화기를 이용하여 신호를 더욱 효율적으로 부호화할 수 있는지 여부를 결정하는 방법에 대한 실시예들을 설명하기로 한다.

<166> 결정부(122)는 상기 분할된 신호의 변화량을 측정하여, 상기 측정된 변화량이 미리 설정된 기준 값보다 큰 경우 상기 신호를 음성 부호화기로 부호화하는 것으로 결정할 수도 있다.

<167> 또한, 결정부(122)는 상기 분할된 신호의 특정 영역에 포함된 토널(tonal) 성분을 측정하여, 상기 측정된 특정 영역의 토널 성분이 미리 설정된 기준 값보다 강하게 나타내는 경우 상기 신호를 오디오 부호화기를 이용하여 부호화하는 것으로 결정할 수 있다.

<168> 도 11은 도 5에 도시된 결정부(122)의 구성에 대한 실시예를 블록도로 도시한 것으로, 도시된 결정부는 음성 부호화/복호화부(500), 제1 필터뱅크(510), 제2 필터뱅크(520), 판단부(530) 및 심리음향모델링부(540)를 포함하여 이루어진다.

<169> 도 11에 도시된 바와 같은 구조를 가지는 결정부(122)를 이용하여, 음성 부호화기와 오디오 부호화기 중 어느 부호화기를 이용하여 상기 분할된 신호를 더욱 효율적으로 부호화할 수 있는지 여부를 결정할 수 있다.

<170> 도 11을 참조하면, 입력되는 신호는 음성 부호화/복호화부(500)에 포함된 음성 부호화기를 통해 부호화된 후, 다시 음성 복호화기를 통해 복호화되어 복원된다. 음성 부호화/복호화부(500)는 AMR-WB speech coder/decoder를 포함할 수 있으며, 상기 AMR-WB speech coder/decoder는 CELP(Code-Excited Linear Predictive) 구조를 사용하여 구성될 수 있다.

<171> 상기 입력되는 신호는 다운 샘플링(downsampling)된 후 음성 부호화/복호화부(500)로 입력될 수 있으며, 음성 부호화/복호화부(500)로부터 출력된 신호는 업 샘플링(up sampling)되어 원 신호로 복원될 수 있다.

<172> 또한, 입력되는 신호는 제1 필터뱅크(510)를 통과하여 주파수 변환된다.

<173> 상기 음성 부호화/복호화부(500)로부터 출력된 신호는 제2 필터뱅크(520)를 통과하여 주파수 영역으로 변환된다. 제1 필터뱅크(510) 또는 제2 필터뱅크(520)는 상기 신호에 대해 코사인 변환, 예를 들어 MDCT(Modified Discrete Transform)를 수행할 수 있다.

<174> 제1 필터뱅크(510)로부터 출력된 원본 신호의 주파수 성분과 제2 필터뱅크(520)로부터 출력된 음성 부호화/복호화를 거친 신호의 주파수 성분이 판단부(530)로 입력된다. 판단부(530)는 상기 입력된 두 주파수 성분을 이용하여, 음성 부호화기와 오디오 부호화기 중 어느 부호화기가 상기 입력 신호를 더욱 효율적으로 부호화할 수 있는지 여부를 판단할 수 있다.

<175> 판단부(530)는 다음의 수학적 식 15와 같이 계산되는 지각적 엔트로피를 이용하여 입력되는 신호에 더욱 효율적인 부호화기를 판단할 수 있다.

<176> 수학적 식 15

$$PE_i = \sum_{j=j_{low(i)}}^{j_{high(i)}} N(j)$$

where

$$N(j) = \begin{cases} 0 & , x(j) = 0 \\ \log_2 \left(2 \left| \text{nint} \left(\frac{x(j)}{\delta} \right) \right| + 1 \right) & , x(j) \neq 0 \end{cases}$$

<177>

<178> 상기 수학적 식 15에서, $x(j)$ 는 주파수 성분의 계수들이며, j 는 주파수 성분의 인덱스이고, δ 는 양자화 스텝 사이즈(quantization step size)이며, $\text{nint}()$ 는 가장 가까운 정수를 나타내는 함수이고, $j_{low(i)}$ 와 $j_{high(i)}$ 는 각각 스케일팩터 밴드(scalefactor band)에서의 시작 주파수 인덱스와 끝 주파수 인덱스이다.

<179> 판단부(530)는 상기 수학적 식 15를 이용하여 원본 신호의 주파수 성분과 음성 부호화/복호화된 신호의 주파수 성분 각각에 대해 스케일팩터 밴드(scalefactor band)의 지각적 엔트로피(perceptual entropy)를 계산하고, 상기 계산된 두 지각적 엔트로피를 비교하여 상기 신호에 효율적인 부호화기를 판단할 수 있다.

<180> 예를 들어, 판단부(530)는 원본 신호 주파수 성분의 지각적 엔트로피가 적은 값을 가지는 경우 오디오 부호화기가 상기 원본 신호를 더욱 효율적으로 부호화할 수 있는 것으로 판단하며, 음성 부호화/복호화된 신호의 주파수 성분의 지각적 엔트로피가 적은 값을 가지는 경우에는 음성 부호화기가 상기 원본 신호를 더욱 효율적으로 부호화할 수 있는 것으로 판단할 수 있다.

<181> 도 12는 도 1에 도시된 부호화부의 구성에 대한 제1 실시예를 블록도로 도시한 것으로, 음성 부호화기의 구성에 대한 일실시예를 도시한 것이다.

<182> 음성 부호화기는 입력 신호를 프레임 단위로 선형 예측 부호화할 수 있으며, 하나의 프레임마다 Levinson-Durbin 알고리즘을 이용하여 LPC 계수, 예를 들어 16차의 LPC 계수를 추출할 수 있다. 여기 신호는 적응 코드북과 고정 코드북 검색 과정을 통하여 양자화될 수 있으며, 상기 여기 신호는 ACELP(Algebraic Code Excited Linear Prediction) 방법을 이용하여 양자화될 수 있으며, 상기 여기 신호의 이득에 대해서는 켄레 구조(conjugate structure)를 갖는 양자화표를 이용하여 벡터 양자화가 수행될 수 있다.

<183> 도 12에 도시된 음성 부호화 장치는 선형예측분석부(600), 피치추정부(610), 코드북검색부(620), LSP변환부(630) 및 양자화부(640)를 포함하여 이루어진다.

<184> 선형예측분석부(600)는 비대칭 윈도우(window)를 이용해 구한 자기 상관 계수를 사용하여 입력되는 신호에 대해 프레임 단위로 선형 예측 분석을 수행한다. 자기 상관 계수를 구함에 있어, 선형예측분석부(600)는 예견 구간, 예를 들어 상기 비대칭 윈도우가 30ms의 길이를 가지는 경우 5ms의 길이를 가지는 예견 구간을 두어 선형 예측 분석을 수행할 수 있다.

<185> 자기 상관 계수는 Levinson-Durbin 알고리즘을 이용하여 선형 예측 계수로 변환된다. LSP변환부(630)는 양자화와 선형 보간을 위해 상기 선형 예측 계수를 LSP(Line Spectrum Pair)로 변환하고, 양자화부(640)는 상기 LSP로 변환된 선형 예측 계수를 양자화한다.

<186> 피치추정부(610)는 적응 코드북 검색의 복잡도를 줄이기 위해서 개회로 피치를 추정한다. 피치추정부(610)는 프레임마다 가중 음성 신호 도메인에서 개회로 피치 주기를 추정한다. 상기 추정된 피치 주기를 이용하여 하모닉 잡음 형성 필터(harmonic noise shaping filter)가 구성되고, 상기 구성된 하모닉 잡음 형성 필터, 선형 예측 합성 필터, 포맷트 지각 가중 필터를 이용하여 충격 응답이 계산된다. 상기 계산된 충격 응답은 여기 신호(excitation signal)의 양자화를 위한 목적신호(target signal) 생성에 사용될 수 있다.

<187> 코드북검색부(620)는 적응 코드북과 고정 코드북을 검색한다. 적응 코드북은 서브프레임 단위로 검색되고, 폐회로 피치 검색과 과거 여기 신호의 보간에 의한 적응 코드북 벡터 계산에 의해 검색될 수 있다. 적응 코드북 변

수들은 피치 필터의 피치 주기와 이득일 수 있으며, 검색 단계에서 여기 신호는 폐회로 검색을 단순화하기 위해 선형 예측 합성필터에 의해 생성된다.

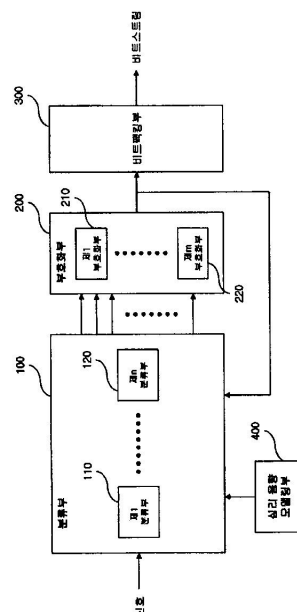
- <188> 고정 코드북의 구조는 ISPP(Interleaved Single Pulse Permutation) 설계에 기반할 수 있다. 64개의 펄스가 위치할 수 있는 위치들로 구성된 코드북 벡터는 16개의 위치들로 구성된 4개의 트랙으로 나누어진다. 전송물에 따라 특정 수의 펄스들이 각 트랙마다 위치하게 된다. 코드북 인덱스(index)는 각 트랙에서의 펄스의 위치와 부호를 나타내게 되므로, 코드북을 저장할 필요가 없이 코드북 인덱스 자체만으로 여기 신호(excitation)의 생성이 가능하다.
- <189> 음성 부호화기는 상기한 바와 같은 부호화 과정들을 시간 영역에서 수행하는 것이 바람직하다. 또한, 입력되는 신호가 도 1에 도시된 분류부(100)에서 선형 예측 부호화되는 경우, 도 12에 도시된 음성 부호화기의 구성 중 선형예측분석부(600)는 생략될 수 있다.
- <190> 도 12를 참조하여 음성 부호화기의 구성에 대한 일 실시예를 설명하였으나, 상기에서 설명한 구조 이외에 음성 신호를 효율적으로 부호화할 수 있는 여러 음성 부호화기들이 사용 가능하다.
- <191> 도 13은 도 1에 도시된 부호화부의 구성에 대한 제2 실시예를 블록도로 도시한 것으로, 오디오 부호화기의 구성에 대한 일 실시예를 도시한 것이다. 도 13에 도시된 오디오 부호화기는 필터뱅크(700), 심리음향모델링부(710) 및 양자화부(720)를 포함하여 이루어진다.
- <192> 필터뱅크(700)는 입력되는 신호를 주파수 영역으로 변환한다. 필터뱅크(700)는 상기 입력되는 신호에 대해 코사인 변환, 예를 들어 MDCT(Modified Discrete Transform)를 수행할 수 있다.
- <193> 심리음향모델링부(710)는 입력되는 신호의 마스킹 임계치 또는 신호 대 마스크 비(SMR, Signal-to-Mask Ratio)를 계산한다. 양자화부(720)는 상기 코사인변환부(700)로부터 출력되는 MDCT 계수들에 대해 상기 마스킹 임계치를 이용하여 양자화를 수행한다. 또한, 양자화부(720)는 주어진 비트율 내에서 양자화된 신호의 가청 왜곡을 최소화하기 위하여 상기 신호 대 마스크 비(SMR)를 사용할 수 있다.
- <194> 오디오 부호화기는 상기한 바와 같은 부호화 과정들을 주파수 영역에서 수행하는 것이 바람직하다.
- <195> 도 13을 참조하여 오디오 부호화기의 구성에 대한 일 실시예를 간략히 설명하였으나, 상기에서 설명한 구조 이외에 AAC(Advanced Audio Coding)와 같은 오디오 신호를 효율적으로 부호화할 수 있는 여러 오디오 부호화기들이 사용 가능하다.
- <196> 상기 AAC에서는 TNS(Temporal Noise Shaping), Intensity/Coupling, Prediction 및 M/S stereo coding 등이 수행되는데, 상기 TNS는 필터뱅크 윈도우 내에서 시간 영역의 양자화 잡음을 적절하게 배치하여 청각적으로 들리지 않도록 하는 역할을 한다. 상기 Intensity/Coupling은 고주파수 대역에 대한 소리의 방향 지각이 주로 에너지의 시간 크기에 의존한다는 사실을 이용해 오디오 신호를 부호화하여, 오디오 신호의 에너지 크기만을 전송함으로써 공간 정보의 전송량을 감소시킬 수 있다.
- <197> 상기 Prediction은 프레임과 프레임 사이에서 스펙트럼 성분들의 연관성(correlation)을 이용하여 통계적 특성이 변화하지 않는 신호의 중복을 제거한다. 상기 M/S stereo coding은 좌-우 신호를 전송하는 대신 신호의 정규화된 합(M, Middle)과 차(S, Side)를 전송한다.
- <198> 상기와 같은 과정을 거친 신호는 심리음향모델에서 얻은 신호 대 마스크 비(SMR)를 이용하여 합성에 의한 분석(AbS, Analysis-by-synthesis)을 수행하는 양자화기에 의해 양자화된다.
- <199> 상기한 바와 같이, 음성 부호화기는 선형 예측 부호화와 같은 모델링 기법을 이용하여 입력되는 신호를 부호화한다. 따라서 결정부(122)는 일정한 기준에 따라 입력되는 신호가 모델링이 잘되는 신호인지 여부를 판단하여, 모델링이 잘되는 신호인 경우 상기 신호를 음성 부호화기를 이용해 부호화하는 것으로 결정하고, 모델링이 잘되지 않는 신호인 경우 상기 신호를 오디오 부호화기를 이용해 부호화하는 것으로 결정할 수 있다.
- <200> 도 14는 본 발명에 따른 부호화 장치의 구성에 대한 제2 실시예를 블록도로 도시한 것, 도 14에 도시된 부호화 장치의 동작들 중 상기 도 1 내지 도 13을 참조하여 설명한 동작과 동일한 것에 대해서는 설명을 생략하기로 한다.
- <201> 분류부(100)는 입력되는 신호를 복수의 신호들(제1 신호 내지 제n 신호)로 분할하고, 상기 분할된 n개의 신호들 각각에 대해 신호의 특성에 따라 부호화할 부호화기를 결정한다.

- <202> 도 14에 도시된 바와 같이, 복수의 부호화기들(230, 240, 250, 260, 270)은 상기 분할된 복수의 신호들을 순차적으로 부호화하는 것이 바람직하다. 또한, 입력되는 신호가 복수의 주파수 대역 신호들로 분할된 경우, 저주파 대역에서 고주파 대역의 순서로 상기 복수의 주파수 대역 신호들을 부호화할 수 있다.
- <203> 분할된 복수의 신호들이 순차적으로 부호화되는 경우, 이전 신호에 대한 부호화 오차가 다음 신호의 부호화에 이용되도록 함으로써, 분할된 신호들에 대해 서로 다른 부호화 방법을 사용함에 따라 발생할 수 있는 왜곡을 막고, bandwidth scalability 제공을 가능하게 할 수 있다.
- <204> 도 14를 참조하면, 첫번째 부호화기(230)는 부호화된 제1 신호를 다시 복호화하고, 상기 복호화된 제1 신호와 원래의 제1 신호의 차인 부호화 오차를 제2 신호를 부호화하기 위한 부호화기(240)로 출력한다. 두번째 부호화기(240)는 상기 제1 신호의 부호화 오차를 이용하여 상기 제2 신호를 부호화한다. 제2 신호 내지 제3 신호에 대해서도 상기한 바와 같이 이전에 부호화된 신호의 부호화 오차를 고려하여 부호화를 수행함으로써, 오차 없는 부호화가 가능하고, 음질을 향상시킬 수 있다.
- <205> 본 발명에 따른 복호화 장치는 상기 도 1 내지 도 14를 참조하여 설명한 부호화 장치의 부호화 과정의 역 과정을 수행함에 의해, 입력되는 비트스트림으로부터 신호를 복원할 수 있다.
- <206> 도 15는 본 발명에 따른 복호화 장치의 구성에 대한 일실시예를 블록도로 도시한 것으로, 도 15에 도시된 복호화 장치는 비트언팩킹부(800), 복호화기결정부(810), 복호화부(820), 합성부(830)를 포함하여 이루어진다.
- <207> 비트언팩킹부(1600)는 입력되는 비트스트림으로부터 부호화된 복수의 신호들 및 상기 신호들을 복호화하기 위한 부가 정보를 추출한다.
- <208> 도 15에 도시된 바와 같이, 복호화부(820)는 서로 다른 방법에 의해 복호화를 수행하는 복수의 복호화부들(821, 822)을 포함한다.
- <209> 상기 부호화된 복수의 신호들 각각에 대해, 복호화기결정부(810)는 복수의 복호화부(821, 822)들 중 상기 신호를 가장 효율적으로 복호화할 복호화부를 결정한다. 복호화기결정부(810)는 상기 본 발명에 따른 부호화 장치에 대한 설명부분에서 기재한 바와 같이 상기 신호들 각각의 특성에 따라 복호화부를 결정할 수도 있으나, 비트스트림으로부터 추출된 부가 정보를 이용하여 복호화부를 결정하는 것이 바람직하다.
- <210> 복호화부 결정에 이용되는 상기 부가 정보는 부호화 장치에서 분류된 상기 신호의 클래스에 대한 정보이거나, 상기 신호를 부호화한 부호화기에 대한 정보 또는 상기 신호를 복호화할 복호화기에 대한 정보일 수 있다.
- <211> 예를 들어, 복호화기결정부(810)는 상기 부가 정보로부터 상기 신호의 클래스를 파악하고, 복수의 복호화부(821, 822)들 중 상기 클래스에 대응되는 복호화부를 상기 신호를 복호화할 복호화부로 결정할 수 있다. 이 경우, 상기 결정된 복호화부는 상기 클래스에 속하는 신호를 가장 효율적으로 복호화하는 구조를 가지는 것이 바람직하다.
- <212> 복호화기결정부(810)는 상기 부가 정보로부터 상기 신호를 부호화한 부호화기를 파악하고, 복수의 복호화부(821, 822)들 중 상기 부호화기에 대응되는 복호화부를 상기 신호를 복호화할 복호화부로 결정할 수 있다. 예를 들어, 상기 신호가 음성 부호화기로 부호화된 경우 복수의 복호화부(821, 822)들 중 음성 복호화기가 상기 신호를 복호화할 복호화기로 결정될 수 있으며, 상기 신호가 오디오 부호화기로 부호화된 경우 복수의 복호화부(821, 822)들 중 오디오 복호화기가 상기 신호를 복호화할 복호화기로 결정될 수 있다.
- <213> 또한, 복호화기결정부(810)는 상기 부가 정보로부터 상기 신호를 복호화할 복호화부를 파악하고, 복수의 복호화부(821, 822)들 중 상기 파악된 복호화부를 상기 신호를 복호화할 복호화부로 결정할 수 있다.
- <214> 그 이외에도, 복호화기결정부(810)는 상기 부가 정보로부터 상기 신호의 특성에 대한 정보를 얻을 수 있으며, 상기 얻어진 특성을 가지는 신호를 가장 효율적으로 복호화할 수 있는 복호화부를 복수의 복호화부(821, 822)들 중에서 선택할 수 있다.
- <215> 상기 부호화된 복수의 신호들 각각은 상기 결정된 복호화부에 의해 복호화되고, 상기 복호화된 복수의 신호들은 합성부(830)에 의해 합성되어 원 신호로 복원된다.
- <216> 비트언팩킹부(800)는 상기 비트스트림에 포함된 복수의 신호들에 대한 분할 정보, 예를 들어 상기 복수의 신호들의 개수, 분할된 주파수 대역 정보 등을 추출하고, 합성부(830)는 상기 추출된 분할 정보를 이용하여 복호화부(820)에서 복호화된 복수의 신호들을 합성할 수 있다.

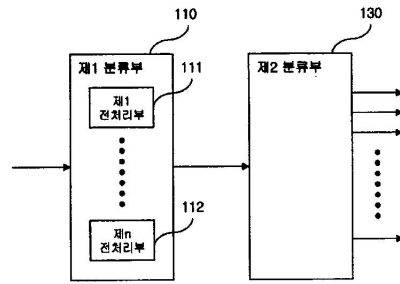
- <217> 도 15에 도시된 바와 같이, 합성부(830)는 복수의 합성부들(831, 832)을 포함할 수 있다. 복수의 합성부들(831, 832) 각각은 상기 복호화된 복수의 신호들에 대해 신호 합성을 수행하거나, 상기 복수의 신호들 전부 또는 일부에 대해 도메인 변환 또는 추가적인 복호화 과정을 수행할 수 있다.
- <218> 상기 복수의 합성부들(831, 832) 중 어느 하나는 부호화 장치에서 수행된 전처리의 역 과정인 후처리를 상기 합성된 신호에 대해 수행할 수 있다. 상기 후처리 수행 여부에 대한 정보 및 상기 후처리에 사용될 복호화 정보는 입력되는 비트스트림으로부터 추출될 수 있다.
- <219> 도 16을 참조하면, 합성부(830)에 포함된 복수의 합성부들 중 어느 하나(833)는 복수의 후처리부들(834, 835)을 포함할 수 있다. 제1 합성부(831)는 상기 복호화된 복수의 신호들을 하나의 신호로 합성하고, 제1 내지 제n 후처리부(834, 835) 중 어느 하나는 상기 합성된 신호에 대해 후처리를 수행한다.
- <220> 복수의 후처리부들(834, 835) 중 상기 합성된 신호에 대해 후처리를 수행할 후처리부에 대한 정보는 입력되는 비트스트림에 포함될 수 있다.
- <221> 복수의 합성부들 중 어느 하나는 비트스트림으로부터 추출된 선형 예측 계수를 이용해 상기 합성된 신호에 대해 선형 예측 복호화를 수행하여, 원 신호를 복원할 수 있다.
- <222> 상기한 본 발명에 따른 부호화/복호화 방법은 컴퓨터에서 실행되기 위한 프로그램으로 제작되어 컴퓨터가 읽을 수 있는 기록 매체에 저장될 수 있으며, 상기한 본 발명에 따른 데이터 구조를 가지는 멀티 미디어 데이터도 컴퓨터가 읽을 수 있는 기록 매체에 저장될 수 있다. 상기 컴퓨터가 읽을 수 있는 기록 매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 저장 장치를 포함한다. 컴퓨터가 읽을 수 있는 기록 매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록 매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다. 그리고, 사용자 추적 방법을 구현하기 위한 기능적인(function) 프로그램, 코드 및 코드 세그먼트들은 본 발명이 속하는 기술분야의 프로그래머들에 의해 용이하게 추론될 수 있다.
- <223> 이상에서는 본 발명의 바람직한 실시예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특정의 실시예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해해서는 안 될 것이다.

도면

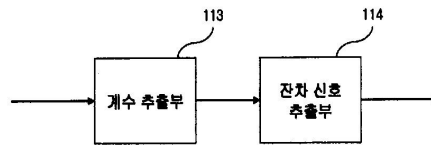
도면1



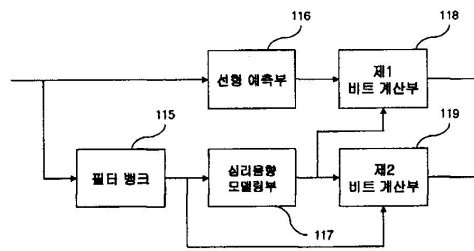
도면2



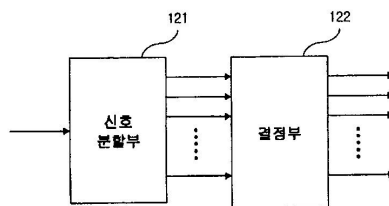
도면3



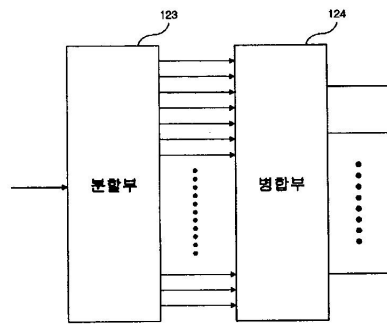
도면4



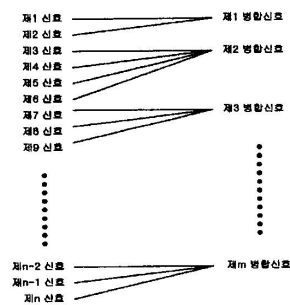
도면5



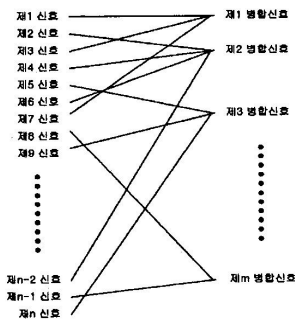
도면6



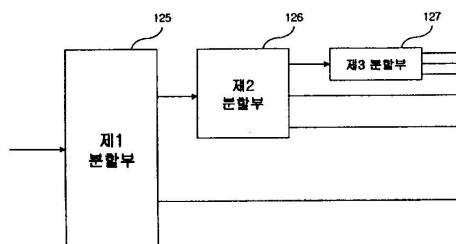
도면7



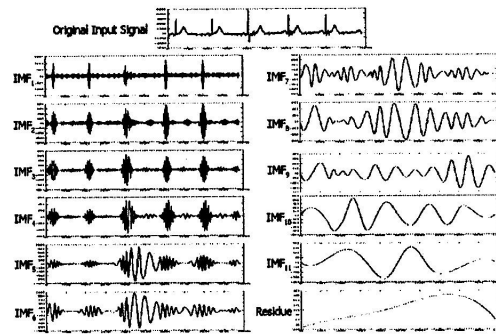
도면8



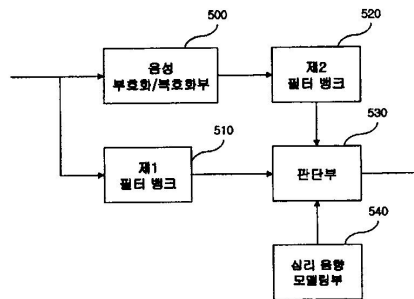
도면9



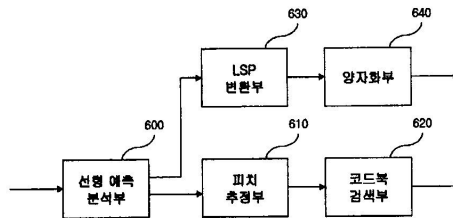
도면10



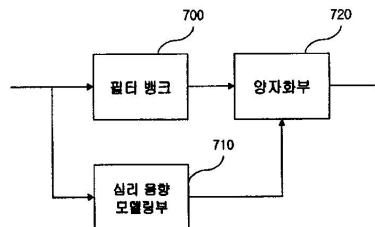
도면11



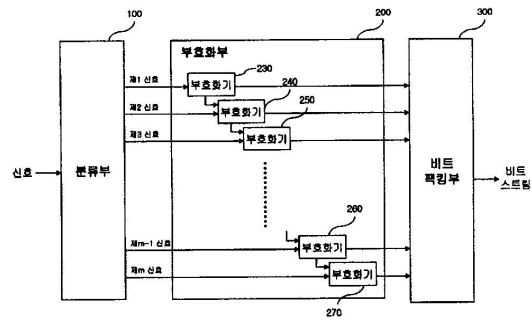
도면12



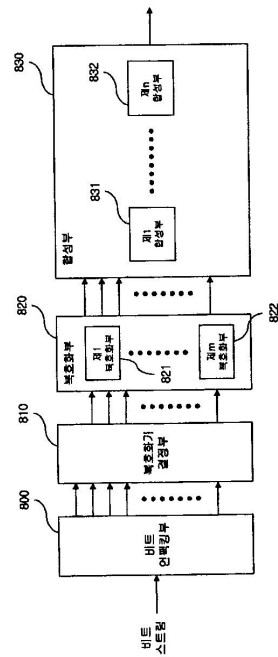
도면13



도면14



도면15



도면16

