



(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl.

G10L 19/02 (2006.01)

G10L 19/14 (2006.01)

(11) 공개번호

10-2007-0077652

(43) 공개일자

2007년07월27일

(21) 출원번호 10-2006-0007341

(22) 출원일자 2006년01월24일

심사청구일자 2006년01월24일

(71) 출원인 삼성전자주식회사
경기도 수원시 영통구 매탄동 416

(72) 발명자 오은미
서울시 서초구 반포4동 미도2차아파트 503동 305호
주기현
서울시 광진구 자양3동 우성1차아파트 105동 1202호
김중희
서울시 강서구 화곡2동 874-1호 4층
손창용
경기도 군포시 금정동 872번지 충무아파트 217동 1401호

(74) 대리인 특허법인무한

전체 청구항 수 : 총 21 항

(54) 적응적 시간/주파수 기반 부호화 모드 결정 장치 및 이를위한 부호화 모드 결정 방법

(57) 요약

적응적 시간/주파수 기반 부호화 모드 결정 장치 및 방법이 개시된다. 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 부호화 모드 결정 장치는 입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성하는 시간 영역 특성 추출부; 상기 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 상기 입력 오디오 신호의 프레임에 상응하는 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성하는 주파수 영역 특성 추출부; 및 상기 시간 영역 특성 및 상기 주파수 영역 특성을 이용하여, 상기 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 모드 결정기를 포함하는 것을 특징으로 한다. 따라서, 다양한 오디오 콘텐츠를 효율적으로 압축할 수 있다.

대표도

도 4

특허청구의 범위

청구항 1.

입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성하는 시간 영역 특성 추출부;

상기 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 상기 입력 오디오 신호의 프레임에 상응하는 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성하는 주파수 영역 특성 추출부; 및

상기 시간 영역 특성 및 상기 주파수 영역 특성을 이용하여, 상기 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 모드 결정기

를 포함하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 2.

제1항에 있어서,

상기 주파수 영역 특성 추출부가 상기 입력 오디오 신호의 현재 프레임의 주파수 영역 신호 분석을 수행할 때, 상기 시간 영역 특성 추출부는 상기 입력 오디오 신호의 현재 또는 다음 프레임의 주파수 영역 신호에 해당하는 시간 영역 신호 분석을 수행하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 3.

제2항에 있어서,

상기 시간 영역 특성은 상기 입력 오디오 신호의 시간 영역 단구간 특성이고, 상기 주파수 영역 특성은 상기 주파수 밴드 각각에 상응하는 주파수 영역 단구간 특성이며,

상기 적응적 시간/주파수 기반 부호화 모드 결정 장치는

상기 시간 영역 단구간 특성 및 상기 주파수 영역 단구간 특성을 분석하여 시간 영역 장구간 특성 및 주파수 영역 장구간 특성을 생성하는 장구간 특성 추출부를 더 포함하고,

상기 모드 결정기는 상기 시간 영역 장구간 특성 및 상기 주파수 영역 장구간 특성을 더 이용하여 부호화 모드를 결정하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 4.

제3항에 있어서,

상기 모드 결정기의 현재 프레임에 대한 모드 결정시 다음 프레임에 대한 시간 영역 분석을 수행한 결과가 프레임 특성 버퍼를 통하여 이전과 현재 그리고 다음 프레임에 대한 단구간/장구간 예측 이득으로 적용되는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 5.

제3항에 있어서,

상기 시간 영역 단구간 특성은 천이 정도 및 단기/장기 예측 이득의 크기 정도를 포함하고, 상기 주파수 영역 단구간 특성은 스펙트럼의 자기상관도를 포함하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 6.

제5항에 있어서,

상기 시간 영역 장구간 특성은 주기성의 연속성 정도, 주파수 스펙트럼의 기울기 정도 및 프레임 에너지 정도를 포함하고, 상기 주파수 영역 장구간 특성은 채널간 상관도를 포함하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 7.

제6항에 있어서,

상기 모드 결정기는

상기 입력 오디오 신호의 스테레오 신호 정도가 소정 레벨 이상이라는 제1 조건, 상기 입력 오디오 신호의 천이 정도가 소정 레벨 미만이라는 제2 조건, 단기/장기 예측 이득이 소정 레벨 미만이라는 제3 조건, 상기 주파수 밴드에 상응하는 스펙트럼의 자기상관도가 소정 레벨 미만이라는 제4 조건 중 어느 하나 이상을 만족하는 경우에 상기 부호화 모드를 주파수 기반 부호화 모드로 결정하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 8.

제7항에 있어서,

상기 모드 결정기는

상기 제1 내지 제4 조건을 모두 만족시키지 아니하고,

상기 입력 오디오 신호의 주기성의 연속성이 소정 구간 이상 지속된다는 제5 조건 및 상기 주파수 스펙트럼의 기울기 정도가 완만하고 높은 프레임 에너지가 일정 구간 이상 연속적으로 지속되는 정도인 음악 연속성이 소정 레벨 이상인 제6 조건을 모두 만족시키지 아니하는 경우에 부호화 모드를 시간 기반 부호화 모드로 결정하고,

상기 제5 조건 및 제6 조건 중 어느 하나 이상이 만족되는 경우에 부호화 모드를 주파수 기반 부호화 모드로 결정하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 9.

제1항에 있어서,

상기 주파수 영역 특성 추출부는 시간 영역의 상기 입력 오디오 신호를 주파수 가변 MLT, MLT 및 FFT 중 어느 하나에 의하여 변환하여, 상기 주파수 영역 신호 분석을 수행하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 장치.

청구항 10.

입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성하는 시간 영역 특성 추출부;

상기 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 상기 입력 오디오 신호의 프레임의 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성하는 주파수 영역 특성 추출부;

상기 시간 영역 특성 및 상기 주파수 영역 특성을 이용하여, 상기 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 모드 결정기;

상기 주파수 밴드 각각에 대해 상기 결정된 부호화 모드로 부호화하는 부호화부; 및

상기 부호화된 데이터에 대한 비트 스트림 처리를 하여 처리된 비트 스트림을 출력하는 비트스트림 출력부를 포함하는 것을 특징으로 하는 적응적 시간/주파수 기반 오디오 부호화 장치.

청구항 11.

제10항에 있어서,

상기 주파수 영역 특성 추출부가 상기 입력 오디오 신호의 현재 프레임의 주파수 영역 신호 분석을 수행할 때, 상기 시간 영역 특성 추출부는 상기 입력 오디오 신호의 현재 또는 다음 프레임의 주파수 영역 신호에 해당하는 시간 영역 신호 분석을 수행하는 것을 특징으로 하는 적응적 시간/주파수 기반 오디오 부호화 장치.

청구항 12.

제11항에 있어서,

상기 시간 영역 특성은 상기 입력 오디오 신호의 시간 영역 단구간 특성이고, 상기 주파수 영역 특성은 상기 주파수 밴드 각각에 상응하는 주파수 영역 단구간 특성이며,

상기 적응적 시간/주파수 기반 오디오 부호화 장치는

상기 시간 영역 단구간 특성 및 상기 주파수 영역 단구간 특성을 분석하여 시간 영역 장구간 특성 및 주파수 영역 장구간 특성을 생성하는 장구간 특성 추출부를 더 포함하고,

상기 모드 결정기는 상기 시간 영역 장구간 특성 및 상기 주파수 영역 장구간 특성을 더 이용하여 부호화 모드를 결정하는 것을 특징으로 하는 적응적 시간/주파수 기반 오디오 부호화 장치.

청구항 13.

입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성하는 단계;

상기 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 상기 입력 오디오 신호의 프레임에 상응하는 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성하는 단계; 및

상기 시간 영역 특성 및 상기 주파수 영역 특성을 이용하여, 상기 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 단계

를 포함하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 14.

제13항에 있어서,

상기 주파수 영역 특성을 생성하는 단계가 상기 입력 오디오 신호의 현재 프레임의 주파수 영역 신호 분석을 수행할 때, 상기 시간 영역 특성을 추출하는 단계는 상기 입력 오디오 신호의 현재 또는 다음 프레임의 주파수 영역 신호에 해당하는 시간 영역 신호 분석을 수행하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 15.

제14항에 있어서,

상기 시간 영역 특성은 상기 입력 오디오 신호의 시간 영역 단구간 특성이고, 상기 주파수 영역 특성은 상기 주파수 밴드 각각에 상응하는 주파수 영역 단구간 특성이며,

상기 적응적 시간/주파수 기반 부호화 모드 결정 방법은

상기 시간 영역 단구간 특성 및 상기 주파수 영역 단구간 특성을 분석하여 시간 영역 장구간 특성 및 주파수 영역 장구간 특성을 생성하는 단계

를 더 포함하고,

상기 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 단계는 상기 시간 영역 장구간 특성 및 상기 주파수 영역 장구간 특성을 더 이용하여 부호화 모드를 결정하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 16.

제15항에 있어서,

상기 주파수 기반 부호화 모드를 결정하는 단계는 현재 프레임에 대한 모드 결정시 다음 프레임에 대한 시간 영역 분석을 수행한 결과가 프레임 특성 버퍼를 통하여 이전과 현재 그리고 다음 프레임에 대한 단구간/장구간 예측 이득으로 적용되는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 17.

제16항에 있어서,

상기 시간 영역 단구간 특성은 천이 정도 및 단기/장기 예측 이득의 크기 정도를 포함하고, 상기 주파수 영역 단구간 특성은 스펙트럼의 자기상관도를 포함하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 18.

제17항에 있어서,

상기 시간 영역 장구간 특성은 주기성의 연속성 정도, 주파수 스펙트럼의 기울기 정도 및 프레임 에너지 정도를 포함하고, 상기 주파수 영역 장구간 특성은 채널간 상관도를 포함하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 19.

제18항에 있어서,

상기 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 단계는

상기 입력 오디오 신호의 스테레오 신호 정도가 소정 레벨 이상이거나, 상기 입력 오디오 신호의 천이 정도가 소정 레벨 미만이거나, 단기/장기 예측 이득이 소정 레벨 미만이거나, 상기 주파수 밴드에 상응하는 스펙트럼의 자기상관도가 소정 레벨 미만인 경우에 상기 부호화 모드를 주파수 기반 부호화 모드로 결정하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 20.

제19항에 있어서,

상기 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 단계는 상기 입력 오디오 신호의 주기성의 연속성이 소정 구간 이상 지속되지 아니함과 동시에, 상기 주파수 스펙트럼의 기울기 정도가 소정 레벨 이상이거나 소정 레벨 이상의 프레임 에너지가 일정 구간 이상 연속적으로 지속되지 아니하는 경우에 부호화 모드를 시간 기반 부호화 모드로 결정하는 것을 특징으로 하는 적응적 시간/주파수 기반 부호화 모드 결정 방법.

청구항 21.

제13항 내지 제20항 중 어느 한 항의 방법을 실행하기 위한 프로그램이 기록되어 있는 것을 특징으로 하는 컴퓨터에서 판독 가능한 기록 매체.

명세서

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 오디오 부호화/복호화 장치 및 방법에 관한 것으로, 특히 입력 오디오 데이터의 특성에 따라 시간 기반 부호화 또는 주파수 기반 부호화를 적응적으로 선택하여 입력 오디오 데이터를 부호화함으로써 두 가지 부호화 방식의 부호화 이득을 최대한 이용하여 고압축 효율을 얻을 수 있는 적응적 시간/주파수 기반 오디오 부호화 장치 및 부호화 모드 결정 방법에 관한 것이다.

기존의 음성/음악 압축 방식은 크게 오디오 코덱과 음성 코덱으로 분류된다. aacPlus 등의 오디오 코덱은 주파수 영역상의 신호를 압축하는 알고리즘으로, 심리 음향 모델을 적용하는데, 압축 대상이 오디오 신호가 아니라 음성 신호인 경우에는 동일한 부호화 데이터량으로 고려할 때, 음성 코덱 방식의 압축 결과물보다 음질이 크게 저하되며, 특히 어택(attack) 신호에는 음질 저하가 더욱 크다. 반면에, AMR-WB 등의 코덱은 시간 영역 상의 신호를 압축하는 알고리즘으로, 음성 발생 모델을 적용하는데, 압축 대상이 음성 신호가 아니라 오디오 신호인 경우에는 동일한 부호화 데이터량으로 고려할 때, 오디오 코덱 방식의 압축 결과물보다 음질이 크게 저하되는 단점이 있다.

상술한 특성을 고려하여, 음성/음악 압축을 동시에 효율적으로 수행하기 위한 종래의 기술로서 AMR-WB+ 방식(3GPP TS 26.290)이 있는데, 이는 음성 압축 방식으로 ACELP(Algebraic Code Excited Linear Prediction)를, 오디오 압축 방식으로 TCX(Transform Coded Excitation)를 사용한다. 특히, 이 방식은 시간축상의 프레임별로 ACELP 방식을 적용할지, TCX 방식을 적용할지를 결정하여 부호화하는데, 압축 대상이 음성 신호에 가까운 경우에는 효율적으로 동작하나, 오디오 신호에 가까운 경우에는 처리 단위별 부호화 과정으로 인한 음질 또는 압축율의 열화가 발생하는 문제점이 있다.

따라서, 압축 방식을 선택적으로 적용하여 입력 오디오 데이터를 부호화함에 있어서, 부호화 모드 결정 단위를 어떻게 결정할 것인지 및 어떤 기준에 의해서 해당 결정 단위에 대한 부호화 모드를 결정할 것인지는 부호화 성능에 큰 영향을 미치는 매우 중요한 요소이다.

발명이 이루고자 하는 기술적 과제

본 발명은 상술한 바와 같은 종래기술의 문제점을 해결하기 위해 안출된 것으로서, 입력 오디오 신호에 대한 부호화 모드를 주파수 대역별로 판단하여 시간 기반 부호화 또는 주파수 기반 부호화함으로써, 두 가지 부호화 방식의 부호화 이득을 효율적으로 이용하여 고압축 성능을 얻을 수 있도록 하는 것을 목적으로 한다.

또한, 본 발명은 입력 오디오 신호를 시간 영역 및 주파수 영역별로 장구간 특성 및 단구간 특성을 추출하여 주파수 대역별로 적절한 부호화 모드를 결정하도록 하여 적응적 시간/주파수 기반 오디오 부호화의 성능을 최적화하는 것을 목적으로 한다.

또한, 본 발명은 개루프(open loop) 결정 방식을 사용하여 낮은 복잡도를 가지면서도 효과적으로 부호화 모드를 결정하는 것을 목적으로 한다.

발명의 구성

상기의 목적을 달성하고 종래기술의 문제점을 해결하기 위하여, 본 발명의 적응적 시간/주파수 기반 부호화 모드 결정 장치는, 입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성하는 시간 영역 특성 추출부; 상기 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 상기 입력 오디오 신호의 프레임에 상응하는 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성하는 주파수 영역 특성 추출부; 및 상기 시간 영역 특성 및 상기 주파수 영역 특성을 이용하여, 상기 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 모드 결정기를 포함하는 것을 특징으로 한다.

또한, 본 발명의 적응적 시간/주파수 기반 오디오 부호화 장치는 입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성하는 시간 영역 특성 추출부; 상기 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 상기 입력 오디오 신호의 프레임의 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성하는 주파수 영역 특성 추출부; 상기 시간 영역 특성 및 상기 주파수 영역 특성을 이용하여, 상기 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 모드 결정기; 상기 주파수 밴드 각각에 대해 상기 결정된 부호화 모드로 부호화하는 부호화부; 및 상기 부호화된 데이터에 대한 비트 스트림 처리를 하여 처리된 비트 스트림을 출력하는 비트스트림 출력부를 포함하는 것을 특징으로 한다.

이 때, 주파수 영역 특성 추출부가 입력 오디오 신호의 현재 프레임(current frame)의 주파수 영역 신호 분석을 수행할 때, 시간 영역 특성 추출부는 입력 오디오 신호의 현재 또는 다음 프레임(next frame)의 주파수 영역 신호에 해당하는 시간 영역 신호 분석을 수행할 수 있다.

이 때, 시간 영역 특성은 입력 오디오 신호의 시간 영역 단구간 특성이고, 주파수 영역 특성은 주파수 밴드 각각에 상응하는 주파수 영역 단구간 특성이며, 적응적 시간/주파수 기반 오디오 부호화 장치는 시간 영역 단구간 특성 및 주파수 영역 단구간 특성을 분석하여 시간 영역 장구간 특성 및 주파수 영역 장구간 특성을 생성하는 장구간 특성 추출부를 더 포함하고, 모드 결정기는 시간 영역 장구간 특성 및 주파수 영역 장구간 특성을 더 이용하여 부호화 모드를 결정할 수 있다.

또한, 본 발명의 적응적 시간/주파수 기반 부호화 모드 결정 방법은 입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성하는 단계; 상기 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 상기 입력 오디오 신호의 프레임에 상응하는 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성하는 단계; 및 상기 시간 영역 특성 및 상기 주파수 영역 특성을 이용하여, 상기 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정하는 단계를 포함하는 것을 특징으로 한다.

본 발명에서 시간 기반 부호화 방식은 CELP(Code Excited Linear Prediction) 등 시간축 상에서 압축을 행하는 음성 압축 알고리즘을 의미하며, 주파수 기반 부호화 방식은 TCX(Transform Coded Excitation), AAC(Advanced Audio Codec) 등 주파수 축상에서 압축을 행하는 오디오 압축 알고리즘을 의미한다.

이하, 본 발명에 따른 바람직한 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.

도 1은 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 오디오 부호화 장치의 블록도이다.

도 1을 참조하면, 적응적 시간/주파수 기반 오디오 부호화 장치는 변환/모드 결정부(110), 부호화부(120) 및 비트스트림 출력부(130)를 포함한다.

변환/모드 결정부(110)는 입력 오디오 신호(IN)를 프레임 단위로 주파수 변환(frequency transform)하고, 변환된 주파수 영역을 복수 개의 주파수 영역으로 분할하여 생성된 주파수 밴드 각각에 대하여 시간 기반 부호화 모드 또는 주파수 기반 부호화 모드를 결정한다. 이 과정을 통하여, 변환/모드 결정부(110)는 시간 기반 부호화 모드로 결정된 주파수 영역 신호(S1), 주파수 기반 부호화 모드로 결정된 주파수 영역 신호(S2), 주파수 영역 분할에 관한 정보(S3), 주파수 밴드 각각의 부호화 모드 정보(S4)를 출력한다. 이 때, 주파수 영역을 일관되게 분할하는 경우라면, 부호화 단계에서 분할 정보를 필요로 하지 않을 수도 있으므로 주파수 영역 분할에 관한 정보(S3)는 이용되지 않을 수도 있다.

부호화부(120)는 시간 기반 부호화 모드로 결정된 주파수 영역 신호(S1)를 시간 기반 부호화하고, 주파수 기반 부호화 모드로 결정된 주파수 영역 신호(S2)를 주파수 기반 부호화하고, 시간 기반 부호화된 데이터(S5) 및 주파수 기반 부호화된 데이터(S6)를 출력한다.

비트스트림(bit stream) 출력부(130)는 부호화된 데이터(S5, S6)에 대한 비트 스트림 처리를 하여 처리된 비트 스트림을 출력한다. 이 때, 비트스트림 출력부(130)는 주파수 영역 분할에 관한 정보(S3) 및 주파수 밴드 각각의 부호화 모드 정보(S4)를 이용하여 비트스트림 처리를 할 수 있다. 이 때, 비트스트림은 엔트로피 부호화 등의 데이터 압축 과정을 거칠 수도 있다.

도 2는 주파수 영역 변환된 신호의 분할 및 부호화 모드 결정 과정을 나타낸 개념도이다.

도 2를 참조하면, 입력 오디오 신호에는 22000Hz의 주파수 성분까지 포함되어 있으며, 5개의 주파수 밴드로 분할되어 있다. 분할된 주파수 밴드는 상응하는 부호화 모드가 저주파에서 고주파 순으로 시간 기반 부호화 모드, 주파수 기반 부호화 모드, 시간 기반 부호화 모드, 주파수 기반 부호화 모드 및 주파수 기반 부호화 모드로 결정되어 있음을 알 수 있다. 이 때, 입력 오디오 신호는 소정 시간(예를 들어, 약 20ms) 동안의 오디오 프레임이며, 도 2에 도시된 그래프는 소정 시간 동안의 오디오 프레임에 대하여 주파수 변환한 것이다. 도 2에 도시된 바와 같이, 오디오 프레임은 sf1, sf2, sf3, sf4 및 sf5의 5개의 주파수 대역으로 나뉘어진다.

도 2에 도시된 바와 같이, 시간 영역에서의 하나의 프레임에 상응하는 주파수 영역이 분할된 주파수 밴드 각각에 대하여 적절한 부호화 모드 할당을 하는 것은 매우 중요하다. 이 때, 입력 오디오 신호의 시간 영역 특성 및 주파수 영역 특성을 이용하여 각각의 주파수 밴드 별로 적절한 부호화 모드 결정이 수행될 수 있다. 각각의 주파수 밴드 별 부호화 모드를 결정에 대해서는 이후에 상술하기로 한다.

도 3은 도 1에 도시된 변환/모드 결정부(110)의 일 예를 나타낸 블록도이다.

도 3을 참조하면, 변환/모드 결정부는 주파수 영역 변환부(310), 부호화 모드 판단부(320) 및 출력부(330)를 포함한다.

주파수 영역 변환부(310)는 입력 오디오 신호(IN)를 도 2에 도시된 주파수 스펙트럼과 같은 주파수 영역 신호(S7)로 변환한다. 예를 들어, 주파수 영역 변환부(310)는 입력 오디오 신호(IN)에 대하여 MLT(Modulated Lapped Transform)를 할 수 있다.

특히, 주파수 영역 변환부(310)는 입력 오디오 신호(IN)에 대한 주파수 가변(Frequency Varying) MLT를 수행할 수 있다. 주파수 가변 MLT는 M. Purat와 P. Noll이 저술한 "A New Orthonormal Wavelet Packet Decomposition for Audio Coding Using Frequency-Varying Modulated Lapped Transform, IEEE Workshop on Application of Signal Processing to Audio and Acoustics, Oct, 1995"에 상세하게 기술되어 있다.

주파수 가변 MLT를 이용하면, 주파수 변환된 주파수 영역 신호의 일부 주파수 밴드에 대해서는 주파수 기반 부호화를 수행하고, 다른 주파수 밴드에 대해서는 역 MLT 변환을 수행하여 시간 영역 신호로 바꾼 후 시간 기반 부호화를 수행할 수 있다. 이후에, 시간 기반 부호화된 신호가 다시 MLT 변환된 후 주파수 기반 부호화된 주파수 밴드 신호와 합해지면, 모든 주파수 대역에 대하여 부호화된 신호가 얻어진다.

부호화 모드 판단부(320)는 시간 영역 신호인 입력 오디오 신호(IN) 및 입력 오디오 신호(IN)가 주파수 변환된 주파수 영역 신호(S7)를 신호 분석하여 각각의 주파수 밴드 별로 시간 기반 부호화 모드 및 주파수 기반 부호화 모드 중 하나의 부호화 모드를 결정한다. 이 때, 부호화 모드 판단부(320)는 시간 영역 신호인 입력 오디오 신호(IN)의 현재 또는 다음 프레임(next frame)의 주파수 영역 신호 분석을 수행할 때, 주파수 영역 신호(S7)의 현재 프레임(current frame)의 주파수 영역 신호 분석을 수행할 수 있다.

현재 프레임의 모드 결정시 다음 프레임의 특성을 반영함으로써, 모드가 한 프레임 간격으로 자주 스위칭되는 것을 억제하여 모드 변화를 부드럽게 할 수 있다. 예를 들어, 과거, 현재 및 다음 특성값의 평균값을 사용하거나, 과거와 현재의 특징으로 현재 프레임의 모드를 결정한 후에 다음 프레임의 특징 값에 따라 스위칭되는 것을 보류하고 판단을 다음 프레임으로 넘기는 방식으로 부호화 모드 판단부(320)를 구현할 수 있다.

출력부(330)는 부호화 모드 판단부(320)의 판단 결과에 따라 주파수 영역 신호(S7)를 시간 기반 부호화 모드로 결정된 주파수 영역 신호(S1), 주파수 기반 부호화 모드로 결정된 주파수 영역 신호(S2), 주파수 영역 분할에 관한 정보(S3) 및 부호화 모드 정보(S4)를 출력한다.

도 4는 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 부호화 모드 결정 장치의 블록도이다.

도 4를 참조하면, 적응적 시간/주파수 기반 부호화 모드 결정 장치는 시간 영역 특성 추출부(410), 주파수 영역 특성 추출부(420), 모드 결정기(430), 장구간 특성 추출부(440) 및 프레임 특성 버퍼(450)를 포함한다.

도 4에 도시된 적응적 시간/주파수 기반 부호화 모드 결정 장치는 도 3에 도시된 부호화 모드 판단부(320)로 사용될 수 있다.

시간 영역 특성 추출부(410)는 입력 오디오 신호(IN)의 시간 영역 신호 분석을 수행하여 시간 영역 특성을 생성한다. 이 때, 시간 영역 특성은 특히 시간 영역 단구간 특성(time domain short-term feature)일 수 있다. 예를 들어, 시간 영역 단구간 특성은 천이(transient) 정도 및 단기/장기 예측 이득(prediction gain)의 크기 정도를 포함할 수 있다.

주파수 영역 특성 추출부(420)는 입력 오디오 신호(IN)의 주파수 영역 신호 분석을 수행하여 입력 오디오 신호(IN)의 하나의 프레임에 상응하는 주파수 영역이 복수 개의 주파수 영역으로 분할되어 생성된 주파수 밴드 각각에 상응하는 주파수 영역 특성을 생성한다. 이 때, 주파수 영역 특성 추출부(420)는 도 3에 도시된 주파수 영역 변환부(310)로부터 입력 오디오 신호(IN)의 주파수 영역 신호(S7)를 수신하여 주파수 분석을 할 수 있다. 이 때, 주파수 영역 특성은 주파수 영역 단구간 특성(frequency domain short-term feature)일 수 있다. 예를 들어, 주파수 영역 단구간 특성은 스펙트럼의 자기상관도(auto correlation)를 포함할 수 있다.

이 때, 주파수 영역 특성 추출부(420)가 입력 오디오 신호(IN)의 현재 프레임(current frame)에 대한 주파수 영역 신호 분석을 수행할 때, 시간 영역 특성 추출부(410)는 입력 오디오 신호(IN)의 현재 또는 다음 프레임(next frame)의 주파수 영역 신호 분석에 해당하는 시간 영역 신호 분석을 수행할 수 있다. 이 때, 주파수 영역 특성 추출부(420)는 현재 프레임과 함께 이전 프레임(previous frame)의 일부를 윈도우링(windowing)할 수도 있다.

장구간 특성 추출부(440)는 시간 영역 단구간 특성 및 주파수 영역 단구간 특성을 분석하여 시간 영역 장구간 특성 및 주파수 영역 장구간 특성을 생성한다.

이 때, 시간 영역 장구간 특성은 주기성(periodicity)의 연속성(continuity) 정도, 주파수 스펙트럼의 기울기(spectral tilt) 정도 및 프레임 에너지 정도를 포함할 수 있다. 이 때, 주기성의 연속성은 피치 랙(pitch lag)의 변화가 적고, 피치 상관도(pitch correlation)가 높은 프레임이 일정 구간 이상 연속적으로 지속되는 정도일 수 있다. 또한, 주기성의 연속성은 첫 번째 포먼트 주파수(1st formant frequency)가 매우 낮고, 피치 상관도가 높은 프레임이 일정 구간 이상 연속적으로 지속되는 정도일 수 있다.

이 때, 주파수 영역 장구간 특성은 채널간 상관도(correlation)를 포함할 수 있다.

프레임 특성 버퍼(450)는 시간 영역 특성 추출부(410)로부터 시간 영역 단구간 특성을 제공 받아 이를 저장한다. 따라서, 시간 영역 특성 추출부(410)가 다음 프레임(next frame)에 상응하는 시간 영역 단구간 특성을 출력할 때, 프레임 특성 버퍼(450)는 현재 프레임(current frame)에 상응하는 시간 영역 단구간 특성을 출력할 수 있다.

모드 결정기(430)는 상기 시간 영역 단구간 특성, 상기 주파수 영역 단구간 특성, 시간 영역 장구간 특성 및 상기 주파수 영역 장구간 특성을 이용하여, 상기 주파수 밴드 각각에 대한 부호화 모드를 시간 기반 부호화 모드 및 주파수 기반 부호화 모드 중 하나로 결정한다. 이 때, 모드 결정기(430)는 이전 프레임(previous frame), 현재 프레임(current frame) 및 다음 프레임(next frame)의 시간 영역 신호 분석 결과 및 이전 프레임(previous frame) 및 현재 프레임(current frame)의 주파수 영역 신호 분석 결과를 이용하여 주파수 밴드 별로 부호화 모드를 결정할 수 있다.

시간 기반 부호화 방식이 효과적인 경우의 예는, 선형 예측(Linear Prediction)으로 예측 이득(Prediction Gain)이 큰 경우 또는 음성 신호처럼 높은 피치를 가진 신호(highly pitched signal)가 입력 오디오 신호인 경우를 들 수 있다. 이에 반하여, 주파수 기반 부호화 방식이 효과적인 경우의 예는, 사인파 신호(sinusoidal signal), 입력 오디오 신호 중 부수적인 고주파 신호가 포함된 경우, 신호 간의 차폐현상(masking effect)이 큰 경우를 들 수 있다.

하기 표 1은 주파수 기반 부호화가 효율적인 입력 오디오 신호의 특성의 예를 나타낸 표이다.

[표 1]

	시간 영역 특성	주파수 영역 특성
단구간 특성	-천이 정도가 약한 신호 -단기/ 장기 예측 이득이 낮은 신호	-다중 밴드의 음성 확률(스펙트럼의 자기 상관도)가 낮은 신호
장구간 특성	-높은 주기성이 연속으로 장구간 유지되는 신호 -완만한 주파수 스펙트럼의 기울기를 갖고 높은 프레임 에너지를 갖는 신호	-스테레오 정도가 강한 신호(채널간 상관도가 낮은 신호)

하기 표 2는 시간 기반 부호화가 효율적인 입력 오디오 신호 특성의 예를 나타낸 표이다.

[표 2]

	시간 영역 특성	주파수 영역 특성
단구간 특성	-천이 정도가 강한 신호 -단기/ 장기 예측 이득이 높은 신호	-다중 밴드의 음성 확률(스펙트럼의 자기 상관도)가 높은 신호
장구간 특성	-연속적인 프레임에 대해 가파른 주파수 스펙트럼의 기울기를 갖고 선형 예측 필터의 스펙트럼 변화가 적은 신호	-스테레오 정도가 약한 신호(채널간 상관도가 높은 신호)

예를 들어, 모드 결정기(430)는 상기 시간 영역 단구간 특성, 상기 주파수 영역 단구간 특성, 시간 영역 장구간 특성 및 상기 주파수 영역 장구간 특성을 이용하여 상기 표 1의 조건에 가까운 경우에는 부호화 모드를 주파수 기반 부호화 모드로 결정하고, 상기 표 2의 조건에 가까운 경우에는 부호화 모드를 시간 기반 부호화 모드로 결정할 수 있다.

도 5는 도 4에 도시된 모드 결정기(430)의 동작을 나타낸 동작 흐름도이다.

도 5를 참조하면, 모드 결정기는 입력 오디오 신호의 스테레오 신호 정도가 소정 레벨 이상인지 여부를 판단한다(S510).

단계(S510)의 판단 결과, 입력 오디오 신호의 채널간 상관도가 낮아서 스테레오 신호 정도가 소정 레벨 이상인 경우에, 모드 결정기는 부호화 모드를 주파수 기반 부호화 모드로 결정한다(S570).

단계(S510)의 판단 결과, 입력 오디오 신호의 채널간 상관도가 높아서 스테레오 신호 정도가 소정 레벨 미만인 경우에, 모드 결정기는 입력 오디오 신호의 천이(transient) 정도가 소정 레벨 이상인지 여부를 판단한다(S520).

단계(S520)의 판단 결과, 입력 오디오 신호의 천이 정도가 소정 레벨 미만인 경우에, 모드 결정기는 부호화 모드를 주파수 기반 부호화 모드로 결정한다(S570).

단계(S520)의 판단 결과, 입력 오디오 신호의 천이 정도가 소정 레벨 이상인 경우에, 모드 결정기는 입력 오디오 신호의 장기/단기 예측 이득이 소정 레벨 이상인지 여부를 판단한다(S530).

단계(S530)의 판단 결과, 입력 오디오 신호의 장기/단기 예측 이득이 소정 레벨 미만인 경우에, 모드 결정기는 부호화 모드를 주파수 기반 부호화 모드로 결정한다(S570).

단계(S530)의 판단 결과, 입력 오디오 신호의 장기/단기 예측 이득이 소정 레벨 이상인 경우에, 모드 결정기는 해당 주파수 밴드에 상응하는 스펙트럼의 자기상관도가 소정 레벨 이상인지 여부를 판단한다(S540).

단계(S540)의 판단 결과, 해당 주파수 밴드에 상응하는 스펙트럼의 자기상관도가 소정 레벨 미만인 경우에, 모드 결정기는 부호화 모드를 주파수 기반 부호화 모드로 결정한다(S570).

단계(S540)의 판단 결과, 해당 주파수 밴드에 상응하는 스펙트럼의 자기상관도가 소정 레벨 이상인 경우에, 모드 결정기는 입력 오디오 신호의 주기성의 연속성이 소정 구간 이상 지속되는지 여부를 판단한다(S550). 이 때, 단계(S550)는 피치 랙(pitch lag)의 변화가 적고, 피치 상관도(pitch correlation)가 높은 프레임이 일정 구간 이상 연속적으로 지속되는지 여부 또는 첫 번째 포먼트 주파수(1st formant frequency)가 매우 낮고, 피치 상관도가 높은 프레임이 일정 구간 이상 연속적으로 지속되는지 여부를 판단할 수도 있다.

단계(S550)의 판단 결과, 입력 오디오 신호의 주기성의 연속성이 소정 구간 이상 지속되는 경우에, 모드 결정기는 부호화 모드를 주파수 기반 부호화 모드로 결정한다(S570).

단계(S550)의 판단 결과, 입력 오디오 신호의 주기성의 연속성이 소정 구간 이상 지속되지 아니하는 경우에, 모드 결정기는 주파수 스펙트럼의 기울기(spectral tilt)가 완만하고 높은 프레임 에너지가 일정 구간 이상 연속적으로 지속되는 정도인 음악 연속성(music continuity)이 소정 레벨 이상인지 여부를 판단한다(S560).

단계(S560)의 판단 결과, 주파수 스펙트럼의 기울기가 완만하고 높은 프레임 에너지가 일정 구간 이상 연속적으로 지속되는 정도가 소정 레벨 이상인 경우에, 모드 결정기는 부호화 모드를 주파수 기반 부호화 모드로 결정한다(S570).

단계(S560)의 판단 결과, 주파수 스펙트럼의 기울기가 완만하고 높은 프레임 에너지가 일정 구간 이상 연속적으로 지속되는 정도가 소정 레벨 미만인 경우에, 모드 결정기는 부호화 모드를 시간 기반 부호화 모드로 결정한다(S580).

도 6은 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 부호화 모드 결정 방법을 나타낸 동작 흐름도이다.

도 6을 참조하면, 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 부호화 모드 결정 방법은 입력 오디오 신호의 시간 영역 신호 분석을 수행하여 시간 영역 단구간 특성을 생성한다(S610).

이 때, 시간 영역 단구간 특성은 입력 오디오 신호의 천이(transient) 정도 및 단기/장기 예측 이득(prediction gain)의 크기 정도를 포함할 수 있다.

또한, 적응적 시간/주파수 기반 부호화 모드 결정 방법은 입력 오디오 신호의 주파수 영역 신호 분석을 수행하여 주파수 밴드 각각에 상응하는 주파수 영역 단구간 특성을 생성한다(S620).

이 때, 주파수 영역 단구간 특성은 스펙트럼의 자기상관도를 포함할 수 있다.

이 때, 단계(S620)가 입력 오디오 신호의 현재 프레임(current frame)에 대한 주파수 영역 신호 분석을 수행할 때, 단계(S610)는 입력 오디오 신호의 현재 또는 다음 프레임(next frame)의 주파수 영역 신호에 상응하는 시간 영역 신호 분석을 수행할 수 있다. 이 때, 단계(S620)는 현재 프레임과 함께 이전 프레임(previous frame)의 일부를 윈도우(windowing)할 수도 있다.

또한, 적응적 시간/주파수 기반 부호화 모드 결정 방법은 시간 영역 단구간 특성 및 상기 주파수 영역 단구간 특성을 분석하여 시간 영역 장구간 특성 및 주파수 영역 장구간 특성을 생성한다(S630).

이 때, 시간 영역 장구간 특성은 주기성(periodicity)의 연속성(continuity) 정도, 주파수 스펙트럼의 기울기(spectral tilt) 정도 및 프레임 에너지 정도를 포함할 수 있다. 이 때, 주기성의 연속성은 피치 락(pitch lag)의 변화가 적고, 피치 상관도(pitch correlation)가 높은 프레임이 일정 구간 이상 연속적으로 지속되는 정도일 수 있다. 또한, 주기성의 연속성은 첫 번째 포먼트 주파수(1st formant frequency)가 매우 낮고, 피치 상관도가 높은 프레임이 일정 구간 이상 연속적으로 지속되는 정도일 수 있다.

이 때, 주파수 영역 장구간 특성은 채널간 상관도(correlation)를 포함할 수 있다.

또한, 적응적 시간/주파수 기반 부호화 모드 결정 방법은 시간 영역 특성 및 주파수 영역 특성을 이용하여, 주파수 밴드 각각에 대한 부호화 모드를 시간 기반 부호화 모드 및 주파수 기반 부호화 모드 중 하나로 결정한다(S640).

이와 같은 과정을 통하여 적응적으로 시간 기반 부호화 모드 및 주파수 기반 부호화 모드 중 하나를 선택적으로 적용하여 부호화를 수행함으로써, 다양한 오디오 콘텐츠에 대하여 효율적인 부호화를 수행할 수 있고, 개루프(open loop) 방식으로 부호화 모드를 선택함으로써 폐루프(closed loop) 방식에 비하여 낮은 복잡도를 갖는 부호화기를 구현할 수 있다.

본 발명에 따른 적응적 시간/주파수 기반 부호화 모드 결정 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 본 발명을 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 상기 매체는 프로그램 명령, 데이터 구조 등을 지정하는 신호를 전송하는 반송파를 포함하는 광 또는 금속선, 도파관 등의 전송 매체일 수도 있다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 본 발명의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

이상과 같이 본 발명은 비록 한정된 실시예와 도면에 의해 설명되었으나, 본 발명은 상기의 실시예에 한정되는 것은 아니며, 본 발명이 속하는 분야에서 통상의 지식을 가진 자라면 이러한 기재로부터 다양한 수정 및 변형이 가능하다.

그러므로, 본 발명의 범위는 설명된 실시예에 국한되어 정해져서는 아니 되며, 후술하는 특허청구범위뿐 아니라 이 특허청구범위와 균등한 것들에 의해 정해져야 한다.

발명의 효과

본 발명의 적응적 시간/주파수 기반 오디오 부호화 장치 및 부호화 모드 결정 방법은 입력 오디오 신호에 대한 부호화 모드를 주파수 대역별로 판단하여 시간 기반 부호화 또는 주파수 기반 부호화함으로써, 두 가지 부호화 방식의 부호화 이득을 효율적으로 이용하여 고압축 성능을 얻을 수 있다.

또한, 본 발명은 입력 오디오 신호를 시간 영역 및 주파수 영역별로 장구간 특성 및 단구간 특성을 추출하여 주파수 대역별로 적절한 부호화 모드를 결정하도록 하여 적응적 시간/주파수 기반 오디오 부호화의 성능을 최적화할 수 있다.

또한, 본 발명은 개루프(open loop) 결정 방식을 사용하여 낮은 복잡도를 가지면서도 효과적으로 부호화 모드를 결정할 수 있다.

또한, 본 발명은 현재 프레임의 모드 결정시 다음 프레임의 특성을 반영함으로써, 모드가 한 프레임 간격으로 자주 스위칭 되는 것을 억제하여 모드 변화를 부드럽게 할 수 있다.

도면의 간단한 설명

도 1은 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 오디오 부호화 장치의 블록도이다.

도 2는 주파수 영역 변환된 신호의 분할 및 부호화 모드 결정 과정을 나타낸 개념도이다.

도 3은 도 1에 도시된 변환/모드 결정부의 일 예를 나타낸 블록도이다.

도 4는 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 부호화 모드 결정 장치의 블록도이다.

도 5는 도 4에 도시된 모드 결정기의 동작을 나타낸 동작 흐름도이다.

도 6은 본 발명의 일 실시예에 따른 적응적 시간/주파수 기반 부호화 모드 결정 방법을 나타낸 동작 흐름도이다.

<도면의 주요 부분에 대한 부호의 설명>

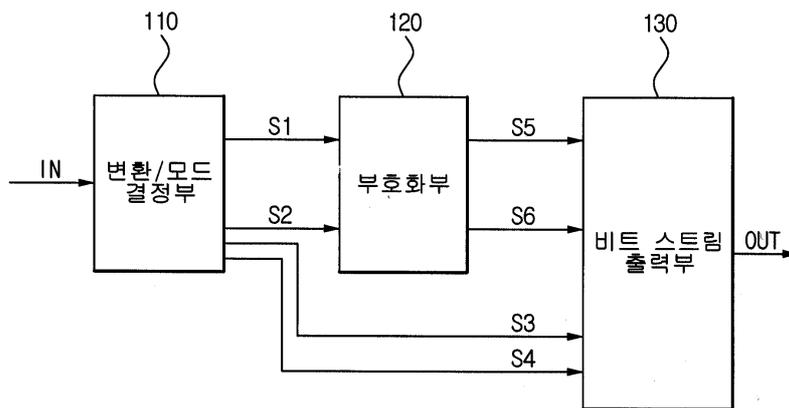
410: 시간 영역 특성 추출부 420: 주파수 영역 특성 추출부

430: 모드 결정기 440: 장구간 특성 추출부

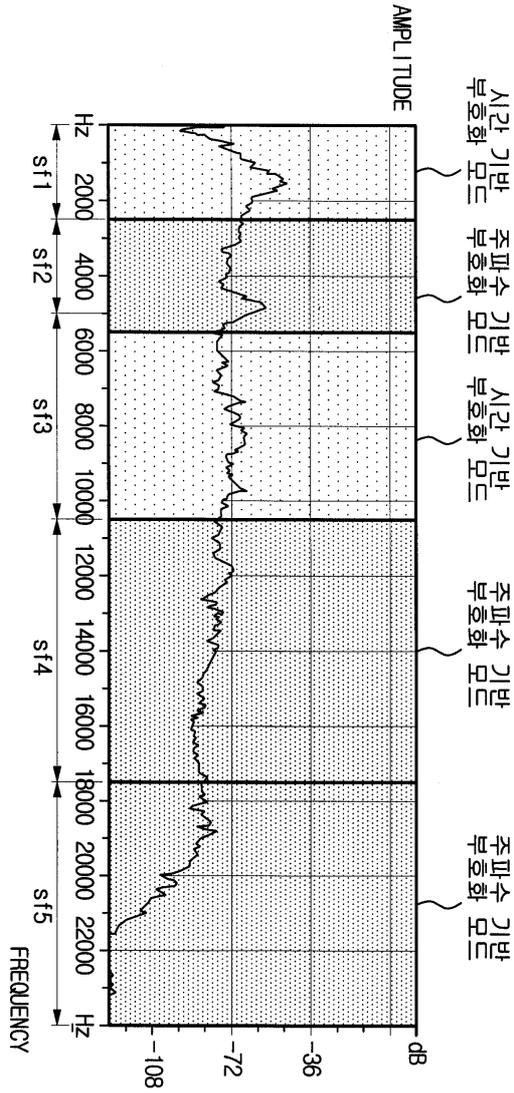
450: 프레임 특성 버퍼

도면

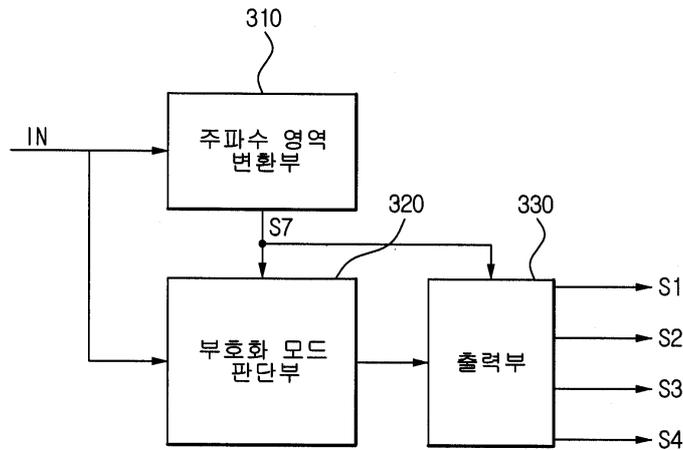
도면1



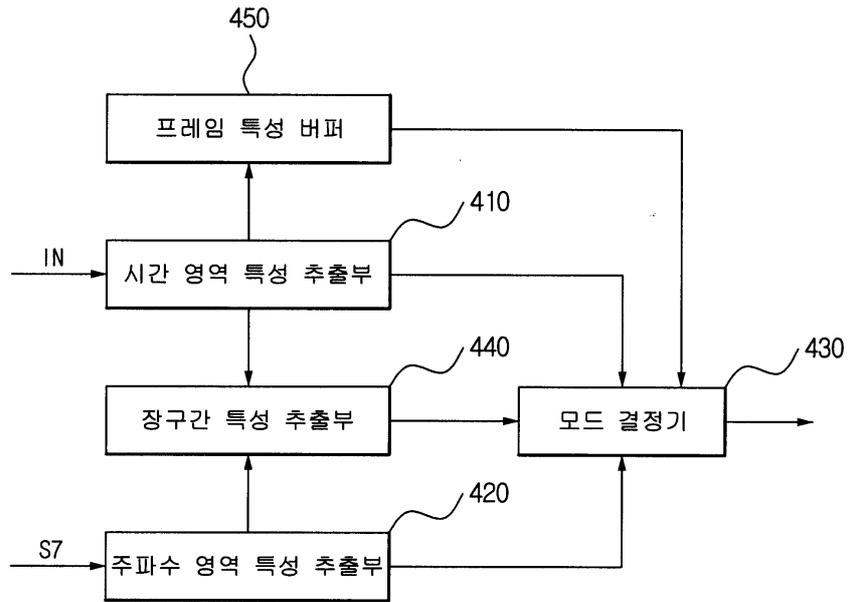
도면2



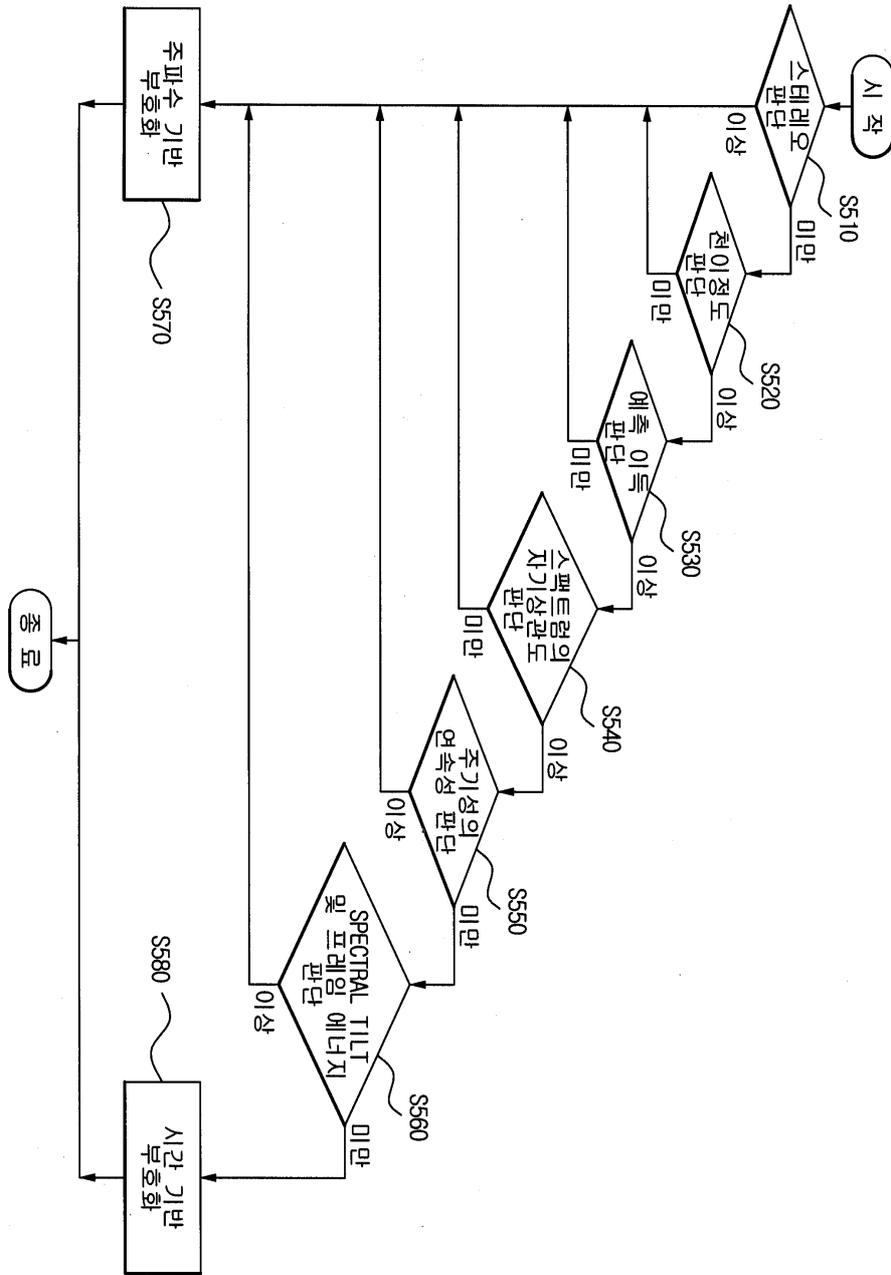
도면3



도면4



도면5



도면6

