

(19)대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。 Int. Cl. G11B 20/10 (2006.01)	(45) 공고일자 (11) 등록번호 (24) 등록일자	2006년08월02일 10-0608062 2006년07월26일
---	-------------------------------------	--

(21) 출원번호	10-2004-0061423	(65) 공개번호	10-2006-0012783
(22) 출원일자	2004년08월04일	(43) 공개일자	2006년02월09일

(73) 특허권자 삼성전자주식회사
 경기도 수원시 영통구 매탄동 416

(72) 발명자 오운학
 경기도 수원시 영통구 매탄3동 1242-6번지 302호

 이혁재
 서울특별시 서초구 서초2동 무지개아파트 1동 1107호

(74) 대리인 리앤목특허법인
 이혜영

(56) 선행기술조사문헌

KR1020030062338 A	KR1020030068716 A
US6115689 B	WO0252545 A1
WO9116769 A1	

* 심사관에 의하여 인용된 문헌

심사관 : 안병일

(54) 오디오 데이터의 고주파수 복원 방법 및 그 장치

요약

디코더내에서 mp3 압축 오디오 신호의 고주파수 복원 방법 및 그 장치가 개시되어 있다. 본 발명은 입력되는 비트스트림 으로부터 추출되는 MDCT 계수로부터 윈도우 타입에 따른 저주파수 영역의 필터 बैं크 값을 생성하는 과정, 윈도우 타입을 근거로 프레임의 트랜지언트(transient)정보를 추출하고, 그 트랜지언트 정보에 따른 웨이트 계수를 선택하는 과정, 저주 파수 영역의 필터 बैं크값으로부터 손실된 고주파수 영역의 필터 बैं크 값을 복원하는 과정, 선택된 웨이트 계수를 바탕으로 상기 과정에서 복원된 고주파수 성분들의 필터 बैं크 값을 조정하는 과정을 포함한다.

대표도

도 2

명세서

도면의 간단한 설명

도 1은 기존의 SBR 방식의 mp3PRO 디코딩 블록도이다.

도 2는 본 발명에 따른 고주파수 복원 방식을 적용한 mp3 디코더의 전체 블록도이다.

도 3a 내지 도 3d는 본 발명에 따른 고주파수 성분을 복원하는 과정을 보이는 그래프이다.

도 4는 본 발명에 따른 오디오 데이터 고주파수 복원 방법을 보이는 흐름도이다.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 오디오 압축/복호 시스템에 관한 것이며, 특히 디코더내에서 mp3 압축 오디오 신호의 고주파수 복원 방법 및 그 장치에 관한 것이다.

일반적으로 디지털 엠펙(MPEG:Moving Picture Experts Group) 오디오로 통칭되고 있는 디지털 오디오는 고품질,고능률 스테레오 부호화를 위한 국제 표준화기구(ISO/IEC)의 표준방식이다. 상기 엠펙 오디오는 엠펙 비디오와 조합됨으로써 고능률의 멀티미디어 정보압축을 실현 가능하게 하며, 최근에는 디지털 티브이(DTV),디브이디(DVD), 디지털 음악방송(DAB:Digital Audio Broadcasting) 및 엠펙-3 플레이어등 다양한 응용 제품이 등장하고 있는 상황이다. mp3오디오는 최근 널리 사용되고 있는 .mp3 확장자를 갖는 방식으로, 엠펙-1 오디오 계층 3의 방식으로 인코딩된 것을 의미한다. 또한, 엠펙 오디오의 압축원리는 인간의 감각특성을 이용해서 감도가 낮은 세부의 정보를 생략하여 부호량을 절감시키는 "지각 부호화(Perceptual Coding)" 방법을 이용한다.

그러나 mp3 오디오 데이터는 압축을 많이 할수록 고주파수 영역을 손실하게된다. 이러한 고주파 영역의 손실로 인해 음색이 바뀌고 명료도가 저하되며 억눌리거나 무딘 소리가 나게 된다. 따라서 손실 고주파수 성분들을 복원하기 위해 후처리 음질 개선을 적용한 SBR(SBR(Spectral Band Replication)방식의 mp3PRO 포맷을 이용하고 있다.

도 1은 기존의 SBR 방식의 mp3PRO 디코딩 블록도이다.

도 1을 참조하면, 디코더부(110)는 mp3PRO 비트스트림이 입력되면 시간 차원의 PCM 오디오 데이터와 보조 데이터로 디코딩한다. 이때 PCM 오디오 데이터는 좌측채널 오디오 데이터와 우측채널 오디오 데이터로 분리되고, 보조 데이터는 엔벨로프(envelope) 정보를 포함한다. QMF 분석부(120)는 PCM 오디오 데이터를 32 밴드들의 저주파수 영역의 신호로 변환한다. 고주파수 발생부(130)는 QMF 분석부(120)에서 변환된 저주파수 영역의 성분들과 하모닉(harmonic)한 관계를 가지도록 엔벨로프 정보에 따른 고주파수 성분들을 생성한다. 엔벨로프 조정부(140)는 저주파수 영역의 스펙트럼을 이용하여 엔벨로프 정보에 따라 고주파수 성분들의 에너지를 조정한다. QMF 합성부(150)는 엔벨로프 조정부(140)에서 조정된 고주파수 성분들의 에너지와 QMF 분석부(120)에서 분석된 저주파수 영역의 신호를 합성하여 고주파수 성분이 복원된 시간차원의 오디오 데이터를 출력한다. 채널분리부(160)는 디코더(110)에서 발생하는 보조 데이터에 따라 좌측 채널과 우측 채널을 분리한 오디오 데이터를 출력한다.

결국, 기존에서 디코더부(110)에서 디코딩된 mp3 오디오 데이터는 후처리 장치들 즉, QMF 분석부(120), 고주파발생부(130), 엔벨로프조절부(140), QMF 합성부(150)에 의해 고주파수 성분이 복원된다. 따라서 SBR(Spectral Band Replication)방식은 후처리를 이용함으로써 다음과 같은 두 가지 문제점이 있다.

첫째로, 디코딩된 파일을 주파수 영역으로 변환하고 존재하는 주파수 성분들로부터 고주파 성분들을 추정한다. 추정된 고주파 성분들은 다시 시간 차원으로 변환되고 디코딩 파일에 더해져서 출력된다. 기존의 SBR(Spectral Band Replication) 방식의 mp3 디코딩 방법은 시간 차원에서 주파수 차원으로, 주파수 차원에서 시간 차원으로 변환하는 두가지 과정이 필요하게 된다. 따라서 기존의 SBR(Spectral Band Replication)방식의 mp3 디코딩 방법은 차원 변환 과정에서 과도한 계산량이 요구된다.

두 번째로, SBR(Spectral Band Replication)방식의 mp3PRO 디코더는 주파수 차원에서 고주파수 영역을 복원하기 위해 인코더단에서 구한 스펙트럼 엔벨로프(spectrum envelope) 정보를 이용하므로 기존의 mp3 인코더가 그대로 사용되지 않고 수정된다. 즉, SBR(Spectral Band Replication)방식의 mp3PRO 디코더는 기존의 mp3 파일에 대해 고주파수 성분들을 복원할 수 없다.

발명이 이루고자 하는 기술적 과제

본 발명이 이루고자하는 기술적 과제는 mp3 디코딩 과정에서 손실된 고주파수 성분들을 복원함으로써 기존의 오디오 코덱 방식에 의해 손실된 고주파수 성분들로 인해 저하된 원음의 음색을 재생하고 명료도를 높이는 오디오 데이터의 고주파수 복원 방법을 제공하는 데 있다.

본 발명이 이루고자하는 다른 기술적 과제는 오디오 데이터의 고주파수 복원 방법을 적용한 오디오 데이터의 고주파수 복원 장치를 제공하는 데 있다.

상기의 기술적 과제를 해결하기 위하여, 본 발명은 압축 오디오 신호의 고주파수 성분 복원 방법에 있어서,

- (a) 입력되는 비트스트림으로부터 추출되는 MDCT 계수로부터 윈도우 타입에 따른 저주파수 영역의 필터 बैं크 값을 생성하는 과정;
- (b) 상기 윈도우 타입을 근거로 프레임의 트랜지언트(transient)정보를 추출하고, 그 트랜지언트 정보에 따른 웨이트 계수를 선택하는 과정;
- (c) 상기 생성된 저주파수 영역의 필터 बैं크값으로부터 손실된 고주파수 영역의 필터 बैं크 값을 복원하는 과정;
- (d) 상기 과정에서 선택된 웨이트 계수를 바탕으로 상기 과정에서 복원된 고주파수 성분들의 필터 बैं크 값을 조정하는 과정을 포함하는 것을 특징으로 한다.

상기의 다른 기술적 과제를 해결하기 위하여, 본 발명은 압축 오디오 신호의 고주파수 성분 복원 장치에 있어서,

입력되는 압축 오디오 비트 스트림을 역양자화하여 MDCT를 추출하는 역 양자화부;

상기 역 양자화부 추출된 MDCT 계수로부터 저주파수 영역의 필터 बैं크 값을 생성하는 역 MDCT부;

상기 역 MDCT부에서 사용하는 윈도우 타입을 근거로 프레임의 트랜지언트(transient)정보를 추출하고, 그 트랜지언트 정보를 바탕으로 고주파수 성분들의 크기를 조정하는 웨이트 계수를 선택하는 웨이트 계수 추출부;

상기 역 DCT부에서 생성된 저주파수 영역의 필터 बैं크값으로부터 고주파수 영역의 필터 बैं크값을 복원하는 고주파수 영역 생성부;

상기 웨이트 계수 추출부에서 선택된 웨이트 계수를 상기 고주파수영역 생성부에서 복원된 고주파수 영역의 필터 बैं크 값에 곱하는 곱셈부를 포함하는 것을 특징으로 한다.

발명의 구성 및 작용

이하 첨부된 도면을 참조하여 본 발명의 바람직한 실시예를 설명하기로 한다.

먼저, 본 발명에 따른 mp3 디코더에 입력되는 mp3 비트 스트림은 다음과 같은 과정을 통해 형성된다. 먼저, PCM(Pulse coded Modulation) 형태의 오디오 데이터를 입력한다. 이어서, 입력된 PCM 오디오 데이터를 각 그라눌(granule)마다 576개 샘플들로 나눈다. 이어서, 그 샘플들에 대해서 MPEG1 레이어3(MP3)에서 심리 음향 모델을 적용하여 지각 에너지(perceptual energy)를 구한다. 이어서, 심리음향 모델에서 구해진 지각 에너지와 임계치를 비교하여 윈도우잉을 스위칭하면서 MDCT를 수행한다. MDCT 윈도우 사이즈의 일부 또는 전부는 임계치에 따라 스위칭 될 수 있다. 즉, 지각 에너지의 레벨이 임계치보다 크면 에너지 레벨이 급격하게 증가하는 어택(attack) 상태 신호에 해당하므로 쇼트 윈도우를 선택하고, 임계치보다 적으면 에너지 레벨이 일정한 상태의 신호에 해당하므로 롱 윈도우를 선택하며, 이어서 그 선택된 각 윈도우

우 범위에 해당하는 오디오 샘플들을 MDCT 처리하여 주파수 도메인상의 데이터로 변환한다. 이때 스타트 윈도우 또는 스톱 윈도우는 롱 윈도우에서 쇼트 윈도우로 스위칭하기 위해 사용된다. 또한 윈도우잉 타입은 MPEG 1 레이어3에서 롱 윈도우, 시작 윈도우, 쇼트 윈도우, 스톱 윈도우등으로 개시되어 있다. 그리고 각 윈도우들은 엘리머싱을 방지하기 위해 서로 오버랩된다. 이어서, MDCT가 수행된 주파수 도메인상의 데이터를 할당된 비트수에 따라 양자화한다. 이어서, 양자화된 데이터를 호프만 코딩(huffman coding)을 이용하여 mp3 비트스트림을 형성한다. 이때 mp3 비트스트림은 프레임 단위로 형성된다. mp3 프레임 포맷은 헤더, 사이드 정보 및 메인 데이터로 이루어진다. 사이드 정보는 스케일 팩터, 윈도우 타입과 같이 메인 데이터를 디코드하기 위한 필요 정보를 포함한다.

도 2는 본 발명에 따른 고주파수 복원 방식을 적용한 mp3 디코더의 전체 블록도이다.

도 2의 mp3 디코더는 역양자화부(210), 사이드 정보 분석부(220), 역MDCT부(230), 고주파 영역 분석부(250), 고주파 영역 생성부(260), 웨이트 계수 추출부(240), 곱셈부(270), 합성부(280), 역 필터뱅크부(290)로 구성되며, 웨이트 계수 추출부(240)는 트랜지언트 정보 검출부(242), 웨이트 테이블 선택부(244)를 구비한다.

먼저, 역양자화부(210)는 입력되는 mp3 비트 스트림으로부터 MDCT(Modified DCT) 계수를 추출한다. 이때 역양자화된 MDCT 계수는 고주파수 대역이 손실된 저주파수 대역으로 분포한다.

사이드 정보 분석부(220)는 입력되는 mp3 비트 스트림으로부터 사이드 정보(side information)를 분석하여 윈도우 타입을 추출한다.

역MDCT부(230)는 주파수역양자화부(210)에서 추출된 MDCT 계수로부터 사이드 정보 분석부(220)에서 추출된 윈도우 타입을 이용하여 필터뱅크값(filter bank value)을 생성한다.

트랜지언트 정보 검출부(242)는 역MDCT부(230)에서 사용하던 윈도우 타입으로부터 현재 프레임의 트랜지언트 정보를 검출한다. 즉, 윈도우 타입이 롱(long)인 경우 현재 프레임은 논-트랜지언트(non-transient) 영역이고, 상기 윈도우 타입이 쇼트(short)인 경우 현재 프레임은 트랜지언트(transient) 영역이고, 상기 윈도우 타입이 스타트(start) 또는 엔드(end)인 경우 현재 프레임이 트랜지션(transition) 영역이다.

웨이트 테이블 선택부(244)는 트랜지언트 정보 검출부(242)에서 검출된 트랜지언트 정보로부터 고주파수 성분들의 웨이트를 조정하기 위한 웨이트 계수를 선택한다. 예컨대, 트랜지언트(transient) 영역에서는 높은 웨이트를 갖는 고주파 성분(harmonic component), 논-트랜지언트(non-transient) 영역에서는 낮은 웨이트를 갖는 고주파 성분(harmonic component), 트랜지션(transition) 영역에서는 중간 웨이트를 갖는 고주파 성분을 갖는다.

고주파수 영역 분석부(250)는 역MDCT부(230)에서 생성된 필터뱅크값을 분석하여 손실된 고주파수 영역을 검출한다. 예를 들어 도 3a를 보면, 96kbps의 mp3 파일인 경우 32개의 필터뱅크값들중 11.025kHz이상의 주파수 성분들이 손실된다. 128kbps의 mp3 파일인 경우 15kHz 32개의 필터뱅크값들중 15kHz이상의 주파수 성분들이 손실된다.

고주파수 영역 생성부(260)는 고주파수 영역 분석부(250)에서 손실된 고주파수 성분들을 복원한다. 도 3b를 참조하여 96kbps의 mp3 파일에 대해서 설명하면, 32개의 필터뱅크값들중 11.025kHz이상의 주파수 성분들이 손실되므로 "0"값을 가지는 16번째 बैं크 이상의 필터뱅크값들을 8 - 15번째 필터뱅크 값들로부터 복원해야한다. 예컨대, 16번째 बैं드는 8번째 बैं드와 하모닉(harmonic) 관계를 가지므로 8번째의 필터뱅크값이 복사된다. 마찬가지로 18번째 बैं드는 9번째 बैं드의 필터뱅크값이 복사된다. 또한 사람의 인지 특성상 고주파수 영역에서 같은 주파수라고 인지하는 대역폭이 넓어지므로 19번째 बैं드는 복원된 18번째 बैं드를 복사한다. 이때 음성은 6kHzs이내의 주파수 성분을 갖는다. 음성이 포함된 저주파 성분들로부터 고주파 성분들을 생성할 경우 고주파수 영역에 음성에 해당하는 주파수 성분들이 나타나는 문제점이 있다. 따라서 5.5kHz이내의 저주파수 영역의 1 - 7번째 필터뱅크값은 고주파수 복원용으로 이용되지 않는다.

곱셈부(270)는 웨이트 테이블 선택부(244)에서 선택된 웨이트 계수를 고주파수 성분들에 곱하여 도 3c 및 도 3d의 그래프와 같은 고주파수 성분들의 크기를 조정한다. 도 3c는 현재 프레임이 트랜지언트 영역인 경우 복원된 고주파 성분들을 보이는 그래프이다. 도 3c를 보면, 트랜지언트 영역에서는 높은 웨이트들을 갖는 고주파 성분들이 생성된다. 도 3d는 현재 프레임이 논-트랜지언트 영역인 경우 복원된 고주파 성분들을 보이는 그래프이다. 도 3d를 보면, 논-트랜지언트 영역에서는 낮은 웨이트들을 갖는 고주파 성분들이 생성된다.

합성부(280)는 역 MDCT부(230)에서 생성된 저주파수 영역의 필터뱅크 값과 곱셈부(270)에서 생성되는 고주파수 영역의 필터뱅크 값을 합성한다.

역다상필터뱅크부(290)는 합성부(280)에서 고주파수 성분이 복원된 필터 뱅크 값들을 서브밴드로 통합한 후, 이 통합된 서브밴드를 합성 필터로 통과시켜 PCM(Pulse Coded Modulation)오디오 데이터로 복원한다.

도 4는 본 발명에 따른 오디오 데이터 고주파수 복원 방법을 보이는 흐름도이다.

먼저, 프레임 단위의 mp3 비트스트림을 입력한다(410).

이때 입력되는 압축 오디오 비트 스트림을 역양자화하여 MDCT를 추출한다(420). 동시에 사이드 정보(side information)를 분석하여 윈도우 타입을 추출한다.

이어서, MDCT 계수를 윈도우 타입에 따라 역 MDCT하여 저주파수 영역의 필터 뱅크 값을 생성한다(430). 이때 윈도우 타입을 근거로 프레임의 트랜지언트(transient)정보를 추출하고(424), 그 트랜지언트 정보를 바탕으로 고주파수 성분들의 크기를 조정하는 웨이트 계수를 계수 테이블에서 선택한다(426).

이어서, 저주파수영역의 필터 뱅크값을 분석하여 손실된 고주파수 영역을 검출한다(440).

이어서, 저주파수 영역의 필터 뱅크값으로 부터 고주파수 영역의 필터 뱅크값을 복원한다(450).

이어서, 계수 테이블에서 선택된 웨이트 계수를 복원된 고주파수 영역의 필터 뱅크 값에 곱하여 고주파수 성분의 크기를 조정한다(460).

이어서, 역 MDCT를 통해 생성된 저주파수 영역의 필터 뱅크 값과 조정된 고주파수 영역의 필터 뱅크 값을 합성한다(470).

이어서, 고주파수 성분이 복원된 필터 뱅크 값들을 서브밴드로 통합한 후, 이 통합된 서브밴드를 합성 필터로 통과시켜 PCM(Pulse Coded Modulation)오디오 데이터로 복원한다(480).

본 발명은 상술한 실시예에 한정되지 않으며, 본 발명의 사상내에서 당업자에 의한 변형이 가능함은 물론이다. 즉, 본 발명은 MP3재생기, 노트북 PC 등과 같은 오디오를 재생하는 모든 기기에 오디오 데이터의 고주파수 성분을 복원하는 기술에 적용할 수 있다.

발명의 효과

상술한 바와 같이 본 발명에 의하면, 기존의 mp3 인코더에 대한 수정이 전혀 필요없고, mp3 디코딩 과정에서 손실된 고주파수 성분들을 복원함으로써 기존에 사용하던 차원 변환이 필요하지 않으므로 적은 계산량으로 mp3 음질을 개선할 수 있다.

(57) 청구의 범위

청구항 1.

압축 오디오 신호의 고주파수 성분 복원 방법에 있어서,

(a) 입력되는 비트스트림으로부터 추출되는 MDCT 계수로부터 윈도우 타입에 따른 저주파수 영역의 필터 뱅크 값을 생성하는 과정;

(b) 상기 윈도우 타입을 근거로 프레임의 트랜지언트(transient)정보를 추출하고, 그 트랜지언트 정보에 따른 웨이트 계수를 선택하는 과정;

(c) 상기 생성된 저주파수 영역의 필터 뱅크값으로부터 손실된 고주파수 영역의 필터 뱅크 값을 복원하는 과정;

(d) 상기 (b)과정에서 선택된 웨이트 계수를 바탕으로 상기 과정에서 복원된 고주파수 성분들의 필터 बैं크 값을 조정하는 과정을 포함하는 오디오 데이터의 고주파수 복원 방법.

청구항 2.

제1항에 있어서, (b) 과정은

(b-1) 역 MDCT에서 사용되는 윈도우 타입을 참조하여 현재 프레임에 대한 트랜지언트 정보를 추출하는 과정;

(b-2) 상기 (b-1)과정에서 추출된 트랜지언트 정보에 따라 상기 고주파 성분의 필터 बैं크값의 웨이트를 조정하는 웨이트 계수를 소정의 계수 테이블로부터 선택하는 과정을 구비하는 것을 특징으로 하는 오디오 데이터의 고주파수 복원 방법.

청구항 3.

제2항에 있어서, 상기 트랜지언트 정보는 트랜지언트 영역 정보, 논-트랜지언트 영역 정보, 트랜지션 영역 정보임을 특징으로 하는 오디오 데이터의 고주파수 복원 방법.

청구항 4.

제2항에 있어서, 상기 윈도우가 롱(long) 타입인 경우 현재 프레임은 논-트랜지언트(non-transient) 영역이고, 상기 윈도우가 쇼트(short) 타입인 경우 현재 프레임은 트랜지언트(transient) 영역이고, 상기 윈도우가 스타트(start) 또는 엔드(end)타입인 경우 현재 프레임이 트랜지션(transition) 영역임을 특징으로 하는 오디오 데이터의 고주파수 복원 방법.

청구항 5.

제1항에 있어서, 상기 (c)과정은 상기 트랜지언트 정보에 따라 선택된 웨이트 계수와 고주파 성분들의 필터 बैं크값을 곱하는 것임을 특징으로 하는 오디오 데이터의 고주파수 복원 방법.

청구항 6.

압축 오디오 신호의 고주파수 성분 복원 장치에 있어서,

입력되는 압축 오디오 비트 스트림을 역양자화하여 MDCT를 추출하는 역 양자화부;

상기 역 양자화부 추출된 MDCT 계수로부터 저주파수 영역의 필터 बैं크 값을 생성하는 역 MDCT부;

상기 역 MDCT부에서 사용하는 윈도우 타입을 근거로 프레임의 트랜지언트(transient)정보를 추출하고, 그 트랜지언트 정보를 바탕으로 고주파수 성분들의 크기를 조정하는 웨이트 계수를 선택하는 웨이트 계수 추출부;

상기 역 DCT부에서 생성된 저주파수 영역의 필터 बैं크값으로부터 고주파수 영역의 필터 बैं크값을 복원하는 고주파수 영역 생성부;

상기 웨이트 계수 추출부에서 선택된 웨이트 계수를 상기 고주파수영역 생성부에서 복원된 고주파수 영역의 필터 बैं크 값에 곱하는 곱셈부를 포함하는 오디오 데이터의 고주파수 복원 장치.

청구항 7.

제6항에 있어서, 상기 역 MDCT부에서 생성된 주파수 영역의 필터 뱅크 값과 상기 곱셈부에서 생성되는 고주파수 영역의 필터 뱅크 값을 합성하는 합성부를 더 포함하는 것을 특징으로 하는 오디오 데이터의 고주파수 복원 장치.

청구항 8.

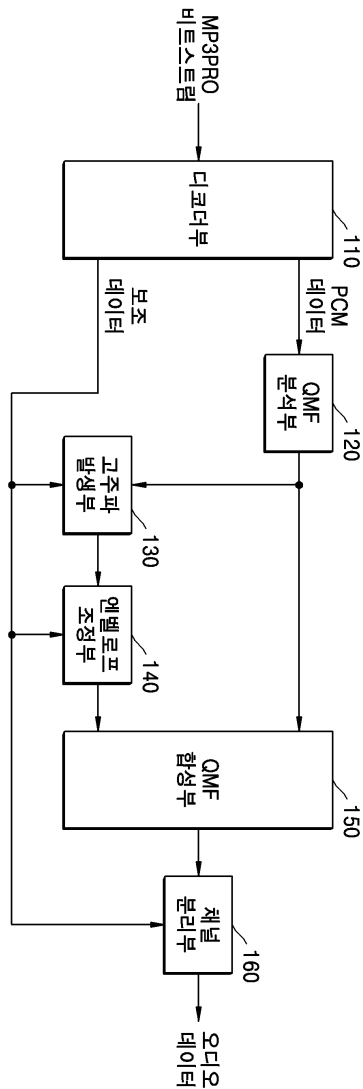
제6항에 있어서, 상기 웨이트 계수 추출부는

상기 역 MDCT에서 사용되는 윈도우 타입으로부터 현재 프레임에 대한 트랜지언트정보를 검출하는 트랜지언트 정보 검출부;

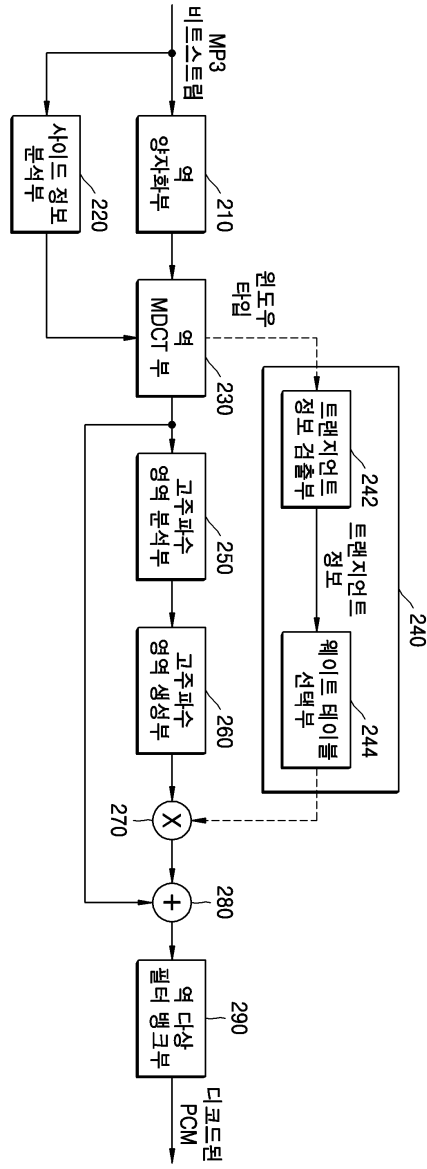
상기 트랜지언트 정보 검출부에서 검출된 트랜지언트 정보에 상응하는 웨이트를 미리 정해진 계수 테이블로부터 선택하는 웨이트 계수 선택부를 포함하는 것을 특징으로 하는 오디오 데이터의 고주파수 복원 장치.

도면

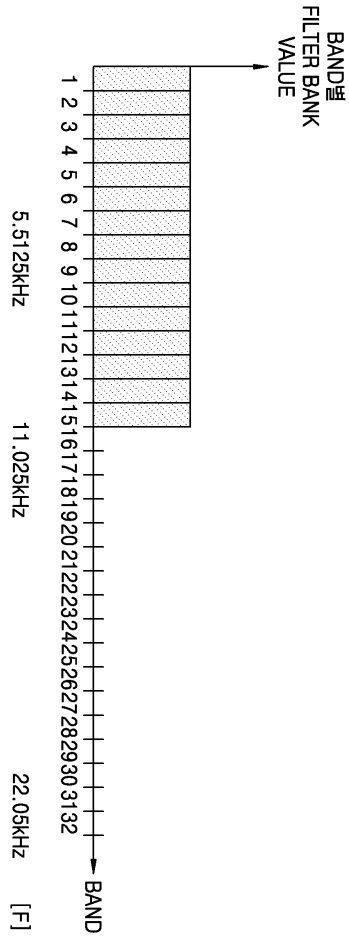
도면1



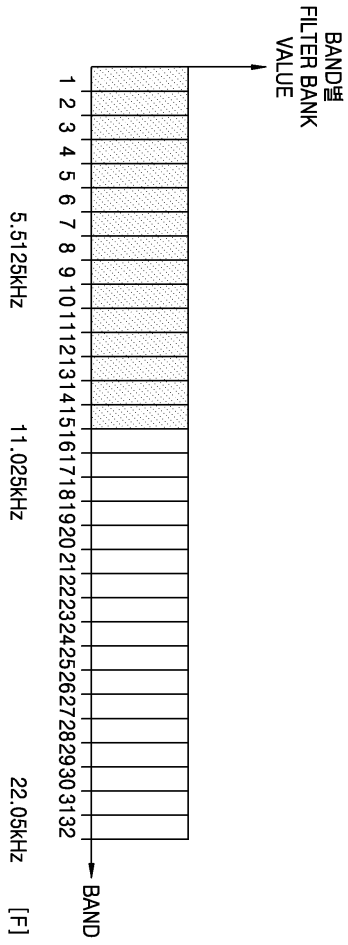
도면2



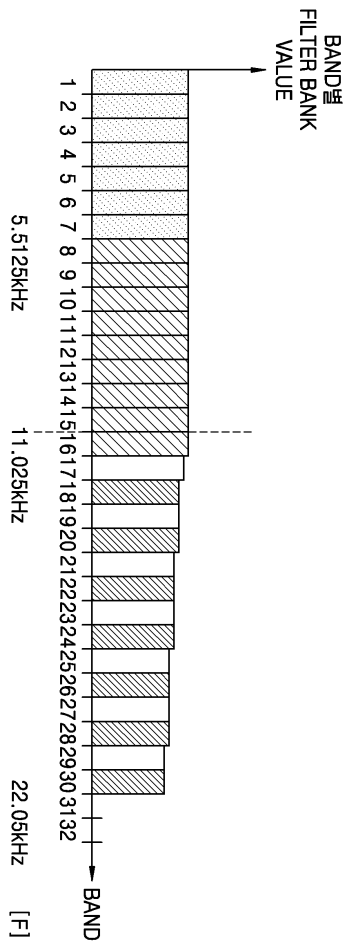
도면3a



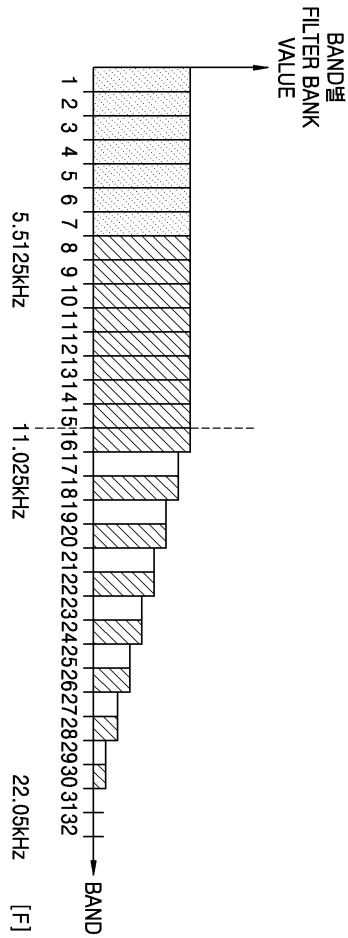
도면3b



도면3c



도면3d



도면4

