



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2011-0036906
(43) 공개일자 2011년04월12일

(51) Int. Cl.

G10L 19/14 (2006.01)

(21) 출원번호 10-2011-7000728
(22) 출원일자(국제출원일자) 2009년06월26일
심사청구일자 2011년01월11일
(85) 번역문제출일자 2011년01월11일
(86) 국제출원번호 PCT/EP2009/004652
(87) 국제공개번호 WO 2010/003564
국제공개일자 2010년01월14일

(30) 우선권주장
08017663.9 2008년10월08일
유럽특허청(EPO)(EP)
(뒷면에 계속)

(71) 출원인

프라온호퍼 게젤샤프트 쭈르 피르데롱 데어 안겐
반텐 포르숨 에. 베.

독일 80686 뮌헨 한자슈트라쎄 27 체

보이세지 코포레이션

캐나다 퀘백 에이치3알 2에이치6 빌레 몬트-로얄
케민 루션 750 스위트 250

(72) 발명자

그릴, 번하드

독일 91207 라우프, 피터-헨레인슈트라쎄 7

르퀘브르, 로흐

캐나다 제이1케이 5알9 퀘백, 칸톤 데 마곡, 259
애비뉴 데 라 불가드

(뒷면에 계속)

(74) 대리인

윤의섭, 김수진

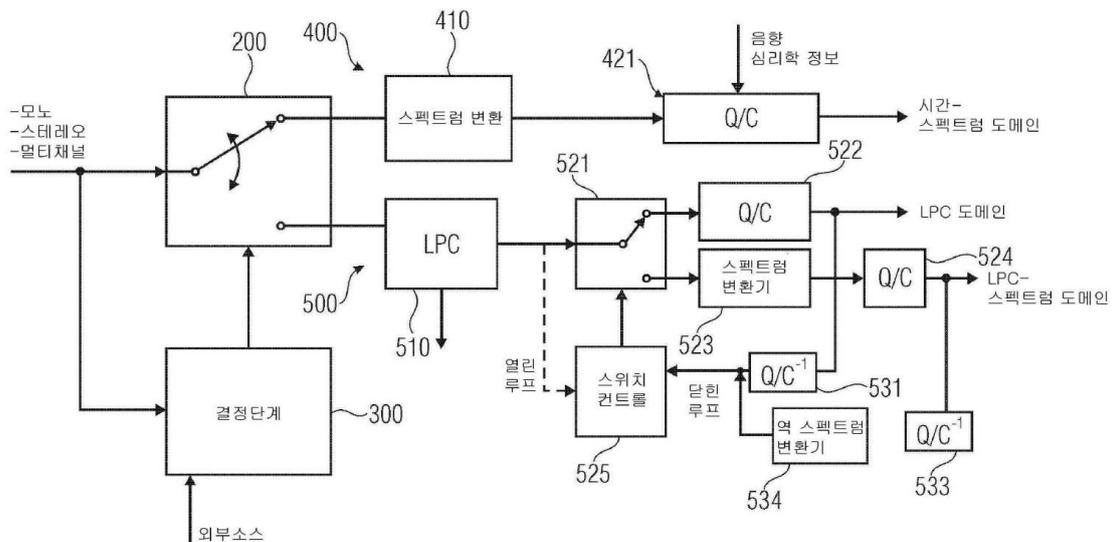
전체 청구항 수 : 총 25 항

(54) 캐스케이드 된 스위치를 구비하는 저 비트레이트 오디오 인코딩/디코딩 기법

(57) 요약

스펙트럼 도메인 인코딩 브랜치같은 제1 정보 싱크 기원 인코딩 브랜치, LPC-도메인 인코딩 브랜치같은 제2 정보 소스 또는 SNR 기원 인코딩 브랜치 및 상기 제1 인코딩 브랜치와 상기 제2 인코딩 브랜치 간의 전환을 수행하는 스위치를 포함하고, 여기서 상기 제2 인코딩 브랜치는 여자 신호를 생성하기 위한 LPC 분석 스테이지같은 스펙트럼 도메인과 다른 특수한 도메인으로의 변환기를 포함하고, 상기 제2 인코딩 브랜치는, 추가적으로 LPC 도메인 처리 브랜치같은 특수한 도메인 코딩 브랜치 및 LPC 스펙트럼 도메인 프로세싱 브랜치같은 특수한 스펙트럼 도메인 코딩 브랜치를 포함하며 상기 특수한 도메인 코딩 브랜치 및 상기 특수한 스펙트럼 도메인 코딩 브랜치 간의 전환을 위한 스위치를 포함한다. 오디오 디코더는, 스펙트럼 도메인 디코딩 브랜치같은 제1 도메인 디코더, 제2 도메인에서 여기 신호같은 신호를 디코딩 하기 위한 LPC 도메인 디코딩 브랜치같은 제2 도메인 디코더, LPC 스펙트럼 디코더 브랜치같은 제3 도메인 디코더 및 상기 디코더 간의 전환을 위한 두 개의 캐스케이드된 스위치를 포함한다.

대표도



(72) 발명자

베셋, 브루노

캐나다 제이1엔 4지5 퀘백, 서브룩, 1600 루 무틸로

라피에르, 지미

캐나다 제이1지 5에이5 퀘백, 서브룩, 2-1793, 루데 폰텐블로

구르네이, 필립

캐나다 제이1엘 0에이2 퀘백, 서브룩, 3012 루 드사우비그논

살라미, 레드완

캐나다 에이치4알 2와이3 퀘백, 생로랑, 4045 알버트 드룩스 플레이스

바이어, 스테판

독일 90425 뉘른베르크, 도트문더슈트라쎬 14

푸흐스, 구일라우메

독일 90409 뉘른베르크, 파크슈트라쎬 12

귀일스버거, 스테판

독일 97076 뷔르츠부르크, 오토-로쓰 - 슈트라쎬 90

가이거, 랄프

독일90409 뉘른베르크, 막스토그라벤 29

힐퍼트, 요한네스

독일90411 뉘른베르크, 헨후테슈트라쎬 46

크라에머, 올리흐

독일 70195 슈투트가르트, 밀로에커슈트라쎬 5

르콩트, 제레미

독일 90489 뉘른베르크, 술츠바허슈트라쎬 39

몰트루스, 마르쿠스

독일 90469 뉘른베르크 에츠라우프웨그 7

네우엔들프, 맥스

독일 90402 뉘른베르크, 테아터가쎬 17

포프, 하랄드

독일 90587 투첸바흐, 오버미켈바허슈트라쎬 18

레텔바흐, 리콜라우스

독일 90427 뉘른베르크, 스페쎬트슈트라쎬 38

(30) 우선권주장

09002271.6 2009년02월18일

유럽특허청(EPO)(EP)

61/079,854 2008년07월11일 미국(US)

특허청구의 범위

청구항 1

제1 도메인에 존재하는 오디오 입력 신호(195)를 인코딩하는 오디오 인코더에 있어서,

제1 인코딩된 신호를 얻기 위하여 제1 코딩 알고리즘을 사용하여 오디오 신호를 인코딩하는 제1 코딩 브랜치(400);

제2 인코딩된 신호를 얻기 위하여, 상기 제1 코딩 알고리즘과는 다른 제2 코딩 알고리즘을 사용하여 오디오 신호를 인코딩하는 제2 인코딩 브랜치(500); 및

오디오 입력 신호의 일부에 대하여 상기 제1 인코딩된 신호 또는 상기 제2 인코딩된 신호가 인코더 출력 신호가 되도록, 상기 제1 코딩 브랜치 및 상기 제2 코딩 브랜치 간의 전환을 수행하는 제1 스위치(200);

를 포함하고,

상기 제2 코딩 브랜치는,

오디오 신호를, 상기 제1 도메인과 다른 제1 도메인으로 변환하는 컨버터(510);

제1 처리된 신호를 구하기 위해 제2 도메인 내에서 오디오 신호를 처리하는 제1 처리 브랜치(522);

제2 처리된 신호를 얻기 위하여, 신호를 상기 제1 도메인 및 상기 제2 도메인과 다른 제3 도메인으로 변환하고, 상기 제3 도메인에서 신호를 처리하는 제2 처리 브랜치(523, 524); 및

상기 제2 코딩 브랜치로 입력되는 오디오 신호의 일부에 대하여, 상기 처리된 신호 또는 상기 제2 처리된 신호가 상기 제2 인코딩된 신호 내에 있도록 하기 위하여 상기 제1 처리 브랜치(522) 및 상기 제2 처리 브랜치(523, 524) 간의 전환을 수행하는 제2 스위치(521)인 것을 특징으로 하는 오디오 인코더.

청구항 2

청구항 1에 있어서,

상기 제1 코딩 브랜치(400) 내의 제1 코딩 알고리즘은 정보 싱크 모델(information sink model)을 기반으로 하거나 또는 상기 제2 코딩 브랜치(500) 내의 상기 제2 코딩 알고리즘은 정보 소스 또는 SNR(signal to noise ratio) 모델을 기반으로 하는 것을 특징으로 하는 오디오 인코더.

청구항 3

청구항 1 또는 2에 있어서,

상기 제1 코딩 브랜치는, 오디오 입력 신호를, 상기 제1 도메인, 상기 제2 도메인 및 상기 제3 도메인과는 다른 제4 도메인으로 변환하는 컨버터(410);를 포함하고 있는 것을 특징으로 하는 오디오 인코더.

청구항 4

청구항 1 내지 3에 있어서,

상기 제1 도메인은 시간 도메인이고, 상기 제2 도메인은 상기 제1 도메인 신호를 LPC 필터링(filtering)함으로써 얻어진 LPC 도메인이고, 상기 제3 도메인은 LPC 필터링된 신호를 스펙트럼 도메인으로 변환하여 얻어진 LPC 스펙트럼 도메인이고, 상기 제4 도메인은 상기 제1 도메인 신호를 변환하는 주파수 도메인에 의하여 얻어진 스펙트럼 도메인인 것을 특징으로 하는 오디오 인코더.

청구항 5

청구항 1 내지 4에 있어서,
 상기 제1 스위치(200) 또는 신호 적용 방법에 있어서 상기 제2 스위치(521)를 제어하는 제어기(300, 525);
 를 더 포함하고,
 상기 제어기는,
 상기 제1 스위치(200)에 입력되거나 또는 상기 제1 코딩 브랜치에 의해 출력되는 신호 또는 타겟 함수에 대하여
 상기 제1 코딩 브랜치 또는 제2 코딩 브랜치의 출력 신호를 디코딩함으로써 얻어진 신호를 분석하거나, 또는
 상기 제2 스위치(200)에 입력되거나 또는 상기 제1 제어 브랜치에 의해 출력되는 신호 또는 타겟 함수에 대하여
 상기 제1 제어 브랜치 또는 제2 제어 브랜치의 출력 신호를 역처리함으로써 얻어진 신호를 분석하는 것을 특징
 으로 하는 오디오 인코더.

청구항 6

청구항 1 내지 5에 있어서,
 제1 코딩 브랜치(400) 또는 제2 코딩 브랜치(500)의 제2 처리 브랜치(523, 524)는,
 에일리어싱(aliasing) 도입 시간/주파수 컨버터를 안내하는 및 양자화기/엔트로피 코더 스테이지(421)를 포함
 하고,
 여기서 상기 제2 코딩 브랜치 내의 상기 제1 처리 브랜치는, 에일리어싱 도입 변환을 포함하지 않는 양자화기
 또는 엔트로피 인코더를 포함하고 있는 것을 특징으로 하는 오디오 인코더.

청구항 7

청구항 6에 있어서,
 상기 에일리어싱 도입 시간/주파수 컨버터는 분석창을 적용하기 위한 윈도우어(windower) 및 MDCT(modified
 discrete cosine transform) 알고리즘을 포함하고,
 상기 윈도우어는, 상기 윈도우어 내부로 입력되는 신호의 샘플이 적어도 두 개의 이어지는 프레임(frame) 내에
 발생하도록 오버래핑의 방법으로 이어지는 프레임에 윈도우 함수를 적용하는 작용을 하는 것을 특징으로 하
 는 오디오 인코더.

청구항 8

청구항 1 내지 7에 있어서,
 제1 처리 브랜치(522)는, ACELP(algebraic code excited linear prediction) 코더의 LPC 여자 코딩을 포함하는
 제1 처리 브랜치(522) 및 MDCT 스펙트럼 컨버터 및 양자화된 스펙트럼 요소를 구하기 위해 스펙트럼 요소를 양
 자화하기 위한 양자화기를 포함하는 제2 처리 브랜치를 포함하고,
 상기 각각의 양자화된 스펙트럼 요소는 0이거나 또는 다수의 양자화 인덱스에 대한 하나의 양자화 인덱스에 의
 해 정의되는 것을 특징으로 하는 오디오 인코더.

청구항 9

청구항 5에 있어서,
 상기 제어기는, 열린 루프 방식에서 제1 스위치(200)를 제어하고, 닫힌 루프 방식에서는 제2 스위치(521)를 제

어하기 위해 작동하는 것을 특징으로 하는 오디오 인코더.

청구항 10

청구항 1 내지 9에 있어서,

상기 제1 코딩 브랜치 및 상기 제2 코딩 브랜치는, 블록에 관한 방식에서 오디오 신호를 인코딩하고,

상기 제1 스위치 및 상기 제2 스위치는, 최소한 기정의된 신호의 샘플이 대응되는 스위치(521, 200)에 대하여 프레임 길이를 형성한 후에 전환 행위를 하도록, 블록에 관한 방식에서 전환을 행하는 것을 특징으로 하는 오디오 인코더.

청구항 11

청구항 10에 있어서,

상기 제1 스위치에 대한 프레임 길이는 적어도 상기 제2 스위치에 대한 프레임 길이의 두 배인 것을 특징으로 하는 오디오 인코더.

청구항 12

청구항 5에 있어서,

상기 제어기는, 상기 제1 스위치에 대해서 프레임의 50% 이하의 부분은 발화이고, 상기 제1 스위치에 대해서 상기 프레임의 50% 이상의 부분은 음악일 때조차 발화에 대한 결정을 하도록, 음악에 대한 결정에 대하여 발화에 대한 결정이 선호되는 방법으로, 발화/음악을 차별하도록 작동하는 것을 특징으로 하는 오디오 인코더.

청구항 13

청구항 5 또는 12에 있어서,

제2 스위치에 대한 프레임은, 제1 스위치에 대한 프레임보다 작고, 상기 제어기 (525, 300)은, 제2 프레임의 길이의 50% 이상보다 긴 길이를 구비한 제1 프레임의 부분이 음악을 포함하고 있음이 발견되었을 때 발화에 대한 결정을 하도록 작동하는 것을 특징으로 하는 오디오 인코더.

청구항 14

청구항 1 내지 13에 있어서,

제1 인코딩 브랜치(400) 또는 제2 코딩 브랜치의 제2 제어 브랜치는 다양한 시간 워핑(warping) 기능을 포함하고 있는 것을 특징으로 하는 오디오 인코더.

청구항 15

제1 도메인 내의 오디오 입력 신호(195)를 인코딩하는 방법에 있어서,

제1 인코딩된 신호를 얻기 위하여 제1 코딩 알고리즘을 사용하여 오디오 신호를 인코딩(400)하는 단계;

제2 인코딩된 신호를 얻기 위하여 상기 제1 코딩 알고리즘과 다른 제2 코딩 알고리즘을 사용하여 오디오 신호를 인코딩(500)하는 단계; 및

상기 오디오 입력 신호에 대하여, 제1 인코딩된 신호 또는 제2 인코딩된 신호가 인코딩된 출력 신호 이도록, 상기 제1 코딩 알고리즘을 사용한 인코딩 및 상기 제2 코딩 알고리즘을 사용한 인코딩 간에 전환(switching)을 수

행하는 단계;
 를 포함하고,
 상기 제2 코딩 알고리즘을 사용한 인코딩(500)은,
 상기 오디오 신호를 제1 도메인과는 다른 제2 도메인으로 변환하는 단계(510);
 제1 처리된 신호를 얻기 위하여 상기 제2 도메인 내에서 오디오 신호가 처리되는 단계(522);
 신호를 상기 제1 도메인 및 상기 제2 도메인과는 다른 제3 도메인으로 변환하는 단계;(523)
 제2 처리된 신호를 얻기 위하여 상기 제3 도메인 내에서 신호를 처리하는 단계;(524) 및
 오디오 신호의 처리(522) 및 변환(523) 및 처리(524) 간의 전환을 행함으로써 상기 제2 코딩 알고리즘을 사용하여 인코딩된 오디오 신호의 일부에 있어서, 제1 처리된 신호 또는 제2 처리된 신호가 제2 인코딩된 신호가 되도록 하는 단계;
 를 포함하는 것을 특징으로 하는 인코딩 방법.

청구항 16

제1 코딩된 신호, 제2 도메인 내의 제1 처리된 신호 및 제3 도메인 내의 제2 처리된 신호를 포함하는 인코딩된 오디오 신호를 디코딩하고, 상기 제1 코딩된 신호, 상기 제1 처리된 신호 및 상기 제2 처리된 신호는 디코딩된 오디오 신호의 다른 시간 부분에 관련되어 있고, 제1 도메인, 상기 제2 도메인 및 상기 제3 도메인은 서로 간에 다른 디코더에 있어서,
 제1 코딩 알고리즘에 의한 제1 인코딩된 신호를 디코딩하는 제1 디코딩 브랜치(431, 440);
 상기 제1 처리된 신호 또는 상기 제2 처리된 신호를 디코딩하는 제2 디코딩 브랜치; 및
 제1 도메인에서 변환된 출력 신호를 얻기 위하여 제1 도메인에서 변환된 신호 및 상기 제1 디코딩 브랜치에 의해 출력되는 제1 디코딩된 신호를 결합하는 제2 결합자(combiner);
 를 포함하고,
 상기 제2 디코딩 브랜치는,
 상기 제2 도메인 내에서 제1 역 처리된 신호를 얻기 위하여 제1 처리된 신호를 역처리하는 제1 역처리 브랜치(531);
 상기 제2 도메인 내에서 제2 역처리된 신호를 구하기 위하여 상기 제2 처리된 신호를 역처리하는 제2 역처리 도메인(533, 534);
 상기 제2 도메인 내에서 결합된 신호를 얻기 위하여 상기 제1 역처리된 신호 및 상기 제2 역처리된 신호를 결합하는 제1 결합자(532); 및
 결합된 신호를 상기 제1 도메인으로 변환하는 컨버터(540);
 를 포함하는 것을 특징으로 하는 디코더.

청구항 17

청구항 16에 있어서,
 상기 제1 결합자(532) 및 상기 제2 결합자(600)는 크로스페이드(cross fade) 기능을 구비한 스위치를 포함하고 있는 것을 특징으로 하는 디코더.

청구항 18

청구항 16 또는 17에 있어서,

상기 제1 도메인은 시간 도메인이거나, 상기 제2 도메인은 LPC 도메인이거나, 상기 제3 도메인은 LPC 스펙트럼 도메인이거나 또는 상기 제1 도메인에서 신호를 시간/주파수 변환을 통해 얻어진 시간 스펙트럼 도메인인 제4 도메인 내에서 상기 제1 인코딩된 신호가 인코딩되는 것을 특징으로 하는 디코더.

청구항 19

청구항 1 내지 18에 있어서,

상기 제1 디코딩 브랜치(431, 440)는 역코더, 역양자화기(de-quantizer) 및 주파수 도메인 시간 도메인 컨버터(440)를 포함하거나 또는

상기 제2 디코딩 브랜치는, 상기 제1 역처리 브랜치 내에 역코더(inverse coder) 및 역양자화기를 포함하거나 또는 상기 제2 역처리 브랜치 내에 역코더 및 역양자화기 및 LPC 도메인 컨버터에 대한 LPC 스펙트럼 도메인을 포함하고 있는 것을 특징으로 하는 디코더.

청구항 20

청구항 19에 있어서,

상기 제1 디코딩 브랜치 또는 상기 제2 역처리 브랜치는, 시간 도메인 엘리어싱 취소 기능을 수행하기 위해 오버랩 추가자(overlap adder)를 포함하고 있는 것을 특징으로 하는 디코더.

청구항 21

청구항 16 내지 20에 있어서,

제1 디코딩 브랜치 또는 제2 역처리 브랜치는 상기 인코딩된 오디오 신호 내에 포함된 워핑된 특징에 의하여 제어되는 역와퍼(de-warper)를 포함하고 있는 것을 특징으로 하는 디코더.

청구항 22

청구항 16 내지 21에 있어서,

상기 인코딩된 신호는 사이드 정보(4a)로써, 코딩된 신호가 제1 인코딩 브랜치, 제2 인코딩 브랜치, 제2 인코딩 브랜치의 제1 처리 브랜치 또는 제2 인코딩 브랜치의 제2 처리 브랜치 중 어느 브랜치에 의해서 코딩되었는지에 대한 지시를 포함하고,

또한 추가적으로, 상기 사이드 정보(4a)를 기초로 하여, 코딩된 신호가, 제1 디코딩 브랜치, 제2 디코딩 브랜치, 제2 디코딩 브랜치의 제1 역처리 브랜치 또는 제2 디코딩 브랜치의 제2 역처리 브랜치 중 어느 브랜치에서 처리될 것인지를 결정하는 인코딩된 신호를 파싱하는 파서(parser)를 포함하고 있는 것을 특징으로 하는 디코더.

청구항 23

제1 코딩된 신호, 제2 도메인 내의 제1 처리된 신호 및 제3 도메인 내의 제2 처리된 신호를 포함하는 인코딩된 오디오 신호를 디코딩하며, 상기 제1 코딩된 신호, 상기 제1 처리된 신호 및 상기 제2 처리된 신호는 디코딩된 오디오 신호의 서로 다른 시간 부분에 관련되어 있고, 제1 도메인, 상기 제2 도메인 및 상기 제3 도메인은 서로 다른 디코딩 방법에 있어서,

제1 코딩 알고리즘에 기반하여 제1 인코딩된 신호를 디코딩하는 단계(431, 440);

제1 처리된 신호 또는 제2 처리된 신호를 디코딩하는 단계; 및
 상기 제1 도메인 내에서 디코딩된 출력 신호를 구하기 위하여, 상기 제1 도메인 내의 변환된 신호 및 디코딩된 제1 신호를 결합하는 단계;
 를 포함하고,
 상기 제1 처리된 신호 또는 상기 제2 처리된 신호를 디코딩하는 단계는,
 상기 제2 도메인 내에서 제1 역처리 신호를 구하기 위해 상기 제1 처리된 신호를 역처리하는 단계;(531)
 상기 제2 도메인 내에서 제2 역처리 신호를 구하기 위해서 상기 제2 처리된 신호를 역처리하는 단계;(533, 534)
 및
 상기 제2 도메인 내에서 결합된 신호를 얻기 위해 상기 제1 역처리된 신호 및 상기 제2 역처리된 신호를 결합하는 단계;
 를 포함하는 것을 특징으로 하는 디코딩 방법.

청구항 24

제1 코딩 알고리즘을 사용하여 인코딩 또는 디코딩되는 제1 코딩된 신호;
 제2 도메인 내의 제1 처리된 신호, 제3 도메인 내의 제2 처리된 신호; 및
 인코딩된 신호가 상기 제1 코딩된 신호, 상기 제1 처리된 신호 또는 상기 제2 처리된 신호인지 여부를 지시하는 사이드 정보(4a);
 를 포함하고,
 상기 제1 처리된 신호 및 상기 제2 처리된 신호는 제2 코딩 알고리즘을 이용하여 인코딩되고,
 상기 제1 코딩된 신호, 상기 제1 처리된 신호 및 상기 제2 처리된 신호는, 디코딩된 오디오 신호의 다른 시간 부분에 관련되고,
 제1 도메인, 상기 제2 도메인 및 상기 제3 도메인은 서로 다른 것을 특징으로 하는 인코딩된 오디오 신호.

청구항 25

컴퓨터에서 실행될 때, 청구항 15에 따른 오디오 신호를 인코딩하는 방법 또는 청구항 23에 따라 인코딩된 오디오 신호를 디코딩하는 방법을 수행하는 컴퓨터 프로그램.

명세서

기술분야

[0001] 본 발명은 오디오 인코딩 및 특히 저 비트레이트 오디오 인코딩 기법에 관한 것이다.

배경기술

[0002] 종래 기술에 있어서 MP3 또는 AAC와 같은 주파수 영역 인코딩 기법(frequency domain coding scheme)이 알려져 있다. 이러한 주파수 영역 인코더는 시간영역/주파수영역 전환, 이후 양자화 스테이지(이 스테이지에서는 상기 양자화 오류는 음향 심리학적 모듈의 정보를 이용하여 제어된다) 및 인코딩 스테이지(이 스테이지에서는 상기 양자화된 스펙트럼 변수 및 대응변 정보는 코드 테이블을 이용하여 엔트로피 인코딩된다)에 기초를 두고 있다.

[0003] 다른 한편으로는 3GPP TS 26.290.에서 설명된 바와 같이, 적응성 다중 레이트 광대역(AMR-WB+)같은 음성처리에 최적인 인코더도 있다. 이러한 음성 코딩 기법은 시간영역 신호에 대한 선형 예측 필터링(linear predictive

filtering, 이하 LP필터링)을 수행한다. 상기 LP 필터링은 상기 입력 시간 영역 신호에 대한 선형 예측 분석에서 과생된 것이다. 그 후 상기 생성된 LP 필터링의 계수값은, 수치화/인코딩되고, 사이드 정보로써 전송된다. 상기 과정은 선형 예측 인코딩 방법(Linear Prediction coding)으로 알려져 있다. 상기 필터의 출력에서, 여파 신호로 알려진 상기 예측 잔류 신호 또는 예측 오류 신호는, 상기 ACELP(Algebraic Code Excited Linear Predictive) 인코더의 합성에 의한 분석 스테이지를 통하여 인코딩되거나, 또는 이와는 다르게, 중첩에 의한 푸리에 변환을 사용하는 변환 인코더를 사용하여 인코딩된다. 폐쇄 루프 알고리즘을 사용하거나 또는 개방 루프 알고리즘을 사용함으로써 상기 ACELP 인코딩 및 상기 TCX(Transform Coded eXcitation) 인코딩 중 어느 하나가 결정된다.

[0004] 고품질 인코딩 기법 및 스펙트럼 대역 복사 기술을 조합하는 고효율-고품질 디코딩 기법 같은 주파수 영역 인코딩 기법은, 또한 조인트 스테레오 또는 "MPEG 서라운드"라는 명칭으로 알려진 다중채널 인코딩 도구가 조합된 것일 수도 있다.

[0005] 한편으로는, 적응성 다중 레이트 광대역 같은 음성 인코더는 또한 고주파 향상 스테이지 및 스테레오 가능성을 구비하고 있다.

[0006] 주파수 영역 인코딩 기법은, 음악 신호에 있어서 낮은 비트레이트에서 높은 품질을 보여준다는 점에서 장점이 있다. 그러나 낮은 비트레이트에서 음성 신호의 품질에 있어서는 문제가 있다.

[0007] 음성 인코딩 기법은 낮은 비트레이트에서도 음성 신호에 있어서 높은 품질을 보여준다. 그러나 낮은 비트레이트에서의 음악 신호에 대해서는 낮은 품질을 보여준다.

발명의 내용

해결하려는 과제

[0008] 본 발명의 목적은, 발전된 인코딩/디코딩 개념을 제공하는 것에 있다.

과제의 해결 수단

[0009] 이러한 목적은, 청구항 1에 부합되는 오디오 인코더, 청구항 15에 부합되는 오디오 인코딩 방법, 청구항 16에 부합되는 디코더, 청구항 23에 부합되는 디코딩 방법, 청구항 24에 부합되는 인코딩된 신호 또는 청구항 25에 부합되는 컴퓨터 프로그램을 통하여 달성될 수 있다.

[0010] 본 발명의 일실시예에 있어서 제1 도메인에 위치하는 음성 입력 신호를 인코딩하는 음성 인코더는, 제1 인코딩된 신호를 구하기 위한 제1 인코딩 알고리즘을 사용하여 음성 신호를 인코딩하는 제1 인코딩 브랜치; 상기 제1 인코딩 알고리즘과는 차별되고, 제1 인코딩된 신호를 구하기 위한 제2 인코딩 알고리즘을 사용하여 음성 신호를 인코딩하는 제2 인코딩 브랜치; 및 상기 음성 입력 신호에 대하여, 상기 제1 인코딩된 신호 또는 상기 제2 인코딩된 신호 중 어느 하나가 인코더 출력 신호에 포함되도록, 상기 제1 인코딩 브랜치 및 상기 제2 인코딩 브랜치 간의 전환을 수행하는 제1 스위치; 를 포함하고, 상기 제2 인코딩 브랜치는, 상기 음성 신호를, 상기 제1 도메인과는 다른 제2 도메인으로 변환하여 입력하는 컨버터(converter); 제1 처리된 신호를 얻기 위해 상기 제2 도메인 내에서 음성 신호를 처리하는 제1 처리 브랜치; 상기 제1 도메인 및 상기 제2 도메인과는 차별되는 제3 도메인으로 신호를 변환하여 입력하고 제2 처리된 신호를 구하기 위하여 상기 제3 도메인에서 상기 신호를 처리하는 제2 처리 브랜치; 및 상기 제2 인코딩 브랜치에 입력되는 음성 신호에 대하여, 상기 제1 처리된 신호 또는 상기 제2 처리된 신호 중 어느 하나가 상기 제2 인코딩된 신호에 포함되도록, 상기 제1 처리 및 상기 제2 처리

브랜치 간의 전환을 수행하는 제2 스위치; 를 포함하는 것을 특징으로 한다.

[0011] 본 발명의 다른 일례는 인코딩된 음성 신호를 디코딩하는 디코더로, 상기 인코딩된 음성 신호는 제1 인코딩된 신호, 제2 도메인에서 제1 처리된 신호 및 제3 도메인에서의 제2 처리된 신호이고, 여기서 상기 제1 인코딩된 신호, 제1 처리된 신호 및 상기 제2 처리된 신호는, 디코딩된 오디오 신호의 다른 시간 부분(time portion)에 관련되어 있고, 여기서 제1 도메인, 상기 제2 도메인 및 상기 제3 도메인은 서로 동일하지 않은 것을 특징으로 하고, 상기 디코더는, 제1 인코딩 알고리즘에 기반하여 제1 인코딩된 신호를 디코딩하는 제1 디코딩 브랜치; 상기 제1 처리된 신호 또는 상기 제2 처리된 신호를 디코딩하고, 상기 제2 도메인에서 제1 역처리된 신호를 구하기 위하여 상기 제1 처리된 신호를 역처리하는 제1 역처리 브랜치를 포함하는 제2 디코딩 브랜치; 상기 제2 도메인에서 제2 역처리된 신호를 구하기 위한 상기 제2 처리된 신호를 역처리하기 위한 제2 역처리 브랜치; 상기 제2 도메인에서 결합된 신호를 구하기 위해 상기 제1 역처리된 신호 및 상기 제2 역처리된 신호를 결합하는 제1 결합자(combiner); 상기 결합된 신호를 상기 제1 도메인으로 변환하는 컨버터; 및 상기 제1 도메인에서 디코딩된 출력 신호를 구하기 위하여 상기 제1 도메인의 상기 변환된 신호 및 상기 제1 디코딩 브랜치에 의해 출력된 상기 디코딩된 제1 신호를 결합하는 제2 결합자;를 포함하는 것을 특징으로 한다.

[0012] 본 발명의 최신 실시예에 있어서, 두 개의 스위치는 순차적인 순서로 제공된다. 여기서 제1 스위치는 주파수 도메인 인코더를 사용하여 스펙트럼 도메인 내에서 코딩할 것인지, 아니면 예를 들어 LPC 분석 스테이지에서의 출력 시의 신호를 처리하는 것 같이 LPC 도메인 내에서 코딩할 것인지를 결정한다. 상기 제2 스위치는, ACELP 코더를 사용하거나 또는 LPC 스펙트럼 도메인 내에서 상기 LPC 도메인 신호를 코딩하는 것과 같이 상기 LPC 도메인 내에서 상기 LPC 도메인 신호를 인코딩하기 위한 상기 LPC 도메인 내에서의 전환의 기능을 제공한다. 여기서 상기 LPC 스펙트럼 도메인은, 상기 LPC 도메인 신호를, 스펙트럼 도메인과는 차별되는 LPC 스펙트럼 도메인으로 변환하기 위한 컨버터를 요구한다. 왜냐하면, 상기 LPC 스펙트럼 도메인은 상기 시간 도메인 신호의 상기 스펙트럼보다는 LPC 필터링 된 신호의 스펙트럼을 보여주기 때문이다.

[0013] 상기 제1 스위치는 두 개의 처리 브랜치 중 어느 하나를 결정하는데, 여기서 하나의 처리 브랜치는 주로 청각 마스킹 같은 싱크모델 및/또는 심리 음향 모델에 의하여 자극받는다. 그리고 다른 처리 브랜치는 주로 소스 모델 및 세그먼트 신호 대 잡음비(Segmental SNR) 계산에 의해서 주로 자극받는다. 전형적으로 하나의 처리 브랜치는 주파수 도메인 인코더를 구비하고 다른 처리 브랜치는 발성 코더 같이 LPC에 기반한 인코더를 구비하고 있다. 상기 소스 모델은 통상 상기 발성 처리이고, 그러므로 LPC은 일반적으로 사용된다.

[0014] 상기 제2 스위치는 다시, 상기 "외부의" 제1 처리 브랜치 도메인과는 다른 도메인 내에서 두 개의 처리 브랜치 중 어느 하나를 결정한다. 다시 하나의 "내부의" 처리 브랜치는 주로 소스 모델 또는 신호 대 잡음비 계산에 의하여 자극받는다. 그리고 다른 "내부의" 처리 브랜치는, 싱크 모델 및/또는 음향 심리학 모델 즉, 마스킹에 의하여 영향받거나 또는 적어도 주파수/스펙트럼 도메인 코딩 측면을 포함한다. 전형적으로 하나의 "내부의" 처리 브랜치는 주파수 도메인 인코더/스펙트럼 컨버터를 구비하고, 다른 처리 브랜치는, LPC 도메인같은 다른 도메인에서 코딩하는 인코더를 구비한다. 여기서 이러한 인코더는 예를 들어, 스펙트럼 변환 없이 입력 신호를 처리하는 CELP 또는 ACELP 양자화기/스케일러이다.

[0015] 더욱 바람직하게는, 오디오 인코더는 스펙트럼 도메인 인코딩 브랜치 같은 제1 정보 싱크 기반의 인코딩 브랜치, LPC 도메인 인코딩 브랜치같은 제2 정보 소스 또는 SNR 기반의 인코딩 브랜치 및 상기 제1 인코딩 브랜치 및 상기 제2 인코딩 브랜치 사이에서 전환하는 스위치를 포함한다. 여기서 상기 제2 인코딩 브랜치는 여자(勵磁) 신호를 생성하는 LPC 분석 스테이지 같은 시간 도메인과는 차별되는 특정 도메인으로서의 컨버터를 포함한다. 그리고 상기 제2 인코딩 브랜치는, 추가적으로 LPC 도메인 처리 브랜치같은 특정 도메인 및 LPC 스펙트럼 도메인 처리 브랜치 같은 특정 스펙트럼 도메인을 포함할 수 있다. 그리고 추가적으로 상기 특정 도메인 인코딩 브랜치 및 상기 특정 스펙트럼 도메인 인코딩 브랜치를 서로 전환시키는 스위치를 포함한다.

[0016] 본 발명의 다른 일실시예에 있어서 오디오 디코더는, 스펙트럼 도메인 디코딩 브랜치의 제1 도메인, 도메인 내에서 여자 신호 등의 신호를 디코딩하기 위한 LPC 도메인 인코딩 브랜치 등의 제2 도메인 및 LPC 스펙트럼 도메인 등의 도메인 내의 여자 신호 등의 신호를 디코딩하기 위한 LPC 스펙트럼 디코더 같은 제3 도메인 등을 포함하는 것을 특징으로 한다. 여기서 상기 제3 도메인은 상기 제2 도메인으로부터 주파수 변환을 수행하여 구해지고, 여기서 상기 제2 도메인 신호 및 상기 제3 도메인 신호를 위한 제1 스위치가 제공되며, 상기 제1 도메인 디코더 및 상기 제2 도메인 또는 상기 제3 도메인에 대한 상기 디코더 간의 변환을 위한 제2 스위치가 제공된다.

발명의 효과

[0017] 이와 같은 방법을 통해서, 인코딩 및 디코딩의 개념이 개량되었다.

도면의 간단한 설명

[0018] 도 1a는, 본 발명의 일실시예에 있어서 첫 번째 관점에 따르는 인코딩 기법에 대한 블록도이다.
 도 1b는, 본 발명의 일실시예에 있어서 첫 번째 관점에 따르는 디코딩 기법에 대한 블록도이다.
 도 1c는, 본 발명의 일실시예에 있어서 다른 관점에 따르는 인코딩 기법에 대한 블록도이다.
 도 2a는, 본 발명의 일실시예에 있어서 두 번째 관점에 따르는 인코딩 기법에 대한 블록도이다.
 도 2b는, 본 발명의 일실시예에 있어서 두 번째 관점에 따르는 디코딩 기법에 대한 블록도이다.
 도 2c는, 본 발명의 일실시예에 있어서 다른 관점에 따르는 인코딩 기법에 대한 블록도이다.
 도 3a는, 본 발명의 일실시예에 있어서 다른 관점에 따르는 인코딩 기법에 대한 블록도를 도시한 것이다.
 도 3b는, 본 발명의 일실시예에 있어서 다른 관점에 따르는 인코딩 기법에 대한 블록도를 도시한 것이다.
 도 3c는, 캐스캐이드된 스위치를 이용한 인코딩 장치/방법을 도식적으로 표현한 것이다.
 도 3d는, 캐스캐이드된 결합자가 사용되는, 디코딩을 위한 장치 또는 방법에 대한 도식적인 도면을 도시한 것이다.
 도 3e는, 시간 도메인 신호 및 이에 상응하는, 양쪽의 인코딩된 신호에 포함되는 짧은 크로스페이드되는 지역을 도시한, 인코딩된 신호에 대한 표현을 도시한 것이다.
 도 4a는, 인코딩 브랜치 앞에 위치한 스위치를 포함하는 블록도를 도시한 것이다.
 도 4b는, 인코딩 브랜치 뒤에 위치하는 스위치를 이용한 인코딩 기법에 대한 블록도를 도시한 것이다.
 도 4c는, 선회되는 결합자의 실시예에 대한 블록도를 도시한 것이다.
 도 5a는, 유사 주기적 또는 맥박형 신호 세그먼트로써 시간 도메인 발화 세그먼트의 파동 형상을 도시한 것이다.
 도 5b는, 도 5a의 세그먼트에 대한 스펙트럼을 도시한 것이다.
 도 5c는, 노이즈형 세그먼트에 대한 일례로써 무성 발화의 시간 도메인 발화 세그먼트를 도시한 것이다.
 도 5d는, 도 5c의 파동형 스펙트럼을 도시한 것이다.
 도 6은, 합성 CELP 인코더에 의한 분석에 대한 블록도를 도시한 것이다.
 도 7a 내지 7d는, 맥박형 신호에 대한 일례로써 유성/무성 여자 신호를 도시한 것이다.
 도 7e는, 단기 예측 정보 및 예측 오류 (여자) 신호를 제공하는 인코더 측 LPC 스테이지를 도시한 것이다.
 도 7f는, 가중 신호를 생성하기 위해 LPC 장치에 대한 다른 일례를 도시한 것이다.
 도 7g는, 역가중 작업 및 도 2b의 컨버터(537)에서 요구되는 후속 여자 분석을 적용함으로써 가중 신호를 여자 신호로 변형하는 방법을 도시한 것이다.

도 8은, 본 발명의 일실시예에 있어서 조인트 멀티 채널 알고리즘의 블록도를 도시한 것이다.

도 9는, 대역폭 확장 알고리즘의 최선 실시예를 도시한 것이다.

도 10a는, 열린 루프 결정을 수행할 때 스위치에 대한 상세한 설명을 도시한 것이다.

도 10b는, 닫힌 루프 결정 모드에서 작동될 때 스위치에 대한 도면을 도시한 것이다.

발명을 실시하기 위한 구체적인 내용

[0019] 도 1a는, 본 발명의 일실시예에 있어서 두 개의 캐스케이드된 스위치를 구비된 것에 대해 도시하고 있다. 모노 신호, 스테레오 신호 또는 멀티 채널 신호가 스위치(200) 내로 입력된다. 상기 스위치(200)는 결정 스테이지(300)에 의해서 제어된다. 상기 결정 스테이지는, 입력값으로써 블록(200)으로 입력된 신호를 받는다. 대안적으로 상기 결정 스테이지(300)는 또한 모노 신호, 스테레오 신호 또는 멀티 채널 신호 내에 포함되어 있거나 또는 적어도 이러한 신호와 연결되어 있는 사이드 정보(side information, 부가정보)를 입력받을 수 있다. 여기서 정보는 상기 모노 신호, 상기 스테레오 신호 또는 상기 멀티 채널 신호를 최초로 생산할 때, 생성된 것일 수 있다.

[0020] 상기 결정 스테이지(300)는, 도 1의 상단의 브랜치에 도시된 주파수 인코딩 부분(400) 또는 도 1a의 하단 브랜치에 도시된 LPC-도메인 인코딩 부분(500) 중 어느 하나에서 신호를 공급하기 위해 상기 스위치(200)를 작동시킨다. 상기 주파수 도메인 인코딩 브랜치의 주요 구성요소는 스펙트럼 변환 블록(410)이고, 이것은 일반적인 처리 전 단계 출력 신호(추후 기술한다)를 스펙트럼 도메인으로 변환시키는 기능을 한다. 상기 스펙트럼 변환 블록은 MDCT 알고리즘, QMF, FFT 알고리즘, 웨이블릿 분석(Wavelet analysis) 또는 일정한 숫자의 필터뱅크 채널을 구비한 임계적으로 샘플링된 필터뱅크 같은 필터뱅크를 포함할 수 있다. 여기서 이러한 필터뱅크 내의 서브밴드는, 실수값 신호 또는 복소수값 신호일 수도 있다. 상기 스펙트럼 변환 블록(410)의 출력은 스펙트럼 오디오 인코더(421)를 사용하여 인코딩될 수 있다. 여기서 상기 인코더(421)는 AAC 코딩 기법으로부터 알려진 처리 블록을 포함할 수 있다.

[0021] 일반적으로 브랜치 400 내에서의 처리는 인식 기반 모델 또는 정보 싱크 모델 내에서 처리하는 것이다. 그래서 이러한 브랜치는 소리를 받는 인간 청각 시스템을 모델로 한다. 이와는 반대로, 상기 브랜치 500 내에서의 처리는 여자, 잔여 또는 LPC 도메인 내에서 신호를 생성한다. 일반적으로 브랜치 500 내에서의 처리는 발화 모델 또는 정보 생성 모델 내에서 처리된다. 발화 신호에 있어서 이러한 모델은, 소리를 생성하는 인간 발화/소리 생성 시스템에 대한 모델이다. 그러나 만약 다른 소리 생성 모델을 요구로 하는 다른 소스로부터 발생된 소리가 인코딩되어야 한다면, 그때 상기 브랜치 500 내에서의 처리는 다를 수 있다.

[0022] 상기 하단의 인코딩 브랜치 500 내에서, 주요 구성요소는 LPC 장치(510)이고, 이는 LPC 필터의 특징을 제어하기 위해 사용되는 LPC 정보를 출력한다. 이러한 LPC 정보는 디코더로 전송된다. LPC 스테이지(510) 출력 신호는 여자 신호 및/또는 가중된 신호를 포함하는 LPC 도메인 신호이다.

[0023] 상기 LPC 장치는 일반적으로 LPC 도메인 신호를 출력하는데, 상기 LPC 도메인 신호는 도 7e의 여자 신호 또는 도 7f의 가중된 신호 또는 오디오 신호에 대한 LPC 필터 계수를 적용함으로써 생성되는 다른 신호 같은 LPC 도메인 내의 어떠한 신호일 수도 있다. 더구나 LPC 장치는 또한 이러한 계수를 결정하고 이러한 계수를 양자화/인코딩할 수 있다.

[0024] 상기 결정 스테이지에서의 결정은, 상기 결정 스테이지가 음악/발화를 차별하고, 음악 신호는 상기 상단의 브랜치(400)로 입력되고 발화 신호는 상기 하단의 브랜치(500)로 입력되도록 스위치(200)를 제어하기 위해서 신호 적응성을 지닌다. 일실시예에 있어서, 상기 결정 스테이지는, 정확한 디코딩 작업을 수행하기 위해서 디코더가 결정 정보를 사용할 수 있도록, 상기 결정 정보를 출력 비트스트림으로 제공한다.

[0025] 이러한 디코더가 도 1b에 도시되어 있다. 상기 스펙트럼 오디오 인코더(421)에 의해 출력된 신호는, 전송된 후, 스펙트럼 오디오 디코더(431)로 입력된다. 상기 스펙트럼 오디오 디코더(431)의 출력은 시간 도메인 컨버터(440)로 입력된다. 유사하게, 도 1a의 상기 LPC 도메인 인코딩 브랜치(500)는 디코더 측에서 수신되고, LPC 여자 신호를 얻기 위하여 구성요소 531, 533, 534 및 532에 의해 처리된다. 상기 LPC 여자 신호는 LPC 통합 스테이지(540)로 입력된다. 여기서 상기 스테이지(540)는, 추가적인 입력으로 상응하는 LPC 분석 스테이지(510)에 의해 생성된 LPC 정보를 수신받을 수 있다. 상기 시간 도메인 컨버터(440) 및/또는 상기 LPC 통합 단계(540)의 출력은 스위치(600)로 입력된다. 상기 스위치(600)는, 예를 들어 결정 스테이지(300)에 의해 생성되거나 또는 원래의 모노 신호, 스테레오 신호 또는 멀티 채널 신호의 생산자에 의해 외부로부터 제공되는 스위치 제어 신호를 통해 제어된다. 상기 스위치(600)의 출력은 완전히 모노 신호, 스테레오 신호 또는 멀티채널 신호이다.

[0026] 상기 스위치(200) 및 상기 결정 스테이지(300)로 입력되는 신호는 모노 신호, 스테레오 신호, 멀티채널신호 또는 일반적인 오디오 신호일 수 있다. 스위치(200) 입력 신호 또는 스테이지 200에 입력되는 신호의 기반인 원래의 오디오 신호의 생산자 같은 외부 소스로부터 추론된 결정에 의존하여, 상기 스위치는 상기 주파수 인코딩 브랜치(400) 및 상기 LPC 인코딩 브랜치(500) 간의 전환을 수행한다. 상기 주파수 인코딩 브랜치(400)는, 스펙트럼 전환 스테이지(410) 및 그 뒤에 연결되는 양자화/코딩 스테이지(421)를 포함한다. 상기 양자화/코딩 단계는 AAC 인코더같은 근대 주파수 도메인 인코더로부터 알려진 기능들을 포함한다. 더구나 상기 양자화/코딩 스테이지(421) 내에서의 양자화 과정은, 상기 주파수에 대한 음향심리학적 한계점 등의 음향심리학적 정보를 생산하는 음향심리학적 모듈에 의해서 제어될 수 있다. 여기서 상기 정보는 스테이지 421에 입력된다.

[0027] LPC 인코딩 브랜치에서, 상기 스위치 출력 신호는, LPC 사이드 정보 및 LPC 도메인 신호를 생산하는 LPC 분석 스테이지(510)를 통해 처리된다. 상기 여자 인코더는 다양한 방법으로, LPC 도메인 내의 양자화/코딩 작동(522) 또는 LPC 스펙트럼 도메인 내에서 값들을 처리하는 양자화/코딩 스테이지(524) 사이에서 LPC 도메인 신호의 추가적인 처리를 전환하기 위한 추가적인 스위치를 비교한다. 마지막으로 스펙트럼 컨버터(523)는 양자화/코딩 스테이지(524)가 입력될 때, 제공된다. 상기 스위치(521)는, 예를 들어 AMR-WB+ 같은 기술적 설명 내에 묘사된 구체적인 설정에 의하여, 기술적 열린 루프 방식 또는 닫힌 루프 방식 내에서 제어된다.

[0028] 상기 닫힌 루프 제어 모드에 있어서, 상기 인코더는 추가적으로 LPC 도메인 신호에 대하여 역양자화기/코더(531), LPC 스펙트럼 도메인 신호를 위한 역양자화기/코더(533) 및 아이템(533)의 출력에 대한 역스펙트럼 컨버터(534)를 포함한다. 상기 제2 인코딩 브랜치의 처리 브랜치 내에서의 인코딩된 신호 및 다시 디코딩된 신호 양자 모두, 상기 스위치 제어 장치(525)에 입력된다. 상기 스위치 제어 장치(525) 내에서, 상기 두 개의 출력신호들은 서로서로 비교되거나 또는 타겟 함수와 비교되고, 또는 상기 스위치(521)가 내리는 결정을 위해서 더 낮은 왜곡을 가진 신호가 사용될 수 있도록 상기 타겟 함수는 양 신호 내의 왜곡을 비교하는 것에 의해 계산될 수 있다. 그렇지 않으면 양 브랜치가 고정적이지 않은 비트 레이트를 제공하는 경우에 있어서, 더 낮은 비트 레이트를 제공하는 브랜치는, 상기 브랜치의 노이즈(noise)에 대한 신호의 비율이 다른 브랜치의 노이즈에 대한 신호의 비율보다 낮은 경우에서조차 선택될 수 있다. 그렇지 않으면, 상기 타겟 함수는, 입력값으로써, 구체적인 목적을 위한 최선의 결정을 검색하기 위해 각각의 신호의 노이즈에 대한 신호의 비율 및 각각의 신호 및/또는 추가적인 기준(criteria)의 비트 레이트를 사용할 수 있다. 예를 들어 만약 상기의 목적이 가능한 한 비트 레이트가 낮도록 하는 것이라면, 그때 상기 타겟 함수는 상기 구성요소 531 및 534에 의한 두 개의 출력되는 신호의 비트 레이트에 매우 의존할 수도 있다. 그러나 상기 주된 목적이 특정 비트레이트에 대하여 최선의 질을 구비하는 것이라면, 상기 스위치 제어 525는, 예를 들어 허용된 비트 레이트 이상의 각각의 신호는 폐기하고, 모든 신호가 허용된 비트 레이트 미만인 경우에는 상기 스위치 제어는 더 좋은 노이즈에 대한 신호 비율을 구비한 신호, 예를 들어 더 작은 양자화/코딩 왜곡을 구비한 신호를 선택할 수 있다.

[0029] 본 발명에 있어서 상기 디코딩 기법은 도 1b에 도시되어 있다. 세 가지 가능한 출력 신호 종류 각각에 있어서, 구체적인 디코딩/재양자화 스테이지 431, 531 또는 533이 있다. 상기 스테이지 431이 주파수/시간 컨버터 440를 사용하여 시간 도메인으로 변환되는 시간 스펙트럼을 출력하는 반면에, 상기 스테이지 531은 LPC 도메인 신호를

출력하고, 아이템 533은 LPC 스펙트럼을 출력한다. 스위치 532에 입력되는 신호가 모두 LPC 도메인 내에 있다는 것을 확실히 하기 위하여 상기 LPC 스펙트럼/LPC 컨버터 534가 제공된다. 상기 스위치(532)의 출력 데이터는 LPC 통합 스테이지(540)를 이용하여 시간 도메인으로 다시 변형된다. 이는 LPC 정보를 생성하고 전송하는 인코더-사이드(encoder-side)를 통하여 제어된다. 그래서 블록 540 뒤에 이어지는, 양 브랜치는, 도 1a의 인코딩 기법 내로 입력되는 신호에 따라 모노 신호, 스테레오 신호 또는 멀티채널 신호 같은 오디오 신호를 최종적으로 얻기 위하여 스위치 제어 신호에 따라 전환되는 시간 도메인 정보를 구비하고 있다.

[0030] 도 1c는 도 4b의 원리와 유사한 스위치 521의 다른 배열에 대한 또 다른 실시예를 도시한 것이다.

[0031] 도 2a는 본 발명의 두 번째 관점에 따른, 최선 인코딩 기법에 대해 도시한 것이다. 상기 스위치(200) 입력에 관련된 일반적인 전처리 기법은, 출력값으로써, 조인트 스테레오 매개변수 및 둘 또는 그 이상의 채널을 구비한 신호인 입력 신호를 다운믹싱함으로써 생성된 모노 출력 신호를 생성하는, 서라운드/조인트 스테레오 블록 101을 포함하는 것을 특징으로 한다. 일반적으로 상기 블록 101의 출력에서의 상기 신호는 또한 더 많은 채널을 구비한 신호일 수도 있다. 그러나 블록 101의 다운믹싱 기능에 기인하여 블록 101의 출력 시의 채널의 숫자는 블록 101로 입력된 채널의 숫자보다 작을 것이다.

[0032] 상기 일반적인 전처리 기법은, 블록 101 대신에 또는 블록 101에 더하여 대역폭 확장 스테이지(102)를 포함할 수 있다. 도 2a에 도시된 바와 같이 일실시예에 있어서 블록(101)의 출력값은, 도 2a의 인코더 내에 도시된 것처럼, 출력이 대역 신호 또는 저 패스 신호 같은 대역이 제한된 신호를 출력하는 대역폭 확장 블록(120)으로 입력된다. 바람직하게는, 이러한 신호는 마찬가지로 (예를 들어 인자 2에 의해) 다운샘플링될 수 있다. 더구나 블록 102로 입력되는 고 대역의 신호에 있어서, MPEG-4의 HE-AAC 프로파일로 알려진 스펙트럼 인벨롭 매개변수(spectral envelope parameter), 역 필터링 매개변수, 노이즈 플로어 매개변수(noise floor parameter) 같은 대역폭 확장 매개변수는 생성되고 비트스트림 멀티플렉서(multiplexer, 800)로 이동된다.

[0033] 바람직하게는, 상기 결정 스테이지(300)는, 예를 들어 음악 모드 또는 발화 모드를 결정하기 위하여 블록 101로 입력되거나 또는 블록 102로 입력되는 신호를 수신한다. 음악 모드에서는, 상단의 인코딩 브랜치(400)가 선택되고, 반면에 발화 모드에서는 하단의 인코딩 브랜치(500)가 선택된다. 바람직하게는, 상기 결정 스테이지는, 추가적으로 구체적인 신호에 이러한 블록들의 기능을 적용하기 위해 조인트 스테레오 블록(101) 및/또는 대역폭 확장 블록(102)을 제어할 수 있다. 그러므로 상기 결정 스테이지가 상기 입력 신호의 어떤 시간 부분이, 예를 들어 음악 모드같은 제1 모드라고 결정한 경우, 블록 101 및/또는 블록 102의 구체적인 특징은 상기 결정 스테이지(300)에 의해서 제어될 수 있다. 그렇지 않으면, 상기 결정 스테이지(300)가 상기 신호는 발화 모드 또는, 일반적으로, 제2 LPC 도메인 모드라고 결정한 경우, 블록 101 및 102의 구체적인 특징은 상기 결정 스테이지 출력값에 따라서 제어될 수 있다.

[0034] 바람직하게는 상기 코딩 브랜치(400)의 상기 스펙트럼 변환은 MDCT 작동을 통하여 행해진다. 더욱 바람직하게는 상기 MDCT 작동은 시간 워핑(warping)된 MDCT 작동일 수 있다. 여기서 강도 또는, 일반적으로 상기 워핑 강도는 0에서 높은 워핑 강도 사이에서 제어될 수 있다. 워핑 강도가 0인 경우에 블록 411 내의 상기 MDCT 작동은 당 업계에서 알려진 직진 MDCT 작동(straight-forward MDCT operation)이다. 시간 워핑 사이트 정보와 더불어 상기 시간 워핑 강도는, 상기 비트스트림 멀티플렉서(800)에 사이트 정보로써 전송 및 입력된다.

[0035] LCP 인코딩 브랜치에 있어서, 상기 LPC 도메인 인코더는 피치 이득(pitch gain), 피치 지연(pitch lag) 및/또는 코드북 인덱스 및 이득 같은 코드북 정보를 계산하는 ACELP 코어(526)를 포함할 수 있다. 3GPP TS 26.290에서 알려진 TCX 모드는, 상기 변환 도메인의 지각적으로 가중된 신호의 처리를 초래한다. 푸리에 변환 가중 신호는, 노이즈 인자 양자화를 포함하는 나뉜 멀티 레이트 래티스 양자화(split multi-rate lattice quantization, 대수학적 VQ)를 사용하여 양자화된다. 변환은 1024, 512 또는 256의 샘플 윈도우 내에서 계산된다. 상기 여자 신호는, 역(逆) 가중 필터를 통하여 상기 양자화된 가중 신호를 역 필터링함으로써 복구된다. 상기 제1 코딩 브

랜치(400)에서, 스펙트럼 컨버터는, 바람직하게는 단일 벡터 양자화 스테이지를 포함하는 양자화/엔트로피 인코딩이 이어지는 어떤 윈도우 함수를 포함하는, 구체적으로 적용되는 MDCT 작용을 포함하나, 바람직하게는, 예를 들어 도 2a의 아이템 421 내의 주파수 도메인 코딩 브랜치 내의 양자화/코더와 유사한 결합된 스칼라 양자화/엔트로피 코더일 수 있다.

[0036] 제2 코딩 브랜치에서, 스위치 521이 이어지고, 다시 ACELP 블록 526 또는 TCX 블록 527이 뒤따르는 LPC 블록 510이 존재한다. ACELP는 3GPP TS 26.190에서 설명되고, TCX는 3GPP TS 26.290에서 설명된다. 일반적으로 상기 ACELP 블록 526은, 도 7e에 설명된 바와 같은 과정에 의해 계산된 LPC 여자신호를 수신한다. 상기 TCX 블록 527은 도 7f에 의해 생성된 가중된 신호를 수신한다.

[0037] TCX에서, 상기 변환은 LPC 기반의 가중 필터를 통하여 입력된 신호를 필터링함으로써 연산된 가중 신호에 적용된다. 본 발명의 바람직한 실시예에 있어서 상기 가중 필터는 $(1-A(z/\gamma)) / (1-\mu z^{-1})$ 에 따라 주어진다.

[0038] 그러므로, 상기 가중 신호는 LPC 도메인 신호이고, 그것의 변형은 LPC 스펙트럼 도메인이다. 상기 ACELP 블록 (526)에 의해 처리된 상기 신호는 여자 신호이고, 상기 블록 527에 의하여 처리된 신호와는 다르나, 양 신호 모두 LPC 도메인 내에 있다.

[0039] 도 2b에 도시된 디코더의 면에서, 블록 537 내의 역 스펙트럼 변환 후에, 상기 가중 필터의 역은 $(1-\mu z^{-1}) / (1-A(z/\gamma))$ 와 같은 함수로 주어진다.

[0040] 그러면, 상기 신호는 $(1-A(z))$ 를 통과하여 필터링되고 LPC 여자 도메인으로 이동한다. 그러면, 상기 LPC 도메인 블록(534) 및 상기 TCX⁻¹ 블록(537)에 대한 변환은, 상기 가중 도메인은 여자 도메인으로 변환하기 위하여 역변환 및 다음의 수학적식을 통한 필터링을 포함한다.

$$\frac{(1-\mu z^{-1})}{(1-A(z/\gamma))} (1-A(z))$$

[0041]

[0042] 도 1a, 1c, 2a, 2c의 아이템 510은 단일 블록을 도시하고 있음에도 불구하고, 블록 510은, 이러한 신호가 상기 LPC 도메인 내에 있는 한 다른 신호를 출력한다. 여자 신호 모드 또는 가중 신호 모드 같은 블록 510의 현 모드는 현재의 스위치의 상태에 의존한다. 그렇지 않으면 상기 블록 510은 두 개의 평행한 처리 장치를 구비할 수 있다. 여기서 하나의 장치는 도 7e에 도시된 바와 유사하게 수행되고, 다른 장치는 도 7f에 도시된 바와 같이 수행된다. 이런 이유로 510의 출력에서 상기 LPC 도메인은 LPC 여자 신호, LPC 가중 신호 또는 다른 LPC 도메인 신호를 나타낸다.

[0043] 도 2a 또는 도 2c의 제2 인코딩 브랜치 (ACELP/TCX)에서, 상기 신호는 바람직하게는 인코딩 전에 필터 $1-0.68z^{-1}$ 를 통과하여 선(先)강조된다. 도 2b에 도시된 ACELP/TCX 디코더에서, 통합적인 신호는 상기 필터 $1/(1-0.68z^{-1})$ 에 의해서 역(逆)강조(deemphasize)된다. 상기 선강조는, LPC 분석 및 양자화 전에 상기 신호가 선강조되는, 상기 LPC 블록의 일부분이 될 수 있다. 유사하게, 역강조는 LPC 통합 블록 LPC⁻¹ 540의 일부일 수 있다.

- [0044] 도 2c는 도 2a의 구현에 있어서 다른 실시예를 도시하고 있다. 그러나 도 4의 원리와 유사하게 상기 스위치 (521)의 다른 배열을 갖는다.
- [0045] 바람직하게는, 상기 제1 스위치 200(도 1a 또는 도 2a 참조)은 열린 루프 결정(도 4a에 도시됨)을 통해 제어될 수 있고, 상기 제2 스위치는 닫힌 루프 결정(도 4b에 도시됨)을 통해 제어될 수 있다.
- [0046] 예를 들어, 도 2c에서는, 도 4b와 마찬가지로 상기 제2 스위치는 ACELP 및 TCX 브랜치 뒤에 위치한다. 그러면 제1 처리 브랜치에서, 상기 제1 LPC 도메인은 상기 LPC 여자를 나타내고, 제2 처리 브랜치에서 상기 제2 LPC 도메인은 상기 LPC 가중된 신호를 나타낸다. 즉, 상기 제1 LPC 도메인 신호는, 상기 LPC 잔여 도메인에 대하여 변환하기 위해서, $(1-A(z))$ 를 통한 필터링에 의하여 구해진다. 반면에 상기 제2 LPC 도메인 신호는, LPC 가중 도메인에 대해 변환하기 위해서, 필터 $(1-A(z/\gamma)) / (1-\mu z^{-1})$ 를 통한 필터링에 의해서 구해진다.
- [0047] 도 2b는 상기 도 2a에 도시된 인코딩 기법에 상응하는 디코딩 기법을 도시하고 있다. 상기 도 2a의 비트 스트림 멀티플렉서(800)로부터 생성된 상기 비트스트림은 비스트림 디멀티플렉서(demultiplexer, 900)으로 입력된다. 일례로 모드 탐색 블록(601)에 의해 비트스트림으로부터 추출된 정보에 의존하여, 디코더 쪽 스위치(600)은 상기 상단의 브랜치로부터 전신 신호 또는 대역폭 확장 블록 701으로의 하단의 브랜치 방향으로 제어된다. 상기 대역폭 확장 블록(701)은, 상기 비트스트림 디멀티플렉서(900)으로부터 사이드 정보를 수신받고, 사이드 정보 및 모드 결정(601)의 출력값을 기반으로 하여, 스위치(600)에 의해 출력되는 저 대역을 기반으로 하는 고(高)대역을 재구성한다.
- [0048] 상기 블록 701에 의해 생산되는 전체 대역 신호는 상기 조인트 스테레오/서라운드 처리 스테이지(702)로 입력되는데, 이 스테이지는 두 개의 스테레오 채널 또는 몇 개의 멀티 채널을 재구성한다. 일반적으로 블록 702는, 상기 블록에 입력된 것보다 더 많은 채널을 출력한다. 상기 사용에 의존하여 상기 블록 702으로의 입력은 스테레오 모드 같이 두 개의 채널을 포함할 수 있고, 상기 블록에 의한 출력이 상기 블록으로 입력된 것보다 더 많은 채널을 구비하는 한 더 많은 채널을 포함할 수 있다.
- [0049] 상기 스위치(200)는, 오직 하나의 브랜치는 처리를 위한 신호를 수신하고, 다른 브랜치는 처리하기 위한 신호를 수신하지 않도록 양 브랜치를 서로 전환시키는 것을 보여주고 있다. 그러나 이와 다른 실시예에서는, 상기 스위치는 또한, 예를 들어 오디오 인코더 421 및 여자 인코더 522, 523, 524 다음에 배열될 수 있다. 이는 양 브랜치 400, 500은 평행하게 동일한 신호를 처리한다는 것을 의미한다. 그러나 비트레이트가 두 배가 되지 않도록 하기 위해서 이러한 인코딩된 브랜치 400 또는 500 중의 하나에서 출력된 신호만은 선택되고, 출력 비트레이트 내로 기록된다. 상기 결정 단계는 그때, 상기 비트스트림 내로 기록된 신호를 특정한 코스트 함수(cost function)으로 최소화하기 위하여 작동할 것이다. 여기서 상기 코스트 함수는, 생성된 비트레이트 또는 생성된 지각의 왜곡 또는 결합된 레이트/왜곡 코스트 함수이다. 그러므로 이러한 모드 또는 도면에 도시된 모드에 있어서, 상기 결정 스테이지는 또한, 최종적으로 인코딩 브랜치 출력만이, 주어진 지각 왜곡을 위하여 최소 비트레이트를 구비하거나, 또는 주어진 비트레이트에 대해서 최소의 지각 왜곡을 구비하는 비트스트림 내로 쓰여진다는 것을 확실히 하기 위해서 닫힌 루프 모드에서 작동될 수도 있다. 닫힌 루프 모드에서, 피드백(feedback)된 입력은 도 1a의 상기 세 개의 양자화기/스칼라 블록 421, 522 및 424의 출력으로부터 추론된다.
- [0050] 예를 들어 제1 스위치(200) 및 제2 스위치(521)의 두 개의 스위치를 구비한 구현예에서, 바람직하게는 상기 제1 스위치에 대한 시간 해상도는 상기 제2 스위치에 대한 시간 해상도보다 낮은 것이다. 다르게 말하면, 전환 작업에 의하여 전환할 수 있는 상기 제1 스위치로 입력되는 신호의 블록들은 상기 LPC 도메인 내에서 구동되는 제2 스위치에 의해 전환된 블록보다 더 크다. 모범적으로 상기 주파수 도메인/LPC-도메인 스위치(200)는 1024 샘플의 길이에 대한 블록들을 전환하고, 상기 제2 스위치(521)는 256 샘플 구비한 블록을 각각 전환한다.

- [0051] 도 1a 내지 10b의 일부는 장치로써의 블록도가 도시되어 있지만, 이러한 그림은 동시에, 방법의 단계들에 상기 블록의 기능들이 대응되는, 방법에 대한 도면이다.
- [0052] 도 3a는 제1 인코딩 브랜치(400) 및 제2 인코딩 브랜치(500)의 출력으로써 인코딩된 오디오 신호를 생성하기 위한 오디오 인코더를 도시한다. 더구나, 상기 인코딩된 오디오 신호는 바람직하게는, 일반적인 선처리 스테이지로부터 얻어진 선처리 매개변수 또는 이전의 도면을 참조하여 설명된 스위치 제어 정보 같은 사이드 정보를 포함한다.
- [0053] 바람직하게는, 상기 제1 인코딩 브랜치는 제1 코딩 알고리즘에 따라 오디오 중간 신호(195)를 인코딩하기 위하여 작동한다. 여기서 상기 제1 코딩 알고리즘은 정보 싱크 모델을 구비한다. 상기 제1 인코딩 브랜치(400)는, 상기 오디오 중간 신호(195)의 인코딩된 스펙트럼 정보 표현인, 제1 인코더 출력 신호를 생성한다.
- [0054] 더구나, 제2 인코딩 브랜치(500)는, 제2 코딩 알고리즘에 따라 상기 오디오 중간 신호(195)를 인코딩하는데 적용될 수 있다. 여기서 상기 제2 코딩 알고리즘은 정보 소스 모델을 구비하며, 제2 인코더 출력 신호 내에서 상기 중간 오디오 신호를 나타내는 정보 소스 모델에 대한 인코딩된 매개 변수를 생성한다.
- [0055] 더구나 상기 오디오 인코더는, 상기 오디오 중간 신호(195)를 얻기 위해 오디오 입력 신호(99)를 선처리하기 위한 일반적인 선처리 단계를 포함한다. 구체적으로, 상기 일반적인 선처리 단계는, 상기 오디오 중간 신호(195), 예를 들어 일반적인 선처리 알고리즘의 출력은 상기 오디오 입력 신호의 압축된 버전이기 위하여 상기 오디오 입력 신호(99)를 처리하기 위해 작동한다.
- [0056] 인코딩된 오디오 신호를 생성하기 위한 오디오 인코딩 방법의 바람직한 방법은, 제1 코딩 알고리즘에 따라 오디오 중간 신호(195)를 인코딩하는 단계(400), 제2 코딩 알고리즘에 따라 오디오 중간 신호(195)를 인코딩하는 단계(500) 및 상기 오디오 중간신호(195)를 얻기 위해서 오디오 입력 신호(99)를 일반적으로 선처리하는 단계;를 포함하고, 여기서 상기 제1 코딩 알고리즘은 정보 싱크 모델을 구비하고, 제1 출력 신호 내에서, 오디오 신호를 나타내는 인코딩된 스펙트럼 정보를 생산하고, 상기 제2 코딩 알고리즘은 정보 소스 모델을 구비하고, 제2 출력 신호 내에서 상기 중간 신호(195)를 나타내는 정보 소스 모델에 대한 인코딩된 매개변수를 생산하며, 상기 일반적으로 선처리되는 과정에서 상기 오디오 입력 신호(99)는, 상기 오디오 중간 신호(195)가 오디오 입력 신호(99)의 압축 버전이 되기 위해서 처리되고, 상기 인코딩된 오디오 신호는, 상기 오디오 신호의 일정 부분에서 상기 제1 출력신호 또는 상기 제2 출력 신호를 포함하고 있는 것을 특징으로 한다. 바람직하게는 상기 방법은, 추가적으로 제1 코딩 알고리즘 또는 상기 제2 코딩 알고리즘을 사용하여 상기 오디오 중간 신호의 일정부분을 인코딩하거나 또는 상기 두 개의 알고리즘을 사용하여 인코딩하고, 인코딩된 신호 내에서 상기 제1 코딩 알고리즘의 결과 또는 상기 제2 코딩 알고리즘의 결과를 출력하는 단계;를 포함하는 것을 특징으로 할 수 있다.
- [0057] 일반적으로, 상기 제1 인코딩 브랜치(400) 내에서 사용되는 상기 오디오 인코딩 알고리즘은 오디오 싱크 내의 상황을 반영하고 모델로 한다. 오디오 정보의 상기 싱크는 일반적으로 인간의 귀이다. 상기 인간의 귀는 주파수 분석기의 모델이 될 수 있다. 그러므로 상기 제1 인코딩 브랜치는 인코딩된 스펙트럼 정보를 출력한다. 바람직하게는, 상기 제1 인코딩 브랜치는, 더 나아가 음향심리학적 임계값(psychoacoustic masking threshold)을 추가적으로 적용하기 위한 음향 심리학적 모델을 포함할 수 있다. 이러한 음향심리학적 임계값은, 바람직하게는 상기 음향심리학적 임계값 아래에 숨겨진 스펙트럼 오디오 값을 양자화함으로써 양자화 잡음이 도입되도록 상기 양자화가 수행되는, 오디오 스펙트럼 값을 양자화할 때 사용된다.
- [0058] 상기 제2 인코딩 브랜치는 정보 소스 모델을 나타낸다. 이는 오디오 소리의 발생을 반영한다. 그러므로 정보 소스 모델은 LPC 분석 스테이지, 즉 시간 도메인 신호를 LPC 도메인으로 전송하고 이어서 여자 신호 같은 LPC 잔여 신호를 처리하는 스테이지에 의해 반영되는 발화 모델을 포함한다. 그러나 다른 소리 소스 모델은 어떤 악기

또는 실제 세계에 존재하는 구체적인 소리 소스 같은 소리 생산자를 나타내는 소리 소스 모델이다. 다른 소리 소스 모델 사이에서의 선택은, 예를 들어 SNR 계산, 즉 상기 소스 모델에 대한 계산이 하나의 오디오 신호의 어떤 시간 일부 및/또는 주파수 일부에 적합한 최선인 계산을 기반하는, 몇 개의 소리 소스 모델이 가능한 경우에 수행될 수 있다. 그러나 바람직하게는 상기 인코딩 브랜치 간의 전환은 시간 도메인 내에서 수행된다. 즉, 어떤 시간의 일부가 하나의 모델을 이용하여 인코딩되고, 중간 신호의 어떤 다른 시간의 일부는 다른 인코딩 브랜치를 이용하여 인코딩된다.

[0059] 정보 소스 모델은 특정한 매개 변수에 의해서 표현된다. 발화 모델을 고려하여, 상기 매개 변수는, AMR-WB+같은 근대적인 발화 코더를 고려하였을 때, LPC 매개변수 및 코딩된 여자 매개변수이다. 상기 AMR-WB+는 CELP 인코더 및 TCX 인코더를 포함한다. 이 경우에 있어서 상기 코딩된 여자 매개변수는 글로벌 이득, 노이즈 플로어 및 다양한 길이의 코드들일 수 있다.

[0060] 도 3b는 도 3a에 도시된 인코더에 대응하는 디코더가 도시되어 있다. 일반적으로 도 3b는, 디코딩된 오디오 신호(799)를 구하기 위한 인코딩된 오디오 신호를 디코딩하기 위한 오디오 디코더를 도시한다. 상기 디코더는 정보 싱크 모델을 구비한 제1 코딩 알고리즘에 따라 인코딩된 신호를 디코딩하기 위한 제1 디코딩 브랜치(400)를 포함한다. 상기 오디오 디코더는, 더구나 정보 소스 모델을 구비한 제2 코딩 알고리즘에 따른 인코딩된 정보 신호를 디코딩하기 위한 제2 디코딩 브랜치(400)를 포함한다. 상기 오디오 디코더는, 더구나 결합된 신호를 얻기 위해 상기 제1 디코딩 브랜치(540) 및 상기 제2 디코딩 브랜치(550)로부터 출력된 신호를 결합하기 위한 결합자를 포함할 수 있다. 도 3b에 도시된 바와 같이 상기 결합된 신호는, 디코딩된 오디오 중간 신호(699)로써, 상기 디코딩된 오디오 중간 신호(699)를 사후 처리하기 위한 일반적인 사후 처리 스테이지로 입력된다. 여기서 상기 결합된 신호는, 상기 일반적인 전처리 스테이지의 출력 신호가 상기 결합된 신호의 확장된 버전이도록, 상기 결합자(600)에 의해 출력된, 결합 신호이다. 그러므로 상기 디코딩된 오디오 신호(799)는 상기 디코딩된 오디오 중간 신호(699)에 비교하여 강화된 정보 콘텐츠를 구비하고 있다. 이러한 정보의 확장은, 인코더로부터 디코더로 전송될 수 있거나 또는 상기 디코딩된 오디오 중간 신호 그 자체로부터 추론될 수 있는 전처리, 사후처리 매개변수의 도움으로 상기 일반적인 사후 처리 스테이지에 의해 제공된다. 그러나 바람직하게는 전처리, 사후처리 매개변수는, 이러한 과정이 디코딩되는 오디오 신호의 개선된 질을 가능하게 하는 까닭에, 인코더에서 디코더로 전송된다.

[0061] 도 3c는, 오디오 입력 신호(195)를 인코딩하기 위한 오디오 인코더를 도시하고 있다. 이는 본 발명의 바람직한 일 실시예에 따르면, 도 3a의 중간 오디오 신호(195)와 동일하다. 상기 오디오 입력 신호(195)는, 예를 들어 시간 도메인이지만, 주파수 도메인, LPC 도메인, LPC 스펙트럼 도메인 또는 다른 종류의 도메인일 수도 있는 제1 도메인 내에 존재한다. 일반적으로, 하나의 도메인에서 다른 도메인으로의 변환은, 잘 알려진 시간/주파수 변환 알고리즘 또는 주파수/시간 변환 알고리즘 같은 변환 알고리즘에 의해서 수행된다.

[0062] 예를 들어 상기 LPC 도메인 내에서 상기 시간 도메인으로부터의 또 다른 변형은, LPC 잔여 신호 또는 여자 신호가 되는, 시간 도메인 신호의 LPC 필터링에 대한 결과이다. 상기 변형 전에 상당한 숫자의 신호 샘플에 대해 충격을 갖는 필터링된 신호를 생산하는 또 다른 필터링 작업은, 상황이 가능한 한 변환 알고리즘으로써 사용될 수 있다. 그러므로 LPC 기반의 가중 필터를 사용하여 오디오 신호를 가중하는 것은 더욱 변형적이고, 이는 상기 LPC 도메인 내에서 신호를 생성한다. 시간/주파수 변형에서, 상기 단일 스펙트럼 값의 수정은, 상기 변형 전의 모든 시간 도메인값에 충격을 가한다. 유사하게, 어떠한 시간 도메인 샘플의 변형은 각각의 주파수 도메인 샘플에 충격을 준다. 비슷하게, LPC 도메인 상황 내에서의 여자 신호의 샘플의 수정은, 상기 LPC 필터의 길이 때문에, 상기 LPC 필터 앞의 상당한 숫자의 샘플에 영향을 준다. 비슷하게, LPC 변형 앞의 샘플에 대한 수정은, 상기 LPC 필터의 내재하는 메모리 효과 때문에 이러한 LPC 변형에 의해 확인되는 많은 샘플에 충격을 준다.

[0063] 도 3c에 도시된 상기 오디오 인코더는, 제1 인코딩된 신호를 생성하는 제1 코딩 브랜치(400)를 포함한다. 이러한 제1 인코딩된 신호는, 바람직하게는, 시간 스펙트럼 도메인, 즉 시간/주파수 변환에 따라 시간 도메인 신호

가 처리될 때 획득되는 도메인인 제4 도메인일 수 있다.

- [0064] 그러므로, 상기 오디오 신호의 인코딩을 위한 제1 코딩 브랜치(400)는, 제1 인코딩된 신호를 얻기 위하여 제1 코딩된 알고리즘을 사용한다. 여기서 제1 코딩 알고리즘은 시간/주파수 변환 알고리즘을 포함할 수 있고, 포함하지 않을 수도 있다.

- [0065] 뿐만 아니라 상기 오디오 인코더는, 오디오 입력 신호의 일부에 있어서 블록 400의 출력에서의 제1 인코딩된 신호 또는 제2 인코딩 브랜치의 출력에서의 제2 인코딩된 신호가 하나의 인코더의 출력 신호에 포함될 수 있도록, 상기 제1 코딩 브랜치(400) 및 상기 제2 코딩 브랜치(500) 간의 전환을 수행하기 위한 제1 스위치(200)를 포함한다. 그러므로 상기 오디오 입력 신호(195)의 일부에 있어서 상기 제4 도메인 내의 상기 제1 인코딩된 신호가 상기 인코더의 출력 신호에 포함될 때, 상기 제2 도메인 내의 상기 제1 처리된 신호이거나 또는 상기 제3 도메인 내의 상기 제2 처리된 신호인 상기 제2 인코딩된 신호는 상기 인코더 출력 신호에 포함되지 않는다. 이는 이러한 인코더의 비트 레이트에 있어서 효율적임을 확실하게 한다. 실시예에 있어서, 두 개의 다른 인코딩된 신호 내부에 포함된 오디오 신호의 특정 시간 일부는 하나의 프레임의 프레임 길이에 비교하여 작다. 이는 도 3e를 참조하여 서술할 것이다. 이는 작은 부분은, 어떤 크로스페이드 없이 발생할 수 있는 인공 구조를 줄이기 위해서 전환 이벤트의 경우에, 하나의 인코딩된 신호에서 다른 인코딩된 신호로의 크로스페이드에 유용하다. 그러므로 상기 크로스페이드 영역을 제외하고 각각의 시간 도메인 블록은 오직 단일 도메인의 인코딩된 신호에 의해 표현된다.

- [0066] 도 3c에 도시된 바와 같이, 상기 제2 코딩 브랜치(500)는 제1 도메인 내에서의 오디오 신호, 즉, 제2 도메인로의 신호 195를 변환하기 위한 컨버터(510)를 포함하고 있다. 뿐만 아니라 상기 제2 코딩 브랜치(500)는, 상기 제1 처리 브랜치(522)가 도메인 변화를 수행하지 않도록, 바람직하게는 상기 제2 도메인 내에도 위치하는 제1 처리된 신호를 얻기 위하여 상기 제2 도메인 내의 오디오 신호를 처리하기 위한 제1 처리 브랜치(522)를 포함한다.

- [0067] 뿐만 아니라 상기 제2 인코딩 브랜치(500)는, 제2 도메인 내의 상기 오디오 신호를 제3 도메인으로 변환하는 제2 처리 브랜치(523, 524)를 포함하고 있는데, 여기서 상기 제3 도메인은 상기 제1 도메인 및 상기 제2 도메인과 다르며, 상기 제2 처리 브랜치(523, 524)의 출력에서 제2 처리된 신호를 얻기 위해 제3 도메인 내에서 오디오 신호를 처리한다.

- [0068] 더구나, 상기 제2 코딩 브랜치는, 상기 제2 코딩 브랜치로 입력되는 상기 오디오 신호의 일부에 대하여 상기 제2 도메인 내에서 상기 제1 처리된 신호 또는 상기 제3 도메인 내에서 상기 제2 처리된 신호가 상기 제2 인코딩된 신호이기 위해서, 상기 제1 처리 브랜치(522) 및 상기 제2 처리 브랜치(523, 524) 간의 변환을 위한 제2 스위치(521)를 포함하는 것을 특징으로 한다.

- [0069] 도 3d는 도 3c의 인코더에 의해 생성된 인코딩된 오디오 신호를 디코딩하기 위한 상응하는 디코더를 도시하고 있다. 일반적으로 제1 도메인의 각각의 블록에서 오디오 신호는, 중요한 샘플링 한계에서 가능한 한 많은 시스템을 얻기 위해서, 바람직하게는 하나의 프레임의 길이에 비해 짧은, 선택적인 크로스페이드 영역을 제외하는, 제2 도메인 신호, 제3 도메인 신호 또는 제4 도메인 인코딩된 신호에 의해 표현된다. 상기 인코딩된 오디오 신호는 제1 코딩된 신호, 제2 도메인 내의 제2 코딩된 신호 및 제3 도메인 내의 제3 코딩된 신호를 포함하고, 여기서 상기 제1 코딩된 신호, 상기 제2 코딩된 신호 및 상기 제3 코딩된 신호 모두는 상기 디코딩된 오디오 신호의 다른 시간 부분에 관련되어 있으며, 여기서 디코딩된 오디오 신호를 위한 상기 제2 도메인, 상기 제3 도메인 및 상기 제4 도메인은 서로서로 차별된다.

- [0070] 상기 디코더는, 제1 코딩 알고리즘을 기반으로 하여 디코딩하기 위한 제1 디코딩 브랜치를 포함한다. 상기 제1

디코딩 브랜치는 도 3d의 431, 440에 표현되어 있고, 바람직하게는 주파수/시간 컨버터를 포함하고 있다. 상기 제1 코딩된 신호는 바람직하게는 제4 도메인 내에 있으며 디코딩된 출력 신호를 위한 도메인인 제1 도메인으로 변환된다.

[0071] 뿐만 아니라 도 3d의 디코더는 몇몇 구성요소를 포함하는 제2 디코딩 브랜치를 포함하고 있다. 이러한 구성요소는 블록 531의 출력에서 상기 제2 도메인 내의 제1 역처리 신호를 구하기 위해 상기 제2 코딩된 신호를 역처리하기 위한 제1 역처리 브랜치(531)이다. 상기 제2 디코딩 브랜치는, 뿐만 아니라, 상기 제2 도메인에서 제2 역처리된 신호를 구하기 위하여 제3 코딩된 신호를 역처리하기 위한 제2 역처리 브랜치 533, 534를 포함한다. 여기서 상기 제2 역처리 브랜치는 상기 제3 도메인을 제2 도메인으로 변환하기 위한 컨버터를 포함하고 있다.

[0072] 상기 제2 디코딩 브랜치는, 또한, 상기 제2 도메인에서 신호를 얻기 위하여 상기 제 역처리된 신호 및 상기 제2 역처리된 신호를 결합하기 위한 제1 결합자(532)를 포함하고 있다. 여기서 상기 결합되는 신호는, 처음 일순간에는 오직 상기 제1 역처리 신호에 의해서 영향받고, 그 후에는, 오직 상기 제2 역처리 신호에 의해서 영향받는다.

[0073] 상기 제2 디코딩 브랜치는, 뿐만 아니라 상기 결합된 신호를 상기 제1 도메인으로 변환하기 위한 컨버터(540)를 포함하고 있다.

[0074] 마지막으로 도 3d에 도시된 디코더는, 상기 제1 도메인 내에서 디코딩된 출력 신호를 얻기 위하여 블록 431, 440으로부터 얻어진 상기 디코딩된 제1 신호 및 컨버터(540)에서 출력된 신호를 결합하는 제2 결합자(600)를 포함하고 있다. 다시, 상기 제1 도메인에서 디코딩된 출력 신호는, 처음에는 오직 상기 컨버터(540)에서 출력된 신호에 의해서 영향받고, 그 후에는 오직 블록 431, 440에서 출력된 제1 디코딩된 신호에 의해서 영향받는다.

[0075] 인코더의 관점에서 이러한 상황이 도 3e에 도시되어 있다. 도 3e의 상단 부분은 개념적 표현에 의하여, 시간 도메인 오디오 신호 등의 제1 도메인 오디오 신호를 도시하고 있다. 여기서 타임 인덱스는 좌측에서 우측으로 갈수록 증가하며 아이템 3은 도 3c의 상기 신호 195를 나타내는 오디오 샘플의 스트림으로 고려될 수 있다. 도 3e는, 도 3e의 아이템 4에서 도시되어있는 바와 같이 제1 인코딩 신호, 제1 처리된 신호 및 제2 처리된 신호 간의 전환에 의하여 생성되는 프레임 3a, 3b, 3c, 3d를 도시한 것이다. 상기 제1 인코딩된 신호, 제1 처리된 신호 및 제2 처리된 신호는 모두 다른 도메인이며, 상기 다른 도메인 간의 전환이 디코딩 면에서 인공물이 되지 않도록 하기 위해서, 시간 도메인 신호의 프레임 3a, 3b는 크로스페이드 영역으로서 지시되는 겹치는 구간을 포함한다. 그리고 크로스페이드 영역은 프레임 3b 및 3c에 위치한다. 그러나 어떠한 크로스페이드 영역도 프레임 3d, 3c 간에는 존재하지 않으며, 이는 프레임 3d는 또한 제2 처리된 신호, 즉 제3 도메인의 신호에 의해 표현되며, 프레임 3c 및 3d 사이에는 어떠한 도메인 변화도 없음을 의미한다. 그러므로 일반적으로 어떠한 도메인 변화가 없는 크로스페이드 영역을 제공하지 않고, 도메인 변화, 즉 두 개의 스위치 중 하나에 의한 전환 행동이 있을 때 두 개의 이어지는 코딩된/처리된 신호에 의해 인코딩된 오디오 신호의 일부인 크로스페이드 영역을 제공하는 것이 선호된다. 바람직하게는 다른 도메인 변화에서도 크로스페이드는 수행된다.

[0076] 예를 들어 50 퍼센트의 중첩을 구비한 MDCT 실행에 의해 생성된 상기 제1 인코딩된 신호 또는 상기 제2 처리된 신호에 대한 실시예에 있어서, 각각의 시간 도메인 샘플은 두 개의 이어지는 프레임 내에 포함된다. 그러나 MDCT의 특성 때문에 이것은 오버헤드(overhead)하는 결과를 낳지는 않는다. 왜냐하면, 상기 MDCT는 임계적으로 샘플링된 시스템이기 때문이다. 이 콘텍스트에서, 임계적으로 샘플링되었다는 것은 스펙트럼 값의 숫자가 시간 도메인 값의 숫자와 동일하다는 것을 의미한다. 상기 MDCT는, 하나의 MCDT 블록으로부터 다음 MCDT 블록까지의 크로스오버가, 상기 임계적으로 샘플링된 요건 사항을 위반하는 어떠한 오버헤드없이 제공되기 위하여 구체적인 크로스오버 영역 없이 상기 크로스오버 효과가 얻어진다는 점에서 이점이 있다.

- [0077] 바람직하게 상기 제1 코딩 브랜치 내의 상기 제1 코딩 알고리즘은 정보 싱크 모델을 기반으로 하고, 상기 제2 코딩 브랜치 내의 상기 제2 코딩 알고리즘은 정보 소스 또는 SNR 모델을 기반으로 한다. 하나의 SNR 모델은 구체적인 소리 생성 메커니즘에 구체적으로 관련된 것은 아니지만, 닫힌 루프 결정 등을 기반으로 하는, 다수의 코딩 모드 사이에서 선택된 하나의 코딩 모드일 수 있다. 그러므로 하나의 SNR 모델은 가능한 코딩 모델이나, 상기 소리 생성 수단의 물리적 구조와 반드시 관계될 필요는 없고, 상기 정보 싱크 모델과 다른 매개변수화 한 코딩 모델이어야 한다. 그리고 그것은 닫힌 루프 결정에 의해서, 구체적으로는 다른 모델과 SNR 결과의 차이를 비교함으로써 선택될 수 있어야 한다.
- [0078] 도 3c에 도시된 바에 따르면, 제어기(300, 525)가 제공된다. 이러한 제어기는 도 1a의 결정 스테이지(300)의 기능을 포함할 수 있고, 추가적으로 도 1a의 스위치 제어 장치(525)의 기능을 포함할 수 있다. 일반적으로, 상기 제어기는, 신호 적용 방법에 있어서 상기 제1 스위치 및 상기 제2 스위치를 제어하기 위한 것이다. 상기 제어기는, 타겟 함수에 관하여 상기 제1 스위치로 입력되는 신호 또는 상기 제1 또는 제2 코딩 브랜치로부터 출력되는 신호, 또는 상기 제1 및 제2 인코딩 브랜치로부터 인코딩 및 디코딩됨으로써 구해진 신호를 분석하기 위해 작동한다. 그렇지 않다면 또는 추가적으로, 타겟 함수에 관하여 상기 제어기는 상기 제2 스위치에 입력되는 신호, 상기 제1 처리 브랜치 또는 상기 제2 처리 브랜치에 의해 출력되는 신호 또는 상기 제1 처리 브랜치 및 상기 제2 처리 브랜치로부터 처리 또는 역처리됨으로써 얻어진 신호를 분석하기 위해 작동한다.
- [0079] 일 실시예에 있어서, 상기 제1 코딩 브랜치 또는 상기 제2 코딩 브랜치는, MDCT 또는 MDST 알고리즘 같은 에일리어싱 도입 시간/주파수 변환 알고리즘을 포함할 수 있다. 이는 어떠한 에일리어싱 효과를 도입하지 못하는 간단한 FFT 변형과는 다르다. 이뿐만 아니라, 하나 또는 두 개의 브랜치는 양자화기/엔트로피 코드 블록을 포함할 수 있다. 구체적으로 상기 제2 코딩 브랜치의 상기 제2 처리 브랜치만이 에일리어싱 작동을 도입하는 시간/주파수 컨버터를 포함하고, 상기 제2 코딩 브랜치의 상기 제1 처리 브랜치는 양자화기 및/또는 엔트로피 인코더를 포함하고 있고 어떠한 에일리어싱 효과도 도입하지 않는다. 상기 에일리어싱 도입 시간/주파수 컨버터는, 바람직하게는 분석 윈도우 및 MDCT 변형 알고리즘을 적용하기 위한 윈도우어(windower)를 포함할 수 있다. 구체적으로 상기 윈도우어는, 적어도 두 개의 이어지는 윈도우화 된 프레임에서 윈도우화된 신호의 샘플이 발생하도록, 상기 윈도우 기능을, 중첩된 연속되는 프레임에 적용하도록 동작한다.
- [0080] 일 실시예에 있어서, 상기 제1 처리 브랜치는 ACELP 코더를 포함하고 제2 처리 브랜치는, 양자화된 스펙트럼 요소를 얻기 위하여 MDCT 스펙트럼 컨버터 및 스펙트럼 요소를 양자화하기 위한 양자화기를 포함하고 있다. 여기서 각각의 양자화된 스펙트럼 요소는 0 또는 다수의 다른 가능한 양자화 인덱스의 일 양자화 인덱스에 의해 정의된다.
- [0081] 뿐만 아니라, 바람직하게는 상기 제1 스위치(200)는 열린 루프 방식에서 작동하고, 상기 제2 스위치는 닫힌 루프 방식에서 작동한다.
- [0082] 기상술한 바에 따라, 양 코딩 브랜치는 블록의 면에서 오디오 신호를 인코딩하는데 이용된다. 여기서 상기 제1 스위치 또는 상기 제2 스위치는, 블록의 면에서, 전환 행동이, 신호의 샘플에 대한 기정의된 숫자의 블록 뒤에, 최소한으로 발생하도록, 전환한다. 여기서 상기 기정의된 숫자는 상응하는 스위치에 대한 프레임 길이를 형성한다. 그러므로 상기 제1 스위치에 의해 전환되기 위한 소부분(granule)은, 예를 들어 2048 또는 1028 샘플에 대한 블록일 수 있고, 상기 프레임 길이는, 상기 제1 스위치(200)가 전환하는 것을 기반으로 하면서, 변화될 수 있으나, 바람직하게는 꽤 긴 기간 동안 고정되어 있다.
- [0083] 그것에 반대되도록, 상기 제2 스위치(521)에 대한 블록 길이는, 즉 상기 제2 스위치(521)가 하나의 모드에서 다른 모드로 전환할 때, 상기 제1 스위치가 블록 길이보다 상당히 작다. 바람직하게는, 상기 스위치에 대한 양 블록의 길이는, 더 짧은 블록 길이의 정수곱이 되도록 선택된다. 바람직한 실시예에 있어서 제1 스위치에 대한 블록 길이는 2048 또는 1024이고, 제2 스위치의 블록 길이는 1024 또는 더 바람직하게는 512이고, 더욱 바람직하

게는 256이며, 더욱더 바람직하게는 128 샘플이다. 이는, 상기 제1 스위치가 오직 한번 전환할 때, 상기 제2 스위치가 최대한 16번 전환할 수 있도록 하기 위함이다. 그러나 바람직한 최대 블록 길이 비율은 4:1이다.

[0084] 다른 실시예에 있어서 상기 제어기(300, 525)는, 음악에 대한 결정에 관하여 발화에 대한 결정이 선호되는 것과 같은 방법으로 상기 제1 스위치에 대한 발화, 음악 차별을 수행한다. 이러한 실시예에서, 상기 제1 스위치에 대한 프레임의 50%더욱 낮은 비율이 발화이고, 상기 프레임의 50% 이상의 비율이 음악인 경우에서조차 발화로 결정하도록 할 수 있다.

[0085] 뿐만 아니라, 상기 제어기는, 상기 제1 프레임의 상당 작은 부분이 발화일 때, 구체적으로 상기 더 작은 제2 프레임의 길이의 50%인 상기 제1 프레임의 일부가 발화일 때, 발화 모드로 기 전환하도록 작동할 수 있다. 그러므로 바람직한 발화/선호의 전환 결정은, 예를 들어 상기 제1 스위치의 프레임 길이에 대응하는 블록의 오직 6% 또는 12%만이 발화인 경우에조차, 발화로 이미 전환된다.

[0086] 이러한 절차는, 일 실시예에 있어서 유성 발화 코어를 구비하는 상기 제1 처리 브랜치의 비트레이트 저장 가능 용량을 완전하게 이용한다는 점 및, 제2 처리 브랜치가 컨버터를 구비하고 그러므로 마찬가지로 발화가 아닌 신호를 구비한 오디오 신호에 대해서 이용된다는 사실 때문에, 발화가 아닌 큰 제1 프레임의 나머지에 대해서 품질을 떨어뜨리지 않는다는 점에서 바람직하다. 바람직하게는 이러한 제2 처리 브랜치는, 임계적으로 샘플링되고, 에일리어싱 취소가 디코더의 면에서 중첩 및 추가됨으로써 실행되는 동안에 시간 도메인 덕분에 작은 윈도우 사이즈에서조차 고효율의 에일리어싱에서 자유로운 작동을 제공하는 중첩되는 MDCT를 포함한다. 뿐만 아니라 바람직하게는 AAC 유사한 MDCT 인코딩 브랜치인 상기 제1 인코딩 브랜치에 대한 큰 블록 길이가 유용하다. 왜냐하면 발화가 아닌 신호는 일반적으로 상당히 정적이며, 긴 변형 윈도우는 음향 심리학적으로 제어되는 양자화 모듈 덕분에 고주파수 해상도 및 이에 따른 고품질을 제공하고, 추가적으로 비트레이트의 효율성을 제공하기 때문이다. 여기서 상기 양자화 모듈은, 상기 제2 코딩 브랜치의 제2 처리 브랜치 내의 상기 변형 기반의 코딩 모듈에도 적용될 수 있다.

[0087] 도 3d의 디코더 도식을 참조하면, 바람직하게는 상기 전송된 신호는 도 3e에 도시된 것과 같이 사이드 정보 4a로써 명백한 지시자를 포함한다. 이러한 사이드 정보 4a는, 도 3d에 도시된 제1 디코딩 브랜치, 제1 역처리 브랜치 또는 제2 역처리 브랜치같은 정확한 프로세서에 대한 상응한 제1 인코딩된 신호, 제1 처리된 신호 또는 제2 처리된 신호를 전송하기 위하여, 도 3d에는 도시되지 않은 비트 스트림 파서(parser)에 의해 추출될 수 있다. 그러므로, 하나의 인코딩된 신호는 상기 인코딩된/처리된 신호를 구비하고 있을 뿐만 아니라 이러한 신호에 관련된 사이드 정보도 포함하고 있다. 그러나 다른 일 실시예에 있어서, 디코더 면의 비트 스트림 파서가 상기 신호들 간에 차별하도록 하는 내포된 신호화가 있을 수 있다. 도 3e를 참조하면 상기 제1 처리된 신호 또는 상기 제2 처리된 신호는 상기 제2 코딩 브랜치 및 이에 따른 제2 코딩된 신호이라는 것에 대해 설명되어 있다.

[0088] 바람직하게는 상기 제1 디코딩 브랜치 및/또는 상기 제2 역처리 브랜치는, 상기 스펙트럼 도메인으로부터 상기 시간 도메인으로 변환하기 위한 MDCT 변형을 포함할 수 있다. 마지막에는 중첩 애더(첨가기, adder)가, 동시에, 블록의 인공물을 회피하기 위해 크로스페이드 효과를 제공하는 시간 도메인 에일리어싱 취소 기능을 수행하도록 제공될 수 있다. 일반적으로, 상기 제1 디코딩 브랜치는, 상기 제4 도메인에서 인코딩된 신호를 상기 제1 도메인으로 변환한다. 반면에 상기 제2 역처리 브랜치는 상기 제3 도메인으로부터 상기 제2 도메인으로의 변환을 수행한다. 그리고 상기 제1 결합자 다음으로 연결되어 있는 상기 컨버터는 상기 제2 도메인으로부터 상기 제1 도메인에 대한 변환을 제공한다. 이는 결합자(600)에 대한 입력에서 도 3d의 실시예에서 디코딩된 출력 신호만을 나타내는 오직 제1 도메인 신호만이 존재하도록 하기 위함이다.

[0089] 도 4a 및 4b는 두 개의 다른 실시예를 도시하고 있다. 이들은 상기 스위치(200)의 위치에서 차이가 있다. 도 4a에서 상기 스위치(200)는 일반적인 전처리 스테이지(100)의 출력 및 두 개의 인코딩된 브랜치(400, 500)의 입력 사이에 위치한다. 도 4a의 실시예는 단일한 인코딩 브랜치로만 상기 오디오 신호가 입력되고, 상기 일반적인 전

처리 스테이지의 출력과 연결되지 않은 다른 인코딩 브랜치는 작동하지 않으며, 그러므로 스위치는 꺼지고 수면 모드에 있는 것을 확실히 하고 있다. 이러한 실시예에서는 바람직하게도, 상기 비활동적인 인코딩 브랜치는, 특히 배터리로 전원을 공급받고, 그럼으로써 파워의 소모에 있어서 일반적인 한계를 가진 모바일 장치에 필요한, 전력 및 컴퓨팅에 필요한 자원을 소비하지 않는다.

[0090] 그러나 한편으로는 전력 소비가 별다른 문제가 되지 않을 때에는, 상기 도 4b의 실시예가 선호될 수 있다. 이러한 실시예에서는 양 인코딩 브랜치 (400, 500)은 언제나 작동하며, 어떤 시간 부분 및/또는 어떤 주파수 부분에 대한 선택된 인코딩 브랜치의 출력만이, 비트 스트림 멀티플렉서(800)으로서 구현되는 상기 비트스트림 포맷터 (formatter)로 전송된다. 그러므로 다른 선택되지 않은 인코딩 브랜치(400)의 출력이 폐기되고, 인코딩된 오디오 신호 등의 상기 출력 비트스트림에 삽입되지 않는 동안, 도 4b의 실시예에 있어서 인코딩 브랜치 둘 모두 언제나 작동하며, 결정 스테이지 300에 의해 선택된 인코딩 브랜치의 출력은 상기 출력 비트스트림으로 삽입된다.

[0091] 도 4c는 최선의 디코더의 다른 관점을 도시한 것이다. 상기 제1 디코더가 시간-에일리어싱 생성하는 디코더이거나 또는 일반적으로 언급되는 주파수 도메인 디코더이고 상기 제2 디코더는 시간 도메인 장치인 상황에서, 구체적으로 들을 수 있는 인공물을 피하기 위하여 제1 디코더(450) 및 상기 제2 디코더(550)에 의해 출력되는 블록 또는 프레임 간의 경계는, 구체적으로 전환되는 상황에서 완전히 계속적이지 않다. 그래서 상기 제1 디코더 (450)의 제1 블록이 출력되고, 이어지는 시간 비율에 있어서 상기 제2 디코더의 블록이 출력될 때, 크로스페이드 블록(670)에 의해 도시된 것처럼 크로스페이드 작동을 수행하는 것이 바람직하다. 마지막으로, 상기 크로스페이드 블록(607)은 도 4c에 도시된 바와 같이, 607a, 607b 및 607c로 구현될 수 있다. 각각의 브랜치는 일반화된 스케일에 따라서 0과 1 사이의 가중 인자 m_1 을 구비한 웨이터(weighter)를 포함할 수 있다. 여기서 상기 가중 인자는 플롯 609에서 제시된 바와 같이 변화할 수 있다. 상기 크로스페이드 규칙은 추가적으로 사용자가 어떠한 음의 고저 변화도 감지하지 않도록 하는 지속적이고 부드러운 크로스페이드가 발생하도록 한다. \sin^2 크로스페이드 규칙같은 비선형 크로스페이드 규칙은 선형 크로스페이드 규칙 대신에 적용될 수 있다.

[0092] 일례에 있어서, 상기 제1 디코더의 마지막 블록은 윈도우를 사용해서 생성될 수 있고, 여기서 상기 윈도우는 이 블록의 페이드 아웃을 수행한다. 이러한 경우에서 블록 607a 내의 상기 가중 인자 m_1 은 1과 동일하고, 실질적으로 이러한 브랜치에는 어떠한 가중치도 요청되지 않는다.

[0093] 상기 제2 디코더에서 상기 제1 디코더로의 전환이 발생할 때, 그리고 상기 제2 디코더가 상기 블록의 말단에서 상기 출력을 실질적으로 페이드아웃시키는 윈도우를 포함할 때, 그때 " m_2 "로 지시되는 상기 웨이터는 요청되지 않거나 또는 상기 가중 매개변수는 전체적인 크로스페이드 구역을 통하여 1로 설정된다.

[0094] 스위치 후에 위치하는 제1 블록은, 윈도우 작동을 통하여 생성되고, 이러한 윈도우가 작동 중 실질적으로 페이드(fade)를 수행할 때, 상기 상응하는 가중 인자는 또한 1로 설정되고 웨이터는 불필요하다. 그러므로, 상기 디코더에 의해서 페이드 아웃하기 위하여 상기 마지막 블록이 윈도우화되고, 상기 스위치 뒤의 제1 블록이 페이드-인을 제공하기 위하여 상기 디코더를 이용하여 윈도우화 될 때, 상기 웨이터 607a, 607b는 전혀 요구되지 않으며, 애더(adder, 607c)에 의한 첨가 작용으로 충분하다.

[0095] 이러한 경우에 있어서 상기 마지막 프레임의 페이드-아웃되는 부분과 다음 프레임의 페이드-인되는 부분은, 블록 609에서 지시되는 상기 크로스페이드 영역을 정의한다. 뿐만 아니라 그러한 상황에서는 바람직하게는 하나의 디코더의 마지막 블록은 다음 디코더의 첫번째 블록과 일정하게 시간상으로 중첩된다.

[0096] 만약 크로스페이드 작업이 요청되지 않거나, 불가능하거나 또는 바람직하지 않고, 만약 하나의 디코더에서 다른 디코더로 전환하는 어려운 스위치만이 위치한다면, 그러한 스위치는, 상기 오디오 신호의 조용한 통로(passage)

또는 최소한 낮은 에너지가 존재하는 오디오 신호의 통로, (즉, 조용하거나 또는 거의 조용한 것으로 감지된다)에서 수행되는 것이 바람직하다. 바람직하게는 상기 결정 단계(300)는 일실시예에 있어서, 상기 스위치 이벤트 뒤의 상응하는 시간 부분, 예를 들어 오디오 신호의 평균 에너지보다 낮고, 더 바람직하게는, 예를 들어 둘 또는 그 이상의, 상기 오디오 신호에 대한 시간 부분/ 프레임에 관련되어 있는 오디오 신호의 평균 에너지의 50%보다 낮은 에너지를 구비할 때, 상기 스위치(200)가 오직 활성화되도록 확실히 할 수 있다.

[0097] 바람직하게는 상기 제2 인코딩 규칙/디코딩 규칙은 LPC 기반의 코딩 알고리즘이다. LPC 기반의 발화 코딩에서, 준(準)주기적인 맥박형 여자 신호 세그먼트 또는 신호의 부분 및 잡음형의 여자 신호 세그먼트 또는 신호 부분 간의 차이가 생긴다. 이는 도 7b에 도시된 바와 같이 매우 낮은 비트 레이트의 LPC 보코더(vocoder, 2.4 kbps)에 대하여 행해진다. 그러나 중간 레이트의 CELP 코더에서는 상기 여자는, 조정되는 코드북 및 고정된 코드북으로부터 스케일 된 벡터의 추가로서 얻어진다.

[0098] 준주기적인 맥박형 여자 신호 세그먼트, 즉 특유의 피치를 구비한 신호 세그먼트는, 잡음형의 여자 신호 세그먼트와는 다른 메커니즘을 통하여 코딩된다. 준주기적인 맥박형 여자 신호가 유성 발화에 관련되어 있는 반면에, 잡음형 신호는 무성 발화에 관련되어 있다.

[0099] 전형적으로 도 5a에 도 5d까지는 참고 자료가 도시되어 있다. 여기서, 준주기적인 맥박형 신호 세그먼트 또는 신호 부분 및 잡음형 신호 세그먼트 또는 신호 부분이 전형적으로 논의된다. 구체적으로 도 5a에 도시된 시간 도메인 내의 유성 발화 및 도 5b에 도시된 주파수 도메인 내의 유성 발화는 준주기적인 맥박형 신호 부분의 일례로써 토론되고, 도 5c 및 도 5d를 참조하여 무성 발화 세그먼트는 잡음형 신호 부분의 일례로써 토론된다. 발화는 일반적으로 유성, 무성 또는 혼합으로 분류된다. 샘플링된 유성 및 무성 세그먼트에 대한 시간-및-주파수 도메인의 도면이 도 5a에서 도 5d에 걸쳐 도시되어 있다. 무성 발화가 임의적이고 브로드밴드인 반면에, 유성 발화는 시간 도메인 내에서 준주기적이고 상기 주파수 도메인 내에서 화음으로 구축된다. 상기 유성 발화의 단 시간 스펙트럼은 잘 조화된 포먼트(formant) 구조에 의해서 특성화된다. 상기 화음 구조는 준주기적인 발화의 결과이고, 진동하는 보컬 화음(vocal chord)에 기여한다. 상기 포먼트 구조(스펙트럼 인벨롭)은 상기 소스 및 성도 간의 상호작용에 기인한다. 상기 성도는 인두 및 구개로 구성된다. 유성 발화의 단 시간 스펙트럼에 적합한 스펙트럼 인벨롭의 형상은, 성문음의 펄스 때문에 성도 및 스펙트럼 틸트(6 dB / 옥타브)의 전이 특성에 관련되어 있다. 상기 스펙트럼 인벨롭은 포먼트라고 불리는 일련의 피크(peak)에 의해서 특징지어진다. 상기 포먼트는 성도의 공명 모드이다. 평균적인 성도에 대해서 5 kHz 이하의 3 ~ 5의 포먼트가 있다. 3 kHz 이하에서 종종 발생하는, 제1 세 포먼트의 진폭 및 위치는 발화 통합 및 인지의 양면에서 상당히 중요하다. 더 높은 포먼트는 또한 넓은 대역 및 무성 발화 표현을 위해서 중요하다. 발화의 특징은, 다음에 설명하는 것과 같이 물리적인 발화 생성 시스템에 관련되어 있다. 유성 발화는, 상기 진동하는 보컬 화음에 의해 생성되는 준주기적인 성문음의 공기 펄스로써 상기 성도를 흥분시킴으로써 생성된다. 상기 주기적인 펄스의 주파수는, 기본적인 주파수 또는 피치라고 언급된다. 무성 발화는 상기 성도의 수축을 통하여 공기에 힘을 가함으로써 생성된다. 코의 소리는, 음성관에 코의 관을 음향학적으로 결합하는 것에 기인한다. 그리고 파열음은, 상기 관이 폐쇄된 후에 생성되는 공기 압력을 갑자기 유출시킴으로써 생성된다.

[0100] 그러므로 오디오 신호의 노이즈 형의 부분은, 도 5c 및 도 5d에 도시된 바와 같이 어떠한 맥박 형 시간 도메인 구조 또는 화음 주파수 도메인 구조도 보이지 않는다. 그리고 그것은 도 5a 및 도 5b에 일례로 도시된 준주기적인 맥박형 부분과도 다르다. 그러나 후술하는 바와 같이 상기 노이즈형 부분 및 준주기적인 맥박형 부분 간의 차이는 또한 상기 여자 신호에 대한 LPC 이후에 관찰될 수 있다. 상기 LPC는 상기 성도를 모델로 하고, 신호로부터 성도의 여자를 추출하는 방법이다.

[0101] 뿐만 아니라, 준주기적인 맥박형 부분 및 노이즈형의 부분은 적절한 때에 발생할 수 있다. 이는 어떤 시간의 오디오 신호의 부분은 노이즈이고, 어떤 시간의 다른 오디오 신호의 부분은 준주기적, 즉 음조인 것을 의미한다. 그렇지 않으면 또는 이에 더하여 신호의 특징은 다른 주파수 대역 내에서 다를 수 있다. 그러므로 오디오 신호가 노이즈 또는 음조인지(tonal) 여부에 대한 결정은 또한, 특정 주파수 대역 또는 몇몇 주파수 대역이 노이즈

한 것으로 고려되고, 다른 주파수 대역은 음조인 것으로 고려될 수 있도록 주파수 선택을 수행할 수 있다. 이 경우에 있어서 오디오 신호의 어떤 시간 부분은 음조 요소 및 노이즈 요소를 포함할 수 있다.

[0102] 도 7a는 발화 생성 시스템의 선형 모델을 도시하고 있다. 이러한 모델은, 도 7c에 지시된 바와 같이 유성 발화에 대한 맥박열(列) 및 도 7d에 지시된 바와 같이 무성 발화에 대한 랜덤 노이즈 등의 두 개의 스테이지의 여자를 추정한다. 도 7c 및 도 7d의 맥박을 처리하고, 성문음 모델(72)에 의해 생성되는 올폴 필터(all-pole filter)로써 모델링된다. 그러므로 도 7a의 시스템은 개인 스테이지(77), 전진 통로(78) 피드백 통로(79) 및 추가 스테이지(80)를 구비한 도 7b의 올폴필터 모델로 축소될 수 있다. 피드백 통로(79)에는 예측 필터(81)가 있고, 도 7b에 도시된 전체적인 소스 모델 합성 시스템은 아래와 같은 z -도메인 함수에 의해 표현될 수 있다.

[0103]
$$S(z) = g / (1-A(z)) \cdot X(z)$$

[0104] 여기서 g 는 개인(gain)을 의미하고, $A(z)$ 는 LP 분석에 의해 결정된 예측 필터를 의미하며, $X(z)$ 는 여자 신호를 의미하고, $S(z)$ 는 합성 발화 출력을 의미한다.

[0105] 도 7c 및 도 7d에서는, 선형 소스 시스템 모델을 이용한 유성 또는 무성 발화 합성에 대한, 시각적인 시간 도메인에 대한 묘사를 볼 수 있다. 이러한 시스템 및 상기 방정식 내의 여자 매개변수는 알려지지 않았으나, 유한 개의 발화 샘플로부터 결정되어야만 한다. 상기 $A(z)$ 의 계수는 입력 신호의 선형 예측 및 필터 계수의 양자화를 통하여 얻어진다. 선형 예측기 쪽으로의 p 번째 순서에서, 상기 발화 시퀀스의 본 샘플은 p 번 통과한 샘플의 선형 결합으로부터 예측된다. 상기 예측 계수는, 레빈슨-더빈 알고리즘(Levinson-Durbin Algorithm)같은 잘 알려진 알고리즘에 의해서 또는 일반적으로 자기 상관 수단 또는 반영 수단 등에 의해서 결정될 수 있다.

[0106] 도 7e는 LPC 분석 블록(510)에 대한 더 자세한 구현을 도시한 것이다. 상기 오디오 신호는 상기 필터 정보 $A(z)$ 를 결정하는 필터 결정 블록 내로 입력된다. 이러한 정보는 디코더를 위해 필요한 단기 예측 정보로써 출력된다. 상기 단기 예측 정보는, 실제의 예측 필터(85)에 의해서 요구된다. 서브트랙터(86)에서 오디오 신호의 현재의 샘플은 입력되고, 현재의 샘플에 대한 예측값은, 이러한 샘플에 있어서 상기 예측 오류 신호가 라인 84에서 생성되도록, 제하게 된다. 상기 일련의 예측 오류 신호 샘플이 도 7c 또는 도 7d에 매우 도식적으로 도시되어 있다. 그러므로 도 7a, 7b는 교정된 맥박형 신호로써 고려될 수 있다.

[0107] 도 7e는 상기 여자 신호를 계산하는 바람직한 방법에 대해 도시하고 있는 반면에, 도 7f는 상기 가중된 신호를 계산하는 바람직한 방법에 대해 도시하고 있다. 도 7e와는 반대로, g 가 1과 다를 때, 상기 필터(85)는 다르다. 1보다 작은 값은 g 를 위해서 바람직하다. 일반적으로 도 7e 및 도 7f의 구성요소들은 3GPP TS 26.190 또는 3GPP TS 26.290.에서 구현될 수 있다.

[0108] 도 7g는 도 2b에 도시된 구성요소(537) 내에서 같이 디코더의 측면에서 적용될 수 있는 역처리를 도시하고 있다. 특히, 블록 88은 상기 가중된 신호로부터 가중되지 않은 신호를 생성하고, 블록 89는 가중되지 않은 신호로부터 여자 신호를 계산한다. 일반적으로 도 7g의 가중되지 않은 신호를 제외한 모든 신호는 LPC 도메인 내에 있으나, 상기 여자 신호 및 상기 가중 신호는 동일 도메인 내에서 다른 신호이다. 블록 99는 블록 536의 출력과 함께 사용될 수 있는 여자 신호를 출력한다. 그리고, 상기 일반적인 역 LPC 변형은 도 2b의 블록 504에서 수행될 수 있다.

[0109] 이어서, 이러한 알고리즘에 적용되는 수정을 도시하기 위해서 도 6을 참조하여 합성에 의한 분석 CELP 인코더(analysis-by-synthesis CELP encoder)를 설명한다. 상기 CELP 인코더는, 구체적으로는 "발화 코딩: 튜토리얼 리뷰"(안드레 스페니아스(Andreas Spanias) 저, Proceedings of the IEEE, Vol. 82, No. 10, October 1994,

pages 1541-1582)에서 토론되고 있다. 도 6에 도시된 바와 같이 상기 CELP 인코더는 장기 예측 요소(60) 및 단기 예측 요소(62)를 포함하고 있다. 더구나, 64에 도시된 코드북이 사용되고 있다. 각각 가중 필터 $W(z)$ 는 66에 구현되어 있고, 오류 최소화 제어기는 68에서 제공되고 있다. $s(n)$ 는 시간 도메인 입력 신호이다. 지각적으로 가중된 이후에, 상기 가중된 신호는 감산기(subtractor, 69)로 입력된다. 상기 감산기는 블록 66의 출력에서 상기 가중된 합성 신호 및 상기 원래의 가중 신호 $S_w(n)$ 사이의 오류를 연산한다. 일반적으로 상기 단기 예측 필터 계수 $A(z)$ 는 LP 분석 스테이지에 의해서 계산된다. 그리고 그의 계수는, 도 7e에 도시된 바와 같이 $A'(z)$ 내에서 양자화된다. 장기 예측 계인 g 및 코드북 레퍼런스 같은 벡터 양자화 인덱스 등을 포함하고 있는 상기 장기 예측 정보 $A_L(z)$ 는, 도 7e의 10a로써 언급된 LPC 분석 스테이지의 출력에서 예측된 에러 신호에 따라 연산된다. 상기 LTP 매개변수는 피치 딜레이 및 계인이다. CELP에서 이것은 상기 과거 여자 신호(잔여 신호는 아니다)를 포함하는 적용되는 코드북으로서 종종 구현된다. 상기 적용되는 CB 딜레이 및 계인은 평균 제공의 가중 에러(달힌 루프 피치 검색)을 최소화함으로써 구해진다.

[0110] 그리고 상기 CELP 알고리즘은, 예를 들어 가우스 시퀀스같은 코드북을 사용한 단기 및 장기 예측 이후에 얻어진 잔여 신호를 인코딩한다. 상기 ACELP 알고리즘, 여기서 "A"는 "대수"를 상징한다,은 특수한 대수적으로 디자인된 코드북을 구비한다.

[0111] 하나의 코드북은 과부족의 벡터를 포함하고, 여기서 각각의 벡터는 몇몇 긴 샘플이다. 계인 인자 g 는 상기 코드북을 조정하고, 상기 계인된 코드는 상기 장기 예측 합성 필터 및 단기 예측 합성 필터에 의해 필터링된다. 상기 "최적의" 코드 벡터는, 상기 감산기(69)의 출력에서 지각적으로 가중된 평균 제공 오류가 최소화되도록 선택된다. 상기 CELP의 검색 과정은 도 6에 도시된 합성에 의한 분석 최적화 방법에 의해 행해질 수 있다.

[0112] 하나의 프레임이 유성 또는 무성 발화의 혼합이거나 또는 음악과 겹치는 발화가 발생한 때 같은 구체적인 경우에 있어서, TCX 코딩은 LPC 도메인 내의 여자를 코딩하기 위해 더욱 적합할 수 있다. 상기 TCX 코딩은, 주파수 도메인 내에서 상기 가중된 신호를 여자 생산에 대한 어떠한 추론 없이도 처리할 수 있다. 상기 TCX는, 그리고, 더욱 CELP 코딩보다 포괄적이고, 여자에 대한 유성 또는 무성 소스 모델에 제한되지 않는다. TCX는 여전히, 발화 유사한 신호의 포먼트를 모델링하기 위한 선형 예측 필터를 사용하여 코딩하는 소스-필터 모델이다.

[0113] AMR-WB+ 같은 코딩에서, 상기 AMR-WB+ 설명으로부터 알려진 바와 같이 다른 TCX 모델 및 ACELP 사이에서의 선택이 발생한다. 상기 TCX 모델은, 상기 블록 와이즈 디크리트 푸리에 변환(Discrete Fourier Transform)의 길이가 다른 모드에 따라 다르고, 최선의 모드는 합성 접근 또는 직접적인 "피드포워드(feedforward)" 모드에 의해서 분석됨으로써 선택된다는 점에서 다르다.

[0114] 도 2a 및 도 2b를 참조하여 설명하면 상기 일반적인 전처리 스테이지(100)는, 바람직하게는 조인트 멀티 채널(서라운드/조인트 스테레오 장치, 101) 및, 추가적으로 대역폭 확장 스테이지(102)를 포함한다. 상응하여 상기 디코더는 대역폭 확장 스테이지(701) 및 이후에 연결되는 조인트 멀티채널 스테이지(702)를 포함한다. 바람직하게 상기 조인트 멀티채널 스테이지(10)는, 인코더와 관련하여, 대역폭 확장 스테이지(102)의 앞에 연결되고, 디코더의 측면에서, 상기 대역폭 확장 스테이지(701)는, 신호 처리 방향과 관련하여 조인트 멀티채널 스테이지(702) 앞에 연결된다. 그러나 그렇지 않으면 상기 일반적인 전처리 스테이지는 순차적으로 연결되는 대역폭 확장 스테이지 없는 조인트 멀티채널 스테이지를 포함하거나 또는 연결된 조인트 멀티채널 스테이지가 없는 대역폭 확장 스테이지를 포함하고 있다.

[0115] 인코더 측면 101a 및 디코더 측면 702a 및 702b에서의 조인트 멀티채널 스테이지의 바람직한 일례가 도 8의 콘텍스트에 도시되어 있다. 다수의 E 원래의 입력 채널은 다운믹서 101a로 입력되고, 상기 다운믹서는 다수의 K 전송된 채널을 생성한다. 여기서 상기 숫자 K는 1 이상이고, E 이하이다.

- [0116] 바람직하게 상기 E 입력 채널은, 매개 변수 정보를 생성하는 조인트 멀티 채널 매개 변수 분석기 (101b)내로 입력된다. 상기 매개변수 정보는 바람직하게는, 다른 인코딩 및 차후의 허프만 인코딩(Huffman encoding) 또는 이와는 다르게 차후의 산술 인코딩에 의해서 엔트로피 인코딩된다. 상기 블록 101b에 의해 출력되는 인코딩된 매개 변수 정보는, 도 2b의 아이템 702의 일부인 매개 변수 디코더(702b)로 전송된다. 상기 매개 변수 디코더 (702b)는 상기 전송된 매개변수 정보를 디코딩하고, 상기 디코딩된 매개변수 정보를 업믹서 702a로 보낸다. 상기 업믹서 702a는 상기 K 전송된 채널을 수신하고, 다수의 L 출력 채널을 생성한다. 여기서 상기 L은 K보다는 같거나 크며, E보다는 작거나 같다.
- [0117] 매개 변수 정보는, BCC 테크닉으로부터 알려지거나 MPEG 서라운드 스탠다드에서 상세히 묘사되고 알려진, 인터 채널 레벨 차이, 인터 채널 시간 차이, 인터 채널 페이스 차이 및/또는 인터 채널 일관성 수단을 포함할 수 있다. 상기 전송된 채널의 숫자는 매우 작은 비트레이트(ultra-low bit rate) 실행을 위한 단일 모노 채널일 수도 있고, 호환되는 스테레오 실행을 포함할 수 있으며 또는 2 채널 등의 호환되는 스테레오 신호일 수도 있다. 전형적으로, 상기 E 입력 채널의 숫자는 5 개일 수 있고, 이보다 더 클 수도 있다. 그렇지 않으면 E 입력 채널의 숫자는, SAOC(spatial audio object coding) 내의 컨텍스트 내에서 알려진 E 오디오 오브젝트(object)일 수도 있다.
- [0118] 일 실시예에 있어서, 상기 다운믹서는 원래의 E 입력 채널에 대한 가중된 또는 가중되지 않은 추가 또는 E 입력 오디오 오브젝트의 추가를 수행한다. 입력 채널로써의 오디오 오브젝트인 경우에 있어서, 상기 조인트 멀티채널 매개변수 분석기(101b)는, 바람직하게는 각각의 시간 부분을 위해서, 더욱 바람직하게는 각각의 주파수 대역을 위해서 오디오 오브젝트 간의 상관 계수 행렬 같은 오디오 오브젝트 매개변수를 계산할 것이다. 마지막으로 상기 전체적인 주파수 범위는 최소한 10, 바람직하게는 32 또는 64 주파수 대역으로 나누어질 수도 있다.
- [0119] 도 9는 도 2a에 도시된 대역폭 확장 스테이지(102) 및 도 2b에 도시된 상응하는 대역폭 확장 스테이지(701)의 구현에 대한 바람직한 일 실시예에 대해서 도시한 것이다. 인코더의 면에서 상기 대역폭 확장 블록(102)은 바람직하게는 저패스필터링블록(102b), 상기 저패스를 뒤따르거나 또는 QMF 대역의 오직 절반에서 활용하는 역 QMF의 부분인 다운샘플러 블록 및 고대역 분석자(102a)를 포함한다. 상기 대역폭 확장 블록(102)에 입력된 원래의 오디오 신호는 저패스 필터링되고, 인코딩 브랜치 및/또는 스위치로 입력되는 저 대역 신호를 생성한다. 상기 저패스 필터는 3 kHz에서 10 kHz 범위 내에 있을 수 있는 주파수를 절단한다. 뿐만 아니라 상기 대역폭 확장 블록(102)은 스펙트럼 인벨롭 매개 변수 정보, 노이즈 플로어 매개변수 정보, 역필터링 매개변수 정보, 고대역 내의 어떤 화음 라인에 관련된 추가적인 매개변수 정보 및 스펙트럼 밴드 복제에 관련된 챔터 내에서 MPEG-4 스탠다드 내에서 상세히 설명된 추가적인 매개 변수 등의 대역폭 확장 매개변수를 계산하기 위한 고대역 분석기를 포함한다.
- [0120] 디코더 측면에서 상기 대역폭 확장 블록(701)은 패처(701a), 조정자(adjuster, 701b) 및 결합자(701c)를 포함한다. 상기 결합자(701c)는, 상기 조정자(701b)에 의해서 출력되는 디코딩된 저대역 신호 및 재구축되고 조정된 고대역 신호를 결합한다. 상기 조정자 701b내로의 입력은, 스펙트럼 밴드 복제 또는, 일반적으로 대역폭 확장에 의하여 저 대역 신호로부터 고 대역 신호를 추출하는 패처에 의해 제공된다. 상기 패처 701a에 의해 수행되는 패칭작업은 화음 방법 또는 화음이 아닌 방법에 의해 수행된다. 상기 패처 701a에 의해 생성된 신호, L 이어서 상기 조정자 701b에 의해 상기 전송된 매개 변수의 대역폭 확장 정보를 이용하여 조정된다.
- [0121] 도 8 및 도 9에 도시된 바와 같이, 상기 설명된 블록들은 바람직하게는 모드를 제어하는 입력을 구비할 수 있다. 이러한 모드를 제어하는 입력은 상기 결정 단계(300)의 출력 신호로부터 추론된다. 바람직하게는, 상응하는 블록의 특성은, 바람직하게는 오시오 신호의 어떤 시간 부분에 대하여 발화에 대한 결정이 만들어질지 또는 음악에 대한 결정이 만들어질지 여부에 대한 상기 결정 단계의 출력에 적용될 수 있다. 바람직하게는 상기 모드 제어는 이러한 블록의 일 이상의 블록에 관련될 뿐, 상기 블록의 모든 기능에 관련되지는 않는다. 예를 들어 상기 결정은 오직 패처(701a)에만 영향을 미치고 도 9의 다른 블록들에는 어떠한 영향도 미치지 않는다. 또는 예

를 들어 오직 상기 조인트 멀티채널 매개 변수 분석기(도 8의 101b)에만 영향을 미치고 도 8에 도시된 다른 블록에 대해서는 영향을 미치지 않는다. 이러한 구현으로 바람직하게, 상기 일반적인 전처리 스테이지에 유연성을 제공함으로써 더 우수한 유연성, 고품질 및 저비트레이트의 출력 신호를 얻을 수 있게 된다. 그러나 한편으로는 양 종류의 신호에 대한 일반적인 전처리 스테이지에서의 알고리즘의 사용은 효과적인 인코딩/디코딩 기법을 구현할 수 있게 한다.

[0122] 도 10a 및 도 10b은, 상기 결정 단계(300)의 두 개의 다른 구현예를 도시한다. 도 10a에서는 열린 루프 결정이 제시되어 있다. 여기서 결정 스테이지에서의 상기 신호 분석기(300a)는, 입력 신호의 시간 부분 또는 주파수 부분이 상기 제1 인코딩된 브랜치(400) 또는 상기 제2 인코딩된 브랜치(500)에 의해서 신호의 일부가 인코딩될 것을 요청하는지에 대한 특성을 구비하였는지 여부에 대한 결정을 내리기 위하여 특정한 규칙을 구비하고 있다. 마지막으로, 상기 신호 분석기(300a)는 상기 전처리 스테이지로 입력되는 오디오 신호를 분석하거나 또는 상기 전처리 단계에 의해 출력되는 오디오 신호, 즉 오디오 중간 신호를 분석하거나 또는 상기 전처리 스테이지 내에서, 모노 신호 또는 도 8에 나타난 k 채널을 구비한 신호일 수 있는 다운 믹스 신호의 출력 같은 중간 신호를 분석할 수 있다. 출력의 면에서, 상기 신호 분석기(300a)는 인코더의 면에서 스위치(200) 및 디코딩 면에서의 이에 상응하는 스위치(600) 및 결합자(600)를 제어하기 위한 전환 결정을 생성할 수 있다.

[0123] 상기 제2 스위치(521)에 대해서는 상세히 설명하지 않았으나, 상기 제2 스위치(521)는 도 4a 및 도 4b를 참조하여 설명한 바와 같이 상기 제1 스위치(200)와 유사한 방법으로 위치될 수 있음이 강조되어야 한다. 그러므로 도 3c 내에서의 스위치(521)의 변형된 위치는 양 처리 브랜치 522, 523, 524의 출력 지점이다. 그럼으로써 양 처리 브랜치는 평행하게 작동하고, 하나의 처리 브랜치의 출력 신호만이 도 3c에 도시된 비트스트림 포머(former)를 통해서 비트스트림에 쓰여질 수 있다.

[0124] 뿐만 아니라, 상기 제2 결합자(600)은, 도 4c를 통해 상술한 바와 같이 특유의 크로스페이드 기능을 구비할 수 있다. 그렇지 않으면 또는 추가적으로, 상기 제1 결합자(532)는 동일한 크로스페이드 기능을 구비할 수 있다. 뿐만 아니라 양 결합자는 동일한 크로스페이드 기능을 구비할 수 있고, 다른 크로스페이드 기능을 구비할 수도 있으며, 양 결합자가 어떠한 추가적인 크로스페이드 기능 없이 전환되도록 어떠한 크로스페이드 기능도 구비하지 않을 수도 있다.

[0125] 기 상술한 바와 같이 양 스위치는, 도 10a 및 도 10b를 통해 설명된 열린 루프 결정 또는 닫힌 루프 결정에 의해서 제어될 수 있다. 여기서 도 3c의 제어기 300, 500은 양 스위치에 대해 차이가 있거나 또는 동일한 기능을 구비할 수 있다.

[0126] 그뿐만 아니라, 신호 적용에 대한 시간 워핑 기능은, 제1 인코딩 브랜치 또는 제1 디코딩 브랜치 뿐만 아니라, 디코더의 측면뿐만 아니라 인코더의 측면에서도, 제2 처리 코딩 브랜치 내의 제2 처리 브랜치에도 존재할 수 있다. 처리 신호에 의존하여, 양 시간 워핑 기능은, 제1 도메인 또는 제2 도메인 내에서 동일한 시간 워프가 신호에 적용될 수 있도록 하기 위해서 동일한 시간 워핑 정보를 구비할 수 있다. 이는 처리 로드를 저장하며, 차후의 블록이 유사한 시간 워핑 시간 특징을 구비한 경우 같은 일례에서는 유용하다. 그러나 다른 실시예에 있어서, 바람직하게는 제1 코딩 브랜치 및 상기 제2 코딩 브랜치 내의 상기 제2 처리 브랜치에 대한, 독립적인 시간 워핑 평가자(estimator)를 구비할 수 있다.

[0127] 상기 본 발명에 의해 인코딩된 오디오 신호는 디지털 저장 매체에 저장될 수 있고, 무선 전송 매체 또는 인터넷 같은 유선 전송 매체 등의 전송 매체를 통하여 전송될 수 있다.

[0128] 다른 실시예에 있어서 도 1a 또는 2a의 스위치(200)는 두 코딩 브랜치 400 및 500 사이의 전환을 수행한다. 다른 실시예에 있어서, 제3 인코딩 브랜치 또는 제4 인코딩 브랜치 또는 그 이상의 인코딩 브랜치 등의 추가적인

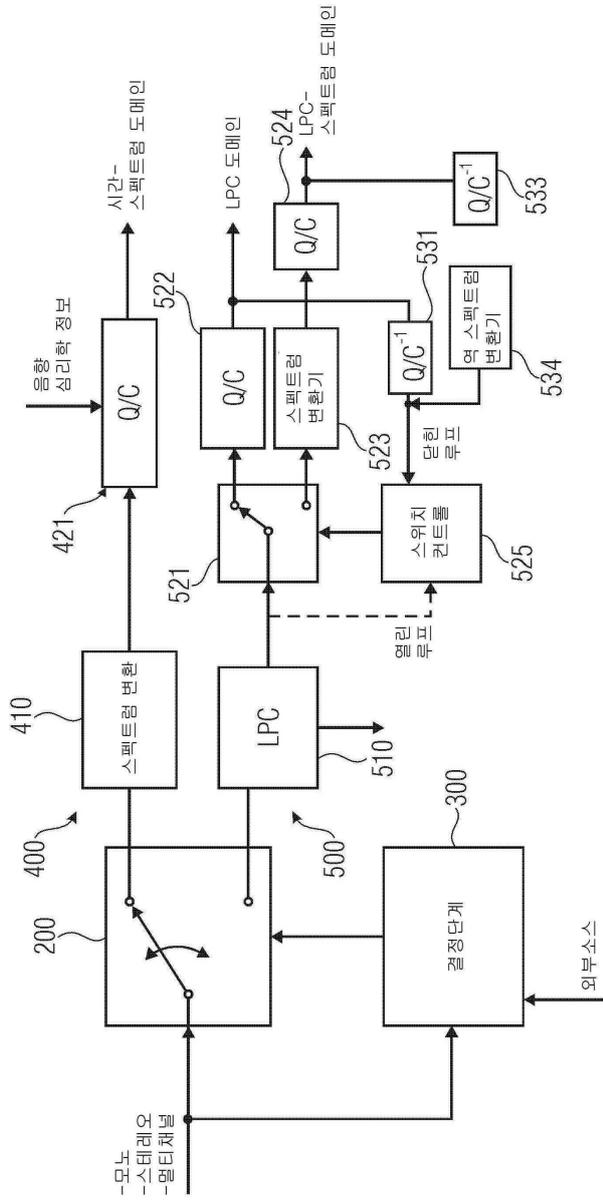
인코딩 브랜치가 있을 수 있다. 디코더의 측면에서, 도 1b 및 2b의 스위치(600)는 상기 두 개의 디코딩 브랜치 431, 440 및 531, 532, 533, 534, 540 간의 전환을 수행한다. 다른 실시예에서는, 제3 디코딩 브랜치, 제4 디코딩 브랜치 또는 그 이상의 디코딩 브랜치 같은 추가적인 디코딩 브랜치가 있을 수 있다. 유사하게, 상기 추가적인 코딩/디코딩 브랜치가 제공될 때, 다른 스위치 521 또는 532도 둘 이상의 다른 코딩 알고리즘을 전환할 수 있다.

[0129] 기상술한 실시예들은 단지 본 발명의 개념을 도시하기 위한 것이다. 여기서 설명된 상기 일례 및 상세한 내용에 대한 수정 및 변형은 당업자에게는 자명하다고 이해된다. 그러므로 본 발명의 범위는 오직 첨부된 특허청구범위에 의해서 판단되어야지 여기서 설명되고 묘사된 실시예의 특수한 세부내용에 의해서 제한되어서는 안 된다.

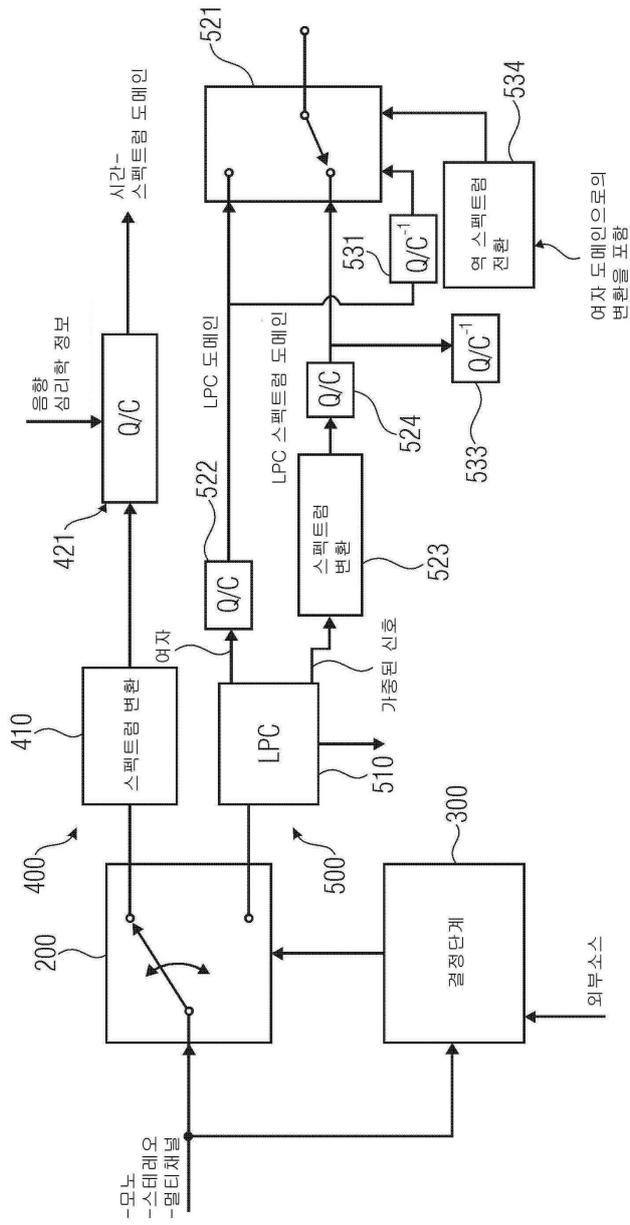
[0130] 상기 본 발명의 방법에 대한 구현에 필요한 사항에 따라, 본 방법은 하드웨어 또는 소프트웨어에서도 구현될 수 있다. 상기 구현은 디지털 저장 매체, 특히 전기적으로 읽힐 수 있는 제어 신호가 저장되고, 본 방법이 수행될 수 있는 프로그래밍이 가능한 컴퓨터 시스템과 같이 작동하는 디스크, DVD 또는 CD 등을 사용하여 수행될 수 있다. 그러므로 일반적으로 본 발명은 기계 장치에 의해 읽힐 수 있는 매체에 의해 저장될 수 있는 프로그램을 포함하는 컴퓨터 프로그램 제품이다. 여기서 상기 프로그램 코드는, 상기 컴퓨터 프로그램이 컴퓨터 내에서 실행될 때, 본 발명을 수행하기 위하여 구동된다. 그러므로 다시 말하면 본 방법의 방법은 컴퓨터 프로그램이 컴퓨터 내에서 구동될 때, 상기 본 방법의 발명 중의 적어도 하나 이상이 수행되는 프로그램 코드를 구비한 컴퓨터 프로그램이다.

도면

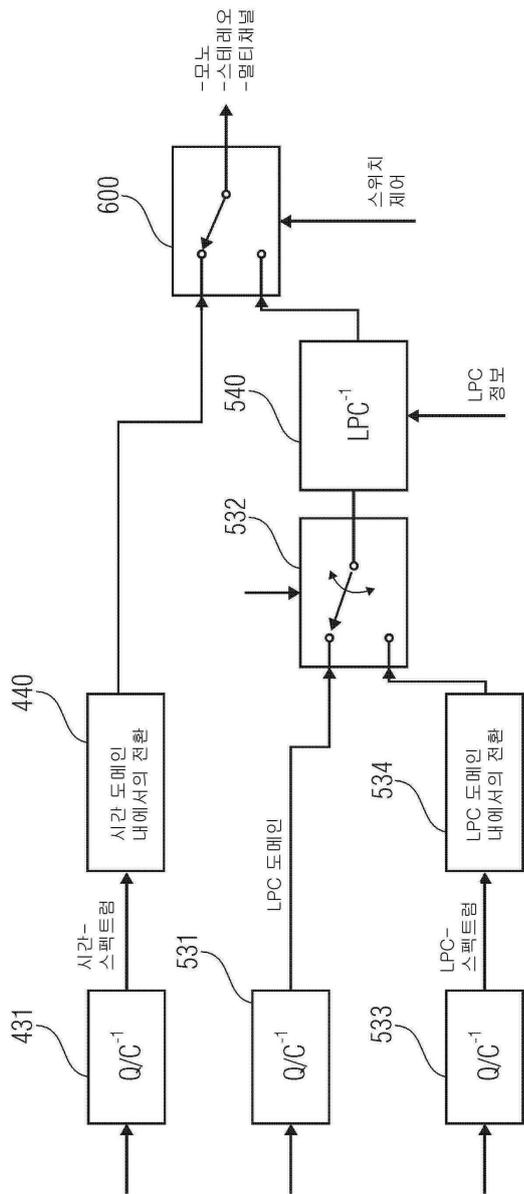
도면1a



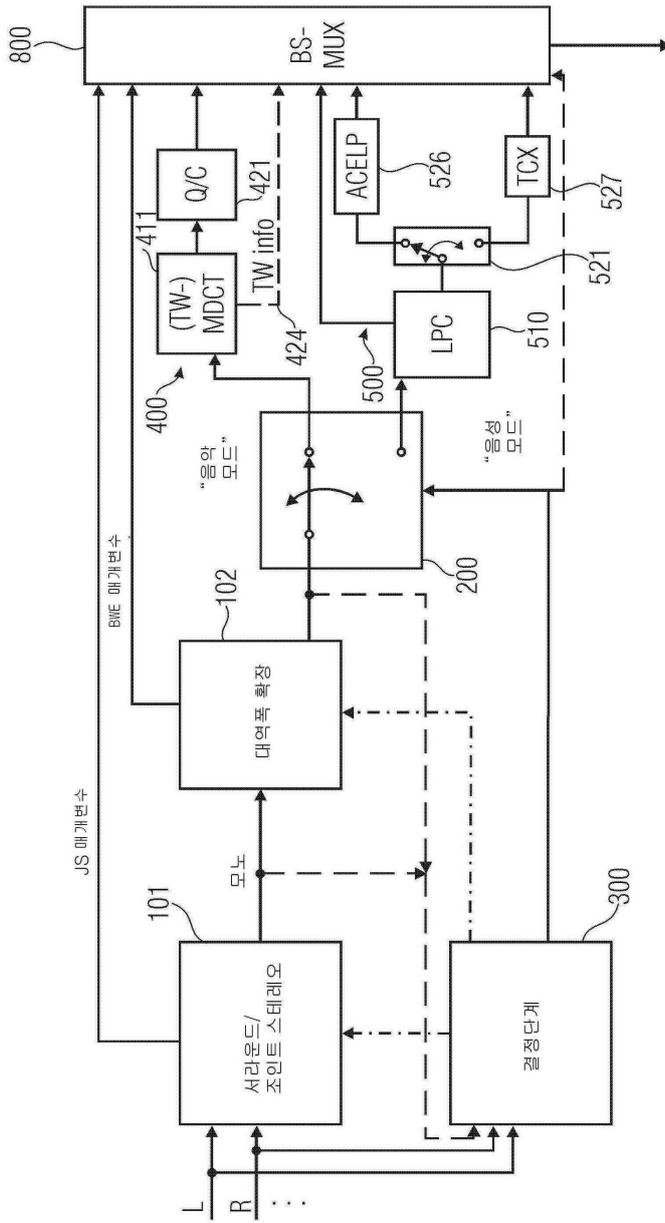
도면1b



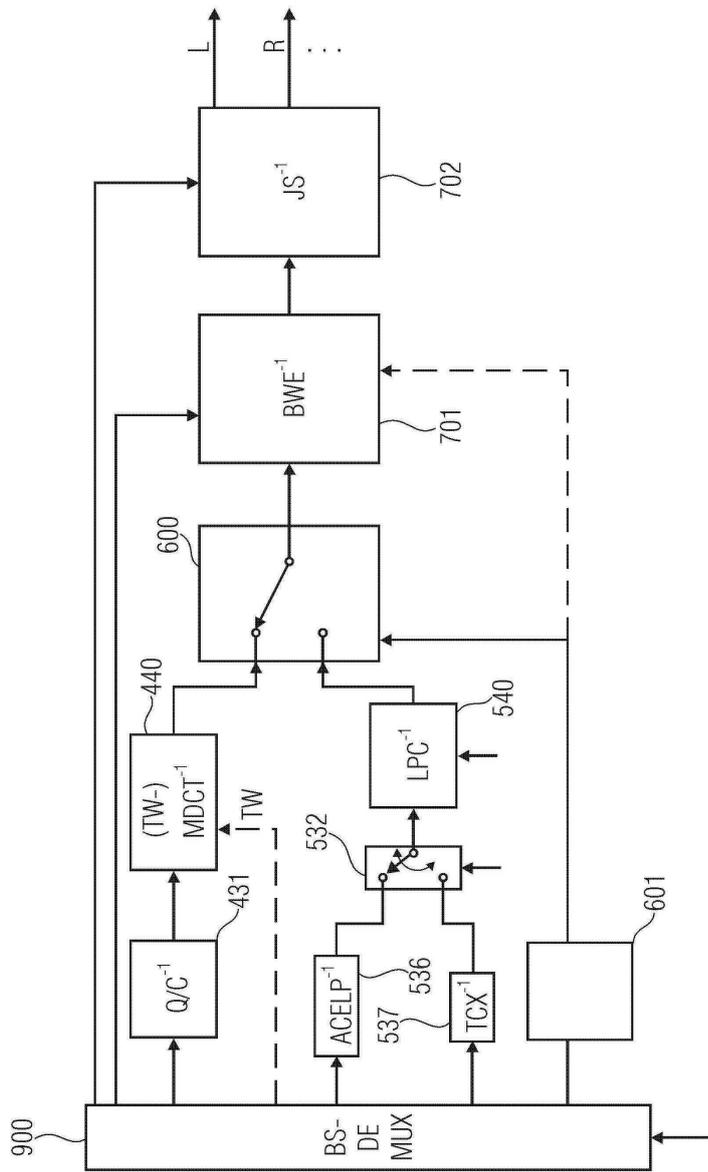
도면1c



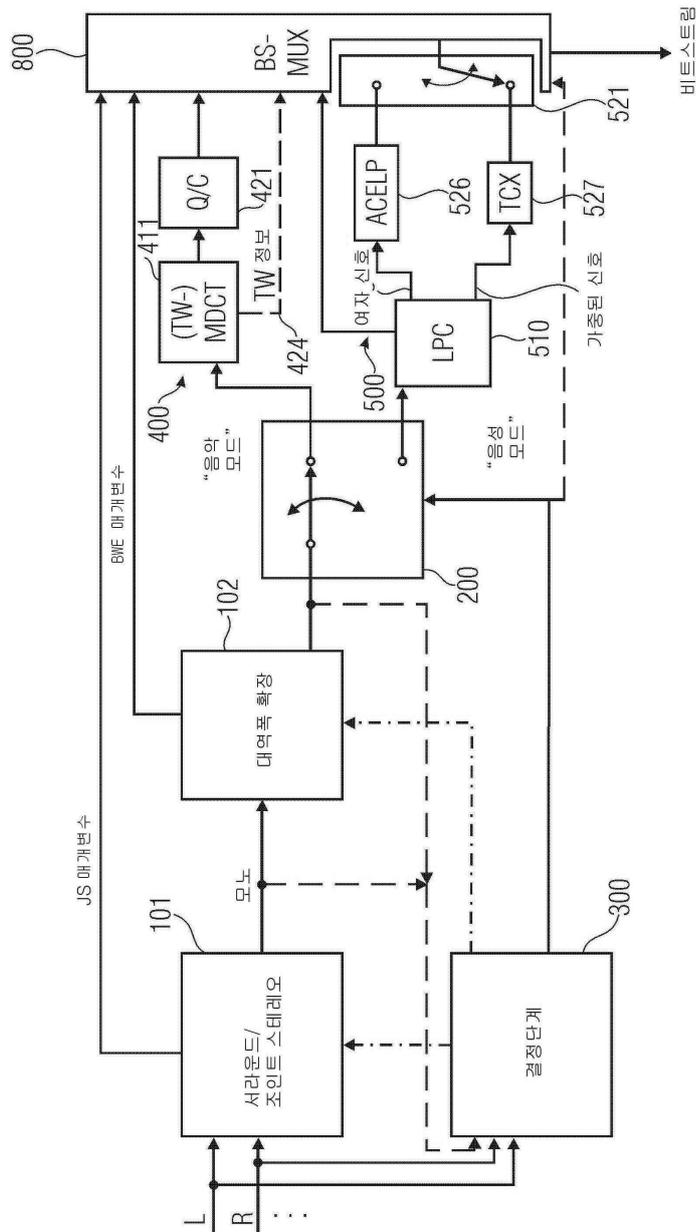
도면2a



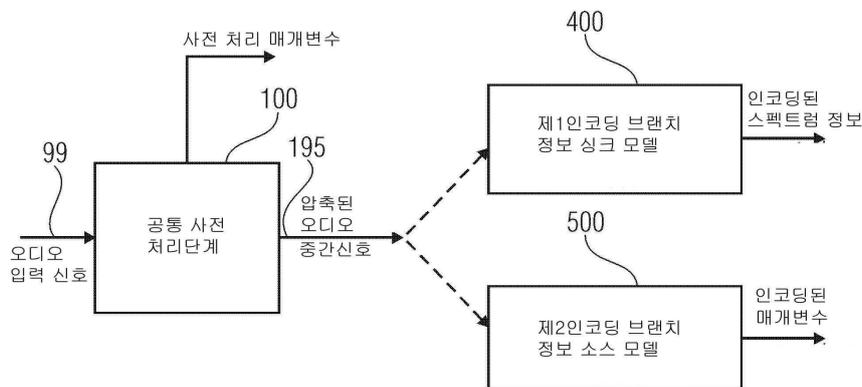
도면2b



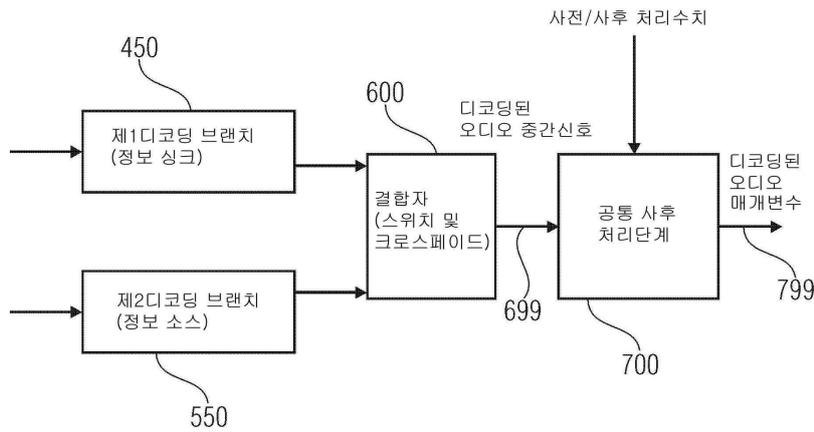
도면2c



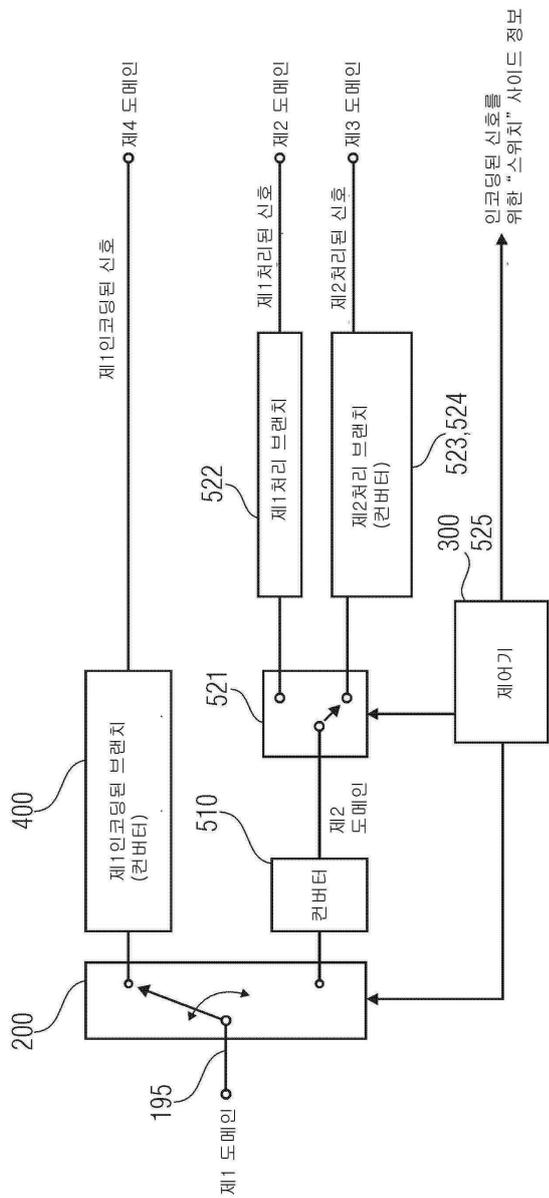
도면3a



도면3b

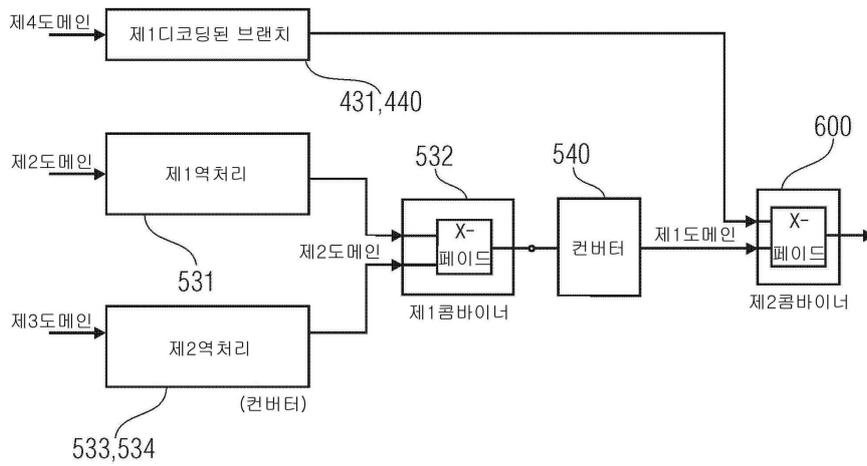


도면3c

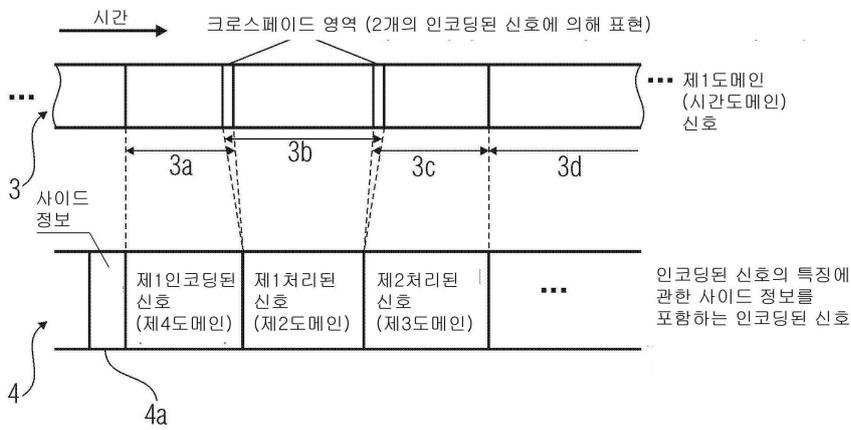


- 제1도메인 오디오 신호의 각각의 블록은 선택적인 크로스 오버 영역과는 별도로 제2도메인, 제3도메인, 제4도메인 인코딩된 신호에 의해서 나타난다.

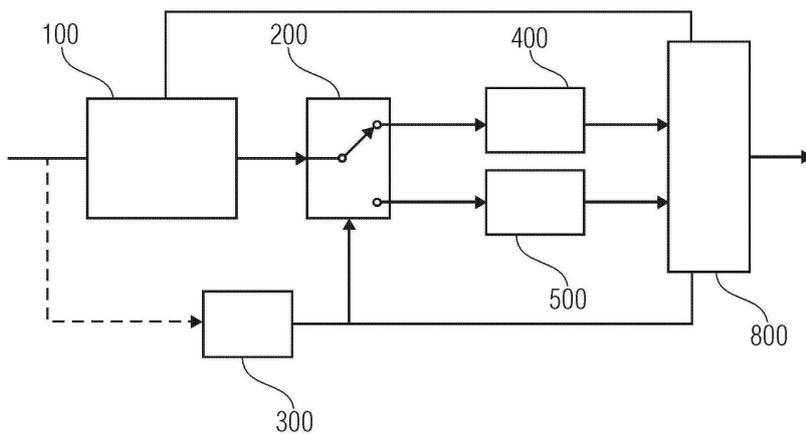
도면3d



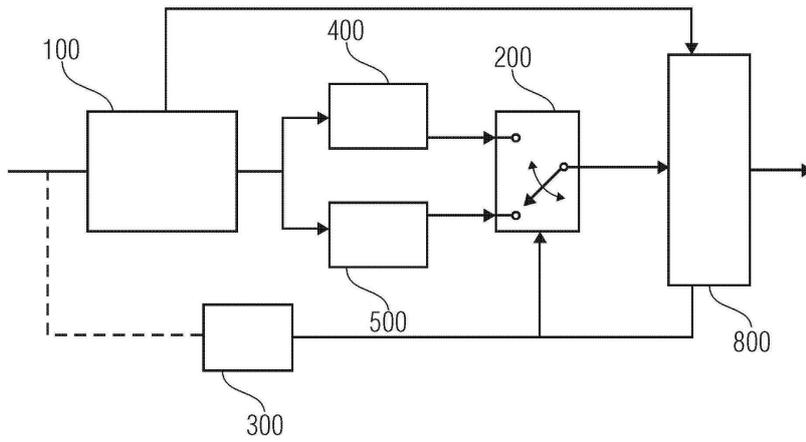
도면3e



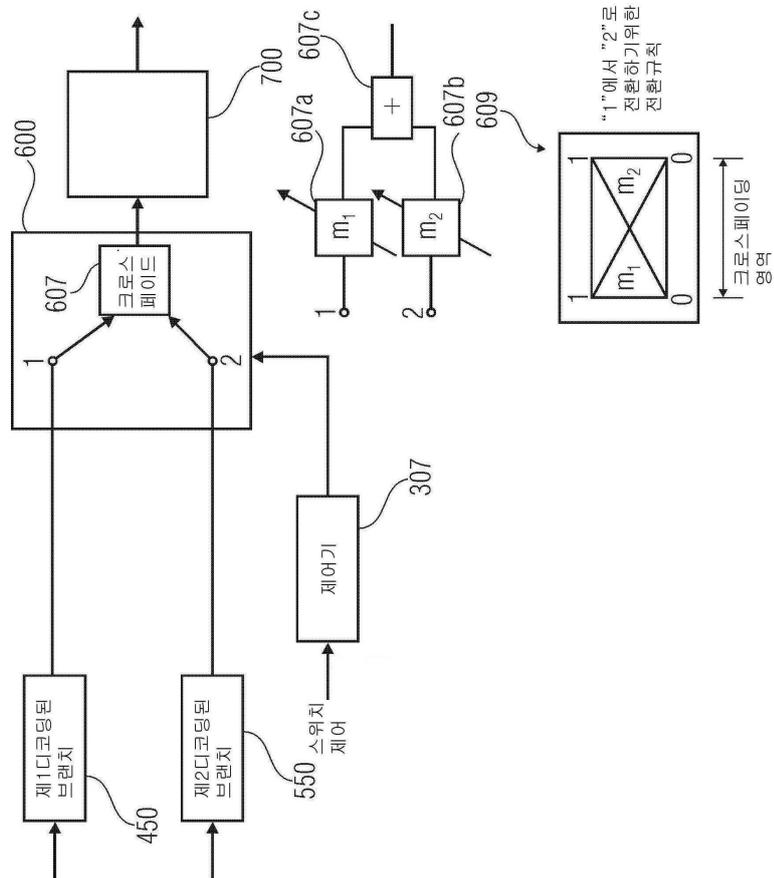
도면4a



도면4b

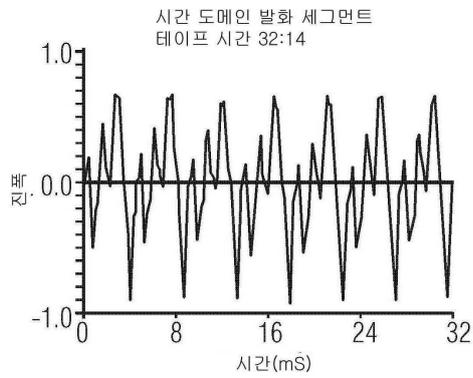


도면4c

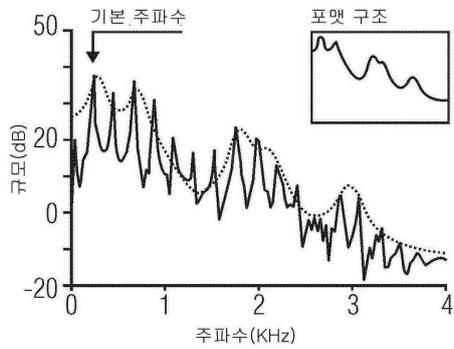


도면5a

맥박형 신호세그먼트 (예를 들어 음성 발화)

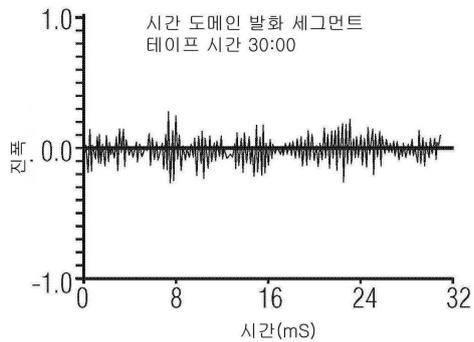


도면5b

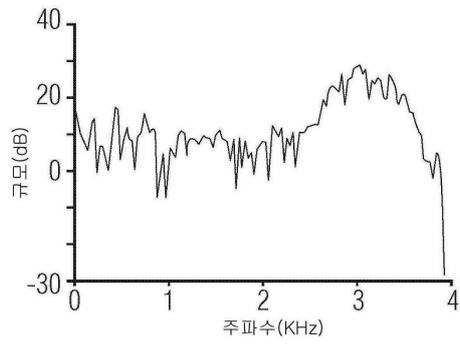


도면5c

정지 세그먼트 (예를 들어 무성 발화)

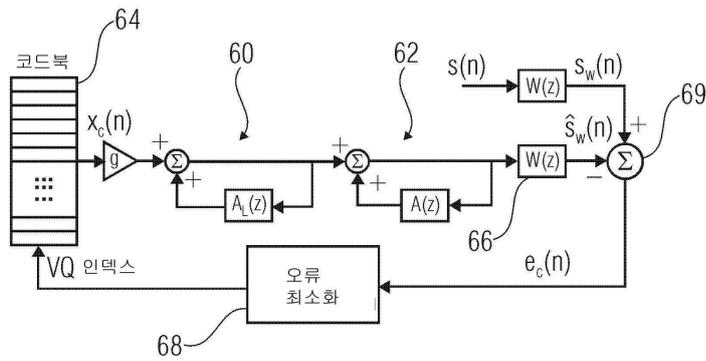


도면5d



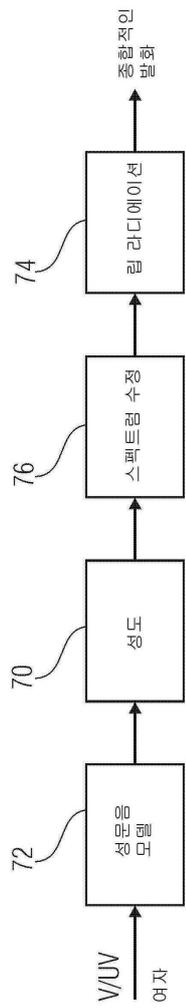
도면6

analysis-by-synthesis CELP

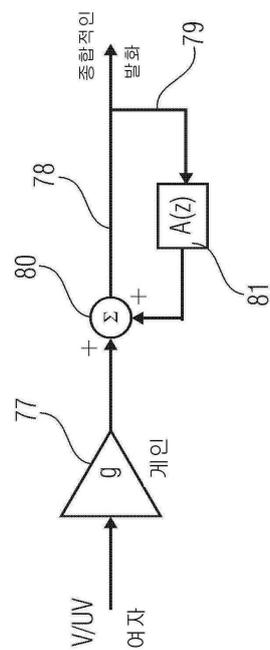


- $A_L(z)$: 장기 예측
≒ 피치(파인) 구조
- $A(z)$: 단기 예측
≒ 포맷구조/ 스펙트럼 인벨롭

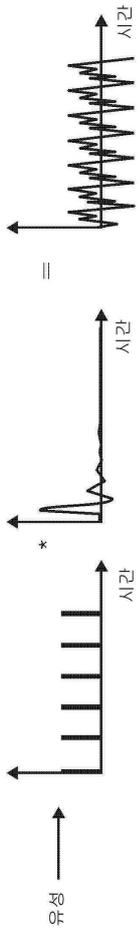
도면7a



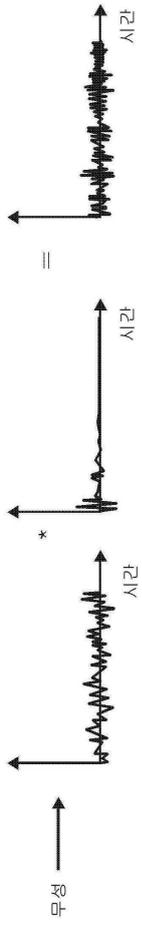
도면7b



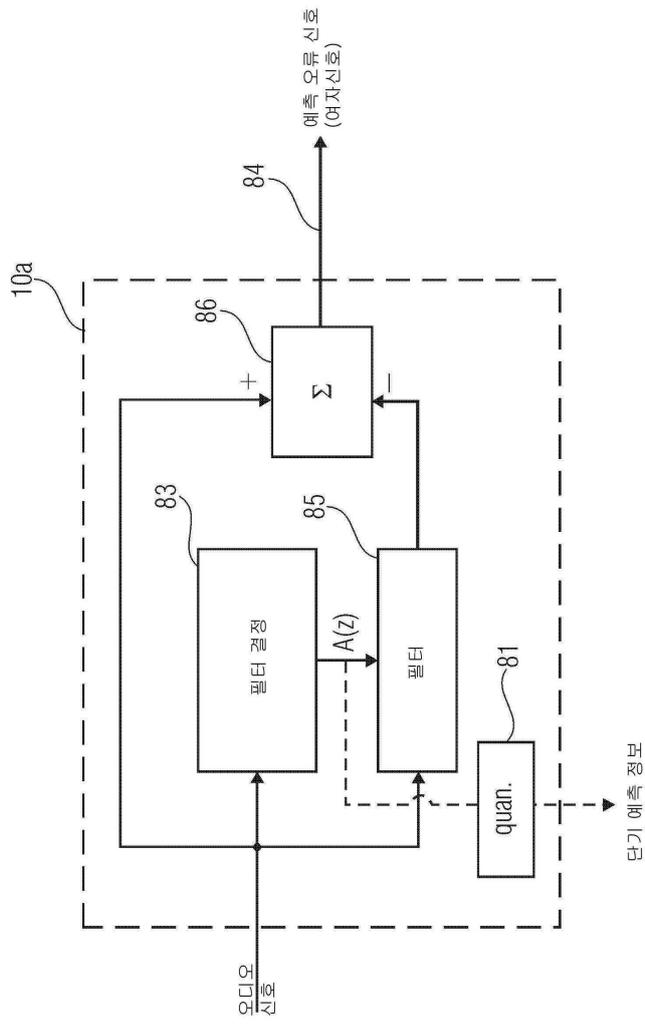
도면7c



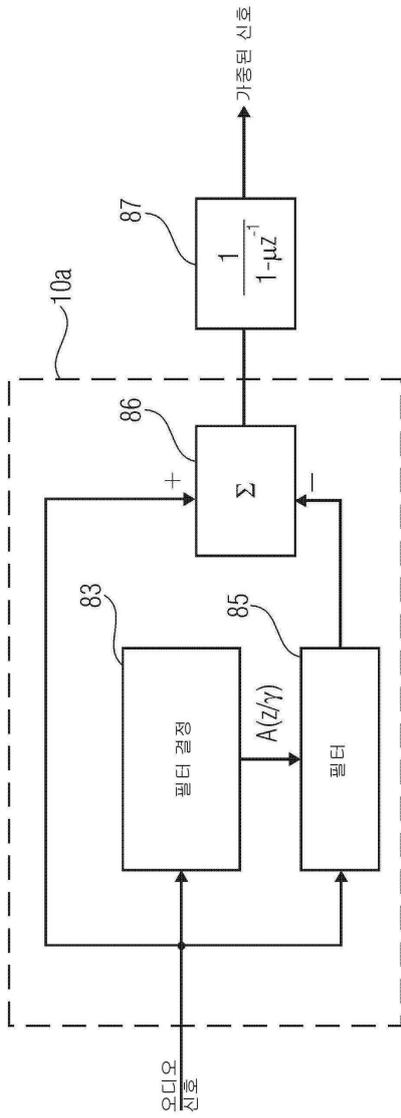
도면7d



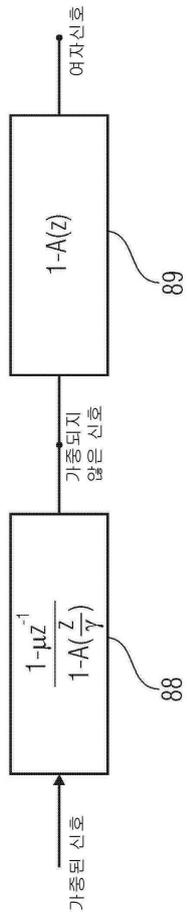
도면7e



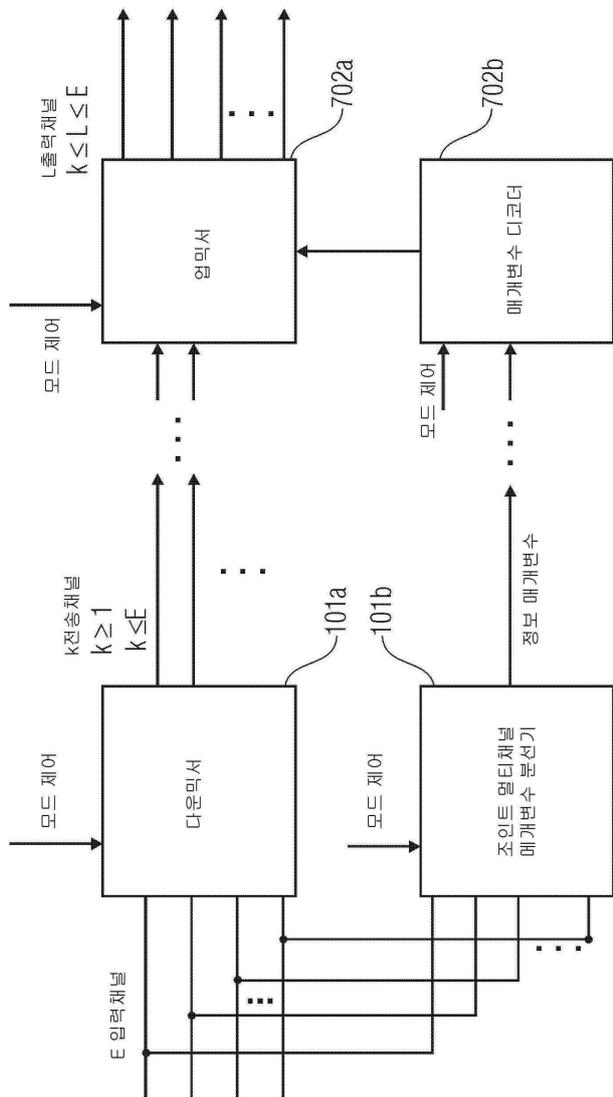
도면7f



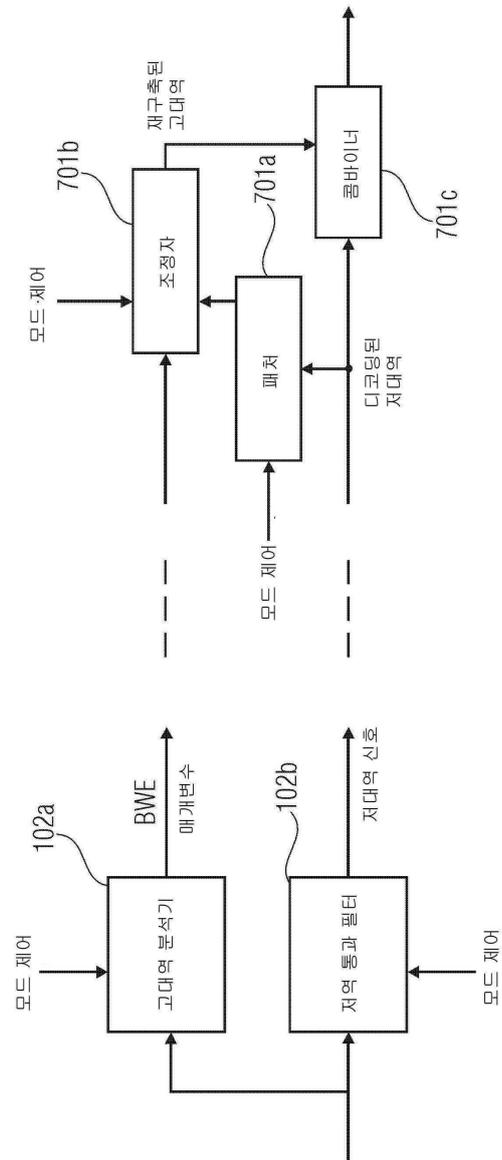
도면7g



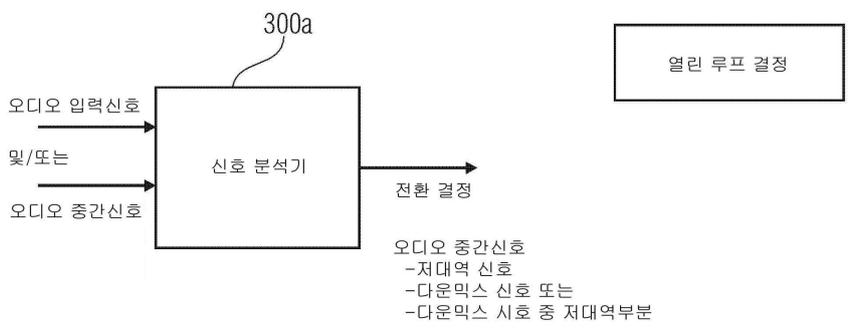
도면8



도면9



도면10a



도면10b

