



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2012년12월10일
(11) 등록번호 10-1210797
(24) 등록일자 2012년12월04일

(51) 국제특허분류(Int. Cl.)
G10L 19/00 (2006.01) H04R 5/00 (2006.01)
(21) 출원번호 10-2007-7011845
(22) 출원일자(국제) 2005년10월28일
심사청구일자 2010년10월27일
(85) 번역문제출일자 2007년05월25일
(65) 공개번호 10-2007-0084552
(43) 공개일자 2007년08월24일
(86) 국제출원번호 PCT/US2005/038961
(87) 국제공개번호 WO 2006/050112
국제공개일자 2006년05월11일
(30) 우선권주장
10/975,841 2004년10월28일 미국(US)
60/622,922 2004년10월28일 미국(US)
(56) 선행기술조사문헌
JP평성09505193 A
US20020071574 A1
전체 청구항 수 : 총 14 항

(73) 특허권자
디티에스 워싱턴, 엘엘씨
미국 91301 캘리포니아주 아고라 힐스 클라레튼
드라이브 5171
(72) 발명자
리암즈, 로버트, 더블유.
미국 98012 워싱턴주 밀 크릭 21번 드라이브 사우
쓰이스트 14415
투스, 제프리, 케이.
미국 98011 워싱턴주 보텔 92번 애비뉴 노쓰이스트
19511
워너, 아론
미국 98102 워싱턴주 시애틀 이스트 해리슨 스트
리트 넘버 102316
(74) 대리인
백만기, 양영준

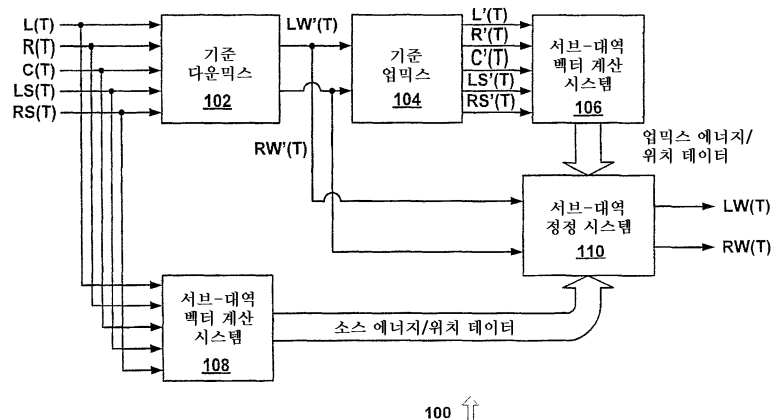
심사관 : 이정숙

(54) 발명의 명칭 오디오 공간 환경 엔진

(57) 요약

서로 다른 포맷의 오디오 데이터 간에 변환을 하는 오디오 공간 환경 엔진이 제공된다. 오디오 공간 환경 엔진은 N-채널 데이터와 M-채널 데이터 간의 유연한 변환 및 M-채널 데이터로부터 다시 N'-채널 데이터로의 변환(단, N, M 및 N'은 정수이고, N이 N'과 꼭 같을 필요는 없음)을 가능하게 해준다. 예를 들어, 이러한 시스템은 스테레오 사운드 데이터용으로 설계된 네트워크 또는 기반구조를 통해 서라운드 사운드 데이터를 전송 또는 저장하는데 사용될 수 있다. 오디오 공간 환경 엔진은 개량된 동적 다운-믹싱 유닛 및 고분해능 주파수 대역 업-믹싱 유닛으로 인해 서로 다른 공간 환경 간의 개선되고 유연한 변환을 제공한다. 동적 다운-믹싱 유닛은 많은 다운-믹싱 방법에 공통된 스펙트럼, 시간 및 공간 부정확성을 정정할 수 있는 지능적인 분석 및 정정 루프를 포함한다. 업-믹싱 유닛은 서로 다른 주파수 성분의 공간 배치를 도출하기 위해 고분해능 주파수 대역들에 걸쳐 중요한 채널간 공간 큐(inter-channel spatial cue)의 추출 및 분석을 이용한다. 다운-믹싱 유닛 및 업-믹싱 유닛은, 개별적으로 또는 시스템으로서 사용될 때, 개선된 사운드 품질 및 공간 구분을 제공한다.

대표도 - 도1



특허청구의 범위

청구항 1

N 채널 오디오 시스템으로부터 M 채널 오디오 시스템으로(단, M 및 N은 정수이고, N은 M보다 큼) 변환하는 오디오 공간 환경 엔진으로서,

N개 채널의 오디오 데이터를 수신하고 상기 N개 채널의 오디오 데이터를 M개 채널의 오디오 데이터로 변환하는 기준 다운-믹서,

상기 M개 채널의 오디오 데이터를 수신하고 상기 M개 채널의 오디오 데이터를 N'개 채널의 오디오 데이터로 변환하는 기준 업-믹서, 및

상기 M개 채널의 오디오 데이터, 상기 N개 채널의 오디오 데이터, 및 상기 N'개 채널의 오디오 데이터를 수신하고 상기 N개 채널의 오디오 데이터와 상기 N'개 채널의 오디오 데이터 간의 차이에 기초하여 상기 M개 채널의 오디오 데이터를 정정하는 정정 시스템을 포함하는 오디오 공간 환경 엔진.

청구항 2

제1항에 있어서, 상기 정정 시스템은,

상기 N개 채널의 오디오 데이터를 수신하고 제1 복수의 서브-대역의 오디오 공간 이미지 데이터를 발생하는 제1 서브-대역 벡터 교정 유닛,

상기 N'개 채널의 오디오 데이터를 수신하고 제2 복수의 서브-대역의 오디오 공간 이미지 데이터를 발생하는 제2 서브-대역 벡터 교정 유닛

을 더 포함하고,

상기 정정 시스템은, 상기 제1 복수의 서브-대역의 오디오 공간 이미지 데이터 및 상기 제2 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 상기 제1 복수의 서브-대역의 오디오 공간 이미지 데이터와 상기 제2 복수의 서브-대역의 오디오 공간 이미지 데이터 간의 차이에 기초하여 상기 M개 채널의 오디오 데이터를 정정하는 오디오 공간 환경 엔진.

청구항 3

제2항에 있어서, 상기 제1 복수의 서브-대역의 오디오 공간 이미지 데이터 및 상기 제2 복수의 서브-대역의 오디오 공간 이미지 데이터 각각은 연관된 에너지 값 및 위치 값을 갖는 것인 오디오 공간 환경 엔진.

청구항 4

제3항에 있어서, 상기 위치 값 각각은 2차원 공간에서의 연관된 서브-대역의 오디오 공간 이미지 데이터의 중심의 겉보기 위치(apparent location)를 나타내고,

상기 중심의 좌표는 N개의 스피커 각각과 연관된 에너지 값과 상기 N개의 스피커 각각의 좌표의 벡터 합에 의해 결정되는 것인 오디오 공간 환경 엔진.

청구항 5

제1항에 있어서, 상기 기준 다운-믹서는 복수의 분수 힐버트 스테이지(fractional Hilbert stage)를 더 포함하며,

각각의 스테이지는 상기 N개 채널의 오디오 데이터 중 하나를 수신하고 미리 정해진 위상 천이를 연관된 채널의 오디오 데이터에 적용하는 것인 오디오 공간 환경 엔진.

청구항 6

제5항에 있어서, 상기 기준 다운-믹서는, 상기 M개 채널의 오디오 데이터를 발생하기 위해 미리 정해진 방식으로 상기 힐버트 스테이지로부터의 출력을 결합시키는, 복수의 분수 힐버트 스테이지에 연결된 복수의 합산 스테이지를 더 포함하는 것인 오디오 공간 환경 엔진.

청구항 7

제1항에 있어서, 상기 기준 업-믹서는,

상기 M개 채널의 오디오 데이터를 수신하고 복수의 서브-대역의 오디오 공간 이미지 데이터를 발생하는 시간 영역-주파수 영역 변환 스테이지,

상기 M개 채널의 상기 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 N'개 채널의 복수의 서브-대역의 오디오 공간 이미지 데이터를 발생하는 필터 발생기,

상기 N'개 채널의 상기 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 각각의 서브-대역의 하나 이상의 인접한 서브-대역과의 평균을 구하는 평활화 스테이지,

상기 M개 채널의 상기 복수의 서브-대역의 오디오 공간 이미지 데이터 및 평활화된 N'개 채널의 상기 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 스케일링된 N'개 채널의 상기 복수의 서브-대역의 오디오 공간 이미지 데이터를 발생하는, 상기 평활화 스테이지에 연결되어 있는 합산 스테이지, 및

상기 스케일링된 N'개 채널의 상기 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 상기 N'개 채널의 오디오 데이터를 발생하는 주파수 영역-시간 영역 변환 스테이지를 더 포함하는 것인 오디오 공간 환경 엔진.

청구항 8

제1항에 있어서, 상기 정정 시스템은 제1 서브-대역 벡터 교정 스테이지를 더 포함하고,

상기 제1 서브-대역 벡터 교정 스테이지는,

상기 N개 채널의 오디오 데이터를 수신하고 제1 복수의 서브-대역의 오디오 공간 이미지 데이터를 발생하는 시간 영역-주파수 영역 변환 스테이지,

상기 제1 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 각각의 서브-대역에 대한 제1 에너지 값을 발생하는 제1 서브-대역 에너지 스테이지, 및

상기 제1 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 각각의 서브-대역에 대한 제1 위치 벡터를 발생하는 제1 서브-대역 위치 스테이지를 포함하는 것인 오디오 공간 환경 엔진.

청구항 9

제8항에 있어서, 상기 정정 시스템은 제2 서브-대역 벡터 교정 스테이지를 더 포함하고,

상기 제2 서브-대역 벡터 교정 스테이지는,

제2 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 각각의 서브-대역에 대한 제2 에너지 값을 발생하는 제2 서브-대역 에너지 스테이지, 및

상기 제2 복수의 서브-대역의 오디오 공간 이미지 데이터를 수신하고 각각의 서브-대역에 대한 제2 위치 벡터를 발생하는 제2 서브-대역 위치 스테이지를 포함하는 것인 오디오 공간 환경 엔진.

청구항 10

N 채널 오디오 시스템으로부터 M 채널 오디오 시스템으로(단, N 및 M은 정수이고, N은 M보다 큼) 변환하는 방법으로서,

N개 채널의 오디오 데이터를 M개 채널의 오디오 데이터로 변환하는 단계,

상기 M개 채널의 오디오 데이터를 N'개 채널의 오디오 데이터로 변환하는 단계, 및

상기 N개 채널의 오디오 데이터와 상기 N'개 채널의 오디오 데이터 간의 차이에 기초하여 상기 M개 채널의 오디오 데이터를 정정하는 단계를 포함하는 변환 방법.

청구항 11

제10항에 있어서, 상기 N개 채널의 오디오 데이터를 상기 M개 채널의 오디오 데이터로 변환하는 상기 단계는,

미리 정해진 위상 천이를 연관된 채널의 오디오 데이터에 적용하기 위해 상기 N개 채널의 오디오 데이터 중 하나 이상을 분수 힐버트 함수로 처리하는 단계, 및

상기 M개 채널의 오디오 데이터 각각에서의 상기 N개 채널의 오디오 데이터 중 하나 이상의 결합이 미리 정해진 위상 관계를 갖도록, 상기 분수 힐버트 함수로 처리 후에 상기 N개 채널의 오디오 데이터 중 하나 이상을 결합하여 상기 M개 채널의 오디오 데이터를 생성하는 단계를 포함하는 것인 변환 방법.

청구항 12

제10항에 있어서, 상기 M개 채널의 오디오 데이터를 상기 N'개 채널의 오디오 데이터로 변환하는 상기 단계는,

상기 M개 채널의 오디오 데이터를 시간 영역으로부터 주파수 영역의 복수의 서브-대역으로 변환하는 단계,

상기 M개 채널의 상기 복수의 서브-대역을 필터링하여 N개 채널의 복수의 서브-대역을 발생하는 단계,

각각의 서브-대역의 하나 이상의 인접한 대역과의 평균을 구함으로써 상기 N개 채널의 복수의 서브-대역을 평활화하는 단계,

상기 N개 채널의 상기 복수의 서브-대역 각각을 상기 M개 채널의 대응하는 서브-대역 중 하나 이상과 곱하는 단계, 및

상기 N개 채널의 상기 복수의 서브-대역을 상기 주파수 영역으로부터 상기 시간 영역으로 변환하는 단계를 포함하는 것인 변환 방법.

청구항 13

제10항에 있어서, 상기 N개 채널의 오디오 데이터와 상기 N'개 채널의 오디오 데이터 간의 차이에 기초하여 상기 M개 채널의 오디오 데이터를 정정하는 상기 단계는,

상기 N개 채널의 오디오 데이터의 복수의 서브-대역 각각에 대한 에너지 및 위치 벡터를 결정하는 단계,

상기 N'개 채널의 오디오 데이터의 복수의 서브-대역 각각에 대한 에너지 및 위치 벡터를 결정하는 단계, 및

상기 N개 채널의 오디오 데이터 및 상기 N'개 채널의 오디오 데이터의 대응하는 서브-대역들에 대한 에너지 및 위치 벡터의 차이가 허용 공차보다 큰 경우 상기 M개 채널의 오디오 데이터의 하나 이상의 서브-대역을 정정하는 단계를 포함하는 것인 변환 방법.

청구항 14

제13항에 있어서, 상기 M개 채널의 오디오 데이터의 하나 이상의 서브-대역을 정정하는 상기 단계는, 상기 M개 채널의 오디오 데이터의 조정된 서브-대역이 상기 N'개 채널의 오디오 데이터의 복수의 서브-대역 각각에 대한 미조정된 에너지 및 위치 벡터보다 상기 N개 채널의 오디오 데이터의 상기 서브-대역의 에너지 및 위치 벡터에 더 가까운 하나 이상의 서브-대역 에너지 및 위치 벡터를 갖는 조정된 N'개 채널의 오디오 데이터로 변환되도록, 상기 M개 채널의 오디오 데이터의 상기 서브-대역에 대한 에너지 및 위치 벡터를 조정하는 단계를 포함하는 것인 변환 방법.

청구항 15

삭제

청구항 16

삭제

청구항 17

삭제

청구항 18

삭제

청구항 19

삭제

청구항 20

삭제

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

명세서

기술분야

[0001] 관련 출원

[0002] 본 출원은 2004년 10월 28일자로 출원된 발명의 명칭이 "2-N 렌더링(2-to-N Rendering)"인 미국 가특허 출원 제 60/622,922호, 2004년 10월 28일자로 출원된 발명의 명칭이 "오디오 공간 환경 엔진(Audio Spatial Environment Engine)"인 미국 특허 출원 제10/975,841호, 본 출원과 함께 출원된 발명의 명칭이 "오디오 공간 환경 다운-믹서(Audio Spatial Environment Down-Mixer)"인 미국 특허 출원 제_____호(대리인 문서 번호 13646.0014), 및 본 출원과 함께 출원된 발명의 명칭이 "오디오 공간 환경 업-믹서(Audio Spatial Environment Up-Mixer)"인 미국 특허 출원 제_____호(대리인 문서 번호 13646.0012)를 우선권 주장하며, 이들 각각은 공동 소유되어 있고 또 여기에 인용함으로써 그 전체 내용이 본 명세서에 포함된다.

[0003] 본 발명은 오디오 데이터 처리 분야에 관한 것으로서, 보다 상세하게는 서로 다른 포맷의 오디오 데이터 간에 변환을 하는 시스템 및 방법에 관한 것이다.

배경 기술

[0004] 오디오 데이터를 처리하는 시스템 및 방법이 기술 분야에 공지되어 있다. 이들 시스템 및 방법의 대부분은, 2-채널 스테레오 환경, 4-채널 쿼드라포닉(quadrasonic) 환경, 5 채널 서라운드 사운드 환경(5.1 채널 환경이라고도 함) 또는 다른 적당한 포맷 또는 환경 등의, 기지의 오디오 환경에 대한 오디오 데이터를 처리하는 데 사용된다.

[0005] 포맷 또는 환경의 수의 증가로 야기되는 한가지 문제는 제1 환경에서 최적의 오디오 품질을 위해 처리된 오디오 데이터가 종종 다른 오디오 환경에서 즉시 사용될 수 없다는 것이다. 이 문제의 한 예는 스테레오 사운드 데이터용으로 설계된 네트워크 또는 기반구조를 통해 서라운드 사운드 데이터를 전송 또는 저장하는 것이다. 스테레오 2-채널 전송 또는 저장을 위한 기반구조가 서라운드 사운드 포맷에 대한 추가 채널의 오디오 데이터를 지원할 수 없기 때문에, 기존의 기반구조에서 서라운드 사운드 포맷을 전송 또는 이용하는 것이 어렵거나 불가능하다.

발명의 상세한 설명

- [0006] 본 발명에 따르면, 공간 오디오 환경들 간에 변환을 하는 것에서의 기지의 문제점을 극복하는 오디오 공간 환경 엔진에 대한 시스템 및 방법이 제공된다.
- [0007] 상세하게는, N-채널 데이터와 M-채널 데이터 간의 변환 및 M-채널 데이터로부터 다시 N'-채널 데이터로의 변환 (단, N, M 및 N'는 정수이고, N이 N'과 반드시 같을 필요는 없음)을 가능하게 해주는 오디오 공간 환경 엔진에 대한 시스템 및 방법이 제공된다.
- [0008] 본 발명의 예시적인 실시예에 따르면, N-채널 오디오 시스템으로부터 M-채널 오디오 시스템으로, 다시 N'-채널 오디오 시스템으로(단, N, M 및 N'은 정수이고, N이 N'과 꼭 같을 필요는 없음) 변환하는 오디오 공간 환경 엔진이 제공된다. 이 오디오 공간 환경 엔진은 N개 채널의 오디오 데이터를 수신하고 이 N개 채널의 오디오 데이터를 M개 채널의 오디오 데이터로 변환하는 동적 다운-믹서를 포함한다. 이 오디오 공간 환경 엔진은 또한 M개 채널의 오디오 데이터를 수신하고 이 M개 채널의 오디오 데이터를 N'개 채널의 오디오 데이터로 변환하는 업-믹서를 포함한다. 이 시스템의 한 예시적인 응용은 스테레오 사운드 데이터용으로 설계된 네트워크 또는 기반구조를 통해 서라운드 사운드 데이터를 전송 또는 저장하는 것이다. 동적 다운-믹싱 유닛은 전송 또는 저장을 위해 서라운드 사운드 데이터를 스테레오 사운드 데이터로 변환하고, 업-믹싱 유닛은 재생, 처리 또는 어떤 다른 적당한 사용을 위해 스테레오 사운드 데이터를 서라운드 사운드 데이터로 복원한다.
- [0009] 본 발명은 많은 중요한 기술적 이점을 제공한다. 본 발명의 한가지 중요한 기술적 이점은 개량된 동적 다운-믹싱 유닛 및 고분해능 주파수 대역 업-믹싱 유닛으로 인해 서로 다른 공간 환경 간의 개선되고 유연한 변환을 제공하는 시스템이다. 동적 다운-믹싱 유닛은 많은 다운-믹싱 방법에 공통된 스펙트럼, 시간 및 공간 부정확성을 정정하는 지능적인 분석 및 정정 루프를 포함한다. 업-믹싱 유닛은 서로 다른 주파수 성분의 공간 배치를 도출하기 위해 고분해능 주파수 대역들에 걸쳐 중요한 채널간 공간 큐(inter-channel spatial cue)의 추출 및 분석을 이용한다. 다운-믹싱 유닛 및 업-믹싱 유닛은, 개별적으로 또는 시스템으로서 사용될 때, 개선된 사운드 품질 및 공간 구분을 제공한다.
- [0010] 당업자라면, 도면과 관련하여 이하에 기술된 상세한 설명을 읽어보면, 본 발명의 이점 및 우수한 특징을, 본 발명의 다른 중요한 측면과 함께, 잘 알 것이다.

실시예

- [0025] 이하의 설명에서, 명세서 및 도면 전체에 걸쳐 유사한 부분은 유사한 참조 번호로 표시되어 있다. 도면은 축척대로 되어 있지 않을 수 있으며, 어떤 구성요소는 일반화된 또는 개략적인 형태로 도시될 수 있고 또 명확함 및 간결함을 위해 상업적 명칭에 의해 식별될 수도 있다.
- [0026] 도 1은 본 발명의 예시적인 실시예에 따른, 분석 및 정정 루프를 갖는, N-채널 오디오 포맷으로부터 M-채널 오디오 포맷으로 동적 다운-믹싱을 하는 시스템을 나타낸 도면이다. 시스템(100)은 5.1 채널 사운드(즉, N = 5)를 사용하고 이 5.1 채널 사운드를 스테레오 사운드(즉, M = 2)로 변환하지만, 다른 적당한 수의 입력 및 출력 채널이 그에 부가하여 또는 다른 대안으로서 사용될 수 있다.
- [0027] 시스템(100)의 동적 다운-믹스 프로세스는 기준 다운-믹스(102), 기준 업-믹스(104), 서브-대역 벡터 계산 시스템(106, 108), 및 서브-대역 정정 시스템(110)을 사용하여 구현된다. 분석 및 정정 루프는 업-믹스 프로세스를 시뮬레이트하는 기준 업-믹스(104), 시뮬레이트된 업-믹스 및 원래의 신호의 주파수 대역별로 에너지 및 위치 벡터를 계산하는 서브-대역 벡터 계산 시스템(106, 108), 및 시뮬레이트된 업-믹스 및 원래의 신호의 에너지 및 위치 벡터를 비교하고 어떤 불일치를 정정하기 위해 다운-믹싱된 신호의 채널간 공간 큐를 수정하는 서브-대역 정정 시스템(110)을 통해 실현된다.
- [0028] 시스템(100)은 수신된 N-채널 오디오를 M-채널 오디오로 변환하는 정적 기준 다운-믹스(102)를 포함한다. 정적 기준 다운-믹스(102)는 5.1 사운드 채널, 좌측 L(T), 우측 R(T), 중앙 C(T), 좌측 서라운드 LS(T) 및 우측 서라운드 RS(T)를 수신하고 이 5.1 채널 신호를 스테레오 채널 신호, 좌측 위터마크 LW'(T) 및 우측 위터마크 RW'(T)로 변환한다.
- [0029] 좌측 위터마크 LW'(T) 및 우측 위터마크 RW'(T) 스테레오 채널 신호는 그 다음에 기준 업-믹스(104)에 제공되고, 이 기준 업-믹스는 스테레오 사운드 채널을 5.1 사운드 채널로 변환한다. 기준 업-믹스(104)는 5.1 사운드 채널, 좌측 L'(T), 우측 R'(T), 중앙 C'(T), 좌측 서라운드 LS'(T) 및 우측 서라운드 RS'(T)를 출력한다.
- [0030] 기준 업-믹스(104)로부터 출력되는 업-믹싱된 5.1 채널 사운드 신호는 이어서 서브-대역 벡터 계산 시스템(10

6)에 제공된다. 서브-대역 벡터 계산 시스템(106)으로부터의 출력은 업-믹싱된 5.1 채널 신호, $L'(T)$, $R'(T)$, $C'(T)$, $LS'(T)$ 및 $RS'(T)$ 에 대한 복수의 주파수 대역에 대한 업-믹싱된 에너지 및 이미지 위치 데이터이다. 이와 유사하게, 원래의 5.1 채널 사운드 신호는 서브-대역 벡터 계산 시스템(108)에 제공된다. 서브-대역 벡터 계산 시스템(108)으로부터의 출력은 원래의 5.1 채널 신호, $L(T)$, $R(T)$, $C(T)$, $LS(T)$ 및 $RS(T)$ 에 대한 복수의 주파수 대역에 대한 소스 에너지 및 이미지 위치 데이터이다. 서브-대역 벡터 계산 시스템(106, 108)에 의해 계산되는 에너지 및 위치 벡터는 이상적인 청취 조건 하에서 듣는 사람에 대한 주어진 주파수 성분의 인식된 세기 및 소스 위치를 나타내는 주파수 대역별 2-차원 벡터 및 총 에너지 측정치로 이루어져 있다. 예를 들어, 오디오 신호는, 유한 임펄스 응답(FIR) 필터 뱅크, 직교 미러 필터(QMF) 뱅크, 이산 푸리에 변환(DFT), 시간 영역 엘리머싱 소거(TDAC) 필터 뱅크, 또는 다른 적당한 필터 뱅크 등의, 적절한 필터 뱅크를 사용하여 시간 영역으로부터 주파수 영역으로 변환될 수 있다. 이 필터 뱅크 출력은 주파수 대역별 총 에너지 및 주파수 대역별 정규화된 이미지 위치 벡터를 결정하기 위해 추가적으로 처리된다.

[0031] 서브-대역 벡터 계산 시스템(106, 108)으로부터 출력되는 에너지 및 위치 벡터 값은 서브-대역 정정 시스템(110)에 제공되며, 이 서브-대역 정정 시스템은 원래의 5.1 채널 사운드에 대한 소스 에너지 및 위치를 이 5.1 채널 사운드에 대한 업-믹싱된 에너지 및 위치와 함께 분석하는데, 그 이유는 원래의 5.1 채널 사운드가 좌측 위터마크 $LW'(T)$ 및 우측 위터마크 $RW'(T)$ 스테레오 채널 신호로부터 발생되기 때문이다. 그 다음에, 스테레오 채널 신호가 차후에 업-믹싱될 때 보다 정확한 다운-믹싱된 스테레오 채널 신호 및 보다 정확한 5.1 표현을 제공하기 위해, $LW(T)$ 및 $RW(T)$ 를 생성하는 좌측 위터마크 $LW'(T)$ 및 우측 위터마크 $RW'(T)$ 신호에 대해 소스 에너지 및 위치 벡터와 업-믹싱된 에너지 및 위치 벡터 간의 차이가 서브-대역별로 식별되고 정정된다. 정정된 좌측 위터마크 $LW(T)$ 및 우측 위터마크 $RW(T)$ 신호는 전송, 스테레오 수신기에 의한 수신, 업-믹스 기능을 갖는 수신기에 의한 수신, 또는 다른 적당한 사용을 위해 출력된다.

[0032] 동작을 설명하면, 시스템(100)은, 다운-믹스/업-믹스 시스템 전체의 시뮬레이션, 분석 및 정정으로 이루어져 있는 지능적 분석 및 정정 루프를 통해, 5.1 채널 사운드를 스테레오 사운드로 동적으로 다운-믹싱한다. 이 방법은 통계적으로 다운-믹싱된 스테레오 신호 $LW'(T)$ 및 $RW'(T)$ 를 발생하고, 후속하는 업-믹싱된 신호 $L'(T)$, $R'(T)$, $C'(T)$, $LS'(T)$ 및 $RS'(T)$ 를 시뮬레이트하며, 좌측 위터마크 $LW'(T)$ 및 우측 위터마크 $RW'(T)$ 스테레오 신호 또는 차후에 업-믹싱되는 서라운드 채널 신호의 품질에 영향을 줄 수 있는 임의의 에너지 또는 위치 벡터 차이를 서브-대역별로 식별 및 정정하기 위해 상기 업-믹싱된 신호를 원래의 5.1 채널 신호와 함께 분석함으로써 달성된다. 좌측 위터마크 $LW(T)$ 및 우측 위터마크 $RW(T)$ 스테레오 신호를 생성하는 서브-대역 정정 프로세싱은, $LW(T)$ 및 $RW(T)$ 가 업-믹싱될 때, 그 결과 얻어지는 5.1 채널 사운드가 향상된 정확도로 원래의 입력 5.1 채널 사운드와 일치하도록 수행된다. 이와 유사하게, 임의의 적당한 수의 입력 채널이 적당한 수의 위터마크된 출력 채널로 변환될 수 있도록 해주기 위해, 예를 들어, 7.1 채널 사운드를 위터마크된 스테레오로, 7.1 채널 사운드를 위터마크된 5.1 채널 사운드로, 커스텀 사운드 채널(자동차 사운드 시스템 또는 극장 등)을 스테레오로, 또는 다른 적당한 변환을 위해, 부가적인 프로세싱이 수행될 수 있다.

[0033] 도 2는 본 발명의 예시적인 실시예에 따른 정적 기준 다운-믹스(200)를 나타낸 도면이다. 정적 기준 다운-믹스(200)는 도 1의 기준 다운-믹스(102)로서 또는 다른 적당한 방식으로 사용될 수 있다.

[0034] 기준 다운-믹스(200)는 N 채널 오디오를 M 채널 오디오로 변환하며, 여기서 N 및 M은 정수이고 N은 M보다 크다. 기준 다운-믹스(200)는 입력 신호 $X_1(T)$, $X_2(T)$, ..., $X_N(T)$ 를 수신한다. 각각의 입력 채널 i에 대해, 입력 신호 $X_i(T)$ 가 신호의 90° 위상 천이를 야기하는 힐버트 변환 유닛(202 내지 206)에 제공된다. 90° 위상 천이를 달성하는 힐버트 필터 또는 전역 통과 필터 회로망 등의 다른 프로세싱이 그에 부가하여 또는 다른 대안으로서 힐버트 변환 유닛 대신에 사용될 수 있다. 각각의 입력 채널 i에 대해, 힐버트 변환된 신호 및 원래의 입력 신호는 이어서 곱셈기(208 내지 218)의 제1 스테이지에 의해 미리 정해진 스케일링 상수 C_{i11} 및 C_{i12} 와 각각 곱해지며, 여기서 첫번째 첨자는 입력 채널 번호 i를 나타내고, 두번째 첨자는 곱셈기의 제1 스테이지를 나타내며, 세번째 첨자는 스테이지별 곱셈기 수를 나타낸다. 곱셈기(208 내지 218)의 출력은 이어서 합산기(220 내지 224)에 의해 합산되어, 분수 힐버트 신호 $X'_i(T)$ 를 발생한다. 곱셈기(220 내지 224)로부터 출력된 분수 힐버트 신호 $X'_i(T)$ 는 대응하는 입력 신호 $X_i(T)$ 에 대해 가변적인 위상 천이량을 갖는다. 위상 천이량은 스케일링 상수 C_{i11} 및 C_{i12} 에 의존하며, 여기서 0° 위상 천이는 $C_{i11} = 0$ 및 $C_{i12} = 1$ 에 대응하여 가능하고, $\pm 90^\circ$ 위상 천이는 $C_{i11} = \pm 1$ 및 $C_{i12} = 0$ 에 대응하여 가능하다. 임의의 중간 위상 천이량은 C_{i11} 및 C_{i12} 의 적절한 값으로 가능하다.

[0035] 각각의 입력 채널 i에 대한 각각의 신호 $X'_i(T)$ 는 이어서 제2 스테이지의 곱셈기(226 내지 242)에 의해 미리 정

해진 스케일링 상수 C_{i2j} 와 곱해지며, 여기서 첫번째 첨자는 입력 채널 번호 i 를 나타내고, 두번째 첨자는 제2 스테이지의 곱셈기를 나타내며, 세번째 첨자는 출력 채널 번호 j 를 나타낸다. 곱셈기(226 내지 242)의 출력은 이어서 합산기(244 내지 248)에 의해 적절히 합산되어 각각의 출력 채널 j 에 대한 대응하는 출력 신호 $Y_j(T)$ 를 발생한다. 각각의 입력 채널 i 및 출력 채널 j 에 대한 스케일링 상수 C_{i2j} 는 각각의 입력 채널 i 및 출력 채널 j 의 공간 위치에 의해 결정된다. 예를 들어, 좌측 입력 채널 i 및 우측 출력 채널 j 에 대한 스케일링 상수 C_{i2j} 는 공간 구분(spatial distinction)을 유지하기 위해 0에 가깝게 설정될 수 있다. 이와 유사하게, 전방 입력 채널 i 에 대한 스케일링 상수 C_{i2j} 는 공간 배치(spatial placement)를 유지하기 위해 1에 가깝게 설정될 수 있다.

[0036] 동작을 설명하면, 기준 다운-믹스(200)는, 출력 신호가 수신기에 수신될 때, 입력 신호들 간의 공간 관계가 임의적으로 관리 및 추출될 수 있게 해주는 방식으로, N 개의 사운드 채널을 M 개의 사운드 채널로 결합한다. 게다가, 도시된 바와 같은 N 개 채널 사운드의 결합은 M 채널 오디오 환경에서 청취하고 있는 듣는 사람에게 만족스러운 품질을 갖는 M 개 채널 사운드를 발생한다. 따라서, N 채널 사운드를, M 채널 수신기, 적당한 업-믹서를 갖는 N 채널 수신기, 또는 다른 적당한 수신기에서 사용될 수 있는 M 채널 사운드로 변환하기 위해 기준 다운-믹스(200)가 사용될 수 있다.

[0037] 도 3은 본 발명의 예시적인 실시예에 따른 정적 기준 다운-믹스(300)를 나타낸 도면이다. 도 3에 나타난 바와 같이, 정적 기준 다운-믹스(300)는, 5.1 채널 시간 영역 데이터들 스테레오 채널 시간 영역 데이터로 변환하는 도 2의 정적 기준 다운-믹스(200)의 구현이다. 정적 기준 다운-믹스(300)는 도 1의 기준 다운-믹스(102)로서 또는 다른 적당한 방식으로 사용될 수 있다.

[0038] 기준 다운-믹스(300)는, 소스 5.1 채널 사운드의 좌채널 신호 $L(T)$ 를 수신하고 그 시간 신호에 대해 힐버트 변환을 수행하는 힐버트 변환(302)을 포함한다. 이 힐버트 변환은 신호의 90° 위상 천이를 야기하며, 이 신호는 이어서 곱셈기(310)에 의해 미리 정해진 스케일링 상수 C_{L1} 와 곱해진다. 90° 위상 천이를 달성하는 힐버트 필터 또는 전역-통과 필터 회로망 등의 다른 프로세싱이 그에 부가하여 또는 다른 대안으로서 힐버트 변환 유닛 대신에 사용될 수 있다. 원래의 좌채널 신호 $L(T)$ 는 곱셈기(312)에 의해 미리 정해진 스케일링 상수 C_{L2} 와 곱해진다. 곱셈기(310, 312)의 출력은 합산기(320)에 의해 합산되어 분수 힐버트 신호 $L'(T)$ 를 발생한다. 이와 유사하게, 소스 5.1 채널 사운드로부터의 우채널 신호 $R(T)$ 는 힐버트 변환(304)에 의해 처리되고 곱셈기(314)에 의해 미리 정해진 스케일링 상수 C_{R1} 와 곱해진다. 원래의 우채널 신호 $R(T)$ 는 곱셈기(316)에 의해 미리 정해진 스케일링 상수 C_{R2} 와 곱해진다. 곱셈기(314, 316)의 출력은 합산기(322)에 의해 합산되어 분수 힐버트 신호 $R'(T)$ 를 발생한다. 곱셈기(320, 322)로부터 출력되는 분수 힐버트 신호 $L'(T)$ 및 $R'(T)$ 는 각각 대응하는 입력 신호 $L(T)$ 및 $R(T)$ 에 대해 가변적인 위상 천이량을 갖는다. 위상 천이량은 스케일링 상수 C_{L1} , C_{L2} , C_{R1} 및 C_{R2} 에 의존하며, 여기서 0° 위상 천이는 $C_{L1} = 0$ 및 $C_{L2} = 1$, $C_{R1} = 0$ 및 $C_{R2} = 1$ 에 대응하여 가능하고, $\pm 90^\circ$ 위상 천이는 $C_{L1} = \pm 1$ 및 $C_{L2} = 0$, $C_{R1} = \pm 1$ 및 $C_{R2} = 0$ 에 대응하여 가능하다. 임의의 중간의 위상 천이량은 C_{L1} , C_{L2} , C_{R1} 및 C_{R2} 의 적절한 값으로 가능하다. 소스 5.1 채널 사운드로부터 입력되는 중앙 채널은 곱셈기(318)에 분수 힐버트 신호 $C'(T)$ 로서 제공되며, 이는 중앙 채널 입력 신호에 대해 위상 천이가 수행되지 않음을 의미한다. 곱셈기(318)는 $C'(T)$ 를, 3 데시벨 정도의 감쇄 등의 미리 정해진 스케일링 상수 C_3 와 곱한다. 합산기(320, 322) 및 곱셈기(318)의 출력은 적절히 합산되어 좌측 위터마크 채널 $LW'(T)$ 및 우측 위터마크 채널 $RW'(T)$ 로 된다.

[0039] 소스 5.1 채널 사운드로부터의 서라운드 좌채널 $LS(T)$ 는 힐버트 변환(306)에 제공되고, 소스 5.1 채널 사운드로부터의 서라운드 우채널 $RS(T)$ 은 힐버트 변환(308)에 제공된다. 힐버트 변환(306, 308)의 출력은 분수 힐버트 신호 $LS'(T)$ 및 $RS'(T)$ 이고, 이는 $LS(T)$ 및 $LS'(T)$ 쌍과 $RS(T)$ 및 $RS'(T)$ 쌍 간에 전체 90° 위상 천이가 존재함을 암시한다. $LS'(T)$ 는 이어서 곱셈기(324, 326)에 의해 미리 정해진 스케일링 상수 C_{LS1} 및 C_{LS2} 와 각각 곱해진다. 이와 마찬가지로, $RS'(T)$ 는 곱셈기(328, 330)에 의해 미리 정해진 스케일링 상수 C_{RS1} 및 C_{RS2} 와 각각 곱해진다. 곱셈기(324 내지 330)의 출력은 좌측 위터마크 채널 $LW'(T)$ 및 우측 위터마크 채널 $RW'(T)$ 에 적절히 제공된다.

[0040] 합산기(332)는 합산기(320)로부터 출력된 좌채널, 곱셈기(318)로부터 출력된 중앙 채널, 곱셈기(324)로부터 출력된 서라운드 좌채널, 및 곱셈기(328)로부터 출력된 서라운드 우채널을 수신하고, 이들 신호를 가산하여 좌측

위터마크 채널 $LW'(T)$ 을 형성한다. 이와 유사하게, 합산기(334)는 곱셈기(318)로부터 출력된 중앙 채널, 합산기(322)로부터 출력된 우채널, 곱셈기(326)로부터 출력된 서라운드 좌채널, 및 곱셈기(330)로부터 출력된 서라운드 우채널을 수신하고, 이들 신호를 가산하여 우측 위터마크 채널 $RW'(T)$ 을 형성한다.

[0041] 동작을 설명하면, 기준 다운-믹스(300)는, 좌측 위터마크 채널 및 우측 위터마크 채널 스테레오 신호가 수신기에 수신될 때, 5.1 입력 채널 간의 공간 관계가 유지되고 추출될 수 있게 해주는 방식으로, 소스 5.1 사운드 채널들을 결합한다. 게다가, 도식된 바와 같이 5.1 채널 사운드를 결합하면 서라운드 사운드 업-믹스를 수행하지 않는 스테레오 수신기를 사용하는 듣는 사람에게 만족할만한 품질을 갖는 스테레오 사운드를 발생한다. 따라서, 기준 다운-믹스(300)는 5.1 채널 사운드를 스테레오 수신기, 적당한 업-믹서를 갖는 5.1 채널 수신기, 적당한 업-믹서를 갖는 7.1 채널 수신기, 또는 다른 적당한 수신기에서 사용될 수 있는 스테레오 사운드로 변환하는 데 사용될 수 있다.

[0042] 도 4는 본 발명의 예시적인 실시예에 따른 서브-대역 벡터 계산 시스템(400)을 나타낸 도면이다. 서브-대역 벡터 계산 시스템(400)은 복수의 주파수 대역에 대한 에너지 및 위치 벡터 데이터를 제공하고, 도 1의 서브-대역 벡터 계산 시스템(106, 108)으로서 사용될 수 있다. 5.1 채널 사운드가 도식되어 있지만, 다른 적당한 채널 구성이 사용될 수 있다.

[0043] 서브-대역 벡터 계산 시스템(400)은 시간-주파수 분석 유닛(402 내지 410)을 포함한다. 5.1 시간 영역 사운드 채널 $L(T)$, $R(T)$, $C(T)$, $LS(T)$ 및 $RS(T)$ 는 시간-주파수 분석 유닛(402 내지 410)에 각각 제공되고, 이 시간-주파수 분석 유닛은 시간 영역 신호를 주파수 영역 신호로 변환한다. 이들 시간-주파수 분석 유닛은, 유한 임펄스 응답(FIR) 필터 뱅크, 직교 미리 필터(QMF) 뱅크, 이산 푸리에 변환(DFT), 시간-영역 엘리머싱 소거(TDAC) 필터 뱅크, 또는 다른 적당한 필터 뱅크 등의 적절한 필터 뱅크일 수 있다. $L(F)$, $R(F)$, $C(F)$, $LS(F)$ 및 $RS(F)$ 에 대해 주파수 대역별 크기 또는 에너지 값이 시간-주파수 분석 유닛(402 내지 410)으로부터 출력된다. 이들 크기/에너지 값은 각각의 대응하는 채널의 각각의 주파수 대역 성분에 대한 크기/에너지 측정치로 이루어져 있다. 크기/에너지 측정치는 합산기(412)에서 합산되고, 이 합산기는 $T(F)$ 를 출력하며, 여기서 $T(F)$ 는 주파수 대역별 입력 신호의 총 에너지이다. 이 값은 이어서 분할 유닛(414 내지 422)에 의해 채널 크기/에너지 값 각각으로 분할되어, 대응하는 정규화된 채널간 레벨 차이(inter-channel level difference, ICLD) 신호 $M_L(F)$, $M_R(F)$, $M_C(F)$, $M_{LS}(F)$ 및 $M_{RS}(F)$ 를 발생하며, 여기서 이들 ICLD 신호는 각각의 채널에 대한 정규화된 서브-대역 에너지 추정치로 볼 수 있다.

[0044] 5.1 채널 사운드는 도식된 바와 같이 가로축과 깊이축으로 이루어진 2차원 평면 상에 예시적인 위치를 갖는 정규화된 위치 벡터에 매핑된다. 도식된 바와 같이, (X_{LS}, Y_{LS}) 에 대한 위치값은 원점에 할당되고, (X_{RS}, Y_{RS}) 에 대한 값은 $(0, 1)$ 에 할당되며, (X_L, Y_L) 의 값은 $(1, 1-C)$ 에 할당되고, 여기서 C 는 방의 후방으로부터 좌측 및 우측 스피커에 대한 후퇴 거리(setback distance)를 나타내는 1과 0 사이의 값이다. 이와 유사하게, (X_R, Y_R) 의 값은 $(1, 1-C)$ 이다. 마지막으로, (X_C, Y_C) 의 값은 $(0.5, 1)$ 이다. 이 좌표는 예시적인 것이며, 스피커 좌표가 방의 크기, 방의 형상 또는 다른 인자에 기초하여 다른 경우와 같이, 서로에 대한 스피커의 실제의 정규화된 위치 또는 구성을 반영하기 위해 변경될 수 있다. 예를 들어, 7.1 사운드 또는 다른 적당한 사운드 채널 구성이 사용되는 경우, 방에서의 스피커의 위치를 반영하는 부가적인 좌표값이 제공될 수 있다. 이와 유사하게, 이러한 스피커 위치는 자동차, 방, 강당, 경기장에서의 스피커의 실제 분포에 기초하여 또는 다른 적당한 방식으로 조정될 수 있다.

[0045] 추정된 이미지 위치 벡터 $P(F)$ 는 이하의 벡터식에 기술된 바와 같이 서브-대역별로 계산될 수 있다.

[0046]
$$P(F) = M_L(F) * (X_L, Y_L) + M_R(F) * (X_R, Y_R) + M_C(F) * (X_C, Y_C) + i \cdot (M_{LS}(F) * (X_{LS}, Y_{LS}) + M_{RS}(F) * (X_{RS}, Y_{RS}))$$

[0047] 따라서, 각각의 주파수 대역에 대해, 그 주파수 대역에 대한 겉보기 주파수 소스의 인식된 세기 및 위치를 정의하는 데 사용되는 총 에너지 $T(F)$ 및 위치 벡터 $P(F)$ 의 출력이 제공된다. 이와 같이, 예를 들어, 서브-대역 정정 시스템(110)에서 사용하기 위해 또는 다른 적당한 목적을 위해, 주파수 성분의 공간 이미지가 국소화될 수 있다.

[0048] 도 5는 본 발명의 예시적인 실시예에 따른 서브-대역 정정 시스템을 나타낸 도면이다. 서브-대역 정정 시스템이 도 1의 서브-대역 정정 시스템(110)으로서 또는 다른 적당한 목적으로 사용될 수 있다. 서브-대역 정정 시스템은 좌측 위터마크 $LW'(T)$ 및 우측 위터마크 $RW'(T)$ 스테레오 채널 신호를 수신하고 기준 다운-믹싱 또는 다른 적당한 방법의 결과로서 생성될 수 있는 각각의 주파수 대역에 대한 신호 부정확성을 보상하기 위해 위터마크

크된 신호에 에너지 및 이미지 정정을 수행한다. 서브-대역 정정 시스템은, 각각의 서브-대역에 대해, 소스의 총 에너지 신호 $T_{SOURCE}(F)$ 및 차후의 업-믹싱된 신호 $T_{UMIX}(F)$ 및 소스에 대한 위치 벡터 $P_{SOURCE}(F)$ 및 차후의 업-믹싱된 신호 $P_{UMIX}(F)$ (도 1의 서브-대역 벡터 계산 시스템(106, 108)에 의해 발생하는 것 등)를 수신하고 이용한다. 이들 총 에너지 신호 및 위치 벡터는 수행할 적절한 정정 및 보상을 결정하는 데 사용된다.

[0049] 서브-대역 정정 시스템은 위치 정정 시스템(500) 및 스펙트럼 에너지 정정 시스템(502)을 포함한다. 위치 정정 시스템(500)은 좌측 위터마크 스테레오 채널 $LW'(T)$ 및 우측 위터마크 스테레오 채널 $RW'(T)$ 에 대한 시간 영역 신호를 수신하고, 이들 신호는 시간-주파수 분석 유닛(504, 506)에 의해 시간 영역으로부터 주파수 영역으로 각각 변환된다. 이들 시간-주파수 분석 유닛은, 유한 임펄스 응답(FIR) 필터 뱅크, 직교 미러 필터(QMF) 뱅크, 이산 푸리에 변환(DFT), 시간 영역 엘리머싱 소거(TDAC) 필터 뱅크, 또는 다른 적당한 필터 뱅크 등의, 적절한 필터 뱅크일 수 있다.

[0050] 시간-주파수 분석 유닛(504, 506)의 출력은 주파수 영역 서브-대역 신호 $LW'(F)$ 및 $RW'(F)$ 이다. 채널간 차이(ICLD) 및 채널간 코히런스(ICC)의 관련 공간 큐는 신호 $LW'(F)$ 및 $RW'(F)$ 에서 서브-대역별로 수정된다. 예를 들어, 이들 큐는, $LW'(F)$ 및 $RW'(F)$ 의 절대값 및 $LW'(F)$ 및 $RW'(F)$ 의 위상각으로 나타내어진, $LW'(F)$ 및 $RW'(F)$ 의 크기 또는 에너지의 조작을 통해 수정될 수 있다. ICLD의 정정은 곱셈기(508)에 의해 $LW'(F)$ 의 크기/에너지 값을 이하의 식에 의해 발생된 값과 곱함으로써 수행된다.

[0051] $[X_{MAX} - P_{X,SOURCE}(F)]/[X_{MAX} - P_{X,UMIX}(F)]$

[0052] 여기서, X_{MAX} = 최대 X 좌표 경계이고,

[0053] $P_{X,SOURCE}(F)$ = 소스 벡터로부터의 추정된 서브-대역 X 위치 좌표이며,

[0054] $P_{X,UMIX}(F)$ = 차후의 업-믹스 벡터로부터의 추정된 서브-대역 X 위치 좌표이다.

[0055] 이와 마찬가지로, $RW'(F)$ 의 크기/에너지는, 곱셈기(510)에 의해, 이하의 식에 의해 발생된 값과 곱해진다.

[0056] $[P_{X,SOURCE}(F) - X_{MIN}] / [P_{X,UMIX}(F) - X_{MIN}]$

[0057] 여기서, X_{MIN} = 최소 X 좌표 경계이다.

[0058] ICC의 정정은, 가산기(512)에 의해, $LW'(F)$ 의 위상각을 이하의 식에 의해 발생된 값과 가산함으로써 행해진다.

[0059] $+/- \Pi * [P_{Y,SOURCE}(F) - P_{Y,UMIX}(F)]/[Y_{MAX} - Y_{MIN}]$

[0060] 여기서, $P_{Y,SOURCE}(F)$ = 소스 벡터로부터의 추정된 서브-대역 Y 위치 좌표이고,

[0061] $P_{Y,UMIX}(F)$ = 차후의 업-믹스 벡터로부터의 추정된 서브-대역 Y 위치 좌표이며,

[0062] Y_{MAX} = 최대 Y 좌표 경계이고,

[0063] Y_{MIN} = 최소 Y 좌표 경계이다.

[0064] 이와 유사하게, $RW'(F)$ 에 대한 위상각은, 가산기(514)에 의해, 이하의 식에 의해 발생된 값에 가산된다.

[0065] $-/+ \Pi * [P_{Y,SOURCE}(F) - P_{Y,UMIX}(F)]/[Y_{MAX} - Y_{MIN}]$

[0066] 유의할 점은 $LW'(F)$ 및 $RW'(F)$ 에 가산된 각도 성분이 동일한 값을 갖지만 반대 극성을 가지며, 여기서 그 결과 얻어지는 극성이 $LW'(F)$ 와 $RW'(F)$ 간의 앞선 위상각에 의해 결정된다는 것이다.

[0067] 정정된 $LW'(F)$ 크기/에너지 및 $LW'(F)$ 위상각은 가산기(516)에 의해 재결합되어 각각의 서브-대역에 대한 복소 값 $LW(F)$ 을 형성하고 이어서 주파수-시간 합성 유닛(520)에 의해 좌측 위터마크 시간 영역 신호 $LW(T)$ 로 변환된다. 이와 마찬가지로, 정정된 $RW'(F)$ 크기/에너지 및 $RW'(F)$ 위상각은 가산기(518)에 의해 재결합되어 각각의 서브-대역에 대한 복소 값 $RW(F)$ 을 형성하고 이어서 주파수-시간 합성 유닛(522)에 의해 우측 위터마크 시간 영역 신호 $RW(T)$ 로 변환된다. 주파수-시간 합성 유닛(520, 522)은 주파수 영역 신호를 다시 시간 영역 신호로 변환할 수 있는 적당한 합성 필터 뱅크일 수 있다.

[0068] 이 예시적인 실시예에 나타난 바와 같이, 위터마크 좌측 및 우측 채널 신호의 각각의 스펙트럼 성분

널간 공간 큐는 ICLD 및 ICC 공간 큐를 적절히 수정하는 위치 정정(500)을 사용하여 정정될 수 있다.

[0069] 스펙트럼 에너지 정정 시스템(502)은 다운-믹싱된 신호의 전체적인 스펙트럼 균형(total spectral balance)이 원래의 5.1 신호의 전체적인 스펙트럼 균형과 일치하고 따라서, 예를 들어, 콤 필터링(comb filtering)에 의해 야기되는 스펙트럼 편이를 보상하도록 보장하는 데 사용될 수 있다. 좌측 위터마크 시간 영역 신호 및 우측 위터마크 시간 영역 신호 $LW'(T)$ 및 $RW'(T)$ 는 시간-주파수 분석 유닛(524, 526)을 사용하여 시간 영역으로부터 주파수 영역으로 각각 변환된다. 이들 시간-주파수 분석 유닛은, 유한 임펄스 응답(FIR) 필터 뱅크, 직교 미리 필터(QMF) 뱅크, 이산 푸리에 변환(DFT), 시간-영역 엘리머싱 소거(TDAC) 필터 뱅크, 또는 다른 적당한 필터 뱅크 등의, 적절한 필터 뱅크일 수 있다. 시간-주파수 분석 유닛(524, 526)으로부터의 출력은 $LW'(F)$ 및 $RW'(F)$ 주파수 서브-대역 신호이며, 이들 신호는 곱셈기(528, 530)에 의해 $T_{SOURCE}(F)/T_{UMIX}(F)$ 와 곱해지며, 여기서

$$T_{SOURCE}(F) = |L(F)| + |R(F)| + |C(F)| + |LS(F)| + |R(F)|$$

$$T_{UMIX}(F) = |L_{UMIX}(F)| + |R_{UMIX}(F)| + |C_{UMIX}(F)| + |LS_{UMIX}(F)| +$$

$$|RS_{UMIX}(F)|$$

[0073] 곱셈기(528, 530)로부터의 출력은 이어서 주파수-시간 합성 유닛(532, 534)에 의해 다시 주파수 영역으로부터 시간 영역으로 변환되어 $LW(T)$ 및 $RW(T)$ 를 발생한다. 주파수-시간 합성 유닛은 주파수 영역 신호를 다시 시간 영역 신호로 변환할 수 있는 적당한 합성 필터 뱅크일 수 있다. 이와 같이, 원래의 5.1 신호에 충실한 좌측 및 우측 위터마크 채널 신호 $LW(T)$ 및 $RW(T)$ 를 생성하기 위해, 위치 및 에너지 정정이 다운-믹싱된 스테레오 채널 신호 $LW'(F)$ 및 $RW'(F)$ 에 적용될 수 있다. $LW(T)$ 및 $RW(T)$ 는, 원래의 5.1 채널 사운드에 존재하는 임의적인 내용 요소의 스펙트럼 성분 위치 또는 에너지를 그다지 변경시키지 않고, 스테레오로 재생될 수 있거나 다시 5.1 채널로 또는 다른 적당한 수의 채널로 업-믹싱될 수 있다.

[0074] 도 6은 본 발명의 예시적인 실시예에 따른, M개 채널로부터 N개 채널로 데이터를 업-믹싱하는 시스템(600)을 나타낸 도면이다. 시스템(600)은 스테레오 시간 영역 데이터를 N 채널 시간 영역 데이터로 변환한다.

[0075] 시스템(600)은 시간-주파수 분석 유닛(602, 604), 필터 발생 유닛(606), 평활화 유닛(608), 및 주파수-시간 합성 유닛(634 내지 638)을 포함한다. 시스템(600)은, 고분해능 주파수 대역 프로세싱을 가능하게 해주는 확장가능한 주파수 영역 아키텍처를 통해, 또한 업-믹싱된 N 채널 신호에서의 주파수 성분의 공간적 배치를 도출하기 위해 주파수 대역별로 중요한 채널간 공간 큐를 추출 및 분석하는 필터 발생 방법을 통해, 업-믹스 프로세스에서 향상된 공간 구분 및 안정성을 제공한다.

[0076] 시스템(600)은 시간-주파수 분석 유닛(602, 604)에서 좌채널 스테레오 신호 $L(T)$ 및 우채널 스테레오 신호 $R(T)$ 를 수신하고, 이 시간-주파수 분석 유닛은 시간 영역 신호를 주파수 영역 신호로 변환한다. 이들 시간-주파수 분석 유닛은, 유한 임펄스 응답(FIR) 필터 뱅크, 직교 미리 필터(QMF) 뱅크, 이산 푸리에 변환(DFT), 시간-영역 엘리머싱 소거(TDAC) 필터 뱅크, 또는 다른 적당한 필터 뱅크 등의, 적절한 필터 뱅크일 수 있다. 시간-주파수 분석 유닛(602, 604)으로부터의 출력은, 분석 필터 뱅크 서브-대역 대역폭이 심리 음향적 임계 대역(psych-acoustic critical band), 등가 장방형 대역폭(equivalent rectangular bandwidth), 또는 어떤 다른 인지 특성에 근사하도록 처리될 수 있는 경우 0 내지 20 kHz 주파수 범위 등의, 사람의 청각 시스템의 충분한 주파수 범위를 포함하는 일련의 주파수 영역 값이다. 이와 유사하게, 다른 적당한 수의 주파수 대역 및 범위가 사용될 수 있다.

[0077] 시간-주파수 분석 유닛(602, 604)으로부터의 출력은 필터 발생 유닛(606)에 제공된다. 한 예시적인 실시예에서, 필터 발생 유닛(606)은 주어진 환경을 위해 출력되어야만 하는 채널의 수에 관한 외부 선택을 수신할 수 있다. 예를 들어, 2개의 전방 스피커 및 2개의 후방 스피커가 있는 경우 4.1 사운드 채널이 선택될 수 있거나, 2개의 전방 스피커 및 2개의 후방 스피커 및 1개의 전방 중앙 스피커가 있는 경우, 5.1 사운드 시스템이 선택될 수 있거나, 2개의 전방 스피커, 2개의 측방 스피커, 2개의 후방 스피커, 및 1개의 전방 중앙 스피커가 있는 경우, 7.1 사운드 시스템이 선택될 수 있거나, 다른 적당한 사운드 시스템이 선택될 수 있다. 필터 발생 유닛(606)은 주파수 대역별로 채널간 레벨 차이(ICLD) 및 채널간 코히런스(ICC) 등의 채널간 공간 큐를 추출 및 분석한다. 이들 관련 공간 큐는 이어서 업-믹싱된 음장(sound field)에서의 주파수 대역 성분의 공간 배치를 제어하는 적응 채널 필터를 발생하기 위한 파라미터로서 사용된다. 아주 빠르게 변할 수 있는 경우에 짜증나는 변동 효과를 야기할 수 있는 필터 변동성을 제한하기 위해 시간 및 주파수 둘다에 걸쳐 평활화 유닛(608)에 의해 채널 필터가 평활화된다. 도 6에 나타난 예시적인 실시예에서, 좌측 및 우측 채널 $L(F)$ 및 $R(F)$

주파수 영역 신호가, 평활화 유닛(608)에 제공되는 N 채널 필터 신호 $H_1(F)$, $H_2(F)$, ..., $H_N(F)$ 를 생성하는 필터 발생 유닛(606)에 제공된다.

[0078] 평활화 유닛(608)은 N 채널 필터의 각각의 채널에 대한 주파수 영역 성분을 시간 및 주파수 차원 둘다에 걸쳐 평균을 구한다. 시간 및 주파수에 걸친 평활화는 채널 필터 신호의 빠른 변동을 제어하는 데 도움이 되며, 따라서 듣는 사람에게 짜증나는 것일 수 있는 지터 아티팩트 및 불안정성을 감소시킨다. 한 예시적인 실시예에서, 시간 평활화는 1차 저역-통과 필터를 현재 프레임으로부터의 각각의 주파수 대역 및 이전의 프레임으로부터의 대응하는 주파수 대역에 적용함으로써 실현될 수 있다. 이것은 프레임마다 각각의 주파수 대역의 변동성을 감소시키는 효과를 갖는다. 다른 예시적인 실시예에서, 스펙트럼 평활화는 사람의 청각 시스템의 임계 대역 간격에 가깝도록 모델링되는 주파수 빈(frequency bin)의 그룹들에 걸쳐 수행될 수 있다. 예를 들어, 균일한 간격의 주파수 빈을 갖는 분석 필터 뱅크가 이용되는 경우, 서로 다른 수의 주파수 빈이 그룹화되고 주파수 스펙트럼의 서로 다른 부분에 대해 평균될 수 있다. 예를 들어, 0부터 5 kHz까지는, 5개의 주파수 빈이 평균될 수 있거나, 5 kHz부터 10 kHz까지는, 7개의 주파수 빈이 평균될 수 있거나, 10 kHz부터 20 kHz까지는, 9개의 주파수 빈이 평균될 수 있거나, 다른 적당한 수의 주파수 빈 및 대역폭 범위가 선택될 수 있다. $H_1(F)$, $H_2(F)$, ..., $H_N(F)$ 의 평활화된 값은 평활화 유닛(608)으로부터 출력된다.

[0079] N 개의 출력 채널 각각에 대한 소스 신호 $X_1(F)$, $X_2(F)$, ..., $X_N(F)$ 가 M 개 입력 채널의 적응적 결합으로서 발생된다. 도 6에 나타난 예시적인 실시예에서, 주어진 출력 채널 i 에 대해, 합산기(614, 620, 626)로부터 출력된 채널 소스 신호 $X_i(F)$ 는 $L(F)$ 를 적응적 스케일링 신호 $G_i(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_i(F)$ 와 곱한 것의 합산으로서 발생된다. 곱셈기(610, 612, 616, 618, 622, 624)에 의해 사용되는 적응적 스케일링 신호 $G_i(F)$ 는 주파수 대역별로 출력 채널 i 의 의도된 공간 위치와 $L(F)$ 및 $R(F)$ 의 동적 채널간 코히런스 추정치에 의해 결정된다. 이와 유사하게, 합산기(614, 620, 626)에 제공되는 신호의 극성은 출력 채널 i 의 의도된 공간 위치에 의해 결정된다. 예를 들어, 합산기(614, 620, 626)에서의 적응적 스케일링 신호 $G_i(T)$ 및 극성은, 종래의 매트릭스 업-믹싱 방법에서 통상적인 바와 같이, 전방 중앙 채널에 $L(F)+R(F)$ 결합을, 좌측 채널에 $L(F)$ 를, 우측 채널에 $R(F)$ 를, 그리고 후방 채널에 $L(F)-R(F)$ 결합을 제공하도록 설계될 수 있다. 적응적 스케일링 신호 $G_i(F)$ 는 또한 출력 채널 쌍들(이들이 가로 또는 깊이 채널 쌍인지에 상관없음) 간의 상관을 동적으로 조정하는 방법을 제공할 수 있다.

[0080] 채널 소스 신호 $X_1(F)$, $X_2(F)$, ..., $X_N(F)$ 는 곱셈기(628 내지 632)에 의해 평활화된 채널 필터 $H_1(F)$, $H_2(F)$, ..., $H_N(F)$ 와 각각 곱해진다.

[0081] 곱셈기(628 내지 632)로부터의 출력은 이어서 주파수-시간 합성 유닛(634 내지 638)에 의해 주파수 영역으로부터 시간 영역으로 변환되어 출력 채널 $Y_1(T)$, $Y_2(T)$, ..., $Y_N(T)$ 를 발생한다. 이와 같이, 좌측 및 우측 스테레오 신호는 N 개 채널 신호로 업-믹싱되고, 여기서 자연적으로 존재하거나, 도 1의 다운-믹싱 위터마크 프로세스 또는 다른 적당한 프로세스 등에 의해, 좌측 및 우측 스테레오 신호로 의도적으로 인코딩된 채널간 공간 큐는 시스템(600)에 의해 생성되는 N 채널 음장 내에서의 주파수 성분의 공간적 배치를 제어하는 데 사용될 수 있다. 이와 유사하게, 스테레오에서 7.1 사운드로, 5.1에서 7.1로, 또는 다른 적당한 조합 등의, 다른 적당한 조합의 입력 및 출력이 사용될 수 있다.

[0082] 도 7은 본 발명의 예시적인 실시예에 따른, M 개 채널을 N 개 채널로 데이터를 업-믹싱하는 시스템(700)을 나타낸 도면이다. 시스템(700)은 스테레오 시간 영역 데이터를 5.1 채널 시간 영역 데이터로 변환한다.

[0083] 시스템(700)은 시간-주파수 분석 유닛(702, 704), 필터 발생 유닛(706), 평활화 유닛(708), 및 주파수-시간 합성 유닛(738 내지 746)을 포함한다. 시스템(700)은, 고분해능 주파수 대역 프로세싱을 가능하게 해주는 확장 가능한 주파수 영역 아키텍처의 사용을 통해, 또한 업-믹싱된 5.1 채널 신호에서의 주파수 성분의 공간적 배치를 도출하기 위해 주파수 대역별로 중요한 채널간 공간 큐를 추출 및 분석하는 필터 발생 방법을 통해, 업-믹스 프로세스에서 향상된 공간 구분 및 안정성을 제공한다.

[0084] 시스템(700)은 시간-주파수 분석 유닛(702, 704)에서 좌채널 스테레오 신호 $L(T)$ 및 우채널 스테레오 신호 $R(T)$ 를 수신하고, 이 시간-주파수 분석 유닛은 시간 영역 신호를 주파수 영역 신호로 변환한다. 이들 시간-주파수 분석 유닛은, 유한 임펄스 응답(FIR) 필터 뱅크, 직교 미리 필터(QMF) 뱅크, 이산 푸리에 변환(DFT), 시간-영역 엘리머싱 소거(TDAC) 필터 뱅크, 또는 다른 적당한 필터 뱅크 등의, 적절한 필터 뱅크일 수 있다. 시간-

주파수 분석 유닛(702, 704)으로부터의 출력은, 분석 필터 बैं크 서브-대역 대역폭이 심리 음향적 임계 대역 (psycho-acoustic critical band), 등가 장방형 대역폭(equivalent rectangular bandwidth), 또는 어떤 다른 인지 특성에 근사하도록 처리될 수 있는 경우 0 내지 20 kHz 주파수 범위 등의, 사람의 청각 시스템의 충분한 주파수 범위를 포함하는 일련의 주파수 영역 값이다. 이와 유사하게, 다른 적당한 수의 주파수 대역 및 범위가 사용될 수 있다.

[0085] 시간-주파수 분석 유닛(702, 704)으로부터의 출력은 필터 발생 유닛(706)에 제공된다. 한 예시적인 실시예에서, 필터 발생 유닛(706)은 주어진 환경을 위해 출력되어야만 하는 채널의 수에 관한 외부 선택을 수신할 수 있다. 예를 들어, 2개의 전방 스피커 및 2개의 후방 스피커가 있는 경우 4.1 사운드 채널이 선택될 수 있거나, 2개의 전방 스피커 및 2개의 후방 스피커 및 1개의 전방 중앙 스피커가 있는 경우에, 5.1 사운드 시스템이 선택될 수 있거나, 2개의 전방 스피커 및 1개의 전방 중앙 스피커가 있는 경우에, 3.1 사운드 시스템이 선택될 수 있거나, 다른 적당한 사운드 시스템이 선택될 수 있다. 필터 발생 유닛(706)은 주파수 대역별로 채널 간 레벨 차이(ICLD) 및 채널간 코히런스(ICC) 등의 채널간 공간 큐를 추출 및 분석한다. 이들 관련 공간 큐는 이어서 업-믹싱된 음장(sound field)에서의 주파수 대역 성분의 공간 배치를 제어하는 적응 채널 필터를 발생하기 위한 파라미터로서 사용된다. 아주 빠르게 변할 수 있는 경우에 짜증나는 변동 효과를 야기할 수 있는 필터 변동성을 제한하기 위해 시간 및 주파수 둘다에 걸쳐 평활화 유닛(708)에 의해 채널 필터가 평활화된다. 도 7에 나타난 예시적인 실시예에서, 좌측 및 우측 채널 $L(F)$ 및 $R(F)$ 주파수 영역 신호가, 평활화 유닛(708)에 제공되는 5.1 채널 필터 신호 $H_L(F)$, $H_R(F)$, $H_C(F)$, $H_{LS}(F)$, 및 $H_{RS}(F)$ 를 생성하는 필터 발생 유닛(706)에 제공된다.

[0086] 평활화 유닛(708)은 5.1 채널 필터의 각각의 채널에 대한 주파수 영역 성분을 시간 및 주파수 차원 둘다에 걸쳐 평균을 구한다. 시간 및 주파수에 걸친 평활화는 채널 필터 신호의 빠른 변동을 제어하는 데 도움이 되며, 따라서 듣는 사람에게 짜증나는 것일 수 있는 지터 아티팩트 및 불안정성을 감소시킨다. 한 예시적인 실시예에서, 시간 평활화는 1차 저역-통과 필터를 현재 프레임으로부터의 각각의 주파수 대역 및 이전의 프레임으로부터의 대응하는 주파수 대역에 적용함으로써 실현될 수 있다. 이것은 프레임마다 각각의 주파수 대역의 변동성을 감소시키는 효과를 갖는다. 다른 예시적인 실시예에서, 스펙트럼 평활화는 사람의 청각 시스템의 임계 대역 간격에 가깝도록 모델링되는 주파수 빈(frequency bin)의 그룹들에 걸쳐 수행될 수 있다. 예를 들어, 균일한 간격의 주파수 빈을 갖는 분석 필터 बैं크가 이용되는 경우, 서로 다른 수의 주파수 빈이 그룹화되고 주파수 스펙트럼의 서로 다른 부분에 대해 평균될 수 있다. 이 예시적인 실시예에서, 0부터 5 kHz까지는, 5개의 주파수 빈이 평균될 수 있거나, 5 kHz부터 10 kHz까지는, 7개의 주파수 빈이 평균될 수 있거나, 10 kHz부터 20 kHz까지는, 9개의 주파수 빈이 평균될 수 있거나, 다른 적당한 수의 주파수 빈 및 대역폭 범위가 선택될 수 있다. $H_L(F)$, $H_R(F)$, $H_C(F)$, $H_{LS}(F)$, 및 $H_{RS}(F)$ 의 평활화된 값은 평활화 유닛(708)으로부터 출력된다.

[0087] 5.1 출력 채널 각각에 대한 소스 신호 $X_L(F)$, $X_R(F)$, $X_C(F)$, $X_{LS}(F)$, 및 $X_{RS}(F)$ 가 스테레오 입력 채널의 적응적 결합으로서 발생된다. 도 7에 나타난 예시적인 실시예에서, $X_L(F)$ 는 단지 $L(F)$ 로서 제공되어 있으며, 이는 모든 주파수 대역에 대해 $G_L(F) = 1$ 임을 암시한다. 이와 마찬가지로, $X_R(F)$ 는 단지 $R(F)$ 로서 제공되어 있으며, 이는 모든 주파수 대역에 대해 $G_R(F) = 0$ 임을 암시한다. 합산기(714)로부터 출력되는 $X_C(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_C(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_C(F)$ 와 곱한 것의 합산으로서 발생된다. 합산기(720)로부터 출력되는 $X_{LS}(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_{LS}(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_{LS}(F)$ 와 곱한 것의 합산으로서 계산된다. 이와 유사하게, 합산기(726)로부터 출력되는 $X_{RS}(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_{RS}(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_{RS}(F)$ 와 곱한 것의 합산으로서 계산된다. 유의할 점은, 모든 주파수 대역에 대해 $G_C = 0.5$, $G_{LS}(F) = 0.5$ 및 $G_{RS}(F) = 0.5$ 인 경우, 종래의 매트릭스 업-믹싱 방법에서 통상적인 바와 같이, 전방 중앙 채널은 $L(F)+R(F)$ 결합으로부터 비롯되고, 서라운드 채널은 스케일링된 $L(F)-R(F)$ 결합으로부터 비롯된다는 것이다. 적응적 스케일링 신호 $G_C(F)$, $G_{LS}(F)$ 및 $G_{RS}(F)$ 는 또한 인접한 출력 채널 쌍들(이들이 가로 또는 깊이 채널 쌍인지에 상관없음) 간의 상관을 동적으로 조정하는 방법을 제공할 수 있다. 채널 소스 신호 $X_L(F)$, $X_R(F)$, $X_C(F)$, $X_{LS}(F)$, 및 $X_{RS}(F)$ 는 곱셈기(728 내지 732)에 의해 평활화된 채널 필터 $H_L(F)$, $H_R(F)$, $H_C(F)$, $H_{LS}(F)$, 및 $H_{RS}(F)$ 와 각각 곱해진다.

[0088] 곱셈기(728 내지 736)로부터의 출력은 이어서 주파수-시간 합성 유닛(738 내지 746)에 의해 주파수 영역으로부

터 시간 영역으로 변환되어 출력 채널 $Y_L(T)$, $Y_R(T)$, $Y_C(T)$, $Y_{LS}(T)$ 및 $Y_{RS}(T)$ 를 발생한다. 이와 같이, 좌측 및 우측 스테레오 신호는 5.1 채널 신호로 업-믹싱되고, 여기서 자연적으로 존재하거나, 도 1의 다운-믹싱 워터마크 프로세스 또는 다른 적당한 프로세스 등에 의해, 좌측 및 우측 스테레오 신호로 의도적으로 인코딩된 채널간 공간 큐는 시스템(700)에 의해 생성되는 5.1 채널 음장 내에서의 주파수 성분의 공간적 배치를 제어하는 데 사용될 수 있다. 이와 유사하게, 스테레오에서 4.1 사운드로, 4.1에서 5.1 사운드로, 또는 다른 적당한 조합 등의, 다른 적당한 조합의 입력 및 출력이 사용될 수 있다.

[0089] 도 8은 본 발명의 예시적인 실시예에 따른, M개 채널을 N개 채널로 데이터를 업-믹싱하는 시스템(800)을 나타낸 도면이다. 시스템(800)은 스테레오 시간 영역 데이터를 7.1 채널 시간 영역 데이터로 변환한다.

[0090] 시스템(800)은 시간-주파수 분석 유닛(802, 804), 필터 발생 유닛(806), 평활화 유닛(808), 및 주파수-시간 합성 유닛(854 내지 866)을 포함한다. 시스템(800)은, 고분해능 주파수 대역 프로세싱을 가능하게 해주는 확장가능한 주파수 영역 아키텍처의 사용을 통해, 또한 업-믹싱된 7.1 채널 신호에서의 주파수 성분의 공간적 배치를 도출하기 위해 주파수 대역별로 중요한 채널간 공간 큐를 추출 및 분석하는 필터 발생 방법을 통해, 업-믹스 프로세스에서 향상된 공간 구분 및 안정성을 제공한다.

[0091] 시스템(800)은 시간-주파수 분석 유닛(802, 804)에서 좌채널 스테레오 신호 $L(T)$ 및 우채널 스테레오 신호 $R(T)$ 를 수신하고, 이 시간-주파수 분석 유닛은 시간 영역 신호를 주파수 영역 신호로 변환한다. 이들 시간-주파수 분석 유닛은, 유한 임펄스 응답(FIR) 필터 뱅크, 직교 미러 필터(QMF) 뱅크, 이산 푸리에 변환(DFT), 시간-영역 엘리밍 소거(TDAC) 필터 뱅크, 또는 다른 적당한 필터 뱅크 등의, 적절한 필터 뱅크일 수 있다. 시간-주파수 분석 유닛(802, 804)으로부터의 출력은, 분석 필터 뱅크 서브-대역 대역폭이 심리 음향적 임계 대역(psycho-acoustic critical band), 등가 장방형 대역폭(equivalent rectangular bandwidth), 또는 어떤 다른 인지 특성에 근사하도록 처리될 수 있는 경우 0 내지 20 kHz 주파수 범위 등의, 사람의 청각 시스템의 충분한 주파수 범위를 포함하는 일련의 주파수 영역 값이다. 이와 유사하게, 다른 적당한 수의 주파수 대역 및 범위가 사용될 수 있다.

[0092] 시간-주파수 분석 유닛(802, 804)으로부터의 출력은 필터 발생 유닛(806)에 제공된다. 한 예시적인 실시예에서, 필터 발생 유닛(806)은 주어진 환경을 위해 출력되어야만 하는 채널의 수에 관한 외부 선택을 수신할 수 있다. 예를 들어, 2개의 전방 스피커 및 2개의 후방 스피커가 있는 경우 4.1 사운드 채널이 선택될 수 있거나, 2개의 전방 스피커 및 2개의 후방 스피커 및 1개의 전방 중앙 스피커가 있는 경우에, 5.1 사운드 시스템이 선택될 수 있거나, 2개의 전방 스피커, 2개의 측방 스피커, 2개의 후방 스피커, 및 1개의 전방 중앙 스피커가 있는 경우에, 7.1 사운드 시스템이 선택될 수 있거나, 다른 적당한 사운드 시스템이 선택될 수 있다. 필터 발생 유닛(806)은 주파수 대역별로 채널간 레벨 차이(ICLD) 및 채널간 코히런스(ICC) 등의 채널간 공간 큐를 추출 및 분석한다. 이들 관련 공간 큐는 이어서 업-믹싱된 음장(sound field)에서의 주파수 대역 성분의 공간 배치를 제어하는 적응 채널 필터를 발생하기 위한 파라미터로서 사용된다. 아주 빠르게 변할 수 있는 경우에 짜증나는 변동 효과를 야기할 수 있는 필터 변동성을 제한하기 위해 시간 및 주파수 둘다에 걸쳐 평활화 유닛(808)에 의해 채널 필터가 평활화된다. 도 8에 나타난 예시적인 실시예에서, 좌측 및 우측 채널 $L(F)$ 및 $R(F)$ 주파수 영역 신호가, 평활화 유닛(808)에 제공되는 7.1 채널 필터 신호 $H_L(F)$, $H_R(F)$, $H_C(F)$, $H_{LS}(F)$, $H_{RS}(F)$, $H_{LB}(F)$ 및 $H_{RB}(F)$ 를 생성하는 필터 발생 유닛(806)에 제공된다.

[0093] 평활화 유닛(808)은 7.1 채널 필터의 각각의 채널에 대한 주파수 영역 성분을 시간 및 주파수 차원 둘다에 걸쳐 평균을 구한다. 시간 및 주파수에 걸친 평활화는 채널 필터 신호의 빠른 변동을 제어하는 데 도움이 되며, 따라서 듣는 사람에게 짜증나는 것일 수 있는 지터 아티팩트 및 불안정성을 감소시킨다. 한 예시적인 실시예에서, 시간 평활화는 1차 저역-통과 필터를 현재 프레임으로부터의 각각의 주파수 대역 및 이전의 프레임으로부터의 대응하는 주파수 대역에 적용함으로써 실현될 수 있다. 이것은 프레임마다 각각의 주파수 대역의 변동성을 감소시키는 효과를 갖는다. 한 예시적인 실시예에서, 스펙트럼 평활화는 사람의 청각 시스템의 임계 대역 간격에 가깝도록 모델링되는 주파수 빈(frequency bin)의 그룹들에 걸쳐 수행될 수 있다. 예를 들어, 균일한 간격의 주파수 빈을 갖는 분석 필터 뱅크가 이용되는 경우, 서로 다른 수의 주파수 빈이 그룹화되고 주파수 스펙트럼의 서로 다른 부분에 대해 평균될 수 있다. 이 예시적인 실시예에서, 0부터 5 kHz까지는, 5개의 주파수 빈이 평균될 수 있거나, 5 kHz부터 10 kHz까지는, 7개의 주파수 빈이 평균될 수 있거나, 10 kHz부터 20 kHz까지는, 9개의 주파수 빈이 평균될 수 있거나, 다른 적당한 수의 주파수 빈 및 대역폭 범위가 선택될 수 있다. $H_L(F)$, $H_R(F)$, $H_C(F)$, $H_{LS}(F)$, $H_{RS}(F)$, $H_{LB}(F)$ 및 $H_{RB}(F)$ 의 평활화된 값은 평활화 유닛(808)으로부터 출력된

다.

[0094] 7.1 출력 채널 각각에 대한 소스 신호 $X_L(F)$, $X_R(F)$, $X_C(F)$, $X_{LS}(F)$, $X_{RS}(F)$, $X_{LB}(F)$ 및 $X_{RB}(F)$ 가 스테레오 입력 채널의 적응적 결합으로서 발생된다. 도 8에 나타난 예시적인 실시예에서, $X_L(F)$ 는 단지 $L(F)$ 로서 제공되어 있으며, 이는 모든 주파수 대역에 대해 $G_L(F) = 1$ 임을 암시한다. 이와 마찬가지로, $X_R(F)$ 는 단지 $R(F)$ 로서 제공되어 있으며, 이는 모든 주파수 대역에 대해 $G_R(F) = 0$ 임을 암시한다. 합산기(814)로부터 출력되는 $X_C(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_C(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_C(F)$ 와 곱한 것의 합산으로서 발생된다. 합산기(820)로부터 출력되는 $X_{LS}(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_{LS}(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_{LS}(F)$ 와 곱한 것의 합산으로서 계산된다. 이와 유사하게, 합산기(826)로부터 출력되는 $X_{RS}(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_{RS}(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_{RS}(F)$ 와 곱한 것의 합산으로서 계산된다. 이와 유사하게, 합산기(832)로부터 출력되는 $X_{LB}(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_{LB}(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_{LB}(F)$ 와 곱한 것의 합산으로서 계산된다. 이와 유사하게, 합산기(838)로부터 출력되는 $X_{RB}(F)$ 는 신호 $L(F)$ 를 적응적 스케일링 신호 $G_{RB}(F)$ 와 곱한 것과 $R(F)$ 를 적응적 스케일링 신호 $1-G_{RB}(F)$ 와 곱한 것의 합산으로서 계산된다. 유의할 점은, 모든 주파수 대역에 대해 $G_C = 0.5$, $G_{LS}(F) = 0.5$, $G_{RS}(F) = 0.5$, $G_{LB}(F) = 0.5$ 및 $G_{RB}(F) = 0.5$ 인 경우, 종래의 매트릭스 업-믹싱 방법에서 통상적인 바와 같이, 전방 중앙 채널은 $L(F)+R(F)$ 결합으로부터 비롯되고, 측방 및 후방 채널은 스케일링된 $L(F)-R(F)$ 결합으로부터 비롯된다는 것이다. 적응적 스케일링 신호 $G_C(F)$, $G_{LS}(F)$, $G_{RS}(F)$, $G_{LB}(F)$ 및 $G_{RB}(F)$ 는 또한 인접한 출력 채널 쌍들(이들이 가로 또는 깊이 채널 쌍인지에 상관없음) 간의 상관을 동적으로 조정하는 방법을 제공할 수 있다. 채널 소스 신호 $X_L(F)$, $X_R(F)$, $X_C(F)$, $X_{LS}(F)$, $X_{RS}(F)$, $X_{LB}(F)$ 및 $X_{RB}(F)$ 는 곱셈기(840 내지 852)에 의해 평활화된 채널 필터 $H_L(F)$, $H_R(F)$, $H_C(F)$, $H_{LS}(F)$, $H_{RS}(F)$, $H_{LB}(F)$ 및 $H_{RB}(F)$ 와 각각 곱해진다.

[0095] 곱셈기(840 내지 852)로부터의 출력은 이어서 주파수-시간 합성 유닛(854 내지 866)에 의해 주파수 영역으로부터 시간 영역으로 변환되어 출력 채널 $Y_L(T)$, $Y_R(T)$, $Y_C(T)$, $Y_{LS}(T)$, $Y_{RS}(T)$, $Y_{LB}(T)$ 및 $Y_{RB}(T)$ 를 발생한다. 이와 같이, 좌측 및 우측 스테레오 신호는 7.1 채널 신호로 업-믹싱되고, 여기서 자연적으로 존재하거나, 도 1의 다운-믹싱 워터마크 프로세스 또는 다른 적당한 프로세스 등에 의해, 좌측 및 우측 스테레오 신호로 의도적으로 인코딩된 채널간 공간 큐는 시스템(800)에 의해 생성되는 7.1 채널 음장 내에서의 주파수 성분의 공간적 배치를 제어하는 데 사용될 수 있다. 이와 유사하게, 스테레오에서 5.1 사운드로, 5.1에서 7.1 사운드로, 또는 다른 적당한 조합 등의, 다른 적당한 조합의 입력 및 출력이 사용될 수 있다.

[0096] 도 9는 본 발명의 예시적인 실시예에 따른, 주파수 영역 응용에 대한 필터를 발생하는 시스템(900)을 나타낸 도면이다. 이 필터 발생 프로세스는 M 채널 입력 신호의 주파수 영역 분석 및 프로세싱을 이용한다. M 채널 입력 신호의 각각의 주파수 대역에 대해 관련 채널간 공간 큐가 추출되고, 각각의 주파수 대역에 대해 공간 위치 벡터가 발생된다. 이 공간 위치 벡터는 이상적인 청취 조건 하에서 듣는 사람의 그 주파수 대역에 대한 인지된 소스 위치로서 해석된다. 이어서, 업-믹싱된 N 채널 출력 신호에서의 주파수 성분이 채널간 큐와 일치하게 재생되도록 각각의 채널 필터가 발생된다. 채널간 레벨 차이(ICLD) 및 채널간 코히런스(ICC)의 추정치가 공간 위치 벡터를 생성하기 위한 채널간 큐로서 사용된다.

[0097] 시스템(900)에 나타난 예시적인 실시예에서, 서브-대역 크기 또는 에너지 성분은 채널간 레벨 차이를 추정하는 데 사용되고, 서브-대역 위상각 성분은 채널간 코히런스를 추정하는 데 사용된다. 좌측 및 우측 주파수 영역 입력 $L(F)$ 및 $R(F)$ 는 크기 또는 에너지 성분 및 위상각 성분으로 변환되고, 여기서 크기/에너지 성분은 분할기(904, 906)에 의해 각각의 주파수 대역에 대한 좌측 채널 $M_L(F)$ 및 우측 채널 $M_R(F)$ 의 크기/에너지 값을 각각 정규화하기 위해 나중에 사용되는 총 에너지 신호 $T(F)$ 를 계산하는 합산기(902)에 제공된다. 정규화된 가로 좌표 신호 $LAT(F)$ 는 이어서 $M_L(F)$ 및 $M_R(F)$ 로부터 계산되고, 여기서 주파수 대역에 대한 정규화된 가로 좌표는 다음과 같이 계산된다.

[0098]
$$LAT(F) = M_L(F) * X_{MIN} + M_R(F) * X_{MAX}$$

[0099] 이와 유사하게, 정규화된 깊이 좌표는 다음과 같이 입력의 위상각 성분으로부터 계산된다.

[0100] $DEP(F) = Y_{MAX} - 0.5 * (Y_{MAX} - Y_{MIN}) * \sqrt{[\cos(\angle L(F)) - \cos(\angle R(F))]^2 + [\sin(\angle L(F)) - \sin(\angle R(F))]^2}$

[0101] 정규화된 깊이 좌표는 기본적으로 위상각 성분 $\angle L(F)$ 및 $\angle R(F)$ 간의 스케일링되고 천이된 거리 측정치로부터 계산된다. 위상각 $\angle L(F)$ 및 $\angle R(F)$ 이 단위원 상에서 서로에 가까워짐에 따라 $DEP(F)$ 의 값은 1에 가까워지고, 위상각 $\angle L(F)$ 및 $\angle R(F)$ 이 단위원의 반대쪽에 가까워짐에 따라 $DEP(F)$ 의 값은 0에 가까워진다. 각각의 주파수 대역에 대해, 정규화된 가로 좌표 및 깊이 좌표는, 이하의 도 10a 내지 도 10e에 도시된 것 등의, 2차원 채널 맵에 입력되어 각각의 채널 i 에 대한 필터 값 $H_i(F)$ 를 생성하는 2차원 벡터($LAT(F)$, $DEP(F)$)를 형성한다. 각각의 채널 i 에 대한 이들 채널 필터 $H_i(F)$ 는, 도 6의 필터 발생 유닛(606), 도 7의 필터 발생 유닛(706) 및 도 8의 필터 발생 유닛(806) 등의, 필터 발생 유닛으로부터 출력된다.

[0102] 도 10a는 본 발명의 예시적인 실시예에 따른 좌측 전방 신호에 대한 필터 맵을 나타낸 도면이다. 도 10a에서, 필터 맵(1000)은 0 내지 1의 범위에 있는 정규화된 가로 좌표 및 0 내지 1의 범위에 있는 정규화된 깊이 좌표를 받고 0 내지 1의 범위에 있는 정규화된 필터 값을 출력한다. 회색의 음영은, 필터 맵(1000)의 우측 상의 스케일로 나타내어져 있는 바와 같이, 최대 1에서 최소 0까지의 크기 변동을 나타내는 데 사용된다. 이 예시적인 좌측 전방 필터 맵(1000)에 있어서, (0, 1)에 가까워지는 정규화된 가로 및 깊이 좌표는 1.0에 가까워지는 최고 필터값을 출력하는 반면, 대략 (0.6, Y) 내지 (1.0, Y)의 범위에 있는 좌표(단, Y는 0 내지 1의 숫자임)는 기본적으로 0의 필터값을 출력한다.

[0103] 도 10b는 예시적인 우측 전방 필터 맵(1002)을 나타낸 도면이다. 필터 맵(1002)은 필터 맵(1000)과 동일한 정규화된 가로 좌표 및 정규화된 깊이 좌표를 받지만 출력 필터값은 정규화된 레이아웃의 우측 전방 부분에 유리하다.

[0104] 도 10c는 예시적인 중앙 필터 맵(1004)을 나타낸 도면이다. 이 예시적인 실시예에서, 중앙 필터 맵(1004)에 대한 최대 필터값은 정규화된 레이아웃의 중앙에서 발생하며, 좌표가 레이아웃의 전방 중앙으로부터 레이아웃의 후방쪽으로 멀어짐에 따라 상당한 크기 저하가 있다.

[0105] 도 10d는 예시적인 좌측 서라운드 필터 맵(1006)을 나타낸 도면이다. 이 예시적인 실시예에서, 좌측 서라운드 필터 맵(1006)에 대한 최대 필터값은 정규화된 레이아웃의 후방 좌측 좌표 근방에서 발생하며, 좌표가 레이아웃의 전방 우측으로 이동함에 따라 크기 저하가 있다.

[0106] 도 10e는 예시적인 우측 서라운드 필터 맵(1008)을 나타낸 도면이다. 이 예시적인 실시예에서, 우측 서라운드 필터 맵(1008)에 대한 최대 필터값은 정규화된 레이아웃의 후방 우측 좌표 근방에서 발생하며, 좌표가 레이아웃의 전방 좌측으로 이동함에 따라 크기 저하가 있다.

[0107] 이와 마찬가지로, 다른 스피커 레이아웃 또는 구성이 사용되는 경우, 기존의 필터 맵이 수정될 수 있고, 새로운 청취 환경에서의 변화를 반영하기 위해 새로운 스피커 위치에 대응하는 새로운 필터 맵이 발생될 수 있다. 한 예시적인 실시예에서, 7.1 시스템은, 좌측 서라운드 및 우측 서라운드가 깊이 좌표 차원에서 위쪽으로 이동해 있고 좌측 후방 및 우측 후방 위치가 필터 맵(1006, 1008)과 유사한 필터 맵을 각각 갖는, 2개의 부가적인 필터 맵을 포함하게 된다. 필터 계수가 감소하는 비율은 다른 수의 스피커에 대응하기 위해 변경될 수 있다.

산업상 이용 가능성

[0108] 본 발명의 시스템 및 방법의 예시적인 실시예가 본 명세서에 상세히 기술되어 있지만, 당업자라면 첨부된 청구 범위의 범위 및 정신을 벗어나지 않고 본 시스템 및 방법에 여러가지 치환 및 수정이 행해질 수 있다는 것도 잘 알 것이다.

도면의 간단한 설명

[0011] 도 1은 본 발명의 예시적인 실시예에 따른, 분석 및 정정 루프를 갖는 동적 다운-믹싱 시스템을 나타낸 도면.

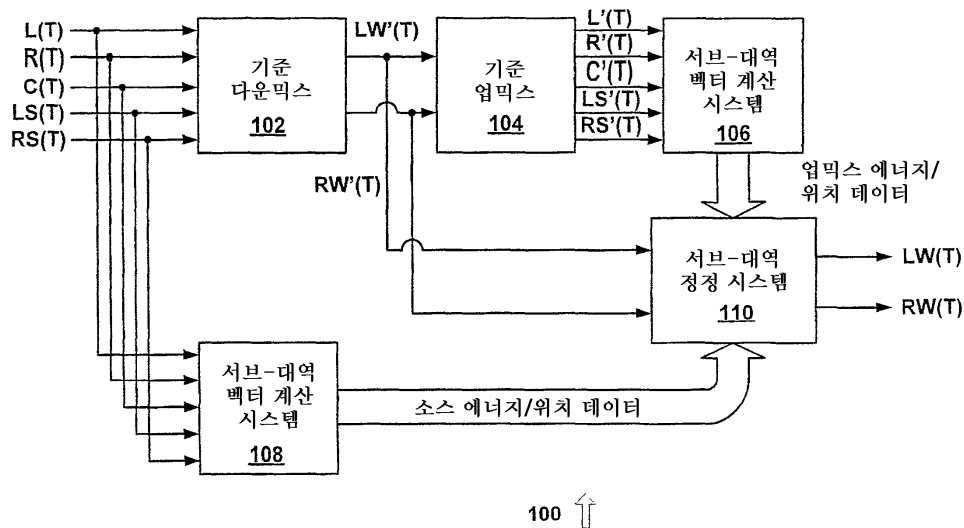
[0012] 도 2는 본 발명의 예시적인 실시예에 따른, N개 채널로부터 M개 채널로 데이터를 다운-믹싱하는 시스템을 나타낸 도면.

[0013] 도 3은 본 발명의 예시적인 실시예에 따른, 5개 채널로부터 2개 채널로 데이터를 다운-믹싱하는 시스템을 나타낸 도면.

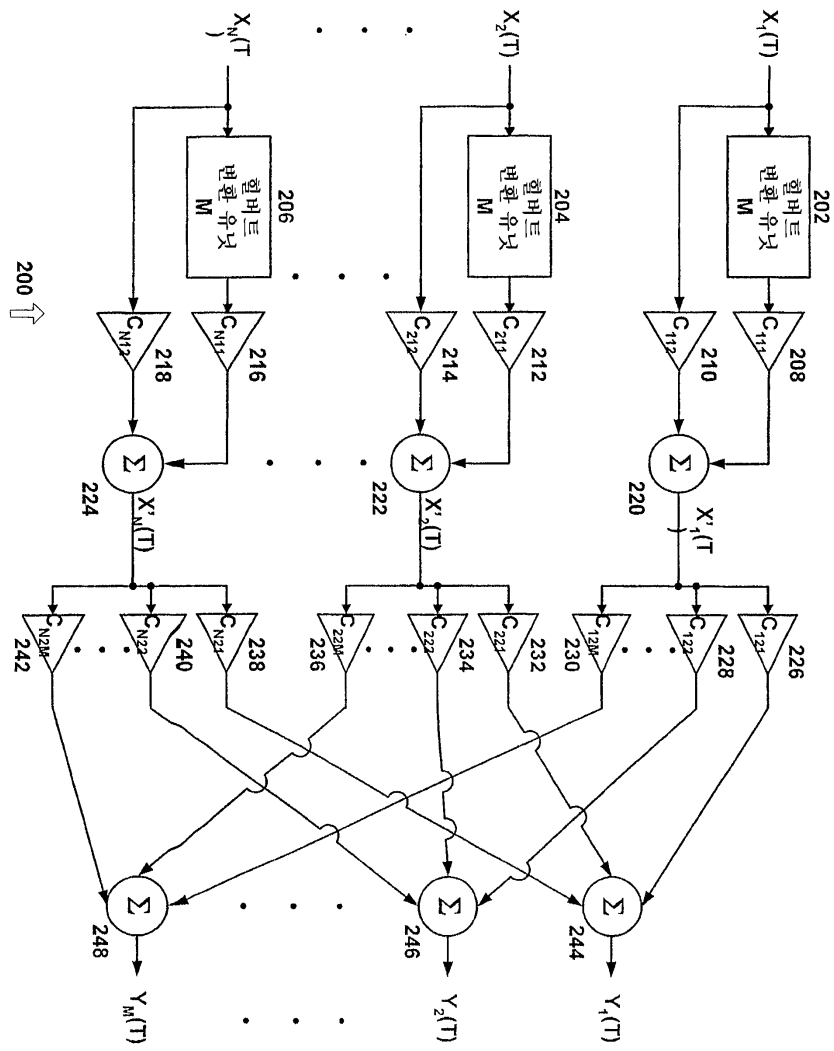
- [0014] 도 4는 본 발명의 예시적인 실시예에 따른 서브-대역 벡터 계산 시스템을 나타낸 도면.
- [0015] 도 5는 본 발명의 예시적인 실시예에 따른 서브-대역 정정 시스템을 나타낸 도면.
- [0016] 도 6은 본 발명의 예시적인 실시예에 따른, M개 채널로부터 N개 채널로 데이터를 업-믹싱하는 시스템을 나타낸 도면.
- [0017] 도 7은 본 발명의 예시적인 실시예에 따른, 2개 채널로부터 5개 채널로 데이터를 업-믹싱하는 시스템을 나타낸 도면.
- [0018] 도 8은 본 발명의 예시적인 실시예에 따른, 2개 채널로부터 7개 채널로 데이터를 업-믹싱하는 시스템을 나타낸 도면.
- [0019] 도 9는 본 발명의 예시적인 실시예에 따른, 채널간 공간 큐를 추출하고 주파수 영역 응용을 위한 공간 채널 필터를 발생하는 방법을 나타낸 도면.
- [0020] 도 10a는 본 발명의 예시적인 실시예에 따른 예시적인 좌전방 채널 필터 맵을 나타낸 도면.
- [0021] 도 10b는 예시적인 우전방 채널 필터 맵을 나타낸 도면.
- [0022] 도 10c는 예시적인 중앙 채널 필터 맵을 나타낸 도면.
- [0023] 도 10d는 예시적인 서라운드 좌채널 필터 맵을 나타낸 도면.
- [0024] 도 10e는 예시적인 서라운드 우채널 필터 맵을 나타낸 도면.

도면

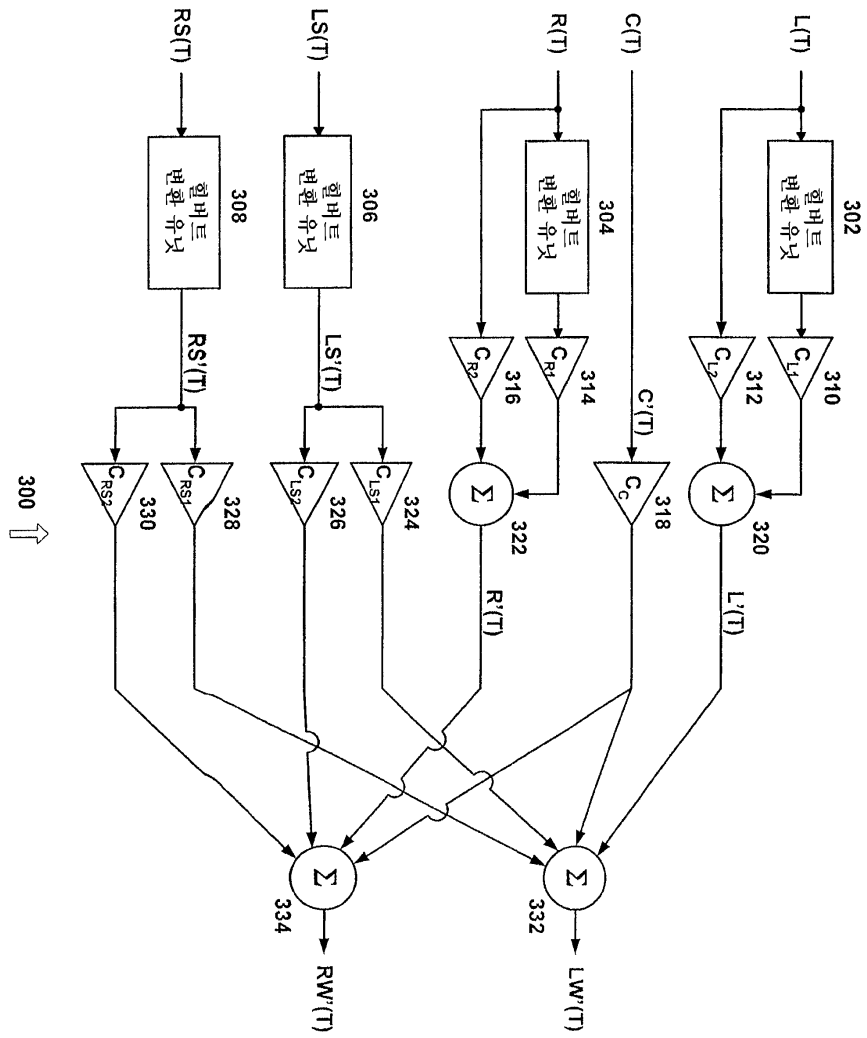
도면1

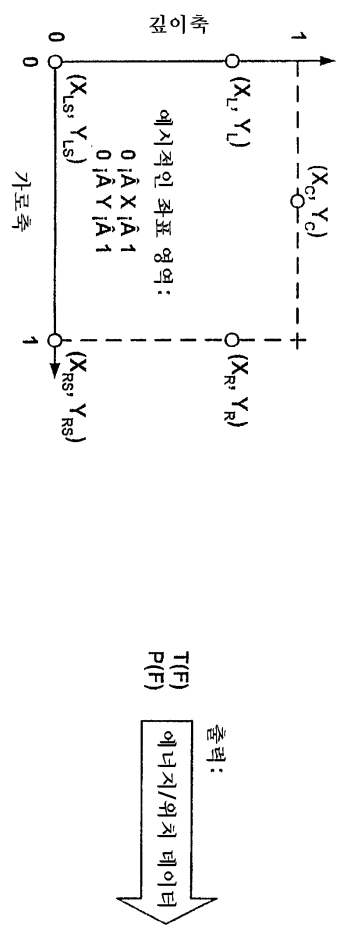
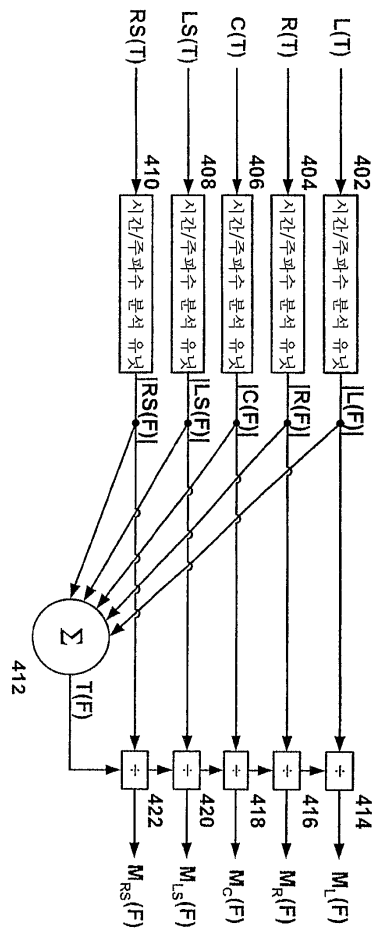


도면2

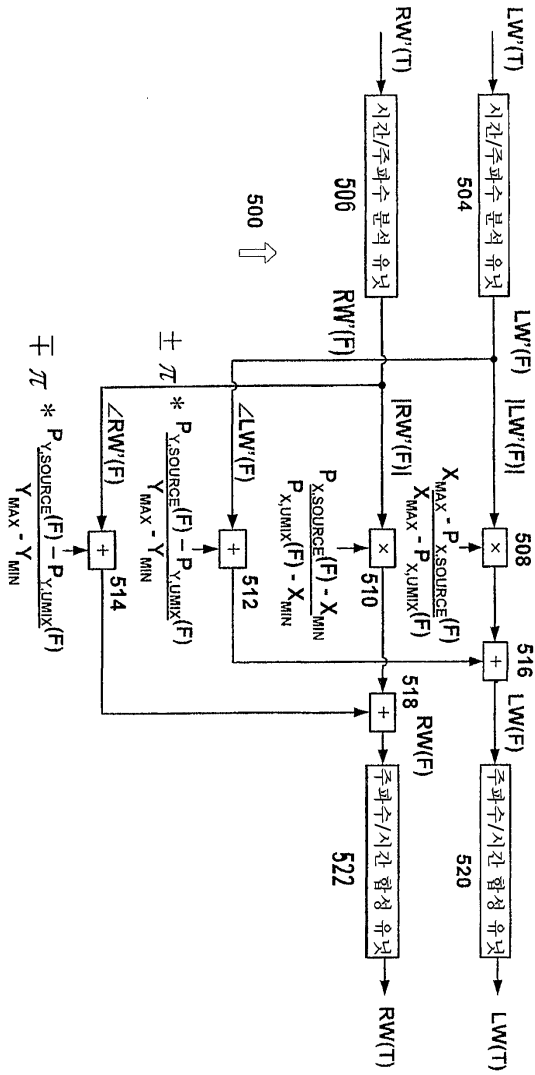
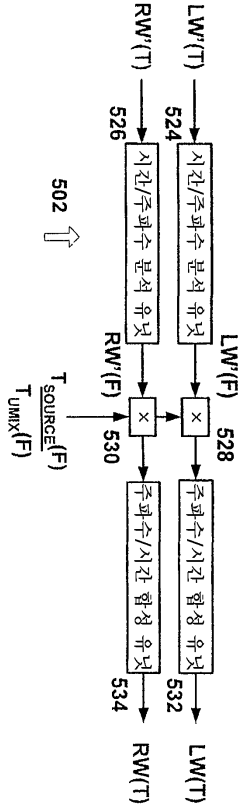


도면3



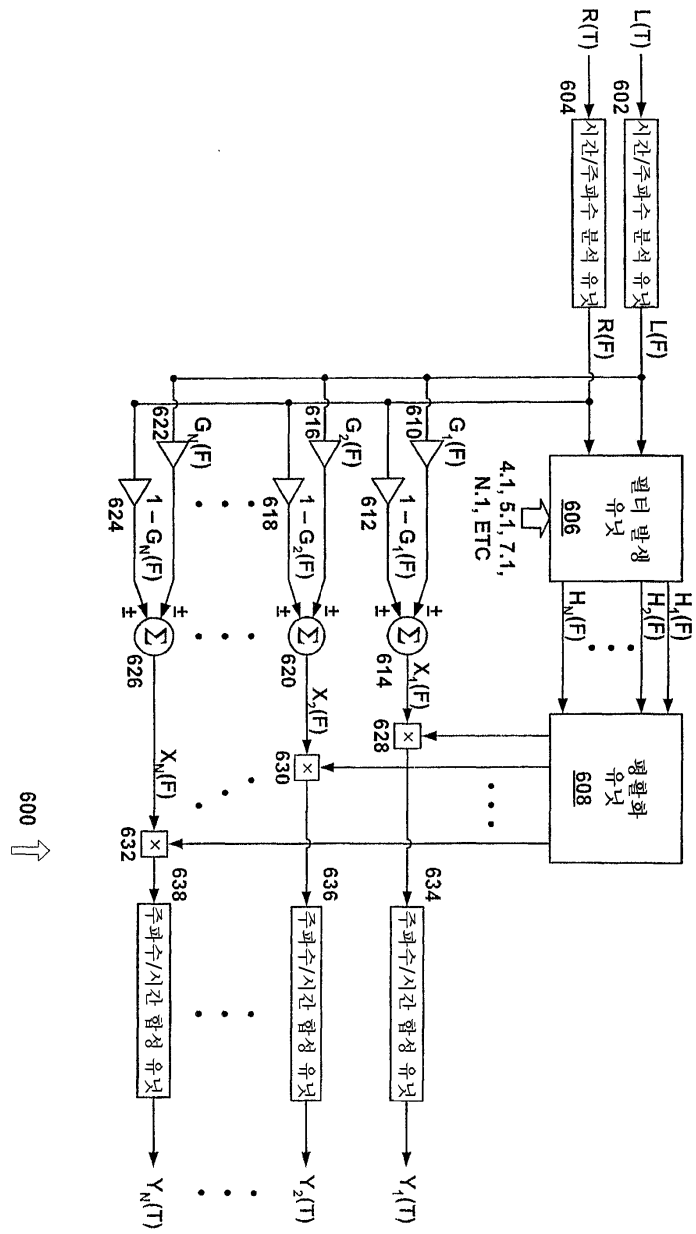


도면4

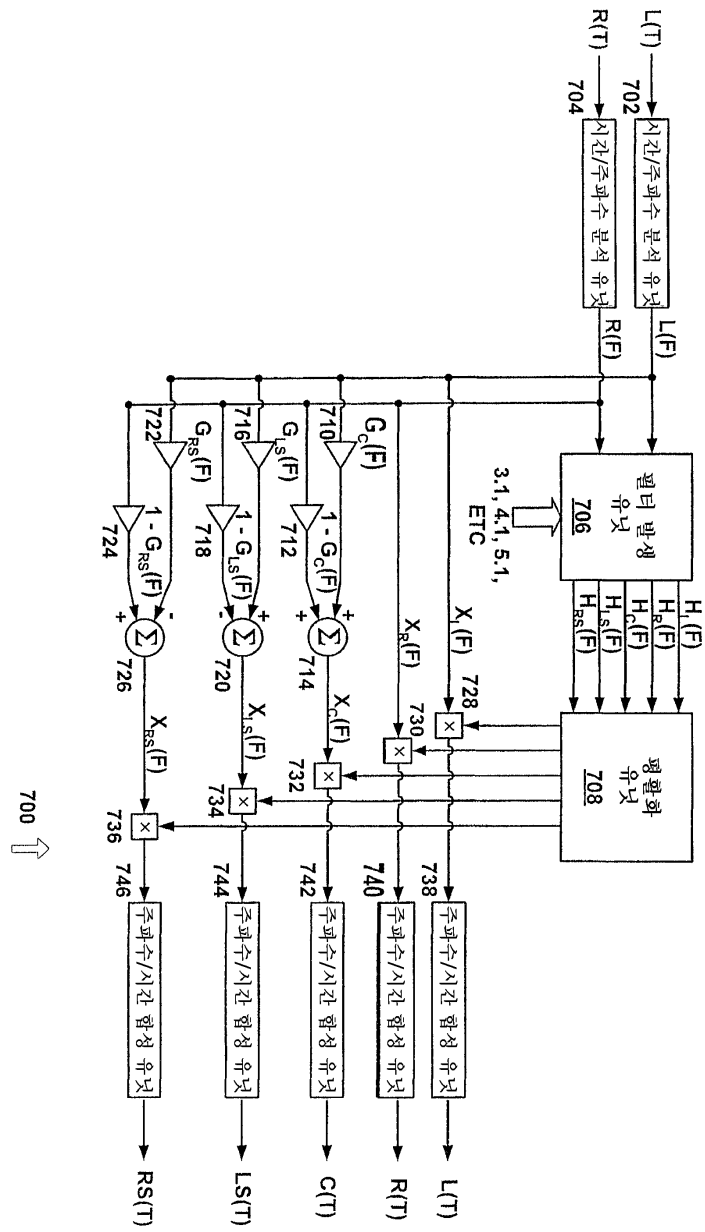


도면5

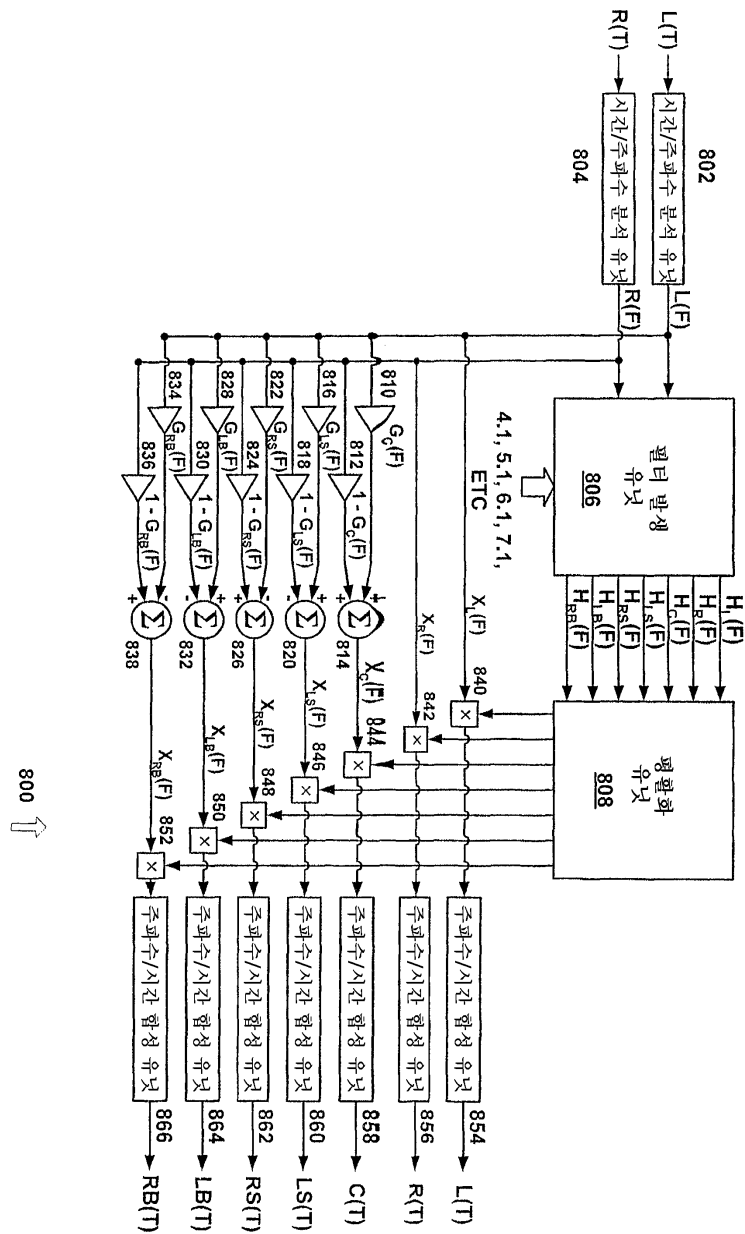
도면6



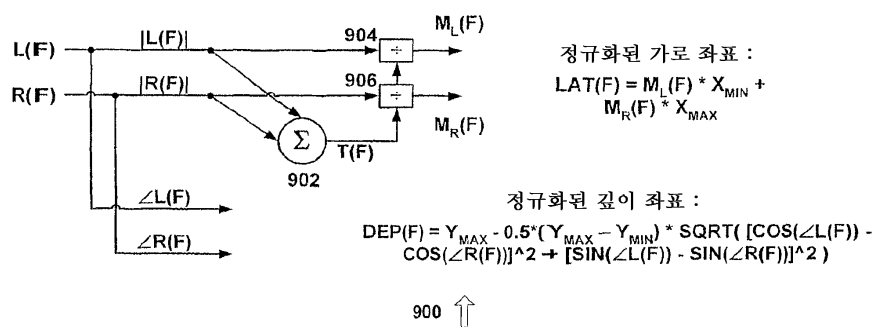
도면7



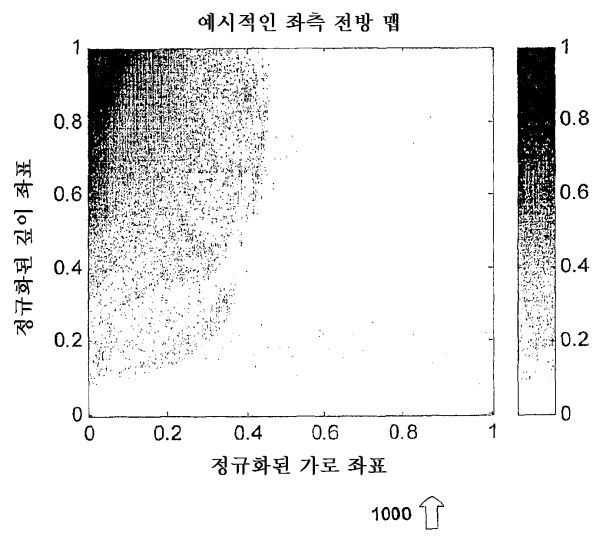
도면8



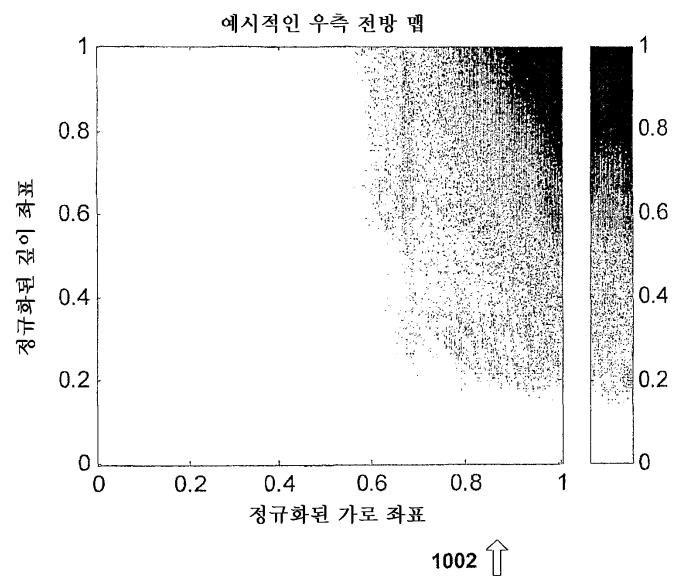
도면9



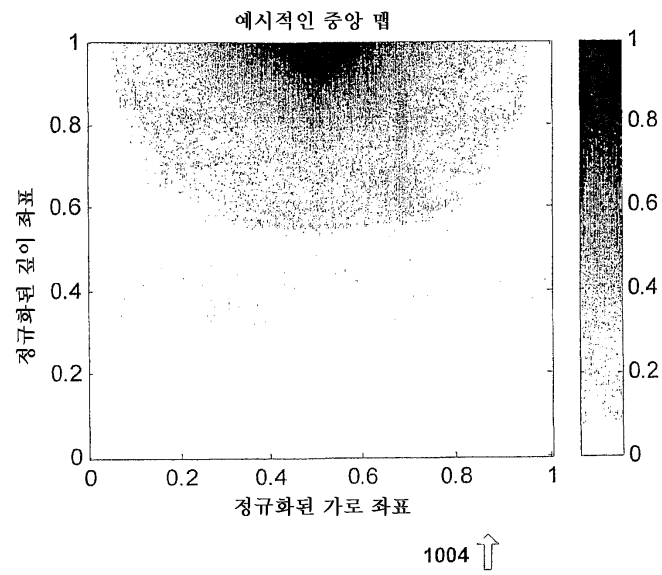
도면10a



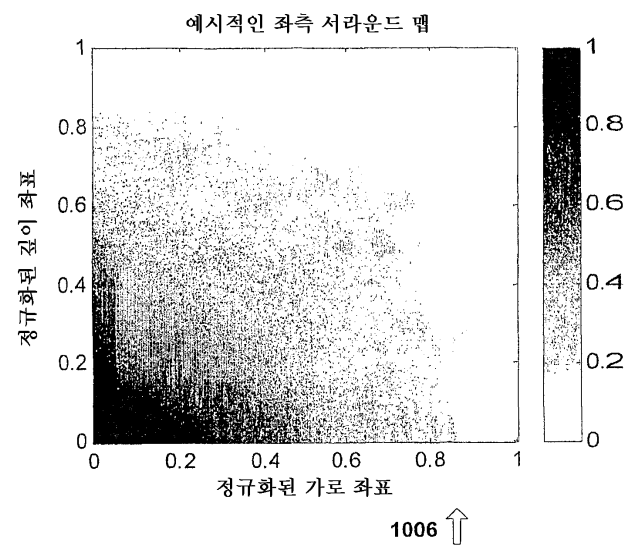
도면10b



도면10c



도면10d



도면10e

