(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2013-0133242
(43) 공개일자 2013년12월06일

(51) 국제특허분류(Int. Cl.)
*H04S 3/00* (2006.01)
(21) 출원번호 10-2013-7017810
(22) 출원일자(국제) 2011년11월22일
심사청구일자 없음
(85) 번역문제출일자 2013년07월08일
(86) 국제출원번호 PCT/EP2011/070702
(87) 국제공개번호 WO 2012/076332
국제공개일자 2012년06월14일
(30) 우선권주장
11165742.5 2011년05월11일
유럽특허청(EPO)(EP)
61/421,927 2010년12월10일 미국(US)

(71) 출원인
**프라운호퍼 게젤샤프트 쭈르 푀르데룽 데어 안겐반텐 포르슝 에. 베.**
독일 80686 뮌헨 한자슈트라쎄 27 체
(72) 발명자
**왈터, 안드레아스**
스위스, 씨에이치-1023 크리시어, 슈민 디 타우트-벤트 2
(74) 대리인
**특허법인이상**

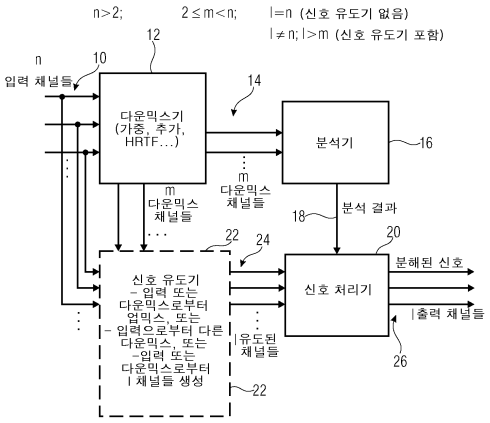
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **다운믹스기를 이용한 입력 신호 분해 장치 및 방법**

## (57) 요 약

적어도 세 개의 채널들을 가지는 신호를 분해하는 장치는 적은 채널들의 수를 가지는 다운믹스된 신호를 획득하기 위해 입력 신호들을 다운믹싱하는 다운믹스기(12)를 포함한다. 추가로, 분석 결과를 유도하기 위해 상기 다운믹스된 신호를 분석하는 분석기(16)가 제공되며, 그리고 분석 결과(18)는 분해된 신호(26)를 획득하기 위해 입력 신호로부터 유도된 신호 또는 상기 입력 신호를 처리하는 신호 처리기(20)에 제공된다.

*대 표 도* - 도1



## *특허청구의 범위*

**청구항 1**

적어도 세 개의 입력 채널들을 가지는 입력 신호(10)를 분해하는 장치로서,

다운믹스 신호를 획득하기 위한 입력 신호를 다운믹싱하며, 다운믹스된 신호(14)의 다수의 다운믹스 채널들이 적어도 2개이고 입력 채널들의 수보다 적도록 다운믹싱하도록 구성된, 다운믹스기(12);

분석 결과(18)를 유도하기 위해 상기 다운믹스된 신호를 분석하는 분석기(16); 및

상기 분석 결과(18)를 사용하여 상기 입력 신호(10) 또는 상기 입력 신호로부터 유도된 신호(24), 또는 상기 입력 신호가 유도된 신호를 처리하며, 분해된 신호(26)를 획득하기 위해 상기 분석 결과를 상기 입력 신호로부터 유도된 신호의 채널들 또는 상기 입력 신호의 상기 입력 채널들에 적용하도록 구성된, 신호 처리기(processor)(20)를 포함하는, 입력 신호를 분해하는 장치.

**청구항 2**

청구항 1에 있어서,

상기 입력 채널들을 채널 주파수 표현들(representations)(각각의 입력 채널 주파수 표현은 다수의 서브밴드들(subbands)을 가짐)의 시간 시퀀스(sequence)로 변환하는 시간/주파수 변환기(32)를 더 포함하며, 또는 상기 다운믹스기(12)가 상기 다운믹스된 신호를 변환하기 위한 시간/주파수 변환기를 포함하며,

상기 분석기(16)는 개별의 서브밴드들을 위한 분석 결과(18)를 생성하도록 구성되고,

상기 신호 처리부(20)는 개별의 분석 결과들을 상기 입력 신호로부터 유도된 상기 신호 또는 상기 입력 신호의 상응하는 서브밴드들에 적용하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 3**

청구항 1 또는 2에 있어서,

상기 분석기(16)는 상기 분석 결과로서 가중치 요소들 $(W(m, i))$ 을 생성하도록 구성되고,

상기 신호 처리기(20)는 상기 가중치 요소들을, 상기 가중치 요소들로 가중함에 의해 상기 입력 신호로부터 유도된 상기 신호 또는 상기 입력 신호에 적용하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 4**

청구항 1 내지 3 중 어느 하나에 있어서,

상기 다운믹스기는 적어도 두 개의 다운믹스 채널들이 서로 다르도록 다운믹스 규칙에 따라 가중된 또는 가중되지 않은 입력 신호들을 추가하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 5**

청구항 1 내지 4 중 어느 하나에 있어서,

상기 다운믹스기(12)는 실내 임펄스(impulse) 응답들-기반 필터들(filters), 바이노럴 실내 임펄스 응답들(binaural room impulse responses, BRIR)-기반 필터들 또는 머리-관련 전달 함수(head-related transfer function, HRTF)-기반 필터들을 사용하여 상기 입력 신호(10)를 필터링하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 6**

청구항 1 내지 5 중 어느 하나에 있어서,

상기 처리기(20)는 위너(Wiener) 필터를 상기 입력 신호로부터 유도된 상기 신호 또는 상기 입력 신호에 적용하

도록 구성되고,

상기 분석기(16)는 상기 다운믹스 채널들로부터 유도된 기대 값들을 사용하여 상기 위너 필터를 계산하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 7**

청구항 1 내지 6 중 어느 하나에 있어서,

상기 입력 신호로부터 유도된 상기 신호가 상기 다운믹스 신호 또는 상기 입력 신호에 비해 다른 채널들의 수를 가지도록 상기 입력 신호로부터 상기 신호를 유도하는 신호 유도기(22)를 더 포함하는, 입력 신호를 분해하는 장치.

**청구항 8**

청구항 1 내지 7 중 어느 하나에 있어서,

상기 분석기(20)는 미리 알려진 참조 신호들에 의해 생성 가능한 두 개의 신호들 간의 주파수-종속 유사도를 나타내는 미리-저장된 주파수-종속 유사도 곡선을 사용하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 9**

청구항 1 내지 8 중 어느 하나에 있어서,

상기 분석기는, 신호들이 알려진 유사도 특성을 가지고 신호들이 알려진 라우드스피커(loudspeaker) 위치들에서 라우드스피커들에 의해 발생 가능하다는 가정하에서, 청취자 위치에서 둘 이상의 신호들 간의 주파수-종속 유사도를 나타내는 미리-저장된 주파수-종속 유사도 곡선을 사용하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 10**

청구항 1 내지 7 중 어느 하나에 있어서,

상기 분석기는 상기 입력 채널들의 주파수-종속 짧은-시간 파워를 사용하여 단일-종속 주파수-종속 유사도 곡선을 계산하도록 구성되는, 입력 신호를 분해하는 장치.

**청구항 11**

청구항 8 내지 10 중 어느 하나에 있어서,

상기 분석기(16)는, 주파수 서브밴드에서 상기 다운믹스된 채널의 유사도를 계산하고(80), 참조 곡선에 의해 표현된 유사도를 가지는 유사도 값을 비교하고(82, 83), 상기 분석결과로서 압축의 결과를 기반으로 가중치 요소를 생성하도록 구성되고,

또는 상기 동일한 주파수 서브밴드를 위한 상기 참조 곡선에 의해 표현된 유사도 및 상응하는 결과 간의 거리를 계산하고 상기 분석 결과로서 상기 거리를 기반으로 가중치 요소를 추가로 계산하도록 구성된, 입력 신호를 분해하는 장치.

**청구항 12**

청구항 1 내지 11 중 어느 하나에 있어서,

상기 분석기(16)는 사람의 귀에 대한 주파수 해상도(resolustion)에 의해 결정된 서브밴드들에서 상기 다운믹스 채널들을 분석하도록 구성된, 입력 신호를 분해하는 장치.

**청구항 13**

청구항 1 내지 12 중 어느 하나에 있어서,

상기 분석기(16)는 다이렉트(direct) 엠비언스(ambience) 분해를 허용하는 분석 결과를 생성하기 위해 상기 다운믹스된 신호를 분석하도록 구성되고,

상기 신호 처리기(20)는 상기 분석 결과를 사용하여 다이렉트 부분 또는 엠비언스 부분을 추출하도록 구성된,

입력 신호를 분해하는 장치.

**청구항 14**

적어도 세 개의 입력 채널들(channels)을 가지는 입력 신호(10)를 분해하는 방법으로서,

다운믹스된(downmixed) 신호(14)에 대한 다수의 다운믹스 채널들이 적어도 2 및 입력 채널들의 수보다 적도록, 다운믹스 신호를 획득하기 위해 상기 입력 신호를 다운믹싱하는(12) 단계;

분석 결과(18)를 유도하기 위해 상기 다운믹스된 신호를 분석하는(16) 단계; 및

상기 분석 결과(18)를 사용하여, 상기 입력 신호(10) 또는 상기 입력 신호로부터 유도된 신호(24) 또는 상기 입력 신호가 유도된 신호를 처리하는(20) 단계를 포함하며,

상기 분석 결과는 분해된 신호(26)를 획득하기 위해 상기 입력 신호로부터 유도된 신호의 채널들 또는 상기 입력 신호의 상기 입력 채널들에 적용되는, 입력 신호를 분해하는 방법.

**청구항 15**

컴퓨터(computer) 또는 프로세서(processor)에 의하여 실행될 때, 청구항 14의 방법을 수행하는 컴퓨터 프로그램.

# 명 세 서

## 기 술 분 야

[0001] 본 발명은 오디오 처리에 관한 것으로, 특히 지각적으로 구별되는 요소들과 같은 다른 요소들로 오디오를 신호 분해하는 처리에 관한 것이다.

## 배 경 기 술

[0002] 인간의 청각(auditory) 시스템은 모든 방향으로부터 사운드(sound)를 감지한다. 지각된 청각(사운드라는 단어가 물리적 현상을 묘사하는 데 사용되는 반면, 청각이라는 용어는 지각되는 것을 나타냄) 환경은 발생한 사운드 이벤트 및 주변 공간의 음향(acoustic) 특성의 인상을 생성한다. 특정 사운드 필드(field)에서 지각된 청각 인상은 (적어도 부분적으로) 귀의 입구에서 신호들의 상응하는 세 가지의 다른 유형들로 모델링(modeling) 될 수 있다. 신호들의 상응하는 세 가지의 다른 유형들은 다이렉트(direct) 사운드, 초기 반사들(early reflections) 및 확산 반사들(diffuse reflections)이다. 이들 신호들은 지각된 청각 공간 이미지의 구성에 기여한다.

[0003] 다이렉트 사운드는 방해 없이 사운드 소스(source)로부터 청취자(listener)에게 직접 처음으로 도달하는 각 사운드 이벤트의 파동들을 나타낸다. 이것은 사운드 소스에 대한 특성(characteristic)이고 사운드 이벤트의 입사 방향에 관한 최소 손상된(least-compromised) 정보를 제공한다. 평면에서 사운드 소스의 방향의 추정을 위한 주요 큐들(cues)은 두 귀의 시간차(interaural time differences, ITDs)와 두 귀의 레벨차(interaural level differences, ILDs)로 불리는 좌측 및 우측 귀 입력 신호들 간의 차이들이다. 다음에, 다이렉트 사운드의 다수의 반사들이 다른 상대적 시간 지연 및 레벨들을 가지고 다른 방향들로부터 귀들에 도착한다. 다이렉트 사운드에 상대적으로, 시간 지연이 증가할수록 반사들의 밀도는 반사들이 통계적 클러터(statistical clutter)를 형성할 때까지 증가한다.

[0004] 반사된 사운드는 거리 지각(distance perception)과 적어도 두 개의 요소들로 구성된 청각 공간 인상(auditory spatial impression)에 기여한다. 두 개의 요소들은 어페런트 소스 너비(apparent source width, ASW)(ASW를 위해 일반적으로 사용되는 다른 용어는 청각 공간감(auditory spaciousness)) 및 청취자 인벨로프먼트(listener envelopment, LEV)이다. ASW는 사운드 소스의 어페런트 너비의 확대로 정의되고, 주로 초기 측면 반사들(early lateral reflections)에 의해 결정된다. LEV는 사운드로 포위되는 청취자의 감각(sense)을 의미하며, 주로 지연-도착하는 반사들(late-arriving reflections)에 의해 결정된다. 전기음향 스테레오 사운드 재생(electroacoustic stereophic sound reproduction)의 목표는 만족스러운 청각 공간 이미지의 지각을 일으키는 것이다. 이는 자연적 또는 건축적 참조(예를 들어, 홀 내에서의 콘서트 녹음)가 될 수 있거나, 또는 현실에 존재하지 않는 사운드 필드일 수 있다(예를 들어, 전기음향(electroacoustic) 음악).

[0005] 콘서트 홀 음향의 필드로부터, 주관적으로 만족스런 사운드 필드를 획득하기 위해, 중요한 부분인 LEV와 함께

청각 공간 인상의 강한 감각은 중요하다고 알려져 있다. 확산 사운드 필드를 재생하는 수단에 의해 포위하는 사운드 필드를 재생하는 라우드스피커 설정의 능력이 관심사가 된다. 합성 사운드 필드에서, 전용 변환기들(dedicated transducers)을 사용하여 자연적으로 발생하는 모든 반사들을 재생하는 것은 불가능하다. 이것은 확산 이후 반사들의 경우 특히 사실이다. 확산 반사들의 타이밍 및 레벨 특성은 라우드스피커들 피드(feeds)와 같은 "잔향된(reverberated)" 신호들을 사용하여 시뮬레이션(simulation) 될 수 있다. 충분히 비상관된다면, 재생에 사용되는 라우드스피커들의 숫자와 위치는 사운드 필드가 확산된 것으로 지각되는 지를 결정한다. 변환기들(transducers)의 이산 숫자만을 사용하여 연속적인 확산 사운드 필드를 지각하는 것이 목적이다. 사운드의 방향이 예측될 수 없고 특히 하나의 변환기 위치도 정해질 수 없는 상황에서 사운드 필드를 생성하는 것이다. 합성 사운드 필드의 주관적인 확산도는 주관적인 시험들에서 평가된다.

[0006] 스테레오(stereophonic) 사운드 재생들은 변환기(transducer)의 이산 숫자만 사용하여 연속적인 사운드 필드의 지각을 일으키는 것을 목적으로 한다. 가장 요구되는 기능들은 위치된(localized) 소스들의 방향 안정성과 서라운딩(surrounding) 청각 환경의 현실적인 렌더링(rendering)이다. 현재 스테레오 녹음의 저장 또는 전송에 주로 사용되는 형태는 채널 기반이다. 각 채널은 특정 위치에서 연관된 라우드스피커(loudspeaker)를 통해 재생되도록 의도된 신호를 전송한다. 특정 청각 이미지는 녹음 또는 믹싱(mixing) 과정에서 설계된다. 이 이미지는 재생에 사용되는 라우드스피커의 설정이 녹음을 위해 설계된 목표 설정과 유사하면 정확하게 재생된다.

[0007] 가능한 전송 및 재생 채널들의 수는 지속적으로 증가하고 모든 새로운 오디오 재생 형식과 함께 실제 재생 시스템을 통한 기존 형식의 콘텐츠의 렌더링을 위한 요구가 나타난다. 업믹스 알고리즘들은 이러한 요구에 대한 해결방안으로, 기존 신호로부터 더 많은 채널들을 가지고 신호를 계산한다. 다수의 스테레오 업믹스 알고리즘들이 문헌들로 제안되어 왔으며, 예를 들어 카를로스 아벤다노(Carlos Avendano)와 진-마크 잣(Jean-Marc Jot)의 "다중 채널 업믹스의 주파수 도메인 접근(A frequency-domain approach to multichannel upmix)", 오디오 엔지니어링 학회 저널(Journal of the Audio Engineering Society), vol. 52, no. 7/8, pp. 740-749, 2004; 크리스토프 폴러, "스테레오 신호들의 다중-라우드스피커 재생(Multiple-loudspeaker playback of stereo signals)", 오디오 엔지니어링 학회 저널(Journal of the Audio Engineering Society), vol. 54, no. 11, pp. 1051-1064, November 2006; 존 어셔랜드 제이콥 베네스티, "공간 사운드 품질의 향상 : 새로운 잔향-추출 오디오 업믹서(Enhancement of spatial sound quality : A new reverberation-extraction audio upmixer)", 오디오, 음성 및 언어처리 국제 전기전자학회 트랜잭션(IEEE Transactions on Audio, Speech, and Language Processing) , vol. 15, no. 7, pp. 2141-2150, 2007년 9월이 있다. 이들 알고리즘들의 대부분은 목표 라우드스피커 설정에 적응된 랜더링에 따른 다이렉트/엠비언트(direct/ambient) 신호 분해를 기반으로 한다.

[0008] 기술된 다이렉트/엠비언트 신호 분해는 다중-채널 서라운드 신호들에 바로 적용할 수 없다. N 오디오 채널들로부터 상응하는 N 다이렉트 사운드와 N 엠비언트 사운드 채널들을 획득하기 위한 필터링 및 신호 모델의 공식화는 쉽지 않다. 다이렉트 사운드가 모든 채널과 상관 관계에 있는 것을 가정하는, 예를 들어, 크리스토프 폴러의 "스테레오 신호들의 다중-라우드스피커 재생(Multiple-loudspeaker playback of stereo signals)", 오디오 엔지니어링 학회 저널(Journal of the Audio Engineering Society), vol. 54, no. 11, pp. 1051-1064, 2006년 11월,에서 볼 수 있는, 스테레오 경우에서 사용되는 간단한 신호 모델은 서라운드 신호 채널들 간에 존재할 수 있는 채널 관계들의 다양성을 캡쳐(capture)할 수 없다.

[0009] 스테레오 사운드 재생의 일반적인 목적은 제한된 수의 변환기들 및 전송 채널들만을 이용하여 연속적인 사운드 필드의 지각을 일으키는 것이다. 두 개의 라우드스피커들은 공간 사운드 재생을 위한 최소 요구 사항이다. 현대의 소비자 시스템들은 종종 많은 수의 재생 채널들을 제공한다. 기본적으로, 스테레오 신호들(채널들의 수와 무관한)은 각 소스에 대하여 다이렉트 사운드가 어페런트 소스 너비와 청취자 인벨로프먼트를 결정하는 큐들(cues)인 특정 방향 큐들과 반사된 독립 신호들을 가진 다수의 채널들에 코히어런트(=종속)하도록 녹음되거나 믹스된다. 의도된 청각 이미지의 정확한 지각은 일반적으로 녹음이 의도되었던 재생 설정에서 이상적인 관측점에서만 가능하다. 주어진 라우드스피커 설정에 추가의 스피커들을 더하면 일반적으로 자연스런 사운드 필드의 더 현실적인 재구성/시뮬레이션이 가능하다. 입력 신호들이 다른 형식으로 주어지는 경우 확장된 라우드스피커 설정의 전체 이점의 사용 또는 입력 신호의 지각적으로 구별되는 부분들의 조작은, 개별적으로 접근 가능해야 한다. 이 규격은 임의의 수의 입력 채널들을 포함하는 스테레오 녹음들의 종속 및 독립 요소들을 분리하는 방법을 아래에 설명한다.

[0010] 오디오 신호들을 지각적으로 구별되는 요소들로 분해하는 것은 고품질 신호 수정, 향상, 적응 재생 그리고 지각 코딩에 필요하다. 두-채널 입력 신호들로부터 지각 구별 신호 요소들의 추출 및/또는 조작을 허용하는 다수의 방법이 최근에 제안되어 왔다. 두 채널들 이상을 가지는 입력 신호들이 점점 일반화되면서, 기술된 조작방법들

은 다중채널 입력 신호들을 위해 또한 바람직하다. 그러나, 두 채널의 입력을 위해 기술된 대부분의 개념은 임의의 채널 수를 가진 입력 신호들에 확장 적용이 쉽지 않다.

[0011] 예를 들어, 좌측 채널, 중심 채널, 우측 채널, 좌측 서라운드 채널, 우측 서라운드 채널, 그리고 저주파 개선(서브우퍼(subwoofer))을 가지는 5.1 채널 서라운드 신호를 이용해 다이렉트 및 엠비언스 부분들로 신호 분석을 수행해야 하는 경우, 다이렉트/엠비언스 신호 분석을 적용하는 방법은 간단하지 않다. 결국 최대 15개의 다른 비교 동작들을 가지는 계층적 처리를 야기하는 여섯 개 채널들의 각 쌍의 비교를 고려할 수 있다. 그 후, 각 채널이 모든 다른 채널들과 비교되는 이들 15개의 비교 동작들이 모두 수행된 후에, 15개의 결과를 어떻게 평가할 것인지 결정할 수 있다. 이것은 시간 소비적이고, 결과는 해석하기 어려우며, 상당한 양의 처리 자원들 때문에 예를 들어, 다이렉트/엠비언스 분리의 실시간 응용들 또는, 일반적인, 예를 들어 업믹스 콘텍스트 또는 다른 오디오 처리 동작들에 사용될 수 있는, 신호 분해에 사용할 수 없다.

[0012] 2007년 ICASSP 2007 학회의 M.M. Goodwin과 J.M.Jot의 "공간 오디오 코딩 및 향상을 위한 벡터-기반 지역화 및 우선-주변 신호 분해(Primary-ambient signal decomposition and vector-based localization for spatial audio coding and enhancement)"에서, 주요 요소 분석(principal component analysis)이 주요(=다이렉트) 및 엠비언트 신호 분해를 수행하기 위한 입력 채널 신호들에 적용된다.

[0013] 오디오 엔지니어링 학회 저널(Journal of the Audio Engineering Society), vol. 54, no. 11, pp. 1051-1064, 2006년 11월의 크리스토프 폴러"스테레오 신호들의 다중-라우드스피커 재생(Multiple-loudspeaker playback of stereo signals)"과 프리프린트 123회 컨벤션 오디오 엔지니어링 학회, 2007년 10월의 크리스토프 폴러, "고도 지향성 2-캡슐 기반 마이크로폰 시스템(A highly directive 2-capsule based microphone system)"에서 사용된 모델은 스테레오 및 마이크로폰 신호들에서 비-상관 또는 일부 상관된 확산 신호를 각각 가정한다. 논문들은 이 가정에서 주어진 확산/엠비언트 신호 추출을 위한 필터들을 유도한다. 이 접근법들은 하나 및 두 채널의 오디오 신호들에 한정된다.

[0014] 추가 참조 논문으로 2004년 vol.52, no 7/8, pp.740-749, 오디오 엔지니어링 학회, C. 아벤다노와 J.-M. Jot의 "다중 채널 업믹스로의 주파수 도메인 접근(A frequency-domain approach to multichannel upmix)"이 있다. 2007년 ICASSP 2007 학회의 M.M. 굿윈과 J.M.Jot의 참조 논문 "공간 오디오 코딩 및 향상을 위한 벡터-기반 지역화 및 우선-주변 신호 분해(Primary-ambient signal decomposition and vector-based localization for spatial audio coding and enhancement)"에서 아벤다노와 잣의 논문에 대해서 다음과 같은 주석을 달았다. 참조 논문은 스테레오 입력 신호에서 엠비언스를 추출하는 시간-주파수 마스크의 생성을 포함하는 접근법을 제공한다. 마스크는 좌측과 우측 채널 간의 상호-상관(cross-correlation)을 기반으로 하나, 임의의 다중 채널 입력으로부터 엠비언스 추출의 문제 때문에 이 접근은 즉시 적용할 수 없다. 이러한 고차(higher-order)의 경우에서 이 같은 어떤 상관-기반 방법의 사용은 상당한 계산 비용을 요구하는 계층적 쌍 상관 분석 또는 다중채널 상관의 일부 대체 측정을 요구한다.

[0015] 공간 임펄스 응답 렌더링(Spatial Impulse Response Rendering, SIRR)(주하 메리마(Juha Merimaa)와 빌레 풀키(Ville Pulkki), "공간 임펄스 응답 렌더링(Spatial Impulse Response Rendering)", 2004년 제 7회 디지탈 오디오 효과에 관한 국제 학회(Proc. of the 7th Int. Conf. on Digital Audio Effects, DAFx'04))은 B-형식 임펄스 응답들에서 방향 및 확산 사운드를 가진 다이렉트 사운드를 예측한다. SIRR과 매우 유사한, 방향성 오디오 코딩(Directional Audio Coding, DirAC)(빌레 풀키(Ville Pulkki), "방향성 오디오 코딩을 가지는 공간 사운드 재생(Spatial sound reproduction with directional audio coding)", 오디오 엔지니어링 협회 저널, vol.55, no.6, pp.503-516, 2007년 6월)은 B-형식의 연속적인 오디오 신호들에 대한 유사한 다이렉트 및 확산 사운드 분석을 구현한다.

[0016] 줄리아 자카(Julia Jakka), 다중채널 오디오 업믹스에 대한 입체 음향(Binaural to Multichannel Audio Upmix), 박사 학위 논문 및 석사학위 논문, 헬싱키 기술 대학 2005에서 제시하는 접근법은 입력으로서 바이노럴(binaural) 신호들을 사용한 업믹싱을 설명한다.

[0017] 보아즈 라펠리의 참조 논문("잔향 사운드 필드에서 공간적 최적 위너 필터(Spatially Optimal Wiener Filtering in a Reverberant Sound Field), 2001 오디오와 음향의 신호 처리의 응용 IEEE 워크샵(IEEE Workshop on Applications of Signal Processing to Audio and Acoustics 2001), 2001년 10월 21일부터 24일, 뉴 팔즈(New Paltz), 뉴욕(New York)"은 잔향 사운드 필드를 위한 공간적으로 최적인 위너 필터들의 유도를 설명한다. 잔향 실내들에서 두-마이크로폰 잡음 제거에 대한 응용이 주어진다. 확산 사운드 필드들의 공간 상관에서 유도된 최적의 필터들은 사운드 필드들의 지역 동작을 캡처하고, 따라서 잔향 실내들에서 전통적인 적응 잡

음 제거 필터보다 잠재적으로 더욱 공간적으로 강건하고 저차원이다. 비제약 및 인과 제약 최적 필터들에 대한 공식들이 제시되고 두-마이크로폰 스피치 향상에 대한 응용 예는 컴퓨터 시뮬레이션을 이용하여 증명된다.

## 발명의 내용

### 해결하려는 과제

[0018] 본 발명의 목적은 입력 신호 분해를 위한 향상된 개념을 제공하는 것이다.

### 과제의 해결 수단

[0019] 본 발명의 목적은 청구항 1항에 따른 입력 신호 분해 장치, 청구항 14에 따른 입력 신호 분해 방법 또는 청구항 15에 따른 컴퓨터 프로그램으로 달성할 수 있다.

[0020] 본 발명은 다중-채널의 분해에 대한 결과를 기반으로 한다. 입력 신호, 즉, 적어도 세 개의 입력 채널들을 가진 신호를 직접적으로 분석을 수행하지 않는 것이 유리한 접근법이다. 대신에 적어도 세 개의 입력 채널을 가진 다중-채널 입력 신호는 다운믹스된 신호를 획득하기 위하여 입력신호를 다운믹싱하는 다운믹스기에 의해 처리된다. 다운믹스된 신호는 입력 채널의 수보다 적은 다운믹스 채널의 수, 바람직하게는 2를 가진다. 그 후에, 입력 신호 분석은 직접 입력 신호보다 다운믹스된 신호에 대해 수행되고, 분석은 분석 결과를 얻는다. 그러나, 이 분석 결과는 다운믹스된 신호에 적용되지 않으며, 입력 신호 또는, 대안으로, 입력 신호로부터 유도된 신호에 적용되고, 입력 신호로부터 유도된 신호는 업믹스 신호 또는, 입력 신호의 채널 수에 따라, 다운믹스 신호도 될 수 있으나, 입력 신호에서 유도된 이 신호는 분석이 수행된 다운믹스된 신호와는 다를 것이다.

[0021] 예를 들어, 입력 신호가 5.1 채널 신호로 고려되는 경우, 그러면 다운믹스 신호는 두 개의 채널들을 가진 스테레오 다운믹스가 될 수 있으며, 분석은 다운믹스 신호에서 수행된다. 분석 결과는 직접 5.1 입력 신호, 7.1 출력 신호처럼 높은 업믹스, 또는 (단지 세 개 채널의 오디오 렌더링 장치가 사용 가능할 때) 예를 들어 좌측 채널, 중심 채널, 그리고 우측 채널의 단지 세 개 채널을 가진 입력 신호의 다중-채널 다운믹스에 적용된다. 그러나, 어떤 경우든 분석 결과들이 신호 처리기에 의해 적용되는 신호는 분석이 수행된 다운믹스된 신호와 다르고 일반적으로 다운믹스된 신호보다 많은 채널을 가지며, 다운믹스된 신호 상에서 신호 요소들에 대해 분석이 수행된다.

[0022] 다운믹스가 전형적으로 다른 방식에서 입력 채널들의 추가로 구성되므로, 각각의 입력 채널들에서 임의의 신호 요소들이 다운믹스된 채널에서도 발생한다고 가정할 수 있다는 사실 때문에 소위 "간접" 분석/처리가 가능하다.

[0023] 간단한 다운믹스는, 예를 들어, 다운믹스 규칙 또는 다운믹스 매트릭스에 의한 요구로서 개별 입력 채널들을 가중하고, 그 후 개별 입력 채널을 함께 결합한다. 다른 다운믹스는 HRTF 필터들과 같은 특정 필터들과 함께 입력 채널을 필터링하는 것으로 구성되고, 다운믹스는 필터링된 신호들, 즉 기존에 알려진 HRTF 필터들에 의해 필터링된 신호들을 사용함으로써 수행된다. 다섯 개 채널 입력 신호를 위해 10개의 HRTF 필터들이 필요하고, 좌측 부분/좌측 귀를 위한 HRTF 필터의 출력들이 함께 결합 되고, 우측 채널 필터들을 위한 HRTF 필터들의 출력은 우측 귀를 위하여 함께 결합 된다. 다른 다운믹스들이 신호 분석기에서 처리되어야 하는 채널들의 수를 줄이기 위하여 적용될 수 있다.

[0024] 이에, 본 발명의 실시예는 분석 신호를 고려함으로써 임의의 입력 신호들로부터 지각적으로 구별되는 요소들을 추출하는 새로운 개념을 설명하며, 반면에 분석의 결과는 입력 신호에 적용된다. 이와 같은 분석 신호는 예를 들어, 귀에 대한 라우드스피커 신호들 또는 채널들의 전파 모델을 고려함으로써 획득될 수 있다. 이것은 인간의 청각 시스템이 사운드 필드들을 평가하기 위해 오로지 두 개의 센서들(좌측 및 우측 귀)을 또한 사용한다는 사실에 의해 일부 동기된다.

[0025] 따라서, 지각적으로 구별되는 요소들의 추출은 기본적으로 아래에서 다운믹스로 표시되는 분석 신호의 고려를 위해 감소된다. 본 명세서를 통해서, 다운믹스라는 용어는 분석 신호(예를 들어, 전파 모델, HRTFs, BRIRs, 단순 상호-요소 다운믹스를 포함할 수 있음)를 야기하는 다중채널 신호의 어느 사전 처리에 사용된다.

[0026] 주어진 입력의 형식과 추출될 신호의 요구 특성을 알면, 이상적인 채널간 관계들은 다운믹스된 형식을 위해 정의될 수 있고, 이 분석 신호의 분석은 다중채널 신호들의 분해를 위한 가중치 마스크(또는 다중 가중치 마스크들)를 생성하는 데 충분하다.

[0027] 일 실시예에서, 다중-채널 문제는 서라운드 신호의 스테레오 다운믹스를 사용하고, 다이렉트/엠비언트 분석을

다운믹스에 적용함으로써 단순화된다. 결과, 즉, 다이렉트 및 엠비언스 사운드들의 숏-타임 파워 스펙트럼들 예측들에 기반하여, N-채널 신호를 N 직접 사운드와 N 주변 사운드 채널들로 분해하기 위한 필터들이 유도된다.

[0028] 바람직한 실시예에서, 신호 분해의 목적을 위한 특정 효율은, 신호 분석이 참조 곡선으로서 미리-계산된 주파수-종속적인 유사도 곡선을 기반으로 수행되는 경우 획득된다. 유사도라는 용어는 상관(correlation)과 코히어런스(coherence)를 포함한다. 여기서, 엄격한 수학적 의미에서 상관은 추가적인 시간 쉬프트 없이 두 신호 간의 계산 결과이며, 코히어런스는 신호들이 최대 상관을 가지고 주파수에 대한 실제 상관이 시간/위상 쉬프트의 적용과 함께 계산되도록 시간/위상에서 두 신호들을 쉬프팅하여 계산된다. 본 명세서에서, 유사도, 상관과 일관성은 같은 의미 즉, 두 신호 간의 유사도의 정량 정도로 고려되며, 예를 들어 유사도의 더 높은 절대 값은 두 신호가 더욱 유사함을 의미하고, 유사도의 더 낮은 절대 값은 두 신호가 덜 유사함을 의미한다.

[0029] 참조 곡선으로서 이 같은 유사도 곡선의 사용은 매우 효과적으로 구현가능한 분석을 허용하는 것을 볼 수 있는데, 이것은 이 곡선이 쉽게 비교 동작 및/또는 가중치 계산에 사용될 수 있기 때문이다. 미리 계산된 주파수 종속 유사도 곡선의 사용은 더 복잡한 위너 필터 동작보다 간단한 계산들의 수행을 허용한다. 더욱이, 주파수 종속 유사도 곡선의 응용(application)은, 문제에 대한 해답을 얻기 위하여 현재 설정으로부터 가능한 많은 정보들이 도입되기 때문에, 문제가 통계적인 관점이 아니라 분석 방법에 있다는 사실로 인하여 특히 유용하다. 추가로, 참조 곡선은 수많은 다른 방법으로 얻을 수 있기 때문에 이 절차의 유연성은 매우 높다. 실제로 특정 설정에서 둘 또는 그 이상의 신호들을 측정하고 측정된 신호들로부터 주파수 상에서 유사도 곡선을 계산하는 것이 하나의 방법이다. 그러므로, 다른 스피커들로부터의 독립 신호들 또는 미리 알려진 특정 종속도를 가지는 신호들을 발산할 수 있다.

[0030] 바람직한 다른 방법은 독립적인 신호들이라는 가정하에 간단하게 유사도 곡선을 계산하는 것이다. 이 경우에는, 결과는 신호-독립적이므로, 어떤 신호도 실제로 필요하지 않다.

[0031] 신호 분석을 위한 참조 곡선을 사용하는 신호 분해는 스테레오 처리, 즉, 스테레오 신호의 분해에 적용될 수 있다. 또한, 이 절차는 또한 다중채널 신호 분해를 위한 다운믹스기와 함께 구현될 수 있다. 또한, 이 절차는 계층적 방법으로 신호의 쌍 평가를 구상중이라면 다운믹스기의 사용 없이 다중채널 신호를 위하여 역시 구현될 수 있다.

## *발명의 효과*

[0032] 본 발명은 신호 분석이, 필요한 처리 시간이 급격히 단축되는, 보다 적은 수의 채널들에 적용된다는 사실에 장점이 있고, 따라서, 본 발명의 개념은 업믹싱 또는 다운믹싱 또는 신호의 지각적으로 다른 요소들과 같은 다른 요소들이 필요한 다른 어떤 신호 처리 동작을 위한 실시간 어플리케이션에도 적용될 수 있다.

[0033] 본 발명의 또 다른 장점은 다운믹스를 수행함에도 불구하고 입력 신호에서 지각적으로 구별되는 요소의 검출 능력이 저하되지 않는다는 것이다. 다르게 말하면, 입력 채널들이 다운믹스되는 경우에도, 개별 신호 요소들은 큰 범위로 분리될 수 있다. 게다가, 다운믹스는 모든 입력 채널들의 모든 신호 요소들의 두 채널로의 "수집"과 같이 동작하고 이들 "수집된" 다운믹스된 신호들에 적용된 단일 분석은 더 이상 해석될 필요 없는 유일한 결과를 제공하고 신호 처리를 위해 직접 사용될 수 있다.

## *도면의 간단한 설명*

[0034] 도 1은 다운믹스기를 사용한 입력 신호 분해 장치를 나타내는 블록도이다.

도 2는 본 발명의 추가 측면에 따른 미리-계산된 주파수 종속 상관 곡선을 가진 분석기를 사용하여, 적어도 세 개의 입력 채널들을 가진 신호를 분해하는 장치 구현을 나타내는 블록도이다.

도 3은 다운믹스, 분석 및 신호 처리에 대하여 주파수-도메인 처리하는 본 발명의 더 바람직한 구현 예를 나타낸다.

도 4는 도 1 또는 도 2에 표시된 분석을 위한 참조 곡선에 대한 미리-계산된 주파수 종속 상관 곡선의 예를 나타낸다.

도 5는 독립 요소들을 추출하기 위한 추가 처리를 나타내는 블록도이다.

도 6은 독립 확산, 독립 다이렉트 및 다이렉트 요소들이 추출되는 추가 처리에 대한 추가 구현을 나타내는 블록도이다.

도 7은 분석 신호 발생기로서 다운믹스기 구현을 나타내는 블록도이다.

도 8은 도 1 또는 도 2의 신호 분석기에서의 바람직한 처리 방식을 나타내는 흐름도를 나타낸다.

도 9a 내지 도 9e는 (라우드스피커들과 같은) 다른 개수 및 위치를 갖는 사운드 소스들의 여러 다른 설정들을 위한 참조 곡선들로서 사용될 수 있는 다른 미리-계산된 주파수 종속 상관 곡선들을 나타낸다.

도 10은 확산 요소들이 분해될 요소들인 확산 예측의 다른 실시예를 나타내는 블록도이다.

도 11a 및 도 11b는 주파수-종속 상관 곡선 없이, 위너(wiener) 필터링 접근법에 의존하는 신호 분석에 적용하는 수식 예이다.

## *발명을 실시하기 위한 구체적인 내용*

[0035] 본 발명의 바람직한 실시예들이 첨부된 도면들을 참조하여 아래에서 논의된다.

[0036] 도 1은 적어도 세 개의 입력 채널들 또는 일반적으로 N 입력 채널들을 갖는 입력 신호(10)를 분해하는 장치를 나타낸다. 이들 입력 채널들은 다운믹스된 신호(14)를 획득하기 위하여 입력 신호를 다운믹싱하는 다운믹스기(12)로 입력되며, 여기서 다운믹스기(12)는 "m"으로 표시되는 다운믹스된 신호(14)의 다수의 다운믹스 채널들이 적어도 둘이고 입력 신호(10)의 입력 채널들 수보다 적게 다운믹싱하도록 구성된다. m 다운믹스 채널들은 분석 결과(18)를 유도하기 위하여 다운믹스된 신호를 분석하는 분석기(16)로 입력된다. 분석 결과(18)는 신호 처리기(20)로 입력되며, 여기서 신호 처리기는 분석 결과들을 사용하여 신호 유도기(22)에 의해 입력 신호로부터 유도된 신호 또는 입력 신호(10)를 처리하도록 구성되며, 여기서 신호 처리기(20)는 분해된 신호(26)를 획득하기 위해 분석결과를 입력 신호로부터 유도된 신호(24)의 채널들 또는 입력 채널들에 적용하도록 구성된다.

[0037] 도 1에서 나타나는 실시예에서, 입력 신호 대신 유도된 신호가 신호 처리기에 의해 처리되는 경우, 입력 채널들의 수는 n, 다운믹스 채널들의 수는 m, 유도된 채널들의 수는 l, 그리고 출력 채널들의 수는 1과 동일하다. 그 대안으로, 신호 유도기(22)가 존재하지 않아서, 입력 신호가 신호 처리기에 의해서 바로 처리되면, 도 1에서 "l"로 표시된 분해된 신호(26)의 채널 수는 n과 동일할 것이다. 따라서, 도 1은 다른 두 가지 예를 나타낸다. 하나의 예는 신호 유도기(22)가 존재하지 않고, 입력 신호가 신호 처리기(20)에 바로 적용된다. 다른 하나의 예는 신호 유도기(22)가 구현되어 있으며, 입력 신호(10) 대신 유도된 신호(24)가 신호 처리기(20)에 의해 처리된다. 신호 유도기는 예를 들어, 더욱 많은 출력 채널들의 발생을 위한 업믹서와 같은 오디오 채널 믹서가 될 수 있다. 이 경우 l은 n 보다 클 수 있다. 다른 실시예에서, 신호 유도기는 가중치, 지연 또는 입력 채널에 대한 어떤 처리를 수행하는 다른 오디오 처리기가 될 수 있고, 이 경우에는 신호 유도기(22)의 출력 채널들의 수 l은 입력 채널들의 수 n과 같아질 수 있다. 추가 실시예에서, 신호 유도기는 입력 신호로부터 유도된 신호로의 채널들의 수를 감소시키는 다운믹스기(downmixer)가 될 수 있다. 이 실시예에서, 본 발명의 장점 중 하나(즉, 신호 분석이 더 적은 수의 채널 신호들에 적용된다)를 가지기 위해 수 l은 다운믹스된 채널들의 수 m보다 여전히 큰 것이 바람직하다.

[0038] 분석기는 지각적으로 구별되는 요소들에 대하여 다운믹스된 신호를 분석하도록 동작된다. 이들 지각적으로 구별되는 요소들은 한편으로는 개별적인 채널들에서 독립적 요소들이 될 수 있고 다른 한편으로는 종속적인 요소들이 될 수 있다. 본 발명에서 분석되는 대체 신호 요소들은 한편으로는 다이렉트(direct) 요소들이고 다른 한편으로는 앰비언트(ambient) 요소들이다. 음악(music) 요소들로부터의 음성(speech) 요소들, 음성 요소들로부터의 잡음 요소들, 음악 요소들로부터의 잡음 요소들, 저주파수 잡음 요소들에 대한 고주파 잡음 요소들, 다중-피치 신호들에서 다른 악기들에 의해 제공되는 요소들 등과 같은 본 발명에 의해 분리될 수 있는 많은 다른 요소들이 있다. 이는 도 11a, 11b의 관점에서 논의되는 위너 필터링(Wiener filtering) 또는 예를 들어 본 발명에 따른 도 8의 관점에서 논의되는 주파수-종속 상관 곡선을 이용하는 다른 분석 절차들과 같은 강력한 분석 도구들이 있기 때문이다.

[0039] 도 2는 분석기(16)가 미리-계산된 주파수-종속 상관 곡선을 이용하는 것으로 구현된 다른 측면을 도시한다. 따라서, 다수의 채널을 가지는 신호(28)를 분해하는 장치는 예를 들어, 도 1의 관점에서 도시된 다운믹싱 동작에 의해 입력 신호와 동일하거나 입력 신호와 관련된 분석 신호의 두 채널들 간의 상관(correlation)을 분석하는 분석기(16)를 포함한다. 분석기(16)에 의해 분석된 분석 신호는 적어도 두 개의 분석 채널을 가지며, 분석기(16)는 분석 결과(18)를 결정하기 위하여 참조 곡선으로 미리-계산된 주파수 종속 상관 곡선을 사용하도록 구성된다. 신호 처리기(20)는 도 1의 관점에서 논의된 동일한 방식으로 동작할 수 있으며, 신호 유도기(22)에 의한

분석 신호로부터 유도된 신호 또는 분석 신호를 처리하도록 구성된다. 여기서 신호 유도기(22)는 도 1의 신호 유도기(22)의 관점에서 논의된 방식과 유사하게 구현될 수 있다. 그 대안으로, 신호 처리기가 신호를 처리할 수 있으며, 분석 신호가 신호로부터 유도되고 신호 처리는 분해된 신호를 획득하기 위해 분석 결과를 사용한다. 따라서, 도 2의 실시예에서, 입력 신호는 분석 신호와 동일할 수 있고, 이 경우, 분석 신호는 또한 도 2에 나타낸 바와 같이 단지 두 개의 채널들을 가진 스테레오 신호가 될 수 있다. 그 대안으로, 분석 신호는 도 1의 관점에서 묘사된 다운믹싱과 같은 어떤 종류의 처리 또는 업믹싱과 같은 어떤 다른 처리에 의한 입력 신호로부터 유도될 수 있다. 이에 더하여, 신호 처리기(20)는 분석기에 입력되는 동일한 신호에 신호 처리를 적용하는데 유용할 수 있거나, 신호 처리기는 도 1의 관점에서 표시된 것과 같이 신호 처리를 분석 신호가 유도된 신호에 적용할 수 있거나, 신호 처리기는 업믹싱 등에 의해 분석 신호로부터 유도된 신호에 신호처리를 적용할 수 있다.

[0040] 따라서, 신호 처리기에 다른 가능성들이 존재하고, 이 모든 가능성들은 분석 결과를 결정하기 위하여 참조 곡선으로 미리-계산된 주파수-종속 상관 곡선을 사용하는 분석기의 고유한 동작 때문에 이점을 가진다.

[0041] 이후에, 추가 실시예들이 논의된다. 도 2의 관점에서 논의된 것처럼, (다운믹스 없이) 두-채널 분석 신호의 사용도 고려되는 것이 주목된다. 따라서, 도 1과 도 2의 관점에서 다른 측면들로 논의된 본 발명은 함께 또는 별도의 측면들로 사용될 수 있으며, 다운믹스는 분석기 또는 아마도 다운믹스에 의해서 생성되지 않은 두-채널 신호에 의해 처리되고, 본 발명은 미리-계산된 참조 곡선을 사용하는 신호 분석기에 의해 처리될 수 있다. 이러한 관점에서, 특정 기능이 두 개가 아닌 하나의 측면에서 기술된 경우라도 구현 측면들의 후속 설명이 도 1과 도 2에 도식적으로 설명된 양 측면에 적용될 수 있다는 점에 주목해야 한다. 예를 들어, 도 3이 고려된다면, 도 3의 주파수-도메인 특성들은 도 1에서 나타난 측면의 관점에서 기술되는 것이 명백하나, 도 3에서 대해 후속되어 기술되는 것처럼 시간/주파수 변환과 역변환은 다운믹스기가 없으나 미리-계산된 주파수 종속 상관 곡선을 사용하는 명시된 분석기를 가진, 도 2의 구현에도 적용될 수 있다는 것은 명백하다.

[0042] 특히, 시간/주파수 변환기는 분석 신호가 분석기에 입력되기 전에 분석 신호를 변환하기 위해 배치될 것이고, 주파수/시간 변환기는 처리된 신호를 다시 시간 도메인으로 변환하기 위해 신호 처리기의 출력에 배치될 것이다. 신호 유도기가 존재할 때, 시간/주파수 변환기는 신호 유도기, 분석기, 그리고 신호 처리기가 모두 주파수/서브밴드 도메인에서 동작하도록 신호 유도기의 입력에 배치될 수 있다. 이러한 관점에서, 주파수와 서브밴드는 기본적으로 주파수 표현의 주파수에서 일부분을 의미한다.

[0043] 도 1에서 분석기는 많은 다른 방식으로 구현될 수 있으나, 이 분석기는 또한 일 실시예에서, 도 2에서 논의된 분석기 즉, 위너 필터링 또는 어떤 다른 분석 방법의 대안으로 미리-계산된 주파수-종속 상관 곡선을 사용하는 분석기로 구현되는 것이 더욱 명백하다.

[0044] 도 3의 일 실시예는 두-채널 표현을 얻기 위하여 임의의 입력 신호에 다운믹스 절차를 적용한다. 시간-주파수 도메인에서 분석이 수행되고, 가중치 마스크들은 도 3에서 나타내어진 것처럼 입력 신호의 시간 주파수 표현으로 배가되어 계산된다.

[0045] 이 그림에서, T/F는 시간 주파수 변환을 의미한다; 보통은 숏-타임 푸리에 변환(SFTF)이다. iT/F는 각각의 역변환을 의미한다. $[x_1(n),\cdots,x_N(n)]$ 는 시간 도메인의 입력 신호들이며, 여기서 n은 시간 인덱스이다. $[X_1(m,i),\cdots,X_N(m,i)]$ 은 주파수 분해의 계수들을 나타내며, 여기서 m은 분해 시간 인덱스이고, i는 분해 주파수 인덱스이다. $[D_1(m,i),D_2(m,i)]$ 은 다운믹스된 신호의 두 채널들이다.

*수학식 1*

[0046]
$$\begin{pmatrix} D_1(m,i) \\ D_2(m,i) \end{pmatrix} = \begin{pmatrix} H_{11}(i) & H_{12}(i) & \cdots & H_{1N}(i) \\ H_{21}(i) & H_{22}(i) & \cdots & H_{2N}(i) \end{pmatrix} \begin{pmatrix} X_1(m,i) \\ X_2(m,i) \\ \vdots \\ X_N(m,i) \end{pmatrix}$$

[0047] $W(m,i)$ 는 계산된 가중치다. $[Y_1(m,i),\ldots,Y_N(m,i)]$ 는 각 채널의 가중된 주파수 분해들이다. $H_{ij}(i)$ 는 다운믹스 계수들이며, 이는 실수 또는 복소수가 될 수 있고 계수들은 시간 상수 또는 시간-변수일 수 있다. 따라서, 다운믹스 계수들은 단순 상수들 또는 HRTF 필터들, 잔향 필터들 또는 유사 필터들과 같은 필터들일 수 있다.

*수학식 2*

[0048] $$Y_j(m,i)=W_j(m,i)\cdot X_j(m,i),\ \text{where}\ j=(1,2,\ldots,N)$$

[0049] 도 3에서, 모든 채널들에 같은 가중치를 적용하는 경우를 도시하였다.

*수학식 3*

[0050] $$Y_j(m,i)=W(m,i)\cdot X_j(m,i)$$

[0051] $[y_1(n),\ldots,y_N(n)]$ 는 추출된 신호 요소을 포함하는 시간-도메인의 출력 신호이다. (입력 신호는 임의의 대상 재생 라우드스피커 설정을 위하여 제작된 임의의 수의 채널들(N)을 가질 수 있다. 다운믹스는 귀-입력-신호들, 청각 필터들의 시뮬레이션 등을 획득하기 위한 HRTFs를 포함할 수 있다. 다운믹스는 또한 시간 도메인에서 수행될 수 있다.)

[0052] 일 실시예에서, 주파수 함수 $(c_{ref}(\omega))$ 로서 참조 상관(이 문서를 통해서, 상관(correlation)이라는 용어는 상호-채널의 유사도(similarity)와 동의어로 사용되고, 또한 코히어런스(coherence)라는 용어가 평소 사용되는 시간 쉬프트(*shifts*)의 평가들을 포함할 수 있다. 시간-쉬프트가 평가될지라도, 결과 값은 부호를 가질 수 있다. 일반적으로, 코히어런스는 단지 양의 값들을 가지는 것으로 정의된다.)과 다운믹스된 입력 신호 $(c_{sig}(\omega))$ 의 실제 상관의 차이가 계산된다. 참조 곡선으로부터의 실제 곡선의 편차에 따라서, 종속 요소 또는 독립 요소를 포함하는 것을 나타내는 각 시간-주파수 타일(time-frequency tile)에 대한 가중치 요소가 계산된다. 획득된 시간-주파수 가중치는 독립 요소들을 가리키고, 구별과 확산 중에서 하나로 지각될 수 있는 독립 부분들을 포함한 다중채널(multichannel) 신호(채널들의 수는 입력 채널들의 수와 동일한)를 이끌어 내기 위하여 입력 신호의 각 채널에 이미 적용될 수 있다.

[0053] 참조 곡선은 다른 방식들로 정의될 수 있다. 예를 들면:

[0054] · 독립 요소들로 구성된 이상화된(idealized) 2- 또는 3-차원 확산 사운드 필드를 위한 이상적 이론 참조 곡선.

[0055] · 주어진 입력 신호를 위한 참조 대상 라우드스피커 설정으로 성취할 수 있는 이상 곡선(예를 들어, 방위각 ±30° 를 가진 표준 스테레오 설정, 또는 방위각 (0° ,±30° , ±110° )를 가지는 ITU-R BS.775에 따른 표준 다섯 개 채널 설정).

[0056] · 실제의 현재의 라우드스피커 설정을 위한 이상 곡선(실제 위치들은 측정되거나 사용자-입력으로 알려질 수 있다. 참조 곡선은 주어진 라우드스피커들을 통해 독립 신호들의 재생을 가정하고 계산될 수 있다).

[0057] · 각 입력 채널의 실제 주파수-종속 숏 타임 파워(short time power)는 참조 계산에 포함될 수 있다.

[0058] 주파수 종속 참조 곡선 $(c_{ref}(\omega))$ 이 주어지면, 상한 $(c_{hi}(\omega))$ 그리고 하한 $(c_{lo}(\omega))$ 이 정의될 수 있다(도 4 참조). 임계 곡선들은 참조 곡선과 일치할 수 $(c_{ref}(\omega) = c_{hi}(\omega) = c_{lo}(\omega))$ 있거나, 검출 임계값을 가정하여 정의하거나, 경험적으로 도출될 수 있다.

[0059] 참조 곡선으로부터 실제 곡선의 편차가 임계값에 의해 주어진 경계 안에 있으면, 실제 빈(bin)은 독립 요소들을 가리키는 가중치를 갖는다. 상한의 위, 또는 하한의 밑이라면, 빈(bin)은 종속을 가리킨다. 이 표시는 이진 또는 점진적이다(즉, 연판정(soft-decision) 함수를 따름). 특히, 상한과 하한이 참조 곡선과 일치하면, 적용된 가중치는 참조 곡선으로부터의 편차에 직접 연관된다.

[0060] 도 3을 참조하면, 참조 번호 32는 시간/주파수 변환기를 나타낸다. 이 변환기는 숏-타임 푸리에 변환 또는 QMF 필터 뱅크와 같은 서브밴드 신호들을 발생하는 어떤 종류의 필터뱅크로 구현될 수 있다. 시간/주파수 변환기(32)의 자세한 구현과 무관하게, 시간/주파수 변환기의 출력은 각 입력 채널 $x_i$에 대한, 입력 신호의 각 시간 주기에 대한 스펙트럼이다. 따라서, 시간/주파수 처리기(32)는 항상 개별 채널 신호의 입력 샘플들 블록을 획득하기 위해 구현될 수 있고, 그리고 저주파부터 고주파까지 확장된 스펙트럼 라인들(lines)을 가지는 FFT 스펙트럼과 같은 주파수 표현을 계산하기 위해 구현될 수 있다. 그 후에, 다음 시간 블록에 대하여, 결국에 숏 타임 스펙트럼들의 시퀀스가 각 입력 채널 신호에 대하여 계산되도록 동일 절차는 수행된다. 입력 채널의 입력 샘플들의 특정 블록과 관련된 특정 스펙트럼의 특정 주파수 범위는 "시간/주파수 타일"로 불리고, 바람직하게는, 분석기(16)의 분석은 이들 시간/주파수 타일들에 기반해서 수행된다. 그러므로, 분석기는, 하나의 시간/주파수 타일을 위한 입력으로, 제1 다운믹스 채널 $D_1$ 의 입력 샘플의 특정 블록을 위한 첫번째 주파수에 대한 스펙트럼 값을 수신하고, 제2 다운믹스 채널 $D_2$ 의 같은 주파수와 같은 블록(시간상)에 대한 값을 수신한다.

[0061] 그 다음에, 도 8에 도시된 예와 같이, 분석기(16)는 서브밴드 및 시간 블록마다 두 개의 입력 채널들 간의 상관 값, 즉 시간/주파수 타일을 위한 상관 값을 결정(80)하기 위해 구성된다. 그 다음에, 분석기(16)는, 도 2 또는 도 4에서 설명한 실시예에서, 참조 상관 곡선으로부터 상응하는 서브밴드에 대한 상관 값(82)을 획득한다. 예를 들어, 서브밴드가 도 4의 40으로 표시된 서브밴드일때, 단계 82는 -1과 +1 사이의 상관을 표시하는 값(41)이 된다. 그리고 값(41)은 획득된 상관 값이다. 그 다음 단계 83에서, 단계 80으로부터 결정된 상관 값과 단계 82에서 얻은 획득된 상관 값(41)을 사용하는 서브밴드의 결과는 비교와 그에 이은 결정에 의해 수행되거나 실제 차이 값을 계산하여 수행된다. 결과는, 이전에 논의된 바와 같이, 다운믹스/분석 신호에서 고려된 실제 시간/주파수 타일이 독립 요소들을 가진다는 것을 말하는 이진 결과가 될 수 있다. 실제로 결정된 상관 값(단계 80에서)이 참조 상관 값과 같거나 참조 상관 값과 매우 근접할 때, 이러한 결정은 이루어질 것이다.

[0062] 그러나, 결정된 상관 값이 참조 상관 값보다 높은 절대 상관 값을 가리키는 것으로 결정될 경우, 고려중인 시간/주파수 타일은 종속 요소들을 포함하는 것으로 결정된다. 따라서, 다운믹스 또는 분석 신호의 시간/주파수 타일의 상관이 참조 곡선보다 높은 절대 상관 값을 나타낼 경우, 시간/주파수 타일에서 요소들은 서로 종속적이라고 말할 수 있다. 하지만, 상관이 참조 곡선에 매우 근접하게 표시된다면, 요소들은 독립적이라고 말할 수 있다. 종속 요소들은 1과 같은 제1 가중치 값을 받을 수 있고, 독립 요소들은 0과 같은 제2 가중치 값을 받을 수 있다. 바람직하게는, 도 4에 나타난 바와 같이 참조선으로부터 떨어져 있는 상한과 하한은 참조 곡선을 단독으로 사용하는 것보다 더 좋은 결과를 제공하기 위해서 사용된다.

[0063] 더욱이, 도 4를 참조하면, 상관은 +1과 -1 사이에서 변할 수 있음을 알 수 있다. 음의 부호를 가지는 상관은 추가적으로 신호들간의 180° 위상 쉬프트를 나타낸다. 그러므로, 단순히 0과 1 사이에 확장된 다른 상관들은 상관의 음수 부분을 간단히 양수로 만들어서 적용될 수 있다. 그러면, 이 절차에서 상관 결정의 목적을 위한 시간 쉬프트 또는 위상 쉬프트는 무시할 수 있다.

[0064] 결과를 계산하는 다른 방법은 실제로 블록 80에서 결정된 상관 값과 블록 82에서 얻은 획득된 상관 값의 거리를

계산하고, 거리에 기반을 둔 가중치 요소로서 0과 1 사이에서 메트릭을 결정한다. 도 8에서 첫 번째 대안(1)이 오직 0 또는 1의 값들을 가지는 결과인 반면에, 0과 1사이의 값들을 결과로 가지는 가능성(2)은 일부 구현들에서 바람직하다.

[0065] 도 3에서 신호 처리기(20)는 증배기들(multipliers)로 나타내지고, 분석 결과들은 도 8의 84에서 나타내진 바와 같이 분석기에서 신호 처리기로 전달되는 단지 결정된 가중치 요소이고, 그러면 분석 결과들은 입력 신호(10)의 상응하는 시간/주파수 타일에 적용된다. 예를 들어, 실제 고려된 스펙트럼이 스펙트럼들의 시퀀스에서 20번째 스펙트럼이고, 실제 고려된 주파수 빈이 이 20번째 스펙트럼에서 5번째 주파수 빈일 때, 시간/주파수 타일은 (20,5)와 같이 나타낼 수 있으며, 여기서 첫 번째 숫자는 시간에서 블록의 번호를 나타내고, 두 번째 숫자는 이 스펙트럼의 주파수 빈을 나타낸다. 그 후, 시간/주파수 타일(20,5)에 대한 분석 결과는 도 3의 입력 신호의 각 채널의 상응하는 시간/주파수 타일(20,5)에 적용되거나, 도 1에서 나타내진 신호 유도기가 구현되었을 경우 유도된 신호의 각 채널의 상응하는 시간/주파수 타일에 적용된다.

[0066] 이어서, 참조 곡선의 계산이 좀 더 자세히 논의된다. 그러나, 본 발명에서 참조 곡선이 유도되는 방법은 기본적으로 중요하지 않다. 예를 들어, 다운믹스 신호 D 또는 도 2의 관점에서는 분석 신호에서 입력 신호들 $x_j$ 의 이상적 또는 요구된 관계를 나타내는 룩업 테이블의 값들 또는 임의의 곡선이 될 수 있다. 다음의 도출은 예제이다.

[0067] 사운드 필드의 물리적 확산은, 다음 수학식 4에 나타내진 바와 같이 두 공간적으로 분리된 점들에서의 평면파들(plane waves)의 정상 상태 사운드 압력의 상관 계수(r)를 활용하여, 쿡(Cook)등이 소개한 방법(리차드 K. 쿨, R.V. 워터하우스, R.D 버렌트, 시무어 에델만과 Jr.M.C. 톰슨, "잔향 사운드 필드들에서 상관 계수들의 측정(Measurement of correlation coefficients in reverberant sound fields)",미국 음향 학회 저널(Journal of The Acoustical Society of America), vol.27, no. 6, pp.1072-1077, 1995년 11월)으로 평가될 수 있다.

*수학식 4*

$$r=\frac{<p_1(n)\cdot p_2(n)>}{\left[<p_1^2(n)>\cdot<p_2^2(n)>\right]^{\frac{1}{2}}}$$

[0068]

[0069] 여기서, $p_1(n)$ 와 $p_2(n)$ 는 두 점들에서의 사운드 압력 측정들이고, n은 시간 인덱스이고, $<\cdot>$ $<\cdot>$ 는 시간 평균이다. 정상 상태의 사운드 필드에서, 다음 관계식들이 유도될 수 있다.

*수학식 5*

$$r(k,d)=\frac{\sin(kd)}{kd}\,(\text{for three}-\text{dimensional sound fields})$$

[0070]

*수학식 6*

$$r(k,d)=J_0(kd)\,(\text{for two}-\text{dimensional soundfields})$$

[0071]

[0072] 여기서, d는 두 측정점들 간의 거리이며, $k=\frac{2\pi}{\lambda}$ 는 파장 $\lambda$ 을 가지는 파수(wavenumber)이다.(물리적 참조 곡선 r(k,d)는 추가 처리를 위한 $c_{ref}$ 로 사용될 수 있다.)

[0073] 사운드 필드의 지각 확산도를 위한 측정은 사운드 필드에서 측정된 귀간 상호 상관 계수(interaural cross correlation coefficient)( $\rho$ )이다. $\rho$ 측정은 압력 센서들(다시 말해 귀들) 간의 반지름이 고정되어 있음을 의미한다. 이런 제한을 포함하여, r은 라디안 주파수 $\omega=kc$ 를 갖는 주파수의 함수가 된다. 여기서 c는 공기중에서의 사운드의 속도이다. 더욱이, 압력 신호들은 청취자의 귓바퀴, 머리 그리고 몸통에 의한 반사, 회절 및 굽힘-효과들 때문에 이전에 고려된 자유 필드 신호들과 다르다. 공간 청취에 대해 중요한 이러한 효과는 머리-관련 전달 함수들(head-related transfer functions, HRTFs)에 의해 설명된다. 이러한 영향들을 고려하면, 귀 입구에서의 결과 압력 신호들은 $p_L(n,\omega)$ 와 $p_R(n,\omega)$ 이다. 계산과 관련하여서는, 측정된 HRTF 데이터가 사용될 수 있거나, 근사값들이 분석 모델(예를 들어, 리차드.O. 두다와 윌리엄 L.마틴스(Richard O. Duda and William L. Martens), 미국 음향 협회 저널(Journal of The Acoustical Society of America) vol.104, no.5, pp.3048-3058, 11월 1998의 "구형 머리 모델의 응답에 대한 범위 의존(Range depedence of the response of a spherical head model)")을 사용하여 획득될 수 있다.

[0074] 인간의 청각 시스템은 제한된 주파수 선택도를 가진 주파수 분석기처럼 작동하기 때문에, 더욱이 이 주파수 선택도는 통합될 수 있다. 청각 필터들은 중첩된 밴드패스 필터들처럼 동작하는 것으로 가정된다. 다음 예의 설명에서, 중요 밴드 접근법은 사각 필터들에 의하여 이들 중첩된 밴드패스들을 근사화하도록 사용된다. 등가 사각 대역폭(Equivalent Rectangular Bandwidth, ERB)은 중심 주파수의 함수로서 계산될 수 있다(브라이언 R. 글라스버그와 브라이언 C. J. 무어의 청취 연구(Hearing Research) vol. 47, pp.103-138, 1990에서 "노치-잡음 데이터로부터 청각 필터 형태들의 유도(Derivation of auditory filter shapes from notched-noise data)"). 청각 필터링 이후의 바이노럴(binaural) 처리를 고려하면, $\rho$ 은 아래의 주파수 종속 압력 신호들을 산출하는, 분리된 주파수 채널에 대하여 계산되어야 한다.

*수학식 7*

$$p_{\hat{L}}(n,\omega)=\frac{1}{b(\omega)}\int_{\omega-\frac{b(\omega)}{2}}^{\omega+\frac{b(\omega)}{2}}p_L(n,\omega)d\omega$$

[0075]

*수학식 8*

$$p_{\hat{R}}(n,\omega)=\frac{1}{b(\omega)}\int_{\omega-\frac{b(\omega)}{2}}^{\omega+\frac{b(\omega)}{2}}p_R(n,\omega)d\omega$$

[0076]

[0077] 여기서, 적분 한계들은 실제 중심 주파수 ω에 따른 중요 밴드의 경계들에 의하여 주어진다. 인자 $1/b\ (w)$ 는 수학식 7 과 수학식 8에 사용되거나 사용되지 않을 수 있다.

[0078] 사운드 압력 측정들 중 하나가 주파수 독립 시간 차이에 의해서 선행되거나 지연되면, 신호들의 코히어런스

(coherence)가 평가될 수 있다. 인간의 청각 시스템은 시간 정렬 속성 같은 것을 사용할 수 있다. 보통, 귀간의 코히어런스는 ±1ms 이내에서 계산된다. 가능한 처리 능력에 따라, 계산들은 (낮은 복잡도를 위한) 지연-제로 값 또는 (높은 복잡도가 가능할 경우) 시간 선행과 지연을 가지는 코히어런스를 사용하여 구현될 수 있다. 아래에서는, 두 경우 간의 구별은 이루어지지 않는다.

[0079] 이상적인 동작은 모든 방향들로 전파되는 동등한 세기와 비상관된 평면 파들로 구성된 파 필드(wave field)로 이상화된(즉, 전파의 균일하게 분포된 방향들과 임의의 위상 관계들을 가진 무한한 수의 전파 평면 파들의 중첩), 이상적인 확산 사운드 필드를 고려할 때 얻어진다. 라우드스피커에 의해 방사되는 신호는 충분히 멀리 떨어진 곳의 청취자에 대한 평면파로 고려될 수 있다. 이 평면파 가정은 라우드스피커들을 통한 스테레오 재생에서는 일반적이다. 따라서, 라우드스피커들에 의해 재생되는 합성 사운드 필드는 한정된 수의 방향들로부터의 평면 파들의 기여로 구성된다.

[0080] 주어진 N채널들의 입력 신호가 라우드스피커의 위치들이 $[l_1, l_2, l_3, \ldots, l_N]$ 인 설정을 통한 재생을 위해 생성된다. (수평만의 재생 설정의 경우, $l_i$ 는 방위각을 표시한다. 일반적인 경우, $l_i = (azimuth,\ elevation)$ 는 청취자의 머리에 관하여 라우드스피커의 위치를 나타낸다. 청취 룸의 현재 설정이 참조 설정과 다르면, $l_i$ 는 또한 실제 재생 설정의 라우드스피커의 위치들을 표현할 수 있다.) 이 정보를 이용하여 확산 필드 시뮬레이션에 대한 귀간 코히어런스 참조 곡선 $\rho_{ref}$ 이 독립 신호들이 각 라우드스피커로 공급된다는 가정하의 설정에 대하여 계산될 수 있다. 각 시간-주파수 타일에서 각 입력 채널에 의해 기여된 신호 파워는 참조 곡선의 계산에 포함될 수 있다. 예제 구현에서, $\rho_{ref}$ 은 $c_{ref}$로서 사용된다.

[0081] 주파수-종속 참조 곡선들 또는 상관 곡선들의 예들로서 다른 참조 곡선들이 도면들에 표시된 다른 머리 방향들 및 사운드 소스들의 다른 방향들에서 다른 수의 사운드 소스들을 위해 도 9a 내지 9e에서 도시된다.

[0082] 이후에 참조 곡선들에 기반을 둔 도 8의 관점에서 논의된 분석 결과들의 계산이 더욱 자세하게 논의된다.

[0083] 목표는 만일 다운믹스 채널들의 상관이 모든 라우드스피커에서 재생되는 독립 신호들이라는 가정에서 계산된 참조 상관과 같으면 가중치를 1로 유도하는 것이다. 다운믹스의 상관이 +1 또는 -1과 같으면, 유도된 가중치는 독립 요소들이 없는 것을 나타내는 0이 되어야 한다. 이러한 극단적인 경우들 사이에서, 가중치는 독립 표시(W=1) 또는 완전한 종속 표시(W=0) 간에 적절한 전환을 표현해야 한다.

[0084] 참조 상관 곡선 $c_{ref}(\omega)$ 과 실제 재생 설정을 통하여 재생된 실제 입력 신호의 상관/코히어런스의 예측 $c_{sig}(\omega)$ ( $c_{sig}$ 는 각 다운믹스의 코히어런스에 상응하는 상관이다)이 주어지면, $c_{ref}(\omega)$ 로부터 $c_{sig}(\omega)$ 의 편차가 계산될 수 있다.

[0085] (가능한 상한과 하한을 포함하는) 이러한 편차는 독립 요소들의 분리를 위해 모든 입력 채널들에 적용되는 가중치 $W(m,i)$ 를 획득하기 위하여 범위[0;1]에 매핑된다.

[0086] 다음 예제는 임계값들이 참조 곡선에 상응할 때, 가능한 매핑을 나타낸다.

[0087] 참조 곡선 $c_{ref}$ 으로부터 실제 곡선 $c_{sig}$ 의 편차 크기(Δ로 표시되는)는 다음과 같이 주어진다.

수학식 9

$$\Delta(\omega) = | c_{sig}(\omega) - c_{ref}(\omega) |$$

[0088]

[0089] 상관/코히어런스는 [-1;+1] 사이에서 제한되도록 주어지면, 각 주파수에 대한 +1 또는 -1에 대한 최대로 가능한 편차는 다음과 같이 주어진다.

수학식 10

$$\overline{\Delta}_{+}(\omega) = 1 - c_{ref}(\omega)$$

[0090]

수학식 11

$$\overline{\Delta}_{-}(\omega) = c_{ref}(\omega) + 1$$

[0091]

[0092] 각 주파수에 대한 가중치는 다음으로부터 획득된다.

수학식 13

$$W(\omega) = \begin{cases} 1 - \dfrac{\Delta(\omega)}{\overline{\Delta}_{+}(\omega)} & c_{sig}(\omega) \geq c_{ref}(\omega) \\ 1 - \dfrac{\Delta(\omega)}{\overline{\Delta}_{-}(\omega)} & c_{sig}(\omega) < c_{ref}(\omega) \end{cases}$$

[0093]

[0094] 시간 종속과 주파수 분해의 제한된 주파수 분해도를 고려하면, 가중치 값들은 다음과 같이 도출된다(여기서, 시간에 따라 변화할 수 있는 참조 곡선의 일반적 경우가 주어진다. 시간-독립적인 참조 곡선(즉, $c_{ref}(i)$)또한 가능하다).

수학식 14

$$W(m,i) = \begin{cases} 1 - \dfrac{\Delta(m,i)}{\overline{\Delta}_{+}(m,i)} & c_{sig}(m,i) \geq c_{ref}(m,i), \\ 1 - \dfrac{\Delta(m,i)}{\overline{\Delta}_{-}(m,i)} & c_{sig}(m,i) < c_{ref}(m,i) \end{cases}$$

[0095]

[0096] 이 같은 처리는, 계산적 복잡도를 이유로 그리고, 더 짧은 임펄스 응답들을 가지는 필터들을 획득하기 위해, 지각적으로 동기가 부여된 서브밴드들로 그룹핑되는 주파수 계수들을 가지는 주파수 분해에서 수행될 수 있다. 추가로, 스무딩(smoothing) 필터들이 적용될 수 있고, 압축 함수들(즉, 최소 및/또는 최대 가중치 값을 추가로 도입하여, 원하는 방식으로 왜곡한 가중치)이 적용될 수 있다.

[0097] 도 5는 본 발명의 추가 실시예를 도시하며, 여기서 다운믹스기는 HRTF와 청각필터들을 사용하여 구현된다. 더욱이, 도 5는 추가적으로 분석기(16)에 의한 출력인 분석 결과들은 각 시간/주파수 빈에 대한 가중치 요소들이고, 신호 처리기(20)는 독립 요소들을 추출하기 위한 추출기로 도시된다. 다음에, 처리기(20)의 출력은 다시, N 채널들이나, 각 채널은 현재 단지 독립 요소들을 포함하고, 더 이상의 종속 요소들은 포함하지 않는다. 이 구현예에서, 분석기는, 도 8의 첫 번째 구현에서 독립 요소가 가중치 값 1을 수신하고 종속 요소가 가중치 값 0을 수신하도록, 가중치들을 계산할 것이다. 다음에, 처리기(20)에 의해 처리된 원시 N 채널들에서 시간/주파수 타일은 0으로 설정되는 종속 요소들을 가진다.

[0098] 도 8에서 0과 1 사이의 가중치 값들을 가지는 다른 대안에서, 분석기는, 참조 곡선에 대해 작은 거리를 가지는 시간/주파수 타일이 높은 값(1에 더 가까움)을 수신하고 참조 곡선에 대해 큰 거리를 가지는 시간/주파수 타일이 작은 가중치 요소(0에 더 가깝게 존재함)를 수신하도록, 가중치를 계산할 것이다. 예를 들어, 도 3의 20에 도시된 다음의 가중치에서, 독립 요소들은 증폭될 것이고, 반면 종속 요소들은 감쇠될 것이다.

[0099] 그러나, 신호 처리기(20)가 독립 요소들을 추출하지 않고 종속 요소들을 추출하도록 구현되는 경우, 가중치들은, 도 3에 도시된 증배기들(multipliers)(20)에서 가중이 수행되는 경우 독립 요소들이 감쇠되고 종소 요소들이 증폭되도록, 반대로 할당될 것이다. 따라서, 실제 추출된 신호 요소들의 결정은 실제 가중치 값들의 할당에 의해서 결정되므로, 각 신호 처리기는 신호 요소들을 추출하도록 적용될 수 있다.

[0100] 도 6은 처리기(20)의 다른 구현으로, 발명의 개념에 대한 추가 구현을 나타낸다. 도 6의 실시예에서, 처리기(20)는 독립 확산 부분들, 독립 다이렉트 부분들 그리고 다이렉트 부분들/요소들 자체를 추출하도록 구현된다.

[0101] 분리된 독립 요소들($Y_1, \cdots, Y_N$)로부터 인벨로핑/엠비언트 사운드 필드(enveloping/ambient sound field)의 지각에 기여하는 부분들을 획득하기 위하여, 추가 제한 사항들이 고려되어야 한다. 하나의 제한 사항은 인벨로핑 엠비언스 사운드(enveloping ambience sound)가 각 방향으로부터 동등한 세기를 가지는 가정이 될 수 있다. 따라서, 즉, 독립 사운드 신호들의 모든 채널에서 각 시간-주파수 타일의 최소 에너지는 인벨로핑 엠비언트(enveloping ambient) 신호(이는 높은 수의 엠비언스(ambience) 채널들을 얻기 위해 추가 처리될 수 있음)를 얻기 위해서 추출될 수 있다. 예:

*수학식 15*

[0102]
$$\tilde{Y}_j(m,i) = g_j(m,i) \cdot Y_j(m,i), \text{ with } g_j(m,i) = \sqrt{\frac{\min_{1 \le k \le N} \{P_{Y_k}(m,i)\}}{P_{Y_j}(m,i)}}$$

[0103] 여기서 P는 숏-타임 파워 추정치를 나타낸다.(이 예는 가장 간단한 경우를 보여준다. 적용할 수 없는 명백히 예외적인 하나의 경우는, 채널들 중 하나가 채널의 파워가 매우 낮거나 0인 동안 신호 멈춤을 포함하고 있을 때이다.)

[0104] 어떤 경우들에서는 오직 이러한 추출된 스펙트럼들을 사용하여 가중치를 계산하고 모든 입력 채널들의 동일한 에너지 부분들을 추출하는 것이 유리하다.

*수학식 16*

$$\tilde{X}_j(m,i) = g_j(m,i) \cdot X_j(m,i), \text{with } g_j(m,i) = \sqrt{\frac{\min_{1 \le k \le N} \{P_{X_k}(m,i)\}}{P_{X_j}(m,i)}}$$

[0105]

[0106] 추출된 종속들(예를 들어, $Y_{dependent} = Y_j(m,i) - X_j(m,i)$ 부분들과 같이 유도될 수 있다)은 예를 들어 재전개(repanning)와 같은 추가 처리들을 허용함으로써, 입력 신호에 내재된 방향 큐들의 추정과 채널 종속을 위해 사용될 수 있다.

[0107] 도 7은 일반적 개념의 변형을 묘사한다. N-채널 입력 신호가 분석 신호 발생기(analysis signal generator, ASG)에 공급된다. M-채널 분석 신호의 발생은 예를 들어, 채널들/라우드스피커들로부터 귀로의 전파 모델 또는 본 문서를 통해 다운믹스로 표시된 다른 방법들을 포함할 수 있다. 구별 요소들의 표시는 분석 신호에 기반을 둔다. 다른 요소들을 표시하는 마스크들은 입력 신호들에 적용된다(A 추출/D 추출(20a, 20b)). 가중된 입력 신호들은 특정한 특성을 가진 출력 신호들을 이끌어내기 위해서 추가 처리(A 포스트/D 포스트 (70a, 70b))될 수 있다. 이 예제에서 지정자들 "A"와 "D"는 추출되기 위한 요소들이 "엠비언스(Ambience)" 및 "다이렉트 사운드(Direct Sound)"가 될 수 있는 것을 나타내도록 선택되었다.

[0108] 이후에, 도 10이 설명된다. 사운드 에너지의 방향 분포가 방향에 종속되지 않으면, 정적 사운드 필드들은 확산이라 불린다. 방향 에너지 분포는 높은 지향성 마이크로폰(microphone)을 사용하여 모든 방향들에서 측정함으로써 평가될 수 있다. 실내 음향에서, 포위된 곳에서의 잔향 사운드 필드는 종종 확산 필드로 모델링 된다. 확산 사운드 필드는, 동일한 크기로 구성된 파동 필드로서, 모든 방향들로 전파되는 비상관된 평면파들로 이상화될 수 있다. 이와 같은 사운드 필드는 등방성이며 등질(homogeneous)하다.

[0109] 에너지 분포의 균일성에 특별히 관심을 갖는다면, 두 공간적으로 분리된 점들에서 정상상태의 사운드 압력들 $p_1(t)$ 과 $p_2(t)$ 의 점대점 상관 계수인

$$r = \frac{< p_1(t) \cdot p_2(t) >}{\left[ < p_1^2(t) > \cdot < p_2^2(t) > \right]^{\frac{1}{2}}}$$

[0110]

[0111] 는 사운드 필드의 물리적 확산을 평가하는 데 사용될 수 있다. 가정된 이상적 3차원과 이차원의 정상 상태 확산 사운드 필드들은 사인곡선의 소스(sinusoidal source)에 의해 유도되고, 다음 관계식들이 도출될 수 있다:

[0112] $r_{3D} = \frac{\sin(kd)}{kd}$ ,

[0113] $r_{2D} = J_0(kd)$ 과 ,

$$k = \frac{2\pi}{\lambda}(with\ \lambda = wavelength)$$

[0114] 여기서, 는 파동 번호이고, d는 측정점들 간의 거리이다. 주어진 이 같은 관계들에서, 사운드 필드의 확산은 참조 곡선들에 대한 측정 데이터를 비교함으로써 평가될 수 있다. 이상 관계들이 충분 조건들이 아닌 단지 필요 조건들이기 때문에, 마이크로폰들을 연결하는 축들의 다른 방향들에서의 다수의 측정이 고려될 수 있다.

[0115] 사운드 필드에서 청취자를 고려하면, 사운드 압력 측정들은 귀 입력 신호들 $p_l(t)$와 $p_r(t)$에 의해 주어진 $f = \frac{kc}{2\pi}$ 다. 따라서, 측정점들 간의 가정된 거리 d는 고정되고, r은 단지 를 가진 주파수의 함수가 된다. 여기서, c는 공기 중 사운드의 속도이다. 귀 입력 신호들은 청취자의 귓바퀴, 머리 그리고 몸통에 의한 효과들의 영향으로 이전에 고려된 자유 필드 신호와 다르다. 공간 청취에 중요한 이런 효과들은 머리 관련 전달 함수들(head related transfer functions, HRTFs)에 의해 묘사된다. 측정된 HRTF 데이터는 이런 효과들을 통합하는데 사용될 수 있다. 우리는 HRTF들의 근사치를 시뮬레이션하기 위한 분석 모델을 사용한다. 머리는 방위각 ± 100° , 고도 0° 의 귀 위치들과 반지름 8.75cm의 단단한 구형으로 모델링된다. 주어진 이상적 확산 사운드 필드에서 r의 이론적 동작과 HRTF들의 영향에서, 확산 사운드 필드들을 위한 주파수 종속 귀간 상호-상관(interaural cross-correlation) 참조 곡선을 결정하는 것이 가능하다.

[0116] 확산 예측은 시뮬레이트된 큐들과 가정된 확산 필드 참조 신호의 비교를 기반으로 한다. 이 비교는 인간의 청각 한계들에 종속된다. 청각 시스템에서 바이노럴(binaural) 처리는 외이, 중이 및 내이로 구성된 청각계를 따른다. 구형-모델(sphere-model)(예를 들어 귓바퀴 모양, 외이도(ear-canal))에서 근사되지 않는 외이의 효과들과 중이의 효과들은 고려되지 않는다. 내이의 스펙트럼 선택도는 중첩된 밴드패스 필터들(도 10에서 청각 필터들로 표시된)의 뱅크로 모델링된다. 중요 밴드 접근법은 사각 필터들에 의한 이들 중첩된 밴드패스들을 근사화하기 위해 사용된다. 등가 사각 밴드폭(equivalent rectangular bandwidth, ERB)은

$$b(f_c) = 24.7 \cdot (0.00437 \cdot f_c + 1)$$

[0117]

[0118] 에 따른 중심 주파수의 함수로 계산된다.

[0119] 인간의 청각 시스템은 일관된 신호 요소들을 검출하기 위한 시간 정렬을 수행할 수 있다고 가정되고, 상호-상관 분석은 복합 사운드들이 존재하는 상태에서 (ITD에 상응하는) 정렬 시간 τ의 예측에 사용된다고 가정된다. 고주파에서 포락선 상호-상관(envelop cross-correlation)이 관련 큐가 되는 반면, 최대 약 1- 1.5kHz 까지, 반송 신호의 타임 쉬프트들은 파형 상호 상관(waveform cross-correlation)을 사용하여 평가된다. 다음에서, 우리는 이 구별을 하지 않는다. 귀간 코히어런스(interaural coherence, IC) 예측은 정규화된 귀간 상호-상관도 함수

$$IC = \max_{\tau} \left| \frac{< p_L(t) \cdot p_R(t+\tau) >}{\left[ < p_L^2(t) > \cdot < p_R^2(t) > \right]^{\frac{1}{2}}} \right|$$

[0120]

[0121] 의 최대 절대값으로 모델링 된다.

[0122] 바이노럴 지각의 일부 모델들은 러닝 귀간 상호-상관 분석(running interaural cross-correlation)을 고려한다. 정적 신호들을 고려하기 때문에, 시간에 대한 종속을 고려하지 않는다. 중요 밴드 처리의 영향을 모델링하기 위하여, 주파수 종속 정규화된 상호-상관 함수를 다음과 같이 계산한다

$$IC(f_c)=\frac{<A>}{\left[<B>\cdot<C>\right]^{\frac{1}{2}}}$$

[0123]

[0124] 여기서 A는 중요 밴드당 상호-상관도 함수이고, B와 C는 중요 밴드당 자기상관 함수들이다. 밴드패스 상호-스펙트럼과 밴드패스 자기-스펙트럼들에 의한 주파수 도메인에 대한 관계를 다음과 같이 공식화할 수 있다:

$$A=\max_{\tau}\left|2\,\mathrm{Re}\left(\int_{f_-}^{f^+}L^*(f)R(f)e^{j2\pi f(t-\tau)}df\right)\right|,$$

$$B=\left|2\left(\int_{f_-}^{f^+}L^*(f)L(f)e^{j2\pi ft}df\right)\right|,$$

$$C=\left|2\left(\int_{f_-}^{f^+}R^*(f)R(f)e^{j2\pi ft}df\right)\right|,$$

[0125]

[0126] 여기서 $L(f)$ 와 $R(f)$ 은 귀 입력 신호들의 푸리에 변환들이고, $f^{\pm}=f_c\pm\frac{b(f_c)}{2}$ 은 실제 중심 주파수에 따른 중요 밴드의 적분의 상한과 하한을, *는 공액 복소수를 표시한다.

[0127] 다른 각도들에의 둘 이상의 소스들로부터의 신호들이 겹쳐지면, 변동하는 ILD 및 ITD 큐들이 유발된다. 시간 및/또는 주파수의 함수와 같은 ILD 와 ITD의 변화들은 공간감을 발생할 수 있다. 그러나, 긴 시간 평균에서 확산 사운드 필드의 ILD들과 ITD들은 없어야 한다. 0의 평균 ITD는 신호들 간의 상관이 시간 정렬에 의해 증가될 수 없음을 의미한다. ILD들은 원칙적으로 전체 가청 주파수 범위에서 평가될 수 있다. 머리는 저주파에서 장애물을 구성하지 않기 때문에, ILD들은 중간 및 높은 주파수들에서 가장 효과적이다.

[0128] 이어서, 도 11a와 도 11b가 도 10 또는 도 4의 관점에서 논의된 참조 곡선을 사용하지 않는 분석기의 대체 구현을 나타내기 위하여 논의된다.

[0129] 숏-타임 푸리에 변환(Short-time Fourier transform, STFT)은 각각 숏-타임 스펙트럼들 $X_1(m,i)$ 내지 $X_N(m,i)$ 을 산출하는 입력 서라운드 오디오 채널들 $x_1(n)$ 내지 $x_N(n)$ 에 적용된다. 여기서, m은 스펙트럼(시간) 인덱스며, i는 주파수 인덱스이다. $\overline{X}_1(m,i)$ 과 $\overline{X}_2(m,i)$ 로 표시된 서라운드 입력 신호의 스테레오 다운믹스 스펙트럼들이 계산된다. 5.1 서라운드에서, ITU 다운믹스는 수학식 (1)에 적합하다. $X_1(m,i)$ 내지 $X_5(m,i)$ 는 좌(L), 우(R), 중앙(C), 좌측 서라운드(LS), 우측 서라운드(RS) 채널들에 순서대로 상응한다. 아래에서, 시간과 주파수 인덱스들은 표기법의 간단화를 위하여 대체로 생략된다.

[0130] 다운믹스 스테레오 신호에 기반을 둔, 필터 $W_D$ 와 $W_A$ 이 수학식 (2)와 (3)에서 추정되는 다이렉트 및 엠비

언트 사운드 서라운드 신호를 얻기 위해 계산된다.

[0131] 엠비언트 사운드 신호가 모든 입력 채널들 간에 상관성이 없다는 가정이 주어지면, 우리는 다운믹스 계수들을 선택하며 그래서 이러한 가정은 다운믹스 채널들을 위해 또한 유지된다. 따라서, 수학식 4에서 다운믹스 신호 모델을 수립할 수 있다.

[0132] $D_1$ 와 $D_2$ 는 상관된 다이렉트 사운드 STFT 스펙트럼들을 대표하고, $A_1$ 와 $A_2$ 는 비상관된 엠비언스 사운드를 대표한다. 각 채널에서 다이렉트 사운드와 엠비언스 사운드는 서로 비상관되는 것을 추가로 가정한다.

[0133] 최소 자승법의 감각에서 다이렉트 사운드의 예측은 엠비언스를 억제하기 위해 원시 서라운드 신호에 위너 필터를 적용하여 얻을 수 있다. 모든 입력 채널들에 적용될 수 있는 하나의 필터를 도출하기 위해, 수학식 (5)에서와 같이 좌 그리고 우 채널을 위한 동일 필터를 사용하여 다운믹스에서 다이렉트 요소들을 예측한다.

[0134] 이 예측을 위한 공동 평균 제곱 오류 함수(joint mean square error function)은 수학식 (6)으로 주어진다.

[0135] $E\{\cdot\}$ 은 기대 연산자이고 $P_D$ 와 $P_A$ 은 엠비언스 요소들 및 다이렉트의 단기 파워 예측들의 합이다(수학식 7).

[0136] 오류 함수(6)는 자신의 도함수를 0으로 설정함으로써 최소화된다. 다이렉트 사운드의 예측을 위한 결과 필터는 수학식 8에 있다.

[0137] 유사하게, 엠비언트 사운드를 위한 예측 필터는 도 9에서와 같이 유도할 수 있다.

[0138] 다음으로, $W_D$ 와 $W_A$ 를 계산하기 위해 필요한 $P_D$ 와 $P_A$ 를 위한 예측이 유도된다. 다운믹스의 상호-상관도는 수학식 10에 의하여 주어진다.

[0139] 여기서, 다운믹스 신호 모델(40)로 주어진 참조는 (11)로 구성된다.

[0140] 다운믹스의 엠비언스 요소들이 좌우측 다운믹스 채널에서 같은 에너지를 갖는다고 더 가정하면, 수학식 12로 나타낼 수 있다.

[0141] 수학식 10의 마지막 줄에 수학식 12를 대입하고, 수학식 13을 고려하면, 수학식 (14)와 수학식 (15)를 얻는다.

[0142] 도 4의 관점에서 논의된 바와 같이, 최소 상관을 위한 참조 곡선들의 생성은 재생 설정에서 둘 이상의 다른 사운드 소스들로 대체하고 이 재생 설정에서 특정 위치의 청취자의 머리로 대체함으로써 추측될 수 있다. 다음, 완벽히 독립적인 신호들은 다른 라우드스피커에 의해 방출된다. 두 스피커 설정에서, 어떤 상호-믹싱 결과물이 없는 경우에 두 채널들은 상관도가 0으로 완벽히 비상관되어야 한다. 그러나, 이들 상호-믹싱 결과물은 사람의 청취 시스템의 좌측에서 우측으로의 상호-커플링과 또한 실내 잔향 때문에 발생하는 다른 상호 커플링 때문에 발생한다. 그러므로, 도 4 또는 도 9a 내지 9d에 나타내진 결과 참조 곡선들은 항상 0은 아니나, 이 시나리오에서 추측된 참조 신호들이 완전히 독립적일지라도 0으로부터 특히 다른 값들을 가진다. 그러나, 실제로 이들 신호들이 필요하지 않다는 것을 이해하는 것은 중요하다. 참조 곡선을 계산할 때 둘 이상의 신호들간의 충분한 독립성의 가정은 또한 충분하다. 그러나, 이 관점에서 다른 참조 곡선들은, 예를 들면, 완벽히 독립적이지 않으나, 특정한 미리 알고 있는 종속성 또는 서로 간의 종속성 정도를 가진 신호들을 사용하여 다른 시나리오들을 위해 계산될 수 있는 것을 주목해야 한다. 이와 같은 다른 참조 곡선이 계산될 때, 해석 또는 가중치 요소들의 제공은 완전히 독립적인 신호들이 가정되는 참조 곡선과는 다를 것이다.

[0143] 비록 몇몇의 측면들은 장치의 관점에서 설명되었지만, 이러한 측면들은 상응하는 방법의 설명을 또한 나타내는 것을 명확하며, 여기서 블록 또는 장치는 방법의 단계 또는 방법의 단계의 특징에 대응한다. 유사하게, 방법의 단계의 관점에서 설명된 측면들은 상응하는 장치에 대해 상응하는 블록 또는 아이템 또는 특징의 설명을 또한 나타낸다.

[0144] 발명의 분해된 신호는 디지털 저장 매체에 저장될 수 있거나, 인터넷과 같은 유선 전송 매체 및 무선 전송 매체와 같은 전송 매체를 통하여 전송될 수 있다.

[0145] 특정 구현 요구들에 따르면, 발명의 실시예들은 하드웨어 또는 소프트웨어에서 구현될 수 있다. 구현은 디지털 저장 매체(예를 들어, 플로피 디스크(floppy disk), DVD, 블루레이(blue-ray), CD, ROM, PROM, EPROM, EEPROM

또는 플래시 메모리(FLASH memory))를 사용하여 실행될 수 있으며, 디지털 저장 매체는 그것들에 저장된 전자적으로 판독가능한 제어 신호들을 가지며, 이는 프로그래밍 가능한 컴퓨터 시스템과 협력하며(또는 협력하는 것이 가능하며), 그로 인해 각각의 방법이 수행된다.

[0146] 발명에 따른 몇몇 실시예들은 전자적으로 판독가능한 제어 신호들을 가지는 비-일시적인 데이터 운반자(carrier)를 포함하며, 이는 프로그래밍 가능한 컴퓨터 시스템과 협력하는 것이 가능하며, 그로 인해 명세서에 설명된 방법들 중 하나는 수행된다.

[0147] 일반적으로, 컴퓨터 프로그램 제품이 컴퓨터에서 동작하는 경우, 본 발명의 실시예들은 프로그램 코드, 방법들 중 하나의 수행을 위해 동작하는 프로그램 코드를 가지는 컴퓨터 프로그램 제품으로서 수행될 수 있다. 예를 들어, 프로그램 코드는 기계 판독가능한 운반자 상에 저장될 수 있다.

[0148] 다른 실시예들은 명세서에 설명된 방법들 중 하나의 수행을 위한 컴퓨터 프로그램을 포함하며, 이는 기계 판독가능한 운반자 상에 저장된다.

[0149] 즉, 컴퓨터 프로그램이 컴퓨터상에서 동작하는 경우, 따라서 방법 발명의 실시예는 명세서에 설명된 방법들 중 하나의 수행을 위한 프로그램 코드를 가지는 컴퓨터 프로그램을 포함한다.

[0150] 따라서, 방법 발명의 추가 실시예는 데이터 운반자(또는 디지털 저장 매체, 또는 컴퓨터 판독가능한 매체)이며, 데이터 운반자는, 자신에 기록되는, 명세서에 설명된 방법들 중 하나의 수행을 위한 컴퓨터 프로그램을 포함한다.

[0151] 따라서, 방법 발명의 추가 실시예는 데이터 스트림 또는 명세서에 설명된 방법들 중 하나를 수행하기 위한 컴퓨터 프로그램을 표현하는 신호들의 시퀀스(sequence)이다. 예를 들어, 데이터 스트림 또는 신호들의 시퀀스는 데이터 통신 연결(예를 들어, 인터넷)을 통해 전송되기 위해 구성될 수 있다.

[0152] 추가 실시예는 명세서에 설명된 방법들 중 하나의 실행을 위해 적응되거나 또는 구성되는 처리부(예를 들어, 컴퓨터 또는 프로그래밍 가능한 논리 디바이스(device))를 포함한다.

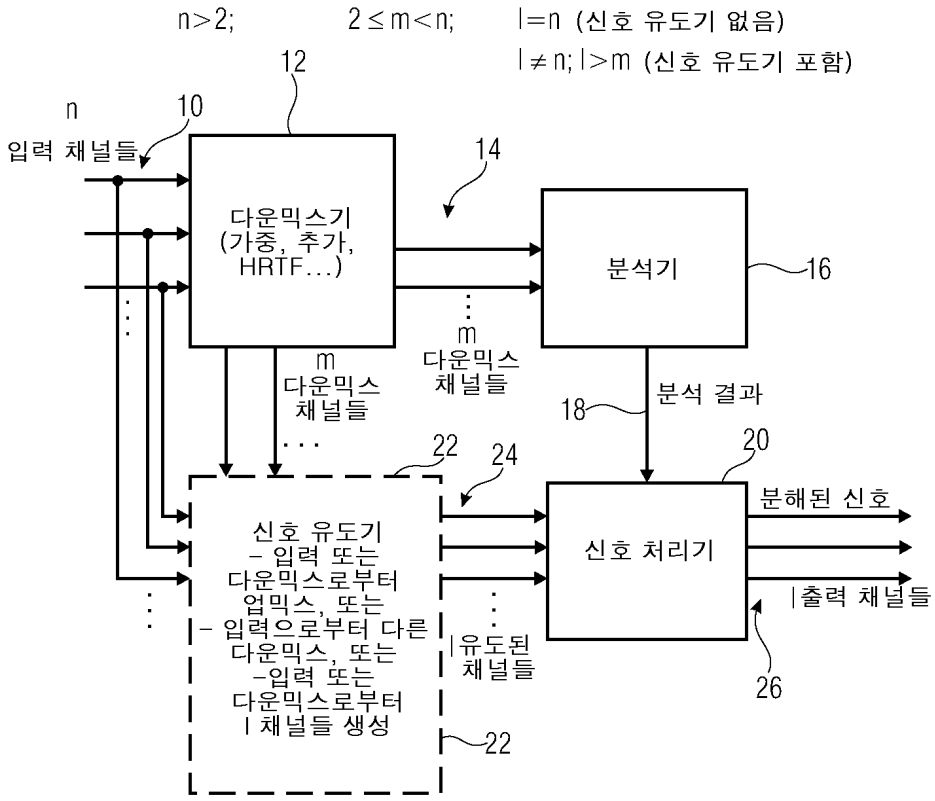
[0153] 추가 실시예는 명세서에 설명된 방법들 중 하나의 실행을 위해 설치된 컴퓨터 프로그램을 가지는 컴퓨터를 포함한다.

[0154] 몇몇 실시예들에서, 프로그래밍 가능한 논리 디바이스(예를 들어, 파일된 프로그래밍 가능한 게이트 어레이(filed programmable gate array, FPGA))는 명세서에 설명된 방법들의 몇몇 또는 모든 기능들의 수행을 위해 사용될 수 있다. 몇몇 실시예들에서, 파일된 프로그래밍 가능한 게이트 어레이는 명세서에 설명된 방법들 중 하나를 수행하기 위해 마이크로프로세서와 협력할 수 있다. 일반적으로, 방법들은 어떤 하드웨어 장치에 의해 바람직하게 수행된다.

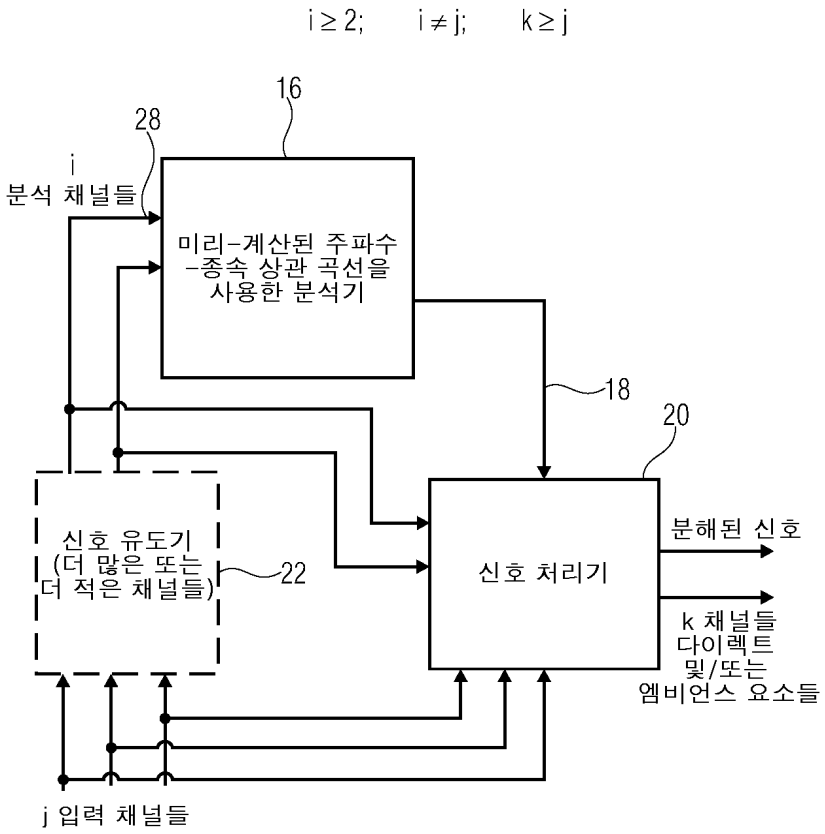
[0155] 상기에서 설명된 실시예들은 본 발명의 원칙들을 위해 단지 설명되었다. 수정들 및 방식들의 변화들 및 명세서에 기재된 상세한 설명들은 당업자에게 자명할 것으로 이해된다. 따라서, 이것은 임박한(impending) 특허 청구항들의 범위에 의해 오직 제한되고 실시예들에 대한 기재 및 설명에 의해 제시된 특정 상세한 설명들에 의해 제한되지 않음을 의도한다.

## 도면

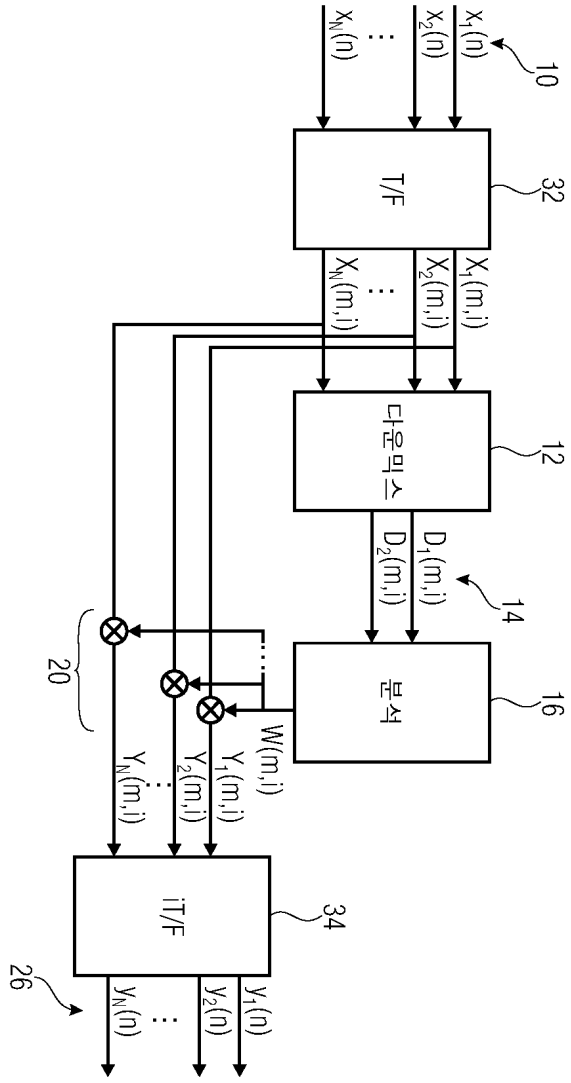
### 도면1


n>2;
2≤m<n;
l=n (신호 유도기 없음)
l≠n; l>m (신호 유도기 포함)
12
n
10
입력 채널들
다운믹스기
(가중, 추가,
HRTF...)
14
분석기
16
m
다운믹스
채널들
m
다운믹스
채널들
18
분석 결과
20
22
24
분해된 신호
신호 유도기
- 입력 또는
다운믹스로부터
업믹스, 또는
- 입력으로부터 다른
다운믹스, 또는
-입력 또는
다운믹스로부터
l 채널들 생성
신호 처리기
l출력 채널들
l유도된
채널들
26
22

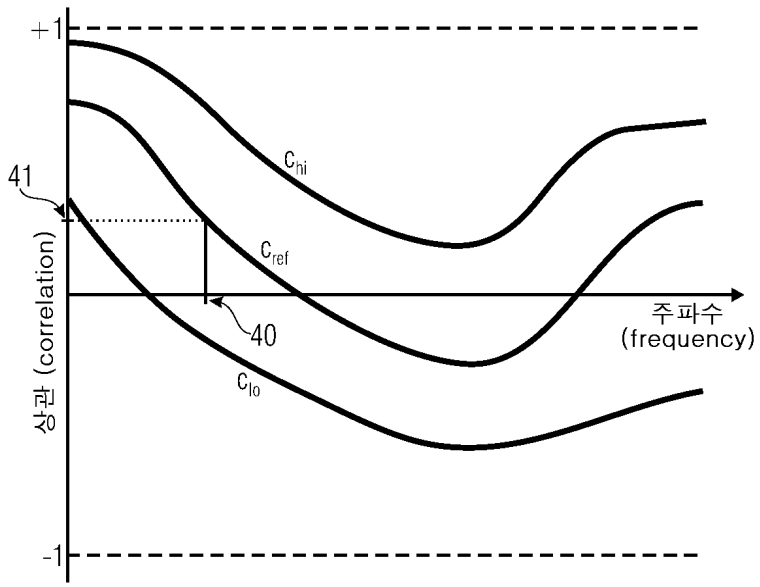
도면2


i ≥ 2;  i ≠ j;  k ≥ j
16
28
i
분석 채널들
미리-계산된 주파수
-종속 상관 곡선을
사용한 분석기
18
20
신호 유도기
(더 많은 또는
더 적은 채널들)
22
신호 처리기
분해된 신호
k 채널들
다이렉트
및/또는
엠비언스 요소들
j 입력 채널들

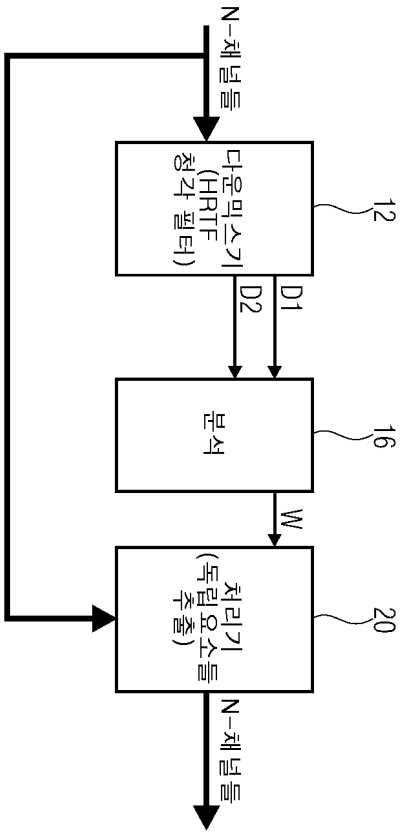
도면3


$x_N(n)$
$x_2(n)$
$x_1(n)$
10
T/F
32
$X_N(m,i)$
$X_2(m,i)$
$X_1(m,i)$
다운믹스
12
$D_2(m,i)$
$D_1(m,i)$
14
분석
16
20
$W(m,i)$
$Y_N(m,i)$
$Y_2(m,i)$
$Y_1(m,i)$
iT/F
34
26
$y_N(n)$
$y_2(n)$
$y_1(n)$

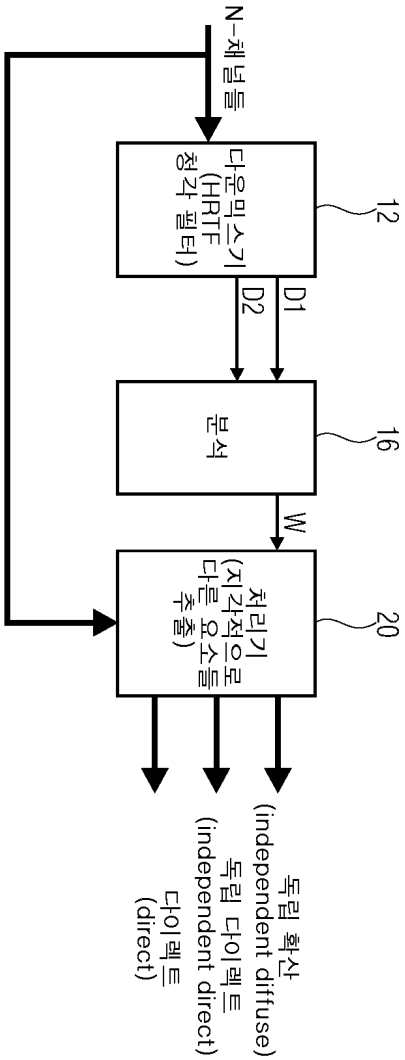
도면4


+1
41
$c_{hi}$
$c_{ref}$
40
$c_{lo}$
-1
상관 (correlation)
주파수
(frequency)

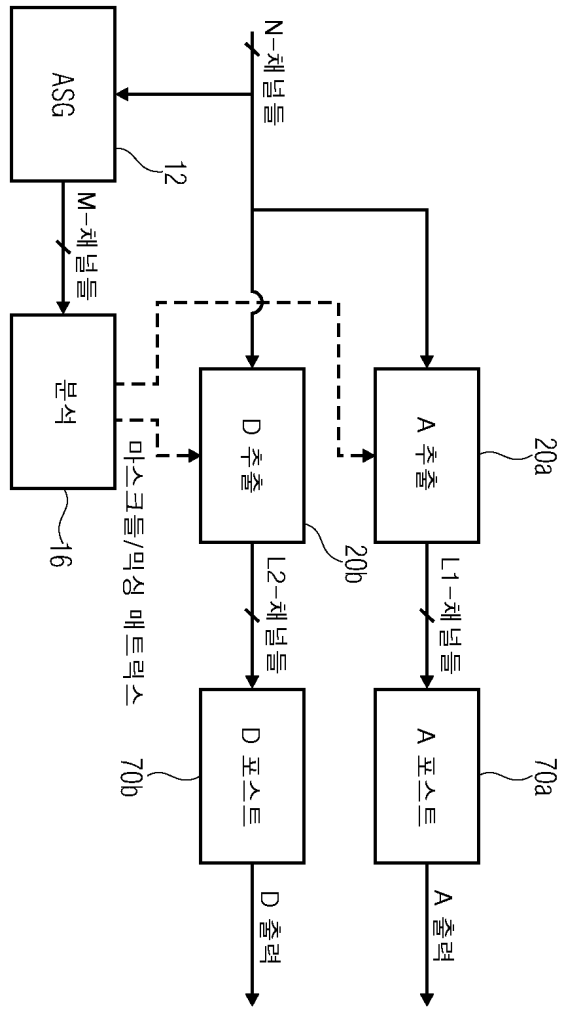
도면5



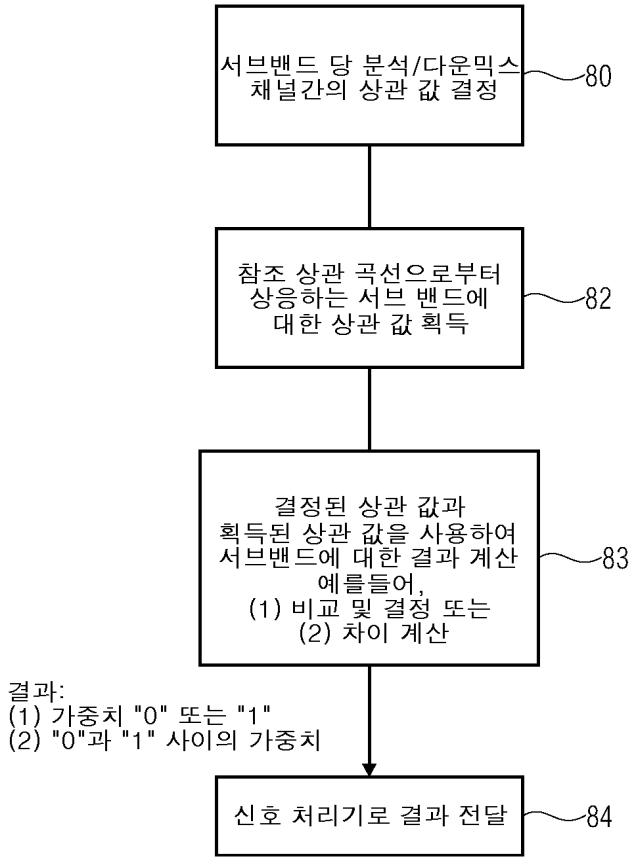
도면6


N-채널들
다운믹스기
(HRTF
청각 필터)
12
D1
D2
분석
16
W
처리기
(지각적으로
다른 요소들
추출)
20
독립 확산
(independent diffuse)
독립 다이렉트
(independent direct)
다이렉트
(direct)

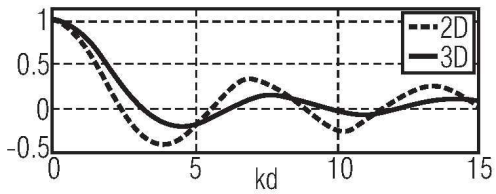
도면7


ASG
12
N-채널들
M-채널들
분석
16
마스크들/믹싱 매트릭스
D 추출
20b
A 추출
20a
L2-채널들
L1-채널들
D 포스트
70b
A 포스트
70a
D 출력
A 출력

도면8

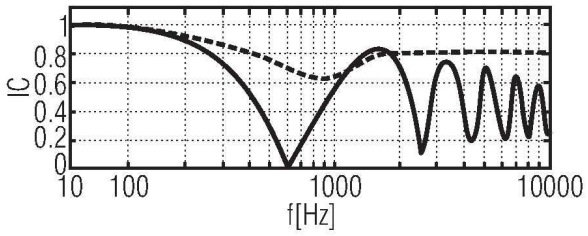



도면9a



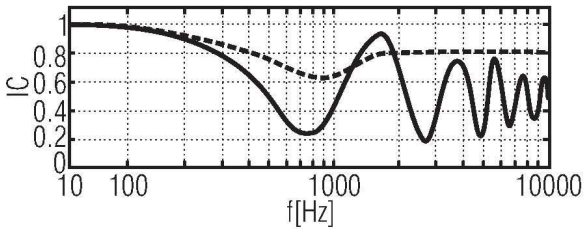

2차원 및 3차원 이상적 확산 사운드 필드들에서 kd의 함수인 상호-상관(r)

도면9b



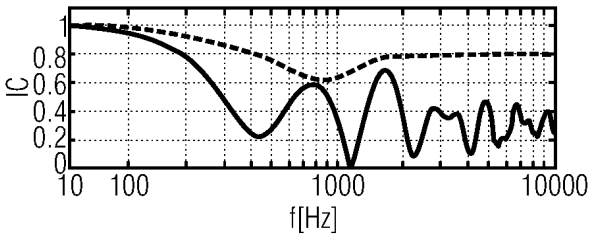

머리방향 0°를 가진 방위각 ±30°에서의 2개 사운드 소스들에 대한 $IC(f_c)$

도면9c



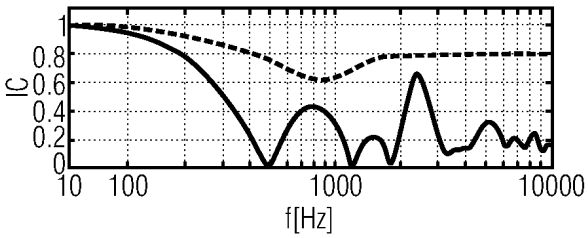

머리방향 25°를 가진 방위각 ±30°에서의 2개 사운드 소스들

도면9d



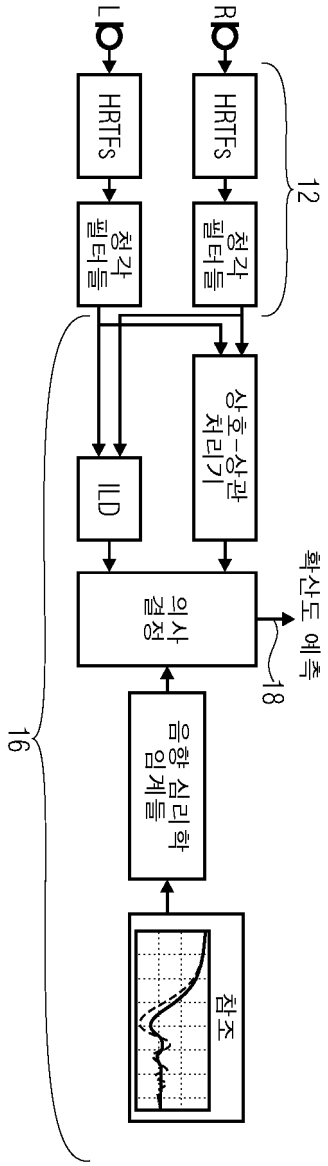

머리방향 25°를 가진 방위각 {±30°, ±110°}에서의 4개 사운드 소스들

도면9e




머리방향 25°를 가진 방위각 {0, ±45°, ±90°, ±135°, ±180°} 에서의 8개 사운드 소스들

도면10


L
R
HRTFs
HRTFs
청각 필터들
청각 필터들
12
상호-상관 처리기
ILD
확산도 예측
의사 결정
18
16
음향 심리학 임계들
참조

도면11a

$$\begin{pmatrix}\bar{X}_1(m,i)\\ \bar{X}_2(m,i)\end{pmatrix}=\begin{pmatrix}1 & 0 & \frac{1}{\sqrt{2}} & 1 & 0\\ 0 & 1 & \frac{1}{\sqrt{2}} & 0 & 1\end{pmatrix}\begin{pmatrix}X_1(m,i)\\ \vdots\\ X_5(m,i)\end{pmatrix}, \quad (1)$$

$$\begin{pmatrix}Y_{1D}(m,i)\\ \vdots\\ Y_{ND}(m,i)\end{pmatrix}=W_D(m,i)\begin{pmatrix}X_1(m,i)\\ \vdots\\ X_N(m,i)\end{pmatrix} \quad (2)$$

$$\begin{pmatrix}Y_{1A}(m,i)\\ \vdots\\ Y_{NA}(m,i)\end{pmatrix}=W_A(m,i)\begin{pmatrix}X_1(m,i)\\ \vdots\\ X_N(m,i)\end{pmatrix}, \quad (3)$$

$$\begin{aligned}\bar{X}_1&=D_1+A_1\\ \bar{X}_2&=D_2+A_2,\end{aligned} \quad (4)$$

$$\begin{aligned}\hat{D}_1&=W_D\cdot\bar{X}_1\\ \hat{D}_2&=W_D\cdot\bar{X}_2.\end{aligned} \quad (5)$$

$$\begin{aligned}J&=E\{|D_1-\hat{D}_1|^2\}+E\{|D_2-\hat{D}_2|^2\}\\ &=(W_d-1)^2P_D+W_d^2P_A,\end{aligned} \quad (6)$$

$$\begin{aligned}P_D&=E\{|D_1|^2\}+E\{|D_2|^2\}\\ P_A&=E\{|A_1|^2\}+E\{|A_2|^2\}.\end{aligned} \quad (7)$$

$$W_D=\frac{P_D}{P_{D+}P_A}. \quad (8)$$

$$W_A=\frac{P_A}{P_{D+}P_A}=1-W_d. \quad (9)$$

분석을 위한 위너 필터링

도면11b

$$\begin{aligned}\mathrm{Re}(E\{\check{X}_1\check{X}_2^*\}) &= \mathrm{Re}(E\{D_1+A_1)(D_2^*+A_2^*)\}) \\ &= \mathrm{Re}(E\{D_1D_2^*\}) \\ &= \sqrt{E\{|D_1|^2\}E\{|D_2|^2\}}. \end{aligned} \quad (10)$$

$$\begin{aligned}E\{|D_1|^2\} &= E\{|\bar{X}_1|^2\}-E\{|A_1|^2\} \\ E\{|D_2|^2\} &= E\{|\bar{X}_2|^2\}-E\{|A_2|^2\}. \end{aligned} \quad (11)$$

$$\begin{aligned}E\{|D_1|^2\} &= E\{|\bar{X}_1|^2\}-\frac{P_A}{2} \\ E\{|D_2|^2\} &= E\{|\bar{X}_2|^2\}-\frac{P_A}{2}. \end{aligned} \quad (12)$$

$$P_D+P_A = E\{|\bar{X}_1|^2\}+E\{|\bar{X}_2|^2\}, \quad (13)$$

$$P_D=\sqrt{(E\{|\bar{X}_1|^2\}-E\{|\bar{X}_2|^2\})^2+4\mathrm{Re}(E\{\bar{X}_1\bar{X}_2^*\})^2} \quad (14)$$

$$P_A=E\{|\bar{X}_1|^2\}+E\{|\bar{X}_2|^2\}-P_D. \quad (15)$$

분석을 위한 위너 필터링