



(19) 대한민국특허청(KR)  
 (12) 등록특허공보(B1)

(45) 공고일자 2019년06월27일  
 (11) 등록번호 10-1993587  
 (24) 등록일자 2019년06월21일

- (51) 국제특허분류(Int. Cl.)  
*H04S 7/00* (2006.01) *H04H 20/89* (2008.01)  
*H04L 29/06* (2006.01) *H04S 3/00* (2006.01)
- (52) CPC특허분류  
*H04S 7/30* (2013.01)  
*H04H 20/89* (2013.01)
- (21) 출원번호 10-2017-7017001(분할)
- (22) 출원일자(국제) 2013년06월18일  
 심사청구일자 2018년06월18일
- (85) 번역문제출일자 2017년06월20일
- (65) 공개번호 10-2017-0075025
- (43) 공개일자 2017년06월30일
- (62) 원출원 특허 10-2015-7004013  
 원출원일자(국제) 2013년06월18일  
 심사청구일자 2016년05월18일
- (86) 국제출원번호 PCT/US2013/046369
- (87) 국제공개번호 WO 2014/014600  
 국제공개일자 2014년01월23일
- (30) 우선권주장  
 61/671,789 2012년07월15일 미국(US)  
 13/844,447 2013년03월15일 미국(US)
- (56) 선행기술조사문헌  
 US20040247134 A1  
 EP2469741 A  
 US20060045275 A1

(73) 특허권자  
 웰컴 인코포레이티드  
 미국 92121-1714 캘리포니아주 샌 디에고 모어하우스 드라이브 5775

(72) 발명자  
 센 디판잔  
 미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775  
 샹 페이  
 미국 92121-1714 캘리포니아주 샌디에고 모어하우스 드라이브 5775

(74) 대리인  
 특허법인코리아나

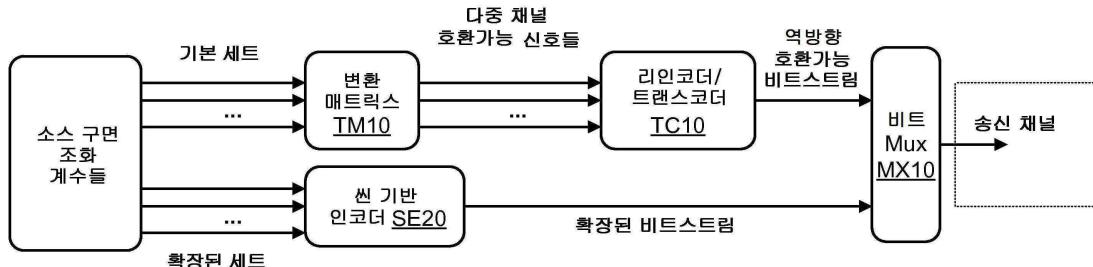
전체 청구항 수 : 총 24 항

심사관 : 권영학

(54) 발명의 명칭 역방향 호환가능 오디오 코딩을 위한 시스템, 방법, 장치 및 컴퓨터 판독가능 매체

**(57) 요약**

사운드 필드를 기술하는 기저 함수 계수들의 세트의 역방향 호환가능 코딩을 위한 시스템, 방법 및 장치가 제시된다.

**대 표 도**

(52) CPC특허분류

*H04L 65/601* (2013.01)

*H04S 3/008* (2013.01)

*H04S 2400/03* (2013.01)

*H04S 2400/15* (2013.01)

*H04S 2420/11* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

오디오 데이터를 디코딩하는 디바이스로서,  
비트스트림을 수신하도록 구성된 수신기, 및

상기 비트스트림에 포함된 역방향 호환가능 비트스트림과 상기 비트스트림에 포함된 추가의 인코딩된 비트들에  
기초하여, 복수의 출력 확성기 피드들을 렌더링하도록 구성된 디코더 또는 렌더러를 포함하고,

상기 역방향 호환가능 비트스트림은 상기 추가의 인코딩된 비트들을 사용하지 않고 디코딩가능하고, 상기 추가  
의 인코딩된 비트들이 구면 조화 계수 (SHC) 들을 특정하며, 상기 SHC 들은 구면 조화 기저 함수들의 계수들인,  
오디오 데이터를 디코딩하는 디바이스.

#### 청구항 2

제 1 항에 있어서,

상기 역방향 호환가능 비트스트림은 다중 채널 오디오 포맷에 따라 포맷팅되는, 오디오 데이터를 디코딩하는 디  
바이스.

#### 청구항 3

제 2 항에 있어서,

상기 다중 채널 오디오 포맷에 대응하는 확성기 지오메트리와 상이한 로컬 확성기 지오메트리에 따라 배열되는  
복수의 확성기들을 더 포함하고,

상기 복수의 확성기들은 상기 복수의 출력 확성기 피드들을 수신하도록 구성되고,

상기 수신기는 상기 복수의 출력 확성기 피드들이 상기 로컬 확성기 지오메트리에 적응되도록 상기 복수의 출력  
확성기 피드들을 렌더링하도록 구성되는, 오디오 데이터를 디코딩하는 디바이스.

#### 청구항 4

제 3 항에 있어서,

상기 다중 채널 오디오 포맷에 대응하는 상기 확성기 지오메트리 내의 확성기들은 공면 (coplanar) 에 있고 상  
기 로컬 확성기 지오메트리 내의 확성기들은 공면에 있지 않은, 오디오 데이터를 디코딩하는 디바이스.

#### 청구항 5

제 1 항에 있어서,

상기 비트스트림은 패킷들을 포함하고, 상기 추가의 인코딩된 비트들은 상기 패킷들의 확장된 부분들로 패킹되  
는, 오디오 데이터를 디코딩하는 디바이스.

#### 청구항 6

제 1 항에 있어서,

상기 디바이스는 상기 역방향 호환가능 비트스트림과 상기 추가의 인코딩된 비트들을 디멀티플렉싱하도록 구성  
된 디멀티플렉서를 포함하는, 오디오 데이터를 디코딩하는 디바이스.

#### 청구항 7

제 1 항에 있어서,

상기 복수의 출력 확성기 피드들을 수신하는 하나 이상의 확성기들을 더 포함하는, 오디오 데이터를 디코딩하는

디바이스.

### 청구항 8

제 1 항에 있어서,

상기 역방향 호환가능 비트스트림을 다중 채널 오디오 신호들로 컨버팅하도록 구성되는 트랜스코더를 더 포함하는, 오디오 데이터를 디코딩하는 디바이스.

### 청구항 9

오디오 데이터를 인코딩하는 디바이스로서,

사운드 필드를 표현하는 복수의 오디오 입력 신호들에 기초하여, 역방향 호환가능 비트스트림과 추가의 인코딩된 비트들을 포함하는, 비트스트림을 생성하도록 구성되는 인코더를 포함하고,

상기 역방향 호환가능 비트스트림은, 상기 사운드 필드를 표현하는 복수의 출력 확성기 피드들을 생성하기 위해 상기 추가의 인코딩된 비트들을 사용하지 않고 디코딩가능하고,

상기 추가의 인코딩된 비트들은, 상기 사운드 필드를 표현하는 구면 조화 계수 (SHC) 들의 세트에서 상위 차수 구면 조화 계수들의 세트를 표현하되, 상기 사운드 필드를 표현하는 SHC 들의 세트에서 하위 차수 구면 조화 계수들을 표현하지는 않고,

상기 SHC 들은 구면 조화 기저 함수들의 계수들이고, 상기 상위 차수 구면 조화 계수들 각각은 상기 하위 차수 구면 조화 계수들 중 어느 하나보다도 높은 차수를 갖는, 오디오 데이터를 인코딩하는 디바이스.

### 청구항 10

제 9 항에 있어서,

상기 디바이스는 상기 오디오 입력 신호들을 수신하도록 구성되는 2 이상의 마이크로폰들을 포함하는, 오디오 데이터를 인코딩하는 디바이스.

### 청구항 11

제 9 항에 있어서,

상기 역방향 호환가능 비트스트림과 상기 추가의 인코딩된 비트들을 저장하도록 구성된 메모리를 더 포함하는, 오디오 데이터를 인코딩하는 디바이스.

### 청구항 12

제 9 항에 있어서,

상기 비트스트림은 패킷들을 포함하고, 그리고 상기 추가의 인코딩된 비트들은 사용자 정의된 데이터를 위해 지정된 패킷들의 필드들로 패킹되는, 오디오 데이터를 인코딩하는 디바이스.

### 청구항 13

제 9 항에 있어서,

상기 인코더는 상기 하위 차수 구면 조화 계수들을 상기 역방향 호환가능 비트스트림으로 컨버팅하도록 구성되는, 오디오 데이터를 인코딩하는 디바이스.

### 청구항 14

오디오 데이터를 디코딩하는 방법으로서,

디코더에 의해, 비트스트림에 포함된 역방향 호환가능 비트스트림과 상기 비트스트림에 포함된 추가의 인코딩된 비트들에 기초하여, 복수의 출력 확성기 피드들을 렌더링하는 단계를 포함하고,

상기 역방향 호환가능 비트스트림은 상기 추가의 인코딩된 비트들을 사용하지 않고 디코딩가능하고, 상기 추가의 인코딩된 비트들이 구면 조화 계수 (SHC) 들을 특정하며, 상기 SHC 들은 구면 조화 기저 함수들의 계수들인,

오디오 데이터를 디코딩하는 방법.

### 청구항 15

제 14 항에 있어서,

상기 역방향 호환가능 비트스트림은 다중 채널 오디오 포맷에 따라 포맷팅되는, 오디오 데이터를 디코딩하는 방법.

### 청구항 16

제 15 항에 있어서,

상기 다중 채널 오디오 포맷에 대응하는 확성기 지오메트리 내의 모든 확성기들은 단일 2 차원 평면에 있고, 상기 디코더의 로컬 확성기 지오메트리 내의 확성기들은 3 차원 공간에 걸친, 오디오 데이터를 디코딩하는 방법.

### 청구항 17

제 15 항에 있어서,

상기 다중 채널 오디오 포맷은, 5.1 확성기 어레이 포맷, 7.1 확성기 어레이 포맷, 또는 22.2 확성기 어레이 포맷 중 하나인, 오디오 데이터를 디코딩하는 방법.

### 청구항 18

제 14 항에 있어서,

상기 복수의 출력 확성기 피드들 중 각 출력 확성기 피드는 공연 방향들의 세트 내의 각각의 방향에 대응하는, 오디오 데이터를 디코딩하는 방법.

### 청구항 19

제 14 항에 있어서,

상기 SHC 들은 SHC 들의 세트의 제 1 부분이고,

상기 복수의 출력 확성기 피드들을 렌더링하는 단계는,

상기 역방향 호환가능 비트스트림을 SHC 들의 상기 세트의 제 2 부분으로 컨버팅하는 단계; 및

상기 복수의 출력 확성기 피드들을 생성하기 위해 SHC 들의 상기 세트를 렌더링하는 단계를 포함하는, 오디오 데이터를 디코딩하는 방법.

### 청구항 20

오디오 데이터를 인코딩하는 방법으로서,

복수의 입력 오디오 신호들을 구면 조화 계수 (SHC) 들의 세트로 컨버팅하는 단계로서, 상기 SHC 들은 구면 조화 기저 함수들의 계수들이고, SHC 들의 상기 세트는 SHC 들의 제 1 서브세트 및 SHC 들의 제 2 서브세트로 이루어지는, 상기 컨버팅하는 단계; 및

역방향 호환가능 비트스트림과 추가의 인코딩된 비트들을 포함하는, 비트스트림을 발생시키는 단계로서, 상기 역방향 호환가능 비트스트림은 다중 채널 오디오 포맷에서 SHC 들의 상기 제 1 서브세트를 표현하고, 상기 추가의 인코딩된 비트들은 SHC 들의 상기 제 2 서브세트를 표현하는, 상기 발생시키는 단계

를 포함하는 오디오 데이터를 인코딩하는 방법.

### 청구항 21

제 20 항에 있어서,

SHC 들의 상기 제 1 서브세트는 상기 역방향 호환가능 비트스트림으로부터 복구가능한, 오디오 데이터를 인코딩 하는 방법.

**청구항 22**

제 20 항에 있어서,

상기 복수의 입력 오디오 신호들은 오디오 오브젝트를 포함하는, 오디오 데이터를 인코딩하는 방법.

**청구항 23**

제 20 항에 있어서,

SHC 들의 상기 제 1 서브세트의 각 SHC 는 SHC 들의 상기 제 2 서브세트의 어느 SHC 보다도 하위 차수 SHC 인, 오디오 데이터를 인코딩하는 방법.

**청구항 24**

제 20 항에 있어서,

상기 역방향 호환가능 비트스트림은 초고정세 텔레비전 표준 (Ultra High Definition Television standard) 의 채널 기반 포맷과 호환가능한, 오디오 데이터를 인코딩하는 방법.

**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

## 발명의 설명

### 기술 분야

[0001]

본원은 2012년 7월 15일에 출원된 명칭이 "PROVIDING BACKWARD COMPATIBILITY WHEN USING SPHERICAL HARMONICS COEFFICIENTS FOR SPATIAL AUDIO REPRESENTATION"이고, 본원의 양수인에게 양도된 가출원 제 61/671,789호에 대해 우선권을 주장한다.

[0002]

본 개시물은 공간 오디오 코딩에 관한 것이다.

### 배경 기술

[0003]

서라운드 사운드의 진화는 요즈음 엔터테인먼트에 대해 이용 가능한 많은 출력 포맷들을 만들고 있다. 시중에서 서라운드 사운드 포맷들의 범위는 대중적인 5.1 홈 씨어터 시스템 포맷을 포함하며, 이 홈 씨어터 시스템 포맷은 스테레오를 넘어 거실에 영향을 미친다는 면에서 가장 성공적인 것이다. 이 포맷은 다음의 6개의 채널들을 포함한다: 전방 좌측(L), 전방 우측(R), 중앙 또는 전방 중앙(C), 후방 좌측 또는 서라운드 좌측(Ls), 후방 우측 또는 서라운드 우측(Rs) 및 저주파수 효과들(LFE). 서라운드 사운드 포맷들의 다른 예들은, 예를 들어, 초고정세 텔레비전 표준으로 사용하기 위해 NHK(Nippon Hoso Kyokai 또는 Japan Broadcasting Corporation)에 의해 개발된 미래의 22.2 포맷 및 성장 중인 7.1 포맷을 포함한다. 서라운드 사운드 포맷은 2차원에서 및/또는 3차원에서 오디오를 인코딩하는 것이 바람직할 수도 있다.

### 발명의 내용

#### 해결하려는 과제

#### 과제의 해결 수단

[0004]

일반적인 구성에 따라, 복수의 기저 함수 계수들을 프로세싱하는 방법이 기재된다. 복수의 기저 함수 계수들은 시간 간격 동안 사운드 필드를 기술한다. 이 방법은 복수의 채널 신호들을 생성하기 위해 복수의 기저 함수 계수들의 제1 그룹에 대해 가역(reversible) 변환을 수행하는 단계를 포함하며, 복수의 채널 신호들의 각각은 공간의 대응하는 상이한 영역과 연관된다. 이 방법은 또한 복수의 채널 신호들에 기초하여, (A) 복수의 기저 함수 계수들의 제2 그룹의 표현으로서, 제2 그룹은 제1 그룹과는 상이한, 제2 그룹의 표현(representation), 및 (B) 제2 그룹의 표현과 별개인 복수의 채널 신호들의 표현을 포함하는 데이터 구조를 생성하는 단계를 포함한다. 또한, 유형의 피쳐들을 관리하는 머신으로 하여금 그러한 방법을 수행하게 하는

유형의 피쳐들을 갖는 컴퓨터 판독가능 저장 매체 (예를 들어, 비일시적 매체) 가 개시된다.

[0005] 일반적인 구성에 따라, 복수의 기저 함수들을 프로세싱하는 장치가 기재된다. 복수의 기저 함수들은 시간 간격 동안 사운드 필드를 기술한다. 이 장치는 복수의 채널 신호들을 생성하기 위해 복수의 기저 함수들의 제 1 그룹에 대해 가역 변환을 수행하는 수단을 포함하고, 복수의 채널 신호들의 각각은 공간의 대응하는 상이한 영역과 연관된다. 이 장치는 또한 복수의 채널 신호들에 기초하여, (A) 복수의 기저 함수 계수들의 제 2 그룹의 표현으로서, 제 2 그룹은 제 1 그룹과는 상이한, 제 2 그룹의 표현, 및 (B) 제 2 그룹의 표현과 별개인 복수의 채널 신호들의 표현을 포함하는 데이터 구조를 생성하는 수단을 포함한다.

[0006] 또 다른 일반적인 구성에 따라, 복수의 기저 함수 계수들을 프로세싱하는 장치가 기재된다. 복수의 기저 함수들은 시간 간격 동안 사운드 필드를 기술한다. 이 장치는 복수의 채널 신호들을 생성하기 위해 복수의 기저 함수들의 제 1 그룹에 대해 가역 변환을 수행하도록 구성된 계산기를 포함하고, 복수의 채널 신호들의 각각은 공간의 대응하는 상이한 영역과 연관된다. 이 장치는 또한 복수의 채널 신호들에 기초하여, (A) 복수의 기저 함수 계수들의 제 2 그룹의 표현으로서, 제 2 그룹은 제 1 그룹과는 상이한, 제 2 그룹의 표현, 및 (B) 제 2 그룹의 표현과 별개인 복수의 채널 신호들의 표현을 포함하는 데이터 구조를 생성하도록 구성된 데이터 포맷터를 포함한다.

### 도면의 간단한 설명

[0007] 도 1a 는 역방향 호환성을 고려하지 않은 씬 (scene) 기반 코딩의 일 예를 나타낸다.

도 1b 는 MPEG 코덱을 사용한 표준화를 위한 일반적인 구조를 도시한다.

도 1c 는 모노/스테레오에 대한 역방향 호환가능 예를 나타낸다.

도 2a 는 역방향 호환가능 설계에 의한 인코딩 프로세스의 일 예를 나타낸다.

도 2b 는 씬 기반 데이터를 디코딩할 수 없는 종래 디코더에 대한 디코딩 프로세스의 일 예를 나타낸다.

도 3a 는 씬 기반 데이터를 핸들링할 수 있는 신규 디바이스에 의한 디코딩 프로세스의 일 예를 나타낸다.

도 3b는 일반적인 구성에 따른 오디오 신호 프로세싱의 방법 (M100) 의 플로우챠트를 나타낸다.

도 3c 는 일반적인 구성에 따른 오디오 신호 프로세싱을 위한 장치 (MF100) 의 블록 다이어그램을 나타낸다.

도 3d 는 또 다른 일반적인 구성에 따른 오디오 신호 프로세싱을 위한 장치 (A100) 의 블록 다이어그램을 나타낸다.

도 4 는 차수 0 및 1 의 구면 조화 기저 함수들의 크기들의 표면 메시 플롯들의 예들을 나타낸다.

도 5 는 차수 2 의 구면 조화 기저 함수들의 크기들의 표면 메시 플롯들의 예들을 나타낸다.

도 6 은 5.1 시스템에 대한 확성기 각도들의 표준 컴플라이언트 예를 나타낸다.

도 7 은 5.1 타겟 시스템에 대한 확성기 각도들의 균등하게 이격된 배열의 일 예를 나타낸다.

도 8 은 7.1 시스템에 대한 확성기 각도들의 표준 컴플라이언트 예를 나타낸다.

도 9 는 7.1 타겟 시스템에 대한 확성기 각도들의 균등하게 이격된 배열의 일 예를 나타낸다.

도 10a 는 일반적인 구성에 따른 오디오 신호 프로세싱의 방법 (M200) 의 플로우챠트를 나타낸다.

도 10b 는 방법 (M200) 의 구현 (M210) 의 플로우챠트를 나타낸다.

도 11a 는 일반적인 구성에 따른 오디오 신호 프로세싱을 위한 장치 (MF200) 의 블록 다이어그램을 나타낸다.

도 11b 는 장치 (MF100) 의 구현 (MF210) 의 블록 다이어그램을 나타낸다.

도 11c 는 또 다른 일반적인 구성에 따른 오디오 신호 프로세싱을 위한 장치 (A200) 의 블록 다이어그램을 나타낸다.

도 11d 는 장치 (A100) 의 구현 (A210) 의 블록 다이어그램을 나타낸다.

도 11e 는 또 다른 일반적인 구성에 따른 오디오 신호 프로세싱을 위한 장치 (P100) 의 블록 다이어그램을 나타낸다.

## 발명을 실시하기 위한 구체적인 내용

[0008]

맥락에 의해 명확히 제한되지 않는 한, 본 명세서에서 용어 "신호"는 유선, 버스, 또는 다른 송신 매체에 대해 표현되는 바와 같이 메모리 위치 (또는 메모리 위치들의 세트)의 상태를 포함하여, 그 보통의 의미들 중 어느 것을 표시하기 위해 사용된다. 맥락에 의해 명확히 제한되지 않는 한, 본 명세서에서 용어 "발생시키는" (generating)은 컴퓨팅하는 또는 그렇지 않으면 생성하는 (producing)과 같은, 그 보통의 의미들 중 어느 것을 표시하기 위해 사용된다. 맥락에 의해 명확히 제한되지 않는 한, 용어 "계산하는"은 복수의 값들로부터 컴퓨팅하는, 평가하는, 추정하는, 및/또는 선택하는과 같은 보통의 의미들 중 어느 것을 표시하기 위해 사용된다. 맥락에 명확히 제한되지 않는 한, 용어 "획득하는"은 본 명세서에서 계산하는, 도출하는, 수신하는 (예를 들어, 외부 디바이스로부터) 및/또는 취출하는 (예를 들어, 저장 엘리먼트들의 어레이로부터)과 같은 보통의 의미들 중 어느 것을 표시하기 위해 사용된다. 맥락에 의해 명확히 제한되지 않는 한, 본 명세서에서 용어 "선택하는"은 2 이상의 세트 중 적어도 하나 및 전부 보다 적은 것을 식별하는, 표시하는, 적용하는, 및/또는 사용하는과 같은 보통의 의미들 중 어느 것을 표시하기 위해 사용된다. 용어 "포함하는"이 본 상세한 설명 및 청구항들에서 사용되며, 이는 다른 엘리먼트들 또는 동작들을 배제하지 않는다. 용어 "에 기초하는" ("A 는 B 에 기초한다"와 같이)은, (i) "로부터 도출된" (예를 들어, "B 는 A 의 전구체이다"), (ii) "에 적어도 기초하는" (예를 들어, "A 는 적어도 B 에 기초한다"), 및 특정 맥락에서 적절하다면, (iii) "와 동등하다" (예를 들어, "A 는 B 와 동등하다")의 경우들을 포함하여, 그 보통의 의미들 중 어느 것을 표시하기 위해 사용된다. 유사하게, 용어 "에 응답하여"는 "에 적어도 응답하여"를 포함하여, 그 보통의 의미들 중 어느 것을 표시하기 위해 사용된다.

[0009]

다중 마이크로폰 오디오 센싱 디바이스의 마이크로폰의 "위치"에 대한 언급은 맥락에 의해 별도 표시된 사항이 없으면, 마이크로폰의 음향적 감응 면의 중앙의 위치를 표시한다. 용어 "채널"은 특정 맥락에 따라, 가끔은 신호 경로를 표시하기 위해 사용되고, 다른 때에는 그러한 경로에 의해 반송된 신호를 표시하기 위해 사용된다. 별도 표시된 사항이 없으면, 용어 "시리즈들"은 2 이상의 아이템들의 시퀀스를 표시하기 위해 사용된다. 용어 "로그 (logarithm)"는 십진 로그를 표시하기 위해 사용되지만, 그러한 다른 베이스들에 대한 연산의 확장들이 본 개시물의 범위 내에 있다. 용어 "주파수 성분"은 신호의 주파수 대역들 또는 주파수들의 세트, 예컨대 신호의 주파수 도메인 표현의 샘플 (예를 들어, 고속 푸리에 변환에 의해 생성되는 바와 같음) 또는 신호의 서브 대역 (예를 들어, 박크 스케일 (Bark scale) 또는 멜 스케일 (mel scale) 서브 대역) 중 하나를 표시하기 위해 사용된다.

[0010]

별도 표시된 사항이 없으면, 특정 피쳐를 갖는 장치의 동작의 임의의 개시는 또한 유사한 피쳐를 갖는 방법을 개시하는 것으로 명확히 의도되며 (그리고 그 역도 또한 같음), 특정 구성에 따른 장치의 동작의 임의의 개시는 또한 유사한 구성에 따른 방법을 개시하는 것으로 명확히 의도된다 (그리고 그 역도 또한 같음). 용어 "구성"은 특정 맥락에 의해 표시된 바와 같이 방법, 장치, 및/또는 시스템과 관련하여 사용될 수도 있다. 용어들 "방법", "프로세스", "절차", 및 "기법"은 특정 맥락에 의해 별도 표시된 사항이 없으면 일반적으로 교환 가능하게 사용된다. 용어 "장치" 및 "디바이스"는 또한 특정 맥락에 의해 별도 표시된 사항이 없으면 일반적으로 교환 가능하게 사용된다. 용어 "엘리먼트" 및 "모듈"은 통상 많은 구성의 부분을 표시하기 위해 사용된다. 맥락에 의해 명확히 제한되지 않는 한, 용어 "시스템"은 본 명세서에서 "공통 목적을 제공하기 위해 상호작용하는 엘리먼트들의 그룹"을 포함하여, 그 보통의 의미들 중 어느 것을 표시하기 위해 사용된다.

[0011]

문헌의 부분의 참조에 의한 임의의 통합은 또한 그 부분 내에 참조되는 용어들 또는 변수들의 정의들을 통합하도록 이해되어야 할 것이며, 그러한 정의들은 통합된 부분에서 참조된 임의의 도면들 뿐만 아니라 문헌에서 어디가에 나타난다. 초기에 정판사로 도입되지 않으면, 청구항 엘리먼트를 조정하기 위해 사용된 서수의 용어 (예를 들어, "제 1", "제 2", "제 3" 등)는 그 자체가 청구항 엘리먼트의 임의의 우선순위 또는 차수를 표시하는 것이 아니라, 단지 그 청구항 엘리먼트를 동일한 명칭을 갖는 또 다른 청구항 엘리먼트와 구별하는 것이다 (서수 용어의 사용이 아니라면). 맥락에 의해 명확히 제한되지 않는 한, 용어 "복수의" 및 "세트"는 본 명세서에서 1 보다 큰 정수 양을 표시하기 위해 사용된다.

[0012]

컨슈머 오디오에서의 기술의 현 상태는 채널 기반 서라운드 사운드를 사용한 공간 코딩이며, 이는 사전 특정된 포지션들에서 확성기들을 통해 플레이되는 것을 의미한다. 공간 오디오 코딩에 대한 또 다른 접근은 오브젝트 기반 오디오인데, 이는 공간에서의 오브젝트들의 위치 좌표들을 포함하는 연관된 메타데이터를 갖는 단일 오디오 오브젝트들에 대한 이산 펄스 코드 변조 (PCM) 데이터를 수반한다.

[0013]

공간 오디오 코딩에 대한 (예를 들어, 서라운드 사운드 코딩에 대한) 추가적인 접근은 씬 기반 오디오이며, 이

는 구면 조화 기저 함수들의 계수들을 사용하여 사운드 필드를 표현하는 것을 수반한다. 그러한 계수들은 또한 "구면 조화 계수들" 또는 SHC 로도 불린다. 도 1a 는 씬 기반 접근에 의한 간단한 인코딩 및 디코딩 프로세스를 도시한다. 본 예에서, 씬 기반 인코더 (SE10) 는 씬 기반 디코더 (SD10) 에서 송신되고 (및/또는 저장되고) 디코딩되는 SHC 의 디스크립션 (description) 을 생성하여 렌더링을 위한 SHC 를 수신한다 (예를 들어, SH 렌더러 (SR10) 에 의해). 그러한 인코딩은 대역폭 압축을 위한 하나 이상의 손실 또는 무손실 코딩 기법들, 예컨대 (예를 들어, 하나 이상의 코드북 인덱스들로의) 양자화, 에러 정정 코딩, 리던던시 코딩 등을 포함할 수도 있다. 부가적으로 또는 대안으로, 그러한 인코딩은 앰비소닉스 (Ambisonic) 포맷, 예컨대 B-포맷, G-포맷, 또는 고차원 앰비소닉스 (HOA) 로 오디오 채널들을 인코딩하는 것을 포함할 수도 있다. 일 반적으로, 인코더 (SE10) 는 (손실 또는 무손실 코딩 중 어느 하나에 대한) 계수들 및/또는 무관한 것들 중에서 리던던시들을 이용하는 기법들을 사용하여 SHC 를 인코딩할 수도 있다.

[0014]

공간 오디오 정보의 표준 비트 스트림으로의 인코딩 및 렌더러의 위치에서 스피커 지오메트리 및 음향 조건들에 적응 가능하고 불가지론적인 (agnostic) 후속 디코딩을 제공하는 것이 바람직할 수도 있다. 그러한 접근은 결국 재생을 위해 사용되는 특정 셋업에 관계없이 균일한 청취 경험의 목적을 제공할 수도 있다. 도 1b 는 MPEG 코덱을 사용하는 그러한 표준화를 위한 일반적인 구조를 도시한다. 본 예에서, 인코더 (MP10) 에 대한 입력 오디오 소스들은 다음 중 어느 하나 이상을 포함할 수도 있다, 예를 들어: 채널 기반 소스들 (예를 들어, 1.0 (모노포닉), 2.0 (스테레오포닉), 5.1, 7.1, 11.1, 22.2), 오브젝트 기반 소스들, 및 씬 기반 소스들 (예를 들어, 고차원 구면 조화, 앰비소닉스). 유사하게, 디코더 (및 렌더러)(MP20) 에 의해 생성된 오디오 출력은 다음 중 어느 하나 이상을 포함할 수도 있다, 예를 들어: 모노포닉, 스테레오포닉, 5.1, 7.1, 및/또는 22.2 확성기 어레이들을 위한 피드들; 불규칙하게 분산된 확성기 어레이들을 위한 피드들; 헤드폰들을 위한 피드들; 상호작용 오디오.

[0015]

오디오 자료가 (예를 들어, 컨텐츠 크리에이터에 의해) 한번 제작되고 후속하여 상이한 출력들 및 확성기 셋업들에 대해 디코딩되고 렌더링될 수 있는 포맷들로 인코딩되는 '한번 제작, 많은 사용 (creat-once, use-many)' 을 따르는 것이 또한 바람직할 수도 있다. 예를 들어, 헐리우드 스튜디오와 같은 컨텐츠 크리에이터는 통상 영화에 대한 사운트트랙을 한번 생성하면 각각의 가능한 확성기 구성에 대해 그것을 리믹스하기 위해 노력하지 않으려 한다.

[0016]

도 1b 의 표준화된 예에서 MPEG 인코더 (MP10) 에 대한 입력은 옵션으로 채널 기반 포맷, 오브젝트 기반 포맷, 및 씬 기반 포맷 중 하나이다. 씬 기반 포맷을 사용하는 것에 대해 많은 이점들이 있다. 하지만, 그러한 포맷을 사용하는 것의 하나의 가능한 단점은 기존 컨슈머 오디오 시스템들에 대한 역방향 호환성의 부족이다. SHC 를 사용할 때 역방향 호환성 부족은 SHC 가 PCM 데이터가 아니라는 사실에 기인한다. 신규 씬 기반 디코더가 없는 기존 수신기들 (또한 "레거시" 시스템들로도 불림) 은 그러한 신호들을 디코딩할 수 없고 그 프로그램을 플레이할 수 없게 된다. 본 개시물은 사운드 필드를 표현하기 위해 계층적 계수들의 세트 (예를 들어, SHC, 또는 다른 기저 함수 세트들의 계수들) 를 사용할 때 이러한 역방향 호환성 부족을 해결하기 위해 사용될 수도 있는 방법, 시스템, 및 장치를 기재한다.

[0017]

본 명세서에 기재된 접근은 사운드 필드들의 SHC 기반 표현의 사용에 있어서 잠재적인 단점에 대한 해결책을 제공한다. 이러한 해결책 없이, SHC 기반 표현은 수많은 기존의 레거시 플레이백 시스템들에서 기능성을 가질 수 없는 것에 의해 부과되는 상당한 단점으로 인해 전개에 어려움이 있게 된다.

[0018]

호환성을 보유하기 위해 레거시 모노포닉 플레이백 시스템들에 필요했던 것과 같이, 스테레오포닉 포맷이 최초로 도입되지 않았을 때에도 역방향 호환성은 관심사였다. 이 경우, 모노-스테레오 역방향 호환성은 매트릭싱을 사용하여 보유되었다. 스테레오 'M-중간' 및 'S-측면' 포맷은 단지 M 채널만을 사용하는 것에 의해 모노-가능 시스템들과의 호환성을 보유할 수 있다. 도 1c 에 나타낸 바와 같이, 스테레오 가능 시스템은 간단한  $2 \times 2$  매트릭스 연산을 수행하여 'L-좌측' 및 'R-우측' 채널들을 디코딩한다.

[0019]

M-S 신호는 (동일하게 발생되는) 위의 매트릭스의 역을 사용하는 것에 의해 L-R 신호로부터 컴퓨팅될 수 있다. 이러한 방식으로, 레거시 모노포닉 플레이어는 기능성을 보유하며, 스테레오포닉 플레이어는 좌측 및 우측 채널들을 정확하게 디코딩할 수 있다. 유사한 방식으로, 역방향 호환성을 보유하는 제 3 채널이 부가되어, 모노포닉 플레이어 및 스테레오포닉 플레이어의 기능성을 보존하고, 3-채널 플레이어의 기능성을 부가할 수 있다.

[0020]

부가 채널들을 포함하는 종래의 5.1 채널 기반 매트릭스형 오디오 접근의 확장들은 확장된 채널 포맷의 세트로서 5.1 샘플들을 포함하는 것에 의해 역방향 호환성의 문제를 우회할 수 있다. 예를 들어, 대부분의 기존

컨슈머 오디오 시스템들은 5.1 채널 입력을 수용한다. 역방향 호환가능 확장된 비트스트림에서, 5.1 샘플들은 레거시 시스템들에 의해 인식된 위치에 있으며, 여분의 채널들이, 예를 들어 채널 샘플들 모두를 포함하는 프레임 패킷의 확장된 부분에 위치된다. 그러한 접근의 또 다른 예는 돌비 디지털 플러스 비트스트림 (Dolby Laboratories, Inc., 캘리포니아 샌프란시스코)이며, 이는 5.1 믹스를 갖는 코어 패킷 및 7.1 믹스의 4 개의 서라운드 채널들을 갖는 확장 패킷을 포함한다. 대안으로, 디코더는 다수의 채널들에 대해 매트릭스 연산을 수행하는 것에 의해 5.1 채널 데이터를 결정하도록 구성될 수도 있다.

[0021] 오브젝트 기반 포맷에서의 역방향 호환성 이슈를 해결하기 위한 하나의 제안된 접근은 오브젝트들과 함께 다운 믹스된 5.1 채널 신호를 전송하는 것이다. 그러한 시나리오에서, 레거시 채널 기반 시스템들은 다운믹스된 5.1 채널 기반 오디오를 플레이하게 되지만, 보다 진보된 렌더러들은 5.1 오디오 및 개별 오디오 오브젝트들의 조합 또는 단지 개별 오브젝트들만을 사용하여 사운드 필드를 렌더링하게 된다. 불행히도, 역방향 호환성에 대한 그러한 접근은 2 개의 상이한 포맷들에서 동일한 오디오 정보를 전송하는 것에 의해 대역폭을 비효율적으로 사용할 수도 있다.

[0022] 본 개시물은 사운드 필드를 기술하는 계층적 엘리먼트들의 세트 (예를 들어, SHC 의 세트) 의 2 개의 부분들이 상이하게 프로세싱되는 시스템, 방법 및 장치의 기재들을 포함한다. 이러한 접근에 있어서, 엘리먼트들의 서브세트는 컨버팅되어 (예를 들어, 하기에 상세되는 변환 매트릭스를 사용하여) 다중 채널 오디오 신호들의 세트를 획득한다. 그러한 접근은 계층적 세트의 서브세트 ("기본 세트") 를 종래의 다중 채널 오디오 포맷과 호환가능한 오디오의 다중 채널들로 컨버팅하기 위해 사용될 수도 있다. 그러한 접근은 또한 역방향 호환성을 유지하기 위해 임의의 원하는 수의 채널들을 획득하도록 구성될 수도 있다. 그러한 역방향 호환가능 시스템의 양태들의 예들은 인코더 및 디코더 구조들의 양자에 대한 설명들과 함께, 도 2a, 도 2b 및 도 3a 의 시스템 다이어그램들에서 요약된다.

[0023] 도 2a 는 그러한 인코딩 접근의 일반적인 예를 나타낸다. 하나의 특정 사용은 (그의 역방향 호환성이 이슈가 아니었다면 사용될 수도 있는) 완전한 SHC 의 세트의 기본 세트를 종래의 다중 채널 오디오 포맷 (예를 들어, 변환 매트릭스 (TM10) 및 리인코더/트랜스코더 (TC10) 에 의해 생성된 바와 같음) 을 표현하는 다중 채널들로 컨버팅하는 것이다. SHC 의 세트의 나머지 (예를 들어, 서브세트의 부분이 아니었던 계수들) 는 컨버팅되지 않은 확장된 세트로서 간주되며 (예를 들어, 대역폭 압축을 위한 씬 기반 인코더 (SE10) 의 구현 (SE20) 에 의해) 송신 채널 (예를 들어, 유선 및/또는 무선 채널) 을 통한 송신을 위해, 및/또는 저장을 위해, 역방향 호환가능 다중 채널 오디오 인코딩된 신호들과 함께 (예를 들어, 본 명세서에 기재되고 비트스트림 멀티플렉서 (MX10) 에 의해 본 예에서 생성되는 패킷 또는 비트스트림에서) 별도로 인코딩될 수도 있다. 예를 들어, 이러한 인코딩된 비트들은 비트스트림의 확장된 부분 또는 프레임에 대한 패킷의 확장된 부분 (예를 들어, 사용자 정의된 부분) 으로 패킹될 수도 있다.

[0024] 수신단에서, 비트스트림은 디멀티플렉싱되고 (예를 들어, 디멀티플렉서 (DX10) 에 의해), 종래 디코더들 (예를 들어, 레거시 디코더 (LD10)) 만을 지원하는 레거시 시스템들은 역방향 호환가능 다중 채널 오디오 컨텐츠만을 사용하게 되고, 이로써 기능성을 보유하며, 도 2b 에 나타낸 바와 같은 비트스트림 또는 프레임 패킷의 확장된 부분들을 무시하게 된다.

[0025] 씬 기반 디코딩을 지원하는 수신기는 도 3a 에 나타낸 일반적인 예에 따라 디코딩 프로세스를 수행할 수도 있는 데, 이는 도 2b 에 나타낸 바와 같은 인코딩 접근과 반대이다. 채널들이 리인코딩되거나 트랜스코딩된다면, 디코딩의 중간 단계가 수행될 수도 있다. 예를 들어, 트랜스 디코더 (TC20) 는 역방향 호환가능 비트스트림을 다중 채널 오디오 신호들로 컨버팅하는데 사용될 수도 있다. 후속하여 역 변환 매트릭스 (IM10) 가 다중 채널 오디오 신호들을 계층적 세트의 원래 서브세트 (예를 들어, SHC 의 기본 세트) 로 컨버팅하는데 사용될 수도 있다. 계층적 세트의 나머지 (예를 들어, SHC 의 확장된 세트) 는 씬 기반 디코더 (SD10) 의 구현 (SD20) 에 의해 복구된다 (예를 들어, 비트스트림 또는 패킷의 확장된 부분으로부터). 이러한 방식으로, 완전한 계층적 세트 (예를 들어, SHC 의 완전한 세트) 는 SH 렌더러 (SR10) 에 의해 복구되고 프로세싱되어 사운드 필드 렌더링의 다양한 유형들이 발생하는 것을 허용할 수 있다.

[0026] 사운드 필드를 표현하기 위해 SHC 의 세트를 사용하는 것은 사운드 필드를 표현하기 위해 계층적 엘리먼트들의 세트를 사용하는 일반적인 접근의 특정 예이다. SHC 의 세트와 같은 계층적 엘리먼트들의 세트는, 하위 차수의 엘리먼트들의 기본 세트가 모델링된 사운드 필드의 전체 표현을 제공하도록 엘리먼트들이 차수화되는 세트이다. 세트가 상위 차수 엘리먼트들을 포함하기 위해 확장됨에 따라, 공간에서의 사운드 필드의 표현은 더 상세해지게 된다.

[0027] 소스 SHC (예를 들어, 도 3a에 나타낸 바와 같음)는 씬 기반 가능 레코딩 스튜디오에서 믹싱 엔지니어들에 의해 믹싱되는 바와 같은 소스 신호들일 수도 있다. 소스 SHC는 또한 확성기들의 서라운드 어레이에 의해 소닉 제시의 레코딩으로부터 또는 마이크로폰 어레이에 의해 캡처된 신호들로부터 생성될 수도 있다. PCM 스트림 및 연관된 위치 정보 (예를 들어, 오디오 오브젝트)의 SHC의 소스 세트로의 컨버전이 또한 고려된다.

[0028] 사운드 필드는 다음과 같은 수식을 사용하여 SHC에 관하여 표현될 수도 있다:

$$p_i(t, r_r, \theta_r, \varphi_r) = \sum_{\omega=0}^{\infty} \left[ 4\pi \sum_{n=0}^{\infty} j_n(kr_r) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta_r, \varphi_r) \right] e^{j\omega t}, \quad (1)$$

[0029]

[0030] 이 수식은 사운드 필드의 임의의 포인트  $\{r_r, \theta_r, \varphi_r\}$ 에서 압력  $p_i$ 은 SHC  $A_n^m(k)$ 에 의해 고유하게 표현될 수 있음을 나타낸다. 여기서, 과수  $k = \frac{\omega}{c}$ ,  $c$ 는 사운드의 속도 (~343 m/s),  $\{r_r, \theta_r, \varphi_r\}$ 는 레퍼런스의 포인트 (또는 관측 포인트),  $j_n(\cdot)$ 은 차수  $n$ 의 구면 베셀 함수이고,  $Y_n^m(\theta_r, \varphi_r)$ 는 차수  $n$  및 서브보더  $m$  (정도로서의 SHC 라벨  $n$  (즉, 대응하는 르장드르 다항식) 및 차수로서의  $m$ 의 일부 디스크립션들)의 구면 조화 기저 함수들이다. 각 괄호에서의 용어는 이산 푸리에 변환 (DFT), 이산 코사인 변환 (DCT), 또는 웨이블릿 변환과 같은, 다양한 시간 주파수 변환들에 의해 근사화될 수 있는 신호 (즉,  $S(\omega, r_r, \theta_r, \varphi_r)$ )의 주파수 도메인 표현이라는 것을 알 수 있다.

[0031] 도 4는 차수 0 및 1의 구면 조화 기저 함수들의 크기들의 표면 메시 플롯들의 예들을 나타낸다. 함수  $Y_0^0$ 의 크기는 구면이고 전방향성이다. 함수  $Y_1^{-1}$ 은 +y 및 -y 방향들에서 각각 확장하는 포지티브 및 네거티브 구면 로브(lobe)들을 갖는다. 함수  $Y_1^0$ 은 +z 및 -z 방향들에서 각각 확장하는 포지티브 및 네거티브 구면 로브들을 갖는다. 함수  $Y_1^1$ 은 +x 및 -x 방향들에서 각각 확장하는 포지티브 및 네거티브 구면 로브들을 갖는다.

[0032] 도 5는 차수 2의 구면 조화 기저 함수들의 크기들의 표면 메시 플롯들의 예들을 나타낸다. 함수들  $Y_2^{-2}$  및  $Y_2^2$ 는 x - y 평면에서 확장하는 로브들을 갖는다. 함수  $Y_2^{-1}$ 은 y - z 평면에서 확장하는 로브들을 가지며, 함수  $Y_2^1$ 은 x - z 평면에서 확장하는 로브들을 갖는다. 함수  $Y_2^0$ 은 +z 및 -z 방향들에서 확장하는 포지티브 로브들 및 x - y 평면에서 확장하는 도넛형 네거티브 로브를 갖는다.

[0033] 세트에서의 SHC의 총 수는 다양한 인자들에 의존할 수도 있다. 씬 기반 오디오에 대하여, 예를 들어, SHC의 총 수는 레코딩 어레이에서의 마이크로폰 트랜스듀서들의 수에 의해 제약될 수도 있다. 채널 및 오브젝트 기반 오디오에 대하여, SHC의 총 수는 가용 대역폭에 의해 결정될 수도 있다. 일 예에서, 각각의 주파수에 대하여 25개의 계수들을 수반하는 4차수 표현 (즉,  $0 \leq n \leq 4$ ,  $-n \leq m \leq +n$ )이 사용된다. 본 명세서에 기재된 접근으로 사용될 수도 있는 계층적 세트들의 다른 예들은 웨이블릿 변환 계수들의 세트들 및 다중 해상도 기저 함수들의 다른 세트들을 포함한다.

[0034]  $A_n^m(k)$ 는 사면체 또는 구면 마이크로폰 어레이와 같은 다양한 마이크로폰 어레이 구성들 중 어느 것을 사용하여 물리적으로 획득 (예를 들어, 레코딩) 되는 신호들로부터 도출될 수 있다. 이러한 형태의 입력은 제안된 인코더에 대한 씬 기반 오디오 입력을 표현한다. 한정이 아닌 예에서, SHC 인코더에 대한 입력들은 Eigenmike<sup>R</sup> (mh acoustics LLC, 캘리포니아 샌프란시스코)와 같은 마이크로폰 어레이의 상이한 출력 채널들이라고 가정된다. Eigenmike<sup>R</sup> 어레이의 일 예는 em32 어레이인데, 이는 직경 8.4 센티미터의 구의 표면 상에 배열된 32개의 마이크로폰들을 포함하여, 출력 신호들  $p_i(t)$  ( $i = 1$  내지 32)의 각각은 시간 샘플 ( $t$ )에서 마

이크로폰 (i)에 의해 기록되는 압력이다. 대안으로, SHC  $A_n^m(k)$ 는 사운드 필드의 채널 기반 또는 오브젝트 기반 디스크립션들로부터 도출될 수 있다. 예를 들어, 개별 오디오 오브젝트에 대응하는 사운드 필드에 대한 계수들  $A_n^m(k)$ 는 다음과 같이 수식화될 수도 있다.

$$[0035] A_n^m(k) = g(\omega)(-4\pi ik)h_n^{(2)}(kr_s)Y_n^{m*}(\theta_s, \varphi_s), \quad (2)$$

[0036] 식 중, i는  $\sqrt{-1}$ 이고,  $h_n^{(2)}(\cdot)$ 는 차수 n의 (제 2 종의) 구면 한켈 함수이고,  $\{r_r, \theta_r, \varphi_r\}$ 는 오브젝트의 위치이며,  $g(\omega)$ 는 주파수의 함수로서의 소스 에너지이다.

[0037] 주파수의 함수로서의 소스 에너지  $g(\omega)$ 를 알면, 각각의 PCM 오브젝트 및 그 위치를 SHC  $A_n^m(k)$ 로 컨버팅할 수 있다. 이 소스 에너지는 예를 들어 PCM 스트림에 대해 고속 푸리에 변환 (예를 들어, 256-, -512-, 또는 1024- 포인트 FET)을 수행하는 것 등에 의해 시간 주파수 분석 기법들을 사용하여 획득될 수도 있다. 추가

로, (상기가 선형이고 수직 분해이기 때문에) 각각의 오브젝트에 대한  $A_n^m(k)$  계수들이 가산적이라는 것이 나타날 수 있다. 이러한 방식으로, 다수의 PCM 오브젝트들은  $A_n^m(k)$  계수들에 의해 (예를 들어, 개별 오브젝트들에 대한 계수 벡터들의 합으로서) 표현될 수 있다. 본질적으로, 이러한 계수들은 사운드 필드에 관한 정보 (3D 좌표들의 함수로서의 압력)를 포함하며, 상기는 개별 오브젝트들로부터 관측 포인트  $\{r_r, \theta_r, \varphi_r\}$ 의 근처에서, 전반적인 사운드 필드의 표현으로의 변환을 나타낸다.

[0038] 당업자는 방사상 성분을 포함하지 않는 표현들과 같은 수식 (2)에 나타낸 표현 이외에 계수들  $A_n^m$  (또는 동등하게 대응하는 시간 도메인 계수들  $a_n^m$ )의 표현들이 사용될 수도 있음을 알게 된다. 당업자는 구면 조화 기저 함수들의 몇몇 다소 상이한 정의들이 알려져 있다는 것 (예를 들어, 실수, 복소수, 정규화 (예를 들어, N3D), 반정규화 (예를 들어, SN3D), Furse-Malham (FuMa 또는 FMH) 등)과, 결과적으로 수식 (1)(즉, 사운드 필드의 구면 조화 분해) 및 수식 (2)(즉, 포인트 소스에 의해 생성된 사운드 필드의 구면 조화 분해)는 문헌에서 다소 상이한 형태로 나타날 수도 있다는 것을 알게 된다. 본 기재는 구면 조화 기저 함수들의 임의의 특정 형태에 한정되지 않으며 실제로 다른 계층적 엘리먼트들의 세트에도 일반적으로 적용가능하다.

[0039] 도 3b는 시간 간격 동안 사운드 필드를 기술하는 복수의 기저 함수 계수들을 프로세싱하기 위한 방법 (M100)의 플로우챠트를 나타낸다. 방법 (M100)은 태스크들 (T100 및 T200)을 포함한다. 태스크 (T100)는 복수의 채널 신호들을 생성하기 위해 복수의 기저 함수 계수들의 제 1 그룹 ("기본 세트")에 대해 가역 변환을 수행하며, 복수의 채널 신호들의 각각은 공간의 대응하는 상이한 영역과 연관된다. 복수의 채널 신호들에 기초하여, 태스크 (T200)는 (A) 복수의 기저 함수 계수들의 제 2 그룹 ("확장된 세트")의 표현으로서, 제 2 그룹은 제 1 그룹과 상이한, 제 2 그룹의 표현, 및 (B) 제 2 그룹의 상기 표현과 별개인 복수의 채널 신호들의 표현을 포함하는 데이터 구조를 생성한다.

[0040] 방법 (M100)의 구현의 설계는 다중 채널 오디오로 (예를 들어, 종래 포맷으로) 컨버팅될 원래의 계층적 세트 (즉, 기본 세트)의 서브세트를 선택하는 단계를 포함할 수도 있다. 그러한 접근은 적어도 종래 5.1 서라운드/홈 씨어터 능력과의 혼합성을 유지하도록 구현될 것임이 예상될 수도 있다. 5.1 포맷에 대하여, 다중 채널 오디오 채널들은 전방 좌측 (L), 중앙 (C), 전방 우측 (R), 좌측 서라운드 (Ls), 우측 서라운드 (Rs), 및 저주파수 효과들 (LFE)이다. 7.1 포맷은 좌측 후방 (Lb) 및 우측 후방 (Rb) 채널들을 부가한다. 하기의 논의들에서, 5.1 포맷은 통상의 타겟 다중 채널 오디오 포맷으로서 사용될 것이며, 일 예의 접근이 설명될 것이다. 여기에서는 동일한 방법론이 다른 다중 채널 오디오 포맷들 (예를 들어, 7.1, 22.2 등)로 일반화되는 것이 고려되고 개시된다.

[0041] 기본 세트를 선택하는 다양한 접근들이 가능하다. 5 개의 신호들 (특정된 위치들로부터 전대역 오디오에 대응) 이 5.1 포맷 (플러스 LFE 신호 - 표준화된 위치를 갖지 않으며 5 개의 채널들을 로우패스 필터링하는 것에 의해 결정될 수 있음) 에서 이용가능하기 때문에, 하나의 접근은 5.1 포맷으로 컨버팅하기 위해 SHC 의 5 개를 사용한다는 것이다. 추가로, 대응하는 기저 함수들의 공간 배향들에 따라 기본 세트를 선택하는 것이 바람직할 수도 있다. 5.1 포맷은 예를 들어 2D 렌더링만을 할 수 있기 때문에, 기본 세트는 일부 수평 정보를 반송하는 SHC 만을 포함하는 것이 바람직할 수도 있다. 즉, 확성기 위치들을 포함하는 평면 (예를 들어, x - y 평면) 내에서 적어도 일 방향을 따라 에너지가 집중되는 기저 함수에 대응하는 SHC 를 선택하는 것이 바람직할 수도 있다. 도 4 에서 보여지는 바와 같이, 계수들  $A_1^0(k)$  은 예를 들어 수평 방향성에 대한 매우 적은 정보를 반송하며, 이로써 이러한 서브세트로부터 배제될 수 있다.  $A_2^1(k)$  의 실수 또는 허수부 중 어느 하나에 대해서도 마찬가지이다 (예를 들어,  $Y_2^1$  및  $Y_2^{-1}$  로서 도 5 에 나타낸 기저 함수들 참조). 기본 세트에 대하여 선택된 특정 계수들은, 예를 들어 구현에서 채택된 구면 조화 기저 함수들의 정의에 의존하여 달라질 수도 있다 (문헌에 다양한 정의들이 있다: 실수, 허수, 복소수 또는 조합들).

[0042] 그러한 방식으로, 5 개의 계수들은 컨버전을 위해 채택될 수 있다. 계수  $A_0^0(k)$  는 전방향성 정보를 반송하기 때문에, 이러한 계수를 항상 사용하는 것이 바람직할 수도 있다. 유사하게,  $A_1^{-1}(k)$  의 허수부 및  $A_1^1(k)$  의 실수부는 상당한 수평 방향성 정보를 반송하기 때문에, 이들을 포함하는 것이 바람직할 수도 있다. 마지막 2 개의 계수들에 대하여, 가능한 후보들은  $A_2^2(k)$  의 실수 및 허수부를 포함한다. 다양한 다른 조합들이 또한 가능하다. 예를 들어, 기본 세트는 단지 3 개의 계수들,  $A_0^0(k)$ ,  $A_1^1(k)$  의 실수부, 및  $A_1^{-1}(k)$  의 허수부만을 포함하도록 선택될 수도 있다. 기본 세트가 하기에서 논의되는 바와 같이 원하는 채널들의 수에 의존하여 그리고 또한 변환 매트릭스에 대한 가역성 제약에 의존하여 많거나 적은 계수들을 포함하는 것이 또한 가능하다.

[0043] 따라서 본 명세서에 기재된 바와 같이 5.1 호환가능 시스템에 대한 설계의 상세들은 상이한 타겟 포맷들에 대해 조정될 수도 있다. 일 예로서, 7.1 시스템들에 대한 호환성을 가능하게 하기 위해, 2 개의 여분의 오디오 컨텐츠 채널들이 호환가능한 요건들에 부가되고, 2 이상의 SHC 가 기본 세트에 부가되어, 결과의 변환 매트릭스가 정사각형이며, 이로써 가역가능일 수도 있다. 7.1 시스템들에 대한 대다수의 확성기 배열 (예를 들어, Dolby TrueHD) 은 여전히 수평 평면 상에 있기 때문에, SHC 의 선택은 여전히 높이 정보에 의한 것들을 배제할 수 있다. 이러한 식으로, 수평 평면 신호 렌더링은 렌더링 시스템에서 부가된 확성기 채널들로부터 이익을 얻을 수 있다. 높이 다이버시티에 의한 확성기들을 포함하는 시스템 (예를 들어, 9.1, 11.1 및 22.2 시스템들) 에서, 기본 세트에 높이 정보를 갖는 SHC 를 포함하는 것이 바람직할 수도 있다.

[0044] 다중 채널들로 일반화하고 있지만, 현재의 시장에서의 주요 강점은, 셋톱 박스들과 같은 레거시 컨슈머 오디오 시스템들의 기능성을 보장하는 '최소 공통 분모' 인 것으로서 5.1 채널에 대해서이다. 스테레오 및 모노와 같은 적은 수의 채널들에 대하여, 많은 종래 기술들에서의 기존 5.1 해결책들은 컨텐츠 정보를 유지하기 위해 다운믹스를 컨버팅하기에 충분하여야 한다. 이러한 경우들 (즉, 1.0 및 2.0) 은 중요하지 않은 것으로 고려되며 본 개시물에서 더 논의되지 않는다.

[0045] 일어날 수 있는 또 다른 이슈는, 얼마나 많은 에러가 (예를 들어, SHC 의) 기본 세트로부터 다중 채널 오디오로의 그리고 다시 기본 세트로의 순방향 및 역방향 컨버전에서 생성되는지이다. 태스크 (T100) 는 기저 함수 계수들의 기본 세트에 대해 가역 변환을 수행하여, 각각이 공간의 대응하는 상이한 영역 (예를 들어, 대응하는 상이한 확성기 위치) 과 연관되는 복수의 채널 신호들을 생성한다. SHC 의 기본 세트 (예를 들어, 상술한 선택된 5 개의 계수들) 를 5.1 포맷에서의 5 개의 전대역 오디오 신호들로 컨버팅하기 위해 가역 매트릭스를 적용하도록 태스크 (T100) 를 수행하는 것이 바람직할 수도 있다. 가역성에 대한 요망은 5 개의 전대역 오디오 신호들을 해상도의 손실이 약간 있거나 거의 없는 SHC 의 기본 세트로 역 컨버전하는 것을 허용하는 것이다.

[0046] 이러한 변환 매트릭스 (예를 들어, 도 2a 에 나타낸 바와 같은 매트릭스 (TM10)) 를 결정하기 위한 하나의 가능한 방법은 '모드 매칭' 으로서 알려진 동작이다. 여기서, 확성기 피드들은 각각의 확성기가 구면파를 생성

한다고 가정하는 것에 의해 컴퓨팅된다. 그러한 시나리오에서,  $\ell$  번째 확성기에 기인하는 소정의 포지션  $r, \theta, \varphi$ 에서 (주파수의 함수로서) 압력은 다음과 같이 주어진다.

$$P_l(\omega, r, \theta, \varphi) = g_l(\omega) \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n (-4\pi ik) h_n^{(2)}(kr_l) Y_n^{m*}(\theta_l, \varphi_l) Y_n^m(\theta, \varphi), \quad (3)$$

[0047]

[0048] 식 중,  $\{r_l, \theta_l, \varphi_l\}$  는  $\ell$  번째 확성기의 포지션을 나타내고,  $g_l(\omega)$ 는 (주파수 도메인에서)  $\ell$  번째 확성기의 확성기 피드이다. 따라서, 5 개의 스피커들 모두에 기인하는 총 압력  $P_t$ 은 다음과 같이 주어진다.

$$\begin{aligned} P_t(\omega, r, \theta, \varphi) &= \sum_{l=1}^5 g_l(\omega) \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n (-4\pi ik) h_n^{(2)}(kr_l) Y_n^{m*}(\theta_l, \varphi_l) Y_n^m(\theta, \varphi). \quad (4) \end{aligned}$$

[0049]

[0050] 5 개의 SHC 에 관하여 총 압력이 식에 의해 주어진다는 것을 또한 알게 된다.

$$P_t(\omega, r, \theta, \varphi) = 4\pi \sum_{n=0}^{\infty} j_n(kr) \sum_{m=-n}^n A_n^m(k) Y_n^m(\theta, \varphi). \quad (5)$$

[0051]

[0052] 위의 2 개의 식들을 동등화하는 것은 SHC 에 관하여 확성기 피드들을 수식화하기 위해 다음과 같이 변환 매트릭스를 사용하는 것을 허용한다:

$$\begin{bmatrix} A_0^0(\omega) \\ A_1^1(\omega) \\ A_1^{-1}(\omega) \\ A_2^2(\omega) \\ A_2^{-2}(\omega) \end{bmatrix} = -ik \begin{bmatrix} h_0^{(2)}(kr_1) Y_0^{0*}(\theta_1, \varphi_1) & h_0^{(2)}(kr_2) Y_0^{0*}(\theta_2, \varphi_2) & \dots & \dots & g_1(\omega) \\ h_1^{(2)}(kr_1) Y_1^{1*}(\theta_1, \varphi_1) & \dots & \dots & \dots & g_2(\omega) \\ \vdots & \vdots & \ddots & \ddots & g_3(\omega) \\ \vdots & \vdots & \ddots & \ddots & g_4(\omega) \\ \vdots & \vdots & \ddots & \ddots & g_5(\omega) \end{bmatrix}. \quad (6)$$

[0053]

[0054] 이러한 수식은 5 개의 확성기 피드들과 채택된 SHC 사이의 직접적인 관계가 있다는 것을 나타낸다. 변환 매트릭스는, 예를 들어 계수들이 서브세트 (예를 들어, 기본 세트)에서 사용되었었는지 그리고 구면 조화 기저 함수들의 정의가 사용되는지에 의존하여 달라질 수도 있다. 유사한 방식으로, 선택된 기본 세트로부터 상이한 채널 포맷 (예를 들어, 7.1, 22.2)으로 컨버팅하는 변환 매트릭스가 구성될 수도 있다. 위의 변환 매트릭스는 '모드 매칭' 기준으로부터 도출되었지만, 대안의 변환 매트릭스들이 압력 매칭, 에너지 매칭 등과 같은 다른 기준으로부터도 또한 도출될 수 있다.

[0055]

[0055] 위의 수식에서처럼 변환 매트릭스가 확성기 피드들로부터 SHC 로의 컨버전을 허용하면서, 그 매트릭스가 또한 가역성이 되도록, SHC 로 시작하여, 5 개의 채널 피드들을 산출할 수 있고, 그 후 디코더에서, 옵션으로 SHC 로 역 컨버팅할 수 있다 (어드밴스드 (즉, 비-레거시) 렌더러들이 존재할 때). 예를 들어, 변환 매트릭스가 정수 입력들에 응답하여 정수 출력들을 생성하는 것이 바람직할 수도 있다. 매트릭스의 가역성을 보장하기 위해 위의 프레임워크를 조정하는 다양한 방식들이 활용될 수 있다. 이들은 규칙화 기법들 (예를 들어, 주파수 의존 규칙화) 와 풀 랭크 (full rank) 및 잘 정의된 고유값들을 보장하기 위해 종종 작업하는 다른 매트릭스 조작 기법들과 같은 수학적 기법들을 포함하지만 이에 한정되지 않는다.

[0056]

[0056] 가역성을 보장하기 위한 기법들은 확성기들의 하나 이상의 포지션들을 변화시키는 것을 포함하지만 이에 한정되지 않는다. 그러한 기법들은, 예를 들어 5.1 시스템의 5 개의 확성기를 중 하나 이상의 포지션들을 조정하는 것을 포함하여, 이들이 특정된 각도 공차를 고수하도록 한다 (예를 들어, ITU-R BS.775-1 추천에 의해 특정된 바와 같음, International Telecommunication Union, 스위스 제네바). 도 6 은 사용될 수도 있는 5.1 시스템에 대한 확성기 각도들의 캠플라이언트 예를 나타낸다. T-설계를 고수하는 것들과 같은 트랜스듀서들의 규칙적인 스페이싱들이 통상적으로 행해진다.

- [0057] 대안으로, 태스크 (T100) 는 평면에서 균일하게 이격되는 확성기 각도들에 대응하는 채널 신호들의 세트로 기본 세트를 컨버팅하는 변환 매트릭스를 적용하도록 구현될 수도 있다. 가역성은 통상 그러한 변환 매트릭스에 대해 보장된다. 도 7 은 5.1 타겟 시스템에 대한 그러한 확성기 각도들의 배열의 예를 나타낸다. 그러한 균일하게 이격된 배열은 5.1 표준에 정확하게 순응하지 않을 수도 있지만, 적절한 역방향 호환가능 확성기 피드들의 세트를 여전히 제공하며, 또한 계수들의 기본 세트를 복구하기 위한 가역 변환 매트릭스를 회득하는 신뢰성있는 방법을 제공한다. 태스크 (T100) 의 7-채널 구현들을 위해, 도 8 및 도 9 는 7.1 시스템에 대한 확성기 각도들의 표준 컴플라이언트 세트 및 평면에서 균일하게 이격된 대응하는 확성기 각도들의 세트의 유사한 예를 각각 나타낸다.
- [0058] 5.1, 7.1 및 22.2 와 같은 적어도 하나의 저주파수 효과 (LFE) 채널을 포함하는 타겟 채널 포맷에 대하여, 채널 신호들의 세트를 생성하여 그러한 채널을 또한 포함하도록 방법 (M100) 을 구현하는 것이 바람직할 수도 있다. 일 예에서, 방법 (M100) 은 태스크 (T100) 에 의해 채널 신호들을 로우 패스 필터링하고 그 결과의 로우 패스 신호들을 합산하는 것에 의해 LFE 채널을 생성하는 태스크를 포함한다. 다중 LFE 채널들 (예를 들어, 22.2 포맷에 대한 좌측 및 우측 LFE 채널들) 이 필요한 경우에 대하여, 그러한 태스크는 각각의 타겟 위치에 공간적으로 가장 근접하는 로우 패스 신호들을 합산하는 것에 의해 이러한 채널들을 생성하도록 구현될 수도 있다. 이러한 방식으로 생성된 LFE 채널은 중복이지만, 여전히 레거시 LFE 채널에 대해 적당한 신호를 제공하며, 그러한 접근은 또한 변환 매트릭스에 대한 임의의 변화들을 회피한다.
- [0059] 모든 조작 후, 수정된 매트릭스가 정확하고 및/또는 수용가능한 확성기 피드들을 확실히 생성하는 것을 보장하기 위해 변환 매트릭스에 의해 렌더링된 채널 신호들을 테스트 (예를 들어, 청감적으로 (psycho-acoustically)) 하는 것이 바람직할 수도 있다. 가역성이 보존되는 한, SHC 에 대한 정확한 디코딩을 보장하는 반대의 문제는 통상적으로 이슈가 아니다. 기본 세트 (예를 들어, SHC 서브세트) 와 종래 다중 채널 오디오 사이에서 정보를 허용하는 매트릭스가 도출될 수 있고 또한 (다중 채널 오디오의 충실도를 감소시키지 않는) 조작 후, 또한 가역성인 다소 수정된 매트릭스가 또한 형성될 수 있는 것이 통상 충분하다.
- [0060] 복수의 채널 신호들에 기초하여, 태스크 (T200) 는, (A) 복수의 기저 함수 계수들 (예를 들어, "확장된 세트") 의 제 2 그룹의 표현으로서, 제 2 그룹은 제 1 그룹과 상이한, 제 2 그룹의 표현, 및 (B) 제 2 그룹의 상기 표현과 별개인 복수의 채널 신호들의 표현을 포함하는 데이터 구조를 생성한다. 데이터 구조는 시간 내의 구조 (예를 들어, 하나 이상의 라이브 비트스트림들 또는 패킷들의 스트림들을 포함) 일 수도 있고 및/또는 공간 내의 구조 (예를 들어, 저장 매체 상에 저장된 하나 이상의 비트스트림 또는 베퍼 내에 저장된 하나 이상의 패킷들을 포함) 일 수도 있다.
- [0061] 수식 (6) 에 나타낸 바와 같이, 태스크 (T100) 는 주파수 도메인에서 채널 신호들을 생성하기 위해 각각의 주파수에서 기본 세트에 변환 매트릭스를 적용하도록 구현될 수도 있다. 방법 (M100) 은 이러한 주파수 도메인 채널 신호들의 각각에 대해 역 주파수 변환 (예를 들어, 역 FFT) 을 수행하여 대응하는 시간 도메인 채널 신호 (예를 들어, 선형 PCM 스트림) 를 생성하는 작업을 포함하도록 구현될 수도 있다.
- [0062] 채널 신호들은 HDMI 인터페이스 (고정세 멀티미디어 인터페이스, HDMI Licensing, LLC, 캘리포니아 씨니밸리) 에 의해 선형 PCM 스트림들로서 반송될 수도 있다. 또 다른 예에서, 채널 신호들은 광학 디스크, 예컨대 CD, DVD, DVD-오디오, 또는 블루-레이 디스크 상에 선형 PCM 스트림들로서 저장될 수도 있다. 블루-레이 디스크 (예를 들어, 블루-레이 디스크 어플리케이션 정의 BD-J 에 따르는 광학 데이터 저장 매체, 2005년 3월, 블루-레이 디스크 협회, www-dot-blu-raydisc-dot-com) 는 MPEG-2 전송 스트림을 포함하는 파일 'zzzzz.m2ts' 를 포함할 수도 있으며, 여기서 'zzzzz' 는 AV 스트림 파일을 클립 형성 파일과 연관시키는 5 자리 숫자이다. 스트림 파일 'zzzzz.m2ts' 은 다중의 기초적인 오디오 스트림들을 포함할 수도 있다. 태스크 (T200) 는 LPCM 스트림들로서 태스크 (T100) 에 의해 생성된 채널 신호들의 시간 도메인 베타들을 포함하는 그러한 스트림 파일을 생성하도록 구현될 수도 있다.
- [0063] 대역폭 및/또는 저장 리소스들의 사용을 감소시키기 위해, LPCM 채널 스트림들을 압축하도록 태스크 (T200) 를 구현하는 것이 바람직할 수도 있다. SHC 의 기본 세트의 복구성을 보장하기 위해, 그러한 압축을 무손실 압축 스킴을 사용하여 수행하는 것이 바람직할 수도 있다. 일 예에서, 태스크 (T200) 는 DVD-오디오에 순응하는 비트스트림을 생성하기 위해 머리디언 무손실 압축 (MLP) 을 사용하여 PCM 스트림들을 인코딩하도록 구현된다. 또 다른 예에서, 태스크 (T200) 는 PCM 스트림들을 MPEG-4 SLS (Scalable to Lossless) 무손실 확장을 사용하여 AAC 코어 코덱으로 인코딩하도록 구현된다. 추가 예에서, 태스크 (T200) 는, MLP의 개선된 베타를 사용하여 7.1 오디오를 인코딩하는 Dolby TrueHD 및/또는 무손실 옵션으로 7.1 오디오를 또한 인코딩하는 DTS-

HD 마스터 오디오 (DTS, Inc., 캘리포니아 칼라바사스) 를 사용하여, PCM 스트림들을 무손실로 인코딩하는 것에 의해 생성된 기초적인 오디오 스트림들을 포함하는 스트림 파일 (예를 들어, 상술한 블루-레이-컴플라이언트 m2ts 파일) 을 생성하도록 구현된다.

[0064] 그렇지 않으면 태스크 (T200) 는 (예를 들어, 트랜스코더 (TC10) 의 동작에 의해 도 2a 에 예시된 바와 같은) 채널 신호들을 기술하는 역방향 호환가능 코딩된 신호들로 채널 신호들을 인코딩하도록 구현된다. 그러한 인코딩은 채널 신호들에 대해 손실 압축 스킴을 수행하는 것을 포함할 수도 있다. 태스크 (T200) 의 그러한 구현들에서 사용될 수도 있는 역방향 호환가능 코덱들의 예들은, AC3 (예를 들어, ATSC 표준에 기재된 바와 같음: 디지털 오디오 압축, Doc. A/52:2012, 2012년 3월 23일, 어드밴스드 텔레비전 시스템 위원회, 워싱턴 DC; 또한 ATSC A/52 또는 돌비 디지털로도 칭하며, 손실 MDCT 압축을 사용함), Dolby TrueHD (손실 압축 옵션들을 포함), DTS-HD 마스터 오디오 (손실 압축 옵션들을 또한 포함), 및 MPEG 서라운드 (MPS, ISO/IEC 14496-3, 고효율 어드밴스드 오디오 코딩 또는 HeAAC 로도 칭하며, 여기에서는 5.1 포맷 신호의 6 채널들이 렌더러에서 채널들의 나머지의 합성을 허용하는, 대응하는 측면 정보, 예컨대 두 귀의 레벨차, 두 귀의 시간차 및 내부 채널 일관성으로, 모노 또는 스테레오 PCM 스트림으로 다운믹스됨) 를 포함한다. 이러한 코덱들을 통상 입력으로서 시간 도메인 채널 신호들 (예를 들어, 선형 PCM 스트림들의 세트) 을 수용한다. 그러한 트랜스코딩은 채널 신호들이 많은 컨슈머 디바이스들 및 셋톱 박스들에 있는 AC3 디코더들과의 역방향 호환성을 보유하는 것을 허용한다. 예를 들어, 인코딩된 채널들은 소망하는 대응 채널 기반 포맷에 순응하는 패킷의 대응 부분으로 패킹될 수도 있다.

[0065] 그러한 경우들에서, 방법 (M100) 은 트랜스코딩된 비트스트림 (예를 들어, AC3 패킷의 "옥스데이터 (auxdata)" 부분) 의 하나 이상의 확장된 부분들에서 송신 (및/또는 저장) 을 위해 별도로 계층적 세트의 나머지 (예를 들어, 서브세트의 부분이 아니었던 SHC) 를 인코딩하도록 구현될 수도 있다. 그러한 동작은 예를 들어 비트스트림 멀티플렉서 (MX10) 의 동작에 의해 도 2a 에 도시되어 있다. 그러한 방법 (M100) 의 구현은 다중 채널 신호를 상이한 각각의 포맷 (예를 들어, AC3 트랜스코딩 및 Dolby TrueHD 트랜스코딩) 으로 각각 코딩하는, 2 이상의 상이한 트랜스코딩 동작들을 포함하여, 송신 및/또는 저장을 위한 2 개의 상이한 역방향 호환가능 비트스트림들을 생성하는 것이 또한 가능하다.

[0066] 위에서 주시된 바와 같이, 돌비 디지털 플러스 비트스트림은 5.1 믹스에 의한 코어 (A) 패킷 및 7.1 믹스의 4 서라운드 채널들에 의한 확장 (B) 패킷을 포함한다. 그러한 코덱과의 호환성을 위한 방법 (M100) 의 구현은 7 개의 SHC 의 기본 세트 (각각의 주파수에 대하여) 를 그러한 A 및 B 패킷들로의 인코딩을 위한 7 채널들로 변환하도록 구성된 태스크 (T100) 의 구현을 포함할 수도 있다. 그러한 경우, 태스크 (T200) 는 확장된 세트에서 8 번째 및 그 상위 SHC 를 포함하도록 구현될 수도 있고, 또는 대안으로 확장된 세트에서 6 번째 및 그 상위 SHC 를 포함할 수도 있으며, 확장된 세트는 하나 이상의 추가 확장 패킷들로 인코딩될 수도 있다.

[0067] 상이한 길이들의 시간 간격들에 대응하는 패킷들을 생성하도록 태스크 (T200) 를 구현하는 것이 바람직할 수도 있다. 예를 들어, 과도 (transient) 이벤트들을 오디오 입력으로 인코딩하기 위해 짧은 간격이 사용될 수도 있고, 정적 신호들을 인코딩하기 위해 긴 간격이 사용될 수도 있다. 그러한 일 예에서, 태스크 (T200) 는 역방향 호환성을 위해 PCM 채널들에서 일정한 시간 간격을 인코딩하도록 구현되지만, 확장된 세트에서 상이한 간격 길이들을 인코딩하도록 구현된다. 그러한 또 다른 예에서, 태스크 (T200) 는 대응 시간 간격의 지속 기간의 표시와 함께, 상이한 길이들의 간격들을 PCM 스트림들로 인코딩하도록 구현된다. AC-3 코딩이 사용되는 경우에 대하여, 예를 들어, 태스크 (T200) 는 256 및 512 샘플들의 2 개의 블록 길이 옵션들 사이에서 스위칭하도록 구현될 수도 있다.

[0068] 도 2a 에 도시된 바와 같이, 계수들의 확장된 세트는 송신 및/또는 저장을 위해 인코딩된 채널 신호들과 연관되기 (예를 들어, 멀티플렉싱되기) 전에 인코딩될 수도 있다. 그러한 인코딩은 대역폭 압축을 포함할 수도 있다. 예를 들어, 확장된 세트는 하나 이상의 손실 또는 무손실 코딩 기법들, 예컨대 양자화 (예를 들어, 하나 이상의 코드북 인덱스들로), 여러 정정 코딩, 리던던시 코딩 등, 및/또는 패킷화 (packetization) 를 적용하는 것에 의해 (예를 들어, 인코더 (SE20)) 에 의해) 인코딩될 수도 있다. 부가적으로 또는 대안으로, 그러한 인코딩은 앰비소닉스 포맷, 예컨대 B-포맷, G-포맷, 또는 고차원 앰비소닉스 (HOA) 로 인코딩하는 것을 포함할 수도 있다. 일 예에서, 계수들의 확장된 세트는 HOA B-포맷으로 인코딩되고, 그 후 B-포맷 신호들이 어드밴스드 오디오 코딩 (AAC; 예를 들어 ISO/IEC 14496-3:2009, 표준화를 위한 "information technology -- Coding of audio-visual objects -- Part 3: Audio" 에 정의된 바와 같음, 국제 표준화 기구, 스위스 제네바) 을 사용하여 인코딩된다. 확장된 세트에 대해 (예를 들어, 인코더 (SE20) 에 의해) 수행될 수도 있는 SHC 의 세트들을 인코딩하는 다른 방법들의 기재들은, 예를 들어 미국 특허출원 공보 제 2012/0155653 A1 호 (Jax

등) 및 제 2012/0314878 A1 호 (Daniel 등)에서 찾을 수 있다. 계수들의 확장된 세트는, 예를 들어 상이한 차수들의 계수들 간 차들 및/또는 상이한 시간들에서 동일한 차수의 계수들 간 차들로서 인코딩될 수도 있다.

[0069] 도 1b에 나타낸 바와 같은 MPEG 인코더 (MP10)는 본 명세서에서 기재된 바와 같이, (예를 들어, 스트리밍, 브로드캐스트, 멀티캐스트, 저장, 및/또는 미디어 마스터링 (예를 들어, CD, DVD, 및/또는 Blu-Ray™ 디스크의 마스터링)을 위한 비트스트림을 생성하기 위한) 방법 (M100)의 구현을 수행하는 것이 바람직할 수도 있다.

[0070] 역방향 호환가능 데이터 구조로부터 기저 함수 계수들의 완전한 세트를 획득하는 것은, 채널 신호들을 획득하기 위해 임의의 중재 코딩 계층들 및/또는 도메인 변환들을 (예를 들어, 디멀티플렉서 (DX10), 트랜스 디코더 (TC20), 및 썬 기반 디코더 (SD20)을 통해) 제거하고, 그 후 상술한 가역 변환의 역 (도 3a에 나타낸 바와 같은 매트릭스 (IM10))을 채널 신호들에 적용하는 것에 의해, 유사한 방식으로 수행될 수도 있다. 도 10a는 일반적인 구성에 따라 시간 간격 동안 사운드 필드를 기술하는 복수의 기저 함수 계수들을 획득하는 방법 (M200)의 플로우챠트를 나타낸다. 상술한 바와 같이, 복수의 기저 함수 계수들의 각각은 수직 기저 함수들의 세트의 고유한 하나 (예를 들어, 구면 조화 기저 함수들의 세트의 고유한 하나)에 대응할 수도 있다.

[0071] 방법 (M200)은 태스크들 (T300 및 T400)을 포함한다. 데이터 구조로부터, 태스크 (T300)는 (A) 복수의 기저 함수 계수들의 제 2 그룹 ("확장된 세트")의 표현, 및 (B) 제 2 그룹의 상기 표현과 별개인 복수의 채널 신호들의 표현을 획득한다. 복수의 채널 신호들의 서브세트의 각각은 공간의 대응하는 상이한 영역 (예를 들어, 5.1 또는 7.1 포맷의 공연 방향들)과 연관되며, 복수의 채널 신호들은 또한 LFE 채널들과 같은, 하나 이상의 비방향성 (또는 전방향성) 신호들을 포함할 수도 있다. 주파수 도메인에서 복수의 채널 신호들의 서브세트를 획득하기 위해 LPCM 스트림들의 세트에 대해 FFT를 수행하는 것이 바람직하도록 태스크 (T300)를 구현하는 것이 바람직할 수도 있다. 태스크 (T400)는 복수의 기저 함수 계수들의 제 1 그룹 ("기본 세트")을 생성하기 위해 복수의 채널 신호들의 서브세트에 대해 변환을 수행하며, 제 1 그룹은 제 2 그룹과 상이하다.

계수들의 확장된 세트의 표현이 데이터 구조에 제시되는 것을 검출할 시에만, 그리고 그렇지 않으면 채널 신호들을 계수들의 기본 세트로 변환하지 않으면서 역방향 호환가능 방식으로 채널 신호들을 렌더링하기 위해 (예를 들어, 확성기 피드들로서) 방법 (M200)을 실행하는 것이 바람직할 수도 있다.

[0072] 방법 (M200)은 완전한 기저 함수 계수들의 세트를 기본 세트가 복구되었던 복수의 채널 신호들과 연관된 하나와 상이한 확성기 지오메트리로 렌더링하도록 구현될 수도 있다. 도 10b는 태스크 (T500)를 포함하는 방법 (M200)의 그러한 구현 (M210)의 플로우챠트를 나타낸다. 복수의 기저 함수 계수들에 기초하여, 태스크 (T500)는 제 2 복수의 채널 신호들을 생성한다. 이 경우, 복수의 채널 신호들의 서브세트의 각 신호는 공연 방향들의 세트의 대응하는 상이한 하나와 연관되고, 제 2 복수의 채널 신호들의 각각은 3 차원 공간에 걸친 방향들의 세트의 대응하는 상이한 하나와 연관된다. 예를 들어, 제 2 복수의 채널 신호들은 위의 레퍼런스 평면인 또는 그렇지 않으면 높이 정보를 제공하는 확성기 위치들에 대응하는 하나 이상의 채널들을 포함할 수도 있다. 일 예에서, 태스크 (T500)는 유사한 방식으로 획득되는 특정 확성기 지오메트리에 대한 렌더링 매트릭스를, 계수들의 기본 세트를 상술한 역방향 호환가능 채널 신호들로 컨버팅하기 위한 변환 매트릭스의 도출에 적용하도록 구현된다 (예를 들어, 수식 (3) 내지 (6)을 참조하지만, 통상적으로 어떠한 가역성 제약은 없음).

[0073] 따라서, 상기는 계층적 엘리먼트들의 세트 (예를 들어, SHC의 세트)와 다중 오디오 채널들 사이에서 컨버팅하는 무손실 메커니즘을 표현한다. 다중 채널 오디오 신호들이 추가적인 코딩 노이즈를 받게되지 않는 한 어떠한 에러들도 발생되지 않는다. 그 신호들이 코딩 노이즈를 받게 되는 경우 (예를 들어, 손실 압축 동작을 통해), SHC에 대한 컨버전은 에러들을 발생할 수도 있다. 하지만, 계수들의 값들을 모니터링하고 적절한 액션을 취하여 그 효과를 감소시키는 것에 의해 이 에러들을 설명하는 것이 가능하다. 이러한 방법들은 SHC 표현에서 고유 리던던시를 포함하여, SHC의 특성들을 고려할 수도 있다.

[0074] 도 3c는 시간 간격 동안 사운드 필드를 기술하는 복수의 기저 함수 계수들을 프로세싱하는 장치 (MF100)의 블록 다이어그램을 나타낸다. 장치 (MF100)는 복수의 채널 신호들을 생성하기 위해 복수의 기저 함수들의 제 1 그룹 ("기본 세트")에 대해 가역 변환을 수행하기 위한 수단 (F100)을 포함하며, 복수의 채널 신호들의 각각은 공간의 대응하는 상이한 영역과 연관된다 (예를 들어, 변환 매트릭스 (TM10) 및 태스크 (T100)의 구현들을 참조하여 본 명세서에서 기재된 바와 같음). 장치 (MF100)는 또한, 복수의 채널 신호들에 기초하여, (A) 복수의 기저 함수 계수들의 제 2 그룹 ("확장된 세트")의 표현으로서, 제 2 그룹은 제 1 그룹과 상이한, 제 2 그룹의 표현, 및 (B) 제 2 그룹의 상기 표현과 별개인 복수의 채널 신호들의 표현을 포함하는 데

이터 구조를 생성하기 위한 수단 (F200) 을 포함한다 (예를 들어, 태스크 (T200), 트랜스코더 (TC10), 인코더 (SE20), 및 멀티플렉서 (MX10) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음).

[0075] 도 3d 는 시간 간격 동안 사운드 필드를 기술하는 복수의 기저 함수 계수들을 프로세싱하기 위한 장치 (A100) 의 블록 다이어그램을 나타낸다. 장치 (A100) 는 복수의 채널 신호들을 생성하기 위해 복수의 기저 함수 계수들의 제 1 그룹 ("기본 세트") 에 대해 가역 변환을 수행하도록 구성된 계산기 (100) 를 포함하며, 복수의 채널 신호들의 각각은 공간의 대응하는 상이한 영역과 연관된다 (예를 들어, 변환 매트릭스 (TM10) 및 태스크 (T100) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음). 장치 (A100) 는 또한, 복수의 채널 신호들에 기초하여, (A) 복수의 기저 함수 계수들의 제 2 그룹 ("확장된 세트") 의 표현으로서, 제 2 그룹은 제 1 그룹과 상이한, 제 2 그룹의 표현, 및 (B) 제 2 그룹의 상기 표현과 별개인 복수의 채널 신호들의 표현을 포함하는 데 이터 구조를 생성하도록 구성된 데이터 포맷터 (200) 를 포함한다 (예를 들어, 태스크 (T200), 트랜스코더 (TC10), 인코더 (SE20), 및 멀티플렉서 (MX10) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음). 포맷터 (200) 는 패킷화기 (packetizer) 를 포함할 수도 있다.

[0076] 도 11a 는 일반적인 구성에 따라 시간 간격 동안 사운드 필드를 기술하는 복수의 기저 함수 계수들을 획득하기 위한 장치 (MF200) 의 블록 다이어그램을 나타낸다. 장치 (MF200) 는, 데이터 구조로부터, (A) 복수의 기저 함수 계수들의 제 2 그룹 ("확장된 세트") 의 표현, 및 (B) 제 2 그룹의 상기 표현과 별개인 복수의 채널 신호들의 표현을 획득하기 위한 수단 (F300) 을 포함한다 (예를 들어, 태스크 (T300), 디멀티플렉서 (DX10), 트랜스코더 (TC20), 및 디코더 (SD20) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음). 장치 (MF200) 는 또한 복수의 기저 함수 계수들의 제 1 그룹 ("기본 세트") 를 생성하기 위해 복수의 채널 신호들의 서브세트에 대해 변환을 수행하기 위한 수단 (F400) 을 포함하며, 제 1 그룹은 제 2 그룹과 상이하다 (예를 들어, 태스크 (T400) 및 역변환 매트릭스 (IM10) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음). 도 11b 는 복수의 기저 함수 계수들에 기초하여 제 2 복수의 채널 신호들을 생성하기 위한 수단 (F500) 을 포함하는 장치 (MF200) 의 구현 (MF210) 의 플로우챠트를 나타낸다 (예를 들어, 태스크 (T500) 및 렌더러 (SR10) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음).

[0077] 도 11c 는 일반적인 구성에 따른 시간 간격 동안 사운드 필드를 기술하는 복수의 기저 함수 계수들을 획득하기 위한 장치 (A200) 의 블록 다이어그램을 나타낸다. 장치 (200) 는, 데이터 구조로부터, (A) 복수의 기저 함수 계수들의 제 2 그룹 ("확장된 세트") 의 표현, 및 (B) 제 2 그룹의 상기 표현과 별개인 복수의 채널 신호들의 표현을 획득하도록 구성된 데이터 추출기 (300) 를 포함한다 (예를 들어, 태스크 (T300), 디멀티플렉서 (DX10), 트랜스코더 (TC20), 및 디코더 (SD20) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음). 추출기 (300) 는 디패킷화기 (depacketizer) 를 포함할 수도 있다. 장치 (A200) 는 또한 복수의 기저 함수 계수들의 제 1 그룹 ("기본 세트") 를 생성하기 위해 복수의 채널 신호들의 서브세트에 대해 변환하도록 구성된 계산기 (400) 를 포함하며, 제 1 그룹은 제 2 그룹과 상이하다 (태스크 (T400) 및 역변환 매트릭스 (IM10) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음). 도 11d 는 복수의 기저 함수 계수들에 기초하여 제 2 복수의 채널 신호들을 생성하도록 구성된 렌더러 (500) 를 포함하는 장치 (A200) 의 구현 (A210) 의 플로우챠트를 나타낸다 (예를 들어, 태스크 (T500) 및 렌더러 (SR10) 의 구현들을 참조하여 본 명세서에 기재된 바와 같음).

[0078] 추가 예에서, 방법 (M100) 은 기저 함수 계수들의 2 이상의 세트들을 결합하는 태스크를 포함하도록 구현되며, 각각의 세트는 대응하는 시간 간격 동안 대응하는 사운드 필드를 기술하며, 결합된 세트를 입력으로서 태스크 (T100) 에 제공한다. 예를 들어, 그러한 태스크는 결합된 사운드 필드의 디스크립션을 생성하기 위해 복수의 계수들의 세트들을 부가하도록 (예를 들어, SHC 벡터들의 계수 벡터 부가를 수행하도록) 구현될 수도 있다.

기저 함수 계수들의 각각의 세트는, 예를 들어 오디오 오브젝트를 표현할 수도 있다. 그러한 기저 함수 계수들의 세트 (예를 들어, 하나의 오브젝트에 대한 SHC 벡터) 는 기저 함수 계수들의 또 다른 세트 (오브젝트들의 또 다른 것에 대해) 보다 높은 차수 (예를 들어, 긴 길이) 를 가질 수도 있다. 예를 들어, 포어그라운드 (foreground) 에서의 오브젝트 (예를 들어, 주연 배우의 보이스) 는 백그라운드에서의 오브젝트 (예를 들어, 사운드 효과) 보다 높은 차수 세트로 표현될 수도 있다.

[0079] 도 11e 는 일반적인 구성에 따른 장치 (P100) 에 대한 블록 다이어그램을 나타낸다. 장치 (P100) 는 대응하는 데이터 구조를 생성하기 위해 시간 간격 동안 사운드 필드를 기술하는 복수의 기저 함수 계수들에 대하여 본 명세서에 기재된 바와 같은 방법 (M100) 의 구현을 수행하도록 구성된다. 장치 (P100) 는 또한 프로세서 (P10) 에 의해 생성된 데이터 구조를 저장하도록 구성된 메모리 (ME10) 를 포함한다.

- [0080] 본 명세서에 기재된 바와 같은 방법 (예를 들어, 방법 (M100 또는 M200)) 을 음향 신호들을 수신하도록 구성된 2 이상의 마이크로폰들의 어레이를 갖는 포터블 오디오 센싱 디바이스 내에서 수행하는 것이 바람직할 수도 있다. 그러한 어레이를 포함하도록 구현될 수도 있고, 오디오 레코딩 및/또는 보이스 통신 어플리케이션들에 대해 사용될 수도 있는 포터블 오디오 센싱 디바이스의 예들은, 전화기 핸드셋 (예를 들어, 셀룰러 전화기 핸드셋); 유선 또는 무선 헤드셋 (예를 들어, 블루투스 헤드셋); 핸드헬드 오디오 및/또는 비디오 레코더 (예를 들어, 캡코더); 오디오 및/또는 비디오를 레코딩하도록 구성된 개인용 미디어 플레이어; 개인용 휴대 단말기 (PDA) 또는 다른 핸드헬드 컴퓨팅 디바이스; 및 노트북 컴퓨터, 랩탑 컴퓨터, 넷북 컴퓨터, 태블릿 컴퓨터, 또는 다른 포터블 컴퓨팅 디바이스를 포함한다.
- [0081] 본 명세서에 개시된 방법 및 장치는 일반적으로 모바일 또는 그렇지 않으면 그러한 어플리케이션들의 포터블 경우들 및/또는 원거리 필드 소스들 (far-field sources)로부터의 신호 성분들의 센싱을 포함하여, 임의의 트랜시빙 및/또는 오디오 센싱 어플리케이션에 적용될 수도 있다. 예를 들어, 본 명세서에 기재된 구성들의 범위는 코드 분할 다중 액세스 (CDMA) 오버-더-에어 (over-the-air) 인터페이스를 채용하도록 구성된 무선 텔레포니 통신 시스템에 상주하는 통신 디바이스들을 포함한다. 그럼에도 불구하고, 당업자는 본 명세서에 기재된 바와 같은 피쳐들을 갖는 방법 및 장치는 당업자에게 알려진 기술들의 넓은 범위를 채용하는 임의의 다양한 통신 시스템, 예컨대 유선 및/또는 무선 (예를 들어, CDMA, TDMA, FDMA, 및/또는 TD-SCDMA) 송신 채널들을 통한 보이스 오버 IP (VoIP) 를 채용하는 시스템들에 상주할 수도 있다는 것을 이해하게 된다.
- [0082] 본 명세서에 개시된 통신 디바이스들 (예를 들어, 스마트폰들, 태블릿 컴퓨터들) 은 패킷 스위칭되는 네트워크들 (예를 들어, VoIP 와 같은 프로토콜들에 따라 오디오 송신들을 수행하도록 배열된 유선 및/또는 무선 네트워크들) 및/또는 회로-스위칭되는 네트워크들에 사용하기 위해 적응될 수도 있다는 것이 명확히 고려되고 여기에 개시된다. 본 명세서에 개시된 통신 디바이스들은 협대역 코딩 시스템들 (예를 들어, 약 4 또는 5 킬로헤르츠의 오디오 주파수 범위를 인코딩하는 시스템들) 에 사용하기 위해 및/또는 전대역 광대역 코딩 시스템들 및 스플릿 대역 광대역 코딩 시스템들을 포함하는 광대역 코딩 시스템들 (예를 들어, 5 킬로헤르츠보다 큰 오디오 주파수들을 인코딩하는 시스템들) 에 사용하기 위해 적응될 수도 있다는 것이 또한 명확히 고려되고 여기에 개시된다.
- [0083] 상기 기재된 구성들의 제시는 당업자가 본 명세서에 개시된 방법들 및 다른 구조들을 제작하거나 사용하는 것을 가능하게 하기 위해 제공된다. 본 명세서에 나타내고 기재된 플로우차트들, 블록 다이어그램들 및 다른 구조들은 단지 예들일 뿐이며, 또한 개시물의 범위 내에서 이러한 구조들의 다른 변형들이 있다. 이러한 구성들에 대한 다양한 변형들이 가능하며, 본 명세서에 제시된 일반적인 원리들도 또한 다른 구성들에 적용될 수도 있다. 따라서, 본 개시물은 위에 나타낸 구성들에 한정하려고 의도된 것이 아니라 오히려 원래의 개시물의 부분을 형성하는 출원된 첨부된 청구항들에서 포함하는, 본 명세서에서 임의의 방식으로 개시된 신규한 피쳐들 및 원리들과 일치하는 최광 범위와 부합되도록 하는 것이다.
- [0084] 당업자는 정보 및 신호들이 임의의 다양한 상이한 기술들 및 기법들을 사용하여 표현될 수도 있음을 이해하게 된다. 예를 들어, 위의 상세한 설명 전체에 걸쳐 언급될 수도 있는 데이터, 명령, 커맨드, 정보, 신호, 비트, 및 심볼은 전압, 전류, 전자파, 자기장 또는 자기 입자들, 광학 입자들, 또는 그 조합에 의해 표현될 수도 있다.
- [0085] 본 명세서에 개시된 바와 같은 구성의 구현을 위한 중요한 설계 요건들은, 특히 계산 집중적 어플리케이션들, 예컨대 시청각적 정보 또는 압축된 오디오의 플레이백 (예를 들어, 본 명세서에서 식별된 예들 중 하나와 같은, 압축 포맷에 따라 인코딩된 스트림 또는 파일) 또는 광대역 통신들을 위한 어플리케이션들 (예를 들어, 12, 16, 44.1, 48, 또는 192 kHz 등의 8 킬로헤르츠보다 높은 샘플링 레이트에서의 보이스 통신들) 에 대하여, 프로세싱 지연 및/또는 계산 복잡도 (통상적으로 초당 수백만의 명령들에서 측정되는, 또는 MIPS) 를 최소화하는 것을 포함할 수도 있다.
- [0086] 다중 마이크로폰 프로세싱 시스템의 목표는, 전반적인 노이즈 감소에 있어서 10 내지 12 dB 을 달성하는 것, 원하는 스피커의 이동 동안 보이스 레벨 및 컬러를 보존하는 것, 공격적인 노이즈 제거 대신 백그라운드로 노이즈가 이동되게 하는 퍼셉션 (perception) 을 획득하는 것, 스피치의 탈반향 (dereverberation), 및/또는 더 공격적인 노이즈 감소를 위한 포스트 프로세싱의 옵션을 가능하게 하는 것을 포함할 수도 있다.
- [0087] 본 명세서에 개시된 바와 같은 장치 (예를 들어, 장치 (A100, A200, A210, MF100, MF200, MF210, 및 P100) 중 어느 것) 는 의도된 어플리케이션에 적합하다고 여겨지는, 소프트웨어와 및/또는 펌웨어와 하드웨어의 임의의 조합으로 구현될 수도 있다. 예를 들어, 그러한 장치의 엘리먼트들은, 예를 들어 칩셋에서 2 이상의 칩들

중에 또는 동일한 칩 상에 상주하는 전자 및/또는 광학 디바이스들로서 제조될 수도 있다. 그러한 디바이스의 일 예는 트랜지스터들 또는 로직 게이트들과 같은 로직 엘리먼트들의 고정된 또는 프로그램가능 어레이이며, 이러한 엘리먼트들 중 어느 것은 하나 이상의 그러한 어레이들로서 구현될 수도 있다. 장치의 엘리먼트들 중 임의의 2 이상 또는 심지어 전부 동일한 어레이 또는 어레이들 내에서 구현될 수도 있다. 그러한 어레이 또는 어레이들은 하나 이상의 칩들 내에 (예를 들어, 2 이상의 칩들을 포함하는 칩셋 내에) 구현될 수도 있다.

[0088] 본 명세서에 개시된 장치의 다양한 구현들 중 하나 이상의 엘리먼트들 (예를 들어, 장치 (A100, A200, A210, MF100, MF200, MF210, 및 P100) 중 어느 것) 은 또한, 하나 이상의 고정된 또는 프로그램가능한 로직 엘리먼트들의 어레이들, 예컨대 마이크로프로세서들, 임베딩된 프로세서들, IP 코어들, 디지털 신호 프로세서들, FPGA들 (필드 프로그램가능 게이트 어레이들), ASSP들 (주문형 표준 제품들), 및 ASIC들 (주문형 접적 회로들) 상에서 실행하도록 배열된 하나 이상의 명령들의 세트로서 부분적으로 또는 전체로 구현될 수도 있다. 본 명세서에 개시된 장치의 구현의 다양한 엘리먼트들 중 어느 것은 또한 하나 이상의 컴퓨터들 (예를 들어, "프로세서"라고도 또한 칭하는, 명령들의 시퀀스들 또는 하나 이상의 세트들을 실행하도록 프로그래밍된 하나 이상의 어레이들을 포함하는 머신들) 로서 구현될 수도 있고, 이러한 엘리먼트들 중 임의의 2 이상 또는 심지어 전부가 동일한 그러한 컴퓨터 또는 컴퓨터들 내에서 구현될 수도 있다.

[0089] 본 명세서에 개시된 바와 같은 프로세싱을 위한 프로세서 또는 다른 수단 (예를 들어, 프로세서 (P10)) 은, 예를 들어 칩셋에서 하나 이상의 칩들 중에 또는 동일한 칩 상에 상주하는 하나 이상의 전자 및/또는 광학 디바이스들로서 제조될 수도 있다. 그러한 디바이스의 일 예는 고정된 또는 프로그램가능한 로직 엘리먼트들의 어레이, 예를 들어 트랜지스터들 또는 로직 게이트들이며, 이러한 엘리먼트들 중 어느 것은 하나 이상의 그러한 어레이들로서 구현될 수도 있다. 그러한 어레이 또는 어레이들은 하나 이상의 칩들 내에 (예를 들어, 2 이상의 칩들을 포함하는 칩셋 내에) 구현될 수도 있다. 그러한 어레이들의 예들은 고정된 또는 프로그램가능한 로직 엘리먼트들의 어레이들, 예를 들어 마이크로프로세서들, 임베딩된 프로세서들, IP 코어들, DSP들, FPGA들, ASSP들, 및 ASIC들을 포함한다. 본 명세서에 개시된 프로세서 또는 다른 수단은 하나 이상의 컴퓨터들 (예를 들어, 명령들의 시퀀스들 또는 하나 이상의 세트들을 실행하도록 프로그래밍된 하나 이상의 어레이들을 포함하는 머신들) 또는 다른 프로세서들로서 구현될 수도 있다. 본 명세서에 기재된 바와 같은 프로세서는, 프로세서가 임베딩되는 디바이스 또는 시스템 (예를 들어, 오디오 센싱 디바이스) 의 또 다른 동작에 관련하는 태스크와 같은, 본 명세서에 기재된 바와 같은 오디오 코딩 절차와 직접적으로 관련되지 않은 명령들의 다른 세트들을 실행하거나 태스크들을 수행하기 위해 사용되는 것이 가능하다. 또한 본 명세서에 개시된 방법의 부분은 오디오 센싱 디바이스의 프로세서에 의해 수행되고 방법의 또 다른 부분은 하나 이상의 다른 프로세서들의 제어 하에서 수행되는 것이 가능하다.

[0090] 본 명세서에 개시된 구성들과 관련하여 기재되는 다양한 예시적인 모듈들, 로직 블록들, 회로들, 및 테스트들 그리고 다른 동작들은 전자 하드웨어, 컴퓨터 소프트웨어, 또는 양자의 조합으로서 구현될 수도 있다는 것을 당업자는 알게 된다. 그러한 모듈들, 로직 블록들, 회로들 및 동작들은 범용 프로세서, 디지털 신호 프로세서 (DSP), ASIC 또는 ASSP, FPGA 또는 다른 프로그램가능 로직 디바이스, 이산 게이트 또는 트랜지스터 로직, 이산 하드웨어 컴포넌트들 또는 본 명세서에 개시된 바와 같은 구성을 생성하도록 설계된 그 임의의 조합으로 구현되거나 수행될 수도 있다. 예를 들어, 그러한 구성은 하드 와이어 회로로서, 주문형 접적 회로 내에 제조된 회로 구성으로서, 또는 비휘발성 스토리지 내에 로딩되는 펌웨어 프로그램 또는 머신 판독가능 코드로서 데이터 저장 매체로부터 또는 이 매체 내에 로딩되는 소프트웨어 프로그램으로서, 적어도 부분적으로 구현될 수도 있으며, 그러한 코드는 범용 프로세서 또는 다른 디지털 신호 프로세싱 유닛과 같은 로직 엘리먼트들의 어레이에 의해 실행가능한 명령들이다. 범용 프로세서는 마이크로프로세서일 수도 있지만, 대안적으로, 프로세서는 임의의 종래 프로세서, 제어기, 마이크로제어기 또는 상태 머신일 수도 있다. 프로세서는 또한 컴퓨팅 디바이스들의 조합, 예를 들어 DSP 및 마이크로프로세서의 조합, 복수의 마이크로프로세서들, DSP 코더와 협력하는 하나 이상의 마이크로프로세서들, 또는 임의의 다른 그러한 구성으로서 구현될 수도 있다. 소프트웨어 모듈은 비일시적 저장 매체, 예컨대 RAM (random-access memory), ROM (read-only memory), 비휘발성 RAM (NVRAM), 예컨대 플래시 RAM, 소거가능 프로그램가능 ROM (EPROM), 전기적 소거가능 프로그램가능 ROM (EEPROM), 레지스터들, 하드 디스크, 탈착가능 디스크, 또는 CD-ROM 에 상주할 수도 있고 또는 종래에 알려진 저장 매체의 임의의 다른 형태에 상주할 수도 있다. 예시적인 저장 매체는 프로세서에 커플링되고, 그러한 프로세서는 저장 매체로부터 정보를 판독하고 저장 매체에 정보를 기록할 수 있다. 대안으로, 저장 매체는 프로세서에 통합될 수도 있다. 프로세서 및 저장 매체는 ASIC 에 상주할 수도 있다. ASIC 은 사용자 단말기에 상주할 수도 있다. 대안으로, 프로세서 및 저장 매체는 사용자 단말기에 이산 컴포넌트들로서 상주할 수도 있다.

[0091]

본 명세서에 개시된 다양한 방법들 (예를 들어, 방법들 (M100, M200 및 M210) 중 어느 것)은 프로세서와 같은 로직 엘리먼트들의 어레이에 의해 수행될 수도 있고, 본 명세서에 기재된 장치의 다양한 엘리먼트들은 그러한 어레이 상에서 실행하도록 설계된 모듈들로서 구현될 수도 있음을 유의한다. 본 명세서에서 사용된 바와 같이, 용어 "모듈" 및 "서브 모듈"은 임의의 방법, 장치, 디바이스, 유닛, 또는 소프트웨어, 하드웨어 또는 펌웨어 형태로 컴퓨터 명령들 (예를 들어, 로직적 표출)을 포함하는 컴퓨터 판독가능 저장 매체를 지칭할 수 있다.

다중 모듈들 또는 시스템들은 하나의 모듈 또는 시스템으로 결합될 수 있고, 하나의 모듈 또는 시스템은 동일한 기능들을 수행하기 위해 다중 모듈들 또는 시스템들로 분리될 수 있음을 이해하게 된다. 소프트웨어 또는 다른 컴퓨터 실행가능 명령들에서 구현될 때, 프로세스의 엘리먼트들은 본질적으로, 예컨대 루틴들, 프로그램들, 오브젝트들, 컴포넌트들, 데이터 구조들 등으로, 관련된 테스크들을 수행하는 코드 세그먼트들이다.

용어 "소프트웨어"는 소스 코드, 어셈블리 언어 코드, 머신 코드, 바이너리 코드, 펌웨어, 매크로코드, 마이크로코드, 로직 엘리먼트들의 어레이에 의해 실행가능한 명령들의 어느 하나 이상의 세트들 또는 시퀀스들, 및 그러한 샘플들의 조합을 포함하는 것을 이해해야 한다. 프로그램 또는 코드 세그먼트들은 송신 매체 또는 통신 링크를 통해 반송파에서 구현되는 컴퓨터 데이터 신호에 의해 송신되거나 프로세서 판독가능 저장 매체에 저장될 수 있다.

[0092]

본 명세서에 개시된 방법들, 스킵들 및 기법들의 구현들은 또한, 로직 엘리먼트들의 어레이를 포함하는 머신 (예를 들어, 프로세서, 마이크로프로세서, 마이크로제어기, 또는 다른 유한 상태 머신)에 의해 판독가능하고 및/또는 실행가능한 명령들의 하나 이상의 세트들로서 (예를 들어, 본 명세서에 열거된 바와 같은 하나 이상의 컴퓨터 판독가능 매체에서) 유형으로 구현될 수도 있다. 용어 "컴퓨터 판독가능 매체"는 휘발성, 비휘발성, 탈착가능 및 비탈착가능 매체를 포함하여, 정보를 저장하거나 전달할 수 있는 임의의 매체를 포함할 수도 있다. 컴퓨터 판독가능 매체의 예들은 전자 회로, 반도체 메모리 디바이스, ROM, 플래시 메모리, 소거 가능 ROM (EROM), 플로피 디스크 또는 다른 자기 스토리지, CD-ROM/DVD 또는 다른 광학 스토리지, 하드 디스크, 광섬유 매체, 라디오 주파수 (RF) 링크, 또는 원하는 정보를 저장하는데 사용될 수 있고 액세스될 수 있는 임의의 다른 매체를 포함한다. 컴퓨터 데이터 신호는 전자 네트워크 채널들, 광섬유들, 에어, 전자기, RF 링크들 등과 같은 송신 매체를 통해 전파할 수 있는 임의의 신호를 포함할 수도 있다. 코드 세그먼트들은 인터넷 또는 인트라넷과 같은 컴퓨터 네트워크들을 통해 다운로드될 수도 있다. 임의의 경우, 본 개시물의 범위가 그러한 실시형태들에 의해 한정되는 것으로서 이해되지 않아야 한다.

[0093]

본 명세서에 기재된 방법들의 각각은 하드웨어에서 직접 구현될 수도 있고, 프로세서에 의해 실행되는 소프트웨어에서 구현될 수도 있으며 또는 둘의 조합에서 구현될 수도 있다. 본 명세서에 개시된 방법의 구현의 통상의 어플리케이션에서, 로직 엘리먼트들의 어레이 (예를 들어, 로직 게이트들)는 방법의 다양한 태스크들 중 하나, 하나 초과, 또는 전부를 수행하도록 구성된다. 태스크들의 하나 이상 (가능하다면 전부)은 또한, 로직 엘리먼트들의 어레이를 포함하는 머신 (예를 들어, 프로세서, 아미크로프로세서, 마이크로제어기, 또는 다른 유한 상태 머신)에 의해 판독가능하고 및/또는 실행가능한, 컴퓨터 프로그램 제품 (예를 들어, 디스크들, 플래시 또는 다른 비휘발성 메모리 카드들, 반도체 메모리 칩들 등과 같은 하나 이상의 저장 매체)에서 구현되는 코드 (예를 들어, 명령들의 하나 이상의 세트들)로서 구현될 수도 있다. 본 명세서에 개시된 방법의 구현의 태스크들은 또한 하나 초과의 그러한 어레이 또는 머신에 의해 수행될 수도 있다.

이들 또는 다른 구현들에서, 태스크들은 그러한 통신 능력들을 갖는 셀룰러 전화기 또는 다른 디바이스와 같은 무선 통신들을 위한 디바이스 내에서 수행될 수도 있다. 그러한 디바이스는 (예를 들어, VoIP 와 같은 하나 이상의 프로토콜들을 사용하여) 회로 스위칭되는 및/또는 패킷 스위칭되는 네트워크들과 통신하도록 구성될 수도 있다. 예를 들어, 그러한 디바이스는 인코딩된 프레임들을 수신하고 및/또는 송신하도록 구성된 RF 회로를 포함할 수도 있다.

[0094]

본 명세서에 개시된 다양한 방법들은 포터블 통신 디바이스, 예컨대 핸드셋, 헤드셋 또는 포터블 디지털 보조기 (PDA)에 의해 수행될 수도 있으며, 본 명세서에 기재된 다양한 장치는 그러한 디바이스 내에 포함될 수도 있다는 것이 명확히 개시된다. 통상 실시간 (예를 들어, 온라인) 어플리케이션은 그러한 모바일 디바이스를 사용하여 행해지는 전화 대화이다.

[0095]

하나 이상의 예시적인 실시형태들에서, 본 명세서에 기재된 동작들은 하드웨어, 소프트웨어, 펌웨어 또는 그 임의의 조합에서 구현될 수도 있다. 소프트웨어에서 구현되는 경우, 그러한 동작들은 하나 이상의 명령들 또는 코드로서 컴퓨터 판독가능 매체 상에 저장되거나 이를 통해 송신될 수도 있다. 용어 "컴퓨터 판독가능 매체"는 컴퓨터 판독가능 저장 매체 및 통신 (예를 들어, 송신) 매체의 양자를 포함한다. 한정이 아닌 예로서, 컴퓨터 판독가능 저장 매체는 저장 엘리먼트들의 어레이, 예컨대 반도체 메모리 (동적 또는 정적 RAM,

ROM, EEPROM, 및/또는 플래시 RAM 을 제한없이 포함할 수도 있음) 또는 강유전성, 자기저항식, 오보닉(ovonic), 중합성, 또는 상변화 메모리; CD-ROM 또는 다른 광학 디스크 스토리지; 및/또는 자기 디스크 스토리지 또는 다른 자기 스토리지 디바이스들을 포함할 수 있다. 그러한 저장 매체는 컴퓨터에 의해 액세스될 수 있는 데이터 구조들 또는 명령들의 형태로 정보를 저장할 수도 있다. 통신 매체는 하나의 장소에서 다른 곳으로 컴퓨터 프로그램의 전달을 용이하게 하는 임의의 매체를 포함하는, 컴퓨터에 의해 액세스될 수 있는 데이터 구조들 또는 명령들의 형태로 원하는 프로그램 코드를 반송하는데 사용될 수 있는 임의의 매체를 포함할 수 있다. 또한, 임의의 접속은 컴퓨터 판독가능 매체를 적절하게 일컫는다. 예를 들어, 소프트웨어가 웹사이트, 서버, 또는 동축 케이블, 광섬유 케이블, 트위스트된 페어, 디지털가입자 라인(DSL), 또는 무선 기술, 예컨대 적외선, 라디오, 및/또는 마이크로파를 사용하는 다른 원격 소스로부터 송신되면, 광축 케이블, 광섬유 케이블, 트위스트된 페어, DSL, 또는 무선 기술, 예컨대 적외선, 라디오, 및/또는 마이크로파가 매체의 정의에 포함된다. 본 명세서에서 사용되는 바와 같이, 디스크(disk) 및 디스크(disc) 는, 컴팩 디스크(CD), 광학 디스크, 디지털 다기능 디스크(DVD), 플로피 디스크, 및 블루-레이 디스크™(블루-레이 디스크 협회, 캐리포니아 유니버셜 시티)을 포함하며, 여기서 디스크(disk)들은 보통 데이터를 자기적으로 재생시키는 한편, 디스크(disc)들은 레이저를 이용하여 데이터를 광학적으로 재생시킨다. 상기의 조합들이 또한 컴퓨터 판독가능 매체들의 범위 내에 포함되어야 한다.

[0096]

본 명세서에 기재된 바와 같이 음향 신호 프로세싱 장치(예를 들어, 장치(A100 또는 MF100))는, 소정의 동작들을 제어하기 위해서 스피치 입력을 수용하는 전자 디바이스에 통합될 수도 있고, 또는 그렇지 않으면 통신 디바이스들과 같은 백그라운드 노이즈들로부터의 원하는 노이즈들의 분리로부터 이익을 얻을 수도 있다. 많은 어플리케이션들은 다중 방향들로부터 유발되는 백그라운드 사운드들로부터 분명한 원하는 사운드를 강화시키거나 분리시키는데 유리할 수도 있다. 그러한 어플리케이션들은 보이스 인식 및 검출, 스피치 강화 및 분리, 보이스 활성화 제어 등과 같은 능력들을 통합하는 전자 또는 컴퓨팅 디바이스들에서 휴면 멀티 인터페이스들을 포함할 수도 있다. 그러한 음향 신호 프로세싱 장치는 제한된 프로세싱 능력들만을 제공하는 디바이스들에 적합하도록 구현하는 것이 바람직할 수도 있다.

[0097]

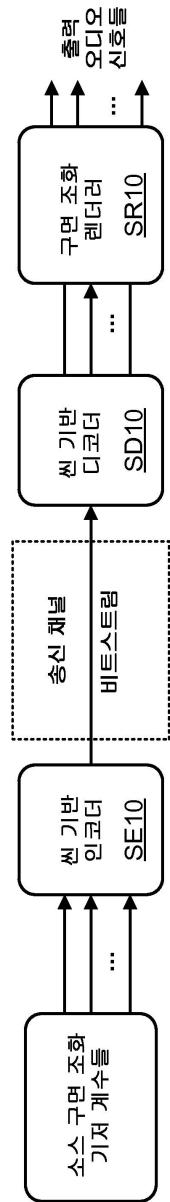
본 명세서에 기재된 모듈들, 엘리먼트들 및 디바이스들의 다양한 구현들의 엘리먼트들은, 예를 들어 칩셋에서 2 이상의 칩들 중에 또는 동일한 칩 상에 상주하는 전자 및/또는 광학 디바이스들로서 제조될 수도 있다. 그러한 디바이스의 일 예는, 트랜지스터들 또는 게이트들과 같은 로직 엘리먼트들의 고정된 또는 프로그램가능한 어레이이다. 본 명세서에 기재된 장치의 다양한 구현들의 하나 이상의 엘리먼트들은 또한 마이크로프로세서들, 임베딩된 프로세서들, IP 코어들, 디지털 신호 프로세서들, FPGA들, ASSP들, 및 ASIC들과 같은 로직 엘리먼트들의 하나 이상의 고정된 또는 프로그램가능한 어레이들 상에서 실행하도록 배열된 명령들의 하나 이상의 세트들로서 전체로 또는 부분으로 구현될 수도 있다.

[0098]

본 명세서에 기재된 바와 같이 장치의 구현의 하나 이상의 엘리먼트들은, 장치가 임베딩되는 디바이스 또는 시스템의 또 다른 동작과 관련된 태스크와 같은, 장치의 동작과 직접 관련되지 않는 다른 명령들의 세트를 실행하거나 태스크들을 수행하기 위해 사용되는 것이 가능하다. 그러한 장치의 구현의 하나 이상의 엘리먼트들은 공통인 구조(예를 들어, 상이한 시간들에서 상이한 엘리먼트들에 대응하는 코드의 부분들, 상이한 시간들에서 상이한 엘리먼트들에 대응하는 태스크들을 수행하기 위해 실행된 명령들의 세트, 또는 상이한 시간들에서 상이한 엘리먼트들에 대한 동작들을 수행하는 전자 및/또는 광학 디바이스들의 배열을 실행하기 위해 사용된 프로세서)를 갖는 것이 가능하다.

도면

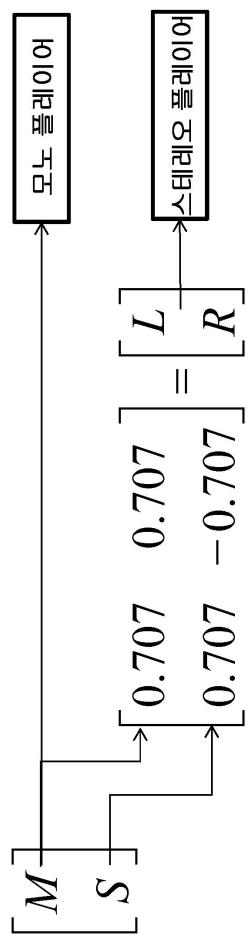
도면 1a



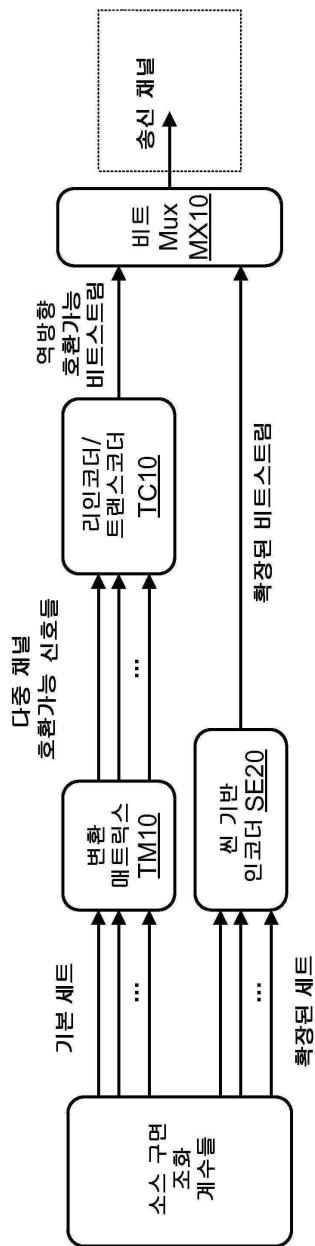
도면 1b



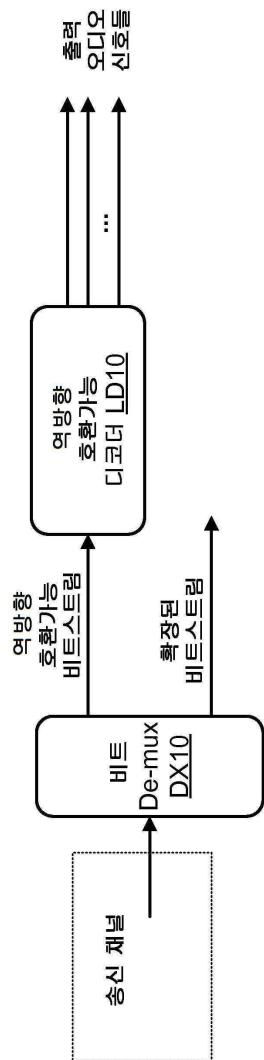
## 도면 1c



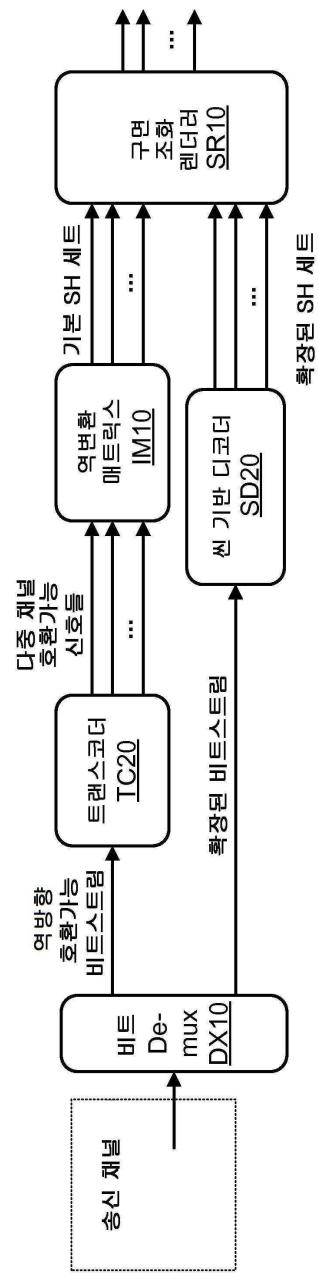
도면2a



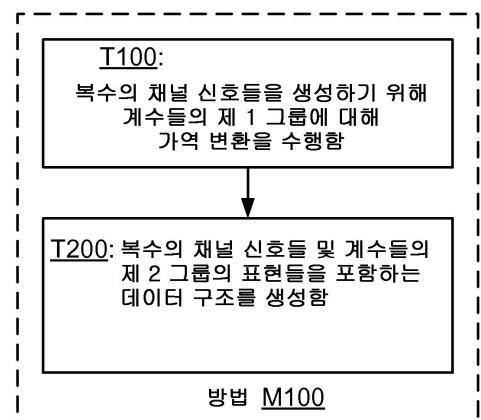
도면2b



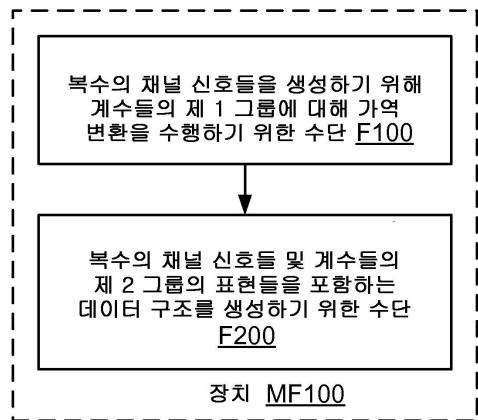
도면3a



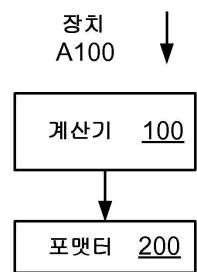
도면3b



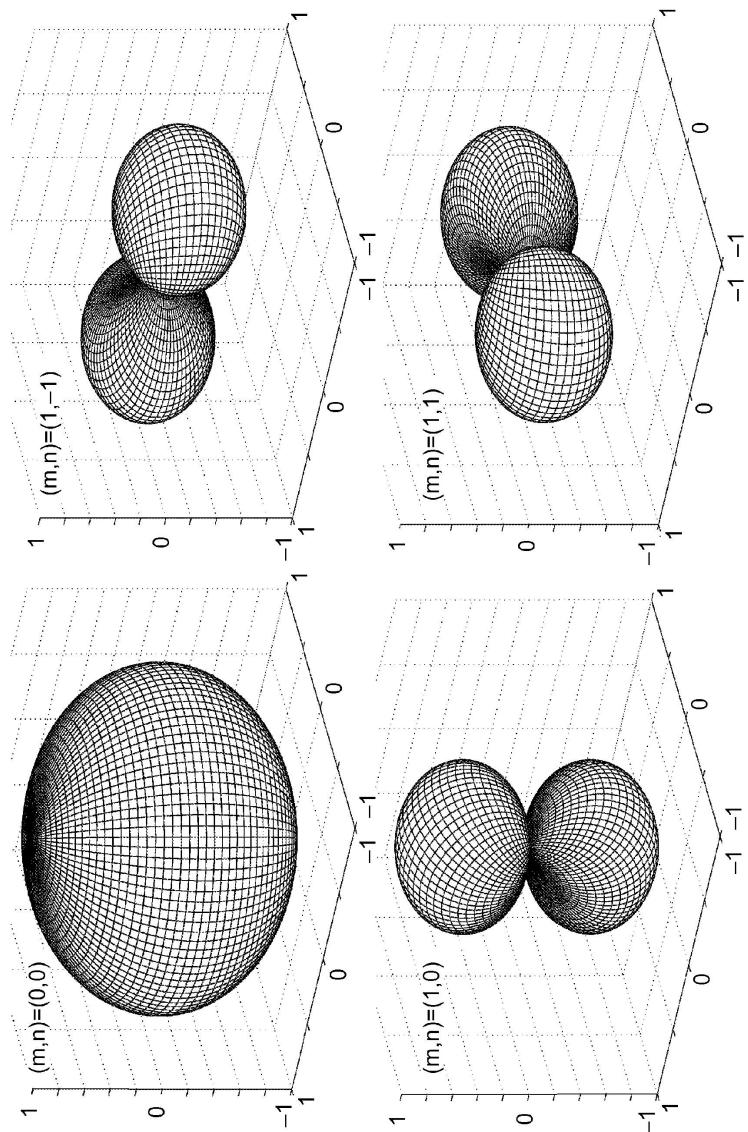
도면3c



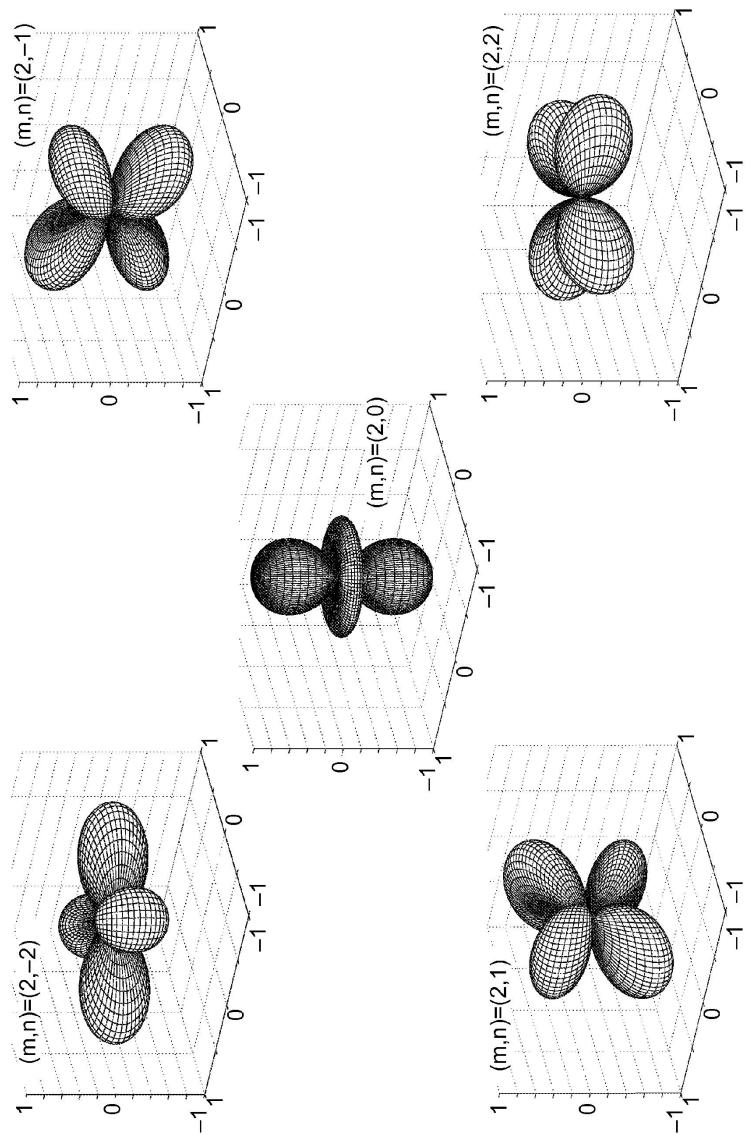
도면3d



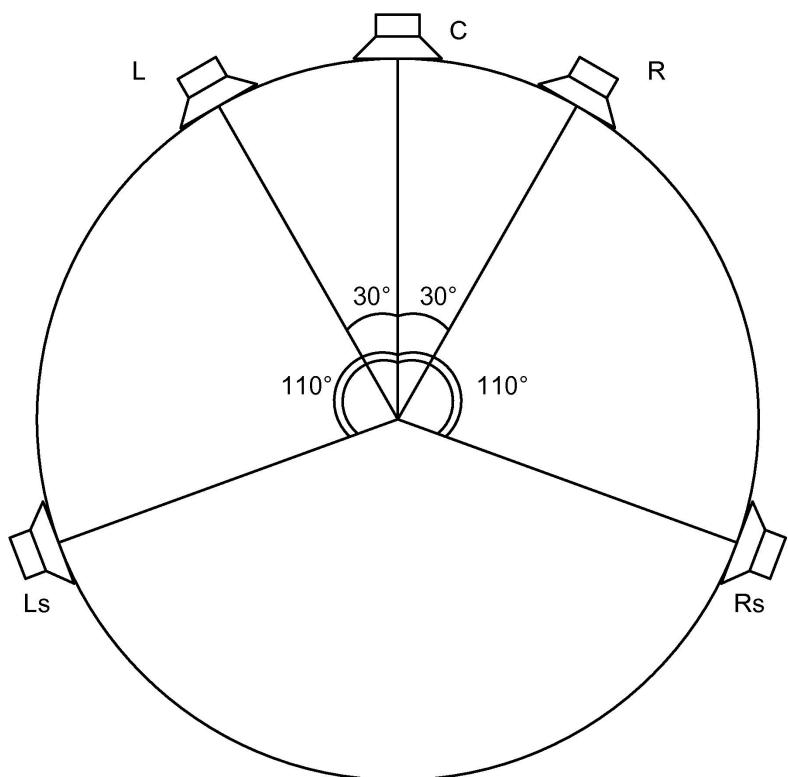
## 도면4



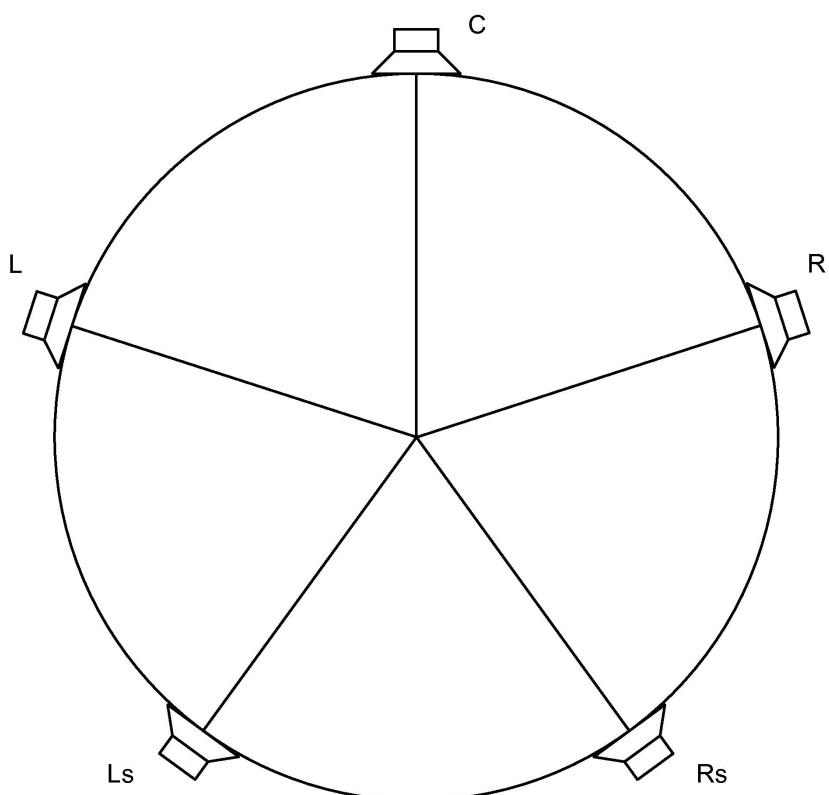
## 도면5



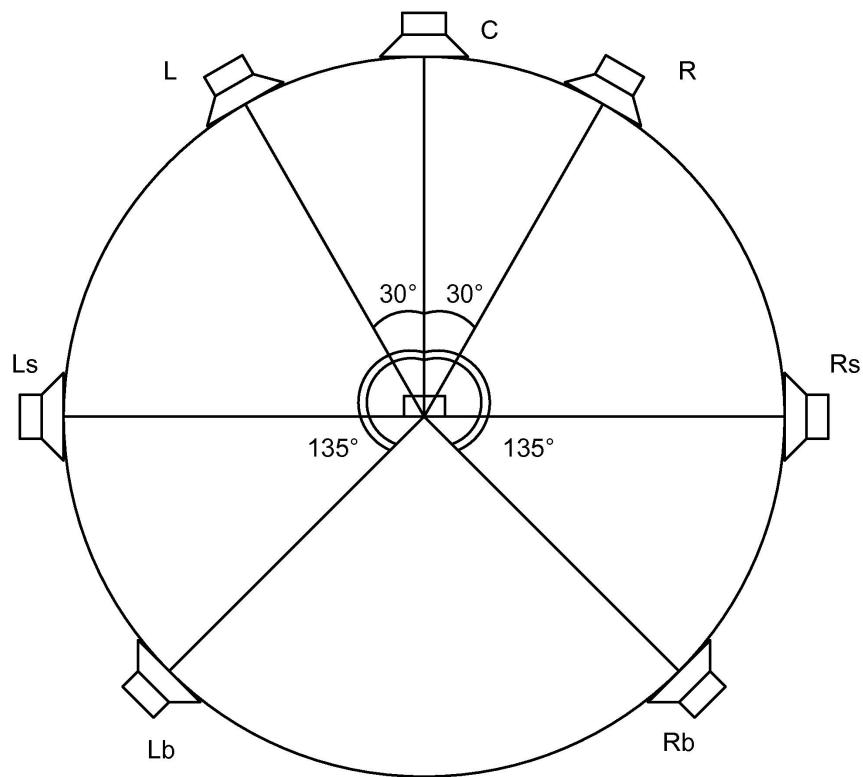
도면6



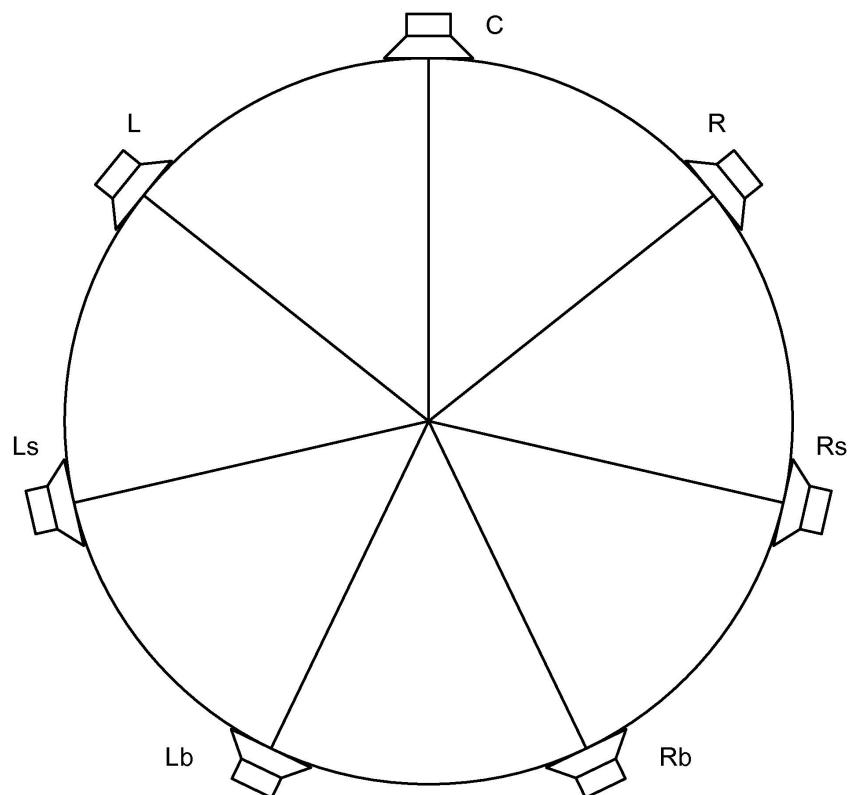
도면7



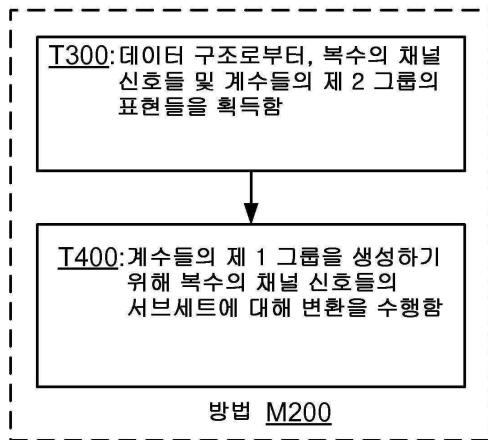
도면8



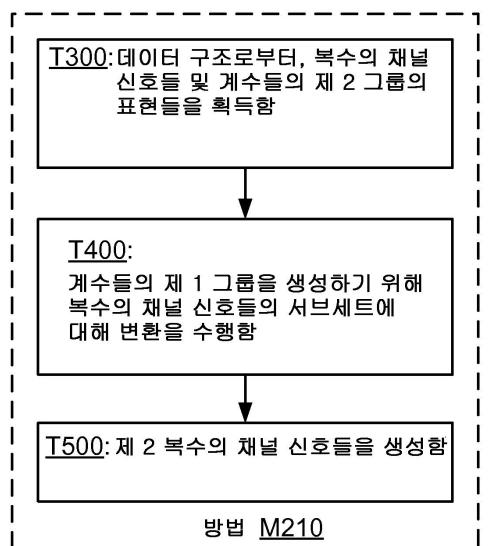
도면9



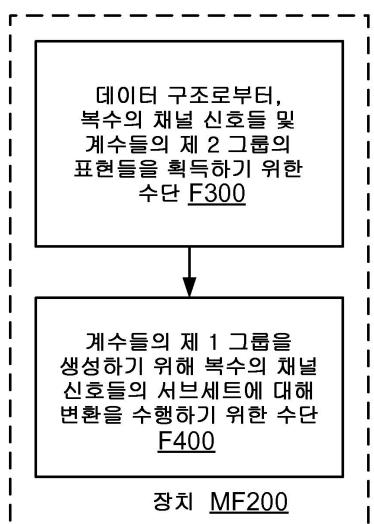
## 도면10a



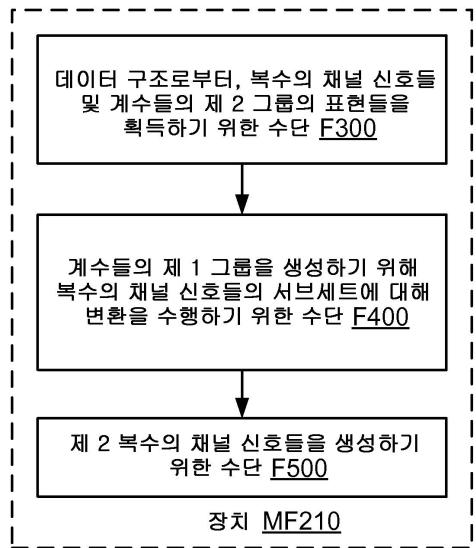
## 도면10b



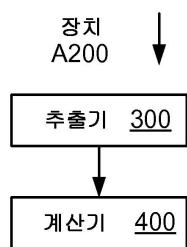
## 도면11a



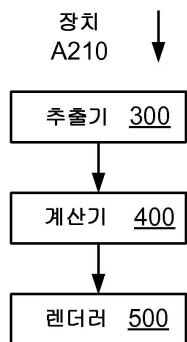
## 도면11b



## 도면11c



## 도면11d



## 도면11e

