



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년11월30일
(11) 등록번호 10-2332739
(24) 등록일자 2021년11월25일

- (51) 국제특허분류(Int. Cl.)
H04S 7/00 (2006.01)
- (52) CPC특허분류
H04S 7/302 (2013.01)
H04S 3/008 (2020.05)
- (21) 출원번호 10-2018-7033606
- (22) 출원일자(국제) 2017년05월17일
심사청구일자 2020년03월25일
- (85) 번역문제출일자 2018년11월20일
- (65) 공개번호 10-2019-0013758
- (43) 공개일자 2019년02월11일
- (86) 국제출원번호 PCT/JP2017/018500
- (87) 국제공개번호 WO 2017/208821
국제공개일자 2017년12월07일
- (30) 우선권주장
JP-P-2016-107043 2016년05월30일 일본(JP)
- (56) 선행기술조사문헌
US20100328423 A1*
US20140369506 A1*
*는 심사관에 의하여 인용된 문헌

- (73) 특허권자
소니그룹주식회사
일본국 도쿄도 미나토쿠 코난 1-7-1
- (72) 발명자
혼마 히로유키
일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니
주식회사 내
츠지 미노루
일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니
주식회사 내
- (74) 대리인
장수길, 이중희

전체 청구항 수 : 총 10 항

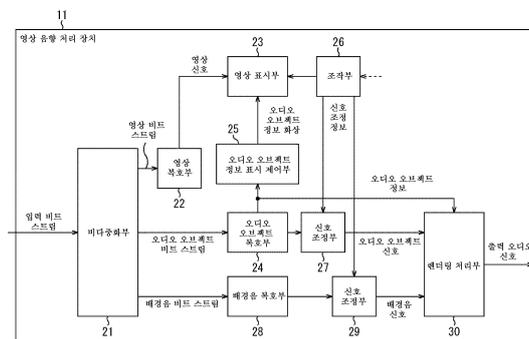
심사관 : 권영학

(54) 발명의 명칭 음향 처리 장치 및 방법, 그리고 프로그램

(57) 요약

본 기술은, 더 간단하게 음향 특성을 조정할 수 있도록 하는 음향 처리 장치 및 방법, 그리고 프로그램에 관한 것이다. 음향 처리 장치는 오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키는 표시 제어부와, 1 또는 복수의 오디오 오브젝트 중에서, 소정의 오디오 오브젝트를 선택하는 선택부를 구비한다. 본 기술은 영상 음향 처리 장치에 적용할 수 있다.

대표도



(52) CPC특허분류

H04S 7/305 (2020.05)

H04S 2400/13 (2013.01)

H04S 2420/03 (2013.01)

명세서

청구범위

청구항 1

음향 처리 장치로서,

오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키도록 구성된 표시 제어부와,

1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하도록 구성된 선택부와,

상기 선택부에 의해 선택된 상기 오디오 오브젝트의 음향에 관한 파라미터를 설정하도록 구성된 파라미터 설정부와,

상기 파라미터에 기초하여, 배경음의 배경음 신호에 대하여 상기 오디오 오브젝트의 음성의 음향 특성을 조정하는 처리를 행하도록 구성된 신호 조정부

를 구비하고,

상기 처리는, 오디오 오브젝트 신호에 대해 배경음의 음량을 낮춤으로써 상대적으로 상기 오디오 오브젝트의 음성의 음량을 증가시키도록 행해지는 것을 특징으로 하는, 음향 처리 장치.

청구항 2

제1항에 있어서, 상기 파라미터는 음량 조정 또는 음질 조정을 위한 파라미터인, 음향 처리 장치.

청구항 3

제1항에 있어서, 상기 오디오 오브젝트의 오디오 오브젝트 신호의 렌더링 처리를 행하도록 구성된 렌더링 처리부를 더 구비하는, 음향 처리 장치.

청구항 4

제3항에 있어서, 상기 파라미터는 상기 오디오 오브젝트의 위치를 지정하는 파라미터이고,

상기 렌더링 처리부는 상기 파라미터에 기초하여 상기 렌더링 처리를 행하는, 음향 처리 장치.

청구항 5

제1항에 있어서, 상기 표시 제어부는, 상기 표시부에 표시되어 있는, 상기 오디오 오브젝트의 음성이 부수되는 영상 상에 있어서의 상기 오브젝트 위치 정보에 의해 정해지는 위치에 상기 오디오 오브젝트 정보 화상을 중첩 표시시키는, 음향 처리 장치.

청구항 6

제5항에 있어서, 상기 표시 제어부는, 상기 오브젝트 위치 정보에 의해 정해지는 위치가 상기 표시부의 표시 화면 밖에 있는 경우, 상기 표시 화면의 단 부분에 상기 오디오 오브젝트 정보 화상을 표시시키는, 음향 처리 장치.

청구항 7

제5항에 있어서, 상기 선택부는, 유저에 의한 상기 오디오 오브젝트 정보 화상의 위치에서의 지정 조작에 따라, 상기 오디오 오브젝트를 선택하는, 음향 처리 장치.

청구항 8

제1항에 있어서, 오디오 오브젝트 비트 스트림을 복호하고, 상기 오디오 오브젝트의 오디오 오브젝트 신호와, 상기 오브젝트 위치 정보를 얻도록 구성된 오디오 오브젝트 복호부를 더 구비하는, 음향 처리 장치.

청구항 9

음향 처리 방법으로서,

오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키는 단계,

1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하는 단계,

선택된 상기 오디오 오브젝트의 음향에 관한 파라미터를 설정하는 단계, 및

상기 파라미터에 기초하여, 배경음의 배경음 신호에 대하여 상기 오디오 오브젝트의 음성의 음향 특성을 조정하는 처리를 행하기 위해 신호를 조정하는 단계

를 포함하고,

상기 처리는, 오디오 오브젝트 신호에 대해, 배경음의 음량을 낮춤으로써 상대적으로 상기 오디오 오브젝트의 음성의 음량을 증가시키도록 행해지는 것을 특징으로 하는, 음향 처리 방법.

청구항 10

오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키는 단계,

1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하는 단계,

선택된 상기 오디오 오브젝트의 음향에 관한 파라미터를 설정하는 단계, 및

상기 파라미터에 기초하여, 배경음의 배경음 신호에 대하여 상기 오디오 오브젝트의 음성의 음향 특성을 조정하는 처리를 행하기 위해 신호를 조정하는 단계 - 상기 처리는, 오디오 오브젝트 신호에 대해, 배경음의 음량을 낮춤으로써 상대적으로 상기 오디오 오브젝트의 음성의 음량을 증가시키도록 행해짐 -

를 포함하는 처리를 컴퓨터에 실행시키는 프로그램을 기록한 컴퓨터 판독가능한 기록 매체.

청구항 11

삭제

청구항 12

삭제

발명의 설명

기술 분야

[0001] 본 기술은 음향 처리 장치 및 방법, 그리고 프로그램에 관한 것으로, 특히, 더 간단하게 음향 특성을 조정할 수 있도록 한 음향 처리 장치 및 방법, 그리고 프로그램에 관한 것이다.

배경 기술

[0002] 근년, 영화나 게임 등에서 오브젝트 오디오 기술이 사용되고, 오브젝트 오디오를 취급할 수 있는 부호화 방식도 개발되고 있다. 예를 들어 오브젝트 오디오의 부호화에 관한 규격으로서, 국제 표준 규격인 MPEG(Moving Picture Experts Group)-H Part 3: 3D audio 규격 등이 알려져 있다(예를 들어, 비특허문헌 1 참조).

[0003] 이와 같은 부호화 방식에서는, 종래의 2채널 스테레오 방식이나, 5.1채널 등의 멀티 채널 스테레오 방식과 함께, 이동하는 음원 등을 독립된 오디오 오브젝트로서 취급하고, 오디오 오브젝트의 신호 데이터와 함께 오디오 오브젝트의 위치 정보를 메타데이터로서 부호화할 수 있다. 이와 같이 함으로써, 종래의 부호화 방식에서는 곤란했던, 재생 시에 있어서의 특정한 음원의 가공을 용이하게 행할 수 있다. 구체적으로는, 예를 들어 특정한 음원의 가공으로서, 음량 조정이나 이펙트의 추가 등을 오디오 오브젝트마다 행하는 것이 가능하다.

선행기술문헌

비특허문헌

[0004] (비특허문헌 0001) INTERNATIONAL STANDARD ISO/IEC 23008-3 First edition 2015-10-15 Information technology-High efficiency coding and media delivery in heterogeneous environments-Part 3: 3D audio

발명의 내용

해결하려는 과제

[0005] 그러나, 이와 같은 오브젝트 오디오 부호화 기술로 부호화된 오디오 오브젝트 비트 스트림을 포함하는 영상을 재생할 때에는, 오디오 오브젝트의 음량 조정 등을 하는 것은 곤란했다.

[0006] 구체적으로는, 복수의 오디오 오브젝트 중에서 원하는 오디오 오브젝트를 선택하여 음량 조정 등을 행하는 경우, 출력된 음성을 듣는 것만으로는, 목적으로 하는 오디오 오브젝트를 정확하게 선택하는 것은 매우 곤란하다.

[0007] 본 기술은 이와 같은 상황을 감안하여 이루어진 것이고, 더 간단하게 음향 특성을 조정할 수 있도록 하는 것이다.

과제의 해결 수단

[0008] 본 기술의 일측면의 음향 처리 장치는, 오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키는 표시 제어부와, 1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하는 선택부를 구비한다.

[0009] 음향 처리 장치에는, 상기 선택부에 의해 선택된 상기 오디오 오브젝트의 음향에 관한 파라미터를 설정하는 파라미터 설정부를 더 설치할 수 있다.

[0010] 음향 처리 장치에는, 상기 파라미터에 기초하여, 상기 오디오 오브젝트의 오디오 오브젝트 신호 및 배경음의 배경음 신호의 적어도 어느 한쪽에 대하여, 상기 오디오 오브젝트의 음성의 음향 특성을 조정하기 위한 처리를 행하는 신호 조정부를 더 설치할 수 있다.

[0011] 상기 파라미터를, 음량 조정 또는 음질 조정을 위한 파라미터로 할 수 있다.

[0012] 음향 처리 장치에는, 상기 오디오 오브젝트의 오디오 오브젝트 신호의 렌더링 처리를 행하는 렌더링 처리부를 더 설치할 수 있다.

[0013] 상기 파라미터를, 상기 오디오 오브젝트의 위치를 지정하는 파라미터라고 하고, 상기 렌더링 처리부에는 상기 파라미터에 기초하여 상기 렌더링 처리를 행하게 할 수 있다.

[0014] 상기 표시 제어부에는, 상기 표시부에 표시되어 있는, 상기 오디오 오브젝트의 음성이 부수되는 영상 상에 있어서의 상기 오브젝트 위치 정보에 의해 정해지는 위치에 상기 오디오 오브젝트 정보 화상을 중첩 표시시킬 수 있다.

[0015] 상기 표시 제어부에는, 상기 오브젝트 위치 정보에 의해 정해지는 위치가 상기 표시부의 표시 화면 밖에 있는 경우, 상기 표시 화면의 단 부분에 상기 오디오 오브젝트 정보 화상을 표시시킬 수 있다.

[0016] 상기 선택부에는, 유저에 의한 상기 오디오 오브젝트 정보 화상의 위치에서의 지정 조작에 따라, 상기 오디오 오브젝트를 선택시킬 수 있다.

[0017] 음향 처리 장치에는, 오디오 오브젝트 비트 스트림을 복호하고, 상기 오디오 오브젝트의 오디오 오브젝트 신호와, 상기 오브젝트 위치 정보를 얻는 오디오 오브젝트 복호부를 더 설치할 수 있다.

[0018] 본 기술의 일측면의 음향 처리 방법 또는 프로그램은, 오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키고, 1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하는 스텝을 포함한다.

[0019] 본 기술의 일측면에 있어서는, 오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상이 표시부에 표시되고, 1 또는 복수의 상기 오디오 오브젝트 중에서,

소정의 상기 오디오 오브젝트가 선택된다.

발명의 효과

- [0020] 본 기술의 일측면에 의하면, 더 간단하게 음향 특성을 조정할 수 있다.
- [0021] 또한, 여기에 기재된 효과는 반드시 한정되는 것은 아니고, 본 개시 중에 기재된 어느 효과여도 된다.

도면의 간단한 설명

- [0022] 도 1은 영상 음향 처리 장치의 구성예를 나타내는 도면이다.
- 도 2는 콘텐츠 재생 처리를 설명하는 흐름도이다.
- 도 3은 오디오 오브젝트 정보에 대하여 설명하는 도면이다.
- 도 4는 오브젝트 위치 정보의 좌표계에 대하여 설명하는 도면이다.
- 도 5는 영상 표시부의 화각 정보와 표시 화면의 관계에 대하여 설명하는 도면이다.
- 도 6은 오디오 오브젝트의 음향 특성의 조정에 대하여 설명하는 도면이다.
- 도 7은 음향 처리 장치의 구성예를 나타내는 도면이다.
- 도 8은 오디오 오브젝트 정보 화상의 표시에 대하여 설명하는 도면이다.
- 도 9는 콘텐츠 재생 처리를 설명하는 흐름도이다.
- 도 10은 컴퓨터의 구성예를 나타내는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0023] 이하, 도면을 참조하여, 본 기술을 적용한 실시 형태에 대하여 설명한다.
- [0024] <제1 실시 형태>
- [0025] <영상 음향 처리 장치의 구성예>
- [0026] 본 기술은 오디오 오브젝트 비트 스트림 중 오브젝트 위치 정보에 기초하여, 표시 장치의 표시 화면 상의 대응하는 위치에 오디오 오브젝트가 존재하고 있는 것을 직사각형 프레임 등의 화상을 중첩 표시함으로써, 오디오 오브젝트의 위치 정보를 가시화하도록 한 것이다. 또한, 본 기술에서는, 오디오 오브젝트가 표시 화면의 표시 범위 외에 있는 경우에는, 표시 화면 외측 프레임의 대응하는 방향으로 범위 외인 것을 나타내는 정보와 함께 화상을 중첩 표시함으로써, 오디오 오브젝트의 위치 정보를 가시화하도록 했다. 이에 의해, 기기 사용자는 표시된 정보에 기초하여 오디오 오브젝트의 선택을 행하고, 음량 조정 등의 조작을 용이하게 행할 수 있게 된다.
- [0027] 도 1은 본 기술을 적용한 영상 음향 처리 장치의 일 실시 형태의 구성예를 나타내는 도면이다.
- [0028] 도 1에 나타내는 영상 음향 처리 장치(11)는 비다중화부(21), 영상 복호부(22), 영상 표시부(23), 오디오 오브젝트 복호부(24), 오디오 오브젝트 정보 표시 제어부(25), 조작부(26), 신호 조정부(27), 배경음 복호부(28), 신호 조정부(29) 및 렌더링 처리부(30)를 갖고 있다.
- [0029] 이 영상 음향 처리 장치(11)에는 영상과 음성을 포함하는 콘텐츠를 재생하기 위한 입력 비트 스트림이 공급된다. 보다 상세하게는, 입력 비트 스트림으로부터 얻어지는 콘텐츠는, 영상과 그 영상에 부수되는 오디오 오브젝트의 음성 및 배경음을 포함하는 것이 된다.
- [0030] 비다중화부(21)는 외부로부터 공급된 입력 비트 스트림을, 영상 비트 스트림과, 오디오 오브젝트 비트 스트림과, 배경음 비트 스트림으로 비다중화한다.
- [0031] 여기서, 영상 비트 스트림은 콘텐츠의 영상(화상)을 재생하기 위한 영상 신호가 포함되어 있는 비트 스트림이고, 비다중화부(21)는 비다중화에 의해 얻어진 영상 비트 스트림을 영상 복호부(22)에 공급한다.
- [0032] 또한, 오디오 오브젝트 비트 스트림은 콘텐츠의 영상에 부수되는 음성을 재생하기 위한 오디오 신호 중, 오디오 오브젝트의 음성을 재생하기 위한 오디오 오브젝트 신호와, 그 오디오 오브젝트의 메타데이터인 오디오 오브젝트 정보가 포함되어 있는 비트 스트림이다.

- [0033] 비다중화부(21)는 비다중화에 의해 얻어진 오디오 오브젝트 비트 스트림을, 오디오 오브젝트 복호부(24)에 공급한다.
- [0034] 또한, 배경음 비트 스트림은 콘텐츠의 영상에 부수되는 음성을 재생하기 위한 오디오 신호 중, 오디오 오브젝트의 음성 이외의 음성, 즉 배경음을 재생하기 위한 배경음 신호가 포함되어 있는 비트 스트림이다. 비다중화부(21)는 비다중화에 의해 얻어진 배경음 비트 스트림을 배경음 복호부(28)에 공급한다.
- [0035] 영상 복호부(22)는 비다중화부(21)로부터 공급된 영상 비트 스트림을 복호하고, 그 결과 얻어진 영상 신호를 영상 표시부(23)에 공급한다.
- [0036] 영상 표시부(23)는, 예를 들어 액정 표시 패널 등의 표시 디바이스를 포함하고, 영상 복호부(22)로부터 공급된 영상 신호에 기초하여, 콘텐츠의 영상(화상)을 표시한다.
- [0037] 오디오 오브젝트 복호부(24)는 비다중화부(21)로부터 공급된 오디오 오브젝트 비트 스트림을 복호하고, 오디오 오브젝트 정보 및 오디오 오브젝트 신호를 얻는다.
- [0038] 오디오 오브젝트 복호부(24)는 복호에 의해 얻어진 오디오 오브젝트 정보를 오디오 오브젝트 정보 표시 제어부(25) 및 렌더링 처리부(30)에 공급함과 함께, 복호에 의해 얻어진 오디오 오브젝트 신호를 신호 조정부(27)에 공급한다.
- [0039] 오디오 오브젝트 정보 표시 제어부(25)는 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 정보에 기초하여, 오디오 오브젝트의 위치를 나타내는 화상 정보인 오디오 오브젝트 정보 화상을 생성하고, 영상 표시부(23)에 공급한다. 영상 표시부(23)는 영상 복호부(22)로부터 공급된 영상 신호에 기초하여 표시하고 있는 콘텐츠의 영상에, 오디오 오브젝트 정보 표시 제어부(25)로부터 공급된 오디오 오브젝트 정보 화상을 중첩 표시함으로써, 기기 사용자에게 대하여 오디오 오브젝트의 위치를 시각적으로 제시한다.
- [0040] 조작부(26)는, 예를 들어 리모트 컨트롤러로부터의 신호를 수신하는 수신부나, 영상 표시부(23)에 중첩하여 설치된 터치 패널, 버튼, 마우스, 키보드 등을 갖고, 유저의 조작을 접수하고, 유저의 조작에 따른 신호를 출력한다.
- [0041] 구체적으로는, 기기 사용자는 영상 표시부(23)에 표시된 오디오 오브젝트 정보 화상을 보면서 조작부(26)를 조작하고, 오디오 오브젝트를 선택하거나, 그 선택한 오디오 오브젝트의 음성의 음량 조정 등의 음향 특성의 조정 조작을 하거나 한다. 조작부(26)는 유저에 의한 음향 특성의 조정 조작을 받으면, 그 조작에 따라 음향 특성을 조정하기 위한 신호 조정 정보를 생성하고, 신호 조정부(27) 또는 신호 조정부(29)에 공급한다.
- [0042] 또한, 이하에는, 조작부(26)는 영상 표시부(23)와 일체적으로 설치된, 즉 영상 표시부(23)의 표시 화면에 중첩하여 설치된 터치 패널을 갖는 것으로 하여 설명을 계속한다.
- [0043] 신호 조정부(27)는 조작부(26)로부터 공급된 신호 조정 정보에 기초하여, 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 신호의 진폭 등을 조정함으로써 음량 조정이나 음질 조정 등의 음향 특성의 조정을 행하고, 그 결과 얻어진 오디오 오브젝트 신호를 렌더링 처리부(30)에 공급한다.
- [0044] 예를 들어 음향 특성으로서의 음량을 조정하는 경우에는, 오디오 오브젝트 신호의 진폭 조정이 행해진다. 또한, 음향 특성으로서의 음질을 조정하는 경우에는, 예를 들어 오디오 오브젝트 신호의 주파수 대역(밴드)마다 필터 계수를 사용한 필터 처리에 의해 게인 조정을 행함으로써, 오디오 오브젝트 신호에 기초하는 음성에 이펙트가 추가된다.
- [0045] 배경음 복호부(28)는 비다중화부(21)로부터 공급된 배경음 비트 스트림을 복호하고, 그 결과 얻어진 배경음 신호를 신호 조정부(29)에 공급한다.
- [0046] 신호 조정부(29)는 조작부(26)로부터 공급된 신호 조정 정보에 기초하여, 배경음 복호부(28)로부터 공급된 배경음 신호의 진폭 등을 조정함으로써 음량 조정이나 음질 조정 등의 음향 특성의 조정을 행하고, 그 결과 얻어진 배경음 신호를 렌더링 처리부(30)에 공급한다.
- [0047] 영상 음향 처리 장치(11)에서는, 신호 조정부(27) 및 신호 조정부(29)를 포함하는 신호 조정부에 의해, 오디오 오브젝트 신호 및 배경음 신호의 적어도 어느 한쪽에 대하여, 음량이나 음질 등의 음향 특성을 조정하기 위한 처리가 행해진다. 이에 의해, 오디오 오브젝트의 음성의 음향 특성이 조정된다.
- [0048] 오디오 오브젝트의 음성의 음향 특성을 조정할 때에는, 예를 들어 배경음의 음량을 낮춤으로써, 상대적으로 오

디오 오브젝트의 음성의 음량을 증가시키는 등, 적절히, 오디오 오브젝트 신호에 대한 처리와, 배경음 신호에 대한 처리가 조합되고, 오디오 오브젝트의 음성의 음향 특성 조정이 실현된다.

- [0049] 렌더링 처리부(30)는 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 정보에 기초하여, 신호 조정부(27)로부터 공급된 오디오 오브젝트 신호에 대한 렌더링 처리를 행한다.
- [0050] 또한, 렌더링 처리부(30)는 렌더링 처리에 의해 얻어진 오디오 오브젝트 신호와, 신호 조정부(29)로부터 공급된 배경음 신호를 합성하는 믹싱 처리를 행하고, 그 결과 얻어진 출력 오디오 신호를 출력한다. 이 출력 오디오 신호의 공급을 받은 스피커는 출력 오디오 신호에 기초하여, 콘텐츠의 음성을 재생한다. 이때, 콘텐츠의 음성으로서 오디오 오브젝트의 음성과 배경음이 재생된다.
- [0051] <콘텐츠 재생 처리의 설명>
- [0052] 이어서, 영상 음향 처리 장치(11)의 동작에 대하여 설명한다. 즉, 이하, 도 2의 흐름도를 참조하여, 영상 음향 처리 장치(11)에 의해 행해지는 콘텐츠 재생 처리에 대하여 설명한다.
- [0053] 스텝 S11에 있어서, 비다중화부(21)는 외부로부터 공급된 입력 비트 스트림을 비다중화하여, 영상 비트 스트림, 오디오 오브젝트 비트 스트림 및 배경음 비트 스트림을 얻는다.
- [0054] 비다중화부(21)는 비다중화에 의해 얻어진 영상 비트 스트림, 오디오 오브젝트 비트 스트림 및 배경음 비트 스트림을, 각각 영상 복호부(22), 오디오 오브젝트 복호부(24) 및 배경음 복호부(28)에 공급한다.
- [0055] 스텝 S12에 있어서, 영상 복호부(22)는 비다중화부(21)로부터 공급된 영상 비트 스트림을 복호하고, 그 결과 얻어진 영상 신호를 영상 표시부(23)에 공급한다. 영상 표시부(23)는 영상 복호부(22)로부터 공급된 영상 신호에 기초하여, 콘텐츠의 화상(영상)을 표시시킨다. 즉, 콘텐츠의 영상을 재생한다.
- [0056] 스텝 S13에 있어서, 배경음 복호부(28)는 비다중화부(21)로부터 공급된 배경음 비트 스트림을 복호하고, 그 결과 얻어진 배경음 신호를 신호 조정부(29)에 공급한다.
- [0057] 스텝 S14에 있어서, 오디오 오브젝트 복호부(24)는 비다중화부(21)로부터 공급된 오디오 오브젝트 비트 스트림을 복호하고, 오디오 오브젝트 정보 및 오디오 오브젝트 신호를 얻는다.
- [0058] 여기서, 오디오 오브젝트 신호는 오디오 오브젝트의 음성의 파형 신호이고, 오디오 오브젝트 비트 스트림의 복호에 의해, 1 또는 복수의 오디오 오브젝트마다 오디오 오브젝트 신호가 얻어진다. 예를 들어 오디오 오브젝트 신호는 PCM(Pulse Code Modulation) 신호 등이 된다.
- [0059] 한편, 오디오 오브젝트 정보란, 음원인 오디오 오브젝트가 공간 상의 어느 위치에 존재하는지를 나타내는 정보 등을 포함하는 메타데이터이고, 예를 들어 도 3에 나타내는 포맷으로 부호화되어 있다.
- [0060] 도 3에 나타내는 오디오 오브젝트 정보에 있어서, 「num_objects」는 오디오 오브젝트 비트 스트림에 포함되어 있는 오디오 오브젝트의 수를 나타내고 있다.
- [0061] 또한, 「tcimbsf」는 「Two's complement integer, most significant(sign) bit first」의 줄임이고, 부호 비트가 선두의 2의 보수를 나타내고 있다. 「uimbsf」는 「Unsigned integer, most significant bit first」의 줄임이고, 최상위 비트가 선두의 부호없이 정수를 나타내고 있다.
- [0062] 또한, 「gain_factor[i]」는 오디오 오브젝트 비트 스트림에 포함되어 있는 i번째의 오디오 오브젝트의 게인을 나타내고 있다.
- [0063] 「position_azimuth[i]」, 「position_elevation[i]」 및 「position_radius[i]」는 각각 오디오 오브젝트 비트 스트림에 포함되어 있는 i번째의 오디오 오브젝트의 위치 정보를 나타내고 있다.
- [0064] 구체적으로는, 「position_azimuth[i]」는 구면 좌표계에 있어서의 오디오 오브젝트의 위치의 방위각을 나타내고 있고, 「position_elevation[i]」는 구면 좌표계에 있어서의 오디오 오브젝트의 위치의 양각을 나타내고 있다. 또한, 「position_radius[i]」는 구면 좌표계에 있어서의 오디오 오브젝트의 위치까지의 거리, 즉 반경을 나타내고 있다.
- [0065] 이하에는 오디오 오브젝트 정보에 포함되는 「position_azimuth[i]」, 「position_elevation[i]」 및 「position_radius[i]」를 포함하는, 오디오 오브젝트의 위치를 나타내는 정보를, 오브젝트 위치 정보라고도 칭하는 것으로 한다.

- [0066] 또한, 이하, 오디오 오브젝트 정보에 포함되는 「gain_factor[i]」인 오디오 오브젝트의 계인을 나타내는 정보를 계인 정보라고도 칭하는 것으로 한다.
- [0067] 이와 같은 각 오디오 오브젝트의 오브젝트 위치 정보나 계인 정보가 포함되는 오디오 오브젝트 정보는, 오디오 오브젝트의 메타데이터이다.
- [0068] 여기서 구면 좌표계와 3차원 직교 좌표계의 관계는 도 4에 나타내는 관계가 되어 있다.
- [0069] 도 4에서는 원점 0를 지나고, 서로 수직인 X축, Y축 및 Z축이 3차원 직교 좌표계의 축이 되어 있다. 예를 들어 3차원 직교 좌표계에서는, 공간 상의 오디오 오브젝트 OB11의 위치는, X축 방향의 위치를 나타내는 X좌표인 X1, Y축 방향의 위치를 나타내는 Y좌표인 Y1 및 Z축 방향의 위치를 나타내는 Z좌표인 Z1이 사용되고 (X1, Y1, Z1)이라고 표현된다.
- [0070] 이에 비해 구면 좌표계에서는, 방위각 position_azimuth, 양각 position_elevation 및 반경 position_radius가 사용되고 공간 상의 오디오 오브젝트 OB11의 위치가 표현된다.
- [0071] 지금, 원점 0와, 공간 상의 오디오 오브젝트 OB11의 위치를 연결하는 직선을 직선 r이라고 하고, 이 직선 r을 XY 평면 상에 투영하여 얻어진 직선을 직선 L이라고 한다.
- [0072] 이때, X축과 직선 L이 이루는 각 θ 가 오디오 오브젝트 OB11의 위치를 나타내는 방위각 position_azimuth가 된다. 또한, 직선 r과 XY 평면이 이루는 각 ϕ 가 오디오 오브젝트 OB11의 위치를 나타내는 양각 position_elevation이 되고, 직선 r의 길이가 오디오 오브젝트 OB11의 위치를 나타내는 반경 position_radius가 된다.
- [0073] 예를 들어, 원점 0의 위치는 콘텐츠의 영상(화상)을 보는 유저의 위치가 되고, X방향(X축 방향)의 정의 방향, 즉 도 4 중, 전방 방향이 유저로부터 본 정면 방향이 되고, Y방향(Y축 방향)의 정의 방향, 즉 도 4 중, 우측 방향이 유저로부터 본 좌측 방향이 된다.
- [0074] 이와 같이 오디오 오브젝트 정보에 있어서는, 각 오디오 오브젝트의 위치가 구면 좌표에 의해 표현되어 있다.
- [0075] 이와 같은 오디오 오브젝트 정보에 의해 나타나는 오디오 오브젝트의 위치와 계인은 소정의 시간 구간마다 변화되는 물리량이다. 콘텐츠의 재생 시에는, 오디오 오브젝트 정보의 변화에 따라, 오디오 오브젝트의 음상 정위 위치를 이동시킬 수 있다.
- [0076] 도 2의 흐름도의 설명으로 돌아가, 오디오 오브젝트 복호부(24)는 오디오 오브젝트 비트 스트림을 복호함으로써, 오디오 오브젝트 정보 및 오디오 오브젝트 신호를 얻는다.
- [0077] 오디오 오브젝트 복호부(24)는 복호에 의해 얻어진 오디오 오브젝트 정보를 오디오 오브젝트 정보 표시 제어부(25) 및 렌더링 처리부(30)에 공급함과 함께, 복호에 의해 얻어진 오디오 오브젝트 신호를 신호 조정부(27)에 공급한다.
- [0078] 스텝 S15에 있어서, 오디오 오브젝트 정보 표시 제어부(25)는 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 정보에 기초하여, 오디오 오브젝트의 표시 화면 상에 있어서의 위치를 산출한다.
- [0079] 예를 들어 「INTERNATIONAL STANDARD ISO/IEC 23008-3 First edition 2015-10-15 Information technology-High efficiency coding and media delivery in heterogeneous environments-Part 3: 3D audio」(이하, 참조 문헌 1이라고도 칭함)의 규격에 있어서는, 비트 스트림 중에 영상 제작자측이 상정한 재생 장치의 화면에 관한 정보를 수평 방향의 화각 정보 및 수직 방향의 화각 정보로서 기술할 수 있지만, 이들의 화각 정보가 기술되어 있지 않은 경우에는, 화각 정보로서 디폴트값을 사용하게 되어 있다.
- [0080] 여기서, 예로서 화각 정보의 디폴트값을 사용하여 오디오 오브젝트의 표시 화면 상에 있어서의 위치를 산출하는 경우에 대하여 설명한다.
- [0081] 예를 들어, 영상 표시부(23)에 대하여, 공간 상의 원점 0에서 본 영상 표시부(23)의 수평 방향, 즉 Y방향의 화각을 나타내는 화각 정보를 screen_azimuth라고 하고, 공간 상의 원점 0에서 본 영상 표시부(23)의 수직 방향, 즉 Z방향의 화각을 나타내는 화각 정보를 screen_elevation이라고 한다.
- [0082] 이때, 수평 방향의 화각 정보 screen_azimuth의 디폴트값과, 수직 방향의 화각 정보 screen_elevation의 디폴트값은 다음 식(1)에 나타내는 바와 같이 된다.

$$\begin{cases} \text{screen_azimuth} = 29.0^\circ \\ \text{screen_elevation} = 17.5^\circ \end{cases} \dots (1)$$

[0083]

[0084]

또한, 이와 같은 화각 정보 screen_azimuth 및 화각 정보 screen_elevation과, 영상 표시부(23)의 표시 화면의 관계는 도 5에 나타낸 바와 같이 된다.

[0085]

이 예에서는, 영상 표시부(23)의 표시 화면의 중심 위치 0'과, 공간 상의 유저의 위치인 원점 0의 도면 중, 중 방향의 위치가 동일한 위치로 되어 있다.

[0086]

또한, 중심 위치 0'을 원점으로 하고, 도면 중, 우측 방향을 x방향의 정의 방향으로 하고, 도면 중, 상측 방향을 y방향의 정의 방향으로 하는 2차원 직교 좌표계를 xy 좌표계로 하고, 그 xy 좌표계 상의 위치를, x좌표와 y좌표를 사용하여 (x, y)로 나타내는 것으로 한다.

[0087]

또한, 영상 표시부(23)의 표시 화면의 x방향의 폭(길이)을 screen_width라고 하고, 영상 표시부(23)의 표시 화면의 y방향의 폭(길이)을 screen_height라고 한다. 또한, 이들의 screen_width나 screen_height는 기지의 값이다.

[0088]

이때, 영상 표시부(23)의 표시 화면의 도면 중, 우측의 단의 위치 PS11과 좌측의 단의 위치 PS12는 각각 xy 좌표에 의해 위치 PS11=(screen_width/2, 0) 및 위치 PS12=(-screen_width/2, 0)이 된다.

[0089]

마찬가지로, 영상 표시부(23)의 표시 화면의 도면 중, 상측의 단의 위치 PS13과 하측의 단의 위치 PS14는, 각각 xy 좌표에 의해 위치 PS13=(0, screen_height/2) 및 위치 PS14=(0, -screen_height/2)가 된다.

[0090]

또한 원점 0를 시점으로 하고, 중심 위치 0'을 종점으로 하는 벡터 VB31과, 원점 0를 시점으로 하고, 위치 PS11을 종점으로 하는 벡터 VB32가 이루는 각 AG31이 -screen_azimuth가 된다. 마찬가지로 벡터 VB31과, 원점 0를 시점으로 하고, 위치 PS12를 종점으로 하는 벡터 VB33이 이루는 각 AG32가 screen_azimuth가 된다.

[0091]

또한, 벡터 VB31과, 원점 0를 시점으로 하고, 위치 PS13을 종점으로 하는 벡터 VB34가 이루는 각 AG33이 screen_elevation이 된다. 벡터 VB31과, 원점 0를 시점으로 하고, 위치 PS14를 종점으로 하는 벡터 VB35가 이루는 각 AG34가 -screen_elevation이 된다.

[0092]

오디오 오브젝트 정보 표시 제어부(25)는 디폴트값인 화각 정보 screen_azimuth 및 화각 정보 screen_elevation과, 기지의 값인 영상 표시부(23)의 길이 screen_width 및 screen_height와, 오디오 오브젝트 정보에 포함되는 오브젝트 위치 정보로부터 다음 식 (2)를 계산하고, 영상 표시부(23)의 표시 화면 상에 있어서의 오디오 오브젝트의 위치를 산출한다.

$$\begin{cases} x = -\frac{\text{screen_width}}{2} \frac{\sin(\text{position_azimuth})}{\sin(\text{screen_azimuth})}, \\ \quad |\text{position_azimuth}| < |\text{screen_azimuth}| \\ \dots (2) \\ y = \frac{\text{screen_height}}{2} \frac{\sin(\text{position_elevation})}{\sin(\text{screen_elevation})}, \\ \quad |\text{position_elevation}| < |\text{screen_elevation}| \end{cases}$$

[0093]

[0094]

또한, 식 (2)에 있어서, position_azimuth 및 position_elevation은 오브젝트 위치 정보를 구성하는, 오디오 오브젝트의 위치를 나타내는 방위각 및 양각을 나타내고 있다.

[0095]

이와 같이 하여 식 (2)의 계산을 행하면, 영상 표시부(23)의 표시 화면 상, 즉 콘텐츠의 화상 상에 있어서의 오디오 오브젝트의 위치를 나타내는 x좌표와 y좌표가 구해지게 된다. 이하, 이와 같이 하여 구해진 영상 표시부(23)의 표시 화면 상에 있어서의 오디오 오브젝트의 위치를, 오브젝트 화면 위치라고도 칭하는 것으로 한다.

[0096]

또한, 오디오 오브젝트 정보 표시 제어부(25)는 식 (2)에 나타내는 제약 조건, 즉

$|position_azimuth| < |screen_azimuth|$ 및 $|position_elevation| < |screen_elevation|$ 을 만족시키지 않는 오디오 오브젝트에 대해서는, 예를 들어 이하와 같이 하여 오브젝트 화면 위치를 구한다.

- [0097] 즉, $|position_azimuth| < |screen_azimuth|$ 가 만족되지 않는 경우, 오브젝트 화면 위치에 의해 나타나는 오디오 오브젝트의 x좌표는, $position_azimuth$ 가 정의 값일 때에는 $x = -screen_width/2$ 가 되고, $position_azimuth$ 가 부의 값일 때에는 $x = screen_width/2$ 가 된다.
- [0098] 또한, $|position_elevation| < |screen_elevation|$ 이 만족되지 않는 경우, 오브젝트 화면 위치에 의해 나타나는 오디오 오브젝트의 y좌표는, $position_elevation$ 이 정의 값일 때에는 $y = screen_height/2$ 가 되고, $position_elevation$ 이 부의 값일 때에는 $y = -screen_height/2$ 가 된다.
- [0099] 따라서, 식 (2)에 나타내는 제약 조건을 만족시키지 않는 오디오 오브젝트에 대해서는, 그 오디오 오브젝트의 오브젝트 화면 위치는, 영상 표시부(23)의 표시 화면의 단 부분의 위치가 된다. 즉, 오브젝트 화면 위치는, 영상 표시부(23)의 화각 정보에 의해 나타나는 위치가 된다.
- [0100] 이와 같은 식 (2)의 제약 조건을 만족시키지 않는 오디오 오브젝트는, 콘텐츠의 화상 상에서 관찰되지 않는, 화상 밖에 위치하는 오브젝트, 즉 영상 표시부(23)의 표시 화면 밖에 있는 오브젝트이다.
- [0101] 따라서, 오브젝트 위치 정보로부터 구해지는 오브젝트 화면 위치가, 영상 표시부(23)의 표시 화면 밖이 되는 오디오 오브젝트에 대해서는, 그 오브젝트 화면 위치에 가장 가까운, 영상 표시부(23)의 표시 화면의 단 부분의 위치가, 최종적인 오브젝트 화면 위치가 되게 된다. 또한, 그 밖에, 오브젝트 화면 위치가 표시 화면 밖이 되는 경우, 그 오브젝트 화면 위치 및 중심 위치 '0'을 연결하는 직선과, 영상 표시부(23)의 표시 화면의 단의 교점 위치를, 최종적인 오브젝트 화면 위치로 하거나 해도 된다.
- [0102] 도 2의 흐름도의 설명으로 돌아가, 스텝 S16에 있어서, 오디오 오브젝트 정보 표시 제어부(25)는 오브젝트 화면 위치에 기초하여 영상 표시부(23)를 제어하고, 콘텐츠의 화상(영상) 상에, 오디오 오브젝트가 존재하는 것을 나타내는 오디오 오브젝트 정보 화상을 중첩 표시시킨다.
- [0103] 이 오디오 오브젝트 정보 화상의 표시 위치는, 오브젝트 화면 위치, 즉 오브젝트 위치 정보에 의해 정해지는 영상 표시부(23)의 표시 화면 상의 위치가 된다. 바꾸어 말하면, 오디오 오브젝트 정보 화상은, 오브젝트 위치 정보에 의해 정해지는, 콘텐츠의 화상(영상) 상의 위치에 표시된다.
- [0104] 예를 들어, 오디오 오브젝트 정보 표시 제어부(25)는 스텝 S15의 처리에서 얻어진 오브젝트 화면 위치에 기초하여, 그 오브젝트 화면 위치를 중심으로 하는 소정의 크기의 직사각형 프레임의 화상인 직사각형 프레임 화상을 표시시키기 위한 화상 정보를, 오디오 오브젝트 정보 화상의 화상 정보로서 생성한다. 예를 들어, 직사각형 프레임 화상의 크기는, 미리 정해진 크기로 되어도 되고, 오브젝트 위치 정보로서의 반경 $position_radius$ 에 의해 정해지는 크기로 되어도 된다.
- [0105] 단, 상술한 식 (2)의 제약 조건을 만족시키지 않은 오디오 오브젝트에 대한 직사각형 프레임 화상은 제약 조건을 만족시키는 오디오 오브젝트의 직사각형 프레임 화상과는 상이한 직사각형 프레임 화상이 된다. 여기서, 상이한 직사각형 프레임 화상이란, 예를 들어 직사각형 프레임의 형상이나 크기가 상이한 것이 되지만, 색 등의 표시 형식이 상이한 것이 되어도 된다.
- [0106] 오디오 오브젝트 정보 표시 제어부(25)는, 이와 같이 하여 생성한 오디오 오브젝트 정보 화상을 영상 표시부(23)에 공급하고, 콘텐츠의 화상 상에 오디오 오브젝트 정보 화상을 중첩하여 표시시킨다.
- [0107] 이때, 식 (2)의 제약 조건을 만족시키지 않는, 즉 오브젝트 위치 정보로부터 구해지는 오브젝트 화면 위치가, 영상 표시부(23)의 표시 화면 밖이 되는 오디오 오브젝트의 오디오 오브젝트 정보 화상은, 그 오브젝트 화면 위치에 가장 가까운, 영상 표시부(23)의 표시 화면의 단 부분의 위치에 표시되게 된다. 즉, 오디오 오브젝트 정보 화상이 콘텐츠의 화상 상의 단 부분에 표시되게 된다.
- [0108] 스텝 S16의 처리가 행해지면, 예를 들어 영상 표시부(23)에는, 도 6에 나타내는 화상이 표시된다.
- [0109] 도 6에 나타내는 예에서는, 영상 표시부(23)에 표시되어 있는 콘텐츠의 화상 상에는 3명의 인물 HM11 내지 인물 HM13이 오디오 오브젝트로서 표시되어 있다.
- [0110] 또한, 이들의 인물 HM11 내지 인물 HM13의 얼굴 영역의 부분에는 각각 오디오 오브젝트 정보 화상으로서의 직사각형 프레임 화상 FR11 내지 직사각형 프레임 화상 FR13이 중첩 표시되어 있다. 따라서, 유저는 이들의 직사각형 프레임 화상 FR11 내지 직사각형 프레임 화상 FR13을 봄으로써, 간단하게 오디오 오브젝트를 인식할 수

있다.

- [0111] 또한, 도 6에서는 식 (2)의 제약 조건을 만족시키지 않는 오디오 오브젝트, 즉 표시 화면 밖에 있는 오디오 오브젝트가 존재하는 것을 나타내는 직사각형 프레임 화상 FR14가, 영상 표시부(23)의 표시 화면단에 표시되어 있다.
- [0112] 특히, 이 예에서는 직사각형 프레임 화상 FR14에 대응하는 오디오 오브젝트가 표시 화면 밖에 있는 것을 나타내기 때문에, 직사각형 프레임 화상 FR14가 점선 표시되어 있다. 즉, 직사각형 프레임 화상 FR14는 다른 직사각형 프레임 화상 FR11 내지 직사각형 프레임 화상 FR13과는 상이한 표시 형식으로 표시되어 있다.
- [0113] 이 예에서는, 직사각형 프레임 화상 FR11과 직사각형 프레임 화상 FR13도 점선 표시되어 있지만, 직사각형 프레임 화상 FR14의 점선 표시는, 이들의 직사각형 프레임 화상 FR11 및 직사각형 프레임 화상 FR13의 점선 표시와는 상이한 표시 형식으로 되어 있고, 구별할 수 있도록 되어 있다.
- [0114] 직사각형 프레임 화상 FR14에 대응하는 오디오 오브젝트는, 표시 화면 밖에 위치하고 있으므로, 유저는 그 오디오 오브젝트를 콘텐츠의 화상 상에서 확인할 수는 없다. 그러나 유저는 직사각형 프레임 화상 FR14를 봄으로써, 표시 화면 밖에 오디오 오브젝트가 존재하는 것을 알 수 있다. 예를 들어 유저는, 직사각형 프레임 화상 FR14를 보면, 자신으로부터 볼 때 표시 화면 밖의 좌측에 보이지 않는 오디오 오브젝트가 있는 것을 인식할 수 있다.
- [0115] 이와 같은 상태에서, 예를 들어 기기 사용자인 유저가 인물 HM12의 얼굴 영역의 부분에 표시되어 있는 직사각형 프레임 화상 FR12를 손가락으로 누르거나 하여 선택하는 것으로 한다. 이 선택 조작에 의해, 직사각형 프레임 화상 FR12가 표시되어 있는 인물 HM12가 선택되게 된다.
- [0116] 그러면, 선택된 인물 HM12에 대하여 표시되어 있는 직사각형 프레임 화상 FR12가 하이라이트 표시된다. 이 예에서는, 직사각형 프레임 화상 FR12가 실선으로 그려져 있고, 직사각형 프레임 화상 FR12가 하이라이트 표시되어 있는 것을 나타내고 있다. 이에 의해, 유저는 어느 오디오 오브젝트를 선택했는지를 시각적으로 파악할 수 있다.
- [0117] 이에 비해, 선택되어 있지 않은 인물의 직사각형 프레임 화상 FR11, 직사각형 프레임 화상 FR13 및 직사각형 프레임 화상 FR14는 점선으로 그려져 있고, 하이라이트 표시되어 있지 않은, 즉 통상의 표시로 되어 있는 것을 나타내고 있다. 따라서 직사각형 프레임 화상 FR12가 선택되면, 직사각형 프레임 화상 FR12는 점선으로 그려진 통상 표시의 상태로부터, 실선으로 그려진 하이라이트 표시된 상태로 표시 상태가 변화된다.
- [0118] 또한, 이 예에서는, 직사각형 프레임 화상 FR12가 선택되어 하이라이트 표시된 것에 수반하여, 직사각형 프레임 화상 FR12 근방에는, 선택된 인물 HM12의 음성의 음향 특성의 조정을 행하기 위한 조정 지시 화상 CT11이 표시되어 있다.
- [0119] 여기서, 조정 지시 화상 CT11의 예로서, 인물 HM12의 음성의 음량 조정을 행하기 위한 화상이 표시되어 있다. 또한, 조정 지시 화상에 대한 조작에 의해, 음량 조정뿐만 아니라, 음질의 조정 등도 행할 수 있지만, 여기서는 설명을 간단하게 하기 위해, 음량 조정을 예로 들어 설명을 계속한다.
- [0120] 기기 사용자인 유저는 이 조정 지시 화상 CT11에 나타나는 화살표 부분에 대한 조작을 행함으로써, 더 간단하고 또한 직감적으로 오디오 오브젝트의 음성의 음량을 조정할 수 있다. 구체적으로는, 유저는 조정 지시 화상 CT11의 화살표 부분에 있어서의 도면 중, 상측의 부분에 접촉함으로써 음량을 높이도록 할 수 있고, 반대로 화살표 부분에 있어서의 도면 중, 하측의 부분에 접촉함으로써 음량을 낮추도록 할 수 있다. 또한, 예를 들어 화살표 부분에 접촉한 횟수나 접촉한 시간에 따라 음량을 높이거나 낮추거나 하는 양이 결정된다.
- [0121] 또한, 유저는, 조정 지시 화상 CT11이 표시되어 있는 상태에서, 다시, 직사각형 프레임 화상 FR12를 손가락으로 누르거나 함으로써 인물 HM12의 선택을 해제하고, 조정 지시 화상 CT11이 표시되어 있지 않은 상태로 표시를 복귀시킬 수 있다.
- [0122] 도 2의 흐름도의 설명으로 돌아가, 스텝 S17에 있어서, 조작부(26)는 기기 사용자인 유저에 의한 조작에 따라, 음향 특성의 조정을 행하는 오디오 오브젝트를 선택한다.
- [0123] 예를 들어, 유저에 의한 오디오 오브젝트의 선택은 그 오디오 오브젝트에 대하여 표시되어 있는 직사각형 프레임 화상, 즉 오디오 오브젝트 정보 화상을 유저가 지정하는 것에 의해 행해진다. 이때, 유저는, 1 또는 복수의 오디오 오브젝트 중에서, 1개만 오디오 오브젝트를 선택하여 음향 특성을 조정할 수도 있고, 복수의 오디오 오

브젝트를 차례로 선택하여 음향 특성을 조정할 수도 있다.

- [0124] 조작부(26)는 유저에 의한 오디오 오브젝트 정보 화상의 지정 조작에 따라, 오디오 오브젝트를 선택한다.
- [0125] 구체적으로는, 예를 들어 도 6의 예에 있어서, 유저가 직사각형 프레임 화상 FR12에 손가락으로 접촉하는 조작을 행한 경우, 조작부(26)는 유저의 조작부(26)에 대한 조작에 따라 발생한 신호에 기초하여, 직사각형 프레임 화상 FR12에 대응하는 인물 HM12를, 음향 특성의 조정을 행하는 오디오 오브젝트로서 선택한다.
- [0126] 이와 같이 하여, 직사각형 프레임 화상에 대응하는 인물을 선택하는 조작부(26)는, 유저의 조작에 따라 오디오 오브젝트를 선택하는 오디오 오브젝트 선택부로서 기능한다고 할 수 있다.
- [0127] 또한, 조작부(26)는 음향 특성의 조정을 행하는 오디오 오브젝트를 선택하면, 영상 표시부(23)를 제어하고, 선택한 오디오 오브젝트에 대응하는 직사각형 프레임 화상(오디오 오브젝트 정보 화상)을 하이라이트 표시시킴과 함께, 그 직사각형 프레임 화상 근방에 조정 지시 화상을 표시시킨다. 이에 의해, 도 6의 예에서는, 직사각형 프레임 화상 FR12가 하이라이트 표시됨과 함께, 조정 지시 화상 CT11이 표시된다.
- [0128] 그러면, 기기 사용자인 유저는 조정 지시 화상에 대한 조작을 행하여, 오디오 오브젝트의 음성의 음향 특성의 조정을 지시한다. 또한, 오디오 오브젝트의 음성뿐만 아니라, 배경음의 음향 특성도 조정할 수 있도록 해도 된다.
- [0129] 스텝 S18에 있어서, 조작부(26)는 유저의 조정 지시 화상에 대한 조작에 따라 발생한 신호에 기초하여, 선택되어 있는 오디오 오브젝트의 음성의 음향 특성을 조정하기 위한 신호 조정 정보를 생성한다.
- [0130] 예를 들어 조작부(26)는, 오디오 오브젝트의 음성의 음량 조정이 지시된 경우, 그 지시된 분만큼 음량을 낮추거나, 또는 높이는 것을 지시하는 신호 조정 정보를 생성한다. 이 경우, 예를 들어 신호 조정 정보에는, 음량을 증가 또는 감소시키는 양, 즉 음량의 조정량을 나타내는 정보가 파라미터로서 포함되게 된다.
- [0131] 또한, 예를 들어 조작부(26)는 오디오 오브젝트의 음성의 음질 조정이 지시된 경우, 그 지시에 따른 이펙트를 부가하기 위한 필터 처리에 사용하는 필터 계수를 선택하고, 선택한 필터 계수를 나타내는 정보를 파라미터로서 포함하는 신호 조정 정보를 생성한다.
- [0132] 이와 같이 하여 생성되는 신호 조정 정보는, 음량의 조정량을 나타내는 정보나, 필터 계수를 나타내는 정보 등의, 음향에 관한 파라미터, 즉 음향 특성을 조정할 때의 조정 정도를 나타내는 파라미터를 포함하는 것이다. 따라서, 조작부(26)는 유저의 조작에 따라, 음향 특성을 조정할 때의 파라미터를 설정하고, 설정된 파라미터를 포함하는 신호 조정 정보를 생성하는 파라미터 설정부로서도 기능한다고 할 수 있다.
- [0133] 조작부(26)는 이상과 같이 하여 생성한 신호 조정 정보를, 신호 조정부(27), 신호 조정부(29), 또는 신호 조정부(27) 및 신호 조정부(29)에 공급한다.
- [0134] 스텝 S19에 있어서, 신호 조정부(27) 또는 신호 조정부(29)는 조작부(26)로부터 공급된 신호 조정 정보에 기초하여 음향 특성을 조정한다.
- [0135] 예를 들어, 신호 조정부(27)에만 신호 조정 정보가 공급된 경우, 신호 조정부(27)는 조작부(26)로부터 공급된 신호 조정 정보에 기초하여, 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 신호에 대하여 음량 조정이나 음질 조정 등의 음향 특성의 조정을 행한다. 그리고, 신호 조정부(27)는 음향 특성이 조정된 오디오 오브젝트 신호를 렌더링 처리부(30)에 공급한다. 또한, 신호 조정부(29)는 배경음 복호부(28)로부터 공급된 배경음 신호를 그대로 렌더링 처리부(30)에 공급한다.
- [0136] 구체적으로는, 예를 들어 신호 조정부(27)는 신호 조정 정보에 기초하여, 오디오 오브젝트 신호의 진폭을 증폭 또는 감쇠시킴으로써 음량 조정을 행한다. 또한, 예를 들어 신호 조정부(27)는 신호 조정 정보에 의해 나타나는 필터 계수를 사용하여, 오디오 오브젝트 신호에 대하여 필터 처리를 실시함으로써 음질 조정을 행하고, 음성에 이펙트를 부가한다.
- [0137] 이에 비해, 예를 들어 신호 조정부(29)에만 신호 조정 정보가 공급된 경우, 신호 조정부(29)는 조작부(26)로부터 공급된 신호 조정 정보에 기초하여, 배경음 복호부(28)로부터 공급된 배경음 신호에 대하여 음량 조정이나 음질 조정 등의 음향 특성의 조정을 행한다. 그리고, 신호 조정부(29)는 음향 특성이 조정된 배경음 신호를 렌더링 처리부(30)에 공급한다. 또한, 신호 조정부(27)는, 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 신호를 그대로 렌더링 처리부(30)에 공급한다.

- [0138] 또한, 예를 들어 신호 조정부(27)와 신호 조정부(29)의 양쪽에 신호 조정 정보가 공급된 경우에는, 신호 조정부(27) 및 신호 조정부(29)에 있어서, 각각 오디오 오브젝트 신호 및 배경음 신호에 대하여 음향 특성의 조정이 행해진다. 그리고, 음향 특성이 조정된 오디오 오브젝트 신호 및 배경음 신호가 렌더링 처리부(30)에 공급된다.
- [0139] 여기서, 음향 특성의 조정 방법은 다양한 방법이 있고, 유저에 의해 지정된 오디오 오브젝트의 음성의 음향 특성의 조정이 실현되면, 어떤 방법에 의해 음향 특성의 조정이 행해져도 된다. 예를 들어, 선택된 오디오 오브젝트의 오디오 오브젝트 신호 이외의 모든 오디오 오브젝트 신호 및 배경음 신호의 진폭을 감소시킴으로써, 선택된 오디오 오브젝트의 음성의 음량을 상대적으로 증가시키도록 해도 된다.
- [0140] 또한, 오디오 오브젝트 신호나 배경음 신호의 진폭 등을 직접 조정하는 것은 아니고, 도 3에 나타낸 오디오 오브젝트 정보에 포함되는 개인 정보 gain_factor[i]를 변경함으로써, 렌더링 처리부(30)에 있어서 음향 특성의 조정이 행해지도록 해도 된다.
- [0141] 그와 같은 경우, 조작부(26)는 개인 정보 gain_factor[i]의 변화량을 나타내는 정보를 파라미터로서 포함하는 신호 조정 정보를 생성하고, 렌더링 처리부(30)에 공급한다. 이와 같은 개인 정보의 변화량을 나타내는 정보는, 음성의 음량을 조정하기 위한 정보이기 때문에, 오디오 오브젝트의 음향에 관한 파라미터라고 할 수 있다.
- [0142] 렌더링 처리부(30)는 조작부(26)로부터의 신호 조정 정보에 기초하여, 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 정보에 포함되는 개인 정보를 변경하고, 변경 후의 개인 정보를 사용하여, 후술하는 스텝 S20의 처리를 행한다.
- [0143] 스텝 S20에 있어서, 렌더링 처리부(30)는, 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 정보에 기초하여, 신호 조정부(27)로부터 공급된 오디오 오브젝트 신호의 렌더링 처리를 행한다.
- [0144] 또한, 렌더링 처리부(30)는, 렌더링 처리에 의해 얻어진 오디오 오브젝트 신호와, 신호 조정부(29)로부터 공급된 배경음 신호를 합성하는 믹싱 처리를 행하고, 그 결과 얻어진 출력 오디오 신호를 출력하고, 콘텐츠 재생 처리는 종료된다.
- [0145] 예를 들어, 영상 음향 처리 장치(11)에서는, 배경음 신호에 대해서는, 종래의 2채널이나 5.1채널 등의 소위 멀티 채널 스테레오 방식으로 재생이 행해진다.
- [0146] 이에 비해 오디오 오브젝트 신호에 대해서는, 참조 문헌 1에 오디오 오브젝트의 렌더링 방식이 기술되어 있고, 이 렌더링 방식에 의해 렌더링 처리부(30)에서 렌더링이 행해지고 재생된다. 구체적으로는, VBAP(Vector Base Amplitude Panning)라고 불리는 방식으로 재생 환경의 스피커에 각 오디오 오브젝트 신호가 매핑되어 재생된다.
- [0147] 즉, 렌더링 처리부(30)에서는, 도 3에 나타낸 오디오 오브젝트 정보에 포함되는 개인 정보가 오디오 오브젝트 신호에 승산되어 개인 조정이 행해지고, 개인 조정이 행해진 오디오 오브젝트 신호에 기초하여 VBAP에서의 처리가 행해진다.
- [0148] VBAP에서는, 도 3에 나타낸 오디오 오브젝트 정보에 포함되는 오브젝트 위치 정보에 의해 나타나는, 공간 상에 있어서의 오디오 오브젝트의 위치에 가장 가까운 3개의 스피커에, 스피커마다 정한 계인으로 오디오 오브젝트 신호가 매핑된다.
- [0149] 이와 같이 VBAP는 오브젝트 위치 정보에 의해 나타나는, 공간 상에 있어서의 오디오 오브젝트의 위치에 가장 가까운 위치에 있는 3개의 스피커의 출력을 사용하여, 오디오 오브젝트의 공간 상의 위치에 소리를 정위시키는 기술이다.
- [0150] VBAP에 대해서는, 예를 들어 「Virtual Sound Source Positioning Using Vector Base Amplitude Panning, AES Volume 45 Issue 6 pp.456-266, June 1997」(이하, 참조 문헌 2이라고도 칭함)에 상세하게 기재되어 있다. 또한, 참조 문헌 1이나 참조 문헌 2에서는, 스피커의 개수는 3개이지만 4개 이상의 스피커 수로 소리를 정위시키는 것도 물론 가능하다.
- [0151] 또한, 이상에 있어서는, 배경음 신호는 멀티 채널 스테레오 방식으로 부호화 및 복호되는 예에 대하여 설명했지만, 참조 문헌 1의 규격에 기술되어 있는 HOA(Higher Order Ambisonic)라고 불리는 시청자 근방의 공간을, 구면 조화 함수를 직교 기저로서 표현하는 방식을 사용하도록 해도 된다.
- [0152] 이상과 같이 하여 영상 음향 처리 장치(11)는 오디오 오브젝트 정보에 기초하여 오디오 오브젝트 정보 화상을

생성하고, 콘텐츠의 화상에 중첩 표시함과 함께, 유저의 조작에 따라 신호 조정 정보를 생성하고, 음성의 음향 특성의 조정을 행한다.

- [0153] 이와 같이 콘텐츠의 화상에, 오디오 오브젝트 정보를 사용하여 얻어지는 오디오 오브젝트 정보 화상을 중첩 표시시킴으로써, 유저는 더 간단하고 또한 직감적으로 오디오 오브젝트를 선택하고, 음량 조정 등의 음향 특성의 조정을 행할 수 있다.
- [0154] <제2 실시 형태>
- [0155] <음향 처리 장치의 구성예>
- [0156] 그런데, 상술한 제1 실시 형태에서는, 콘텐츠의 화상(영상)에 오디오 오브젝트의 오브젝트 위치 정보를 사용하여 얻어진 오디오 오브젝트 정보 화상을 중첩 표시하는 예에 대하여 설명했다. 그러나, 본 기술은, 콘텐츠로서 영상을 수반하지 않는 경우라도 적용 가능하다.
- [0157] 제2 실시 형태에서는, 입력 비트 스트림에 영상 비트 스트림이 포함되지 않는 경우, 즉 입력 비트 스트림에 오디오 오브젝트 비트 스트림과 배경음 비트 스트림만이 포함되는 경우에 대하여 설명한다.
- [0158] 그와 같은 경우, 본 기술을 적용한 음향 처리 장치는, 예를 들어 도 7에 나타내는 구성이 된다. 또한, 도 7에 있어서, 도 1에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 적절히 생략한다.
- [0159] 도 7에 나타내는 음향 처리 장치(81)는, 비다중화부(21), 영상 표시부(23), 오디오 오브젝트 복호부(24), 오디오 오브젝트 정보 표시 제어부(25), 조작부(26), 신호 조정부(27), 배경음 복호부(28), 신호 조정부(29) 및 렌더링 처리부(30)를 갖고 있다.
- [0160] 이 음향 처리 장치(81)의 구성은, 영상 복호부(22)가 설치되어 있지 않은 점에서 영상 음향 처리 장치(11)의 구성과 상이하고, 그 밖의 점에서는 영상 음향 처리 장치(11)와 동일한 구성으로 되어 있다.
- [0161] 단, 음향 처리 장치(81)에서는, 오디오 오브젝트 복호부(24)에서 얻어진 오디오 오브젝트 정보가 조작부(26)에도 공급된다. 또한, 조작부(26)는 기기 사용자인 유저의 조작에 따라, 적절히, 오디오 오브젝트의 오브젝트 위치 정보를 변경하고, 변경 후의 오브젝트 위치 정보를 렌더링 처리부(30)에 공급한다.
- [0162] 예를 들어, 영상을 수반하지 않는 순수한 음성만의 콘텐츠를 재생하는 경우, 오디오 오브젝트의 위치를 임의의 위치로 변경할 수 있다. 이는, 콘텐츠에 영상이 포함되어 있을 때에는, 오디오 오브젝트의 위치를 이동시키면, 오디오 오브젝트의 위치와, 그 오디오 오브젝트에 대응하는 영상 오브젝트의 위치에 어긋남이 발생하지만, 음성만의 콘텐츠에서는 그와 같은 것은 발생하지 않기 때문이다.
- [0163] 이 경우에 있어서도, 오디오 오브젝트 비트 스트림에는, 오디오 오브젝트 정보가 포함되어 있기 때문에, 영상 표시부(23)에 오디오 오브젝트 정보 화상을 표시시킬 수 있다. 따라서, 기기 사용자인 유저는, 이 오디오 오브젝트 정보 화상을 봄으로써, 오디오 오브젝트의 위치를 시각적으로 확인하면서, 콘텐츠의 가공이나 편집을 행할 수 있다. 이와 같은 실시 형태는, 예를 들어 스튜디오에서 음성만을 포함하는 콘텐츠의 편집 작업을 행하는 경우에 적합하다.
- [0164] 음향 처리 장치(81)에서는, 예를 들어 도 8에 나타낸 바와 같이 오디오 오브젝트 정보 화상의 표시가 행해진다.
- [0165] 도 8에 나타내는 예에서는, 영상 표시부(23)의 표시 화면에는, 오브젝트 위치 표시 영역 R11, 오브젝트 메타데이터 표시 영역 R12 및 오브젝트 위치 시간 천이 표시 영역 R13이 설치되어 있다.
- [0166] 오브젝트 위치 표시 영역 R11에는 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상이 표시된다. 이 실시 형태에서는, 오디오 오브젝트 정보 화상은 오브젝트 위치 정보에 의해 나타나는 위치에 표시된다.
- [0167] 이 예에서는, 오브젝트 위치 표시 영역 R11에는, 화살표 A11에 나타내는 3차원 직교 좌표계의 각 축이 표시되어 있음과 함께, 2개의 오디오 오브젝트의 각각의 위치를 나타내는 오디오 오브젝트 정보 화상 FR31 및 오디오 오브젝트 정보 화상 FR32가 표시되어 있다. 여기서, 화살표 A11에 나타내는 3차원 직교 좌표계는, 도 4에 나타낸 X축, Y축 및 Z축을 각 축으로 하는 3차원 직교 좌표계이다.
- [0168] 또한, 이 예에서는, 오디오 오브젝트 정보 화상 FR31이 점선으로 그려져 있는 것에 비해, 오디오 오브젝트 정보 화상 FR32는 실선으로 그려져 있고, 오디오 오브젝트 정보 화상 FR32가 선택 상태가 되고 하이라이트 표시되어 있는 모습이 표현되어 있다. 즉, 각 오디오 오브젝트 정보 화상은, 선택 상태로 되어 있을 때와, 그렇지 않을 때에 상이한 표시 형식으로 표시된다. 또한, 도 8에서는, 오디오 오브젝트 정보 화상 FR32가 이동되는 모습이

표현되어 있다.

- [0169] 기기 사용자인 유저는, 오브젝트 위치 표시 영역 R11에 표시되는 오디오 오브젝트 정보 화상을 봄으로써, 공간 상에 있어서의 오디오 오브젝트의 위치를 시각적으로 확인할 수 있다.
- [0170] 오브젝트 메타데이터 표시 영역 R12에는 오디오 오브젝트 비트 스트림으로부터 추출된, 선택 상태로 되어 있는 오디오 오브젝트의 메타데이터, 즉 오디오 오브젝트 정보에 포함되어 있는 정보가 표시된다. 예를 들어 오디오 오브젝트 정보에 포함되어 있는 정보로서, 오브젝트 위치 정보나 계인 정보가 표시된다.
- [0171] 오브젝트 위치 시간 천이 표시 영역 R13에는, 선택 상태가 된 오디오 오브젝트 정보 화상, 즉 선택 상태가 된 오디오 오브젝트의 각 시각에 있어서의 공간 상의 위치가 표시된다.
- [0172] 구체적으로는, 오브젝트 위치 시간 천이 표시 영역 R13에는 X좌표 표시 영역 R21, Y좌표 표시 영역 R22 및 Z좌표 표시 영역 R23이 설치되어 있고, 그들의 X좌표 표시 영역 R21 내지 Z좌표 표시 영역 R23의 도면 중, 횡방향은 시간 방향을 나타내고 있다.
- [0173] X좌표 표시 영역 R21에는, 오브젝트 위치 표시 영역 R11에 표시되어 있는 오디오 오브젝트 정보 화상에 대응하는 오디오 오브젝트 중, 선택 상태로 되어 있는 오디오 오브젝트의 각 시각에 있어서의 공간 상의 X축 방향의 위치인 X좌표를 나타내는 위치 천이 정보 PL11이 표시된다. 바꾸어 말하면, 위치 천이 정보 PL11은 오디오 오브젝트의 X좌표의 시간 천이를 나타내는 정보이다.
- [0174] 마찬가지로, Y좌표 표시 영역 R22에는, 선택 상태로 되어 있는 오디오 오브젝트의 각 시각에 있어서의 공간 상의 Y축 방향의 위치인 Y좌표를 나타내는 위치 천이 정보 PL12가 표시된다. Z좌표 표시 영역 R23에는, 선택 상태로 되어 있는 오디오 오브젝트의 각 시각에 있어서의 공간 상의 Z축 방향의 위치인 Z좌표를 나타내는 위치 천이 정보 PL13이 표시된다.
- [0175] 또한, 이들의 X좌표 표시 영역 R21 내지 Z좌표 표시 영역 R23에 있어서의 시간축 상의 1시각의 위치에는, 커서 CR11이 표시된다.
- [0176] 이 예에서는, 기기 사용자인 유저가 조작부(26)를 조작함으로써, 오브젝트 위치 표시 영역 R11에 표시되어 있는 원하는 오디오 오브젝트 정보 화상을 선택하면, 그 오디오 오브젝트 정보 화상의 오디오 오브젝트의 각 시각에 있어서의 위치를 나타내는 위치 천이 정보 PL11 내지 위치 천이 정보 PL13이 표시된다. 도 8의 예에서는, 오디오 오브젝트 정보 화상 FR32에 대응하는 오디오 오브젝트가 선택 상태로 되어 있다.
- [0177] 또한, 유저는 커서 CR11을 시간축 방향의 원하는 위치로 이동시킴으로써, 소정의 시각을 지정할 수 있다. 커서 CR11에 의해 시각이 지정되면, 오브젝트 위치 표시 영역 R11에는 커서 CR11에 의해 나타나는 시각에 있어서의 각 오디오 오브젝트의 공간 상의 위치에, 그들의 오디오 오브젝트의 오디오 오브젝트 정보 화상이 표시된다.
- [0178] 따라서, 이 예에서는, 위치 천이 정보 PL11 내지 위치 천이 정보 PL13에 있어서의 커서 CR11이 위치하고 있는 시각의 X좌표, Y좌표 및 Z좌표에 의해 나타나는 공간 상의 위치에, 선택 상태의 오디오 오브젝트의 오디오 오브젝트 정보 화상 FR32가 표시되어 있다.
- [0179] 이와 같은 상태에서, 도 8에 나타나 있는 바와 같이, 유저가 선택 상태로 되어 있는 오디오 오브젝트 정보 화상 FR32를 이동시키면, 위치 천이 정보 PL11 내지 위치 천이 정보 PL13에 있어서의 커서 CR11이 위치하고 있는 시각의 X좌표, Y좌표 및 Z좌표가 변경된다. 즉, 유저에 의한 오디오 오브젝트 정보 화상 FR32의 이동 조작이 위치 천이 정보 PL11 내지 위치 천이 정보 PL13에 대하여 반영된다.
- [0180] 또한, 유저가 선택 상태로 하는 오디오 오브젝트 정보 화상을, 오디오 오브젝트 정보 화상 FR31로 변경하면, 오브젝트 메타데이터 표시 영역 R12 및 오브젝트 위치 시간 천이 표시 영역 R13의 표시도, 새롭게 선택 상태로 된 오디오 오브젝트 정보 화상 FR31의 것으로 갱신된다.
- [0181] 또한, 이 예에서는, 오디오 오브젝트의 공간 상의 위치의 변경 조작을 용이하게 행할 수 있도록, 유저가 화살표 A11에 나타내는 3차원 직교 좌표계 자체를 회전시키거나, 확대 축소시키거나 할 수 있도록 되어 있다.
- [0182] 이상과 같이, 음향 처리 장치(81)에 의하면, 입력 비트 스트림에 포함되는 오디오 오브젝트의 가공이나 편집의 작업을 시각적으로 확인하면서 용이하게 행할 수 있다.
- [0183] 또한, 도 8에 나타내는 예에서는, 도 6을 참조하여 설명한 경우와 마찬가지로, 오디오 오브젝트를 선택하여 조정 지시 화상 등을 표시시키고, 그 선택한 오디오 오브젝트에 대하여, 음량 조정이나 음질 조정 등의 음향 특성

의 조정을 행할 수도 있다.

- [0184] <콘텐츠 재생 처리의 설명>
- [0185] 이어서, 음향 처리 장치(81)의 동작에 대하여 설명한다. 즉, 이하, 도 9의 흐름도를 참조하여, 음향 처리 장치(81)에 의해 행해지는 콘텐츠 재생 처리에 대하여 설명한다.
- [0186] 콘텐츠 재생 처리가 개시되면, 스텝 S51의 처리가 행해지지만, 이 처리는 도 2의 스텝 S11의 처리와 마찬가지로, 그 설명은 생략한다. 단, 스텝 S51에서는, 입력 비트 스트림이 오디오 오브젝트 비트 스트림과 배경음 비트 스트림에 비다중화된다.
- [0187] 또한, 입력 비트 스트림이 비다중화되면, 스텝 S52 및 스텝 S53의 처리가 행해지지만, 이들의 처리는 도 2의 스텝 S13 및 스텝 S14의 처리와 마찬가지로, 그 설명은 생략한다. 단, 스텝 S53에서는, 오디오 오브젝트 비트 스트림의 복호에 의해 얻어진 오디오 오브젝트 정보가, 오디오 오브젝트 정보 표시 제어부(25), 조작부(26) 및 렌더링 처리부(30)에 공급된다.
- [0188] 스텝 S54에 있어서, 오디오 오브젝트 정보 표시 제어부(25)는, 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 정보에 기초하여, 영상 표시부(23)를 제어하고, 오디오 오브젝트 정보 화상을 표시시킨다.
- [0189] 즉, 오디오 오브젝트 정보 표시 제어부(25)는 오디오 오브젝트 정보에 기초하여 오디오 오브젝트 정보 화상을 생성하고, 그 오디오 오브젝트 정보 화상을 영상 표시부(23)에 공급하여 표시시킨다.
- [0190] 이에 의해, 영상 표시부(23)에는, 예를 들어 도 8에 나타난 화면이 표시된다. 즉, 스텝 S54의 처리에 의해, 영상 표시부(23)에는 오디오 오브젝트 정보에 포함되는 오브젝트 위치 정보에 의해 나타나는 위치에 오디오 오브젝트 정보 화상이 표시됨과 함께, 오디오 오브젝트의 메타데이터나 위치 천이 정보도 표시된다.
- [0191] 이와 같이 하여 오디오 오브젝트 정보 화상이 표시되면, 기기 사용자인 유저는, 조작부(26)를 조작하여, 오디오 오브젝트의 위치를 변경하거나, 음량 조정이나 음질 조정을 행하거나 한다.
- [0192] 스텝 S55에 있어서, 조작부(26)는 유저의 조작에 따라, 오디오 오브젝트의 오브젝트 위치 정보를 변경한다.
- [0193] 예를 들어, 도 8에 나타난 예에 있어서, 유저가 오디오 오브젝트 정보 화상 FR32를 이동시킨 경우, 조작부(26)는, 그 오디오 오브젝트 정보 화상 FR32의 이동에 따라, 대응하는 오디오 오브젝트의 오브젝트 위치 정보를 변경한다.
- [0194] 오브젝트 위치 정보는 렌더링 처리에 사용되고, 공간 상에 있어서의 오디오 오브젝트의 위치, 즉 공간 상에 있어서의 오디오 오브젝트의 음상 정위 위치를 지정하는 정보이다. 따라서, 오브젝트 위치 정보를 변경하는 처리는 오디오 오브젝트의 음향에 관한 파라미터를 설정하는 처리라고 할 수 있다.
- [0195] 스텝 S56에 있어서, 조작부(26)는 유저의 조작에 따라, 신호 조정 정보를 생성한다. 스텝 S56에서는, 도 2의 스텝 S18과 동일한 처리가 행해진다. 또한, 스텝 S56에서는, 오디오 오브젝트 정보 화상의 위치의 이동에 따라, 음향 특성의 조정을 위한 파라미터가 설정되고, 그 파라미터를 포함하는 신호 조정 정보가 생성되도록 해도 된다.
- [0196] 조작부(26)는, 이상과 같이 하여 생성한 신호 조정 정보를, 신호 조정부(27), 신호 조정부(29) 또는 신호 조정부(27) 및 신호 조정부(29)에 공급한다. 또한, 조작부(26)는 스텝 S55의 처리에 의해 얻어진, 변경 후의 오브젝트 위치 정보를 렌더링 처리부(30)에 공급한다.
- [0197] 신호 조정 정보가 생성되면, 그 후, 스텝 S57 및 스텝 S58의 처리가 행해져 콘텐츠 재생 처리는 종료되지만, 이들의 처리는 도 2의 스텝 S19 및 스텝 S20의 처리와 마찬가지로, 그 설명은 생략한다.
- [0198] 단, 스텝 S58에서는, 렌더링 처리부(30)는 조작부(26)로부터 공급된, 변경 후의 오브젝트 위치 정보와, 오디오 오브젝트 복호부(24)로부터 공급된 오디오 오브젝트 정보에 포함되는 개인 정보를 사용하여 렌더링 처리를 행한다.
- [0199] 이상과 같이 하여, 음향 처리 장치(81)는 오디오 오브젝트 정보에 기초하여, 오디오 오브젝트 정보 화상을 생성하여 표시함과 함께, 유저의 조작에 따라 신호 조정 정보를 생성하여 음성의 음향 특성의 조정을 행하거나, 오브젝트 위치 정보를 변경하거나 한다. 이에 의해, 유저는, 더 간단하고 또한 직감적으로 오디오 오브젝트를 선택하고, 음향 특성의 조정을 행하거나, 오디오 오브젝트의 위치를 이동시키거나 할 수 있다. 즉, 오디오 오브젝트의 가공이나 편집의 작업을 시각적으로 확인하면서 용이하게 행할 수 있다.

- [0200] 그런데, 상술한 일련의 처리는, 하드웨어에 의해 실행할 수도 있고, 소프트웨어에 의해 실행할 수도 있다. 일련의 처리를 소프트웨어에 의해 실행하는 경우에는, 그 소프트웨어를 구성하는 프로그램이 컴퓨터에 인스톨된다. 여기서, 컴퓨터에는 전용의 하드웨어에 내장되어 있는 컴퓨터나, 각종 프로그램을 인스톨함으로써, 각종 기능을 실행하는 것이 가능한, 예를 들어 범용의 퍼스널 컴퓨터 등이 포함된다.
- [0201] 도 10은 상술한 일련의 처리를 프로그램에 의해 실행하는 컴퓨터의 하드웨어의 구성예를 나타내는 블록도이다.
- [0202] 컴퓨터에 있어서, CPU(Central Processing Unit)(501), ROM(Read Only Memory)(502), RAM(Random Access Memory)(503)은 버스(504)에 의해 서로 접속되어 있다.
- [0203] 버스(504)에는 입출력 인터페이스(505)가 더 접속되어 있다. 입출력 인터페이스(505)에는 입력부(506), 출력부(507), 기록부(508), 통신부(509) 및 드라이브(510)가 접속되어 있다.
- [0204] 입력부(506)는 키보드, 마우스, 마이크로폰, 촬상 소자 등을 포함한다. 출력부(507)는 디스플레이, 스피커 등을 포함한다. 기록부(508)는 하드 디스크나 불휘발성의 메모리 등을 포함한다. 통신부(509)는 네트워크 인터페이스 등을 포함한다. 드라이브(510)는 자기 디스크, 광 디스크, 광자기 디스크, 또는 반도체 메모리 등의 리무버블 기록 매체(511)를 구동한다.
- [0205] 이상과 같이 구성되는 컴퓨터에서는, CPU(501)가, 예를 들어 기록부(508)에 기록되어 있는 프로그램을, 입출력 인터페이스(505) 및 버스(504)를 통해, RAM(503)에 로드하여 실행함으로써, 상술한 일련의 처리가 행해진다.
- [0206] 컴퓨터(CPU(501))가 실행하는 프로그램은, 예를 들어 패키지 미디어 등으로서의 리무버블 기록 매체(511)에 기록하여 제공할 수 있다. 또한, 프로그램은 로컬 에어리어 네트워크, 인터넷, 디지털 위성 방송 등의, 유선 또는 무선의 전송 매체를 통해 제공할 수 있다.
- [0207] 컴퓨터에서는, 프로그램은 리무버블 기록 매체(511)를 드라이브(510)에 장착함으로써, 입출력 인터페이스(505)를 통해, 기록부(508)에 인스톨할 수 있다. 또한, 프로그램은 유선 또는 무선의 전송 매체를 통해, 통신부(509)에서 수신하고, 기록부(508)에 인스톨할 수 있다. 그 밖에, 프로그램은 ROM(502)이나 기록부(508)에 미리 인스톨해 둘 수 있다.
- [0208] 또한, 컴퓨터가 실행하는 프로그램은, 본 명세서에서 설명하는 순서에 따라 시계열로 처리가 행해지는 프로그램이어도 되고, 병렬로, 혹은 호출이 행해졌을 때 등의 필요한 타이밍에 처리가 행해지는 프로그램이어도 된다.
- [0209] 또한, 본 기술의 실시 형태는 상술한 실시 형태에 한정되는 것은 아니고, 본 기술의 요지를 이탈하지 않는 범위 내에 있어서 다양한 변경이 가능하다.
- [0210] 예를 들어, 본 기술은 하나의 기능을 네트워크를 통해 복수의 장치에서 분담, 공동하여 처리하는 클라우드 컴퓨팅의 구성을 취할 수 있다.
- [0211] 또한, 상술한 흐름도에서 설명한 각 스텝은 하나의 장치에서 실행하는 것 외에, 복수의 장치에서 분담하여 실행할 수 있다.
- [0212] 또한, 하나의 스텝에 복수의 처리가 포함되는 경우에는, 그 하나의 스텝에 포함되는 복수의 처리는 하나의 장치에서 실행하는 것 외에, 복수의 장치에서 분담하여 실행할 수 있다.
- [0213] 또한, 본 기술은 이하의 구성으로 하는 것도 가능하다.
- [0214] (1)
- [0215] 오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키는 표시 제어부와,
- [0216] 1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하는 선택부
- [0217] 를 구비하는 음향 처리 장치.
- [0218] (2)
- [0219] 상기 선택부에 의해 선택된 상기 오디오 오브젝트의 음향에 관한 파라미터를 설정하는 파라미터 설정부를 더 구비하는
- [0220] (1) 에 기재된 음향 처리 장치.

- [0221] (3)
- [0222] 상기 파라미터에 기초하여, 상기 오디오 오브젝트의 오디오 오브젝트 신호 및 배경음의 배경음 신호의 적어도 어느 한쪽에 대하여, 상기 오디오 오브젝트의 음성의 음향 특성을 조정하기 위한 처리를 행하는 신호 조정부를 더 구비하는
- [0223] (2)에 기재된 음향 처리 장치.
- [0224] (4)
- [0225] 상기 파라미터는 음량 조정 또는 음질 조정을 위한 파라미터인
- [0226] (3)에 기재된 음향 처리 장치.
- [0227] (5)
- [0228] 상기 오디오 오브젝트의 오디오 오브젝트 신호의 렌더링 처리를 행하는 렌더링 처리부를 더 구비하는
- [0229] (2) 내지 (4) 중 어느 한 항에 기재된 음향 처리 장치.
- [0230] (6)
- [0231] 상기 파라미터는 상기 오디오 오브젝트의 위치를 지정하는 파라미터이고,
- [0232] 상기 렌더링 처리부는 상기 파라미터에 기초하여 상기 렌더링 처리를 행하는
- [0233] (5)에 기재된 음향 처리 장치.
- [0234] (7)
- [0235] 상기 표시 제어부는, 상기 표시부에 표시되어 있는, 상기 오디오 오브젝트의 음성이 부수되는 영상 상에 있어서의 상기 오브젝트 위치 정보에 의해 정해지는 위치에 상기 오디오 오브젝트 정보 화상을 중첩 표시시키는
- [0236] (1) 내지 (6) 중 어느 한 항에 기재된 음향 처리 장치.
- [0237] (8)
- [0238] 상기 표시 제어부는, 상기 오브젝트 위치 정보에 의해 정해지는 위치가 상기 표시부의 표시 화면 밖에 있는 경우, 상기 표시 화면의 단 부분에 상기 오디오 오브젝트 정보 화상을 표시시키는
- [0239] (7)에 기재된 음향 처리 장치.
- [0240] (9)
- [0241] 상기 선택부는, 유저에 의한 상기 오디오 오브젝트 정보 화상의 위치에서의 지정 조작에 따라, 상기 오디오 오브젝트를 선택하는
- [0242] (7) 또는 (8)에 기재된 음향 처리 장치.
- [0243] (10)
- [0244] 오디오 오브젝트 비트 스트림을 복호하고, 상기 오디오 오브젝트의 오디오 오브젝트 신호와, 상기 오브젝트 위치 정보를 얻는 오디오 오브젝트 복호부를 더 구비하는
- [0245] (1) 내지 (9) 중 어느 한 항에 기재된 음향 처리 장치.
- [0246] (11)
- [0247] 오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트 정보 화상을 표시부에 표시시키고,
- [0248] 1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하는
- [0249] 스텝을 포함하는 음향 처리 방법.
- [0250] (12)
- [0251] 오디오 오브젝트의 오브젝트 위치 정보에 기초하여, 상기 오디오 오브젝트의 위치를 나타내는 오디오 오브젝트

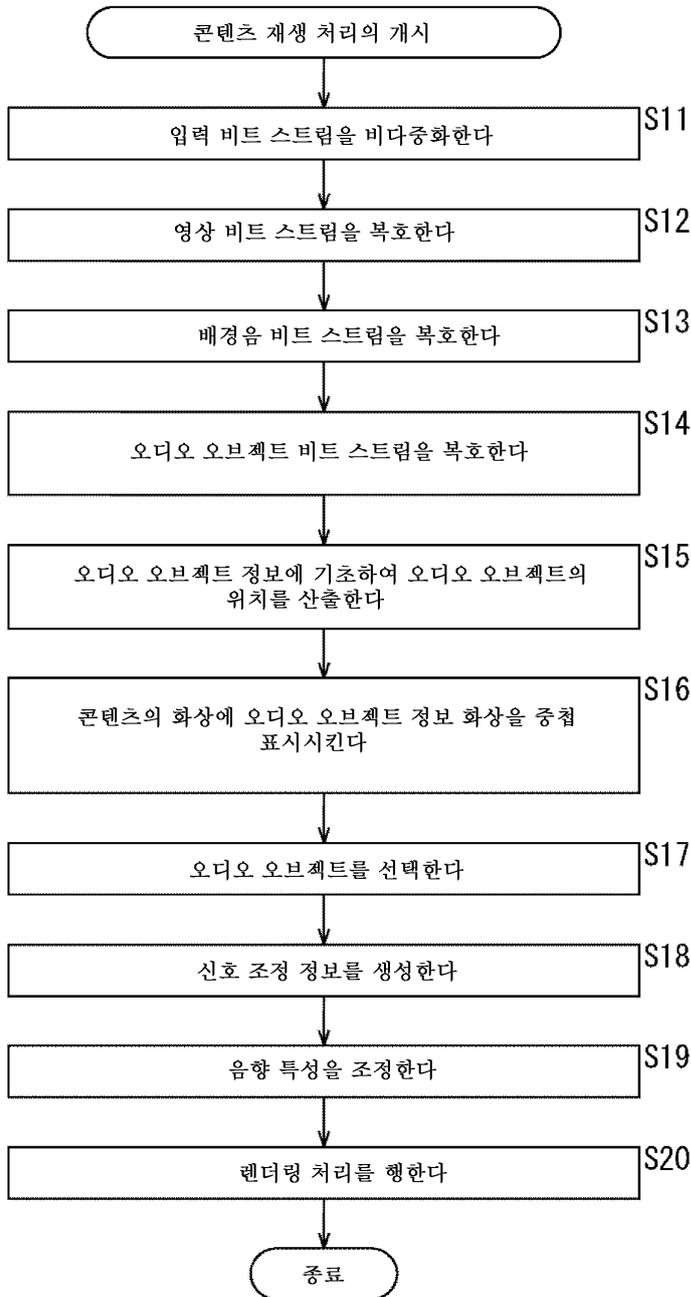
정보 화상을 표시부에 표시시키고,

- [0252] 1 또는 복수의 상기 오디오 오브젝트 중에서, 소정의 상기 오디오 오브젝트를 선택하는
- [0253] 스텝을 포함하는 처리를 컴퓨터에 실행시키는 프로그램.

부호의 설명

- [0254] 11 : 영상 음향 처리 장치
- 21 : 비다중화부
- 23 : 영상 표시부
- 24 : 오디오 오브젝트 복호부
- 25 : 오디오 오브젝트 정보 표시 제어부
- 26 : 조작부
- 27 : 신호 조정부
- 28 : 배경음 복호부
- 29 : 신호 조정부
- 30 : 렌더링 처리부

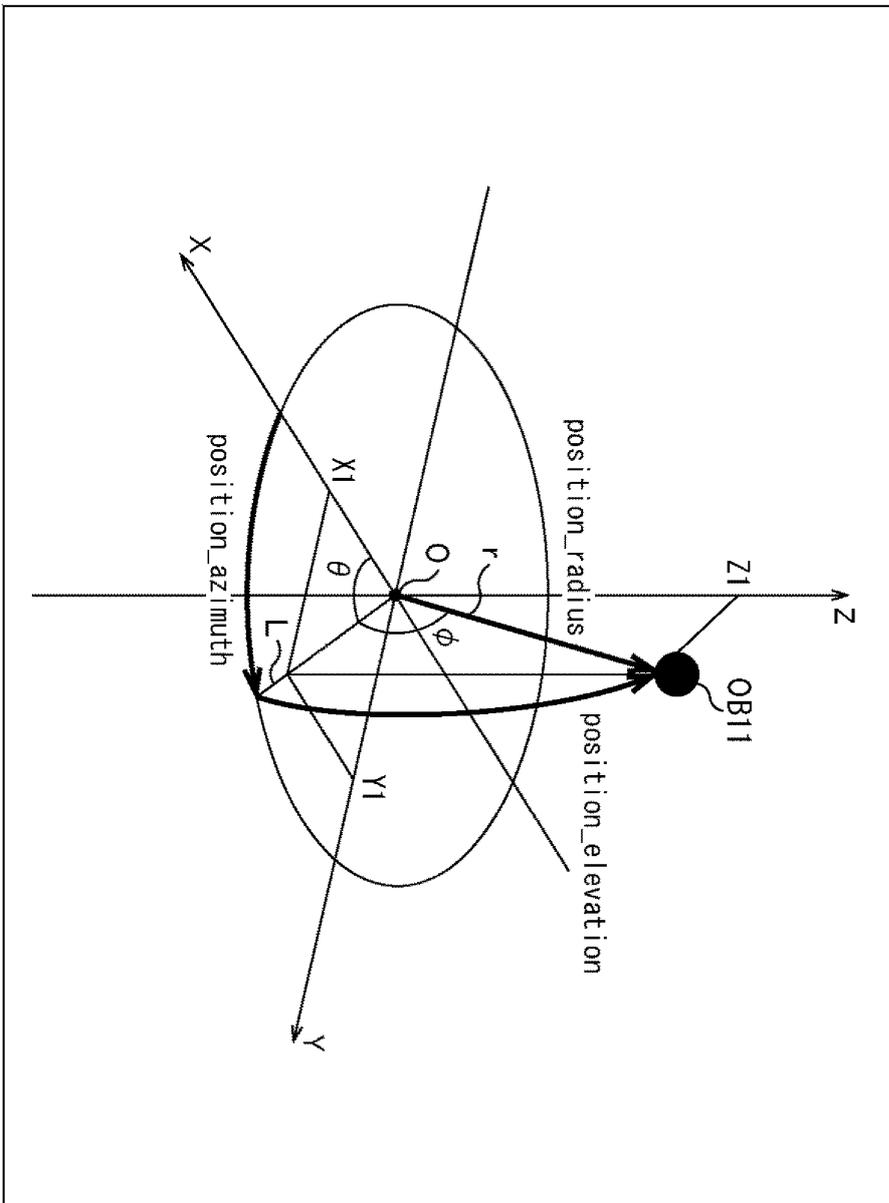
도면2



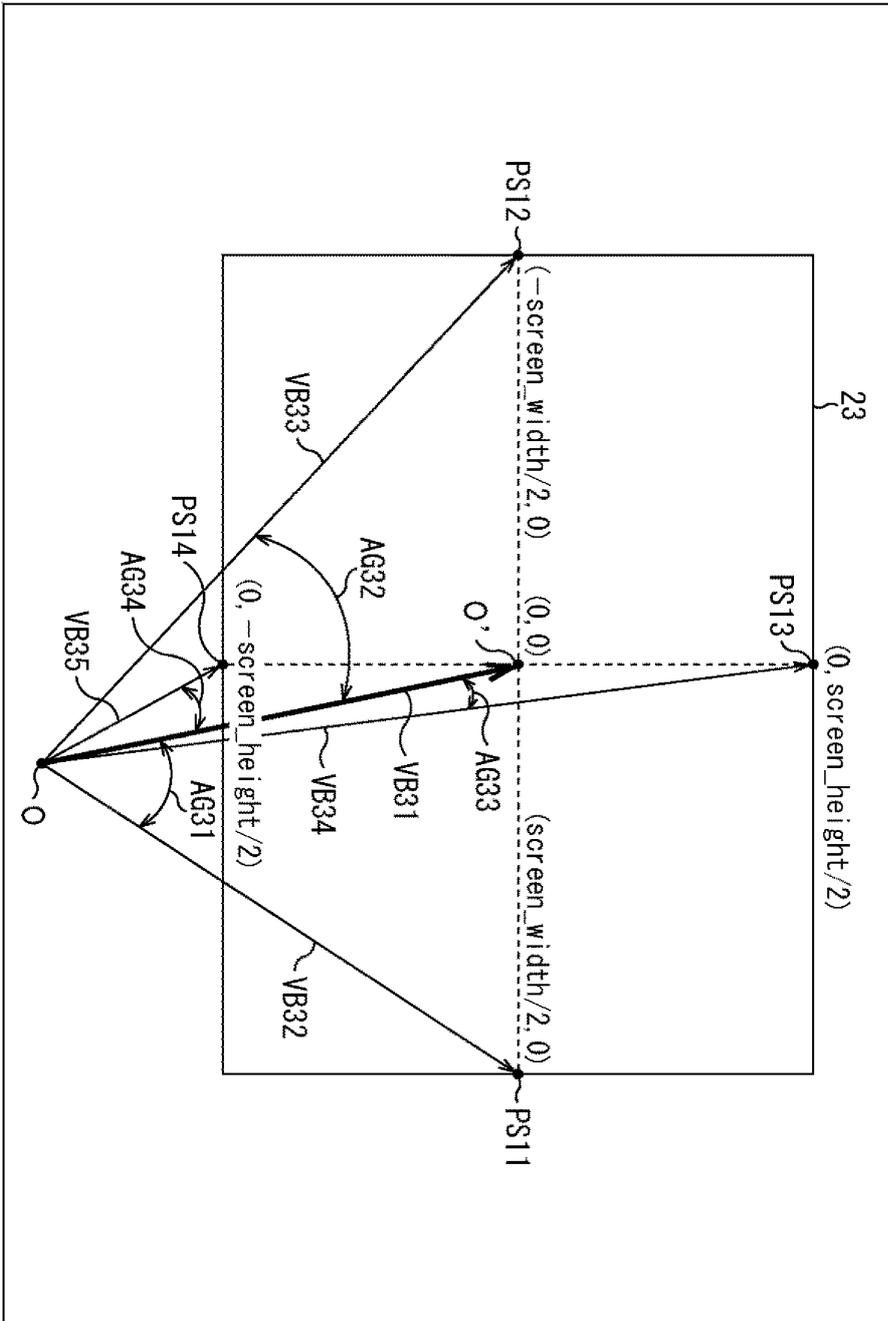
도면3

Syntax	No. of bits	Mnemonic
<pre> object_metadata () { for (i=0; i < num_objects; i++) { position_azimuth[i]; position_elevation[i]; position_radius[i]; gain_factor[i]; } } </pre>	<p>8</p> <p>6</p> <p>4</p> <p>7</p>	<p>tcimbsf</p> <p>tcimbsf</p> <p>uimbsf</p> <p>tcimbsf</p>

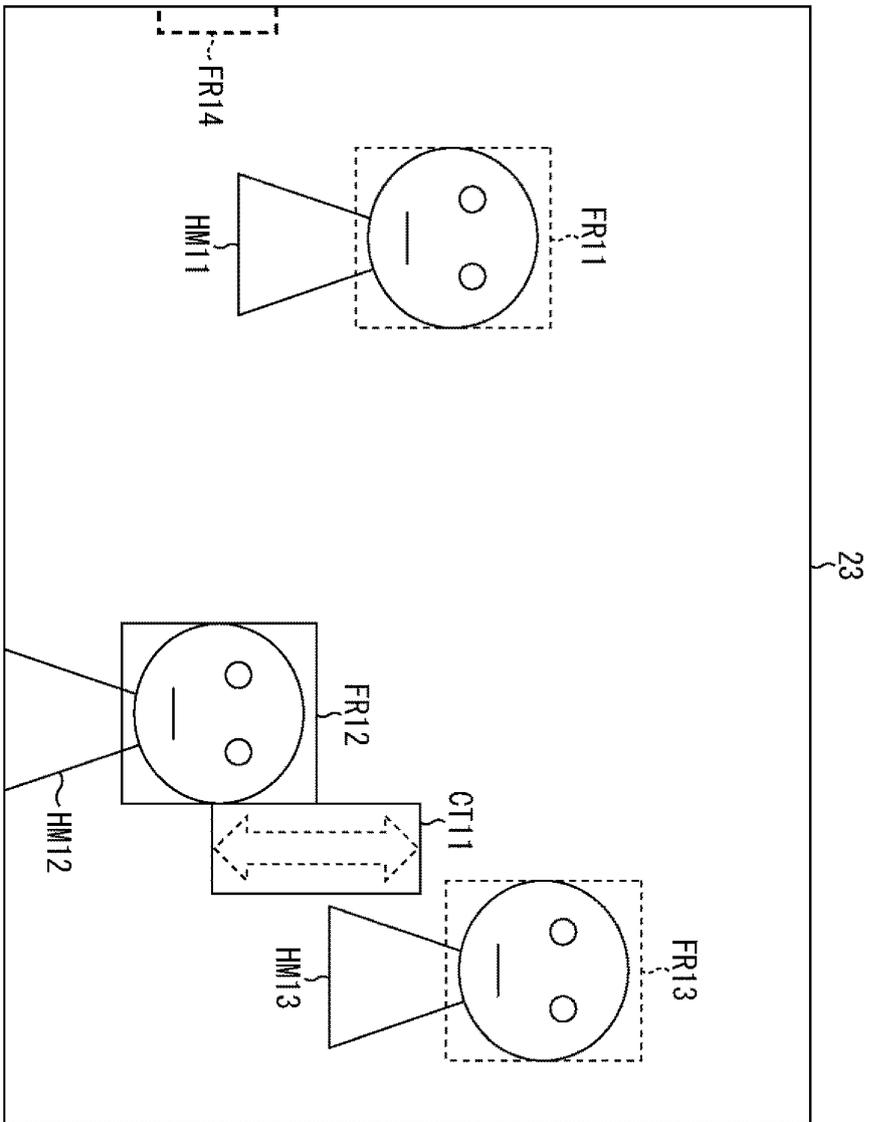
도면4



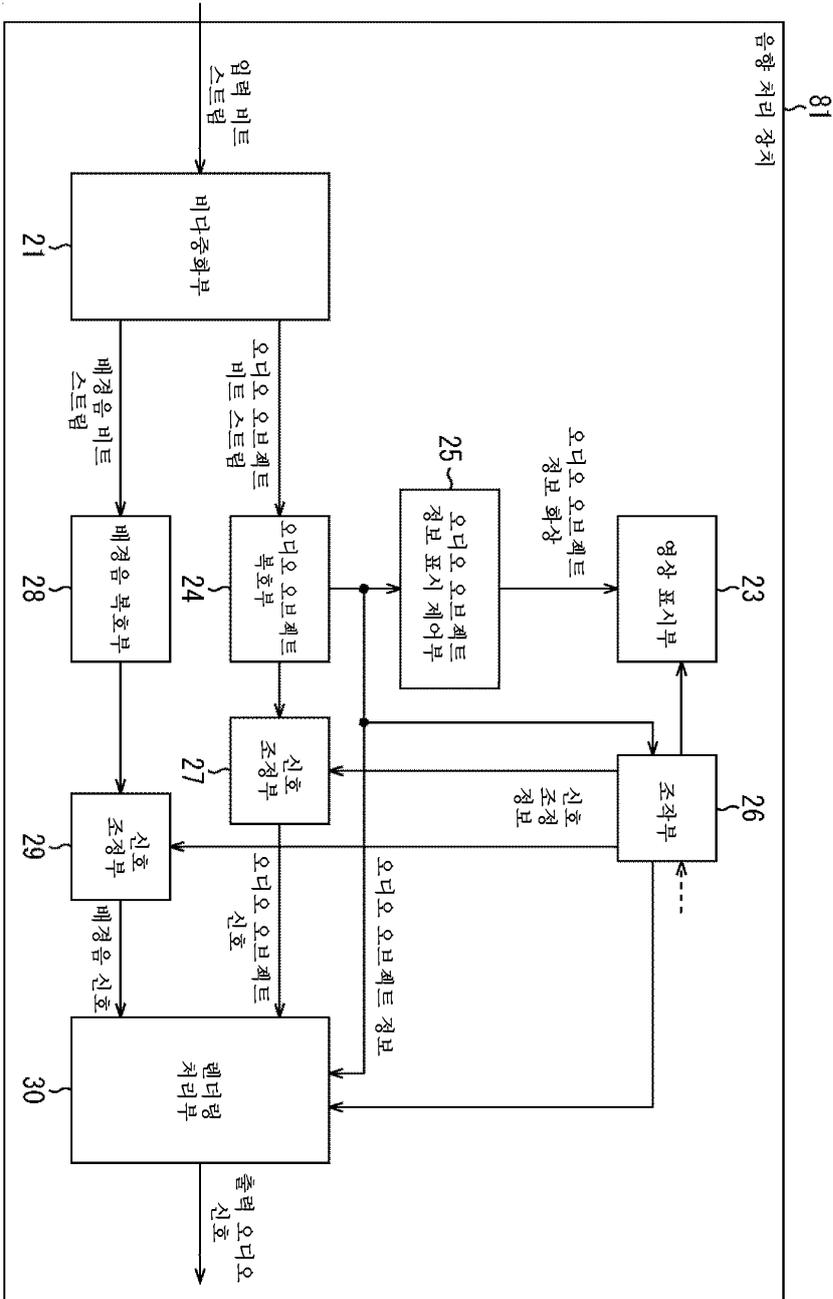
도면5



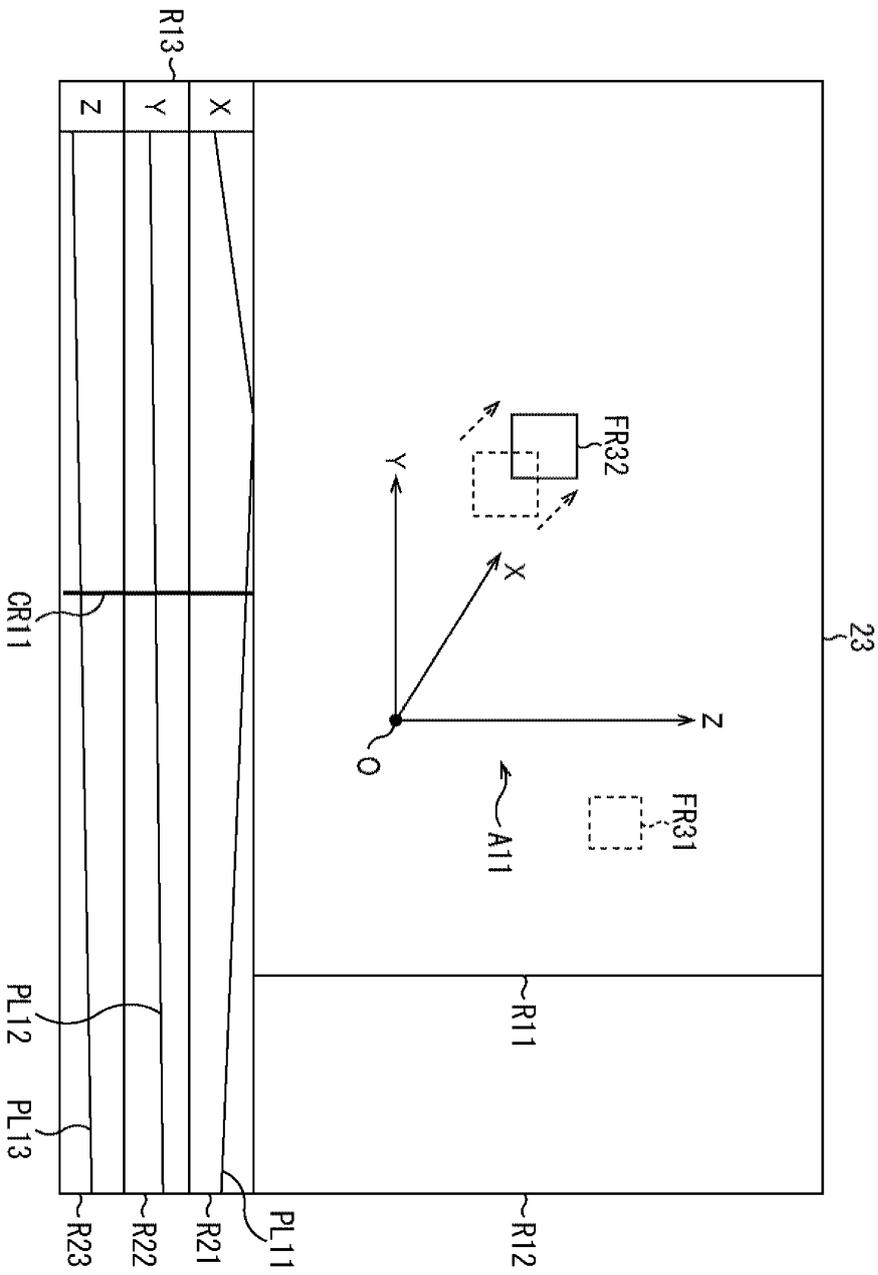
도면6



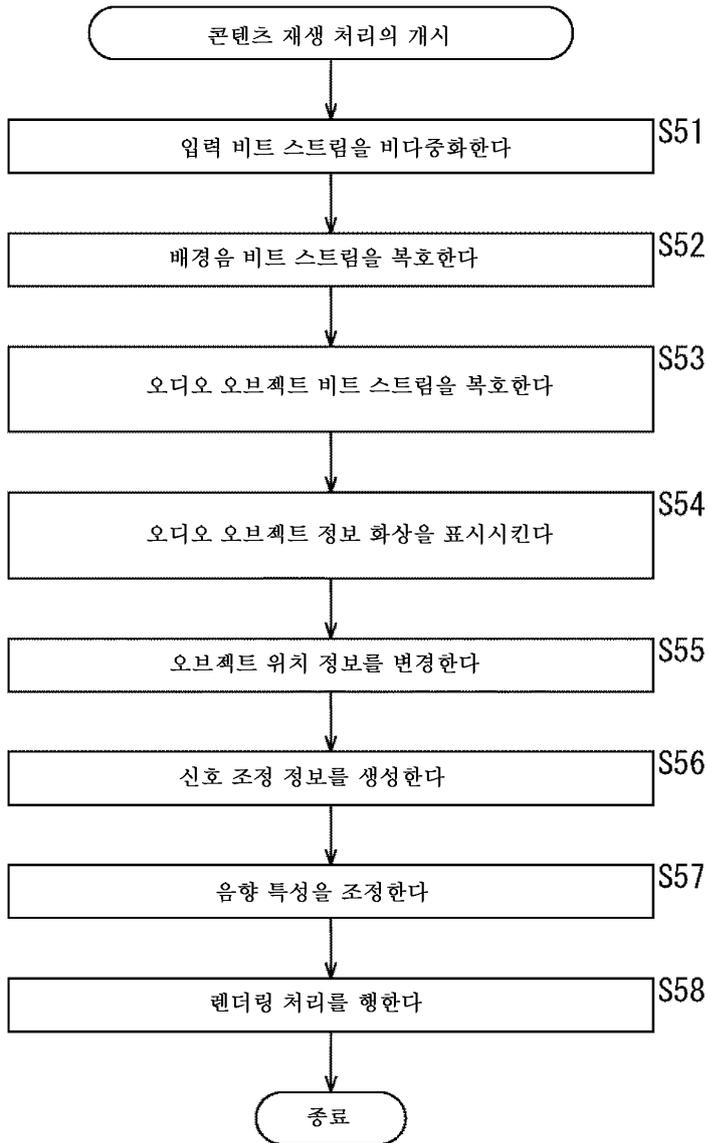
도면7



도면8



도면9



도면10

