



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2018년12월18일

(11) 등록번호 10-1930671

(24) 등록일자 2018년12월12일

- (51) 국제특허분류(Int. Cl.)
G10L 19/008 (2014.01) H04S 5/02 (2006.01)

- (52) CPC특허분류
G10L 19/008 (2013.01)
H04S 5/02 (2013.01)

(21) 출원번호 10-2017-7035890

(22) 출원일자(국제) 2016년06월09일

심사청구일자 2018년01월05일

(85) 번역문제출일자 2017년12월13일

(65) 공개번호 10-2018-0008609

(43) 공개일자 2018년01월24일

(86) 국제출원번호 PCT/JP2016/067195

(87) 국제공개번호 WO 2016/208406

국제공개일자 2016년12월29일

(30) 우선권주장

JP-P-2015-126650 2015년06월24일 일본(JP)

JP-P-2015-148683 2015년07월28일 일본(JP)

(56) 선행기술조사문헌

HERRE, Jurgen, et al. MPEG-H audio - the new standard for universal spatial/3D audio coding. Journal of the Audio Engineering Society, 2015.01.05. Vol.62,No.12. pp.821-830.

Draft International Standard ISO/IEC DIS
23008-3. ISO/IEC JTC 1/SC 29/WG 11.
2014.07.25.*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

소니 주식회사

일본국 도쿄도 미나토쿠 코난 1-7-1

(72) 발명자

야마모토 유키

일본 1080075 도쿄도 미나토쿠 코난 1-7-1 소니
주식회사 내

치넨 도루

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 16 항

심사관 : 이남숙

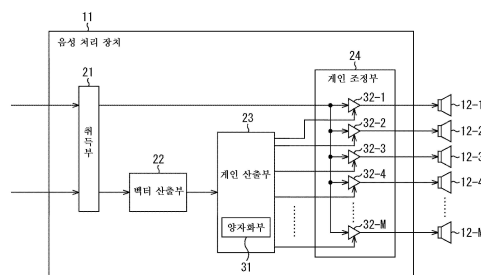
(54) 발명의 명칭 음성 처리 장치 및 방법, 그리고 기록 매체

(57) 요약

본 기술은, 보다 고품질의 음성을 얻을 수 있도록 하는 음성 처리 장치 및 방법, 그리고 프로그램에 관한 것이다. 취득부는, 오브젝트의 오디오 신호와 메타데이터를 취득한다. 백터 산출부는, 오브젝트의 메타데이터에 포함되어 있는, 음상의 범위를 나타내는 수평 방향 각도 및 수직 방향 각도에 기초하여, 음상의 범위를 나타

(뒷면에 계속)

대표도



내는 영역 내의 위치를 나타내는 spread 벡터를 산출한다. 게인 산출부는, spread 벡터에 기초하여, VBAP에 의해 각 스피커에 대하여 오디오 신호의 VBAP 게인을 산출한다. 본 기술은 음성 처리 장치에 적용할 수 있다.

명세서

청구범위

청구항 1

오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하는 취득부와,

상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도의 비에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하는 벡터 산출부와,

상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 계인을 산출하는 계인 산출부

를 구비하는 음성 처리 장치.

청구항 2

삭제

청구항 3

제1항에 있어서, 상기 벡터 산출부는, 미리 정해진 개수의 상기 spread 벡터를 산출하는 음성 처리 장치.

청구항 4

제1항에 있어서, 상기 벡터 산출부는, 가변인 임의의 개수의 상기 spread 벡터를 산출하는 음성 처리 장치.

청구항 5

제1항에 있어서, 상기 음상 정보는, 상기 영역의 중심 위치를 나타내는 벡터인 음성 처리 장치.

청구항 6

제1항에 있어서, 상기 음상 정보는, 상기 영역의 중심으로부터의 음상의 범위 정도를 나타내는 2차원 이상의 벡터인 음성 처리 장치.

청구항 7

제1항에 있어서, 상기 음상 정보는, 상기 위치 정보에 의해 나타나는 위치로부터 본 상기 영역의 중심 위치의 상대적인 위치를 나타내는 벡터인 음성 처리 장치.

청구항 8

제1항에 있어서, 상기 계인 산출부는,

각 상기 음성 출력부에 대해서, 상기 spread 벡터마다 상기 계인을 산출하고,

상기 음성 출력부마다, 각 상기 spread 벡터에 대하여 산출한 상기 계인의 가산값을 산출하고,

상기 음성 출력부마다, 상기 가산값을 2차 이상의 계인으로 양자화하고,

상기 양자화된 상기 가산값에 기초하여, 상기 음성 출력부마다 최종적인 상기 계인을 산출하는 음성 처리 장치.

청구항 9

제8항에 있어서, 상기 계인 산출부는, 3개의 상기 음성 출력부에 의해 둘러싸이는 영역인 메쉬이며, 상기 계인의 산출에 사용하는 메쉬의 수를 선택하고, 상기 메쉬의 수의 선택 결과와 상기 spread 벡터에 기초하여, 상기 spread 벡터마다 상기 계인을 산출하는

음성 처리 장치.

청구항 10

제9항에 있어서, 상기 계인 산출부는, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화 시에 있어서의 상기 가산값의 양자화수를 선택하고, 그 선택 결과에 따라서 상기 최종적인 상기 계인을 산출하는

음성 처리 장치.

청구항 11

제10항에 있어서, 상기 계인 산출부는, 상기 오디오 오브젝트의 수에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택하는

음성 처리 장치.

청구항 12

제10항에 있어서, 상기 계인 산출부는, 상기 오디오 오브젝트의 중요도에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택하는

음성 처리 장치.

청구항 13

제12항에 있어서, 상기 계인 산출부는, 상기 중요도가 높은 상기 오디오 오브젝트에 가까운 위치에 있는 상기 오디오 오브젝트일수록, 상기 계인의 산출에 사용하는 상기 메쉬의 수가 많아지도록, 상기 계인의 산출에 사용하는 상기 메쉬의 수를 선택하는

음성 처리 장치.

청구항 14

제10항에 있어서, 상기 계인 산출부는, 상기 오디오 오브젝트의 오디오 신호의 음압에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택하는

음성 처리 장치.

청구항 15

제9항에 있어서, 상기 계인 산출부는, 상기 메쉬의 수의 선택 결과에 따라, 복수의 상기 음성 출력부 중, 서로 다른 높이에 위치하는 상기 음성 출력부를 포함하는 3 이상의 상기 음성 출력부를 선택하고, 선택한 상기 음성 출력부로 형성되는 1개 또는 복수의 상기 메쉬에 기초하여 상기 계인을 산출하는

음성 처리 장치.

청구항 16

오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하고,

상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도의 비

에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하고,

상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 계인을 산출하는

시스템을 포함하는 음성 처리 방법.

청구항 17

오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하고,

상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도의 비에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하고,

상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 계인을 산출하는

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

발명의 설명

기술 분야

[0001] 본 기술은 음성 처리 장치 및 방법, 그리고 프로그램에 관한 것으로서, 특히, 보다 고품질의 음성을 얻을 수 있도록 한 음성 처리 장치 및 방법, 그리고 프로그램에 관한 것이다.

배경 기술

[0002] 종래, 복수의 스피커를 사용하여 음상의 정위를 제어하는 기술로서, VBAP(Vector Base Amplitude Panning)가 알려져 있다(예를 들어, 비특허문헌 1 참조).

[0003] VBAP에서는, 3개의 스피커로부터 소리를 출력함으로써, 그들 3개의 스피커로 구성되는 삼각형의 내측의 임의의 1점에 음상을 정위시킬 수 있다.

[0004] 그러나, 실세계에서는, 음상은 1점에 정위되는 것이 아니고, 어느 정도의 범위를 갖는 공간에 정위된다고 생각된다. 예를 들어, 인간의 목소리는 성대로부터 발해지지만, 그 진동은 얼굴이나 몸 등에 전반하여, 그 결과, 인간의 몸 전체라고 하는 부분 공간으로부터 음성이 발해진다고 생각된다.

[0005] 이러한 부분 공간에 소리를 정위시키는 기술, 즉 음상을 확장하는 기술로서 MDAP(Multiple Direction Amplitude Panning)가 일반적으로 알려져 있다(예를 들어, 비특허문헌 2 참조). 또한, 이 MDAP는 MPEG(Moving Picture Experts Group)-H 3D Audio 규격의 렌더링 처리부에서도 사용되고 있다(예를 들어, 비특허문헌 3 참조).

선행기술문헌

비특허문헌

[0006] (비특허문헌 0001) Ville Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", Journal of AES, vol.45, no.6, pp.456-466, 1997

(비특허문헌 0002) Ville Pulkki, "Uniform Spreading of Amplitude Panned Virtual Sources", Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, Oct. 17-20, 1999

(비특허문헌 0003) ISO/IEC JTC1/SC29/WG11 N14747, August 2014, Sapporo, Japan, "Text of ISO/IEC 23008-3/DIS, 3D Audio"

발명의 내용

해결하려는 과제

- [0007] 그러나, 상술한 기술로는, 충분히 고품질의 음성을 얻을 수 없었다.
- [0008] 예를 들어 MPEG-H 3D Audio 규격에서는, 오디오 오브젝트의 메타데이터에 spread라고 불리는 음상의 범위 정도를 나타내는 정보가 포함되어 있고, 이 spread에 기초하여 음상을 확장하는 처리가 행해진다. 그런데, 음상을 확장하는 처리에서는, 오디오 오브젝트의 위치를 중심으로 하여 음상의 범위가 상하 좌우 대칭이라고 하는 제약이 있다. 그 때문에, 오디오 오브젝트로부터의 음성의 지향성(방사 방향)을 고려한 처리를 행할 수 없어, 충분히 고품질의 음성을 얻을 수 없었다.
- [0009] 본 기술은, 이러한 상황을 감안하여 이루어진 것이며, 보다 고품질의 음성을 얻을 수 있도록 하는 것이다.

과제의 해결 수단

- [0010] 본 기술의 일 측면의 음성 처리 장치는, 오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하는 취득부와, 상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하는 벡터 산출부와, 상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 계인을 산출하는 계인 산출부를 구비한다.
- [0011] 상기 벡터 산출부에는, 상기 수평 방향 각도와 상기 수직 방향 각도의 비에 기초하여, 상기 spread 벡터를 산출시킬 수 있다.
- [0012] 상기 벡터 산출부에는, 미리 정해진 개수의 상기 spread 벡터를 산출시킬 수 있다.
- [0013] 상기 벡터 산출부에는, 가변인 임의의 개수의 상기 spread 벡터를 산출시킬 수 있다.
- [0014] 상기 음상 정보를, 상기 영역의 중심 위치를 나타내는 벡터로 할 수 있다.
- [0015] 상기 음상 정보를, 상기 영역의 중심으로부터의 음상의 범위 정도를 나타내는 2차원 이상의 벡터로 할 수 있다.
- [0016] 상기 음상 정보를, 상기 위치 정보에 의해 나타나는 위치로부터 본 상기 영역의 중심 위치의 상대적인 위치를 나타내는 벡터로 할 수 있다.
- [0017] 상기 계인 산출부에는, 각 상기 음성 출력부에 대해서, 상기 spread 벡터마다 상기 계인을 산출시키고, 상기 음성 출력부마다, 각 상기 spread 벡터에 대하여 산출한 상기 계인의 가산값을 산출시키고, 상기 음성 출력부마다, 상기 가산값을 2차 이상의 계인으로 양자화시키고, 상기 양자화된 상기 가산값에 기초하여, 상기 음성 출력부마다 최종적인 상기 계인을 산출시킬 수 있다.
- [0018] 상기 계인 산출부에는, 3개의 상기 음성 출력부에 의해 둘러싸이는 영역인 메쉬이며, 상기 계인의 산출에 사용하는 메쉬의 수를 선택시켜, 상기 메쉬의 수의 선택 결과와 상기 spread 벡터에 기초하여, 상기 spread 벡터마다 상기 계인을 산출시킬 수 있다.
- [0019] 상기 계인 산출부에는, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화 시에 있어서의 상기 가산값의 양자화수를 선택시키고, 그 선택 결과에 따라서 상기 최종적인 상기 계인을 산출시킬 수 있다.
- [0020] 상기 계인 산출부에는, 상기 오디오 오브젝트의 수에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택시킬 수 있다.
- [0021] 상기 계인 산출부에는, 상기 오디오 오브젝트의 중요도에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택시킬 수 있다.
- [0022] 상기 계인 산출부에는, 상기 중요도가 높은 상기 오디오 오브젝트에 가까운 위치에 있는 상기 오디오 오브젝트 일수록, 상기 계인의 산출에 사용하는 상기 메쉬의 수가 많아지도록, 상기 계인의 산출에 사용하는 상기 메쉬의 수를 선택시킬 수 있다.
- [0023] 상기 계인 산출부에는, 상기 오디오 오브젝트의 오디오 신호의 음압에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택시킬 수 있다.

[0024] 상기 개인 산출부에는, 상기 메쉬의 수의 선택 결과에 따라, 복수의 상기 음성 출력부 중, 서로 다른 높이에 위치하는 상기 음성 출력부를 포함하는 3개 이상의 상기 음성 출력부를 선택시키고, 선택한 상기 음성 출력부로 형성되는 1개 또는 복수의 상기 메쉬에 기초하여 상기 개인을 산출시킬 수 있다.

[0025] 본 기술의 일 측면의 음성 처리 방법 또는 프로그램은, 오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하고, 상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하고, 상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 개인을 산출하는 스텝을 포함한다.

[0026] 본 기술의 일 측면에 있어서는, 오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터가 취득되고, 상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터가 산출되고, 상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 개인이 산출된다.

발명의 효과

[0027] 본 기술의 일 측면에 의하면, 보다 고품질의 음성을 얻을 수 있다.

[0028] 또한, 여기에 기재된 효과는 반드시 한정되는 것은 아니며, 본 개시 중에 기재된 어느 효과여도 된다.

도면의 간단한 설명

[0029] 도 1은 VBAP에 대하여 설명하는 도면이다.

도 2는 음상의 위치에 대하여 설명하는 도면이다.

도 3은 spread 벡터에 대하여 설명하는 도면이다.

도 4는 spread 중심 벡터 방식에 대하여 설명하는 도면이다.

도 5는 spread 방사 벡터 방식에 대하여 설명하는 도면이다.

도 6은 음성 처리 장치의 구성예를 도시하는 도면이다.

도 7은 재생 처리를 설명하는 흐름도이다.

도 8은 spread 벡터 산출 처리를 설명하는 흐름도이다.

도 9는 spread 3차원 벡터에 기초하는 spread 벡터 산출 처리를 설명하는 흐름도이다.

도 10은 spread 중심 벡터에 기초하는 spread 벡터 산출 처리를 설명하는 흐름도이다.

도 11은 spread 단부 벡터에 기초하는 spread 벡터 산출 처리를 설명하는 흐름도이다.

도 12는 spread 방사 벡터에 기초하는 spread 벡터 산출 처리를 설명하는 흐름도이다.

도 13은 spread 벡터 위치 정보에 기초하는 spread 벡터 산출 처리를 설명하는 흐름도이다.

도 14는 메쉬수의 전환에 대하여 설명하는 도면이다.

도 15는 메쉬수의 전환에 대하여 설명하는 도면이다.

도 16은 메쉬의 형성에 대하여 설명하는 도면이다.

도 17은 음성 처리 장치의 구성예를 도시하는 도면이다.

도 18은 재생 처리를 설명하는 흐름도이다.

도 19는 음성 처리 장치의 구성예를 도시하는 도면이다.

도 20은 재생 처리를 설명하는 흐름도이다.

도 21은 VBAP 게인 산출 처리를 설명하는 흐름도이다.

도 22는 컴퓨터의 구성예를 도시하는 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0030] 이하, 도면을 참조하여, 본 기술을 적용한 실시 형태에 대하여 설명한다.
- [0031] <제1 실시 형태>
- [0032] <VBAP과 음상을 확장하는 처리에 대해서>
- [0033] 본 기술은, 오디오 오브젝트의 오디오 신호와, 그 오디오 오브젝트의 위치 정보 등의 메타데이터를 취득하여 렌더링을 행하는 경우에, 보다 고품질의 음성을 얻을 수 있도록 하는 것이다. 또한, 이하에서는, 오디오 오브젝트를, 간단히 오브젝트라고도 칭하기로 한다.
- [0034] 이하에서는, 먼저 VBAP, 및 MPEG-H 3D Audio 규격에 있어서의 음상을 확장하는 처리에 대하여 설명한다.
- [0035] 예를 들어, 도 1에 도시한 바와 같이, 음성이 있는 동화상이나 악곡 등의 콘텐츠를 시청하는 유저(U11)가, 3개의 스피커(SP1) 내지 스피커(SP3)로부터 출력되는 3 채널의 음성을 콘텐츠의 음성으로서 듣고 있다고 하자.
- [0036] 이러한 경우에, 각 채널의 음성을 출력하는 3개의 스피커(SP1) 내지 스피커(SP3)의 위치를 나타내는 정보를 사용하여, 위치 p에 음상을 정위시키는 것을 생각한다.
- [0037] 예를 들어, 유저(U11)의 헤드부 위치를 원점 0로 하는 3차원 좌표계에 있어서, 위치 p를, 원점 0를 시점으로 하는 3차원의 벡터(이하, 벡터 p라고도 칭한다)에 의해 나타내기로 한다. 또한, 원점 0를 시점으로 하여, 각 스피커(SP1) 내지 스피커(SP3)의 위치의 방향을 향하는 3차원의 벡터를 벡터 l_1 내지 벡터 l_3 이라 하면, 벡터 p는 벡터 l_1 내지 벡터 l_3 의 선형합에 의해 나타낼 수 있다.
- [0038] 즉, $p = g_1 l_1 + g_2 l_2 + g_3 l_3$ 으로 할 수 있다.
- [0039] 여기서, 벡터 l_1 내지 벡터 l_3 에 승산되어 있는 계수 g_1 내지 계수 g_3 을 산출하고, 이들 계수 g_1 내지 계수 g_3 을, 스피커(SP1) 내지 스피커(SP3) 각각으로부터 출력하는 음성의 게인으로 하면, 위치 p에 음상을 정위시킬 수 있다.
- [0040] 이와 같이 하여, 3개의 스피커(SP1) 내지 스피커(SP3)의 위치 정보를 사용하여 계수 g_1 내지 계수 g_3 을 구하고, 음상의 정위 위치를 제어하는 방법은, 3차원 VBAP라고 부르고 있다. 특히, 이하에서는, 계수 g_1 내지 계수 g_3 과 같이 스피커마다 구해진 계인을, VBAP 게인이라고 칭하기로 한다.
- [0041] 도 1의 예에서는, 스피커(SP1), 스피커(SP2), 및 스피커(SP3)의 위치를 포함하는 구면 상의 삼각형의 영역 TR11 내의 임의의 위치에 음상을 정위시킬 수 있다. 여기서, 영역 TR11은, 원점 0를 중심으로 하여, 스피커(SP1) 내지 스피커(SP3)의 각 위치를 통과하는 구의 표면 상의 영역이며, 스피커(SP1) 내지 스피커(SP3)에 의해 둘러싸이는 3각형의 영역이다.
- [0042] 이러한 3차원 VBAP를 사용하면, 공간 상의 임의의 위치에 음상을 정위시킬 수 있게 된다. 또한, VBAP에 대해서는, 예를 들어 「Ville Pulkki, "Virtual Sound Source Positioning Using Vector Base Amplitude Panning", Journal of AES, vol.45, no.6, pp.456-466, 1997」 등에 상세하게 기재되어 있다.
- [0043] 이어서, MPEG-H 3D Audio 규격에서의 음상을 확장하는 처리에 대하여 설명한다.
- [0044] MPEG-H 3D Audio 규격에서는, 부호화 장치로부터는, 각 오브젝트의 오디오 신호를 부호화하여 얻어진 부호화 오디오 데이터와, 각 오브젝트의 메타데이터를 부호화하여 얻어진 부호화 메타데이터를 다중화하여 얻어진 비트스트림이 출력된다.
- [0045] 예를 들어, 메타데이터에는, 오브젝트의 공간 상의 위치를 나타내는 위치 정보, 오브젝트의 중요도를 나타내는 중요도 정보, 및 오브젝트의 음상의 범위 정도를 나타내는 정보인 spread가 포함되어 있다.
- [0046] 여기서, 음상의 범위 정도를 나타내는 spread는, 0° 부터 180° 까지의 임의의 각도로 되고, 부호화 장치에서는, 각 오브젝트에 대해서, 오디오 신호의 프레임마다 상이한 값의 spread를 지정하는 것이 가능하다.

- [0047] 또한, 오브젝트의 위치는 수평 방향 각도 azimuth, 수직 방향 각도 elevation, 및 거리 radius에 의해 표현된다. 즉, 오브젝트의 위치 정보는 수평 방향 각도 azimuth, 수직 방향 각도 elevation, 및 거리 radius의 각 값을 포함한다.
- [0048] 예를 들어, 도 2에 도시한 바와 같이, 도시하지 않은 스피커로부터 출력되는 각 오브젝트의 음성을 듣고 있는 시청자의 위치를 원점 0으로 하고, 도면 중, 우상측 방향, 좌상측 방향, 및 상측 방향을 서로 수직한 x축, y축, 및 z축의 방향으로 하는 3차원 좌표계를 생각한다. 이때, 하나의 오브젝트 위치를 위치 OBJ11이라 하면, 3차원 좌표계에 있어서의 위치 OBJ11에 음상을 정위시키면 된다.
- [0049] 또한, 위치 OBJ11과 원점 0을 연결하는 직선을 직선 L이라 하면, xy 평면 상에 있어서 직선 L과 x축이 이루는 도면 중, 수평 방향의 각도 Θ (방위각)가 위치 OBJ11에 있는 오브젝트의 수평 방향 위치를 나타내는 수평 방향 각도 azimuth로 되고, 수평 방향 각도 azimuth는 $-180^\circ \leq \text{azimuth} \leq 180^\circ$ 을 충족하는 임의의 값으로 된다.
- [0050] 예를 들어 x축 방향의 정의 방향이 azimuth= 0° 로 되고, x축 방향의 부의 방향이 azimuth= $+180^\circ = -180^\circ$ 로 된다. 또한, 원점 0를 중심으로 반시계 방향이 azimuth의 +방향으로 되고, 원점 0를 중심으로 시계 방향이 azimuth의 -방향으로 된다.
- [0051] 또한, 직선 L과 xy 평면이 이루는 각도, 즉 도면 중, 수직 방향의 각도 γ (양각)가 위치 OBJ11에 있는 오브젝트의 수직 방향의 위치를 나타내는 수직 방향 각도 elevation이 되고, 수직 방향 각도 elevation은 $-90^\circ \leq \text{elevation} \leq 90^\circ$ 을 충족하는 임의의 값으로 된다. 예를 들어 xy 평면의 위치가 elevation= 0° 로 되고, 도면 중, 상측 방향이 수직 방향 각도 elevation의 +방향으로 되고, 도면 중, 하측 방향이 수직 방향 각도 elevation의 -방향으로 된다.
- [0052] 또한, 직선 L의 길이, 즉 원점 0부터 위치 OBJ11까지의 거리가 시청자까지의 거리 radius로 되고, 거리 radius는 0 이상의 값으로 된다. 즉, 거리 radius는, $0 \leq \text{radius} < \infty$ 을 충족하는 값으로 된다. 이하에서는, 거리 radius를 반경 방향의 거리라고도 칭한다.
- [0053] 또한, VBAP에서는 모든 스피커나 오브젝트로부터 시청자까지의 거리 radius가 동일해서, 거리 radius를 1로 정규화하여 계산을 행하는 것이 일반적인 방식이다.
- [0054] 이렇게 메타데이터에 포함되는 오브젝트의 위치 정보는, 수평 방향 각도 azimuth, 수직 방향 각도 elevation, 및 거리 radius의 각 값을 포함한다.
- [0055] 이하에서는, 수평 방향 각도 azimuth, 수직 방향 각도 elevation, 및 거리 radius를, 간단히 azimuth, elevation, 및 radius라고도 칭하기로 한다.
- [0056] 또한, 부호화 오디오 데이터와 부호화 메타데이터가 포함되는 비트 스트림을 수신한 복호 장치에서는, 부호화 오디오 데이터와 부호화 메타데이터의 복호가 행해진 후, 메타데이터에 포함되어 있는 spread의 값에 따라, 음상을 확장하는 렌더링 처리가 행해진다.
- [0057] 구체적으로는, 먼저 복호 장치는, 오브젝트의 메타데이터에 포함되는 위치 정보에 의해 나타나는 공간 상의 위치를 위치 p라 한다. 이 위치 p는, 상술한 도 1의 위치 p에 대응한다.
- [0058] 계속해서, 복호 장치는, 예를 들어 도 3에 도시한 바와 같이 위치 p=중심 위치 p0으로 하고, 중심 위치 p0을 중심으로 하여 단위 구면 상에서 상하 좌우 대칭이 되도록, 18개의 spread 벡터 p1 내지 spread 벡터 p18을 배치한다. 또한, 도 3에 있어서, 도 1에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 적절히 생략한다.
- [0059] 도 3에서는, 원점 0를 중심으로 하는 반경(1)의 단위 구의 구면 상에 5개의 스피커(SP1) 내지 스피커(SP5)가 배치되어 있고, 위치 정보에 의해 나타나는 위치 p가, 중심 위치 p0으로 되어 있다. 이하에서는, 위치 p를 특히 오브젝트 위치 p라고도 칭하고, 원점 0를 시점으로 하고, 오브젝트 위치 p를 중점으로 하는 벡터를 벡터 p라고도 칭하기로 한다. 또한, 원점 0를 시점으로 하고, 중심 위치 p0을 중점으로 하는 벡터를 벡터 p0이라고도 칭하기로 한다.
- [0060] 도 3에서는, 원점 0를 시점으로 하는, 점선으로 그려진 화살표가 spread 벡터를 나타내고 있다. 단, 실제로는 spread 벡터는 18개 있지만, 도 3에서는, 도면을 보기 쉽게 하기 위해서 spread 벡터가 8개만 그려져 있다.
- [0061] 여기서, spread 벡터 p1 내지 spread 벡터 p18 각각은, 그 중점 위치가 중심 위치 p0을 중심으로 하는 단위 구면 상의 원의 영역 R11 내에 위치하는 벡터로 되어 있다. 특히, 영역 R11로 표현되는 원의 원주 상에 중점 위

치가 있는 spread 벡터와, 벡터 p0과의 이루는 각도가 spread에 의해 나타나는 각도가 된다.

- [0062] 따라서, 각 spread 벡터의 중점 위치는, spread의 값이 커질수록 중심 위치 p0으로부터 이격된 위치에 배치되게 된다. 즉, 영역 R11은 커진다.
- [0063] 이 영역 R11은, 오브젝트의 위치로부터의 음상의 범위를 표현하고 있다. 바꾸어 말하면, 영역 R11은, 오브젝트의 음상이 확장되는 범위를 나타내는 영역으로 되어 있다. 더욱 상세히 설명하면, 오브젝트의 음성은, 오브젝트 전체로부터 발해진다고 생각되므로, 영역 R11은 오브젝트의 형상을 나타내고 있다고도 할 수 있다. 이하에서는, 영역 R11과 같이, 오브젝트의 음상이 확장되는 범위를 나타내는 영역을, 음상의 범위를 나타내는 영역이라고도 칭하기로 한다.
- [0064] 또한, spread의 값이 0일 경우에는, 18개의 spread 벡터 p1 내지 spread 벡터 p18의 각각의 중점 위치는, 중심 위치 p0과 동등해진다.
- [0065] 또한, 이하, spread 벡터 p1 내지 spread 벡터 p18의 각각의 중점 위치를, 특히 위치 p1 내지 위치 p18이라고도 칭하기로 한다.
- [0066] 이와 같이 하여, 단위 구면 상에 있어서 상하 좌우 대칭인 spread 벡터가 정해지면, 복호 장치는, 벡터 p와 각 spread 벡터에 대해서, 즉 위치 p와 위치 p1 내지 위치 p18 각각에 대해서, VBAP에 의해 각 채널의 스피커마다 VBAP 계인을 산출한다. 이때, 위치 p나 위치 p1 등, 그들 각 위치에 음상이 정위하도록 스피커마다의 VBAP 계인이 산출된다.
- [0067] 그리고, 복호 장치는 각 위치에 대하여 산출한 VBAP 계인을 스피커마다 가산한다. 예를 들어 도 3의 예에서는, 스피커(SP1)에 대하여 산출된 위치 p 및 위치 p1 내지 위치 p18의 각각의 VBAP 계인이 가산된다.
- [0068] 또한, 복호 장치는, 스피커마다 구해진 가산 처리 후의 VBAP 계인을 정규화한다. 즉, 전체 스피커의 VBAP 계인의 2승합이 1로 되도록 정규화가 행해진다.
- [0069] 그리고, 복호 장치는, 정규화에 의해 얻어진 각 스피커의 VBAP 계인을, 오브젝트의 오디오 신호에 승산하고, 그들 스피커마다의 오디오 신호로 하고, 스피커마다 얻어진 오디오 신호를 스피커에 공급하여 음성을 출력시킨다.
- [0070] 이에 의해, 예를 들어 도 3의 예에서는, 영역 R11 전체로부터 음성이 출력되어 있도록 음상이 정위된다. 즉, 음상이 영역 R11 전체에 확장되게 된다.
- [0071] 도 3에서는, 음상을 확장하는 처리를 행하지 않는 경우에는, 오브젝트의 음상은 위치 p에 정위하므로, 이 경우에는, 실질적으로 스피커(SP2)와 스피커(SP3)로부터 음성이 출력된다. 이에 반해, 음상을 확장하는 처리가 행해진 경우에는, 음상이 영역 R11 전체에 확장되므로, 음성 재생 시에는, 스피커(SP1) 내지 스피커(SP4)로부터 음성이 출력된다.
- [0072] 그런데, 이상과 같은 음상을 확장하는 처리를 행하는 경우에는, 음상을 확장하는 처리를 행하지 않는 경우에 비하여, 렌더링 시의 처리량이 많아진다. 그렇게 하면, 복호 장치로 취급할 수 있는 오브젝트의 수가 줄어들거나, 하드 규모가 작은 렌더러가 탑재된 복호 장치로는 렌더링을 행할 수 없게 되거나 하는 경우가 발생해버린다.
- [0073] 그래서, 렌더링 시에 음상을 확장하는 처리를 행하는 경우에는, 더 적은 처리량으로 렌더링을 행할 수 있도록 하는 것이 바람직하다.
- [0074] 또한, 상술한 18개의 spread 벡터는, 중심 위치 p0=위치 p를 중심으로 하여, 단위 구면 상에서 상하 좌우 대칭이라고 하는 제약이 있기 때문에, 오브젝트의 소리의 지향성(방사 방향)이나 오브젝트의 형상을 고려한 처리를 할 수 없다. 그 때문에, 충분히 고품질의 음성을 얻을 수 없었다.
- [0075] 또한, MPEG-H 3D Audio 규격에서는, 렌더링 시에 음상을 확장하는 처리로서, 처리가 1가지밖에 규정되어 있지 않기 때문에, 렌더러의 하드 규모가 작은 경우에는, 음상을 확장하는 처리를 행할 수 없었다. 즉, 음성의 재생을 행할 수 없었다.
- [0076] 또한, MPEG-H 3D Audio 규격에서는, 렌더러의 하드 규모에서 허용되는 처리량내에서, 최대의 품질의 음성을 얻을 수 있도록, 처리를 전환하여 렌더링을 행할 수 없었다.
- [0077] 이상과 같은 상황을 감안하여, 본 기술에서는, 렌더링 시의 처리량을 삭감할 수 있도록 하였다. 또한, 본 기술에서는, 오브젝트의 지향성이나 형상을 표현함으로써 충분히 고품질의 음성을 얻을 수 있도록 하였다. 또한,

본 기술에서는, 렌더링의 하드 규모 등에 따라서 렌더링 시의 처리로서 적절한 처리를 선택하고, 허용되는 처리량의 범위에서 가장 높은 품질의 음성을 얻을 수 있도록 하였다.

- [0078] 이하, 본 기술의 개요에 대하여 설명한다.
- [0079] <처리량의 삭감에 대해서>
- [0080] 먼저, 렌더링 시의 처리량의 삭감에 대하여 설명한다.
- [0081] 음상을 확장하지 않는 통상의 VBAP 처리(렌더링 처리)에서는, 구체적으로 이하에 나타내는 처리 A1 내지 처리 A3이 행해진다.
- [0082] (처리 A1)
- [0083] 3개의 스피커에 대해서, 오디오 신호에 승산하는 VBAP 계인을 산출한다
- [0084] (처리 A2)
- [0085] 3개의 스피커의 VBAP 계인의 2승합이 1로 되도록 정규화를 행한다
- [0086] (처리 A3)
- [0087] 오브젝트의 오디오 신호에 VBAP 계인을 승산한다
- [0088] 여기서, 처리 A3에서는, 3개의 스피커마다, 오디오 신호에 대한 VBAP 계인의 승산 처리가 행해지기 때문에, 이러한 승산 처리는 최대로 3회 행해지게 된다.
- [0089] 이에 반해, 음상을 확장하는 처리를 행하는 경우의 VBAP 처리(렌더링 처리)에서는, 구체적으로 이하에 나타내는 처리 B1 내지 처리 B5가 행해진다.
- [0090] (처리 B1)
- [0091] 벡터 p 에 대해서, 3개의 각 스피커의 오디오 신호에 승산하는 VBAP 계인을 산출한다
- [0092] (처리 B2)
- [0093] 18개의 각 spread 벡터에 대해서, 3개의 각 스피커의 오디오 신호에 승산하는 VBAP 계인을 산출한다
- [0094] (처리 B3)
- [0095] 스피커마다, 각 벡터에 대하여 구한 VBAP 계인을 가산한다
- [0096] (처리 B4)
- [0097] 전체 스피커의 VBAP 계인의 2승합이 1로 되도록 정규화를 행한다
- [0098] (처리 B5)
- [0099] 오브젝트의 오디오 신호에 VBAP 계인을 승산한다
- [0100] 음상을 확장하는 처리를 행한 경우, 음성을 출력하는 스피커의 수는 3 이상이 되므로, 처리 B5에서는 3회 이상 승산 처리가 행해지게 된다.
- [0101] 따라서, 음상을 확장하는 처리를 행하는 경우와 행하지 않는 경우를 비교하면, 음상을 확장하는 처리를 행하는 경우에는, 특히 처리 B2와 처리 B3의 분만큼 처리량이 많아지고, 또한 처리 B5에서도 처리 A3보다도 처리량이 많아진다.
- [0102] 그래서, 본 기술에서는, 스피커마다 구해진, 각 벡터의 VBAP 계인의 합을 양자화함으로써, 상술한 처리 B5의 처리량을 삭감할 수 있도록 하였다.
- [0103] 구체적으로는, 본 기술에서는, 이하와 같은 처리가 행해진다. 또한, 이하에서는, 스피커마다 구해지는, 벡터 p 나 spread 벡터 등의 각 벡터마다 구한 VBAP 계인의 합(가산값)을 VBAP 계인 가산값이라고도 칭하기로 한다.
- [0104] 먼저, 처리 B1 내지 처리 B3이 행해지고, 스피커마다 VBAP 계인 가산값이 얻어지면, 그 VBAP 계인 가산값이 2치화된다. 2치화에서는, 예를 들어 각 스피커의 VBAP 계인 가산값이 0 또는 1 중 어느 값으로 된다.
- [0105] VBAP 계인 가산값을 2치화하는 방법은, 예를 들어 반올림, 실링(절상), 플로어링(잘라 버림), 역치 처리 등, 어

편 방법이어도 된다.

- [0106] 이와 같이 하여 VBAP 게인 가산값이 2치화되면, 그 후, 2치화된 VBAP 게인 가산값에 기초하여, 상술한 처리 B4가 행해진다. 그렇게 하면, 그 결과, 각 스피커의 최종적인 VBAP 게인은, 0을 제외하면 1가지가 된다. 즉, VBAP 게인 가산값을 2치화하면, 각 스피커의 최종적인 VBAP 게인의 값은 0이거나, 또는 소정값 중 어느 것이 된다.
- [0107] 예를 들어 2치화의 결과, 3개의 스피커의 VBAP 게인 가산값이 1이 되고, 다른 스피커의 VBAP 게인 가산값이 0이 되었다고 하면, 그들 3개의 스피커의 최종적인 VBAP 게인의 값은 $1/3^{(1/2)}$ 이 된다.
- [0108] 이와 같이 하여 각 스피커의 최종적인 VBAP 게인이 얻어지면, 그 후에는 상술한 처리 B5 대신에, 처리 B5'로서, 각 스피커의 오디오 신호에, 최종적인 VBAP 게인을 승산하는 처리가 행해진다.
- [0109] 상술한 바와 같이 2치화를 행하면, 각 스피커의 최종적인 VBAP 게인의 값은 0이거나 소정값 중 어느 것이 되므로, 처리 B5'에서는 1번의 승산 처리를 행하면 되게 되어, 처리량을 삭감할 수 있다. 즉, 처리 B5에서는 3회 이상의 승산 처리를 해야만 했던 것을, 처리 B5'에서는 1회의 승산 처리를 행하기만 해도 되게 된다.
- [0110] 또한, 여기에서는 VBAP 게인 가산값을 2치화하는 경우를 예로 들어 설명했지만, VBAP 게인 가산값이 3값 이상의 값으로 양자화되게 해도 된다.
- [0111] 예를 들어 VBAP 게인 가산값이 3개의 값 중 어느 것으로 될 경우, 상술한 처리 B1 내지 처리 B3이 행해지고, 스피커마다 VBAP 게인 가산값이 얻어지면, 그 VBAP 게인 가산값이 양자화되어, 0, 0.5, 또는 1 중 어느 값으로 된다. 그리고, 그 후에는 처리 B4와 처리 B5'가 행해진다. 이 경우, 처리 B5'에 있어서의 승산 처리의 횟수는 최대 2회가 된다.
- [0112] 이와 같이, VBAP 게인 가산값을 x치화하면, 즉 2 이상의 x개의 게인 중 어느 것이 되도록 양자화하면, 처리 B5'에 있어서의 승산 처리의 횟수는 최대 (x-1)회가 된다.
- [0113] 또한, 이상에 있어서는, 음상을 확장하는 처리를 행하는 경우에, VBAP 게인 가산값을 양자화하여 처리량을 삭감하는 예에 대하여 설명했지만, 음상을 확장하는 처리를 행하지 않는 경우에 있어서는, 동일하게 하여 VBAP 게인을 양자화함으로써, 처리량을 삭감할 수 있다. 즉, 벡터 p에 대하여 구한 각 스피커의 VBAP 게인을 양자화하면, 정규화 후의 VBAP 게인의 오디오 신호에의 승산 처리의 횟수를 삭감할 수 있다.
- [0114] <오브젝트의 형상 및 소리의 지향성을 표현하는 처리에 대해서>
- [0115] 이어서, 본 기술에 의해, 오브젝트의 형상과, 오브젝트의 소리의 지향성을 표현하는 처리에 대하여 설명한다.
- [0116] 이하에서는, spread 3차원 벡터 방식, spread 중심 벡터 방식, spread 단부 벡터 방식, spread 방사 벡터 방식, 및 임의 spread 벡터 방식의 5가지의 방식에 대하여 설명한다.
- [0117] (spread 3차원 벡터 방식)
- [0118] 먼저, spread 3차원 벡터 방식에 대하여 설명한다.
- [0119] spread 3차원 벡터 방식에서는, 비트 스트림 내에 3차원 벡터인 spread 3차원 벡터가 저장되어서 전송된다. 여기에서는, 예를 들어 오브젝트마다의 각 오디오 신호의 프레임 메타데이터에 spread 3차원 벡터가 저장된다고 하자. 이 경우, 메타데이터에는, 음상의 범위 정도를 나타내는 spread는 저장되지 않는다.
- [0120] 예를 들어 spread 3차원 벡터는, 수평 방향의 음상의 범위 정도를 나타내는 s3_azimuth, 수직 방향의 음상의 범위 정도를 나타내는 s3_elevation, 및 음상의 반경 방향의 깊이를 나타내는 s3_radius의 3가지의 요소를 포함하는 3차원 벡터로 된다.
- [0121] 즉, spread 3차원 벡터=(s3_azimuth, s3_elevation, s3_radius)이다.
- [0122] 여기에서 s3_azimuth는, 위치 p로부터의 수평 방향, 즉 상술한 수평 방향 각도 azimuth의 방향으로의 음상의 범위 각도를 나타내고 있다. 구체적으로는, s3_azimuth는 원점 0로부터 음상의 범위를 나타내는 영역의 수평 방향측의 단부를 향하는 벡터와, 벡터 p(벡터 p0)가 이루는 각도를 나타내고 있다.
- [0123] 마찬가지로 s3_elevation은, 위치 p로부터의 수직 방향, 즉 상술한 수직 방향 각도 elevation의 방향으로의 음상의 범위 각도를 나타내고 있다. 구체적으로는, s3_elevation은 원점 0로부터 음상의 범위를 나타내는 영역의 수직 방향측의 단부를 향하는 벡터와, 벡터 p(벡터 p0)가 이루는 각도를 나타내고 있다. 또한, s3_radius는,

상술한 거리 radius의 방향, 즉 단위 구면의 법선 방향의 깊이를 나타내고 있다.

[0124] 또한, 이들 s3_azimuth, s3_elevation, 및 s3_radius는 0 이상의 값으로 된다. 또한, 여기에서는 spread 3차원 벡터가, 오브젝트의 위치 정보에 의해 나타나는 위치 p에 대한 상대 위치를 나타내는 정보로 되어 있지만, spread 3차원 벡터는 절대 위치를 나타내는 정보로 되도록 해도 된다.

[0125] spread 3차원 벡터 방식에서는, 이러한 spread 3차원 벡터가 사용되어서 렌더링이 행해진다.

[0126] 구체적으로는, spread 3차원 벡터 방식에서는, spread 3차원 벡터에 기초하여, 이하의 식 (1)을 계산함으로써, spread의 값이 산출된다.

$$\text{spread} : \max(s3_azimuth, s3_elevation) \quad \cdot \cdot \cdot (1)$$

[0128] 또한, 식 (1)에 있어서 $\max(a, b)$ 는 a와 b 중 큰 값을 돌려주는 함수를 나타내고 있다. 따라서, 여기에서는 s3_azimuth와 s3_elevation 중 큰 쪽의 값이 spread의 값으로 되게 된다.

[0129] 그리고, 이와 같이 하여 얻어진 spread의 값과, 메타데이터에 포함되어 있는 위치 정보에 기초하여, MPEG-H 3D Audio 규격에 있어서의 경우와 마찬가지로 18개의 spread 벡터 p1 내지 spread 벡터 p18이 산출된다.

[0130] 따라서, 메타데이터에 포함되어 있는 위치 정보에 의해 나타나는 오브젝트의 위치 p가 중심 위치 p0으로 되어, 중심 위치 p0을 중심으로 하여 단위 구면 상에서 상하 좌우 대칭이 되도록, 18개의 spread 벡터 p1 내지 spread 벡터 p18이 구해진다.

[0131] 또한, spread 3차원 벡터 방식에서는, 원점 0를 시점으로 하고, 중심 위치 p0을 중점으로 하는 벡터 p0이 spread 벡터 p0으로 된다.

[0132] 또한, 각 spread 벡터는, 수평 방향 각도 azimuth, 수직 방향 각도 elevation, 및 거리 radius에 의해 표현된다. 이하에서는, 특히 spread 벡터 pi(단, i=0 내지 18)의 수평 방향 각도 azimuth 및 수직 방향 각도 elevation을, a(i) 및 e(i)라고 나타내기로 한다.

[0133] 이와 같이 하여 spread 벡터 p0 내지 spread 벡터 p18이 얻어지면, 그 후, s3_azimuth와 s3_elevation의 비에 기초하여, 그들 spread 벡터 p1 내지 spread 벡터 p18이 변경(보정)되어, 최종적인 spread 벡터로 된다.

[0134] 즉, s3_azimuth가 s3_elevation보다도 큰 경우, 이하의 식 (2)의 계산이 행해지고, spread 벡터 p1 내지 spread 벡터 p18의 각각의 elevation인 e(i)가 e'(i)로 변경된다.

$$e'(i) = e(0) + (e(i) - e(0)) \times s3_elevation / s3_azimuth \quad \cdot \cdot \cdot (2)$$

[0135] 또한, spread 벡터 p0에 대해서는, elevation의 보정은 행해지지 않는다.

[0137] 이에 반해, s3_azimuth가 s3_elevation 미만인 경우, 이하의 식 (3)의 계산이 행해지고, spread 벡터 p1 내지 spread 벡터 p18의 각각의 azimuth인 a(i)가 a'(i)로 변경된다.

$$a'(i) = a(0) + (a(i) - a(0)) \times s3_azimuth / s3_elevation \quad \cdot \cdot \cdot (3)$$

[0138] 또한, spread 벡터 p0에 대해서는, azimuth의 보정은 행해지지 않는다.

[0140] 이상과 같이 해서 s3_azimuth와 s3_elevation 중의 큰 쪽을 spread로 하고, spread 벡터를 구하는 처리는, 단위 구면 상에 있어서의 음상의 범위를 나타내는 영역을, 우선 s3_azimuth와 s3_elevation 중 큰 쪽의 각도에 의해 정해지는 반경의 원으로 하여, 종래와 동일한 처리로 spread 벡터를 구하는 처리이다.

[0141] 또한, 그 후, s3_azimuth와 s3_elevation의 대소 관계에 따라, 식 (2)나 식 (3)에 의해 spread 벡터를 보정하는 처리는, 단위 구면 상에 있어서의 음상의 범위를 나타내는 영역이, spread 3차원 벡터에 의해 지정된 본래의 s3_azimuth와 s3_elevation에 의해 정해지는 영역이 되도록, 음상의 범위를 나타내는 영역, 즉 spread 벡터를 보정하는 처리이다.

- [0142] 따라서, 결국에는 이들 처리는, spread 3차원 벡터, 즉 $s3_azimuth$ 와 $s3_elevation$ 에 기초하여, 단위 구면 상에 있어서의 원형 또는 타원형인 음상의 범위를 나타내는 영역에 대한 spread 벡터를 산출하는 처리가 된다.
- [0143] 이와 같이 하여 spread 벡터가 얻어지면, 그 후, spread 벡터 $p0$ 내지 spread 벡터 $p18$ 이 사용되어서 상술한 처리 B2, 처리 B3, 처리 B4, 및 처리 B5'가 행해져서, 각 스피커에 공급되는 오디오 신호가 생성된다.
- [0144] 또한, 처리 B2에서는, spread 벡터 $p0$ 내지 spread 벡터 $p18$ 의 19개의 각 spread 벡터에 대하여 스피커마다의 VBAP 게인이 산출된다. 여기서, spread 벡터 $p0$ 은 벡터 p 이기 때문에, spread 벡터 $p0$ 에 대하여 VBAP 게인을 산출하는 처리는, 처리 B1을 행하는 것이라고도 할 수 있다. 또한, 처리 B3 후, 필요에 따라 VBAP 게인 가산값의 양자화가 행해진다.
- [0145] 이렇게 spread 3차원 벡터에 의해, 음상의 범위를 나타내는 영역을 임의의 형상의 영역으로 함으로써, 오브젝트의 형상이나, 오브젝트의 소리의 지향성을 표현할 수 있게 되어, 렌더링에 의해, 보다 고품질의 음성을 얻을 수 있다.
- [0146] 또한, 여기에서는 $s3_azimuth$ 와 $s3_elevation$ 중 큰 쪽의 값이 spread의 값으로 되는 예에 대하여 설명했지만, $s3_azimuth$ 와 $s3_elevation$ 중 작은 쪽의 값이 spread의 값으로 되게 해도 된다.
- [0147] 이 경우, $s3_azimuth$ 가 $s3_elevation$ 보다도 클 때에는, 각 spread 벡터의 azimuth인 $a(i)$ 가 보정되고, $s3_azimuth$ 가 $s3_elevation$ 미만일 때에는, 각 spread 벡터의 elevation인 $e(i)$ 가 보정된다.
- [0148] 또한, 여기에서는 spread 벡터 $p0$ 내지 spread 벡터 $p18$, 즉 미리 정해진 19개의 spread 벡터를 구하고, 그들 spread 벡터에 대하여 VBAP 게인을 산출하는 예에 대하여 설명했지만, 산출되는 spread 벡터의 개수를 가변으로 하게 해도 된다.
- [0149] 그러한 경우, 예를 들어 $s3_azimuth$ 와 $s3_elevation$ 의 비에 따라, 생성되는 spread 벡터의 개수가 결정되도록 할 수 있다. 이러한 처리에 의하면, 예를 들어 오브젝트가 가로로 길고, 오브젝트의 소리의 수직 방향으로의 확장이 적은 경우에, 수직 방향으로 배열되는 spread 벡터를 생략하고, 각 spread 벡터가 대략 가로 방향으로 배열되도록 함으로써, 수평 방향으로의 소리의 확장을 적절하게 표현할 수 있게 된다.
- [0150] (spread 중심 벡터 방식)
- [0151] 계속해서, spread 중심 벡터 방식에 대하여 설명한다.
- [0152] spread 중심 벡터 방식에서는, 비트 스트림 내에 3차원 벡터인 spread 중심 벡터가 저장되어서 전송된다. 여기에서는, 예를 들어 오브젝트마다의 각 오디오 신호의 프레임 메타데이터에, spread 중심 벡터가 저장된다고 하자. 이 경우, 메타데이터에는, 음상의 범위 정도를 나타내는 spread도 저장되어 있다.
- [0153] spread 중심 벡터는, 오브젝트의 음상의 범위를 나타내는 영역의 중심 위치 $p0$ 을 나타내는 벡터이며, 예를 들어 spread 중심 벡터는, 중심 위치 $p0$ 의 수평 방향 각도를 나타내는 azimuth, 중심 위치 $p0$ 의 수직 방향 각도를 나타내는 elevation, 및 중심 위치 $p0$ 의 반경 방향의 거리를 나타내는 radius의 3가지의 요소를 포함하는 3차원 벡터로 된다.
- [0154] 즉, spread 중심 벡터=(azimuth, elevation, radius)이다.
- [0155] 렌더링 처리 시에는, 이 spread 중심 벡터에 의해 나타나는 위치가 중심 위치 $p0$ 으로 되어, spread 벡터로서 spread 벡터 $p0$ 내지 spread 벡터 $p18$ 이 산출된다. 여기서, spread 벡터 $p0$ 은, 예를 들어 도 4에 도시한 바와 같이, 원점 0를 시점으로 하고, 중심 위치 $p0$ 을 중점으로 하는 벡터 $p0$ 이다. 또한, 도 4에 있어서, 도 3에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 적절히 생략한다.
- [0156] 또한, 도 4에서는, 점선으로 그려진 화살표가 spread 벡터를 나타내고 있고, 도 4에 있어서도 도면을 보기 쉽게 하기 위해서 spread 벡터가 9개만 그려져 있다.
- [0157] 도 3에 도시한 예에서는, 위치 p =중심 위치 $p0$ 으로 되어 있었지만, 도 4에 도시하는 예에서는, 중심 위치 $p0$ 은, 위치 p 와는 다른 위치로 되어 있다. 이 예에서는, 중심 위치 $p0$ 을 중심으로 하는 음상의 범위를 나타내는 영역 R21은, 오브젝트의 위치인 위치 p 에 대하여 도 3의 예보다도 도면 중, 좌측으로 어긋나 있음을 알 수 있다.
- [0158] 이렇게 음상의 범위를 나타내는 영역의 중심 위치 $p0$ 으로서, spread 중심 벡터에 의해 임의의 위치를 지정할 수 있도록 하면, 오브젝트의 소리의 지향성을 더욱 정확하게 표현할 수 있게 된다.

- [0159] spread 중심 벡터 방식에서는, spread 벡터 p0 내지 spread 벡터 p18이 얻어지면, 그 후, 벡터 p에 대하여 처리 B1이 행해지고, spread 벡터 p0 내지 spread 벡터 p18에 대하여 처리 B2가 행해진다.
- [0160] 또한, 처리 B2에서는, 19개의 각 spread 벡터에 대하여 VBAP 게인이 산출되게 해도 되고, spread 벡터 p0을 제외한 spread 벡터 p1 내지 spread 벡터 p18에 대해서만 VBAP 게인이 산출되게 해도 된다. 이하에서는, spread 벡터 p0에 대해서도 VBAP 게인이 산출되는 것으로 하여 설명을 계속한다.
- [0161] 또한, 각 벡터의 VBAP 게인이 산출되면, 그 후에는 처리 B3, 처리 B4, 및 처리 B5'가 행해져서, 각 스피커에 공급되는 오디오 신호가 생성된다. 또한, 처리 B3 후, 필요에 따라 VBAP 게인 가산값의 양자화가 행해진다.
- [0162] 이상과 같은 spread 중심 벡터 방식에서도, 렌더링에 의해, 충분히 고품질의 음성을 얻을 수 있다.
- [0163] (spread 단부 벡터 방식)
- [0164] 이어서, spread 단부 벡터 방식에 대하여 설명한다.
- [0165] spread 단부 벡터 방식에서는, 비트 스트림 내에 5차원 벡터인 spread 단부 벡터가 저장되어서 전송된다. 여기에서는, 예를 들어 오브젝트마다의 각 오디오 신호의 프레임 메타데이터에, spread 단부 벡터가 저장된다고 하자. 이 경우, 메타데이터에는, 음상의 범위 정도를 나타내는 spread는 저장되지 않는다.
- [0166] 예를 들어 spread 단부 벡터는, 오브젝트의 음상의 범위를 나타내는 영역을 나타내는 벡터이며, spread 단부 벡터는, spread 좌단 azimuth, spread 우단 azimuth, spread 상단 elevation, spread 하단 elevation, 및 spread 용 radius의 5가지의 요소 등을 포함하는 벡터이다.
- [0167] 여기서, spread 단부 벡터를 구성하는 spread 좌단 azimuth 및 spread 우단 azimuth는, 각각 음상의 범위를 나타내는 영역에서의, 수평 방향의 좌단 및 우단이 절대적인 위치를 나타내는 수평 방향 각도 azimuth의 값을 나타내고 있다. 바꾸어 말하면, spread 좌단 azimuth 및 spread 우단 azimuth는, 각각 음상의 범위를 나타내는 영역의 중심 위치 p0으로부터의 좌측 방향 및 우측 방향으로의 음상의 범위 정도를 나타내는 각도를 나타내고 있다.
- [0168] 또한, spread 상단 elevation 및 spread 하단 elevation은, 각각 음상의 범위를 나타내는 영역에서의, 수직 방향의 상단 및 하단의 절대적인 위치를 나타내는 수직 방향 각도 elevation의 값을 나타내고 있다. 바꾸어 말하면, spread 상단 elevation 및 spread 하단 elevation은, 각각 음상의 범위를 나타내는 영역의 중심 위치 p0으로부터의 상측 방향 및 하측 방향으로의 음상의 범위 정도를 나타내는 각도를 나타내고 있다. 또한, spread 용 radius는, 음상의 반경 방향의 깊이를 나타내고 있다.
- [0169] 또한, 여기에서는 spread 단부 벡터는, 공간에 있어서의 절대적인 위치를 나타내는 정보로 되어 있는데, spread 단부 벡터는, 오브젝트의 위치 정보에 의해 나타나는 위치 p에 대한 상대 위치를 나타내는 정보로 되도록 해도 된다.
- [0170] spread 단부 벡터 방식에서는, 이러한 spread 단부 벡터가 사용되어서 렌더링이 행해진다.
- [0171] 구체적으로는, spread 단부 벡터 방식에서는, spread 단부 벡터에 기초하여, 이하의 식 (4)를 계산함으로써, 중심 위치 p0이 산출된다.

$$\begin{aligned} \text{azimuth} &: (\text{spread 좌단 azimuth} + \text{spread 우단 azimuth}) / 2 \\ \text{elevation} &: (\text{spread 상단 elevation} + \text{spread 하단 elevation}) / 2 \\ \text{radius} &: \text{spread 용 radius} \end{aligned} \quad \cdot \cdot \cdot (4)$$

- [0172]
- [0173] 즉, 중심 위치 p0을 나타내는 수평 방향 각도 azimuth는, spread 좌단 azimuth와 spread 우단 azimuth의 중간(평균)의 각도로 되고, 중심 위치 p0을 나타내는 수직 방향 각도 elevation은, spread 상단 elevation과 spread 하단 elevation의 중간(평균)의 각도로 된다. 또한, 중심 위치 p0을 나타내는 거리 radius는, spread 용 radius로 된다.
- [0174] 따라서, spread 단부 벡터 방식에서는, 중심 위치 p0은, 위치 정보에 의해 나타나는 오브젝트의 위치 p와는 다른 위치가 되는 경우도 있다.

[0175] 또한, spread 단부 벡터 방식에서는, 다음 식 (5)를 계산함으로써, spread의 값이 산출된다.

$$\text{spread} : \max((\text{spread 좌단 azimuth}-\text{spread 우단 azimuth})/2, \\ (\text{spread 상단 elevation}-\text{spread 하단 elevation})/2) \\ \dots (5)$$

[0176]

[0177] 또한, 식 (5)에 있어서 $\max(a, b)$ 는 a와 b 중 큰 값을 돌려주는 함수를 나타내고 있다. 따라서, 여기에서는 spread 단부 벡터에 의해 나타나는 오브젝트의 음상의 범위를 나타내는 영역에서의, 수평 방향의 반경에 대응하는 각도인 $(\text{spread 좌단 azimuth}-\text{spread 우단 azimuth})/2$ 와, 수직 방향의 반경에 대응하는 각도인 $(\text{spread 상단 elevation}-\text{spread 하단 elevation})/2$ 중 큰 쪽의 값이 spread의 값으로 되게 된다.

[0178] 그리고, 이와 같이 하여 얻어진 spread의 값과, 중심 위치 p0(벡터 p0)에 기초하여, MPEG-H 3D Audio 규격에 있어서의 경우와 마찬가지로 18개의 spread 벡터 p1 내지 spread 벡터 p18이 산출된다.

[0179] 따라서, 중심 위치 p0을 중심으로 하여 단위 구면 상에서 상하 좌우 대칭이 되도록, 18개의 spread 벡터 p1 내지 spread 벡터 p18이 구해진다.

[0180] 또한, spread 단부 벡터 방식에서는, 원점 0를 시점으로 하고, 중심 위치 p0을 종점으로 하는 벡터 p0이 spread 벡터 p0으로 된다.

[0181] spread 단부 벡터 방식에 있어서도, spread 3차원 벡터 방식에 있어서의 경우와 마찬가지로, 각 spread 벡터는, 수평 방향 각도 azimuth, 수직 방향 각도 elevation, 및 거리 radius에 의해 표현된다. 즉, spread 벡터 pi (단, $i=0$ 내지 18)의 수평 방향 각도 azimuth 및 수직 방향 각도 elevation이, 각각 a(i) 및 e(i)로 된다.

[0182] 이와 같이 하여 spread 벡터 p0 내지 spread 벡터 p18이 얻어지면, 그 후, $(\text{spread 좌단 azimuth}-\text{spread 우단 azimuth})$ 와 $(\text{spread 상단 elevation}-\text{spread 하단 elevation})$ 의 비에 기초하여, 그들 spread 벡터 p1 내지 spread 벡터 p18이 변경(보정)되어, 최종적인 spread 벡터가 구해진다.

[0183] 즉, $(\text{spread 좌단 azimuth}-\text{spread 우단 azimuth})$ 가 $(\text{spread 상단 elevation}-\text{spread 하단 elevation})$ 보다도 큰 경우, 이하의 식 (6)의 계산이 행해지고, spread 벡터 p1 내지 spread 벡터 p18의 각각의 elevation인 e(i)가 e'(i)로 변경된다.

$$e'(i) = e(0) + (e(i) - e(0)) \times (\text{spread 상단 elevation}-\text{spread} \\ \text{하단 elevation}) / (\text{spread 좌단 azimuth}-\text{spread 우단 azimuth}) \\ \dots (6)$$

[0184]

[0185] 또한, spread 벡터 p0에 대해서는, elevation의 보정은 행해지지 않는다.

[0186] 이에 반해, $(\text{spread 좌단 azimuth}-\text{spread 우단 azimuth})$ 가 $(\text{spread 상단 elevation}-\text{spread 하단 elevation})$ 미만인 경우, 이하의 식 (7)의 계산이 행해지고, spread 벡터 p1 내지 spread 벡터 p18의 각각의 azimuth인 a(i)가 a'(i)로 변경된다.

$$a'(i) = a(0) + (a(i) - a(0)) \times (\text{spread 좌단 azimuth}-\text{spread} \\ \text{우단 azimuth}) / (\text{spread 상단 elevation}-\text{spread 하단 elevation}) \\ \dots (7)$$

[0187]

[0188] 또한, spread 벡터 p0에 대해서는, azimuth의 보정은 행해지지 않는다.

[0189] 이상에 있어서 설명한 spread 벡터의 산출 방법은, 기본적으로는 spread 3차원 벡터 방식에 있어서의 경우와 마찬가지로이다.

[0190] 따라서, 결국에는 이들의 처리는, spread 단부 벡터에 기초하여, 그 spread 단부 벡터에 의해 정해지는 단위 구면 상에 있어서의 원형 또는 타원형인 음상의 범위를 나타내는 영역에 대한 spread 벡터를 산출하는 처리가 된다.

- [0191] 이와 같이 하여 spread 벡터가 얻어지면, 그 후, 벡터 p와, spread 벡터 p0 내지 spread 벡터 p18이 사용되어서 상술한 처리 B1, 처리 B2, 처리 B3, 처리 B4, 및 처리 B5'가 행해져서, 각 스피커에 공급되는 오디오 신호가 생성된다.
- [0192] 또한, 처리 B2에서는, 19개의 각 spread 벡터에 대하여 스피커마다의 VBAP 계인이 산출된다. 또한, 처리 B3 후, 필요에 따라 VBAP 계인 가산값의 양자화가 행해진다.
- [0193] 이렇게 spread 단부 벡터에 의해, 음상의 범위를 나타내는 영역을, 임의의 위치를 중심 위치 p0으로 하는 임의의 형상의 영역으로 함으로써, 오브젝트의 형상이나, 오브젝트의 소리의 지향성을 표현할 수 있게 되어, 렌더링에 의해, 보다 고품질의 음성을 얻을 수 있다.
- [0194] 또한, 여기에서는 (spread 좌단 azimuth-spread 우단 azimuth)/2와 (spread 상단 elevation-spread 하단 elevation)/2 중 큰 쪽의 값이 spread의 값으로 되는 예에 대하여 설명했지만, 그들 중의 작은 쪽의 값이 spread의 값으로 되게 해도 된다.
- [0195] 또한, 여기에서는 spread 벡터 p0에 대하여 VBAP 계인을 산출하는 경우를 예로 들어 설명했지만, spread 벡터 p0에 대해서는 VBAP 계인을 산출하지 않도록 해도 된다. 이하에서는, spread 벡터 p0에 대해서도 VBAP 계인이 산출되는 것으로 하여 설명을 계속한다.
- [0196] 또한, spread 3차원 벡터 방식에 있어서의 경우와 마찬가지로, 예를 들어 (spread 좌단 azimuth-spread 우단 azimuth)와 (spread 상단 elevation-spread 하단 elevation)의 비에 따라, 생성되는 spread 벡터의 개수가 결정되게 해도 된다.
- [0197] (spread 방사 벡터 방식)
- [0198] 또한, spread 방사 벡터 방식에 대하여 설명한다.
- [0199] spread 방사 벡터 방식에서는, 비트 스트림 내에 3차원 벡터인 spread 방사 벡터가 저장되어서 전송된다. 여기에서는, 예를 들어 오브젝트마다의 각 오디오 신호의 프레임 메타데이터에, spread 방사 벡터가 저장된다고 하자. 이 경우, 메타데이터에는, 음상의 범위 정도를 나타내는 spread도 저장되어 있다.
- [0200] spread 방사 벡터는, 오브젝트의 위치 p에 대한, 오브젝트의 음상의 범위를 나타내는 영역의 중심 위치 p0의 상대적인 위치를 나타내는 벡터이다. 예를 들어 spread 방사 벡터는, 위치 p로부터 본, 중심 위치 p0까지의 수평 방향 각도를 나타내는 azimuth, 중심 위치 p0까지의 수직 방향 각도를 나타내는 elevation, 및 중심 위치 p0의 반경 방향의 거리를 나타내는 radius의 3가지의 요소를 포함하는 3차원 벡터로 된다.
- [0201] 즉, spread 방사 벡터=(azimuth, elevation, radius)이다.
- [0202] 렌더링 처리 시에는, 이 spread 방사 벡터와 벡터 p를 가산하여 얻어지는 벡터에 의해 나타나는 위치가 중심 위치 p0으로 되어, spread 벡터로서 spread 벡터 p0 내지 spread 벡터 p18이 산출된다. 여기서, spread 벡터 p0은, 예를 들어 도 5에 도시한 바와 같이, 원점 0를 시점으로 하고, 중심 위치 p0을 종점으로 하는 벡터 p0이다. 또한, 도 5에 있어서, 도 3에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 적절히 생략한다.
- [0203] 또한, 도 5에서는, 점선으로 그려진 화살표가 spread 벡터를 나타내고 있고, 도 5에 있어서도 도면을 보기 쉽게 하기 위해서 spread 벡터가 9개만 그려져 있다.
- [0204] 도 3에 도시한 예에서는, 위치 p=중심 위치 p0으로 되어 있었지만, 도 5에 도시하는 예에서는, 중심 위치 p0은, 위치 p와는 다른 위치로 되어 있다. 이 예에서는, 벡터 p와, 화살표 B11에 의해 나타나는 spread 방사 벡터를 벡터 가산하여 얻어지는 벡터의 종점 위치가 중심 위치 p0으로 되어 있다.
- [0205] 또한, 중심 위치 p0을 중심으로 하는 음상의 범위를 나타내는 영역 R31은, 오브젝트의 위치인 위치 p에 대하여 도 3의 예보다도 도면 중, 좌측으로 어긋나 있음을 알 수 있다.
- [0206] 이렇게 음상의 범위를 나타내는 영역의 중심 위치 p0으로서, spread 방사 벡터와 위치 p를 사용하여 임의의 위치를 지정할 수 있도록 하면, 오브젝트의 소리의 지향성을 더욱 정확하게 표현할 수 있게 된다.
- [0207] spread 방사 벡터 방식에서는, spread 벡터 p0 내지 spread 벡터 p18이 얻어지면, 그 후, 벡터 p에 대하여 처리 B1이 행해지고, spread 벡터 p0 내지 spread 벡터 p18에 대하여 처리 B2가 행해진다.

- [0208] 또한, 처리 B2에서는, 19개의 각 spread 벡터에 대하여 VBAP 게인이 산출되게 해도 되고, spread 벡터 p0을 제외한 spread 벡터 p1 내지 spread 벡터 p18에 대해서만 VBAP 게인이 산출되게 해도 된다. 이하에서는, spread 벡터 p0에 대해서도 VBAP 게인이 산출되는 것으로 하여 설명을 계속한다.
- [0209] 또한, 각 벡터의 VBAP 게인이 산출되면, 그 후에는 처리 B3, 처리 B4, 및 처리 B5'가 행해져서, 각 스피커에 공급되는 오디오 신호가 생성된다. 또한, 처리 B3 후, 필요에 따라 VBAP 게인 가산값의 양자화가 행해진다.
- [0210] 이상과 같은 spread 방사 벡터 방식에서도, 렌더링에 의해, 충분히 고품질의 음성을 얻을 수 있다.
- [0211] (임의 spread 벡터 방식)
- [0212] 이어서, 임의 spread 벡터 방식에 대하여 설명한다.
- [0213] 임의 spread 벡터 방식에서는, 비트 스트림 내에 VBAP 게인을 산출하는 spread 벡터의 수를 나타내는 spread 벡터수 정보와, 각 spread 벡터의 종점 위치를 나타내는 spread 벡터 위치 정보가 저장되어서 전송된다. 여기에서는, 예를 들어 오브젝트마다의 각 오디오 신호의 프레임 메타데이터에, spread 벡터수 정보와 spread 벡터 위치 정보가 저장된다고 하자. 이 경우, 메타데이터에는, 음상의 범위 정도를 나타내는 spread는 저장되지 않는다.
- [0214] 렌더링 처리 시에는, 각 spread 벡터 위치 정보에 기초하여, 원점 0를 시점으로 하고, spread 벡터 위치 정보에 의해 나타나는 위치를 종점으로 하는 벡터가 spread 벡터로서 산출된다.
- [0215] 그 후, 벡터 p에 대하여 처리 B1이 행해지고, 각 spread 벡터에 대하여 처리 B2가 행해진다. 또한, 각 벡터의 VBAP 게인이 산출되면, 그 후에는 처리 B3, 처리 B4, 및 처리 B5'가 행해져서, 각 스피커에 공급되는 오디오 신호가 생성된다. 또한, 처리 B3 후, 필요에 따라 VBAP 게인 가산값의 양자화가 행해진다.
- [0216] 이상과 같은 임의 spread 벡터 방식에서는, 임의로 음상을 확장하는 범위와 그 형상을 지정하는 것이 가능하므로, 렌더링에 의해, 충분히 고품질의 음성을 얻을 수 있다.
- [0217] <처리의 전환에 대해서>
- [0218] 본 기술에서는, 렌더러의 하드 규모 등에 따라서 렌더링 시의 처리로서 적절한 처리를 선택하고, 허용되는 처리량의 범위에서 가장 높은 품질의 음성을 얻을 수 있도록 하였다.
- [0219] 즉, 본 기술에서는, 복수의 처리의 전환을 가능하게 하기 위해서, 처리를 전환하기 위한 인덱스가 비트 스트림에 저장되어서 부호화 장치로부터 복호 장치에 전송된다. 즉, 처리를 전환하기 위한 인덱스 index가 비트 스트림 선택스에 추가 된다.
- [0220] 예를 들어 인덱스 index의 값에 따라, 이하와 같은 처리가 행해진다.
- [0221] 즉, 인덱스 index=0일 때에는, 복호 장치, 보다 상세하게는 복호 장치 내의 렌더러에서는, 종래의 MPEG-H 3D Audio 규격에 있어서의 경우와 동일한 렌더링이 행해진다.
- [0222] 또한, 예를 들어 인덱스 index=1일 때에는, 종래의 MPEG-H 3D Audio 규격에 있어서의 18개의 각 spread 벡터를 나타내는 인덱스의 조합 중, 소정의 조합의 각 인덱스가 비트 스트림에 저장되어서 송신된다. 이 경우, 렌더러에서는, 비트 스트림에 저장되어서 전송되어 온 각 인덱스에 의해 나타나는 spread 벡터에 대하여 VBAP 게인이 산출된다.
- [0223] 또한, 예를 들어 인덱스 index=2일 때에는, 처리에 사용하는 spread 벡터의 수를 나타내는 정보와, 처리에 사용하는 spread 벡터가, 종래의 MPEG-H 3D Audio 규격에 있어서의 18개의 spread 벡터 중 어느 spread 벡터인지를 나타내는 인덱스가 비트 스트림에 저장되어서 송신된다.
- [0224] 또한, 예를 들어 인덱스 index=3일 때에는, 상술한 임의 spread 벡터 방식으로 렌더링 처리가 행해지고, 예를 들어 인덱스 index=4일 때에는, 렌더링 처리에 있어서 상술한 VBAP 게인 가산값의 2치화가 행해진다. 또한, 예를 들어 인덱스 index=5일 때에는, 상술한 spread 중심 벡터 방식으로 렌더링 처리가 행해지거나 하게 된다.
- [0225] 또한, 부호화 장치에 있어서 처리를 전환하기 위한 인덱스 index를 지정하는 것이 아니고, 복호 장치 내의 렌더러에 있어서, 처리가 선택되게 해도 된다.
- [0226] 그러한 경우, 예를 들어 오브젝트의 메타데이터에 포함되어 있는 중요도 정보에 기초하여, 처리를 전환하는 것이 생각된다. 구체적으로는, 예를 들어 중요도 정보에 의해 나타나는 중요도가 높은(소정값 이상임) 오브젝트

에 대해서는, 상술한 인덱스 index=0에 의해 나타나는 처리가 행해지고, 중요도 정보에 의해 나타나는 중요도가 낮은(소정값 미만임) 오브젝트에 대해서는, 상술한 인덱스 index=4에 의해 나타나는 처리가 행해지는 등으로 할 수 있다.

- [0227] 이와 같이, 적절히, 렌더링 시의 처리를 전환함으로써, 렌더러의 하드 규모 등에 따라, 허용되는 처리량의 범위에서 가장 높은 품질의 음성을 얻을 수 있다.
- [0228] <음성 처리 장치의 구성예>
- [0229] 계속해서, 이상에 있어서 설명한 본 기술의 보다 구체적인 실시 형태에 대하여 설명한다.
- [0230] 도 6은, 본 기술을 적용한 음성 처리 장치의 구성예를 도시하는 도면이다.
- [0231] 도 6에 도시하는 음성 처리 장치(11)에는, M개의 각 채널에 대응하는 스피커(12-1) 내지 스피커(12-M)가 접속되어 있다. 음성 처리 장치(11)는 외부로부터 공급된 오브젝트의 오디오 신호와 메타데이터에 기초하여, 각 채널의 오디오 신호를 생성하고, 그들 오디오 신호를 스피커(12-1) 내지 스피커(12-M)에 공급하여 음성을 재생시킨다.
- [0232] 또한, 이하, 스피커(12-1) 내지 스피커(12-M)를 특별히 구별할 필요가 없는 경우, 간단히 스피커(12)라고도 칭하기로 한다. 이들 스피커(12)는 공급된 오디오 신호에 기초하여 음성을 출력하는 음성 출력부이다.
- [0233] 스피커(12)는 콘텐츠 등을 시청하는 유저를 둘러싸도록 배치되어 있다. 예를 들어, 각 스피커(12)는 상술한 단위 구면 상에 배치되어 있다.
- [0234] 음성 처리 장치(11)는 취득부(21), 벡터 산출부(22), 계인 산출부(23), 및 계인 조정부(24)를 갖고 있다.
- [0235] 취득부(21)는 외부로부터 오브젝트의 오디오 신호와, 각 오브젝트의 오디오 신호의 프레임마다의 메타데이터를 취득한다. 예를 들어 오디오 신호 및 메타데이터는, 부호화 장치로부터 출력된 비트 스트림에 포함되어 있는 부호화 오디오 데이터 및 부호화 메타데이터를, 복호 장치로 복호함으로써 얻어진 것이다.
- [0236] 취득부(21)는 취득한 오디오 신호를 계인 조정부(24)에 공급함과 함께, 취득한 메타데이터를 벡터 산출부(22)에 공급한다. 여기서, 메타데이터에는, 예를 들어 오브젝트의 위치를 나타내는 위치 정보나, 오브젝트의 중요도를 나타내는 중요도 정보, 오브젝트의 음상의 범위 정도를 나타내는 spread 등이 필요에 따라서 포함되어 있다.
- [0237] 벡터 산출부(22)는 취득부(21)로부터 공급된 메타데이터에 기초하여 spread 벡터를 산출하여 계인 산출부(23)에 공급한다. 또한, 벡터 산출부(22)는 필요에 따라, 메타데이터에 포함되는 위치 정보에 의해 나타나는 오브젝트의 위치 p, 즉 위치 p를 나타내는 벡터 p도 계인 산출부(23)에 공급한다.
- [0238] 계인 산출부(23)는 벡터 산출부(22)로부터 공급된 spread 벡터나 벡터 p에 기초하여, VBAP에 의해 각 채널에 대응하는 스피커(12)의 VBAP 계인을 산출하고, 계인 조정부(24)에 공급한다. 또한, 계인 산출부(23)는 각 스피커의 VBAP 계인을 양자화하는 양자화부(31)를 구비하고 있다.
- [0239] 계인 조정부(24)는 계인 산출부(23)로부터 공급된 각 VBAP 계인에 기초하여, 취득부(21)로부터 공급된 오브젝트의 오디오 신호에 대한 계인 조정을 행하고, 그 결과 얻어진 M개의 각 채널의 오디오 신호를 스피커(12)에 공급한다.
- [0240] 계인 조정부(24)는 증폭부(32-1) 내지 증폭부(32-M)를 구비하고 있다. 증폭부(32-1) 내지 증폭부(32-M)는, 취득부(21)로부터 공급된 오디오 신호에, 계인 산출부(23)로부터 공급된 VBAP 계인을 승산하고, 그 결과 얻어진 오디오 신호를 스피커(12-1) 내지 스피커(12-M)에 공급하고, 음성을 재생시킨다.
- [0241] 또한, 이하, 증폭부(32-1) 내지 증폭부(32-M)를 특별히 구별할 필요가 없는 경우, 간단히 증폭부(32)라고도 칭한다.
- [0242] <재생 처리의 설명>
- [0243] 계속해서, 도 6에 도시한 음성 처리 장치(11)의 동작에 대하여 설명한다.
- [0244] 음성 처리 장치(11)는 외부로부터 오브젝트의 오디오 신호와 메타데이터가 공급되면, 재생 처리를 행하여 오브젝트의 음성을 재생시킨다.
- [0245] 이하, 도 7의 흐름도를 참조하여, 음성 처리 장치(11)에 의한 재생 처리에 대하여 설명한다. 또한, 이 재생 처리는, 오디오 신호의 프레임마다 행해진다.

- [0246] 스텝 S11에 있어서, 취득부(21)는 외부로부터 오브젝트의 1 프레임분의 오디오 신호 및 메타데이터를 취득하고, 오디오 신호를 증폭부(32)에 공급함과 함께, 메타데이터를 벡터 산출부(22)에 공급한다.
- [0247] 스텝 S12에 있어서, 벡터 산출부(22)는 취득부(21)로부터 공급된 메타데이터에 기초하여 spread 벡터 산출 처리를 행하고, 그 결과 얻어진 spread 벡터를 게인 산출부(23)에 공급한다. 또한, 벡터 산출부(22)는 필요에 따라 벡터 p도 게인 산출부(23)에 공급한다.
- [0248] 또한, spread 벡터 산출 처리의 상세는 후술하겠지만, 이 spread 벡터 산출 처리에서는, 상술한 spread 3차원 벡터 방식, spread 중심 벡터 방식, spread 단부 벡터 방식, spread 방사 벡터 방식, 또는 임의 spread 벡터 방식에 의해 spread 벡터가 산출된다.
- [0249] 스텝 S13에 있어서, 게인 산출부(23)는 미리 보유하고 있는 각 스피커(12)의 배치 위치를 나타내는 배치 위치 정보와, 벡터 산출부(22)로부터 공급된 spread 벡터 및 벡터 p에 기초하여, 각 스피커(12)의 VBAP 게인을 산출한다.
- [0250] 즉, spread 벡터나 벡터 p의 각 벡터에 대해서, 각 스피커(12)의 VBAP 게인이 산출된다. 이에 의해, spread 벡터나 벡터 p라고 하는 벡터마다, 오브젝트의 위치 근방, 보다 상세하게는 벡터에 의해 나타나는 위치 근방에 위치하는 1 이상의 스피커(12)의 VBAP 게인이 얻어진다. 또한, spread 벡터의 VBAP 게인은 반드시 산출되지만, 스텝 S12의 처리에 의해, 벡터 산출부(22)로부터 게인 산출부(23)에 벡터 p가 공급되지 않은 경우에는, 벡터 p의 VBAP 게인은 산출되지 않는다.
- [0251] 스텝 S14에 있어서, 게인 산출부(23)는 스피커(12)마다, 각 벡터에 대하여 산출한 VBAP 게인을 가산하여 VBAP 게인 가산값을 산출한다. 즉, 동일한 스피커(12)에 대하여 산출된 각 벡터의 VBAP 게인의 가산값(총합)이 VBAP 게인 가산값으로서 산출된다.
- [0252] 스텝 S15에 있어서, 양자화부(31)는 VBAP 게인 가산값의 2치화를 행할지 여부를 판정한다.
- [0253] 예를 들어 2치화를 행할지 여부는, 상술한 인덱스 index에 기초하여 판정되어도 되고, 메타데이터로서의 중요도 정보에 의해 나타나는 오브젝트의 중요도에 기초하여 판정되도록 해도 된다.
- [0254] 인덱스 index에 기초하여 판정이 행해지는 경우에는, 예를 들어 비트 스트림으로부터 판독된 인덱스 index가 게인 산출부(23)에 공급되도록 하면 된다. 또한, 중요도 정보에 기초하여 판정이 행해지는 경우에는, 벡터 산출부(22)로부터 게인 산출부(23)에 중요도 정보가 공급되도록 하면 된다.
- [0255] 스텝 S15에 있어서 2치화를 행한다고 판정된 경우, 스텝 S16에 있어서, 양자화부(31)는 스피커(12)마다 구해진 VBAP 게인의 가산값, 즉 VBAP 게인 가산값을 2치화하고, 그 후, 처리는 스텝 S17로 진행한다.
- [0256] 이에 반해, 스텝 S15에 있어서 2치화를 행하지 않는다고 판정된 경우에는, 스텝 S16의 처리는 스킵되어, 처리는 스텝 S17로 진행한다.
- [0257] 스텝 S17에 있어서, 게인 산출부(23)는 모든 스피커(12)의 VBAP 게인의 2승합이 1로 되도록, 각 스피커(12)의 VBAP 게인을 정규화한다.
- [0258] 즉, 스피커(12)마다 구한 VBAP 게인의 가산값에 대해서, 그들 모든 가산값의 2승합이 1로 되도록 정규화가 행해진다. 게인 산출부(23)는 정규화에 의해 얻어진 각 스피커(12)의 VBAP 게인을, 그들 스피커(12)에 대응하는 증폭부(32)에 공급한다.
- [0259] 스텝 S18에 있어서, 증폭부(32)는 취득부(21)로부터 공급된 오디오 신호에, 게인 산출부(23)로부터 공급된 VBAP 게인을 승산하고, 스피커(12)에 공급한다.
- [0260] 그리고, 스텝 S19에 있어서 증폭부(32)는 공급한 오디오 신호에 기초하여 스피커(12)에 음성을 재생시키고 재생 처리는 종료한다. 이에 의해, 재생 공간에 있어서의 원하는 부분 공간에 오브젝트의 음상이 정위된다.
- [0261] 이상과 같이 하여 음성 처리 장치(11)는 메타데이터에 기초하여 spread 벡터를 산출하고, 스피커(12)마다 각 벡터의 VBAP 게인을 산출함과 함께, 그들 스피커(12)마다 VBAP 게인의 가산값을 구하여 정규화한다. 이렇게 spread 벡터에 대하여 VBAP 게인을 산출함으로써, 오브젝트의 음상의 범위, 특히 오브젝트의 형상이나 소리의 지향성을 표현할 수 있어, 보다 고품질의 음성을 얻을 수 있다.
- [0262] 게다가, 필요에 따라 VBAP 게인의 가산값을 2치화함으로써, 렌더링 시의 처리량을 삭감할 수 있을 뿐 아니라, 음성 처리 장치(11)의 처리 능력(하드 규모)에 따라서 적절한 처리를 행하여, 가능한 한 고품질의 음성을 얻을

수 있다.

[0263] <spread 벡터 산출 처리의 설명>

[0264] 여기서, 도 8의 흐름도를 참조하여, 도 7의 스텝 S12의 처리에 대응하는 spread 벡터 산출 처리에 대하여 설명한다.

[0265] 스텝 S41에 있어서, 벡터 산출부(22)는 spread 3차원 벡터에 기초하여 spread 벡터를 산출할 지 여부를 판정한다.

[0266] 예를 들어, 어떤 방법에 의해 spread 벡터를 산출할지는, 도 7의 스텝 S15에 있어서의 경우와 마찬가지로, 인덱스 index에 기초하여 판정되어도 되고, 중요도 정보에 의해 나타나는 오브젝트의 중요도에 기초하여 판정되도록 해도 된다.

[0267] 스텝 S41에 있어서, spread 3차원 벡터에 기초하여 spread 벡터를 산출한다고 판정된 경우, 즉, spread 3차원 벡터 방식에 의해 spread 벡터를 산출한다고 판정된 경우, 처리는 스텝 S42로 진행한다.

[0268] 스텝 S42에 있어서, 벡터 산출부(22)는 spread 3차원 벡터에 기초하는 spread 벡터 산출 처리를 행하고, 얻어진 벡터를 계인 산출부(23)에 공급한다. 또한, spread 3차원 벡터에 기초하는 spread 벡터 산출 처리의 상세는 후술한다.

[0269] spread 벡터가 산출되면, spread 벡터 산출 처리는 종료되고, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.

[0270] 이에 반해, 스텝 S41에 있어서 spread 3차원 벡터에 기초하여 spread 벡터를 산출하지 않는다고 판정된 경우, 처리는 스텝 S43으로 진행한다.

[0271] 스텝 S43에 있어서, 벡터 산출부(22)는 spread 중심 벡터에 기초하여 spread 벡터를 산출할 지 여부를 판정한다.

[0272] 스텝 S43에 있어서, spread 중심 벡터에 기초하여 spread 벡터를 산출한다고 판정된 경우, 즉, spread 중심 벡터 방식에 의해 spread 벡터를 산출한다고 판정된 경우, 처리는 스텝 S44로 진행한다.

[0273] 스텝 S44에 있어서, 벡터 산출부(22)는 spread 중심 벡터에 기초하는 spread 벡터 산출 처리를 행하고, 얻어진 벡터를 계인 산출부(23)에 공급한다. 또한, spread 중심 벡터에 기초하는 spread 벡터 산출 처리의 상세는 후술한다.

[0274] spread 벡터가 산출되면, spread 벡터 산출 처리는 종료되고, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.

[0275] 한편, 스텝 S43에 있어서 spread 중심 벡터에 기초하여 spread 벡터를 산출하지 않는다고 판정된 경우, 처리는 스텝 S45로 진행한다.

[0276] 스텝 S45에 있어서, 벡터 산출부(22)는 spread 단부 벡터에 기초하여 spread 벡터를 산출할 지 여부를 판정한다.

[0277] 스텝 S45에 있어서, spread 단부 벡터에 기초하여 spread 벡터를 산출한다고 판정된 경우, 즉, spread 단부 벡터 방식에 의해 spread 벡터를 산출한다고 판정된 경우, 처리는 스텝 S46으로 진행한다.

[0278] 스텝 S46에 있어서, 벡터 산출부(22)는 spread 단부 벡터에 기초하는 spread 벡터 산출 처리를 행하고, 얻어진 벡터를 계인 산출부(23)에 공급한다. 또한, spread 단부 벡터에 기초하는 spread 벡터 산출 처리의 상세는 후술한다.

[0279] spread 벡터가 산출되면, spread 벡터 산출 처리는 종료되고, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.

[0280] 또한, 스텝 S45에 있어서 spread 단부 벡터에 기초하여 spread 벡터를 산출하지 않는다고 판정된 경우, 처리는 스텝 S47로 진행한다.

[0281] 스텝 S47에 있어서, 벡터 산출부(22)는 spread 방사 벡터에 기초하여 spread 벡터를 산출할 지 여부를 판정한다.

[0282] 스텝 S47에 있어서, spread 방사 벡터에 기초하여 spread 벡터를 산출한다고 판정된 경우, 즉, spread 방사 벡터 방식에 의해 spread 벡터를 산출한다고 판정된 경우, 처리는 스텝 S48로 진행한다.

[0283] 스텝 S48에 있어서, 벡터 산출부(22)는 spread 방사 벡터에 기초하는 spread 벡터 산출 처리를 행하고, 얻어진

백터를 게인 산출부(23)에 공급한다. 또한, spread 방사 백터에 기초하는 spread 백터 산출 처리의 상세는 후술한다.

- [0284] spread 백터가 산출되면, spread 백터 산출 처리는 종료되고, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.
- [0285] 또한, 스텝 S47에 있어서 spread 방사 백터에 기초하여 spread 백터를 산출하지 않는다고 판정된 경우, 즉 임의 spread 백터 방식에 의해 spread 백터를 산출한다고 판정된 경우, 처리는 스텝 S49로 진행한다.
- [0286] 스텝 S49에 있어서, 백터 산출부(22)는 spread 백터 위치 정보에 기초하는 spread 백터 산출 처리를 행하고, 얻어진 백터를 게인 산출부(23)에 공급한다. 또한, spread 백터 위치 정보에 기초하는 spread 백터 산출 처리의 상세는 후술한다.
- [0287] spread 백터가 산출되면, spread 백터 산출 처리는 종료되고, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.
- [0288] 이상과 같이 하여 음성 처리 장치(11)는 복수의 방식 중 적절한 방식에 의해 spread 백터를 산출한다. 이렇게 적절한 방식에 의해 spread 백터를 산출함으로써, 렌더러의 하드 규모 등에 따라, 허용되는 처리량의 범위에서 가장 높은 품질의 음성을 얻을 수 있다.
- [0289] <spread 3차원 백터에 기초하는 spread 백터 산출 처리의 설명>
- [0290] 이어서, 도 8을 참조하여 설명한 스텝 S42, 스텝 S44, 스텝 S46, 스텝 S48, 및 스텝 S49의 각 처리에 대응하는 처리의 상세에 대하여 설명한다.
- [0291] 먼저, 도 9의 흐름도를 참조하여, 도 8의 스텝 S42에 대응하는 spread 3차원 백터에 기초하는 spread 백터 산출 처리에 대하여 설명한다.
- [0292] 스텝 S81에 있어서, 백터 산출부(22)는 취득부(21)로부터 공급된 메타데이터에 포함되는 위치 정보에 의해 나타나는 위치를, 오브젝트 위치 p로 한다. 즉, 위치 p를 나타내는 백터가 백터 p로 된다.
- [0293] 스텝 S82에 있어서, 백터 산출부(22)는 취득부(21)로부터 공급된 메타데이터에 포함되는 spread 3차원 백터에 기초하여 spread를 산출한다. 구체적으로는, 백터 산출부(22)는 상술한 식 (1)을 계산함으로써, spread를 산출한다.
- [0294] 스텝 S83에 있어서, 백터 산출부(22)는 백터 p와 spread에 기초하여, spread 백터 p0 내지 spread 백터 p18을 산출한다.
- [0295] 여기에서는, 백터 p가 중심 위치 p0을 나타내는 백터 p0으로 됨과 함께, 백터 p가 그대로 spread 백터 p0으로 된다. 또한, spread 백터 p1 내지 spread 백터 p18에 대해서는, MPEG-H 3D Audio 규격에 있어서의 경우와 마찬가지로, 중심 위치 p0을 중심으로 하는, 단위 구면 상의 spread에 나타나는 각도에 의해 정해지는 영역 내에 있어서, 상하 좌우 대칭이 되도록 각 spread 백터가 산출된다.
- [0296] 스텝 S84에 있어서, 백터 산출부(22)는 spread 3차원 백터에 기초하여, $s3_azimuth \geq s3_elevation$ 인지 여부, 즉 $s3_azimuth$ 가 $s3_elevation$ 보다도 큰지 여부를 판정한다.
- [0297] 스텝 S84에 있어서 $s3_azimuth \geq s3_elevation$ 이라고 판정된 경우, 스텝 S85에 있어서, 백터 산출부(22)는 spread 백터 p1 내지 spread 백터 p18의 elevation을 변경한다. 즉, 백터 산출부(22)는 상술한 식 (2)의 계산을 행하고, 각 spread 백터의 elevation을 보정하고, 최종적인 spread 백터로 한다.
- [0298] 최종적인 spread 백터가 얻어지면, 백터 산출부(22)는 이들 spread 백터 p0 내지 spread 백터 p18을 게인 산출부(23)에 공급하고, spread 3차원 백터에 기초하는 spread 백터 산출 처리는 종료한다. 그렇게 하면, 도 8의 스텝 S42의 처리가 종료되므로, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.
- [0299] 이에 반해, 스텝 S84에 있어서 $s3_azimuth \geq s3_elevation$ 이 아니라고 판정된 경우, 스텝 S86에 있어서, 백터 산출부(22)는 spread 백터 p1 내지 spread 백터 p18의 azimuth를 변경한다. 즉, 백터 산출부(22)는 상술한 식 (3)의 계산을 행하고, 각 spread 백터의 azimuth를 보정하고, 최종적인 spread 백터로 한다.
- [0300] 최종적인 spread 백터가 얻어지면, 백터 산출부(22)는 이들 spread 백터 p0 내지 spread 백터 p18을 게인 산출부(23)에 공급하고, spread 3차원 백터에 기초하는 spread 백터 산출 처리는 종료한다. 그렇게 하면, 도 8의 스텝 S42의 처리가 종료되므로, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.
- [0301] 이상과 같이 하여 음성 처리 장치(11)는 spread 3차원 백터 방식에 의해 각 spread 백터를 산출한다. 이에 의

해, 오브젝트의 형상이나, 오브젝트의 소리의 지향성을 표현할 수 있게 되어, 보다 고품질의 음성을 얻을 수 있다.

- [0302] <spread 중심 벡터에 기초하는 spread 벡터 산출 처리의 설명>
- [0303] 이어서, 도 10의 흐름도를 참조하여, 도 8의 스텝 S44에 대응하는 spread 중심 벡터에 기초하는 spread 벡터 산출 처리에 대하여 설명한다.
- [0304] 또한, 스텝 S111의 처리는, 도 9의 스텝 S81의 처리와 동일하므로, 그 설명은 생략한다.
- [0305] 스텝 S112에 있어서, 벡터 산출부(22)는 취득부(21)로부터 공급된 메타데이터에 포함되는 spread 중심 벡터와 spread에 기초하여, spread 벡터 p0 내지 spread 벡터 p18을 산출한다.
- [0306] 구체적으로는, 벡터 산출부(22)는 spread 중심 벡터에 의해 나타나는 위치를 중심 위치 p0으로 하고, 그 중심 위치 p0을 나타내는 벡터를 spread 벡터 p0으로 한다. 또한, 벡터 산출부(22)는 중심 위치 p0을 중심으로 하는, 단위 구면 상의 spread에 나타나는 각도에 의해 정해지는 영역 내에 있어서, 상하 좌우 대칭이 되도록 spread 벡터 p1 내지 spread 벡터 p18을 구한다. 이들 spread 벡터 p1 내지 spread 벡터 p18은, 기본적으로는 MPEG-H 3D Audio 규격에 있어서의 경우와 동일하게 하여 구해진다.
- [0307] 벡터 산출부(22)는 이상의 처리에 의해 얻어진 벡터 p와, spread 벡터 p0 내지 spread 벡터 p18을 게인 산출부(23)에 공급하고, spread 중심 벡터에 기초하는 spread 벡터 산출 처리는 종료한다. 그렇게 하면, 도 8의 스텝 S44의 처리가 종료되므로, 그 후, 처리는 도 7의 스텝 S13으로 진행된다.
- [0308] 이상과 같이 하여 음성 처리 장치(11)는 spread 중심 벡터 방식에 의해 벡터 p와 각 spread 벡터를 산출한다. 이에 의해, 오브젝트의 형상이나, 오브젝트의 소리의 지향성을 표현할 수 있게 되어, 보다 고품질의 음성을 얻을 수 있다.
- [0309] 또한, spread 중심 벡터에 기초하는 spread 벡터 산출 처리에서는, spread 벡터 p0은 게인 산출부(23)에 공급하지 않도록 해도 된다. 즉, spread 벡터 p0에 대해서는 VBAP 게인을 산출하지 않도록 해도 된다.
- [0310] <spread 단부 벡터에 기초하는 spread 벡터 산출 처리의 설명>
- [0311] 또한, 도 11의 흐름도를 참조하여, 도 8의 스텝 S46에 대응하는 spread 단부 벡터에 기초하는 spread 벡터 산출 처리에 대하여 설명한다.
- [0312] 또한, 스텝 S141의 처리는, 도 9의 스텝 S81의 처리와 동일하므로, 그 설명은 생략한다.
- [0313] 스텝 S142에 있어서, 벡터 산출부(22)는 취득부(21)로부터 공급된 메타데이터에 포함되는 spread 단부 벡터에 기초하여 중심 위치 p0, 즉 벡터 p0을 산출한다. 구체적으로는, 벡터 산출부(22)는 상술한 식 (4)를 계산함으로써 중심 위치 p0을 산출한다.
- [0314] 스텝 S143에 있어서, 벡터 산출부(22)는 spread 단부 벡터에 기초하여 spread를 산출한다. 구체적으로는, 벡터 산출부(22)는 상술한 식 (5)를 계산함으로써, spread를 산출한다.
- [0315] 스텝 S144에 있어서, 벡터 산출부(22)는 중심 위치 p0과 spread에 기초하여, spread 벡터 p0 내지 spread 벡터 p18을 산출한다.
- [0316] 여기에서는, 중심 위치 p0을 나타내는 벡터 p0이 그대로 spread 벡터 p0으로 된다. 또한, spread 벡터 p1 내지 spread 벡터 p18에 대해서는, MPEG-H 3D Audio 규격에 있어서의 경우와 마찬가지로, 중심 위치 p0을 중심으로 하는, 단위 구면 상의 spread에 나타나는 각도에 의해 정해지는 영역 내에 있어서, 상하 좌우 대칭이 되도록 각 spread 벡터가 산출된다.
- [0317] 스텝 S145에 있어서, 벡터 산출부(22)는 $(\text{spread 좌단 azimuth} - \text{spread 우단 azimuth}) \geq (\text{spread 상단 elevation} - \text{spread 하단 elevation})$ 인지 여부, 즉 $(\text{spread 좌단 azimuth} - \text{spread 우단 azimuth})$ 가 $(\text{spread 상단 elevation} - \text{spread 하단 elevation})$ 보다도 큰지 여부를 판정한다.
- [0318] 스텝 S145에 있어서 $(\text{spread 좌단 azimuth} - \text{spread 우단 azimuth}) \geq (\text{spread 상단 elevation} - \text{spread 하단 elevation})$ 이라고 판정된 경우, 스텝 S146에 있어서, 벡터 산출부(22)는 spread 벡터 p1 내지 spread 벡터 p18의 elevation을 변경한다. 즉, 벡터 산출부(22)는 상술한 식 (6)의 계산을 행하고, 각 spread 벡터의 elevation을 보정하고, 최종적인 spread 벡터로 한다.

- [0319] 최종적인 spread 벡터가 얻어지면, 벡터 산출부(22)는 이들 spread 벡터 p0 내지 spread 벡터 p18과 벡터 p를 게인 산출부(23)에 공급하고, spread 단부 벡터에 기초하는 spread 벡터 산출 처리는 종료한다. 그렇게 하면, 도 8의 스텝 S46의 처리가 종료되므로, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.
- [0320] 이에 반해, 스텝 S145에 있어서 (spread 좌단 azimuth-spread 우단 azimuth) $\geq$ (spread 상단 elevation-spread 하단 elevation)이 아니라고 판정된 경우, 스텝 S147에 있어서, 벡터 산출부(22)는 spread 벡터 p1 내지 spread 벡터 p18의 azimuth를 변경한다. 즉, 벡터 산출부(22)는 상술한 식 (7)의 계산을 행하고, 각 spread 벡터의 azimuth를 보정하고, 최종적인 spread 벡터로 한다.
- [0321] 최종적인 spread 벡터가 얻어지면, 벡터 산출부(22)는 이들 spread 벡터 p0 내지 spread 벡터 p18과 벡터 p를 게인 산출부(23)에 공급하고, spread 단부 벡터에 기초하는 spread 벡터 산출 처리는 종료한다. 그렇게 하면, 도 8의 스텝 S46의 처리가 종료되므로, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.
- [0322] 이상과 같이 하여 음성 처리 장치(11)는 spread 단부 벡터 방식에 의해 각 spread 벡터를 산출한다. 이에 의해, 오브젝트의 형상이나, 오브젝트의 소리의 지향성을 표현할 수 있게 되어, 보다 고품질의 음성을 얻을 수 있다.
- [0323] 또한, spread 단부 벡터에 기초하는 spread 벡터 산출 처리에서는, spread 벡터 p0은 게인 산출부(23)에 공급하지 않도록 해도 된다. 즉, spread 벡터 p0에 대해서는 VBAP 게인을 산출하지 않도록 해도 된다.
- [0324] <spread 방사 벡터에 기초하는 spread 벡터 산출 처리의 설명>
- [0325] 이어서, 도 12의 흐름도를 참조하여, 도 8의 스텝 S48에 대응하는 spread 방사 벡터에 기초하는 spread 벡터 산출 처리에 대하여 설명한다.
- [0326] 또한, 스텝 S171의 처리는, 도 9의 스텝 S81의 처리와 동일하므로, 그 설명은 생략한다.
- [0327] 스텝 S172에 있어서, 벡터 산출부(22)는 오브젝트 위치 p와, 취득부(21)로부터 공급된 메타데이터에 포함되는 spread 방사 벡터 및 spread에 기초하여, spread 벡터 p0 내지 spread 벡터 p18을 산출한다.
- [0328] 구체적으로는, 벡터 산출부(22)는 오브젝트 위치 p를 나타내는 벡터 p와 spread 방사 벡터를 가산하여 얻어지는 벡터에 의해 나타나는 위치를 중심 위치 p0으로 한다. 이 중심 위치 p0을 나타내는 벡터가 벡터 p0이며, 벡터 산출부(22)는 벡터 p0을 그대로 spread 벡터 p0으로 한다.
- [0329] 또한, 벡터 산출부(22)는 중심 위치 p0을 중심으로 하는, 단위 구면 상의 spread에 나타나는 각도에 의해 정해지는 영역 내에 있어서, 상하 좌우 대칭이 되도록 spread 벡터 p1 내지 spread 벡터 p18을 구한다. 이들 spread 벡터 p1 내지 spread 벡터 p18은, 기본적으로는 MPEG-H 3D Audio 규격에 있어서의 경우와 동일하게 하여 구해진다.
- [0330] 벡터 산출부(22)는 이상의 처리에 의해 얻어진 벡터 p와, spread 벡터 p0 내지 spread 벡터 p18을 게인 산출부(23)에 공급하고, spread 방사 벡터에 기초하는 spread 벡터 산출 처리는 종료한다. 그렇게 하면, 도 8의 스텝 S48의 처리가 종료되므로, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.
- [0331] 이상과 같이 하여 음성 처리 장치(11)는 spread 방사 벡터 방식에 의해 벡터 p와 각 spread 벡터를 산출한다. 이에 의해, 오브젝트의 형상이나, 오브젝트의 소리의 지향성을 표현할 수 있게 되어, 보다 고품질의 음성을 얻을 수 있다.
- [0332] 또한, spread 방사 벡터에 기초하는 spread 벡터 산출 처리에서는, spread 벡터 p0은 게인 산출부(23)에 공급하지 않도록 해도 된다. 즉, spread 벡터 p0에 대해서는 VBAP 게인을 산출하지 않도록 해도 된다.
- [0333] <spread 벡터 위치 정보에 기초하는 spread 벡터 산출 처리의 설명>
- [0334] 이어서, 도 13의 흐름도를 참조하여, 도 8의 스텝 S49에 대응하는 spread 벡터 위치 정보에 기초하는 spread 벡터 산출 처리에 대하여 설명한다.
- [0335] 또한, 스텝 S201의 처리는, 도 9의 스텝 S81의 처리와 동일하므로, 그 설명은 생략한다.
- [0336] 스텝 S202에 있어서, 벡터 산출부(22)는 취득부(21)로부터 공급된 메타데이터에 포함되는 spread 벡터수 정보와 spread 벡터 위치 정보에 기초하여, spread 벡터를 산출한다.
- [0337] 구체적으로는, 벡터 산출부(22)는 원점 0를 시점으로 하고, spread 벡터 위치 정보에 의해 나타나는 위치를 중

점으로 하는 벡터를 spread 벡터로서 산출한다. 여기에서는, spread 벡터수 정보에 의해 나타나는 수만큼 spread 벡터가 산출된다.

[0338] 벡터 산출부(22)는 이상의 처리에 의해 얻어진 벡터 p와, spread 벡터를 게인 산출부(23)에 공급하고, spread 벡터 위치 정보에 기초하는 spread 벡터 산출 처리는 종료한다. 그렇게 하면, 도 8의 스텝 S49의 처리가 종료되므로, 그 후, 처리는 도 7의 스텝 S13으로 진행한다.

[0339] 이상과 같이 하여 음성 처리 장치(11)는 임의 spread 벡터 방식에 의해 벡터 p와 각 spread 벡터를 산출한다. 이에 의해, 오브젝트의 형상이나, 오브젝트의 소리의 지향성을 표현할 수 있게 되어, 보다 고품질의 음성을 얻을 수 있다.

[0340] <제2 실시 형태>

[0341] <렌더링 처리의 처리량 삭감에 대해서>

[0342] 그런데, 상술한 바와 같이, 복수의 스피커를 사용하여 음상의 정위를 제어하는, 즉 렌더링 처리를 행하는 기술로서 VBAP가 알려져 있다.

[0343] VBAP에서는, 3개의 스피커로부터 소리를 출력함으로써, 그들 3개의 스피커로 구성되는 삼각형의 내측의 임의의 1점에 음상을 정위시킬 수 있다. 이하에서는, 특히, 이러한 3개의 스피커로 구성되는 삼각형을 메쉬라 칭하기로 한다.

[0344] VBAP에 의한 렌더링 처리는, 오브젝트마다 행해지기 때문에, 예를 들어 게임 등, 오브젝트의 수가 많은 경우에는, 렌더링 처리의 처리량이 많아져버린다. 그로 인해, 하드 규모가 작은 렌더러에서는, 모든 오브젝트에 대하여 렌더링할 수 없어, 그 결과, 한정된 수의 오브젝트 소리밖에 재생되지 않는 경우가 있다. 그렇게 하면, 음성 재생 시에 임장감이나 음질이 손상되어버리는 경우가 있다.

[0345] 그래서, 본 기술에서는, 임장감이나 음질의 열화를 억제하면서 렌더링 처리의 처리량을 저감시킬 수 있도록 하였다.

[0346] 이하, 이러한 본 기술에 대하여 설명한다.

[0347] 통상의 VBAP 처리, 즉 렌더링 처리에서는, 오브젝트마다 상술한 처리 A1 내지 처리 A3의 처리가 행해져서, 각 스피커의 오디오 신호가 생성된다.

[0348] 실질적으로 VBAP 게인이 산출되는 스피커의 수는 3개이며, 각 스피커의 VBAP 게인은 오디오 신호를 구성하는 샘플마다 산출되므로, 처리 A3에 있어서의 승산 처리에서는, (오디오 신호의 샘플수×3)회의 승산이 행해지게 된다.

[0349] 이에 반해 본 기술에서는, VBAP 게인에 대한 게인 처리, 즉 VBAP 게인의 양자화 처리, 및 VBAP 게인 산출 시에 사용하는 메쉬수를 변경하는 메쉬수 전환 처리를, 적절히 조합하여 행함으로써 렌더링 처리의 처리량을 저감하도록 하였다.

[0350] (양자화 처리)

[0351] 먼저, 양자화 처리에 대하여 설명한다. 여기에서는, 양자화 처리의 예로서, 2치화 처리와 3치화 처리에 대하여 설명한다.

[0352] 양자화 처리로서 2치화 처리가 행해지는 경우, 처리 A1이 행해진 후, 그 처리 A1에 의해 각 스피커에 대하여 얻어진 VBAP 게인이 2치화된다. 2치화에서는, 예를 들어 각 스피커의 VBAP 게인이 0 또는 1 중 어느 값으로 된다.

[0353] 또한, VBAP 게인을 2치화하는 방법은, 예를 들어 반올림, 실링(절상), 플로어링(잘라 버림), 역치 처리 등, 어떤 방법이어도 된다.

[0354] 이와 같이 하여 VBAP 게인이 2치화되면, 그 후에는 처리 A2 및 처리 A3이 행해져서, 각 스피커의 오디오 신호가 생성된다.

[0355] 이때, 처리 A2에서는, 2치화된 VBAP 게인에 기초하여 정규화가 행해지므로, 상술한 spread 벡터의 양자화 시와 동일하도록 각 스피커의 최종적인 VBAP 게인은, 0을 제외하면 1가지가 된다. 즉, VBAP 게인을 2치화하면, 각 스피커의 최종적인 VBAP 게인의 값은 0이거나, 또는 소정값 중 어느 것이 된다.

- [0356] 따라서, 처리 A3에 있어서의 승산 처리에서는, (오디오 신호의 샘플수×1)회의 승산을 행하면 되므로, 렌더링 처리의 처리량을 대폭으로 삭감할 수 있다.
- [0357] 마찬가지로, 처리 A1 후, 각 스피커에 대하여 얻어진 VBAP 계인을 3치화하도록 해도 된다. 그러한 경우에는, 처리 A1에 의해 각 스피커에 대하여 얻어진 VBAP 계인이 3치화되어서 0, 0.5, 또는 1 중 어느 값으로 된다. 그리고, 그 후에는 처리 A2 및 처리 A3이 행해져서, 각 스피커의 오디오 신호가 생성된다.
- [0358] 따라서, 처리 A3에 있어서의 승산 처리에서의 승산 횟수는, 최대로 (오디오 신호의 샘플수×2)회가 되므로, 렌더링 처리의 처리량을 대폭으로 삭감할 수 있다.
- [0359] 또한, 여기에서는 VBAP 계인을 2치화 또는 3치화하는 경우를 예로 들어 설명하지만, VBAP 계인을 4 이상의 값으로 양자화하도록 해도 된다. 일반화하면, 예를 들어 VBAP 계인을 2 이상의 x개의 계인 중 어느 것이 되도록 양자화하면, 즉 VBAP 계인을 양자화수 x로 양자화하면, 처리 A3에 있어서의 승산 처리의 횟수는 최대로 (x-1)회가 된다.
- [0360] 이상과 같이 VBAP 계인을 양자화함으로써, 렌더링 처리의 처리량을 저감시킬 수 있다. 이렇게 렌더링 처리의 처리량이 적어지면, 오브젝트수가 많은 경우일지라도 모든 오브젝트의 렌더링을 행하는 것이 가능하게 되므로, 음성 재생 시에 있어서의 임장감이나 음질의 열화를 작게 억제할 수 있다. 즉, 임장감이나 음질의 열화를 억제하면서 렌더링 처리의 처리량을 저감시킬 수 있다.
- [0361] (메쉬수 전환 처리)
- [0362] 이어서, 메쉬수 전환 처리에 대하여 설명한다.
- [0363] VBAP에서는, 예를 들어 도 1을 참조하여 설명한 바와 같이, 처리 대상의 오브젝트 음상 위치 p를 나타내는 벡터 p가, 3개의 스피커(SP1) 내지 스피커(SP3)의 방향을 향하는 벡터 l₁ 내지 벡터 l₃의 선형합으로 표현되고, 이들 벡터에 승산되어 있는 계수 g₁ 내지 계수 g₃이 각 스피커의 VBAP 계인으로 된다. 도 1의 예에서는, 스피커(SP1) 내지 스피커(SP3)에 의해 둘러싸이는 삼각형의 영역 TR11이 하나의 메쉬가 되어 있다.
- [0364] VBAP 계인의 산출 시에는, 구체적으로는 다음 식 (8)에 의해, 삼각 형상의 메쉬의 역행렬 L₁₂₃⁻¹과 오브젝트의 음상 위치 p로부터 3개의 계수 g₁ 내지 계수 g₃을 계산에 의해 구할 수 있다.
- $$[g_1 \ g_2 \ g_3] = pL_{123}^{-1} = [p_1 \ p_2 \ p_3] \begin{bmatrix} l_{11} & l_{12} & l_{13} \\ l_{21} & l_{22} & l_{23} \\ l_{31} & l_{32} & l_{33} \end{bmatrix}^{-1} \quad \dots (8)$$
- [0365]
- [0366] 또한, 식 (8)에 있어서 p₁, p₂, 및 p₃은, 오브젝트의 음상 위치 p를 나타내는 직교 좌표계, 즉 도 2에 도시한 3차원 좌표계상의 x 좌표, y 좌표, 및 z 좌표를 나타내고 있다.
- [0367] 또한 l₁₁, l₁₂, 및 l₁₃은, 메쉬를 구성하는 첫번째의 스피커(SP1)를 향하는 벡터 l₁을 x축, y축, 및 z축의 성분으로 분해한 경우에 있어서의 x 성분, y 성분, 및 z 성분의 값이며, 첫번째의 스피커(SP1)의 x 좌표, y 좌표, 및 z 좌표에 상당한다.
- [0368] 마찬가지로, l₂₁, l₂₂, 및 l₂₃은, 메쉬를 구성하는 두번째 스피커(SP2)를 향하는 벡터 l₂를 x축, y축, 및 z축의 성분으로 분해한 경우에 있어서의 x 성분, y 성분, 및 z 성분의 값이다. 또한, l₃₁, l₃₂, 및 l₃₃은, 메쉬를 구성하는 세번째 스피커(SP3)를 향하는 벡터 l₃을 x축, y축, 및 z축의 성분으로 분해한 경우에 있어서의 x 성분, y 성분, 및 z 성분의 값이다.
- [0369] 또한, 위치 p의 3차원 좌표계의 p₁, p₂, 및 p₃으로부터, 구좌표계의 좌표 θ, γ, 및 r로의 변환은 r=1일 경우에는 다음 식 (9)에 도시하는 바와 같이 정의되어 있다. 여기서 θ, γ, 및 r은, 각각 상술한 수평 방향 각도 azimuth, 수직 방향 각도 elevation, 및 거리 radius이다.

$$[p_1 \ p_2 \ p_3] = [\cos(\theta) \times \cos(\gamma) \ \sin(\theta) \times \cos(\gamma) \ \sin(\gamma)]$$

. . . (9)

[0370]

[0371]

상술한 바와 같이 콘텐츠 재생측의 공간, 즉 재생 공간에서는, 단위 구 상에 복수의 스피커가 배치되어 있고, 그들 복수의 스피커 중 3개의 스피커로부터 하나의 메쉬가 구성된다. 그리고, 기본적으로는 단위 구의 표면 전체가 복수의 메쉬에 의해 간극 없이 덮여 있다. 또한, 각 메쉬는 서로 겹치지 않도록 정해진다.

[0372]

VBAP에서는, 단위 구의 표면 상에 배치된 스피커 중, 오브젝트의 위치 p 를 포함하는 하나의 메쉬를 구성하는 2개 또는 3개의 스피커로부터 음성을 출력하면, 음상을 위치 p 에 정위시킬 수 있으므로, 그 메쉬를 구성하는 스피커 이외의 VBAP 계인은 0이 된다.

[0373]

따라서, VBAP 계인의 산출 시에는, 오브젝트의 위치 p 를 포함하는 하나의 메쉬를 특정하고, 그 메쉬를 구성하는 스피커의 VBAP 계인을 산출하면 되게 된다. 예를 들어, 소정의 메쉬가 위치 p 를 포함하는 메쉬인지 여부는, 산출한 VBAP 계인으로부터 판정할 수 있다.

[0374]

즉, 메쉬에 대하여 산출된 3개의 각 스피커의 VBAP 계인이 모두 0 이상의 값이라면, 그 메쉬는 오브젝트의 위치 p 를 포함하는 메쉬이다. 반대로, 3개의 각 스피커의 VBAP 계인 중 1개라도 음의 값으로 된 경우에는, 오브젝트의 위치 p 는, 그들 스피커를 포함하는 메쉬 밖에 위치하고 있게 되므로, 산출된 VBAP 계인은 올바른 VBAP 계인이 아니다.

[0375]

그래서, VBAP 계인의 산출 시에는, 각 메쉬가 하나씩 차례로 처리 대상의 메쉬로서 선택되어 가고, 처리 대상의 메쉬에 대하여 상술한 식 (8)의 계산이 행해져서, 메쉬를 구성하는 각 스피커의 VBAP 계인이 산출된다.

[0376]

그리고, 그들 VBAP 계인의 산출 결과로부터, 처리 대상의 메쉬가 오브젝트의 위치 p 를 포함하는 메쉬인지가 판정되어, 위치 p 를 포함하지 않는 메쉬라고 판정된 경우에는, 다음 메쉬가 새로운 처리 대상의 메쉬로 되어 동일한 처리가 행해진다.

[0377]

한편, 처리 대상의 메쉬가 오브젝트의 위치 p 를 포함하는 메쉬라고 판정된 경우에는, 그 메쉬를 구성하는 스피커의 VBAP 계인이, 산출된 VBAP 계인으로 되고, 그 이외의 다른 스피커의 VBAP 계인은 0으로 된다. 이에 의해, 전체 스피커의 VBAP 계인이 얻어지게 된다.

[0378]

이렇게 렌더링 처리에서는, VBAP 계인을 산출하는 처리와, 위치 p 를 포함하는 메쉬를 특정하는 처리가 동시에 행해진다.

[0379]

즉, 올바른 VBAP 계인을 얻기 위해서, 메쉬를 구성하는 각 스피커의 VBAP 계인이 모두 0 이상의 값으로 되는 것이 얻어질 때까지, 처리 대상으로 하는 메쉬를 선택하고, 그 메쉬의 VBAP 계인을 산출하는 처리가 반복하여 행해진다.

[0380]

따라서 렌더링 처리에서는, 단위 구의 표면에 있는 메쉬의 수가 많을수록, 위치 p 를 포함하는 메쉬를 특정하기에, 즉 올바른 VBAP 계인을 얻기에 필요하게 되는 처리의 처리량이 많아진다.

[0381]

그래서, 본 기술에서는, 실제의 재생 환경의 스피커 모두를 사용하여 메쉬를 형성(구성)하는 것은 아니고, 전체 스피커 중 일부의 스피커만을 사용하여 메쉬를 형성하도록 함으로써, 메쉬의 총 수를 저감시키고, 렌더링 처리시의 처리량을 저감시키도록 하였다. 즉, 본 기술에서는, 메쉬의 총 수를 변경하는 메쉬수 전환 처리를 행하도록 하였다.

[0382]

구체적으로는, 예를 들어 22 채널의 스피커 시스템에서는, 도 14에 도시한 바와 같이 단위 구의 표면 상에 각 채널의 스피커로서, 스피커(SPK1) 내지 스피커(SPK22)의 합계 22개의 스피커가 배치된다. 또한, 도 14에 있어서, 원점 0는 도 2에 도시한 원점 0에 대응하는 것이다.

[0383]

이렇게 단위 구의 표면 상에 22개의 스피커가 배치된 경우, 그들 22개 모든 스피커를 사용하여 단위 구 표면을 덮도록 메쉬를 형성하면, 단위 구 상의 메쉬의 총 수는 40개가 된다.

[0384]

이에 반해, 예를 들어 도 15에 도시한 바와 같이 스피커(SPK1) 내지 스피커(SPK22)의 합계 22개의 스피커 중, 스피커(SPK1), 스피커(SPK6), 스피커(SPK7), 스피커(SPK10), 스피커(SPK19), 및 스피커(SPK20)의 합계 6개의 스피커만을 사용하여 메쉬를 형성한 것으로 한다. 또한, 도 15에 있어서 도 14에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 적절히 생략한다.

- [0385] 도 15의 예에서는, 22개의 스피커 중 합계 6의 스피커만이 사용되어서 메쉬가 형성되어 있으므로, 단위 구 상의 메쉬의 총 수는 8개가 되어, 대폭으로 메쉬의 총 수를 저감시킬 수 있다. 그 결과, 도 15에 도시하는 예에서는, 도 14에 도시한 22개의 스피커 모두를 사용하여 메쉬를 형성하는 경우와 비하여, VBAP 계인을 산출할 때의 처리량을 8/40배로 할 수 있어, 대폭으로 처리량을 저감시킬 수 있다.
- [0386] 또한, 이 예에 있어서도 단위 구의 표면 전체가 8개의 메쉬에 의해, 간극 없이 덮여 있으므로, 단위 구의 표면 상의 임의의 위치에 음상을 정위시키는 것이 가능하다. 단, 단위 구 표면에 설치된 메쉬의 총 수가 많을수록, 각 메쉬의 면적은 작아지므로, 메쉬 총 수가 많을수록, 보다 고정밀도로 음상의 정위를 제어하는 것이 가능하다.
- [0387] 메쉬수 전환 처리에 의해 메쉬 총 수가 변경된 경우, 변경 후의 수의 메쉬를 형성는 데에 사용하는 스피커를 선택하는데 있어서는, 원점 0에 있는 유저로부터 보아서 수직 방향(상하 방향), 즉 수직 방향 각도 elevation의 방향의 위치가 다른 스피커를 선택하는 것이 바람직하다. 바꾸어 말하면, 서로 다른 높이에 위치하는 스피커를 포함하는, 3 이상의 스피커를 사용하여, 변경 후의 수의 메쉬가 형성되도록 하는 것이 바람직하다. 이것은, 음성의 입체감, 즉 입장감의 열화를 억제하기 위해서이다.
- [0388] 예를 들어 도 16에 도시한 바와 같이, 단위 구 표면에 배치된 5개의 스피커(SP1) 내지 스피커(SP5)의 일부 또는 전부를 사용하여 메쉬를 형성하는 경우를 생각한다. 또한, 도 16에 있어서 도 3에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 생략한다.
- [0389] 도 16에 도시하는 예에 있어서, 5개의 스피커(SP1) 내지 스피커(SP5) 모두를 사용하여, 단위 구 표면이 덮이는 메쉬를 형성하는 경우, 메쉬의 수는 3개가 된다. 즉, 스피커(SP1) 내지 스피커(SP3)에 의해 둘러싸이는 삼각형의 영역, 스피커(SP2) 내지 스피커(SP4)에 의해 둘러싸이는 삼각형의 영역, 및 스피커(SP2), 스피커(SP4), 및 스피커(SP5)에 의해 둘러싸이는 삼각형의 영역 3개의 각 영역이 메쉬로 된다.
- [0390] 이에 반해, 예를 들어 스피커(SP1), 스피커(SP2), 및 스피커(SP5)만을 사용하면 메쉬가 삼각형이 아니고 2차원의 원호가 되어버린다. 이 경우, 단위 구에 있어서의, 스피커(SP1)와 스피커(SP2)를 연결하는 호 상, 또는 스피커(SP2)와 스피커(SP5)를 연결하는 호 상에밖에 오브젝트의 음상을 정위시킬 수 없게 된다.
- [0391] 이렇게 메쉬를 형성하는 데에 사용하는 스피커를, 모두 수직 방향에 있어서의 동일한 높이, 즉 동일한 레이어의 스피커로 하면, 전체 오브젝트의 음상 정위 위치의 높이가 동일한 높이가 되어버리기 때문에, 입장감이 열화되어버린다.
- [0392] 따라서, 수직 방향(연직 방향)의 위치가 서로 다른 스피커를 포함하는 3 이상의 스피커를 사용하여 1개 또는 복수의 메쉬를 형성하여, 입장감의 열화를 억제할 수 있도록 하는 것이 바람직하다.
- [0393] 도 16의 예에서는, 예를 들어 스피커(SP1) 내지 스피커(SP5) 중, 스피커(SP1) 및 스피커(SP3) 내지 스피커(SP5)를 사용하면, 단위 구 표면 전체를 덮도록 2개의 메쉬를 형성할 수 있다. 이 예에서는, 스피커(SP1) 및 스피커(SP5)와, 스피커(SP3) 및 스피커(SP4)가 서로 다른 높이에 위치하고 있다.
- [0394] 이 경우, 예를 들어 스피커(SP1), 스피커(SP3), 및 스피커(SP5)에 의해 둘러싸이는 삼각형의 영역과, 스피커(SP3) 내지 스피커(SP5)에 의해 둘러싸이는 삼각형의 영역의 2개의 영역이 각각 메쉬로 된다.
- [0395] 기타, 이 예에서는, 스피커(SP1), 스피커(SP3), 및 스피커(SP4)에 의해 둘러싸이는 삼각형의 영역과, 스피커(SP1), 스피커(SP4), 및 스피커(SP5)에 의해 둘러싸이는 삼각형의 영역의 2개의 영역을 각각 메쉬로 하는 것도 가능하다.
- [0396] 이들 2가지의 예에서는, 어느 경우에도 단위 구 표면 상의 임의의 위치에 음상을 정위시킬 수 있으므로, 입장감의 열화를 억제할 수 있다. 또한, 단위 구 표면 전체가 복수의 메쉬로 덮이도록 메쉬를 형성하기 위해서는, 유저의 바로 위에 위치하는, 소위 톱 스피커가 반드시 사용되도록 하면 된다. 예를 들어 톱 스피커는, 도 14에 도시한 스피커(SPK19)이다.
- [0397] 이상과 같이 메쉬수 전환 처리를 행하여 메쉬의 총 수를 변경함으로써, 렌더링 처리의 처리량을 저감시킬 수 있고, 또한 양자화 처리의 경우와 마찬가지로 음성 재생 시에 있어서의 입장감이나 음질의 열화를 작게 억제할 수 있다. 즉, 입장감이나 음질의 열화를 억제하면서 렌더링 처리의 처리량을 저감시킬 수 있다.
- [0398] 이러한 메쉬수 전환 처리를 행할지 여부나, 메쉬수 전환 처리에서 메쉬의 총 수를 몇으로 할지를 선택하는 것은, VBAP 계인을 산출하는 데에 사용하는 메쉬의 총 수를 선택하는 것이라고 하는 것이 가능하다.

- [0399] (양자화 처리와 메쉬수 전환 처리의 조합)
- [0400] 또한, 이상에 있어서는 렌더링 처리의 처리량을 저감시키는 방법으로서, 양자화 처리와 메쉬수 전환 처리에 대하여 설명하였다.
- [0401] 렌더링 처리를 행하는 렌더러측에서는, 양자화 처리나 메쉬수 전환 처리로서 설명한 각 처리 중 어느 것이 고정적으로 사용되게 해도 되고, 이들 처리가 전환되거나, 이들 처리가 적절히 조합되거나 해도 된다.
- [0402] 예를 들어 어떤 처리를 조합하여 행할지는, 오브젝트의 총 수(이하, 오브젝트수라고 칭한다)나, 오브젝트의 메타데이터에 포함되어 있는 중요도 정보, 오브젝트의 오디오 신호의 음압 등에 기초하여 정해지게 하면 된다. 또한, 처리의 조합, 즉 처리의 전환은, 오브젝트마다나, 오디오 신호의 프레임마다 행해지도록 하는 것이 가능하다.
- [0403] 예를 들어 오브젝트수에 따라서 처리의 전환을 행하는 경우, 다음과 같은 처리를 행하도록 할 수 있다.
- [0404] 예를 들어 오브젝트수가 10 이상인 경우, 모든 오브젝트에 대해서, VBAP 계인에 대한 2치화 처리가 행해지도록 한다. 이에 반해, 오브젝트수가 10 미만인 경우, 모든 오브젝트에 대해서, 종래대로 상술한 처리 A1 내지 처리 A3만이 행해지도록 한다.
- [0405] 이와 같이, 오브젝트수가 적을 때에는 종래대로의 처리를 행하고, 오브젝트수가 많을 때에는 2치화 처리를 행하도록 함으로써, 하드 규모가 작은 렌더러로도 충분히 렌더링을 행할 수 있고, 또한 가능한 한 품질이 높은 음성을 얻을 수 있다.
- [0406] 또한, 오브젝트수에 따라서 처리의 전환을 행하는 경우, 오브젝트수에 따라서 메쉬수 전환 처리를 행하여, 메쉬의 총 수를 적절하게 변경하도록 해도 된다.
- [0407] 이 경우, 예를 들어 오브젝트수가 10 이상이라면 메쉬의 총 수를 8개로 하고, 오브젝트수가 10 미만이라면 메쉬의 총 수를 40개로 하거나 할 수 있다. 또한, 오브젝트수가 많을수록 메쉬의 총 수가 적어지도록, 오브젝트수에 따라서 다단계로 메쉬의 총 수가 변경되도록 해도 된다.
- [0408] 이렇게 오브젝트수에 따라서 메쉬의 총 수를 변경함으로써, 렌더러의 하드 규모에 따라서 처리량을 조정하여, 가능한 한 품질이 높은 음성을 얻을 수 있다.
- [0409] 또한, 오브젝트의 메타데이터에 포함되는 중요도 정보에 기초하여, 처리의 전환이 행해지는 경우, 다음과 같은 처리를 행하도록 할 수 있다.
- [0410] 예를 들어 오브젝트의 중요도 정보가 가장 높은 중요도를 나타내는 최고값일 경우에는, 종래대로 처리 A1 내지 처리 A3만이 행해지도록 하고, 오브젝트의 중요도 정보가 최고값 이외의 값일 경우에는, VBAP 계인에 대한 2치화 처리가 행해지도록 한다.
- [0411] 기타, 예를 들어 오브젝트의 중요도 정보의 값에 따라서 메쉬수 전환 처리를 행하고, 메쉬의 총 수를 적절하게 변경하도록 해도 된다. 이 경우, 오브젝트의 중요도가 높을수록, 메쉬의 총 수가 많아지게 하면 되고, 다단계로 메쉬의 총 수가 변경되도록 할 수 있다.
- [0412] 이들 예에서는, 각 오브젝트의 중요도 정보에 기초하여, 오브젝트마다 처리를 전환할 수 있다. 여기서 설명한 처리에서는, 중요도가 높은 오브젝트에 대해서는 음질이 높아지도록 하고, 또한 중요도가 낮은 오브젝트에 대해서는 음질을 낮게 하여 처리량을 저감시키도록 할 수 있다. 따라서, 여러가지 중요도의 오브젝트의 음성을 동시에 재생하는 경우에, 가장 청감상의 음질 열화를 억제하여 처리량을 적게 할 수 있어, 음질의 확보와 처리량 삭감의 균형이 잡힌 방법이라고 할 수 있다.
- [0413] 이와 같이, 오브젝트의 중요도 정보에 기초하여 오브젝트마다 처리의 전환을 행하는 경우, 중요도가 높은 오브젝트일수록 메쉬의 총 수가 많아지도록 하거나, 오브젝트의 중요도가 높을 때에는 양자화 처리를 행하지 않도록 하거나 할 수 있다.
- [0414] 또한, 이것에 추가로 중요도가 낮은 오브젝트, 즉 중요도 정보의 값이 소정값 미만인 오브젝트에 대해서도, 중요도가 높은, 즉 중요도 정보의 값이 소정값 이상인 오브젝트에 가까운 위치에 있는 오브젝트일수록, 메쉬의 총 수가 많아지도록 하거나, 양자화 처리를 행하지 않도록 하거나 하는 등 해도 된다.
- [0415] 구체적으로는, 중요도 정보가 최고값인 오브젝트에 대해서는 메쉬의 총 수가 40개가 되게 되고, 중요도 정보가 최고값이 아닌 오브젝트에 대해서는, 메쉬의 총 수가 적어지게 되는 것으로 한다.

[0416] 이 경우, 중요도 정보가 최고값이 아닌 오브젝트에 대해서는, 그 오브젝트와, 중요도 정보가 최고값인 오브젝트의 거리가 짧을수록, 메쉬의 총 수가 많아지게 하면 된다. 통상, 유저는 중요도가 높은 오브젝트의 소리를 특히 주의하여 듣기 때문에, 그 오브젝트의 근처에 있는 다른 오브젝트의 소리의 음질이 낮으면, 유저는 콘텐츠 전체의 음질이 좋지 않은 것 같이 느끼게 된다. 그래서, 중요도가 높은 오브젝트에 가까운 위치에 있는 오브젝트에 대해서도, 가능한 한 좋은 음질이 되도록 메쉬의 총 수를 정함으로써 청감 상의 음질의 열화를 억제할 수 있다.

[0417] 또한, 오브젝트의 오디오 신호의 음압에 따라서 처리를 전환하게 해도 된다. 여기서, 오디오 신호의 음압은, 오디오 신호의 렌더링 대상을 포함하는 프레임 내의 각 샘플의 샘플값의 2승 평균값의 평방근을 계산함으로써 구할 수 있다. 즉, 음압 RMS는 다음 식 (10)의 계산에 의해 구할 수 있다.

$$\text{RMS} = 20 \times \log_{10} \left(\sqrt{\frac{1}{N} \sum_{n=0}^{N-1} (X_n)^2} \right) \quad \dots (10)$$

[0418]

[0419] 또한, 식 (10)에 있어서 N은 오디오 신호의 프레임을 구성하는 샘플의 수를 나타내고 있고, x_n 은 프레임 내의 n 번째(단, $n=0, \dots, N-1$)의 샘플의 샘플값을 나타내고 있다.

[0420] 이와 같이 하여 얻어지는 오디오 신호의 음압 RMS에 따라서 처리를 전환하는 경우, 다음과 같은 처리를 행하도록 할 수 있다.

[0421] 예를 들어 음압 RMS의 폴스케일인 0dB에 대하여 오브젝트의 오디오 신호의 음압 RMS가 -6dB 이상인 경우에는, 종래대로 처리 A1 내지 처리 A3만이 행해지도록 하고, 오브젝트의 음압 RMS가 -6dB 미만인 경우에는, VBAP 게인에 대한 2치화 처리가 행해지도록 한다.

[0422] 일반적으로, 음압이 큰 음성은 음질의 열화가 두드러지기 쉽고, 또한, 그러한 음성은 중요도가 높은 오브젝트의 음성인 경우가 많다. 그래서, 여기에서는 음압 RMS가 큰 음성의 오브젝트에 대해서는 음질이 열화되지 않도록 하고, 음압 RMS가 작은 음성의 오브젝트에 대해서 2치화 처리를 행하여, 전체적으로 처리량을 삭감하도록 하였다. 이에 의해, 하드 규모가 작은 렌더러로도 충분히 렌더링을 행할 수 있고, 또한 가능한 한 품질이 높은 음성을 얻을 수 있다.

[0423] 또한, 오브젝트의 오디오 신호의 음압 RMS에 따라서 메쉬수 전환 처리를 행하고, 메쉬의 총 수를 적절하게 변경하도록 해도 된다. 이 경우, 예를 들어 음압 RMS가 큰 오브젝트일수록, 메쉬의 총 수가 많아지게 하면 되고, 다단계로 메쉬의 총 수가 변경되도록 할 수 있다.

[0424] 또한, 오브젝트수, 중요도 정보, 및 음압 RMS에 따라, 양자화 처리나 메쉬수 전환 처리의 조합을 선택하도록 해도 된다.

[0425] 즉, 오브젝트수, 중요도 정보, 및 음압 RMS에 기초하여, 양자화 처리를 행할지 여부, 양자화 처리에 있어서 VBAP 게인을 몇개의 게인으로 양자화할지, 즉 양자화 처리 시에 있어서의 양자화수, 및 VBAP 게인의 산출에 사용하는 메쉬의 총 수를 선택하고, 그 선택 결과에 따른 처리에 의해 VBAP 게인을 산출해도 된다. 그러한 경우, 예를 들어 다음과 같은 처리를 행하도록 할 수 있다.

[0426] 예를 들어 오브젝트수가 10 이상인 경우, 모든 오브젝트에 대해서, 메쉬의 총 수가 10개가 되도록 하고, 또한 2치화 처리가 행해지도록 한다. 이 경우, 오브젝트수가 많으므로, 메쉬의 총 수를 적게 함과 함께 2치화 처리를 행하도록 함으로써 처리량을 저감시킨다. 이에 의해, 렌더러의 하드 규모가 작은 경우에도 모든 오브젝트의 렌더링을 행할 수 있게 된다.

[0427] 또한, 오브젝트수가 10 미만이고, 또한 중요도 정보의 값이 최고값일 경우에는, 종래대로 처리 A1 내지 처리 A3만이 행해지도록 한다. 이에 의해, 중요도가 높은 오브젝트에 대해서는 음질을 열화시키지 않고 음성을 재생할 수 있다.

[0428] 오브젝트수가 10 미만이고, 또한 중요도 정보의 값이 최고값이 아니고, 또한 음압 RMS가 -30dB 이상인 경우에는, 메쉬의 총 수가 10개가 되도록 하고, 또한 3치화 처리가 행해지도록 한다. 이에 의해, 중요도는 낮지만 음압이 큰 음성에 대해서, 음성의 음질 열화가 눈에 띄지 않을 정도로 렌더링 처리 시의 처리량을 저감시킬 수 있다.

- [0429] 또한, 오브젝트수가 10 미만이고, 또한 중요도 정보의 값이 최고값이 아니고, 또한 음압 RMS가 -30dB 미만인 경우에는, 메쉬의 총 수가 5개가 되도록 하고, 또한 2치화 처리가 행해지도록 한다. 이에 의해, 중요도가 낮고 음압도 작은 음성에 대해서, 렌더링 처리 시의 처리량을 충분히 저감시킬 수 있다.
- [0430] 이렇게 오브젝트수가 많을 때에는 렌더링 처리의 처리량을 적게 하여 전체 오브젝트의 렌더링을 행할 수 있도록 하고, 오브젝트수가 어느 정도 적은 경우에는, 오브젝트마다 적절한 처리를 선택하고, 렌더링을 행하도록 한다. 이에 의해, 오브젝트마다 음질의 확보와 처리량 삭감의 균형을 잡으면서, 전체적으로 적은 처리량으로 충분한 음질로 음성을 재생할 수 있다.
- [0431] <음성 처리 장치의 구성예>
- [0432] 이어서, 이상에 있어서 설명한 양자화 처리나 메쉬수 전환 처리 등을 적절히 행하면서 렌더링 처리를 행하는 음성 처리 장치에 대하여 설명한다. 도 17은, 그러한 음성 처리 장치의 구체적인 구성예를 도시하는 도면이다. 또한, 도 17에 있어서 도 6에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 적절히 생략한다.
- [0433] 도 17에 도시하는 음성 처리 장치(61)는 취득부(21), 게인 산출부(23), 및 게인 조정부(71)를 갖고 있다. 게인 산출부(23)는 취득부(21)로부터 오브젝트의 메타데이터와 오디오 신호의 공급을 받고, 각 오브젝트에 대하여 스피커(12)마다의 VBAP 게인을 산출하고, 게인 조정부(71)에 공급한다.
- [0434] 또한, 게인 산출부(23)는 VBAP 게인의 양자화를 행하는 양자화부(31) 구비하고 있다.
- [0435] 게인 조정부(71)는 각 오브젝트에 대해서, 게인 산출부(23)로부터 공급된 스피커(12)마다의 VBAP 게인을, 취득부(21)로부터 공급된 오디오 신호에 승산함으로써, 스피커(12)마다의 오디오 신호를 생성하고, 스피커(12)에 공급한다.
- [0436] <재생 처리의 설명>
- [0437] 계속해서, 도 17에 도시된 음성 처리 장치(61)의 동작에 대하여 설명한다. 즉, 도 18의 흐름도를 참조하여, 음성 처리 장치(61)에 의한 재생 처리에 대하여 설명한다.
- [0438] 또한, 이 예에서는, 취득부(21)에는, 1개 또는 복수의 오브젝트에 대해서, 오브젝트의 오디오 신호와 메타데이터가 프레임마다 공급되고, 재생 처리는, 각 오브젝트에 대하여 오디오 신호의 프레임마다 행해지는 것으로 한다.
- [0439] 스텝 S231에 있어서, 취득부(21)는 외부로부터 오브젝트의 오디오 신호 및 메타데이터를 취득하고, 오디오 신호를 게인 산출부(23) 및 게인 조정부(71)에 공급함과 함께, 메타데이터를 게인 산출부(23)에 공급한다. 또한, 취득부(21)는 처리 대상으로 되어 있는 프레임에서 동시에 음성을 재생하는 오브젝트의 수, 즉 오브젝트수를 나타내는 정보도 취득하여 게인 산출부(23)에 공급한다.
- [0440] 스텝 S232에 있어서, 게인 산출부(23)는 취득부(21)로부터 공급된 오브젝트수를 나타내는 정보에 기초하여, 오브젝트수가 10 이상인지 여부를 판정한다.
- [0441] 스텝 S232에 있어서 오브젝트수가 10 이상이라고 판정된 경우, 스텝 S233에 있어서, 게인 산출부(23)는 VBAP 게인 산출 시에 사용하는 메쉬의 총 수를 10으로 한다. 즉, 게인 산출부(23)는 메쉬의 총 수로서 10을 선택한다.
- [0442] 또한, 게인 산출부(23)는 선택한 메쉬의 총 수에 따라, 그 총 수만큼 단위 구 표면 상에 메쉬가 형성되도록, 전체 스피커(12) 중에서, 소정 개수의 스피커(12)를 선택한다. 그리고, 게인 산출부(23)는 선택한 스피커(12)로 형성되는 단위 구 표면 상의 10개의 메쉬를, VBAP 게인 산출 시에 사용하는 메쉬로 한다.
- [0443] 스텝 S234에 있어서, 게인 산출부(23)는 스텝 S233에 있어서 정해진 10개의 메쉬를 구성하는 각 스피커(12)의 배치 위치를 나타내는 배치 위치 정보와, 취득부(21)로부터 공급된 메타데이터에 포함되는, 오브젝트의 위치를 나타내는 위치 정보에 기초하여, VBAP에 의해 각 스피커(12)의 VBAP 게인을 산출한다.
- [0444] 구체적으로는, 게인 산출부(23)는 스텝 S233에 있어서 정해진 메쉬를 차례로 처리 대상의 메쉬로서 식 (8)의 계산을 행해 감으로써, 각 스피커(12)의 VBAP 게인을 산출한다. 이때, 상술한 바와 같이, 처리 대상의 메쉬를 구성하는 3개의 스피커(12)에 대하여 산출된 VBAP 게인이 모두 0 이상의 값으로 될 때까지, 새로운 메쉬가 처리 대상의 메쉬로 되고, VBAP 게인이 산출되어 간다.
- [0445] 스텝 S235에 있어서, 양자화부(31)는 스텝 S234에서 얻어진 각 스피커(12)의 VBAP 게인을 2치화하고, 그 후, 처

리는 스텝 S246으로 진행한다.

- [0446] 또한, 스텝 S232에 있어서 오브젝트수가 10 미만이라고 판정된 경우, 처리는 스텝 S236으로 진행한다.
- [0447] 스텝 S236에 있어서, 게인 산출부(23)는 취득부(21)로부터 공급된 메타데이터에 포함되는 오브젝트의 중요도 정보의 값이 최고값인지 여부를 판정한다. 예를 들어 중요도 정보의 값이, 가장 중요도가 높은 것을 나타내는 수치 「7」 일 경우, 중요도 정보가 최고값이라고 판정된다.
- [0448] 스텝 S236에 있어서 중요도 정보가 최고값이라고 판정된 경우, 처리는 스텝 S237로 진행한다.
- [0449] 스텝 S237에 있어서, 게인 산출부(23)는 각 스피커(12)의 배치 위치를 나타내는 배치 위치 정보와, 취득부(21)로부터 공급된 메타데이터에 포함되는 위치 정보에 기초하여, 각 스피커(12)의 VBAP 게인을 산출하고, 그 후, 처리는 스텝 S246으로 진행한다. 여기에서는, 모든 스피커(12)로 형성되는 메쉬가 차례로 처리 대상의 메쉬로 되어 가고, 식 (8)의 계산에 의해 VBAP 게인이 산출된다.
- [0450] 이에 반해, 스텝 S236에 있어서 중요도 정보가 최고값이 아니라고 판정된 경우, 스텝 S238에 있어서, 게인 산출부(23)는 취득부(21)로부터 공급된 오디오 신호의 음압 RMS를 산출한다. 구체적으로는, 처리 대상으로 되어 있는 오디오 신호의 프레임에 대해서, 상술한 식 (10)의 계산이 행해지고, 음압 RMS가 산출된다.
- [0451] 스텝 S239에 있어서, 게인 산출부(23)는 스텝 S238에서 산출한 음압 RMS가 -30dB 이상인지 여부를 판정한다.
- [0452] 스텝 S239에 있어서, 음압 RMS가 -30dB 이상이라고 판정된 경우, 그 후, 스텝 S240 및 스텝 S241의 처리가 행해진다. 또한, 이들 스텝 S240 및 스텝 S241의 처리는, 스텝 S233 및 스텝 S234의 처리와 동일하므로, 그 설명은 생략한다.
- [0453] 스텝 S242에 있어서, 양자화부(31)는 스텝 S241에서 얻어진 각 스피커(12)의 VBAP 게인을 3치화하고, 그 후, 처리는 스텝 S246으로 진행한다.
- [0454] 또한, 스텝 S239에 있어서 음압 RMS가 -30dB 미만이라고 판정된 경우, 처리는 스텝 S243으로 진행한다.
- [0455] 스텝 S243에 있어서, 게인 산출부(23)는 VBAP 게인 산출 시에 사용하는 메쉬의 총 수를 5로 한다.
- [0456] 또한, 게인 산출부(23)는 선택한 메쉬의 총 수 「5」에 따라, 전체 스피커(12) 중에서, 소정 개수의 스피커(12)를 선택하고, 선택한 스피커(12)로 형성되는 단위 구 표면 상의 5개의 메쉬를, VBAP 게인 산출 시에 사용하는 메쉬로 한다.
- [0457] VBAP 게인 산출 시에 사용하는 메쉬가 정해지면, 그 후, 스텝 S244 및 스텝 S245의 처리가 행해져서 처리는 스텝 S246으로 진행한다. 또한, 이들 스텝 S244 및 스텝 S245의 처리는, 스텝 S234 및 스텝 S235의 처리와 동일하므로, 그 설명은 생략한다.
- [0458] 스텝 S235, 스텝 S237, 스텝 S242, 또는 스텝 S245의 처리가 행해져서, 각 스피커(12)의 VBAP 게인이 얻어지면, 그 후, 스텝 S246 내지 스텝 S248의 처리가 행해져서 재생 처리는 종료한다.
- [0459] 또한, 이들 스텝 S246 내지 스텝 S248의 처리는, 도 7을 참조하여 설명한 스텝 S17 내지 스텝 S19의 처리와 동일하므로, 그 설명은 생략한다.
- [0460] 단, 보다 상세하게는, 재생 처리는 각 오브젝트에 대하여 대략 동시에 행해지고, 스텝 S248에서는, 오브젝트마다 얻어진 각 스피커(12)의 오디오 신호가, 그들 스피커(12)에 공급된다. 즉, 스피커(12)에서는, 각 오브젝트의 오디오 신호를 가산하여 얻어진 신호에 기초하여 음성이 재생된다. 그 결과, 전체 오브젝트의 음성이 동시에 출력되게 된다.
- [0461] 이상과 같이 하여 음성 처리 장치(61)는 오브젝트마다, 적절히, 양자화 처리나 메쉬수 전환 처리를 선택적으로 행한다. 이렇게 함으로써, 임장감이나 음질의 열화를 억제하면서 렌더링 처리의 처리량을 저감시킬 수 있다.
- [0462] <제2 실시 형태의 변형예 1>
- [0463] <음성 처리 장치의 구성예>
- [0464] 또한, 제2 실시 형태에서는, 음상을 확장하는 처리를 행하지 않는 경우에 양자화 처리나 메쉬수 전환 처리를 선택적으로 행하는 예에 대하여 설명했지만, 음상을 확장하는 처리를 행하는 경우에도 양자화 처리나 메쉬수 전환 처리를 선택적으로 행하게 해도 된다.

- [0465] 그러한 경우, 음성 처리 장치(11)는 예를 들어 도 19에 도시하는 바와 같이 구성된다. 또한, 도 19에 있어서, 도 6 또는 도 17에 있어서의 경우와 대응하는 부분에는 동일한 부호를 부여하고 있고, 그 설명은 적절히 생략한다.
- [0466] 도 19에 도시하는 음성 처리 장치(11)는 취득부(21), 벡터 산출부(22), 게인 산출부(23), 및 게인 조정부(71)를 갖고 있다.
- [0467] 취득부(21)는 1개 또는 복수의 오브젝트에 대해서, 오브젝트의 오디오 신호와 메타데이터를 취득하고, 취득한 오디오 신호를 게인 산출부(23) 및 게인 조정부(71)에 공급함과 함께, 취득한 메타데이터를 벡터 산출부(22) 및 게인 산출부(23)에 공급한다. 또한, 게인 산출부(23)는 양자화부(31)를 구비하고 있다.
- [0468] <재생 처리의 설명>
- [0469] 이어서, 도 20의 흐름도를 참조하여, 도 19에 도시된 음성 처리 장치(11)에 의해 행해지는 재생 처리에 대하여 설명한다.
- [0470] 또한, 이 예에서는, 취득부(21)에는, 1개 또는 복수의 오브젝트에 대해서, 오브젝트의 오디오 신호와 메타데이터가 프레임마다 공급되고, 재생 처리는, 각 오브젝트에 대하여 오디오 신호의 프레임마다 행해지는 것으로 한다.
- [0471] 또한, 스텝 S271 및 스텝 S272의 처리는 도 7의 스텝 S11 및 스텝 S12의 처리와 동일하므로, 그 설명은 생략한다. 단, 스텝 S271에서는, 취득부(21)에 의해 취득된 오디오 신호는 게인 산출부(23) 및 게인 조정부(71)에 공급되고, 취득부(21)에 의해 취득된 메타데이터는, 벡터 산출부(22) 및 게인 산출부(23)에 공급된다.
- [0472] 이들 스텝 S271 및 스텝 S272의 처리가 행해지면, spread 벡터, 또는 spread 벡터 및 벡터 p가 얻어진다.
- [0473] 스텝 S273에 있어서, 게인 산출부(23)는 VBAP 게인 산출 처리를 행하여 스피커(12)마다 VBAP 게인을 산출한다. 또한, VBAP 게인 산출 처리의 상세에 대해서는 후술하는데, VBAP 게인 산출 처리에서는, 적절히, 양자화 처리나 메쉬수 전환 처리가 선택적으로 행해지고, 각 스피커(12)의 VBAP 게인이 산출된다.
- [0474] 스텝 S273의 처리가 행해져서 각 스피커(12)의 VBAP 게인이 얻어지면, 그 후, 스텝 S274 내지 스텝 S276의 처리가 행해져서 재생 처리는 종료하는데, 이들 처리는, 도 7의 스텝 S17 내지 스텝 S19의 처리와 동일하므로, 그 설명은 생략한다. 단, 보다 상세하게는, 재생 처리는 각 오브젝트에 대하여 대략 동시에 행해지고, 스텝 S276에서는, 오브젝트마다 얻어진 각 스피커(12)의 오디오 신호가, 그들 스피커(12)에 공급된다. 그로 인해, 스피커(12)에서는, 전체 오브젝트의 음성이 동시에 출력되게 된다.
- [0475] 이상과 같이 하여 음성 처리 장치(11)는 오브젝트마다, 적절히, 양자화 처리나 메쉬수 전환 처리를 선택적으로 행한다. 이렇게 함으로써, 음상을 확장하는 처리를 행하는 경우에 있어서도, 임장감이나 음질의 열화를 억제하면서 렌더링 처리의 처리량을 저감시킬 수 있다.
- [0476] <VBAP 게인 산출 처리의 설명>
- [0477] 계속해서, 도 21의 흐름도를 참조하여, 도 20의 스텝 S273의 처리에 대응하는 VBAP 게인 산출 처리에 대하여 설명한다.
- [0478] 또한, 스텝 S301 내지 스텝 S303의 처리는, 도 18의 스텝 S232 내지 스텝 S234의 처리와 동일하므로, 그 설명은 생략한다. 단, 스텝 S303에서는, spread 벡터, 또는 spread 벡터 및 벡터 p의 각 벡터에 대해서, 스피커(12)마다 VBAP 게인이 산출된다.
- [0479] 스텝 S304에 있어서, 게인 산출부(23)는 스피커(12)마다, 각 벡터에 대하여 산출한 VBAP 게인을 가산하고, VBAP 게인 가산값을 산출한다. 스텝 S304에서는, 도 7의 스텝 S14와 동일한 처리가 행해진다.
- [0480] 스텝 S305에 있어서, 양자화부(31)는 스텝 S304의 처리에 의해 스피커(12)마다 얻어진 VBAP 게인 가산값을 2치화하여 VBAP 게인 산출 처리는 종료되고, 그 후, 처리는 도 20의 스텝 S274로 진행된다.
- [0481] 또한, 스텝 S301에 있어서 오브젝트수가 10 미만이라고 판정된 경우, 스텝 S306 및 스텝 S307의 처리가 행해진다.
- [0482] 또한, 이들 스텝 S306 및 스텝 S307의 처리는, 도 18의 스텝 S236 및 스텝 S237의 처리와 동일하므로, 그 설명은 생략한다. 단, 스텝 S307에서는, spread 벡터, 또는 spread 벡터 및 벡터 p의 각 벡터에 대해서, 스피커

(12)마다 VBAP 게인이 산출된다.

- [0483] 또한, 스텝 S307의 처리가 행해지면, 스텝 S308의 처리가 행해져서 VBAP 게인 산출 처리는 종료되고, 그 후, 처리는 도 20의 스텝 S274로 진행하는데, 스텝 S308의 처리는 스텝 S304의 처리와 동일하므로, 그 설명은 생략한다.
- [0484] 또한, 스텝 S306에 있어서, 중요도 정보가 최고값이 아니라고 판정된 경우, 그 후, 스텝 S309 내지 스텝 S312의 처리가 행해지는데, 이들 처리는 도 18의 스텝 S238 내지 스텝 S241의 처리와 동일하므로, 그 설명은 생략한다. 단, 스텝 S312에서는, spread 벡터, 또는 spread 벡터 및 벡터 p의 각 벡터에 대해서, 스피커(12)마다 VBAP 게인이 산출된다.
- [0485] 이와 같이 하여, 각 벡터에 대하여 스피커(12)마다의 VBAP 게인이 얻어지면, 스텝 S313의 처리가 행해져서 VBAP 게인 가산값이 산출되는데, 스텝 S313의 처리는 스텝 S304의 처리와 동일하므로, 그 설명은 생략한다.
- [0486] 스텝 S314에 있어서, 양자화부(31)는 스텝 S313의 처리에 의해 스피커(12)마다 얻어진 VBAP 게인 가산값을 3치화하여 VBAP 게인 산출 처리는 종료되고, 그 후, 처리는 도 20의 스텝 S274로 진행한다.
- [0487] 또한, 스텝 S310에 있어서 음압 RMS가 -30dB 미만이라고 판정된 경우, 스텝 S315의 처리가 행해져서 VBAP 게인 산출 시에 사용하는 메쉬의 총 수가 5로 된다. 또한, 스텝 S315의 처리는, 도 18의 스텝 S243의 처리와 동일하므로, 그 설명은 생략한다.
- [0488] VBAP 게인 산출 시에 사용하는 메쉬가 정해지면, 스텝 S316 내지 스텝 S318의 처리가 행해져서 VBAP 게인 산출 처리는 종료되고, 그 후, 처리는 도 20의 스텝 S274로 진행한다. 또한, 이들 스텝 S316 내지 스텝 S318의 처리는, 스텝 S303 내지 스텝 S305의 처리와 동일하므로, 그 설명은 생략한다.
- [0489] 이상과 같이 하여 음성 처리 장치(11)는 오브젝트마다, 적절히, 양자화 처리나 메쉬수 전환 처리를 선택적으로 행한다. 이렇게 함으로써, 음상을 확장하는 처리를 행하는 경우에 있어서도, 임장감이나 음질의 열화를 억제하면서 렌더링 처리의 처리량을 저감시킬 수 있다.
- [0490] 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.
- [0491] 도 22는, 상술한 일련의 처리를 프로그램에 의해 실행하는 컴퓨터의 하드웨어 구성예를 도시하는 블록도이다.
- [0492] 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.
- [0493] 버스(504)에는, 또한, 입출력 인터페이스(505)가 접속되어 있다. 입출력 인터페이스(505)에는, 입력부(506), 출력부(507), 기록부(508), 통신부(509), 및 드라이브(510)가 접속되어 있다.
- [0494] 입력부(506)는 키보드, 마우스, 마이크로폰, 촬상 소자 등을 포함한다. 출력부(507)는 디스플레이, 스피커 등을 포함한다. 기록부(508)는 하드 디스크나 불휘발성이 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.
- [0495] 이상과 같이 구성되는 컴퓨터에서는, CPU(501)가, 예를 들어, 기록부(508)에 기록되어 있는 프로그램을, 입출력 인터페이스(505) 및 버스(504)를 통하여, RAM(503)에 로드하여 실행함으로써, 상술한 일련의 처리가 행해진다.
- [0496] 컴퓨터(CPU(501))가 실행하는 프로그램은, 예를 들어, 패키지 미디어 등으로서의 리무버블 기록 매체(511)에 기록하여 제공할 수 있다. 또한, 프로그램은, 로컬에어리어 네트워크, 인터넷, 디지털 위성 방송과 같은, 유선 또는 무선의 전송 매체를 통하여 제공할 수 있다.
- [0497] 컴퓨터에서는, 프로그램은, 리무버블 기록 매체(511)를 드라이브(510)에 장착함으로써, 입출력 인터페이스(505)를 통하여 기록부(508)에 인스톨할 수 있다. 또한, 프로그램은, 유선 또는 무선의 전송 매체를 통하여, 통신부(509)로 수신하고, 기록부(508)에 인스톨할 수 있다. 기타, 프로그램은, ROM(502)이나 기록부(508)에 미리 인스톨해 둘 수 있다.
- [0498] 또한, 컴퓨터가 실행하는 프로그램은, 본 명세서에서 설명하는 순서를 따라서 시계열로 처리가 행해지는 프로그램

램이어도 되고, 병렬로, 또는 호출이 행하여졌을 때 등의 필요한 타이밍에 처리가 행해지는 프로그래머도 된다.

- [0499] 또한, 본 기술의 실시 형태는, 상술한 실시 형태에 한정되는 것은 아니며, 본 기술의 요지를 일탈하지 않는 범위에서 다양한 변경이 가능하다.
- [0500] 예를 들어, 본 기술은, 하나의 기능을 네트워크를 통하여 복수의 장치에 분담, 공동하여 처리하는 클라우드 컴퓨팅의 구성을 취할 수 있다.
- [0501] 또한, 상술한 흐름도에서 설명한 각 스텝은, 하나의 장치로 실행하는 외에, 복수의 장치에 분담하여 실행할 수 있다.
- [0502] 또한, 하나의 스텝에 복수의 처리가 포함되는 경우에는, 그 하나의 스텝에 포함되는 복수의 처리는, 하나의 장치로 실행하는 외에, 복수의 장치에 분담하여 실행할 수 있다.
- [0503] 또한, 본 기술은, 이하의 구성으로 하는 것도 가능하다.
- [0504] (1)
- [0505] 오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하는 취득부와,
- [0506] 상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하는 벡터 산출부와,
- [0507] 상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 게인을 산출하는 게인 산출부를 구비하는 음성 처리 장치.
- [0509] (2)
- [0510] 상기 벡터 산출부는, 상기 수평 방향 각도와 상기 수직 방향 각도의 비에 기초하여, 상기 spread 벡터를 산출하는
- [0511] (1)에 기재된 음성 처리 장치.
- [0512] (3)
- [0513] 상기 벡터 산출부는, 미리 정해진 개수의 상기 spread 벡터를 산출하는
- [0514] (1) 또는 (2)에 기재된 음성 처리 장치.
- [0515] (4)
- [0516] 상기 벡터 산출부는, 가변인 임의의 개수의 상기 spread 벡터를 산출하는
- [0517] (1) 또는 (2)에 기재된 음성 처리 장치.
- [0518] (5)
- [0519] 상기 음상 정보는, 상기 영역의 중심 위치를 나타내는 벡터인
- [0520] (1)에 기재된 음성 처리 장치.
- [0521] (6)
- [0522] 상기 음상 정보는, 상기 영역의 중심으로부터의 음상의 범위 정도를 나타내는 2차원 이상의 벡터인
- [0523] (1)에 기재된 음성 처리 장치.
- [0524] (7)
- [0525] 상기 음상 정보는, 상기 위치 정보에 의해 나타나는 위치로부터 본 상기 영역의 중심 위치의 상대적인 위치를 나타내는 벡터인
- [0526] (1)에 기재된 음성 처리 장치.

- [0527] (8)
- [0528] 상기 계인 산출부는,
- [0529] 각 상기 음성 출력부에 대해서, 상기 spread 벡터마다 상기 계인을 산출하고,
- [0530] 상기 음성 출력부마다, 각 상기 spread 벡터에 대하여 산출한 상기 계인의 가산값을 산출하고,
- [0531] 상기 음성 출력부마다, 상기 가산값을 2차 이상의 계인으로 양자화하고,
- [0532] 상기 양자화된 상기 가산값에 기초하여, 상기 음성 출력부마다 최종적인 상기 계인을 산출하는
- [0533] (1) 내지 (7) 중 어느 한 항에 기재된 음성 처리 장치.
- [0534] (9)
- [0535] 상기 계인 산출부는, 3개의 상기 음성 출력부에 의해 둘러싸이는 영역인 메쉬이며, 상기 계인의 산출에 사용하는 메쉬의 수를 선택하고, 상기 메쉬의 수의 선택 결과와 상기 spread 벡터에 기초하여, 상기 spread 벡터마다 상기 계인을 산출하는
- [0536] (8)에 기재된 음성 처리 장치.
- [0537] (10)
- [0538] 상기 계인 산출부는, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화 시에 있어서의 상기 가산값의 양자화수를 선택하고, 그 선택 결과에 따라서 상기 최종적인 상기 계인을 산출하는
- [0539] (9)에 기재된 음성 처리 장치.
- [0540] (11)
- [0541] 상기 계인 산출부는, 상기 오디오 오브젝트의 수에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택하는
- [0542] (10)에 기재된 음성 처리 장치.
- [0543] (12)
- [0544] 상기 계인 산출부는, 상기 오디오 오브젝트의 중요도에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택하는
- [0545] (10) 또는 (11)에 기재된 음성 처리 장치.
- [0546] (13)
- [0547] 상기 계인 산출부는, 상기 중요도가 높은 상기 오디오 오브젝트에 가까운 위치에 있는 상기 오디오 오브젝트일 수록, 상기 계인의 산출에 사용하는 상기 메쉬의 수가 많아지도록, 상기 계인의 산출에 사용하는 상기 메쉬의 수를 선택하는
- [0548] (12)에 기재된 음성 처리 장치.
- [0549] (14)
- [0550] 상기 계인 산출부는, 상기 오디오 오브젝트의 오디오 신호의 음압에 기초하여, 상기 계인의 산출에 사용하는 상기 메쉬의 수, 상기 양자화를 행할지 여부, 및 상기 양자화수를 선택하는
- [0551] (10) 내지 (13) 중 어느 한 항에 기재된 음성 처리 장치.
- [0552] (15)
- [0553] 상기 계인 산출부는, 상기 메쉬의 수의 선택 결과에 따라, 복수의 상기 음성 출력부 중, 서로 다른 높이에 위치하는 상기 음성 출력부를 포함하는 3 이상의 상기 음성 출력부를 선택하고, 선택한 상기 음성 출력부로 형성되는 1개 또는 복수의 상기 메쉬에 기초하여 상기 계인을 산출하는
- [0554] (9) 내지 (14) 중 어느 한 항에 기재된 음성 처리 장치.

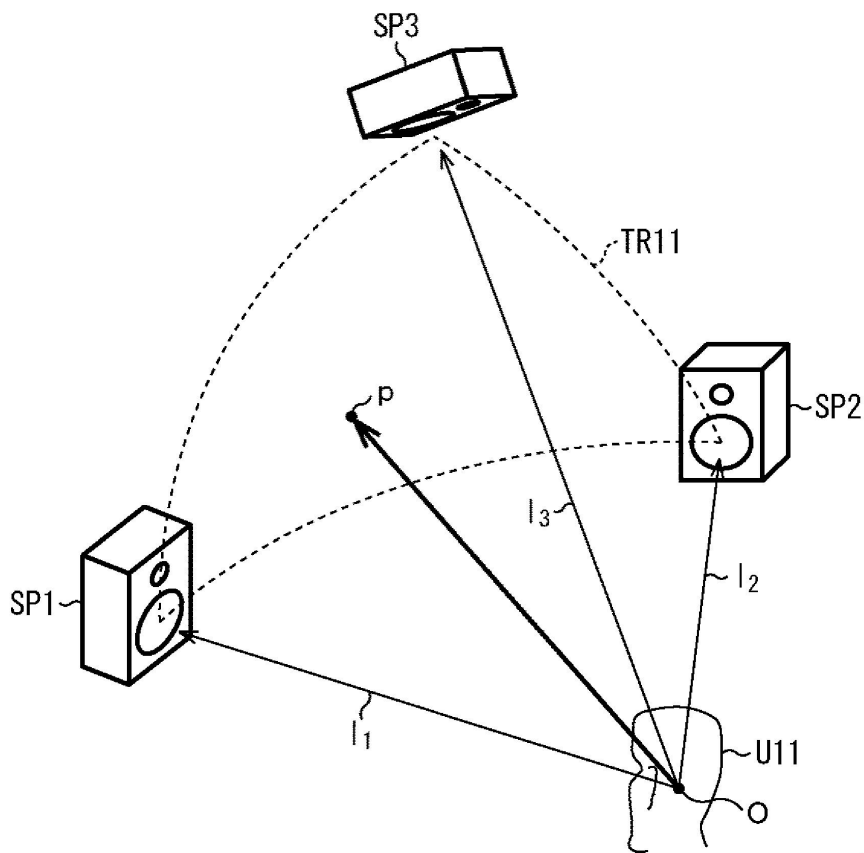
- [0555] (16)
- [0556] 오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하고,
- [0557] 상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하고,
- [0558] 상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 계인을 산출하는
- [0559] 스텝을 포함하는 음성 처리 방법.
- [0560] (17)
- [0561] 오디오 오브젝트의 위치를 나타내는 위치 정보와, 적어도 2차원 이상의 벡터를 포함하는, 상기 위치로부터의 음상의 범위를 나타내는 음상 정보를 포함하는 메타데이터를 취득하고,
- [0562] 상기 음상 정보에 의해 정해지는 음상의 범위를 나타내는 영역에 관한 수평 방향 각도 및 수직 방향 각도에 기초하여, 상기 영역 내의 위치를 나타내는 spread 벡터를 산출하고,
- [0563] 상기 spread 벡터에 기초하여, 상기 위치 정보에 의해 나타나는 상기 위치 근방에 위치하는 2 이상의 음성 출력부에 공급되는 오디오 신호의 각각의 계인을 산출하는
- [0564] 스텝을 포함하는 처리를 컴퓨터에 실행시키는 프로그램.
- [0565] (18)
- [0566] 오디오 오브젝트의 위치를 나타내는 위치 정보를 포함하는 메타데이터를 취득하는 취득부와,
- [0567] 3개의 음성 출력부에 의해 둘러싸이는 영역인 메쉬이며, 상기 음성 출력부에 공급되는 오디오 신호의 계인 산출에 사용하는 메쉬의 수를 선택하고, 상기 메쉬의 수의 선택 결과와 상기 위치 정보에 기초하여, 상기 계인을 산출하는 계인 산출부
- [0568] 를 구비하는 음성 처리 장치.

부호의 설명

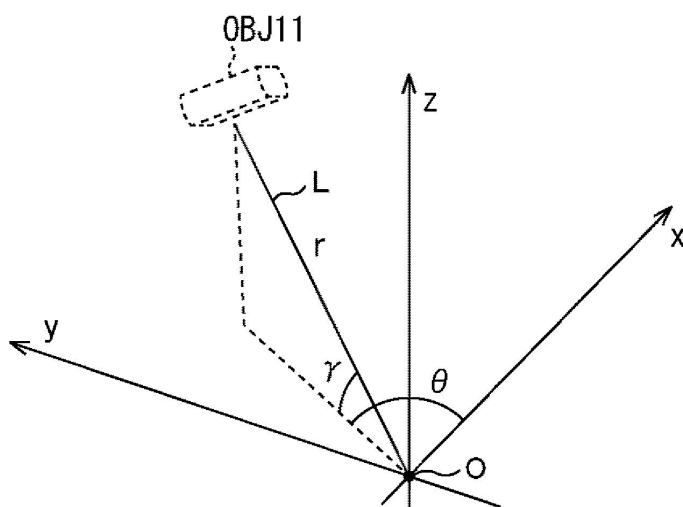
- [0569] 11: 음성 처리 장치
- 21: 취득부
- 22: 벡터 산출부
- 23: 계인 산출부
- 24: 계인 조정부
- 31: 양자화부
- 61: 음성 처리 장치
- 71: 계인 조정부

도면

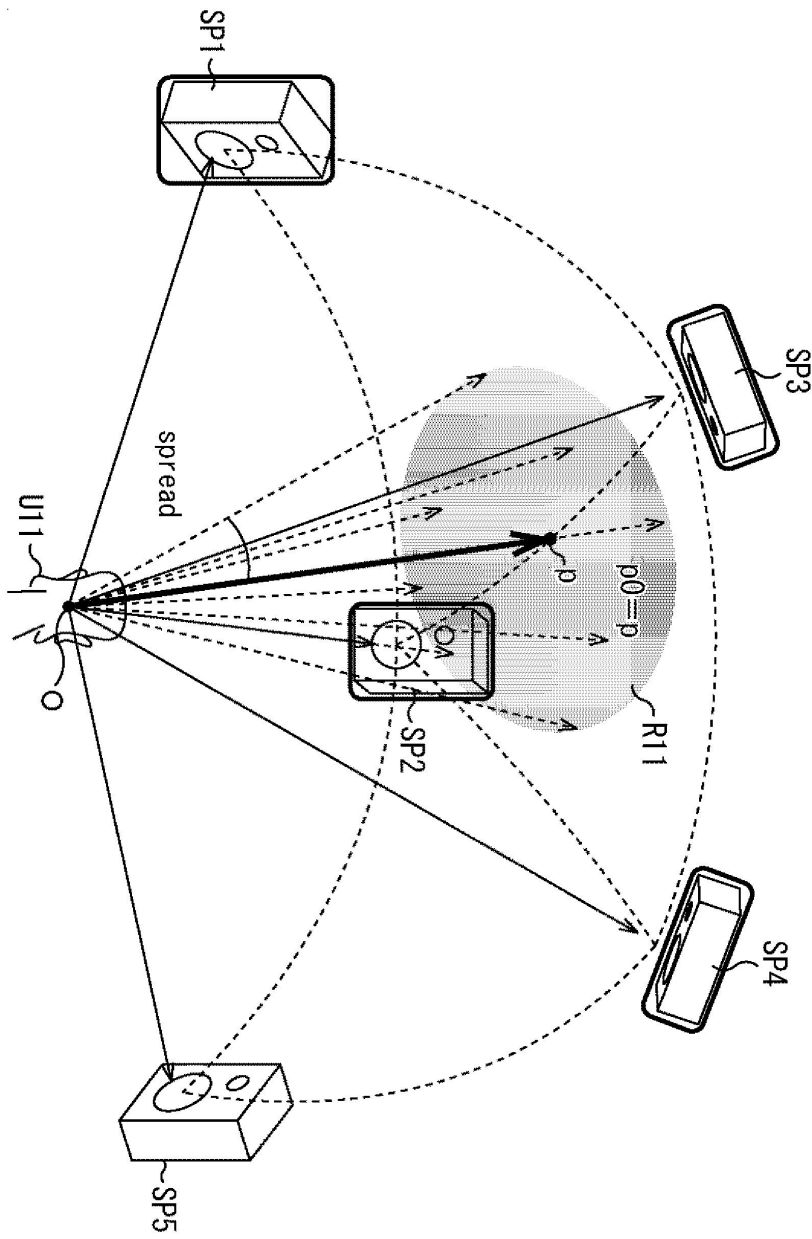
도면1



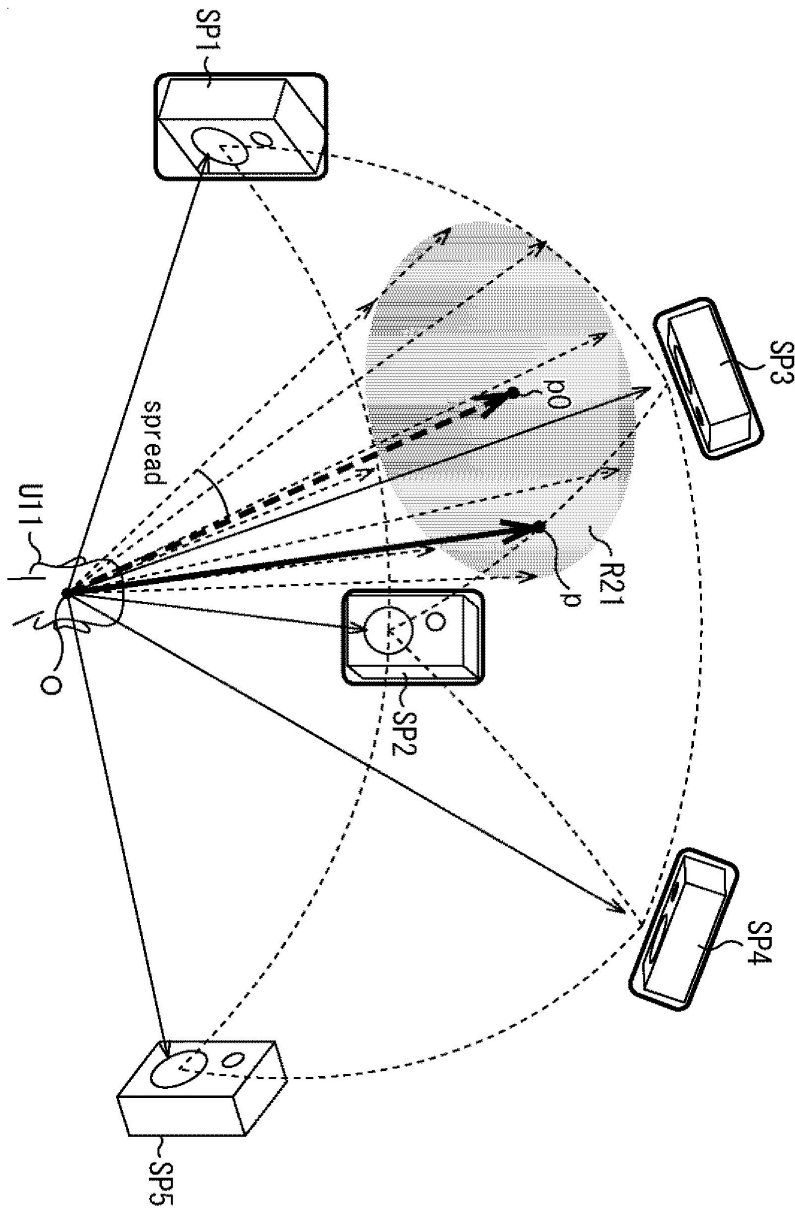
도면2



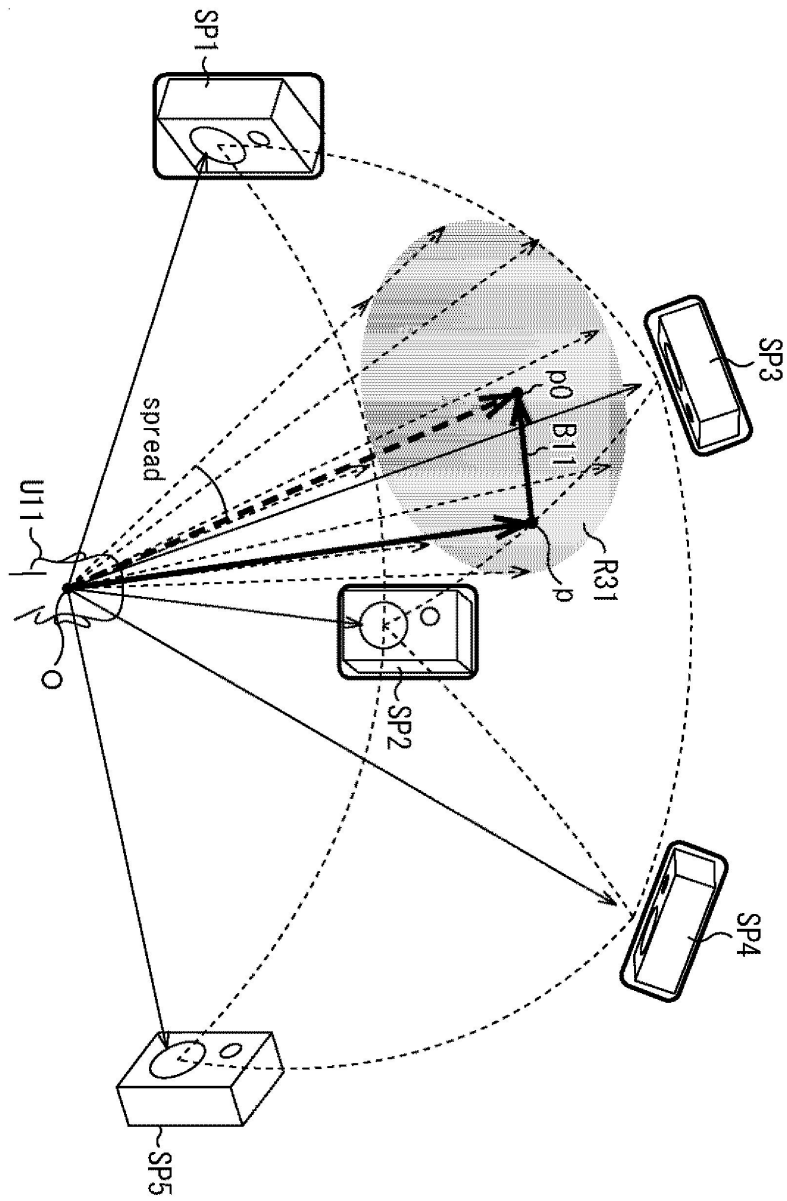
도면3



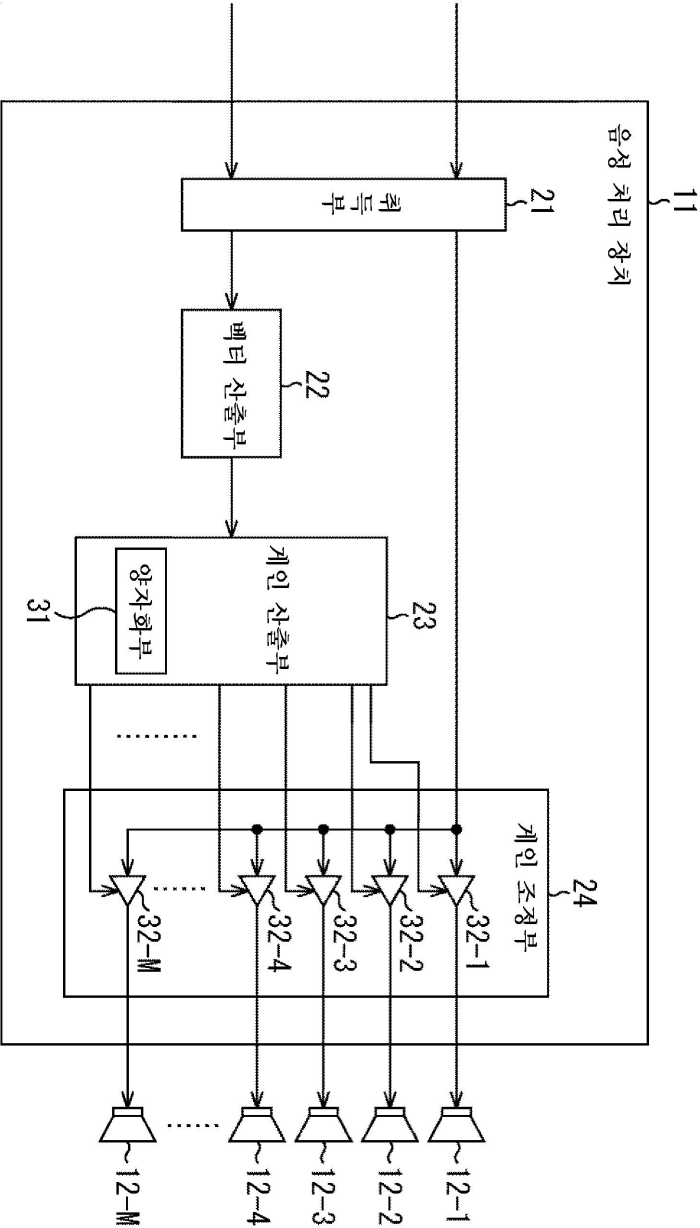
도면4



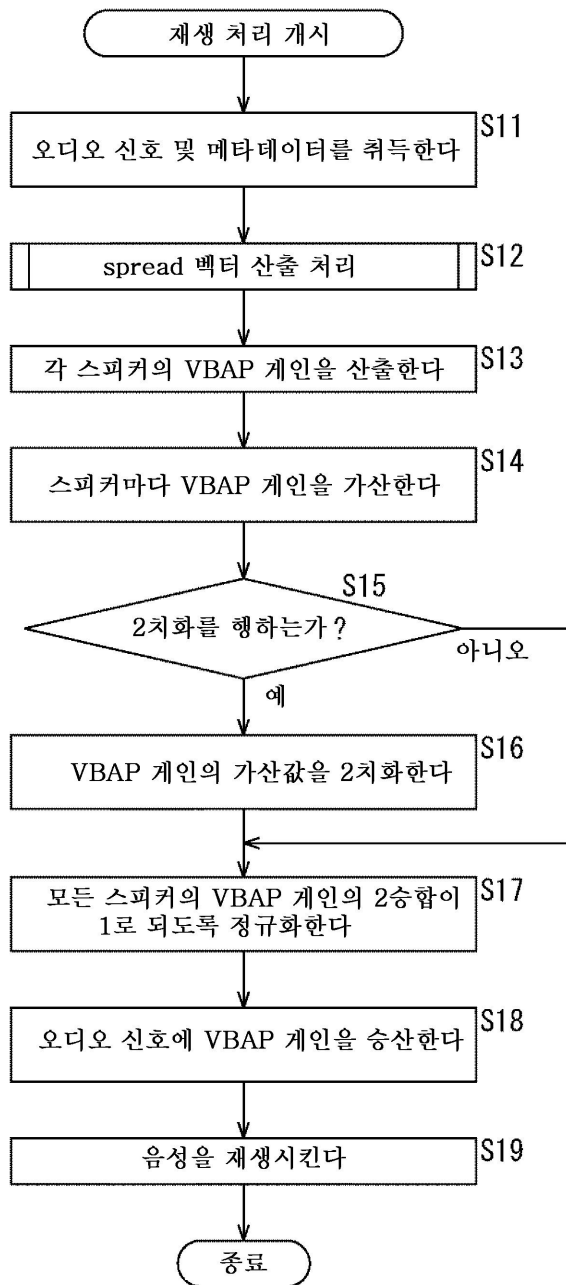
도면5



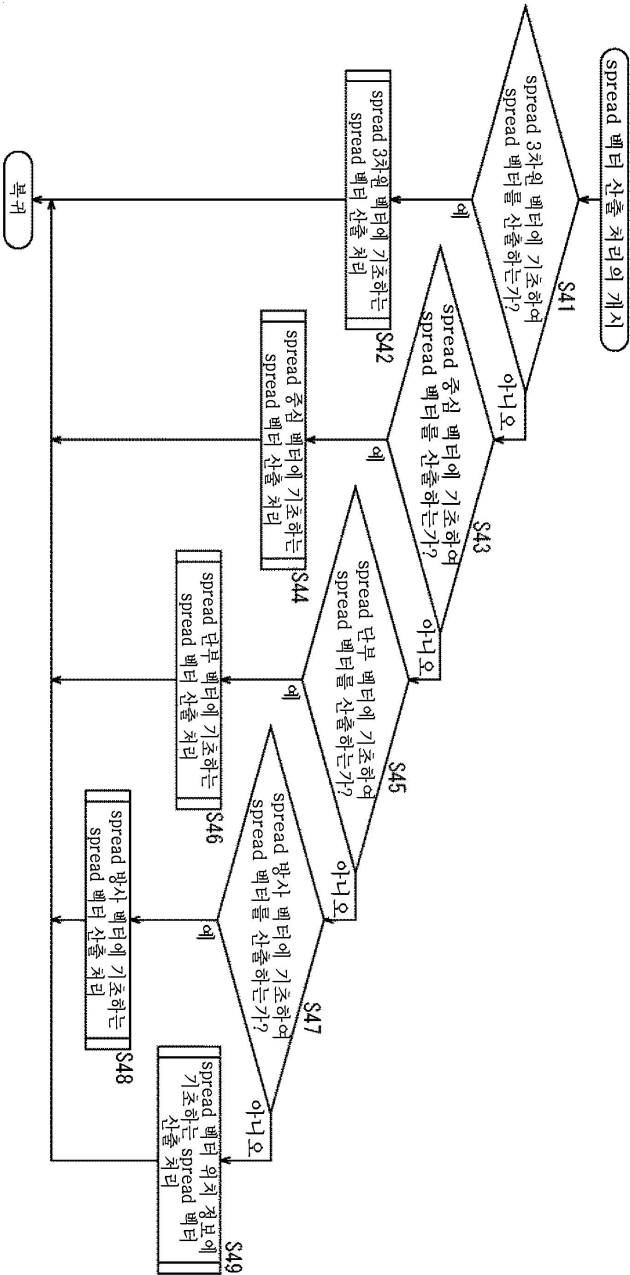
도면6



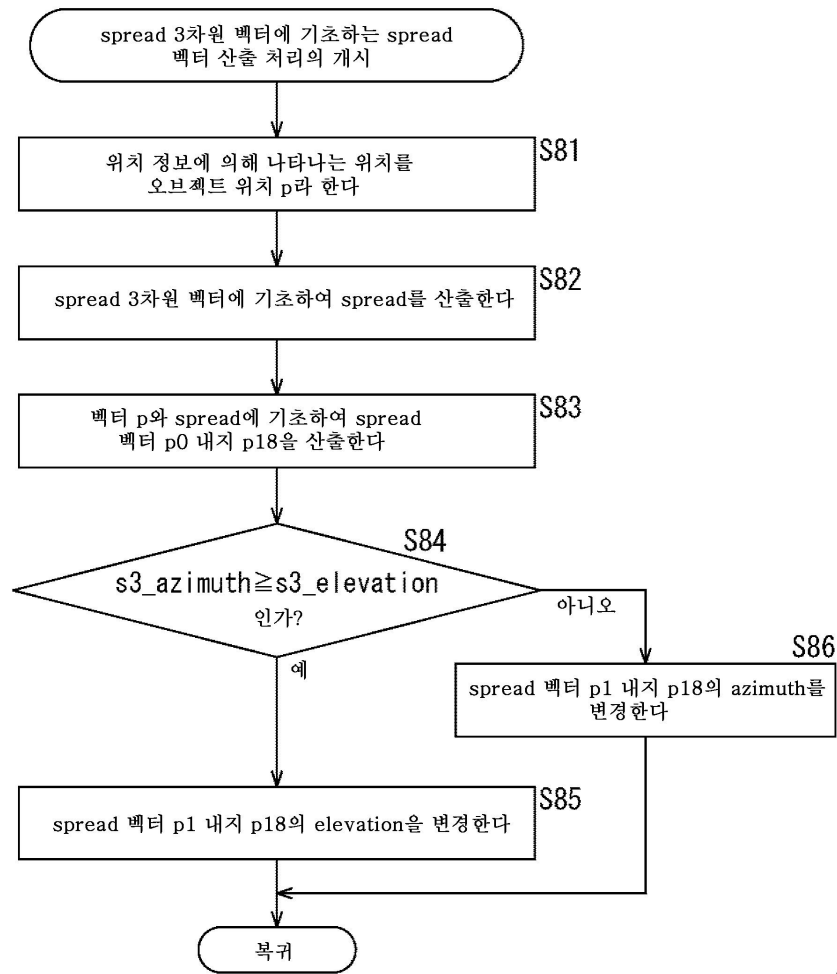
도면7



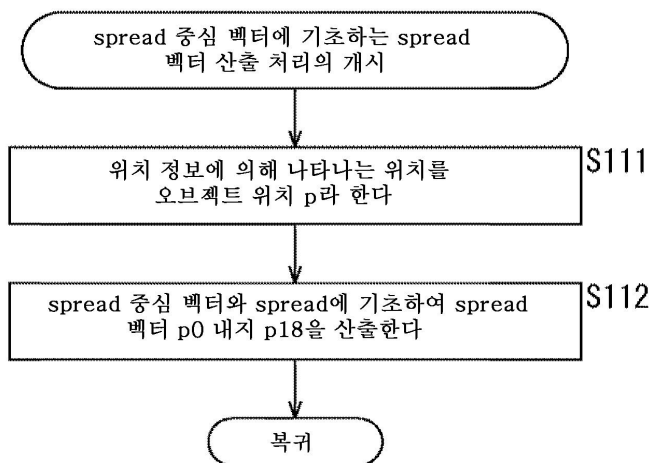
도면8



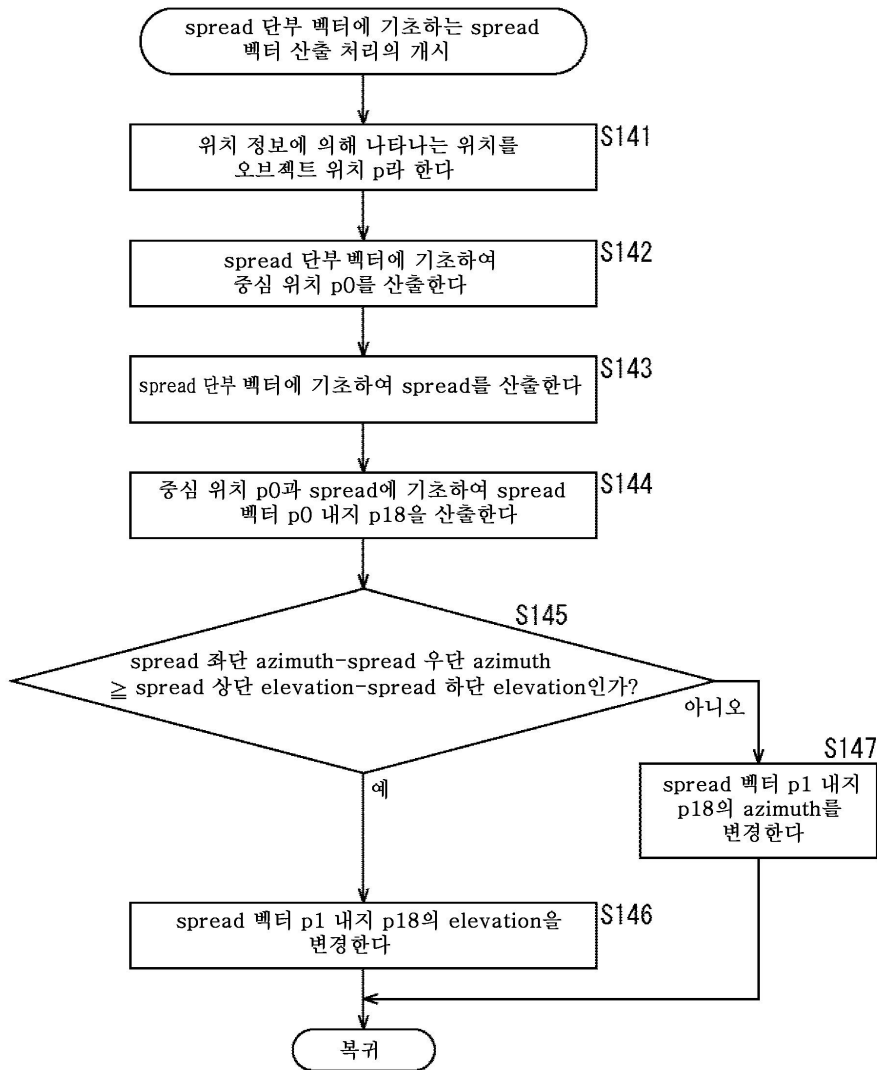
도면9



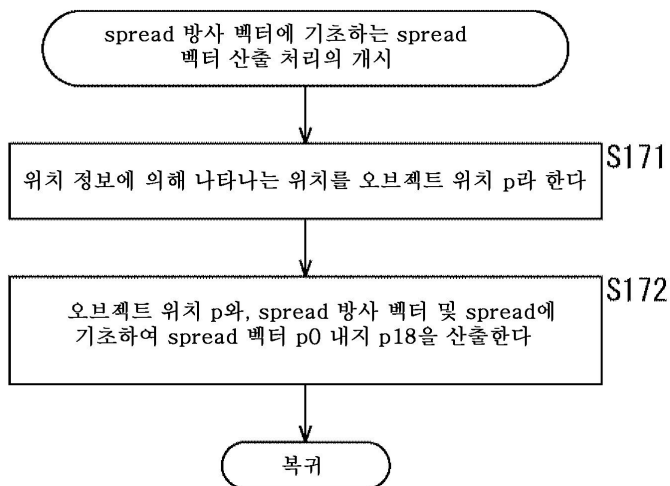
도면10



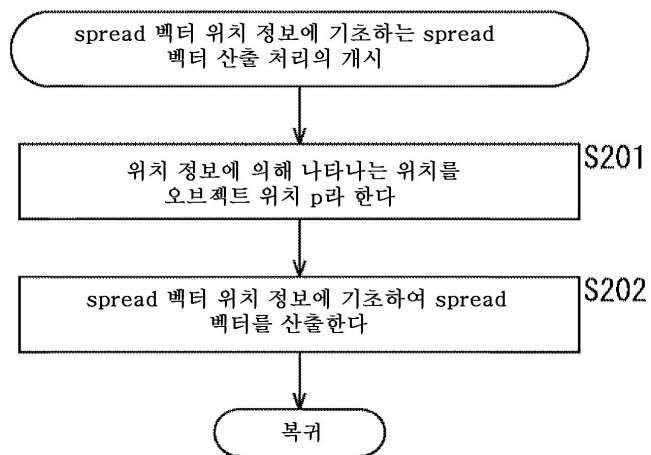
도면11



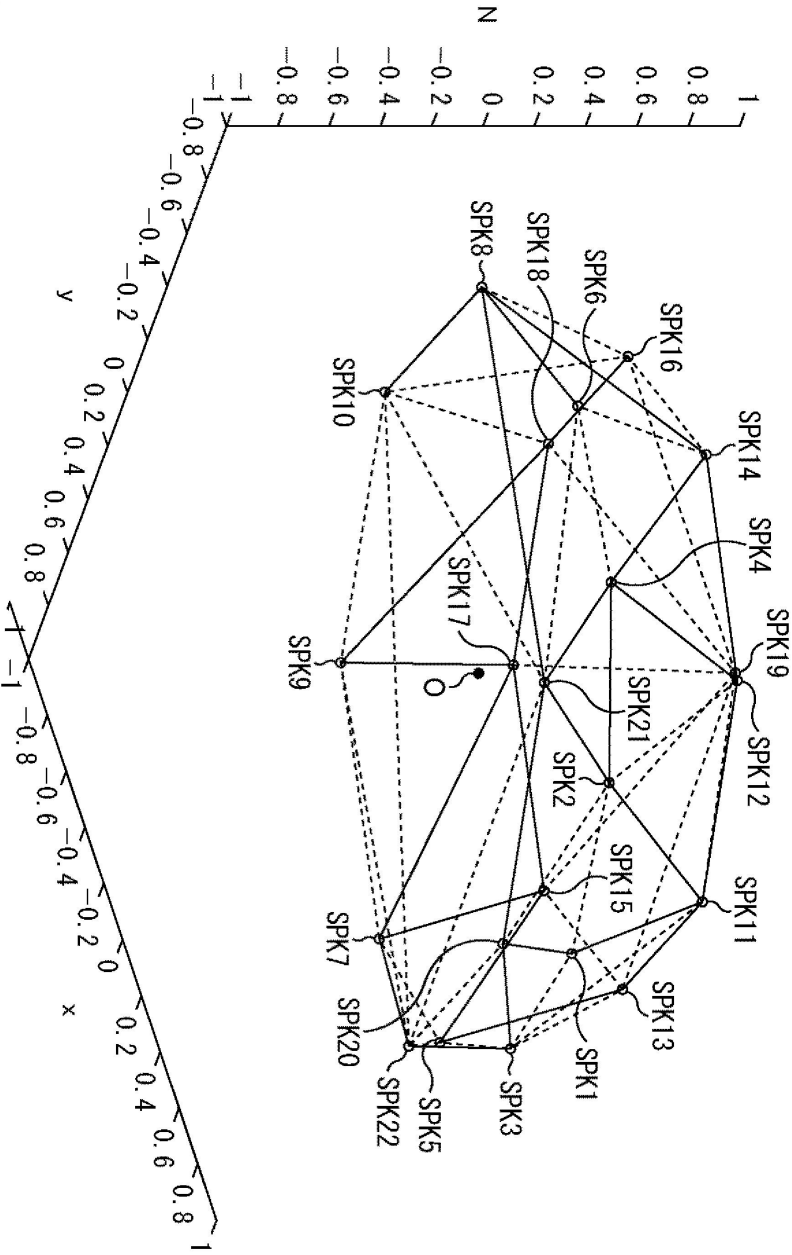
도면12



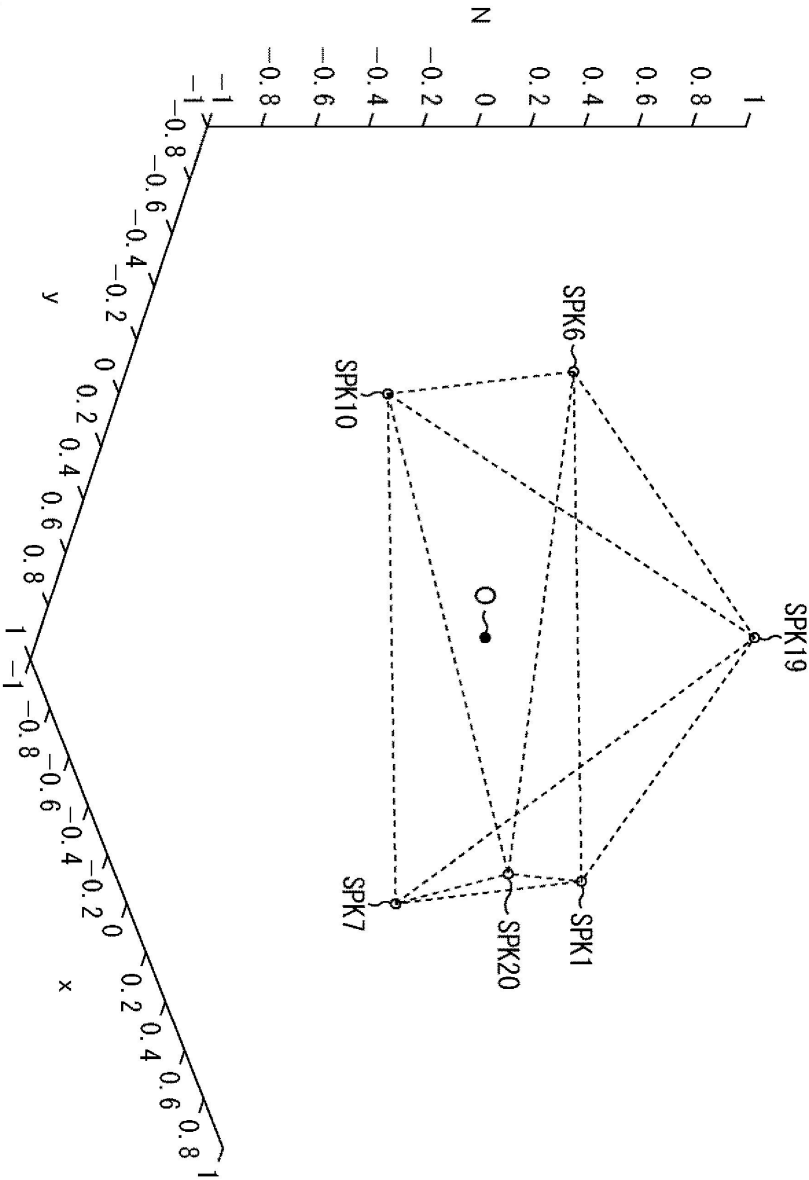
도면13



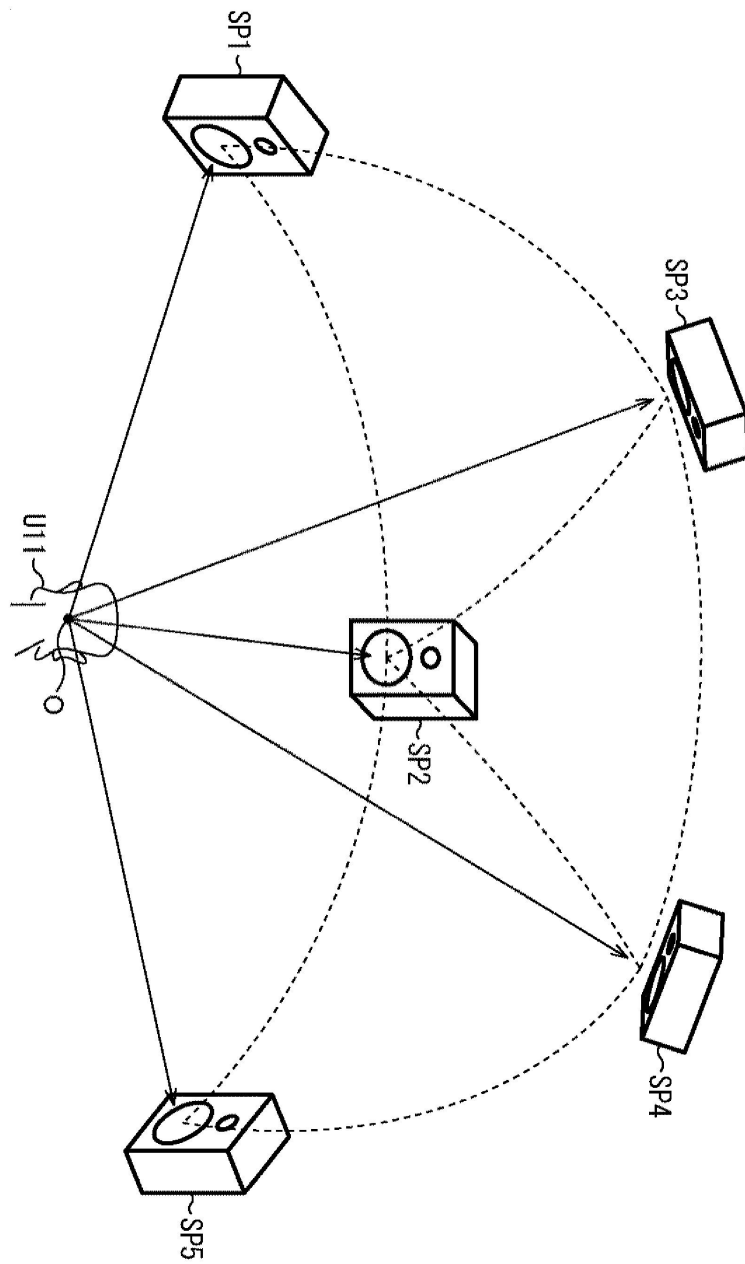
도면14



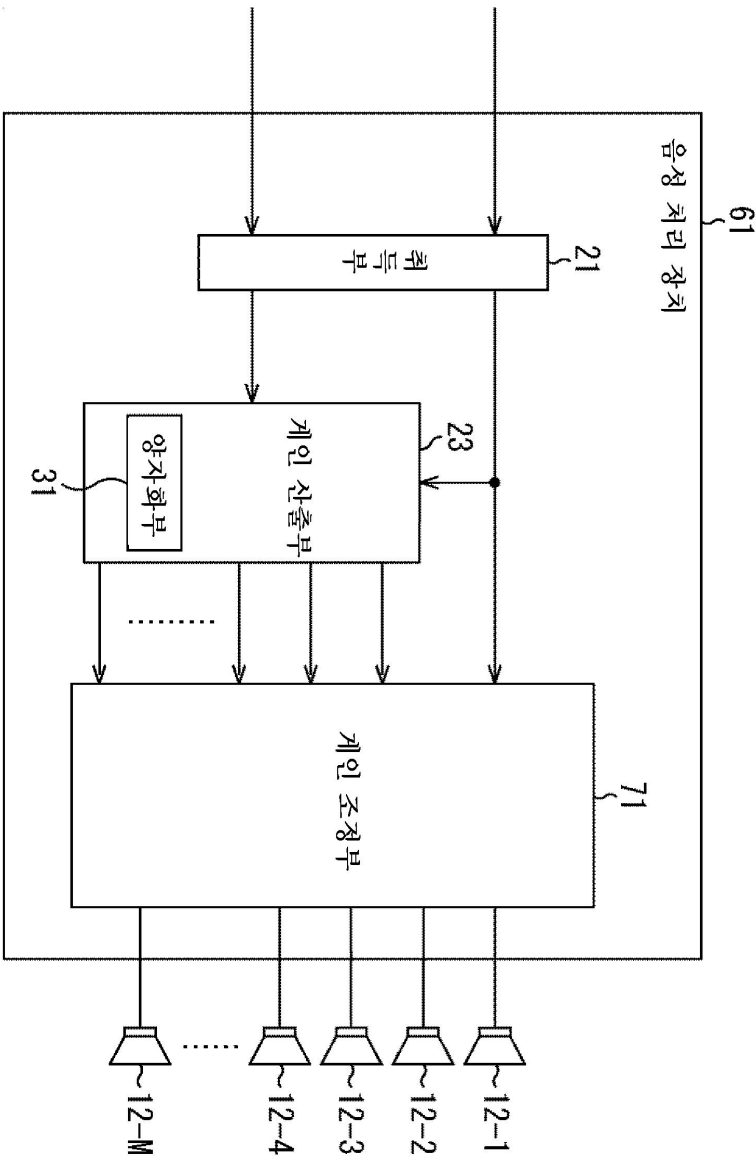
도면15



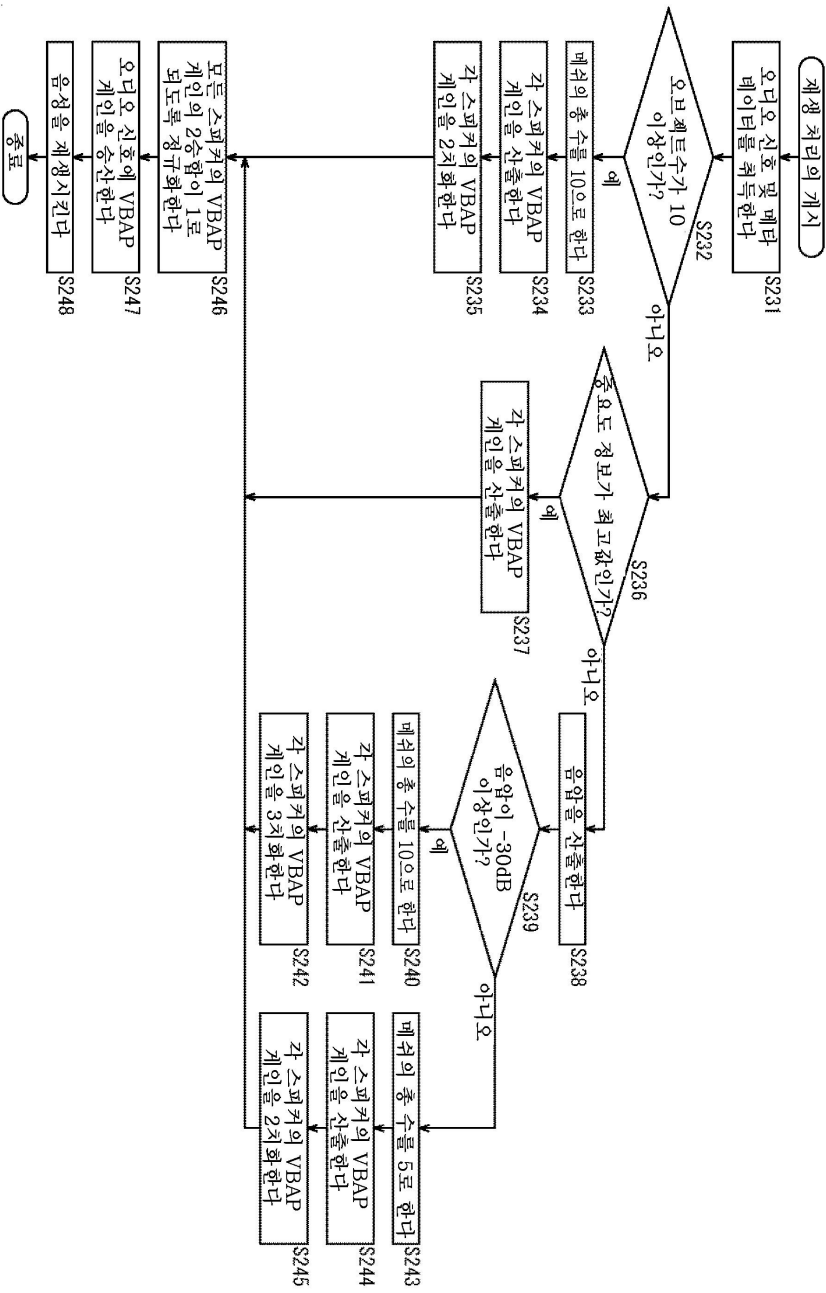
도면16



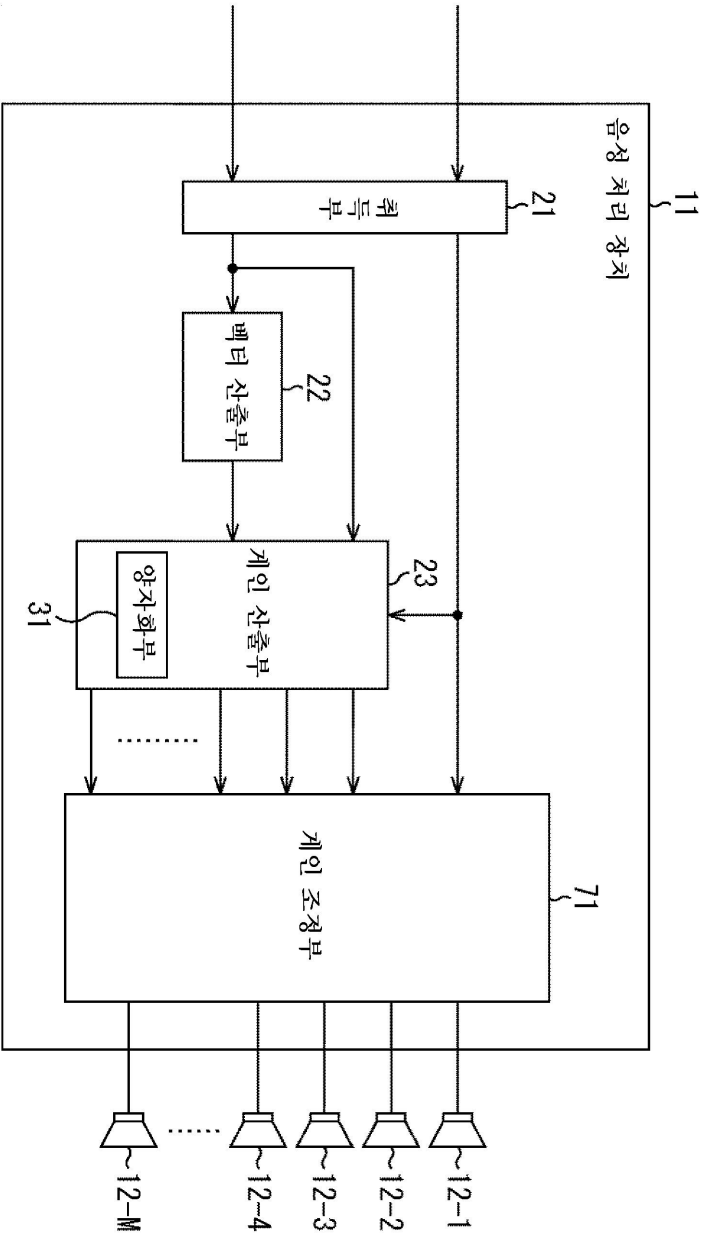
도면17



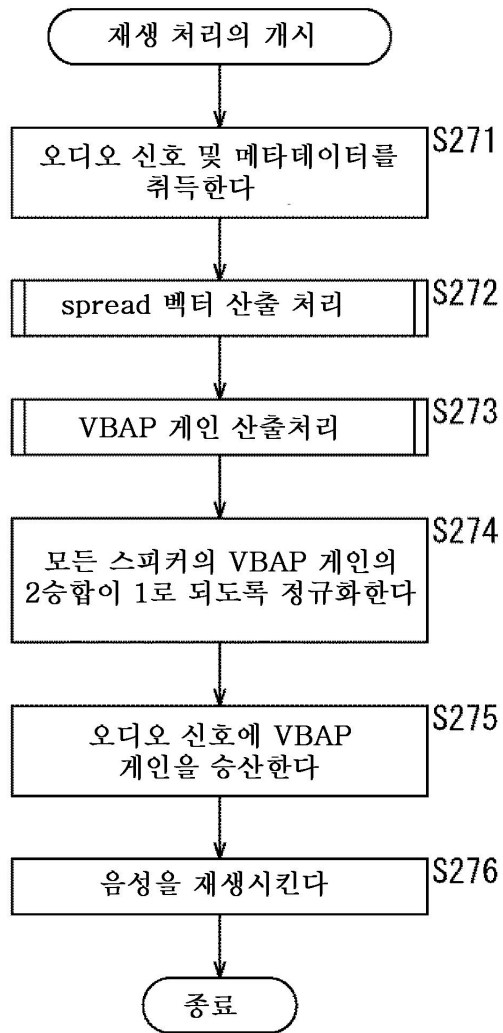
도면18



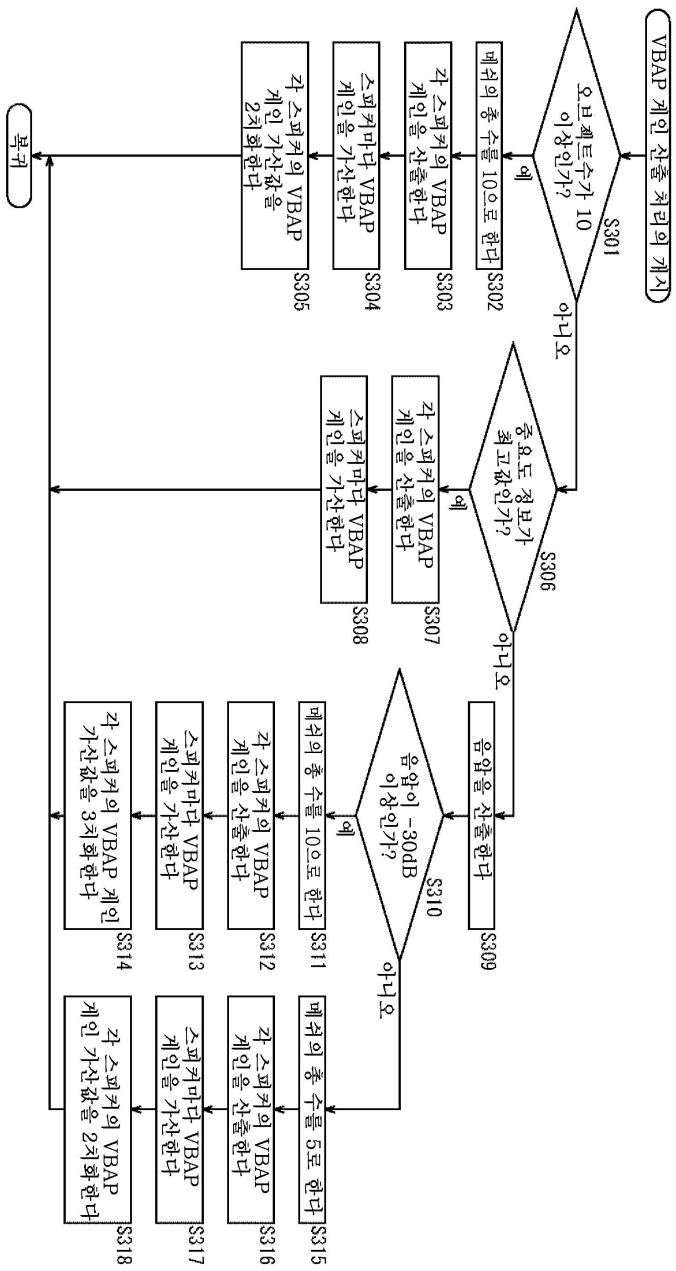
도면19



도면20



도면21



도면22

