



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년11월09일
(11) 등록번호 10-2323529
(24) 등록일자 2021년11월02일

(51) 국제특허분류(Int. Cl.)
H04S 7/00 (2006.01) H04S 3/00 (2006.01)
(52) CPC특허분류
H04S 7/30 (2020.05)
H04S 3/002 (2013.01)
(21) 출원번호 10-2018-0163511
(22) 출원일자 2018년12월17일
심사청구일자 2020년09월23일
(65) 공개번호 10-2020-0074757
(43) 공개일자 2020년06월25일
(56) 선행기술조사문헌
KR1020170024584 A

(73) 특허권자
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
유재현
대전광역시 유성구 죽동로 251, 305동 2303호 (죽동, 푸르지오아파트)
이용주
대전광역시 서구 도안동로 183, 1508동 1404호 (도안동, 대전 도안 아이파크)
(74) 대리인
특허법인 무한
(뒷면에 계속)

전체 청구항 수 : 총 19 항

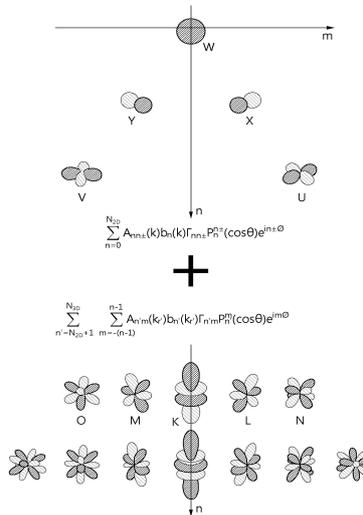
심사관 : 권영학

(54) 발명의 명칭 복합 차수 앰비소닉을 이용한 오디오 신호 처리 방법 및 장치

(57) 요약

본 발명은 오디오 신호 처리 방법으로서, VR 환경에서 사용자가 위치하는 임의의 가상 공간에서 재현되는 VR 콘텐츠에 따른 음장을 재현하기 위해, 미리 설정된 기준거리 이내에서 제공되는 음원과 상기 기준거리 밖에서 제공되는 음원을 식별하는 단계; 상기 미리 설정된 기준거리 이내에서 제공되는 음원의 높이 정보와 수평 정보를 고려한 3D 앰비소닉(ambisonics)을 이용하여 음장(sound field)을 처리하는 단계; 상기 기준거리 밖에서 제공되는 음원의 수평 정보를 고려한 2D 앰비소닉을 이용하여 음장을 처리하는 단계를 포함할 수 있다.

대표도 - 도4



- (52) CPC특허분류
H04S 2420/11 (2013.01)
H04S 2420/13 (2013.01)

장대영

대전광역시 유성구 장대로71번길 34, 110동 601호
 (장대동, 장대푸르지오아파트)

- (72) 발명자

이태진

대전광역시 유성구 진잠로149번길 24, 103동 1502호
 (교촌동, 제이파크아파트)

이 발명을 지원한 국가연구개발사업

과제고유번호	18ZR1100
부처명	과학기술정보통신부
과제관리(전문)기관명	한국전자통신연구원
연구사업명	정부출연금사업(기관고유사업)
연구과제명	초실감 공간미디어 원천기술 개발
기여율	1/1
과제수행기관명	한국전자통신연구원
연구기간	2018.01.01 ~ 2018.12.31

명세서

청구범위

청구항 1

VR 환경에서 사용자가 위치하는 임의의 가상 공간에서 재현되는 VR 콘텐츠에 따른 음장을 재현하기 위해, 미리 설정된 기준거리 이내에서 제공되는 음원과 상기 기준거리 밖에서 제공되는 음원을 식별하는 단계;

상기 미리 설정된 기준거리 이내에서 제공되는 음원의 높이 정보와 수평 정보를 고려한 3D 앰비소닉(ambisonics)을 이용하여 음장(sound field)을 처리하는 단계;

상기 기준거리 밖에서 제공되는 음원의 수평 정보를 고려한 2D 앰비소닉을 이용하여 음장을 처리하는 단계를 포함하는, 오디오 신호 처리 방법.

청구항 2

제1항에 있어서,

상기 3D 앰비소닉을 이용하여 처리된 음장과 상기 2D 앰비소닉을 이용하여 처리된 음장을 조합하여, 상기 임의의 가상 공간에 위치한 사용자에게 재현되는 상기 VR 콘텐츠에 적합한 음장을 재현하는 단계

를 더 포함하는, 오디오 신호 처리 방법.

청구항 3

제1항에 있어서,

상기 기준거리는,

높이 정보보다 수평 정보에 민감한 인간의 청취 특성을 고려하여 설정되는, 오디오 신호 처리 방법.

청구항 4

제1항에 있어서,

상기 앰비소닉은 차수 n 과 degree m 을 포함하는 구면 조화(spherical harmonics)로 표현되는 방향성분 정보를 이용하여 음장을 처리하는 방식을 나타내는, 오디오 신호 처리 방법.

청구항 5

제1항에 있어서,

상기 기준거리 이외의 경우 차수(order) $n=0 \sim N_{2D}$ 인 2D 앰비소닉을 이용하며, 상기 기준거리 이내의 경우 차수 $n=N_{2D}+1 \sim N_{3D}$ 인 3D 앰비소닉을 이용하는, 오디오 신호 처리 방법.

청구항 6

제5항에 있어서,

상기 기준거리 이외의 경우 degree $m=\pm n$ 을 이용하며, 상기 기준거리 이내의 경우 degree $m=\pm n$ 을 제외한

degree $m=-(n-1) \sim (n-1)$ 를 이용하는, 오디오 신호 처리 방법.

청구항 7

제1항에 있어서,

상기 앰비소닉은,

푸리에 계수 $A_{nm}(k)$, 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)인 방사 함수(radial function)를 나타내는 $b_n(k)$, 정규화 상수(normalization constant)인 Γ_{nm} , 버금 르장드르 함수(Associated Legendre function)인 $P_n^m(x)$, 방위 조화(azimuthal harmonics)인 $e^{im\phi}$ 를 포함하는, 오디오 신호 처리 방법

청구항 8

제7항에 있어서,

상기 푸리에 계수 $A_{nm}(k)$ 는 자유음장(free field)에서 획득된 평면파를 통해 결정될 수 있고, 상기 $b_n(k)$ 는 자유음장(free field)의 음원 입사와 마이크로폰으로 인한 스캐터링(scattering)을 반영하는, 오디오 신호 처리 방법.

청구항 9

제7항에 있어서,

상기 방사 함수(radial function) $b_n(k)$ 로 사용되는 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)는 상기 임의의 가상 공간에 위치한 사용자로부터 근거리에서의 크기가 원거리에서의 크기보다 큰, 오디오 신호 처리 방법.

청구항 10

하드웨어와 결합되어 제1항 내지 제9항 중 어느 하나의 항의 방법을 실행시키기 위하여 매체에 저장된 컴퓨터 프로그램.

청구항 11

프로세서 및 컴퓨터에서 읽을 수 있는 명령어를 포함하는 메모리를 포함하고,

상기 명령어가 상기 프로세서에서 실행되면, 상기 프로세서는 VR 환경에서 사용자가 위치하는 임의의 가상 공간에서 재현되는 VR 콘텐츠에 따른 음장을 재현하기 위해, 미리 설정된 기준거리 이내에서 제공되는 음원과 상기 기준거리 밖에서 제공되는 음원을 식별하고,

상기 미리 설정된 기준거리 이내에서 제공되는 음원의 높이 정보와 수평 정보를 고려한 3D 앰비소닉(ambisonics)을 이용하여 음장(sound field)을 처리하고,

상기 기준거리 밖에서 제공되는 음원의 수평 정보를 고려한 2D 앰비소닉을 이용하여 음장을 처리하는, 오디오 신호 처리 장치.

청구항 12

제11항에 있어서,

상기 프로세서는,

상기 3D 앰비소닉을 이용하여 처리된 음장과 상기 2D 앰비소닉을 이용하여 처리된 음장을 조합하여, 상기 임의의 가상 공간에 위치한 사용자에게 재현되는 상기 VR 콘텐츠에 적합한 음장을 재현하는, 오디오 신호 처리 장치.

청구항 13

제11항에 있어서,

상기 기준거리는,

높이 정보보다 수평 정보에 민감한 인간의 청취 특성을 고려하여 설정되는, 오디오 신호 처리 장치.

청구항 14

제11항에 있어서,

상기 앰비소닉은 차수 n 과 degree m 을 포함하는 구면 조화(spherical harmonics)로 표현되는 방향성분 정보를 이용하여 음장을 처리하는 방식을 나타내는, 오디오 신호 처리 장치.

청구항 15

제11항에 있어서,

상기 기준거리 이외의 경우 차수(order) $n=0 \sim N_{2D}$ 인 2D 앰비소닉을 이용하며, 상기 기준거리 이내의 경우 차수 $n=N_{2D}+1 \sim N_{3D}$ 인 3D 앰비소닉을 이용하는, 오디오 신호 처리 장치.

청구항 16

제15항에 있어서,

상기 기준거리 이외의 경우 degree $m=\pm n$ 을 이용하며, 상기 기준거리 이내의 경우 degree $m=\pm n$ 을 제외한 degree $m=-(n-1) \sim (n-1)$ 를 이용하는, 오디오 신호 처리 장치.

청구항 17

제11항에 있어서,

상기 앰비소닉은,

푸리에 계수 $A_{nm}(k)$, 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)인 방사 함수(radial function)를 나타내는 $b_n(k)$, 정규화 상수(normalization constant)인 Γ_{nm} , 버금 르장드르 함수(Associated Legendre function)인 $P_n^m(x)$, 방위 조화(azimuthal harmonics)인 $e^{im\phi}$ 를 포함하는, 오디오 신호 처리 장치.

청구항 18

제17항에 있어서,

상기 푸리에 계수 $A_m(k)$ 는 자유음장(free field)에서 획득된 평면파를 통해 결정될 수 있고, 상기 $b_n(k)$ 는 자유음장(free field)의 음원 입사와 마이크로폰으로 인한 스퀀터링(scattering)을 반영하는, 오디오 신호 처리 장치.

청구항 19

제17항에 있어서,

상기 방사 함수(radial function) $b_n(k)$ 로 사용되는 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)는 상기 임의의 가상 공간에 위치한 사용자로부터 근거리에서의 크기가 원거리에서의 크기보다 큰, 오디오 신호 처리 장치.

발명의 설명

기술 분야

[0001] 아래 실시예들은 복합 차수 앰비소닉(COA)를 이용한 오디오 신호 처리를 통해 유효 청취 영역을 제공하는 기술에 관한 것이다.

배경 기술

[0003] 오디오 재생 환경은 모노에서 스테레오, 5.1 채널을 지나 10.2 채널, 22.2 채널과 같은 사운드 시스템으로 발전하고 있다. 또한, 콘텐츠 재생 환경도 TV나 영화관에서 스마트폰과 HMD를 이용한 VR 환경으로 발전하고 있다.

[0004] 특히 6 dof VR 환경에서 사용자가 가상 공간을 임의로 움직이며 다양한 인터랙션을 할 경우, 제공되는 VR 콘텐츠에 따른 적절한 오디오 렌더링 방식이 필요하다. 이때, 인터랙션을 반영한 VR 콘텐츠에 적합한 오디오를 제공할 때, 적은 연산량을 유지하면서 사용자에게 효과적인 유효 청취 영역(sweet spot)을 제공하는 기술이 필요하다.

발명의 내용

해결하려는 과제

[0006] 일 실시예에 따르면, 사용자의 위치가 자유롭게 변화하는 VR 환경에서 적은 연산량의 앰비소닉을 이용하여, 사용자에게 제공되는 VR 콘텐츠에 적합한 음장을 처리할 수 있다.

과제의 해결 수단

[0008] 일 측면에 따르면, VR 환경에서 사용자가 위치하는 임의의 가상 공간에서 재현되는 VR 콘텐츠에 따른 음장을 재현하기 위해, 미리 설정된 기준거리 이내에서 제공되는 음원과 상기 기준거리 밖에서 제공되는 음원을 식별하는 단계; 상기 미리 설정된 기준거리 이내에서 제공되는 음원의 높이 정보와 수평 정보를 고려한 3D 앰비소닉(ambisonics)을 이용하여 음장(sound field)을 처리하는 단계; 상기 기준거리 밖에서 제공되는 음원의 수평 정보를 고려한 2D 앰비소닉을 이용하여 음장을 처리하는 단계를 포함하는, 오디오 신호 처리 방법일 수 있다.

[0009] 상기 3D 앰비소닉을 이용하여 처리된 음장과 상기 2D 앰비소닉을 이용하여 처리된 음장을 조합하여, 상기 임의의 가상 공간에 위치한 사용자에게 재현되는 상기 VR 콘텐츠에 적합한 음장을 재현하는 단계를 더 포함하는, 오디오 신호 처리 방법일 수 있다.

[0010] 상기 기준거리는, 높이 정보보다 수평 정보에 민감한 인간의 청취 특성을 고려하여 설정되는, 오디오 신호 처리 방법일 수 있다.

- [0011] 상기 앰비소닉은 차수 n 과 degree m 을 포함하는 구면 조화(spherical harmonics)로 표현되는 방향성분 정보를 이용하여 음장을 처리하는 방식을 나타내는, 오디오 신호 처리 방법일 수 있다.
- [0012] 상기 기준거리 이외의 경우 차수(order) $n=0 \sim N_{2D}$ 인 2D 앰비소닉을 이용하며, 상기 기준거리 이내의 경우 차수 $n=N_{2D}+1 \sim N_{3D}$ 인 3D 앰비소닉을 이용하는, 오디오 신호 처리 방법일 수 있다.
- [0013] 상기 기준거리 이외의 경우 degree $m=\pm n$ 을 이용하며, 상기 기준거리 이내의 경우 degree $m=\pm n$ 을 제외한 degree $m=-(n-1) \sim (n-1)$ 를 이용하는, 오디오 신호 처리 방법일 수 있다.
- [0014] 상기 앰비소닉은, 푸리에 계수 $A_{nm}(k)$, 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)인 방사 함수(radial function)를 나타내는 $b_n(k)$, 정규화 상수(normalization constant)인 Γ_{nm} , 버금 르장드르 함수(Associated Legendre function)인 $P_n^m(x)$, 방위 조화(azimuthal harmonics)인 $e^{im\phi}$ 를 포함하는, 오디오 신호 처리 방법일 수 있다.
- [0015] 상기 푸리에 계수 $A_{nm}(k)$ 는 자유음장(free field)에서 획득된 평면파를 통해 결정될 수 있고, 상기 $b_n(k)$ 는 자유음장(free field)의 음원 입사와 마이크로폰으로 인한 스캐터링(scattering)을 반영하는, 오디오 신호 처리 방법일 수 있다.
- [0016] 상기 방사 함수(radial function) $b_n(k)$ 로 사용되는 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)는 상기 임의의 가상 공간에 위치한 사용자로부터 근거리에서의 크기가 원거리에서의 크기보다 큰, 오디오 신호 처리 방법일 수 있다.
- [0017] 일 측면에 따르면, 프로세서 및 컴퓨터에서 읽을 수 있는 명령어를 포함하는 메모리를 포함하고, 상기 명령어가 상기 프로세서에서 실행되면, 상기 프로세서는 VR 환경에서 사용자가 위치하는 임의의 가상 공간에서 재현되는 VR 콘텐츠에 따른 음장을 재현하기 위해, 미리 설정된 기준거리 이내에서 제공되는 음원과 상기 기준거리 밖에서 제공되는 음원을 식별하고, 상기 미리 설정된 기준거리 이내에서 제공되는 음원의 높이 정보와 수평 정보를 고려한 3D 앰비소닉(ambisonics)을 이용하여 음장(sound field)을 처리하고, 상기 기준거리 밖에서 제공되는 음원의 수평 정보를 고려한 2D 앰비소닉을 이용하여 음장을 처리하는, 오디오 신호 처리 장치일 수 있다.
- [0018] 상기 프로세서는, 상기 3D 앰비소닉을 이용하여 처리된 음장과 상기 2D 앰비소닉을 이용하여 처리된 음장을 조합하여, 상기 임의의 가상 공간에 위치한 사용자에게 재현되는 상기 VR 콘텐츠에 적합한 음장을 재현하는, 오디오 신호 처리 장치일 수 있다.
- [0019] 상기 기준거리는, 높이 정보보다 수평 정보에 민감한 인간의 청취 특성을 고려하여 설정되는, 오디오 신호 처리 장치일 수 있다.
- [0020] 상기 앰비소닉은 차수 n 과 degree m 을 포함하는 구면 조화(spherical harmonics)로 표현되는 방향성분 정보를 이용하여 음장을 처리하는 방식을 나타내는, 오디오 신호 처리 장치일 수 있다.
- [0021] 상기 기준거리 이외의 경우 차수(order) $n=0 \sim N_{2D}$ 인 2D 앰비소닉을 이용하며, 상기 기준거리 이내의 경우 차수 $n=N_{2D}+1 \sim N_{3D}$ 인 3D 앰비소닉을 이용하는, 오디오 신호 처리 장치일 수 있다.
- [0022] 상기 기준거리 이외의 경우 degree $m=\pm n$ 을 이용하며, 상기 기준거리 이내의 경우 degree $m=\pm n$ 을 제외한 degree $m=-(n-1) \sim (n-1)$ 를 이용하는, 오디오 신호 처리 장치일 수 있다.
- [0023] 상기 앰비소닉은, 푸리에 계수 $A_{nm}(k)$, 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)인 방사 함수(radial function)를 나타내는 $b_n(k)$, 정규화 상수(normalization constant)인 Γ_{nm} , 버금 르장드르 함수(Associated Legendre function)인 $P_n^m(x)$, 방위 조화(azimuthal harmonics)인 $e^{im\phi}$ 를 포함하는, 오디오 신호 처리 장치일 수 있다.
- [0024] 상기 푸리에 계수 $A_{nm}(k)$ 는 자유음장(free field)에서 획득된 평면파를 통해 결정될 수 있고, 상기 $b_n(k)$ 는 자유음장(free field)의 음원 입사와 마이크로폰으로 인한 스캐터링(scattering)을 반영하는, 오디오 신호 처리 장치일 수 있다.

[0025] 상기 방사 함수(radial function) $b_n(k)$ 로 사용되는 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)는 상기 임의의 가상 공간에 위치한 사용자로부터 근거리에서의 크기가 원거리에서의 크기보다 큰, 오디오 신호 처리 장치일 수 있다.

발명의 효과

[0027] 일 실시예에 따르면, 복합 차수 앰비소닉을 이용한 오디오 신호 처리 방법으로서, 사용자의 위치가 자유롭게 변화하는 VR 환경에서 적은 연산량의 앰비소닉을 이용하여, 사용자에게 제공되는 VR 콘텐츠에 적합한 음장을 처리하는 오디오 신호 처리 방법일 수 있다.

도면의 간단한 설명

[0029] 도 1은 일 실시예에 따른, HMD를 통해 가상 공간에 있는 사용자에게 VR 콘텐츠를 제공하는 도면을 나타낸다.
 도 2는 일 실시예에 따른, $n=4$ 인 경우를 나타낸 도면이다.
 도 3은 일 실시예에 따른, 거리 r 과 거리 r' 에 대한 상대적인 음장 처리를 나타낸 도면이다.
 도 4는 일 실시예에 따른, $N_{2D}=2$, $N_{3D}=4$ 인 경우의 관계를 나타내고 있다.
 도 5는 일 실시예에 따른, 사용하거나 사용하지 않는 구성요소(component)를 나타내고 있다.
 도 6은 일 실시예에 따른, 방사 함수로 사용되는 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)의 크기가 사용자로부터 거리에 따른 변화를 나타낸다.
 도 7은 일 실시예에 따른, 오디오 신호 처리 장치가 수행하는 오디오 신호 처리 방법을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0030] 이하, 실시예들을 첨부된 도면을 참조하여 상세하게 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.

[0031] 아래 설명하는 실시예들에는 다양한 변경이 가해질 수 있다. 아래 설명하는 실시예들은 실시 형태에 대해 한정하려는 것이 아니며, 이들에 대한 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0032] 제1 또는 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 이런 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 이해되어야 한다. 예를 들어, 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소는 제1 구성요소로도 명명될 수 있다.

[0033] 실시예에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 실시예를 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수 개의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서 상에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.

[0034] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 실시예가 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

[0035] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조 부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 실시예를 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 실시예의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.

[0036] 이하, 본 발명의 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.

[0038] 도 1은 일 실시예에 따른, HMD를 통해 가상 공간에 있는 사용자에게 VR 콘텐츠를 제공하는 도면을 나타낸다.

[0039] VR 콘텐츠는 HMD(Head Mounted Device)를 착용한 사용자에게 가상 공간에서 제공될 수 있다. 이때, 사용자는

입력의 가상 공간을 이동하면서 다양한 인터랙션을 할 수 있고, VR 콘텐츠는 인터랙션에 반응하여 사용자에게 제공될 수 있다.

[0040] 여기서, VR 콘텐츠를 제공할 때, 다양한 오디오 신호 처리 방법이 적용될 수 있다. 일례로서, 제공되는 VR 콘텐츠의 장면(scene)에 따라 음장을 재현하기 위해, 장면 기반(scene-based) 오디오 렌더링 방식인 앰비소닉(Ambisonics)이 사용될 수 있다.

[0041] 앰비소닉은 구면 조화 함수(spherical harmonics function)로 나타내어지는 방향 성분 정보를 이용하여 음장(sound field)을 처리하는 오디오 신호 처리 방식이다. 즉, 앰비소닉은 음장을 여러 개의 기본 함수의 합으로 인식하고, 음장을 나타내기 위한 각각의 기본 함수의 기여도를 확인하여 음장을 처리하는 오디오 신호 처리 방식이다. 따라서, 앰비소닉은 음장을 표현하기 위해 3차원 구면 조화 함수를 이용할 수 있다.

[0042] 이때, 앰비소닉의 차수가 높아지면 처리해야 하는 데이터 양이 많아져, 실시간으로 인터랙션을 반영하는 VR 콘텐츠를 제공할 때 많은 연산량이 필요할 수 있다. 그러므로, 사용자의 위치가 자유롭게 변화하는 VR 환경에서 적은 연산량의 앰비소닉을 이용하여, 음장을 나타내는 방법이 필요하다.

[0044] 도 2는 일 실시예에 따른, n=4인 경우를 나타낸 도면이다.

[0045] 앰비소닉은 구면 조화 함수(spherical harmonics function)로 나타내어지는 방향 성분 정보를 이용하여 음장(sound field)을 처리하는 오디오 신호 처리 방식으로, 앰비소닉 신호는 아래의 수학적 식 1과 같이 표현될 수 있다.

수학적 식 1에서 Γ^{mm} 는 아래의 수학적 식 2와 같이 표현될 수 있다.

수학적 식 1

$$p(r, \theta, \phi, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n A_{nm}(k) b_n(k) \Gamma_{nm} P_n^m(\cos \theta) e^{im\phi}$$

[0046]

수학적 식 2

$$\Gamma_{nm} = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}}$$

[0047]

[0048] 여기서, n과 m은 차수(order)과 degree를 나타내며, $A_{nm}(k)$ 는 푸리에 계수(Fourier coefficient)이고, $b_n(k)$ 는 방사 함수(radial function)로서 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)를 나타내며, Γ^{mm} 는 정규화 상수(normalization constant)이고, $P_n^m(x)$ 은 버금 르장드르 함수(Associated Legendre function)를 나타내며, $e^{im\phi}$ 는 방위 조화(azimuthal harmonics)를 나타낸다. 이때, $\Gamma_{nm} P_n^m(\cos x) e^{im\phi}$ 는 구면 조화(spherical harmonics)를 나타낸다.

[0049] 수학적 식 1에서 확인가능하듯이, VR 콘텐츠에 대응하는 정확한 음장을 표현하기 위해 앰비소닉을 사용할 경우, n=0 ~ 무한대 차수까지 관련된 정보를 이용할 수 있다.

[0050] 여기서, 각 차수(n)별 구면 조화는 2n+1 종류를 가질 수 있으며, 특정 차수 N 까지는 (N+1)² 만큼의 채널 수가 존재할 수 있다. 예를 들면, n=1차인 경우, W/X/Y/Z의 총 4개 채널이 존재할 수 있으며, 이는 (1+1)²만큼의 채널 수를 나타낸다. 즉, 0차에서 2*0+1개 채널, 1차에서 2*1+1개 채널이 존재하므로, 총 4개의 채널이 존재할 수 있다. 다른 예를 들면, n=10차 인 경우, 정확한 음장을 나타내기 위해서는 총 (10+1)²=121개의 채널(즉, (2*0+1) + (2*1+1) + (2*2+1) + ~ + (2*10+1))에 대한 연산을 수행할 필요가 있다.

[0052] 도 3은 일 실시예에 따른, 거리 r과 거리 r'에 대한 상대적인 음장 처리를 나타낸 도면이다.

[0053] 일 실시예에 따르면, 차수가 높아지면 처리해야 하는 데이터 양이 많아지므로, 인간의 청취 특성을 고려하면 approximation을 통해 적은 연산을 수행할 수 있다. 음원은 높이 정보와 수평 정보를 포함할 수 있다. 인간의 청취 특성은 높이 정보 보다 수평 정보에 민감할 수 있다. 따라서, 이와 같은 인간의 청취 특성을 고려하여, 먼 거리의 음원에 대해서는 높이 정보가 제공되지 않더라도 수평 정보만 이용하여 음장이 재현될 수 있으며, 사용자 주변의 가까운 거리의 음원에 대해서는 수평 정보뿐만 아니라 높이 정보를 이용하여 음장이 재현될 수 있다.

[0054] 이때, 2차원 정보로 먼 거리까지의 음장을 처리하고 3차원 정보로 가까운 거리의 음장을 처리하기 위하여, 평면 상에 배치되어 2차원 음장을 획득하는 마이크로폰은 3차원 마이크로폰 보다 멀리 배치될 수 있다. 예를 들면, 사각형 공간의 녹음 공간인 경우, 녹음 공간의 정중앙에 3차원 마이크로폰을 배치하고 각 가장자리(네군데 모서리)에 수평면 마이크로폰을 배치하여 2차원 음장을 획득할 수 있다.

[0055] 또는, 거리 구분 없이 구형의 마이크로폰 상의 표면에 2차원 마이크로폰과 3차원 마이크로폰이 배치될 수 있다. 이때, 2차원 마이크로폰의 경우 상대적으로 적은 개수 및 넓은 간격으로 수평면 상에 배치되고, 3차원 마이크로폰의 경우 상대적으로 많은 개수 및 좁은 간격으로 수평면을 제외한 모든 공간에 배치될 수 있다. 따라서, 2차원 정보보다 더욱 정밀한 3차원 정보가 획득될 수 있다.

[0056] 이와 같이, 사용자로부터의 거리에 따라 음원에 대한 높이 정보를 이용할 지 여부를 결정하는 오디오 렌더링 방식을 복합 차수 앰비소닉(Composited Order Ambisonics, COA)이라 하며, 아래의 수학적 식 3과 같이 표현될 수 있다.

수학적 식 3

$$p'(r, \theta, \phi, \omega) = \sum_{n=0}^{N_{2D}} A_{nm}(k) b_n(k) \Gamma_{nm} P_n^{nz}(\cos \theta) e^{mz\phi} + \sum_{n'=N_{2D}+1}^{N_{3D}} \sum_{m=-(n-1)}^{n-1} A_{n'm}(k_{r'}) b_{n'}(k_{r'}) \Gamma_{n'm} P_n^m(\cos \theta) e^{mz\phi}$$

[0057] ...

[0058] 거리 r에 대응하는 전체 영역의 경우, n=0~N_{2D} 차수를 가지는 2D 앰비소닉을 이용하여 음장을 처리할 수 있다. 또는, r'까지 거리의 음장을 표현하기 위해, n'= N_{2D} +1~N_{3D}차수의 높이 정보를 가지는 3D 앰비소닉을 이용하여 음장을 처리할 수 있다. 여기서, r은 사용자로부터 원거리를 나타내며, r'은 사용자로부터 근거리를 나타낸다. r과 r'은 상이한 거리를 나타내지만, 특별한 경우 r과 r'는 동일할 수도 있다.

[0059] 수학적 식 3은 degree m=±n인 경우와 degree m=±n을 제외한 m=-(n-1)~(n-1)인 경우의 수식을 포함하고 있다. 수학적 식 3을 보다 구체적으로 표현할 경우, 아래의 수학적 식 4와 같을 수 있다.

수학적 식 4

$$p'(r, \theta, \phi, \omega) = \sum_{n=0}^{N_{2D}} A_{nm}(k) \left(j_n(kr) - \frac{j'_n(kr)}{h'_n(kr)} h_n(kr) \right) \Gamma_{nm} P_n^m(\cos \theta) e^{mz\phi} + \sum_{n'=N_{2D}+1}^{N_{3D}} \sum_{m=-(n-1)}^{n-1} A_{n'm}(k_{r'}) \left(j_{n'}(k_{r'}) - \frac{j'_{n'}(k_{r'})}{h'_{n'}(k_{r'})} h_{n'}(k_{r'}) \right) \Gamma_{n'm} P_n^m(\cos \theta) e^{mz\phi}$$

[0060] ...

[0061] 수학적 식 4에서, $A_{nm}(k) = 4\pi i^n Y_n^m(\theta, \phi)^*$ 로서 자유 음장(free field)에서 획득된 평면파를 통해 결정될 수 있고, 구면 베셀 함수(spherical Bessel function, j_n(kr))은 자유 음장에서 방사 함수(radial function)에 해당하고, 한켈 함수(Hankel function, h_n(kr))는 rigid환경에서처럼 스캐터링(scattering)이 발생하는 경우를 나타내는 방사 함수(radial function)에 해당한다. 즉, 자유 음장의 음원 입사와 마이크로폰으로 인한 스캐터

링(scattering)을 반영하는 $b_n(kr)$ 이 적용된다.

[0063] 도 4는 일 실시예에 따른, $N_{2D}=2$, $N_{3D}=4$ 인 경우의 관계를 나타내고 있다. 또한, 도 5는 일 실시예에 따른, 사용하지 않거나 사용하지 않는 구성요소(component)를 나타내고 있다. 보다 n 이 고차인 경우, COA를 이용하면 사용하지 않는 구성요소들은 증가할 수 있다.

[0064] 일례로, 앰비소닉을 사용하는 시스템은 아래의 표 1과 같은 파라미터로서 N_{2D} , N_{3D} , COA flag 를 메타데이터로 전송하거나 수신하여 사용할 수 있다.

표 1

parameter		value
Type	HOA	0
	COA	1
	reserved...	2...
N for HOA		5
N _{2D} for COA		2
N _{3D} for COA		4

[0065]

[0067] 도 6은 일 실시예에 따른, 방사 함수로 사용되는 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)의 크기가 사용자로부터 거리에 따른 변화를 나타낸다. 도 6에서 확인 가능하듯이, 구면 베셀 함수(spherical Bessel function, $j_n(kr)$) 또는 한켈 함수(Hankel function, $h_n(kr)$)는 사용자로부터 근거리에서 음장 재현에 있어 영향이 크고, 원거리에서 음장 재현에 있어 영향이 작다는 것을 나타낸다. 따라서, 사용자로부터 먼 위치의 경우, 저차 성분만 사용함으로써 연산량을 줄일 수 있다. 아래의 수학적 5는 구면 베셀 함수를 나타내고, 수학적 6은 한켈 함수를 나타낸다.

수학적 5

$$j_n(x) = (-1)^n x^n \left(\frac{1}{x} \frac{d}{dx} \right)^n \frac{\sin(x)}{x}$$

[0068]

수학적 6

$$h_n^{(1)}(x) = -i(-1)^n x^n \left(\frac{1}{x} \frac{d}{dx} \right)^n \frac{e^{ix}}{x}$$

[0069]

[0070] 일 실시예에 따르면, VR 콘텐츠가 재생되는 VR 환경에서 앰비소닉의 차수를 무한대까지 이용하지 않고, 사용자로부터의 거리에 따른 제한된 차수의 2D 앰비소닉으로 음장 처리와 3D 앰비소닉으로 음장 처리를 이용하여 적은 연산량으로 VR 콘텐츠에 적합한 음장을 재현할 수 있다.

[0072] 도 7은 일 실시예에 따른, 오디오 신호 처리 장치가 수행하는 오디오 신호 처리 방법을 나타낸다.

[0073] 단계(710)에서, 오디오 신호 처리 장치는 VR 환경에서 사용자가 위치하는 임의의 가상 공간에서 재현되는 VR 콘텐츠에 따른 음장을 재현하기 위해, 미리 설정된 기준거리 이내에서 제공되는 음원과 기준거리 밖에서 제공되는

음원을 식별할 수 있다. 이때, 기준거리는, 높이 정보보다 수평 정보에 민감한 인간의 청취 특성을 고려하여 설정될 수 있다.

[0074] 단계(720)에서, 오디오 신호 처리 장치는 미리 설정된 기준거리 이내에서 제공되는 음원의 높이 정보와 수평 정보를 고려한 3D 앰비소닉을 이용하여 음장을 나타낼 수 있다.

[0075] 단계(730)에서, 오디오 신호 처리 장치는 기준거리 밖에서 제공되는 음원의 수평 정보를 고려한 2D 앰비소닉을 이용하여 음장을 나타낼 수 있다. 따라서, 오디오 신호 처리 장치는 3D 앰비소닉을 이용하여 처리된 음장과 2D 앰비소닉을 이용하여 처리된 음장을 조합하여, 임의의 가상 공간에 위치한 사용자에게 재현되는 VR 콘텐츠에 적합한 음장을 재현할 수 있다.

[0076] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 콘트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPA(field programmable array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 콘트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.

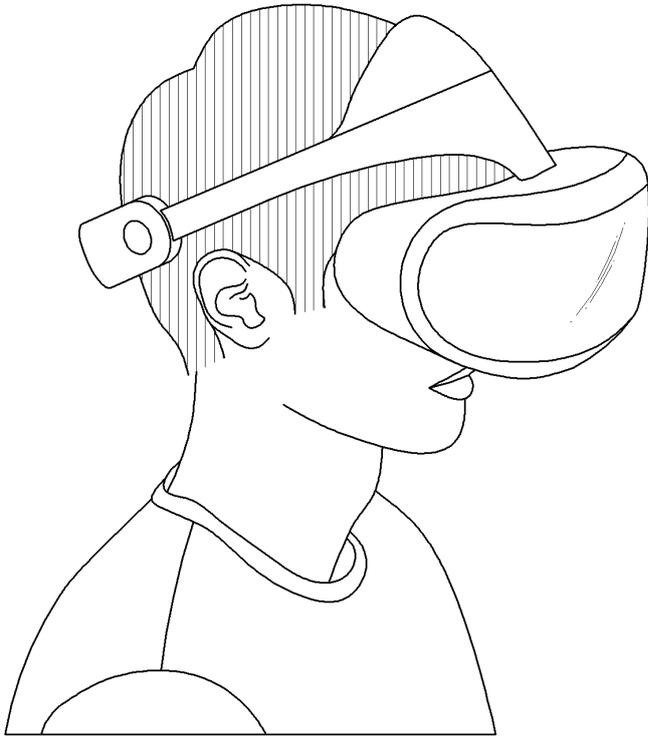
[0077] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상 장치(virtual equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(signal wave)에 영구적으로, 또는 일시적으로 구체화(embodiment)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0078] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

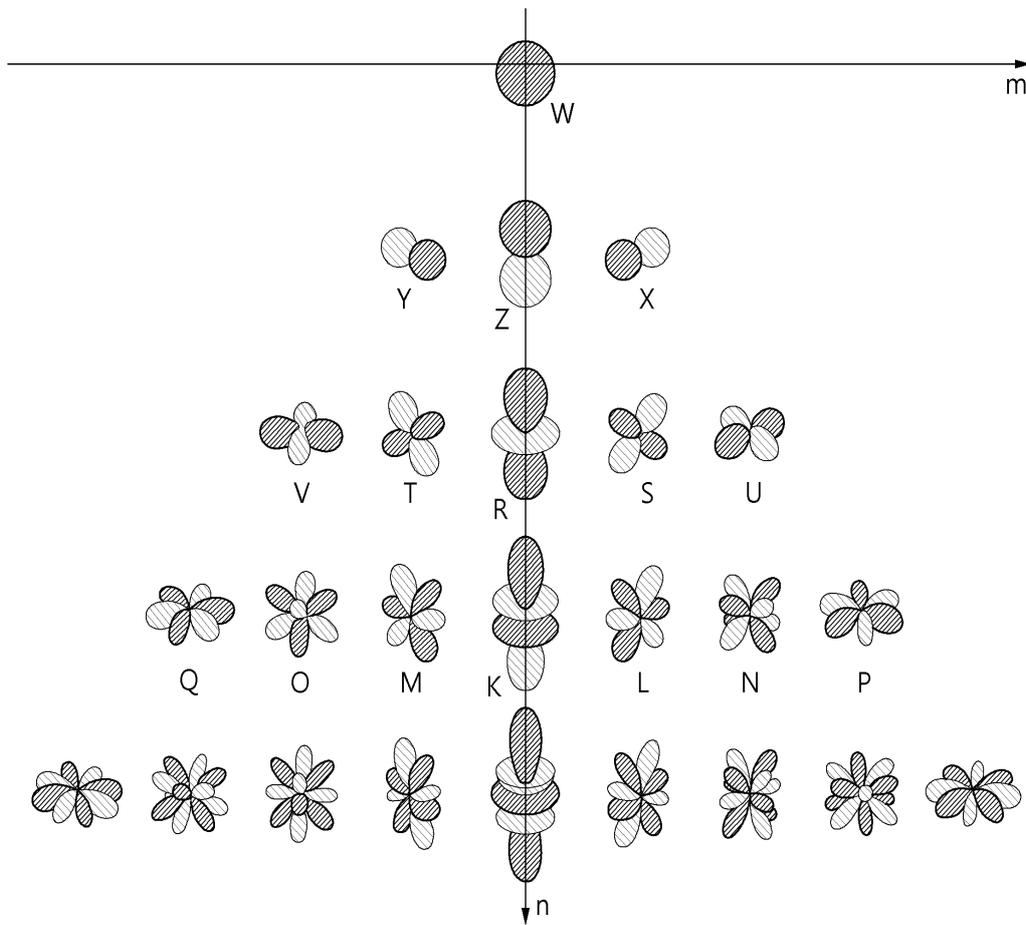
[0079] 이상과 같이 실시예들이 비록 한정된 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기를 기초로 다양한 기술적 수정 및 변형을 적용할 수 있다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

도면

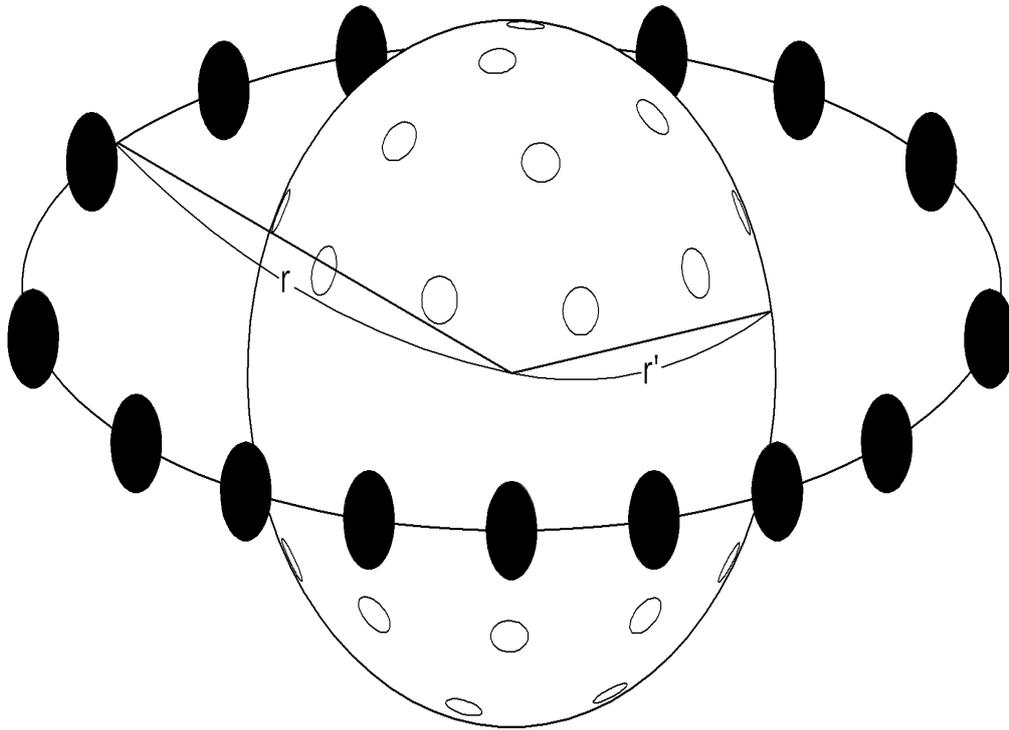
도면1



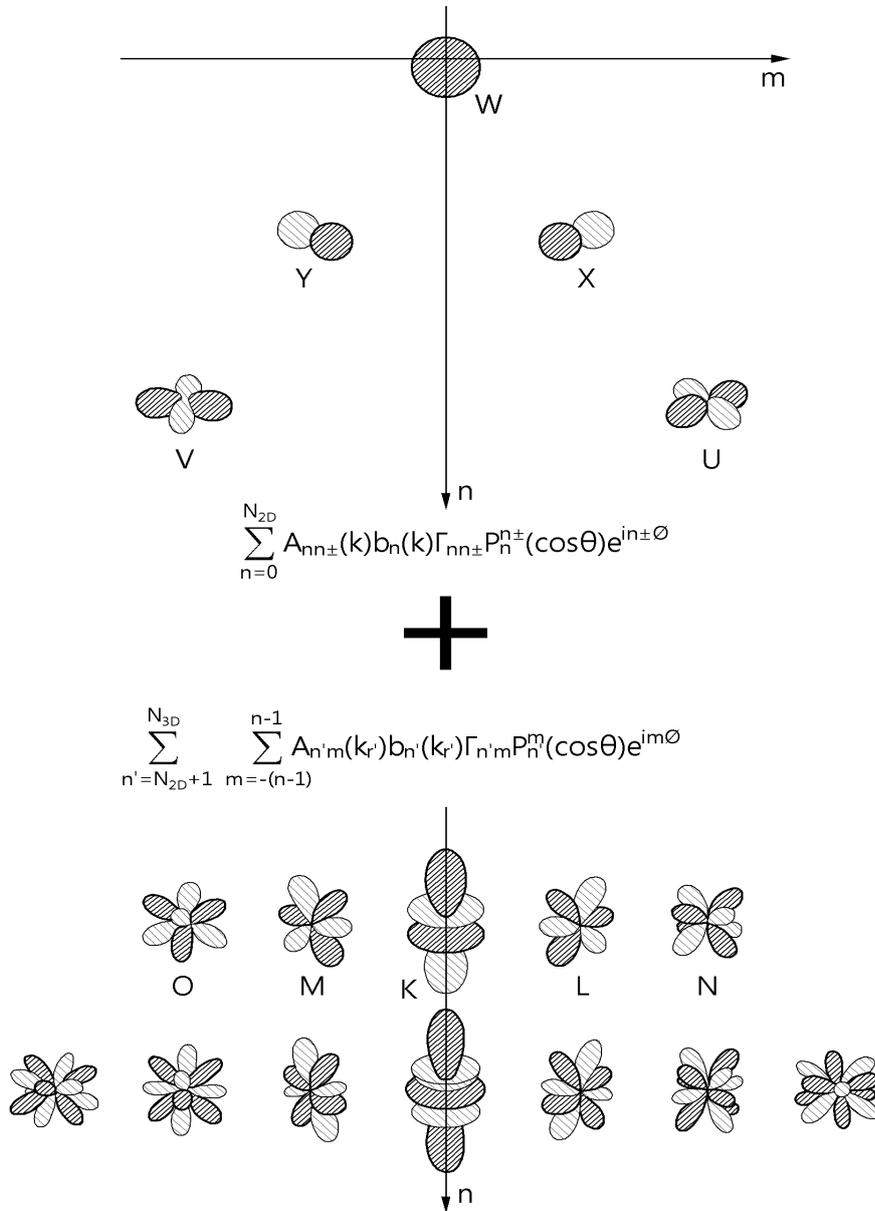
도면2



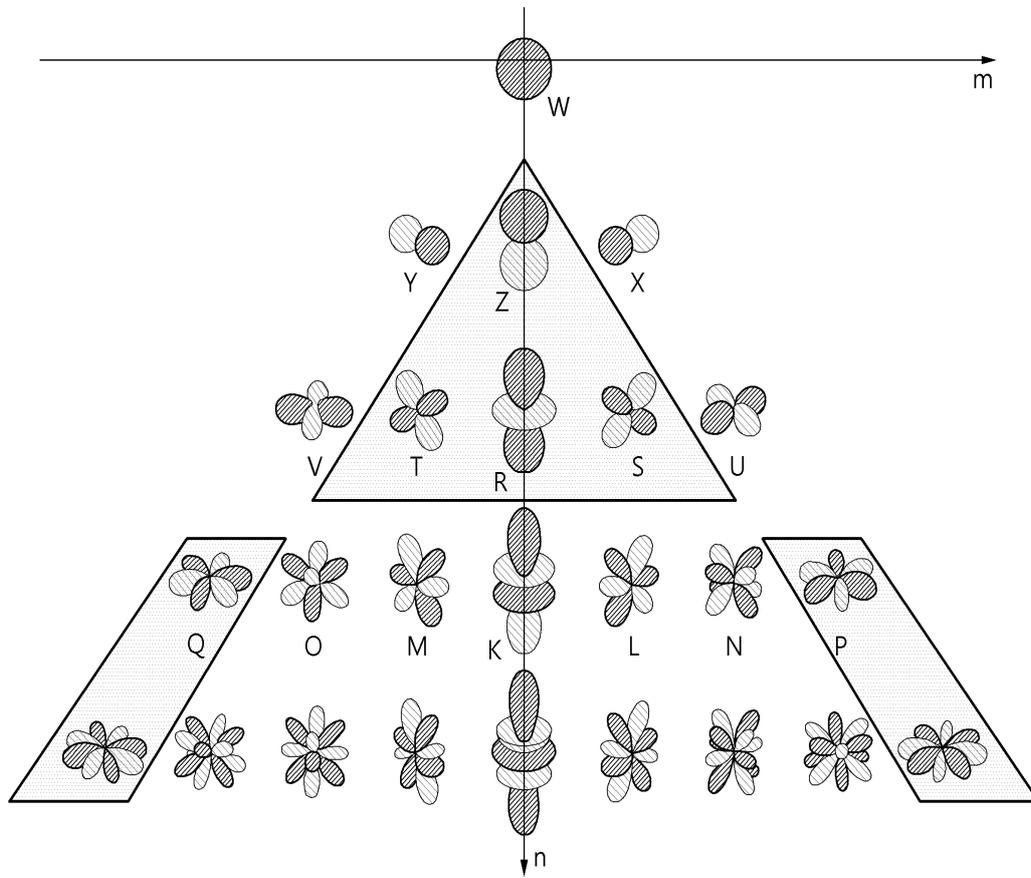
도면3



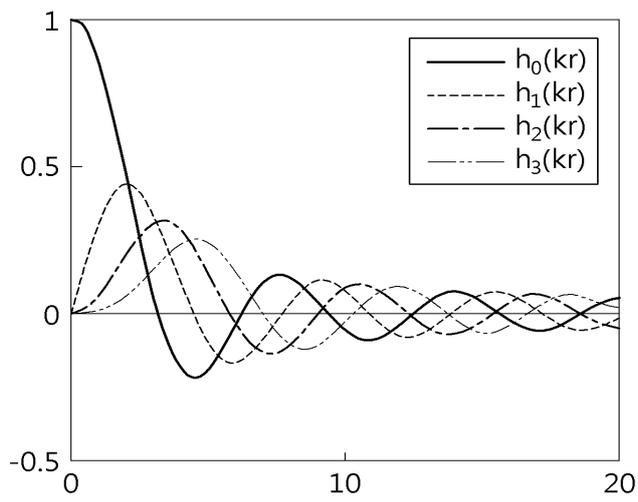
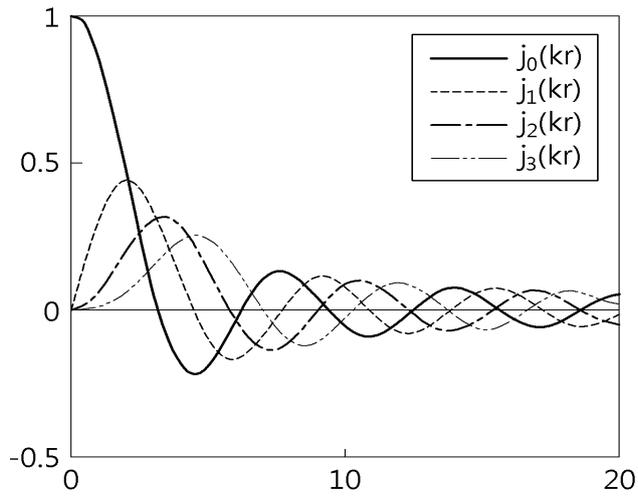
도면4



도면5



도면6



도면7

