



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2018년12월18일
(11) 등록번호 10-1930657
(24) 등록일자 2018년12월12일

- (51) 국제특허분류(Int. Cl.)
G06T 19/00 (2011.01) *G06F 3/01* (2006.01)
G06F 3/14 (2006.01) *G06F 3/16* (2018.01)
G06K 9/00 (2006.01) *G06K 9/46* (2006.01)
G06T 7/73 (2017.01) *H04N 5/225* (2006.01)
H04N 5/33 (2006.01)
- (52) CPC특허분류
G06T 19/006 (2013.01)
G06F 3/017 (2013.01)
- (21) 출원번호 10-2017-7013982
- (22) 출원일자(국제) 2015년10월23일
 심사청구일자 2017년05월23일
- (85) 번역문제출일자 2017년05월23일
- (65) 공개번호 10-2017-0095834
- (43) 공개일자 2017년08월23일
- (86) 국제출원번호 PCT/US2015/000116
- (87) 국제공개번호 WO 2016/064435
 국제공개일자 2016년04월28일
- (30) 우선권주장
 62/068,423 2014년10월24일 미국(US)
- (56) 선행기술조사문헌
 US20140002442 A1*
 *는 심사관에 의하여 인용된 문헌

- (73) 특허권자
유센스, 인코퍼레이티드
 미국 캘리포니아주 95110 새너제이 스위트 130 테크놀로지 드라이브 1741
- (72) 발명자
페이 유에
 미국 캘리포니아주 95134 새너제이 누네스 드라이브 2039
허 안리
 미국 캘리포니아주 95134 새너제이 누네스 드라이브 2039
- (74) 대리인
특허법인아주김장리

전체 청구항 수 : 총 20 항

심사관 : 장석환

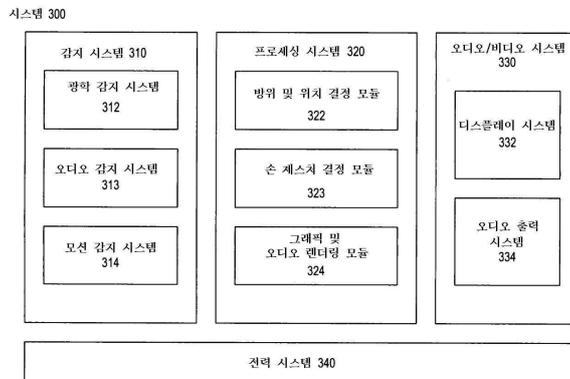
(54) 발명의 명칭 **몰입식 및 대화식 멀티미디어 생성을 위한 시스템 및 방법**

(57) 요약

장치가 개시된다. 장치는 적어도 하나의 카메라를 포함하는 광학 감지 시스템을 포함하며, 상기 적어도 하나의 카메라는 물리적 환경의 이미지를 획득하도록 구성된다. 장치는 프로세싱 시스템을 추가로 포함한다. 프로세싱 시스템은 이미지로부터 두드러진 특징을 검출하며 상기 검출된 두드러진 특징에 기초하여 물리적 환경에 대하여

(뒷면에 계속)

대표도 - 도3



장치의 방위 및/또는 위치에서의 변화를 결정하도록 구성된 방위 및 위치 결정 모듈을 포함한다. 상기 프로세싱 시스템은 이미지에 및 장치의 방위 및/또는 위치에서의 결정된 변화에 기초하여 물리적 환경의 렌더링을 결정하며, 물리적 환경의 렌더링에 관련된 데이터를 디스플레이 시스템에 제공하도록 구성된 렌더링 모듈을 또한 포함한다.

(52) CPC특허분류

G06F 3/14 (2013.01)

G06F 3/16 (2013.01)

G06K 9/00355 (2013.01)

G06K 9/4676 (2013.01)

G06T 7/74 (2017.01)

H04N 5/2258 (2013.01)

H04N 5/332 (2013.01)

G06T 2207/10012 (2013.01)

G06T 2207/30244 (2013.01)

명세서

청구범위

청구항 1

장치로서,

적어도 하나의 카메라를 포함한 광학 감지 시스템으로서, 상기 적어도 하나의 카메라는, 상기 장치가 물리적 환경에서 제1 방위 및/또는 위치에 있을 때, 상기 물리적 환경의 하나 이상의 제1 이미지를 획득하고, 상기 장치가 상기 물리적 환경에서 제2 방위 및/또는 위치에 있을 때, 상기 물리적 환경의 하나 이상의 제2 이미지를 획득하도록 구성되는, 상기 광학 감지 시스템; 및

프로세싱 시스템을 포함하되,

상기 프로세싱 시스템은, 방위 및 위치 결정 모듈 및 렌더링 모듈을 포함하며;

상기 방위 및 위치 결정 모듈은,

상기 하나 이상의 제1 이미지로부터 하나 이상의 제1 두드러진 특징을 검출하고,

상기 제1 두드러진 특징과 연관된 제1 세트의 좌표를 결정하며,

상기 하나 이상의 제2 이미지로부터 하나 이상의 제2 두드러진 특징을 검출하고,

상기 제2 두드러진 특징과 연관된 제2 세트의 좌표를 결정하며,

상기 하나 이상의 제2 이미지와 연관된 하나 이상의 2차원(2D) 공간에 상기 제1 세트의 좌표를 재투사하고,

상기 재투사된 제1 세트의 좌표를 상기 하나 이상의 2차원 공간에서의 상기 제2 세트의 좌표와 비교하여, 상기 하나 이상의 제1 두드러진 특징에 대응하는 상기 하나 이상의 제2 두드러진 특징을 결정하며, 그리고

결정된 대응하는 제1 및 제2 두드러진 특징에 기초하여 상기 물리적 환경에 대한 상기 장치의 상기 제1 방위 및/또는 위치와 상기 제2 방위 및/또는 위치 사이의 변화를 결정하도록 구성되고;

상기 렌더링 모듈은,

상기 장치의 방위 및/또는 위치에서의 상기 결정된 변화와, 상기 하나 이상의 제1 이미지와, 상기 하나 이상의 제2 이미지에 기초하여 상기 물리적 환경의 렌더링을 결정하고, 그리고

상기 물리적 환경의 렌더링에 관련된 데이터를 디스플레이 시스템에 제공하도록 구성된, 장치.

청구항 2

제1항에 있어서,

상기 광학 감지 시스템은 제1 카메라 및 제2 카메라를 포함하되, 상기 제1 카메라는 상기 물리적 환경의 제1 좌측 이미지 및 제2 좌측 이미지를 획득하도록 구성되고, 상기 제2 카메라는 상기 물리적 환경의 제1 우측 이미지 및 제2 우측 이미지를 획득하도록 구성되며, 상기 제1 이미지는 상기 제1 좌측 이미지 및 상기 제1 우측 이미지를 포함하고, 상기 제2 이미지는 상기 제2 좌측 이미지 및 상기 제2 우측 이미지를 포함하며,

상기 하나 이상의 제1 두드러진 특징에 대응하는 상기 하나 이상의 제2 두드러진 특징을 결정하기 위해, 방위 및 위치 결정 모듈은, 위치 근접성, 특징 근접성, 및 스테레오 제약에서 상기 재투사된 제1 좌표 세트를 상기 제2 좌표 세트와 비교하도록 구성되고,

상기 방위 및 위치 결정 모듈은, 상기 제1 및 제2 두드러진 특징이 서로에 대응한다는 결정에 응답하여, 상기 제1 및 제2 세트의 좌표에 기초하여 상기 물리적 환경에 대해서 상기 장치의 방위 및/또는 위치에서의 변화를 결정하도록 구성되며;

상기 렌더링 모듈은 상기 제1 및 제2 이미지에 및 상기 장치의 방위 및/또는 위치에서의 상기 결정된 변화에 기

초하여 상기 물리적 환경의 입체 렌더링을 결정하도록 구성되는, 장치.

청구항 3

제2항에 있어서, 상기 제1 및 제2 카메라의 각각은 적색-녹색-청색-적외선(RGB-IR) 픽셀 센서를 포함하는, 장치.

청구항 4

제1항에 있어서, 상기 광학 감지 시스템은 제1 카메라, 제2 카메라, 및 미러를 포함하며, 상기 제1 카메라는 적외선(IR) 이미지를 획득하도록 구성되고, 상기 제2 카메라는 적색-녹색-청색(RGB) 이미지를 획득하도록 구성되며, 상기 미러는 상기 제2 카메라로부터 멀리 IR 광을 반사하도록 및 상기 RGB 이미지와 상기 IR 이미지를 동조시키도록 구성되는, 장치.

청구항 5

제4항에 있어서, 상기 제1 카메라는 비행시간(time-of-flight) 카메라인, 장치.

청구항 6

제4항에 있어서, 상기 제1 카메라는 IR 카메라인, 장치.

청구항 7

제1항에 있어서, 상기 광학 감지 시스템은 상기 물리적 환경으로 하나 이상의 패턴을 투사하도록 구성된 IR 조명기를 포함하며; 검출된 두드러진 특징은 상기 하나 이상의 패턴을 포함하는, 장치.

청구항 8

제1항에 있어서,

상기 프로세싱 시스템은 손 제스처 결정 모듈을 더 포함하되, 상기 손 제스처 결정 모듈은,

이미지로부터 손 제스처를 검출하고,

상기 손 제스처와 연관된 사용자 지시를 결정하도록 구성되며,

상기 렌더링 모듈은 상기 결정된 사용자 지시에 기초하여 상기 물리적 환경의 렌더링을 결정하도록 구성되는, 장치.

청구항 9

제8항에 있어서,

상기 손 제스처 결정 모듈은 상기 사용자 지시가 렌더링되는 오브젝트의 선택과 연관됨을 결정하도록 구성되고;

상기 렌더링 모듈은 상기 사용자 지시에 기초하여 상기 오브젝트의 렌더링에서의 변화를 결정하도록 구성되는, 장치.

청구항 10

제9항에 있어서, 상기 오브젝트는 가상 오브젝트인, 장치.

청구항 11

제9항에 있어서, 상기 오브젝트는 상기 물리적 환경에 존재하는 물리적 오브젝트인, 장치.

청구항 12

제1항에 있어서, 상기 렌더링 모듈은 상기 물리적 환경의 렌더링에서 상기 물리적 환경에 위치한 물리적 오브젝트의 이미지와 가상 오브젝트의 이미지를 합성하도록 구성되며; 상기 합성은 상기 가상 오브젝트와 연관된 제1 세트의 좌표 및 상기 물리적 오브젝트와 연관된 제2 세트의 좌표에 기초하는, 장치.

청구항 13

제12항에 있어서, 상기 합성은 상기 물리적 환경에서 비롯된 오디오 신호 및 상기 가상 오브젝트와 연관된 오디오 데이터의 렌더링을 포함하는, 장치.

청구항 14

제1항에 있어서, 스마트폰에 연결하도록 구성된 커넥터를 더 포함하는, 장치.

청구항 15

제14항에 있어서, 상기 스마트폰은 상기 프로세싱 시스템의 적어도 부분을 포함하는, 장치.

청구항 16

제14항에 있어서,

상기 스마트폰은 추가적인 콘텐츠를 제공하도록 구성되고;

상기 렌더링 모듈은 하나 이상의 이미지 및 상기 추가적인 콘텐츠에 기초하여 상기 물리적 환경의 렌더링을 결정하도록 구성되는, 장치.

청구항 17

시스템으로서,

프로세서; 및

명령들을 저장하는 비-일시적 컴퓨터 판독 가능한 저장 매체를 포함하되,

상기 명령들은 상기 프로세서에 의해 실행될 때, 상기 시스템으로 하여금,

상기 시스템이 물리적 환경에서 제1 방위 및/또는 위치에 있을 때, 상기 물리적 환경의 하나 이상의 제1 이미지를 획득하고,

상기 하나 이상의 제1 이미지로부터 하나 이상의 제1 두드러진 특징을 검출하며,

상기 제1 두드러진 특징과 연관된 제1 세트의 좌표를 결정하고,

상기 시스템이 상기 물리적 환경에서 제2 방위 및/또는 위치에 있을 때, 상기 물리적 환경의 하나 이상의 제2 이미지를 획득하며,

상기 하나 이상의 제2 이미지로부터 하나 이상의 제2 두드러진 특징을 검출하고,

상기 제2 두드러진 특징과 연관된 제2 세트의 좌표를 결정하며,

상기 하나 이상의 제2 이미지와 연관된 하나 이상의 2차원(2D) 공간에 상기 제1 세트의 좌표를 채투사하고,

상기 채투사된 제1 세트의 좌표를 상기 하나 이상의 2차원 공간에서의 상기 제2 세트의 좌표와 비교하여, 상기 하나 이상의 제1 두드러진 특징에 대응하는 상기 하나 이상의 제2 두드러진 특징을 결정하며,

결정된 대응하는 제1 및 제2 두드러진 특징에 기초하여 상기 물리적 환경에 대한 상기 시스템의 상기 제1 방위 및/또는 위치와 상기 제2 방위 및/또는 위치 사이의 변화를 결정하고,

상기 시스템의 방위 및/또는 위치에서의 상기 결정된 변화와, 상기 하나 이상의 제1 이미지와, 상기 하나 이상의 제2 이미지에 기초하여 상기 물리적 환경의 렌더링을 결정하고, 그리고

상기 물리적 환경의 렌더링에 관련된 데이터를 디스플레이 시스템에 제공하도록 하는, 시스템.

청구항 18

제17항에 있어서,

상기 명령들은 상기 시스템으로 하여금 추가적으로,

하나 이상의 이미지로부터 손 제스처를 검출하고,
 상기 손 제스처와 연관된 사용자 지시를 결정하며,
 상기 결정된 사용자 지시에 기초하여 상기 물리적 환경의 렌더링을 결정하도록 하는, 시스템.

청구항 19

방법으로서,
 시스템이 물리적 환경에서 제1 방위 및/또는 위치에 있을 때, 상기 물리적 환경의 하나 이상의 제1 이미지를 획득하는 단계;
 상기 하나 이상의 제1 이미지로부터 하나 이상의 제1 두드러진 특징을 검출하는 단계;
 상기 제1 두드러진 특징과 연관된 제1 세트의 좌표를 결정하는 단계;
 상기 시스템이 상기 물리적 환경에서 제2 방위 및/또는 위치에 있을 때, 상기 물리적 환경의 하나 이상의 제2 이미지를 획득하는 단계;
 상기 하나 이상의 제2 이미지로부터 하나 이상의 제2 두드러진 특징을 검출하는 단계;
 상기 제2 두드러진 특징과 연관된 제2 세트의 좌표를 결정하는 단계;
 상기 하나 이상의 제2 이미지와 연관된 하나 이상의 2차원(2D) 공간에 상기 제1 세트의 좌표를 재투사하는 단계;
 상기 재투사된 제1 세트의 좌표를 상기 하나 이상의 2차원 공간에서의 상기 제2 세트의 좌표와 비교하여, 상기 하나 이상의 제1 두드러진 특징에 대응하는 상기 하나 이상의 제2 두드러진 특징을 결정하는 단계;
 결정된 대응하는 제1 및 제2 두드러진 특징에 기초하여 상기 물리적 환경에 대한 상기 시스템의 상기 제1 방위 및/또는 위치와 상기 제2 방위 및/또는 위치 사이의 변화를 결정하는 단계;
 상기 시스템의 방위 및/또는 위치에서의 상기 결정된 변화와, 상기 하나 이상의 제1 이미지와, 상기 하나 이상의 제2 이미지에 기초하여 상기 물리적 환경의 렌더링을 결정하는 단계; 그리고
 상기 물리적 환경의 렌더링에 관련된 데이터를 디스플레이 시스템에 제공하는 단계를 포함하는, 방법.

청구항 20

제19항에 있어서,
 하나 이상의 이미지로부터 손 제스처를 검출하는 단계;
 상기 손 제스처와 연관된 사용자 지시를 결정하는 단계;
 상기 결정된 사용자 지시에 기초하여 상기 물리적 환경의 렌더링을 결정하는 단계를 추가적으로 포함하는, 방법.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은 2014년 10월 24일자로 출원된, 미국 가출원 번호 제62/068,423호의 이득을 주장하며, 그 내용은 전체적으로 참고로 본 명세서에 편입된다.

[0003] 분야

[0004] 본 개시는 인간-컴퓨터 상호 작용의 기술적 분야에 관한 것이며, 특히 몰입식 및 대화식 멀티미디어 생성에 관한 것이다.

배경 기술

- [0005] 몰입식 멀티미디어는 통상적으로 멀티미디어 데이터를 수신하는 사람이 환경에 물리적으로 존재하는 경험을 가질 수 있게 하는 상기 환경에 관련된 멀티미디어 데이터를 (오디오 및 비디오의 형태로) 제공하는 것을 포함한다. 몰입식 멀티미디어의 생성은 통상적으로 대화형이며, 따라서 사람에게 제공된 멀티미디어 데이터는 예를 들면, 사람의 물리적 위치, 사람에 의해 수행된 활동 등에 기초하여 자동으로 업데이트될 수 있다. 대화형 몰입식 멀티미디어는 예를 들면, 경험을 보다 실제처럼 만듦으로써 사용자 경험을 개선할 수 있다.
- [0006] 두 개의 주요 유형의 대화형 몰입식 멀티미디어가 있다. 제1 유형은 가상 현실(VR)이며, 여기에서 멀티미디어 데이터는 예를 들면, 실세계 또는 이미징된 세계에서의 장소에서 물리적 존재를 시뮬레이션하는 환경을 복제한다. 환경의 렌더링은 또한 사용자에 의해 수행된 동작을 반영하며, 그에 의해 사용자가 환경과 상호 작용할 수 있게 한다. 사용자의 동작(예로서, 신체 움직임)은 통상적으로 모션 센서에 의해 검출될 수 있다. 가상 현실은 시각, 청각, 촉각 등을 포함할 수 있는 감각 경험을 인위적으로 생성한다.
- [0007] 제2 유형의 대화형 몰입식 멀티미디어는 증강 현실(AR)이며, 여기에서 멀티미디어 데이터는 사람이 위치되는 물리적 환경의 실시간 그래픽 이미지, 뿐만 아니라 부가적인 디지털 정보를 포함한다. 부가적인 디지털 정보는 통상적으로 실시간 그래픽 이미지의 최상부 상에 놓이지만, 물리적 환경의 실시간 그래픽 이미지의 렌더링을 변경하거나 또는 강화하지 않는다. 부가적인 디지털 정보는 또한 가상 오브젝트의 이미지일 수 있지만, 통상적으로 가상 오브젝트의 이미지는 사실적인 렌더링을 생성하기 위해 물리적 환경과 합성되는 대신에, 실시간 그래픽 이미지의 최상부 상에 놓인다. 물리적 환경의 렌더링은 또한 상호 작용을 가능하게 하기 위해 사용자에 의해 수행된 동작 및/또는 사람의 위치를 반영할 수 있다. 사용자의 동작(예로서, 신체 움직임)은 통상적으로 모션 센서에 의해 검출될 수 있지만, 사람의 위치는 그래픽 이미지로부터 물리적 환경의 특징을 검출하고 추적함으로써 결정될 수 있다. 증강 현실은 부가적인 디지털 정보를 사람에게 동시에 제공하면서, 물리적 환경에 존재하는 동안 사람의 감각 경험의 일부를 복제할 수 있다.
- [0008] 현재 가상 오브젝트의 이미지 및 물리적 환경의 이미지의 사실적인 합성을 생성하는 가상 현실 및 증강 현실의 조합을 제공할 수 있는 시스템이 없다. 게다가, 현재 증강 현실 시스템은 사용자의 감각 경험을 복제할 수 있지만, 이러한 시스템은 통상적으로 사용자의 감지 능력을 강화시킬 수 없다.

발명의 내용

- [0009] 본 개시의 실시예의 부가적인 양상 및 이점이 다음의 설명에서 부분적으로 제공될 것이고, 다음의 설명으로부터 부분적으로 명백해지거나, 또는 본 개시의 실시예의 실시로부터 학습될 것이다.
- [0010] 몇몇 실시예에 따르면, 장치는 적어도 하나의 카메라를 포함하는 광학 감지 시스템을 포함하며, 상기 적어도 하나의 카메라는 물리적 환경의 이미지를 획득하도록 구성된다. 상기 장치는 프로세싱 시스템을 추가로 포함한다. 상기 프로세싱 시스템은 이미지로부터 두드러진 특징을 검출하며, 상기 검출된 두드러진 특징에 기초하여 물리적 환경에 대하여 상기 장치의 방위 및/또는 위치에서의 변화를 결정하도록 구성된 방위 및 위치 결정 모듈을 포함한다. 상기 프로세싱 시스템은 또한 이미지에 및 장치의 방위 및/또는 위치에서의 상기 결정된 변화에 기초하여 상기 물리적 환경의 렌더링을 결정하며, 상기 물리적 환경의 렌더링에 관련된 데이터를 디스플레이 시스템에 제공하도록 구성된 렌더링 모듈을 포함한다.
- [0011] 몇몇 실시예에 따르면, 광학 감지 시스템은 제1 카메라 및 제2 카메라를 포함하며, 상기 제1 카메라는 상기 물리적 환경의 제1 이미지를 획득하도록 구성되고, 상기 제2 카메라는 상기 물리적 환경의 제2 이미지를 획득하도록 구성된다. 상기 방위 및 위치 결정 모듈은 상기 장치가 물리적 환경에서 제1 위치에 있을 때 상기 제1 및 제2 이미지로부터 제1 두드러진 특징을 검출하고, 상기 제1 두드러진 특징과 연관된 제1 세트의 좌표를 결정하고, 상기 장치가 상기 물리적 환경에서 제2 위치에 있을 때 상기 제1 및 제2 이미지로부터 제2 두드러진 특징을 검출하고, 상기 제2 두드러진 특징과 연관된 제2 세트의 좌표를 결정하며, 상기 제1 및 제2 두드러진 특징이 서로에 대응한다는 결정에 응답하여, 상기 제1 및 제2 세트의 좌표에 기초하여 상기 물리적 환경에 대하여 상기 장치의 방위 및/또는 위치에서의 변화를 결정하도록 구성된다. 상기 렌더링 모듈은 또한 상기 제1 및 제2 이미지에 및 상기 장치의 방위 및/또는 위치에서의 상기 결정된 변화에 기초하여 상기 물리적 환경의 입체적 렌더링을 결정하도록 구성된다. 몇몇 실시예에 따르면, 제1 및 제2 카메라의 각각은 적색-녹색-청색-적외선(RGB-IR) 픽셀 센서를 포함한다.
- [0012] 몇몇 실시예에 따르면, 상기 광학 감지 시스템은 제1 카메라, 제2 카메라, 및 미러를 포함하며, 상기 제1 카메라는 적외선(IR) 이미지를 획득하도록 구성되고, 상기 제2 카메라는 적색-녹색-청색(RGB) 이미지를 획득하도록 구성되며, 상기 미러는 상기 제2 카메라로부터 멀리 IR 광을 반사하며 상기 RGB 이미지와 IR 이미지를 동조시키

도록 구성된다. 몇몇 실시예에서, 상기 제1 카메라는 비행시간(time-of-flight) 카메라이다. 몇몇 실시예에서, 상기 제1 카메라는 IR 카메라이다.

[0013] 몇몇 실시예에 따르면, 상기 광학 감지 시스템은 하나 이상의 패턴을 물리적 환경으로 투사하도록 구성된 IR 조명기를 포함하며; 여기에서 상기 검출된 두드러진 특징은 상기 하나 이상의 패턴을 포함한다.

[0014] 몇몇 실시예에 따르면, 상기 프로세싱 시스템은 이미지로부터 손 제스처를 검출하며, 상기 손 제스처와 연관된 사용자 지시를 결정하도록 구성된 손 제스처 결정 모듈을 추가로 포함한다. 상기 렌더링 모듈은 상기 결정된 사용자 지시에 기초하여 상기 물리적 환경의 렌더링을 결정하도록 구성된다. 몇몇 실시예에서, 상기 손 제스처 결정 모듈은 상기 사용자 지시가 렌더링되는 오브젝트의 선택과 연관됨을 결정하기 위해 구성되며, 상기 렌더링 모듈은 상기 사용자 지시에 기초하여 상기 오브젝트의 렌더링에서의 변화를 결정하도록 구성된다. 몇몇 실시예에서, 상기 오브젝트는 가상 오브젝트이다. 몇몇 실시예에서, 상기 오브젝트는 물리적 환경에 존재하는 물리적 오브젝트이다.

[0015] 몇몇 실시예에 따르면, 상기 렌더링 모듈은 상기 물리적 환경의 렌더링에서 상기 물리적 환경에 위치한 물리적 오브젝트의 이미지와 가상 오브젝트의 이미지를 합성하도록 구성되며; 여기에서 상기 합성은 상기 가상 오브젝트와 연관된 제1 세트의 좌표 및 상기 물리적 오브젝트와 연관된 제2 세트의 좌표에 기초한다. 몇몇 실시예에서, 상기 합성은 상기 물리적 환경에서 비롯된 오디오 신호 및 가상 오브젝트와 연관된 오디오 데이터의 렌더링을 포함한다.

[0016] 몇몇 실시예에 따르면, 상기 장치는 스마트폰에 연결하도록 구성된 커넥터를 추가로 포함한다. 몇몇 실시예에서, 상기 스마트폰은 상기 프로세싱 시스템의 적어도 부분을 포함한다. 몇몇 실시예에서, 상기 스마트폰은 부가적인 콘텐츠를 제공하도록 구성되며, 상기 렌더링 모듈은 상기 이미지 및 상기 부가적인 콘텐츠에 기초하여 상기 물리적 환경의 렌더링을 결정하도록 구성된다.

도면의 간단한 설명

[0017] 참조가 이제 본 출원의 예시적인 실시예를 도시하는 첨부한 도면에 대해 이루어질 것이다:

도 1은 본 개시의 실시예가 구현될 수 있는 대표적인 컴퓨팅 디바이스의 블록도이다.

도 2a 및 도 2b는 본 개시의 실시예와 일치하는, 몰입식 멀티미디어 생성을 예시한 대표적인 렌더링의 예시이다.

도 3은 본 개시의 실시예와 일치하는, 몰입식 및 대화식 멀티미디어 생성을 위한 대표적인 시스템의 블록도이다.

도 4a 내지 도 4e는 본 개시의 실시예와 일치하는, 몰입식 및 대화식 멀티미디어 생성을 지원하기 위한 대표적인 카메라 시스템의 개략도이다.

도 5는 본 개시의 실시예와 일치하는, 몰입식 및 대화식 멀티미디어 생성을 지원하기 위해 카메라의 위치 및 포즈를 감지하기 위한 대표적인 방법의 흐름도이다.

도 6은 본 개시의 실시예와 일치하는 손 제스처에 기초하여 멀티미디어 렌더링을 업데이트하기 위한 대표적인 방법의 흐름도이다.

도 7a 및 도 7b는 본 개시의 실시예와 일치하는 물리적 환경의 실시간 그래픽 이미지로의 3D 가상 오브젝트의 이미지의 합성의 예시이다.

도 8은 본 개시의 실시예와 일치하는 물리적 환경의 실시간 그래픽 이미지로의 3D 가상 오브젝트의 이미지의 합성을 위한 대표적인 방법의 흐름도이다.

도 9a 및 도 9b는 본 개시의 실시예와 일치하는 대표적인 헤드-장착 대화형 몰입식 멀티미디어 생성을 예시한 개략도이다.

발명을 실시하기 위한 구체적인 내용

[0018] 참조가 이제 실시예에 대해 상세히 이루어질 것이며, 그 예는 첨부한 도면에서 예시된다. 가능하다면, 동일한 참조 숫자는 도면 전체에 걸쳐 동일한 또는 유사한 부분을 나타내기 위해 사용될 것이다.

- [0019] 실시예의 설명은 단지 대표적이며, 제한적이지도록 의도되지 않는다.
- [0020] 도 1은 본 개시의 실시예가 구현될 수 있는 대표적인 컴퓨팅 디바이스(100)의 블록도이다. 도 1에 도시된 바와 같이, 컴퓨팅 디바이스(100)는 프로세서(121) 및 메인 메모리(122)를 포함한다. 프로세서(121)는 메인 메모리(122)로부터 인출된 지시에 응답하며 이를 프로세싱하는 임의의 논리 회로일 수 있다. 프로세서(221)는 단일 또는 다수의 범용 마이크로프로세서, 필드-프로그램 가능한 게이트 어레이(FPGA), 또는 메모리(예로서, 메인 메모리(122))에 저장된 지시를 실행할 수 있는 디지털 신호 프로세서(DSP), 또는 애플리케이션 특정 집적 회로(ASIC)일 수 있으며, 따라서 프로세서(221)는 특정한 태스크를 수행하도록 구성된다.
- [0021] 메모리(122)는 플렉시블 디스크, 하드 디스크, CD-ROM(컴팩트 디스크 판독-전용 메모리), MO(자기-광학) 드라이브, DVD-ROM(디지털 다목적 디스크 판독-전용 메모리), DVD-RAM(디지털 다목적 디스크 랜덤-액세스 메모리), 플래시 드라이브, 플래시 메모리, 레지스터, 캐시, 또는 반도체 메모리와 같은, 유형의 및/또는 비-일시적 컴퓨터-판독 가능한 매체를 포함한다. 메인 메모리(122)는 데이터를 저장하며 임의의 저장 위치가 프로세서(121)에 의해 직접 액세스되도록 허용할 수 있는 하나 이상의 메모리 칩일 수 있다. 메인 메모리(122)는 임의의 유형의 랜덤 액세스 메모리(RAM), 또는 여기에 설명된 바와 같이 동작할 수 있는 임의의 다른 이용 가능한 메모리 칩일 수 있다. 도 1에 도시된 대표적인 실시예에서, 프로세서(121)는 시스템 버스(150)를 통해 메인 메모리(122)와 통신한다.
- [0022] 컴퓨팅 디바이스(100)는 운영 시스템 및 다른 관련 소프트웨어를 저장하기 위해, 애플리케이션 소프트웨어 프로그램을 저장하기 위해, 및 애플리케이션 소프트웨어 프로그램에 의해 사용될 애플리케이션 데이터를 저장하기 위해, 하나 이상의 하드 디스크 드라이브와 같은, 저장 디바이스(128)를 추가로 포함할 수 있다. 예를 들면, 애플리케이션 데이터는 멀티미디어 데이터를 포함할 수 있지만, 소프트웨어는 멀티미디어 데이터를 렌더링하도록 구성된 렌더링 엔진을 포함할 수 있다. 소프트웨어 프로그램은 하나 이상의 지시를 포함할 수 있으며, 이것은 프로세서(121)에 의해 프로세싱되도록 저장 장치(128)로부터 메모리(122)로 인출될 수 있다. 소프트웨어 프로그램은 예로서, 소프트웨어 구성요소, 오브젝트-지향 소프트웨어 구성요소, 클래스 구성요소 및 태스크 구성요소, 프로세스, 함수, 필드, 프로시저, 서브루틴, 프로그램 코드의 세그먼트, 드라이버, 펌웨어, 마이크로코드, 회로, 데이터, 데이터베이스, 데이터 구조, 테이블, 어레이, 및 변수와 같은, 구성요소를 포함할 수 있는, 상이한 소프트웨어 모듈을 포함할 수 있다.
- [0023] 일반적으로, 여기에서 사용된 바와 같이, 단어 "모듈"은 하드웨어 또는 펌웨어로 구체화된 로직, 또는 예를 들면, 자바, Lua, C 또는 C++과 같은, 프로그래밍 언어로 기록된, 가능하게는 입구 및 출구 포인트를 갖는, 소프트웨어 지시의 모음을 나타낸다. 소프트웨어 모듈은 컴파일링되며 실행 가능한 프로그램으로 링킹되고, 동적 링크 라이브러리에 설치되거나, 또는 예를 들면, BASIC, Perl, 또는 Python과 같은 인터프리터식 프로그래밍 언어로 기록될 수 있다. 소프트웨어 모듈은 다른 모듈로부터 또는 자체로부터 호출 가능할 수 있으며, 및/또는 검출된 이벤트 또는 인터럽트에 응답하여 호출될 수 있다는 것이 이해될 것이다. 컴퓨팅 디바이스 상에서의 실행을 위해 구성된 소프트웨어 모듈은 콤팩트 디스크, 디지털 비디오 디스크, 플래시 드라이브, 자기 디스크, 또는 임의의 다른 유형의 매체와 같은, 컴퓨터 판독 가능한 매체상에, 또는 디지털 다운로드로서 제공될 수 있다(및 원래 실행 이전에 설치, 압축 해제, 또는 복호화를 요구하는 압축된 또는 설치 가능한 포맷으로 저장될 수 있다). 이러한 소프트웨어 코드는 컴퓨팅 디바이스에 의한 실행을 위해, 실행하는 컴퓨팅 디바이스의 메모리 디바이스 상에, 부분적으로 또는 전체적으로, 저장될 수 있다. 소프트웨어 지시는 EPROM과 같은, 펌웨어에 내장될 수 있다. 하드웨어 모듈(예로서, 프로세서(221)가 ASIC인 경우에)은 게이트 및 플립-플롭과 같은, 연결된 논리 유닛으로 구성될 수 있으며, 및/또는 프로그램 가능한 게이트 어레이 또는 프로세서와 같은, 프로그램 가능한 유닛으로 구성될 수 있다는 것이 추가로 이해될 것이다. 여기에서 설명된 모듈 또는 컴퓨팅 디바이스 기능은 바람직하게는 소프트웨어 모듈로 구현되지만, 하드웨어 또는 펌웨어로 나타내어질 수 있다. 일반적으로, 여기에서 설명된 모듈은 그것들의 물리적 조직 또는 저장 장치에도 불구하고 다른 모듈과 조합되거나 또는 서브-모듈로 분할될 수 있는 논리 모듈을 나타낸다.
- [0024] 여기에서 사용된 바와 같이, 용어 "비-일시적 미디어"는 기계가 특정 방식으로 동작하게 하는 데이터 및/또는 지시를 저장하는 임의의 비-일시적 미디어를 나타낸다. 이러한 비-일시적 미디어는 비-휘발성 미디어 및/또는 휘발성 미디어를 포함할 수 있다. 비-휘발성 미디어는, 예를 들면, 저장 장치(128)를 포함할 수 있다. 휘발성 미디어는, 예를 들면, 메모리(122)를 포함할 수 있다. 비-일시적 미디어의 보통의 형태는, 예를 들면, 플로피 디스크, 플렉시블 디스크, 하드 디스크, 고체 상태 드라이브, 자기 테이프, 또는 임의의 다른 자기 데이터 저장 매체, CD-ROM, 임의의 다른 광학 데이터 저장 매체, 홀의 패턴을 가진 임의의 물리적 매체, RAM, PROM, 및

EPROM, FLASH-EPROM, NVRAM, 임의의 다른 메모리 칩 또는 카트리지, 및 이것들의 네트워킹된 버전을 포함한다.

- [0025] 컴퓨팅 디바이스(100)는 또한 하나 이상의 입력 디바이스(123) 및 하나 이상의 출력 디바이스(124)를 포함할 수 있다. 입력 디바이스(123)는, 예를 들면, 카메라, 마이크로폰, 모션 센서 등을 포함할 수 있지만, 출력 디바이스(124)는, 예를 들면, 디스플레이 유닛 및 스피커를 포함할 수 있다. 입력 디바이스(123) 및 출력 디바이스(124) 양쪽 모두는 I/O 제어기(125)를 통해 시스템 버스(150)에 연결되어, 프로세서(121)가 입력 디바이스(123) 및 출력 디바이스(124)와 통신할 수 있게 한다. 프로세서(121) 및 입력 디바이스(123) 및 출력 디바이스(124) 중에서의 통신은, 예를 들면, 메모리(122)로부터 인출된 지시를 실행하는 프로세서(121)에 의해 수행될 수 있다.
- [0026] 몇몇 실시예에서, 프로세서(121)는 또한 I/O 제어(125)를 통해 하나 이상의 스마트 디바이스(130)와 통신할 수 있다. 스마트 디바이스(130)는 멀티미디어 데이터를 프로세싱하며 생성하는 능력을 포함하는 시스템(예로서, 스마트폰)을 포함할 수 있다. 몇몇 실시예에서, 프로세서(121)는 입력 디바이스(123)로부터 데이터를 수신하고, 프로세싱을 위해 상기 데이터를 스마트 디바이스(130)로 인출하고, 프로세싱의 결과로서 멀티미디어 데이터(예를 들면, 오디오 신호, 비디오 신호 등의 형태로)를 스마트 디바이스(130)로부터 수신하며, 그 후 멀티미디어 데이터를 출력 디바이스(124)에 제공할 수 있다. 몇몇 실시예에서, 스마트 디바이스(130)는 멀티미디어 콘텐츠의 소스로서 동작하며 멀티미디어 콘텐츠에 관련된 데이터를 프로세서(121)에 제공할 수 있다. 프로세서(121)는 그 후 출력 디바이스(124)로 제공될 데이터를 출력하기 위해 스마트 디바이스(130)로부터 수신된 멀티미디어 콘텐츠를 부가할 수 있다. 프로세서(121) 및 스마트 디바이스(130) 사이에서의 통신은 예를 들면, 메모리(122)로부터 인출된 지시를 실행하는 프로세서(121)에 의해 구현될 수 있다.
- [0027] 몇몇 실시예에서, 컴퓨팅 디바이스(100)는 가상 현실, 증강 현실, 또는 양쪽 모두의 조합을 포함하여, 대화식 및 몰입식 멀티미디어를 생성하도록 구성될 수 있다. 예를 들면, 저장 장치(128)는 가상 현실 경험의 생성을 위해 그래픽 이미지 및 오디오 효과의 렌더링을 위해 멀티미디어 데이터를 저장할 수 있으며, 프로세서(121)는 가상 현실 경험을 생성하기 위해 출력 디바이스(124)를 통해 멀티미디어 데이터의 적어도 부분을 제공하도록 구성될 수 있다. 프로세서(121)는 또한 프로세서(121)가 예를 들면, 사용자의 위치에서의 변화, 사용자에 의해 수행된 동작(예로서, 신체 움직임) 등을 결정할 수 있게 하는 입력 디바이스(123)(예로서, 모션 센서)로부터 수신된 데이터를 수신할 수 있다. 프로세서(121)는 결정에 기초하여, 사용자에게 대한 대화식 경험을 생성하기 위해, 출력 디바이스(124)를 통해 멀티미디어 데이터를 렌더링하도록 구성될 수 있다.
- [0028] 게다가, 컴퓨팅 디바이스(100)는 또한 증강 현실을 제공하도록 구성될 수 있다. 예를 들면, 입력 디바이스(123)는 사용자가 위치되는 물리적 환경의 그래픽 이미지를 캡처하도록 구성된 하나 이상의 카메라, 및 물리적 환경으로부터 오디오 신호를 캡처하도록 구성된 하나 이상의 마이크로폰을 포함할 수 있다. 프로세서(121)는 카메라로부터 캡처된 그래픽 이미지 및 오디오 정보를 나타내는 데이터를 수신할 수 있다. 프로세서(121)는 사용자에게 제공될 부가적인 콘텐츠를 나타내는 데이터를 또한 프로세싱할 수 있다. 부가적인 콘텐츠는 예를 들면, 물리적 환경의 그래픽 이미지로부터 검출된 하나 이상의 오브젝트에 관련된 정보일 수 있다. 프로세서(121)는 증강 현실 경험을 생성하기 위해, 출력 디바이스(124)를 통해, 캡처된 그래픽 이미지, 오디오 정보, 뿐만 아니라 부가적인 콘텐츠를 포함하는 멀티미디어 데이터를 렌더링하도록 구성될 수 있다. 부가적인 콘텐츠를 나타내는 데이터는 저장 장치(128)에 저장될 수 있거나, 또는 외부 소스(예로서, 스마트 디바이스(130))에 의해 제공될 수 있다.
- [0029] 프로세서(121)는 또한 예를 들면, 사용자 동작에 대한 정보를 획득함으로써 사용자에게 대한 대화식 경험을 생성하도록 구성될 수 있으며 출력 디바이스(124)를 통한 멀티미디어 데이터의 렌더링은 사용자 동작에 기초하여 이루어질 수 있다. 몇몇 실시예에서, 사용자 동작은 사용자의 위치의 변화를 포함할 수 있으며, 이것은 예를 들면, 모션 센서로부터의 데이터, 및 그래픽 이미지로부터의 두드러진 특징의 추적에 기초하여 프로세서(121)에 의해 결정될 수 있다. 몇몇 실시예에서, 사용자 동작은 또한 손 제스처를 포함할 수 있으며, 이것은 카메라에 의해 캡처된 손 제스처의 이미지에 기초하여 프로세서(121)에 의해 결정될 수 있다. 프로세서(121)는 위치 정보 및/또는 손 제스처 정보에 기초하여, 대화식 경험을 생성하기 위해 멀티미디어 데이터의 렌더링을 업데이트하도록 구성될 수 있다. 몇몇 실시예에서, 프로세서(121)는 또한 사용자의 손 제스처에 기초하여, 예를 들면, 물리적 환경에서 특정 위치로 줌하는 것, 상기 특정 위치에서 비롯된 오디오 신호의 볼륨을 증가시키는 것 등에 의해 사용자의 감지 능력을 강화하기 위해 멀티미디어 데이터의 렌더링을 업데이트하도록 구성될 수 있다.
- [0030] 참조가 이제 도 2a 및 도 2b에 대해 이루어지며, 이것은 본 개시의 실시예와 일치하는 증강 현실을 제공하기 위한 대표적인 멀티미디어 렌더링(200a 및 200b)을 예시한다. 도 2a 및 도 2b에 도시된 바와 같이, 렌더링(200a

및 200b)은 사용자가 위치되는 물리적 환경의 그래픽 표현을 반영한다. 몇몇 실시예에서, 렌더링(200a 및 200 b)은 하나 이상의 카메라(예로서, 입력 디바이스(123))에 의해 캡처된 그래픽 이미지에 기초하여 컴퓨팅 디바이스(100)의 프로세서(221)에 의해 구성될 수 있다. 프로세서(221)는 또한 그래픽 이미지로부터 손 제스처를 검출하며, 손 제스처에 관련된 부가적인 콘텐츠를 포함하기 위해 렌더링을 업데이트하도록 구성될 수 있다. 예시적인 예로서, 도 2a 및 도 2b에 도시된 바와 같이, 렌더링(200a 및 200b)은, 각각, 손 제스처의 생성에 수반되는 손가락의 움직임을 나타내는 점선(204a 및 204b)을 포함할 수 있다. 몇몇 실시예에서, 검출된 손 제스처는 사용자의 감지 능력(예로서, 시각)을 강화하기 위해 그래픽 이미지의 부가적인 프로세싱을 트리거할 수 있다. 예시적인 예로서, 도 2a에 도시된 바와 같이, 렌더링(200a)에서 렌더링된 물리적 환경은 오브젝트(204)를 포함한다. 오브젝트(204)는 제1 손 제스처의 검출, 및 제1 손 제스처를 생성하는 손가락의 움직임 사이에서의 중첩(예로서, 점선(202a)에 의해 표시된 바와 같이)에 기초하여 선택될 수 있다. 중첩은, 예를 들면, 물리적 환경을 나타내는 3D 맵에서 점선(202a)의 3D 좌표 및 오브젝트(204)의 3D 좌표 사이의 관계에 기초하여 결정될 수 있다.

[0031] 오브젝트(204)가 선택된 후, 사용자는 제2 손 제스처(점선(202b)에 의해 표시된 바와 같이)를 제공할 수 있으며, 이것은 또한 프로세서(221)에 의해 검출될 수 있다. 프로세서(221)는, 시간적 및 공간적으로 근접하여 발생하는 두 개의 손 제스처의 검출에 기초하여, 제2 손 제스처가 물리적 환경의 렌더링에서 오브젝트(204)의 확장되고 확대된 이미지를 제공하도록 프로세서(221)에 지시하는 것임을 결정할 수 있다. 이것은 렌더링(200b)을 야기할 수 있으며, 여기에서 오브젝트(204)의 확장되고 확대된 이미지를 나타내는 이미지(206)는 사용자가 위치되는 물리적 환경과 함께, 렌더링된다. 오브젝트의 확대된 이미지를 사용자에게 제공하며, 그에 의해 사용자로 하여금 그 또는 그녀가 물리적 환경 내에서의 동일한 위치에서 육안으로 지각하는 것보다 오브젝트에 대해 더 상세히 지각하도록 허용함으로써, 사용자의 감각 능력이 강화될 수 있다.

[0032] 몇몇 실시예에서, 오브젝트(204)는 또한 물리적 환경의 렌더링에서 삽입된 가상 오브젝트일 수 있으며, 이미지(206)는 오브젝트(204)의 선택 및 점선(202b)에 의해 표현된 손 제스처의 검출에 응답하여 제공된 임의의 이미지(또는 물리적 환경의 렌더링의 최상부 상에 오버레이한 텍스트)일 수 있다.

[0033] 다시 도 1을 참조하면, 몇몇 실시예에서, 컴퓨팅 디바이스(100)는 또한 이에 제한되지 않지만, 표준 전화 라인, LAN 또는 WAN 링크(예로서, 802.11, T1, T3, 56 kb, X.25), 광대역 링크(예로서, ISDN, 프레임 릴레이, ATM), 무선 연결(Wi-Fi, 블루투스, Z-Wave, 지그비(Zigbee)), 또는 상기 중 임의의 것 또는 모두의 몇몇 조합을 포함하는, 다양한 링크를 통해 LAN, WAN, MAN, 또는 인터넷으로의 인터페이스에 대한 네트워크 인터페이스(140)를 포함할 수 있다. 네트워크 인터페이스(140)는 내장형 네트워크 어댑터, 네트워크 인터페이스 카드, PCMCIA 네트워크 카드, 카드 버스 네트워크 어댑터, 무선 네트워크 어댑터, USB 네트워크 어댑터, 모뎀 또는 통신이 가능한 임의의 유형의 네트워크로 컴퓨팅 디바이스(100)를 인터페이싱하며 여기에서 설명된 동작을 수행하는데 적합한 임의의 다른 디바이스를 포함할 수 있다. 몇몇 실시예에서, 프로세서(121)는 네트워크 인터페이스(140)를 통해 출력 디바이스(124)로 뿐만 아니라 또한 다른 디바이스(예로서, 또 다른 컴퓨팅 디바이스(100))로 생성된 멀티미디어 데이터를 송신할 수 있다.

[0034] 도 3은 본 개시의 실시예와 일치하는, 몰입식 및 대화식 멀티미디어 생성을 위한 대표적인 시스템(300)의 블록도이다. 도 3에 도시된 바와 같이, 시스템(300)은 감지 시스템(310), 프로세싱 시스템(320), 오디오/비디오 시스템(330), 및 전력 시스템(340)을 포함한다. 몇몇 실시예에서, 시스템(300)의 적어도 부분은 도 1의 컴퓨팅 디바이스(100)로 구현된다.

[0035] 몇몇 실시예에서, 감지 시스템(310)은 대화식 및 몰입식 멀티미디어의 생성을 위한 데이터를 제공하도록 구성된다. 감지 시스템(310)은 이미지 감지 시스템(312), 오디오 감지 시스템(313), 및 모션 감지 시스템(314)을 포함한다.

[0036] 몇몇 실시예에서, 광학 감지 시스템(312)은 물리적 환경으로부터 반사되거나 또는 방출된 다양한 파장의 광(가시적 및 비가시적인 광 양쪽 모두를 포함한다)을 수신하도록 구성될 수 있다. 몇몇 실시예에서, 광학 감지 시스템(312)은, 예를 들면, 하나 이상의 그레이스케일-적외선(그레이스케일 IR) 카메라, 하나 이상의 적색-녹색-청색(RGB) 카메라, 하나 이상의 RGB-IR 카메라, 하나 이상의 비행시간(TOF) 카메라, 또는 그것들의 조합을 포함한다. 카메라의 출력에 기초하여, 시스템(300)은 물리적 환경의 이미지 데이터(예로서, RGB 픽셀 및 IR 픽셀의 형태로 표현된)를 획득할 수 있다. 광학 감지 시스템(312)은 동일한 카메라의 쌍(예로서, RGB 카메라의 쌍, IR 카메라의 쌍, RGB-IR 카메라의 쌍 등)를 포함할 수 있으며, 각각의 카메라는 좌측 눈 또는 우측 눈의 시점을 캡처한다. 이하에서 논의될 바와 같이, 각각의 카메라에 의해 캡처된 이미지 데이터는 그 후 물리적 환경

의 입체 3D 렌더링을 생성하기 위해 시스템(300)에 의해 조합될 수 있다.

- [0037] 몇몇 실시예에서, 광학 감지 시스템(312)은 오브젝트를 조명하도록 구성된 IR 투사기를 포함할 수 있다. 조명은 레인지 이미징(range imaging)을 지원하기 위해 사용될 수 있으며, 이것은 시스템(300)으로 하여금, 스테레오 매칭 알고리즘에 또한 기초하여, 물리적 환경에서 오브젝트의 상이한 부분 및 카메라 사이의 거리를 결정할 수 있게 한다. 거리 정보에 기초하여, 오브젝트의 3-차원(3D) 깊이 맵, 뿐만 아니라 물리적 환경의 3D 맵이 생성될 수 있다. 이하에서 논의될 바와 같이, 오브젝트의 깊이 맵은 오브젝트를 나타내는 3D 포인트 클라운드를 생성하기 위해 사용될 수 있고; 오브젝트의 RGB 데이터는, RGB 카메라에 의해 캡처된 대로, 그 후 가상 현실 및 증강 현실 효과를 생성하기 위해 오브젝트의 3D 렌더링을 생성하도록 3D 포인트 클라운드에 매핑될 수 있다. 다른 한편으로, 물리적 환경의 3D 맵은 상호 작용적 경험을 생성하기 위해 위치 및 방위 결정을 위해 사용될 수 있다. 몇몇 실시예에서, 비행시간 카메라는 또한 레인지 이미징을 위해 포함될 수 있으며, 이것은 카메라 및 오브젝트의 다양한 부분 사이의 거리가 결정되도록 허용하며, 물리적 환경의 깊이 맵은 거리 정보에 기초하여 생성될 수 있다.
- [0038] 몇몇 실시예에서, IR 투사기는 또한 물리적 환경의 하나 이상의 표면으로 특정한 패턴(예로서, 바 코드, 코너 패턴 등)을 투사하도록 구성된다. 이하에서 논의될 바와 같이, 사용자의 모션(카메라의 모션에 의해 반영됨)은 카메라에 의해 캡처된 다양한 두드러진 특징 포인트를 추적함으로써 결정될 수 있으며, 알려진 패턴의 투사(그 후 카메라에 의해 캡처되며 시스템에 의해 추적되는)는 효율적이며 강력한 추적을 가능하게 한다.
- [0039] 참조가 이제 도 4a 내지 도 4f에 대해 이루어지며, 이것은 각각, 본 개시의 실시예와 일치하는 대표적인 카메라 시스템(400, 420, 440, 460, 480 및 494)을 예시한 개략도이다. 도 4a 내지 도 4f의 각각의 카메라 시스템은 도 3의 광학 감지 시스템(312)의 부분일 수 있다.
- [0040] 도 4a에 도시된 바와 같이, 카메라 시스템(400)은 RGB 카메라(402), IR 카메라(404), 및 IR 조명기(406)를 포함하며, 그 모두는 보드(408)로 부착된다. 상기 논의된 바와 같이, RGB 카메라(402)는 RGB 이미지 데이터를 캡처하도록 구성되고, IR 카메라(404)는 IR 이미지 데이터를 캡처하도록 구성되지만, IR 카메라(404) 및 IR 조명기(406)의 조합은 이미징되는 오브젝트의 깊이 맵을 생성하기 위해 사용될 수 있다. 이전에 논의된 바와 같이, 오브젝트의 3D 렌더링 동안, RGB 이미지 데이터는 깊이 맵으로부터 생성된 오브젝트의 3D 포인트 클라운드를 표현에 매핑될 수 있다. 그러나, 몇몇 경우에서, RGB 카메라 및 IR 카메라 사이의 위치 차이로 인해, RGB 이미지 데이터에서의 RGB 픽셀 모두가 3D 포인트 클라운드에 매핑될 수 있는 것은 아니다. 그 결과, 부정확성 및 불일치가 오브젝트의 3D 렌더링에서 도입될 수 있다.
- [0041] 도 4b는 RGB-IR 카메라(422) 및 IR 조명기(424)를 포함하는, 카메라 시스템(420)을 예시하며, 그 모두는 보드(426)로 부착된다. RGB-IR 카메라(422)는 픽셀 그룹을 형성하기 위해 함께 섞인 RGB 및 IR 픽셀 센서를 포함하는 RGB-IR 센서를 포함한다. 실질적으로 같은 장소에 배치된 RGB 및 IR 픽셀 센서를 갖고, RGB 및 IR 센서 사이의 위치 차이의 앞서 언급한 효과가 제거될 수 있다. 그러나, 몇몇 경우에서, RGB 스펙트럼의 부분 및 IR 스펙트럼의 부분의 중첩으로 인해, 같은 장소에 배치된 RGB 및 IR 픽셀 센서를 갖는 것은 RGB 픽셀 센서에 의해 생성된 컬러 이미지 품질뿐만 아니라 RGB 픽셀 센서의 컬러 생성의 열화를 야기할 수 있다.
- [0042] 도 4c는 IR 카메라(442), RGB 카메라(444), 미러(446), 및 IR 조명기(448)를 포함하는, 카메라 시스템(440)을 예시하며, 그 모두는 보드(450)에 부착된다. 몇몇 실시예에서, 미러(446)는 IR 반사성 코팅(452)을 갖고 반투명이다. 광(가시 광, 및 IR 조명기(448)에 의해 조명된 오브젝트에 의해 반사된 IR 광을 포함한)이 미러(446)에 입사함에 따라, IR 광은 미러(446)에 의해 반사되며 IR 카메라(442)에 의해 캡처될 수 있지만, 가시 광은 미러(446)를 통과하며 RGB 카메라(444)에 의해 캡처될 수 있다. IR 카메라(442), RGB 카메라(444), 및 미러(446)는 IR 카메라(442)에 의해 캡처된(IR 반사성 코팅에 의한 반사에 의해 야기된) IR 이미지 및 RGB 카메라(444)에 의해 캡처된(미러(446)를 통과하는 가시 광으로부터) RGB 이미지가 IR 카메라(442) 및 RGB 카메라(444) 사이의 위치 차이의 효과를 제거하기 위해 정렬될 수 있도록 배치될 수 있다. 게다가 IR 광은 RGB 카메라(444)로부터 멀리 반사되므로, RGB 카메라(444)에 의해 생성된 컬러 이미지 품질뿐만 아니라 컬러 제품이 개선될 수 있다.
- [0043] 도 4d는 RGB 카메라(462), TOF 카메라(464), 및 IR 조명기(466)를 포함하는 카메라 시스템(460)을 예시하며, 그 모두는 보드(468)로 부착된다. 카메라 시스템(400, 420, 및 440)과 유사하게, RGB 카메라(462)는 RGB 이미지 데이터를 캡처하도록 구성된다. 다른 한편으로, TOF 카메라(464) 및 IR 조명기(466)는 이미지-레이징을 수행하기 위해 동기화되며, 이것은 이미징되는 오브젝트의 깊이 맵을 생성하기 위해 사용될 수 있으며, 그로부터 오브젝트의 3D 포인트 클라운드가 생성될 수 있다. 카메라 시스템(400)과 유사하게, 몇몇 경우에서, RGB 카메라 및

TOF 카메라 사이에서의 위치 차이로 인해, RGB 이미지 데이터에서의 RGB 픽셀 모두가 TOF 카메라의 출력에 기초하여 생성된 3D 포인트 클라우드에 매핑될 수 있는 것은 아니다. 그 결과, 부정확성 및 불일치가 오브젝트의 3D 렌더링에 도입될 수 있다.

[0044] 도 4e는 TOF 카메라(482), RGB 카메라(484), 미러(486), 및 IR 조명기(488)를 포함하는 카메라 시스템(480)을 예시하며, 그 모두는 보드(490)에 부착된다. 몇몇 실시예에서, 미러(486)는 IR 반사성 코팅(492)을 갖고 반투명이다. 광(가시 광, 및 IR 조명기(488)에 의해 조명된 오브젝트에 의해 반사된 IR 광을 포함한)이 미러(446)에 입사함에 따라, IR 광은 미러(446)에 의해 반사되며 TOF 카메라(482)에 의해 캡처될 수 있지만, 가시 광은 미러(486)를 통과하며 RGB 카메라(484)에 의해 캡처될 수 있다. TOF 카메라(482), RGB 카메라(484), 및 미러(486)는 TOF 카메라(442)에 의해 캡처된(IR 반사성 코팅에 의한 반사에 의해 야기된) IR 이미지 및 RGB 카메라(484)에 의해 캡처된(미러(486)를 통과하는 가시 광으로부터) RGB 이미지가 TOF 카메라(482) 및 RGB 카메라(484) 사이에서의 위치 차이의 효과를 제거하기 위해 정렬될 수 있도록 배치될 수 있다. 게다가, IR 광은 RGB 카메라(484)로부터 멀리 반사되므로, RGB 카메라(484)에 의해 생성된 컬러 이미지 품질뿐만 아니라 컬러 제품이 또한 개선될 수 있다.

[0045] 도 4f는, 각각이 인간 눈의 시점을 모방하도록 구성되는, 두 개의 RGB-IR 카메라(495 및 496)를 포함하는 카메라 시스템(494)을 예시한다. RGB-IR 카메라(495 및 496)의 조합은, 이하에서 논의될 바와 같이, 입체 이미지를 생성하기 위해 및 물리적 환경에서 오브젝트의 깊이 정보를 생성하기 위해 사용될 수 있다. 카메라의 각각은 같은 장소에 배치된 RGB 및 IR 픽셀을 가지므로, 픽셀 매핑에서의 열화를 야기하는 RGB 카메라 및 IR 카메라 사이에서의 위치 차이의 효과가 완화될 수 있다. 카메라 시스템(494)은 상기 논의된 다른 IR 조명기와 유사한 기능을 가진 IR 조명기(497)를 추가로 포함한다. 도 4f에 도시된 바와 같이, RGB-IR 카메라(495 및 496) 및 IR 조명기(497)는 보드(498)에 부착된다.

[0046] 다시 도 3을 참조하면, 감지 시스템(310)은 또한 오디오 감지 시스템(313) 및 모션 감지 시스템(314)을 포함한다. 오디오 감지 시스템(313)은 물리적 환경에서 비롯된 오디오 신호를 수신하도록 구성될 수 있다. 몇몇 실시예에서, 오디오 감지 시스템(313)은, 예를 들면, 하나 이상의 마이크로폰 어레이를 포함한다. 모션 감지 시스템(314)은 사용자의(및 시스템이 사용자에게 부착된다면, 시스템의) 모션 및/또는 포즈를 검출하도록 구성될 수 있다. 몇몇 실시예에서, 모션 감지 시스템(314)은, 예를 들면, 관성 모션 센서(IMU)를 포함할 수 있다. 몇몇 실시예에서, 감지 시스템(310)은 도 1의 입력 디바이스(123)의 부분일 수 있다.

[0047] 몇몇 실시예에서, 프로세싱 시스템(320)은 광학 감지 시스템(312)으로부터의 그래픽 이미지 데이터, 오디오 감지 시스템(313)으로부터의 오디오 데이터, 및 모션 감지 시스템(314)으로부터의 모션 데이터를 프로세싱하며, 가상 현실 및/또는 증강 현실 경험을 생성하기 위해 물리적 환경을 렌더링하기 위한 멀티미디어 데이터를 생성하도록 구성된다. 프로세싱 시스템(320)은 방위 및 위치 결정 모듈(322), 손 제스처 결정 시스템 모듈(323), 및 그래픽 및 오디오 렌더링 엔진 모듈(324)을 포함한다. 이전에 논의된 바와 같이, 이들 모듈의 각각은 프로세서(예로서, 도 1의 프로세서(121))에 의해 실행되는 소프트웨어 모듈, 또는 특정 기능을 수행하도록 구성된 하드웨어 모듈(예로서, ASIC)일 수 있다.

[0048] 몇몇 실시예에서, 방위 및 위치 결정 모듈(322)은 감지 시스템(310)의 출력 중 적어도 일부에 기초하여 사용자의 방위 및 위치를 결정할 수 있으며, 그것에 기초하여 멀티미디어 데이터가 가상 현실 및/또는 증강 현실 효과를 생성하기 위해 렌더링될 수 있다. 시스템(300)이 사용자에게 의해 착용되는 경우에(예로서, 고글), 방위 및 위치 결정 모듈(322)은 시스템(예로서, 카메라)의 부분의 방위 및 위치를 결정할 수 있으며, 이것은 사용자의 방위 및 위치를 추론하기 위해 사용될 수 있다. 결정된 방위 및 위치는 움직임이 발생하기 전에 사용자의 이전 방위 및 위치에 상대적일 수 있다.

[0049] 참조가 이제 도 5에 대해 이루어지며, 이것은 본 개시의 실시예와 일치하는 카메라의(예로서, 감지 시스템(310)) 쌍의 방위 및 위치를 결정하기 위한 대표적인 방법(500)을 예시하는 흐름도이다. 예시된 절차는 단계를 삭제하거나 또는 부가적인 단계를 추가로 포함하기 위해 변경될 수 있다는 것이 쉽게 이해될 것이다. 방법(500)은 프로세서(예로서, 방위 및 위치 결정 모듈(322))에 의해 수행되는 것으로 설명되지만, 방법(500)은 단독으로 또는 프로세서와 조합하여 다른 디바이스에 의해 수행될 수 있다는 것이 이해된다.

[0050] 단계(502)에서, 프로세서는 제1 카메라로부터 제1 좌측 이미지를 및 제2 카메라로부터 제1 우측 이미지를 획득할 수 있다. 좌측 카메라는, 예를 들면, 도 4f의 RGB-IR 카메라(495)일 수 있지만, 우측 카메라는, 예를 들면, 도 4f의 RGB-IR 카메라(496)일 수 있다. 제1 좌측 이미지는 사용자의 좌측 눈으로부터의 물리적 환경의 시점을 나타낼 수 있지만, 제1 우측 이미지는 사용자의 우측 눈으로부터의 물리적 환경의 시점을 나타낼 수 있다. 양쪽

이미지 모두는 IR 이미지, RGB 이미지, 또는 양쪽 모두의 조합(예로서, RGB-IR)일 수 있다.

- [0051] 단계(504)에서, 프로세서는 제1 좌측 이미지 데이터로부터 및 우측 이미지 데이터로부터 제1 두드러진 특징 포인트의 세트를 식별할 수 있다. 몇몇 경우에서, 두드러진 특징들은 물리적 환경에서 이전부터 존재하는 물리적 특징일 수 있으며(예로서, 벽 상에서의 특징 표시, 의류의 특징 등), 두드러진 특징은 이들 특징과 연관된 RGB 픽셀 및/또는 IR 픽셀에 기초하여 식별된다. 몇몇 경우에서, 두드러진 특징은 물리적 환경의 하나 이상의 표면으로 특정 IR 패턴(예로서, 점)을 투사하는 IR 조명기(예로서, 도 4f의 IR 조명기(497))에 의해 생성될 수 있다. 하나 이상의 표면은 카메라로 IR을 반사할 수 있다. 이전에 논의된 바와 같이, 이들 IR 패턴은 고르게 분포되는 것과 같은, 효율적인 검출 및 추적을 위해 설계될 수 있으며 날카로운 에지 및 코너를 포함할 수 있다. 몇몇 경우에서, 두드러진 특징은 물리적 환경 내에서의 특징한 위치에 고정되며 환경 내에서 IR 패턴을 투사하는 하나 이상의 IR 투사기를 위치시킴으로써 생성될 수 있다.
- [0052] 단계(506)에서, 프로세서는 삼각측량에 대한 스테레오 제약에 기초하여 식별된 제1 두드러진 특징으로부터 대응하는 쌍을 찾을 수 있다. 스테레오 제약은, 예를 들면, 스테레오 속성, 시차에 대한 허용량 한계 등에 기초하여 제1 두드러진 특징의 대응하는 쌍에 대한 각각의 이미지 내에서의 탐색 범위를 제한하는 것을 포함할 수 있다. 대응하는 쌍의 식별은 후보 특징의 IR 픽셀, 후보 특징의 RGB 픽셀, 및/또는 양쪽 모두의 조합에 기초하여 이루어질 수 있다. 제1 두드러진 특징의 대응하는 쌍이 식별된 후, 좌측 및 우측 이미지 내에서의 그것들의 위치 차이가 결정될 수 있다. 위치 차이 및 제1 및 제2 카메라 사이에서의 거리에 기초하여, 제1 두드러진 특징(그것들이 물리적 환경에서 나타나는 바와 같이) 및 제1 및 제2 카메라 사이에서의 거리가 선형 삼각측량을 통해 결정될 수 있다.
- [0053] 단계(508)에서, 선형 삼각측량에 의해 결정된 제1 두드러진 특징 및 제1 및 제2 카메라 사이에서의 거리, 및 좌측 및 우측 이미지에서의 제1 두드러진 특징의 위치에 기초하여, 프로세서는 제1 두드러진 특징의 하나 이상의 3D 좌표를 결정할 수 있다.
- [0054] 단계(510)에서, 프로세서는 물리적 환경을 나타내는 3D 맵에서, 단계(508)에서 결정된 제1 두드러진 특징의 3D 좌표 및 제1 두드러진 특징에 대한 정보를 추가하거나 또는 업데이트할 수 있다. 업데이트는, 예를 들면, 동시 위치 및 매핑 알고리즘(SLAM)에 기초하여 수행될 수 있다. 저장된 정보는, 예를 들면, 제1 두드러진 특징과 연관된 IR 픽셀 및 RGB 픽셀 정보를 포함할 수 있다.
- [0055] 단계(512)에서, 카메라의 움직임(예로서, 카메라를 운반하는 사용자의 움직임에 의해 야기된) 후, 프로세서는 제2 좌측 이미지 및 제2 우측 이미지를 획득하며, 움직임 후 제2 좌측 및 우측 이미지로부터 제2 두드러진 특징을 식별할 수 있다. 식별 프로세스는 단계(504)와 유사할 수 있다. 식별되는 제2 두드러진 특징은 제2 좌측 이미지와 연관된 제1 2D 공간 내에서 및 제2 우측 이미지와 연관된 제2 2D 공간 내에서 2D 좌표와 연관된다.
- [0056] 단계(514)에서, 프로세서는 제1 두드러진 특징의 3D 좌표(단계(508)에서 결정된)를 제1 및 제2 2D 공간으로 재투사할 수 있다.
- [0057] 단계(516)에서, 프로세서는 예를 들면, 위치 근접성, 특징 근접성, 및 스테레오 제약에 기초하여 제1 두드러진 특징에 대응하는 제2 두드러진 특징 중 하나 이상을 식별할 수 있다.
- [0058] 단계(518)에서, 프로세서는 제1 두드러진 특징의 재투사된 위치 및 제1 및 제2 2D 공간의 각각에서 제2 두드러진 특징의 2D 좌표 사이에서의 거리를 결정할 수 있다. 움직임 전 및 후 제1 및 제2 카메라의 상대적 3D 좌표 및 방위는 그 후 예를 들면, 그에 따라 결정된 3D 좌표 및 방위의 세트가 제1 및 제2 2D 공간의 양쪽 모두에서의 거리를 최소화하도록 거리에 기초하여 결정될 수 있다.
- [0059] 몇몇 실시예에서, 방법(500)은 프로세서가 좌측 및 우측 이미지 사이의 두드러진 특징의 위치 차이를 최소화하기 위해 3D 맵에서 두드러진 특징의 좌표의 번들 조정을 수행할 수 있는 단계(도 5에 도시되지 않음)를 추가로 포함한다. 조정은 방법(500)의 단계 중 임의의 것과 동시에 수행될 수 있으며, 단지 키프레임에 대해서만 수행될 수 있다.
- [0060] 몇몇 실시예에서, 프로세서는 또한 방법(500)의 수행을 용이하게 하기 위해 우리의 입력 디바이스로부터의 데이터를 사용할 수 있다. 예를 들면, 프로세서는 하나 이상의 모션 센서(예로서, 모션 감지 시스템(314))로부터 데이터를 획득할 수 있으며, 그로부터 프로세서는 카메라의 모션이 발생하였음을 결정할 수 있다. 이러한 결정에 기초하여, 프로세서는 단계(512)를 실행할 수 있다. 몇몇 실시예에서, 프로세서는 또한 단계(518)에서 카메라의 위치 및 방위의 산출을 용이하게 하기 위해 모션 센서로부터의 데이터를 사용할 수 있다.

- [0061] 다시 도 3을 참조하면, 프로세싱 시스템(320)은 손 제스처 결정 모듈(323)을 추가로 포함한다. 몇몇 실시예에서, 손 제스처 결정 모듈(323)은 광학 감지 시스템(312)으로부터의 그래픽 이미지 데이터로부터 손 제스처를 검출할 수 있다. 손 제스처 정보의 기술은 2013년 9월 23일에 출원된 미국 출원 번호 제14/034,286호, 및 2014년 8월 18일에 출원된, 미국 출원 번호 제14/462,324호에서 설명된 것에 관련된다. 상기 참조된 출원은 참조로서 여기에 통합된다. 그에 따라 결정된 손 제스처 정보는, 도 2a 및 도 2b에서 전에 논의된 바와 같이, 부가적인 콘텐츠를 제공하기 위해 및/또는 사용자의 감각 능력을 강화하기 위해 물리적 환경의 렌더링(그래픽 및 오디오 양쪽 모두)을 업데이트하기 위해 사용될 수 있다. 예를 들면, 몇몇 실시예에서, 손 제스처 결정 모듈(323)은 손 제스처와 연관된 해석을 결정하며(예로서, 줌 인을 위한 오브젝트를 선택하기 위해), 그 후 렌더링을 업데이트하기 위해 다운스트림 로직(예로서, 그래픽 및 오디오 렌더링 모듈(324))으로 해석 및 다른 관련 정보를 제공할 수 있다.
- [0062] 참조가 이제 도 6에 대해 이루어지며, 이것은 본 개시의 실시예와 일치하는 검출된 손 제스처에 기초하여 멀티미디어 렌더링을 업데이트하기 위한 대표적인 방법(600)을 예시하는 흐름도이다. 예시된 절차는 단계를 삭제하거나 또는 부가적인 단계를 추가로 포함하기 위해 변경될 수 있다는 것이 쉽게 이해될 것이다. 방법(600)은 프로세서(예로서, 손 제스처 결정 모듈(323))에 의해 수행되는 것으로 설명되지만, 방법(600)은 단독으로 또는 프로세서와 조합하여 다른 디바이스에 의해 수행될 수 있다는 것이 이해된다.
- [0063] 단계(602)에서, 프로세서는 하나 이상의 카메라(예로서, 광학 감지 시스템(312)의)로부터 이미지 데이터를 수신할 수 있다. 카메라가 그레이-스케일 IR 카메라인 경우에, 프로세서는 IR 카메라 이미지를 획득할 수 있다. 카메라가 RGB-IR 카메라인 경우에, 프로세서는 IR 픽셀 데이터를 획득할 수 있다.
- [0064] 단계(604)에서, 프로세서는 상기 논의된 기술에 기초하여 이미지 데이터로부터 손 제스처를 결정할 수 있다. 결정은 또한 손 제스처의 유형(특정 명령을 나타낼 수 있는) 및 손가락의 궤적의 3D 좌표(손 제스처를 생성할 때) 양쪽 모두의 결정을 포함한다.
- [0065] 단계(606)에서, 프로세서는 검출된 손 제스처에 관련되는, 몰입식 멀티미디어 데이터의 부분으로서 렌더링되는, 오브젝트를 결정할 수 있다. 예를 들면, 손 제스처가 선택을 시그널링하는 경우에, 손 제스처에 의해 선택되는 렌더링된 오브젝트가 결정된다. 결정은 손 제스처의 궤적의 3D 좌표 및 손 제스처의 특정한 부분이 사용자의 관점 내에서 오브젝트의 적어도 부분과 중첩함을 나타내는 3D 맵에서의 오브젝트의 3D 좌표 사이의에서의 관계에 기초할 수 있다.
- [0066] 단계(608)에서, 프로세서는 단계(604)에서 결정된 손 제스처 및 단계(608)에서 결정된 오브젝트에 대한 정보에 기초하여, 멀티미디어 데이터의 렌더링을 변경하도록 결정할 수 있다. 예시적인 예로서, 단계(604)에서 검출된 손 제스처가 줌 동작을 위한 오브젝트를 선택하기 위한 명령과 연관된다는 결정에 기초하여(그것이 물리적 환경에 위치한 실제 오브젝트인지, 또는 렌더링에 삽입되는 가상 오브젝트인지), 프로세서는 렌더링을 위해 다운스트림 로직(예로서, 그래픽 및 오디오 렌더링 모듈(324))으로 오브젝트의 확대 이미지를 제공할 수 있다. 또 다른 예시적 예로서, 손 제스처가 오브젝트에 대한 부가적인 정보를 디스플레이하기 위한 명령과 연관된다면, 프로세서는 렌더링을 위해 부가적인 정보를 그래픽 및 오디오 렌더링 모듈(324)에 제공할 수 있다.
- [0067] 다시 도 3을 참조하면, 카메라의 방위 및 위치에 대한 정보(예로서, 방위 및 위치 결정 모듈(322)에 의해 제공된) 및 검출된 손 제스처에 대한 정보(예를 들면, 손 제스처 결정 모듈(323)에 의해 제공된)에 기초하여, 그래픽 및 오디오 렌더링 모듈(324)은 대화식 가상 현실 및/또는 증강 현실 경험을 생성하기 위해 몰입식 멀티미디어 데이터(그래픽 및 오디오 모두)를 렌더링할 수 있다. 다양한 방법이 렌더링을 위해 사용될 수 있다. 몇몇 실시예에서, 그래픽 및 오디오 렌더링 모듈(324)은 좌측 눈에 대한 이미지를 캡처하는 제1 카메라와 연관된 제1 3D 메시(평면이거나 또는 곡선일 수 있는), 및 우측 눈에 대한 이미지를 캡처하는 제2 카메라와 연관된 제2 3D 메시(또한 평면이거나 또는 곡선일 수 있는)를 생성할 수 있다. 3D 메시는 카메라로부터의 특정한 가상 거리에 위치될 수 있으며, 3D 메시의 크기는 그것들이 상기 가상 거리에서 카메라의 시야 절두체의 크기에 맞도록 결정될 수 있다. 그래픽 및 오디오 렌더링 모듈(324)은 그 후 제1 3D 메시에 좌측 이미지(제1 카메라에 의해 획득된)를 매핑시키며, 제2 3D 메시에 우측 이미지(제2 카메라에 의해 획득된)를 매핑시킬 수 있다. 그래픽 및 오디오 렌더링 모듈(324)은 좌측 눈에 대한 장면을 렌더링할 때 제1 3D 메시(및 그것에 매핑된 콘텐츠)만을 보여주도록, 및 우측 눈에 대한 장면을 렌더링할 때 제2 3D 메시(및 그것에 매핑된 콘텐츠)만을 보여주도록 구성될 수 있다.
- [0068] 몇몇 실시예에서, 그래픽 및 오디오 렌더링 모듈(324)은 또한 3D 포인트 클라우드를 사용하여 렌더링을 수행할 수 있다. 이전에 논의된 바와 같이, 위치 및 방위의 결정 동안, 물리적 환경 내에서 두드러진 특징(및 연관된

오브젝트)의 깊이 맵은 IR 픽셀 데이터에 기초하여 결정될 수 있다. 물리적 환경의 3D 포인트 클라우드는 그 후 깊이 맵에 기초하여 생성될 수 있다. 그래픽 및 오디오 렌더링 모듈(324)은 환경의 3D 렌더링을 생성하기 위해 3D 포인트 클라우드에 물리적 환경의 RGB 픽셀 데이터(예로서, RGB 카메라, 또는 RGB-IR 센서의 RGB 픽셀에 의해 획득된)를 매핑시킬 수 있다.

[0069] 몇몇 실시예에서, 3D 가상 오브젝트의 이미지가 물리적 환경의 실시간 그래픽 이미지와 합성되는 경우에, 그래픽 및 오디오 렌더링 모듈(324)은 가상 3D 오브젝트 및 물리적 환경의 깊이 정보, 뿐만 아니라 카메라의 위치 및 방위에 기초하여 렌더링을 결정하도록 구성될 수 있다. 참조가 이제 도 7a 및 도 7b에 대해 이루어지며, 이것은 본 개시의 실시예와 일치하는, 물리적 환경의 실시간 그래픽 이미지로의 3D 가상 오브젝트의 이미지의 합성을 예시한다. 도 7a에 도시된 바와 같이, 환경(700)은 물리적 오브젝트(702) 및 물리적 오브젝트(706)를 포함한다. 그래픽 및 오디오 렌더링 모듈(324)은 환경(700)을 렌더링할 때 물리적 오브젝트(702) 및 물리적 오브젝트(706) 사이에 가상 오브젝트(704)를 삽입하도록 구성된다. 환경(700)의 그래픽 이미지는 위치(A)에서 위치(B)로의 루트(710)를 따라 카메라(708)에 의해 캡처된다. 위치(A)에서, 물리적 오브젝트(706)는 렌더링된 환경 내에서 가상 오브젝트(704)에 대해 카메라(708)에 더 가까우며, 가상 오브젝트(704)의 부분을 보기 어렵게 하지만, 위치(B)에서, 가상 오브젝트(704)는 렌더링된 환경 내에서 물리적 오브젝트(706)에 대해 카메라(708)에 더 가깝다.

[0070] 그래픽 및 오디오 렌더링 모듈(324)은 그것들의 깊이 정보, 뿐만 아니라 카메라의 위치 및 방위에 기초하여 가상 오브젝트(704) 및 물리적 오브젝트(706)의 렌더링을 결정하도록 구성될 수 있다. 참조가 이제 도 8에 대해 이루어지며, 이것은 본 개시의 실시예와 일치하는, 물리적 환경의 그래픽 이미지와 가상 오브젝트 이미지를 합성하기 위한 대표적인 방법(800)을 예시하는 흐름도이다. 방법(800)이 프로세서(예로서, 그래픽 및 오디오 렌더링 모듈(324))에 의해 수행되는 것으로 설명되지만, 방법(800)은 단독으로 또는 프로세서와 조합하여 다른 디바이스에 의해 수행될 수 있다는 것이 이해된다.

[0071] 단계(802)에서, 프로세서는 가상 오브젝트(예로서, 도 7a의 가상 오브젝트(704))의 제1 이미지의 픽셀과 연관된 깊이 정보를 수신할 수 있다. 깊이 정보는 예를 들면, 도 3의 방위 및 위치 결정 모듈(322)에 의해 결정된 카메라(708)의 위치 및 방위에 기초하여 생성될 수 있다. 예를 들면, 3D 맵 내에서 가상 오브젝트의 미리 결정된 위치 및 상기 3D 맵에서의 카메라의 위치에 기초하여, 프로세서는 카메라 및 가상 오브젝트 사이에서의 거리를 결정할 수 있다.

[0072] 단계(804)에서, 프로세서는 물리적 오브젝트(예로서, 도 7a의 물리적 오브젝트(706))의 제2 이미지의 픽셀과 연관된 깊이 정보를 결정할 수 있다. 깊이 정보는 예를 들면, 도 3의 방위 및 위치 결정 모듈(322)에 의해 결정된 카메라(708)의 위치 및 방위에 기초하여 생성될 수 있다. 예를 들면, (예로서, SLAM 알고리즘을 이용해서) 3D 맵 내에서 물리적 오브젝트의 이전 결정된 위치 및 상기 3D 맵에서의 카메라의 위치에 기초하여, 카메라 및 물리적 오브젝트 사이에서의 거리가 결정될 수 있다.

[0073] 단계(806)에서, 프로세서는 두 개의 픽셀의 깊이 정보를 비교하며, 그 후 단계(808)에서, 비교 결과에 기초하여 픽셀 중 하나를 렌더링하도록 결정할 수 있다. 예를 들면, 프로세서가 물리적 오브젝트의 픽셀이 (예로서, 도 7b의 위치(A)에서) 가상 오브젝트의 픽셀보다 카메라에 더 가깝다고 결정하면, 프로세서는 가상 오브젝트의 픽셀이 물리적 오브젝트의 픽셀에 의해 가려짐을 결정하며, 물리적 오브젝트의 픽셀을 렌더링하도록 결정할 수 있다.

[0074] 다시 도 3을 참조하면, 몇몇 실시예에서, 그래픽 및 오디오 렌더링 모듈(324)은 또한 렌더링을 위한 오디오 데이터를 제공할 수 있다. 오디오 데이터는 예로서, 오디오 감지 시스템(313)(마이크로폰 어레이와 같은)으로부터 수집될 수 있다. 몇몇 실시예에서, 강화된 감각 능력을 제공하기 위해, 오디오 데이터의 일부가 사용자 지시(예로서, 손 제스처를 통해 검출된)에 기초하여 확대될 수 있다. 예를 들면, 마이크로폰 어레이를 사용하여, 그래픽 및 오디오 렌더링 모듈(324)은 오디오 데이터의 소스의 위치를 결정할 수 있으며, 사용자 지시에 기초하여 상기 특정한 소스와 연관된 오디오 데이터의 볼륨을 증가시키거나 또는 감소시키도록 결정할 수 있다. 오디오 데이터의 가상 소스가 물리적 환경에서 비롯된 오디오 신호와 합성되는 경우에, 그래픽 및 오디오 렌더링 모듈(324)은, 방법(800)과 유사한 방식으로, 마이크로폰 및 가상 소스 사이에서의 거리, 및 마이크로폰 및 물리적 오브젝트 사이에서의 거리를 또한 결정할 수 있다. 거리에 기초하여, 그래픽 및 오디오 렌더링 모듈(324)은 가상 소스로부터의 오디오 데이터가 물리적 오브젝트에 의해 차단되는지를 결정하며, 그에 따라 오디오 데이터의 렌더링을 조정할 수 있다.

[0075] 렌더링된 그래픽 및 오디오 데이터를 결정한 후, 그래픽 및 오디오 렌더링 모듈(324)은 그 후, 렌더링된 그래픽

데이터를 디스플레이하도록 구성된 디스플레이 시스템(332)(예로서, 디스플레이 스크린), 및 렌더링된 오디오 데이터를 플레이하도록 구성된 오디오 출력 시스템(334)(예로서, 스피커)을 포함하는, 오디오/비디오 시스템(330)으로 그래픽 및 오디오 데이터를 제공할 수 있다. 그래픽 및 오디오 렌더링 모듈(324)은 또한 저장 장치(예로서, 도 1의 저장 장치(128))에 그래픽 및 오디오 데이터를 저장하거나, 또는 렌더링을 위해 또 다른 디바이스로 송신된 데이터를 네트워크 인터페이스(예로서, 도 1의 네트워크 인터페이스(140))에 제공할 수 있다.

[0076] 또한, 시스템(300)은 또한 전력 시스템(340)을 포함하며, 이것은 통상적으로 배터리 및 전력 관리 시스템(도 3에는 도시 생략)을 포함한다.

[0077] 시스템(300)의 구성요소(소프트웨어 또는 하드웨어)의 일부는 상이한 플랫폼에 걸쳐 분포될 수 있다. 예를 들면, 도 1에 논의된 바와 같이, 컴퓨팅 시스템(100)(이에 기초하여 시스템(300)이 구현될 수 있음)은 스마트 디바이스(130)(예로서, 스마트폰)에 연결될 수 있다. 스마트 디바이스(130)는 프로세싱 시스템(320)의 기능 중 일부를 수행하도록 구성될 수 있다. 예를 들면, 스마트 디바이스(130)는 그래픽 및 오디오 렌더링 모듈(324)의 기능을 수행하도록 구성될 수 있다. 예시적인 예로서, 스마트 디바이스(130)는 방위 및 위치 결정 모듈(322)로부터 카메라의 방위 및 위치에 대한 정보, 및 손 제스처 결정 모듈(323)로부터 손 제스처 정보, 뿐만 아니라 감지 시스템(310)으로부터 물리적 환경에 대한 그래픽 및 오디오 정보를 수신하며, 그 후 그래픽 및 오디오의 렌더링을 수행할 수 있다. 또 다른 예시적인 예로서, 스마트 디바이스(130)는 동작하는 또 다른 소프트웨어(예로서, 앱)일 수 있으며, 이것은 멀티미디어 렌더링에 부가될 부가적인 콘텐츠를 생성할 수 있다. 스마트 디바이스(130)는 그 후 (그래픽 및 오디오 렌더링 모듈(324)을 통해 렌더링을 수행하는) 시스템(300)에 부가적인 콘텐츠를 제공할 수 있거나, 또는 그래픽 및 오디오 데이터의 렌더링에 부가적인 콘텐츠를 부가할 수 있다.

[0078] 도 9a 및 도 9b는 본 개시의 실시예와 일치하는, 대표적인 헤드-장착 대화형 몰입식 멀티미디어 생성 시스템(900)을 예시한 개략도이다. 몇몇 실시예에서, 시스템(900)은 컴퓨팅 디바이스(100), 시스템(300), 및 도 4f의 카메라 시스템(494)의 실시예를 포함한다.

[0079] 도 9a에 도시된 바와 같이, 시스템(900)은 개구(904)의 쌍, 헤드 밴드(906)를 가진 하우징(902)을 포함한다. 하우징(902)은 대화형 몰입식 멀티미디어 데이터를 생성하도록 구성된 하나 이상의 하드웨어 시스템을 유지하도록 구성된다. 예를 들면, 하우징(902)은 (도 9b에 예시된 바와 같이) 회로 보드(950)를 유지할 수 있으며, 이것은 한 쌍의 카메라(954a 및 954b), 하나 이상의 마이크로폰(956), 프로세싱 시스템(960), 모션 센서(962), 전력 관리 시스템(964), 하나 이상의 커넥터(968), 및 IR 투사기/조명기(970)를 포함한다. 카메라(954a 및 954b)는 물리적 환경의 그래픽 데이터를 생성하도록 구성되며 각각 RGB-IR 카메라를 포함할 수 있다. 마이크로폰(956)은 몰입식 멀티미디어 데이터의 부분으로서 렌더링된 오디오 데이터를 환경으로부터 수집하도록 구성된다. 프로세싱 시스템(960)은 도 3의 프로세싱 시스템(300)의 기능 중 적어도 일부를 수행하도록 구성되는, 범용 프로세서, FPGA, ASIC 등일 수 있다. 커넥터(968)는 부가적인 능력을 제공하기 위해(예로서, 오디오 및 그래픽 데이터를 렌더링하기 위해, 렌더링을 위한 부가적인 콘텐츠를 제공하기 위해 등) 도 1의 스마트 디바이스(130)로서 동작하는 이동 디바이스(예로서, 스마트폰)에 시스템(900)을 연결하도록 구성되며, 따라서 프로세싱 시스템(960)은 이동 디바이스와 통신할 수 있다. 이러한 경우에, 하우징(902)은 또한 이동 디바이스를 유지하기 위해 내부 공간을 제공한다. 하우징(902)은 또한 이동 디바이스에 의해 및/또는 프로세싱 시스템(960)에 의해 렌더링된 입체 3D 이미지를 디스플레이하도록 구성된 한 쌍의 렌즈(도시 생략) 및 선택적으로 디스플레이 디바이스(이동 디바이스에 의해 제공될 수 있음)를 포함한다. 하우징(902)은 또한 카메라(954)가 물리적 환경 시스템(900)이 위치되는 이미지를 캡처할 수 있는 개구(904)를 포함한다.

[0080] 도 9a에 도시된 바와 같이, 시스템(900)은 헤드 밴드의 세트(906)를 추가로 포함한다. 헤드 밴드는 사람이 그녀의 머리에 시스템(900)을 착용하도록 허용하기 위해 구성될 수 있으며, 그녀의 눈은 디스플레이 디바이스 및 렌즈에 노출된다. 몇몇 실시예에서, 배터리는 헤드 밴드에 위치될 수 있으며, 이것은 또한 배터리 및 하우징(902)에 하우징된 시스템 사이에서 전기적 연결을 제공할 수 있다.

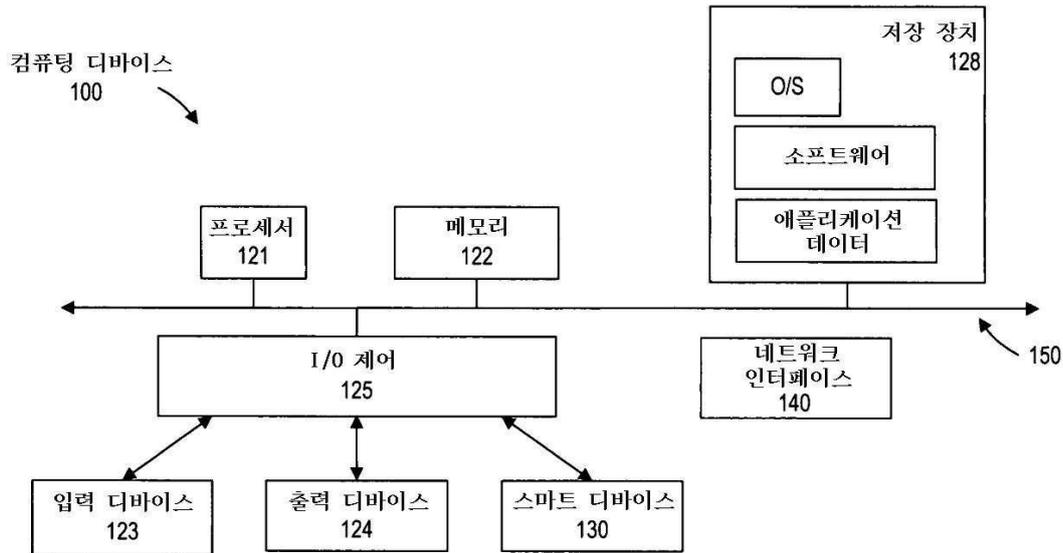
[0081] 본 개시의 실시예에 의하면, 사용자(및 카메라)의 3D 위치 및 방위의 정확한 추적이 제공될 수 있다. 사용자의 위치 및 방위 정보에 기초하여, 대화형 몰입식 멀티미디어 경험이 제공될 수 있다. 정보는 또한 가상 오브젝트의 이미지 및 물리적 환경의 이미지의 사실적인 합성이 증강 현실 및 가상 현실의 조합된 경험을 생성할 수 있게 한다. 본 개시의 실시예는 또한 사용자가 사용자의 감각 능력을 강화하기 위해 물리적 환경의 부분의 그래픽 및 오디오 렌더링을 효율적으로 업데이트할 수 있게 한다.

[0082] 앞서 말한 명세서에서, 실시예는 구현마다 달라질 수 있는 다수의 특정 세부사항을 참조하여 설명되었다. 설명된 실시예의 특정한 적응화 및 수정이 이루어질 수 있다. 다른 실시예는 여기에 개시된 본 발명의 명세서 및 실

시의 고려로부터 이 기술분야의 숙련자에게 명백할 수 있다. 명세서 및 예는, 단지 예로서 고려되도록 의도되며, 본 발명의 진정한 범위와 사상은 다음의 청구범위에 의해 나타나 있다. 도면에 도시된 단계의 수순은 단지 예시적인 목적을 위한 것이며 단계의 임의의 특정한 수순에 제한되도록 의도되지 않는다는 것이 또한 의도 된다. 이와 같이, 당업자라면 이들 단계가 동일한 방법을 구현하는 동안 상이한 순서로 수행될 수 있다는 것을 이해할 수 있다.

도면

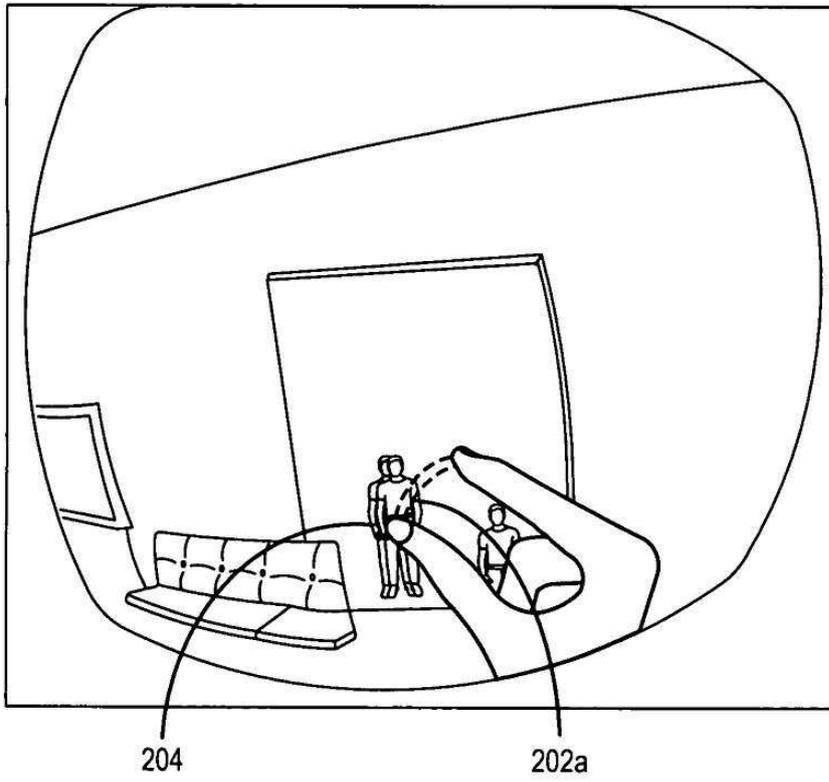
도면1



도면2a

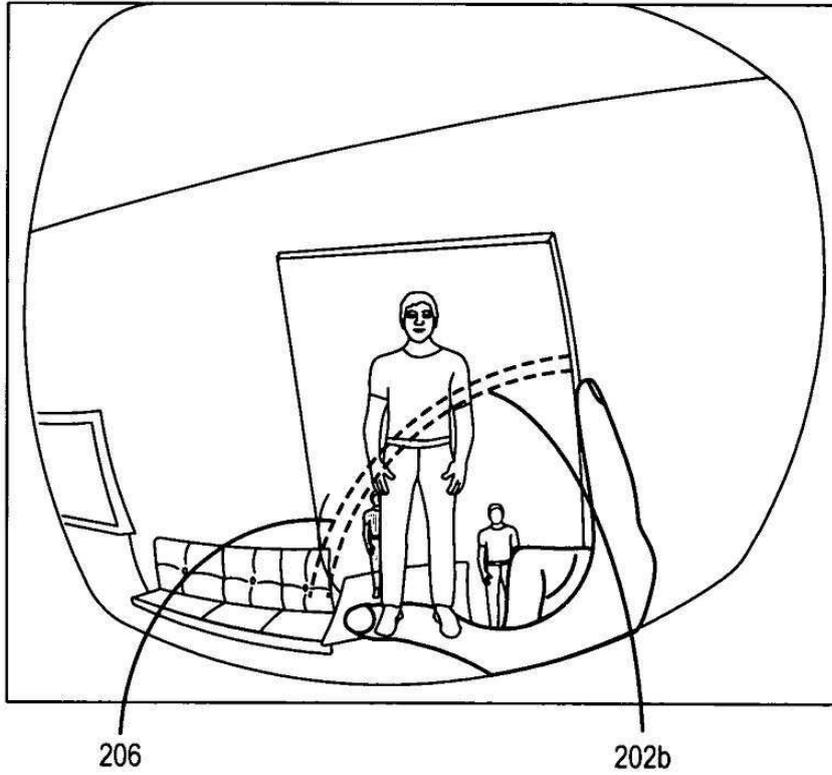
200

렌더링
200a



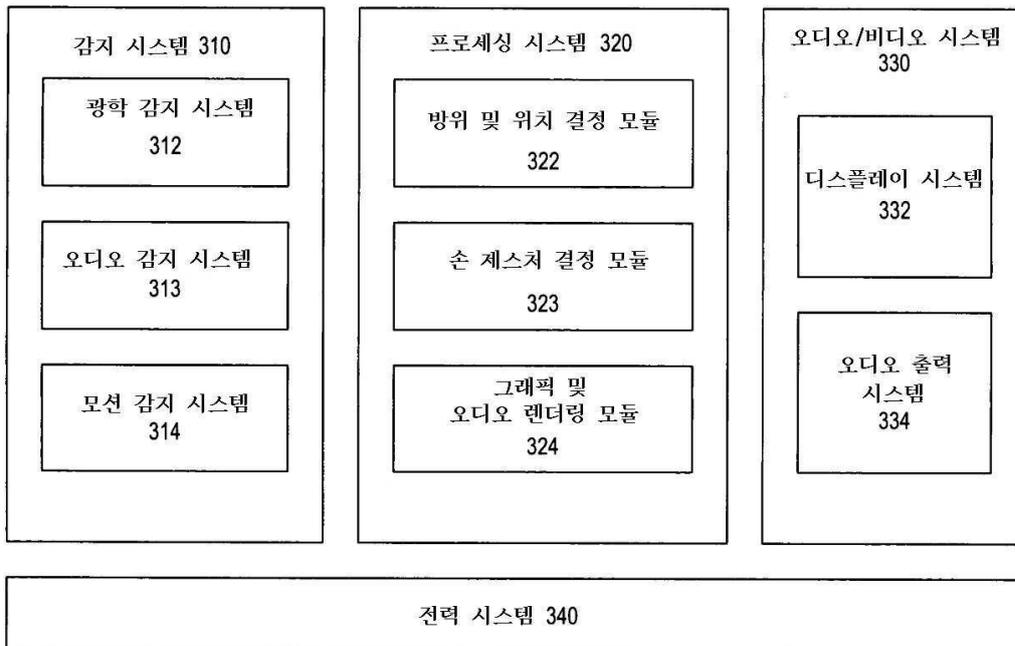
도면2b

렌더링
200b



도면3

시스템 300



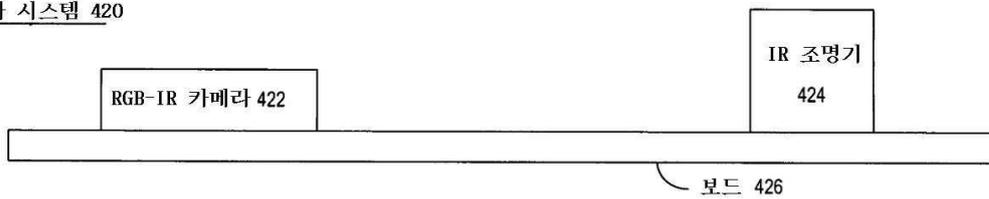
도면4a

카메라 시스템 400



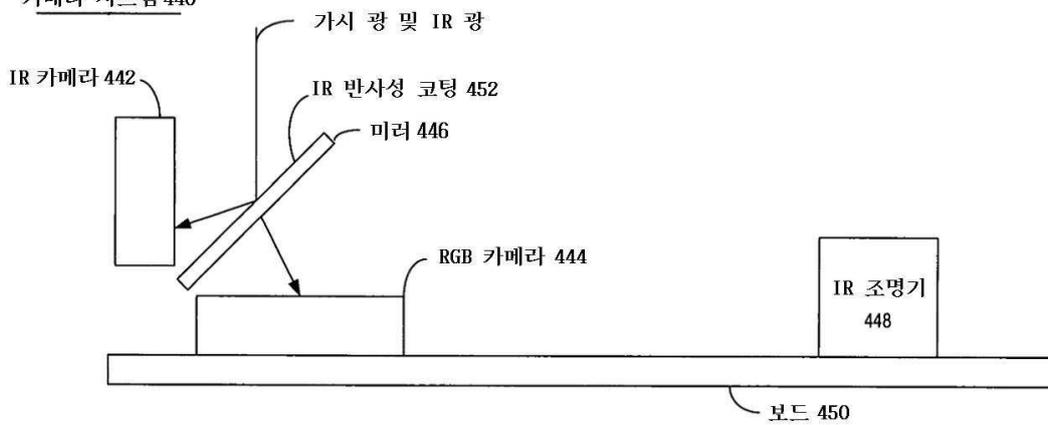
도면4b

카메라 시스템 420



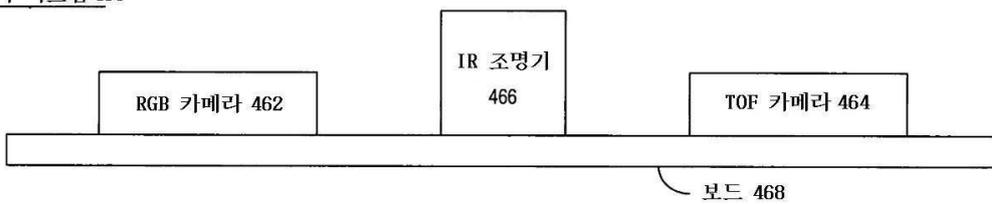
도면4c

카메라 시스템 440

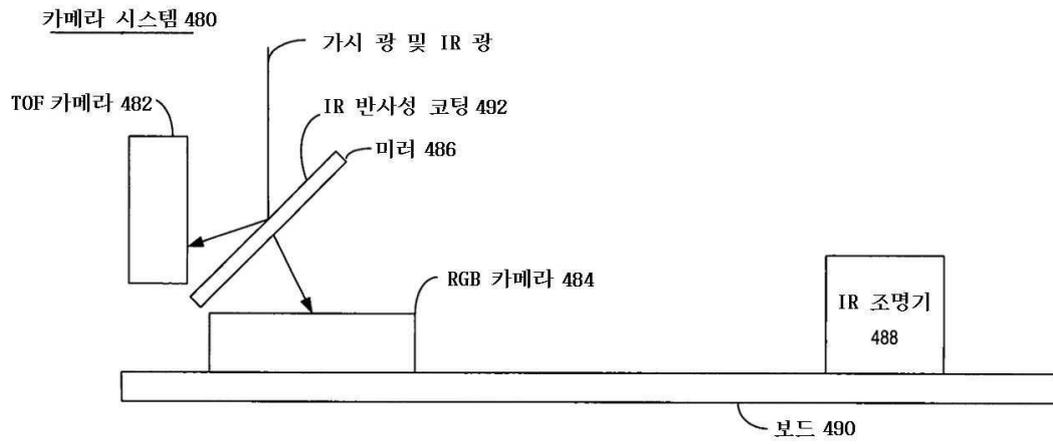


도면4d

카메라 시스템 460



도면4e



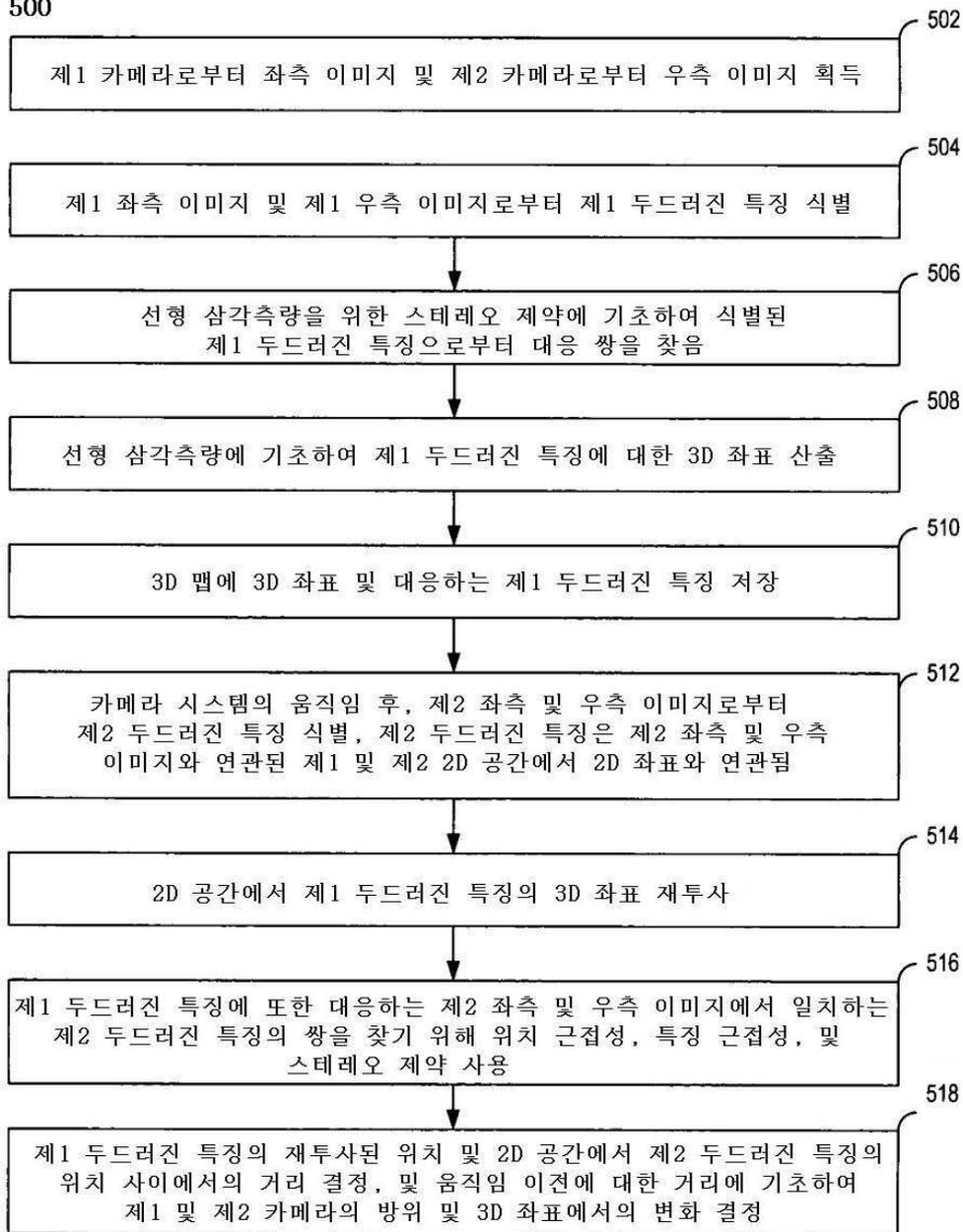
도면4f

카메라 시스템 494



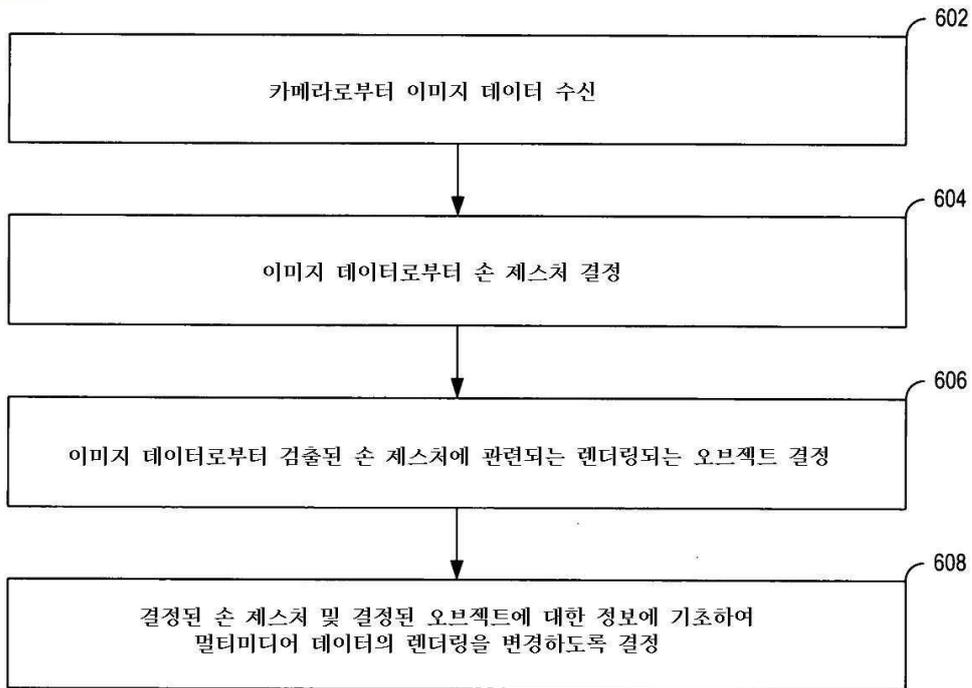
도면5

방법 500



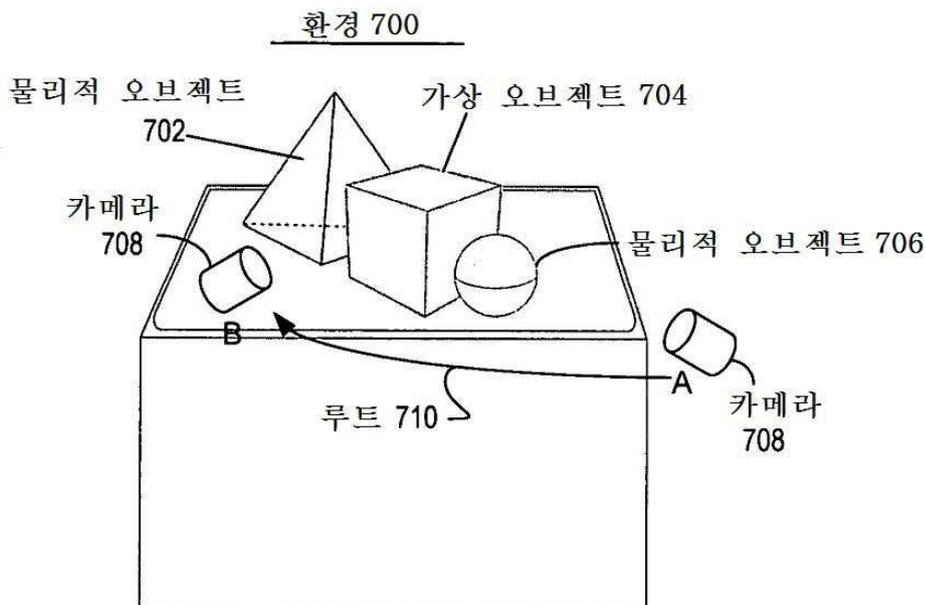
도면6

방법 600



도면7a

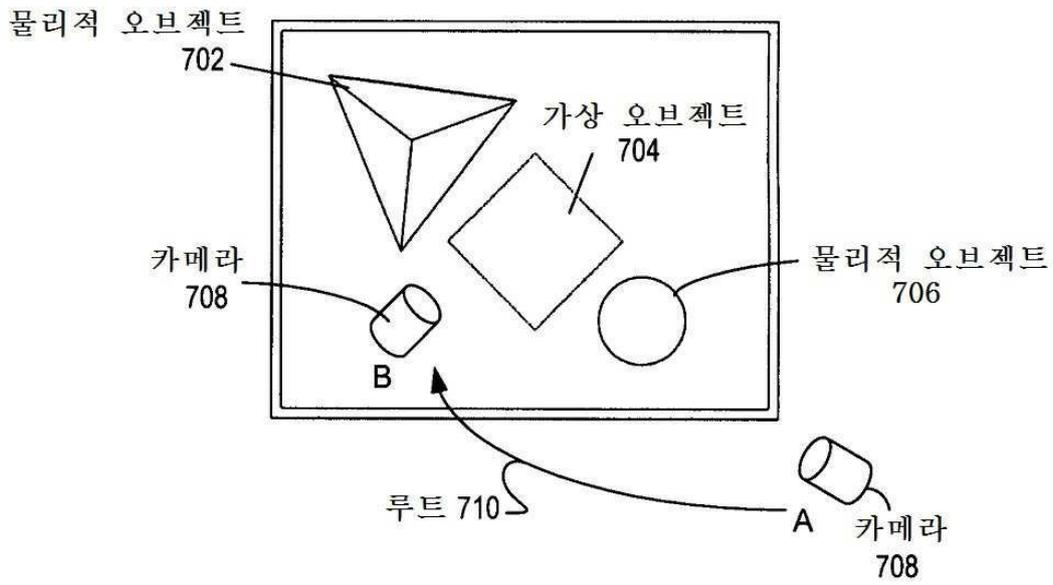
측면도



도면7b

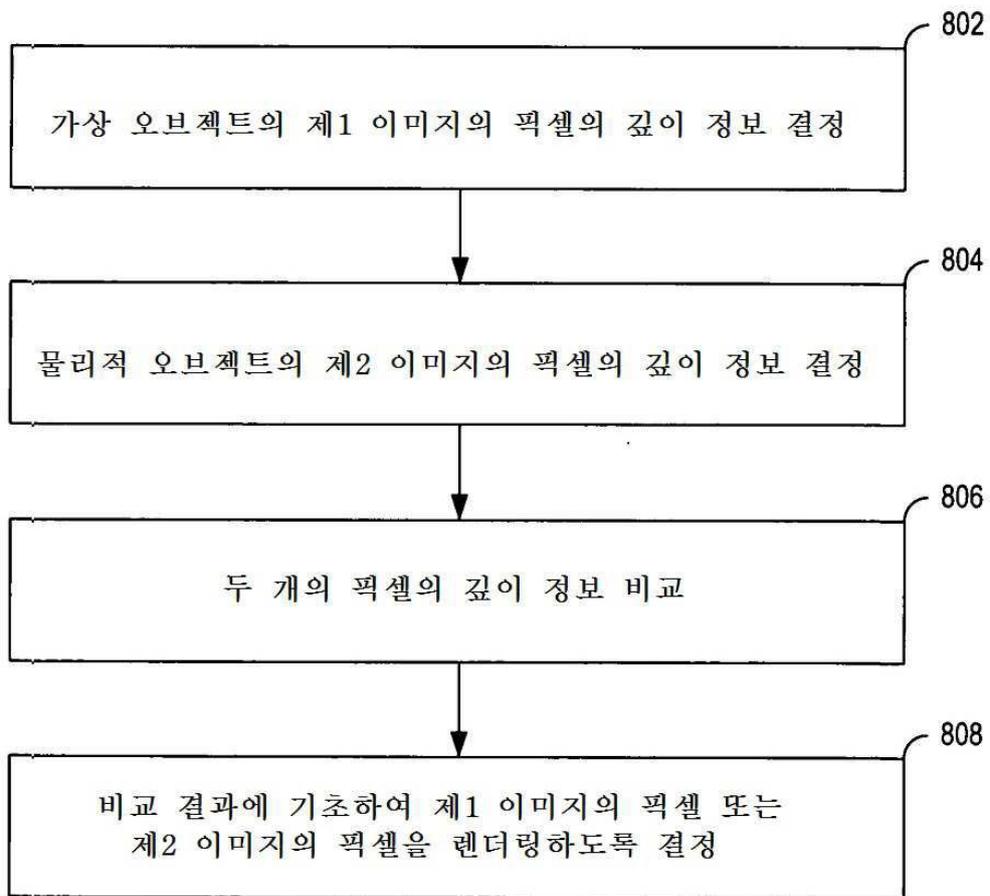
평면도

환경 700



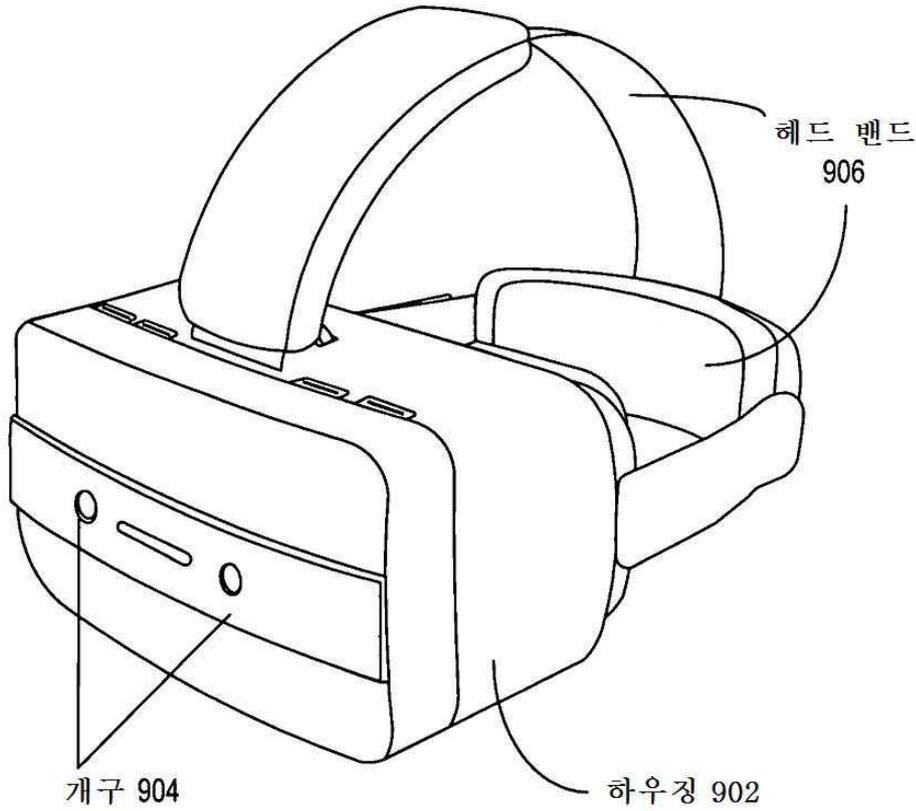
도면8

방법 800



도면9a

시스템 900



도면9b

회로 보드 950

