



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0093831
(43) 공개일자 2015년08월18일

(51) 국제특허분류(Int. Cl.)
G06F 3/01 (2006.01) G06T 19/00 (2011.01)
(52) CPC특허분류
G06F 3/011 (2013.01)
G06T 19/006 (2013.01)
(21) 출원번호 10-2015-7018669
(22) 출원일자(국제) 2013년12월12일
심사청구일자 없음
(85) 번역문제출일자 2015년07월10일
(86) 국제출원번호 PCT/US2013/074636
(87) 국제공개번호 WO 2014/093608
국제공개일자 2014년06월19일
(30) 우선권주장
13/713,910 2012년12월13일 미국(US)

(71) 출원인
마이크로소프트 테크놀로지 라이선싱, 엘엘씨
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이
(72) 발명자
마고리스 제프리 엔
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이
크로소프트 코포레이션 내
보트 벤자민 아이
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이
크로소프트 코포레이션 내
(뒷면에 계속)
(74) 대리인
김태홍, 김진희

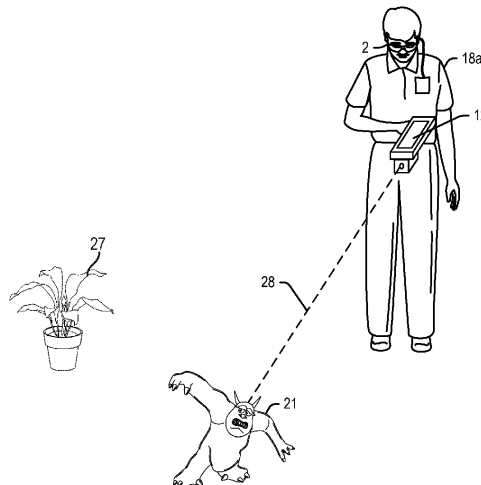
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 혼합 현실 환경에 대한 직접 상호작용 시스템

(57) 요약

핸드헬드 물체와 같은 액세서리를 사용하여 가상 환경 내의 가상 물체와 상호작용하는 시스템 및 방법이 개시되어 있다. 가상 물체는 디스플레이 디바이스를 사용하여 볼 수 있다. 디스플레이 디바이스 및 핸드헬드 물체는 가상 환경의 장면 맵을 결정하기 위해 협력할 수 있고, 디스플레이 디바이스 및 핸드헬드 물체는 장면 맵에 등록되어 있다.

대표도 - 도1



(72) 발명자

키프만 알렉스 아벤-아다르

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이크로소프트 코포레이션 내

클레인 조지

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이크로소프트 코포레이션 내

샤팔리츠키 프레데릭

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이크로소프트 코포레이션 내

니스터 데이비드

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이크로소프트 코포레이션 내

맥마킨 루스

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이크로소프트 코포레이션 내

반즈 덕

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 패턴즈 마이크로소프트 코포레이션 내

명세서

청구범위

청구항 1

현실 세계 공간과 동일 공간에 걸쳐 있는 가상 환경을 제시하는 시스템으로서,

하나 이상의 가상 물체들을 포함하는 장면 맵(scene map)을 결정하는 데 적어도 부분적으로 도움을 주는 디스플레이 디바이스 - 상기 디스플레이 디바이스는 상기 가상 환경에서의 상기 하나 이상의 가상 물체들 중 한 가상 물체를 디스플레이하는 디스플레이 유닛을 포함함 -; 및

상기 현실 세계 공간에서 상기 디스플레이 디바이스와 독립적으로 이동될 수 있는 액세서리 - 상기 액세서리는 상기 디스플레이 디바이스와 동일한 장면 맵 내에 등록되어 있음 - 를 포함하는, 가상 환경을 제시하는 시스템.

청구항 2

제1항에 있어서, 상기 액세서리는 핸드헬드 디바이스인 것인, 가상 환경을 제시하는 시스템.

청구항 3

제2항에 있어서, 상기 핸드헬드 디바이스는, 상기 핸드헬드 디바이스가 상기 현실 세계 공간에 이동될 때, 상기 핸드헬드 디바이스의 가속도 또는 속도 데이터 중 적어도 하나를 제공하는 관성 측정 유닛을 포함하는 것인, 가상 환경을 제시하는 시스템.

청구항 4

제1항에 있어서, 상기 액세서리는 촬영 디바이스 및 픽(puck)을 포함하는 것인, 가상 환경을 제시하는 시스템.

청구항 5

제4항에 있어서, 상기 픽은 셀룰러 진화를 포함하는 것인, 가상 환경을 제시하는 시스템.

청구항 6

현실 세계 공간과 동일 공간에 걸쳐 있는 가상 환경을 제시하는 시스템으로서,

하나 이상의 가상 물체들을 포함하는 장면 맵을 결정하는 데 적어도 부분적으로 도움을 주는 디스플레이 디바이스 - 상기 디스플레이 디바이스는 상기 가상 환경에서의 상기 하나 이상의 가상 물체들 중 한 가상 물체를 디스플레이하는 디스플레이 유닛을 포함함 -; 및

상기 디스플레이 디바이스와 동일한 장면 맵 내에 등록되어 있는 액세서리 - 상기 액세서리는 상기 가상 물체와 상호작용할 수 있음 - 를 포함하는, 가상 환경을 제시하는 시스템.

청구항 7

제6항에 있어서, 상기 액세서리는 상기 디스플레이 디바이스 상에 디스플레이되는 가상 광선을 사용하여 상기 가상 물체를 선택하는 것에 의해 상기 가상 물체와 상호작용하고, 상기 가상 광선은 상기 액세서리로부터 상기 가상 물체까지 뻗어있는 것으로 디스플레이되는 것인, 가상 환경을 제시하는 시스템.

청구항 8

제6항에 있어서, 상기 액세서리는 상기 액세서리가 상기 가상 물체의 표면과 접촉하거나 상기 가상 물체의 내부 내에 배치될 때 상기 가상 물체를 선택하는 것에 의해 상기 가상 물체와 상호작용하는 것인, 가상 환경을 제시하는 시스템.

청구항 9

현실 세계 공간과 동일 공간에 걸쳐 있는 가상 환경 내의 가상 물체들과 직접 상호작용하는 방법으로서,

- (a) 상기 가상 환경에 대한 장면 맵 - 가상 물체의 위치가 상기 장면 맵 내에 규정되어 있음 - 을 규정하는 단계;
- (b) 상기 가상 물체를 디스플레이 디바이스 - 상기 디스플레이 디바이스의 위치가 상기 장면 맵 내에 등록되어 있음 - 를 통해 디스플레이하는 단계; 및
- (c) 핸드헬드 디바이스 - 상기 핸드헬드 디바이스의 위치가 상기 장면 맵 내에 등록되어 있음 - 를 사용하여 상기 디스플레이 디바이스에 의해 디스플레이되는 상기 가상 물체와 직접 상호작용하는 단계를 포함하는, 가상 환경 내의 가상 물체들과 직접 상호작용하는 방법.

청구항 10

제9항에 있어서, 상기 가상 물체와 직접 상호작용하는 상기 단계 (c)는,

상기 핸드헬드 물체로부터 방출하는 것으로 상기 디스플레이 디바이스에 의해 디스플레이되는 가상 광선을 사용하여 상기 가상 물체를 선택하고, 상기 디스플레이 디바이스가 상기 가상 광선을 상기 가상 디바이스와 교차하는 것으로 디스플레이하도록 상기 핸드헬드 물체를 조작하는 단계, 또는

상기 가상 물체가 디스플레이되는 현실 세계 공간에 상기 핸드헬드 물체를 배치하는 것에 의해 상기 가상 물체를 선택하는 단계 중 하나를 포함하는 것인, 가상 환경 내의 가상 물체들과 직접 상호작용하는 방법.

발명의 설명

배 경 기 술

[0001]

혼합 현실(mixed reality)은 가상 영상이 현실 세계의 물리적 환경과 혼합될 수 있게 하는 기술이다. 사용자의 시야에 디스플레이되는 실제 물체들과 가상 물체들의 혼합 영상을 보기 위해 투시형 헤드 마운트드 혼합 현실 디스플레이 디바이스(see-through, head mounted, mixed reality display device)가 사용자에게 의해 착용될 수 있다. 헤드 마운트드 디스플레이 디바이스는 가상 물체들 및 실제 물체들이 보일 수 있는 주변의 3차원 맵을 생성할 수 있다. 사용자는 가상 물체들을 선택하는 것에 의해, 예를 들어, 가상 물체를 쳐다보는 것에 의해, 가상 물체들과 상호작용할 수 있다. 일단 선택되면, 사용자는 그 후에, 예를 들어, 가상 물체를 잡아 이동시키는 것 또는 그 물체에 대해 어떤 다른 사전 정의된 제스처를 수행하는 것에 의해 가상 물체를 조작하거나 이동시킬 수 있다.

[0002]

이 유형의 간접적 상호작용은 단점들이 있다. 예를 들어, 헤드 마운트드 디스플레이 디바이스에 의해 생성된 장면 맵 내에서 사용자의 손의 위치가 추정되고, 추정된 위치가 시간에 따라 드리프트(drift)할 수 있다. 이 결과, 잡고 있는 가상 물체가 사용자의 손을 벗어나 디스플레이되어 버릴 수 있다. 머리 움직임을 사용하여 물체를 선택하는 것이 또한 때로는 반직관적일 수 있다.

발명의 내용

[0003]

본 기술의 실시예들은 가상 환경 내의 3차원 가상 물체들과 상호작용하는 시스템 및 방법에 관한 것이다. 가상 환경 내에 가상 물체들을 생성하는 시스템은 하나 이상의 처리 유닛들에 결합된 투시형 헤드 마운트드 디스플레이 디바이스를 부분적으로 포함할 수 있다. 처리 유닛들은, 헤드 마운트드 디스플레이 유닛(들)과 협력하여, 가상 환경 내의 가상 물체들의 장면 맵(scene map)을 규정할 수 있다.

[0004]

본 시스템은 헤드 마운트드 디스플레이 디바이스와 독립적으로 움직이는 핸드헬드 디바이스와 같은 액세서리를 추가로 포함할 수 있다. 실시예들에서, 핸드헬드 디바이스는, 핸드헬드 디바이스가 헤드 마운트드 디스플레이 디바이스에 의해 사용되는 동일한 장면 맵에 등록될 수 있도록, 헤드 마운트드 디스플레이 디바이스 및/또는 처리 유닛(들)과 협력할 수 있다.

[0005]

핸드헬드 물체는 펙(puck)에 고정된 카메라를 포함할 수 있다. 펙은, 예를 들어, 사용자가 헤드 마운트드 디스플레이 디바이스에 의해 디스플레이되는 가상 물체와 상호작용하기 위한 명령들을 입력 패드 상에서 선택할 수 있게 하는 용량성 터치 스크린을 포함하는 입력 패드를 가질 수 있다. 카메라는 헤드 마운트드 디스플레이 디바이스 상의 하나 이상의 영상 포착 디바이스들에 의해 분간되는 지점들과 공통으로 그의 시야 내의 지점들을 분간할 수 있다. 이 공통의 지점들은 핸드헬드 디바이스에 대한 헤드 마운트드 디스플레이 디바이스의 위치를 결정(resolve)하고, 양 디바이스를 동일한 장면 맵에 등록시키는 데 사용될 수 있다. 헤드 마운트드 디스플레이

이 디바이스의 장면 맵에 핸드헬드 디바이스를 등록시키는 것은 핸드헬드 디바이스가 헤드 마운티드 디스플레이 디바이스에 의해 디스플레이되는 가상 물체들과 직접 상호작용하는 것을 가능하게 한다.

[0006]

한 예에서, 본 기술은 현실 세계 공간과 동일 공간에 걸쳐 있는 가상 환경을 제시하는 시스템에 관한 것으로서, 이 시스템은 하나 이상의 가상 물체들을 포함하는 장면 맵을 결정하는 데 적어도 부분적으로 도움을 주는 디스플레이 디바이스 - 디스플레이 디바이스는 가상 환경에서의 하나 이상의 가상 물체들 중 한 가상 물체를 디스플레이하는 디스플레이 유닛을 포함함 -; 및 현실 세계 공간에서 디스플레이 디바이스와 독립적으로 이동될 수 있는 액세서리 - 액세서리는 디스플레이 디바이스와 동일한 장면 맵 내에 등록되어 있음 - 를 포함한다.

[0007]

다른 예에서, 본 기술은 현실 세계 공간과 동일 공간에 걸쳐 있는 가상 환경을 제시하는 시스템에 관한 것으로서, 이 시스템은 하나 이상의 가상 물체들을 포함하는 장면 맵을 결정하는 데 적어도 부분적으로 도움을 주는 디스플레이 디바이스 - 디스플레이 디바이스는 가상 환경에서의 하나 이상의 가상 물체들 중 한 가상 물체를 디스플레이하는 디스플레이 유닛을 포함함 -; 및 디스플레이 디바이스와 동일한 장면 맵 내에 등록되어 있는 액세서리 - 액세서리는 가상 물체와 상호작용할 수 있음 - 를 포함한다.

[0008]

추가적인 예에서, 본 기술은 현실 세계 공간과 동일 공간에 걸쳐 있는 가상 환경 내의 가상 물체들과 직접 상호작용하는 방법에 관한 것으로서, 이 방법은 (a) 가상 환경에 대한 장면 맵을 규정하는 단계 - 가상 물체의 위치가 장면 맵 내에 규정되어 있음 -; (b) 가상 물체를 디스플레이 디바이스를 통해 디스플레이하는 단계 - 디스플레이 디바이스의 위치가 장면 맵 내에 등록되어 있음 -; 및 (c) 핸드헬드 디바이스를 사용하여 디스플레이 디바이스에 의해 디스플레이되는 가상 물체와 직접 상호작용하는 단계 - 핸드헬드 디바이스의 위치가 장면 맵 내에 등록되어 있음 - 를 포함한다.

[0009]

이 발명의 내용은 이하에서 발명을 실시하기 위한 구체적인 내용에 추가로 기술되는 개념들 중 선택된 것을 간략화된 형태로 소개하기 위해 제공된다. 이 발명의 내용은 청구된 발명 요지의 중요 특징들 또는 필수적인 특징들을 확인하기 위한 것이 아니며, 청구된 발명 요지의 범주를 정하는 데 보조 수단으로 사용되기 위한 것도 아니다.

도면의 간단한 설명

[0010]

도 1은 하나 이상의 사용자에게 가상 환경을 제시하는 시스템의 하나의 실시예의 예시적인 구성요소들을 예시한 도면.

도 2는 헤드 마운티드 디스플레이 유닛의 하나의 실시예의 사시도.

도 3은 헤드 마운티드 디스플레이 유닛의 하나의 실시예의 일부분의 측면도.

도 4는 헤드 마운티드 디스플레이 유닛의 구성요소들의 하나의 실시예의 블록도.

도 5는 헤드 마운티드 디스플레이 유닛의 포착 디바이스 및 처리 유닛의 구성요소들의 하나의 실시예의 블록도.

도 6은 헤드 마운티드 디스플레이 유닛과 연관된 처리 유닛의 구성요소들의 하나의 실시예의 블록도.

도 7은 본 개시 내용의 실시예들에 따른, 핸드헬드 디바이스의 사시도.

도 8은 본 개시 내용의 실시예들에 따른, 핸드헬드 디바이스의 일부로서 제공되는 펌웨어의 블록도.

도 9는 사용자가 핸드헬드 디바이스를 사용하여 가상 물체와 상호작용하는 가상 환경의 일례를 예시한 도면.

도 10은 본 시스템의 하나 이상의 처리 유닛들, 헤드 마운티드 디스플레이 디바이스 및 핸드헬드 디바이스의 동작 및 협동(collaboration)을 도시한 플로우차트.

도 11은 도 10의 플로우차트의 단계(608)의 보다 상세한 플로우차트.

발명을 실시하기 위한 구체적인 내용

[0011]

이제부터, 일반적으로 혼합 현실 환경에서 가상 물체들과 직접 상호작용하는 시스템 및 방법에 관한 본 기술의 실시예들이 도 1 내지 도 11을 참조하여 설명될 것이다. 실시예들에서, 시스템 및 방법은 헤드 마운티드 디스플레이 디바이스에 의해 발생하는 3차원 장면 맵에서 그 자신을 추적하고 등록시킬 수 있는 핸드헬드 디바이스를 사용할 수 있다. 핸드헬드 디바이스 및/또는 헤드 마운티드 디스플레이 디바이스는 각자의 디바이스들에 결합되거나 그 내에 통합되어 있는 모바일 처리 유닛은 물론, 사용자 주위의 시야를 포착하는 카메라를 포함할 수

있다.

- [0012] 각각의 사용자는 디스플레이 요소를 포함하는 헤드 마운티드 디스플레이 디바이스를 착용할 수 있다. 사용자가 디스플레이 요소를 통해 사용자의 시야(FOV) 내의 현실 세계 물체들을 볼 수 있도록 디스플레이 요소는 어느 정도 투명하다. 가상 영상들이 또한 현실 세계 물체들과 함께 나타날 수 있도록, 디스플레이 요소는 또한 가상 영상들을 사용자의 FOV 내에 투사하는 기능을 제공한다. 시스템이 사용자의 FOV에서 어디에 가상 영상을 삽입해야 하는지를 결정할 수 있도록, 시스템은 사용자가 어디를 보고 있는지를 자동으로 추적한다. 시스템이 가상 영상을 어디에 투사해야 하는지를 알게 되면, 영상이 디스플레이 요소를 사용하여 투사된다.
- [0013] 실시예들에서, 헤드 마운티드 디스플레이 디바이스 및/또는 핸드헬드 디바이스는 방 또는 다른 환경에 사용자, 현실 세계 물체들 및 가상 3차원 물체들의 6 자유도(x, y, z, 피치(pitch), 요(yaw) 및 롤(roll) 위치들)를 포함하는 환경의 모델을 구축하기 위해 협력할 수 있다. 환경에서 사용자들에 의해 착용되는 각각의 헤드 마운티드 디스플레이 디바이스의 위치들이 환경의 모델에 따라 그리고 서로에 대해 교정(calibrate)될 수 있다. 이것은 시스템이 환경의 각각의 사용자의 시선 및 FOV를 결정할 수 있게 한다. 이와 같이, 가상 영상이 각각의 사용자에게 디스플레이될 수 있지만, 시스템은, 환경에 있는 다른 물체들로부터의 또는 그들에 의한 임의의 폐색(occlusion)들 및 시차(parallax)에 대해 가상 영상을 조절하여, 각각의 사용자의 시점(perspective)으로부터의 가상 영상의 디스플레이를 결정한다. 환경의 모델(본 명세서에서 장면 맵이라고 지칭됨)은 물론, 환경에서의 각각의 사용자의 FOV 및 물체들의 추적이 협력하여 또는 개별적으로 동작하는 하나 이상의 처리 유닛들에 의해 발생될 수 있다.
- [0014] 본 기술의 측면들에 따르면, 핸드헬드 디바이스가 또한 환경의 모델에 따라 교정되고 환경의 모델 내에 등록될 수 있다. 이하에서 설명되는 바와 같이, 이것에 의해, 환경의 모델(장면 맵이라고도 지칭됨) 내에서 핸드헬드 디바이스의 위치 및 움직임(병진 및 회전)을 정확하게 알 수 있게 된다.
- [0015] 본 시스템에 의해 제공되는 가상 환경은 현실 세계 공간과 동일 공간에 걸쳐 있을 수 있다. 환언하면, 가상 환경이 현실 세계 공간 상에 오버레이되거나 그와 동일한 영역을 공유할 수 있다. 가상 환경이 방 또는 다른 현실 세계 공간의 범위 내에 들어갈 수 있다. 다른 대안으로서, 가상 환경이 현실 세계 물리적 공간의 범위보다 더 클 수 있다.
- [0016] 현실 세계 공간에서 이리저리 움직이는 사용자는 또한 동일 공간에 걸쳐 있는 가상 환경에서 이리저리 움직일 수 있고, 상이한 시점들 및 유리한 지점(vantage point)들로부터 가상 물체들 및/또는 실제 물체들을 볼 수 있다. 하나의 유형의 가상 환경은 가상 환경이 가상 물체들 및 현실 세계 물체들 둘 다를 포함하는 혼합 현실 환경이다. 다른 유형의 가상 환경은 단지 가상 물체들만을 포함한다.
- [0017] 이하에서 설명되는 바와 같이, 핸드헬드 물체는 가상 환경 내의 가상 물체들을 선택하고 그들과 직접 상호작용하는 데 사용될 수 있다. 그렇지만, 사용자는 다른 신체 및/또는 구두 제스처와 결합하여 핸드헬드 물체를 사용하여 가상 물체들과 상호작용할 수 있다. 따라서, 핸드헬드 디바이스 상의 버튼들 및/또는 터치 스크린을 작동시키는 것에 부가하여, 신체 제스처들은 혼합 현실 시스템에 의해 사전 정의된 동작을 수행하라는 시스템에 대한 사용자 요청으로서 인식되는 손가락, 손 및/또는 다른 신체 부위를 사용하여 사전 정의된 제스처를 수행하는 것을 추가로 포함할 수 있다. 물리적 상호작용은 핸드헬드 디바이스 또는 사용자의 다른 부위들에 의한 가상 물체와의 접촉을 추가로 포함할 수 있다. 예를 들어, 사용자는 핸드헬드 물체를 가상 물체와 접촉하게 또는 가상 물체 내에 둘 수 있고, 그 후에, 가상 물체를 밀거나 그와 부딪칠 수 있다.
- [0018] 사용자는, 다른 대안으로서 또는 그에 부가하여, 예를 들어, 사전 정의된 동작을 수행하라는 시스템에 대한 사용자 요청으로서 혼합 현실 시스템에 의해 인식되는 발화된 단어 또는 문구(spoken word or phrase)와 같은, 구두 제스처와 함께 핸드헬드 디바이스를 사용하여 가상 물체들과 상호작용할 수 있다. 가상 환경 내의 하나 이상의 가상 물체들과 상호작용하기 위해 구두 제스처들이 신체 제스처들과 함께 사용될 수 있다.
- [0019] 도 1은 가상 콘텐츠(21)와 실제 콘텐츠(27)를 사용자의 FOV 내에 융합(fuse)시키는 것에 의해 혼합 현실 경험을 제공하기 위한 시스템(10)을 예시한 것이다. 도 1은 사용자(18)가 헤드 마운티드 디스플레이 디바이스(2) - 하나의 실시예에서, 사용자가 디스플레이를 투시할 수 있고 그에 의해 사용자의 전방에 있는 공간을 실제로 직접 볼 수 있도록 안경의 형상으로 되어 있음 - 을 착용하고 있는 것을 도시하고 있다. "실제로 직접 본다"는 용어의 사용은, 물체들의 생성된 영상 표현들을 보는 것이 아니라, 현실 세계 물체들을 사람의 눈으로 직접 볼 수 있다는 것을 지칭한다. 예를 들어, 안경을 통해 방을 보는 것은 사용자가 방을 실제로 직접 볼 수 있게 하는 것인 반면, 텔레비전에서 방의 비디오를 보는 것은 방을 실제로 직접 보는 것이 아니다. 헤드 마운티드 디스플레이

레이 디바이스(2)의 추가의 상세들이 이하에서 제공된다.

- [0020] 본 기술의 측면들은 사용자에게 의해 휴대될 수 있는 핸드헬드 디바이스(12)를 추가로 포함할 수 있다. 실시예들에서 핸드헬드 디바이스라고 불리우고 도 1에 그러한 것으로서 도시되어 있지만, 디바이스(12)는 보다 광의적으로 헤드 마운티드 디스플레이 디바이스와 독립적으로 이동되고 헤드 마운티드 디스플레이 디바이스의 장면 맵 내에 등록될 수 있는 액세서리라고 지칭될 수 있다. 액세서리가 사용자의 손에 보유되어 있지 않으면서 조작될 수 있다. 액세서리가 사용자의 팔 또는 다리에 묶여 있을 수 있거나, 환경 내의 실제 물체 상에 배치될 수 있다.
- [0021] 도 2 및 도 3에서 보는 바와 같이, 각각의 헤드 마운티드 디스플레이 디바이스(2)는 유선(6)을 통해 그 자신의 처리 유닛(4)과 통신한다. 다른 실시예들에서, 헤드 마운티드 디스플레이 디바이스(2)는 무선 통신을 통해 처리 유닛(4)과 통신한다. 하나의 실시예에서, 처리 유닛(4)은, 예를 들어, 사용자의 손목에 착용되거나 사용자의 주머니 내에 보관되는 소형 휴대용 디바이스이다. 처리 유닛이, 예를 들어, 셀룰러 전화의 크기 및 폼 팩터(form factor)일 수 있지만, 추가의 예들에서, 다른 형상을 및 크기들일 수 있다. 추가의 실시예에서, 처리 유닛(4)은 헤드 마운티드 디스플레이 디바이스(2) 내에 통합될 수 있다. 처리 유닛(4)은 헤드 마운티드 디스플레이 디바이스(2)를 동작시키는 데 사용되는 컴퓨팅 능력의 상당 부분을 포함할 수 있다. 실시예들에서, 처리 유닛(4)은 핸드헬드 디바이스(12)와 무선으로(예컨대, WiFi, 블루투스, 적외선, 또는 다른 무선 통신 수단으로) 통신한다. 추가의 실시예들에서, 처리 유닛(4)이 그 대신에 핸드헬드 디바이스(12) 내에 통합되는 것이 생각되고 있다.
- [0022] 도 2 및 도 3은 헤드 마운티드 디스플레이 디바이스(2)의 사시도 및 측면도를 도시한 것이다. 도 3은 안경다리(102) 및 코 받침(104)을 가지는 디바이스의 일부분을 포함하는 헤드 마운티드 디스플레이 디바이스(2)의 우측을 도시한 것이다. 이하에서 기술되는 바와 같이, 소리를 녹음하고 그 오디오 데이터를 처리 유닛(4)으로 전송하기 위한 마이크(110)가 코 받침(104) 내에 내장되어 있다. 헤드 마운티드 디스플레이 디바이스(2)의 전방에는 비디오 영상 및 정지 영상을 포착할 수 있는 하나 이상의 방을 향해 있는 포착 디바이스들(125)이 있다. 이하에서 기술되는 바와 같이, 그 영상들은 처리 유닛(4)으로 전송된다.
- [0023] 헤드 마운티드 디스플레이 디바이스(2)의 프레임의 일부분은 (하나 이상의 렌즈들을 포함하는) 디스플레이를 둘러싸고 있을 것이다. 헤드 마운티드 디스플레이 디바이스(2)의 구성요소들을 보여주기 위해, 디스플레이를 둘러싸고 있는 프레임의 일부분이 도시되어 있지 않다. 디스플레이는 도광체 광학 요소(light-guide optical element)(115), 불투명도 필터(opacity filter)(114), 투시형 렌즈(see-through lens)(116) 및 투시형 렌즈(118)를 포함한다. 하나의 실시예에서, 불투명도 필터(114)는 투시형 렌즈(116)의 후방에서 그와 정렬되어 있고, 도광체 광학 요소(115)는 불투명도 필터(114)의 후방에서 그와 정렬되어 있으며, 투시형 렌즈(118)는 도광체 광학 요소(115)의 후방에서 그와 정렬되어 있다. 투시형 렌즈들(116 및 118)은 안경에서 사용되는 보통의 렌즈이고, 임의의 처방에 따라(처방 없음을 포함함) 제조될 수 있다. 하나의 실시예에서, 투시형 렌즈들(116 및 118)은 가변 처방 렌즈(variable prescription)로 대체될 수 있다. 일부 실시예들에서, 헤드 마운티드 디스플레이 디바이스(2)는 하나의 투시형 렌즈를 포함하거나 투시형 렌즈를 포함하지 않을 수 있다. 다른 대안에서, 처방 렌즈가 도광체 광학 요소(115) 내부에 들어갈 수 있다. 불투명도 필터(114)는 가상 영상의 콘트라스트를 향상시키기 위해 (픽셀별로 또는 균일하게) 자연 광을 필터링 제거한다. 도광체 광학 요소(115)는 인공 광(artificial light)을 눈으로 보낸다.
- [0024] (하나의 실시예에서) 가상 영상을 투사하기 위한 마이크로디스플레이(120) 및 마이크로디스플레이(120)로부터의 영상들을 도광체 광학 요소(115) 내로 지향시키기 위한 렌즈(122)를 포함하는 영상 소스가 안경다리(102)에 또는 그의 내부에 탑재되어 있다. 하나의 실시예에서, 렌즈(122)는 평행화 렌즈(collimating lens)이다.
- [0025] 제어 회로(136)는 헤드 마운티드 디스플레이 디바이스(2)의 다른 구성요소들을 지원하는 다양한 전자 회로를 제공한다. 제어 회로(136)의 부가의 상세들은 도 4와 관련하여 이하에서 제공된다. 이어폰(130), 관성 측정 유닛(132) 및 온도 센서(138)가 안경다리(102) 내부에 있거나 안경다리(102)에 탑재되어 있다. 도 4에 도시된 하나의 실시예에서, 관성 측정 유닛(132)[또는 IMU(132)]은 3축 자력계(132A), 3축 자이로(132B) 및 3축 가속도계(132C)와 같은 관성 센서들을 포함한다. 관성 측정 유닛(132)은 헤드 마운티드 디스플레이 디바이스(2)의 위치, 배향, 및 가속도(피치, 롤 및 요)를 감지한다. IMU(132)는, 자력계(132A), 자이로(132B) 및 가속도계(132C)에 부가하여 또는 그 대신에, 다른 관성 센서들을 포함할 수 있다.
- [0026] 마이크로디스플레이(120)는 영상을 렌즈(122)를 통해 투사한다. 마이크로디스플레이(120)를 구현하기 위해 사용될 수 있는 다른 영상 발생 기술들이 있다. 예를 들어, 마이크로디스플레이(120)는 광원이 광학적 활성 물질

(optically active material)로 변조되고 백색광으로 백라이팅되는 투과형 투사 기술을 사용하여 구현될 수 있다. 이들 기술은 보통 강력한 백라이트 및 높은 광 에너지 밀도를 갖는 LCD 유형 디스플레이를 사용하여 구현된다. 마이크로디스플레이(120)는 또한 외부 광이 반사되어 광학적 활성 물질에 의해 변조되는 반사 기술을 사용하여 구현될 수 있다. 기술에 따라, 백색 광원 또는 RGB 광원 중 어느 하나에 의해, 조명이 전방으로 비춰진다. DLP(digital light processing), LCOS(liquid crystal on silicon), 및 Qualcomm, Inc.의 Mirasol® 디스플레이 기술은 대부분의 에너지가 변조된 구조물(modulated structure)로부터 멀어지는 쪽으로 반사되기 때문에 효율적인 반사 기술의 예이고 본 시스템에서 사용될 수 있다. 그에 추가하여, 마이크로디스플레이(120)가 광이 디스플레이에 의해 발생하는 발광 기술(emissive technology)을 사용하여 구현될 수 있다. 예를 들어, Micro vision, Inc.의 PicoPTM 디스플레이 엔진은 레이저 신호를 마이크로 미러 조정(micro mirror steering)에 의해 투과 요소로서 기능하는 소형 스크린 상으로 또는 눈 안으로 직접 비춰지게(예컨대, 레이저) 방출한다.

[0027]

도광체 광학 요소(115)는 마이크로디스플레이(120)로부터의 광을 헤드 마운티드 디스플레이 디바이스(2)를 착용하고 있는 사용자의 눈(140) 쪽으로 투과시킨다. 도광체 광학 요소(115)는 또한, 화살표(142)로 도시된 바와 같이, 헤드 마운티드 디스플레이 디바이스(2)의 전방으로부터의 광이 도광체 광학 요소(115)를 통해 눈(140) 쪽으로 투과될 수 있게 하고, 그에 의해, 마이크로디스플레이(120)로부터 가상 영상을 수신하는 것에 추가하여, 사용자가 헤드 마운티드 디스플레이 디바이스(2)의 전방에 있는 공간을 실제로 직접 볼 수 있게 한다. 이와 같이, 도광체 광학 요소(115)의 벽은 투시형이다. 도광체 광학 요소(115)는 제1 반사면(124)(예컨대, 거울 또는 다른 표면)을 포함한다. 마이크로디스플레이(120)로부터의 광은 렌즈(122)를 통과하여 반사면(124)에 입사하게 된다. 반사면(124)은, 광이 내부 반사에 의해 도광체 광학 요소(115)를 구성하는 평면 기관 내부에 포획되도록, 마이크로디스플레이(120)로부터의 입사광을 반사시킨다. 기관의 표면으로부터의 몇번의 반사 후에, 포획된 광파들은 선택적 반사면들(126)의 어레이에 도달한다. 유의할 점은, 도면의 혼잡을 방지하기 위해, 5 개의 반사면들 중 하나가 126으로 표시되어 있다는 것이다. 반사면들(126)은 기관으로부터 그 반사면들에 입사하는 광파들을 사용자의 눈(140) 안으로 결합시킨다.

[0028]

상이한 광선들이 진행하여 기관의 내부로부터 상이한 각도들로 반사하기 때문에, 상이한 광선들이 다양한 반사면들(126)에 상이한 각도들로 충돌할 것이다. 따라서, 상이한 광선들이 기관으로부터 반사면들 중 상이한 것들에 의해 반사될 것이다. 어느 광선들이 기관으로부터 어느 반사면(126)에 의해 반사될 것인지의 선택은 반사면들(126)의 적절한 각도를 선택하는 것에 의해 이루어진다. 도광체 광학 요소에 대한 부가의 상세들은 2008년 11월 20일자로 출원된, 발명의 명칭이 "기관 도파 광학 디바이스(Substrate-Guided Optical Devices)"인 미국 특허 공개 제2008/0285140호(참조 문헌으로서 그 전체가 본 명세서에 포함됨)에서 찾아볼 수 있다. 도광체 광학 요소(115)가, 도파로(waveguide)를 통한 반사 대신에 또는 그에 추가하여, 투사 광학계에 의해 동작할 수 있다는 것을 잘 알 것이다. 하나의 실시예에서, 각각의 눈은 그 자신의 도광체 광학 요소(115)를 가질 것이다. 헤드 마운티드 디스플레이 디바이스(2)가 두 개의 도광체 광학 요소들을 가질 때, 각각의 눈은 양 눈에 동일한 영상을 디스플레이하거나 두 개의 눈에 상이한 영상들을 디스플레이할 수 있는 그 자신의 마이크로디스플레이(120)를 가질 수 있다. 다른 실시예에서, 광을 양 눈 안으로 반사시키는 하나의 도광체 광학 요소가 있을 수 있다.

[0029]

도광체 광학 요소(115)와 정렬되어 있는 불투명도 필터(114)는 자연 광을, 균일하게 또는 픽셀별로, 도광체 광학 요소(115)를 통과하지 못하게 선택적으로 차단시킨다. 불투명도 필터(114)의 일례의 상세들은 2010년 9월 21일자로 출원된, 발명의 명칭이 "투시형 탑재 디스플레이용 불투명도 필터(Opacity Filter For See-Through Mounted Display)"인 Bar-Zeev 등의 미국 특허 공개 제2012/0068913호(참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 제공되어 있다. 그렇지만, 일반적으로, 불투명도 필터(114)의 일 실시예는 불투명도 필터로서 역할할 수 있는, 투시형 LCD 패널, 전기 변색 필름, 또는 유사한 디바이스일 수 있다. 불투명도 필터(114)는 픽셀들의 조밀한 격자를 포함할 수 있고, 여기서 각각의 픽셀의 광 투과율은 최소 투과율과 최대 투과율 사이에서 개별적으로 제어가능하다. 0부터 100%까지의 투과율 범위가 이상적이지만, 예를 들어, 픽셀별로 약 50%부터 90%까지와 같은 더 제한된 범위들도 허용가능하다.

[0030]

현실 세계 물체들에 대한 대응물들로 z-버퍼링(z-buffering)을 한 후에, 알파 값들의 마스크가 렌더링 파이프라인으로부터 사용될 수 있다. 시스템이 증강 현실 디스플레이를 위한 장면을 렌더링할 때, 이하에서 설명되는 바와 같이, 시스템은 어느 현실 세계 물체들이 어느 가상 물체들의 전방에 있는지에 주목한다. 가상 물체가 현실 세계 물체의 전방에 있는 경우, 불투명도가 가상 물체의 커버리지 영역(coverage area)에 대해 온일 수 있다. 가상 물체가 (가상적으로) 현실 세계 물체의 후방에 있는 경우, 불투명도가 오프일 수 있음은 물론, 그 픽셀에 대한 임의의 색도 오프일 수 있으며, 따라서 사용자는 실제 광의 그 대응하는 영역(한 픽셀 이상의 크기

임)에 대한 현실 세계 물체를 볼 것이다. 커버리지는 픽셀 단위로(on a pixel-by-pixel basis) 되어 있을 것이고, 따라서 시스템은 가상 물체의 일부가 현실 세계 물체의 전방에 있는 경우, 가상 물체의 일부가 현실 세계 물체의 후방에 있는 경우, 및 가상 물체의 일부가 현실 세계 물체와 일치하는 경우를 처리할 수 있을 것이다. 낮은 비용, 전력 및 중량으로 0%부터 100% 불투명도에 이를 수 있는 디스플레이가 이 용도에 가장 바람직하다. 더욱이, 불투명도 필터가, 예컨대, 컬러 LCD에 의해 또는 유기 LED 등의 다른 디스플레이에 의해, 컬러로 렌더링될 수 있다.

[0031]

헤드 마운티드 디스플레이 디바이스(2)는 또한 사용자의 눈의 위치를 추적하는 시스템을 포함한다. 이하에서 설명될 것인 바와 같이, 시스템이 사용자의 FOV를 결정할 수 있도록 시스템은 사용자의 위치 및 배향을 추적할 것이다. 그렇지만, 사람은 그의 전방에 있는 모든 것을 인지하지는 않는다. 그 대신에, 사용자의 눈은 환경의 일부로 지향될 것이다. 따라서, 하나의 실시예에서, 시스템은 사용자의 FOV의 측정을 개선(refine)시키기 위해 사용자의 눈의 위치를 추적하기 위한 기술을 포함할 것이다. 예를 들어, 헤드 마운티드 디스플레이 디바이스(2)는 눈 추적 조명 디바이스(134A) 및 눈 추적 카메라(134B)(도 4)를 가지는 눈 추적 어셈블리(134)(도 3)를 포함한다. 하나의 실시예에서, 눈 추적 조명 디바이스(134A)는 눈 쪽으로 적외선(IR) 광을 방출하는 하나 이상의 IR 방출기를 포함한다. 눈 추적 카메라(134B)는 반사된 IR 광을 감지하는 하나 이상의 카메라들을 포함한다. 각막의 반사를 검출하는 공지의 영상 기법들에 의해 동공의 위치가 식별될 수 있다. 예를 들어, 2008년 7월 22일자로 특허된, 발명의 명칭이 "헤드 마운티드 눈 추적 및 디스플레이 시스템(Head Mounted Eye Tracking and Display System)"인 미국 특허 제7,401,920호(참조 문헌으로서 본 명세서에 포함됨)를 참고하기 바란다. 이러한 기법은 추적 카메라에 대한 눈의 중앙의 위치를 찾아낼 수 있다. 일반적으로, 눈 추적은 눈의 영상을 획득하는 것 및 컴퓨터 비전(computer vision) 기법들을 사용하여 안와(eye socket) 내의 동공의 위치를 결정하는 것을 포함한다. 하나의 실시예에서, 양 눈이 보통 일제히 움직이기 때문에, 하나의 눈의 위치를 추적하는 것으로 충분하다. 그렇지만, 각각의 눈을 개별적으로 추적하는 것이 가능하다.

[0032]

하나의 실시예에서, 시스템은, 헤드 마운티드 디스플레이 디바이스(2)의 렌즈의 각각의 코너에 하나의 IR LED 및 IR 광 검출기가 있도록, 직사각형 배열로 있는 4 개의 IR LED들 및 4 개의 IR 광 검출기들을 사용할 것이다. LED들로부터의 광이 눈으로부터 반사한다. 4 개의 IR 광 검출기들 각각에서 검출되는 적외선 광의 양은 동공 방향을 결정한다. 즉, 눈에서의 백색 대 흑색의 양은 그 특징의 광 검출기에 대한 눈으로부터 반사되는 광의 양을 결정할 것이다. 이와 같이, 광 검출기는 눈에서의 백색 또는 흑색의 양을 측정할 것이다. 4 개의 샘플들로부터, 시스템은 눈의 방향을 결정할 수 있다.

[0033]

다른 대안은 앞서 논의된 바와 같은 4 개의 적외선 LED들 및 헤드 마운티드 디스플레이 디바이스(2)의 렌즈의 측면에 있는 하나의 적외선 CCD를 사용하는 것이다. CCD가 안경 프레임으로부터 보이는 눈의 최대 75%를 촬영할 수 있도록, CCD는 소형 거울 및/또는 렌즈[어안 렌즈(fish eye lens)]를 사용할 것이다. 앞서 논의된 것과 거의 유사하게, CCD는 이어서 영상을 감지하고 컴퓨터 비전을 사용하여 영상을 찾아낼 것이다. 이와 같이, 도 3이 하나의 IR 송신기를 갖는 하나의 어셈블리를 도시하고 있지만, 도 3의 구조는 4 개의 IR 송신기들 및/또는 4 개의 IR 센서들을 갖도록 조절될 수 있다. 4 개 초과 또는 미만의 IR 송신기들 및/또는 4 개 초과 또는 미만의 IR 센서들이 또한 사용될 수 있다.

[0034]

눈의 방향을 추적하는 다른 실시예는 전하 추적(charge tracking)에 기초하고 있다. 이 개념은 망막이 측정 가능한 양전하를 지니고 각막이 음전하를 가진다는 관찰에 기초하고 있다. 센서들은 눈이 이리저리 움직이는 동안 전위(electrical potential)를 검출하고 눈이 무엇을 하고 있는지를 실시간으로 효과적으로 판독하기 위해 사용자의 귀에[이어폰(130) 근방에] 탑재되어 있다. 눈을 추적하기 위한 다른 실시예들이 또한 사용될 수 있다.

[0035]

도 3은 헤드 마운티드 디스플레이 디바이스(2)의 절반을 도시한 것이다. 전체 헤드 마운티드 디스플레이 디바이스는 다른 투시형 렌즈 세트, 다른 불투명도 필터, 다른 도광체 광학 요소, 다른 마이크로디스플레이(120), 다른 렌즈(122), 방을 향해 있는 카메라, 눈 추적 어셈블리, 마이크로 디스플레이, 이어폰, 및 온도 센서를 포함할 수 있다.

[0036]

도 4는 헤드 마운티드 디스플레이 디바이스(2)의 다양한 구성요소들을 도시한 블록도이다. 도 5는 처리 유닛(4)의 다양한 구성요소들을 나타낸 블록도이다. 헤드 마운티드 디스플레이 디바이스(2) - 그의 구성요소들은 도 4에 도시되어 있음 - 는 하나 이상의 가상 영상들을 현실 세계의 사용자의 뷰(view)와 매끄럽게 융합시키는 것에 의해 혼합 현실 경험을 사용자에게 제공하기 위해 사용된다. 그에 부가하여, 도 4의 헤드 마운티드 디스플레이 디바이스 구성요소들은 다양한 조건들을 추적하는 많은 센서들을 포함한다. 헤드 마운티드 디스플레이

디바이스(2)는 처리 유닛(4)으로부터 가상 영상에 관한 명령어들을 수신할 것이고 센서 정보를 다시 처리 유닛(4)에 제공할 것이다. 처리 유닛(4) - 그의 구성요소들은 도 4에 도시되어 있음 - 은 헤드 마운티드 디스플레이 디바이스(2)로부터 감각 정보를 수신할 것이다. 그 정보 및 데이터에 기초하여, 처리 유닛(4)은 가상 영상을 언제 어디서 사용자에게 제공할지를 결정하고 그에 따라 명령어들을 도 4의 헤드 마운티드 디스플레이 디바이스로 송신할 것이다.

[0037]

도 4는 전력 관리 회로(202)와 통신하는 제어 회로(200)를 도시하고 있다. 제어 회로(200)는 프로세서(210), 메모리(214)(예컨대, DRAM)와 통신하는 메모리 제어기(212), 카메라 인터페이스(216), 카메라 버퍼(218), 디스플레이 구동기(display driver)(220), 디스플레이 포맷터(display formatter)(222), 타이밍 발생기(226), 디스플레이 출력 인터페이스(display out interface)(228), 및 디스플레이 입력 인터페이스(display in interface)(230)를 포함한다.

[0038]

하나의 실시예에서, 제어 회로(200)의 구성요소들은 전용 라인(dedicated line)들 또는 하나 이상의 버스들을 통해 서로 통신한다. 다른 실시예에서, 제어 회로(200)의 구성요소들은 프로세서(210)와 통신한다. 카메라 인터페이스(216)는 영상 포착 디바이스들(125)에 대한 인터페이스를 제공하고, 영상 포착 디바이스들로부터 수신되는 영상들을 카메라 버퍼(218)에 저장한다. 디스플레이 구동기(220)는 마이크로디스플레이(120)를 구동할 것이다. 디스플레이 포맷터(222)는 마이크로디스플레이(120) 상에 디스플레이되는 가상 영상에 관한 정보를, 불투명도 필터(114)를 제어하는 불투명도 제어 회로(224)에 제공한다. 타이밍 발생기(226)는 시스템에 대한 타이밍 데이터를 제공하기 위해 사용된다. 디스플레이 출력 인터페이스(228)는 영상 포착 디바이스들(125)로부터의 영상들을 처리 유닛(4)에 제공하기 위한 버퍼이다. 디스플레이 입력 인터페이스(230)는 마이크로디스플레이(120) 상에 디스플레이될 가상 영상과 같은 영상들을 수신하기 위한 버퍼이다. 디스플레이 출력 인터페이스(228) 및 디스플레이 입력 인터페이스(230)는 처리 유닛(4)에 대한 인터페이스인 대역 인터페이스(band interface)(232)와 통신한다.

[0039]

전력 관리 회로(202)는 전압 조절기(234), 눈 추적 조명 구동기(236), 오디오 DAC 및 증폭기(238), 마이크 전치 증폭기 및 오디오 ADC(240), 온도 센서 인터페이스(242) 및 클럭 발생기(244)를 포함한다. 전압 조절기(234)는 처리 유닛(4)으로부터 대역 인터페이스(232)를 통해 전력을 수신하고 그 전력을 헤드 마운티드 디스플레이 디바이스(2)의 다른 구성요소들에 제공한다. 눈 추적 조명 구동기(236)는, 앞서 기술한 것과 같은, 눈 추적 조명을 위한 IR 광원(134A)을 제공한다. 오디오 DAC 및 증폭기(238)는 오디오 정보를 이어폰(130)에 출력한다. 마이크 전치 증폭기 및 오디오 ADC(240)는 마이크(110)에 대한 인터페이스를 제공한다. 온도 센서 인터페이스(242)는 온도 센서(138)에 대한 인터페이스이다. 전력 관리 회로(202)는 또한 3축 자력계(132A), 3축 자이로(132B) 및 3축 가속도계(132C)에 전력을 제공하고 그로부터 다시 데이터를 수신한다.

[0040]

헤드 마운티드 디스플레이(2)는 사용자의 환경의 장면 맵 및 3차원 모델의 구성을 가능하게 하기 위해 사용자의 FOV의 RGB 및 깊이 영상들을 포착하기 위한 복수의 포착 디바이스들(125)을 추가로 포함할 수 있다. 도 3은 2개의 이러한 포착 디바이스들(125) - 하나는 헤드 마운티드 디스플레이(2)의 전방을 향해 있고 다른 하나는 측면을 향해 있음 - 을 개략적으로 도시하고 있다. 깊이 정보를 발생시키기 위해 결정될 수 있는 시각 입체 데이터(visual stereo data)를 획득하기 위해 상이한 각도들로부터 장면을 보기 위한 4개의 포착 디바이스들(125)을 제공하기 위해, 반대쪽 측면이 동일한 구성을 포함할 수 있다. 추가의 실시예들에서, 더 많은 또는 더 적은 포착 디바이스들이 있을 수 있다.

[0041]

예시적인 실시예에 따르면, 포착 디바이스(125)는, 예를 들어, 비행 시간(time-of-flight), 구조화된 광(structured light), 입체 영상 등을 비롯한 임의의 적당한 기법을 통해, 깊이 값들을 포함할 수 있는 깊이 영상을 비롯한 깊이 정보를 갖는 비디오를 포착하도록 구성될 수 있다. 하나의 실시예에 따르면, 포착 디바이스(125)는 깊이 정보를 "Z층들", 즉 깊이 카메라로부터 그의 시선(line of sight)을 따라 뻗어 있는 Z축에 수직일 수 있는 층들로 편성할 수 있다.

[0042]

포착 디바이스(125)의 개략적인 표현이 도 5에 도시되어 있다. 포착 디바이스(125)는, 실시예들에서, 장면의 깊이 영상을 포착할 수 있는 깊이 카메라일 수 있거나 그를 포함할 수 있는 카메라 구성요소(423)를 가질 수 있다. 깊이 영상은 포착된 장면의 2차원(2D) 픽셀 영역을 포함할 수 있고, 여기서 2D 픽셀 영역 내의 각각의 픽셀은 포착된 장면 내의 물체와 카메라 간의 거리와 같은 깊이 값(단위: 예를 들어, 센티미터, 밀리미터 등)을 나타낼 수 있다.

[0043]

카메라 구성요소(423)는 장면의 깊이 영상을 포착하는 데 사용될 수 있는 적외선(IR) 광 구성요소(425), 3차원(3D) 카메라(426), 및 RGB(시각 영상) 카메라(428)를 포함할 수 있다. 예를 들어, 비행 시간 분석에서, 포착

디바이스(125)의 IR 광 구성요소(425)는 장면 상으로 적외선 광을 방출할 수 있고, 이어서, 예를 들어, 3D 카메라(426) 및/또는 RGB 카메라(428)를 사용하여 장면 내의 하나 이상의 목표물들 및 물체들의 표면으로부터 후방 산란된 광을 검출하기 위해 센서들(일부 실시예들에서, 도시되지 않은 센서들을 포함함)을 사용할 수 있다. 추가의 실시예들에서, 3D 카메라 및 RGB 카메라는, 예를 들어, 진보된 컬러 필터 패턴들을 이용하는 동일한 센서 상에 존재할 수 있다. 일부 실시예들에서, 나가는 광 펄스(outgoing light pulse)와 대응하는 들어오는 광 펄스(incoming light pulse) 사이의 시간이 측정될 수 있고 포착 디바이스(125)로부터 장면 내의 목표물들 또는 물체들 상의 특징의 위치까지의 물리적 거리를 결정하는 데 사용될 수 있도록 펄스형 적외선 광(pulsed infrared light)이 사용될 수 있다. 그에 부가하여, 다른 예시적인 실시예들에서, 위상 천이를 결정하기 위해, 나가는 광파의 위상이 들어오는 광파의 위상과 비교될 수 있다. 위상 천이는 이어서 포착 디바이스로부터 목표물들 또는 물체들 상의 특징의 위치까지의 물리적 거리를 결정하는 데 사용될 수 있다.

[0044] 다른 예시적인 실시예에 따르면, 예를 들어, 셔터 방식의 광 펄스 촬영(shuttered light pulse imaging)을 비롯한 다양한 기법들을 통해 시간에 따라 반사된 광 빔의 세기를 분석함으로써 포착 디바이스(125)로부터 목표물들 또는 물체들 상의 특징의 위치까지의 물리적 거리를 간접적으로 결정하기 위해 비행 시간 분석이 사용될 수 있다.

[0045] 다른 예시적인 실시예에서, 포착 디바이스(125)는 깊이 정보를 포착하기 위해 구조화된 광을 사용할 수 있다. 이러한 분석에서, 패턴화된 광(즉, 격자 패턴, 줄무늬 패턴, 또는 다른 패턴과 같은 기지의 패턴으로서 디스플레이되는 광)이, 예를 들어, IR 광 구성요소(425)를 통해 장면 상에 투사될 수 있다. 장면 내의 하나 이상의 목표물들 또는 물체들의 표면에 부딪칠 때, 패턴이 그에 응답하여 변형될 수 있다. 패턴의 이러한 변형은, 예를 들어, 3D 카메라(426) 및/또는 RGB 카메라(428)(및/또는 다른 센서)에 의해 포착될 수 있고, 이어서 포착 디바이스로부터 목표물들 또는 물체들 상의 특징의 위치까지의 물리적 거리를 결정하기 위해 분석될 수 있다. 일부 구현예들에서, IR 광 구성요소(425)가 카메라들(426 및 428)로부터 변위되어 있고, 따라서 카메라들(426 및 428)로부터의 거리를 결정하기 위해 삼각측량이 사용될 수 있다. 일부 구현예들에서, 포착 디바이스(125)는 IR 광을 감지하기 위한 전용 IR 센서, 또는 IR 필터를 갖는 센서를 포함할 것이다.

[0046] 예시적인 실시예에서, 포착 디바이스(125)는 카메라 구성요소(423)와 통신할 수 있는 프로세서(432)를 추가로 포함할 수 있다. 프로세서(432)는, 예를 들어, 깊이 영상을 수신하고, 적절한 데이터 포맷(data format)(예컨대, 프레임)을 발생시키며, 데이터를 처리 유닛(4)으로 전송하기 위한 명령어들을 비롯한 명령어들을 실행할 수 있는 표준화된 프로세서, 특수 프로세서, 마이크로프로세서 등을 포함할 수 있다.

[0047] 포착 디바이스(125)는 프로세서(432)에 의해 실행될 수 있는 명령어들, 3D 카메라 및/또는 RGB 카메라에 의해 포착된 영상들 또는 영상들의 프레임들, 또는 임의의 다른 적당한 정보, 영상들 등을 저장할 수 있는 메모리(434)를 추가로 포함할 수 있다. 예시적인 실시예에 따르면, 메모리(434)는 랜덤 액세스 메모리(RAM), 판독 전용 메모리(ROM), 캐시, 플래시 메모리, 하드 디스크, 또는 임의의 다른 적당한 저장 구성요소를 포함할 수 있다. 추가의 실시예들에서, 프로세서(432) 및/또는 메모리(434)는 헤드 마운티드 디스플레이 디바이스(2)의 제어 회로(도 4) 또는 처리 유닛(4)의 제어 회로(도 6) 내에 통합될 수 있다.

[0048] 포착 디바이스(125)는 통신 링크(436)를 통해 처리 유닛(4)과 통신할 수 있다. 통신 링크(436)는, 예를 들어, USB 연결, Firewire 연결, 이더넷 케이블 연결 등을 비롯한 유선 연결 및/또는 무선 802.11b, g, a, 또는 n 연결과 같은 무선 연결일 수 있다. 하나의 실시예에 따르면, 처리 유닛(4)은, 예를 들어, 장면을 언제 포착해야 하는지를 결정하는 데 사용될 수 있는 클럭[도 6의 클럭 발생기(360) 등]을 통신 링크(436)를 통해 포착 디바이스(125)에 제공할 수 있다. 그에 부가하여, 포착 디바이스(125)는, 예를 들어, 3D 카메라(426) 및/또는 RGB 카메라(428)에 의해 포착된 깊이 정보 및 시각(예컨대, RGB) 영상들을 통신 링크(436)를 통해 처리 유닛(4)에 제공한다. 하나의 실시예에서, 깊이 영상들 및 시각 영상들은 초당 30 프레임으로 전송되지만; 다른 프레임 레이트들이 사용될 수 있다. 처리 유닛(4)은 이어서 모델, 깊이 정보, 및 포착된 영상들을 생성하고, 예를 들어, 가상 물체들의 발생을 포함할 수 있는 응용을 제어하기 위해, 그를 사용할 수 있다.

[0049] 처리 유닛(4)은 골격 추적 모듈(skeletal tracking module)(450)을 포함할 수 있다. 모듈(450)은, 각각의 사용자가 장면에서 이리저리 움직일 때 포착 디바이스(125)의 FOV 내의 사용자(18)(또는 다른 사람들)의 대표적인 모델을 개발하기 위해, 포착 디바이스(125)로부터의 그리고 어쩌면 하나 이상의 헤드 마운티드 디스플레이 디바이스들(2) 상의 카메라들로부터의 각각의 프레임에서 획득된 깊이 영상들을 사용한다. 이 대표적인 모델은 이하에서 기술되는 골격 모델일 수 있다. 처리 유닛(4)은 장면 매핑 모듈(452)을 추가로 포함할 수 있다. 장면 매핑 모듈(452)은, 사용자(18)가 존재하는 장면의 맵 또는 모델을 개발하기 위해, 포착 디바이스(125)로부터 획득

득된 깊이 그리고 어쩌면 RGB 영상 데이터를 사용한다. 장면 맵은 골격 추적 모듈(450)로부터 획득된 사용자들의 위치들을 추가로 포함할 수 있다. 처리 유닛(4)은 장면 내의 하나 이상의 사용자들에 대한 골격 모델 데이터를 수신하고 사용자가 처리 유닛(4)에서 실행 중인 애플리케이션에 영향을 미치는 사전 정의된 제스처 또는 애플리케이션 제어 움직임을 수행하고 있는지를 판정하기 위한 제스처 인식 엔진(454)을 추가로 포함할 수 있다.

[0050] 제스처 인식 엔진(454)에 관한 추가 정보는 2009년 4월 13일자로 출원된, 발명의 명칭이 "제스처 인식기 시스템 아키텍처(Gesture Recognizer System Architecture)"인 미국 특허 출원 제12/422,661호(참조 문헌으로서 그 전체가 본 명세서에 포함됨)에서 찾아볼 수 있다. 제스처를 인식하는 것에 관한 추가 정보는 또한 2009년 2월 23일자로 출원된, 발명의 명칭이 "표준 제스처(Standard Gestures)"인 미국 특허 출원 제12/391,150호; 및 2009년 5월 29일자로 출원된, 발명의 명칭이 "제스처 도구(Gesture Tool)"인 미국 특허 출원 제12/474,655호(둘 다 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에서 찾아볼 수 있다.

[0051] 포착 디바이스(125)는 RGB 영상들(또는 다른 포맷으로 된 또는 다른 색 공간에서의 시각 영상들) 및 깊이 영상들을 처리 유닛(4)에 제공한다. 깊이 영상은 복수의 관찰된 픽셀들일 수 있고, 여기서 각각의 관찰된 픽셀은 관찰된 깊이 값을 가진다. 예를 들어, 깊이 영상은 캡처된 장면의 2차원(2D) 픽셀 영역을 포함할 수 있고, 여기서 2D 픽셀 영역 내의 각각의 픽셀은 포착된 장면 내의 물체와 포착 디바이스 간의 거리와 같은 깊이 값을 가질 수 있다. 처리 유닛(4)은, 사용자의 골격 모델을 개발하기 위해 그리고 사용자 또는 다른 물체의 움직임을 추적하기 위해, RGB 영상들 및 깊이 영상들을 사용할 것이다. 깊이 영상들로 사람의 골격을 모델링하고 추적하기 위해 사용될 수 있는 많은 방법들이 있다. 깊이 영상을 사용하여 골격을 추적하는 하나의 적당한 예는 2009년 10월 21일자로 출원된, 발명의 명칭이 "자세 추적 파이프라인(Pose Tracking Pipeline)"인 미국 특허 출원 제12/603,437호(이후부터 '437 출원이라고 지칭됨)(참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 제공되어 있다.

[0052] '437 출원의 프로세스는 깊이 영상을 획득하는 단계, 데이터를 다운샘플링하는 단계, 고분산의 잡음이 많은 데이터(high variance noisy data)를 제거 및/또는 평활화하는 단계, 배경을 식별하고 제거하는 단계, 그리고 전경 픽셀들 각각을 신체의 상이한 부위들에 할당하는 단계를 포함한다. 그 단계들에 기초하여, 시스템은 데이터에 따라 모델을 근사시키고 골격을 생성할 것이다. 골격은 관절들 및 관절들 사이의 연결들의 그룹을 포함할 것이다. 사용자 모델링 및 추적을 위한 다른 방법들이 또한 사용될 수 있다. 적당한 추적 기술들이 또한 이하의 4 개의 미국 특허 출원들(이들 모두는 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 개시되어 있다: 2009년 5월 29일자로 출원된, 발명의 명칭이 "시간에 따라 다수의 사람들을 식별 및 추적하는 디바이스(Device for Identifying and Tracking Multiple Humans Over Time)"인 미국 특허 출원 제12/475,308호; 2010년 1월 29일자로 출원된, 발명의 명칭이 "시각 기반 신원 추적(Visual Based Identity Tracking)"인 미국 특허 출원 제12/696,282호; 2009년 12월 18일자로 출원된, 발명의 명칭이 "깊이 영상을 사용하는 움직임 검출(Motion Detection Using Depth Images)"인 미국 특허 출원 제12/641,788호; 및 2009년 10월 7일자로 출원된, 발명의 명칭이 "사람 추적 시스템(Human Tracking System)"인 미국 특허 출원 제12/575,388호.

[0053] 도 6은 처리 유닛(4)의 다양한 구성요소들을 나타낸 블록도이다. 도 6은 전력 관리 회로(306)와 통신하는 제어 회로(304)를 도시하고 있다. 제어 회로(304)는 CPU(central processing unit)(320), GPU(graphics processing unit)(322), 캐시(324), RAM(326), 메모리(330)(예컨대, DRAM)와 통신하는 메모리 제어기(328), 플래시 메모리(334)(또는 다른 유형의 비휘발성 저장소)와 통신하는 플래시 메모리 제어기(332), 대역 인터페이스(302) 및 대역 인터페이스(232)를 통해 헤드 마운티드 디스플레이 디바이스(2)와 통신하는 디스플레이 출력 버퍼(display out buffer)(336), 대역 인터페이스(302) 및 대역 인터페이스(232)를 통해 헤드 마운티드 디스플레이 디바이스(2)와 통신하는 디스플레이 입력 버퍼(display in buffer)(338), 마이크에 연결하기 위해 외부 마이크 커넥터(342)와 통신하는 마이크 인터페이스(340), 무선 통신 디바이스(346)에 연결하기 위한 PCI Express 인터페이스, 및 USB 포트(들)(348)를 포함한다. 하나의 실시예에서, 무선 통신 디바이스(346)는 Wi-Fi 지원 통신 디바이스, 블루투스 통신 디바이스, 적외선 통신 디바이스 등을 포함할 수 있다. USB 포트는, 데이터 또는 소프트웨어를 처리 유닛(4)에 로드하는 것은 물론 처리 유닛(4)을 충전시키기 위해, 처리 유닛(4)을 컴퓨팅 디바이스(도시 생략)에 도킹시키는 데 사용될 수 있다. 하나의 실시예에서, CPU(320) 및 GPU(322)는 가상 3차원 물체들을 언제, 어디서 그리고 어떻게 사용자의 뷰에 삽입해야 하는지를 결정하기 위한 주된 장비이다. 추가의 상세들이 이하에서 제공된다.

[0054] 전력 관리 회로(306)는 클럭 발생기(360), 아날로그-디지털 변환기(262), 배터리 충전기(364), 전압 조절기(366), 헤드 마운티드 디스플레이 전원(376), 및 온도 센서(374)[어쩌면 처리 유닛(4)의 손목 밴드 상에 위치해

있음]와 통신하는 온도 센서 인터페이스(372)를 포함한다. 아날로그-디지털 변환기(362)는 배터리 전압, 온도 센서를 모니터링하고 배터리 충전 기능을 제어하는 데 사용된다. 전압 조절기(366)는 시스템에 전력을 공급하기 위해 배터리(368)와 통신한다. 배터리 충전기(364)는 충전 잭(370)으로부터 전력을 받을 때 [전압 조절기(366)를 통해] 배터리(368)를 충전시키는 데 사용된다. HMD 전원(376)은 전력을 헤드 마운티드 디스플레이 디바이스(2)에 제공한다.

[0055]

앞서 기술한 헤드 마운티드 디스플레이 디바이스(2) 및 처리 유닛(4)은, 가상 3차원 물체가 현실 세계의 뷰를 보강 및/또는 대체하도록, 가상 3차원 물체를 하나 이상의 사용자들의 FOV에 삽입할 수 있다. 살펴본 바와 같이, 처리 유닛(4)은 부분적으로 또는 전체적으로 헤드 마운티드 디스플레이(2)에 통합될 수 있고, 따라서 장면에 대한 깊이 맵을 발생시키기 위한 앞서 기술한 계산이 헤드 마운티드 디스플레이(2) 내에서 수행된다. 추가의 실시예들에서, 장면에 대한 깊이 맵을 발생시키기 위한 앞서 기술한 계산의 일부 또는 전부가, 다른 대안으로서 또는 그에 추가하여, 핸드헬드 디바이스(12) 내에서 수행될 수 있다.

[0056]

하나의 예시적인 실시예에서, 헤드 마운티드 디스플레이(2) 및 처리 유닛(4)은 하나 이상의 사용자들이 있는 환경의 장면 맵 또는 모델을 생성하고 그 환경 내의 다양한 움직이는 물체들을 추적하기 위해 협력한다. 그에 추가하여, 헤드 마운티드 디스플레이(2) 및 처리 유닛(4)은, 헤드 마운티드 디스플레이 디바이스(2)의 위치 및 배향을 추적하는 것에 의해, 사용자(18)에 의해 착용된 헤드 마운티드 디스플레이 디바이스(2)의 FOV를 추적할 수 있다. 헤드 마운티드 디스플레이 디바이스(2)에 의해 획득된 센서 정보는 처리 유닛(4)으로 전송되고, 하나의 실시예에서, 처리 유닛(4)은 이어서 장면 모델을 업데이트할 수 있다. 처리 유닛(4)은 이어서 헤드 마운티드 디스플레이 디바이스(2)로부터 수신하는 부가의 센서 정보를 사용하여, 사용자의 FOV를 개선시키고, 가상 3차원 물체를 언제, 어디서 그리고 어떻게 삽입할지에 관한 명령어들을 헤드 마운티드 디스플레이 디바이스(2)에 제공한다. 포착 디바이스(125) 내의 카메라들로부터의 센서 정보에 기초하여, 장면 모델 및 추적 정보가 이하에서 설명되는 바와 같이 페루프 피드백 시스템에서 헤드 마운티드 디스플레이(2)와 처리 유닛(4) 사이에서 주기적으로 업데이트될 수 있다.

[0057]

도 1 및 도 7 내지 도 9를 참조하면, 본 개시 내용은 장면 내에 투사되는 가상 물체들과 직접 상호작용하기 위해 사용될 수 있는 핸드헬드 디바이스(12)를 추가로 포함한다. 핸드헬드 디바이스(12)의 위치 및 움직임(병진 및/또는 회전)이 매 프레임마다 업데이트될 수 있도록, 핸드헬드 디바이스(12)가 이하에서 설명되는 바와 같이 헤드 마운티드 디스플레이 디바이스(2) 및 처리 유닛(4)에 의해 발생하는 장면 맵 내에 등록될 수 있다. 이것은 핸드헬드 디바이스(12)와 장면 내의 가상 물체들 간의 직접 상호작용을 가능하게 한다. "직접" 대 "간접"은, 본 명세서에서 사용되는 바와 같이, 사용자의 손과 같은 장면 내의 비등록된 물체들의 위치가 포착된 깊이 데이터 및 신체 부위들을 식별하는 데 사용되는 골격 추적 소프트웨어에 기초하여 추정된다는 사실을 말한다. 때때로, 손 또는 다른 신체 부위를 추적할 때, 정확한 배향을 도출하거나 깊이 맵에 따라 정확한 손 모델을 신뢰성있게 근사시키는 것이 어려울 수 있다. 그에 따라, 사용자의 손과 같은 비등록된 물체들의 위치를 "직접" 알지 못한다. 사용자가 손을 사용하여 가상 물체들과 상호작용할 때, 이 상호작용은, 손 위치의 상기 추정에 기초하여, 간접적이라고 말해진다.

[0058]

이와 달리, 핸드헬드 디바이스의 위치가 헤드 마운티드 디스플레이(2) 및 처리 유닛(4)에 의해 발생한 동일한 장면 맵 내에 등록되어 있다[디바이스(2) 및 유닛(4)이 때로는 모두 합하여 본 명세서에서 모바일 디스플레이 디바이스라고 지칭될 수 있다]. 이하에서 설명되는 바와 같이, 하나의 예에서, 핸드헬드 디바이스(12)는 모바일 디스플레이 디바이스에 의해 작성된 장면 맵 내의 동일한 지점들과 동일시될 수 있는 지점들을 식별할 수 있는 카메라를 포함한다. 그 공통의 지점들이 식별되면, 모바일 디스플레이 디바이스의 장면 맵 내에서 핸드헬드 디바이스(12)의 위치를 식별하고 등록시키기 위해 다양한 방법들이 사용될 수 있다.

[0059]

도 7은 핸드헬드 디바이스(12)의 사시도를 도시한 것이다. 디바이스(12)는 일반적으로 영상 포착 디바이스(2)에 고정되게 탑재되거나 그와 일체로 형성된 껍(20)을 포함할 수 있다. 껍(20)은 다수의 기능들을 수행할 수 있다. 하나의 이러한 기능은 사용자가 장면 내의 가상 물체들과의 상호작용을 제어할 수 있게 하는 입력/피드백 디바이스이다. 상세하게는, 껍(20)은 사용자 입력을 수신하기 위한 입력 패드(24)를 포함할 수 있다. 하나의 예에서, 입력 패드(24)는 용량성 또는 기타 터치 감응 스크린을 포함할 수 있다. 이러한 예들에서, 입력 패드(24)는 그래픽 버튼, 휠(wheel), 슬라이드 또는 기타 컨트롤 - 각각은 가상 물체와의 상호작용을 용이하게 하기 위한 사전 정의된 명령들과 연관되어 있음 - 을 디스플레이하는 하나 이상의 화면들을 디스플레이할 수 있다. 공지된 바와 같이, 이러한 예에서의 주어진 명령은 그래픽 버튼, 휠, 슬라이드 등을 작동시키기 위해 스크린과의 사용자 접촉에 의해 발생할 수 있다. 추가의 실시예들에서, 터치 감응 스크린 대신에, 입력 패드가 앞서 기술한 바와 같이 명령을 수행하기 위해 작동될 수 있는 실제 버튼, 휠, 슬라이드 또는 기타 컨트롤로 형

성될 수 있다.

[0060] 많은 가능한 예들 중 하나로서, 사용자는, 도 1에 도시된 바와 같이, 핸드헬드 디바이스(12)로부터 광선을 뺏어 나가게 하기 위해 입력 패드(24) 상의 컨트롤을 작동시킬 수 있다. 해당 컨트롤을 작동시킬 때, 핸드헬드 디바이스(12)의 전방으로부터 뺏어 있는 가상 광선(28)이 발생되고 모바일 디스플레이 디바이스를 통해 사용자에게 디스플레이될 수 있다. 광선(28)의 사용은 이하에서 설명된다. 다른 예로서, 사용자는 가상 물체를 잡기 위해 입력 패드(24) 상의 컨트롤을 작동시킬 수 있다. 이러한 예에서, 시스템은 가상 물체의 표면 상에서의 또는 가상 물체 내에서의 핸드헬드 디바이스(12)의 접촉을 검출하고, 그 후에 가상 물체의 위치를 핸드헬드 디바이스(12)에 관련시킬 수 있다. 사용자는 그 후에 컨트롤을 놓거나(release) 입력 패드(24) 상의 다른 컨트롤을 작동시키는 것에 의해 가상 물체를 놓을 수 있다. 예를 들어, 다음과 같은 것들을 비롯하여, 각종의 다른 명령을 수행하기 위해 추가의 버튼, 휠 및 슬라이드가 사용될 수 있다:

- [0061] • 가상 물체를 핸드헬드 디바이스(12)로부터 멀어지게 미는 것,
- [0062] • 가상 물체를 핸드헬드 디바이스(12)에 더 가까워지게 당기는 것,
- [0063] • 가상 물체를 뒤로, 앞으로, 좌측으로, 우측으로, 위로 또는 아래로 움직이는 것,
- [0064] • 가상 물체를 크기 조정하는 것,
- [0065] • 가상 물체를 회전시키는 것,
- [0066] • 가상 물체를 복사 및/또는 붙여넣기하는 것,
- [0067] • 가상 물체를 제거하는 것,
- [0068] • 가상 물체의 색, 텍스처 또는 형상을 변경하는 것,
- [0069] • 물체를 사용자 정의 방식으로 가상 환경에서 이리저리 움직이도록 애니메이션화하는 것.

[0070] 다른 명령들이 생각되고 있다. 이 상호작용들은 입력 패드(24) 상에서 해당 명령을 선택하는 것에 의해 개시될 수 있다. 추가의 실시예들에서, 이 상호작용들은 입력 패드(24) 상에서 명령들을 선택하는 것과 어떤 다른 사전 정의된 제스처(신체 제스처 및/또는 구두 제스처)의 수행의 조합에 의해 개시될 수 있다. 추가의 실시예들에서, 앞서 기술한 상호작용들 중 적어도 일부는 입력 패드(24)와 무관한 신체 제스처들의 수행에 의해 수행될 수 있다.

[0071] 펍(20)은 사용자에게 피드백을 추가로 제공할 수 있다. 이 피드백은 입력 패드(24)를 통해 사용자에게 시각적으로 디스플레이되고 그리고/또는 펍(20) 상에 제공된 스피커를 통해 사용자에게 청각적으로 재생될 수 있다. 추가의 실시예들에서, 펍(20)은 사용자에게 햅틱 응답을 제공하는 진동 모터(519)(도 8)를 구비하고 있을 수 있다. 실시예들에서, 핸드헬드 디바이스는, 적어도 때때로, 사용자가 핸드헬드 디바이스가 아니라 장면을 보고 있도록 사용될 수 있다. 이와 같이, 사용자가 이하에서 설명되는 바와 같이 물체를 선택하고 있는 경우, 펍(20)은 사용자가 물체를 추적(lock onto)하거나 어떤 다른 의도된 동작을 성공적으로 수행했을 때를 나타내는 햅틱 응답을 제공할 수 있다.

[0072] 펍(20)의 다른 기능은 핸드헬드 디바이스(12)의 각 가속도 및 위치 정보 및/또는 병진 가속도 및 위치 정보를 제공하는 것이다. 펍(20)은 헤드 마운티드 디스플레이 유닛에서의 IMU(132)와 유사하거나 동일할 수 있는 IMU(511)(도 8)를 포함할 수 있다. 이러한 IMU는, 예를 들어, 핸드헬드 디바이스(12)의 위치, 방향, 및 가속도(피치, 롤 및 요)를 감지하는 3축 자력계, 3축 자이로 및 3축 가속도계와 같은 관성 센서를 포함할 수 있다. 앞서 살펴보고 이하에서 설명되는 바와 같이, 핸드헬드 디바이스(12)의 x, y 및 z 위치 및 방향이 핸드헬드 디바이스(12)와 모바일 디스플레이 디바이스의 협력을 통해 장면 맵에 등록된다. 그렇지만, 핸드헬드 디바이스(12) 내의 IMU에 의해 제공되는 데이터는 모바일 디스플레이 디바이스의 장면 맵에서의 핸드헬드 디바이스의 위치 및/또는 방향을 확인 및/또는 보완할 수 있다. 추가의 실시예들에서, 핸드헬드 디바이스(12) 내의 IMU가 생략될 수 있는 것이 생각되고 있다.

[0073] 도 8은 펍(20)의 내부에 있는 하드웨어 구성요소들의 일부의 하나의 예의 블록도를 도시한 것이다. 하나의 예에서, 펍(20)은 종래의 셀룰러 전화일 수 있다. 이러한 실시예들에서, 펍(20)은 셀룰러 전화에 대한 종래의 하

드웨어 구성을 가질 수 있고, 셀룰러 전화에 대해 종래에 알려진 기능들을 수행하기 위해 동작할 수 있다. 그
에 부가하여, 전화가 본 기술에 따라 동작할 수 있게 하기 위해, 소프트웨어 애플리케이션 프로그램 및 기타 소
프트웨어 구성요소들이 펍(20)에 로드될 수 있다. 추가의 실시예들에서, 펍(20)은 본 기술에 따라 동작하도록
커스터마이징된 전용 하드웨어 디바이스일 수 있다.

[0074] 펍(20)은 펍(20)의 동작 및 모바일 디스플레이 디바이스와 상호작용을 제어하기 위한 프로세서(502)를 포함할
수 있다. 앞서 살펴본 바와 같이, 펍(20)의 하나의 기능은 펍(20)에 관한 가속도 및 위치 정보를 제공하는 것
이다. 이 정보는 IMU(511)를 통해 프로세서(502)에 제공될 수 있다. 펍(20)은 프로세서(502)에 의해 실행되는
소프트웨어 코드, 그리고 가속도 및 위치 데이터, 영상 데이터 및 장면 맵과 같은 데이터를 저장하기 위한 메모
리(514)를 추가로 포함할 수 있다.

[0075] 펍(20)은 LCD 화면(520) 및 터치스크린(512) - 이들은 함께 앞서 기술된 입력 패드(24)로서 기능함 - 을 포함하
는 사용자 인터페이스를 추가로 포함할 수 있다. LCD 화면(520) 및 터치스크린(512)은, 각각, LCD 제어기(522)
및 터치스크린 제어기(513)를 통해 프로세서(502)와 통신할 수 있다. 터치스크린(512)은 LCD 화면(520)에 오버
레이된 용량성 표면일 수 있다. 그렇지만, 앞서 살펴본 바와 같이, 터치스크린(512)은, 추가의 실시예들에서,
LCD 화면(520)과 함께 각종의 물리적 액추에이터들 중 임의의 것에 의해 대체될 수 있다. 종래의 전화인 경우,
물리적 액추에이터들 중 적어도 일부는 앞서 기술한 바와 같이 사용자 입력을 제어하기 위한 기능들을 할당받을
수 있다.

[0076] 펍(20)은 펍(20)을, 예를 들어, 컴퓨팅 디바이스(도시 생략)와 같은 다른 디바이스에 연결시키기 위한 연결
(516)을 추가로 포함할 수 있다. 연결(516)은 USB 연결일 수 있지만, 직렬, 병렬, SCSI 및 IEEE
1394("Firewire") 연결들을 비롯한 다른 유형의 연결들이 제공될 수 있다는 것을 잘 알 것이다.

[0077] 펍(20)은 기술 분야에 공지된 바와 같이 카메라(518)를 추가로 포함할 수 있다. 카메라(518)는 카메라(22)와
관련하여 이하에서 기술되는 것들 중 일부, 전부 및/또는 부가의 구성요소들을 가질 수 있다. 실시예들에서,
펍(20)은 카메라(518) 또는 카메라(22)에 의해 포착되는 FOV를 디스플레이할 수 있다.

[0078] 앞서 살펴본 바와 같이, 펍(20)은 햅틱 피드백을 제공할 수 있는 진동 모터(519) 및 오디오를 제공하기 위한 스피
커(530)를 비롯한 다양한 피드백 구성요소들을 포함할 수 있다. 음성 명령들을 수신하기 위한 공지된 구성의
마이크(532)가 추가로 제공될 수 있다.

[0079] 펍(20)은 펍(20)과 모바일 디스플레이 디바이스와 같은 다른 구성요소들 간의 통신을 가능하게 하는 구성요소들
을 추가로 포함할 수 있다. 이들 구성요소는 처리 유닛(4)의 무선 통신 디바이스(346)를 통해 안테나(542)를
거쳐 모바일 디스플레이 디바이스와 무선 통신을 할 수 있는 통신 인터페이스(540)를 포함한다. 펍(20)은 이하
에서 기술하는 바와 같이 카메라(22)에 하드와이어링(hardwire)될 수 있지만, 추가의 실시예들에서, 통신 인터
페이스(540)를 통해 무선으로 연결되어 통신할 수 있다.

[0080] 더욱이, 통신 인터페이스(540)는, 본 기술의 실시예들에서, 모바일 디스플레이 디바이스 및 카메라(22) 이외의
구성요소들로/로부터 전송을 송신 및 수신할 수 있다. 예를 들어, 펍(20)은 사진 및 비디오 영상과 같은 데이
터는 물론, 애플리케이션 프로그램, API, 업데이트, 패치 등과 같은 소프트웨어를 전송하기 위해 호스트 컴퓨터
와 통신할 수 있다. 통신 인터페이스(540)는 또한, 본 기술의 실시예들에 따라, 핸드헬드 컴퓨터, PDA 및 기타
모바일 디바이스를 비롯한 핸드헬드 컴퓨팅 디바이스와 같은 다른 디바이스들과 통신하기 위해 사용될 수 있다.
통신 인터페이스(540)는 펍(20) 및 카메라(22)를, LAN(local area network), WAN(wide area network) 및 인터
넷을 비롯한 각종의 네트워크들에 연결하기 위해 사용될 수 있다.

[0081] 비록 불가결한 것은 아니지만, 펍(20)은 수신된 디지털 및 아날로그 신호를 처리하기 위한 디지털 기저대역 및/
또는 아날로그 기저대역을 추가로 포함할 수 있다. 아날로그 음성 신호와 같은 아날로그 신호를 안테나(510)를
통해 수신 및 전송하기 위해 RF 송수신기(506) 및 스위치(508)가 제공될 수 있다. 실시예들에서, 송수신기
(504)는 직교 변조 및 복조는 물론, 이중 대역(800 및 1900 MHz) RF와 기저대역 간의 업 컨버전 및 다운 컨버전
을 수행할 수 있다. 본 명세서에 기술된 다양한 통신 인터페이스들은 송수신기(506) 및 스위치(508)와 같은 송
수신기 및/또는 스위치를 포함할 수 있다.

[0082] 본 기술의 대안의 실시예들에서, 펍(20)이 각종의 다른 구성들 및 부가의 또는 대안의 구성요소들을 가질 수 있
다는 것을 잘 알 것이다.

[0083] 다시 도 7을 참조하면, 카메라(22)는, 실시예들에서, 포착 디바이스(125)와 유사한 디바이스일 수 있고, 따라서
포착 디바이스(125)에 대한 이상의 설명이 마찬가지로 카메라(22)에 적용된다. 추가의 실시예들에서, 카메라

(22)는 그 대신에 단지 정지 영상 및 비디오 영상을 포착할 수 있는 보통의 시판 중인 카메라일 수 있다.

[0084] 카메라(22)가, 도시된 바와 같이, 꺾(20) 아래에 고정되어 있을 수 있지만, 추가의 실시예들에서, 카메라(22)가 꺾(20)의 전방에, 측면에, 또는 심지어 후방에 고정되어 있을 수 있다. 카메라(22)가 브라켓(bracket)(30) 및 패스너(fastener)를 통해 꺾(20)에 고정될 수 있지만, 추가의 실시예들에서, 카메라(22)가 꺾(20)과 일체로 형성될 수 있다. 도시된 예에서, 카메라는 전방을 향해 있다. 이것은 카메라가 사용자의 전방의 FOV를 포착할 수 있는 반면, 입력 패드(24)가 입력 패드(24)를 보기 쉽도록 사용자 쪽으로 위로 향해 있다는 장점을 제공한다. 그렇지만, 추가의 실시예들에서, 카메라 렌즈가 입력 패드(24)의 표면에 일반적으로 평행하도록 카메라가 위쪽으로 향해 있을 수 있다. 추가의 실시예들에서, 카메라 렌즈는 입력 패드의 표면에 대해 얼마간 경사진 각도로 있을 수 있다. 카메라(22)가 생략될 수 있고 꺾(20) 내의 카메라(518)가 카메라(22)의 기능을 수행할 수 있는 것이 추가로 생각되고 있다.

[0085] 앞서 살펴본 바와 같이, 핸드헬드 디바이스(12) 및 모바일 디스플레이는 핸드헬드 디바이스(12)의 정확한 위치를, 앞서 기술한 바와 같이 모바일 디스플레이 디바이스에 의해 결정된 FOV의 x, y, z 장면 맵에 등록시키기 위해 협력할 수 있다. 하나의 등록 방법이 도 11의 플로우차트와 관련하여 이하에서 기술된다. 그렇지만, 다른 등록 방법들이 가능하다.

[0086] 꺾(20)의 특정의 구성이 도 7에 도시되어 있지만, 꺾(20)이 각종의 상이한 구성들을 가지며 앞서 기술한 기능을 제공할 수 있다는 것을 잘 알 것이다. 추가의 실시예에서, 카메라(22)가 생략될 수 있고, 모든 추적 기능이 꺾(20) 내에 제공된 IMU(511)에 의해 수행될 수 있다.

[0087] 앞서 기술한 구성요소들을 사용하여, 사용자들은 가상 영상들을 발생시키는 모바일 디스플레이 디바이스에 의해 사용되는 동일한 장면 맵 내에 등록되어 있는 핸드헬드 디바이스(12)를 사용하여 가상 환경 내의 가상 물체들과 직접 상호작용할 수 있다. 하나의 예가 도 1에 도시되어 있다. 사용자는 디바이스(12)로부터 광선을 뺏어나가게 하는 것에 의해 핸드헬드 디바이스(12)를 사용하여 물체를 선택하고자 하는 바램을 나타낼 수 있다. 꺾(20)의 입력 패드(24) 상의 해당 명령을 선택할 때, 모바일 디스플레이 디바이스는 핸드헬드 디바이스(12)의 일부 부분으로부터(예컨대, 전방으로부터) 뺏어나가는 가상 광선(28)을 디스플레이한다. 사용자가 입력 패드(24)와의 상호작용 이외의 제스처를 수행하는 것에 의해 광선(28)이 나타날 수 있다는 것을 잘 알 것이다. 모바일 디스플레이 디바이스 및 핸드헬드 디바이스로 이루어진 시스템(10)이 핸드헬드 디바이스의 정확한 위치 및 배향을 알고 있기 때문에, 디바이스(12)가 회전되거나 이리저리 움직일 때 광선(28)이 핸드헬드 디바이스(12) 상의 고정점으로부터 방출하는 것처럼 디스플레이될 수 있다. 더욱이, 디바이스(12)가 회전되거나 이리저리 움직일 때, 광선이 디바이스(12)와 일대일 관계로 움직인다.

[0088] 사용자는 광선(28)을 사용하여 실제 또는 가상 물체를 가리킬 수 있고, 광선(28)은 실제 또는 가상 물체와 교차할 때까지 뺏어갈 수 있다. 사용자는 광선(28)을 가상 물체로 지향시키는 것에 의해 가상 물체와 직접 상호작용할 수 있다. 광선이 도 1의 가상 물체(21)와 같은 가상 물체와 교차하면, 그 가상 물체의 선택을 나타내기 위해 피드백이 사용자에게 제공될 수 있다. 앞서 살펴본 바와 같이, 피드백은 시각적, 청각적 및/또는 햅틱일 수 있다. 실시예들에서, 사용자는 물체들의 원치 않는 선택을 방지하기 위해 물체가 선택된 것으로 간주되기 전에 어떤 미리 결정된 기간 동안 광선(28)이 가상 물체를 겨누고 있는 채로 유지할 필요가 있을 수 있다. 사용자가 다른 물체(실제 물체 또는 가상 물체)에 의해 가려져 있는 물체를 선택하고자 하는 경우가 있을 수 있고, 본 시스템의 이동성에 의해, 원하는 물체로의 명확한 시선(line of sight)이 있을 때(이 시점에서 사용자는 물체를 선택할 수 있음)까지 사용자가 환경에서 이리저리 움직일 수 있다.

[0089] 선택되면, 사용자는 임의의 수의 방식으로 물체와 상호작용할 수 있다. 사용자는 가상 물체를 광선을 따라 더 가까워지게 또는 더 멀어지게 이동시킬 수 있다. 사용자는, 그에 부가하여 또는 다른 대안으로서, 물체가 광선에 고정된 상태에서, 광선을 위치 변경하여 가상 물체를 정확히 원하는 장소에 위치시킬 수 있다. 부가의 잠재적인 상호작용들이 앞서 기술되었다.

[0090] 도 1은 사용자가 입력 패드(24) 상의 해당 명령을 선택할 때 핸드헬드 디바이스(12)로부터 뺏어나가는 가상 광선을 통해 가상 물체(21)가 선택되는 상호작용을 예시하고 있다. 예를 들어, 도 9에 도시된 추가의 실시예들에서, 사용자는 물체와 핸드헬드 디바이스(12)를 물리적으로 접촉시키는 것에 의해 가상 물체와 상호작용할 수 있다. 이러한 실시예들에서, 사용자는 가상 물체(21)를 선택하기 위해 핸드헬드 디바이스(12)의 일부분을 가상 물체(21)의 표면과 접촉하게 또는 가상 물체(21)의 내부에 위치시킬 수 있다. 그 후에, 사용자는 가상 물체(21)와 상호작용하기 위해 입력 패드(24) 상의 컨트롤을 선택하거나 신체 제스처를 수행할 수 있다. 앞서 살펴본 바와 같이, 이 상호작용은 물체를 새로운 위치로 운반하여 내려 놓는 것, 물체를 복제하는 것, 물체를 제거

하는 것 등과 같은 각종의 상호작용들 중 임의의 것일 수 있다.

- [0091] 사용자 접촉 시에 물체를 잡는 것 대신에, 물체가 그 대신에 물체와의 충돌의 결과로서 "튀어" 나갈 수 있다. 충돌에 대한 물체의 반응은 물리학에 의해 규정될 수 있고 정확할 수 있다. 즉, 충돌 시의 핸드헬드 디바이스(12)의 속도를 IMU 및 기타 데이터로부터 정확하게 알 수 있기 때문에, 가상 물체가 정확한 속도로 튀어나갈 수 있다. 이 속도는 물리학 및 가상 물체에 대해 규정된 일련의 변형 및 탄성 특성들에 의해 구해질 수 있다.
- [0092] 이하에서 설명되는 바와 같이, 장면 맵에서의 가상 물체들의 위치들을, 예를 들어, 처리 유닛(4)이 알고 있다. 핸드헬드 디바이스(12)를 동일한 장면 맵 내에 등록시키는 것에 의해, 사용자는 장면 맵 내의 가상 물체들과 직접 상호작용할 수 있거나, 새로운 가상 물체들을 장면 맵에 생성할 수 있다 - 그 가상 물체들이 이어서 헤드 마운티드 디스플레이 디바이스(12)를 통해 디스플레이됨 -. 이러한 직접적인 상호작용은 가상 환경 내의 정확한 위치들에서 그리고 정확한 방식으로 가상 물체들의 상호작용 및/또는 생성을 가능하게 한다.
- [0093] 더욱이, 본 시스템은 장비가 설치되지 않은 환경(non-instrumented environment)에서 동작한다. 즉, 어떤 종래 기술의 시스템들은 영상 포착 디바이스들의 FOV 내의 물체들의 위치들을 결정하기 위해 링 또는 다른 구성의 고정된 영상 포착 디바이스들을 사용한다. 그렇지만, 모바일 디스플레이 디바이스 및 핸드헬드 디바이스(12) 둘 다 사용자와 함께 움직일 수 있기 때문에, 본 기술은 사용자가 움직이는 임의의 환경에서 동작할 수 있다. 환경을 사전에 설치(set up)할 필요가 없다.
- [0094] 꺾(20)의 특징의 구성이 도 7에 도시되어 있지만, 꺾(20)이 각종의 상이한 구성들을 가지며 앞서 기술한 기능을 제공할 수 있다는 것을 잘 알 것이다. 하나의 추가의 실시예에서, 꺾(20)은 가상 물체들을 표적으로 하는 게임 애플리케이션에서 사용하기 위한 총 또는 사격할 수 있는 어떤 다른 물체로서 구성될 수 있다. 핸드헬드 디바이스(12)의 위치 및 배향이 정확하게 알려져 있고 가상 표적들을 디스플레이하는 모바일 디스플레이 유닛의 기준 프레임(frame of reference) 내에 등록되어 있기 때문에, 정확한 사격 재현이 제공될 수 있다. 추가의 실시예들에서, 꺾(20)이 다른 응용들에서 사용될 수 있다.
- [0095] 도 10은 단일의 영상 데이터 프레임을 발생시키고 렌더링하며 각각의 사용자에게 디스플레이하는 데 걸리는 시간과 같은 이산 기간(discrete time period) 동안 처리 유닛(4), 헤드 마운티드 디스플레이 디바이스(2), 및 핸드헬드 디바이스(12)의 동작 및 상호작용의 상위 레벨 플로우차트이다. 실시예들에서, 처리 유닛(4), 헤드 마운티드 디스플레이 디바이스(2) 및 핸드헬드 디바이스(12)에서 일어나는 프로세스들이 병행하여(in parallel) 행해질 수 있지만, 추가의 실시예들에서는, 그 단계들이 순차적으로(serially) 행해질 수 있다. 더욱이, 각각의 구성요소 내에서의 단계들이 단계별로 순차적으로 행해지는 것으로 도시되어 있지만, 구성요소 내에서의 단계들 중 하나 이상이 서로 병행하여 행해질 수 있다. 예를 들어, 처리 유닛(4)에서의 장면 맵의 결정 단계, 가상 영상 위치의 평가 단계 및 영상 렌더링 단계(각각이 이하에서 설명됨) 모두가 서로 병행하여 행해질 수 있다.
- [0096] 추가로, 상이한 구성요소들 내에서 또는 동일한 구성요소 내에서 행해지는 병행 단계들이 상이한 프레임 레이트로 행해질 수 있다는 것을 잘 알 것이다. 실시예들에서, 디스플레이된 영상이 60 Hz의 레이트로 리프레시(refresh)될 수 있지만, 추가의 실시예들에서, 더 빈번히 또는 덜 빈번히 리프레시될 수 있다. 달리 언급하지 않는 한, 도 10의 이하의 설명에서, 단계들은 헤드 마운티드 디스플레이 디바이스(2) 내의 하나 이상의 프로세서들이 단독으로 동작하는 것, 처리 유닛(4) 내의 하나 이상의 프로세서들이 단독으로 동작하는 것, 핸드헬드 디바이스(12) 내의 하나 이상의 프로세서들이 단독으로 동작하는 것, 또는 디바이스(2), 유닛(4) 및 디바이스(12) 중 2 개 이상으로부터의 프로세서들의 조합이 협력하여 동작하는 것에 의해 수행될 수 있다.
- [0097] 일반적으로, 시스템은 환경 및 환경 내의 물체들(사용자, 현실 세계 물체 및 가상 물체 등)의 x, y, z 좌표들을 가지는 장면 맵을 발생시킨다. 시스템은 또한 각각의 사용자의 FOV를 추적한다. 사용자들이 어떠한 장면의 동일한 측면들을 보고 있을 수 있지만, 사용자들은 그들을 상이한 시점들로부터 보고 있다. 이와 같이, 시스템은 가상 또는 현실 세계 물체들의 상이한 관찰 시점(viewing perspective), 시차 및 폐색 - 이들이 또다시 각각의 사용자에게 대해 상이할 수 있음 - 에 대해 조절하기 위해 장면의 각각의 사람의 FOV를 발생시킨다.
- [0098] 주어진 영상 데이터 프레임에 대해, 사용자의 뷰는 하나 이상의 실제 및/또는 가상 물체들을 포함할 수 있다. 사용자가 머리를, 예를 들어, 좌에서 우로 또는 위아래로, 돌릴 때, 사용자의 FOV에서의 현실 세계 물체들의 상대 위치가 본질적으로 사용자의 FOV 내에서 움직인다. 예를 들어, 도 1에서의 식물(27)이 처음에는 사용자의 FOV의 우측에 나타날 수 있다. 그러나, 사용자가 이어서 머리를 오른쪽으로 돌리는 경우, 식물(27)은 결국 사용자의 FOV의 좌측에 있게 되어 버릴 수 있다.

- [0099] 그렇지만, 사용자가 머리를 움직일 때 가상 물체들을 사용자에게 디스플레이하는 것은 더 어려운 문제이다. 사용자가 그의 FOV 내의 가상 물체를 보고 있는 일례에서, 사용자가 FOV를 왼쪽으로 이동시키기 위해 머리를 왼쪽으로 움직이는 경우, 가상 물체의 디스플레이가 사용자의 FOV 천이의 양만큼 오른쪽으로 천이될 수 있고, 따라서 순 효과는 가상 물체가 FOV 내에서 정지한 채로 있다.
- [0100] 단계들(604 및 606)에서, 모바일 디스플레이 디바이스 및 핸드헬드 디바이스(12)는 장면으로부터 데이터를 수집한다. 이것은 카메라(22) 및/또는 포착 디바이스들(125)의 깊이 카메라(426) 및 RGB 카메라(428)에 의해 감지되는 영상 데이터일 수 있다. 이것은 눈 추적 어셈블리(134)에 의해 감지된 영상 데이터일 수 있고, 이것은 IMU(132) 및 IMU(511)에 의해 감지된 가속도/위치일 수 있다.
- [0101] 단계(606)에서, 장면 데이터가, 예를 들어, 처리 유닛(4)과 같은, 시스템(10) 내의 처리 유닛들 중 하나 이상에 의해 수집된다. 프로세스가 처리 유닛(4)에 의해 수행되는 것으로 기술되는 이하의 설명에서, 프로세스가 시스템(10) 내의 프로세서들 중 하나 이상에 의해 수행될 수 있다는 것을 잘 알 것이다. 단계(608)에서, 처리 유닛(4)은 포착 디바이스(125) 및 카메라(22)의 영상 데이터의 조정(coordination)을 가능하게 하는 다양한 설정 동작(setup operation)들을 수행한다. 상세하게는, 단계(608)에서, 모바일 디스플레이 디바이스 및 핸드헬드 디바이스(12)는 모바일 디스플레이 디바이스의 기준 프레임에 핸드헬드 디바이스(12)의 위치를 등록시키기 위해 협력할 수 있다. 단계(608)의 추가 상세들이 이제부터 도 11의 플로우차트를 참조하여 설명될 것이다. 이하의 설명에서, 포착 디바이스들(125) 및 카메라(22)는 모두 합하여 촬영 디바이스라고 지칭될 수 있다.
- [0102] 단계(608)의 하나의 동작은 단계(670)에서 시스템(10) 내의 다양한 촬영 디바이스들의 클럭 오프셋들을 결정하는 것을 포함할 수 있다. 상세하게는, 시스템 내의 촬영 디바이스들 각각으로부터의 영상 데이터를 조정하기 위해, 조정되는 영상 데이터가 동일한 시각으로부터인 것이 확인될 수 있다. 영상 데이터의 클럭 오프셋 및 동기를 결정하는 것에 관한 상세들은 2010년 5월 3일자로 출원된, 발명의 명칭이 "이기종 영상 센서 동기화(Heterogeneous Image Sensor Synchronization)"인 미국 특허 출원 제12/772,802호, 및 2010년 6월 3일자로 출원된, 발명의 명칭이 "다수의 오디오비주얼 소스들로부터의 정보의 합성(Synthesis Of Information From Multiple Audiovisual Sources)"인 미국 특허 출원 제12/792,961호(이들 출원은 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 개시되어 있다. 일반적으로, 포착 디바이스(125)로부터의 영상 데이터 및 카메라(22)로부터 들어오는 영상 데이터가, 예를 들어, 처리 유닛(4) 내의 단일의 마스터 클럭으로부터 타임 스탬핑(time stamp)된다. 주어진 프레임의 이러한 데이터에 대한 타임 스탬핑은 물론, 촬영 디바이스들 각각에 대한 기지의 해상도를 사용하여, 처리 유닛(4)은 시스템 내의 촬영 디바이스들 각각에 대한 시간 오프셋을 결정할 수 있다. 이것으로부터, 각각의 촬영 디바이스들로부터 수신되는 영상들 간의 차 및 그 영상들에 대한 조절이 결정될 수 있다.
- [0103] 단계(608)는 장면의 x , y , z 직교 좌표 공간(Cartesian space)에서 촬영 디바이스들의 위치들을 서로에 대해 교정하는 동작을 추가로 포함한다. 이 정보를 알게 되면, 시스템(10) 내의 하나 이상의 프로세서들은 장면 맵 또는 모델을 형성할 수 있고, 장면의 기하메트리(geometry) 및 장면 내의 물체들(사용자들을 포함함)의 기하메트리 및 위치들을 식별할 수 있다. 촬영 디바이스들의 영상 데이터를 서로에 대해 교정할 때, 깊이 및/또는 RGB 데이터가 사용될 수 있다. RGB 정보만을 사용하여 카메라 뷰들을 교정하는 기술은 2007년 5월 17일자로 공개된, 발명의 명칭이 "영상 기반 기하학적 정렬 및 물체 기반 제어를 사용한 영상 탐색 Navigating Images Using Image Based Geometric Alignment and Object Based Controls"인 미국 특허 공개 제2007/0110338호(이 공개 공보는 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 기술되어 있다.
- [0104] 시스템(10) 내의 촬영 디바이스들 각각은 어떤 렌즈 왜곡을 가질 수 있고, 이 렌즈 왜곡은 상이한 촬영 디바이스들로부터의 영상들을 교정하기 위해 보정될 수 있다. 시스템 내의 다양한 촬영 디바이스들로부터의 영상 데이터가 단계(604)에서 수신되면, 단계(674)에서 다양한 촬영 디바이스들에 대한 렌즈 왜곡을 참작하기 위해 영상 데이터가 조절될 수 있다. 주어진 촬영 디바이스(깊이 또는 RGB)의 왜곡은 카메라 제조업체에 의해 제공되는 기지의 특성일 수 있다. 그렇지 않은 경우, 예를 들어, 체커 보드 패턴과 같은 기지의 치수의 물체를 카메라의 FOV 내의 상이한 위치들에서 촬영하는 것을 포함하는 촬영 디바이스의 왜곡을 계산하는 알고리즘들이 공지되어 있다. 그 영상 내의 지점들의 카메라 뷰 좌표(camera view coordinate)들의 편차는 카메라 렌즈 왜곡의 결과일 것이다. 렌즈 왜곡의 정도를 알게 되면, 왜곡이 공지된 역행렬 변환에 의해 보정될 수 있고, 그 결과 주어진 카메라에 대한 점 구름(point cloud) 내의 점들의 균일한 촬영 디바이스 시야 맵(imaging device view map)이 얻어진다.
- [0105] 시스템은 그 다음에 단계(678)에서 각각의 촬영 디바이스에 의해 포착되는 왜곡 보정된 영상 데이터 점들을 카

메라 뷰로부터 직교 3D 세계 뷰(orthogonal 3-D world view)로 변환할 수 있다. 이 직교 3D 세계 뷰는 직교 x, y, z 직교 좌표계(orthogonal x, y, z Cartesian coordinate system)에서의 포착 디바이스(125) 및 카메라(22)에 의해 포착된 영상 데이터의 점 구름 맵이다. 카메라 뷰를 직교 3D 세계 뷰로 변환하기 위해 행렬 변환식을 사용하는 방법들은 공지되어 있다. 예를 들어, David H. Eberly의 "3d Game Engine Design: A Practical Approach To Real-Time Computer Graphics," Morgan Kaufman Publishers (2000)(이 간행물은 참조 문헌으로서 그 전체가 본 명세서에 포함됨)을 참고하기 바란다. 또한, 미국 특허 출원 제12/792,961호(참조 문헌으로서 이전에 포함됨)를 참고하기 바란다.

[0106]

시스템(10) 내의 각각의 촬영 디바이스는, 단계(678)에서, 직교 3D 세계 뷰를 구성할 수 있다. 주어진 촬영 디바이스로부터의 데이터 점들의 x, y, z 세계 좌표들은 여전히 단계(678)의 끝에서 그 촬영 디바이스의 시점으로부터의 것이고, 시스템(10) 내의 다른 촬영 디바이스들로부터의 데이터 점들의 x, y, z 세계 좌표들에 아직 상관되어 있지 않다. 그 다음 단계는 상이한 촬영 디바이스들의 다양한 직교 3D 세계 뷰들을 시스템(10) 내의 촬영 디바이스들에 의해 공유되는 단일의 전체 3D 세계 뷰로 변환하는 것이다.

[0107]

이것을 달성하기 위해, 시스템의 실시예들은 그 다음에, 단계(682)에서, 각자의 촬영 디바이스들의 세계 뷰들의 점 구름들에서 중요한 불연속(discontinuity)들 또는 단서(cue)들을 찾아낼 수 있다. 발견되면, 시스템은, 단계(684)에서, 상이한 촬영 디바이스들의 상이한 점 구름들 간에 동일한 단서들을 식별한다. 시스템이 2 개의 상이한 촬영 디바이스들의 2 개의 세계 뷰들이 동일한 단서들을 포함하는 것으로 판정할 수 있다면, 시스템은, 단계(688)에서, 서로 및 단서들에 대한 2 개의 촬영 디바이스들의 위치, 배향 및 초점 거리를 결정할 수 있다. 실시예들에서, 포착 디바이스들(125) 및 카메라(22)는 동일한 공통의 단서들을 공유하지 않을 것이다. 그렇지만, 그들이 적어도 하나의 공유 단서를 가지는 한, 시스템은 서로 및 단일의 전체 3D 세계 뷰에 대한 포착 디바이스들(125) 및 카메라(22)의 위치, 배향 및 초점 거리를 결정할 수 있을 것이다.

[0108]

영상 점 구름으로부터 단서들을 식별하는 다양한 공지된 알고리즘들이 존재한다. 이러한 알고리즘들은, 예를 들어, Mikolajczyk, K.와 Schmid, C의 "A Performance Evaluation of Local Descriptors," IEEE Transactions on Pattern Analysis & Machine Intelligence, 27, 10, 1615-1630. (2005)(이 논문은 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 기재되어 있다. 영상 데이터에서 단서들을 검출하는 추가의 방법은 SIFT(Scale-Invariant Feature Transform) 알고리즘이다. SIFT 알고리즘은 2004년 3월 23일자로 특허된, 발명의 명칭이 "영상에서 스케일 불변 특징들을 식별하는 방법 및 장치와 영상에서 물체를 찾아내는 데 그를 사용하는 것(Method and Apparatus for Identifying Scale Invariant Features in an Image and Use of Same for Locating an Object in an Image)"인 미국 특허 제6,711,293호(이 특허는 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 기술되어 있다. 다른 단서 검출 방법은 MSER(Maximally Stable Extremal Regions) 알고리즘이다. MSER 알고리즘은, 예를 들어, J. Matas, O. Chum, M. Urba, 및 T. Pajdla의 논문, "Robust Wide Baseline Stereo From Maximally Stable Extremal Regions," Proc. of British Machine Vision Conference, pages 384-396 (2002)(이 논문은 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 기술되어 있다.

[0109]

단계(684)에서, 촬영 디바이스들로부터의 점 구름들 사이에 공유되는 단서들이 식별된다. 개념적으로, 제1 카메라와 제1 카메라의 직교 좌표계에서의 한 그룹의 단서들 간에 제1 그룹의 벡터들이 존재하고 제2 카메라와 제2 카메라의 직교 좌표계에서의 그 동일한 그룹의 단서들 간에 제2 그룹의 벡터들이 존재하는 경우, 2 개의 좌표계가 서로에 대해 결정되어, 양 카메라를 포함하는 단일의 직교 좌표계로 될 수 있다. 2 개 이상의 카메라로부터의 점 구름들 간에 공유 단서들을 찾아내기 위한 다수의 공지된 기법들이 존재한다. 이러한 기법들은, 예를 들어, Arya, S., Mount, D.M., Netanyahu, N.S., Silverman, R., 및 Wu, A.Y.의 "An Optimal Algorithm For Approximate Nearest Neighbor Searching Fixed Dimensions," Journal of the ACM 45, 6, 891-923 (1998)(이 논문은 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 나와 있다. 앞서 포함된 Arya 등의 근사 최근방 이웃 해법(approximate nearest neighbor solution) 대신에 또는 그에 부가하여, 해싱 또는 상황 의존 해싱(context-sensitive hashing)(이들로 제한되지 않음)을 비롯한 다른 기법들이 사용될 수 있다.

[0110]

2 개의 상이한 촬영 디바이스들로부터의 점 구름들이 많은 충분한 수의 정합된 단서들을 공유하는 경우, 2 개의 점 구름들을 서로 상관시키는 행렬이, 예를 들어, RANSAC(Random Sampling Consensus) 또는 각종의 다른 추정 기법들에 의해, 추정될 수 있다. 복구된 기본 행렬(fundamental matrix)에 대한 이상치(outlier)인 정합들이 이어서 제거될 수 있다. 한 쌍의 점 구름들 간의 추정되는 기하학적으로 일관성있는 정합들을 찾아낸 후에, 정합들은 각자의 점 구름들에 대한 트랙들의 그룹으로 편성될 수 있고, 여기서 트랙은 점 구름들 간의 상호 정합하는 단서들의 그룹이다. 그룹에서의 제1 트랙은 제1 점 구름에서의 각각의 공통의 단서의 투영(projection)을 포함할 수 있다. 그룹에서의 제2 트랙은 제2 점 구름에서의 각각의 공통의 단서의 투영을 포함할 수 있다. 상

이한 카메라들로부터의 점 그룹들이 결정되어 단일의 직교 3D 현실 세계 뷰에서의 단일의 점 그룹으로 될 수 있다.

- [0111] 촬영 디바이스들의 위치들 및 배향들이 이 단일의 점 그룹 및 단일의 직교 3D 현실 세계 뷰와 관련하여 교정된다. 2 개의 점 그룹을 함께 결정하기 위해, 2 개의 점 그룹에 대한 트랙들의 그룹에서의 단서들의 투영들이 분석된다. 이들 투영으로부터, 시스템은 단서들과 관련한 포착 디바이스들(125)의 시점을 결정할 수 있고, 또한 단서들과 관련한 카메라(22)의 시점을 결정할 수 있다. 그로부터, 시스템은 점 그룹들을 결정하여, 이들 점 그룹 둘 다로부터의 단서들 및 다른 데이터 점들을 포함하는 단일의 점 그룹 및 단일의 직교 3D 현실 세계 뷰의 추정치를 얻을 수 있다. 이것이 행해지면, 시스템은 단일의 직교 3D 현실 세계 뷰 및 서로에 대한 촬영 디바이스들의 상대 위치 및 배향을 결정할 수 있다. 시스템은 단일의 직교 3D 현실 세계 뷰에 대한 각각의 카메라의 초점 거리를 추가로 결정할 수 있다.
- [0112] 이상에서 헤드 마운티드 디스플레이 디바이스(2) 및 핸드헬드 디바이스(12)를 단일의 장면 맵에 등록시키기 위한 하나의 방법을 기술하고 있지만, 추가의 실시예들에서, 헤드 마운티드 디스플레이 디바이스(2) 및 핸드헬드 디바이스(12)의 상대 위치들이 다른 방법들에 의해 결정될 수 있다는 것을 잘 알 것이다. 하나의 추가의 예로서, 헤드 마운티드 디스플레이 디바이스(2) 및 핸드헬드 디바이스(12) 중 하나 또는 둘 다는 상대방 디바이스의 FOV에서 한 번 상대방 디바이스에 의해 검출되고 추적될 수 있는 마커(marker)들을 포함할 수 있다.
- [0113] 다시 도 10을 참조하면, 단계(608)에서, 시스템이 교정되면, 단계(610)에서, 장면의 지오메트리는 물론, 장면 내의 물체들의 지오메트리 및 위치들을 식별해주는 장면 맵이 개발될 수 있다. 실시예들에서, 주어진 프레임에서 발생한 장면 맵은 장면 내의 사용자들, 현실 세계 물체들 및 가상 물체들의 x, y 및 z 위치들을 포함할 수 있다. 이 정보는 영상 데이터 수집 단계들(604 및 606) 동안 획득되고, 단계(608)에서 함께 교정된다. 단계들(608 및 610)에서 결정된 정보를 사용하여, 핸드헬드 디바이스(12)는, 단계(614)에서, 장면 맵에서의 그의 위치를 결정할 수 있다.
- [0114] 단계(614)에서, 시스템은 시스템(10) 내의 사용자들에 대한 각각의 헤드 마운티드 디스플레이 디바이스(2)의 x, y 및 z 위치, 배향 및 FOV를 결정한다. 단계(614)의 추가의 상세들은 발명의 명칭이 "가상 환경 내에서의 가상 물체 발생(Virtual Object Generation Within a Virtual Environment)"인 미국 특허 출원 제13/525,700호(이 출원은 참조 문헌으로서 그 전체가 본 명세서에 포함됨)에 제공되어 있다.
- [0115] 단계(660)에서, 핸드헬드 디바이스(12) 또는 처리 유닛(4)은 앞서 기술한 바와 같이 핸드헬드 디바이스(12)를 사용하는 가상 물체와의 사용자 상호작용이 있는지를 검사할 수 있다. 이러한 상호작용이 검출되는 경우, 단계(664)에서, 영향을 받는 가상 물체의 새로운 위치 및/또는 모습이 결정되고 저장되며, 단계(618)에서, 처리 유닛(4)에 의해 사용된다.
- [0116] 단계(618)에서, 시스템은 사용자 위치, FOV 및 핸드헬드 디바이스(12)와 가상 물체들 간의 상호작용의 장면 맵을 사용하여, 현재 시각에서의 가상 물체들의 위치 및 모습을 결정할 수 있다. 가상 물체의 디스플레이된 모습의 이들 변화가 시스템에 제공되고, 시스템은 이어서 단계(618)에서 사용자의 시점으로부터의 가상 3차원 물체의 배향, 모습 등을 업데이트할 수 있다.
- [0117] 단계(634)에서, 처리 유닛(4)[또는 시스템(10) 내의 다른 프로세서]은, 헤드 마운티드 디스플레이 디바이스(2)의 최종 FOV 내에 어찌면 나타날 수 있을 그 가상 물체들만이 렌더링되도록, 렌더링 동작들을 선별(cull)할 수 있다. 다른 가상 물체들의 위치들이 여전히 추적될 수 있지만, 그들은 렌더링되지 않는다. 추가의 실시예들에서, 단계(634)가 완전히 생략될 수 있고 전체 영상이 렌더링되는 것이 또한 생각될 수 있다.
- [0118] 처리 유닛(4)은 그 다음에, 단계들(610, 612 및 614)에서 결정된 장면 맵 및 FOV를 사용하여 렌더링 설정 동작들이 수행되는 렌더링 설정 단계(638)를 수행할 수 있다. 가상 물체 데이터가 수신되면, 처리 유닛은 FOV에 렌더링되어야 하는 가상 물체들에 대해 단계(638)에서의 렌더링 설정 동작들을 수행할 수 있다. 단계(638)에서의 렌더링 설정 동작은 최종 FOV에 디스플레이될 가상 물체(들)와 연관된 공통의 렌더링 작업들을 포함할 수 있다. 이들 렌더링 작업은, 예를 들어, 그림자 맵(shadow map) 발생, 조명, 및 애니메이션화를 포함할 수 있다. 실시예들에서, 렌더링 설정 단계(638)는 예측된 최종 FOV에 디스플레이될 가상 물체들에 대한 버텍스 버퍼(vertex buffer), 텍스처 및 상태와 같은 알맞은 그리기 정보(draw information)를 편집하는 것을 추가로 포함할 수 있다.
- [0119] 시스템은 그 다음에, 단계(644)에서, 사용자의 FOV에서의 페색 및 음영(shading)을 결정할 수 있다. 상세하게는, 장면 맵은 움직이는 및 움직이지 않는 물체들 그리고 가상 물체들을 비롯한 장면 내의 물체들의 x, y 및 z

위치들을 가진다. FOV에서의 사용자의 위치 및 물체로의 사용자의 시선을 알면, 처리 유닛(4)(또는 다른 프로세서)은 가상 물체가 눈에 보이는 현실 세계 물체의 사용자의 뷰를 부분적으로 또는 전체적으로 폐색하는지를 판정할 수 있다. 그에 부가하여, 처리 유닛(4)은 눈에 보이는 현실 세계 물체가 가상 물체의 사용자의 뷰를 부분적으로 또는 전체적으로 폐색하는지를 판정할 수 있다. 폐색은 사용자에게 관련(user-specific)되어 있을 수 있다. 가상 물체가 제2 사용자가 아니라 제1 사용자의 뷰에서 가리거나(block) 가려질 수 있다. 그에 따라, 처리 유닛(4)에서 각각의 사용자의 폐색 판정이 수행될 수 있다.

[0120] 단계(646)에서, 처리 유닛(4)의 GPU(322)는 그 다음에 사용자에게 디스플레이될 영상을 렌더링할 수 있다. 렌더링 동작들의 일부분이 렌더링 설정 단계(638)에서 이미 수행되고 주기적으로 업데이트되었을 수 있다.

[0121] 단계(650)에서, 처리 유닛(4)은 렌더링된 영상을 헤드 마운티드 디스플레이 디바이스(2)로 송신할 때가 되었는지, 또는 핸드헬드 디바이스(12) 및/또는 헤드 마운티드 디스플레이 디바이스(2)로부터의 보다 최근의 위치 피드백 데이터를 사용하여 영상의 추가적인 개선을 위한 시간이 여전히 있는지를 검사한다. 60 헤르츠 프레임 리프레시 레이트(frame refresh rate)를 사용하는 시스템에서, 단일의 프레임은 약 16 ms이다.

[0122] 업데이트된 영상을 디스플레이할 때가 되면, 하나 이상의 가상 물체들에 대한 영상들이, 시점 및 폐색을 참작하여, 적절한 픽셀에 디스플레이되도록 마이크로디스플레이(120)로 송신된다. 이 때, 불투명도 필터에 대한 제어 데이터가 또한, 불투명도 필터(114)를 제어하기 위해, 처리 유닛(4)으로부터 헤드 마운티드 디스플레이 디바이스(2)로 전송된다. 헤드 마운티드 디스플레이는 이어서, 단계(658)에서, 영상을 사용자에게 디스플레이할 것이다.

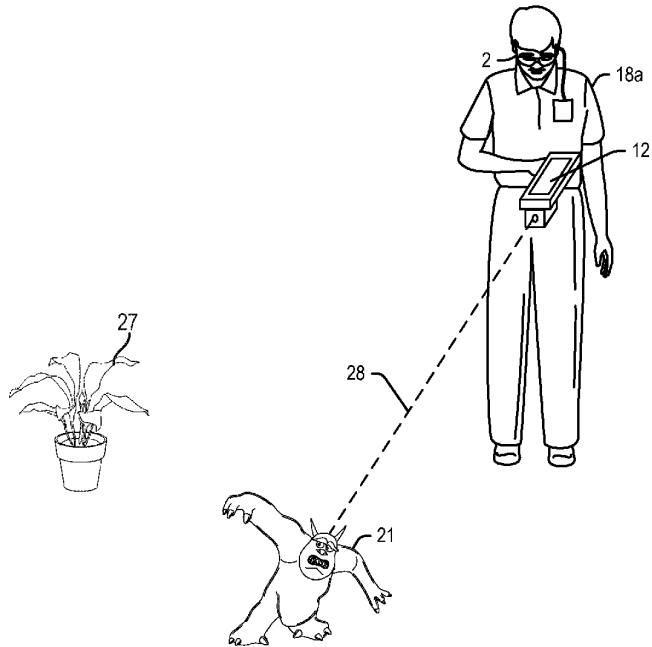
[0123] 다른 한편으로, 단계(650)에서, 디스플레이될 영상 데이터의 프레임을 송신할 때가 아직 되지 않은 경우, 처리 유닛은 최종 FOV 및 최종 FOV에서의 물체들의 최종 위치들의 예측들을 추가로 개선시키기 위해 더 많은 업데이트된 데이터를 위해 루프백할 수 있다. 상세하게는, 단계(650)에서, 여전히 시간이 있는 경우, 처리 유닛(4)은 헤드 마운티드 디스플레이 디바이스(2) 및 핸드헬드 디바이스(12)로부터 더 많은 최근의 센서 데이터를 얻기 위해 단계들(604 및 606)로 되돌아갈 수 있다.

[0124] 처리 단계들(604 내지 668)은 이상에서 단지 예로서 기술되어 있다. 추가의 실시예들에서, 이들 단계 중 하나 이상이 생략될 수 있거나, 단계들이 상이한 순서로 수행될 수 있거나, 부가의 단계들이 추가될 수 있다는 것을 잘 알 것이다.

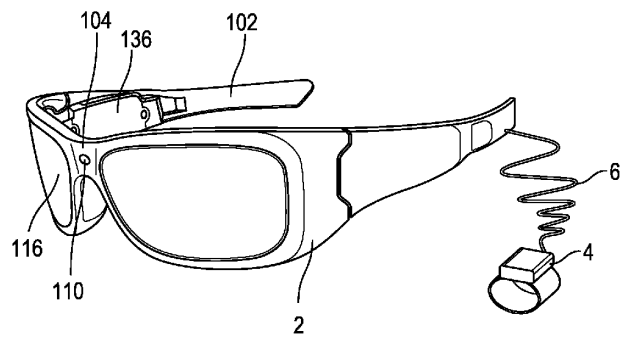
[0125] 발명 요지가 구조적 특징들 및/또는 방법 동작들과 관련하여 기술되어 있지만, 첨부된 청구범위에 한정되어 있는 발명 요지가 이상에서 기술한 구체적인 특징들 또는 동작들로 꼭 제한되는 것은 아님을 잘 알 것이다. 오히려, 이상에서 기술한 구체적인 특징들 및 동작들은 청구항들을 구현하는 예시적인 형태로서 개시되어 있다. 본 발명의 범위는 본 명세서에 첨부된 특허청구범위에 의해 한정되는 것으로 보아야 한다.

도면

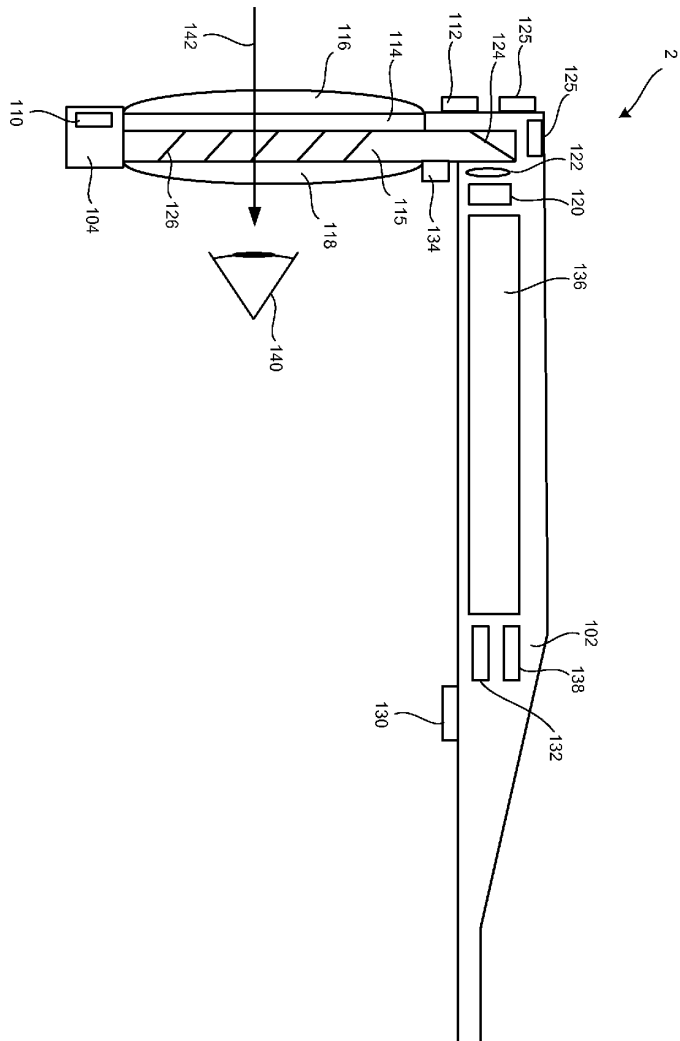
도면1



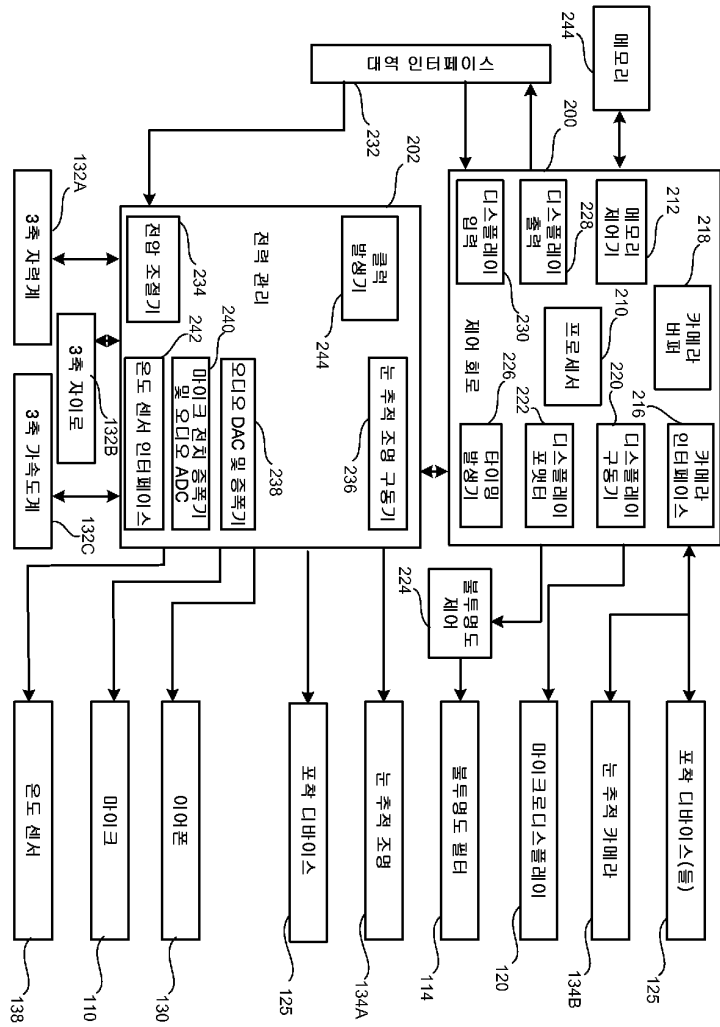
도면2



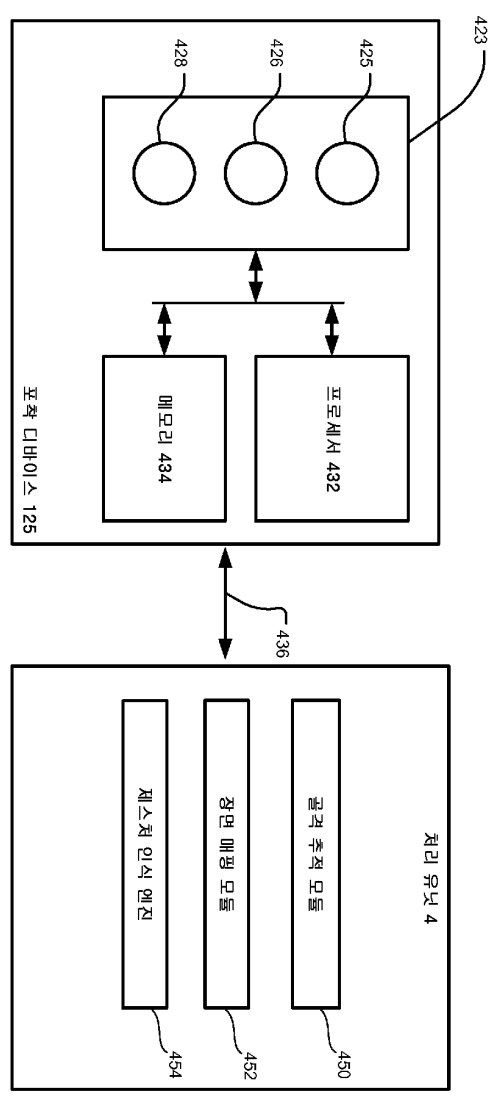
도면3



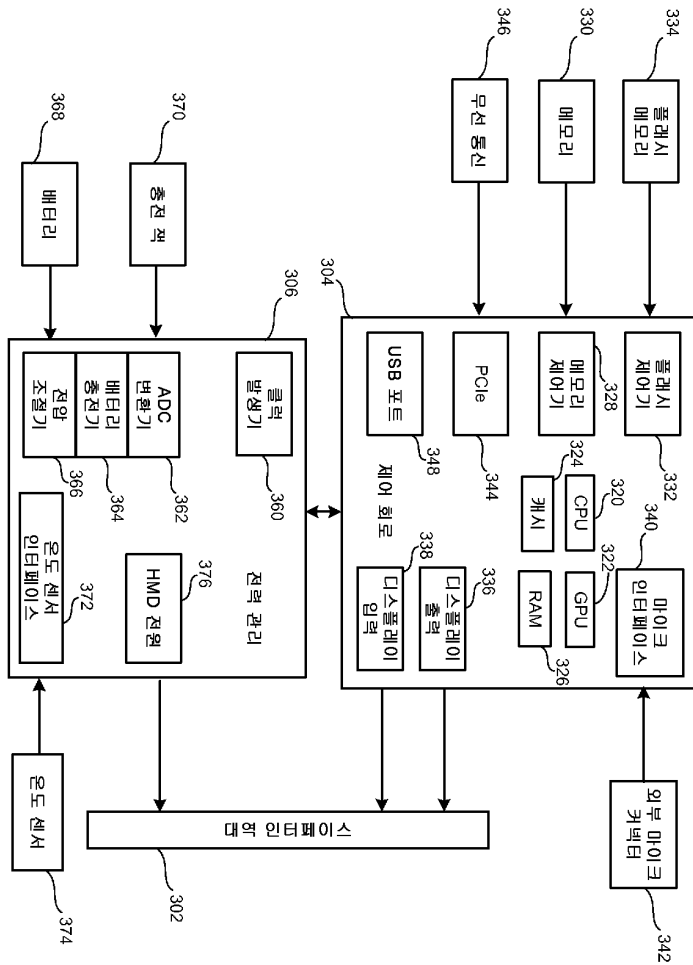
도면4



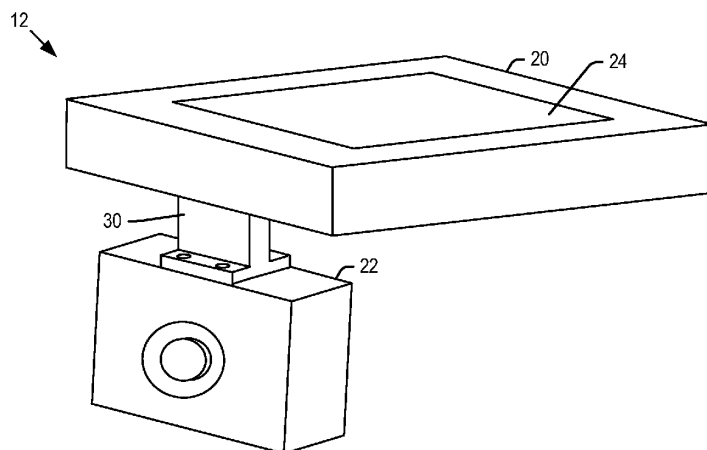
도면5



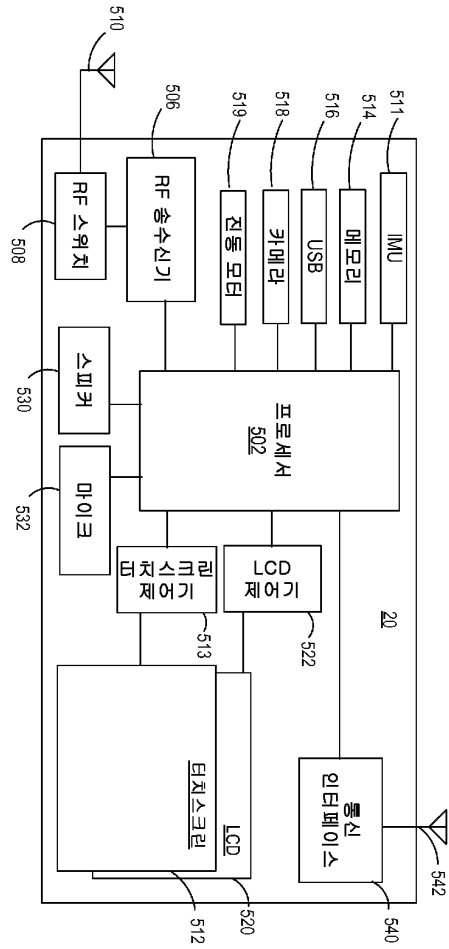
도면6



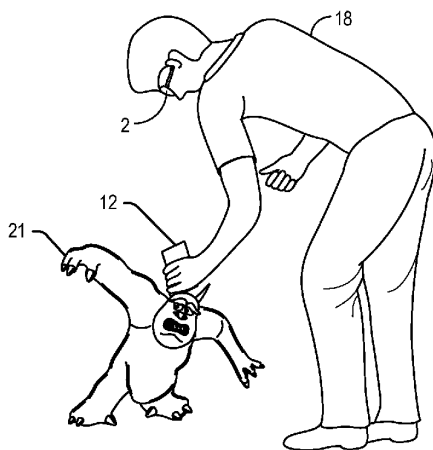
도면7



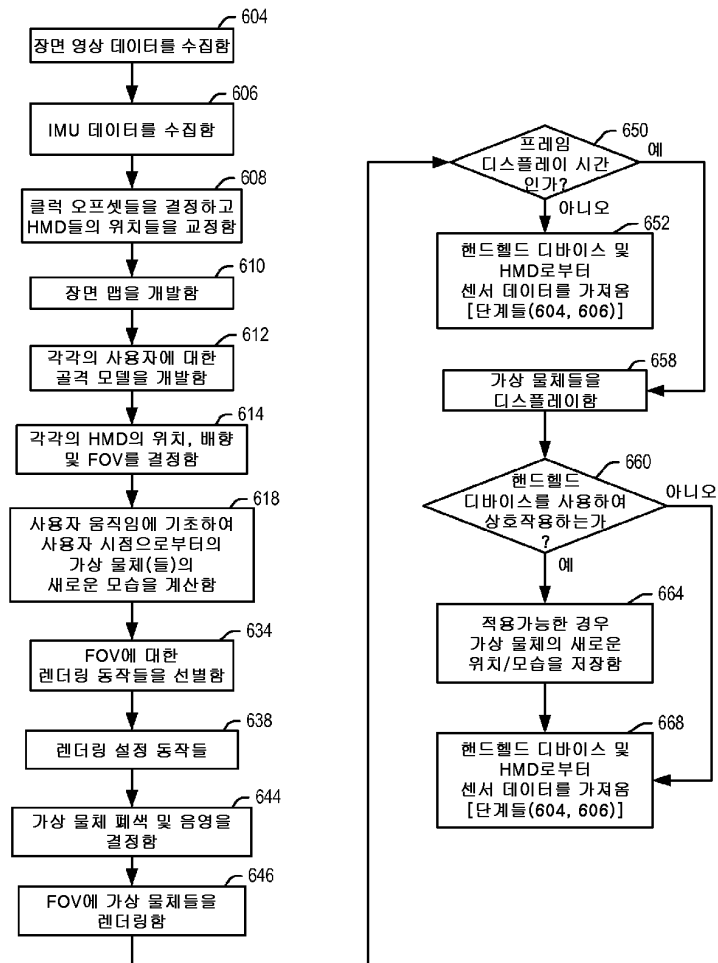
도면8



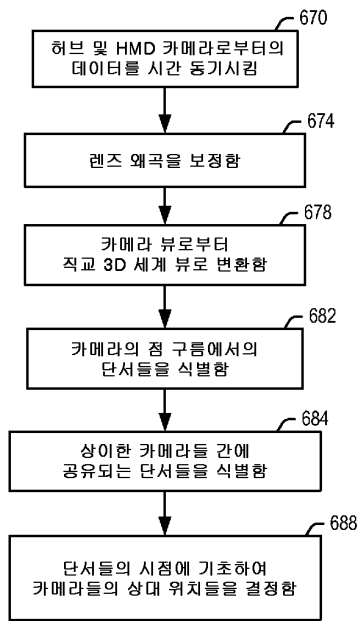
도면9



도면10



도면11



(단계 608)