



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2019년01월03일

(11) 등록번호 10-1935061

(24) 등록일자 2018년12월27일

(51) 국제특허분류(Int. Cl.)

G02B 27/01 (2006.01) G06T 17/00 (2006.01)

H04N 5/262 (2006.01)

(21) 출원번호 10-2013-7015226

(22) 출원일자(국제) 2011년12월05일

심사청구일자 2016년11월03일

(85) 번역문제출일자 2013년06월13일

(65) 공개번호 10-2013-0130740

(43) 공개일자 2013년12월02일

(86) 국제출원번호 PCT/US2011/063350

(87) 국제공개번호 WO 2012/082444

국제공개일자 2012년06월21일

(30) 우선권주장

12/970,695 2010년12월16일 미국(US)

(56) 선행기술조사문현

JP09204287 A\*

(뒷면에 계속)

(73) 특허권자

마이크로소프트 테크놀로지 라이센싱, 엘엘씨

미국 워싱턴주 (우편번호 : 98052) 레드몬드 원  
마이크로소프트 웨이

(72) 발명자

페레즈 케이티 스톤

미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마  
이크로소프트 코포레이션

바-지브 아비

미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마  
이크로소프트 코포레이션

카츠 사기

미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마  
이크로소프트 코포레이션

(74) 대리인

김태홍

심사관 : 이정호

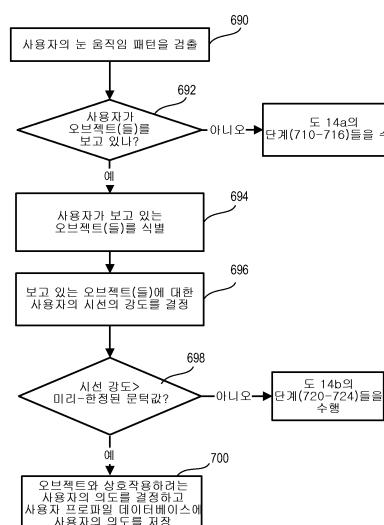
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 중강 현실 디스플레이를 위한 이해 및 의도-기반 컨텐트

### (57) 요 약

시스루 디스플레이 장치 또는 헤드 마운트 디스플레이 장치와 같은 니어 아이 디스플레이 장치를 사용할 때 사용자의 경험을 향상시키는 방법 및 시스템이 제공된다. 장면 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도가 결정된다. 사용자의 의도에 기초하여 사용자를 위한 최적 이미지가 생성된다. 최적 이미지는 시스루 디스플레이 장치를 통해 사용자에게 디스플레이된다. 최적 이미지는 장면 내에서 사용자가 상호작용하려고 의도하는 오브젝트의 외관을 시각적으로 강화하고 장면 내에서 사용자가 상호작용하려고 의도하지 않는 오브젝트의 외관을 약화시킨다. 최적 이미지는 사용자의 이해를 증가시키는 오브젝트의 외관을 시각적으로 강화할 수 있다.

### 대 표 도



(56) 선행기술조사문현

JP2002041234 A\*

JP2002163670 A\*

JP2003132068 A\*

JP2003508808 A\*

JP2010218405 A\*

\*는 심사관에 의하여 인용된 문현

---

## 명세서

### 청구범위

#### 청구항 1

증강 현실 시스템에서 사용자 의도에 기초하여 최적화된 이미지를 생성하는 방법에 있어서,

시스템(see-through) 디스플레이 디바이스의 시야(field of view) 내의 장면 내에서 복수의 현실 세계 오브젝트들을 결정하는 단계;

상기 장면 내에서 상기 시스템 디스플레이 디바이스를 통해 사용자의 시야를 결정하고, 상기 시야 내에서 상기 사용자의 포커스 영역(focal region)을 결정하는 단계;

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 결정하는 단계;

상기 사용자의 의도에 기초하여, 최적화된 이미지를 생성하는 단계; 및

상기 최적화된 이미지를 상기 사용자에게 디스플레이하는 단계

를 포함하고,

상기 최적화된 이미지는, 상기 사용자의 포커스 영역 내에서의 상기 복수의 현실 세계 오브젝트들 중 적어도 하나의 현실 세계 오브젝트의 시각적 강화, 및 상기 복수의 오브젝트들 중, 상기 사용자의 포커스 영역 외부에 있지만 상기 시스템 디스플레이 디바이스를 통한 상기 사용자의 시야 내에 있는 하나 이상의 오브젝트들의 동시적으로 축소된 외관을 포함한 것인 최적화된 이미지 생성 방법.

#### 청구항 2

제1항에 있어서,

상기 복수의 현실 세계 오브젝트들을 결정하는 단계는 상기 장면의 3차원 모델을 생성하는 단계를 포함하며, 상기 방법은,

상기 3차원 모델을, 상기 장면을 구성하는 하나 이상의 오브젝트들로 분할(segment)하고, 상기 장면 내에서 상기 하나 이상의 오브젝트들을 식별하는 단계

를 더 포함하는 최적화된 이미지 생성 방법.

#### 청구항 3

제2항에 있어서,

상기 동시적으로 축소된 외관은, 나머지 상기 복수의 현실 세계 오브젝트들을 포함한 상기 시야 내에서의 증가된 불투명성을 포함한 것인 최적화된 이미지 생성 방법.

#### 청구항 4

제1항에 있어서,

상기 사용자의 포커스 영역은 상기 사용자의 안구들의 파눔의 융합(Panum's fusional) 영역인 것인 최적화된 이미지 생성 방법.

#### 청구항 5

제1항에 있어서,

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 결정하는 단계는,

상기 사용자의 포커스 영역 내에서의 상기 사용자의 안구 움직임 패턴들을 검출하는 단계;

상기 사용자의 포커스 영역 내에서 상기 사용자가 바라보고 있는 하나 이상의 오브젝트들에 대한 상기 사용자의 시선의 세기를 결정하는 단계; 및

상기 안구 움직임 패턴들 및 상기 사용자의 시선의 세기에 기초하여, 상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 추론하는 단계

를 더 포함한 것인 최적화된 이미지 생성 방법.

## 청구항 6

제5항에 있어서,

상기 추론하는 단계는 상호작용하려는 상기 사용자의 의도에 대해 결정력을 가지지는 않으며,

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 결정하는 단계는,

상기 사용자와 관련된 사용자 특유의 정보에 액세스하는 단계;

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들을 상기 사용자 특유의 정보와 상관시키는 단계; 및

상기 상관에 기초하여, 상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 자동으로 추론하는 단계

를 더 포함한 것인 최적화된 이미지 생성 방법.

## 청구항 7

제6항에 있어서,

상기 자동적으로 추론하는 단계는 상호작용하려는 상기 사용자의 의도에 대해 결정력을 가지지는 않으며,

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 결정하는 단계는,

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 의도를 지정할 것을 상기 사용자에게 프롬프트하는 단계;

적어도 하나의 사용자 물리적 행동을 통해 상기 사용자의 의도를 수신하는 단계; 및

상기 사용자의 물리적 행동에 기초하여 상기 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 결정하는 단계

를 더 포함하는 것인 최적화된 이미지 생성 방법.

## 청구항 8

제1항에 있어서,

상기 최적화된 이미지를 생성하는 단계는,

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 관련된 증강 콘텐츠 또는 오디오 정보 중 적어도 하나를 검색하는(retrieve) 단계; 및

상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들 위 또는 옆에 상기 증강 콘텐츠 또는 상기 오디오 정보를 투사하는 단계

를 더 포함하는 최적화된 이미지 생성 방법.

## 청구항 9

제8항에 있어서,

상기 하나 이상의 오브젝트들의 외관의 시작적 강화는, 상기 하나 이상의 오브젝트들의 에지(edge)들을 하이라

이트하는 것, 또는 상기 하나 이상의 오브젝트들이 위치해 있는 영역 내에 가상 박스, 원 또는 가상 지시자를 디스플레이하는 것 중 적어도 하나를 더 포함한 것인 최적화된 이미지 생성 방법.

### 청구항 10

제1항에 있어서,

상기 복수의 현실 세계 오브젝트들을 결정하는 단계는,

오브젝트로부터 상기 시스루 디스플레이 디바이스로 방출된 신호를 검출하는 단계; 및

거리 검출 기술을 이용하여 상기 시스루 디스플레이 디바이스로부터의 가능한 오브젝트들의 거리를 결정하는 단계

중 적어도 하나를 포함한 것인 최적화된 이미지 생성 방법.

### 청구항 11

제8항에 있어서,

상기 최적화된 이미지 내에서 디스플레이되는 증강 콘텐츠와 상호작용하려는 상기 사용자의 의도를 결정한 것에 기초하여, 추가적인 증강 콘텐츠를 사용자에게 디스플레이하는 단계

를 더 포함하고,

상기 추가적인 증강 콘텐츠를 디스플레이하는 단계는,

상기 최적화된 이미지와 연관된 메뉴를 디스플레이하는 단계;

상기 메뉴에 의해 디스플레이된 하나 이상의 선택사항들과 상호작용하려는 상기 사용자의 의도를 결정하는 단계;

상기 사용자의 의도에 기초하여, 상기 메뉴 내 선택사항의 선택을 수신하는 단계; 및

상기 선택사항과 관련된 추가적인 증강 콘텐츠를 검색하는 단계; 및

상기 최적화된 이미지 내의 상기 추가적인 증강 콘텐츠를 상기 사용자에게 투사하는 단계

를 더 포함한 것인 최적화된 이미지 생성 방법.

### 청구항 12

시스루 디스플레이 및 캡처 디바이스를 포함한 장치에 있어서, 상기 장치는 하나 이상의 프로세서로 판독가능한 저장 디바이스들을 포함하고, 상기 하나 이상의 프로세서로 판독가능한 저장 디바이스들 상에는 프로세서로 판독가능한 코드가 포함되어 있고, 상기 프로세서로 판독가능한 코드는 방법을 수행하도록 하나 이상의 프로세서들을 프로그래밍하기 위한 것이며, 상기 방법은,

장면 내 복수의 관심대상 현실 세계 오브젝트들을 결정하기 위해 상기 캡처 디바이스를 이용하여 상기 장면을 스캐닝하는 단계;

상기 장면 내에서 상기 시스루 디스플레이 디바이스의 시야 및, 상기 시야 내의 상기 현실 세계 오브젝트들을 결정하는 단계;

상기 시야 내의 사용자의 포커스 영역을 결정하는 단계;

상기 포커스 영역 내에 있는 상기 복수의 현실 세계 오브젝트들과, 상기 복수의 현실 세계 오브젝트들 중 환경 내에서 상기 포커스 영역 외부에 있는 나머지 복수의 현실 세계 오브젝트들, 중 적어도 하나를 식별하는 단계;

상기 시야 내에서 최적화된 이미지 - 상기 최적화된 이미지는, 상기 사용자의 포커스 영역 내에서의 상기 복수의 현실 세계 오브젝트들 중 적어도 하나의 현실 세계 오브젝트의 시작적 강화, 및 상기 사용자의 포커스 영역 외부에 있는 임의의 상기 나머지 복수의 현실 세계 오브젝트들의 동시적으로 축소된 외관을 포함함 - 를 생성하는 단계; 및

상기 시스루 디스플레이 디바이스의 시야 내에서 상기 최적화된 이미지를 디스플레이하는 단계

를 포함한 것인 장치.

### 청구항 13

제12항에 있어서,

상기 포커스 영역 내에 있는 상기 복수의 현실 세계 오브젝트들을 식별하는 단계는 상기 환경에 관한 정보를 실시간으로 수집하는 단계를 포함하며, 상기 포커스 영역 내에 있는 상기 복수의 현실 세계 오브젝트들은 상기 환경의 사용자 이해를 강화시키는 것인 장치.

### 청구항 14

제12항에 있어서,

상기 포커스 영역 내에 있는 상기 복수의 현실 세계 오브젝트들을 식별하는 단계는 상기 사용자에 관한 사용자 특유의 정보에 기초한 것인 장치.

### 청구항 15

제12항에 있어서,

상기 최적화된 이미지를 생성하는 단계는, 상기 환경 내의 이해를 강화하는 하나 이상의 오브젝트들의 외관을 시각적으로 강화시키는 단계와, 상기 나머지 복수의 현실 세계 오브젝트들을 포함한 상기 시야 내에서의 불투명성을 증가시키는 단계를 더 포함한 것인 장치.

### 청구항 16

제15항에 있어서,

상기 하나 이상의 오브젝트들의 외관을 시각적으로 강화시키는 단계는, 상기 하나 이상의 오브젝트들의 에지들을 하이라이트하는 단계, 또는 상기 하나 이상의 오브젝트들이 위치해 있는 영역 내에 가상 박스, 원 또는 가상 지시자를 디스플레이하는 단계 중 적어도 하나를 더 포함한 것인 장치.

### 청구항 17

제16항에 있어서,

상기 방법은,

상기 최적화된 이미지 내 복수의 현실 세계 오브젝트들 중의 적어도 하나의 현실 세계 오브젝트와 상호작용하려는 사용자 의도를 결정하는 단계

를 더 포함하고,

상기 사용자 의도를 결정하는 단계는,

상기 시각적 강화에 대한 상기 사용자의 시선의 세기를 결정하는 단계;

상기 사용자와 관련된 사용자 특유의 정보에 액세스하는 단계; 및

상기 시선의 세기 및 상기 사용자 특유의 정보 중 적어도 하나에 기초하여, 상기 시각적 강화와 상호작용하려는 상기 사용자의 의도를 결정하는 단계

를 포함한 것인 장치.

### 청구항 18

증강 현실 시스템에 있어서,

시야를 갖는 시스루 디스플레이 디바이스;

상기 시스루 디스플레이 디바이스의 시야 내의 하나 이상의 오브젝트들과 상호작용하려는 사용자 의도를 결정한 것에 기초하여, 최적화된 이미지를 사용자에게 디스플레이하도록 적응된, 상기 시스루 디스플레이 디바이스에 부착된 마이크로 디스플레이 조립체;

상기 마이크로 디스플레이 조립체에 통신 가능하게 결합된 프로세서; 및  
상기 프로세서에 의해 액세스가능하며 소프트웨어 및 데이터를 저장하는 메모리  
를 포함하고,

소프트웨어의 제어 하에서,

상기 프로세서는 상기 시야 내의 복수의 현실 세계 오브젝트들을 결정하고,

상기 프로세서는, 제일먼저, 상기 시스루 디스플레이 디바이스를 통해 상기 현실 세계 오브젝트에 대한 상기 사용자의 시선의 세기를 결정하고, 상기 세기가 사용자 의도를 결정하지 않는 경우, 상기 사용자와 관련된 사용자 특유의 정보에 액세스하며, 상기 시야 내 상기 현실 세계 오브젝트들과 상기 사용자 특유의 정보 간의 상관에 기초하여 상호작용하려는 상기 사용자의 의도를 결정함으로써, 상기 시야 내에서 상기 사용자의 포커스 영역 내의 상기 복수의 현실 세계 오브젝트들 중 하나와 상호작용하려는 사용자 의도를 결정하고,

상기 프로세서는 상기 시야 내의 상기 복수의 현실 세계 오브젝트들 중 하나의 현실 세계 오브젝트의 최적화된 이미지를 생성하며,

상기 최적화된 이미지는, 상기 포커스 영역 내의 상기 복수의 현실 세계 오브젝트들 중 하나의 현실 세계 오브젝트의 시각적 강화, 및 상기 사용자의 포커스 영역 외부에 있지만 상기 시야 내에 있는 상기 복수의 현실 세계 오브젝트들 중의 임의의 현실 세계 오브젝트들의 동시적으로 축소된 외관을 포함한 것인 증강 현실 시스템.

## 청구항 19

제18항에 있어서,

상기 프로세서는, 상기 사용자의 안구 움직임 패턴들, 상기 사용자의 시선의 세기 및 상기 사용자와 관련된 사용자 특유의 정보 중 적어도 하나에 기초하여, 상기 사용자의 포커스 영역 내의 하나 이상의 오브젝트들과 상호작용하려는 상기 사용자의 의도를 결정하는 것인 증강 현실 시스템.

## 청구항 20

제18항에 있어서,

상기 마이크로 디스플레이 조립체는, 상기 최적화된 이미지에서, 상기 사용자가 상호작용하려는 하나 이상의 오브젝트들을 하이라이트시키고, 나머지 상기 복수의 현실 세계 오브젝트들을 포함한 상기 시야 내에서 불투명성을 증가시키는 것인, 증강 현실 시스템.

## 발명의 설명

### 기술 분야

[0001] 증강 현실(augmented reality)이란 현실 세계의 물리적인 환경이나 공간과 혼합되는 가상 영상을 가능하게 하는 기술이다. 통상적으로, 니어 아이 디스플레이(near eye display)는 사용자가 가상 및 현실의 물체에 대한 혼합된 영상을 볼 수 있도록 하기 위해 착용된다. 일반적으로 니어 아이 디스플레이는 공간 내에 가상 영상을 포커싱하기 위하여 광학과 입체 영상을 이용한다.

[0002] 특정 상황에서, 니어 아이 디스플레이 장치를 통해 사용자에게 디스플레이되는 가상 영상은 매우 자세하게 묘사된 그래픽을 포함하는 가상 이미지나 오브젝트를 포함할 수 있다. 그렇지만, 일반적으로 모바일인 니어-아이 디스플레이 장치는 컴퓨터 자원에 있어 제한적인 경향이 있어 사용자에게 정확하게 가상 영상을 표현하지 못할 수 있다. 덧붙여서, 니어-아이 디스플레이 장치를 착용한 사용자에게는 통상적으로 사용자가 딱히 보고 싶어하는 것이 아닌 많은 양의 정보가 표현된다.

## 발명의 내용

[0003] 본 명세서에서, 사용자에게 디스플레이되는 시각화된 정보를 최적화함으로써 시스루 디스플레이(see-through display) 장치 또는 헤드 마운트 디스플레이(head mounted display) 장치와 같은 니어 아이 디스플레이 장치를 사용할 때 사용자의 경험을 향상시키는 방법 및 시스템이 개시된다. 시각화된 정보는, 사용자에게 디스플레이된

시각화된 정보에 우선순위를 매김으로써 최적화된다. 하나의 접근방식에 있어서, 시각화된 정보는 사용자의 환경 내의 하나 이상의 오브젝트와 상호작용하는 사용자의 추측된 또는 표현된 의도를 결정함으로써 우선순위가 매겨진다. 일 실시예에서, 최적 이미지(optimized image)는 우선순위가 매겨진 시각 정보에 기초하여 생성된다. 최적 이미지는 사용자가 사용자의 환경과 상호작용하려고 의도하는 오브젝트의 외관을 시각적으로 강화시키며 및/또는 사용자가 사용자의 환경과 상호작용하지 않으려고 의도하는 오브젝트의 외관을 약화시킨다. 하나의 접근방식에 있어서, 사용자의 의도는 사용자 눈의 움직임, 및 사용자의 환경 내의 오브젝트에 대한 사용자의 시선을 분석함으로써 결정된다. 다른 접근방식에 있어서, 사용자의 의도는 사용자와 관련된 사용자 특정 정보에 기초하여 자동으로 결정된다. 일 실시예에서, 최적 이미지는, 시각화된 정보를 처리하고 사용자에게 표현하는 동안 헤드 마운트 디스플레이 장치에 의해 요구되는 컴퓨터 자원을 감소시킨다.

[0004] 개시된 기술의 다른 실시예에서, 사용자의 환경 내에서 시각화된 정보에 대한 사용자의 이해는 사용자의 환경 내에 있는 하나 이상의 오브젝트를 자동으로 강화시킴으로써 증가된다. 일 실시예에서, 사용자의 이해를 증가시키는 하나 이상의 오브젝트의 외관을 시각적으로 강화시킨 최적 이미지가 생성된다. 다른 실시예에서, 최적 이미지 내의 하나 이상의 시각적으로 강화된 오브젝트에 대한 사용자의 의도가 결정된다.

[0005] 일 실시예에서, 중강 현실 시스템에서 사용자 의도에 기초하여 최적 이미지를 생성하기 위한 방법이 제공된다. 하나의 접근방식에서, 상기 방법은 사용자가 시스루 디스플레이 장치를 사용하여 보고 있는 장면에서 흥미있는 오브젝트 또는 사람을 결정하는 과정을 포함한다. 사용자의 뷔 필드(field of view)가 결정된다. 뷔 필드는 사용자가 바라보는 공간이나 환경의 일부이다. 뷔 필드 내에서 사용자의 포커스 영역(focal region)이 결정된다. 사용자의 포커스 영역 내에 있는 하나 이상의 오브젝트나 사람과 상호작용하려는 사용자의 의도가 결정된다. 일 실시예에서, 사용자의 의도에 기초하여 사용자를 위해 최적 이미지가 생성된다. 최적 이미지는 시스루 디스플레이 장치를 통해 사용자에게 디스플레이된다.

[0006] 이 발명의 내용은 아래의 발명을 실시하기 위한 구체적인 내용에서 더 상세히 설명되는 개념들을 개략적인 형태로 소개하기 위하여 제공된다. 이 발명의 내용은 청구된 발명 주제들에 대해 중요한 특징이나 본질적인 특징들을 식별하도록 의도되지 않으며, 또한 청구된 발명 주제들의 범위를 결정하기 위한 보조수단으로서 사용되는 것도 역시 의도되지 않는다. 더 나아가, 청구된 발명 주제들은 이 개시의 어떠한 부분에서 언급된 단점들 중 임의의 단점이나 모든 단점들을 해결하는 실시예들로 제한되지 않는다.

### 도면의 간단한 설명

[0007] 도 1은 사용자 의도에 기초하여 최적화된 컨텐트를 생성하기 위한 시스템의 일 실시예의 예시적인 컴포넌트들을 도시하는 블록도이다.

도 2는 헤드 마운트 디스플레이 유닛의 일 실시예의 일부의 평면도이다.

도 2a 내지 도 2e는 불투명 필러의 동작을 도시하는 예시적인 이미지들을 제공한다.

도 3은 헤드 마운트 디스플레이 유닛의 컴포넌트들의 일 실시예의 블록도이다.

도 4는 헤드 마운트 디스플레이 유닛에 연관된 프로세싱 유닛의 컴포넌트들의 일 실시예의 블록도이다.

도 5는 헤드 마운트 디스플레이 유닛과 함께 사용된 허브 컴퓨팅 시스템의 컴포넌트들의 일 실시예의 블록도이다.

도 6은 본 명세서에서 기술된 허브 컴퓨팅 시스템을 구현하기 위하여 사용될 수 있는 컴퓨팅 시스템의 일 실시예의 블록도이다.

도 7은 사용자 의도에 기초하여 최적화된 컨텐트를 생성하기 위한 멀티-유저 시스템을 도시하는 블록도이다.

도 8은 사용자 의도에 기초하여 최적화된 컨텐트를 생성하기 위한 프로세스의 일 실시예를 기술한다.

도 9는 사용자의 공간의 모델을 만들어내기 위한 프로세스의 일 실시예를 기술한다.

도 10은 공간 모델을 오브젝트들로 분할하는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 11a는 오브젝트를 식별하는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 11b는 이동하는 오브젝트에 반응하여 도 8의 프로세스에 의해 만들어진 모델을 업데이트하는 프로세스의 일

실시예를 기술하는 흐름도이다.

도 12는 사용자의 뷰 필드를 추적하고 사용자의 포커스 영역을 결정하는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 13a는 도 12의 프로세스에서 사용되는 추적 정보를 제공하기 위하여 허브 컴퓨팅 시스템에 의해 수행되는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 13b는 그 결과가 도 12의 프로세스에 의해 사용되는, 눈을 추적하는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 14는 사용자의 포커스 영역 내에 있는 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정하는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 14a는 사용자의 포커스 영역 내에 있는 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정하는 프로세스의 다른 실시예를 기술하는 흐름도이다.

도 14b는 하나 이상의 오브젝트와 상호작용하는 의도를 특정하도록 사용자에게 프롬프팅함으로써 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정하는 프로세스의 다른 실시예를 기술하는 흐름도이다.

도 15a는 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정하는 과정에 기초하여 최적 이미지를 생성하고 사용자에게 디스플레이하는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 15b는 최적 이미지 내에서 디스플레이된 증강 컨텐트와 상호작용하려는 사용자의 의도를 결정하는 과정에 기초하여, 사용자에게 추가적인 증강 컨텐트를 디스플레이하는 프로세스의 일 실시예를 기술하는 흐름도이다.

도 15c는 사용자의 환경 내에서 시각화된 정보에 대한 사용자의 이해가 강화되도록 하는 프로세스를 기술한다.

도 16a 내지 도 16c는 사용자의 환경 내 하나 이상의 오브젝트와의 사용자의 상호작용 및 사용자 상호작용에 기초하여 최적 이미지의 생성의 일 실시예를 도시한다.

도 17a 내지 도 17d는 사용자의 환경 내 하나 이상의 오브젝트와의 사용자의 상호작용 및 사용자 상호작용에 기초하여 최적 이미지의 생성의 다른 실시예를 도시한다.

도 18a 내지 도 18c는 사용자의 환경 내 하나 이상의 오브젝트와의 사용자의 상호작용 및 사용자 상호작용에 기초하여 최적 이미지의 생성의 다른 실시예를 도시한다.

### 발명을 실시하기 위한 구체적인 내용

[0008]

니어 아이 디스플레이를 사용할 때 사용자의 경험이 향상되는 기술이 개시된다. 사용자는 헤드 마운트 디스플레이 장치와 같은 니어 아이 디스플레이를 통해 장면을 바라본다. 사용자가 관찰할 수 있는 환경 또는 공간의 일부인 사용자의 뷰 필드가 결정된다. 뷰 필드 내의 흥미 있을 만한 오브젝트들도 역시 결정된다. 뷰 필드는 사용자의 포커스 영역, 즉 사용자가 뷰 필드 내에서 실제로 포커싱하고 있는 것을 포함할 수 있다. 사용자의 포커스 영역은 뷰 필드 내에서 사용자의 눈의 위치를 추적함으로써 결정된다. 그런 다음 사용자의 포커스 영역 내의 하나 이상의 오브젝트와 상호작용하는 사용자의 추측된 또는 표현된 의도가 결정된다. 한 접근방식에서, 사용자의 포커스 영역 내 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도는 사용자의 포커스 영역 내에서 사용자 눈의 이동 패턴을 검출하고 사용자의 포커스 영역 내에서 사용자가 보고 있는 하나 이상의 오브젝트에 대한 사용자의 시선의 강도를 결정함으로써 결정된다. 다른 접근방식에서, 사용자의 포커스 영역 내 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도는 사용자와 관련된 사용자-특정 정보에 액세스함으로써 자동으로 결정된다. 다른 접근방식에서, 사용자의 포커스 영역 내 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도는, 음성 입력, 키보드 입력, 터치 감응 장치 또는 제스처와 같은 사용자의 물리적인 행위를 통해 하나 이상의 오브젝트와 상호작용하는 의도를 사용자가 특정하도록 프롬프팅(prompting)함으로써 결정된다. 또 다른 실시예에서, 하나 이상의 오브젝트와 상호작용하는 사용자의 추측된 또는 표현된 의도는 환경 및 시간, 위치, 및 외부 입력과 같은 그 외 다른 인자들에 의하여 결정된다. 사용자의 필요 또는 의도에 기초하여 최적 이미지가 생성된다. 최적 이미지는 헤드 마운트 디스플레이 장치를 통해 사용자에게 디스플레이된다. 최적 이미지는 사용자의 포커스 영역 내의 강화된 외관을 가진 오브젝트, 사용자의 뷰 필드 내에 있지만 사용자의 포커스 영역 외부에 있는 약화된 외관을 가진 오브젝트, 및 사용자의 포커스 영역 내에서 사용자가 상호작용하려고 의도하는 오브젝트와 관련된 가상 이미지 또는 가상 오브젝트와 같은 증강 컨텐트 중 하나 이상을 포함할 수 있다. 이러한 향상의 임

의의 하나 이상은 조합하여 사용될 수 있다.

[0009] 도 1은 사용자 의도에 기초하여 최적 이미지를 생성하기 위한 시스템(10)의 일 실시예의 예시적인 컴포넌트들을 도시하는 블록도이다. 시스템(10)은 유선(6)으로 프로세싱 유닛(4)과 통신하는 니어-아이, 헤드 마운트 디스플레이 장치(2)와 같은 시스루 디스플레이 장치를 포함한다.

[0010] 그 외 다른 실시예들에서, 헤드 마운트 디스플레이 장치(2)는 무선 통신을 통해 프로세싱 유닛(4)과 통신한다. 헤드 마운트 디스플레이 장치(2)는, 일 실시예에서 안경의 형태를 하는데, 사용자에 의하여 착용됨으로써 사용자가 디스플레이를 통하여 볼 수 있게 하고 이에 의하여 사용자 앞에 있는 공간을 실제적이고 직접적으로 볼 수 있게 한다. "실제적이고 직접적으로 본다"는 표현의 사용은, 만들어진 오브젝트의 이미지 표현을 보는 것이 아니라, 인간의 눈으로 직접 실제 세계의 오브젝트를 보는 능력을 나타낸다. 예를 들어, 방에서 안경을 통해 바라보는 것은 사용자로 하여금 방을 실제적이고 직접적으로 볼 수 있게 허용하는데, 다른 한편으로 텔레비전 사에서 방에 대한 비디오를 보는 것은 방을 실제적이고 직접적으로 보는 것이 아니다. 아래에 헤드 마운트 디스플레이 장치(2)에 대한 더 상세한 설명이 제공된다.

[0011] 일 실시예에서, 프로세싱 유닛(14)은 사용자에 의해 착용되고 헤드 마운트 디스플레이 장치(2)를 동작하기 위하여 사용되는 컴퓨팅 능력의 대부분을 포함한다. 프로세싱 유닛(4)은 하나 이상의 허브 컴퓨팅 시스템(hub computing system)(12)과 무선(예컨대, WiFi, 블루투스, 적외선, 또는 그 외 다른 무선 통신 수단)으로 통신한다.

[0012] 허브 컴퓨팅 시스템(12)은 컴퓨터, 게임 시스템이나 콘솔 등일 수 있다. 예시적인 일 실시예에 따라, 허브 컴퓨팅 시스템(12)은 허브 컴퓨팅 시스템(12)이 게임 애플리케이션, 비-게임 애플리케이션 등과 같은 애플리케이션을 실행하기 위해 사용될 수 있도록 하기 위한 하드웨어 컴포넌트들 및/또는 소프트웨어 컴포넌트들을 포함할 수 있다. 일 실시예에서, 허브 컴퓨팅 시스템(12)은 본 명세서에서 기술된 프로세스들을 수행하도록 프로세서 판독가능 저장 장치 상에 저장된 인스트럭션(instructions)을 실행할 수 있는 표준화된 프로세서, 특수 목적의 프로세서, 마이크로프로세서 등과 같은 프로세서를 포함할 수 있다.

[0013] 허브 컴퓨팅 시스템(12)은 캡처 장치(20A 및 20B)와 같은 하나 이상의 캡처 장치를 더 포함한다. 그 외 다른 실시예들에서, 2개보다 많거나 적은 수의 캡처 장치가 사용될 수 있다. 예시적인 일 실시예에서, 캡처 장치(20A 및 20B)는 방의 서로 다른 부분을 캡처할 수 있도록 서로 다른 방향을 향하게 있다. 2개의 캡처 장치의 뷰 필드는 허브 컴퓨팅 시스템(12)이 캡처 장치들의 뷰 필드가 서로에 대해 어떻게 관계되어 있는지를 이해할 수 있도록 하기 위하여 약간 중첩되는 것이 유리할 수 있다. 이런 방식으로, 전체 방(또는 그 외 다른 공간)를 보기 위하여 다수의 캡처 장치가 사용될 수 있다. 대안적으로, 만약 뷰 필드들이 캡처 장치에 의하여 시간에 걸쳐 보여질 수 있도록 캡처 장치가 동작시 상하좌우로 움직일(panned) 수 있다면 하나의 캡처 장치가 사용될 수 있다. 비-중첩 캡처 장치들도 또한, 사용자에 의해 착용된 헤드 마운트 디스플레이의 위치와 같은 다른 알려진 포인트에 관련하여 뷰 필드를 결정할 수 있다.

[0014] 캡처 장치(20A 및 20B)는, 예컨대, 애플리케이션 내에서 하나 이상의 컨트롤 또는 행위를 수행하기 위해 및/또는 아바타 또는 온-스크린 캐릭터를 애니메이팅하기 위해 주변 공간의 구조뿐만 아니라 하나 이상의 사용자에 의해 수행되는 제스처 및/또는 움직임이 캡처되고, 분석되고, 및 추적될 수 있는 방식으로, 하나 이상의 사용자 및 주변 공간을 시각적으로 감시하는 카메라일 수 있다.

[0015] 허브 컴퓨팅 시스템(12)은, 게임 또는 애플리케이션 비주얼을 제공할 수 있는 텔레비전, 모니터, 고해상도 텔레비전(HDTV) 등과 같은 오디오비주얼 장치(16)에 연결될 수 있다. 예를 들어, 허브 컴퓨팅 시스템(12)은, 게임 애플리케이션, 비-게임 애플리케이션 등과 연관되는 오디오비주얼 신호를 제공할 수 있는 그래픽 카드와 같은 비디오 어댑터 및/또는 사운드 카드와 같은 오디오 어댑터를 포함할 수 있다. 오디오비주얼 장치(16)는 허브 컴퓨팅 시스템(12)으로부터 오디오비주얼 신호를 수신할 수 있고 그런 다음 오디오비주얼 신호와 연관된 게임 또는 애플리케이션 비주얼 및/또는 오디오를 출력할 수 있다. 일 실시예에 따라, 오디오비주얼 장치(16)는 예컨대, S-Video 케이블, 동축 케이블, HDMI 케이블, DVI 케이블, VGA 케이블, 컴포넌트 비디오 케이블, RCA 케이블 등을 통해 허브 컴퓨팅 시스템(12)에 연결될 수 있다. 일 예에서, 오디오비주얼 장치(16)는 내장 스피커를 포함한다. 그 외 다른 실시예들에서, 오디오비주얼 장치(16), 별도의 스테레오 또는 허브 컴퓨팅 시스템(12)은 외부 스피커(22)에 연결된다.

[0016] 허브 컴퓨팅 장치(10)는, 캡처 장치(20A 및 20B)를 구비하는데, 인간(및 그 외 다른 타입의) 타겟을 인식하고, 분석하며, 및/또는 추적하기 위하여 사용될 수 있다. 예를 들어, 헤드 마운트 디스플레이 장치(2)를 착용한 사

용자는, 사용자의 제스처 및/또는 움직임이 아바타 또는 온-스크린 캐릭터를 애니메이트하기 위해 캡처될 수 있거나 및/또는 허브 컴퓨팅 시스템(12)에 의해 실행되고 있는 애플리케이션에 영향을 미치기 위하여 사용될 수 있는 컨트롤로서 해설될 수 있도록 하는 방식으로, 캡처 장치(20A 및 20B)를 사용하여 추적될 수 있다.

[0017] 개시되었고 또한 아래에서 상세하게 논의될 기술의 일 실시예에서, 시스템(10)은 사용자의 환경 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정하는 것에 기초하여 사용자에 대하여 최적 이미지를 생성한다. 최적 이미지는 사용자가 상호작용하려고 의도하는 오브젝트의 강화된 외관, 사용자가 상호작용하려고 의도하지 않는 오브젝트의 약화된 외관, 또는 양자 모두를 포함할 수 있다. 최적 이미지는 헤드 마운트 디스플레이 장치(2)를 통해 사용자에게 디스플레이된다.

[0018] 도 2는 안경다리(102) 및 노즈 브릿지(nose bridge)(104)를 포함하여 프레임 일부를 포함하는 헤드 마운트 디스플레이 장치(2)의 일부의 평면도를 도시한다. 헤드 마운트 디스플레이 장치(20)의 우측만이 도시되어 있다. 아래에서 기술되는 바와 같이, 음향을 기록하고 프로세싱 유닛(4)으로 오디오 데이터를 전송하기 위한 마이크로폰(110)이 노즈 브릿지(104)에 내장된다. 헤드 마운트 디스플레이 장치(2)의 전면에 비디오 및 정지 이미지를 찍을 수 있는 방 대면 비디오 카메라(113)가 있다. 이들 이미지는 아래에서 기술되는 바와 같이 프로세싱 유닛(4)으로 전송된다.

[0019] 헤드 마운트 디스플레이 장치(2)의 프레임 일부는 디스플레이(하나 이상의 렌즈를 포함하는)을 둘러쌀 것이다. 헤드 마운트 디스플레이 장치(2)의 컴포넌트들을 보여주기 위하여, 디스플레이를 둘러싸는 프레임의 일부는 도시되어 있지 않다. 일 실시예에서, 디스플레이는 광 안내 광학 요소(112), 불투명 필터(114), 시스루 렌즈(116) 및 시스루 렌즈(118)를 포함한다. 일 실시예에서, 불투명 필터(114)는 시스루 렌즈(116)의 뒤에서 그것에 정렬되어 있으며, 광 안내 광학 요소(112)는 불투명 필터(114)의 뒤에서 그것에 정렬되어 있고, 또한 시스루 렌즈(118)는 광 안내 광학 요소(112)의 뒤에서 그것에 정렬되어 있다. 시스루 렌즈(116 및 118)는 안경에서 사용되는 표준 렌즈이며 임의의 처방에 따라(무 처방을 포함하여) 만들어질 수 있다. 일 실시예에서, 시스루 렌즈(116 및 118)는 가변 처방 렌즈에 의해 대체될 수 있다. 일부 실시예들에서, 헤드 마운트 디스플레이 장치(2)는 오직 하나의 시스루 렌즈를 포함할 것이며 또는 아무런 시스루 렌즈도 포함하지 않을 것이다. 다른 대안에 있어서, 처방 렌즈는 광 안내 광학 요소(112) 내부에 배치될 수 있다. 불투명 필터(114)는 가상 영상의 콘트라스트를 향상시키기 위하여 자연광(픽셀 기반으로 또는 균일하게 중 어느 한 방식으로)을 필터링하여 제거한다. 광 안내 광학 요소(112)는 인공적인 광을 눈으로 전달한다. 불투명 필터(114)와 광 안내 광학 요소(112)에 대한 더 상세한 설명은 아래에 제공된다. 대안적인 실시예들에서, 불투명 필터(114)는 사용되지 않을 수 있다.

[0020] 이미지 소스가 안경다리(102)에 탑재되거나 내장되는데, 이는 (일 실시예에서) 가상 이미지를 투사하기 위한 마이크로 디스플레이 조립체(120) 및 마이크로 디스플레이(120)로부터 광 안내 광학 요소(112) 내로 이미지를 향하게 하기 위한 렌즈(122)를 포함한다. 일 실시예에서, 렌즈(122)는 콜러메이팅 렌즈(collimating lens)이다.

[0021] 컨트롤 회로(136)는 헤드 마운트 디스플레이 장치(2)의 그 외 다른 컴포넌트들을 지원하는 다양한 전자부품을 제공한다. 컨트롤 회로(136)에 대해서는 도 3을 참조하여 아래에서 더 상세히 설명된다. 이어폰(130), 관성 센서(132) 및 온도 센서(138)가 안경다리(102)의 내부에 또는 그 위에 탑재된다. 일 실시예에서 관성 센서(132)는 3축 자력계(132A), 3축 자이로(gyro)(132B) 및 3축 가속도계(132C)를 포함한다(도 3 참조). 관성 센서는 헤드 마운트 디스플레이 장치(2)의 위치, 배향 및 갑작스러운 가속을 감지하기 위한 것이다.

[0022] 마이크로 디스플레이(120)는 렌즈(122)를 통해 이미지를 투사한다. 마이크로 디스플레이(120)를 구현하기 위하여 사용될 수 있는 여러 가지 이미지 생성 기술들이 존재한다. 예컨대, 마이크로 디스플레이(120)는, 광원이 광활성 물질에 의하여 조절되고 백색광에 의해 백라이트되는 투과성 프로젝션 기술을 사용하여 구현될 수 있다. 이를 기술은 통상적으로 강력한 백라이트와 높은 광에너지 밀도를 가진 LCD 타입 디스플레이를 사용하여 구현된다. 마이크로 디스플레이(120)는 또한 광 활성 물질에 의해 외부 광이 반사되고 조절되는 반사성 기술을 사용하여 구현될 수도 있다. 순방향으로의 조명은 기술에 따라 백색 소스 또는 RGB 소스 중 어느 하나에 의해 이루어진다. DGP(digital light processing), LCOS(liquid crystal on silicon) 및 퀼컴사의 Mirasol® 디스플레이 기술은 대부분의 에너지가 모두 조절 구조로부터 반사되어 제거되기 때문에 효율적인 반사성 기술의 예이며 본 명세서에서 기술된 시스템에서 사용될 수 있다. 덧붙여서, 마이크로 디스플레이(120)는 광이 디스플레이에 의해 생성되는 발광성 기술을 사용하여 구현될 수 있다. 예를 들어, 마이크로비전사의 PicoP™ 디스플레이 엔진은 투과성 요소로서 행동하는 작은 스크린 상으로 또는 눈으로 직접 향해지도록(예컨대, 레이저) 조향되는 마이크로 미러를 이용하여 레이저 신호를 방출한다.

[0023] 광 안내 광학 요소(112)는 마이크로 디스플레이(120)로부터 사용자가 착용한 헤드 마운트 디스플레이 장치(2)의

눈(140)으로 광을 전달한다. 광 안내 광학 요소(112)는 또한 광이 헤드 마운트 디스플레이 디바이스(2)의 전면에서부터 광 안내 광학 요소(112)를 통과하여 눈(140)으로 전달될 수 있게 허용하며, 이에 의해 사용자로 하여금 마이크로 디스플레이(120)로부터 가상 영상을 수신하는 것에 덧붙여서 헤드 마운트 디스플레이 장치(2)의 전면에 있는 공간을 실제적이고 직접적으로 볼 수 있도록 허용한다. 따라서, 광 안내 광학 요소(112)의 벽은 시스루 형태이다. 광 안내 광학 요소(112)는 제 1 반사면(124)(예컨대, 미러 또는 그 외 다른 표면)을 포함한다. 마이크로 디스플레이(120)로부터의 광은 렌즈(122)를 통과하고 반사면(124)으로 입사하게 된다. 반사면(124)은 마이크로 디스플레이(120)로부터의 입사광을 반사하여, 광이 광 안내 광학 요소(112)를 구성하는 평면적인 기판 내부에서 내부 반사에 의해 간접도록 한다. 이 기판의 표면들에서 몇 번인가 반사된 이후, 간접 광 웨이브는 선택적으로 반사면(126)들의 어레이에 도달한다. 도면이 지저분하게 되지 않도록 하기 위해 5개의 반사면들 중 오직 하나에만 126이 표기되어 있음을 주목하라. 반사면(126)들은 이들 반사면들로 입사한 광 웨이브를 기판의 외부 사용자의 눈(140)으로 전달되도록 한다. 서로 다른 광선은 서로 다른 각도에서 기판의 내부를 이동하고 반사될 것이기 때문에, 서로 다른 광선은 서로 다른 각도에서 다양한 반사면(126)에 부딪칠 것이다. 그러므로, 서로 다른 광선은 반사면들 중 서로 다른 것들에 의해 기판의 외부로 반사될 것이다. 어떤 면(126)에 의해 어떤 광선이 기판 외부로 반사될 것인가의 선택은 면(126)의 적절한 각도를 선택함에 의하여 구현된다. 일 실시예에서, 각각의 눈은 그것 자체의 광 안내 광학 요소(112)를 가질 것이다. 헤드 마운트 디스플레이 장치가 2개의 광 안내 광학 요소를 가지는 경우, 각각의 눈은 양 눈에 동일한 이미지를 디스플레이할 수 있거나 또는 2개의 눈에 서로 다른 이미지를 디스플레이할 수 있는 그것 자체의 마이크로 디스플레이(120)를 가질 수 있다. 다른 실시예에서, 양 눈에 광을 반사할 수 있는 하나의 광 안내 광학 요소가 존재할 수 있다.

[0024] 불투명 필터(114)는, 광 안내 광학 요소(112)에 대해 정렬되어 있으며, 광 안내 광학 요소(112)를 통과하여 오는 자연광을 균일하게 또는 픽셀-당 기반 중 어느 하나의 방식으로 선택적으로 차단한다. 일 실시예에서, 불투명 필터는 시스루 LCD 패널, 전기크롬 필름(electro chromic film), 또는 불투명 필터로서 이용될 수 있는 유사한 장치일 수 있다. 이러한 시스루 LCD 패널은 종래의 LCD로부터 다양한 층들의 기판, 백라이트 및 디퓨저들을 제거함으로써 얻어질 수 있다. LCD 패널은 광이 액정을 통과하도록 허용하는 하나 이상의 광-투과와 LCD 칩을 포함할 수 있다. 이러한 칩은 예를 들어 LCD 프로젝터에서 사용된다.

[0025] 불투명 필터(114)는 고밀도의 픽셀 그리드를 포함하는데, 여기서 각 픽셀의 광 투과도는 최소 투과도 및 최대 투과도 사이에서 개별적으로 제어가능하다. 0 - 100%의 투과도 범위가 이상적이기는 하지만, 더 제한된 범위도 역시 수용가능하다. 예를 들어, 2개 이하의 편광 필터를 가진 단색 LCD 패널은, LCD의 해상도에 이르기까지 픽셀 당 약 50% 내지 90%의 불투명 범위를 제공하기 위해서 충분하다. 최소값인 50%에서, 렌즈는 약간 어두운 외관을 가질 것이지만, 허용될 수 있다. 100%의 투과도는 완전히 깨끗한 렌즈를 나타낸다. "알파" 스케일은 0 - 100%로부터 정의될 수 있는데, 여기서 0%는 아무런 광의 통과도 허용하지 않으면 100%는 모든 광이 통과되도록 허용한다. 알파의 값은 아래에서 기술된 불투명 필터 컨트롤 회로(224)에 의해 각 픽셀에 대해 설정될 수 있다.

[0026] 알파 값의 마스크는 현실-세계 오브젝트에 대한 프록시(대리)를 이용한 z-버퍼링 이후, 파이프라인 렌더링으로부터 사용될 수 있다. 시스템이 중강 현실 디스플레이를 위한 장면을 렌더링할 때, 시스템은 어떤 현실-세계 오브젝트가 어떤 가상 오브젝트 앞에 있는지를 알아차린다. 만약 어떤 하나의 가상 오브젝트가 어떤 하나의 현실-세계 오브젝트 앞에 있다면, 불투명도는 가상 오브젝트의 커버리지 영역을 위해 온(on) 되어야 할 것이다. 만약 가상 오브젝트가 (가상적으로) 현실-세계 오브젝트 뒤에 있다면, 불투명도는 오프(off)되어야 할 것이며, 뿐만 아니라, 해당 픽셀을 위한 임의의 칼러도 역시 오프되어야 하며, 따라서 사용자는 실제 빛의 그 대응하는 영역(하나 이상의 픽셀 크기로)에 대해 오직 현실-세계 오브젝트만을 볼 것이다. 커버리지는 픽셀-바이-픽셀 기반으로 될 것이며, 따라서 시스템은 어떤 하나의 가상 오브젝트의 일부가 어떤 하나의 현실-세계 오브젝트의 앞에 있고, 이 가상 오브젝트의 일부는 이 현실-세계 오브젝트의 뒤에 있으며, 또한 이 가상 오브젝트의 일부가 이 현실-세계 오브젝트와 합치되어 있는 경우를 취급할 수 있다. 낮은 비용, 전력 및 중량에서 0%에서부터 100%로 갈 수 있는 디스플레이가 이러한 용도를 위해 가장 바람직하다. 더 나아가, 불투명 필터는 넓은 뷰 필드를 제공하기 위하여 컬러 LCD를 이용하여 또는 유기 LED와 같은 그 외 다른 디스플레이를 이용하여 컬러로 렌더링될 수 있다. 불투명 필터에 대한 더 상세한 설명은, 본 명세서에 그 전체가 참조로 병합되어 있는, 2010년 9월 21일 출원된, 미국 특허 출원 번호 12/887,426, "Opacity Filter For See-Through Mounted Display"에서 제공된다.

[0027] LCD와 같은 불투명 필터는 일반적으로 눈에 대해 이러한 근접 거리에서는 초점에서 벗어날 수 있기 때문에 본 명세서에서 기술된 바와 같은 시스루 렌즈와 함께 사용되지 않는다. 그렇지만, 일부 경우에, 이러한 결과는 바람직할 수 있다. 사용자는, 초점이 맞도록 설계되어 있는, 추가적인 컬러를 사용하는 정상적인 HMD 디스플레이를 통해 명확한 컬러 그래픽을 가진 가상 이미지를 본다. LCD 패널은 이러한 디스플레이 "뒤에" 배치됨으로써,

흐릿한 검은 경계가 임의의 가상 컨텐트를 둘러싸게 하여 원하는 만큼 불명료하게 만들 수 있다. 본 시스템은 자연스러운 흐릿함이라는 결함을 안티-에일리어싱 및 대역폭 감소의 특징을 편리하게 획득하는 것으로 변환한다. 이것은 더 낮은 해상도 및 초점을 벗어난 이미지를 사용하는 자연스러운 결과이다. 디지털-샘플링된 이미지의 효과적으로 스무딩(smoothing)이 존재한다. 임의의 디지털 이미지는 에일리어싱을 겪게 되는데, 이 경우 샘플링의 비연속성이라는 성질은, 광의 파장 주변에서, 자연적인 아날로그 및 연속 신호에 반하는 에러를 야기시킨다. 스무딩이란 시각적으로 이상적인 아날로그 신호에 더 가깝운 것을 의미한다. 비록 저 해상도로 갈 때 유실된 정보는 회복되지 않지만, 그 결과 발생하는 에러는 눈에 덜 뛴다.

[0028] 일 실시예에서, 디스플레이 및 불투명 필터는 동시에 렌더링되며 각도-오프셋 문제를 보상하기 위하여 공간 내에서 사용자의 정확한 위치로 조정된다. 눈 추적은 뷰 필드의 말단에서 정확한 이미지 오프셋을 계산하기 위해 채용될 수 있다. 일부 실시예들에서, 불투명도의 양에 있어서 시간적인 또는 공간적인 페이드(fade, 점차 약해짐)가 불투명 필터에서 사용될 수 있다. 유사하게, 가상 이미지 내에서 시간적인 또는 공간적인 페이드가 사용될 수 있다. 하나의 접근방식에 있어서, 불투명 필터의 불투명도의 양에 있어서 시간적인 페이드는 가상 이미지에서의 시간적인 페이드에 대응한다. 다른 접근방식에 있어서, 불투명 필터의 불투명도의 양에 있어서 공간적인 페이드는 가상 이미지에서의 공간적인 페이드에 대응한다.

[0029] 예시적인 하나의 접근방식에서, 증가되는 불투명도는 식별된 사용자의 눈의 위치의 관점으로부터, 가상 이미지의 뒤에 있는 불투명 필터의 픽셀들에 대해 제공된다. 이런 방식에서, 가상 이미지 뒤의 픽셀들은 어두워짐으로써 현실 세계 장면의 대응하는 부분으로부터의 광이 사용자의 눈에 도달하는 것을 막는다. 이것은 가상 이미지가 실제적이며 완전한 컬러 범위와 명도를 나타내도록 허용한다. 더 나아가, 증강 현실 방출기에 의한 전력 소비는 감소되는데 왜냐하면 가상 이미지가 더 낮은 명도에서 제공될 수 있기 때문이다. 불투명 필터가 없다면, 가상 이미지는 현실 세계 장면의 대응하는 부분보다 더 밝은 충분히 높은 명도에서 제공되어야 할 필요가 있는데, 왜냐하면 가상 이미지는 구별되어야 하고 투명하지 않기 때문이다. 불투명 필터의 픽셀들을 어둡게 하는데 있어서, 일반적으로, 가상 이미지의 닫힌 가장자리를 따르는 픽셀들이 가장자리 내부의 픽셀들과 함께 어두워진다. 가장자리의 바로 바깥에 있는 일부 픽셀들 및 가장자리를 둘러싸고 있는 일부 픽셀들도 또한 (가장자리 내부의 픽셀들보다 덜 어둡거나 또는 동일한 수준으로 어둡게 하는 정도로) 어두워질 수 있도록 부분적인 중첩을 제공하는 것이 바람직할 수 있다. 가장자리의 바로 바깥에 있는 이들 픽셀들은 가장자리 내부의 어두움으로부터 가장자리 바깥 불투명의 전체 양에 이르기까지 페이드(예컨대, 불투명에서의 점차적인 전이)를 제공할 수 있다.

[0030] 불투명 필터의 동작의 일 예를 제공하기 위하여, 도 2a는 헤드 마운트 디스플레이 장치(2)를 통하여 바라보고 있는 인간 눈에 보여질 수 있는 숲의 현실-세계 장면의 예를 도시한다. 도 2b는 가상 이미지를 도시하는데, 이 경우 돌고래이다. 도 2c는 도 2b의 가상 이미지의 형상에 기초한 불투명 필터의 구성 예를 도시한다. 불투명 필터는 돌고래가 렌더링될 위치에 증가된 불투명도를 가진 어두운 영역을 제공한다. 증가된 불투명도는 일반적으로 광이 덜 통과되도록 하는 기능을 가질 수 있는 픽셀들의 어두워짐을 지칭한다. 단색 구조에 있어서, 여러 가지 그레이 레벨들(또는 흑색)로 어두워짐, 또는 컬러 구조에 있어서 여러 가지 컬러 레벨들로 어두워짐이 사용될 수 있다. 도 2d는 사용자가 본 예시적인 이미지를 도시하며 사용자의 뷰 필드에 돌고래의 가상 이미지를 투사하고 돌고래의 가상 이미지의 위치에 대응하는 픽셀들에 대해 광을 제거하기 위하여 불투명 필터를 사용한 결과이다. 볼 수 있는 바와 같이, 배경은 돌고래를 통과해서는 보이지 않는다. 비교 목적으로, 도 2e는 불투명 필터를 사용하지 않고 현실 이미지에 가상 이미지를 삽입한 것을 보여준다. 볼 수 있는 바와 같이, 현실 배경은 가상 이미지인 돌고래를 통과하여 보여질 수 있다.

[0031] 헤드 마운트 디스플레이 장치(2)는 또한 사용자의 눈의 위치를 추적하기 위한 시스템을 포함한다. 아래에서 설명될 것과 같이, 시스템은 사용자의 뷰 필드를 결정할 수 있게 하기 위하여 사용자의 위치와 배향을 추적할 것이다. 그렇지만, 인간은 그의 앞에 있는 모든 것을 지각하지 않을 것이다. 그 대신, 사용자의 눈은 환경 중의 임의의 서브세트(subset)로 향해질 것이다. 그러므로, 일 실시예에서, 시스템은 사용자의 뷰 필드의 측정을 정제하기 위하여 사용자의 눈의 위치를 추적하는 기술을 포함할 것이다. 예를 들어, 헤드 마운트 디스플레이 장치(2)는 눈 추적 조립체(134)(도 2 참조)를 포함하는데, 이는 눈 추적조명 장치(134A) 및 눈 추적 카메라(134B)를 포함할 것이다(도 3 참조). 일 실시예에서, 눈 추적 조명 소스(134A)는 하나 이상의 적외선(IR) 방출기를 포함하는데, 이는 눈을 향해 IR 광을 방출한다. 눈 추적 카메라(134B)는 반사된 IR 광을 감지하는 하나 이상의 카메라를 포함한다. 동공의 위치는 각막의 반사를 검출하는 알려진 이미징 기술에 의해 식별될 수 있다. 예를 들어, 본 명세서에 그 전체가 참조로서 병합된, 2008년 7월 22일에 Kranz 등에게 허여된 미국 특허 제7,401,920호 발명의 명칭 "Head Mounted Eye Tracking and Display System"을 참조하라. 이러한 기술은 추적 카메라에 대하여 눈 중심의 위치를 알아낼 수 있다. 일반적으로 눈 추적은 눈의 이미지 획득 및 눈 소켓 내에서 동공의 위치를

결정하기 위해 컴퓨터 비전 기술 사용을 포함한다. 일 실시예에서, 눈은 통상적으로 함께 움직이기 때문에 하나의 눈의 위치를 추적하는 것으로 충분하다. 그렇지만 각각의 눈을 별도로 추적하는 것도 가능하다. 대안적으로, 눈 추적 카메라는, 조명 소스가 있는 경우 또는 없는 경우에, 위치를 검출하기 위하여 눈의 이미지에 기초하여 임의의 움직임을 이용하는 추적 카메라의 대안적인 형태일 수 있다.

[0032] 일 실시예에서, 헤드 마운트 디스플레이 장치(2)의 렌즈의 각 코너에 하나의 IR LED 및 IR광 검출기가 존재할 수 있도록 하기 위하여 시스템은 4각형 배열로 된 4개의 IR RED 및 4개의 IR광 검출기를 사용할 것이다. LED에서 방출된 광은 눈에서 반사된다. 4개의 IR광 검출기의 각각에서 검출된 적외선의 양은 동공 방향을 결정한다. 즉, 눈의 흑색 대비 백색의 양은 그 특정 광 검출기에 있어서 눈에서 반사된 광의 양을 결정할 것이다. 따라서, 광 검출기는 눈에서 백색의 양 또는 흑색의 양에 대해 측정할 것이다. 4개의 샘플로부터, 시스템은 눈의 방향을 결정할 수 있다.

[0033] 다른 대안은, 위에서 논의한 바와 같이 4개의 적외선 LED를 사용하지만, 헤드 마운트 디스플레이 장치(2)의 측부 상에 오직 하나의 적외선 CCD를 사용하는 것이다. CCD는 CCD가 안경 프레임으로부터 보여질 수 있는 눈의 75%에 이르기까지 활성화할 수 있는 하나의 작은 미러 및/또는 렌즈(어안)를 사용할 것이다. 그런 다음 CCD는 이미지를 감지하고, 위에 논의된 바와 매우 유사하게, 이 이미지를 발견하기 위하여 컴퓨터 비전을 사용할 것이다. 따라서, 비록 도 2가 하나의 IR 방출기를 가진 하나의 조립체를 보여주고 있지만, 도 2의 구조는 4개의 IR 송신기 및/또는 4개의 IR 센서를 가지도록 조정될 수 있다. 4개의 IR 송신기 및/또는 4개의 IR 센서보다 더 많거나 더 적은 수도 또한 사용될 수 있다.

[0034] 눈의 방향을 추적하기 위한 다른 실시예는 전하 추적에 기초한다. 이 개념은 망막이 측정 가능한 양의 전하를 가지며 각막은 음의 전하를 가진다는 관찰에 기초하고 있다. 눈이 이리 저리 움직이는 동안 전위를 검출하기 위하여 센서들이 사용자의 귀(이어폰(130) 근처)에 탑재되며 눈이 하고 있는 것을 실시간으로 효과적으로 읽어낸다. 안경 내부에 탑재된 작은 카메라와 같이, 눈 추적을 위한 그 외 다른 실시예들이 또한 사용될 수 있다.

[0035] 하나 이상의 추가 검출기들이 헤드 마운트 디스플레이 장치(2) 상에 포함될 수 있다는 점이 이해되어야만 한다. 이러한 검출기는 장치의 착용자가 바라볼 수 있는 정보를 검출하기 위하여 배치된 SONAR, LIDAR, Structured Light, 및/또는 Time of Flight 거리 검출기를, 제한없이, 포함할 수 있다.

[0036] 도 2는 오직 헤드 마운트 디스플레이 장치(2)의 절반만을 보여준다. 전체 헤드 마운트 디스플레이 장치는 또 하나의 씨 스루 렌즈들의 세트, 또 하나의 불투명 필터, 또 하나의 광 안내 광학 요소, 또 하나의 마이크로 디스플레이(136), 또 하나의 렌즈(122), 방 대면 카메라, 눈 추적 조립체, 마이크로 디스플레이, 이어폰, 및 온도 센서를 포함할 것이다. 헤드 마운트 디스플레이(2)에 대한 추가 설명은 본 명세서에 그 전체가 참조로서 병합되어 있는, 2010년 10월 15일에 출원된 미국 특허 출원 일련 번호 12/905952, 발명의 명칭 "Fusing Virtual Content Into Real Content"에 예시되어 있다.

[0037] 도 3은 헤드 마운트 디스플레이 장치(2)의 다양한 컴포넌트들을 도시하는 블록도이다. 도 4는 프로세싱 유닛(4)의 다양한 컴포넌트들을 기술하는 블록도이다. 그 컴포넌트들이 도 3에 도시되어 있는 헤드 마운트 디스플레이 장치(2)는, 사용자의 환경 내에 있는 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도 결정에 기초하여, 사용자에게 최적 이미지를 디스플레이하기 위하여 사용된다. 추가적으로, 도 3의 헤드 마운트 디스플레이 장치 컴포넌트들은 다양한 조건을 추적하는 많은 센서를 포함한다. 헤드 마운트 디스플레이 장치(2)는 프로세싱 유닛(4)으로부터 가상 이미지에 관한 인스트럭션(instructions)을 수신할 것이며 또한 프로세싱 유닛(4)에게 센서 정보를 제공할 것이다. 프로세싱 유닛(4)은, 그 컴포넌트들이 도 3에 도시되어 있는데, 헤드 마운트 디스플레이 장치(2)로부터 및 또한 허브 컴퓨터 장치(12)(도 1 참조)로부터 센서 정보를 수신할 것이다. 이런 정보에 기초하여, 프로세싱 유닛(4)은 사용자에게 가상 이미지를 제공하는 위치 및 시간을 결정하고 그에 따라 도 3의 헤드 마운트 디스플레이 장치에 인스트럭션을 송신한다.

[0038] 도 3의 컴포넌트들 중 일부(예컨대, 후방 대면 카메라(113), 눈 추적 카메라(134B), 마이크로 디스플레이(120), 불투명 필터(114), 눈 추적 조명(134A), 이어폰(130), 및 온도 센서(138))는, 이를 장치들 각각이 헤드 마운트 디스플레이 장치(2)의 좌측에 하나 그리고 우측에 하나씩 모두 2개 존재하는 것을 나타내기 위하여, 그림자를 가지는 방식으로 도시되었음에 주목하라. 도 3은 전력 관리 회로(202)와 통신하는 컨트롤러 회로(200)를 도시한다. 컨트롤러 회로(200)는 프로세서(210), 메모리(214)(예컨대, D-RAM)와 통신하는 메모리 컨트롤러(212), 카메라 인터페이스(216), 카메라 버퍼(218), 디스플레이 드라이버(220), 디스플레이 포맷터(222), 타이밍 생성기(226), 디스플레이 출력 인터페이스(228), 및 디스플레이 입력 인터페이스(230)를 포함한다. 일 실시예에서, 컨트롤러 회로(200)의 모든 컴포넌트들이 전용 라인이나 하나 이상의 버스를 통해 상호 통신한다. 다른 실시예에

서, 컨트롤 회로(200)의 각각의 컴포넌트가 프로세서(210)와 통신한다. 카메라 인터페이스(216)는 2개의 방 대면 카메라(113)에 인터페이스를 제공하며 방 대면 카메라로부터 수신된 이미지를 카메라 버퍼(218)에 저장한다. 디스플레이 드라이버(220)는 마이크로 디스플레이(120)를 구동할 것이다. 디스플레이 포맷터(222)는, 불투명 필터(114)를 제어하는 불투명 컨트롤 회로(224)에게, 마이크로 디스플레이(120) 상에 디스플레이되고 있는 가상 이미지에 관한 정보를 제공한다. 타이밍 생성기(226)는 시스템을 위한 타이밍 데이터를 제공하기 위해 사용된다. 디스플레이 출력 인터페이스(228)는 방 대면 카메라(113)로부터의 이미지를 프로세싱 유닛(4)에게 제공하기 위한 버퍼이다. 디스플레이 입력(230)은 마이크로 디스플레이(120) 상에 디스플레이될 가상 이미지와 같은 이미지를 수신하기 위한 버퍼이다. 디스플레이 출력(228) 및 디스플레이 입력(230)은 프로세싱 유닛(4)에 대한 인터페이스인 밴드 인터페이스(232)와 통신한다.

[0039] 전력 관리 회로(202)는 전압 레귤레이터(234), 눈 추적 조명 드라이버(236), 오디오 DAC 및 증폭기(238), 마이크로폰 전치증폭기 오디오 ADC(240), 온도 센서 인터페이스(242) 및 클럭 발생기(244)를 포함한다. 전압 레귤레이터(234)는 밴드 인터페이스(232)를 통해 프로세싱 유닛(4)으로부터 전력을 수신하고 이 전력을 헤드 마운트 디스플레이 장치(2)의 그 외 다른 컴포넌트들에게 제공한다. 눈 추적 조명 드라이버(236)는 위에서 기술된 바와 같이 눈 추적 조명(134A)을 위해 IR 광원을 제공한다. 오디오 DAC 및 증폭기(238)는 이어폰(130)으로부터 오디오 정보를 수신한다. 마이크로폰 전치증폭기 및 오디오 ADC(240)는 마이크로폰(110)에 대한 인터페이스를 제공한다. 온도 센서 인터페이스(242)는 온도 센서(138)를 위한 인터페이스이다. 전력 관리 유닛(202)은 또한 3축 자력계(132A), 3축 자이로(132B) 및 3축 가속도계(132C)에 전력을 제공하며 그들로부터 데이터를 수신한다.

[0040] 도 4는 프로세싱 유닛(4)의 다양한 컴포넌트들을 기술하는 블록도이다. 도 4는 전력 관리 회로(306)와 통신하는 컨트롤 회로(304)를 보여준다. 컨트롤 회로(304)는 중앙 프로세싱 유닛(CPU)(320), 그래픽 프로세싱 유닛(GPU)(322), 캐시(324), RAM(326), 메모리(330)(예컨대, D-RAM)와 통신하는 메모리 컨트롤(328), 플래시 메모리(334)(또는 그 외 다른 타입의 비-휘발성 저장장치)와 통신하는 플래시 메모리 컨트롤러(332), 밴드 인터페이스(302) 및 밴드 인터페이스(232)를 통해 헤드 마운트 디스플레이 장치(2)와 통신하는 디스플레이 출력 버퍼(336), 밴드 인터페이스(302) 및 밴드 인터페이스(232)를 통해 헤드 마운트 디스플레이 장치(2)와 통신하는 디스플레이 입력 버퍼(338), 마이크로폰에 연결하기 위한 외부 마이크로폰 커넥터(342)와 통신하는 마이크로폰 인터페이스(340), 무선 통신 장치(346)에 연결하기 위한 PCI 익스프레스 인터페이스, 및 USB 포트(들)(348)을 포함한다. 일 실시예에서, 무선 통신 장치(346)는 Wi-Fi 구동 통신 장치, 블루투스 통신 장치, 전외선 통신 장치 등을 포함할 수 있다. USB 포트는 프로세싱 유닛(4)을 충전하는 것 뿐만 아니라 프로세싱 유닛 상에 데이터나 소프트웨어를 로딩하기 위하여 프로세싱 유닛(4)을 허브 컴퓨팅 장치(12)에 도킹시키기 위해 사용될 수 있다. 일 실시예에서, CPU(320) 및 GPU(322)는 사용자의 뷰 필드에 가상 이미지를 언제, 어디서, 어떻게 삽입할 것인가를 결정하기 위한 주요 장치들이다. 아래에서 더 상세히 설명된다.

[0041] 전력 관리 회로(306)는 클럭 생성기(360), 아날로그 디지털 컨버터(362), 배터리 충전기(364), 전압 레귤레이터(366), 헤드 마운트 디스플레이 전력 소스(376), 및 온도 센서(374)(프로세싱 유닛(4)의 손목 밴드 상에 위치된)와 통신하는 온도 센서 인터페이스(372)를 포함한다. 아날로그 디지털 컨버터(362)는 AC 공급을 수신하고 시스템을 위한 DC 공급을 생성하기 위해 충전 잭(370)에 연결된다. 전압 레귤레이터(366)는 시스템에 전력을 공급하기 위한 배터리(368)와 통신한다. 배터리 충전기(364)는 충전 잭(370)으로부터 전력을 수신하면 (전압 레귤레이터(366)를 통해) 배터리(368)를 충전하기 위해 사용된다. HMD 전력 소스(376)는 헤드 마운트 디스플레이 장치(2)에게 전력을 제공한다.

[0042] 위에 기술된 시스템은 가상 이미지가 현실 세계 오브젝트의 뷰를 대체할 수 있도록 하기 위하여 사용자의 뷰 필드 내로 가상 이미지를 삽입하도록 구성될 것이다. 대안적으로, 가상 이미지는 현실 세계 오브젝트의 이미지를 대체하지 않고도 삽입될 수 있다. 다양한 실시예들에서, 가상 이미지는 대체되는 오브젝트 또는 그 이미지가 삽입되는 환경에 기초하여 적절한 배향, 크기 및 형상과 합치하도록 조정될 것이다. 덧붙여서, 가상 이미지는 반사도 및 그림자를 포함하도록 조정될 수 있다. 일 실시예에서, 헤드 마운트 디스플레이 장치(12), 프로세싱 유닛(4) 및 허브 컴퓨팅 장치(12)는, 장치들 각각이 가상 이미지를 어디서, 언제 및 어떻게 삽입하여야 할지를 결정하기 위한 데이터를 획득하기 위하여 사용되는 센서들의 서브세트를 포함하고 있으므로 협력한다. 일 실시예에서, 어디서, 어떻게 및 언제 가상 이미지를 삽입할지를 결정하는 계산은 허브 컴퓨팅 장치(12)에 의하여 수행된다. 다른 실시예에서, 이를 계산은 프로세싱 유닛(4)에 의해 수행된다. 다른 실시예에서 계산 중 일부는 허브 컴퓨팅 장치(12)에 의하여 수행되는 한편 그 외 다른 계산은 프로세싱 유닛(4)에 의하여 수행된다. 그 외 다른 실시예들에서 계산은 헤드 마운트 디스플레이 장치(12)에 의하여 수행될 수 있다.

[0043] 예시적인 일 실시예에서, 허브 컴퓨팅 장치(12)는 사용자가 그 안에 존재하는 환경의 모델을 생성하고 그 환경

내에서 이동하는 다양한 오브젝트들을 추적할 것이다. 대안적인 실시예에서, 환경 내에서 흥미로운 오브젝트 및 사용자들은 그 외 다른 수단에 의해 결정된다. 덧붙여서, 허브 컴퓨팅 장치(12)는 헤드 마운트 디스플레이 장치(2)의 위치와 배향을 추적함으로써 헤드 마운트 디스플레이 장치(2)의 뷰 필드를 추적한다. 모델 및 추적 정보는 허브 컴퓨팅 장치(12)로부터 프로세싱 유닛(4)으로 제공된다. 헤드 마운트 장치(2)에 의해 획득된 센서 정보는 프로세싱 유닛(4)으로 전송된다. 이후 프로세싱 유닛(4)은 사용자의 뷰 필드를 정제하고 가상 이미지를 어떻게, 어디서, 및 언제 삽입할지를 대한 인스트럭션을 헤드 마운트 디스플레이 장치(2)에게 제공하기 위하여 헤드 마운트 디스플레이 장치(2)로부터 수신한 추가적인 센서 정보를 이용한다.

[0044] 도 5는 캡처 장치를 가진 허브 컴퓨팅 시스템(12)의 예시적인 일 실시예를 도시한다. 일 실시예에서, 캡처 장치(20A 및 20B)는 동일 구조이며, 따라서, 도 5는 오직 캡처 장치(20A)만을 보여준다. 예시적인 일 실시예에 따라, 캡처 장치는 예를 들어 타임오브플라이트(time-of-flight), 구조화된 광(structured light), 스테레오 이미지 등을 포함하는 임의의 적당한 기술을 통해 깊이 값들을 포함할 수 있는 깊이 이미지를 포함하여 깊이 정보를 가진 비디오를 캡처하도록 구성될 수 있다. 일 실시예에 따라, 캡처 장치(20A)는 깊이 정보를 "Z 충들" 즉 그 시선을 따라 깊이 카메라에서부터 연장하는 Z 축에 대해 수직일 수 있는 충들로 조직화할 수 있다.

[0045] 도 5에 도시된 바와 같이, 캡처 장치(20A)는 카메라 컴포넌트(423)를 포함할 수 있다. 예시적인 일 실시예에 따라, 카메라 컴포넌트(423)는 임의 장면의 깊이 이미지를 캡처할 수 있는 깊이 카메라이거나 깊이 카메라를 포함할 수 있다. 깊이 이미지는 캡처된 장면의 2-차원(2-D) 픽셀 영역을 포함할 수 있는데, 여기서 2-D 픽셀 영역 내의 각각의 픽셀은 카메라로부터 캡처된 장면 안에 있는 오브젝트의, 예컨대 센터미터, 밀리미터 등과 같은 단위의 길이와 같은, 깊이 값을 나타낼 수 있다.

[0046] 카메라 컴포넌트(23)는 장면의 깊이 이미지를 캡처하기 위하여 사용될 수 있는 적외(IR) 광 컴포넌트(425), 3-차원(3-D) 카메라(426), 및 RGB(비주얼 이미지) 카메라(428)를 포함할 수 있다. 예를 들어, 타임-오브-플라이트 분석에서, 캡처 장치(20A)의 IR 광 컴포넌트(425)는 장면에 적외광을 방출할 수 있고 그런 다음, 예컨대 3-D 카메라(426) 및/또는 RGB 카메라(428)를 사용하여 그 장면에서 하나 이상의 타겟 및 오브젝트의 표면에서 반사되어 흘어지는 광을 검출(일부 실시예들에서, 도시되지 않는 센서들을 사용하여)하기 위해 센서들을 사용할 수 있다. 일부 실시예들에서, 펄스형 적외광이 사용될 수 있는데, 인출되는 광 펄스와 대응하는 인입하는 광 펄스 사이의 시간이 측정되어 캡처 장치(20A)로부터 해당 장면에서 타겟이나 오브젝트 상의 특정 위치에 이르는 물리적인 거리를 결정하기 위하여 사용될 수 있다. 덧붙여서, 그 외 다른 예시적인 실시예들에서, 위상 시프트를 결정하기 위하여 인출되는 광 펄스파의 위상이 인입하는 광 펄스파의 위상과 비교될 수 있다. 이후 위상 시프트는 캡처 장치로부터 타겟이나 오브젝트 상의 특정 위치에 이르는 물리적인 거리를 결정하기 위하여 사용될 수 있다.

[0047] 다른 예시적인 일 실시예에 따라, 예컨대 셔터드(shuttered) 광 펄스 이미징을 포함하는 다양한 기술을 통해 시간에 걸쳐 반사된 광 빔의 강도를 분석함으로써 캡처 장치(20A)로부터 타겟이나 오브젝트 상의 특정 위치에 이르는 물리적인 거리를 간접적으로 결정하기 위하여 타임-오브-플라이트 분석이 사용될 수 있다.

[0048] 다른 예시적인 일 실시예에서, 캡처 장치(20A)는 깊이 정보를 캡처하기 위하여 구조화된 광을 사용할 수 있다. 이러한 분석에 있어서, 패턴화된 광(즉, 격자 패턴, 줄무늬 패턴, 또는 상이한 패턴과 같은 알려진 패턴으로 디스플레이된 광)은 예컨대 IR 광 컴포넌터(424)를 통해 장면 상에 투사될 수 있다. 장면 내의 하나 이상의 타겟이나 오브젝트의 표면에 부딪치는 것에 반응하여 패턴은 변형되어질 수 있다. 이러한 패턴의 변형은 예컨대 3-D 카메라(426) 및/또는 RGB 카메라(428)(및/또는 그 외 다른 센서)에 의해 캡처될 수 있으며 그런 다음 캡처 장치로부터 타겟이나 오브젝트 상의 특정 위치까지의 물리적인 거리를 결정하기 위하여 분석될 수 있다. 일부 구현 예들에서, 카메라들(425 및 426)로부터 거리를 결정하기 위하여 삼각측량이 사용될 수 있도록 IR 광 컴포넌트(425)는 카메라들(425 및 426)로부터 탈거될 수 있다. 일부 구현예들에 있어서, 캡처 장치(20A)는 IR 광을 감지하기 위한 전용 IR 센서, 또는 IR 필터를 가진 센서를 포함할 것이다.

[0049] 다른 실시예에 따라, 캡처 장치(20A)는 깊이 정보를 생성하기 위하여 결합될 수 있는 비주얼 스테레오 데이터를 획득하기 위하여 서로 다른 각도에서 하나의 장면을 볼 수 있는 2개 이상의 물리적으로 떨어져 있는 카메라를 포함할 수 있다. 그 외 다른 타입의 깊이 이미지 센서들도 또한 깊이 이미지를 생성하기 위하여 사용될 수 있다.

[0050] 캡처 장치(20A)는 마이크로폰(430)을 더 포함할 수 있는데, 이 마이크로폰은 음향을 수신하여 전기적인 신호로 변환시킬 수 있는 트랜스듀서 또는 센서를 포함한다. 마이크로폰(430)은 또한 허브 컴퓨팅 시스템(12)에 의해 제공될 수 있는 오디오 신호를 수신하기 위하여 사용될 수 있다.

[0051] 예시적인 일 실시예에서, 캡처 장치(20A)는 이미지 카메라 컴포넌트(423)와 통신할 수 있는 프로세서(432)를 더 포함할 수 있다. 프로세서(432)는, 예컨대 깊이 이미지를 수신하고, 적절한 데이터 포맷(예컨대, 프레임)을 생성하고 또한 그 데이터를 허브 컴퓨팅 시스템(12)으로 전송하기 위한 인스트럭션을 포함하는 인스트럭션을 실행할 수 있는 표준화된 프로세서, 특수 목적의 프로세서, 마이크로프로세서 등을 포함할 수 있다.

[0052] 캡처 장치(20A)는 프로세서(432)에 의해 실행되는 인스트럭션, 3-D 카메라 및/또는 RGB 카메라에 의해 캡처된 이미지 또는 이미지의 프레임, 또는 임의의 적절한 정보, 이미지 등을 저장할 수 있는 메모리(434)를 더 포함할 수 있다. 예시적인 일 실시예에 따라, 메모리(434)는 RAM(random access memory), ROM(read only memory), 캐시, 플래시 메모리, 하드 디스크, 또는 임의의 그 외 적당한 저장 컴포넌트를 포함할 수 있다. 도 5에 도시된 바와 같이, 일 실시예에서, 메모리(434)는 이미지 캡처 컴포넌트(423) 및 프로세서(432)와 통신하는 별도의 컴포넌트일 수 있다. 다른 실시예에 따라, 메모리(434)는 프로세서(432) 및/또는 이미지 캡처 컴포넌트(422) 내에 병합될 수 있다.

[0053] 캡처 장치(20A 및 20B)는 통신 링크(436)를 통해 허브 컴퓨팅 시스템(12)과 통신한다. 통신 링크(436)는 예를 들어 USB 연결, 파이어와이어(Firewire) 연결, 이더넷 연결 등을 포함하는 유선 연결 및/또는 무선 802.11b, g, a 또는 n 연결과 같은 무선 연결일 수 있다. 일 실시예에 따라, 허브 컴퓨팅 시스템(12)은, 캡처 장치(20A)에, 예컨대, 통신 링크(436)를 통해 임의의 장면을 캡처하는 때를 결정하기 위하여 사용될 수 있는 클럭(clock)을 제공할 수 있다. 덧붙여서, 캡처 장치(20A)는 예컨대 3-D 카메라(426) 및/또는 RGB 카메라(428)에 의해 캡처된 깊이 정보 및 비주얼(예컨대, RGB) 이미지를 통신 링크(436)를 통해 허브 컴퓨팅 시스템(12)에 제공한다. 일 실시예에서, 깊이 이미지 및 비주얼 이미지는 초당 30 프레임으로 전송되지만; 그러나 그 외 다른 프레임 속도가 사용될 수 있다. 그런 다음 허브 컴퓨팅 시스템(12)은 모델, 깊이 정보, 및 캡처된 이미지를 생성하고 사용하여, 예컨대 게임, 또는 워드프로세서와 같은 애플리케이션을 제어하거나 및/또는 아바타 또는 온-스크린 캐릭터를 애니메이팅할 수 있다.

[0054] 허브 컴퓨팅 시스템(12)은 깊이 이미지 프로세싱 및 골격 추적 모듈(450)을 포함하며, 이것은 캡처 장치(20A)의 깊이 카메라 기능에 의해 검출가능한 하나 이상의 사람을 추적하기 위해 깊이 이미지를 사용한다. 깊이 이미지 프로세싱 및 골격 추적 모듈(450)은, 비디오 게임, 생산성 애플리케이션 통신 애플리케이션 또는 그 외 다른 소프트웨어 애플리케이션 등일 수 있는, 애플리케이션(453)에 추적 정보를 제공한다. 오디오 데이터 및 비주얼 이미지 데이터도 또한 애플리케이션(452) 및 깊이 이미지 프로세싱 및 골격 추적 모듈(450)에 제공된다. 애플리케이션(452)은 추적 정보, 오디오 데이터 및 비주얼 이미지 데이터를 인식기 엔진(454)에 제공한다. 다른 실시예에서, 인식기 엔진(454)은 깊이 이미지 프로세싱 및 골격 추적 모듈(450)로부터 직접 추적 정보를 수신하고 캡처 장치(20A 및 20B)로부터 직접 오디오 데이터 및 비주얼 이미지 데이터를 수신한다.

[0055] 인식기 엔진(454)은 일단의 필터들(460, 462, 464, ..., 466)과 연관되어 있으며, 이들 필터 각각은 캡처 장치(20A 또는 20B)에 의해 검출가능한 임의의 사람이나 오브젝트에 의해 수행될 수 있는 제스처, 행동 또는 조건에 관한 정보를 포함한다. 예컨대, 캡처 장치(20A)로부터의 데이터는 사용자 또는 일단의 사용자들이 하나 이상의 제스처 또는 그 외 다른 행동을 수행한 때를 식별하기 위하여 필터들(460, 462, 464, ..., 466)에 의해 프로세싱될 수 있다. 이들 제스처는 다양한 컨트롤, 오브젝트, 또는 애플리케이션(452)의 조건과 연관될 수 있다. 따라서 허브 컴퓨팅 시스템(12)은 오브젝트(사람을 포함하여)의 움직임을 해석하고 추적하기 위하여, 필터와 함께, 인식기 엔진(454)을 사용할 수 있다.

[0056] 캡처 장치(20A 및 20B)는 허브 컴퓨팅 시스템(12)에 RGB 이미지(또는 그 외 다른 포맷이나 컬러 공간의 비주얼 이미지) 및 깊이 이미지를 제공한다. 깊이 이미지는 복수의 관찰된 픽셀들일 수 있는데, 여기서 각각의 관찰된 픽셀은 하나의 관찰된 깊이 값을 가진다. 예를 들어, 깊이 이미지는 캡처된 장면의 2-차원(2-D) 픽셀 영역을 포함할 수 있는데, 여기서 2-D 픽셀 영역 내의 각각의 픽셀이 캡처 장치에 의해 캡처된 장면 내의 오브젝트의 거리 같은, 깊이 값을 가질 수 있다. 허브 컴퓨팅 시스템(12)은 사용자의 움직임이나 오브젝트의 움직임을 추적하기 위하여 RGB 이미지 및 깊이 이미지를 사용할 것이다. 예를 들어, 시스템은 깊이 이미지를 사용하여 사람의 골격을 추적할 것이다. 깊이 이미지를 사용하여 사람의 골격을 추적하기 위해 사용될 수 있는 많은 방법들이 있다. 깊이 이미지를 사용하여 골격을 추적하는 하나의 적합한 예는 본 명세서에 그 전체가 참조로서 병합되어 있는, Craig 등이 2009년 10월 21일에 출원한 미국 특허 출원 12/603,437 "Pose Tracking Pipeline"(이하, '437 애플리케이션으로 지칭한다)에서 제공된다. '437 애플리케이션의 프로세스는 깊이 이미지를 획득하는 과정, 데이터를 다운 샘플링하는 과정, 높은 변동 잡음 데이터를 제거 및/또는 스무딩하는 과정, 배경을 식별하고 제거하는 과정, 및 전경 픽셀들의 각각을 몸의 서로 다른 부분으로 할당하는 과정을 포함한다. 이들 단계에 기초하여, 시스템은 데이터에 모델을 맞추고 골격을 생성할 것이다. 골격은 한 세트의 관절들과 관절 사이의 연결부

들을 포함할 것이다. 그 외 다른 추적 방법들도 또한 사용될 수 있다. 적합한 추적 기술들이 다음의 4개의 미국 특허 출원들, 즉 2009년 5월 29일에 출원된 미국 특허 출원 12/475,308, "Device for Identifying and Tracking Multiple Humans Over Time"; 2010년 1월 29일에 출원된 미국 특허 출원 12/696,282, "Visual Based Identity Tracking"; 2009년 12월 18일에 출원된 미국 특허 출원 12/641,788, "Motion Detection Using Depth Images"; 및 2009년 10월 7일에 출원된 미국 특허 출원 12/575,388, "Human Tracking System"에도 또한 개시되어 있는데, 이들은 모두 그 전체가 본 명세서에 참조로서 병합되어 있다.

[0057]

인식기 엔진(454)은 제스처 또는 행동을 결정하기 위하여 다수의 필터(460, 462, 464, ..., 466)를 포함한다. 필터는 해당 제스처, 행동 또는 조건에 대하여 파라미터, 또는 메타데이터와 함께 제스처, 액션 또는 조건을 한정하는 정보를 포함한다. 예를 들어, 던지기(throw)는, 깊이 카메라에 의해 그 움직임이 캡처되어지는 바와 같이, 몸의 후방의 뒤에서부터 몸의 전방의 앞을 지나는 손 하나의 움직임을 포함하는데, 몸의 후방의 뒤에서부터 몸의 전방의 앞을 지나는 사용자의 손 하나의 움직임을 나타내는 정보를 포함하는 제스처로서 구현될 수 있다. 그런 다음 이 제스처에 대해 파라미터들이 설정될 수 있다. 제스처가 던지기인 경우에, 파라미터는 손이 도달해야만 하는 문턱 속도, 손이 이동하는 거리(절대적인, 또는 사용자의 크기를 전체로 하는 상대적인 둘 중 어느 하나), 및 인식기 엔진에 의하여 그 제스처가 발생되는 신뢰도(confidence rating)일 수 있다. 제스처에 대한 이들 파라미터는 애플리케이션마다 변화될 수 있으며, 하나의 단일 애플리케이션의 맥락들 사이에서 변화될 수 있으며, 또는 하나의 애플리케이션의 하나의 맥락 내에서도 시간에 따라 변화될 수 있다.

[0058]

필터들은 모듈형이거나 교체가능형일 수 있다. 일 실시예에서, 하나의 필터는 다수의 입력(이들 입력 각각이 하나의 타입을 가지는) 및 다수의 출력(이들 출력 각각이 하나의 타입을 가지는)을 가진다. 제 1 필터는 인식기 엔진 아키텍처의 임의의 그 외 다른 양상을 변화시킬 수 있는 제 1 필터와 동일한 수와 타입의 입력과 출력을 가지는 제 2 필터로 대체될 수 있다. 예를 들어, 입력으로서 골격 데이터를 취하고 필터와 연관된 제스처가 발생하는 신뢰도 및 조향 각도를 출력하는 드라이빙을 위한 제 1 필터가 존재할 수 있다. 이 제 1 드라이빙 필터를 제 2 드라이빙 필터로 대체하고 싶은 경우 - 아마도 제 2 드라이빙 필터가 더욱 효율적이며 더 적은 프로세싱 자원을 요구하기 때문에 - 제 2 필터가 동일한 입력과 출력 - 골격 데이터 타입의 한가지 입력 및 신뢰도 타입 및 각도 타입의 2개의 출력 - 을 가지는 한 단순히 제 1 필터를 제 2 필터로 대체함으로써 그렇게 할 수 있다.

[0059]

필터는 반드시 파라미터를 가질 필요가 없다. 예를 들어, 사용자의 키를 리턴하는 "사용자 키" 필터는 튜닝될 수 있는 임의의 파라미터를 허용하지 않을 수 있다. 대안적인 "사용자 키" 필터는 - 사용자의 키를 결정하는데 있어 사용자의 신발, 헤어스타일, 모자 및 자세를 고려하는지 여부와 같은 - 튜닝가능한 파라미터를 가질 수 있다.

[0060]

필터에 대한 입력은 사용자의 관절 위치에 관한 관절 데이터, 관절에서 만나는 뼈들에 의해 형성된 각도, 장면으로부터의 RGB 데이터, 및 사용자의 모습의 변화 속도와 같은 것들을 포함할 수 있다. 필터로부터의 출력은 주어진 제스처가 만들어지는 신뢰도, 제스처 움직임이 이루어지는 속도, 및 제스처 움직임이 만들어지는 시간과 같은 것들을 포함할 수 있다.

[0061]

인식기 엔진(454)은 필터들에 기능성을 제공하는 하나의 기본 인식기 엔진을 가질 수 있다. 일 실시예에서, 인식기 엔진(454)이 구현하는 기능성에는 인식된 제스처 및 그 외 다른 입력을 추적하는 시간-결친-입력 아카이브, Hidden Markov Model 구현(여기서 모델링된 시스템은 미지의 파라미터를 가진 Markov 프로세스 - 여기서 현재 상태는 미래 상태를 결정하기 위하여 사용된 임의의 과거 상태 정보를 포함하고, 따라서 그 외 다른 과거 상태 정보는 이런 목적으로 유지되지 않아야만 한다 - 라고 간주되며, 숨겨진 파라미터들은 관찰가능한 데이터로부터 결정된다), 및 제스처 인식의 특정 예들을 해결하기 위하여 사용된 그 외 다른 기능성이 포함된다.

[0062]

필터들(460, 462, 464, ..., 466)은 인식기 엔진(454)의 상부에 로딩되어 실행되며 모든 필터(460, 462, 464, ..., 466)에게 인식기 엔진(454)에 의해 제공된 서비스들을 이용할 수 있다. 일 실시예에서, 인식기 엔진(454)은 데이터를 수신하여 임의의 필터(460, 462, 464, ..., 466)의 요구사항에 합치되는지 여부를 결정한다. 입력을 파싱(parsing)하는 것과 같은, 이를 제공된 서비스들이 각각의 필터(460, 462, 464, ..., 466)에 의해서라기 보다는 인식기 엔진(454)에 의해 한 번 제공되기 때문에, 이러한 서비스는 시간 주기 동안 필터마다 한 번씩이 아니라 해당 주기 내에 한 번만 프로세싱되는 것이 필요하고, 따라서 제스처를 결정하기 위하여 사용되는 프로세싱은 감소된다.

[0063]

애플리케이션(452)은 인식기 엔진(454)에 의해 제공된 필터들(460, 462, 464, ..., 466)을 사용할 수 있으며, 또는 인식기 엔진(454)으로 플러그인하는 그것 자체의 필터를 제공할 수도 있다. 일 실시예에서, 모든 필터들은 이 플러그-인 특성을 인에이블하기 위한 하나의 공통 인터페이스를 가진다. 더 나아가, 모든 필터들은 파라미터

들을 이용할 수 있고, 따라서 아래의 하나의 단일한 제스처 툴은 전체 필터 시스템을 디버그하고 투닝하기 위하여 사용될 수 있다.

[0064] 인식기 엔진(454)에 관한 더 많은 정보는, 본 명세서에 그 전체가 참조로서 병합되어 있는, 2009년 4월 13일에 출원된 미국 특허 출원 12/422,661, "Gesture Recognizer System Architecture"에서 발견될 수 있다. 제스처 인식에 대한 더 많은 정보는 2009년 2월 23일에 출원된 미국 특허 출원 12/391,150, "Standard Gestures"; 및 2009년 5월 29일에 출원된 미국 특허 출원 12/474,655, "Gesture Tool"에서 발견될 수 있으며, 이들 둘 모두는 그 전체가 본 명세서에 참조로서 병합되어 있다.

[0065] 일 실시예에서, 컴퓨팅 시스템(12)은 허브 컴퓨팅 시스템(12)과 상호작용하는 하나 이상의 사용자에 관련된 사용자-특정 정보를 포함하는 사용자 프로파일 데이터베이스(470)를 포함한다. 일 예에서, 사용자-특정 정보는 사용자의 표현된 선호, 사용자의 친구 리스트, 사용자의 선호하는 활동, 사용자의 리마인더(reminder) 리스트, 사용자의 소셜 그룹, 사용자의 현재 위치, 사용자의 환경에 있는 오브젝트와 상호작용하는 사용자의 과거의 의도와 같은 사용자에 관련된 정보 및 사용자의 사진, 이미지 및 기록 비디오와 같은 그 외에 사용자가 생성한 컨텐트를 포함한다. 일 실시예에서, 사용자-특정 정보는 사용자의 소셜 네트워킹 사이트, 주소록, 이메일 데이터, 인스턴트 메시징 데이터, 사용자 프로파일 또는 그 외 다른 인터넷 상의 소스들과 같은 하나 이상의 데이터 소스로부터 획득될 수 있다. 하나의 접근 방식에 있어서, 그리고 아래에서 상세히 논의될 것과 같이, 사용자-특정 정보는 사용자의 환경 내에 있는 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 자동으로 결정하기 위하여 사용된다.

[0066] 도 6은 허브 컴퓨팅 시스템(12)을 구현하기 위하여 사용될 수 있는 컴퓨팅 시스템의 예시적인 일 실시예를 도시한다. 도 6에 도시된 바와 같이, 멀티미디어 콘솔(500)은 레벨 1 캐시(502), 레벨 2 캐시(504), 및 플래시 ROM(Read Only Memory)(506)을 가진 중앙 프로세싱 유닛(CPU)(501)을 가진다. 레벨 1 캐시(502) 및 레벨 2 캐시(504)는 데이터를 일시적으로 저장하며 따라서 메모리 액세스 사이클의 수를 감소시키고, 그에 의해 프로세싱 속도와 스루풋을 향상시킨다. CPU(501)는 하나보다 많은 코어를 가지며, 또한 따라서, 추가적인 레벨 1 및 레벨 2 캐시들(502 및 504)를 가지도록 제공될 수 있다. 플래시 ROM(506)은 멀티미디어 콘솔(500)에 전원이 들어올 때 부팅 프로세스의 초기 시기 동안에 로딩되는 실행가능한 코드를 저장할 수 있다.

[0067] 그래픽 프로세싱 유닛(GPU)(508) 및 비디오 인코더/비디오 코덱(코더/디코더)(514)은 고속 및 고해상도 그래픽 프로세싱을 위한 비디오 프로세싱 파이프라인을 형성한다. 그래픽 프로세싱 유닛(508)로부터 버스를 통해 비디오 인코더/비디오 코덱(514)으로 데이터가 운반된다. 비디오 프로세싱 파이프라인은 텔레비전이나 그 외 다른 디스플레이로 전송되도록 하기 위하여 A/V(audio/video) 포트(540)으로 데이터를 출력한다. 메모리 컨트롤러(510)는 RAM(Random Access Memory)와 같은, 그러나 이에만 국한되는 것이 아닌, 다양한 타입의 메모리(512)에 대한 프로세서 액세스를 용이하게 하기 위하여 GPU(508)에 연결된다.

[0068] 멀티미디어 콘솔(500)은 I/O 컨트롤러(520), 시스템 관리 컨트롤러(522), 오디오 프로세싱 유닛(523), 네트워크 인터페이스(524), 제 1 USB 호스트 컨트롤러(526), 제 2 USB 컨트롤러(528) 및 바람직하게는 모듈(518) 상에 구현되어 있는 전면 패널 I/O 서브조립체(530)를 포함한다. USB 컨트롤러(526 및 528)는 주변장치 컨트롤러(542(1)-542(2)), 무선 어댑터(548), 및 외부 메모리 장치(546)(예컨대, 플래시 메모리, 외장 CD/DVD ROM 드라이브, 탈거가능 미디어 등)을 위한 호스트로서 이용된다. 네트워크 인터페이스(524) 및/또는 무선 어댑터(548)는 네트워크(예컨대, 인터넷, 홈 네트워크 등)에의 액세스를 제공하며 이더넷 카드, 모뎀, 블루투스 모듈, 케이블 모뎀, 등을 포함하는 다양한 유선 또는 무선 어댑터 컴포넌트들의 넓은 다양성 중 임의의 것일 수 있다.

[0069] 시스템 메모리(543)는 부팅 프로세스 동안에 로딩되는 애플리케이션 데이터를 저장하기 위하여 제공된다. 미디어 드라이브(544)가 제공되며 DVD/CD 드라이브, 블루투스 드라이브, 하드 디스크 드라이브, 또는 그 외 다른 탈거가능 미디어 드라이브 등을 포함할 수 있다. 미디어 드라이브(144)는 멀티미디어 콘솔(500)에 대해 내장형이거나 외장형일 수 있다. 애플리케이션 데이터는 멀티미디어 콘솔(500)에 의한 실행, 재생 등을 위하여 미디어 드라이브(544)를 통해 액세스될 수 있다. 미디어 드라이브(544)는 직렬 ATA 버스 또는 그 외 고속 연결(예컨대, IEEE 1394)과 같은 버스를 통해 I/O 컨트롤러(520)에 연결된다.

[0070] 시스템 관리 컨트롤러(522)는 멀티미디어 콘솔(500)의 이용가능성을 보장하기 위하여 다양한 관련 서비스 기능을 제공한다. 오디오 프로세싱 유닛(523) 및 오디오 코덱(532)은 높은 충실도와 스테레오 프로세싱을 가진 대응 오디오 프로세싱 파이프라인을 형성한다. 오디오 데이터는 통신 링크를 통해 오디오 프로세싱 유닛(523) 및 오디오 코덱(532) 사이에서 운반된다. 오디오 프로세싱 파이프라인은 외부의 오디오 사용자나 오디오 성능을 가지는 장치에 의한 재생을 위해 A/V 포트(540)로 데이터를 출력한다.

- [0071] 전면 패널 I/O 서브조립체(530)는 전원 버튼(550) 및 이젝트 버튼(552)의 기능성뿐만 아니라 멀티미디어 콘솔(100)의 외부 표면 상에 노출된 임의의 LED(light emitting diode)들이나 그 외 다른 지시자들을 지원한다. 시스템 전력 공급 모듈(536)은 멀티미디어 콘솔(100)의 컴포넌트들에게 전력을 공급한다. 팬(538)은 멀티미디어 콘솔(500) 내의 화로를 냉각시킨다.
- [0072] CPU(501), GPU(508), 메모리 컨트롤러(510), 및 멀티미디어 콘솔(500) 내의 다양한 그 외 다른 컴포넌트들은, 직렬 및 병렬 버스, 메모리 버스, 주변장치 버스, 및 다양한 버스 아키텍처 중 임의의 것을 사용하는 프로세서나 로컬 버스를 포함하는, 하나 이상의 버스들을 통해 상호연결된다. 예시적으로 말하자면, 이러한 아키텍처는 PCI(Peripheral Component Interconnects) 버스, PCI-Express 버스 등을 포함할 수 있다.
- [0073] 멀티미디어 콘솔(500)이 전원 온 되는 때, 애플리케이션 데이터는 시스템 메모리(543)로부터 메모리(512) 및/또는 캐시(502, 504)로 로딩되고 CPU(501) 상에서 실행될 수 있다. 애플리케이션은 멀티미디어 콘솔(500) 상에서 이용 가능한 여러 가지 미디어 타입들에 대해 네비게이팅할 때 일관된 사용자 경험을 제공하는 그래픽 유저 인터페이스를 표현할 수 있다. 동작시, 애플리케이션 및/또는 미디어 드라이브(544) 내에 담겨진 그 외 다른 미디어는 멀티미디어 콘솔(500)에게 추가적인 기능성을 제공하기 위하여 미디어 드라이브(544)로부터 런칭되거나 재생될 수 있다.
- [0074] 멀티미디어 콘솔(500)은 시스템을 텔레비전이나 그 외 다른 디스플레이에 단순히 연결하는 독립 시스템으로서 동작될 수 있다. 이러한 독립 모드에서, 멀티미디어 콘솔(500)은 하나 이상의 사용자가 시스템과 상호작용하거나, 영화를 시청하거나, 또는 음악을 청취할 수 있게 허용한다. 그렇지만, 네트워크 인터페이스(524) 또는 무선 어댑터(548)를 통해 이용 가능하게 만들어지는 광대역 연결성을 통합한 경우, 멀티미디어 콘솔(500)은 더 넓은 네트워크 커뮤니티 내의 참가자로서 더 동작될 수 있다. 덧붙여서, 멀티미디어 콘솔(500)은 무선 어댑터(548)를 통해 프로세싱 유닛(4)과 통신할 수 있다.
- [0075] 멀티미디어 콘솔(500)이 전원 온 된 때에, 하드웨어 자원의 한 세트의 양은 멀티미디어 콘솔 운영 시스템에 의한 시스템 사용을 위해 예약된다. 이를 자원은 메모리, CPU 및 GPU 사이클, 네트워킹 대역폭 등의 예약을 포함할 수 있다. 이를 자원이 시스템 부팅 시간에 예약되기 때문에, 예약된 자원은 애플리케이션의 관점에서는 존재하지 않는다. 특히, 메모리 예약은 바람직하게 런치 커널, 협력 시스템 애플리케이션들 및 드라이버들을 포함하기에 충분히 크다. CPU 예약은 바람직하게 만약 예약된 CPU 용도가 시스템 애플리케이션에 의해 사용되지 않는 경우 아이들 스레드(idle thread)가 임의의 사용되지 아니한 사이클을 소비하는 방식으로 지속된다.
- [0076] GPU 예약에 관련하여, 시스템 애플리케이션에 의해 생성된 가벼운 메시지(예컨대, 팝업)는 오버레이로 팝업을 렌더링하도록 코드를 스케줄링하는 GPU 인터럽트를 사용함으로써 디스플레이된다. 오버레이를 위해 요구되는 메모리의 양은 오버레이 영역 크기에 따라 다르며 오버레이는 바람직하게 스크린 해상도에 따라 스케일링된다. 전체적인 사용자 인터페이스가 협력 시스템 애플리케이션에 의해 사용되는 경우, 애플리케이션 해상도에 대해 독립적인 해상도를 사용하는 것이 바람직하다. 주파수를 변화시킬 필요 및 TV를 재동기화(resync)하여야 할 필요를 제거하는 방식으로 이러한 해상도를 설정하기 위해 스케일러가 사용될 수 있다.
- [0077] 멀티미디어 콘솔(500)이 부팅되고 시스템 자원이 예약된 이후, 협력 시스템 애플리케이션들이 실행되어 시스템 기능성을 제공한다. 시스템 기능성은 위에서 기술된 예약 시스템 자원들 내에서 실행되는 한 세트의 시스템 애플리케이션들로 대표된다. 운영 시스템 커널은 게임 애플리케이션 스레드를 대비 시스템 애플리케이션 스레드들인 스레드들을 식별한다. 시스템 애플리케이션들은 바람직하게 애플리케이션에 대해 일관된 시스템 자원 관점을 제공하기 위하여 미리 결정된 시간과 간격에서 CPU(501) 상에서 실행되도록 스케줄링된다. 스케줄링은 콘솔에서 실행되는 게임 애플리케이션에 대해 캐시 장애를 최소화하기 위한 것이다.
- [0078] 협력 시스템 애플리케이션이 오디오를 요구할 때, 오디오 프로세싱은 시간 민감도에 기인하여 게임 애플리케이션에 대해 비동기적으로 스케줄링된다. 멀티미디어 콘솔 애플리케이션 관리자(아래에서 기술되는)는 시스템 애플리케이션이 활성일 때 게임 애플리케이션 오디오 레벨(예컨대, 소거, 감쇠)을 제어한다.
- [0079] 선택적인 입력 장치(예컨대 컨트롤러(542(1) 및 542(2)))는 게임 애플리케이션과 시스템 애플리케이션에 의해 공유된다. 입력 장치는 예약된 자원이 아니지만 각각 장치의 포커스를 가지는 방식으로 시스템 애플리케이션과 게임 애플리케이션 사이에서 스위칭된다. 애플리케이션 관리자는 바람직하게 게임 애플리케이션의 지식을 알지 못한 상태로 입력 스트림의 스위칭을 제어하며, 드라이버는 포커스 스위칭에 관한 상태 정보를 유지한다. 캡처 장치(20A 및 20B)는 USB 컨트롤러(526) 또는 그 외 다른 인터페이스를 통해 콘솔(500)을 위한 추가 입력 장치를 한정할 수 있다. 그 외 다른 실시예들에서, 허브 컴퓨팅 시스템(12)은 그 외 다른 하드웨어 아키텍처를 사용하

여 구현될 수 있다. 어떠한 하드웨어 아키텍처도 필수적인 것은 아니다.

[0080] 도 1은 하나의 허브 컴퓨팅 장치(12)(허브라고 지칭된다)와 통신하는 하나의 헤드 마운트 디스플레이 장치(2) 및 프로세싱 유닛(4)(집합적으로 모바일 디스플레이 장치라고 지칭된다)를 도시한다. 다른 실시예에서, 다수의 모바일 디스플레이 장치들이 하나의 단일 허브와 통신할 수 있다. 각각의 모바일 디스플레이 장치는 위에서 기술된 바와 같이 무선 통신을 사용하여 허브와 통신할 것이다. 이러한 일 실시예에서, 모든 모바일 디스플레이 장치들에게 유용한 정보의 대부분은 허브에서 계산되어 저장되고 각각의 모바일 디스플레이 장치들에게 전송되어질 것으로 설계된다. 예컨대, 허브는 환경의 모델을 생성하고 허브와 통신하고 있는 모든 모바일 디스플레이 장치들에게 이 모델을 제공할 것이다. 덧붙여서, 허브는 모바일 디스플레이 장치들 및 실내에 있는 이동하는 오브젝트의 위치와 배향을 추적할 수 있고, 그런 다음 해당 정보를 각각의 모바일 디스플레이 장치들에게 전송할 수 있다.

[0081] 다른 실시예에서, 시스템은 다수의 허브를 포함할 수 있으며, 각각의 허브는 하나 이상의 모바일 디스플레이 장치들을 포함할 수 있다. 허브들은 직접 또는 인터넷(또는 그 외 다른 네트워크)를 통해 서로 통신할 수 있다. 예컨대, 도 7은 허브(560, 562 및 564)를 보여준다. 허브(560)는 허브(562)와 직접 통신한다. 허브(560)는 인터넷을 통해 허브(564)와 통신한다. 허브(560)는 모바일 디스플레이 장치(570, 572, ..., 574)와 통신한다. 허브(562)는 모바일 디스플레이 장치(578, 580, ..., 582)와 통신한다. 허브(564)는 모바일 디스플레이 장치(584, 586, ..., 588)와 통신한다. 각각의 모바일 디스플레이 장치는 위에서 논의한 바와 같이 무선 통신을 통해 각자의 개별 허브와 통신한다. 만약 이들 허브가 하나의 공통 환경 내에 있다면, 각각의 허브는 이 환경의 모델의 일부를 제공하거나, 또는 하나의 허브가 그 외 다른 허브들을 위해 모델을 생성할 수 있다. 각각의 허브는 이동하는 오브젝트들의 서브세트를 추적하고 해당 정보를 그 외 다른 허브들과 공유할 것이며, 그 외 다른 허브들은 차례로 그 정보를 적절한 모바일 디스플레이 장치들과 공유할 것이다. 모바일 디스플레이 장치를 위한 센서 정보는 각자의 개별 허브로 제공될 것이며 그런 다음 결국 그 외 다른 모바일 디스플레이 장치들에게 공유되도록 하기 위하여 그 외 다른 허브들에게 공유될 것이다. 따라서, 허브들 사이에서 공유된 정보는 골격 추적, 모델들에 관한 정보, 애플리케이션의 다양한 상태, 및 그 외 다른 추적을 포함할 수 있다. 허브들과 그들 각자의 모바일 디스플레이 장치들 사이에서 소통되는 정보에는 움직이는 오브젝트의 추적 정보, 세계 모델을 위한 상태 및 물리 업데이트, 지리 및 텍스처 정보, 비디오 및 오디오, 그리고 본 명세서에서 기술된 동작을 수행하기 위해 사용되는 그 외 다른 정보가 포함된다.

[0082] 도 8은 헤드 마운트 디스플레이 장치를 착용한 사용자에게 표현되는 시각화된 정보의 디스플레이를 최적화하기 위한 프로세스의 일 실시예를 기술한다. 단계(600)에서, 시스템(10)이 구성된다. 예컨대, 애플리케이션(도 5의 애플리케이션(452))은 최적 이미지가 설계된 위치에서 장면의 3차원 모델 안으로 삽입되어져야 함을 지시하도록 시스템을 구성할 수 있다. 다른 예에서, 허브 컴퓨팅 시스템(12) 상에서 실행되는 애플리케이션은 중장 컨텐트(특정 가상 이미지 또는 가상 오브젝트와 같은)가 비디오 게임이나 그 외 다른 프로세스의 일부로서 장면 안으로 삽입되어야 함을 지시할 것이다.

[0083] 단계(605)에서, 사용자의 환경 내의 흥미있는 오브젝트가 결정될 것이다. 일 실시예에서, 이것은 단계들(602 ~ 606)에서 윤곽이 그려진 바와 같이 사용자의 공간에 대한 3차원 모델을 생성함으로써 수행될 수 있다. 단계(602)에서, 시스템은 헤드 마운트 디스플레이 장치(2)가 위치하고 있는 공간의 부피 분석적 모델을 생성할 것이다. 일 실시예에서, 예컨대, 허브 컴퓨팅 장치(12)는 헤드 마운트 디스플레이 장치(2)가 위치하고 있는 환경이나 장면의 3차원 모델을 생성하기 위하여 하나 이상의 깊이 카메라로부터의 깊이 이미지를 사용할 것이다. 단계(604)에서, 상기 모델은 하나 이상의 오브젝트로 분할된다. 예컨대, 만약 허브 컴퓨팅 장치(12)가 어떤 방의 3차원 모델을 생성하면, 그 방은 그 안에 다수의 오브젝트를 가지고 있을 것이다. 방 안에 있을 수 있는 오브젝트의 예에는 사람, 의자, 테이블, 소파 등이 포함된다. 단계(604)는 서로 구별된 오브젝트를 결정하는 과정을 포함한다. 단계(606)에서, 시스템은 오브젝트를 식별할 것이다. 예컨대, 허브 컴퓨팅 장치(12)는 특정 오브젝트는 테이블이고 다른 오브젝트는 의자라는 것을 식별할 수 있다.

[0084] 대안적으로, 단계(605)는 임의 개수의 수단에 의하여 수행될 수 있다. 일 실시예에서, 센서들이 방 안에 있는 하나 이상의 오브젝트 또는 개인들에 내장될 수 있다. 흥미있는 오브젝트를 식별하거나 또는 오브젝트의 타입 및 특정 애플리케이션에 관하여 오브젝트에게 가질 수 있는 흥미의 수준과 같은 추가 정보를 제공하는 검출 가능한 신호를 방출하는 센서들이 사용자의 환경 내에 있는 오브젝트들을 식별하기 위하여 사용될 수 있다. 대안적으로, 그 외 다른 센서 기술들이 위에 개시된 캡처 장치 중 어느 하나의 범위 내에서 오브젝트들을 식별하기 위하여 사용될 수 있는데, 환경 내에 또는 헤드 마운트 디스플레이 장치 상에 배치된 그 외 다른 거리 검출 센서 기술을 채용하는 그 외 다른 센서 장치들(SONAR, LIDAR, 구조화된 광, 및/또는 타임 오브 플라이트 거리 검출기

와 같은)은 모두 가능한 포커싱 오브젝트 및 그들의 형상을 제공하기 위하여 사용될 수 있다.

[0085] 다른 대안은 한번에 하나의 프레임씩, 환경의 스테레오 재구성을 수행하기 위하여 위에 기술된 깊이 카메라들 중 하나 이상을 사용하는 것이다. 또 다른 대안은, 오브젝트 스스로가 헤드 마운트 디스플레이를 향하는 광 필스를 방출할 수 있고 그래서 그 오브젝트가 흥미있는 것임과 그 오브젝트를 포함하는 디스플레이의 해당 부분이 하이라이트되어야 함을 디스플레이에게 지시할 수 있다.

[0086] 도 8의 단계(608)에서, 시스템은 모델 또는 위의 기술에 기초하여 사용자의 뷔 필드를 결정한다. 즉, 시스템은 환경이나 공간에서 사용자가 바라보고 있는 부분을 결정한다. 일 실시예에서, 단계(608)는 허브 컴퓨팅 장치(12), 프로세싱 유닛(4) 및 헤드 마운트 디스플레이 장치(2)를 사용하여 수행될 수 있다. 하나의 예시적인 구현 예에서, 허브 컴퓨팅 장치(12)는 헤드 마운트 디스플레이 장치(2)의 위치와 배향에 대한 예비적인 결정을 제공하기 위하여 사용자와 헤드 마운트 디스플레이 장치(2)를 추적할 것이다. 헤드 마운트 디스플레이 장치(2) 상의 센서들은 결정된 배향을 정제하기 위하여 사용될 것이다. 예컨대, 위에 기술된 관성 센서들(34)은 헤드 마운트 디스플레이 장치(2)의 배향을 정제하기 위하여 사용될 수 있다. 덧붙여서, 아래에서 기술되는 눈 추적 프로세스는, 특히 사용자가 바라보고 있는 곳 즉 뷔 필드 중에 사용자의 포커스 영역 또는 깊이 포커스라고 알려져 있는 곳에 대응하는, 초기에 결정된 뷔 필드의 서브세트를 식별하기 위하여 사용될 수 있다. 더 상세한 사항은 도 12, 도 13a 및 도 13b를 참조하여 아래에서 기술될 것이다.

[0087] 단계(610)에서, 프로세싱 유닛(4)에서 실행되는 소프트웨어와 같은 시스템은 사용자의 뷔 필드 내에서 사용자의 현재 포커스 영역을 결정한다. 도 13a 및 도 13b에서 아래에 더 논의되는 바와 같이, 각각의 눈에 대해 눈 추적 카메라(134)에 의해 캡처된 데이터에 기초하는 눈 추적 프로세싱은 사용자의 현재 포커스 영역을 제공할 수 있다. 예컨대 동공들 사이의 수렴지점과 사용자의 얼굴 위치를 지시하는 데이터는, 그것으로부터 포커스 영역, 파눔의 융합 영역(Panum's fusion area)이 계산될 수 있는 포커싱 곡선, 단시체적(Horopter) 상의 하나의 포커싱 포인트에 대해 삼각측량하기 위하여 사용될 수 있다. 파눔의 융합 영역은 인간 눈에 의해 사용되는 양안 입체영상(binocular stereopsis)을 위한 단일 비전의 영역이다.

[0088] 단계(612)에서, 프로세싱 유닛(4)은 홀로 또는 허브 컴퓨팅 장치(12)(소프트웨어의 제어 하에)와 협력하여 사용자의 포커스 영역 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정한다. 단계(614)에서, 소프트웨어의 제어 하에, 프로세싱 유닛(4)은 홀로 또는 허브 컴퓨팅 장치(12)와 협력하여 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도에 기초하여 최적 이미지를 생성하고 이 최적 이미지를 사용자에게 디스플레이 한다.

[0089] 단계(616)에서, 헤드 마운트 디스플레이 장치(2)의 사용자는 헤드 마운트 디스플레이 장치(2)에서 디스플레이된 최적 이미지에 기초하여 허브 컴퓨팅 장치(12)(또는 다른 컴퓨팅 장치) 상에서 실행되는 애플리케이션과 상호작용한다. 도 8의 프로세싱 단계들(608-616)은, 사용자의 뷔 필드와 포커스 영역이 사용자가 그의 또는 그녀의 머리를 움직일 때 갱신되며, 사용자의 포커스 영역 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도가 결정되고, 또한 사용자의 의도에 기초하여 최적 이미지가 사용자에게 디스플레이되는 방식으로, 시스템의 동작 동안에 연속적으로 수행될 수 있다. 단계들(604-614) 각각은 아래에서 더 상세히 기술된다.

[0090] 도 9는 사용자의 공간의 모델을 생성하기 위한 프로세스의 일 실시예를 기술한다. 예컨대, 도 9의 프로세스는 도 8의 단계(602)를 구현하는 일 예이다. 단계(620)에서, 허브 컴퓨팅 시스템(12)은, 헤드 마운트 디스플레이 장치가 있는 한경의, 도 1에 도시된 것들과 같은, 다수의 관점에 대해 하나 이상의 깊이 이미지를 수신한다. 예컨대, 허브 컴퓨팅 장치(12)는 다수의 깊이 카메라들로부터 깊이 이미지들을 획득할 수 있고 또는 서로 다른 방향으로 카메라를 포인팅하거나 또는 모델이 세워질 환경이나 공간의 전체 뷔를 허용하는 렌즈를 가진 깊이 카메라를 사용함으로써 동일한 카메라로부터 다수의 깊이 이미지들을 획득할 수 있다. 단계(622)에서, 다양한 깊이 이미지들로부터의 깊이 데이터는 하나의 공통 좌표 시스템에 기초하여 결합된다. 예컨대, 만약 이 시스템이 다수의 카메라들로부터 깊이 이미지들을 수신한다면, 시스템은 하나의 공통 좌표 시스템을 가지도록 2개의 이미지를 상관시킬 것이다(예컨대, 이미지들을 라인업). 단계(624)에서, 깊이 데이터를 사용하여 공간의 부피 분석적 설명이 생성된다.

[0091] 도 10은 공간의 모델을 오브젝트들로 분할하는 프로세스의 일 실시예를 기술하는 흐름도이다. 예컨대, 도 10의 프로세스는 도 8의 단계(604)의 예시적인 일 구현예이다. 도 10의 단계(626)에서, 시스템은 위에서 논의한 바와 같이 하나 이상의 깊이 카메라들로부터 하나 이상의 깊이 이미지들을 수신할 것이다. 대안적으로, 시스템은 이미 수신한 하나 이상의 깊이 이미지들에 액세스할 수 있다. 단계(628)에서, 시스템은 위에서 기술된 카메라들로부터 하나 이상의 비주얼 이미지들을 수신할 것이다. 대안적으로, 시스템은 이미 수신된 하나 이상의 비주얼 이

미지들을 액세스할 수 있다. 단계(630)에서, 허브 컴퓨팅 시스템은 깊이 이미지들 및/또는 비주얼 이미지들에 기초하여 하나 이상의 사람을 검출할 것이다. 예컨대, 시스템은 하나 이상의 골격을 인식할 것이다. 단계(632)에서, 허브 컴퓨팅 장치는 깊이 이미지 및/또는 비주얼 이미지에 기초하여 모델 내에서 에지(edge)들을 검출할 것이다. 단계(634)에서, 허브 컴퓨팅 장치는 검출된 에지를 사용하여 서로 구별된 오브젝트들을 식별할 것이다. 예컨대, 에지는 오브젝트 사이의 경계선이라고 가정한다. 단계(636)에서, 도 9의 프로세스를 사용하여 생성된 모델은 서로 다른 오브젝트에 연된되어 있는 모델 부분을 보여주기 위하여 업데이트될 것이다.

[0092] 도 11a는 오브젝트를 식별하기 위한 프로세스의 일 실시예를 기술하는 흐름도이다. 예컨대 도 11a의 프로세스는 도 8의 단계(606)의 예시적인 일 구현예이다. 단계(640)에서, 허브 컴퓨팅 장치(12)는 식별된 사람을 사용자 아이덴티티와 매칭시킬 것이다. 예컨대, 사용자 프로파일은 캡쳐 장치에 의해 수신된 검출 오브젝트의 이미지와 매칭될 수 있는 비주얼 이미지를 포함할 수 있다. 대안적으로, 사용자 프로파일은 깊이 이미지나 비주얼 이미지에 기초하여 매칭될 수 있는 사람의 특징을 기술할 수 있다. 다른 실시예에서, 사용자는 시스템에 로그인할 수 있고 허브 컴퓨팅 장치(12)는 로그인 프로세스를 사용하여 특정 사용자를 식별하고 본 명세서에서 기술된 상호 작용 전체에 걸쳐 해당 사용자를 추적할 수 있다. 단계(642)에서, 허브 컴퓨팅 장치(12)는 형상 데이터베이스를 액세스할 것이다. 단계(644)에서, 허브 컴퓨팅 장치는 모델 내의 오브젝트를 떨 수 있는 한 많이 이 데이터베이스 내의 형상들에 매칭시킬 것이다. 단계(646)에서, 매칭되지 아니한 형상들은 하이라이트되어 사용자에게 디스플레이될 것이다(예컨대 모니터(16)를 사용하여). 단계(648)에서, 허브 컴퓨팅 장치(12)는 하이라이트된 형상들의 각각(또는 서브세트)를 식별하는 사용자 입력을 수신할 것이다. 예컨대, 사용자는 각각의 미식별된 형상이 무엇인지를 지시하기 위하여 키보드, 마우스, 음성 입력, 또는 그 외 다른 타입의 입력을 사용할 수 있다. 단계(650)에서, 형상 데이터베이스는 단계(648)의 사용자 입력에 기초하여 업데이트된다. 단계(652)에서, 단계(602)에서 생성되고, 단계(632)에서 업데이트된 환경 모델은, 오브젝트 각각에 대한 메타데이터를 추가함으로써 더 업데이트된다. 메타데이터는 오브젝트를 식별한다. 예컨대, 메타데이터는 특정 오브젝트가 등글고 광나는 테이블, 존 도우, 녹색의 가죽 소파 등임을 지시할 수 있다.

[0093] 도 11b는 움직이는 오브젝트(예컨대, 움직이는 사람이나 그 외 다른 타입의 오브젝트)에 반응하여 도 8의 프로세스에 의해 생성된 모델을 업데이트하는 프로세스의 일 실시예를 기술하는 흐름도이다. 단계(654)에서, 시스템은 오브젝트가 움직이는 것을 결정한다. 예컨대, 시스템은 연속적으로 깊이 이미지를 수신할 것이다. 만약 깊이 이미지가 시간에 걸쳐 변화하면, 오브젝트가 움직이고 있는 것이다. 만약 움직이는 오브젝트가 검출되지 않는다면, 시스템은 깊이 이미지 수신을 계속할 것이고 움직이는 오브젝트를 찾는 것을 계속할 것이다.

[0094] 만약 움직이는 오브젝트가 존재한다면, 단계(654)에서 시스템은 움직이는 오브젝트를 식별할 것이다. 프레임 디페レン싱(frame differencing) 또는 임의의 다양한 추적 기술이 움직이는 오브젝트를 인식하기 위해 그리고 인식된 오브젝트를 도 8의 단계(606)에서 식별된 오브젝트 중 하나에 상관시키기 위해 사용될 수 있다. 일부 오브젝트는 움직일 때 형상을 변화시킬 것이다. 예컨대, 인간은 그 인간이 걷거나 달릴 때 형상이 변화할 수 있다. 단계(658)에서, 움직이는 오브젝트의 새로운 형상이 식별되어 저장된다. 단계(660)에서, 이전에 생성되었던 공간 모델은 움직이는 오브젝트의 새로운 위치와 형상에 기초하여 업데이트된다. 도 11b의 프로세스는 허브 컴퓨팅 장치(12)의 프로세싱 유닛(4)에 의해 수행될 수 있다.

[0095] 도 12는, 도 8의 단계(608)의 예시적인 일 구현예인, 사용자의 뷔 필드를 결정하기 위한 프로세스, 및 도 8의 단계(610)의 예시적인 일 구현예인, 사용자의 포커스 영역을 결정하기 위한 프로세스의 일 실시예를 기술하는 흐름도이다. 도 12의 프로세스는 허브 컴퓨팅 장치(12)로부터의 정보 및 위에 기술된 눈 추적 기술에 의지한다. 도 13a는 도 12의 프로세스에서 사용된 추적 정보를 제공하기 위하여 허브 컴퓨팅 시스템에 의해 수행되는 프로세스의 일 실시예를 기술하는 흐름도이다. 도 13b는, 그 결과가 도 12의 프로세스에서 사용되는, 눈을 추적하는 프로세스의 일 실시예를 기술하는 흐름도이다.

[0096] 도 13a의 단계(686)에서, 허브 컴퓨팅 장치(12)는 사용자의 위치를 추적할 것이다. 예컨대, 허브 컴퓨팅 장치(12)는 사용자를 추적하기 위하여(예컨대, 골격 추적을 사용하여) 하나 이상의 깊이 이미지 및 하나 이상의 비주얼 이미지를 사용할 것이다. 하나 이상의 깊이 이미지 및 하나 이상의 비주얼 이미지는 단계(688)에서 헤드 마운트 디스플레이 장치(2)의 위치 및 헤드 마운트 디스플레이 장치(2)의 배향을 결정하기 위하여 사용될 수 있다. 단계(690)에서, 사용자와 헤드 마운트 디스플레이 장치(2)의 위치 및 배향은 허브 컴퓨팅 장치(12)로부터 프로세싱 유닛(4)으로 전송된다. 단계(692)에서 위치 및 배향 정보가 프로세싱 유닛(4)에서 수신된다. 도 13a의 프로세싱 단계들은 사용자가 연속적으로 추적되는 방식으로 시스템의 동작 동안에 연속적으로 수행될 수 있다.

[0097] 도 13b는 환경 내에서 사용자의 눈의 위치를 추적하기 위한 일 실시예를 기술하는 흐름도이다. 단계(662)에서,

눈이 조명된다. 예컨대 눈은 눈 추적 조명(134A)으로부터의 적외광을 사용하여 조명될 수 있다. 단계(664)에서, 하나 이사의 눈 추적 카메라(134B)를 사용하여 눈으로부터의 반사가 검출된다. 단계(665)에서, 반사 데이터는 헤드 마운트 디스플레이 장치(2)로부터 프로세싱 유닛(4)으로 전송된다. 단계(668)에서, 프로세싱 유닛(4)은 위에서 논의된 바와 같이 반사 데이터에 기초하여 눈의 위치를 결정할 것이다.

[0098] 도 12는 사용자의 뷰 필드를 결정하는(예컨대, 도 8의 단계(608)) 및 사용자의 포커스 영역을 결정하는(예컨대, 도 8의 단계(610)) 프로세스의 일 실시예를 기술하는 흐름도이다. 단계(670)에서, 프로세싱 유닛(4)은 허브로부터 수신된 최신 위치 및 배향 정보를 액세스할 것이다. 도 12의 프로세스는 단계(686)에서부터 단계(690)까지 화살표로 표시된 바와 같이 연속적으로 수행될 수 있고; 따라서, 프로세싱 유닛(4)은 허브 컴퓨팅 장치(12)로부터 업데이트된 위치 및 배향 정보를 주기적으로 수신할 것이다. 그렇지만, 프로세싱 유닛(4)은 그것이 허브 컴퓨팅 장치(12)로부터 업데이트된 정보를 수신하는 것보다 더 자주 가상 이미지를 그려야 할 필요가 있을 것이다. 그러므로, 프로세싱 유닛(4)은 허브 컴퓨팅 장치(12)로부터의 샘플들 사이에 배향에 대한 업데이트를 제공하기 위하여 국소적으로(예컨대, 헤드 마운트 장치(2)로부터) 감지된 정보에 의지할 필요가 있을 것이다.

[0099] 단계(672)에서, 프로세싱 유닛(4)은 3축 자로(132B)로부터의 데이터를 액세스할 것이다. 단계(674)에서, 프로세싱 유닛(4)은 3축 가속도계(132C)로부터의 데이터를 액세스할 것이다. 단계(676)에서, 프로세싱 유닛(4)은 3축 자력계(132A)로부터의 데이터를 액세스할 것이다. 단계(678)에서, 프로세싱 유닛(4)은 자이로, 가속도계 및 자력계로부터의 데이터로 허브 컴퓨팅 장치(12)로부터의 위치 및 배향 데이터를 정제(또는 즉 업데이트)할 것이다. 단계(680)에서, 프로세싱 유닛(4)은 헤드 마운트 디스플레이 장치의 위치와 배향에 기초하여 잠재적인 뷰 포인트를 결정할 것이다.

[0100] 단계(682)에서, 프로세싱 유닛(4)은 최신 눈 위치 정보를 액세스할 것이다. 단계(684)에서, 프로세싱 유닛(4)은 눈 위치에 기초하여, 잠재적인 뷰 포인트의 서브세트로서, 사용자가 보고 있는 모델 부분을 결정할 것이다. 예컨대, 사용자는 벽을 대면하고 있을 수 있으며, 따라서 헤드 마운트 디스플레이에 있어서 뷰 포인트는 벽을 따라 임의의 곳을 포함할 수 있다. 그렇지만, 만약 사용자의 눈이 오른쪽을 포인팅하고 있으면, 단계(684)는 사용자의 뷰 필드가 오직 벽의 오른쪽 부분이라고 결론내릴 것이다. 단계(684)에서의 결론에서, 프로세싱 유닛(4)은 헤드 마운트 디스플레이(2)를 통해 사용자의 뷰 포인트를 결정하였다. 그런 다음, 프로세싱 유닛(4)은 가상 이미지를 삽입하기 위한 해당 뷰 필드 내의 위치를 식별하고 불투명 필터를 사용하여 광을 차단할 수 있다. 도 12의 프로세싱 단계들은, 사용자가 그의 또는 그녀의 머리를 움직임에 따라 사용자의 뷰 필드 및 포커스 영역이 연속적으로 업데이트되는 방식으로, 시스템의 동작 동안에 연속적으로 수행될 수 있다.

[0101] 도 14는 사용자의 포커스 영역 내 하나 이상의 오브젝트와 상호작용하는 사용자의 추측된 또는 표현된 의도를 결정하기 위한 프로세스의 일 실시예를 기술하는 흐름도이다. 예컨대, 도 14의 프로세스는 도 8의 단계(612)의 예시적인 일 구현예이다. 도 14의 단계들은, 소프트웨어의 제어 하에, 허브 컴퓨팅 장치(12) 또는 프로세싱 유닛(4)에 의해 수행될 수 있다. 도 14는 사용자의 포커스 영역 내에서 사용자의 눈 움직임 패턴을 검출하는 것 및 사용자의 포커스 영역 내에서 사용자가 보고 있는 하나 이사의 오브젝트에 대한 사용자 시선의 강도를 결정하는 것에 기초하여 사용자의 포커스 영역 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도가 결정되는 프로세스를 기술한다. 도 14a는 사용자와 관련된 사용자-특정 정보에 기초하여 하나 이상의 오브젝트나 사람과 상호작용하려는 사용자의 의도가 추측되는 프로세스를 기술한다. 도 14b는 하나 이상의 오브젝트와 상호작용하는 사용자의 표현된 의도에 기초하여 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도가 결정되는 프로세스가 기술된다.

[0102] 도 14의 단계(690)에서, 사용자의 포커스 영역 내에서 사용자의 눈 움직임 패턴이 검출된다. 알게 될 것과 같이, 눈 움직임은 통상적으로 각각 눈 시선이 특정 위치에서 정지할 때, 및 다른 위치로 이동할 때인 정착(fixations) 및 단속이동(saccades)으로 분리된다. 그러므로 눈 움직임 패턴은 이미지나 시각적인 장면을 볼 때 사용자의 눈이 따라가는 일련의 정착 및 단속이동으로서 한정될 수 있다. 결과적인 일련의 정착 및 단속이동은 일반적으로 스캔경로(scanpath)라고 지칭되는데, 이는 시각적인 장면을 바라볼 때 사용자의 눈이 따라가는 경로이다. 덧붙여서, 눈으로부터의 대부분의 정보는 정착 동안에 이용가능하게 된다. 시각적인 각도의 중심의 1도 또는 2도(중심(fovea))는 시각적 정보의 대부분을 제공하며 더 크게 멀어진 각도(주변)로부터의 입력은 정보가 적다. 따라서, 스캔경로를 따르는 정착 위치들은 시각적인 장면을 볼 때 사용자의 눈에 의해 처리되었던 자극에 대한 정보 궤적을 보여준다. 일 실시예에서, 스캔경로를 따르는 정착 위치들은 사용자의 포커스 영역 내에서 사용자가 보고 있는 하나 이상의 오브젝트를 검출하기 위하여 사용될 수 있다.

[0103] 단계(692)에서, 사용자가 사용자의 포커스 영역 내의 하나 이상의 오브젝트를 보고 있는지 여부가 결정된다. 위

에서 논의된 바와 같이, 일 실시예에서, 스캔경로를 따르는 정착 위치들은 사용자가 사용자의 포커스 영역 내의 하나 이상의 오브젝트를 보고 있는지 여부를 검출하기 위하여 사용될 수 있다. 만약 사용자가 어떤 오브젝트도 보고 있지 않다고 결정되면, 일 실시예에서 도 14a에 기술된 단계들(710-716)이 수행된다.

[0104] 만약 사용자가 사용자의 포커스 영역 내의 하나 이상의 오브젝트를 보고 있다고 결정된다면, 단계(694)에서, 사용자가 보고 있는 오브젝트가 식별된다. 하나 이상의 오브젝트는 도 11에 기술되어 있는 프로세스를 이용함으로써 식별될 수 있다. 예컨대, 오브젝트는 벽시계, 둥글고 광나는 테이블, 존 도우, 녹색 가죽 소파 등으로 식별될 수 있다. 단계(696)에서, 사용자의 포커스 영역에서 사용자가 보고 있는 하나 이상의 오브젝트에 대한 사용자의 시선의 강도가 결정된다. 일 예에서, 사용자의 시선의 강도는 어떤 시간 윈도우 내에 오브젝트에 대한 사용자 시선(또는 정착)의 지속시간을 결정하는 것에 기초하여 결정된다.

[0105] 단계(698)에서, 사용자의 시선 강도가 미리-한정된 문턱 값보다 더 큰지 여부가 결정된다. 일 예에서, 미리-한정된 문턱 값은 10초이다. 만약 시선 강도가 미리-한정된 문턱 값보다 더 크다면, 단계(700)에서 사용자의 포커스 영역 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도가 추측된다. 일 실시예에서, 사용자의 환경 내의 오브젝트 각각은 오브젝트와 상호작용하려는 사용자의 의도를 나타내는 파라미터를 포함할 수 있다. 일 예에서, 시스템(10)은, 오브젝트와 상호작용하려는 사용자의 의도를 결정하면, 오브젝트의 파라미터에 이진 값을 할당할 수 있는데, 여기서 이진 값 1은 오브젝트와 상호작용하려는 사용자의 의도를 지시한다. 단계(700)에서, 사용자의 의도는 사용자 프로파일 데이터베이스(470)에 저장된다.

[0106] 다른 예에서, 사용자의 시선의 강도(단계(696)에서 결정된)는 또한 오브젝트를 보는 동안 사용자에 의해 수행되고 있는 활동의 타입을 결정할 수 있다. 예컨대, 사용자의 시선의 강도로부터 유도된 단속이동 시퀀스의 길이는, 사용자가 오브젝트를 조사/브라우징하고 있는지, 오브젝트에 의해 디스플레이된 텍스트의 한 줄을 읽고 있는지, 오브젝트에 관해 생각하고 있는지 또는 오브젝트를 강렬하게 바라보고 있는지를 결정하기 위하여 사용될 수 있다. 일 실시예에서, 오브젝트에 대한 사용자 시선의 강도에 의해 결정되는 것과 같은 사용자에 의해 수행되고 있는 활동의 타입은 또한 오브젝트와 상호작용하려는 사용자의 의도를 추측하기 위해서도 사용될 수 있다.

[0107] 단계(698)에서, 만약 사용자 시선의 강도가 미리-한정된 문턱보다 더 크지 않다고 결정되면, 일 실시예에서, 도 14b에 기술된 프로세스의 단계들(720-724)이 수행된다.

[0108] 도 14a는 사용자의 포커스 영역 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정하기 위한 프로세스의 다른 실시예를 기술하는 흐름도이다. 도 14a에 기술된 프로세스는, 사용자의 눈 움직임 패턴이 사용자가 사용자의 포커스 영역 내의 어떤 특별한 오브젝트도 보고 있지 않음을 나타내는 때(예컨대, 도 14의 단계(692)) 수행된다. 예컨대, HMD 장치를 착용한 사용자가 정상적인 시력보다 낮은 시력을 가지며 그의 또는 그녀의 환경에 있는 오브젝트나 사람을 매우 명확하게 볼 수 없는 예시적인 상황을 고려해보자. 일 접근방식에서, 하나 이상의 오브젝트나 사람과 상호작용하려는 사용자의 의도는 사용자와 관련된 사용자-특정 정보를 액세스함으로써 자동으로 추측될 수 있다. 사용자의 의도를 자동으로 결정하기 위한 프로세스는 아래의 단계들(710-716)로 기술된다.

[0109] 단계(710)에서, 사용자와 관련된 사용자-특정 정보는 사용자 프로파일 데이터베이스(470)로부터 액세스된다. 사용자-특정 정보에는 사용자의 표현된 선호도, 사용자의 친구 리스트, 사용자의 선호된 활동, 사용자의 소셜 그룹, 사용자의 현재 위치, 사용자의 과거의 오브젝트와 상호작용하려는 의도, 사용자의 리마인더와 같은 사용자와 관련된 정보, 및 사용자의 사진, 이미지 및 기록 비디오와 같은 그 외 다른 사용자가 만든 컨텐트가 포함될 수 있다.

[0110] 단계(712)에서, 사용자의 포커스 영역 내의 어떤 오브젝트나 사람이 사용자-특정 정보에 상관되는지 여부가 결정된다. 예컨대, 사용자가 파티에서 그의 또는 그녀의 친구들을 발견하려고 시도하고 있는 예시적인 상황에서, 사용자의 포커스 영역 내에 나타나는 사람들은, 사람들 중 한 명 이상이 사용자의 친구들 중 임의의 친구와 대응하는지 여부를 결정하기 위하여 사용자와 관련된 사용자-특정 정보와 상관될 수 있다. 예컨대, 사용자의 포커스 영역 내의 한 명 이상의 사람이 사용자의 친구들 중 임의의 친구와 대응하는지 여부를 결정하기 위하여, 사용자의 포커스 영역 내의 사람들의 이미지를 사용자 프로파일 데이터베이스(470) 내에 저장되어 있는 사용자의 친구들의 시각적인 이미지와 상관시키는 얼굴 인식 기술이 이용될 수 있다.

[0111] 단계(712)에서, 만약 사용자의 포커스 영역 내의 오브젝트나 사람 중 어느 것도 사용자-특정 정보와 상관되지 않다고 결정된다면, 단계(704)에서, 프로세스는 도 8의 단계(608)에서 사용자의 뷰 필드를 추적하는 과정으로

되돌아간다. 일 예에서, 만약 사용자의 포커스 영역 내의 오브젝트나 사람 중 어느 것도 사용자-특정 정보와 상관되지 않다고 결정된다면, "당신의 포커스 영역 내에는 현재 흥미로운 오브젝트나 사람이 없습니다"와 같은 가상 텍스트 메시지가, 사용자의 HMD를 통해, 사용자에게 디스플레이될 수 있다.

[0112] 단계(714)에서, 하나 이상의 오브젝트나 사람은 상관 관계(correlation)에 기초하여 식별된다. 예컨대, 사람은 사용자의 친구들, 준 도우 및 샐리 도우로서 식별될 수 있다. 다른 실시예에서, 사용자의 열쇠들이 장면 내에 위치된 것으로서 식별될 수 있다. 단계(716)에서, 식별된 오브젝트나 사람과 상호작용하려는 사용자의 의도는 상관 관계에 기초하여 자동 추측된다. 위에서 논의된 바와 같이, 일 예에서, 시스템은 오브젝트나 사람과 상호작용하려는 사용자의 의도를 결정하면 하나 이상의 식별된 오브젝트나 사람에 의해 대표되는 파라미터에 이진 값을 할당한다. 단계(716)에서, 하나 이상의 식별된 오브젝트와 상호작용하려는 사용자의 의도는 사용자 프로파일 데이터베이스(470)에 저장된다. 단계(716)는 사용자 특정 인자(factor) 또는 외부 인자에 기초하여 의도를 자동 추측하는 과정을 포함할 수 있다. 이러한 인자들에는 사용자의 캘린더 또는 스케줄 정보, 연락처 데이터베이스 또는 소셜 그래프에 저장된 사용자의 친구 정보, 지리적인 위치 정보, 일시 등이 포함될 수 있지만, 이에만 제한되는 것은 아니다. 예컨대, 사용자 특정 정보는, 월요일에 사용자가 회의에 참석하도록 예정되어 있음을 지시할 수 있으며 사용자의 랩탑이 뷰 필드 내에 있다. 도 15c에서 아래에 제공되는 설명에 따라, 시스템은 사용자가 회의에 랩탑을 가지고 갈 것을 기억하도록 유도하기 위하여 랩탑을 하이라이트할 수 있다.

[0113] 도 14b는 하나 이상의 오브젝트와 상호작용하려는 사용자의 표현된 의도에 기초하여 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 결정하기 위한 프로세스의 다른 실시예를 기술하는 흐름도이다. 일 실시예에서, 예컨대 사용자가 사용자의 포커스 영역 내의 하나 이상의 오브젝트를 보고 있지만 사용자 시선의 강도는 충분히 길지 않을 때(즉, 도 14의 단계(698)에서 논의된 바와 같이, 미리-한정된 문턱보다 크지 않을 때), 도 14b에 기술된 프로세스가 수행된다. 단계(720)에서, 사용자는 사용자의 포커스 영역 내의 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도를 특정하거나 표현하도록 프롬프트된다. 일 예에서, 시스템은 하나 이상의 오브젝트와 상호작용하려는 의도를 특정하도록 사용자에게 프롬프트하기 위하여 사용자의 HMD 장치를 통해 사용자에게 가상 텍스트를 디스플레이할 수 있다. 예컨대, 사용자는 음성 입력, 키보드 기입, 터치 또는 제스처와 같은 사용자의 물리적인 행동을 통해 의도를 특정하도록 프롬프트될 수 있다. 음성 입력은 말해진 단어, 회파람, 외침 및 그 외 다른 발화와 같이 사용자에 의해 발화된 음성 명령일 수 있다. 예컨대, 사용자가 파티에서 그의 또는 그녀의 친구들을 발견하려고 시도하는 예시적인 상황에서, 사용자는 "내 친구들 찾게 도와줘!"와 같은 명령을 발화할 수 있다.

[0114] 손뼉을 치는 것과 같은 비-음성적 음향도 또한 사용자로부터 수신될 수 있다. 일 예에서, 음성 입력 또는 비-음성적 음향은 도 3에 도시된 헤드 마운트 디스플레이 내의 마이크로폰(110)에 의해 또는 도 5에 도시된 캡쳐 장치(423) 내의 마이크로폰(430)에 의해 캡쳐될 수 있다. 사용자로부터의 제스처는 미소, 웃음 또는 환호와 같이 사용자에 의해 수행된 얼굴 표정을 포함할 수 있다. 일 예에서, 제스처는 허브 컴퓨팅 장치(12) 내의 인식기 엔진(454)에 의해 캡쳐될 수 있다.

[0115] 단계(722)에서, 하나 이상의 오브젝트와 상호작용하려는 의도가 사용자에 의해 수신되었는지 여부가 결정된다. 만약 사용자의 의도가 수신되지 않았다면, 단계(704)에서, 프로세스는 도 8의 단계(608)에서 사용자의 뷰 필드를 추적하는 과정으로 되돌아간다. 만약 사용자의 의도가 수신되었다면(물리적인 행동을 통해), 하나 이상의 오브젝트와 상호작용하려는 사용자의 의도는 결정되고 사용자의 특정된 또는 표현된 의도는 사용자 프로파일 데이터베이스(700) 내에 저장된다.

[0116] 도 15a는 하나 이상의 오브젝트와 상호작용하려는 사용자의 추측된 또는 표현된 의도를 결정하는 것에 기초하여 최적 이미지를 생성하고 사용자에게 이 최적 이미지를 디스플레이하기 위한 프로세스의 일 실시예를 기술한 흐름도이다. 일 실시예에서, 최적 이미지를 생성하는 과정은 사용자가 상호작용하려고 의도하는 사용자의 포커스 영역 내의 오브젝트의 외관을 시각적으로 강화시키는 과정 및 사용자가 상호작용하려고 의도하지 않는 사용자의 포커스 영역 외부에 있지만 사용자의 뷰 필드 내에 있는 오브젝트의 외관을 약화시키는 과정을 포함한다. 예컨대, 도 15a의 프로세스는 도 8의 단계(614)의 예시적인 일 구현예이다. 도 15a의 단계들은 소프트웨어의 제어 하에 허브 컴퓨팅 장치(12) 또는 프로세싱 유닛(4)에 의해 수행될 수 있다.

[0117] 단계(730)에서, 사용자의 포커스 영역 외부에 있지만 사용자의 뷰 필드 내에 있는 오브젝트의 외관이 약화된다. 일 실시예에서, HMD 장치(2) 내의 불투명 필터(114)는, 사용자의 포커스 영역 외부에 있는 오브젝트의 외관을 약화하기 위하여, 사용자의 포커스 영역 외부에 있는 오브젝트를 차단하거나 어둡게 만들기 위하여 이용된다. 따라서, 사용자가 흥미있어 하지 않는 오브젝트를 포함하는 현실-세계 장면의 부분은 불투명 필터(114)에 의해

차단되어 사용자의 눈에 도달되지 않을 수 있고, 이에 따라 사용자가 사용자의 포커스 영역 내에서 상호작용하려고 의도하는 오브젝트는 사용자가 명확하게 볼 수 있게 된다.

[0118] 단계(732)에서, 사용자가 사용자의 포커스 영역 내에서 상호작용하려고 의도하는 오브젝트는 시각적으로 강화된다. 일 실시예에서, HMD 장치(2) 내의 마이크로 디스플레이 조립체(120)는 사용자의 포커스 영역 내의 오브젝트를 시각적으로 강화시키기 위하여 이용된다. 한가지 접근방식에서, 오브젝트는 오브젝트의 에지를 하이라이트함으로써 시각적으로 강화된다. 다른 접근방식에서, 오브젝트는 오브젝트가 위치되는 영역 내에 가상 박스, 원 또는 시각적인 지시자를 디스플레이함으로써 시각적으로 강화된다. 덧붙여서, 컬러도 또한 오브젝트를 시각적으로 강화시키기 위해 사용될 수 있다.

[0119] 단계(734)에서, 하나 이상의 오브젝트와 관련된 증강 컨텐트가 회수된다. 일 예에서, 증강 컨텐트는 사용자 프로파일 데이터베이스(470)로부터 회수된 사용자-특정 정보를 포함할 수 있다. 다른 예에서, 증강 컨텐트는 사용자의 소셜 네트워킹 사이트, 주소록, 이메일 데이터, 인스턴트 메시징 데이터, 사용자 프로파일 또는 인터넷 상의 그 외 다른 소스와 같은 하나 이상의 데이터 소스로부터 실시간으로 회수되는 사용자-특정 정보를 포함할 수 있다.

[0120] 단계(736)에서, 식별된 오브젝트에 관련된 오디오 컨텐트가 추출된다. 단계(736)는 선택적이다. 예컨대, 만약 사용자가 사용자의 거실 내의 벽시계를 바라보고 있고 사용자가 벽시계와 상호작용하려고 의도한다고 결정되면, 사용자는 시간에 관한 오디오 정보를 들을 수 있다. 또는 예컨대, 만약 사용자가 그의 또는 그녀의 친구들과 파티에 있고, 사용자가 파티에서 특정 그룹의 친구들과 상호작용하려고 의도한다고 결정되면, 사용자는 특정 그룹의 친구들로부터의 음성만을 들을 수 있다.

[0121] 단계(738)에서, 증강 컨텐트는 사용자의 포커스 영역 내의 하나 이상의 오브젝트 위에 또는 옆에 투사된다. 일 예에서, 증강 컨텐트는 사용자에게 디스플레이되는 하나 이상의 가상 오브젝트 또는 가상 텍스트를 포함하는 가상 이미지이다. 다른 예에서, 증강 컨텐트는 하나 이상의 선택 사항을 가진 메뉴와 같은 가상 오브젝트를 포함할 수 있다.

[0122] 단계(740)에서, 최적 이미지가 헤드 마운트 디스플레이 장치(2)를 통해 사용자에게 디스플레이된다. 최적 이미지는 사용자의 포커스 영역 내의 오브젝트의 강화된 외관, 사용자의 포커스 영역 외부의 오브젝트의 약화된 외관을 포함하며 또한 선택적으로 하나 이상의 오브젝트와 관련된 증강 컨텐트를 포함할 수 있다. 최적 이미지를 생성하고 이 최적 이미지를 사용자에게 디스플레이하는 방식의 예시들이 도 16a 내지 도 16c, 도 17a 내지 도 17d 및 도 18a 내지 도 18c에 도시된다.

[0123] 일 예에서, 프로세싱 유닛(4)은 증강 컨텐트의 마이크로 디스플레이(120) 상에서의 디스플레이를 위해 컨트롤허브(136)의 디스플레이 드라이버(220)에 인스트럭션을 송신한다. 이후 웬즈 시스템(122)은 마이크로 디스플레이(120)로부터 수신된 최적 이미지를 사용자가 보도록 하기 위해 반사면(124) 상에 그리고 사용자의 눈을 향해 또는 광 안내 광학 요소(112) 내로 투사한다. 일 구현예에서, 사용자가 헤드 마운트 디스플레이 장치(예컨대, 광 안내 광학 요소(112))를 통하여 보고 있는 디스플레이는 픽셀들로 분리된다. 최적 이미지는 픽셀들에 의해 타겟 위치에 디스플레이된다. 덧붙여서, 불투명 필터(114)는 비현실적인 시각 효과를 방지하기 위하여 사용될 수 있다. 예컨대, 최적 이미지 내의 배경의 가상 오브젝트가 전경의 가상 오브젝트를 통하여 보여지지 않도록 하기 위하여, 불투명 필터(114)는 광을 광 안내 광학 요소(112) 내의 픽셀들로 수정할 수 있다.

[0124] 도 15b는 최적 이미지 내에서 디스플레이된 증강 컨텐트와 상호작용하려는 사용자의 추측된 또는 표현된 의도를 결정하는 것에 기초하여, 사용자에게 추가적인 증강 컨텐트를 디스플레이하기 위한 프로세스의 일 실시예를 기술하는 흐름도이다. 일 예에서, 도 15a의 단계(740)에서 기술된 바와 같이 최적 이미지를 디스플레이하면, 사용자는 추가적인 증강 정보를 제공받을 수 있다. 예컨대, 사용자가 사용자의 거실 내의 테이블 상에 놓여진 사과 바구니를 바라보고 있으며 사용자가 사과 바구니와 상호작용하려고 의도한다는 것이 결정된 상황을 고려해보자. 사과 바구니의 외관이 시각적으로 증가된 최적 이미지가 사용자에게 디스플레이된다. 최적 이미지는 또한 사용자에게 선택 사항들을 가진 메뉴와 같은 가상 오브젝트를 포함하는 증강 컨텐트를 투사한다는 것을 가정하자.

[0125] 단계(756)에서, 메뉴에 디스플레이되어 있는 하나의 선택 사항과 상호작용하려는 사용자의 의도가 결정된다. 사용자의 의도는 도 14에서 기술된 프로세스에 의해 논의된 바와 같이 결정될 수 있다. 단계(760)에서, 사용자의 의도에 기초하여, 메뉴의 선택 사항의 사용자의 선택이 수신된다. 단계(762)에서, 메뉴 선택 사항과 관련된 증강 컨텐트 및 선택적인 오디오 정보가 회수된다. 단계(762)에서, 증강 컨텐트는 최적 이미지 내에서 시각적으로 강화된 오브젝트 위에 또는 옆에 투사된다. 메뉴 선택 사항을 포함하는 최적 이미지를 디스플레이하고 메모 선

택 사항과 상호작용하려는 사용자의 의도를 결정하는 방식의 예시가 도 17a 내지 17d에 도시되어 있다.

[0126] 개시된 기술의 다른 실시예에서, 사용자의 환경 내의 시각화된 정보의 사용자의 이해는 사용자의 환경 내의 특정 오브젝트를 자동으로 강화함으로써 증가된다. 도 15c는 사용자의 환경 내의 시각화된 정보의 사용자의 이해가 증가되는 프로세스를 기술한다. 단계(770)에서, 사용자의 환경의 사용자의 이해를 증가시키는 사용자의 포커스 영역 내의 하나 이상의 오브젝트가 식별된다. 일 예에서, 사용자의 이해를 증가시키는 오브젝트는 실시간으로 사용자의 환경에 관한 정보를 수집함으로써 식별될 수 있다. 예컨대, 정보는 사용자의 환경 내의 비상사태 표시, 또는 사용자의 환경 내에 있는 위험한 오브젝트 또는 날카로운 오브젝트의 존재를 포함할 수 있다. 다른 예에서, 사용자의 이해를 증가시키는 오브젝트는 도 14b에 관련하여 위에서 주목한 바와 같이 사용자에 관련된 사용자-특정 정보에 기초하여 식별될 수 있다. 예컨대, 사용자에 관련된 사용자-특정 정보는, 사과 바구니와 같은 오브젝트가 사용자의 거실에서 식별되는 때에, 사용자에게 간식을 먹을 때임을 특정하는 리마인더를 포함할 수 있다.

[0127] 단계(772)에서, 사용자의 이해를 증가시키는 하나 이상의 오브젝트를 포함하는 최적 이미지가 생성된다. 일 실시예에서, 최적 이미지는 사용자의 환경 내에서 사용자의 이해를 증가시키는 하나 이상의 오브젝트의 외관을 시각적으로 강화시킨다. 하나의 접근방식에서, 오브젝트는, 오브젝트의 애지를 하이라이트함으로써, 오브젝트가 위치된 영역에 가상 박스, 원 또는 가상 지시자를 디스플레이함으로써 시각적으로 강화된다.

[0128] 단계(774)에서, 최적 이미지 내의 하나 이상의 시각적으로 강화된 오브젝트와 상호작용하려는 사용자의 의도가 결정될 수 있다. 일 예에서, 사용자의 의도는 하나 이상의 시각적으로 강화된 오브젝트와 상호작용하려는 사용자의 의도를 추측(도 14 또는 도 14a에서 논의된 바와 같이)함으로써 결정될 수 있다. 다른 예에서, 사용자의 의도는 또한 하나 이상의 시각적으로 강화된 오브젝트와 상호작용하려는 사용자의 표현된 의도(도 14b에서 논의된 바와 같이)에 기초하여 결정될 수 있다.

[0129] 도 16a 내지 도 16c는 사용자 상호작용에 기초하는 사용자의 환경 내의 하나 이상의 오브젝트와의 사용자의 상호작용 및 최적 이미지의 생성의 일 실시예를 도시한다. 도 16a는 사용자가 HMD 장치(2)를 사용하여 방(1100) 안의 하나 이상의 오브젝트를 보는 환경을 도시한다. 방(1100)에는 전면 벽(1102), 측벽(1104) 및 바닥(1108), 및 전등(1106), 의자(1107), 벽시계(1118) 및 테이블(1120)과 같은 예시적인 가구가 포함된다. 비디오 디스플레이 스크린(1110)은 본 예에서 벽(1102)에 고정되어 있고, 허브(1116)는 테이블(1120) 상에 놓여져 있다. 예시적인 상황에서, 사용자(1112)는 HMD 장치(2)를 통해 전면 벽(1102)에 배치되어 있는 벽시계(1118)와 같은 오브젝트를 바라본다. 1121은 사용자의 뷔 필드를 나타내고 1122는 사용자의 포커스 영역을 나타낸다.

[0130] 도 16b는, 벽시계 오브젝트(1118)와 상호작용하려는 사용자의 의도가 결정되면, 도 16a의 HMD 장치의 카메라에 의해 생성된 최적 이미지를 도시한다. 일 실시예에서, 사용자의 의도는 도 14에서 기술된 프로세스에 의해 논의된 바와 같이 결정될 수 있다. 도 16b에 도시된 바와 같이, 최적 이미지(1124)는 사용자의 포커스 영역 내의 강화된 외관의 벽시계 오브젝트(1118) 및 사용자의 포커스 영역 외부에 있지만 사용자의 뷔 필드 내에 있는 약화된 외관의 전등(1106), 허브(1116) 및 테이블(1120)을 포함한다. 예시적인 구현예에서, 벽시계 오브젝트(1118)는 그것의 외관을 강화하기 위하여 하이라이트되었다. 오브젝트(1106, 110, 1116 및 1120)을 둘러싸는 점선들은 그들의 약화된 외관을 지시한다. 덧붙여서, 최적 이미지는 벽시계 오브젝트(1118) 옆에 디지털 포맷으로 일시를 보여주는 중강 컨텐트(1126)를 디스플레이한다. 일 예에서 일시에 관한 오디오 정보도 또한 사용자가 들을 수 있다.

[0131] 도 16c는 HMD 장치를 통해 사용자가 볼 수 있는 도 16b의 최적 이미지를 도시한다. 최적 이미지는 HMD 장치(2)의 렌즈들(116 및 118)의 각각에 의하여 각각 제공된다. 열린 영역들(1127 및 1128)은 디스플레이 스크린으로부터의 광이 사용자의 눈으로 입사하는 위치들을 지시한다.

[0132] 도 17a 내지 도 17d는 사용자의 상호작용에 기초하여 사용자의 환경 내의 하나 이상의 오브젝트와의 사용자 상호작용 및 최적 이미지의 생성의 다른 실시예를 도시한다. 도 17a는 사용자(1112)가 HMD 장치(2)를 사용하여 방(1100) 안의 하나 이상의 오브젝트를 바라보는 환경(예컨대, 도 16a에서 도시된 바와 같은)을 도시한다. 도 17a의 예에서, 사용자(1112)는 테이블(1130) 상에 놓여진 사과 바구니(1132)를 바라본다. 1121은 사용자의 뷔 필드를 나타내고 1122는 사용자의 포커스 영역을 나타낸다.

[0133] 도 17b는 사과 바구니(1132)와 상호작용하려는 사용자의 의도가 결정되면 도 17a의 HMD 장치의 카메라에 의해 생성되는 최적 이미지를 도시한다. 일 실시예에서, 사용자의 의도는 도 14에 기술된 프로세스에 의해 논의된 바와 같이 결정될 수 있다. 최적 이미지(1133)는 사용자의 포커스 영역 내에 있는 사과 바구니(1132)의 강화된 외

관 및 사용자의 포커스 영역 외부에 있지만 사용자의 뷔 필드 내부에 있는 테이블(1120)의 약화된 외관을 포함한다. 다른 예에서, 사과 바구니의 외관은 사과 바구니가 사용자의 환경의 사용자의 이해를 증가한다고 결정하는 것에 기초하여 자동으로 강화될 수 있다. 도 15c에 기술된 프로세스에 따라, 사용자의 포커스 영역 내의 사과 바구니(1132)의 강화된 외관을 포함하는 최적 이미지가 생성될 수 있다. 최적 이미지(1133)는 또한 중강 컨텐트(1134)를 디스플레이할 수 있다. 도시된 예에서, 중강 컨텐트(1134)는 사용자(1112)에게 디스플레이되는 선택사항 "애플파이를 만들고 싶습니까?(Do you wish to make an apple pie?)" 또는 "리마인더를 보고 싶습니까?(Do you wish to view your reminder?)"를 가진 메뉴이다.

[0134] 도 17c는 메뉴에서 디스플레이된 선택 사항 중 하나와 상호작용하려는 사용자의 의도가 결정되면 사용자에게 디스플레이되는 다른 최적 이미지를 도시한다. 일 실시예에서, 사용자의 의도는 도 14에 기술된 프로세스에 의해 논의된 바와 같이 결정될 수 있다. 도시된 예에서, 최적 이미지(1133)는, 사용자가 메뉴 선택 사항 "애플파이를 만들고 싶습니까?"와 상호작용하려고 의도한다고 결정된 때에, 애플 파이를 만들기 위한 재료를 포함하는 중강 컨텐트(1134)를 디스플레이한다.

[0135] 도 17d는 메뉴에 디스플레이된 선택 사항들 중 하나와 상호작용하려는 사용자의 의도가 결정되면 사용자에게 디스플레이되는 최적 이미지를 도시한다. 도시된 예에서, 최적 이미지(1133)는, 사용자가 메뉴 선택 사항 "리마인더를 보고 싶습니까?"와 상호작용하려고 의도한다고 결정된 때에, 사용자가 그의 또는 그녀의 간식을 먹을 때임을 사용자에게 알려주는 리마인더를 포함하는 중강 컨텐트(1136)를 디스플레이한다.

[0136] 도 18a 내지 도 18c는 사용자의 상호작용에 기초하여 사용자의 환경 내의 하나 이상의 오브젝트와의 사용자 상호작용 및 최적 이미지의 생성의 다른 실시예를 도시한다. 도 18a는 사용자(1112)가 사용자의 뒤틀에서 열리는 모임에서 다수의 사람을 만나고 있는 환경을 도시한다. 도 18a에서 도시된 예시적인 상황에서, 사용자(1112)는 정상 시력보다 낮은 시력을 가질 수 있으며 HMD 장치(2)를 통해 모임 내의 모든 사람들을 명확하게 볼 수 없을 수 있다. 또는 사용자는 사용자가 그들의 이름을 기억하지 않기 때문에 모임 내에 있는 사람들과 곧바로 상호작용하고 싶어하지 않을 수 있다.

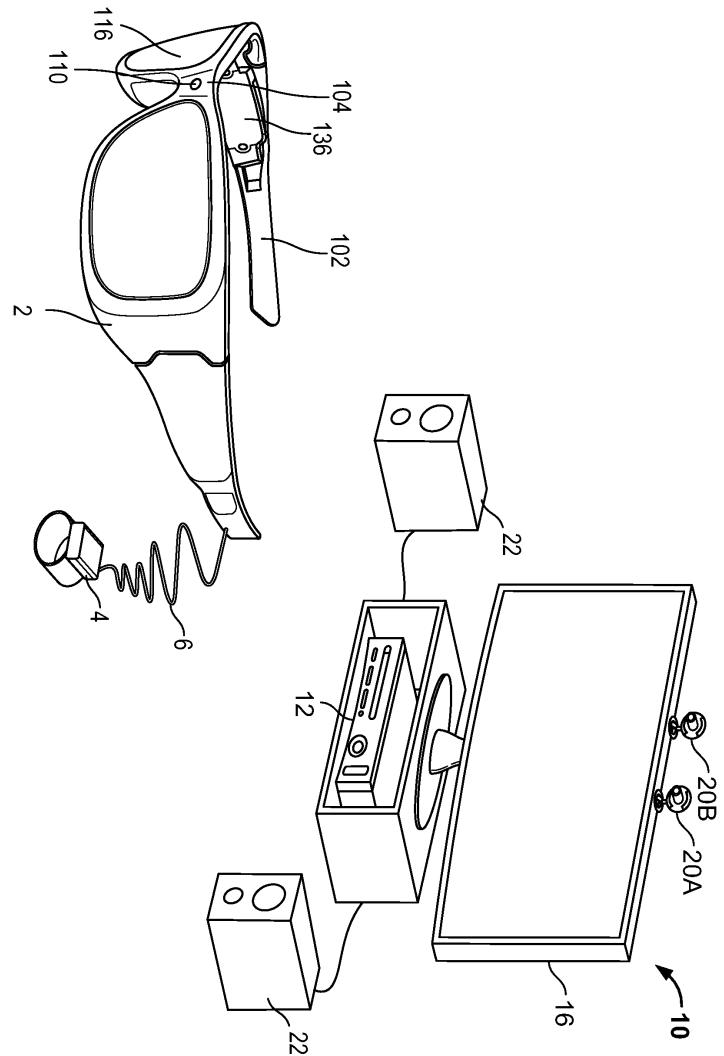
[0137] 도 18b는 모임 내의 한 명 이상의 사람과 상호작용하려는 사용자의 의도를 자동으로 결정함으로써 사용자에게 디스플레이되는 최적 이미지를 도시한다. 한가지 접근 방식에서, 그리고 도 14a에서 논의된 바와 같이, 사용자의 포커스 영역 내의 한 명 이상의 사람들과 상호작용하려는 사용자의 의도는 사용자와 관련된 사용자-특정 정보를 액세스함으로써 자동으로 추측될 수 있다. 다른 접근 방식에서, 그리고 도 14b에서 논의된 바와 같이, 사용자의 포커스 영역 내의 한 명 이상의 사람들과 상호작용하려는 사용자의 의도는 또한 "내 친구 찾는 걸 도와줘!"와 같은 사용자로부터의 음성 명령을 수신함으로써 결정될 수 있다. 최적 이미지(1136)는 사용자의 친구 존 도우(1136) 및 셀리 도우(1138)의 강화된 외관과 사용자의 포커스 영역 외부에 있으나 사용자의 뷔 필드 내에 있는 그 파티 내의 그 외 다른 사람(1138)의 약화된 외관을 포함한다. 최적 이미지(1133)는 또한 각각 "여기가 존 도우!(Here is John Doe!)" 및 "여기가 셀리 도우!(Here is Sally Doe!)"와 같은 가상 텍스트 메시지를 사용하여 사용자의 친구 존 도우(1136) 및 셀리 도우(1138)를 식별하는 중강 컨텐트를 디스플레이한다. 사용자가 상호작용하고자 의도할 수 있는 개인들의 결정은 사용자의 소셜 그래프(사용자-특정 정보 내에 저장된)의 분석을 통해, 또는 동시에 사용자를 찾고 있을 수 있는 사용자(그렇지만 결정된)의 친구들에 대한 참조에 의하여 수행될 수 있다.

[0138] 도 18c는 HMD 장치를 통해 사용자가 보는 도 18b의 최적 이미지를 도시한다. 최적 이미지는 HMD 장치(2)의 렌즈들(116 및 118)의 각각에 의하여 각각 제공된다. 개방된 영역(1127 및 1128)은 비디오 디스플레이 스크린으로부터의 광이 사용자의 눈으로 입사하는 위치를 지시한다.

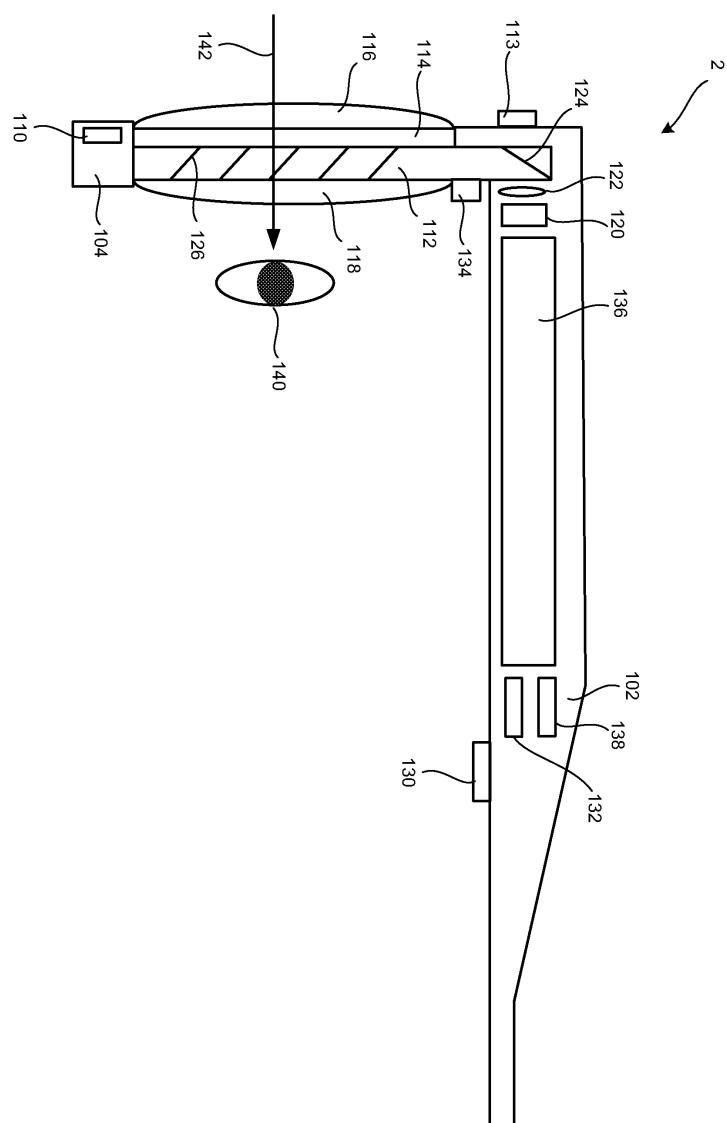
[0139] 비록 발명 주체가 구조적 특징 및/또는 방법론적 행동에 특정적인 언어도 기술되었으나, 첨부된 청구범위에서 한정된 발명 주체는 위에 기술된 특정 특징이나 행동으로 반드시 한정되어야 할 필요가 없다는 점이 이해되어야 한다. 반대로, 위에 기술된 특정 특징과 행동은 청구범위를 구현하는 예시적인 형태로서 개시된다. 본 발명의 범위는 본 명세서에 첨부된 청구범위에 의해서 한정된다.

## 도면

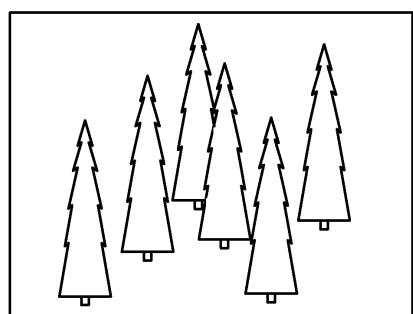
## 도면1



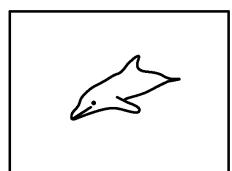
도면2



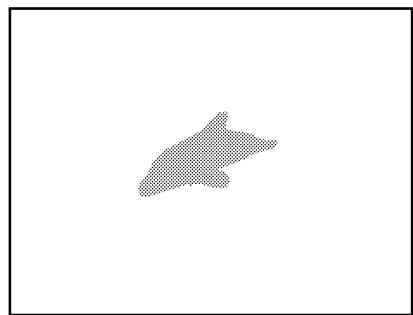
도면2a



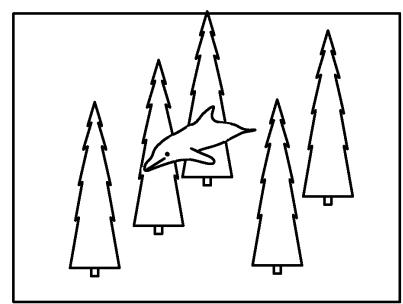
도면2b



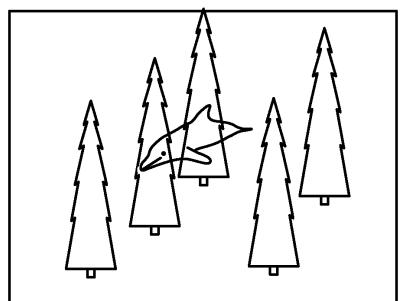
도면2c



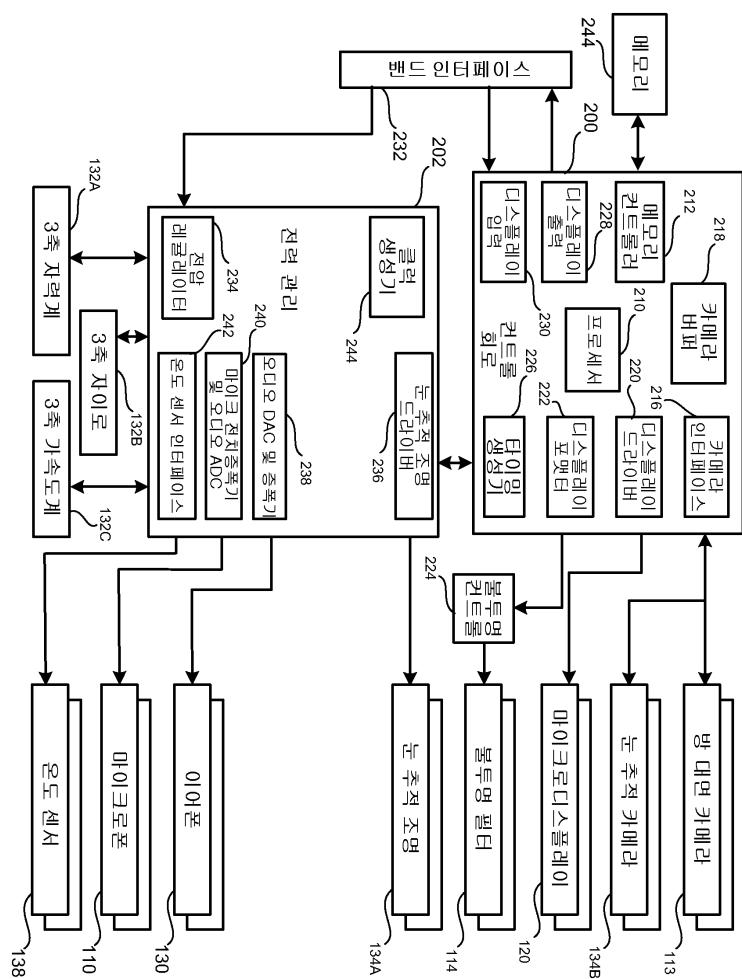
도면2d



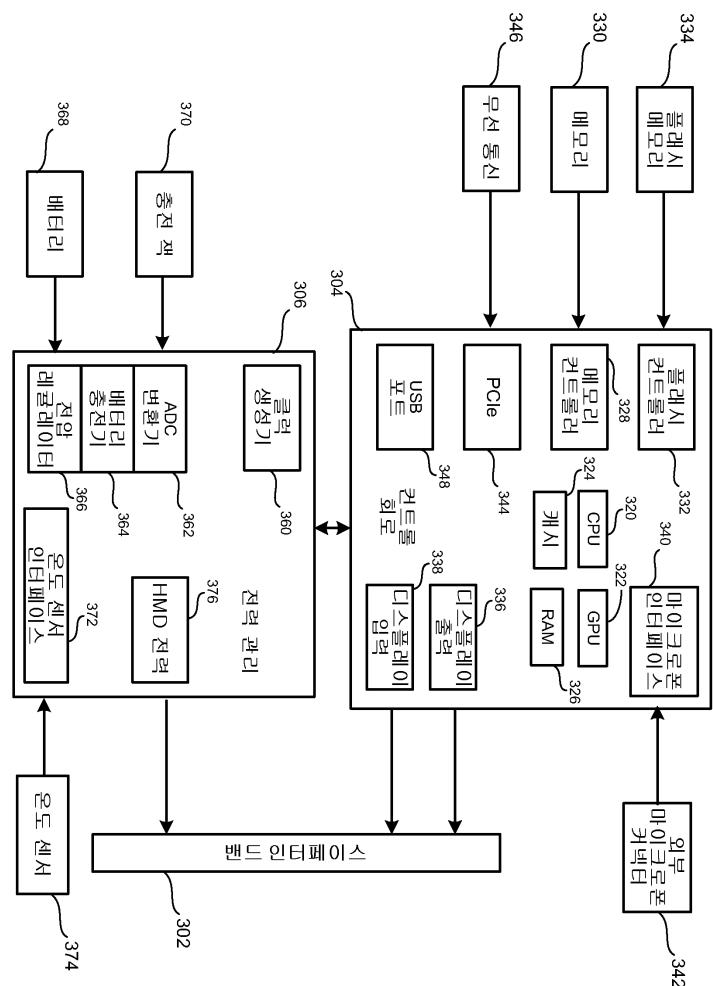
도면2e



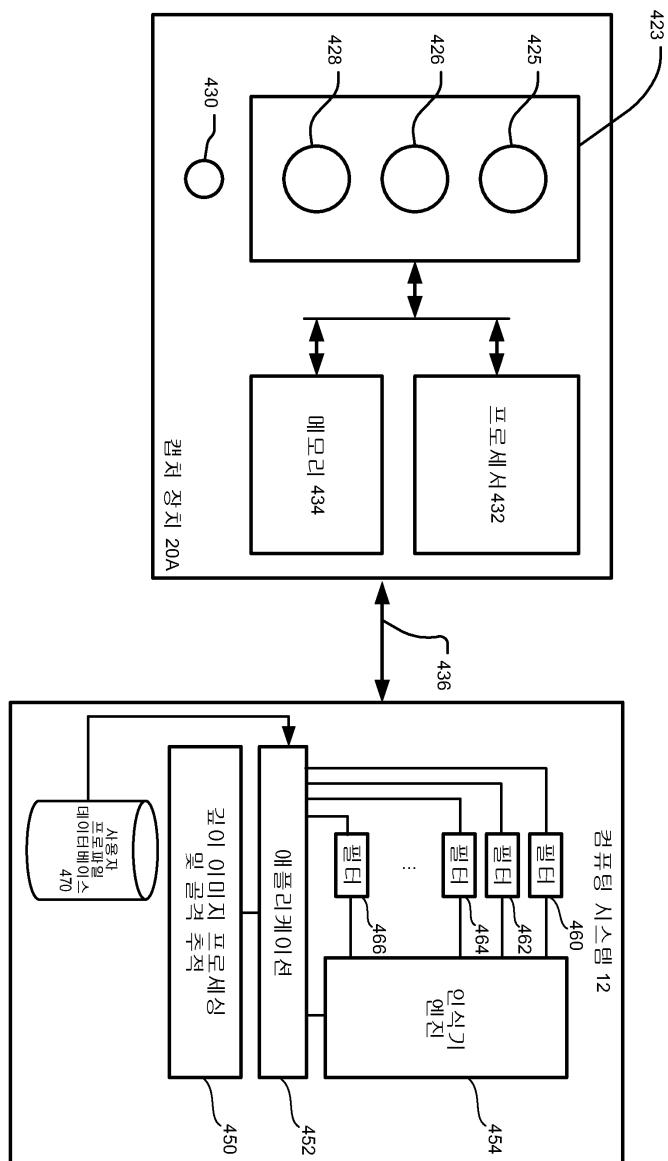
도면3



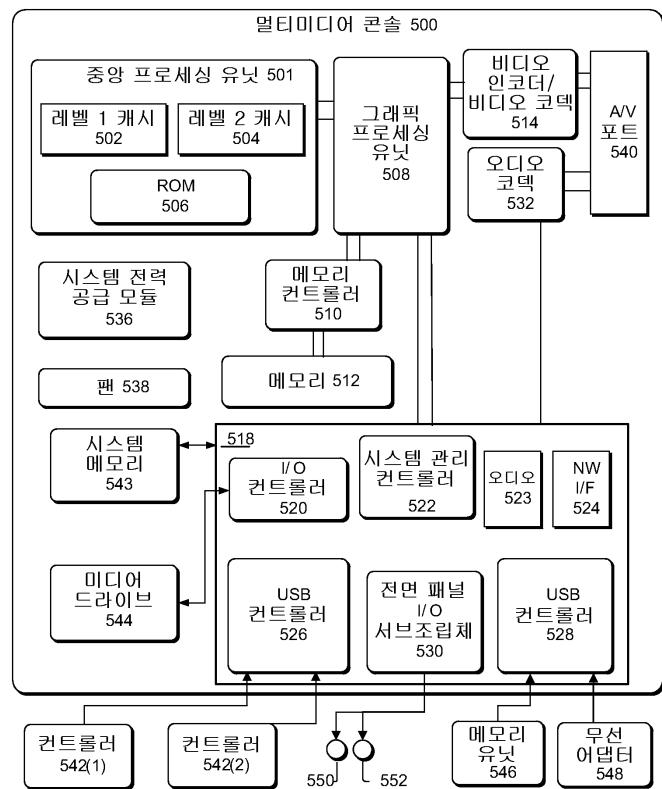
## 도면4



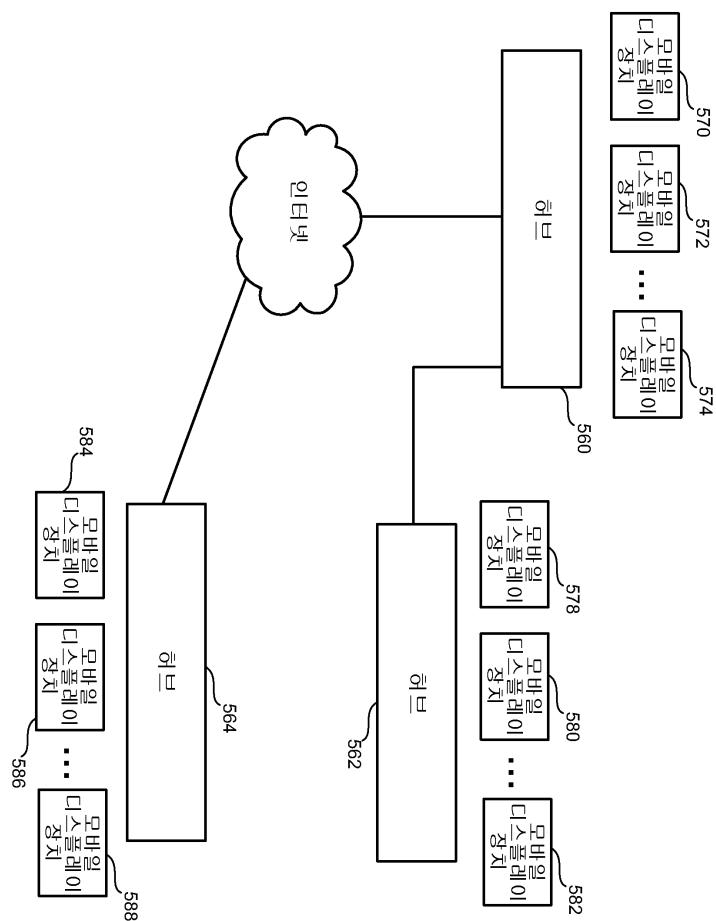
도면5



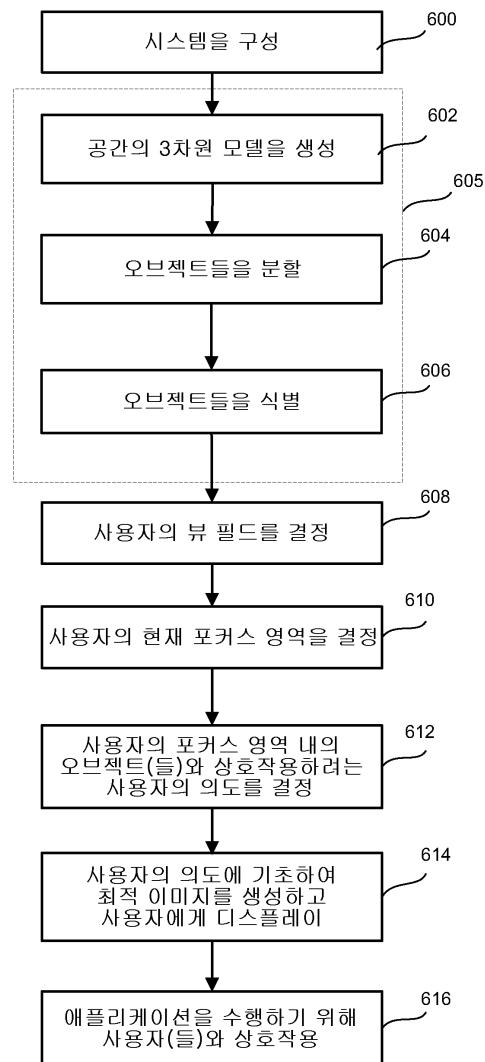
## 도면6



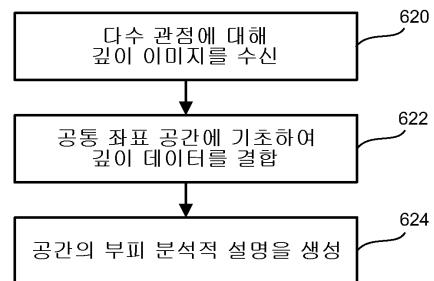
## 도면7



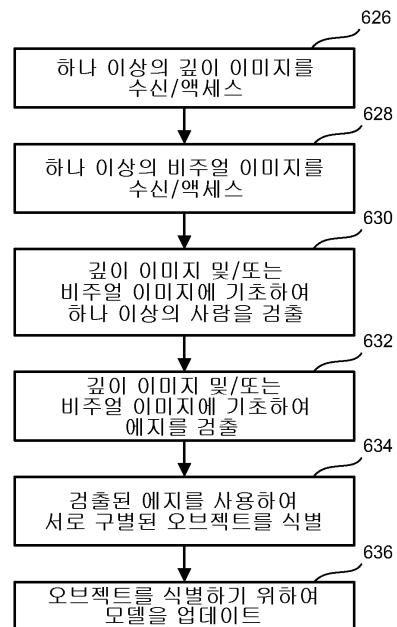
## 도면8



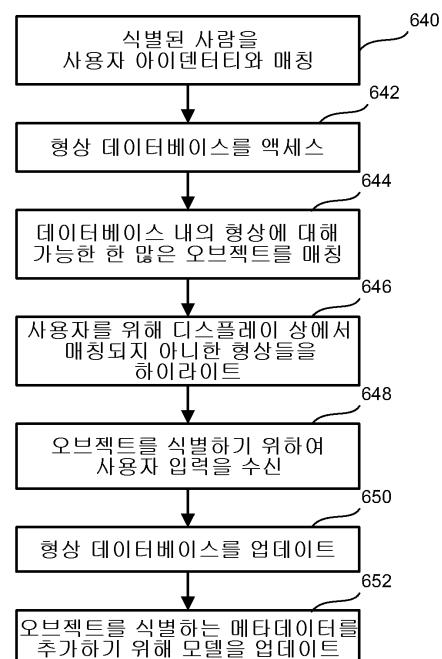
## 도면9



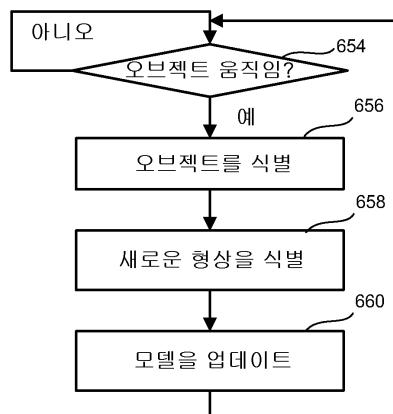
## 도면10



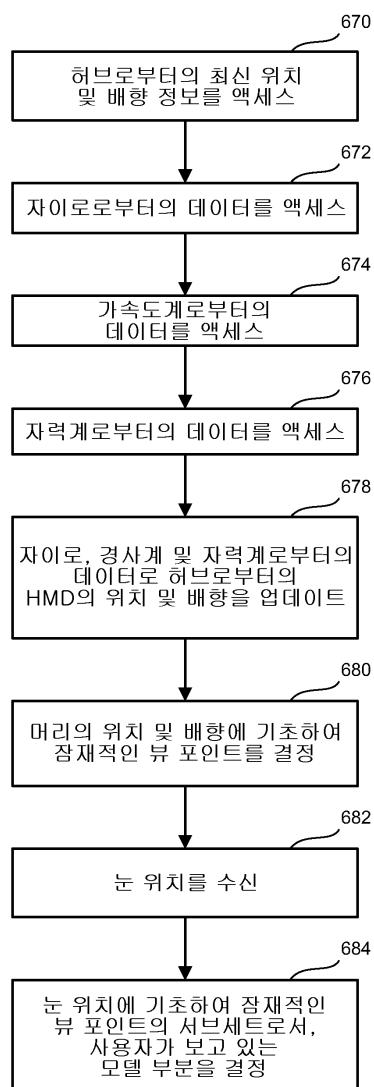
## 도면11a



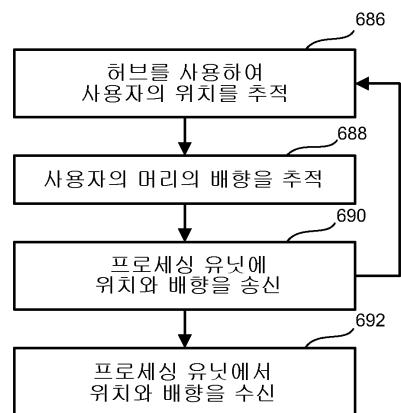
## 도면11b



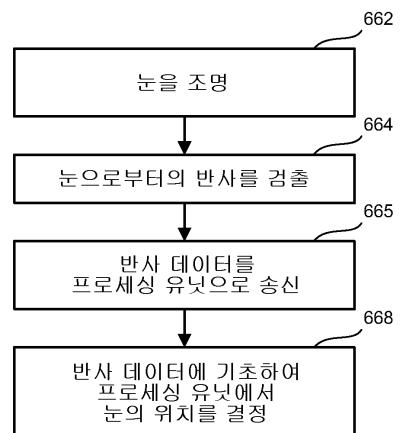
## 도면12



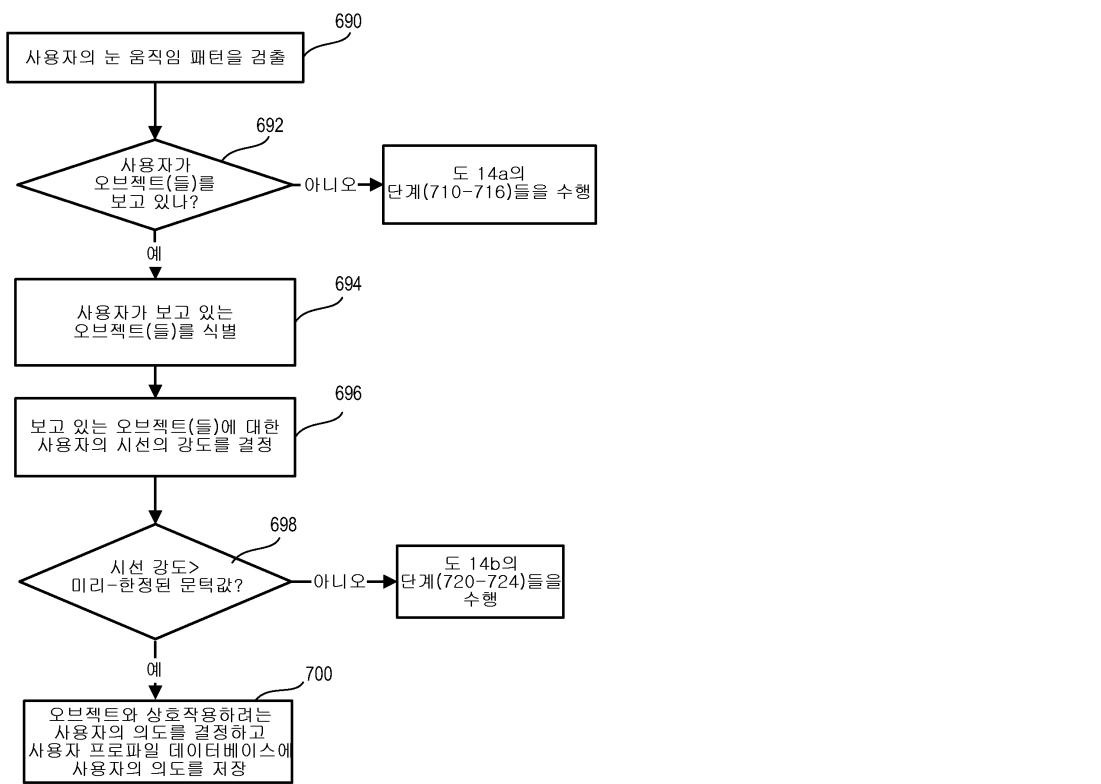
## 도면13a



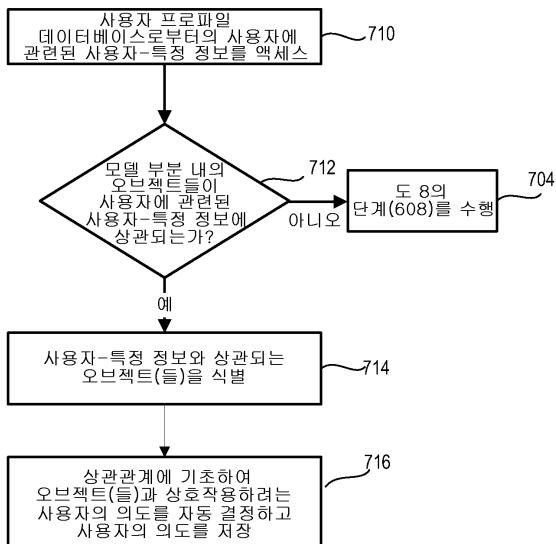
## 도면13b



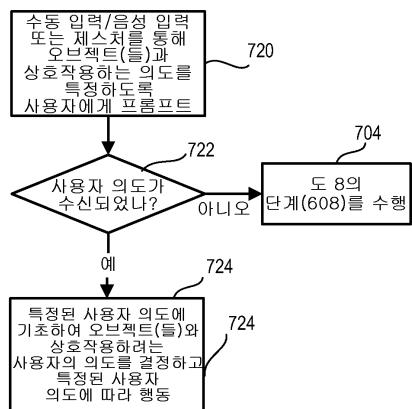
## 도면14



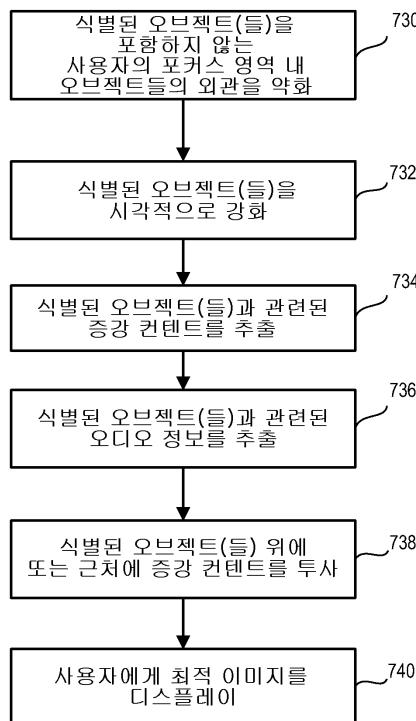
## 도면14a



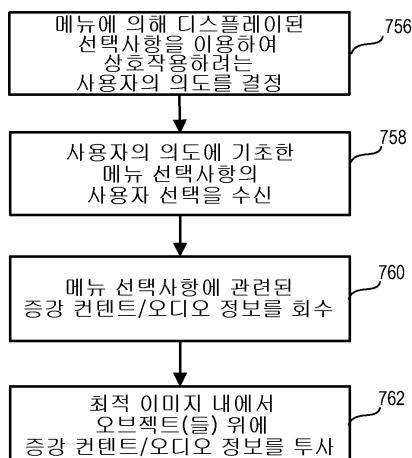
## 도면14b



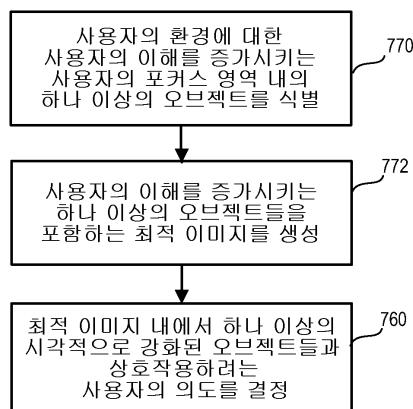
## 도면15a



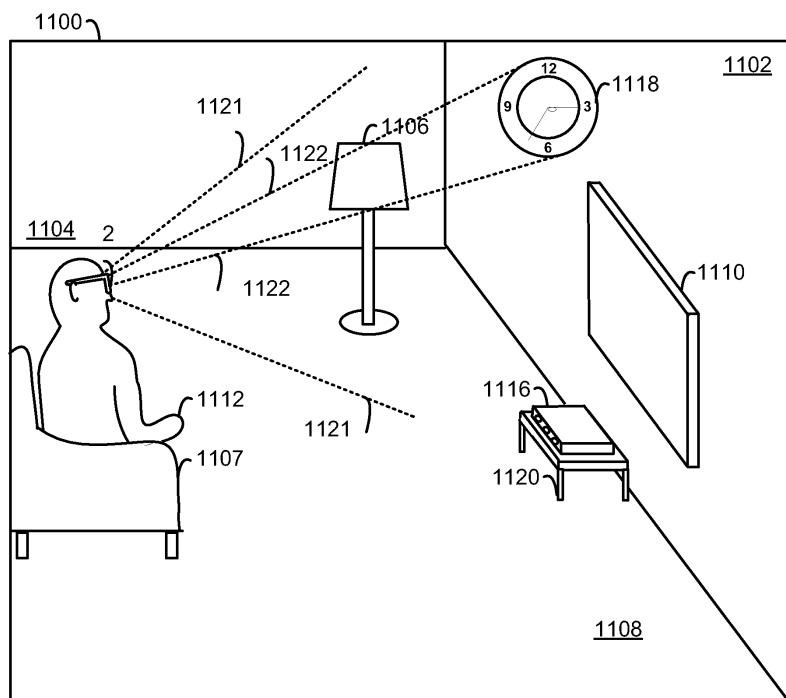
## 도면15b



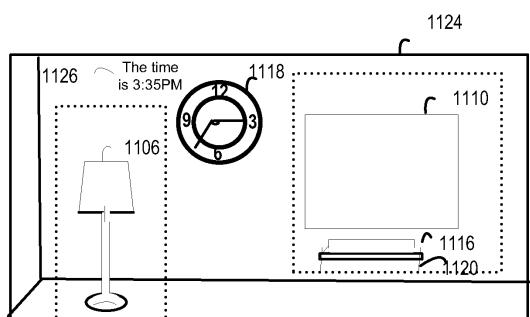
## 도면15c



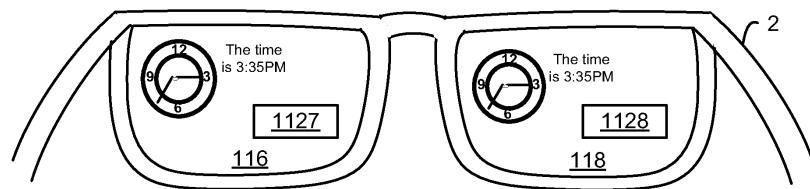
## 도면16a



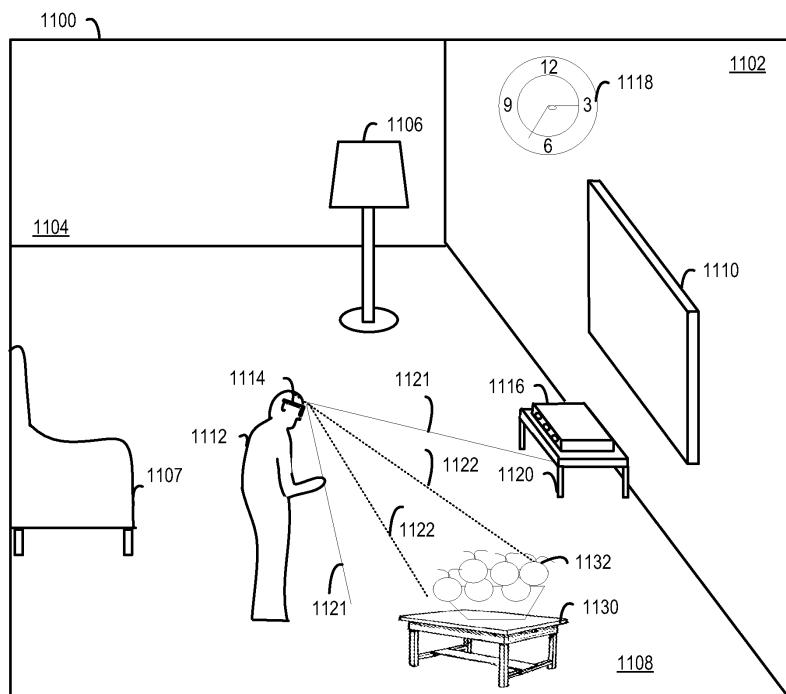
## 도면16b



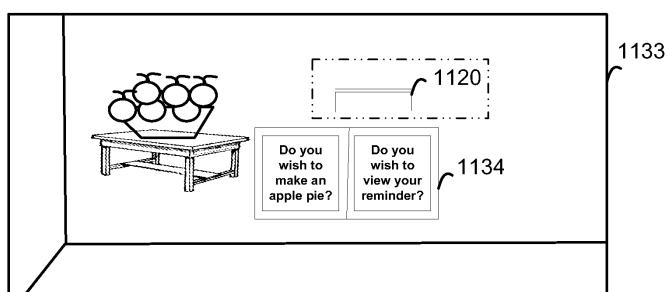
도면16c



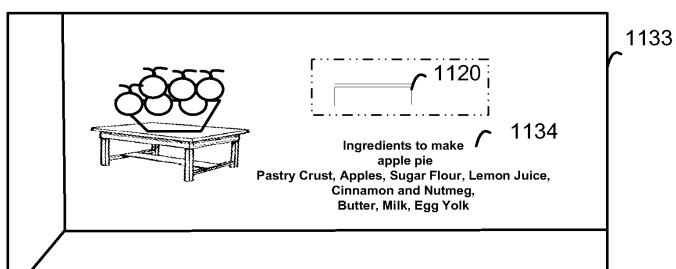
도면17a



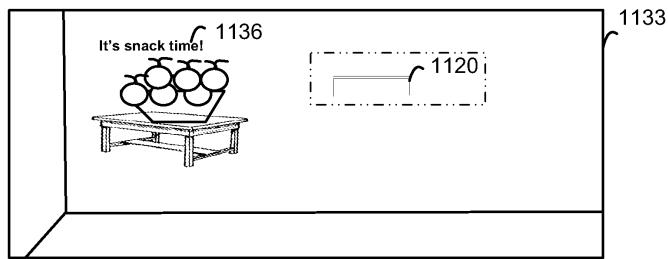
도면17b



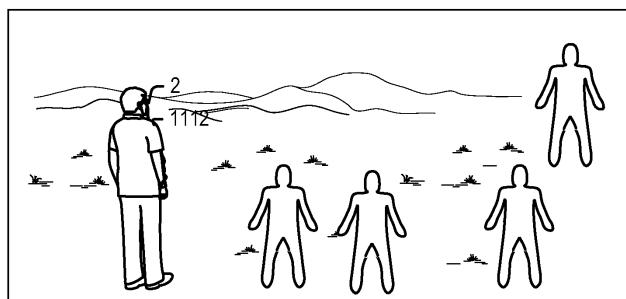
도면17c



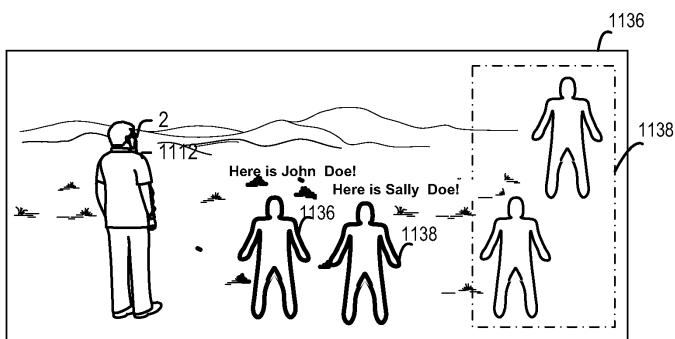
도면17d



도면18a



도면18b



도면18c

