



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2018년07월24일  
 (11) 등록번호 10-1881620  
 (24) 등록일자 2018년07월18일

(51) 국제특허분류(Int. Cl.)  
 A63F 13/213 (2014.01) A63F 13/525 (2014.01)  
 (21) 출원번호 10-2013-7020253  
 (22) 출원일자(국제) 2013년01월09일  
 심사청구일자 2016년11월28일  
 (85) 번역문제출일자 2013년07월30일  
 (65) 공개번호 10-2014-0008510  
 (43) 공개일자 2014년01월21일  
 (86) 국제출원번호 PCT/US2012/020687  
 (87) 국제공개번호 WO 2012/106070  
 국제공개일자 2012년08월09일  
 (30) 우선권주장  
 13/017,729 2011년01월31일 미국(US)  
 (56) 선행기술조사문헌  
 KR1020020011851 A  
 KR1020070119018 A  
 US20020024517 A1  
 US20090231425 A1

(73) 특허권자  
**마이크로소프트 테크놀로지 라이선싱, 엘엘씨**  
 미국 워싱턴주 (우편번호 : 98052) 레드몬드 원  
 마이크로소프트 웨이  
 (72) 발명자  
**이자디 샤람**  
 미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
 소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마  
 이크로소프트 코포레이션  
**몰리니우스 데이비드**  
 미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
 소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마  
 이크로소프트 코포레이션  
 (뒷면에 계속)  
 (74) 대리인  
**김태홍**

전체 청구항 수 : 총 20 항

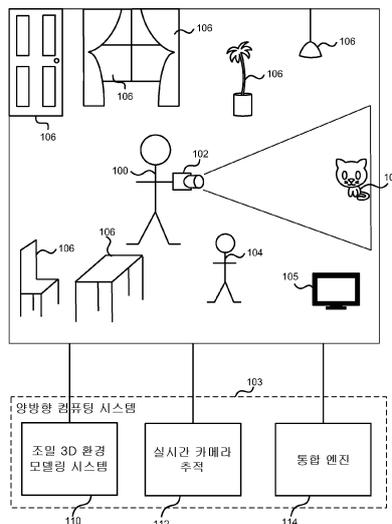
심사관 : 신유식

(54) 발명의 명칭 **게임플레이에서의 3차원 환경 모델 사용**

**(57) 요약**

게임플레이에서 3D 환경 모델의 사용이 설명된다. 일 실시형태에서, 이동 깊이 카메라가 돌아다니는 때에 이동 깊이가 카메라가 깊이 이미지의 시리즈를 캡처하는데 사용되고, 이 깊이가 이미지의 시리즈로부터 환경의 조밀 3D 모델이 생성된다. 이 조밀 3D 모델은 게임과 같은 양방향 애플리케이션 내에 통합된다. 이동 깊이가 카메라는 그 후, 일부 예에서 게임플레이인 양방향 페이지를 위해 정지 위치에 배치되고, 시스템은 카메라에 의해 캡처된 깊이 이미지의 제2 시리즈로부터 환경의 일부 내의 사용자의 움직임을 검출한다. 이 움직임은 게임과 같은 양방향 애플리케이션으로 사용자 입력을 제공한다. 다른 실시형태에서, 3D 모델 내 물체의 자동 인식 및 식별이 수행될 수 있고, 그러면 이들 식별된 물체는 양방향 애플리케이션이 동작하는 방식을 변경한다.

**대표도 - 도1**



(72) 발명자

**힐리게스 오트마**

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

**킴 데이비드**

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

**쇼튼 제이미 다니엘 요셉**

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

**폴리 퍼쉬미트**

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

**피츠기븐 앤드류**

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

**하지스 스티븐 에드워드**

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

**버틀러 데이비드 알렉산더**

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

## 명세서

### 청구범위

#### 청구항 1

양방향(interactive) 컴퓨팅 시스템에 있어서,

움직이는 때에 이동 깊이 카메라(mobile depth camera)로부터 깊이 이미지의 제1 시리즈를, 그리고 정지한 때에 상기 이동 깊이 카메라로부터 깊이 이미지의 제2 시리즈를 수신하기 위한 입력과,

움직이는 때에 상기 이동 깊이 카메라의 위치를 추적하기 위한 실시간 카메라 추적 엔진과,

깊이 이미지의 상기 시리즈 중 어느 한 시리즈와 상기 실시간 카메라 추적 엔진으로부터의 위치 데이터로부터 환경의 조밀(dense) 3D 모델을 생성하도록 구성된 조밀 3D 환경 모델링 시스템과,

상기 조밀 3D 모델의 적어도 일부를 애플리케이션으로 통합하고, 깊이 이미지의 상기 제2 시리즈를 이용하여 상기 환경의 일부에서 실시간 움직임을 추적함으로써 상기 애플리케이션에 사용자 입력을 제공하도록 구성된 통합(integration) 엔진

을 포함하고,

상기 통합 엔진은 또한,

상기 조밀 3D 모델을 분석하여 상기 모델 내에서 물체(object)를 식별하고,

물체가 식별되는 경우, 물체 유형 정보를 상기 애플리케이션에 입력하며,

입력된 물체 유형 정보에 기초하여 상기 애플리케이션의 동작을 조정하도록

구성되며,

상기 물체는 능동 물체(active object)를 포함하고, 상기 통합 엔진은 또한 상기 능동 물체의 출력을 조정하도록 구성되거나,

상기 물체는 디스플레이를 포함하고, 상기 통합 엔진은 또한, 상기 디스플레이와 상기 이동 깊이 카메라의 검출된 상대 위치에 기초하여 상기 디스플레이에 대한 물체의 위치를 자동으로 교정하도록 구성되거나,

상기 통합 엔진은 또한, 다른 검출된 물체의 출력을 조정하도록 구성되거나,

상기 물체는 다른 깊이 카메라를 포함하고, 상기 통합 엔진은 또한, 각각의 깊이 카메라의 상대 위치에 기초하여 상기 시스템을 교정하도록 구성되는 것인 양방향 컴퓨팅 시스템.

#### 청구항 2

제 1 항에 있어서,

상기 환경의 상기 조밀 3D 모델을 분석하여 상기 모델 내에서 물체를 식별하도록 구성된 물체 인식 모듈을 더 포함하는 양방향 컴퓨팅 시스템.

#### 청구항 3

제 2 항에 있어서,

상기 통합 엔진은 또한, 상기 조밀 3D 모델 내에서 식별된 물체에 기초하여 상기 애플리케이션을 적응시키도록 구성되는 것인 양방향 컴퓨팅 시스템.

#### 청구항 4

제 3 항에 있어서,

상기 애플리케이션을 적응시키는 것은, 디스플레이와 상기 이동 깊이 카메라의 검출된 상대 위치에 기초하여 물

체에 대한 위치 교정 정보를 자동으로 결정하는 것을 포함하는 것인 양방향 컴퓨팅 시스템.

**청구항 5**

제 4 항에 있어서,

상기 물체는 포인팅 장치나 상기 시스템의 사용자를 포함하는 것인 양방향 컴퓨팅 시스템.

**청구항 6**

제 1 항에 있어서,

상기 이동 깊이 카메라를 더 포함하는 양방향 컴퓨팅 시스템.

**청구항 7**

제 6 항에 있어서,

복수의 추가 깊이 카메라를 더 포함하는 양방향 컴퓨팅 시스템.

**청구항 8**

제 3 항에 있어서,

상기 식별된 물체는 프로젝터(projector)이고, 상기 통합 엔진은 상기 3D 모델을 이용하여 상기 프로젝터의 출력에 영향을 미치도록 구성되는 것인 양방향 컴퓨팅 시스템.

**청구항 9**

제 1 항에 있어서,

상기 시스템은 게임 시스템(gaming system)이고, 상기 애플리케이션은 게임인 것인 양방향 컴퓨팅 시스템.

**청구항 10**

양방향 컴퓨팅 시스템을 동작시키는 방법에 있어서,

환경 주위에서 움직이는 이동 깊이 카메라로부터 깊이 이미지의 시리즈를 수신하는 단계와,

깊이 이미지의 상기 시리즈로부터 상기 환경의 조밀 3D 모델을 생성하는 단계와,

상기 조밀 3D 모델의 적어도 일부를 애플리케이션으로 통합하는 단계

를 포함하고,

상기 조밀 3D 모델의 적어도 일부를 애플리케이션으로 통합하는 단계는,

상기 조밀 3D 모델을 분석하여 상기 모델 내에서 물체를 자동으로 식별하는 단계와,

물체가 식별되는 경우, 물체 유형과 물체 위치 정보를 상기 애플리케이션에 입력하는 단계와,

입력된 물체 유형 또는 위치 정보에 기초하여 상기 애플리케이션의 동작을 조정하는 단계와,

움직이지 않는 때에 상기 이동 깊이 카메라로부터 깊이 이미지의 제2 시리즈를 수신하는 단계와,

깊이 이미지의 상기 제2 시리즈를 이용하여 상기 환경의 영역에서의 3D 실시간 움직임을 추적하는 단계

를 포함하며,

상기 물체는 능동 물체를 포함하고, 상기 애플리케이션의 동작을 조정하는 단계는 상기 능동 물체의 출력을 조정하는 단계를 포함하거나,

상기 물체는 디스플레이를 포함하고, 상기 애플리케이션의 동작을 조정하는 단계는, 상기 디스플레이와 상기 이동 깊이 카메라의 검출된 상대 위치에 기초하여 상기 디스플레이에 대한 물체의 위치를 자동으로 교정하는 단계를 포함하거나,

상기 애플리케이션의 동작을 조정하는 단계는 다른 검출된 물체의 출력을 조정하는 단계를 포함하거나,

상기 물체는 다른 깊이 카메라를 포함하고, 상기 애플리케이션의 동작을 조정하는 단계는, 각각의 깊이 카메라의 상대 위치에 기초하여 상기 시스템을 교정하는 단계를 포함하는 것인 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 11**

제 10 항에 있어서,

깊이 이미지의 상기 제2 시리즈를 이용하여 상기 조밀 3D 모델을 생성하는 단계를 더 포함하는 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 12**

제 11 항에 있어서,

깊이 이미지의 상기 제2 시리즈를 이용하여 상기 조밀 3D 모델을 업데이트하는 단계를 더 포함하는 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 13**

제 10 항에 있어서,

상기 조밀 3D 모델의 적어도 일부를 애플리케이션으로 통합하는 단계는,

상기 조밀 3D 모델을 분석하여 상기 모델 내에서 물체를 자동으로 식별하는 단계와,

입력된 물체 유형 및 위치 정보에 기초하여 상기 애플리케이션의 동작을 조정하는 단계

를 포함하는 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 14**

제 10 항에 있어서,

상기 물체는 상기 능동 물체를 포함하고, 상기 애플리케이션의 동작을 조정하는 단계는 상기 능동 물체의 출력을 조정하는 단계를 포함하는 것인 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 15**

제 10 항에 있어서,

상기 물체는 상기 디스플레이를 포함하고, 상기 애플리케이션의 동작을 조정하는 단계는, 상기 디스플레이와 상기 이동 깊이 카메라의 검출된 상대 위치에 기초하여 상기 디스플레이에 대한 상기 물체의 위치를 자동으로 교정하는 단계를 포함하는 것인 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 16**

제 10 항에 있어서,

상기 애플리케이션의 동작을 조정하는 단계는 다른 검출된 물체의 출력을 조정하는 단계를 포함하는 것인 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 17**

제 10 항에 있어서,

상기 물체는 다른 깊이 카메라이고, 상기 애플리케이션의 동작을 조정하는 단계는 각각의 깊이 카메라의 상대 위치에 기초하여 상기 시스템을 교정하는 단계를 포함하는 것인 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 18**

제 10 항에 있어서,

하나 이상의 추가 깊이 카메라의 각각으로부터 깊이 이미지의 시리즈를 수신하는 단계와,

깊이 이미지의 각각의 시리즈로부터 상기 환경의 조밀 3D 모델을 생성하는 단계와,  
 복수의 조밀 3D 모델에서 공통 특성(common feature)을 식별하는 단계와,  
 식별된 공통 특성에 기초하여 카메라의 상대 위치를 계산하는 단계  
 를 더 포함하는 양방향 컴퓨팅 시스템을 동작시키는 방법.

**청구항 19**

제 10 항에 있어서,  
 상기 양방향 컴퓨팅 시스템은 게임 시스템이고, 상기 애플리케이션은 게임인 것인 양방향 컴퓨팅 시스템을 동작  
 시키는 방법.

**청구항 20**

동작의 제1 페이즈(phase) 및 제2 페이즈를 갖는 게임 시스템에 있어서,  
 동작의 상기 제1 페이즈에서 이동 깊이 카메라로부터 깊이 이미지의 제1 시리즈를, 그리고 동작의 상기 제2 페  
 이즈에서 상기 이동 깊이 카메라로부터 깊이 이미지의 제2 시리즈를 수신하기 위한 입력으로서, 상기 이동 깊이  
 카메라는 동작의 상기 제1 페이즈에서는 움직이고 동작의 상기 제2 페이즈에서는 정지하는 것인 상기 입력과,  
 동작의 상기 제1 페이즈에서 상기 이동 깊이 카메라의 위치를 추적하도록 구성된 실시간 카메라 추적 엔진과,  
 각각의 깊이 이미지가 캡처되었을 때에 적어도 깊이 이미지의 제1 시리즈와 상기 이동 깊이 카메라의 위치를 이  
 용하여 실내(room)의 3D 모델을 생성하도록 구성된 조밀 3D 모델링 시스템과,  
 상기 조밀 3D 모델의 적어도 일부를 게임으로 통합하고, 상기 게임에 대한 사용자 입력으로서 깊이 이미지의 상  
 기 제2 시리즈를 이용하여 상기 실내의 일부에서 실시간 움직임을 추적하도록 구성된 게임 엔진  
 을 포함하고,

상기 게임 엔진은 또한,

- 상기 조밀 3D 모델을 분석하여 상기 모델 내에서 물체를 식별하고,
- 물체가 식별되는 경우, 물체 유형 정보를 상기 게임에 입력하며,
- 입력된 물체 유형 정보에 기초하여 상기 게임의 동작을 조정하도록

구성되며,

상기 물체는 능동 물체를 포함하고, 상기 게임 엔진은 또한 상기 능동 물체의 출력을 조정하도록 구성되거나,  
 상기 물체는 디스플레이를 포함하고, 상기 게임 엔진은 또한, 상기 디스플레이와 상기 이동 깊이 카메라의 검출  
 된 상대 위치에 기초하여 상기 디스플레이에 대한 물체의 위치를 자동으로 교정하도록 구성되거나,  
 상기 게임 엔진은 또한, 다른 검출된 물체의 출력을 조정하도록 구성되거나,  
 상기 물체는 다른 깊이 카메라를 포함하고, 상기 게임 엔진은 또한, 각각의 깊이 카메라의 상대 위치에 기초하  
 여 상기 시스템을 교정하도록 구성되는 것인 게임 시스템.

**발명의 설명**

**기술 분야**

**배경 기술**

사용자의 움직임 또는 사용자가 잡고 있는 제어기의 움직임을 검출하여 게임과의 사용자의 상호작용을 가능하게  
 하기 위해 비디오 카메라 또는 깊이 카메라와 같은 카메라를 이용하는 게임 시스템이 개발되어 왔다. 사용자의

[0001]

위치/움직임을 검출하기 위해, 카메라에 의해 캡처되는 이미지는 전경(사용자)을 선택하고 배경을 제거하기 위해 분할된다. 그러면 사용자는 게임 자체 내에서 생성되고 사용자의 실제 환경과는 완전히 구별되는 환경에서 게임 내 캐릭터나 물체와 상호작용할 수 있다. 아래에 설명되는 실시형태는 알려진 게임 시스템의 단점의 전부 또는 여하한 것을 해결하는 구현으로 제한되지 않는다.

**발명의 내용**

**해결하려는 과제**

**과제의 해결 수단**

- [0002] 이하는 독자에게 기본적 이해를 제공하기 위해 본 개시의 단순화된 개요를 제공한다. 본 개요는 본 개시의 광범위한 개요는 아니고 본 발명의 중요/핵심 요소를 식별하거나 본 발명의 범위를 경계 짓지 않는다. 그 유일한 목적은 이후에 제시되는 더 상세한 설명의 전주로서 여기에 개시된 개념 중 선택된 것을 단순화된 형태로 제시하는 것이다.
- [0003] 게임플레이에서의 3D 환경 모델의 사용이 설명된다. 일 실시형태에서, 이동 깊이 카메라가 움직임에 따라 깊이 이미지의 시리즈를 캡처하는데 이동 깊이가 카메라가 사용되고 환경의 조밀(dense) 3D 모델이 깊이 이미지의 이 시리즈로부터 생성된다. 이 조밀 3D 모델은 게임과 같은 양방향 애플리케이션 내에 통합된다. 그러면 이동 깊이가 카메라가 일부 예에서 게임플레이인 양방향 페이지를 위하여 정지 위치에 배치되고, 시스템은 카메라에 의해 캡처된 깊이 이미지의 제2 시리즈로부터 환경의 일부 내의 사용자의 움직임을 검출한다. 이 움직임은 게임과 같은 양방향 애플리케이션으로의 사용자 입력을 제공한다. 다른 실시형태에서, 3D 모델 내 물체의 자동 인식 및 식별이 수행될 수 있고 이들 식별된 물체는 양방향 애플리케이션이 동작하는 방식을 변경한다.
- [0004] 많은 수반되는 특성이 첨부된 도면과 함께 아래의 상세한 설명을 참조함으로써 더 잘 이해되므로 더 용이하게 인식될 것이다.

**도면의 간단한 설명**

- [0005] 본 설명은 첨부된 도면에 비추어 아래의 상세한 설명으로부터 더 잘 이해될 것이다.
- 도 1은 방안에 서서 방의 3D 모델을 생성하는데 사용될 수 있는 이동 깊이가 카메라를 잡고 있는 사람의 개략도이다.
- 도 2는 이동 깊이가 카메라를 잡고 있는 사람에 의해 탐험되고 있는 건물의 바닥의 평면도를 도시한다.
- 도 3은 도 1에 도시된 양방향 컴퓨팅 시스템의 동작의 예의 흐름도이다.
- 도 4와 5는 양방향 컴퓨팅 시스템에 접속된 이동 깊이가 카메라의 개략도이다.
- 도 6은 도 5에 도시된 물체 인식 모듈의 사용이 관여되는 양방향 컴퓨팅 시스템의 동작의 일부의 흐름도를 도시한다.
- 도 7은 포인팅 장치를 교정하는 알려진 방법 및 자동 교정 방법 모두의 개략도를 도시한다.
- 도 8은 도 1에 도시된 것과 유사하지만 다수의 깊이 카메라를 갖는 배열의 개략도이다.
- 도 9는 다수의 깊이 카메라의 자동 교정의 예시적인 방법의 흐름도이다.
- 도 10은 여기 설명된 방법의 실시형태가 구현될 수 있는 예시적인 컴퓨팅 기반 장치를 도시한다.
- 도 11은 프레임 정렬 엔진에서의 예시적인 반복 프로세스의 흐름도이다.
- 도 12는 대응 지점의 쌍을 계산하는 단계를 더 자세히 도시하는 흐름도이다.
- 도 13은 3D 환경 모델을 생성하기 위한 병렬화가능한 프로세스의 예의 흐름도이다.

도 14는 예시적인 병렬화가 가능한 서명된 거리 함수 계산을 도시한다.

첨부된 도면에서 유사한 참조 부호는 유사한 부분을 지시하는데 사용된다.

**발명을 실시하기 위한 구체적인 내용**

- [0006] 첨부된 도면과 관련하여 아래에 제공되는 상세한 설명은 본 발명의 설명으로 의도된 것이고 본 예가 구성되거나 사용되는 유일한 형태를 나타내려는 것이 아니다. 본 설명은 예의 기능 및 예를 구성하고 운영하기 위한 단계의 시퀀스를 제시한다. 그러나, 동일 또는 균등한 기능 및 시퀀스가 다른 예에 의해 달성될 수 있다.
- [0007] 도 1은 방 안에 서서 게임 시스템과 같은 양방향 컴퓨팅 시스템(103)과 통신하는 (그리고 그 일부를 이룰 수 있는) 이동 깊이 카메라(102)를 잡고 있는 사람(100)의 개략도이다. 방은 다양한 물체(106)(의자, 테이블, 문, 창, 식물과 조명 등), 다른 사람(104), 디스플레이(105) 및 고양이(108)를 포함한다. 디스플레이(105)는 예를 들어 텔레비전이나 컴퓨터 모니터일 수 있거나, 다르게는, 디스플레이나 프로젝션 디스플레이를 통합하는 아이웨어(eyewear)를 포함하는 양방향 컴퓨팅 시스템(103)의 GUI(graphical user interface)를 표시할 수 있는 디스플레이의 여하한 다른 형태일 수 있다. 많은 물체(106)가 정지 상태이지만 사람(100, 104)과 고양이(108)과 같은 일부 물체가 움직일 수 있다. 이동 깊이 카메라(102)를 잡고 있는 사람(100)이 방을 돌아다니거나 가만히 서서 이동 깊이 카메라(102)를 움직임에 따라, 이동 깊이 카메라는 방과 물체의 깊이 이미지를 캡처한다.
- [0008] 이동 깊이 카메라(102)는 양방향 컴퓨팅 시스템(103)의 일부인 조밀(dense) 3D 환경 모델링 시스템(110)(이 경우 환경은 방)에 있을 수 있다. 이 예에서 "조밀"은 표면을 암시적으로 묘사하기 위해 인입 깊이 맵으로부터의 지점의 모두 또는 많은 것을 사용함으로써 나오는 모델의 고도의 해상도 및 정확성을 지칭한다. 이는 계산 속도를 높이고 메모리 풋프린트(footprint)를 감소시키기 위해 지점의 부분집합만을 취하는 "성긴(sparse)" 모델에 대비될 수 있다. 예를 들어, 이동 깊이 카메라(102)에 의해 캡처된 이미지는 사람이 방을 돌아다니거나 이동 깊이 카메라(102)를 움직임에 따라 환경의 조밀 3D 모델을 형성하고 구축하는데 사용된다.
- [0009] 실시간 카메라 추적 시스템(112)은 방안의 카메라(102)의 위치 및 방향을 모니터링한다. 실시간 카메라 추적 시스템(112)은 이동 깊이 카메라(102)와 통합될 수 있거나, 양방향 컴퓨팅 시스템(103)의 일부와 같이 다른 위치에 있을 수 있다. 실시간 카메라 추적 시스템(112)은 어디에 배치되는지와 무관하게, 직접 또는 간접으로 이동 깊이 카메라(102)로부터 통신을 수신할 수 있다. 일 예에서, 실시간 추적 시스템(112)은 양방향 컴퓨팅 시스템(103)의 일부인 컴퓨팅 장치에서 제공될 수 있고, 이동 깊이 카메라(102)와 무선으로 통신할 수 있다. 다른 예에서, 실시간 카메라 추적 시스템(112)은 건물 내 다른 곳에 있거나 여하한 적당한 유형의 통신 네트워크를 이용하는 이동 깊이 카메라(102)와 통신하는 다른 원격 장소에 있을 수 있다.
- [0010] 실시간 카메라 추적 시스템(112)은 개별 깊이 이미지가 전체 3D 모델로 구축될 수 있도록 하기 위해, 조밀 3D 모델링 시스템으로 입력을 제공한다. 실시간 카메라 추적 시스템(112)은 환경의 3D 모델과 관련하여 카메라의 위치를 또한 추적할 수 있다. 카메라 추적과 3D 모델링의 조합은 SLAM(simultaneous localization and mapping)으로 알려져 있다.
- [0011] 실시간 카메라 추적 시스템(112)과 조밀 3D 모델링 시스템(110)의 출력은 통합 엔진(114)에 의해 사용될 수 있는데, 통합 엔진은 게임 엔진, 증강 현실 엔진 또는 환경의 조밀 3D 모델의 적어도 일부의 애플리케이션으로의 통합을 가능하게 하는 여하한 엔진을 포함할 수 있다. 예를 들어, 통합 엔진은 양방향 애플리케이션(이는 게임일 수 있다)과의 사용자 상호작용을 가능하게 할 수 있다. 예를 들어, 현실 물체가 게임 (또는 기타 양방향) 환경에 포함될 수 있고 이는 아래에서 더 상세히 설명된다. 다른 예에서, 이동 깊이 카메라(102)의 프로젝터가 실시간 카메라 추적 시스템(112)과 3D 모델링 시스템(110)의 출력에 의존하여 이미지를 투사하도록 배열될 수 있다.
- [0012] 다른 예로서, 도 2는 건물의 바닥(200)의 평면도를 도시한다. 이동 깊이 카메라(204)를 잡고 있는 사람(202)이 점선 화살표(208)에 의해 나타내는 바와 같이 바닥을 돌아다니고 있다. 사람은 방과 가구(210)를 지나 복도를 따라 걷는다. 실시간 카메라 추적 시스템(112)은 이동 깊이 카메라(204)가 움직이는 때에 그 위치를 추적할 수 있고 3D 모델링 시스템은 바닥의 맵이나 3D 모델을 생성한다. 사람(202)이 이동 깊이 카메라(204)를 운반하는 것이 중요한 것이 아니다. 다른 예에서 이동 깊이 카메라(204)는 로봇이나 차량에 설치된다. 이는 도 1의 예에도 적용된다.
- [0013] 양방향 컴퓨팅 시스템(103)의 동작의 예의 흐름도가 도 3에 도시된다. 이동 깊이 카메라(102)가 도 1에 도시된 방과 같은 환경을 돌아다님에 따라 깊이 이미지의 시리즈가 이동 깊이 카메라(102)로부터 수신된다(블록 302).

그 후 이 깊이 이미지의 시리즈는 환경의 조밀 3D 모델을 생성하는데 사용되고(블록 304) 이 3D 모델 (또는 모델의 일부)가 양방향 컴퓨팅 시스템에서 실행되고 사용자가 상호작용할 수 있는 양방향 애플리케이션 내에 통합된다(블록 306). 양방향 컴퓨팅 시스템(103)이 게임 시스템인 경우, 양방향 애플리케이션은 게임이고 3D 모델 (또는 그 일부)는 게임 내에 통합된다. 일단 3D 모델이 게임 내에 통합되면, 모델은 사용자에게 보일 수 있거나 게임 내에서 사용되지만 사용자에게 보이지 않을 수 있다. 조밀 3D 배경 모델이라고도 지칭될 수 있는 환경의 조밀 3D 모델의 생성은 양방향 컴퓨팅 시스템의 동작의 제1 페이지(점선 박스 31로 표시됨)로 생각될 수 있고, 오프라인에서, 애플리케이션(예를 들어, 게임)의 일부로서 실시간으로 또는 별개의 활동으로서(예를 들어, 게임 또는 게임 시스템을 위한 설정 페이지로서) 수행될 수 있다.

[0014] 카메라가 정지한 때에 깊이 이미지의 제2 시리즈가 후속하여 이동 깊이가 카메라(102)로부터 수신되고(블록 308), 깊이 이미지의 이 제2 시리즈는 카메라의 시야(FOV; field of view) 내에서 3D 실시간 움직임을 동적으로 추적하는데 사용되는데, 카메라의 시야는 제1 페이지에서 모델링된 환경의 일 영역(예를 들어, 방의 일부)만을 나타낼 것이다(블록 310). 환경의 이 영역은 '관심 영역(region of interest)'이라고 지칭될 수 있으며, 일부 구현에서 시스템의 동작에 있어서 이 단계 동안 이동 깊이가 카메라에 대한 정의된 고정 위치(예를 들어 도 1의 디스플레이 위나 아래)가 있을 수 있다. 환경의 일부에서의 실시간 감지(블록 308과 310)는 양방향 컴퓨팅 시스템의 동작의 제2 페이지(점선 박스 32로 표시됨)라고 생각될 수 있다. 식별된 사용자 제스처에 관한 데이터와 같은, 로(raw) 데이터나 처리된 데이터를 포함할 수 있는 실시간 움직임 데이터는 양방향 애플리케이션으로의 입력으로 제공된다(블록 312). 깊이 이미지의 제2 시리즈는 또한 (실시간으로 또는 후속하여) 이전에 생성된 3D 모델을 업데이트하는데 사용될 수 있고(블록 314), 그러면 업데이트된 모델(또는 업데이트 모델의 태양)이 애플리케이션으로 다시 피드백될 수 있다(블록 314에서 블록 306으로의 점선 화살표로 나타냄).

[0015] 게임과 같은 양방향 애플리케이션으로의 환경(배경이라고 생각될 수도 있다)의 상세한 모델의 통합은 전체 사용자 경험을 개선하고 상호작용의 현실성을 증가시킬 수 있다. 예를 들어, 사용자가 공과 같은 양방향 애플리케이션 내의 가상 물체와 상호작용하는 경우, 공의 움직임과 행태는 조밀 3D 모델에 기초하여 적용될 수 있다. 사용자가 벽이나 기타 물체를 향해 가상 공을 차면, 양방향 애플리케이션은, 공이 벽에서 되튀어 나오는 것으로 보이도록 조밀 3D 모델에 기초하여 (애플리케이션의 GUI에서 보여지는 바와 같은) 가상 공의 행태를 수정할 수 있다. 다른 예에서, 사용자가 가상 공을 방 안의 더 작은 (실제) 물체를 향해 차면, 공은 반드시 되튀어야 하는 것은 아닐 수 있고 가상 공의 궤적은 실제 물체에 의한 가상적인 충격을 따라 변경될 수 있다. 후술하는 바와 같이, 특정한 더 작은 실제 물체의 특성이 결과적인 가상 공의 움직임을 결정하는데 사용될 수 있다. 광원(예를 들어, 도 1의 램프나 창)으로부터의 조명, 물체(예를 들어, 도 1의 식물)의 그림자 등과 같은 환경의 다른 태양이 또한 양방향 애플리케이션으로 통합될 수 있다.

[0016] 또한, 환경의 상세한 모델의 양방향 애플리케이션(게임 등)으로의 통합은 새로운 사용자 경험을 제공할 수 있다. 예를 들어, 작은 스케일로(예를 들어, 사용자가 시뮬레이션된 미니어처 비행기를 타고 비행) 사용자의 집(예를 들어, 제1 페이지에서 캡처된 거실)에서 플레이되는 격투 시뮬레이션 게임. 다른 예에서, 양방향 애플리케이션(게임 등)은, 예를 들어, 소파에서 폭포가 떨어지는 등, 방을 정글이나 기타 다른 환경으로 변환할 수 있다. 다른 예에서, 양방향 애플리케이션(게임 등)은 사용자가 그의 집을 시뮬레이션 게임으로 통합시킬 수 있도록 할 수 있다. 또 다른 예에서, 양방향 애플리케이션으로의 상세한 모델의 통합은 사용자가 실제로 변경을 가하거나(예를 들어, 벽을 다른 색으로 칠하거나, 카펫의 색을 바꾸거나, 벽을 없애는 등) 새로운 가구 등을 사기 전에 방의 가상적 표현 내에서 홈 데코레이션, 새로운 가구 등의 변경을 시도해볼 수 있도록 할 수 있다. 일 구현에서, 애플리케이션 내에 통합되는 모델을 생성하는데 사용될 깊이 데이터를 가져오고(import) 그 후 상점에서 가상 장면으로 아이템을 가져올 수 있도록, 양방향 애플리케이션은 홈 퍼니싱 스토어에 의해 (예를 들어 웹사이트를 통해) 제공될 수 있다.

[0017] 도 1에 도시된 시스템은 단일 깊이 카메라(이동 깊이가 카메라(102))를 포함하고, 도 3에 도시된 방법은 환경의 상세한 3D 모델을 생성하고 또한 그 더 큰 환경의 일 영역 내에서 3D 실시간 움직임을 추적하기 위해 단일 깊이 감지 장치가 사용될 수 있도록 한다. 이는 다수의 깊이 감지 장치를 포함하는 시스템보다 저비용 시스템을 제공한다. 도 3의 방법은 다수의 깊이 카메라를 포함하는 시스템과 결합하여 사용될 수도 있고 이는 도 8을 참조하여 아래에서 더 상세히 설명된다.

[0018] 도 4는 실시간 카메라 추적기(416), 조밀 모델 형성 시스템(424) 및 통합 엔진(432)(게임 엔진일 수 있음)과 함께 사용하기 위한 이동 환경 캡처 장치(400)의 개략도이다. 이동 환경 캡처 장치(400)는 장면의 깊이 이미지의 시퀀스 (또는 시리즈)를 캡처하도록 배열되는 깊이 카메라(402)를 포함한다. 이후 깊이 카메라의 언급은 독립 깊이 카메라 또는 도 4에 도시된 이동 환경 캡처 장치의 일부인 깊이 카메라를 지칭할 수 있다. 깊이 감지를

위해 사용되는 기술에 따라, 깊이 카메라는 추가의 컴포넌트(예를 들어, 깊이 카메라(402)가 적외선 구조광 패턴을 검출하는 경우 에미터(404)가 요구된다)를 요할 수 있지만 이러한 추가 컴포넌트는 반드시 깊이 카메라와 함께 위치되어야 하는 것은 아니다.

- [0019] 각각의 깊이 이미지(414) (또는 깊이 맵 프레임)는, 각각의 이미지 요소(즉, 픽셀)가 카메라로부터 그 이미지 요소가 발생하도록 하는 캡처된 장면 내 물체까지의 거리 또는 길이와 같은 깊이 값을 포함하는 2차원 이미지를 포함한다. 이 깊이 값은 미터나 센티미터와 같은 특정 척도 단위로 제공되는 절대값일 수 있거나 상대적 깊이 값일 수 있다. 각각의 캡처된 깊이 이미지(414)에, 각각 깊이 값을 갖는 약 300,000 이상의 이미지 요소가 있을 수 있다. 깊이 카메라(402)의 프레임 레이트는, 깊이 이미지가 로보틱스, 컴퓨터 게임 또는 기타 애플리케이션을 작동하기 위해 사용될 수 있도록 하기에 충분히 높다(예를 들어, 초당 20프레임 이상).
- [0020] 깊이 정보는, 비행시간(time of flight), 구조광(structured light) 및 스테레오 이미지(stereo image)를 포함하지만 이에 제한되지 않는 여하한 적당한 기술을 이용하여 획득될 수 있다. 이동 환경 캡처 장치(400)는 깊이 정보가 깊이 카메라(402)에 의해 확인될 수 있도록 하는 방식으로 장면을 조명하도록 배열된 에미터(404)를 또한 포함할 수 있다.
- [0021] 예를 들어, 깊이 카메라(402)가 적외선(IR) 비행시간 카메라인 경우, 에미터(404)는 IR 광을 장면에 발산하고, 깊이 카메라(402)는 장면 내 하나 이상의 물체의 표면으로부터 후방산란(backscatter)되는 광을 검출하도록 배열된다. 일부 예에서, 인출 광 펄스와 대응하는 인입 광 펄스 사이의 시간이 깊이 카메라에 의해 검출되고 측정되어 환경 캡처 장치(400)로부터 장면 내 물체 상의 위치까지의 물리적 거리를 결정하는데 사용되도록 펄스 적외선 광이 에미터(404)로부터 발산될 수 있다. 또한, 일부 예에서, 에미터(404)로부터의 인출 광파의 위상이 깊이 카메라(402)에서의 인입 광파의 위상과 비교되어 위상 변이가 결정될 수 있다. 그러면, 예를 들어, 셔터 광 펄스 이미징(shuttered light pulse imaging)을 포함하는 다양한 기술을 통해 시간에 걸쳐 반사된 광선의 강도를 분석함으로써, 위상 변이는 이동 환경 캡처 장치(400)로부터 물체 상의 위치까지의 물리적 거리를 결정하는데 사용될 수 있다.
- [0022] 다른 예에서, 이동 환경 캡처 장치(400)는 깊이 정보를 캡처하기 위해 구조광을 사용할 수 있다. 이러한 기술에서, 패턴닝된 광(예를 들어, 점, 격자(grid) 또는 띠(stripe) 패턴과 같은 알려진 패턴으로 표시되고, 또한 시변(time-varying)일 수 있는 광)이 에미터(404)를 이용하여 장면에 투사될 수 있다. 장면 내 물체의 표면에 부딪히면, 패턴은 그 패턴을 캡처하는 깊이 카메라(402)의 시점에서 변형된다. 캡처된 패턴의 변형은 깊이 카메라(402)로부터 장면 내 물체까지의 절대 또는 상대 거리를 결정하기 위해 분석된다.
- [0023] 다른 예에서, 상대 깊이 정보를 생성하기 위해 시각 스테레오 데이터가 획득되고 분해(resolve)되도록 깊이 카메라(402)는 한 쌍의 스테레오 카메라를 포함한다. 이 경우 에미터(404)는 장면을 조명하기 위해 사용되거나 생략될 수 있다.
- [0024] 일부 예에서, 깊이 카메라(402)에 추가하여, 이동 환경 캡처 장치(400)는 RGB 카메라(406)라고 불리는 컬러 비디오 카메라를 포함한다. RGB 카메라(406)는 가시광 주파수에서 장면의 이미지의 시퀀스를 캡처하도록 배열된다.
- [0025] 이동 환경 캡처 장치(400)는 관성 측정 유닛(IMU; inertial measurement unit), 가속도계, 자이로스코프, 나침반, 또는 기타 방향 센서(408)와 같은 방향 센서(408)를 포함할 수 있다. 그러나, 방향 센서를 이용하는 것은 중요하지 않다. 이동 환경 캡처 장치(400)는 GPS와 같은 위치 추적 장치를 포함할 수 있지만, 이는 중요하지 않다.
- [0026] 이동 환경 캡처 장치는 상술한 바와 같이 프로젝터(412)를 포함할 수 있지만, 이는 중요하지 않다. 이동 환경 캡처 장치는 또한, 아래에 더 설명되는 바와 같이, 하나 이상의 프로세서, 메모리 및 통신 인프라스트럭처를 포함한다. 이동 환경 캡처 장치는, 사용자가 손으로 잡거나 사용자가 입도록 모양 및 사이즈가 정해지는 하우징 내에 제공될 수 있다. 다른 예에서, 이동 환경 캡처 장치는 차량, 장난감 또는 기타 움직일 수 있는 장치 상에 통합되거나 설치되도록 크기와 모양이 정해진다. 이동 환경 카메라는 또한, 고정 (또는 정지) 위치에서 동작하는 때에(예를 들어, 도 3의 제2 페이지, 박스 32) 표면이나 브래킷(bracket) 상에 배치될 수 있도록 모양이 정해질 수 있다.
- [0027] 이동 환경 캡처 장치(400)는 실시간 추적기(416)에 접속된다. 이 접속은 물리적 유선 접속일 수 있고 무선 통신을 이용할 수 있다. 일부 예에서, 이동 환경 캡처 장치(400)가 인터넷과 같은 하나 이상의 통신 네트워크를 통해 간접적으로 실시간 추적기(416)로 접속된다.

- [0028] 실시간 추적기(416)는 하나 이상의 그래픽 처리 유닛을 제어하는 범용 마이크로소프트나 기타 병렬 컴퓨팅 유닛을 이용하여 컴퓨터-구현된다. 이는 프레임 정렬 엔진(418) 및 선택적으로 루프 폐쇄 엔진(420)과 재로컬화(relocalization) 엔진(422)을 포함한다. 실시간 추적기(416)는 깊이 카메라로부터 깊이 이미지 프레임을, 또한 선택적으로 이동 환경 캡처 장치(400)로부터의 입력을, 그리고 선택적 맵 데이터(434)를 취한다. 깊이 카메라(402)의 여섯 개의 자유도(degree of freedom)(6DOF) 포즈 추정의 실시간 시리즈(428)를 산출하기 위해 실시간 추적기(416)는 공간 정렬(spatial alignment)로 깊이 이미지 프레임을 배치하도록 동작한다. 또한 깊이 이미지 프레임의 쌍 사이의 변환을 위해 변환 파라미터를 산출할 수 있다. 일부 예에서, 실시간 추적기는 깊이 카메라로부터의 깊이 이미지 프레임(414)의 쌍에 대해 동작한다. 다른 예에서, 실시간 추적기(416)는 단일 깊이 이미지(414)를 취하고 다른 깊이 이미지가 아니라 환경의 조밀 3D 모델(426)과 이를 정렬한다. 실시간 추적의 방법의 일 예는, 본 출원과 같은 날 출원되고 같이 계류 중인 미국 특허 출원 'Real-time Camera Tracking Using Depth Maps'에서 설명된다. 이 방법은 또한 도 11과 12를 참조하여 아래에서 간략하게 설명된다.
- [0029] 실시간 추적기(416)는, 이동 환경 캡처 장치(400)가 움직이고 있는 환경 또는 장면의 조밀 3D 모델(426)을 형성 및 저장하기 위해 깊이 이미지 프레임과 함께 그 정보를 이용하는 조밀 3D 모델 생성 시스템(424)으로의 출력으로 카메라 포즈를 제공한다. 예를 들어, 도 1의 경우에, 3D 모델은 방안의 물체 및 표면의 3D 모델이 된다. 도 2의 경우에, 3D 모델은 건물의 바닥의 3D 모델이 된다. 조밀 3D 모델은 GPU 메모리에 저장될 수 있다. 조밀 3D 모델 생성의 방법의 일 예는 본 출원과 같은 날 출원되고 같이 계류 중인 미국 특허 출원 'Three-Dimensional Environment Reconstruction'에서 설명된다. 이 방법은 또한 도 13과 14를 참조하여 간략하게 후술된다. 조밀 3D 모델 생성의 다른 예는 SIGGRAPH '96에 게재된 Curless와 Levoy의 논문 'A Volumetric Method for Building Complex Models from Range Images'에 설명된다.
- [0030] 실시간 추적기(416) 및/또는 조밀 3D 모델 형성 시스템(424)에 의해 수행되는 처리는, 일 예에서, 이동 환경 캡처 장치(400)의 위치로부터 원격으로 실행될 수 있다. 예를 들어, 이동 환경 캡처 장치(400)는 상대적으로 낮은 처리 전력을 갖고, 통신 네트워크를 통해 서버로 깊이 이미지를 스트리밍하는 컴퓨팅 장치에 접속될 수 (또는 이를 포함할 수) 있다. 서버는 상대적으로 높은 처리 전력을 갖고, 실시간 추적기(416) 및/또는 조밀 3D 모델 형성 시스템(424)의 연산적으로 복잡한 태스크를 수행한다. 서버는 사용자에게 양방향 경험을 제공하기 위해 프레임마다 조밀 재구성의 렌더링된 이미지를 반환할 수 있고, 또한, 나중의 로컬 사용(예를 들어, 게임에서)을 위해 모델의 완성 시에 대한 최종 조밀 3D 재구성을 반환할 수 있다. 이러한 배열은 사용자가 고 전력 로컬 컴퓨팅 장치를 보유할 필요가 없게 한다.
- [0031] 사용자가 조밀 3D 모델을 구축하는 것을 돕기 위해, 조밀 3D 모델의 실시간 렌더링의 형태로 사용자에게 피드백이 제공될 수 있다. 이는 사용자가 구성되는 바와 같이 조밀 모델의 상태를 볼 수 있도록 하고, 그에 의해, 놓친 영역을 캡처하는데 있어 사용자를 돕는다. 환경의 어떤 영역이 더 많은 데이터를 필요로 하는지를 사용자에게 나타내기 위해 색 및/또는 텍스처가 렌더링에 중첩(superimpose)될 수 있다. 이는 사용자를 위한 더 양방향성의 경험을 제공한다. 조밀 3D 모델에서 낮은 품질이거나 놓친 영역으로 사용자를 안내하기 위해 사용자에게 오디오 또는 시각적 큐가 사용자에게 제공될 수도 있다.
- [0032] 이동 환경 캡처 장치(400)는 디스플레이 장치(430)에 접속되는 통합 엔진(432)(이는 게임 엔진일 수 있다)과 결합하여 사용될 수 있다. 예를 들어, 게임은 FPS(first-person shooter) 게임, 골프 게임, 복싱 게임, 모터 레이싱 게임 또는 기타 유형의 컴퓨터 게임일 수 있다. 조밀 3D 모델은, 도 3을 참조하여 상술한 바와 같이, 통합 엔진(432) 및 게임으로 통합된 모델의 태양으로 제공될 수 있다. 예를 들어, 3D 모델은 방 안의 물체의 위치와 모양을 결정하는데 사용될 수 있는데, 그러면 방 안의 물체는 플레이어가 상호작용할 수 있는 게임 내 물체로서 게임 자체에 통합될 수 있거나, 게임 내의 가상 물체가 환경 내의 물체와 (가상적으로) 상호작용할 수 있다(예를 들어, 플레이어가 찬 가상 공은 환경 내의 방의 벽 또는 기타 물체를 때린 후 되튀는 것으로 보일 수 있다). 관심 영역에서의 3D 실시간 움직임을 추적하고 이 움직임을 게임과 같은 애플리케이션으로의 입력으로 이용하기 위해, 통합 엔진(432)은 또한 제2 페이지 동안(예를 들어, 통합 엔진이 게임 시스템인 경우 게임 플레이 동안) 캡처된 깊이 이미지를 이용한다. 깊이 이미지(414)를 이용한 실시간 움직임의 추적은 (예를 들어, 2009.5.1. 출원된 공개 번호 US-2010-0278384-A1 같이 계류 중인 미국 특허 출원 'Human Body Pose Estimation'에 설명된 것과 같은) 배경 제거 후 예-기반 신체 인식(exemplar-based body part recognition)과 같은 알려진 기술을 이용하여 수행될 수 있다. 게임 상태나 게임에 대한 메타데이터와 같은 통합 엔진(432)으로부터의 데이터가 또한 실시간 추적기(416)로 제공될 수 있다.
- [0033] 실시간 추적기(416)에게 맵 데이터(434)가 선택적으로 사용가능할 수 있다. 예를 들어, 이는 건축가의 환경(예를 들어, 건물의 방 또는 바닥)에 대한 도면, 환경에서 알려진 랜드마크의 위치, 다른 소스로부터 사용가능할

환경의 지도일 수 있다.

- [0034] 실시간 추적기의 프레임 정렬 엔진(418)은 깊이 이미지 프레임의 쌍 또는 깊이 이미지 프레임과 조밀 3D 모델로부터의 깊이 이미지 프레임의 추정을 정렬하도록 배열된다. 이는 프레임 정렬 엔진이 실시간으로 동작하도록 하나 이상의 그래픽 처리 유닛을 이용하여 구현되는 반복(Iterative) 프로세스를 이용한다. 루프 폐쇄 엔진(420)은 이동 환경 캡처 장치가 루프에서 움직인 때를 검출하도록 배열되어서, 현재 깊이 프레임에서 그려진 장면이 적어도 부분적으로 이전의 깊이 프레임의 것과 오버랩하고 있다. 예를 들어, 이는 사용자가 도 2의 건물의 전체 바닥을 걸어다닌 후에 다시 시작점에 도달하는 때에 일어날 수 있다. 이는 또한 사용자가 어떤 가구 뒤에서 방을 돌아다니다가 원래의 시작점이나 그 원래의 시작점에 가까운 곳으로 다시 나오는 때에도 일어난다. 재로컬화(relocalization) 엔진(422)은, 실시간 추적기가 이동 환경 캡처 장치(400)의 현재 위치를 상실하고 현재 위치를 다시 찾거나 재로컬화하는 상황을 다루도록 배열된다.
- [0035] 전문한 바와 같이, 실시간 추적기(416) 및/또는 조밀 3D 모델 생성 시스템(424)에 의해 수행되는 처리는, 일 예에서, 이동 환경 캡처 장치(400)의 위치로부터 떨어져서 실행될 수 있다. 예를 들어, 이동 환경 캡처 장치(400)는 상대적으로 낮은 처리 전력을 갖고 통신 네트워크를 통해 서버로 깊이 이미지를 스트리밍하는 컴퓨팅 장치에 접속될 수(또는 이를 포함할 수) 있다. 서버는 상대적으로 높은 처리 전력을 갖고, 실시간 추적기(416) 및/또는 조밀 3D 모델 생성 시스템(424)의 연산적으로 복잡한 태스크를 수행한다. 서버는 사용자에게 양방향 경험을 제공하기 위해 프레임마다 조밀 재구성의 렌더링된 이미지를 반환할 수 있고, 또한, 나중의 로컬 사용(예를 들어, 게임에서)을 위해 모델의 완성 시에 대한 최종 조밀 3D 재구성을 반환할 수 있다. 이러한 배열은 사용자가 고전력 로컬 컴퓨팅 장치를 보유할 필요가 없게 한다.
- [0036] 상술한 시스템 및 방법에서, 이동 깊이 카메라가 움직이는 동안(페이지 1, 도 3의 박스 31)캡처된 깊이 이미지에 기초하여 구성된 조밀 3D 모델은 양방향 컴퓨팅 시스템 상에서 실행되는 애플리케이션으로 통합되고, 그러면 사용자는 이동 깊이 카메라가 안정되어서 정지되면(페이지 2, 도 3의 박스 32) 검출되는 움직임을 통해 애플리케이션(그리고 환경)과 상호작용할 수 있다. 일부 예에서, 페이지 1에서 이동 깊이 카메라를 어떻게 움직이는 지에 대하여 시스템에 의해 사용자에게 명령이 제공될 수 있고, 다른 예에서, 사용자는 어떠한 방식으로든 이동 깊이 카메라를 자유롭게 움직일 수 있다. 상술한 시스템과 방법의 변형에서, 양방향 컴퓨팅 시스템은 또한 도 5의 개략도에서 도시되는 바와 같은 물체 인식 모듈(502)도 포함할 수 있다. 이 모듈(502)에 의해 3D 모델 내로부터 인식되는 물체는 디스플레이 장치(430) 또는, 스피커, 프로젝터, 광원 등과 같은 다른 주변장치(504) 등의 능동 물체와 가구와 같은 수동 물체를 포함할 수 있다.
- [0037] 도 6은 물체 인식 모듈(502)의 사용이 관여되는 양방향 컴퓨팅 시스템의 동작의 일부의 흐름도를 도시한다. 이 방법은, 예를 들어, 도 3에서 애플리케이션으로 3D 모델을 통합하는 단계(블록 306)의 일부를 형성하거나 이 단계 후에 통합 엔진에 의해 수행될 수 있다. (조밀 3D 모델 생성 시스템(424)에 의해 생성되는 바와 같은)3D 모델은 인식가능한 물체를 찾기 위해 분석된다(블록 602). 이는, 예를 들어, 생성된 모델에서 특정 특성을 찾거나 저장된 템플릿과 모델의 부분들을 비교함으로써 수행될 수 있는데, 여기서 이들 템플릿 또는 특정 특성은 물체 데이터베이스(506)에 저장될 수 있다. 일 예에서, 텔레비전의 근사적 모양(예를 들어, 깊이 약 10cm이고 약 16:9의 폭-높이 비를 가짐)을 정의하는 일반 '텔레비전' 템플릿이 저장될 수 있다. 여하한 적당한 이미지 분석이나 머신-러닝(machine-learning) 방법(블록 602에서)이 분석을 수행하기 위해 사용될 수 있고, 일부 예에서, 트레이닝 데이터가 물체 데이터베이스(506)에 저장될 수 있다.
- [0038] 조밀 3D 모델의 해상도에 따라, 물체 인식 모듈(502)은 물체 인식을 돕기 위해 양각 또는 음각된 제품 로고를 식별할 수 있을 수 있다(예를 들어, DELL(상표) 로고가 컴퓨터나 컴퓨터 디스플레이를 나타낼 수 있다). 이동 환경 캡처 장치(400)가 깊이 카메라(402)에 추가하여 선택적인 RGB 카메라(406)(도 4에 도시)를 포함하는 경우, RGB 이미지는 물체 인식을 수행하기 위해 3D 모델과 결합하여 사용될 수 있다(예를 들어, 로고 검출을 위해 또는 바코드나 제품 ID를 읽기 위해). 인식될 물체가 텔레비전, 컴퓨터 모니터나 이동 전화와 같은 디스플레이를 포함하는 경우, 스캔하는 동안 알려진 특정 패턴이 장치상에 표시될 수 있어서(그리고, 장치가 시스템의 제어하에 있는 경우, 이 알려진 패턴은 시스템에 의해 장치로 출력될 수 있다), RGB에서 쉽게 검출될 수 있다. 유사하게, 프로젝터에 대해, 프로젝터는 시스템의 제어 하에서 또는 독립적으로 알려진 패턴(선택적으로 시변임)을 출력할 수 있고, 현실의 RGB 패턴을 찾음으로써 방 안에서 프로젝터의 위치가 결정될 수 있다(예를 들어, 깊이 재구성으로부터 알려진 방 수치를 이용하여). 일부 시스템에서, 마커(예를 들어, 적외선 반사 마커)가 물체 인식/분류를 돕기 위해 사용될 수 있다.
- [0039] 물체가 인식된 때에(블록 604에서 '예'), 물체에 대한 데이터가 통합 엔진(432)에서 실행되는 애플리케이션으로

입력되는데(블록 606), 애플리케이션은 예를 들어 게임 시스템에서 실행되는 게임일 수 있다. 이 물체에 대한 데이터는, 예를 들어, 물체 유형(예를 들어, '텔레비전'이나 '디스플레이') 및 (예를 들어, 이동 깊이 카메라 또는 이동 환경 캡처 장치(400)의 현재 위치에 대한)위치 정보를 포함할 수 있다. 그러면 애플리케이션 성능(또는 동작)이 식별된 물체에 기초하여 조정된다(블록 608).

[0040] (블록 608에서) 식별된 물체에 기초하여 애플리케이션 성능(또는 동작)이 조정될 수 있는 많은 상이한 방법이 존재하고, 이는 식별된 물체의 유형에 의존할 수 있다. 물체가 스피커와 같은 능동 물체인 경우, 양방향 컴퓨팅 시스템의 오디오 출력이 스피커의 위치와 간격을 고려하여 조정될 수 있다(블록 681). 이러한 방식으로, 시스템은 서라운드 사운드 시스템의 효과를 모방(simulate) 가능할 수 있거나 진짜 서라운드 시스템이 그 출력을 교정 및 최적화하기 위해 전체 3D 방 재구성을 이용가능할 수 있다. 일부 예에서, 다수의 스피커가 식별되는 경우, 양방향 컴퓨팅 시스템 및 검출된 스피커 물체로부터의 오디오 출력 사이의 구별을 위해 각각의 스피커에서 순서대로 소리를 발산하는 것과 같은 추가 단계를 수행하는 것이 필요할 수 있다(이동 환경 캡처 장치(400)는 선택적으로 도 4에 도시되지 않은 마이크로폰을 포함할 수 있다). 일부 상황에서, 이 정보는 검출된 스피커 물체의 모양/크기 및/또는 스피커의 상대적 위치로부터 추론가능할 수 있으므로, 이는 필요하지 않을 수 있다. 스피커의 검출 및 후속하는 서라운드 사운드의 모방은, 예를 들어, 게임 또는 증강 현실 환경의 현실성을 높이기 위해 사용될 수 있다.

[0041] 능동 물체의 다른 예는 디스플레이 장치(430) 자체이고, 카메라와 디스플레이의 상대 위치에 대한 지식을 이용하여, 도 7을 참조하여 설명할 수 있는 바와 같이 포인팅 장치의 자동 교정이 수행될 수 있다(블록 682). 많은 알려진 시스템에서, 디스플레이 장치(702)에 대해 포인팅 장치를 교정하기 위해, 도 7의 위 그림에서 도시된 바와 같이, 디스플레이의 각각의 모서리에 표시되는 레티클(704) 또는 기타 마커)를 지시하는 것(그리고 가능하게는 지시하는 동안 버튼을 누르는 것)이 필요하다. 그러나 여기 설명된 시스템에서, 통합 엔진은 깊이 카메라(712), 디스플레이(714) 및 포인팅 장치(716)(공 게임을 하는 때에는 배트 또는 FPS 게임을 하는 때에는 모형 무기 등)의 상대적인 위치를 알고 있으며, 그래서 사용자가 지시하는 디스플레이 상의 위치(도 7의 아래 그림에서 화살표(718)로 표시)가 초기 교정 페이즈(도 7의 위 그림에 도시됨)를 요하지 않고 (예를 들어, 삼각법을 이용하여)계산될 수 있다. 또한, 생성된 환경의 조밀 3D 모델에 기초하여, 심지어 사용자가 디스플레이 지시하고 있지 않은 경우에도, 통합 엔진은 사용자가 환경 내에서 무엇을 지시하는지를 또한 결정할 수 있다. 이는, 예를 들어, 사용자가 정지 위치의 (페이즈 2에서) 이동 깊이 카메라의 시야 밖에 있는 실제 물체를 선택하고(예를 들어, 그것을 가리킴으로써) 게임 내 물체로서 게임으로 가져올 수 있도록 할 수 있다.

[0042] 도 7에 도시되고 상술한 것의 변형에서, 방법은 사용자가 환경 내 어디를 보고 있는지를 결정하는데 사용될 수 있다. 시스템이 깊이 카메라와 디스플레이의 상대적 위치뿐만 아니라 사용자의 머리와 눈의 상대적 위치를 알고 있으므로, 사용자가 머리를 움직여 GUI를 둘러볼 수 있도록 시스템은 머리 추적을 수행할 수 있다. 이는 게임 또는 증강 현실 애플리케이션 내에서 가상 환경을 탐험하기 위한 매우 현실적이고 직관적인 인터페이스를 제공한다.

[0043] 다른 예에서, 애플리케이션의 GUI는 식별된 물체에 기초하여 조정될 수 있는데(블록 683), 예를 들어, (도 1에 도시된 램프 또는 창과 같은)광원이 인식되는 경우, GUI는 광원의 식별된 위치로부터의 표시된 장면 내의 추가적 광을 모방하도록 조정될 수 있다. 검출된 물체가 창인 경우, (예를 들어, 날짜 및 시간에 기초하여) 창을 통한 햇빛으로부터 예상될 수 있는 여하한 반짝임을 수용하거나 이전의 예에서와 같이 추가 광을 모방하도록 GUI가 적응될 수 있다.

[0044] 스피커와 디스플레이는 능동 물체의 2가지 예일 뿐이고, (블록 608에서) 시스템은 다른 능동 물체를 인식하고 애플리케이션으로부터 그 능동 물체로의 출력을 조정할 수 있음을 인식할 것이다. (블록 685에서) 능동 물체로의 출력은 그 물체의 검출된 위치 및/또는 다른 물체의 검출된 위치에 기초하여 조정될 수 있다.

[0045] 다른 예에서, 장치는 프로젝터와 같은 식별된 능동 물체일 수 있고 프로젝터의 출력은 환경 내에서 식별된 다른 물체에 대한 그의 결정된 위치에 기초하여 조정될 수 있다. 이는 교정 단계의 필요성을 제거하고, 시스템이 식별된 능동 물체의 검출된 위치에 기초하여 자신을 교정하도록 자동으로 적응할 수 있음을 의미한다. 프로젝터는 가시적 이미지(또는 기타 GUI)를 투사할 수 있고, 예를 들어 편평하지 않거나 프로젝터에 대해 각이 진 표면을 보상하기 위해, 이들은 프로젝터의 알려진 위치와 이미지가 투사될 표면에 대한 상대적 위치에 기초하여 조정될 수 있다. 이미지는 또한, 사용자에게 대해 더 실감나는 환경을 제공하고 메인 디스플레이로부터 디스플레이를 확장하기 위해 프로젝터의 알려진 위치에 기초하여 투사를 위해 생성될 수 있는데, 메인 디스플레이는 투사된 디스플레이이거나 텔레비전, 컴퓨터 모니터 또는 기타 디스플레이 스크린에 보여질 수 있다. 구체적으로,

3D 효과가 달성되고 이미지가 임의의 표면 모양에 투사될 수 있도록, 투사된 이미지는 실시간으로 사용자의 위치(예를 들어, 얼굴과 눈의 위치)에 적응될 수 있다.

[0046] 다른 예에서, 프로젝터는 깊이 검출에 사용하기 위한 구성광 패턴을 투사할 수 있고 (블록 685에서) 투사된 구성광 패턴은 예를 들어 이동 깊이 카메라의 현재 위치에 대한 프로젝터의 알려진 위치에 기초하여 변경될 수 있다. 투사가 각을 가지고 편평한 평면에 투사되는 경우, 이미지는 투사 이미지의 여하한 관점 왜곡(perspective distortion)(키스톤(keystone) 효과라고도 함)을 보상하기 위해 사전 왜곡될 수 있다. 일 예에서, 이미지는 두 단계로 사전 왜곡된다. 먼저, 캡처된 3D 모델에서, 투사 표면의 중심에 직교하는 위치로부터의 2D 투사 이미지로 표면이 프로젝티브 텍스처링된다(projectively textured). 그러면 텍스처링된 3D 표면이 프로젝터와 동일한 외부 및 내부 렌즈를 공유하는 가상 카메라를 이용하여 다시 2D 이미지로 투사된다. 투사는 시스템이 임의의 비편평 표면 상에서의 왜곡을 정정할 수 있도록 하는 편평한 표면이 아니라 사용자에게 대해 왜곡되지 않을 수 있다. 상술한 기술은 캡처된 3D 모델에서 사용자의 추정된 눈 위치로부터 표면(비편평이라도)을 프로젝티브 텍스처링함으로써 변형될 수 있다.

[0047] 상술한 예는 모두 능동 물체였다; 그러나, 애플리케이션 성능은 가구와 같은 수동 물체 또는 실제 다른 움직임은 물체(도 1에 도시된 다른 사람(104)이나 고양이(108) 등)의 검출에 기초하여 조정될 수도 있다. 사용자가 공과 같은 가상 물체와 상호작용하고 있고 이 가상 물체를 던지는/차는 경우, 통합 엔진 내의 물리 엔진이 가상 물체와 실제 물체 사이의 상호작용을 더 정확히 모방하도록 조정될 수 있다(블록 684). 가상 공이 벽에서 튀어나오는 경우, 공은 소파(침상이라고도 함)과 같은 더 부드러운 물체로부터 튕겨나오는 경우에 비해 더 빨리 움직일 수 있다. 가상 공이 우연히 램프(106)나 실제 고양이(108)와 접촉하게 되면, 통합 엔진에 의해 적당한 음향 효과가 시뮬레이션되고 사용자에게 재생될 수 있다(예를 들어 유리 깨지는 소리나 고양이의 '야옹' 소리).

[0048] 물체 인식이 게임 플레이 (또는 기타 페이지 2에서의 사용자 상호작용) 시작 전에 수행되는 것으로 설명되지만, 추가적으로 또는 대신에, 물체 인식은 깊이 카메라가 정지한 때에 생성되는 (도 3의 블록 308에서 수신되는 바와 같은) 깊이 이미지의 제2 시리즈에 기초하여 수행될 수 있다. 이러한 예에서, 사용자가 물체와 상호작용하는 방식이 물체를 인식 또는 분류하는데 사용될 수 있고 이는 예-기반 신체 인식 방법 또는 기타 방법을 이용할 수 있다. 예를 들어, 사용자가 앉아 있는 물체는 물체 인식 모듈(502)에 의해 의자나 소파로 식별될 수 있다.

[0049] (블록 608에서) 상술한 식별된 물체에 기초하여 애플리케이션 성능이 조정 또는 적응될 수 있는 방식은 단지 일부 예를 제공하는 것임을 인식할 것이다. 일단 조밀 3D 모델로부터 물체가 식별되고 이 정보가 적당한 엔진/모듈로 공급되면, 사용자 경험을 개선하고 애플리케이션과의 사용자 상호작용의 현실성을 높이기 위하여 조정이 이루어질 수 있는 많은 방식이 있다.

[0050] 상술한 시스템은 단일 깊이 카메라, 이동 깊이 카메라(102, 204, 402)만을 포함하지만, 도 8에 도시된 바와 같이 다른 예시적 시스템은 하나 이상의 추가 깊이 카메라(802-806)도 포함할 수 있고, 이런 예시적 시스템에서 상술한 방법(예를 들어, 도 3 및 6에 도시됨)도 사용될 수 있다. 시스템의 각각의 깊이 카메라(102, 802-806)는 동일할 수 있거나, 또는 그들은 상이할 수 있다(예를 들어, 그들은 상이한 시야 모양을 가질 수 있거나, 광각(wide-angle), 숏스로우(short-throw) 및 장거리(long-range) 깊이 카메라의 조합을 포함할 수 있거나, 및/또는 그들은 구성광, 비행시간, 스테레오, 레이저 범위 파인더(laser range finder) 등과 같은 상이한 기술을 이용하는 카메라들을 포함할 수 있다). 추가의 깊이 카메라(802-806)는 정지 카메라일 수 있고 따라서 이동 깊이 카메라(102)와 상이한 모양의 하우징을 가질 수 있거나, 추가 깊이 카메라(802-806) 중 하나 이상도 이동 깊이 카메라일 수 있다. 각각의 깊이 카메라는 유선 또는 무선 접속을 이용하여 양방향 컴퓨팅 시스템에 접속된다.

[0051] 이러한 시스템의 구현에서, 이동 깊이 카메라(102) 단독으로 (도 3의 블록 304에서) 환경의 조밀 3D 모델을 생성하는데 사용될 수 있거나, 또는 다수 깊이 카메라로부터 수신된 이미지의 시리즈가 사용될 수 있다. 또한, 제2 페이지(박스 32)에서, 모든 깊이 카메라(102, 802-806) 또는 카메라의 부분집합으로부터의 이미지가 (도 3의 블록 310에서) 3D 실시간 움직임을 추적하는데 사용될 수 있다. 일 예에서, 별도의 실시간 신체 인식이 각각의 깊이 카메라에 대해 실행되고, 가구, 다른 사용자 또는 그들 자신으로부터 유발되는 사용자의 가림(occlusion)을 보상하기 위해 최고 신뢰도 값을 갖는 추적기의 결과가 사용된다.

[0052] 도 8에 도시된 바와 같이 다수의 깊이 카메라를 포함하는 시스템의 구현에서, 추가의 깊이 카메라(802-806)의 존재 및 위치를 검출하고 (도 6의 블록 608에서) 그들의 결정된 위치에 기초하여 애플리케이션 성능을 조정하기 위해 도 6에 도시된 방법이 사용될 수 있다. 사용자에게 대한 상대적 방향을 결정하기 위해 추가의 각 정지 카메라가 사용자를 개별적으로 추적할 수 있다. 카메라의 다른 정지 카메라에 대한 위치에 관한 여하한 모호성을

해결하기 위해 이 지식이 사용될 수 있다. 이 예에서, 애플리케이션 성능을 조정하는 것은 알려진 카메라 위치에 기초하여 시스템의 자동 교정을 수행하는 것을 포함할 수 있고, 이 교정은 (다수의 깊이 카메라로부터의 이미지의 시리즈가 사용되는 도 3의 동작의 제2 페이지, 박스 32에서) 실시간으로 움직임을 검출하는 때에 사용될 수 있다.

[0053] 깊이 카메라들이 동일하지 않은 경우, (도 6의) 물체 인식이 카메라의 존재와 위치 외에 카메라의 유형을 검출하기 위해 사용될 수 있다. 상술한 바와 같이, 이동 깊이 카메라가 RGB 카메라(도 4에 도시)도 포함하는 이동 환경 캡처 장치의 일부인 경우, RGB 카메라에 의해 캡처된 이미지도 카메라의 유형을 식별하는데 사용될 수 있다. 다른 예에서, 카메라의 유형을 식별하기 위해 장치 상의 특정 마커나 레이블(예를 들어, IR 반사 마커)과 같은 다른 특징이 사용될 수 있다.

[0054] 다른 변형에서, 다수의 깊이 카메라의 결정된 위치(또는 상대 위치)에 기초한 자동 교정이 3D 모델의 비교에 의해 수행될 수 있다. 도 9는, (블록 902에서) 각각의 깊이 카메라로부터 수신되는 깊이 이미지의 시리즈를 이용하여 (블록 904에서) 별개의 조밀 3D 모델이 생성되는 다수의 깊이 카메라의 자동 교정의 예시적 방법의 흐름도인데, 각각의 깊이 카메라는 이동 깊이 카메라가 있다면 이동 깊이 카메라를 포함한다. 모델은 시리즈 내의 모든 이미지를 이용하여, 또는 (예를 들어 요구되는 연산을 감소시키기 위해) 이미지의 부분집합으로부터 생성될 수 있다. 일단 생성되면, 하나 이상의 모델에서 보이는 공통 특성 세트를 식별하기 위해 개별 모델이 비교되고 (블록 906), 2 이상의 모델 내 공통 특성 세트의 식별에 기초하여, 각각의 모델과 연관된 카메라의 상대 위치가 계산될 수 있다(블록 908). 일 예에서, 환경이 여러 물체가 있는 커피 테이블과 같은 어수선한 표면을 포함하는 경우, 이러한 특성은 카메라의 상대 위치를 계산하는데 좋은 기준점을 제공하는데, 물체의 배열에 대칭성이 적거나 없을 가능성이 높기 때문이다. 도 9에 도시된 방법의 구현에서, 공통 특성 세트의 식별은 제1 모델에서의 특정 특성의 배열(어수선한 커피 테이블 등)을 식별하고 모델에서 동일한 특성의 배열이 보이는지 여부를 결정하기 위해 각각의 다른 모델을 조사함으로써 구현될 수 있다. 예시적 구현에서, 설명자 벡터(descriptor vectors)가 각각의 모델로부터 드문드문한(sparse) 관심 지점에서 추출될 수 있고, 그 후 이들이 상이한 모델 사이의 초기 대응을 위해 매칭될 수 있다. 최종적으로 기준 프레임 사이(예를 들어, 2개의 모델이 비교되고 있는 경우 2개의 기준 프레임 사이)의 변환에 대하여 그들을 정렬하기 위한 약간의 최적화가 수행된다.

[0055] 도 10은 여기 설명된 방법의 실시형태가 구현될 수 있는, 여하한 컴퓨팅 및/또는 전자 장치의 형태로 구현될 수 있는 예시적인 컴퓨팅-기반(computing-based) 장치(1000)의 다양한 컴포넌트를 도시한다.

[0056] 컴퓨팅 기반 장치(1000)는, 마이크로프로세서, 컨트롤러 또는, 깊이 이미지의 제1 시리즈로부터 환경의 조밀 3D 모델을 생성하고, 애플리케이션에서 그 모델의 적어도 일부를 이용하며, 애플리케이션으로의 입력으로 사용하기 위해 깊이 이미지의 제2 시리즈로부터 환경의 일부에서 실시간 움직임을 추적하도록 장치의 동작을 제어하는 컴퓨터 실행가능 명령을 처리하기 위한 여하한 기타 적당한 유형의 프로세서일 수 있는 하나 이상의 프로세서(1002)를 포함한다. 일부 예에서, 시스템 온 칩(system on chip) 아키텍처가 사용되는 예에서, 프로세서(1002)는 여기서 설명되는 방법의 일부를 (소프트웨어나 펌웨어가 아니라) 하드웨어에서 구현하는 하나 이상의 고정 블록(가속도계라고도 지칭)을 포함할 수 있다.

[0057] 컴퓨팅 기반 장치(1000)는 또한 그래픽 처리 시스템(1004)도 포함할 수 있는데, 이는 통신 인터페이스(1006)를 통해 프로세서(1002)와 통신하고 병렬 스트레드 연산을 빠르고 효율적인 방식으로 수행하도록 배열된 하나 이상의 그래픽 처리 유닛(GPU)(1008)을 포함한다. 그래픽 처리 시스템(1004)은 또한 GPU(1008)로부터의 빠른 병렬 액세스를 가능하게 하도록 배열되는 메모리 장치(1010)를 포함한다. 예에서, GPU는 아래에서 더 설명되는 바와 같이 조밀 3D 모델을 생성하는데 사용될 수 있다.

[0058] 컴퓨팅 기반 장치(1000)는 또한, 이동 환경 캡처 장치(깊이 카메라 포함), 깊이 카메라, 그리고 선택적으로 하나 이상의 사용자 입력 장치(예를 들어, 게임 컨트롤러, 마우스, 키보드 등)과 같은 하나 이상의 장치로부터의 입력을 수신하고 처리하도록 배열된 입력/출력(I/O) 인터페이스(1012)를 포함한다. I/O 인터페이스(1012)는 통신 인터페이스로도 동작할 수 있는데, 통신 인터페이스는 하나 이상의 통신 네트워크(예를 들어 인터넷)와 통신하도록 배열될 수 있다.

[0059] 컴퓨팅 기반 장치(1000)와 별개이거나 이와 통합될 수 있는 디스플레이 장치로 디스플레이 정보를 출력하도록 배열되는 디스플레이 인터페이스(1014)가 또한 제공된다. 디스플레이 정보는 그래픽 사용자 인터페이스, 예를 들어, 조밀 3D 모델의 적어도 일부가 통합된 애플리케이션에 대한 그래픽 사용자 인터페이스를 제공할 수 있다. 일 실시형태에서, 디스플레이 장치가 터치 감지 장치인 경우 디스플레이 장치는 사용자 입력 장치로도 동작할 수 있고, 이 경우, I/O 인터페이스(1012)와 디스플레이 인터페이스(1014)는 단일의 인터페이스로 결합될 수 있다.

거나 디스플레이 장치가 두 인터페이스 모두를 통해 접속될 수 있다.

[0060] 컴퓨팅 기반 장치(1000)에 의해 액세스가능한 여하한 컴퓨터 판독가능 매체를 이용하여 컴퓨터 실행가능 명령이 제공될 수 있다. 컴퓨터 판독가능 매체는, 예를 들어, 메모리(1016)와 같은 컴퓨터 저장 매체 및 통신 매체를 포함할 수 있다. 메모리(1016)와 같은 컴퓨터 저장 매체는, 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈 또는 기타 데이터와 같은 정보의 저장을 위한 여하한 방법이나 기술로 구현되는 휘발성 및 비휘발성, 제거가능(removable) 및 비제거가능 매체를 포함한다. 컴퓨터 저장 매체는, RAM, ROM, EPROM, EEPROM, 플래시 메모리나 기타 메모리 기술, CD-ROM, DVD(digital versatile disks)나 기타 광 저장소, 자기 카세트, 자기 테이프, 자기 디스크 저장소 또는 기타 자기 저장 장치, 또는 컴퓨팅 장치에 의한 액세스를 위하여 정보를 저장하는데 사용될 수 있는 여하한 기타 비전송(non-transmission) 매체를 포함하지만 이에 제한되지 않는다. 반면, 통신 매체는 컴퓨터 판독가능 명령, 데이터 구조, 프로그램 모듈 또는 기타 데이터를 반송파나 기타 전송 매커니즘과 같은 변조된 데이터 신호에 포함할 수 있다. 여기서 정의되는 때에, 컴퓨터 저장 매체는 통신 매체를 포함하지 않는다. 컴퓨터 저장 매체(메모리(1016))가 컴퓨팅 기반 장치(1000) 안에서 도시되지만, 저장소는 분산되거나 원격 배치되고 네트워크나 기타 통신 링크를 통해 (예를 들어, I/O 인터페이스(1012)를 이용하여) 액세스될 수 있음을 인식할 것이다.

[0061] 운영 체제(1018)를 포함하는 플랫폼 소프트웨어(1018)이나 여하한 기타 적당한 플랫폼 소프트웨어가 애플리케이션 소프트웨어(1020)가 장치 상에서 실행될 수 있게 하기 위해 컴퓨터 기반 장치에서 제공될 수 있다. 메모리(1016)는 또한 조밀 모델 통합 엔진(1022), 물체 인식 모듈(1024), 통합 엔진(1026)(예를 들어, 게임 엔진) 및 카메라 추적 엔진(1028) 중 하나 이상의 기능을 구현하기 위한 실행가능 명령을 저장할 수 있다. 메모리는, 예를 들어, 깊이 이미지의 수신된 시리즈를 저장하기 위해, 생성된 조밀 3D 모델을 저장하기 위해, 물체 인식을 수행함에 있어 사용되는 데이터를 저장하기 위해 (예를 들어, 도 5에 도시된 물체 데이터베이스를 제공하기 위해) 등으로, 프로세서(1002)에 의해 사용되는 데이터를 위한 데이터 저장소를 제공하기 위해 사용되는 데이터 저장소(1030)도 제공할 수 있다.

[0062] 도 4의 논의에서 언급한 본 출원과 같은 날 출원된 같이 계류 중인 미국 특허 출원 'Real-time Camera Tracking Using Depth Maps'가 깊이 카메라에 의해 생성된 깊이 맵을 이용하는 실시간 카메라 추적 방법을 설명하며, 이 방법의 태양이 도 4, 11 및 12를 참조하여 설명될 수 있다. 일 예시적 구현에서, 도 4의 프레임 정렬 엔진(418)은 하나 이상의 GPU(416) 또는 기타 병렬 컴퓨팅 유닛을 갖는 컴퓨팅 장치에서 구현되는 컴퓨터이다. 예를 들어, 병렬 컴퓨팅 유닛은 벡터 프로세서, SIMD(single instruction multiple data) 아키텍처, 그래픽 처리 유닛 또는 기타 병렬 컴퓨팅 장치일 수 있다. 이는 반복 최근점 프로세스(iterative closest point process) 및 선택적인 평면 추출 컴포넌트(plane extraction component)를 포함한다. 반복 최근점 프로세스는 프로젝티브 데이터 연관 및 점-대-면(point-to-plane) 에러 메트릭을 이용한다. 프레임 정렬 엔진은 깊이 카메라로부터 현재 깊이 맵을 수신한다. 이는 목적 깊이 맵(destination depth map)이라고도 지칭된다. 일부 예에서, 프레임 정렬 엔진은 깊이 카메라로부터의 이전 깊이 맵인 소스 깊이 맵도 수신한다. 다른 예에서, 프레임 정렬 엔진은 소스 깊이 맵의 조밀 표면 모델 추정을 취한다. 프레임 정렬 엔진의 출력은 현재 및 소스 프레임(또는 프레임 추정)을 정렬하기 위한 변환의 등록 파라미터(registration parameter)의 세트이고, 이들 파라미터는 도 11에 도시된 방법을 이용하여 계산될 수 있다. 일부 예에서, 이들 등록 파라미터는 SE3 매트릭스(아래에 더 자세히 설명된다)로서 제공된다; 그러나, 등록 파라미터는 여하한 적당한 형태로 제공될 수 있다. 이들 등록 파라미터는 실시간 추적기(416)에 의해 깊이 카메라의 6 자유도 포즈 추정의 실시간 시리즈를 산출하는데 사용된다.

[0063] 도 11은 프레임 정렬 엔진에서의 예시적인 반복 프로세스(iterative process)의 흐름도이다. 등록 파라미터의 초기 추정이 형성된다(1100). 이들은 현재 및 소스 프레임을 정렬하기 위한 변환의 등록 파라미터이다. 이들 초기 추정은 여하한 적당한 방법으로 형성된다. 예를 들어, 하나 이상의 다음 정보원이 초기 추정을 형성하는데 사용될 수 있다; 게임 상태, 게임 메타데이터, 맵 데이터, RGB 카메라 출력, 방향 센서 출력, GPS 데이터 등. 다른 예에서, 초기 추정은 카메라가 카메라의 이전 움직임 경로에 관한 정보를 어디서 이용하고 있는지를 예측함으로써 형성된다. 예를 들어, 카메라는 일정한 속도나 일정한 가속도를 갖는 것으로 가정될 수 있다. 시간 0에서 시간 t-1까지의 카메라의 움직임 경로가, 시간 t에서 카메라가 어디에 있을지를 추정하고 그에 따라 등록 파라미터의 추정을 획득하는데 사용될 수 있다.

[0064] 초기 추정을 이용하여, 현재 및 소스 프레임(깊이 맵 또는 추정 깊이 맵) 사이의 대응 지점의 쌍들이 계산된다(1102). 대응 지점의 쌍은 하나의 깊이 맵으로부터의 지점과 다른 깊이 맵으로부터의 지점인데, 이들 지점은 장면 내의 동일한 현실 지점으로부터 나타난 것으로 추정된다. "지점"이라는 용어는 여기서 픽셀 또는 인접하

는 픽셀의 그룹이나 패치를 지칭하는 것으로 사용된다. 수많은 가능한 지점들의 조합 때문에 이 대응 문제는 매우 어렵다. 색 또는 그레이 스케일 이미지를 이용하는 종래의 접근은 각각의 이미지 내의 선, 에지, 코너 등과 같은 모양을 식별하고 이미지 쌍 사이에서 이들 모양을 매칭하려고 시도함으로써 이 문제를 다루었다. 반면, 도 12에 도시되고 아래에 설명되는 방법은 깊이 맵에서 모양을 찾을 필요 없이 대응 지점을 식별한다. 연산된 대응 지점에 적용되는 에러 메트릭을 최적화하는 등록 파라미터의 업데이트된 추정치가 계산된다(1104). 수렴에 도달하였는지 여부에 대해 평가하기 위해 확인이 이루어진다(1106). 그렇다면, 업데이트된 추정에서 변화가 매우 적거나 없고 등록 파라미터가 출력된다(1108). 그렇지 않다면, 도 11에 도시된 바와 같이 반복 프로세스가 반복된다.

[0065] 도 12를 참조하면, 이제, 대응 지점의 쌍이 어떻게 계산되는지에 대한 추가의 세부사항이 주어진다. 각각의 사용가능한 지점에 대한 법선(surface normal)이 계산되고(블록 1202), 일 예에서, 이는 주어진 지점에 대하여, 깊이 맵에서 2개 (또는 그 이상)의 최근접 이웃 지점들을 찾고 이들 이웃과 지점 그 자체를 포함하는 표면 패치를 연산함으로써 달성된다. 그 표면 패치에 대한 법선이 그 지점의 위치에서 계산된다.

[0066] 그러면 지점의 대응하는 쌍을 찾는 프로세스가 뒤따른다(블록 1204). 이는 이제, 소스와 현재 깊이 맵이 조밀 3D 모델의 사용 없이 사용가능한 경우에서 설명된다. 소스 깊이 맵으로부터의 각각의 샘플링된 소스 지점에 대해, 소스 깊이 맵과 연관된 카메라 위치로부터 샘플링된 소스 지점을 통과하여 목적 깊이 맵의 목적 지점으로 선이 투사된다(블록 1206). 일부 경우에, 목적 지점은 투사된 선을 따라 샘플링된 소스 지점의 앞에 있을 수 있다. 이 투사 프로세스는 "프로젝티브 데이터 연관(projective data association)"이라고 지칭될 수 있다. 그 후 목적 지점의 주위 및 그 지점을 포함하는 후보 대응 지점에 대한 검색이 이루어진다(블록 1208). 예를 들어, 검색은 목적 지점의 특정 유클리드 거리(Euclidean distance) 내에 있고 샘플링된 소스 지점의 법선과 호환가능한(compatible) 법선을 갖는 지점에 대한 것이다. 법선은 서로 특정 범위 내에 있으면 호환가능하다고 말한다. 예를 들어, 이 특정 범위 및 유클리드 거리는 사용자가 구성가능하거나 및/또는 관심대상인 특정 애플리케이션 조건에 관한 경험적 데이터를 이용하여 설정될 수 있다.

[0067] (블록 1208에서) 하나 이상의 후보 대응 지점이 이 검색의 결과로 찾아진다. (블록 1210에서) 이들 대응 지점으로부터 소스 지점과 쌍을 형성할 단일 지점이 선택된다. 이 선택은 거리 메트릭에 기초하여 이루어진다. 예를 들어, 소스 지점과 각각의 후보 대응 지점 사이의 유클리드 거리가 계산된다. 그러면 최소 유클리드 거리를 가져오는 쌍이 선택된다. 그 후, 각각의 샘플링된 소스 지점에 대해, 또는 샘플링이 이루어지지 않은 경우, 소스 깊이 맵의 사용가능한 지점 각각에 대해 블록 604의 프로세스가 반복된다.

[0068] (블록 1212에서) 일부 실시형태에서 대응 지점의 쌍 각각에 가중치가 할당된다. 예를 들어, 가중치는 RGB 카메라나 기타 센서와 같은 다른 소스로부터의 정보를 이용하여 할당될 수 있다. 일 예에서, 방사상(radial) 렌즈 왜곡 및/또는 깊이 의존 에러(depth dependent error)와 같은 깊이 카메라의 측정 특성에 관한 가중치가 계산되고 대응 지점의 쌍 각각과 함께 저장된다. 다른 예에서, 에지 검출 알고리즘을 이용하여 깊이 맵에서 검출된 에지에 있는 지점을 포함하는 쌍이 다른 쌍보다 더 높은 가중치를 받는다. 이들 가중치는 결과의 품질을 개선하기 위해 에러 매트릭을 적용하는 프로세스 동안 사용될 수 있다. 예를 들어, 깊이 의존 에러에 관한 가중치는 정확성 부족으로 인하여 크게 요동하기 쉬운 높은 깊이 값이 고려될 수 있도록 한다.

[0069] 일부 실시형태에서, 깊이 맵 경계 상에 또는 그 근처에 있는 지점을 포함하는 쌍이 거부된다(블록 1214). 이는 2개의 깊이 맵 사이의 중첩이 단지 부분적인 경우에 에러를 피하는 것을 돕는다. 다른 기준이 쌍을 거부하는데 또한 사용될 수 있다. 예를 들어, 일 실시형태에서 평면 추출이 수행되고, 그 경우, 추적기가 큰 평면에 의해 편향되어서 깊이 맵 안의 작지만 고유한 부분을 무시하는 것을 막기 위해, 같은 평면에 있는 쌍이 거절될 수 있다.

[0070] 일부 실시형태에서, (블록 1200에서) 현재 및 소스 깊이 맵 중 하나 또는 둘 모두로부터 샘플 포인트가 취해지고, 이들 샘플 포인트는 대응 지점의 쌍을 찾기 위한 후보로서 사용된다. 이러한 실시형태에서, (블록 1202에서) 각각의 사용가능한 지점이 아니라 각각의 샘플링된 지점에 대해 법선이 계산되고, 그 후 블록 1204의 프로세스가 각각의 샘플링된 소스 지점에 대해 구현된다. 샘플링은 지점의 특정 비율을 임의로 선택함으로써 이루어질 수 있다. 다른 실시형태에서 (블록 1200에서) 지점의 법선을 고려하는 방식으로 샘플링이 이루어지고 따라서 블록 1202 후에 수행될 수 있다. 예를 들어, (블록 1202에서) 각각의 지점과 법선 값의 상이한 범위에 대한 복수의 빈(bin)으로 생성된 히스토그램에 대해 법선이 계산된다. 빈들에 걸쳐 균일한 샘플링이 달성되도록 샘플링이 수행된다.

[0071] 도 4의 논의에서 언급한 본 출원과 같은 날 출원된 같이 계류 중인 US 특허 출원 'Three-dimensional

environment reconstruction'은 깊이 카메라에 의해 생성된 깊이 맵과 카메라 위치 및 방향을 설명하는 데이터로부터 실제 환경의 3D 모델을 구축하는 방법을 설명하고, 이 방법의 태양이 도 4, 13 및 14를 참조하여 설명될 수 있다.

[0072] 도 13은 3D 환경 모델을 생성하기 위한 병렬화가능한(prallelizable) 프로세스의 일 예의 흐름도를 도시한다. 일 예에서, 도 13의 프로세스는 GPU 프로세서, 멀티-코어 프로세서 또는 고도의 실행 병렬성(parallelism)을 허용하는 기타 유형의 프로세서 상에서 실행된다. 다른 예에서, 충분히 빠르다면 단일 코어 프로세서가 사용될 수 있다. 먼저, (블록 1300에서) 모델을 저장하기 위한 메모리 장치 상에서 3D 부피가 생성된다. 일 예에서, 3D 부피는 그래픽 처리 시스템(예를 들어, 도 10의 시스템(1004)의 메모리 상에서 생성되는데, 이는 GPU로부터의 빠른 병렬 액세스를 가능하게 한다.

[0073] 3D 부피가 메모리의 직육면체로서 시각화될 수 있는데, 여기서 각각의 메모리 위치는 모델링되는 환경의 공간 내 지점을 나타내는 복셀(voxel)이다. 그러므로, 3D 부피는 현실 환경의 공간 비율을 직접 나타낸다. 3D 부피가 실제 부피에 직접 대응되므로, 고정 크기 메모리에서 나타내어지는 실제 부피의 크기가 모델 해상도를 결정한다. 예를 들어, 큰 현실 부피가 모델링되어야 하는 경우, 메모리의 각각의 복셀은 실제 공간 내 더 큰 영역을 나타내고, 그래서 더 작은 실제 부피가 모델링되는 경우에 비해 해상도가 낮다. 그러나 더 많은 메모리가 사용가능하면, 큰 현실 부피가 더 높은 해상도에서 모델링될 수 있다.

[0074] (블록 1300에서) 메모리에서 3D 부피가 생성되었으면, (블록 1302에서) 깊이 이미지를 캡처하는 때에 이동 환경 캡처 장치(400)로부터 깊이 이미지(414)와 깊이 카메라(402)의 6DOF 포즈 추정을 수신함으로써 모델링 프로세스가 시작한다. 6DOF 포즈 추정은 깊이 카메라(402)의 위치와 방향을 나타내고, 실시간 추적기(416)에 의해 제공될 수 있다. 일 예에서, 6DOF 포즈 추정은 실제 좌표에 대한 깊이 카메라(402)의 회전과 번역을 설명하는 SE3 매트릭스의 형태일 수 있다. 더 형식적으로는, 이 변환 매트릭스는 다음과 같이 표현될 수 있다:

[0075] 
$$T_k = \begin{bmatrix} R_k & t_k \\ 0^T & 1 \end{bmatrix} \in SE_3$$

[0076] 여기서,  $T_k$ 는 깊이 이미지 프레임  $k$ 에 대한 변환 매트릭스이고,  $R_k$ 는 프레임  $k$ 에 대한 카메라 회전이고,  $t_k$ 는 프레임  $k$ 에서의 카메라 번역이고, 유클리드 그룹  $SE_3 := \{R, t \mid R \in SO_3, t \in \mathbb{R}^3\}$ 이다. 카메라 공간에서의(즉, 카메라 관점으로부터의) 좌표는 이 변환 매트릭스를 곱함으로써 실제 좌표로 매핑될 수 있다.

[0077] (블록 1304에서) 별개 실행 스레드가 3D 부피의 평면 내 각각의 복셀에 할당될 수 있다. 예를 들어, 부피의 평면이  $z$  평면에 있으면, 실행 스레드가 3D 부피의  $x$  및  $y$  좌표의 각 조합에 할당된다. 그러면, 도 13의 점선 박스(1306)에 의해 나타내어지는 바와 같이, 각각의 별개 실행 스레드가 동일한 동작을 수행한다. 각각의 스레드에 의해 수행되는 동작은 도 14에 그림으로 예시된다. 도 14는 3D 부피(1400)의 상면도를 도시한다(즉,  $x$  축(1420)과  $z$  축(1422)을 보여주며, 명확성을 위해  $y$  축은 무시된다). 실행 스레드는,  $T_1(1402)$ ,  $T_2(1404)$ , ...,  $T_{n-1}(1406)$ ,  $T_n(1408)$ 으로 표시되는 부피의  $z$  평면 상의 각각의 좌표에 할당된다. 박스(1306)에서 각각의 스레드에 의해 수행되는 동작이 먼저  $T_1(1402)$ 를 참조하여 예시된다.

[0078] (블록 1308에서) 실행 스레드에 대한  $z$  좌표가 0으로 설정되고, (블록 1310에서) 스레드와 연관된 복셀에 대한  $x$ ,  $y$ ,  $z$  좌표가 실제 좌표로 변환된다. 예를 들어, 스레드  $T_1(1402)$ 의 경우에, 스레드와 연관된  $x$ ,  $y$  좌표는 (0,0)이고, 그래서  $z$  좌표가 0으로 설정되는 때에 이는 실제 좌표로 변환될 수 있는 (0,0,0)(즉, 도 14의 복셀 1410)의 좌표를 준다. 일단 복셀 1410에 대한 실제 좌표가 결정되었으면, 그들은 이 복셀에 대한 깊이 이미지(414) 내의 대응 위치를 결정하기 위해 깊이 카메라의 6DOF 위치 및 방향을 이용하여 변환될 수 있다. 이는, 스레드와 연관되는 현재 복셀에 대한 실제 좌표는 깊이 카메라의 투사를 통해 투사되는 관점이라는 것을 의미한다. 달리 말하면, 실제로 이 지점을 나타내는 복셀을 그 위치와 방향을 갖는 깊이 카메라에서 보면, 이는 그 지점에 대하여 깊이 이미지(414)에서 어떤 좌표인지를 결정한다. 복셀에 대한 깊이 카메라 관점 좌표는  $pX$ ,  $pY$ ,  $pZ$ 로 표시된다. 깊이 카메라 관점 좌표는  $u = pX/pZ$ 와  $v = pY/pZ$ 에 의해 깊이 이미지 깊이 좌표( $u$ ,  $v$ 로 표시)로 변환될 수 있다.

[0079] 깊이 카메라(402)의 관점은 도 14에서 깊이 카메라(402)의 프러스텀(1412)을 나타내는 점선으로 도시된다. 도 14는 깊이 카메라로부터 출력되는 깊이 이미지(414) 내의 표면(1414)을 또한 도시한다.

[0080] (블록 1314에서) 현재 스레드와 연관되는 복셀에 대한 깊이 이미지 좌표 ( $u$ ,  $v$ )는, 복셀이 깊이 카메라의 프러

스텝(1412) 내에 위치되는지 여부를 결정하는데 사용된다. 그렇지 않다면, 이 복셀이 스킵될 수 있다. 도 14를 참조하면, 스텝 T<sub>1</sub>의 경우에 현재 복셀(1410)은 카메라의 프러스텝(1412) 내에 있음을 유의하라. 이는, 예를 들어, 카메라 프러스텝(1412) 밖에 있어서 스킵되는 스텝 T<sub>n-1</sub>에 의해 고려되는 제1 복셀에 대비될 수 있다.

[0081] 현재 복셀이 카메라 프러스텝(1412) 내에 있다고 결정되면, 복셀과, 카메라의 관점에서 복셀에 대응되는 위치의 환경 내 지점 사이의 거리에 관한 인자가 결정된다. 대응 위치의 환경 내 지점이 (u,v)에서 깊이 이미지에서의 깊이 값에 의해 주어질 수 있다. 그러므로, 도 14를 참조하면, 이 인자는 복셀(1410)과 복셀(1410)로 매핑되는 깊이 이미지(414) 내 표면(1414) 상의 지점 사이의 거리(1416)를 고려한다.

[0082] 일 예에서, (블록 1320에서) 계산되는 인자는 서명된(signed) 거리 함수이다. 서명된 거리 함수 계산은 현재 복셀과 깊이 이미지(414) 내의 대응 지점 사이의 거리의 값을 주고, (카메라의 관점에서) 깊이 이미지 내 대응 지점 바깥의 (즉, 외부의) 복셀에는 양의 거리가 주어지고 (카메라의 관점에서) 깊이 이미지 내 대응 지점 안의 (즉, 내부의) 복셀에는 음의 거리가 주어지도록 서명된다. 0 값은 연관되는 복셀이 대응 지점과 정확히 일치하는(coincident) 것을 나타낸다. 서명된 거리 함수는 깊이 이미지 내 (u,v)에서의 깊이 값으로부터 pZ를 빼서 용이하게 계산될 수 있다. 예를 들어, 도 14의 복셀(1410)은 표면(1414) 상의 대응 지점의 앞에 있고, 따라서 양의 서명된 거리 함수 값을 갖는다. 역으로, 도 14의 복셀(1418)은 표면(1414) 상의 그의 대응 지점 뒤에 있고, 그래서 음의 서명된 거리 함수 값을 갖는다.

[0083] 그러면 (블록 1322에서) 서명된 거리 함수 값은 사전정의된 거리 값으로 정규화된다. 일 예에서, 이 사전정의된 값은 5cm와 같은 작은 거리일 수 있지만, 여하한 적당한 값이 사용될 수 있다. 그러면 (블록 1324에서) 정규화된 거리가 (서명된 거리가 양인 경우) 양의 문턱 값보다 크지 또는 (서명된 거리가 음인 경우) 음의 문턱 값보다 작는지 여부가 결정된다. 그렇다면, (블록 1326에서) 값을 최대화 또는 최소화하기 위해 서명된 거리 함수 값이 절단(truncated)된다. 예를 들어, 정규화된 거리가 양의 문턱 값보다 크면, 값은 +1(정규화 후의 양의 문턱 값)에서 절단될 수 있고, 정규화된 거리가 음의 문턱 값보다 작으면, 값은 -1(정규화 후의 음의 문턱 값)에서 절단될 수 있다. 이 계산의 결과는 절단된 서명된 거리 함수(truncated signed distance function)(TSDF)로 알려져 있다.

[0084] 예를 들어, 도 14를 참조하면, 복셀(1410)은 표면(1414) 밖의 상대적으로 큰 거리에 있고, 그래서 +1로 절단될 수 있다. 역으로, 복셀(1418)은 표면(1414) 안의 상대적으로 짧은 거리에 있고, 그래서 절단되지 않고 그 정규화된 거리(예를 들어, -0.6)를 유지할 수 있다.

[0085] 그러면 정규화된 (그리고, 적당하다면 절단된) 서명된 거리 함수 값은 현재 복셀에서 저장된 여하한 이전 값과 결합된다. 이것이 3D 부피로 통합되는 첫 번째 깊이 이미지인 경우, 이전의 값이 없다. 그러나, 깊이 카메라로부터의 추가의 프레임이 수신되고 통합됨에 따라 복셀에 값이 이미 존재할 수 있다.

[0086] 일 예에서, 서명된 거리 함수 값은 평균에 의해 이전 값과 결합된다(블록 1328). 이는 움직이는 물체를 갖는 환경의 모델을 구축하는 것을 도울 수 있는데, 이는 추가된 측정이 오래된 것이 되고 더 최근의 측정과 평균되면서 움직인 물체가 시간에 따라 사라지도록 할 수 있기 때문이다. 예를 들어, 지수적으로 감소하는(exponentially decaying) 이동 평균이 사용될 수 있다. 다른 예에서, 평균은 깊이 카메라로부터의 연관 복셀의 거리와 관련한 가중 함수를 이용하는 가중 평균일 수 있다. (블록 1330에서) 그후 평균된 서명된 거리 함수 값이 현재 복셀에 저장될 수 있다.

[0087] 다른 예에서, 각각의 복셀에 2개의 값이 저장될 수 있다. 서명된 거리 함수 값의 가중 합이 계산되고 저장될 수 있고, 가중치의 합도 계산되고 저장될 수 있다. 그러면 (가중 합)/(가중치 합)으로 가중 평균이 계산될 수 있다.

[0088] 그러면 (블록 1316에서) 3D 부피의 z 평면 모두가 이 스텝에 의해 고려되었는지 여부가 결정된다. 그렇지 않다면, z 좌표가 증가되고(블록 1318), 프로세스는 z 방향에서 다음 복셀에 대해 반복된다. 달리 말해, 부피의 z 방향을 따라 균등한 복셀(즉, 동일한 x, y 좌표를 갖는 것)에 걸쳐 반복되도록 스텝이 시작되어, 절단된 서명된 거리 함수 값을 결정하고 그들을 이전의 값과 평균한다.

[0089] 이는 모든 z 평면에 걸쳐 스텝이 반복될 때까지 반복된다. 이는 각각의 스텝에 의해 수행되는데, 스텝의 각각은 z 평면 상의 상이한 x, y 좌표에 해당된다. 도 5의 메모리 레이아웃에 대해, 한번에 부피의 평면 하나의 시퀀스는 양호한 메모리 액세스 효율 특성을 갖는데, 이는 하나의 z 평면이 선형 피치 어레이(linear

pitched array)의 작은 연속적인 영역에 대응하기 때문이다. 그래서, 모든 스투드에 의해 이것이 완료되면, 전체 3D 부피가 분석되었고, 깊이 이미지(414)에 대한 각각의 복셀에 대해 절단된 서명된 거리 함수가 계산되었다. 그 후 프로세스는 깊이 카메라로부터 다음 깊이 이미지 프레임이 수신되기를 기다리고, 다시 시작하여 가능한 경우, 3D 부피에 추가의 값을 더한다.

[0090] 이 프로세스의 결과는 -1과 +1 사이의 평균 값을 갖는 복셀을 포함하는 3D 부피이다(그러나 이 평균 값은 위의 "가중 합" 및 "가중치의 합" 값과 같은 다수의 값으로 나타내어질 수 있음을 유의하라). 0의 평균값은 그 위치에 환경 내 표면이 존재함을 나타낸다. 깊이 이미지의 다수의 프레임이 시간에 걸쳐 부피에 더해지기 때문에, 부피에 저장된 모델은 점진적으로 더 상세하게 구축된다. 환경의 동일한 영역의 반복된 측정은 추가의 세부사항을 더하고, 깊이 카메라에 의해 제공되는 정보의 해상도를 효과적으로 증가시킨다. 또한, 잡음이 또한 깊이 이미지로부터 효과적으로 평균되어 제거되어(averaged out), 더 부드러운 표면을 남긴다. 이는 개별 깊이 이미지로부터 볼 수 없는 세부사항이 모델에서 보일 수 있도록 한다.

[0091] 이러한 방식의 별개 실행 스투드의 사용은 전체 3D 부피에 걸쳐 연산적으로 효과적인 처리를 가능하게 된다. 예를 들어, GPU 프로세서가 효과적으로 병렬 프로그램 스투드를 실행하도록 배열되는데, 각각의 스투드가 동일한 동작을 수행하고 있다. 이는, 모델이 깊이 카메라의 프레임 레이트에서 구성되도록, 도 13을 참조하여 상술한 모델 생성 프로세스가 실시간으로 수행될 수 있도록 한다.

[0092] 본 예가 게임 시스템에서 구현되는 것으로 설명되고 도시되지만, 설명된 시스템은 제한이 아니라 예로서 제공된다. 당업자가 인식하는 바와 같이, 본 예는 증강 현실 시스템과 로보틱스 시스템, 텔레프레즌스(telepresence)와 시뮬레이션을 포함하는 다양한 상이한 유형의 양방향 컴퓨팅 시스템에서의 적용에 적합하다. 다른 예는 3D TV나 3D 프리젠테이션의 원격 디스플레이를 위한 세트의 스캐닝, 가상 여행을 위한 콘텐츠 생성, 원격 협업, 응급 서비스 및 응급 구조원(first responder)을 위한 장면의 스캐닝을 포함한다.

[0093] '컴퓨터'라는 용어는 여기서 명령을 실행할 수 있도록 처리 기능을 갖는 여하한 장치를 지칭한다. 당업자는 이러한 처리 기능이 많은 상이한 장치로 통합되고 따라서 '컴퓨터'라는 용어는 PC, 서버, 이동 전화, PDA(personal digital assistants) 및 많은 기타 장치를 포함함을 깨달을 것이다.

[0094] 여기 설명된 방법은, 예를 들어, 프로그램이 컴퓨터 상에서 실행되는 때에 여기 설명된 방법 중 여하한 것의 단계 전부를 수행하도록 구성된 컴퓨터 프로그램 코드 수단을 포함하는 컴퓨터 프로그램의 형태로, 유형(tangible) 저장 매체 상에서 기계 판독가능 형태의 소프트웨어에 의해 수행될 수 있고, 컴퓨터 프로그램은 컴퓨터 판독가능 매체 상에 포함될 수 있다. 유형(또는 비일시적(non-transitory)) 저장 매체의 예는 디스크, 섬드라이브(thumb drive), 메모리 등을 포함하고 전파되는 신호는 포함하지 않는다. 소프트웨어는 방법 단계가 여하한 적당한 순서로 또는 동시에 수행될 수 있도록 병렬 프로세스나 직렬 프로세서 상에서 실행하기에 적합할 수 있다.

[0095] 이는 소프트웨어가 가치있고 별도로 거래가능한 상품일 수 있음을 인정한다. 이는, 희망 함수를 실행하기 위해 "둔한(dumb)" 또는 표준 하드웨어 상에서 실행되거나 이를 제어하는 소프트웨어를 포함하려는 것이다. 이는 또한, 희망 기능을 수행하도록 실리콘 칩을 설계하거나 유니버설 프로그램가능 칩을 구성하기 위해 사용되는 바와 같은 HDL(hardware description language)과 같이, 하드웨어의 구성을 "설명(describe)"하거나 정의하는 소프트웨어를 포함하려는 것이다.

[0096] 당업자는 프로그램 명령을 저장하기 위해 사용되는 저장 장치가 네트워크에 걸쳐 분산될 수 있음을 깨달을 것이다. 예를 들어, 원격 컴퓨터는 소프트웨어로 설명되는 프로세스의 예를 저장할 수 있다. 로컬 또는 단말 컴퓨터는 원격 컴퓨터에 액세스하여 프로그램을 실행하기 위하여 소프트웨어의 전부나 일부를 다운로드할 수 있다. 다르게는, 로컬 컴퓨터가 필요에 따라 소프트웨어의 일부를 다운로드하거나, 로컬 터미널에서 일부 소프트웨어 명령을 실행하고 일부는 원격 컴퓨터(또는 컴퓨터 네트워크)에서 실행할 수 있다. 당업자는 또한, 당업자에게 알려진 종래의 기술을 이용함으로써, 소프트웨어 명령의 전부 또는 일부가 DSP, 프로그램가능 로직 어레이 등과 같은 전용 회로에 의해 수행될 수 있음을 깨달을 것이다.

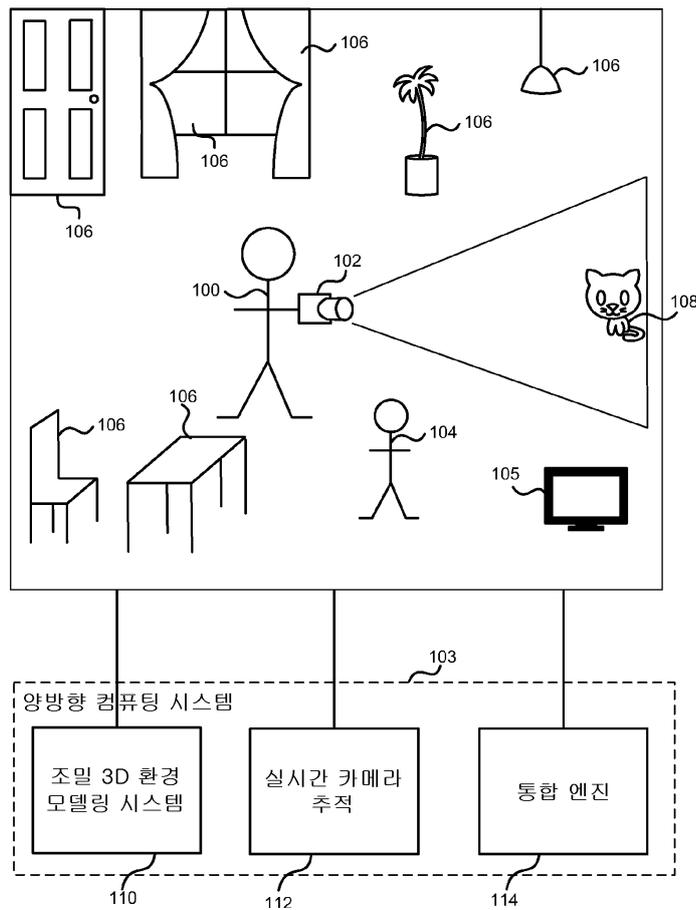
[0097] 당업자에게 명백한 바와 같이, 여기 주어진 여하한 범위 또는 장치 값은 의도한 효과를 잃지 않고 확장 또는 변경될 수 있다.

[0098] 주제가 구조적 특성 및/또는 방법론적 동작에 특정한 언어로 설명되었지만, 첨부된 청구범위에 정의된 주제는 상술한 특정 특성이나 동작으로 제한되어야 하는 것이 아님을 이해할 것이다. 오히려, 상술한 구체적인 특성 및 동작은 청구범위를 구현하는 예시적 형태로 개시된다.

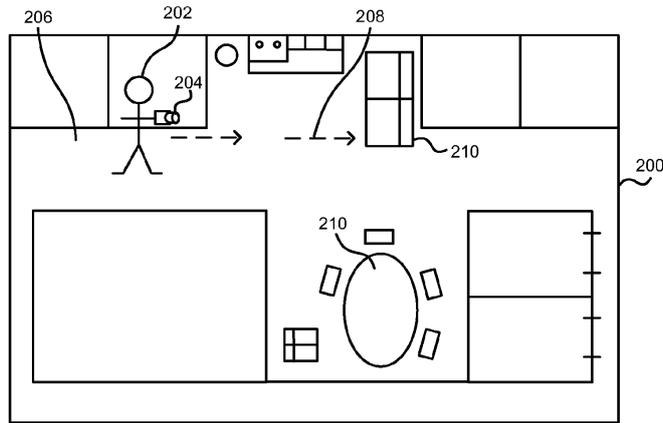
- [0099] 상술한 이점과 장점은 일 실시형태에 관한 것이거나 수개의 실시형태에 관한 것일 수 있음을 이해할 것이다. 실시형태는 서술된 문제 중 여하한 것 또는 그 전부를 해결하는 것이나 서술된 이점과 장점 중 여하한 것 또는 그 전부를 갖는 것으로 제한되지 않는다. 또한 하나의 아이템의 언급은 하나 이상의 그들 아이템을 지칭할 수 있음을 이해할 것이다.
- [0100] 여기 설명된 방법의 단계는 여하한 적당한 순서로, 또는 적절한 경우에는 동시에 수행될 수 있다. 또한, 개개의 블록은 여기 설명된 주제의 사상 및 범위를 벗어남이 없이 방법 중 여하한 것으로부터 삭제될 수 있다. 상술한 예 중 여하한 것의 태양이 설명된 다른 예의 여하한 것의 태양과 결합되어 희망 효과를 잃지 않고 다른 예를 형성할 수 있다.
- [0101] '포함한다'는 용어는 여기서, 식별된 방법 블록 또는 요소를 포함하지만, 이러한 블록이나 요소가 배타적 목록을 구성하는 것이 아니고 방법이나 장치는 추가의 블록이나 요소를 포함할 수 있음을 의미하는 것으로 사용된다.
- [0102] 바람직한 실시형태에 대한 위의 설명은 오직 예로 주어진 것이고 당업자에 의해 많은 변형이 이루어질 수 있음이 이해될 것이다. 상술한 명세, 예 및 데이터는 본 발명의 예시적 실시형태의 사용 및 구조의 완전한 설명을 제공한다. 위에서 본 발명의 다양한 실시형태가 특정 정도의 구체성으로, 또는 하나 이상의 개별 실시형태를 참조하여 설명되었으나, 당업자는 본 발명의 사상과 범위를 벗어남이 없이 개시된 실시형태에 수많은 변형을 가할 수 있을 것이다.

**도면**

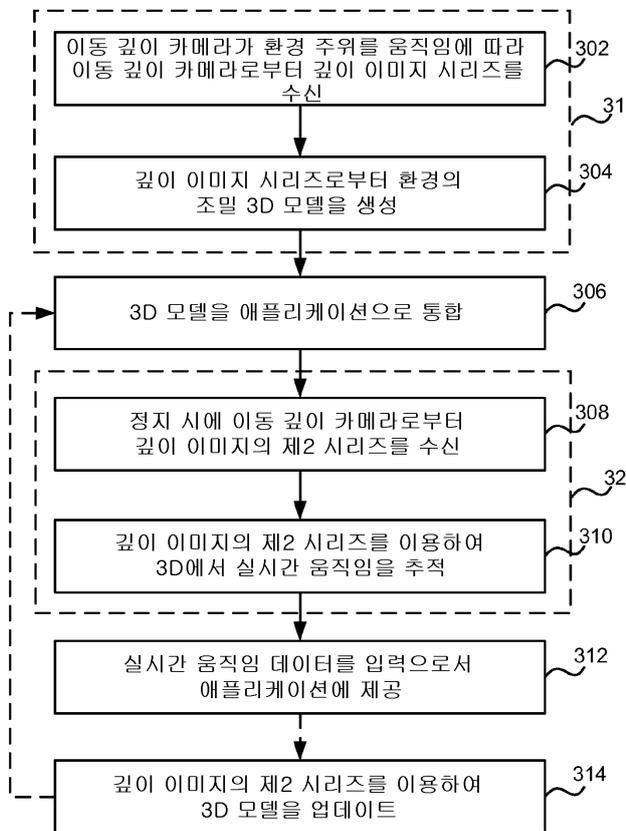
**도면1**



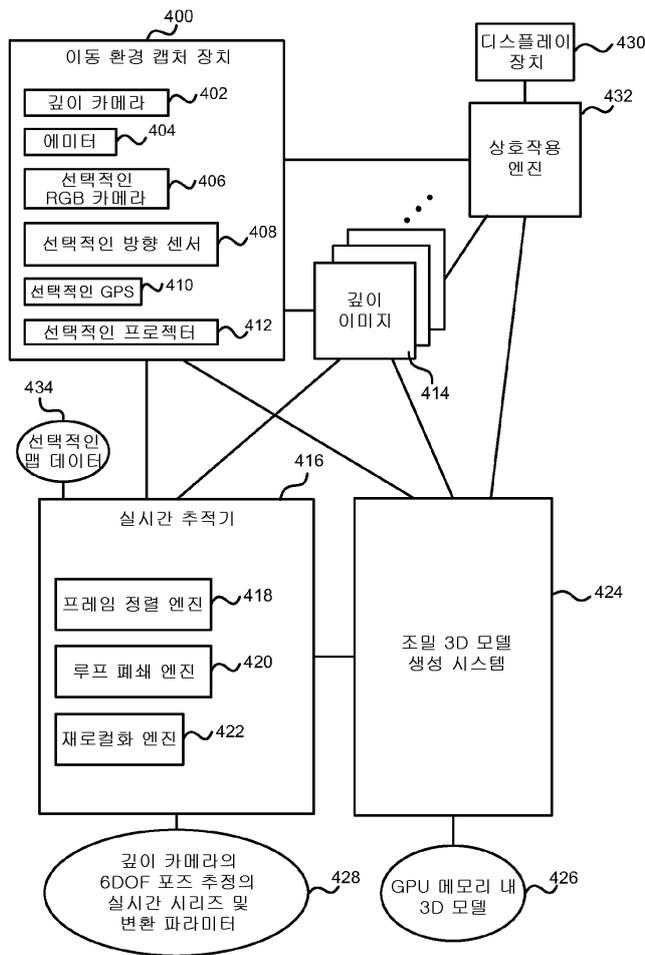
도면2



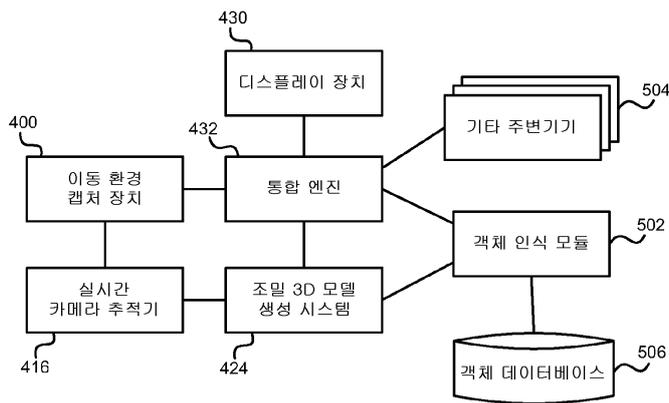
도면3



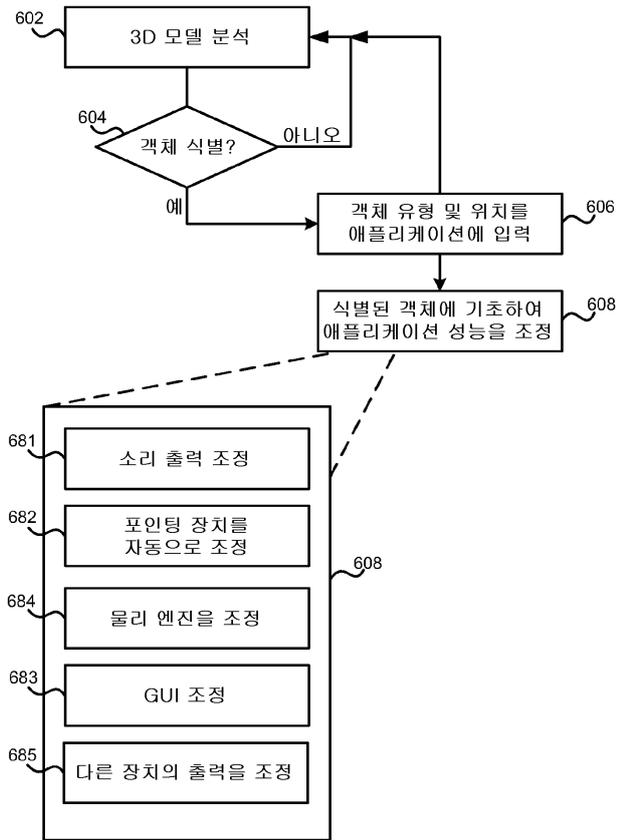
도면4



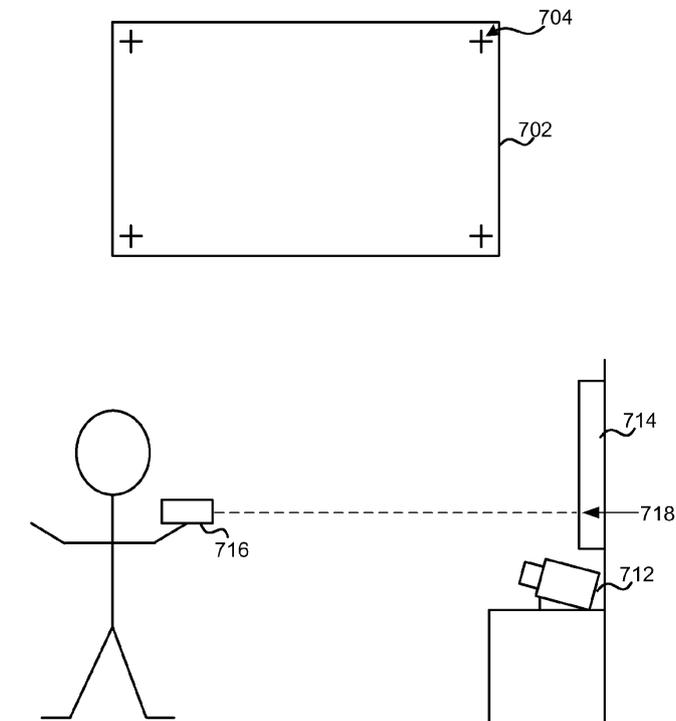
도면5



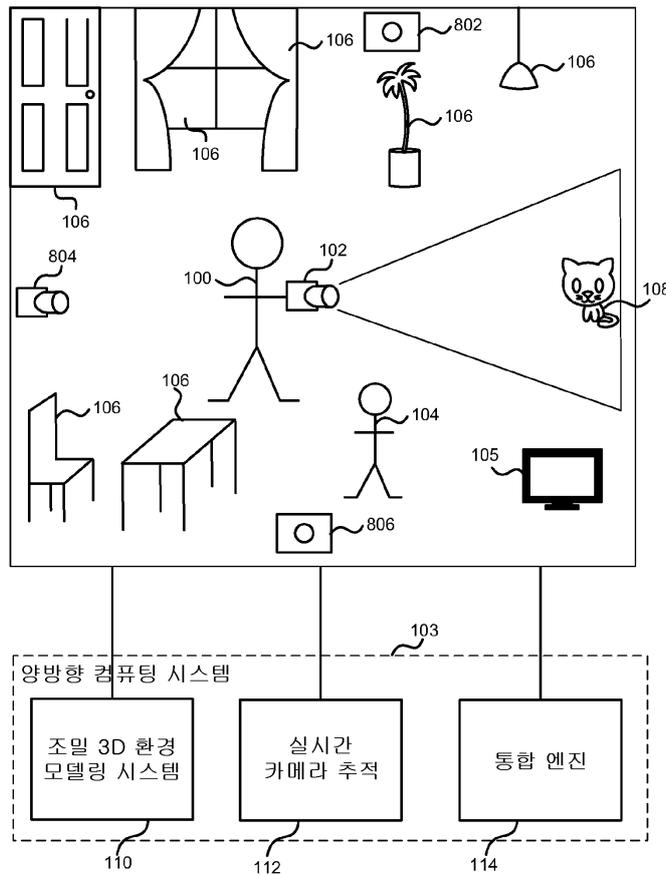
도면6



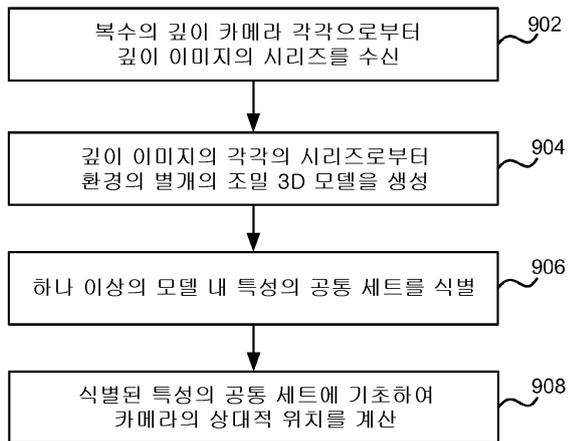
도면7



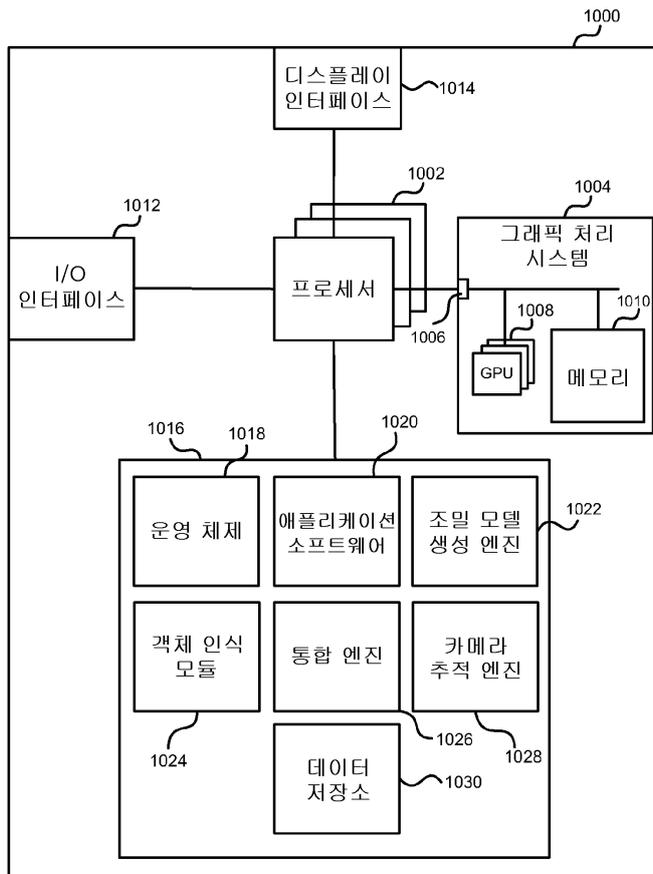
도면8



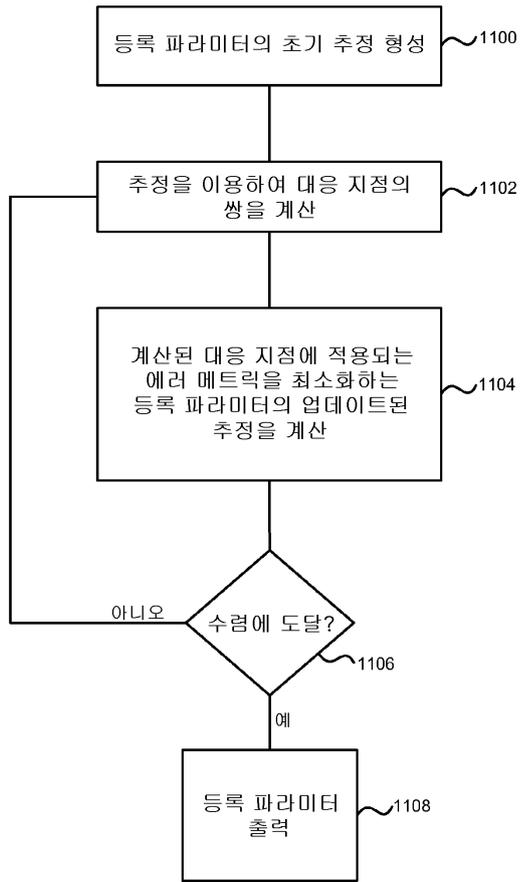
도면9



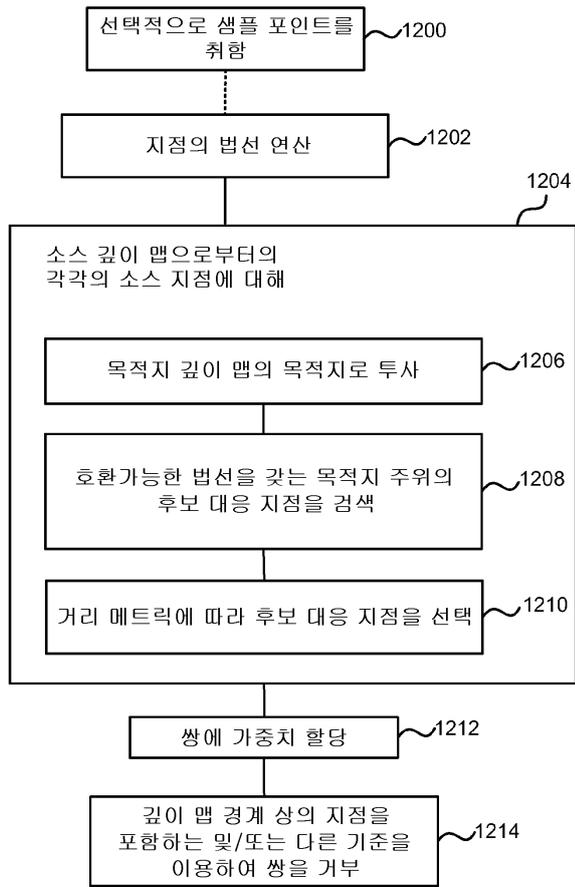
도면10



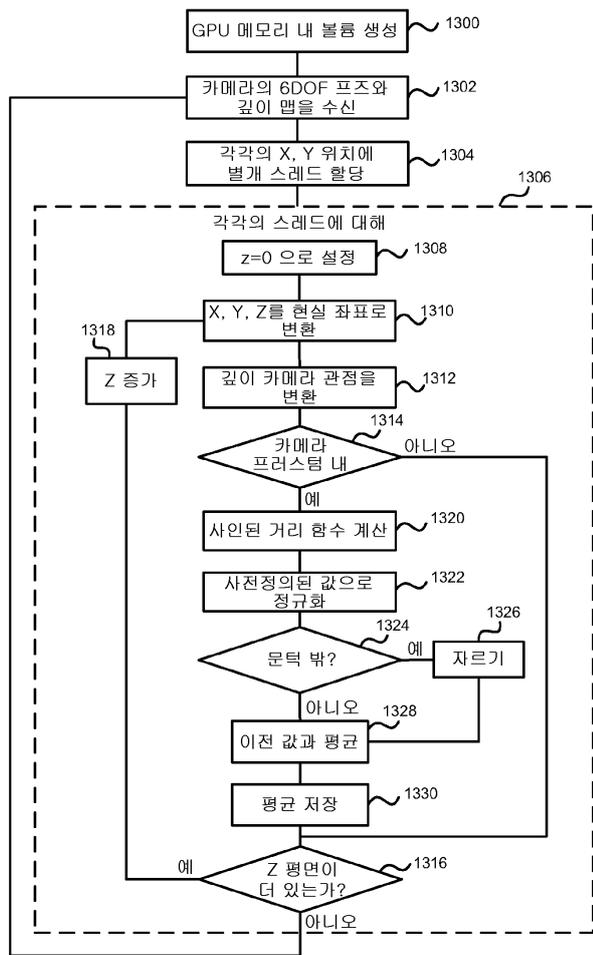
도면11



도면12



도면13



도면14

