



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0051659
(43) 공개일자 2012년05월22일

- | | |
|---|--|
| <p>(51) 국제특허분류(Int. Cl.)
A63F 13/00 (2006.01) A63F 13/06 (2006.01)
G06T 19/00 (2011.01)</p> <p>(21) 출원번호 10-2012-7002332</p> <p>(22) 출원일자(국제) 2010년07월27일
심사청구일자 없음</p> <p>(85) 번역문제출일자 2012년01월27일</p> <p>(86) 국제출원번호 PCT/US2010/043291</p> <p>(87) 국제공개번호 WO 2011/014467
국제공개일자 2011년02월03일</p> <p>(30) 우선권주장
12/511,850 2009년07월29일 미국(US)</p> | <p>(71) 출원인
마이크로소프트 코포레이션
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원 마이크로소프트 웨이</p> <p>(72) 발명자
페레즈 캐서린 스톤
미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이지즈 마이크로소프트 코포레이션
킵맨 알렉스
미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이지즈 마이크로소프트 코포레이션
(뒷면에 계속)</p> <p>(74) 대리인
제일특허법인</p> |
|---|--|

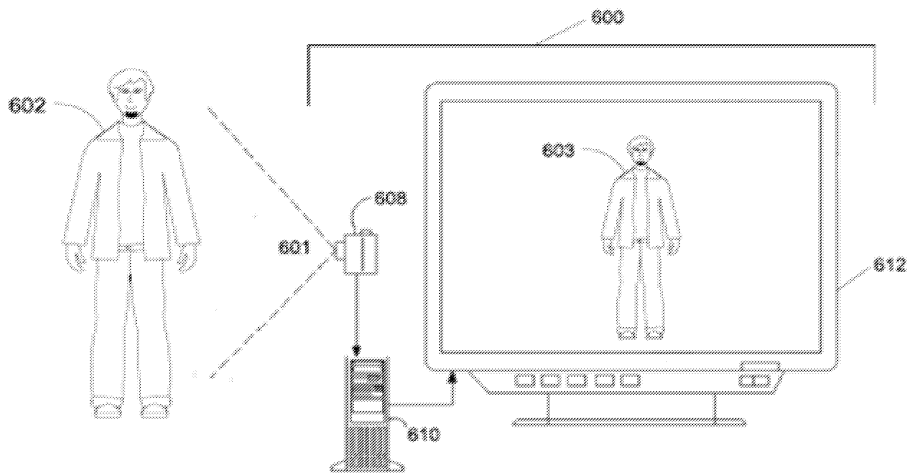
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 **비주얼 형상의 자동 생성 방법**

(57) 요약

타겟의 비주얼 형상을 자동으로 생성하는 기법은 타겟의 비주얼 형상을 생성하기 위해 요구되는 수동적인 입력을 줄이거나 제거할 수 있다. 예를 들어, 캡처 장치를 구비한 시스템은 물리적인 공간에서 사용자의 다양한 특징들을 탐지하고, 탐지된 특징들을 기초로 하여 비주얼 형상 특징 옵션 라이브러리로부터 특징을 선택할 수 있다. 시스템은 탐지된 특징들을 기초로 하여 사용자의 비주얼 형상에 선택 내역을 자동으로 적용할 수 있다. 또는, 시스템은 사용자가 선택할 수 있는 특징에 대한 옵션의 수를 좁히는 선택을 할 수도 있다. 시스템은 실시간으로 사용자에게 선택 내역을 적용하고, 선택되어 타겟의 비주얼 형상에 적용된 특징들을 실시간으로 업데이트할 수 있다.

대표도



(72) 발명자

버튼 니콜라스 디

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

월슨 앤드류

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

특허청구의 범위

청구항 1

타겟의 비주얼 형상을 생성하는 방법에 있어서,

상기 방법은

장면의 데이터를 수신하는 단계(802)와 - 상기 데이터는 물리적인 공간의 상기 타겟을 나타내는 데이터를 포함함 -,

상기 데이터로부터 적어도 하나의 타겟 특징을 탐지하는 단계(806)와,

상기 탐지된 적어도 하나의 타겟 특징을 비주얼 형상 특징 옵션들과 비교하는 단계(806)와 - 상기 비주얼 형상 특징 옵션들은 상기 타겟의 비주얼 형상에 적용되도록 구성되는 선택가능한 옵션들을 포함함 -,

상기 비주얼 형상 특징 옵션들로부터 비주얼 형상 특징을 선택하는 단계(810)와,

상기 타겟의 비주얼 형상에 상기 비주얼 형상 특징을 적용하는 단계(816)와,

상기 비주얼 형상을 렌더링하는 단계를 포함하는 방법.

청구항 2

제 1 항에 있어서,

사용자(18, 602)에 의한 수동 선택 없이 상기 비주얼 형상 특징이 선택되도록, 상기 탐지된 적어도 하나의 타겟 특징과 상기 비주얼 형상 특징 옵션들과의 비교로부터 상기 비주얼 형상이 자동으로 생성되는 방법.

청구항 3

제 1 항에 있어서,

상기 비주얼 형상 특징을 선택하는 단계는 상기 탐지된 적어도 하나의 타겟 특징과 유사한 상기 비주얼 형상 특징을 선택하는 단계(810)를 포함하는 방법.

청구항 4

제 1 항에 있어서,

상기 비주얼 형상 특징은 얼굴 특징, 신체 부위, 색상, 크기, 키, 너비, 형태, 액세스리 또는 옷 항목 중 적어도 하나인 방법.

청구항 5

제 1 항에 있어서,

상기 방법은

상기 비주얼 형상 특징 옵션들로부터 상기 비주얼 형상 특징에 관한 비주얼 형상 특징 옵션의 서브셋(702)을 생성하는 단계(810)와,

상기 비주얼 형상에 적용할 상기 비주얼 형상 특징의 사용자(18, 602)의 선택을 위해 상기 생성된 특징 옵션들의 서브셋을 제공하는 단계(810)를 더 포함하는 방법.

청구항 6

제 5 항에 있어서,

상기 생성된 비주얼 형상 특징 옵션들의 서브셋은 상기 탐지된 적어도 하나의 타겟 특징과 유사한 복수의 비주얼 형상 특징 옵션들을 포함하는 방법.

청구항 7

제 5 항에 있어서,

상기 방법은

상기 생성된 특징 옵션들의 서브셋으로부터 상기 비주얼 형상 특징의 상기 사용자(18, 602)의 선택을 수신하는 단계(812)를 더 포함하며,

상기 비주얼 형상 특징 옵션들로부터 상기 비주얼 형상 특징을 선택하는 단계는 상기 사용자(18, 602)의 선택에 해당하는 상기 비주얼 형상 특징을 선택하는 단계를 포함하는 방법.

청구항 8

제 1 항에 있어서,

상기 비주얼 형상 특징을 갖는 상기 비주얼 형상은 실시간으로 렌더링되는 방법.

청구항 9

제 1 항에 있어서,

상기 방법은

상기 타겟을 모니터링하고, 상기 탐지된 적어도 하나의 타겟 특징의 변화를 탐지하는 단계(818)와,

상기 탐지된 적어도 하나의 타겟 특징의 변화를 기초로 하여, 상기 비주얼 형상에 적용되는 상기 비주얼 형상 특징을 업데이트함으로써 상기 타겟의 상기 비주얼 형상을 실시간으로 업데이트하는 방법.

청구항 10

제 1 항에 있어서,

상기 방법은

상기 타겟이 인간 타겟이라면,

사용자(18, 602)의 양쪽 눈, 입, 코 또는 양쪽 눈썹 중 적어도 하나의 위치(position)를 탐지하고, 그 위치를 사용하여 대응 비주얼 형상 특징을 상기 비주얼 형상에 맞추는(align) 단계를 더 포함하는 방법.

청구항 11

제 1 항에 있어서,

상기 방법은

원하는 변화를 줄 수 있는 세팅을 기초로 하여, 상기 선택된 비주얼 형상 특징을 변화시키는 단계(816)를 더 포함하는 방법.

청구항 12

제 11 항에 있어서,

상기 변화는 상기 비주얼 형상 특징에 대한 다양한 변화 레벨을 제공할 수 있는 슬라이딩 스케일을 기반으로 하는 방법.

청구항 13

장면의 데이터를 수신하는 캡처 장치(20)와 - 상기 데이터는 물리적인 공간에서 타겟을 나타내는 데이터를 포함함 -,

컴퓨터 실행가능 인스트럭션들을 실행하는 프로세서(32)를 포함하며,

상기 컴퓨터 실행가능 인스트럭션들은

상기 데이터로부터 적어도 하나의 타겟 특징을 탐지하고(806),

상기 탐지된 적어도 하나의 타겟 특징을 비주얼 형상 특징 옵션들과 비교하고(806) - 상기 비주얼 형상 특징 옵션들은 비주얼 형상에 적용되도록 구성되는 선택가능한 옵션들을 포함함 -,

상기 비주얼 형상 특징 옵션들로부터 비주얼 형상 특징을 선택하고(810),

상기 타겟의 비주얼 형상에 상기 비주얼 형상 특징을 적용하는(816) 인스트럭션들을 포함하는 장치.

청구항 14

제 13 항에 있어서,

상기 장치는

상기 비주얼 형상을 실시간으로 렌더링하는(816) 디스플레이 장치(193)를 더 포함하며,

상기 프로세서(32)는 사용자(18, 602)에 의한 수동적인 선택 없이 상기 비주얼 형상 특징이 선택되도록, 상기 탐지된 적어도 하나의 타겟 특징과 상기 비주얼 형상 특징 옵션들과의 비교로부터 상기 비주얼 형상을 자동으로 생성하는 장치.

청구항 15

제 13 항에 있어서,

상기 컴퓨터 실행가능 인스트럭션들은

상기 비주얼 형상 특징 옵션들로부터 상기 비주얼 형상 특징에 관한 비주얼 형상 특징 옵션의 서브셋(702)을 생성하고(810),

상기 비주얼 형상에 적용할 상기 비주얼 형상 특징의 사용자(18, 602)의 선택을 위해 상기 생성된 특징 옵션들의 서브셋을 디스플레이 장치에 제공하는(810) 인스트럭션들을 더 포함하는 장치.

명세서

배경 기술

보통, 애플리케이션에서 사용자가 리모콘의 버튼을 선택하거나 특정 방식으로 컨트롤러를 움직이는 등의 특정 행위를 통해 제어할 수 있는, 사용자에게 대응하는 비주얼 형상이 디스플레이된다. 이러한 비주얼 형상은 아바타, 상상의 캐릭터, 만화 이미지 또는 동물, 커서, 손 등의 형태일 수 있다. 비주얼 형상은 컴퓨터 게임, 비디오 게임, 채팅, 포럼, 커뮤니티, 인스턴트 메시징 서비스 등의 다양한 애플리케이션에서 일반적으로 2D(two-dimensional) 또는 3D(three-dimensional) 모델의 형태를 갖는 컴퓨터 형상이다. 컴퓨터 게임, 멀티미

[0001]

디어 애플리케이션, 오피스 애플리케이션을 비롯한 많은 컴퓨팅 애플리케이션에서는, 사용자의 아바타로써 애플리케이션에서 사용되도록 선택되는 일련의 사전 정의된 애니메이션화된 캐릭터들을 제공한다.

[0002] 아바타의 생성을 허용하는 대부분의 시스템에서는 그 아바타에 적용할 수 있는 선택가능한 특징들에 대한 데이터베이스를 제공함으로써 캐릭터의 외관의 커스터마이제이션(customization)도 허용한다. 예를 들어, 사용자는 애플리케이션에서 이용가능한 옷 및 액세서리 보관소에 액세스하여, 아바타의 모습을 변화시킬 수 있다. 흔히, 사용자는 자신의 특징과 가장 유사한 특징들을 선택할 것이다. 예를 들어, 사용자는 자신과 유사한 신체 구조를 갖는 아바타를 선택한 후에, 특징들의 카탈로그로부터 자신과 유사한 눈, 코, 입, 머리카락 등을 선택할 수 있다. 그러나, 특징 및 각각의 특징의 옵션은 압도적으로 많은 수의 선택 옵션을 낳게 되어, 사용자의 비주얼 형상의 생성이 더욱 힘들어질 수 있다. 시스템은 사용자에게 요구되는 수고를 줄이기 위해 선택 가능한 특징의 수를 제한할 수 있지만, 이는 사용자가 독특한 아바타를 생성하기 위해 이용가능한 특징들을 제한하게 되는 바람직하지 않은 결과를 낳는다.

발명의 내용

과제의 해결 수단

[0003] 애플리케이션이나 시스템이 사용자의 비주얼 형상의 특징을 사용자를 대신하여 선택하는 것이 바람직하다. 시스템은 선택된 특징들을 사용하여 사용자의 비주얼 형상을 자동 생성할 수 있다. 예를 들어, 시스템은 사용자의 다양한 특징들을 탐지하고, 이 탐지된 특징들을 기초로 하여 특징 선택을 할 수 있다. 시스템은 탐지된 특징을 기초로 하여 사용자의 비주얼 형상에 선택 내역을 자동으로 적용할 수 있다. 또는, 시스템은 사용자가 선택할 수 있는 특징에 대한 옵션의 수를 좁히는 선택을 할 수도 있다. 시스템이 사용자를 대신하여 결정을 해줄 수 있다면, 사용자가 많은 결정을 내리거나 많은 옵션에서 선택할 필요가 없게 된다. 따라서, 본원에 개시된 기법은 사용자의 수고를 많이 덜어주고, 사용자를 대신하여 선택을 하여, 이를 사용자의 비주얼 형상에 적용할 수 있게 해준다.

[0004] 일 실시예에서, 시스템은 신체 스캔을 실행하고, 안면 인식 기법 및/또는 신체 인식 기법을 사용하여 사용자의 특징들을 식별할 수 있다. 시스템은 사용자의 식별된 특징들과 가장 유사한 특징들을 사용자의 비주얼 형상을 위해 선택한다. 다른 실시예에서, 시스템은 선택을 비주얼 형상에 적용하기 이전에 선택을 수정할 수 있다. 사용자는 시스템이 사용자의 비주얼 형상에 선택을 적용하기 전에 수정하도록 지시할 수 있다. 예를 들어, 사용자가 과체중이면, 사용자는 사용자의 비주얼 형상을 위해 시스템이 날씬한 신체 크기를 선택하도록 지시할 수 있다.

[0005] 시스템은 그 선택 내역을 사용자에게 실시간으로 적용할 수 있다. 또한, 시스템은 물리적인 공간에서 데이터를 캡처하고, 사용자의 특성을 식별하며, 사용자의 비주얼 형상의 특징을 실시간으로 업데이트하는 것이 바람직하다.

[0006] 본 요약은 아래의 상세한 설명에서 추가적으로 설명되는 일련의 컨셉을 간략화된 형태로 소개하기 위한 것이다. 본 요약은 특허청구된 대상의 핵심적인 특징 또는 필수적인 특징을 밝히기 위한 것이 아니며, 특허청구된 대상의 범위를 결정하는 데 일조하기 위해 사용되는 것도 아니다. 또한 특허청구된 대상은 본 내용에서 언급된 임의의 또는 모든 단점들을 해결하는 구현에 한정되지 않는다.

도면의 간단한 설명

[0007] 본 명세서에 따라 특징을 선택하여 비주얼 형상을 자동 생성하는 시스템, 방법 및 컴퓨터 판독가능 매체는 첨부된 도면을 참조하여 설명된다.

도 1은 게임을 하는 사용자가 있는 타겟 인식, 분석 및 트래킹 시스템의 일 실시예가 도시된다.

도 2는 타겟 인식, 분석 및 트래킹 시스템에서 사용되고, 연쇄 및 애니메이션 혼합 기법(chaining and animation blending techniques)을 사용하는 캡처 장치의 일 실시예를 도시한다.

도 3은 본원에서 설명된 애니메이션 기법들이 구현될 수 있는 컴퓨팅 환경의 일 실시예를 도시한다.

도 4는 본원에서 설명된 애니메이션 기법들이 구현될 수 있는 컴퓨팅 환경의 다른 실시예를 도시한다.

도 5는 깊이 이미지(depth image)로부터 생성된 사용자의 골격 매핑(skeletal mapping)을 도시한다.

도 6a 및 6b는 각각 타겟 인식, 분석 및 트래킹 시스템의 일례와 자동 생성된 비주얼 형상의 실시예들을 도시한다.

도 7은 타겟의 비주얼 형상에 적용가능한 특징 옵션의 서브셋을 제공하는 타겟 인식, 분석 및 트래킹 시스템의 일례를 도시한다.

도 8은 비주얼 형상 또는 비주얼 형상에 적용가능한 특징 옵션의 서브셋을 자동 생성하는 방법에 관한 순서도의 일례를 도시한다.

도 9는 타겟 디지털화 기법(target digitization techniques)을 사용하여 물리적인 공간의 타겟을 식별하는 타겟 인식, 분석 및 트래킹 시스템의 일례를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0008] 물리적인 공간의 사용자 또는 비인간 객체 등의 타겟의 비주얼 형상을 제공하는 기법이 본원에 개시된다. 예를 들어, 사용자의 비주얼 형상은 아바타, 스크린 상의 커서, 손 또는 물리적 공간에서의 사용자에게 대응하는 기타 임의의 가상 객체의 형태일 수 있다. 사람의 골격 또는 메쉬(mesh) 모델의 양태들이 캡처 장치에서 캡처된 이미지 데이터를 기반으로 하여 생성되고, 사용자의 특성을 탐지하기 위해 평가될 수 있다. 캡처 장치는 사용자의 특징을 탐지하고, 얼굴 표정, 머리카락 색상 및 타입, 피부 색 및 타입, 옷, 체형(body type), 키(height), 몸무게 등과 같이 탐지된 특징들과 닮은 특징들을 특징 카탈로그로부터 선택하여, 사용자의 비주얼 형상을 자동 생성할 수 있다. 예를 들어, 시스템은 안면 인식 기법 및 제스처/자세 인식 기법을 사용하여, 인식된 특징들에 대응하는 특징들을 특징 옵션 카탈로그 또는 데이터베이스로부터 자동으로 선택할 수 있다. 실시간으로, 시스템은 사용자의 비주얼 형상에 선택된 특징들을 적용하고 그 특징들을 업데이트할 수 있다. 유사하게, 시스템은 물리적인 공간의 비인간 타겟의 특징들을 탐지하고, 가상 객체에 대한 특징 옵션 카탈로그로부터 특징들을 선택할 수 있다. 시스템은 탐지된 특징들에 대응하는 가상 객체를 디스플레이할 수 있다.

[0009] 컴퓨팅 환경은, 인식되어 시스템에 의해 자동 생성된 비주얼 형상에 매핑된 사용자의 제스처들을 기반으로 하여, 컴퓨팅 환경에서 실행되는 애플리케이션에서 어떤 제어를 실행할지를 결정할 수 있다. 따라서, 가상 사용자가 디스플레이되고, 사용자는 물리적인 공간에서 제스처를 취함으로써 가상 사용자의 모션을 제어할 수 있다. 캡처된 모션은 카메라와 같은 캡처 장치로 캡처된 물리적 공간에서의 임의의 모션일 수 있다. 이 캡처된 모션은 사용자 또는 사물(object)과 같은 물리적 공간에서의 타겟의 모션을 포함할 수 있다. 캡처된 모션은 운영 체제 또는 애플리케이션에서의 제어로 해석되는 제스처를 포함할 수 있다. 모션은 달리는 모션과 같이 동적일 수도 있고, 또는 거의 움직임이 없는 자세를 취하는 사용자와 같이 정적일 수도 있다.

[0010] 탐지가능한 사용자 특성을 기초로 하여 비주얼 형상을 선택하는 시스템, 방법, 기법 및 컴포넌트들은 게임 콘솔과 같은 멀티미디어 콘솔 또는, 예를 들어 그러나 이에 제한되지 않는, 위성 수신기, 셋톱 박스, 아케이드 게임(arcade games), PC(personal computers), 휴대 전화, PDA(personal digital assistants) 및 기타 핸드헬드 장치들을 포함하는 타겟의 비주얼 형상을 디스플레이하기 위한 기타 임의의 컴퓨팅 장치에서 구현될 수 있다.

[0011] 도 1은 사용자의 특징을 아바타에 적용하는 기법을 사용하는 타겟 인식, 분석 및 트래킹 시스템(10)의 구성에 대한 일 실시예를 도시한다. 일 실시예에서, 사용자(18)는 복싱 게임을 하고 있다. 일 실시예에서, 시스템(10)은 사용자(18)와 같은 인간 타겟을 인식, 분석 및/또는 트래킹할 수 있다. 시스템(10)은 물리적 공간의 사용자의 모션, 얼굴 표정, 바디 랭귀지, 감정 등에 관련된 정보를 수집할 수 있다. 예를 들어, 시스템은 인간 타겟(18)을 식별하고 스캔할 수 있다. 시스템(10)은 자세 인식 기법을 사용하여 인간 타겟(18)의 체형을 식별할 수 있다. 시스템(10)은 사용자(18)의 신체 부위와, 이들이 어떻게 움직이는지를 식별할 수 있다. 시스템(10)은 탐지된 사용자 특징들과 선택가능한 비주얼 형상 특징들의 카탈로그를 비교할 수 있다.

[0012] 도 1에 도시된 바와 같이, 타겟 인식, 분석 및 트래킹 시스템(10)은 컴퓨팅 환경(12)을 포함할 수 있다. 컴퓨팅 환경(12)은 컴퓨터, 게임 시스템 또는 콘솔 등일 수 있다. 일 실시예에 따르면, 컴퓨팅 환경(12)은 게임 애플리케이션, 비게임 애플리케이션 등의 애플리케이션들을 실행하는 데에 사용될 수 있도록 하드웨어 컴포넌트 및/또는 소프트웨어 컴포넌트들을 포함할 수 있다.

[0013] 도 1에 도시된 바와 같이, 타겟 인식, 분석 및 트래킹 시스템(10)은 캡처 장치(20)를 더 포함할 수 있다. 캡처 장치(20)는 예를 들어, 사용자(18)를 비롯한 한 명 이상의 사용자들을 시각적으로 모니터링하여, 이 사용

자들이 행하는 제스처를 캡처, 분석 및 트래킹해서 애플리케이션 내의 여러 제어나 액션을 실행할 수 있는 카메라일 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다.

- [0014] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템(10)은 사용자(18)와 같은 사용자에게 게임 또는 애플리케이션 비주얼 및/또는 오디오를 제공하는 텔레비전, 모니터, HDTV(high-definition television) 등의 오디오 비주얼 장치(16)에 연결될 수 있다. 예를 들어, 컴퓨팅 환경(12)은 게임 애플리케이션, 비게임 애플리케이션 등과 관련된 오디오비주얼 신호를 제공하는 그래픽 카드 등의 비디오 어댑터 및/또는 사운드 카드 등의 오디오 어댑터를 포함할 수 있다. 오디오비주얼 장치(16)는 컴퓨팅 환경(12)으로부터 오디오비주얼 신호를 수신하여, 이 오디오비주얼 신호와 관련된 게임 또는 애플리케이션의 비주얼 및/또는 오디오를 사용자(18)에게 제공한다. 일 실시예에 따르면, 오디오비주얼 장치(16)는 예를 들어, S-Video 케이블, 동축 케이블, HDMI 케이블, DVI 케이블, VGA 케이블 등을 통해 컴퓨팅 환경(12)에 연결될 수 있다.
- [0015] 도 1에 도시된 바와 같이, 타겟 인식, 분석 및 트래킹 시스템(10)은 사용자(18) 등의 인간 타겟을 인식, 분석 및/또는 트래킹하기 위해 사용될 수 있다. 예를 들어, 사용자(18)는 캡처 장치(20)를 통해 트래킹되어, 사용자(18)의 움직임이 컴퓨터 환경(12)에서 실행 중인 애플리케이션에 영향을 주는 제어로써 해석될 수 있다. 따라서, 일 실시예에 따르면, 사용자(18)는 자신의 신체를 움직임으로써 애플리케이션을 제어할 수 있다. 시스템(10)은 사용자의 신체 및, 애플리케이션, 운영 체제 등의 시스템의 양태를 제어하는 제스처를 포함하는, 사용자의 신체로 하는 모션들을 트래킹할 수 있다.
- [0016] 시스템(10)은 캡처 장치(20)로의 입력을 애니메이션으로 변환하여, 애니메이션이 이 입력에 의해 구동되게 할 수 있으며, 여기서 입력은 사용자의 모션을 나타낸다. 따라서, 사용자의 모션을 아바타(40)에 매핑시켜, 물리적 공간에서의 사용자의 모션들이 아바타(40)에 의해 실행되도록 할 수 있다. 사용자의 모션들은 애플리케이션에서 제어에 해당하는 제스처들일 수 있다. 도 1에 도시된 바와 같이, 일 실시예에서, 컴퓨팅 환경(12)에서 실행되는 애플리케이션은 사용자(18)가 하고 있는 복싱 게임일 수 있다.
- [0017] 컴퓨팅 환경(12)은 오디오비주얼 장치(16)를 사용하여, 사용자(18)가 그 자신의 움직임으로 제어할 수 있는 플레이어 아바타(40)의 비주얼 형상을 제공할 수도 있다. 시스템은, 사용자의 탐지된 특징들에 기반하여 시스템에 의해 자동 생성된 자동 생성 비주얼 형상일 수 있는 사용자의 비주얼 형상에 모션 및/또는 제스처를 적용할 수 있다. 예를 들어, 사용자(18)는 물리적인 공간에서 펀치를 날려, 플레이어 아바타(40)가 게임 공간에서 펀치를 날리게 할 수 있다. 플레이어 아바타(40)는 캡처 장치(20)에 의해 식별된 사용자 특성을 가질 수 있고, 또는 시스템(10)은 사용자의 모션에 일치하는 비주얼 형상에 유명한 복싱 선수의 특징을 사용하거나 프로 복싱 선수의 체격을 묘사할 수 있다. 시스템(10)은 사용자를 트래킹하여, 물리적인 공간에서 탐지가능한 사용자 특징을 기반으로 하여 사용자 아바타의 특성을 바꿀 수 있다. 또한, 컴퓨팅 환경(12)은 오디오비주얼 장치(16)를 사용하여 사용자(18)에게 복싱 상대(38)의 비주얼 형상을 제공할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템(10)의 컴퓨팅 환경(12) 및 캡처 장치(20)가 물리적 공간에서의 사용자(18)의 펀치를 인식하고 분석하는 데에 사용되어, 그 펀치가 게임 공간에서 플레이어 아바타(40)의 게임 제어로 해석될 수 있다. 다수의 사용자들이 원격 위치에서 서로 인터랙션할 수도 있다. 예를 들어, 복싱 상대(38)의 비주얼 형상은 사용자(18)와 같은 물리적 공간에서의 제 2 사용자 또는 제 2 물리적 공간에서의 네트워크 연결된 사용자 등 다른 사용자를 나타낼 수 있다.
- [0018] 또한, 사용자(18)의 다른 움직임들은 예를 들어, 보빙(bob), 위빙(weave), 셔플(shuffle), 블록(block), 잭(jab)을 하거나 또는 다양한 다른 파워 펀치들을 날리기 위한 제어와 같이 다른 제어나 액션로 해석될 수 있다. 나아가, 일부 움직임들은 플레이어 아바타(40)의 제어가 아닌 액션들에 해당하는 제어로 해석될 수 있다. 예를 들어, 플레이어가 움직임을 통해 게임의 종료, 일시 중지 또는 저장, 레벨 선택, 최고점 보기, 친구와의 통신 등을 할 수 있다. 나아가, 사용자(18)의 모든 움직임은 애플리케이션과 인터랙션하도록 적절한 방식으로 이용 가능, 사용 및 분석될 수 있다.
- [0019] 실시예들에서, 사용자(18)와 같은 인간 타겟은 사물을 가질 수 있다. 이러한 실시예에서는, 전자 게임의 사용자가 사물을 쥐어, 플레이어와 사물의 모션들이 게임의 파라미터들을 조정 및/또는 제어하는 데에 사용될 수 있다. 예를 들어, 라켓을 쥐는 플레이어의 모션이 트래킹되어, 전자 스포츠 게임의 온스크린 라켓을 제어하는 데에 이용될 수 있다. 또 다른 실시예에서, 사물을 쥐고 있는 플레이어의 모션이 트래킹되어, 전자 컴뱃 게임에서 온스크린 무기를 제어하는 데에 이용될 수 있다.
- [0020] 사용자의 제스처 또는 모션은 플레이어 아바타(40)의 제어가 아닌 액션들에 해당하는 제어들로 해석될 수도 있다. 예를 들어, 플레이어가 움직임을 통해 게임의 종료, 일시 중지 또는 저장, 레벨 선택, 최고점 보기, 친

구와의 통신 등을 할 수 있다. 플레이어는 움직임 을 통해 아바타를 변화시킬 수 있다. 예를 들어, 사용자가 물리적인 공간에서 자신의 팔을 흔들면, 이는 아바타의 팔을 길게 만들어 달라는 요청으로 시스템(10)이 인식 하는 제스처일 수 있다. 운영 체제 및/또는 애플리케이션의 제어가능한 사실상 모든 양태가 사용자(18)와 같 은 타겟의 움직임에 의해 제어될 수 있다. 다른 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템(10)은 타겟 움직임들을 게임 세상 밖의 운영 체제 및/또는 애플리케이션의 양태를 제어하는 것으로 해석할 수 있다.

[0021] 사용자의 제스처는 운영 체제, 게임의 비게임 양태 또는 비게임 애플리케이션에 적용되는 제어들일 수 있다. 사용자의 제스처는 사용자 인터페이스의 제어와 같이 사물 조작으로 해석될 수 있다. 예를 들어, 왼쪽에서 오른쪽으로 수직으로 정렬된 블레이드 또는 탭 인터페이스를 갖는 사용자 인터페이스에서, 각 블레이드 또는 탭 을 선택하면 애플리케이션 또는 시스템 내의 다양한 제어들에 대한 옵션들이 펼쳐진다. 시스템은 사용자 손의 움직임을 탭의 움직임으로 파악하며, 여기서 물리적 공간의 사용자 손은 애플리케이션 공간의 탭과 가상으로 맞춰져 있다. 잠시 멈춤, 잡는 모션 및 손을 왼쪽으로 내미는 것을 포함하는 제스처는 탭의 선택 및 그 탭을 지나쳐 다음 탭을 여는 것으로 해석될 수 있다.

[0022] 도 2는 타겟 인식, 분석 및 트래킹에 사용되는 캡처 장치(20)의 일 실시예를 도시하며, 여기서 타겟은 사용자 또는 사물일 수 있다. 일 실시예에 따르면, 캡처 장치(20)는 비행 시간(time-of-flight), 구조광(structured light), 스테레오 이미지 등을 포함하는 적절한 임의 기술을 통해, 깊이 값을 포함하는 깊이 이미지 등 깊이 정보가 있는 비디오를 캡처하도록 구성될 수 있다. 일 실시예에 따르면, 캡처 장치(20)는 "Z 레이어들", 또는 깊이 카메라로부터 시선을 따라 연장되는 Z 축에 수직인 레이어들로 깊이 정보를 체계화시킬 수 있다.

[0023] 도 2에 도시된 바와 같이 캡처 장치(20)는 이미지 카메라 컴포넌트(22)를 포함할 수 있다. 일 실시예에 따르면, 이미지 카메라 컴포넌트(22)는 장면의 깊이 이미지를 캡처하는 깊이 카메라일 수 있다. 깊이 이미지는 캡 처된 장면의 2-D(two-dimensional) 화소 영역을 포함할 수 있고, 2-D 화소 영역의 각 화소는 카메라로부터 캡 처된 장면에서 사물의 길이 또는 거리 등의 깊이 값을 센티미터, 밀리미터 등으로 나타낼 수 있다.

[0024] 도 2에 도시된 바와 같이, 일 실시예에 따르면, 이미지 카메라 컴포넌트(22)는 장면의 깊이 이미지를 캡처하 기 위해 사용되는 적외선 컴포넌트(IR light component, 24), 3-D(three-dimensional) 카메라(26) 및 RGB 카 메라(28)를 포함할 수 있다. 예를 들면, 비행 시간 분석법에서는, 캡처 장치(20)의 적외선 컴포넌트(24)가 장 면으로 적외선을 내보내고, 센서들(도시되지 않음)을 사용하여 3-D 카메라(26) 및/또는 RGB 카메라(28)를 사 용한 장면의 하나 이상의 타겟 및 사물들의 표면으로부터 후방 산란되는 빛을 검출한다. 몇몇 실시예에서는, 펄스 적외선(pulsed infrared light)을 사용하여, 나가는 광 펄스와 이에 대응하여 들어오는 광 펄스 간의 시 간을 측정하고, 이를 사용하여 캡처 장치(20)로부터 장면의 타겟 또는 사물들의 특정 지점까지의 물리적인 거 리를 결정할 수 있다. 또한, 또 다른 실시예에서는, 나가는 광파의 위상과 들어오는 광파의 위상을 비교하여 위상 변위(phase shift)를 결정할 수 있다. 이후, 이 위상 변위를 사용하여 캡처 장치(20)로부터 타겟 또는 사물의 특정 지점까지의 물리적인 거리를 결정할 수 있다.

[0025] 다른 실시예에 따르면, 비행 시간 분석법을 사용하여 셔터 광 펄스 이미징(shuttered light pulse imaging)을 포함하는 다양한 기술을 통해 시간에 따른 반사광의 강도를 분석함으로써 캡처 장치(20)로부터 타겟 또는 사 물의 특정 지점까지의 물리적인 거리를 간접적으로 결정할 수 있다.

[0026] 또 다른 실시예에서, 캡처 장치(20)는 구조광을 사용하여 깊이 정보를 캡처할 수 있다. 이러한 분석법에서는, 패턴화된 광(즉, 그리드 패턴이나 스트라이프 패턴과 같이 공지된 패턴으로 디스플레이되는 광)이 적외선 컴 포넌트(24)를 통해 장면에 투사될 수 있다. 그 장면의 하나 이상의 타겟 또는 사물의 표면에 부딪치면, 패턴 이 그에 따라 변형되게 된다. 이러한 패턴의 변형이 예컨대, 3-D 카메라(26) 및/또는 RGB 카메라(28)에 의해 캡처된 후에 분석되어, 캡처 장치(20)로부터 타겟 또는 사물의 특정 지점까지의 물리적인 거리를 결정할 수 있다.

[0027] 또 다른 실시예에 따르면, 캡처 장치(20)는, 깊이 정보의 생성을 위해 분석되는 비주얼 스테레오 데이터를 얻 기 위해 각기 다른 각도에서 장면을 볼 수 있는 물리적으로 분리된 둘 이상의 카메라들을 포함할 수 있다.

[0028] 또 다른 실시예에서, 캡처 장치(20)는 포인트 클라우드 데이터(point cloud data) 및 타겟 디지털화 기법을 사용하여 사용자의 특징들을 탐지할 수 있다. 이러한 기법에 대해서는 도 6a와 관련하여 아래에서 보다 상세 하게 설명된다.

[0029] 캡처 장치(20)는 마이크(30) 또는 마이크 어레이를 더 포함할 수 있다. 마이크(30)는 소리를 수신하여 이를 전기 신호로 변환하는 변환기(transducer) 또는 센서를 포함할 수 있다. 일 실시예에 따르면, 마이크(30)는

타겟 인식, 분석 및 트래킹 시스템(10)의 캡처 장치(20)와 컴퓨팅 환경(12) 사이의 피드백을 감소시키기 위해 사용될 수 있다. 게다가, 컴퓨팅 환경(12)에서 실행가능한 게임 애플리케이션, 비게임 애플리케이션 등의 애플리케이션들을 제어하기 위해 사용자로부터 제공되는 오디오 신호를 수신하기 위해 마이크(30)가 사용될 수 있다.

[0030] 일 실시예에서, 캡처 장치(20)는 이미지 카메라 컴포넌트(22)와 통신 가능한 논리 프로세서(32)를 더 포함할 수 있다. 프로세서(32)는 깊이 이미지를 수신하고, 그 깊이 이미지에 적절한 타겟의 포함 여부를 결정하고, 적절한 타겟을 타겟의 골격 형상 또는 모델로 변환하는 인스트럭션을 포함하는 인스트럭션들 또는 기타 임의의 적절한 인스트럭션을 실행하는 표준화된 프로세서, 특화된 프로세서, 마이크로 프로세서를 포함할 수 있다.

[0031] 예를 들어, 컴퓨터 관독가능 매체는 장면의 데이터를 수신하기 위한 컴퓨터 실행가능 인스트럭션들을 포함하고, 여기서 데이터는 물리적인 공간의 타겟을 나타내는 데이터를 포함한다. 인스트럭션은 그 데이터로부터 적어도 하나의 타겟 특징을 탐지하고, 이 탐지된 특징을 특징 라이브러리(features library, 197)의 비주얼 형상 특징 옵션들과 비교하기 위한 인스트럭션들을 포함할 수 있다. 비주얼 형상 특징 옵션들은 비주얼 형상에 적용될 수 있도록 구성된 선택가능한 옵션들을 포함할 수 있다. 비주얼 형상 특징 옵션들로부터 비주얼 형상 특징을 선택하고, 이 비주얼 형상 특징을 타겟의 비주얼 형상에 적용하여, 비주얼 형상을 렌더링하기 위해 추가적인 인스트럭션들이 제공된다. 비주얼 형상은 탐지된 적어도 하나의 특징을 비주얼 형상 특징 옵션들과 비교함으로써 자동생성되어, 사용자의 수동 선택(manual selection) 없이 비주얼 형상이 선택될 수 있다.

[0032] 비주얼 형상 특징의 선택에는 탐지된 타겟 특징과 유사한 비주얼 형상 특징을 선택하는 것이 포함될 수 있다. 비주얼 형상 특징은 얼굴 특징, 신체 부위, 색상, 크기, 키, 너비, 형태, 액세서리 또는 옷 중 적어도 하나를 포함할 수 있다. 비주얼 형상 특징 옵션들로부터 비주얼 형상 특징에 대한 서브셋(subset)을 생성하고, 이 생성된 특징 옵션들의 서브셋을 비주얼 형상에 적용되는 비주얼 형상 특징의 사용자 선택을 위해 제공하는 인스트럭션들이 제공될 수 있다. 생성된 비주얼 형상 특징 옵션들의 서브셋은 탐지된 타겟 특징과 유사한 다수의 비주얼 형상 특징 옵션들을 포함할 수 있다. 생성된 특징 옵션들의 서브셋으로부터 사용자가 선택한 비주얼 형상을 수신하기 위한 인스트럭션이 제공될 수 있으며, 여기서 비주얼 형상 특징 옵션들로부터 비주얼 형상 특징을 선택하는 것은 사용자 선택에 대응하는 비주얼 형상 특징을 선택하는 것을 포함한다. 비주얼 형상 특징을 갖는 비주얼 형상이 실시간으로 렌더링될 수 있다. 또한, 타겟을 모니터링하고, 탐지된 타겟 특징의 변화를 탐지하며, 탐지된 타겟 특징의 변화를 기초로 하여, 실시간으로, 비주얼 형상에 적용되는 비주얼 형상 특징을 업데이트함으로써 타겟의 비주얼 형상을 업데이트하기 위한 인스트럭션들이 제공될 수 있다.

[0033] 캡처 장치(20)는, 프로세서(32)로 실행가능한 인스트럭션, 3-D 카메라(26) 또는 RGB 카메라(28)로 캡처된 이미지 또는 이미지 프레임, 또는 기타 임의의 적절한 정보, 이미지 등을 저장하는 메모리 컴포넌트(34)를 더 포함할 수 있다. 일 실시예에 따르면, 메모리 컴포넌트(34)는 RAM(random access memory), ROM(read only memory), 캐시(cache), 플래시 메모리, 하드 디스크 또는 기타 적절한 저장 컴포넌트를 포함한다. 도 2에 도시된 바와 같이, 일 실시예에서, 메모리 컴포넌트(34)는 이미지 카메라 컴포넌트(22) 및 프로세서(32)와 통신하는 별개의 컴포넌트일 수 있다. 또 다른 실시예에 따르면, 메모리 컴포넌트(34)가 프로세서(32) 및/또는 이미지 카메라 컴포넌트(22)에 통합될 수도 있다.

[0034] 도 2에 도시된 바와 같이, 캡처 장치(20)는 통신 링크(36)를 통해 컴퓨팅 환경(12)과 통신할 수 있다. 통신 링크(36)는 USB 연결, 파이어와이어(Firewire) 연결, 이더넷 케이블 연결 등을 포함하는 유선 연결 및/또는 무선 802.11 b, g, a 또는 n 연결 등의 무선 연결일 수 있다. 일 실시예에 따르면, 컴퓨팅 환경(12)은 장면을 언제 캡처할 지를 결정하는 데에 사용되는 클럭을 통신 링크(36)를 통해 캡처 장치(20)로 제공할 수 있다.

[0035] 또한, 캡처 장치(20)는 예를 들어, 3-D 카메라(26) 및/또는 RGB 카메라(28)에서 캡처되는 깊이 정보 및 이미지들, 및/또는 캡처 장치(20)로부터 생성되는 골격 모델을 통신 링크(36)를 통해 컴퓨팅 환경(12)으로 제공할 수 있다. 이후에, 컴퓨팅 환경(12)은 골격 모델, 깊이 정보 및 캡처된 이미지를 사용하여, 예컨대, 게임 또는 워드 프로세서 등의 애플리케이션을 제어할 수 있다. 예를 들어, 도 2에 도시된 바와 같이, 컴퓨팅 환경(12)은 제스처 라이브러리(192)를 포함할 수 있다.

[0036] 도 2에 도시된 바대로, 컴퓨팅 환경(12)은 제스처 라이브러리(192) 및 제스처 인식 엔진(190)을 포함할 수 있다. 제스처 인식 엔진(190)은 제스처 필터(191)의 컬렉션을 포함할 수 있다. 필터는 제스처를 인식할 수 있는 코드 및 관련 데이터를 포함하고, 깊이, RGB 또는 골격 데이터를 처리할 수 있다. 각 필터(191)는 제스처를 정의하는 정보와 그 제스처에 대한 파라미터 또는 메타데이터를 함께 포함할 수 있다. 예를 들어, 한 손이 몸

뒤로부터 몸 앞을 지나는 모션을 포함하는 던지기 동작은 사용자의 한 손이 몸 뒤로부터 몸 앞을 지나는 움직임 나타내는 정보를 포함하는 제스처 필터(191)로써 구현될 수 있으며, 이 움직임은 깊이 카메라에 의해 캡처될 것이다. 이후에, 그 제스처에 대한 파라미터들이 설정된다. 제스처가 던지기 동작이면, 파라미터는 손이 도달해야 하는 임계 속도(threshold velocity), 손이 이동해야 하는 (절대적인 또는 전체적으로 사용자의 신체 크기에 따라 상대적인) 거리 및 제스처가 발생했다는 인식 엔진의 신뢰도 평가일 수 있다. 제스처에 대한 이러한 파라미터들은 애플리케이션들 간에, 단일 애플리케이션의 컨텍스트들 간에, 또는 단일 애플리케이션의 단일 컨텍스트 내에서 시간에 따라 달라질 수 있다.

[0037] 제스처 인식 엔진(190)은 제스처 필터 컬렉션을 포함하며, 필터는 코드를 포함하거나 또는, 깊이, RGB 또는 골격 데이터를 처리하는 컴포넌트를 가리킬 수 있지만, 필터의 사용이 필터에 대한 분석을 제한하기 위한 의도는 아니다. 필터는 시스템에서 수신한 장면의 데이터를 분석하는 컴포넌트의 일례 또는 코드의 일부분을 나타내며, 데이터를 제스처를 나타내는 기본 정보와 비교한다. 분석 결과로서, 시스템은 입력 데이터가 제스처에 해당하는지에 대한 출력을 생성한다. 제스처를 나타내는 기본 정보는 사용자의 캡처 모션을 나타내는 데이터의 과거 기록에서 반복되는 특징에 상응하도록 조정될 수 있다. 예를 들어, 기본 정보는 전술한 제스처 필터의 일부일 수 있다. 그러나, 입력 데이터 및 제스처 데이터를 분석하는 임의의 적절한 방식이 고려될 수 있다.

[0038] 일 실시예에서, 제스처는 변화 모드(modification mode)로 들어가기 위한 트리거(trigger)로 인식될 수 있으며, 사용자는 시스템에서 자동 생성된 비주얼 형상을 변화시킬 수 있다. 예를 들어, 제스처 필터(191)는 변화 트리거 제스처를 인식하는 정보를 포함할 수 있다. 변화 트리거 제스처가 인식되면, 애플리케이션은 변화 모드로 들어간다. 변화 트리거 제스처는 애플리케이션마다, 시스템마다, 사용자마다 다를 수 있다. 예를 들어, 테니스 게임 애플리케이션과 볼링 게임 애플리케이션에서 동일한 제스처가 다른 변화 트리거 제스처일 수 있다. 사용자가 자신의 몸 앞에 검지 손가락이 위쪽을 향한 채 오른손으로 원을 그리는 모션을 포함하는 변화 트리거 제스처를 예를 들어 생각해보자. 변화 트리거 제스처의 파라미터 셋이 사용되어, 사용자의 손이 몸 앞에 있고, 사용자의 검지 손가락이 위쪽을 가리키고, 검지 손가락이 원을 그리는 모션으로 움직이고 있음을 식별할 수 있다.

[0039] 특정 제스처들은 변화 모드로 들어가기 위한 요청으로써 식별될 수 있으며, 애플리케이션이 현재 실행 중이라면, 변화 모드는 애플리케이션의 현 상태를 중단하고 변화모드로 들어가게 된다. 변화 모드는 애플리케이션을 일시 중지시키며, 사용자가 변화 모드에서 나갈 때, 중지된 시점에서 애플리케이션이 다시 시작될 수 있다. 또는, 변화 모드는 애플리케이션을 중지시키지 않고, 사용자가 변화시키는 동안에도 애플리케이션이 계속 실행될 수 있다.

[0040] 카메라(26, 28) 및 장치(20)에 의해 캡처된 골격 모델 및 그에 관련된 움직임 형식의 데이터가 제스처 라이브러리(192)의 제스처 필터들(191)과 비교되어, (골격 모델로 표현되는) 사용자가 하나 이상의 제스처들을 언제 행하였는지를 확인할 수 있다. 따라서, 필터(191) 등의 필터로의 입력은 사용자의 관절 위치에 관한 관절 데이터, 관절에서 만나는 뼈가 형성하는 각도, 장면의 RGB 색상 데이터 및 사용자 양태의 변화율 등을 포함할 수 있다. 언급한 대로, 제스처에 대한 파라미터들이 설정될 수 있다. 필터(191)의 출력에는 특정 제스처가 취해짐에 대한 신뢰도, 제스처 모션이 행해진 속도 및 제스처가 발생한 시각 등이 포함될 수 있다.

[0041] 컴퓨팅 환경(12)은 깊이 이미지를 처리하여, 방에 있는 사용자(18) 또는 사물과 같이 장면에서 어떤 타겟이 있는지를 결정할 수 있는 프로세서(195)를 포함할 수 있다. 이는, 예를 들어 유사한 거리 값을 갖는 깊이 이미지의 화소들을 함께 그룹화함으로써 실행될 수 있다. 또한, 이미지가 파싱되어(parsed) 관절 및 관절 사이의 조직 등의 특징이 식별되는 사용자의 골격 형상을 생성하게 된다. 깊이 카메라로 사람을 캡처하는 골격 매핑 기법이 있으며, 이것으로부터 사용자의 골격 상의 다양한 지점들, 손, 손목, 팔꿈치, 무릎, 코, 발목, 어깨 및 골반이 척추를 만나는 곳의 관절들을 결정한다. 다른 기법들은 이미지를 사람의 신체 모델 형상으로 전환하고, 이미지를 사람의 메쉬 모델 형상으로 전환하는 것이 있다.

[0042] 일 실시예에서, 이러한 처리는 캡처 장치(20) 자체에서 실행되며, 깊이 및 색상 값의 실제(raw) 이미지 데이터는(캡처 장치(20)가 3D 카메라(26)를 포함) 링크(36)를 통해 컴퓨팅 환경(12)으로 전송된다. 또 다른 실시예에서는, 이러한 처리가 이미지 카메라 컴포넌트(22)에 연결된 프로세서(32)에서 실행되고, 파싱된 이미지 데이터가 컴퓨팅 환경(12)으로 전송된다. 또 다른 실시예에서, 실제 이미지 데이터 및 파싱된 이미지 데이터 모두가 컴퓨팅 환경(12)으로 전송된다. 컴퓨팅 환경(12)은 파싱된 이미지를 수신하지만, 현 프로세스 또는 애플리케이션을 실행하기 위해 실제 데이터도 수신할 수 있다. 예를 들어, 장면의 이미지가 컴퓨터 네트워크를 통해 다른 사용자에게 전송되는 경우, 컴퓨팅 환경(12)은 다른 컴퓨팅 환경에서의 처리를 위해 실제 데이터도

전송할 수 있다.

- [0043] 프로세서는 특징 비교 모듈(196)을 포함할 수 있다. 특징 비교 모듈(196)은 타겟의 탐지된 특징들을 특징 라이브러리(197)의 옵션들과 비교할 수 있다. 특징 라이브러리(197)는 색상 옵션, 얼굴 특징 옵션, 체형 옵션, 크기 옵션 등의 비주얼 형상 특징 옵션들을 제공하며, 이 옵션들은 인간 및 비인간 타겟에 따라 다를 수 있다. 라이브러리는 비주얼 형상의 특징들을 저장하는 카탈로그, 데이터베이스, 메모리 등일 수 있다. 라이브러리는 조직화되거나 또는 조직화되지 않은 특징 옵션들의 컬렉션일 수 있다. 시스템 또는 사용자는 카탈로그에 특징들을 추가할 수 있다. 예를 들어, 애플리케이션이 사전 패키징된 특징 옵션 셋을 갖거나, 시스템이 이용가능한 디폴트 특징들을 갖고 있을 수 있다. 추가 특징 옵션들이 특징 라이브러리(197)에 추가되거나 업데이트될 수 있다. 예를 들어, 사용자는 가상 장터(marketplace)에서 추가 특징 옵션들을 구매하거나, 다른 사용자에게 특징 옵션들을 선물하거나, 또는 시스템이 사용자의 탐지된 특징들의 스냅샷을 찍어 특징 옵션들을 생성할 수도 있다.
- [0044] FCM(196)은 특징 옵션 카탈로그로부터 등, 타겟의 탐지된 특징들과 가장 닮은 특징을 선택할 수 있다. 시스템은 탐지된 특징들을 갖는 가상 객체를 자동 생성할 수 있다. 예를 들어, 물리적인 공간에서 빨간 2인용 의자(two-seater couch)를 탐지하는 경우를 생각해보자. 시스템은 의자의 탐지된 타겟 특징들과 닮은 특징들을, 단독으로 또는 다른 특징들과 조합하여 특징 라이브러리(197)로부터 식별한다. 일 실시예에서, 특징 라이브러리(197)로부터 선택하는 것은 단순하게는 물리적인 타겟의 적어도 하나의 특징을 갖는 가상 객체를 선택하는 것일 수도 있다. 예를 들어, 특징 라이브러리(197)는 가구에 관한 다양한 특징 옵션들을 갖고 있으며, 빨간 2인용 의자의 가상 이미지 또는 그림을 포함할 수 있다. 이러한 특징들은 사전에 패키징되어 애플리케이션 또는 시스템과 함께 제공될 수 있다. 다른 예로, 시스템은 물리적인 의자의 스냅샷을 찍어, 그 물리적인 의자의 형태를 취하는 만화 또는 가상 이미지를 만들 수 있다. 따라서, 그 선택된 특징이 시스템에 의해 이전에 찍힌 물리적인 의자의 스냅샷으로부터 온 것으로, 특징 라이브러리(197)에 추가될 수 있다.
- [0045] 시스템은 탐지된 타겟 특징들을 기초로 하여, 선택된 특징의 색상, 위치 또는 크기(scale)을 조정할 수 있다. 예를 들어, 시스템은 탐지된 타겟의 특징들과 닮은 특징을 특징 라이브러리(197)로부터 선택하거나 여러 특징들을 조합할 수 있다. 시스템은 탐지된 타겟과 완전히 닮게 하기 위해, 선택된 특징 또는 가상 이미지에 특징들을 추가할 수 있다. 탐지된 의자의 예시에서, 시스템은 특징 라이브러리(197)에서 특징을 찾고, 물리적인 의자의 특징과 닮은 적어도 하나의 특징을 갖는 의자에 대한 가상 프레임을 식별할 수 있다. 예를 들어, 처음에 시스템은 탐지된 물리적인 의자와 형상이 닮은 가상 의자를 선택할 수 있다. 가상 2인용 의자가 이용가능한 특징 옵션이라면, 시스템은 가상 2인용 의자를 선택할 수 있다. 색상도 시스템이 선택가능한 특징 옵션이 될 수 있다. 본 예시에서, 빨간 의자가 명확하게 특징 라이브러리(197)의 옵션이 아니라면, 시스템은 특징 라이브러리(197)로부터 색상을 선택하고 이를 선택된 가상 프레임에 적용할 수 있다. 시스템은 물리적인 의자의 탐지된 빨간 색상과 닮은 특징 라이브러리(197)에 존재하는 색상을 선택하거나, 또는 물리적인 의자의 색상의 스냅샷을 찍어, 이를 특징 옵션으로써 특징 라이브러리에 추가할 수 있다. 시스템은 선택된 빨간 색 특징을 가상 의자 이미지에 적용할 수 있다.
- [0046] 다른 실시예에서, 시스템은 탐지된 타겟을 닮은 가상 객체를 생성하기 위해, 특징 라이브러리의 특징들을 조합할 수 있다. 예를 들어, 시스템은 팔, 다리, 앉는 부분, 쿠션, 등, 뼈대(spine) 등 특징 라이브러리(197)의 의자 특징 옵션들을 선택하고, 선택된 특징들을 조립하여 2인용 의자를 생성할 수 있다.
- [0047] 또 다른 실시예에서, 사용자가 타겟이며, 시스템은 사용자의 특징들, 예컨대, 눈 색상, 크기 및 모양, 머리 색상, 타입 및 길이 등을 탐지한다. 시스템은 탐지된 특징들을 특징 옵션 카탈로그와 비교하고, 선택된 특징들을 비주얼 형상에 적용할 수 있다. 전술한 바와 같이, 시스템은 특징들을 결합하고, 이들 특징들을 변경할 수 있다(alter). 예를 들어, 색상, 위치 또는 스케일링(scaling)을 타겟에 적용하여 특징들을 변경할 수 있다. 색상 등 특징 라이브러리(197)로부터 추가적인 특징들을 선택하거나, 또는 타겟의 스냅샷의 이미지 데이터를 사용하여 특징들을 변경할 수 있다. 예를 들어, 애플리케이션은 특징 라이브러리(197)에 단색의 바지, 티셔츠 및 신발 종류 등의 일반적인 셋을 제공할 수 있다. 시스템은 일반적인 옷 특징을 선택하고, 시스템에 의해 탐지된 타겟의 옷의 색상을 반영하도록 그 옷에 색상을 적용함으로써 선택된 옷 특징을 변경할 수 있다.
- [0048] 또 다른 예를 들면, 시스템은 사용자의 특징들과 닮은 특징들의 서브셋을 특징 라이브러리(197)에서 식별하고, 사용자가 선택할 수 있는 이 서브셋을 제공할 수 있다. 따라서, 특정한 특징에 대해 사용자에게 제공되는 옵션의 수가 지적으로(intelligently) 필터링되어, 사용자가 비주얼 형상을 커스터마이징하기 쉬워지게 된다.

- [0049] 특징 라이브러리는 애플리케이션에 적용되거나, 또는 시스템 전반에서 이용될 수 있다. 예를 들어, 게임 애플리케이션은 게임에 적용가능한 다양한 정서를 나타내는 특징들을 정의한다. 특징 옵션들은 특정 또는 일반적인 특징들을 포함할 수 있다. 룩업 테이블 또는 데이터베이스에 관한 언급은 예시적인 것임을 알 것이며, 본원에 개시된 기법들에 관련된 특징 옵션의 준비는 적절한 임의의 방식으로 액세스, 저장, 포장, 제공, 생성될 수 있다.
- [0050] 컴퓨팅 환경(12)은 제스처 라이브러리(192)를 사용하여 골격 모델의 움직임을 해석하고, 이 움직임을 기초로 하여 애플리케이션을 제어할 수 있다. 컴퓨팅 환경(12)은 디스플레이 장치(193) 등의 디스플레이에 아바타 또는 포인터의 형태로 사용자의 형상을 모델링하고 디스플레이할 수 있다. 디스플레이 장치(193)는 컴퓨터 모니터, TV 스크린 또는 기타 적절한 디스플레이 장치를 포함할 수 있다. 예를 들어, 카메라로 제어되는 컴퓨터 시스템(camera-controlled computer system)은 사용자 이미지 데이터를 캡처하고 사용자의 제스처에 매핑되는 사용자 피드백을 TV 스크린에 디스플레이한다. 사용자 피드백은 도 1에 도시된 바와 같이 스크린에 아바타로서 디스플레이될 수 있다. 아바타의 움직임을 사용자의 움직임에 매핑시킴으로써 아바타의 모션이 직접 제어될 수 있다. 사용자 제스처는 애플리케이션의 특정 양태들을 제어하는 것으로 해석될 수 있다.
- [0051] 일 실시예에 따르면, 타겟은 가상 스크린의 생성, 사용자를 저장된 여러 프로파일들과 비교 및/또는 컴퓨팅 환경(12)과 같은 컴퓨팅 환경에서 타겟에 대한 프로파일 정보(198)를 저장하도록 스캔, 트래킹, 모델링 및/또는 평가되는, 서 있거나 앉아 있는 등 임의의 자세의 인간 타겟, 사물을 갖는 인간 타겟, 둘 이상의 인간 타겟, 하나 이상의 인간 타겟의 하나 이상의 부속물 등일 수 있다. 프로파일 정보(198)는 사용자 프로파일, 개인 프로파일, 애플리케이션 프로파일, 시스템 프로파일 또는 추후 액세스를 위해 데이터를 저장하는 기타 적절한 방법의 형태일 수 있다. 예를 들어, 프로파일 정보(198)는 애플리케이션을 통해 액세스 가능하거나, 시스템 전체에 걸쳐 이용될 수 있다. 프로파일 정보(198)는 특정 사용자 프로파일 정보를 로딩하는 룩업 테이블들을 포함할 수 있다. 가상 스크린은 도 1과 관련하여 진술한 컴퓨팅 환경(12)에서 실행되는 애플리케이션과 인터랙션할 수 있다.
- [0052] 시스템은 사용자의 프로파일에 저장된 정보를 기초로 하여 비주얼 형상을 자동 생성함으로써, 사용자와 같은 타겟의 비주얼 형상을 렌더링할 수 있다. 실시예들에 따르면, 룩업 테이블은 사용자 전용 프로파일 정보를 포함할 수 있다. 일 실시예에서, 컴퓨팅 환경(12)과 같은 컴퓨팅 환경에서는 하나 이상의 사용자에 대한 프로파일 데이터(198)가 룩업 테이블에 저장될 수 있다. 저장된 프로파일 데이터(198)는 타겟의 스캔되거나 추정된 신체 크기, 골격 모델, 목소리 샘플 또는 비밀 번호, 타겟의 성별, 타겟의 나이, 이전 제스처들, 타겟 제한 사항(target limitations) 및 예컨대, 앉는 경향, 왼손 또는 오른손잡이 또는 캡처 장치에 인접해서 서는 경향 등 시스템의 타겟에 의한 표준 사용법을 포함한다. 이러한 정보는 캡처 장면의 타겟과 하나 이상의 사용자 프로파일(198) 간에 일치 여부를 결정하는 데에 사용되어, 일 실시예에 따르면, 시스템이 가상 스크린을 사용자에게 맞게 조정하거나 또는, 컴퓨팅 또는 게임 경험의 다른 요소들을 프로파일(198)에 따라 조정할 수 있게 된다.
- [0053] 타겟의 비주얼 형상에 대해 이전에 선택된 특징들은 프로파일에 저장될 수 있다. 예를 들어, 사용자 전용 프로파일은, 선택되고 적용되어 사용자의 비주얼 형상을 자동 생성한 특징들을 저장할 수 있다. 장소 전용 프로파일(location-specific profile)이, 선택되고 적용되어 물리적인 공간과 유사한 가상 장면을 자동 생성하고 디스플레이하는 특징들을 저장할 수 있다. 예를 들어, 방안의 가구 등 물리적인 공간의 객체에 대응하는 가상 객체는 특징 라이브러리(197)의 옵션들을 선택함으로써 생성될 수 있다. 색상이 탐지되고, 이용가능한 색상이 특징 라이브러리(197)로부터 선택될 수 있다. 시스템에서 인식하고 초기화하면, 장소 전용 프로파일이 로딩되어, 그 장소에 대응하는 가구 및 색상들이 디스플레이된다.
- [0054] 하나 이상의 개인 프로파일(198)이 컴퓨터 환경(12)에 저장되고 다수의 사용자 세션에서 사용될 수 있으며, 또는 단일 세션만을 위해 하나 이상의 개인 프로파일이 생성될 수도 있다. 사용자는 목소리 또는 신체 스캔, 나이, 개인 선호도, 왼손 또는 오른손잡이, 아바타, 이름 등의 정보를 시스템에 제공하는 프로파일을 설정하는 옵션을 가질 수 있다. 캡처 공간에 들어간 다음에도 시스템에 어떤 정보도 제공하지 않는 "게스트"에 대해서 개인 프로파일이 제공될 수도 있다. 하나 이상의 게스트에 대해 임시 개인 프로파일이 설정될 수 있다. 게스트 세션 끝에, 게스트 개인 프로파일이 저장되거나 삭제될 것이다.
- [0055] 제스처 라이브러리(192), 제스처 인식 엔진(190), 특징 라이브러리(197), 특징 비교기(features comparer, 196) 및 프로파일(198)은 하드웨어, 소프트웨어 또는 이들의 조합으로 구현될 수 있다. 예를 들어, 제스처 라이브러리(192) 및 제스처 인식 엔진(190)은 컴퓨팅 환경(12)의 프로세서(195)와 같은 프로세서(또는 도 3의

처리 장치(101)나 도 4의 처리 장치(259))에서 실행되는 소프트웨어로 구현될 수 있다.

[0056] 이하에서 설명될 도 3 및 4에 도시된 블록 다이어그램은 예시적인 것으로 특정한 구현을 의미하는 것은 아님이 강조된다. 따라서, 도 2의 프로세서(195 또는 32), 도 3의 처리 장치(101) 및 도 4의 처리 장치(259)는 단일 프로세서 또는 다수의 프로세서들도 구현될 수 있다. 다수의 프로세서들은 분산되거나 중앙에 위치할 수 있다. 예를 들어, 제스처 라이브러리(192)는 캡처 장치의 프로세서(32)에서 실행되는 소프트웨어로 구현되거나, 또는 컴퓨팅 환경(12)의 프로세서(195)에서 실행되는 소프트웨어로 구현될 수 있다. 본원에 개시된 기법들을 실행하기에 적합한 프로세서들의 임의의 조합이 고려될 수 있다. 다수의 프로세서들은 무선으로, 또는 유선(hard wire)으로 이들의 조합으로 통신할 수 있다.

[0057] 또한 본원에서는, 컴퓨팅 환경(12)은 단일 컴퓨팅 장치 또는 컴퓨팅 시스템을 의미한다. 컴퓨팅 환경은 비-컴퓨팅 컴포넌트들을 포함할 수 있다. 컴퓨팅 환경은 도 2에 도시된 디스플레이 장치(193)와 같은 디스플레이 장치를 포함할 수 있다. 예컨대, 디스플레이 장치는 컴퓨팅 환경과는 분리되지만 컴퓨팅 환경에 연결된 엔티티이며, 디스플레이 장치는 디스플레이 및 처리를 하는 컴퓨팅 장치이다. 따라서, 컴퓨팅 시스템, 컴퓨팅 장치, 컴퓨팅 환경, 컴퓨터, 프로세서 또는 기타 컴퓨팅 컴포넌트는 서로 교체 사용이 가능하다.

[0058] 애플리케이션 또는 애플리케이션의 컨텍스트에 따라 제스처 라이브러리 및 필터 파라미터들이 제스처 툴에 의해 튜닝될 수 있다. 컨텍스트는 문화적인 컨텍스트일 수 있으며, 환경적인 컨텍스트일 수도 있다. 문화적인 컨텍스트란 시스템을 사용하는 사용자의 문화를 의미한다. 다른 문화들은 유사한 제스처를 사용하여 현저하게 다른 의미를 전할 수 있다. 예를 들면, 다른 사용자에게 "보도록" 또는 "눈을 사용하도록" 말하고 싶은 미국인 사용자는 자신의 관자놀이 근처에 검지 손가락을 둔다. 반면, 이탈리아 사용자는 이러한 제스처를 마피아를 언급하는 것으로 해석한다.

[0059] 유사하게, 단일 애플리케이션의 상이한 환경에서 상이한 컨텍스트들이 존재할 수 있다. 자동차 운전을 포함하는 일인칭 슈터 게임(first-user shooter game)을 예를 든다. 사용자가 걷는 경우, 땅을 향해 손가락으로 주먹을 쥐고 그 주먹을 앞으로 몸에서 멀어지게 뻗는 것은 펀치 제스처를 나타낸다. 사용자가 운전하는 컨텍스트에서는, 동일한 모션은 "기어 전환" 제스처를 나타낸다. 비주얼 형상에 대한 변화에 관해서는, 환경에 따라 다른 제스처들이 다른 변화를 트리거할 수 있다. 애플리케이션 전용 변화 모드 및 시스템 전반 변화 모드로 들어가는 데에 다른 변화 트리거 제스처가 사용될 수 있다. 각 변화 모드는, 그 변화 모드에 대응하는 독립적인 제스처 세트를 수반하며, 그 변화 트리거 제스처의 결과에 따라 해당 모드로 들어갈 수 있다. 예를 들어, 볼링 게임에서, 팔을 흔드는 모션은 가상 볼링장에서 볼링 공을 내려놓기 위해 흔드는 것으로 식별되는 제스처이다. 그러나, 다른 애플리케이션에서는, 팔을 흔드는 모션은 스크린에 디스플레이되는 사용자의 아바타의 팔을 늘리는 요청으로 식별되는 제스처일 수 있다. 또한, 사용자가 자신의 게임을 저장하거나, 자신의 캐릭터의 장비들을 선택하는 등, 직접적인 게임 플레이를 포함하지 않는 유사한 액션을 취할 수 있는 하나 이상의 메뉴 환경이 마련될 수도 있다. 이러한 환경에서는, 이 동일한 제스처가 임의의 것을 선택하거나 또는 다른 스크린으로 진행되도록 하는 등 제 3의 의미를 가질 수 있다.

[0060] 제스처들은 각 장르의 애플리케이션에서 사용될 수 있는 연관 제스처들의 장르 패키지로 그룹화될 수 있다. 연관 제스처들은 - 통상적으로 함께 사용되기 때문에, 또는 제스처의 파라미터의 변경이 다른 제스처의 파라미터를 변경시킨다는 점에서 연관된 - 장르 패키지들로 함께 그룹화될 수 있다. 이러한 패키지들은 적어도 하나의 패키지를 사용하는 애플리케이션으로 제공된다. 애플리케이션은 애플리케이션 특유의 양태들에 가장 적합하도록 제스처 또는 제스처 필터(191)의 파라미터를 튜닝 또는 변화시킬 수 있다. 파라미터가 튜닝될 때, 그 제스처 또는 제 2 제스처의 (상호의존적인 면에서의) 제 2 연관 파라미터도 파라미터들이 연관되도록 튜닝된다. 비디오 게임의 장르 패키지들은 일인칭 슈터, 액션, 드라이빙 및 스포츠와 같은 장르를 포함할 수 있다.

[0061] 도 3은 타겟 인식, 분석 및 트래킹 시스템에서 여러 제스처들을 해석하는 데에 사용되는 컴퓨팅 환경의 일 실시예를 도시한다. 도 1 및 2와 관련하여 기술한 컴퓨팅 환경(12) 등의 컴퓨팅 환경은 게임 콘솔과 같은 멀티미디어 콘솔(100)일 수 있다. 도 3에 도시된 바와 같이, 멀티미디어 콘솔(100)은, 레벨 1 캐시(102), 레벨 2 캐시(104) 및 플래시 ROM(Read Only Memory, 106)을 포함하는 CPU(central processing unit, 101)를 갖는다. 레벨 1 캐시(102) 및 레벨 2 캐시(104)는 데이터를 임시로 저장하여 메모리 액세스 사이클의 수를 감소시킴으로써, 처리 속도와 처리율을 향상시킨다. CPU(101)에는 하나 이상의 코어, 따라서 추가적인 레벨 1 및 레벨 2 캐시(102 및 104)가 제공될 수 있다. 플래시 ROM(106)은 멀티미디어 콘솔(100)의 전원이 켜질 때 부팅 프로세스(boot process)의 초기 단계 동안 로딩되는 실행가능 코드를 저장할 수 있다.

- [0062] 그래픽 처리 장치(GPU, 108)와 비디오 인코더/비디오 코덱(코더/디코더, 114)은 고속 및 고해상도 그래픽 처리를 위해 비디오 처리 파이프라인을 형성한다. 데이터는 버스를 통해 그래픽 처리 장치(108)에서 비디오 인코더/비디오 코덱(114)으로 전달된다. 비디오 처리 파이프라인은 텔레비전이나 다른 디스플레이로의 전송을 위해 A/V(오디오/비디오) 포트(140)로 데이터를 출력한다. 메모리 제어기(110)는 GPU(108)에 접속되어, RAM(Random Access Memory)과 같은, 하지만 이에 제한되지는 않는, 다양한 유형의 메모리(112)로의 프로세서 액세스를 용이하게 한다.
- [0063] 멀티미디어 콘솔(100)은, 모듈(118)에서 바람직하게 구현되는 I/O 제어기(120), 시스템 관리 제어기(122), 오디오 처리 장치(123), 네트워크 인터페이스 제어기(124), 제 1 USB 호스트 제어기(126), 제 2 USB 제어기(128), 및 전면 패널 I/O 서브어셈블리(130)를 포함한다. USB 제어기들(126 및 128)은 주변 제어기들(142(1) ~ 142(2)), 무선 어댑터(148) 및 외부 메모리 장치(146, 예를 들어, 플래시 메모리, 외부 CD/DVD ROM 드라이브, 이동식 매체 등)를 위한 호스트들로서 기능한다. 네트워크 인터페이스(124) 및/또는 무선 어댑터(148)는 네트워크(예를 들어, 인터넷, 홈 네트워크 등)로의 액세스를 제공하고, 이더넷 카드, 모뎀, 블루투스 모듈, 케이블 모뎀 등을 포함하는 아주 다양한 유무선 어댑터 컴포넌트들 중 임의의 것일 수 있다.
- [0064] 시스템 메모리(143)는 부팅 프로세스 동안에 로딩되는 애플리케이션을 저장하기 위해 제공된다. 미디어 드라이브(144)가 제공되고, 이는 DVD/CD 드라이브, 하드 드라이브, 또는 다른 이동식 미디어 드라이브 등을 포함할 수 있다. 미디어 드라이브(144)는 멀티미디어 콘솔(100)의 내부 또는 외부에 있을 수 있다. 애플리케이션 데이터는 멀티미디어 콘솔(100)에 의한 실행, 플레이백 등을 위해 미디어 드라이브(144)를 통해 액세스될 수 있다. 미디어 드라이브(144)는, 직렬 ATA 버스 또는 기타 고속 접속(예를 들어, IEEE 1394)과 같은 버스를 통해 I/O 제어기(120)에 접속된다.
- [0065] 시스템 관리 제어기(122)는 멀티미디어 콘솔(100)의 가용성의 보장과 관련된 다양한 서비스 기능들을 제공한다. 오디오 처리 장치(123)와 오디오 코덱(132)은 고신뢰성과 스테레오 처리를 갖는 대응 오디오 처리 파이프라인을 형성한다. 오디오 데이터는 통신 링크를 통해 오디오 처리 장치(123)와 오디오 코덱(132) 사이에서 전달된다. 오디오 처리 파이프라인은 오디오 기능이 있는 외부 오디오 플레이어 또는 장치에 의한 재생을 위해 A/V 포트(140)로 데이터를 출력시킨다.
- [0066] 전면 패널 I/O 서브어셈블리(130)는 전원 버튼(150), 배출 버튼(eject button, 152), 및 멀티미디어 콘솔(100)의 외부 표면에 노출된 임의의 LED들(light emitting diodes) 또는 기타 인디케이터들의 기능을 지원한다. 시스템 전원 모듈(136)은 멀티미디어 콘솔(100)의 컴포넌트들에 전력을 공급한다. 팬(fan, 138)은 멀티미디어 콘솔(100) 내의 회로를 냉각시킨다.
- [0067] 멀티미디어 콘솔(100) 내의 CPU(101), GPU(108), 메모리 제어기(110) 및 기타 다양한 컴포넌트들은 직렬 및 병렬 버스들, 메모리 버스, 주변장치 버스, 또는 다양한 버스 아키텍처들 중의 임의의 것을 사용하는 프로세서나 로컬 버스를 포함하는 하나 이상의 버스들을 통해 상호접속된다. 예를 들어, 이러한 아키텍처에는 PCI(Peripheral Component Interconnects) 버스, PCI 익스프레스(PCI-Express) 버스 등이 포함될 수 있다.
- [0068] 멀티미디어 콘솔(100)의 전원이 켜질 때, 시스템 메모리(143)로부터 애플리케이션 데이터가 메모리(112) 및/또는 캐시(102, 104)로 로딩되어, CPU(101)에서 실행될 수 있다. 애플리케이션은 멀티미디어 콘솔(100)에서 이용가능한 다른 미디어 유형들로 네비게이트할 때 일관된 사용자 경험을 제공하는 그래픽 사용자 인터페이스를 제공할 수 있다. 동작 시에는, 미디어 드라이브(144) 내에 포함된 애플리케이션들 및/또는 기타 미디어들이 미디어 드라이브(144)로부터 시작되거나 재생되어, 멀티미디어 콘솔(100)에 추가 기능들을 제공할 수 있다.
- [0069] 멀티미디어 콘솔(100)은 텔레비전이나 기타 디스플레이에 시스템을 단순히 접속시킴으로써 독립형 시스템으로서 동작할 수 있다. 이 독립형 모드에서, 멀티미디어 콘솔(100)은 한 명 이상의 사용자들이 시스템과 인터랙트하고, 영화를 보고, 음악을 듣게 하도록 할 수 있다. 그러나, 네트워크 인터페이스(124) 또는 무선 어댑터(148)를 통해 이용가능하게 된 통합 광대역 접속으로 인해, 멀티미디어 콘솔(100)은 보다 큰 네트워크 커뮤니티의 참가자로서 동작할 수도 있다.
- [0070] 멀티미디어 콘솔(100)의 전원이 켜지면, 설정량의 하드웨어 리소스들이 멀티미디어 콘솔 운영 체제에 의한 시스템 사용을 위해 예약된다. 이러한 리소스들은 메모리(예컨대, 16 MB), CPU 및 GPU 사이클(예컨대, 5%), 네트워크 대역폭(예컨대, 8 kbs) 등의 예약을 포함할 수 있다. 이러한 리소스들은 시스템 부팅 시간에 예약되기 때문에, 애플리케이션의 관점에서는 예약된 리소스들이 존재하지 않는다.

- [0071] 특히, 개시 커널(launch kernel), 동시(concurrent) 시스템 애플리케이션 및 드라이버를 포함할 수 있을 정도로 메모리 예약이 충분히 큰 것이 바람직하다. 예약된 CPU 사용량이 시스템 애플리케이션에서 사용되지 않는 경우, 휴지 스레드(idle thread)가 미사용 사이클들을 쓸 수 있도록 CPU 예약이 일정한 것이 바람직하다.
- [0072] GPU 예약과 관련하여, 팝업이 오버레이되도록 코드를 스케줄링하는 GPU 인터럽트를 사용하여, 시스템 애플리케이션에서 생성되는 간단한 메시지(예컨대, 팝업)가 디스플레이된다. 오버레이에 필요한 메모리량은 오버레이 영역 크기에 따르며, 오버레이는 스크린 해상도에 맞춰 스케일링되는 것이 바람직하다. 동시 시스템 애플리케이션이 풀 유저 인터페이스(full user interface)를 사용하는 경우에는, 애플리케이션 해상도와는 별개의 해상도를 사용하는 것이 바람직하다. 주파수를 변경하고 TV를 재동기화시킬 필요가 없도록 이 해상도를 설정하기 위해, 스케일러가 사용될 수 있다.
- [0073] 멀티미디어 콘솔(100)이 부팅되고 시스템 리소스가 예약된 후에, 동시 시스템 애플리케이션이 실행되어 시스템 기능들을 제공한다. 시스템 기능들은, 상기에서 설명한 예약된 시스템 리소스들 내에서 실행되는 일련의 시스템 애플리케이션에서 캡슐화되어 있다. 운영 체제 커널은 시스템 애플리케이션 스레드인지 게임 애플리케이션 스레드인지를 식별한다. 일관적인 시스템 리소스 뷰를 애플리케이션에 제공하기 위해, 시스템 애플리케이션은 사전 설정된 시간 및 간격으로 CPU(101)에서 실행되도록 스케줄링되는 것이 바람직하다. 스케줄링은 콘솔에서 실행되는 게임 애플리케이션에 대한 캐시 중단을 최소화하기 위한 것이다.
- [0074] 동시 시스템 애플리케이션이 오디오를 필요로 할 때, 오디오 처리는 시간에 대한 민감도로 인해 게임 애플리케이션과 비동기적으로 스케줄링된다. 시스템 애플리케이션이 활성화될 때, 멀티미디어 콘솔 애플리케이션 관리자(이하에서 설명됨)는 게임 애플리케이션 오디오 레벨(예컨대, 음소거(mute), 감쇠(attenuate))을 제어한다.
- [0075] 게임 애플리케이션 및 시스템 애플리케이션은 입력 장치들(예컨대, 제어기(142(1) 및 142(2)))을 공유한다. 입력 장치들은 예약된 리소스들이 아니지만, 각 시스템 애플리케이션 및 게임 애플리케이션이 입력 장치의 포커스를 갖도록 애플리케이션들 사이에서 스위칭될 것이다. 애플리케이션 관리자는 게임 애플리케이션에 대한 정보 없이 입력 스트림의 스위칭을 제어하는 것이 바람직하며, 드라이버는 포커스 스위치에 관한 상태 정보를 보유한다. 카메라(26, 28) 및 캡처 장치(20)는 콘솔(100)을 위한 추가적인 입력 장치들이다.
- [0076] 도 4는 타겟 인식, 분석 및 트래킹 시스템에서의 여러 체스처들의 해석, 및/또는 타겟 인식, 분석 및 트래킹 시스템에 의해 디스플레이되는 아바타, 온스크린 캐릭터, 온스크린 객체 등의 가상 캐릭터의 애니메이션화를 위해 사용되는 도 1 및 도 2에 도시된 컴퓨팅 환경(12)의 또 다른 실시예인 컴퓨팅 환경(220)을 도시한다. 컴퓨팅 시스템 환경(220)은 적합한 컴퓨팅 환경의 일례일 뿐, 본원에 개시된 대상의 사용이나 기능의 범위를 제한하는 것은 아니다. 컴퓨팅 환경(220)은, 예시적인 운영 환경(220)에 도시된 임의의 컴포넌트 또는 컴포넌트 조합에 관한 의존성이나 요구 사항을 가지는 것으로 해석되어서는 안 된다. 몇몇 실시예에서, 도시된 다양한 컴퓨팅 구성요소들은 본 개시의 특정 양태들에 대한 예를 들어 설명하는 회로를 포함할 수 있다. 예를 들어, 본 개시에서 사용되는 회로는 펌웨어나 스위치로 기능(들)을 실행하도록 구성되는 특화된 하드웨어 컴포넌트들을 포함할 수 있다. 또 다른 실시예에서, 회로는 기능(들)을 실행하는 논리 동작을 구현하는 소프트웨어 인스트럭션으로 구성되는 범용 처리 장치(general purpose processing unit), 메모리 등을 포함할 수 있다. 하드웨어와 소프트웨어의 조합을 포함하는 회로의 실시예들에 있어서, 구현자는 논리를 구현하는 소스 코드를 작성하고, 이 소스 코드는 범용 처리 장치에서 처리될 수 있는 기계 판독가능 코드로 컴파일될 수 있다. 실시예들에서, 당업자라면 하드웨어, 소프트웨어 또는 하드웨어/소프트웨어의 조합 간에 차이가 거의 없는 방향으로 선행 기술이 발전해왔다는 점을 알 수 있으므로, 특정 기능을 구현하기 위한 하드웨어 대 소프트웨어의 선택은 구현자에게 달려 있는 설계 선택에 불과하다. 특히, 당업자라면 소프트웨어 프로세스가 동등한 하드웨어 구조로 변환가능하고, 하드웨어 구조 또한 동등한 소프트웨어 프로세스로 변환가능하다는 점을 알 수 있다. 따라서, 하드웨어 구현 대 소프트웨어 구현의 선택은 설계 선택으로, 구현자에게 달려 있다.
- [0077] 도 4에서, 컴퓨팅 환경(220)은, 일반적으로 다양한 컴퓨터 판독가능 매체를 포함하는 컴퓨터(241)를 포함한다. 컴퓨터 판독가능 매체는 컴퓨터(241)가 액세스할 수 있으며, 휘발성 및 비휘발성 매체, 이동식 및 비이동식 매체 모두를 포함하는 임의의 가용 매체일 수 있다. 시스템 메모리(222)는 ROM(read only memory, 223) 및 RAM(random access memory, 260)과 같은 휘발성 및/또는 비휘발성 메모리의 형태인 컴퓨터 저장 매체를 포함한다. 스타트업 동안과 같이 컴퓨터(241) 내의 구성요소들 사이의 정보 전송을 돕는 기본 루틴을 포함하는 기본 입출력 시스템(BIOS, 224)은 일반적으로 ROM(223)에 저장된다. RAM(260)은 일반적으로 처리 장치(259)에 의해 즉시 액세스 가능 및/또는 바로 동작되는 데이터 및/또는 프로그램 모듈들을 포함한다. 예를 들어, 도 4는 운영 체제(225), 애플리케이션 프로그램(226), 다른 프로그램 모듈(227), 및 프로그램 데이터

(228)를 도시하고 있으며, 이들로서 제한되는 것은 아니다.

[0078] 또한, 컴퓨터(241)는 다른 이동식/비이동식, 휘발성/비휘발성 컴퓨터 저장 매체를 포함할 수 있다. 예를 들어, 도 4는 비이동식, 비휘발성 자기 매체에 관독 또는 기록하는 하드 디스크 드라이브(238), 이동식, 비휘발성 자기 디스크(254)에 관독 또는 기록하는 자기 디스크 드라이브(239), 및 CD ROM 또는 다른 광학 매체와 같은 이동식, 비휘발성 광 디스크(253)에 관독 또는 기록하는 광 디스크 드라이브(240)를 도시한다. 예시적인 운영 환경에서 사용가능한 다른 이동식/비이동식, 휘발성/비휘발성 컴퓨터 저장 매체는, 자기 테이프 카세트, 플래시 메모리 카드, DVD, 디지털 비디오 테이프, 고체 상태 RAM, 고체 상태 ROM 등을 포함하지만, 이들로서 제한되는 것은 아니다. 하드 디스크 드라이브(238)는 일반적으로 인터페이스(234)와 같은 비이동식 메모리 인터페이스를 통해 시스템 버스(221)에 연결되며, 자기 디스크 드라이브(239) 및 광 디스크 드라이브(240)는 일반적으로 인터페이스(235)와 같은 이동식 메모리 인터페이스에 의해 시스템 버스(221)에 연결된다.

[0079] 앞서 논의되었으며 도 4에 도시된 드라이브 및 그 관련 컴퓨터 저장 매체는 컴퓨터(241)를 위한 컴퓨터 관독 가능 인스트럭션, 데이터 구조, 프로그램 모듈 및 기타 데이터를 저장한다. 도 4에서, 예를 들어, 하드 디스크 드라이브(238)는 운영 체제(258), 애플리케이션 프로그램(257), 다른 프로그램 모듈(256) 및 프로그램 데이터(255)를 저장하는 것으로 도시되어 있다. 이러한 컴포넌트들은 운영 체제(225), 애플리케이션 프로그램(226), 다른 프로그램 모듈(227) 및 프로그램 데이터(228)와 동일할 수도 또는 다를 수도 있음을 알 것이다. 운영 체제(258), 애플리케이션 프로그램(257), 다른 프로그램 모듈(256), 및 프로그램 데이터(255)는, 최소한 이들이 상이한 카피들을 나타내기 위해 본원에서 상이한 번호가 부여된다. 사용자는 키보드(251), 및 일반적으로 마우스, 트랙볼, 또는 터치 패드로 불리는 포인팅 장치(252)를 비롯한 입력 장치들을 통해 명령어 및 정보를 컴퓨터(241)에 입력할 수 있다. 다른 입력 장치들(도시되지 않음)은 마이크, 조이스틱, 게임 패드, 위성 접시, 스캐너, 리모콘 등을 포함할 수 있다. 이들 및 다른 입력 장치들은 보통 시스템 버스에 연결된 사용자 입력 인터페이스(236)를 통해 처리 장치(259)에 접속되지만, 병렬 포트, 게임 포트, 또는 USB(universal serial bus)를 비롯한 다른 인터페이스 및 버스 구조에 의해 접속될 수도 있다. 카메라(26, 28) 및 캡처 장치(20)는 콘솔(100)의 추가 입력 장치에 해당한다. 모니터(242) 또는 다른 형태의 디스플레이 장치도 비디오 인터페이스(232)와 같은 인터페이스를 통해 시스템 버스(221)에 접속된다. 모니터뿐만 아니라, 컴퓨터들은, 출력 주변 기기 인터페이스(233)를 통해 접속될 수 있는 스피커(244) 및 프린터(243)를 비롯한 다른 주변 기기 출력 장치들을 포함할 수 있다.

[0080] 컴퓨터(241)는 원격 컴퓨터(246)와 같은 하나 이상의 원격 컴퓨터들에 대한 논리적인 연결들을 사용하여 네트워크 환경에서 동작할 수 있다. 원격 컴퓨터(246)는 개인용 컴퓨터, 서버, 라우터, 네트워크 PC, 피어 장치(peer device) 또는 다른 공통 네트워크 노드일 수 있으며, 단지 메모리 저장 장치(247)만이 도 4에 도시되어 있지만, 일반적으로 컴퓨터(241)와 관련하여 전술한 다수의 또는 모든 구성요소들을 포함한다. 도 4에 도시된 논리적인 연결은 LAN(local area network, 245) 및 WAN(wide area network, 249)을 포함하지만, 또한 다른 네트워크들도 포함할 수 있다. 이러한 네트워크 환경들은 사무실, 기업(enterprise-wide) 컴퓨터 네트워크, 인터넷 및 인터넷에 혼다.

[0081] LAN 네트워크 환경에서 사용될 때, 컴퓨터(241)는 네트워크 인터페이스 또는 어댑터(237)를 통해서 LAN(245)에 연결된다. WAN 네트워크 환경에서 사용될 때, 컴퓨터(241)는 일반적으로 모뎀(250) 또는 인터넷과 같이 WAN(249)을 통해 통신을 구축하기 위한 다른 수단을 포함한다. 내부 또는 외부에 존재할 수 있는 모뎀(250)은 사용자 입력 인터페이스(236), 또는 다른 적절한 메카니즘을 통해 시스템 버스(221)에 연결될 수 있다. 네트워크 환경에서, 컴퓨터(241) 또는 그 일부분에 관련하여 도시된 프로그램 모듈들이 원격 메모리 저장 장치 내에 저장될 수 있다. 예를 들면, 도 4는 메모리 장치(247) 상에 존재하는 원격 애플리케이션 프로그램들(248)을 도시하고 있다. 도시된 네트워크 연결들은 예시적인 것이며, 컴퓨터들 간에 통신 링크를 구축하는 다른 수단이 사용될 수 있음을 이해할 것이다.

[0082] 컴퓨터 관독가능 저장 매체는 비주얼 형상을 변화시키는 컴퓨터 관독가능 인스트럭션들을 포함할 수 있다. 이 인스트럭션들은, 비주얼 형상을 렌더링하고, 물리적 공간에서 사용자의 변화 제스처를 나타내는 데이터를 포함하는 장면의 데이터를 수신하며, 사용자의 변화 제스처를 기초로 하여 비주얼 형상을 변화시키는 인스트럭션들을 포함할 수 있으며, 변화 제스처는 비주얼 형상의 특성을 변화시키기 위한 제어에 매핑되는 제스처이다.

[0083] 도 5는 캡처 장치(20)에서 캡처되는 이미지 데이터로부터 생성되는 사용자의 골격 매핑의 일례를 도시한다. 본 실시예에서, 각 손(502), 각 팔뚝(504), 각 팔꿈치(506), 이두박근(508), 각 어깨(510), 각 엉덩이(512), 각 허벅지(514), 각 무릎(516), 각 종아리(518), 각 발(520), 머리(522), 몸통(524), 척추 상부(526)와 하부

(528) 및 허리(530) 등 다양한 관절들과 뼈들이 식별된다. 더 많은 포인트들이 트래킹되면, 손가락 또는 발가락의 뼈 및 관절들, 또는 코 및 눈과 같은 얼굴 각각의 특징 등의 추가 특징들이 식별될 것이다.

[0084] 사용자는 자신의 신체를 움직여서 제스처를 생성할 수 있다. 제스처는 이미지 데이터로 캡처되어 의미를 위해 파싱되는 사용자의 모션 또는 포즈를 포함한다. 제스처는 동적일 수 있으며, 공 던지기를 흉내내는 것과 같은 모션을 포함한다. 제스처는 자신의 몸통(524) 앞에 팔뚝(504)을 교차하고 있는 등 정적인 포즈일 수 있다. 또한, 제스처는 모형 칼을 흔드는 것처럼 소품을 포함할 수도 있다. 제스처는 양손(502)으로 손뼉을 치는 것과 같이 하나 이상의 신체 부위를 포함하거나, 입술을 오므리는 것과 같은 세밀한 모션을 포함할 수 있다.

[0085] 사용자의 제스처는 일반적인 컴퓨팅 컨텍스트에서 입력으로 사용될 수 있다. 예를 들어, 양손(502) 또는 다른 신체 부위들의 다양한 모션들은 계층적 리스트에서 위아래로 탐색, 파일 열기, 파일 닫기 및 파일 저장을 비롯한 통상의 시스템 전반에 걸친 태스크들에 대응한다. 예를 들어, 사용자가 손가락은 위를 가리키고 손바닥은 캡처 장치(20)를 향한 채 손을 들고 있다. 다음에, 사용자가 손바닥쪽으로 손가락을 말아 주먹을 쥐면, 이는 창 기반(window-based) 사용자 인터페이스 환경에서 포커스된 창이 닫혀져야 함을 나타내는 제스처일 수 있다. 또한, 제스처는 비디오 게임 전용 컨텍스트에서 게임에 따라 다르게 사용될 수 있다. 예를 들어, 드라이빙 게임에서는, 양손(502)과 양발(520)의 다양한 모션들이 자동차 방향 조종, 기어 전환, 가속(accelerating) 및 감속(braking)에 대응할 수 있다. 따라서, 제스처는 비디오 게임, 문서 편집기, 워드 프로세싱, 데이터 관리 등의 매우 다양한 애플리케이션에서 디스플레이되는 사용자 형상에 매핑되는 매우 다양한 모션들을 나타낼 수 있다.

[0086] 사용자는 제자리에서 걷거나 뚝뚝 걸거나 뛰는 것에 해당하는 제스처를 생성할 수 있다. 예를 들어, 사용자는 양쪽 다리(512 - 520)를 번갈아 들어올렸다 내림으로써, 이동하지 않고 걸음을 흉내낼 수 있다. 시스템은 각 엉덩이(512) 및 각 허벅지(514)를 분석하여 이러한 제스처를 파싱할 수 있다. 한쪽의 엉덩이-허벅지 각도(수직선에 대해 상대적으로 측정되며, 서있는 다리는 0°의 엉덩이-허벅지 각도를, 수평으로 앞으로 뻗은 다리는 90°의 엉덩이-허벅지 각도를 가짐)가 다른쪽 허벅지에 대하여 특정 임계치를 넘을 때 걸음으로 인식된다. 다리를 번갈아서 임의의 수의 연속적인 걸음 후에 걷기 또는 뛰기로 인식될 수 있다. 가장 최근의 두 걸음 사이의 시간이 주기라고 간주된다. 임계각이 충족되지 않는 임의의 수의 주기가 지난 후에, 시스템은 걷거나 뛰는 제스처가 중지되었음을 결정하게 된다.

[0087] "걷기 또는 뛰기" 제스처가 제공되면, 애플리케이션은 이러한 제스처에 관련된 파라미터들에 대한 값을 설정한다. 이들 파라미터는 상기의 임계각, 걷기 또는 뛰기 제스처의 시작을 위해 필요한 걸음의 수, 및 제스처가 걷기인지 뛰기인지를 결정하는 임계 주기를 포함할 수 있다. 사용자가 다리를 재빨리 움직이기 때문에, 빠른 주기는 뛰기에 해당하고, 느린 주기는 걷기에 해당할 것이다.

[0088] 제스처는 처음에는 애플리케이션이 자체 파라미터들로 오버라이드할 수 있는 일련의 디폴트 파라미터들에 연관될 수 있다. 이러한 시나리오에서, 애플리케이션은 파라미터를 제공하도록 요구되지 않으며, 애플리케이션에서 정의된 파라미터들 없이 제스처가 인식되도록 일련의 디폴트 파라미터들을 대신 사용할 수 있다. 제스처에 관련된 정보는 사전 제작된 애니메이션(pre-canned animation)을 위해 저장된다.

[0089] 제스처에 관련된 다양한 출력이 있다. 제스처의 발생 여부에 대해 기준 "예 또는 아니오"가 있을 수도 있다. 또한, 트래킹된 사용자의 움직임이 제스처에 해당할 가능성에 대응하는 신뢰도도 있을 수 있다. 이는 0 과 1을 포함한 그 사이의 부동 소수점 수치들에 분포하는 선형 스케일일 수 있다. 이러한 제스처 정보를 수신하는 애플리케이션이 거짓 긍정(false-positives)을 입력으로 받아들이지 않으려면, 적어도 0.95의 높은 신뢰도로 인식된 제스처들만을 사용할 것이다. 애플리케이션이 거짓 긍정이 있음에도 모든 제스처를 인식해야 하는 경우에는, 0.2보다 조금 큰 정도의 상당히 낮은 신뢰도를 갖는 제스처들을 사용할 것이다. 제스처는 가장 최근의 두 걸음 사이의 시간에 대한 출력을 가질 수 있으며, 오직 첫번째 걸음이 등록되고, (임의의 두 걸음 사이의 시간은 포지티브여야 하므로) 이는 기본값, 예컨대 -1로 설정될 수 있다. 또한 제스처는 가장 최근의 걸음에서 도달한 가장 큰 허벅지 각에 대한 출력을 가질 수도 있다.

[0090] 또 다른 제스처의 예로 "발뒤꿈치 들기 점프"가 있다. 여기서는, 사용자는 자신의 발가락을 디딘 채 발뒤꿈치를 지면에서 들어올림으로써 제스처를 취한다. 또는, 사용자는 자신의 양발(520) 전체가 지면을 떠나 공중으로 뛰어오를 수도 있다. 시스템은 양 어깨(510), 엉덩이(512) 및 무릎(516)의 각도 관계를 분석하여 이들이 똑바로 선 자세와 동일한 정렬 위치에 있는 지를 살핌으로써 이러한 제스처에 대한 골격을 파싱한다. 이후에, 임의의 위쪽을 향한 가속에 대해 이러한 포인트들과 척추 상부(526) 및 하부(528) 포인트들이 모니터링될 것이다. 충분한 가속 조합은 점프 제스처를 트리거할 수 있다. 특정 제스처가 있는 충분한 가속 조합은 전환 포

인트(transition point)의 파라미터들을 만족시킬 수 있다.

- [0091] 이러한 "발뒤꿈치 들기 점프" 제스처가 제공되면, 애플리케이션은 이 제스처에 관련된 파라미터들에 대한 값을 설정할 수 있다. 파라미터들은 제스처를 트리거하기 위해 사용자의 양 어깨(510), 엉덩이(512) 및 무릎(516)의 임의의 조합이 얼마나 빨리 위로 움직여야만 하는 지를 결정하는 상기의 가속 임계치(acceleration threshold)뿐만 아니라, 점프가 트리거될 수 있는 양 어깨(510), 엉덩이(512) 및 무릎(516)의 정렬의 최대 각도를 포함할 수 있다. 출력은 신뢰도를 비롯하여 점프 시의 사용자 신체 각도까지 포함할 수 있다.
- [0092] 제스처를 정확하게 식별하기 위해서는 제스처를 수신할 애플리케이션의 세부 사항을 기초로 하여 제스처에 대한 파라미터들을 설정하는 것이 중요하다. 제스처와 사용자 의도를 적절하게 식별하는 것은 긍정적인 사용자 경험을 만드는 데 크게 도움이 된다.
- [0093] 애플리케이션은 사전 제작된 애니메이션들이 사용되는 포인트들을 식별하기 위해서 다양한 전환 포인트들에 관련된 파라미터들의 값을 설정할 수 있다. 전환 포인트들은 다양한 파라미터들, 예컨대 특정 제스처의 식별, 속도, 타겟 또는 객체의 각도 또는 이들의 임의의 조합에 의해 정의될 수 있다. 전환 포인트가 적어도 부분적으로는 특정 제스처를 식별하는 것에 의해 정의된다면, 제스처를 적절하게 식별하는 것이 전환 포인트의 파라미터들이 충족시킨 신뢰도를 증가시키는 데에 일조하게 된다.
- [0094] 제스처에 대한 또 다른 파라미터로 움직인 거리를 들 수 있다. 사용자의 제스처들이 가상 환경의 아바타의 액션들을 제어하는 경우, 아바타는 공으로부터 팔 길이만큼 떨어져 있다. 사용자가 이 볼과 인터랙션하여 볼을 잡기를 원한다면, 사용자는 잡는 제스처를 취하면서 자신의 팔(520 - 510)을 쪽 내밀어야 한다. 이러한 상황에서, 사용자가 자신의 팔(520 - 510)을 조금만 내밀어 잡는 유사 제스처는 공과 인터랙션하는 결과를 얻지 못할 것이다. 마찬가지로, 전환 포인트의 파라미터가 잡는 제스처를 식별할 수 있으며, 만일 사용자가 자신의 팔(520 - 510)을 조금만 내민다면 공과 인터랙션하는 결과를 얻지 못하고, 사용자의 제스처는 전환 포인트의 파라미터들을 만족시키지 못할 것이다.
- [0095] 제스처 또는 그 일부는 이들이 발생하는 공간의 부피를 파라미터로 가질 수 있다. 일반적으로 이 공간의 부피는 제스처가 신체의 움직임으로 이루어지는 신체에 대하여 표현될 수 있다. 예를 들어, 오른손잡이 사용자가 풋볼 던지는 제스처는 오른쪽 어깨(510a)보다 낮지 않은 공간의 부피와 던지는 팔(502a - 510a)과 같은 편에 있는 머리 부위(522)에서만 인식될 수 있다. 이런 던지는 제스처의 공간의 부피의 모든 경계들을 정의할 필요는 없으며, 몸에서 먼 바깥쪽 경계는 정의되지 않고, 부피는 무제한으로 또는 모니터링되는 장면의 가장자리까지 확장될 수 있다.
- [0096] 도 6a 및 6b는 캡처 장치(608), 컴퓨팅 장치(610) 및 디스플레이 장치(512)를 포함하는 시스템(600)을 도시한다. 예를 들어, 캡처 장치(608), 컴퓨팅 장치(610) 및 디스플레이 장치(512)는 도 1 내지 5와 관련하여 설명한 장치들을 비롯하여 원하는 기능을 실행하는 적절한 임의의 장치를 각각 포함할 수 있다. 단일 장치로 시스템(600)의 모든 기능을 실행할 수도 있고, 적절한 장치의 조합으로 원하는 기능들을 실행할 수도 있다. 예를 들어, 컴퓨팅 장치(610)는 도 2에 도시된 컴퓨팅 환경(12) 또는 도 3의 컴퓨터에 관해 설명한 기능을 제공할 수 있다. 도 2에 도시된 바와 같이, 컴퓨팅 환경(12)은 디스플레이 장치 및 프로세서를 포함할 수 있다. 컴퓨팅 장치(610)는 자체 카메라 컴포넌트를 포함하거나, 캡처 장치(608)와 같이 카메라 컴포넌트가 있는 장치에 연결될 수도 있다.
- [0097] 본 예시들에서, 깊이 카메라(608)는 사용자(602)가 존재하는 물리적 공간(601)의 장면을 캡처한다. 깊이 카메라(608)는 깊이 이미지를 처리하고, 또는 컴퓨터(610)와 같은 컴퓨터로 깊이 이미지를 제공한다. 사용자(602)의 비주얼 형상의 디스플레이를 위해 깊이 이미지가 해석될 수 있다. 예를 들어, 깊이 카메라(608) 또는, 도시된 바와 같이, 깊이 카메라(608)가 연결된 컴퓨팅 장치(610)가 디스플레이(612)로 출력될 수 있다. 이미지 데이터의 프레임들이 캡처되어 디스플레이되는 속도에 의해 비주얼 형상이 디스플레이되는 모션의 연속성 레벨이 결정될 수 있다. 이미지 데이터의 추가적인 프레임들이 캡처되고 디스플레이될 수 있으며, 도 6a 및 6b에 도시된 프레임은 예를 들기 위해 선택되었다. 또한, 비주얼 형상은 물리적 공간(601)의 또 다른 사용자나 비인간 객체와 같이 또 다른 타겟일 수 있으며, 비주얼 형상이 부분적으로 또는 전체적으로 가상 객체일 수도 있음을 알 것이다.
- [0098] 타겟의 탐지된 특징과 닮은 특징을 갖는 타겟의 비주얼 형상을 자동 생성하는 시스템의 기능이 본원에 개시된다. 또는, 시스템은 사용자가 고를 수 있는, 선택가능한 특징들의 서브셋을 제공할 수 있다. 시스템은 타겟의 탐지된 특징들을 기초로 하여 특징들을 선택하고, 선택된 특징들을 타겟의 비주얼 형상에 적용할 수 있다. 또는, 시스템은 사용자가 고를 수 있는 옵션의 수를 줄이는 선택을 할 수 있다. 시스템이 사용자를 대신하여 결

정을 내릴 수 있다면, 사용자는 많은 결정을 하도록 요구되지도 않고, 많은 옵션에서 선택을 할 필요도 없다. 따라서, 개시된 기법에 따르면 사용자의 수고를 크게 줄일 수 있다. 예를 들어, 시스템이 사용자를 대신하여 선택을 하고, 이들 선택을 사용자의 비주얼 형상에 적용할 수 있다.

[0099] 도 6a에 도시된 바와 같이, 시스템은 물리적인 공간(601)의 사용자(602)에 대응하는 비주얼 형상(603)을 렌더링한다. 본 예시에서, 시스템은 사용자(602)의 특징들을 탐지하고, 탐지된 특징들을 특징 옵션 라이브러리와 비교하여, 사용자(602)의 탐지된 특징들과 닮은 특징 옵션들을 선택하고, 이들을 사용자의 비주얼 형상(603)에 자동으로 적용함으로써 비주얼 형상(603)을 자동 생성했다. 비주얼 형상의 자동 생성으로 인해, 사용자(602)의 수고가 덜어지고, 수월하게 게임 또는 애플리케이션 경험으로 이동하게 되므로, 사용자(602)는 마법과 같은 경험을 하게 된다.

[0100] 또한, 실시간으로 비주얼 형상을 디스플레이하고, 비주얼 형상에 적용되는 특징 선택을 실시간으로 업데이트하는 기법이 도시된다. 시스템은 물리적인 공간의 사용자를 시간에 걸쳐 트래킹하여, 비주얼 형상에 적용되는 특징들의 변화 또는 업데이트를 실시간으로 적용할 수 있다. 예를 들어, 시스템은 사용자를 트래킹하여, 사용자가 스웨터를 벗은 것을 식별할 수 있다. 시스템은 사용자 신체 움직임을 식별하고, 사용자의 옷 종류 또는 색상의 변화를 인식할 수 있다. 시스템은 사용자의 식별된 임의의 특성을 사용하여 특징 선택 프로세스를 돕고, 특징 라이브러리로부터 선택된 특징들을 업데이트하고 비주얼 형상에 적용시킬 수 있다. 따라서, 다시, 시스템은 수월하게 사용자를 애플리케이션 경험으로 이동시키고, 사용자의 탐지된 특징들이 변함에 따라, 실시간으로 그 특징들에 대응하도록 비주얼 형상을 업데이트할 수 있다.

[0101] 일 실시예에서, 사용자의 특징들을 탐지하고 탐지된 특징들을 사용하여 비주얼 형상의 특징에 관한 옵션들을 선택하기 위해, 시스템은 사용자의 모델을 생성할 수 있다. 모델을 생성하기 위해, 캡처 장치는 장면의 이미지를 캡처하고, 장면의 타겟 또는 객체들을 스캔할 수 있다. 일 실시예에 따르면, 이미지 데이터는 깊이 카메라(608) 및/또는 RGB 카메라로부터의 깊이 이미지 또는 이미지, 또는 기타 임의의 탐지기에서의 이미지를 포함할 수 있다. 시스템(600)은 장면으로부터 깊이 정보, 이미지 정보, RGB 데이터 등을 캡처할 수 있다. 그 장면 안의 타겟 또는 객체가 인간 타겟에 대응하는지를 결정하기 위해, 각각의 타겟이 플러드 필딩되어 인간 신체 모델의 패턴과 비교될 수 있다. 인간 패턴에 일치하는 각 타겟이나 객체가 스캔되어 관련 골격 모델, 플러드 모델, 메쉬 인간 모델 등의 모델을 생성할 수 있다. 이후에, 골격 모델을 트래킹하여 그 골격 모델에 연관된 아바타를 렌더링하는 컴퓨팅 환경으로 골격 모델이 공급될 수 있다.

[0102] 이미지 데이터 및/또는 깊이 정보를 사용하여 타겟의 특징들을 식별할 수 있다. 예를 들어, 인간 타겟에 관한 이러한 타겟 특징들은 키(height) 및/또는 팔길이를 포함할 수 있고, 또한 신체 스캔, 골격 모델, 사용자(602)의 화소 영역 범위 또는 기타 임의의 적절한 프로세스나 데이터를 기반으로 하여 얻어질 수 있다. 예를 들어, 인간 타겟과 연관있는 다수의 관찰된 화소들의 깊이 값 및, 키, 머리너비, 어깨너비 등 인간 타겟의 하나 이상의 양태들을 사용하여, 인간 타겟의 크기를 결정할 수 있다. 카메라(608)는 이미지 데이터를 처리하고, 이를 사용하여 사용자의 머리카락, 옷 등을 포함하는 사용자의 다양한 부위의 형태, 색상 및 크기를 결정할 수 있다. 특징 라이브러리(197)의 비주얼 형상 특징 옵션과 같이 비주얼 형상에 적용되기 위한 특징 옵션 카탈로그와 탐지된 특징들이 비교될 수 있다.

[0103] 다른 실시예에서, 사용자의 특성을 식별하고, 그 식별된 특성을 사용하여 비주얼 형상의 특징을 선택하기 위해, 시스템은 도 2와 관련해 전술한 것과 같이 타겟 디지털화 기법을 사용할 수 있다. 이 기법은 깊이 판독 장치(depth sensing device)와 같은 캡처 장치로부터 얻어지는 조직화되지 않은 포인트 클라우드로부터 표면, 질감(texture), 객체 면적(object dimensions)을 식별할 수 있다. 타겟 디지털화 기법을 사용하는 것은, 표면 추출(surface extraction), 포인트 클라우드의 포인트 식별, 법선 벡터 레이블링(labeling surface normals), 객체 속성 계산(computing object properties), 시간에 걸쳐 객체 속성의 변화 트래킹, 및 추가 프레임이 캡처됨에 따라 객체 경계 및 아이덴티티의 신뢰도(confidence) 증가 등을 포함할 수 있다. 예를 들어, 물리적인 공간의 객체와 관련된 데이터 포인트들의 포인트 클라우드가 수신되거나 관찰될 수 있다. 다음으로, 포인트 클라우드가 분석되어 포인트 클라우드가 객체를 포함하는 지를 결정하게 된다. 포인트 클라우드들의 집합이 객체로써 식별되고, 하나로 합쳐져 단일 객체를 나타낼 수 있다. 포인트 클라우드의 표면은 식별된 객체로부터 추출될 수 있다.

[0104] 알고 있는/모르는 객체, 사람, 및 장면의 배경 양태(예컨대, 바닥, 벽)를 스캔하는 기능을 제공하는 임의의 공지된 기법 또는 본원에 기술된 기법이 사용되어 물리적인 공간의 타겟의 특징들을 탐지할 수 있다. 깊이 및 RGB 데이터의 조합을 포함할 수 있는, 각 스캔 데이터가 사용되어 객체의 3 차원 모델을 제작할 수 있다. 모델의 해당 영역에 RGB 데이터가 적용된다. 프레임 간의 시간 트래킹(temporal tracking)은 신뢰도를 높이고,

실시간으로 객체 데이터를 조정할 수 있다. 따라서, 객체 속성 및 시간에 걸친 객체 속성의 변화 트래킹을 통해, 프레임 간에 자세 및 방향이 변하는 객체들을 실시간으로 신뢰도 있게 트래킹할 수 있다. 캡처 장치는 인터랙션 속도(interactive rates)로 데이터를 캡처함으로써, 데이터의 정확도(fidelity)를 높이고, 또한 개선된 기법으로 실제 깊이 데이터를 처리하고, 장면의 객체들을 디지털화하고, 객체의 표면 및 질감을 추출하며, 이러한 임의의 기법을 실시간으로 실행하여, 디스플레이가 장면의 실시간 묘사를 할 수 있게 한다.

[0105] 카메라 인식 기술을 사용하여 사용자(602)의 특성과 가장 닮은 특징 라이브러리(197)의 엘리먼트들을 결정할 수 있다. 시스템은 안면 인식 및/또는 신체 인식 기법을 사용하여 사용자(602)의 특징을 탐지할 수 있다. 예를 들어, 시스템은 이미지 데이터, 포인트 클라우드 데이터, 깊이 데이터 등으로부터 모델을 생성하는 것을 기초로 하여 사용자의 특징을 탐지할 수 있다. 얼굴이 스캔되고, 시스템은 사용자의 얼굴 특징 및 RGB 데이터에 관해 캡처된 데이터를 처리할 수 있다. 일 실시예에서, 다섯 개의 핵심 데이터 포인트(즉, 양쪽 눈, 입의 양끝 포인트 및 코)의 위치에 근거하여, 시스템은 플레이어에 대한 얼굴을 추천할 수 있다. 얼굴 추천에는 적어도 하나의 선택된 얼굴 특징, 얼굴 특징 전체 셋이 포함되거나, 또는 특징 라이브러리(197)의 얼굴 특징에 대한 옵션의 서브셋으로 좁혀질 수도 있다. 시스템은 신체 인식 기법을 실행하여, 신체 스캔을 통해 다양한 신체 부위/타입을 식별할 수 있다. 예를 들어, 사용자의 신체 스캔을 통해 사용자의 키(height)가 제안될 수 있다. 이러한 임의의 스캔의 경우, 사용자는 최상의 스캔 결과를 낳는 자세로 물리적인 공간에 서도록 제시받을 수 있다.

[0106] 캡처된 데이터로부터 다른 특징들이 탐지될 수 있다. 예를 들어, 시스템은 사용자 및/또는 사용자의 모델을 분석하여 색상 데이터 및 옷 데이터를 탐지할 수 있다. 시스템은 이러한 사용자 특성의 아이덴티티에 기반하여 사용자에게 옷을 추천할 수 있다. 옷 추천은 사용자의 옷장에 있는 옷 또는 가상 세계 장터에서 구매할 수 있는 옷일 수 있다. 예를 들어, 사용자는 특정 비주얼 형상과 관련하여 소유한 아이템들의 저장소인 개인 옷장을 가질 수 있다. 개인 옷장에는 사용자가 옷을 보고 수정가능하게 하는 인터페이스 및 사용자의 비주얼 형상에 적용되는 기타 아이템들이 포함될 수 있다. 예를 들어, 액세서리, 신발 등이 수정될 수 있다. 사용자의 성별은 캡처된 데이터를 기초로 하여 결정되거나, 또는 사용자에 관련된 프로파일에 액세스하여 결정될 수 있다.

[0107] 시스템은 적어도 하나의 사용자 특징을 탐지하고, 탐지된 특징을 나타내는 특징을 특징 라이브러리(197)에서 선택할 수 있다. 시스템은 선택된 특징을 사용자의 비주얼 형상(603)에 자동으로 적용할 수 있다. 따라서, 사용자의 비주얼 형상(603)은 시스템에서 선택한 대로 사용자와 닮게 된다. 예를 들어, 특징 추출 기법으로 사용자의 얼굴 특징들을 매핑시키고, 특징 라이브러리에서 선택된 특징 옵션들을 사용하여 사용자의 만화 형상을 제작할 수 있다. 특징 라이브러리에서 선택된, 사용자의 탐지된 특징과 유사한 특징으로 비주얼 형상(603)이 자동으로 생성되지만, 본 예시에서 비주얼 형상은 사용자(602)의 만화 버전이다. 비주얼 형상은 사용자(602)의 머리카락, 눈, 코, 옷(예컨대, 청바지, 재킷, 신발), 자세 및 타입의 만화 버전을 갖는다. 시스템은 특징을 적용하고 자동 생성된 비주얼 형상(603)을 렌더링하여 제작된 비주얼 형상(603)을 사용자(602)에게 보여준다. 사용자(602)는 자동 생성된 비주얼 형상(603)을 수정하거나, 비주얼 형상에 적용하기 위한 선택을 계속할 수 있다.

[0108] 또한, 물리적인 공간(601)에서 탐지된 사용자의 비주얼 형상은 애니메이션, 캐릭터, 아바타 등의 다른 형태를 취할 수 있다. 도 6b에 도시된 비주얼 형상의 예시는 원숭이 캐릭터(605)이다. 사용자(602)는 사용자의 온스크린 표현을 위해 시스템이나 애플리케이션에서 제공되는 다양한 저장 모델에서 선택할 수 있다. 예를 들어, 야구 게임 애플리케이션에서, 사용자(602)를 시각적으로 표현하도록 이용가능한 저장 모델은 유명한 야구 선수의 형상에서부터 태피 조각 또는 코끼리, 커서 또는 손 기호와 같은 상상 캐릭터나 기호 등을 포함할 수 있다. 도 6b에 도시된 예시에서, 원숭이 캐릭터(605)는 시스템 또는 애플리케이션에서 제공되는 저장 모델 형상일 수 있다. 저장 모델은 프로그램과 함께 패키징된 것과 같이 애플리케이션 전용이거나, 또는 애플리케이션 간에 또는 시스템 전반에서 이용가능할 수 있다.

[0109] 비주얼 형상은 사용자(602) 특징들과, 애니메이션 또는 저장 모델의 결합일 수 있다. 예를 들어, 원숭이 형상(605)은 원숭이의 저장 모델로 초기화되지만, 특징 라이브러리(197)에서와 같이, 특징 옵션 카탈로그로부터 시스템(600)이 선택한, 사용자를 닮은 특징들로 다양한 원숭이 특징들이 수정될 수 있다. 시스템은 저장 모델로 비주얼 형상을 초기화한 후, 사용자의 특징들을 탐지하고, 탐지된 특징들과 특징 라이브러리(197)를 비교하고, 사용자와 유사한 특징들을 선택하고, 선택된 특징들을 원숭이 캐릭터(605)에 적용하게 된다. 따라서, 원숭이(605)는 원숭이의 몸을 가지지만, 눈썹, 눈 및 코 등의 사용자의 얼굴 특징을 갖게 된다. 사용자의 표정, 자세, 한 말 또는 기타 임의의 탐지가능한 특성이 가상 원숭이(605)에게 적용되고, 적절하다면 수정될 수

있다. 예를 들어, 사용자가 물리적 공간에서 얼굴을 찌푸리고 있다. 시스템은 이러한 얼굴 표정을 탐지하고, 사용자의 찌푸림과 가장 닮은 찌푸림을 특징 라이브러리로부터 선택하고, 선택된 찌푸림을 원숭이에게 적용하여, 가상 원숭이도 얼굴을 찌푸리게 한다. 또한, 그 자세의 원숭이의 몸 형태와 크기에 대응하도록 수정되는 것을 제외하고, 원숭이는 사용자와 유사한 자세로 앉아 있다. 시스템(600)은 탐지된 타겟 체형 특징을 체형에 관한 일련의 비주얼 형상 특징들을 저장한 특징 라이브러리(197)와 비교할 수 있다. 시스템은 특징 라이브러리의 원숭이 특징들의 서브셋으로부터 특징들을 선택할 수 있다. 예를 들어, 애플리케이션은 그 애플리케이션과 함께 사전에 패키지화된 저장 모델 원숭이 캐릭터 옵션에 해당하는, 특징 라이브러리의 원숭이 전용 특징 옵션들을 제공할 수 있다. 시스템 또는 사용자는 사용자의 탐지된 특징들과 가장 닮은, 원숭이 특징에 관한 옵션을 선택할 수 있다.

[0110] 시스템이 특징 라이브러리(197)의 특징들의 서브셋을 제공하는 것이 바람직하다. 예를 들어, 특징 라이브러리(197)의 하나 이상의 옵션이 사용자의 탐지된 특징과 유사할 수 있다. 시스템은 사용자가 선택가능한 특징들의 작은 서브셋을 제공할 수 있다. 사용자가 수십, 수백, 수천 개의 특징 옵션으로부터 수동으로 고르는 대신, 시스템이 좁혀진 옵션의 서브셋을 제공할 수 있다. 예를 들어, 도 7은 도 6a 및 6b에서와 같은 시스템(600)을 도시한다. 디스플레이(612)에서, 시스템은 비주얼 형상의 머리에 관한 특징 옵션들의 예시적인 셋, 옵션 1 - 10을 디스플레이한다. 도 6a에서, 시스템은 사용자의 비주얼 형상에 적용되도록 머리 옵션 #5를 자동으로 선택하였다. 그러나, 도 7에 도시된 예시에서는, 시스템은 사용자의 탐지된 머리 특징들과 가장 유사한 머리 옵션들의 서브셋(702)을 선택했다. 따라서, 사용자는 사용자의 비주얼 형상에 적용하기 위해 옵션들의 서브셋(702)으로부터 선택할 수 있다.

[0111] 본 예시에서, 머리에 관한 특징 옵션들의 서브셋(702)은 사용자의 머리 모양, 색상 및 타입을 포함하는, 신체 및 얼굴 스캔으로부터 탐지된 사용자의 특징과 가장 유사한 선택들을 포함할 수 있다. 선택가능한 압도적인 수의 머리 옵션들 대신에, 시스템은 사용자의 머리 모양, 색상 및 타입과 가장 유사한 머리 옵션들에 관한 짧은 목록의 옵션들을 제공할 수 있다. 시스템이 비주얼 형상을 자동 생성할 수도 있지만, 사용자가 선택할 수 있는 하나 이상의 옵션을 제공하여 사용자가 자신을 가장 만족시키는 특징 옵션들 간에 마지막 세밀한 선택을 할 수 있도록 설계될 수도 있다. 옵션의 서브셋은 모든 옵션을 평가해야 하는 사용자의 수고를 덜어준다.

[0112] 사용자 또는 애플리케이션은 사용자의 특성에 대응하는 특정 특징들을 비주얼 형상에 적용하기 전에 이들을 수정하는 세팅을 가질 수 있다. 예를 들어, 시스템은 캡처된 데이터(예컨대, 체형/크기)를 기초로 하여 특정 몸무게 영역을 탐지할 수 있다. 그러나, 사용자가 실제 사용자의 몸무게 영역이 아닌 특정 몸무게 영역에서 디스플레이되도록, 사용자가 설정하거나 또는 시스템 자체에서 디폴트 값을 설정할 수 있다. 따라서, 예를 들어, 과체중인 사람의 경우, 사용자보다 더 좋게 보이는 비주얼 형상이 디스플레이될 수 있다. 다른 예시에서, 사용자의 얼굴 특징들이 탐지되고, 사용자의 비주얼 형상에 적용되는 특징들은 이 탐지된 특징들에 대응하여, 비주얼 형상의 얼굴 특징들이 크기, 비율, 머리의 공간 배치 등에서 사용자의 특징들과 닮도록 할 수 있다. 사용자는 특징들을 바꿈으로써 안면 인식 기법의 실제 효과를 수정할 수 있다. 예를 들어, 사용자는 슬라이딩 스케일(sliding scale)을 변경하여 특징들을 수정할 수 있다. 사용자는 슬라이딩 스케일을 변경하여, 비주얼 형상에 적용할 몸무게를 수정하거나, 또는 비주얼 형상에 적용되는 코의 크기를 변경할 수 있다. 따라서, 시스템에 의해 선택된 일부 특징들이 적용되고, 다른 특징들은 수정된 후 적용될 수 있다.

[0113] 디스플레이를 위해 시스템에서 탐지하는 특정 타겟 특성들이 수정될 수 있다. 예를 들어, 비주얼 형상의 형태, 애플리케이션, 애플리케이션의 상태 등에 상응하도록 타겟 특성들이 수정될 수 있다. 예를 들어, 비주얼 형상이 상상 캐릭터라면, 특정 특성들이 사용자의 비주얼 형상에 직접적으로 매핑되지 않을 수 있다. 사용자(602)와 유사한 신체 비율이, 아바타 (603)와 같은 사용자의 임의의 비주얼 형상 또는 원숭이(605)와 같은 사용자의 캐릭터 형상에 제공되지만, 특정 캐릭터에 대해서는 수정될 수 있다. 예를 들어, 사용자(602)와 비슷한 키(height)가 원숭이 형상(605)에 제공되지만, 원숭이의 팔은 사용자의 팔보다 상당히 길다. 시스템에서 식별된 대로, 원숭이(605)의 양팔의 움직임은 사용자의 양팔의 움직임에 대응하지만, 시스템은 원숭이의 양팔이 움직이는 방식을 반영하도록 원숭이의 양팔의 애니메이션을 수정할 수 있다.

[0114] 시스템은 스캔된 데이터, 이미지 데이터 또는 깊이 정보 등의 캡처된 데이터를 사용하여 다른 타겟 특성들을 식별할 수 있다. 타겟 특성은 타겟의 기타 임의의 특징, 예를 들어, 눈 크기, 유형 및 색상; 머리카락 길이, 유형 및 색상; 피부색; 옷 및 옷 색상을 포함할 수 있다. 예를 들어, 색상은 대응 RGB 이미지에 근거하여 식별된다. 또한, 시스템은 이 탐지가능한 특징들을 비주얼 형상에 매핑할 수 있다. 예를 들어, 시스템은 사용자가 안경을 쓰고 빨간 셔츠를 입고 있음을 탐지하고, 본 예시에서의 사용자의 비주얼 형상인 가상 원숭이(605)에 안경과 빨간 셔츠를 적용할 수 있다.

- [0115] 또한, 깊이 이미지 및 타겟 특성은 특정한 제스처, 목소리 인식 정보 등 특정 사용자(602)와 관련된 정보를 포함하는 추가적인 정보와 결합될 수 있다. 모델을 트래킹하고, 그 모델에 관련된 비주얼 형상을 렌더링하며, 모델을 기반으로 하여 컴퓨팅 환경(610)에서 실행되는 애플리케이션에서 어떤 제어를 실행할지를 결정하는 컴퓨팅 환경(610)으로 모델이 제공될 수 있다.
- [0116] 도 8은 특징 선택 내역을 사용자에게 제공하는 방법의 일례를 도시한다. 특징 선택 내역은, 특징이 적용된 비주얼 형상의 디스플레이 또는 사용자가 선택할 수 있는 좁혀진 옵션 서브셋을 갖는 특징 라이브러리의 서브셋에 의해 제공될 수 있다. 예를 들어, 802에서, 시스템은 사용자 또는 비인간 객체 등의 타겟을 포함하는 물리적인 공간으로부터 데이터를 수신한다.
- [0117] 전술한 바와 같이, 캡처 장치는 장면의 깊이 이미지와 같은 장면의 데이터를 캡처하고 장면의 타겟들을 스캔할 수 있다. 캡처 장치는 장면의 하나 이상의 타겟이 사용자와 같은 인간 타겟에 해당하는지를 결정한다. 예를 들어, 장면의 타겟 또는 객체가 인간 타겟에 해당하는 지를 결정하기 위해, 각 타겟이 플러드 필딩되어 사용자 신체 모델의 패턴과 비교될 수 있다. 다음으로, 인간 신체 모델과 일치하는 각 타겟 또는 객체가 스캔되어, 그에 연관된 골격 모델을 생성하게 된다. 예를 들어, 사람으로 식별된 타겟이 스캔되어 관련된 골격 모델을 생성할 수 있다. 이후에, 이 골격 모델은 골격 모델을 트래킹하고 그에 연관된 비주얼 형상을 렌더링하는 컴퓨팅 환경에 제공된다.
- [0118] 804에서, 시스템은 물리적인 공간의 타겟의 특징들을 식별하기 위해, 신체 스캔, 포인트 클라우드 모델, 골격 모델, 플러드 필딩 기법 등 적절한 임의의 기법을 사용하여 캡처된 데이터를 해석한다.
- [0119] 806에서, 시스템은 타겟의 특성들을 탐지하고, 이들을 특징 라이브러리의 특징 옵션 등의 특징 옵션들과 비교한다. 특징 옵션들은 타겟의 다양한 특징들에 관한 옵션들의 컬렉션이다. 예를 들어, 사용자에게 관한 특징 옵션들로 눈썹 옵션, 머리카락 옵션, 코 옵션 등을 들 수 있다. 방안에 있는 가구에 관한 특징 옵션들은 크기 옵션, 형태 옵션, 하드웨어 옵션 등을 포함할 수 있다.
- [0120] 일 실시예에서, 시스템은 사용자의 탐지된 특징과 유사한, 비주얼 형상에 적용가능한 몇몇 특징을 탐지할 수 있다. 806에서, 시스템은 사용자의 특징을 탐지하고, 사용자의 비주얼 형상에 적용하기 위해 이 탐지된 특징을 특징 라이브러리(197)와 비교하고, 810에서 시스템은 탐지된 특징에 기반하여 특징 옵션들의 서브셋을 선택한다. 시스템은 사용자의 탐지된 특징과 특징 라이브러리(197)의 특징들의 유사점을 비교하여, 그 특징들과 같은 서브셋을 선택할 수 있다. 간혹, 특징이 아주 유사하지만, 801에서 시스템은 여전히 선택가능한 옵션의 서브셋을 사용자에게 제공할 수 있다. 이러한 방식으로, 사용자는 사용자의 해당 특징과 적어도 비슷한 특징을 특징 라이브러리로부터 선택할 수 있으며, 그 서브셋으로부터 더 좋아 보이는 특징을 선택할 수도 있다. 812에서, 시스템은 옵션의 서브셋에서의 사용자의 선택을 수신한다. 따라서, 사용자는 자신과 유사한 특징을 위해, 특정 특징에 관한 옵션의 라이브러리 전체를 필터링할 필요가 없다. 시스템은 옵션의 라이브러리를 필터링하여, 사용자에게 선택할 수 있는 특징들의 서브셋을 제공할 수 있다.
- [0121] 814에서 시스템은 사용자의 비주얼 형상을 자동 생성한다. 따라서, 타겟의 탐지된 특징들과 특징 라이브러리의 옵션들의 비교에 따라, 시스템은 비주얼 형상에 적용될 특징들을 자동으로 선택함으로써 타겟의 비주얼 형상을 자동 생성할 수 있다. 시스템이 타겟의 탐지된 특징들과 유사한 특징들을 특징 라이브러리로부터 자동으로 선택하여, 사용자에게 대응하는 비주얼 형상을 자동으로 렌더링할 때, 타겟은 쉽게 시스템 또는 소프트웨어 경험을 하게 된다.
- [0122] 비주얼 형상은 자동으로 선택된 특징들과 시스템에서 제공한 옵션들의 서브셋에 기반하여 사용자가 선택한 특징들의 결합을 포함할 수 있다. 따라서, 비주얼 형상의 일부가 생성되고, 일부는 사용자가 커스터마이징할 수 있다.
- [0123] 816에서, 시스템 및/또는 사용자의 선택이 타겟의 비주얼 형상에 적용된다. 시스템은 사용자에게 비주얼 형상을 렌더링한다. 818에서, 시스템은 시간에 걸쳐 타겟의 탐지가능한 특징들을 트래킹하면서, 물리적인 공간의 타겟을 계속 모니터링한다. 타겟의 탐지된 특징들에 대한 임의의 변화를 반영하기 위해 타겟의 비주얼 형상이 실시간으로 변화할 수 있다. 예를 들어, 사용자가 타겟이고, 물리적인 공간에서 사용자가 스웨터를 벗으면, 시스템은 새로운 셔츠 스타일 및/또는 색상을 탐지하여, 사용자의 셔츠와 가장 유사한 옵션을 특징 라이브러리로부터 자동으로 선택할 수 있다.
- [0124] 선택된 옵션이 실시간으로 사용자의 비주얼 형상에 적용될 수 있다. 따라서, 이전 단계들이 실시간으로 처리되어 디스플레이가 물리적인 공간에 실시간으로 대응할 수 있다. 이러한 방식으로, 물리적인 공간의 객체, 사

용자 또는 모션이 디스플레이를 위해 실시간으로 해석되어, 사용자가 실행중인 애플리케이션과 실시간으로 인터랙션할 수 있다.

- [0125] 사용자의 탐지된 특징, 시스템에서 선택된 특징 및 사용자가 선택한 임의의 특징들은 프로파일의 일부가 된다. 예를 들어, 프로파일은 특정 물리적인 공간 또는 사용자 전용일 수 있다. 사용자의 특징들을 포함하는 아바타 데이터는 사용자의 프로파일의 일부가 된다. 프로파일은 사용자가 캡처 장면에 등장하면 액세스될 수 있다. 프로파일이 비밀번호, 사용자의 선택, 신체 크기, 목소리 인식 등에 기반하여 사용자와 일치하면, 그 프로파일은 사용자의 비주얼 형상을 결정하는 데에 사용될 수 있다. 사용자의 프로파일에 정보를 저장하면서 사용자의 과거 데이터가 모니터링된다. 예를 들어, 시스템은 얼굴 특징, 체형 등 사용자 특유의 특징들을 탐지할 수 있다. 시스템은 타겟의 비주얼 형상에 적용되고 타겟 프로파일에 저장되기 위한, 탐지된 특징들과 유사한 특징들을 선택할 수 있다.
- [0126] 도 9는 물리적인 공간(601)의 타겟에 관하여 수신된 정보를 처리하고, 타겟 디지털화 기법을 사용하여 타겟을 식별할 수 있는 도 6a 및 6b의 시스템(600)의 일례를 도시한다. 캡처된 타겟들은 가상 환경의 그 타겟들의 비주얼 형상에 매핑될 수 있다. 본 예시에서, 물리적인 장면은 도 9에 도시된 물리적인 공간에 보이는 공(102), 상자(104), 창의 차양(window shade, 106), 벽 레일(wall rail, 108), 벽 #1(110), 벽 #2(112) 및 바닥(115)을 포함한다. 이 장면에는 사용자(602)도 있다. 일 실시예에서, 시스템(600)은 이러한 객체들(102, 104, 106, 108, 110, 112 및 115), 및 사용자(602)와 같은 인간 타겟 등 기타 타겟들을 인식, 분석 및/또는 트래킹할 수 있다. 시스템(600)은 각각의 객체(102, 104, 106, 108, 110, 112 및 115)에 관련된 정보 및/또는 물리적인 공간에서의 사용자(602)의 제스처를 수집할 수 있다. 또한, 사용자(602)와 같은 사용자가 물리적인 공간에 들어갈 수 있다.
- [0127] 물리적인 공간(601)에서 임의의 객체 또는 사용자가 타겟이 될 수 있다. 예를 들어, 캡처 장치(608)는 물리적인 공간(601)에 있는 사람(602) 또는, 공(102), 종이 상자(104)와 같은 비인간 객체를 스캔할 수 있다. 본 예시에서, 시스템(600)은 캡처 장치(608)를 사용하여 물리적인 공간(601)을 스캔하여 타겟을 캡처한다. 예를 들어, 깊이 카메라(608)가 실제 깊이 데이터를 수신한다. 시스템(600)은 실제 깊이 데이터를 처리하고, 이 깊이 데이터를 포인트 클라우드 데이터로 해석하여, 이 포인트 클라우드 데이터를 법선 벡터로 변환한다. 예를 들어, 깊이 버퍼(depth buffer)가 캡처되어 정해진 포인트 클라우드로 변환된다.
- [0128] 깊이 버퍼는 렌더링된 각 픽셀의 깊이를 기록하는 버퍼이다. 깊이 버퍼는 렌더링된 추가 픽셀들을 기록하며, 렌더링된 다른 픽셀들의 깊이 간의 관계를 결정할 수 있다. 예를 들어, 깊이 버퍼는 은면 제거(hidden surface removal)를 실행하고, 렌더링된 각 픽셀과 프레임 버퍼의 그 위치에 이미 있는 픽셀과 비교한다. 깊이 버퍼는 z 버퍼라고도 하며, 캡처 장치와 캡처된 이미지의 각 가시 포인트(visible point)까지의 측정 거리를 저장하는 프레임 버퍼로 이루어질 수 있다.
- [0129] 시스템(600)은 식별된 포인트 클라우드 및 법선 벡터들을 기초로 하여, 장면에서 파싱된 객체들을 레이블링하고(label), 노이즈를 제거하고, 각 객체에 대한 방향을 계산할 수 있다. 객체 주위에 경계 상자(bounding box)가 형성될 수 있다. 이후에, 이 객체는 질감 추출을 위해 프레임마다 트래킹된다.
- [0130] 일 실시예에 따르면, 이미지 데이터는 깊이 카메라 및/또는 RGB 카메라로부터의 깊이 이미지 또는 이미지, 또는 기타 임의의 탐지기에서의 이미지를 포함할 수 있다. 예를 들어, 카메라(608)는 이미지 데이터를 처리하고, 이를 사용하여 타겟의 형태, 색상 및 크기를 결정할 수 있다. 본 예시에서, 물리적인 공간(601)의 타겟들(602, 102, 104, 106, 108, 110, 112 및 115)은 깊이 정보를 처리하는 깊이 카메라(608)에 의해 캡처되고, 컴퓨터(610)와 같은 컴퓨터에 깊이 정보를 제공하게 된다.
- [0131] 디스플레이(612)에 비주얼 형상을 디스플레이하기 위해 깊이 정보가 해석된다. 시스템은 물리적인 공간의 타겟들에 대응하는 가상 객체들을 생성하기 위해, 이 정보를 사용하여 특정 라이브러리(197)로부터 옵션들을 선택할 수 있다. 인간 패턴에 일치하는 각 타겟이나 객체가 스캔되어, 관련 골격 모델, 메쉬 인간 모델 등의 모델을 생성할 수 있다. 알고 있는 객체의 라이브러리와 일치하는 각 타겟 또는 객체가 스캔되어 그 특정 객체에 대해 이용가능한 모델을 생성할 수 있다. 또한, 모르는 객체도 스캔되어, 포인트 클라우드 데이터, RGB 데이터, 법선 벡터, 방향, 경계 상자 및 모르는 객체에 해당하는 실제 깊이 데이터의 기타 임의의 처리 등에 대응하는 모델을 생성할 수 있다.
- [0132] 이미지 데이터의 프레임들이 캡처되어 디스플레이되는 비율에 의해 비주얼 형상 디스플레이의 연속성 레벨이 결정될 수 있다. 또한, 시간이 갈수록, 프레임간 이미지(frame-to-frame images)의 수는, 포인트 클라우드 데이터가 따로따로 레이블링된 객체들로 파싱되는 방식의 신뢰도를 증가시킨다. 객체의 움직임은 법선 벡터 및

방향에 관한 깊이 정보를 제공할 수 있다. 또한, 시스템(600)은 바람직한 포인트 데이터와 노이즈를 구분할 수 있다. 또한, 시스템(600)은 캡처 데이터의 단일 프레임에서 또는 연속 프레임들에 걸쳐 사용자(602)의 자세를 평가함으로써 사용자(602)의 모션으로부터 제스처를 식별할 수 있다.

[0133] 시스템(600)은 디스플레이(612)의 비주얼 형상이 물리적인 공간(601)에서 캡처된 타겟들(602, 102, 104, 106, 108, 110, 112 및 115)과 이들의 모션들에 매핑되도록, 물리적인 공간(601)의 임의의 타겟(602, 102, 104, 106, 108, 110, 112 및 115)을 트래킹할 수 있다. 캡처 장치가 캡처하고 스캔하여 도 2의 특징 라이브러리(197)와 같은 특징 라이브러리의 특징 옵션들과 비교할 수 있는 특성들을 물리적인 공간의 객체가 가질 수 있다. 시스템은 타겟의 탐지된 특징들과 가장 유사한 특징들을 특징 라이브러리로부터 선택할 수 있다.

[0134] 타겟 디지털화의 구현에 관련된 컴퓨터 비전(computer vision) 기법들이 본원에서 기술된다. 이러한 기법들을 사용함으로써, 시스템은 특징 라이브러리로부터 선택된 타겟 특징들과 가장 유사한 특징들과 캡처된 특징들을 비교할 수 있게 된다. 컴퓨터 비전은, 실제 깊이 또는 이미지 데이터 등의 캡처된 데이터로부터 물리적인 공간의 객체의 모델을 생성함으로써 장면의 내용을 이해하는 개념을 가리킨다. 예를 들어, 기법들로는 표면 추출, 법선 벡터를 복구하기 위해 근접도(proximity)에 기반한 포인트 클라우드의 포인트 해석, 객체 속성 계산, 시간에 걸친 객체 속성 트래킹, 시간에 걸친 객체 식별 및 형태의 신뢰도 향상, 및 인간 또는 알고 있는/모르는 객체 스캐닝을 들 수 있다.

[0135] 캡처 장치는 물리적인 공간을 스캔하고, 물리적인 공간(601)에 있는 다양한 객체에 관한 영역 데이터(range data)를 수신할 수 있다. 스캔에는 객체의 표면 스캔 또는 전체 입체(entire solid) 스캔이 포함될 수 있다. 이차원 깊이 버퍼의 형태로 실제 깊이 데이터를 취함으로써, 임의의 적절한 컴퓨팅 장치로 객체의 표면에 있는 다수의 포인트들을 해석하고 포인트 클라우드를 출력할 수 있다. 포인트 클라우드란 x , y , z 좌표에 의해 정의된 데이터 포인트들과 같이, 삼차원 좌표계에서 정의된 일련의 데이터 포인트들이다. 포인트 클라우드 데이터는 스캔된 물리적인 공간의 객체들의 가시적인 표면을 나타낸다. 따라서, 객체는 장면의 객체들을 일련의 불연속 포인트들로서 표현하여 디지털화될 수 있다. 포인트 클라우드 데이터는 이차원 데이터 셋으로써 데이터 파일에 저장될 수 있다.

[0136] 영역 데이터는 깊이 카메라 또는 깊이 감지 장치(depth sensing device)를 비롯한 캡처 장치를 사용하여 실시간으로 캡처될 수 있다. 예를 들어, 데이터 프레임이 깊이 버퍼 형태의 깊이 감지 카메라를 사용하여 적어도 20 hertz의 주파수로 캡처된다. 데이터는 조직화된 샘플 포인트 클라우드로 해석되고, 각 포인트는 위치, 방향, 법선 벡터, 색상 또는 질감 특성 등 관련 타겟의 특성들을 포함한다. 포인트 클라우드 데이터는 이차원 데이터 셋에 저장될 수 있다. 캡처 장치의 광학 특성이 알려지면, 영역 데이터가 모든 삼차원 포인트 클라우드에 투사되고, 이에 의해서 정형화된(regularized) 데이터 구조로 저장될 수 있다. 삼차원 포인트 클라우드는 객체 표면의 토폴로지(topology)를 나타낸다. 예를 들어, 표면의 인접한 부분들 사이의 관계는 클라우드의 이웃한 포인트들로부터 결정된다. 포인트 클라우드 데이터는 표면으로 변환될 수 있고, 포인트 클라우드 데이터에 의해 표현되는 그 객체의 표면은 포인트 클라우드 데이터의 표면 상의 법선 벡터를 평가함으로써 추출될 수 있다. 정형화된 데이터 구조는 이차원 깊이 버퍼와 유사할 수 있다.

[0137] 포인트 클라우드는 물리적인 공간의 다양한 객체들에 관련된 많은 데이터 포인트들을 포함할 수 있다. 전술한 바와 같이 포인트 클라우드 데이터는 캡처 장치에서 수신하거나 관찰할 수 있다. 이후에 포인트 클라우드가 분석되어 객체 또는 객체들의 셋을 포함하는지를 결정할 수 있다. 데이터가 객체를 포함하면, 객체의 모델이 생성될 수 있다. 프레임들이 캡처됨에 따라, 객체 식별 시에 신뢰도가 향상하게 된다. 특정 객체에 관련된 모델의 피드백이 생성되어 실시간으로 사용자에게 제공될 수 있다. 또한, 객체의 모델이 물리적인 공간에서의 객체의 움직임에 대응하여 트래킹되어, 그 모델이 객체의 움직임을 흉내내도록 조정될 수 있게 된다.

[0138] 이 모든 것은 결과의 실시간 디스플레이 처리 속도로 실행될 수 있다. 실시간 디스플레이란, 물리적인 공간의 제스처의 수행과 동시에 혹은 거의 동시에 디스플레이되는, 제스처의 비주얼 형상의 디스플레이 또는 비주얼 보조(visual assistance)의 디스플레이를 말한다. 예를 들어, 시스템이 사용자와 사용자의 환경을 그대로 따라하는 디스플레이를 제공하는 디스플레이의 업데이트 속도는 20 Hz 이상이어서, 처리 지연은, 디스플레이 지연이 최소화되거나 또는 사용자가 전혀 알아챌 수 없을 정도로 근소하다. 따라서, 실시간이란, 자동 데이터 처리를 위해 필요한 시간만큼 딜레이된, 데이터의 적시성(timeliness)에 관련된 근소한 지연을 포함한다.

[0139] 캡처 장치는 인터랙션 속도로 데이터를 캡처함으로써, 데이터의 정확도를 높이고 또한, 개시된 기법으로 실제 깊이 데이터를 처리하고, 장면의 객체들을 디지털화하고, 객체의 표면 및 질감을 추출하며, 이러한 임의의 기법을 실시간으로 실행하여, 디스플레이에서 장면의 실시간 묘사를 할 수 있게 한다. 클라우드 포인트들의 그

를 임의의 프레임의 장면에서 불연속(discrete) 객체들로 클러스터링하기 위해(cluster), 깊이 버퍼는 왼쪽에서 오른쪽으로, 다음에 위쪽에서 아랫쪽으로 스캔 선을 따라 이동할 수 있다. 클라우드의 해당 포인트 또는 포인트들의 클러스터는 스캔 시에 처리될 수 있다.

[0140] 카메라는 깊이 및 색상 데이터를 캡처하고, 그 색상 데이터에 대응하는 포인트 클라우드의 색을 지정할 수 있다. 캡처 장치는 카메라의 시각으로 바라보기 때문에, 카메라는 삼차원으로 물리적인 공간을 표현하도록 깊이 데이터를 해석할 수 있다. 포인트들이 포인트 클라우드가 될 수 있도록, 삼차원 포인트 클라우드 데이터가 합쳐지고 연결된 뒤, 클라우드의 포인트들의 서브셋이 특정 객체로 레이블링될 수 있다. 이렇게 레이블링된 포인트 클라우드로부터, 각 레이블링된 객체 및 생성된 대응 메쉬 모델에 대한 삼차원 데이터가 복구될 수 있다. 색상 정보는 깊이 정보와 관련이 있기 때문에, 각 객체의 질감 및 표면도 추출될 수 있다. 이러한 타겟 디지털화는 운영 체제 또는 소프트웨어 애플리케이션과 같은 게임 애플리케이션 또는 비게임 애플리케이션에서 유용하다. 캡처와 실시간으로 디스플레이 장치에 피드백을 제공하고 깊이 데이터를 처리함으로써, 게임을 하는 등의 만족스러운 인터랙티브한 경험을 제공할 수 있다.

[0141] 도 9에 도시된 예시에서, 물리적인 공간에는 벽, 천장 및 바닥이 있다. 캡처 장치에서 수신된 실제 깊이 데이터를 처리하여 얻어지는 포인트 클라우드 데이터의 분석을 통해, 시스템은 벽과 바닥을 레이블링할 수 있다. 이후에, 방의 형태와 같은 물리적인 장면에 대한 추가 정보가 추출될 수 있다. 물리적인 공간에 대한 기본 정보를 사용하여, 시스템은 특징 라이브러리에서 선택하여 그 물리적인 공간에 대응하는 가상 공간을 생성할 수 있다. 예를 들어, 특징 라이브러리는 다양한 특징들의 만화 그림을 포함하고, 따라서 자동 생성되는 가상 공간이 물리적인 공간의 만화 버전일 수 있다.

[0142] 표면과 실제 깊이 데이터로부터 식별된 객체를 분리하기 위해 깊이 버퍼의 정보가 사용될 수 있다. 깊이 버퍼에 의한 제 1 단계 워크(pass walk)는 포인트 클라우드로부터 얻어진 법선 벡터를 기초로 하여 깊이 버퍼에 대한 법선 맵(normal map)을 계산하는 데 사용될 수 있다. 따라서, 공간의 개개의 포인트들 대신에, 시스템은 표면이 향하는 방향을 얻을 수 있다. 시스템은 깊이 버퍼로부터 법선 벡터들을 복구하고, 이들 법선 벡터와 연관된 클라우드의 포인트들과 함께 법선 벡터들을 저장할 수 있다. 법선 벡터는 객체의 형태 및 윤곽을 식별하는 데 사용될 수 있다. 예를 들어, 구체(sphere)의 전체 표면에서 법선의 방향은 완만한 일정량의 변화를 가진다. 다양한 객체의 법선 벡터는 장면에서 탐지된 법선 벡터와 비교되는 다양한 객체 필터들에 따라 달라질 수 있다.

[0143] 법선 벡터 계산 및 법선 맵 계산이 포인트 클라우드 데이터로부터 표면을 식별하기 위해 본원에 기술된 통상적인 기법이지만, 휴 변환(Hough Transforms), 법선 매핑(normal mapping), 푸리에 변환(Fourier transforms), 커블릿 변환(Curvelet transforms) 등의 임의의 적절한 표면 분리 또는 추출 기법이 사용될 수도 있다. 예를 들어, 포인트 클라우드로부터 표면을 분리 및/또는 추출하는 계산은 평면일 경우 휴 변환을 통해 실행될 수 있다. 이 경우에 법선 맵은 필요하지 않으며, 포인트 클라우드의 휴 변환이 실행될 수 있다. 따라서, 클라우드의 포인트들이 객체에 합쳐지고 레이블링될 때, 각 포인트에 대한 휴 공간의 평가는 포인트가 이웃한 포인트들과 한 평면에 있는지 여부를 나타내고, 시스템이 특정 객체를 구성하는 특정 평면들을 개별적으로 레이블링할 수 있게 된다. 적절한 임의의 분리/추출 기법이 사용될 수 있고, 시나리오에 따라 전반적인 레이블링 성능 및 특성에 맞춰 조정될 수 있다. 다양한 표면 분리/추출 기법을 사용하여 레이블링 휴리스틱(heuristics)을 바꿀 수도 있으며, 이러한 식별 및 레이블링을 위해 적절한 임의의 기법이 사용될 수 있고, 사용자에게 대한 디스플레이를 실시간으로 생성하고 리프레시하기 위해 시스템이 깊이 데이터를 실시간으로 처리할 수 있게 한다.

[0144] 노이즈는 사용된 깊이 센서의 유형에 기인할 수 있다. 제 1 워크 단계는 실제 데이터에 대한 노이즈 억제 단계(noise suppression pass)를 포함할 수 있다. 예를 들어, 법선 맵에서 노이즈를 제거하기 위해 스무딩 단계(smoothing pass)가 실행될 수 있다.

[0145] 클라우드의 포인트들은 데이터 셋의 이차원 스캔 단계로 레이블링될 수 있고, 서로 가까이 있고 유사하다고 식별된 표면을 갖는 옵션들은 동일한 객체에 속하는 것으로 레이블링될 수 있다. 예를 들어, 표면 분리 기법이 법선 맵 생성을 포함한다면, 서로 가까이 있고 유사하다고 식별된 표면을 갖는 데이터 셋들은 동일한 객체에 속하는 것으로 레이블링된다. 바닥 및 벽과 같이 공간적으로 연결되거나 분리된 표면들은 따로 레이블링되며, 레이블링으로 인해 평면과 완만한 곡면이 구별된다. 이웃하는 포인트들과 연결된 포인트들은 그 포인트들 간의 거리 및 유사한 방향을 향하는 해당 법선 벡터들을 기초로 하여 레이블링될 수 있다. 거리 임계치(distance threshold) 및 법선 유사성 임계치(normal similarity threshold)를 튜닝하여, 다른 크기 및 곡률(curvature)을 갖는 객체와 표면들이 불연속적인 레이블을 갖게 할 수 있다. 알고 있는 객체의 임계치 및 기

대치가 그 객체 필터에 저장될 수 있다.

- [0146] 가까운 거리에 있는 포인트 클라우드 데이터 및 포인트 클라우드들의 컬렉션으로부터 식별된 법선 벡터들의 평가를 통해 공과 상자가 구분될 수 있다. 그에 따라, 각 객체(102, 104)는 레이블링될 수 있다. 레이블링은 단순히 고유 식별자일 수도 있다. 클라우드의 포인트 위치와 법선 벡터들의 조합은 표면에 있는 객체들과 객체를 구성하는 객체들을 서로 구별하는 데 유용하다. 예를 들어, 컵이 상자(104) 위에 있는 경우, 포인트 클라우드 데이터로부터 객체들의 분리 여부가 결정되지 않았기 때문에, 컵은 상자에 부여된 것과 같은 고유 ID로 레이블링된다. 그러나, 이후에 법선 벡터를 계산함으로써, 시스템은 법선들 간에 90도 차이가 있음을 알 수 있고, 포인트들과 포인트 클라우드들의 근접도를 기초로 하여 객체들은 따로따로 레이블링되어야 한다고 결정할 수 있다. 따라서, 구조적인 표면 구성 요소와 일치하는 포인트 클라우드의 데이터 포인트 그룹들이 결합되어 레이블링될 수 있다.
- [0147] 시스템은 다양한 포인트 클라우드의 결정된 표면 방향을 재투사하고(re-project), 마치 평면에 있는 것처럼 질감을 재배치할 수 있다. 이러한 기법으로 시스템은 객체를 좀더 정확하게 질감을 재생할 수 있게 된다. 예를 들어, 사용자가 인쇄된 글이 있는 잡지를 들고 있을 때, 사용자가 캡처 장치에 대해 잡지를 들고 있는 방향에 대한 제한은 없다. 캡처 장치는 잡지 표면의 캡처된 질감을 재투사할 수 있고, 색상 정보, 글 및 기타 질감을 포함하여 질감을 재투사할 수 있다.
- [0148] 레이블링되어 일련의 계산된 파라미터를 갖는 객체에 대해, 시스템은 가상 장면에 대해 향상된 정확도, 구성(organization), 구조를 위해 분석을 실행하거나 계속할 수 있다. 예를 들어, 최적합 경계 상자(best fit bounding box)는 특정 객체를 구별할 수 있는 보다 정확한 방법이다. 최적합 경계 상자는 특정 프레임에서 객체의 방향을 제공할 수 있다. 예를 들어, 커피 컵이 위에 놓여 있는 상자에 대해 상자의 포인트 클라우드 및 커피 컵을 나타내는 포인트 클라우드를 모두 포함하는 경계 상자가 처음에 제공된다. 각 프레임에서, 객체들이 공간적으로 마지막 프레임에서와 같은 위치에 있는지를 평가하고, 방향이 유사한지 여부를 결정할 수 있다. 커피 컵은 프레임과 프레임 사이에서 이동하고, 시스템은 이 컵이 상자와 별개임을 식별하여, 컵에 대한 새로운 경계 상자를 생성하고, 종이 상자에 대한 경계 상자를 재정의할 수 있다.
- [0149] 간혹, 방안의 사소한 입자 또는 객체로 인해 또는 사용된 센서 유형에 따라 시스템에 노이즈가 들어온다. 예를 들어, 클라우드의 포인트 셋이 파리를 나타내거나, 사용된 센서의 유형이 불필요하고 무관한 포인트들을 야기시킨다. 노이즈를 줄이기 위해서, 클리닝 단계(cleaning phase)가 실행되어 센서 데이터를 클리닝하거나, 아주 작은 객체들 및 적은 수의 구성 포인트 샘플들을 갖는 객체들을 제거할 수 있다. 예를 들어, 장면에서 먼지 입자나 파리가 캡처되지만, 파리를 나타내는 적은 수의 구성 포인트 샘플들은 그 포인트 클라우드와 관련된 법선 벡터들의 아이덴티티를 트리거할 정도로 의미가 있지는 않다. 따라서, 파리를 나타내는 적은 수의 구성 포인트 샘플들이 분석에서 제외될 수 있다. 포인트 클라우드 데이터의 초기 단계(initial pass)는 큰 무리의 객체를 얻기 위해, 공간적으로 관련된 객체에 있는 포인트들을 함께 사용할 수 있다. 예를 들어, 다수의 포인트가 의자라서, 특정 ID로 레이블링되고, 또 다른 객체는 바닥일 수 있다. 분석에서 제외되어야 하는 포인트들의 셋을 식별하기 위해 특정 임계치가 설정될 수 있다. 예를 들어, 한 객체에 대해 단지 20 포인트만이 식별되고, 그 20 포인트가 물리적인 공간이나 장면의 다른 객체들과 비교하여 상대적으로 작은 영역에 공간적으로 배치되는 경우, 시스템은 그 20 포인트를 제거할 수 있다.
- [0150] 객체가 차지하는 전체 부피/공간에 대한 빠른 측정 방법으로 축 정렬 경계 상자(axis aligned bounding box)가 사용될 수 있다. 축 정렬이란 공간에서 물체의 축이 아니라, X, Y 또는 Z와 같이 특정한 축을 의미한다. 예를 들어, 시스템은 표면이 복잡하거나 단순한지를 계산할 수 있다(예컨대, 구체나 잡지는 단순한 표면을 가지고, 인형이나 식물은 복잡한 표면을 갖는다). 객체의 회전은 시스템이 객체의 정교한 특성을 분석하고 결정하는 데에 유용하게 쓰일 수 있다. 캡처 장치는 부피 판단을 위해 객체의 입체 스캔을 실행할 수 있다. 또한, 캡처 장치는 포인트 클라우드와 장면에 있는 객체들 간의 레퍼런스를 제공하여, 물리적인 공간과 관련하여 객체의 특정 위치를 식별할 수 있게 한다.
- [0151] 객체의 속성을 계산하고 시간에 걸쳐 이들 변화를 트래킹함으로써, 프레임간에 위치나 방향이 변하는 객체들을 실시간으로 트래킹할 수 있는 신뢰도 있는 기법이 확립될 수 있다. 변화를 캡처하기 위해 임시 정보를 사용함으로써, 보다 많은 프레임들이 캡처됨에 따라 장면의 객체를 파싱, 식별 및 레이블링하는 신뢰도를 높일 수 있다. 640 x 480 포인트와 같은 일반적인 데이터 셋의 크기로 인해, 본원에 개시된 기법을 사용하여 보다 복잡한 처리도 가능해진다. 데이터는 적어도 20 Hertz의 주파수로 프레임 시퀀스에서 캡처될 수 있다.
- [0152] 객체 파라미터들은 이전 프레임의 파라미터들과 비교되어, 움직이는 객체들은 실시간으로 트래킹되도록 다시

레이블링되고, 정적인 객체는 계속 레이블을 유지한다. 각 객체에 대한 신뢰도가 계산되고, 신뢰도 요소는 시간이 감에 따라 증가할 수 있다. 따라서, 정적인 객체는 오클루전(occlusion)으로 인해 시야에서 들락날락하는 한편, 객체의 신뢰도는 높게 유지될 수 있다. 시간 분석법(temporal analysis)은 바로 이전 프레임과 현재 프레임의 평가를 포함할 수 있다. 객체가 각 프레임에서 동일하다면, 객체는 이전 프레임에서 갖고 있던 레이블로 다시 레이블링되어, 프레임 간에 레이블 및 객체에 일관성을 제공할 수 있다. 카메라 주변 환경에 관련된 통계 데이터를 수집하고 깊이 카메라의 방향을 추정하기 위해 객체 및 표면 방향 및 위치가 사용될 수 있다. 예를 들어, 여러 경우에 주요 평면 위치는 벽과 바닥에 해당될 것이다.

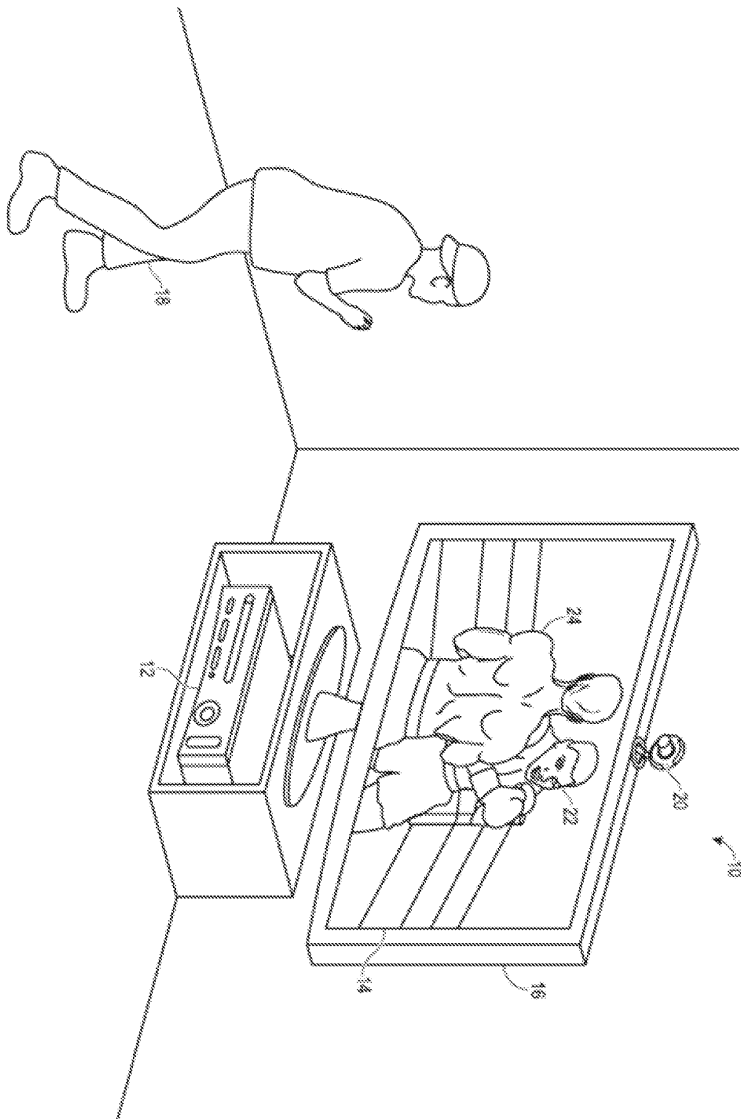
[0153] 본원에 기술된 구성 및/또는 접근 방법은 예시적인 것으로, 이러한 특정 실시예들이나 예시들로 제한되는 것은 아님을 이해할 것이다. 본원에 기술된 특정 루틴 또는 방법들은 하나 이상인 임의의 수의 처리 전략을 나타낸다. 따라서, 도시된 다양한 행위들은 도시된 순서대로, 다른 순서로 또는 동시에 실행될 수 있다. 마찬가지로, 전술한 프로세스의 순서도 변경될 수 있다.

[0154] 또한, 본 개시는 특정 양태들과 관련되어 설명되었지만, 다양한 도면에 도시된 바와 같이, 본 개시를 벗어나지 않고 그 동일한 기능을 실행하도록 다른 유사한 양태들이 사용되거나 또는, 설명된 양태들에 대해 변경 및 추가될 수 있다. 본 개시의 대상은, 본원에 기술된 다양한 프로세스, 시스템 및 구성의 신규하고 자명하지 않은 모든 컴비네이션 및 서브 컴비네이션, 및 기타 특징, 기능, 행위 및/또는 특성들뿐만 아니라 그들의 임의의 모든 등가물을 포함한다. 따라서, 개시된 실시예들의 방법 및 장치, 또는 특정 양태들이나 그 일부분은, 플로피 디스켓, CD-ROM, 하드 드라이브 또는 기타 임의의 기계 판독가능 저장 매체 등의 실제 매체에 구현되는 프로그램 코드(즉, 인스트럭션)의 형태를 취할 수 있다. 프로그램 코드가 컴퓨터와 같은 기계에 로딩되어 실행된다면, 그 기계는 개시된 실시예를 실시하도록 구성된 장치가 된다.

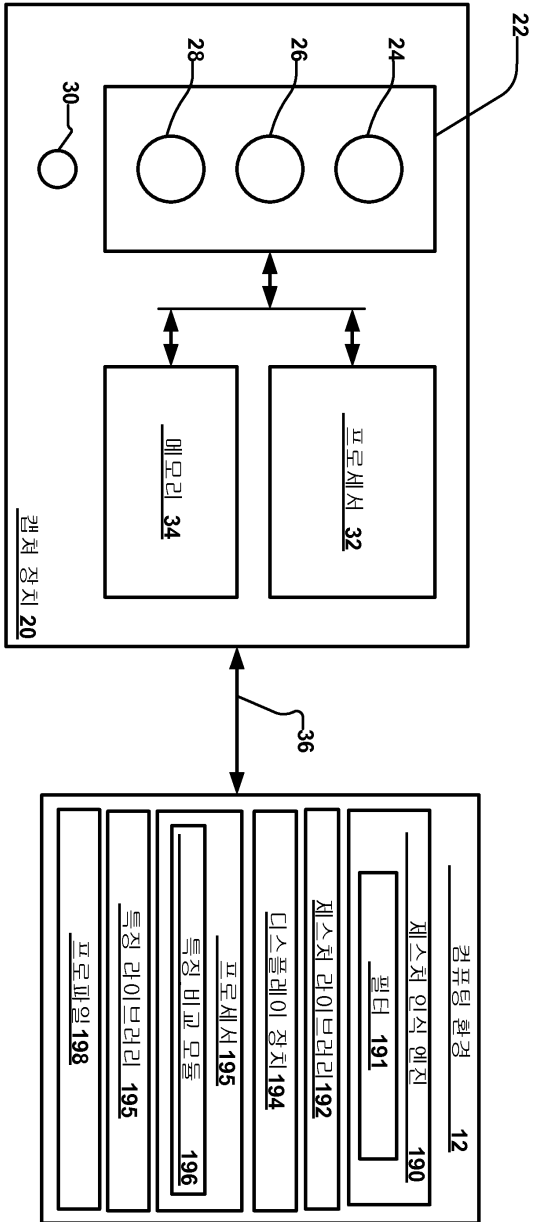
[0155] 본원에 명시된 특정 구현예뿐만 아니라, 본원에 기술된 상세한 설명을 참고하여 그 밖의 양태들 및 구현예들도 당업자에게 자명할 것이다. 따라서, 본 개시는 임의의 단일 양태에 제한되지 않고, 첨부된 특허청구범위에 따른 폭과 범위로 해석되어야 한다. 예를 들어, 본원에서 설명된 다양한 절차들은 하드웨어 또는 소프트웨어 또는 이들의 조합으로 구현될 수 있다.

도면

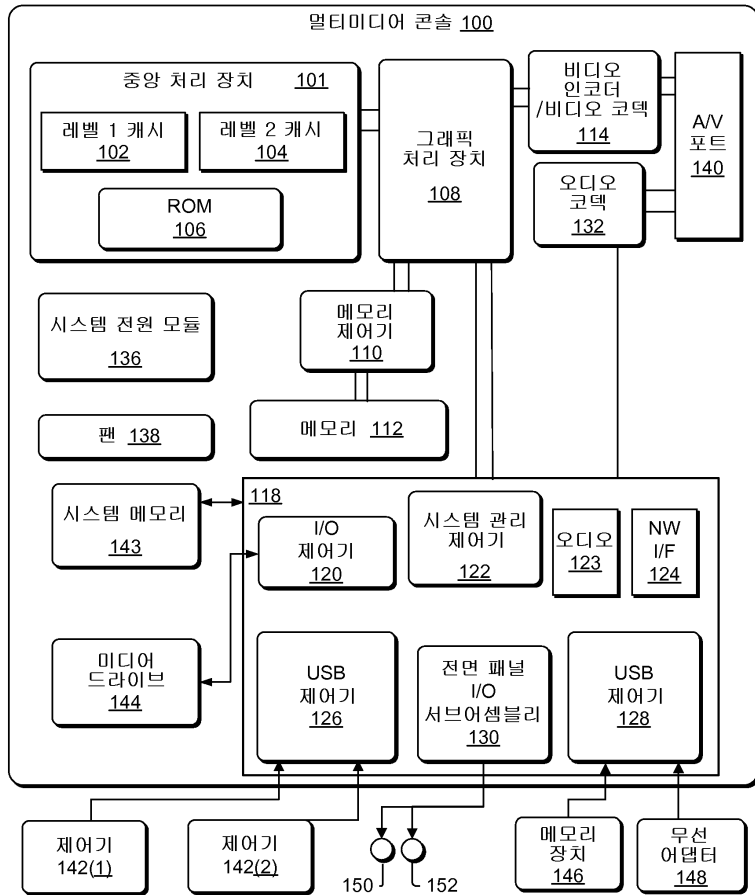
도면1



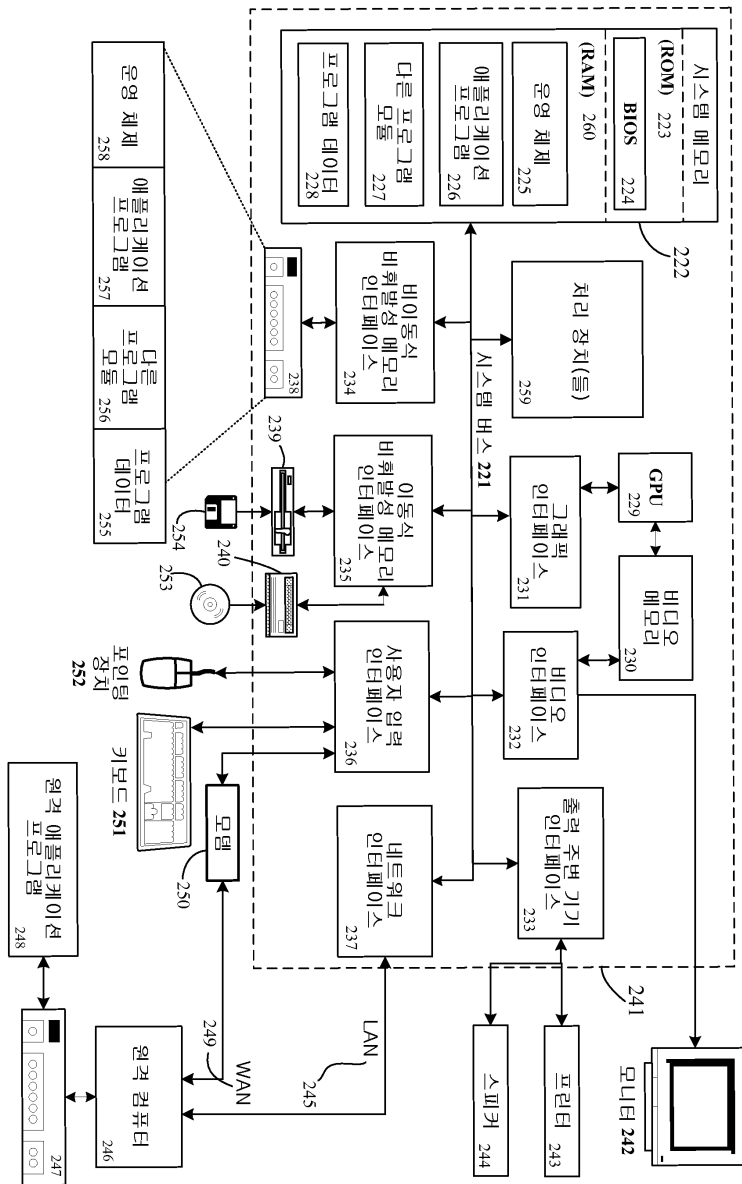
도면2



도면3

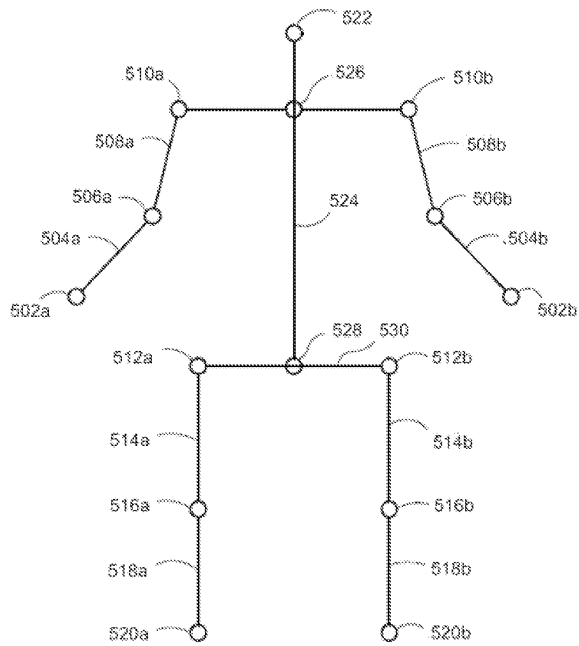


도면4

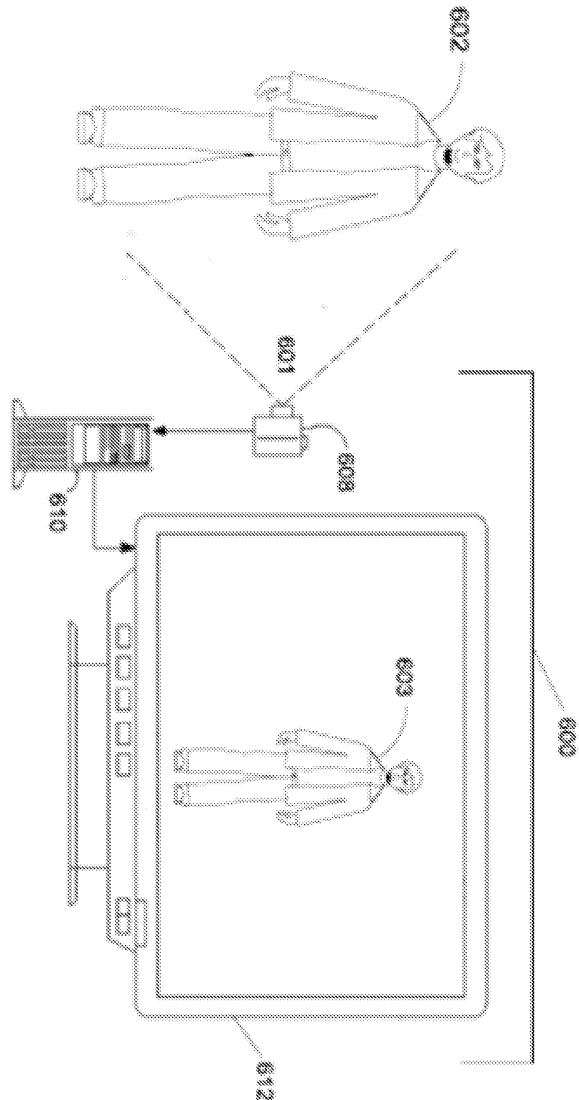


컴퓨터링 환경 220

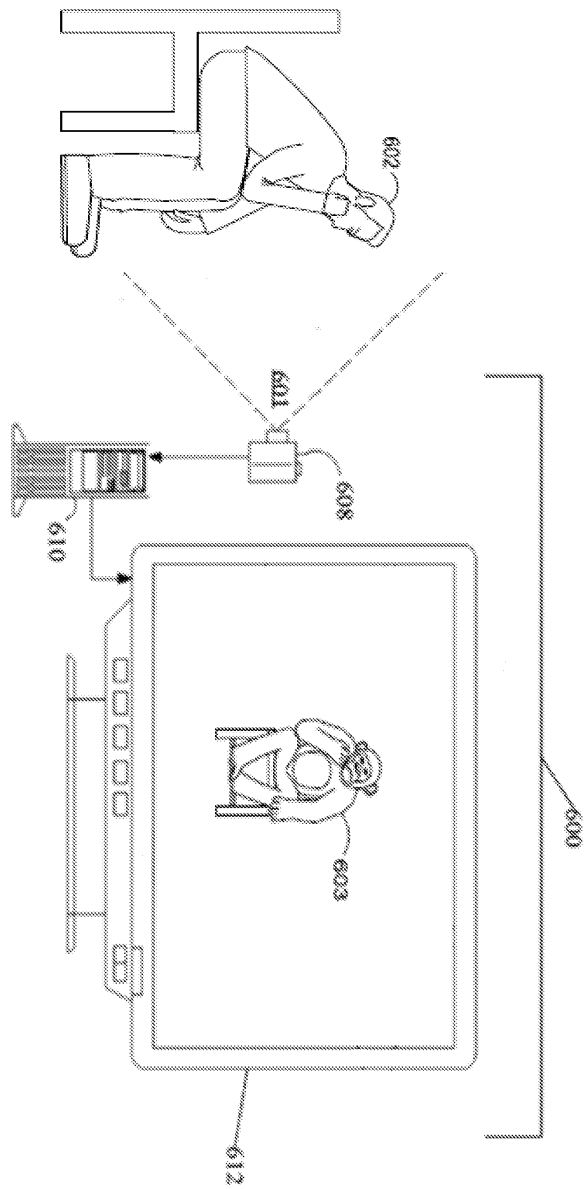
도면5



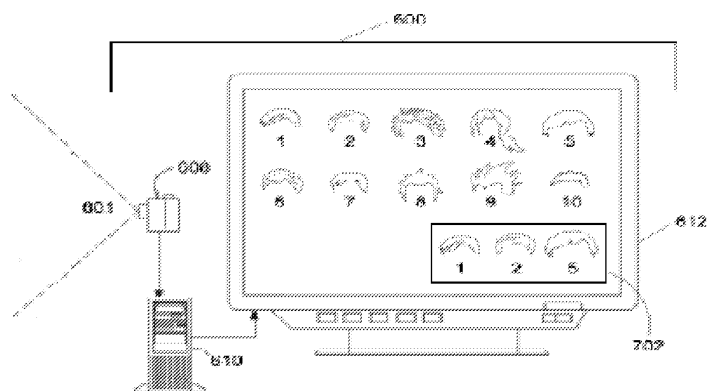
도면6a



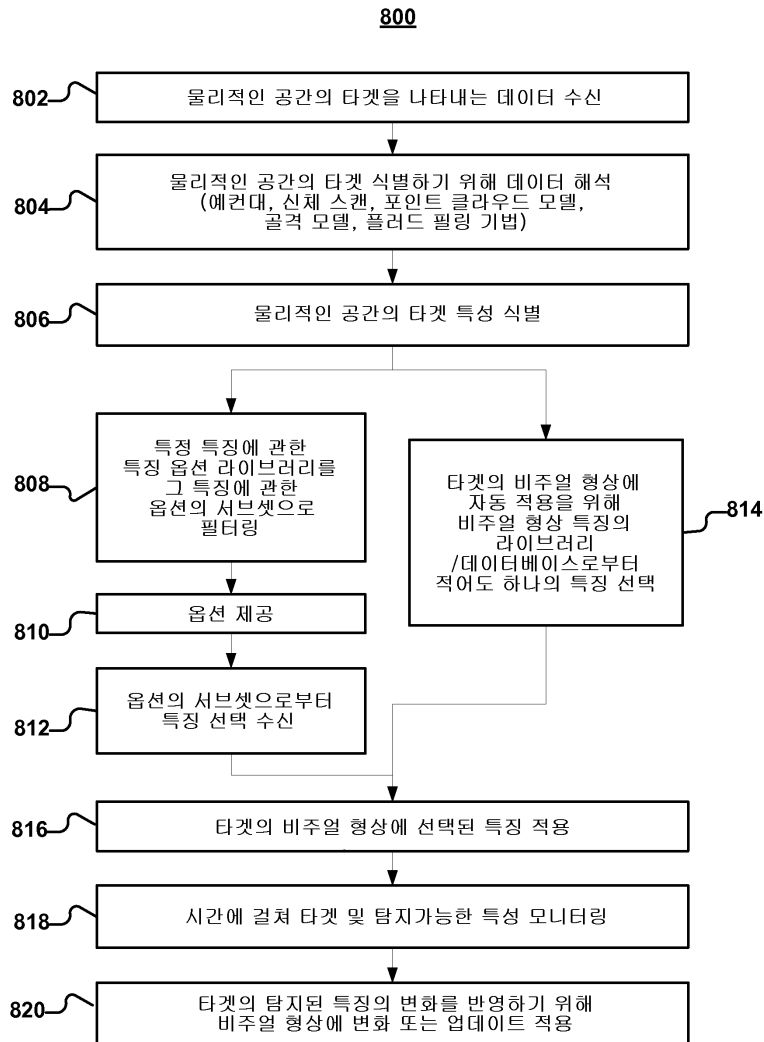
도면6b



도면7



도면8



도면9

