



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2017년02월08일  
(11) 등록번호 10-1704848  
(24) 등록일자 2017년02월02일

(51) 국제특허분류(Int. Cl.)  
A63F 13/213 (2014.01) A63F 13/23 (2014.01)  
A63F 13/42 (2014.01) G06F 3/01 (2006.01)  
G06T 13/80 (2011.01)  
(21) 출원번호 10-2012-7000471  
(22) 출원일자(국제) 2010년07월06일  
심사청구일자 2015년06월05일  
(85) 번역문제출일자 2012년01월06일  
(65) 공개번호 10-2012-0049218  
(43) 공개일자 2012년05월16일  
(86) 국제출원번호 PCT/US2010/041097  
(87) 국제공개번호 WO 2011/005784  
국제공개일자 2011년01월13일  
(30) 우선권주장  
12/500,251 2009년07월09일 미국(US)  
(56) 선행기술조사문헌  
KR1020070021003 A\*  
KR1020050033918 A\*  
KR100884467 B1\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
마이크로소프트 테크놀로지 라이선싱, 엘엘씨  
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원  
마이크로소프트 웨이  
(72) 발명자  
페레즈 캐스린 스톤  
미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
소프트 웨이 엘씨에이 - 인터내셔널 페이턴즈 마  
이크로소프트 코포레이션  
킴맨 알렉스  
미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
소프트 웨이 엘씨에이 - 인터내셔널 페이턴즈 마  
이크로소프트 코포레이션  
(뒷면에 계속)  
(74) 대리인  
김태홍

전체 청구항 수 : 총 19 항

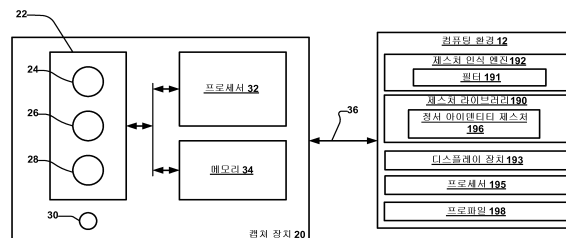
심사관 : 임지환

(54) 발명의 명칭 플레이어의 표현에 기반하는 비주얼 표현의 표현 방법

(57) 요약

시스템은 얼굴 인식 및 제스처/자세 인식 기법을 사용하여, 사용자의 비주얼 표현을 통해 사용자의 감정 및 태도를 자연스럽게 전달할 수 있다. 본 기법은 탐지가능한 특성들에 근거하여 사용자의 비주얼 표현을 커스터마이징하고, 탐지된 특성들로부터 사용자의 기질을 추론하고, 그 기질을 나타내는 속성들을 비주얼 표현에 실시간으로 적용하는 단계들을 포함한다. 또한, 본 기법은 물리적 공간에서 사용자의 특성 변화를 처리하고, 실시간으로 비주얼 표현을 업데이트하는 단계를 포함한다. 예를 들어, 시스템은 사용자의 얼굴 표정 및 신체 움직임을 트래킹하여 기질을 식별하고, 그 기질을 나타내는 속성을 비주얼 표현에 적용한다. 따라서, 아바타 또는 가상 캐릭터와 같은 사용자의 비주얼 표현은 사용자의 표정 및 기분을 실시간으로 반영할 수 있다.

대표도



(72) 발명자

**버튼 니콜라스 디**

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

**월슨 앤드류**

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

---

## 명세서

### 청구범위

#### 청구항 1

시스템 내에서 수행되는, 사용자의 기질(temperament)을 나타내는 속성들을 비주얼 표현(visual representation)에 적용하는 방법에 있어서,

상기 시스템은 프로세서와, 상기 시스템이 동작 중일 때 상기 프로세서에 통신 가능하게 연결된 메모리를 포함하고, 상기 방법은 상기 프로세서에 의해 수행되며, 상기 방법은

사용자의 상기 비주얼 표현을 렌더링하는 단계;

물리적 공간 내의 상기 사용자를 나타내는 제1 데이터를 수신하는 단계;

상기 사용자의 기질을 추론하기 위해 상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 데이터를 분석하는 단계로서,

상기 사용자의 특성, 상기 사용자의 물리적 특징, 상기 사용자의 행동, 상기 사용자의 말투, 상기 사용자의 목소리, 제스처, 또는 이력 데이터를 분석하는 단계; 및

상기 사용자의 상기 제1 데이터를 상기 메모리에 저장된 테이블 - 상기 테이블은 특성들을 특정 기질에 연관시킴 - 과 비교하는 단계

를 포함하는, 상기 제1 데이터를 분석하는 단계;

상기 사용자의 기질에 기초해서 디스플레이할 상기 비주얼 표현의 사전 제작된(pre-canned) 애니메이션을 결정하는 단계;

디스플레이 장치 상에 상기 비주얼 표현을 위한 상기 사전 제작된 애니메이션을 디스플레이하는 단계;

상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 데이터가 수신되었을 때 실행되는 애플리케이션의 상태를 결정하는 단계;

상기 물리적 공간 내의 상기 사용자를 나타내는 제2 데이터를 수신하는 단계;

상기 실행되는 애플리케이션이 상기 상태로 복귀했다고 결정하는 단계;

상기 애플리케이션이 상기 상태로 복귀했기 때문에, 상기 제2 데이터와는 독립적으로 상기 사전 제작된 애니메이션을 디스플레이할 것을 결정하는 단계; 및

디스플레이 장치 상에 상기 사전 제작된 애니메이션을 디스플레이하는 단계를

포함하는, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 2

제1항에 있어서, 상기 비주얼 표현은 아바타, 캐릭터, 커서, 또는 저장 모델(stock model) 중 적어도 하나인 것인, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 3

삭제

#### 청구항 4

제1항에 있어서, 상기 물리적 공간 내의 상기 사용자를 나타내는 제1 데이터는 상기 물리적 공간 내의 사용자의 특성들 중 적어도 하나를 나타내는 것인, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 5

제1항에 있어서, 상기 디스플레이할 상기 비주얼 표현의 사전 제작된 애니메이션을 결정하는 단계는, 사용자의 특성들 중 적어도 하나를 상기 비주얼 표현에 적용하는 단계를 포함하는, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 6

제1항에 있어서, 상기 물리적 공간 내의 상기 사용자를 나타내는 제1 데이터를 수신하는 단계는 캡처 장치(capture device)를 통해 데이터를 수신하는 단계를 포함하는 것인, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 7

제1항에 있어서, 상기 사용자의 기질은 부정적인, 긍정적인, 양면성을 가진(ambivalent), 지루한, 행복한, 슬픈, 좌절스러운, 흥분한 또는 화가 난 것 중 적어도 하나를 포함하는 것인, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 8

제1항에 있어서, 상기 애플리케이션이 상기 상태로 복귀하지 않은 경우 상기 방법은, 상기 제2 데이터와 상기 제1 데이터 간의 차이에 기초해서 상기 사용자의 기질의 변화를 결정하는 단계; 상기 사용자의 기질의 변화에 기초해서 상기 비주얼 표현을 위한 제2 사전 제작된 애니메이션을 결정하는 단계; 및 디스플레이 장치 상에 상기 비주얼 표현을 위한 상기 제2 사전 제작된 애니메이션을 디스플레이하는 단계를 더 포함하는, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 9

제1항에 있어서, 상기 디스플레이할 상기 비주얼 표현의 사전 제작된 애니메이션을 결정하는 단계는, 상기 사용자의 기질에 대응하는 복수의 사전 제작된 애니메이션들로부터 상기 사전 제작된 애니메이션을 선택하는 단계를 포함하는, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

#### 청구항 10

사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하기 위한 시스템에 있어서,  
제1 프로세서; 및  
상기 시스템이 동작 중일 때 상기 프로세서에 통신 가능하게 연결된 메모리를 포함하고,  
상기 메모리는 상기 프로세서 상에서 실행될 때 상기 시스템으로 하여금 적어도,  
상기 사용자의 비주얼 표현을 렌더링하고;  
물리적 공간 내의 상기 사용자를 나타내는 제1 데이터를 수신하고;  
상기 사용자의 기질을 추론하기 위해 상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 데이터를 분석하고 - 상기 분석은 상기 사용자의 특성, 상기 사용자의 물리적 특징, 상기 사용자의 행동, 상기 사용자의 말투, 상기 사용자의 목소리, 제스처, 또는 이력 데이터를 분석하고, 상기 사용자의 상기 제1 데이터를, 상기 메모리에 저장되고 특성들을 특정 기질에 연관시키는 테이블과 비교하는 것을 포함함 - ;  
상기 사용자의 기질에 기초해서 디스플레이할 상기 비주얼 표현의 사전 제작된 애니메이션을 결정하고;  
디스플레이 장치 상에 상기 비주얼 표현을 위한 상기 사전 제작된 애니메이션을 디스플레이하고;  
상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 데이터가 수신되었을 때 실행되는 애플리케이션의 상태를 결정하고;  
상기 물리적 공간 내의 상기 사용자를 나타내는 제2 데이터를 수신하고;

상기 실행되는 애플리케이션이 상기 상태로 복귀했다고 결정하고;

상기 애플리케이션이 상기 상태로 복귀했기 때문에, 상기 제2 데이터와는 독립적으로 상기 사전 제작된 애니메이션을 디스플레이할 것을 결정하며;

디스플레이 장치 상에 상기 사전 제작된 애니메이션을 디스플레이하게 하는 프로세서-실행가능 명령어들을 저장하는 것인, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하기 위한 시스템.

#### 청구항 11

제10항에 있어서, 상기 명령어들은, 상기 프로세서 상에서 실행될 때, 상기 시스템으로 하여금 적어도 디스플레이 장치 상에서 상기 비주얼 표현을 위한 상기 사전 제작된 애니메이션을 디스플레이하게 하고, 또한 상기 시스템으로 하여금 적어도,

상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 및 제2 데이터의 수신에 대해 실시간으로 디스플레이 장치 상에서 상기 비주얼 표현을 위한 상기 사전 제작된 애니메이션을 디스플레이하게 하는 것인, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하기 위한 시스템.

#### 청구항 12

삭제

#### 청구항 13

컴퓨터-실행가능 명령어들을 저장한, 기질을 나타내는 속성들을 비주얼 표현에 적용하기 위한 유형의(tangible) 컴퓨터-판독가능 저장 장치에 있어서,

상기 명령어들은 컴퓨터 상에서 실행될 때, 상기 컴퓨터로 하여금,

사용자의 상기 비주얼 표현을 렌더링하는 것;

물리적 공간 내의 상기 사용자를 나타내는 제1 데이터를 수신하는 것;

상기 사용자의 기질을 추론하기 위해 상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 데이터를 분석하는 것으로서,

상기 사용자의 특성, 상기 사용자의 물리적 특징, 상기 사용자의 행동, 상기 사용자의 말투, 상기 사용자의 목소리, 제스처, 또는 이력 데이터를 분석하는 것; 및

상기 사용자의 상기 제1 데이터를 메모리에 저장된 테이블 - 상기 테이블은 특성들을 특정 기질에 연관시킴 - 과 비교하는 것을 포함하는, 상기 제1 데이터를 분석하는 것;

상기 사용자의 기질에 기초해서 디스플레이할 상기 비주얼 표현의 사전 제작된 애니메이션을 결정하는 것;

디스플레이 장치 상에 상기 비주얼 표현을 위한 상기 사전 제작된 애니메이션을 디스플레이하는 것;

상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 데이터가 수신되었을 때 실행되는 애플리케이션의 상태를 결정하는 것;

상기 물리적 공간 내의 상기 사용자를 나타내는 제2 데이터를 수신하는 것;

상기 실행되는 애플리케이션이 상기 상태로 복귀했다고 결정하는 것;

상기 애플리케이션이 상기 상태로 복귀했기 때문에, 상기 제2 데이터와는 독립적으로 상기 사전 제작된 애니메이션을 디스플레이할 것을 결정하는 것; 및

디스플레이 장치 상에 상기 사전 제작된 애니메이션을 디스플레이하는 것을

포함하는 동작들을 수행하게 하는 것인, 컴퓨터-판독가능 저장 장치.

#### 청구항 14

제13항에 있어서, 디스플레이 장치 상에 상기 비주얼 표현을 위해 상기 사전 제작된 애니메이션을 디스플레이하는 것은 상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 및 제2 데이터의 수신에 대해 실시간으로 수

행되는 것인, 컴퓨터-판독가능 저장 장치.

#### 청구항 15

제13항에 있어서, 상기 비주얼 표현은 아바타, 캐릭터, 커서, 또는 저장 모델 중 적어도 하나인 것인, 컴퓨터-판독가능 저장 장치.

#### 청구항 16

제13항에 있어서, 상기 컴퓨터 상에서 실행될 때, 상기 컴퓨터로 하여금, 사용자의 특성들 중 적어도 하나의 특성을 상기 비주얼 표현에 적용하는 것을 포함하는 동작들을 수행하게 하는 컴퓨터-실행가능 명령어들을 또한 저장하고, 상기 제1 및 제2 데이터는 상기 사용자의 특성들을 나타내는 데이터를 포함하는 것인, 컴퓨터-판독가능 저장 장치.

#### 청구항 17

제13항에 있어서, 상기 물리적 공간 내의 상기 사용자를 나타내는 데이터를 수신하는 것은 캡처 장치를 통해 데이터를 수신하는 것을 포함하는 것인, 컴퓨터-판독가능 저장 장치.

#### 청구항 18

제13항에 있어서, 상기 사용자의 기질은 부정적인, 긍정적인, 양면성을 가진, 지루한, 행복한, 슬픈, 좌절스러운, 흥분한 또는 화가 난 것 중 적어도 하나를 포함하는 것인, 컴퓨터-판독가능 저장 장치.

#### 청구항 19

제13항에 있어서, 상기 컴퓨터 상에서 실행될 때, 상기 컴퓨터로 하여금,

기질 아이덴티티 제스처를 나타내는 필터 - 상기 필터는 상기 기질 아이덴티티 제스처에 대한 기본 정보를 포함함 - 를 제공하는 것; 및

상기 필터를 상기 제1 및 제2 데이터에 적용하고, 상기 기질 아이덴티티 제스처에 대한 상기 기본 정보로부터 출력을 결정하는 것을

포함하는 동작들을 수행하게 하는 컴퓨터-실행가능 명령어들을 또한 저장하며,

상기 기질 아이덴티티 제스처는 기질을 나타내는 속성을 상기 사용자의 비주얼 표현에 적용하는 제어에 매핑하는 제스처인 것인, 컴퓨터-판독가능 저장 장치.

#### 청구항 20

제13항에 있어서,

상기 사용자의 기질에 대응하는 복수의 사전 제작된 애니메이션들로부터 상기 사전 제작된 애니메이션을 선택하는 것을 더 포함하는, 컴퓨터-판독가능 저장 장치.

#### 청구항 21

제1항에 있어서, 상기 사용자의 기질에 기초해서 디스플레이할 상기 비주얼 표현의 사전 제작된 애니메이션을 결정하는 단계는,

상기 사용자의 기질에 기초해서 디스플레이할 상기 비주얼 표현의 사전 제작된 애니메이션을 결정하는 단계를 포함하고, 상기 사전 제작된 애니메이션은 상기 물리적 공간 내의 상기 사용자를 나타내는 상기 제1 데이터에서 캡처된 상기 사용자의 모션(motion) 또는 포즈(pose)와는 상이한 것인, 사용자의 기질을 나타내는 속성들을 비주얼 표현에 적용하는 방법.

### 발명의 설명

### 배경 기술

보통, 다양한 애플리케이션에서 사용자가, 예컨대 리모콘의 버튼을 선택하거나 특정 방식으로 컨트롤러를 움직

[0001]

여서 제어하는 특정 행위를 통해, 사용자에게 대응하는 비주얼 표현을 디스플레이한다. 이러한 비주얼 표현은 아바타, 상상의 캐릭터, 만화 이미지 또는 동물, 커서, 손 등의 형태일 수 있다. 비주얼 표현은 컴퓨터 게임, 비디오 게임, 채팅, 포럼, 커뮤니티, 인스턴트 메시징 서비스 등의 다양한 애플리케이션에서 일반적으로 2D(two-dimensional) 또는 3D(three-dimensional) 모델의 형태를 갖는, 사용자에게 대응하는 컴퓨터 표현이다. 컴퓨터 게임, 멀티미디어 애플리케이션, 오피스 애플리케이션을 비롯한 많은 컴퓨팅 애플리케이션에서는, 사용자의 아바타로써 애플리케이션에서 사용되도록 선택되는 일련의 사전 정의된 애니메이션화된 캐릭터들을 제공한다. 일부 시스템에서는 사용자의 사진을 찍는 기능을 가진 카메라를 포함하고, 그 데이터 프레임의 특징들을 식별할 수 있다. 그러나 이러한 시스템은 사용자 특징 캡처, 이미지 처리, 비실시간 처리 환경에서의 캐릭터에 적용을 필요로 하므로, 적용된 특징들은 보통 사용자의 단일 스냅샷을 기반으로 하게 되어 낮은 정확도(fidelity)를 갖게 된다.

## 발명의 내용

### 과제의 해결 수단

- [0002] 사용자의 탐지된 특징들을 기반으로 하여 사용자의 비주얼 표현을 커스터마이징하고, 이 특징들을 실시간으로 비주얼 표현에 적용하는 것이 바람직하다. 또한 시스템은 물리적 공간에서의 사용자 특징에 대한 변화를 처리하여, 실시간으로 비주얼 표현을 업데이트할 수 있는 것이 바람직하다. 이러한 특징들 중에서, 시스템은 사용자의 기질(temperament)을 식별하고, 이러한 기질을 나타내는 속성(attributes)을 사용자의 비주얼 표현에 적용하는 것이 바람직하다.
- [0003] 본원에는 사용자의 기질을 실시간으로 반영할 수 있는, 아바타 또는 상상 캐릭터와 같은 사용자의 비주얼 표현을 제공하는 기법이 개시된다. 시스템은 얼굴 인식 및 제스처/자세 인식 기법을 사용하여 사용자의 기질을 추론할 수 있다. 시스템은 사용자의 기질의 속성을 사용자의 비주얼 표현에 적용함으로써 사용자의 감정 및 태도를 자연스럽게 전달할 수 있다. 또한, 본원에는 물리적 공간의 사용자를 시간에 걸쳐 트래킹하여 실시간으로 비주얼 표현에 변화 또는 업데이트를 적용하는 기법이 개시된다. 예를 들어, 시스템은 사용자의 얼굴 표정 및 신체 움직임의 트래킹을 기질을 식별한 다음에, 이러한 기질을 나타내는 속성을 비주얼 표현에 적용할 수 있다. 시스템은 비주얼 표현에 적용되기 위한 사용자의 기질을 평가하기 위해, 탐지가능한 모든 특징들을 사용할 수 있다.
- [0004] 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

- [0005] 본 명세서에 따라 비주얼 표현을 변화시키는 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 5(a)는 깊이 이미지(depth image)로부터 생성된 사용자의 골격 매핑을 도시한다.
- 도 5(b)는 도 2에 도시된 제스처 인식 아키텍처의 세부 사항을 도시한다.
- 도 6은 타겟 인식, 분석 및 트래킹 시스템의 일례 및 물리적 공간에서의 사용자 및 사용자의 비주얼 표현 디스플레이의 일 실시예를 도시한다.

도 7은 사용자의 기질을 나타내는 속성을 비주얼 표현에 적용하는 방법에 대한 순서도의 일례를 도시한다.

도 8은 사용자의 기질을 추론하기 위한 룩업 테이블의 일례를 도시한다.

도 9는 타겟 인식, 분석 및 트래킹 시스템의 다른 예시 및 물리적 공간에서의 사용자의 실시예들 및 사용자의 비주얼 표현 디스플레이의 실시예들을 도시한다.

### 발명을 실시하기 위한 구체적인 내용

- [0006] 아바타와 같이 사용자의 기질을 반영가능한 사용자의 비주얼 표현을 제공하는 기법이 본원에 개시된다. 사용자의 비주얼 표현은 캐릭터, 애니메이션, 아바타, 스크린 상의 커서(cursor), 손 또는 물리적 공간에서의 사용자에게 대응하는 기타의 비주얼 표현의 형태일 수 있다. 시스템은 안면 인식 및 제스처/신체 자세 인식 기법을 사용하여, 사용자의 비주얼 표현을 통해 사용자의 감정 및 태도를 자연스럽게 전달할 수 있다. 예를 들어, 캡처 장치는 사용자의 얼굴을 식별하고, 감정(emotions), 표정(expressions) 및 기분(moods) 등 그 식별된 얼굴에 기반하여 사용자의 비주얼 표현을 커스터마이즈할 수 있다. 일 실시예에서, 캡처 장치에서 캡처된 이미지 데이터를 기반으로 하여 사람의 골격 또는 메쉬(mesh) 모델의 양태들을 생성 및 사용하고, 신체 인식 기법을 사용하여 사용자의 기질을 결정한다.
- [0007] 또한 실시간으로 비주얼 표현을 디스플레이하고, 사용자의 기질을 나타내는 속성을 실시간으로 비주얼 표현에 적용하는 기법이 개시된다. 시스템은 시간이 흐름에 따라 물리적 공간의 사용자를 트래킹하고, 비주얼 표현에 변화 또는 업데이트를 실시간으로 적용한다. 시스템은 사용자의 특징, 제스처, 애플리케이션 상태 등을 비롯한 탐지가능한 특징들을 트래킹하여, 사용자의 기질을 추론할 수 있다. 얼굴 표정 및 신체 움직임과 같은 사용자의 특징을 사용하여 기질을 추론하고, 비주얼 표현이 사용자의 기질을 반영하도록 그 기질의 속성이 비주얼 표현에 적용될 수 있다. 예를 들어, 캡처 장치는 사용자의 행동 및 매너리즘(mannerisms), 감정, 말투(speech patterns), 이력 데이터(history data) 등을 식별하여 사용자의 기질을 결정하고, 이들을 사용자의 비주얼 표현에 적용할 수 있다. 시스템은 임의의 탐지가능한 특징들을 사용하여, 비주얼 표현에 적용되기 위한 사용자의 기질을 평가할 수 있다.
- [0008] 물리적 공간의 타겟 또는 사물의 모델을 생성하기 위해, 캡처 장치는 장면의 깊이 이미지를 캡처하고, 그 장면의 타겟 또는 사물들을 스캔할 수 있다. 타겟은 물리적 공간에서 사용자와 같은 인간 타겟일 수 있다. 따라서, 본원에서 사용된 것처럼, 타겟과 사용자는 서로 교체 사용이 가능함을 이해할 것이다. 일 실시예에서, 캡처 장치는 장면의 하나 이상의 타겟 또는 사물들이 사용자와 같은 인간 타겟에 대응하는지를 결정할 수 있다. 장면의 타겟 또는 사물이 인간 타겟에 대응하는지 여부를 결정하기 위해, 각 타겟이 플러드 필링되어(flood filled) 인체 모델의 패턴과 비교된다. 인체 모델에 부합하는 각 타겟 또는 사물이 스캔되어, 그에 관련된 골격 모델을 생성한다. 예를 들어, 사람이라고 식별된 타겟이 스캔되어, 그에 관련된 골격 모델을 생성한다. 다음에, 이 골격 모델은, 골격 모델을 트래킹하여 그 골격 모델에 관련된 비주얼 표현을 렌더링하는(rendering) 컴퓨팅 환경으로 제공된다. 컴퓨팅 환경은, 인식되고 골격 모델로 매핑된 사용자의 제스처들을 기반으로 하여 컴퓨팅 환경에서 실행되는 애플리케이션에서 어떤 제어를 실행할지를 결정할 수 있다. 따라서, 사용자 피드백이 스크린 상에 아바타를 통해서 디스플레이될 수 있고, 사용자는 물리적 공간에서 제스처를 취함으로써 아바타의 모션(motion)을 제어할 수 있다.
- [0009] 비주얼 표현의 모션을 물리적 공간의 사용자의 모션에 일치시킴으로써, 비주얼 표현의 모션을 제어할 수 있다. 예를 들어, 타겟은 물리적 공간에서 모션을 취하거나 제스처를 취하는 사용자일 수 있다. 타겟의 비주얼 표현은 스크린에 디스플레이되는 아바타이며, 아바타의 모션이 사용자의 모션에 대응할 수 있다. 물리적 공간에서의 모션은 가상 공간 및/또는 게임 공간 등의 시스템 또는 애플리케이션 공간에서의 제어로 해석될 수 있다. 예를 들어, 사용자의 모션들은 트래킹, 모델링 및 디스플레이되고, 사용자의 제스처들은 운영 체제 또는 실행 중인 애플리케이션의 특정 양태들을 제어할 수 있다. 사용자의 제스처들은 기질을 나타내는 속성을 비주얼 표현에 적용하는 시스템 또는 애플리케이션 공간에서의 제어로 해석될 수 있다.
- [0010] 캡처된 모션은 카메라와 같은 캡처 장치로 캡처된 물리적 공간에서의 임의의 모션일 수 있다. 이 캡처된 모션은 사용자 또는 사물과 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.
- [0011] 본원에서 설명된, 사용자의 태도 및 감정을 전달하는 안면 및 신체 인식 시스템, 방법 및 컴포넌트들은 게임 콘솔과 같은 멀티미디어 콘솔 또는, 예를 들어 그러나 이에 제한되지 않는, 위성 수신기, 셋톱 박스, 아케이드 게



임(arcade games), PC(personal computers), 휴대 전화, PDA(personal digital assistants) 및 기타 핸드헬드 장치들을 포함하는 타겟의 비주얼 표현을 디스플레이하기 위한 기타 임의의 컴퓨팅 장치에서 구현될 수 있다.

[0012] 도 1은 사용자의 특징을 비주얼 표현에 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 기법(body posture recognition techniques)을 사용하여 인간 타겟(18)의 기질을 식별할 수 있다. 예를 들어, 사용자가 구부정하게, 자신의 가슴 위에 팔짱을 끼고, 자신의 머리를 힘없이 옆으로 움직인다면, 시스템(10)은 사용자(18)의 신체 부위와, 이들이 어떻게 움직이는지를 식별할 수 있다. 시스템(10)은 이 모션들을 감정, 기분, 태도, 표정 등의 목록과 비교하여 사용자의 기질을 해석할 수 있다.

[0013] 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.

[0014] 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 사용자들이 행하는 제스처를 캡처, 분석 및 트래킹해서 애플리케이션 내의 여러 제어나 액션을 실행할 수 있는 카메라일 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다.

[0015] 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 오디오비주얼 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.

[0016] 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. 시스템은 사용자의 자세, 얼굴 표정, 목소리 표현 및 어조(tone), 주시 시선(directed gazes) 등을 비교하여 사용자의 기질 또는 태도를 결정하고, 그 기질 또는 태도의 특성을 아바타에 적용할 수 있다.

[0017] 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 비주얼 표현(40)에 매핑시켜, 물리적 공간에서의 사용자의 모션들이 아바타(40)에 의해 실행되도록 할 수 있다. 사용자의 모션들은 애플리케이션에서 제어에 해당되는 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.

[0018] 컴퓨팅 환경(12)은 오디오비주얼 장치(16)를 사용하여, 사용자(18)가 그 자신의 움직임으로 제어할 수 있는 플레이어 아바타(40)의 비주얼 표현을 제공할 수도 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 특징을 가질 수 있고, 또는 시스템(10)은 사용자의 모션에 일치하는 비주얼 표현에 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 특징을 바꿀 수 있다. 또한, 컴퓨팅 환경(12)은 오디오비주얼 장치(16)를 사용하여 사용자(18)에게 복싱 상대(38)의 비주얼 표현을 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 표현은 사용자(18)와 같은 물리적 공간에서의 제 2 사용자 또는 제 2 물리적 공간에서의 네트워크 연결된 사용자 등 다른 사용자를 나타낼 수 있다.

[0019] 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.

[0020] 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 쥔 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.

[0021] 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, 친구와의 통신 등을 할 수 있다. 플레이어는 움직임을 통해 기질을 나타내는 속성을 사용자의 비주얼 표현에 적용할 수 있다. 운영 체제 및/또는 애플리케이션의 제어가능한 사실상 모든 양태가 사용자(18)와 같은 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.

[0022] 사용자 특징을 비주얼 표현에 적용하거나 사용자의 특정 감정 또는 태도를 탐지하는 것은, 사용자 제스처에 의해 제어되거나 사용자 제스처로부터 인식될 수 있는 운영 체제 및/또는 애플리케이션의 양태이다. 예를 들어, 사용자가 자신의 가슴에 팔짱을 끼는 제스처는 좌절감으로 인식되는 제스처이다. 얼굴을 찌푸리는 것과 같은 사용자의 표정과 함께 사용자가 좌절하고 있음을 나타내는 제스처를 시스템이 인식함으로써 좌절된 기질을 반영하는 비주얼 표현으로 나타난다.

[0023] 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 탭을 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.

[0024] 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 임의의 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.

[0025] 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 캡처된 장면의 2-D(two-dimensional) 화소 영역을 포함할 수 있고, 2-D 화소 영역의 각 화소는 카메라로부터 캡처된 장면에서 사물의 길이 또는 거리 등의 깊이 값을 센티미터, 밀리미터 등으로 나타낼 수 있다.

[0026] 도 2에 도시된 바와 같이, 일 실시예에 따르면, 이미지 카메라 컴포넌트(22)는 장면의 깊이 이미지를 캡처하기 위해 사용되는 적외선 컴포넌트(IR light component, 24), 3-D(three-dimensional) 카메라(26) 및 RGB 카메라(28)를 포함할 수 있다. 예를 들면, 비행 시간 분석법에서는, 캡처 장치(20)의 적외선 컴포넌트(24)가 장면으로 적외선을 내보내고, 센서들(도시되지 않음)을 사용하여 3-D 카메라(26) 및/또는 RGB 카메라(28)를 사용한 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 시간을 측정하고, 이를 사용하여 캡처 장치(20)로부터 장면의 타겟 또는 사물들의 특정 지점까지의 물리적인 거리를 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.

[0027] 또 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 사물의 특정 지점까지의 물리적인 거리를 간접적으로 결정할 수 있다.

[0028] 또 다른 실시예에 따르면, 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 컴포넌트(24)를 통해 장면에 투사될 수 있다. 그 장면의 하나 이상의 타겟 또는 사물의 표면에 부딪치면, 패턴이 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.

[0029] 또 다른 실시예에 따르면, 캡처 장치(20)는, 깊이 정보의 생성을 위해 분석되는 비주얼 스테레오 데이터를 얻기 위해 각기 다른 각도에서 장면을 볼 수 있는 물리적으로 분리된 둘 이상의 카메라들을 포함할 수 있다.

[0030] 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.

[0031] 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 표현 또는 모델로 변환하는 인스트럭션을 포함하는 인스트럭션들 또는 기타 임의의 적절한 인스트럭션을 실행하는 표준화된 프로세서, 특화된 프로세서, 마이크로 프로세서를 포함할 수 있다.

[0032] 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.

[0033] 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.

[0034] 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(190)를 포함할 수 있다.

[0035] 도 2에 도시된 바대로, 컴퓨팅 환경(12)은 제스처 라이브러리(190) 및 제스처 인식 엔진(192)을 포함할 수 있다. 제스처 인식 엔진(192)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 움직임은 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.

[0036] 제스처 인식 엔진은 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 내력에서 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.

[0037] 제스처는 기질 아이덴티티 제스처(temperament identity gesture)로 인식될 수 있다. 일 실시예에서, 물리적 공간에서의 모션은, 특정 기질의 속성을 타겟의 비주얼 표현에 적용하는 요청으로 인식되는 제스처를 나타낼 수 있다. 복수의 제스처들 각각이 특정 기질 아이덴티티 제스처를 나타낸다. 따라서, 사용자는 기질 아이덴티티 제스처로 인식되는 제스처를 물리적 공간에서 취함으로써 비주얼 표현의 유형을 제어할 수 있다. 예를 들어, 전술한 바와 같이, 사용자의 모션이 도 2의 제스처 필터(191) 등의 제스처 필터와 비교된다. 제스처 필터(191)는 제스처 라이브러리(190)의 기질 아이덴티티 제스처들(196)에 대한 정보를 포함할 수 있다.

[0038] 복수의 기질 아이덴티티 제스처들 각각은 스크린의 비주얼 표현에 적용될 속성을 갖는 기질을 나타낸다. 예를 들어, "흥분된" 아이덴티티 제스처는 공중으로 사용자의 팔을 들어올리고 필쩍 뛰는 것을 포함하는 사용자 모션의 아이덴티티로부터 인식될 수 있다. 그 결과, 사용자의 모션에 직접적으로 매핑되는 속성 및/또는 그 사용자의 모션뿐만 아니라 애니메이션도 사용자의 비주얼 표현에 적용될 것이다.

[0039] 카메라(26, 28) 및 장치(20)에 의해 캡처된 골격 모델 및 그에 관련된 움직임 형식의 데이터가 제스처 라이브러리(190)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.

[0040] 컴퓨팅 환경(12)은 깊이 이미지를 처리하여, 방에 있는 사용자(18) 또는 사물과 같이 장면에 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 표현을 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 기법들로는 이미지를 사람의 신체 모델 표현으로 전환하고, 이미지를 사람의 메쉬 모델 표현으로 전환하는 것이 있다.

[0041] 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.

[0042] 컴퓨팅 환경(12)은 제스처 라이브러리(190)를 사용하여, 골격 모델의 움직임을 해석하고 그 움직임을 기초로 하여 애플리케이션을 제어한다. 컴퓨팅 환경(12)은 디스플레이 장치(193) 등의 디스플레이에 아바타 또는 포인트의 형태로 사용자의 표현을 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.

[0043] 전술한 바와 같이, 기질의 속성을 타겟의 비주얼 표현에 적용시키는 것이 바람직하다. 예를 들어, 사용자는 자신의 행복함을 보여주기 위해 사용자의 비주얼 표현이 스크린 상에서 춤을 추기를 원한다. 사용자는 특정 기질 아이덴티티 제스처를 취함으로써 이러한 속성의 적용을 시작한다.

[0044] 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 록업 테이블들을 포함할 수 있다. 가상 스크린은 도 1과 관련하여 전술한 컴퓨팅 환경(12)에서 실행되는 애플리케이션과 인터랙션할 수 있다.

[0045] 실시예들에 따르면, 록업 테이블은 사용자 전용 프로파일 정보를 포함할 수 있다. 일 실시예에서, 컴퓨팅 환경(12)과 같은 컴퓨팅 환경에서는 하나 이상의 사용자에게 대한 프로파일 데이터(198)가 록업 테이블에 저장될 수 있다. 저장된 프로파일 데이터(198)는 타겟의 스캔되거나 추정된 신체 크기, 골격 모델, 목소리 샘플 또는 비밀 번호, 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.

[0046] 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.

[0047] 제스처 라이브러리(190), 제스처 인식 엔진(192) 및 프로파일(198)은 하드웨어, 소프트웨어 또는 이들의 조합으로 구현될 수 있다. 예를 들어, 제스처 라이브러리(190) 및 제스처 인식 엔진(192)은 컴퓨팅 환경(12)(또는 도 3의 처리 장치(101)나 도 4의 처리 장치(259))의 프로세서(195)와 같은 프로세서에서 실행되는 소프트웨어로 구현될 수 있다.

[0048] 도 2 및 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(190)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.

[0049] 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.

[0050] 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 의

미를 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.

[0051] 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 앞쪽으로 몸에서 멀어지게 뻗는 것은 펀치 제스처를 나타낸다. 사용자가 운전하는 컨텍스트에서는, 동일한 모션은 "기어 전환" 제스처를 나타낸다. 비주얼 표현에 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.

[0052] 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.

[0053] 도 3은 타겟 인식, 분석 및 트래킹 시스템에서 여러 제스처들을 해석하는 데에 사용되는 컴퓨팅 환경의 일 실시 예를 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.

[0054] 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.

[0055] 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.

[0056] 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

어 드라이브(144)는, 직렬 ATA 버스 또는 기타 고속 접속(예를 들어, IEEE 1394)과 같은 버스를 통해 I/O 제어기(120)에 접속된다.

- [0057] 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.
- [0058] 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.
- [0059] 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.
- [0060] 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 이용 가능한 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.
- [0061] 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.
- [0062] 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.
- [0063] 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 쓰레드(idle thread)가 미사용 사이클들을 쓸 수 있도록 CPU 예약이 일정한 것이 바람직하다.
- [0064] 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.
- [0065] 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 쓰레드인지 게임 애플리케이션 쓰레드인지를 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.
- [0066] 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.

[0067] 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.

[0068] 도 4는 타겟 인식, 분석 및 트래킹 시스템에서의 여러 제스처들의 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 구성 요소들은 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.

[0069] 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.

[0070] 또한, 컴퓨터(241)는 다른 이동식/비이동식, 휘발성/비휘발성 컴퓨터 저장 매체를 포함할 수 있다. 예를 들어, 도 4는 비이동식, 비휘발성 자기 매체에 판독 또는 기록하는 하드 디스크 드라이브(238), 이동식, 비휘발성 자기 디스크(254)에 판독 또는 기록하는 자기 디스크 드라이브(239), 및 CD ROM 또는 다른 광학 매체와 같은 이동식, 비휘발성 광 디스크(253)에 판독 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.

[0071] 앞서 논의되었으며 도 4에 도시된 드라이브 및 그 관련 컴퓨터 저장 매체는 컴퓨터(241)를 위한 컴퓨터 판독가능 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 카피들임을 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.

[0072] 컴퓨터(241)는 원격 컴퓨터(246)와 같은 하나 이상의 원격 컴퓨터들에 대한 논리적인 연결들을 사용하여 네트워크된 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 네트워킹 환경들은 사무실, 기업(enterprise-wide) 컴퓨터 네트워크, 인터넷 및 인터넷에 혼다.

[0073] LAN 네트워킹 환경에서 사용될 때, 컴퓨터(241)는 네트워크 인터페이스 또는 어댑터(237)를 통해서 LAN(245)에 연결된다. WAN 네트워킹 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. 네트워킹 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.

[0074] 컴퓨터 판독가능 저장 매체는 비주얼 표현을 변화시키는 컴퓨터 판독가능 인스트럭션들을 포함할 수 있다. 이 인스트럭션들은, 비주얼 표현을 렌더링하고, 물리적 공간에서 사용자의 기질 아이덴티티 체스처를 나타내는 데이터를 포함하는 장면의 데이터를 수신하며, 사용자의 기질 아이덴티티 체스처를 기초로 하여 비주얼 표현을 변화시키는 인스트럭션들을 포함할 수 있으며, 기질 아이덴티티 체스처는 기질을 나타내는 속성을 사용자의 비주얼 표현에 적용하는 제어에 매핑되는 체스처이다.

[0075] 도 5(a)는 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.

[0076] 사용자는 자신의 신체를 움직여서 체스처를 생성할 수 있다. 체스처는 이미지 데이터로 캡처되어 의미를 위해 파싱되는 사용자의 모션 또는 포즈를 포함한다. 체스처는 동적일 수 있으며, 공 던지기를 흉내내는 것과 같은 모션을 포함한다. 체스처는 자신의 몸통(524) 앞에 팔뚝(504)을 교차하고 있는 등 정적인 포즈일 수 있다. 또한, 체스처는 모형 칼을 흔드는 것처럼 소품을 포함할 수도 있다. 체스처는 양손(502)으로 손뼉을 치는 것과 같이 하나 이상의 신체 부위를 포함하거나, 입술을 오므리는 것과 같은 세밀한 모션을 포함할 수 있다.

[0077] 사용자의 체스처는 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 체스처일 수 있다. 또한, 체스처는 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, 체스처는 비디오 게임, 문서 편집기, 워드 프로세싱, 데이터 관리 등의 매우 다양한 애플리케이션에서 디스플레이되는 사용자 표현에 매핑되는 매우 다양한 모션들을 나타낼 수 있다.

- [0078] 사용자는 물리적 공간에서 제자리에서 걷거나 뛰으로써 걷거나 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 파악할 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.
- [0079] "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.
- [0080] 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.
- [0081] 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.
- [0082] 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 파악한다. 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 포인트(transition point)의 파라미터들을 만족시킬 수 있다.
- [0083] 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 임계값(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.
- [0084] 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.
- [0085] 애플리케이션은 사전 제작된 애니메이션들이 사용되는 포인트들을 식별하기 위해서 다양한 전환 포인트들(transition points)에 관련된 파라미터들의 값을 설정할 수 있다. 전환 포인트들은 다양한 파라미터들, 예컨대 특정 제스처의 식별, 속도, 타겟 또는 사물의 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.
- [0086] 제스처에 대한 또 다른 파라미터로 움직인 거리를 들 수 있다. 사용자의 제스처들이 가상 환경의 비주얼 표현의 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.

[0087] 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.

[0088] 도 5(b)는 도 2의 제스처 인식 엔진(192)의 일 실시예에 대해 추가적인 세부 사항을 제공한다. 도시된 바와 같이, 제스처 인식 엔진(192)은 제스처 또는 제스처들을 결정하기 위해 적어도 하나의 필터(519)를 포함할 수 있다. 필터(519)는 제스처(526, 이하 "제스처"라고 함)를 정의하는 정보를 포함하고, 또한 그 제스처(526)에 대해 적어도 하나의 파라미터(528) 또는 메타데이터를 포함할 수 있다. 예를 들어, 한 손이 몸 뒤로부터 몸 앞을 지나가는 모션을 포함하는 던지기 동작은 사용자의 한 손이 몸 뒤로부터 몸 앞을 지나가는 움직임을 나타내는 정보를 포함하는 제스처(526)로 구현될 수 있으며, 이 움직임은 깊이 카메라에 의해 캡처될 것이다. 이후에, 그 제스처(526)에 대한 파라미터(528)들이 설정된다. 제스처(526)가 던지기 동작이면, 파라미터(528)는 손이 도달해야 하는 임계 속도, 손이 이동해야 하는 (절대적인 또는 전체적으로 사용자의 신체 크기에 따라 상대적인) 거리 및 제스처(526)가 발생했다는 인식 엔진(192)의 신뢰도 평가될 수 있다. 제스처(526)에 대한 이러한 파라미터(528)들은 애플리케이션들 간에, 단일 애플리케이션의 컨텍스트들 간에, 또는 단일 애플리케이션의 단일 컨텍스트 내에서 시간에 따라 달라질 수 있다.

[0089] 필터들은 모듈식(modular)이거나 상호 교체 가능하다. 일 실시예에서, 필터는 각각 유형이 있는 다수의 입력 및 출력을 가진다. 이런 상황에서, 제 1 필터는 인식 엔진(192) 구조의 나머지 다른 양태의 변경 없이 제 1 필터와 동일한 수 및 유형의 입출력을 갖는 제 2 필터로 대체될 수 있다. 예를 들어, 드라이빙에 관한 제 1 필터는 골격 데이터를 입력으로 받고, 그 필터에 관련된 제스처(526)가 발생하였다는 신뢰도(confidence) 및 조종 각도(angle of steering)를 출력한다. 제 2 드라이빙 필터가 보다 효율적이고 적은 프로세싱 리소스를 필요로 하기 때문에 제 1 드라이빙 필터를 제 2 드라이빙 필터로 교체하기를 바라는 경우, 제 2 필터가 동일한 입력 및 출력, 즉 골격 데이터 유형의 하나의 입력 및 신뢰도 유형 및 각도 유형의 두 개의 출력을 가진다면 제 1 필터를 제 2 필터로 간단하게 교체할 수 있다.

[0090] 필터가 파라미터(528)를 가질 필요는 없다. 예를 들어, 사용자의 키(user's height)를 출력하는 "사용자 키" 필터는 튜닝가능한 어떤 파라미터들도 포함하지 않을 수 있다. 또 다른 "사용자 키" 필터는, 예컨대, 사용자의 키를 결정 시 사용자의 신발, 헤어스타일, 모자 및 자세의 설명 여부에 관한 튜닝가능한 파라미터들을 포함할 수 있다.

[0091] 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. 필터의 출력에는 특정 제스처가 취해짐에 대한 신뢰도, 제스처 모션이 행해진 속도 및 제스처가 발생한 시각 등이 포함될 수 있다.

[0092] 컨텍스트는 문화적인 컨텍스트일 수 있으며, 환경적인 컨텍스트일 수도 있다. 문화적인 컨텍스트란 시스템을 사용하는 사용자의 문화를 의미한다. 다른 문화들이 유사한 제스처를 사용하여 현저하게 다른 의미를 전할 수 있다. 예를 들면, 다른 사용자에게 "보도록" 또는 "눈을 사용하도록" 말하고 싶은 미국인 사용자는 관자놀이 근처에 검지 손가락을 위치시킨다. 반면, 이탈리아 사용자는 이러한 제스처를 마피아를 언급하는 것으로 해석한다.

[0093] 유사하게, 단일 애플리케이션의 상이한 환경에서 상이한 컨텍스트들이 존재할 수 있다. 자동차 운전을 포함하는 일인칭 슈터 게임을 예를 든다. 사용자가 걷는 경우, 땅을 향해 손가락으로 주먹을 쥐고 그 주먹을 앞쪽으로 몸에서 멀어지게 펼치는 것은 펀치 제스처를 나타낸다. 사용자가 운전하는 컨텍스트에서는, 동일한 모션은 "기어 전환" 제스처를 나타낸다. 또한, 사용자가 자신의 게임을 저장하고, 자신의 캐릭터의 장비들 중에서 선택하고, 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.



- [0094] 제스처 인식 엔진(190)은 제스처 필터(519)에 기능을 제공하는 베이스 인식기 엔진(base recognizer engine, 517)을 가질 수 있다. 일 실시예에서, 인식기 엔진(517)이 구현하는 기능은 인식된 제스처들 및 기타 입력을 트래킹하는 입력-오버-타임 저장소(input-over-time archive), (모델링된 시스템이 미지의 파라미터들을 가진 마르코프 프로세스 - 현 상태에 미래 상태를 결정하는 데에 필요한 임의의 과거 상태 정보가 압축되어 있고, 이러한 목적으로 그 밖의 어떤 과거 상태 정보도 존재해서는 안 됨 - 라고 가정하고, 관찰가능한 데이터로부터 히든 파라미터들이 결정되는) 히든 마르코프 모델(Hidden Markov Mode) 구현, 및 특정한 제스처 인식 사례를 해결하기 위해 요구되는 기능을 포함한다.
- [0095] 필터들(519)은 베이스 인식기 엔진(517) 상에 로딩되어 구현되고, 엔진(517)로부터 모든 필터들(519)에 제공되는 서비스들을 이용할 수 있다. 일 실시예에서, 베이스 인식기 엔진(517)은 수신된 데이터를 처리하여 이것이 임의의 필터(519)의 요구사항에 부합하는 지를 결정한다. 입력 과정과 같이 제공된 이들 서비스는 각 필터(519) 대신 베이스 인식기 엔진(517)에서 한 번 제공되므로, 이러한 서비스는 한 주기 동안 각 필터(519)에서 한 번씩 처리되는 것이 아니라 그 주기 동안 한 번만 처리되면 되므로, 제스처를 결정하는 데 요구되는 프로세싱이 줄어든다.
- [0096] 애플리케이션은 인식 엔진(192)에서 제공하는 필터들(519)을 사용하거나, 또는 베이스 인식기 엔진(517)에 플러그인 된 애플리케이션 자체 필터(519)를 제공할 수 있다. 일 실시예에서, 모든 필터들(519)은 이러한 플러그인 특징을 이용가능하도록 공통의 인터페이스를 가진다. 또한, 모든 필터들(519)이 파라미터들(528)을 이용할 수 있고, 따라서 이후에 설정될 단일 제스처 툴을 사용하여 전체 필터 시스템(519)을 디버깅하고 튜닝할 수 있다.
- [0097] 이들 파라미터(528)는 애플리케이션 또는 애플리케이션의 컨텍스트에서 제스처 툴(521)로 튜닝할 수 있다. 일 실시예에서, 제스처 툴(521)은 복수의 슬라이더(523)를 포함하며, 슬라이더(523)는 각각 파라미터(528) 및 신체(524)의 그림 표현에 대응한다. 파라미터(528)는 대응 슬라이더(523)에 의해 조정되므로, 신체(524)는 이들 파라미터(528)가 있는 제스처로 인식되는 액션들 및 이들 파라미터(528)가 있는 제스처로 인식되지 않는 액션들 모두를 보여줄 것이다. 제스처의 파라미터들(528)의 이와 같은 시각화는 제스처를 디버깅하고 미세 튜닝하기 위한 효과적인 수단을 제공한다.
- [0098] 도 6은 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 5(b)와 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.
- [0099] 본 예시에서, 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 표현의 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.
- [0100] 물리적 공간(601)의 사용자(602)의 비주얼 표현은 애니메이션, 캐릭터, 아바타를 비롯하여 어떤 형태도 취할 수 있다. 예를 들어, 사용자(602)와 같은 타겟의 비주얼 표현은 처음에는 사용자(602)가 원하는 모양 및 크기로 조각할 수 있는 디지털 진흙 덩어리이거나, 또는 디스플레이 장치(612)에 도시된 원숭이(603)와 같은 캐릭터 표현일 수 있다. 비주얼 표현은 사용자(602)의 특징과, 애니메이션 또는 저장 모델(stock model)의 결합일 수 있다. 비주얼 표현은 시스템(600) 또는 애플리케이션에서 제공되는 저장 모델일 수 있다. 예를 들어, 사용자(602)는 게임 애플리케이션에서 제공되는 다양한 저장 모델로부터 선택을 할 수 있다. 예를 들어, 야구 게임 애플리케이션에서, 사용자(602)를 시각적으로 표현하는 옵션들은 유명한 야구 선수의 표현에서부터 태피 조각 또는 코끼리, 커서 또는 손 기호와 같은 상상 캐릭터나 기호 등 임의의 형태를 취할 수 있다. 저장 모델은 시스템에서 탐지되는 사용자의 특징으로 바뀔 수 있다. 비주얼 표현은 프로그램과 함께 포장된 것과 같이 애플리케이션 전용이거나, 또는 애플리케이션에서 또는 시스템 전반에서 이용가능할 수도 있다.
- [0101] 도 6에 도시된 비주얼 표현의 예시는 디스플레이 장치(612)에 도시된 바와 같이 원숭이 캐릭터(603)이다. 이미지 데이터의 추가적인 프레임들이 캡처되고 디스플레이될 수 있으며, 도 6에 도시된 프레임은 예를 들기 위해 선택되었다. 이미지 데이터의 프레임들이 캡처되어 디스플레이되는 비율에 의해 비주얼 표현이 디스플레이되는 모션의 연속성 레벨이 결정될 수 있다. 또한, 다른 또는 추가적인 비주얼 표현이 물리적 공간(601)의 또 다른

타겟, 예를 들어, 또 다른 사용자나 비인간 객체에 대응할 수 있으며, 비주얼 표현이 부분적으로 또는 전체적으로 가상 객체일 수도 있음을 알 것이다.

[0102] 시스템(600)은 깊이 정보, 이미지 정보, RGB 데이터 등 물리적 공간(601)에 관한 정보를 캡처할 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. 예를 들어, 카메라(608)는 이미지 데이터를 처리하고, 이를 사용하여 타겟의 형태, 색상 및 크기를 결정할 수 있다. 인간 패턴에 일치하는 각 타겟이나 사물이 스캔되어 관련 골격 모델, 플러드 모델, 메쉬 인간 모델 등의 모델을 생성할 수 있다. 예를 들어, 전술한 바와 같이, 깊이 정보를 사용하여 도 5(a)에 도시된 것처럼 사용자의 골격 모델을 생성할 수 있으며, 시스템은 머리 및 팔다리 등의 사용자의 신체 부위를 식별한다. 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.

[0103] 시스템(600)은 캡처된 데이터를 분석하여 이를 골격 모델로 변환시켜 사용자의 팔다리의 움직임을 트래킹할 수 있다. 이후에, 시스템(600)은 골격 모델을 트래킹하고, 각 신체 부위의 움직임을 비주얼 표현의 각 부분에 매핑시킬 수 있다. 예를 들어, 사용자(602)가 자신의 팔을 흔들면, 시스템이 이 모션을 캡처하고 이를 가상 원숭이(603)의 팔에 적용하여 가상 원숭이도 팔을 흔들게 할 수 있다. 또한, 시스템(600)은 캡처 데이터의 단일 프레임에서 또는 연속 프레임들에 걸쳐 사용자의 위치를 평가함으로써 사용자의 모션으로부터 제스처를 식별하고, 이 제스처를 비주얼 표현에 적용할 수 있다.

[0104] 시스템은 스캔된 데이터와 같은 캡처된 데이터, 이미지 데이터 또는 깊이 정보를 사용하여 특성들을 탐지할 수 있다. 탐지가능한 특성들로 시스템(600)에 의해 탐지될 수 있는 사용자 또는 물리적 공간에 관한 임의의 특성을 포함할 수 있다. 예를 들어, 탐지가능한 특성으로 타겟 특징(예컨대, 사용자의 안면 특징, 머리 색, 목소리 분석 등), 제스처(즉, 사용자가 행하고 시스템(600)이 인식하는 제스처), 이력 데이터(시스템에 의해 탐지되고 저장되는 사용자 경향 데이터 등의 데이터), 애플리케이션 상태(예컨대, 게임 애플리케이션에서 실패/성공) 또는 사용자의 기질을 나타내거나 사용자의 기질을 추론하기 위해 사용될 수 있는, 시스템에서 탐지가능한 기타 임의의 특징을 포함할 수 있다.

[0105] 시스템은 하나 이상의 탐지가능한 특성들을 분석하여 사용자의 기질을 추론할 수 있다. 추론(deduction)은 추정(inference) 또는 가정을 기초로 하거나, 또는 기질 및 연관 특성의 연구 결과와 같은 과학적인 방법에 기반할 수 있다. 따라서, 추론은 특정 기질, 특정 기질을 나타내는 제스처의 아이덴티티, 탐지가능한 특징과 심도 있는 심리학 분석과의 비교, 다양한 기질에 연관된 특성을 비롯한 전형적인 특성의 간단한 분석에 근거할 수 있다.

[0106] 타겟 특성은 특정 사용자(602)와 관계있는 정보, 예컨대, 행동, 말투, 얼굴 표정, 한 말(words spoken), 골격 움직임, 이력 데이터, 목소리 인식 정보 등을 포함할 수 있다. 타겟 특성은 타겟의 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. 인간 타겟에 관한 그 밖의 타겟 특성은 예를 들어, 키 및/또는 팔 길이를 포함할 수 있고, 이 특성은 신체 스캔, 골격 모델, 화소 영역 상 사용자(602)의 범위 또는 기타 임의의 적절한 프로세스 또는 데이터 등에 근거하여 얻어질 수 있다. 컴퓨팅 시스템(610)은 신체 인식 기법을 사용하여 이미지 데이터를 해석하고, 사용자(602)의 부속물의 크기, 모양 및 깊이에 따라서 사용자(602)의 비주얼 표현의 크기와 모양을 정할 수 있다.

[0107] 전술한 바와 같이, 시스템(600)은 사용자의 기질을 포함하는 물리적인 공간의 데이터를 식별할 수 있다. 예를 들어, 시스템(600)은 물리적 공간의 사용자의 모션, 얼굴 표정, 바디 랭귀지, 감정 등에 관련된 정보를 수집한다. 시스템(600)은 자세 인식 기법을 사용하여 인간 타겟(18)의 감정 또는 기질을 식별할 수 있다. 예를 들어, 시스템(600)은 사용자의 골격 모델을 분석하고 트래킹하여 사용자가 어떻게 움직이는지를 결정한다. 시스템(600)은 사용자의 신체 및, 애플리케이션, 운영 체제 등의 시스템의 양태를 제어하는 제스처를 포함하는, 사용자의 신체로 하는 모션들을 트래킹할 수 있다. 시스템은 사용자의 자세, 얼굴 표정, 목소리 표현 및 어조, 주시 시선 등을 식별할 수 있다. 사용자의 목소리 표현으로 사용자의 기질을 나타낼 수 있다. 예를 들어, 사용된 언어, 어조, 피치(pitch), 크기(volume) 등이 사용자의 기질에 대한 느낌을 전달할 수 있다. 예를 들어, 격한 어조는 분노 또는 공격성으로 해석될 수 있다. 다른 어조로는 긴장된 발성의(tense), 보통 발성의(modal), 숨이 새는(breathy), 속삭이는(whispery), 쥐어 짜내는 발성의(creaky), 차분한, 흥분한, 행복한 또는 기타 임의의 어조일 수 있다. 따라서, 사용자의 특성은 사용자의 기질의 유익한 지표이다.

[0108] 시스템은 시스템(600)에 의해 캡처된, 사용자의 탐지된 타겟 특성 중 적어도 하나를 사용자의 비주얼 표현에 적

용할 수 있다. 예를 들어, 시스템은 사용자가 안경을 쓰고 빨간 셔츠를 입고 있음을 탐지하고, 안경과 빨간 셔츠를 본 예시에서의 사용자의 비주얼 표현인 가상 원숭이(603)에 적용할 수 있다. 시스템은 사용자의 눈썹의 움직임 및/또는 찌푸리거나 웃는 표정을 비롯한 사용자의 얼굴 움직임을 식별할 수 있다. 시스템은 사용자가 한 말 및 사용자의 어조, 또는 사용자의 자세 등을 탐지할 수 있다. 예를 들어, 시스템은 사람의 오른팔을 탐지할 수 있고, 팔의 윗부분, 팔의 아랫부분, 손가락, 엄지 손가락, 손가락의 관절들 등을 구별하는 정확도를 가질 수 있다. 시스템은 사용자의 위아래 팔에 해당하는 사용자의 셔츠의 색상을 식별하고, 이 색상을 비주얼 표현에 적절하게 적용할 수 있다. 시스템은 손가락의 반지 또는 사용자의 손에 있는 문신을 식별하고, 시스템에서 생성되는 사용자의 모델에 따라, 물리적 공간의 사용자의 특징을 흉내내도록 비주얼 표현에 탐지된 타겟 특성을 적용할 수 있다. 비주얼 표현은 사용자처럼 보이고, 사용자처럼 움직이고, 사용자의 옷과 비슷한 옷을 입을 수 있다.

[0109] 시스템에 의해 탐지되고 사용자의 기질을 추론하는 데 사용되는 특정 타겟 특성은 사용자에게 직접적으로 적용되지 않고, 디스플레이 목적으로 바뀔 수 있다. 사용자의 특성은 비주얼 표현의 형태, 애플리케이션, 애플리케이션의 상태 등에 상응하도록 수정될 수 있다. 비주얼 표현이 가상 캐릭터라면, 특정 특성이 사용자의 비주얼 표현에 직접적으로 매핑되지 않을 수 있다. 예를 들어, 사용자(602)와 유사한 신체 비율이 디스플레이 장치(612)에 보이는 원숭이(603)와 같은 사용자의 캐릭터 표현에 제공되지만, 특정 캐릭터에 대해서는 수정될 수 있다. 사용자(602)와 유사한 키(height)가 원숭이 표현(603)에 제공되지만, 원숭이의 팔은 사용자의 팔보다 상당히 길다. 시스템에서 식별된 대로, 원숭이(603)의 양팔의 움직임은 사용자의 양팔의 움직임에 대응하지만, 시스템은 원숭이의 양팔이 움직이는 방식을 반영하도록 원숭이의 양팔의 애니메이션을 수정할 수 있다.

[0110] 도 6에 도시된 예시에서, 사용자는 머리를 한쪽으로 기울이고, 오른쪽 팔꿈치를 무릎에 대고, 사용자의 오른손으로 머리를 받치고 앉아 있다. 사용자의 표정, 자세, 한 말 또는 기타 임의의 탐지가능한 특성이 가상 원숭이(603)에게 적용되고, 적절하다면 수정될 수 있다. 예를 들어, 사용자가 물리적 공간에서 얼굴을 찌푸리고 있다. 시스템은 이러한 얼굴 표정을 탐지하여 찌푸림을 원숭이에게 적용하여, 가상 원숭이도 얼굴을 찌푸리게 한다. 또한, 그 자세의 원숭이의 신체 형태와 크기에 대응하도록 수정되는 것을 제외하고, 원숭이는 사용자와 유사한 자세로 앉아 있다. 유사하게, 시스템은 사용자의 타겟 특성을 사용하여 사용자의 기질을 추론한 후에, 그 기질을 나타내지만 사용자의 특성에 직접적으로 매핑될 수도 또는 매핑되지 않을 수도 있는 속성을 사용자의 비주얼 표현에 적용한다.

[0111] 시스템(600)은 탐지된 타겟 특성들을 가능한 기질의 라이브러리와 비교하고, 사용자의 비주얼 표현에 어떤 속성이 적용되어야 하는 지를 결정할 수 있다. 예를 들어, 컴퓨터(610)는 기질 정보를 모은 룩업 테이블들을 저장하며, 이에 대해서는 도 7 및 8과 관련하여 아래에서 설명될 것이다. 룩업 테이블은 특정한 또는 일반적인 기질 정보를 포함할 수 있다. 탐지된 특성들은 룩업 테이블과 비교되어 사용자의 기질을 추론한다. 분석에는 탐지된 자세, 얼굴 표정, 어조 및 말, 제스처, 이력 데이터 등의 비교가 포함된다.

[0112] 도 7은 사용자의 기질을 추론하고, 그 기질에 대응하는 비주얼 표현의 디스플레이를 위해 그 기질을 나타내는 속성을 선택하는 방법의 일례를 도시한다. 예를 들어, 702에서, 시스템은 사용자를 포함하는 물리적 공간으로부터 데이터를 수신한다. 전송한 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 해당하는 지를 결정한다. 인간 신체 모델과 일치하는 각 타겟 또는 사물이 스캔되어, 그에 연관된 골격 모델을 생성하게 된다. 이후에, 이 골격 모델은 골격 모델을 트래킹하고 그에 연관된 비주얼 표현을 렌더링하는 컴퓨팅 환경에 제공된다.

[0113] 704에서, 시스템은 사용자의 비주얼 표현을 렌더링한다. 예를 들어, 비주얼 표현은 모델을 기초로 한다. 물리적 공간(601)의 타겟의 비주얼 표현은 애니메이션, 캐릭터, 아바타 등 임의의 형태를 취할 수 있다. 비주얼 표현은 처음에는 사용자(602)가 원하는 모양 및 크기로 조각할 수 있는 디지털 진흙 덩어리이거나, 원숭이(603)와 같은 캐릭터 표현일 수 있다. 비주얼 표현은 캡처 장치에 의해 탐지된 사용자의 특징들을 기초로 하여 직접 모델링되거나, 또는 사용자의 선택된 특징들을 가진 가상 캐릭터일 수 있다. 비주얼 표현은 사용자(602)의 특징들 및 애니메이션 또는 저장 모델의 결합일 수 있다.

[0114] 706에서, 시스템은 사용자를 트래킹하고, 사용자의 기질을 나타내는 사용자의 특징들을 탐지한다. 예를 들어, 시스템은 사용자의 얼굴 표정과 신체 움직임을 트래킹하여 기질을 식별하고, 아바타가 사용자의 감정을 반영하도록 그 기질을 적용한다. 시스템은 임의의 탐지가능한 특징들을 사용하여 비주얼 표현에 적용되기 위한 사용자의 기질을 평가한다. 708에서, 시스템은 탐지된 특징들을 분석하고, 사용자의 기질을 추론한다. 예를 들어, 시



시스템의 프로세서는 기질 정보가 있는 록업 테이블 또는 데이터베이스를 저장한다. 사용자의 탐지된 특징들이 데이터베이스 또는 록업 테이블의 다양한 기질을 나타내는 특징들과 비교된다. 예를 들어, 록업 테이블은 "슬픈" 기질을 나타내는 특징을 정의한다. 이러한 특징들은 찌푸림, 눈물, 낮고 조용한 어조 및 가슴에 팔짱을 끼는 것 등이 있다. 물리적 공간에서 사용자의 임의의 또는 모든 이러한 특징들이 탐지되면, 프로세서는 사용자가 "슬픈" 기질을 드러내고 있다고 추론할 것이다.

[0115] 예를 들어, 록업 테이블 또는 데이터베이스는 애플리케이션에 적용되거나, 또는 시스템 전반에서 이용될 수 있다. 예를 들어, 게임 애플리케이션은 게임에 적용가능한 다양한 기질을 나타내는 특징들을 정의한다. 정의된 기질은 특정 또는 일반적인 기질을 포함하고, 하나 이상의 입력(즉, 탐지된 특징들)을 각 기질을 정의하는 특징들과 비교하여 기질을 식별한다. 록업 테이블 또는 데이터베이스에 관한 언급은 예시적인 것임을 알 것이며, 본원에 개시된 기법들에 관련된 기질 정보는 적절한 임의의 방식으로 액세스, 저장, 포장, 제공, 생성될 수 있다.

[0116] 이를 대신하여 또는 이에 결합하여, 710에서 시스템은 사용자에게 관해 캡처된 데이터로부터 기질 요청 제스처를 식별한다. 예를 들어, 사용자는 특정 제스처가 사용자의 비주얼 표현에 적용되게 하는 요청을 하는 제스처를 취할 수 있다.

[0117] 712에서, 시스템은 사용자의 제스처로부터 추론되거나 식별되는 기질을 반영하는, 사용자의 비주얼 표현에 적용될 속성을 선택한다. 특정 기질에 적용가능한 속성들도 록업 테이블 또는 데이터베이스 안에 존재한다. 선택된 속성들은 캡처 장치에 의해 탐지된 사용자의 특징이거나, 기질을 반영하는 애니메이션일 수 있다. 예를 들어, 시스템이 사용자가 "슬픈" 기질을 나타내는 특징들을 드러내고 있음을 추론한다면, 록업 테이블은 이러한 기질을 반영하는 다양한 애니메이션들을 보여줄 것이다. 시스템은 이러한 속성들 중 임의의 것을 선택하고 이를 사용자의 비주얼 표현에 적용하게 된다.

[0118] 714에서는, 속성들이 비주얼 표현에 실시간으로 적용된다. 따라서, 신체 인식 분석 등과 함께 사용자의 기분 또는 감정에 관련해 캡처된 데이터가 실시간으로 실행되어, 사용자의 비주얼 표현에 실시간으로 적용된다. 따라서 사용자는 사용자의 감정이나 기질의 실시간 디스플레이를 볼 수 있다.

[0119] 716에서 시스템은 시간이 감에 따라 물리적 공간의 사용자 및 모션을 계속해서 트래킹하고, 718에서 변화나 업데이트를 비주얼 표현에 적용하여 기질 변화를 반영한다. 예를 들어, 업데이트는 사용자의 탐지된 특징들의 변화 및 이력 데이터에 기반한다. 언제든지, 캡처 장치는 사용자의 행동 및 매너리즘, 감정, 말투 등을 식별하여 사용자의 기질을 결정하고, 이들을 사용자의 비주얼 표현에 적용할 수 있다. 업데이트는 실시간으로 비주얼 표현에 적용될 수 있다. 예를 들어, 시스템은 시간의 흐름에 따라 사용자의 표정을 캡처하고 흉내내어, 비주얼 표현을 통해 사용자의 기질을 반영하는 것이 바람직하다.

[0120] 도 8은 사용자의 기질을 추론하기 위해 사용되는 록업 테이블(800)의 일례를 도시한다. 도 8에 도시된 기질 록업 테이블(800)은 탐지가능한 특성들의 카테고리, 예를 들어, 얼굴 표정(802), 어조(804), 목소리 크기(808), 말(808), 자세(810), 제스처(812), 애플리케이션 결과(814), 및 이력 데이터(816)를 포함한다. 탐지된 특징 또는 특성들은 탐지가능한 타겟 특성, 애플리케이션 상태 등 시스템이 캡처 장치를 통해 정보를 캡처할 수 있는 물리적 공간의 임의의 특징을 포함한다. 록업 테이블(800)의 카테고리는 예시적인 것으로, 임의의 수 및 종류의 카테고리들도 사용자 기질 분석의 일부가 될 수 있다. 예를 들어, 카테고리는 다른 사용자 또는 사물들과의 탐지된 인터랙션, 사용자가 입은 옷의 종류 분석, 사용자의 신체에 있는 기타 아이템들 등을 더 포함할 수 있다. 시스템(600)에 의해 사용자의 태도 또는 기질의 분석의 일환으로 사용될 수 있는 방식으로 캡처되는 사용자의 탐지가능한 임의의 특징 또는 특성이 적용가능할 것이다.

[0121] 차트(800)에는 세 명의 사용자에게 대한 탐지된 특성들의 세 가지 예들이 도시되며, 각각의 행 A, B 및 C는 탐지된 특성들을 나타낸다. 테이블의 첫번째 부분(850)은 장면에서 캡처된 타겟의 탐지가능한 특성들을 나타낸다. 테이블의 두번째 부분(860)은 탐지가능한 다른 특성들, 예를 들어, 사용자가 취하는 제스처의 식별, 애플리케이션의 상태 및 그 결과, 사용자 또는 애플리케이션 전용의 이력 데이터를 나타낸다. 테이블의 마지막 부분(870)은 이용할 수 있는 탐지가능한 특징들의 분석 결과 사용자의 기질에 대한 시스템의 추론을 보여준다. 앞서 언급한 바와 같이, 테이블(800)은 단지 예시를 위한 것이며, 탐지가능한 추가 특성들을 포함할 수 있다.

[0122] A 행은 시스템에서 탐지되는 특성들의 일 실시예를 나타낸다. A 행에서, 시스템은, 제 1 사용자가 찌푸리는 얼굴 표정을 하고, 애플리케이션의 결과가 실패이며, 사용자가 실패한 결과 다음에는 찌푸리는 경향이 있음을 제 1 사용자에게 대한 이력 데이터에서 보여주고 있음을 탐지한다. 이러한 탐지된 특징들의 시스템 분석은 제 1 사용자의 기질이 "보통 부정적임"을 나타낼 것이다. 탐지가능한 추가 특징들로 보다 구체적인 기질이 제공될 수 있

지만, 이용가능한 데이터를 가지고도, 시스템은 보다 일반적으로, 보통 부정적인 기질을 추론한다.

[0123] 제 2 사용자의 경우, B 행에 제시된 탐지가능한 특징들을 가지고, 시스템은 찌푸린 얼굴 표정, 간결한 어조, 조용한 목소리, 말이 없음을 탐지하지만, 사용자의 자세는 뒤로 기댄 자세, 머리를 한쪽으로 떨어뜨려 한 손으로 받치고 있는 자세를 포함한다. 시스템은 이러한 특징들로부터 사용자의 기질이 보통 부정적이거나, 아니면 지루하고, 피곤하며, 화가 나고, 슬픈 것일 수도 있음을 결정한다. 제 2 사용자와 관련하여, 시스템은 게임 애플리케이션에서 다른 사용자가 플레이하는 순서이고, 다른 사용자의 순서가 오래 지속되었음을 탐지하고, 사용자의 이력 데이터의 분석에 의해 이러한 상황에서 이 사용자의 기질 경향을 탐지할 수 있다. 이러한 데이터로, 시스템은 제 2 사용자의 기질이 일반적으로 부정적일 뿐만 아니라 구체적으로 지루하거나 관심이 없다고 결정할 것이다. 예를 들어, 시스템은 제 2 사용자가 게임 애플리케이션에서 게임하는 플레이어가 아닐 때, 제 2 사용자가 "지루한" 기질에 해당하는 얼굴 표정, 어조, 자세 등을 가지는 경향을 식별할 수 있다.

[0124] 예를 들어, 찌푸리는 얼굴 표정은 많은 기질들에 해당할 수 있다. 테이블(800)에서 도시되는 기질 및 그 각각의 특정 기질을 나타내는 특징들은 오로지 예시적인 것이다. 탐지가능한 각각의 특징을 사용하여 보다 구체적인 태도 또는 기분으로 기질의 범위를 좁힐 수 있고, 또는 시스템은 예컨대, 일반적으로 부정적이거나 긍정적으로, 단순히 일반적인 태도를 식별할 수 있다.

[0125] C 행에 도시된 제 3 사용자의 탐지가능한 특성들은 미소짓는 얼굴 표정, 큰 목소리의 행복한 어조, "그래" 및 "평장해"와 같은 말, 및 손을 들고 펄쩍펄쩍 뛰는 자세를 포함한다. 펄쩍 뛰는 모션은 제 3 사용자의 성공적인 게임 결과가 나온 애플리케이션에 적용가능한 제스처를 나타내기도 한다. 이런 탐지가능한 특성들과 사용자의 이력 데이터의 비교 또한 상기 정보를 기초로 하여 제 3 사용자의 그럴듯한 기질을 나타낼 수 있다. 본 예시에서, 시스템은 탐지가능한 특성들에 기반하여 사용자의 기질이 "흥분했음"을 추론한다.

[0126] 시스템은 단순히 사용자의 실제 특성을 비주얼 표현에 매핑시킬 수 있다. 비주얼 표현이 직접적으로 사용자의 탐지된 특징들에 매핑되는 일 실시예에서는, 비주얼 표현에 사용자의 탐지된 특징들을 반영하므로 사용자의 기질이 비주얼 표현에 의해 자연적으로 나타난다. 그러나, 비주얼 표현이 항상 사용자의 직접적인 표현인 것은 아니기 때문에, 시스템은 비주얼 표현의 형태에 대응하도록 기질을 수정할 것이다. 사용자의 기질이 추론되면, 시스템은 그 기질을 반영하기 위해서 사용자의 비주얼 표현에 적용될 적절한 애니메이션들을 결정한다.

[0127] 예를 들어, 도 6은 원숭이 캐릭터의 대응 특징들을 나타내도록 수정되는, 사용자의 얼굴 표정, 자세 등의 사용자의 비주얼 표현(603)으로의 적용에 대해 도시한다. 원숭이가 찌푸리지만, 원숭이의 입이 사용자의 입에 직접적으로 매핑되지는 않고, 대신에 시스템이 원숭이가 찌푸릴 때처럼 보이도록 탐지된 찌푸림을 가상 원숭이의 입에 적용한다. 사용자의 기질을 사용자의 비주얼 표현으로 변환하는 것은 다양한 형태를 취할 수 있고, 입의의 수의 애니메이션을 포함할 수 있다. 예를 들어, 사용자의 비주얼 표현이 "집"이면, 집은 얼굴 특징들로 애니메이션화되지 않는다. 따라서, 시스템은 사용자의 기질을 새로운 형태로 변환하여 집에 매핑시킨다. 예를 들어, 시스템이 사용자의 얼굴 표정이나 자세에 근거하여 사용자의 "슬픈" 기질을 탐지하면, 가상 집의 가상 창문들을 축 늘어지게 디스플레이하고, 집을 부풀렸다가 앞문으로 공기를 내보내서 집이 한숨을 쉰 것처럼 보이도록 집을 애니메이션화하여, 이러한 기질을 집으로 변환할 수 있다.

[0128] 시스템은 탐지가능한 특성들에 근거하여, 사용자의 기분 또는 태도인 기질을 추론할 수 있다. 기질은 사용자의 느낌이나 생각을 표출하는 사용자의 감정적인 반응에 대한 입의의 표현을 포함할 수 있다. 식별된 기질은 일반적으로 긍정적이거나 부정적이며, 또는 양면성을 가질(ambivalent) 수 있다. 식별된 태도는 행복한, 화가 난, 좌절스러운, 지루한, 슬픈 등등 보다 구체적일 수 있다. 태도의 특이성은 태도/감정/기분의 라이브러리에 따라 달라지며, 시스템(600)은 일반적인 것부터 구체적인 것까지 사용자 태도의 범위를 식별한다. 예를 들어, 시스템은 사용자의 깨끗한 자세 및 즐거운 어조의 탐지가능한 특징들로부터 사용자가 일반적으로 긍정적인 태도를 가지고 있음을 결정한다. 또는, 시스템은 보다 구체적으로는, 깨끗한 자세가 펄쩍 뛰는 것과 올린 양팔을 포함하고, 이렇게 탐지되는 특성들은 흥분된 기질을 표시한다고 사용자의 이력 데이터에서 나타내기 때문에, 사용자가 흥분함을 결정할 수 있다. 다른 애플리케이션들은 일반적인 또한 구체적인 기분과 기질에 대한 방대한 데이터베이스를 가질 수 있으며, 기타 애플리케이션들도 예를 들어, 보통 긍정적인 또는 보통 부정적인 일반적인 기질을 추론할 수 있다.

[0129] 더 많은 탐지가능한 특징들로 사용자 태도에 대한 시스템 분석의 정확도를 높일 수 있다. 사용자의 자세의 변화는 사용자의 기질의 확실한 지표가 된다. 사용자의 자세는 사용자 신체의 자세, 사용자가 서고, 앉고, 자신의 가슴을 받치는 방식, 사용자가 자신의 팔, 다리 및 발을 어디에 두는 지를 포함한다. 예를 들어, 사용자가 머리를 한쪽으로 떨어뜨리고 한 손으로 받치면서 뒤로 기대있으면, 시스템은 사용자의 기질을 지루하거나 관심이 없



다고 식별한다. 또는, 예를 들어, 사용자가 머리를 세우고 가슴에 팔짱을 끼고 입을 오므리고 콧코이 앉아 있다면, 시스템은 의견이 불일치하거나, 방어적이거나 좌절스러운 것 중 하나인 것으로 사용자의 기질을 식별한다. 일반적으로, 사용자의 아바타에는 부정적인 의미(negative connotation)가 반영될 수 있다. 시스템은 목 또는 어깨 근육의 조임에 의한 사용자의 자세 변화를 탐지한다. 가끔, 사용자의 구부정한 자세는 단순히 사용자가 편히 쉬고 있거나 또는 나쁜 자세를 하고 있음을 나타낸다. 사용자의 머리의 자세도 사용자의 기질을 나타낼 수 있다. 시스템은 사용자가 턱을 경직시키거나 미간을 찡그리는 것을 탐지할 수 있다.

[0130] 도 9는 도 6에 도시된 시스템(600)을 도시하며, 시스템은 사용자의 탐지가능한 특징들을 트래킹하여 기질을 추론한다. 사용자의 탐지가능한 특징들을 비주얼 표현에 매핑시킴으로써, 사용자의 비주얼 표현에서 기질이 반영될 수 있다. 또한, 특정 기질에 대응하는 애니메이션을 사용자의 비주얼 표현에 적용시킴으로써 기질이 반영될 수 있다. 도 9는 세 시점에서 물리적 공간(601)의 사용자(602)를 도시하며, 901a, 901b 및 901c은 불연속적인 세 시점에서의 물리적 공간을 나타낸다. 각 시점에서, 사용자(602)는 자세를 바꾸고, 얼굴 표정을 바꾸고, 다른 모션을 하고, 또는 신체의 위치를 이동시킨다. 시스템(600)은 물리적 공간(601)에서 타겟인 사용자(602)를 각 시점에서 캡처하고, 901a, 901b 및 901c에 보이는 사용자의 탐지가능한 특징들을 캡처할 수 있다. 사용자(602)의 비주얼 표현의 디스플레이 결과의 두 가지 예시들이 디스플레이(912a) 및 디스플레이(902b)에서 도시된다.

[0131] 앞서 논의된 바와 같이, 사용자의 비주얼 표현은 임의의 애니메이션, 캐릭터, 아바타 등일 수 있다. 도 9에 도시되는 비주얼 표현의 예로 (디스플레이 장치(912a)에 도시되는) 아바타(905) 또는 (디스플레이 장치(912b)에 도시되는) 캐릭터(907)가 있다. 예를 들어, 아바타(905)는 물리적 공간의 사용자와 비슷한 표현으로, 사용자의 자세, 머리 색상, 옷 등에 매핑된다. 예를 들어, 캐릭터(907)는 도시된 원숭이와 같은 캐릭터 표현이다. 또한, 캐릭터(907)는 시스템(600)에서 캡처한 사용자의 특성들을 가질 수 있다. 예컨대, 얼굴 표정, 옷 등이 캐릭터 표현에 매핑된다.

[0132] 시스템(600)은 사용자의 기질 표시를 포함하는 물리적 공간의 데이터를 식별한다. 시스템(600)은 기질을 나타내는 속성을 사용자의 비주얼 표현에 적용시킴으로써 사용자의 기질을 비주얼 표현에 적용할 수 있다. 또한, 시스템(600)은 캡처 데이터의 단일 프레임에서 또는 연속 프레임들에 걸쳐 사용자의 위치를 평가함으로써 사용자의 모션으로부터 चे스처를 식별할 수 있다. 시스템(600)은 각 데이터 프레임으로부터의 정보, 데이터 프레임 간에 또는 시간에 따른 캡처된 데이터의 변화 정보, 캡처된 데이터로부터 식별된 चेсч터 및, 목소리 데이터와 같은 기타 이용가능한 정보의 조합을 사용하여 사용자의 기질이나 감정을 식별할 수 있다.

[0133] 일 실시예에서, 이미지 데이터의 분석을 통해 결정된 특성들이 아바타(905)에 제공된다. 사용자(602)는 자신의 특징들에 매핑되는 비주얼 표현을 선택할 수 있고, 물리적이거나 그렇지 않은 사용자(602) 자신의 특성이 비주얼 표현에 의해 표현된다. 아바타(905)와 같이, 아바타로도 불리는 사용자(602)의 비주얼 표현은 자세, 얼굴 표정 등을 비롯한 사용자(602)의 특성에 기반하여 초기화된다. 예를 들어, 골격 모델은 사용자(602)의 비주얼 표현의 생성을 위한 기본 모델이 될 수 있으며, 사용자(602)의 비율, 팔다리의 길이, 무게 등에 따라 모델링된다. 이후에, 사용자(602)의 머리 색, 피부, 옷 및 기타 탐지된 특성들이 비주얼 표현에 매핑될 것이다.

[0134] 비주얼 표현이 변화에 대해 조정될 수 있으므로, 사용자의 모션의 매핑이 사용자 움직임의 직접적인 변환이 아닐 수도 있다. 예를 들어, 사용자의 비주얼 표현이 얼굴 특징이 없는 상상 캐릭터일 수 있다. 시스템은 비주얼 표현의 형태에 적용가능한 다른 방식으로 사용자의 기질을 반영한다. 따라서, 사용자의 모션들은 몇몇 추가된 애니메이션을 가지고 비주얼 표현에 매핑되도록 변환되어, 비주얼 표현의 형태를 반영하게 된다. 예를 들어, 도 9에서, 디스플레이 장치(912b)에 도시된 사용자의 비주얼 표현은 원숭이 캐릭터(907)이다. 사용자(602)의 비주얼 표현(907)이 사용자의 물리적인 구조를 표현하지 않기 때문에, 사용자(602)의 모션 및/또는 기질은 시각 특성(907)이 취하는 형태와 일치하도록 변환될 것이다. 본 예시에서, 예를 들어, 탐지된 특징들 및/또는 기질은 원숭이(907)의 특징들과 일치하도록 변환된다.

[0135] 사용자의 기질을 나타내기도 하는 사용자의 특성이 탐지가능한 특성들의 시스템 분석에 기반하여 비주얼 표현으로 매핑됨으로써, 물리적 공간에서의 사용자의 모습 및/또는 움직임을 흉내낼 수 있다. 본 예시에서, 시스템은 세 시점(901a, 901b 및 901c)에서 물리적 공간의 사용자의 탐지가능한 특징들을 트래킹한다. 시스템은 제 위치에 있는 사용자(902a)가 한쪽으로 머리를 기대고 손으로 받친 채 앉아 있음을 탐지한다. 사용자(902a)는 찌푸리고 있고, 지루하거나 좌절한 기질을 나타내는 소리를 내거나 말을 하고 있다. 따라서, 시스템은 한동안 탐지가능한 특징들을 분석하여, 사용자의 기질을 추론할 수 있다.

[0136] 본 예시에서, 시스템은 사용자의 "지루한" 기질을 추론한다. 시스템은 시점(901a)에서 물리적 공간으로부터 캡처된 데이터로부터 사용자의 기질을 추론할 수 있다. 시스템은 사용자의 탐지가능한 특징들을 트래킹하고, 물리

적 공간(901b 및 901c)은 다른 시점에서의 사용자의 예들을 나타낸다. 시스템은 물리적 공간의 장면(901a)으로부터 캡처된 데이터와 같이, 캡처된 데이터의 단일 프레임을 기초로 하거나 또는, 세 장면 모두(901a, 901b, 901c)로부터 캡처된 데이터와 같이, 캡처된 데이터의 다수의 프레임 동안의 시간에 걸쳐, 추론된 기질을 나타내는 속성을 적용한다. 시스템은 단일 프레임 기반 및/또는 시간에 걸쳐 추론된 기질을 나타내는 속성들을 적용할 수 있다. 추론된 기질의 신뢰도는 사용자의 탐지가능한 특성들의 지속적인 분석에 따라 증가할 것이다. 또는, 시스템은 탐지가능한 특성의 변화에 따라 다른 기질을 탐지하거나 추론할 수 있다.

[0137] 시스템은 탐지된 특성들을 사용자의 비주얼 표현에 적용함으로써 이 특성들을 실시간으로 디스플레이할 수 있다. 따라서, 도 6에 도시된 바와 같이, 비주얼 표현(603)은 사용자의 탐지된 많은 특성(예컨대, 얼굴 표정, 자세 등)을 묘사한다. 유사하게, 시스템은 사용자의 타겟 특성을 사용하여 사용자의 기질을 추론한 후에, 그 기질을 나타내지만 사용자의 특성에 직접적으로 매핑될 수도 또는 매핑되지 않을 수도 있는 속성을 사용자의 비주얼 표현에 적용한다. 예를 들어, 시스템은 탐지된 특성들로부터 사용자가 "흥분하거나 행복한" 기질을 가질 것임을 추론한다. 이러한 기질을 나타내는 탐지된 특성들로는 펄쩍 뛰는 모션, 흥분해서 소리지르기, 게임 애플리케이션에서 성공 및 미소 등이 있다. 시스템은 예를 들어, 사용자의 기질을 추론하기 위해서, 이러한 특성들을 다양한 기질을 나타내는 특성을 가진 데이터베이스와 비교할 수 있다. 시스템은 타겟의 특성들이, 기질을 나타내는 속성들의 좋은 예가 되기 때문에, 이들 특성을 비주얼 표현에 직접 적용할 수 있다. 그러나, 시스템은, 적용된 속성들이 사용자의 특성의 직접적으로 매핑하는지 여부에 상관없이, 기질을 나타내는 속성들을 번갈아 또는 추가적으로 적용할 수 있다. 예를 들어, 시스템이 사용자의 탐지가능한 특징들로부터 "흥분하거나 행복한" 기질을 추론하는 경우, 시스템은 사용자의 비주얼 표현이 스크린 상에서 춤을 추게 애니메이션화 하거나 또는, 하늘로 뛰어올라 별을 잡게 애니메이션화할 수 있다. 시스템은 그 기질을 나타내는 다른 속성들, 예를 들어, (예컨대, "나는 정말 행복해" 또는 익살스럽거나 우스꽝스러운) 변칙이는 말을 디스플레이 장치에 적용할 수도 있다.

[0138] 도 9에서, 사용자의 탐지가능한 많은 특성들을 가진 아바타(905)의 애니메이션의 예로 벽에 머리를 대고 서서 "지루해"라고 말하는 아바타(905)가 있다. 사용자(602)는 시스템에 의해 캡처되는 어떤 시점에서도 이러한 행위를 하지 않고, 이러한 말을 하지 않지만, 시스템은 이 속성들이 "지루한" 기질을 나타내기 때문에 이들을 사용자의 비주얼 표현에 적용한다. 유사하게, 디스플레이 장치(912b)는 비주얼 표현의 디스플레이의 일례를 도시하며, 여기서 자신의 팔을 끌어당겨 아주 천천히 원숭이 소리 "우, 우, 아, 아"를 내는 원숭이 캐릭터(907)가 보인다. 원숭이에 적용되는 속성들은 지루한 기질을 나타낸다. 이 속성들은 예를 들어, 특정 테이블을 기초로 하여 시스템에 의해 식별되고, 원숭이와 같은 캐릭터 전용이거나, 또는 이 속성들은 많은 종류의 비주얼 표현들에 일반적으로 적용될 수 있다.

[0139] 아바타(905) 및 원숭이 표현(907)은 디스플레이될 수 있는 상이한 두 비주얼 표현의 예시이며, 예시된 디스플레이 장치(912a 및 912b)에서 보여진다. 각 비주얼 표현(905, 907) 및 사용자(602)의 기질을 나타내는 속성의 적용은 시점(901a)에서 물리적 공간에서 캡처된 것과 같이, 단일 세트의 캡처된 데이터를 기초로 할 것이다. 또는, 각 비주얼 표현(905, 907)의 디스플레이들의 예들은 시스템이 시간에 따라 사용자(602)를 모니터링한 결과일 수 있다. 시스템은 시간에 따라 캡처 데이터를 사용하여 사용자의 기질을 업데이트하고, 비주얼 표현에 특징들을 추가하고, 보다 구체적인 기질을 나타내는 속성들을 적용할 수 있다.

[0140] 사용자(602)는 특정 기질을 나타내는 속성을 사용자의 비주얼 표현에 적용하게 하는 제스처를 실행할 수 있다. 기질 아이덴티티 제스처는 특정 기질을 나타내는 속성들을 사용자의 비주얼 표현에 적용하는 요청으로 해석되는 제스처이다. 예를 들어, 도 9에서 시스템이 사용자의 "지루한" 기질을 탐지하는 것은 시스템이 "지루한" 기질을 나타내는 물리적 공간에서의 사용자의 제스처를 인식한 결과이다. 이 제스처는 예를 들어, 가슴에 팔짱을 끼고 있는 902c의 사용자의 자세를 포함할 수 있다. 이러한 모션과 단순히 이런 식으로 서있는 사용자의 모션을 구분하기 위해서, 제스처는 과장되게 팔짱을 끼는 것 또는 가슴에 팔짱을 천천히 끼는 것을 포함할 수 있다. 도 5(b)에 관해 설명된 제스처 인식 엔진(192)과 같은 제스처 인식 엔진은 사용자의 모션을 제스처 라이브러리(190)의 제스처에 대응하는 제스처 필터들과 비교할 것이다. 예를 들어, 사용자(602)의 캡처된 모션은 제스처 라이브러리(190)의 기질 아이덴티티 제스처(196)에 대응한다. 따라서, 비주얼 표현으로의 이러한 속성의 적용은 사용자의 제스처에 의해 제어되거나 사용자의 제스처로부터 인식될 수 있는 운영 체제 및/또는 애플리케이션의 일 양태이다.

[0141] 기질 아이덴티티 제스처는, 일반적으로 특정 기질에 연관된 특성들을 포함하거나 포함하지 않을 수 있다. 예를 들어, "슬픈" 기질에 관한 제스처가 한 손의 움직임일 수도 있는데, 이 손 움직임은 사람이 "슬픈" 기질일 때 일반적으로 하는 특성이 아닐 수 있다. 그러나, 손 움직임은, 사용자가 시스템이 "슬픈" 기질을 나타내는 속성

들을 비주얼 표현에 적용시키게 하는 제스처일 수 있다. 따라서, 사용자는 물리적 공간에서 제스처를 취함으로써 사용자의 비주얼 표현의 기질을 제어할 수 있다. 사용자는 기질에 대응하는 제스처를 의도적으로 또는 무심코 실행할 수 있다. 예를 들어, 사용자의 자신의 가슴에 팔짱을 끼는 제스처는 좌절의 기질로 인식되는 제스처이며, 사용자가 좌절스럽기 때문에, 단순히 이 제스처에 해당하는 모션을 행할 수 있다.

[0142] 찌푸림과 같은 사용자의 표정과 함께, 사용자가 좌절하고 있음을 나타내는 제스처의 시스템 인식 결과 비주얼 표현이 좌절의 기질을 반영하게 된다. 또는, 사용자는 특정 기질이 사용자의 비주얼 표현에 적용되도록 의도적으로 물리적 공간에서 제스처를 취할 수 있다. 예를 들어, 사용자가 방금 게임에서 이겼건, 어떤 일에 성공했다고 하자. "행복한" 기질에 관한 제스처는 사용자가 팔을 들어올리며 펄쩍 뛰는 모션을 포함할 수 있다. 사용자는, 시스템이 타겟 특성 및/또는 임의의 수의 "행복한" 특성을 사용자의 비주얼 표현에 적용시키게 하는 "행복한" 기질 제스처를 취할 수 있다. 예를 들어, 전술한 바와 같이, 사용자의 비주얼 표현은 옆으로 재주 넘기, 춤추기 또는, 시스템이 행복한 기질 표출에 결부시킨 기타 임의의 액티비티를 할 수 있다. 따라서, 가상 공간의 제스처들은 전자 게임과 같은 애플리케이션의 제어로서 동작하는 한편, 이 제스처들은 시스템이 사용자의 비주얼 표현에 특정 기질을 반영시키도록 하는 사용자의 요청에 대응할 수도 있다.

[0143] 시스템(600)은 탐지가능한 특성들을 모니터링하여 사용자의 기질을 사용자의 비주얼 표현에 업데이트할 수 있다. 시스템(600)은 시점(901a, 901b, 901c)에서 사용자로부터 캡처된 것과 같이 각 데이터 프레임으로부터의 정보, 데이터 프레임 간에 또는 시간에 따른 캡처된 데이터의 변화 정보, 캡처된 데이터로부터 식별된 제스처, 타겟 특성 및 타겟 특성의 변화 및, 얼굴 표정, 자세, 목소리 데이터와 같은 기타 이용가능한 정보의 조합을 사용하여, 사용자의 비주얼 표현에 의해 반영되는 기질을 식별하고 업데이트할 수 있다.

[0144] 물리적 공간의 사용자와 관련 타겟 특성들은 프로파일의 일부가 된다. 예를 들어, 프로파일은 특정한 물리적 공간 또는 사용자 전용일 수 있다. 사용자의 특징들을 포함하는 아바타 데이터는 사용자의 프로파일의 일부가 된다. 프로파일은 사용자가 캡처 장면에서 등장하면 액세스될 수 있다. 프로파일이 비밀 번호, 사용자의 선택, 신체 크기, 목소리 인식 등에 기반하여 사용자와 일치하면, 그 프로파일은 사용자의 비주얼 표현을 결정하는 데에 사용될 수 있다.

[0145] 사용자의 이력 데이터가 모니터링되고, 사용자의 프로파일에 정보를 저장한다. 예를 들어, 시스템은 사용자의 행동, 말투, 감정, 소리 등 사용자 특유의 특징들을 탐지할 수 있다. 시스템은 기질을 비주얼 표현에 적용시킬 때, 사용자의 비주얼 표현에 이러한 특징들을 적용시킨다. 예를 들어, 시스템이 사용자의 기질을 식별하고, 그 기질을 반영하는 말투를 포함하는 속성을 선택하면, 비주얼 표현의 목소리는 사용자의 말투로부터 패턴화되거나, 아니면 녹음된 사용자의 목소리일 수 있다.

[0146] 또한, 사용자 전용 정보는 한 명 이상의 사용자의 플레이 모드의 경향을 포함할 수 있다. 사용자가 특정 방식으로 행동하거나 반응하는 경향이 있으면, 시스템은 이러한 사용자의 경향을 트래킹하여 보다 정확하게 사용자의 기질을 추론할 수 있다. 예를 들어, 시스템이 "화난" 기질을 나타내는 사용자의 자세를 탐지하고, 사용자가 애플리케이션(예컨대, 게임)에서 실패할 때마다 유사한 방식으로 행동하는 경향이 있으면, 시스템은 이러한 정보를 트래킹할 수 있다. 따라서, 시스템은 사용자의 경향을 트래킹하기 시작하고, 이 정보를 사용하여 사용자의 기질을 보다 정확하게 추정할 수 있다.

[0147] 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.

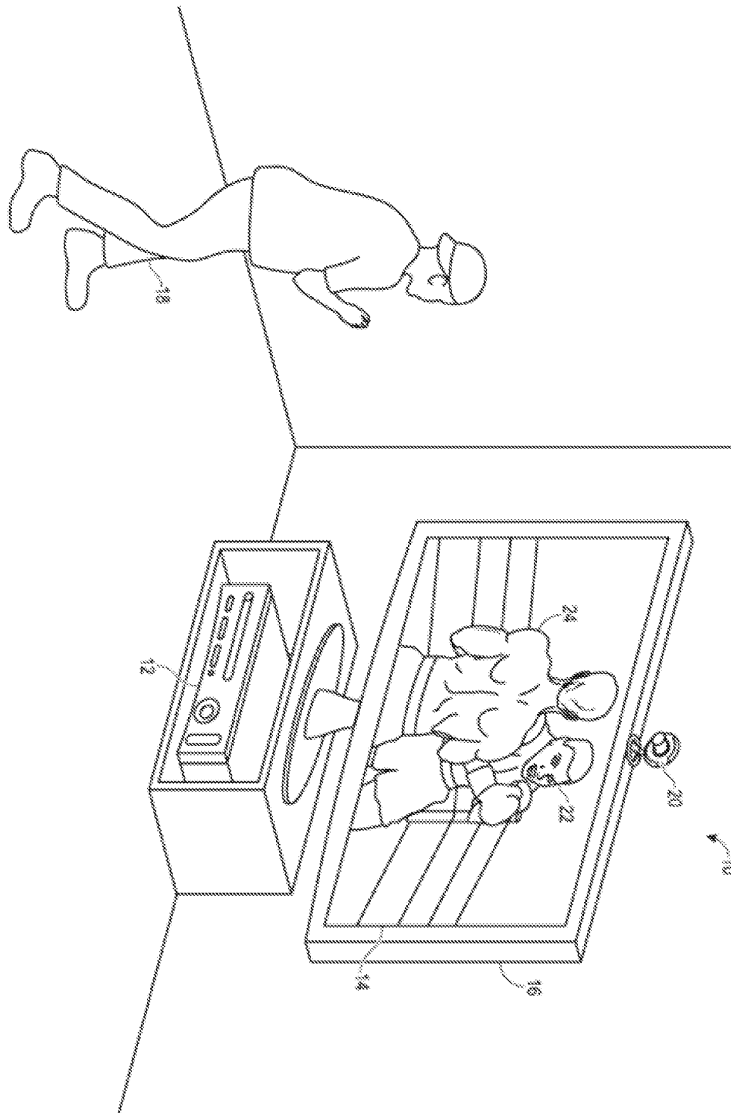
[0148] 또한, 본 개시는 특정 양태들과 관련되어 설명되었지만, 다양한 도면에 도시된 바와 같이, 본 개시를 벗어나지 않고 그 동일한 기능을 실행하도록 다른 유사한 양태들이 사용되거나 또는, 설명된 양태들에 대해 변경 및 추가될 수 있다. 본 개시의 대상은, 본원에 기술된 다양한 프로세스, 시스템 및 구성의 신규하고 자명하지 않은 모든 컴비네이션 및 서브 컴비네이션, 및 기타 특징, 기능, 행위 및/또는 특성들뿐만 아니라 그들의 임의의 모든 등가물을 포함한다. 따라서, 개시된 실시예들의 방법 및 장치, 또는 특정 양태들이나 그 일부분은, 플로피 디스크, 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.

[0149] 본원에 명시된 특정 구현예뿐만 아니라, 본원에 기술된 상세한 설명을 참고하여 그 밖의 양태들 및 구현예들도

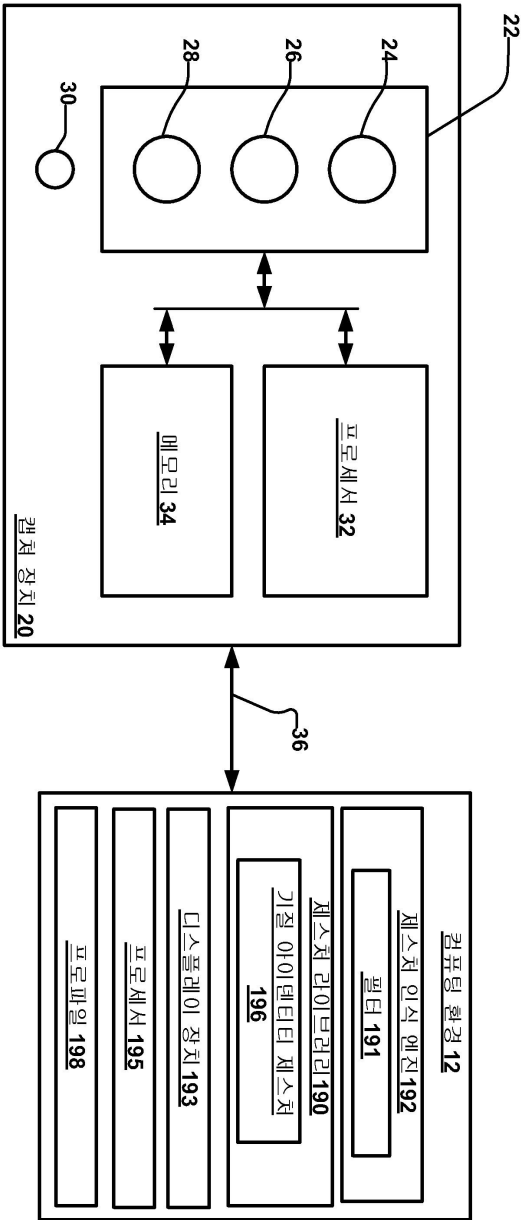
당업자에게 자명할 것이다. 따라서, 본 개시는 임의의 단일 양태에 제한되지 않고, 첨부된 특허청구범위에 따른 폭과 범위로 해석되어야 한다. 예를 들어, 본원에서 설명된 다양한 절차들은 하드웨어 또는 소프트웨어 또는 이들의 조합으로 구현될 수 있다.

## 도면

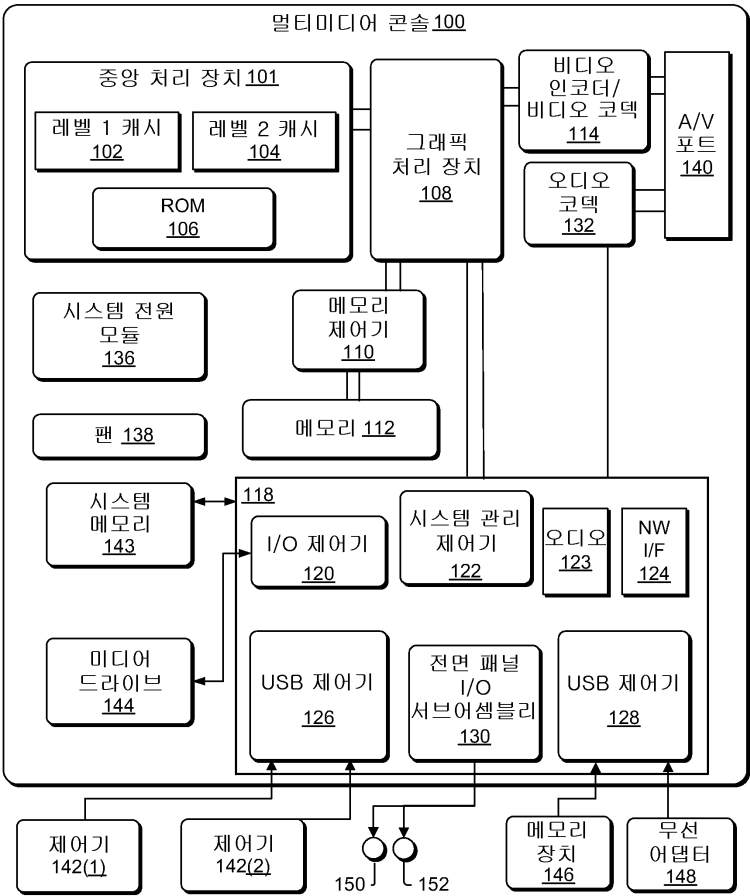
### 도면1



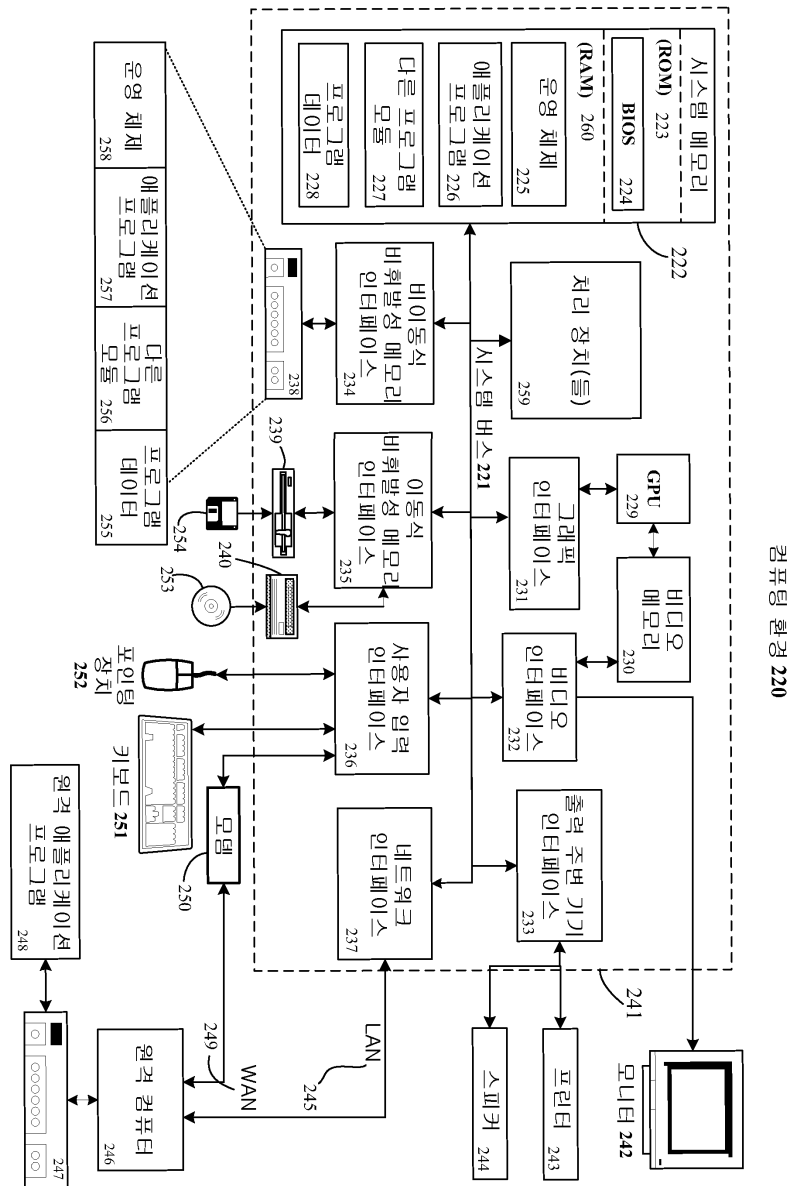
도면2



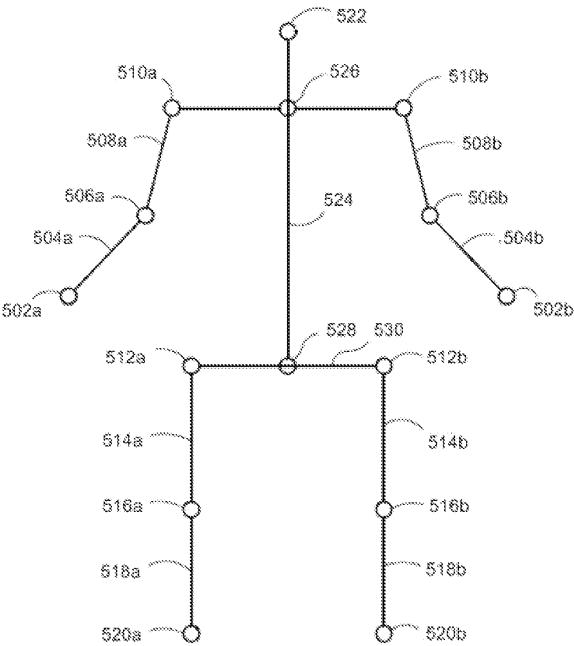
도면3



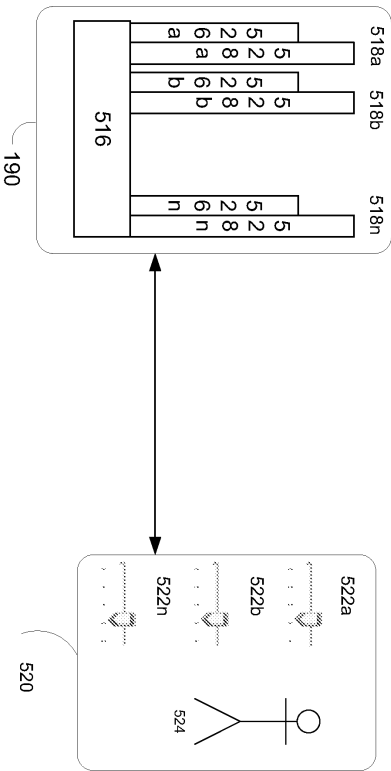
도면4



도면5a

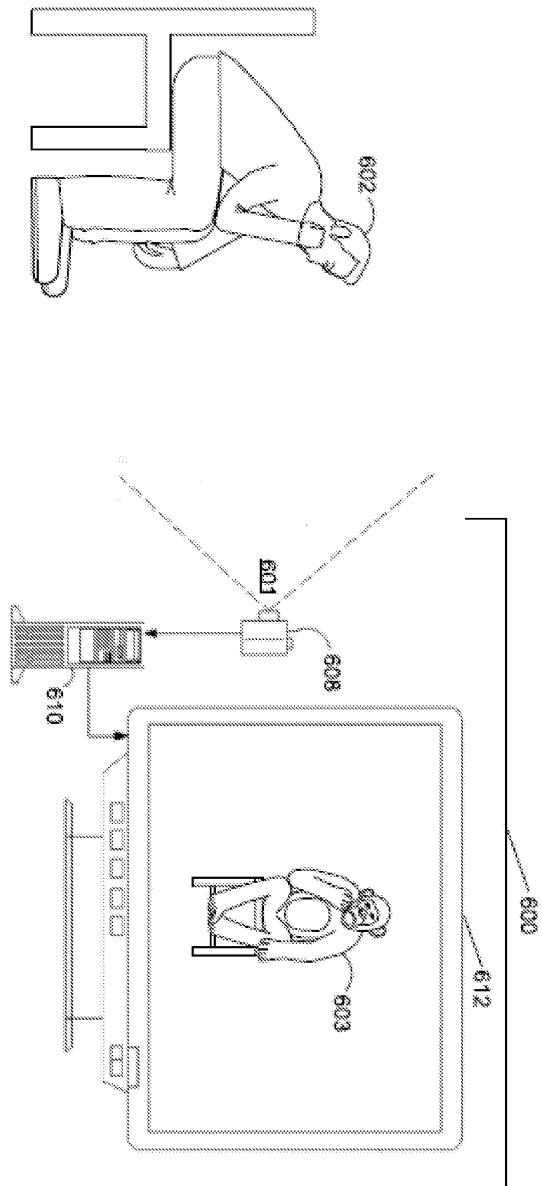


도면5b

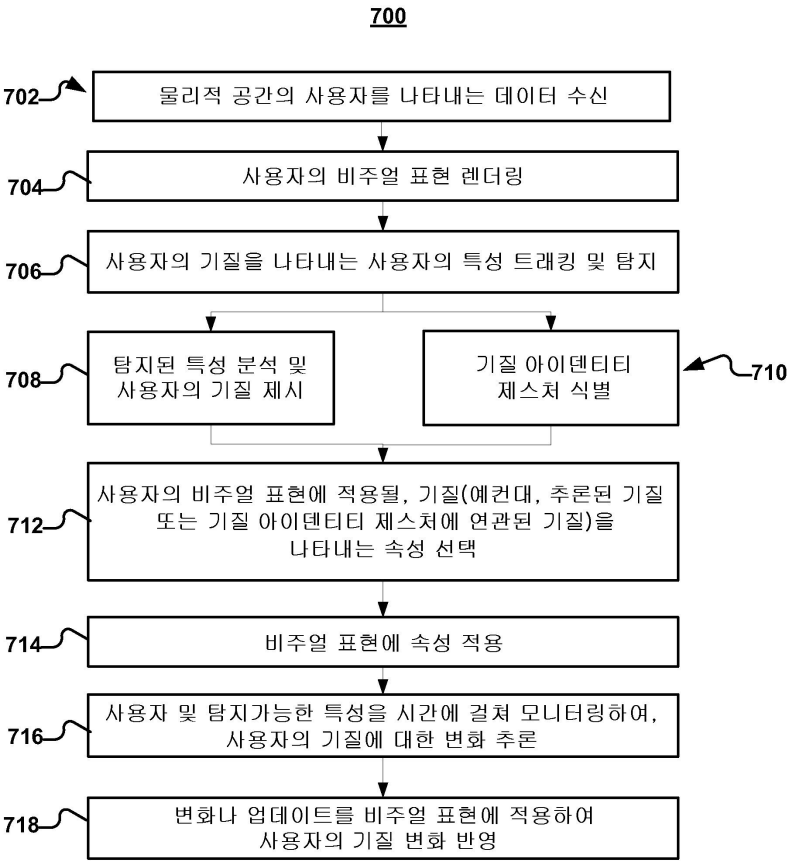




도면6



도면7



도면8

	열권 표정	아조	목소리 크기	말	자세	제스처	애플리케이션 결과	내력 데이터	기질
A	찌푸림	N/A	N/A	N/A	N/A	N/A	실패한 게임 결과	실패 결과 후 보통 찌푸림	보통 부정적인
B	찌푸림	간결함	조용함		뒤로 기대고, 머리를 한쪽으로 떨어뜨려 한 손으로 받침	발로 치는 모션	실패한 게임 결과	카테고리와 특정 사용자 내력 데이터의 비교	지루함, 관심 없음
C	미소지음	행복함	큼	그래, 굉장해!	손을 들고 펄쩍 뛸	펄쩍 뛸	성공한 게임 결과	카테고리와 특정 사용자 내력 데이터의 비교	충만함

도면9

