



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년05월30일
 (11) 등록번호 10-1741864
 (24) 등록일자 2017년05월24일

- (51) 국제특허분류(Int. Cl.)
 G06F 3/01 (2006.01) G06F 3/00 (2006.01)
 G06K 9/00 (2006.01) G06T 7/20 (2017.01)
 G10L 15/00 (2006.01)
- (21) 출원번호 10-2012-7018437
- (22) 출원일자(국제) 2010년12월31일
 심사청구일자 2015년12월03일
- (85) 번역문제출일자 2012년07월13일
- (65) 공개번호 10-2012-0126070
- (43) 공개일자 2012년11월20일
- (86) 국제출원번호 PCT/US2010/062655
- (87) 국제공개번호 WO 2011/087890
 국제공개일자 2011년07월21일
- (30) 우선권주장
 12/688,808 2010년01월15일 미국(US)
- (56) 선행기술조사문헌
 US20090315740 A1
 WO2009042579 A1
 US20090221368 A1
 US20090267894 A1

- (73) 특허권자
 마이크로소프트 테크놀로지 라이선싱, 엘엘씨
 미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
 마이크로소프트 웨이
- (72) 발명자
 마코빅 헬자
 미국 워싱턴주 98052-6399 레드몬드 원 마이크로
 소프트 웨이 엘씨에이 - 인터내셔널 페이턴즈 마
 이크로소프트 코포레이션
 라타 스티븐 지
 미국 워싱턴주 98052-6399 레드몬드 원 마이크로
 소프트 웨이 엘씨에이 - 인터내셔널 페이턴즈 마
 이크로소프트 코포레이션
 (뒷면에 계속)
- (74) 대리인
 김태홍

전체 청구항 수 : 총 12 항

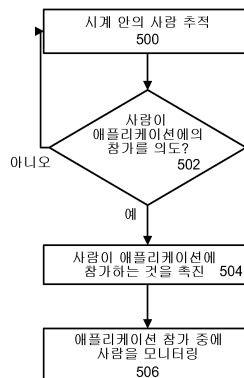
심사관 : 김창원

(54) 발명의 명칭 모션 캡처 시스템에서 사용자 의도 인식

(57) 요약

모션 캡처 시스템 안의 애플리케이션과의 상호작용을 촉진하는 기법들은 사람이 수동적 설정 없이 상호작용을 용이하게 시작할 수 있게 한다. 깊이 카메라 시스템은 물리적 공간 내 사람을 추적하며 애플리케이션에 참가하고자 하는 그 사람의 의도를 평가한다. 위치, 자세, 움직임 및 음성 데이터가 평가될 수 있다. 깊이 카메라의 시계 안의 절대 위치 및 다른 사람에 대한 상대 위치가 평가될 수 있다. 자세는 상호작용에 대한 의지를 나타내는 깊이 카메라 마주보기를 포함할 수 있다. 움직임은 물리적 공간 안에서 중심 영역을 향하거나 그로부터 멀어지는 동작, 시계를 통과하여 걷기, 및 팔 흔들기, 제스처 취하기, 또는 무게중심을 한 발에서 다른 발로 옮기는 것과 같이 일반적으로 한 위치에 서 있으면서 일어나는 움직임들을 포함할 수 있다. 음성 데이터는 볼륨뿐 아니라 음성 인식에 의해 인식되는 음성을 포함할 수 있다.

대표도



(72) 발명자

게이스너 케빈 에이

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

스티드 조나단 티

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

베네트 대런 에이

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

반스 아모스 디

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

명세서

청구범위

청구항 1

모션 캡처 시스템에 있어서,

시야(field of view)를 가지는 깊이(depth) 카메라 시스템과,

디스플레이와,

상기 깊이 카메라 시스템과 상기 디스플레이와 통신하는 적어도 하나의 프로세서로서, 상기 적어도 하나의 프로세서는 상기 모션 캡처 시스템에서의 애플리케이션을 구현하고 이미지들을 디스플레이 하기 위하여 상기 디스플레이에 신호를 제공하기 위한 명령어들을 실행하는 것인 상기 적어도 하나의 프로세서

를 포함하며,

상기 시야 내의 제1 사람 및 제2 사람을 추적하기 위한 상기 깊이 카메라 시스템 및 적어도 하나의 프로세서는, 상기 시야 내의 상기 제1 사람의 신체 및 상기 제2 사람의 신체를 구별하고, 상기 제2 사람은 상기 디스플레이 상의 가상 공간 내의 아바타를 제어하기 위하여 상기 제2 사람의 신체를 제어함으로써 상기 애플리케이션과 관련되며, 상기 제2 사람은 상기 아바타에 결속(bind)되는 반면에 상기 제1 사람은 상기 애플리케이션과 관련되고자 하는 의도를 가지는 것으로 인식되지 않으며,

상기 추적에 기초하여, 상기 적어도 하나의 프로세서는, 상기 시야 내의 상기 제2 사람의 신체의 위치에 대한 상기 제1 사람의 신체의 위치에 기초하여, 상기 적어도 하나의 프로세서가 상기 제1 사람이 상기 애플리케이션과 관련되고자 하는 의도를 가졌다고 결정할 때 상기 제1 사람으로 하여금 상기 애플리케이션과 관련되게 하는 것인 모션 캡처 시스템.

청구항 2

제1항에 있어서, 상기 추적에 기초하여, 상기 추적은 상기 제2 사람에 의해 상기 애플리케이션과 분리되기 위한 의도를 결정하며, 상기 적어도 하나의 프로세서는, 상기 제2 사람 대신에 상기 제1 사람으로 하여금 상기 애플리케이션과 관련되게 하며, 상기 제1 사람의 신체를 이동시킴으로써 상기 제1 사람으로 하여금 상기 디스플레이 상의 상기 가상 공간 내의 상기 아바타를 제어하게 하기 위하여, 상기 제2 사람 대신에 상기 제1 사람을 상기 아바타에 결속시키는 것을 포함하는 것인 모션 캡처 시스템.

청구항 3

제1항에 있어서, 상기 추적에 기초하여, 상기 추적은, 상기 제2 사람에 의해 상기 애플리케이션과 관련되어 남겨지도록 하는 의도를 결정하며,

상기 적어도 하나의 프로세서는, 상기 제1 사람으로 하여금 자동 생성에 의해 상기 애플리케이션과 관련되게 하며, 상기 제1 사람으로 하여금 상기 디스플레이 상의 상기 가상 공간 내의 별개의 아바타를 제어하게 하는 반면에 상기 제2 사람은 상기 디스플레이 상의 상기 가상 공간 내의 그 아바타에 결속되어 그 아바타를 제어하도록 상기 제1 사람을 별개의 아바타에 결속시키는 것인 모션 캡처 시스템.

청구항 4

제1항에 있어서, 상기 제1 사람으로 하여금 상기 애플리케이션과 관련되게 하고, 상기 적어도 하나의 프로세서는 상기 디스플레이를 적어도 제1 및 제2 영역으로 분할하고, 상기 제1 영역은 상기 제1 사람의 아바타를 제공하며, 상기 제1 사람으로 하여금 시점(view)의 제1 포인트로부터 상기 가상 공간을 조망하게 하고, 상기 제2 영역은 상기 제2 사람의 아바타를 제공하며, 상기 제2 사람으로 하여금 상이한 시점의 제2 포인트로부터 상기 가상 공간을 조망하게 하는 것인 모션 캡처 시스템.

청구항 5

제4항에 있어서, 상기 적어도 하나의 프로세서는, 제2 사람의 아바타를 제1 영역으로, 상기 제1 사람의 아바타

를 상기 제2 영역으로 스왑(swap)하고, 상기 추적은 상기 제1 및 제2 사람이 상기 시야 내의 스왑된 위치들을 가지므로, 상기 제1 영역은, 제2 사람의 아바타를 제공하며, 상기 제2 사람으로 하여금 시점의 제2 포인트로부터 가상 공간을 조망하게 하며, 상기 제2 영역은 상기 제1 사람의 아바타를 제공하며, 상기 제1 사람으로 하여금 시점의 제1 포인트로부터 상기 가상 공간을 조망하게 하는 것인 모션 캡처 시스템.

청구항 6

제4항에 있어서, 상기 적어도 하나의 프로세서는, 상기 추적에 기초하여, 상기 시야 내의 상기 제1 사람 및 상기 제2 사람의 상대 위치에 기초하여 상기 디스플레이 상의 상기 제1 및 제2 영역의 상대 위치를 설정하는 것인 모션 캡처 시스템.

청구항 7

모션 캡처 시스템에서 방법을 수행하기 위하여 적어도 하나의 프로세서를 프로그래밍하기 위해 구현된 컴퓨터 판독가능한 소프트웨어를 가진 유형의(tangible) 컴퓨터 판독가능한 저장 장치에 있어서,

상기 방법은,

상기 모션 캡처 시스템의 시야 내의 제1 사람의 신체를 포함하는 장면과 연관되는 이미지들을 수신하는 단계와,

상기 이미지들에 기초하여, 상기 시야 내의 상기 제1 사람 - 상기 제1 사람은 디스플레이 상의 가상 공간 내의 아바타를 제어하기 위하여 상기 제1 사람의 신체의 움직임에 의해 애플리케이션과 상호작용함 - 의 신체를 구별하고, 상기 장면 내의 상기 제1 사람의 신체의 움직임을 식별하고 상기 시야 내의 적어도 하나의 추가적인 사람 - 상기 적어도 하나의 추가적인 사람은 상기 디스플레이 상의 상기 가상 공간 내의 아바타를 제어하지 않음 - 을 구별하는 단계와,

상기 적어도 하나의 추가적인 사람의 행위에 관한 미리 정해진 기준이 충족될 때, 상기 애플리케이션의 시각적 및 가청(audible) 출력 중 적어도 하나를 변경하는 단계를 포함하는 유형의 컴퓨터 판독가능한 저장 장치.

청구항 8

제7항에 있어서, 상기 미리 정해진 기준은, 상기 적어도 하나의 추가적인 사람이 상기 애플리케이션에 대한 관심의 수준이 증가했다는 것을 나타내는 방식으로 이동하는 것을 포함하는 것인 유형의 컴퓨터 판독가능한 저장 장치.

청구항 9

제7항에 있어서, 상기 미리 정해진 기준은, 상기 적어도 하나의 추가적인 사람이 상기 애플리케이션에 대한 관심의 수준이 증가했다는 것을 나타내는 사운드를 생성하는 것을 포함하는 것인 유형의 컴퓨터 판독가능한 저장 장치.

청구항 10

모션 캡처 시스템에서의 프로세서 구현 방법에 있어서,

상기 모션 캡처 시스템의 시야 내에 제1 사람의 신체를 포함하는 장면과 연관되는 이미지들을 수신하는 프로세서 구현 단계와,

상기 이미지들에 기초하여, 상기 시야 내의 상기 제1 사람 - 상기 제1 사람은 디스플레이 상의 가상 공간 내의 아바타를 제어하기 위하여 상기 제1 사람의 신체의 움직임에 의해 애플리케이션과 상호작용함 - 의 신체를 구별하고, 상기 장면 내의 상기 제1 사람의 신체의 움직임을 식별하고 상기 시야 내의 적어도 하나의 추가적인 사람 - 상기 적어도 하나의 추가적인 사람은 상기 디스플레이 상의 상기 가상 공간 내의 아바타를 제어하지 않음 - 을 구별하는 프로세서 구현 단계와,

상기 적어도 하나의 추가적인 사람의 행위에 관한 미리 정해진 기준이 충족될 때, 상기 애플리케이션의 시각적 및 가청 출력 중 적어도 하나를 변경하는 프로세서 구현 단계를 포함하는 프로세서 구현 방법.

청구항 11

제10항에 있어서, 상기 미리 정해진 기준은, 상기 적어도 하나의 추가적인 사람이 상기 애플리케이션에 대한 관

심의 수준이 증가했다는 것을 나타내는 방식으로 이동하는 것을 포함하는 것인 프로세서 구현 방법.

청구항 12

제10항에 있어서, 상기 미리 정해진 기준은, 상기 적어도 하나의 추가적인 사람이 상기 애플리케이션에 대한 관심의 수준이 증가했다는 것을 나타내는 사운드를 생성하는 것을 포함하는 것인 프로세서 구현 방법.

청구항 13

삭제

청구항 14

삭제

청구항 15

삭제

발명의 설명

배경 기술

[0001] 모션 캡처 시스템은 물리적 공간 안에서 사람이나 다른 대상의 위치 및 움직임에 관한 데이터를 획득하여, 그 데이터를 컴퓨팅 시스템 내 임의의 애플리케이션으로의 입력으로서 사용할 수 있다. 군사용, 오락용, 스포츠용 및 의료용과 같은 많은 애플리케이션들이 가능하다. 예를 들어, 사람의 모션은 3차원 인체 골격 모델로 매핑되어 움직이는 캐릭터나 아바타를 생성하는데 사용될 수 있다. 가시 및 비가시 (예를 들어 적외선) 광 등을 이용하는 것을 포함하는 광학 시스템은 시계(a field of view) 안의 사람의 존재를 검출하기 위해 카메라를 이용한다. 마커 없는 시스템들 역시 개발되어 왔으나, 검출을 돕기 위해 마커(marker)가 사람 위에 놓여질 수 있다. 일부 시스템들은 움직임을 검출하기 위해 사람에 의해 휴대되거나 부착되는 관성 센서들을 사용한다. 예를 들어, 일부 비디오 게임 애플리케이션에서, 사용자는 게임을 하면서 움직임을 검출할 수 있는 무선 제어기를 보유한다. 그러나, 사람이 애플리케이션과 보다 자연스럽게 상호작용할 수 있게 하는 더욱 정교한 발명이 필요하다.

발명의 내용

과제의 해결 수단

[0002] 모션 캡처 시스템을 이용하는 물리적 공간 내의 애플리케이션과 상호작용하고자 하는 사용자의 의도를 인식하기 위한 프로세서 구현 방법, 모션 캡처 시스템 및 유형적인 컴퓨터 판독가능 저장부가 제공된다.

[0003] 모션 캡처 시스템에 의해 제공되는 엔터테인먼트나 다른 체험에 대한 접근성을 극대화하기 위해, 간단한 자동적 기법이 제공되어 사람들이 소수의 명시적 액션 또는 명시적 액션 없이 예컨대 그 체험에 로그인하거나 그로부터 로그아웃함으로써 체험으로 들어오거나 빠져 나가게 한다. 사람은 그냥 걸어 들어와 모션 캡처 시스템의 플레이를 시작하고, 플레이를 마쳤을 때 떠나며, 다시 돌아와 그가 떠났던 곳으로부터 플레이를 재개할 수 있다. 해당 애플리케이션에 참가하거나 그로부터 참가 해제하고자 하는 사람의 의도가 사람의 위치, 움직임, 자세 및/또는 음성 데이터에 기반하여 판단된다. 또한, 모션 캡처 시스템의 시계 안에 있는 관람자들과 같은 다른 사람들의 의도 역시 검출될 수 있다.

[0004] 일 실시예에서 모션 캡처 시스템의 어떤 애플리케이션에 참가하고자 하는 사람의 의도를 인식하기 위한 프로세서 구현 방법이 제공된다. 그 방법은 모션 캡처 시스템의 시계 안에서 그 사람의 몸을 추적하는 것을 포함한다. 추적은 그 사람의 움직임을 식별하기 위하여 골격 모델을 이용하는 것과 같은 것을 통해, 시계 안의 사람을 분간한다. 추적에 기반하여, 상기 방법은 그 사람이 첫 번째 시점에서는 애플리케이션에 참가하고자 의도하지 않지만 두 번째 시점에서는 그 애플리케이션에 참가하고자 한다고 판단한다. 그 사람이 애플리케이션에 참가하고자 의도할 때, 그가 애플리케이션에 참가하는 것이 허용된다. 참가를 용이하게 하기 위해 다양한 단계들이 취하여지고, 그에 따라 그 사람이 키보드나 다른 수동 입력 기기 상에 패스워드와 사용자 로그인 아이디어를

입력하는 것과 같은 수동적 입력을 제공해야 할 어떤 필요성을 완화할 수 있다. 구체적으로, 프로파일 및 아바타가 자동으로 생성되어 그 사람과 연관될 수 있다. 아바타는 디스플레이 상의 가상 공간 안에 디스플레이될 수 있다. 디스플레이는 사람이 자신의 몸을 움직여 아바타를 제어함으로써 애플리케이션에 참가할 때, 시계 안에서 그 사람의 신체에 대한 뒤이은 추적에 기반하여 업데이트된다.

[0005] 이 요약은 이하의 설명에 자세히 기술되는 개념들의 선택을 간략한 형식으로 소개하기 위해 주어진다. 이 요약은 청구된 발명 대상의 주요 특징이나 필수적 특징을 확인하도록 의도되거나 청구된 주제의 범위를 한정하는 데 사용되도록 의도된 것이 아니다.

도면의 간단한 설명

[0006] 도 1a 및 1b는 사용자가 복싱 경기를 시뮬레이션하는 애플리케이션과 상호작용하는 모션 캡처 시스템의 전형적 실시예를 도시한다.

도 2는 도 1a의 모션 캡처 시스템(10)의 전형적 블록도를 도시한다.

도 3은 도 1a의 모션 캡처 시스템에서 사용될 수 있는 컴퓨팅 환경의 전형적 블록도를 도시한다.

도 4는 도 1a의 모션 캡처 시스템에서 사용될 수 있는 컴퓨팅 환경의 또 하나의 전형적 블록도를 도시한다.

도 5는 사람이 모션 캡처 시스템과 상호작용할 수 있게 하는 방법을 도시한다.

도 6a는 도 5의 단계(500)에 기술된 것과 같이 사람의 움직임의 추적을 전형적 방법을 도시한다.

도 6b는 도 6a의 단계(608)에 기술된 것과 같이 사람의 전형적 모델을 도시한다.

도 6c는 도 6a의 단계(608)에 기술된 것과 같이 사람의 또 다른 전형적 모델을 도시한다.

도 7a는 도 5의 단계(502)에 기술된 것과 같이 사람이 어떤 애플리케이션에 참가하고자 의도하는지의 여부를 판단하는 전형적 방법을 도시한다.

도 7b는 도 7a의 단계(700)에 논의된 바와 같이, 소정 위치에 대해 상대적으로 어떤 사람의 위치가 평가되는 시계의 예를 도시한다.

도 7c는 도 7a의 단계(700)에 논의된 바와 같이, 또 다른 사람에 대해 상대적으로 어떤 사람의 위치가 평가되는 시계의 예를 도시한다.

도 7d는 도 7a의 단계(704)에 논의된 바와 같이, 사람의 움직임이 평가되는 시계의 예를 도시한다.

도 8a는 도 5의 단계(504)에 기술된 것과 같이 사람이 애플리케이션에 참가하는 것을 촉진시키게 하는 전형적 방법을 도시한다.

도 8b는 또 다른 사용자가 현재 어떤 애플리케이션에 참가 중에 도 5의 단계(504)에 기술된 것과 같이 어떤 사람이 그 애플리케이션에 참가하는 것을 촉진시키게 하는 전형적 방법을 도시한다.

도 9a는 도 5의 단계(506)에 기술된 것과 같이 애플리케이션에 참가 중에 사람을 모니터링하는 전형적 방법을 도시한다.

도 9b는 사람이 애플리케이션으로부터 참가를 해제하고 이후 다시 그 애플리케이션에 참가하는 경우, 도 5의 단계(506)에 기술된 것과 같이 애플리케이션에 참가 중에 그 사람을 모니터링하는 전형적 방법을 도시한다.

도 9c는 플레이어가 모션 캡처 시스템의 시계 안에서 위치를 서로 바꾼 경우, 도 5의 단계(506)에 기술된 것과 같이, 그리고 도 8b의 방법에 뒤이어서 애플리케이션에 참가 중에 사람을 모니터링하는 전형적인 방법을 도시한다.

도 9d는 어떤 애플리케이션에 참가하고 있지 않은 추가된 사람에 대한 추적에 기초하여 애플리케이션의 출력이 수정되는 경우, 도 5의 단계(506)에 기술된 바와 같이 그 애플리케이션에 참가 중에 어떤 사람을 모니터링하는 전형적인 방법을 도시한다.

도 9e는 어떤 애플리케이션에 참가 시 한 사람이 다른 사람을 대신하는 경우, 도 5의 단계(506)에 기술된 것과 같이 애플리케이션에 참가 중에 사람을 모니터링하는 전형적 방법을 도시한다.

도 10a는 한 사람이 어떤 애플리케이션에 참가하고 있고, 다른 사람들은 그 애플리케이션에 참가하고 있지 않은 예를 도시한다.

도 10b는 두 사람이 한 애플리케이션에 참가하고 있고 디스플레이 상의 가상 공간에 대한 공통 시점을 가지는 전형적 디스플레이를 도시한다.

도 10c는 두 사람이 한 애플리케이션에 참가하고 있고 디스플레이 상의 가상 공간에 대한 서로 다른 시점을 가지는 전형적 스플릿 스크린 디스플레이를 도시한다.

도 10d는 도 10c의 디스플레이 영역들과 관련하여, 사람들이 물리적 공간에서 위치를 서로 바꿀 때 디스플레이 영역들이 맞아꾸어지는 전형적인 스플릿 스크린 디스플레이를 도시한다.

도 10e는 첫 번째 사람이 프로파일 및 아바타를 이용하는 애플리케이션에 참가하는 전형적 디스플레이를 도시한다.

도 10f는 프로파일 및 아바타를 이용하는 애플리케이션에 참가 시 두 번째 사람이 첫 번째 사람을 대신할 때 도 10e에 뒤이은 전형적 디스플레이를 도시한다.

발명을 실시하기 위한 구체적인 내용

- [0007] 한 사람이나 사람들의 그룹이 모션 캡처 시스템의 어떤 애플리케이션과 쉽게 상호작용할 수 있게 하기 위한 다양한 기법들이 제공된다. 깊이 카메라 시스템은 물리적 공간 안에서 어떤 사람의 위치 및 움직임을 추적하며 그것들을 평가하여, 그 사람이 애플리케이션에 참가하고자 하는지 여부, 예컨대 애플리케이션과 상호작용하고자 하는지 여부를 판단할 수 있다. 물리적 공간으로부터 오디오 입력과 같은 추가 정보 역시 사용될 수 있다. 어떤 경우, 어떤 사람의 움직임이나 위치는 시계 안의 하나 이상의 다른 사람들을 기준으로 평가된다. 또한 애플리케이션에서 어떤 사람의 아바타를 언제 활성화할지, 어떤 사람이 현재 그 애플리케이션에 참가하고 있는지 아닌지 여부를 반영하기 위해 디스플레이를 언제 수정할지를 판단하기 위한 기법들이 제공된다. 어떤 사람이 서로 다른 시간대에 어떤 애플리케이션에 참가하고 있을 때, 그 사람이 연속성의 상실 없이 그 관계를 지속할 수 있도록 정보가 저장된다. 예를 들어, 어떤 사람이 점수를 획득하거나 게임에 소비한 시간과 그의 진척상태에 기반하여 소정 과위를 얻는 게임과 같은 어떤 애플리케이션에서, 그 사람은 앞서 만들어졌던 진척상태를 상실하지 않고 플레이를 재개할 수 있다.
- [0008] 도 1a 및 1b는 어떤 사람(18)이 복싱 경기를 시뮬레이션하는 애플리케이션과 상호작용하는 모션 캡처 시스템(10)의 전형적 실시예를 도시한다. 모션 캡처 시스템(10)은 사용자나 플레이어라고 일컬어질 수도 있는 사람(18)과 같은 인간 타깃을 인식, 분석, 및/또는 추적하는 데 사용될 수 있다.
- [0009] 도 1a에 도시된 바와 같이, 모션 캡처 시스템(10)은 컴퓨터, 게임 시스템 또는 콘솔 등과 같은 컴퓨팅 환경(12)을 포함할 수 있다. 컴퓨팅 환경(12)은 교육 목적 및/또는 오락 목적과 같은 애플리케이션들을 실행하기 위한 하드웨어 구성요소들 및/또는 소프트웨어 구성요소들을 포함할 수 있다.
- [0010] 모션 캡처 시스템(10)은 깊이(depth) 카메라 시스템(20)을 더 포함할 수 있다. 깊이 카메라 시스템(20)은 예컨대 사람(18)과 같은 한 명 이상의 사람들을 시각적으로 모니터링하는데 사용될 수 있는 카메라일 수 있으므로, 이하에서 보다 상세히 기술되는 것처럼, 아바타나 온 스크린 캐릭터를 애니메이션하는 것과 같이 어떤 애플리케이션 안에서 하나 이상의 컨트롤이나 액션을 수행하기 위해 사람들에게 의해 행해진 제스처 및/또는 움직임들이 캡처, 분석, 및 추적될 수 있다.
- [0011] 모션 캡처 시스템(10)은 사용자에게 시청각적 출력을 제공하는 텔레비전, 모니터, 고선명도 텔레비전(HDTV)과 같은 시청각 장치(16)에 연결될 수 있다. 오디오 출력은 별개의 장치를 통해 제공될 수도 있다. 시청각 장치(16)를 구동하기 위해, 컴퓨팅 환경(12)은 애플리케이션과 관련된 시청각 신호들을 제공하는 그래픽 카드 같은 비디오 어댑터 및/또는 사운드 카드 같은 오디오 어댑터를 포함할 수 있다. 시청각 장치(16)는 가령 S-비디오 케이블, 동축 케이블, HDMI 케이블, DVI 케이블, VGA 케이블 등을 통해 컴퓨팅 환경(12)에 연결될 수 있다.
- [0012] 사람(18)의 제스처 및/또는 움직임이 캡처되어, 아바타나 온 스크린 캐릭터를 애니메이션하는데 사용되고/거나 컴퓨터 환경(12)에 의해 실행되고 있는 애플리케이션에 대한 입력 컨트롤들로서 사용되도록, 깊이 카메라 시스템(20)을 사용하여 사람(18)이 추적될 수 있다. 따라서, 일 실시예에 따르면, 사용자(18)는 애플리케이션을 제어하고/하거나 아바타나 다른 온 스크린 캐릭터를 애니메이션하기 위해 자신의 신체를 움직일 수 있다.
- [0013] 예로서, 애플리케이션은 사람(18)이 참여하고 시청각 장치(16)가 그 사람(18)에게 복싱 상대(38)에 대한 시각적

표현을 제공하는 복싱 게임일 수 있다. 컴퓨팅 환경(12)은 또한 시청각 장치(16)를 사용하여, 사람을 표현하고 사용자(18)가 자신의 신체 움직임을 통해 제어할 수 있는 플레이어 아바타(40)에 대한 시각적 표현을 제공할 수도 있다.

- [0014] 예를 들어, 도 1b에 도시된 바와 같이, 사람(18)이 물리적 공간, 예컨대 그 사람이 서 있는 방 안에서 펀치를 날려, 플레이어 아바타(40)가 복싱 링을 포함하는 가상 공간 안에서 펀치를 날리게 할 수 있다. 따라서, 전형적인 일 실시예에 따르면, 모션 캡처 시스템(10)의 컴퓨터 환경(12) 및 깊이 카메라 시스템(20)은 물리적 공간에서의 사람(18)의 펀치를 인식 및 해석하여 그 펀치가 가상 공간에서 플레이어 아바타(40)를 제어하기 위해 복싱 경기를 시뮬레이션하는 애플리케이션에 대한 입력으로서 해석될 수 있도록 사용될 수 있다.
- [0015] 사람(18)에 의한 다른 움직임들 역시, 다른 컨트롤들이나 액션들로서 해석되고/거나, 빠르게 움직이거나 비틀거리거나 질질끌거나 막거나 꺾을 날리거나 다른 다양한 펀치를 날리는 컨트롤들 같이, 플레이어 아바타를 애니메이션하는데 사용될 수 있다. 또한 일부 움직임들은 플레이어 아바타(40)를 컨트롤하는 것 이외의 액션들에 해당할 수 있는 컨트롤들로서 해석될 수 있다. 예를 들어, 일 실시예에서, 플레이어는 게임을 종료하거나 일시 정지하거나 저장하고, 레벨을 선택하고, 고득점을 확인하며, 친구와 통신하는 등의 움직임들을 이용할 수 있다. 플레이어는 메인 사용자 인터페이스로부터 게임이나 다른 애플리케이션을 선택하기 위해 움직임들을 이용할 수 있다. 그에 따라, 애플리케이션과 상호작용하도록 사용자(18)의 모든 모션이 어떤 적절한 방식으로 이용 가능하고, 사용되며 분석될 수 있다.
- [0016] 사람은 애플리케이션과 상호동작할 때 지지대(prop)와 같은 객체를 잡을 수 있다. 그러한 실시예들에서 사람과 객체의 움직임이 애플리케이션을 제어하는 데 사용될 수 있다. 예를 들어, 테니스 게임을 시뮬레이션하는 애플리케이션에서, 라켓을 든 플레이어의 모션이 온 스크린 라켓을 제어하는데 추적되고 사용될 수 있다. 또 다른 전형적 실시예에서, 플라스틱 칼과 같은 장난감 무기를 든 사용자의 모션이 해석선을 제공하는 애플리케이션의 가상 공간 안에서 해당 무기를 제어하는 데 추적되고 이용될 수 있다.
- [0017] 모션 캡처 시스템(10)은 운영체제 및/또는 게임 영역 밖에 있는 애플리케이션 컨트롤들 및 오락과 레저에 의도된 다른 애플리케이션들로서 타겟 움직임들을 해석하는데 추가로 사용될 수 있다. 예를 들어, 운영체제 및/또는 애플리케이션의 가상적인 어떤 제어가능한 양태가 사람(18)의 움직임에 의해 제어될 수 있다.
- [0018] 도 2는 도 1a의 모션 캡처 시스템(10)의 전형적 블록도를 도시한다. 깊이 카메라 시스템(20)은 예컨대 비행 시간(time-of-flight), 구조화된 광(structured light), 스테레오 이미지 등을 포함하는 어떤 적절한 기법을 통해 깊이 값들을 포함할 수 있는 깊이 이미지를 구비한 깊이 정보와 함께 비디오를 캡처하도록 구성될 수 있다. 깊이 카메라 시스템(20)은 깊이 정보를 "Z 계층들", 또는 깊이 인식 카메라로부터 그 시선을 따라 연장되는 Z 축에 수직일 수 있는 계층들 안에 조직할 수 있다.
- [0019] 깊이 카메라 시스템(20)은 물리적 공간 내 어떤 장면에 대한 깊이 이미지를 캡처하는 깊이 카메라와 같은 이미지 카메라 컴포넌트(22)를 포함할 수 있다. 깊이 이미지는 캡처된 장면의 이차원(2D) 픽셀 영역을 포함할 수 있으며, 2D 픽셀 영역 내 각각의 픽셀은 이미지 카메라 컴포넌트(22)로부터 선형 거리를 나타내는 관련 깊이 값을 가진다.
- [0020] 이미지 카메라 컴포넌트(22)는 장면의 깊이 이미지를 캡처하는데 사용될 수 있는 적외선(IR) 광 컴포넌트(24), 삼차원(3D) 카메라(26), 및 적녹청(RGB) 카메라(28)를 포함할 수 있다. 예를 들어, 비행 시간(TOF) 분석시, 깊이 카메라 시스템(20)의 IR 광 컴포넌트(24)는 물리적 공간 상에 적외선 광을 발산하고, 센서들(미도시)을 사용하여 가령 3D 카메라(26) 및/또는 RGB 카메라(28)를 이용하여 물리적 공간 내 한 개 이상의 타겟 및 오브젝트의 표면으로부터 후방산란된 광을 검출할 수 있다. 일부 실시예에서, 펄싱되는 적외선 광은 나가는 광 펄스와 그에 대응하여 들어오는 광 펄스 사이의 시간이 측정되어 깊이 카메라 시스템(20)으로부터 물리적 공간 내 타겟이나 오브젝트들 상의 특정 위치까지의 물리적 거리를 판단하는데 사용될 수 있도록 이용될 수 있다. 위상 쉬프트를 판단하기 위해 나가는 광 파동의 위상이 들어오는 광 파동의 위상과 비교될 수 있다. 그런 다음 위상 쉬프트가 깊이 카메라 시스템으로부터 타겟이나 오브젝트들 상의 특정 위치까지의 물리적 거리를 판단하는데 사용될 수 있다.
- [0021] 가령 서터링되는 광 펄스 이미징을 포함하는 다양한 기법들을 통하여 시간에 따른 광의 반사 빔 강도를 분석함으로써, TOF 분석은 깊이 카메라 시스템(20)으로부터 타겟이나 오브젝트들 상의 특정 위치까지의 물리적 거리를 간접적으로 판단하는데 사용될 수 있다.
- [0022] 또 다른 전형적 실시예에서, 깊이 카메라 시스템(20)은 깊이 정보를 캡처하기 위해 구조화된 광을 사용할 수 있

다. 그러한 분석시, 패턴화된 광(즉, 그리드 패턴이나 스트라이프 패턴 같이 알려진 패턴으로서 디스플레이되는 광)이 가령 IR 광 컴포넌트(24)를 통해 장면 위로 투사될 수 있다. 장면 내 한 개 이상의 타깃이나 오브젝트의 표면을 칠 때, 그에 응하여 패턴이 변형될 수 있다. 그러한 패턴의 변형은 가령 3D 카메라(26) 및/또는 RGB 카메라(28)에 의해 캡처될 수 있고 그런 다음 깊이 카메라 시스템으로부터 타깃이나 오브젝트들 상의 특정 위치까지의 물리적 거리를 판단하기 위해 분석될 수 있다.

[0023] 또 다른 실시예에 따르면, 깊이 카메라 시스템(20)은 서로 다른 각도로부터 장면을 볼 수 있는 둘 이상의 물리적으로 분리된 카메라들을 포함함으로써, 깊이 정보를 산출하기 위해 결정될 수 있는 비주얼 스테레오 데이터를 획득할 수 있다.

[0024] 깊이 카메라 시스템(20)은 예컨대 소리 파동을 수신하여 전기 신호로 변환하는 트랜스듀서나 센서를 포함하는 마이크로폰(30)을 더 포함할 수 있다. 또한, 마이크로폰(30)은 컴퓨팅 환경(12)에 의해 실행되는 애플리케이션을 제어하기 위해 어떤 사람에 의해 주어진 소리들과 같은 오디오 신호를 수신하는데 사용될 수 있다. 오디오 신호는 소리낸 말, 휘파람, 외침 및 기타 발성음과 같은 음성 사운드뿐 아니라 손뼉치기나 발로 쿵쿵대기와 같은 비음성 사운드를 포함할 수 있다.

[0025] 깊이 카메라 시스템(20)은 이미지 카메라 컴포넌트(22)와 통신하는 프로세서(32)를 포함할 수 있다. 프로세서(32)는 예를 들어 이하에서 보다 상세히 기술되는 바와 같이, 깊이 이미지를 수신하고; 깊이 이미지에 기반하여 복셀(voxel)들의 그리드를 생성하고; 복셀들의 그리드에 포함된 배경을 제거하여 인간 타깃과 관련된 한 개 이상의 복셀들을 분리하고; 분리된 인간 타깃의 한 개 이상의 손발의 위치나 자리를 판단하고; 한 개 이상의 손발의 위치나 자리, 또는 어떤 다른 적절한 명령에 기반하여 모델을 조정하는 명령어들을 포함하는 명령어들을 실행할 수 있는 위치표준형 프로세서, 특수형 프로세서, 마이크로프로세서 등을 포함할 수 있다.

[0026] 깊이 카메라 시스템(20)은 프로세서(32)에 의해 실행되는 명령어들을 저장할 뿐 아니라, 3D 카메라나 RGB 카메라에 의해 캡처되는 이미지들이나 이미지들의 프레임들, 또는 어떤 다른 적절한 정보, 이미지 등을 저장할 수 있는 메모리 컴포넌트(34)를 더 포함할 수 있다. 전형적인 일 실시예에 따르면, 메모리 컴포넌트(34)는 RAM(random access memory), ROM(read only memory), 캐시, 플래시 메모리, 하드 디스크, 또는 어떤 다른 적절한 유형의 컴퓨터 저장가능 저장 컴포넌트를 포함할 수 있다. 메모리 컴포넌트(34)는 버스(21)를 통해 이미지 캡처 컴포넌트(22) 및 프로세서(32)와 통신하는 별도의 컴포넌트일 수 있다. 또 다른 실시예에서, 메모리 컴포넌트(34)는 프로세서(32) 및/또는 이미지 캡처 컴포넌트(22) 안에 병합될 수 있다.

[0027] 깊이 카메라 시스템(20)은 통신 링크(36)를 통해 컴퓨팅 환경(12)과 통신할 수 있다. 통신 링크(36)는 유선 및/또는 무선 접속일 수 있다. 일 실시예에 따르면, 컴퓨팅 환경(12)은 통신 링크(36)를 통해 깊이 카메라 시스템(20)으로, 깊이 카메라 시스템(20)의 시계 안에 있는 물리적 공간으로부터 언제 이미지 데이터를 캡처할지를 지시하는 클럭 신호를 제공할 수 있다.

[0028] 또한 깊이 카메라 시스템(20)은 가령 3D 카메라(26) 및/또는 RGB 카메라(28)에 의해 캡처된 깊이 정보 및 이미지들, 및/또는 이미지 카메라 시스템(20)에 의해 생성될 수 있는 골격 모델을 통신 링크(36)를 통해 컴퓨팅 환경(12)으로 제공할 수 있다. 그러면 컴퓨팅 환경(12)은 애플리케이션을 제어하기 위해 모델, 깊이 정보, 및 캡처된 이미지들을 이용할 수 있다. 예를 들어 도 2에 도시된 바와 같이, 컴퓨팅 환경(12)은 제스처 필터들의 모음과 같은 제스처 라이브러리(190)를 포함할 수 있으며, 각각의 필터는 (사용자가 움직이면서) 골격 모델에 의해 수행될 수 있는 제스처에 관한 정보를 가진다. 예를 들어 제스처 필터는 한 팔이나 양 팔을 위나 옆으로 올리기, 양 팔을 등글게 돌리기, 새처럼 양팔을 펼치기, 앞뒤나 옆으로 기울이기, 또는 위로 점프하기, 뒤꿈치를 들어올려 발가락으로 서 있기, 제자리에서 걷기, 시계/물리적 공간 안에서 다른 위치로 걸어가기 등의 동작들 각각을 제공받을 수 있다. 검출된 모션을 각각의 필터와 비교함으로써, 어떤 사람에 의해 행해진 특정 제스처나 움직임이 식별될 수 있다. 움직임이 어느 정도까지 수행되었는지 역시 판단될 수 있다.

[0029] 깊이 카메라 시스템(20)에 의해 캡처된 골격 모델 형태의 데이터와 그와 관련된 움직임들은 사용자(골격 모델에 의해 표현됨)가 언제 하나 이상의 특정 움직임을 행했는지를 식별하기 위해 제스처 라이브러리(190) 안의 제스처 필터들과 비교될 수 있다. 그러한 움직임들은 애플리케이션의 다양한 컨트롤과 관련될 수 있다.

[0030] 컴퓨팅 환경은 또한, 디스플레이 장치(196)로 시청각적 출력 신호를 제공하고 여기 기술된 바와 같은 다른 기능을 수행하기 위해 메모리(194)에 저장된 명령어들을 실행하는 프로세서(192)를 포함할 수 있다.

[0031] 도 3은 도 1a의 모션 캡처 시스템에서 사용될 수 있는 컴퓨팅 환경의 전형적 블록도를 도시한다. 컴퓨팅 환경은 하나 이상의 제스처들이나 다른 움직임들을 해석하고, 그에 대응하여 디스플레이 상의 시각적 공간을 업데이트

트하는 데 사용될 수 있다. 도 1a, 1b, 및 2와 관련하여 위에서 기술된 컴퓨팅 환경(12)과 같은 컴퓨팅 환경은 게임 콘솔과 같은 멀티미디어 콘솔(100)을 포함할 수 있다. 멀티미디어 콘솔(100)은 레벨 1 캐시(102), 레벨 2 캐시(104), 및 플래시 ROM(Read Only Memory)(106)을 가진 중앙 프로세싱 유닛(CPU)(101)을 가진다. 레벨 1 캐시(102) 및 레벨 2 캐시(104)는 임시로 데이터를 저장하고 그에 따라 메모리 액세스 사이클 횟수를 줄임으로써 프로세싱 속도 및 처리율을 향상시킬 수 있다. CPU(101)는 한 개를 넘는 코어, 그에 따른 추가적 레벨 1 및 레벨 2 캐시들(102 및 104)을 가지는 것으로 제공될 수 있다. 플래시 ROM(106)은 멀티미디어 콘솔(100)이 파워 온(ON) 될 때 부팅 프로세스의 초기 단계 중에 로드되는 실행가능 코드를 저장할 수 있다.

[0032] 그래픽 프로세싱 유닛(GPU)(108) 및 비디오 인코더/비디오 코덱(코더/디코더)(114)이 고속 고해상도의 그래픽 프로세싱을 위한 비디오 프로세싱 파이프라인을 형성한다. 데이터는 버스를 통해 그래픽 프로세싱 유닛(108)으로부터 비디오 인코더/비디오 코덱(114)으로 전달된다. 비디오 프로세싱 파이프라인은 데이터를 텔레비전이나 다른 디스플레이로 전송하기 위해 A/V(오디오/비디오) 포트(140)로 출력한다. 메모리 제어기(110)가 GPU(108)에 연결되어 RAM(Random Access Memory)과 같은 다양한 타입의 메모리(112)에 대한 프로세서 액세스를 돕는다.

[0033] 멀티미디어 콘솔(100)은 모듈(118) 상에 바람직하게 구현되는 I/O 제어기(120), 시스템 관리 제어기(122), 오디오 프로세싱 유닛(123), 네트워크 인터페이스 제어기(124), 제1USB 호스트 제어기(126), 제2USB 제어기(128) 및 전면 패널 I/O 하위 어셈블리(130)를 포함한다. USB 제어기들(126 및 128)은 주변 제어기들(142(1)142(2)), 무선 어댑터(148) 및 외부 메모리 장치(146)(가령, 플래시 메모리, 외부 CD/DVD-ROM 드라이브, 착탈가능 매체 등)에 대한 호스트로서의 역할을 한다. 네트워크 인터페이스(124) 및/또는 무선 어댑터(148)는 네트워크(가령, 인터넷, 홈 네트워크 등)로의 액세스를 제공하며, 이더넷 카드, 모뎀, 블루투스 모듈, 케이블 모뎀 등을 포함하는 광범위한 유무선 어댑터 컴포넌트들 중 어느 것일 수 있다.

[0034] 시스템 메모리(143)가 제공되어 부팅 프로세스 중에 로드되는 애플리케이션 데이터를 저장한다. 매체 드라이브(144)가 제공되며, DVD/CD 드라이브, 하드 드라이브, 또는 다른 착탈형 미디어 드라이브를 포함할 수 있다. 매체 드라이브(144)는 멀티미디어 콘솔(100)의 내부나 외부에 있을 수 있다. 애플리케이션 데이터가 멀티미디어 콘솔(100)에 의한 실행, 재생 등을 위해 매체 드라이브(144)를 통해 액세스될 수 있다. 매체 드라이브(144)는 직렬 ATA 버스나 기타 고속 접속과 같은 버스를 통해 I/O 제어기(120)에 연결된다.

[0035] 시스템 관리 제어기(122)는 멀티미디어 콘솔(100)의 가용성을 보장하는 것과 관련된 다양한 서비스 기능들을 제공한다. 오디오 프로세싱 유닛(123) 및 오디오 코덱(132)은 높은 충실도를 가지고 스테레오 처리를 하는 해당 오디오 프로세싱 파이프라인을 형성한다. 오디오 데이터는 통신 링크를 통해 오디오 프로세싱 유닛(123) 및 오디오 코덱(132) 사이에서 전송된다. 오디오 프로세싱 파이프라인은 데이터를 외부 오디오 플레이어나 오디오 사양을 가진 장치에 의한 재생을 위해 A/V 포트(140)로 출력한다.

[0036] 전면 패널 I/O 하위 어셈블리(130)는 파워 버튼(150)과 이젝트(eject) 버튼(152)의 기능 뿐 아니라 멀티미디어 콘솔(100)의 외부 면 상에 노출되는 어떤 LED(light emitting diodes)나 기타 표시자들의 기능을 지원한다. 시스템 전력 공급 모듈(136)은 멀티미디어 콘솔(100)의 구성요소들로 전력을 제공한다. 팬(138)은 멀티미디어 콘솔(100) 안의 회로를 쿨링한다.

[0037] CPU(101), GPU(108), 메모리 제어기(110), 및 멀티미디어 콘솔(100) 안의 다른 다양한 구성요소들은 다양한 버스 구조들 중 어느 하나를 이용하는 직병렬 버스, 메모리 버스, 주변기기 버스, 및 프로세서나 로컬 버스를 포함하는 한 개 이상의 버스들을 통해 서로 연결된다.

[0038] 멀티미디어 콘솔(100)이 파워 온 될 때, 애플리케이션 데이터가 시스템 메모리(143)로부터 메모리(112) 및/또는 캐시들(102, 104)로 로드되고 CPU(101) 상에서 실행된다. 애플리케이션은 멀티미디어 콘솔(100)에서 사용가능한 각종 미디어 유형들을 탐색할 때 일관된 사용자 경험을 제공하는 그래픽 사용자 인터페이스를 제공할 수 있다. 동작시, 애플리케이션들 및/또는 매체 드라이브(144) 안에 포함되는 기타 매체가 매체 드라이브(144)로부터 시동되거나 재생되어 멀티미디어 콘솔(100)로 추가 기능들을 제공하도록 할 수 있다.

[0039] 멀티미디어 콘솔(100)은 단순히 시스템을 텔레비전이나 다른 디스플레이로 연결함으로써 단독형 시스템으로서 작동될 수 있다. 이러한 단독형 모드에서, 멀티미디어 콘솔(100)은 한 명 이상의 사용자들이 시스템과 상호작용하거나 영화를 보거나 음악을 듣게 할 수 있다. 그러나, 네트워크 인터페이스(124)나 무선 어댑터(148)를 통하여 가능한 광대역 접속의 통합을 이용하여, 멀티미디어 콘솔(100)은 더 큰 네트워크 커뮤니티 안의 참여자로서 추가 운영될 수 있다.

[0040] 멀티미디어 콘솔(100)이 파워 온될 때, 특정 분량의 하드웨어 자원들이 멀티미디어 콘솔 운영체제에 의한 시스

템 사용을 위해 예비된다. 그 자원들은 메모리(가령 16MB), CPU 및 GPU 사이클(가령 5%)의 비축, 네트워킹 대역폭(가령 8kbs) 등을 포함할 수 있다. 이러한 자원들은 시스템 부팅 시점에 준비되기 때문에, 예비 자원들은 애플리케이션의 시점으로부터 존재하지 않는다.

[0041] 특히, 메모리 예비는 시동 커널(launch kernel), 동시적 시스템 애플리케이션들 및 드라이버들을 포함하기 충분할 만큼 많은 것이 바람직하다. CPU 예비는 예비된 CPU 사용이 시스템 애플리케이션들에 의해 이용되지 않는 경우 아이들(idle) 스레드가 어떤 미사용 사이클을 소비하도록 일정한 것이 바람직하다.

[0042] GPU 예비와 관련하여, 시스템 애플리케이션(가령, 팝업)에 의해 생성된 경량의 메시지가 팝업을 오버레이로 렌더링할 코드를 스케줄링하는 GPU 인터럽트를 사용해 디스플레이된다. 오버레이에 필요한 메모리 양은 오버레이 영역 사이즈에 좌우되며, 오버레이는 스크린 해상도에 따라 스케일링됨이 바람직하다. 동시발생적 시스템 애플리케이션에 의해 전면적 사용자 인터페이스가 사용되는 경우, 애플리케이션 해상도와 무관한 해상도를 사용함이 바람직하다. 주파수를 변경 및 TV 재동기를 유도해야 할 필요성이 없도록 그러한 해상도를 설정하는데 스케일러가 사용될 수 있다.

[0043] 멀티미디어 콘솔(100)이 부팅하고 시스템 자원들이 예비된 뒤, 동시발생적 시스템 애플리케이션들이 시스템 기능을 제공하도록 실행된다. 시스템 기능들은 상술한 예비 시스템 자원들 안에서 실행되는 시스템 애플리케이션들의 집합 안에 캡슐화된다. 운영체제 커널이 시스템 애플리케이션 스레드 대 게임 애플리케이션 스레드인 스레드들을 식별한다. 시스템 애플리케이션들은 애플리케이션에 일관된 시스템 자원 보기를 제공하기 위해 소정 시점 및 인터벌로 CPU(101)를 실행하도록 스케줄링됨이 바람직하다. 스케줄링은 콘솔 상에서 실행되는 게임 애플리케이션에 대한 캐시 혼란을 최소화하도록 해야 한다.

[0044] 동시적 시스템 애플리케이션이 오디오를 필요로 할 때, 시간에 대한 민감도로 인해 게임 애플리케이션과는 비동기적으로 오디오 프로세싱이 스케줄링된다. 시스템 애플리케이션들이 액티브될 때, 멀티미디어 콘솔 애플리케이션 관리자(이하에 기술됨)가 게임 애플리케이션 오디오 레벨(가령, 묵음(mute), 감쇠(attenuate))를 제어한다.

[0045] 입력 장치들(가령, 제어기들(141(1) 및 142(2))이 게임 애플리케이션 및 시스템 애플리케이션들에 의해 공유된다. 입력 장치들은 예비 자원들이 아니지만, 각각이 장치의 포커스를 가질 수 있도록 시스템 애플리케이션 및 게임 애플리케이션들 사이에서 스위칭되어야 한다. 애플리케이션 관리자는 게임 애플리케이션의 지식 없이 입력 스트림의 스위칭을 제어하고 드라이버가 포커스 스위치에 관한 상태 정보를 관리한다. 콘솔(100)은 카메라들(26 및 28)을 포함하는 도 2의 깊이 카메라 시스템(20)으로부터 추가 입력들을 수신할 수 있다.

[0046] 도 4는 도 1a의 모션 캡처 시스템에서 사용될 수 있는 컴퓨팅 환경의 또 하나의 전형적 블록도를 도시한다. 컴퓨팅 환경은 하나 이상의 제스처들이나 다른 움직임들을 해석하고, 그에 대응하여 디스플레이 상의 시각적 공간을 업데이트하는 데 사용될 수 있다. 컴퓨팅 환경(220)은 통상적으로 다양한 유형의 컴퓨터 판독가능 저장 매체를 포함하는 컴퓨터(241)를 구비한다. 그것은 컴퓨터(241)에 의해 액세스될 수 있는 어떤 가용 매체일 수 있으며, 휘발성 및 비휘발성 매체 양자 모두와 착탈형 및 비착탈형 매체를 포함한다. 시스템 메모리(222)는 ROM(read only memory)(223) 및 RAM(random access memory)(260) 같은 휘발성 및/또는 비휘발성 메모리의 형식으로 된 컴퓨터 저장 매체를 포함한다. 가령 시동 중에, 컴퓨터(241) 내 구성요소들 사이에서 정보를 전달하는 것을 돕는 기본 루틴들을 포함하는 기본 입출력 시스템(224)(BIOS)은 통상적으로 ROM(223)에 저장된다. RAM(260)은 통상적으로 즉시 액세스가능하고/하거나 프로세싱 유닛(259)에 의해 현재 운영되고 있는 데이터 및/또는 프로그램 모듈들을 포함한다. 한정하는 것이 아닌 예로서, 도 4는 운영체제(225), 애플리케이션 프로그램(226), 기타 프로그램 모듈(227), 및 프로그램 데이터(228)를 예시한다.

[0047] 컴퓨터(241)는 다른 착탈형/비착탈형, 휘발성/비휘발성 컴퓨터 저장 매체, 예컨대, 비착탈형, 비휘발성 마그네틱 매체로부터/로 읽고/쓰기를 하는 하드 디스크 드라이브(238), 착탈형, 비휘발성 마그네틱 디스크(254)로부터/로 읽고/쓰기를 하는 마그네틱 디스크 드라이브(239), 및 CD ROM 또는 다른 광 매체와 같은 착탈형, 비휘발성 광 디스크(253)로부터/로 읽기/쓰기를 하는 광 디스크 드라이브(240)를 또한 포함할 수 있다. 전형적인 동작 환경에 사용될 수 있는 다른 착탈형/비착탈형, 휘발성/비휘발성 유형의 컴퓨터 판독가능 저장 매체로는 마그네틱 테이프 카세트, 플래시 메모리 카드, DVD(digital versatile disks), 디지털 비디오 테이프, 고체 상태 RAM, 고체 상태 ROM 등이 포함되나 그러한 것에 국한되지 않는다. 하드 디스크 드라이브(238)는 통상적으로 인터페이스(234)와 같은 비착탈형 메모리 인터페이스를 통해 시스템 버스(221)에 연결되고, 마그네틱 디스크 드라이브(239) 및 광 디스크 드라이브(240)는 통상적으로 인터페이스(235)와 같은 착탈형 메모리 인터페이스를 통해 시스템 버스(221)에 연결된다.

- [0048] 위에서 논의되고 도 4에서 예시된 그러한 드라이브들 및 그 관련 컴퓨터 저장 매체는 컴퓨터(241)에 컴퓨터 판독가능 명령어, 데이터 구조, 프로그램 모듈 및 기타 데이터의 저장을 지원한다. 예를 들어 하드 디스크 드라이브(238)는 운영체제(258), 애플리케이션 프로그램(257), 기타 프로그램 모듈(256), 및 프로그램 데이터(255)를 저장하는 것으로 묘사된다. 이 구성요소들은 운영체제(225), 애플리케이션 프로그램(226), 기타 프로그램 모듈(227), 및 프로그램 데이터(228)와 같을 수도 다를 수도 있다는 것을 알아야 한다. 운영체제(258), 애플리케이션 프로그램(257), 기타 프로그램 모듈(256), 및 프로그램 데이터(255)에는 여기서 최소한 그들이 다른 사본들임을 묘사하기 위해 다른 부호들이 주어진다. 사용자는 키보드(251) 및 전형적으로는 마우스, 트랙볼 또는 터치 패드라 일컫는 포인팅 장치(252)와 같은 입력 장치들을 통해 컴퓨터(241) 안에 명령과 정보를 입력할 수 있다. 다른 입력 장치들(미도시)에는 마이크로폰, 조이스틱, 게임 패드, 위성 접시, 스캐너 등이 포함될 수 있다. 이들 및 기타 입력 장치들은 흔히 시스템 버스에 연결된 사용자 입력 인터페이스(236)를 통해 프로세싱 유닛(259)에 연결되나, 병렬 포트나 게임 포트 또는 유니버설 시리얼 버스(USB) 같은 다른 인터페이스 및 버스 구조에 의해 연결될 수도 있다. 카메라들(26 및 28)을 포함하는 도 2의 깊이 카메라 시스템(20)은 콘솔(100) 콘솔(100)에 대한 추가 입력 장치들을 규정할 수 있다. 모니터(242) 또는 다른 유형의 디스플레이 장치 역시 비디오 인터페이스(232)와 같은 인터페이스를 통해 시스템 버스(221)에 연결된다. 모니터 외에, 컴퓨터들은 출력 주변기기 인터페이스(233)를 통해 연결될 수 있는 스피커(244) 및 프린터(243)와 같은 다른 주변기기 출력 장치들을 포함할 수도 있다.
- [0049] 컴퓨터(241)는 원격 컴퓨터(246)와 같은 한 개 이상의 원격 컴퓨터로의 로직 연결을 이용하여, 네트워킹 환경 안에서 동작할 수 있다. 원격 컴퓨터(246)는 퍼스널 컴퓨터, 서버, 라우터, 네트워크 PC, 피어 장치 또는 다른 전형적 네트워크 노드일 수 있으며, 도 4에는 메모리 저장 장치(247) 만이 도시되었지만 통상적으로 컴퓨터(241)와 관련해 위에서 기술된 구성요소들 중 다수나 전부를 포함한다. 논리 접속들은 LAN(local area network)(245) 및 WAN(wide area network)(249)을 포함하지만, 다른 네트워크들 역시 포함할 수 있다. 그러한 네트워킹 환경은 사무소, 기업 전체의 컴퓨터 네트워크, 인트라넷 및 인터넷에서 일반적이다.
- [0050] LAN 네트워킹 환경에서 사용될 때, 컴퓨터(241)는 네트워크 인터페이스나 어댑터(237)를 통해 LAN(245)에 연결된다. WAN 네트워킹 환경에서 사용될 때, 컴퓨터(241)는 통상적으로, 인터넷 같은 WAN(249)을 통한 통신을 설정하기 위한 모뎀(250) 또는 다른 수단을 포함한다. 내장형 또는 외장형일 수 있는 모뎀(250)이 사용자 입력 인터페이스(236)나 다른 알맞은 메커니즘을 통해 시스템 버스(221)에 연결될 수 있다. 네트워킹 환경에서, 컴퓨터(241)와 관련해 묘사된 프로그램 모듈들이나 그 일부는 원격 메모리 저장 장치에 저장될 수 있다. 한정하는 것이 아닌 예로서, 도 4는 원격 애플리케이션 프로그램(248)을 메모리 장치(247) 상에 상주하는 것으로서 묘사한다. 도시된 네트워크 접속은 본보기적인 것이며 컴퓨터들 사이에 통신 링크를 설정하는 다른 수단 역시 사용될 수 있다는 것을 알 수 있을 것이다.
- [0051] 도 5는 사람이 모션 캡처 시스템과 상호작용할 수 있게 하는 방법을 도시한다. 단계(500)는 모션 캡처 시스템의 시계 안에서 어떤 사람의 움직임과 위치를 추적하는 것을 포함한다. 단계(500)의 보다 세부적인 내용은 도 6a-c와 관련하여 제공된다. 결정 단계(502)는 그 사람이 애플리케이션에 참가하고자 의도하는지를 판단한다. 단계(502)의 보다 세부적인 내용은 도 7a-d 및 10a와 관련하여 제공된다. 일반적으로 단계(502)는 한 개 이상의 변수들에 기초하여 어떤 사람이 특정 의도를 가질 가능성을 판단하고, 그 가능성이 충분히 높을 때 그에 따른 액션을 일으킬 수 있다.
- [0052] 그 사람이 참가하고자 하는 의도가 없으면, 단계 500의 추적이 계속된다. 결정 단계(502)가 참이면, 단계(504)는 그 사람이 애플리케이션에 참가하는 것을 돕는다. 단계(504)의 보다 세부적인 내용은 도 8a, 8b 및 10b-10f와 관련하여 제공된다. 단계(506)는 애플리케이션 참가 중에 그 사람을 모니터링하는 일을 포함한다. 단계(506)의 보다 세부적인 내용은 도 9a-9d와 관련하여 제공된다.
- [0053] 예를 들어 도 5의 방법은 어떤 사람이 최초로 시계 안에 진입할 때와 같은 최초 시점에는 그 사람이 애플리케이션에 참가하고자 하는 의도가 없지만, 그 사람이 몇 초 간 시계 안에 있었을 때와 같이 이후 두 번째 시점에서 애플리케이션에 참가하고자 한다고 판단할 수 있다. 그 사람이 최초 시계 진입시 자신의 마음 속에 참가하고자 하는 의도를 형성할 수는 있지만, 그 사람이 아직 충분히 긴 시간 동안 관찰되지 않았기 때문에 모션 캡처 시스템이 그 의도를 아직 판단하지 않았다는 것을 알아야 한다. 애플리케이션이 최초에 비활성 모드 중에 있다면, 사람이 그 애플리케이션에 참가하고자 한다는 판단에 따라 두 번째에는 액티브 모드로 세팅될 수 있다. 비활성 모드에서 애플리케이션의 일반적인 시각적이거나 청각적인 출력은 예컨대, 빈 스크린이나 스크린 세이버 디스플레이를 제공함으로써 억제될 수 있고, 액티브 모드 중에는 일반적인 시각적이거나 청각적인 출력이 제공될 수

있다.

- [0054] 도 6a는 도 5의 단계(500)에 기술된 것과 같이 사람의 움직임 추적하는 전형적 방법을 도시한다. 이 전형적인 방법은 도 2-4와 관련하여 논의된 바와 같이, 예컨대 깊이 카메라 시스템(20) 및/또는 컴퓨팅 환경(12, 100 또는 220)을 이용하여 구현될 수 있다. 사람의 골격 모델, 메쉬(mesh) 인체 모델, 또는 어떤 다른 적절한 표현과 같은 모델을 생성하기 위해 한 명 이상의 사람들이 스캐닝될 수 있다. 그런 다음 컴퓨팅 환경에 의해 실행되는 애플리케이션과 상호작용하는데 그 모델이 사용될 수 있다. 모델을 생성하기 위한 스캐닝은 애플리케이션이 시작 또는 시동될 때 일어날 수 있고, 그렇지 않으면 스캐닝되는 사람에 대해 애플리케이션이 제어하는 다른 시간대에 일어날 수 있다.
- [0055] 사람(58)의 물리적 움직임이나 모션이 애플리케이션의 파라미터들을 조정 및/또는 제어하는 실시간 사용자 인터페이스로서 작용할 수 있도록 추적될 수 있는 골격 모델을 생성하기 위해 사람이 스캐닝될 수 있다. 예를 들어, 추적된 사람의 움직임들이 사용되어 전자적 롤 플레이(role-playing) 게임에서 아바타나 다른 온 스크린 캐릭터를 움직이거나; 전자 레이싱 게임에서 온 스크린 자동차를 제어하거나; 가상 환경에서 객체들의 생성이나 구성을 제어하거나; 애플리케이션의 어떤 다른 적절한 제어를 수행할 수 있다.
- [0056] 일 실시예에 따르면, 단계(600)에서 깊이 정보가 예컨대 깊이 카메라 시스템으로부터 수신된다. 깊이 카메라 시스템은 한 개 이상의 타깃들을 포함할 수 있는 시계를 캡처 또는 관찰할 수 있다. 한 전형적인 실시예에서, 깊이 카메라 시스템은 논의된 것과 같은 TOF 분석, 구조화된 광 분석, 스테레오 비전 분석 등과 같은 어떤 적절한 기법을 이용하여 캡처 영역 내 한 개 이상의 타깃들과 관련된 깊이 정보를 획득할 수 있다. 깊이 정보는 복수의 관찰 픽셀들을 가지는 깊이 이미지를 포함할 수 있으며, 각각의 관찰 픽셀은 논의된 바와 같이 깊이 값을 가진다.
- [0057] 깊이 이미지는 보다 적은 컴퓨팅 오버헤드를 가지고 보다 쉽게 사용 및 처리될 수 있도록 보다 낮은 프로세싱 해상도로 다운샘플링될 수 있다. 또한, 한 개 이상의 고 변동 및/또는 요란한 깊이 값들이 깊이 이미지로부터 제거 및/또는 완화될 수 있고; 누락 및/또는 제거된 깊이 정보의 일부가 채워지고/거나 재구성될 수 있고; 도 6b 및 6c와 관련하여 논의된 골격 모델과 같은 모델을 생성하기 위해 깊이 정보가 사용될 수 있도록 수신된 깊이 정보 상에서 어떤 다른 적절한 처리가 수행될 수 있다.
- [0058] 결정 단계(604)에서, 깊이 이미지가 인간 타깃을 포함하는지 여부에 대한 판단이 이루어진다. 이것은 깊이 이미지가 인간 타깃을 포함하는지 여부를 판단하기 위해 각각의 타깃이나 객체를 어떤 패턴과 비교하는, 깊이 이미지 내 각각의 타깃이나 객체의 플러드 필링(flood filling)을 포함할 수 있다. 예를 들어, 위에서 기술한 바와 같이 타깃들이나 객체들을 규정할 수 있는 에지(edge)들을 판단하기 위해 깊이 이미지의 선택 영역이나 선택 포인트 안에 있는 픽셀들의 다양한 깊이 값들이 비교될 수 있다. Z 계층들의 있을 법한 Z 값들이 상기 판단된 에지들에 기반하여 플러드 필링될 수 있다. 예를 들어, 판단된 에지들과 관련된 픽셀들 및 에지들 안에 있는 영역의 픽셀들이 서로와 연관되어, 어떤 패턴과 비교될 수 있는 캡처 영역 안의 타깃이나 객체를 규정하도록 할 수 있으며, 이것은 이하에서 보다 상세히 설명될 것이다.
- [0059] 결정 단계(604)가 참이면, 단계(606)이 수행된다. 결정 단계(604)가 거짓이면, 단계(600)에서 추가 깊이 정보가 수신된다.
- [0060] 각각의 타깃이나 객체와 비교되는 패턴은 사람의 통상적 신체를 집합적으로 규정하는 변수들의 집합을 가진 한 개 이상의 데이터 구조들을 포함할 수 있다. 인간 타깃을 식별하기 위해, 예컨대 시계 안에 있는 어떤 인간 타깃 및 비인간 타깃의 픽셀들과 관련된 정보가 그 변수들과 비교될 수 있다. 일 실시예에서, 상기 집합 내 변수들 각각은 신체 부분에 기초해 가중될 수 있다. 예를 들어, 패턴 안에서 머리 및/또는 어깨와 같은 여러 신체 부분들은 다리와 같은 다른 신체 부분들보다 더 클 수 있는 가중치와 연관되어 있을 수 있다. 일 실시예에 따르면, 가중치들은 타깃이 인간인지 여부 및 타깃들 중 어느 것이 인간인지를 판단하기 위해 그 타깃을 변수들과 비교할 때 사용될 수 있다. 예를 들어 변수들과 타깃 사이에서 보다 큰 가중치들을 가질 수 있는 매치들이 보다 작은 가중치들을 가진 매치들보다, 타깃이 인간일 보다 큰 가능성을 낳을 수 있다.
- [0061] 단계(606)은 신체 부분들에 대한 인간 타깃의 스캐닝을 포함한다. 인간 타깃이 스캐닝되어, 사람의 정확한 모델을 제공하기 위해 그 사람의 하나 이상의 신체 부분들과 관련된 길이, 너비 등과 같은 측정치들을 제공할 수 있다. 한 전형적인 실시예에서, 인간 타깃이 분리될 수 있고, 한 개 이상의 신체 부분들에 대해 스캔하기 위해 그 인간 타깃의 비트마스크(bitmask)가 생성될 수 있다. 비트마스크는 예컨대, 인간 타깃이 캡처 영역 요소들 안에서 다른 타깃들이나 객체들과 분리될 수 있도록 그 인간 타깃을 플러드 필링함으로써 생성될 수 있다. 그

런 다음 인간 타깃의 골격 모델, 메쉬 모델 등과 같은 모델을 생성하기 위해 한 개 이상의 신체 부분들에 대하여 비트마스크가 분석될 수 있다. 예를 들어, 일 실시예에 따르면 스캐닝된 비트마스크에 의해 결정된 측정값들은 도 6b 및 6c와 관련하여 논의된 골격 모델의 한 개 이상의 관절들을 규정하는 데 사용될 수 있다. 한 개 이상의 관절들은 인간의 신체 부분에 대응할 수 있는 한 개 이상의 뼈대를 정의하는 데 사용될 수 있다.

[0062] 예를 들어, 인간 타깃의 비트마스크의 최상부는 머리의 최상부 위치와 연관될 수 있다. 머리의 상부를 결정한 후, 비트마스크가 아래 방향으로 스캐닝되어 목의 위치, 어깨의 위치 등을 결정하도록 할 수 있다. 가령 스캐닝되고 있는 위치에서 비트마스크의 너비는 예컨대 목, 어깨 등과 연관된 통상적 너비의 문턱치와 비교될 수 있다. 다른 대안적 실시예에서, 스캐닝되어 비트마스크 내 어느 신체 부분과 연관된 이전 위치로부터의 거리, 목, 어깨 등의 위치를 결정하는 데 사용될 수 있다. 다리, 발 등과 같은 일부 신체 부분들은 예컨대 다른 신체 부분들의 위치에 기반하여 산출될 수 있다. 어떤 신체 부분의 값들을 결정할 때, 그 신체 부분의 측정치들을 포함하는 데이터 구조가 생성된다. 데이터 구조는 깊이 카메라 시스템에 의해 여러 시간 상의 시점들에서 제공된 여러 깊이 이미지들로부터 평균을 낸 스캔 결과들을 포함할 수 있다.

[0063] 단계(608)은 인간 타깃의 모델을 생성하는 것을 포함한다. 일 실시예에서, 스캐닝된 비트마스크에 의해 결정된 측정값들은 골격 모델의 한 개 이상의 관절들을 규정하는 데 사용될 수 있다. 한 개 이상의 관절들은 인간의 신체 부분에 대응할 수 있는 한 개 이상의 뼈대를 정의하는 데 사용된다. 예를 들어 도 6b는 도 6a의 단계(608)에 기술된 것과 같은 사람의 전형적 모델(620)을 묘사하고, 도 6c는 도 6a의 단계(608)에 기술된 것과 같은 사람의 또 다른 전형적 모델(630)을 묘사한다.

[0064] 일반적으로, 각각의 신체 부분은 골격 모델의 관절들과 뼈대들을 규정하는 수학적 벡터로서 특징지을 수 있다. 신체 부분들은 관절들에서 다른 부분과 관련하여 움직일 수 있다. 예를 들어, 팔뚝 부분(628)은 관절들(626 및 629)와 연결되고 상완 부분(624)은 관절들(622 및 626)과 연결된다. 팔뚝 부분(628)은 상완 부분(624)과 관련하여 움직일 수 있다.

[0065] 한 개 이상의 관절들은 그들이 사람의 한 관절 및 신체 부분 사이의 통상적 거리들의 범위 안에 있게 될 때까지 조정되어, 보다 정확한 골격 모델을 만들 수 있도록 할 수 있다. 모델은 예컨대 인간 타깃의 키에 기반하여 추가로 조정될 수 있다.

[0066] 단계(610)에서, 초 당 수차례 그 사람의 위치를 업데이트함으로써 모델이 추적된다. 사용자가 물리적 공간 안에서 움직이면서, 깊이 카메라 시스템으로부터의 정보가 사용되어 골격 모델이 사람을 표현하도록 골격 모델을 조정한다. 특히, 하나 이상의 힘이 골격 모델의 하나 이상의 힘을 받는 부분들로 인가되어, 그 골격 모델이 물리적 공간 내 인간 타깃의 포즈에 보다 가깝게 대응하는 포즈를 취하도록 조정할 수 있다.

[0067] 일반적으로, 사람의 움직임을 추적하는 어떤 알려진 기법이 사용될 수 있다.

[0068] 도 7a는 도 5의 단계(502)에 기술된 것과 같이 사람이 어떤 애플리케이션에 참가하고자 의도하는지를 판단하는 전형적 방법을 도시한다. 사람의 의도를 판단할 때 다양한 요인들이 고려될 수 있다. 그 요인들 중 한 개 이상이 이용될 수 있다. 또한, 제공된 요인들은 애플리케이션에 참가하고자 하는 의도나, 관련되었지만 상반된 개념들로서 참가하지 않고자 하는(또는 참가 해제하고자 하는) 의도를 가리킬 수 있다. 예를 들어, 낮은 가능성의 참가하고자 하는 의도는 높은 가능성의 참가하지 않고자 하는 의도와 균등할 수 있다. 그 사람의 의도에 대한 보다 나은 파악을 하기 위해 그 요인들이 예컨대, 수 초와 같은 일정 시간 주기 동안 초 당 수 차례, 빈번하게 평가될 수 있다.

[0069] 단계(700)은 가령 방 안과 같은 물리적 공간 안에서 시계 안의 사람의 위치를 평가한다. 예를 들어, 도 7b는 소정 위치에 대한 어떤 사람의 위치가 평가되는 시계의 예를 도시한다. 포인트(700)는 깊이 카메라를 나타내고, 라인(722 및 724)은 시계(726)의 경계를 나타낸다. 시계가 통상적으로 광학 축에 대해 대칭적인 경우, 깊이 카메라의 광학 축 또는 z 축, 예컨대 깊이 카메라가 물리적 공간을 보는 방향의 중심으로 뻗어나간 라인을 포함하는 좌표 시스템이 규정될 수 있다. 깊이 카메라가 디스플레이와 동일한 방향을 향하는 경우, 광학 축은 일반적으로 디스플레이 평면에 수직이 된다. z 축에 직교하는 교차 축(y 축)이 정의될 수 있다. 제1준(729)은 원형으로 z 축을 따라 반경 r1, 주변둘레(728), 및 중심점(727)을 가진다. 제2준(731)은 고리 또는 링 모양이고, 내부 반경 r1, 외부 반경 r2, 안쪽 둘레(728), 바깥쪽 둘레(730)를 가지고, 제1준(729)과 같은 중심을 가지며 제1준에 인접한다. 원형 및 동심 준들의 사용은 단지 예일 뿐이며, 타원형, 직사각형, 정사각형 등의 다른 준 모양들도 사용될 수 있다. 중심점(734)을 가진 원(732)은 애플리케이션에 참가하고자 하는 의도가 평가되고 있는 사람을 포함하는 바운딩 실린더를 나타낸다. 그 사람의 중심점(734)은 중심점(727)로부터 거리

r3이다. 바운딩 실린더에 대한 대안으로서 바운딩 박스 또는 직사각형이 사용될 수 있다.

- [0070] 예를 들어, 사람이 어떤 위치, 가령 애플리케이션과 상호작용하려는 의도의 가능성이 높은 제1존(729) 안이나, 가령 애플리케이션과 상호작용하려는 의도의 가능성이 중간 정도인 제2존(731) 안이나, 존들(729 및 731) 밖에 있으면서 애플리케이션과 상호작용하려는 의도의 가능성이 낮은 시계 안의 한 위치에 있을 수 있다. 애플리케이션과 상호작용하고자 하는 의도의 가능성이 가장 높은 위치, 예컨대 존(729)에 대한 사람의 위치에 기반하여 위치 기반의 가능성 평가가 이루어질 수 있다.
- [0071] 사람의 위치는 시계 안에서 현재 애플리케이션과 연관된 다른 사람의 위치를 기준으로 평가될 수도 있다. 예를 들어, 애플리케이션에 참가하고자 의도한 어떤 사람이 가령 게임의 잠정적 팀 메이트나 경쟁자로서 이미 그 애플리케이션에 참가한 다른 사람 옆에 서 있을 수 있다. 그 사람이 시계 안에서 애플리케이션에 참가하고 있는 다른 사람의 옆에 있는지, 그 다른 사람과 특정 거리 안에 있는지에 대한 판단이 이루어질 수 있다. 예를 들어, 도 7c는 다른 사람에 대해 상대적인 어떤 사람의 위치가 평가되는 시계의 예를 묘사한다. 여기서, 중심점(738)을 가진 원(736)은 현재 애플리케이션에 참가하고 있는 사람을 포함하는 바운딩 실린더를 나타내고, 중심점(734)을 가진 원(732)은 애플리케이션에 참가하고자 하는 의도가 평가되고 있는 사람을 포함하는 바운딩 실린더를 나타낸다. 축(744)은 y 축과 평행하며 중심점(736)을 통과한다. 라인들(740 및 742)은 축(744)을 중심으로 +/- 각도로 중심점(738)을 통과한다. 영역(746)은 어떤 사람이 원(736)으로 나타낸 사람의 옆이나 바로 옆에 있다고 간주되는 라인들(740 및 742) 사이라고 정의될 수 있다. 비슷한 영역이 그 사람(736)의 또 다른 측면 상에 정의될 수 있다. d1은 중심점(738)에서 중심점(734)까지의 거리이며, 원들(736 및 732)로 묘사된 각자의 사람들 사이의 거리를 나타낸다. 대안으로서 거리가 원들(736 및 732)의 외주들 사이의 최단 거리로 정의될 수도 있다.
- [0072] 여기서, 참가하고자 하는 의도의 보다 높은 가능성은 원(732)으로 나타낸 사람이 영역(746) 안에 있지 않을 때 보다 원(732)으로 나타낸 사람이 영역(746) 안(원(736)으로 나타낸 사람의 측면)에 있을 때와 관련될 수 있다. 마찬가지로, 참가하고자 하는 의도의 보다 높은 가능성은 원(732)으로 나타낸 사람이 원(738)으로 나타낸 사람으로부터 영역(746) 안에서 상대적으로 먼 거리에 있을 때보다는 원(732)으로 나타낸 사람이 원(736)으로 나타낸 사람과 영역(746) 안에서 상대적으로 가까운 거리에 있을 때와 관련될 수 있다.
- [0073] 단계(702)는 사람의 자세나 포즈를 평가한다. 예를 들어 사람의 신체가 깊이 카메라를 마주하면서 깊이 카메라 쪽으로 어깨를 펴고 있으면(가령, 도 10b의 사람들(1022 및 1024)), 그 사람의 몸이 카메라를 마주하지 않으면서 어깨를 다른 방향으로 펴고 있을 때(가령, 도 10a의 사람(1026 또는 1028)) 보다 그 사람이 애플리케이션에 참가하고자 의도할 가능성이 더 많다. 위에서 후자의 경우, 깊이 카메라는 사람의 측면을 보게 된다. 사람이 앉아 있거나(도 10a의 사람(1028)) 누워 있는 포즈도 마찬가지로 애플리케이션에 참가하고자 할 의도를 나타내지 않을 가능성이 많다. 마찬가지로, 가슴에 팔짱을 낀 자세는 참가할 의도가 없다고 신호할 수 있다.
- [0074] 애플리케이션 상황에 대해 특정되는 참가 의도의 예로서, 요가 수행과 관련된 애플리케이션을 고려할 수 있으며, 이 경우 앉아 있는 자세는 벗어날 의도가 아니라 참가할 의도를 신호하는 것일 수 있다. 또한, 참가나 참가 해제 의도를 판단하는 기준은 나이, 성별, 신체 건강 수준이나 장애와 같은 요인들에 기반하여 조정될 수 있을 것이다. 예를 들어, 느린 사람의 움직임은 참가하고자 하는 의도일 수 있는 반면 빠른 사람의 움직임은 그렇지 않을 수 있다.
- [0075] 특정 자세가 얼마나 오래 유지되는지를 판단하기 위해 타이머들이 사용될 수 있다. 보다 긴 지속시간이 참가 또는 참가 해체에 대한 보다 큰 의도와 연관될 수 있다.
- [0076] 애플리케이션과 상호작용하고자 하는 의도의 가능성이 가장 높은 자세에 대해 상대적인 사람의 자세에 기초하여 자세 기반 가능성 평가가 이루어질 수 있다.
- [0077] 단계(704)는 시계 안에서의 사람의 움직임을 평가한다. 예를 들어, 사람이 시계를 지나 빠르게 이동하고 있으면(가령, 도 10a의 시계(1021) 안의 사람(1026)이나 도 7d의 원(732)으로 표현된 사람의 화살표 방향(750)으로 표현됨), 그 사람은 아마도 시계를 그냥 통과해 걸어가는 것이고 애플리케이션에 참가하고자 의도하는 것이 아닐 것이다. 또는, 사람이 디스플레이로부터(도 7d의 원(732)으로 표현된 사람에 대한 화살표(752)의 방향으로) 멀리 이동하고 있으면, 사용자는 아마도 애플리케이션에 참가하고자 의도하는 것이 아닐 것이다. 또한, 사람이 일반적으로 한 위치에 있으면서 자신의 팔을 펼칠 때(도 10a의 사람(1024)으로 표현됨)나 한 발에서 다른 발로 무게중심을 이동할 때와 같이 움직이는 것 같으면, 이것이 애플리케이션에의 보다 높은 가능성의 참가 의도와 관련되어 있을 수 있다. 마찬가지로, 사람이 일반적으로 한 위치에 있으면서 완전히 정지 상태인 것으로 보여

지만, 이것은 애플리케이션으로의 보다 낮은 가능성의 의도와 관련되어 있을 수 있다. 또한, 움직임과 위치 요인들이 관련될 수 있다, 예컨대 참가 의도에 대한 보다 높은 가능성을 가진 위치를 향한 움직임이 참가 의도에 대한 보다 높은 가능성의 표시일 수 있다.

- [0078] 사람이 취하는 제스처 또한 움직임의 한 형태로서 간주될 수 있다. 예를 들어 사용자가 자신의 팔을 들어 올리고 깊이 카메라를 향해 흔들어 그 사람이 참가하고자 한다는 것을 가리킬 수 있다. 또한, 의도와 관련된 어떤 제스처가 애플리케이션 상황에 특정한 것일 수 있다, 예컨대 상황에 민감한 제스처들일 수 있다. 예를 들어, 참가나 참가 해제 의도를 신호하는 것으로서 미리 규정되는 제스처들이 있을 수 있다. 움직임(가령, 사람의 팔을 옆으로 올리고 그런 다음 팔을 내리는 것을 검출한 경우)이나 정적인 위치(가령, 사람의 팔이 옆으로 올려진 것을 검출한 경우)에 기반하여 제스처가 식별될 수 있다.
- [0079] 애플리케이션과 상호작용하고자 하는 의도의 가능성이 가장 높은 움직임에 대해 상대적인 사람의 움직임에 기초하여 움직임 기반 가능성 평가가 이루어질 수 있다.
- [0080] 단계(706)은 사람의 음성 데이터를 평가한다. 일반적으로 사람의 음성은 소리 낸 말, 휘파람, 외치기 및 다른 발성들을 포함할 수 있다. 손뼉치기와 같은 비음성적 소리들 역시 검출될 수 있다. 소리가 검출되는 방향을 검출하고 그것을 검출된 사람의 위치와 상관시켜 그 사람이 애플리케이션에 참가하고자 의도할 가능성에 대해 훨씬 더 안정적인 평가를 제공하도록, 방향성 마이크로폰이 옵션으로서 사용될 수 있다. 사람이 시계 밖에 있을 때에도 그 사람의 의도를 판단하려면 음성 데이터만으로도 충분할 수 있다.
- [0081] 음성 데이터의 존재는 어떤 사람이 애플리케이션에 참가하고자 의도하는 높은 가능성과 상관될 수 있다. 또한 음성 데이터의 볼륨이나 크기가 어떤 사람이 애플리케이션에 참가하고자 의도하는 높은 가능성과 상관될 수 있다. 또한 "시작"이나 "준비"와 같은 명령들이 애플리케이션에의 참가 의도를 가리키도록, 음성이 검출될 수 있다. 그에 따라, 사람이 애플리케이션에 참가하고자 한다는 판단은 사람의 음성 데이터를 처리하는 일과 애플리케이션에의 참가 의도를 나타내는 음성을 검출하는 일 및/또는 애플리케이션에의 참가 의도를 나타내는 음성 볼륨을 검출하는 일을 포함할 수 있다.
- [0082] 사소한 게임 쇼 애플리케이션의 예에서, 플레이어들은 보통 상대적으로 정적이고, 아마도 소파에 앉아 있으면서 질문에 답한다. 이 경우, 새로운 음성의 존재 검출은 새로운 사람에 의한 참가 의도로서 평가될 수 있다. 마찬가지로, 어떤 사람의 앞서 인식된 음성이 소정 시간 동안 부재하는 것은 그 현재 침묵하는 사람에 의해 참가 해제의 의도로서 평가될 수 있다.
- [0083] 애플리케이션과 상호작용하고자 하는 의도의 가능성이 가장 높은 음성 데이터에 대해 상대적인 사람의 음성 데이터에 기초하여 음성 기반 가능성 평가가 이루어질 수 있다.
- [0084] 일단 한 사람이 애플리케이션에 참가하고자 한다고 판단되면, 다음에 논의되는 바와 같이 그러한 참가가 촉진된다. 어떤 사람이 애플리케이션에 참가하고자 한다고 판단될 때, 또는 참가 해제 시와 같은 다른 적절한 시점에, 상태 정보가 피드백으로서 제공될 수 있다는 것을 알아야 한다. 예를 들어, 어떤 사람이 애플리케이션에 참가할 의도를 가지고 있다고 판단될 때 도어 벨 소리와 같은 오디오 메시지가 애플리케이션에 의해 생성될 수 있고, 어떤 사람이 애플리케이션으로부터 참가 해제할 의도를 가지고 있다고 판단될 때 문이 닫히는 소리가 애플리케이션에 의해 만들어질 수 있다.
- [0085] 아이콘과 같은 시각적 메시지 역시, 디스플레이의 방해받지 않는 쪽이나 하부 영역과 같은 디스플레이 상에 주어질 수 있다. 시각적 메시지는 현재 참가한 사람들의 수뿐 아니라, 어떤 사람의 참가가 정지상태에 있을 때(도 9a 참조)나 사람이 참가 해제할 때(도 9b 참조)와 같은 상태 변화를 나타낼 수 있다. 가상 공간에 참여할 때 새로운 아바타가 페이드 인되고 떠날 때 아바타가 페이드 아웃되도록 시각적 효과들이 제공될 수 있다. 사람들이 언제 애플리케이션에 구속되고 언제 구속 해제되는지를 그들이 이해하도록 명확한 피드백이 주어진다. 모든 플레이어들이 참가를 해제하는 경우, 애플리케이션이 정지되거나 비활성화될 수 있다.
- [0086] 도 8a는 도 5의 단계(504)에 기술된 것과 같이 사람이 애플리케이션 참가를 촉진시키는 전형적 방법을 도시한다. 단계(800)은 아바타 및 프로파일 데이터를 자동으로 생성하는 일을 포함한다. 단계(802)는 프로파일과 아바타를 사람과 자동으로 연관시키는 일을 포함한다. 단계(804)는 디스플레이 상에 아바타를 디스플레이 하는 일을 포함하고, 단계(806)은 사람의 후속 추적에 기반하여 가령 실시간으로 디스플레이를 업데이트하는 일을 포함한다. 사람이 현재 애플리케이션에 참가하고 있는 유일한 사람이면, 그 애플리케이션은 스탠바이 모드에서, 사용자의 움직임에 기반하는 입력들을 수신할 준비가 된 레디(ready) 모드로 전환될 수 있다.
- [0087] 사람이 애플리케이션에 참가하기 위해 어떤 유형의 수동적 설정 프로세스를 거쳐야 할 필요는 없으며, 그에 따

라 지연 없이 "그냥 걸어 와서 플레이하는" 접근 방식으로 애플리케이션 참가를 시작할 수 있다. 키보드, 마우스, 또는 핸드헬드 제어기와 같은 하드웨어가 참가 시작을 위해 필요로 된다.

[0088] 자동으로 생성되어 사람과 연관된 아바타는 키, 너비, 골격 모델, 그 사람의 옷가지로부터의 RGB 컬러 등과 같이 검출된 그 사람의 특징들, 및 옵션으로서 음성 톤과 볼륨과 같은 청각적 특징들에 기반할 수 있다. 아니면 디폴트의 일반적 아바타가 제공될 수 있다. 사람이 적절한 사용자 인터페이스를 사용하여 자동 생성된 아바타를 수정하는 것 역시 가능하다. 아바타를 설정하기 위해 사람이 어떤 유형의 수동적 설정 프로세스를 거쳐야 할 필요는 없고, 그에 따라 지연 없이 아바타를 이용하여 애플리케이션 참가를 시작할 수 있다. 일부 애플리케이션들에서, 아바타는 가상 공간을 탐험하면서 그 외관이나 다른 특징들 면에서 변경될 수 있다. 아바타의 특징들은 가상 공간에서 얻어지거나 상실되는 능력들(가령, "파워")과도 관련될 수 있다. 어떤 경우, 사람이 실제 또는 게임 머니를 써서 아바타의 특징들을 구매하고, 가상 공간 안에서 아바타 액션에 의해 특징들을 따내거나 잃는다. 프로파일 데이터는 또한 아바타의 특징들이나 다른 특색들을 포함할 수 있다. 어떤 사람이 애플리케이션 참가를 중지하고 나중에 돌아올 때, 프로파일 데이터가 액세스되어 그 사람과 그의 아바타를 그 사람이 참가를 해제하였을 때 존재했던 상태로 복구시킬 수 있다.

[0089] 생성된 프로파일은 키, 너비, 골격 모델, 그 사람의 옷가지로부터의 RGB 컬러 등과 같이 검출된 그 사람의 특징들, 및 옵션으로서 음성 톤과 볼륨과 같은 청각적 특징들을 포함할 수 있다. 사람은 자신의 특징들을 애플리케이션에 연관시킴으로써 애플리케이션에 구속될 수 있다. 일반적으로, 어떤 사람이 시계 안으로 진입할 때 골격 데이터와 같은 특징들이 그 사람과 연관될 수 있다. 따라서 프로파일 데이터는 깊이 카메라에 의해 정해진 물리적 특징들을 포함하여, 그 사람의 신분증명을 포함할 수 있다.

[0090] 프로파일 데이터는 또한, 애플리케이션과 그 사람의 참가를 나타내는 정보를 포함할 수 있다. 예를 들어 게임 애플리케이션에서, 프로파일 데이터는 그 사람이 달성했던 점수나 진척상태, 및 그 사람의 아바타의 가상 공간 내 위치를 나타낼 수 있다. 가상 공간은 컴퓨터 모니터, 텔레비전 스크린과 같은 디스플레이 상에 제공되거나 벽에 투영된다. 가상 공간은 어떤 실체나 허상, 실내나 실외 위치를 나타낼 수 있다. 가상 공간은 그 사람을 나타내고 그 사람의 추적된 움직임들에 기초하여 움직이는 아바타를 포함할 수 있다. 아바타는 사용자를 실사 방식으로 묘사할 수 있고, 아니면 동물, 차량 또는 다른 캐릭터나 객체가 될 수 있다.

[0091] 옵션으로서, 네트워킹 접근방식에 따라 여러 사람들이 여러 위치들로부터 동시에 애플리케이션에 참가한다. 예를 들어, 제1 및 제2친구들이 자신들의 집에서 각기 별개의 제1 및 제2모션 캡처 시스템을 가질 수 있다. 그들은 동시에 가상 공간 안에서 상호작용을 시작하도록 약속할 수 있다. 제1인이 자신의 몸을 움직여 시계 안에서 제1아바타를 제어함으로써 애플리케이션에 참가할 때, 제1모션 캡처 시스템은 제1모션 캡처 시스템의 시계 안에 있는 제1인의 몸을 추적한다. 제1모션 캡처 시스템은 또한 제2모션 캡처 시스템으로부터 네트워킹을 통해 데이터를 수신하는데, 제2모션 캡처 시스템은 제2인이 자신의 몸을 움직여 가상 공간 안에서 제2아바타를 제어함으로써 원격 위치로부터 애플리케이션에 참가할 때, 제2모션 캡처 시스템의 시계 안에 있는 제2인의 몸을 추적한다. 상기 데이터에 대한 응답으로서, 제1애플리케이션은 제1아바타와 함께 디스플레이 상의 가상 공간 안에 제2아바타를 디스플레이하고, 제2인이 자신의 몸을 움직여 제2아바타를 제어함으로써 애플리케이션에 계속 참가할 때 디스플레이를 업데이트한다. 제2모션 캡처 시스템도 마찬가지로 네트워킹을 통해 제1모션 캡처 시스템으로부터 데이터를 수신하여 제2인이 자신의 디스플레이 상에서 두 아바타들을 볼 수 있게 할 수 있다. 그에 따라, 두 사람이 동일한 시청각적 출력을 경험한다.

[0092] 아바타 및/또는 가상 공간의 디스플레이는 다음에 논의되는 것과 같이, 다른 사람이 현재 애플리케이션에 참가되어 있는지 여부에 기반하여 조정될 수 있다.

[0093] 도 8b는 또 다른 사용자가 현재 어떤 애플리케이션에 참가하고 있을 때 도 5의 단계(504)에 기술된 것과 같이 어떤 사람이 그 애플리케이션에 참가하는 것을 촉진시키게 하는 전형적 방법을 도시한다. 단계(810)에서, 또 다른 사람이 현재 어떤 애플리케이션에 참가하고 있다는 판단이 이루어지고 그의 아바타가 디스플레이 상에 나타난다. 단계(812)는 프로파일과 아바타를 새로운 사람, 예컨대 애플리케이션에 참가하고자 하는 의도를 가진다고 새롭게 인식된 사람과 연관시킨다. 새로운 사람의 아바타를 묘사하기 위해 디스플레이를 업데이트함에 있어 다양한 옵션들이 있을 수 있다. 한 접근법으로서 단계 814에서, 아바타들이 디스플레이의 동일 영역에 디스플레이되어 그들이 동일한 시점으로부터 보여지게 된다. 이하에 더 논의되는 한 예가 도 10b에 제공된다. 또 다른 접근법으로서 단계(816)에서, 아바타들이 디스플레이의 다른 영역들에 디스플레이되어 서로 다른 각자의 시점들로부터 보여질 수 있다. 이하에 더 논의되는 한 예가 도 10c에 제공된다.

[0094] 도 9a는 도 5의 단계(506)에 기술된 것과 같이 애플리케이션에 참가하고 있는 동안 사람을 모니터링하는 전형적

방법을 도시한다. 사람이 애플리케이션에 참가하고 있을 때 다양한 시나리오들을 만날 수 있다. 예를 들어, 시계 안에서 사람이 다른 사람이나 객체에 의해 가려지거나, 그 사람이 시계를 벗어나거나, 그 사람이 어떤 다른 이유로 검출되지 않음에 따라, 그 사람이 더 이상 검출되지 않을 수 있다. 그러한 시나리오들을 다루기 위한 전형적인 방법에 있어서, 결정 단계 900은 어떤 사람이 예컨대 어떤 검출 사이클 수에 해당할 수 있는 문턱 시간 주기 동안 더 이상 검출되지 않는지를 판단한다. 통상적으로 초 당 여러 검출 사이클들이 존재한다. 시험을 통해 적절한 문턱 시간 주기가 결정될 수 있다. 결정 단계(900)가 거짓이면, 단계(902)에서 그 사람은 애플리케이션에 계속 참가한다. 예를 들어 아바타가 디스플레이 상에서 유지될 수 있다.

[0095] 결정 단계(900)가 참이면, 단계(904)는 그 사람이 애플리케이션에 참가하는 것을 중단한다. 이것은 적절한 시청각적 메시지를 제공하는 일을 포함할 수 있다. 또한, 그 사람의 아바타가 가상 공간에서 제거되도록 비활성화될 수 있다. 대신 아바타가 그것이 일시적으로 비활성 상태를 나타내는 디스플레이의 별도 영역 안에 디스플레이될 수 있다. 또는 아바타가 가상 공간에 잔존할 수는 있지만 그것이 활성화되지 않아서 예컨대 움직일 수 없다는 것을 가리키는 회색 외관을 가질 수 있다. 여러 참가 인원들이 존재하고 각각의 사람에게 별개의 스크린 영역이 할당되는 경우(도 10c), 중단은 중단된 사람의 스크린 영역이 제거되도록(도 10b) 디스플레이를 업데이트하는 일을 포함할 수 있을 것이다.

[0096] 옵션으로서, 결정 단계(906)은 그 사람이 시계를 벗어나는 것이 추적되는가를 판단한다. 이것은, 예컨대 그 사람이 시계의 경계 방향으로 이동하였고 이어서 더 이상 검출될 수 없는지를 판단하기 위해 그 사람의 시계 안에서의 움직임을 추적함으로써 판단될 수 있다. 결정 단계(906)가 거짓이면, 그 사람이 아직 시계 안에 있지만 아마도 다른 사람이나 객체, 예컨대 가구에 의해 가려져서 깊이 카메라에 의해 검출되지 못할 수 있다고 추정된다. 이 경우, 단계(908)에서 그 사람은 나중에 다시 검출된다. 앞서 검출되었다가 더 이상 검출되지 않다가 다시 검출된 동일한 사람은 골격 모델과 같은 그 사람의 물리적 특징들에 기반하여 식별될 수 있다.

[0097] 결정 단계(912)는 사람이 타임 아웃되는지(이는 특정 시간, 예컨대 수 분보다 길게 참가 정지되었음을 의미)를 판단한다. 결정 단계(912)가 거짓이면, 단계(916)에서 참가 의도를 재판단하지 않고 그 사람의 참가가 계속된다. 앞서서와 같이, 적절한 시각 및/또는 청각적 메시지가 애플리케이션에 의해 제공될 수 있다. 참가의 지속은 예를 들어 아바타를 재활성화하는 일을 포함할 수 있다. 아직 참가 중인 한 명 이상의 다른 사람들이 존재하는 경우, 중단은 별도의 스크린 영역이 참가가 계속되었던 사람에게 제공되도록(도 10c) 디스플레이를 업데이트하는 일을 포함할 수 있을 것이다. 결정 단계(912)가 참이면, 그 사람이 다시 참가할 수 있게 되기 전에, 단계(914)에서 애플리케이션에 참가하고자 하는 그 사람의 의도가 새롭게 재판단된다. 그에 따라, 그 사람은 다시 참가하고자 하는 의도를 보여줄 때까지 참가 해제 상태에 있는 것으로 다뤄진다.

[0098] 결정 단계(906)가 참이면, 그 사람은 시계를 벗어났다. 이 경우, 단계(910)에서 그 사람이 나중에 다시 검출될 때, 그 사람이 다시 참가할 수 있기 전에, 단계(914)에서 애플리케이션에 참가하고자 하는 그 사람의 의도가 새롭게 재판단된다.

[0099] 도 5의 예는 어떤 사람이 최초로 시계 안에 진입할 때와 같이, 최초 시점에는 그 사람이 애플리케이션에 참가하고자 의도하지 않지만, 이 삼 초 뒤와 같이 두 번째 시점에서는 애플리케이션에 참가하고자 한다는 것을 나타낼 수 있다. 도 9b의 예에서, 프로파일과 아바타가 자동으로 저장되고 그 사람의 식별자에 인덱싱됨에 따라, 두 번째 시점 이후의 세 번째 시점에 그 사람은 애플리케이션으로부터 참가를 해제하려고 의도한다. 또한, 그 사람이 저장된 프로파일 및 아바타에 기반하여 애플리케이션에 재참가함에 따라, 세 번째 시점 이후의 네 번째 시점에 그 사람이 애플리케이션에 재참가하고자 하는 의도를 나타낸다.

[0100] 도 9b는 사람이 애플리케이션으로부터 참가를 해제하고 이후 다시 그 애플리케이션에 참가하는 경우, 도 5의 단계(506)에 기술된 것과 같이 애플리케이션에 참가하고 있는 동안 그 사람을 모니터링하는 전형적 방법을 도시한다. 이 경우, 한 개 이상의 요인들이 그 사람이 애플리케이션으로부터 참가를 해제하고자 한다는 것을 나타낼 때(단계(910)), 그 사람이 시계 안에서 현재 검출될 수 있고 애플리케이션에 현재 참가하고 있을 수 있다. 일반적으로 도 7a와 관련하여 논의된 동일한 요인들이 참가 및 참가 해제 의도를 판단하는데 사용될 수 있다. 예를 들어, 그 사람이 낮은 가능성의 참가 의도와 연관된 시계의 한 위치에 있으면, 이것은 높은 가능성의 참가 해제 의도에 해당한다. 그 사람이 낮은 가능성의 참가 의도와 연관된 자세를 가지면, 이것은 높은 가능성의 참가 해제 의도에 해당한다. 그 사람이 낮은 가능성의 참가 의도와 연관된 움직임을 가지면, 이것은 높은 가능성의 참가 해제 의도에 해당한다. 참가 해제 의도를 가리키는 움직임들은 디스플레이로부터 멀리 걸어 나가기, 디스플레이로부터 등을 돌리기(그대로 서 있거나 걸어 나가면서), 참가 의도 가능성이 가장 높은 위치로부터 걸어 나가기 및/또는 시계의 경계를 향해, 그런 다음 시계 밖으로 걸어 나가기를 포함한다. 특정 시간 동안 어떤

사람으로부터의 음성 데이터가 부족한 것 역시 높은 가능성의 참가 해제 의도와 연관될 수 있다.

- [0101] 애플리케이션 상황에 특정한 참가 해제 의도의 일례로서, 어떤 사람에 의해 물리적 공간에서 실질적 앞뒤, 전방 및 후위 움직임들을 요하는 애플리케이션을 고려할 수 있다. 이 경우, 참가 해제 의도의 기준은 그러한 보다 큰 움직임들을 고려하도록 조정될 수 있다. 실질적 움직임들을 요하지 않는 애플리케이션에서, 실질적 움직임은 참가 해제 의도의 한 징후라고 간주될 수 있을 것이다. 또는, 요가 수련과 관련된 애플리케이션에 있어서, 앉아 있는 포즈가 기대될 수 있고 참가 해제 의도의 징후로 간주되어서는 안된다.
- [0102] 단계 912는 사람의 식별자에 인덱싱된 프로파일 데이터 및 아바타 데이터를 저장한다. 식별자는 키와 너비, 및 골격 모델, 목소리와 다른 요인들에 기반하는 특정인과 관련될 수 있는 음성 데이터, 옷 색깔과 같은 컬러 데이터 등과 같이 그 사람의 물리적 특징들과 연관될 수 있다. 사람은 "잘 가"나 "정지"와 같은 인지 명령을 말할 수도 있을 것이다. 몇 시간 후, 단계 914는 그 사람이 애플리케이션에 재참가하고자 한다는 것을 판단한다. 단계 916은 저장된 프로파일 및 아바타 데이터를 액세스하고, 그에 따라 사람은 그의 경험의 계속성을 유지하기 위해 참가를 해제할 때 자신이 떠났던 동일한 장소에서 애플리케이션에 재참가할 수 있다.
- [0103] 도 9c는 플레이어들이 모션 캡처 시스템의 시계 안에서 위치를 서로 바꾼 경우, 도 5의 단계 506에 기술된 것과 같이, 그리고 도 8b의 방법에 뒤이어서 애플리케이션 참가 중에 사람을 모니터링하는 전형적인 방법을 도시한다. 결정 단계 920은 플레이어들이 시계 안에서 위치를 서로 바꾸는지를 판단한다. 그들이 위치를 서로 바꾸면, 단계 922에서 그 사람 고유의 디스플레이 영역들 역시 맞바꾸어진다. 예를 들어 도 10c에 도시된 바와 같이, 디스플레이(1004)의 영역(1005)은 사람(1022)의 아바타(1023)를 제공하고, 디스플레이(1004)의 영역(1006)은 사람(1024)의 아바타(1025)를 제공한다. 그에 따라, 디스플레이 영역들은 물리적 공간에서 그 사람들에 대응하여 위치되며, 왼쪽 디스플레이 영역은 왼쪽 사람을 위한 것이고, 오른쪽 디스플레이 영역은 오른쪽 사람을 위한 것이다. 애플리케이션과 상호작용할 때, 사람들은 위치를 서로 바꿀 수 있다, 예컨대 오른쪽에서 왼쪽으로 왼쪽에서 오른쪽으로 바꿀 수 있다. 예를 들어, 어떤 애플리케이션들은 물리적 공간 안에서 사람들에게 의한 상당한 이동을 포함하고, 그에 따라 고정 위치가 유지되지 않는다. 따라서 사람은 물리적 공간 안에서 자연스럽게 옆으로 이동할 수 있다. 어떤 경우, 한 사람이 디스플레이에 대한 다른 사람의 보기를 가로 막아서, 가로막힌 사람이 옆으로 이동하고 위치가 맞바뀌지게 된다.
- [0104] 도 10d에 묘사된 바와 같이, 사람들(1022 및 1024)이 위치를 맞바꾸어 사람(1024)이 왼쪽에 있고 사람(1022)이 오른쪽에 있다면, 그에 따라 디스플레이 영역들(1005 및 1006)이 맞바뀌어서 디스플레이 영역(1005)은 오른쪽으로 이동하고 디스플레이 영역(1006)은 왼쪽으로 이동한다.
- [0105] 일반적으로, 한 개의 디스플레이 영역으로부터 둘 이상의 디스플레이 영역들로의 변환 및 그 반대 변환이 자동적으로 제공될 수 있다. 둘이 넘는 디스플레이 영역들을 포함하는 변환 역시 가능하다. 예를 들어, 각자의 디스플레이 영역들을 가진 두 플레이어들에 제3의 플레이어가 합류될 때, 두 디스플레이 영역들에서 세 디스플레이 영역들로의 변환이 가능하다. 각각의 디스플레이 영역은 물리적 공간 내 플레이어들의 위치에 대응하여 위치될 수 있다. 또한, 디스플레이 영역들이 수직으로 쌓여, 한 영역이 다른 영역 위에 놓여지는 것이 가능하다. 예를 들어 한 플레이어는 시계 안에서 더 높은 표면 위에서 서 있음으로 인해 다른 플레이어보다 더 높을 수 있으므로, 그 더 높은 플레이어의 디스플레이 영역이 더 낮은 플레이어의 것 위에 제공되어진다.
- [0106] 디스플레이 영역들이나 맞바뀌지는 디스플레이 영역들의 너무 잦은 변경을 피하기 위해 주의가 필요할 수 있다. 한 번의 변경 뒤에, 다음 변경을 허용하기 위한 최소한의 시간이 설정될 수 있다. 어떤 경우들에서 있어서, 참가하는 사람이 다른 참가하는 사람이나 관람자에 의해 일시적으로 가려질 수 있고, 그에 따라 추적되지 못할 수 있다. 그러한 경우들에서의 연속성을 위해, 디스플레이 영역들은 가능한 한 보통때 처럼 계속 응답을 해야 할 것이다.
- [0107] 도 9d는 어떤 애플리케이션에 참가하고 있지 않은 추가된 사람에 대한 추적이 기초하여 애플리케이션의 출력이 수정되는 경우, 도 5의 단계(506)에 기술된 바와 같이 그 애플리케이션에 참가 중에 어떤 사람을 모니터링하는 전형적인 방법을 도시한다. 애플리케이션에 참가하고자 하는 의도가 있는지를 판단하기 위해 사람을 추적하는 것 외에, 시계 안의 다른 사람의 의도를 판단하여 애플리케이션으로 해당 입력을 제공하는 것이 있을 수 있다. 예를 들어, 한 명 이상의 사람들이 다른 사람을 응원하든가 하기 위해 그 다른 사람이 애플리케이션에 참가하는 것을 지켜보는 관람자들로서 시계 안에 있을 수 있다. 단계(930)은 애플리케이션에 현재 참가하고 있지 않은, 예컨대 디스플레이 상의 가상 공간 안에서 아바타를 제어하고 있지 않은 시계 안의 추가 인원을 추적하는 일을 포함한다. 결정 단계(932)는 그 추가 인원이 애플리케이션에 높은 관심을 가지는지를 판단한다. 이것은 도 7a와 관련하여 논의된 것들을 포함하는 다양한 요인들을 사용하여 판단될 수 있다. 또한, 관심을 가진 관람자가

수행할 수 있는 액션들과 관련된 다른 요인들 역시 검출될 수 있다. 예를 들어, 한 명 이상의 사람들이 참가 중인 사람의 뒤나 옆의 어떤 거리에서 검출될지도 모른다. 그 한 명 이상의 사람들은 서 있거나, 응원하기 위해 그들의 머리 위로 팔을 들어 올리거나, 응원하는 큰 소리를 내는 것과 같은 액션을 수행할지 모른다.

- [0108] 결정 단계(932)가 거짓이면, 단계(930)의 추적이 계속된다. 결정 단계(932)가 참이면, 단계(934)는 애플리케이션의 시각 및/또는 청각적 출력을 수정한다. 예를 들어, 사람의 아바타가 달리거나 운전하고 있는 애플리케이션에서, 사람을 응원하는 관중에 대한 적절한 반응은 즐거운 경험을 고양시키기 위해 그 아바타에게 속도 증가나 좀 더 요란하게 플레이하기와 같은 강화된 능력이나 빠른 박자의 배경 음악을 제공하는 것일수가 있다.
- [0109] 전형적인 시나리오에는 제1인이 제1인의 몸을 움직여 디스플레이 상의 가상 공간에서 아바타를 제어함으로써 애플리케이션에 참가하는 경우, 모션 캡처 시스템의 시계 안에서 제1인의 몸을 구별하는 일을 포함하여, 시계 안에서 (이미 애플리케이션에 참가하고 있는) 제1인의 움직임을 추적하는 일을 포함한다. 제1인의 움직임을 추적하면서, 단계(930)는 애플리케이션에 참가하고 있지 않아서 디스플레이 상의 가상 공간 안에서 아바타를 제어하고 있지 않은 시계 안의 적어도 한 명의 추가 인원을 추적하는 일을 수반한다. 결정 단계(932)에서 적어도 한 명의 추가 인원 추적이 그 적어도 한 명의 추가 인원의 행동에 관한 소정 기준을 만족시킨다는 것을 나타낼 때, 단계(934)에서 애플리케이션의 시각 및/또는 청각적 출력이 수정된다. 소정 기준은 적어도 한 명의 추가 인원이 애플리케이션에 대한 높은 수준의 관심을 나타내는 방식으로 움직이거나(가령, 서 있거나, 팔을 들어 올림), 애플리케이션에 대한 높은 관심을 나타내는 소리(가령, 함성, 손뼉치기)를 내는 것을 포함할 수 있다.
- [0110] 도 9e는 어떤 애플리케이션에 참가 시 한 사람이 다른 사람을 대신하는 경우, 도 5의 단계(506)에 기술된 것과 같이 애플리케이션에 참가하고 있는 동안 사람을 모니터링하는 전형적 방법을 도시한다. 단계(940)에서, 제1인이 프로파일 및 아바타를 이용하여 애플리케이션에 참가한다. 단계(942)에서, 제2인은 참가 의도를 가지고 제1인은 참가 해제 의도를 가진다. 단계(944)에서, 제2인이 제1인에 의해 사용된 것과 동일한 프로파일 및 아바타를 이용하여 애플리케이션에 참가한다. 이 경우, 가상 공간 안에서 동일한 상태의 아바타 제어가 제2인에 의해 인수된다. 예를 들어 게임 애플리케이션에서, 제1인이 도달한 같은 점수나 진척 수준이 제2인에게 넘겨져서 연속성을 유지한다. 이것은 제2인에 대해 상이한 아바타와 프로파일을 사용하는 것과는 대조적이다. 점수나 진척 레벨과 같은 다른 프로파일 데이터의 연속성은 유지하면서 제2인에게 새로운 아바타가 제공되는 것 역시 가능하다. 다른 가능성은 새로운 아바타의 연속성은 유지하면서 다른 프로파일 데이터를 사용하는 것이다. 예를 들어 제2인은 리셋된 점수로 새로운 게임을 시작하면서 제1인과 동일한 아바타를 사용할 수 있다.
- [0111] 예를 들어 도 10e를 참조할 때, 사람(1024)이 디스플레이(1040) 상의 아바타(1042)를 제어함으로써(가령, 사람(1024)이 왼팔을 들어 올리면 아바타(1042)가 대응하는 팔을 들어 올림) 애플리케이션에 참가하고 있다. 예로서, 아바타는 그 사람(1024)의 특정 외형에 대응하지 않는 일반적인 외관을 가진다. 제2인(1022)이 참가 의도를 나타내면서 시계의 뒤쪽에 서 있다. 도 10f에서, 제1인(1024)이 뒤로 나가며 참가 해제 의도를 나타내고, 제2인은 앞으로 나오면서 참가 의도를 나타냄으로써, 이제 제2인의 움직임들이 디스플레이(1050) 내 아바타(1042)를 제어하는 데 사용된다(가령 사람(1022)이 오른팔을 들어 올리면 아바타(1042)가 상응하는 팔을 들어 올림).
- [0112] 이 예에서는 한 번에 한 명의 참가하는 플레이어가 존재하며, 플레이어 교체가 일어난다. 다른 예에서는 한 번에 여러 참가하는 플레이어가 존재하고, 한 명 이상의 플레이어들에 대한 플레이어 교체가 일어난다.
- [0113] 애플리케이션 자체도 참가할 사람의 의도를 판단하는 데 있어 한 요인이 될 수 있다. 예를 들어, 한 명의 플레이어가 있는 애플리케이션에서는 새로운 아바타가 기존 아바타에 합류하지 못할 것이다. 제2인이 제1인을 대신하기 위해서는, 아마도 제2인에 의해 보다 가능한 참가 의도 및 제1인에 의해 참가 해체에 대한 보다 강한 의도가 필요로 될 것이다. 어떤 경우 제2인이 보다 강한 참가 의도를 가질 수 있지만 제1인이 참가 해체를 가지지 않는 한 무시된다.
- [0114] 도 10a는 한 사람이 어떤 애플리케이션에 참가하고 있고, 다른 사람들은 그 애플리케이션에 참가하지 않는 예를 도시한다. 앞서 논의한 바와 같이, 시계 안의 사람들이 애플리케이션에 참가하고자 반드시 의도하지는 않는다. 또한, 시계 밖의 사람들이 애플리케이션에 참가하고자 의도할 수 있다. 그러나, 이러한 의도는 사람이 시계 안에 있기 까지는 그 사람의 이미지 데이터로부터 판단될 수 없다. 이 경우 오디오 데이터가 의도를 판단하는 데 사용될 수도 있다.
- [0115] 디스플레이(1000)는 아바타(1025)가 길(1008) 위에 서 있는 가상 세계의 예를 제공한다. 아바타(1025)는 사람(1024)의 외관에 해당하는 모양을 가진다. 점선들(1019 및 1020) 사이의 해당 시계(1021)와 일치하는 대표 렌즈(1012)를 가진 깊이 카메라 시스템(1010)이 제공된다. 한 사람(1022)이 시계 밖에 있고, 그에 따라 깊이 카

메라 시스템(1010)에 의해서 시각적으로 검출되지 못한다. 또 다른 사람(1028)이 의자(1030)에 앉아 있다. 또 다른 사람(1024)은 깊이 카메라 시스템을 마주하여 서 있다. 또 다른 사람(1026)은 시계를 지나 걸어가고 있다. 앞서 논의된 다양한 고려사항들에 기초하여, 사람(1022)은 시계 밖에 있는 그의 위치에 기반하여 애플리케이션에 참가할 의도를 가진다고 간주 될 가능성이 없을 것이며, 심지어 존재한다고 인식되지도 못할 것이다. 사람(1028)은 시계(1021)의 뒤나 옆을 향하는 그의 위치, 및 몸의 옆모습이 깊이 카메라 시스템을 향하며 앉아 있는 자세인 자세에 기반하여 애플리케이션에 참가하고자 하는 의도를 가진다고 간주될 가능성이 없을 것이다.

[0116] 사람(1024)은 시계(1021)의 중앙에 있는 그의 위치, 및 어깨를 펴고 깊이 카메라 시스템을 마주하는 자세, 그리고 한 팔을 옆으로 들어 올리는 움직임에 기반하여 애플리케이션에 참가하고자 하는 의도를 가진다고 간주될 가능성이 있을 것이다. 사람(1024)이 참가될 때, 그 사람(1024)의 움직임들에 상응하여 아바타(1026)의 움직임들이 애플리케이션에 의해 정해질 수 있다.

[0117] 사람(1026)은 시계(1021)의 옆을 향하는 그의 위치, 몸이 깊이 카메라 시스템으로부터 멀리 향하는 자세, 및 사람이 시계를 지나고 시계의 경계(1020)를 향해 이동하는 움직임에 기반하여 애플리케이션에 참가하고자 하는 의도를 가진다고 간주될 가능성이 없을 것이다.

[0118] 깊이 카메라 시스템 및/또는 관련 컴퓨팅 환경은 한 명 또는 여러 명의 사람들의 의도를 반복적으로, 가령 초당 여러 번 판단하여, 적절한 변화들이 빠르게 검출되고 애플리케이션이 그에 따라 맞춰지게 할 수 있다. 예컨대 너무 빈번한 변화들을 피하기 위해, 결과들을 완만하게 하기 위해 필터링 및 평균내기가 사용될 수 있다.

[0119] 도 10b는 두 사람이 한 애플리케이션에 참가하고 있고 디스플레이 상의 가상 공간에 대한 공통 시점을 가지는 전형적 디스플레이를 도시한다. 디스플레이(1002)에서 도 10a로부터의 사람(1024)에게 앞에서 시계 밖에 있었던 사람(1022)이 합류한다. 앞서 논의된 다양한 요인들에 기반하여, 사람(1022)은 애플리케이션에 참가하고자 하는 의도를 가진다고 판단된다. 그 결과, 해당 아바타(1023)가 디스플레이(1002) 상에서 사람(1024)의 아바타(1025) 옆에 제공된다. 이 경우, 디스플레이는 도 8b의 단계(814)와 관련하여 앞서 논의된 바와 같이, 사람들(1022 및 1024)에게 가상 공간의 공통 시점을 제공한다.

[0120] 도 10c는 두 사람이 한 애플리케이션에 참가하고 있고 디스플레이 상의 가상 공간에 대한 서로 다른 시점을 가지는 전형적 스플릿 스크린(split screen) 디스플레이를 도시한다. 도 8b의 단계(816)와 관련하여 앞서 논의된 바와 같이, 디스플레이(1004)에는 아바타들(1023 및 1025)에 별개의 디스플레이 영역들(1005 및 1006)이 제공되고, 각각의 디스플레이 영역에서 상이한 시점이나 카메라 각도로부터 가상 공간이 보여진다. 예를 들어, 디스플레이 영역(1006)의 시점은 도 10b에서와 거의 같다. 그러나, 디스플레이 영역(1005)의 시점은 도 10b의 디스플레이 영역(1002)과 매우 상이하다. 아바타들(1023 및 1025)은 여전히 동일한 가상 공간에 있다는 것을 알아야 한다. 예를 들어, 그들은 모두 길(1008) 위에 서 있다. 그러나, 디스플레이 영역(1005)의 시점은 언덕 위에 집(1007)이 있는 길가를 보고, 디스플레이 영역(1006)의 시점은 계속해서 길을 본다.

[0121] 옵션으로서, 디스플레이 영역들은 각각의 디스플레이 영역들 안에 완전히 상이한 가상 공간들이나 구획된 장면들을 제공할 수 있다. 디스플레이 영역들은 심지어 상이한 애플리케이션들의 가상 공간들을 제공할 수 있다.

[0122] 논의된 바와 같이, 디스플레이 영역들(1005 및 1006)의 상대적 위치들은 사람들(1022 및 1024)의 상대적 위치들에 상응할 수 있다. 또 다른 옵션은 어떤 사람이 시계로 진입하는 방향에 기반하여 그 사람에게 디스플레이 영역을 제공하는 것이다. 예를 들어, 사람(1022)은 왼편으로부터 시계 안에 진입할 수 있으며, 이 경우 도 10c의 설정이 제공된다.

[0123] 도 10d는 도 10c의 디스플레이 영역들과 관련하여, 사람들이 물리적 공간에서 위치를 서로 바꿀 때 디스플레이 영역들이 맞교환되는 전형적인 스플릿 스크린 디스플레이를 도시한다. 도 9c와 관련하여 앞서 논의된 바와 같이, 사람(1022)이 사람(1024)의 왼편으로부터 오른편으로 이동할 때, 디스플레이 영역들(1005 및 1006)은 그에 맞춰 맞교환된다. 이것은 사람들이 시계 안에서 돌아다닐 때에도 그들이 애플리케이션에 참가하는 것을 용이하게 하는데, 이는 사람들이 그들이 상호작용하고 있는 디스플레이 영역에 물리적으로 가까우면서 그와 연계되기 때문이다.

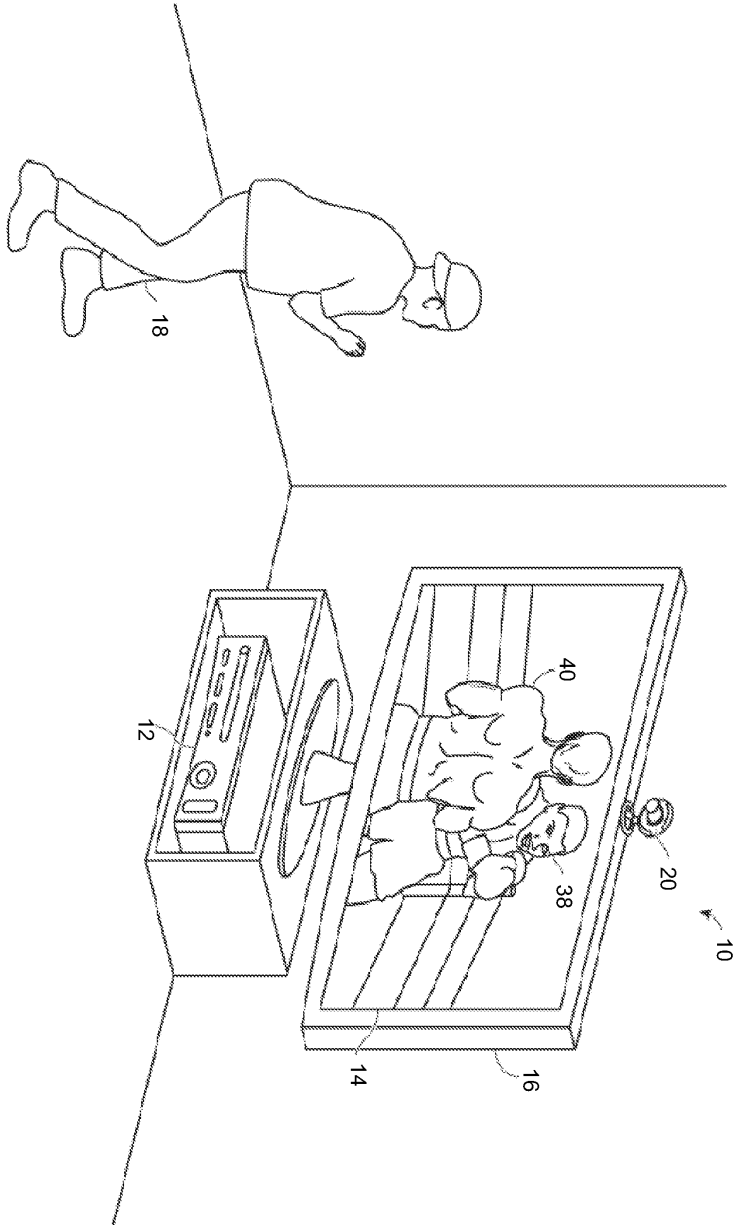
[0124] 도 10a-10f는 추적이 시계 안의 사람의 신체를 구별하는 것을 포함하는 경우, 모션 캡처 시스템의 시계 안에서 추적되는 사람들의 신체에 대한 예들을 제공한다.

[0125] 여기서의 기술에 대한 상기 상세한 설명은 예시 및 설명의 목적으로 제시되었다. 그것이 포괄적이도록 의도되거나 개시된 정확한 형태로 기술을 제한하고자 의도된 것은 아니다. 상기 교시에 비추어 많은 수정 및 변경이 가능하다. 상기 기술된 실시예들이 기술의 원리와 그의 실용적 응용을 가장 잘 설명하도록 선택되어, 그에 의

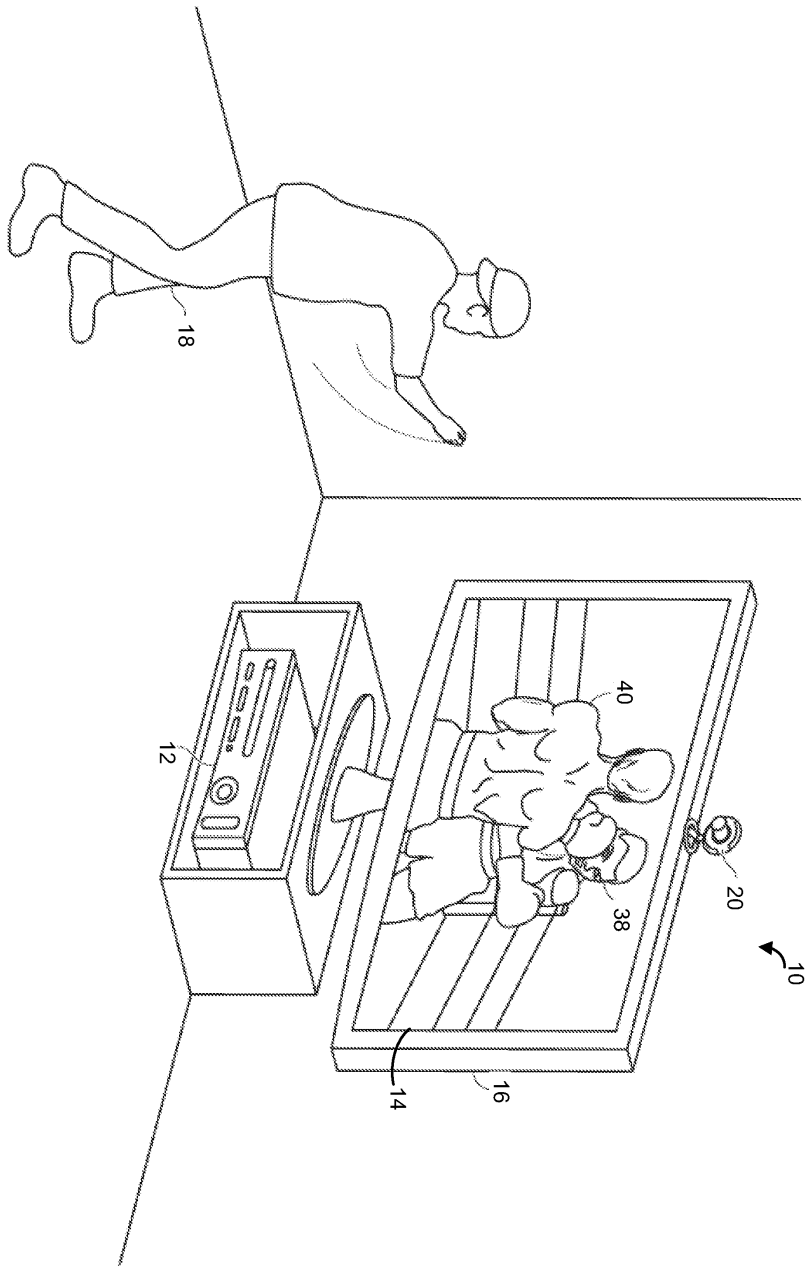
해 당업자가 고려된 특정 용도에 적합한 다양한 실시예에서 그리고 다양한 변형을 이용하여 그 기술을 가장 잘 활용하도록 해준다. 기술의 범위는 여기 첨부된 청구범위에 의해 정의되도록 의도된다.

도면

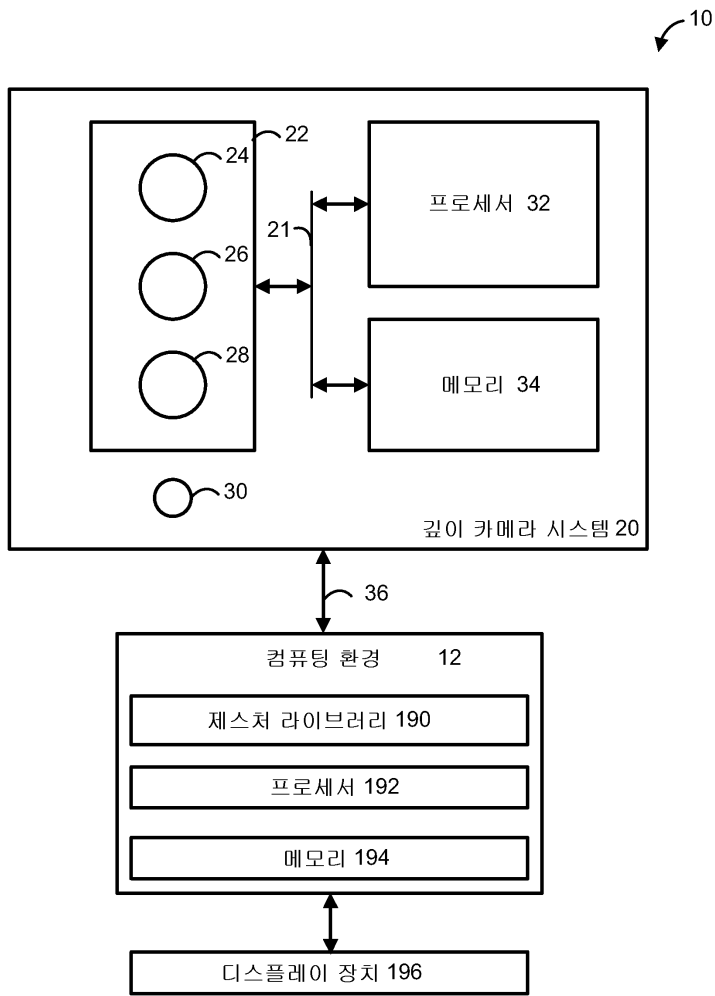
도면1a



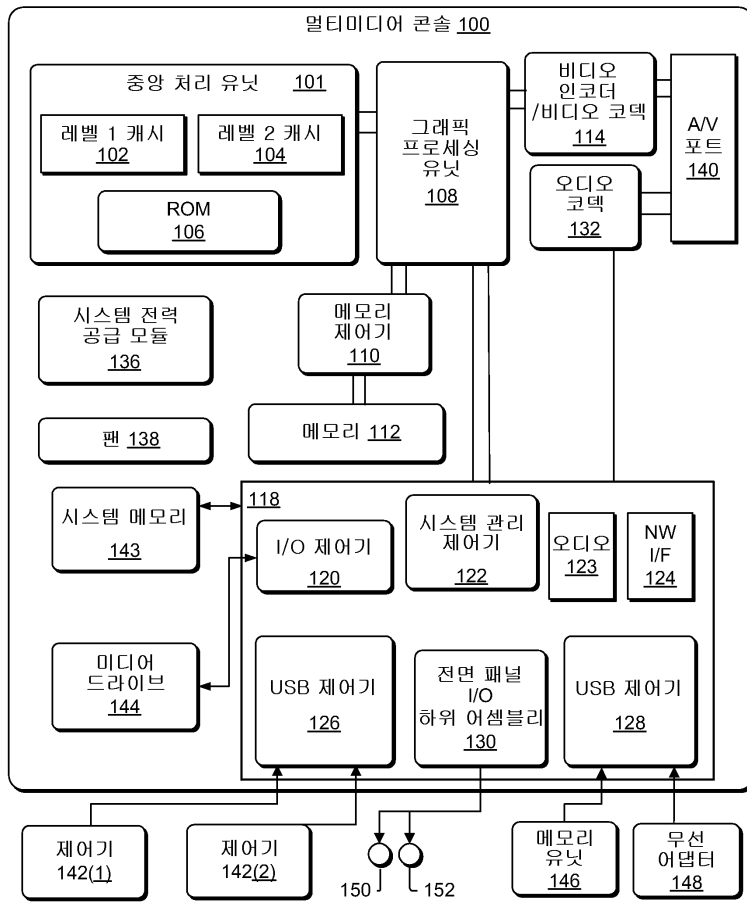
도면1b



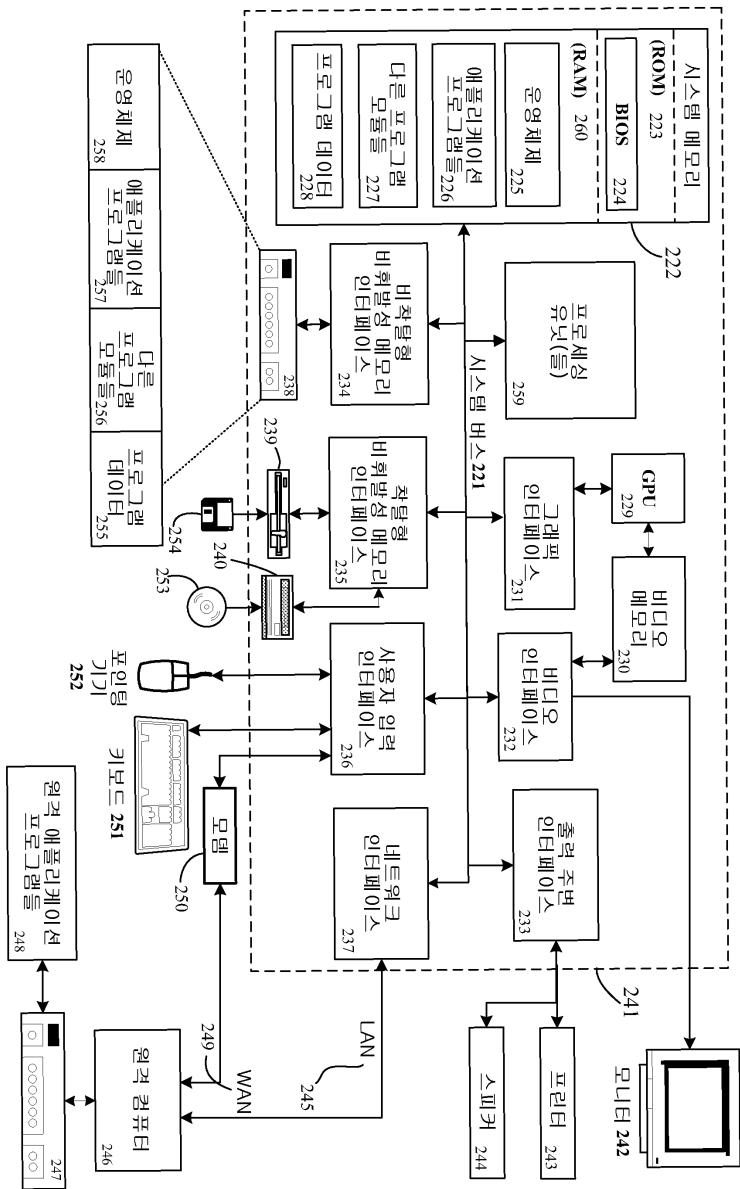
도면2



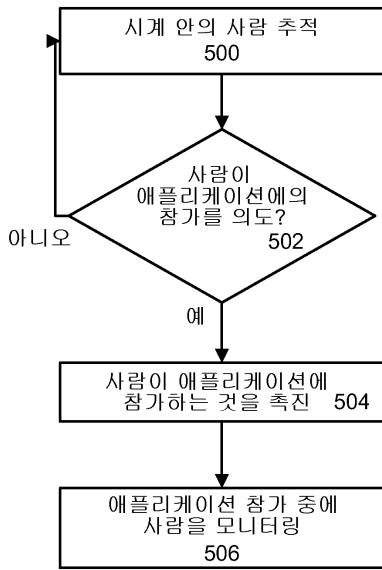
도면3



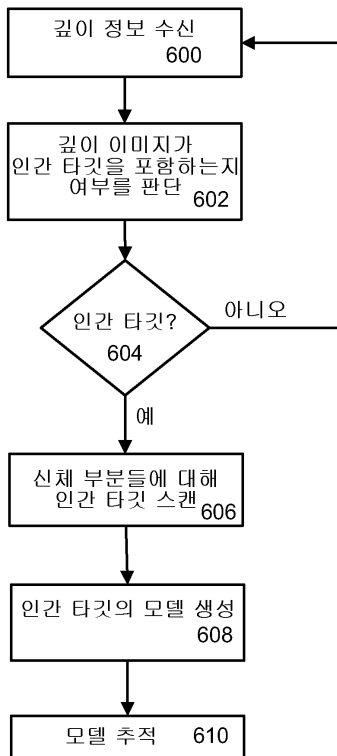
도면4



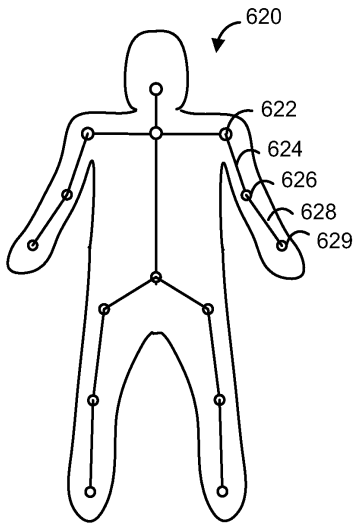
도면5



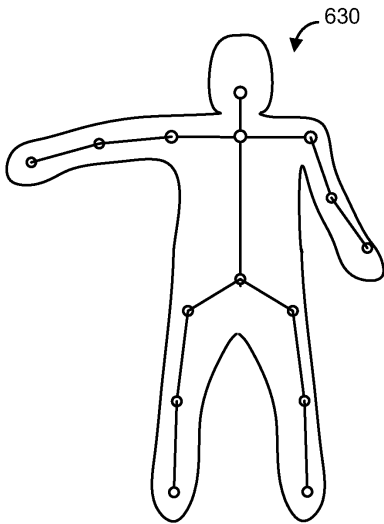
도면6a



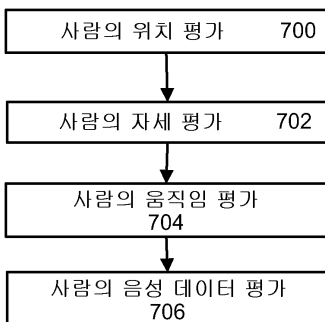
도면6b



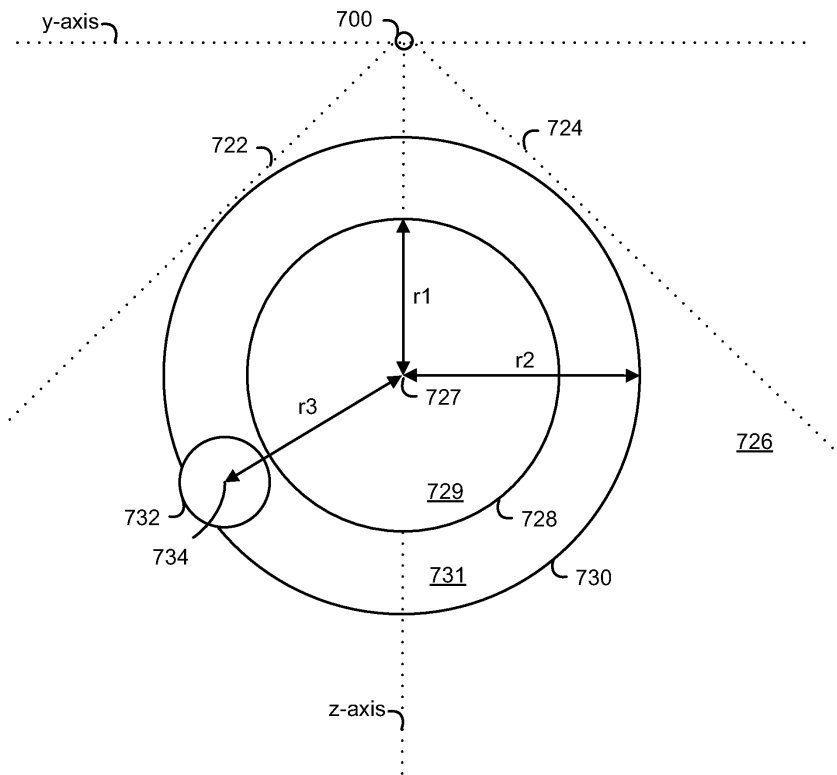
도면6c



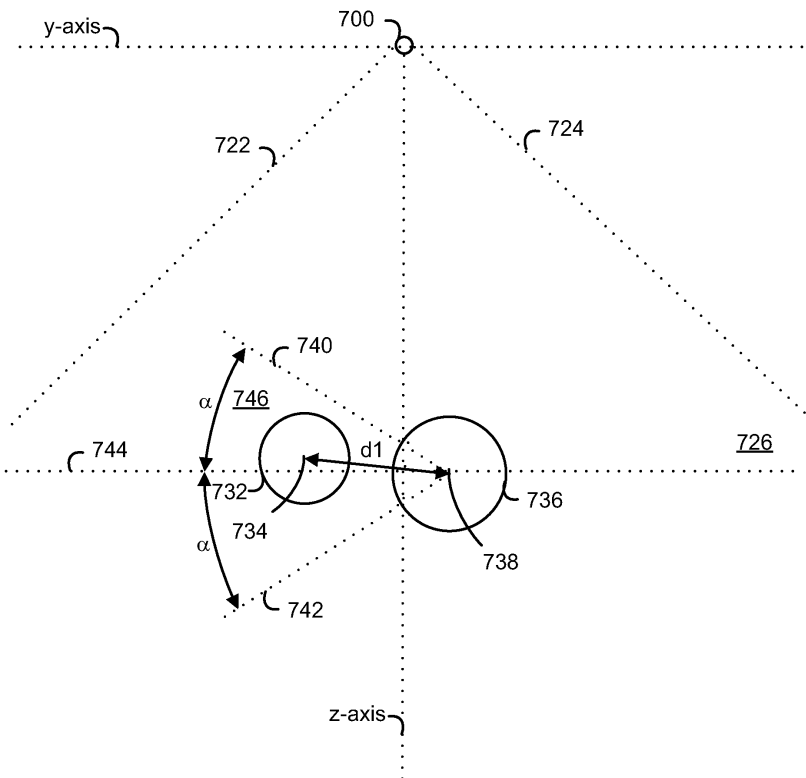
도면7a



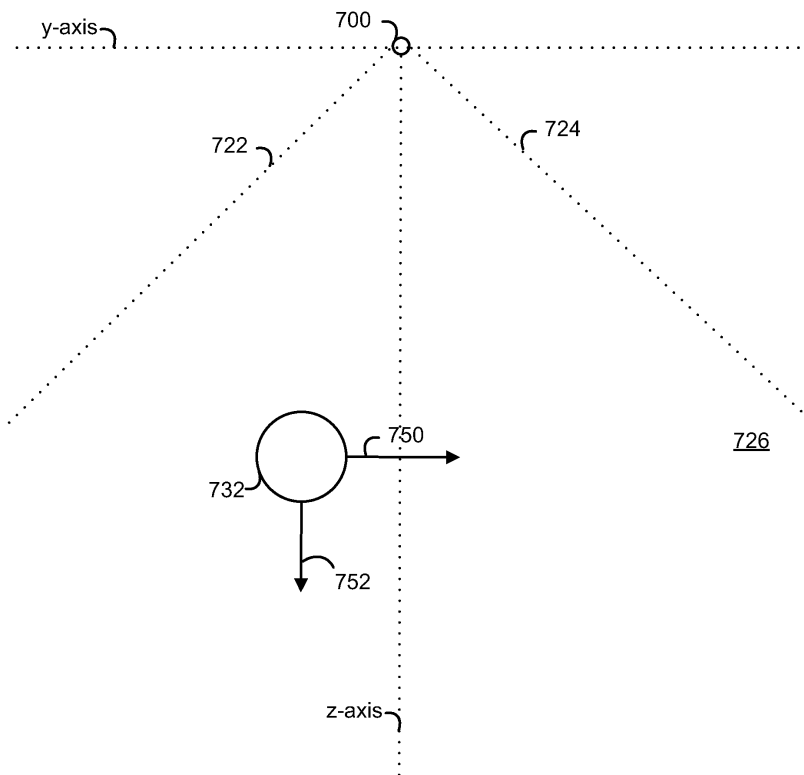
도면7b



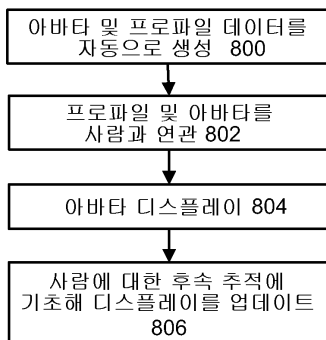
도면7c



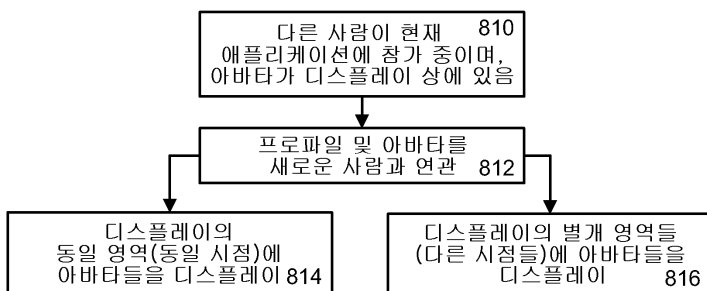
도면7d



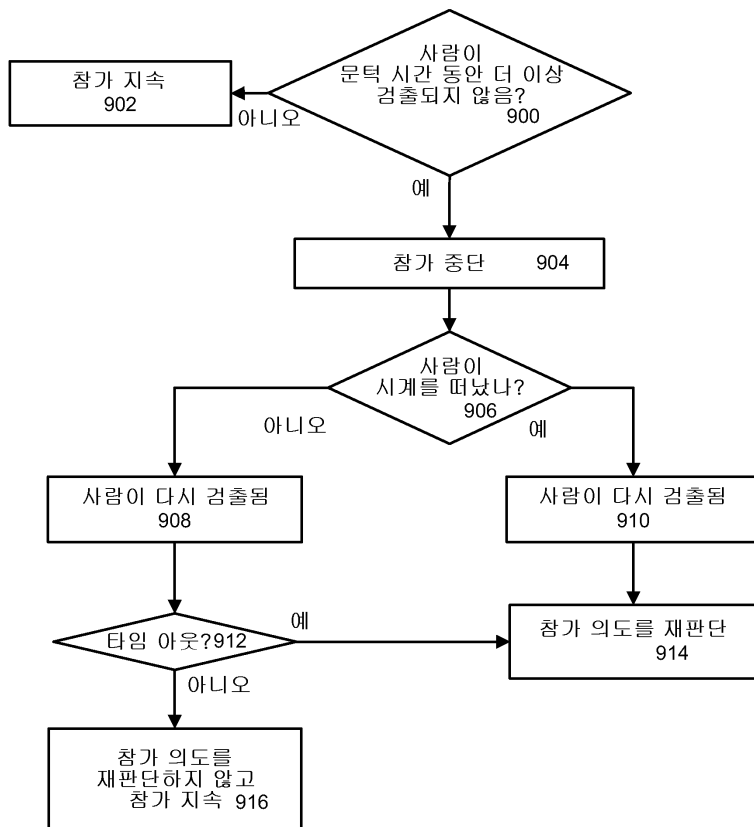
도면8a



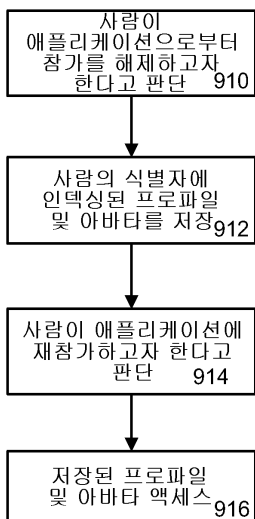
도면8b



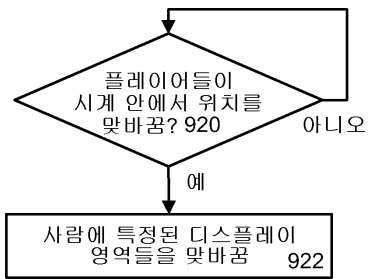
도면9a



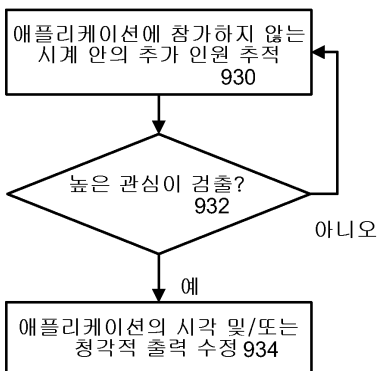
도면9b



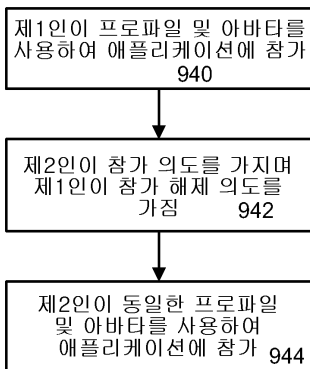
도면9c



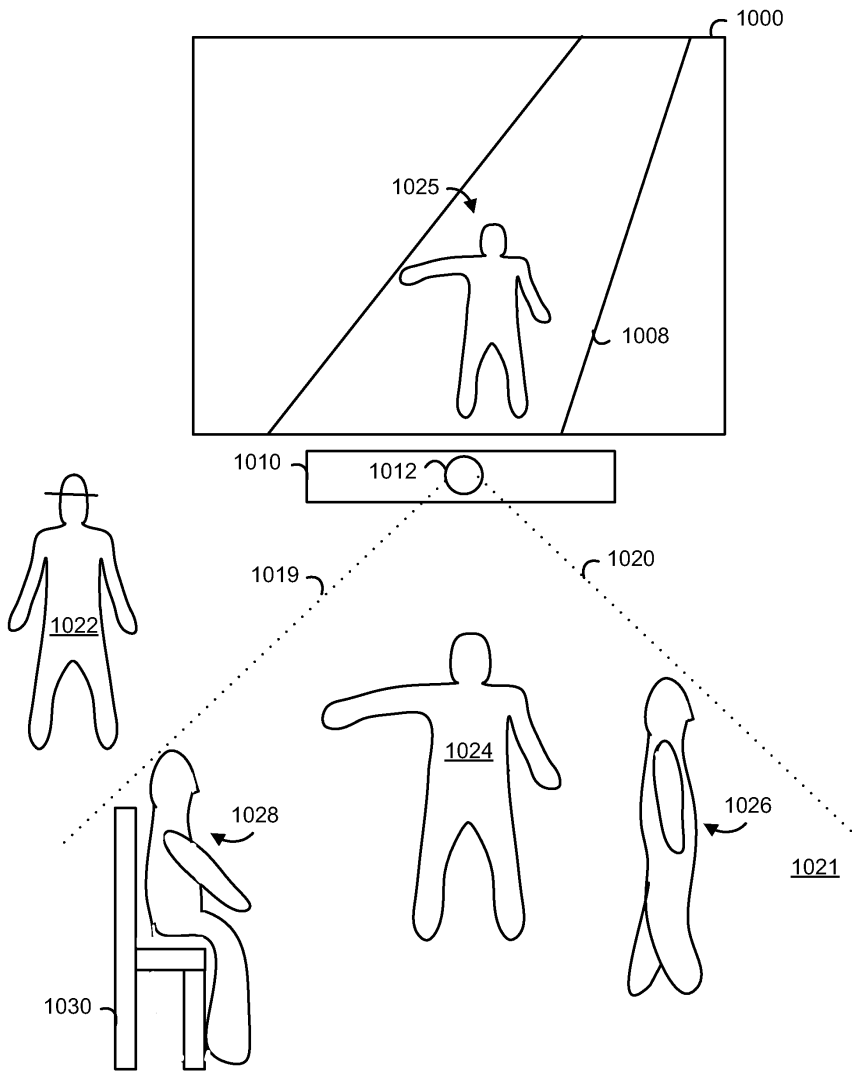
도면9d



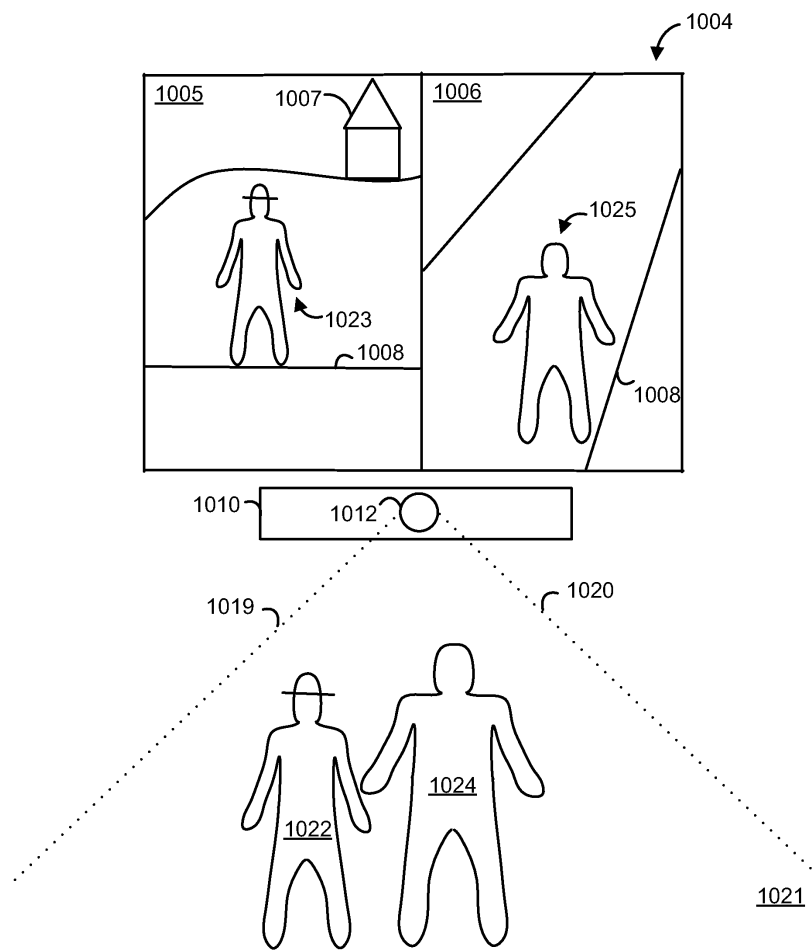
도면9e



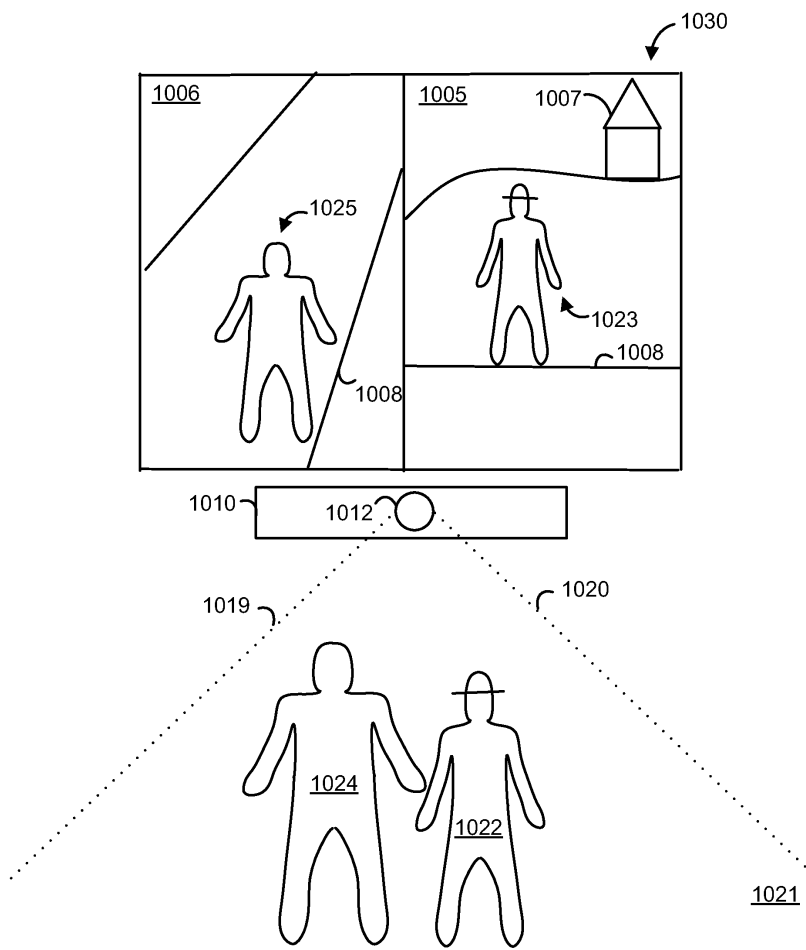
도면10a



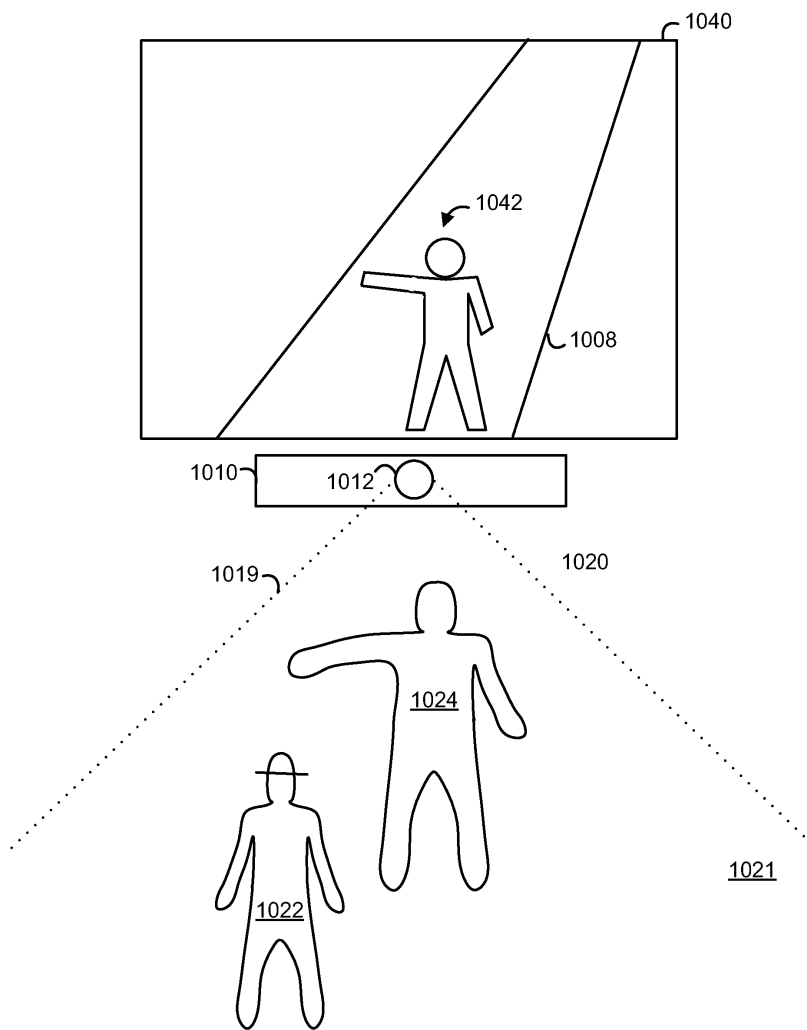
도면10c



도면10d



도면10e



도면10f

