



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년11월27일
(11) 등록번호 10-1802125
(24) 등록일자 2017년11월21일

(51) 국제특허분류(Int. Cl.)
G06T 7/73 (2017.01) A63F 13/213 (2014.01)
G06F 3/01 (2006.01) G06K 9/00 (2006.01)
G06T 7/194 (2017.01) G06T 7/579 (2017.01)

(52) CPC특허분류
G06T 7/75 (2017.01)
A63F 13/213 (2015.01)

(21) 출원번호 10-2017-7008298(분할)

(22) 출원일자(국제) 2010년10월06일

심사청구일자 2017년03월29일

(85) 번역문제출일자 2017년03월27일

(65) 공개번호 10-2017-0038101

(43) 공개일자 2017년04월05일

(62) 원출원 특허 10-2012-7008854

원출원일자(국제) 2010년10월06일

심사청구일자 2015년09월07일

(86) 국제출원번호 PCT/US2010/051662

(87) 국제공개번호 WO 2011/044257

국제공개일자 2011년04월14일

(30) 우선권주장
12/575,388 2009년10월07일 미국(US)

(56) 선행기술조사문헌
Mikic et al. Human body model acquisition and tracking using voxel data. International Journal of Computer Vision, Vol. 53, No. 3, 2003년, pp. 199-223.

KR1020050099547 A

Werghi. Segmentation and modeling of full human body shape from 3-D scan data: A survey. IEEE Transactions on Systems Man. and Cybernetics, Vol. 37, No. 6, 2007년, pp. 1122-1136.

GB2389500 A

(73) 특허권자
마이크로소프트 테크놀로지 라이선싱, 엘엘씨
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이

(72) 발명자
레이밴드 토머
미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 페이턴즈 마
이크로소프트 코포레이션

리 조니

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

(뒷면에 계속)

(74) 대리인
김태홍, 김진희

전체 청구항 수 : 총 20 항

심사관 : 김창원

(54) 발명의 명칭 인간 트래킹 시스템

(57) 요약

장면의 깊이 이미지 등의 이미지가 장치에 의해 수신, 관찰 또는 캡처될 수 있다. 다음으로 복셀 격자가 깊이 이
미지에 근거하여 생성되어 깊이 이미지가 다운샘플링될 수 있다. 복셀 격자에 포함된 배경이 제거되어, 인간 타
겟 등의 전경 사물에 연관된 하나 이상의 복셀을 분리할 수 있다. 분리된 인간 타겟의 하나 이상의 말단의 위치
또는 포지션이 결정되고, 그 하나 이상의 말단의 위치 또는 포지션에 근거하여 모델이 조정될 수 있다.

대표도



(52) CPC특허분류

G06F 3/017 (2013.01)
G06K 9/00342 (2013.01)
G06K 9/00362 (2013.01)
G06K 9/00711 (2013.01)
G06T 7/194 (2017.01)
G06T 7/579 (2017.01)
A63F 2300/6045 (2013.01)
G06T 2207/10028 (2013.01)
G06T 2207/30196 (2013.01)

(72) 발명자

스테치니악 시몬

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

피퍼 크레이그

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

리우 샤오

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

명세서

청구범위

청구항 1

장면에서 사용자의 관절의 치수를 추정하는 방법에 있어서,
 상기 장면에서 상기 사용자를 나타내는 깊이 이미지를 수신하는 단계;
 상기 깊이 이미지에서 상기 사용자의 상기 관절에 대한 복수의 제시된 치수들을 결정하는 단계;
 상기 복수의 제시된 치수들에 기초하여 상기 관절의 치수를 결정하는 단계; 및
 상기 사용자에게 대응하는 컴퓨터-유지 모델(computer-maintained model)을 상기 관절의 치수를 갖도록 업데이트하는 단계를 포함하고,
 상기 관절의 치수는 길이와 폭을 포함하는 것인, 장면에서 사용자의 관절의 치수를 추정하는 방법.

청구항 2

제1항에 있어서,
 상기 제시된 치수들의 각각에 대해 평균 및 표준 편차를 계산하는 단계를 더 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 방법.

청구항 3

제1항에 있어서,
 상기 복수의 제시된 치수들 중
 정의된 비율 편차(percentage deviation) 밖인 하나 이상의 제시된 치수를 버리는(rejecting) 단계를 더 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 방법.

청구항 4

제1항에 있어서,
 표준 편차와 상기 복수의 제시된 치수들의 개수 간의 가장 높은 비율(highest ratio)을 갖는 예측치에 기초하여 상기 관절의 치수를 결정하는 단계를 더 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 방법.

청구항 5

제1항에 있어서,
 상기 관절의 치수는 길이 또는 폭 중 적어도 하나를 포함하는 것인, 장면에서 사용자의 관절의 치수를 추정하는 방법.

청구항 6

제1항에 있어서,
 하나 이상의 휴리스틱(heuristics)에 기초하여 상기 관절의 치수가 맞는지 여부를 결정하는 단계를 더 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 방법.

청구항 7

장면에서 사용자의 관절의 치수를 추정하는 시스템에 있어서,
 컴퓨터 실행가능한 명령어들이 저장되어 있는 하나 이상의 메모리; 및
 상기 컴퓨터 실행가능한 명령어들을 실행시키도록 동작가능한 하나 이상의 프로세서를 포함하고,

상기 컴퓨터 실행가능한 명령어들은

상기 장면에서 상기 사용자를 나타내는 깊이 이미지를 수신하는 동작;

상기 깊이 이미지에서 상기 사용자의 상기 관절에 대한 복수의 제시된 치수들을 결정하는 동작;

상기 복수의 제시된 치수들에 기초하여 상기 관절의 치수를 결정하는 동작; 및

상기 사용자에게 대응하는 컴퓨터-유지 모델(computer-maintained model)을 상기 관절의 치수를 갖도록 업데이트하는 동작

을 포함하는 동작들을 수행하기 위한 명령어들을 포함하고,

상기 관절의 치수는 길이와 폭을 포함하는 것인, 장면에서 사용자의 관절의 치수를 추정하는 시스템.

청구항 8

제7항에 있어서,

상기 동작들은 또한, 상기 복수의 제시된 치수들 중 정의된 비율 편차(percentage deviation) 밖인 하나 이상의 제시된 치수를 버리는 동작을 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 시스템.

청구항 9

제7항에 있어서,

상기 동작들은 또한, 표준 편차와 상기 복수의 제시된 치수들의 개수 간의 가장 높은 비율(highest ratio)을 갖는 예측치에 기초하여 상기 관절의 치수를 결정하는 동작을 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 시스템.

청구항 10

제7항에 있어서,

상기 관절의 치수는 길이 또는 폭 중 적어도 하나를 포함하는 것인, 장면에서 사용자의 관절의 치수를 추정하는 시스템.

청구항 11

제7항에 있어서,

상기 동작들은 또한, 하나 이상의 휴리스틱(heuristics)에 기초하여 상기 관절의 치수가 맞는지 여부를 결정하는 동작을 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 시스템.

청구항 12

제7항에 있어서,

상기 동작들은 또한, 상기 제시된 치수들의 각각에 대해 평균을 계산하는 동작을 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 시스템.

청구항 13

제7항에 있어서,

상기 동작들은 또한, 상기 제시된 치수들의 각각에 대해 표준 편차를 계산하는 동작을 포함하는, 장면에서 사용자의 관절의 치수를 추정하는 시스템.

청구항 14

일시적인 신호로만 구성되지 않은 컴퓨터 판독가능 저장 장치에 있어서, 상기 저장 장치는 장면에서 사용자의 관절의 치수를 추정하기 위한 컴퓨터 실행가능한 명령어들이 저장되어 있고, 상기 컴퓨터 실행가능한 명령어들은

상기 장면에서 상기 사용자를 나타내는 깊이 이미지를 수신하는 동작;
 상기 깊이 이미지에서 상기 사용자의 상기 관절에 대한 복수의 제시된 치수들을 결정하는 동작;
 상기 복수의 제시된 치수들에 기초하여 상기 관절의 치수를 결정하는 동작; 및
 상기 사용자에게 대응하는 컴퓨터-유지 모델(computer-maintained model)을 상기 관절의 치수를 갖도록 업데이트하는 동작
 을 포함하는 동작들을 수행하기 위한 명령어들을 포함하고,
 상기 관절의 치수는 길이와 폭을 포함하는 것인, 컴퓨터 판독가능 저장 장치.

청구항 15

제14항에 있어서,
 상기 동작들은 또한, 상기 복수의 제시된 치수들 중 정의된 비율 편차(percentage deviation) 밖인 하나 이상의 제시된 치수를 버리는 동작을 포함하는, 컴퓨터 판독가능 저장 장치.

청구항 16

제14항에 있어서,
 상기 동작들은 또한, 표준 편차와 상기 복수의 제시된 치수들의 개수 간의 가장 높은 비율(highest ratio)을 갖는 예측치에 기초하여 상기 관절의 치수를 결정하는 동작을 포함하는, 컴퓨터 판독가능 저장 장치.

청구항 17

제14항에 있어서,
 상기 관절의 치수는 길이 또는 폭 중 적어도 하나를 포함하는 것인, 컴퓨터 판독가능 저장 장치.

청구항 18

제14항에 있어서,
 상기 동작들은 또한, 하나 이상의 휴리스틱(heuristics)에 기초하여 상기 관절의 치수가 맞는지 여부를 결정하는 동작을 포함하는, 컴퓨터 판독가능 저장 장치.

청구항 19

제14항에 있어서,
 상기 동작들은 또한, 상기 제시된 치수들의 각각에 대해 평균을 계산하는 동작을 포함하는, 컴퓨터 판독가능 저장 장치.

청구항 20

제14항에 있어서,
 상기 동작들은 또한, 상기 제시된 치수들의 각각에 대해 표준 편차를 계산하는 동작을 포함하는, 컴퓨터 판독가능 저장 장치.

발명의 설명

배경 기술

[0001] 컴퓨터 게임, 멀티미디어 애플리케이션을 비롯한 각종 컴퓨팅 애플리케이션에서 사용자는 컨트롤(controls)을 통해 게임 캐릭터나 애플리케이션의 기타 양태들을 조종한다. 통상적으로 이런 컨트롤은 컨트롤러, 리모콘, 키보드, 마우스 등을 사용하여 입력된다. 불행히도, 이런 컨트롤은 배우기가 어려워, 사용자와 이런 게임이나 애플리케이션 간에 벽을 생기게 할 수 있다. 또한, 이러한 컨트롤은 이 컨트롤이 사용된 실제 게임 동작이나 기타

애플리케이션 동작과 다를 수 있다. 예를 들어, 게임 캐릭터가 야구 방망이를 휘두르게 하는 게임 제어가 실제로 야구 방망이를 휘두르는 모션과 일치하지 않을 수 있다.

발명의 내용

과제의 해결 수단

- [0002] 한 장면에서 사용자를 트래킹하는 시스템 및 방법이 본원에 개시된다. 예를 들어, 장면의 깊이 이미지(depth image) 등의 이미지를 수신하고 관찰한다. 이후에, 깊이 이미지에 기반하여 복셀 격자(grid of voxels)를 생성하여, 깊이 이미지를 다운샘플링할 수 있다(downsample). 예를 들어, 깊이 이미지는 부분 또는 블록으로 나뉘는 복수 개의 화소를 포함할 수 있다. 이후에 각 부분 또는 블록에 대해 복셀을 생성하여, 수신된 깊이 이미지를 복셀 격자로 다운샘플링할 수 있다.
- [0003] 일 실시예에 따르면, 복셀 격자에 포함된 배경을 제거하여, 인간 타겟과 같은 전경(foreground) 사물에 관련된 하나 이상의 복셀을 분리할 수 있다(isolate). 분리된 인간 타겟의 중심(centroid) 또는 중앙, 머리, 어깨, 엉덩이, 팔, 손, 팔꿈치, 다리, 발, 무릎 등 하나 이상의 말단(extremities)의 위치나 포지션을 결정할 수 있다. 또한, 말단의 폭, 길이 등을 포함하는 치수를 비롯한 크기(dimensions)를 결정할 수 있다.
- [0004] 다음으로, 하나 이상의 말단 및/또는 그에 따라 결정된 크기를 기초로 하여 모델을 조정할 수 있다. 예를 들어, 모델은 관절 및/또는 뼈를 포함하는 골격 모델일 수 있다. 하나 이상의 관절이 그에 해당하는 하나 이상의 말단의 위치나 포지션에 할당되고, 또는 그 사이에 정의된 뼈가 그에 해당하는 하나 이상의 말단의 크기로 조정되도록, 모델의 하나 이상의 관절을 조정할 수 있다.
- [0005] 조정된 모델을 처리할 수 있다. 예를 들어, 일 실시예에서, 조정된 모델을 아바타(avatar)나 게임 캐릭터에 매핑하여, 그 아바타나 게임 캐릭터가 사용자를 흉내 내도록 애니메이션화하거나, 또는 모델의 각종 신체 부위의 포지션에 기반하여 애플리케이션에서 수행될 컨트롤을 결정하는 데 사용가능한 컴퓨팅 환경의 제스처 라이브러리에 조정된 모델을 제공할 수 있다.
- [0006] 본 요약은 아래의 상세한 설명에서 추가적으로 설명되는 일련의 개념을 간략화된 형태로 소개하기 위한 것이다. 본 요약은 특허청구된 대상의 핵심적인 특징 또는 필수적인 특징을 밝히기 위한 것이 아니며, 특허청구된 대상의 범위를 결정하는 데 일조하기 위해 사용되는 것도 아니다. 또한, 특허청구된 대상은 본 내용에서 언급된 임의의 또는 모든 단점들을 해결하는 구현에 한정되지 않는다.

도면의 간단한 설명

- [0007] 도 1a 및 1b는 게임을 하는 사용자가 있는 타겟 인식, 분석 및 트래킹 시스템의 일 실시예가 도시된다.
- 도 2는 타겟 인식, 분석 및 트래킹 시스템에 사용가능한 캡처 장치의 예시적인 실시예를 도시한다.
- 도 3은 타겟 인식, 분석 및 트래킹 시스템에서 하나 이상의 제스처를 해석하고, 또한 타겟 인식, 분석 및 트래킹 시스템에 의해 디스플레이되는 아바타나 온-스크린(on-screen) 캐릭터를 애니메이션화하는 데 사용가능한 컴퓨팅 환경의 예시적인 실시예를 도시한다.
- 도 4는 타겟 인식, 분석 및 트래킹 시스템에서 하나 이상의 제스처를 해석하고, 또한 타겟 인식, 분석 및 트래킹 시스템에 의해 디스플레이되는 아바타나 온-스크린(on-screen) 캐릭터를 애니메이션화하는 데 사용가능한 컴퓨팅 환경의 다른 실시예를 도시한다.
- 도 5는 한 장면의 사용자를 트래킹하는 예시적인 방법의 순서도를 도시한다.
- 도 6은 캡처되거나 관찰되는 깊이 이미지의 예시적인 실시예를 도시한다.
- 도 7a 및 7b는 다운샘플링되는 깊이 이미지의 일부의 예시적인 실시예를 도시한다.
- 도 8은 인간 타겟에 대해 추정되는 중심 또는 중앙의 예시적인 실시예를 도시한다.
- 도 9는 중심부 부피(core volume)를 결정하기 위해 정의되는 경계 상자(bounding box)의 예시적인 실시예를 도시한다.
- 도 10은 머리 후보(head candidate)의 점수를 매기기 위해(score) 생성되는 머리 원기둥(head cylinder) 및 상체 원기둥(torso cylinder)의 예시적인 실시예를 도시한다.

도 11은 머리와 인간 타겟의 중심 또는 중앙에 기반한 헤드-투-센터(head-to-center) 벡터의 예시적인 실시예를 도시한다.

도 12는 헤드-투-센터 벡터에 근거하여 결정된 어깨 부피 상자(volume box) 및 엉덩이 부피 상자의 예시적인 실시예를 도시한다.

도 13은 어깨 부피 상자 및 엉덩이 부피 상자에 근거하여 계산된 어깨 및 엉덩이의 일 실시예를 도시한다.

도 14는 중심부 부피를 나타내는 원기둥의 예시적인 실시예를 도시한다.

도 15(a) - 15(c)는 앵커 포인트(anchor points)에 근거하여 결정되는 손의 예시적인 실시예를 도시한다.

도 16은 팔과 다리의 평균 포지션(average positions) 및/또는 앵커 포인트에 근거하여 결정되는 손과 발의 예시적인 실시예를 도시한다.

도 17은 생성가능한 모델의 예시적인 실시예를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0008] 도 1a 및 1b는 복싱 게임을 하는 사용자(18)가 있는 타겟 인식, 분석 및 트래킹 시스템(10)의 구성에 대한 예시적인 실시예를 도시한다. 예시적인 실시예에서, 타겟 인식, 분석 및 트래킹 시스템(10)을 사용하여 사용자(18) 등의 인간 타겟을 인식, 분석 및/또는 트래킹할 수 있다.

[0009] 도 1a에 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. 예시적인 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. 일 실시예에서, 컴퓨팅 환경(12)은 예를 들어, 깊이 이미지를 수신하고, 그 깊이 이미지에 근거하여 복셀 격자를 생성하고, 복셀 격자에 포함된 배경을 제거하여 인간 타겟에 관련된 하나 이상의 복셀을 분리하고, 분리된 인간 타겟의 하나 이상의 말단의 위치나 포지션을 결정하여, 그 하나 이상의 말단의 위치나 포지션에 근거하여 모델을 조정하기 위한 인스트럭션 또는 그 밖의 임의의 적절한 인스트럭션을 포함하는 인스트럭션을 실행할 수 표준화된 프로세서, 특화된 프로세서, 마이크로프로세서 등의 프로세서를 포함할 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다.

[0010] 도 1a에 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 제스처 및/또는 움직임을 캡처, 분석 및 트래킹하여 애플리케이션 내의 여러 컨트롤이나 동작을 실행하거나 또는 아바타나 온-스크린 캐릭터를 애니메이션화할 수 있는 카메라일 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다.

[0011] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템(10)은 사용자(18) 등의 사용자에게 게임 또는 애플리케이션 비주얼 및/또는 오디오를 제공하는 텔레비전, 모니터, HDTV(high-definition television) 등의 오디오비주얼 장치(audiovisual device, 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.

[0012] 도 1a 및 1b에 도시된 바와 같이, 타겟 인식, 분석 및 트래킹 시스템(10)을 사용하여 사용자(18) 등의 인간 타겟을 인식, 분석 및/또는 트래킹할 수 있다. 예를 들어, 사용자(18)를 캡처 장치(20)를 통해 트래킹하여, 사용자(18)의 제스처 및/또는 움직임을 캡처하여 아바타나 온-스크린 캐릭터를 애니메이션화할 수 있고 또는 이러한 제스처 및/또는 움직임이 컴퓨터 환경(12)에서 실행 중인 애플리케이션에 영향을 주는 제어로써 해석될 수 있다. 따라서, 일 실시예에 따르면, 사용자(18)는 자신의 몸을 움직여 애플리케이션을 제어하거나, 또는 아바타나 온-스크린 캐릭터를 애니메이션화할 수 있다.

[0013] 도 1a 및 1b에 도시된 바와 같이, 예시적인 실시예에서, 컴퓨팅 환경(12)에서 실행되는 애플리케이션이 사용자(18)가 플레이하고 있는 복싱 게임일 수 있다. 예를 들어, 컴퓨팅 환경(12)은 오디오비주얼 장치(16)를 사용하여 사용자에게 복싱 상대(38)의 비주얼 형상(visual representation)을 제공할 수 있다. 컴퓨팅 환경(12)은 오

디오비주얼 장치(16)를 사용하여, 사용자(18)가 그 자신의 움직임으로 제어할 수 있는 플레이어 아바타(40)의 비주얼 형상을 제공할 수도 있다. 예를 들어, 도 1b에 도시된 바와 같이, 사용자(18)는 물리적 공간에서 펀치를 날려, 플레이어 아바타(40)가 게임 공간에서 펀치를 날리게 할 수 있다. 따라서, 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템(10)의 컴퓨팅 환경(12) 및 캡처 장치(20)를 사용하여 물리적 공간에서의 사용자(18)의 펀치를 인식하고 분석하여, 그 펀치가 게임 공간에서 플레이어 아바타(40)의 게임 제어로 해석되거나 또는, 펀치의 모션이 게임 공간에서 플레이어 아바타(40)를 애니메이션화하는 데 사용될 수 있다.

[0014] 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 컨트롤과 같이, 다른 컨트롤이나 동작으로 해석되거나 또는, 플레이어 아바타를 애니메이션화하는 데 사용될 수 있다. 나아가, 일부 움직임들은 플레이어 아바타(40)의 제어가 아닌 동작들에 해당하는 컨트롤로 해석될 수 있다. 예를 들어, 플레이어가 움직임을 통해 게임의 종료, 일시 중지 또는 저장, 레벨 선택, 최고점 보기, 친구와의 통신 등을 할 수 있다. 다른 실시예에 따르면, 플레이어는 움직임을 사용하여 메인 사용자 인터페이스로부터 게임 또는 다른 애플리케이션을 선택할 수 있다. 또한, 사용자(18)의 모든 움직임이 애플리케이션과 인터랙션하도록 적절한 방식으로 이용 가능, 사용 및 분석될 수 있다.

[0015] 예시적인 실시예에서, 사용자(18) 등의 인간 타겟은 사물(object)을 가질 수 있다. 이러한 실시예에서는, 전자 게임의 사용자가 사물을 쥐어, 플레이어와 사물의 모션이 게임의 파라미터를 조정 및/또는 제어하는 데에 사용될 수 있다. 예를 들어, 라켓을 쥔 플레이어의 모션이 트래킹되어, 전자 스포츠 게임의 온-스크린 라켓을 제어하는 데에 이용될 수 있다. 또 다른 실시예에서, 사물을 쥐고 있는 플레이어의 모션이 트래킹되어, 전자 컴뱃 게임에서 온-스크린 무기를 제어하는 데에 이용될 수 있다.

[0016] 다른 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템(10)은 타겟 움직임을 게임 세상 밖에 있는 운영 체제 및/또는 애플리케이션의 컨트롤로 해석할 수 있다. 예를 들어, 운영 체제 및/또는 애플리케이션의 제어가능한 사실상 모든 양태가 사용자(18) 등의 타겟의 움직임에 의해 제어될 수 있다.

[0017] 도 2는 타겟 인식, 분석 및 트래킹 시스템(10)에서 사용되는 캡처 장치(20)의 예시적인 실시예를 도시한다. 예시적인 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 시선(line of sight)을 따라 연장되는 Z 축에 수직인 레이어들로 깊이 정보를 체계화시킬 수 있다.

[0018] 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.

[0019] 도 2에 도시된 바와 같이, 예시적인 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 예시적인 실시예에서는, 나가는 광파의 위상과 들어오는 광파의 위상을 비교하여 위상 변위(phase shift)를 결정할 수 있다. 이후, 이 위상 변위를 사용하여 캡처 장치로부터 타겟 또는 사물의 특정 지점까지의 물리적인 거리를 결정할 수 있다.

[0020] 또 다른 예시적인 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.

[0021] 또 다른 예시적인 실시예에 따르면, 캡처 장치(20)는 구조광을 사용하여 깊이 정보를 캡처할 수 있다. 이러한 분석법에서는, 패턴화된 광(즉, 격자 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.

[0022] 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.

[0023] 캡처 장치(20)는 마이크(30)를 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 사용할 수 있다. 게다가, 컴퓨팅 환경(12)에서 실행가능한 게임 애플리케이션, 비게임 애플리케이션 등의 애플리케이션들을 제어하기 위해 사용자로부터 제공되는 오디오 신호를 수신하기 위해 마이크(30)를 사용할 수 있다.

[0024] 예시적인 실시예에서, 캡처 장치(20)는 이미지 카메라 컴포넌트(22)와 통신 가능한 논리 프로세서(32)를 더 포함할 수 있다. 프로세서(32)는 예를 들어, 깊이 이미지를 수신하고, 그 깊이 이미지에 근거하여 복셀 격자를 생성하고, 복셀 격자에 포함된 배경을 제거하여 인간 타겟에 관련된 하나 이상의 복셀을 분리하고, 분리된 인간 타겟의 하나 이상의 말단의 위치나 포지션을 결정하여, 그 하나 이상의 말단의 위치나 포지션에 근거하여 모델을 조정하기 위한 인스트럭션 또는 그 밖의 임의의 적절한 인스트럭션을 포함하는 인스트럭션들을 실행할 수 표준화된 프로세서, 특화된 프로세서, 마이크로프로세서 등을 포함할 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다.

[0025] 캡처 장치(20)는, 프로세서(32)로 실행가능한 인스트럭션, 3-D 카메라 또는 RGB 카메라로 캡처된 이미지 또는 이미지 프레임, 또는 기타 임의의 적절한 정보, 이미지 등을 저장하는 메모리 컴포넌트(34)를 더 포함할 수 있다. 예시적인 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.

[0026] 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 캡처할지를 결정하는 데에 사용되는 클럭을 통신 링크(36)를 통해 캡처 장치(20)로 제공할 수 있다.

[0027] 또한, 캡처 장치(20)는 예를 들어, 3-D 카메라(26) 및/또는 RGB 카메라(28)에서 캡처되는 깊이 정보 및 이미지, 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 모델, 깊이 정보 및 캡처된 이미지를 사용하여, 예컨대, 게임 또는 워드 프로세서 등의 애플리케이션을 제어하고 아바타 또는 온-스크린 캐릭터를 애니메이션화할 수 있다. 예를 들어, 도 2에 도시된 바와 같이, 컴퓨팅 환경(12)은 제스처 라이브러리(190)를 포함할 수 있다. 제스처 라이브러리(190)는 (사용자가 움직임에 따라) 골격 모델에 의해 실행되는 제스처에 관한 정보를 각각 포함하는 일련의 제스처 필터를 포함할 수 있다. 카메라(26, 28)와 캡처 장치(20)에 의해 캡처된 골격 모델 및 그에 연관된 움직임 형태의 데이터를 제스처 라이브러리(190)의 제스처 필터와 비교하여 (골격 모델로 표현된 바와 같이) 사용자가 하나 이상의 제스처를 언제 수행했는지를 식별할 수 있다. 이들 제스처는 애플리케이션의 각종 컨트롤에 연관될 수 있다. 따라서, 컴퓨팅 환경(12)은 제스처 라이브러리(190)를 사용하여 골격 모델의 움직임을 해석하고 그 움직임에 기반하여 애플리케이션을 제어할 수 있다.

[0028] 도 3은 타겟 인식, 분석 및 트래킹 시스템에서 여러 제스처들을 해석하고 이 시스템에 의해 디스플레이되는 아바타 또는 온-스크린 캐릭터를 애니메이션화하는 데에 사용되는 컴퓨팅 환경의 예시적인 실시예를 도시한다. 도 1a - 2와 관련하여 진술한 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 임시로 데이터를 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.

- [0029] 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 프로세서의 액세스를 용이하게 한다.
- [0030] 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.
- [0031] 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, 재생 등을 위해 미디어 드라이브(144)를 통해 액세스될 수 있다. 미디어 드라이브(144)는, 직렬 ATA 버스 또는 기타 고속 접속(예를 들어, IEEE 1394)과 같은 버스를 통해 I/O 제어기(120)에 접속된다.
- [0032] 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.
- [0033] 전면 패널 I/O 서브어셈블리(130)는 전원 버튼(150), 배출 버튼(eject button, 152), 및 멀티미디어 콘솔(100)의 외부 표면에 노출된 LED들(light emitting diodes) 또는 기타 인디케이터의 기능을 지원한다. 시스템 전원 모듈(136)은 멀티미디어 콘솔(100)의 컴포넌트들에 전력을 공급한다. 팬(fan, 138)은 멀티미디어 콘솔(100) 내의 회로를 냉각시킨다.
- [0034] 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.
- [0035] 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 네비게이션할 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.
- [0036] 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.
- [0037] 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.
- [0038] 특히, 개시 커널(launch kernel), 동시(concurrent) 시스템 애플리케이션 및 드라이버를 포함할 수 있을 정도로

메모리 예약이 충분히 큰 것이 바람직하다. 예약된 CPU 사용량이 시스템 애플리케이션에서 사용되지 않는 경우, 휴지 스레드(idle thread)가 미사용 사이클들을 쓸 수 있도록 CPU 예약이 일정한 것이 바람직하다.

[0039] GPU 예약과 관련하여, 팝업(popup)이 오버레이(overlay)로 렌더링되도록 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.

[0040] 멀티미디어 콘솔(100)이 부팅되고 시스템 리소스가 예약된 후에, 동시 시스템 애플리케이션이 실행되어 시스템 기능들을 제공한다. 시스템 기능들은, 상기에서 설명한 예약된 시스템 리소스들 내에서 실행되는 일련의 시스템 애플리케이션에서 캡슐화되어 있다. 운영 체제 커널은 시스템 애플리케이션 스레드인지 게임 애플리케이션 스레드인지를 식별한다. 일관적인 시스템 리소스 뷰를 애플리케이션에 제공하기 위해, 시스템 애플리케이션은 사전 설정된 시간 및 간격으로 CPU(101)에서 실행되도록 스케줄링되는 것이 바람직하다. 스케줄링은 콘솔에서 실행되는 게임 애플리케이션에 대한 캐시 중단을 최소화하기 위한 것이다.

[0041] 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.

[0042] 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.

[0043] 도 4는 타겟 인식, 분석 및 트래킹 시스템에서의 여러 체크처들의 해석, 및/또는 타겟 인식, 분석 및 트래킹 시스템에 의해 디스플레이되는 아바타 또는 온-스크린 캐릭터의 애니메이션화를 위해 사용되는 도 1a - 2에 도시된 컴퓨팅 환경(12)의 또 다른 실시예인 컴퓨팅 환경(220)을 도시한다. 컴퓨팅 시스템 환경(220)은 적합한 컴퓨팅 환경의 일례일 뿐, 본원에 개시된 대상의 사용이나 기능의 범위를 제한하는 것은 아니다. 컴퓨팅 환경(220)은, 예시적인 운영 환경(220)에 도시된 임의의 컴포넌트 또는 컴포넌트 조합에 관한 의존성이나 요구 사항을 가지는 것으로 해석되어서는 안 된다. 몇몇 실시예에서, 도시된 다양한 컴퓨팅 구성요소들은 본 개시의 특정 양태들에 대한 예를 들어 설명하는 회로를 포함할 수 있다. 예를 들어, 본 개시에서 사용되는 회로라는 용어는 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 예시적인 실시예에 있어서, 구현자는 논리를 구현하는 소스 코드를 작성하고, 이 소스 코드는 범용 처리 장치에서 처리될 수 있는 기계 판독가능 코드로 컴파일될 수 있다. 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.

[0044] 도 4에서, 컴퓨팅 환경(220)은, 통상적으로 각종 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 통상적으로 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.

- [0045] 또한, 컴퓨터(241)는 다른 이동식/비이동식, 휘발성/비휘발성 컴퓨터 저장 매체를 포함할 수 있다. 예를 들어, 도 4는 비이동식, 비휘발성 자기 매체에 관독 또는 기록하는 하드 디스크 드라이브(238), 이동식, 비휘발성 자기 디스크(254)에 관독 또는 기록하는 자기 디스크 드라이브(239), 및 CD ROM 또는 다른 광학 매체와 같은 이동식, 비휘발성 광 디스크(253)에 관독 또는 기록하는 광 디스크 드라이브(240)를 도시한다. 예시적인 운영 환경에서 사용가능한 다른 이동식/비이동식, 휘발성/비휘발성 컴퓨터 저장 매체는, 자기 테이프 카세트, 플래시 메모리 카드, DVD, 디지털 비디오 테이프, 고체 상태 RAM, 고체 상태 ROM 등을 포함하지만, 이들로서 제한되는 것은 아니다. 하드 디스크 드라이브(238)는 일반적으로 인터페이스(234)와 같은 비이동식 메모리 인터페이스를 통해 시스템 버스(221)에 연결되며, 자기 디스크 드라이브(239) 및 광 디스크 드라이브(240)는 일반적으로 인터페이스(235)와 같은 이동식 메모리 인터페이스에 의해 시스템 버스(221)에 연결된다.
- [0046] 앞서 논의되었으며 도 4에 도시된 드라이브 및 그 관련 컴퓨터 저장 매체는 컴퓨터(241)를 위한 컴퓨터 관독가능 인스트럭션, 데이터 구조, 프로그램 모듈 및 기타 데이터를 저장한다. 도 4에서, 예를 들어, 하드 디스크 드라이브(238)는 운영 체제(258), 애플리케이션 프로그램(257), 다른 프로그램 모듈(256) 및 프로그램 데이터(255)를 저장하는 것으로 도시되어 있다. 이러한 컴포넌트들은 운영 체제(225), 애플리케이션 프로그램(226), 다른 프로그램 모듈(227) 및 프로그램 데이터(228)와 동일할 수도 또는 다를 수도 있음을 알 것이다. 운영 체제(258), 애플리케이션 프로그램(257), 다른 프로그램 모듈(256), 및 프로그램 데이터(255)는, 최소한 이들이 상이한 카피들을 나타내기 위해 본원에서 상이한 번호가 부여된다. 사용자는 키보드(251), 및 일반적으로 마우스, 트랙볼, 또는 터치 패드로 불리는 포인팅 장치(252)를 비롯한 입력 장치들을 통해 명령어 및 정보를 컴퓨터(241)에 입력할 수 있다. 다른 입력 장치들(도시되지 않음)은 마이크, 조이스틱, 게임패드, 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.
- [0047] *컴퓨터(241)는 원격 컴퓨터(246)와 같은 하나 이상의 원격 컴퓨터들에 대한 논리적인 연결들을 사용하여 네트워크 환경에서 동작할 수 있다. 원격 컴퓨터(246)는 개인용 컴퓨터, 서버, 라우터, 네트워크 PC, 피어 장치(peer device) 또는 다른 공통 네트워크 노드일 수 있으며, 단지 메모리 저장 장치(247)만이 도 4에 도시되어 있지만, 일반적으로 컴퓨터(241)와 관련하여 전술한 다수의 또는 모든 구성요소들을 포함한다. 도 4에 도시된 논리적인 연결은 LAN(local area network, 245) 및 WAN(wide area network, 249)을 포함하지만, 또한 다른 네트워크들도 포함할 수 있다. 이러한 네트워크 환경들은 사무실, 기업(enterprise-wide) 컴퓨터 네트워크, 인트라넷 및 인터넷에 흔하다.
- [0048] LAN 네트워크 환경에서 사용될 때, 컴퓨터(241)는 네트워크 인터페이스 또는 어댑터(237)를 통해서 LAN(245)에 연결된다. WAN 네트워크 환경에서 사용될 때, 컴퓨터(241)는 통상적으로 모뎀(250) 또는 인터넷과 같이 WAN(249)을 통해 통신을 구축하기 위한 다른 수단을 포함한다. 내부 또는 외부에 존재할 수 있는 모뎀(250)은 사용자 입력 인터페이스(236), 또는 다른 적절한 메카니즘을 통해 시스템 버스(221)에 연결될 수 있다. 네트워크 환경에서, 컴퓨터(241) 또는 그 일부분에 관련하여 도시된 프로그램 모듈들이 원격 메모리 저장 장치 내에 저장될 수 있다. 예를 들면, 도 4는 메모리 장치(247) 상에 존재하는 원격 애플리케이션 프로그램들(248)을 도시하고 있다. 도시된 네트워크 연결들은 예시적인 것이며, 컴퓨터들 간에 통신 링크를 구축하는 다른 수단이 사용될 수 있음을 이해할 것이다.
- [0049] 도 5는 한 장면의 사용자를 트래킹하는 예시적인 방법(300)의 순서도를 도시한다. 예를 들어, 예시된 방법(300)은 도 1a - 4에 관해 설명된 타겟 인식, 분석 및 트래킹 시스템(10)의 캡처 장치(20) 및/또는 컴퓨팅 환경(12)을 사용하여 구현될 수 있다. 예시적인 실시예에서, 예시된 방법(300)은 도 1a - 4에 관해 설명된 타겟 인식, 분석 및 트래킹 시스템(10)의 캡처 장치(20) 및/또는 컴퓨팅 환경(12)에서 실행가능한 프로그램 코드(즉, 인스트럭션)의 형태를 취할 수 있다.
- [0050] 일 실시예에 따르면, 305에서, 깊이 이미지를 수신할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 도 1a - 2에 관해 전술한 캡처 장치(20)와 같은 캡처 장치를 포함할 수 있다. 캡처 장치는 하나 이상의 타겟을 포함할 수 있는 장면을 캡처하거나 관찰할 수 있다. 예시적인 실시예에서, 캡처 장치는 비행 시간 분석, 구조광 분석, 스테레오 이미지 분석 등 적절한 임의의 기술을 사용하여 장면의 깊이 이미지와 같은 이미지를 얻도록 구

성된 깊이 카메라일 수 있다.

- [0051] 깊이는 이미지는 각각의 화소가 관찰된 깊이 값을 갖는 관찰된 복수 개의 화소일 수 있다. 예를 들어, 깊이는 이미지는 캡처된 장면의 2-D(two-dimensional) 화소 영역을 포함할 수 있고, 2-D 화소 영역의 각 화소는 캡처된 장면의 사물의 센티미터, 밀리미터 등으로 나타낸 캡처 장치로부터의 길이 또는 거리 등의 깊이 값을 가질 수 있다.
- [0052] 도 6은 305에서 수신가능한 깊이 이미지(400)의 예시적인 실시예를 도시한다. 예시적인 실시예에 따르면, 깊이는 이미지(400)는 예컨대, 도 2에 관해 기술한 캡처 장치(20)의 3-D 카메라(26) 및/또는 RGB 카메라(28)로 캡처한 장면의 이미지 또는 프레임일 수 있다. 도 6에 도시된 바와 같이, 깊이는 이미지(400)는 캡처된 장면에도 1a 및 1b에 관해 기술한 사용자(18) 등의 사용자에게 대응하는 인간 타겟(402a)과 벽, 테이블, 모니터 등의 하나 이상의 비인간 타겟(404)을 포함할 수 있다. 기술한 바와 같이, 깊이는 이미지(400)는 관찰된 복수 개의 화소를 포함하고, 관찰된 각 화소는 그에 연관된 관찰된 깊이 값을 갖는다. 예를 들어, 깊이는 이미지(400)는 캡처된 장면의 2-D(two-dimensional) 화소 영역을 포함할 수 있고, 2-D 화소 영역의 특정 X-값 및 Y-값에서의 각 화소는 캡처된 장면의 타겟 또는 사물의 센티미터, 밀리미터 등으로 나타낸 캡처 장치로부터의 길이 또는 거리 등의 깊이 값을 가질 수 있다.
- [0053] 일 실시예에서는, 깊이는 이미지(400)는 그 깊이는 이미지의 화소들의 상이한 색상들이 캡처 장치로부터의 인간 타겟 및 비인간 타겟의 상이한 거리들에 상응하거나 또는 이를 시각적으로 표현하도록 컬러화될 수 있다. 예를 들어, 깊이는 이미지에서 캡처 장치에 가장 가까이에 있는 타겟에 관련된 화소들은 적색 및/또는 주황색을 내는 한편, 멀리 떨어진 타겟에 관련된 화소들은 깊이 이미지에서 녹색 및/또는 청색을 내게 된다.
- [0054] 다시 5를 참조하면, 일 실시예에서, 305에서 이미지를 수신하면, 깊이는 이미지에서 높은 분산(high-variance) 및/또는 노이즈가 있는 하나 이상의 깊이 값을 제거하거나 스무딩하고(smooth), 누락 및/또는 제거된 깊이 정보 부분을 채우거나 복원하며, 또는 수신된 깊이는 이미지에 기타 적절한 임의의 처리를 하여, 깊이는 이미지에 관련된 깊이 정보를 사용하여 골격 모델 등의 모델을 생성할 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다.
- [0055] 예시적인 실시예에 따르면, 310에서, 수신된 깊이는 이미지에 기반하여 하나 이상의 복셀이 있는 격자를 생성할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 수신된 깊이는 이미지에 포함된 정보를 사용하여 하나 이상의 복셀을 생성함으로써 수신된 깊이는 이미지를 다운샘플링하여, 다운샘플링된 깊이는 이미지를 생성할 수 있다. 일 실시예에서, 하나 이상의 복셀은 서브-샘플링된 격자 상에서, 수신된 깊이는 이미지에 포함된 정보 데이터 또는 값을 나타내는 부피 요소(volume elements)일 수 있다.
- [0056] 예를 들어, 기술한 바와 같이, 깊이는 이미지는 캡처된 장면의 2-D 화소 영역을 포함할 수 있고, 각 화소는 그에 연관된 X-값, Y-값 및 깊이 값(또는 Z-값)을 가질 수 있다. 일 실시예에서, 2-D 화소 영역의 화소들을 하나 이상의 복셀 격자로 줄임으로써 깊이는 이미지를 다운샘플링할 수 있다. 예를 들어, 깊이는 이미지는 4x4 화소 블록, 5x5 화소 블록, 8x8 화소 블록, 10x10 화소 블록 등의 화소 부분 또는 블록으로 나뉠 수 있다. 현실 세계 공간에서 2-D 깊이는 이미지의 화소에 연관된 부분이나 블록의 위치를 나타내는 깊이는 이미지에 대한 복셀을 생성하도록 각 부분 또는 블록을 처리한다. 예시적인 실시예에 따르면, 복셀이 나타내는 부분 또는 블록의 화소들에 대한 유효(valid) 또는 영이 아닌(non-zero) 깊이 값의 평균 깊이 값, 복셀이 나타내는 부분 또는 블록의 화소의 최소, 최대 및/또는 중간(median) 깊이 값, 복셀이 나타내는 부분 또는 블록의 유효 깊이 값을 갖는 화소들에 대한 X-값 및 Y-값의 평균, 또는 깊이는 이미지로부터 제공되는 그 밖의 임의의 적절한 정보에 기반하여 각 복셀의 위치를 얻을 수 있다. 따라서, 예시적인 실시예에 따르면, 각각의 복셀은, 복셀이 나타내는 부분 또는 블록의 화소들에 대한 유효 또는 영이 아닌 깊이 값의 평균 깊이 값, 복셀이 나타내는 부분 또는 블록의 화소의 최소, 최대 및/또는 중간 깊이 값, 복셀이 나타내는 부분 또는 블록의 유효 깊이 값을 갖는 화소들에 대한 X-값 및 Y-값의 평균, 또는 305에서 수신되는 깊이는 이미지의 해당 부분 또는 블록의 X-값, Y-값 및 깊이 값에 근거하여 깊이는 이미지로부터 제공되는 그 밖의 임의의 적절한 정보 등의 값을 갖는 깊이는 이미지의 서브-블록 부분 또는 블록을 나타낼 수 있다.
- [0057] 일 실시예에서, 다운샘플링된 깊이는 이미지 안에서 하나 이상의 복셀의 격자가 층을 이룰 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 기술한 대로 복셀을 생성할 수 있다. 이후에 타겟 인식, 분석 및 트래킹 시스템은 격자에서, 생성된 하나 이상의 다른 복셀 위에 생성된 복셀을 쌓을 수도 있다.
- [0058] 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 깊이는 이미지에서 캡처될 수 있는 장면의 사물의 테두리(edges) 주변의 격자에서 복셀을 쌓을 수 있다. 예를 들어, 305에서 수신된 깊이는 이미지는 인간 타겟

과 벽 등의 비인간 타겟을 포함할 수 있다. 인간 타겟은 예컨대, 인간 타겟의 테두리에서 벽 등의 비인간 타겟에 오버랩될 수 있다. 일 실시예에서, 오버랩된 테두리는 깊이 이미지에서 캡처될 수 있는 인간 타겟과 비인간 타겟에 연관된 깊이 값, X-값, Y-값 등의 정보를 포함할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 오버랩된 테두리에서 인간 타겟에 연관된 복셀 및 비인간 타겟에 연관된 복셀을 생성하여, 복셀들이 쌓이고 격자가 오버랩 테두리의 깊이 값, X-값, Y-값 등의 정보를 보유할 수 있게 된다.

[0059] 또 다른 실시예에 따르면, 305에서 수신되는 깊이 이미지의 화소에 대한 깊이 값, X-값, Y-값 등의 정보를 3-D(three-dimensional) 공간에 투사함으로써, 310에서 하나 이상의 복셀이 있는 격자를 생성할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 카메라, 이미지 또는 원근법 변환(perspective transform) 등의 변환을 통해 깊이 이미지의 화소에 대한 깊이 값, X-값, Y-값 등의 정보를 3-D 공간의 3-D 포인트에 매핑시킴으로써, 정보가 3-D 공간에서 사다리꼴이나 피라미드형으로 변환될 수 있다. 일 실시예에서, 사다리꼴이나 피라미드형의 3-D 공간은 복셀 격자를 생성할 수 있는 큐브(cubes) 등의 블록으로 분할될 수 있고, 블록 또는 큐브 각각은 격자의 복셀을 나타낼 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 깊이 이미지의 사물에 대응하는 3-D 포인트 위에 3-D 격자를 중첩시킬 수 있다. 그리고 나서, 타겟 인식, 분석 및 트래킹 시스템은 격자를 복셀을 나타내는 블록으로 분할하거나 쪼개어, 깊이 이미지를 낮은 해상도(resolution)로 다운샘플링할 수 있다. 예시적인 실시예에 따르면, 격자 내의 각 복셀은, 복셀이 나타내는 격자의 3-D 공간에 연관된 화소들의 유효 또는 영이 아닌 깊이 값의 평균 깊이 값, 복셀이 나타내는 격자의 3-D 공간에 연관된 화소들의 최소 및/또는 최대 깊이 값, 복셀이 나타내는 격자의 3-D 공간에 연관된 유효 깊이 값을 갖는 화소들에 대한 X-값 및 Y-값의 평균, 또는 깊이 이미지로부터 제공되는 그 밖의 임의의 적절한 정보를 포함할 수 있다.

[0060] 도 7a 및 7b는 다운샘플링된 이미지의 일부의 예시적인 실시예를 도시한다. 예를 들어, 도 7a에 도시된 대로, 도 6에 관해서 전술한 깊이 이미지(400)의 일부(410)가 복수의 화소(420)를 포함하고, 각 화소는 그에 연관된 X-값, Y-값 및 깊이 값(또는 Z-값)을 포함할 수 있다. 일 실시예에 따르면, 전술한 바와 같이, 2-D 화소 영역의 화소들을 하나 이상의 복셀이 있는 격자로 줄임으로써 깊이 이미지(400) 등의 깊이 이미지를 다운샘플링할 수 있다. 예를 들어, 도 7a에 도시된 대로, 깊이 이미지(400)의 일부(410)는 화소(420)의 8x8 블록과 같이 화소(420)의 부분 또는 블록(430)으로 분할될 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 부분 또는 블록(430)을 처리하여, 도 7a 및 7b에 도시된 바와 같이 현실 세계 공간의 화소(420)에 연관된 부분 또는 블록(430)의 위치를 나타내는 복셀(440)을 생성할 수 있다.

[0061] 다시 도 5를 참조하면, 315에서, 다운샘플링된 깊이 이미지로부터 배경을 제거할 수 있다. 예를 들어, 다운샘플링된 깊이 이미지의 비인간 타겟 또는 사물 등의 배경을 제거하여, 사용자와 연관된 인간 타겟 등의 전경 사물을 분리시킬 수 있다. 일 실시예에서, 전술한 바와 같이, 타겟 인식, 분석 및 트래킹 시스템은 캡처되거나 관찰된 깊이 이미지에 대한 하나 이상의 복셀이 있는 격자를 생성함으로써, 캡처되거나 관찰된 깊이 이미지를 다운샘플링할 수 있다. 타겟 인식, 분석 및 트래

[0062] 킹 시스템은 다운샘플링된 깊이 이미지의 각각의 복셀을 분석하여, 복셀이 깊이 이미지의 하나 이상의 비인간 타겟 등의 배경 사물의 연관 여부를 결정할 수 있다. 복셀이 배경 사물과 연관이 있으면, 다운샘플링된 깊이 이미지로부터 그 복셀을 제거하거나 버려, 인간 타겟과 같은 전경 사물 및 전경 사물에 연관된 격자의 하나 이상의 복셀을 분리시킬 수 있다.

[0063] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 각 복셀을 분석하여 그에 연관된 사물을 결정할 수 있다. 예를 들어, 전술한 바와 같이, 도 6에서 전술한 깊이 이미지(400) 등의 깊이 이미지로써 305에서 관찰되거나 캡처될 수 있는 장면은 복수의 사물을 포함할 수 있다. 이들 사물로는 하나 이상의 인간 타겟 및/또는 벽, 테이블, 소파, 램프 등 하나 이상의 비인간 타겟을 들 수 있다. 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 격자의 각 복셀을 분석하여, 복셀이 장면의 어떤 사물에 연관되었는지를 결정함으로써, 315에서 장면의 각 사물에 연관된 복셀들을 식별할 수 있다. 따라서, 예시적인 실시예에 따르면, 인간 타겟 또는 사람이 장면의 벽 앞에서 있는 경우, 타겟 인식, 분석 및 트래킹 시스템은 각 복셀을 분석하여 각 복셀이 인간 타겟 또는 벽에 연관이 있는지를 결정할 수 있다.

[0064] 복셀이 장면의 어떤 사물과 연관이 있는지를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 복셀이 나타내는 블록 또는 부분의 화소들에 대한 유효 또는 영이 아닌 깊이 값의 평균 깊이 값, 복셀이 나타내는 부분 또는 블록의 화소의 최소 및/또는 최대 깊이 값, 복셀이 나타내는 부분 또는 블록의 유효 깊이 값을 갖는 화소들에 대한 X-값 및 Y-값의 평균, 또는 인접한 또는 근처의 복셀의 기타 임의의 적절한 정보와 같은 값들을 비교할 수 있다. 예를 들어, 일 실시예에서, 격자에서 분석 중인 특정 복셀에 연관된 평균 깊이 값을 그 특정 복셀에

인접한 각 복셀의 평균 깊이 값과 비교할 수 있다. 분석 중인 특정 복셀의 평균 깊이 값과 인접 복셀의 평균 깊이 값의 차이가 임계치보다 작으면, 특정 복셀과 인접 복셀은 동일한 사물에 속하는 것으로 볼 수 있다. 분석 중인 특정 복셀의 평균 깊이 값과 인접 복셀의 평균 깊이 값의 차이가 임계치보다 크면, 특정 복셀과 인접 복셀이 별개의 사물에 속하는 것으로 볼 수 있다. 예시적인 실시예에 따르면, 임계치는 타겟 인식, 분석 및 트래킹 시스템에서 생성되는 사전결정된 값일 수 있고, 이는 복셀이 동일한 사물의 일부일 가능성 또는 확률에 기반할 수 있다. 따라서, 예시적인 실시예에 따르면, 깊이 이미지에서 캡처되거나 관찰된 장면에서 인간 타겟이나 사람이 벽 앞에 서 있는 경우, 타겟 인식, 분석 및 트래킹 시스템은 깊이 이미지에 대해 생성되는 각 복셀을 분석하여 복셀이 인간 타겟 또는 벽에 연관되었는지를 결정할 수 있다.

[0065] 타겟 인식, 분석 및 트래킹 시스템은 수신된 깊이 이미지의 장면에서 사물과 그에 연관된 복셀을 식별한 후에, 식별된 각각의 사물에 연관된 정보를 계산할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 식별된 각각의 사물에 대한 최대 세계 공간 포지션, 최소 세계 공간 포지션, 평균 세계 공간 포지션 등을 계산할 수 있다.

[0066] 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 315에서, 장면에서 식별된 하나 이상의 사물이 그 장면의 다른 사물들과 합쳐져야만(merged) 하는지도 결정할 수 있다. 예를 들어, 305에서 수신된 깊이 이미지에서 사물의 일부 또는 부분이 그 사물의 다른 일부나 부분과 분리될 수 있다. 일 실시예에 따르면, 사물의 일부 또는 부분이 예컨대, 장면의 그 사물, 다른 사물 등으로 인해 드리워진 적외선 그림자에 의해 그 사물의 다른 일부나 부분과 분리될 수 있다. 다른 실시예에서, 사물의 일부 또는 부분은 예컨대, 그 사물과 연관된 색상, 질감, 패턴 등에 의해 사물의 다른 일부 또는 부분과 분리될 수 있다. 예를 들어, 인간 타겟의 머리는 예컨대, 안면 털, 다양한 의류 품목 등에 의해 Y-방향으로 Y-면을 따라 인간 타겟의 상체와 분리될 수 있다.

[0067] 장면에서 식별된 사물이 실제로 장면에서 식별된 다른 사물의 일부 또는 부분인지 여부를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 사물에 연관된 복셀들의 X-값 및 깊이 값과 근처 사물에 연관된 복셀들의 X-값 및 깊이 값을 비교할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 장면에서 식별된 제 1 사물에 연관된 하나 이상의 복셀의 X-값 및 깊이 값과 제 1 사물 근처에 또는 그에 인접한 제 2 사물에 연관된 하나 이상의 복셀의 X-값 및 깊이 값을 비교할 수 있다. 따라서, 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 장면의 복셀들을 분석하여, 제 1 및 제 2 사물이 X-방향으로 정의된 X-면 및/또는 Z-방향으로 정의된 Z-면을 따라서 오버랩되어, 제 1 및 제 2 사물이 합쳐져 동일한 사물의 일부 또는 부분으로 식별되는지를 결정할 수 있다.

[0068] 일 실시예에 따르면, 제 1 사물과 연관된 하나 이상의 복셀의 X-값 및 깊이 값과 제 2 사물과 연관된 하나 이상의 복셀의 X-값 및 깊이 값이 오버랩되는 경우, 타겟 인식, 분석 및 트래킹 시스템은 제 1 및 제 2 사물을 합쳐, 제 1 및 제 2 사물을 공통 사물의 일부 또는 부분으로 식별할 수 있다. 예를 들어, 제 1 사물에 연관된 제 1 복셀이 X-방향을 따라 X-값으로 5 및 제 1 사물의 오른쪽 바깥 테두리에서 깊이 값으로 10 mm를 갖고 있고, 제 2 사물에 연관된 제 2 복셀이 X-방향을 따라 X-값으로 3 및 제 2 사물의 왼쪽 바깥 테두리에서 깊이 값으로 10 mm를 갖고 있는 경우, 타겟 인식, 분석 및 트래킹 시스템은 제 1 및 제 2 사물이 오버랩한다고 결정할 수 있다. 그리고 나서, 타겟 인식, 분석 및 트래킹 시스템은 제 1 및 제 2 사물을 합쳐, 제 1 및 제 2 사물을 동일한 사물의 일부 또는 부분으로 식별할 수 있다.

[0069] 또한, 장면에서 식별된 사물이 실제로 장면에서 식별된 다른 사물의 일부 또는 부분인지 여부를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 사물에 대해 정의된 경계 상자가 그 장면의 다른 사물의 경계 상자와 오버랩되는지를 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 식별된 사물 각각에 대해 경계 상자를 정의할 수 있다. 이후에 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 전술한 바와 같이 그 안에 포함된 하나 이상의 복셀의 X-값, Y-값 및/또는 깊이 값에 근거하여 하나 이상의 사물의 경계 상자가 오버랩되는지 여부를 결정할 수 있다.

[0070] 다른 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 사물에 포함된 복셀들의 X-값, Y-값 및 깊이 값의 평균을 내어 각 사물의 중앙 또는 중심을 결정할 수 있다. 다음으로, 타겟 인식, 분석 및 트래킹 시스템은 장면의 사물들의 중심 또는 중앙 간의 거리를 결정하여, 장면에서 식별된 사물이 실제로 그 장면에서 식별된 다른 사물의 일부 또는 부분인지 여부를 결정할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 사물 간의 거리에 근거하여 하나 이상의 사물을 합칠 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 제 1 사물의 중심 또는 중앙과 제 2 사물의 중앙 또는 중심 간의 거리를 결정할 수 있다. 제 1 사물과 제 2 사물의 중심 또는 중앙 간의 거리가 제 1 및 제 2 사물이 합쳐져야만 함을 가리키는 사전결정된 범위 내에

있다면, 타겟 인식, 분석 및 트래킹 시스템은 사물들을 합쳐, 제 1 및 제 2 사물을 동일한 사물의 일부 또는 부분으로 식별할 수 있다.

[0071] 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 315에서 장면에서 식별된 하나 이상의 사물이 분리되어야 하는지도 결정할 수 있다. 예를 들어, 315에서 장면에서 식별된 사물이 실제로 별개의 두 사물일 수 있다. 장면의 사물이 분리되어야 하는지를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 이전에 수신된 프레임에서 결정된 각 사물의 중심의 위치를 확인할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 이전에 수신된 프레임의 사물들로부터 결정된 중심의 위치에서부터 시작하여, 305에서 수신된 프레임의 깊이 이미지에 대해 생성된 장면의 복셀들을 동시에 플러드 필링할 수 있다(floodfill). 그 후에, 타겟 인식, 분석 및 트래킹 시스템은 사물들에 대한 이전 위치를 사용하여 플러드 필링된 복셀들이 이전에 수신된 프레임의 어떤 사물에 더 가까운지를 결정할 수 있다. 플러드 필링된 복셀들이 이전에 수신된 프레임에서 식별된 다른 사물에 더 가깝다면, 타겟 인식, 분석 및 트래킹 시스템은 315에서 사물을 나눌 수 있다.

[0072] 315에서, 타겟 인식, 분석 및 트래킹 시스템은 식별된 사물이 비인간 타겟 등의 배경 사물인지 또는 인간 타겟 등의 전경 사물인지를 결정할 수 있다. 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 식별된 사물의 이동 또는 움직임 여부에 근거하여 식별된 사물이 배경 사물인지 또는 전경 사물인지를 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 각 복셀에 대한 비-모션(non-motion) 깊이 정보를 포함하는, 장면의 기준 이미지 등의 기준 판(reference plate)를 포함할 수 있다. 예시적인 실시예에 따르면, 기준 판은, 일련의 프레임에 걸쳐 결정되는 격자의 복셀에 대한 최소 X-값, Y-값 및 깊이 값 등의 복셀의 최소 세계 공간 포지션, 일련의 프레임에 걸쳐 결정되는 격자의 복셀에 대한 최대 X-값, Y-값 및 깊이 값 등의 복셀의 최대 세계 공간 포지션, 일련의 프레임에 걸쳐 결정되는 격자의 복셀에 대한 평균 X-값, Y-값 및 깊이 값 등의 복셀의 평균 세계 공간 포지션 또는 기타 임의의 적절한 기준 판을 포함할 수 있다. 다른 실시예에서, 기준 판은 장면의 각 복셀에 연관된 이동 평균(moving average)을 포함할 수 있다. 예를 들어, 이동 평균은 이전에 수신된 일련의 프레임에 걸쳐 결정된 복셀의 평균 깊이 값을 포함할 수 있다.

[0073] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 305에서 수신된 깊이 이미지의 장면에서 식별된 사물들에 연관된 각각의 복셀의 최대 깊이 값, 평균 깊이 값, 최소 깊이 값 등을 비롯한 깊이 정보와 기준 판에 포함된 각 대응 복셀의 비-모션 깊이 정보를 비교할 수 있다. 깊이 정보 및 기준 판의 대응 복셀의 비-모션 깊이 정보의 비교에 따라, 타겟 인식, 분석 및 트래킹 시스템은 복셀을 움직이는 것으로 식별할 수 있다. 예를 들어, 일 실시예에서, 복셀의 최소 깊이 값, 최대 깊이 값 및/또는 평균 깊이 값을 비롯한 깊이 값이 기준 판의 대응 복셀의 이동 평균보다 작아서 복셀이 이동 평균의 앞에 있다면, 복셀은 움직이는 것으로 식별될 수 있다. 다른 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 복셀 및 기준 판의 대응 복셀과 연관된 값들 간의 차이를 계산할 수 있다. 예를 들어, 복셀의 평균 깊이 값, 최대 깊이 값 및/또는 최소 깊이 값 등의 깊이 값과 기준 판의 대응 복셀의 비-모션 정보에 포함된 깊이 값들 간의 차이가 모션 임계치보다 큰 경우, 타겟 인식, 분석 및 트래킹 시스템은 그 복셀을 움직이는 것으로 식별할 수 있다.

[0074] 또 다른 예시적인 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 복셀 및 인접 복셀들의 최대 깊이 값, 평균 깊이 값, 최소 깊이 값 등의 깊이 정보와 기준 판에 포함된 각 대응 복셀의 비-모션 깊이 정보를 비교할 수 있다. 예를 들어, 테두리 노이즈를 처리하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 특정 복셀 및 인접 복셀들의 최소 깊이 값과 기준 판의 대응 복셀을 비교하여, 복셀 및/또는 그에 연관된 사물의 움직임 여부를 결정할 수 있다. 예컨대, 특정 복셀 및 인접 복셀들의 최소 깊이 값과 기준 판의 대응 복셀들의 비-모션 정보에 포함된 최소 깊이 값 간의 차이가 모션 임계치보다 크면, 타겟 인식, 분석 및 트래킹 시스템은 그 특정 복셀을 움직이는 것으로 식별할 수 있다.

[0075] 다음으로, 타겟 인식, 분석 및 트래킹 시스템은 움직이는 복셀들의 평균에 근거하여 각 식별된 사물에 대한 전경 점수(foreground score)를 계산할 수 있다. 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 움직이는 것으로 식별된 섬(island)에 포함된 복셀의 개수를 섬에 포함된 복셀의 전체 개수로 나누어 전경 점수를 계산할 수 있다.

[0076] 타겟 인식, 분석 및 트래킹 시스템은 점수 임계치를 넘는 전경 점수를 갖는 사물을 분리시킬 수 있다. 타겟 인식, 분석 및 트래킹 시스템에 의해 정의된, 값 또는 퍼센트인 점수 임계치는 움직이는 사물을 나타낼 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 전경 점수에 기반하여 다운샘플링된 깊이 이미지로부터 움직이지 않는 배경 사물을 제거하거나 버려, 다운샘플링된 깊이 이미지에서 점수 임계치를 초과하는 전경 점수를 갖는 인간 타겟 등의 전경 사물을 분리할 수 있다. 예시적인 실시예에 따르면, 움직이지 않는 사물을 제거하거나 버리기

위해, 타겟 인식, 분석 및 트래킹 시스템은 X-값, Y-값 및/또는 깊이 값을 제로 값 또는 복셀이 유효하지 않음을 나타내는 다른 적절한 지표나 표시로 대체함으로써 움직이지 않는 사물에 연관된 복셀들을 제거하거나 버릴 수 있다.

[0077] 320에서, 인간 타겟과 같은 분리된 전경 사물에 대해 하나 이상의 신체 부위 등의 하나 이상의 말단을 결정할 수 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 분리된 인간 타겟에 하나 이상의 휴리스틱(heuristics) 기법 또는 규칙을 적용하여, 예컨대, 분리된 인간 타겟에 관련된 중심 또는 중앙, 머리, 어깨, 상체, 팔, 다리 등을 결정할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 말단의 결정을 기초로 하여, 분리된 인간 타겟의 모델을 생성 및/또는 조정할 수 있다. 예를 들어, 305에서 수신된 깊이 이미지가 도 1a - 2에 관해 전술한 캡처 장치(20) 등의 캡처 장치에 의해 관찰되거나 캡처된 최초 프레임에 포함된다면, 골격 모델의 관절을 말단들의 결정된 위치에 할당함으로써 320에서 결정된 중심, 머리, 어깨, 팔, 손, 다리 등의 말단의 위치에 기반하여 모델을 생성할 수 있고, 이는 아래에서 보다 상세하게 설명될 것이다. 또는, 깊이 이미지가 캡처 장치에 의해 관찰되거나 캡처된 다음의 또는 최초가 아닌 프레임에 포함된다면, 320에서 결정된 중심, 머리, 어깨, 팔, 손, 다리 등의 말단의 위치에 기반하여 이전에 생성된 모델을 조정할 수 있다.

[0078] 예시적인 실시예에 따르면, 315에서 인간 타겟 등의 전경 사물이 분리되면, 타겟 인식, 분석 및 트래킹 시스템은 320에서 인간 타겟의 복셀의 평균을 계산하여, 인간 타겟의 중심 또는 중앙을 추정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 중심 또는 중앙의 추정치를 제공하는 인간 타겟에 포함된 복셀의 평균 위치를 계산할 수 있다. 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 그 복셀들에 연관된 X-값, Y-값 및 깊이 값에 근거하여 인간 타겟에 연관된 복셀들의 평균 위치를 계산할 수 있다. 예를 들어, 전술한 바와 같이, 타겟 인식, 분석 및 트래킹 시스템은 복셀에 연관된 화소들의 X-값의 평균을 냄으로써 복셀에 대한 X-값을 계산하고, 복셀에 연관된 화소들의 Y-값의 평균을 냄으로써 복셀에 대한 Y-값을 계산하며, 복셀에 연관된 화소들의 깊이 값의 평균을 냄으로써 복셀에 대한 깊이 값을 계산할 수 있다. 320에서, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟에 포함된 복셀들의 X-값, Y-값 및 깊이 값의 평균을 내어, 인간 타겟의 중심 또는 중앙의 추정치를 제공하는 평균 위치를 계산할 수 있다.

[0079] 도 8은 인간 타겟(402b)에 대해 추정된 중심 또는 중앙의 예시적인 실시예를 도시한다. 예시적인 실시예에 따르면, 중심 또는 중앙의 위치 또는 포지션(802)은 전술한 바와 같이 분리된 인간 타겟(402b)에 연관된 복셀들의 평균 위치 또는 장소에 기반할 수 있다.

[0080] 다시 도 5에서, 타겟 인식, 분석 및 트래킹 시스템은 320에서 인간 타겟에 대한 경계 상자를 정의하여, 예컨대, 인간 타겟의 머리 및/또는 상체를 포함하는 인간 타겟의 중심부 부피를 결정할 수 있다. 예를 들어, 인간 타겟의 중심 또는 중앙의 추정치가 결정되면, 타겟 인식, 분석 및 트래킹 시스템은 X-방향을 따라 수평으로 검색하여 중심부 부피에 관련된 경계 상자를 정의하는 데 사용하는 인간 타겟의 폭(width)을 결정할 수 있다. 일 실시예에 따르면, X-방향을 따라 수평으로 검색하여 인간 타겟의 폭을 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 복셀에 연관된 깊이 값을 포함하지 않는 복셀이나 장면에서 식별된 다른 사물에 연관된 복셀과 같이 유효하지 않은(invalid) 복셀에 도달할 때까지, 중심 또는 중앙으로부터 X-축을 따라 좌우 방향으로 검색할 수 있다. 예를 들어, 전술한 바와 같이, 315에서 배경에 연관된 복셀들을 제거하여, 인간 타겟 및 그에 연관된 복셀들을 분리할 수 있다. 전술한 바와 같이, 예시적인 실시예에 따르면, 315에서 복셀들을 제거하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 배경 사물의 복셀들에 연관된 X-값, Y-값 및/또는 깊이 값을 제로 값 또는 복셀이 유효하지 않음을 나타내는 다른 적절한 지표나 표시로 대체할 수 있다. 320에서, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 왼쪽에 있는 제 1 유효하지 않은 복셀에 도달할 때까지 인간 타겟의 중앙으로부터 왼쪽 방향으로 검색할 수 있으며, 또한 인간 타겟의 오른쪽에 있는 제 2 유효하지 않은 복셀에 도달할 때까지 인간 타겟의 중앙으로부터 오른쪽 방향으로 검색할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 왼쪽 방향에서 도달한 제 1 유효하지 않은 복셀 근처의 제 1 유효 복셀의 X-값과 오른쪽 방향에서의 제 2 유효하지 않은 복셀 근처의 제 2 유효 복셀 간의 차이에 근거하여, 길이를 계산하거나 측정할 수 있다.

[0081] 다음으로, 타겟 인식, 분석 및 트래킹 시스템은 Y-방향을 따라 수직으로 검색하여 중심부 부피에 관련된 경계 상자를 정의하는 데 사용하는 인간 타겟의 키(height), 예컨대 머리부터 엉덩이까지를 결정할 수 있다. 일 실시예에 따르면, Y-방향을 따라 수직으로 검색하여 인간 타겟의 키를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 유효하지 않은(invalid) 복셀이나 또는 복셀에 연관된 X-값, Y-값 또는 깊이 값을 포함하지 않는 복셀에 도달할 때까지, 중심 또는 중앙으로부터 Y-축을 따라 상하 방향으로 검색할 수 있다. 예를 들어, 320에서, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 윗부분에 있는 제 3 유효하지 않은 복셀에 도달할 때까지 인

간 타겟의 중앙으로부터 위쪽 방향으로 검색할 수 있으며, 또한 인간 타겟의 아랫부분에 있는 제 4 유효하지 않은 복셀에 도달할 때까지 인간 타겟의 중앙으로부터 아래쪽 방향으로 검색할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 위쪽 방향에서 도달한 제 3 유효하지 않은 복셀 근처의 제 3 유효 복셀의 Y-값과 아래쪽 방향에서의 제 4 유효하지 않은 복셀 근처의 제 4 유효 복셀 간의 차이에 근거하여, 키를 계산하거나 측정할 수 있다.

[0082] 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 X- 및 Y-축 상에서 X- 및 Y-방향을 따라 30 도, 45 도, 60 도 등 여러 각도에서 대각선으로 검색하여, 중심부 부피에 관련된 경계 상자를 정의하는 데 사용하는 그 밖의 거리 및 값들을 결정할 수 있다.

[0083] 또한, 타겟 인식, 분석 및 트래킹 시스템은 거리 또는 값들의 비율에 근거하여 중심부 부피에 관련된 경계 상자를 정의할 수도 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 전술한 바와 같이 결정된 키와 0.2, 0.25, 0.3 등의 상수 또는 그 밖의 적절한 임의의 값의 곱에 기반하여 경계 상자의 폭을 정의할 수 있다.

[0084] 타겟 인식, 분석 및 트래킹 시스템은 X-축에 따른 수평 방향 검색으로 결정되는 제 1 및 제 2 유효 복셀, Y-축에 따른 수직 방향 검색으로 결정되는 제 3 및 제 4 유효 복셀 또는 예컨대, 대각선 검색으로 결정되는 그 밖의 거리 및 값들에 근거하여 중심부 부피를 나타내는 경계 상자를 정의할 수 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 제 1 유효 복셀의 X-값에서 Y-축을 따라 경계 상자의 제 1 수직선을 생성하고, 제 2 유효 복셀의 X-값에서 Y-축을 따라 경계 상자의 제 2 수직선을 생성할 수 있다. 게다가, 타겟 인식, 분석 및 트래킹 시스템은 제 3 유효 복셀의 Y-값에서 X-축을 따라 경계 상자의 제 1 수평선을 생성하고, 제 4 유효 복셀의 Y-값에서 X-축을 따라 경계 상자의 제 2 수평선을 생성할 수 있다. 예시적인 실시예에 따르면, 제 1 및 제 2 수평선은 제 1 및 제 2 수직선과 교차하여, 인간 타겟의 중심부 부피에 연관된 경계 상자를 나타내는 직사각형 또는 정사각형을 형성할 수 있다.

[0085] 도 9는 중심부 부피를 결정하기 위해 정의된 경계 상자(804)의 예시적인 실시예를 도시한다. 도 9에 도시된 바와 같이, 경계 상자(804)는 제 1 수직선(VL1) 및 제 2 수직선(VL2)와 제 1 수평선(HL1) 및 제 2 수평선(HL2)의 교차점에 기반하여 직사각형을 형성할 수 있다.

[0086] 다시 도 5에서, 타겟 인식, 분석 및 트래킹 시스템은 320에서 인간 타겟의 머리를 결정할 수 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 중심부 부피를 결정하고 그에 연관된 경계 상자를 정의한 후에 인간 타겟의 머리의 위치 또는 포지션을 결정할 수 있다.

[0087] 머리의 포지션 또는 위치를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 머리에 적합한 포지션 또는 위치에서 여러 후보들을 검색하고, 그 여러 후보들에 대해 점수를 매긴 후, 그 점수에 기반하여 여러 후보들로부터 머리의 위치를 선택할 수 있다. 예를 들어, 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 절대적으로 가장 높이 있는 복셀(absolute highest voxel) 및/또는 이 절대적으로 가장 높이 있는 복셀에 인접하거나 근접한 복셀들, 이전 프레임에서 결정된 머리의 위치에 근거하여 증가하는 하나 이상의 복셀, 예컨대, 중심 또는 중앙으로부터 수직으로 연장되는 상향 벡터(upward vector) 상에서 가장 높이 있는 복셀 및/또는 이전 프레임으로부터 결정되는 이전 상향 벡터 상에서 가장 높이 있는 복셀에 인접하거나 근접한 복셀들, 이전 프레임에 대해 결정된 중심 및 가장 높이 있는 복셀 사이의 이전 상향 벡터 상에 있는 가장 높이 있는 복셀, 또는 머리에 대한 후보가 될 수 있는 그 밖의 적절한 임의의 복셀들을 검색할 수 있다.

[0088] 다음으로, 타겟 인식, 분석 및 트래킹 시스템은 후보들의 점수를 매길 수 있다. 일 실시예에 따르면, 3-D 패턴 매칭에 따라 후보들의 점수를 매길 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 머리 원기둥 및 어깨 원기둥을 생성할 수 있다. 그 다음에, 타겟 인식, 분석 및 트래킹 시스템은 머리 원기둥에 포함된, 후보들에 연관된 복셀의 개수에 따라 후보들에 대한 점수를 계산할 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다.

[0089] 도 10은 머리에 관련된 후보들의 점수를 매기기 위해 생성되는 머리 원기둥(806) 및 어깨 원기둥(808)의 예시적인 실시예를 도시한다. 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 머리 원기둥(806) 및 어깨 원기둥(808)에 포함된, 머리 후보들에 연관된 복셀의 개수에 기반하여 후보들의 점수를 매길 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 머리 후보들에 관련된 복셀들의 위치에 기반하여 머리 원기둥(806) 및/또는 어깨 원기둥(808) 안에 있는 머리 후보들의 전체 개수와, 머리 후보들에 관련된 복셀들에 기반하여 머리 원기둥(806)(예컨대, 영역(807) 내) 및/또는 어깨 원기둥(808) 밖에 있는 머리 후보들의 전체 개수를

결정할 수 있다. 또한, 타겟 인식, 분석 및 트래킹 시스템은 어깨 원기둥(808)의 왼쪽 절반(LH)에 있는 머리 후보들의 개수와 어깨 원기둥(808)의 오른쪽 절반(RH)에 있는 머리 후보들의 개수의 차이의 절대값의 함수를 기초로 하여 대칭 메트릭(symmetric metric)을 계산할 수 있다. 예시적인 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 머리 원기둥(806) 및/또는 어깨 원기둥(808) 안에 있는 후보들의 전체 개수에서 머리 원기둥(806) 및/또는 어깨 원기둥(808) 밖에 있는 후보들의 전체 개수를 빼고, 머리 원기둥(806) 및/또는 어깨 원기둥(808) 안팎에 있는 후보들의 전체 개수의 차이에서 대칭 메트릭을 뺌으로써 후보들의 점수를 계산할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 뺄셈 전에, 머리 원기둥(806) 및/또는 어깨 원기둥(808) 안팎에 있는 후보들의 전체 개수에 타겟 인식, 분석 및 트래킹 시스템에 의해 결정된 상수를 곱할 수 있다.

[0090] 다시 도 5를 참조하여, 일 실시예에 따르면, 어느 후보에 관한 점수가 머리 임계 점수를 초과하면, 타겟 인식, 분석 및 트래킹 시스템은 320에서 후보와 연관된 복셀들에 기반하여 머리의 포지션 또는 위치를 결정할 수 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 가장 높이 있는 포인트, 예컨대, 중심 또는 중앙으로부터 수직으로 연장되는 상향 벡터 상에서 가장 높이 있는 복셀 및/또는 이전 프레임에서 결정된 이전 상향 벡터 상에서 가장 높이 있는 복셀에 인접하거나 근접한 복셀들, 이전 프레임에 대해 결정된 중심 및 가장 높이 있는 복셀 사이의 이전 상향 벡터 상에서 가장 높이 있는 복셀, 이전 프레임의 머리의 포지션 또는 위치 주변의 상자, 큐브 등의 영역 내에 있는 모든 복셀들의 평균 위치 등에 기반하여 머리의 포지션 또는 위치를 선택할 수 있다. 다른 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 머리 임계 점수를 초과하는 후보에 연관된 복셀들의 X-값, Y-값 및 깊이 값 등의 값들의 평균을 계산하여, 머리의 포지션 또는 위치를 결정할 수 있고, 또는 타겟 인식, 분석 및 트래킹 시스템은 머리 임계 점수를 초과하는 후보에 포함된 복셀들의 선 적합성(line fit) 또는 최적선(line of best fit)에 기반하여 머리의 포지션 또는 위치를 선택할 수 있다.

[0091] 또한, 일 실시예에서, 둘 이상의 후보가 머리 임계 점수를 초과하는 경우, 타겟, 인식 및 트래킹 시스템은 최고 점수를 갖는 후보를 선택한 후, 그 후보에 관련된 복셀들에 기반하여 머리의 포지션 또는 위치를 결정할 수 있다. 전술한 바와 같이, 타겟 인식, 분석 및 트래킹 시스템은 최고 점수를 갖는 후보에 관련된 복셀들의 X-값, Y-값 및 깊이 값 등의 값들의 평균에 기반하여 머리의 포지션 또는 위치를 선택할 수 있다.

[0092] 일 실시예에 따르면, 후보와 연관된 점수들 중 어느 하나도 머리 임계 점수를 초과하지 않는 경우, 타겟 인식, 분석 및 트래킹 시스템은 머리 점수가 머리 임계 점수를 초과한 이전 프레임의 깊이 값에 연관된 인간 타겟에 포함된 복셀들에 대해 결정된 머리의 이전 포지션 또는 위치를 사용할 수 있고, 305에서 수신된 깊이 이미지가 캡처 장치에서 캡처되거나 관찰된 최초 프레임이라면, 타겟 인식, 분석 및 트래킹 시스템은 T-포즈, 자연스럽게 서 있는 포즈 등과 같이 인간 타겟의 디폴트 포즈에서의 머리에 대한 디폴트 포지션 또는 위치를 사용할 수 있다.

[0093] 다른 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 머리 모양에 연관된 하나 이상의 2-D(two-dimensional) 패턴을 포함할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 후보에 연관된 복셀들이 하나 이상의 2-D 패턴의 머리 모양일 가능성에 기반하여 머리에 관련된 후보들의 점수를 매길 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 머리 모양의 정의를 나타내는 인접 또는 근처 복셀들의 깊이 값을 샘플링할 수 있다. 머리 모양의 정의를 나타내는 복셀들 중 어느 하나의 샘플링된 깊이 값이 머리 모양의 복셀들의 하나 이상의 예상 또는 사전 정의된 깊이 값에서 벗어나는 경우, 타겟 인식, 분석 및 트래킹 시스템은 디폴트 점수나 초기 점수를 낮춰, 그 복셀이 머리가 아님을 나타낼 수 있다. 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 최고치를 갖는 점수를 선택하여, 그 최고 점수를 갖는 후보에 연관된 복셀의 위치 또는 포지션에 기반하여 머리의 위치 또는 포지션을 배치할 수 있다.

[0094] 또 다른 실시예에 따르면, 디폴트 점수 또는 초기 점수는 전술한 바와 같이 머리 및/또는 어깨 원기둥을 사용하여 계산된 머리에 연관된 후보에 대한 점수일 수 있다. 후보가 하나 이상의 2-D 패턴에 연관된 머리 모양이 아닐 경우, 타겟 인식, 분석 및 트래킹 시스템은 이러한 점수를 낮출 수 있다. 전술한 바와 같이, 타겟 인식, 분석 및 트래킹 시스템은 머리 임계 점수를 초과하는 후보의 점수를 선택하여, 그 후보의 위치나 포지션에 기반하여 머리의 위치나 포지션을 배치할 수 있다.

[0095] 나아가, 타겟 인식, 분석 및 트래킹 시스템은 320에서 인간 타겟의 어깨와 엉덩이를 결정할 수 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 머리의 위치나 포지션을 결정한 후에, 인간 타겟의 어깨와 엉덩이의 위치나 포지션을 결정할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 어깨와 엉덩이의 회전(rotation)이나 각도 등 어깨와 엉덩이의 오리엔테이션(orientation)도 결정할 수 있다.

- [0096] 예시적인 실시예에 따르면, 어깨와 엉덩이의 위치나 포지션을 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 머리의 위치 또는 포지션과 인간 타겟의 중심 또는 중앙에 기반해서 헤드-투-센터 벡터를 정의할 수 있다. 예를 들어, 헤드-투-센터 벡터는 머리 포인트의 위치나 포지션에서의 X-값, Y-값 및 깊이 값(또는 Z-값)과 중심 또는 중앙 포인트의 위치나 포지션에서의 깊이 값(또는 Z-값) 사이에서 정의된 벡터 또는 선(line)일 수 있다.
- [0097] 도 11은 머리와 인간 타겟의 중심 또는 중앙에 기반한 헤드-투-센터 벡터의 예시적인 실시예를 도시한다. 전술한 바와 같이, 머리의 위치 또는 포지션(810) 등의 위치 또는 포지션을 결정할 수 있다. 도 11에 도시된 바와 같이, 타겟 인식, 분석 및 트래킹 시스템은 머리의 위치나 포지션(810)과 중심이나 중앙의 위치 또는 포지션(802) 사이에서 헤드-투-센터 벡터(812)를 정의할 수 있다.
- [0098] 도 5를 다시 참조하면, 타겟 인식, 분석 및 트래킹 시스템은 320에서 헤드-투-센터 벡터에 기반하여 어깨 부피 상자 및 엉덩이 부피 상자를 정의할 수 있다. 예를 들어, 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 머리 또는 중앙이나 중심에 연관된 포지션이나 위치 등의 신체 랜드마크로부터의 길이와 같은 변위(displacement)에 기반하여 어깨 및 엉덩이의 위치나 포지션을 정의하거나 결정할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 머리 또는 중앙이나 중심에 연관된 포지션이나 위치 등의 신체 랜드마크로부터의 변위 주위에 어깨 부피 상자 및 엉덩이 부피 상자를 정의할 수 있다.
- [0099] 도 12는 헤드-투-센터 벡터(812)에 근거하여 결정된 어깨 부피 상자(SVB) 및 엉덩이 부피 상자(HVB)의 예시적인 실시예를 도시한다. 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 머리에 연관된 위치나 포지션(810) 또는 중앙이나 중심에 연관된 위치나 포지션(802) 등의 신체 랜드마크로부터의 길이와 같은 변위에 기반하여 어깨 및 엉덩이의 대략적인 위치나 포지션을 정의하거나 결정할 수 있다. 그 후에 타겟 인식, 분석 및 트래킹 시스템은 신체 랜드마크로부터의 변위 값 주변에서 어깨 부피 상자(SVB) 및 엉덩이 부피 상자(HVB)를 정의할 수 있다.
- [0100] 도 5를 다시 참조하면, 타겟 인식, 분석 및 트래킹 시스템은 헤드-투-센터 벡터에 따른, 머리 등의 신체 랜드마크로부터의 길이와 같은 변위 값에 기반하여 어깨 및 엉덩이의 중심을 계산할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 헤드-투-센터 벡터를 따라 변위 값만큼 위아래로 이동하여, 어깨 및 엉덩이의 중심을 계산할 수 있다.
- [0101] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 어깨 및 엉덩이의 각도 등의 오리엔테이션도 결정할 수 있다. 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 어깨 부피 상자 및 엉덩이 부피 상자 내의 깊이 값들의 선 적합성을 계산하여, 어깨 및 엉덩이의 각도 등의 오리엔테이션을 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 어깨 부피 상자 및 엉덩이 부피 상자에 연관된 복셀들의 X-값, Y-값 및 깊이 값들에 기반하여 최적선을 계산하여, 어깨의 중심을 통과하는 어깨뼈를 정의하는 벡터의 어깨 경사(slope)와 엉덩이의 중심을 통과하는 엉덩이의 관절 사이의 엉덩이뼈를 정의하는 벡터의 엉덩이 경사를 계산할 수 있다. 어깨 경사 및 엉덩이 경사는 어깨 및 엉덩이의 각도 등 각각의 오리엔테이션을 정의할 수 있다.
- [0102] 다른 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 320에서 복셀들의 깊이 값이 헤드-투-센터 벡터에 기반하여 인간 타겟의 중심 주변을 반영하도록, 인간 타겟의 깊이 값을 반영할(mirror) 수 있다. 따라서, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 헤드-투-센터 벡터 등의 중심축 벡터(pivot vector), 어깨 및/또는 엉덩이 경계 상자로부터 계산된 중심점(pivot point) 등의 주변에서 인간 타겟의 복셀들의 깊이 값을 반영 시킴으로써 신체의 뒤쪽을 보완할 수 있다.
- [0103] 타겟 인식, 분석 및 트래킹 시스템은 예컨대, 어깨 부피 상자 및 엉덩이 부피 상자 안의 반영된 깊이 값들을 포함하는 깊이 값의 선 적합성을 계산하여, 어깨 및 엉덩이 각도 등의 오리엔테이션을 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 어깨 부피 상자 및 엉덩이 부피 상자에 연관된 반영된 깊이 값을 포함하는 복셀들의 X-값, Y-값 및 깊이 값들에 기반하여 최적선을 계산하여, 어깨의 중심을 통과하는 어깨뼈를 정의하는 벡터의 어깨 경사와 엉덩이의 중심을 통과하는 엉덩이의 관절 사이의 엉덩이뼈를 정의하는 벡터의 엉덩이 경사를 계산할 수 있다. 어깨 경사 및 엉덩이 경사는 어깨 및 엉덩이의 각도 등 각각의 오리엔테이션을 정의할 수 있다.
- [0104] 도 13은 어깨 부피 상자(SVB) 및 엉덩이 부피 상자(HVB)에 근거하여 계산된 어깨 및 엉덩이의 일 실시예를 도시한다. 도 13에 도시된 대로, 어깨의 위치 또는 포지션(816a-b)과 엉덩이의 위치 또는 포지션(818a-b)이 전술한 바와 같이 각각 어깨 부피 상자(SVB) 및 엉덩이 부피 상자(HVB)에 근거하여 결정될 수 있다.
- [0105] 다시 도 5를 참조하면, 320에서, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 상체를 결정할 수 있다. 일

실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 어깨 및 엉덩이를 결정한 후에, 머리, 어깨, 중심 및 엉덩이에 연관되고 이를 둘러싼 복셀을 포함하는 상체 부피를 생성하거나 만들 수 있다. 상체 부피는 중심, 머리, 어깨 및/또는 엉덩이의 위치 또는 포지션에 기반한 원기둥, 둥근 말단을 갖는 원기둥같은 캡슐 형태(pill shape) 동일 수 있다.

[0106] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 어깨, 머리, 엉덩이, 중심 등에 기반한 크기를 갖는 중심부 부피를 나타내는 원기둥을 만들 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 어깨의 폭에 기반한 폭 또는 직경과, 머리와 엉덩이 간의 거리에 기반한 높이를 갖는 원기둥을 만들 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 헤드-투-센터 벡터를 따른 상체 부피를 나타내는 원기둥의 방향과 각을 정해서, 상체 부피가 인간 타겟의 상체의 각도 등의 오리엔테이션을 반영하도록 할 수 있다.

[0107] 도 14는 중심부 부피를 나타내는 원기둥(820)의 예시적인 실시예를 도시한다. 도 14에 도시된 대로, 원기둥(820)은 어깨의 폭에 기반한 폭 또는 직경과, 머리와 엉덩이 간의 거리에 기반한 높이를 가질 수 있다. 원기둥(820)은 헤드-투-센터 벡터(812)를 따라 방향과 각을 정할 수 있다.

[0108] 다시 도 5를 참조하면, 320에서, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 팔다리를 추정하거나 결정할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 상체 부피를 생성하거나 또는 만들어진 후에, 팔다리와 같이 상체 부피 밖에 있는 복셀들을 대략적으로 레이블링할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 상체 부피 밖에 있는 각각의 복셀을 식별하여 그 복셀들을 팔다리(limb)의 일부로 레이블링할 수 있다.

[0109] 타겟 인식, 분석 및 트래킹 시스템은 상체 부피 밖에 있는 복셀들에 연관된 오른팔 및 왼팔, 오른손 및 왼손, 오른쪽 다리 및 왼쪽 다리, 오른발 및 왼발 등의 실제 팔다리를 결정할 수 있다. 일 실시예에서, 실제 팔다리를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 오른팔, 왼팔, 왼쪽 다리, 오른쪽 다리의 이전 포지션 또는 위치 등 식별된 팔다리의 이전 포지션 또는 위치를 상체 부피 바깥의 복셀들의 포지션 또는 위치와 비교할 수 있다. 예시적인 실시예에 따르면, 이전에 식별된 팔다리의 이전 위치 또는 포지션은, 이전 프레임에서 수신된 깊이 이미지에서의 팔다리의 위치 또는 포지션, 이전 움직임에 기반한 투사된 신체 부위 위치 또는 포지션, 또는 인간 타겟의 온전히 이어진 골격 또는 볼륨 모델(fully articulated skeleton or volumetric model)과 같은 인간 타겟의 표상의 기타 적절한 임의의 이전 위치 또는 포지션일 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 이러한 비교에 기반하여, 상체 부피 밖에 있는 복셀들과 이전에 식별된 가장 가까운 팔다리를 연계시킬 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 상체 부피 밖에 있는 각각의 복셀의 X-값, Y-값 및 깊이 값을 포함하는 포지션 또는 위치를, 이전에 식별된 왼팔, 오른팔, 왼쪽 다리, 오른쪽 다리 등 이전에 식별된 팔다리의 X-값, Y-값 및 깊이 값을 포함하는 이전 포지션 또는 위치와 비교할 수 있다. 그 후에 타겟 인식, 분석 및 트래킹 시스템은 이 비교 결과에 따라 상체 볼륨 밖에 있는 각각의 복셀을 가장 가까운 위치나 포지션에 있는 이전에 식별된 팔다리와 연계시킬 수 있다.

[0110] 일 실시예에서, 실제 팔다리를 결정하기 위해, 타겟 인식, 분석 및 트래킹 시스템은 인간 타겟의 표상의 디폴트 포즈에서 오른팔, 왼팔, 오른쪽 다리, 왼쪽 다리 등의 식별된 팔다리의 디폴트 포지션 또는 위치를 상체 부피 밖에 있는 복셀들의 포지션 또는 위치와 비교할 수 있다. 예를 들어, 305에서 수신된 깊이 이미지는 캡처 장치에서 캡처되거나 관찰된 최초 프레임에 포함될 수 있다. 305에서 수신된 깊이 이미지가 최초 프레임에 포함된 경우에, 타겟 인식, 분석 및 트래킹 시스템은 오른팔, 왼팔, 왼쪽 다리, 오른쪽 다리 등의 디폴트 포지션 또는 위치를 비롯한 팔다리의 디폴트 포지션 또는 위치를 상체 부피 밖의 복셀들의 포지션 또는 위치와 비교할 수 있다. 예시적인 실시예에 따르면, 식별된 팔다리의 디폴트 위치 또는 포지션은 디폴트 포즈인 인간 타겟의 온전히 이어진 골격 또는 볼륨 모델을 비롯한 인간 타겟 표상의 T-포즈, 다빈치 포즈, 자연스러운 포즈 등의 디폴트 포즈에서의 팔다리의 위치 또는 포지션일 수 있다. 그 후에 타겟 인식, 분석 및 트래킹 시스템은 이 비교 결과에 따라 상체 볼륨 밖에 있는 각각의 복셀을 디폴트 포즈에 관련된 가장 가까운 팔다리와 연계시킬 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 상체 부피 밖에 있는 각각의 복셀의 X-값, Y-값 및 깊이 값을 포함하는 포지션 또는 위치를 디폴트 왼팔, 오른팔, 왼쪽 다리, 오른쪽 다리 등의 디폴트 팔다리의 X-값, Y-값 및 깊이 값을 포함하는 디폴트 포지션 또는 위치와 비교할 수 있다. 그 후에 타겟 인식, 분석 및 트래킹 시스템은 이 비교 결과에 따라 상체 볼륨 밖에 있는 각각의 복셀을 가장 가까운 위치나 포지션에 있는 디폴트 팔다리와 연계시킬 수 있다.

[0111] 또한 타겟 인식, 분석 및 트래킹 시스템은 추정된 팔다리에 기반하여 상체 부피 내의 복셀들을 다시 레이블링할 수 있다. 예를 들어, 일 실시예에서, 왼쪽 팔뚝과 같이 팔의 적어도 일부분이 인간 타겟의 상체의 앞에 위치할

수 있다. 식별된 팔의 이전 포지션 또는 위치에 기반하여, 타겟 인식, 분석 및 트래킹 시스템은 전술한 바와 같이 그 부분을 팔과 연관된다고 결정하거나 추정할 수 있다. 예를 들어, 이전에 식별된 팔다리의 이전 포지션 또는 위치는, 인간 타겟의 팔 등의 특정 팔다리의 하나 이상의 복셀이 상체 부피 안에 있음을 나타낼 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 이전에 식별된 왼팔, 오른팔, 왼쪽 다리, 오른쪽 다리 등 이전에 식별된 팔다리의 X-값, Y-값 및 깊이 값을 포함하는 이전 포지션 또는 위치를 상체 부피에 포함된 복셀들의 포지션 또는 위치와 비교할 수 있다. 그 후에 타겟 인식, 분석 및 트래킹 시스템은 이 비교 결과에 따라 상체 볼륨 안의 각각의 복셀을 가장 가까운 위치나 포지션에 있는 이전에 식별된 팔다리와 연계시키고 다시 레이블링할 수 있다.

[0112] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 팔다리에 연관된 복셀들을 레이블링을 한 후에, 320에서 레이블링된 팔다리 부위들의 위치 또는 포지션을 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 왼팔, 오른팔, 왼쪽 다리 및/또는 오른쪽 다리에 연관된 복셀들을 레이블링한 후에, 손 및/또는 오른팔 및 왼팔의 팔꿈치, 무릎 및/또는 발, 팔꿈치 등의 위치나 포지션을 결정할 수 있다.

[0113] 타겟 인식, 분석 및 트래킹 시스템은 각 팔다리에 대한 팔다리 평균에 기반하여 손, 팔꿈치, 발, 무릎 등의 부위들의 위치나 포지션을 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 왼팔에 연관된 각각의 복셀에 대한 X-값, 왼팔에 연관된 각각의 복셀에 대한 Y-값, 왼팔에 연관된 각각의 복셀에 대한 깊이 값을 추가하고, 추가된 X-값, Y-값 및 깊이 값 각각의 합을 왼팔에 연관된 복셀들의 전체 개수로 나누어 왼팔 평균 위치를 계산할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 왼쪽 어깨와 왼팔 평균 위치 사이의 벡터 또는 선을 정의하여, 그 벡터 또는 선으로 왼손에 대한 제 1 검색 방향을 정의할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 어깨에서 벡터 또는 선에 의해 정의된 제 1 검색 방향을 따라 마지막 유효 복셀 또는 유효 X-값, Y-값 및/또는 깊이 값을 갖는 마지막 복셀을 검색하여, 마지막 유효 복셀의 위치 또는 포지션을 왼손과 연계시킬 수 있다.

[0114] 다른 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 앵커 포인트(anchor point)를 계산할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 앵커 포인트와 왼팔 평균 위치 사이의 벡터 또는 선을 정의하여, 그 벡터 또는 선으로 왼손에 대한 제 2 검색 방향을 정의할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 앵커 포인트로부터 벡터 또는 선에 의해 정의된 제 2 검색 방향을 따라 마지막 유효 복셀 또는 유효 X-값, Y-값 및/또는 깊이 값을 갖는 마지막 복셀을 검색하여, 마지막 유효 복셀의 위치 또는 포지션을 왼손과 연계시킬 수 있다.

[0115] 예시적인 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 머리, 엉덩이, 어깨 등 결정된 다른 말단들로부터의 하나 이상의 오프셋에 기반하여 앵커 포인트의 위치 또는 포지션을 계산할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 어깨의 위치나 포지션에 연관된 X-값 및 깊이 값의 절반만큼 어깨의 위치 또는 포지션을 X-방향 및 Z-방향으로 각각 연장시켜 앵커 포인트에 대한 X-값 및 깊이 값을 계산할 수 있다. 그리고 나서, 타겟 인식, 분석 및 트래킹 시스템은 연장된 위치 또는 포지션 주변의 앵커 포인트에 대한 X-값 및 깊이 값의 위치나 포지션을 반영시킬 수 있다.

[0116] 타겟 인식, 분석 및 트래킹 시스템은 머리 및/또는 엉덩이로부터의 왼팔 평균 위치의 변위에 근거하여 앵커 포인트에 대한 Y-값을 계산할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 머리의 Y-값 및 왼팔 평균의 Y-값 간의 변위 또는 차이를 계산할 수 있다. 그 다음에, 타겟 인식, 분석 및 트래킹 시스템은 그 변위 또는 차이를 예컨대, 엉덩이의 중심의 Y-값에 더하여 앵커 포인트의 Y-값을 계산할 수 있다.

[0117] 도 15(a) - 15(c)는 앵커 포인트(828a-828c)에 근거하여 결정되는 손의 예시적인 실시예를 도시한다. 도 15(a) - 15(c)에 도시된 바와 같이, 다른 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 앵커 포인트(828a-828c)를 계산할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 앵커 포인트(828a-828c)와 왼팔 평균 위치(826a-826c) 사이의 벡터 또는 선을 정의하여, 그 벡터 또는 선으로 왼손에 대한 제 2 검색 방향을 정의할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 앵커 포인트(828a-828c)로부터 벡터 또는 선에 의해 정의된 제 2 검색 방향을 따라 마지막 유효 복셀 또는 유효 X-값, Y-값 및/또는 깊이 값을 갖는 마지막 복셀을 검색하고, 마지막 유효 복셀의 위치 또는 포지션을 왼손과 연계시킬 수 있다.

[0118] 전술한 바와 같이, 예시적인 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 머리, 엉덩이, 어깨 등 결정된 다른 말단들로부터의 하나 이상의 오프셋(offset)에 기반하여 앵커 포인트(828a-828c)의 위치 또는 포지션을 계산할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 어깨의 위치나 포지션에 연관된 X-값 및 깊이 값의 절반만큼 어깨의 위치 또는 포지션을 X-방향 및 Z-방향으로 각각 연장시켜 앵커 포인트(828a-828c)에 대한 X-값 및 깊이 값을 계산할 수 있다. 그리고 나서, 타겟 인식, 분석 및 트래킹 시스템은 연장된 위치 또는 포지션 주변의 앵커 포인트(828a-828c)에 대한 X-값 및 깊이 값의 위치나 포지션을 반영시킬 수 있다.

- [0119] 타겟 인식, 분석 및 트래킹 시스템은 머리 및/또는 엉덩이로부터의 왼팔 평균 위치의 변위에 근거하여 앵커 포인트(828a-828c)에 대한 Y-값을 계산할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 머리의 Y-값 및 왼팔 평균(826a-826c)의 Y-값 간의 변위 또는 차이를 계산할 수 있다. 그 다음에, 타겟 인식, 분석 및 트래킹 시스템은 그 변위 또는 차이를 예컨대, 엉덩이의 중심의 Y-값에 더하여 앵커 포인트(828a-828c)의 Y-값을 계산할 수 있다.
- [0120] 다시 도 5를 참조하여, 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 320에서 오른손의 위치 또는 포지션을 결정하는 데 사용되는, 전술한 바와 같은 제 1 및 제 2 검색 방향 등의 검색 방향을 정의하기 위해 사용되는 오른팔 평균 위치를 결정할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 왼발 및 오른발을 결정하는 데 사용되는, 전술한 바와 같은 검색 방향을 정의하기 위해 사용되는 왼발 평균 위치 및 오른발 평균 위치도 계산할 수 있다.
- [0121] 도 16은 팔과 다리의 평균 위치 및/또는 앵커 포인트에 근거하여 계산되는 손과 발의 예시적인 실시예를 도시한다. 도 16에 도시된 바와 같이, 양손의 위치 또는 포지션(822a-b)과 양발의 위치 또는 포지션(824a-b)은 전술한 바와 같이 각각의 팔 및 다리 평균 포지션 및/또는 앵커 포인트에 의해 결정되는 제 1 및 제 2 검색 방향에 기반하여 결정될 수 있다.
- [0122] 도 6을 다시 참조하면, 320에서, 타겟 인식, 분석 및 트래킹 시스템은 오른팔 및 왼팔의 평균 위치 및 오른쪽 및 왼쪽 다리의 평균 위치, 어깨, 엉덩이, 머리 등에 기반하여 팔꿈치와 무릎의 위치 또는 포지션을 결정할 수 있다. 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 왼팔 평균 위치의 X-값, Y-값 및 깊이 값을 리파인함으로써(refining) 왼쪽 팔꿈치의 위치 또는 포지션을 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 왼팔에 연관된 테두리를 정의하는 가장바깥쪽 복셀들을 결정할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 왼팔 평균 위치의 X-값, Y-값 및 깊이 값이 테두리의 중심에 또는 테두리로부터 동일한 거리에 있도록 조정할 수 있다.
- [0123] 타겟 인식, 분석 및 트래킹 시스템은 320에서, 분리된 인간 타겟에 대해 추가적인 관심 포인트를 결정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 신체 중심으로부터 가장 먼 복셀, 카메라에 가장 가까운 복셀, 예컨대, 어깨 각도 등의 오리엔테이션에 근거하여 인간 타겟의 가장 앞쪽에 있는 복셀을 결정할 수 있다.
- [0124] 타겟 인식, 분석 및 트래킹 시스템은 320에서 머리, 어깨, 엉덩이, 손, 발 등의 말단에 대해 결정된 하나 이상의 위치 또는 포지션이 인간 타겟의 실제 말단들에 대한 정확한 위치 또는 포지션이 아닌지를 결정할 수 있다. 예를 들어, 일 실시예에서, 오른손의 위치 또는 포지션이 부정확하여, 오른손의 위치 또는 포지션이 어깨 또는 엉덩이의 위치나 포지션에 또는 그 가까이에 붙어있을 수 있다.
- [0125] 예시적인 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 말단의 부정확한 위치나 포지션을 나타내는, 각종 말단에 대한 부피 마커(volume marker) 목록을 포함하거나 저장할 수 있다. 예를 들면, 목록은 손에 연관된 어깨와 엉덩이 주변의 부피 마커들을 포함할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 목록에 있는 손에 연관된 부피 마커들에 근거하여 손에 대한 위치나 포지션이 정확한지 여부를 결정할 수 있다. 예를 들어, 손의 위치나 포지션이 목록의 손에 연관된 어느 하나의 부피 마커 안에 있다면, 타겟 인식, 분석 및 트래킹 시스템은 손의 위치나 포지션이 부정확하다고 결정할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 손의 위치나 포지션을 이전 프레임의 손의 정확한 이전 위치 또는 손의 현재 위치 또는 포지션으로 조정할 수 있다.
- [0126] 325에서, 타겟 인식, 분석 및 트래킹 시스템은 320에서 결정된 말단들에 기반하여 생성된 모델을 트래킹할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 하나 이상의 관절과 그 관절 사이에 정의된 뼈를 갖는 골격 모델을 비롯한 모델을 생성 및/또는 포함할 수 있다.
- [0127] 도 17은 생성된 골격 모델 등의 모델(900)의 예시적인 실시예를 도시한다. 예시적인 실시예에 따르면, 모델(900)은 예컨대, 인간의 3-D 모델을 나타내는 하나 이상의 데이터 구조를 포함할 수 있다. 각 신체 부위는 모델(900)의 관절과 뼈를 정의하는 X, Y 및 Z 값을 갖는 수학적 벡터로 표현될 수 있다.
- [0128] 도 17에 도시된 바와 같이, 모델(900)은 하나 이상의 관절(j1-j16)을 포함할 수 있다. 예시적인 실시예에 따르면, 각각의 관절(j1-j16)에 의해 그들 사이에서 정의된 여러 신체 부위가 다른 여러 신체 부위에 대해 상대적으로 움직일 수 있게 된다. 예를 들어, 인간 타겟을 나타내는 모델은 인접 뼈들의 교차점에 위치하는 관절(j1-j16)을 갖는 "뼈"와 같은 하나 이상의 구조 부재에 의해 정의되는 복수의 딱딱한(rigid) 및/또는 변형가능한 신체 부위를 포함할 수 있다. 관절(j1-j16)을 통해, 뼈와 관절(j1-j16)과 연관된 각종 신체 부위들이 서로 독립적

으로 움직일 수 있게 된다. 예를 들어, 도 17에 도시된 대로, 관절(j10)과 관절(j12) 사이에 정의된 뼈는, 종아리에 해당하는 관절들(j14 및 j16) 사이에서 정의된 뼈와는 독립적으로 움직이는 팔뚝에 해당한다.

- [0129] 도 5를 다시 참조하면, 325에서, 타겟 인식, 분석 및 트래킹 시스템은 320에서 생성된 모델을 말단들에 대해 결정된 위치 또는 포지션에 기반하여 조정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 머리에 연관된 관절(j2)을 320에서 결정된 머리에 대한 위치 또는 포지션(810) 등의 포지션 또는 위치에 대응하도록 조정할 수 있다. 따라서, 예시적인 실시예에서, 전술한 바와 같이 머리에 대해 결정된 위치 또는 포지션(810)에 연관된 X-값, Y-값 및 깊이 값을 관절(j2)에 할당할 수 있다. 예컨대, 전술한 부피 마커 목록에 근거하여 하나 이상의 말단이 부정확한 경우, 타겟 인식, 분석 및 트래킹 시스템은 그 부정확한 관절을 이전 프레임에 근거한 이전 위치나 포지션에 유지시킬 수 있다.
- [0130] *나아가, 타겟 인식, 분석 및 트래킹 시스템이 말단들의 위치 또는 포지션을 갖고 있지 않은 경우, T-포즈, 다빈치 포즈 등의 디폴트 포즈에 기반한 디폴트 위치나 포지션을 사용할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 모델의 하나 이상의 관절(j1-j16)이 디폴트 포즈에서의 가장 가까운 복셀들의 X-값, Y-값 및/또는 깊이 값에 연계되도록 이 관절들을 끌어들이거나 조정할 수 있다.
- [0131] 325에서, 타겟 인식, 분석 및 트래킹 시스템은 330에서 결정된 하나 이상의 신체 측정치에 기반하여, 모델의 관절 사이에서 정의된 하나 이상의 뼈의 측정치를 조정할 수 있으며, 이는 아래에서 보다 상세하게 설명될 것이다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 330에서 인간 타겟의 왼쪽 팔뚝의 길이를 결정할 수 있다. 325에서, 타겟 인식, 분석 및 트래킹 시스템은 330에서 인간 타겟에 대해 결정된 왼쪽 팔뚝의 길이를 모방하도록, 왼쪽 팔뚝에 연관된 뼈의 길이를 조정할 수 있다. 예를 들어, 타겟 인식, 분석 및 트래킹 시스템은 관절(j10 및 j12)의 X-값, Y-값 및 깊이 값(또는 Z-값) 중에서 하나 이상을 조정하여 그 사이에서 정의된 뼈가 330에서 인간 타겟의 왼쪽 팔뚝에 대해 결정된 길이와 동일하게 할 수 있다.
- [0132] 325에서, 타겟 인식, 분석 및 트래킹 시스템은 조정된 모델에 연관된 관절들의 유효하지 않은 위치 또는 포지션을 확인할 수 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 관절(j10) 등의 관절이 튀어나와 모델이 닭춤 포즈에 빠져있는지를 결정하기 위해 확인할 수 있다. 따라서, 325에서, 타겟 인식, 분석 및 트래킹 시스템은 팔꿈치에 연관된 관절을 비롯하여 부적절한 방식으로 무너진 모델의 알려져 있는 위치 또는 포지션에 관해 모델을 확인할 수 있다.
- [0133] 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 305에서 수신된 다운샘플링되지 않은 깊이 이미지의 2-D 화소 영역에서의 X-값, Y-값 및 깊이 값에 기반하여 관절의 위치나 포지션을 리파인할 수도 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 다운샘플링되지 않은 깊이 이미지의 데이터를 사용하여 예컨대, 무너진 모델의 관절들의 위치나 포지션을 리파인할 수 있다.
- [0134] 또한, 타겟 인식, 분석 및 트래킹 시스템은 다운샘플링되지 않은 깊이 이미지의 데이터를 사용하여 자주 사용되는 제스처에 관련된 모델의 관절들의 위치나 포지션을 리파인할 수 있다. 예를 들어, 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 손에 관련된 관절들에 우선 순위를 둘 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 305에서 수신된 다운샘플링되지 않은 깊이 이미지에서 손 주변의 데이터를 localize(국부화)하여, 305에서 수신된 다운샘플링되지 않은 깊이 이미지의 높은 해상도 데이터를 사용하여 320에서 결정된 손의 위치 또는 포지션을 수정할 수 있다.
- [0135] 전술한 바와 같이, 330에서, 타겟 인식, 분석 및 트래킹 시스템은 분리된 인간 타겟에 연관된 복셀들을 스캔하여 그와 관련된 말단들의 크기를 결정할 수 있다. 예를 들어, 분리된 인간 타겟을 스캔하여, 예컨대, 팔, 다리, 머리, 어깨, 엉덩이, 상체 등의 말단과 연관된 길이, 폭 등의 치수를 결정할 수 있다.
- [0136] 크기를 결정하기 위해, 330에서, 타겟 인식, 분석 및 트래킹 시스템은 각각의 말단에 대한 제시된 관절 치수의 예측치(estimator)를 생성할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 예측치를 사용하여 제시된 각 관절 치수에 대한 평균 및 표준 편차를 계산할 수 있다. 타겟 인식, 분석 및 트래킹 시스템은 정의된 비율 편차(percentage deviation) 내의 제시된 관절 치수를 추가하고, 벗어난 값(outliers)이나 정의된 비율 편차 밖의 제시된 관절 치수는 버려진다. 타겟 인식, 분석 및 트래킹 시스템은 표준 편차와 제시된 관절 치수의 개수 간의 가장 높은 비율(highest ratio)을 갖는 예측치에 기반하여 말단들의 치수를 결정할 수 있다.
- [0137] 스캔을 통해 결정된 말단에 관련된 치수를 330에서 업데이트할 수 있다. 일 실시예에 따르면, 타겟 인식, 분석 및 트래킹 시스템은 스캔에 의해 결정된 치수가 정확한지를 결정하기 위해 하나 이상의 휴리스틱 기법 또는 규칙을 포함할 수 있다. 예를 들어, 타겟 인식 분석, 및 트래킹 시스템은 치수를 조정하는 데 사용되는, 대칭 관

절들 간의 유클리디안 거리(Euclidean distance)가 거의 동일한지를 결정하는 휴리스틱 기법 또는 규칙, 손 및/또는 팔꿈치가 신체 근처에 있는지를 결정하는 휴리스틱 기법 또는 규칙, 머리가 특정 포지션이나 위치에 있는지를 결정하는 휴리스틱 기법 및/또는 규칙, 손이 머리 근처에 있는지를 결정하는 휴리스틱 기법 및/또는 규칙 등을 포함할 수 있다. 전술한 바와 같이, 330에서 결정된 치수를 사용하여 325에서 다음 프레임에 대해 트래킹된 모델을 조정할 수 있다.

[0138] 335에서, 조정된 모델을 처리할 수 있다. 예를 들어, 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 조정된 모델에 적용된 하나 이상의 모션 또는 움직임을 아바타나 게임 캐릭터에 매핑시켜, 그 아바타나 게임 캐릭터가 애니메이션화되어 도 1a 및 1b에서 전술한 사용자(18) 등의 사용자를 흉내내게 함으로써 조정된 모델을 처리할 수 있다. 예를 들어, 온-스크린 캐릭터의 시각적인 모습은 조정된 모델의 변화에 따라 바뀔 수 있다.

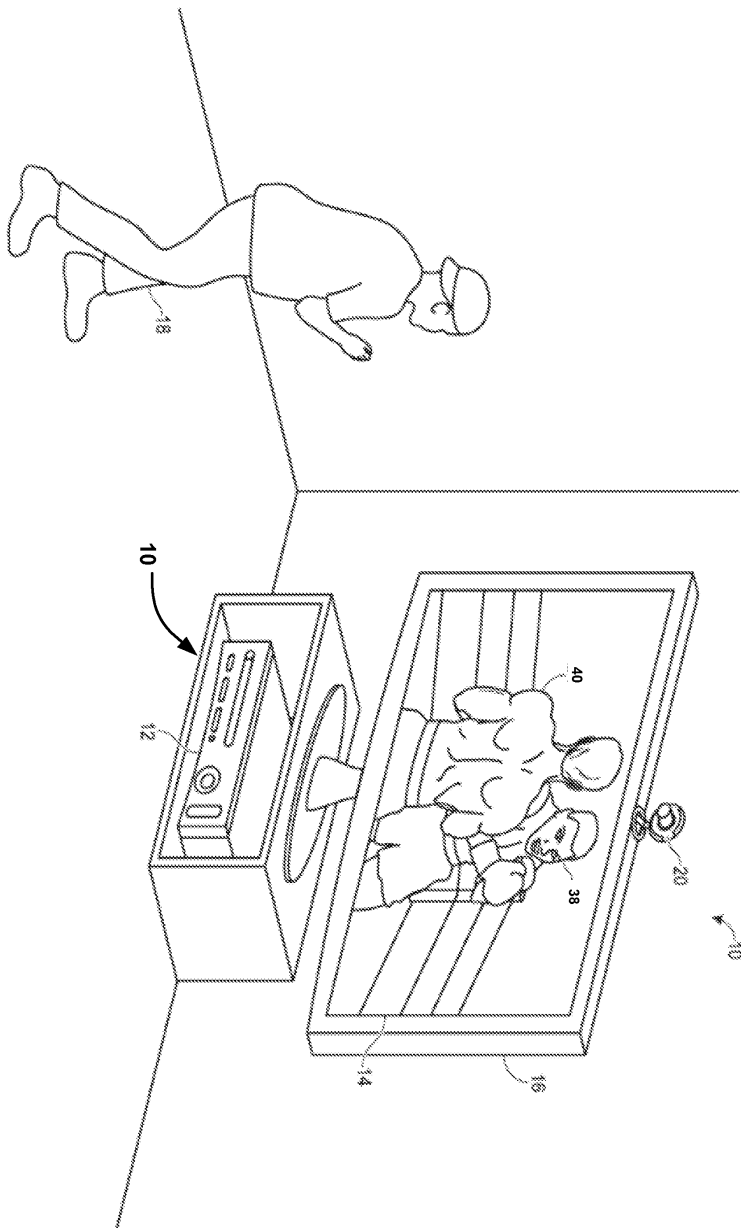
[0139] 일 실시예에서, 타겟 인식, 분석 및 트래킹 시스템은 도 1a - 4에서 전술한 컴퓨팅 환경(12) 등의 컴퓨팅 환경의 제스처 라이브러리에 조정된 모델을 제공함으로써 조정된 모델을 처리할 수 있다. 골격 모델의 각종 신체 부위의 포지션에 기반하여 애플리케이션에서 실행할 컨트롤을 결정하는 데 제스처 라이브러리를 사용할 수 있다.

[0140] 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 동작들은 도시된 순서대로, 다른 순서로 또는 동시에 실행될 수 있다. 마찬가지로, 전술한 프로세스의 순서가 변경될 수도 있다.

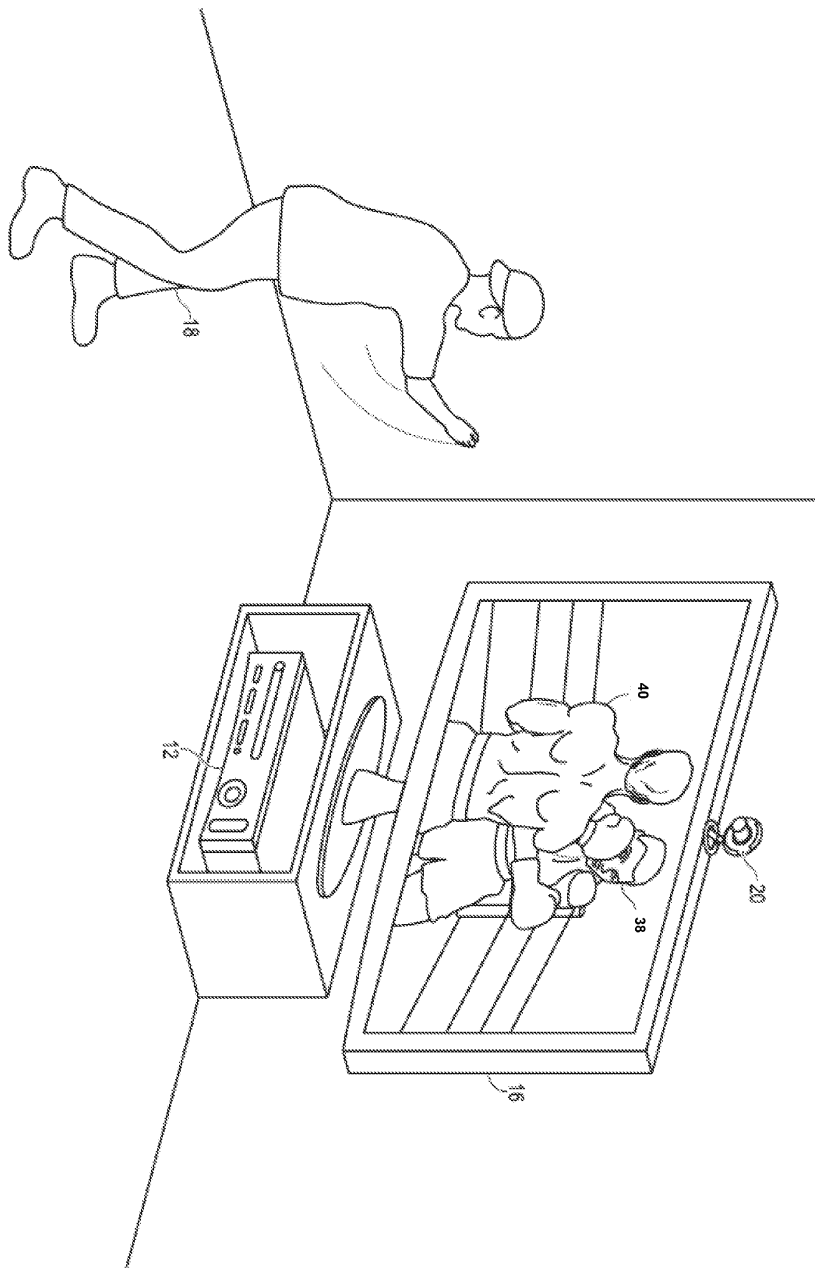
[0141] 본 개시의 대상은 다양한 프로세스, 시스템 및 구성, 및 기타 특징, 기능, 동작 및/또는 특성의 모든 신규하고 비자명한 조합 및 서브ком비네이션뿐만 아니라, 임의의 그리고 모든 등가물도 포함한다.

도면

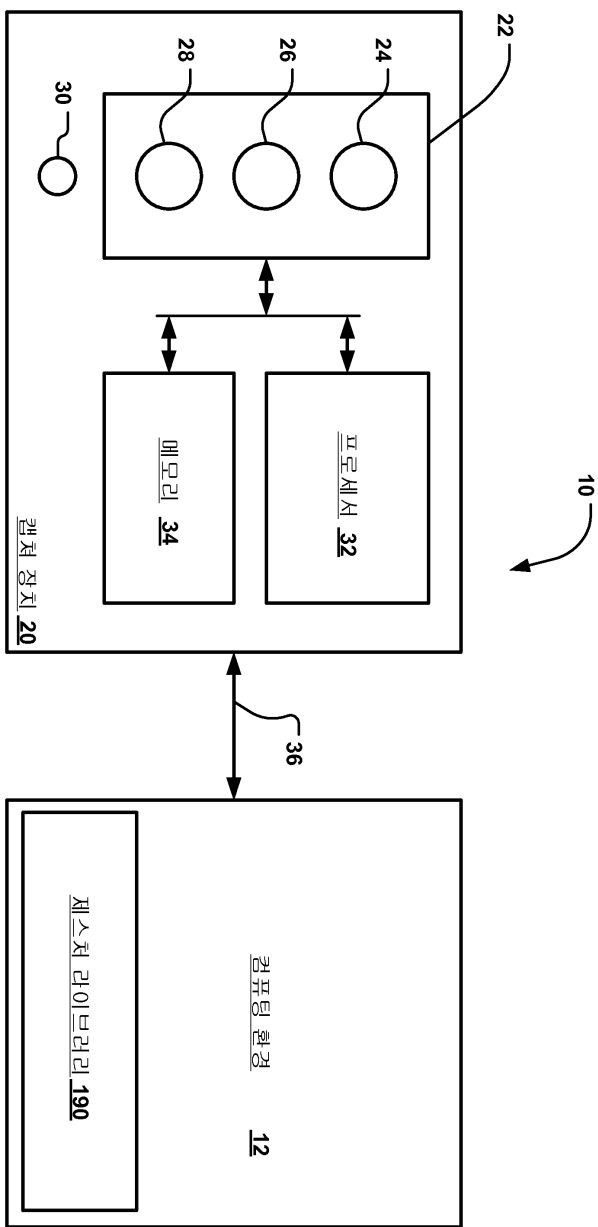
도면1a



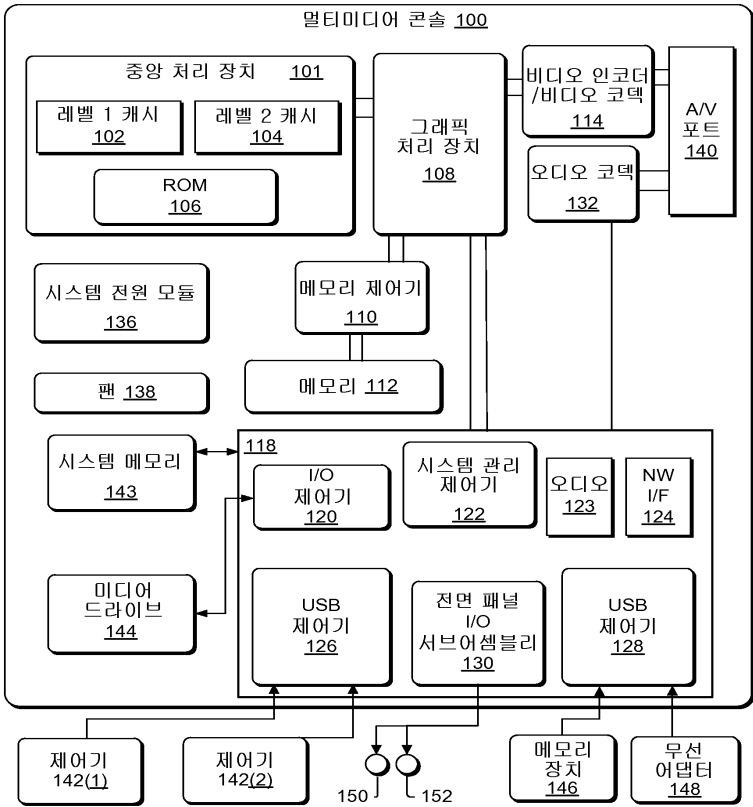
도면1b



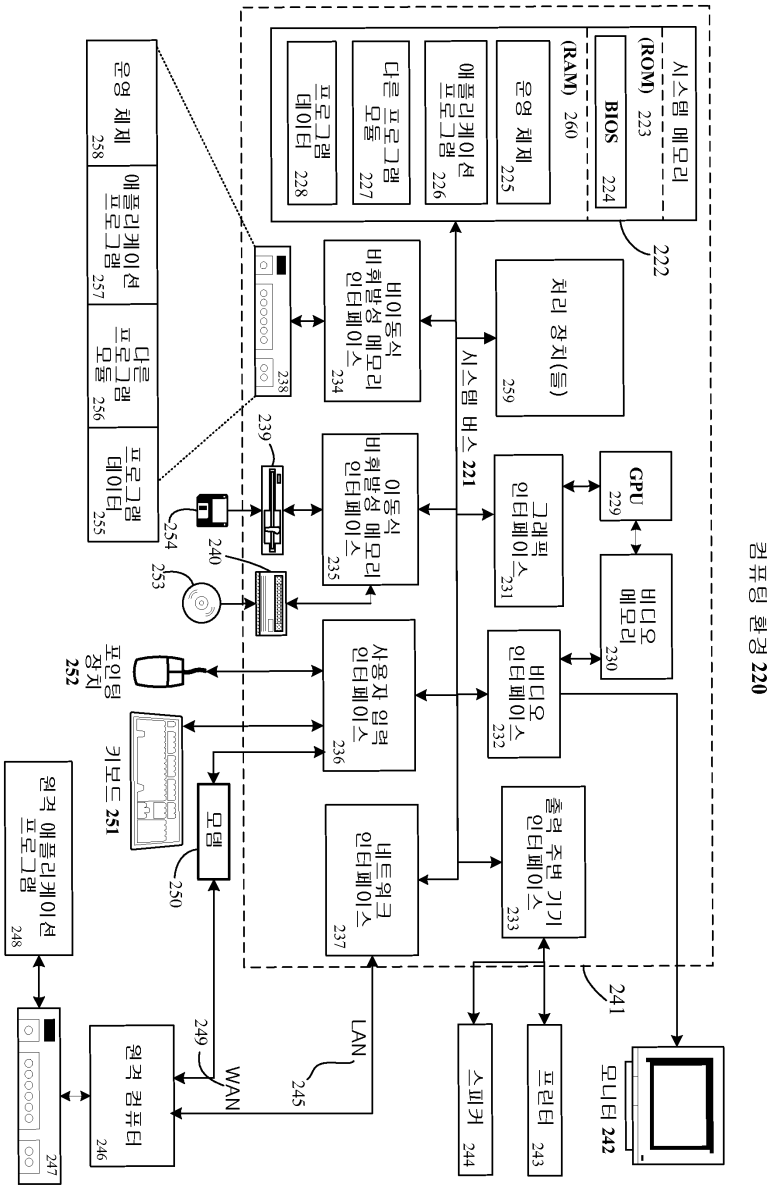
도면2



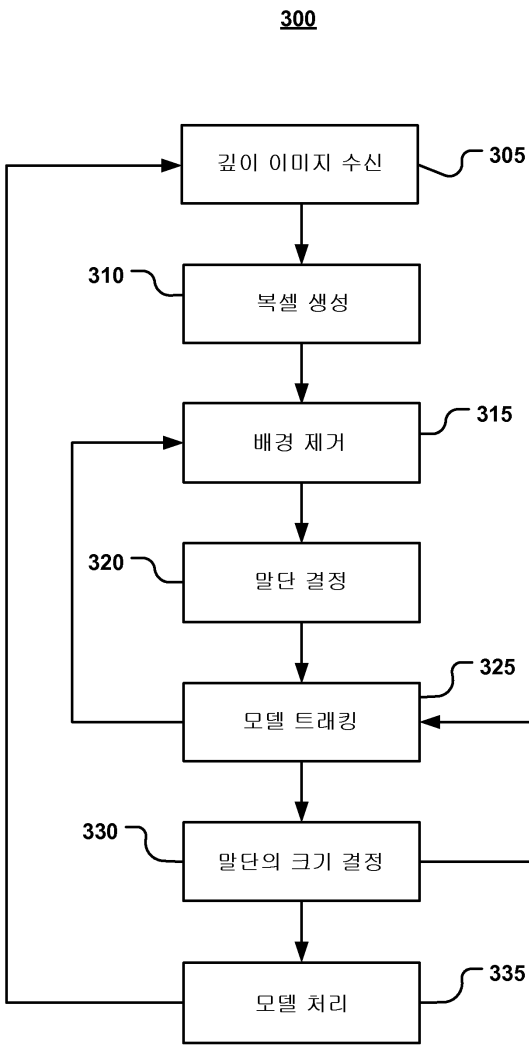
도면3



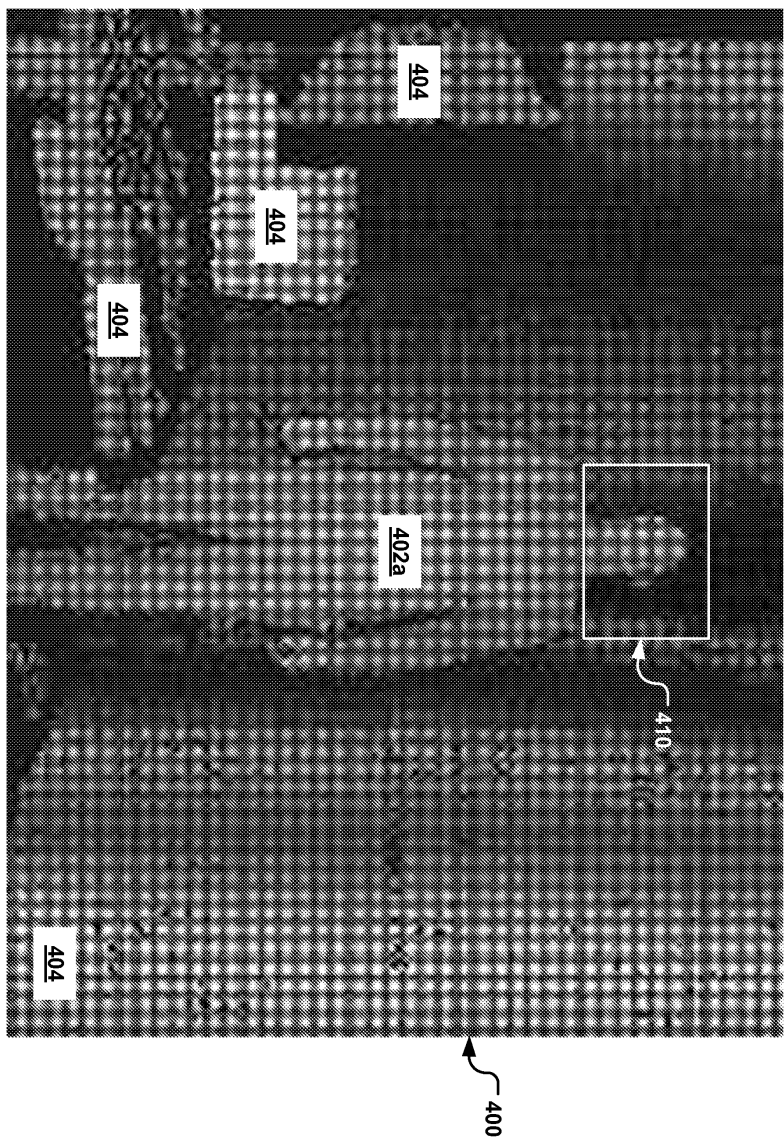
도면4



도면5



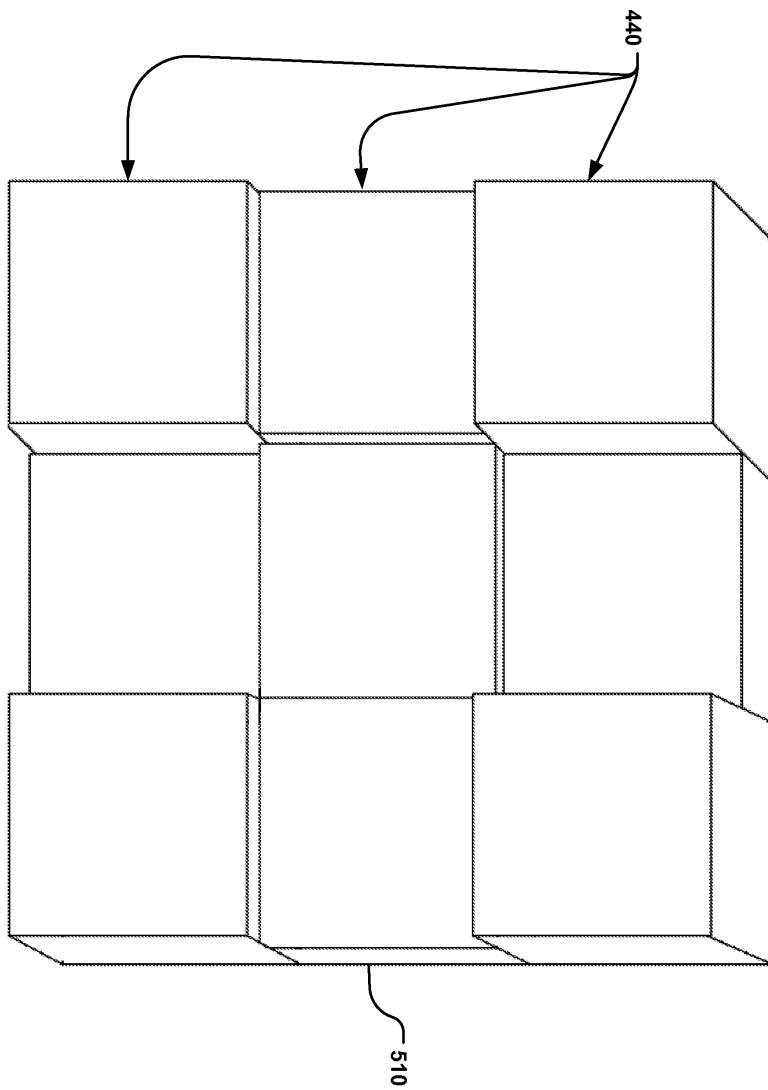
도면6



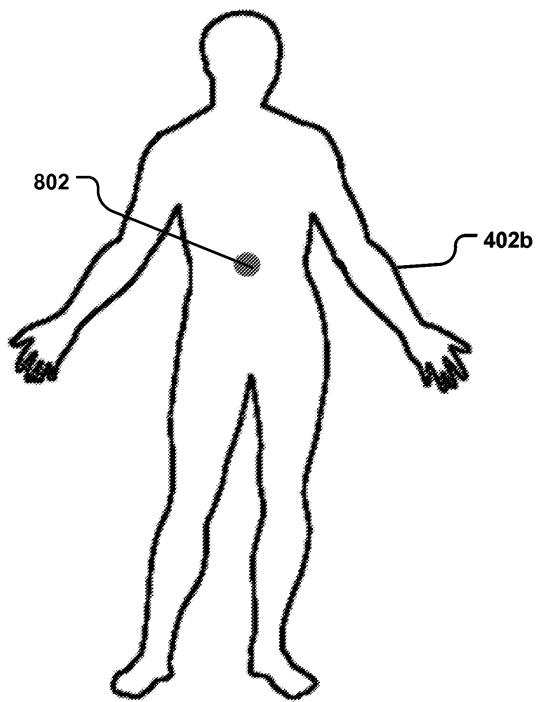
도면7a

[illegible]

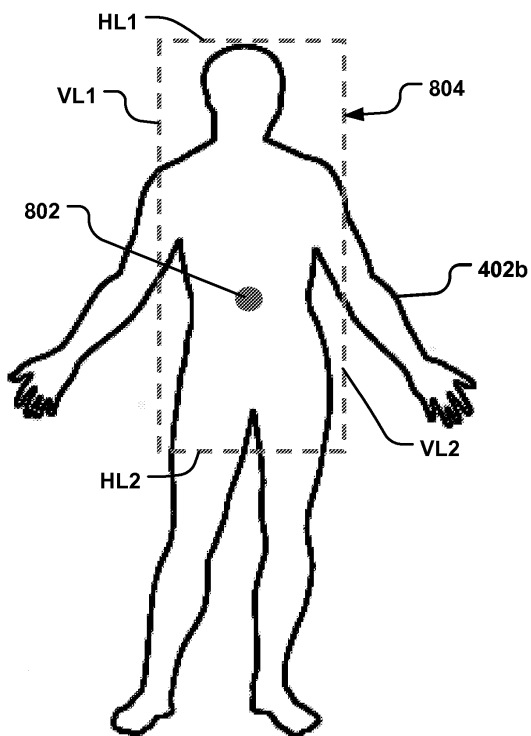
도면7b



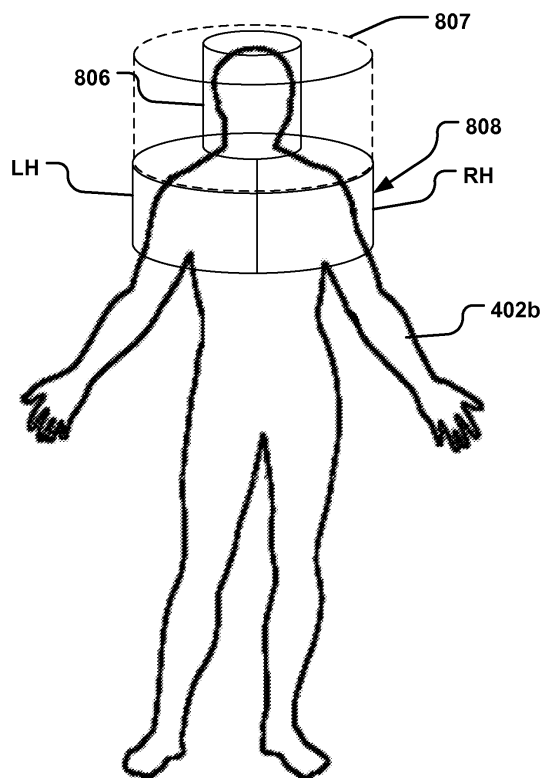
도면8



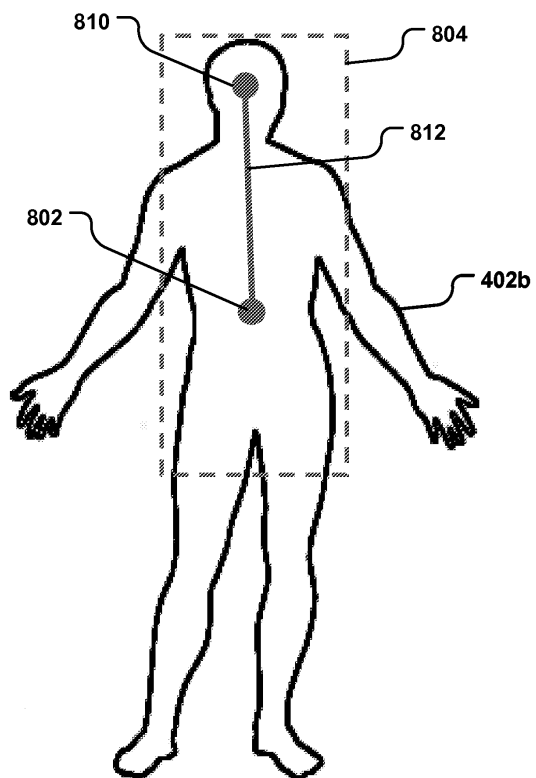
도면9



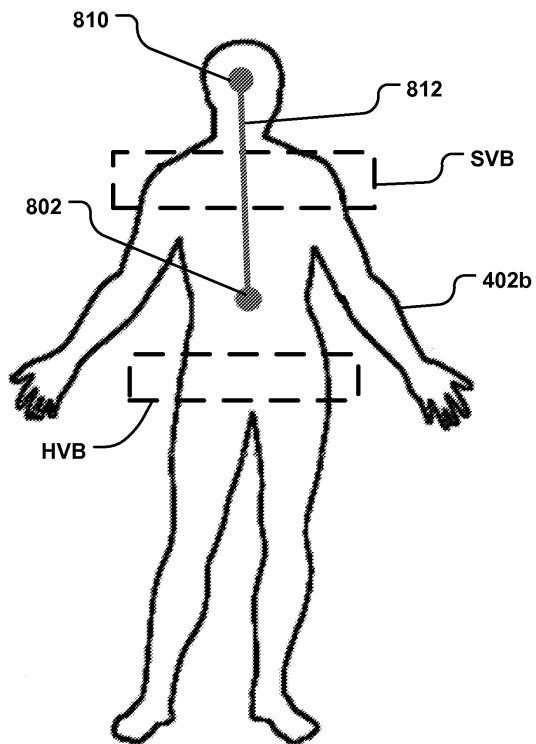
도면10



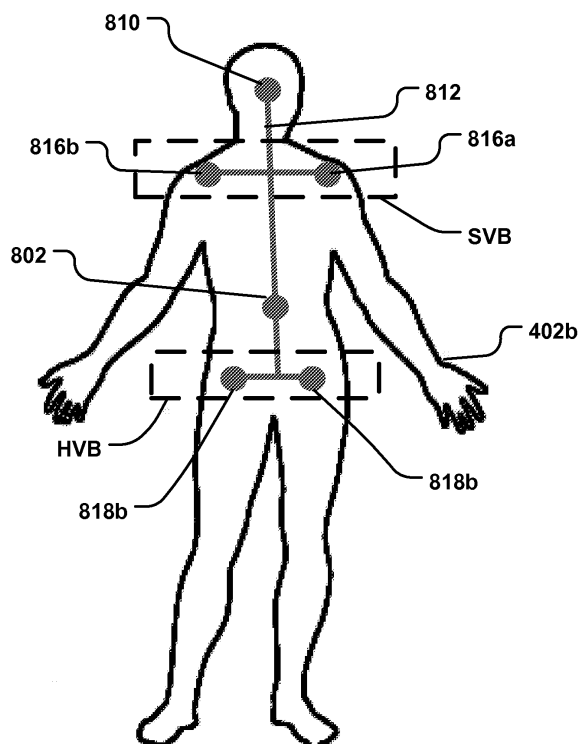
도면11



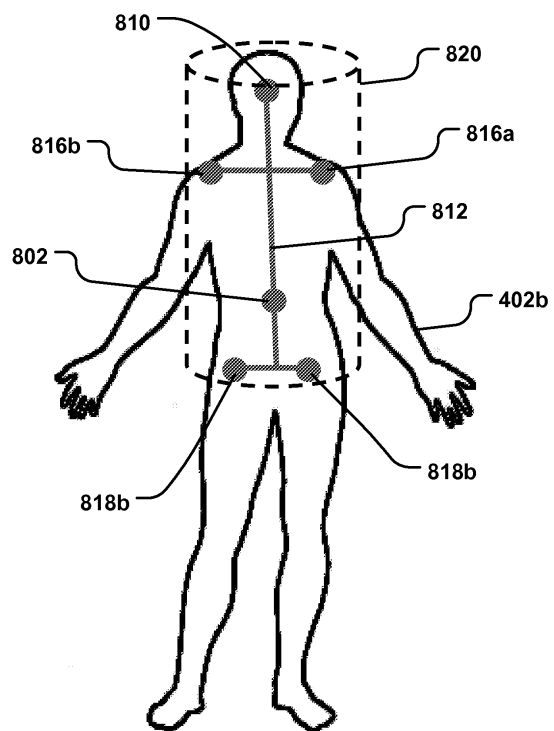
도면12



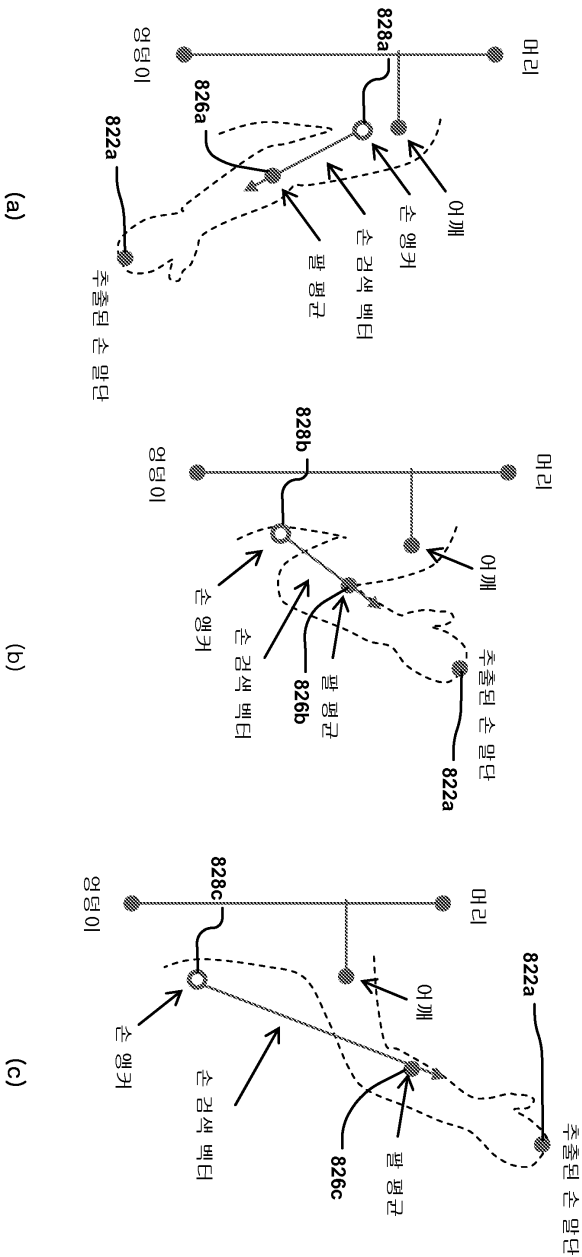
도면13



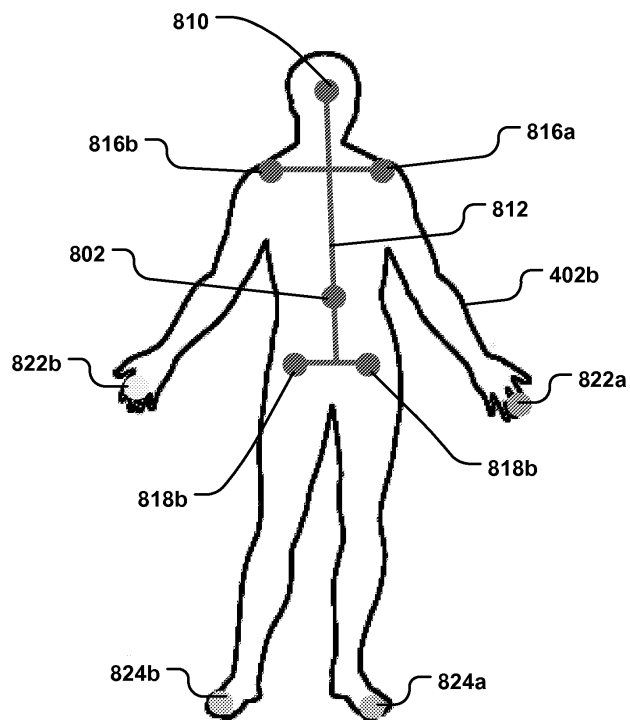
도면14



도면15



도면16



도면17

