



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0007878
(43) 공개일자 2012년01월25일

(51) Int. Cl.

G06T 7/20 (2006.01) G06T 17/00 (2006.01)

(21) 출원번호 10-2010-0068656

(22) 출원일자 2010년07월15일

심사청구일자 2010년07월15일

(71) 출원인

경희대학교 산학협력단

경기도 용인시 기흥구 덕영대로 1732, 국제캠퍼스 내 (서천동, 경희대학교)

(72) 발명자

이승룡

경기도 성남시 분당구 구미로144번길 25, 102동 102호 (구미동, 삼환빌라)

이영구

대전광역시 서구 만년로 25, 105동 904호 (만년동, 강변아파트)

(뒷면에 계속)

(74) 대리인

특허법인 신지

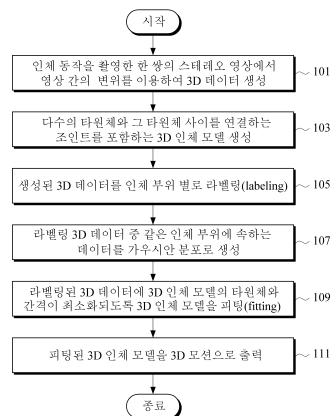
전체 청구항 수 : 총 8 항

(54) 인체 동작의 3D 모션 구현방법

(57) 요약

인체 동작의 영상으로부터 3D 모션을 캡처하는 기술에 관한 것으로, 본 발명의 인체 동작 영상의 3D 모션 구현방법은 인체 동작을 촬영한 한 쌍의 스테레오 영상에서, 그 영상 간의 변위를 이용하여 생성된 3D 데이터를 다수의 타원체와 그 타원체 사이를 연결하는 조인트를 포함하여 이루어지는 3D 인체 모델과 정합하기 위해, 생성된 3D 데이터를 각 신체 부위별로 라벨링(labeling)하는 단계와 라벨링된 3D 데이터와 상기 3D 인체 모델의 타원체와의 간격이 최소화되도록 3D 인체 모델을 피팅(fitting)하는 단계를 포함함으로써, 모션 자세에 대한 방대한 데이터 없이도 3D 모션 구현할 수 있다.

대표도 - 도1



(72) 발명자

김태성

경기도 수원시 팔달구 권광로 373, 106동 601호 (우만동, 월드메르디앙 아파트)

뉴엔덕탕

경기도 용인시 기흥구 덕영대로 1732, 전자정보대학 351호 (서천동, 경희대학교)

특허청구의 범위

청구항 1

인체 동작을 촬영한 한 쌍의 스테레오 영상에서, 그 영상 간의 변위를 이용하여 생성된 3D 데이터를 다수의 타원체와 그 타원체 사이를 연결하는 조인트를 포함하여 이루어지는 3D 인체 모델과 정합하여 3D 모션을 구현하는 방법에 있어서,

상기 생성된 3D 데이터를 각 인체 부위별로 라벨링(labeling)하는 단계; 및

상기 라벨링된 3D 데이터와 상기 3D 인체 모델의 타원체와의 간격이 최소화되도록 상기 3D 인체 모델을 피팅(fitting)하는 단계;

를 포함하는 인체 동작의 3D 모션 구현방법.

청구항 2

제 1항에 있어서, 상기 3D 모션 구현방법은 :

상기 라벨링된 3D 데이터 중 같은 인체 부위에 속하는 데이터를 가우시안 분포(Gaussian distribution)로 생성하는 단계;

를 더 포함하는 인체 동작의 3D 모션 구현방법.

청구항 3

제 1항에 있어서,

상기 3D 인체 모델의 조인트는 1 에서 3 사이의 회전 자유도(degree of freedom)를 가지는 인체 동작의 3D 모션 구현방법.

청구항 4

제 1항에 있어서, 상기 라벨링 단계는,

2D 영상을 이용하여 인체의 머리 부분과 몸통 부분을 구분하는 단계;를 더 포함하고, 상기 생성된 3D 데이터를 상기 구분된 인체의 머리 부분과 몸통 부분에 대응되도록 라벨링하는 인체 동작의 3D 모션 구현방법.

청구항 5

제 4항에 있어서,

상기 2D 영상은 RGB(red-green-blue) 영상, 이진 영상(binary image) 중 적어도 하나인 인체 동작의 3D 모션 구현방법.

청구항 6

제 4항 또는 제 5항에 있어서, 상기 구분하는 단계는 :

인체 동작을 촬영한 이진 영상(binary image)에서 인체 실루엣을 특정하는 단계; 및

인체 동작을 촬영한 RGB 영상에서 상기 특정된 인체 실루엣 영역의 피부색깔을 이용하여 머리 부분을 특정하는 단계;

를 포함하는 인체 동작의 3D 모션 구현방법.

청구항 7

제 6항에 있어서, 상기 구분하는 단계는 :

상기 특정된 머리 부분의 중심점과 상기 특정된 인체 실루엣의 중심점의 가운데를 중심점으로 하는 토르소(torso) 부분을 특정하는 단계;

를 더 포함하는 인체 동작의 3D 모션 구현방법.

청구항 8

제 1항에 있어서, 상기 피팅(fitting)하는 단계는 :

상기 라벨링된 3D 데이터의 각 점과 그에 대응되는 상기 인체 모델의 타원체의 표면과의 거리를 계산하는 단계;
및

상기 계산된 거리가 설정된 범위 내에 해당할 때까지 상기 인체 모델의 위치를 조정하는 단계;

를 포함하는 인체 동작의 3D 모션 구현방법.

명세서

기술 분야

- [0001] 인체 동작의 비디오 영상으로부터 3D 모션을 캡처하는 기술에 관한 것으로, 보다 구체적으로는 마커나 센서의 사용없이 인체 동작을 촬영한 2D 및 3D 스테레오 영상으로부터 추출한 3D 데이터를 이용하여 인체 모션을 3D로 캡처하는 기술이 개시된다.

배경 기술

- [0002] 오늘날 게임이나 영화와 같은 디지털 콘텐츠에는 사람의 몸 동작을 입체적으로 표현하는 3D(three dimension) 모션 캡처 기술이 널리 사용되고 있다. 게임의 경우에는 사용자가 움직임 감지센서가 부착된 리모콘을 조작하여 화면상에 나타나는 캐릭터의 움직임을 제어할 수 있다. 또한, 영화의 경우에는 인체 부착 센서 정보를 이용하여 캐릭터가 실제 사람이 움직이는 것과 같이 입체적인 효과를 만들어 내고 있다.
- [0003] 이러한 사람의 동작을 3D 모션으로 구현하기 위해서는 실제 사람의 몸에 마커(marker)를 부착하거나 센서가 부착된 옷을 입고 연기함으로써 움직임 데이터를 컴퓨터에 전달하게 된다. 그러나 이러한 방법은 3D 모션을 획득하기 위해 실제 사람들이 마커나 센서가 부착된 옷을 입어야하는 불편이 있고, 다중 카메라를 사용해야 하며, 제한된 공간에서만 촬영을 해야한다는 공간의 제약이 있었다.
- [0004] 위와 같은 문제점을 해결하기 위해, 많은 연구가 진행되고 있다. 일 예로 스테레오 카메라를 이용하는 인체 모션 캡처링(human motion capturing) 시스템이 개발되었다. 이는 촬영된 비디오 영상으로부터 구현된 3D 정보를 데이터 베이스에 저장된 3D 비교함으로써 인체 모션을 획득하는 방법이다. 그러나, 이러한 방법은 3D 인체 모델에 대한 방대한 데이터를 관리하고, 그로부터 검색하는데 불편이 있었다.
- [0005] 따라서, 본 발명자는 단일 스테레오 카메라를 이용하여 취득한 2D 및 3D 스테레오 영상으로부터 획득한 3D 데이터와 3D 인체 모델을 피팅(fitting)하여 3D 인체 모션을 획득하는 기술을 개발하게 되었다.

발명의 내용

해결하려는 과제

- [0006] 인체 동작의 영상으로부터 다중 카메라 및 인체 부착 센서 없이 3D 인체 모델 기반으로 3D 모션을 트래킹(tracking)할 수 있다. 또한, 3D 데이터를 가우시안 분포로 생성하여 라벨링(labeling)할 수 있다. 또한, 2D 및 3D 스테레오 영상으로부터 획득한 3D 데이터와 3D 인체 모델을 피팅(fitting)할 수 있다.

과제의 해결 수단

- [0007] 본 발명의 일 실시예에 따른 인체 동작의 3D 모션 구현방법은 인체 동작을 촬영한 한 쌍의 스테레오 영상에서, 그 영상 간의 변위를 이용하여 생성된 3D 데이터를 다수의 타원체와 그 타원체 사이를 연결하는 조인트를 포함하여 이루어지는 3D 인체 모델과 정합하기 위해, 생성된 3D 데이터를 각 신체 부위별로 라벨링(labeling)하는 단계와 라벨링된 3D 데이터와 상기 3D 인체 모델의 타원체와의 간격이 최소화되도록 3D 인체 모델을 피팅(fitting)하는 단계를 포함한다.
- [0008] 본 발명의 일 실시예에 따른 인체 동작의 3D 모션 구현방법은 라벨링된 3D 데이터 중 같은 신체 부위에 속하는 데이터를 가우시안 분포(Gaussian distribution)로 생성하는 단계를 더 포함하고, 라벨링 단계는 2D 영상을 이

용하여 인체의 머리 부분과 몸통 부분을 구분하는 단계와 구분된 신체의 머리 부분과 몸통 부분을 대응되는 생성된 3D 데이터를 라벨링한다.

[0009] 본 발명의 일 실시예에 따른 인체 동작의 3D 모션 구현방법 중 라벨링(labeling)하는 단계는 2D 영상을 이용하여 인체의 머리 부분과 몸통 부분을 구분하는 단계를 더 포함하고, 구분된 신체의 머리 부분과 몸통 부분을 대응되는 상기 생성된 3D 데이터를 라벨링하며, 인체 동작을 촬영한 이진 영상(binary image)에서 신체 실루엣을 특정하는 단계를 포함한다.

[0010] 본 발명의 일 실시예에 따른 인체 동작의 3D 모션 구현방법 중 라벨링(labeling)하는 단계에서 인체 동작을 촬영한 RGB 영상에서 특정된 신체 실루엣 영역의 피부색깔을 이용하여 머리 부분을 특정하고, 특정된 머리 부분의 중심점과 특정된 신체 실루엣의 중심점의 가운데를 중심점으로 하는 토르소(torso) 부분을 특정하여 신체 일부를 구분한다.

발명의 효과

[0011] 모델 기반의 3D 모션을 구현함에 있어서, 다중 카메라, 인체 부착 센서 및 인체 동작 자세에 대한 방대한 데이터 없이도 3D 모션을 구현할 수 있다. 또한, 3D 데이터에서 각 신체 부위별 위치를 인식하고, 서로 다른 신체 부분을 기 설정된 인체 모델과 연결하는 작업이 동시에 이뤄질 수 있다. 또한, 2D 영상을 이용함으로써 보다 정확하고 간편하게 3D 데이터에서 신체 부위를 라벨링할 수 있다.

도면의 간단한 설명

[0012] 도 1은 일 실시예에 따른 인체 동작의 3D 모션 구현방법의 흐름도,
 도 2는 일 실시예에 따른 스테레오 영상으로부터 3D 데이터를 추출하는 예시도,
 도 3은 일 실시예에 따른 인체 모델의 구성도,
 도 4는 일 실시예에 따른 이진 영상을 획득하는 예시도,
 도 5는 일 실시예에 따른 실루엣을 이용한 인체의 토르소 영역 특정 예시도,
 도 6은 일 실시예에 따른 3D 데이터의 지오데식 거리를 설정하는 예시도,
 도 7은 일 실시예에 따른 최종 구현된 3D 인체 모션의 예시도이다.

발명을 실시하기 위한 구체적인 내용

[0013] 이하, 첨부된 도면들을 참조하여 본 발명의 실시예를 상세하게 설명한다. 사용되는 용어들은 실시예에서의 기능을 고려하여 선택된 용어들로서, 그 용어의 의미는 사용자, 운용자의 의도 또는 판례 등에 따라 달라질 수 있다. 그러므로 후술하는 실시예들에서 사용된 용어의 의미는, 본 명세서에 구체적으로 정의된 경우에는 그 정의에 따르며, 구체적인 정의가 없는 경우는 당업자들이 일반적으로 인식하는 의미로 해석되어야 할 것이다.

[0014] 도 1은 일 실시예에 따른 인체 동작의 3D 모션 구현방법의 흐름도이다. 도 1을 참조하면, 먼저 스테레오 카메라(stereo camera)를 이용하여 인체 동작을 촬영하고, 촬영된 영상으로부터 3D 데이터를 생성한다(101). 스테레오 카메라는 동시에 2개의 화상을 얻을 수 있도록 고안된 특수카메라이다. 2개의 촬영용 렌즈를 일정 간격을 띄어 놓고 같은 물체를 촬영하기 위해 사용된다. 이렇게 얻은 스테레오 영상을 스테레오 뷰어를 이용하여 보면 입체적인 이미지를 얻을 수 있다.

[0015] 도 2는 일 실시예에 따른 스테레오 영상으로부터 3D 데이터를 추출하는 예시도이다. 도 2를 참조하면, 스테레오 카메라로부터 획득한 한 쌍의 스테레오 영상 간에 변위를 나타내는 불일치 영상(disparity image)(201)을 획득할 수 있다. 이러한 불일치는 각도를 달리하여 촬영된 영상의 변위가 다름에 따라 픽셀(pixel)별로 깊이 값(depth-value)이 다르게 계산된다. 깊이 값을 계산하는 방법은 빠른 스테레오 매칭 알고리즘인 GCS(Growing Correspondent Seeds)를 이용할 수 있다.

수학식 1

$$Z = \frac{fb}{d}$$

[0016]

[0017] 수학식 1은 GCS 알고리즘을 이용하여 z축에 대한 깊이 값을 구하는 공식이다. Z는 z축의 깊이 값이고, f는 포커스 길이, b는 베이스 라인, d는 불일치 값에 해당한다. 따라서, z축 방향의 깊이 값은 스테레오 카메라의 포커스 길이와 베이스 라인의 곱을 불일치 값으로 나눈 값이 된다.

수학식 2

$$X = \frac{uZ}{f}, \quad Y = \frac{vZ}{f}$$

[0018]

[0019] 수학식 2는 GCS 알고리즘을 이용하여 x축, y축에 대한 깊이 값을 구하는 공식이다. 이는 수학식 1에서 구한 Z축의 깊이 값인 Z를 이용하고, u, v는 불일치 영상의 픽셀의 행렬 값을 나타낸다. 이처럼, 불일치 영상의 깊이 값을 구하는데 있어서, GCS 알고리즘을 이용하면 계산식을 단순화할 수 있어 계산 속도와 정확성을 향상시킬 수 있다.

[0020] 도 2에서, 불일치 영상(201)에 대한 픽셀별 계산을 줄이기 위해 그리드(grid) 단위로 샘플링할 수 있다. 샘플링된 불일치 영상(203)은 계산이 필요한 픽셀의 숫자를 줄임으로써 광범위한 계산을 피할 수 있다. 샘플링된 불일치 영상(203)의 각 픽셀(pixel)에 계산된 깊이 값을 적용하면 3D 영상(205)을 얻을 수 있다. 획득된 3D 영상의 3D 데이터는 x축, y축, z축에 대한 공간 정보를 포함하게 된다.

[0021] 다시, 도 1을 참조하면, 다음단계로 3D 인체 모델을 생성한다(103). 3D 인체 모델은 모델 기반의 3D 모션 구현 방법에서 사용된다. 3D 인체 모델이란, 인체 부위를 형상화한 가상의 모델을 말한다. 3D 인체 모델은 다수의 타원체와 그 타원체 사이를 연결하는 조인트를 포함하여 구성될 수 있다. 이와 관련하여, 도 3을 참조하여 설명하도록 한다.

[0022] 도 3은 일 실시예에 따른 인체 모델의 구성도이다. 도 3을 참조하면, 3D 인체 모델을 생성하기 위해, 먼저 뼈대 구조 모델(301)을 생성할 수 있다. 뼈대 구조 모델(301)은 예를 들어 14개의 세그먼트(303)와 그 세그먼트를 연결하는 9개의 조인트(305)로 구성될 수 있다. 조인트(305)는 회전 자유도(degree of freedom)를 최대 3까지 가질 수 있다. 뼈대 구조 모델(301)의 경우 각 조인트에서 자유도 3을 가지고, 엉덩이 부분에서 자유도 6을 가지므로, 전체에서 33의 자유도를 가질 수 있다.

[0023] 이러한 뼈대 구조 모델(301)의 세그먼트(303)를 타원체로 대체하여 타원체 모델(307)을 생성할 수 있다. 뼈대 구조 모델(301)은 다수의 타원체(309)로 구성되며, 각각의 타원체는 다음과 같은 수학식을 이용하여 생성될 수 있다.

수학식 3

$$(X_i - m_i)^T R_i^T(\theta) \Lambda_{v_i} R_i(\theta) (X_i - m_i) = 1$$

[0024]

[0025]

수학식 3은 인체모델의 타원체 모델(307)을 계산하는 공식이다. 수학식 3에서, m_i 는 타원체의 중심이며, $R_i(\theta)$ 는 타원체의 회전형렬이며, 상수행렬인 Λ_{v_i} 는 타원체의 사이즈를 결정한다. 각 타원체의 회전은 각 조인트에서의 변환의 연속에 의해 결정되기 때문에 $R_i(\theta)$ 는 단일 파라미터에 의존하는 독립적인 회전 행렬들의 시리즈로 분리될 수 있다.

수학식 4

$$R_i(\theta) = R_x(\theta_1) R_y(\theta_2) R_z(\theta_3) R_x(\theta_4) \dots R_y(\theta_{k-1}) R_z(\theta_k)$$

[0026]

[0027]

수학식 4에서, $R_x(\theta_n)$, $R_y(\theta_n)$, $R_z(\theta_n)$ 는 각각 x, y, z 축에서 θ_n 각도로 기울어진 회전 매트릭스를 나타낸다.

[0028]

다시 도 1을 참조하면, 다음으로 생성된 3D 데이터를 인체 부위별로 라벨링(labeling)한다(105). 라벨링(labeling)이란 인접한 화소에 같은 라벨(label)을 붙이고, 연결되지 않거나 다른 성분의 화소에는 다른 라벨(label)을 부여하는 것을 말한다. 라벨링된 후에는 서로 연결된 인접화소 영역은 픽셀(pixel)들의 값이 동일한 하나의 번호로 라벨링되어 나타나며, 다른 인접영역의 성분은 다른 번호로 나타난다. 이에 따라, 라벨링된 영상을 이용하면 대상을 쉽게 인식할 수 있다.

[0029]

일 실시예에 따른 라벨링 단계는 2D 영상을 이용하여 인체의 머리 부분과 몸통 부분을 구분할 수 있다. 예를 들어, 2D 영상은 RGB(red-green-blue) 영상, 이진 영상(binary image)일 수 있다. 이러한 2D 영상은 3D 데이터에서 인체 부위를 확인하는데 추가적인 정보를 제공할 수 있다. 특히, 얼굴 부위는 RGB 영상에서 피부색깔을 확인하여 인식할 수 있으며, 토르소(torso) 부위는 이진 영상에서 인체의 실루엣을 통해 인식할 수 있다.

수학식 5

$$p_{i_{step}+1}(v_i) = \frac{1}{Z_{i_{step}}(v_i)} \exp \left\{ \sum_{i=1}^M \sum_{j \in N_i} p_{i_{step}}(v_j) f_{ij}(v_i, v_j) + \sum_{i=1}^M \sum_{j=1}^{M_c} p_{i_{step}}(v_j) g_{ij}(v_i, v_j) + f_i(v_i) + g_i(v_i) \right\}$$

[0030]

[0031]

수학식 5는 픽셀 i가 라벨 v_i 를 가질 확률($P_{i_{step}}$)을 계산하는 공식이다. 여기서 $Z_{i_{step}}(v_i)$ 는 정규화하기 위한 팩터이고, N_i 는 포인트 X_i 의 인접한 점들의 집합이며, M_c 는 샘플링된 셀들의 숫자를 나타낸다.

$[1/Z_{i0}(v_i)]\exp\{f_i(v_i)+g_i(v_i)\}$ 를 $P_{i0}(v_i)$ 의 초기값으로 설정할 수 있다. 수학식 5를 반복적으로 계산함에 따라 픽셀 i 가 라벨 v_i 를 가질 확률이 업데이트될 수 있다.

[0032] 3D 데이터에서 인체 중 얼굴 부위는 RGB 영상으로부터 특정될 수 있다. 피부색깔은 다른 신체부위와 색깔이 다르므로 얼굴로 지정된 피부색깔을 가지는 부위는 얼굴 부위로 특정된다. 얼굴은 머리의 일부로 볼 수 있기 때문에 얼굴에 해당하는 부위는 머리로 간주될 수 있다. 인체 부위별 가능성 $f_i(v_i)$ 는 피부색깔인 $c>0$ 이면 v_i 를 머리로 라벨링될 수 있다.

[0033] 또한, 얼굴 부위를 구분하여 라벨링하기 위해 HSV 색 공간이나 열 적외선 영역을 이용할 수 있다. 사람의 인체는 고유의 열 적외선을 방출한다. 그러나 인체의 대부분은 옷이나 신발로 인해 열이 감지되는 영역이 다를 수 있다. 다만, 얼굴에는 마스크를 착용하는 경우가 아닌 이상 인체의 열 적외선이 그대로 방출되므로 인체의 다른 부위와 쉽게 구분할 수 있다.

[0034] 한편, 인체의 부위 중 토르소(torso) 부위를 라벨링하기 위해서는 이진 영상(binary image)을 이용할 수 있다. 이와 관련해서는 도 4와 도 5를 통해 설명하기로 한다.

[0035] 도 4는 일 실시예에 따른 이진 영상을 획득하는 예시도이다. 도 4를 참조하면, 사람의 몸 동작이 포함된 인체 영상(401)을 획득하고, 이를 이진 영상(403)으로 변환할 수 있다. 이진 영상(binary image)이란, 이미지를 2가지 색으로 단순화하여 나타낸 영상을 말한다. 예를 들어, 물체는 검은색으로 나타내고 배경은 흰색으로 나타낼 수도 있다. 이에 따라, 이진 영상을 이용하여 대상을 쉽게 인식하므로 라벨링 작업이 더 효율적일 수 있다.

[0036] 변환된 이진 영상(403)에는 배경에 속하는 부분이 인체와 동일한 색깔로 표시되는 경우가 발생할 수 있다. 정확한 인체 실루엣을 획득하기 위해 인체 실루엣의 외곽영역에 있는 부분은 삭제영역(407)으로 분리될 수 있다. 또한, 인체의 일부분이라고 볼 수 없는 부위도 삭제영역(407)으로 분리된다. 이러한 삭제과정을 거치면 수정된 이진 영상(405)을 획득할 수 있다.

[0037] 도 5는 일 실시예에 따른 실루엣을 이용한 인체의 토르소 영역을 특정하는 예시도이다. 도 5를 참조하면, 인체의 실루엣(501)을 통해 토르소 영역을 특정할 수 있다. 머리영역은 앞서 설명한 RGB 영상 등을 통해 특정할 수 있으며, 특정된 머리 영역(503)의 중심점(507)과 전체 실루엣의 중심점(511) 사이의 가운데 지점을 중심점(509)으로 설정하여 토르소 영역(505)을 인식할 수 있다. 이에 따라, 인체 부위 중 머리 부위와 토르소 부위를 인식할 수 있다.

[0038] 토르소 부위 해당 가능성 $f_i(v_i)$ 는 점 X_i 로부터 토르소 부위의 중심점까지의 대수적인 거리를 나타내는 $d(x_i)$ 를 이용하여 특정할 수 있다. 만약 $d(x_i) \leq 1$ 일 때, 가능성 $f_i(v_i)$ 가 $1 - d(x_i)$ 이면 v_i 를 토르소로 라벨링될 수 있다. 값이 0인 경우에는 다른 부위로 인식하게 된다. 이에 따라, 3D 데이터의 라벨링시 보다 정확하게 인체 부위별 라벨링이 이뤄질 수 있다.

[0039] 수학식 5에서 $g_i(v_i)$ 는 복원(reconstruction) 에러와 관련된다. 각 포인트에서의 복원 에러는 다음과 같은 수학식 6으로 표현될 수 있다.

수학식 6

$$g_i(v_i) = -\frac{d^2(X_i, \theta, v_i)}{\sigma^2}$$

[0040]

[0041] 수학식 6은 복원 에러를 계산하는 공식이다. 수학식 6에서, σ^2 는 분산을 나타내고, θ 는 인체 모델의 운동 파라미터를 나타낸다. $d(X_i, \theta, v_i)$ 는 하나의 3D 포인트 X_i 와 v_i 라벨을 갖는 타원체 사이의 거리를 나타낸다. $d(X_i, \theta, v_i)$ 는 다음의 수학식 7로 나타낼 수 있다.

수학식 7

$$d(X_i, \theta, v_i) = \sqrt{d_1 \left(1 - \frac{1}{d_2} \right)}$$

[0042]

[0043] 수학식 7에서 $d_1 = (X_i - m_i)^T * R_i^T(\theta) * R_i(\theta) * (X_i - m_i)$ 이고, $d_2 = (X_i - m_i)^T * R_i^T(\theta) * \Lambda_{v_i} * R_i(\theta) * (X_i - m_i)$ 이다. 이러한 복원에러를 계산하여 보정함으로써 라벨링이 보다 정확하게 이뤄질 수 있다.

[0044] 수학식 5로 돌아가서, $f_{ij}(v_i, v_j)$ 는 평탄(smoothness) 에너지를 나타내며, v_i 는 픽셀 i 의 라벨을 v_j 는 픽셀 j 의 라벨을 나타낸다. 여기서, 평탄 에너지 $f_{ij}(v_i, v_j)$ 는 $v_i = v_j$ 이면 γ ($\gamma > 0$)이고, $v_i \neq v_j$ 이면 1의 값을 갖게 된다.

[0045] 도 6은 일 실시예에 따른 3D 데이터의 지오데식 거리를 설정하는 예시도이다. 수학식 5에서 지오데식(geodesic) 제한인 $g_{ij}(v_i, v_j)$ 는 지오데식 거리에 기초하여 잘못 라벨이 매겨진 것을 제한하기 위한 것이다. 도 6에서 두 점들 간의 지오데식 거리는 디스트라 알고리즘(Dijkstra's algorithm)을 사용하여 그래프상 가장 짧은 경로를 추정한다. 계산을 간소화하기 위해, 근접한 픽셀들을 셀(cell)이라 불려지는 그룹으로 할당할 수 있다. 같은 셀에 속하는 모든 점들은 동일한 지오데식 제한을 받게된다.

[0046] 이에 따르면, 지오데식 제한은 오직 셀들 간의 쌍대-비교(pair-wise) 관계만 문제된다. 지오데식 제한은 다음과 같이 정의된다.

수학식 8

$$g_{ij_c}(v_i, v_{j_c}) = \begin{cases} -\alpha & d(v_{i_c}, v_{j_c}) < d_{\min}(v_{i_c}, v_{j_c}) \\ -\beta & d(v_{i_c}, v_{j_c}) > d_{\max}(v_{i_c}, v_{j_c}) \end{cases}$$

[0047]

[0048] 수학식 8에서, i_c 는 점 i 를 포함하는 셀이며, $d(v_{i_c}, v_{j_c})$ 는 셀 i_c 와 셀 j_c 간의 지오데식 거리이며, α , β 는 양의 상수이다. $d_{\min}(v_{i_c}, v_{j_c})$ 과 $d_{\max}(v_{i_c}, v_{j_c})$ 는 한 쌍의 라벨 간의 지오데식 거리를 위한 하위, 상위 경계로 정의된다.

[0049] 다시, 도 1로 돌아가면, 다음으로 라벨링된 3D 데이터 중 같은 인체 부위에 속하는 데이터를 가우시안 분포로 생성할 수 있다(107). 이는 라벨링된 3D 데이터를 3D 인체 모델에 피팅하기 위한 계산과정을 간소화하기 위한 것이다. 같은 라벨의 점들을 가우시안 분포로 그룹화할 수 있다. 같은 라벨 i 를 부여받은 포인트 집합인 X_i 에

서 raw 가우시안 분포는 $P(x_i | D) \propto N(m_{i0}, \Sigma_{i0}^{-1})$ 이다. 평균 m_{i0} 는 $\frac{1}{N_i} \sum_i x_i$ 이고, 공분산 Σ_{i0} 는

$\frac{1}{N_i - 1} \sum_i (x_i - m_{i0})(x_i - m_{i0})^T$ 이다. 여기서, N_i 는 인체 부위 i 에 속하는 픽셀들의 숫자이다.

[0050]

수정된 추정 가우시안 분포는 $\tilde{P}(x_i) \propto \mathcal{N}(m_i, \Sigma_i^{-1})$ 이다. 이 때, 공분산은

$\Sigma_i^{-1} = \tilde{R}_i^T \Lambda_{i0} \tilde{R}_i$ 이다. 상수행렬 $\Lambda_{i0} = \mu \cdot \text{diag}\{a^{-2}, b^{-2}, c^{-2}\}$ 는 타원체의 사이즈에 비례한다.

평균 m_i 는 기울기 하강법(gradient descent)에 의해 추정된다.

수학식 9

$$\begin{aligned} m_i &\leftarrow m_i - \gamma \left((m_i - m_{i0})^T + \sum_{j \in S_i} \lambda_{ij} (m_i - m_j + \tilde{R}_i^T a_{ij} - \tilde{R}_j^T b_{ji})^T \right) \\ \tilde{R}_i &\leftarrow \tilde{R}_i + \gamma \left(\tilde{R}_i - \tilde{R}_i \Sigma_i \Sigma_i^{-1} - \sum_{j \in S_i} \lambda_{ij} a_{ij} (m_i - m_j + \tilde{R}_i^T a_{ij} - \tilde{R}_j^T b_{ji})^T \right) \\ \lambda_{ij} &\leftarrow \lambda_{ij} + \eta \|m_i - m_j + \tilde{R}_i^T a_{ij} - \tilde{R}_j^T b_{ji}\|^2 \end{aligned}$$

[0051]

[0052] 수학식 9에서, S_i 는 인체 부위 i 에 연결된 다른 인체 부위들 j 의 집합이다. γ , η 는 학습률(learning rate)이다. λ_{ij} 는 라그랑제(Lagrange) 곱셈 연산자이다. a_{ij} , b_{ji} 는 벡터이다. 이와 같이 라벨링 단계에서 가우시안 분포를 이용함으로써 각 인체 부위별로 정확하고 신속하게 작업을 할 수 있다.

[0053] 다음으로 라벨링된 3D 데이터에 3D 인체 모델의 타원체와의 간격이 최소화되도록 3D 인체 모델을 피팅(fitting)

한다(109). 앞서 설명한 각각의 가우시안 분포의 파라미터인 $(\tilde{R}_i, m_i)$ 이 추정된 후에는 확장된 칼만 필터(Kalman Filter)에 의해 $R_i(\theta) = R_i(\theta_1, \theta_2, \dots, \theta_n)$ 와 $\tilde{R}_i$ 사이를 최소화하기 위해 운동 파라미터 $\theta_1, \theta_2, \dots, \theta_n$ 의 값들을 계산할 수 있다. 이러한 운동 파라미터의 값을 추정하기 위한 공식은 다음과 같다.

수학식 10

$$\begin{aligned} R_i(\theta) &= R_x(\theta_1)R_y(\theta_2)R_z(\theta_3)R_x(\theta_4)...R_y(\theta_{k-1})R_z(\theta_k) \\ dR_i &= \tilde{R}_i - R_i(\theta) \\ J_i &= [J_i^1, J_i^2, \dots, J_i^k], J_i^j = \text{Vec}(\partial R_i / \partial \theta_j) \\ P_{t/t-1} &= Q_\sigma + P_{t-1/t-1} \\ S_t &= J_t P_{t/t-1} J_t^T + R_\varsigma \\ K_t &= P_{t/t-1} J_t^T S_t^{-1} \\ \theta &= \theta + K_t dR_i \\ P_{t/t} &= P_{t/t-1} - K_t J_t P_{t/t-1} \end{aligned}$$

[0054]

[0055] 수학식 10에서, $Q_\sigma = \sigma \cdot I$ 이고, $R_\varsigma = \varsigma \cdot I$ 은 가우시안 노이즈의 공분산 매트릭스를 표현하는 대각행렬이다. 여기서, σ , ς 은 양의 상수이고, I 는 단위행렬이다.

[0056] 이상과 같은 과정을 거치면, 피팅된 3D 인체 모델을 3D 모션으로 출력한다(111). 이렇게 출력되는 3D 모션의

예는 도 7에 나타나 있다.

[0057] 도 7은 최종 구현된 3D 인체 모션의 예시도이다. 이는 도 2와 도 3의 영상을 3D 인체 모션으로 구현한 것이다. 최종 구현된 3D 인체 모션의 경우 각 인체 부분을 타원체의 형상에서 인체 부분에 더욱 가까운 수퍼 2차 곡면 함수(super quadrics)로 표현한 예이다. 이는 앞에서 설명한 라벨링 과정을 거침으로써 보다 인체에 가까운 형상으로 수정되기 때문이다. 이에 따라, 각각의 인체 동작에 따른 3D 인체 모델의 데이터 없이도 보다 정교한 3D 인체 모션을 획득할 수 있다.

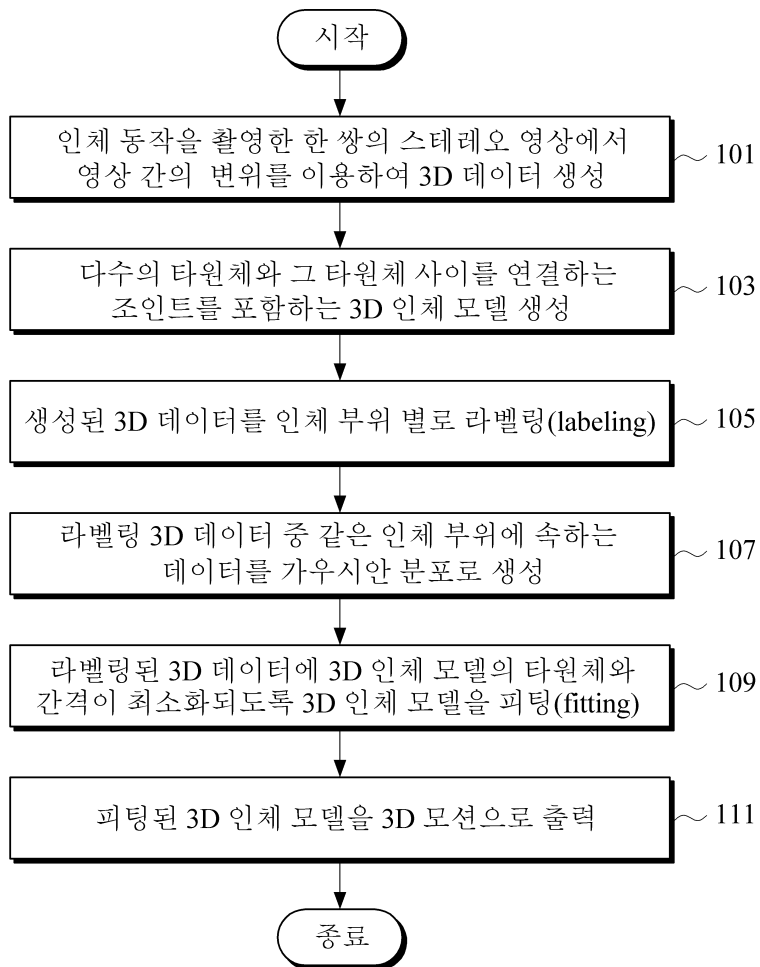
[0058] 이상에서 본 발명은 도면을 참조하면서 기술되는 바람직한 실시예를 중심으로 설명되었지만 이에 한정되는 것은 아니다. 따라서 본 발명은 기재된 실시예로부터 도출 가능한 자명한 변형예를 포괄하도록 의도된 특허청구범위의 기재에 의해 해석되어야 한다.

부호의 설명

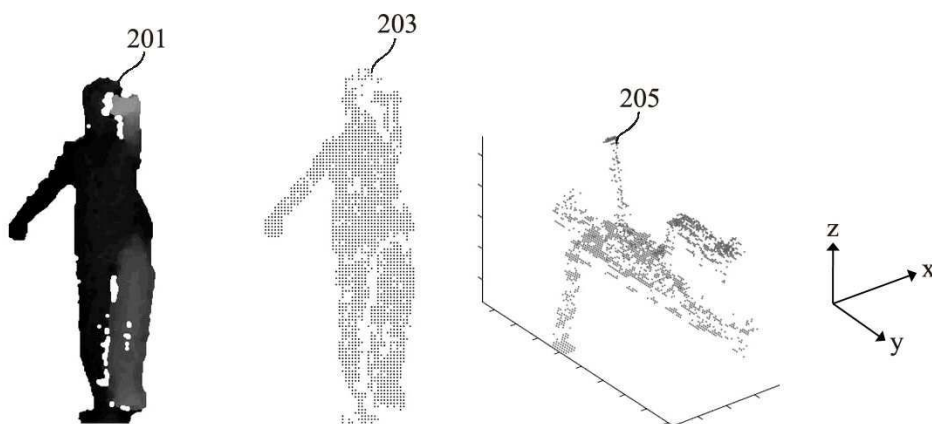
[0059] 201 : 불일치 영상(disparity image)
 203 : 샘플링된 불일치 영상
 205 : 3D 영상
 301 : 뼈대 구조 모델
 303 : 세그먼트
 305 : 조인트
 307 : 타원체 모델
 309 : 타원체
 401 : 인체 영상
 403 : 이진 영상
 405 : 수정된 이진 영상
 407 : 삭제영역
 501 : 실루엣
 503 : 머리 영역
 505 : 토르소 영역
 507 : 머리 영역의 중심점
 509 : 토르소 영역의 중심점
 511 : 실루엣의 중심점
 601 : 3D 영상
 603 : 최단경로

도면

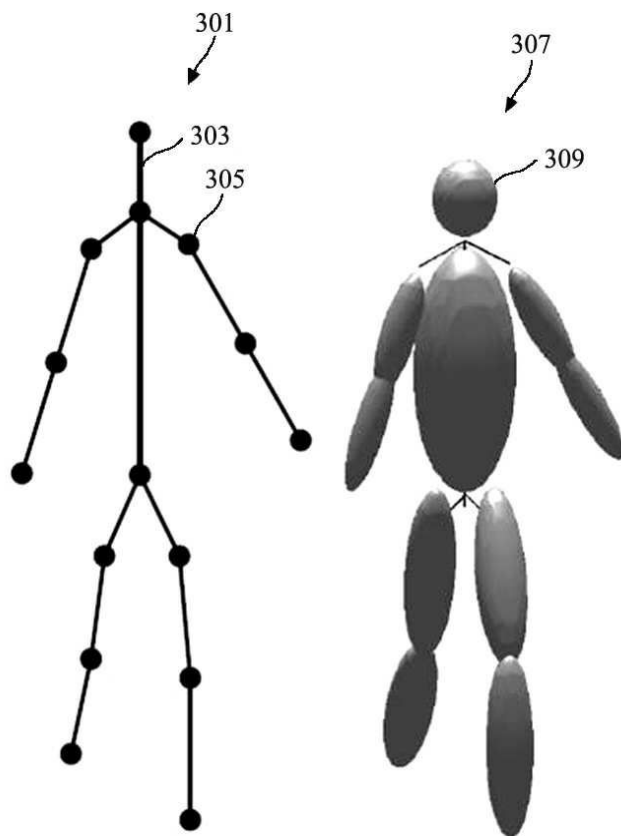
도면1



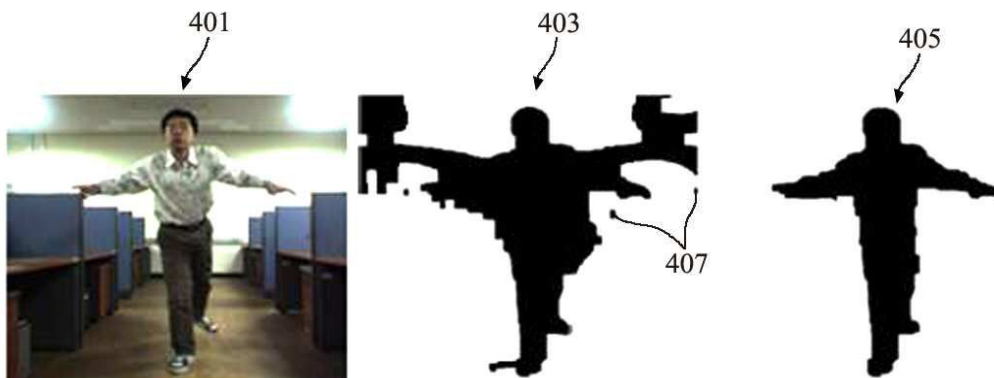
도면2



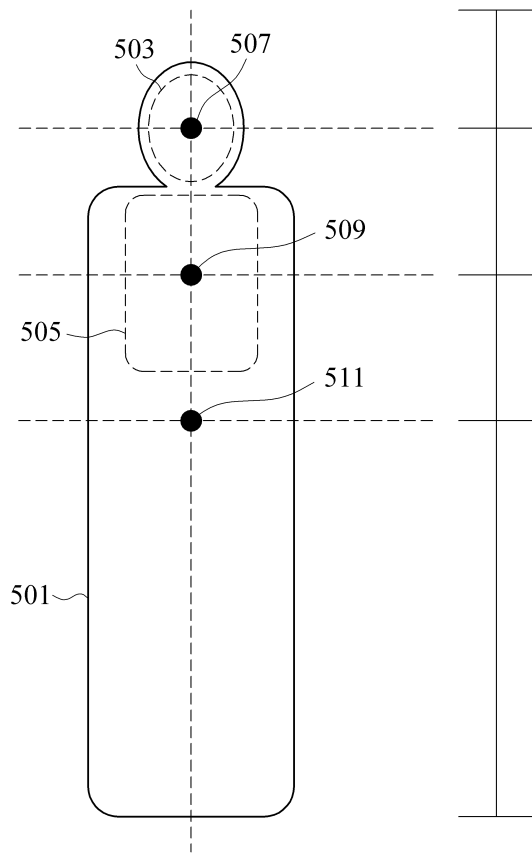
도면3



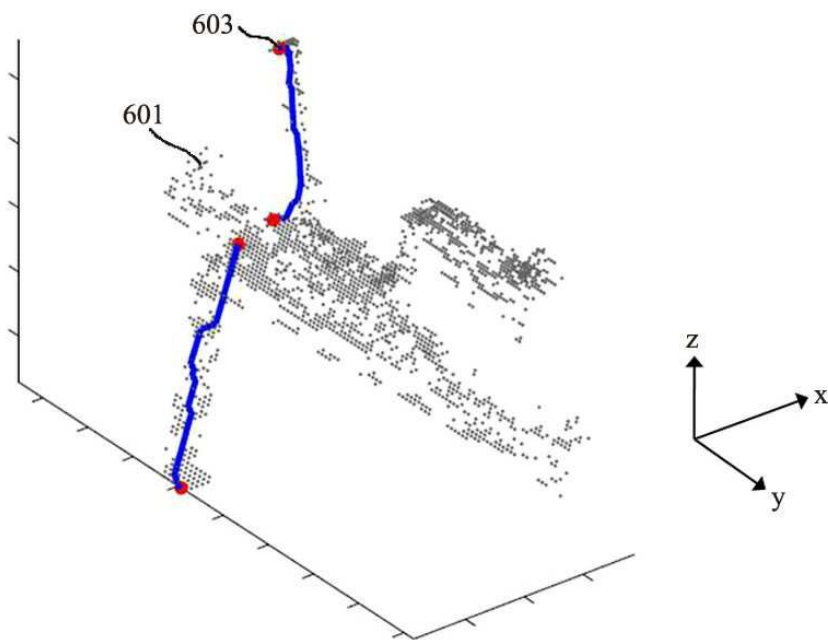
도면4



도면5



도면6



도면7

