



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0033608
(43) 공개일자 2025년03월10일

(51) 국제특허분류(Int. Cl.)
G06T 13/40 (2011.01) G06T 17/20 (2006.01)
G06T 7/246 (2017.01)
(52) CPC특허분류
G06T 13/40 (2013.01)
G06T 17/205 (2013.01)
(21) 출원번호 10-2023-0115997
(22) 출원일자 2023년09월01일
심사청구일자 2023년09월01일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암
동5가)
(72) 발명자
최성준
경기도 하남시 위례대로6길 20, 7107호 403호
채윤병
서울특별시 성북구 서경로13길 27, 302호
(74) 대리인
김동용

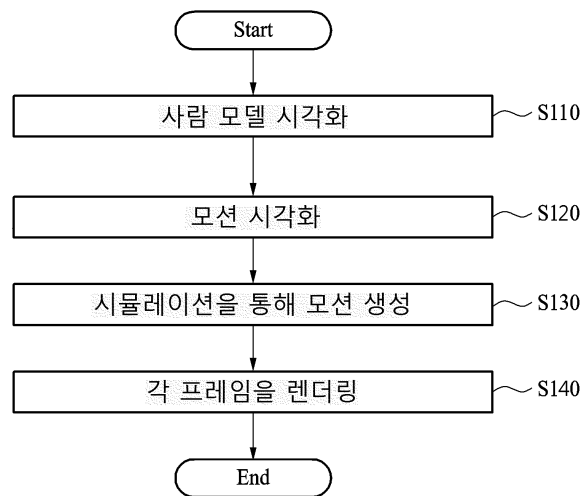
전체 청구항 수 : 총 6 항

(54) 발명의 명칭 사람 동작 데이터와 사람의 이미지를 합성하여 동영상을 생성하는 방법

(57) 요약

사람 동작 데이터와 사람의 이미지를 합성하여 동영상을 생성하는 방법이 개시된다. 상기 동영상 생성 방법은 적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되고, 제1 동영상으로부터 상기 제1 동영상에 등장하는 사람에 대한 SMPL(Skinned Multi-Person Linear Model) 메시(Mesh)를 생성하는 단계, 모션 데이터로부터 모션을 시각화하는 단계, 시각화 결과를 이용하여 상기 모션 데이터의 모션을 시뮬레이션하는 단계, 및 렌더링 기법을 이용하여 상기 제1 동영상에 등장하는 사람이 상기 모션 데이터의 모션을 수행하는 제2 동영상을 생성하는 단계를 포함한다.

대표도 - 도1



(52) CPC특허분류

G06T 7/251 (2017.01)

G06T 2207/30196 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711193973
과제번호	2019-0-00079-005
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성
연구과제명	인공지능대학원지원(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2023.01.01 ~ 2023.12.31

명세서

청구범위

청구항 1

적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 동영상 생성 방법에 있어서,

제1 동영상으로부터 상기 제1 동영상에 등장하는 사람에 대한 SMPL(Skinned Multi-Person Linear Model) 메시(Mesh)를 생성하는 단계;

모션 데이터로부터 모션을 시각화하는 단계;

시각화 결과를 이용하여 상기 모션 데이터의 모션을 시뮬레이션하는 단계; 및

렌더링 기법을 이용하여 상기 제1 동영상에 등장하는 사람이 상기 모션 데이터의 모션을 수행하는 제2 동영상을 생성하는 단계를 포함하는 동영상 생성 방법.

청구항 2

제1항에 있어서,

상기 시각화하는 단계는 자세 추정(Pose Estimation) 기법이나 데이터셋을 활용하여 모션 데이터를 얻고, 모션 데이터로부터 사람의 모션을 3D 직교 작업 공간에 시각화하는,

동영상 생성 방법.

청구항 3

제2항에 있어서,

상기 시뮬레이션하는 단계는 상기 SMPL 메시로부터 측정된 치수에 기초하여 스켈레톤을 생성하는,

동영상 생성 방법.

청구항 4

제3항에 있어서,

상기 시뮬레이션하는 단계는 Isaac 시뮬레이터를 이용하여 상기 시각화 결과로부터 모션 시뮬레이션을 수행하는,

동영상 생성 방법.

청구항 5

제4항에 있어서,

상기 제2 동영상을 생성하는 단계는 상기 제1 동영상의 사람이 상기 모션 데이터의 모션을 수행하는 동영상인,

동영상 생성 방법.

청구항 6

제5항에 있어서,

상기 제2 동영상을 생성하는 단계는 NeRF를 이용하여 생성되는,

동영상 생성 방법.

발명의 설명

기술 분야

[0001] 본 발명은 딥러닝 모델 기반 동영상 생성 방법에 관한 것으로, 특히 모션 (동작) 데이터를 입력으로 주었을 때 해당 모션을 수행하는 사람이 포함된 (총 시간)*(높이)*(너비) 차원의 동영상을 생성하는 장치 및 방법에 관한 것이다.

배경 기술

[0002] AMP(Adversarial Motion Prior)를 활용하여 시뮬레이터 내에서 모방학습을 통해 물리적으로 가능한 모션을 생성할 수 있지만, 이러한 모션은 단순히 스켈레톤이 움직이는 모션이기 때문에 물리적으로 가능하지만 실제 사람이 움직이는 것처럼 렌더링하지 못한다.

[0003] 그리고, NeRF나 이를 활용한 NeuMan은 사실적으로 사람을 렌더링할 수 있지만 이 모션을 새로운 배경에 옮길 경우 사람이 보기에 부자연스럽게 나올 수 있다. 이는 물리적으로 가능하지 않은 모션이 나오기 때문일 수 있는데, 사람의 발이 바닥으로부터 떨어져 있거나 사람의 신체 일부분이 서로 관통할 수 있기 때문에 발생한다.

[0004] 본 발명에서는, 상술한 두 기술의 장점을 합치고 단점을 보완하기 위해 물리적으로 가능하면서(Physically Plausible) 사진처럼 사실적인(Photo-realistic) 동영상 생성 기법을 제안한다.

선행기술문헌

특허문헌

[0005] (특허문헌 0001) 대한민국 공개특허 제2020-0032614호 (2020.03.26. 공개)

(특허문헌 0002) 대한민국 공개특허 제2023-0023804호 (2023.02.17. 공개)

비특허문헌

[0006] (비특허문헌 0001) Caroline Chan, Shiry Ginossar, Tinghui Zhou, and Alexei A. Efros, "Everybody Dance Now", arXiv:1808.07371v2, 27 Aug 2019.

발명의 내용

해결하려는 과제

[0007] 본 발명이 이루고자 하는 기술적인 과제는 사람 동작 데이터와 사람의 이미지를 합성하여 동영상을 생성하는 장치 및 방법을 제공하는 것이다.

과제의 해결 수단

[0008] 본 발명의 일 실시예에 따른 동영상 생성 방법은 적어도 프로세서(processor)를 포함하는 컴퓨팅 장치에 의해 수행되는 동영상 생성 방법으로써, 제1 동영상으로부터 상기 제1 동영상에 등장하는 사람에 대한 SMPL(Skinned Multi-Person Linear Model) 메시(Mesh)를 생성하는 단계, 모션 데이터의 모션을 시각화하는 단계, 시각화 결과를 이용하여 상기 모션 데이터의 모션을 시뮬레이션하는 단계, 및 렌더링 기법을 이용하여 상기 제1 동영상에 등장하는 사람이 생성하고자 하는 모션 데이터의 모션을 수행하는 제2 동영상을 생성하는 단계를 포함한다.

발명의 효과

[0009] 동작을 기반으로 시뮬레이션을 통과한 후 동영상을 생성하는 파이프라인은 아직 완성된 적이 없으며, 하나의 파이프라인 안에서 원하는 모션을 선택하고 사람이 사실적으로 움직이는 동영상을 생성할 수 있다.

[0010] 단순히 사람의 영상이 나오는 것이 아니라 공을 차거나 벽에 부딪히는 것과 같은 상호작용 등 또한 시뮬레이션을 통해 표현할 수 있기 때문에 사용자는 더 다양한 상황에 대한 영상을 만들 수 있고, 영상의 시청자들은 영상을 자연스럽게 느낄 수 있다.

[0011] 이는 위급 상황 시 대처 요령과 같이 글만으로는 부족한 정보들에 동영상 보조 자료 생성에 기여할 수 있으며, 콘티와 시나리오와 같이 동작 정보가 포함됐지만 텍스트만으로 표현된 매체를 가지고 이해를 도울 수 있는 동영상

상 매체를 생성하는 데에 사용할 수 있다.

도면의 간단한 설명

- [0013] 본 발명의 상세한 설명에서 인용되는 도면을 보다 충분히 이해하기 위하여 각 도면의 상세한 설명이 제공된다.
- 도 1은 본 발명의 일 실시예에 따른 동영상 생성 방법을 설명하기 위한 흐름도이다.
- 도 2는 제1 동영상으로부터 생성된 SMPL 메시지를 도시한다.
- 도 3은 생성하고자 하는 모션 데이터의 모션에 포함된 프레임들 중 어느 하나의 프레임에 대한 시각화 결과를 도시한다.
- 도 4는 시물레이션을 위해 생성된 스켈레톤 MJCF를 도시한다.
- 도 5는 리타게팅의 개략적인 과정을 설명하기 위한 도면이다.
- 도 6은 모션 시물레이션 결과를 도시한다.
- 도 7은 SMPL 시각화 결과를 도시한다.
- 도 8은 도 1에 도시된 동영상 생성 방법을 설명하기 위한 전체 개략도이다.
- 도 9는 Everybody Dance Now 논문에 있는 실험 결과를 도시한다.
- 도 10은 자세 추정 기술을 적용하여 획득된 포즈를 도시한다.
- 도 11은 사진으로부터 포즈를 얻는 모델 P와 포즈로부터 그 사진(프레임)을 다시 생성하는 모델 G를 포함하는 모델을 도시한다.
- 도 12는 본 발명에서 이용하는 자세 제작 리그를 도시한다.
- 도 13은 시뮬레이터에 의해 재현된 모션을 도시한다.
- 도 14는 NeRF를 설명하기 위한 개념도이다.
- 도 15는 SMPL의 입력과 출력을 도시한다.
- 도 16은 NeuMan 기법의 개략도이다.
- 도 17은 NeuMan 기법을 설명하기 위한 도면이다.

발명을 실시하기 위한 구체적인 내용

- [0014] 본 명세서에 개시되어 있는 본 발명의 개념에 따른 실시예들에 대해서 특정한 구조적 또는 기능적 설명들은 단지 본 발명의 개념에 따른 실시예들을 설명하기 위한 목적으로 예시된 것으로서, 본 발명의 개념에 따른 실시예들은 다양한 형태로 실시될 수 있으며 본 명세서에 설명된 실시예들에 한정되지 않는다.
- [0015] 본 발명의 개념에 따른 실시예들은 다양한 변경들을 가할 수 있고 여러 가지 형태들을 가질 수 있으므로 실시예들을 도면에 예시하고 본 명세서에서 상세하게 설명하고자 한다. 그러나, 이는 본 발명의 개념에 따른 실시예들을 특정한 개시 형태들에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물, 또는 대체물을 포함한다.
- [0016] 제1 또는 제2 등의 용어는 다양한 구성 요소들을 설명하는데 사용될 수 있지만, 상기 구성 요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성 요소를 다른 구성 요소로부터 구별하는 목적으로만, 예컨대 본 발명의 개념에 따른 권리 범위로부터 벗어나지 않은 채, 제1 구성 요소는 제2 구성 요소로 명명될 수 있고 유사하게 제2 구성 요소는 제1 구성 요소로도 명명될 수 있다.
- [0017] 어떤 구성 요소가 다른 구성 요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성 요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성 요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성 요소가 다른 구성 요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는 중간에 다른 구성 요소가 존재하지 않는 것으로 이해되어야 할 것이다. 구성 요소들 간의 관계를 설명하는 다른 표현들, 즉 "~사이에"와 "바로 ~사이에" 또는 "~에 이웃하는"과 "~에 직접 이웃하는" 등

도 마찬가지로 해석되어야 한다.

- [0018] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로서, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 본 명세서에 기재된 특징, 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성 요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0019] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가진다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥상 가지는 의미와 일치하는 의미를 갖는 것으로 해석되어야 하며, 본 명세서에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0020] 이하, 본 명세서에 첨부된 도면들을 참조하여 본 발명의 실시예들을 상세히 설명한다. 그러나, 특허출원의 범위가 이러한 실시예들에 의해 제한되거나 한정되는 것은 아니다. 각 도면에 제시된 동일한 참조 부호는 동일한 부재를 나타낸다.
- [0021] 우선, 본 발명의 배경이 되는 기술에 대해 설명한다.
- [0022] 1) Everybody Dance Now(<https://arxiv.org/abs/1808.07371>)
- [0023] 사람의 영상이 주어졌을 때 그 사람이 입은 옷과 그 사람의 체형을 기반으로 동작을 합성하여 새로운 동작을 하는 사람 동영상 생성하는 기법으로, 도 8에는 Everybody Dance Now 논문에서 실험 결과가 도시되어 있다. 해당 논문에서는 일반인이 따라하기 힘든 발레리나의 모션을 다른 사람이 따라할 수 있도록 실험을 진행하였다. 동영상의 각 프레임은 곧 사진이다. 각 프레임 별로 사람의 전신 형태가 있으면 자세 추정(Pose estimation) 기술을 적용하여 도 10의 사진과 같은 포즈를 얻을 수 있다.
- [0024] 도 11의 P는 사진으로부터 포즈를 얻는 모델이고 G는 포즈로부터 그 사진(프레임)을 다시 생성하는 모델이다. 목표 비디오로부터 배경과 사람을 생성하는 G를 학습한다. 전문가의 영상 프레임으로부터 포즈를 얻고 그 포즈를 목표 사람의 신체 모양에 맞게 조절(Norm)한다. 조절된 포즈를 G 모델에 통과시키면 목표 비디오의 배경과 사람이 생성된다. 해당 논문은 사람 모션이 포함된 비디오 생성(합성)의 시작점이 되는 논문이다.
- [0025] 2) GPU 연산으로 Articulated Body Algorithm을 계산하여 한 번에 수천 개의 에이전트의 물리 시뮬레이션이 가능한 Isaac Simulator
- [0026] 본 발명에서는 로봇을 사용하지 않고 사람의 신체 구조를 대표할 수 있는 자체 제작 리그(Rig)를 사용한다. 자체 제작된 리그는 도 12에 도시되어 있다.
- [0027] 3) 시뮬레이터 내에서 모방학습을 통한 물리적으로 가능한 모션 생성
- [0028] 기존 동작에서 비디오를 생성하는 기술들은 물리 시뮬레이션을 하지 않고 단순히 동작을 사람 모델에 대입하여 영상을 생성하기 때문에 신체 구조에 따라 발이 땅에 제대로 붙어있지 않거나 팔과 다리와 같은 신체 부분들이 서로 관통하는 문제가 있었다. 물리 시뮬레이터에서 모션을 모방하는 학습 기술(Imitation learning, AMP(Adversarial Motion Prior) 논문 참조)을 통해 시뮬레이터에서 원본 모션을 재현할 수 있다. 도 13은 물리 시뮬레이터 안에서 생성된 모션이기 때문에 물리적으로 가능한 모션이다. 이러한 재현된 모션은 물리 시뮬레이터에서 연산된 결과이기 때문에 발이 떠있거나 신체 부분들이 서로 통과하는 결과가 나오지 않는다. 또 너무 무거운 팔이나 다리의 경우 조인트에 걸리는 토크가 너무 세게 걸리지 않도록 할 수 있다. 단순히 원본 모션을 따라하는 것뿐만 아니라 특정 작업을 수행하도록 강화학습의 보상함수를 작성하여 작업을 수행하는 모션을 만들 수도 있다.
- [0029] 4) NeRF 기반 비디오 생성
- [0030] NeRF는 카메라의 위치와 광선의 방향을 주었을 때 그 광선에 해당하는 RGB 색과 밀도(투명도)를 출력하는 모델이다. 임의의 카메라의 위치와 광선의 방향을 줄 때 RGB 색과 밀도를 잘 예측할 수 있기 때문에 view synthesis 분야에서 가장 대표적인 기술이다.
- [0031] 5) SMPL & SMPLify

- [0032] SMPL은 사람의 형태(shape)와 자세(pose)가 주어졌을 때, 그 사람의 피부 메시(Mesh)를 사실적으로 표현할 수 있는 모델이다. 모델의 입력과 출력은 도 15에 도시되어 있다. SMPLify는 사람의 사진이 주어졌을 때 그 사람의 형태와 포즈를 가장 잘 대변하는 변수들을 찾는 최적화 기법이다.
- [0033] 6) SMPL과 NeRF를 이용한 비디오 생성(합성)
- [0034] NeuMan은 Apple에서 발표한 논문으로 비디오를 생성할 때 사람의 형태를 SMPLify를 통해 가져오고, NeRF에 카메라의 위치와 방향을 주었을 때 SMPL 안에 있는 영역에만 색깔을 입혀 사람의 형태 안에서만 RGB 색과 밀도를 가져온다. 도 16에는 NeuMan 기법의 개략도가 도시되어 있다. 영상으로부터 사람의 형태나 인상 착의에 대한 정보를 추출할 수 있다. 영상에서 보지 않은 새로운 자세가 주어졌을 경우에 SMPL을 통해 자세를 만들고 추출한 정보로부터 사실적인 렌더링을 NeRF가 수행할 수 있다. 배경에 대한 정보를 NeRF가 따로 학습하기 때문에 사람의 정보를 학습한 NeRF를 다른 영상으로부터 학습한 뒤 대체할 경우 다른 영상의 인물을 이 영상에 대입할 수 있다. 반대로 인물은 그대로 두고 배경만 다른 영상으로 대체하는 것도 가능하다.
- [0035] 도 1은 본 발명의 일 실시예에 따른 동영상 생성 방법을 설명하기 위한 흐름도이다.
- [0036] 도 1을 참조하면, 동영상 생성 방법은 적어도 프로세서(processor) 및/또는 메모리(memory)를 포함하는 컴퓨팅 장치에 의해 수행될 수 있다. 다시 말하면, 동영상 생성 방법을 이루는 단계들 중 적어도 일부는 컴퓨팅 장치에 포함되는 프로세서의 동작으로 이해될 수도 있다. 따라서, 컴퓨팅 장치는 동영상 생성 장치로 명명될 수 있으며, PC(Personal Computer), 서버(server), 랩탑 컴퓨터, 태블릿 PC 등을 포함할 수 있다.
- [0037] 우선, 사람 모델 최적화 동작이 수행된다(S110). 구체적으로, 생성되는 동영상에 등장시킬 사람의 움직임을 포함하고 있는 제1 동영상으로부터 SMPL 메시(Mesh)가 생성될 수 있다. 예시적으로, SMPL 메시는 SMPL(A Skinned Multi-Person Linear Model) 및/또는 SMPLify를 이용하여 생성될 수 있다. 생성된 SMPL 메시로부터 각 치수가 산출될 수 있다. 각 치수는 SMPL 메시의 정점(Vertex)으로부터 획득될 수 있다. 여기서, (신체) 치수라 함은 SMPL 메시의 크기(키), 너비, 관절과 관절을 잇는 링크의 길이 등을 포함할 수 있다. 생성된 SMPL 메시는 도 2에 도시되어 있다.
- [0038] 모션 시각화 동작이 수행된다(S120). 생성된 동영상에서의 모션을 포함하고 있는 모션 데이터 내의 모션에 대한 시각화 동작을 의미한다. 모션 데이터 내의 사람의 (회전) 조인트(joint)와 조인트와 조인트 사이를 잇는 링크(link)를 시각화하고 모션 데이터에 따라 회전 조인트가 돌아가면서 모션을 3D 직교(3D Cartesian) 작업 공간(Task space)에서 시각화할 수 있다. 도 3에는 모션 데이터에 포함된 프레임들 중 어느 하나의 프레임에 대한 시각화 결과가 도시되어 있다. 다시 말하면, 모션 데이터 내의 각 프레임에서 포즈(pose)를 얻을 수 있다.
- [0039] 시뮬레이션을 통해 모션이 생성된다(S130). SMPL은 시뮬레이션을 위한 모델이 아니기 때문에, 시뮬레이터에 호환되는 MJCF로 바뀌어 주어야 한다. SMPL 메시의 정점(Vertex)로부터 측정된 치수로부터 (사람 형태의) 스켈레톤 MJCF가 생성될 수 있다. 생성된 스켈레톤 MJCF는 도 4에 도시되어 있다.
- [0040] 실시예에 따라, 모션 데이터에 대한 리타게팅이 수행될 수도 있다. 리타게팅은 원본 모션과 원본 신체 형태가 있을 때, 이 모션을 다른 신체 형태로 옮기기 위한 방법론이다. 모션 데이터를 측정된 치수의 신체 형태에 적합하도록 모션을 리타게팅하여 만들 수 있다. 즉, 모션 시각화 결과에 측정된 치수를 적용함으로써 리타게팅이 수행될 수 있다. 리타게팅에 대한 개략적인 과정은 도 5에 도시되어 있다.
- [0041] 실시예에 따라, 강화학습을 적용하여 모션을 모방하는 시뮬레이션을 통해 모션이 생성될 수 있다. 예시적인, 시뮬레이션 결과는 도 6에 도시되어 있다. 또한, 본 발명에서 사용하는 시뮬레이터는 Isaac Simulator일 수 있다.
- [0042] 각 프레임에 대한 렌더링 동작이 수행될 수 있다(S140). 시뮬레이터에서 적당한 모션이 잘 시뮬레이션 되는 경우, 모방 학습을 멈추고, SMPL 시각화를 위한 자세(pose) 변수가 획득될 수 있다. 자세 변수는 시뮬레이터(즉, 시뮬레이션 결과)로부터 획득될 수 있다. SMPL 시각화 결과는 도 7에 도시되어 있다.
- [0043] 예시적인 렌더링 모델인 NeRF는 배경과 사람에 대한 정보를 추출하기 때문에 SMPL 메시가 만드는 모션을 통해 각 프레임을 렌더링할 수 있다. 공과 같은 물체 또한 NeRF로 원하는 형태와 색깔로 렌더링할 수 있다. 이를 통해, 제1 동영상에 등장하는 사람이 모션 데이터의 모션을 수행하는 새로운 동영상이 생성될 수 있다. 도 8에는 전체적인 과정에 대한 개략도가 도시되어 있다.
- [0044] 이상에서 설명된 장치는 하드웨어 구성 요소, 소프트웨어 구성 요소, 및/또는 하드웨어 구성 요소 및 소프트웨어 구성 요소의 집합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성 요소는, 예를 들어, 프로세서, 컨트롤러, ALU(Arithmetic Logic Unit), 디지털 신호 프로세서(Digital Signal Processor), 마이크

로컴퓨터, FPA(Field Programmable array), PLU(Programmable Logic Unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(Operation System, OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술 분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(Processing Element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(Parallel Processor)와 같은, 다른 처리 구성(Processing Configuration)도 가능하다.

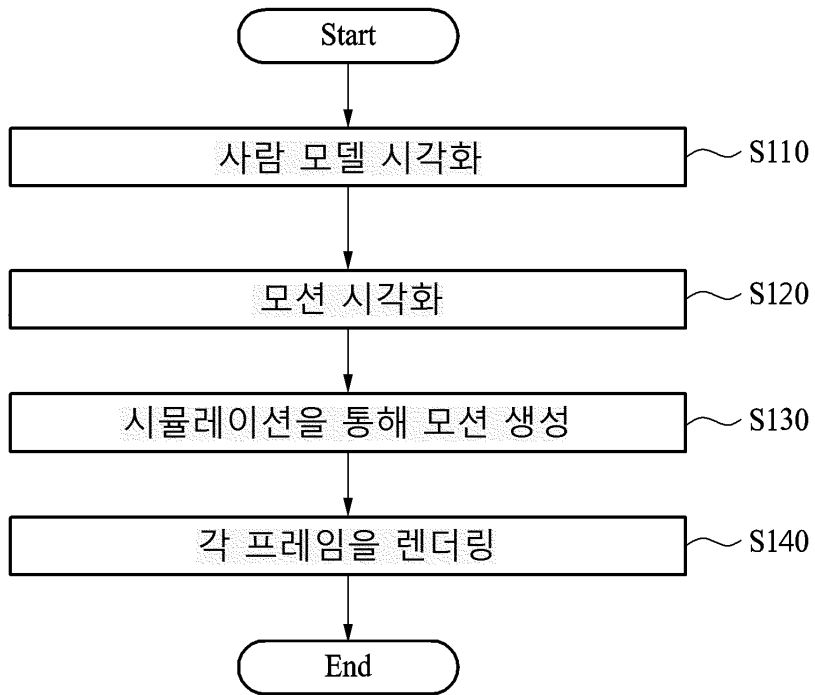
[0045] 소프트웨어는 컴퓨터 프로그램(Computer Program), 코드(Code), 명령(Instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(Collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성 요소(Component), 물리적 장치, 가상 장치(Virtual Equipment), 컴퓨터 저장 매체 또는 장치, 또는 전송되는 신호 파(Signal Wave)에 영구적으로, 또는 일시적으로 구체화(Embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.

[0046] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(Magnetic Media), CD-ROM, DVD와 같은 광기록 매체(Optical Media), 플롭티컬 디스크(Floptical Disk)와 같은 자기-광 매체(Magneto-optical Media), 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다. 상기된 하드웨어 장치는 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지이다.

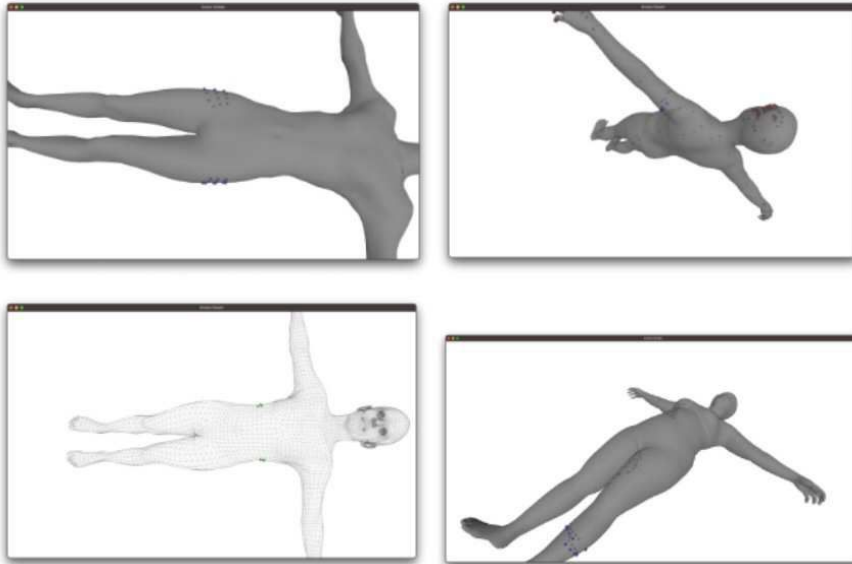
[0047] 본 발명은 도면에 도시된 실시예를 참고로 설명되었으나 이는 예시적인 것에 불과하며, 본 기술 분야의 통상의 지식을 가진 자라면 이로부터 다양한 변형 및 균등한 타 실시예가 가능하다는 점을 이해할 것이다. 예를 들어, 설명된 기술들이 설명된 방법과 다른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성 요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성 요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다. 따라서, 본 발명의 진정한 기술적 보호 범위는 첨부된 등록청구범위의 기술적 사상에 의해 정해져야 할 것이다.

도면

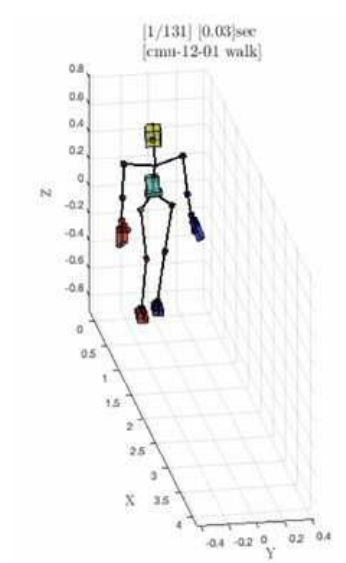
도면1



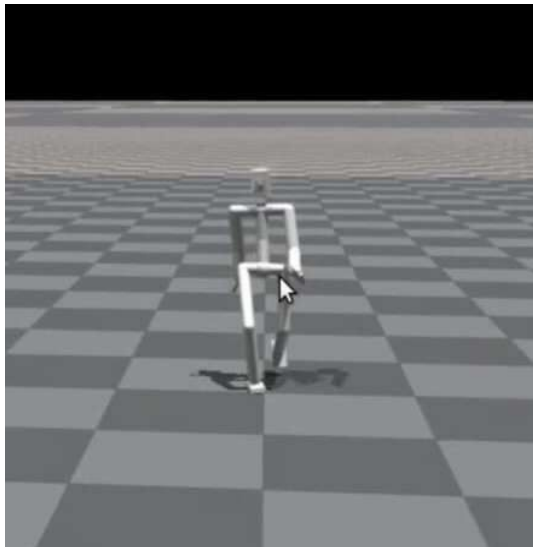
도면2



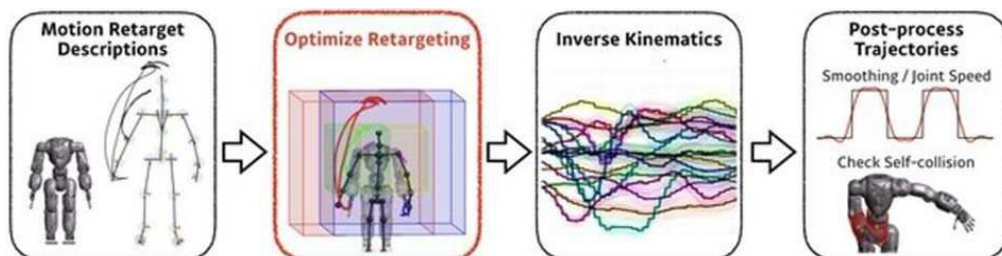
도면3



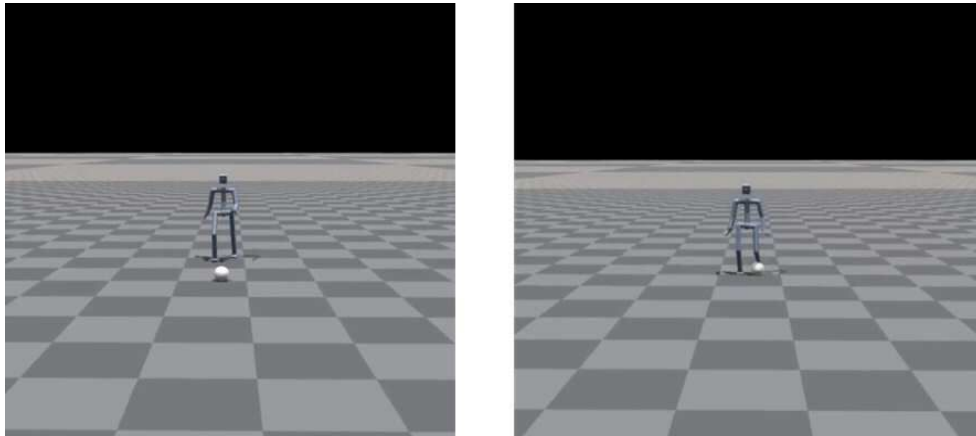
도면4



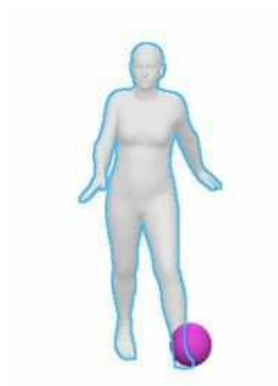
도면5



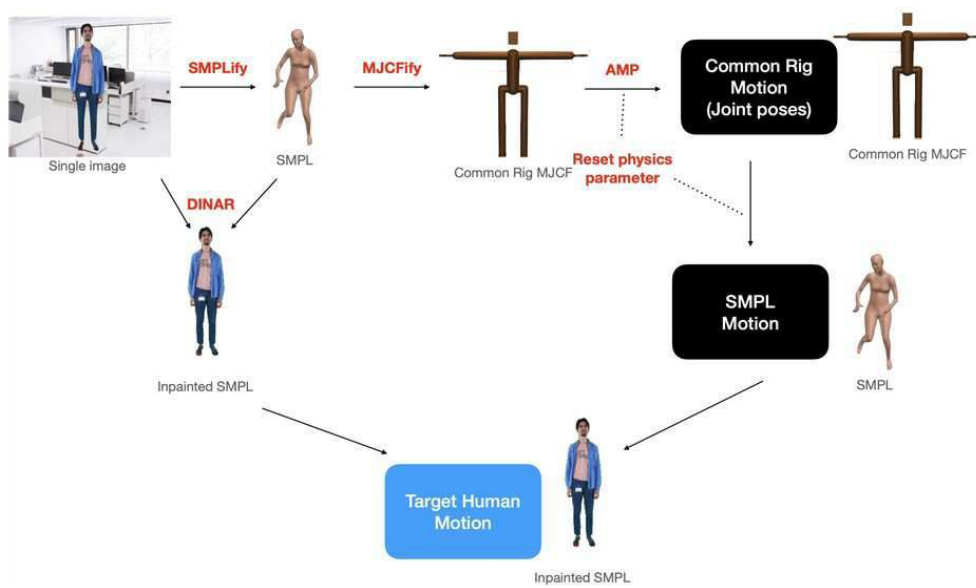
도면6



도면7



도면8



도면9

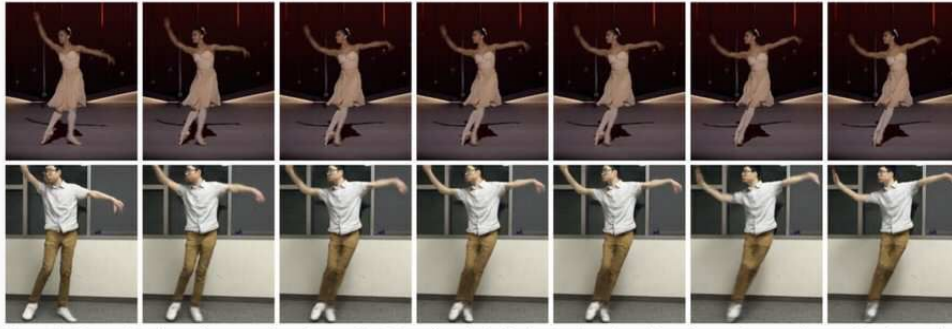
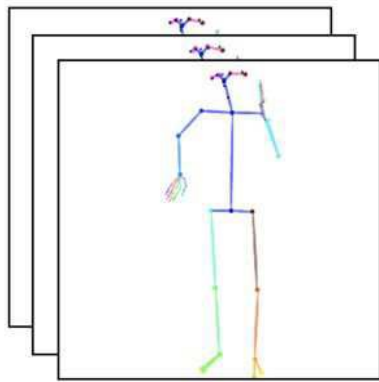
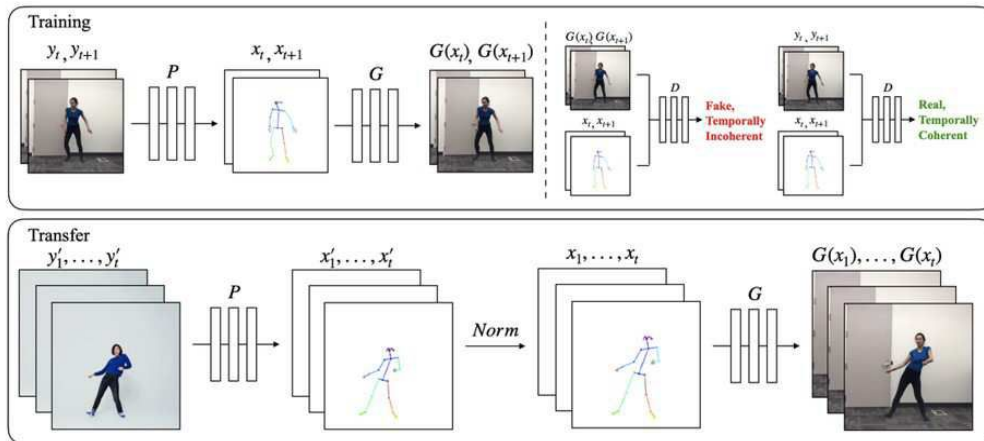


Figure 1: “Do as I Do” motion transfer: given a YouTube clip of a ballerina (top), and a video of a graduate student performing various motions, our method transfers the ballerina’s performance onto the student (bottom). Video: <https://youtu.be/mSaIrz81M1U>

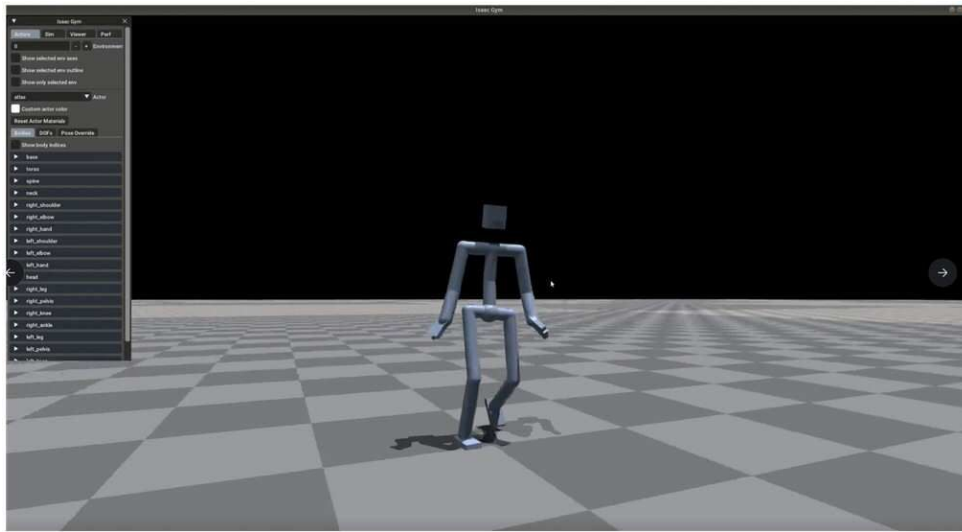
도면10



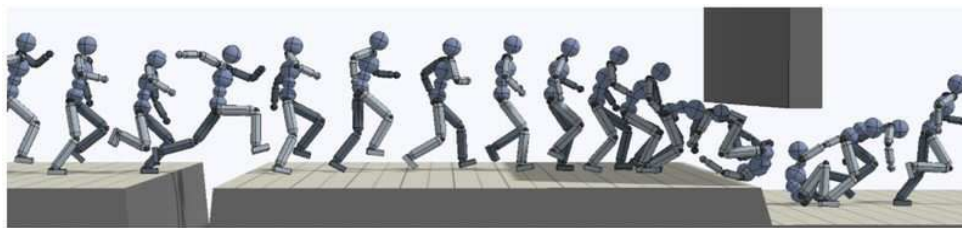
도면11



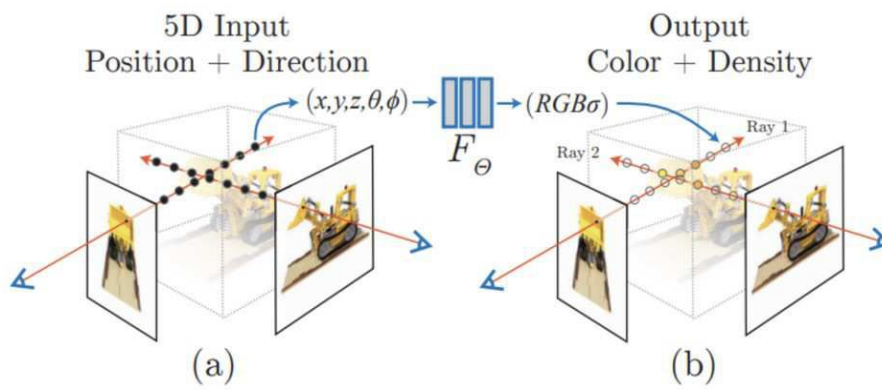
도면12



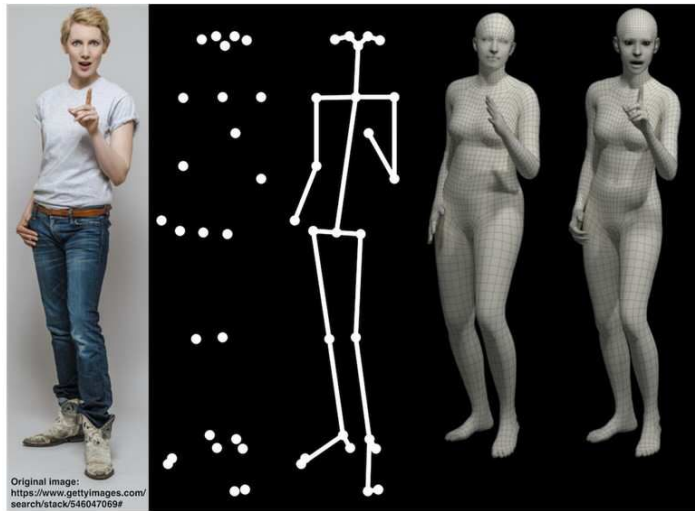
도면13



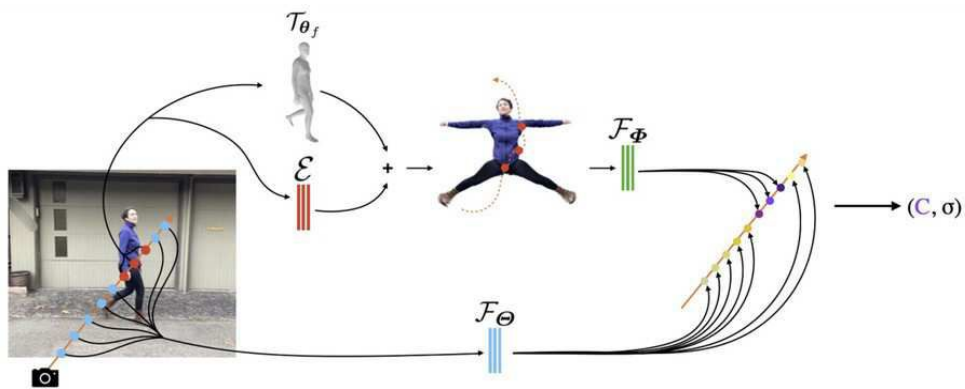
도면14



도면15



도면16



도면17

