



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0095224
(43) 공개일자 2023년06월29일

(51) 국제특허분류(Int. Cl.)
G06T 5/00 (2019.01) G06T 17/20 (2006.01)
G06T 7/70 (2017.01)
(52) CPC특허분류
G06T 5/001 (2013.01)
G06T 17/205 (2013.01)
(21) 출원번호 10-2021-0184479
(22) 출원일자 2021년12월22일
심사청구일자 2021년12월22일

(71) 출원인
고려대학교 산학협력단
서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)
(72) 발명자
김승룡
서울특별시 성북구 오패산로 46, 103동 1804호(하월곡동, 월곡두산위브아파트)
김미라
서울특별시 성북구 고려대로28길 17-3, 해랑빌딩 402호(안암동5가)
(74) 대리인
송인호, 윤형근, 최관락

전체 청구항 수 : 총 7 항

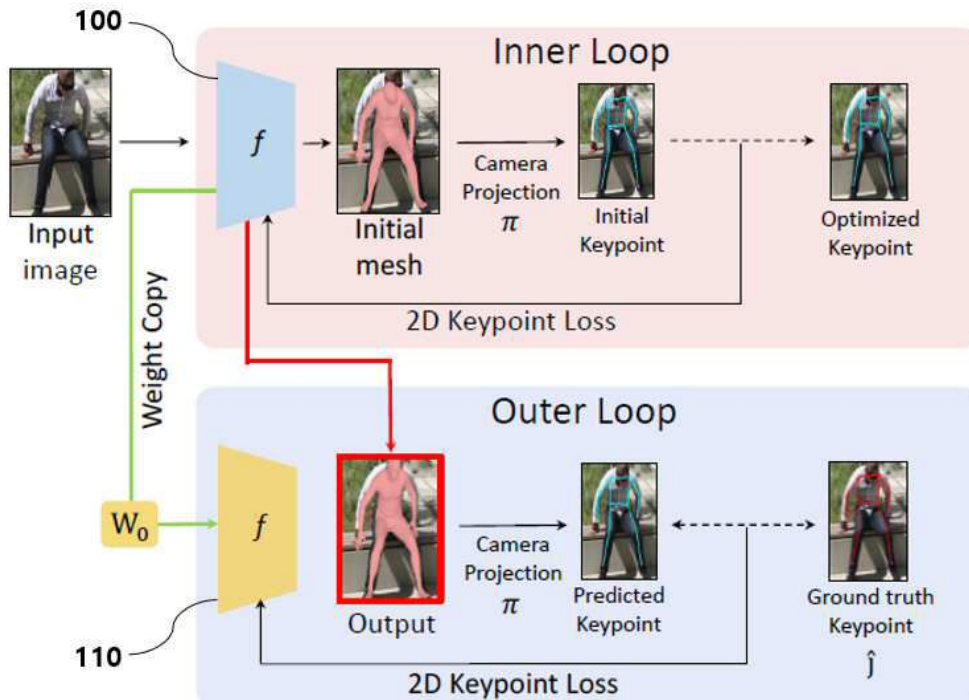
(54) 발명의 명칭 메타 학습을 통한 삼차원 사람 메쉬 복원 방법 및 장치

(57) 요약

본 발명은 메타 학습을 통한 삼차원 사람 메쉬 복원 방법 및 장치를 개시한다. 본 발명에 따르면, 프로세서; 및 상기 프로세서에 연결되는 메모리를 포함하되, 상기 메모리는, 내부 루프의 파라메트릭 모델 기반 제1 신경망에 실제 사람 이미지를 입력하여 상기 실제 사람 이미지에 포함된 사람의 삼차원 메쉬를 추정하고, 상기 삼차원 메

(뒷면에 계속)

대표도 - 도1



위에 관절 회귀 모델을 적용하여 상기 삼차원 메쉬에 상응하는 삼차원 키포인트를 추정하고, 카메라 파라미터를 통해 상기 삼차원 키포인트를 이차원 공간으로 투영하여 이차원 키포인트를 추정하고, 상기 추정된 이차원 키포인트와 이차원 참값에 기초한 손실 함수를 통해 상기 제1 신경망의 네트워크 파라미터를 업데이트하고, 상기 제1 신경망의 제1 초기 네트워크 파라미터가 복사된 외부 루프의 삼차원 파라메트릭 모델 기반 제2 신경망의 제1 초기 네트워크 파라미터를 소정 횟수의 네트워크 파라미터 업데이트가 수행된 제1 신경망에서 출력하는 삼차원 메쉬를 이용하여 제2 초기 네트워크 파라미터로 업데이트하도록, 상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하는 삼차원 사람 메쉬 복원 장치가 제공된다.

(52) CPC특허분류

G06T 7/70 (2017.01)

G06T 2207/20081 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711126325
과제번호	2020-0-01819-002
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	정보통신방송혁신인재양성(R&D)
연구과제명	ICT명품인재양성(고려대학교)
기 여 율	1/1
과제수행기관명	고려대학교산학협력단
연구기간	2021.01.01 ~ 2021.12.31

명세서

청구범위

청구항 1

메타 학습을 통한 삼차원 사람 메쉬 복원 장치로서,

프로세서; 및

상기 프로세서에 연결되는 메모리를 포함하되,

상기 메모리는,

내부 루프의 파라메트릭 모델 기반 제1 신경망에 실제 사람 이미지를 입력하여 상기 실제 사람 이미지에 포함된 사람의 삼차원 메쉬를 추정하고,

상기 삼차원 메쉬에 관절 회귀 모델을 적용하여 상기 삼차원 메쉬에 상응하는 삼차원 키포인트를 추정하고,

카메라 파라미터를 통해 상기 삼차원 키포인트를 이차원 공간으로 투영하여 이차원 키포인트를 추정하고,

상기 추정된 이차원 키포인트와 이차원 참값에 기초한 손실 함수를 통해 상기 제1 신경망의 네트워크 파라미터를 업데이트하고,

상기 제1 신경망의 제1 초기 네트워크 파라미터가 복사된 외부 루프의 삼차원 파라메트릭 모델 기반 제2 신경망의 제1 초기 네트워크 파라미터를 소정 횟수의 네트워크 파라미터 업데이트가 수행된 제1 신경망에서 출력하는 삼차원 메쉬를 이용하여 제2 초기 네트워크 파라미터로 업데이트하도록,

상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하는 삼차원 사람 메쉬 복원 장치.

청구항 2

제1항에 있어서,

상기 제1 신경망은 상기 실제 사람 이미지를 입력으로 하여 결정된 포즈 파라미터 및 형상 파라미터를 이용하여 상기 삼차원 메쉬를 추정하는 삼차원 사람 메쉬 복원 장치.

청구항 3

제1항에 있어서,

상기 프로그램 명령어들은,

상기 제1 신경망에서 출력하는 삼차원 메쉬에 관절 회귀 모델을 적용하여 제2 삼차원 키포인트를 추정하고,

카메라 파라미터를 통해 상기 제2 삼차원 키포인트를 이차원 공간으로 투영하여 제2 이차원 키포인트를 추정하고,

상기 제2 이차원 키포인트와 이차원 참값에 기초한 손실 함수를 통해 상기 제2 신경망의 제1 초기 네트워크 파라미터를 상기 제2 초기 네트워크 파라미터로 업데이트하는 삼차원 사람 메쉬 복원 장치.

청구항 4

제1항에 있어서,

상기 프로그램 명령어들은,

경사 하강법을 이용하여 상기 제1 신경망의 학습 및 상기 제2 초기 네트워크 파라미터로 업데이트를 수행하는 삼차원 사람 메쉬 복원 장치.

청구항 5

제1항에 있어서,

상기 프로그램 명령어들은,

상기 제2 초기 네트워크 파라미터로 업데이트 과정은 매 배치(batch)에 대해 반복 수행하는 삼차원 사람 메쉬 복원 장치.

청구항 6

제1항에 있어서,

상기 프로그램 명령어들은,

모든 배치에 대해 상기 제2 초기 네트워크 파라미터로의 업데이트가 완료된 이후, 테스트 시간에 업데이트가 완료된 제2 초기 네트워크 파라미터를 테스트 신경망에 적용하여 실제 사람 이미지로부터 삼차원 사람 메쉬를 복원하는 삼차원 사람 메쉬 복원 장치.

청구항 7

프로세서 및 메모리를 포함하는 장치에서 메타 학습을 통해 삼차원 사람 메쉬를 복원하는 방법으로서,

내부 루프의 파라메트릭 모델 기반 제1 신경망에 실제 사람 이미지를 입력하여 상기 실제 사람 이미지에 포함된 사람의 삼차원 메쉬를 추정하는 단계;

상기 삼차원 메쉬에 관절 회귀 모델을 적용하여 상기 삼차원 메쉬에 상응하는 삼차원 키포인트를 추정하는 단계;

카메라 파라미터를 통해 상기 삼차원 키포인트를 이차원 공간으로 투영하여 이차원 키포인트를 추정하는 단계;

상기 추정된 이차원 키포인트와 이차원 참값에 기초한 손실 함수를 통해 상기 제1 신경망의 네트워크 파라미터를 업데이트 하는 단계; 및

상기 제1 신경망의 제1 초기 네트워크 파라미터가 복사된 외부 루프의 삼차원 파라메트릭 모델 기반 제2 신경망의 제1 초기 네트워크 파라미터를 소정 횟수의 네트워크 파라미터 업데이트가 수행된 제1 신경망에서 출력하는 삼차원 메쉬를 이용하여 제2 초기 네트워크 파라미터로 업데이트 하는 단계를 포함하는 삼차원 사람 메쉬 복원 방법.

발명의 설명

기술 분야

[0001] 본 발명은 메타 학습을 통한 삼차원 사람 메쉬 복원 방법 및 장치에 관한 것으로서, 보다 상세하게는 메타 학습을 통해 테스트 시간에서 더 빠르게 최적화할 수 있는 복원 기술에 관한 것이다.

배경 기술

[0002] 파라메트릭 모델 기반의 삼차원 사람 복원 기술은 사람을 찍은 이차원 이미지로부터 파라메트릭 모델의 파라미터를 예측하여 삼차원으로 복원하는 AI 시스템으로, 크게 회귀 기반(Regression-based) 접근법과 최적화 기반(Optimization-based) 접근법이 있다.

[0003] 그 중 최적화 기반 접근법은 테스트 시간에서 주어진 이미지에 대해 파라메트릭 모델의 파라미터를 직접 최적화하는 시스템으로, SMPLify, EFT 등의 네트워크가 대표적이다.

[0004] 이러한 최적화 기반 접근법은 파라메트릭 모델의 초기 네트워크 파라미터에 따라 최적화 이후 파라메트릭 모델의 결과값이 크게 좌우된다는 문제가 제기되어 왔다. 이 문제를 해결하기 위하여 테스트 시간에 주어진 이미지에 효과적으로 최적화하기 위한 다양한 시도들이 있었다.

[0005] 삼차원 사람 복원을 위한 네트워크를 학습하기 위해서는 삼차원 참값(ground truth)이 필요하다.

[0006] 그러나 테스트 시간에서 접할 실제 세계의 다양한 배경, 다이내믹한 포즈와는 달리 실험실에서는 단조로운 배경의 제한된 포즈를 가진 이미지만 획득할 수 있으므로 도메인 격차가 발생한다.

- [0007] 테스트 시간 최적화는 이러한 도메인 격차로 인해 잘 예측되지 못한 샘플들에 대해 테스트 시간 때 주어지는 이차원 키폰트 참조값으로부터 투영 손실값(reprojection loss)를 사용해 삼차원 결과값을 이미지에 정렬시킨다.
- [0008] 그러나 최적화 이후의 결과 값이 초기 네트워크 파라미터에 지나치게 의존한다는 문제점이 주의해야 할 사항으로 지적된다. 이러한 문제를 해결하기 위해 더 좋은 초기 네트워크 파라미터의 필요성이 대두되었다.

선행기술문헌

특허문헌

- [0009] (특허문헌 0001) KR 등록특허 10-2261869

발명의 내용

해결하려는 과제

- [0010] 상기한 종래기술의 문제점을 해결하기 위해, 본 발명은 실제 세계의 데이터 분포를 따르는 이미지와 이미지에 대응되는 이차원 키폰트 참조값만을 사용한 투영 손실로 파라메트릭 모델을 메타 학습(meta-learning)하여, fine-tuning보다 더 좋은 초기 네트워크 파라미터를 테스트 시간에 제공할 수 있는 메타 초기화를 통한 삼차원 사람 메쉬 복원 방법 및 장치를 제안하고자 한다.

과제의 해결 수단

- [0011] 상기한 바와 같은 목적을 달성하기 위하여, 본 발명의 일 실시예에 따르면, 메타 학습을 통한 삼차원 사람 메쉬 복원 장치로서, 프로세서; 및 상기 프로세서에 연결되는 메모리를 포함하되, 상기 메모리는, 내부 루프의 파라메트릭 모델 기반 제1 신경망에 실제 사람 이미지를 입력하여 상기 실제 사람 이미지에 포함된 사람의 삼차원 메쉬를 추정하고, 상기 삼차원 메쉬에 관절 회귀 모델을 적용하여 상기 삼차원 메쉬에 상응하는 삼차원 키폰트를 추정하고, 카메라 파라미터를 통해 상기 삼차원 키폰트를 이차원 공간으로 투영하여 이차원 키폰트를 추정하고, 상기 추정된 이차원 키폰트와 이차원 참조값에 기초한 손실 함수를 통해 상기 제1 신경망의 네트워크 파라미터를 업데이트하고, 상기 제1 신경망의 제1 초기 네트워크 파라미터가 복사된 외부 루프의 삼차원 파라메트릭 모델 기반 제2 신경망의 제1 초기 네트워크 파라미터를 소정 횟수의 네트워크 파라미터 업데이트가 수행된 제1 신경망에서 출력하는 삼차원 메쉬를 이용하여 제2 초기 네트워크 파라미터로 업데이트하도록, 상기 프로세서에 의해 실행 가능한 프로그램 명령어들을 저장하는 삼차원 사람 메쉬 복원 장치가 제공된다.
- [0012] 상기 제1 신경망은 상기 실제 사람 이미지를 입력으로 하여 결정된 포즈 파라미터 및 형상 파라미터를 이용하여 상기 삼차원 메쉬를 추정할 수 있다.
- [0013] 상기 프로그램 명령어들은, 상기 제1 신경망에서 출력하는 삼차원 메쉬에 관절 회귀 모델을 적용하여 제2 삼차원 키폰트를 추정하고, 카메라 파라미터를 통해 상기 제2 삼차원 키폰트를 이차원 공간으로 투영하여 제2 이차원 키폰트를 추정하고, 상기 제2 이차원 키폰트와 이차원 참조값에 기초한 손실 함수를 통해 상기 제2 신경망의 제1 초기 네트워크 파라미터를 상기 제2 초기 네트워크 파라미터로 업데이트할 수 있다.
- [0014] 상기 프로그램 명령어들은, 경사 하강법을 이용하여 상기 제1 신경망의 학습 및 상기 제2 초기 네트워크 파라미터로 업데이트를 수행할 수 있다.
- [0015] 상기 프로그램 명령어들은, 상기 제2 초기 네트워크 파라미터로 업데이트 과정은 매 배치(batch)에 대해 반복 수행할 수 있다.
- [0016] 상기 프로그램 명령어들은, 모든 배치에 대해 상기 제2 초기 네트워크 파라미터로의 업데이트가 완료된 이후, 테스트 시간에 업데이트가 완료된 제2 초기 네트워크 파라미터를 테스트 신경망에 적용하여 실제 사람 이미지로부터 삼차원 사람 메쉬를 복원할 수 있다.
- [0017] 본 발명의 다른 측면에 따르면, 프로세서 및 메모리를 포함하는 장치에서 메타 학습을 통해 삼차원 사람 메쉬를 복원하는 방법으로서, 내부 루프의 파라메트릭 모델 기반 제1 신경망에 실제 사람 이미지를 입력하여 상기 실제 사람 이미지에 포함된 사람의 삼차원 메쉬를 추정하는 단계; 상기 삼차원 메쉬에 관절 회귀 모델을 적용하여 상기 삼차원 메쉬에 상응하는 삼차원 키폰트를 추정하는 단계; 카메라 파라미터를 통해 상기 삼차원 키폰트를

이차원 공간으로 투영하여 이차원 키포인트를 추정하는 단계; 상기 추정된 이차원 키포인트와 이차원 참값에 기초한 손실 함수를 통해 상기 제1 신경망의 네트워크 파라미터를 업데이트 하는 단계; 및 상기 제1 신경망의 제1 초기 네트워크 파라미터가 복사된 외부 루프의 삼차원 파라메트릭 모델 기반 제2 신경망의 제1 초기 네트워크 파라미터를 소정 횟수의 네트워크 파라미터 업데이트가 수행된 제1 신경망에서 출력하는 삼차원 메쉬를 이용하여 제2 초기 네트워크 파라미터로 업데이트 하는 단계를 포함하는 삼차원 사람 메쉬 복원 방법이 제공된다.

발명의 효과

[0018] 본 발명에 따르면, 기존에 존재하는 학습 네트워크의 구조와 손실함수를 그대로 유지하면서도 추가적인 삼차원 참값 없이 fine-tuning보다 더 나은 성능을 이룰 수 있는 장점이 있다.

[0019] 또한, 테스트 시간에 더 나은 파라미터로 모델을 초기화할 수 있으며, 테스트 시간 최적화 이후 파라메트릭 모델의 예측값도 기존 모델에 비해 더 정확하다는 장점이 있으며, 이를 통해 삼차원 사람 복원의 고질적인 문제인 도메인 격차 문제를 효과적으로 해결할 수 있다.

도면의 간단한 설명

[0020] 도 1은 본 발명의 바람직한 일 실시예에 따른 메타 학습을 통해 삼차원 사람 메쉬 복원 프레임워크를 도시한 도면이다.

도 2는 본 실시예에 따른 삼차원 사람 복원을 위한 메타 학습 과정의 알고리즘을 나타낸 도면이다.

도 3은 본 실시예에 따른 메타 학습을 통한 삼차원 사람 메쉬 복원 장치의 구성을 도시한 도면이다.

발명을 실시하기 위한 구체적인 내용

[0021] 본 발명은 다양한 변형을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세하게 설명하고자 한다.

[0022] 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.

[0024] 본 실시예는 파라메트릭 모델의 초기 네트워크 파라미터를 업데이트하여, 즉 메타 초기화를 통해 테스트 시간 최적화 이후 파라메트릭 모델의 예측값의 정확도를 높이는 것이다.

[0025] 여기서, 초기 네트워크 파라미터는 파라메트릭 모델에 대한 모델 네트워크 또는 가중치로 표현될 수 있다.

[0026] 도 1은 본 발명의 바람직한 일 실시예에 따른 메타 학습을 통해 삼차원 사람 메쉬 복원 프레임워크를 도시한 도면이다.

[0027] 도 1을 참조하면, 본 실시예에 따른 프레임워크는 내부 루프(Inner Loop) 및 외부 루프(Outer Loop)를 포함한다.

[0028] 내부 루프 및 외부 루프 각각에는 파라메트릭 모델 기반 신경망(100,110)이 제공된다.

[0029] 여기서, 파라메트릭 모델은 SMPL(Skinned Multi-Person Linear model)일 수 있다.

[0030] 메타 학습 과정에서, 내부 루프의 제1 신경망(100)에는 j번째 배치에 속하는 실제 사람 이미지(I)가 입력되고, 제1 신경망(100)은 실제 사람 이미지에 포함된 사람의 삼차원 메쉬를 추정한다.

[0031] 하나의 배치에는 복수의 실제 사람 이미지가 포함될 수 있고, 이하의 과정은 하나의 배치에 속하는 복수의 실제 사람 이미지가 순차적으로 입력되면서 수행될 수 있다.

[0032] 제1 신경망(100)은 신체 관절의 회전 각도를 나타내는 포즈 파라미터와 신체 형상 비율을 나타내는 형상 파라미터를 추정하고, 이를 통해 삼차원 메쉬 정점들(vertices)을 추정할 수 있다.

[0033] 도 1의 내부 루프에서 초기 메쉬(initial mesh)는 메타 학습 초기에 제1 신경망(100)을 통해 추정되는 삼차원 메쉬이다.

[0034] 내부 루프에서 삼차원 메쉬에 관절 회귀 모델을 적용하여 삼차원 메쉬에 상응하는 삼차원 키포인트를 추정한다.

[0035] 또한, 내부 루프는 카메라 파라미터(투영 함수, π)를 통해 삼차원 키포인트를 이차원 공간으로 투영하여 이차원 키포인트를 추정한다.

[0036] 도 1의 내부 루프에서 초기 키포인트(initial keypoint)는 메타 학습 초기에 상기와 같이 추정된 이차원 키포인트이다.

[0037] 이후, 추정된 이차원 키포인트와 이차원 참값(ground truth)에 기초한 손실 함수를 통해 제1 신경망(100)의 학습이 수행된다.

[0038] 여기서, 제1 신경망(100)의 학습을 위한 손실 함수 $L(W)$ 는 다음과 정의될 수 있다.

수학식 1

$$L(W) = \left\| \pi(M(f(I; W))) - \hat{j} \right\|_2 + \lambda \|\beta\|_2,$$

[0039]

[0040] 여기서, W 는 네트워크 파라미터, π 는 투영 함수, $M(f(I; W))$ 는 관절 회귀 모델, I 는 실제 사람 이미지, $\hat{j}$ 는 참값, λ 는 가중치, β 는 메타 학습 스텝 사이즈이다.

[0041] 두 번째 항은 형상 파라미터가 합리적인 분포가 되도록 정규화된다.

[0042] 상기한 손실 함수를 최소화함으로써 테스트 이미지에 대한 최적 네트워크 파라미터(W^*)를 추정할 수 있으며, 최적 네트워크 파라미터를 다음과 같이 파라메트릭 모델의 포즈 파라미터와 형상 파라미터(Θ^*)를 추정하는데 사용할 수 있다.

수학식 2

$$\Theta^* = f(I; W^*) \quad \text{where} \quad W^* = \operatorname{argmin}_W L(W).$$

[0043]

[0044] 본 실시예에 따르면, 손실 함수 최소화를 위해 제1 신경망(100)에 대해 사전 훈련된 네트워크 파라미터를 적용하고, 경사 하강법(gradient descent)을 이용한다.

[0045] 본 실시예에 따르면, 경사 하강법을 통해 내부 루프의 반복 과정에서 예측된 이차원 키포인트와 참값과의 손실에 대해 그동안 쌓였던 손실의 기울기 값을 역전파하여 네트워크 파라미터를 업데이트한다.

[0046] 이하에서는 경사 하강법을 상세하게 설명한다.

[0047] W_0 와 W_m 은 각각 초기 시점과 m번째 업데이트 이후의 네트워크 파라미터를 나타낸다. 이때, 경사 하강법은 다음 규칙을 통해 진행된다.

수학식 3

$$\mathbf{W}_{m+1} = \mathbf{W}_m - \alpha \nabla L(\mathbf{W}_m),$$

여기서, α 는 학습률이다.

테스트 시간에서, m번째 업데이트 이후, 초기 네트워크 파라미터 $\mathbf{W}_0$ 에 따라 $\mathbf{W}_m$ 이 달라질 수 있다. 초기 네트워크 파라미터 및 주어진 입력에 대한 파라메트릭 모델의 종속성은 $\mathbf{W}_m(\mathbf{W}_0, \mathbf{I})$ 로 표현할 수 있다.

테스트 시간 동안 m번째 업데이트 이후 파라메트릭 모델 결과 $f(\mathbf{W}_m)$ 의 품질을 향상시키기 위해 다음과 같이 설명할 수 있는 더 나은 초기 네트워크 파라미터 $\mathbf{W}_0^*$ 를 찾는 것이 바람직할 수 있다.

이를 위해, 도 1에 도시된 바와 같이, 내부 루프에서 m번째 네트워크 파라미터 업데이트가 완료된 이후, 외부 루프에서 초기 파라미터 업데이트 과정이 수행될 수 있다.

보다 상세하게, 외부 루프의 제2 신경망(110)에는 제1 신경망(100)의 초기 네트워크 파라미터(제1 초기 네트워크 파라미터)가 복사된다.

제2 신경망(110)의 제1 초기 네트워크 파라미터($\mathbf{W}_0$)는 상기 m번째 업데이트가 수행된 제1 신경망에서 출력하는 삼차원 메쉬를 이용하여 제2 초기 네트워크 파라미터($\mathbf{W}_0^*$)로 업데이트된다.

수학식 4

$$\mathbf{W}_0^* = \operatorname{argmin}_{\mathbf{W}_0} \sum_l L(\mathbf{W}_m(\mathbf{W}_0, \mathbf{I}_l)),$$

여기서 $\mathbf{I}_l$ 은 l 번째 훈련 이미지이다.

본 실시예에 따른 제2 초기 네트워크 파라미터로의 업데이트는 매 배치마다 수행되며, j번째 배치의 외부 루프는 다음과 같이 초기 네트워크 파라미터를 학습한다.

수학식 5

$$\mathbf{W}_0^{j+1} = \mathbf{W}_0^j - \beta \left(\nabla_{\mathbf{W}} \sum_l L(\mathbf{W}_m(\mathbf{W}, \mathbf{I}_l)) \right)_{\mathbf{W}=\mathbf{W}_0^j}$$

여기서, β 는 메타 학습 단계 크기이다.

$$W_m(W, I_l)$$

[0060] 상기한 수식은 내부 루프 최적화 동안 얻은 네트워크 파라미터의 손실에 경사 하강법을 적용한 것을 나타내며, 학습된 초기 네트워크 파라미터는 삼차원 사람 복원을 위한 테스트 시간 최적화를 위한 테스트 신경망에 삼차원 사람 메쉬를 복원하는데 더 좋은 성능을 나타내도록 할 수 있다.

[0061] 도 2는 본 실시예에 따른 삼차원 사람 복원을 위한 메타 학습 과정의 알고리즘을 나타낸 도면이다.

[0062] 아래의 표 1은 fine-tuning결과와 본 실시예에 따른 메타 학습 결과를 비교한 것이다.

표 1

	Fine-tuning	Meta-L1	Meta-L2	Meta-L3	Meta-L4
Outer 0	58.40	58.16	58.08	58.04	58.02
Outer 1	58.38	58.06	57.98	57.94	58.00
Outer 2	58.35	58.11	57.90	57.86	57.97
Outer 3	58.31	58.08	57.88	57.91	57.94

[0063]

[0064] 사용된 메트릭값은 PA-MPJPE로, 삼차원 키포인트들의 참값과 예측값이 Procrustes method를 통해 정렬된 후의 각 키포인트 별 유클리디안 거리의 평균이 계산된 결과로, 수치가 작을수록 더 나은 결과이다. Meta-L1은 내부 루프가 1인 메타 학습 모델을 의미한다. Outer 옆에 있는 숫자는 테스트 시간 때 최적화 루프의 수를 의미한다. Outer0, 즉 테스트 시간에서의 최적화 이전 네트워크 초기값이 내부 루프를 더 많이 돌린 메타 학습으로 학습되었을 때 최적화 단계마다 더 나은 결과를 보이는 것을 확인할 수 있다.

[0065] 도 3은 본 실시예에 따른 메타 학습을 통한 삼차원 사람 메쉬 복원 장치의 구성을 도시한 도면이다.

[0066] 본 실시예에 따른 삼차원 사람 메쉬 복원 장치는 프로세서(300) 및 메모리(302)를 포함할 수 있다.

[0067] 프로세서(300)는 컴퓨터 프로그램을 실행할 수 있는 CPU(central processing unit)나 그 밖에 가상 머신 등을 포함할 수 있다.

[0068] 메모리(302)는 고정식 하드 드라이브나 착탈식 저장 장치와 같은 불휘발성 저장 장치를 포함할 수 있다. 착탈식 저장 장치는 콤팩트 플래시 유닛, USB 메모리 스틱 등을 포함할 수 있다. 메모리(302)는 각종 랜덤 액세스 메모리와 같은 휘발성 메모리도 포함할 수 있다.

[0069] 메모리(302)에는 메타 학습을 통해 삼차원 사람 메쉬를 복원하기 위한 프로그램 명령어들이 저장된다.

[0070] 본 실시예에 따른 프로그램 명령어들은, 내부 루프의 파라메트릭 모델 기반 제1 신경망에 실제 사람 이미지를 입력하여 상기 실제 사람 이미지에 포함된 사람의 삼차원 메쉬를 추정하고, 상기 삼차원 메쉬에 관절 회귀 모델을 적용하여 상기 삼차원 메쉬에 상응하는 삼차원 키포인트를 추정하고, 카메라 파라미터를 통해 상기 삼차원 키포인트를 이차원 공간으로 투영하여 이차원 키포인트를 추정하고, 상기 추정된 이차원 키포인트와 이차원 참값에 기초한 손실 함수를 통해 상기 제1 신경망의 네트워크 파라미터를 업데이트하고, 상기 제1 신경망의 제1 초기 네트워크 파라미터가 복사된 외부 루프의 삼차원 파라메트릭 모델 기반 제2 신경망의 제1 초기 네트워크 파라미터를 소정 횟수의 네트워크 파라미터 업데이트가 수행된 제1 신경망에서 출력하는 삼차원 메쉬를 이용하여 제2 초기 네트워크 파라미터로 업데이트할 수 있다.

[0071] 또한, 프로그램 명령어들은, 상기 제1 신경망에서 출력하는 삼차원 메쉬에 관절 회귀 모델을 적용하여 제2 삼차원 키포인트를 추정하고, 카메라 파라미터를 통해 상기 제2 삼차원 키포인트를 이차원 공간으로 투영하여 제2 이차원 키포인트를 추정하고, 상기 제2 이차원 키포인트와 이차원 참값에 기초한 손실 함수를 통해 상기 제2 신경망의 제1 초기 네트워크 파라미터를 상기 제2 초기 네트워크 파라미터로 업데이트한다.

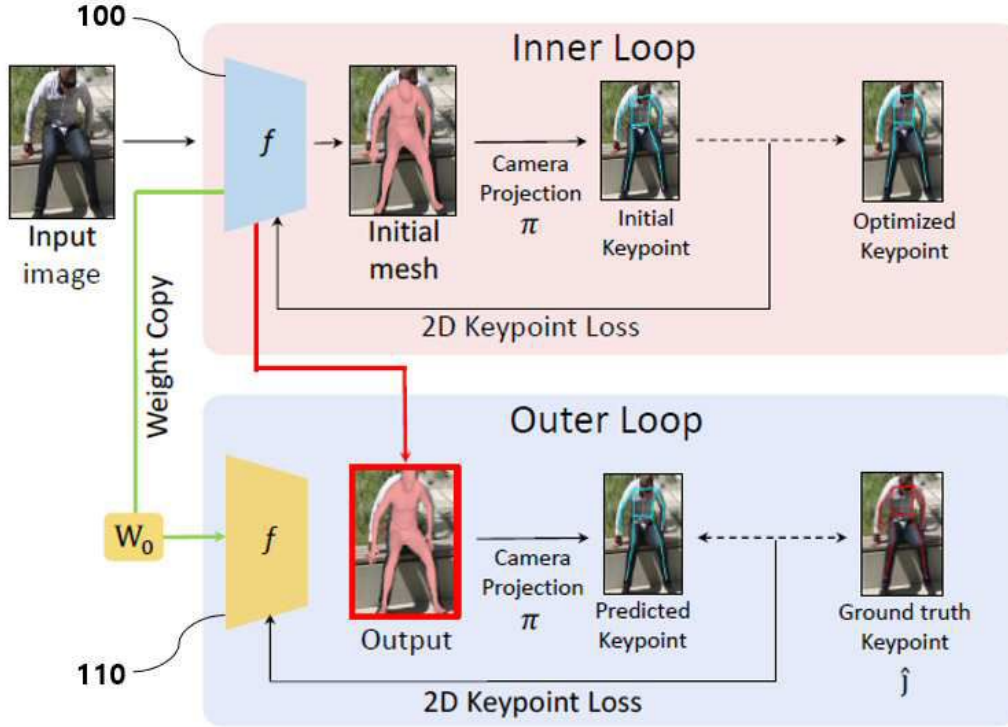
[0072] 이후, 프로그램 명령어들은, 모든 배치에 대해 상기 제2 초기 네트워크 파라미터로의 업데이트가 완료된 이후, 테스트 시간에 업데이트가 완료된 제2 초기 네트워크 파라미터를 테스트 신경망에 적용하여 실제 사람 이미지로부터 삼차원 사람 메쉬를 복원한다.

[0073] 상기한 본 발명의 실시예는 예시의 목적을 위해 개시된 것이고, 본 발명에 대한 통상의 지식을 가지는 당업자라면 본 발명의 사상과 범위 안에서 다양한 수정, 변경, 부가가 가능할 것이며, 이러한 수정, 변경 및 부가는 하

기의 특허청구범위에 속하는 것으로 보아야 할 것이다.

도면

도면1



도면2

Algorithm 1 Meta-Learning for 3D Human Recovery

$p(\mathcal{T})$: uniform distribution over dataset

α, β : step size hyperparameters

$\mathbf{W}^*$: pre-trained model weights

- 1: Initialize model parameters $\mathbf{W}$ as $\mathbf{W}^*$
- 2: **while** not converged **do**
- 3: Sample batch of images $\mathcal{T}_l \sim p(\mathcal{T})$
- 4: **for each** $\mathcal{T}_l$ **do**
- 5: Sample images $\mathbf{I}_l$ from $\mathcal{T}_l$
- 6: Compute $L(\mathbf{W}_m(\mathbf{W}, \mathbf{I}_l))$ using $L(\mathbf{W})$ in eq 1.
- 7: Evaluate $\nabla_{\mathbf{W}} L(\mathbf{W}_m(\mathbf{W}, \mathbf{I}_l))$
- 8: Update $\mathbf{W}_l \leftarrow \mathbf{W} - \alpha \nabla_{\mathbf{W}} L(\mathbf{W}_m(\mathbf{W}, \mathbf{I}_l))$
- 9: **end for**
- 10: Update $\mathbf{W} \leftarrow \mathbf{W} - \beta \nabla_{\mathbf{W}} \sum_l L(\mathbf{W}_m(\mathbf{W}_l, \mathbf{I}_l))$
- 11: **end while**

도면3

