



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2024년12월11일  
(11) 등록번호 10-2741821  
(24) 등록일자 2024년12월09일

(51) 국제특허분류(Int. Cl.)  
G06T 13/40 (2011.01) G06N 20/00 (2019.01)  
G06T 17/20 (2006.01) G06T 3/00 (2024.01)  
G06T 7/73 (2017.01)  
(52) CPC특허분류  
G06T 13/40 (2013.01)  
G06N 20/00 (2021.08)  
(21) 출원번호 10-2022-0090998  
(22) 출원일자 2022년07월22일  
심사청구일자 2022년07월22일  
(65) 공개번호 10-2024-0013454  
(43) 공개일자 2024년01월30일  
(56) 선행기술조사문헌  
US20210335039 A1\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
포항공과대학교 산학협력단  
경상북도 포항시 남구 청암로 77 (지곡동)  
(72) 발명자  
이승용  
경상북도 포항시 남구 청암로 77  
정유철  
경상북도 포항시 남구 청암로 77  
(74) 대리인  
인비전 특허법인

전체 청구항 수 : 총 11 항

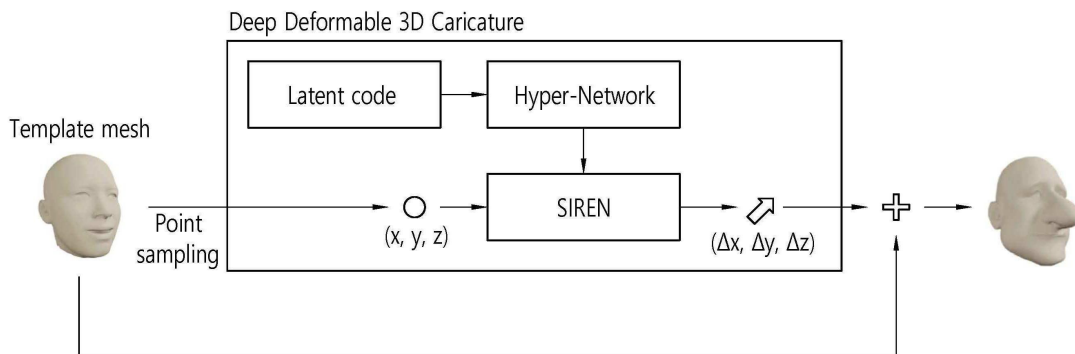
심사관 : 전세운

(54) 발명의 명칭 영상처리 장치 및 영상처리 장치를 위한 3차원 형상 잠재공간의 기계 학습 방법

(57) 요약

본 발명에 따른 영상처리 장치는 3차원 형상의 잠재공간(Latent space)을 학습하여 템플릿 메쉬의 변형을 수행하는 영상처리 장치에 있어서, 상기 템플릿 메쉬가 저장되는 저장부 및 상기 템플릿 메쉬를 처리하는 처리부를 포함하고, 상기 처리부는 상기 템플릿 메쉬의 3차원 좌표를 포인트 샘플링하고, 상기 포인트 샘플링된 상기 3차원 좌표를 다층 퍼셉트론(Multi Layer Perceptron, MLP)에 적용하여 상기 3차원 좌표 상에 추가되는 3차원 변형값을 출력하며 상기 잠재공간의 원소인 잠재코드(Latent code)를 하이퍼 네트워크에 적용하여 상기 다층 퍼셉트론에 대한 매개변수를 출력한다.

대표도



(52) CPC특허분류

*G06T 17/20* (2013.01)

*G06T 3/18* (2024.01)

*G06T 7/73* (2017.01)

*G06T 2207/20081* (2013.01)

(72) 발명자

**장원종**

경상북도 포항시 남구 청암로 77

**김승진**

경상북도 포항시 남구 청암로 77

이 발명을 지원한 국가연구개발사업

과제고유번호 1711153053

과제번호 2015-0-00174-008

부처명 과학기술정보통신부

과제관리(전문)기관명 한국연구재단

연구사업명 정보통신기획평가원

연구과제명 (SW 스타랩) 빅 비주얼 데이터 기반의 고품질 사진 메이크업 SW 개발

기 여 율 1/1

과제수행기관명 포항공과대학교 산학협력단

연구기간 2022.01.01 ~ 2022.12.31

공지예외적용 : 있음

---

## 명세서

### 청구범위

#### 청구항 1

3차원 형상의 잠재공간(Latent space)을 학습하여 템플릿 메쉬의 변형을 수행하는 영상처리 장치에 있어서,  
 상기 템플릿 메쉬가 저장되는 저장부; 및  
 상기 템플릿 메쉬를 처리하는 처리부를 포함하고,  
 상기 처리부는  
 상기 템플릿 메쉬의 3차원 좌표를 포인트 샘플링하고,  
 상기 포인트 샘플링된 상기 3차원 좌표를 다층 퍼셉트론(Multi Layer Perceptron, MLP)에 적용하여 상기 3차원 좌표 상에 추가되는 3차원 변형값을 출력하며 상기 잠재공간의 원소인 잠재코드(Latent code)를 하이퍼 네트워크에 적용하여 상기 다층 퍼셉트론에 대한 매개변수를 출력하며,  
 상기 처리부에서는 다양한 3차원 형상을 하나의 잠재공간 상에 모델링하기 위하여 상기 잠재코드를 3차원 형상으로 변환하고,  
 상기 하이퍼네트워크는 상기 잠재코드를 3차원 형상으로 변환하는 다층 퍼셉트론을 포함하고,  
 상기 하이퍼네트워크의 학습에서는 Auto-decoder 프레임워크가 적용되어 상기 잠재코드와 상기 하이퍼네트워크를 동시에 학습하는 것을 특징으로 하는 영상처리 장치.

#### 청구항 2

제1 항에 있어서,  
 상기 하이퍼 네트워크는  
 상기 잠재코드를 입력받아 상기 매개변수를 출력하는 다층 퍼셉트론을 포함하고,  
 상기 처리부는  
 상기 매개변수를 기반으로 상기 템플릿 메쉬의 3차원 형상 변형이 발생되도록 하는 것을 특징으로 하는 영상처리 장치.

#### 청구항 3

제1 항에 있어서,  
 상기 다층 퍼셉트론은  
 상기 템플릿 메쉬의 포인트를 변위 벡터로 매핑하는 SIREN 다층 퍼셉트론을 포함하는 것을 특징으로 하는 영상처리 장치.

#### 청구항 4

제1 항에 있어서,  
 상기 처리부는  
 상기 템플릿 메쉬의 정점(Vertices) 및 상기 템플릿 메쉬의 삼각지역(Triangle)을 기반으로 상기 3차원 좌표의 포인트 샘플링을 수행하는 것을 특징으로 하는 영상처리 장치.

#### 청구항 5

제1 항에 있어서,

상기 다층 퍼셉트론은

고정된 상기 템플릿 메쉬 상의 3차원 좌표를 변형된 3차원 좌표로 매핑하여 변형된 3차원 형상을 표현하는 것을 특징으로 하는 영상처리 장치.

#### 청구항 6

삭제

#### 청구항 7

삭제

#### 청구항 8

제1 항에 있어서,

상기 하이퍼네트워크의 학습에서 손실함수로는

상기 잠재코드의 Squared L2 norm과 Ground-truth 형상과의 평균 제곱 오차(Mean Square Error, MSE)의 합이 사용되는 것을 특징으로 하는 영상처리 장치.

#### 청구항 9

제8 항에 있어서,

상기 평균 제곱 오차는

Ground-truth 메쉬의 표면에서 균일하게 추출한 3차원 좌표와 Ground-truth 메쉬 정점의 3차원 좌표의 합집합에 대하여 계산되는 것을 특징으로 하는 영상처리 장치.

#### 청구항 10

제1 항에 있어서,

상기 처리부는

상기 잠재공간을 기반으로 편집을 수행하여 2차원 캐릭처치의 랜드마크로부터 3차원 캐릭처치를 생성하는 것을 특징으로 하는 영상처리 장치.

#### 청구항 11

제1 항에 있어서,

상기 처리부는

상기 잠재공간을 기반으로 의미론적 편집(Semantic editing) 및 포인트-핸들-기반 편집(Point-handle-based editing) 중 적어도 어느 하나를 수행하는 것을 특징으로 하는 영상처리 장치.

#### 청구항 12

제1 항에 있어서,

상기 처리부는

상기 잠재공간을 기반으로 일반 3차원 얼굴 이미지로부터 자동으로 3차원 캐리커처를 생성하는 것을 특징으로 하는 영상처리 장치.

### 청구항 13

3차원 형상의 잠재공간을 학습하여 템플릿 메쉬의 변형을 수행하는 영상처리 장치의 기계 학습 방법에 있어서, 상기 템플릿 메쉬의 3차원 좌표를 포인트 샘플링하는 단계; 및

상기 포인트 샘플링된 상기 3차원 좌표를 다층 퍼셉트론(Multi Layer Perceptron, MLP)에 적용하여 상기 3차원 좌표 상에 추가되는 3차원 변형값을 출력하며 상기 잠재공간의 원소인 잠재코드(Latent code)를 하이퍼 네트워크에 적용하여 상기 다층 퍼셉트론에 대한 매개변수를 출력하고,

상기 잠재코드는 다양한 3차원 형상을 하나의 잠재공간 상에 모델링하기 위하여 3차원 형상으로 변환되고,

상기 하이퍼네트워크는 상기 잠재코드를 3차원 형상으로 변환하는 다층 퍼셉트론을 포함하고,

상기 하이퍼네트워크의 학습에서는 Auto-decoder 프레임워크가 적용되어 상기 잠재코드와 상기 하이퍼네트워크를 동시에 학습하는 것을 특징으로 하는 영상처리 장치의 기계 학습 방법.

### 발명의 설명

#### 기술 분야

[0001] 본 발명은 영상처리 장치 및 영상처리 장치를 위한 3차원 형상 잠재공간의 기계 학습 방법에 관한 것으로 3차원 캐리커처를 처리하는 영상처리 장치 및 영상처리 장치를 위한 3차원 형상 잠재공간의 기계 학습 방법에 관한 것이다.

#### 배경 기술

[0002] 일반적으로 3차원 캐리커처는 사람의 얼굴을 과장된 3D 이미지로 묘사한 것으로, 모델링을 통해 원본 이미지를 3차원 캐리커처로 변환하는 기술은 그래픽 및 영상처리 분야 등 다양한 산업 분야에 적용되고 있다.

[0003] 종래의 3차원 캐리커처 생성에 대한 기술은 대한민국 공개특허공보 제10-2004-0032452호(3차원 캐리커처 생성 장치 및 방법, 2004.04.17.)에 의해 공개되어 있다. 상기 공개발명은 모델링을 통해 이미지로부터 특징점을 추출하고, 추출된 특징점을 이용하여 3차원 캐리커처를 생성하는 것을 특징으로 한다.

[0004] 다만, 특정 도메인에 대한 3차원 형상의 잠재공간(Latent Space) 분포를 나타내는 생성 모델은 3차원 복원 및 3차원 형상 조작 문제에 필수적이다. 이에, 종래 기술은 얼굴 간 분산이 작은 일반 3차원 얼굴 도메인에서는 작동하지만 과장된 표현으로 얼굴 간 분산이 큰 3차원 캐리커처 도메인에서는 작동이 어려운 문제점이 있었다. 따라서, 분산이 큰 3차원 얼굴 도메인에서 효과적으로 잠재공간을 학습하기 위한 기법이 요구되고 있다.

### 선행기술문헌

#### 특허문헌

[0005] (특허문헌 0001) 대한민국 공개특허공보 제10-2004-0032452호(3차원 캐리커처 생성 장치 및 방법, 2004.04.17.)

### 발명의 내용

#### 해결하려는 과제

[0006] 본 발명의 목적은 분산이 큰 3차원 얼굴 도메인에서 잠재공간을 학습할 수 있는 영상처리 장치 및 영상처리 장

치를 위한 3차원 형상 잠재공간의 기계 학습 방법을 제공하기 위한 것이다.

**과제의 해결 수단**

- [0007] 본 발명에 따른 영상처리 장치는 3차원 형상의 잠재공간(Latent space)을 학습하여 템플릿 메쉬의 변형을 수행하는 영상처리 장치에 있어서, 상기 템플릿 메쉬가 저장되는 저장부 및 상기 템플릿 메쉬를 처리하는 처리부를 포함하고, 상기 처리부는 상기 템플릿 메쉬의 3차원 좌표를 포인트 샘플링하고, 상기 포인트 샘플링된 상기 3차원 좌표를 다층 퍼셉트론(Multi Layer Perceptron, MLP)에 적용하여 상기 3차원 좌표 상에 추가되는 3차원 변형값을 출력하며 상기 잠재공간의 원소인 잠재코드(Latent code)를 하이퍼 네트워크에 적용하여 상기 다층 퍼셉트론에 대한 매개변수를 출력한다.
- [0008] 상기 하이퍼 네트워크는 상기 잠재코드를 입력받아 상기 매개변수를 출력하는 다층 퍼셉트론을 포함하고, 상기 처리부는 상기 매개변수를 기반으로 상기 템플릿 메쉬의 3차원 형상 변형이 발생되도록 할 수 있다.
- [0009] 상기 다층 퍼셉트론은 상기 템플릿 메쉬의 포인트를 변위 벡터로 매핑하는 SIREN 다층 퍼셉트론을 포함할 수 있다.
- [0010] 상기 처리부는 상기 템플릿 메쉬의 정점(Vertex) 및 상기 템플릿 메쉬의 삼각지역(Triangle)을 기반으로 상기 3차원 좌표의 포인트 샘플링을 수행할 수 있다.
- [0011] 상기 다층 퍼셉트론은 고정된 상기 템플릿 메쉬 상의 3차원 좌표를 변형된 3차원 좌표로 매핑하여 변형된 3차원 형상을 표현할 수 있다.
- [0012] 상기 처리부는 다양한 3차원 형상을 하나의 잠재공간 상에 모델링하기 위하여 상기 잠재코드를 3차원 형상으로 변환할 수 있다.
- [0013] 상기 하이퍼네트워크는 상기 잠재코드를 3차원 형상으로 변환하는 다층 퍼셉트론을 포함하고, 상기 하이퍼네트워크의 학습에서는 Auto-decoder 프레임워크가 적용되어 상기 잠재코드와 상기 하이퍼네트워크를 동시에 학습할 수 있다.
- [0014] 상기 하이퍼네트워크의 학습에서 손실함수로는 상기 잠재코드의 Squared L2 norm과 Ground-truth 형상과의 평균 제곱 오차(Mean Square Error, MSE)의 합이 사용될 수 있다.
- [0015] 상기 평균 제곱 오차는 상기 Ground-truth 메쉬의 표면에서 균일하게 추출한 3차원 좌표와 Ground-truth 메쉬 정점의 3차원 좌표의 합집합에 대하여 계산될 수 있다.
- [0016] 상기 처리부는 상기 잠재공간을 기반으로 편집을 수행하여 2차원 캐릭처의 랜드마크로부터 3차원 캐릭처를 생성할 수 있다.
- [0017] 상기 처리부는 상기 잠재공간을 기반으로 의미론적 편집(Semantic editing) 및 포인트-핸들-기반 편집(Point-handle-based editing) 중 적어도 어느 하나를 수행할 수 있다.
- [0018] 상기 처리부는 상기 잠재공간을 기반으로 일반 3차원 얼굴 이미지로부터 자동으로 3차원 캐릭처를 생성할 수 있다.
- [0019] 한편, 본 발명에 따른 영상처리 장치의 기계 학습 방법은 3차원 형상의 잠재공간을 학습하여 템플릿 메쉬의 변형을 수행하는 영상처리 장치의 기계 학습 방법에 있어서, 상기 템플릿 메쉬의 3차원 좌표를 포인트 샘플링하는 단계 및 상기 포인트 샘플링된 상기 3차원 좌표를 다층 퍼셉트론(Multi Layer Perceptron, MLP)에 적용하여 상기 3차원 좌표 상에 추가되는 3차원 변형값을 출력하며 상기 잠재공간의 원소인 잠재코드(Latent code)를 하이퍼 네트워크에 적용하여 상기 다층 퍼셉트론에 대한 매개변수를 출력한다.

**발명의 효과**

- [0020] 본 발명에 따른 영상처리 장치 및 영상처리 장치를 위한 3차원 형상 잠재공간의 기계 학습 방법은 3차원 형상을 템플릿 변형 함수 표현방식으로 구현하여 잠재공간의 학습 중 형상의 얼굴다움을 보장하는 매개변수화를 제공하며, 매개변수를 잠재 코드(Latent code)로부터 생성하는 다층 퍼셉트론 기반 하이퍼 네트워크를 통해 효율적인 학습을 수행하는 효과가 있다.
- [0021] 이상과 같은 본 발명의 기술적 효과는 이상에서 언급한 효과로 제한되지 않으며, 언급되지 않은 또 다른 기술적 효과들은 아래의 기재로부터 당업자에게 명확하게 이해될 수 있을 것이다.

**도면의 간단한 설명**

- [0022] 도 1은 본 실시예에 따른 3차원 형상 잠재공간의 기계 학습 방법을 나타낸 흐름도이고, 도 2는 본 실시예에 따른 3차원 형상 잠재공간의 기계 학습 방법을 나타낸 개념도이고, 도 3은 본 실시예에 따른 영상처리 장치를 개략적으로 나타낸 구성도이고, 도 4는 본 실시예에 따른 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수 표현방식과 메쉬 정점의 3차원 좌표배열 방식을 적용하였을 때를 비교한 도면이고, 도 5는 본 실시예에 따른 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수 표현방식과, DeepSDF 방식 및 DIF-NET 방식을 적용하였을 때를 비교한 도면이고, 도 6은 본 실시예에 따른 영상처리 장치에서 2차원 캐리커처의 랜드마크를 기반으로 3차원 캐리커처를 재구성한 도면이고, 도 7은 본 실시예에 따른 영상처리 장치에서 의미론적 편집을 수행하는 것을 나타낸 도면이고, 도 8은 본 실시예에 따른 영상처리 장치에서 포인트-헨들-기반 편집을 수행하는 것을 나타낸 도면이고, 도 9는 본 실시예에 따른 영상처리 장치에서 자동으로 3차원 캐리커처를 생성하는 것을 나타낸 도면이다.

**발명을 실시하기 위한 구체적인 내용**

- [0023] 이하 첨부된 도면을 참조하여 본 발명의 실시예를 상세히 설명한다. 그러나 본 실시예는 이하에서 개시되는 실시예에 한정되는 것이 아니라 서로 다양한 형태로 구현될 수 있으며, 단지 본 실시예는 본 발명의 개시가 완전하도록 하며, 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위해 제공되는 것이다. 도면에서의 요소의 형상 등은 보다 명확한 설명을 위하여 과장되게 표현된 부분이 있을 수 있으며, 도면 상에서 동일 부호로 표시된 요소는 동일 요소를 의미한다.
- [0024] 일반적으로, 특정 도메인에 대한 3차원 형상의 잠재공간(Latent space) 분포를 나타내는 생성 모델은 3차원 복원 및 3차원 형상 조작 문제에 필수적이다. 영상 기반 3차원 얼굴복원 및 얼굴 조작 문제를 해결하는 종래 기술은 얼굴 간 분산이 작은 일반 3차원 얼굴 도메인에서는 작동하지만, 과장된 표현이 많아 얼굴 간 분산이 큰 3차원 카툰(Cartoon) 및 3차원 캐리커처 도메인에서는 작동하지 않는 문제점이 있었다. 이에, 본 발명에 따른 3차원 형상 잠재공간의 기계 학습 방법에서는 분산이 큰 3차원 얼굴 도메인에서 효과적으로 잠재공간을 학습하기 위한 기법을 제시한다.
- [0025] 도 1은 본 실시예에 따른 3차원 형상 잠재공간의 기계 학습 방법을 나타낸 흐름도이고, 도 2는 본 실시예에 따른 3차원 형상 잠재공간의 기계 학습 방법을 나타낸 개념도이다. 그리고 도 3은 본 실시예에 따른 영상처리 장치를 개략적으로 나타낸 구성도이다.
- [0026] 도 1 내지 도 3에 도시된 바와 같이, 본 실시예에 따른 3차원 형상 잠재공간의 기계 학습 방법(이하, 기계 학습 방법이라 칭한다.)은 분산이 큰 3차원 얼굴 도메인에서 효과적으로 잠재공간을 학습할 수 있다. 이러한 기계 학습 방법은 3차원 캐리커처 생성기술 등에 적용될 수 있으나, 적용 대상은 한정하지 않는다.
- [0027] 일반적으로 2차원 캐리커처는 숙련된 아티스트에 의해 생성 가능하며, 3차원 캐리커처는 제작 시 더 많은 전문 지식이 요구된다. 이에, 기계 학습 방법은 3차원 캐리커처의 변형을 모델링하여 과장된 3차원 얼굴을 처리하는데 유리할 수 있다.
- [0028] 이러한 기계 학습 방법은 컴퓨터 시스템 등과 같은 영상처리 장치(100)에서 수행될 수 있으며, 영상처리 장치(100)는 저장부(110) 및 처리부(120)를 포함할 수 있다.
- [0029] 저장부(110)에는 기계 학습을 위해 외부로부터 제공되는 데이터 세트가 저장될 수 있으며, 3차원 형상 잠재공간의 기계 학습을 위한 프레임워크(Framework)가 저장될 수 있다. 그리고 처리부(120)는 데이터 세트를 기반으로 학습을 진행하여 3차원 형상의 변형이 이루어지도록 할 수 있다.
- [0030] 일례로, 데이터 세트는 3차원 아티스트에 의해 생성된 '3D CaricShop'의 데이터 세트일 수 있다.
- [0031] 일례로, 기계 학습 방법에서는 1,409개의 3차원 캐리커처 메쉬를 획득하였고, 목 부분을 제외하고 눈과 입의 구멍을 채운 메쉬를 사용하였다. 그리고 1,268개의 메쉬를 훈련 데이터 세트로 적용하고, 14개의 메쉬를 검증 세

트로, 127개의 메쉬를 테스트 세트로 활용하였다. 그리고 데이터 세트와 함께 3D CaricShop의 정점 연결을 갖는 'FaceWarehouse'의 평균 페이스(Mean face)를 획득하여, 평균 페이스를 템플릿 메쉬로 사용하였다.

- [0032] 이에, 영상처리 장치(100)는 3차원 캐리커처의 데이터 세트에서 변형 가능한 모델을 학습하여 3차원 캐리커처의 제어 가능한 3차원 캐리커처용 툴킷(Toolkit)을 제공할 수 있다.
- [0033] 이를 위해 기계 학습 방법에서는 일대일 대응이 존재하는 3차원 메쉬 데이터 세트의 잠재공간과, 잠재공간의 원소인 잠재코드를 3차원 형상으로 변환하는 매핑을 설계한다. 이러한 기계 학습 방법에서는 제1 설계 단계(S100) 및 제2 설계 단계(S200)를 포함할 수 있다.
- [0034] 제1 설계 단계(S100)에서는 템플릿의 포인트를 변위 벡터로 매핑하기 위한 SIREN 다층 퍼셉트론(Multi Layer Perceptron, MLP)이 설계되고, 제2 설계 단계(S100)에서는 SIREN 다층 퍼셉트론에 대한 매개변수를 생성하여 모양의 변형을 발생되도록 하는 하이퍼네트워크가 설계될 수 있다. 즉, 영상처리 장치(100)는 SIREN 다층 퍼셉트론이 포함되는 모듈과, 하이퍼네트워크가 포함되는 모듈, 2개로 모듈을 갖도록 구성될 수 있다.
- [0035] 제1 설계 단계(S100)에서는 하나의 3차원 형상을 나타내는 표현방식으로 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수를 사용하도록 할 수 있다. 그리고 제2 설계 단계(S200)에서는 잠재코드를 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수의 매개변수로 변환하는 하이퍼네트워크를 설계할 수 있다.
- [0036] 우선, 제1 설계 단계(S100)에 대하여 설명하면, 영상처리 장치(100)로 제공되는 템플릿 메쉬 표면 상에 존재하는 3차원 좌표를 포인트 샘플링하여 좌표 상에 더해질 3차원 변형값을 출력하도록 다층 퍼셉트론을 구성한다.
- [0037] 즉, 기계 학습을 위해서는 템플릿 메쉬와 해당 변위에 대한 포인트를 샘플링해야 한다. 포인트 샘플링을 수행하는 방식 중 하나는 템플릿 메쉬의 정점(Vertex)을 이용하는 것이다.
- [0038] 정점은 눈과 같은 세부사항이 필요한 위치에 대한 중요한 샘플이 포함된다. 그러나 템플릿 메쉬의 삼각지역(Triangle)은 서로 크기가 상이하므로, 주변의 모양(예를 들어, 볼)은 표면에 균일한 샘플링을 통해 정확하게 캡처할 수 있다.
- [0039] 이에, 템플릿 메쉬의 포인트 샘플링에서는 하이브리드 방식의 포인트 샘플링 알고리즘이 적용되어 정점과 삼각지역을 기반으로 더 정확한 형태를 생성할 수 있다. 특히, 하이브리드 방식의 포인트 샘플링 알고리즘은 정점 방식만 사용하는 경우의 0.0188의 오류와 비교하여 0.0171의 오류를 보였다.
- [0040] 한편, 제1 설계 단계(S100)에서 구성된 다층 퍼셉트론은 하이퍼네트워크로부터 제공되는 매개변수의 조작에 기인하여 3차원 형상의 변형이 발생되도록 할 수 있다.
- [0041] 일반적으로 기계 학습의 결과는 기계 학습이 다루고자 하는 데이터의 표현 방식 설계에 큰 영향을 받는다. 3차원 형상의 잠재공간 학습에서는 3차원 형상을 나타내는 설계가 중요하다. 이에, 기계 학습 방법에서는 종래 3차원 형상 표현방식의 문제점을 해결하는 방식으로 잠재공간의 학습을 수행할 수 있다.
- [0042] 종래의 잠재공간 학습에서는 3차원 형상 표현방식으로 유한요소 기반 표현방식과 음함수 기반 표현방식이 사용되고 있다.
- [0043] 유한요소 기반 표현방식은 3차원 형상의 표면의 3차원 좌표 표본인 점 집합 또는 그래프를 통해 형상을 나타낸다. 이에, 3차원 좌표 표본을 사용하는 유한요소 기반 표현방식을 생성 모델에서 사용하는 종래의 잠재공간 학습은 일반 얼굴 도메인에서는 잘 작동하지만, 3차원 캐리커처와 같이 각 얼굴 간의 차이가 크고 다양한 도메인에서는 학습이 효과적으로 되지 않는다.
- [0044] 그리고 음함수 기반 표현방식은 거리함수(Signed Distance Function, SDF)를 다층 퍼셉트론으로 매개변수화하여 형상 간 차이가 큰 의자 등 도메인에서는 효과적으로 생성 모델을 학습할 수 있다. 그러나 다층 퍼셉트론으로 매개변수화된 거리함수는 형상의 디테일을 잘 표현하지 못하는 문제점이 있다. 이에, 음함수 기반 표현방식은 게임 엔진 등 상용 프레임워크에서 생성된 3차원 형상을 사용하려면 음함수에서 메쉬를 뽑아내는 추가적인 단계가 요구된다.
- [0045] 그러나 제1 설계 단계(S100)에서 적용된 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수 표현방식(이하, 메쉬 변형 함수 표현방식이라 칭한다.)은 다층 퍼셉트론을 통해 고정된 템플릿 메쉬 상의 3차원 좌표를 변형된 3차원 좌표로 매핑하여 3차원 형상 하나를 나타낸다. 메쉬 변형 함수 표현방식은 다층 퍼셉트론의 표현력을 사용하면서도 생성되는 형상이 추가적인 과정 필요없이 메쉬로 구성될 수 있다.



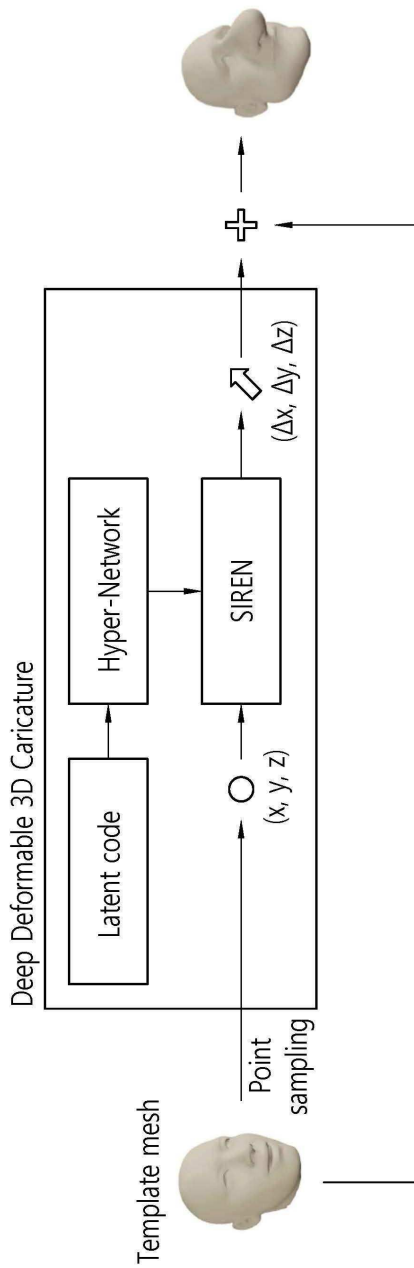
- [0046] 즉, 메쉬 변형 함수 표현방식은 상용 프레임워크에서 바로 사용할 수 있는 메쉬 표현방식의 장점과 다양한 현상을 효과적으로 학습할 수 있는 다층 퍼셉트론 매개변수화의 장점을 가질 수 있다.
- [0047] 한편, 이러한 메쉬 변형 함수 표현방식은 다른 표현방식과 비교해서 테스트데이터에 대한 재구성 오류가 낮아지는 것을 확인할 수 있다.
- [0048] 도 4는 본 실시예에 따른 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수 표현방식과 메쉬 정점의 3차원 좌표배열 방식을 적용하였을 때를 비교한 도면이다.
- [0049] 도 4에 도시된 바와 같이, 본 실시예에 따른 메쉬 변형 함수 표현방식을 Ground-truth 대비 정점 위치(Vertex position) 배열 방식과 비교하며, 메쉬 변형 함수 표현방식의 에러율이 현저하게 낮아지는 것을 알 수 있다.
- [0050] 정점 위치 배열 방식은 모든 조밀한 대응이 데이터 세트에서 제공될 때에 간단히 접근할 수 있는 방식이다. 다만, 정점 위치 배열 방식은 정점 배열의 생성에서 느린 수렴과 높은 재구성 오류를 보인다. 그리고 정점 위치 배열 방식을 이용한 경우에 표면의 샘플링 위치를 정점으로만 제한하게 된다.
- [0051] 또한, 본 실시예에 따른 정점 위치 배열 방식은 평균오차가 0.032 및 0.027으로 나타나 동일한 Ground-truth 대비 0.017 및 0.016의 평균오차를 갖는 메쉬 변형 함수 표현방식 보다 높은 오류를 보이는 것으로 나타났다.
- [0052] 도 5는 본 실시예에 따른 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수 표현방식과, DeepSDF 방식 및 DIF-NET 방식을 적용하였을 때를 비교한 도면이다.
- [0053] 도 5에 도시된 바와 같이, 본 실시예에 따른 메쉬 변형 함수 표현방식을 Ground-truth 대비 거리함수 방식 및 DIF-NET 방식과 비교하면, 메쉬 변형 함수 표현방식의 표현이 거리함수 방식 및 DIF-NET 방식 보다 Ground-truth에 유사하다는 것을 알 수 있다.
- [0054] 여기서, 거리함수 방식은 자동 디코더 프레임워크를 적용하여 3차원 모형을 표현하는 거리함수를 학습하기 위해 다층 퍼셉트론을 적용하고, DIF-NET 방식은 생성된 거리함수 간의 조밀한 대응을 제공하는 다층 퍼셉트론을 적용하여 템플릿의 거리함수 및 최적 변형 함수를 학습하는 방식일 수 있다.
- [0055] 이와 같이, 본 실시예에 따른 메쉬 변형 함수 표현방식을 적용하면, 얼굴 도메인에서 눈 또는 입 주변의 세부적인 형태가 더 잘 구현되는 것을 확인하였다.
- [0056] 한편, 다시 도 1 및 도 2를 참조하면, 다양한 3차원 형상을 하나의 잠재공간 상에 모델링하기 위해서는 잠재코드를 3차원 형상으로 변환하는 것이 요구된다. 메쉬 변형 함수 표현방식은 하나의 3차원 형상이 다층 퍼셉트론의 매개변수로 표현된다. 이에, 다양한 3차원 형상을 학습하기 위해서는 잠재코드를 다층 퍼셉트론의 매개변수로 변화하는 추가적인 요소가 요구된다.
- [0057] 이에, 제2 설계단계(S100)와 같이, 잠재코드를 다층 퍼셉트론 기반 템플릿 메쉬 변형 함수의 매개변수로 변환하는 하이퍼네트워크를 적용한다. 여기서, 하이퍼네트워크는 다층 퍼셉트론의 매개변수를 생성하는 다층 퍼셉트론으로 구성되어 기계 학습이 이루어지도록 한다.
- [0058] 하이퍼네트워크를 학습하기 위한 프레임워크로는 auto-decoder 프레임워크를 적용하여 학습 예시 각각에 해당하는 잠재코드와 하이퍼네트워크를 동시에 학습할 수 있다. 이때, 손실함수로는 잠재코드의 Squared L2 norm과 Ground-truth 형상과의 평균 제곱 오차(Mean Square Error, MSE)의 합을 사용한다. 평균 제곱 오차는 Ground-truth 메쉬의 표면에서 균일하게 추출한 3차원 좌표와 Ground-truth 메쉬 정점의 3차원 좌표의 합집합에 대해 계산될 수 있다.
- [0059] 한편, 이하에서는 본 실시예에 따른 기계 학습 방법에 따라 학습된 잠재공간을 사용하여 영상처리 장치(100)의 다양한 어플리케이션 수행에 대하여 설명하도록 한다.
- [0060] 도 6은 본 실시예에 따른 영상처리 장치에서 2차원 캐리커처의 랜드마크를 기반으로 3차원 캐리커처를 재구성한 도면이다.
- [0061] 도 6에 도시된 바와 같이, 본 실시예에 따른 영상처리 장치(100)는 2차원 캐리커처의 랜드마크에서 3차원 캐리커처를 재구성할 수 있다.
- [0062] 3차원 캐리커처의 재구성은 잠재공간을 사용하여 쉽게 편집될 수 있으며, 2차원 캐리커처의 랜드마크를 사용하여 랜드마크의 제약 조건에 맞게 하이퍼네트워크의 잠재코드를 최적화할 수 있다. 이에, 2차원 랜드마크의 위치와 3차원 랜드마크 정점이 주어지면 반복 최적화를 통해 피팅을 수행하여 2차원 캐리커처의 랜드마크를 기반으

로 3차원 캐리커처를 재구성할 수 있다.

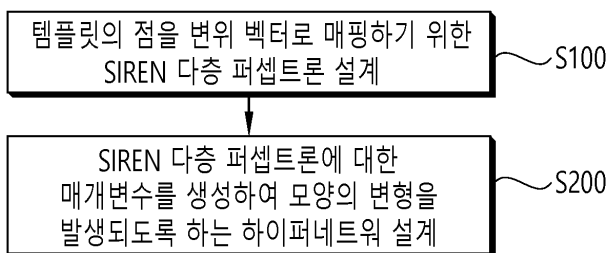
- [0063] 도 7은 본 실시예에 따른 영상처리 장치에서 의미론적 편집을 수행하는 것을 나타낸 도면이고, 도 8은 본 실시예에 따른 영상처리 장치에서 포인트-핸들-기반 편집을 수행하는 것을 나타낸 도면이다.
- [0064] 도 7 및 도 8에 도시된 바와 같이, 본 실시예에 따른 영상처리 장치(100)는 의미론적 편집(Sementic editing) 및 포인트-핸들-기반 편집(Point-handle-based editing)을 수행할 수 있다.
- [0065] 의미론적 편집에서 영상처리 장치(100)는 잠재코드를 조작하여 의미론적 편집을 수행할 수 있다. 즉, 의미론적 편집에서는 잠재공간을 사용하여 편집을 수행할 수 있으며, 의미론적 표현에 맞게 하이퍼네트워크의 잠재코드를 최적화할 수 있다. 이에, 영상처리 장치는 편집 백터를 더하거나 감소하는 것에 의해 3차원 캐리커처의 변형이 발생되도록 할 수 있다. 일례로, 의미론적 편집에서는 웃음 레이블 편집, 이마 및 코 등의 변화가 발생되도록 할 수 있으나, 이는 본 실시예를 설명하기 위한 것으로 의미론적 편집 대상을 한정하지는 않는다.
- [0066] 또한, 포인트-핸들-기반 편집에서 영상처리 장치는 학습된 잠재공간을 사용하여 3차원 캐리커처 모양을 변화시킬 수 있다. 이때, 영상처리 장치(100)는 편집을 위한 입력이 하나 또는 두 개의 점과 같이 매우 적더라도 로컬 편집을 완료하기 위해 자연스러운 변형이 발생되도록 할 수 있다.
- [0067] 여기서, 영상처리 장치(100)는 초기 3차원 형상에 대한 잠재코드, 핸들에 대한 정점 인덱스 목록 및 각 핸들에 대한 3차원 변위 백터가 제공되며 최적화를 통해 포인트-핸들-기반 편집을 수행할 수 있다. 일례로, 포인트-핸들-기반 편집에서는 코의 한 포인트를 선택하면 포인트를 앞으로 이동하거나, 턱 아래 쪽의 한 포인트를 선택하면 포인트를 아래쪽으로 이동하거나, 뺨의 양쪽에서 한 포인트를 선택하면 양 쪽 뺨 각각의 두 포인트를 옆으로 늘리거나, 귀의 양쪽에서 한 포인트를 선택하면 양쪽 귀 각각의 두 포인트를 늘릴 수 있다.
- [0068] 도 9는 본 실시예에 따른 영상처리 장치에서 자동으로 3차원 캐리커처를 생성하는 것을 나타낸 도면이다.
- [0069] 도 9에 도시된 바와 같이, 본 실시예에 따른 영상처리 장치(100)는 자동으로 3차원 캐리커처를 생성할 수 있다. 이때, 영상처리 장치(100)는 의미론적 편집을 통해 3차원 얼굴에 대해 자동으로 캐리커처를 생성할 수 있다.
- [0070] 일례로, 영상처리 장치(100)는 3차원 캐리커처와 일반 3차원 얼굴로 모델을 훈련하여 3차원 캐리커처와 일반 얼굴 모두에 연계된 모델을 획득할 수 있다. 이때, 영상처리 장치(100)는 일반 3차원 얼굴에서 3차원 캐리커처 얼굴로의 편집 방향을 계산하여 일반 얼굴의 잠재코드에 편집 작업을 적용할 수 있다.
- [0071] 일례로, 영상처리 장치(100)는 자동 3차원 캐리커처 생성에서 2차원 캐리커처를 생성하고, 생성된 2차원 캐리커처에 표시된 랜드마크를 사용하여 하이퍼네트워크의 잠재코드를 최적화하여 3차원 캐리커처를 생성할 수 있다.
- [0072] 이와 같이, 본 발명에 따른 영상처리 장치 및 영상처리 장치를 위한 3차원 형상 잠재공간의 기계 학습 방법은 3차원 형상을 템플릿 변형 함수 표현방식으로 구현하여 잠재공간의 학습 중 형상의 얼굴다움을 보장하는 매개변수화를 제공하며, 매개변수를 잠재 코드(Latent code)로부터 생성하는 다층 퍼셉트론 기반 하이퍼 네트워크를 통해 효율적인 학습을 수행하는 효과가 있다.
- [0073] 앞에서 설명되고, 도면에 도시된 본 발명의 일 실시예는 본 발명의 기술적 사상을 한정하는 것으로 해석되어서는 안 된다. 본 발명의 보호범위는 청구범위에 기재된 사항에 의하여만 제한되고, 본 발명의 기술분야에서 통상의 지식을 가진 자는 본 발명의 기술적 사상을 다양한 형태로 개량 변경하는 것이 가능하다. 따라서 이러한 개량 및 변경은 통상의 지식을 가진 자에게 자명한 것인 한 본 발명의 보호범위에 속하게 될 것이다.

도면

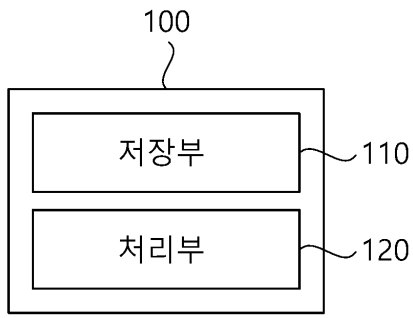
도면1



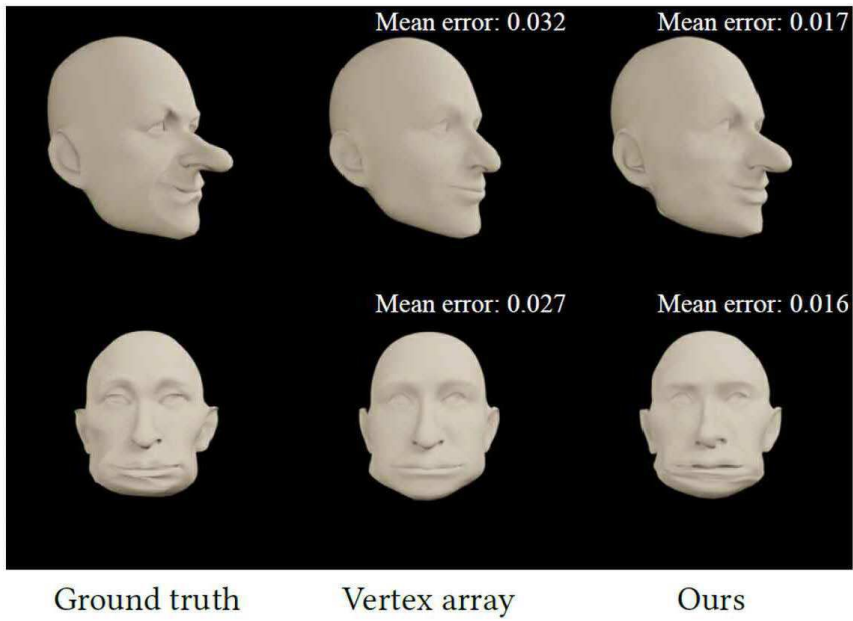
도면2



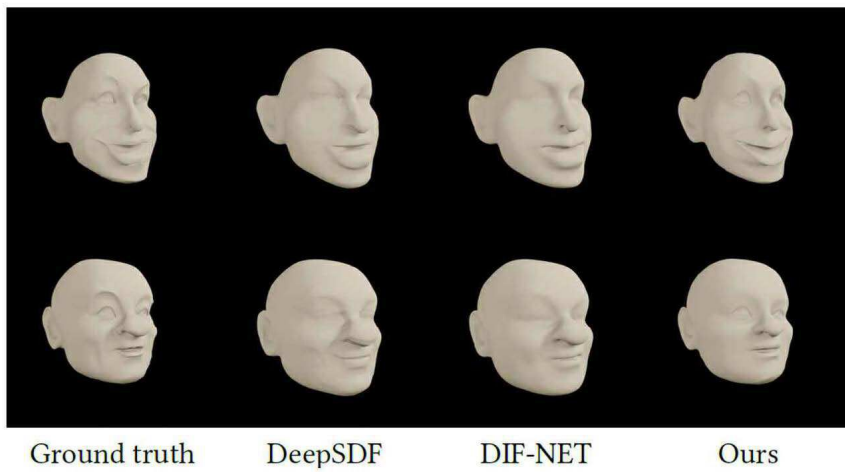
도면3



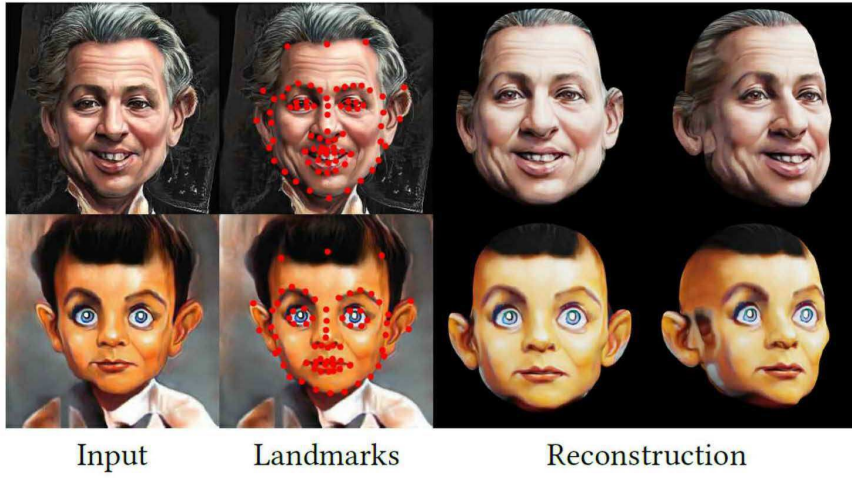
도면4



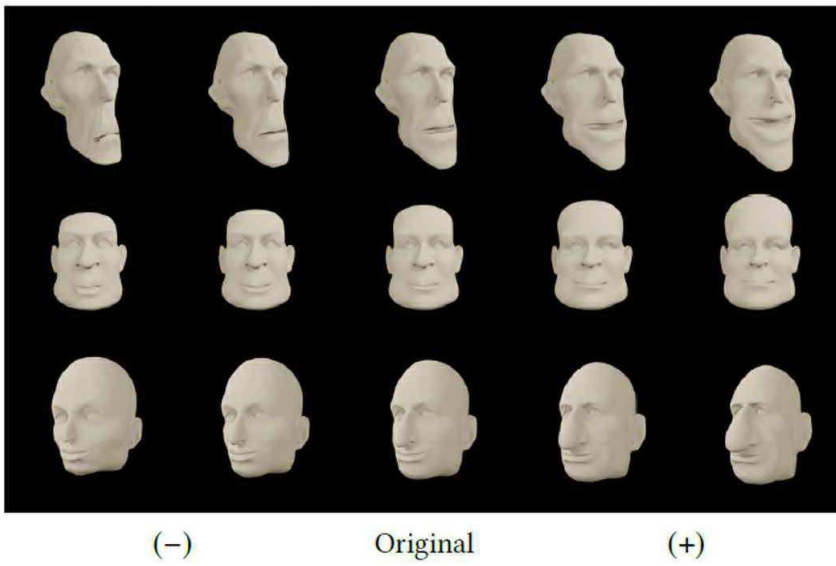
도면5



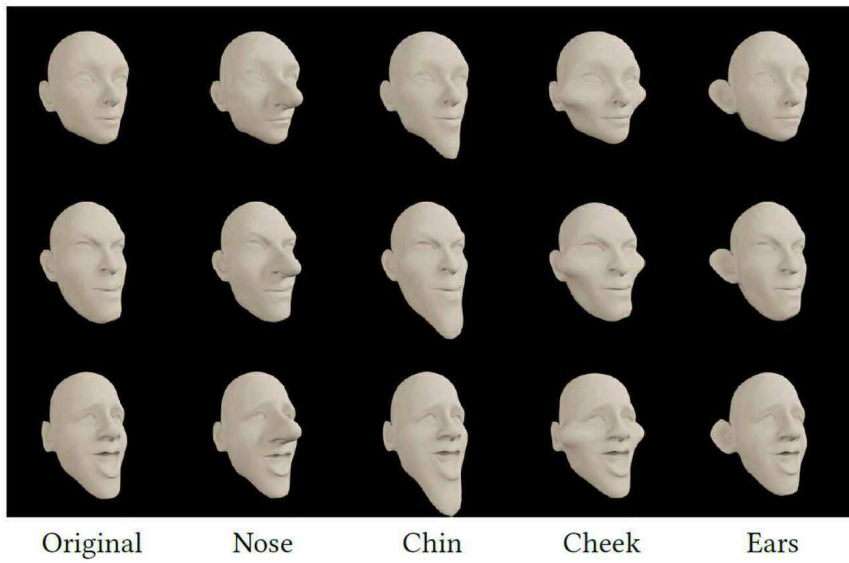
도면6



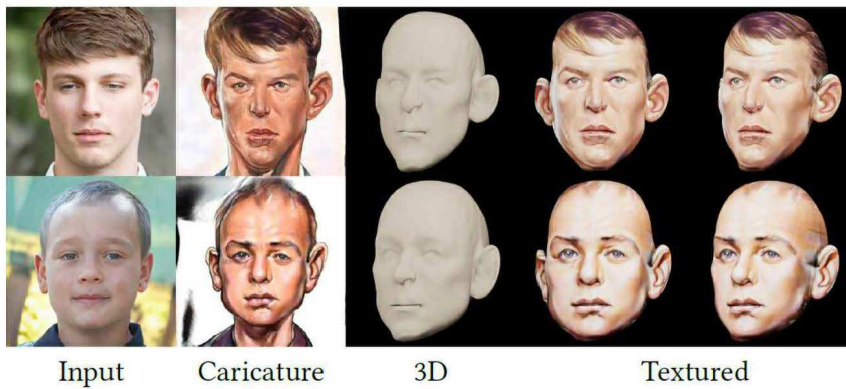
도면7



도면8



도면9



【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 9

【변경전】

제8 항에 있어서,

상기 평균 제곱 오차는

상기 Ground-truth 메쉬의 표면에서 균일하게 추출한 3차원 좌표와 Ground-truth 메쉬 정점의 3차원 좌표의 합집합에 대하여 계산되는 것을 특징으로 하는 영상처리 장치.

【변경후】

제8 항에 있어서,

상기 평균 제곱 오차는

Ground-truth 메쉬의 표면에서 균일하게 추출한 3차원 좌표와 Ground-truth 메쉬 정점의 3차원 좌표의 합집합에 대하여 계산되는 것을 특징으로 하는 영상처리 장치.

**【직권보정 2】**

**【보정항목】** 청구범위

**【보정세부항목】** 청구항 13

**【변경전】**

3차원 형상의 잠재공간을 학습하여 템플릿 메쉬의 변형을 수행하는 영상처리 장치의 기계 학습 방법에 있어서,

상기 템플릿 메쉬의 3차원 좌표를 포인트 샘플링하는 단계; 및

상기 포인트 샘플링된 상기 3차원 좌표를 층 퍼셉트론(Multi Layer Perceptron, MLP)에 적용하여 상기 3차원 좌표 상에 추가되는 3차원 변형값을 출력하며 상기 잠재공간의 원소인 잠재코드(Latent code)를 하이퍼 네트워크에 적용하여 상기 다층 퍼셉트론에 대한 매개변수를 출력하고,

상기 잠재코드는 다양한 3차원 형상을 하나의 잠재공간 상에 모델링하기 위하여 3차원 형상으로 변환되고,

상기 하이퍼네트워크는 상기 잠재코드를 3차원 형상으로 변환하는 다층 퍼셉트론을 포함하고,

상기 하이퍼네트워크의 학습에서는 Auto-decoder 프레임워크가 적용되어 상기 잠재코드와 상기 하이퍼네트워크를 동시에 학습하는 것을 특징으로 하는 영상처리 장치의 기계 학습 방법.

**【변경후】**

3차원 형상의 잠재공간을 학습하여 템플릿 메쉬의 변형을 수행하는 영상처리 장치의 기계 학습 방법에 있어서,

상기 템플릿 메쉬의 3차원 좌표를 포인트 샘플링하는 단계; 및

상기 포인트 샘플링된 상기 3차원 좌표를 다층 퍼셉트론(Multi Layer Perceptron, MLP)에 적용하여 상기 3차원 좌표 상에 추가되는 3차원 변형값을 출력하며 상기 잠재공간의 원소인 잠재코드(Latent code)를 하이퍼 네트워크에 적용하여 상기 다층 퍼셉트론에 대한 매개변수를 출력하고,

상기 잠재코드는 다양한 3차원 형상을 하나의 잠재공간 상에 모델링하기 위하여 3차원 형상으로 변환되고,

상기 하이퍼네트워크는 상기 잠재코드를 3차원 형상으로 변환하는 다층 퍼셉트론을 포함하고,

상기 하이퍼네트워크의 학습에서는 Auto-decoder 프레임워크가 적용되어 상기 잠재코드와 상기 하이퍼네트워크를 동시에 학습하는 것을 특징으로 하는 영상처리 장치의 기계 학습 방법.