



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0071008
(43) 공개일자 2020년06월18일

(51) 국제특허분류(Int. Cl.)
G06T 17/20 (2006.01) G06T 1/00 (2006.01)
G06T 3/40 (2006.01) G06T 7/11 (2017.01)
(52) CPC특허분류
G06T 17/20 (2013.01)
G06T 1/00 (2013.01)
(21) 출원번호 10-2019-0154559
(22) 출원일자 2019년11월27일
심사청구일자 없음
(30) 우선권주장
2018143561 2018년12월10일 러시아(RU)

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
(72) 발명자
글라지스토브 이반 빅토로비치
러시아 433100 울리야놉스크 베스카야마 라보차야
스트리트 33 아파트 5
카라차로브 이반 올레고비치
러시아 117292 모스크바 이바나 바부스키나 스트
리트 15 빌딩 3 아파트 9
(뒷면에 계속)
(74) 대리인
리엔목특허법인

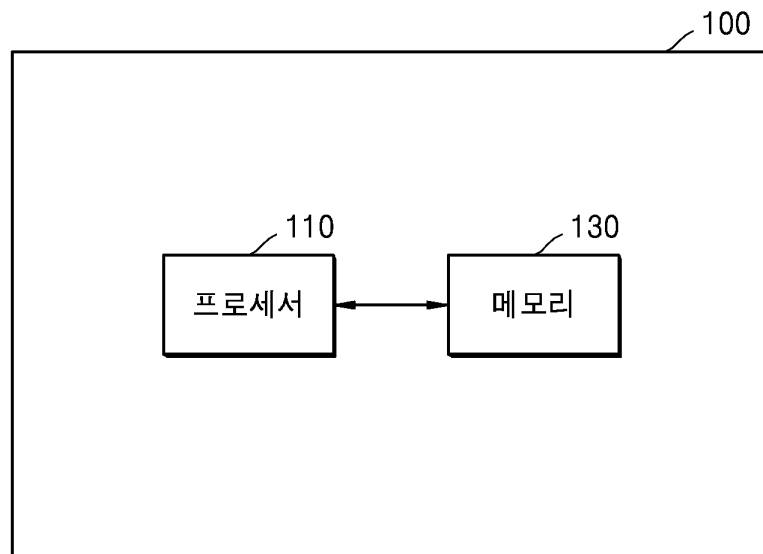
전체 청구항 수 : 총 17 항

(54) 발명의 명칭 2차원 이미지 처리 방법 및 이 방법을 실행하는 디바이스

(57) 요약

2차원 이미지를 처리하는 방법에 있어서, 2차원 이미지를 획득하는 단계, 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하는 단계 및 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 기초로 3차원 얼굴 모델을 생성하는 단계를 포함하는 방법을 제공한다.

대표도 - 도1



(52) CPC특허분류

G06T 3/40 (2013.01)

G06T 7/11 (2017.01)

G06T 2207/20081 (2013.01)

G06T 2207/20084 (2013.01)

G06T 2207/30201 (2013.01)

G06T 2210/12 (2013.01)

(72) 발명자

슈체르비닌 앤드리 유리비치

러시아 111399 모스크바 페테라티브니지 프로스펙
트 스트리트 9 빌딩 아파트 19

쿠릴린 일리야 바실리에비치

러시아 129594 모스크바 토라야 얼리타 메리노이
로스치 스트리트 20 아파트 106

명세서

청구범위

청구항 1

2차원 이미지를 획득하는 단계;

상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하는 단계; 및
상기 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 기초로 3차원 얼굴 모델을 생성하는 단계를 포함하는 2차원 이미지 처리 방법.

청구항 2

제1항에 있어서, 상기 트레이닝된 CNN에 의하여 처리하는 단계는,

상기 2차원 이미지의 크기를 조정하는 단계;

상기 크기가 조정된 2차원 이미지로부터 얼굴을 감지하는 단계;

상기 감지된 얼굴이 포함된 영역을 바운딩 박스로 표시하는 단계;

상기 2차원 이미지로부터 상기 바운딩 박스로 표시된 얼굴 영역의 이미지를 잘라내는 단계;

상기 잘라낸 얼굴 영역 이미지의 크기를 조정하여 제2 이미지를 생성하는 단계;

상기 제2 이미지로부터 상기 얼굴에 대한 적어도 하나의 랜드마크를 마킹하는 단계; 및

상기 적어도 하나의 랜드마크와 가장 근접하게 매치되는 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하는 단계를 포함하는 2차원 이미지 처리 방법.

청구항 3

제1항에 있어서,

상기 적어도 하나의 카메라 파라미터는 비선형 최적화를 이용하여 획득되고, 상기 적어도 하나의 얼굴 모델 파라미터는 선형 최적화를 이용하여 획득되는 2차원 이미지 처리 방법.

청구항 4

제1항에 있어서, 상기 CNN은 다양한 방식으로 트레이닝되는 2차원 이미지 처리 방법.

청구항 5

제1항에 있어서,

상기 3차원 얼굴 모델은 상기 적어도 하나의 카메라 파라미터 및 상기 적어도 하나의 얼굴 모델 파라미터를 3차원 변형 가능 얼굴 모델(3차원 morphable face model)에 적용함으로써 생성되는 2차원 이미지 처리 방법.

청구항 6

제1항에 있어서,

상기 생성된 3차원 얼굴 모델을 상기 2차원 이미지의 얼굴 영역에 오버레이(overlay)하는 단계를 더 포함하는 2차원 이미지 처리 방법.

청구항 7

제6항에 있어서, 상기 오버레이는 핀 홀 카메라 모델을 이용하여 이루어지는 2차원 이미지 처리 방법.

청구항 8

제1항에 있어서,

상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원 이미지로부터 적어도 하나의 감정 파라미터를 획득하는 단계를 더 포함하는 2차원 이미지 처리 방법.

청구항 9

하나 이상의 인스트럭션을 저장하는 메모리;

상기 하나 이상의 인스트럭션을 실행하는 프로세서를 포함하고,

상기 프로세서는, 2차원 이미지를 획득하고, 상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하고, 상기 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 기초로 3차원 얼굴 모델을 생성하는 2차원 이미지 처리 디바이스.

청구항 10

제9항에 있어서, 상기 프로세서는,

상기 2차원 이미지의 크기를 조정하고, 상기 크기가 조정된 2차원 이미지로부터 얼굴을 감지하고, 상기 감지된 얼굴이 포함된 영역을 바운딩 박스로 표시하고, 상기 2차원 이미지로부터 상기 바운딩 박스로 표시된 얼굴 영역의 이미지를 잘라내고, 상기 잘라낸 얼굴 영역 이미지의 크기를 조정하여 제2 이미지를 생성하고, 상기 제2 이미지로부터 상기 얼굴에 대한 적어도 하나의 랜드마크를 마킹하고, 상기 적어도 하나의 랜드마크와 가장 근접하게 매치되는 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득함으로써 상기 2차원 이미지를 상기 트레이닝된 CNN에 의하여 처리하는 2차원 이미지 처리 디바이스.

청구항 11

제9항에 있어서,

상기 적어도 하나의 카메라 파라미터는 비선형 최적화를 이용하여 획득되고, 상기 적어도 하나의 얼굴 모델 파라미터는 선형 최적화를 이용하여 획득되는 2차원 이미지 처리 디바이스.

청구항 12

제9항에 있어서, 상기 CNN은 다양한 방식으로 트레이닝되는 2차원 이미지 처리 디바이스.

청구항 13

제9항에 있어서,

상기 3차원 얼굴 모델은 상기 적어도 하나의 카메라 파라미터 및 상기 적어도 하나의 얼굴 모델 파라미터를 3차원 변형 가능 얼굴 모델(3차원 morphable face model)에 적용함으로써 생성되는 2차원 이미지 처리 디바이스.

청구항 14

제9항에 있어서,

상기 생성된 3차원 얼굴 모델을 상기 2차원 이미지의 얼굴 영역에 오버레이(overlay)하는 2차원 이미지 처리 디바이스.

청구항 15

제14항에 있어서, 상기 오버레이는 핀 홀 카메라 모델을 이용하여 이루어지는 2차원 이미지 처리 디바이스.

청구항 16

제9항에 있어서, 상기 프로세서는,

상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원

원 이미지로부터 적어도 하나의 감정 파라미터를 획득하는 2차원 이미지 처리 디바이스.

청구항 17

제 1항 내지 8 항 중, 어느 한 항의 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

발명의 설명

기술 분야

[0001] 본 개시는 2차원 이미지 처리 방법 및 이 방법을 실행하는 디바이스에 관한 것이다. 좀 더 구체적으로 본 개시는 2차원 이미지 상의 얼굴에 대한 3차원 모델의 생성 방법 및 그 방법을 실행하는 디바이스에 관한 것이다.

배경 기술

[0002] 2차원 이미지로부터 얼굴을 추출하여 3차원 얼굴 모델을 생성하는 기술은 다방면에서 이용되고 있다.

[0003] 그러나, 알고리즘의 복잡성 등으로 인하여 고사양의 하드웨어가 필요하므로, 추가적인 하드웨어의 지원 없이 모바일 디바이스 만으로 실행하는 데에는 어려움이 있었다.

발명의 내용

해결하려는 과제

과제의 해결 수단

[0004] 상술한 기술적 과제를 달성하기 위한 기술적 수단으로서, 본 개시의 일 실시예에 따르면, 2차원 이미지를 획득하는 단계, 상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하는 단계 및 상기 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 기초로 3차원 얼굴 모델을 생성하는 단계를 포함하는 2차원 이미지 처리 방법 방법이 제공될 수 있다.

[0005] 위 2차원 이미지 처리 방법 방법에 있어서, 상기 트레이닝된 CNN에 의하여 처리하는 단계는, 상기 2차원 이미지의 크기를 조정하는 단계, 상기 크기가 조정된 2차원 이미지로부터 얼굴을 감지하는 단계, 상기 감지된 얼굴이 포함된 영역을 바운딩 박스로 표시하는 단계, 상기 2차원 이미지로부터 상기 바운딩 박스로 표시된 얼굴 영역의 이미지를 잘라내는 단계, 상기 잘라낸 얼굴 영역 이미지의 크기를 조정하여 제2 이미지를 생성하는 단계, 상기 제2 이미지로부터 상기 얼굴에 대한 적어도 하나의 랜드마크를 마킹하는 단계 및 상기 적어도 하나의 랜드마크와 가장 근접하게 매치되는 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하는 단계를 포함할 수 있다.

[0006] 또한, 위 2차원 이미지 처리 방법 방법에 있어서, 상기 적어도 하나의 카메라 파라미터는 비선형 최적화를 이용하여 획득되고, 상기 적어도 하나의 얼굴 모델 파라미터는 선형 최적화를 이용하여 획득될 수 있다.

[0007] 위 2차원 이미지 처리 방법 방법에 있어서, 상기 CNN은 다양한 방식으로 트레이닝될 수 있다.

[0008] 위 2차원 이미지 처리 방법 방법에 있어서, 상기 3차원 얼굴 모델은 상기 적어도 하나의 카메라 파라미터 및 상기 적어도 하나의 얼굴 모델 파라미터를 3차원 변형 가능 얼굴 모델(3차원 morphable face model)에 적용함으로써 생성될 수 있다.

[0009] 위 2차원 이미지 처리 방법 방법에 있어서, 상기 생성된 3차원 얼굴 모델을 상기 2차원 이미지의 얼굴 영역에 오버레이(overlay)하는 단계를 더 포함할 수 있다.

[0010] 이 때, 상기 오버레이는 핀 홀 카메라 모델을 이용하여 이루어질 수 있다.

- [0011] 위 2차원 이미지 처리 방법 방법은, 상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원 이미지로부터 적어도 하나의 감정 파라미터를 획득하는 단계를 더 포함할 수 있다.
- [0012] 본 개시 일 실시예의 또 다른 측면에 따르면, 하나 이상의 인스트럭션을 저장하는 메모리, 상기 하나 이상의 인스트럭션을 실행하는 프로세서를 포함하고, 상기 프로세서는, 2차원 이미지를 획득하고, 상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하고, 상기 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 기초로 3차원 얼굴 모델을 생성하는 2차원 이미지 처리 디바이스가 제공될 수 있다.
- [0013] 상기 프로세서는, 상기 2차원 이미지의 크기를 조정하고, 상기 크기가 조정된 2차원 이미지로부터 얼굴을 감지하고, 상기 감지된 얼굴이 포함된 영역을 바운딩 박스로 표시하고, 상기 2차원 이미지로부터 상기 바운딩 박스로 표시된 얼굴 영역의 이미지를 잘라내고, 상기 잘라낸 얼굴 영역 이미지의 크기를 조정하여 제2 이미지를 생성하고, 상기 제2 이미지로부터 상기 얼굴에 대한 적어도 하나의 랜드마크를 마킹하고, 상기 적어도 하나의 랜드마크와 가장 근접하게 매치되는 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득함으로써 상기 2차원 이미지를 상기 트레이닝된 CNN에 의하여 처리할 수 있다.
- [0014] 위 2차원 이미지 처리 디바이스에 있어서, 상기 적어도 하나의 카메라 파라미터는 비선형 최적화를 이용하여 획득되고, 상기 적어도 하나의 얼굴 모델 파라미터는 선형 최적화를 이용하여 획득될 수 있다.
- [0015] 위 2차원 이미지 처리 디바이스에 있어서, 상기 CNN은 다양한 방식으로 트레이닝될 수 있다.
- [0016] 위 2차원 이미지 처리 디바이스에 있어서, 상기 3차원 얼굴 모델은 상기 적어도 하나의 카메라 파라미터 및 상기 적어도 하나의 얼굴 모델 파라미터를 3차원 변형 가능 얼굴 모델(3차원 morphable face model)에 적용함으로써 생성될 수 있다.
- [0017] 위 2차원 이미지 처리 디바이스에 있어서, 프로세서는 상기 생성된 3차원 얼굴 모델을 상기 2차원 이미지의 얼굴 영역에 오버레이(overlay)할 수 있다.
- [0018] 위 2차원 이미지 처리 디바이스에 있어서, 상기 오버레이는 핀 홀 카메라 모델을 이용하여 이루어질 수 있다.
- [0019] 상기 프로세서는, 상기 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 상기 2차원 이미지로부터 적어도 하나의 감정 파라미터를 획득할 수 있다.
- [0020] 본 개시 일 실시예의 또 다른 측면에 따르면, 실행되었을 때, 위 방법들의 실행을 야기하는 컴퓨터 프로그램 제품이 제공될 수 있다.
- [0021] 본 개시 일 실시예의 또 다른 측면에 따르면, 위 컴퓨터 프로그램 제품을 기록한 컴퓨터로 읽을 수 있는 기록 매체가 제공될 수 있다.

도면의 간단한 설명

- [0022] 도1은 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 디바이스의 구조도이다.
- 도2는 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 방법에 대한 흐름도이다.
- 도3은 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 트레이닝된 CNN에 의하여 데이터를 처리하는 방법에 대한 흐름도이다.
- 도4a는 본 개시의 일 실시예에 따라 트레이닝된 CNN이 데이터를 처리하는 방법의 흐름도이다.
- 도4b는 본 개시의 일 실시예에 따라 이용되는 CNN의 구조를 도시한 것이다.
- 도4c는 본 개시의 일 실시예에 따라 이용되는 CNN의 다른 구조를 도시한 것이다.
- 도5는 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 방법이 기존의 방법과 다른 점에 대한 예시이다.
- 도6은 본 개시의 일 실시예에 따라 CNN이 트레이닝 되는 방법에 대한 흐름도이다.
- 도7은 본 개시의 일 실시예에 따라 감정 파라미터가 추가로 획득되는 경우에 대한 흐름도이다.

도8은 본 개시의 일 실시예에 따라 2차원 이미지를 기초로 3차원 아바타를 생성하는 예시이다.

도9는 본 개시의 일 실시예에 따라 2차원 이미지 처리 디바이스가 행인을 촬영한 이미지로부터 감정 파라미터를 획득하여, 적절한 광고를 제공하는 예시이다.

도10은 본 개시의 일 실시예에 따라 애완로봇이 사용자를 촬영한 이미지로부터 감정 파라미터를 획득하여 사용자와 교감하는 예시이다.

도11은 본 개시의 일 실시예에 따라 생성된 3차원 얼굴 모델을 기초로 원근 왜곡을 교정하는 예시이다.

발명을 실시하기 위한 구체적인 내용

- [0023] 본 개시의 이점 및 특징, 그리고 그것들을 달성하는 방법은 첨부되는 도면과 함께 후술되어 있는 실시예들을 참조하면 명확해질 것이다. 그러나 본 개시가 이하에서 제시되는 실시예들에 한정되는 것이 아니라 서로 다른 다양한 형태로 구현될 수 있으며, 단지 본 실시예들은 본 개시가 속하는 기술분야에서 통상의 지식을 가진 자에게 발명의 범주를 알려주기 위해 제공되는 것이며, 본 개시는 청구항의 범주에 의해 정의될 뿐이다.
- [0024] 본 개시에서 사용되는 용어는 본 개시에서의 기능을 고려하면서 가능한 현재 널리 사용되는 일반적인 용어들을 선택하였으나, 이는 당 분야에 종사하는 기술자의 의도 또는 관례, 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 특정한 경우는 출원인이 임의로 선정한 용어도 있으며, 이 경우 해당되는 발명의 설명 부분에서 상세히 그 의미를 기재할 것이다. 따라서 본 개시에서 사용되는 용어는 단순한 용어의 명칭이 아닌, 그 용어가 가지는 의미와 본 개시의 전반에 걸친 내용을 토대로 정의되어야 한다.
- [0025] 명세서 전체에서 어떤 부분이 어떤 구성요소를 "포함"한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있음을 의미한다. 또한, 명세서에서 사용되는 "부"라는 용어는 소프트웨어, FPGA 또는 ASIC과 같은 하드웨어 구성요소를 의미하며, "부"는 어떤 역할을들을 수행한다. 그렇지만 "부"는 소프트웨어 또는 하드웨어에 한정되는 의미는 아니다. "부"는 어드레싱할 수 있는 저장 매체에 있도록 구성될 수도 있고 하나 또는 그 이상의 프로세서들을 재생시키도록 구성될 수도 있다. 따라서, 일 예로서 "부"는 소프트웨어 구성요소들, 객체지향 소프트웨어 구성요소들, 클래스 구성요소들 및 태스크 구성요소들과 같은 구성요소들과, 프로세스들, 함수들, 속성들, 프로시저들, 서브루틴들, 프로그램 코드의 세그먼트들, 드라이버들, 펌웨어, 마이크로 코드, 회로, 데이터, 데이터베이스, 데이터 구조들, 테이블들, 어레이들 및 변수들을 포함한다. 구성요소들과 "부"들 안에서 제공되는 기능은 더 작은 수의 구성요소들 및 "부"들로 결합되거나 추가적인 구성요소들과 "부"들로 더 분리될 수 있다.
- [0026] 아래에서는 첨부한 도면을 참고하여 본 개시의 실시예에 대하여 본 개시가 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 그리고 도면에서 본 개시를 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략한다.
- [0027] 도1은 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 디바이스의 구조도이다.
- [0028] 본 개시의 일 실시예에 따른 2차원 이미지 처리 디바이스(100)는 프로세서(110) 및 메모리(130)를 포함할 수 있다.
- [0029] 프로세서(110)는 2차원 이미지를 획득하고, 획득된 2차원 이미지를 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득할 수 있다.
- [0030] 2차원 이미지는 실시간으로 사용자의 모바일 컴퓨팅 디바이스의 카메라에 의하여 캡처된 비디오 스트림의 일부이거나 혹은 미리 저장된 비디오 스트림의 일부일 수 있다.
- [0031] 일부 실시예에서 2차원 이미지 처리 디바이스(100)는 비디오 스트림을 캡처하는 카메라와 같은 촬영부를 더 포함할 수 있다.
- [0032] 본 개시에서 CNN(convolutional neural network)은 입력 층(input layer)과 출력 층(output layer) 사이에 다중의 은닉 층(hidden layer)을 포함하는 인공 신경망인 심층 신경망(DNN: Deep Neural Network)의 한 종류로, 하나 또는 여러 개의 컨볼루션 레이어(convolutional layer)와 통합 레이어(pooling layer), 완전하게 연결된 레이어(fully connected layer)들로 구성된 신경망일 수 있다.
- [0033] CNN은 이미지와 같은 2차원 데이터의 학습에 적합한 구조를 가지고 있으며, 역전달(Backpropagation

algorithm)을 통해 훈련될 수 있다. CNN은 영상 내 객체 분류, 객체 탐지 등 다양한 응용 분야에 폭넓게 활용되는 DNN의 대표적 모델 중 하나이다. 본 개시에서는 다양한 방식으로 트레이닝된 CNN이 이용될 수 있다.

- [0034] 프로세서(110)가 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득하는 구체적인 방법에 대하여는 도3 등에서 후술한다.
- [0035] 프로세서(110)는 상기 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 기초로 3차원 얼굴 모델을 생성할 수 있다.
- [0036] 메모리(130)는 프로세서(110)의 실행을 야기하는 프로그램 인스트럭션 등을 저장할 수 있다. 메모리(130)는 프로세서에 의하여 실행될 때, 프로세서가 읽을 수 있고, 실행할 수 있는 인스트럭션을 저장하여, 프로세서가 2차원 이미지 처리 방법에 포함된 단계들을 실행할 수 있도록 한다.
- [0037] 일 실시예에서 메모리(130)는 트레이닝된 CNN을 저장할 수 있다.
- [0038] 본 개시의 일부 실시예에서 2차원 이미지 처리 디바이스(100)는 복수의 메모리를 포함할 수 있다.
- [0039] 본 실시예에서 프로세서(110) 및 메모리(130)는 별도의 구성 단위로 표현되어 있으나, 본 개시의 일부 실시예에서는 프로세서(110) 및 메모리(130)가 합쳐져 동일한 구성 단위로 구현될 수도 있다.
- [0040] 또한 본 실시예에서 프로세서(110) 및 메모리(130)는 2차원 이미지 처리 디바이스(100) 내부에 인접하여 위치한 구성 단위로 표현되었지만, 프로세서(110) 및 메모리(130)의 각 기능을 담당하는 장치는 반드시 물리적으로 인접할 필요는 없으므로, 실시 예에 따라 프로세서(110) 및 메모리(130)가 분산되어 있을 수 있다.
- [0041] 또한, 2차원 이미지 처리 디바이스(100)는 물리적 장치에 한정되지 않으므로, 2차원 이미지 처리 디바이스(100)의 기능 중 일부는 하드웨어가 아닌 소프트웨어로 구현될 수도 있다.
- [0042] 본 개시의 일부 실시예에 따르면 2차원 이미지 처리 디바이스(100)는 출력 부, 통신 인터페이스, 촬영부 등을 더 포함할 수 있다.
- [0043] 본 명세서에서 기술된 구성요소들 각각은 하나 또는 그 이상의 부품(component)으로 구성될 수 있으며, 해당 구성요소의 명칭은 디바이스의 종류에 따라서 달라질 수 있다. 다양한 실시예에서, 디바이스는 본 명세서에서 기술된 구성요소 중 적어도 하나를 포함하여 구성될 수 있으며, 일부 구성요소가 생략되거나 또는 추가적인 다른 구성요소를 더 포함할 수 있다. 또한, 다양한 실시예에 따른 디바이스의 구성요소들 중 일부가 결합되어 하나의 개체(entity)로 구성됨으로써, 결합되기 이전의 해당 구성요소들의 기능을 동일하게 수행할 수 있다.
- [0044] 디바이스(100)는 스마트 폰, 태블릿, 가상 현실 안경, 증강 현실 안경, PC, 랩탑, 스마트 워치 등과 같은 도1에 도시되지 않은 다른 구성 요소들을 포함할 수 있다.
- [0045] 다른 실시예에서 사용자 컴퓨팅 디바이스는 분리된 하드웨어 유닛들을 포함할 수 있다. 이러한 실시예에서, 각 하드웨어 유닛은 본 개시 방법의 각 단계 혹은 서브 단계를 책임질 수 있다.
- [0046] 도2는 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 방법에 대한 흐름도이다.
- [0047] S210 단계에서 디바이스(100)는 2차원 이미지를 획득할 수 있다. 2차원 이미지는 실시간으로 사용자의 모바일 컴퓨팅 디바이스의 카메라에 의하여 캡처된 비디오 스트림의 일부이거나 혹은 미리 저장된 비디오 스트림의 일부일 수 있다. 또한, 2차원 이미지는 인터넷이나 클라우드와 같은 네트워크로부터 다운로드된 이미지 시퀀스에 포함된 하나 혹은 그 이상의 이미지일 수 있다.
- [0048] 단계 S230에서 디바이스(100)는 캡처된 2차원 이미지를 미리 트레이닝된 CNN(convolutional neural network)에 의하여 처리함으로써, 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득할 수 있다.
- [0049] CNN은 도4B 또는 도4C에 도시된 구조와 같을 수 있다. CNN은 카메라 파라미터들 및 얼굴 모델 파라미터들을 추론하도록 구성될 수 있다. 구체적으로 CNN은 얼굴을 감지하고, 감지된 얼굴에 따라 위 파라미터들을 추론하여 획득할 수 있다.
- [0050] 이미지에서 얼굴을 감지하는 추가적인 능력을 가진 CNN의 제공은 얼굴 감지에 대하여 자원 집약적인 과정을 회피할 수 있게 하였다. 본 개시의 CNN에 의한 이미지 상의 얼굴 감지는 선행 기술과 비교하여 상대적으로 덜 자원 집약적인 작업이다.

- [0051] 단계 S250에서 디바이스(100)는 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 기초로 3차원 얼굴 모델을 생성할 수 있다.
- [0052] 일 실시예에서 디바이스(100)는 적어도 하나의 카메라 파라미터 및 상기 적어도 하나의 얼굴 모델 파라미터를 3차원 변형 가능 얼굴 모델(3D morphable face model)에 적용함으로써 3차원 얼굴 모델을 생성할 수 있다.
- [0053] 본 개시에서 변형 가능 얼굴 모델(Morphable face model)은 통계적인 얼굴 PCA(Primary Component Analysis) 모델을 의미할 수 있다.
- [0054] 이 모델은 $M = N + B$ 로 표현될 수 있고, 이 때 N 은 중성 모형(neutral model)이고, 기본 3차원 얼굴 모델로 참조되며, e 는 얼굴 모양 파라미터이고, B 는 기초 모양(shape basis)이다. 이 모델은 또한 N 포인트에 센터를 가지고, 공분산 행렬의 고유 벡터 B 의 집합을 가지는 다차원 정규 분포(multidimensional normal distribution)로 표현될 수 있다. 각 벡터 행렬 B 의 열은 이 벡터의 정보상의 중요성을 책임지는 고유치(eigenvalue) $eigen_i$ 에 대응할 수 있다.
- [0055] 감정 반영 모델을 설명하기 위하여, 위 공식 M 은 $\hat{M} = N + B^e + A\alpha$ 로 수정될 수 있다. 이 때, A 는 기초 감정(emotion basis)이고, a 는 감정 파라미터 벡터일 수 있다.
- [0056] 일부 실시예에서 디바이스(100)는 생성된 3차원 얼굴 모델을 2차원 이미지 상의 얼굴 영역에 오버레이하는 단계를 더 포함할 수 있다. 이러한 오버레이 과정에서는 간소화된 핀 홀 카메라 모델이 이용될 수 있다.
- [0057] 본 개시에서 핀 홀 카메라 모델은 3차원 공간에서 각 좌표점들 사이의 수학적 관계를 영상 홀로그램으로 투영한 것이다.
- [0058] 이것은 $U = (t_x, t_y, t_z, a, b, g, f)$ 와 같이 표현될 수 있고, 이 때, (t_x, t_y, t_z) 는 각 중심축에 대한 변형(translation)을 의미하고, (a, b, g) 는 회전 각을 의미하며, f 는 초점 거리를 의미할 수 있다.
- [0059] 투영은 $\dot{x} = \frac{f \cdot (R(\alpha, \beta, \gamma)x + t_x)}{(R(\alpha, \beta, \gamma)x)_z + t_z}$ 와 같이 도출될 수 있으며, 이 때 x 는 오리지널 좌표 점을 의미하고, $\dot{x}$ 는 투영된 좌표 점을 의미하며, f 는 초점 거리를 의미하고, $R(\alpha, \beta, \gamma)$ 는 회전 행렬을 의미할 수 있다.
- [0060] 변형 가능 페이스 모델의 파라미터 t 와 간소화된 핀 홀 카메라 모델의 파라미터들이 고정된 값이라면, 3차원 공간 $R(\alpha, \beta, \gamma)M(\epsilon) + t$ 에서 일련의 포인트들이 도출될 것이고, 그것들은 3차원 얼굴 모델로 참조될 것이다.
- [0061] 도3은 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 트레이닝된 CNN에 의하여 데이터를 처리하는 방법에 대한 흐름도이다.
- [0062] S310 단계에서 CNN은 도2의 단계 S210에서 획득된 2차원 이미지의 크기를 조정할 수 있다. CNN은 2차원 이미지를 확대하거나 축소할 수 있다.
- [0063] S320 단계에서 CNN은 크기가 조정된 2차원 이미지로부터 얼굴을 감지할 수 있다. S330 단계에서 CNN은 감지된 얼굴이 포함된 영역을 바운딩 박스로 표시할 수 있다. 각 감지된 얼굴의 바운딩 박스는 이미지 공간에서 얼굴 위치를 결정하는 것으로 암시될 수 있다. 일부 실시예에서 바운딩 박스는 얼굴보다 약간 크거나 작은 박스의 크기일 수 있다.
- [0064] S340 단계에서 CNN은 바운딩 박스로 표시된 얼굴 영역의 이미지를 잘라낼 수 있다. 본 실시예에서는 바운딩 박스를 기초로 얼굴 영역의 이미지를 잘라내지만, 이 단계는 다른 방법으로 실행될 수도 있다.
- [0065] S350 단계에서 CNN은 잘라낸 얼굴 영역 이미지의 크기를 조정하여 제2 이미지를 생성할 수 있다. CNN은 잘라낸 얼굴 영역 이미지의 크기를 확대하거나 축소하여 제2 이미지를 생성할 수 있다.
- [0066] 일 실시예에서, CNN은 잘라낸 얼굴 영역 이미지의 크기를 소정의 크기 $n \times n$ 즉, 512×512 , 256×256 , 128×128 혹은 다른 크기로 축소하여 제2 이미지를 생성할 수 있다.
- [0067] S360 단계에서 CNN은 제2 이미지로부터 얼굴에 대한 적어도 하나의 랜드마크를 마킹할 수 있다. 랜드마크

크는 예를 들어, 눈 또는 입의 모서리, 혹은 코의 끝 부분, 얼굴, 코, 눈 및 입의 윤곽을 밝히는 인위적인 포인트와 같은, 얼굴의 특징들을 의미한다. 그러나 랜드마크가 이에 한정되는 것은 아니다.

[0068] 일 실시예에서, 랜드마크들은 수동으로 마크될 수 있다. 다른 일 실시예에서 랜드마크들의 마킹은 액티브 셰이프 모델 알고리즘(active shape model algorithms), 캐스케이드 알고리즘(cascade algorithms) 혹은 신경망 네트워크 접근(neural network approach)등을 이용하여 자동으로 이루어질 수도 있다. 그러나 랜드마크 자동 마킹에 이용되는 알고리즘들이 위 예시에 한정되는 것은 아니다.

[0069] 다른 실시예에서 이 단계는 S320 단계 및/또는 S330단계와 결합될 수 있다. 얼굴 영역 감지 및 그에 대한 바운딩 박스 생성은 마킹된 얼굴 랜드마크로부터 암시될 수 있기 때문이다.

[0070] S370 단계에서 CNN은 적어도 하나의 랜드마크와 가장 근접하게 매치되는 적어도 하나의 카메라 파라미터 및 적어도 하나의 얼굴 모델 파라미터를 획득할 수 있다.

[0071] 이 단계에서 CNN은 비선형 최적화를 통하여 적어도 하나의 카메라 파라미터를 획득할 수 있다. 비선형 최적화에는 뉴턴-가우스(Newton-Gauss), 레벤베르크-마르콥르트(Levenberg-Marquardt) 알고리즘, 기울기 하강(gradient descent)이 이용될 수 있지만, 이에 한정되는 것은 아니다.

[0072] 또한, CNN은 선형 최적화를 통하여 적어도 하나의 얼굴 모델 파라미터를 획득할 수 있다. 적어도 하나의 얼굴 모델 파라미터는 모양 파라미터를 포함할 수 있다. CNN은 적어도 하나의 얼굴 모델 파라미터를 획득함에 있어서 선형 최적화를 이용함으로써 필요한 하드웨어 자원을 감소시킬 수 있다. 따라서, 본 개시의 일 실시예에 따른 2차원 이미지 처리 디바이스(100)는 추가적인 하드웨어의 도움 없이 모바일 디바이스에서 3차원 얼굴 모델을 생성할 수 있다. 이에 대한 상세한 내용은 도7에서 후술한다.

[0073] 다른 실시예에서 CNN은 얼굴 랜드마크들과 가장 근접하게 매치되는 3차원 얼굴 모델을 생성하는 단계를 더 포함할 수 있다.

[0074] 도4a는 본 개시의 일 실시예에 따라 트레이닝된 CNN 이 데이터를 처리하는 방법의 흐름도이다.

[0075] 도4a는 도3의 흐름도를 간소화하여 도시한 도면이고, 도4B는 본 개시의 일 실시예에 따라 도4a의 S420 단계에서 CNN의 구조를 도시한 도면이며, 도4C는 본 개시의 일 실시예에 따라 도4a의 S440 단계에서의 CNN 구조를 도시한 도면이다.

[0076] 도4a의 실시예에서 CNN은 복수의 2차원 이미지를 입력 받아 S410 단계, S420 단계, S430 단계를 실행할 수 있다. 이 경우, 복수의 이미지들 및 이들에 대한 바운딩 박스가 S440 단계에 입력될 수 있다.

[0077] 도4b의 실시예에서 CNN은 입력 레이어(450), 컨볼루션 레이어(460) 및 완전하게 연결된 레이어(470)를 포함할 수 있다.

[0078] CNN은 96*96 크기의 이미지를 입력 레이어(450)에 입력할 수 있다.

[0079] 컨볼루션 레이어(convolutional layers)(460)는 2차원 이미지의 특징을 도출하도록 디자인될 수 있다.

[0080] CNN은 스스로 새로운 레이어를 추가할 수 있다. 새로운 레이어가 추가되는 경우 CNN의 얼굴 감지 능력이 향상될 수 있으므로, 다른 추가적인 알고리즘 없이 얼굴을 감지하는 것이 가능할 수 있다.

[0081] 도4c에서 CNN은 입력 레이어(450), 컨볼루션 레이어(460) 및 완전하게 연결된 레이어(470)를 포함할 수 있다.

[0082] CNN은 128*128 크기의 이미지를 입력 레이어(450)에 입력할 수 있다.

[0083] 또한, CNN은 도4c에서 도시된 바와 같이 완전하게 연결된 레이어(470)의 끝에서 두 번째에 바운딩 박스를 추가로 입력할 수 있다. 입력된 바운딩 박스는 완전하게 연결된 레이어(470)의 끝에서 두 번째의 입력 데이터와 결합(concatenate)될 수 있다. 컨볼루션 레이어(convolutional layers)(460)는 2차원 이미지의 특징을 도출하도록 디자인될 수 있다. 그러나 바운딩 박스와 같은 거시 특징(macro-features)들과 작용하도록 디자인되지는 않을 수 있다.

[0084] 본 실시예에서 바운딩 박스들은 완전하게 연결된 레이어(470)에서만 입력되지만, 이에 한정되는 것은 아니다.

[0085] 다른 실시예에서, 완전하게 연결된 레이어(470)의 수가 증가되어 비선형성이 증가되면, 이 CNN 구조에

의하여 예정될 수 있는 기능의 다양성이 증가될 수 있다.

- [0086] 입력을 실행하기 위하여, 적어도 하나의 새로운 완전하게 연결된 레이어(470)가 CNN 구조에 추가되어 CNN에 바운딩 박스 파라미터들을 이용하는 능력을 제공할 수 있게 됨으로써, CNN은 카메라 파라미터들 $U=(tx, ty, tz, a, b, g, f)$ 에 대하여 더 정확한 추론을 할 수 있다.
- [0087] 디바이스(100)는 도4b 및 도4c와 같은 CNN을 이용하여 2차원 이미지로부터 얼굴의 위치, 얼굴의 방향 및 얼굴 랜드마크의 위치 등을 획득할 수 있다.
- [0088] 도5는 본 개시의 일 실시예에 따라 2차원 이미지를 처리하는 방법이 기존의 방법과 다른 점에 대한 예시이다.
- [0089] (a)는 기존의 방식에 따른 얼굴 정렬(face alignment) 최적화 알고리즘을 도시하고, (b)는 본 개시의 일 실시예에 따른 얼굴 정렬(face alignment) 최적화 알고리즘을 도시한다.
- [0090] (a)에서는 3차원 변형 가능 얼굴 모델(3D morphable face model)을 투사(projection)하여 생성된 2차원 모델과 2차원 이미지로부터 감지된 랜드마크들을 비교하는 것과 비교하여, (b)에서는 투사 과정 없이 2차원 이미지에서 직접 감지된 랜드마크들과 3차원 변형 가능 얼굴 모델(3D morphable face model)을 직접 비교할 수 있다.
- [0091] 디바이스(100)는 (b)와 같은 방법을 이용함으로써, 최적화 알고리즘의 수행 속도를 향상시킬 수 있다.
- [0092] 도6은 본 개시의 일 실시예에 따라 CNN이 트레이닝 되는 방법에 대한 흐름도이다.
- [0093] CNN은 다양한 방식으로 트레이닝될 수 있다.
- [0094] 일 실시예에서 CNN은 3차원 변형 가능 얼굴 모델, 일련의 이미지들 및 일련의 이미지들에 대응하는 바운딩 박스, 랜드마크와 가장 근접하게 매치되는 카메라 파라미터들 및 모델 파라미터들과 같은 트레이닝 데이터에 의하여 트레이닝 될 수 있다.
- [0095] 이 경우, CNN에 인위적인 2차원 이미지 입력 능력을 제공하여, 인위적으로 입력된 2차원 이미지로부터 적어도 하나의 카메라 파라미터 및 적어도 하나의 모델 파라미터를 실측 데이터로 추출할 수 있다(610).
- [0096] 또한, 디바이스(100)는 2차원 이미지에 대하여 도2 및 도3의 방법에 따른 CNN의 실행에 의하여 적어도 하나의 카메라 파라미터 및 적어도 하나의 모델 파라미터를 획득할 수 있다(630).
- [0097] 만일 실측데이터(620)와 CNN이 알고리즘에 의하여 획득된 데이터(630)의 차이가 크다면, 손실 함수(loss function)가 CNN 러닝의 단계에서 추가적으로 이용될 수 있다.
- [0098] 실측데이터와 CNN이 알고리즘에 의하여 획득된 데이터의 차이가 큰지 여부에 대한 판단은, CNN 트레이닝으로 입력된 몇몇 인위적인 얼굴 이미지에 대하여 각 얼굴 이미지의 랜드마크와 가장 근접하게 매치되는 적어도 하나의 카메라 파라미터 및 적어도 하나의 모델 파라미터에 대하여 생성된 3차원 얼굴 모델과, 각 이미지에 대하여 CNN에 의하여 획득된 적어도 하나의 카메라 파라미터 및 적어도 하나의 모델 파라미터를 기초로 생성된 3차원 얼굴 모델을 L2 표준의 차이를 이용하여 비교함으로써 이루어질 수 있다.
- [0099] 다른 실시예에서, CNN은 손실 함수로서의 얼굴 랜드마크의 투영들 사이의 차이에 대한 표준을 이용하여 트레이닝될 수 있다.
- [0100] 또 다른 실시예에서 CNN은 아담 최적화 알고리즘(Adam optimization algorithm)을 이용하여 추가적으로 트레이닝될 수 있다.
- [0101] 또는, CNN이 카메라 파라미터들에서의 차이에 대한 L2 표준 및 모델 파라미터들에서의 차이에 대한 L2 표준을 이용하여 트레이닝될 수도 있다. 이 때, L2 표준은 티칭(teaching)의 매크로 파라미터인 가중치 차이로 추가될 수 있다.
- [0102] 도7은 본 개시의 일 실시예에 따라 감정 파라미터가 추가로 획득되는 경우에 대한 흐름도이다.
- [0103] S710 단계에서 디바이스(100)는 비선형 최적화를 이용하여 적어도 하나의 카메라 파라미터를 획득할 수 있다.
- [0104] S730 단계에서 디바이스(100)는 선형 최적화를 이용하여 적어도 하나의 얼굴 모델 파라미터를 획득할 수 있다.

[0105] 모델 파라미터들을 획득하기 위한 선형 최적화 문제는 아래의 함수와 같이 수식(1)로 도식화될 수 있다.

[0106]

$$F(\hat{\phi}_i(\epsilon)) = \sum_{i=0}^n \hat{\phi}_i(\epsilon) \rightarrow \min_{\epsilon}$$

[0107]

$$\hat{\phi}_i(\epsilon) = \left\| f \cdot \left((RM_i(\epsilon))_{x,y} + t_{x,y} \right) - p_i \cdot \left((RM_i(\epsilon))_z + t_z \right) \right\|_2^2,$$

[0108] 여기서 $\hat{\phi}_i(\epsilon)$ 는 L2 표준이고, $p_i = (p_x, p_y)^T$ 는 $M=N+B^i$ 식에 의하여 설명되는 모델의 I 번째 포인트이며, 이 때 N은 중립적인 3차원 얼굴 모델이고, B는 기초 모양(shape basis)이며, i 는 얼굴 모양 파라미터이고, $\epsilon \in \mathbb{R}^{55}$ 는 카메라 파라미터들일 수 있다.

[0109] 또한, R은 파라미터 (a, b, g)를 가지는 회전 행렬이고, $t_{x,y,z}$ 는 변형(translation)이며, f는 초점 거리일 수 있다.

[0110] 전통적인 비선형 최적화 문제 대신에 모델 파라미터 검색을 위한 선형 최적화 문제에 대한 해결책은 특히 기존의 방법을 실행하는데 필요한 시스템 자원의 사양에 대한 요구를 감소시킬 수 있다.

[0111] 모델 파라미터들은 모델을 생성하는 데 사용된 선형 조합 계수 ϵ 이다. 이 문제의 해결을 위하여, 카메라 파라미터들은 상수로 간주된다.

[0112] 최적화 문제를 해결하기 위하여, 수식(1)을 이용하는 선형 등식의 시스템 구성 및 최소 제곱법(least squares)로 문제를 푸는 것이 필수이다. 선형 등식의 시스템에 대한 해결책은 최적화 문제에 대한 해답이 될 것이다.

[0113] 일 실시예에서, 변형 가능 얼굴 모델의 바람직한 파라미터들에 대한 규칙화(regularization)가 더 수행될 수 있다. 규칙화는 다음의 공식에 따라 수행될 수 있다.

[0114]

$$\text{reg}(\epsilon) = \frac{1}{2} \sum_{i=0}^n \frac{\epsilon_i}{\text{eigen}_i^2}$$

[0115] 여기서 ϵ_i 는 최적화될 i 번째 파라미터이고, eigen_i 는 i 번째 베이스 벡터의 고유치이다. 그럼 완성된 최적화 공식(수식 (1))은 다음과 같을 수 있다.

[0116]

$$F(\hat{\phi}(\epsilon)) = \sum_{i=0}^n \hat{\phi}_i(\epsilon) + \text{reg}(\epsilon)$$

[0117] 다른 일 실시예에서, 비디오 스트림 및 변형 가능 얼굴 모델의 파라미터들을 처리하는 기존의 방법은 하나의 이미지 보다는 복수개의 이미지들에서 발견될 수 있다.

[0118] 이 경우, 수식 (1)은 아래와 같이 쓰여질 수 있다.

[0119]

$$\hat{F}(\phi(\epsilon)) = \sum_{k=0}^m \sum_{i=0}^n \hat{\phi}^{(k)}_i(\epsilon) + \text{reg}(\epsilon),$$

여기서, k는 이미지 번호일 수 있다.

[0120] S750 단계에서 디바이스(100)는 선형 최적화를 이용하여 적어도 하나의 감정 파라미터를 획득할 수 있다.

[0121] 3차원 얼굴 모델을 생성하는 단계에서 랜드마크들과 가장 근접하게 매치하는 카메라 파라미터들, 변형 가능 얼굴 모델 파라미터들 및 감정 파라미터들은 아래의 수식 (2)에 따라 선형 최적화 문제를 해결함으로써 이 단계를 수행할 수 있다.

$$F(\phi(\epsilon, \alpha)) = \sum_{i=0}^n \phi_i(\epsilon, \alpha) \rightarrow \min_{\alpha}$$

$$\hat{\phi}_i(\epsilon, \alpha) = \|f \cdot ((R\hat{M}_i(\epsilon, \alpha))_{x,y} + t_{x,y}) - p_i \cdot ((R\hat{M}_i(\epsilon, \alpha))_z + t_z)\|_2^2$$

여기서 $\hat{\phi}_i(\epsilon, \alpha)$ 는 L2 표준이고, $p_i = (p_x, p_y)^T$ 는 i번째 랜드마크이며, $M_i(\epsilon, \alpha) = (M_x, M_y, M_z)^T$ 는

공식 $\hat{M} = N + B + Aa$ (N은 중립적인 3차원 얼굴 모델, B는 얼굴 모형, ϵ 는 얼굴 모형 파라미터들, $\epsilon \in \mathbb{R}^{55}$, A는 감정 기초(emotion basis), a는 감정 파라미터에 대한 벡터)에 의하여 설명되는 감정 반영 모델의 i번째 포인트일 수 있다.

모델 파라미터들은 감정 기초 벡터들의 선형 조합에 대한 계수 ϵ 이고, 모델을 생성하는데 이용된다. 이러한 문제 해결을 위하여, 카메라 파라미터들 및 얼굴 모형 파라미터들은 상수로 간주된다.

최적화 문제 해결을 위하여, 수식(2)를 기초로 선형 등식의 시스템 구성 및 시스템 해결이 필수적이다. 선형 등식의 시스템에 대한 해결책은 최적화 문제에 대한 해답이 될 것이다.

도8은 본 개시의 일 실시예에 따라 2차원 이미지를 기초로 3차원 아바타를 생성하는 예시이다.

디바이스(100)는 2차원 이미지를 획득하고(810), 획득된 2차원 이미지로부터 랜드마크를 감지할 수 있다(820). 디바이스(100)는 도2 및 도3에 개시된 방법에 따라 3차원 얼굴 모델을 생성할 수 있다(830).

나아가 디바이스(100)는 생성된 3차원 얼굴 모델을 기초로 2차원 이미지에서 감지된 얼굴과 유사한 3차원 아바타를 생성할 수 있다(840).

일부 실시예에서 생성된 3차원 아바타는 사용자의 습관, 움직임 등을 흉내낼 수 있다. 일부 실시예에서 디바이스(100)는 생성된 3차원 아바타가 사용자의 움직임을 감지하고, 감지된 사용자의 움직임을 반영하여 유사하게 움직이도록 제어할 수 있다.

다른 실시예에서, 디바이스(100)는 사용자의 습관, 움직임, 표정 등에 대한 데이터를 보안 및 인증 시스템에서 이용할 수 있다. 예를 들어, 보안 시스템은 사용자 고유의 습관, 움직임, 표정 등에 대한 데이터를 출입 문 보안, 문서 보안, 모바일 디바이스 보안 등에 이용할 수 있다.

도9는 본 개시의 일 실시예에 따라 2차원 이미지 처리 디바이스가 행인을 촬영한 이미지로부터 감정 파라미터를 획득하여, 적절한 광고를 제공하는 예시이다.

디바이스(100)는 내장된 촬영부 혹은 외부의 촬영 디바이스를 통하여 사용자를 촬영한 이미지를 획득할 수 있다.

일 실시예에서, 디바이스(100)는 획득된 2차원 이미지에 대하여 도2 등의 실시예에 따른 방법을 수행하여 촬영된 사용자의 얼굴(910)에 대한 추가 정보를 획득하고, 획득된 추가 정보에 따라 상품 또는 서비스에 대한 광고를 제공할 수 있다.

예를 들어, 디바이스(100)는 획득된 2차원 이미지에 대하여 도2 및 도7의 실시예에 따른 방법을 수행하여 촬영된 사용자에게 대한 감정 파라미터를 획득할 수 있다.

디바이스(100)는 획득된 감정 파라미터를 이용하여 사용자의 기분을 결정하고, 결정된 사용자의 기분에 따라 상품 또는 서비스에 대한 광고를 제공할 수 있다.

다른 예를 들어, 디바이스(100)는 획득된 2차원 이미지에 대하여 도2 및 도7의 실시예에 따른 방법을 수행하여 촬영된 사용자의 대략적인 나이, 피부색, 피부 톤, 생김새, 얼굴 크기 등을 추가 정보로 획득할 수 있다.

디바이스(100)는 획득된 추가 정보에 따라 사용자 맞춤형 상품 또는 서비스에 대한 광고를 제공할 수 있다.

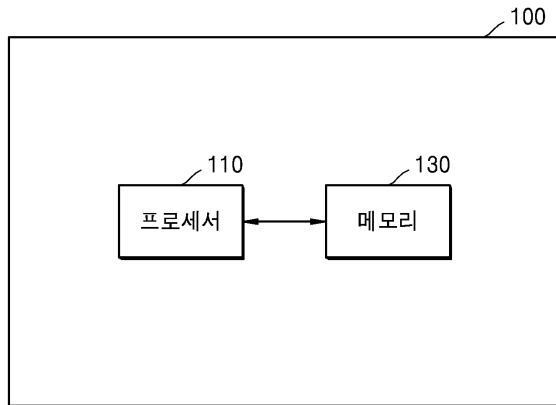
광고의 제공은 디바이스(100) 내부의 디스플레이를 이용하여 이루어지거나, 디바이스(100) 외부의 디스플레이를 이용하여 이루어질 수 있다.

- [0140] 도10은 본 개시의 일 실시예에 따라 애완로봇이 사용자를 촬영한 이미지로부터 감정 파라미터를 획득하여 사용자와 교감하는 예시이다.
- [0141] 다른 실시예에서 디바이스(100)는 애완 로봇(1010)뿐 아니라, 무선 청소기, 드론, 전자 비서 등의 디바이스 또는 이들 디바이스와 연결된 별도의 디바이스일 수 있다.
- [0142] 디바이스(100)는 내장된 촬영부 혹은 외부의 촬영 디바이스를 통하여 사용자를 촬영한 이미지를 획득할 수 있다.
- [0143] 일 실시예에서, 디바이스(100)는 획득된 2차원 이미지에 대하여 도2 및 도7의 실시예에 따른 방법을 수행하여 촬영된 사용자에 대한 감정 파라미터를 획득할 수 있다.
- [0144] 디바이스(100)는 획득된 감정 파라미터를 이용하여 사용자의 기분을 결정하고, 결정된 사용자의 기분에 따라 소정의 기능을 수행함으로써 사용자와 교감할 수 있다.
- [0145] 예를 들어, 사용자의 기분이 좋지 않은 경우, 애완 로봇(1010)은 위로하는 제스처를 취하도록 제어될 수 있다.
- [0146] 도11은 본 개시의 일 실시예에 따라 생성된 3차원 얼굴 모델을 기초로 원근 왜곡을 교정하는 예시이다.
- [0147] 본 실시예에서, 디바이스(100)는 획득된 2차원 이미지(1110)에 대하여 도2 및 도3의 실시예에 따른 방법을 수행하여 촬영된 사용자에 대한 3차원 얼굴 모델을 생성할 수 있다.
- [0148] 디바이스(100)는 생성된 3차원 얼굴 모델을 2차원 이미지의 얼굴 영역에 오버레이함으로써, 기존의 2차원 이미지(1110)에 존재하였던 원근 왜곡(perspective distortion)이 보정된 2차원 이미지(1130)를 획득할 수 있다.
- [0149] 2차원 이미지의 얼굴 영역에 대한 3차원 얼굴 모델의 오버레이는 핀 홀 카메라 모델을 이용하여 이루어질 수 있다.
- [0150] 일 실시예에서 원근 왜곡의 보정은 뷰티 애플리케이션, 디지털 거울 및 디지털 메이크업 등에서 이용될 수 있다.
- [0151] 본 개시의 방법은 프로세서, ASIC(application specific integrated circuit), FPGA(field programmable gate array) 또는 SoC(system-on-chip)에 의하여 실행될 수 있다. 또한, 기재된 방법은 컴퓨터로 실행할 수 있는 인스트럭션을 저장하고, 컴퓨터에 프로세서에 의하여 실행될 때 본 개시의 방법을 실행하도록 할 수 있는 저장매체에 의하여 실행될 수 있다.
- [0152] 기기로 읽을 수 있는 저장매체는, 비일시적(non-transitory) 저장매체의 형태로 제공될 수 있다. 여기서, '비일시적 저장매체'는 실재(tangible)하는 장치이고, 신호(signal)(예: 전자기파)를 포함하지 않는다는 것을 의미할 뿐이며, 이 용어는 데이터가 저장매체에 반영구적으로 저장되는 경우와 임시적으로 저장되는 경우를 구분하지 않는다. 예로, '비일시적 저장매체'는 데이터가 임시적으로 저장되는 버퍼를 포함할 수 있다.
- [0153] 일 실시예에 따르면, 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory (CD-ROM))의 형태로 배포되거나, 또는 어플리케이션 스토어(예: 플레이 스토어™)를 통해 또는 두개의 사용자 장치들(예: 스마트폰들) 간에 직접, 온라인으로 배포(예: 다운로드 또는 업로드)될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품(예: 다운로드할 앱(downloadable app))의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 중계 서버의 메모리와 같은 기기로 읽을 수 있는 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.
- [0154] 전술한 본 개시의 설명은 예시를 위한 것이며, 본 개시가 속하는 기술분야의 통상의 지식을 가진 자는 본 개시의 기술적 사상이나 필수적인 특징을 변경하지 않고서 다른 구체적인 형태로 쉽게 변형이 가능하다는 것을 이해할 수 있을 것이다. 그러므로 이상에서 기술한 실시예들은 모든 면에서 예시적인 것이며 한정적이 아닌 것으로 이해해야만 한다. 예를 들어, 단일 형으로 설명되어 있는 각 구성 요소는 분산되어 실시될 수도 있으며, 마찬가지로 분산된 것으로 설명되어 있는 구성 요소들도 결합된 형태로 실시될 수 있다.
- [0155] 본 개시의 범위는 위 상세한 설명보다는 후술하는 특허청구범위에 의하여 나타내어지며, 특허청구범위의

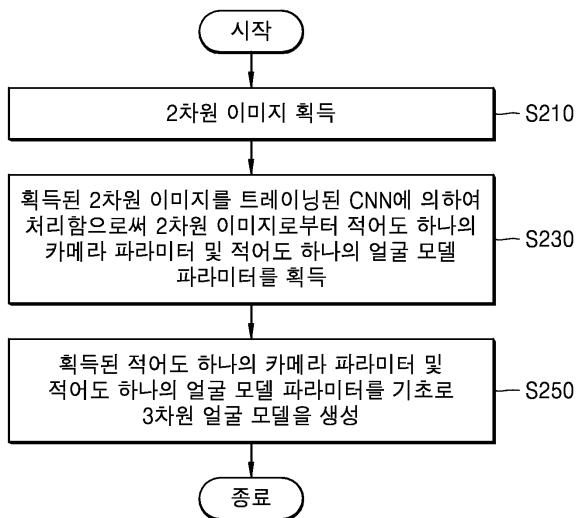
의미 및 범위 그리고 그 균등 개념으로부터 도출되는 모든 변경 또는 변형된 형태가 본 개시의 범위에 포함되는 것으로 해석되어야 한다.

도면

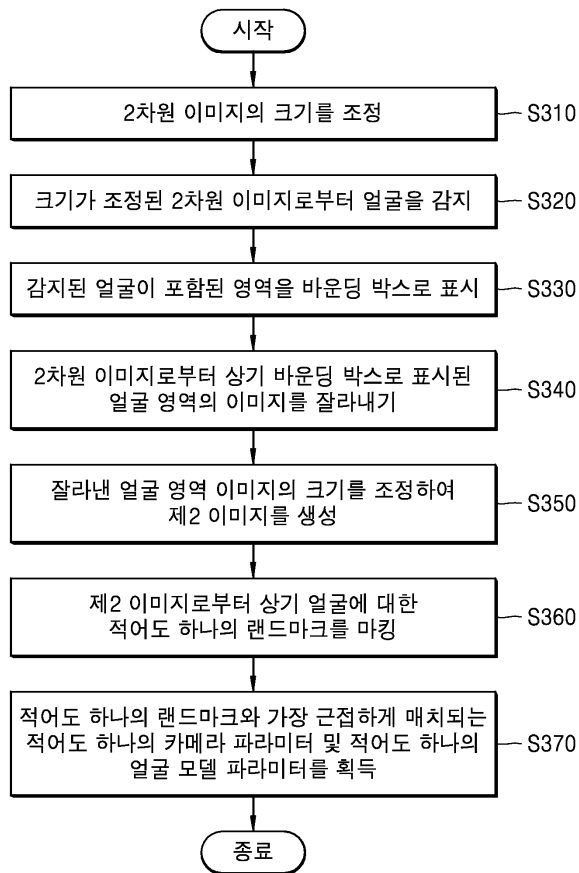
도면1



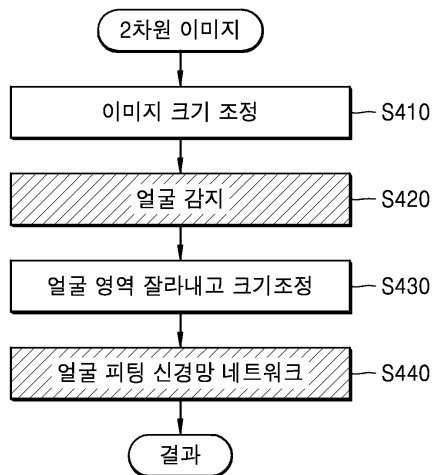
도면2



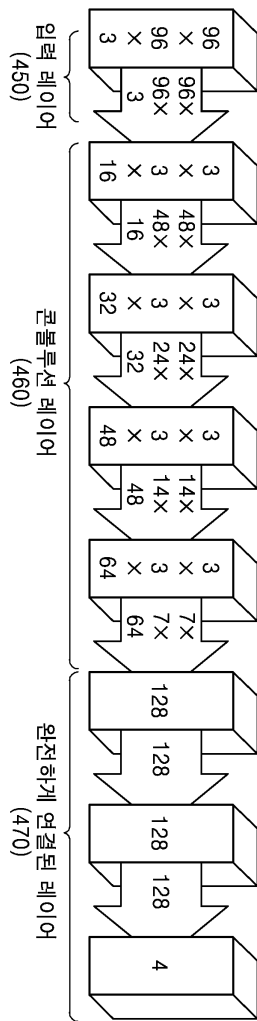
도면3



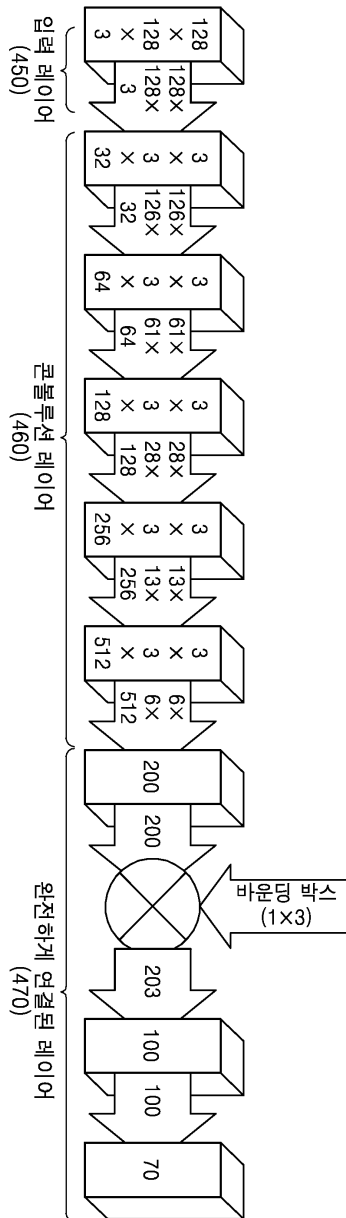
도면4a



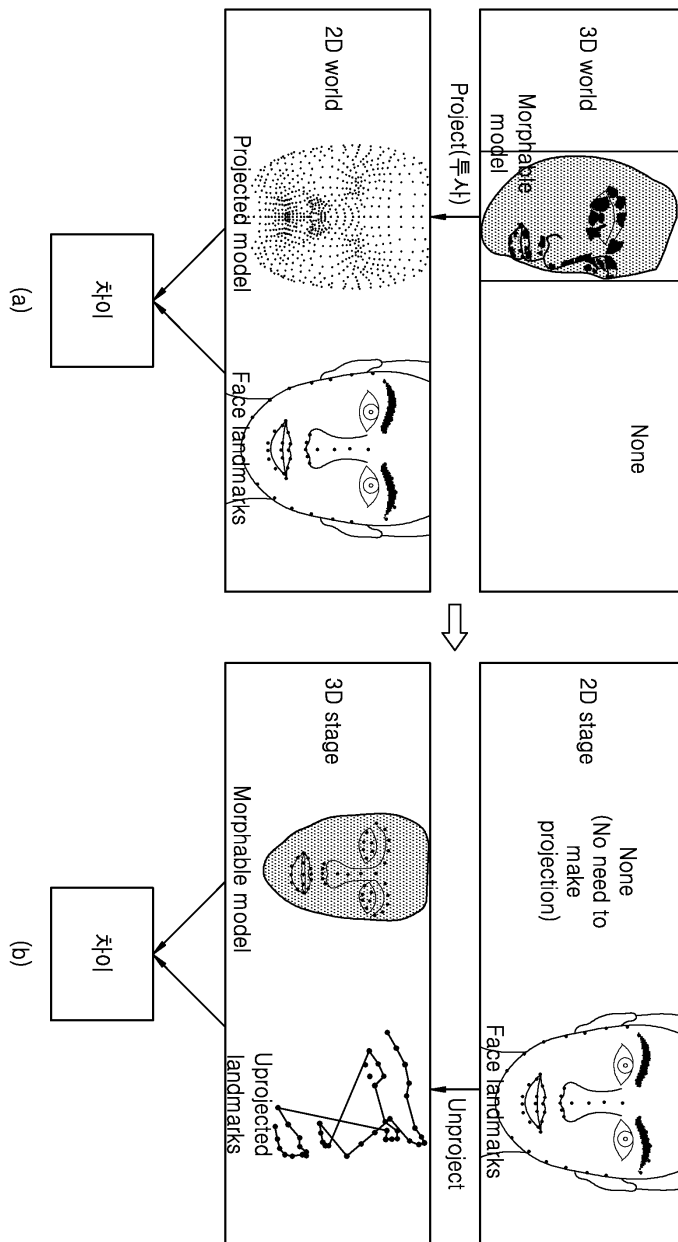
도면4b



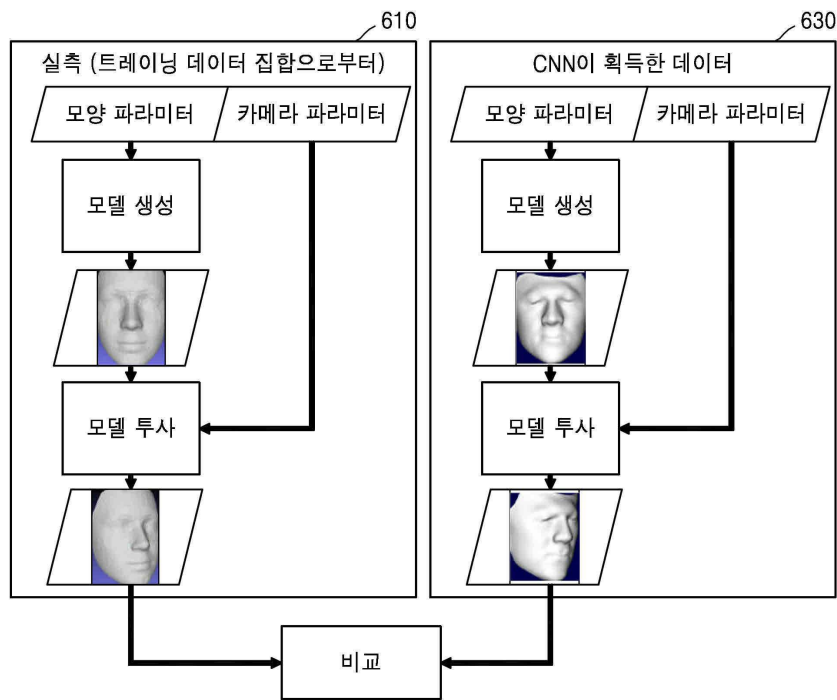
도면4c



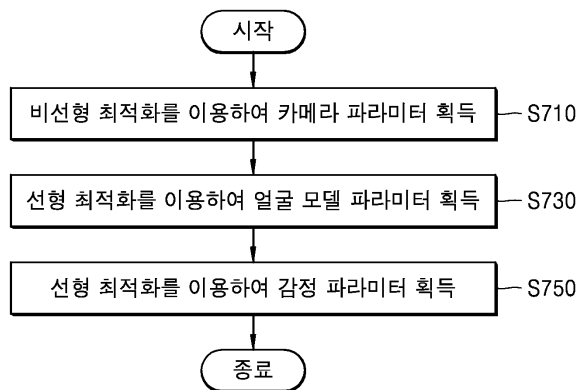
도면5



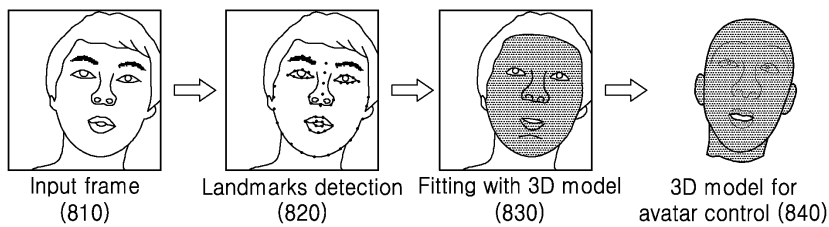
도면6



도면7



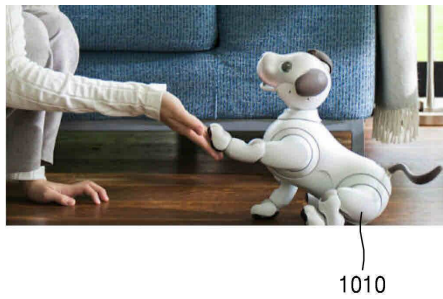
도면8



도면9



도면10



도면11

