



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0092533
 (43) 공개일자 2017년08월11일

(51) 국제특허분류(Int. Cl.)
 G06K 9/00 (2006.01) G06T 7/50 (2017.01)
 G06T 7/70 (2017.01)
 (52) CPC특허분류
 G06K 9/00221 (2013.01)
 G06K 9/00214 (2013.01)
 (21) 출원번호 10-2017-7010873
 (22) 출원일자(국제) 2014년09월23일
 심사청구일자 없음
 (85) 번역문제출일자 2017년04월21일
 (86) 국제출원번호 PCT/EP2014/070282
 (87) 국제공개번호 WO 2016/045711
 국제공개일자 2016년03월31일

(71) 출원인
 키레몬 에스에이
 스위스 1920 마르티그니 뤼 뒤 칼리지 1
 (72) 발명자
 로드리게즈, 안
 스위스 슈바이츠-1941 볼레게스 에르테 데' 티츠 61
 모울린, 프란시스
 스위스 슈바이츠-1941 볼레게스 캄파보이 13
 피칸트, 세바스티앙
 스위스 슈바이츠-1920 마르티그니 뤼 데 라 드란 세 13
 (74) 대리인
 장훈

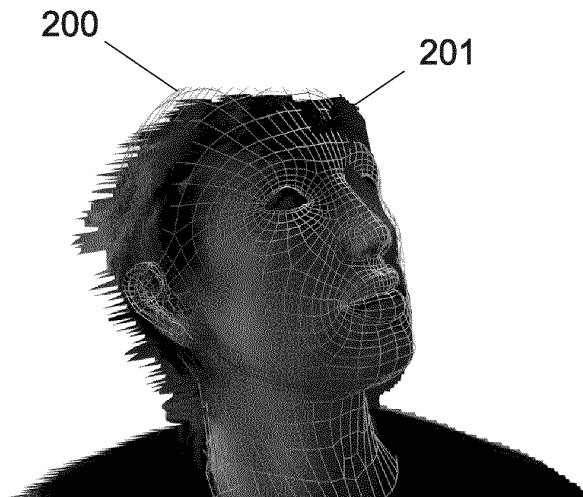
전체 청구항 수 : 총 16 항

(54) 발명의 명칭 **얼굴 포즈 교정 방법 및 장치**

(57) 요약

얼굴(100) 이미지들을 표현하는 데이터에서 포즈(pose)를 교정하기 위한 포즈 교정 방법은: A-2D 근적외선 이미지 데이터, 2D 가시광선 이미지 데이터, 및 깊이 맵을 포함하는 적어도 하나의 테스트 프레임을 획득하는 단계; C-공지된 방향의 머리의 3D 모델에 따라 상기 깊이 맵을 정렬함으로써 상기 테스트 프레임에서 얼굴의 포즈를 추정하는 단계; D-텍스처(texture)된 이미지 데이터를 생성하기 위해, 상기 2D 이미지의 적어도 하나를 깊이 맵 상에 매핑하는 단계; E-포즈 교정된 2D 투영된 이미지를 표현하는 데이터를 생성하기 위해 텍스처된 이미지 데이터를 2D로 투영하는 단계를 포함한다.

대표도 - 도8



(52) CPC특허분류

G06T 7/50 (2017.01)

G06T 7/70 (2017.01)

G06T 2207/10024 (2013.01)

G06T 2207/10028 (2013.01)

G06T 2207/10048 (2013.01)

G06T 2207/20068 (2013.01)

G06T 2207/20182 (2013.01)

G06T 2207/30201 (2013.01)

명세서

청구범위

청구항 1

얼굴(100) 이미지들을 표현하는 데이터에서 포즈(pose)를 교정하기 위한 포즈 교정 방법에 있어서:

A-2D 근적외선 이미지 데이터, 2D 가시광선 이미지 데이터, 및 깊이 맵을 포함하는 적어도 하나의 테스트 프레임임을 획득하는 단계;

C-공지된 방향의 머리의 3D 모델에 따라 상기 깊이 맵을 정렬함으로써 상기 테스트 프레임에서 얼굴의 포즈를 추정하는 단계;

D-텍스처(texture)된 이미지 데이터를 생성하기 위해, 상기 2D 이미지의 적어도 하나를 상기 깊이 맵 상에 매핑하는 단계;

E-포즈 교정된 2D 투영된 이미지를 표현하는 데이터를 생성하기 위해 상기 텍스처된 이미지 데이터를 2D로 투영하는 단계를 포함하는, 포즈 교정 방법.

청구항 2

제 1 항에 있어서,

상기 깊이 맵에서 포인트들의 시간적 및/또는 공간적 스무딩(smoothing) 단계를 포함하는, 포즈 교정 방법.

청구항 3

제 1 항 또는 제 2 항에 있어서,

상기 포즈를 추정하는 상기 단계(C)는 예를 들면, 랜덤 포레스트(random forest)에 기초하여 개략적인 포즈 추정을 수행하는 제 1 단계, 및 상기 포즈의 더 정확한 추정을 결정하는 또 다른 단계를 포함하는, 포즈 교정 방법.

청구항 4

제 1 항 내지 제 3 항 중 어느 한 항에 있어서,

공지된 방향의 머리의 3D 모델에 따라 상기 깊이 맵을 정렬하는 상기 단계(C)는 반복 최근접점(Iterative Closest Points; ICP) 방법을 이용하는, 포즈 교정 방법.

청구항 5

제 1 항 내지 제 4 항 중 어느 한 항에 있어서,

상기 얼굴에 속하지 않는 상기 2D 근적외선 이미지 데이터의, 및/또는 상기 2D 가시광선 이미지 데이터의, 및/또는 상기 깊이 맵의 적어도 일부 부분들을 제거하기 위해, 상기 포즈의 상기 추정(C) 전에 기본 얼굴 검출 단계(B)를 더 포함하는, 포즈 교정 방법.

청구항 6

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 3D 모델은 사용자 독립적인, 포즈 교정 방법.

청구항 7

제 1 항 내지 제 5 항 중 어느 한 항에 있어서,

상기 3D 모델은 사용자 의존적인, 포즈 교정 방법.

청구항 8

제 1 항 내지 제 7 항 중 어느 한 항에 있어서,
상기 3D 모델은 상기 3D 모델을 상기 이용자에 적응시키기 위해 휘어지는(warp), 포즈 교정 방법.

청구항 9

제 1 항 내지 제 8 항 중 어느 한 항에 있어서,
머리의 기존의 3D 모델에 따라 상기 깊이 맵을 정렬하는 상기 단계(C)는 상기 3D 모델을 휘게하는 단계를 포함하는, 포즈 교정 방법.

청구항 10

제 1 항 내지 제 9 항 중 어느 한 항에 있어서,
상기 2D 근적외선 이미지 데이터에 기초하여 상기 2D 가시광선 이미지 데이터의 부분들의 조명을 정정하는 단계를 더 포함하는, 포즈 교정 방법.

청구항 11

제 1 항 내지 제 10 항 중 어느 한 항에 있어서,
상기 깊이 맵 상에서 보이지 않는 부분들에 대응하는 상기 포즈 교정된 2D 투영된 이미지 데이터의 부분들을 표시하는(flagging) 단계를 더 포함하는, 포즈 교정 방법.

청구항 12

제 1 항 내지 제 11 항 중 어느 한 항에 있어서,
상기 깊이 맵의 공지되지 않은 부분들에 대응하는 상기 포즈 교정된 2D 투영된 이미지 데이터의 부분들을 재구성하는 단계를 더 포함하는, 포즈 교정 방법.

청구항 13

제 1 항 내지 제 12 항 중 어느 한 항에 있어서,
상기 2D 투영된 이미지를 분류하는 단계(F)를 더 포함하는, 포즈 교정 방법.

청구항 14

장치로서, 2D 근적외선 이미지 데이터, 2D 가시광선 이미지 데이터, 및 깊이 맵을 포함하는 적어도 하나의 테스트 프레임을 획득하기 위해 배열된 깊이 맵 카메라(101) 뿐만 아니라, 프로그램이 실행될 때 상기 프로세서로 하여금 다음의 단계들을 실행하게 하는 상기 프로그램을 저장하는 메모리를 갖는 프로세서를 포함하는, 상기 장치에 있어서:

상기 단계들은,

C-공지된 방향의 머리의 3D 모델에 따라 상기 깊이 맵을 정렬함으로써 상기 테스트 프레임에서 얼굴의 포즈를 추정하는 단계;

D-텍스처된 이미지 데이터를 생성하기 위해, 상기 2D 이미지의 적어도 하나를 상기 깊이 맵 상에 매핑하는 단계;

E-포즈 교정된 2D 투영된 이미지를 표현하는 데이터를 생성하기 위해 상기 텍스처된 이미지 데이터를 2D로 투영하는 단계인, 장치,

청구항 15

제 1 항 내지 제 14 항 중 어느 한 항의 방법을 수행하기 위한 수단을 포함하는 장치.

청구항 16

컴퓨터 프로그램 제품에 있어서:

A-2D 근적외선 이미지 데이터, 2D 가시광선 이미지 데이터, 및 깊이 맵을 포함하는 적어도 하나의 테스트 프레임 임을 획득하고;

C-공지된 방향의 머리의 3D 모델에 따라 상기 깊이 맵을 정렬함으로써 상기 테스트 프레임에서 얼굴의 포즈를 추정하고;

D-텍스처된 이미지 데이터를 생성하기 위해, 상기 2D 이미지의 적어도 하나를 상기 깊이 맵 상에 매핑하며;

E-포즈 교정된 2D 투영된 이미지를 표현하는 데이터를 생성하기 위해 상기 텍스처된 이미지 데이터를 2D로 투영하도록 실행가능한 지시들을 포함하는 컴퓨터 판독가능한 매체를 포함하는, 컴퓨터 프로그램 제품.

발명의 설명

기술 분야

[0001] 본 발명은 얼굴 포즈(pose) 교정 방법 및 장치에 관한 것이다.

배경 기술

[0002] 얼굴 인식은 인간 얼굴의 테스트 이미지(또는 테스트 이미지들의 세트)를 표현하는 데이터의 분석 및 기준 이미지들의 데이터베이스와의 그것의 비교를 수반한다. 테스트 이미지는 일반적으로, 공통 2D 카메라로 캡처(capture)된 2D 이미지인 반면에, 기준 이미지는 예를 들면, 깊이 카메라로 캡처된 2D 이미지 또는 때때로 3D 이미지이다.

[0003] 얼굴 인식을 위한 이전 접근법들은 일반적으로, 테스트 이미지 및 기준 이미지들이 전부 완전한 정면 뷰로 캡처된다고 가정한다. 따라서, 정면 테스트 이미지들을 분류하고 그들을 대응하는 정면 기준 이미지들과 매칭(matching)하기 위한 다양한 수행 알고리즘들이 개발되었다.

[0004] 인식의 신뢰성은 그러나, 테스트 이미지가 정면이 아닌 시점으로부터 캡처되면 빠르게 감소한다. 이 문제점을 해결하기 위해, 포즈를 교정하는 것 즉, 정면이 아닌 테스트 이미지로부터 합성 정면 테스트 이미지를 생성하기 위해 테스트 이미지를 휘게하는 것(warp)이 종래 기술에서 이미 제안되었다.

[0005] 따라서, 머리 포즈를 추정하고 그 다음, 정정하기 위한 빠르고 강건한(robust) 방법들은 얼굴 인식, 또는 더 일반적으로 얼굴 분류 및 얼굴 프로세싱과 같은 많은 적용들에서 필수적이다.

[0006] US8055028은 정면이 아닌 2D 얼굴 이미지를 정면 얼굴 이미지로 정규화하는 방법에 관한 것이다. 방법은: 객체의 정면이 아닌 이미지의 포즈를 결정하는 단계; 객체의 정면이 아닌 이미지에 관해 스무딩 변환(smoothing transformation)을 수행하고, 그에 의해 스무딩된 객체 이미지를 생성하는 단계; 및 포즈 결정 결과 및 스무딩된 객체 이미지를 이용함으로써 객체의 정면 이미지를 합성하는 단계를 포함한다. 포즈는 정면이 아닌 이미지의 중심 라인의 양쪽 측면에 존재하는 객체 피쳐 포인트(object feature point)들 사이의 제 1 평균 거리를 얻음으로써 결정된다.

[0007] US8199979는 2D 이미지 획득 디바이스로 획득되는 얼굴 영역들을 포함하는 이미지들을 분류하고 보관하기 위한 방법을 개시한다. 얼굴 검출 모듈은 얼굴 영역에 대응하는 픽셀들의 그룹을 식별한다. 얼굴 방향(face orientation) 및 포즈가 그 다음 결정된다. 절반의 옆모습(half-profile) 얼굴의 경우에, 정규화 모듈은 후보 얼굴 영역을 3D 공간으로 변환하고, 그 다음 그것을 적절한 포즈 정정이 행해질 때까지 회전한다. 대안적으로, 2D 얼굴 영역의 텍스처(texture), 컬러 및 피쳐 영역들은 그 다음, 포즈를 정정하기 위해 회전되는 3D 모델 상에 매핑된다. 이것은 그러나, 2D 이미지가 상이한 얼굴의 3D 모델 상에 매핑될 때, 심각한 변형을 야기한다.

[0008] US7929775는 2D 이미지를 3D 클래스 모델(class model)과 매칭하기 위한 방법을 개시한다. 방법은 2D 이미지에서 이미지 피쳐들을 식별하는 단계; 클래스 모델과 이미지 사이의 정렬 변형을 컴퓨팅(computing)하는 단계; 및 정렬 변형 하에서, 클래스 모델의 클래스 부분들을 이미지 피쳐들과 비교하는 단계를 포함한다.

[0009] US7289648은 단일 이미지로부터 얼굴과 같은, 3차원 객체를 자동으로 모델링하기 위한 방법을 개시한다. 본 발명에 따른 시스템 및 방법은 단일 이미지를 이용하여 하나 이상의 3차원(3D) 얼굴 모델들을 구성한다. 그것은 또한, 대부분의 얼굴 인식 시스템들을 트레이닝(training)하기 위해 필요한 다양한 포즈들을 갖는 얼굴들의 데이터베이스를 생성하기 위한 툴(tool)로서 이용될 수 있다.

- [0010] US2013156262는 쌍 피쳐들의 세트를 기하학적 구성요소(geometric primitive)들의 쌍들로서 정의함으로써 객체의 포즈의 추정을 제안하고, 기하학적 구성요소들은 지향된 표면 포인트들, 지향된 경계 포인트들, 및 경계 라인 세그먼트들을 포함한다. 모델 쌍 피쳐들은 객체의 모델에 대한 쌍 피쳐들의 세트에 기초하여 결정된다. 장면 쌍 피쳐들은 3D 센서에 의해 획득된 데이터로부터의 쌍 피쳐들의 세트에 기초하여 결정되고, 그 다음 모델 쌍 피쳐들은 객체의 포즈를 추정하기 위해 장면 쌍 피쳐들과 매칭된다.
- [0011] 2D 이미지로부터의 포즈 교정은 그러나, 어려운 업무이고 교정의 품질은 일반적으로 강건하지 않은데, 이는 그것이 폐색(occlusion)들, 조명, 등에 관해 성취될 진폭 정정에 의존하기 때문이다.
- [0012] 2D 데이터에 기초하여 머리 추정 포즈의 내재하는 문제점들 중 일부를 극복하기 위해, 점점 더 이용가능하고 입수가 가능해지고 있는, 깊이 센서들에 의해 제공된 부가적인 깊이 정보를 이용하는 것이 이미 제안되어 왔다.
- [0013] US8660306은 깊이 데이터로부터 결정된 인간 포즈의 정정을 위한 방법을 개시한다. 방법은 깊이 이미지 데이터를 수신하는 단계, 깊이 이미지 데이터로부터 연계된 객체의 초기 추정된 골격을 얻는 단계, 랜덤 포레스트 부분공간 회귀 함수(random forest subspace regression function)(복수의 랜덤 분할/투영 결정 트리들을 활용하는 함수)를 초기 추정된 골격에 적용하는 단계, 및 랜덤 포레스트 부분공간 회귀를 초기 추정된 골격에 적용하는 결과에 기초하여 포즈의 표현을 결정하는 단계를 포함한다. 방법은 특히, 전체 신체의 포즈의 측정에 적응된다.
- [0014] US6381346은 얼굴 이미지들을 생성하고, 복합 코드들에 의해 그들 이미지들을 인덱싱(indexing)하며, 유사한 2차원 얼굴 이미지들을 검색하기 위한 시스템을 개시한다. 인간 얼굴들의 3D 이미지들은 3D 얼굴 피쳐 표면 형상들의 데이터 저장소로부터 생성된다. 이들 형상들은 얼굴 피쳐 부분들에 의해 조직화된다. 각각의 얼굴 부분에 대한 형상을 조립함으로써, 3D 얼굴 이미지가 형성된다.
- [0015] US8406484는 얼굴 인식 장치에 관한 것이고, 상기 얼굴 인식 장치는: 대상의 2차원 이미지 정보를 획득하기 위한 2차원 정보 획득 유닛; 대상의 3차원 이미지 정보를 획득하기 위한 3차원 정보 획득 유닛; 이용자의 3차원 얼굴 정보 및 이용자의 2차원 얼굴 정보에 대응하는 타원형 모델을 저장하기 위한 이용자 정보 데이터베이스; 및 대상의 2차원 이미지 정보를 이용하여 얼굴 인식을 수행하고, 인식된 얼굴이 이용자의 얼굴인지의 여부를 결정하고, 이용자의 타원형 모델을 3차원 이미지 정보와 매칭하고, 인식된 얼굴이 이용자의 얼굴이라고 결정할 시에 에러를 산출하며, 이용자의 얼굴이 에러에 기초하여 적절하지 않게 이용되는지의 여부를 결정하기 위한 제어 유닛을 포함한다.
- [0016] US7756325는 단일 사진으로부터 검색된 정보에 기초하여, 인간 얼굴과 같은, 3차원 객체의 3D 형상을 추정하기 위한 알고리즘을 개시한다. 픽셀 세기에 더하여, 본 발명은 멀티 피쳐 맞춤 알고리즘(multi-features fitting algorithm)(MFF)에서 다양한 이미지 피쳐들을 이용한다.
- [0017] EP1039417은 3차원 객체의 이미지를 프로세싱하는 방법에 관한 것이고, 상기 방법은 복수의 3차원 이미지들로부터 얻어진 변형가능한 객체 모델을 제공하는 단계, 변형가능한 객체 모델을 적어도 하나의 2D 객체 이미지와 매칭하는 단계, 및 매칭된 변형가능한 객체 모델을 객체의 3D 표현으로서 제공하는 단계를 포함한다.
- [0018] US8553973은 3D 객체들(예를 들면, 인간 얼굴들)을 모델링하기 위한 또 다른 방법들 및 시스템들에 관한 것이다. 게임 콘솔들, 랩탑들, 태블릿들, 스마트폰들과 같은 다양한 소비자 디바이스들, 및 예를 들면, 자동차들에서의 깊이 센서들의 출현으로 인해, 얼굴들의 더욱 더 많은 이미지들이 깊이 맵을 포함할 것이다. RGB-D 데이터세트들을 생성하는 깊이 센서의 대중적인 예는 엑스박스(Xbox) 360, 엑스박스 원 및 윈도우즈 PC(마이크로소프트사의 모든 상표들)를 위해 마이크로소프트에 의해 제안된 키넥트 입력 디바이스이다. 깊이 센서들은 이미지의 각각의 픽셀에 대한 깊이(또는 광원에 대한 거리)의 표시를 포함하지만, 예를 들면 머리의 뒤쪽과 같은, 숨겨지거나 폐색된 요소들에 관해 어떠한 표시도 포함하지 않는 2.5D 이미지 데이터세트들을 생성한다.
- [0019] 테스트 이미지들이 깊이 센서들로 캡처되고, 그 다음 기존의 2D 이미지들과 비교되는 새로운 얼굴 인식 방법들을 제공하는 것이 바람직할 것이다. 이러한 방법은 RGB-D 이미지 데이터의 획득을 위해 최신의 깊이 센서들을 이용하고, 그들을 폭넓게 이용가능한 2D 기준 이미지들과 비교하는 장점을 가질 것이다.
- [0020] 포즈 교정의 업무를 개선하기 위해 깊이 센서들의 능력들 및 RGB-D 데이터세트들을 이용하여 새로운 방법들을 제공하는 것이 또한 바람직할 것이다.
- [0021] 기존의 방법들보다 빠른, 머리 포즈를 평가하기 위한 새로운 방법을 제공하는 것이 또한 바람직할 것이다.

- [0022] 기존 방법들보다 정확한, 머리 포즈를 평가하기 위한 새로운 방법을 제공하는 것이 또한 바람직할 것이다.
- [0023] 넓은 포즈 변형들을 취급할 수 있는, 머리 포즈를 평가하기 위한 새로운 방법을 제공하는 것이 또한 바람직할 것이다.
- [0024] 넓은 포즈 변형들을 취급할 수 있는, 머리 포즈를 평가하기 위한 새로운 방법을 제공하는 것이 또한 바람직할 것이다.

발명의 내용

해결하려는 과제

- [0025] 따라서, 본 발명의 목적은 얼굴 이미지들을 표현하는 2.5D 데이터세트들에서 포즈를 교정하기 위한 새로운 방법을 제안하는 것이다.

과제의 해결 수단

- [0026] 본 발명에 따라, 이들 목적들은 얼굴 이미지들을 표현하는 데이터에서 포즈를 교정하기 위한 포즈 교정 방법에 의해 성취되고, 상기 방법은:
- [0027] A-2D 근적외선 이미지 데이터, 2D 가시광선 이미지 데이터, 및 깊이 맵을 포함하는 적어도 하나의 테스트 프레임임을 획득하는 단계;
- [0028] C-공지된 방향의 머리의 3D 모델에 따라 상기 깊이 맵을 정렬함으로써 상기 테스트 프레임에서 얼굴의 포즈를 추정하는 단계;
- [0029] D-텍스처된 이미지 데이터를 생성하기 위해, 상기 2D 이미지의 적어도 하나를 깊이 맵 상에 매핑하는 단계;
- [0030] E-포즈 교정된 2D 투영된 이미지를 표현하는 데이터를 생성하기 위해 텍스처된 이미지 데이터를 2D로 투영하는 단계를 포함한다.
- [0031] 깊이 맵(즉, 2.5D 데이터세트)의 이용이 이로운데, 이는 그것이 포즈 추정 단계의 정확성 및 강건성을 개선하기 때문이다.
- [0032] 가시광선 데이터의 이용이 이로운데, 이는 그것이 예를 들면, 스킨 컬러 또는 텍스처에 의존하는 피쳐들의 검출을 허용하기 때문이다.
- [0033] 근적외선(NIR) 이미지 데이터의 이용이 이로운데, 이는 그것이 가시광선 데이터보다 조명 조건들에 덜 의존하기 때문이다.
- [0034] 머리 포즈는 깊이 맵의 방향을 추정하기 위해, 기존의 3D 모델에 따라 깊이 맵을 맞춤으로써 추정된다. 이 포즈 추정은 특히 강건하다.
- [0035] 3D 모델은 일반 모델 즉, 사용자 독립 모델일 수 있다.
- [0036] 3D 모델은 사용자 의존 모델 예를 들면, 신원이 검증될 필요가 있는 이용자의 머리의 3D 모델일 수 있다.
- [0037] 3D 모델은 성별 특정 모델, 인종 특정 모델, 또는 나이 특정 모델이고, 성별, 인종성 및/또는 나이의 선택 지식에 기초하여 선택될 수 있다.
- [0038] 기존의 3D 모델은 얼굴의 포즈를 추정하기 위해 이용된다. 그러나, 2D 이미지 데이터는 이 3D 모델 상에 매핑되지 않고, 깊이 맵 상에 매핑된다.
- [0039] 방법은 이미지를 분류하는 예를 들면, 2D 투영된 이미지를 분류하는 또 다른 단계를 포함할 수 있다.
- [0040] 분류는 얼굴 인증, 얼굴 식별, 성별 추정, 나이 추정, 및/또는 다른 얼굴 피쳐들의 검출을 포함할 수 있다.
- [0041] 방법은 2D 투영된 이미지를 또한 프로세싱하는 단계를 포함할 수 있다.
- [0042] 획득 단계는 깊이 센서에 의해 생성된 노이즈(noise)를 제거하기 위해, 깊이 맵에서 포인트들의 시간적 및/또는 공간적 스무딩을 포함할 수 있다.
- [0043] 포즈를 추정하는 단계는 예를 들면, 랜덤 회귀 포레스트 방법에 기초하여 개략적인 포즈 추정을 수행하는 제 1

단계를 포함할 수 있다.

- [0044] 포즈를 추정하는 단계는 미세 포즈 추정의 제 2 단계를 포함할 수 있다. 미세 포즈 추정은 개략적인 포즈 추정의 결과에 기초할 수 있다.
- [0045] 미세 포즈 추정은 엄격한 반복 최근접점(Iterative Closest Point; ICP) 방법들을 이용하여 예를 들면, 3D 모델에 따른 깊이 맵의 정렬에 기초할 수 있다.
- [0046] 방법은 얼굴에 속하지 않는 2D 근적외선 이미지의, 및/또는 상기 2D 가시광선 이미지의, 및/또는 상기 깊이 맵의 적어도 일부 부분들을 제거하기 위해, 상기 포즈 추정 전에 기본 얼굴 검출 단계를 더 포함할 수 있다.
- [0047] 방법은 배경에 속하지 않는 상기 2D 근적외선 이미지의, 및/또는 상기 2D 가시광선 이미지의, 및/또는 상기 2D 근적외선의 부분들을 제거하기 위해 배경 추출 단계를 더 포함할 수 있다.
- [0048] 머리의 기존의 3D 모델에 따라 깊이 맵을 정렬하는 단계는 깊이 맵의 크기를 조정하는(scaling) 단계를 포함할 수 있어서, 그것의 면적 중 일부(예를 들면, 최대 높이)가 3D 모델의 대응하는 면적과 매칭한다.
- [0049] 머리의 기존의 3D 모델에 따라 깊이 맵을 맞추는 단계는 깊이 맵 및/또는 3D 모델을 휘게하는 단계를 포함할 수 있다.
- [0050] 방법은 2D 근적외선 이미지 데이터세트에 기초하여 2D 가시광선 이미지 데이터세트의 조명을 정정하는 또 다른 단계를 포함할 수 있다. 2D 근적외선 이미지 데이터세트에서 나타나지 않는 2D 가시광선 이미지 데이터세트에서의 그림자 또는 밝은 존들이 따라서 정정될 수 있다.
- [0051] 방법은 깊이 맵 상에서 및/또는 2D 이미지 상에서 보이지 않는 부분들에 대응하는 상기 포즈 교정된 2D 투영된 이미지 데이터의 부분들을 표시하는(flagging) 또 다른 단계를 포함할 수 있다.
- [0052] 방법은 깊이 맵의 공지되지 않은 부분들에 대응하는 포즈 교정된 2D 투영된 이미지 데이터의 부분들을 재구성하는 또 다른 단계를 포함할 수 있다.
- [0053] 본 발명은 예로서 주어진 그리고 도면들에 의해 도시된 일 실시예의 설명의 도움으로 더 양호하게 이해될 것이다.

도면의 간단한 설명

- [0054] 도 1은 본 발명의 방법의 흐름도.
- 도 2는 본 발명에 따른 장치의 개략도.
- 도 3은 얼굴의 2D 가시광선 이미지의 일례를 도시한 도면.
- 도 4는 도 3의 얼굴의 2D 적외선 이미지의 일례를 도시한 도면.
- 도 5는 도 3의 얼굴의 깊이 맵의 표현의 일례를 도시한 도면.
- 도 6은 도 3의 얼굴의 텍스처된 깊이 맵의 표현의 일례를 도시한 도면.
- 도 7은 머리 포즈 추정을 위해 이용된 일반 머리 모델의 일례를 도시한 도면.
- 도 8은 3D 모델 내에 깊이 맵을 정렬함으로써 미세 포즈 추정 단계를 도시한 도면.
- 도 9는 깊이 맵의 포즈 교정된 2D 투영을 도시한 도면이고, 깊이 맵에서 존재하지 않는 머리의 부분들에 대응하는 누락 부분들이 표시된다.
- 도 10은 2.5D 데이터세트의 포즈 교정된 2D 투영을 도시한 도면이고, 깊이 맵에서 존재하지 않는 머리의 부분들에 대응하는 누락 부분들이 재구성된다.
- 도 11은 2.5D 데이터세트의 포즈 교정된 2D 투영을 도시한 도면이고, 2.5D에서 존재하지 않는 머리의 부분들에 대응하는 누락 부분들이 재구성되고 표시된다.

발명을 실시하기 위한 구체적인 내용

- [0055] 첨부된 도면들과 관련하여 하기에 제시된 상세한 설명은 본 발명의 다양한 실시예들의 설명으로서 의도되고 본 개시가 실행될 수 있는 유일한 양태들을 표현하도록 의도되지 않는다. 이 개시에서 설명된 각각의 양태는 단지

본 발명의 일례 또는 예시로서 제공되고, 반드시 바람직하거나 필수적인 것으로서 해석되는 것은 아니다.

- [0056] 도 1은 본 발명에 따른 포즈 교정 방법의 일례의 주요 단계들을 개략적으로 도시하는 흐름도이다. 이 방법은 도 2 상에 블록도로서 개략적으로 도시된 장치 또는 시스템으로 실행될 수 있다. 이 예에서, 장치는 사용자(100)의 이미지를 캡처하기 위한 카메라(101)를 포함한다. 카메라(101)는 제한 없이 키넥트 카메라(마이크로소프트의 상표), 비행 시간(time-of-flight) 카메라, 또는 RGB-D 데이터 스트림을 생성할 수 있는 임의의 다른 카메라와 같은, 깊이 카메라일 수 있다.
- [0057] 카메라(101)는 메모리(104)에 액세스하는 프로세서(102)에 접속되고 네트워크 인터페이스(103)를 통해 네트워크에 접속된다. 메모리(104)는 프로세서로 하여금 도 1의 방법의 적어도 일부 단계들을 실행하게 하는 컴퓨터 코드를 저장하기 위한 영구적 메모리 부분을 포함할 수 있다. 메모리(104), 또는 메모리(104)의 부분들은 제거 가능할 수 있다.
- [0058] 본 명세서에서 이용된 바와 같이, 장치(101+102+104)는 모바일 폰, 개인용 네비게이션 디바이스, 개인 정보 관리자(PIM), 자동차 장비, 게이밍 디바이스, 개인 휴대용 정보 단말기(PDA), 랩탑, 태블릿, 노트북 및/또는 핸드헬드 컴퓨터, 스마트 글래스, 스마트 시계, 스마트TV, 다른 웨어러블 디바이스, 등의 형태를 취할 수 있다.
- [0059] 도 1의 획득 단계(A)에서, 테스트 비디오 스트림은 식별되거나 그렇지 않으면 분류될 사용자(100)의 테스트 이미지를 캡처하기 위해 깊이 카메라(101)에 의해 생성된다.
- [0060] 본 설명에서, 표현("테스트 이미지들")은 포즈가 전형적으로, (식별 또는 인증을 위한) 테스트 동안 뿐만 아니라, 등록 동안 교정될 필요가 있는 페이스의 이미지들을 지정한다.
- [0061] 테스트 비디오 스트림의 각각의 프레임은 바람직하게, 3개의 시간적으로 그리고 공간적으로 정렬된 데이터셋들을 포함한다:
- [0062] i) (예를 들면, 그레이스케일(grayscale) 또는 RGB 이미지와 같은) 사용자(100)의 얼굴의 2차원(2D) 가시광선 이미지에 대응하는 제 1(선택적) 데이터셋. 하나의 예가 도 3 상에 도시된다.
- [0063] ii) 사용자(100)의 얼굴의 2D 근적외선(NIR) 이미지를 표현하는 제2(선택적) 데이터셋. 하나의 예가 도 4 상에 도시된다.
- [0064] iii) 각각의 픽셀과 연관된 깊이 발광원의 깊이 즉, 카메라(101)에서 깊이 센서까지의 그것의 거리에 의존하는 깊이 맵(즉, 2.5D 데이터셋). 이러한 깊이 맵의 표현은 도 5 상에 도시된다.
- [0065] 도 6은 프레임의 또 다른 표현(201)이고, 여기서 제 1 RGB 데이터셋이 깊이 맵 상에 투영된다.
- [0066] 도 5 상에서 보여질 수 있는 바와 같이, 많은 저 비용 깊이 센서들은 노이즈 깊이 맵들 즉, 각각의 포인트에 할당된 깊이 값이 노이즈를 포함하는 데이터셋들을 생성한다. 노이즈의 해로운 영향은 깊이 맵을 스무딩함으로써 감소될 수 있다. 스무딩은 공간 및/또는 시간(연속적인 프레임들에 걸친) 도메인에서 깊이 맵의 지역 필터링을 포함할 수 있다.
- [0067] 도 1의 기본 얼굴 검출 단계(B)에서, 사용자 얼굴을 표현하는 각각의 데이터셋의 부분이 검출되고 이미지의 배경 및 다른 요소들로부터 분리된다. 이 검출은 3개의 데이터셋들 중 임의의 데이터셋 또는 그들 전부에 관해 수행되고 그들 데이터셋들 중 임의의 데이터셋 또는 그들 전부에 적용될 수 있다.
- [0068] 하나의 실시예에서, 이 기본 얼굴 검출은 미리 정의된 깊이 범위 예를 들면, 20cm와 100cm 사이에 있지 않은 픽셀들을 배제하기 위해, 깊이 맵의 경계화(thresholding)에 적어도 부분적으로 기초한다. 다른 공지된 알고리즘들은 사용자 얼굴을 표현하는 배경을 추출하고, 배경을 배제하기 위해 이용될 수 있고, 이들은 예를 들면, 컬러 검출에 기초한 알고리즘들을 포함한다.
- [0069] 도 1의 머리 포즈 추정 단계(C)에서, 프레임에서의 사용자의 포즈가 추정된다. 하나의 예에서, 이 추정은 2개이 연속적인 단계들로 수행된다:
- [0070] 머리 포즈 추정 단계의 제 1 부분 동안, 머리 포즈의 개략적인 추정치가 결정된다. 하나의 예에서, 이 개략적인 추정치의 산출은 몇몇 정확도들로 머리 포즈를 빠르게 결정하기 위해, 랜덤 포레스트 알고리즘을 이용한다. 랜덤 회귀 포레스트를 통한 개략적인 머리 포즈 추정 방법은 지. 파넬리(G. Fanelli), 등에 의한, "랜덤 회귀 포레스트들을 통한 실시간 머리 포즈 추정(Real Time Head Pose Estimation with Random Regression Forests)", 컴퓨터 비전 및 패턴 인식(CVPR), 2011 IEEE 회의, 617 내지 624. 바람직하게, 코의 및/또는 얼굴의 다른 키 피

처들의 위치는 또한, 이 단계 동안 결정된다.

- [0071] 머리 포즈 추정의 제 2 단계 동안, 머리 포즈의 더 미세한 추정치가 결정된다. 이 미세한 추정치는 속도 및 강건성을 개선하기 위해, 방향의 그리고 코 위치와 같은, 하나의 포인트의 위치의 이전에 결정된 개략적인 추정치로부터 시작할 수 있다. 미세한 추정치는 깊이 맵에 최상으로 대응하는 머리의 3D 모델(도 7)의 방향을 결정함으로써 컴퓨터될 수 있다. 하나의 예에서, 테스트 깊이 맵에 대한 3D 모델의 매칭은 예를 들면, 반복 최근접점(ICP) 방법을 이용하여 깊이 맵의 포인트들과 3D 모델의 대응하는 포인트들 사이의 거리 함수를 최소화함으로써 수행될 수 있다. 도 8은 3D 모델(200)(이 예시에서 메시(mesh)로서 표현됨) 내에 테스트 깊이 맵(201)(이 예시에서 텍스처됨) 맞추는 이 단계를 개략적으로 도시한다.
- [0072] 이 정렬 단계는 바람직하게, 3D 모델의 조정을 포함하여, 그것의 면적의 적어도 일부가 테스트 깊이 맵에 대응한다.
- [0073] 3D 모델(200)은 일반적, 즉 사용자 독립적일 수 있다. 대안적으로, 3D 모델은 사용자 의존적이고, 예를 들면, 사용자(100)의 가정된 신원에 기초하여 사용자 의존 3D 모델들의 데이터베이스로부터 검색될 수 있다. 복수의 사용자 독립 3D 모델들은 또한, 예를 들면, 사용자의 가정된 성별, 나이, 또는 인종성에 따라 저장되고 선택될 수 있다.
- [0074] 하나의 예에서, 개인 3D 모델은 엄격하지 않은 반복 최근접점(ICP) 방법을 이용하여 생성된다. 3D 모델은 그 다음, 머리의 변형가능한 부분들 예를 들면, 얼굴의 하부에서 일부 현실적이고 제한된 변형을 허용하기 위해, 노드들 사이의 위치 및/또는 관계들에 관한 일부 제약들을 갖는 메시를 포함할 수 있다. 이 경우에, ICP 방법은 모든 가능한 변형들을 고려해볼 때 가장 공산이 큰 방향을 발견하기 위해, 모델의 일부 변형들 또는 모프(morph)를 시도할 수 있다.
- [0075] 머리 포즈 추정 단계의 출력은 주어진 좌표 시스템에 관해서 3개의 회전들을 설명하는 각들(피(phi), 세타(theta), 시(psi))의 세트를 포함할 수 있다.
- [0076] 도 1의 단계(D)에서, 프레임의 2D 텍스처들(가시적 및 NIR 범위들에서의)은 깊이 맵 상에 매핑된다(UV 매핑). 단지 가시적 데이터세트의 그레이스케일 값을 이용하는 것이 가능하다. 이 매핑은 포즈 정정 전후에 수행될 수 있고, 공지된 방향을 갖는 텍스처된 깊이 맵을 생성한다. 숨겨진 부분들 예를 들면, 깊이 맵에서 숨겨지거나 폐색된 이용자의 얼굴의 부분들은 예를 들면, 이용자의 얼굴의 대칭을 가정함으로써 유효하지 않은 것으로서 표시되거나, 재구성된다.
- [0077] 이 단계는 또한, 가시광선 및/또는 NIR 이미지 데이터세트들에서의 조명의 정정을 포함할 수 있다. 조명의 정정은 컬러 이미지들의 경우에 밝기, 콘트라스트, 및/또는 화이트 밸런스의 정정을 포함할 수 있다. 하나의 바람직한 실시예에서, NIR 데이터세트는 가시광선 데이터세트의 부분들에서 나타나지만 NIR 데이터세트들의 대응하는 부분들에서 나타나지 않는 밝기 변화들을 보상함으로써, 그림자를 제거하거나 약화시키기 위해 이용되고/되거나 가시광선 데이터세트에서 반사한다.
- [0078] 도 1의 단계(E)에서, 텍스처된 3D 이미지는 가시적 및/또는 NIR 범위에서, 포즈 교정된 2D 투영된 이미지를 표현하는 적어도 하나의 데이터세트를 생성하기 위해 2D로 투영된다. 다양한 투영들이 고려될 수 있다. 먼저, 카메라에 의해 및/또는 원근법에 의해 야기된 변형들이 바람직하게 고려된다. 그 다음, 하나의 실시예에서, 투영은 정면 얼굴 이미지 즉, 사용자(100) 정면의 뷰로부터 보여지는 바와 같은 2D 이미지를 생성한다. 정면이 아닌 얼굴 이미지, 또는 예를 들면, 하나의 정면 얼굴 투영 및 하나의 또 다른 프로파일 투영과 같은 복수의 2D 투영들을 생성하는 것이 또한 가능하다. 지도 투영들, 또는 얼굴의 차별적 부분들 특히, 눈, 얼굴의 상부 절반을 확대하기 위해 변형들을 도입하고, 입과 같은, 얼굴의 더 변형가능한 부분들의 크기를 감소시키는 투영들을 포함하는 다른 투영들이 고려될 수 있다. 비교를 용이하게 하기 위해, 비교 전에 머리를 일반 모델로 변형시키는 것이 또한 가능하다. 용이하게 표현가능하지 않은 공간 상으로의 투영들과 같은, 순전히 수학적 투영들이 또한 고려될 수 있다.
- [0079] 도 9는 단계(E) 동안 생성된 2D 텍스처된 투영(202)의 일례를 도시한다. 깊이 맵에서 이용가능하지 않은 투영의 부분들(203) 예를 들면, 숨겨지거나 폐색된 부분들은 이와 같이 표시된다.
- [0080] 도 10은 단계(E) 동안 생성된 2D 텍스처된 투영(202)의 또 다른 예를 도시한다. 이 예에서, 깊이 맵에서 이용가능하지 않은 투영의 부분들(204) 예를 들면, 숨겨지거나 폐색된 부분들은 비 텍스처된 이미지로서 재구성된다. 재구성은 예를 들면, 이용자의 얼굴이 대칭이라고 가정함으로써 이미지의 이용가능한 부분들에 기초할 수 있다. 대안적으로, 또는 게다가, 재구성은 머리의 일반 모델을 이용할 수 있다. 대안적으로, 또는 게다가, 재구성은

동일한 비디오 시퀀스에서의 다른 프레임들로부터 이용가능한 이미지 부분 데이터를 이용할 수 있다.

- [0081] 도 10은 단계(E) 동안 생성된 2D 텍스처된 투영(202)의 또 다른 예를 도시한다. 이 예에서, 깊이 맵에서 이용가능하지 않은 투영의 부분들(205) 예를 들면, 숨겨지거나 폐색된 부분들은 텍스처된 이미지로서 재구성된다. 재구성은 예를 들면, 이용자의 얼굴이 대칭이라고 가정함으로써 이미지의 이용가능한 부분들에 기초할 수 있다. 대안적으로, 또는 게다가, 재구성은 머리의 일반 모델을 이용할 수 있다. 대안적으로, 또는 게다가, 재구성은 동일한 비디오 시퀀스에서의 다른 프레임들로부터 이용가능한 이미지 부분 데이터를 이용할 수 있다.
- [0082] 상기 설명된 방법은 따라서, 깊이 카메라로 얻어진 2.5D 테스트 뷰에 기초하여, 이용자의 포즈 정정된 2D 테스트 이미지 데이터셋을 생성한다. 얼굴 프로세싱 단계(F) 동안, 이 데이터셋은 그 다음, 이용자 식별 또는 인증 모듈, 또는 성별 추정 모듈, 나이 추정 모듈, 등과 같은 분류 모듈에 의해 이용될 수 있다. 분류는 단일 프레임 예를 들면, 최고의 신뢰성에 따라 분류될 수 있는 프레임에, 또는 주어진 임계치보다 높은 신뢰성에 따라 분류될 수 있는 제 1 프레임에, 또는 동일한 비디오 스트림의 복수의 연속적인 프레임들에 기초할 수 있다. 부가적으로, 또는 대안적으로, 분류는 또한, 지향된 텍스처된 3D 이미지에 기초할 수 있다. 다른 얼굴 프로세싱 이 단계(F) 동안 적용될 수 있다.
- [0083] 본 명세서에 개시된 방법들은 설명된 방법을 성취하기 위한 하나 이상의 단계들 또는 동작들을 포함한다. 단계들 또는 동작들의 특정 순서가 명시되지 않으면, 특정 단계들 및/또는 동작들의 순서 및/또는 이용은 청구항들의 범위를 벗어나지 않고 변경될 수 있다.
- [0084] 실시예에 의존하여, 본 명세서에서 설명된 방법들 중 임의의 방법의 특정 행위 또는 이벤트(event)들 또는 단계들이 상이한 시퀀스로 수행될 수 있거나, 모두 함께 부가되거나, 합쳐지거나, 배제될 수 있음이 인식될 것이다 (예로서, 설명된 모든 행위가 방법의 실행을 위해 필요한 것은 아니다). 게다가, 특정 실시예들에서, 동작들 또는 이벤트들은 예로서, 멀티 쓰레딩(multi-threading)된 프로세싱, 인터럽트 프로세싱(interrupt processing), 또는 다수의 프로세서들을 통해, 순차적이기보다는 동시에 수행될 수 있다.
- [0085] 상기 설명된 방법들의 다양한 동작들은 대응하는 기능들을 수행할 수 있는 임의의 적합한 수단에 의해 수행될 수 있다. 수단은 본 명세서에서 설명된 방법 단계들을 실행하기 위해 설계된 회로, 주문형 반도체(application specific integrate circuit; ASIC), 프로세서, 필드 프로그래밍가능한 게이트 어레이 신호(FPGA) 또는 다른 프로그래밍가능한 로직 디바이스(PLD), 별개의 게이트 또는 트랜지스터 로직, 별개의 하드웨어 구성요소들 또는 그의 임의의 조합을 포함하지만, 그들로 제한되지 않는, 다양한 하드웨어 및/또는 소프트웨어 구성요소(들) 및/또는 모듈(들)을 포함할 수 있다.
- [0086] 본 명세서에서 이용된 바와 같이, 용어들("결정하는" 및 "추정하는")은 광범위한 동작들을 포함한다. 예를 들면, "결정하는" 및 "추정하는"은 산출하는 것, 컴퓨팅하는 것, 얻는 것, 검색하는 것(예로서, 표, 데이터베이스 또는 또 다른 데이터 구조에서 검색하는 것), 확인하는 것 등을 포함할 수 있다. 또한, "결정하는" 및 "추정하는"은 수신하는 것, 액세스하는 것(예로서, 메모리의 데이터에 액세스하는 것) 등을 포함할 수 있다.
- [0087] 본 개시와 관련하여 설명된 방법 또는 알고리즘의 단계들은 하드웨어로, 프로세서에 의해 실행된 소프트웨어 모듈로, 또는 그 둘의 조합으로 직접적으로 구현될 수 있다. 소프트웨어 모듈은 본 분야에서 공지되는 임의의 형태의 저장 매체에 상주할 수 있다. 이용될 수 있는 저장 매체들의 일부 예들은 랜덤 액세스 메모리(RAM), 판독 전용 메모리(ROM), 플래시 메모리, EPROM 메모리, EEPROM 메모리, 레지스터들, 하드 디스크, 착탈가능한 디스크, CD-ROM 등을 포함한다. 소프트웨어 모듈은 단일 지시, 또는 많은 지시들을 포함할 수 있고, 상이한 프로그램들 중에서, 몇몇 상이한 코드 세그먼트들을 통해, 그리고 다수의 저장 매체에 걸쳐 분산될 수 있다. 저장 매체는 프로세서가 저장 매체로부터 정보를 판독하고, 상기 저장 매체에 정보를 기록할 수 있도록 프로세서에 결합될 수 있다. 대안에서, 저장 매체는 프로세서와 일체형일 수 있다.
- [0088] 따라서, 특정 양태들은 본 명세서에서 제시된 동작들을 수행하기 위한 컴퓨터 프로그램 제품을 포함할 수 있다. 예를 들면, 이러한 컴퓨터 프로그램 제품은 지시들이 저장된(및/또는 인코딩된) 컴퓨터 판독가능한 매체를 포함할 수 있고, 지시들은 본 명세서에서 설명된 동작들을 수행하기 위해 하나 이상의 프로세서들에 의해 실행가능하다.
- [0089] 본 발명의 설명된 실시예들에 대한 다양한 수정들 및 변형들은 첨부된 청구항들에서 정의된 바와 같이 본 발명의 범위를 벗어나지 않고 당업자들에게 명백할 것이다. 본 발명이 특정 바람직한 실시예들과 관련하여 설명되었을지라도, 청구된 바와 같이 본 발명이 이러한 특정 실시예로 지나치게 제한되지 않아야 함이 이해되어야 한다.

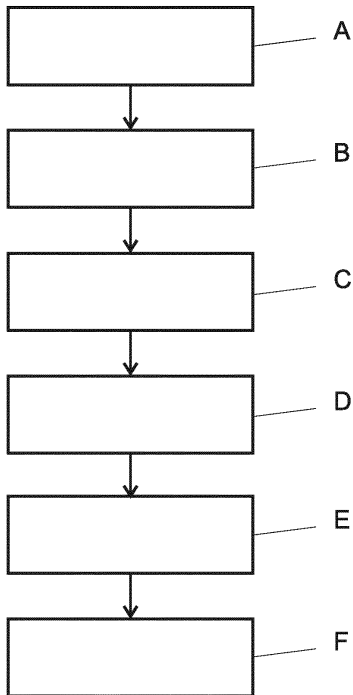
부호의 설명

[0090]

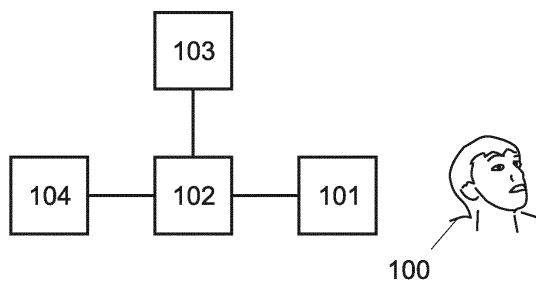
- 100: 이용자 101: 카메라
- 102: 프로세서
- 103: 네트워크 인터페이스 104: 메모리
- 200: 3D 모델 201: 테스트 깊이 맵
- 202: 2D 텍스처된 투영

도면

도면1



도면2



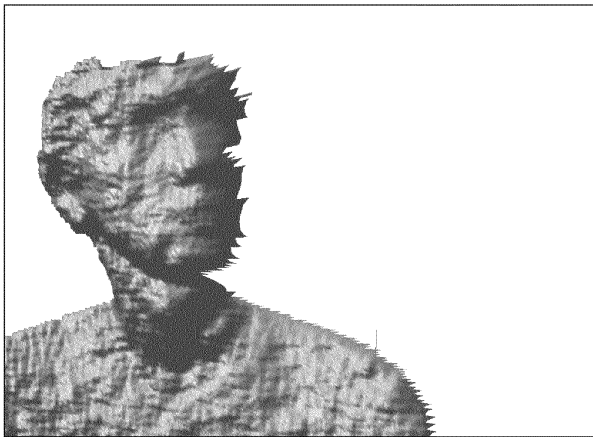
도면3



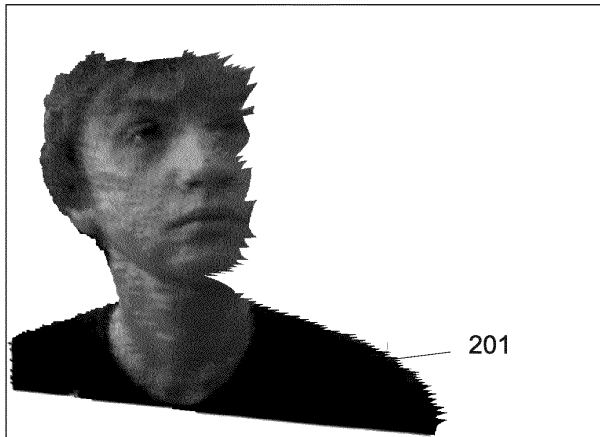
도면4



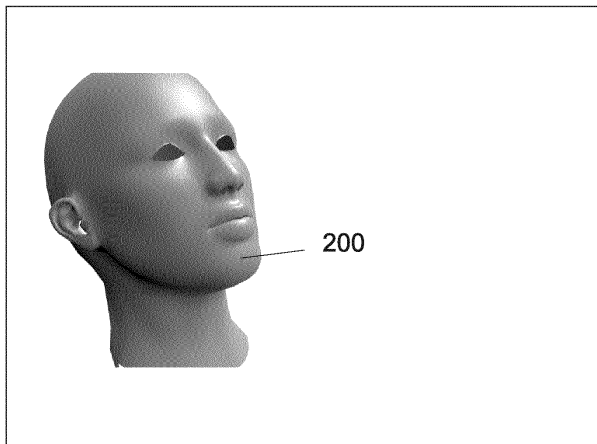
도면5



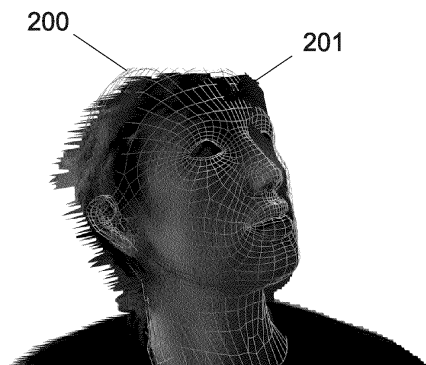
도면6



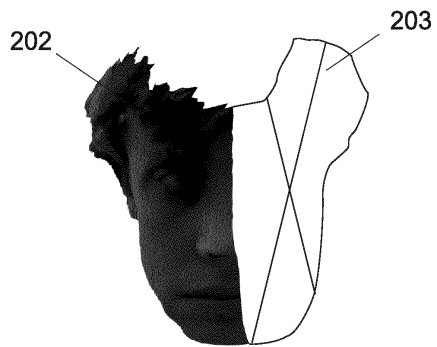
도면7



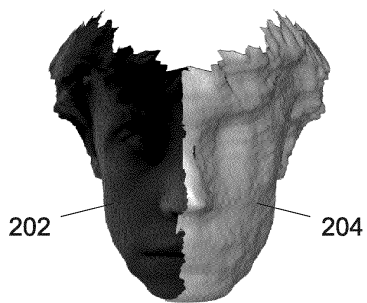
도면8



도면9



도면10



도면11

