



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2021년02월18일

(11) 등록번호 10-2217797

(24) 등록일자 2021년02월15일

(51) 국제특허분류(Int. Cl.)

G06T 17/20 (2006.01) G02B 27/00 (2020.01)

G02B 27/01 (2006.01) G06K 9/00 (2006.01)

G06T 13/40 (2011.01) G10L 21/10 (2013.01)

(52) CPC특허분류

G06T 17/20 (2013.01)

G02B 27/0093 (2013.01)

(21) 출원번호 10-2019-7014144

(22) 출원일자(국제) 2017년11월09일

심사청구일자 2020년11월05일

(85) 번역문제출일자 2019년05월16일

(65) 공개번호 10-2019-0084260

(43) 공개일자 2019년07월16일

(86) 국제출원번호 PCT/US2017/060942

(87) 국제공개번호 WO 2018/089691

국제공개일자 2018년05월17일

(30) 우선권주장

62/421,033 2016년11월11일 미국(US)

(56) 선행기술조사문헌

US09479736 B1

US20060009978 A1

JP2010250419 A

KR1020160091402 A

(73) 특허권자

매직 립, 인코포레이티드

미국 플로리다 플랜타타운 웨스트 선라이즈 블러
바드 7500 (우: 33322)

(72) 발명자

캐홀러, 아드리안

미국 90027 캘리포니아 로스앤젤레스 노스 웨스턴
애비뉴 1940

(74) 대리인

특허법인 남앤남

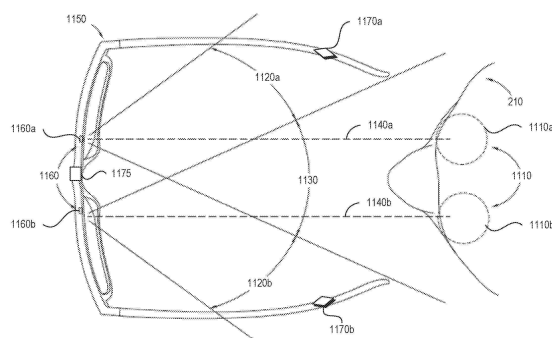
전체 청구항 수 : 총 23 항

심사관 : 강석제

(54) 발명의 명칭 전체 얼굴 이미지의 안구주위 및 오디오 합성

(57) 요약

HMD(head-mounted device)에 의해 얼굴의 이미지를 합성하기 위한 시스템들 및 방법들이 개시된다. HMD는 얼굴의 부분을 관찰할 수 없을 수 있다. 본원에서 설명된 시스템들 및 방법들은 이미징되지 않은 얼굴의 부분의 컨퍼메이션(conformation)로부터 관찰된 얼굴의 부분의 컨퍼메이션으로의 맵핑을 생성할 수 있다. HMD는 얼굴의 부분의 이미지를 수신하고 맵핑을 이용하여 관찰되지 않은 얼굴의 부분의 컨퍼메이션을 결정할 수 있다. HMD는 관찰된 부분 및 관찰되지 않은 부분을 결합하여 전체 얼굴 이미지를 합성할 수 있다.

대표도 - 도11

(52) CPC특허분류

G02B 27/0172 (2013.01)
G06K 9/00275 (2013.01)
G06K 9/00281 (2013.01)
G06K 9/0061 (2013.01)
G06T 13/40 (2013.01)
G10L 21/10 (2013.01)
G02B 2027/0138 (2013.01)
G02B 2027/014 (2013.01)
G02B 2027/0178 (2013.01)

명세서

청구범위

청구항 1

스피치(speech) 동안 사용자의 얼굴을 애니메이션(animating)하기 위한 웨어러블 시스템으로서,

이미지들을 캡처하도록 구성되는 내향 이미징 시스템(inward-facing imaging system) — 상기 내향 이미징 시스템은, 상기 웨어러블 시스템이 상기 사용자에게 의해 착용될 때 상기 사용자의 얼굴의 안구주위 구역이 상기 내향 이미징 시스템에 의해 관찰되고 상기 사용자의 하부 얼굴은 상기 내향 이미징 시스템에 의해 관찰되지 않도록 포지셔닝되는 하나 또는 그 초과 카메라들을 포함함 —;

상기 사용자의 스피치를 수신하도록 구성된 오디오 센서; 및

하드웨어 프로세서를 포함하고,

상기 하드웨어 프로세서는,

상기 웨어러블 시스템이 상기 사용자에게 의해 착용될 때, 상기 내향 이미징 시스템을 통해, 상기 사용자의 안구주위 구역의 이미지를 획득하고;

상기 사용자의 안구주위 구역의 상기 이미지에 적어도 부분적으로 기초하여, 상기 사용자의 적어도 안구주위 구역의 안구주위 컨퍼메이션(periorcular conformation)을 인코딩하는 안구주위 얼굴 파라미터들을 생성하고;

상기 오디오 센서에 의해, 상기 사용자에게 의해 말해진 오디오 스트림을 획득하고;

상기 오디오 스트림에서 음소를 식별하고;

상기 사용자를 포함하지 않는 사람들의 그룹과 연관된 이미지들을 사용하여 생성되었던 기본 모델에 액세스하고;

상기 사용자의 안구주위 구역의 이미지 및 상기 기본 모델에 적어도 부분적으로 기초하여 맵핑을 커스터마이징하고 — 상기 맵핑의 입력은 상기 음소 및 상기 사용자의 안구주위 구역의 이미지를 포함하고, 상기 맵핑의 출력은 상기 사용자의 하부 얼굴의 컨퍼메이션을 인코딩하고 그리고 상기 음소 및 상기 사용자의 안구주위 구역의 이미지의 분석으로부터 추론되는 하부 얼굴 파라미터들을 포함함 —;

상기 하부 얼굴 파라미터들을 생성하도록 상기 사용자의 안구주위 구역의 상기 이미지에 상기 맵핑을 적용하고;

3-차원(3D) 얼굴 모델과 연관된 전체 얼굴 파라미터들을 생성하기 위해 상기 안구주위 얼굴 파라미터들 및 상기 하부 얼굴 파라미터들을 결합하고; 그리고

상기 전체 얼굴 파라미터들에 적어도 부분적으로 기초하여 상기 사용자의 얼굴의 애니메이션을 생성하도록 프로그래밍되는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 2

제1 항에 있어서,

상기 3D 얼굴 모델은 변형 가능한 선형 모델을 포함하고, 상기 안구주위 얼굴 파라미터들 및 상기 하부 얼굴 파라미터들은 상기 사용자가 말할 때 상기 얼굴의 변형을 설명하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 3

제2 항에 있어서,

상기 전체 얼굴 파라미터들을 생성하기 위해, 상기 하드웨어 프로세서는 상기 하부 얼굴 파라미터들 또는 상기 안구주위 얼굴 파라미터들 중 적어도 하나에 대한 업데이트를 반영하도록 상기 3D 얼굴 모델을 업데이트하도록 프로그래밍되는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 4

제1 항에 있어서,

상기 맵핑의 입력은 눈 특정 정보, 신체 움직임 또는 심박수 중 적어도 하나를 더 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 5

제4 항에 있어서,

상기 눈 특정 정보는 상기 사용자의 눈 포즈, 동공 확장 상태, 눈 컬러 또는 눈꺼풀 상태 중 적어도 하나를 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 6

제1 항에 있어서,

상기 하부 얼굴 파라미터들은 상기 오디오 스트림의 음소들을 시각적으로 설명하는 비슴(viseme)들을 인코딩하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 7

제1 항에 있어서,

상기 맵핑을 커스터마이징하기 위해, 상기 하드웨어 프로세서는 상기 안구주위 구역의 이미지에 적어도 부분적으로 기초하여 상기 사용자의 얼굴의 피부 텍스처를 추론하도록 프로그래밍되고, 상기 사용자의 얼굴의 애니메이션은 상기 얼굴의 피부 텍스처를 통합하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 8

제1 항에 있어서,

상기 내향 이미징 시스템은 눈 카메라를 포함하고, 상기 내향 이미징 시스템에 의해 획득된 상기 안구주위 구역의 이미지는 제1 눈에 대한 안구주위 구역의 이미지를 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 9

제8 항에 있어서,

상기 전체 얼굴 파라미터들을 생성하기 위해, 상기 하드웨어 프로세서는,

상기 내향 이미징 시스템에 의해 획득된 상기 안구주위 구역의 이미지에 기초하여 제2 눈에 대한 안구주위 얼굴 파라미터들을 결정하고; 그리고

상기 제2 눈에 대한 안구주위 얼굴 파라미터들을 상기 전체 얼굴 파라미터들에 통합하도록 프로그래밍되는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 10

제1 항에 있어서,

상기 사용자의 얼굴의 애니메이션을 생성하도록 상기 전체 얼굴 파라미터들이 적용되게 하기 위해, 상기 하드웨어 프로세서는 혼합 현실 디스플레이를 포함하는 웨어러블 디바이스에 명령들을 통신하도록 프로그래밍되고,

상기 명령들은 중립 포지션으로부터 상기 3D 얼굴 모델을 변동시키도록 상기 전체 얼굴 파라미터들이 적용되게 하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 11

제1 항에 있어서,

상기 기본 모델은 상기 사람들의 그룹의 이미지들에 기초하여 생성되는 3D 변형 가능한 선형 모델을 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 12

제1 항에 있어서,

상기 사용자의 하부 얼굴은, 상기 웨어러블 시스템이 상기 사용자에게 의해 착용될 때, 상기 웨어러블 시스템과 연관되는 어떠한 카메라에 의해서도 관찰되지 않는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 웨어러블 시스템.

청구항 13

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법으로서,

이미지들을 캡처하도록 구성되는 내향 이미징 시스템에 의해 획득된 사용자의 안구주위 구역의 이미지에 액세스하는 단계 — 상기 내향 이미징 시스템은, 웨어러블 시스템이 상기 사용자에게 의해 착용될 때 상기 사용자의 얼굴의 안구주위 구역이 상기 내향 이미징 시스템에 의해 관찰되고 상기 사용자의 하부 얼굴은 상기 내향 이미징 시스템에 의해 관찰되지 않도록 포지셔닝되는 하나 또는 그 초과 카메라들을 포함함 —;

상기 이미지들에 적어도 부분적으로 기초하여, 상기 사용자의 적어도 안구주위 구역의 안구주위 컨퍼메이션을 인코딩하는 안구주위 얼굴 파라미터들을 결정하는 단계;

오디오 센서에 의해 획득된, 상기 사용자에게 의해 말해진 오디오 스트림에 액세스하는 단계;

상기 오디오 스트림에서 음소를 식별하는 단계;

상기 사용자를 포함하지 않는 사람들의 그룹과 연관된 이미지들을 사용하여 생성되었던 기본 모델에 액세스하는 단계;

상기 사용자의 안구주위 구역의 이미지 및 상기 기본 모델에 적어도 부분적으로 기초하여 맵핑을 커스터마이징하는 단계 — 상기 맵핑의 입력은 상기 음소 및 상기 사용자의 안구주위 구역의 이미지를 포함하고, 상기 맵핑의 출력은 상기 사용자의 하부 얼굴의 컨퍼메이션을 인코딩하고 그리고 상기 음소 및 상기 사용자의 안구주위 구역의 이미지의 분석으로부터 추론되는 하부 얼굴 파라미터들을 포함함 —;

상기 하부 얼굴 파라미터들을 생성하도록 상기 이미지에 상기 맵핑을 적용하는 단계;

3-차원(3D) 얼굴 모델과 연관된 전체 얼굴 파라미터들을 생성하기 위해 상기 안구주위 얼굴 파라미터들 및 상기 하부 얼굴 파라미터들을 결합하는 단계; 및

상기 전체 얼굴 파라미터들에 적어도 부분적으로 기초하여 전체 얼굴 이미지를 생성하는 단계를 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 14

제13 항에 있어서,

상기 3D 얼굴 모델은 변형 가능한 선형 모델을 포함하고, 상기 안구주위 얼굴 파라미터들 및 상기 하부 얼굴 파

라미터들은 상기 사용자가 말할 때 상기 얼굴의 변형을 설명하는,
스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 15

제14 항에 있어서,

상기 전체 얼굴 파라미터들을 생성하기 위해, 하드웨어 프로세서는 상기 하부 얼굴 파라미터들 또는 상기 안구 주위 얼굴 파라미터들 중 적어도 하나에 대한 업데이트를 반영하도록 상기 3D 얼굴 모델을 업데이트하도록 프로 그래밍되는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 16

제13 항에 있어서,

상기 전체 얼굴 파라미터들은 상기 사용자의 얼굴과 연관된 애니메이션을 결정하기 위해 눈 특정 정보와 결합되는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 17

제16 항에 있어서,

상기 눈 특정 정보는 상기 사용자의 눈 포즈, 동공 확장 상태, 눈 컬러 또는 눈꺼풀 상태 중 적어도 하나를 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 18

제13 항에 있어서,

상기 하부 얼굴 파라미터들은 상기 오디오 스트림의 음소들을 시각적으로 설명하는 비슴(viseme)들을 인코딩하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 19

제13 항에 있어서,

상기 전체 얼굴 이미지는 추가로, 상기 내향 이미징 시스템에 의해 획득된 이미지에 적어도 부분적으로 기초하여 결정되는, 상기 사용자의 피부 텍스처들을 통합하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 20

제13 항에 있어서,

상기 맵핑은, 안구주위 얼굴 파라미터가 하부 얼굴 파라미터와 연관되는 가능성을 포함하고, 상기 하부 얼굴 파라미터들은 상기 하부 얼굴 파라미터들이 임계 기준들을 통과한다는 결정에 대한 응답으로 상기 전체 얼굴 이미지를 생성하도록 선택되는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 21

제13 항에 있어서,

상기 이미지는 정지 이미지 또는 비디오 프레임 중 적어도 하나를 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 22

제13 항에 있어서,

혼합 현실 환경에서 상기 전체 얼굴 이미지를 렌더링하도록 머리-장착 디스플레이에 지시하는 단계를 더 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

청구항 23

제13 항에 있어서,

상기 기본 모델은 상기 사람들의 그룹의 이미지들에 기초하여 생성되는 3D 변형 가능한 선형 모델을 포함하는,

스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법.

발명의 설명

기술 분야

[0001] 본 출원은, 2016년 11월 11일에 출원되고 발명의 명칭이 "PERIOULAR AND AUDIO SYNTHESIS OF A FULL FACE IMAGE"인 미국 가출원 번호 제62/421033호를 35 U.S.C. § 119(e) 하에서 우선권으로 주장하며, 그리하여, 이 출원의 개시내용은 그 전체가 인용에 의해 본원에 포함된다.

[0002] 본 발명은 혼합 현실 이미징 및 시각화 시스템들에 관한 것으로, 보다 구체적으로 오디오 데이터 및 안구주위(periocular) 이미지들을 이용한 얼굴 이미지의 합성 및 생성에 관한 것이다.

배경 기술

[0003] 현대 컴퓨팅 및 디스플레이 기술들은 소위 "가상 현실", "증강 현실", 또는 "혼합 현실" 경험들을 위한 시스템들의 개발을 용이하게 했으며, 여기서 디지털적으로 재생된 이미지들 또는 이미지들의 부분들은, 그들이 실제인 것으로 보이거나, 실제로서 지각될 수 있는 방식으로 사용자에게 제시된다. 가상 현실, 또는 "VR" 시나리오는 통상적으로 다른 실제 실세계 시각적 입력에 대한 투명성(transparency) 없이 디지털 또는 가상 이미지 정보의 프리젠테이션(presentation)을 수반하고; 증강 현실, 또는 "AR" 시나리오는 통상적으로 사용자 주위 실제 세계의 시각화에 대한 증강으로서 디지털 또는 가상 이미지 정보의 프리젠테이션을 수반하고; 혼합 현실, 또는 "MR"은, 실제 및 가상 세계들을 병합하여 물리적 및 가상 객체들이 공존하고 실시간으로 상호작용하는 새로운 환경을 생성하는 것과 관련된다. 밝혀진 바와 같이, 인간 시각적 지각 시스템은 매우 복잡하고, 다른 가상 또는 실세계 이미저리 엘리먼트들 사이에서 가상 이미지 엘리먼트들의 편안하고, 자연스럽고, 풍부한 프리젠테이션을 용이하게 하는 VR, AR 또는 MR 기술을 생성하는 것은 난제이다. 본원에 개시된 시스템들 및 방법들은 VR, AR 및 MR 기술에 관련된 다양한 난제들을 해결한다.

발명의 내용

[0004] HMD(head-mounted device)에 의해 얼굴의 이미지를 합성하기 위한 시스템들 및 방법들이 개시된다. HMD는 얼굴의 부분을 관찰할 수 없을 수 있다. 본원에서 설명된 시스템들 및 방법들은 이미징되지 않은 얼굴의 부분의 컨퍼메이션(conformation)로부터 관찰된 얼굴의 부분의 컨퍼메이션으로의 맵핑을 생성할 수 있다. HMD는 얼굴의 부분의 이미지를 수신하고 맵핑을 이용하여 관찰되지 않은 얼굴의 부분의 컨퍼메이션을 결정할 수 있다. HMD는 관찰된 부분 및 관찰되지 않은 부분을 결합하여 전체 얼굴 이미지를 합성할 수 있다.

[0005] 본 명세서에서 설명되는 청구 대상의 하나 이상의 구현들의 세부사항들은, 아래의 첨부 도면들 및 설명에서 기술된다. 다른 특징들, 양상들, 및 이점들은 설명, 도면들, 및 청구항들로부터 명백해질 것이다. 이 개요 또는 다음의 상세한 설명 어느것도, 본 발명의 청구 대상의 범위를 한정하거나 제한하는 것으로 의도되지 않는다.

도면의 간단한 설명

[0006]

[0006] 도 1은 사람에 의해 관찰되는 소정의 물리적 객체들 및 소정의 가상 현실 객체들을 갖는 혼합 현실 시나리오의 예시를 도시한다.

[0007] 도 2는 웨어러블 시스템의 예를 개략적으로 예시한다.

[0008] 도 3은 다중 깊이 평면들을 사용하여 3-차원 이미저리를 시뮬레이팅하기 위한 접근법의 양상들을 개략적으로 예시한다.

[0009] 도 4는 사용자에게 이미지 정보를 출력하기 위한 도파관 스택의 예를 개략적으로 예시한다.

[0010] 도 5는 도파관에 의해 출력될 수 있는 예시적인 출사 빔들을 도시한다.

[0011] 도 6은, 다중-초점 볼류메트릭 디스플레이, 이미지 또는 광 필드의 생성에 사용되는, 도파관 장치, 도파관 장치로 또는 도파관 장치로부터의 광을 광학적으로 커플링하는 광학 커플러 서브시스템, 및 제어 서브시스템을 포함하는 광학 시스템을 도시하는 개략도이다.

[0012] 도 7은 웨어러블 시스템의 일 예의 블록도이다.

[0013] 도 8은 인지된 객체들과 관련하여 가상 콘텐츠를 렌더링하는 방법의 예의 프로세스 흐름도이다.

[0014] 도 9는 웨어러블 시스템의 다른 예의 블록도이다.

[0015] 도 10은 가상 사용자 인터페이스와 상호작용하기 위한 방법의 예의 프로세스 흐름도이다.

[0016] 도 11은 사용자의 얼굴의 이미지들을 획득할 수 있는 예시적인 웨어러블 디바이스를 예시한다.

[0017] 도 12는 하나의 눈에 대한 안구주위 구역의 예시적인 이미지를 예시한다.

[0018] 도 13은 안구주위 구역의 컨퍼메이션 및 오디오 입력을 하부 얼굴의 컨퍼메이션과 상관시킬 수 있는 머신 러닝 유도 모델을 트레이닝하는 예를 예시한다. 도 14는, 얼굴의 부분이 웨어러블 디바이스에 의해 관찰되지 않을 때 얼굴의 이미지를 생성하는 예를 예시한다.

[0019] 도 15a는 오디오 입력에서의 안구주위 이미지들 및/또는 오디오 입력의 사운드들로부터 하부 얼굴의 이미지들로의 맵핑을 생성하는 예시적인 프로세스를 설명한다.

[0020] 도 15b는 입력으로 안구주위 구역의 얼굴 파라미터들 및/또는 음소들 및 출력으로서 하부 얼굴 파라미터들을 갖는 맵핑을 트레이닝하는 다른 예시적인 프로세스를 예시한다.

[0021] 도 16a는 얼굴의 부분이 웨어러블 디바이스에 의해 관찰되지 않을 때 전체 얼굴 이미지를 합성하는 예시적인 프로세스를 설명한다.

[0022] 도 16b는 제1 사용자의 전체 얼굴 이미지를 제2 사용자의 웨어러블 디바이스에 통신하는 예를 예시한다.

[0023] 도면들 전체에 걸쳐, 참조 번호들은 참조된 엘리먼트들 사이의 대응성(correspondence)을 표시하는 데 재사용될 수 있다. 도면들은 본원에서 설명된 예시적인 실시예들을 예시하기 위해 제공되며 본 개시내용의 범위를 제한하려는 의도는 아니다.

발명을 실시하기 위한 구체적인 내용

[0007]

개요

[0008]

[0024] AR/VR/MR 환경에서, 웨어러블 디바이스(이를테면, 머리-장착 디바이스)는 3차원(3D) 공간에 가상 이미지를 제시할 수 있다. 가상 이미지는 사용자의 3D 가상 아바타를 포함할 수 있다. 3D 가상 아바타는 사용자의 얼굴의 이미지를 이용하여 생성될 수 있다. 그러나, 사용자가 웨어러블 디바이스를 착용할 때, 얼굴의 부분(예컨대, 눈들 주위의 안구주위 구역)은 웨어러블 디바이스에 의해 폐색될 수 있다. 예컨대, 이미지가 외부 카메라 또는 외향 이미징 시스템을 사용하여 획득되는 경우, 안구주위 구역의 움직임들(예컨대, 눈 또는 피부 움직임들)이 이미징되지 않을 수 있다. 다른 예로서, 이미지가 내향 눈 카메라를 사용하여 획득되는 경우, 하부 얼굴 구역은 눈 카메라에 의해 이미징되지 않을 수 있다. 결과적으로, 사용자가 웨어러블 디바이스를 착용 가능한 동안(이를테면, 사용자가 말할 때) 사용자의 얼굴 표정들을 포함하도록 사용자의 얼굴의 3D 아바타를 업데이트하는 것이 어려울 수 있다.

- [0009] [0025] 본원에서 설명된 시스템들 및 방법들은 카메라에 의해 관찰될 수 있는 얼굴의 구역의 이미지들을 사용하여 얼굴의 관찰 불가능한 구역의 외관을 추론함으로써 이러한 난제를 적어도 해결하는 것에 관한 것이다. 웨어러블 디바이스는 이미징된 구역 및 추론된 구역을 사용하여 전체 얼굴 이미지를 합성할 수 있다. 웨어러블 디바이스는 안구주위 구역의 관찰 가능한 컨퍼메이션과 하부 얼굴의 추론된 컨퍼메이션 사이의 맵핑을 사용하여 관찰 불가능한 구역을 추론할 수 있다. 맵핑은 입력들, 이를테면, 사용자에게 의해 만들어진 사운드들(예컨대, 사용자의 스피치에서의 음소들), 사운드들에 대응하는 얼굴 컨퍼메이션들(예컨대, 음소와 연관된 비스듬들), 사용자의 안구주위 구역의 이미지들, 또는 하부 얼굴의 관찰된 이미지들을 단독으로 또는 조합하여 사용하여 생성될 수 있다. 컨퍼메이션들은 얼굴 모델의 얼굴 파라미터들, 사용자의 얼굴의 특징들, 또는 사용자의 얼굴에 대해 생성된 이미지들 등을 포함할 수 있다. 사용자가 말하는 동안, 웨어러블 디바이스는 오디오 출력을 감지하고, 음소들을 식별하고, 오디오 입력에서 감지된 음소들로부터 대응하는 비스듬들의 맵핑을 사용하여 비스듬들을 추론할 수 있다.
- [0010] [0026] 관찰된 이미지들은 전체 얼굴 이미지를 생성하도록 추론된 이미지들과 결합될 수 있다. 전체 얼굴 이미지는 AR/VR/MR 경험에서 사용자의 가상 아바타를 동적으로 업데이트하는 데 사용될 수 있다. 얼굴 또는 뉴럴 네트워크에 대한 DLM(deformable linear model)(또는 다른 형상 또는 텍스처 모델)은 맵핑을 결정하고, 트레이닝하거나 적용하는 데 사용될 수 있다.
- [0011] 웨어러블 시스템의 3D 디스플레이의 예들
- [0012] [0027] 웨어러블 시스템(본원에서 AR(augmented reality) 시스템으로 또한 지칭됨)은 2D 또는 3D 가상 이미지들을 사용자에게 제시하도록 구성될 수 있다. 이미지들은 정지 이미지들, 비디오의 프레임들 또는 비디오, 이들의 조합 등일 수 있다. 웨어러블 시스템의 적어도 일부는 사용자 상호작용을 위해 VR, AR 또는 MR 환경을 단독으로 또는 조합하여 제시할 수 있는 웨어러블 디바이스 상에서 구현될 수 있다. 웨어러블 디바이스는 ARD(AR device)로서 상호 교환 가능하게 사용되는 HMD(head-mounted device)일 수 있다. 또한, 본 개시내용의 목적을 위해, "AR"이란 용어는 "MR"이란 용어와 상호 교환 가능하게 사용된다.
- [0013] [0028] 도 1은 사람에 의해 관찰되는 소정의 물리적 객체들 및 소정의 가상 현실 객체들을 갖는 혼합 현실 시나리오의 예시를 도시한다. 도 1에서, MR 장면(100)이 도시되며, 여기서 MR 기술의 사용자는 배경에 있는 사람들, 나무들, 빌딩들, 및 콘크리트 플랫폼(120)을 특징(featuring)하는 실세계 공원-형 세팅(110)을 본다. 이들 아이템들에 더하여, MR 기술의 사용자는 또한, 그가 실세계 플랫폼(120) 상에 서 있는 로봇 동상(130), 및 호박벌의 의인화인 것으로 보여지는 날고 있는 만화-형 아바타 캐릭터(140)를 보는 것을 지각하더라도, 이들 엘리먼트들은 실세계에 존재하지 않는다.
- [0014] [0029] 3D 디스플레이가 진정한 깊이감(sensation of depth) 및 보다 구체적으로, 시물레이팅된 표면 깊이감을 생성하기 위해, 디스플레이의 시계(visual field)의 각각의 포인트가 그의 가상 깊이에 대응하는 원근조절 응답을 생성하는 것이 바람직할 수 있다. 디스플레이 포인트에 대한 원근조절 응답이 수렴 및 입체시(stereopsis)의 양안 깊이 단서들에 의해 결정된 바와 같은 그 포인트의 가상 깊이에 대응하지 않는 경우, 인간의 눈은 원근조절 충동을 경험할 수 있어, 불안정한 이미징, 유해한 눈의 피로, 두통들, 그리고 원근조절 정보의 부재 시에, 표면 깊이의 거의 완전한 결여를 초래할 수 있다.
- [0015] [0030] VR, AR 및 MR 경험들은 복수의 깊이 평면들에 대응하는 이미지들이 뷰어에게 제공되는 디스플레이들을 갖는 디스플레이 시스템들에 의해 제공될 수 있다. 이미지들은 각각의 깊이 평면마다 상이할 수 있고(예컨대, 장면 또는 객체의 약간 다른 프리젠테이션들을 제공함) 뷰어의 눈들에 의해 별개로 포커싱될 수 있어서, 상이한 깊이 평면 상에 로케이팅되는 장면에 대한 상이한 이미지 특징들에 포커스를 맞추도록 요구되는 눈의 원근조절에 기초하여 또는 상이한 깊이 평면들 상의 상이한 이미지 특징들이 아웃 포커스(out of focus)되는 것을 관찰하는 것에 기초하여 깊이 단서들을 사용자에게 제공하는 것을 돕는다. 본원의 다른 곳에서 논의된 바와 같이, 이러한 깊이 단서들은 깊이의 신뢰할 수 있는 지각들을 제공한다.
- [0016] [0031] 도 2는 AR/VR/MR 장면을 제공하도록 구성될 수 있는 웨어러블 시스템(200)의 예를 예시한다. 웨어러블 시스템(200)은 또한 AR 시스템(200)으로서 지칭될 수 있다. 웨어러블 시스템(200)은 디스플레이(220), 및 디스플레이(220)의 기능을 지원하기 위한 다양한 기계적 및 전자적 모듈들 및 시스템들을 포함한다. 디스플레이(220)는 사용자, 착용자 또는 뷰어(210)에 의해 착용 가능한 프레임(230)에 커플링될 수 있다. 디스플레이(220)는 사용자(210)의 눈들 앞에 포지셔닝될 수 있다. 디스플레이(220)는 AR/VR/MR 콘텐츠를 사용자에게 제시할 수 있다. 디스플레이(220)는 사용자의 머리 상에 착용되는 머리 장착 디스플레이를 포함할 수 있다. 머리-장착 디스플레이는 HUD(heads-up display)를 통해 지각되는 바와 같은 사용자의 시야 내의 미리 결정된 위치

들에 가상 정보를 디스플레이할 수 있는 HUD일 수 있다. 머리-장착 디스플레이는 또한, 가상 객체들이 실세계 객체들과 유사하게 나타나도록 (예컨대, 사용자의 관점으로부터) 관점 정정(perspective correct) 방식으로 3D 객체들을 사용자의 환경 내로 렌더링할 수 있는 SAR(spatial augmented reality) 디스플레이일 수 있다. 가상 객체들을 렌더링하는 데 사용되는 관점은 렌더링 시점으로서 또한 지칭될 수 있다.

[0017] [0032] 일부 실시예들에서, 스피커(240)는 프레임(230)에 커플링되고 사용자의 외이도에 인접하게 포지셔닝된다(일부 실시예들에서, 도시되지 않은 다른 스피커가 사용자의 다른 외이도에 인접하게 포지셔닝되어 스테레오/성형 가능(shapeable) 사운드 제어를 제공함). 디스플레이(220)는 환경으로부터 오디오 스트림을 검출하고 주변 사운드를 캡처하기 위한 오디오 센서(232)(예컨대, 마이크로폰)를 포함할 수 있다. 일부 실시예들에서, 하나 이상의 다른 오디오 센서들(도시되지 않음)이 스테레오 사운드 수신을 제공하도록 포지셔닝된다. 스테레오 사운드 수신은 사운드 소스의 위치를 결정하는 데 사용될 수 있다. 웨어러블 시스템(200)은 오디오 스트림 상에서 음성 또는 스피치 인식을 수행할 수 있다.

[0018] [0033] 웨어러블 시스템(200)은 사용자 주위의 환경의 세계를 관찰하는 외향 이미징 시스템(464)(도 4에 도시됨)을 포함할 수 있다. 웨어러블 시스템(200)은 또한 사용자의 눈 움직임들을 추적할 수 있는 내향 이미징 시스템(462)(도 4에 도시됨)을 포함할 수 있다. 내향 이미징 시스템은 어느 한쪽 눈의 움직임들 또는 양쪽 눈의 움직임들을 추적할 수 있다. 내향 이미징 시스템(462)은 프레임(230)에 부착될 수 있고, 프로세싱 모듈(260 또는 270)과 전기 통신할 수 있으며, 이 프로세싱 모듈(260 또는 270)은 예컨대, 사용자(210)의 동공 직경들 또는 눈들의 배향들, 눈 움직임들 또는 눈 포즈(pose)를 결정하도록 내향 이미징 시스템에 의해 획득된 이미지 정보를 프로세싱할 수 있다. 내향 이미징 시스템(462)은 하나 이상의 카메라들을 포함할 수 있다. 예컨대, 적어도 하나의 카메라가 각각의 눈을 이미징하는 데 사용될 수 있다. 카메라들에 의해 획득된 이미지들은 각각의 눈에 대한 동공 크기 또는 눈 포즈를 개별적으로 결정하는 데 사용될 수 있고, 그리하여 각각의 눈에 대한 이미지 정보의 프리젠테이션이 그 눈에 동적으로 맞춰질 수 있게 한다. 다른 예로서, 단지 하나의 눈의 동공 직경 또는 배향이 (예컨대, 그 눈의 이미지들을 획득하도록 구성된 카메라에 의해 획득된 이미지들에 기초하여) 결정되고, 이 눈에 대해 결정된 눈 특징들이 사용자(210)의 다른 눈과 유사한 것으로 가정된다.

[0019] [0034] 예로서, 웨어러블 시스템(200)은 외향 이미징 시스템(464) 또는 내향 이미징 시스템(462)을 사용하여 사용자의 포즈의 이미지들을 획득할 수 있다. 이미지들은 정지 이미지들, 비디오의 프레임들 또는 비디오일 수 있다.

[0020] [0035] 디스플레이(220)는 이를테면, 유선 리드 또는 무선 연결성에 의해, 다양한 구성들로 장착될 수 있는, 이를테면, 프레임(230)에 고정되게 부착되거나, 사용자에게 의해 착용된 헬멧 또는 모자에 고정되게 부착되거나, 헤드폰들에 내장되거나, 그렇지 않으면 사용자(210)에게 제거 가능하게 부착되는 (예컨대, 백팩(backpack)-스타일 구성으로, 벨트-커플링 스타일 구성으로) 로컬 데이터 프로세싱 모듈(260)에 동작 가능하게 커플링(250)될 수 있다.

[0021] [0036] 로컬 프로세싱 및 데이터 모듈(260)은 하드웨어 프로세서는 물론, 디지털 메모리 이를테면, 비-휘발성 메모리(예컨대, 플래시 메모리)를 포함할 수 있고, 이 둘 모두는 데이터의 프로세싱, 캐싱(caching) 및 저장을 보조하기 위해 활용될 수 있다. 데이터는 a) 센서들(예컨대, 프레임(230)에 동작 가능하게 커플링되거나 그렇지 않으면 사용자(210)에게 부착될 수 있음), 이를테면, 이미지 캡처 디바이스들(예컨대, 내향 이미징 시스템 또는 외향 이미징 시스템의 카메라들), 오디오 센서들(예컨대, 마이크로폰들), IMU(inertial measurement unit)들, 가속도계들, 컴퍼스(compass)들, GPS(global positioning system) 유닛들, 라디오 디바이스들, 또는 자이로스코프들로부터 캡처되고; 또는 b) 원격 프로세싱 모듈(270) 또는 원격 데이터 리포지토리(repository)(280)를 사용하여 획득 또는 프로세싱되는 (가능하게는, 이러한 프로세싱 또는 리트리벌(retrieval) 이후 디스플레이(220)에 전달하기 위한) 데이터를 포함할 수 있다. 로컬 프로세싱 및 데이터 모듈(260)은 통신 링크들(262 또는 264)에 의해, 이를테면, 유선 또는 무선 통신 링크들을 통하여, 원격 프로세싱 모듈(270) 또는 원격 데이터 리포지토리(280)에 동작 가능하게 커플링될 수 있어서, 이들 원격 모듈들은 로컬 프로세싱 및 데이터 모듈(260)에 대한 자원들로서 이용 가능하다. 또한, 원격 프로세싱 모듈(280) 및 원격 데이터 리포지토리(280)는 서로 동작 가능하게 커플링될 수 있다.

[0022] [0037] 일부 실시예들에서, 원격 프로세싱 모듈(270)은 데이터 또는 이미지 정보를 분석 및 프로세싱하도록 구성된 하나 이상의 프로세서들을 포함할 수 있다. 일부 실시예들에서, 원격 데이터 리포지토리(280)는 "클라우드" 자원 구성에서 인터넷 또는 다른 네트워킹 구성을 통하여 이용 가능할 수 있는 디지털 데이터 저장 설비를 포함할 수 있다. 일부 실시예들에서, 모든 데이터는 저장되고 모든 컴퓨테이션들은 로컬 프로세싱 및 데이터

모듈에서 수행되어, 원격 모듈로부터 완전히 자율적인 사용을 허용한다.

- [0023] [0038] 인간 시각 시스템은 복잡하고 현실적인 깊이의 지각을 제공하는 것은 난제이다. 이론에 의해 제한됨이 없이, 객체의 뷰어들은 이접운동 및 원근조절의 조합으로 인해 객체를 "3-차원"인 것으로 지각할 수 있다고 여겨진다. 서로에 대한 두 눈들의 이접운동(vergence) 움직임들(즉, 객체를 응시하기 위해 눈들의 시선들을 수렴하도록 서로를 향하는 또는 서로 멀어지는 동공들의 롤링(rolling) 움직임들)은 눈들의 렌즈들의 포커싱(또는 "원근조절")과 밀접하게 연관된다. 정상 조건들하에서, 하나의 객체로부터 상이한 거리에 있는 다른 객체로 포커스를 변화시키기 위하여, 눈들의 렌즈들의 포커스를 변화시키거나, 또는 눈들을 원근조절하는 것은 "원근조절-이접운동 반사(accommodation-vergence reflex)"로서 알려진 관계하에서, 동일한 거리에 대한 이접운동에서의 매칭하는 변화를 자동으로 유발할 것이다. 마찬가지로, 이접운동의 변화는 정상 조건들하에서, 원근조절의 매칭 변화를 트리거할 것이다. 원근조절과 이접운동 사이의 더 양호한 매칭을 제공하는 디스플레이 시스템들은 3-차원 이미저리의 더 현실적이고 편안한 시뮬레이션들을 형성할 수 있다.
- [0024] [0039] 도 3은 다중 깊이 평면들을 사용하여 3-차원 이미저리를 시뮬레이팅하기 위한 접근법의 양상들을 예시한다. 도 3를 참조하면, z-축 상에서 눈들(302 및 304)로부터 다양한 거리들에 있는 객체들은, 이들 객체들이 인 포커싱(in focus)되도록 눈들(302 및 304)에 의해 원근조절된다. 눈들(302 및 304)은 z-축을 따라 상이한 거리들에 있는 객체들에 포커싱을 맞추게 하는 특정 원근조절된 상태들을 취한다. 결과적으로, 특정 원근조절된 상태는 연관된 초점 거리를 갖는, 깊이 평면들(306) 중 특정한 하나의 깊이 평면과 연관되는 것으로 말해질 수 있어서, 특정 깊이 평면의 객체들 또는 객체들의 부분들은, 눈이 해당 깊이 평면에 대해 원근조절된 상태에 있을 때 인 포커싱된다. 일부 실시예들에서, 3-차원 이미저리는 눈들(302 및 304) 각각에 대해 이미지의 상이한 프리젠테이션들을 제공함으로써, 그리고 또한 깊이 평면들 각각에 대응하는 이미지의 상이한 프리젠테이션들을 제공함으로써 시뮬레이팅될 수 있다. 예시의 명확성을 위해 별개인 것으로 도시되지만, 눈들(302 및 304)의 시야들은 예컨대, z-축을 따른 거리가 증가함에 따라 겹쳐질 수 있다는 것이 인지될 것이다. 게다가, 예시의 용이함을 위해 평평한 것으로 도시되지만, 깊이 평면의 윤곽들은 물리적 공간에서 만곡될 수 있어서, 깊이 평면의 모든 특징들은 특정 원근조절된 상태에서 눈과 인 포커싱된다는 것이 인지될 것이다. 이론에 의해 제한됨이 없이, 인간 눈이 통상적으로 깊이 지각을 제공하기 위하여 유한 수의 깊이 평면들을 해석할 수 있다고 여겨진다. 결과적으로, 지각된 깊이의 매우 믿을 만한 시뮬레이션은, 눈에, 이들 제한된 수의 깊이 평면들 각각에 대응하는 이미지의 상이한 프리젠테이션들을 제공함으로써 달성될 수 있다.
- [0025] 도파관 스택 어셈블리
- [0026] [0040] 도 4는 이미지 정보를 사용자에게 출력하기 위한 도파관 스택의 예를 예시한다. 웨어러블 시스템(400)은 복수의 도파관들(432b, 434b, 436b, 438b, 4400b)을 사용하여 3-차원 지각을 눈/뇌에 제공하기 위하여 활용될 수 있는 도파관들의 스택, 또는 스택된 도파관 어셈블리(480)를 포함한다. 일부 실시예들에서, 웨어러블 시스템(400)은 도 2의 웨어러블 시스템(200)에 대응할 수 있고, 도 4는 그 웨어러블 시스템(200)의 일부 부분들을 더 상세히 개략적으로 보여준다. 예컨대, 일부 실시예들에서, 도파관 어셈블리(480)는 도 2의 디스플레이(220)에 통합될 수 있다.
- [0027] [0041] 도 4를 계속 참조하면, 도파관 어셈블리(480)는 또한 도파관들 사이에 복수의 특징들(458, 456, 454, 452)을 포함할 수 있다. 일부 실시예들에서, 특징들(458, 456, 454, 452)은 렌즈들일 수 있다. 다른 실시예들에서, 특징들(458, 456, 454, 452)은 렌즈들이 아닐 수 있다. 오히려, 이들은 단순히 스페이서들(예컨대, 에어 갭들을 형성하기 위한 클래딩 층들 또는 구조들)일 수 있다.
- [0028] [0042] 도파관들(432b, 434b, 436b, 438b, 440b) 또는 복수의 렌즈들(458, 456, 454, 452)은 다양한 레벨들의 파면 곡률 또는 광선 발산으로 이미지 정보를 눈에 전송하도록 구성될 수 있다. 각각의 도파관 레벨은 특정 깊이 평면과 연관될 수 있고 그 깊이 평면에 대응하는 이미지 정보를 출력하도록 구성될 수 있다. 이미지 주입 디바이스들(420, 422, 424, 426, 428)은 이미지 정보를 도파관들(440b, 438b, 436b, 434b, 432b)에 주입하기 위하여 활용될 수 있고, 이 도파관들 각각은 (도 3의 눈(304)에 대응할 수 있는) 눈(410)을 향하여 출력하도록, 각각의 개별 도파관을 가로질러 인입 광을 분산시키도록 구성될 수 있다. 광은 이미지 주입 디바이스들(420, 422, 424, 426, 428)의 출력 표면을 나가고 도파관들(440b, 438b, 436b, 434b, 432b)의 대응하는 입력 예지에 주입된다. 일부 실시예들에서, 단일 광 빔(예컨대, 시준된 빔)은 특정 도파관과 연관된 깊이 평면에 대응하는 특정 각도들(및 발산의 양들)로 눈(410)을 향하여 지향되는 시준된 클론 빔(cloned collimated beam)들의 전체 필드를 출력하기 위하여 각각의 도파관으로 주입될 수 있다.
- [0029] [0043] 일부 실시예들에서, 이미지 주입 디바이스들(420, 422, 424, 426, 428)은 각각, 대응하는 도파관(440b,

438b, 436b, 434b, 432b)으로의 주입을 위한 이미지 정보를 각각 생성하는 이산 디스플레이들이다. 일부 다른 실시예들에서, 이미지 주입 디바이스들(420, 422, 424, 426, 428)은 예컨대, 이미지 정보를 하나 이상의 광학 도파관들(예컨대, 광섬유 케이블들)을 통하여 이미지 주입 디바이스들(420, 422, 424, 426, 428) 각각에 파이프(pipe)할 수 있는 단일 멀티플렉싱된 디스플레이의 출력 단부들이다.

[0030] [0044] 제어기(460)는 스택된 도파관 어셈블리(480) 및 이미지 주입 디바이스들(420, 422, 424, 426, 428)의 동작을 제어한다. 제어기(460)는 도파관들(440b, 438b, 436b, 434b, 432b)에 대한 이미지 정보의 타이밍 및 프로비전(provision)을 레귤레이팅하는 프로그래밍(예컨대, 비-일시적 컴퓨터-판독 가능 매체의 명령들)을 포함한다. 일부 실시예들에서, 제어기(460)는 단일 통합 디바이스, 또는 유선 또는 무선 통신 채널들에 의해 연결되는 분산 시스템일 수 있다. 제어기(460)는 일부 실시예들에서, 프로세싱 모듈들(260 또는 270)(도 2에 예시됨)의 부분일 수 있다.

[0031] [0045] 도파관들(440b, 438b, 436b, 434b, 432b)은 TIR(total internal reflection)에 의해 각각의 개별 도파관 내에서 광을 전파시키도록 구성될 수 있다. 도파관들(440b, 438b, 436b, 434b, 432b)은 각각 평면형이거나 다른 형상(예컨대, 곡선)을 가질 수 있으며, 주 최상부 및 최하부 표면들 및 이들 주 최상부와 최하부 표면들 사이에서 연장되는 에지들을 갖는다. 예시된 구성에서, 도파관들(440b, 438b, 436b, 434b, 432b)은 이미지 정보를 눈(410)에 출력하기 위해 각각의 개별 도파관 내에서 전파되는 광을 도파관 밖으로 재지향시킴으로써 도파관으로부터 광을 추출하도록 구성된 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)을 각각 포함할 수 있다. 추출된 광은 아웃커플링된 광으로서 또한 지칭될 수 있고, 광 추출 광학 엘리먼트들은 또한 아웃커플링 광학 엘리먼트들로서 지칭될 수 있다. 추출된 광 빔은, 도파관 내에서 전파되는 광이 광 재지향 엘리먼트에 부딪치는 위치들에서 도파관에 의해 출력된다. 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 예컨대, 반사성 및/또는 회절성 광학 특징들일 수 있다. 설명의 용이함 및 도면 명확성을 위하여 도파관들(440b, 438b, 436b, 434b, 432b)의 최하부 주 표면들에 배치된 것으로 예시되지만, 일부 실시예들에서, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 최상부 및/또는 최하부 주 표면들에 배치될 수 있거나 도파관들(440b, 438b, 436b, 434b, 432b)의 볼륨에 직접 배치될 수 있다. 일부 실시예들에서, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 도파관들(440b, 438b, 436b, 434b, 432b)을 형성하기 위해 투명 기판에 부착된 재료 층에 형성될 수 있다. 일부 다른 실시예들에서, 도파관들(440b, 438b, 436b, 434b, 432b)은 재료의 모놀리식 피스(piece)일 수 있고 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 재료의 해당 피스의 표면 상에 그리고/또는 그 내부에 형성될 수 있다.

[0032] [0046] 도 4를 계속 참조하면, 본원에 논의된 바와 같이, 각각의 도파관(440b, 438b, 436b, 434b, 432b)은 특정 깊이 평면에 대응하는 이미지를 형성하기 위해 광을 출력하도록 구성된다. 예컨대, 눈에 가장 가까운 도파관(432b)은, 그러한 도파관(432b)에 주입된 시준된 광을 눈(410)에 전달하도록 구성될 수 있다. 시준된 광은 광학 무한대 초점 평면을 나타낼 수 있다. 다음 위의 도파관(434b)은 시준된 광이 눈(410)에 도달할 수 있기 전에 제1 렌즈(452)(예컨대, 네거티브 렌즈)를 통과하는 시준된 광을 보내도록 구성될 수 있다. 제1 렌즈(452)는 약간의 볼록한 파면 곡률을 생성하도록 구성될 수 있어서, 눈/뇌가 상기 다음 위의 도파관(434b)으로부터 오는 광을 광학 무한대로부터 눈(410)을 향해 안쪽으로 더 가까운 제1 초점 평면으로부터 오는 것으로 해석한다. 유사하게, 세번째 위의 도파관(436b)은 그 출력 광을, 눈(410)에 도달하기 전에 제1 렌즈(452) 및 제2 렌즈(454) 둘 모두를 통과시킨다. 제1 및 제2 렌즈들(452 및 454)의 결합된 광학 전력은 다른 증분 양의 파면 곡률을 생성하도록 구성될 수 있어서, 눈/뇌는 제3 도파관(436b)으로부터 오는 광을, 상기 다음 위의 도파관(434b)으로부터의 광보다는 광학 무한대로부터 사람을 향하여 안쪽으로 훨씬 더 가까운 제2 초점 평면으로부터 오는 것으로 해석한다.

[0033] [0047] 다른 도파관 층들(예컨대, 도파관들(438b, 440b)) 및 렌즈들(예컨대, 렌즈들(456, 458))은 유사하게 구성되는데, 스택에서 가장 높은 도파관(440b)은 자신의 출력을, 사람과 가장 가까운 초점 평면을 나타내는 어그리게이트 초점 전력에 대해 자신과 눈 사이의 렌즈들 모두를 통하여 전송한다. 스택된 도파관 어셈블리(480)의 다른 측 상에서 세계(470)로부터 오는 광을 보거나/해석할 때 렌즈들(458, 456, 454, 452)의 스택을 보상하기 위하여, 보상 렌즈 층(430)은 아래의 렌즈 스택(458, 456, 454, 452)의 어그리게이트 파워를 보상하기 위하여 스택의 최상부에 배치될 수 있다. 이러한 구성은 이용 가능한 도파관/렌즈 쌍들이 존재하는 만큼 많은 지각된 초점 평면들을 제공한다. 도파관들의 광 추출 광학 엘리먼트들 및 렌즈들의 포커싱 양상들 둘 모두는 정적(예컨대, 동적이 아니거나 전자-활성이 아님)일 수 있다. 일부 대안적인 실시예들에서, 어느 하나 또는 둘 모두는 전자-활성 특징들을 사용하여 동적일 수 있다.

[0034] [0048] 도 4를 계속 참조하면, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 자신의 개별 도파

관들로부터 광을 재지향하는 것은 물론, 도파관과 연관된 특정 깊이 평면에 대해 적절한 양의 발산 또는 시준으로 이 광을 출력하도록 구성될 수 있다. 결과로서, 상이한 연관된 깊이 평면들을 가진 도파관들은 상이한 구성들의 광 추출 광학 엘리먼트들을 가질 수 있고, 이러한 광 추출 광학 엘리먼트들은 연관된 깊이 평면에 따라 상이한 양의 발산으로 광을 출력한다. 일부 실시예들에서, 본원에 논의된 바와 같이, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 특정 각도들로 광을 출력하도록 구성될 수 있는 볼류메트릭(volumetric) 또는 표면 특징들일 수 있다. 예컨대, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 볼륨 홀로그램들, 표면 홀로그램들, 및/또는 회절 격자들일 수 있다. 광 추출 광학 엘리먼트들, 이를테면, 회절 격자들은 2015년 6월 25일에 공개된 미국 특허 공개 번호 제2015/0178939호에 설명되며, 이는 그 전체가 본원에 인용에 의해 포함된다.

[0035] [0049] 일부 실시예들에서, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 회절 패턴을 형성하는 회절 특징들 또는 "회절 광학 엘리먼트"(또한 본원에서 "DOE"로서 지칭됨)이다. 바람직하게는, DOE는 비교적 낮은 회절 효율성을 가져서, 빔의 광의 일부만이 DOE의 각각의 교차의 경우 눈(410)을 향하여 편향되지만, 나머지는 내부 전반사를 통하여 도파관을 통해 계속 이동한다. 따라서, 이미지 정보를 전달하는 광은 다수의 위치들에서 도파관을 나가는 다수의 관련된 출사 빔들로 분할될 수 있고 그 결과는 이런 특정 시준된 빔이 도파관 내에서 이리저리 바운딩되기 때문에 눈(304)을 향하는 상당히 균일한 출사 방출 패턴이다.

[0036] [0050] 일부 실시예들에서, 하나 이상의 DOE들은, 그것들이 활발하게 회절시키는 "온" 상태와 그것들이 크게 회절시키지 않는 "오프" 상태 사이에서 스위칭 가능할 수 있다. 예컨대, 스위칭 가능 DOE는, 마이크로액체들이 호스트 매질에서 회절 패턴을 포함하는 폴리머 분산형 액정 층을 포함할 수 있고, 마이크로액체들의 굴절률은 호스트 재료의 굴절률에 실질적으로 매칭하도록 스위칭될 수 있거나(이 경우에 패턴은 입사 광을 현저하게 회절시키지 않음) 또는 마이크로액체는 호스트 매질의 인덱스에 매칭하지 않는 인덱스로 스위칭될 수 있다(이 경우 패턴은 입사 광을 활발하게 회절시킴).

[0037] [0051] 일부 실시예들에서, 피사계 심도 또는 깊이 평면들의 수 및 분포는 뷰어의 눈들의 동공 크기들 또는 배향들에 기초하여 동적으로 변동될 수 있다. 피사계 심도는 뷰어의 동공 크기와 반대로 변경될 수 있다. 결과적으로, 뷰어의 눈들의 동공들의 크기들이 감소함에 따라, 피사계 심도가 증가하여서, 식별 불가능한 하나의 평면의 위치가 눈의 포커스의 깊이를 넘어서기 때문에 그 평면이 식별 가능하게 되고 동공 크기의 감소를 통해 보다 인 포커스로 나타나고 피사계 심도의 증가와 상응할 수 있다. 마찬가지로, 뷰어에게 상이한 이미지들을 제시하는 데 사용되는 이격된 깊이 평면들의 수는 감소된 동공 크기에 따라 감소될 수 있다. 예컨대, 뷰어는 하나의 깊이 평면으로부터 벗어나게 그리고 다른 깊이 평면으로 눈의 원근조절을 조정하지 않고서는, 하나의 동공 크기에서 제1 깊이 평면 및 제2 깊이 평면 둘 모두의 세부사항들을 명확하게 지각할 수 없을 수 있다. 그러나, 이러한 2개의 깊이 평면들은 원근조절을 변경하지 않고도 다른 동공 크기에서 사용자에게 동시에 충분히 인 포커스될 수 있다.

[0038] [0052] 일부 실시예들에서, 디스플레이 시스템은 동공 크기 또는 배향의 결정들에 또는 특정 동공 크기 또는 배향을 나타내는 전기 신호들의 수신에 기초하여 이미지 정보를 수신하는 도파관들의 수를 변동시킬 수 있다. 예컨대, 사용자의 눈들이 2개의 도파관들과 연관된 2개의 깊이 평면들을 구별할 수 없는 경우, 제어기(460)(로컬 프로세싱 및 데이터 모듈(260)의 실시예일 수 있음)은 이들 도파관들 중 하나에 이미지 정보를 제공하는 것을 중단하도록 구성되거나 프로그래밍될 수 있다. 유리하게는, 이는 시스템 상의 프로세싱 부담을 감소시킬 수 있고, 그리하여 시스템의 응답성을 증가시킨다. 도파관에 대한 DOE들이 온 및 오프 상태들 사이에서 스위칭 가능한 실시예들에서, 도파관이 이미지 정보를 수신할 때 DOE들은 오프 상태로 스위칭될 수 있다.

[0039] [0053] 일부 실시예들에서, 출사 빔이 뷰어의 눈의 직경 미만인 직경을 갖는 조건을 충족시키는 것이 바람직할 수 있다. 그러나 이 조건을 충족시키는 것은 뷰어의 동공들의 크기의 변동성을 고려하면 난제가될 수 있다. 일부 실시예들에서, 이 조건은 뷰어의 동공의 크기의 결정에 대한 응답으로 출사 빔의 크기를 변동시킴으로써 광범위한 동공 크기들에 걸쳐 충족된다. 예컨대, 동공 크기가 감소함에 따라, 출사 빔의 크기가 또한 감소할 수 있다. 일부 실시예들에서, 출사 빔 크기는 가변 어퍼처를 사용하여 변동될 수 있다.

[0040] [0054] 웨어러블 시스템(400)은 세계(470)의 일부를 이미지화하는 외향 이미징 시스템(464)(예컨대, 디지털 카메라)을 포함할 수 있다. 이러한 세계(470)의 일부는 세계 카메라의 FOV(field of view)로서 지칭될 수 있고, 이미징 시스템(464)은 때때로 FOV 카메라로서 지칭된다. 세계 카메라의 FOV는, 정해진 시간에 뷰어(210)가 지각하는 세계(470)의 부분을 포함하는, 뷰어(210)의 FOV와 동일할 수 있거나 동일하지 않을 수 있다. 예컨대, 일부 상황들에서, 세계 카메라의 FOV는 웨어러블 시스템(400)의 뷰어(210)의 뷰어(210)보다 클 수 있다. 뷰어

에 의한 이미징 또는 뷰잉을 위해 이용 가능한 전체 구역은 FOR(field of regard)로서 지칭될 수 있다. 착용자가 자신의 신체, 머리 또는 눈들을 움직여 실질적으로 공간의 임의의 방향을 지각할 수 있기 때문에 FOR은 웨어러블 시스템(400)을 둘러싸는 4π 스테라디안(steradian)들의 입체각을 포함할 수 있다. 다른 맥락들에서, 착용자의 움직임들은 더욱 억제될 수 있고, 따라서 착용자의 FOR은 더 작은 입체각을 한정할 수 있다(subtend). 외향 이미징 시스템(464)으로부터 획득된 이미지들은 사용자에게 의해 행해진 제스처들(예컨대, 손 또는 손가락 제스처들)을 추적하고, 사용자 앞의 세계(470)의 객체들을 검출하는 등을 행하는 데 사용될 수 있다.

[0041] [0055] 웨어러블 시스템(400)은 주변 사운드를 캡처하기 위한 오디오 센서(232), 예컨대, 마이크로폰을 포함할 수 있다. 위에서 설명된 바와 같이, 일부 실시예들에서, 하나 이상의 다른 오디오 센서들은 스피치 소스의 위치 결정에 유용한 스테레오 사운드 수신을 제공하도록 포지셔닝될 수 있다. 오디오 센서(232)는 다른 예로서, 오디오 소스가 로케이팅하는 곳에 대한 그러한 유용한 지향성 정보를 또한 제공할 수 있는 지향성 마이크로폰을 포함할 수 있다. 웨어러블 시스템(400)은, 스피치의 소스를 로케이팅하는 데 있어 또는 특정 시간 순간의 활성 스피커를 결정하는 등을 위해 외향 이미징 시스템(464) 및 오디오 센서(230) 둘 모두로부터의 정보를 사용할 수 있다. 예컨대, 웨어러블 시스템(400)은, 스피커의 아이덴티티를 결정하기 위해 음성 인식을 단독으로, 또는 (예컨대, 미리에서 보여지는) 스피커의 반사된 이미지와 함께 사용할 수 있다. 다른 예로서, 웨어러블 시스템(400)은 지향성 마이크로폰들로부터 획득된 사운드에 기초하여 환경 내의 스피커의 포지션을 결정할 수 있다. 웨어러블 시스템(400)은 스피치의 콘텐츠를 결정하기 위해 스피치 인식 알고리즘들로 스피커의 포지션으로부터 나오는 사운드를 파싱(parse)하고 스피커의 아이덴티티(예컨대, 이름 또는 다른 인구학적 정보)를 결정하기 위해 음성 인식 기술들을 사용할 수 있다.

[0042] [0056] 웨어러블 시스템(400)은 또한 눈 움직임들 및 얼굴 움직임들과 같은 사용자의 움직임들을 관찰하는 내향 이미징 시스템(466)(예컨대, 디지털 카메라)을 포함할 수 있다. 내향 이미징 시스템(466)은 눈(304)의 동공의 크기 및/또는 배향을 결정하기 위해 눈(410)의 이미지들을 캡처하는 데 사용될 수 있다. 내향 이미징 시스템(466)은, (예컨대, 홍채 식별을 통해) 사용자의 바이오메트릭 식별(biometric identification)을 위한 이미지들 또는 사용자가 바라보는 방향(예컨대, 눈 포즈)을 결정하는 데 사용하기 위한 이미지들을 획득하는 데 사용될 수 있다. 일부 실시예들에서, 독립적으로, 각각의 눈의 동공 크기 또는 눈 포즈를 별개로 결정하고, 그리하여 각각의 눈에 대한 이미지 정보의 프리젠테이션이 그 눈에 동적으로 맞춰지도록 허용하기 위해, 각각의 눈마다 적어도 하나의 카메라가 활용될 수 있다. 일부 다른 실시예들에서, (예컨대, 한 쌍의 눈들 당 단지 단일 카메라만을 사용하여) 단지 한쪽 눈(410)의 동공 직경 또는 배향이 결정되고 사용자의 양 눈들에 대해 유사한 것으로 가정된다. 내향 이미징 시스템(466)에 의해 획득된 이미지들은 어떤 청각적 또는 시각적 콘텐츠가 사용자에게 제시되어야 하는지를 결정하기 위해 웨어러블 시스템(400)에 의해 사용될 수 있는 사용자의 눈 포즈 또는 분위기(mood)를 결정하도록 분석될 수 있다. 웨어러블 시스템(400)은 또한 IMU들, 가속도계들, 자이로스코프들 등과 같은 센서들을 사용하여 머리 포즈(예컨대, 머리 포지션 또는 머리 배향)를 결정할 수 있다.

[0043] [0057] 웨어러블 시스템(400)은, 사용자가 웨어러블 시스템(400)과 상호작용하도록 제어기(460)에 커맨드들을 입력할 수 있게 하는 사용자 입력 디바이스(466)를 포함할 수 있다. 예컨대, 사용자 입력 디바이스(466)는 트랙패드, 터치스크린, 조이스틱, 다중 DOF(degree-of-freedom) 제어기, 용량성 감지 디바이스, 게임 제어기, 키보드, 마우스, 방향 패드(D-패드), 완드(wand), 햅틱 디바이스, 토템(예컨대, 가상 사용자 입력 디바이스로서 기능함) 등을 포함할 수 있다. 다중-DOF 제어기는 그 제어기의 일부 또는 모든 가능한 병진운동(translation)들(예컨대, 좌/우, 전/후 또는 위/아래) 또는 회전들(예컨대, 요(yaw), 피치 또는 롤)로 사용자 입력을 감지할 수 있다. 병진운동 움직임들을 지원하는 다중-DOF 제어기는 3DOF로서 지칭될 수 있는 반면, 병진운동들 및 회전들을 지원하는 다중-DOF 제어기는 6DOF로서 지칭될 수 있다. 일부 경우들에서, 사용자는 손가락(예컨대, 엄지 손가락)을 사용하여, 웨어러블 시스템(400)에 입력을 제공하도록(예컨대, 웨어러블 시스템(400)에 의해 제공되는 사용자 인터페이스에 사용자 입력을 제공하도록) 터치-감지 입력 디바이스를 누르거나 스와이프(swipe)할 수 있다. 사용자 입력 디바이스(466)는 웨어러블 시스템(400)의 사용 동안 사용자의 손에 의해 보유될 수 있다. 사용자 입력 디바이스(466)는 웨어러블 시스템(400)과 유선 또는 무선 통신할 수 있다.

[0044] [0058] 도 5는 도파관에 의해 출력된 출사 빔들의 예를 도시한다. 하나의 도파관이 예시되지만, 도파관 어셈블리(480) 내의 다른 도파관들이 유사하게 기능할 수 있다는 것이 인지될 것이며, 여기서 도파관 어셈블리(480)는 다수의 도파관들을 포함한다. 광(520)은 도파관(432b)의 입력 에지(432c)에서 도파관(432b)으로 주입되고 TIR에 의해 도파관(432b) 내에서 전파된다. 광(520)이 DOE(432a)에 충돌하는 포인트들에서, 광의 일부는 출사 빔들(510)로서 도파관을 나간다. 출사 빔들(510)은 실질적으로 평행한 것으로 예시되지만, 이들 출사 빔들(510)은 또한 도파관(432b)과 연관된 깊이 평면에 따라, 임의의 각도로 눈(410)으로 전파되도록 재지향될 수 있

다(예컨대, 발산하는 출사 빔들을 형성함). 실질적으로 평행한 출사 빔들은, 눈(410)으로부터 먼 거리(예컨대, 광학적 무한대)에 있는 깊이 평면 상에 세팅된 것으로 보이는 이미지들을 형성하도록 광을 아웃커플링하는 광 추출 광학 엘리먼트들을 갖는 도파관을 나타낼 수 있다는 것이 인지될 것이다. 다른 도파관들 또는 광 추출 광학 엘리먼트들의 다른 세트들은 더 발산하는 출사 빔 패턴을 출력할 수 있고, 이는 눈(410)이 망막 상에 포커싱을 맞추게 하기 위해 더 가까운 거리로 원근조절하는 것을 요구할 것이고 광학적 무한대보다 눈(410)에 더 가까운 거리로부터의 광으로서 뇌에 의해 해석될 것이다.

[0045] [0059] 도 6은, 다중-초점 볼류메트릭 디스플레이, 이미지 또는 광 필드의 생성에 사용되는, 도파관 장치, 도파관 장치로 또는 도파관 장치로부터의 광을 광학적으로 커플링하는 광학 커플러 서브시스템, 및 제어 서브시스템을 포함하는 광학 시스템을 도시하는 개략도이다. 광학 시스템은 도파관 장치, 도파관 장치로 또는 상기 도파관 장치로부터의 광을 광학적으로 커플링하는 광학 커플러 서브시스템, 및 제어 서브시스템을 포함할 수 있다. 광학 시스템은 다중-초점 볼류메트릭, 이미지 또는 광 필드를 생성하는 데 사용할 수 있다. 광학 시스템은 하나 이상의 주 평면 도파관들(632a)(도 6에서 단지 하나만 도시됨) 및 주 도파관들(632a) 중 적어도 일부의 주 도파관들 각각과 연관된 하나 이상의 DOE들(632b)을 포함할 수 있다. 평면 도파관들(632b)은 도 4를 참조하여 논의된 도파관들(432b, 434b, 436b, 438b, 440b)과 유사할 수 있다. 광학 시스템은 제1 축(도 6의 뷰에서 수직 또는 Y-축)을 따라 광을 중계하고 제1 축(예컨대, Y-축)을 따라 광의 유효 출사동을 확장시키기 위해 분배 도파관 장치를 사용할 수 있다. 분배 도파관 장치는, 예컨대, 분배 평면 도파관(622b) 및 분배 평면 도파관(622b)과 연관된 적어도 하나의 DOE(622a)(이중 일점 채선으로 예시됨)를 포함할 수 있다. 분배 평면 도파관(622b)은 그와 상이한 배향을 갖는 주 평면 도파관(632b)과 적어도 일부 면들에서 유사하거나 동일할 수 있다. 마찬가지로, 적어도 하나의 DOE(622a)는 DOE(632a)와 적어도 일부 면들에서 유사하거나 동일할 수 있다. 예컨대, 분배 평면 도파관(622b) 또는 DOE(622a)는 각각, 주 평면 도파관(632b) 또는 DOE(632a)와 동일한 재료들로 구성될 수 있다. 도 6에 도시된 광학 디스플레이 시스템(600)의 실시예들은 도 2에 도시된 웨어러블 시스템(200)에 통합될 수 있다.

[0046] [0060] 중계된 그리고 출사동 확장된 광은 분배 도파관 장치로부터 하나 이상의 주 평면 도파관들(632b)로 광학적으로 커플링될 수 있다. 주 평면 도파관(632b)은 바람직하게는, 제1 축과 직교하는 제2 축(예컨대, 도 6의 뷰에서 수평 또는 X-축)을 따라 광을 중계할 수 있다. 특히, 제2 축은 제1 축에 대해 비-직교 축일 수 있다. 주 평면 도파관(632b)은 제2 축(예컨대, X-축)을 따라 광의 유효 출사동을 확장시킨다. 예컨대, 분배 평면 도파관(622b)은 수직 또는 Y-축을 따라 광을 중계 및 확장시키고, 수평 또는 X-축을 따라 광을 중계 및 확장시킬 수 있는 주 평면 도파관(632b)으로 그 광을 전달할 수 있다.

[0047] [0061] 광학 시스템은 단일 모드 광섬유(640)의 근위 단부에 광학적으로 커플링될 수 있는 컬러 광(예컨대, 적색, 녹색 및 청색 레이저 광)의 하나 이상의 소스들(610)을 포함할 수 있다. 광 섬유(640)의 원위 단부는 압전 재료의 중공 튜브(642)를 통해 수용되거나 스레딩(thread)될 수 있다. 원위 단부는 픽스-프리(fixed-free) 가요성 캔틸레버(644)로서 튜브(642)로부터 돌출한다. 압전 튜브(642)는 4개의 쿼드런트(quadrant) 전극들(예시되지 않음)과 연관될 수 있다. 전극들은 예컨대, 튜브(642)의 외부, 외부 표면 또는 외부 주변부 또는 직경 상에 도금될 수 있다. 코어 전극(예시되지 않음)은 또한 튜브(642)의 코어, 중앙, 내부 주변부 또는 내부 직경에 로케이팅될 수 있다.

[0048] [0062] 예컨대, 와이어들(660)을 통해 전기적으로 커플링된 구동 전자 장치(650)는 2개의 축들에서 압전 튜브(642)를 독립적으로 구부리기 위해 대향하는 전극 쌍들을 구동한다. 광섬유(644)의 돌출 원위 팁은 기계적 공진 모드들을 갖는다. 공진 주파수들은 광섬유(644)의 직경, 길이 및 재료 성질들에 의존할 수 있다. 섬유 캔틸레버(644)의 제1 기계적 공진 모드에 근사하게 압전 튜브(642)를 진동시킴으로써, 섬유 캔틸레버(644)는 진동하게될 수 있고, 큰 편향들을 통해 스위프(sweep)할 수 있다.

[0049] [0063] 2개의 축들에서 공진 진동을 자극함으로써, 섬유 캔틸레버(644)의 팁은 영역 필링 2-차원(2D) 스캔으로 2축 방향으로(biaxially) 스캔된다. 섬유 캔틸레버(644)의 스캔과 동기하여 광 소스(들)(610)의 세기를 변조함으로써, 섬유 캔틸레버(644)로부터 나오는 광은 이미지를 형성할 수 있다. 그러한 셋업에 대한 설명들은 그 전체가 인용에 의해 본원에 포함되는 미국 특허 공보 제2014/0003762호에서 제공된다.

[0050] [0064] 광학 커플러 서브시스템의 컴포넌트는 스캐닝 섬유 캔틸레버(644)로부터 나오는 광을 시준할 수 있다. 시준된 광은 미러 표면(mirrored surface)(648)에 의해 적어도 하나의 DOE(diffractive optical element)(622a)를 포함하는 좁은 분배 평면 도파관(622b)으로 반사될 수 있다. 시준된 광은 TIR에 의해 분배 평면 도파관(622b)을 따라 (도 6의 뷰에 대해) 수직으로 전파될 수 있고, 이렇게 하여, DOE(622a)와 반복적으로

교차한다. DOE(622a)는 바람직하게는, 낮은 회절 효율을 갖는다. 이는, 광의 프랙션(fraction)(예컨대, 10%)이 DOE(622a)와의 각각의 교차 포인트에서 더 큰 주 평면 도파관(632b)의 에지를 향해 회절되게 하고, 광의 프랙션이 TIR을 통해 분배 평면 도파관(622b)의 길이를 따라 그의 원래의 궤적을 지속하게할 수 있다.

[0051] [0065] DOE(622a)와의 각각의 교차 포인트에서, 부가적인 광이 주 도파관(632b)의 입구를 향해 회절될 수 있다. 인입 광을 다수의 아웃커플링된 세트들로 분할함으로써, 광의 출사동은 분배 평면 도파관(622b)에서 DOE(622a)에 의해 수직으로 확장될 수 있다. 분배 평면 도파관(622b) 밖으로 커플링되는 이러한 수직으로 확장된 광은 주 평면 도파관(632b)의 에지에 진입할 수 있다.

[0052] [0066] 주 도파관(632b)에 진입하는 광은 TIR을 통해 주 도파관(632b)을 따라 (도 6의 뷰에 대해) 수평으로 전파될 수 있다. 광이 TIR을 통해 주 도파관(632b)의 길이의 적어도 일부를 따라 수평으로 전파됨에 따라 광은 다수의 포인트들에서 DOE(632a)와 교차한다. DOE(632a)는 유리하게는, 광의 편향 및 포커싱 둘 모두를 생성하도록, 선형 회절 패턴 및 방사상 대칭 회절 패턴의 합인 위상 프로파일(phase profile)을 갖도록 설계 또는 구성될 수 있다. DOE(632a)는 유리하게는, 낮은 회절 효율(예컨대, 10%)을 가질 수 있어서, 빔의 광 중 일부만이 DOE(632a)의 각각의 교차에 의해 뷰어의 눈을 향해 편향되는 반면에, 광의 나머지는 TIR을 통해 주 도파관(632b)을 통해 계속 전파된다.

[0053] [0067] 전파되는 광과 DOE(632a) 사이의 각각의 교차 포인트에서, 광의 프랙션이 주 도파관(632b)의 인접한 면을 향해 회절되어, 광이 TIR을 벗어나 주 도파관(632b)의 면으로부터 나오게 한다. 일부 실시예들에서, DOE(632a)의 방사상 대칭 회절 패턴은 부가적으로 회절된 광에 포커스 레벨을 부여하여, 개별 빔의 광 파면을 성형(예컨대, 곡률을 부여함)하는 것은 물론, 설계된 포커스 레벨과 매칭하는 각도로 빔을 조종한다.

[0054] [0068] 따라서, 이들 상이한 경로들은 상이한 각도들의, 포커스 레벨에서 다수의 DOE들(632a)에 의해, 또는 출사동에서 상이한 필 패턴(fill pattern)들의 산출에 의해 광이 주 평면 도파관(632b) 밖으로 커플링되게할 수 있다. 출사동에서의 상이한 필 패턴들은 다수의 깊이 평면들을 갖는 광 필드 디스플레이를 생성하는데 유리하게 사용될 수 있다. 도파관 어셈블리의 각각의 층 또는 스택의 층들의 세트(예컨대, 3개의 층들)는 각각의 컬러(예컨대, 적색, 청색, 녹색)를 생성하는 데 사용될 수 있다. 따라서, 예컨대, 3개의 인접한 층들의 제1 세트는 제1 초점 깊이의 적색, 청색 및 녹색 광을 각각 생성하도록 사용될 수 있다. 3개의 인접한 층들의 제2 세트는 제2 초점 깊이의 적색, 청색 및 녹색 광을 각각 생성하도록 사용될 수 있다. 다수의 세트들이 다양한 초점 깊이들을 갖는 풀(full) 3D 또는 4D 컬러 이미지 광 필드를 생성하도록 사용될 수 있다.

[0055] 웨어러블 시스템의 다른 컴포넌트

[0056] [0069] 다수의 구현들에서, 웨어러블 시스템은 위에서 설명된 웨어러블 시스템의 컴포넌트들에 부가적으로 또는 대안적으로 다른 컴포넌트들을 포함할 수 있다. 웨어러블 시스템은 예컨대, 하나 이상의 햅틱 디바이스들 또는 컴포넌트들을 포함할 수 있다. 햅틱 디바이스들 또는 컴포넌트들은 사용자에게 촉감을 제공하도록 동작 가능할 수 있다. 예컨대, 햅틱 디바이스들 또는 컴포넌트들은 가상 콘텐츠(예컨대, 가상 객체들, 가상 툴들, 다른 가상 구조들)를 터치할 때 압력 또는 텍스처의 촉감을 제공할 수 있다. 촉감은 가상 객체가 표현하는 물리적 객체의 느낌(feel)을 복제할 수 있거나, 또는 가상 콘텐츠가 표현하는 이미징된 객체 또는 캐릭터(예컨대, 용)의 느낌을 복제할 수 있다. 일부 구현들에서, 햅틱 디바이스들 또는 컴포넌트들(예컨대, 사용자 웨어러블 글러브)은 사용자에 의해 착용될 수 있다. 일부 구현들에서, 햅틱 디바이스들 또는 컴포넌트들은 사용자에게 의해 보유될 수 있다.

[0057] [0070] 웨어러블 시스템은 예컨대, 웨어러블 시스템과의 상호작용 또는 입력을 허용하기 위해 사용자에게 의해 조작 가능한 하나 이상의 물리적 객체들을 포함할 수 있다. 이러한 물리적 객체들은 본원에서 토템(totem)들로서 지칭될 수 있다. 일부 토템들은 예컨대, 금속 또는 플라스틱 조각, 벽, 테이블의 표면과 같은 무생물의 형태를 취할 수 있다. 특정 구현들에서, 토템들은 실제로, 어떠한 물리적 입력 구조들(예컨대, 키들, 트리거들, 조이스틱, 트랙볼, 로커 스위치(rocker switch))도 갖지 않을 수 있다. 대신, 토템은 단순히 물리적 표면을 제공할 수 있고 웨어러블 시스템은 사용자 인터페이스가 토템의 하나 이상의 표면들 상에 있는 것으로 사용자에게 나타나게할 수 있다. 예컨대, 웨어러블 시스템은 컴퓨터 키보드 및 트랙패드의 이미지가 토템의 하나 이상의 표면들에 상주하는 것으로 나타나게할 수 있다. 예컨대, 웨어러블 시스템은 가상 컴퓨터 키보드 및 가상 트랙패드가 토템으로서 역할을 하는 얇은 직사각형 알루미늄 플레이트의 표면 상에 나타나게할 수 있다. 직사각형 플레이트 그 자체는 어떠한 물리적인 키들 또는 트랙패드 또는 센서들도 갖지 않는다. 그러나, 웨어러블 시스템은 가상 키보드 또는 가상 트랙패드를 통해 이루어진 선택들 또는 입력들로서 직사각형 플레이트와의 사용자 조작 또는 상호작용 또는 터치들을 검출할 수 있다. 사용자 입력 디바이스(466)(도 4에 도시됨)는, 트랙패드,

터치 패드, 트리거, 조이스틱, 트랙볼, 로커 또는 가상 스위치, 마우스, 키보드, 다중-자유도 제어기 또는 다른 물리적 입력 디바이스를 포함할 수 있는 토탈의 실시예일 수 있다. 사용자는 토탈을 단독으로 또는 포즈들과 함께 사용하여 웨어러블 시스템 또는 다른 사용자들과 상호작용할 수 있다.

[0058] [0071] 본 개시내용의 웨어러블 디바이스들, HMD 및 디스플레이 시스템들과 함께 사용 가능한 햅틱 디바이스들 및 토탈들의 예들은, 그 전체가 본원에 인용에 의해 포함되는 미국 특허 공보 제2015/0016777호에서 설명된다.

[0059] 웨어러블 시스템들, 환경들 및 인터페이스들의 예들

[0060] [0072] 웨어러블 시스템은 렌더링된 광 필드들에서 높은 피사계 심도를 달성하기 위해 다양한 맵핑 관련 기술들을 사용할 수 있다. 가상 세계의 맵핑 시에, 실제 세계와 관련하여 가상 객체들을 정확하게 묘사하기 위해 현실 세계의 모든 특징들 및 포인트들을 아는 것이 유리하다. 이를 위해, 웨어러블 시스템의 사용자들로부터 캡처된 FOV 이미지들이 실제 세계의 다양한 포인트들 및 특징들에 관한 정보를 전달하는 새로운 사진들을 포함함으로써 세계 모델에 추가될 수 있다. 예컨대, 웨어러블 시스템은 맵 포인트의 세트(이를테면, 2D 포인트들 또는 3D 포인트들)를 수집하고 새로운 맵 포인트들을 발견하여 세계 모델의 보다 정확한 버전을 렌더링할 수 있다. 제1 사용자의 세계 모델이 (예컨대, 클라우드 네트워크와 같은 네트워크를 통해) 제2 사용자에게 통신될 수 있어서, 제2 사용자는 제1 사용자 주위의 세계를 경험할 수 있다.

[0061] [0073] 도 7은 MR 환경(700)의 예의 블록도이다. MR 환경(700)은 하나 이상의 사용자 웨어러블 시스템들(예컨대, 웨어러블 시스템(200) 또는 디스플레이 시스템(220)) 또는 정적 룸(stationary room) 시스템들(예컨대, 룸 카메라들 등)로부터 입력(예컨대, 사용자의 웨어러블 시스템으로부터의 시각적 입력(702), 이를테면, 룸 카메라들로부터의 정적 입력(704), 다양한 센서들로부터의 감각 입력(706), 제스처들, 토탈들, 눈 추적, 사용자 입력 디바이스(466)로부터의 사용자 입력 등)을 수신하도록 구성될 수 있다. 웨어러블 시스템들은 사용자의 환경의 위치 및 다양한 다른 속성들을 결정하기 위해 다양한 센서들 (예컨대, 가속도계들, 자이로스코프들, 온도 센서들, 움직임 센서들, 깊이 센서들, GPS 센서들, 내향 이미징 시스템, 외향 이미징 시스템 등)을 사용할 수 있다. 이 정보는 추가로, 이미지들을 제공할 수 있는 룸의 정적 카메라들로부터의 정보 또는 상이한 관점으로부터의 다양한 단서들로 보완될 수 있다. 카메라들(이를테면, 룸 카메라들 및/또는 외향 이미징 시스템의 카메라들)에 의해 획득된 이미지 데이터는 맵핑 포인트들의 세트로 감소될 수 있다.

[0062] [0074] 하나 이상의 객체 인식기들(708)은 맵 데이터베이스(710)의 도움으로 수신된 데이터(예컨대, 포인트들의 모음)를 크롤링(crawl)하고, 포인트들을 인식 또는 맵핑하고, 이미지들을 태깅하고, 의미 정보를 객체들에 첨부할 수 있다. 맵 데이터베이스(710)는 시간에 걸쳐 수집된 다양한 포인트들 및 그들의 대응하는 객체들을 포함할 수 있다. 다양한 디바이스들 및 맵 데이터베이스는 클라우드에 액세스하기 위해 네트워크(예컨대, LAN, WAN 등)를 통해 서로 연결될 수 있다.

[0063] [0075] 맵 데이터베이스의 이러한 정보 및 포인트들의 모음에 기초하여, 객체 인식기들(708a 내지 708n)은 환경의 객체들을 인식할 수 있다. 예컨대, 객체 인식기들은 얼굴들, 사람들, 창문들, 벽들, 사용자 입력 디바이스들, 텔레비전들, 문서들(예컨대, 본원에서의 보안 예들에서 설명된 바와 같은 여행 티켓들, 운전 면허증, 여권), 사용자의 환경의 다른 객체들 등을 인식할 수 있다. 하나 이상의 객체 인식기들이 소정의 특성들을 갖는 객체에 대해 전문화될 수 있다. 예컨대, 객체 인식기(708a)는 얼굴들을 인식하는 데 사용될 수 있는 반면에, 다른 객체 인식기는 문서들을 인식하는 데 사용될 수 있다.

[0064] [0076] 객체 인식들은 다양한 컴퓨터 비전 기술들을 사용하여 수행될 수 있다. 예컨대, 웨어러블 시스템은 장면 재구성, 이벤트 검출, 비디오 추적, 객체 인식(예컨대, 사람들 또는 문서들), 객체 포즈 추정, (예컨대, 환경 내의 사람 또는 문서 상의 이미지로부터) 얼굴 인식, 러닝(learning), 인텍싱, 모션 추정 또는 이미지 분석(예컨대, 사진들, 서명들, 식별 정보, 여행 정보 등과 같은 문서들 내의 표시들을 식별함) 등을 수행하도록 외향 이미징 시스템(464)(도 4에 도시됨)에 의해 획득된 이미지들을 분석할 수 있다. 이러한 작업들을 수행하기 위해 하나 이상의 컴퓨터 비전 알고리즘들이 사용될 수 있다. 컴퓨터 비전 알고리즘들의 비-제한적인 예들은, SIFT(Scale-invariant feature transform), SURF(speeded up robust features), ORB(oriented FAST and rotated BRIEF), BRISK(binary robust invariant scalable keypoints), FREAK(fast retina keypoint), Viola-Jones 알고리즘, Eigenfaces 접근법, Lucas-Kanade 알고리즘, Horn-Schunk 알고리즘, Mean-shift 알고리즘, vSLAM(visual simultaneous location and mapping) 기술들, 순차적 베이지안 추정기(예컨대, 칼만 필터, 확장된 칼만 필터 등), 번들 조정, 적응형 임계화(Adaptive thresholding)(및 다른 임계화 기술들), ICP(Iterative Closest Point), SGM(Semi Global Matching), SGBM(Semi Global Block Matching), 특징 포인트 히스토그램 (Feature Point Histogram)들, (예컨대, 지원 벡터 머신, k-최근접 이웃 알고리즘, 나이브 베이즈(Naive

Bayes), 뉴럴 네트워크(콘볼루션 또는 딥 뉴럴 네트워크들을 포함함) 또는 다른 감독/비-감독 모델들 등과 같은) 다양한 머신 러닝 알고리즘들을 포함한다.

[0065] [0077] 객체 인식들은 부가적으로 또는 대안적으로, 다양한 머신 러닝 알고리즘들에 의해 수행될 수 있다. 일단 트레이닝되면, 머신 러닝 알고리즘은 HMD에 의해 저장될 수 있다. 머신 러닝 알고리즘들의 일부 예들은, (예컨대, 정규 최소 제곱화 회귀 분석(Ordinary Least Squares Regression)과 같은) 회귀 알고리즘들, (예컨대, 러닝 벡터 양자화(Learning Vector Quantization)와 같은) 인스턴스-기반 알고리즘들, (예컨대, 분류 및 회귀 트리들과 같은) 결정 트리 알고리즘들, (예컨대, 나이브 베이즈(Naive Bayes)와 같은) 베이지안 알고리즘들, (예컨대, k-평균 클러스터링과 같은) 클러스터링 알고리즘들, (예컨대, 선험적 알고리즘들과 같은) 연관 규칙 러닝 알고리즘들, (예컨대, 퍼셉트론(Perceptron)과 같은) 인공 뉴럴 네트워크 알고리즘들, (예컨대, 딥 볼츠만 머신, 또는 딥 뉴럴 네트워크와 같은) 딥 러닝 알고리즘들, (예컨대, 주 성분 분석(Principal Component Analysis)과 같은) 차원성 감소 알고리즘, (예컨대, 누적된 일반화(Stacked Generalization)와 같은) 앙상블 알고리즘들 및/또는 다른 머신 러닝 알고리즘들을 포함하는 감독 또는 비-감독 머신 러닝 알고리즘들을 포함할 수 있다. 일부 실시예들에서, 개별 모델들은 개별 데이터 세트들에 대해 커스터마이징될 수 있다. 예컨대, 웨어러블 디바이스는 기본 모델을 생성하거나 저장할 수 있다. 기본 모델은 데이터 유형(예컨대, 텔레프레전스(telepresence) 세션의 특정 사용자), 데이터 세트(예컨대, 텔레프레전스 세션의 사용자의 획득된 부가적인 이미지들의 세트), 조건부 상황들 또는 다른 변동들에 특정한 부가적인 모델들을 생성하기 위한 시작 포인트로서 사용될 수 있다. 일부 실시예들에서, 웨어러블 HMD는 어그리게이팅된 데이터의 분석을 위한 모델들을 생성하기 위해 복수의 기술들을 활용하도록 구성될 수 있다. 다른 기술들은 미리 정의된 임계치들 또는 데이터 값들을 사용하는 것을 포함할 수 있다.

[0066] [0078] 맵 데이터베이스의 이러한 정보 및 포인트들의 모음에 기초하여, 객체 인식기들(708a 내지 708n)은 객체들을 인식하고 객체들에 생명을 부여하기 위해 의미 정보로 객체들을 보충할 수 있다. 예컨대, 객체 인식기가 포인트들의 세트를 문이라고 인식하는 경우, 시스템은 일부 의미 정보를 첨부할 수 있다(예컨대, 문은 힌지(hinge)를 가지며 힌지를 중심으로 90도 움직임을 가짐). 객체 인식기가 포인트들의 세트를 미러라고 인식하는 경우, 시스템은 미러는 룸 내의 객체들의 이미지들을 반사할 수 있는 반사 표면을 갖는다는 의미 정보를 첨부할 수 있다. 의미 정보는 본원에서 설명된 바와 같이 객체들의 어포던스들(affordance)을 포함할 수 있다. 예컨대, 의미 정보는 객체의 법선을 포함할 수 있다. 시스템은 방향이 객체의 법선을 표시하는 벡터를 할당할 수 있다. 시간이 지남에 따라, 시스템(로컬로 상주할 수도 있거나, 또는 무선 네트워크를 통해 액세스 가능할 수도 있음)이 세계로부터 더 많은 데이터를 누적함에 따라 맵 데이터베이스가 성장한다. 객체들이 인식되면, 정보는 하나 이상의 웨어러블 시스템들에 전송될 수 있다. 예컨대, MR 환경(700)은 캘리포니아에서 일어나는 장면들에 관한 정보를 포함할 수 있다. 환경(700)은 뉴욕의 하나 이상의 사용자들에 전송될 수 있다. FOV 카메라로부터 수신된 데이터 및 다른 입력들에 기초하여, 객체 인식기들 및 다른 소프트웨어 컴포넌트들은 다양한 이미지들로부터 수집된 포인트들을 맵핑하고, 객체들을 인식하는 등을 수행할 수 있어서, 장면이 세계의 상이한 부분에 있을 수 있는 제2 사용자에게 정확히 "전달될" 수 있다. 환경(700)은 또한 로컬화(localization) 목적들을 위해 토폴로지 맵(topological map)을 사용할 수 있다.

[0067] [0079] 도 8은 인지된 객체들과 관련하여 가상 콘텐츠를 렌더링하는 방법(800)의 예의 프로세스 흐름도이다. 방법(800)은 가상 장면이 웨어러블 시스템의 사용자에게 어떻게 제시될 수 있는지를 설명한다. 사용자는 장면으로부터 지리적으로 멀리 있을 수 있다. 예컨대, 사용자는 뉴욕에 있을 수 있지만, 현재 캘리포니아에서 진행되고 있는 장면을 보기를 원할 수 있거나, 또는 캘리포니아에 거주하는 친구와 함께 산책을 하러 가기를 원할 수 있다.

[0068] [0080] 블록(810)에서, 웨어러블 시스템은 사용자 및 다른 사용자들로부터, 사용자의 환경에 관한 입력을 수신할 수 있다. 이는 다양한 입력 디바이스들 및 맵 데이터베이스에 이미 보유된 지식을 통해 달성될 수 있다. 사용자의 FOV 카메라, 센서들, GPS, 눈 추적 등은 블록(810)에서 시스템에 정보를 전달한다. 시스템은 블록(820)에서 이 정보에 기초하여 최소 포인트들을 결정할 수 있다. 최소 포인트들은 사용자의 주변들의 다양한 객체들의 배향 및 포지션을 디스플레이 및 이해하는 데 사용될 수 있는 포즈 데이터(예컨대, 머리 포즈, 눈 포즈, 신체 포즈 또는 손 제스처들)를 결정하는 데 사용될 수 있다. 객체 인식기들(708a-708n)은 블록(830)에서 이러한 수집된 포인트들을 크롤링하고 맵 데이터베이스를 사용하여 하나 이상의 객체들을 인식할 수 있다. 그 후, 이 정보는 블록(840)에서, 사용자의 개별 웨어러블 시스템에 전달될 수 있고, 블록(850)에서, 원하는 가상 장면이 사용자에게 상응하게 디스플레이될 수 있다. 예컨대, 원하는 가상 장면(예컨대, CA의 사용자)은 뉴욕의 사용자의 다양한 객체들 및 다른 주변들과 관련하여 적절한 배향, 포지션 등에서 디스플레이될 수 있다.

- [0069] [0081] 도 9는 웨어러블 시스템의 다른 예의 블록도이다. 이 예에서, 웨어러블 시스템(900)은 세계의 맵 데이터를 포함하는 맵 데이터베이스(710)를 포함할 수 있는 맵(920)을 포함한다. 맵은 부분적으로 웨어러블 시스템 상에 로컬로 상주할 수 있고, (예컨대, 클라우드 시스템에서) 유선 또는 무선 네트워크에 의해 액세스 가능한 네트워크링 저장 위치들에 부분적으로 상주할 수 있다. 포즈 프로세스(910)는 웨어러블 컴퓨팅 아키텍처(예컨대, 프로세싱 모듈(260) 또는 제어기(460)) 상에서 실행되고 웨어러블 컴퓨팅 하드웨어 또는 사용자의 포지션 및 배향을 결정하기 위해 맵(920)으로부터 데이터를 활용할 수 있다. 포즈 데이터는 사용자가 시스템을 경험하고 세계에서 동작할 때 즉석에서 수집된 데이터로부터 컴퓨팅될 수 있다. 데이터는 이미지들, 센서들(이를테면, 일반적으로 가속도계 및 자이로스코프 컴포넌트들을 포함하는 관성 측정 유닛들)로부터의 데이터 및 실제 또는 가상 환경의 객체들과 관련된 표면 정보를 포함할 수 있다.
- [0070] [0082] 희소 포인트 표현은 동시성 로컬화 및 맵핑(예컨대, 입력이 이미지들/시각 전용인 구성을 지칭하는 SLAM 또는 vSLAM) 프로세스의 출력일 수 있다. 시스템은 다양한 컴포넌트들이 세계 어디에 있는지 뿐만 아니라, 세상이 무엇으로 이루어져 있는지를 파악하도록 구성될 수 있다. 포즈는 맵을 업데이트하고 맵으로부터의 데이터를 사용하는 것을 포함해서, 다수의 목표들을 달성하는 빌딩 블록일 수 있다.
- [0071] [0083] 일 실시예에서, 희소 포인트 포지션은 그 자체로 완전히 충분하지 않을 수 있고, 다중 초점 AR, VR 또는 MR 경험을 생성하기 위해 추가의 정보가 필요할 수 있다. 일반적으로 깊이 맵 정보를 지칭하는 밀집된 표현들이 이 겹을 적어도 부분적으로 채우기 위해 활용될 수 있다. 이러한 정보는 스테레오(940)로서 지칭되는 프로세스로부터 컴퓨팅될 수 있으며, 여기서 깊이 정보가 삼각측량 또는 비행-시간 감지와 같은 기술을 사용하여 결정된다. 이미지 정보 및 활성 패턴들(이를테면, 활성 프로젝트들을 사용하여 생성된 적외선 패턴들), 이미지 카메라들로부터 획득된 이미지들, 또는 손 제스처들/토탈(950)은 스테레오 프로세스(940)에 대한 입력으로서 역할을 할 수 있다. 상당한 양의 깊이 맵 정보가 함께 융합될 수 있으며, 그 중 일부는 표면 표현으로 요약될 수 있다. 예컨대, 수학적으로 정의 가능한 표면들은 (예컨대, 대형 포인트 클라우드에 비해) 효율적일 수 있고, 게임 엔진들과 같은 다른 프로세싱 디바이스들에 의해 가능한 입력들일 수 있다. 따라서, 스테레오 프로세스(940)의 출력(예컨대, 깊이 맵)은 융합 프로세스(930)에서 결합될 수 있다. 포즈(910)는 또한 이 융합 프로세스(930)에 대한 입력일 수 있고, 융합(930)의 출력은 맵 프로세스(920)를 업데이트(populating)하기 위한 입력이 된다. 서브-표면들이 이를테면, 토폴로지 맵핑에서 서로 연결되어 더 큰 표면들을 형성할 수 있고 맵은 포인트들 및 표면들의 대형 혼합물이 된다.
- [0072] [0084] 혼합 현실 프로세스(960)에서의 다양한 양상들을 해결하기 위해, 다양한 입력들이 활용될 수 있다. 예컨대, 도 9에 도시된 실시예에서, 게임 파라미터들은, 시스템의 사용자가 다양한 위치들의 하나 이상의 몬스터들, 다양한 조건들(이를테면, 사용자가 몬스터를 쏘는 경우) 하에서 죽거나 도망가는 몬스터들, 다양한 위치들의 벽들 또는 다른 객체들 등을 갖는 몬스터 전투 게임을 플레이하고 있다고 결정하기 위한 입력일 수 있다. 세계 맵은 객체들의 위치에 관한 정보 또는 객체들의 의미 정보를 포함할 수 있고 세계 맵은 혼합 현실에 대한 다른 귀중한 입력일 수 있다. 세계에 대한 포즈가 또한 입력이 되며 거의 모든 상호작용 시스템에 대해 중요한 역할을 한다.
- [0073] [0085] 사용자로부터의 제어들 또는 입력들은 웨어러블 시스템(900)에 대한 다른 입력이다. 본원에서 설명된 바와 같이, 사용자 입력들은 시각적 입력, 제스처들, 토탈들, 오디오 입력, 감각 입력 등을 포함할 수 있다. 예컨대, 주위를 돌아다니거나 게임을 플레이하기 위해, 사용자는 웨어러블 시스템(900)에, 자신이 하기를 원하는 것에 관해 지시할 필요가 있을 수 있다. 단지 공간에서 자신을 움직이는 것 외에도, 활용될 수 있는 다양한 형태들의 사용자 제어들이 존재한다. 일 실시예에서, 토탈(예컨대, 사용자 입력 디바이스) 또는 객체, 이를테면, 장난감 층은 사용자에게 의해 보유되고 시스템에 의해 추적될 수 있다. 시스템은 바람직하게는, 사용자가 아이템을 보유하고 있다는 것을 인지하고 사용자가 아이템과 어떤 종류의 상호작용을 하고 있는지를 이해하도록 구성될 것이다(예컨대, 토탈 또는 객체가 총인 경우, 시스템은 위치 및 배향은 물론, IMU와 같은 센서가 장착될 수 있는 트리거 또는 다른 감지 버튼 또는 엘리먼트를 사용자가 클릭하고 있는지 여부를 이해하도록 구성될 수 있으며, 이러한 센서는 그러한 활동이, 카메라들 중 임의의 것의 시야 내에 있지 않을 때조차도 무슨 일이 일어나고 있는지를 결정하는데 도움을 줄 수 있음).
- [0074] [0086] 손 제스처 추적 또는 인식은 또한 입력 정보를 제공할 수 있다. 웨어러블 시스템(900)은 버튼 누름들에 대한 손 제스처들, 왼쪽 또는 오른쪽, 정지, 잡기, 홀드 등을 제스처링하는 손 제스처들을 추적 및 해석하도록 구성될 수 있다. 예컨대, 일 구성에서, 사용자는 비-게임 환경에서 이메일들 또는 캘린더를 훑어보거나(flip through) 다른 사람이나 플레이어와 "주먹 인사(fist bump)"를 하기를 원할 수 있다. 웨어러블 시스템(900)은 동적일 수 있거나 동적이지 않을 수 있는 최소량의 손 제스처를 레버리지(leverage)하도록 구성될 수

있다. 예컨대, 제스처들은 중지를 위한 펼쳐진 손, ok를 위한 엄지 올리기, not ok를 위한 엄지 내리기; 또는 방향성 커맨드들을 위한 우측 또는 좌측 또는 위/아래로의 손 뒤집기와 같은 단순한 정적 제스처일 수도 있다.

[0075] [0087] 눈 추적(예컨대, 특정 깊이 또는 범위에서 렌더링하도록 디스플레이 기술을 제어하기 위해 사용자가 바라보는 곳을 추적함)은 다른 입력이다. 일 실시예에서, 눈들의 이점운동은 삼각측량을 사용하여 결정될 수 있고, 그 후, 그 특정 사람에 대해 발현되는 이점운동/원근조절 모델을 사용하여, 원근조절이 결정될 수 있다. 눈 시선(예컨대, 한쪽 또는 양쪽 눈들의 방향 또는 배향)을 결정하기 위해 눈 카메라(들)에 의해 눈 추적이 수행될 수 있다. 예컨대, 눈(들) 근처에 배치된 전극들에 의한 전위들의 측정(예컨대, 전기 안구도 기록(electrooculography))과 같은 눈 추적을 위한 다른 기술들이 사용될 수 있다.

[0076] [0088] 스피치 추적은 단독으로 또는 다른 입력들(예컨대, 토템 추적, 눈 추적, 제스처 추적 등)과 함께 사용될 수 있는 다른 입력일 수 있다. 스피치 추적은 스피치 인식, 음성 인식을 단독으로 또는 조합하여 포함할 수 있다. 시스템(900)은 환경으로부터 오디오 스트림을 수신하는 오디오 센서(예컨대, 마이크로폰)를 포함할 수 있다. 시스템(900)은 누가 말하고 있는지(예컨대, 스피치가 ARD의 착용자 또는 다른 사람 또는 음성(예컨대, 환경 내의 라우드스피커에 의해 송신된 레코딩된 음성)으로부터 왔는지)를 결정하기 위한 음성 인식 기술은 물론, 무엇을 말하고 있는지를 결정하기 위한 스피치 인식 기술을 포함할 수 있다. 로컬 데이터 및 프로세싱 모듈(260) 또는 원격 프로세싱 모듈(270)은 예컨대, 히든 마르코프 모델(hidden Markov model)들, DTW(dynamic time warping)-기반 스피치 인식들, 뉴럴 네트워크들, 딥 러닝 알고리즘들, 이를테면, 딥 피드포워드 및 회귀 뉴럴 네트워크(recurrent neural network)들, 단-대-단 자동 스피치 인식들, (도 7을 참조하여 설명된) 머신 러닝 알고리즘들, 또는 음향 모델링 또는 언어 모델링을 사용하는 다른 알고리즘들 등과 같은 다양한 스피치 인식 알고리즘들을 적용함으로써 스피치의 콘텐츠를 식별하기 위해 마이크로폰으로부터의 오디오 데이터(또는, 예컨대, 사용자에 의해 시청되는 비디오 스트림과 같은 다른 스트림의 오디오 데이터)를 프로세싱할 수 있다.

[0077] [0089] 로컬 데이터 및 프로세싱 모듈(260) 또는 원격 프로세싱 모듈(270)은 또한, 스피커의 아이덴티티, 이를테면, 스피커가 웨어러블 시스템(900)의 사용자(210)인지 또는 사용자가 대화중인 다른 사람인지를 식별할 수 있는 음성 인식 알고리즘들을 적용할 수 있다. 일부 예시적인 음성 인식 알고리즘들은 주파수 추정, 숨겨진 마르코프 모델들, 가우스 혼합 모델, 패턴 매칭 알고리즘들, 뉴럴 네트워크들, 매트릭스 표현, 벡터 양자화, 스피커 분리(speaker diarisation), 결정 트리들 및 DTW(dynamic time warping) 기술을 포함할 수 있다. 음성 인식 기술들은 또한, 코호트 모델들 및 세계 모델들과 같은 반-스피커 기술들을 포함할 수 있다. 스펙트럼 특징들은 스피커 특성들을 나타내는 데 사용될 수 있다. 로컬 데이터 및 프로세싱 모듈 또는 원격 데이터 프로세싱 모듈(270)은 음성 인식을 수행하기 위해 도 7을 참조하여 설명된 다양한 머신 러닝 알고리즘들을 사용할 수 있다.

[0078] [0090] 카메라 시스템들과 관련하여, 도 9에 도시된 예시적인 웨어러블 시스템(900)은 3쌍의 카메라들, 즉 사용자의 얼굴의 측면들에 대해 배열되는 비교적 넓은 FOV 또는 수동 SLAM 쌍의 카메라들, 스테레오 이미징 프로세스(940)를 처리하기 위해 그리고 또한, 손 제스처들을 캡처하고 사용자의 얼굴 앞의 토템/객체 추적을 위해 사용자 앞에 배향된 상이한 쌍의 카메라들을 포함할 수 있다. 스테레오 프로세스(940)를 위한 FOV 카메라들 및 카메라들의 쌍은 외향 이미징 시스템(464)(도 4에 도시됨)의 부분일 수 있다. 웨어러블 시스템(900)은 눈 벡터들 및 다른 정보를 삼각 측량하기 위해 사용자의 눈들을 향해 배향되는 눈 추적 카메라들(도 4에 도시된 내향 이미징 시스템(462)의 부분일 수 있음)을 포함할 수 있다. 웨어러블 시스템(900)은 또한 장면의 텍스처를 주입하기 위해 하나 이상의 텍스처링된 광 프로젝터들(이를테면, 적외선 (IR) 프로젝터들)을 포함할 수 있다.

[0079] [0091] 도 10은 가상 사용자 인터페이스와 상호작용하기 위한 방법(1000)의 예의 프로세스 흐름도이다. 방법(1000)은 본원에서 설명된 웨어러블 시스템에 의해 수행될 수 있다. 방법(1000)은 텔레프레젠휰 세션에서 방법(1000)을 수행할 수 있다.

[0080] [0092] 블록(1010)에서, 웨어러블 시스템은 특정 UI를 식별할 수 있다. UI의 유형은 사용자에 의해 미리 결정될 수 있다. 웨어러블 시스템은 사용자 입력(예컨대, 제스처, 시각 데이터, 오디오 데이터, 감각 데이터, 직접 커맨드 등)에 기초하여 특정 UI가 과플레이트(populate)될 필요가 있다고 식별할 수 있다. UI는 텔레프레젠휰 세션에 특정적일 수 있다. 블록(1020)에서, 웨어러블 시스템은 가상 UI에 대한 데이터를 생성할 수 있다. 예컨대, 경계(confine)들, 일반적인 구조, UI의 형상 등과 연관된 데이터가 생성될 수 있다. 또한, 웨어러블 시스템은 사용자의 물리적 위치의 맵 좌표들을 결정할 수 있어서 웨어러블 시스템이 사용자의 물리적 위치와 관련하여 UI를 디스플레이할 수 있다. 예컨대, UI가 신체 중심인 경우, 웨어러블 시스템은, 링 UI가 사용자 주위에

디스플레이될 수 있거나 평면 UI가 벽 상에 또는 사용자 앞에 디스플레이될 수 있도록 사용자의 신체 스탠스, 머리 포즈 또는 눈 포즈의 좌표들을 결정할 수 있다. 텔레프레전스 맥락에서, UI는, 환경 내의 다른 사용자의 존재에 관한 실체감(tangible sense)을 생성하기 위해 UI가 사용자를 둘러싼 것처럼 디스플레이될 수 있다(예컨대, UI는 사용자 주위의 참가자들의 가상 아바타들을 디스플레이할 수 있음). UI가 손 중심인 경우, 사용자의 손들의 맵 좌표들이 결정될 수 있다. 이러한 맵 포인트들은 FOV 카메라들, 감각 입력 또는 임의의 다른 유형의 수집된 데이터를 통해 수신된 데이터를 통해 유도될 수 있다.

[0081] [0093] 블록(1030)에서, 웨어러블 시스템은 클라우드로부터 디스플레이에 데이터를 전송할 수 있거나 데이터는 로컬 데이터베이스로부터 디스플레이 컴포넌트들로 전송될 수 있다. 블록(1040)에서, UI는 전송된 데이터에 기초하여 사용자에게 디스플레이된다. 예컨대, 광 필드 디스플레이는 가상 UI를 사용자의 눈들 중 하나 또는 둘 모두에 프로젝팅할 수 있다. 일단 가상 UI가 생성되면, 웨어러블 시스템은 블록(1050)에서, 가상 UI 상에 보다 많은 가상 콘텐츠를 생성하라는 사용자로부터의 커맨드를 단순히 기다릴 수 있다. 예컨대, UI는 사용자의 신체 또는 사용자의 환경 내의 사람(예컨대, 여행자)의 신체 주위의 신체 중심 링일 수 있다. 그 후, 웨어러블 시스템은 커맨드(제스처, 머리 또는 눈 움직임, 음성 커맨드, 사용자 입력 디바이스로부터의 입력 등)를 기다릴 수 있고, 커맨드가 인식되는 경우(블록 1060), 커맨드와 연관된 가상 콘텐츠가 사용자에게 디스플레이될 수 있다(블록 1070).

[0082] 웨어러블 디바이스의 예

[0083] [0094] 도 11은 사용자의 얼굴의 이미지들을 획득할 수 있는 예시적인 웨어러블 디바이스를 예시한다. 웨어러블 디바이스는 AR, VR 및/또는 MR 콘텐츠들을 디스플레이하도록 구성된 HMD(head-mounted device)일 수 있다. 웨어러블 디바이스에 의해 획득된 이미지들은 정지 이미지들, 애니메이션들, 비디오로부터의 개별 프레임들 또는 비디오를 포함할 수 있다.

[0084] [0095] 도 11의 웨어러블 디바이스(1150)는 도 2를 참조하여 설명된 웨어러블 시스템(200)의 부분일 수 있다. 웨어러블 디바이스(1150)는 사용자(210)의 얼굴을 이미징하도록 구성될 수 있는 이미징 시스템(1160)을 포함할 수 있다. 이미징 시스템(1160)은 도 4에 도시된 내향 이미징 시스템(462)의 예일 수 있다. 예컨대, 이미징 시스템(1160)은 사용자(210)가 웨어러블 디바이스(1150)를 착용하는 동안 사용자의 눈들(1110)의 눈 주위 구역을 이미징하도록 구성된 눈 카메라들(눈 카메라(1160a) 및 눈 카메라(1160b))와 같은 센서들을 포함할 수 있다. 이 예에서, 눈(1110b)은 도 3에 도시된 눈(302)에 대응할 수 있고 눈(1110a)은 눈(304)에 대응할 수 있다.

[0085] [0096] 각각의 눈 카메라는 FOV(field-of-view)를 가질 수 있다. 예컨대, 눈 카메라(1160a)에 대한 FOV는 구역(1120a) 및 구역(1130)을 포함할 수 있다. 눈 카메라(1160b)에 대한 FOV는 구역(1120b) 및 구역(1130)을 포함할 수 있다. 눈 카메라(1160a)의 FOV 및 눈 카메라(1160b)의 FOV는 구역(1130)에서 중첩될 수 있다.

[0086] [0097] 도 11에 도시된 바와 같이, 이미징 시스템(1160)은 사용자(210)의 머리 쪽을 향한다. 눈 카메라(1160a)는 눈(1110a)을 이미징하도록 구성될 수 있는 반면, 눈 카메라(1160b)는 눈(1110b)을 이미징하도록 구성될 수 있다. 이 도면에서, 눈 카메라(1160a)의 광학 축(1140a)은 눈 카메라(1160b)의 광학 축(1140b)과 평행하다.

[0087] [0098] 일부 구현들에서, 눈 카메라들 중 하나 또는 둘 모두는, 2개의 눈 카메라들의 광학 축들이 더 이상 평행하지 않도록 회전될 수 있다. 예컨대, 2개의 눈 카메라들은 (예컨대, 특히, 눈 카메라들이 디바이스(1150)의 프레임의 외측 에지들 근처에 배치되는 경우) 약간 서로를 향할 수 있다. 이 구현은, 이것이 2개의 카메라들 사이의 FOV의 중첩을 증가시킬 뿐만 아니라 2개의 눈 카메라들이 더 가까운 거리에서 얼굴을 이미징하도록 허용할 수 있는 사시(cross eyed) 구성을 생성할 수 있기 때문에 유리할 수 있다.

[0088] [0099] 도 11에 설명된 예가 2개의 눈 카메라들을 예시하지만, 웨어러블 디바이스(1150)는 2개의 눈 카메라들을 갖도록 요구되는 것은 아니다. 일부 실시예들에서, 이미징 시스템(1160)은 사용자의 얼굴을 이미징하는 하나의 눈 카메라를 포함할 수 있다. 한쪽 눈 카메라는 한쪽 눈과 연관된 안구주위 구역 또는 양쪽 눈의 안구주위 구역들을 이미징하도록 구성될 수 있다. 다른 실시예들에서, 웨어러블 디바이스(1150)는 2개 초과인 눈 카메라들을 포함할 수 있다.

[0089] 얼굴 모델의 예들

[0090] [0100] 얼굴 인식, 합성 및 렌더링의 맥락들에서, 인간 얼굴은 삼각형 메쉬들, 파라메트릭 표면(parametric surface)들, 선형 공간 표현들 또는 다른 수학적 표현들과 같은 3D 모델링 기술들을 사용하여 표현될 수 있다. 하나의 예로서, 3D 모델은 DLM(deformable linear model)을 사용하여 구축될 수 있다. DLM은 얼굴의 변형들의

상태를 벡터들의 리스트들로 인코딩할 수 있다. 변형 벡터들은 중립 얼굴과 연관되어 코를 크게 만들고 입을 좁게 만들거나 얼굴을 여성스럽게 만드는 등을 행할 수 있다. 변형 벡터들은 또한 미소 짓기, 찡그리기 또는 찌푸리기 등과 같은 얼굴 표정들과 연관될 수 있다. 얼굴 표정의 강도(이를테면, 큰 소리로 웃거나 부드럽게 미소 지음)를 고려하여, 변형 벡터들은 중립 얼굴 메쉬에 대해 강도 값에 의해 스케일링될 수 있다. 예컨대, 사람이 큰 소리로 웃는 경우, 강도 값은 더 클 수 있는데, 그 이유는 사람이 더 많은 얼굴 움직임들을 갖고 이에 따라 중립 얼굴에 대해 더 큰 변형을 가질 수 있기 때문이다. 반면에, 사람이 부드럽게 미소 짓는 경우, 강도 값은 더 작을 수 있는데, 그 이유는 많은 얼굴 움직임들이 존재하지 않기 때문이다. DLM 모델들이 사용될 수 있지만, 본 개시내용은 DLM 모델들로 제한되지 않으며, 다른 실시예들에서, ASM(active shape model)들, AAM(active appearance model)들, 모핑 가능 모델(morphable model; M)들 또는 임의의 다른 2-차원(2D) 또는 3D 형상 또는 텍스처 모델이 사용될 수 있다.

[0091] [0101] 웨어러블 디바이스(1150)는 이미징 시스템(1160)에 의해 획득된 사용자의 얼굴의 이미지들을 이용하여 사용자의 얼굴의 모델을 구축할 수 있다. 이미지들은 사용자가 디바이스를 착용하거나 벗을 때 이미징 시스템(1160)에 의해 획득될 수 있다. 이미지들은 또한 외향 이미징 시스템(464)(도 4에 도시됨)을 사용하여 사용자의 얼굴을 스캔함으로써 획득될 수 있다. 예컨대, 외향 이미징 시스템(464)을 사용하여 사용자의 얼굴을 스캔하기 위해, 사용자는 외향 이미징 시스템(464)이 (사용자의 환경 보다는) 사용자의 얼굴을 가리키도록 웨어러블 디바이스(1150)를 돌릴 수 있다. 웨어러블 디바이스는, 예컨대, 사용자가 웨어러블 디바이스를 처음 사용할 때 또는 사용자가 웨어러블 디바이스를 턴 온할 때와 같이, 웨어러블 디바이스의 초기화 단계 동안 사용자의 얼굴의 모델을 생성할 수 있다. 이미징 시스템(1160)에 의해 획득된 이미지들을 사용하여 얼굴 모델을 생성하는 예들은 또한, "FACE MODEL CAPTURE BY A WEARABLE DEVICE"란 제목의 미국 특허 출원 제15/717,223호에 설명되며, 이로써 그의 개시내용은 그 전체가 인용에 의해 본원에 포함된다.

[0092] [0102] 사용자의 얼굴의 모델은 기본 모델 및 사용자에 특정한 데이터에 기초하여 생성될 수 있다. 예컨대, 웨어러블 디바이스는 사람들의 그룹과 연관된 데이터로부터 미리 생성된 기본 모델을 사용하고 웨어러블 디바이스에 의해 획득된 이미지들을 분석함으로써 획득된 사용자 특정 정보에 기초하여 기본 모델을 커스터마이징할 수 있다. 일부 구현들에서, 기본 모델은 웨어러블 디바이스의 사용자와 유사한 인구학적 정보(예컨대, 연령, 성별, 민족성 등)를 갖는 사람들의 그룹과 연관될 수 있다. 예컨대, 사용자가 십대 여성인 경우, 웨어러블 디바이스는 통상적인 십대 여성과 연관된 기본 모델에 액세스할 수 있다. 다른 예로서, 사용자가 소정의 성별 또는 민족 그룹에 속하는 경우, 웨어러블 디바이스는 그 성별 또는 민족 그룹에 공통적인 기본 모델에 액세스할 수 있다. 웨어러블 디바이스는 또한 통계 분석을 사용하여 사용자의 얼굴 특징들의 외관을 결정할 수 있다. 예컨대, 웨어러블 디바이스는 사용자 또는 사람들의 그룹과 연관된 이미지들에 대한 통계 분석에 기초하여 얼굴 특징의 소정의 얼굴 외관이 적용될 가능성을 결정할 수 있다. 웨어러블 디바이스는 사용자에게 특정한, 획득된 이미지들에 기초하여 얼굴 특징들의 외관(이를테면, 안구주위 특징들의 외관)을 확인할 수 있다.

[0093] [0103] 웨어러블 디바이스는 또한 사용자 특정 이미지들에 기초하여 얼굴 모델을 구축할 수 있다. 예컨대, 웨어러블 디바이스는 내향 이미징 시스템에 의해 또는 외향 이미징 시스템에 의해 획득된 이미지들로부터만 사용자의 얼굴의 모델을 생성할 수 있다. 일부 구현들에서, 웨어러블 디바이스는 사용자 얼굴의 더 많은 이미지가 획득될 때 사용자의 얼굴 모델을 업데이트할 수 있다. 예컨대, 웨어러블 디바이스는 사용자가 디바이스를 착용할 때 내향 이미징 시스템에 의해 획득된 이미지들에 기초하여 얼굴 모델을 생성할 수 있다. 웨어러블 디바이스는 사용자가 디바이스를 벗을 때 또는 사용자가 디바이스를 재차 착용하는 다음 세션에 획득된 새로운 이미지들에 기초하여 얼굴 모델을 업데이트할 수 있다.

[0094] [0104] 이들 예들은 웨어러블 디바이스를 사용하여 얼굴 모델을 구축하거나 사용자의 얼굴의 맵을 생성하는 것을 지칭하지만, 일부 실시예들은 얼굴 모델을 생성하거나 그렇지 않으면 획득하기 위해 원격 컴퓨팅 디바이스와 통신하는 웨어러블 디바이스를 포함할 수 있다. 예컨대, 웨어러블 디바이스는 사용자의 얼굴의 이미지들을 획득하고, 이미지들(단독으로 또는 예컨대, 사용자의 인구학적 정보와 같은 사용자의 다른 정보와 조합하여)을 원격 컴퓨팅 디바이스(예컨대, 이를테면, 서버)에 전달할 수 있다. 원격 컴퓨팅 디바이스는 이미지들을 분석하고 얼굴 모델을 생성하고 얼굴 모델을 사용자의 웨어러블 디바이스로 전달하거나 또는 (예컨대, 텔레프레젠템 세션 동안) 얼굴 모델을 다른 사용자의 웨어러블 디바이스로 전달할 수 있다.

[0095] [0105] 얼굴 모델의 부분은 비습들, 턱 움직임들 등과 같은 사용자의 하부 얼굴의 특징들을 설명할 수 있다. 얼굴 모델에서 하부 얼굴의 컨퍼메이션들은 사용자가 상이한 사운드들(예컨대, 말)을 생성함에 따라 변할 수 있다. 예컨대, 사운드는 특정 언어의 사운드의 지각 가능하게 별개의 유닛인 음소와 연관될 수 있다. 음소들은 다른 단어로부터 하나의 단어를 구별하는 데 사용된다. 각각의 음소는 추가로, 음소가 발음(sound)될 때 발생

하는 얼굴 컨퍼메이션을 표현할 수 있는 비슴(viseme)에 맵핑될 수 있다. 다수의 언어들에서, 여러 음소들이 하나의 비슴에 대응할 수 있기 때문에(예컨대, 각각의 이러한 음소는 발음될 때 얼굴 상에서 동일하게 보임), 비슴들과 음소들 간의 일대일 대응성이 존재하진 않는다. 비슴은 얼굴의 소정의 컨퍼메이션과 연관된 얼굴 모델 파라미터 값에 의해 표현될 수 있다. 사용자가 상이한 단어를 발음할 때, 비슴이 변할 수 있다. 따라서, 얼굴의 컨퍼메이션이 변할 수 있고, 하부 얼굴을 설명하는 얼굴 모델의 파라미터들이 또한 상응하게 업데이트될 수 있다.

[0096] [0106] 본원에서 설명된 바와 같이, 일부 상황들에서, 얼굴 모델은 사용자가 HMD를 착용할 때 직접 측정에 의해 사용자의 하부 얼굴의 컨퍼메이션들의 업데이트들을 획득할 수 없을 수 있다(예컨대, 그 이유는 내향 눈-추적 카메라는 통상적으로 하부 얼굴을 이미징할 수 없을 수 있기 때문임). 그러나, HMD는 사용자의 안구주위 특징들을 관찰하기 위해 내향 이미징 시스템(예컨대, 이미징 시스템(462 또는 1160))을 사용할 수 있다. 도 13 내지 도 16을 참조하여 추가로 설명되는 바와 같이, HMD는 관찰된 안구주위 특징들을 하부 얼굴의 파라미터들에 맵핑하기 위해 머신 러닝 유도 모델을 사용할 수 있다. 이러한 하부 얼굴 파라미터들은, 안구주위 이미지들로부터 유도된 파라미터들과 함께, 사용자의 전체 얼굴 이미지를 합성하는 데 사용될 수 있다.

[0097] 안구주위 구역의 예들

[0098] [0107] 도 11을 참조하여 설명된 바와 같이, 이미징 시스템(1160)에 의해 획득된 이미지들은 사용자의 안구주위 구역의 부분을 포함할 수 있다. 안구주위 구역은 눈들 및 눈들 주위의 구역들을 포함할 수 있다. 도 12는 하나의 눈에 대한 안구주위 구역의 예시적인 이미지를 예시한다. 이 예에서, 안구주위 구역(1200a)은 눈(1210a)(이를테면, 안와) 및 눈(1210a) 주위의 구역을 포함할 수 있다. 눈(1210a) 주위의 구역은 예컨대, 눈썹(1220a), 코(1230a), 뺨(1240a) 및 이마(1250a)의 부분들을 포함할 수 있다. 안구주위 구역은 하나 이상의 안구주위 특징 또는 안구주위 특징들 중 부분들을 포함할 수 있다. 안구주위 특징들은 예컨대, 눈, 안와, 눈썹, 코, 볼 또는 이마를 포함할 수 있다. 얼굴의 다른 특징들 또는 사용자-특정 세부사항들이 또한 안구주위 특징들로 고려될 수 있다. 그러나 안구주위 구역은 입 또는 눈들로부터 떨어져 있는 코의 부분을 배제할 수 있다. 일부 구현들에서, 안구주위 구역은 키포인트들, 포인트 클라우드들, 벡터들 및 매트릭스들, 또는 다른 유형들의 수학적 표현들에 의해 표현될 수 있다.

[0099] [0108] 예시적인 이미지(1200a)는 사용자의 안구주위 구역(1270)을 이미징하는 HMD의 카메라로부터 획득될 수 있다. 이 예에서, 안구주위 구역(1270)은 특정한 특징들, 이를테면, 눈(1210a), 안와, 눈썹(1220a), 코(1230a), 볼(1240a) 및 이마(1250a)의 부분들을 포함한다. 각각의 안구주위 특징은 안구주위 특징과 연관된 다양한 특성들을 가질 수 있다. 특성들 중 일부는 각각의 안구주위 특징에 특정적일 수 있다. 예컨대, 안구주위 특징 눈썹(1220a)은 눈썹들의 형상, 눈썹의 컬러, 눈썹의 예상 움직임들 또는 움직임 방향들 등을 포함하는 특징들을 가질 수 있다. 안구주위 특징 눈(1210a)은 예컨대, 형상, 크기, 눈꼬리들의 위치, 시선 방향, 동공 위치, 안구 중심의 위치, 눈꺼풀의 형상 및 주름들, 안구주위의 피부 텍스처 등과 같은 특징들을 가질 수 있다. 다수의 다른 특징들이 또한 각각의 안구주위 특징을 식별 및 추적하는 데 사용될 수 있다. 하나 이상의 안구주위 특징의 하나 이상의 특성은 키포인트들, 포인트 클라우드들 또는 다른 유형의 수학적 표현들에 의해 표현될 수 있다.

[0100] [0109] 웨어러블 디바이스는 SIFT(Scale-Invariant Feature Transform), SURF(Speeded Up Robust Features), ORB(Oriented Fast and Rotated BRIEF), BRISK(binary robust invariant scalable keypoints) FREAK(fast retina keypoint) 등과 같은 시각적 키포인트 기술들 또는 뉴럴 네트워크를 사용하여 안구주위 특징들 및 연관된 특징들을 컴퓨팅 및 추적할 수 있다. 일부 실시예들에서, 특정 얼굴 특징은 그 특정한 안구주위 특징에 대해 특별히 설계된 검출기를 사용하여 추적될 수 있다. 예컨대, 눈꼬리들, 코 특징들, 입꼬리들 등과 같은 안구주위 특징 특징들은 다양한 알고리즘들을 사용하여 별개로 식별 및 추적될 수 있다. 이러한 안구주위 특징 특징들 중 하나 이상을 별개로 추적하는 것은 유익할 수 있는데, 그 이유는 사용자가 얼굴 표정을 짓거나 말하는 동안 각각의 안구주위 특징 및/또는 특징이 상당한 모션을 하기 쉽기 때문이다. 이러한 안구주위 특징들 및 특징들을 추적하는 데 사용되는 알고리즘들은 이동 범위를 고려할 수 있다. 예로서, 일부 안구주위 특징들 및/또는 연관된 특성들은 소정의 방향으로 움직일 가능성이 있고 그리고/또는 다른 방향들에서 보다 안정적으로 유지될 가능성이 있다(예컨대, 눈썹들은 왼쪽 또는 오른쪽이 아니라 위 또는 아래로 움직이는 경향이 있음).

[0101] [0110] 웨어러블 디바이스는 통계적으로 안구주위 특징들의 움직임을 분석할 수 있다. 이러한 통계는 얼굴 특징들이 소정의 방향으로 움직일 가능성을 결정하는 데 사용될 수 있다. 얼굴 특징들의 움직임들은 얼굴 모델에서 하나 이상의 파라미터들의 변형들을 계산하는 데 사용될 수 있다. 일부 실시예들에서, 하나 이상의 안구

주위 특징들 또는 특성들은 프로세싱 요건을 감소시키거나 신뢰성을 개선하기 위해 제거되거나 추적되지 않을 수 있다. 신뢰성을 개선하는 것이 바람직한 상황들에서, 다른 것들보다 예러가 발생하기 쉬운 안구주위 특징들 또는 특징들을 무시하거나 마스킹하는 것이 유리할 수 있다. 예컨대, 도 12b를 참조로 설명된 일부 실시예들에서, 웨어러블 디바이스는 안구주위 구역(1270)에서 다른 안구주위 특징들 또는 특징들을 추적할 때 눈 움직임이 HMD에 의해 인식되지 않도록 눈(1210b)의 중심 영역(1212)의 픽셀들을 무시할 수 있다.

[0102] [0111] 웨어러블 디바이스는 또한 순차적 베이지안 추정기(예컨대, 칼만 필터, 확장 칼만 필터 등), 번들 조정 등과 같은 vSLAM(visual simultaneous location and mapping) 기술들을 사용하여 안구주위 특징들 및 특징들을 식별 및 추적할 수 있다. 일부 실시예들에서, 웨어러블 디바이스는 사용자의 깊이 지각들 및 맵핑을 허용하도록 구성될 수 있다. 예컨대, 웨어러블 디바이스는 하나 이상의 카메라들에 의해 획득된 데이터로부터 얼굴 중 적어도 일부를 인코딩하는 밀집 맵을 구성할 수 있다. 키포인트 맵과 대조적으로, 밀집 맵은 3D 형상이 측정되는 얼굴의 패치들 또는 구역을 포함할 수 있다.

[0103] [0112] 웨어러블 디바이스 상의 카메라에 의해 캡처된 안구주위 구역의 크기 및 그 내부의 콘텐츠는 눈 카메라의 FOV에 의존할 수 있다. 일부 구현들에서, 눈 카메라는 캡처된 안구주위 구역 내의 모든 인식 가능한 안구주위 특징들을 피팅하는 큰 FOV를 갖지 않을 수 있다. 예컨대, 눈 카메라에 의해 캡처된 이미지들은 안와를 포함하지만 눈썹은 포함하지 않을 수 있다. 카메라의 기술 사양들은, 어떤 안구주위 특징들이 안구주위 구역의 다수의 캡처된 프레임들에 존재한 채로 남아있을 가능성이 가장 높은지 그리고 어떤 안구주위 특징들이 추적을 위해 가장 신뢰할 수 있는지를 결정할 수 있다.

[0104] [0113] 도 11을 참조하여 설명된 바와 같이, 일부 상황들에서, 각각의 눈 카메라가 눈을 이미징하도록 구성되지만, 2개의 눈 카메라들(하나는 외쪽 눈을 위한 것이고 하나는 오른쪽 눈을 위한 것)은 중첩 안구주위 구역들이 카메라들에 의해 이미징되도록 중첩 FOV(1130)를 가질 수 있다. 이는, 2개의 카메라들의 FOV가 충분히 넓고, 카메라들이 사용자의 얼굴의 중앙을 향하여 내향으로 기울어지고, 카메라들이 서로 근처에 포지셔닝되고, 그리고/또는 2개의 카메라들이 사용자로부터 충분히 멀리 떨어져 있기 때문일 수 있다. 그 결과, 사용자의 얼굴 부분, 통상적으로 중앙 부분(예컨대, 코)이 눈 카메라들 둘 모두에 의해 캡처될 수 있다. 웨어러블 디바이스는 2개의 카메라들로부터 획득된 이미지들을 결합하고, 결합된 이미지들이 안구주위 특징들을 포함하는지를 결정하고, 안구주위 특징들이 이미지들 내에 존재하는 것으로 결정되는 경우, 웨어러블 디바이스는 안구주위 특징들을 식별할 수 있다.

[0105] 안구주위 구역의 이미지로 하부 얼굴의 이미지를 생성하는 예들

[0106] [0114] 안구주위 구역의 이미지들뿐만 아니라 안구주위 구역의 움직임들은, HMD의 이미징 시스템에 의해 직접 관찰될 수 없는 사용자의 하부 얼굴의 이미지를 추론하는 데 사용될 수 있다. 도 13은 안구주위 구역의 컨퍼메이션 및 오디오 입력을 하부 얼굴의 컨퍼메이션과 상관시킬 수 있는 머신 러닝 유도 모델을 트레이닝하는 예를 예시한다. 일부 상황들에서, 컨퍼메이션은 얼굴 모델에서 벡터들에 의해 표현될 수 있다. 머신 러닝 유도 모델은, 안구주위 구역의 컨퍼메이션을 표현하는 하나 이상의 파라미터들과 하부 얼굴의 제2 컨퍼메이션을 표현하는 다른 파라미터들의 맵핑을 포함할 수 있다.

[0107] [0115] 도 13의 흐름도(1300)는 본원에서 설명된 웨어러블 시스템에 의해 수행될 수 있다. 예컨대, 머신 러닝 유도 모델의 트레이닝은 원격 데이터 리포지토리(280)(도 2에 도시됨)에 저장된 트레이닝 데이터를 사용하여 로컬 프로세싱 및 데이터 모듈(260), 원격 프로세싱 모듈(270)에 의해 수행될 수 있다. 예로서, HMD는 이미지들 및 오디오 입력들을 획득하고 모델을 트레이닝하기 위해 네트워크를 통해 이들을 다른 컴퓨팅 디바이스(예컨대, 서버)에 통신할 수 있다.

[0108] [0116] 흐름도(1300)에서, 안구주위 컨퍼메이션 및/또는 오디오 입력과 하부 얼굴(1330)의 컨퍼메이션의 맵핑(1330)이 생성될 수 있다. 소정의 구현들에서, 맵핑(1330)은 얼굴 모델의 안구주위 구역에서의 조정 또는 변형 내지 얼굴 모델에서의 하부 얼굴 구역에 대한 다른 조정 또는 변형을 포함할 수 있다. 조정 또는 변형은 얼굴 모델의 중립 포지션에 대해 결정될 수 있다. 맵핑(1330)은 하부 얼굴의 이미지(1312), 안구주위 이미지(1310), 눈 특정 정보(1316), 오디오 입력(1314)을, 단독으로 또는 조합하여 사용하여 생성될 수 있다. 안구주위 구역의 컨퍼메이션들은 안구주위 이미지들 및/또는 눈 특정 정보를 포함할 수 있다. 안구주위 구역의 컨퍼메이션은 또한, 안구주위 얼굴 파라미터들(예컨대, DLM의 얼굴 공간 파라미터들)에 의해 설명(또는 인코딩)될 수 있다. 하부 얼굴의 컨퍼메이션은 하부 얼굴의 이미지들 또는 특징들을 포함할 수 있다. 하부 얼굴의 컨퍼메이션들은 하부 얼굴 파라미터들에 의해 설명(또는 인코딩)될 수 있다.

- [0109] [0117] 안구주위 이미지(1310), 하부 얼굴의 이미지(1312) 및 오디오 입력(1314)은 HMD의 사용자 또는 사람들의 그룹에 특정적일 수 있다. 맵핑(1330)은 사람들의 그룹으로부터의 정보에 기초한 일반 맵핑 또는 사용자 특정 입력들을 사용한 개별 특정 맵핑일 수 있다. 맵핑(1330)은 또한 사람들의 그룹에 일반적인 정보에 기초할 수 있지만, 개별 특정 정보에 기초하여 추가로 커스터마이징된다.
- [0110] [0118] 안구주위 이미지들(1310) 및 눈 특정 정보(1316)는 내향 이미징 시스템에 의해 획득될 수 있다. 안구주위 이미지(1310)는 안구주위 특징들을 포함할 수 있다. 안구주위 이미지(1310)는 또한 눈 포즈, 동공 확장 상태, 눈 컬러, 눈꺼풀 상태 등과 같은 눈 특정 정보(1316)를 포함할 수 있다. 일부 구현들에서, 얼굴 모델의 얼굴 파라미터들은 눈 특정 정보(1316)를 고려하지 않을 수 있다. 결과적으로, 웨어러블 시스템은 머신 유도 러닝 모델을 트레이닝하기 위해 안구주위 얼굴 파라미터들과 함께, 별개의 입력으로서 눈의 특정 정보(1316)를 입력할 수 있다.
- [0111] [0119] 하부 얼굴의 이미지(1312)는 웨어러블 디바이스 외부의 카메라에 의해 획득될 수 있다. 하부 얼굴의 이미지(1312)는 또한 외향 이미징 시스템에 의해 획득된 바와 같은 사용자의 얼굴의 반사된 이미지들을 사용하여 획득될 수 있다. 일부 실시예들에서, HMD는 (도 4에 도시된 바와 같이) 외향 이미징 시스템 및 내향 이미징 시스템 둘 모두를 포함할 수 있다. 사용자가 반사 표면 근처(이들테면, 미러 앞)에 있을 때, HMD는, 외향 이미징 시스템을 사용하여 사용자 얼굴의 반사된 이미지 및 내향 이미징 시스템을 사용하여 안구주위 구역의 이미지를 동시에 획득할 수 있다. 웨어러블 디바이스는 사용자가 말할 때, 마이크로폰과 같은 오디오 입력 센서를 사용하여 사운드(예컨대, 오디오 입력(1314))를 캡처할 수 있다.
- [0112] [0120] 오디오 입력(1314), 눈 이미지(1310) 및 하부 얼굴의 이미지(1312)는 시간에 걸쳐 그리고/또는 상이한 사용자들로부터 획득될 수 있다. 예컨대, 웨어러블 디바이스는 사용자의 다수의 텔레프레전스 세션에 걸쳐 이미지들 및 오디오들을 레코딩할 수 있다. 웨어러블 디바이스는 또한 일반적인 맵핑(1330)을 생성하기 위해 다수의 사용자들로부터 이들을 획득할 수 있다.
- [0113] [0121] 맵핑(1330)은 다양한 입력들을 사용하여 생성될 수 있다. 일 예로서, 맵핑(1330)은 안구주위 이미지(1310) 및 하부 얼굴의 이미지(1312)를 사용하여 트레이닝될 수 있다. 이 예에서, 맵핑(1330)은 안구주위 컨퍼메이션을 하부 얼굴의 컨퍼메이션과 상관시킬 수 있다. 맵핑(1330)은 또한 오디오 입력(1314), 안구주위 이미지(1310) 및 하부 얼굴의 이미지를 사용하여 생성될 수 있다. 결과적으로, 맵핑(1330)은 사운드 및 대응하는 안구주위 컨퍼메이션을 하부 얼굴의 컨퍼메이션과 상관시킬 수 있다. 일부 상황들에서, 컨퍼메이션은 얼굴 모델의 파라미터와 연관될 수 있다. 예컨대, 파라미터는 중립 얼굴 모델로부터의 편차들에 대한 수학적 표현들을 포함할 수 있다. 맵핑(1330)은 하부 얼굴의 파라미터들에 맵핑된 안구주위 구역의 파라미터들을 포함할 수 있다.
- [0114] [0122] 일부 실시예들에서, 하부 얼굴의 이미지들(1312)은 맵핑(1330)을 트레이닝하는데 이용 가능하지 않을 수 있다. 결과적으로, 트레이닝 프로세스는 오디오 입력(1314)으로부터 음소(1320)를 식별하는 것 그리고 오디오 입력(1314)에서 비음(1318)을 결정하기 위해 음소를 사용하는 것을 포함할 수 있다.
- [0115] [0123] 맵핑(1330)은 안구주위 이미지(1310)와 비음(1318) 사이의 상관성(correlation)을 포함할 수 있다. 부가적으로 또는 대안적으로, 맵핑(1330)은 오디오 입력(1314)의 사운드와 비음(1318)의 맵핑과 연관될 수 있다. 도 14 및 도 16을 참조하여 추가로 설명된 바와 같이, 비음(1318)은 얼굴 이미지를 생성하기 위해 안구주위 이미지(1310)와 결합될 수 있다. 비음은 사람들의 그룹으로부터의 데이터를 사용하여 결정될 수 있다. 비음은 또한 사용자 특정 데이터에 기초하여 결정될 수 있다. 예컨대, 웨어러블 디바이스는 사용자가 말할 때 대응하는 비음들의 이미지들을 획득할 수 있다. 음소들은, 다수의 사운드들이 동일한 비음을 사용하여 발음될 수 있기 때문에, 시각적 비음들에 대한 일대일 맵핑을 갖지 않을 수 있다. 결과적으로, 웨어러블 디바이스는 오디오 입력(1314)이 상이하더라도 동일한 비음을 식별할 수 있다.
- [0116] [0124] 웨어러블 디바이스는 예컨대, 획득 시간에 기초하여, 하부 얼굴의 이미지, 안구주위 구역의 이미지 및 선택적으로 오디오 입력을 연관시킬 수 있다. 소정의 실시예들에서, 웨어러블 디바이스는 또한 사운드에 기초하여 하부 얼굴의 이미지 및 안구주위 구역의 이미지를 연관시킬 수 있다. 일 예로서, 사용자가 말할 때 사용자는 항상 반사 표면 근처에 있진 않을 수 있다. 결과적으로, 안구주위 구역의 이미지들과 상이한 시간에 하부 얼굴의 이미지들이 획득될 수 있다. 웨어러블 디바이스는 사용자가 미러 앞에 있을 때 사용자가 말한 단어뿐만 아니라 단어와 연관된 하부 얼굴 이미지를 식별할 수 있고 사용자가 동일한 단어를 말할 때 획득된 안구주위 이미지들을 식별할 수 있다. 따라서, 웨어러블 디바이스는 단어의 사운드에 기초하여 안구주위 구역의 이미지들

에 하부 얼굴 이미지들을 상응하게 연결할 수 있다.

- [0117] [0125] 맵핑(1330)의 출력은 얼굴의 적어도 부분의 이미지 또는 얼굴 파라미터들을 포함할 수 있다. 예컨대, 하부 얼굴 구역이 웨어러블 디바이스에 의해 관찰 불가능한 경우, 웨어러블 디바이스는 (웨어러블 디바이스에 의해 관찰 가능한) 안구주위 이미지들에 그리고 하부 얼굴의 이미지, 하부 얼굴 파라미터들, 전체 얼굴 이미지 또는 전체 얼굴의 얼굴 파라미터들에 맵핑(1330)을 적용할 수 있다.
- [0118] [0126] 맵핑(1330)은 대응하는 얼굴 파라미터들 또는 하부 얼굴의 이미지들에 대한 확률론적 추정 (probabilistic estimate)을 포함할 수 있다. 예컨대, 사용자의 안구주위 구역은 사용자가 상이한 사운드를 내더라도 유사한 외관을 가질 수 있다. 결과적으로, 맵핑(1330)은 하부 얼굴의 관찰되지 않은 얼굴 파라미터들이 안구주위 구역의 관찰된 얼굴 파라미터들과 연관되는지 여부에 대한 가능성을 제공할 수 있다. 웨어러블 디바이스는 가장 가능성 높은 이미지 또는 얼굴 파라미터들을 고를(pick) 수 있다. 이미지의 가능성은 또한 오디오 입력, 생리적 데이터(이를테면, 동공 확장 상태, 심박수 등) 등과 같은 다른 팩터들과 조합하여 또는 단독으로 기초할 수 있다.
- [0119] [0127] 다양한 머신 러닝 알고리즘들이 이 프로세스에 대해 사용될 수 있다. 모델들을 생성하고 업데이트하는 데 사용될 수 있는 머신 러닝 알고리즘들의 일부 예들은, (예컨대, 정규 최소 제곱화 회귀 분석(Ordinary Least Squares Regression)과 같은) 회귀 알고리즘들, (예컨대, 러닝 벡터 양자화(Learning Vector Quantization)와 같은) 인스턴스-기반 알고리즘들, (예컨대, 분류 및 회귀 트리들과 같은) 결정 트리 알고리즘들, (예컨대, 나이브 베이즈(Naive Bayes)와 같은) 베이지안 알고리즘들, (예컨대, k-평균 클러스터링과 같은) 클러스터링 알고리즘들, (예컨대, 선형적 알고리즘들과 같은) 연관 규칙 러닝 알고리즘들, (예컨대, 퍼셉트론(Perceptron)과 같은) 인공 뉴럴 네트워크 알고리즘들, (예컨대, 딥 볼츠만 머신, 또는 딥 뉴럴 네트워크와 같은) 딥 러닝 알고리즘들, (예컨대, 주 성분 분석(Principal Component Analysis)과 같은) 차원성 감소 알고리즘, (예컨대, 누적된 일반화(Stacked Generalization)와 같은) 앙상블 알고리즘들 및/또는 다른 머신 러닝 알고리즘들을 포함하는 감독 또는 비-감독 머신 러닝 알고리즘들을 포함할 수 있다. 일부 실시예들에서, 개별 모델들은 개별 데이터 세트들에 대해 커스터마이징될 수 있다. 예컨대, 웨어러블 시스템은 기본 모델을 생성할 수 있다. 기본 모델은 데이터 유형(예컨대, 특정 사용자), 데이터 세트(예컨대, 텔레프레젠텐션 세션의 사용자의 획득된 부가적인 이미지들의 세트), 조건부 상황들(예컨대, 게임 플레이 동안의 피팅은 인터넷 브라우징 동안의 피팅과 상이할 수 있음), 또는 다른 변동들에 특정한 부가적인 모델들을 생성하기 위한 시작 포인트로서 사용될 수 있다. 일부 실시예들에서, 웨어러블 디바이스는 어그리게이팅된 데이터의 분석을 위한 모델들을 생성하기 위해 복수의 기술들을 활용하도록 구성될 수 있다. 다른 기술들은 미리 정의된 임계치들 또는 데이터 값들을 사용하는 것을 포함할 수 있다. 시간이 지남에 따라, 웨어러블 디바이스는 머신 러닝 유도 모델(1330)을 계속 업데이트할 수 있다.
- [0120] [0128] 도 13을 참조하면, 기본 모델은 사람들의 그룹으로부터의 데이터를 사용하여 생성된 바와 같이 하부 얼굴에 대한 안구주위 이미지의 맵핑을 포함할 수 있다. 맵핑(1330)은 HMD의 다수의 사용자들 각각에 대해 커스터마이징될 수 있다(예컨대, 가족은 단일 HMD의 사용을 공유할 수 있고 각각의 가족 구성원에 대한 맵핑을 커스터마이징할 수 있음).
- [0121] 얼굴 이미지의 합성의 예들
- [0122] [0129] 도 14는, 얼굴의 부분이 웨어러블 디바이스에 의해 관찰되지 않을 때 얼굴의 이미지를 생성하는 예를 예시한다. 도 14의 흐름도(1400)는 본원에서 설명된 웨어러블 디바이스에 의해 수행될 수 있다. 이 예에서, 웨어러블 디바이스는 사용자의 안구주위 이미지들(1410)을 획득할 수 있는 내향 이미징 시스템을 포함할 수 있다. 웨어러블 디바이스는 또한 사용자로부터 오디오 입력(1414)을 획득할 수 있는 오디오 센서들을 포함할 수 있다. 그러나, 내향 이미징 시스템은 사용자의 하부 얼굴의 이미지를 획득할 수 없을 수도 있다.
- [0123] [0130] 사용자의 얼굴의 이미지를 합성하기 위해, 웨어러블 디바이스는 머신 러닝 유도 모델(1420)을 오디오 입력(1414) 및/또는 안구주위 이미지들(1410)에 적용할 수 있다. 머신 러닝 유도 모델(1420)은 도 13에서 설명된 맵핑(1330)을 생성할 수 있다. 웨어러블 디바이스는 사용자가 말할 때 오디오 센서로부터 오디오 스트림을 수신할 수 있다. 웨어러블 디바이스는 음소와 비음의 맵핑을 제공할 수 있는 머신 러닝 유도 모델(1420)에 액세스할 수 있다. 웨어러블 디바이스는 오디오 스트림을 머신 러닝 유도 모델(1420)에 공급하여 오디오 스트림 내의 사운드들의 비음들(1430)을 획득할 수 있다. 비음들은 얼굴의 이미지(1450)를 생성하기 위해 안구주위 이미지들(1410)과 결합될 수 있다.

- [0124] [0131] 부가적으로 또는 대안적으로, 웨어러블 디바이스는 비습들(1430)을 획득하기 위해 안구주위 이미지들(1410) 및 오디오 입력(1414)의 결합에 머신 러닝 유도 모델(1430)을 적용할 수 있다. 이 예에서, 머신 러닝 유도 모델(1420)은 사운드 및 안구주위 이미지와 비습의 맵핑을 생성할 수 있다.
- [0125] [0132] 다른 예로서, 머신 러닝 유도 모델(1420)은 안구주위 이미지들과 하부 얼굴 이미지들(또는 비습들)의 맵핑을 포함할 수 있다. 웨어러블 디바이스는 내향 이미징 시스템에 의해 획득된 안구주위 이미지들에 머신 러닝 유도 모델(1420)을 적용함으로써 비습을 획득할 수 있다.
- [0126] [0133] 이미지들의 맵핑 외에도 또는 그에 대한 대안으로서, 머신 러닝 유도 모델(1420)은 또한 다른 유형들의 컨퍼메이션들에 대한 맵핑들, 이를테면, 안구주위 파라미터들과 하부 얼굴 파라미터의 맵핑, 안구주위 특징들과 하부 얼굴 파라미터들/이미지들의 맵핑, 또는 안구주위 특징들/이미지들/얼굴 파라미터들과 전체 얼굴 이미지 또는 파라미터들의 맵핑을 제공할 수 있다. 일 예로서, 웨어러블 디바이스는 안구주위 이미지들(1410)에 기초하여 안구주위 얼굴 파라미터들을 계산할 수 있다. 사용자가 말할 때, 웨어러블 디바이스는 사운드 및 안구주위 이미지들(1410)로부터 결정된 바와 같은 안구주위 얼굴 파라미터들을 머신 러닝 유도 모델(1420)에 입력할 수 있다. 머신 러닝 유도 모델(1420)에 의해 생성된 맵핑은 추정된 현재 비습(1430)을 출력할 수 있다. 현재 비습(1430)은 하부 얼굴 파라미터들에 대한 추정을 생성하는 데 사용될 수 있다. 일부 상황들에서, 안구주위 얼굴 파라미터들은 눈 특정 정보(1316)를 고려하지 않을 수 있다. 결과적으로, 웨어러블 디바이스는 비습들(1430)을 결정하기 위한 맵핑의 팩터(이를테면, 안구주위 얼굴 파라미터들과 별개의 팩터)로서 눈 특정 정보(1316)를 머신 러닝 유도 모델(1420) 내로 입력할 수 있다.
- [0127] [0134] 일부 구현들에서, 머신 러닝 유도 모델(1420)은 비습들 외에도 또는 그에 대안적으로, 얼굴의 이미지를 출력할 수 있다. 예컨대, 머신 러닝 유도 모델(1420)은 오디오 입력(1414) 및 안구주위 이미지를 전체 얼굴의 이미지와 연관시킬 수 있다. 따라서, 웨어러블 시스템은 머신 러닝 유도 모델(1420)로부터 출력을 수신한 후에 얼굴의 이미지(1450)를 합성할 필요가 없다. 소정의 구현들에서, 머신 러닝 유도 모델(1420)은 3D 얼굴 모델에서 얼굴 파라미터들을 출력할 수 있으며, 여기서 얼굴 파라미터들은 비습들(1430) 또는 얼굴의 이미지(1450)와 연관될 수 있다.
- [0128] [0135] 웨어러블 시스템은 얼굴의 이미지(1450)를 생성하기 위해 하부 얼굴 및 안구주위 구역의 컨퍼메이션들을 결합할 수 있다. 예컨대, 웨어러블 시스템은 전체 얼굴 이미지를 생성하기 위해 안구주위 이미지를 하부 얼굴의 이미지와 결합할 수 있다. 일부 구현들에서, 웨어러블 시스템은, 사용자가 말할 때 얼굴의 이미지들을 생성하기 위해 관찰된 안구주위 이미지들(1410) 및 추정된 비습들(1430)로 이전에 생성된 얼굴 모델(1440)을 동적으로 업데이트할 수 있다. 얼굴 모델(1440)은 사람들의 그룹으로부터의 데이터에 기초하여 생성된 일반적인 모델일 수 있다. 얼굴 모델(1440)은 또한 사용자 특정적일 수 있다. 예컨대, 얼굴 모델은 사용자가 디바이스를 착용하는 동안 내향 이미징 시스템에 의해 획득된 이미지들로 생성될 수 있다.
- [0129] [0136] 사용자는 또한, 예컨대, 다른 얼굴 특징들 및 텍스처 맵들을 선택함으로써 얼굴 모델을 커스터마이징할 수 있다. 예로서, 사용자는 과학 소설 외계인과 같은 판타지 생물의 외관을 선택할 수 있다.
- [0130] [0137] 웨어러블 디바이스는 사용자의 얼굴 움직임들을 반영하기 위해 사용자가 말할 때 얼굴 모델의 하나 이상의 얼굴 파라미터들을 동적으로 업데이트할 수 있다. 예컨대, 웨어러블 디바이스는 안구주위 구역의 새로운 이미지들에 기초하여 안구주위 얼굴 파라미터들(1410)에 대한 변화들을 계산함으로써 얼굴 모델의 안구주위 구역을 업데이트할 수 있다. 웨어러블 디바이스는 또한 안구주위 이미지들(1410)에 기초하여 추정된 비습들(1430)을 사용하여 사용자의 하부 얼굴을 업데이트할 수 있다.
- [0131] [0138] 부가적으로 또는 대안적으로, 웨어러블 디바이스는 안구주위 구역으로부터의 이미지들(1410), 비습들(1430) 및 사용자 얼굴의 다른 부분들을 결합함으로써 사용자의 얼굴 표정들을 애니메이션할 수 있다. 예컨대, 사용자가 사운드를 낼 때, 안구주위 구역의 이미지(1410) 및 비습(1430)의 이미지는 사운드에 기초하여 변할 수 있다. 그러나 일부 구현들에서, 안구주위 구역의 이미지(1410) 또는 비습(1430)의 이미지 중 하나만이 변할 수 있다. 예컨대, 사용자는 상이한 사운드를 내는 동안 동일한 안구주위 표정을 가질 수 있다.
- [0132] [0139] 일부 상황들에서, 맵핑(1330)은, 중간 스테이지에서, 하부 얼굴의 가능한 컨퍼메이션들의 세트 각각이 관찰된 안구주위 구역의 컨퍼메이션에 매칭할 가능성을 생성할 수 있다. 웨어러블 시스템은 그러한 컨퍼메이션이 임계 기준을 통과하는 경우, 가장 높은 가능성과 연관된 하부 얼굴의 컨퍼메이션을 선택할 수 있다. 웨어러블 디바이스는 또한 수신된 입력에 기초하여 하부 얼굴의 추정된 컨퍼메이션의 정확도를 계산할 수 있다.
- [0133] [0140] 얼굴의 이미지(1450)는 텔레프레젠테이션 세션에서 사용될 수 있다. 예컨대, 두 명의 사용자들이 텔레프레

전스 세션에 있을 때, 사용자의 웨어러블 디바이스는 사용자의 얼굴 파라미터들 또는 얼굴의 이미지에 대한 업데이트들을 다른 사용자의 웨어러블 디바이스에 통신할 수 있다. 따라서, 다른 사용자의 웨어러블 디바이스는, 사용자가 말할 때 보다 생생한 애니메이션을 제공하기 위해 업데이트된 얼굴 파라미터들에 기초하여 사용자의 가상 외관(예컨대, 사용자의 얼굴)의 애니메이션을 업데이트 또는 제공할 수 있다.

[0134] [0141] 일부 구현들에서, 얼굴 파라미터들은 눈 특정 정보(1316)를 포함하지 않을 수 있다. 그러나, 맵핑은 그림에도 불구하고, 내향 이미징 시스템에 의해 관찰된 바와 같은 눈 특정 정보(1316)를 얼굴의 이미지에 통합할 수 있다. 예컨대, 눈 특정 정보는 맵핑이 트레이닝될 때 별개의 입력으로서 사용될 수 있다. 웨어러블 디바이스는 또한, 사용자의 가상 외관에서 사용자의 눈 움직임을 포함 및 업데이트할 수 있는 다른 사용자의 웨어러블 디바이스에 눈 특정 정보(1316)에 대한 파라미터들을 전달할 수 있다.

[0135] [0142] 예들은 하부 얼굴의 컨퍼메이션들을 추론하기 위해 안구주위 이미지들을 사용하는 것을 참조하여 설명되지만, 유사한 기술이 또한, 사용자의 뺨 또는 사용자의 신체 움직임들(예컨대, 사용자가 분노할 때 팔을 흔들음) 등과 같은 사용자의 다른 관찰되지 않은 구역들의 움직임 또는 이미지를 추론하는 데 사용될 수 있다. 예컨대, 소정의 사용자들의 경우, 머리 움직임들과 얼굴 표정들(이를테면, 하부 얼굴 움직임들 또는 안구주위 움직임들) 사이에 강한 상관성이 존재할 수 있다. 웨어러블 디바이스는 도 2에 설명된 IMU를 사용하여 머리 움직임 데이터를 획득하고, 머리 움직임 데이터를 사용하여, 얼굴 파라미터들의 적어도 부분 또는 얼굴 이미지의 부분을 추론할 수 있다. 일부 상황들에서, 안구주위 구역이 이미징되지 않을 수 있는 반면, 사용자의 다른 구역들이 이미징될 수 있다. 예컨대, 웨어러블 디바이스는, 예컨대 외향 이미징 시스템을 사용하여 사용자의 얼굴의 반사된 이미지들을 획득함으로써 사용자의 하부 얼굴의 이미지들을 획득할 수 있다. 다른 예로서, 사용자의 신체 움직임들 및 사용자의 머리의 움직임을 관찰하기 위해 외부 카메라가 사용될 수 있다. 그러나, 안구주위 구역이 웨어러블 디바이스에 의해 폐색되기 때문에, 외향 이미징 시스템 또는 외부 카메라가 사용될 때 안구주위 구역이 이미징되지 않을 수 있다. 웨어러블 디바이스는 유사한 기술들을 사용하여 하부 얼굴의 관찰된 이미지들을 사용하여 안구주위 구역의 컨퍼메이션들을 추론할 수 있다. 웨어러블 디바이스는 하부 얼굴의 컨퍼메이션들과 안구주위 컨퍼메이션의 맵핑을 생성하고, 관찰된 구역들의 이미지들 및/또는 오디오 입력을 사용하여 안구주위 컨퍼메이션들을 결정하기 위해 맵핑을 적용할 수 있다. 예컨대, 맵핑은 비습(1430)으로부터 안구주위 이미지(1410)로 이루어질 수 있다. 웨어러블 디바이스는 비습(1430)을 결정하기 위해 추가로 사용될 수 있는 음소를 결정하는 데 오디오 입력을 사용할 수 있다. 웨어러블 디바이스는 비습(1430) 및/또는 오디오 입력을 사용하여 대응하는 안구주위 이미지/파라미터를 식별할 수 있다.

[0136] [0143] 일부 구현들에서, 획득된 안구주위 이미지는 단지 한쪽 눈에 대한 정보를 포함할 수 있다. 웨어러블 디바이스는 획득된 안구주위 이미지에 기초하여 다른 눈에 대한 안구주위 얼굴 파라미터를 결정하거나 이미지를 생성할 수 있다. 예컨대, 두 눈들의 움직임들은 동일한 방향일 수 있고, 눈 컬러들은 동일할 수 있고, 두 눈들에 대한 형상들이 유사할 수 있고, 두 눈들은 대칭적일 수 있고, 기타 등등이 가능하다. 웨어러블 디바이스는 한쪽 눈(예컨대, 사용자가 단어를 말할 때 눈은 소정의 방향을 보고 있고, 사용자가 흥분할 때 사용자의 동공은 팽창하는 등을 함)의 정보를 사용하여, 다른 눈에 대한 정보를 결정하고 얼굴의 이미지에 양쪽 눈들에 대한 안구주위 이미지들을 통합할 수 있다.

[0137] [0144] 부가적으로 또는 대안적으로, 본원에서 설명된 기술들은 또한 피부 텍스처들을 모델링하는 데 사용될 수 있다. 예컨대, 관찰되지 않은 하부 얼굴의 피부 컬러는 관찰된 안구주위 구역의 피부 컬러와 매칭할 필요가 있을 수 있다. 일부 구현들에서, 사람의 인구학적 특성들(예컨대, 연령, 성별, 인종 등) 뿐만 아니라 피부 상태들이 안구주위 이미지들로부터 결정될 수 있다. 머신 러닝 유도 모델(1420)은 안구주위 특징들의 관찰된 특성들과 하부 얼굴 특징들의 특성들의 맵핑을 포함할 수 있다. 예컨대, 안구주위 이미지가 많은 주름들을 포함하는 경우, 맵핑은 하부 얼굴 이미지가 또한 주름들을 가져야 함을 추론하는 데 사용될 수 있다.

[0138] 안구주위 이미지들을 이용하여 전체 얼굴 이미지를 합성하는 예시적인 프로세스

[0139] [0145] 도 15a는 안구주위 구역의 이미지와 하부 얼굴의 이미지 간의 맵핑을 생성하는 예시적인 프로세스를 설명한다. 프로세스(1500)는 본원에서 설명된 웨어러블 디바이스에 의해 수행될 수 있다.

[0140] [0146] 블록(1510a)에서, 웨어러블 디바이스는 안구주위 이미지 및 사운드에 대한 오디오 입력에 액세스할 수 있다. 안구주위 이미지들 및 오디오 입력은 사람들의 그룹으로부터 획득되거나 사용자에게 특정적일 수 있다.

[0141] [0147] 블록(1520a)에서, 웨어러블 디바이스는 안구주위 이미지들 또는 오디오 입력에 기초하여 비습을 식별할 수 있다. 예컨대, 웨어러블 디바이스는 오디오 입력에서 음소들을 식별할 수 있으며, 여기서 음소들은 추가

로 비습들과 연관될 수 있다. 웨어러블 디바이스는 음소들과 비습들의 맵핑을 생성하기 위해 머신 러닝 기술들을 사용할 수 있다.

[0142] [0148] 웨어러블 디바이스는 또한 머신 러닝 기술들을 사용하여 비습들을 안구주위 이미지들과 연관시킬 수 있다. 예컨대, 사람이 분노할 때, 사람은 소정의 단어들을 말하는 경향이 있을 수 있고(이에 따라 소정의 비습들을 가짐), 안구주위 구역에서 소정의 표정들을 가질 수 있다. 머신 러닝 유도 모델은 블록(1530a)에 도시된 바와 같이 안구주위 이미지들과 비습들의 맵핑을 생성하기 위해 안구주위 이미지들을 비습들과 상관시킬 수 있다.

[0143] [0149] 음소들에 기초하여 비습들을 결정하는 것 외에도, 또는 이에 대안적으로, 웨어러블 디바이스는 예컨대, 사람의 얼굴을 이미징하도록 외부 카메라를 사용함으로써, 또는 외향 이미지 시스템을 사용하여 반사된 이미지들을 획득함으로써, 안구주위 이미지들에 대응하는 하부 얼굴의 이미지들(비습들을 포함함)을 획득할 수 있다. 블록(1530a)에서, 웨어러블 디바이스는 안구주위 이미지들, 하부 얼굴의 이미지들, 및 선택적으로 오디오 입력을 입력하여, 안구주위 이미지들, 오디오 입력, 또는 이들의 결합과 하부 얼굴의 이미지들의 맵핑을 생성할 수 있다. 맵핑은 또한 안구주위 얼굴 파라미터들을 사용하여 생성될 수 있다. 예컨대, 맵핑은 안구주위 얼굴 파라미터들을 하부 얼굴 파라미터들과 상관시킬 수 있다. 안구주위 얼굴 파라미터들은 안구주위 이미지들에 기초하여 계산될 수 있는 반면, 하부 얼굴 파라미터들은 비습들에 기초하여 계산될 수 있다.

[0144] [0150] 블록(1540a)에서, 웨어러블 디바이스는 선택적으로, 맵핑이 사람들의 그룹으로부터 획득된 데이터에 기초하여 생성되는 경우, 맵핑을 특정 사용자에게 맞춤제작(tailor)할 수 있다. 예컨대, 웨어러블 디바이스는 얼굴의 사용자 특정 컨퍼메이션들을 맵핑에 통합하기 위해 부가적인 오디오 정보, 사용자의 안구주위 이미지들 또는 사용자의 하부 얼굴의 이미지들에 액세스할 수 있다. 일부 실시예들에서, 맵핑은 더 많은 사용자 특정 정보가 획득됨에 따라 시간에 걸쳐 트레이닝될 수 있다. 예컨대, 웨어러블 디바이스는 시간에 걸쳐 사용자의 스피치, 안구주위 이미지들 또는 하부 얼굴의 이미지들을 레코딩할 수 있다. 맵핑은 획득된 새로운 정보에 기초하여 업데이트될 수 있다. 예컨대, 웨어러블 디바이스는 매일, 매주, 매월 등으로, 또는 웨어러블 디바이스와 사용자 상호작용에 기초하여 맵핑을 업데이트할 수 있다.

[0145] [0151] 도 15b는 입력으로 안구주위 구역의 얼굴 파라미터들 및/또는 음소들 및 출력으로서 하부 얼굴 파라미터들을 갖는 맵핑을 트레이닝하는 다른 예시적인 프로세스를 예시한다. 프로세스(1500b)는 본원에서 설명된 웨어러블 디바이스에 의해 또는 원격 컴퓨팅 시스템에 의해 수행될 수 있다. 원격 컴퓨팅 시스템은 큰 데이터 세트에 대한 데이터 분석을 수행하도록 구성된 컴퓨터 서버들을 포함할 수 있다.

[0146] [0152] 블록(1510b)에서, 컴퓨터 시스템은 연관된 안구주위 이미지들 및 하부 얼굴 이미지들의 쌍들을 포함하는 트레이닝 데이터에 액세스할 수 있다. 부가적으로 또는 대안적으로, 트레이닝 데이터는 또한, 연관된 오디오 세그먼트들 및 하부 얼굴 이미지들의 쌍들을 포함할 수 있다. 이러한 이미지들 및 오디오 세그먼트들은 동일한 사용자 또는 다수의 사용자들과 연관될 수 있다. 트레이닝 데이터는 도 2에서 설명된 원격 데이터 리포지토리(280)에 저장될 수 있다. 안구주위 이미지들은 내향 이미징 시스템(462)을 사용하여 획득될 수 있다. 하부 얼굴 이미지들은 (예컨대, 사용자가 미러 앞에 있을 때) 외향 이미징 시스템(464)을 사용하여 획득될 수 있다. 오디오 세그먼트들은 마이크로폰과 같은 웨어러블 디바이스의 오디오 센서를 사용하여 획득될 수 있다.

[0147] [0153] 일 경우에, (각각의 맵핑에 입력되는, 즉 입력 안구주위 이미지들 및 오디오 세그먼트들로부터 추출되는) 특징 벡터는 음소에 대응하는 입력 파라미터들의 하나의 서브세트 및 안구주위 구역 유도된 얼굴 파라미터들에 대응하는 입력 파라미터들의 하나의 서브세트를 갖는다. 하부 얼굴 파라미터의 세트 각각에 대해 하나씩 다수의 맵핑들이 존재할 수 있다. 파라미터들의 상이한 리스트를 각각 갖는 상이한 특징 벡터들은 각각의 하부 얼굴 파라미터에 대한 별개의 맵핑들에서 사용될 수 있다.

[0148] [0154] 컴퓨터 시스템은 블록(1520b 및 1530b)에서 각각, 안구주위 이미지들에서 그리고 하부 얼굴 이미지들에서 얼굴 모델의 얼굴 파라미터들(이를테면, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들)을 식별하기 위해 이미지 프로세싱 기술을 이용할 수 있다. 이미지 프로세싱 기술들은 뉴럴 네트워크들, 시각적 키포인트 기술들, 또는 안구주위 특징들 및 하부 얼굴 특징들을 식별하기 위한 다른 얼굴 특징 인식 기술들을 포함할 수 있다.

[0149] [0155] 블록(1540b)에서, 컴퓨터 시스템은 스피치 프로세싱에 의해 오디오 세그먼트들로부터 음소들을 추출할 수 있다. 예컨대, 컴퓨터 시스템은 일정 시간 기간에 걸쳐 웨어러블 디바이스에 의해 레코딩된 오디오 세그먼트들을 수신할 수 있다. 컴퓨터 시스템은 스피치 프로세싱 기술들을 사용하여 스피치 프로세싱에서 음절들을

식별할 수 있다. 일부 실시예들에서, 음소들은 하부 얼굴 파라미터들을 결정하는 데 사용될 수 있는 비슴들과 추가로 연관될 수 있다.

[0150] [0156] 블록(1550b)에서, 컴퓨터 시스템은 머신 러닝 기술들을 사용하여 맵핑을 트레이닝할 수 있다. 안구주위 구역 파라미터들은 맵핑의 입력일 수 있는 반면, 하부 얼굴 파라미터들은 맵핑의 출력일 수 있다. 일부 실시예들에서, 입력은 또한 음소들 또는 오디오 스트림을 포함할 수 있다. 맵핑의 출력은 또한 전체 얼굴 파라미터들의 세트를 획득하기 위해 안구주위 구역 얼굴 파라미터들과 결합될 수 있는 하부 얼굴 파라미터를 포함할 수 있다.

[0151] [0157] 도 16a는 얼굴의 부분이 웨어러블 디바이스에 의해 관찰되지 않을 때 전체 얼굴 이미지를 합성하는 예시적인 프로세스를 설명한다. 프로세스(1600a)는 본원에서 설명된 웨어러블 디바이스에 의해 수행될 수 있다.

[0152] [0158] 블록(1610a)에서, 웨어러블 디바이스는 안구주위 구역의 컨퍼메이션에 액세스할 수 있다. 컨퍼메이션은 (예컨대, 얼굴 특징 벡터들과 같은) 얼굴 파라미터들을 사용하여 인코딩될 수 있다. 컨퍼메이션은 사용자가 말할 때 획득된 안구주위 이미지들에 의해 결정될 수 있다. 웨어러블 디바이스는 하부 얼굴의 컨퍼메이션을 생성하기 위해 머신 러닝 유도 모델에 액세스할 수 있다. 이 예에서, 하부 얼굴의 컨퍼메이션은 사용자가 말할 때 웨어러블 디바이스에 의해 이미징되지 않는다. 머신 러닝 유도 모델은 도 14에서 설명된 머신 러닝 유도 모델(1420)의 예일 수 있다. 예컨대, 머신 러닝 유도 모델은, 단독으로 또는 안구주위 컨퍼메이션과 조합하여, 오디오 입력과 하부 얼굴의 컨퍼메이션의 맵핑을 포함할 수 있다. 맵핑은 입력 및 출력 하부 얼굴 파라미터들로서 하나 이상의 안구주위 얼굴 파라미터들을 수용할 수 있다.

[0153] [0159] 블록(1620a)에서, 웨어러블 디바이스는 오디오 스트림을 수신할 수 있다. 웨어러블 디바이스는 사용자가 말할 때 사용자에게 의해 생성된 사운드를 캡처할 수 있는 오디오 센서(이들테면, 마이크로폰)를 포함할 수 있다.

[0154] [0160] 블록(1630a)에서, 웨어러블 디바이스는 하부 얼굴의 컨퍼메이션을 생성하기 위해 오디오 스트림에 머신 러닝 유도 모델을 적용할 수 있다. 예컨대, 머신 러닝 유도 모델은 음소와 비슴의 맵핑을 포함할 수 있다. 웨어러블 디바이스는 오디오 스트림에서 음소들을 식별하고 오디오 스트림으로부터 대응하는 비슴들을 생성할 수 있다. 비슴들은 또한 하나 이상의 얼굴 파라미터들(이들테면, 하부 얼굴 파라미터들)에 의해 인코딩될 수 있다. 웨어러블 디바이스는 또한 안구주위 구역의 이미지와 함께 오디오 입력을 사용하여 하부 얼굴의 대응하는 이미지를 식별할 수 있다. 일부 실시예들에서, 안구주위 구역의 이미지 또는 음소는 하부 얼굴의 다수의 이미지들과 연관될 수 있으며, 여기서 하부 얼굴의 각각의 이미지는 매칭의 가능성과 연관될 수 있다. 웨어러블 디바이스는 가장 높은 가능성을 갖는 하부 얼굴의 이미지를 고를 수 있다. 매칭의 가능성은 웨어러블 디바이스가 가지고 있는 양 정보에 기초하여 변할 수 있다. 예컨대, 웨어러블 디바이스가 안구주위 이미지만을 제공하는 경우와 비교하여 웨어러블 디바이스가 안구주위 이미지와 함께 사운드를 제공하는 경우, 이미지의 가능성은 증가하거나 감소할 수 있다.

[0155] [0161] 블록(1640a)에서, 웨어러블 디바이스는 전체 얼굴 이미지를 생성하기 위해 안구주위 구역의 컨퍼메이션 및 하부 얼굴의 컨퍼메이션을 결합할 수 있다. 블록(1640a)은 안구주위 이미지들 및 오디오 스트림을 획득하는 웨어러블 디바이스에 의해 선택적으로 수행될 수 있다. 웨어러블 디바이스가 한쪽 눈에 대한 이미지만을 갖는 일부 구현들에서, 웨어러블 디바이스는 다른 눈에 대한 이미지를 생성하고 다른 눈에 대한 이미지를 전체 얼굴 이미지에 통합할 수 있다. 웨어러블 디바이스는 또한 안구주위 구역 및 하부 얼굴의 변화들에 기초하여 기존 얼굴 이미지의 부분을 업데이트할 수 있다.

[0156] [0162] 선택적으로, 블록(1650a)에서, 웨어러블 디바이스는 이들테면, 텔레프레전스 세션 동안 다른 사용자의 웨어러블 디바이스에 얼굴 이미지를 통신할 수 있다. 일부 실시예들에서, 웨어러블 디바이스는 하부 얼굴 컨퍼메이션 및 안구주위 컨퍼메이션을 다른 사용자의 웨어러블 디바이스에 통신할 수 있다. 다른 사용자의 웨어러블 디바이스는 수신된 하부 얼굴 컨퍼메이션 및 안구주위 컨퍼메이션을 사용하여 전체 얼굴 이미지를 생성할 수 있다.

[0157] [0163] 웨어러블 디바이스는 또한 다른 사용자의 웨어러블 디바이스에 얼굴 이미지에 대한 업데이트를 통신할 수 있다. 얼굴 이미지에 대한 업데이트들은 새로고침 레이트(refresh rate)에 따를 수 있다. 예컨대, 사용자의 얼굴 이미지는 다른 사용자의 웨어러블 디바이스의 프레임 레이트, 또는 1/24 초마다, 1/48 초마다, 1 초마다, 2 초마다 등과 같은 미리 정의된 시간 간격으로 업데이트될 수 있다. 다른 사용자의 웨어러블 디바이스는

다른 사용자의 얼굴 이미지에 대한 수신된 업데이트들에 기초하여 사용자의 가상 아바타를 업데이트할 수 있다.

- [0158] [0164] 일부 구현들에서, 웨어러블 디바이스는 전체 얼굴 이미지를 합성하기 위해 다른 웨어러블 디바이스에, 하부 얼굴 및/또는 안구주위 구역의 컨퍼메이션을 인코딩하는 얼굴 파라미터들을 전달할 수 있다. 예컨대, 웨어러블 디바이스는 수신된 얼굴 파라미터들을 사용하여 전체 얼굴 이미지를 합성할 다른 웨어러블 디바이스에 (전체 얼굴 이미지 대신) 얼굴 파라미터들을 통신할 수 있다.
- [0159] [0165] 도 16b는 제1 사용자의 얼굴을 표현하는 파라미터들을 생성하고, 제1 사용자의 얼굴의 가상 표현을 생성하는 데 사용될 수 있는 제2 사용자의 웨어러블 디바이스에 파라미터들을 통신하는 예를 예시한다. 도 16b에 설명된 프로세스(1600b)는 본원에서 설명된 웨어러블 디바이스에 의해 수행될 수 있다.
- [0160] [0166] 블록(1610b)에서, 제1 사용자의 웨어러블 디바이스는 입력으로서 제1 사용자의 안구주위 이미지를 수신할 수 있다. 예컨대, 내향 이미징 시스템(462)은 제1 사용자의 안구주위 구역의 이미지를 획득하고, 블록(1620b)에서 안구주위 구역의 이미지로부터 얼굴 모드 파라미터들을 추출하기 위해 프로세서에 이미지를 통신할 수 있다.
- [0161] [0167] 블록(1630b)에서, 웨어러블 디바이스는 제1 사용자의 마이크로폰으로부터 입력 오디오 스트림을 수신할 수 있다. 마이크로폰은 웨어러블 디바이스의 부분일 수 있다. 블록(1640b)에서, 웨어러블 디바이스는 스피치 프로세싱에 의해 오디오 스트림으로부터 음소들을 추출할 수 있다. 음소들은 또한 비음들과 연관될 수 있다.
- [0162] [0168] 블록(1650b)에서, 안구주위 얼굴 파라미터들 및 음소들은 (예컨대, 블록(1550b)에서 트레이닝된 맵핑과 같은) 머신 러닝 기술들에 의해 트레이닝된 맵핑에 입력될 수 있다. 맵핑은 안구주위 얼굴 파라미터들 및 음소들에 대응하는 하부 얼굴 파라미터들의 세트를 출력할 수 있다.
- [0163] [0169] 블록(1660b)에서, 제1 사용자의 웨어러블 디바이스는 안구주위 얼굴 파라미터 및 하부 얼굴 파라미터들을 제2 사용자의 웨어러블 디바이스에 통신할 수 있다. 제2 사용자의 웨어러블 디바이스는 블록(1670)에서 제1 사용자의 얼굴의 이미지를 생성 및 디스플레이하기 위해 수신된 안구주위 얼굴 파라미터 및 하부 얼굴 파라미터들을 사용할 수 있다. 일부 실시예들에서, 제1 사용자의 웨어러블 디바이스는 제2 사용자의 웨어러블 디바이스에 (얼굴 파라미터들의 전체 세트 대신) 안구주위 얼굴 파라미터 또는 하부 얼굴 파라미터들에 대한 업데이트들을 전송한다. 따라서, 제2 사용자의 웨어러블 디바이스는 업데이트들에 기초하여 제1 사용자의 이미지를 상응하게 조정할 수 있다.
- [0164] 부가적인 양상들
- [0165] [0170] 제1 양상에서, 얼굴의 이미지를 생성하기 위한 방법으로서, 이 방법은, 컴퓨터 프로세서, 사용자의 안구주위 구역을 이미징하도록 구성된 내향 이미징 시스템 및 오디오 센서를 포함하는 HMD(head-mounted device)의 제어 하에서, 내향 이미징 시스템에 의해 사용자의 안구주위 구역의 이미지를 획득하는 단계; 이미지에 적어도 부분적으로 기초하여, 사용자의 적어도 안구주위 구역의 안구주위 컨퍼메이션을 인코딩하는 안구주위 얼굴 파라미터들을 생성하는 단계; 오디오 센서에 의해, 사용자에 의해 말해진 오디오 스트림을 획득하는 단계; 오디오 스트림에서 음소를 식별하는 단계; 맵핑에 액세스하는 단계 — 맵핑의 입력은 음소 및 사용자의 안구주위 구역의 이미지를 포함하고, 맵핑의 출력은 사용자의 적어도 하부 얼굴의 하부 얼굴 컨퍼메이션을 인코딩하는 하부 얼굴 파라미터들을 포함함 — ; 및 전체 얼굴 파라미터들을 생성하기 위해 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들을 결합하는 단계를 포함한다.
- [0166] [0171] 제2 양상에서, 제1 양상의 방법에 있어서, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들은 3-차원(3D) 얼굴 모델의 부분이다.
- [0167] [0172] 제3 양상에서, 제2 양상의 방법에 있어서, 3D 얼굴 모델은 변형 가능한 선형 모델을 포함하고, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들은 사용자가 말할 때 얼굴의 변형을 설명한다.
- [0168] [0173] 제4 양상에서, 제3 양상의 방법에 있어서, 전체 얼굴 파라미터들을 생성하는 단계는, 하부 얼굴 파라미터들 또는 안구주위 얼굴 파라미터들 중 적어도 하나에 대한 업데이트를 반영하도록 3D 얼굴 모델을 업데이트하는 단계를 포함한다.
- [0169] [0174] 제5 양상에서, 제1 양상 내지 제4 양상 중 어느 한 양상의 방법에 있어서, 맵핑의 입력은, 사용자의 눈

포즈, 동공 확장 상태, 눈 컬러 또는 눈꺼풀 상태 중 적어도 하나를 포함하는 눈 특정 정보를 더 포함한다.

- [0170] [0175] 제6 양상에서, 제1 양상 내지 제5 양상 중 어느 한 양상의 방법에 있어서, 하부 얼굴 파라미터들은 오디오 스트림의 음소들을 시각적으로 설명하는 비슴(viseme)들을 인코딩한다.
- [0171] [0176] 제7 양상에서, 제1 양상 내지 제6 양상 중 어느 한 양상의 방법에 있어서, 방법은, 전체 얼굴 파라미터들을 다른 사용자의 웨어러블 디바이스에 통신하는 단계를 더 포함한다.
- [0172] [0177] 제8 양상에서, 제1 양상 내지 제7 양상 중 어느 한 양상의 방법에 있어서, 맵핑은, 안구주위 얼굴 파라미터가 하부 얼굴 파라미터와 연관되는 가능성을 포함하고, 하부 얼굴 파라미터들은 이들이 임계 기준들을 통과한다는 결정에 대한 응답으로 전체 얼굴 이미지를 생성하도록 선택된다.
- [0173] [0178] 제9 양상에서, 제1 양상 내지 제8 양상 중 어느 한 양상의 방법에 있어서, 방법은, 안구주위 구역의 이미지에 적어도 부분적으로 기초하여 사용자의 얼굴의 피부 텍스처를 추론하는 단계를 더 포함한다.
- [0174] [0179] 제10 양상에서, 제1 양상 내지 제9 양상 중 어느 한 양상의 방법에 있어서, 내향 이미징 시스템은 눈 카메라를 포함하고, 내향 이미징 시스템에 의해 획득된 안구주위 구역의 이미지는 제1 눈에 대한 안구주위 구역의 이미지를 포함한다.
- [0175] [0180] 제11 양상에서, 제10 양상의 방법에 있어서, 전체 얼굴 파라미터들을 생성하는 단계는, 내향 이미징 시스템에 의해 획득된 안구주위 구역의 이미지에 기초하여 제2 눈에 대한 안구주위 얼굴 파라미터들을 결정하는 단계; 및 제2 눈에 대한 안구주위 얼굴 파라미터들을 전체 얼굴 파라미터들에 통합하는 단계를 포함한다.
- [0176] [0181] 제12 양상에서, 제1 양상 내지 제11 양상 중 어느 한 양상의 방법에 있어서, 맵핑의 입력은 신체 움직임 또는 심박수를 더 포함한다.
- [0177] [0182] 제13 양상에서, 제1 양상 내지 제12 양상 중 어느 한 양상의 방법에 있어서, 이미지는 사진, 비디오 프레임 또는 비디오 중 적어도 하나를 포함한다.
- [0178] [0183] 제14 양상에서, 얼굴의 이미지를 생성하기 위한 방법으로서, 이 방법은, 컴퓨터 프로세서, 사용자의 안구주위 구역을 이미징하도록 구성된 내향 이미징 시스템을 포함하는 HMD(head-mounted device)의 제어 하에서, 내향 이미징 시스템에 의해 사용자의 안구주위 구역의 안구주위 이미지를 획득하는 단계; 이미지에 적어도 부분적으로 기초하여, 사용자의 적어도 안구주위 구역의 안구주위 컨퍼메이션을 인코딩하는 안구주위 얼굴 파라미터들을 생성하는 단계; 사용자의 적어도 하부 얼굴의 하부 얼굴 컨퍼메이션을 인코딩하는 하부 얼굴 파라미터들을 생성하는 맵핑에 액세스하는 단계 - 맵핑은 사용자의 안구주위 구역의 안구주위 이미지에 적어도 부분적으로 기초함 - ; 전체 얼굴 파라미터들을 생성하기 위해 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들을 결합하는 단계; 및 전체 얼굴 파라미터들에 적어도 부분적으로 기초하여 얼굴의 이미지를 생성하는 단계를 포함한다.
- [0179] [0184] 제15 양상에서, 제14 양상의 방법에 있어서, 하부 얼굴 컨퍼메이션은 비슴을 포함하고, 맵핑은 안구주위 이미지 및 오디오 입력에 적어도 부분적으로 기초하여 하부 얼굴 파라미터들을 생성한다.
- [0180] [0185] 제16 양상에서, 제14 양상 또는 제15 양상의 방법에 있어서, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들은 3-차원(3D) 얼굴 모델의 부분이다.
- [0181] [0186] 제17 양상에서, 제16 양상의 방법에 있어서, 3D 얼굴 모델은 변형 가능한 선형 모델을 포함하고, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들은 사용자가 말할 때 얼굴의 변형을 설명하는,
- [0182] [0187] 제18 양상에서, 제17 양상의 방법에 있어서, 전체 얼굴 파라미터들을 생성하는 단계는, 하부 얼굴 파라미터들 또는 안구주위 얼굴 파라미터들 중 적어도 하나에 대한 업데이트를 반영하도록 3D 얼굴 모델을 업데이트하는 단계를 포함한다.
- [0183] [0188] 제19 양상에서, 제14 양상 내지 제18 양상 중 어느 한 양상의 방법에 있어서, 맵핑의 입력은 눈 특정 정보를 더 포함하고, 눈 특정 정보는 사용자의 눈 포즈, 동공 확장 상태, 눈 컬러 또는 눈꺼풀 상태 중 적어도 하나를 포함한다.
- [0184] [0189] 제20 양상에서, 제14 양상 내지 제19 양상 중 어느 한 양상의 방법에 있어서, 전체 얼굴 파라미터들을 다른 사용자의 웨어러블 디바이스에 통신하는 단계를 더 포함한다.
- [0185] [0190] 제21 양상에서, 제14 양상 내지 제20 양상 중 어느 한 양상의 방법에 있어서, 맵핑은, 안구주위 얼굴

파라미터가 하부 얼굴 파라미터와 연관되는 가능성을 포함하고, 하부 얼굴 파라미터들은 이들이 임계 기준들을 통과한다는 결정에 대한 응답으로 전체 얼굴 이미지를 생성하도록 선택된다.

- [0186] [0191] 제22 양상에서, 제14 양상 내지 제21 양상 중 어느 한 양상의 방법에 있어서, 안구주위 구역의 이미지에 적어도 부분적으로 기초하여 사용자의 얼굴의 피부 텍스처를 추론하는 단계를 더 포함한다.
- [0187] [0192] 제23 양상에서, 제14 양상 내지 제22 양상 중 어느 한 양상의 방법에 있어서, 맵핑의 입력은 신체 움직임 또는 심박수를 더 포함한다.
- [0188] [0193] 제24 양상에서, 제14 양상 내지 제23 양상 중 어느 한 양상의 방법에 있어서, 전체 얼굴 파라미터들을 생성하는 단계는, 내향 이미징 시스템에 의해 획득된 안구주위 구역의 이미지에 기초하여 제2 눈에 대한 안구주위 얼굴 파라미터들을 결정하는 단계; 및 제2 눈에 대한 안구주위 얼굴 파라미터들을 전체 얼굴 파라미터들에 통합하는 단계를 포함한다.
- [0189] [0194] 제25 양상에서, 제14 양상 내지 제24 양상 중 어느 한 양상의 방법에 있어서, 이미지를 생성하는 단계는 미리 정의된 시간 간격에서 사용자의 얼굴 이미지를 업데이트하는 단계를 포함한다.
- [0190] [0195] 제26 양상에서, 제14 양상 내지 제25 양상 중 어느 한 양상의 방법에 있어서, 이미지는 정지 이미지 또는 비디오 프레임 중 적어도 하나를 포함한다.
- [0191] [0196] 제27 양상에서, 얼굴의 이미지를 생성하기 위한 시스템으로서, 이 시스템은, 사용자의 안구주위 구역을 이미징하도록 구성된 내향 이미징 시스템, 오디오 센서, 및 제1 양상 내지 제26 양상의 방법들 중 임의의 하나를 수행하도록 구성된 컴퓨터 프로세서를 포함한다.
- [0192] [0197] 제28 양상에서, 사용자의 얼굴의 안구주위 구역의 제1 컨퍼메이션과 사용자의 얼굴의 제2 컨퍼메이션의 맵핑을 생성하기 위한 방법으로서, 이 방법은, 하드웨어 컴퓨터 시스템의 제어 하에서, 안구주위 구역의 제1 컨퍼메이션과 연관된 제1 복수의 이미지들에 액세스하는 단계; 얼굴의 제2 컨퍼메이션과 연관된 제2 복수의 이미지들을 결정하는 단계; 안구주위 구역의 제1 컨퍼메이션과 얼굴의 제2 컨퍼메이션의 맵핑을 생성하기 위해 제1 복수의 이미지들 및 제2 복수의 이미지들을 사용하여 머신 러닝 유도 모델을 트레이닝하는 단계; 맵핑을 사용자의 증강 현실 디바이스에 출력하는 단계를 포함하고, 증강 현실 디바이스는 사용자의 관찰된 제1 컨퍼메이션을 맵핑에 입력하고 사용자의 예측된 제2 컨퍼메이션의 출력을 수신한다.
- [0193] [0198] 제29 양상에서, 제28 양상의 방법에 있어서, 얼굴의 제2 컨퍼메이션은 사용자의 하부 얼굴의 컨퍼메이션을 포함한다.
- [0194] [0199] 제30 양상에서, 제28 양상 또는 제29 양상의 방법에 있어서, 제1 컨퍼메이션은 안구주위 얼굴 파라미터들에 의해 인코딩되고, 제2 컨퍼메이션은 하부 얼굴 파라미터들 또는 전체 얼굴 파라미터들 중 적어도 하나에 의해 인코딩된다.
- [0195] [0200] 제31 양상에서, 제30 양상의 방법에 있어서, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들은 3-차원(3D) 얼굴 모델의 부분이다.
- [0196] [0201] 제32 양상에서, 제31 양상의 방법에 있어서, 3D 얼굴 모델은 변형 가능한 선형 모델이다.
- [0197] [0202] 제33 양상에서, 제28 양상 내지 제32 양상 중 어느 한 양상의 방법에 있어서, 제2 복수의 이미지들은, 머리-장착 디바이스의 외향 이미징 시스템 또는 외향 이미징 시스템 외부의 카메라 중 적어도 하나에 의해 획득된다.
- [0198] [0203] 제34 양상에서, 제28 양상 내지 제33 양상 중 어느 한 양상의 방법에 있어서, 제1 복수의 이미지들에 대응하는 오디오 스트림에 액세스하는 단계를 더 포함하며, 제2 복수의 이미지들은 오디오 스트림과 연관된 음소들을 식별하고 식별된 음소들과 연관된 비스듬들에 액세스함으로써 결정된 비스듬들을 포함한다.
- [0199] [0204] 제35 양상에서, 제28 양상 내지 제34 양상 중 어느 한 양상의 방법에 있어서, 머신 러닝 유도 모델은 얼굴의 컨퍼메이션이 안구주위 구역의 컨퍼메이션과 매칭하는 가능성을 포함한다.
- [0200] [0205] 제36 양상에서, 제28 양상 내지 제35 양상 중 어느 한 양상의 방법에 있어서, 머신 러닝 유도 모델은 제1 복수의 이미지들 및 제2 복수의 이미지들과 연관된 오디오 스트림, 또는 눈 특정 정보 중 적어도 하나를 사용하여 추가로 트레이닝된다.
- [0201] [0206] 제37 양상에서, 사용자의 얼굴의 제1 구역과 사용자의 얼굴의 제2 구역의 맵핑을 생성하기 위한 방법으

로서, 이 방법은, 하드웨어 컴퓨터 시스템의 제어 하에서, 제1 구역의 제1 컨퍼메이션들과 연관된 제1 복수의 이미지들에 액세스하는 단계; 제2 구역의 제2 컨퍼메이션들과 연관된 제2 복수의 이미지들을 결정하는 단계 - 제1 구역의 제1 컨퍼메이션들 각각은 제2 구역의 제2 컨퍼메이션들 중 적어도 하나와 연관됨 - ; 및 제1 구역의 제1 컨퍼메이션과 제2 구역의 제2 컨퍼메이션의 맵핑을 생성하기 위해 제1 복수의 이미지들 및 제2 복수의 이미지들을 사용하여 머신 러닝 유도 모델을 트레이닝하는 단계를 포함한다.

- [0202] [0207] 제38 양상에서, 제37 양상의 방법에 있어서, 제1 구역은 안구주위 구역이고, 제2 구역은 전체 얼굴 또는 하부 얼굴 구역이다.
- [0203] [0208] 제39 양상에서, 제37 양상의 방법에 있어서, 제1 구역은 전체 얼굴 또는 하부 얼굴 구역이고 제2 영역은 안구주위 구역이다.
- [0204] [0209] 제40 양상에서, 제38 양상 또는 제39 양상의 방법에 있어서, 전체 얼굴 및 하부 얼굴 구역의 이미지들은, 머리-장착 디바이스의 외향 이미징 시스템 또는 외향 이미징 시스템 외부의 카메라 중 적어도 하나에 의해 획득된다.
- [0205] [0210] 제41 양상에서, 제37 양상 내지 제40 양상 중 어느 한 양상의 방법에 있어서, 제1 구역은 머리-장착 디바이스에 의해 관찰 가능하고, 제2 구역은 머리-장착 디바이스에 의해 관찰 불가능하다.
- [0206] [0211] 제42 양상에서, 제37 양상 내지 제41 양상 중 어느 한 양상의 방법에 있어서, 제1 컨퍼메이션 및 제2 컨퍼메이션은 얼굴 파라미터들에 의해 인코딩된다.
- [0207] [0212] 제43 양상에서, 제42 양상의 방법에 있어서, 얼굴 파라미터들은 3-차원(3D) 얼굴 모델의 부분이다.
- [0208] [0213] 제44 양상에서, 제37 양상 내지 제43 양상 중 어느 한 양상의 방법에 있어서, 머신 러닝 유도 모델은 제1 구역의 제1 컨퍼메이션이 제2 구역의 제2 컨퍼메이션과 매칭하는 가능성을 포함한다.
- [0209] [0214] 제45 양상에서, 제37 양상 내지 제44 양상 중 어느 한 양상의 방법에 있어서, 머신 러닝 유도 모델은 제1 복수의 이미지들 및 제2 복수의 이미지들과 연관된 오디오 스트림, 또는 눈 특정 정보 중 적어도 하나를 사용하여 추가로 트레이닝된다.
- [0210] [0215] 제46 양상에서, 제37 양상 내지 제45 양상 중 어느 한 양상의 방법에 있어서, 맵핑은 사용자의 하부 얼굴의 컨퍼메이션을 포함하는 입력 및 안구주위 구역의 컨퍼메이션들을 포함하는 출력을 포함한다.
- [0211] [0216] 제47 양상에서, 제46 양상의 방법에 있어서, 방법은, 사용자의 하부 얼굴의 이미지에 기초하여 안구주위 구역의 이미지를 생성하기 위해 웨어러블 디바이스에 매핑을 통신하는 단계를 더 포함한다.
- [0212] [0217] 제48 양상에서, 사용자의 얼굴의 제1 구역과 사용자의 얼굴의 제2 구역의 맵핑을 생성하기 위한 시스템으로서, 이 시스템은, 제28 양상 내지 제47 양상의 방법들 중 임의의 하나를 수행하도록 구성된 하드웨어 프로세서를 포함한다.
- [0213] [0218] 제49 양상에서, 스피치(speech) 동안 사용자의 얼굴을 애니메이션(animating)하기 위한 웨어러블 시스템으로서, 웨어러블 시스템은, 사용자의 안구주위 구역을 이미징하도록 구성된 내향 이미징 시스템; 사용자의 스피치를 수신하도록 구성된 오디오 센서; 하드웨어 프로세서를 포함하고, 하드웨어 프로세서는, 내향 이미징 시스템을 통해, 사용자의 안구주위 구역의 이미지를 획득하고; 이미지에 적어도 부분적으로 기초하여, 사용자의 적어도 안구주위 구역의 안구주위 컨퍼메이션을 인코딩하는 안구주위 얼굴 파라미터들을 생성하고; 오디오 센서에 의해, 사용자에게 의해 말해진 오디오 스트림을 획득하고; 오디오 스트림에서 음소를 식별하고; 맵핑에 액세스하고 - 맵핑의 입력은 음소 및 사용자의 안구주위 구역의 이미지를 포함하고, 맵핑의 출력은 사용자의 적어도 하부 얼굴의 하부 얼굴 컨퍼메이션을 인코딩하는 하부 얼굴 파라미터들을 포함하고, 하부 얼굴은 사용자에게 의해 착용될 때 내향 이미징 시스템에 의해 관찰되지 않음 - ; 및 하부 얼굴 파라미터들을 생성하도록 이미지에 맵핑을 적용하고; 3-차원(3D) 얼굴 모델과 연관된 전체 얼굴 파라미터들을 생성하기 위해 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들을 결합하고; 그리고 전체 얼굴 파라미터들에 적어도 부분적으로 기초하여 사용자의 얼굴의 애니메이션을 생성하도록 프로그래밍된다.
- [0214] [0219] 제50 양상에서, 제49 양상의 웨어러블 시스템에 있어서, 3D 얼굴 모델은 변형 가능한 선형 모델을 포함하고, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들은 사용자가 말할 때 얼굴의 변형을 설명한다.
- [0215] [0220] 제51 양상에서, 제50 양상의 웨어러블 시스템에 있어서, 전체 얼굴 파라미터들을 생성하기 위해, 하드웨어 프로세서는 하부 얼굴 파라미터들 또는 안구주위 얼굴 파라미터들 중 적어도 하나에 대한 업데이트를 반영

하도록 3D 얼굴 모델을 업데이트하도록 프로그래밍된다.

- [0216] [0221] 제52 양상에서, 제49 양상 내지 제51 양상 중 어느 한 양상의 웨어러블 시스템에 있어서, 맵핑의 입력은 눈 특정 정보, 신체 움직임 또는 심박수 중 적어도 하나를 더 포함한다.
- [0217] [0222] 제53 양상에서, 제52 양상의 웨어러블 시스템에 있어서, 눈 특정 정보는 사용자의 눈 포즈, 동공 확장 상태, 눈 컬러 또는 눈꺼풀 상태 중 적어도 하나를 포함한다.
- [0218] [0223] 제54 양상에서, 제49 양상 내지 제53 양상 중 어느 한 양상의 웨어러블 시스템에 있어서, 하부 얼굴 파라미터들은 오디오 스트림의 음소들을 시각적으로 설명하는 비습들을 인코딩한다.
- [0219] [0224] 제55 양상에서, 제49 양상 내지 제54 양상 중 어느 한 양상의 웨어러블 시스템에 있어서, 하드웨어 프로세서는 추가로, 프로그래밍되는, 안구주위 구역의 이미지에 적어도 부분적으로 기초하여 사용자의 얼굴의 피부 텍스처를 추론하도록, 사용자의 얼굴의 애니메이션은 얼굴의 피부 텍스처를 통합한다.
- [0220] [0225] 제56 양상에서, 제49 양상 내지 제55 양상 중 어느 한 양상의 웨어러블 시스템에 있어서, 내향 이미징 시스템은 눈 카메라를 포함하고, 내향 이미징 시스템에 의해 획득된 안구주위 구역의 이미지는 제1 눈에 대한 안구주위 구역의 이미지를 포함한다.
- [0221] [0226] 제57 양상에서, 제56 양상의 웨어러블 시스템에 있어서, 전체 얼굴 파라미터를 생성하기 위해, 하드웨어 프로세서는, 내향 이미징 시스템에 의해 획득된 안구주위 구역의 이미지에 기초하여 제2 눈에 대한 안구주위 얼굴 파라미터들을 결정하고; 그리고 제2 눈에 대한 안구주위 얼굴 파라미터들을 전체 얼굴 파라미터들에 통합하도록 프로그래밍된다.
- [0222] [0227] 제58 양상에서, 제49 양상 내지 제57 양상 중 어느 한 양상의 웨어러블 시스템에 있어서, 사용자의 얼굴의 애니메이션을 생성하도록 전체 얼굴 파라미터들이 적용되게 하기 위해, 하드웨어 프로세서는 혼합 현실 디스플레이를 포함하는 웨어러블 디바이스에 명령들을 통신하도록 프로그래밍되고, 명령들은 중립 포지션으로부터 3D 얼굴 모델을 변동시키도록 전체 얼굴 파라미터들이 적용되게 한다.
- [0223] [0228] 제59 양상에서, 스피치 동안 사용자의 얼굴을 애니메이션하기 위한 방법으로서, 이 방법은, 사용자의 안구주위 구역을 이미징하도록 구성된 내향 이미징 시스템에 의해 획득된 이미지에 액세스하는 단계; 이미지들에 적어도 부분적으로 기초하여, 사용자의 적어도 안구주위 구역의 안구주위 컨퍼메이션을 인코딩하는 안구주위 얼굴 파라미터들을 결정하는 단계; 오디오 센서에 의해 획득된, 사용자에게 의해 말해진 오디오 스트림에 액세스하는 단계; 오디오 스트림에서 음소를 식별하는 단계; 맵핑에 액세스하는 단계 - 맵핑의 입력은 음소 및 사용자의 안구주위 구역의 이미지를 포함하고, 맵핑의 출력은 사용자의 적어도 하부 얼굴의 하부 얼굴 컨퍼메이션을 인코딩하는 하부 얼굴 파라미터들을 포함함 - ; 하부 얼굴 파라미터들을 생성하도록 이미지에 맵핑을 적용하는 단계; 3-차원(3D) 얼굴 모델과 연관된 전체 얼굴 파라미터들을 생성하기 위해 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들을 결합하는 단계; 및 전체 얼굴 파라미터들에 적어도 부분적으로 기초하여 전체 얼굴 이미지를 생성하는 단계를 포함한다.
- [0224] [0229] 제60 양상에서, 제59 양상의 방법에 있어서, 3D 얼굴 모델은 변형 가능한 선형 모델을 포함하고, 안구주위 얼굴 파라미터들 및 하부 얼굴 파라미터들은 사용자가 말할 때 얼굴의 변형을 설명한다.
- [0225] [0230] 제61 양상에서, 제60 양상의 방법에 있어서, 전체 얼굴 파라미터들을 생성하기 위해, 하드웨어 프로세서는 하부 얼굴 파라미터들 또는 안구주위 얼굴 파라미터들 중 적어도 하나에 대한 업데이트를 반영하도록 3D 얼굴 모델을 업데이트하도록 프로그래밍된다.
- [0226] [0231] 제62 양상에서, 제59 양상 내지 제61 양상 중 어느 한 양상의 방법에 있어서, 전체 얼굴 파라미터는 사용자의 얼굴과 연관된 애니메이션을 결정하기 위해 눈 특정 정보와 결합된다.
- [0227] [0232] 제63 양상에서, 제62 양상의 방법에 있어서, 눈 특정 정보는 사용자의 눈 포즈, 동공 확장 상태, 눈 컬러 또는 눈꺼풀 상태 중 적어도 하나를 포함한다.
- [0228] [0233] 제64 양상에서, 제59 양상 내지 제63 양상 중 어느 한 양상의 방법에 있어서, 하부 얼굴 파라미터들은 오디오 스트림의 음소들을 시각적으로 설명하는 비습들을 인코딩한다.
- [0229] [0234] 제65 양상에서, 제59 양상 내지 제64 양상 중 어느 한 양상의 방법에 있어서, 전체 얼굴 이미지는 추가로, 내향 이미징 시스템에 의해 획득된 이미지에 적어도 부분적으로 기초하여 결정되는 사용자의 피부 텍스처들을 통합한다.

- [0230] [0235] 제66 양상에서, 제59 양상 내지 제65 양상 중 어느 한 양상의 방법에 있어서, 맵핑은, 안구주위 얼굴 파라미터가 하부 얼굴 파라미터와 연관되는 가능성을 포함하고, 하부 얼굴 파라미터들은 이들이 임계 기준들을 통과한다는 결정에 대한 응답으로 전체 얼굴 이미지를 생성하도록 선택된다.
- [0231] [0236] 제67 양상에서, 제59 양상 내지 제66 양상 중 어느 한 양상의 방법에 있어서, 이미지는 정지 이미지 또는 비디오 프레임 중 적어도 하나를 포함한다.
- [0232] [0237] 제68 양상에서, 제59 양상 내지 제67 양상 중 어느 한 양상의 방법에 있어서, 방법은, 혼합 현실 환경에서 전체 얼굴 이미지를 렌더링하도록 머리-장착 디스플레이에 지시하는 단계를 더 포함한다.
- [0233] 다른 고려사항들
- [0234] [0238] 본원에서 설명되고 그리고/또는 첨부 도면들에 도시되는 프로세스들, 방법들 및 알고리즘들 각각은 하나 이상의 물리적 컴퓨팅 시스템들, 하드웨어 컴퓨터 프로세서들, 주문형 회로 및/또는 특유 및 특정 컴퓨터 명령들을 실행하도록 구성된 전자 하드웨어에 의해 실행되는 코드 모듈들로 구현되고, 이 코드 모듈들에 의해 완전히 또는 부분적으로 자동화될 수 있다. 예컨대, 컴퓨팅 시스템들은 특정 컴퓨터 명령들로 프로그래밍된 범용 컴퓨터들(예컨대, 서버들) 또는 특수 목적 컴퓨터들, 특수 목적 회로 등을 포함할 수 있다. 코드 모듈은 실행 가능 프로그램으로 컴파일되어 링크되거나, 동적 링크 라이브러리에 설치될 수 있거나, 또는 인터프리팅된 프로그래밍 언어로 작성될 수 있다. 일부 구현들에서, 특정한 동작들 및 방법들은, 주어진 기능에 특정한 회로에 의해 수행될 수 있다.
- [0235] [0239] 추가로, 본 개시내용의 기능성의 소정의 구현들은 충분히 수학적으로, 계산상으로 또는 기술적으로 복잡하여, (적절한 전문화된 실행 가능한 명령들을 활용하는) 주문형 하드웨어 또는 하나 이상의 물리적 컴퓨팅 디바이스들은 예컨대, 수반되는 계산들의 양(volume) 또는 복잡성으로 인해 또는 실질적으로 실시간으로 결과들을 제공하기 위해 그 기능성들을 수행할 필요가 있을 수 있다. 예컨대, 애니메이션들 또는 비디오는 다수의 프레임들(각각의 프레임은 수백만개의 픽셀들을 가짐)을 포함할 수 있고, 상업적으로 합리적인 시간량 내에 원하는 이미지 프로세싱 태스크 또는 애플리케이션을 제공하기 위해, 특별히 프로그래밍된 컴퓨터 하드웨어가 비디오 데이터를 프로세싱할 필요가 있다.
- [0236] [0240] 코드 모듈들 또는 임의의 유형의 데이터는, 임의의 유형의 비-일시적인 컴퓨터-판독 가능 매체, 이를테면, 하드 드라이브들, 솔리드 스테이트 메모리, RAM(random access memory), ROM(read only memory), 광학 디스크, 휘발성 또는 비-휘발성 저장소, 이들의 조합들 등을 포함하는 물리적 컴퓨터 저장소 상에 저장될 수 있다. 방법들 및 모듈들(또는 데이터)은 또한, 생성된 데이터 신호들로서(예컨대, 반송파 또는 다른 아날로그 또는 디지털 전파 신호의 일부로서) 무선-기반 및 유선/케이블-기반 매체들을 포함하는 다양한 컴퓨터-판독 가능 송신 매체들 상에서 송신될 수 있고, (예컨대, 단일 또는 멀티플렉싱된 아날로그 신호의 일부로서, 또는 다수의 이산 디지털 패킷들 또는 프레임들로서) 다양한 형태들을 취할 수 있다. 개시된 프로세스들 또는 프로세스 단계들의 결과들은 임의의 유형의 비-일시적인 유형의(tangible) 컴퓨터 저장소에 지속적으로 또는 다른 방식으로 저장될 수 있거나, 또는 컴퓨터-판독 가능 송신 매체를 통해 통신될 수 있다.
- [0237] [0241] 본원에서 설명되고 그리고/또는 첨부된 도면들에 도시되는 흐름도들에서의 임의의 프로세스들, 블록들, 상태들, 단계들 또는 기능성들은 프로세스의 단계들 또는 (예컨대, 논리적 또는 산술적) 특정 기능들을 구현하기 위한 하나 이상의 실행 가능 명령들을 포함하는 코드 모듈들, 세그먼트들 또는 코드 부분들을 잠재적으로 나타내는 것으로 이해되어야 한다. 다양한 프로세스들, 블록들, 상태들, 단계들 또는 기능성들은 본원에서 제공된 예시적인 예들에서 조합되거나, 재배열되거나, 이들에 부가되거나, 이들로부터 제거되거나, 수정되거나, 또는 다른 방식으로 변할 수 있다. 일부 실시예들에서, 부가적인 또는 상이한 컴퓨팅 시스템들 또는 코드 모듈들은 본원에서 설명된 기능성들 중 일부 또는 전부를 수행할 수 있다. 본원에 설명된 방법들 및 프로세스들은 또한 임의의 특정 시퀀스로 제한되지 않고, 그에 관련된 블록들, 단계들 또는 상태들은 적절한 다른 시퀀스들로, 예컨대, 직렬로, 병렬로 또는 일부 다른 방식으로 수행될 수 있다. 태스크들 또는 이벤트들은 개시된 예시적인 실시예들에 부가되거나 그로부터 제거될 수 있다. 또한, 본원에서 설명된 구현들에서의 다양한 시스템 컴포넌트들의 분리는 예시 목적들을 위한 것이며, 모든 구현들에서 이러한 분리를 요구하는 것으로 이해되어서는 안 된다. 설명된 프로그램 컴포넌트들, 방법들 및 시스템들은 일반적으로 단일 컴퓨터 제품에 함께 통합되거나 다수의 컴퓨터 제품들로 패키징될 수 있다는 것이 이해되어야 한다. 다수의 구현 변동들이 가능하다.
- [0238] [0242] 프로세스들, 방법들 및 시스템들은 네트워크(또는 분산형) 컴퓨팅 환경에서 구현될 수 있다. 네트워크 환경들은, 전사적(enterprise-wide) 컴퓨터 네트워크들, 인트라넷들, LAN(Local Area Network)들, WAN(Wide

Area Network)들, PAN(Personal Area Network)들, 클라우드 컴퓨팅 네트워크들, 클라우드-소스(crowd-sourced) 컴퓨팅 네트워크들, 인터넷 및 월드 와이드 웹(World Wide Web)을 포함한다. 네트워크는 유선 또는 무선 네트워크 또는 임의의 다른 유형의 통신 네트워크일 수 있다.

[0239] [0243] 본 개시내용의 시스템들 및 방법들 각각은 몇몇 혁신적인 양상들을 가지며, 그 양상들 중 어떠한 단일 양상도 본원에서 개시된 바람직한 속성들을 전적으로 담당하거나 이를 위해 요구되지 않는다. 위에서 설명된 다양한 특징들 및 프로세스들은 서로 독립적으로 사용될 수 있거나, 또는 다양한 방식으로 조합될 수 있다. 모든 가능한 조합들 및 서브조합들은 본 개시내용의 범위 내에 속하는 것으로 의도된다. 본 개시내용에서 설명된 구현들에 대한 다양한 수정들은 당업자들에게 자명할 수 있으며, 본원에서 정의된 일반적인 원리들은 본 개시내용의 사상 또는 범위를 벗어나지 않으면서 다른 구현들에 적용될 수 있다. 따라서, 청구항들은 본 명세서에 도시된 구현들로 제한되도록 의도되는 것이 아니라, 본 명세서에 개시된 본 개시내용, 원리들 및 신규한 특성들과 일치하는 가장 넓은 범위에 부합할 것이다.

[0240] [0244] 별개의 구현들의 맥락에서 본 명세서에 설명된 소정의 특징들은 또한, 단일 구현으로 조합하여 구현될 수 있다. 대조적으로, 단일 구현의 맥락에서 설명된 다양한 특징들은 또한, 별개로 다수의 구현들로 또는 임의의 적절한 서브조합으로 구현될 수 있다. 더욱이, 특징들이 소정의 조합들로 작용하는 것으로 위에서 설명되고 심지어 초기에 이와 같이 청구될 수 있지만, 일부 경우들에서, 청구된 조합으로부터의 하나 이상의 특징들은 그 조합으로부터 제거될 수 있고, 청구된 조합은 서브조합 또는 서브조합의 변동에 관련될 수 있다. 단일 특징 또는 특징들의 그룹이 각각의 그리고 모든 각각의 실시예에 필요하거나 필수적인 것은 아니다.

[0241] [0245] 구체적으로 달리 언급되지 않거나 또는 사용된 맥락 내에서 달리 이해되지 않으면, 본원에서 사용된 조건어, 이를테면, 다른 것들 중에서도, "할 수 있다(can, could, might, may)", "예컨대" 등은 일반적으로, 소정의 실시예들이 소정의 특징들, 엘리먼트들, 및/또는 단계들을 포함하지만 다른 실시예들은 이들을 포함하지 않는다는 것을 전달하도록 의도된다. 따라서, 그러한 조건어는 일반적으로, 특징들, 엘리먼트들, 및/또는 단계들이 하나 이상의 실시예들을 위해 어떤 식으로든 요구된다는 것을, 또는 하나 이상의 실시예들이, 저자 입력 또는 프롬프팅(prompting)을 이용하거나 또는 그러한 것을 이용함이 없이, 이들 특징들, 엘리먼트들, 및/또는 단계들이 임의의 특정 실시예에 포함되는지 또는 임의의 특정 실시예들에서 수행되어야 하는지를 판단하기 위한 로직을 반드시 포함한다는 것을 암시하도록 의도되진 않는다. "포함하는(comprising, including), "갖는(having)" 등의 용어들은 동의어이며, 오픈-엔드(open-ended) 방식으로 포괄적으로 사용되며, 추가적인 엘리먼트들, 특징들, 행동들, 동작들 등을 배제하지 않는다. 또한, "또는"이라는 용어는 (그의 배타적인 의미가 아니라) 그의 포괄적인 의미로 사용되어서, 예컨대, 리스트의 엘리먼트들을 연결하기 위해 사용될 때, "또는"이라는 용어는 리스트 내의 엘리먼트들 중 하나, 일부, 또는 전부를 의미한다. 또한, 본 명세서 및 첨부된 청구항들에서 사용된 바와 같은 단수 표현은 달리 특정되지 않는 한 "하나 이상" 또는 "적어도 하나"를 의미하는 것으로 해석될 것이다.

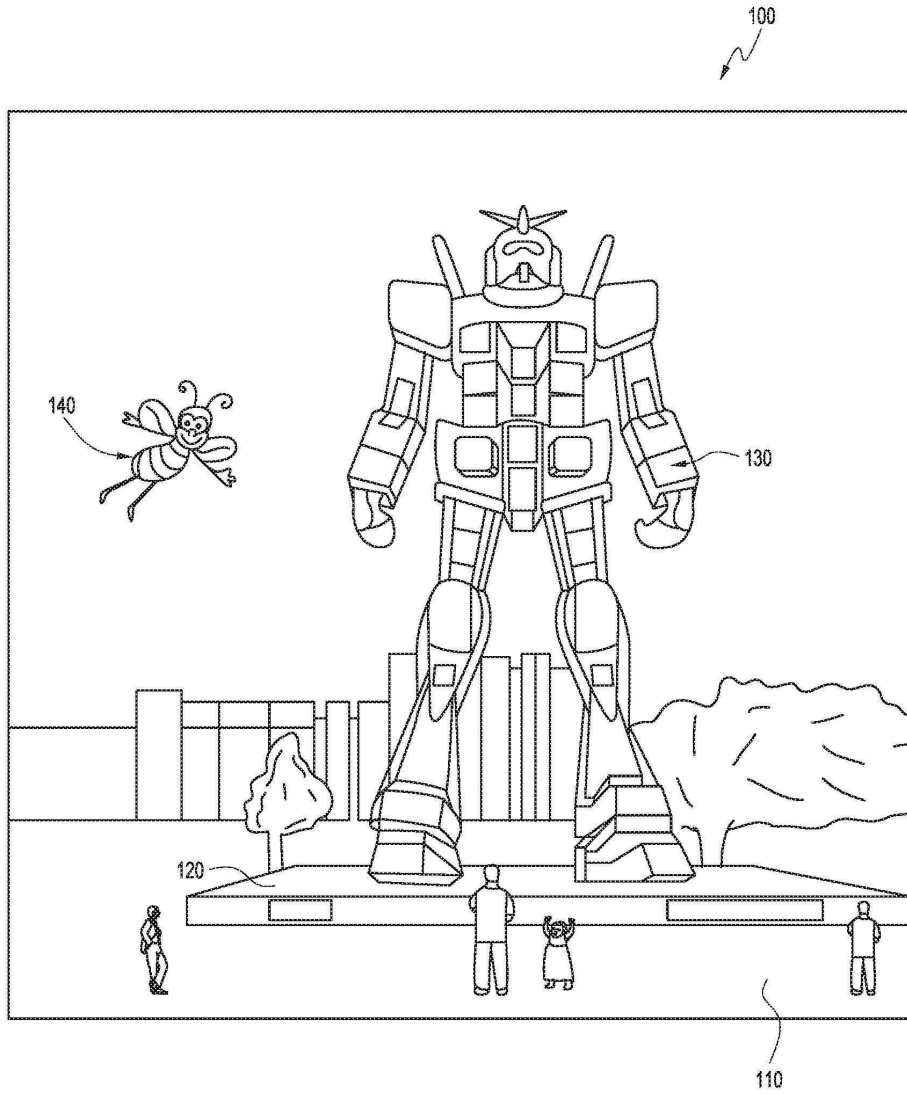
[0242] [0246] 본원에서 사용된 바와 같이, 리스트의 아이템들 “중 적어도 하나”를 지칭하는 어구는 단일 멤버들을 포함하여 그 아이템들의 임의의 조합을 지칭한다. 예로서, "A, B 또는 C 중 적어도 하나"는 A; B; C; A와 B; A와 C; B와 C; 그리고 A와 B와 C를 커버하는 것으로 의도된다. 특정하게 다르게 언급되지 않으면, 어구 "X, Y 또는 Z 중 적어도 하나"와 같은 접속어는,아이템, 용어 등이 X, Y 또는 Z 중 적어도 하나일 수 있다는 것을 전달하기 위해 일반적으로 사용되는 맥락으로 달리 이해된다. 따라서, 이러한 접속어는 일반적으로, 소정의 실시예들이 X 중 적어도 하나, Y 중 적어도 하나 및 Z 중 적어도 하나가 각각 존재할 것을 요구하는 것을 암시하는 것으로 의도되지 않는다.

[0243] [0247] 유사하게, 동작들이 특정한 순서로 도면들에 도시될 수 있지만, 원하는 결과들을 달성하기 위해, 그러한 동작들이 도시된 특정한 순서 또는 순차적인 순서로 수행될 필요가 없거나, 모든 예시된 동작들이 수행될 필요가 없다는 것이 인지될 것이다. 추가로, 도면들은 흐름도의 형태로 둘 이상의 예시적인 프로세스들을 개략적으로 도시할 수 있다. 그러나, 도시되지 않은 다른 동작들이, 개략적으로 예시된 예시적인 방법들 및 프로세스들에 통합될 수 있다. 예컨대, 하나 이상의 추가적인 동작들은, 예시된 동작들 중 임의의 동작 이전에, 이후에, 동시성, 또는 그 중간에 수행될 수 있다. 부가적으로, 동작들은 다른 구현들에서 재배열되거나 재순서화될 수 있다. 소정의 환경들에서, 멀티태스킹 및 병렬 프로세싱이 유리할 수 있다. 또한, 위에서 설명된 구현에서의 다양한 시스템 컴포넌트들의 분리는 모든 구현들에서 그러한 분리를 요구하는 것으로서 이해되지는 않아야 하고, 그리고 설명된 프로그램 컴포넌트들 및 시스템들이 일반적으로, 단일 소프트웨어 제품으로 함께 통합되거나 다수의 소프트웨어 제품들로 패키징될 수 있다는 것이 이해되어야 한다. 부가적으로, 다른 구현들은 다음의 청구항들의 범위 내에 있다. 일부 경우들에서, 청구항들에서 열거된 액션들은, 상이한 순서로 수행될

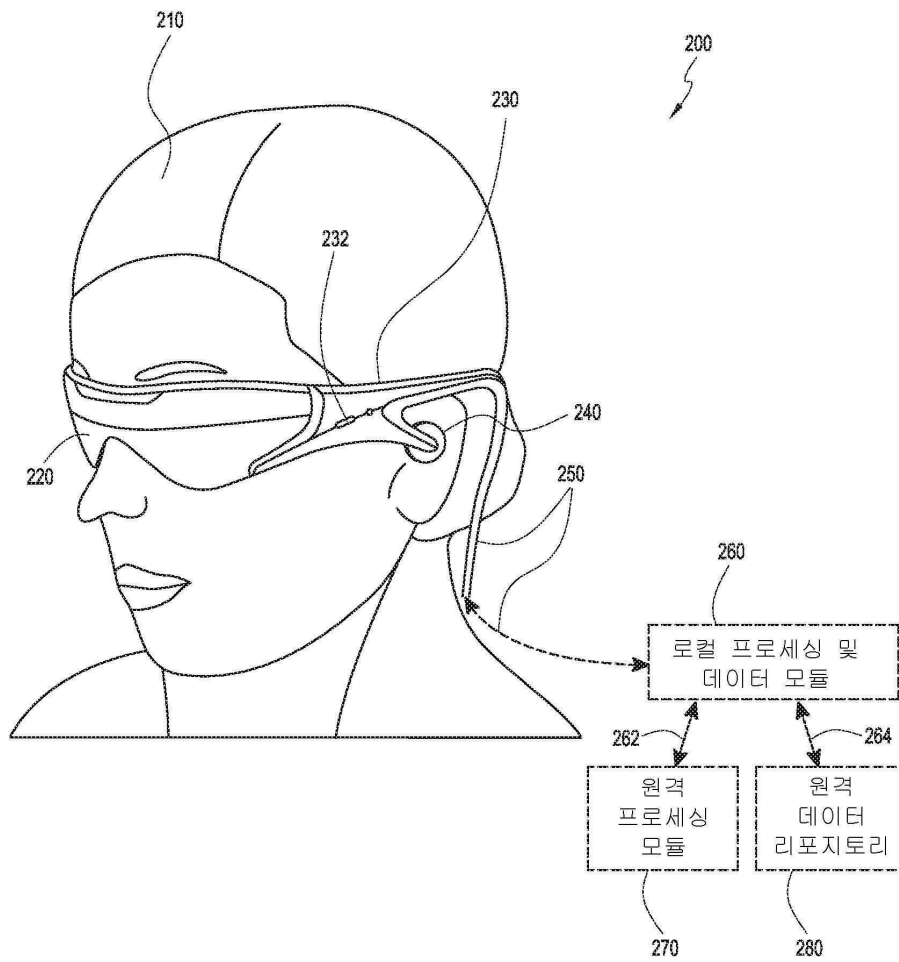
수 있으며, 그럼에도 불구하고 원하는 결과들을 달성할 수 있다.

도면

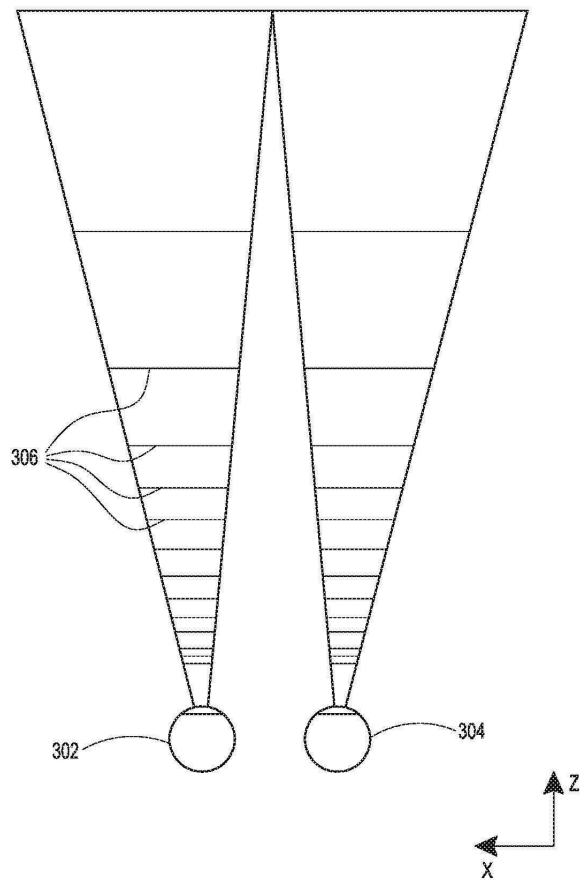
도면1



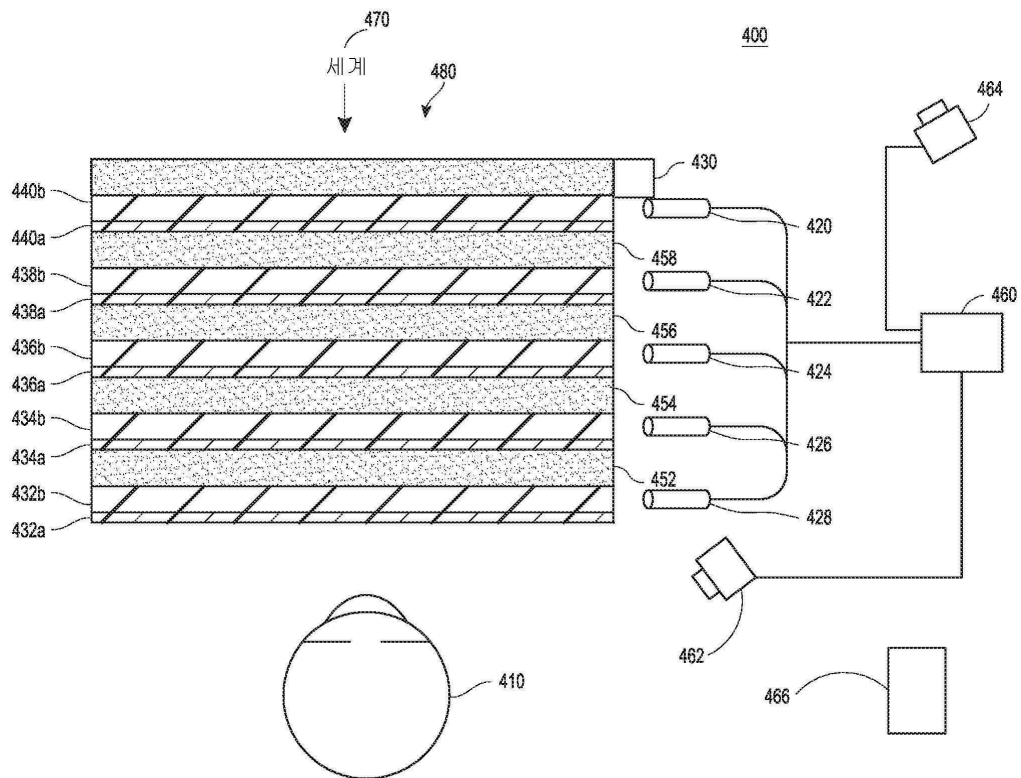
도면2



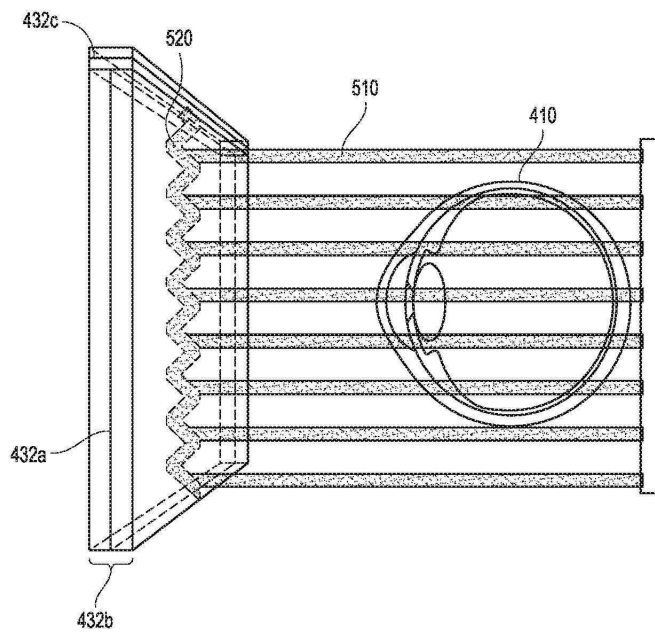
도면3



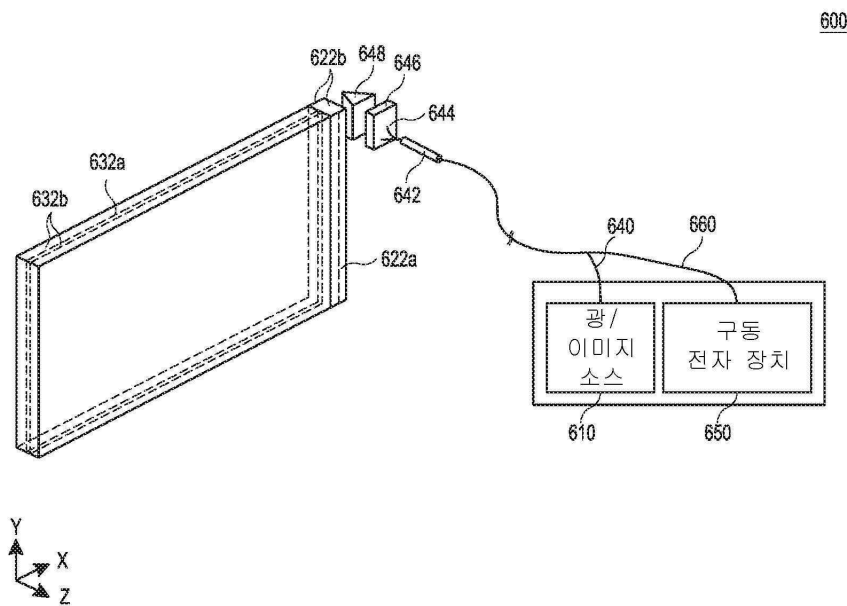
도면4



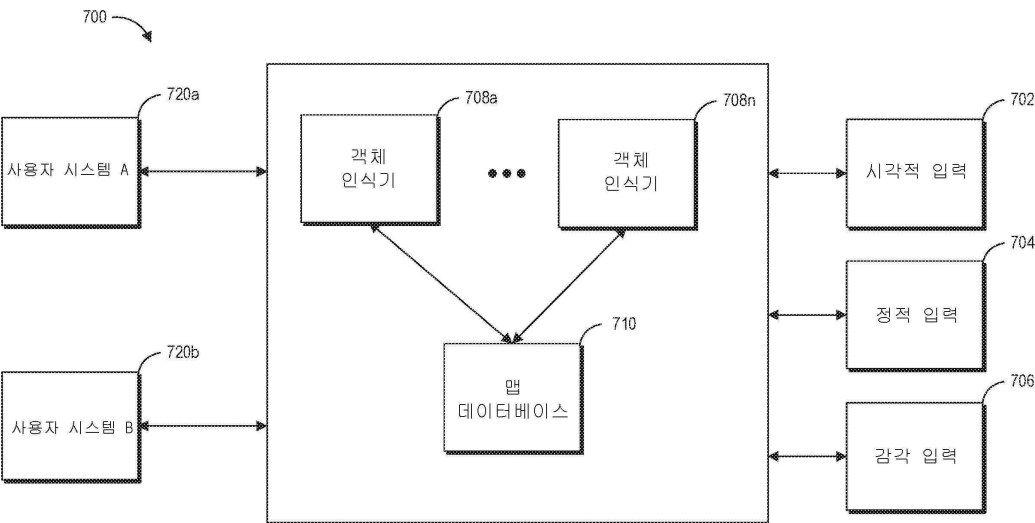
도면5



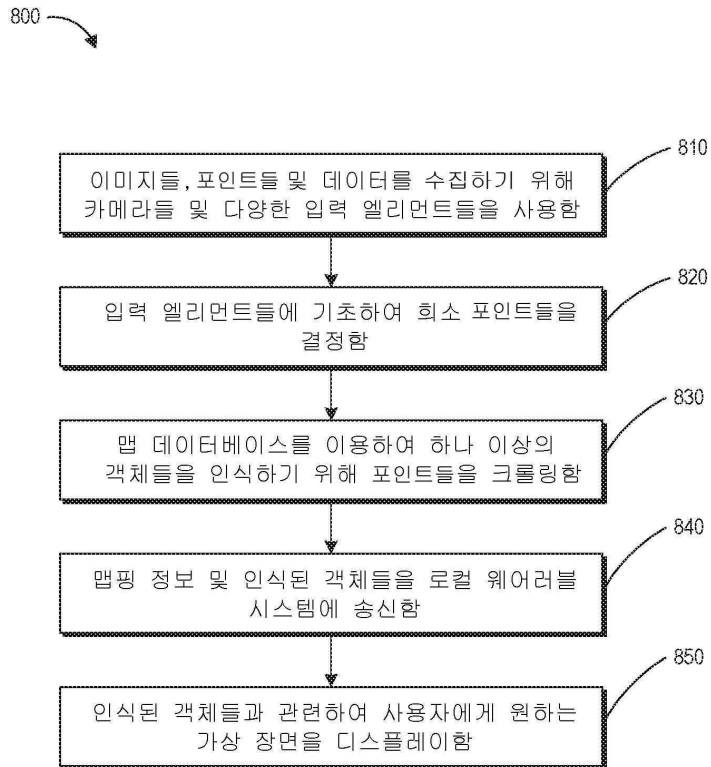
도면6



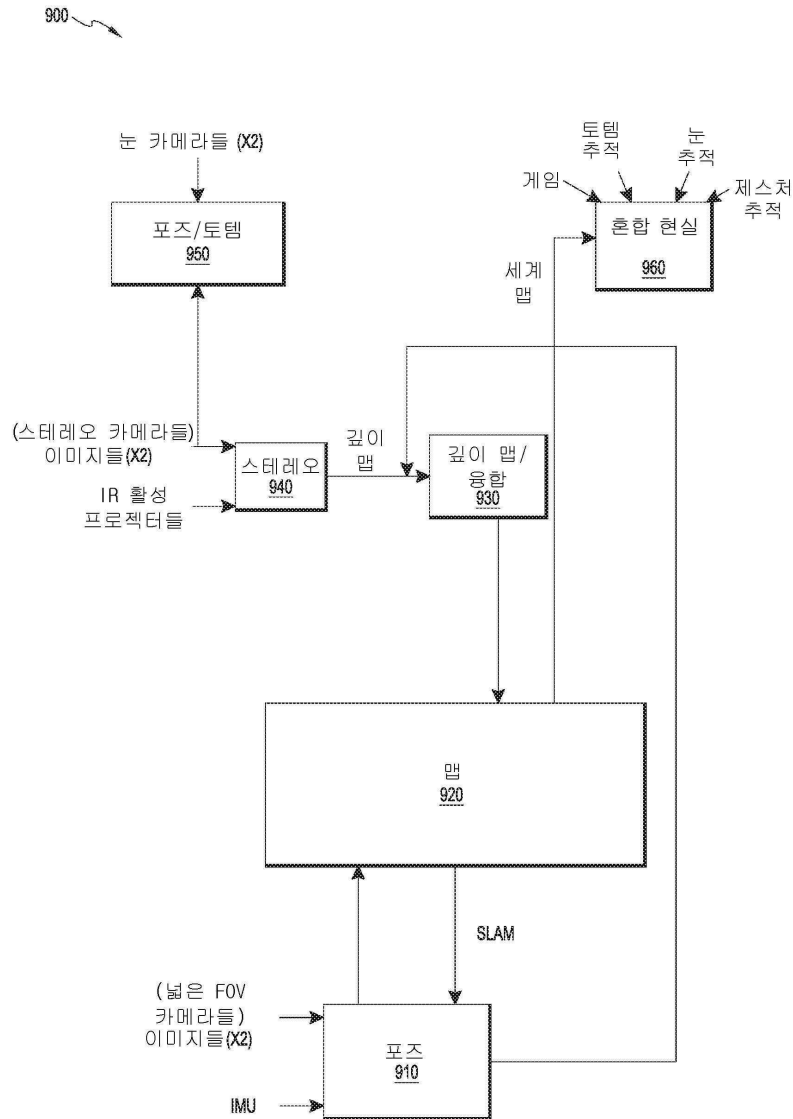
도면7



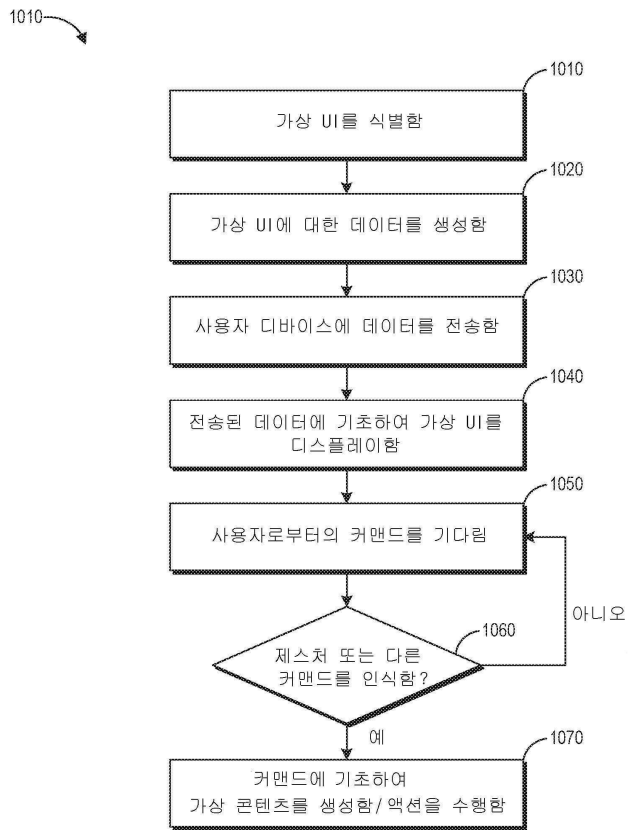
도면8



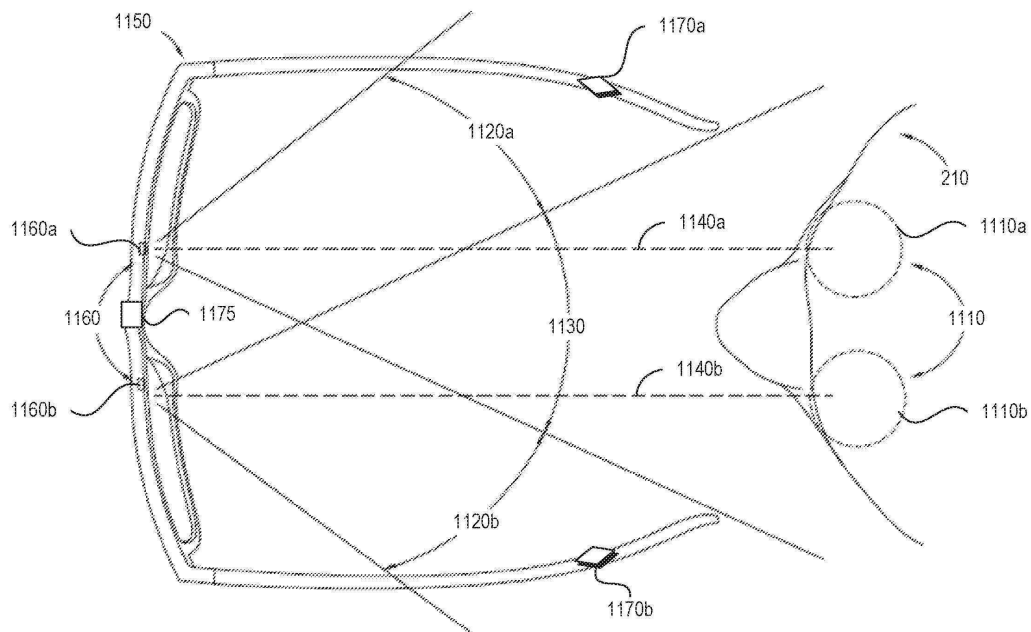
도면9



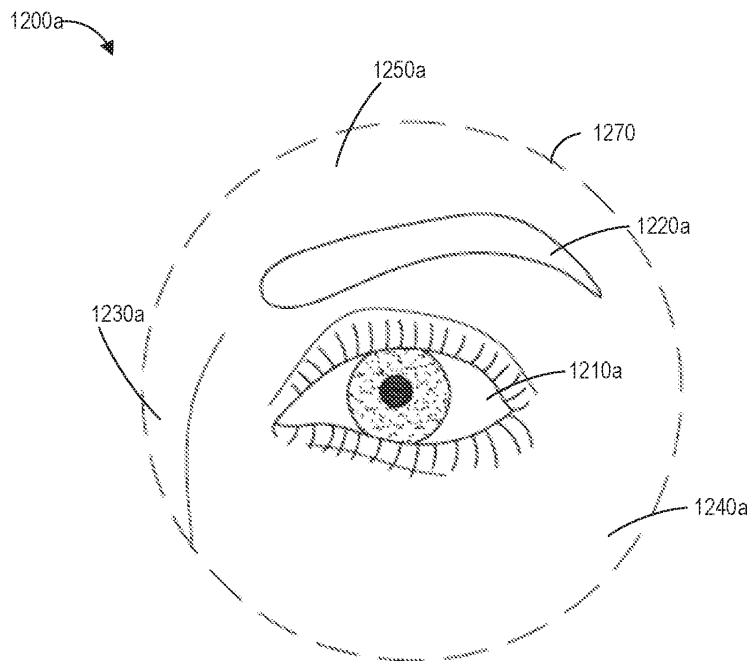
도면10



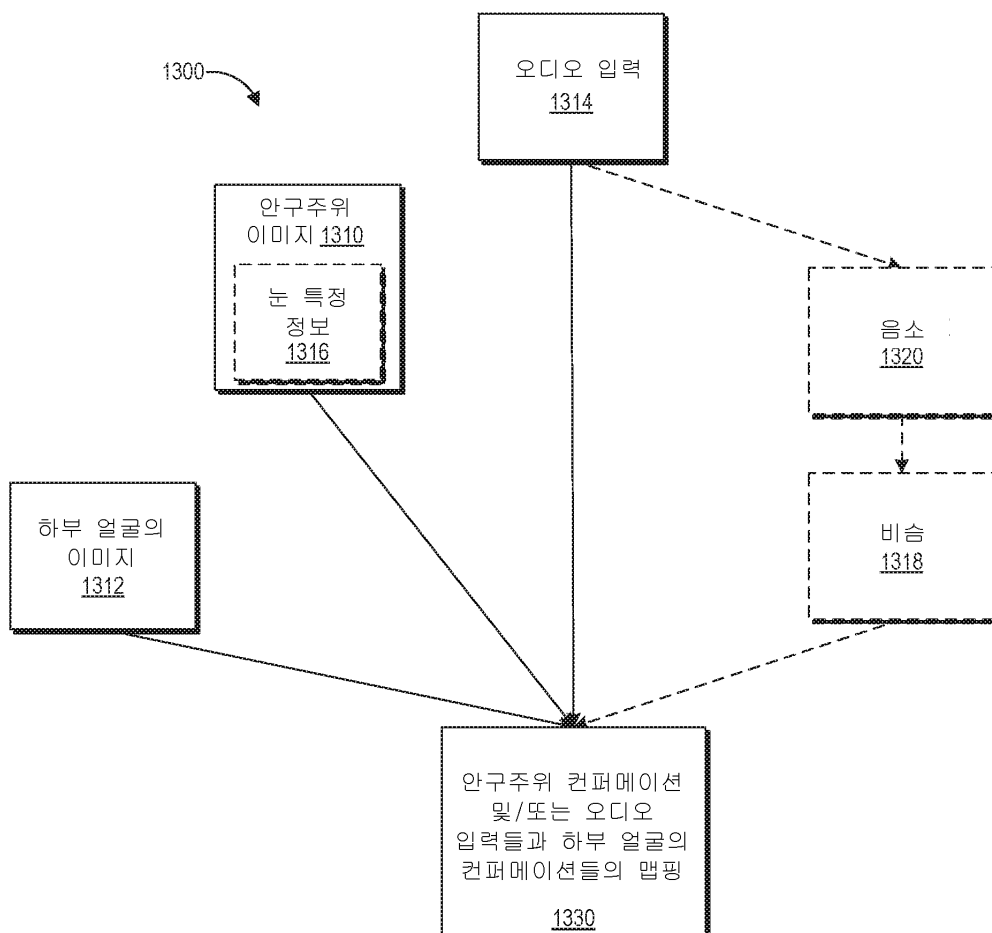
도면11



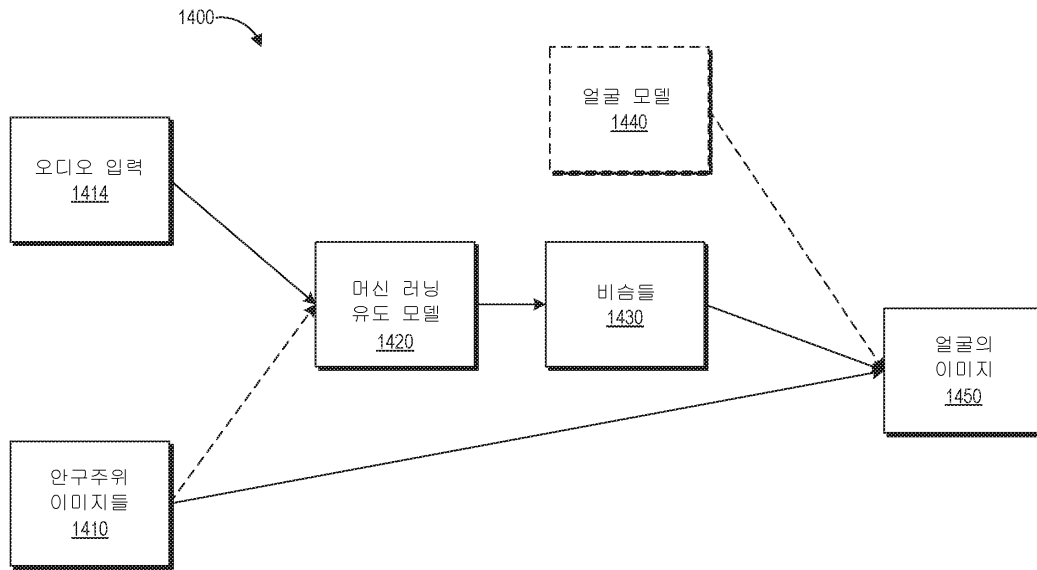
도면12



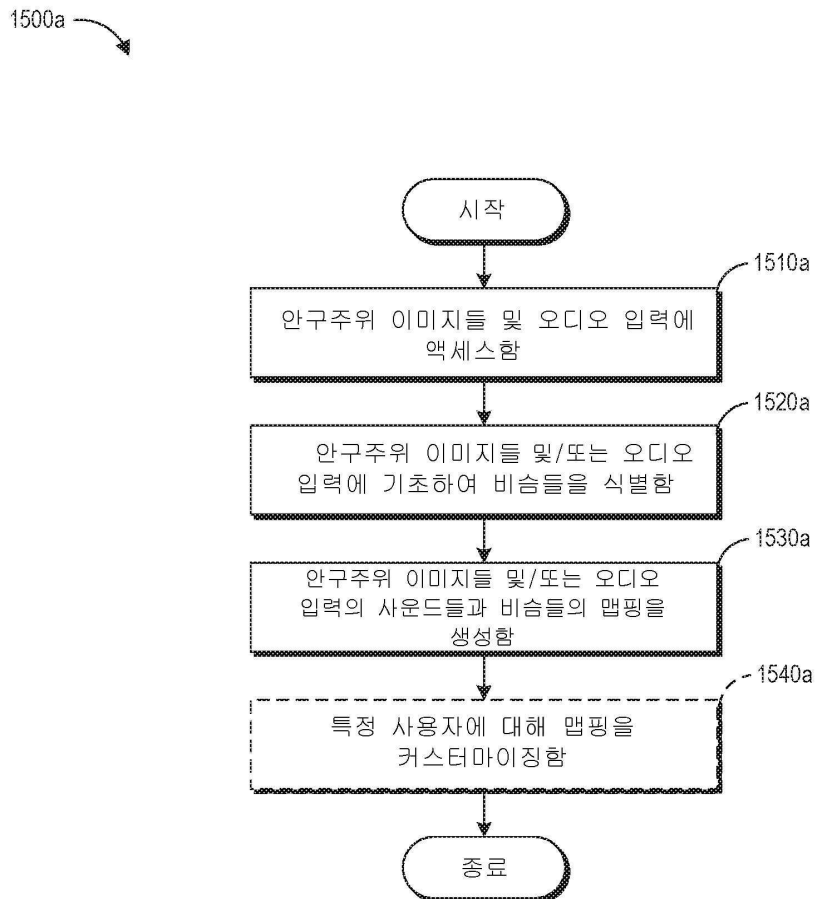
도면13



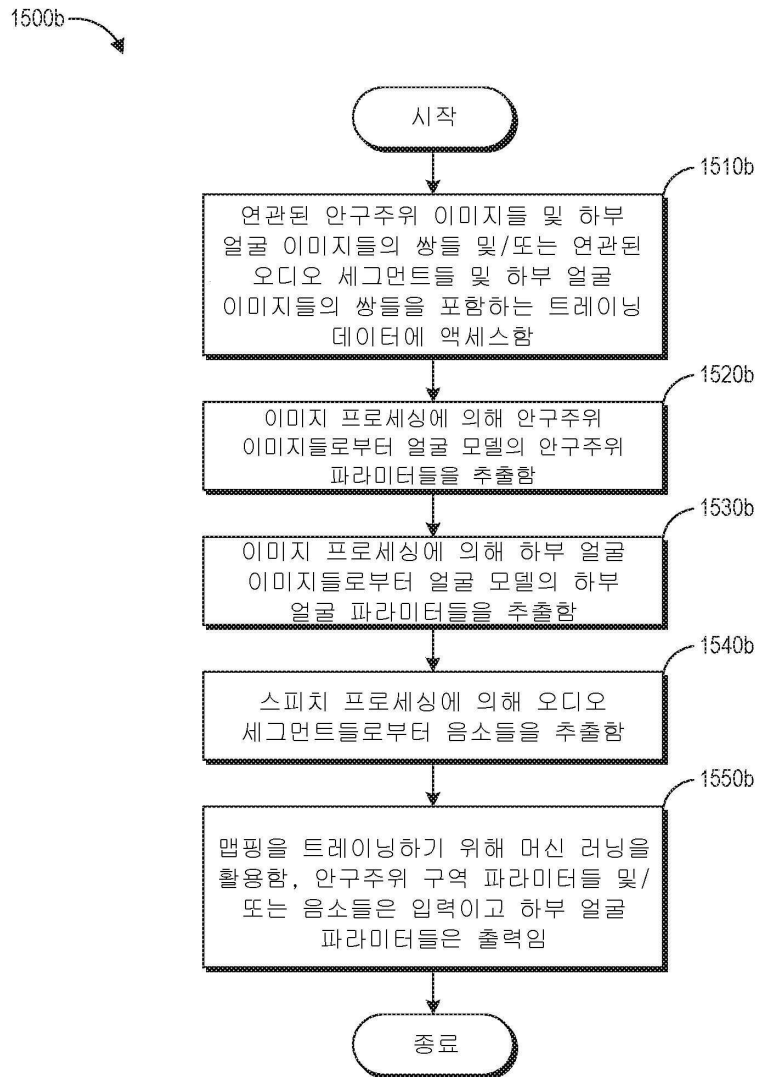
도면14



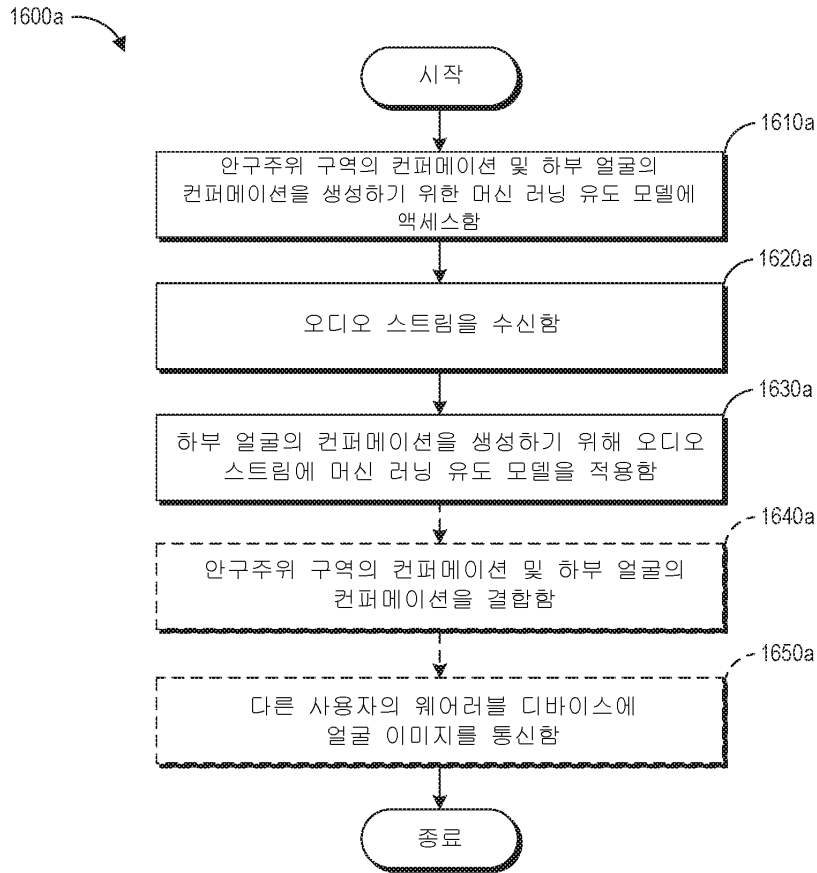
도면15a



도면 15b



도면 16a



도면 16b

1600b

