(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2023년09월06일
(11) 등록번호 10-2576135
(24) 등록일자 2023년09월04일

(51) 국제특허분류(Int. Cl.)
G06F 3/01 (2006.01) *G02B 27/01* (2006.01)
G06F 1/16 (2006.01)
(52) CPC특허분류
G06F 3/017 (2013.01)
G02B 27/0172 (2013.01)
(21) 출원번호 10-2022-7044995(분할)
(22) 출원일자(국제) 2017년09월12일
심사청구일자 2022년12월21일
(85) 번역문제출일자 2022년12월21일
(65) 공개번호 10-2023-0003667
(43) 공개일자 2023년01월06일
(62) 원출원 특허 10-2021-7015218
원출원일자(국제) 2017년09월12일
심사청구일자 2021년06월21일
(86) 국제출원번호 PCT/US2017/051173
(87) 국제공개번호 WO 2018/052901
국제공개일자 2018년03월22일
(30) 우선권주장
62/440,320 2016년12월29일 미국(US)
62/394,013 2016년09월13일 미국(US)
(56) 선행기술조사문헌
US20110301934 A1
US20140081634 A1

(73) 특허권자
매직 립, 인코포레이티드
미국 플로리다 플랜타티온 웨스트 선라이즈 블러바드 7500 (우: 33322)
(72) 발명자
브로위, 에릭
미국 33322 플로리다 플랜타티온 웨스트 선라이즈 블러바드 7500
우즈, 마이클, 자누스즈
미국 33322 플로리다 플랜타티온 웨스트 선라이즈 블러바드 7500
라비노비치, 앤드류
미국 33322 플로리다 플랜타티온 웨스트 선라이즈 블러바드 7500
(74) 대리인
특허법인 남앤남

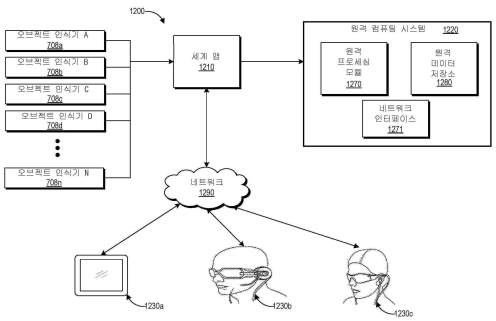
전체 청구항 수 : 총 20 항 심사관 : 이상헌

(54) 발명의 명칭 **감각 안경류**

(57) 요 약

혼합 현실 디바이스를 위한 감각 안경류 시스템은 다른 사람들과의 또는 환경과의 사용자의 상호작용들을 가능하게 할 수 있다. 일 예로서, 감각 안경류 시스템은 수화를 인식 및 해석하고, 변환된 정보를 혼합 현실 디바이스의 사용자에게 제공할 수 있다. 웨어러블 시스템은 또한 사용자의 환경에서 텍스트를 인식하고, (예컨대, 텍스트의 콘텐츠 또는 디스플레이 특징들을 변경함으로써) 텍스트를 수정하며, 오리지널 텍스트를 가리기 위해서 수정된 텍스트를 렌더링할 수 있다.

대 표 도



(52) CPC특허분류
G02B 27/0179 (2013.01)
G06F 1/163 (2013.01)
G06F 1/1686 (2013.01)
G06F 3/013 (2013.01)
G02B 2027/0138 (2013.01)
G02B 2027/0141 (2013.01)
G02B 2027/0178 (2013.01)
G02B 2027/0187 (2013.01)
G06T 2207/20084 (2013.01)

명 세 서

청구범위

청구항 1

수화(sign language) 인식을 위한 웨어러블 시스템으로서,

사용자에게 가상 콘텐트를 제시하도록 구성된 머리-장착 디스플레이;

상기 사용자의 환경으로부터 오디오를 검출하도록 구성된 오디오 센서;

상기 사용자의 환경을 이미징하도록 구성된 외향 이미징 시스템; 및

상기 머리-장착 디스플레이 및 상기 이미징 시스템과 통신하는 하드웨어 프로세서

를 포함하고,

상기 하드웨어 프로세서는:

상기 외향 이미징 시스템에 의해 캡처된 복수의 이미지들을 수신하고;

오브젝트 인식기로 상기 복수의 이미지들 내의 적어도 하나의 세트의 손들을 검출하고;

상기 적어도 하나의 세트의 손들의 상대적 크기를 결정하고;

상기 적어도 하나의 세트의 손들의 상기 상대적 크기에 기초하여 복수의 사람들 중 상기 적어도 하나의 세트의 손들의 소스를 식별하고;

상기 오브젝트 인식기로 상기 복수의 이미지들 내의 상기 적어도 하나의 세트의 손들에 의한 적어도 하나의 제스처를 검출하고;

상기 적어도 하나의 제스처의 소스가 상기 사용자 이외의 사람에게 속한다고 결정하고;

상기 결정에 대한 응답으로, 상기 적어도 하나의 제스처의 의미를 수화로 인식하고;

상기 사용자와 연관된 상황 정보(contextual information)에 기초하여 타겟 언어를 식별하고 - 상기 상황 정보는 상기 오디오 센서에 의해 획득된 오디오 데이터의 분석으로부터 획득된 상기 사용자의 스피치를 포함함 -;

상기 인식된 의미에 기초하여 그리고 상기 적어도 하나의 제스처의 소스가 상기 사용자 이외의 사람에게 속한다는 결정에 추가로 기초하여, 상기 적어도 하나의 제스처를 상기 타겟 언어로 트랜스레이션하고(translate);

상기 제스처를 상기 타겟 언어로 바꾸는 것에 적어도 부분적으로 기초하여 가상 콘텐트를 생성하고; 그리고

상기 머리-장착 디스플레이로 하여금 상기 사용자에게 상기 가상 콘텐트를 렌더링하도록

프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 2

제1항에 있어서,

상기 이미징 시스템은 상기 사용자의 주변을 이미징하도록 구성된 하나 이상의 광각 카메라들을 포함하는,

수화 인식을 위한 웨어러블 시스템.

청구항 3

제1항에 있어서,

상기 하드웨어 프로세서는, 상기 제스처와 연관된 보조 정보에 액세스하도록 추가로 프로그래밍되고,

상기 머리-장착 디스플레이에 의해 렌더링된 상기 가상 콘텐트는 상기 보조 정보를 포함하는,

수화 인식을 위한 웨어러블 시스템.

청구항 4

제1항에 있어서,

상기 사용자와 연관된 상황 정보에 기초하여 타겟 언어를 식별하기 위해, 상기 하드웨어 프로세서는:

상기 타겟 언어를 상기 사용자가 이해하는 언어로 설정하도록 프로그래밍되고, 상기 설정은 상기 웨어러블 시스템의 의해 캡쳐되는 상기 사용자의 스피치, 상기 사용자의 위치, 또는 상기 언어를 상기 타겟 언어로서 선택하는 상기 사용자로부터의 입력 중 적어도 하나에 기초하는,

수화 인식을 위한 웨어러블 시스템.

청구항 5

제1항에 있어서,

상기 하드웨어 프로세서는,

상기 타겟 언어가 발화된 언어인지를 결정하고; 그리고

상기 타겟 언어가 발화된 언어라고 결정하는 것에 대한 응답으로, 상기 트랜스레이션된 제스처와 연관된 스피치의 오디오 스트림을 상기 타겟 언어로 재생하도록

프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 6

제1항에 있어서,

상기 하드웨어 프로세서는,

상기 타겟 언어가 다른 수화인지를 결정하고; 그리고

상기 타겟 언어가 다른 수화라고 결정하는 것에 대한 응답으로, 상기 제스처의 트랜스레이션으로 상기 다른 수화로 다른 제스처의 그래픽을 제시하도록

프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 7

제1항에 있어서,

상기 제스처의 의미를 상기 수화로 인식하기 위해, 상기 하드웨어 프로세서는 상기 이미징 시스템에 의해 캡쳐되는 상기 이미지의 적어도 일부 상에 딥 뉴럴 네트워크 기술을 적용하도록 프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 8

제1항에 있어서,

상기 하드웨어 프로세서는, 상기 사용자의 위치에 적어도 부분적으로 기초하여 후보 수화의 리스트로부터 상기 수화를 식별하도록 추가로 프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 9

제1항에 있어서,

상기 인식된 의미에 기초하여 상기 제스처를 상기 타겟 언어로 트랜스레이션하기 위해, 상기 하드웨어 프로세서는 상기 제스처를 상기 타겟 언어의 텍스트 표현으로 변환(convert)하도록 프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 10

제1항에 있어서,

상기 하드웨어 프로세서는,

상기 검출된 제스처의 소스를 결정하고; 그리고

상기 검출된 제스처의 소스가 상기 웨어러블 시스템의 상기 사용자라고 결정하면, 상기 제스처의 상기 타겟 언어로의 트랜스레이션을 다른 사용자의 웨어러블 디바이스로 통신하도록

프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 11

제1항에 있어서,

상기 하드웨어 프로세서는,

상기 머리-장착 디스플레이에 의한 디스플레이로부터 상기 가상 콘텐트를 거부하기 위한 조건을 검출하고;

상기 조건의 검출에 응답하여 상기 머리-장착 디스플레이에 의한 상기 디스플레이로부터 상기 가상 콘텐트를 제거하도록

프로그래밍되는,

수화 인식을 위한 웨어러블 시스템.

청구항 12

제11항에 있어서,

상기 조건은, 시간 기간, 사용자의 손 제스처, 또는 사용자 입력 디바이스로부터의 입력 중 적어도 하나를 포함하는,

수화 인식을 위한 웨어러블 시스템.

청구항 13

제1항에 있어서,

상기 이미지는 비디오의 하나 이상의 프레임들을 포함하는,

수화 인식을 위한 웨어러블 시스템.

청구항 14

수화 인식을 위한 방법으로서,

웨어러블 시스템의 사용자에게 가상 콘텐트를 제시하도록 구성된 머리-장착 디스플레이를 포함하는 상기 웨어러블 시스템의 제어 하에서,

상기 웨어러블 시스템의 이미징 시스템에 의해 캡처된 이미지를 수신하는 단계;

사람의 제스처를 검출하기 위해 상기 이미지를 분석하는 단계;

상기 사람의 제스처와 연관된 일 세트의 손들의 상대적 크기를 결정하는 단계;

상기 상대적 크기에 기초하여 상기 사람의 제스처의 소스를 식별하는 단계;

상기 검출된 제스처에 적어도 부분적으로 기초하여 수화로된 의사소통의 존재를 검출하는 단계;

상기 사람의 제스처가 상기 사용자의 제스처가 아닌 경우, 상기 제스처의 의미를 상기 수화로 인식하는 단계;

상기 제스처가 트랜스레이션될 타겟 언어를 식별하는 단계 - 상기 타겟 언어를 식별하는 단계는, 상기 사용자의 스피치를 검출하기 위해 상기 웨어러블 시스템의 의해 획득된 오디오 데이터를 분석하는 단계를 포함함 -;

상기 인식된 의미에 기초하여 그리고 상기 사람의 제스처의 소스가 상기 사용자 이외의 사람에게 속한다는 결정에 추가로 기초하여, 상기 제스처를 상기 타겟 언어로 트랜스레이션하는 단계;

상기 제스처의 상기 타겟 언어로의 트랜스레이션에 적어도 부분적으로 기초하여 가상 콘텐트를 생성하는 단계; 및

상기 머리-장착 디스플레이가 상기 사용자에게 상기 가상 콘텐트를 렌더링하게 하는 단계

를 포함하는,

수화 인식을 위한 방법,

청구항 15

제14항에 있어서,

상기 웨어러블 시스템은, 제2 웨어러블 시스템에 의한 렌더링을 위해 제2 가상 콘텐트를 상기 제2 웨어러블 시스템으로 통신하도록 추가로 구성되는,

수화 인식을 위한 방법.

청구항 16

제14항에 있어서,

상기 인식된 의미에 기초하여 상기 제스처를 상기 타겟 언어로 트랜스레이션하는 단계는, 상기 제스처를 상기 타겟 언어의 텍스트 표현으로 변환하는 단계를 포함하는,

수화 인식을 위한 방법.

청구항 17

제14항에 있어서,

상기 가상 콘텐트는, 상기 타겟 언어로 된 텍스트 표현 또는 상기 타겟 언어로 다른 것을 설명하는 그래픽을 포함하는,

수화 인식을 위한 방법.

청구항 18

제14항에 있어서,

상기 제스처의 의미를 상기 수화로 인식하는 단계는, 상기 이미징 시스템에 의해 캡쳐된 상기 이미지의 적어도 일부 상에 딥 뉴럴 네트워크 기술을 적용하는 단계를 포함하는,

수화 인식을 위한 방법.

청구항 19

제14항에 있어서,

상기 수화로된 의사 소통의 존재를 검출하는 단계는:

후보 수화들의 리스트로부터 상기 수화를 식별하는 단계; 및

상기 검출된 제스처가 상기 수화로된 표현에 대응한다고 결정하는 단계

를 포함하는,

수화 인식을 위한 방법.

청구항 20

제19항에 있어서,

상기 검출된 제스처가 상기 수화로된 표현에 대응한다고 결정하는 단계는, 상기 제스처를 하고 있는 상기 사람의 입술 움직임들 및 상기 사람이 상기 제스처를 하는 동안 캡처된 오디오 데이터와 관련하여 상기 제스처를 분석하는 단계를 포함하는,

수화 인식을 위한 방법.

발명의 설명

기 술 분 야

[0001] [0001] 본 출원은, 2016년 9월 13일에 출원되고 발명의 명칭이 "SENSORY EYEWEAR"인 미국 가출원 번호 제62/394,013호, 및 2016년 12월 29일에 출원되고 발명의 명칭이 "SYSTEMS AND METHODS FOR AUGMENTED REALITY"인 미국 가출원 번호 제62/440,320호를 35 U.S.C. § 119(e) 하에서 우선권의 이익을 주장하며, 따라서 이 출원들의 개시내용들은 그 전체가 인용에 의해 본원에 포함된다.

[0002] [0002] 본 개시내용은 가상 현실 및 증강 현실 이미징 및 시각화 시스템들에 관한 것으로, 보다 구체적으로는, 환경에서 수화 또는 텍스트를 인식하고 인식된 수화 또는 텍스트에 기반하여 가상 콘텐츠를 렌더링하는 것에 관한 것이다

배 경 기 술

[0003] [0003] 현대 컴퓨팅 및 디스플레이 기술들은 소위 "가상 현실", "증강 현실", 또는 "혼합 현실" 경험들을 위한 시스템들의 개발을 가능하게 했으며, 여기서 디지털적으로 재생된 이미지들 또는 이미지들의 부분들은, 그들이 실제인 것으로 보이거나, 실제로서 지각될 수 있는 방식으로 사용자에게 제공된다. 가상 현실, 또는 "VR" 시나리오는 전형적으로 다른 실제 실세계 시각적 입력에 대한 투명도(transparency) 없이 디지털 또는 가상 이미지 정보의 프리젠테이션(presentation)을 수반하고; 증강 현실, 또는 "AR" 시나리오는 전형적으로 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 오브젝트들이 공존하고 실시간으로 상호작용하는 새로운 환경을 생성하는 것과 관련된다. 밝혀진 바와 같이, 인간 시각 지각 시스템은 매우 복잡하고, 다른 가상 또는 실세계 이미저리 엘리먼트들 사이에서 가상 이미지 엘리먼트들의 편안하고, 자연스러운 느낌이고, 풍부한 프리젠테이션을 가능하게 하는 VR, AR 또는 MR 기술을 생성하는 것은 난제시된다. 본원에 개시된 시스템들 및 방법들은 VR, AR 및 MR 기술에 관련된 다양한 난제들을 처리한다.

발명의 내용

[0004] [0004] 환경에서 수화 및 텍스트를 인식하기 위한 혼합 현실 시스템의 다양한 실시예들이 개시된다. 이들 실시예들은 다른 능력의 사람(differently-abled person)들 사이에서 떠 뛰어난 상호작용을 유리하게 허가할 수 있다.

[0005] [0005] 혼합 현실 디바이스를 위한 감각 안경류 시스템은 다른 사람들 또는 환경과의 사용자의 상호작용을 가능하게 할 수 있다. 일 예로서, 감각 안경류 시스템은 수화를 인식 및 해석하고, 변환된 정보를 혼합 현실 디바이스의 사용자에게 제공할 수 있다. 웨어러블 시스템은 또한 사용자의 환경 내의 텍스트를 인식하고, (예컨대, 텍스트의 콘텐츠 또는 디스플레이 특징들을 변하게 함으로써) 텍스트를 수정하고, 오리지널 텍스트를 가리도록 수정된 텍스트를 렌더링할 수 있다.

[0006] [0006] 본 명세서에서 설명되는 청구대상의 하나 이상의 구현들의 세부사항들은, 아래의 첨부 도면들 및 설명에서 기술된다. 다른 특징들, 양상들, 및 장점들은 설명, 도면들, 및 청구항들로부터 자명해질 것이다. 이 개요 또는 다음의 상세한 설명 어느 것도, 본 발명의 청구대상의 범위를 정의하거나 제한하는 것으로 의도되지 않는다.

도면의 간단한 설명

[0007] [0007] 도 1은 사람이 보는 특정 물리적 오브젝트들 및 특정 가상 현실 오브젝트들을 갖는 혼합 현실 시나리오의 예시를 묘사한다.

[0008] 도 2a는 감각 안경류 시스템을 구현할 수 있는 웨어러블 시스템의 예를 개략적으로 예시한다.

[0009] 도 2b는 웨어러블 시스템의 다양한 예시적인 컴포넌트들을 개략적으로 예시한다.

[0010] 도 3은 다중 깊이 평면들을 사용하여 3-차원 이미저리를 시뮬레이팅하기 위한 접근법의 양상들을 개략적으로 예시한다.

[0011] 도 4는 사용자에게 이미지 정보를 출력하기 위한 도파관 스택의 예를 개략적으로 예시한다.

[0012] 도 5는 도파관에 의해 출력될 수 있는 예시적인 출사 빔들을 도시한다.

[0013] 도 6은, 다-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.

[0014] 도 7은 웨어러블 시스템의 일 예의 블록 다이어그램이다.

[0015] 도 8은 인지된 오브젝트들에 관하여 가상 콘텐츠를 렌더링하는 방법의 예의 프로세스 흐름도이다.

[0016] 도 9는 감각 안경류 시스템을 포함하는 웨어러블 시스템의 다른 예의 블록 다이어그램이다.

[0017] 도 10은 웨어러블 시스템에 대한 사용자 입력을 결정하는 방법의 예의 프로세스 흐름도이다.

[0018] 도 11은 가상 사용자 인터페이스와 상호작용하기 위한 방법의 예의 프로세스 흐름도이다.

[0019] 도 12는 서로 상호작용하는 다수의 웨어러블 시스템들을 묘사하는 전체 시스템도를 개략적으로 예시한다.

[0020] 도 13a는 감각 안경류 시스템이 (예컨대, 수화 사용자에 의해 제스처링되는) 수화를 해석할 수 있는 감각 안경류 시스템의 예시적인 사용자 경험을 도시한다.

[0021] 도 13b는 타겟 스피치 및 보조 정보 둘 모두가 제공되는 감각 안경류 시스템의 다른 예시적인 사용자 경험을 도시한다.

[0022] 도 13c는 텔레프레전스 세션에서의 감각 안경류 시스템의 예시적인 사용자 경험을 도시한다.

[0023] 도 13d는 수화를 해석하기 위한 예시적인 가상 사용자 인터페이스를 예시한다.

[0024] 도 14a 및 14b는 감각 안경류 시스템으로 개인간 통신들을 가능하게 하기 위한 예시적인 프로세스들을 예시한다.

[0025] 도 14c는 보조 정보를 결정하고 변환된 텍스트와 연관된 보조 정보를 제공하기 위한 예시적인 방법의 프로세스 흐름도이다.

[0026] 도 15는 감각 안경류 시스템으로 개인간 통신을 가능하게 하기 위한 다른 예시적인 프로세스를 예시한다.

[0027] 도 16a-16e는, 환경에서 텍스트를 인식하고, 텍스트의 디스플레이 특징들을 수정하고, 수정된 텍스트를 렌더링하도록 구성된 감각 안경류 시스템에 대한 예시적인 사용자 경험들을 예시한다.

[0028] 도 17은 환경과의 사용자의 상호작용들을 가능하게 하기 위한 감각 안경류의 예시적인 프로세스를 예시한다.

[0029] 도 18은 사이니지의 콘텐츠를 수정함으로써 물리적 환경에서 사용자가 사이니지를 이해하는 것을 보조

하는 예를 예시한다.

[0030] 도 19는 물리적 환경에서 사용자가 사이니지를 이해하는 것을 보조하는 예시적인 프로세스를 예시한다.

[0031] 도면들 전반에 걸쳐, 참조 번호들은 참조된 엘리먼트들 간의 대응성(correspondence)을 표시하는 데 재사용될 수 있다. 도면들은 본원에서 설명된 예시적인 실시예들을 예시하기 위해 제공되며 본 개시내용의 범위를 제한하는 것으로 의도되지 않는다.

발명을 실시하기 위한 구체적인 내용

[0008] 개요

[0009] [0032] AR/VR/MR 콘텐츠를 제공하도록 구성된 웨어러블 시스템은 다른 사람들 또는 환경과의 사용자의 상호작용을 향상시키기 위해 감각 안경류 시스템을 구현할 수 있다. 예시적인 웨어러블 시스템은 머리-장착 디스플레이, 다양한 이미징 센서들 및 하나 이상의 하드웨어 프로세서들을 포함할 수 있다. 디스플레이는 눈 또는 눈들 전방에 착용된 시스루 디스플레이일 수 있다.

[0010] [0033] 다른 사람들과의 사용자의 상호작용 경험을 향상시키기 위해, 웨어러블 시스템은 수화를 캡처 및 해석하도록 구성할 수 있다. 수화는 주로 시각 제스처(예컨대, 손 형상들, 손 배향들, 손, 팔 또는 신체 움직임들, 또는 얼굴 표정들)를 사용하여 통신한다. 전세계에서 사용되는 수백 개의 수화들이 존재한다. 일부 수화들은 다른 것들보다 더 자주 사용될 수 있다. 예컨대, ASL(American sign language)는 미국 및 캐나다에서 널리 사용된다.

[0011] [0034] 다수의 사람들이 어떠한 수화들도 알지 못한다. 언어- 또는 청각-장애 사람들 및 대화 파트너는 동일한 수화에 익숙하지 않을 수 있다. 이는 청각-장애 또는 언어-장애 사람들과의 대화를 방해할 수 있다. 그에 따라서, 대화 파트너가 행한 싸인들(예컨대, 제스처들)을 이미지화하고 싸인들을 텍스트 또는 그래픽 (예컨대, 시스템 사용자가 이해하는 수화의 수화 제스처들의 그래픽)으로 변환하고, 그런 다음, 싸인과 연관된 정보(예컨대, 사용자에 의해 이해되는 언어로의 싸인들의 변환)를 디스플레이할 수 있는 웨어러블 시스템은 사용자와 대화 파트너 간의 통신을 향상시키는 데 크게 도움이 된다. 추가로, 웨어러블 시스템의 사용자에게 최소 레벨의 주의산만함 및 그 사용자에 의한 사소한 레벨의 노력으로 실시간(또는 거의 실시간)으로 수화의 텍스츄얼 또는 그래픽 변환을 제공할 수 있는 웨어러블 시스템을 갖는 것이 바람직할 수 있다.

[0012] [0035] 본 개시내용은 웨어러블 디바이스의 맥락에서 그러한 원하는 시스템들의 예들을 개시한다. 웨어러블 디바이스는 머리-장착 컴포넌트(이를테면, 예컨대, 머리-장착 디스플레이)를 포함할 수 있다. 이러한 디바이스는, 정보가 일반적으로 볼 수 있는 가능한 실제 세계와 나란히 (또는 그 위에) 동시적으로 볼 수 있는 가능한 방식으로, 컴퓨팅 디바이스에 의해 제공되는 정보를 사용자가 시각적으로 수신하게 허용할 수 있다. 그러한 시스템은 통상의 컴퓨터 스크린 상에 디스플레이될 수 있는 임의의 형태의 정보, 이를테면, 캐릭터들, 이미지 효과들, 텍스트, 그래픽들 또는 임의의 종류의 비디오를 디스플레이하는 데 사용될 수 있다.

[0013] [0036] 본원에서 설명된 웨어러블 시스템은 SLR(sign language recognition) 및 웨어러블 디바이스의 디스플레이 능력을 조합하여 검출된 수화에 기반하여 정보를 사용자에게 제공할 수 있다. 예컨대, 웨어러블 디바이스 상의 외향 카메라는 행해진 제스처들을 이미지화하고, 제스처들 사이에서 싸인들을 식별하고, 사용자가 이해하는 언어로 싸인들을 변환하고, 변환을 사용자에게 디스플레이할 수 있다. 검출된 수화의 트랜스크립트(예컨대, 자막 또는 텍스트 버블)는 웨어러블 시스템에 의해 사용자에게 디스플레이될 수 있다. 기계 학습 알고리즘(예컨대, 딥 뉴럴 네트워크)은 이미지들을 수신하고 싸인들의 식별 및 변환을 수행할 수 있다. 사용자에 의해 촉구될 때, 트랜스크립트 내의 단어의 의미 또는 적합한 소스로부터의 관련 정보가 디스플레이될 수 있다. 웨어러블 시스템이 제공할 수 있는 보조 정보의 종류들은 예컨대, 인터넷 상에서 이용가능한 정보 자원들의 거대한 어레이만큼 무제한적일 수 있다.

[0014] [0037] 다른 사람들과의 사용자의 상호작용 경험을 향상시키는 것 이외에도 또는 이에 대한 대안으로서, 감각 안경류 시스템은 또한 환경과의 사용자의 경험을 개선할 수 있다. 환경과의 사용자의 상호작용을 개선하는 예로서, 감각 안경류 시스템을 구현하는 웨어러블 시스템은 환경에서 텍스트 (예컨대, 사이니지 상의 텍스트, 이를테면, 예컨대, 상용 또는 공용 디스플레이 싸인들)를 인식하고, (예컨대, 텍스트의 사이즈를 증가시킴으로써) 텍스트의 디스플레이 특징들을 수정하거나 (예컨대, 텍스트를 다른 언어로 변환함으로써) 텍스트의 콘텐츠를 수정하고, 환경의 물리적 텍스트 위에 수정된 텍스트를 렌더링할 수 있다.

[0015] [0038] 본원에서 추가로 설명된 바와 같이, 웨어러블 시스템은 사용자 환경의 이미지를 수신할 수 있다. 이미

지는 웨어러블 디바이스의 외향 이미징 시스템 또는 웨어러블 디바이스와 연관된 토템에 의해 획득될 수 있다. 웨어러블 시스템은 이미지가 하나 이상의 글자들 또는 캐릭터들을 포함하는지 여부를 결정하고 하나 이상의 글자들 또는 캐릭터들을 텍스트로 변환할 수 있다. 웨어러블 시스템은 예컨대, 기계 학습 알고리즘들 또는 OCR(optical character recognition) 알고리즘들과 같은 다양한 기법들을 사용하여 이미지가 글자들 또는 캐릭터들을 포함하는지를 결정할 수 있다. 웨어러블 시스템은 글자들 및 캐릭터들을 식별하고 이를 텍스트로 변환하기 위해 (예컨대, 도 7에서 설명된) 오브젝트 인식기들을 사용할 수 있다.

[0016] [0039] 특정 실시예들에서, 텍스트는, 웨어러블 디바이스 없이 사용자가 보았을 것과 다르게 사용자에게 디스플레이될 수 있다. 예컨대, 웨어러블 시스템은 머리-장착 디스플레이로 하여금, 오리지널 이미지와 연관된 글자들 또는 캐릭터들과 관련된 폰트 사이즈와 상이한 폰트 사이즈로 텍스트를 디스플레이하게 할 수 있다. 웨어러블 시스템은 또한 텍스트의 디스플레이 품질을 개선할 수 있다. 예컨대, 다양한 환경 팩터들, 이를테면, 안개, 연무, 비, 밝은 조명, 낮은 조명, 글자들 간의 낮은 조명 또는 컬러 콘트라스트 및 주변 이미지 등은 웨어러블 시스템이 없이는 환경에서 텍스트를 사용자가 명확히 보는 것을 방해할 수 있다. 웨어러블 시스템은 텍스트의 명확성을 증가시킬 싸인(예컨대, 증가된 콘트라스트 비율 또는 보다 큰 폰트)을 제공할 수 있다.

[0017] [0040] 웨어러블 시스템은 또한 텍스트 (예컨대, 사이니지 상의 텍스트)를 그의 오리지널 언어로부터 타겟 언어로 변환할 수 있다. 예컨대, 텍스트는 사용자가 이해하지 못하는 언어로부터 사용자가 이해하는 언어로 변환될 수 있다. 변환된 텍스트는 사용자가 이해할 수 있는 언어의 텍스트를 사용자가 쉽게 볼 수 있도록 오리지널 텍스트 위에 렌더링될 수 있다.

[0018] 웨어러블 시스템의 3D 디스플레이의 예들

[0019] [0041] 웨어러블 시스템(본원에서 AR(augmented reality) 시스템으로 또한 지칭됨)은 2D 또는 3D 가상 이미지들을 사용자에게 제공하도록 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 제공할 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.

[0020] [0042] 도 1은 사람이 보는 특정 물리적 오브젝트들 및 특정 가상 현실 오브젝트들을 갖는 혼합 현실 시나리오의 예시를 묘사한다. 도 1에서, MR 장면(100)이 묘사되며, 여기서 MR 기술의 사용자는 배경에 있는 사람들, 나무들, 빌딩들, 및 콘크리트 플랫폼(120)을 특징으로 하는 실세계 공원-형 장소(110)를 본다. 이들 아이템들 외에도, MR 기술의 사용자는 또한, 그가 실세계 플랫폼(120) 상에 서 있는 로봇 동상(130), 및 호박벌의 의인화인 것으로 보여지는 날고 있는 만화형 아바타 캐릭터(140)를 보는 것을 지각하더라도, 이들 엘리먼트들은 실세계에 존재하지 않는다.

[0021] [0043] 3D 디스플레이가 올바른 깊이감(sensation of depth) 및 보다 구체적으로, 시뮬레이팅된 표면 깊이감을 생성하기 위해, 디스플레이의 시계(visual field)의 각각의 포인트가 그의 가상 깊이에 대응하는 원근조절 응답을 생성하는 것이 바람직할 수 있다. 만약 디스플레이 포인트에 대한 원근조절 응답이 수렴 및 입체시(stereopsis)의 양안 깊이 큐들에 의해 결정된 바와 같은 그 포인트의 가상 깊이에 대응하지 않는다면, 인간의 눈은 원근조절 충돌을 경험할 수 있어, 불안정한 이미징, 유해한 눈의 피로, 두통들, 그리고 원근조절 정보의 부재 시에, 표면 깊이의 거의 완벽한 결여를 야기할 수 있다.

[0022] [0044] 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 오브젝트의 약간 상이한 프리젠테이션들을 제공함) 뷰어의 눈들에 의해 별개로 초점이 맞춰질 수 있어서, 상이한 깊이 평면 상에 위치되는 장면에 대한 상이한 이미지 특징들에 초점을 맞추도록 요구되는 눈의 원근조절에 기반하여 또는 상이한 깊이 평면들 상의 상이한 이미지 특징들이 초점에서 벗어나는 것을 관찰하는 것에 기반하여 깊이 큐들을 사용자에게 제공하는 것을 돕는다. 본원의 다른 곳에서 논의된 바와 같이, 이러한 깊이 큐들은 깊이의 신뢰할 수 있는 지각들을 제공한다.

[0023] [0045] 도 2a는 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 기계 및 전자 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 제공할 수 있다. 디스플레이(220)는 사용자의 머리 상에 착용되는 머리 장착 디스플레이를 포함할 수 있다. 일부 실시예들에서, 스피커(240)는 프레임(230)에 커플링되고 사용자의 외이도에 인접하게 포지셔닝된다(일부 실시예들에서, 도시되지 않은 다른 스피커가 사용자의 다른 외이도에 인접하게 포지셔닝되어 입체/성형가능 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 입체 사운드 수신을 제공하도록 포지셔닝된다. 입체 사운드 수신은 사운드 소스의 위치를 결정하는 데 사용될 수 있다. 웨어러블 시스템(200)은 오디오 스트림 상에서 음성 또는 스피치 인식을 수행할 수 있다.

[0024] [0046] 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 눈들의 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.

[0025] [0047] 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.

[0026] [0048] 디스플레이(220)는 이를테면, 유선 리드 또는 무선 연결성에 의해, 다양한 구성들로 장착될 수 있는, 이를테면, 프레임(230)에 고정적으로 부착되거나, 사용자에 의해 착용된 헬멧 또는 모자에 고정적으로 부착되거나, 헤드폰들에 내장되거나, 그렇지 않으면 사용자(210)에게 제거가능하게 부착되는 (예컨대, 백팩(backpack)-스타일 구성으로, 벨트-커플링 스타일 구성으로) 로컬 데이터 프로세싱 모듈(260)에 동작가능하게 커플링(250)될 수 있다.

[0027] [0049] 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 동작가능하게 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 저장소(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 저장소(280)에 동작가능하게 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 저장소(280)는 서로 동작가능하게 커플링될 수 있다.

[0028] [0050] 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 저장소(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.

[0029] [0051] 도 2b는, 디스플레이(220) 및 프레임(230)을 포함할 수 있는 웨어러블 시스템(200)을 도시한다. 확대된 도면(202)은 웨어러블 시스템(200)의 다양한 컴포넌트들을 개략적으로 예시한다. 특정 구현들에서, 도 2b에 예시된 컴포넌트들 중 하나 이상은 디스플레이(220)의 부분일 수 있다. 다양한 컴포넌트들은, 단독으로 또는 조합하여, 웨어러블 시스템(200)의 사용자 또는 사용자의 환경과 연관된 다양한 데이터(이를테면, 예컨대, 오디오 또는 시각 데이터)를 수집할 수 있다. 다른 실시예들이 웨어러블 시스템이 사용되는 애플리케이션에 의존하여 추가적인 또는 더 적은 컴포넌트들을 가질 수 있다는 것이 인지되어야 한다. 그럼에도 불구하고, 도 2b는,

웨어러블 시스템을 통해 수집, 분석 및 저장될 수 있는 데이터의 타입들 및 다양한 컴포넌트들 중 일부의 기본 아이디어를 제공한다.

[0030] [0052] 도 2b는, 디스플레이(220)를 포함할 수 있는 예시적인 웨어러블 시스템(200)을 도시한다. 디스플레이(220)는, 사용자의 머리에 장착될 수 있는 디스플레이 렌즈(106) 또는 프레임(230)에 대응할 수 있는 프레임(108) 또는 하우징을 포함할 수 있다. 디스플레이 렌즈(106)는 사용자의 눈들(302, 304) 앞에 하우징(108)에 의해 포지셔닝된 하나 이상의 투명한 미러들을 포함할 수 있고, 로컬 환경으로부터 적어도 일부 광의 송신을 또한 허용하면서, 투사된 광(38)을 눈들(302, 304)로 바운싱하고, 빔 성형을 가능하게 하도록 구성될 수 있다. 투사된 광빔(38)의 파면은 투사된 광의 원하는 초점 거리와 일치하도록 구부러지거나 초점이 맞춰질 수 있다. 예시된 바와 같이, 2개의 넓은 시야 기계 비전 카메라들(16)(또한 세계 카메라들로서 또한 지칭됨)은 사용자 주변 환경을 이미징하기 위해 하우징(108)에 커플링될 수 있다. 이러한 카메라들(16)은 듀얼 캡처 가시광/비-가시광(예컨대, 적외선) 카메라들일 수 있다. 카메라들(16)은 도 4에 도시된 외향 이미징 시스템(464)의 부분일 수 있다. 세계 카메라들(16)에 의해 획득된 이미지는 포즈 프로세서(36)에 의해 프로세싱될 수 있다. 예컨대, 포즈 프로세서(36)는 사용자 또는 사용자의 환경 내의 다른 사람의 포즈를 식별하거나, 사용자의 환경 내의 물리적 오브젝트를 식별하기 위한 하나 이상의 오브젝트 인식기들(708)(예컨대, 도 7에 도시됨)을 구현할 수 있다.

[0031] [0053] 도 2b를 계속해서 참조하면, 광(38)을 눈들(302, 304)에 투사하도록 구성된 디스플레이 미러들 및 광학기를 갖는 한 쌍의 스캔-레이저 성형 파면(예컨대, 깊이에 대한 것임) 광 투사기 모듈들이 도시된다. 묘사된 도면은 또한 적외선 광 소스들(26)(이를테면, LED(light emitting diode)들)과 페어링된 2개의 미니어처 적외선 카메라들(24)을 도시하고, 이는 렌더링 및 사용자 입력을 지원하기 위해 사용자의 눈들(302, 304)을 추적할 수 있도록 구성된다. 카메라들(24)은 도 4에 도시된 내향 이미징 시스템(462)의 부분일 수 있다. 웨어러블 시스템(200)은, X, Y 및 Z 축 가속도계 능력뿐만 아니라 자기 컴퍼스 및 X, Y 및 Z 축 자이로 능력을 포함할 수 있는 센서 어셈블리(39)를 추가로 특징으로 하여, 바람직하게는 비교적 높은 주파수, 이를테면, 200Hz로 데이터를 제공한다. 센서 어셈블리(39)는 도 2a를 참조하여 설명된 IMU의 부분일 수 있다. 묘사된 시스템(200)은 또한, 캡처 디바이스들(16)로부터 출력된 넓은 시야 이미지 정보로부터 실시간 또는 거의 실시간의 사용자 머리 포즈를 계산하도록 구성될 수 있는 머리 포즈 프로세서(36), 이를테면, ASIC(application specific integrated circuit), FPGA(field programmable gate array) 및/또는 ARM(advanced reduced-instruction-set machine) 프로세서를 포함할 수 있다. 머리 포즈 프로세서(36)는 하드웨어 프로세서일 수 있고, 도 2a에 도시된 로컬 프로세싱 및 데이터 모듈(260)의 부분으로서 구현될 수 있다.

[0032] [0054] 센서 어셈블리(39)로부터의 자이로, 컴퍼스 및/또는 가속도계 데이터로부터 포즈를 도출하기 위해 디지털 및/또는 아날로그 프로세싱을 실행하도록 구성된 다른 프로세서(32)가 또한 도시된다. 프로세서(32)는 도 2a에 도시된 로컬 프로세싱 및 데이터 모듈(260)의 부분일 수 있다. 도 2b에 도시된 웨어러블 시스템(200)은 또한 포즈 및 포지셔닝 분석들을 보조하기 위한 포지션 시스템, 이를테면, 예컨대, GPS(global positioning system)(37)를 포함할 수 있다. 게다가, GPS는 추가로 사용자의 환경에 관한 원격-기반(예컨대, 클라우드-기반) 정보를 제공할 수 있다. 이러한 정보는 사용자의 환경의 오브젝트들 또는 정보를 인식하는 데 사용될 수 있다.

[0033] [0055] 웨어러블 시스템은, 사용자의 환경에 관한 더 많은 정보를 제공할 수 있는 GPS(37) 및 원격 컴퓨팅 시스템(이를테면, 예컨대, 원격 프로세싱 모듈(270), 다른 사용자의 ARD 등)에 의해 획득된 데이터를 조합할 수 있다. 일 예로서, 웨어러블 시스템은 GPS 데이터에 기반하여 사용자의 위치를 결정하고, 사용자의 위치와 연관된 가상 오브젝트들을 포함하는 세계 맵을 (예컨대, 원격 프로세싱 모듈(270)과 통신함으로써) 리트리브할 수 있다. 다른 예로서, 웨어러블 시스템(200)은 (도 4에 도시된 외향 이미징 시스템(464)의 부분일 수 있는) 세계 카메라들(16)을 사용하여 환경을 모니터링할 수 있다. 세계 카메라들(16)에 의해 획득된 이미지들에 기반하여, 웨어러블 시스템(200)은 (예컨대, 도 7에 도시된 하나 이상의 오브젝트 인식기들(708)을 사용함으로써) 환경의 캐릭터들을 검출할 수 있다. 웨어러블 시스템은 캐릭터들을 해석하기 위해 GPS(37)에 의해 획득된 데이터를 추가로 사용할 수 있다. 예컨대, 웨어러블 시스템(200)은, 캐릭터들이 위치된 지리 구역을 식별하고, 지리 구역과 연관된 하나 이상의 언어들을 식별할 수 있다. 그에 따라서, 웨어러블 시스템은 식별된 언어(들)에 기반하여, 예컨대, 식별된 언어(들)와 연관된 신택스, 문법, 문장 구조, 스펠링, 구두법 등에 기반하여 캐릭터들을 해석할 수 있다. 일 예에서, 독일의 사용자(210)는 아우토반에서 운전하면서 교통 표지판을 지각할 수 있다. 웨어러블 시스템(200)은, (단독으로 또는 세계 카메라(16)에 의해 획득된 이미지들과 조합하여) GPS(37)로부터 획득된 데이터에 기반하여, 사용자(210)가 독일에 있고, 이미징된 교통 표지판으로부터의 텍스트가 독일어일 가능

성이 있다고 식별할 수 있다.

[0034] [0056] 일부 상황들에서, 세계 카메라들(16)에 의해 획득된 이미지들은 사용자의 환경의 오브젝트의 불완전한 정보를 포함할 수 있다. 예컨대, 이미지는 흐릿한 대기, 텍스트의 흠집 또는 에러, 낮은 조명, 퍼지 이미지들, 폐색, 세계 카메라들(16)의 제한된 FOV 등으로 인해 불완전한 텍스트(예컨대, 문장, 글자 또는 구절)를 포함할 수 있다. 웨어러블 시스템(200)은 이미지에서 텍스트를 인식하는 데 있어서 상황 단서로서 GPS(37)에 의해 획득된 데이터를 사용할 수 있다.

[0035] [0057] 웨어러블 시스템(200)은 또한, 사용자가 세계를 보도록 스캐너들의 동작 및 사용자의 눈들로의 이미징을 가능하게 하기 위해서 사용자에 로컬인 렌더링 정보를 제공하도록 구성될 수 있는 렌더링 엔진(34)을 포함할 수 있다. 렌더링 엔진(34)은 하드웨어 프로세서(이를테면, 예컨대, 중앙 프로세싱 유닛 또는 그래픽 프로세싱 유닛)에 의해 구현될 수 있다. 일부 실시예들에서, 렌더링 엔진은 로컬 프로세싱 및 데이터 모듈(260)의 부분이다. 렌더링 엔진(34)은 (예컨대, 유선 또는 무선 링크들을 통해) 웨어러블 시스템(200)의 다른 컴포넌트들에 통신가능하게 커플링될 수 있다. 예컨대, 렌더링 엔진(34)은 통신 링크(102)를 통해 눈 카메라들(24)에 커플링되고, 통신 링크(104)를 통해 투사 서브시스템(18)(망막 스캐닝 디스플레이와 유사한 방식으로 스캔된 레이저 어레인지먼트를 통해 사용자의 눈들(302, 304)에 광을 투사할 수 있음)에 커플링될 수 있다. 렌더링 엔진(34)은 또한 링크들(105 및 94)을 통해 다른 프로세싱 유닛들, 이를테면, 예컨대, 센서 포즈 프로세서(32) 및 이미지 포즈 프로세서(36)와 각각 통신할 수 있다.

[0036] [0058] 카메라들(24)(예컨대, 미니 적외선 카메라들)은 렌더링 및 사용자 입력을 지원하기 위해 눈 포즈를 추적하는 데 활용될 수 있다. 일부 예시적인 눈 포즈들은 사용자가 어디를 보고 있는지 또는 사용자가 어떠한 깊이에 초점을 맞추는지(사용자가 눈 이접운동으로 추정될 수 있음)를 포함할 수 있다. GPS(37), 자이로들, 컴퍼스 및 가속도계들(39)은 개략적인(coarse) 또는 빠른 포즈 추정들을 제공하는 데 활용될 수 있다. 카메라들(16) 중 하나 이상은, 로컬 환경을 맵핑하고 사용자 뷰들을 다른 사용자들과 공유하는 데 활용될 수 있는, 연관된 클라우드 컴퓨팅 자원으로부터의 데이터와 연동하여 이미지들 및 포즈를 획득할 수 있다.

[0037] [0059] 도 2b에 묘사된 예시적인 컴포넌트들은 단지 예시 목적들을 위한 것이다. 예시 및 설명을 용이하게 하기 위해서 다수의 센서들 및 다른 기능적 모듈들이 함께 도시된다. 일부 실시예들은 이들 센서들 또는 모듈들 중 하나 또는 서브세트만을 포함할 수 있다. 추가로, 이들 컴포넌트들의 위치들은 도 2b에 묘사된 포지션들로 제한되지 않는다. 일부 컴포넌트들은 다른 컴포넌트들, 이를테면, 벨트-장착된 컴포넌트, 핸드-헬드 컴포넌트, 또는 헬멧 컴포넌트에 장착되거나 그 안에 하우징될 수 있다. 일 예로서, 이미지 포즈 프로세서(36), 센서 포즈 프로세서(32) 및 렌더링 엔진(34)은 벨트팩에 포지셔닝되고, 무선 통신, 이를테면, 초광대역, Wi-Fi, 블루투스 등을 통해, 또는 유선 통신을 통해 웨어러블 시스템의 다른 컴포넌트들과 통신하도록 구성될 수 있다. 묘사된 하우징(108)은 바람직하게는 사용자에 의해 머리 장착가능하고, 착용가능하다. 그러나, 웨어러블 시스템(200)의 일부 컴포넌트들은 사용자의 신체의 다른 부분들에 착용될 수 있다. 예컨대, 스피커(240)는 사운드를 사용자에게 제공하기 위해 사용자의 귀들에 삽입될 수 있다.

[0038] [0060] 사용자의 눈들(302, 304)로의 광(38)의 투사에 관하여, 일부 실시예에서, 카메라들(24)은, 사용자의 눈들(302, 304)의 중심들이 기하학적으로 어디로 향하는지(verge), 일반적으로 어떤 것이 초점의 포지션과 일치하는지, 또는 눈들(302, 304)의 "초점의 깊이"를 측정하는 데 활용될 수 있다. 눈들이 향하는 모든 포인트들의 3-차원 표면은 "호롭터(horopter)"로서 지칭될 수 있다. 초점 거리는 유한수의 깊이들을 받아들일 수 있거나, 무한적으로 가변할 수 있다. 버전스 거리로부터 투사된 광은 피험자 눈(302, 304)에 초점이 맞춰지는 것처럼 보이는 반면, 버전스 거리 앞 또는 뒤의 광은 흐릿해 진다. 본 개시내용의 웨어러블 디바이스들 및 다른 디스플레이 시스템들의 예들은 또한, 인용에 의해 본원에 전체가 통합되는 미국 특허 공보 제2016/0270656호에 설명된다.

[0039] [0061] 인간 시각 시스템은 복잡하고, 깊이의 현실적 지각을 제공하는 것은 난제이다. 오브젝트의 뷰어들은 이접운동과 원근조절의 조합으로 인해 오브젝트를 3-차원인 것으로 지각할 수 있다. 서로에 대해 2개의 눈들의 이접운동 움직임들(예컨대, 오브젝트를 응시하기 위해 눈들의 시선들을 수렴시키기 위해 서로를 향한 또는 서로부터 떨어진 동공들의 롤링 움직임들)은 눈들의 렌즈들의 초점 맞춤(또는 "원근조절")과 근접하게 연관된다. 정상 조건들 하에서, 하나의 오브젝트로부터 상이한 거리에서의 다른 오브젝트로 초점을 변경하기 위해, 눈들의 렌즈들의 초점을 변경하거나, 눈들을 원근조절하는 것은, "원근조절-이접운동 반사작용"으로 알려진 관계 하에서, 동일한 거리에 대한 이접운동에서 매칭 변화를 자동적으로 유발할 것이다. 마찬가지로, 이접운동의 변화는, 정상 조건들 하에서, 원근조절의 매칭 변화를 트리거할 것이다. 원근조절과 이접운동 간의 더 나은 매

치를 제공하는 디스플레이 시스템들은 3-차원 이미저리의 더 현실적이고 편안한 시뮬레이션들을 형성할 수 있다.

[0040] [0062] 약 0.7밀리미터 미만의 빔 직경을 갖는 추가적인 공간적으로 코히어런트 광은 눈이 어디에 초점을 맞추는지에 상관없이 인간 눈에 의해 정확하게 분석될 수 있다. 따라서, 적절한 초점 깊이의 착시효과(illusion)를 생성하기 위해, 눈 이접운동은 카메라들(24)로 추적될 수 있고, 렌더링 엔진(34) 및 투사 서브시스템(18)은 호롭터 상의 또는 호롭터에 근접한 모든 오브젝트들을 초점이 맞게 렌더링하고, (예컨대, 의도적으로 생성된 흐림을 사용하여) 다른 모든 오브젝트들을 다양한 수준으로 초점에서 벗어나게 렌더링하는 데 활용될 수 있다. 바람직하게는, 시스템(220)은 초당 약 60개 이상의 프레임들의 프레임 레이트로 사용자에게 렌더링한다. 앞서 설명된 바와 같이, 바람직하게는, 카메라들(24)은 눈 추적을 위해 활용될 수 있고, 소프트웨어는 사용자 입력들의 역할을 하기 위해 이접운동 기하학적 구조뿐만 아니라 초점 위치 큐들을 선택하도록 구성될 수 있다. 바람직하게는, 이러한 디스플레이 시스템은 낮 또는 밤에 사용하기에 적절한 밝기 및 콘트라스트로 구성된다.

[0041] [0063] 일부 실시예들에서, 디스플레이 시스템은 바람직하게는 시각 오브젝트 정렬에 대해 약 20 밀리초 미만의 레이턴시, 약 0.1도 미만의 각도 정렬, 및 약 1 분각(arc minute)의 해상도를 갖고, 이는, 이론에 의해 제한되지 않고서, 대략 인간 눈의 한계인 것으로 믿겨진다. 디스플레이 시스템(220)은, 포지션 및 포즈 결정을 보조하기 위해 로컬화 시스템과 통합될 수 있고, 로컬화 시스템은 GPS 엘리먼트들, 광학 추적 컴퍼스, 가속도계들, 또는 다른 데이터 소스들을 수반할 수 있고, 로컬화 정보는 적절한 세계의 사용자의 뷰에서 정확한 렌더링을 가능하게 하는 데 활용될 수 있다(예컨대, 이러한 정보는 실세계에 관련하여 사용자들이 어디에 있는지를 안경이 아는 것을 가능하게 할 것임).

[0042] [0064] 일부 실시예들에서, 웨어러블 시스템(200)은 사용자의 눈들의 원근조절에 기반하여 하나 이상의 가상 이미지들을 디스플레이하도록 구성된다. 이미지들이 투사되는 곳에 초점을 맞추도록 사용자를 강제하는 종래의 3D 디스플레이 접근법들과 달리, 일부 실시예들에서, 웨어러블 시스템은, 사용자에게 제공되는 하나 이상의 이미지들의 더 편안한 보기를 허용하기 위해, 투사된 가상 콘텐츠의 초점을 자동으로 변경하도록 구성된다. 예컨대, 만약 사용자의 눈들이 1m의 최신의 초점을 갖는다면, 이미지는 사용자의 초점과 일치하도록 투사될 수 있다. 만약 사용자가 3m의 초점을 시프팅하면, 이미지는 새로운 초점과 일치하도록 투사된다. 따라서, 미리 결정된 초점에 사용자를 강제하기보다는, 일부 실시예들의 웨어러블 시스템(200)은 사용자의 눈이 더 자연스러운 방식으로 기능하게 허용한다.

[0043] [0065] 이러한 웨어러블 시스템(200)은, 가상 현실 디바이스들에 대하여 전형적으로 관찰되는 눈 피로, 두통들, 및 다른 생리적 증상들의 발생들을 제거 또는 감소시킬 수 있다. 이를 달성하기 위해, 웨어러블 시스템(200)의 다양한 실시예들은, 하나 이상의 VFE(variable focus element)들을 통해, 가변 초점 거리들에서 가상 이미지들을 투사하도록 구성된다. 하나 이상의 실시예들에서, 3D 지각은, 사용자로부터 떨어진 고정된 초점 평면들에서 이미지들을 투사하는 다중-평면 초점 시스템을 통해 달성될 수 있다. 다른 실시예들은 가변 평면 초점을 이용하고, 여기서 초점 평면은 사용자의 현재의 초점 상태와 일치하도록 z-방향으로 앞뒤로 이동된다.

[0044] [0066] 다중-평면 초점 시스템들 및 가변 평면 초점 시스템들 둘 모두에서, 웨어러블 시스템(200)은 사용자의 눈들의 이접운동을 결정하고, 사용자의 최신의 초점을 결정하고, 결정된 초점에서 가상 이미지를 투사하기 위해 눈 추적을 이용할 수 있다. 다른 실시예들에서, 웨어러블 시스템(200)은, 파이버 스캐너, 또는 광 생성 소스를 통해, 망막에 걸쳐 래스터 패턴의 가변 초점의 광빔들을 가변적으로 투사하는 광 변조기를 포함한다. 따라서, 가변 초점 거리들에서 이미지들을 투사하기 위한 웨어러블 시스템(200)의 디스플레이의 능력은 사용자가 3D로 오브젝트들을 보기 위한 원근조절을 용이하게 할 뿐만 아니라, 본원에 인용에 의해 전체적으로 통합되는 미국 특허 공보 제2016/0270656호에 추가로 설명되는 바와 같이, 사용자 안구 이상들을 보상하는 데 사용될 수 있다. 일부 다른 실시예들에서, 공간 광 변조기는 다양한 광학 컴포넌트들을 통해 이미지들을 사용자에게 투사할 수 있다. 예컨대, 아래에 추가로 설명되는 바와 같이, 공간 광 변조기는 이미지들을 하나 이상의 도파관들에 투사할 수 있고, 그런 다음 도파관들은 이미지들을 사용자에 송신한다.

[0045] [0067] 도 3은 다중 깊이 평면들을 사용하여 3-차원 이미저리를 시뮬레이팅하기 위한 접근법의 양상들을 예시한다. 도 3을 참조하면, z-축 상에서 눈들(302 및 304)로부터 다양한 거리들에 있는 오브젝트들은, 이들 오브젝트들이 초점을 맞추도록 눈들(302 및 304)에 의해 원근조절된다. 눈들(302 및 304)은 z-축을 따라 상이한 거리들에 있는 오브젝트들에 초점을 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 오브젝트들 또는 오브젝트들의 부분들은, 눈이 해당 깊이 평면에 대해 원근조절된

상태에 있을 때 초점을 맞추게 된다. 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 평탄한 것으로 도시되지만, 깊이 평면의 윤곽들은 물리적 공간에서 만곡될 수 있어서, 깊이 평면의 모든 특징들은 특정 원근조절된 상태에서 눈과 초점을 맞춘다는 것이 인지될 것이다. 이론에 의해 제한됨이 없이, 인간 눈이 전형적으로 깊이 지각을 제공하기 위하여 유한 수의 깊이 평면들을 해석할 수 있다고 여겨진다. 결과적으로, 지각된 깊이의 매우 믿을 수 있는 시뮬레이션은, 눈에, 이들 제한된 수의 깊이 평면들 각각에 대응하는 이미지의 상이한 프리젠테이션들을 제공함으로써 달성될 수 있다.

[0046] 도파관 스택 어셈블리

[0047] [0068] 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 2a의 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 2a의 디스플레이(220)에 통합될 수 있다.

[0048] [0069] 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, 공기 갭들을 형성하기 위한 클래딩 층들 또는 구조들)일 수 있다.

[0049] [0070] 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 눈(410)을 향하여 출력하도록, 각각의 개개의 도파관을 가로질러 인입 광을 분배하도록 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 에지에 주입된다. 일부 실시예들에서, 단일 광빔(예컨대, 시준된 빔)은 특정 도파관과 연관된 깊이 평면에 대응하는 특정 각도들(및 발산의 양들)로 눈(410) 쪽으로 지향되는 시준된 클론 빔(cloned collimated beam)들의 전체 필드를 출력하기 위하여 각각의 도파관으로 주입될 수 있다.

[0050] [0071] 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 도관들(예컨대, 광섬유 케이블들)을 통하여 이미지 주입 디바이스들(420, 422, 424, 426, 428) 각각에 파이핑할 수 있는 단일 멀티플렉싱된 디스플레이의 출력 단부들이다.

[0051] [0072] 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 조절하는 프로그래밍(예컨대, 비-일시적 컴퓨터-판독가능한 매체의 명령들)을 포함한다. 일부 실시예들에서, 제어기(460)는 단일 통합 디바이스, 또는 유선 또는 무선 통신 채널들에 의해 연결되는 분산형 시스템일 수 있다. 제어기(460)는 일부 실시예들에서, 프로세싱 모듈들(260 또는 270)(도 2a에 예시됨)의 부분일 수 있다.

[0052] [0073] 도파관들(440b, 438b, 436b, 434b, 432b)은 TIR(total internal reflection)에 의해 각각의 개개의 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 개개의 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 아웃

커플링 광학 엘리먼트들로서 지칭될 수 있다. 추출된 광빔은, 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 회절 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 계층에 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.

[0053] [0074] 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)에서 나오는 광을 광학 무한대로부터 눈(410)을 향해 안쪽으로 더 가까운 제1 초점 평면에서 나오는 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 조합된 광학 전력은 다른 증분 양의 파면 곡률을 생성하도록 구성될 수 있어서, 눈/뇌는 제3 도파관(436b)에서 나오는 광을, 상기 다음 위의 도파관(434b)으로부터의 광보다는 광학 무한대로부터 사람을 향하여 안쪽으로 훨씬 더 가까운 제2 초점 평면에서 나오는 것으로 해석한다.

[0054] [0075] 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 간의 렌즈들 모두를 통하여 전송한다. 스택된 도파관 어셈블리(480)의 다른 측 상에서 세계(470)에서 나오는 광을 보거나/해석할 때 렌즈들(458, 456, 454, 452)의 스택을 보상하기 위하여, 보상 렌즈 계층(430)은 아래의 렌즈 스택(458, 456, 454, 452)의 어그리게이트 전력을 보상하기 위하여 스택의 최상부에 배치될 수 있다. 이러한 구성은 이용가능한 도파관/렌즈 쌍들이 존재하는 만큼 많은 지각된 초점 평면들을 제공한다. 도파관들의 광 추출 광학 엘리먼트들 및 렌즈들의 초점을 맞추는 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.

[0055] [0076] 도 4를 계속 참조하면, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 자신의 개개의 도파관들로부터 광을 재지향하는 것은 물론, 도파관과 연관된 특정 깊이 평면에 대해 적합한 발산량 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, 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.

[0056] [0077] 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 바운싱되기 때문에 눈(304)을 향하는 상당히 균일한 출사 방출 패턴이다.

[0057] [0078] 일부 실시예들에서, 하나 이상의 DOE들은, 그것들이 활성적으로 회절시키는 "온" 상태와 그것들이 상당히 회절시키지 않는 "오프" 상태 사이에서 스위칭 가능할 수 있다. 예컨대, 스위칭 가능 DOE는, 마이크로액적

들이 호스트 매질에서 회절 패턴을 포함하는 중합체 분산형 액정 계층을 포함할 수 있고, 마이크로액적들의 굴절률은 호스트 재료의 굴절률에 실질적으로 매칭하도록 스위칭될 수 있거나(이 경우에 패턴은 입사 광을 현저하게 회절시키지 않음) 또는 마이크로액적은 호스트 매질의 인덱스에 매칭하지 않는 인덱스로 스위칭될 수 있다(이 경우 패턴은 입사 광을 활성적으로 회절시킴).

[0058] [0079] 일부 실시예들에서, 피사계 심도 또는 깊이 평면들의 수 및 분배는 뷰어의 눈들의 동공 사이즈들 또는 배향들에 기반하여 동적으로 변동될 수 있다. 피사계 심도는 뷰어의 동공 사이즈와 반대로 변할 수 있다. 결과적으로, 뷰어의 눈들의 동공들의 사이즈들이 감소함에 따라, 피사계 심도가 증가하여서, 식별 불가능한 하나의 평면의 위치가 눈의 초점의 깊이를 넘어서기 때문에 그 평면이 식별 가능하게 되고 동공 사이즈의 감소를 통해 보다 초점이 맞게 나타나고 피사계 심도의 증가와 상응할 수 있다. 마찬가지로, 뷰어에게 상이한 이미지들을 제공하는 데 사용되는 이격된 깊이 평면들의 수는 감소된 동공 사이즈에 따라 감소될 수 있다. 예컨대, 뷰어는 하나의 깊이 평면으로부터 벗어나게 그리고 다른 깊이 평면으로 눈의 원근조절을 조정하지 않고서는, 하나의 동공 사이즈에서 제1 깊이 평면 및 제2 깊이 평면 둘 모두의 세부사항들을 명확하게 지각할 수 없을 수 있다. 그러나, 이러한 2개의 깊이 평면들은 원근조절을 변하지 않고도 다른 동공 사이즈에서 사용자에게 동시에 충분히 초점을 맞출 수 있다.

[0059] [0080] 일부 실시예들에서, 디스플레이 시스템은 동공 사이즈 또는 배향의 결정들에 또는 특정 동공 사이즈 또는 배향을 나타내는 전기 신호들의 수신에 기반하여 이미지 정보를 수신하는 도파관들의 수를 변하게 할 수 있다. 예컨대, 만약 사용자의 눈들이 2개의 도파관들과 연관된 2개의 깊이 평면들을 구별할 수 없다면, 제어기(460)(로컬 프로세싱 및 데이터 모듈(260)의 실시예일 수 있음)는 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.

[0060] [0081] 일부 실시예들에서, 출사 빔이 뷰어의 눈의 직경 미만인 직경을 갖는 조건을 충족시키는 것이 바람직할 수 있다. 그러나 이 조건을 충족시키는 것은 뷰어의 동공들의 사이즈의 변동성을 고려하면 난제시될 수 있다. 일부 실시예들에서, 이 조건은 뷰어의 동공의 사이즈의 결정에 대한 응답으로 출사 빔의 사이즈를 변하게 함으로써 넓은 범위의 동공 사이즈들에 걸쳐 충족된다. 예컨대, 동공 사이즈가 감소함에 따라, 출사 빔의 사이즈가 또한 감소할 수 있다. 일부 실시예들에서, 출사 빔 사이즈는 가변 어퍼처를 사용하여 변동될 수 있다.

[0061] [0082] 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 때로는 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)보다 클 수 있다. 뷰어에 의한 이미징 또는 보기를 위해 이용가능한 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 보다 억제될 수 있고, 그에 따라서 착용자의 FOR은 보다 더 작은 입체각을 마주할 수 있다(subtend). 외향 이미징 시스템(464)으로부터 획득된 이미지들은 사용자에 의해 행해진 제스처들(예컨대, 손 또는 손가락 제스처들)을 추적하고, 사용자 앞의 세계(470)의 오브젝트들을 검출하는 등을 행하는 데 사용될 수 있다.

[0062] [0083] 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 입체 사운드 수신을 제공하도록 포지셔닝될 수 있다. 오디오 센서(232)는 다른 예로서, 오디오 소스가 위치되는 곳에 대한 그러한 유용한 지향성 정보를 또한 제공할 수 있는 지향성 마이크로폰을 포함할 수 있다.

[0063] [0084] 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 사이즈 또는 배향을 결정하기 위해 눈(410)의 이미지들을 캡처하는 데 사용될 수 있다. 내향 이미징 시스템(466)은, (예컨대, 홍채 식별을 통해) 사용자의 생체인식 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 사이즈 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 눈들마다 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 획득진 이미지들은 사용자의 눈 포즈 또는 분위기(mood)를 결정하도록 분석될 수 있으며, 이는 어떤 청각적 또는 시각적 콘텐츠가 사용자에게 제공되어야 하는지를 판단하기 위해 웨어러블 시스템(400)에 의해 사용될 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.

[0064] [0085] 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, 전방향/후방향 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 휴대될 수 있다. 사용자 입력 디바이스(466)는 웨어러블 시스템(400)과 유선 또는 무선 통신할 수 있다.

[0065] [0086] 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)에 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 의존하여, 일정 각도로 눈(410)으로 전파되도록 재지향될 수 있다(예컨대, 발산 출사 빔들을 형성함). 실질적으로 평행한 출사 빔들은, 눈(410)으로부터 먼 거리(예컨대, 광학적 무한대)에 있는 깊이 평면 상에 세팅된 것으로 나타나는 이미지들을 형성하도록 광을 아웃커플링하는 광 추출 광학 엘리먼트들을 갖는 도파관을 표시할 수 있다는 것이 인지될 것이다. 다른 도파관들 또는 광 추출 광학 엘리먼트들의 다른 세트들은 더 발산하는 출사 빔 패턴을 출력할 수 있으며, 이는, 눈(410)이 망막 상에 초점을 맞추게 하기 위해 더 근접한 거리로 원근조절하는 것을 요구할 것이고 광학적 무한대보다 눈(410)에 더 근접한 거리로부터의 광으로서 뇌에 의해 해석될 것이다.

[0066] [0087] 도 6은, 다-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 개략적 다이어그램이다. 광학 시스템은 도파관 장치, 도파관 장치로 또는 상기 도파관 장치로부터의 광을 광학적으로 커플링하는 광학 커플러 서브시스템, 및 제어 서브시스템을 포함할 수 있다. 광학 시스템은 다-초점 볼류메트릭, 이미지 또는 광 필드를 생성하는 데 사용할 수 있다. 광학 시스템은 하나 이상의 주 평면형 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. 평면형 도파관들(632b)은 도 4를 참조하여 논의된 도파관들(432b, 434b, 436b, 438b, 440b)과 유사할 수 있다. 광학 시스템은 제1 축(도 6의 도면에서 수직 또는 Y-축)을 따라 광을 중계하고 제1 축(예컨대, Y-축)을 따라 광의 유효 출사 동공을 확장시키기 위해 분배 도파관 장치를 이용할 수 있다. 분배 도파관 장치는, 예컨대, 분배 평면형 도파관(622b) 및 분배 평면형 도파관(622b)과 연관된 적어도 하나의 DOE(622a)(이중 일점 쇄선으로 예시됨)를 포함할 수 있다. 분배 평면형 도파관(622b)은 그와 상이한 배향을 갖는 주 평면형 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 평면형 도파관(622b) 또는 DOE(622a)는 각각, 주 평면형 도파관(632b) 또는 DOE(632a)와 동일한 재료들로 구성될 수 있다. 도 6에 도시된 광학 디스플레이 시스템(600)의 실시예들은 도 2a에 도시된 웨어러블 시스템(200)에 통합될 수 있다.

[0067] [0088] 중계된 그리고 출사-동공 확장된 광은 분배 도파관 장치로부터 하나 이상의 주 평면형 도파관들(632b)에 광학적으로 커플링될 수 있다. 주 평면형 도파관(632b)은 바람직하게는, 제1 축에 직교하는 제2 축(예컨대,

도 6의 도면에서 수평 또는 X-축)을 따라 광을 중계할 수 있다. 특히, 제2 축은 제1 축에 대해 비-직교 축일 수 있다. 주 평면형 도파관(632b)은 그 제2 축(예컨대, X-축)을 따라 광의 유효 출사 동공을 확장시킨다. 예컨대, 분배 평면형 도파관(622b)은 수직 또는 Y- 축을 따라 광을 중계 및 확장시키고, 수평 또는 X-축을 따라 광을 중계 및 확장시킬 수 있는 주 평면형 도파관(632b)으로 그 광을 전달할 수 있다.

[0068] [0089] 광학 시스템은 단일 모드 광섬유(640)의 근단부에 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 원단부는 압전 재료의 중공 튜브(642)를 통해 스레딩(thread)되거나 수용될 수 있다. 원단부는 고정되지 않은 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, 외측 표면 또는 외측 주변부 또는 직경 상에 도금될 수 있다. 코어 전극(예시되지 않음)은 또한 튜브(642)의 코어, 중심, 내부 주변부 또는 내부 직경에 위치될 수 있다.

[0069] [0090] 예컨대, 와이어들(660)을 통해 전기적으로 커플링된 구동 전자 장치(650)는 2개의 축들에서 압전 튜브(642)를 독립적으로 구부리기 위해 전극들의 대향하는 쌍들을 구동한다. 광섬유(644)의 돌출 원단 끝 부분은 기계 공진 모드들을 갖는다. 공진 주파수들은 광섬유(644)의 직경, 길이 및 재료 특성들에 의존할 수 있다. 섬유 캔틸레버(644)의 제1 기계 공진 모드 인근에서 압전 튜브(642)를 진동시킴으로써, 섬유 캔틸레버(644)는 진동하게 될 수 있고, 큰 편향들을 통해 스윕(sweep)할 수 있다.

[0070] [0091] 2개의 축들에서 공진 진동을 자극함으로써, 섬유 캔틸레버(644)의 끝 부분은 영역 채움 2-차원(2D) 스캔으로 2축 방향으로(biaxially) 스캔된다. 섬유 캔틸레버(644)의 스캔과 동기하여 광원(들)(610)의 세기를 변조함으로써, 섬유 캔틸레버(644)로부터 나오는 광은 이미지를 형성할 수 있다. 그러한 셋업에 대한 설명들은 미국 특허 공보 제2014/0003762호에서 제공되며, 그 공보는 그 전체가 본원에 인용에 의해 통합된다.

[0071] [0092] 광학 커플러 서브시스템의 컴포넌트는 스캐닝 섬유 캔틸레버(644)로부터 나오는 광을 시준할 수 있다. 시준된 광은 미러 표면(648)에 의해 적어도 하나의 DOE(diffractive optical element)(622a)를 포함하는 좁은 분배 평면형 도파관(622b)으로 반사될 수 있다. 시준된 광은 TIR에 의해 분배 평면형 도파관(622b)을 따라(도 6의 도면에 대해) 수직으로 전파될 수 있고, 이렇게 하여, DOE(622a)와 반복적으로 교차한다. DOE(622a)는 바람직하게는, 낮은 회절 효율성을 갖는다. 이는, 광의 일부(예컨대, 10%)로 하여금 DOE(622a)와의 각각의 교차점에서 더 큰 주 평면형 도파관(632b)의 에지를 향해 회절되게 하고, 광의 일부로 하여금 TIR을 통해 분배 평면형 도파관(622b)의 길이 아래에서 그의 오리지널 궤적 상에서 계속되게 할 수 있다.

[0072] [0093] DOE(622a)와의 각각의 교차점에서, 추가적인 광이 주 도파관(632b)의 입구를 향해 회절될 수 있다. 인입 광을 다수의 아웃커플링된 세트들로 분할함으로써, 광의 출사 동공은 분배 평면형 도파관(622b)에서 DOE(622a)에 의해 수직으로 확장될 수 있다. 분배 평면형 도파관(622b) 밖으로 커플링되는 이러한 수직으로 확장된 광은 주 평면형 도파관(632b)의 에지에 들어갈 수 있다.

[0073] [0094] 주 도파관(632b)에 들어가는 광은 TIR을 통해 주 도파관(632b)을 따라(도 6의 도면에 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 초점 맞추기 둘 모두를 생성하도록, 선형 회절 패턴 및 방사상 대칭 회절 패턴의 합인 위상 프로파일을 갖도록 설계 또는 구성될 수 있다. DOE(632a)는 유리하게는, 낮은 회절 효율성(예컨대, 10%)을 가질 수 있어서, 빔의 광 중 일부만이 DOE(632a)의 각각의 교차에 의해 뷰의 눈을 향해 편향되는 반면, 광의 나머지는 TIR을 통해 주 도파관(632b)을 통하여 계속 전파된다.

[0074] [0095] 전파되는 광과 DOE(632a) 간의 각각의 교차점에서, 광의 일부가 주 도파관(632b)의 인접한 면을 향해 회절되어, 광이 TIR을 벗어나 주 도파관(632b)의 면으로부터 나오게 허용한다. 일부 실시예들에서, DOE(632a)의 방사상 대칭 회절 패턴은 추가적으로, 회절된 광에 초점 레벨을 부여하여, 개별 빔의 광 파면을 성형(예컨대, 곡률을 부여함)하는 것뿐만 아니라, 설계된 초점 레벨과 매칭하는 각도로 빔을 조종한다.

[0075] [0096] 그에 따라서, 이들 상이한 경로들은 상이한 각도들의 다수의 DOE들(632a), 초점 레벨들, 또는 출사 동공에서 상이한 채움 패턴들을 산출하는 것에 의해 광으로 하여금 주 평면형 도파관(632b) 밖으로 커플링되게 할 수 있다. 출사 동공에서의 상이한 채움 패턴들은 다수의 깊이 평면들을 갖는 광 필드 디스플레이를 생성하는 데 유익하게 사용될 수 있다. 도파관 어셈블리 내의 각각의 계층 또는 스택 내의 한 세트의 계층들(예컨대, 3개의 계층들)은 개개의 컬러(예컨대, 적색, 청색, 녹색)를 생성하는 데 이용될 수 있다. 따라서, 예컨대, 제1 세트의 3개의 인접한 계층들은 제1 초점 깊이로 적색, 청색 및 녹색 광을 각각 생성하는 데 이용될 수 있다.

제2 세트의 3개의 인접한 계층들은 제2 초점 깊이로 적색, 청색 및 녹색 광을 각각 생성하는 데 이용될 수 있다. 다수의 세트들은 다양한 초점 깊이들을 갖는 풀 3D 또는 4D 컬러 이미지 광 필드를 생성하는 데 이용될 수 있다.

[0076] 웨어러블 시스템의 다른 컴포넌트들

[0077] [0097] 다수의 구현들에서, 웨어러블 시스템은 위에서 설명된 웨어러블 시스템의 컴포넌트들에 추가적으로 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 동작가능할 수 있다. 예컨대, 햅틱 디바이스들 또는 컴포넌트들은 가상 콘텐츠(예컨대, 가상 오브젝트들, 가상 도구들, 다른 가상 구조들)를 터치할 때 압력 또는 텍스처의 촉감을 제공할 수 있다. 촉감은 가상 오브젝트가 표현하는 물리적 오브젝트의 느낌을 복제할 수 있거나, 또는 가상 콘텐츠가 표현하는 상상의 오브젝트 또는 캐릭터(예컨대, 용)의 느낌을 복제할 수 있다. 일부 구현들에서, 햅틱 디바이스들 또는 컴포넌트들은 사용자에 의해 착용될 수 있다(예컨대, 사용자 웨어러블 글러브). 일부 구현들에서, 햅틱 디바이스들 또는 컴포넌트들은 사용자에 의해 홀딩될 수 있다.

[0078] [0098] 웨어러블 시스템은 예컨대, 웨어러블 시스템과의 입력 또는 상호작용을 허용하도록 사용자에 의해 조작가능한 하나 이상의 물리적 오브젝트들을 포함할 수 있다. 이러한 물리적 오브젝트들은 본원에서 토템들로서 지칭될 수 있다. 일부 토템들은 무생물 오브젝트들, 예컨대, 이를테면 금속 또는 플라스틱의 피스, 벽, 테이블의 표면의 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 토템의 하나 이상의 표면들 상에 있는 것으로 사용자에게 나타나도록 사용자 인터페이스를 렌더링할 수 있다. 예컨대, 웨어러블 시스템은 토템의 하나 이상의 표면들에 상주하는 것으로 나타나도록 컴퓨터 키보드 및 트랙패드의 이미지를 렌더링할 수 있다. 예컨대, 웨어러블 시스템은 토템으로서 역할을 하는 알루미늄의 얇은 직사각형 플레이트의 표면 상에 나타나도록 가상 컴퓨터 키보드 및 가상 트랙패드를 렌더링할 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 토템의 실시예일 수 있다. 사용자는 웨어러블 시스템 또는 다른 사용자들과 상호작용하기 위해 토템을 단독으로 또는 포즈들과 조합하여 사용할 수 있다.

[0079] [0099] 본 개시내용의 웨어러블 디바이스들, HMD 및 디스플레이 시스템들과 함께 사용가능한 햅틱 디바이스들 및 토템들의 예들은 미국 특허 공보 제2015/0016777호에서 설명되며, 그 공보는 그 전체가 본원에 인용에 의해 통합된다.

[0080] 예시적인 웨어러블 시스템들, 환경들, 및 인터페이스들

[0081] [0100] 웨어러블 시스템은 렌더링된 광 필드들에서 높은 피사계 심도를 달성하기 위해 다양한 맵핑 관련 기법들을 이용할 수 있다. 가상 세계의 맵핑 시에, 실세계에 관하여 가상 오브젝트들을 정확히 묘사하기 위해 실세계의 모든 피처들 및 포인트들을 아는 것이 유리하다. 이를 위해, 웨어러블 시스템의 사용자들로부터 캡처된 FOV 이미지들은, 실세계의 다양한 포인트들 및 피처들에 관한 정보를 운반하는 새로운 사진들을 포함함으로써 세계 모델에 추가될 수 있다. 예컨대, 웨어러블 시스템은, 한 세트의 맵 포인트들(이를테면, 2D 포인트들 또는 3D 포인트들)을 수집하고 새로운 맵 포인트들을 발견하여 세계 모델의 보다 정확한 버전을 렌더링할 수 있다. 제1 사용자의 세계 모델은 (예컨대, 네트워크, 이를테면 클라우드 네트워크를 통해) 제2 사용자에게 통신될 수 있어서, 제2 사용자는 제1 사용자 주변의 세계를 경험할 수 있다.

[0082] [0101] 도 7은 MR 환경(700)의 예의 블록 다이어그램이다. MR 환경(700)은 하나 이상의 사용자 웨어러블 시스템들(예컨대, 웨어러블 시스템(200) 또는 디스플레이 시스템(220)) 또는 고정 룸(stationary room) 시스템들(예컨대, 룸 카메라들 등)로부터 입력(예컨대, 사용자의 웨어러블 시스템으로부터의 시각 입력(702), 이를테면, 룸 카메라들로부터의 고정 입력(704), 다양한 센서들로부터의 감각 입력(706), 제스처들, 토템들, 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 고정 카메라들로부터의 정보 또는 상이한 관점으로부터의 다양한 큐들로 보충될 수 있다. 카메라들(이를테면, 룸 카메라들 또는 외향 이미징 시스템의 카메라들)에 의해 획득된 이미지 데이터는 한 세트의 맵핑 포인트들로 감소될 수 있다.

[0083] [0102] 하나 이상의 오브젝트 인식기들(708)은 맵 데이터베이스(710)의 도움으로, 수신된 데이터(예컨대, 포인트들의 모음)를 크롤링(crawl)하고, 포인트들을 인식 또는 맵핑하고, 이미지들을 태깅하고, 시맨틱 정보를 오브젝트들에 부착할 수 있다. 맵 데이터베이스(710)는 시간이 지남에 따라 수집된 다양한 포인트들 및 그들의 대응하는 오브젝트들을 포함할 수 있다. 다양한 디바이스들 및 맵 데이터베이스는 클라우드에 액세스하기 위해 네트워크(예컨대, LAN, WAN 등)를 통해 서로 연결될 수 있다.

[0084] [0103] 맵 데이터베이스의 이러한 정보 및 포인트들의 모음에 기반하여, 오브젝트 인식기들(708a 내지 708n)은 환경 내의 오브젝트들을 인식할 수 있다. 예컨대, 오브젝트 인식기들은 얼굴들, 사람들, 창문들, 벽들, 사용자 입력 디바이스들, 텔레비전들, 문서들(예컨대, 본원의 보안 예들에서 설명된 바와 같은 여행 티켓들, 운전 면허증, 여권), 사용자의 환경 내의 다른 오브젝트들 등을 인식할 수 있다. 하나 이상의 오브젝트 인식기들은 특정 특징들을 갖는 오브젝트에 대해 전문화될 수 있다. 예컨대, 오브젝트 인식기(708a)는 얼굴들을 인식하는 데 사용될 수 있는 반면, 다른 오브젝트 인식기는 문서들을 인식하는 데 사용될 수 있다.

[0085] [0104] 오브젝트 인식들은 다양한 컴퓨터 비전 기법들을 사용하여 수행될 수 있다. 예컨대, 웨어러블 시스템은 장면 재구성, 이벤트 검출, 비디오 추적, 오브젝트 인식(예컨대, 사람들 또는 문서들), 오브젝트 포즈 추정, (예컨대, 환경 내의 사람 또는 문서 상의 이미지로부터의) 얼굴 인식, 학습, 인덱싱, 모션 추정, 또는 이미지 분석(예컨대, 문서들 내의 표시들, 이를테면 사진들, 서명들, 식별 정보, 여행 정보 등을 식별함) 등을 수행하도록 외향 이미징 시스템(464)(도 4에 도시됨)에 의해 획득된 이미지들을 분석할 수 있다. 하나 이상의 컴퓨터 비전 알고리즘들이 이러한 임무들을 수행하기 위해 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) 기법들, 순차적 베이지안 추정기(예컨대, 칼만 필터, 확장된 칼만 필터 등), 번들 조정, 적응형 임계화(Adaptive thresholding)(및 다른 임계화 기법들), ICP(Iterative Closest Point), SGM(Semi Global Matching), SGBM(Semi Global Block Matching), 피처 포인트 히스토그램(Feature Point Histogram)들, 다양한 기계 학습 알고리즘들(이를테면, 예컨대, 지원 벡터 기계, k-가장 인근의 이웃 알고리즘, 나이브 베이즈(Naive Bayes), 뉴럴 네트워크(콘볼루셔널 또는 딥 뉴럴 네트워크들을 포함함) 또는 다른 감독되는/감독되지 않는 모델들 등) 등을 포함한다.

[0086] [0105] 하나 이상의 오브젝트 인식기들(708)은 또한, 이미지들로부터 텍스트를 식별 및 추출하기 위해 다양한 텍스트 인식 알고리즘들을 구현할 수 있다. 일부 예시적인 텍스트 인식 알고리즘들은, OCR(optical character recognition) 알고리즘들, 심층 학습 알고리즘들(이를테면, 딥 뉴럴 네트워크들), 패턴 매칭 알고리즘들, 사전-프로세싱을 위한 알고리즘들 등을 포함한다.

[0087] [0106] 오브젝트 인식들은 추가적으로 또는 대안적으로, 다양한 기계 학습 알고리즘들에 의해 수행될 수 있다. 일단 트레이닝되면, 기계 학습 알고리즘은 HMD에 의해 저장될 수 있다. 기계 학습 알고리즘들의 일부 예들은, 회귀 알고리즘들(이를테면, 예컨대, 정규 최소 제곱 회귀(Ordinary Least Squares Regression)), 인스턴스-기반 알고리즘들(이를테면, 예컨대, 학습 벡터 양자화(Learning Vector Quantization)), 결정 트리 알고리즘들(이를테면, 예컨대, 분류 및 회귀 트리들), 베이지안 알고리즘들(이를테면, 예컨대, 나이브 베이즈), 클러스터링 알고리즘들(이를테면, 예컨대, k-평균 클러스터링), 연관 규칙 학습 알고리즘들(이를테면, 예컨대, 선험적 알고리즘들), 인공 뉴럴 네트워크 알고리즘들(이를테면, 예컨대, 퍼셉트론(Perceptron)), 심층 학습 알고리즘들(이를테면, 예컨대, 딥 볼츠만 기계, 또는 딥 뉴럴 네트워크), 차원 감소 알고리즘(이를테면, 예컨대, 주요한 컴포넌트 분석(Principal Component Analysis)), 앙상블 알고리즘들(이를테면, 예컨대, 스택된 일반화(Stacked Generalization)), 또는 다른 기계 학습 알고리즘들을 포함하는 감독되는 또는 감독되지 않는 기계 학습 알고리즘들을 포함할 수 있다. 일부 실시예들에서, 개별 모델들은 개별 데이터 세트들에 대해 맞춤화될 수 있다. 예컨대, 웨어러블 디바이스는 기본 모델을 생성하거나 저장할 수 있다. 기본 모델은 데이터 타입(예컨대, 텔레프레전스 세션의 특정 사용자), 데이터 세트(예컨대, 텔레프레전스 세션의 사용자의 획득된 한 세트의 추가적인 이미지들), 조건부 상황들, 또는 다른 변동들에 특정적인 추가적인 모델들을 생성하기 위한 시작 포인트로서 사용될 수 있다. 일부 실시예들에서, 웨어러블 HMD는 어그리게이팅된 데이터의 분석을 위한 모델들을 생성하기 위해 복수의 기법들을 활용하도록 구성될 수 있다. 다른 기법들은 미리 정의된 임계치들 또는 데이터 값들을

사용하는 것을 포함할 수 있다.

[0088] [0107] 맵 데이터베이스의 이러한 정보 및 포인트들의 모음에 기반하여, 오브젝트 인식기들(708a 내지 708n)은 오브젝트들을 인식하고, 오브젝트들에 생기(life)를 제공하기 위해 시맨틱 정보로 오브젝트들을 보충할 수 있다. 예컨대, 만약 오브젝트 인식기가 한 세트의 포인트들을 문이라고 인식하면, 시스템은 일부 시맨틱 정보를 부착할 수 있다(예컨대, 문은 힌지(hinge)를 갖고, 힌지를 중심으로 90도 움직임을 가짐). 만약 오브젝트 인식기가 한 세트의 포인트들을 미러라고 인식하면, 시스템은 미러는 룸 내의 오브젝트들의 이미지들을 반사할 수 있는 반사 표면을 갖는다는 시맨틱 정보를 부착할 수 있다. 시맨틱 정보는 본원에서 설명된 바와 같이 오브젝트들의 어포던스(affordance)들을 포함할 수 있다. 예컨대, 시맨틱 정보는 오브젝트의 법선을 포함할 수 있다. 시스템은 벡터를 할당할 수 있으며, 그의 방향은 오브젝트의 법선을 표시한다. 시간이 지남에 따라, 시스템(로컬로 상주할 수 있거나, 또는 무선 네트워크를 통해 액세스가능할 수 있음)이 세계로부터 더 많은 데이터를 누적함에 따라 맵 데이터베이스가 성장한다. 일단 오브젝트들이 인식되면, 정보는 하나 이상의 웨어러블 시스템들에 송신될 수 있다. 예컨대, MR 환경(700)은 캘리포니아에서 일어나는 장면에 관한 정보를 포함할 수 있다. 환경(700)은 뉴욕의 하나 이상의 사용자들에게 송신될 수 있다. FOV 카메라로부터 수신된 데이터 및 다른 입력들에 기반하여, 오브젝트 인식기들 및 다른 소프트웨어 컴포넌트들은 다양한 이미지들로부터 수집된 포인트들을 맵핑하고, 오브젝트들을 인식하는 등을 수행할 수 있어서, 장면이 세계의 상이한 부분에 있을 수 있는 제2 사용자에게 정확히 "전달"될 수 있다. 환경(700)은 또한 로컬화 목적들을 위해 토폴로지컬 맵을 사용할 수 있다.

[0089] [0108] 도 8은 인지된 오브젝트들에 관하여 가상 콘텐츠를 렌더링하는 방법(800)의 예의 프로세스 흐름도이다. 방법(800)은 가상 장면이 웨어러블 시스템의 사용자에게 어떻게 제공될 수 있는지를 설명한다. 사용자는 장면으로부터 지리적으로 원격일 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.

[0090] [0109] 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. 블록(810)에서, 사용자의 FOV 카메라, 센서들, GPS, 눈 추적 등은 시스템에 정보를 운반한다. 블록(820)에서, 시스템은 이러한 정보에 기반하여 희소 포인트들을 결정할 수 있다. 희소 포인트들은 사용자의 주변들의 다양한 오브젝트들의 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. 블록(830)에서, 오브젝트 인식기들(708a-708n)은 이러한 수집된 포인트들을 크롤링하고, 맵 데이터베이스를 사용하여 하나 이상의 오브젝트들을 인식할 수 있다. 그런 다음, 블록(840)에서, 이러한 정보는 사용자의 개별 웨어러블 시스템으로 운반될 수 있으며, 그에 따라서, 블록(850)에서, 원하는 가상 장면이 사용자에게 디스플레이될 수 있다. 예컨대, 원하는 가상 장면(예컨대, CA의 사용자)은 뉴욕의 사용자의 다양한 오브젝트들 및 다른 주변들에 관하여 적절한 배향, 포지션 등에서 디스플레이될 수 있다.

[0091] [0110] 도 9는 웨어러블 시스템의 다른 예의 블록 다이어그램이다. 이러한 예에서, 웨어러블 시스템(900)은 세계에 대한 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 액세스가능한 네트워킹된 저장 위치들에 부분적으로 상주할 수 있다. 포즈 프로세스(910)는 웨어러블 컴퓨팅 아키텍처(예컨대, 프로세싱 모듈(260) 또는 제어기(460)) 상에서 실행되고 맵(920)으로부터 데이터를 활용하여, 웨어러블 컴퓨팅 하드웨어 또는 사용자의 포지션 및 배향을 결정할 수 있다. 포즈 데이터는 사용자가 시스템을 경험하고 세계에서 동작할 때 즉시(on the fly) 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 오브젝트들에 관련된 표면 정보를 포함할 수 있다.

[0092] [0111] 희소 포인트 표현은 동시적인 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 파퓰레이팅(populate)하고 맵으로부터의 데이터를 사용하는 것을 포함해서, 다수의 목표들을 달성하는 빌딩 블록일 수 있다.

[0093] [0112] 일 실시예에서, 희소 포인트 포지션은 그 자체로 완전히 충분하지 않을 수 있고, 다초점 AR, VR 또는 MR 경험을 생성하기 위해 추가적인 정보가 필요할 수 있다. 일반적으로 깊이 맵 정보를 지칭하는 밀집된 표현

들이 이러한 갭을 적어도 부분적으로 채우기 위해 활용될 수 있다. 그러한 정보는 입체(940)로서 지칭되는 프로세스로부터 컴퓨팅될 수 있으며, 여기서 깊이 정보는 기법, 이를테면 삼각측량 또는 비행-시간 감지를 사용하여 결정된다. 이미지 정보 및 활성 패턴들(이를테면, 활성 투사기들을 사용하여 생성된 적외선 패턴들), 이미지 카메라들로부터 획득된 이미지들, 또는 손 제스처들/토템(950)은 입체 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 정의가능한 표면들은 (예컨대, 큰 포인트 클라우드에 비해) 효율적일 수 있고, 게임 엔진들과 같은 다른 프로세싱 디바이스들에 이해가능한 입력들일 수 있다. 따라서, 입체 프로세스(940)의 출력(예컨대, 깊이 맵)은 융합 프로세스(930)에서 조합될 수 있다. 포즈(910)는 또한 이 융합 프로세스(930)에 대한 입력일 수 있고, 융합(930)의 출력은 맵 프로세스(920)를 파퓰레이팅하기 위한 입력이 된다. 서브-표면들이 이를테면, 토폴리지 맵핑에서 서로 연결되어 더 큰 표면들을 형성할 수 있고, 맵은 포인트들 및 표면들의 큰 하이브리드가 된다.

[0094] [0113] 혼합 현실 프로세스(960)에서의 다양한 양상들을 해결하기 위해, 다양한 입력들이 활용될 수 있다. 예컨대, 도 9에 묘사된 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 오브젝트들 등을 갖는 몬스터 전투 게임을 플레이하고 있다고 결정하기 위한 입력일 수 있다. 세계 맵은 오브젝트들의 위치에 관한 정보 또는 오브젝트들의 시맨틱 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.

[0095] [0114] 사용자로부터의 제어들 또는 입력들은 웨어러블 시스템(900)에 대한 다른 입력이다. 본원에서 설명된 바와 같이, 사용자 입력들은 시각 입력, 제스처들, 토템들, 오디오 입력, 감각 입력 등을 포함할 수 있다. 예컨대, 주위를 돌아다니거나 게임을 플레이하기 위해, 사용자는 자신이 하기를 원하는 것에 관해 웨어러블 시스템(900)에게 명령할 필요가 있을 수 있다. 단지 공간에서 자신을 움직이는 것 외에도, 활용될 수 있는 다양한 형태들의 사용자 제어들이 존재한다. 일 실시예에서, 토템(예컨대, 사용자 입력 디바이스) 또는 오브젝트, 이를테면, 장난감 총은 사용자에 의해 홀딩되고 시스템에 의해 추적될 수 있다. 시스템은 바람직하게는, 사용자가 아이템을 홀딩하고 있다는 것을 알고 사용자가 아이템과 어떤 종류의 상호작용을 하고 있는지를 이해하도록 구성될 것이다(예컨대, 만약 토템 또는 오브젝트가 총이면, 시스템은 위치 및 배향뿐만 아니라, 센서, 이를테면 IMU가 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 때라도 무슨 일이 일어나고 있는지를 결정하는 데 도움을 줄 수 있음).

[0096] [0115] 손 제스처 추적 또는 인식은 또한 입력 정보를 제공할 수 있다. 웨어러블 시스템(900)은 왼쪽 또는 오른쪽, 정지, 잡기, 홀드 등을 제스처링하기 위해 버튼 누름들에 대한 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 수 있다. 손 제스처 추적은, 사용자의 환경 내의 다른 사람들, 이를테면 수화를 이용하여 통신하기 위해 제스처들을 수행하는 다른 사람들에 의해 수행된 제스처들을 추적하는 것을 포함할 수 있다(예컨대, 도 13a 참조).

[0097] [0116] 눈 추적(예컨대, 특정한 깊이 또는 범위로 렌더링하도록 디스플레이 기술을 제어하기 위해 사용자가 바라보는 곳을 추적함)은 다른 입력이다. 일 실시예에서, 눈들의 이접운동은 삼각측량을 사용하여 결정될 수 있으며, 그런 다음, 그 특정 사람을 위해 개발된 이접운동/원근조절 모델을 사용하여, 원근조절이 결정될 수 있다. 눈 추적은, 눈 시선(예컨대, 한쪽 또는 양쪽 눈들의 방향 또는 배향)을 결정하도록 눈 카메라(들)에 의해 수행될 수 있다. 눈 추적, 이를테면, 예컨대 눈(들) 인근에 배치된 전극들에 의한 전기 전위들의 측정(예컨대, 안구전도(electrooculography))을 위한 다른 기법들이 사용될 수 있다.

[0098] [0117] 스피치 추적은 단독으로 또는 다른 입력들(예컨대, 토템 추적, 눈 추적, 제스처 추적 등)과 조합하여 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 착용자로부터의 것인지 또는 다른 사람으로부터의 것인지 또는 음성(예컨대, 환경 내의 라우드스피커에 의해 송신된 레코딩된 음성)으로부터

것인지)를 결정하기 위한 음성 인식 기술뿐만 아니라 무엇을 말하고 있는지를 결정하기 위한 스피치 인식 기술을 통합할 수 있다. 로컬 데이터 및 프로세싱 모듈(260) 또는 원격 프로세싱 모듈(270)은, 다양한 스피치 인식 알고리즘들, 이를테면, 예컨대, 숨겨진 마르코프 모델(hidden Markov model)들, DTW(dynamic time warping)-기반 스피치 인식들, 뉴럴 네트워크들, 심층 학습 알고리즘들, 이를테면, 딥 피드포워드 및 회귀 뉴럴 네트워크들, 단-대-단(end-to-end) 자동 스피치 인식들, (도 7을 참조하여 설명된) 기계 학습 알고리즘들, 또는 음향 모델링 또는 언어 모델링을 사용하는 다른 알고리즘들 등을 적용함으로써 스피치의 콘텐츠를 식별하기 위해 마이크로폰으로부터의 오디오 데이터(또는 다른 스트림의 오디오 데이터, 이를테면, 사용자에 의해 시청되는 비디오 스트림)를 프로세싱할 수 있다.

[0099] [0118] 혼합 현실 프로세스(960)에 대한 다른 입력은 환경에서 사이니지를 추적하는 것을 포함할 수 있다. 사이니지는 상용 또는 공용 디스플레이 싸인들을 포함할 수 있다. 도 16a-19를 참조하여 설명된 바와 같이, 시스템은, 사이니지를 인식하고, 사이니지 내의 텍스트를 식별하고, 텍스트의 특징들을 조정(예컨대, 가독성을 개선하기 위해 텍스트의 폰트 사이즈를 증가)하고, 텍스트의 콘텐츠를 수정(예컨대, 외국어로부터 사용자에 의해 이해되는 언어로 텍스트를 변환)하는 등을 수행할 수 있다.

[0100] [0119] 로컬 데이터 및 프로세싱 모듈(260) 또는 원격 프로세싱 모듈(270)은 또한, 화자의 아이덴티티, 이를테면, 화자가 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, 가우시안 혼합 모델들, 패턴 매칭 알고리즘들, 뉴럴 네트워크들, 매트릭스 표현, 벡터 양자화, 화자 분할, 결정 트리들, 및 DTW(dynamic time warping) 기법을 포함할 수 있다. 음성 인식 기법들은 또한, 반-화자(anti-speaker) 기법들, 이를테면 코호트 모델들 및 세계 모델들을 포함할 수 있다. 스펙트럼 특징들은 화자 특징들을 표현하는 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 기계 학습 알고리즘들을 사용할 수 있다.

[0101] [0120] 시스템(900)은 또한, 다른 사람들 또는 환경과의 사용자의 상호작용들을 가능하게 하기 위한 감각 안경류 시스템(970)을 포함할 수 있다. 감각 안경류 시스템(970)의 구현은 UI를 통한 이러한 사용자 제어들 또는 입력들을 사용할 수 있다. UI 엘리먼트들(예컨대, 제어들, 팝업 윈도우들, 버블들, 데이터 입력 필드들 등)은, 예컨대, 정보, 예컨대 변환된 텍스트, 그래픽, 또는 보조 정보의 디스플레이를 취소하거나 또는 보조 정보의 디스플레이를 요청하기 위해 사용될 수 있다. UI는, 사용자가 이해하는 하나 이상의 언어들의 리스트를 사용자가 입력하게 허가할 수 있어서, 감각 안경류 시스템(970)은 수화에서 대화 파트너에 의해 수행된 싸인들을 변환하는 데 어느 언어를 사용할지를 알게 된다. 그러한 구현들 및 이러한 사용들의 예들은 아래에서 추가로 설명된다.

[0102] [0121] 감각 안경류 시스템(970)은 또한, 텍스트 인식, 수정, 및 렌더링 특징들을 포함할 수 있다. 그러한 특징들은 환경과의 사용자의 상호작용들을 향상시키기 위해 웨어러블 시스템의 다양한 다른 컴포넌트들과 조합될 수 있다. 예컨대, HMD는, 사용자의 물리적 환경의 이미지로부터 식별된 텍스트에 기반하여 (예컨대, 투사된 이미지가 물리적 환경으로부터의 오리지널 텍스트를 가리도록) 이미지를 디스플레이 상으로 투사하도록 구성된 하나 이상의 광원들(11)을 포함할 수 있다. 광학적으로 투과성인 접안레즈(106)는, 하나 이상의 광원들(11)로부터 사용자(210)로 광을 이미지로서 투과하도록 구성될 수 있다. 이미지는 그 이미지가 특정 깊이에 있는 것처럼 나타날 수 있으며, 그 깊이는 HMD 시스템(200)이 이미지를 디스플레이할 수 있는 많은 가능한 깊이들 중 단지 하나일 수 있다. HMD 시스템(100)은 다수의 상이한 깊이들에서 나타나도록 이미지들을 투사할 수 있으며, 그 이미지들은 상이한 깊이 평면들(306)(도 3 참조) 상에 있는 것처럼 나타날 수 있다. 접안레즈(106)가 광학적으로 투과성인 일부 실시예들에서, 접안레즈(106)는 환경으로부터의 광이 사용자의 눈에 들어가게 허용할 수 있다. 따라서, 그러한 실시예들에서, 사용자(210)는 하나 이상의 광원들(11)로부터의 투사된 이미지들과 함께 환경으로부터의 이미지의 일부들을 볼 수 있다.

[0103] [0122] 카메라 시스템들과 관련하여, 도 9에 도시된 예시적인 웨어러블 시스템(900)은 카메라들의 3개의 쌍들, 즉 사용자의 얼굴의 측면들에 대해 배열되는 비교적 넓은 FOV 또는 수동 SLAM 쌍의 카메라들, 입체 이미징 프로세스(940)를 다루기 위해 그리고 또한, 사용자의 얼굴의 앞에서 추적되는 손 제스처들 및 토템/오브젝트를 캡처하기 위해 사용자의 앞에 배향된 카메라들의 상이한 쌍을 포함할 수 있다. 입체 프로세스(940)를 위한 FOV 카메라들 또는 카메라들의 쌍은 또한 카메라들(16)로서 지칭될 수 있다. 입체 프로세스(940)를 위한 FOV 카메라들 및 카메라들의 쌍은 외향 이미징 시스템(464)(도 4에 도시됨)의 일부일 수 있다. 웨어러블 시스템(900)은 눈 벡터들 및 다른 정보를 삼각측량하기 위해 사용자의 눈들을 향해 배향되는 눈 추적 카메라들(또한 눈 카메라들(24)로서 도시되었으며, 도 4에 도시된 내향 이미징 시스템(462)의 일부일 수 있음)을 포함할 수 있다. 웨어

러블 시스템(900)은 또한 장면에 텍스처를 주입하기 위해 하나 이상의 텍스처링된 광 투사기들(이를테면, IR(infrared) 투사기들)을 포함할 수 있다.

[0104] [0123] 도 10은 웨어러블 시스템에 대한 사용자 입력을 결정하기 위한 방법(1000)의 예의 프로세스 흐름도이다. 이러한 예에서, 사용자는 토템과 상호작용할 수 있다. 사용자는 다수의 토템들을 가질 수 있다. 예컨대, 사용자는 소셜 미디어 애플리케이션에 대해 하나의 토템, 게임들을 플레이하기 위해 다른 토템 등을 지정할 수 있다. 블록(1010)에서, 웨어러블 시스템은 토템의 모션을 검출할 수 있다. 토템의 움직임은 외향 이미징 시스템을 통해 인식될 수 있거나, 또는 센서들(예컨대, 햅틱 글러브, 이미지 센서들, 손 추적 디바이스들, 눈-추적 카메라들, 머리 포즈 센서들 등)을 통해 검출될 수 있다.

[0105] [0124] 블록(1020)에서, 토템을 통한 검출된 제스처, 눈 포즈, 머리 포즈, 또는 입력에 적어도 부분적으로 기반하여, 웨어러블 시스템은 참조 프레임에 대하여 토템(또는 사용자의 눈들 또는 머리 또는 제스처들)의 포지션, 배향, 또는 움직임을 검출한다. 레퍼런스 프레임은 한 세트의 맵 포인트들일 수 있으며, 그 세트에 기반하여, 웨어러블 시스템은 토템(또는 사용자)의 움직임을 액션 또는 커맨드로 변환한다. 블록(1030)에서, 토템과의 사용자의 상호작용이 맵핑된다. 레퍼런스 프레임(1020)에 대한 사용자 상호작용의 맵핑에 기반하여, 블록(1040)에서 시스템은 사용자 입력을 결정한다.

[0106] [0125] 예컨대, 사용자는 가상 페이지를 넘기고(turning) 다음 페이지로 이동하거나 또는 하나의 UI(user interface) 디스플레이 스크린으로부터 다른 UI 스크린으로 이동하는 것을 나타내기 위해 토템 또는 물리적 오브젝트를 앞뒤로 이동시킬 수 있다. 다른 예로서, 사용자는 사용자의 FOR에서 상이한 실제 또는 가상 오브젝트들을 보기 위해 그들의 머리 또는 눈들을 이동시킬 수 있다. 만약 특정 실제 또는 가상 오브젝트에서의 사용자의 응시가 임계 시간보다 길면, 실제 또는 가상 오브젝트는 사용자 입력으로서 선택될 수 있다. 일부 구현들에서, 사용자의 눈들의 이접운동이 추적될 수 있으며, 원근조절/이접운동 모델은, 사용자가 초점을 맞추고 있는 깊이 평면에 대한 정보를 제공하는 사용자의 눈들의 원근조절 상태를 결정하기 위해 사용될 수 있다. 일부 구현들에서, 웨어러블 시스템은 어떤 실제 또는 가상 오브젝트들이 사용자의 머리 포즈 또는 눈 포즈의 방향을 따라 존재하는지를 결정하기 위해 광선 캐스팅 기법들을 사용할 수 있다. 다양한 구현들에서, 광선 캐스팅 기법들은, 실질적으로 가로 폭이 거의 없는 얇은 광속 광선들을 캐스팅하는 것 또는 실질적인 가로 폭을 갖는 광선들(예컨대, 원뿔들 또는 절두체들)을 캐스팅하는 것을 포함할 수 있다.

[0107] [0126] 사용자 인터페이스는 본원에서 설명된 바와 같이 디스플레이 시스템(이를테면, 도 2a의 디스플레이(220))에 의해 투사될 수 있다. 그것은 또한, 다양한 다른 기법들, 이를테면 하나 이상의 투사기들을 사용하여 디스플레이될 수 있다. 투사기들은 물리적 오브젝트, 이를테면 캔버스 또는 구체 상으로 이미지들을 투사할 수 있다. 사용자 인터페이스와의 상호작용들은 시스템 또는 시스템의 일부 외부의 하나 이상의 카메라들을 사용하여 (이를테면, 예컨대 내향 이미징 시스템(462) 또는 외향 이미징 시스템(464)을 사용하여) 추적될 수 있다.

[0108] [0127] 도 11은 가상 사용자 인터페이스와 상호작용하기 위한 방법(1100)의 예의 프로세스 흐름도이다. 방법(1100)은 본원에서 설명된 웨어러블 시스템에 의해 수행될 수 있다. 방법(1100)의 실시예들은 웨어러블 시스템의 FOV에서 사람들 또는 문서들을 검출하도록 웨어러블 시스템에 의해 사용될 수 있다.

[0109] [0128] 블록(1110)에서, 웨어러블 시스템은 특정 UI를 식별할 수 있다. UI의 타입은 사용자에 의해 미리 결정될 수 있다. 웨어러블 시스템은, 특정 UI가 사용자 입력(예컨대, 제스처, 시각 데이터, 오디오 데이터, 감각 데이터, 지향 커맨드 등)에 기반하여 파퓰레이팅될 필요가 있다는 것을 식별할 수 있다. UI는 (예컨대, 여행 체크포인트에서) 시스템의 착용자가 착용자에게 문서들을 제공하는 사용자들을 관찰하고 있는 보안 시나리오에 특정적일 수 있다. 블록(1120)에서, 웨어러블 시스템은 가상 UI에 대한 데이터를 생성할 수 있다. 예컨대, UI의 한계들, 일반적인 구조, 형상 등과 연관된 데이터가 생성될 수 있다. 게다가, 웨어러블 시스템이 사용자의 물리적 위치에 관하여 UI를 디스플레이할 수 있도록 웨어러블 시스템은 사용자의 물리적 위치의 맵 좌표들을 결정할 수 있다. 예컨대, 만약 UI가 신체 중심이라면, 웨어러블 시스템은, 링 UI가 사용자 주위에서 디스플레이될 수 있거나 또는 평면 UI가 벽 상에 또는 사용자의 앞에 디스플레이될 수 있도록 사용자의 물리적 자세, 머리 포즈, 또는 눈 포즈의 좌표들을 결정할 수 있다. 본원에서 설명된 보안 상황에서, 시스템의 착용자에게 문서들을 제공하는 여행자를 UI가 둘러싼 것처럼 UI가 디스플레이될 수 있어서, 착용자는 여행자 및 여행자의 문서들을 보면서 UI를 쉽게 볼 수 있다. 만약 UI가 손 중심이라면, 사용자의 손들의 맵 좌표들이 결정될 수 있다. 이들 맵 포인트들은 FOV 카메라들을 통해 수신된 데이터, 감각 입력, 또는 임의의 다른 타입의 수집된 데이터를 통해 도출될 수 있다.

[0110] [0129] 블록(1130)에서, 웨어러블 시스템은 클라우드로부터 디스플레이로 데이터를 전송할 수 있거나 또는 데

이터는 로컬 데이터베이스로부터 디스플레이 컴포넌트들로 전송될 수 있다. 블록(1140)에서, UI는 전송된 데이터에 기반하여 사용자에게 디스플레이된다. 예컨대, 광 필드 디스플레이는 가상 UI를 사용자의 눈들 중 하나 또는 둘 다로 투사될 수 있다. 일단 가상 UI가 생성되면, 블록(1150)에서 웨어러블 시스템은 단순히 가상 UI 상에 더 많은 가상 콘텐츠를 생성하기 위해 사용자로부터의 커맨드를 기다릴 수 있다. 예컨대, UI는 사용자의 신체 또는 사용자의 환경 내의 사람(예컨대, 여행자)의 신체 주위의 신체 중심 링일 수 있다. 그런 다음, 웨어러블 시스템은 커맨드(제스처, 머리 또는 눈 움직임, 음성 커맨드, 사용자 입력 디바이스로부터의 입력 등)를 기다릴 수 있으며, 만약 그것이 인지되면(블록(1160)), 커맨드와 연관된 가상 콘텐츠가 사용자에게 디스플레이될 수 있다(블록(1170)).

[0111] [0130] 웨어러블 시스템들, UI, 및 UX(user experiences)의 추가적인 예들은 미국 특허 공보 제2015/0016777호에 설명되며, 그 공보는 그 전체가 본원에서 인용에 의해 통합된다.

[0112] 다수의 웨어러블 시스템들 간의 예시적인 통신들

[0113] [0131] 도 12는 서로 상호작용하는 다수의 사용자 디바이스들을 묘사하는 전체 시스템 도면을 개략적으로 예시한다. 컴퓨팅 환경(1200)은 사용자 디바이스들(1230a, 1230b, 1230c)을 포함한다. 사용자 디바이스들(1230a, 1230b, 및 1230c)은 네트워크(1290)를 통해 서로 통신할 수 있다. 사용자 디바이스들(1230a-1230c)은 각각, (네트워크 인터페이스(1271)를 또한 포함할 수 있는) 원격 컴퓨팅 시스템(1220)과 네트워크(1290)를 통해 통신하기 위한 네트워크 인터페이스를 포함할 수 있다. 네트워크(1290)는 LAN, WAN, 피어-투-피어 네트워크, 라디오, 블루투스, 또는 임의의 다른 네트워크일 수 있다. 컴퓨팅 환경(1200)은 또한, 하나 이상의 원격 컴퓨팅 시스템들(1220)을 포함할 수 있다. 원격 컴퓨팅 시스템(1220)은 상이한 지리적 위치들에서 클러스터링되고 위치되는 서버 컴퓨터 시스템들을 포함할 수 있다. 사용자 디바이스들(1230a, 1230b, 및 1230c)은 네트워크(1290)를 통해 원격 컴퓨팅 시스템(1220)과 통신할 수 있다.

[0114] [0132] 원격 컴퓨팅 시스템(1220)은 특정한 사용자의 물리적 또는 가상 세계들에 관한 정보를 유지할 수 있는 원격 데이터 저장소(1280)를 포함할 수 있다. 데이터 스토리지(1280)는 감각 안경류에 유용한 정보, 이를테면 수화 사전, 보조 정보 소스 등을 포함할 수 있다. 원격 데이터 저장소는 도 2a에 도시된 원격 데이터 저장소(280)의 실시예일 수 있다. 원격 컴퓨팅 시스템(1220)은 또한, 원격 프로세싱 모듈(1270)을 포함할 수 있다. 원격 프로세싱 모듈(1270)은 도 2a에 도시된 원격 프로세싱 모듈(270)의 실시예일 수 있다. 일부 구현들에서, 원격 컴퓨팅 시스템(1220)은 웨어러블 시스템(200)과는 독립적인 제3 자 시스템일 수 있다.

[0115] [0133] 원격 프로세싱 모듈(1270)은 사용자 디바이스들(1230a, 1230b, 1230c) 및 원격 데이터 저장소(1280)와 통신할 수 있는 하나 이상의 프로세서들을 포함할 수 있다. 프로세서들은 사용자 디바이스들 및 다른 소스들로부터 획득된 정보를 프로세싱할 수 있다. 일부 구현들에서, 프로세싱 또는 스토리지의 적어도 일부는 (도 2a에 도시된 바와 같은) 로컬 프로세싱 및 데이터 모듈(260)에 의해 제공될 수 있다. 원격 컴퓨팅 시스템(1220)은 정해진 사용자가 특정한 사용자 본인의 물리적 또는 가상 세계들에 관한 정보를 다른 사용자와 공유할 수 있게 할 수 있다.

[0116] [0134] 사용자 디바이스는 단독으로 또는 조합하여, 웨어러블 디바이스(이를테면, HMD 또는 ARD), 컴퓨터, 모바일 디바이스, 또는 임의의 다른 디바이스들일 수 있다. 예컨대, 사용자 디바이스들(1230b 및 1230c)은, AR/VR/MR 콘텐츠를 제공하도록 구성될 수 있는, 도 2a에 도시된 웨어러블 시스템(200)(또는 도 4에 도시된 웨어러블 시스템(400))의 실시예일 수 있다.

[0117] [0135] 사용자 디바이스들 중 하나 이상은 도 4에 도시된 사용자 입력 디바이스(466)와 함께 사용될 수 있다. 사용자 디바이스는 (예컨대, 도 4에 도시된 외향 이미징 시스템(464)을 사용하여) 사용자 및 사용자의 환경에 관한 정보를 획득할 수 있다. 사용자 디바이스 또는 원격 컴퓨팅 시스템(1220)은 사용자 디바이스들로부터 획득된 정보를 사용하여 이미지들, 포인트들 및 다른 정보의 모음을 구성, 업데이트, 및 구축할 수 있다. 예컨대, 사용자 디바이스는 획득된 원시 정보를 프로세싱하고, 추가적인 프로세싱을 위해 원격 컴퓨팅 시스템(1220)에 프로세싱된 정보를 전송할 수 있다. 사용자 디바이스는 또한, 프로세싱을 위해 원시 정보를 원격 컴퓨팅 시스템(1220)에 전송할 수 있다. 사용자 디바이스는 원격 컴퓨팅 시스템(1220)으로부터 프로세싱된 정보를 수신하고, 사용자로 투사하기 전에 최종 프로세싱을 제공할 수 있다. 사용자 디바이스는 또한, 획득된 정보를 프로세싱하고, 프로세싱된 정보를 다른 사용자 디바이스들에 전달할 수 있다. 사용자 디바이스는 원격 데이터 저장소(1280)와 통신하면서, 획득된 정보를 프로세싱할 수 있다. 다수의 사용자 디바이스들 또는 다수의 서버 컴퓨터 시스템들은 획득된 이미지들의 구성 또는 프로세싱에 참여할 수 있다.

[0118] [0136] 물리적 세계들에 대한 정보는 시간이 지남에 따라 개발될 수 있으며, 상이한 사용자 디바이스들에 의해 수집된 정보에 기반할 수 있다. 가상 세계들의 모델들은 또한, 시간이 지남에 따라 개발되고 상이한 사용자들의 입력들에 기반할 수 있다. 그러한 정보 및 모델들은 종종, 세계 맵 또는 세계 모델로서 본원에서 지칭될 것이다. 도 7 및 9를 참조하여 설명된 바와 같이, 사용자 디바이스들에 의해 획득된 정보는 세계 맵(1210)을 구성하기 위해 사용될 수 있다. 세계 맵(1210)은 도 9에 설명된 맵(920)의 적어도 일부를 포함할 수 있다. 다양한 오브젝트 인식기들(예컨대, 708a, 708b, 708c … 708n)은 오브젝트들 및 태그 이미지들을 인식할 뿐만 아니라 시맨틱 정보를 오브젝트들에 부착하기 위해 사용될 수 있다. 이러한 오브젝트 인식기들이 또한 도 7에서 설명된다.

[0119] [0137] 원격 데이터 저장소(1280)는 데이터를 저장하고 세계 맵(1210)의 구성을 가능하게 하기 위해 사용될 수 있다. 사용자 디바이스는 지속적으로 사용자의 환경에 관한 정보를 업데이트하고 세계 맵(1210)에 관한 정보를 수신할 수 있다. 세계 맵(1210)은 사용자에 의해 또는 다른 누군가에 의해 생성될 수 있다. 본원에서 논의된 바와 같이, 사용자 디바이스들(예컨대, 1230a, 1230b, 1230c) 및 원격 컴퓨팅 시스템(1220)은 단독으로 또는 조합하여, 세계 맵(1210)을 구성 또는 업데이트할 수 있다. 예컨대, 사용자 디바이스는 원격 프로세싱 모듈(1270) 및 원격 데이터 저장소(1280)와 통신할 수 있다. 사용자 디바이스는 사용자 및 사용자의 환경에 관한 정보를 획득 및/또는 프로세싱할 수 있다. 원격 프로세싱 모듈(1270)은 사용자 및 사용자의 환경에 관한 정보를 프로세싱하기 위해 원격 데이터 저장소(1280) 및 사용자 디바이스들(예컨대, 1230a, 1230b, 1230c)과 통신할 수 있다. 원격 컴퓨팅 시스템(1220)은 사용자 디바이스들(예컨대, 1230a, 1230b, 1230c)에 의해 획득된 정보를 수정할 수 있으며, 이를테면, 예컨대, 사용자의 이미지를 선택적으로 크롭(crop)하거나, 사용자의 배경을 수정하거나, 사용자의 환경에 가상 오브젝트들을 추가하거나, 보조 정보로 사용자의 스피치에 주석을 다는 등을 수행한다. 원격 컴퓨팅 시스템(1220)은 프로세싱된 정보를 동일한 또는 상이한 사용자 디바이스들에 전송할 수 있다.

[0120] [0138] 감각 안경류 시스템의 실시예들의 다양한 기능성들이 아래에서 추가로 설명된다.

[0121] 사용자 상호작용들을 가능하게 하기 위한 예시적인 감각 안경류

[0122] [0139] 웨어러블 시스템(200)은 다른 사람들과 또는 환경과의 사용자의 상호작용들을 가능하게 하기 위한 감각 안경류 시스템(970)을 구현할 수 있다. 다른 사람들과 상호작용하는 일례로서, 웨어러블 시스템(200)은 예컨대 수화를 구성할 수 있는 제스처들을 검출하고, 수화를 다른 언어(예컨대, 다른 수화 또는 발화된 언어)로 변환하며, 그리고 변환된 정보를 웨어러블 디바이스의 사용자에게 제공함으로써 수화를 해석할 수 있다. 다른 예로서, 감각 안경류 시스템(970)은 스피치를 수화로 변환하고 수화를 사용자에게 제공할 수 있다.

[0123] [0140] 웨어러블 시스템(970)은 또한 환경에서 오브젝트들을 인식하고, (가상 환경에서) 오브젝트들의 특징들을 수정하며 그리고 수정된 오브젝트들을 가상 오브젝트들로서 사용자에게 제공함으로써 환경과 사용자의 상호작용을 가능하게 할 수 있다. 예컨대, 웨어러블 시스템(200)은 외향 이미징 시스템(464)에 의해 획득된 이미지들에 기반하여 사용자의 환경에서 싸인들(예컨대, 교통 표지판들, 상점에 대한 싸인들 등)을 인식하고, 사용자의 환경에서 싸인의 특징들을 수정하며 그리고 수정된 싸인을 사용자에게 제공할 수 있다. 수정된 싸인은 오리지널 싸인이 가려질 수 있도록 사용자의 3D 환경 상에 오버레이될 수 있다.

[0124] 개인간 통신들을 위한 도구로서의 예시적인 감각 안경류 시스템

[0125] [0141] 일부 상황들에서, 대화중인 하나 이상의 사람들은 그 자신들을 표현하기 위하여 손 또는 신체 제스처들(이를테면, 예컨대 수화)을 사용할 수 있다. 대화는 텔레프레전스 세션 동안 또는 사람들이 서로 물리적으로 근접해 있을 때 발생할 수 있다. 웨어러블 시스템(200)은 (관찰자로서 또한 지칭되는) 웨어러블 시스템(200)의 사용자가 수화 사용자와 통신할 때 그 사용자에 대한 그 수화 사용자의 수화를 해석할 수 있다. 웨어러블 시스템(200)은 또한 구두 또는 수화 기반 스피치를 그래픽들(이를테면, 예컨대 손 제스처들의 이미지들)로 변환하고, 수화 사용자가 관찰자의 스피치를 이해할 수 있도록 수화 사용자에게 그래픽들을 제공할 수 있다. 예컨대, 머리-장착 디스플레이를 착용한 관찰자는 감소된 시야를 가질 수 있으며, 이에 따라 관찰자는 수화 사용자가 수화를 사용하여 행한 완벽한 제스처들을 관찰하지 못할 가능성이 있을 수 있다. 웨어러블 시스템(200)은 외향 이미징 시스템(464)을 사용하여 수화 사용자에 의한 제스처들을 캡처할 수 있다(왜냐하면, 외향 이미징 시스템(464)은 사용자가 머리-장착 디스플레이를 통해 인식할 수 있는 넓은 시야를 가진 카메라를 가질 수 있기 때문이다). 웨어러블 시스템(200)은 캡처된 제스처들을 가상 그래픽들로서 관찰자에게 보여주거나 또는 관찰자가 수화 사용자 스피치를 이해하는 것을 가능하게 하기 위하여 캡처된 제스처들로부터 변환된 텍스트를 보여줄 수 있다. 추가로, 웨어러블 시스템(200)은 하나의 수화를 다른 수화로 변환하도록 구성될 수 있다. 예

컨대, 대화중인 한 사람은 미국 수화를 사용할 수 있고, 다른 사람은 도곤족 수화를 사용할 수 있다. 웨어러블 시스템(200)은 도곤족 수화를 사용하는 사람의 경우 미국 수화를 도곤족 수화로 변환할 수 있으며, 미국 수화를 사용하는 사람의 경우 도곤족 수화를 미국 수화로 변환할 수 있다.

[0126] 예시적인 수화 캡처링

[0127] [0142] 웨어러블 시스템은 오리지널 스피치를 캡처하고 오리지널 스피치를 타겟 스피치로 변환하기 위해 다양한 기법들을 사용할 수 있다. 스피치는 손 또는 신체 제스처들 또는 가청음들의 형태일 수 있다. 본원에서 설명된 바와 같이, 오리지널 스피치는 수화일 수 있으며, 타겟 스피치는 다른 수화 또는 발화된 언어일 수 있다. 대안적으로, 오리지널 스피치는 발화된 언어일 수 있는 반면에, 타겟 스피치는 수화일 수 있다. 웨어러블 시스템(200)은 외향 이미징 시스템(464), 오디오 센서(232)를 사용하여 오리지널 스피치를 캡처할 수 있거나 또는 스피치의 맥락(예컨대, 스피치가 대면적인지(in-person) 또는 원격통신을 통한 것인지 여부)에 따라 네트워크(1290)를 통해 다른 컴퓨팅 디바이스와 통신함으로써 오리지널 스피치를 캡처할 수 있다.

[0128] [0143] 대면적인 통신 동안 오리지널 스피치를 캡처하는 예로서, 검출된 수화의 수화 사용자가 감각 안경류 시스템에 물리적으로 근접한 경우에, 외향 이미징 시스템(464)은 사용자의 환경의 이미지들을 캡처할 수 있다. 웨어러블 시스템(200)은 이미지 정보로부터, 수화를 구성할 수 있는 제스처들(예컨대, 손/신체 제스처들 또는 입술 움직임들)을 검출할 수 있다. 웨어러블 시스템(200)은 스피커에 의해 행해진 제스처들에 의해 표현되는 싸인들을 인식하기 위하여, 예컨대 딥 뉴럴 네트워크, 히든 마르코프 모델, 동적 프로그래밍 매칭 등과 같은 알고리즘들을 사용하여 제스처들을 인식할 수 있다. 도 7을 참조하여 설명된 바와 같이, 제스처 인식은 하나 이상의 오브젝트 인식기들(708)에 의해 수행될 수 있다.

[0129] [0144] 원격 통신들의 맥락에서 오리지널 스피치를 캡처하는 예로서, 웨어러블 시스템(200)은 원격 컴퓨팅 디바이스(예컨대, 다른 웨어러블 디바이스)로부터 수신된 데이터를 분석함으로써 또는 (예컨대, 오리지널 스피치가 텔레비전 상에서 제공되는 경우에) 외향 이미징 시스템(464)에 의해 캡처되는 데이터를 분석함으로써 오리지널 스피치의 존재를 캡처하여 인식할 수 있다. 일 예에서, 수화 사용자 및 관찰자는 인터넷 비디오 채팅 세션을 통해 대화할 수 있다. 수화 사용자 및 관찰자는 각각 그들의 개개의 HMD들을 착용할 수 있다. HMD들은 네트워크(1290)(도 12에 도시됨)를 통해 서로 통신할 수 있다. 수화 사용자가 반사 표면(예컨대, 미러) 앞에 있는 경우에, 수화 사용자의 HMD는 외향 이미징 시스템(464)을 통해 수화 사용자의 반사된 이미지들을 획득함으로써 수화 사용자의 제스처들을 캡처할 수 있다. 수화 사용자의 반사된 이미지들은 수화들의 인식 및 해석을 위하여 원격 컴퓨팅 시스템(1220) 또는 관찰자의 HMD에 전송될 수 있다. 다른 예로서, 수화 사용자는 텔레비전 또는 인터넷 프로그래밍 등에 제공된 것과 같은 비디오 프로그램의 사람일 수 있다. 수화 사용자의 제스처들이 관찰자의 위치에서 시각적으로 관찰될 수 있는 경우에, 웨어러블 시스템(464)은 (예컨대, 오디오 센서(232) 또는 외향 이미징 시스템(464)을 통해) 대면적인 통신 맥락에서 수화 제스처들을 캡처하는 것과 동일한 방식으로 수화 제스처들을 캡처할 수 있다.

[0130] [0145] 다른 사람에 의해서 웨어러블 시스템(200)의 사용자에게 제스처링된 수화의 텍스트 또는 그래픽 변환을 디스플레이하는 것에 부가하여 또는 이에 대한 대안으로, 웨어러블 시스템(200)의 사용자는 또한 수화로 통신할 수 있다. 이러한 경우에, 웨어러블 시스템은 외향 이미징 시스템(464)에 의해 (일인칭 관점에서) 사용자 그 자체의 수화 제스처들을 캡처할 수 있다. 웨어러블 시스템은 텍스트, 오디오, 이미지들 등의 포맷으로 표현될 수 있는 타겟 스피치로 수화를 변환할 수 있다. 웨어러블 시스템(200)은 다른 사용자에게 프리젠테이션하기 위한 결과를 다른 웨어러블 시스템에 송신할 수 있다. 본원에서 설명된 바와 같이, 오리지널 스피치로부터 타겟 스피치로의 변환은 사용자의 웨어러블 시스템, 다른 사용자의 웨어러블 시스템 또는 원격 컴퓨팅 시스템(1220)을 단독으로 또는 조합하여 사용함으로써 수행될 수 있다. 예컨대, 사용자의 웨어러블 시스템은 사용자의 손의 제스처들을 캡처하고, 캡처된 비디오 또는 이미지(수화 제스처들을 포함함)를 다른 사용자의 웨어러블 시스템 또는 원격 컴퓨팅 시스템(120)에 송신할 수 있으며, 다른 사용자의 웨어러블 시스템 또는 원격 컴퓨팅 시스템(120)은 비디오 또는 이미지로부터 수화를 추출하여, 그 수화를 말하는 언어 또는 다른 수화를 위한 오디오-시각 콘텐츠로 변환할 수 있다. 오디오-시각 콘텐츠는 텍스트, 그래픽, 비디오, 애니메이션들, 사운드 등을 포함할 수 있다.

[0131] 수화 사용자 제스처 거절 및 소스 로컬화

[0132] [0146] 웨어러블 시스템은 다양한 센서들, 이를테면, 예컨대 오디오 센서(232), 외향 이미징 시스템(464), 고정 입력(704) 또는 사용자 환경 내의 다른 센서들을 사용하여 제스처 또는 수화의 소스를 식별할 수 있다. 일 예로서, 웨어러블 시스템은 외향 이미징 시스템(464)에 의해 획득된 데이터로부터 일련의 손 제스처들 뿐만 아

니라 입술 움직임들을 검출할 수 있다. 웨어러블 시스템은 수화 사용자가 또한 대응하는 입술 움직임들을 가지기 때문에 손 제스처들이 수화 사용자에 연관됨을 발견할 수 있다. 다른 예로서, 웨어러블 시스템은 제스처들의 소스를 결정하기 위하여 사용자와 제스처들 간의 거리를 측정할 수 있다. 예컨대, 웨어러블 시스템은 외향 이미징 시스템(464)에 의해 획득된 이미지들에서 손들이 비교적 크게 나타나기 때문에 일련의 제스처들이 사용자로부터 유래한다는 것을 결정할 수 있다. 그러나, 만일 손들이 비교적 작게 나타나면, 웨어러블 시스템은 제스처들이 사용자가 아닌 사람으로부터 유래한다는 것을 발견할 수 있다. 또 다른 예로서, 웨어러블 시스템은 오디오-시각 콘텐츠를 플레이하고 있는 오브젝트를 인식함으로써 (예컨대, 오브젝트 인식기들(708)을 사용하여 텔레비전을 인식함으로써) 제스처들이 (예컨대, 텔레비전의) 오디오-시각 콘텐츠로부터 유래한다는 것을 발견할 수 있다.

[0133] [0147] 제스처들의 소스에 기반하여, 웨어러블 시스템(200)은 특정 사람들로부터의 제스처들을 프로세싱하지 않도록 구성될 수 있다. 예컨대, 웨어러블 시스템은 사용자의 환경의 다수의 사람들로부터의 제스처들을 캡처할 수 있으나, 웨어러블 시스템은 수화 인식을 위해 사용자의 FOV의 중심 외부 사람으로부터의 수화를 프로세싱하지 않도록 구성될 수 있다. 다른 예로서, 웨어러블 시스템은 사용자 그 자체의 수화를 프로세싱하지 않도록 구성될 수 있다.

[0134] [0148] 일부 실시예들에서, 웨어러블 시스템은, 이를테면 예컨대 외향 이미징 시스템(464)이 사용자의 손 제스처들을 캡처하도록 하기 위하여 사용자가 그의 손을 올릴 필요가 없도록 외향 이미징 시스템(464)의 카메라들을 일정 각도로 포지셔닝함으로써 사용자 그 자체의 수화를 검출하도록 센서들을 구성할 수 있다. 센서들은 또한 사용자 그 자체의 수화를 검출하지 않도록 구성될 수 있다. 예컨대, (전형적으로 사용자의 FOV 아래에 있는) 사용자 그 자체의 손들의 방향에서 이미지들을 캡처하지 않거나 또는 이 방향에서 이미지들을 (예컨대, 크로핑에 의해) 필터링함으로써 비-검출이 달성될 수 있다. 따라서, 시스템은 사용자 그 자체의 수화를 다른 사용자들의 수화들로부터 구별할 수 있다.

[0135] 수화로부터 텍스트로의 예시적인 변환

[0136] [0149] 웨어러블 시스템(200)은 캡처된 수화를 텍스트로 변환할 수 있으며, 이 텍스트는 사용자에게 제공되거나 또는 다른 언어로 변환될 수 있다. 수화의 텍스트로의 변환은 (딥 뉴럴 네트워크를 활용할 수 있는) 심층 학습, 히든 마르코프 모델, 동적 프로그래밍 매칭 등과 같은 알고리즘들을 사용하여 수행될 수 있다. 예컨대, 심층 학습 방법(일부 경우들에서 콘볼루셔널 뉴럴 네트워크)은 싸인들을 나타내는 피처들을 결정하고 학습된 피처들에 기반하여 분류 모델을 구축하기 위하여 알려진 싸인들(감독되는 학습)을 포함하는 이미지들 또는 비디오들에 대해 트레이닝될 수 있다. 그런 다음, 이러한 트레이닝된 심층 학습 방법은 웨어러블 시스템(200)의 로컬 프로세싱 및 데이터 모듈(260) 또는 원격 프로세싱 모듈 및 데이터 저장소(270, 280)에 의해, 외향 이미징 서브시스템에 의해 검출된 수화 사용자의 이미지들에 적용될 수 있다.

[0137] [0150] 텍스트 변환 기능성은 로컬 프로세싱 및 데이터 모듈(260), 원격 프로세싱 모듈(270), 원격 데이터 저장소(280) 또는 원격 컴퓨팅 시스템(1220)을 단독으로 또는 조합하여 사용함으로써 구현될 수 있다. 예컨대, 웨어러블 시스템(200)은 HMD 상에서 구현되는 수화-텍스트 기능성을 포함할 수 있다. 일 예로서, 웨어러블 시스템은 로컬 데이터 모듈(260) 또는 원격 데이터 저장소(280)에 수화 사전을 저장할 수 있다. 그에 따라서, 웨어러블 시스템은 검출된 제스처를 텍스트로 변환하기 위하여 수화 사전에 액세스할 수 있다. 다른 예로서, 웨어러블 시스템(200)은 원격 컴퓨팅 시스템(1220)에 의해 구현되는 수화-텍스트 기능성에 액세스할 수 있다. 웨어러블 시스템(200)은 (예컨대, API(application programming interface)를 통해) 상용 수화-텍스트 서비스들 또는 데이터 저장소들에 대해 무선 연결들을 활용할 수 있다. 예컨대, 웨어러블 시스템(200)은 원격 컴퓨팅 시스템(1220)에 캡처된 제스처를 제공하고, 원격 컴퓨팅 시스템(1220)으로부터 대응하는 텍스트를 수신할 수 있다.

[0138] [0151] 변환이 로컬로 수행되든 또는 원격으로 수행되든, 다른 프로세싱 단계들, 이를테면 변환된 텍스트를 디스플레이하는 단계 및 (아래에서 추가로 설명되는) 보조 정보를 리트리브하는 단계가 텍스트 변환이 수행되는 곳과 관계없이 로컬로 또는 원격으로 수행될 수 있다. 예컨대, 만약 수화-텍스트 변환이 원격으로 수행되고 변환된 텍스트가 로컬로 디스플레이된다면(예컨대, 시스템의 사용자가 관찰자이면), 캡처된 비디오 스트림은 네트워크를 통해 변환을 수행하는 원격 서버 또는 원격 프로세싱 모듈(270)로 전송될 수 있으며; 변환된 텍스트 스트링들은 디스플레이를 위해 시스템의 로컬 컴포넌트(예컨대, 로컬 프로세싱 및 데이터 모듈(260))로 리턴된다. 다른 예로서, 만약 수화-텍스트 변환 및 보조 정보 리트리벌이 원격으로 수행된다면, 캡처된 비디오 스트림은 네트워크를 통해 원격 서버에 또는 원격 프로세싱 모듈(270)에 전송될 수 있고, 리트리브된 보조 정보는 시스템

의 로컬 컴포넌트로 리턴될 수 있다. 로컬/원격 프로세싱의 다른 조합들이 또한 실행가능하다.

[0139] [0152] 이들 예들이 싸인을 텍스트로 변환하는 것을 참조하여 설명되었지만, 싸인들은 다양한 다른 포맷들, 이를테면, 예컨대 그래픽들, 애니메이션들, 오디오 또는 다른 타입들의 시청각(audio-visual) 콘텐츠로 변환될 수 있다. 추가로, 싸인들의 변환은 싸인들이 먼저 텍스트로 변환될 것을 요구하지 않는다.

[0140] 하나의 수화를 다른 수화로 변환하는 예들

[0141] [0153] 본원에서 언급되는 바와 같이, 전세계에는 수백 개의 수화들이 존재한다. 그에 따라서, 본원에서 설명된 웨어러블 시스템들은 또한, 대화 파트너들 둘 모두 싸인하고 있지만 상이한 수화 시스템들에서 싸인하고 있을 때 사용될 수 있다. 유리하게, 이러한 각각의 수화 사용자는 사용자 본인의 웨어러블 시스템을 사용하여 다른 수화 사용자의 싸인들을 사용자 본인의 수화 시스템으로 변환할 수 있다. 웨어러블 시스템은 싸인들을 사용자가 이해하는 텍스트로 또는 사용자 본인 수화의 그래픽 표현으로 변환할 수 있다.

[0142] [0154] 웨어러블 시스템(200)은 특정 수화, 예컨대 ASL(American Sign language)를 인식하도록 구성될 수 있다. 웨어러블 시스템(200)은 또한, 복수의 수화들, 예컨대 ASL, 영국 수화, 중국 수화, 도곤족 수화(Dogon Sign Language) 등을 인식하도록 구성될 수 있다. 일부 구현들에서, 웨어러블 시스템(200)은, 예컨대 감각 안경류 시스템의 위치 정보에 기반하여 수화 인식의 재구성을 지원한다. 웨어러블 시스템은, 시스템이 사용자 본인의 또는 선호하는 수화를 인식하는 방법과 유사한 수단을 통해, 예컨대, 외향 이미징 시스템(464)에 의해 지각되는 제스처들을 인식하기 위해 수화 사전을 함께 활용하여 또는 오브젝트 인식기들(708)만을 활용하여 외국 수화(foreign sign language)를 인식할 수 있다. 웨어러블 시스템은 사용자가 지각하는 수화를 사용자의 지배적(dominant) 수화로 변환할 수 있다. 사용자의 지배적 수화는 대화에서 사용자의 우선(first) 수화 또는 사용자가 선호하는 수화일 수 있다. 사용자의 지배적 수화 이외의 수화는 외국 수화로 고려될 수 있다. 웨어러블 시스템은 사용자가 외국 수화의 변환된 텍스트를 선택하게 허용할 수 있다. 예컨대, 사용자는 외국 수화를 선택할 수 있고, 웨어러블 시스템은 외국 수화에서의 제스처들의 의미를 웨어러블 시스템의 사용자에게 텍스트로서 제공할 수 있다.

[0143] [0155] 웨어러블 시스템은 환경 또는 위치 정보에서 발화된 언어의 지원을 통해 외국 수화를 인식할 수 있다. 예컨대, 웨어러블 시스템은, 사용자의 환경에서 이탈리아어가 발화된 것을 검출하거나 또는 GPS에 의해 획득된 데이터에 기반하여 사용자가 이탈리아에 있음을 결정한다. 이 정보에 기반하여, 웨어러블 시스템은 이탈리아 수화를 인식하기 위한 기능들을 자동으로 활성화시킬 수 있다. 또 다른 예로서, 웨어러블 시스템은 웨어러블 시스템이 지원할 수 있는 수화들에 대한 선호도의 순서를 가질 수 있다. 이 예에서, ASL이 이탈리아 수화에 비해 선호될 수 있는데, 이는 사용자가 미국 출신이기 때문이다. 그러나, 사용자가 이탈리아 화자들로 둘러싸여 있다는 것을 또는 물리적으로 이탈리아에 있다는 것을 웨어러블 시스템이 검출하면, 웨어러블 시스템은, 이제 이탈리아 수화가 ASL보다 우위에(before) 있도록 선호도의 순서를 변경할 수 있다. 따라서, 웨어러블 시스템은 이탈리아어 수화를 ASL과 연관된 영어 텍스트 또는 그래픽으로 변환할 수 있다.

[0144] [0156] 감각 안경류 시스템은 사용자가 외국 수화를 이해하는 것을 도울 뿐만 아니라 사용자가 외국 수화에 싸인하는 것을 도울 수 있다. 예컨대, 웨어러블 시스템은 사용자 본인 언어를 외국 수화로 변환하도록 구성될 수 있다. 시스템은 외국 수화 제스처들(예컨대, 변환된 수화)을 디스플레이상에 디스플레이할 수 있다. 사용자는 외국 수화의 제스처들을 보고 제스처들을 모방할 수 있다. 예컨대, 사용자는 난청이 있는 수화 사용자와 대화 중 일 수 있다. 웨어러블 시스템은 사용자의 스피치를 캡처하여 수화 사용자가 이해하는 수화로 해당 제스처들을 사용자에게 디스플레이할 수 있다. 그에 따라서, 사용자는 디스플레이에 의해 제공되는 제스처들이 수화 사용자와 통신하게 할 수 있다. 일부 실시예들에서, 웨어러블 시스템은, 제스처들을 사용자에게 보여주기 보다, 수화 사용자가 사용자의 보컬 스피치(vocal speech)를 이해할 수 있도록, 사용자의 스피치에 대응하는 싸인들을 수화 사용자에게 통신할 수 있다.

[0145] [0157] 웨어러블 시스템은 인식된 수화를 오디오로 제공하기 위해 오디오 증폭기(예컨대, 스피커(240))를 포함할 수 있다. 예컨대, 웨어러블 시스템은 수화 사용자에 의한 수화를 웨어러블 시스템(200)의 사용자에게 플레이백하기 위한 오디오 스트림으로 변환할 수 있다.

[0146] 수화와 연관된 보조 정보를 결정하는 예들

[0147] [0158] 수화를 수반하는 대화를 포함하여, 사람들이 대화에서의 단어들 또는 구절들을 모르거나 이해하지 못하는 것은 드문 일이 아니다. 웨어러블 시스템은 디스플레이된 텍스트의 일부와 연관된 보조 정보를 디스플레이하여 사용자의 이해를 향상시킬 수 있다. 보조 정보는 정보, 이를테면 정의, 변환, 설명 등(이는 정의의 콘텍

스트들을 보강하고 추가함)을 포함할 수 있다. 보조 정보는 다양한 형태들, 이를테면, 예컨대 텍스트, 이미지, 그래픽들, 애니메이션들 또는 다른 오디오 또는 시각 정보로 제공될 수 있다. 시스템은, 예컨대 도 2a의 디스플레이(220)를 통해 보조 정보를 시각적으로 제공할 수 있다. 시스템은, 청각-장애가 없는 사용자에게, 예컨대 도 2a의 오디오 증폭기(240)를 통해 오디오로 보조 정보를 제공할 수 있다. 이러한 단어들 또는 구절들에 대한 정의, 변환, 설명 또는 다른 정보를 제공함으로써, 웨어러블 시스템은 유리하게, 사용자가, 사용자가 관찰하는 수화를 더 잘 이해하게 보조할 수 있다.

[0148] [0159] 보조 정보는 사용자 환경의 상황 정보(contextual information), 스피치의 콘텍스트 등에 기반하여 결정될 수 있다. 예로서, 웨어러블 시스템은 대화 파트너들의 싸인들과 연관된 보조 정보를 디스플레이할지 여부를 결정하는데 있어, 적어도 부분적으로, 사용자 거동을 활용할 수 있다. 예컨대, 사용자는 특정 방향을(예컨대, 수화 사용자 또는 수화 사용자의 손들을 향해) 일시적으로 응시(stare)할 수 있다. 웨어러블 시스템은 (예컨대, 내향 이미징 시스템(462)을 사용하여) 사용자의 시선 방향을 검출할 수 있고, 응답으로, 대화 파트너의 싸인들과 연관된 보조 정보를 리트리브하고 디스플레이할 수 있다.

[0149] [0160] 웨어러블 시스템은 보조 정보의 데이터 저장소(예컨대, 데이터베이스)를 포함할 수 있다. 웨어러블 시스템은, 데이터 저장소에 액세스함으로써, 디스플레이된 텍스트와 연관된 보조 정보를 리트리브할 수 있다. 이러한 정보의 데이터베이스는 웨어러블 디바이스에, 예컨대 도 2a의 데이터 모듈(260) 내에 로컬로, 또는 예컨대 원격 데이터 저장소(270) 내에 원격으로 저장될 수 있다. 웨어러블 시스템은 보조 정보를 결정하기 위해, 공개적으로 액세스가능한 정보, 예컨대 인터넷 상의 정보를 활용할 수 있다. 예컨대, 웨어러블 시스템은 대화에서의 단어/구문에 관한 질의를 인터넷 상의 자원, 이를테면 사전, 백과 사전 또는 다른 유사한 자원에 전송하기 위해 네트워크에 액세스할 수 있다. 그러한 자원들은 일반적인 것(예컨대, 범용 백과 사전, 이를테면 위키피디아) 또는 전문화된 것(예컨대, 광물학 데이터베이스(예컨대, webmineral.com) 또는 rxlist.com 상의 것과 같은 약물들의 인덱스)일 수 있다.

[0150] 변환된 싸인들 또는 보조 정보의 예시적 디스플레이, 취소 및 리콜

[0151] [0161] 웨어러블 시스템은 (예컨대, 텍스트 또는 그래픽 포맷의) 변환된 싸인을 단독으로 또는 보조 정보와 조합하여 웨어러블 시스템(200)의 사용자에게 (예컨대, 텍스트 또는 그래픽 포맷으로) 변환된 싸인들을 제공할 수 있다. 예컨대, 웨어러블 시스템(200)은 수화의 변환된 싸인들과 함께 보조 정보를 디스플레이하도록, 변환된 싸인들 또는 보조 정보를 개별적으로 디스플레이(예컨대, 보조 정보가 디스플레이되는 지속기간 동안 보조 정보만이 디스플레이)하도록, 또는 2개의 디스플레이 모드들 간을 스위칭하도록 구성될 수 있다. 때때로, 변환된 싸인들이 단독으로 또는 보조 정보와 조합되어, 디스플레이 항목으로서 지칭할 수 있다.

[0152] [0162] 변환된 텍스트 또는 보조 정보는 다양한 방식들로 제공될 수 있다. 일 실시예에서, 웨어러블 시스템(200)은 변환된 텍스트 또는 보조 정보를 텍스트 버블들에 배치할 수 있는 데, 예컨대 텍스트는 도 13의 그래픽(1355)에 예시된 바와 같이 수화 사용자 인근에 기하학적으로 로컬화된다. 다른 예로서, 웨어러블 시스템(200)은 검출된 수화의 롤링 트랜스크립트를 디스플레이하도록 구성될 수 있다. 이 구성에서, 예컨대 사용자가 순간적으로 산만했을 경우에, 누락되었던 단어들 또는 심지어 문장들이 신속하게 재판독될 수 있다. 수화 사용자의 변환된 텍스트 트랜스크립트는 영화의 최종 크레딧(end credit) 프리젠테이션과 유사한 롤링 텍스트로서 디스플레이될 수 있다.

[0153] [0163] 변환된 텍스트의 트랜스크립트를 디스플레이하는 시스템은, 일부 방식으로, 예컨대, 밑줄, 컬러화, 굵은 텍스트 등으로, 보조 정보가 요청되는 단어 또는 구절을 강조 표시할 수 있다. 이러한 강조 표시들은 보조 정보가 리트리브되거나 디스플레이되기 전에 디스플레이될 수 있다. 이 디스플레이 모드로 구성된 일부 실시예들은, 사용자가 강조 표시된 텍스트의 요청을 확인 또는 제거하도록 허가할 수 있다. 선택적으로 또는 추가적으로, 그러한 강조 표시들은 보조 정보와 함께 디스플레이될 수 있다. 이 디스플레이 모드는 보조 정보와 연관된 텍스트가 사용자에게 명확해지게 할 수 있다. 아래에서 추가로 설명하는 바와 같이, 시스템은, 사용자가 UI 상호작용을 통해 현재 또는 과거에 변환된 텍스트를 선택하여, 연관된 보조 정보를 불러오거나(bring up) 다시 가져오게(bring back)할 수 있다.

[0154] [0164] 웨어러블 시스템(200)은, UI 상호작용을 통해 정보에 액세스하기 위한 사용자의 눈 움직임을 최소화시키기 위해, 변환된 텍스트 또는 보조 정보를 (예컨대, 텍스트 버블들에 또는 롤링 트랜스크립트로서) 배치할 수 있다. 이런 식으로, UI가 단순화되고, 사용자는 수화 사용자로부터 자신의 주의를 크게 기울일 필요가 없다. 변환된 텍스트 또는 보조 정보는, 판독 액션이 대화 파트너에게 최소한으로 보이도록 배치하고, 이렇게 하여, 변환된 텍스트 또는 보조 정보에 대한 사용자의 액세스를 드러내지 않으면서 주의산만함을 줄여 더 나은 통신을

제공할 수 있다. 예컨대, 수화 사용자의 위치를 결정할 수 있는 구현은, 변환된 텍스트 또는 보조 정보를 수화 사용자 옆에 배치할 수 있다. 예컨대, 외향 이미징 시스템(464)으로부터의 이미지들은, 예컨대 수화 사용자의 얼굴, 제스처 등을 가리지 않는 적절한 배치의 결정을 도울 수 있다. 웨어러블 시스템(200)은, 변환된 텍스트 또는 보조 정보 디스플레이의 배치를 결정하기 위해, 도 8에 예시된 프로세스 흐름을 사용할 수 있다. 예컨대, 블록(850)에서 인식된 오브젝트는, 수화 인식을 위해 수화를 프로세싱할 수화 사용자일 수 있다.

[0155] [0165] 사용자 또는 대화 파트너가 겪게 되는 주의산만함을 감소시키는 다른 예로서, 만약 변환된 텍스트 또는 보조 정보가 오디오로 제공된다면(예컨대, 사용자가 청각-장애가 없고 대화 파트너가 수화를 사용할 경우), 웨어러블 시스템은, 청취를 위해, (언어-장애는 있으나 청각-장애는 없는) 대화 파트너가 아닌 사용자를 위해 충분히 큰 볼륨으로 정보를 제공할 수 있거나, 또는 사용자 또는 대화 파트너가 말하고 있지 않을 때 정보를 제공할 수 있다.

[0156] [0166] 디스플레이된 아이템은 조건이 충족될 때까지 가시적으로 남아 있을 수 있다. 예컨대, 디스플레이된 아이템은, 표시되는 다음 아이템이 표시될 때까지 또는 사용자 액션에 의해 취소될 때까지 고정된 양의 시간 동안 가시적으로 남을 수 있다. 사용자 액션은 수동적일 수 있다(예컨대, 내향 이미징 시스템(462)에 의해 캡처된 눈 움직임들). 웨어러블 시스템은, 사용자가 디스플레이된 아이템을 검토했다고 결정할 때 디스플레이된 아이템을 취소할 수 있다. 예컨대, 디스플레이된 아이템이 텍스트라면, 시스템은 텍스트를 통한 사용자의 눈 움직임들(예컨대, 왼쪽에서 오른쪽으로 또는 최상부에서 최하부로)을 추적할 수 있다. 사용자가 디스플레이된 전체 아이템(또는 디스플레이된 아이템의 대부분)을 살펴보았다고 웨어러블 시스템이 결정하면, 웨어러블 시스템은 디스플레이된 아이템을 그에 따라서 취소할 수 있다. 다른 예로서, 디스플레이된 아이템은, 디스플레이된 아이템에 의해 점유된 영역으로부터 사용자가 눈길을 돌리는것(또는 보지 않는 것)을 시스템이 관찰된 후에 취소될 수 있다. 사용자 액션은 또한 활성일 수 있다(예컨대, 외향 이미징 시스템(464)에 의해 캡처된 손 제스처, 오디오 센서(232)에 의해 수신된 음성 입력, 또는 사용자 입력 디바이스(466)로부터의 입력). 예컨대, 웨어러블 시스템이 사용자에 의한 스와이프 제스처를 검출하면, 웨어러블 시스템은 디스플레이된 아이템을 자동으로 취소할 수 있다.

[0157] [0167] 웨어러블 시스템은 특정 사용자에 대해 사용자 인터페이스(UI) 상호작용들의 맞춤화된 세트를 지원하도록 구성될 수 있다. UI 상호작용들은 손가락으로, 일부 종류의 스타일러스 또는 포인터로, 눈들로 버튼에 대한 시선 및 후속적인 응시에 의해 또는 기타에 의해 작동되는 버튼과 유사한 UI 엘리먼트의 형태를 취할 수 있다. 버튼은 실제 물리적 버튼(예컨대, 키보드) 또는 디스플레이(220)에 의해 디스플레이된 가상 버튼일 수 있다. UI 상호작용들은 예컨대, 도 4와 관련하여 앞서 설명된 바와 같이, 머리 포즈의 형태를 취할 수 있다. UI 상호작용 검출의 예가 도 10과 관련하여 위에서 설명된다.

[0158] [0168] 웨어러블 시스템은 사용자에게 디스플레이된 아이템의 취소를 연기하도록 촉구할 수 있다. 예컨대, 웨어러블 시스템(200)은, 디스플레이된 아이템이 곧, 예컨대, 몇 초 후, 취소될 것임을 사용자에게 통지하기 위해, 디스플레이된 아이템의 밝기를 감소시키거나 컬러 방식을 변화시킬 수 있다. 위에서 설명된 것과 같은 UI 상호작용은 취소를 연기하기 위해 사용될 수 있다. 예컨대, 웨어러블 시스템은, 사용자가 디스플레이된 아이템에서 눈길을 돌렸음을 검출할 수 있다. 따라서, 웨어러블 시스템은, 디스플레이된 아이템의 투명도를 증가시켜서 사용자에게 디스플레이된 아이템이 곧 취소될 것임을 통지할 수 있다. 그러나 눈 추적을 통해, 웨어러블 시스템이 사용자가 디스플레이된 아이템을 다시 보는 것을 검출하면, AR 시스템은 취소를 연기할 수 있다.

[0159] [0169] 위에서 설명된 것과 같은 UI 상호작용은 또한 취소되어 버린, 디스플레이된 아이템을 리콜하기 위해 사용될 수 있다. 예컨대, 사용자 입력 디바이스를 통한 입력 액션(예컨대, 키보드 상의 백 스페이스의 작동)은 가장 최근에 디스플레이된 아이템을 리콜하는 데 사용될 수 있거나, 리콜을 위해 특정 디스플레이된 아이템을 선택하는 데 사용될 수 있다.

[0160] 감각 안경류 시스템의 사용자 경험들의 예

[0161] [0170] 도 13a는, 웨어러블 시스템의 사용자를 위해, 감각 안경류 시스템이 (예컨대, 수화 사용자에 의해 제스처링된) 수화를 해석할 수 있는 감각 안경류 시스템의 예시적인 사용자 경험을 도시한다. 이 예는 감각 안경류 시스템의 사용자가 관찰하는 수화 사용자(1301)를 도시한다. 사용자는 장면(1305, 1310 및 1315)에 도시된 바와 같이, 수화 사용자(1301)가 손 제스처의 시퀀스(1300)를 만들고 있음을 지각할 수 있다. 장면(1305)에서의 손 제스처는 "how"라는 단어를 표현하며; 장면의 손 제스처는 "are"라는 단어를 표현하며; 장면(1315)의 손 제스처는 단어 "you"를 표현한다. 따라서 시퀀스(1300)는 "How are you"로 해석될 수 있다. 시퀀스들(1320 및 1340)은 시퀀스(1300)와 동일한 제스처들을 도시한다. 제스처(1305)는 제스처들(1325, 1345)에 대응하고; 제스

처(1310)는 제스처들(1330 및 1350)에 대응하고; 제스처(1315)는 제스처들(1335, 1355)에 대응한다. 그러나, 시퀀스들(1300, 1320 및 1340)은 이하에서 추가로 설명되는 바와 같이 상이한 사용자 디스플레이 경험을 예시한다.

[0162] [0171] 시퀀스(1300)의 손 제스처를 영어 구절 "How are you"로 변환하기 위해, 웨어러블 시스템(200)의 외향 이미징 시스템(464)은 일련의 이미지들 또는 비디오로서 제스처들의 시퀀스를 캡처할 수 있다. 웨어러블 시스템은 일련의 이미지들 또는 비디오에서 제스처들을 추출할 수 있다. 웨어러블 시스템은 추출된 제스처에 대해, 예컨대, 오브젝트 인식기들(708) 또는 심층 학습 알고리즘을 적용하여 수화 인식을 수행할 수 있다. 수화의 처리 또는 인식에 있어서, 웨어러블 시스템은 로컬 또는 원격 스토리지에 저장된 수화 사전에 액세스할 수 있다. 웨어러블 시스템은 인식된 수화로부터 변환된 텍스트(또는 싸인의 그래픽 표현)를 디스플레이(220)를 통해 사용자(미도시)에게 디스플레이할 수 있다. 감각 안경류 시스템은 또한, 변환된 싸인들과 연관된 보조 정보에 대한 요청을 수신할 수 있으며, 본원에서 설명된 기법들을 사용하여 보조 정보를 리트리브하고 디스플레이할 수 있다.

[0163] [0172] 도 13a에 예시된 그래픽 시퀀스들에서, "how"라는 단어를 표현하는 것은, 예컨대, 그래픽들(1305 및 1310)에 도시된 바와 같이, 두 개의 별개의 제스처들을 취한다. 웨어러블 시스템은 (장면(1305)에서 제스처링된 대로) 단어 "how"를 디스플레이하기 전에, (장면(1310)에서의) 두 번째 제스처 이후까지 기다릴 수 있다. 추가적으로 또는 대안적으로, 웨어러블 시스템은 그래픽 시퀀스(1320)에 도시된 바와 같이 문장 또는 구절이 완료될 때까지 텍스트 변환 또는 디스플레이를 홀딩할 수 있으며, 여기서 "How are you"는 장면(1335)의 끝에 표시된다. 웨어러블 시스템은 변환된 텍스트 또는 보조 정보를, 예컨대 그래프(1355)에 도시된 바와 같이 자막 또는 텍스트 버블로서 디스플레이할 수 있다. 자막 또는 텍스트 버블은, 사용자에 대한 주의산만함을 최소화하기 위해, 사용자의 FOV 내에, 예컨대, 수화 사용자 얼굴의 사용자의 뷰를 가리지 않고 수화 사용자에 근접하게 위치될 수 있다.

[0164] [0173] 도 13b는 타겟 스피치 및 보조 정보가 모두 제공되는, 감각 안경류 시스템의 또 다른 예시적인 사용자 경험을 도시한다. 이 예에서, 사용자(미도시)는 HMD를 착용하고 수화 사용자(1362)를 지각할 수 있다. 수화 사용자는 수화를 사용하여 "Where is the PTO?"라는 질문을 한다(수화 사용자는 질문 끝에 글자 "O"를 제스처하는 것으로 묘사됨). 웨어러블 시스템은 수화 사용자에 의해 행해진 제스처들을 인식하고, 이들을 텍스트로 변환하고, 변환된 텍스트를 텍스트 버블(1360)로 웨어러블 시스템의 사용자에게 디스플레이할 수 있다. 웨어러블 시스템은 "PTO"가 두문자어이고 사용자가 일상적인 스피치에서 흔히 사용하지 않는 단어라는 것을 결정할 수 있다. 예컨대, 웨어러블 시스템은 공통적으로 사용되는 단어들 및 구절들의 사전을 유지하고 "PTO"가 사전에 없다고 결정할 수 있다. "PTO"라는 단어가 사전에 있지 않다고 검출할 시에, 웨어러블 시스템은 구절 "PTO"와 연관된 보조 정보의 액세스를 개시할 수 있다.

[0165] [0174] 웨어러블 시스템은 상황 정보에 기반하여 두문자어에 대한 보조 정보를 리트리브할 수 있다. 이 예에서, 시스템은 자신의 위치 정보에 의존할 수 있는데, 예컨대, 시스템(및 그 사용자)은 현재 버지니아의 알렉산드리아에 있다. 시스템은 약어 "PTO"의 보조 정보로서 "Patent and Trademark Office"를 리트리브한다. 시스템은 디스플레이(220)를 통해 사용자에게 가상 배너(1365)로서 보조 정보를 디스플레이한다. 도 13b에 도시된 변환된 텍스트 및 보조 정보의 디스플레이 모드는 단지 예시를 위한 것이다. 일부 실시예들은 이들을 다르게 디스플레이할 수 있는데, 예컨대, 자막에서 둘 모두가 순차적으로 디스플레이된다.

[0166] [0175] 복수의 착용 가능 시스템의 복수의 사용자는 그들의 각각의 웨어러블 시스템들의 도움을 통해 원격으로 통신할 수 있다. 도 13c는 텔레프레즌스 세션에서의 감각 안경류 시스템의 사용자 경험의 예를 도시한다. 예컨대, 도 13c에 예시된 바와 같이, (사람이 만든 디바이스의 도움없이 직접 서로를 보거나 들을 수 없도록) 두 물리적 위치들(1370a, 1370b)에 있는 두 명의 사용자(1372a, 1372b)는 둘 모두, 각각 웨어러블 디바이스(1374a, 1374b)를 착용할 수 있다. 한 명 또는 둘 모두의 사용자(1372a, 1372b)는 수화를 사용하여 대화할 수 있다. 손 제스처들은 사용자들 개개의 웨어러블 시스템의 이미징 시스템에 의해 캡처되어 네트워크(1290)를 통해 송신될 수 있다. 사용자 A(1372a)의 수화는 사용자 B(1372b)의 디바이스에 변환된 텍스트로 표시될 수 있으며, 그 반대의 경우도 마찬가지이다.

[0167] [0176] 감각 안경류 시스템은 검출된 수화를 로컬로 텍스트로 변환하여 네트워크(1290)를 통해 변환된 텍스트만을 송신할 수 있다. 다른 사용자의 디바이스는 텍스트를 디스플레이하거나, 다른 사용자가 청각-장애가 있지 않은 경우, 텍스트를 가청 스피치로 변환할 수 있다. 이는, 대응하는 이미지, 비디오 또는 오디오를 송신하는 것보다 텍스트를 송신하는 데 더 적은 양의 데이터가 요구되기 때문에, 네트워크(1290)의 대역폭이 제한되는 경

우 유리할 수 있다.

[0168] [0177] 웨어러블 시스템은 또한 디스플레이(220) 상에 제공되는 이미지들을 통해 텔레프레전스 대화를 향상시킬 수 있다. 예컨대, 디스플레이(220)는 참가자의 시각적 감각의 주의를 끌기 위해 변환된 텍스트 또는 보조 정보와 함께 원격 수화 사용자의 아바타를 제공할 수 있다. 예컨대, 내향 이미징 시스템(464)을 갖춘 웨어러블 디바이스는 HMD에 의해 가려진 착용자의 얼굴의 영역을 대체하기 위한 이미지들을 캡처할 수 있으며, 이는 제1 사용자가 텔레프레전스 세션 동안 제2 사용자의 가려지지 않은 얼굴을 볼 수 있도록 사용될 수 있으며, 그 반대의 경우도 마찬가지이다. 제1 사용자와 연관된 월드 맵 정보는 감각 안경류 시스템들을 수반하는 텔레프레전스 세션의 제2 사용자에게 통신될 수 있다. 이것은 HMD 착용자가 볼 원격 사용자의 이미지들의 생성을 통해 사용자 경험을 향상시킬 수 있다.

[0169] [0178] 텔레프레전스 애플리케이션에서, 이미지 정보를 캡처하는 것은 수화 관찰자와 연관된 디바이스가 아닌 (예컨대, 일인칭 관점에서) 수화 사용자와 연관된 디바이스에 의해 수행되며, 이는 대면적인(in-person) 시나리오에서 통상적일 수 있다. 수화의 존재의 검출 및 수화의 텍스트로의 변환은 사용자와 연관된 디바이스 또는 원격 시스템, 예컨대 서버 컴퓨터 시스템(1220)에 의해 수행될 수 있다. 수화의 소스는, 이미지를 캡처하는 디바이스에 기초하여 결정될 수 있는데, 예컨대, 사용자 A의 디바이스가 이미지를 캡처할 때, 사용자 A는 싸인하고 있다.

[0170] [0179] 도 13d는 수화를 해석하기 위한 예시적인 가상 사용자 인터페이스를 예시한다. 이 예에서, 사용자(1392)는 (웨어러블 시스템(200)의 적어도 일부를 포함할 수 있는) 웨어러블 디바이스(1380)를 착용하고 있다. 이 예에서, 사용자(1392)는 카운터 뒤에 있으며, 카운터에 접근하는 사람(1394)을 지각한다. 예컨대, 사용자(1392)는 의료 설비의 간호사 또는 입원자, 손님을 보조하는 호텔 직원(예컨대, 안내인) 등일 수 있다. 사람(1394)은 몸이 편치 않음을 느끼고 약국에 대한 안내와 같은 의료 조치를 요청할 수 있다. 웨어러블 디바이스(1380)는 도 13d에 도시된 바와 같이 사용자(1394)에 의한 손 제스처들을 (예컨대, 외향 이미징 시스템(464)을 통해) 관찰할 수 있다. 웨어러블 디바이스(1380)는 도시된 바와 같은 손 제스처들이 수화 표현인 것을 자동으로(예컨대, 오브젝트 인식기(708)를 사용하여) 검출하고, 손 제스처와 연관된 의미를 인식하고, 그리고 사용자(1392)가 이해하는 타겟 언어(예컨대, 영어)로 손 제스처의 변환을 제공할 수 있다. 웨어러블 디바이스(1380)는 웨어러블 디바이스에 의해 캡처된 입력(1384a)을 나타내기 위한 가상 사용자 인터페이스(1382)를 제공할 수 있는데, 변환(1384b)은 입력(1384a)에 대응한다(예컨대, "Is there a pharmacy nearby? I'm feeling unwell"). 웨어러블 시스템은 또한 가상 사용자 인터페이스(1382) 상에 사용자 입력 엘리먼트들(1384c 및 1384d)을 제공할 수 있다. 예컨대, 사용자(1392)는 사용자 입력 엘리먼트(1384c)를 선택하기 위해 손 제스처(예컨대, 프레스 제스처)를 사용할 수 있다. 사용자 입력 엘리먼트(1384c)의 작동은 웨어러블 디바이스로 하여금 예컨대, 근처의 약국의 위치 또는 "I don't know"와 같은 응답들의 리스트를 제공하게 할 수 있다. 일부 실시예들에서, 웨어러블 디바이스(1380)는 응답들을 위해 수화로 대응하는 그래픽들을 도시할 수 있다. 그에 따라서, 사용자(1392)는 그래픽들에 도시된 바와 같이 손 제스처들을 사용하여 사람(1394)에게 응답할 수 있다. 다른 예로서, 사용자 입력 엘리먼트(1384d)가 작동되는 경우, 웨어러블 시스템은 예컨대, 사용자 인터페이스 엘리먼트(1382)를 취소하거나 도움 요청을 하는 것 등과 같은 옵션들의 리스트를 제공할 수 있다. 일부 실시예들에서, 인터페이스(1382)의 영역(1384a)은, 사용자(1392)가 사람(1394)과 통신하기 위해 수행할 수 있는 수화 제스처들(예컨대, "the pharmacy is across the street"에 대한 싸인들)을 도시하는 출력 그래픽을 포함할 수 있다.

[0171] 개인간 통신을 위한 도구로서 감각 안경류 시스템을 위한 예시적인 프로세스들

[0172] [0180] 도 14a 및 14b는 감각 안경류 시스템에 의해 개인간 통신들을 가능하게 하기 위한 예시적인 프로세스들을 예시한다. 도 14a 및 14b의 예시적인 프로세스들(1400 및 1440)은 도 2a에 도시된 웨어러블 시스템에 의해 수행될 수 있다.

[0173] [0181] 블록(1404)에서, 웨어러블 시스템은 환경에서 이미지 정보를 캡처할 수 있다. 본원에서 설명된 바와 같이, 웨어러블 시스템은 사용자의 주변에서 이미지 정보를 캡처하기 위해 외향 이미징 시스템(464)을 사용할 수 있다. 웨어러블 시스템은 또한, 환경에서 오디오 정보를 캡처할 수 있다. 스피치 또는 제스처의 소스를 결정하거나 수화의 존재를 검출하기 위해, 외향 이미징 시스템(464)에 의해 획득된 데이터와 함께 오디오 정보가 사용될 수 있다.

[0174] [0182] 블록(1408)에서, 웨어러블 시스템은 캡처된 이미지 정보에서 수화의 존재를 검출한다. 이러한 검출 프로세싱은, (예컨대 로컬 프로세싱 모듈(71)에 의해) 로컬로 또는 (예컨대, 원격 프로세싱 모듈(72)에 의해) 원격으로 수행될 수 있다. 웨어러블 시스템은 손 제스처들의 존재를 검출하기 위해 다양한 오브젝트 인식기들을

사용할 수 있다. 예컨대, 웨어러블 시스템은 손 제스처들의 시퀀스가 수화의 구절 또는 문장을 구성할 수 있다는 것을 발견할 수 있다. 다른 예로서, 웨어러블 시스템은 일련의 손 제스처들뿐만 아니라 입술 움직임들을 검출할 수 있다. 웨어러블 시스템은 손 제스처들 및 입술 움직임들이 수화와 연관된다는 것을 발견할 수 있는데, 왜냐하면 그러한 제스처들 및 입술 움직임들에는 오디오 정보가 수반되지 않기 때문이다.

[0175] [0183] 일부 실시예들에서, 웨어러블 시스템은 상황 정보에 기반하여 수화를 검출 및 해석할 수 있다. 예컨대, 웨어러블 시스템은 (예컨대, 저녁 식사 대화의) 오디오 신호들을 수신하고, 이러한 신호들을 언어로 변환하거나, 그 언어로부터 의미를 추출함으로써, (이를테면, 예컨대 토론들의 주제들과 일치하는 방식으로 손 제스처들을 해석하는 것과 같이) 수화를 해석하는 데 사용될 수 있는, 토론의 주제들의 장르(또는 다른 속성)를 추론할 수 있다.

[0176] [0184] 웨어러블 시스템은 사용자 본인의 수화를 검출 또는 무시하도록 구성될 수 있다. 블록(1408)의 기능은 이러한 구성에 기반하여 상이할 수 있는데, 왜냐하면 사용자 본인의 수화는 비교적 가까운 거리에서 일인칭 관점에서 캡처될 수 있기 때문이다. 예컨대, 만약 시스템이 사용자 본인의 수화를 캡처하도록 구성된다면, 사용자의 손들에서 하향으로 지향되는 추가적인 외향 카메라가 턴온될 수 있거나, 또는 외향 이미징 시스템이 광각 모드로 구성되어 사용자의 손들의 이미지들을 캡처할 수 있다.

[0177] [0185] 블록(1412)에서, 시스템은 수화가 검출되는지 여부를 결정한다. 만약 수화가 검출된다면, 프로세스 흐름(1400)은 블록(1416)으로 계속된다. 만약 수화가 검출되지 않는다면, 흐름은 (도시된 바와 같이) 블록(1408)으로 리턴하거나 블록(1404)(미도시)으로 리턴한다.

[0178] [0186] 블록들(1404 내지 1412)의 동작들은, 웨어러블 시스템(그 이미징 시스템들을 포함함)이 턴온될 때 또는 수화 인식 기능이 인에이블링될 때 주기적으로(예컨대, 샘플링 빈도로) 또는 계속해서 수행될 수 있다. 이러한 동작들은 순서도들(1400 및 1440)에서 (예컨대, 타이밍된 인터럽트에 의해 구동되는 배경 임무들로서) 다른 블록들과 병렬로 수행될 수 있다. 이들은 예시의 목적으로 프로세싱 흐름 시퀀스에서 이산 블록들로서 도시된다. 하지만, 이들은 예시된 시퀀스에 의해 제한되지 않는다. 시스템 설계자의 재량으로, 위에서 설명된 예들 이외의 많은 프로세싱 흐름들이 가능하다.

[0179] [0187] 블록(1416)에서, 웨어러블 시스템은 검출된 수화의 소스(예컨대, 수화 사용자)를 결정할 수 있다. 소스는 사용자와 물리적으로 가까이 있는 사람, 사용자, 또는 사용자가 지각하는 시각 콘텐츠의 사람일 수 있다. 수화의 소스는, 예컨대, 만약 시스템이 웨어러블 시스템의 FOV의 중심 또는 그 인근의 사람으로부터의 수화만을 프로세싱하도록 구성되는 경우(예컨대, 다수의 사람들이 수화로 동시에 대화하고 있을 때, FOV의 중심을 벗어난 사람들로부터의 수화는 버려지고 추가로 프로세싱되지 않을 수 있음), 관련될 수 있다. 다른 예로서, 웨어러블 시스템은, FOV의 중심에 있는 사람이거나 또는 그렇지 않을 수도 있는, 사용자가 보고 있는 사람에 대해서만 수화 인식을 위해 제스처들을 프로세싱할 수 있다. 웨어러블 시스템은, 내향 이미징 시스템(462) 및 외향 이미징 시스템(464)에 의해 획득되는 데이터에 기반하여, 사용자가 보고 있는 사람을 식별할 수 있다. 예컨대, 외향 카메라는 사용자에 대한 수화 사용자의 포지션을 포함하는 정보를 제공할 수 있다. 내향 카메라는 사용자가 보고 있는 방향을 포함하는 정보를 제공할 수 있다. 둘 모두의 카메라들로부터의 정보를 사용함으로써, 웨어러블 시스템은 사용자가 보고 있는 사람, 및 그 사람이 수화의 소스인지 여부를 결정할 수 있다.

[0180] [0188] 블록(1420)에서, 시스템은 수화의 소스의 변화가 있었는지 여부를 결정한다. 만약 변화가 있었다면, 흐름(1400)은 블록(1424)을 통해 도 14b에 도시된 블록(1444)으로 계속된다. 만약 수화의 소스의 변화가 없었다면, 흐름은 블록(1428)으로 이동하여 수화 인식 프로세싱을 계속하며, 이러한 수화 인식 프로세싱은 이미지 정보를 캡처하는 것(블록(1404)), 수화의 존재를 검출하는 것(블록(1408))뿐만 아니라, 도 14b에 도시된 프로세싱 단계들을 포함할 수 있다. 예컨대, 만약 계속해서 동일한 수화 사용자로부터 제스처들이 계속해서 비롯된다고 시스템이 결정한다면, 시스템은 이미지 정보를 계속해서 캡처하고 수화를 검출하는 것 외에도, 블록(1448)으로부터 시작되는 기능들을 계속해서 수행할 수 있다.

[0181] [0189] 블록(1448)에서, 웨어러블 시스템은 수화를 사용자가 이해하는 언어로 변환할 수 있다. 예컨대, 시스템은 인식된 수화를 텍스트로 변환할 수 있으며, 이는 시스템에 의해 (예컨대, 텍스트 버블 또는 자막으로서) 디스플레이될 때 사용자에 의해 읽혀질 수 있다. 일부 경우들에서, 만약 사용자가 상이한 수화를 이해한다면, 다른 수화 사용자의 싸인들의 그래픽 표현이, 예컨대, 사용자 본인의 수화의 싸인들로 변환된 싸인들을 도시하는 그래픽스로서 사용자에게 디스플레이될 수 있다.

[0182] [0190] 블록(1452)에서, 예시적인 시스템은, 시스템이 사용자 본인뿐만 아니라 대화 파트너의 수화를 검출하도

록 구성될 때, 검출된 수화가 사용자 본인의 것인지 여부를 결정할 수 있다. 만약 검출된 수화가 사용자 본인의 것이라면, 프로세스는 블록(1484)으로 진행되며, 여기서, 시스템은 변환된 텍스트를 관찰자/대화 파트너의 디스플레이 디바이스에 송신할 수 있다.

[0183] [0191] 블록(1484)으로부터, 시스템은 프로세싱을 계속하기 위해 블록(1488)으로 진행할 수 있다. 시스템이 사용자 본인의 수화를 무시하도록 구성될 때, 블록들(1452 및 1484) 둘 모두는 흐름으로부터 생략될 수 있다. 만약 검출된 수화가 사용자 본인의 것이 아니라면, 흐름은 블록(1456)으로 계속된다.

[0184] [0192] 블록(1456)에서, 앞서 설명된 바와 같이, 웨어러블 시스템은 변환된 텍스트를 디스플레이에 의해 디스플레이할 수 있다. 시스템의 사용자가 청각 장애가 없는 경우, 시각적 디스플레이 외에도 또는 그에 대한 대안으로서, 텍스트는, 예컨대 오디오 증폭기(240)를 통해, 오디오로 제공될 수 있다.

[0185] [0193] 블록(1460)에서, 웨어러블 시스템은 변환된 텍스트에 관한 보조 정보에 대한 요청을 모니터링 할 수 있다. 보조 정보에 대한 요청은, 트리거링 조건의 검출시에 사용자의 웨어러블 디바이스에 의해 전송될 수 있다. 일부 예시적인 트리거링 조건들은 사용자의 표시, 예컨대, 사용자의 제스처 또는 사용자 입력 디바이스(466)의 작동; 또는 사용자가 이해할 수 없는 단어(또는 구절)를 검출할 때를 포함할 수 있다.

[0186] [0194] 블록(1464)에서, 시스템은 요청이 수신되는지 여부를 결정한다. 만약 요청이 수신되지 않는다면, 흐름은 블록(1476)으로 이동하며, 이는 아래에서 추가로 설명된다.

[0187] [0195] 만약 요청이 수신된다면, 블록(1468)에서, 시스템은 변환된 텍스트(또는 그 요청된 부분)와 연관된 보조 정보를 리트리브할 수 있다. 본원에서 설명된 바와 같이, 보조 정보는, 예컨대 사용자의 위치, 스피치의 상황, 또는 본원에서 설명된 다른 타입들의 정보와 같은 상황 정보에 기반하여 결정 및 리트리브될 수 있다.

[0188] [0196] 블록(1472)에서, 웨어러블 시스템은 리트리브된 보조 정보를 웨어러블 시스템의 디스플레이(220)를 통해 디스플레이할 수 있다. 일부 구현들에서, 웨어러블 시스템은 보조 정보를 디스플레이하기 전에, 변환된 텍스트의 디스플레이를 취소할 수 있다.

[0189] [0197] 흐름은 블록(1464 또는 1472)으로부터 블록(1476)으로 들어갈 수 있다. 블록(1476)에서, 시스템은 변환된 텍스트 또는 보조 정보 디스플레이를 취소하기 위한 조건을 검출할 수 있다. 그러한 조건이 검출될 때, 블록(1480)에서, 시스템은 변환된 텍스트 또는 보조 정보의 디스플레이를 취소하고, 블록(1488)으로 계속될 수 있다. 블록(1488)에서, 위의 블록(1428)의 설명과 유사한 방식으로, 수화 인식의 프로세싱이 계속된다.

[0190] [0198] 블록들(1404 내지 1412)에 대하여 위에서 주목한 것과 유사하게, 순서도(1440)에서의 동작들은 순서도들(1400 및 1440)에서의 다른 블록들과 병렬로 수행될 수 있다. 이들은 예시의 목적으로 프로세싱 흐름 시퀀스에서 이산 블록들로서 도시되지만, 이들은 예시된 시퀀스에 의해 제한되지 않는다. 예컨대, 시스템이 (블록(1448)에서) 추가적인 수화를 텍스트로 변환하거나, (블록(1460)에서) 보조 정보 요청 모니터링을 수행하거나, 또는 (블록(1468)에서) 다른 변환된 텍스트에 대한 보조 정보를 리트리브하는 동안, 시스템은 (블록(1472)에서) 변환된 텍스트에 대한 보조 정보를 디스플레이하고 있을 수 있다. 다른 예로서, 시스템이 (블록(1468)에서) (보조 정보에 대해) 사전-요청된(prior-requested) 변환된 텍스트에 대한 보조 정보를 리트리브하는 동안, 시스템은 (블록(1448)에서) 수화를 텍스트로 변환할 수 있다. 시스템 설계자의 재량으로, 많은 다른 프로세싱 흐름들이 가능하다.

[0191] [0199] 도 14c는 보조 정보를 결정하고 그리고 변환된 텍스트와 연관된 보조 정보를 제공하기 위한 예시적인 방법의 프로세스 흐름도이다. 이러한 프로세스(1490)는 본원에서 설명된 웨어러블 시스템(200), 또는 그 자체가 수화 인식 기능성을 갖거나 갖지 않을 수 있는 다른 컴퓨팅 디바이스 상에서 실행가능할 수 있다. 이러한 프로세스(1490)는, 수화를 검출하고, 하나의 감각 안경류 시스템을 사용하여 수화를 텍스트로 변환하고, 그리고 변환된 텍스트를 다른 디바이스 또는 시스템 상에 디스플레이하는 것이 더 유리한 상황들에 적용가능할 수 있다. 예시적인 상황은, 수화 사용자가 두 번째 사람과 원격으로 통신하기를 원하는 상황일 수 있다. 수화 사용자의 웨어러블 디바이스는 수화 사용자 본인의 수화를 텍스트로 변환할 수 있다. 웨어러블 디바이스는 변환된 텍스트를, 두 번째 사람이 볼 수 있는 원격 시스템에 송신할 수 있다. 변환된 텍스트는 대응하는 이미지들 또는 비디오보다 훨씬 더 적은 정보 비트로 송신될 수 있기 때문에, 그러한 프로세스는 유리하게, 송신 매체로부터 훨씬 더 낮은 대역폭을 요구하거나 또는 훨씬 더 신뢰성있는 통신을 초래할 수 있다.

[0192] [0200] 프로세스(1490)는 블록(1492)에서 시작되며, 여기서, 디바이스 또는 시스템은 수화 프로세싱과 관련되거나 관련되지 않을 수 있는 일부 유형의 프로세싱을 수행한다. 블록(1494)에서, 디바이스 또는 시스템은 웨어러블 시스템으로부터 텍스트가 수신되는지 여부를 결정할 수 있다. 만약 텍스트가 수신되지 않는다면, 프로세

스는 블록(1492)으로 리턴할 수 있다. 만약 텍스트가 수신된다면, 프로세스는 블록(1496)으로 진행될 수 있다. 블록(1496)에서, 디바이스 또는 시스템은 웨어러블 시스템으로부터 텍스트를 수신하고 그 텍스트를 렌더링할 수 있다. 그런 다음, 프로세스는 블록(1456)으로 진행될 수 있다. 렌더링 디바이스가 HMD를 포함하는 경우, 렌더링 디바이스는 사용자의 물리적 환경에 오버레이되는 가상 콘텐츠로서 텍스트를 제공할 수 있다. 블록들(1456 내지 1480)에서의 프로세싱은 도 14b와 관련하여 위에서 설명된 것과 유사하게 진행될 수 있다.

[0193] [0201] 도 15는 감각 안경류 시스템에 의해 개인간 통신을 가능하게 하기 위한 다른 예시적인 프로세스를 예시한다. 예시적인 프로세스(1500)는 본원에서 설명된 웨어러블 시스템(200)의 하나 이상의 컴포넌트들에 의해(예컨대, 로컬 프로세싱 및 데이터 모듈(260), 원격 프로세싱 모듈(270)에 의해, 단독으로 또는 조합하여) 수행될 수 있다. 도 12를 참조하여 설명된 바와 같이, 이러한 도 15에서 설명되는 단계들 중 하나 이상은 사용자의 웨어러블 시스템의 일부가 아닌 하나 이상의 컴퓨팅 디바이스들, 이를테면, 예컨대 다른 사용자의 웨어러블 디바이스, 또는 제3 자의 서버 시스템에 의해 수행될 수 있다.

[0194] [0202] 블록(1510)에서, 웨어러블 시스템은 환경에서 스피치를 식별하고 인식할 수 있다. 스피치는 수화의 형태일 수 있다. 예컨대, 웨어러블 시스템은 수화의 일부인 손 제스처들을 식별하기 위해 외향 이미징 시스템(464)에 의해 획득된 데이터를 분석할 수 있다. 웨어러블 시스템은 또한 사용자의 환경에서 사람에 의한 스피치를 포함할 수 있는, 오디오 센서(232)에 의해 획득된 오디오 데이터를 분석할 수 있다. 웨어러블 시스템은 오브젝트 인식기들(708)을 사용하여 스피치를 인식할 수 있다. 예컨대, 웨어러블 시스템은 오브젝트 인식기들을 사용하여 수화의 이미지들을 분석함으로써 구절 또는 단어의 존재를 인식할 수 있다. 웨어러블 시스템은 또한 도 7에서 설명된 다양한 스피치 인식 알고리즘들을 사용하여 오디오 데이터를 인식할 수 있다.

[0195] [0203] 블록(1520)에서, 웨어러블 시스템은 타겟 언어를 식별할 수 있다. 타겟 언어는 웨어러블 시스템의 사용자가 통신하는 데 사용하는 언어일 수 있다. 예컨대, 사용자는 영어를 사용하는 다른 사람들과 통신할 수 있는 한편, (다른 사용자에 의해 사용되는) 인식된 오리지널 스피치는 수화이다. 타겟 언어는 또한 사용자에 의해 또는 웨어러블 시스템에 의해 선택된 언어일 수 있다. 예컨대, 사용자가 다른 언어를 말하더라도 사용자는 수화를 사용하여 다른 사람과 통신하기를 원할 수 있기 때문에, 사용자는 타겟 언어로서 ASL을 선택할 수 있다. 다른 예로서, 웨어러블 시스템은 사용자의 위치에 기반하여 언어를 자동으로 선택할 수 있다. 예컨대 웨어러블 시스템은 사용자가 어느 국가에 있는지를 결정하고 해당 국가의 공식 언어를 타겟 언어로서 선택할 수 있다.

[0196] [0204] 블록(1530)에서, 웨어러블 시스템은 검출된 스피치를 타겟 언어로 변환할 수 있다. 웨어러블 시스템은 본원에서 설명된 다양한 기법들, 이를테면, 예컨대 사전 변환들을 사용하여 그러한 변환을 수행할 수 있다.

[0197] [0205] 블록(1540)에서, 웨어러블 시스템은 웨어러블 시스템의 사용자에게 프리젠테이션하기 위해 변환된 스피치와 연관된 시청각 콘텐츠를 결정할 수 있다. 하나의 예로서, 시청각 콘텐츠는 타겟 언어의 텍스트를 포함할 수 있다. 다른 예로서, 시청각 콘텐츠는 변환된 스피치가 발화된 언어인 타겟 언어의 오디오 스트림일 수 있다. 또 다른 예로서, 만약 타겟 언어가 수화라면, 시청각 콘텐츠는 그래픽들 또는 애니메이션들일 수 있다.

[0198] [0206] 선택적인 블록(1550)에서, 웨어러블 시스템은 프레젠테이션을 위해 머리-장착 디스플레이에 시청각 콘텐츠를 통신할 수 있다. 예컨대, 시청각 콘텐츠는 한 사용자의 웨어러블 디바이스에서 다른 사용자의 웨어러블 디바이스로 통신될 수 있다. 이 예에서, 제1 사용자의 웨어러블 디바이스는 제1 사용자를 캡처하고, 스피치를 타겟 언어로 변환하고, 변환된 스피치를 제2 사용자의 웨어러블 디바이스로 통신할 수 있다.

[0199] 환경과 상호작용하기 위한 도구로서의 예시적인 감각 안경류 시스템

[0200] [0207] 다른 사람에 의해 제스처들을 인식하는 것 외에도 또는 그에 대한 대안으로서, 본원에서 설명된 웨어러블 시스템은 또한 예컨대, 도 7을 참조하여 설명된 다양한 텍스트 인식 알고리즘들로 환경에서 싸인들을 인식할 수 있다. 웨어러블 시스템은 또한 텍스트를 수정(예컨대, 디스플레이 특징들 또는 텍스트의 콘텐츠를 수정)하여 수정된 텍스트를 사용자의 물리적 환경에 렌더링할 수 있다. 예컨대, 수정된 텍스트는 오리지널 텍스트가 아닌 수정된 텍스트를 사용자가 지각하도록 오리지널 텍스트를 오버레이하고 가리도록 렌더링될 수 있다.

[0201] 텍스트의 디스플레이 특징들을 수정하는 예들

[0202] [0208] 도 16a-16e는 환경 내의 텍스트를 인식하고, 텍스트와 연관된 디스플레이 특징들을 수정하고, 수정된 텍스트를 렌더링하도록 구성된 감각 안경류 시스템에 대한 예시적인 사용자 경험들을 예시한다. 도 16a를 참조하면, 사용자(210)는 (도 16a에 도시되지 않은) 웨어러블 디바이스를 착용할 수 있고, 디스플레이(220)를 통해 환경에서 물리적 오브젝트(1606)를 볼 수 있다. 웨어러블 디바이스는 이미지(1602) 내에서 오브젝트(1606)를 포함하는 이미지(1602)를 캡처할 수 있는 외향 이미징 시스템(464)을 포함할 수 있다. 외향 이미징 시스템(464)

외에도 또는 그에 대한 대안으로, 웨어러블 시스템(200)은 다른 센서들 또는 디바이스들을 사용하여 물리적 오브젝트의 이미지를 캡처할 수 있다. 예컨대, 사용자 입력 디바이스(466)(예컨대, 토템)는 이미징 능력들을 가질 수 있고, 오브젝트(1606)의 이미지를 포함하는 이미지(1602)를 캡처할 수 있다. 오브젝트(1606)는 그 위에 또는 그 안에 묘화, 글자들, 기호들 또는 캐릭터들(1610)을 포함할 수 있는 싸인 또는 다른 오브젝트를 포함할 수 있다. 예컨대, 글자들이 오브젝트 상에 쓰여질 수 있거나; 또는 오브젝트로부터 형상화되거나, 오브젝트로 형상화되거나, 또는 오브젝트에 임베딩된다. 텍스트는 또한, 정적 또는 플래싱 광들의 시퀀스; 또는 하나 이상의 물리적 오브젝트들의 어레인지먼트일 수 있다. 도 16a-16e에 도시된 예들에서, 오브젝트(1606)는 교통 정지 표지판이다. 다른 예들에서 그리고 제한 없이, 오브젝트(1606)는 임의의 타입의 사이니지(예컨대, 상용 또는 공용 디스플레이 싸인), 책, 잡지, 한 장의 종이, 컴퓨터 디스플레이 스크린, 텔레비전 스크린 등일 수 있다.

[0203] [0209] 웨어러블 시스템(200)은 이미지(1602)를 분석하고 예컨대, 도 7을 참조하여 설명된 바와 같이, 하나 이상의 오브젝트 인식기들(708)을 사용하여 오브젝트(1606)를 인식할 수 있다. 일례로서, 웨어러블 시스템은 (예컨대, 도 16a의 팔각형인 오브젝트(1606)의 형상에 기반하여) 오브젝트(1606)가 교통 표지임을 인식할 수 있다. 다른 예로서, 웨어러블 시스템은 오브젝트(1606) 내의 텍스트의 존재를 인식할 수 있다. 웨어러블 시스템은 텍스트의 포맷에 관계없이(예컨대, 텍스트가 오브젝트 상에 있든 아니면 텍스트를 투사하는 광들의 시퀀스(예컨대, 네온 광들, LED 광들 등)로 표현되든) 텍스트를 인식할 수 있다.

[0204] [0210] 도 18을 참조하여 추가로 설명되는 바와 같이, 특정 실시예들에서, 웨어러블 시스템(200)은 텍스트의 의미를 인식할 수 있고 텍스트를 오리지널 언어로부터 타겟 언어로 변환할 수 있다. 예컨대, 웨어러블 시스템(200)은 다양한 언어들, 이를테면 예컨대, 영어, 중국어, 스페인어, 독일어, 아랍어, 힌디어 등으로부터 글자들, 기호들 또는 캐릭터들을 식별할 수 있고 텍스트를 오리지널 디스플레이된 언어에서 다른 언어로 변환할 수 있다. 일부 실시예들에서, 이러한 변환은 이전에 특정된 설정들(이를테면, 예컨대 사용자의 선호도 또는 사용자의 인구 통계학적 또는 지리적 정보)에 따라 자동으로 이루어질 수 있다. 일부 실시예들에서, 변환은 사용자로부터의 커맨드(예컨대, 구두 또는 제스처)에 대한 응답으로 이루어질 수 있다.

[0205] [0211] 웨어러블 시스템(200)은 오브젝트 인식기(708)를 사용하여 텍스트(1610)의 특징들을 분석할 수 있다. 예컨대, 웨어러블 시스템(200)은 텍스트(1610)와 연관된 폰트 사이즈 또는 활자체를 인식할 수 있다. 웨어러블 시스템은 텍스트(1610)의 특징들을 조정하여 수정된 텍스트를 생성할 수 있다. 예컨대, 웨어러블 시스템(200)은 텍스트(1610)의 사이즈를 조정하여 텍스트(1610)를 확대 또는 축소할 수 있다. 수정된 텍스트의 사이즈는 눈(210)으로부터 오리지널 텍스트(1610)까지의 거리 또는 사용자의 특징들에 부분적으로 의존할 수 있다. 예컨대, 만약 텍스트(1610)가 사용자로부터 멀리 떨어져 있다면, 웨어러블 시스템은 텍스트(1610)를 확대시킬 수 있다. 다른 예로서, 사용자의 눈 능력에 의존하여, 시스템은 텍스트의 사이즈를 어떻게 조정할지에 대한 결정을 할 수 있다. 웨어러블 시스템은 사용자로부터 이전에 획득된 정보에 기반하여 사람의 눈 능력을 결정할 수 있다. 예컨대, 사용자는 눈들에 임의의 시력 문제점들이 있는지 여부를 입력할 수 있다. 웨어러블 시스템은 또한 사용자에 대한 시력 테스트를 수행하여(예컨대, 상이한 깊이 평면들 및 사이즈들로 가상 오브젝트들을 디스플레이하여, 사용자가 가상 오브젝트들을 명확하게 지각할 수 있는지 여부를 결정함으로써) 사용자의 눈 능력을 결정할 수 있다. 사용자의 눈 능력에 기반하여, 웨어러블 시스템은 사용자가 텍스트의 특징들(예컨대, 거리/위치, 컬러, 사이즈, 폰트 등)에 기반하여 텍스트(1610)를 지각할 가능성이 있을지 여부를 결정할 수 있다. 예컨대, 만약 웨어러블 시스템이 사용자가 텍스트를 명확하게 지각할 수 없다고 결정한다면(예컨대, 텍스트가 초점에서 벗어나는 경우), 웨어러블 시스템은 텍스트를 확대시키거나 굵게 할 수 있다. 만약 사용자가 근시이지만 텍스트가 사용자로부터 멀리 떨어져 있다면, 웨어러블 시스템은 텍스트의 사이즈를 확대시킬 수 있어 사용자가 텍스트를 더 쉽게 지각할 수 있다. 사이즈 조정은 근시 정도에 대응할 수 있다. 만약 사용자의 근시 정도가 크다면 사이즈가 더 큰 증가와 연관될 수 있지만, 만약 사용자의 근시 정도가 작다면 사이즈는 더 작은 증가와 연관될 수 있다. 본원에서 추가로 설명되는 바와 같이, 웨어러블 시스템은 또한 사용자의 눈 능력에 기반하여 수정된 텍스트의 디스플레이 위치를 변경할 수 있다. 도 3을 참조하면, 디스플레이 시스템(220)은 복수의 깊이 평면들을 포함할 수 있으며, 사용자가 원시이지만 텍스트가 사용자에게 근접한 경우, 웨어러블 시스템은 수정된 텍스트가 사용자로부터 멀리 떨어져 나타나도록 오리지널의 깊이 평면보다 사용자로부터 더 멀리 떨어진 깊이 평면(306)에서 수정된 텍스트를 렌더링할 수 있다. 사이즈 조정은 (예컨대, 텍스트가 스트링으로서 인식되는 경우) 텍스트의 정면 사이즈를 변경함으로써 이루어질 수 있다. 사이즈 조정은 또한, (예컨대, 텍스트가 텍스트 스트링이 아닌 이미지로서 분석되는 경우) 텍스트(1610)를 포함하는 이미지(1602)의 부분에 대해 줌인 또는 줌아웃(예컨대, 디지털 줌)함으로써 이루어질 수 있다.

[0206] [0212] 웨어러블 시스템(200)은 사용자에 대해 수정된 텍스트를 렌더링할 수 있다. 계속해서 도 16a를 참조하

면, HMD를 착용하고 있는 사용자는 오브젝트(1606)의 렌더링된 버전(1618)을 포함하는 (HMD에 의해 렌더링된) 가상 이미지(1614)를 볼 수 있다. 일부 구현들에서, 오브젝트(1606)의 렌더링된 버전(1618)은 오리지널 텍스트를 가릴 수 있다. 도 16a에 도시된 바와 같이, 렌더링된 텍스트(1622)는 "STOP"이고 오리지널 텍스트(1610)와 비교할 때 확대된다. HMD는 확대된 텍스트(1622)를 오리지널 텍스트(1610) 상에 오버레이되도록 렌더링할 수 있고, 따라서 사용자는 오리지널 텍스트(1610)를 지각하지 못할 수 있다. 이 예에서, 텍스트 사이즈를 증가시킴으로써, 사용자는 유리하게, 실제로 훨씬 더 작고 지각하기 더 어려울 수 있는 아래에 있는 텍스트(1610)를 보다 쉽게 지각하고, 이해하고, 그에 응답할 수 있다.

[0207] [0213] 도 16b는 사용자의 환경에서 텍스트의 특징들을 수정하는 다른 예를 예시한다. 렌더링된 텍스트(1634)로 도시된 바와 같이, 웨어러블 시스템(200)은 오리지널 텍스트(1610)의 폰트를 굵게 할 수 있다. 폰트를 굵게 하는 것 외에도 또는 그에 대한 대안으로, 오리지널 텍스트(1610)에 대한 다른 변경들, 이를테면 예컨대, 텍스트 컬러 변경, 음영, 외곽선들, 포맷(예컨대, 이탤릭체, 밑줄, 정렬, 자리 맞춤 등) 등이 또한 이루어질 수 있다. 웨어러블 시스템(200)은 텍스트(1610)와 연관된 그래픽 엘리먼트들, 이를테면 렌더링된 텍스트(1634)를 플래시, 스핀 등으로 만드는 것을 추가(또는 수정)할 수 있다.

[0208] [0214] 도 16c는 초점 표시자(1640)와 함께 수정된 텍스트를 렌더링하는 예를 예시한다. 초점 표시자(1640)는 시각 효과들, 이를테면 불스 아이, 십자형, 후광, 컬러, (예컨대, 렌더링된 텍스트로 하여금 더 근접하게 나타나게 하는) 지각된 깊이 변화, 텍스트의 배경의 추가 또는 변화, 애니메이션, 또는 사용자의 주의를 끄는 다른 시각 효과들을 포함할 수 있다. 도 16c에 도시된 예에서, 웨어러블 시스템(200)은 텍스트(1638)가 렌더링되는 배경(1650)으로서 초점 표시자(1640)를 디스플레이하도록 구성될 수 있다. 배경(1650)은 경계 구역(1642) 및 내부 구역(1646)을 포함할 수 있다. 경계 구역(1642)은 내부 구역(1646)을 제한할 수 있다. 도시된 실시예에서는, 가상 글자들(1638)이 내부 구역(1646) 내에 디스플레이된다. 텍스트 배경(1650)이 디스플레이된 이미지(1614)에 렌더링될 수 있어, 텍스트 배경(1650)은 사용자가 HMD 없이 볼 배경과는 상이한 배경이 된다. 일부 실시예들에서, 내부 구역(1646) 및 경계 구역(1642) 중 하나 이상은 단색(예컨대, 백색, 흑색 또는 회색)이다. 오리지널 텍스트(1610)가 아니라 렌더링된 텍스트(1638)가 사용자에 의해 보이도록 시스템이 배경(1650)을 변경할 수 있다. 예컨대, 배경은 투명하지 않을 수 있어, 이는 오리지널 텍스트(1610)를 가릴 수 있다. 프로세싱 전자장치는 또한, 배경(1650)이 이미지(1614)의 나머지에 혼합되게 배경(1650)을 디스플레이하도록 구성될 수 있다. 예컨대, 배경(1650)은 이미지(1614)의 나머지와 동일한 컬러 및 텍스처 효과를 가질 수 있다. 웨어러블 시스템은 또한 텍스트(1638) 또는 배경(1650)을 강조 표시하는, 이를테면 예컨대, 텍스트(1638) 또는 배경(1650) 주위에 후광을 디스플레이하는 식으로 배경(1650) 및 텍스트(1638)를 디스플레이할 수 있다. 그러한 경우들에, 배경(1650)은 이미지(1614)의 나머지에 끊김없이 통합되지 않을 수 있다. 예컨대, 내부 구역(1646)은 배경(1650) 또는 텍스트(1638)를 강조하기 위해 경계 구역(1642)에 의해 윤곽이 잡힐 수 있다.

[0209] [0215] 특정 상황들 하에서, 오리지널 텍스트의 시각 외관은, 예컨대, 사용자와 오브젝트(1606) 간의 환경 효과들(예컨대, 비, 안개) 때문에 명확하지 않을 수 있다. 도 16d는 텍스트의 특징들을 수정하고 수정된 텍스트를 더 판독가능하게 되도록 렌더링하는 예를 예시한다. 이러한 도면에서, 텍스트(1626)는 사용자(210)에게 흐릿하게 나타난다. 텍스트는 다양한 이유들로 흐릿하게 지각될 수 있다. 예컨대, 나쁜 시력을 갖는 사용자는 특정 거리에서 명확하게 텍스트를 보는데 곤란함을 가질 수 있다. 근시를 갖는 사용자는 인근에 있는 텍스트의 이미지들을 비교적 명확하게 나타나는 것으로 발견하는 한편 멀리 나타난 텍스트를 흐릿한 것으로 발견할 수 있다. 유사하게, 원시를 갖는 사용자들은 멀리 나타나는 텍스트를 명확하게 볼 수 있는 한편 인근에 나타나는 텍스트에 초점을 맞추는 데 어려운 시간을 가질 수 있다. 그러나 눈 조건들은, 이미지가 흐릿하게 나타날 수 있는 유일한 이유가 아닐 수 있다. 눈(210)이 원근조절할 수 있는 것보다 더 근접하게 또는 더 멀리 나타나는 텍스트는 또한 흐릿하게 나타날 수 있다. 만약 텍스트가 사용자에 대해 신속하게 이동하고 있는 것으로 나타나면, 텍스트(1626)는 흐릿하게 나타날 수 있다. 위에서 설명된 다른 팩터들, 이를테면 기후 또는 날씨 팩터들 뿐만 아니라 이미지들을 획득한 카메라들의 해상도가 또한 역할을 할 수 있다.

[0210] [0216] 이러한 예에서, 웨어러블 시스템(200)은 흐릿한 텍스트(1626) 또는 달리 더 명확하게 또는 더 판독가능하게 판독하기 어려운 텍스트를 만들 수 있다. 텍스트가 사용자에게 흐릿하게 나타나지만 웨어러블 시스템에 의해 수신된 이미지들에서는 흐릿하지 않은 경우, 웨어러블 시스템은 도 13a를 참조하여 설명된 유사한 기법들을 사용하여 텍스트(1626)를 식별하기 위해 외향 이미징 시스템(464) 또는 다른 디바이스(이를테면, 예컨대, 사용자 입력 디바이스(466) 또는 웨어러블 시스템, 이를테면 대시 캠 외부의 카메라)에 의해 획득된 이미지들을 분석할 수 있다. 웨어러블 시스템은 텍스트(1630)에 의해 도시된 바와 같이 텍스트를 가상으로 렌더링할 수 있다. 특정 구현들에서, 웨어러블 시스템은 사용자 또는 환경의 조건들에 기반하여 가상 텍스트(1630)의 특징들

을 조정할 수 있다. 예컨대, 사용자가 근시인 경우, 웨어러블 시스템은 텍스트(1626)의 폰트를 확대할 수 있거나 사용자에게 더 근접하게 나타나도록 텍스트를 (예컨대, 더 근접한 깊이 평면 상에서) 렌더링할 수 있다. 다른 예로서, 환경이 어두울 때, 웨어러블 시스템은 텍스트(1630)와 가상 이미지(1614)의 다른 구역들 간에 콘트라스트 비율을 증가시킬 수 있다.

[0211] [0217] 일부 상황들에서, (예컨대, 빠른 구동 속도로 인해 또는 카메라의 해상도가 낮을 때) 웨어러블 시스템에 의해 획득된 이미지(1602)가 흐릿하기 때문에 텍스트(1626)는 흐릿하게 나타난다. 본원에서 설명된 바와 같이, 웨어러블 시스템은 흐릿한 텍스트(1626)의 존재를 식별하기 위해 오브젝트 인식기(708)를 사용할 수 있다. 예컨대, 웨어러블 시스템은 오브젝트(1606) 내의 또는 상의 텍스트의 존재에 대한 가능성을 결정할 수 있다. 일부 상황들에서, 만약 가능성이 임계치를 통과하면, 웨어러블 시스템은 흐릿한 텍스트(1626)에 대응할 가능성이 가장 높은 글자들(1630)을 식별하기 위해 예컨대 도 7을 참조하여 설명된 하나 이상의 텍스트 인식 알고리즘들, 이를테면 OCR 알고리즘을 사용할 수 있다.

[0212] [0218] 도 16e는 오리지널 텍스트(1610)가 방해물(1654)로 인해 부분적으로 판독불가능할 때의 시나리오들을 예시한다. 도시된 바와 같이, 방해물(1654)은 오리지널 이미지(1602)에서 오리지널 텍스트(1610)의 일부를 커버한다. 그러나, 방해물(1654)은 하나 이상의 다양한 형태들을 취할 수 있다. 예컨대, 방해물(1654)은 눈(210) 또는 디스플레이와 이미지(1602) 간의 일부 물리적 방해물, 이를테면 예컨대 폴, 빌딩 등일 수 있다. 방해물(1654)은 또한 환경 또는 날씨 방해물, 이를테면 위에서 설명된 것들일 수 있다. 방해물(1654)은 또한 오브젝트(1606) 상에 있을 수 있다(예컨대, 텍스트(1610)의 일부가 싸인(1606) 상의 다른 오브젝트에 의해 가려지거나 텍스트(1610)의 일부가 소거, 누락 또는 스티커에 의해 커버된다). 이는, 예컨대, 먼지 또는 때가 누적된 표면, 기록(1610)이 발견되는 오브젝트(1606)의 표면에 대한 손상, (예컨대, 프린터로부터의) 잉크블롯(inkblot), 오리지널 텍스트(1610)의 왜곡 또는 임의의 다른 유사한 방해물(1654)을 포함할 수 있다.

[0213] [0219] 시스템은 오리지널 텍스트(1610)가 무엇을 말하는지를 결정할 때 상황 정보(또한 때때로 본원에서 상황 단서들로 지칭됨)를 사용할 수 있다. 본원에서 설명된 다양한 상황 단서들은 텍스트(1610)에 대한 풀(full) 텍스트를 결정하기 위해 웨어러블 시스템에 의해 개별적으로 또는 조합하여 사용될 수 있다. 예시적인 상황 단서는 사용자의 위치이다. 예컨대, 앞서 설명된 바와 같이, GPS 시스템(37)(도 2b 참조)은 사용자의 위치 데이터를 획득할 수 있고 위치 데이터에 기반하여, 웨어러블 시스템은 텍스트의 언어가 무엇인지에 대한 초기 추측을 제공할 수 있다. 적용가능한 경우, 일부 실시예들에서, 웨어러블 시스템은 가시적인 스펙트럼 외부의 파장(들)(예컨대, 적외선, 자외선)에서 하나 이상의 광원들로부터 수신된 신호들로부터 부가적 정보를 획득할 수 있다. 예컨대, 웨어러블 시스템은 오직 자외선 광 하에서만 가시적인 사이니지 정보를 드러내기 위해 싸인(1606)을 향해 자외선 광을 방출할 수 있다(또는 사이니지로부터 다른 소스(예컨대, 태양)에 의해 반사된 자외선 광을 검출할 수 있다). 일부 실시예들에서, 시스템은, 시스템이 오리지널 텍스트(1610)의 가시적인 부분을 체크할 수 있는 단어들의 데이터베이스에 대한 액세스를 갖는다. 이러한 예에서, 웨어러블 시스템(200)은 글자들 또는 단어들의 어느 후보들이 가능성이 가장 높은지를 결정할 수 있다. 예컨대, 도 16e에 도시된 바와 같이, 시스템은, 부분적으로 오브젝트(1606)의 팔면체 형상 또는 오브젝트(1606)의 적색 컬러(미도시)로 인해 글자들이 "STOP"으로 스펠링된다고 추론한다.

[0214] [0220] 웨어러블 시스템은 오리지널 텍스트(1610)가 무엇을 말하는지를 결정하기 위해 주변 단어들, 심볼들, 구두점 또는 캐릭터들에 의존할 수 있다. 특정 실시예들에서, 시스템은 예컨대 기계 학습 기법들을 사용하여 위치-특정 상황 단서들을 식별할 수 있다. 예컨대, 시스템은 거리를 운전하고 있는 사용자를 검출할 수 있고 거리 싸인들 상에 빈번하게 사용되는 단어들을 향해 텍스트의 식별을 바이어싱할 수 있다. 웨어러블 시스템은, 로컬 프로세싱 및 데이터 모듈(270) 또는 원격 프로세싱 모듈(280)(예컨대, 도 2a 참조)에 의해 액세스될 수 있는 데이터베이스를 포함할 수 있다. 데이터베이스는 사용자에 의해 관련되는 특정 활동들(예컨대, 스키), 사용자의 지리적 위치들, 사용자의 이동 속도들, 사용자의 고도, 시스템에 의해 수신된 주변 잡음의 볼륨 또는 타입, 시스템에 의해 수신된 영역에서 가시적인 또는 다른 광의 레벨 또는 타입, 시스템을 둘러싸는 온도 또는 기후, 사용자로부터 텍스트의 지각된 거리, 또는 시스템이 픽업하는 다른 당사자에 의해 발화된 단어들의 카테고리 또는 카테고리들과 연관된 단어들의 카테고리들을 저장할 수 있다. 일부 실시예들에서, 웨어러블 시스템은 위에서 설명된 하나 이상의 연관성들에 따라 사용자에 의해 보이는 텍스트의 단어들 또는 언어에 대한 더 가능성이 있는 후보들을 더 정확하게 목표를 좁히기 위한 상황 단서들로서 이러한 정보를 사용할 수 있다. 일부 실시예들에서, 웨어러블 시스템은 현재 상황에 기반하여 다양한 상황들 및 아이덴티티에서 이전 단어들로부터 가능성이 있는 단어를 "학습"하기 위해 기계 학습 알고리즘들(예컨대, 딥 뉴럴 네트워크)을 사용할 수 있다. 그에 따라서, 이러한 학습을 수행함으로써, 웨어러블 시스템(200)은 사용자의 거동에 특정화될 수 있고 텍스트

를 더 신속하게 또는 효율적으로 결정할 수 있다.

[0215] [0221] 도 16a-16e에 설명된 예들에서, 시스템은, 오리지널 글자들이 사용자로부터 나타나는 지각된 거리에 기반하여 텍스트를 디스플레이하기 위한 깊이를 결정할 수 있다. 오리지널 글자들과 사용자 간의 지각된 거리는 이를테면 (예컨대, 외향 이미징 시스템에 의해 획득된 데이터에 대해) 입체 비전 알고리즘을 적용함으로써 또는 깊이 센서(예컨대, 라이다(lidar))에 의해 획득된 데이터를 분석함으로써 다양한 기법들을 사용하여 측정될 수 있다. 입체 비전 알고리즘들은 블록-매칭 알고리즘, 세미-글로벌 매칭 알고리즘, 세미-글로벌 블록-매칭 알고리즘, 디스패리티 맵들, 삼각측량, 깊이 맵들, 뉴럴 네트워크 알고리즘, 동시적인 위치 및 맵핑 알고리즘(예컨대, SLAM 또는 v-SLAM) 등을 포함할 수 있다. 사용자에게 근접하게 지각되는 글자들은 디스플레이 시스템(220) 상에서 인근 깊이에 디스플레이될 수 있다. 일부 실시예들에서, 사용자로부터 제1 거리 임계치(예컨대, 약 800cm)보다 근접하게 나타나는 글자들은 제1 깊이에서 시스템 상에 디스플레이된다. 일부 실시예들에서, 약 200cm보다 근접하게 나타나는 글자들이 제1 깊이에서 디스플레이되도록 제1 거리 임계치는 200cm이다. 일부 실시예들에서, 제1 거리 임계치는 약 80cm이다. 글자들이 제1 깊이에 있는 것처럼 디스플레이되는지 또는 어떤 제1 거리 임계치가 사용되는지 여부는 다수의 팩터들에 의존할 수 있다. 하나의 팩터는 시스템이 얼마나 많은 상이한 깊이들에서 디스플레이할 수 있는지일 수 있다. 예컨대, 만약 실시예가 오직 2개의 상이한 깊이들에서만 오브젝트들을 디스플레이하면 더 짧은 제1 거리 임계치가 사용될 수 있는 한편, 실시예가 더 많은 수의 상이한 깊이들에서 텍스트를 디스플레이할 수 있을 때 더 작은 범위가 사용될 수 있다. 예컨대, 만약 사용자가 신문을 읽고 있다면, 시스템은 사용자에게 근접한 것으로 텍스트를 지각할 것이어서, 신문 상의 글자들은 근접한 깊이에 있는 것처럼 시스템 상에 디스플레이될 것이다. 도 3에 도시된 바와 같이, 디스플레이 시스템(220)은, 가상 오브젝트들로 하여금 사용자로부터 상이한 거리들에 나타나게 할 수 있는 복수의 깊이 평면들(306)을 포함할 수 있다. 특정 구현들에서, 웨어러블 시스템은 사용자의 눈 용량에 기반하여 수정된 텍스트의 렌더링 위치를 조정할 수 있다. 예컨대, 사용자가 근시인 경우, 웨어러블 시스템은 텍스트가 본래 대응하는 깊이 평면보다 사용자에게 더 근접한 깊이 평면에서 수정된 텍스트를 렌더링할 수 있다. 다른 예로서, 사용자가 원시인 경우, 웨어러블 시스템은 오리지널 텍스트가 나타나는 것보다 사용자로부터 더 멀리 있는 깊이 평면에서 수정된 텍스트를 렌더링할 수 있다.

[0216] [0222] 유사하게, 사용자로부터 멀리 있는 것으로 지각되는 글자들은 디스플레이 시스템 상에서 먼 깊이에 디스플레이될 수 있다. 일부 실시예들에서, 사용자로부터 약 제2 거리 임계치보다 더 멀리 있는 것으로 나타나는 글자들은 제1 깊이보다 더 멀리 있는 것으로 나타나는 제2 깊이에서 시스템 상에 디스플레이된다. 일부 실시예들에서, 제2 거리 임계치는 약 300cm이다. 일부 실시예들에서, 제2 거리 임계치는 약 600cm이다. 일부 실시예들에서, 제2 거리 임계치는 약 10m이다. 예컨대, 운전 동안 게시판 상에 보이는 텍스트는 제2 깊이에서 렌더링될 수 있다.

[0217] [0223] 제1 및 제2 거리 임계치들 간의 차이는 다양한 실시예들에서 상이할 수 있다. 차이의 크기는 다수의 팩터들, 이를테면, 예컨대, 시스템이 얼마나 많은 깊이들에서 텍스트를 디스플레이할 수 있는지, 실세계 오브젝트들 또는 텍스트로부터 거리들을 지각하는 시스템의 능력의 정밀도 또는 정확도, 또는 수동 또는 공장 세팅이 무엇인지에 기반할 수 있다. 일부 실시예들에서, 차이는 100m보다 작다. 일부 실시예들에서, 차이는 700cm보다 작다. 일부 실시예들에서, 차이는 30cm보다 작다. 특정 실시예들에서, 차이는 제로이다(예컨대, 제1 거리 임계치 및 제2 거리 임계치는 동일하다).

[0218] [0224] 일부 실시예들에서, 시스템이 네거티브 차이들을 다루는 것이 가능하다. 즉, 여기서 오브젝트 또는 텍스트가 제1 깊이 및 제2 깊이에서 디스플레이되는 것 둘 모두에 대한 표준들에 맞추는 일부 겹침이 존재한다. 이러한 실시예들에서, 웨어러블 시스템은, 어느 깊이가 사용자에게 최대한 끊김없는 보기 경험을 제공할지를 결정하기 위해 상황 단서들을 사용할 수 있다. 예컨대, 초기에 사용자에게 근접하게 나타나지만 사용자로부터 멀리 신속하게 이동하고 있는 오브젝트는 초기에 제1 깊이에서 디스플레이될 표준들에 맞출 수 있다. 그러나, 시스템은 오브젝트의 위치의 궤적 때문에 오브젝트를 제2 깊이에서 디스플레이할 것이라고 결정할 수 있다.

[0219] [0225] 웨어러블 시스템의 일부 시스템들은 텍스트를 3개 이상의 깊이들에서 디스플레이할 수 있다. 이러한 경우들에서, 제1 및 제2 깊이들 간에 제3, 제4 등의 깊이들에 대응하는 거리들의 중간 거리 임계치들 또는 범위들이 포함될 수 있다. 예컨대, 일부 실시예들에서, 텍스트는, 글자들이 디스플레이(220)로부터 멀리 예컨대 약 100cm 내지 300cm 간에 있는 것으로 나타날 때 제3 깊이에서 렌더링될 수 있다.

[0220] [0226] 웨어러블 시스템(200)은 자동으로 또는 사용자 입력에 대한 응답으로 이미지로부터 텍스트를 식별 또는 인식하도록 구성될 수 있다. 텍스트가 자동으로 식별되는 실시예들에서, 사용자는 텍스트를 갖는 이미지를 볼

수 있고, 시스템은 사용자에 의한 커맨드없이 본원에서 설명된 바와 같이 텍스트를 식별하고 디스플레이할 수 있다. 텍스트가 사용자 입력에 대한 응답으로 식별되는 실시예들에서, 사용자는 다양한 커맨드들을 사용하여 텍스트의 식별 또는 디스플레이를 개시할 수 있다. 예컨대, 커맨드는 구두 큐, 손 제스처, 머리 모션(예컨대, 눈인사(nodding)), 눈 움직임(예컨대, 깜박임) 등일 수 있다.

[0221] 텍스트의 디스플레이 특징들을 수정하는 예시적인 프로세스들

[0222] [0227] 도 17은 환경과 사용자의 상호작용들을 가능하게 하기 위한 감각 안경류의 예시적인 프로세스를 예시한다. 프로세스(1700)는 웨어러블 시스템(200)의 하나 이상의 컴포넌트들에 의해(예컨대, 단독으로 또는 조합으로, 로컬 프로세싱 및 데이터 모듈(260), 원격 프로세싱 모듈(270)에 의해) 수행될 수 있다.

[0223] [0228] 블록(1704)에서, 웨어러블 시스템은 하나 이상의 카메라들을 통해 광학 신호를 수신할 수 있다. 하나 이상의 카메라들은 외향 이미징 시스템(464)의 일부이거나 또는 다른 컴퓨팅 디바이스, 이를테면 대시 캠 또는 사용자 입력 디바이스(466)의 일부일 수 있다.

[0224] [0229] 블록(1708)에서, 웨어러블 시스템은 신호로부터 이미지를 식별하는 것을 포함할 수 있다. 예컨대, 웨어러블 시스템은 광학 신호를 인간 판독가능한 이미지로 변환할 수 있다. 일부 실시예들에서, 신호로부터 이미지를 식별하는 것은 또한, 이미지의 콘텐츠를 인식하는 것, 이를테면, 예컨대, 하나 이상의 오브젝트 인식기들(708)을 사용하여 이미지 상에 광학 캐릭터 인식(OCR; optical character recognition)을 수행하는 것을 포함할 수 있다. 특정 실시예들에서, 광학 캐릭터 인식 프로세스는 하나 이상의 글자들 또는 캐릭터들의 텍스트 또는 언어에 대한 가능성이 있는 후보들을 식별하는 것을 포함한다. 광학 캐릭터 인식 프로세스는 다양한 상황 정보(예컨대, 상황 단서들)를 사용하여 인식을 수행할 수 있다. 일부 예시적인 상황 정보는 사용자 또는 사용자 부근의 누군가가 하는 활동, 사용자의 지리적 위치, 사용자의 현재 이동 속도, 사용자의 현재 고도, 시스템에 의해 검출된 주변 잡음의 볼륨 또는 타입, 디스플레이 시스템에 의해 검출된 영역에서의 가시광 또는 다른 광의 레벨 또는 타입, 디스플레이 시스템에 의해 검출된 온도 또는 기후, 사용자로부터의 캐릭터들 또는 글자들의 지각된 거리, 또는 디스플레이에 의해 검출된 단어들의 장르 또는 카테고리를 포함할 수 있다.

[0225] [0230] 계속해서 도 17을 참조하면, 프로세스(1700)는 블록(1712)에 도시된 바와 같이 이미지가 글자들 또는 캐릭터들을 포함하는지 여부를 결정하는 것을 더 포함할 수 있다. 일부 실시예들에서, 만약 프로세스(1700)가, 이미지가 글자들 또는 캐릭터들을 포함하지 않는다고 결정하면, 프로세스는 블록(1704)으로 다시 돌아갈 수 있다. 만약 프로세스(1700)가, 이미지가 글자들 또는 캐릭터들을 포함한다고 결정하면, 방법은 블록(1716)으로 계속된다.

[0226] [0231] 블록(1716)에서, 웨어러블 시스템은 글자들 또는 캐릭터들을 텍스트로 변환할 수 있다. 이는, 예컨대, (도 18 및 19를 참조하여 추가로 설명되는 바와 같이) 텍스트를 제1 언어와 상이한 제2 언어로 디스플레이하는 것을 포함할 수 있다. 일부 실시예들에서, (이미지로부터의) 하나 이상의 글자들 또는 캐릭터들을 텍스트로 변환하는 것은 사용자 또는 다른 사람으로부터 입력 또는 커맨드를 수신하는 것에 대한 응답으로 행해질 수 있다. 이러한 입력 또는 커맨드는 다양한 모드들, 이를테면, 예컨대, 구두 커맨드, 손 제스처, 머리의 모션, 또는 사용자의 하나 이상의 눈들의 움직임을 포함할 수 있다. 이들 예들은 제한하는 것으로서 여겨지지 않아야 한다.

[0227] [0232] 블록(1720)에서, 웨어러블 시스템은 사용자로부터의 복수의 깊이들 중 제1 깊이에 나타나게 텍스트를 렌더링하도록 디스플레이에 명령할 수 있다. 일부 실시예들에서, 텍스트를 디스플레이하는 것은 광학적으로 투과성인 접안렌즈를 통해 이미지로서 사용자에게 광을 송신하는 것을 포함한다. 접안렌즈는 본원에서 설명된 것들 중 임의의 것일 수 있다. 예컨대, 눈에 이미지를 형성하기 위해 광이 사용자의 눈으로 지향될 수 있다. 본원에서 설명된 바와 같이, 웨어러블 시스템은 파이버 스캐닝 투사기 또는 다른 투사기를 사용할 수 있다. 일부 실시예들에서, 방법은 (도 2b를 참조하여 설명된) GPS 시스템(37)으로부터 위치 데이터를 수신할 수 있다. 도 16a-16e를 참조하여 본원에서 추가로 설명되는 바와 같이, 이 위치 데이터는 이미지로부터 추출된 텍스트를 시스템이 추론하는 것을 돕기 위해 사용될 수 있다.

[0228] [0233] 웨어러블 시스템은 또한, 텍스트를 수정하고 수정된 텍스트를 렌더링(예컨대, 디스플레이(220)로부터의 광을 사용자의 눈들을 향해 투사)할 수 있다. 예컨대, 방법은 오리지널 글자들 또는 캐릭터들을 기준으로 상이한 폰트, 폰트 사이즈, 컬러, 배경 또는 배경 컬러, 포맷, 명확성 레벨, 언어, 또는 밝기로 텍스트를 디스플레이할 수 있다. 일부 실시예들에서, 방법은 텍스트를 애니메이팅하는 것 또는 텍스트와 상호작용하는 가상 오브젝트들을 통합하는 것을 포함할 수 있다.

[0229] 사이니지의 콘텐츠를 수정하는 예들

[0230] [0234] 텍스트의 디스플레이 특징들을 수정하는 것 외에도 또는 이에 대한 대안으로, 웨어러블 시스템은 또한, 이를테면, 예컨대, 텍스트를 한 언어로부터 다른 언어로 변환함으로써 텍스트의 콘텐츠를 수정하고, 수정된 텍스트를 디스플레이할 수 있다. 도 18은 사이니지의 콘텐츠를 수정함으로써 사용자가 물리적 환경에서 사이니지를 이해하는 것을 보조하는 예를 예시하며, 여기서, 사이니지는 로컬 언어로부터 웨어러블 시스템의 사용자가 이해할 수 있는 타겟 언어로 변환된다.

[0231] [0235] 도 18은 2개의 장면들(1800a 및 1800b)을 예시한다. 장면(1800a)은 본원에서 설명된 HMD를 착용하지 않고 사용자에 의해 지각될 수 있다. 장면(1800b)은 HMD를 착용하는 동안 (예컨대, 설명될 변환 프로세스 없이 디스플레이(220)를 통해) 사용자에 의해 지각될 수 있다. 예시된 바와 같이, 장면들(1800a 및 1800b) 둘 모두는 거리(1802) 및 보행자들(1804)을 포함한다. 장면(1800a)은 또한, 단순화된 중국어 캐릭터들을 포함하는 거리 싸인들(1810a 및 1820a)을 도시한다. 싸인(1820a)은 또한, 영어 캐릭터들을 포함한다. 그러나, HMD의 사용자(도 18에 도시되지 않음)는 영어 화자일 수 있고, 중국어 캐릭터들을 이해하지 못할 수 있다. 유리하게, 일부 실시예들에서, 웨어러블 시스템은 거리 싸인들(1810a 및 1820b) 상의 텍스트를 자동으로 인식하고, 거리 싸인들의 외국 언어 텍스트 부분을 사용자가 이해하는 언어로 변환할 수 있다. 웨어러블 시스템은 또한, 장면(1800b)에 도시된 바와 같이 물리적 싸인들 위의 가상 이미지로서, 변환된 사이니지를 제공할 수 있다. 그에 따라서, 사용자는 싸인들(1810a, 1820a)에서 중국어 텍스트를 지각하는 것이 아니라, 대신에, 싸인들(1810b, 1820b)에 도시된 영어 텍스트를 지각할 것인데, 그 이유는 HMD가 하부의 중국어 텍스트가 지각되지 않는 충분한 밝기를 이용하여 가상 이미지(영어 텍스트를 가짐)를 디스플레이하기 때문이다.

[0232] [0236] HMD(예컨대, 웨어러블 시스템(200))는 도 16a-17을 참조하여 설명된 것과 유사한 기법들을 사용하여 사용자의 환경에서 싸인을 식별하고 이 싸인을 인식할 수 있다. 일부 상황들에서, 웨어러블 시스템(200)은 싸인의 일부만을 변환하도록 구성될 수 있다. 예컨대, 웨어러블 시스템(200)은 중국어 텍스트를 갖는 싸인(1820a)의 부분만을 변환하지만, 영어 부분이 사용자에 의해 이해될 수 있기 때문에(예컨대, 그것이 사용자의 타겟 언어로 되어 있기 때문에) 영어 텍스트("GOLDSTAR")를 갖는 싸인(1820a)의 부분은 변환하지 않는다. 그러나, 사용자가 영어와 단순화된 중국어 둘 모두를 읽을 수 있도록 사용자가 이중 언어 사용자인 상황들에서, 웨어러블 시스템(200)은 싸인들(1810a 및 1820a) 상의 텍스트 중 어떤 것도 싸인들(1810b 및 1820b)로 변환하지 않도록 구성될 수 있다.

[0233] [0237] 도 16a-16e를 참조하여 설명한 바와 같이, 웨어러블 시스템(200)은 싸인들의 디스플레이 특징들을 조정하도록 구성될 수 있다. 예컨대, 싸인(1820a)의 중국어 부분의 변환으로부터 기인된 텍스트는 싸인(1820a)의 오리지널 중국어 캐릭터들보다 더 길 수 있다. 결과적으로, 웨어러블 시스템은, 렌더링된 텍스트(싸인(1820b)에 도시됨)가 오리지널 싸인의 경계 내에 맞출 수 있도록, 변환된 텍스트(예컨대, "shopping center")의 폰트 사이즈를 감소시킬 수 있다.

[0234] [0238] 도 18이 단순화된 중국어 및 영어 캐릭터들을 도시하지만, 이는 예시를 위한 것이며 제한이 아니다. 웨어러블 디스플레이 시스템(200)의 실시예들에 의해 인식 및 변환되는 언어는 임의의 언어, 이를테면, 예컨대, 영어, 중국어(단순화된 중국어 또는 통상의 중국어), 일본어, 한국어, 프랑스어, 스페인어, 독일어, 러시아어, 아랍어, 로망스 언어, 인도-유럽 언어, 시노-티베트 언어, 아시아-아프리카 언어, 히브리어, 말레이-폴리네시아 언어 등을 포함할 수 있다.

[0235] 사이니지의 콘텐츠를 수정하는 예시적인 프로세스들

[0236] [0239] 도 19는 사용자가 물리적 환경에서 사이니지를 이해하는 것을 보조하는 예시적인 프로세스를 예시한다. 예시적인 프로세스(1900)는 웨어러블 시스템(200)의 하나 이상의 컴포넌트들에 의해(예컨대, 단독으로 또는 조합으로, 로컬 프로세싱 및 데이터 모듈(260), 원격 프로세싱 모듈(270)에 의해) 수행될 수 있다.

[0237] [0240] 블록(1910)에서, 웨어러블 시스템은 사용자의 환경의 이미지들을 수신할 수 있다. 이미지들은 외향 이미징 시스템(464), 사용자 입력 디바이스(466), 또는 웨어러블 시스템의 외부에 있는 다른 디바이스 상의 카메라에 의해 캡처될 수 있다. 이미지들은 정지 이미지들, 비디오의 프레임들, 또는 비디오일 수 있다.

[0238] [0241] 블록(1920)에서, 웨어러블 시스템은 사용자의 환경에서 사이니지를 식별하기 위해 이미지들을 분석할 수 있다. 웨어러블 시스템은 오브젝트 인식기(708)를 사용하여 이러한 식별을 수행할 수 있다. 예컨대, 오브젝트 인식기(708)는 오브젝트 상의 텍스트의 존재를 검출하고 따라서 오브젝트를 싸인으로서 분류할 수 있거나, 또는 사이니지(예컨대, 도 18의 직사각형 싸인들(1810a, 1810b))의 규칙적인 경계를 인식할 수 있다.

[0239] [0242] 블록(1930)에서, 웨어러블 시스템은 사이니지 상의 텍스트를 인식할 수 있다(1930). 예컨대, 웨어러블

시스템은 싸인 상에 어떤 캐릭터들 또는 글자들이 있는지를 결정할 수 있다. 다른 예로서, 웨어러블 시스템은 텍스트가 어떤 언어로 되어 있는지를 결정할 수 있다. 웨어러블 시스템은 사용자 또는 싸인과 연관된 상황 단서들, 이를테면, 예컨대, 사용자의 위치, 구문론, 문법, 텍스트의 철자 등에 기반하여 이러한 결정을 할 수 있다. 블록(1930)에서, 웨어러블 시스템은 (예컨대, 사전을 검색함으로써) 텍스트의 의미를 추가로 결정할 수 있다.

[0240] [0243] 블록(1940)에서, 웨어러블 시스템은 텍스트의 적어도 일부를 타겟 언어로 변환할 수 있다. 타겟 언어는 사용자의 선호도 또는 사용자의 인구통계학적 정보에 기반하여 결정될 수 있다. 예컨대, 타겟 언어는 사용자의 본래의 국가와 연관된 공식 언어, 사용자의 모국어, 사용자가 가장 자주 사용하는 언어, 또는 (예컨대, 웨어러블 시스템에 대한 음성 커맨드에서 또는 다른 사용자와의 대화에서) 사용자가 발화한 언어 등일 수 있다. 타겟 언어는 또한, 사용자의 선호도에 따라 세팅될 수 있다. 예컨대, 사용자의 네이티브어가 프랑스어일지라도 사용자는 영어로 변환될 싸인들을 선호할 수 있다.

[0241] [0244] 선택적인 블록(1950)에서, 웨어러블 시스템은 텍스트와 연관된 디스플레이 특징들을 수정할 수 있다. 예컨대, 웨어러블 시스템은 텍스트(또는 텍스트와 연관된 배경)에 포커스 표시자들을 추가할 뿐만 아니라 텍스트의 폰트 사이즈 또는 컬러를 변경할 수 있다. 디스플레이 특징들의 예시적인 수정들은 도 16a-17을 참조하여 추가로 설명된다.

[0242] [0245] 선택적인 블록(1960)에서, 웨어러블 시스템은 텍스트로 하여금 혼합 현실 디바이스에 의해서 타겟 언어로 렌더링되게 할 수 있다. MR 디바이스는 본원에서 설명된 HMD일 수 있다. 디스플레이 특징들이 수정되는 경우, 웨어러블 시스템은 또한, 수정된 디스플레이 특징들로 하여금 렌더링되게 할 수 있다. 텍스트의 일부만이 타겟 언어로 변환되는 상황들에서, 웨어러블 시스템은 변환된 텍스트의 일부만을 디스플레이하거나, 또는 변환되지 않은 오리지널 텍스트의 부분과 변환된 부분 둘 모두를 디스플레이할 수 있다. 수정된 텍스트는 오리지널 텍스트가 사용자의 뷰로부터 가려질 수 있도록 물리적 사이니지 상의 오리지널 텍스트 위에 렌더링될 수 있다.

[0243] [0246] 도 18 및 19의 예들이 사이니지 상의 텍스트를 변환하는 것을 참조하여 설명되지만, 유사한 기법들이 또한, 다른 타입들의 매체들(이를테면, 책들, 텔레비전, 컴퓨터 모니터 등)에 구현되는 텍스트에 적용될 수 있다.

[0244] 수화와 관련된 추가적인 양상들

[0245] [0247] 수화에서의 감각 안경류의 애플리케이션들의 추가적인 양상들이 아래에서 추가로 제공된다.

[0246] [0248] 제1 양상에서, 증강 현실 시스템을 통해 수화로부터 변환된 텍스트를 제공하기 위한 방법으로서, 방법은: 이미징 시스템을 포함하는 AR(augmented reality) 시스템의 제어 하에: 이미지 시스템을 통해 이미지 정보를 캡처하는 단계; 이미지 정보에서 제스처들을 검출하는 단계 - 제스처들은 수화 인식을 위한 후보들임 -; 검출된 제스처들에서 수화를 인식하는 단계; 인식된 수화를 텍스트로 변환하는 단계; 및 변환된 텍스트를 디스플레이하는 단계를 포함한다.

[0247] [0249] 제2 양상에서, 양상1의 방법은: 변환된 텍스트에 관한 보조 정보에 대한 요청을 수신하는 단계; 요청된 변환된 텍스트와 연관된 보조 정보를 리트리브하는 단계; AR 시스템을 사용하여 보조 정보를 디스플레이하는 단계; 변환된 텍스트 또는 보조 정보의 디스플레이를 취소하기 위한 조건을 검출하는 단계; 및 변환된 텍스트 또는 보조 정보의 디스플레이를 취소하는 단계를 더 포함한다.

[0248] [0250] 제3 양상에서, 양상2의 방법에 있어서, 변환된 텍스트 또는 보조 정보의 디스플레이를 취소하기 위한 조건은 사용자 인터페이스 상호작용에 기반한다.

[0249] [0251] 제4 양상에서, 양상3의 방법에 있어서, 사용자 인터페이스 상호작용은 AR 시스템의 사용자의 눈 움직임들에 적어도 부분적으로 기반한다.

[0250] [0252] 제5 양상에서, 양상2 내지 양상4 중 어느 한 양상의 방법에 있어서, 변환된 텍스트 또는 보조 정보의 디스플레이를 취소하기 위한 조건은 지속기간 시간에 적어도 부분적으로 기반한다.

[0251] [0253] 제6 양상에서, 양상2 내지 양상5 중 어느 한 양상의 방법에 있어서, 변환된 텍스트 또는 보조 정보의 디스플레이를 취소하기 위한 조건은 추가적인 수화 제스처들의 변환 또는 추가적인 보조 정보의 수신에 적어도 부분적으로 기반한다.

[0252] [0254] 제7 양상에서, 양상2 내지 양상6 중 어느 한 양상의 방법은: 변환된 텍스트 또는 보조 정보의 취소된

디스플레이를 재-디스플레이하기 위한 조건을 검출하는 단계; 및 변환된 텍스트 또는 보조 정보의 취소된 디스플레이를 재-디스플레이하는 단계를 더 포함한다.

[0253] [0255] 제8 양상에서, 양상1 내지 양상7 중 어느 한 양상의 방법에 있어서, 인식된 수화를 텍스트로 변환하는 단계는 심층 학습 기법을 적용하는 단계를 포함한다.

[0254] [0256] 제9 양상에서, 양상8의 방법에 있어서, 심층 학습 기법은 뉴럴 네트워크를 포함한다.

[0255] [0257] 제10 양상에서, 양상1 내지 양상9 중 어느 한 양상의 방법에 있어서, AR 시스템은 수화 인식 및 텍스트 변환에서 수화 사전을 사용한다.

[0256] [0258] 제11 양상에서, 양상1 내지 양상10 중 어느 한 양상의 방법에 있어서, AR 시스템은 AR 시스템의 사용자에게 있어서는 외국 수화인 수화를 인식한다.

[0257] [0259] 제12 양상에서, 양상11의 방법에 있어서, AR 시스템은 후보 수화들의 리스트를 통해 작업함으로써 수화를 인식하며, 리스트는 AR 시스템의 위치에 적어도 부분적으로 기반하여 우선순위화된다.

[0258] [0260] 제13 양상에서, 양상11 내지 양상12 중 어느 한 양상의 방법에 있어서, AR 시스템은 후보 수화들의 리스트를 통해 작업함으로써 수화를 인식하며, 리스트는 AR 시스템의 환경에서 검출된 발화된 언어에 적어도 부분적으로 기반하여 우선순위화된다.

[0259] [0261] 제14 양상에서, 수화를 변환하기 위한 AR(augmented reality) 장치로서, 장치는: AR 디스플레이; 이미징 시스템; 컴퓨터-실행가능한 명령들 및 데이터를 저장하도록 구성된 데이터 저장부; 및 데이터 저장부와 통신하는 프로세서를 포함하고, 컴퓨터-실행가능한 명령들은, 실행될 때, 프로세서로 하여금: 이미징 시스템에 의해 캡처된 이미지 정보를 수신하게 하고; 수신된 이미지 또는 비디오 정보에서 제스처들을 검출하게 하고; 검출된 제스처들에서 수화를 인식하게 하고; 인식된 수화를 AR 장치의 사용자에 의해 이해되는 언어로 변환하게 하고; 그리고 AR 디스플레이를 사용하여, 변환된 수화와 연관된 정보를 디스플레이하게 한다.

[0260] [0262] 제15 양상에서, 양상14의 장치에 있어서, 컴퓨터-실행가능한 명령들은, 실행될 때 추가로, 프로세서로 하여금: 변환된 수화에 관한 보조 정보에 대한 요청을 수신하게 하고; 요청된 수화와 관련된 보조 정보를 리트리브하게 하고; 그리고 AR 디스플레이를 사용하여, 리트리브된 보조 정보를 디스플레이하게 한다.

[0261] [0263] 제16 양상에서, 양상14 내지 양상15 중 어느 한 양상의 장치에 있어서, 프로세서는, 원격 프로세서가 제스처들을 검출하고 수화를 인식하도록, 수신된 이미지 정보를 통신 네트워크를 통해 원격 프로세서에 송신함으로써, 제스처들을 검출하고 수화를 인식한다.

[0262] [0264] 제17 양상에서, 양상14 내지 양상16 중 어느 한 양상의 장치에 있어서, 이미징 시스템은 복수의 카메라들 또는 광각 카메라를 포함한다.

[0263] [0265] 제18 양상에서, 양상14 내지 양상17 중 어느 한 양상의 장치에 있어서, 프로세서는 추가로: 검출된 제스처들의 소스를 결정하고; 그리고 검출된 제스처들의 소스를 AR 장치의 사용자인 것으로 결정할 시에, 변환된 수화를 디스플레이를 위해 다른 디바이스에 송신하도록 구성된다.

[0264] [0266] 제19 양상에서, 양상14 내지 양상18 중 어느 한 양상의 장치는, 오디오 증폭기를 더 포함하고, 프로세서는 추가로, 변환된 수화를 오디오 증폭기를 통해 오디오로 제공하도록 프로그래밍된다.

[0265] [0267] 제20 양상에서, 양상19의 장치에 있어서, 프로세서는 추가로, 보조 정보를 오디오 증폭기를 통해 오디오로 제공하도록 구성된다.

[0266] [0268] 제21 양상에서, 양상14 내지 양상20 중 어느 한 양상의 장치에 있어서, AR 장치의 사용자에 의해 이해되는 언어는 인식된 수화와는 상이한 수화를 포함한다.

[0267] [0269] 제22 양상에서, 하나 이상의 수화들을 수반하는 원격 통신을 가능하게 하기 위한 AR(augmented reality) 시스템으로서, AR 시스템은: 복수의 웨어러블 AR 디바이스들 - 복수의 웨어러블 AR 디바이스들 각각은: AR 디스플레이; 이미징 시스템; 및 통신 네트워크를 통해 통신하기 위한 통신 시스템을 포함함 -; 컴퓨터-실행가능한 명령들 및 데이터를 저장하도록 구성된 하나 이상의 데이터 저장부들; 및 데이터 저장부들과 통신하는 하나 이상의 프로세서들을 포함하며, 컴퓨터-실행가능한 명령들은, 실행될 때, 하나 이상의 프로세서들이, 복수의 웨어러블 AR 디바이스들의 제1 웨어러블 AR 디바이스의 이미징 시스템에 의해 캡처된 이미지 정보를 수신하고; 수신된 이미지 정보에서 수화를 검출하고; 검출된 수화를 텍스트로 변환하고; 통신 네트워크를 통해,

변환된 텍스트를 복수의 웨어러블 AR 디바이스들의 제2 웨어러블 AR 디바이스에 송신하고; 그리고 제2 웨어러블 AR 디바이스의 AR 디스플레이 상에서, 변환된 텍스트를 디스플레이하도록, 구성한다.

[0268] [0270] 제23 양상에서, 양상22의 시스템에 있어서, 제2 웨어러블 AR 디바이스는 추가로 제1 사용자의 세계 맵을 디스플레이한다.

[0269] [0271] 제24 양상에서, 양상23의 시스템에 있어서, 제1 사용자의 세계 맵은 제1 사용자의 아바타를 포함한다.

[0270] [0272] 제25 양상에서, 양상22 내지 양상24 중 어느 한 양상의 시스템에 있어서, 복수의 웨어러블 AR 디바이스들 각각은 하나 이상의 데이터 저장부들 및 하나 이상의 프로세서들을 포함하고, 프로세서 기능성들은 로컬 프로세서에 의해 수행된다.

[0271] [0273] 제26 양상에서, 수화 인식을 위한 웨어러블 시스템으로서, 웨어러블 시스템은: 가상 콘텐츠를 사용자에게 제공하도록 구성된 머리-장착 디스플레이; 사용자의 환경을 이미징하도록 구성된 이미징 시스템; 및 머리-장착 디스플레이 및 이미징 시스템과 통신하는 하드웨어 프로세서를 포함하고, 하드웨어 프로세서는: 이미징 시스템에 의해 캡처된 이미지를 수신하도록; 오브젝트 인식기를 이용해 이미지에서 제스처를 검출하도록; 제스처의 의미를 수화로 인식하도록; 사용자와 연관된 상황 정보에 기반하여 타겟 언어를 식별하도록; 인식된 의미에 기반하여 제스처를 타겟 언어로 변환하도록; 제스처를 타겟 언어로 변환하는 것에 적어도 부분적으로 기반하여 가상 콘텐츠를 생성하도록; 그리고 머리-장착 디스플레이로 하여금 사용자에게 가상 콘텐츠를 렌더링하게 하도록 프로그래밍된다.

[0272] [0274] 제27 양상에서, 양상26의 웨어러블 시스템에 있어서, 이미징 시스템은 사용자의 주변을 이미징하도록 구성된 하나 이상의 광각 카메라들을 포함한다.

[0273] [0275] 제28 양상에서, 양상26 내지 양상27 중 어느 한 양상의 웨어러블 시스템에 있어서, 하드웨어 프로세서는 제스처와 연관된 보조 정보에 액세스하도록 추가로 프로그래밍되고; 그리고 머리-장착 디스플레이에 의해 렌더링되는 가상 콘텐츠는 보조 정보를 포함한다.

[0274] [0276] 제29 양상에서, 양상26 내지 양상28 중 어느 한 양상의 웨어러블 시스템에 있어서, 사용자와 연관된 상황 정보에 기반하여 타겟 언어를 식별하기 위해서, 하드웨어 프로세서는: 웨어러블 시스템에 의해서 캡처된 사용자의 스피치, 사용자의 위치, 또는 사용자에 의해서 이해되는 언어를 타겟 언어로서 선택하는 사용자로부터의 입력 중 적어도 하나에 기반하여, 타겟 언어를 사용자에 의해서 이해되는 언어로서 세팅하도록 프로그래밍된다.

[0275] [0277] 제30 양상에서, 양상26 내지 양상29 중 어느 한 양상의 웨어러블 시스템에 있어서, 하드웨어 프로세서는, 타겟 언어가 발화된 언어인지 여부를 결정하도록; 그리고 타겟 언어가 발화된 언어라는 결정에 대한 응답으로, 변환된 제스처와 연관된 스피치의 오디오 스트림을 타겟 언어로 플레이하도록 프로그래밍된다.

[0276] [0278] 제31 양상에서, 양상26 내지 양상29 중 어느 한 양상의 웨어러블 시스템에 있어서, 하드웨어 프로세서는, 타겟 언어가 다른 수화인지 여부를 결정하도록; 그리고 타겟 언어가 다른 수화라는 결정에 대한 응답으로, 다른 수화의 다른 제스처의 그래픽을 제스처의 변환으로서 제공하도록 프로그래밍된다.

[0277] [0279] 제32 양상에서, 양상26 내지 양상31 중 어느 한 양상의 웨어러블 시스템에 있어서, 제스처의 의미를 수화로 인식하기 위해서, 하드웨어 프로세서는 이미징 시스템에 의해 캡처된 이미지의 부분에 딥 뉴럴 네트워크 기법을 적용하도록 프로그래밍된다.

[0278] [0280] 제33 양상에서, 양상26 내지 양상32 중 어느 한 양상의 웨어러블 시스템에 있어서, 하드웨어 프로세서는 사용자의 위치에 적어도 부분적으로 기반하여 후보 수화의 리스트로부터 수화를 식별하도록 추가로 프로그래밍된다.

[0279] [0281] 제34 양상에서, 양상26 내지 양상33 중 어느 한 양상의 웨어러블 시스템에 있어서, 인식된 의미에 기반하여 제스처를 타겟 언어로 변환하기 위해서, 하드웨어 프로세서는 제스처를 타겟 언어의 텍스트 표현으로 변환하도록 프로그래밍된다.

[0280] [0282] 제35 양상에서, 양상26 내지 양상34 중 어느 한 양상의 웨어러블 시스템에 있어서, 하드웨어 프로세서는, 검출된 제스처의 소스를 결정하도록; 그리고 검출된 제스처의 소스를 웨어러블 시스템의 사용자인 것으로 결정할 시에, 타겟 언어로의 제스처의 변환을 다른 사용자의 웨어러블 디바이스에 통신하도록 프로그래밍된다.

[0281] [0283] 제36 양상에서, 양상26 내지 양상35 중 어느 한 양상의 웨어러블 시스템에 있어서, 하드웨어 프로세서는, 머리-장착 디스플레이에 의한 가상 콘텐츠의 디스플레이를 취소하는 것으로부터 조건을 검출하도록, 그리고

조건의 검출에 대한 응답으로 머리-장착 디스플레이에 의한 디스플레이로부터 가상 콘텐츠를 제거하도록 프로그래밍된다.

[0282] [0284] 제37 양상에서, 양상36의 웨어러블 시스템에 있어서, 조건은 지속기간 시간, 사용자의 손 제스처, 또는 사용자 입력 디바이스로부터의 입력 중 적어도 하나를 포함한다.

[0283] [0285] 제38 양상에서, 양상26 내지 양상37 중 어느 한 양상의 웨어러블 시스템에 있어서, 이미지는 비디오의 하나 이상의 프레임들을 포함한다.

[0284] [0286] 제39 양상에서, 수화 인식을 위한 방법으로서, 방법은: 이미징 시스템에 의해 캡처된 이미지를 수신하는 단계; 사용자의 제스처를 검출하기 위해서 이미지를 분석하는 단계; 검출된 제스처에 적어도 부분적으로 기반하여 수화로의 통신의 존재를 검출하는 단계; 제스처의 의미를 수화로 인식하는 단계; 제스처가 변환되어질 타겟 언어를 식별하는 단계; 인식된 의미에 기반하여 제스처를 타겟 언어로 변환하는 단계; 제스처를 타겟 언어로 변환하는 것에 적어도 부분적으로 기반하여 가상 콘텐츠를 생성하는 단계; 및 머리-장착 디스플레이로 하여금 사용자에게 가상 콘텐츠를 렌더링하게 하는 단계를 포함한다.

[0285] [0287] 제40 양상에서, 양상39의 방법에 있어서, 이미지는 혼합 현실 콘텐츠를 제공하도록 구성된 제1 웨어러블 디바이스로부터 수신되는 반면에, 가상 콘텐츠는 렌더링을 위해 제2 웨어러블 디바이스에 통신되고, 제1 웨어러블 디바이스 및 제2 웨어러블 디바이스는 사용자에게 혼합 현실 콘텐츠를 제공하도록 구성된다.

[0286] [0288] 제41 양상에서, 양상39의 방법에 있어서, 인식된 의미에 기반하여 제스처를 타겟 언어로 변환하는 단계는 제스처를 타겟 언어의 텍스트 표현으로 변환하는 단계를 포함한다.

[0287] [0289] 제42 양상에서, 양상39 내지 양상41 중 어느 한 양상의 방법에 있어서, 가상 콘텐츠는 타겟 언어의 텍스트 표현 또는 타겟 언어로 다른 표현을 예시하는 그래픽을 포함한다.

[0288] [0290] 제43 양상에서, 양상39 내지 양상42 중 어느 한 양상의 방법에 있어서, 제스처의 의미를 수화로 인식하는 단계는 이미징 시스템에 의해 캡처된 이미지의 부분에 딥 뉴럴 네트워크 기법을 적용하는 단계를 포함한다.

[0289] [0291] 제44 양상에서, 양상39 내지 양상43 중 어느 한 양상의 방법에 있어서, 수화로의 통신의 존재를 검출하는 단계는: 후보 수화들의 리스트로부터 수화를 식별하는 단계; 및 검출된 제스처가 수화로의 표현에 대응한다고 결정하는 단계를 포함한다.

[0290] [0292] 제45 양상에서, 양상44의 방법에 있어서, 검출된 제스처가 수화로의 표현에 대응한다고 결정하는 단계는 제스처를 행하는 사람의 입술 움직임들 및 사용자가 제스처를 행하고 있는 동안의 오디오 데이터 캡처들과 관련하여 제스처를 분석하는 단계를 포함한다.

[0291] 텍스트 수정과 관련된 추가적인 양상들

[0292] [0293] 감각 안경류에 의해 텍스트의 특징들을 수정하는 추가적인 양상들이 아래에서 추가로 설명된다.

[0293] [0294] 제1 양상에서, 증강 현실 이미지 콘텐츠를 투사하도록 구성된 머리-장착 디스플레이 디바이스로서, 디스플레이 디바이스는: 사용자의 머리 상에 착용가능하도록 구성되고 그리고 사용자의 눈 앞에서의 디스플레이를 지원하도록 구성된 프레임; 광학 신호를 수신하도록 구성된 하나 이상의 카메라들; 하나 이상의 카메라들로부터 신호를 수신하고; 신호로부터 이미지를 식별하고; 이미지가 텍스트(예컨대, 하나 이상의 글자들 또는 캐릭터들)를 포함하는지 여부를 결정하고; 텍스트를 수정된 텍스트로 변환하고; 그리고 수정된 텍스트를 렌더링하도록 디스플레이에 명령하도록 구성된 프로세싱 전자장치를 포함한다.

[0294] [0295] 제2 양상에서, 양상1의 머리-장착 디스플레이 디바이스에 있어서, 디스플레이는 하나 이상의 광원들, 및 사용자의 눈에 이미지들을 형성하기 위해 눈에 광을 지향시키도록 구성된 하나 이상의 도파관 스택들을 포함한다.

[0295] [0296] 제3 양상에서, 양상2의 머리-장착 디스플레이 디바이스에 있어서, 하나 이상의 광원들은 광을 도파관 스택들 내로 지향시키도록 구성된다.

[0296] [0297] 제4 양상에서, 양상2 내지 양상3 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 하나 이상의 광원들은 파이버 스캐닝 투사기를 포함한다.

[0297] [0298] 제5 양상에서, 양상1 내지 양상4 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 하나 이상의 카메라들은 하나 이상의 비디오 카메라들을 포함한다.

[0298] [0299] 제6 양상에서, 양상1 내지 양상5 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는 이미지 내의 하나 이상의 글자들 또는 캐릭터들을 텍스트로 변환하기 위해 광학 캐릭터 인식 알고리즘을 사용하도록 구성된다.

[0299] [0300] 제7 양상에서, 양상6의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는 하나 이상의 글자들 또는 캐릭터들의 텍스트 또는 언어에 대한 가능성이 있는 후보들을 식별하기 위해 데이터베이스에 액세스하도록 구성된다.

[0300] [0301] 제8 양상에서, 양상6 내지 양상7 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는, 사용자가 관여하는 활동, 사용자의 지리적 위치, 사용자의 이동 속도, 사용자의 고도, 디스플레이에 의해 검출된 주변 잡음의 볼륨 또는 타입, 디스플레이에 의해 검출된 영역 내의 가시광 또는 다른 광의 레벨 또는 타입, 디스플레이에 의해 검출된 온도 또는 기후, 사용자로부터의 텍스트의 지각된 거리, 또는 디스플레이에 의해 검출된 단어들의 카테고리 중 하나 이상과 연관된 입력을 수신하도록 구성된다.

[0301] [0302] 제9 양상에서, 양상1 내지 양상8 중 어느 한 양상의 머리-장착 디스플레이 디바이스는 GPS 시스템을 더 포함한다.

[0302] [0303] 제10 양상에서, 양상1 내지 양상9 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 수정된 텍스트는 텍스트의 제1 폰트 사이즈와 상이한 제2 폰트 사이즈이다. 제2 폰트 사이즈는 제1 폰트 사이즈보다 더 클 수 있다.

[0303] [0304] 제11 양상에서, 양상1 내지 양상10 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 수정된 텍스트는 텍스트보다 사용자가 더 판독가능하다.

[0304] [0305] 제12 양상에서, 양상1 내지 양상11 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는 수정된 텍스트를 형성하기 위해 텍스트에 그래픽 엘리먼트들을 부분적으로 추가하도록 구성된다.

[0305] [0306] 제13 양상에서, 양상1 내지 양상12 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는 하나 이상의 글자들 또는 캐릭터들의 제1 폰트와 상이한 제2 폰트로 텍스트의 하나 이상의 글자들 또는 캐릭터들을 디스플레이하도록 구성된다.

[0306] [0307] 제14 양상에서, 양상1 내지 양상13 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는 사용자가 머리-장착 디스플레이 없이 보는 것에 비해 텍스트의 하나 이상의 글자들 또는 캐릭터들을 확대하도록 구성된다.

[0307] [0308] 제15 양상에서, 양상1 내지 양상14 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는 경계 구역을 디스플레이하도록 구성되며, 경계 구역은 내부 구역을 제한한다.

[0308] [0309] 제16 양상에서, 양상15의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는 내부 구역 내에 하나 이상의 글자들 또는 캐릭터들을 디스플레이하도록 구성된다.

[0309] [0310] 제17 양상에서, 양상1 내지 양상16 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 프로세싱 전자장치는, 사용자가 머리-장착 디스플레이없이 하나 이상의 글자들 또는 캐릭터들을 판독하는 제1 배경과 상이한 제2 배경에 대해 텍스트의 하나 이상의 글자들 또는 캐릭터들을 디스플레이하도록 구성된다.

[0310] [0311] 제18 양상에서, 양상17의 머리-장착 디스플레이 디바이스에 있어서, 제2 배경은 단색 배경을 포함한다.

[0311] [0312] 제19 양상에서, 양상18의 머리-장착 디스플레이 디바이스에 있어서, 단색 배경은 백색을 포함한다.

[0312] [0313] 제20 양상에서, 양상17 내지 양상19 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 제1 배경은 사용자가 머리-장착 디스플레이 없이 보는 것을 포함한다.

[0313] [0314] 제21 양상에서, 양상1 내지 양상20 중 어느 한 양상의 머리-장착 디스플레이 디바이스에 있어서, 텍스트는 텍스트 편집기(text editor)에 의해 편집가능하도록 적응된다.

[0314] [0315] 양상1 내지 양상21이 머리-장착 디스플레이를 참조하여 설명되었지만, 이러한 양상들에서 설명되는 유사한 기능들은 또한, 도 2a를 참조하여 설명된 머리-장착 디바이스 또는 웨어러블 시스템으로 구현될 수 있다. 추가로, 디스플레이는 복수의 깊이 평면들을 포함할 수 있고, 머리-장착 디바이스는 사용자의 눈의 능력에 적어도 부분적으로 기반하여, 수정된 텍스트를 렌더링하기 위해 깊이 평면을 식별하도록 구성된다.

[0315] [0316] 제22 양상에서, 머리-장착 디스플레이를 사용하여 증강 현실 이미지 콘텐츠를 투사하기 위한 방법으로서, 방법은: 하드웨어 프로세서의 제어 하에: 하나 이상의 카메라들로부터 광학 신호를 수신하는 단계; 광학 캐릭터 인식 모듈을 사용하여, 신호로부터 이미지를 식별하는 단계; 이미지가 하나 이상의 글자들 또는 캐릭터들을 포함하는지 여부를 결정하는 단계; 하나 이상의 글자들 또는 캐릭터들을 텍스트로 변환하는 단계; 및 머리-장착 디스플레이 상에서 텍스트를 디스플레이하는 단계를 포함하며, 텍스트를 디스플레이하는 단계는 광학적으로 투과성의 접안렌즈를 통해 이미지로서 광을 사용자에게 송신하는 단계를 포함한다.

[0316] [0317] 제23 양상에서, 양상22의 방법은, 사용자의 눈에 이미지들을 형성하기 위해 눈으로 광을 지향시키는 단계를 더 포함한다.

[0317] [0318] 제24 양상에서, 양상22 내지 양상23 중 어느 한 양상의 방법은, 파이버 스캐닝 투사기를 사용하여 접안렌즈로 광을 지향시키는 단계를 더 포함한다.

[0318] [0319] 제25 양상에서, 양상22 내지 양상24 중 어느 한 양상의 방법에 있어서, 광학 캐릭터 인식 모듈을 사용하는 단계는 하나 이상의 글자들 또는 캐릭터들의 텍스트 또는 언어에 대한 가능성이 있는 후보들을 식별하는 단계를 포함한다.

[0319] [0320] 제26 양상에서, 양상22 내지 양상25 중 어느 한 양상의 방법에 있어서, 광학 캐릭터 인식 모듈을 사용하는 단계는 사용자가 관여하는 활동, 사용자의 지리적 위치, 사용자의 이동 속도, 사용자의 고도, 디스플레이에 의해 검출된 주변 잡음의 볼륨 또는 타입, 디스플레이에 의해 검출된 영역 내의 가시광 또는 다른 광의 레벨 또는 타입, 디스플레이에 의해 검출된 온도 또는 기후, 사용자로부터의 하나 이상의 글자들 또는 캐릭터들의 지각된 거리, 또는 디스플레이에 의해 검출된 단어들의 카테고리 중 하나 이상과 연관된 정보를 포함하는 입력을 수신하는 단계를 포함한다.

[0320] [0321] 제27 양상에서, 양상22 내지 양상26 중 어느 한 양상의 방법에 있어서, 하나 이상의 글자들 또는 캐릭터들을 텍스트로 변환하는 단계는 하나 이상의 글자들 또는 캐릭터들과 연관된 제1 언어와 상이한 제2 언어로 텍스트를 디스플레이하는 단계를 포함한다. 제27 양상의 일부 구현들에서, 방법은 텍스트를 제2 언어로 변환하는 단계를 포함한다.

[0321] [0322] 제28 양상에서, 양상22 내지 양상27 중 어느 한 양상의 방법은, GPS 시스템으로부터 위치 데이터를 수신하는 단계를 더 포함한다.

[0322] [0323] 제29 양상에서, 양상22 내지 양상28 중 어느 한 양상의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 하나 이상의 글자들 또는 캐릭터들의 제1 폰트 사이즈와 상이한 제2 폰트 사이즈로 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계를 포함한다.

[0323] [0324] 제30 양상에서, 양상22 내지 양상29 중 어느 한 양상의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 머리-장착 디스플레이가 없는 경우보다 사용자가 더 판독가능하게 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계를 포함한다.

[0324] [0325] 제31 양상에서, 양상22 내지 양상30 중 어느 한 양상의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 머리-장착 디스플레이가 없는 사용자에게 나타나는 것보다 더 큰 폰트 사이즈로 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계를 포함한다.

[0325] [0326] 제32 양상에서, 양상22 내지 양상31 중 어느 한 양상의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 하나 이상의 글자들 또는 캐릭터들의 제1 폰트와 상이한 제2 폰트로 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계를 포함한다.

[0326] [0327] 제33 양상에서, 양상22 내지 양상32 중 어느 한 양상의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 사용자가 머리-장착 디스플레이 없이 보는 것에 비해 하나 이상의 글자들 또는 캐릭터들을 확대하는 단계를 포함한다.

[0327] [0328] 제34 양상에서, 양상22 내지 양상33 중 어느 한 양상의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 경계 구역을 디스플레이하는 단계를 포함하며, 경계 구역은 내부 구역을 제한한다.

[0328] [0329] 제35 양상에서, 양상34의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 내부 구역 내에 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계를 포

함한다.

[0329] [0330] 제36 양상에서, 양상22 내지 양상35 중 어느 한 양상의 방법에 있어서, 머리-장착 디스플레이 상에서 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계는 사용자가 머리-장착 디스플레이 없이 하나 이상의 글자들 또는 캐릭터들을 판독하는 제1 배경과 상이한 제2 배경에 대해 하나 이상의 글자들 또는 캐릭터들을 디스플레이하는 단계를 포함한다.

[0330] [0331] 제37 양상에서, 양상36의 방법에 있어서, 제2 배경은 단색 배경을 포함한다.

[0331] [0332] 제38 양상에서, 양상37의 방법에 있어서, 단색 배경은 백색을 포함한다.

[0332] [0333] 제39 양상에서, 양상36 내지 양상38 중 어느 한 양상의 방법에 있어서, 제1 배경은 사용자가 머리-장착 디스플레이 없이 보는 것을 포함한다.

[0333] [0334] 제40 양상에서, 양상22 내지 양상39 중 어느 한 양상의 방법에 있어서, 텍스트는 텍스트 편집기에 의해 편집가능하도록 적응된다.

[0334] [0335] 제41 양상에서, 양상22 내지 양상40 중 어느 한 양상의 방법에 있어서, 하나 이상의 글자들 또는 캐릭터들을 텍스트로 변환하는 단계는 사용자로부터 입력을 수신하는 단계를 포함한다.

[0335] [0336] 제42 양상에서, 양상41의 방법에 있어서, 사용자로부터 입력을 수신하는 단계는 구두 커맨드, 손 제스처, 머리의 모션, 또는 사용자의 눈들 중 하나 이상의 눈들의 움직임 중 하나 이상을 수신하는 단계를 포함한다.

[0336] [0337] 제43 양상에서, 양상22 내지 양상42 중 어느 한 양상의 방법에 있어서, 하나 이상의 글자들 또는 캐릭터들이 제1 거리 임계치보다 더 근접한 것으로 나타난다면, 텍스트는 제2 깊이보다 더 근접한 것으로 나타나는 제1 깊이에 디스플레이된다.

[0337] [0338] 제44 양상에서, 양상22 내지 양상43 중 어느 한 양상의 방법에 있어서, 하나 이상의 글자들 또는 캐릭터들이 제2 거리 임계치보다 더 먼 것으로 나타난다면, 텍스트는 제1 깊이보다 더 먼 것으로 나타나는 제2 깊이에 디스플레이된다.

[0338] [0339] 제45 양상에서, 양상43 내지 양상44 중 어느 한 양상의 방법에 있어서, 하나 이상의 글자들 또는 캐릭터들이 제1 거리 임계치보다 더 멀고 그리고 제2 거리 임계치보다 더 근접한 것으로 나타난다면, 텍스트는 제1 깊이보다 더 멀지만 제2 깊이보다 더 근접하게 나타나는 제3 깊이에 디스플레이된다.

[0339] [0340] 제46 양상에서, 양상43 내지 양상45 중 어느 한 양상의 방법에 있어서, 제1 거리 임계치는 80 cm이다.

[0340] [0341] 제47 양상에서, 양상43 내지 양상46 중 어느 한 양상의 방법에 있어서, 제2 거리 임계치는 600 cm이다.

[0341] [0342] 제48 양상에서, 양상43 내지 양상47 중 어느 한 양상의 방법에 있어서, 제2 거리 임계치와 제1 거리 임계치 간의 차이는 100 m 미만이다.

[0342] 사이니지 수정과 관련된 추가적인 양상들

[0343] [0343] 제1 양상에서, 증강 현실 시스템은: 외향 이미징 시스템; 외향 이미징 시스템에 의해 획득된 이미지들을 저장하도록 구성된 비-일시적 메모리; 및 하드웨어 프로세서를 포함하며, 하드웨어 프로세서는: 외향 이미징 시스템에 의해 획득된 증강 현실 시스템의 사용자의 환경의 이미지들을 수신하고; 사용자의 환경의 사이니지를 식별하기 위해 이미지들을 분석하고; 사이니지 상의 텍스트를 인식하고; 텍스트의 적어도 일부를 타겟 언어로 변환하고; 그리고 변환된 텍스트를 사용자에게 렌더링하도록 디스플레이에 명령하도록 프로그래밍된다.

[0344] [0344] 제2 양상에서, 양상1의 증강 현실 시스템에 있어서, 하드웨어 프로세서는 텍스트와 연관된 디스플레이 특징들을 수정하도록 프로그래밍된다.

[0345] [0345] 제3 양상에서, 양상1 또는 양상2의 증강 현실 시스템에 있어서, 텍스트의 적어도 일부를 타겟 언어로 변환하기 위해, 하드웨어 프로세서는 사이니지 상의 텍스트의 언어를 식별하고 그리고 언어를 타겟 언어로 변환하도록 프로그래밍된다.

[0346] [0346] 제4 양상에서, 양상1 내지 양상3 중 어느 한 양상의 증강 현실 시스템에 있어서, 하드웨어 프로세서는 사용자의 위치에 적어도 부분적으로 기반하여 타겟 언어를 결정하도록 프로그래밍된다.

[0347] [0347] 제5 양상에서, 양상1 내지 양상4 중 어느 한 양상의 증강 현실 시스템에 있어서, 사이니지 상의 텍스트

를 인식하기 위해, 하드웨어 프로세서는 타겟 언어로 되어 있는 텍스트를 인식하도록 프로그래밍된다.

[0348] [0348] 제6 양상에서, 양상5의 증강 현실 시스템에 있어서, 하드웨어 프로세서는 타겟 언어로 되어 있는 텍스트를 변환하지 않도록 프로그래밍된다.

[0349] 다른 고려사항들

[0350] [0349] 본원에서 설명되거나 또는 첨부 도면들에 묘사된 프로세스들, 방법들 및 알고리즘들 각각은 하나 이상의 물리적 컴퓨팅 시스템들, 하드웨어 컴퓨터 프로세서들, 주문형 회로 또는 특정 그리고 특별 컴퓨터 명령들을 실행하도록 구성된 전자 하드웨어에 의해 실행되는 코드 모듈들로 구현되고, 그리고 이 코드 모듈들에 의해 완전히 또는 부분적으로 자동화될 수 있다. 예컨대, 컴퓨팅 시스템들은 특정 컴퓨터 명령들 또는 특수 목적 컴퓨터들, 특수 목적 회로 등으로 프로그래밍된 범용 컴퓨터들(예컨대, 서버들)을 포함할 수 있다. 코드 모듈은 실행가능 프로그램으로 컴파일링되고 링크되거나, 동적 링크 라이브러리에 설치되거나, 또는 해석형 프로그래밍 언어로 쓰여질 수 있다. 일부 구현들에서, 특정 동작들 및 방법들은 정해진 기능에 특정적인 회로에 의해 수행될 수 있다.

[0351] [0350] 추가로, 본 개시내용의 기능성의 특정 구현들은 충분히 수학적으로, 계산상으로 또는 기술적으로 복잡하여, (적합한 전문화된 실행가능한 명령들을 활용하는) 주문형 하드웨어 또는 하나 이상의 물리적 컴퓨팅 디바이스들은 예컨대 수반된 계산들의 볼륨 또는 복잡성으로 인해 또는 실질적으로 실시간으로 결과들을 제공하기 위해 기능성을 수행할 필요가 있을 수 있다. 예컨대, 애니메이션들 또는 비디오는 많은 프레임들(각각의 프레임은 수 백만개의 픽셀들을 가짐)을 포함할 수 있고, 그리고 상용적으로 합리적인 시간 양에서 원하는 이미지 프로세싱 임무 또는 애플리케이션을 제공하기 위해, 특별하게 프로그래밍된 컴퓨터 하드웨어가 비디오 데이터를 프로세싱할 필요가 있다.

[0352] [0351] 코드 모듈들 또는 임의의 타입의 데이터는 임의의 타입의 비-일시적 컴퓨터-판독가능한 매체, 이를테면 하드 드라이브들, 고체 상태 메모리, RAM(random access memory), ROM(read only memory), 광학 디스크, 휘발성 또는 비-휘발성 스토리지, 이들의 조합들 등을 포함하는 물리적 컴퓨터 스토리지 상에 저장될 수 있다. 방법들 및 모듈들(또는 데이터)은 또한 생성된 데이터 신호들로서(예컨대, 반송파 또는 다른 아날로그 또는 디지털 전파 신호의 부분으로서) 무선-기반 및 유선/케이블-기반 매체들을 포함하는 다양한 컴퓨터-판독가능한 송신 매체들 상에서 송신될 수 있고, 그리고 (예컨대, 단일 또는 멀티플렉싱된 아날로그 신호의 부분으로서, 또는 다수의 이산 디지털 패킷들 또는 프레임들로서) 다양한 형태들을 취할 수 있다. 개시된 프로세스들 또는 프로세스 단계들의 결과들은 임의의 타입의 비-일시적, 유형의 컴퓨터 스토리지에 영구적으로 또는 다른 방식으로 저장될 수 있거나 또는 컴퓨터-판독가능한 송신 매체를 통해 통신될 수 있다.

[0353] [0352] 본원에서 설명되거나 또는 첨부 도면들에 묘사된 흐름도들에서 임의의 프로세스들, 블록들, 상태들, 단계들, 또는 기능성들은 (예컨대, 논리적 또는 산술적) 특정한 기능들 또는 프로세스의 단계들을 구현하기 위한 하나 이상의 실행가능한 명령들을 포함하는 코드 모듈들, 세그먼트들 또는 코드의 부분들을 잠재적으로 표현하는 것으로 이해되어야 한다. 다양한 프로세스들, 블록들, 상태들, 단계들 또는 기능성들은 본원에서 제공된 예시적인 예들로부터 조합되거나, 재배열되거나, 추가되거나, 삭제되거나, 수정되거나 다르게 변경될 수 있다. 일부 실시예들에서, 추가적인 또는 상이한 컴퓨팅 시스템들 또는 코드 모듈들은 본원에서 설명된 기능성들 중 일부 또는 모두를 수행할 수 있다. 본원에서 설명된 방법들 및 프로세스들은 또한 임의의 특정 시퀀스로 제한되지 않고, 이에 관련된 블록들, 단계들 또는 상태들은 적합한 다른 시퀀스들로, 예컨대 직렬로, 병렬로, 또는 일부 다른 방식으로 수행될 수 있다. 임무들 또는 이벤트들은 개시된 예시적인 실시예들에 추가되거나 이들로부터 제거될 수 있다. 게다가, 본원에서 설명된 구현들에서 다양한 시스템 컴포넌트들의 분리는 예시 목적들을 위한 것이고 모든 구현들에서 그런 분리를 요구하는 것으로 이해되지 않아야 한다. 설명된 프로그램 컴포넌트들, 방법들 및 시스템들이 일반적으로 단일 컴퓨터 제품으로 함께 통합되거나 다수의 컴퓨터 제품들로 패키징될 수 있다는 것이 이해되어야 한다. 많은 구현 변동들이 가능하다.

[0354] [0353] 프로세스들, 방법들 및 시스템들은 네트워크(또는 분산형) 컴퓨팅 환경에서 구현될 수 있다. 네트워크 환경들은 전사적 컴퓨터 네트워크들, 인트라넷들, LAN(local area network)들, WAN(wide area network)들, PAN(personal area network)들, 클라우드 컴퓨팅 네트워크들, 크라우드-소스드(crowd-sourced) 컴퓨팅 네트워크들, 인터넷, 및 월드 와이드 웹(World Wide Web)을 포함한다. 네트워크는 유선 또는 무선 네트워크 또는 임의의 다른 타입의 통신 네트워크일 수 있다.

[0355] [0354] 본 개시내용의 시스템들 및 방법들 각각은 몇몇 혁신적인 양상들을 가지며, 이 양상들 중 어떤 단일의

양상이 본원에서 개시된 바람직한 속성들을 위해 전적으로 책임지거나 요구되지 않는다. 위에서 설명된 다양한 특징들 및 프로세스들은 서로 독립적으로 사용될 수 있거나, 또는 다양한 방식들로 조합될 수 있다. 모든 가능한 조합들 및 서브조합들은 본 개시내용의 범위 내에 속하도록 의도된다. 본 개시내용에 설명된 구현들에 대한 다양한 수정들은 당업자들에게 쉽게 자명할 수 있고, 그리고 본원에서 정의된 일반적인 원리들은 본 개시내용의 사상 또는 범위를 벗어나지 않고 다른 구현들에 적용될 수 있다. 따라서, 청구항들은 본원에 도시된 구현들 또는 실시예들로 제한되는 것으로 의도되는 것이 아니라, 본 개시내용, 본원에서 개시된 원리들 및 신규 특징들과 일치하는 가장 넓은 범위에 부합될 것이다.

[0356] [0355] 별개의 구현들 또는 실시예들의 맥락에서 이 명세서에 설명된 특정 특징들은 또한 단일 구현 또는 실시예로 조합하여 구현될 수 있다. 대조적으로, 단일 구현 또는 실시예의 맥락에서 설명된 다양한 특징들은 또한 별도로 다수의 구현들 또는 실시예들로 또는 임의의 적절한 서브조합으로 구현될 수 있다. 게다가, 비록 특징들이 특정 조합들로 작용하는 것으로서 위에서 설명될 수 있고, 심지어 그와 같이 처음에 청구될 수 있지만, 청구된 조합으로부터의 하나 이상의 특징들은 일부 경우들에서 조합으로부터 제거될 수 있고, 그리고 청구된 조합은 서브조합 또는 서브조합의 변형으로 지향될 수 있다. 단일 특징 또는 특징들의 그룹이 각각의 모든 실시예에 필요하거나 필수적인 것은 아니다.

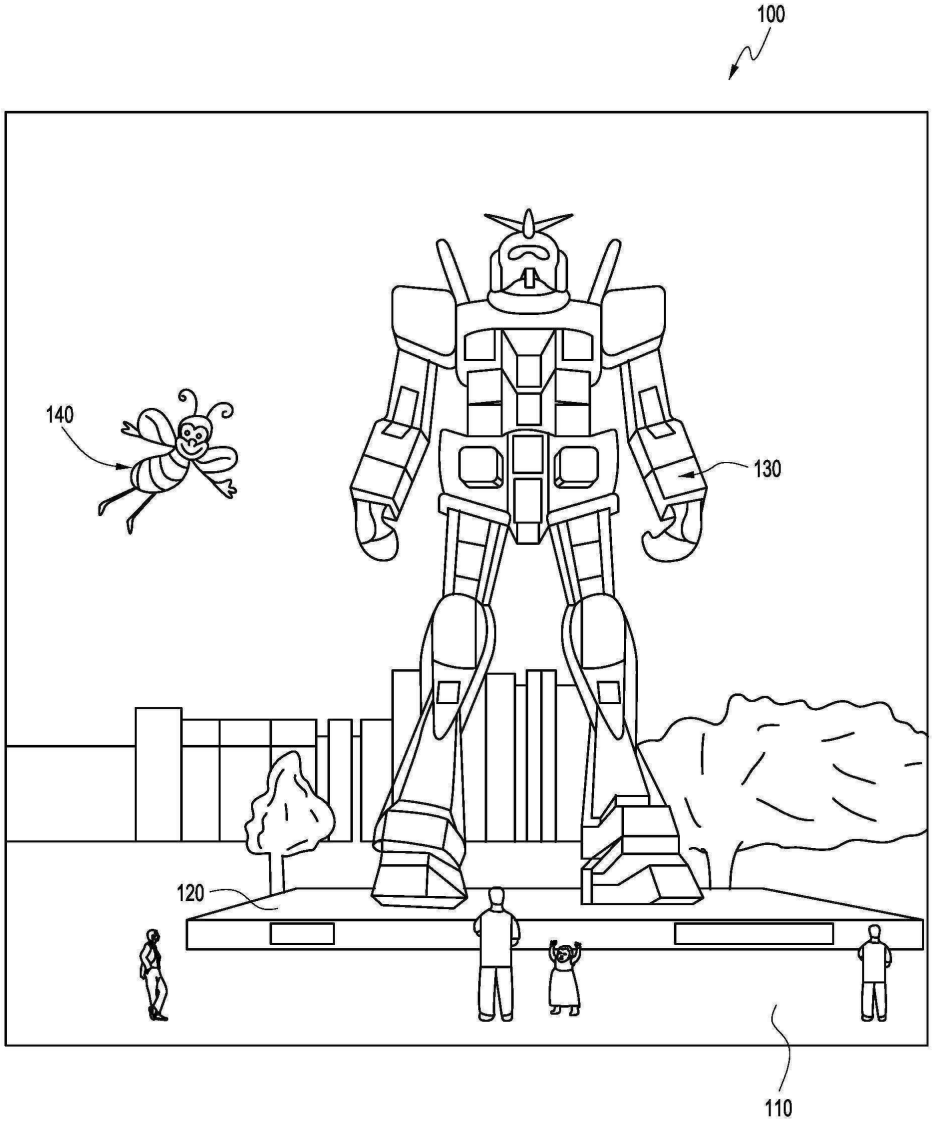
[0357] [0356] 구체적으로 다르게 언급되지 않거나, 사용된 맥락 내에서 다르게 이해되지 않으면, 본원에서 사용된 조건어, 이를테면 특히, "할 수 있다("can", "could", "might", "may")" 및 "예컨대" 등은, 일반적으로 특정 실시예들이 특정 특징들, 엘리먼트들 또는 단계들을 포함하지만, 다른 실시예들은 이들을 포함하지 않는다는 것을 전달하도록 의도된다. 따라서, 그런 조건어는 일반적으로, 특징들, 엘리먼트들 또는 단계들이 어쨌든 하나 이상의 실시예들을 위해 요구된다는 것, 또는 하나 이상의 실시예들이, 저자 입력 또는 프롬프팅을 사용하여 또는 이들을 사용하지 않고, 이들 특징들, 엘리먼트들 또는 단계들이 임의의 특정 실시예에 포함되는지 아니면 이 임의의 특정 실시예에서 수행되는지를 판단하기 위한 로직을 반드시 포함하는 것을 암시하도록 의도되지 않는다. 용어 "포함하는(comprising)", "포함하는(including)", "갖는(having)" 등은 동의어이고 오픈-엔디드(open-ended) 방식으로 포괄적으로 사용되고, 그리고 추가적인 엘리먼트들, 특징들, 작용들, 동작들 등을 배제하지 않는다. 또한, 용어 "또는"은 포괄적인 의미(및 이의 배타적 의미가 아님)로 사용되어, 예컨대 엘리먼트들의 리스트를 연결하기 위해 사용될 때, 용어 "또는"은 리스트 내 엘리먼트들 중 하나, 일부 또는 모두를 의미한다. 게다가, 본 출원 및 첨부된 청구항들에 사용된 단수들은 다르게 특정되지 않으면 "하나 이상" 또는 "적어도 하나"를 의미하는 것으로 이해될 것이다.

[0358] [0357] 본원에서 사용된 바와 같이, 아이템들의 리스트 중 "적어도 하나"를 지칭하는 구절은, 단일 멤버들을 포함하여, 이들 아이템들의 임의의 조합을 지칭한다. 예로서, "A, B 또는 C 중 적어도 하나"는 "A, B, C; A 및 B; A 및 C; B 및 C; 및 A, B 및 C를 커버하도록 의도된다. 구체적으로 다르게 언급되지 않으면, "X, Y 및 Z 중 적어도 하나"라는 구절 같은 결합 어구는, 일반적으로 아이템, 용어 등이 X, Y 또는 Z 중 적어도 하나일 수 있다는 것을 전달하기 위해 사용되는 상황으로 이해된다. 따라서, 그런 접속어는 일반적으로, 특정 실시예들이 X 중 적어도 하나, Y 중 적어도 하나 및 Z 중 적어도 하나가 각각 존재할 것을 요구하는 것을 암시하도록 의도되지 않는다.

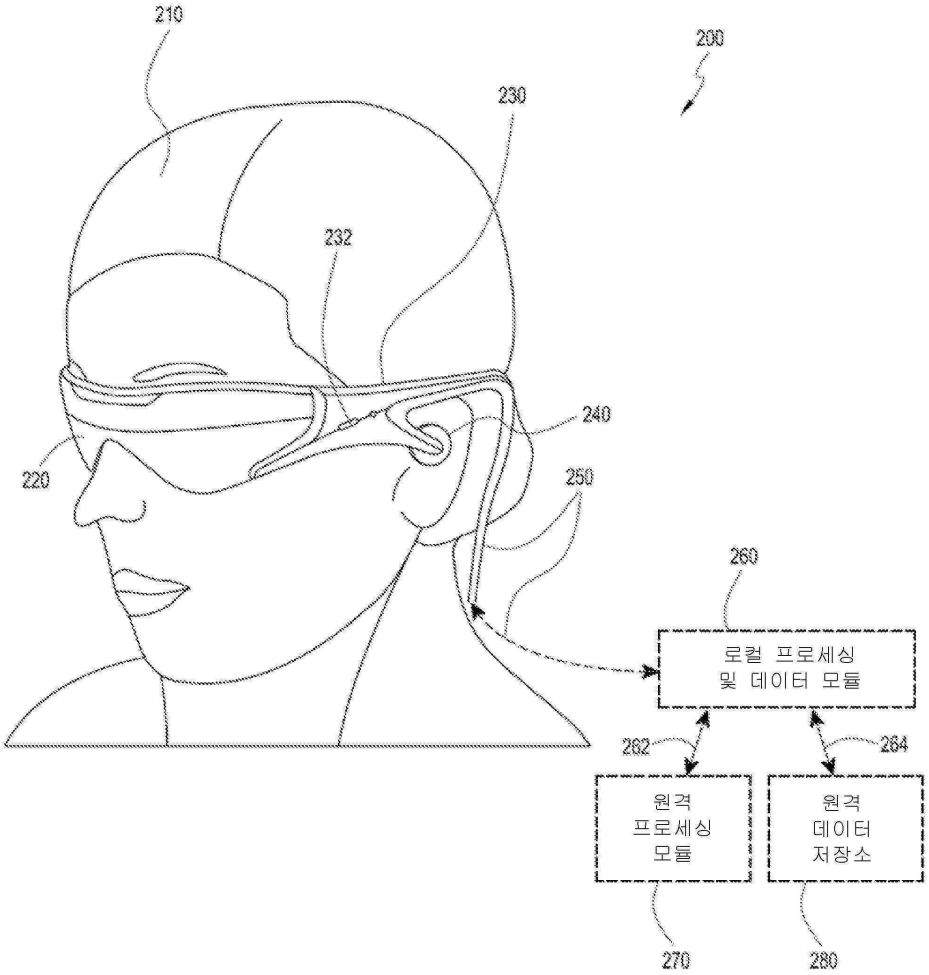
[0359] [0358] 유사하게, 동작들이 특정 순서로 도면들에 묘사될 수 있지만, 원하는 결과들을 달성하기 위해, 그런 동작들이 도시된 특정 순서 또는 순차적 순서로 수행될 필요가 없거나, 또는 모든 예시된 동작들이 수행될 필요가 없다는 것이 인식되어야 한다. 추가로, 도면들은 순서도 형태로 하나 이상의 예시적인 프로세스들을 개략적으로 묘사할 수 있다. 그러나, 묘사되지 않은 다른 동작들은 개략적으로 예시된 예시적인 방법들 및 프로세스들에 통합될 수 있다. 예컨대, 하나 이상의 추가적인 동작들은 예시된 동작들 중 임의의 동작 이전, 이후, 동시에, 또는 중간에 수행될 수 있다. 추가적으로, 동작들은 다른 구현들에서 재배열되거나 재정렬될 수 있다. 특정 상황들에서, 멀티태스킹 및 병렬 프로세싱이 유리할 수 있다. 게다가, 위에서 설명된 구현들에서 다양한 시스템 컴포넌트들의 분리는 모든 구현들에서 그런 분리를 요구하는 것으로 이해되지 않아야 하고, 그리고 설명된 프로그램 컴포넌트들 및 시스템들이 일반적으로 단일 소프트웨어 제품으로 함께 통합될 수 있거나 다수의 소프트웨어 제품들로 패키징될 수 있다는 것이 이해되어야 한다. 추가적으로, 다른 구현들은 다음의 청구항들의 범위 내에 있다. 일부 경우들에서, 청구항들에서 언급된 액션들은 상이한 순서로 수행될 수 있고 여전히 원하는 결과들을 달성할 수 있다.

도면

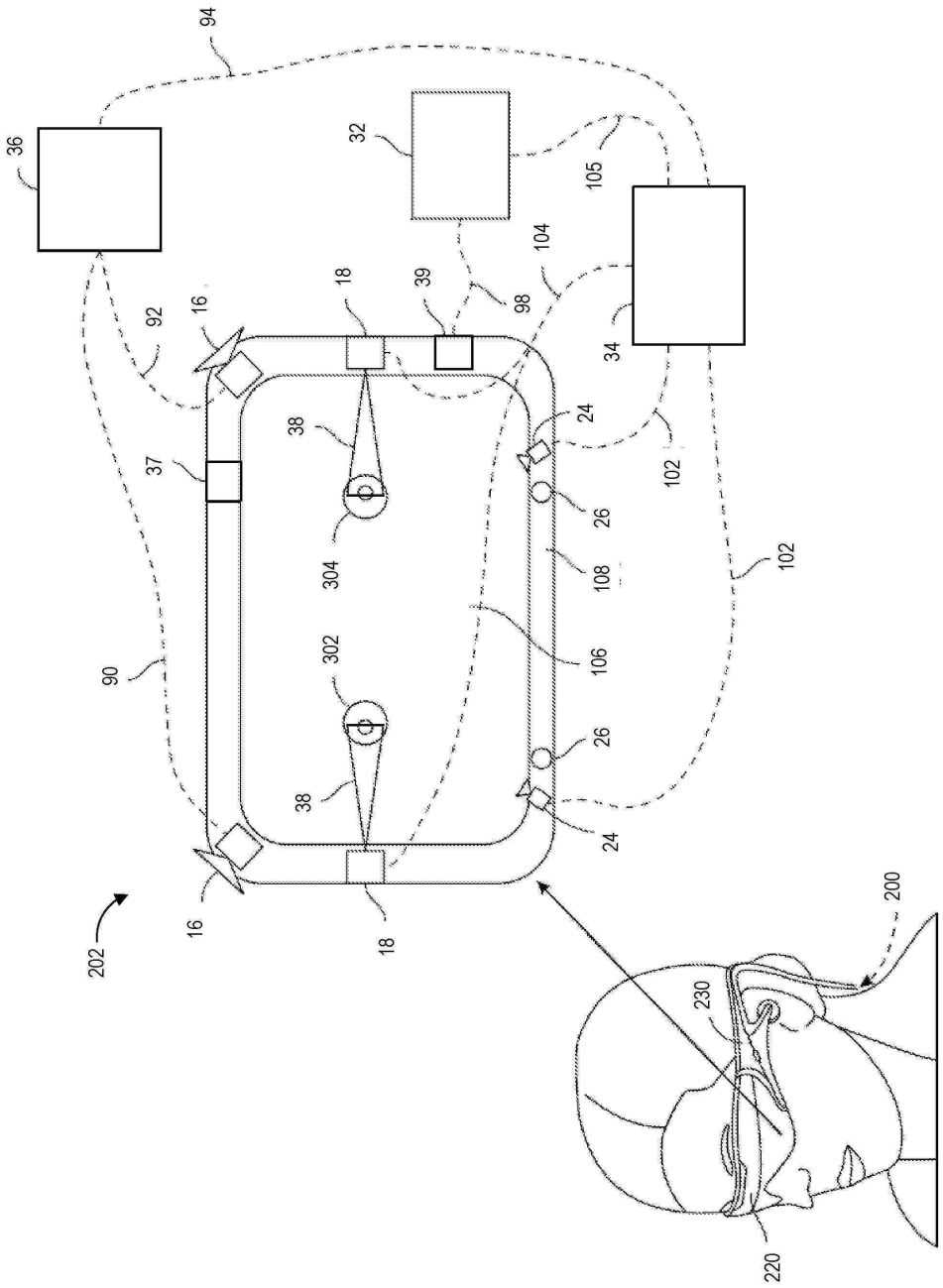
도면1



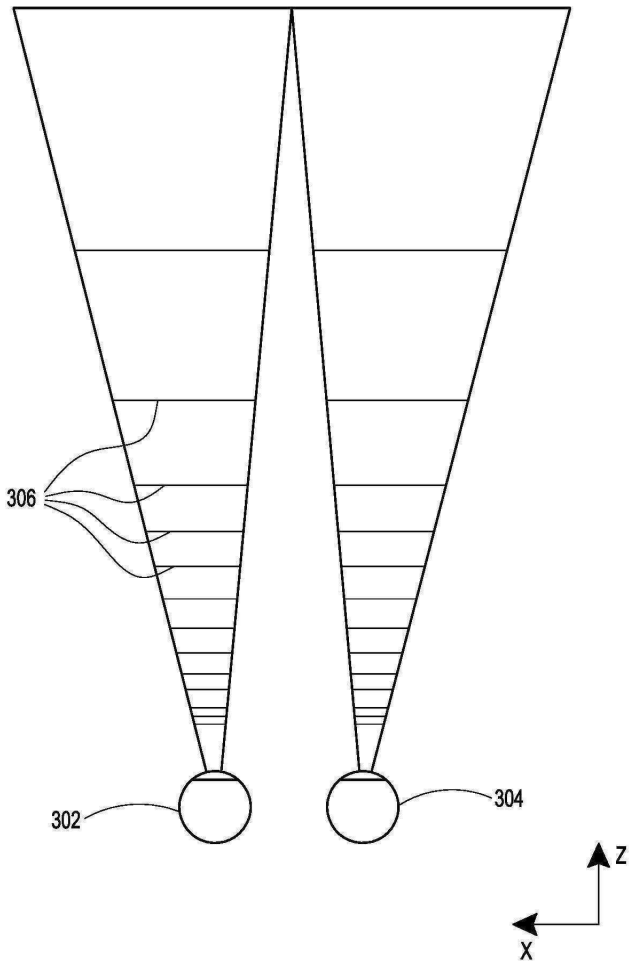
도면2a


210
200
230
232
240
220
250
260
로컬 프로세싱
및 데이터 모듈
262
264
원격
프로세싱
모듈
원격
데이터
저장소
270
280

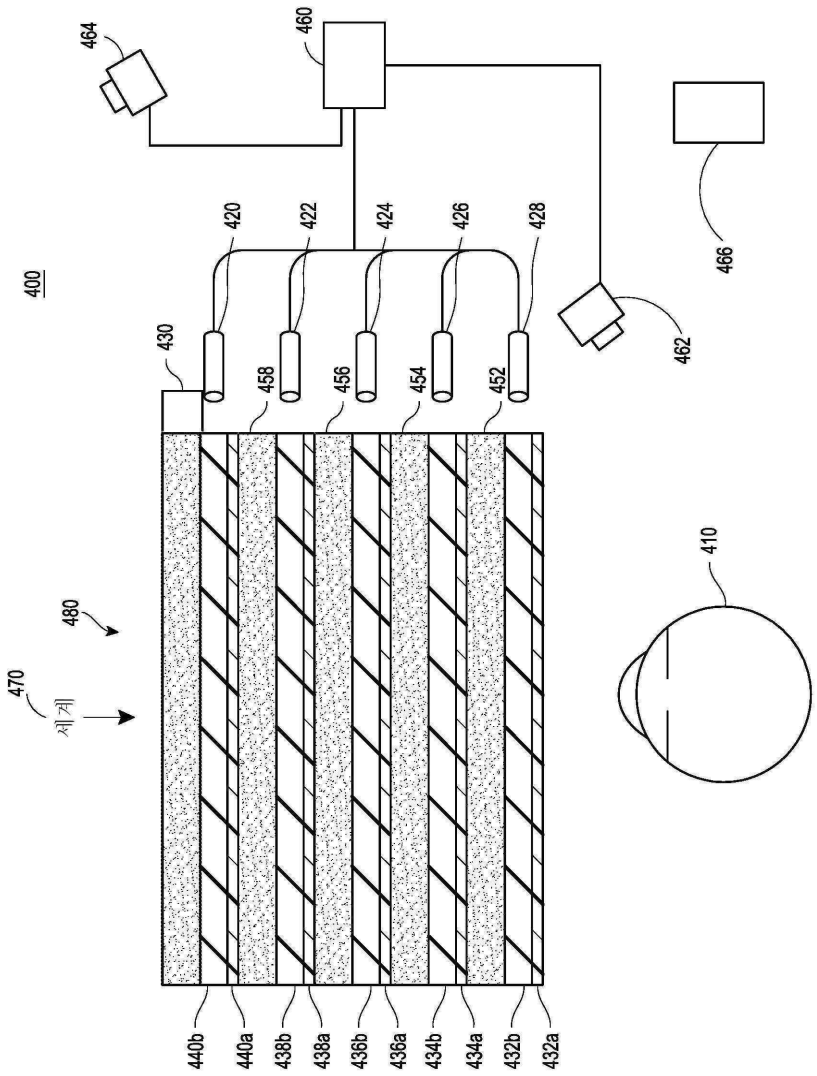
도면2b


94
36
32
105
104
18
39
98
92
16
34
24
38
37
102
26
304
108
102
90
302
106
26
38
24
16
18
202
200
230
220

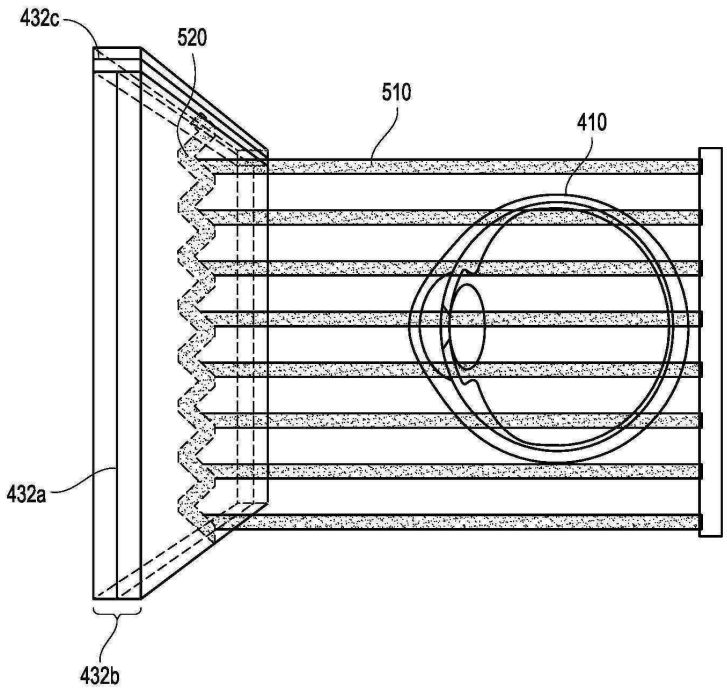
도면3



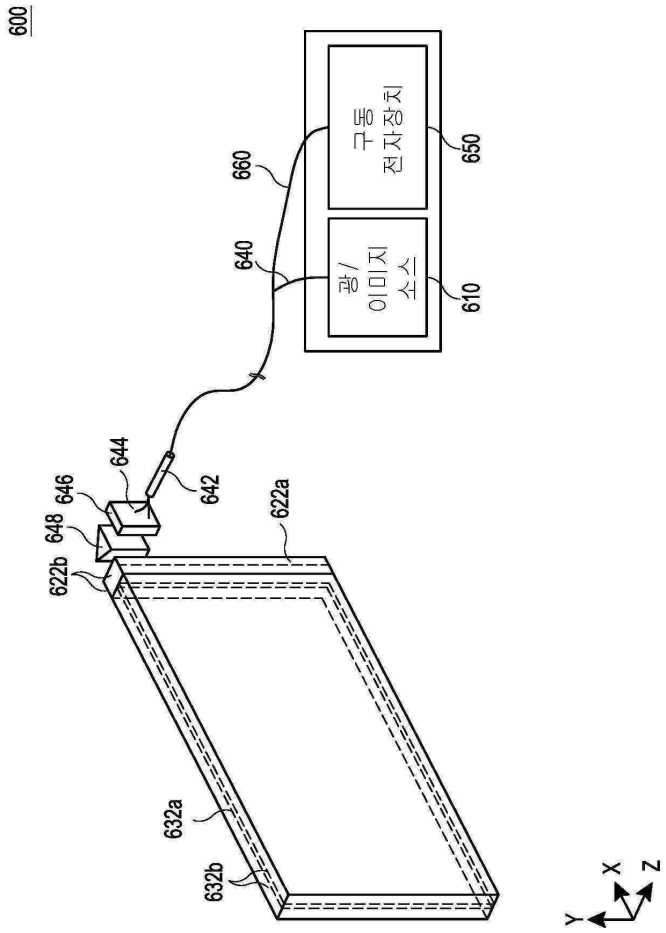
도면4


464
460
466
420
422
424
426
428
400
462
430
458
456
454
452
410
480
470
세계
440b
440a
438b
438a
436b
436a
434b
434a
432b
432a

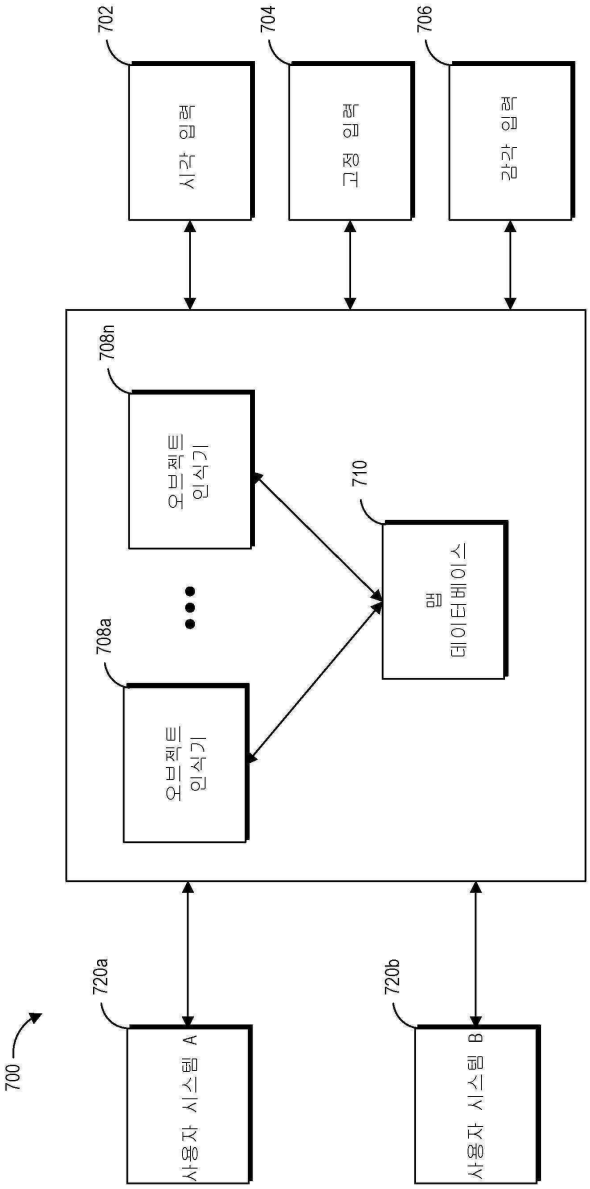
도면5



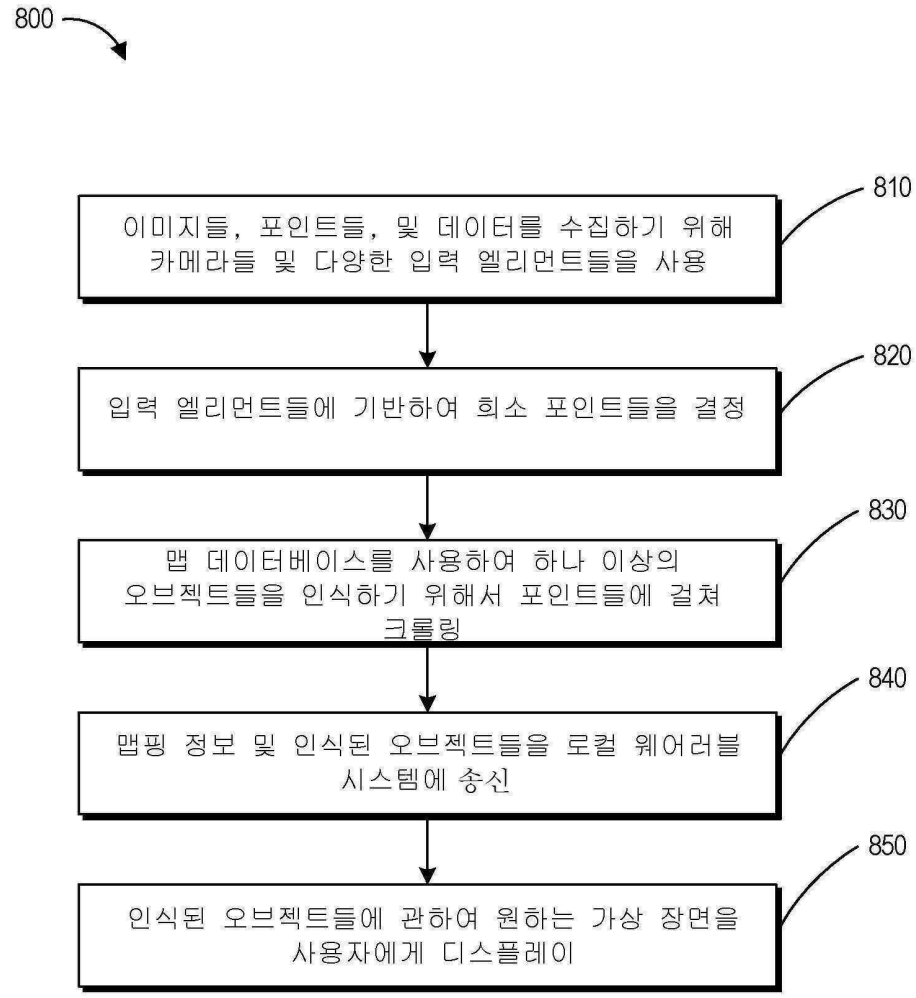
도면6


600
구동 전자장치
650
660
640
광/ 이미지 소스
610
644
646
642
648
622a
622b
632a
632b
X
Z
Y

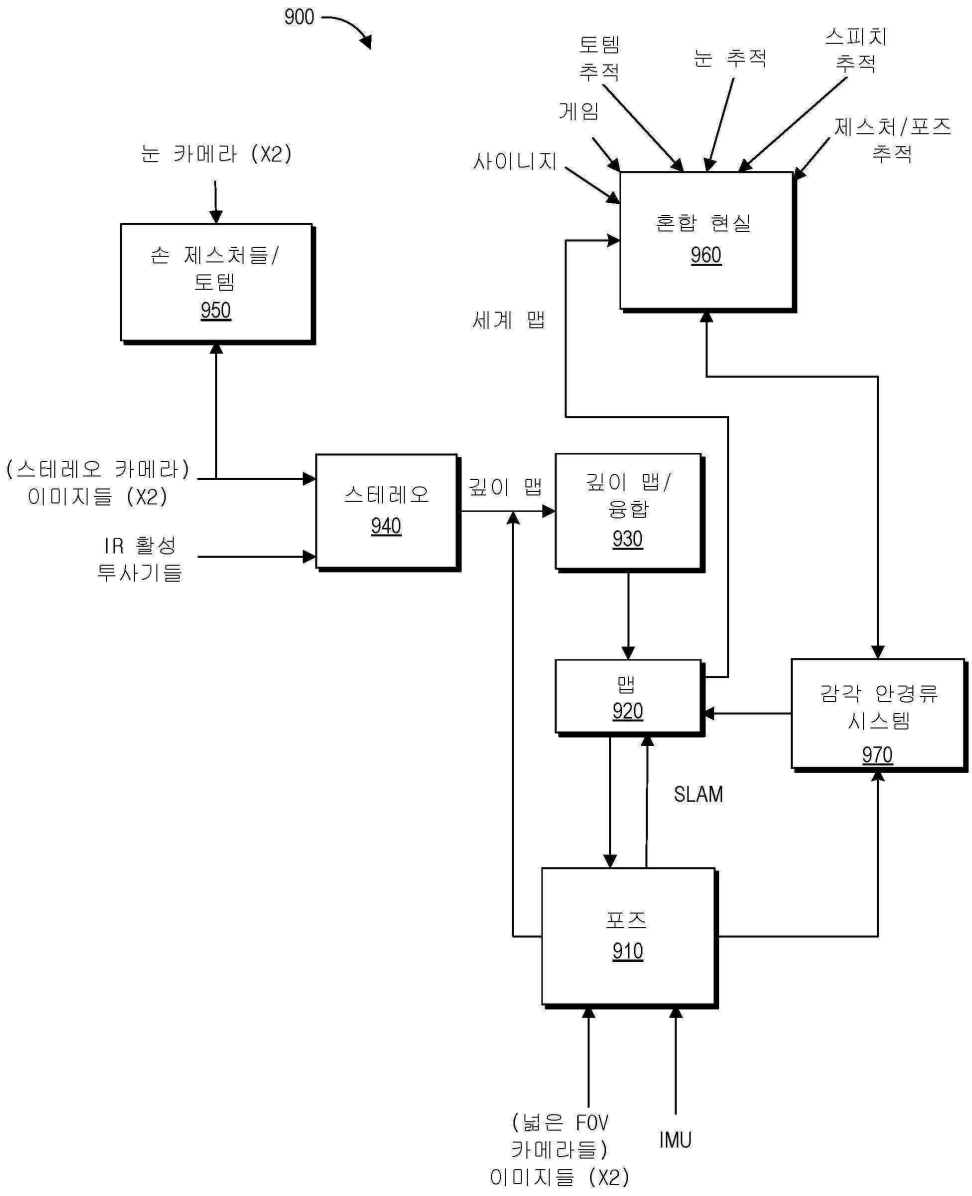
도면7


700
702
시각 입력
704
고정 입력
706
감각 입력
708n
오브젝트 인식기
708a
오브젝트 인식기
710
맵 데이터베이스
720a
사용자 시스템 A
720b
사용자 시스템 B

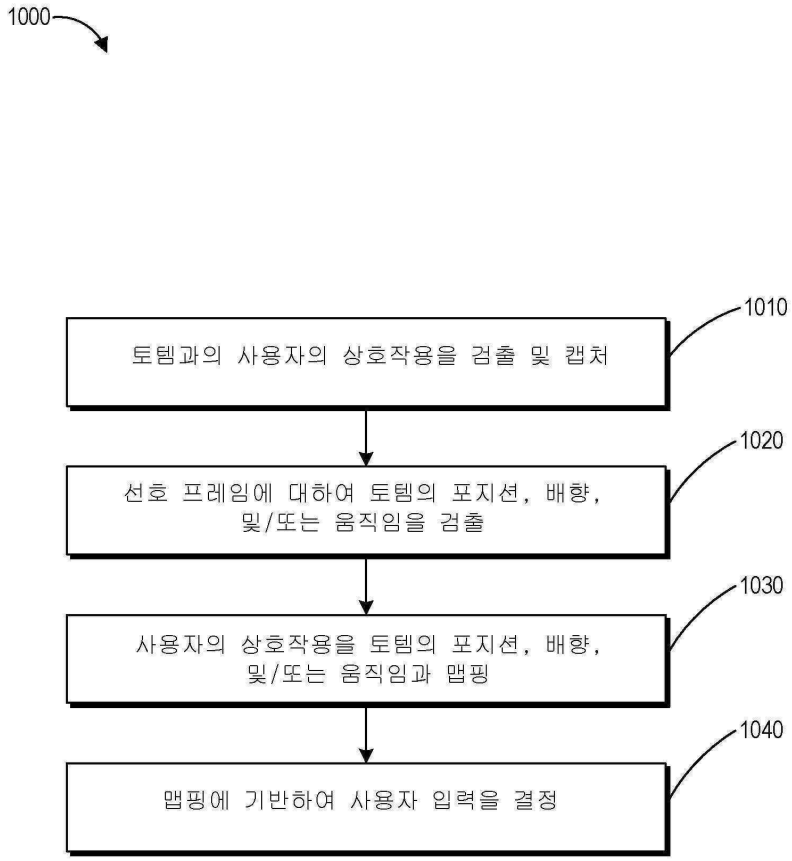
도면8


800
이미지들, 포인트들, 및 데이터를 수집하기 위해 카메라들 및 다양한 입력 엘리먼트들을 사용
810
입력 엘리먼트들에 기반하여 희소 포인트들을 결정
820
맵 데이터베이스를 사용하여 하나 이상의 오브젝트들을 인식하기 위해서 포인트들에 걸쳐 크롤링
830
맵핑 정보 및 인식된 오브젝트들을 로컬 웨어러블 시스템에 송신
840
인식된 오브젝트들에 관하여 원하는 가상 장면을 사용자에게 디스플레이
850

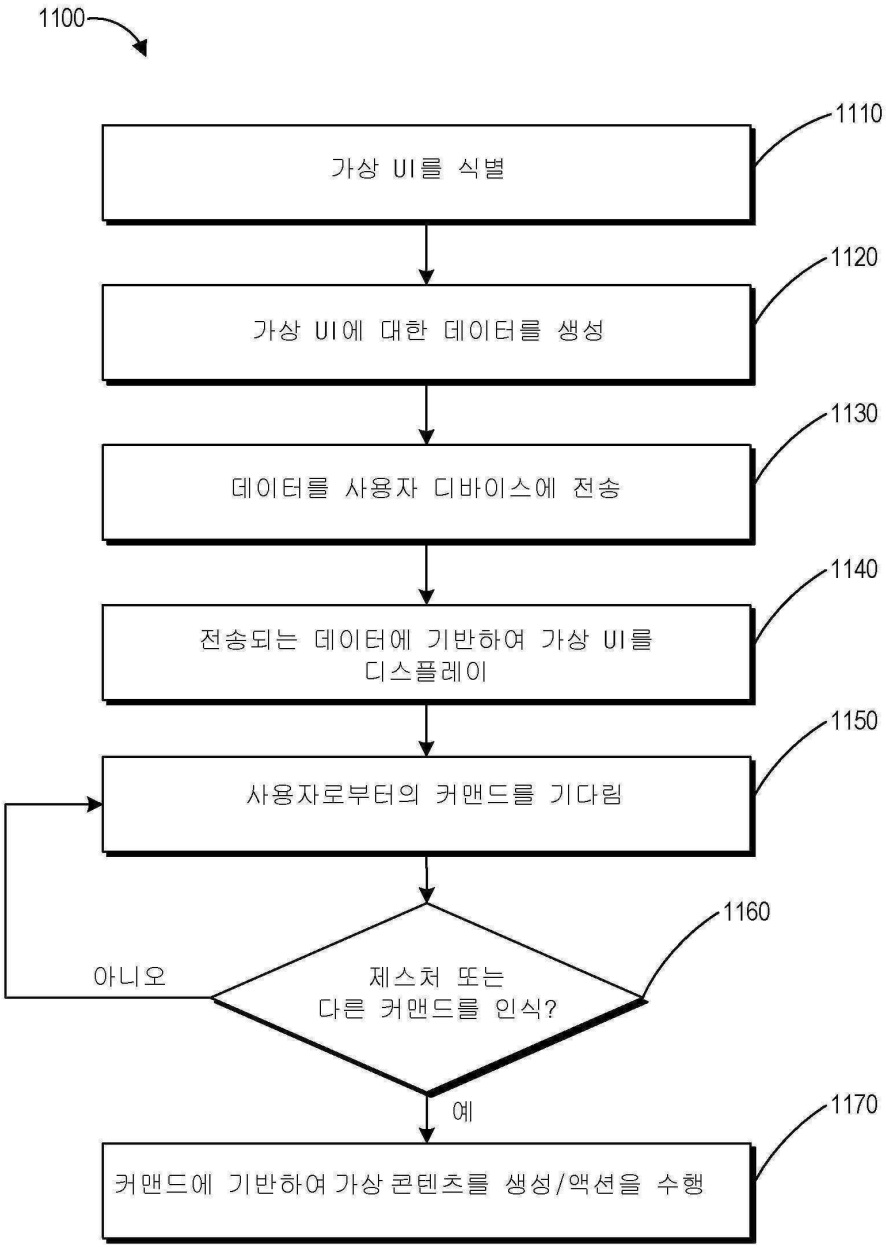
도면9


900
토템 추적
눈 추적
스피치 추적
게임
제스처/포즈 추적
사이니지
눈 카메라 (X2)
혼합 현실 960
손 제스처들/ 토템 950
세계 맵
(스테레오 카메라) 이미지들 (X2)
스테레오 940
깊이 맵
깊이 맵/ 융합 930
IR 활성 투사기들
맵 920
감각 안경류 시스템 970
SLAM
포즈 910
(넓은 FOV 카메라들) 이미지들 (X2)
IMU

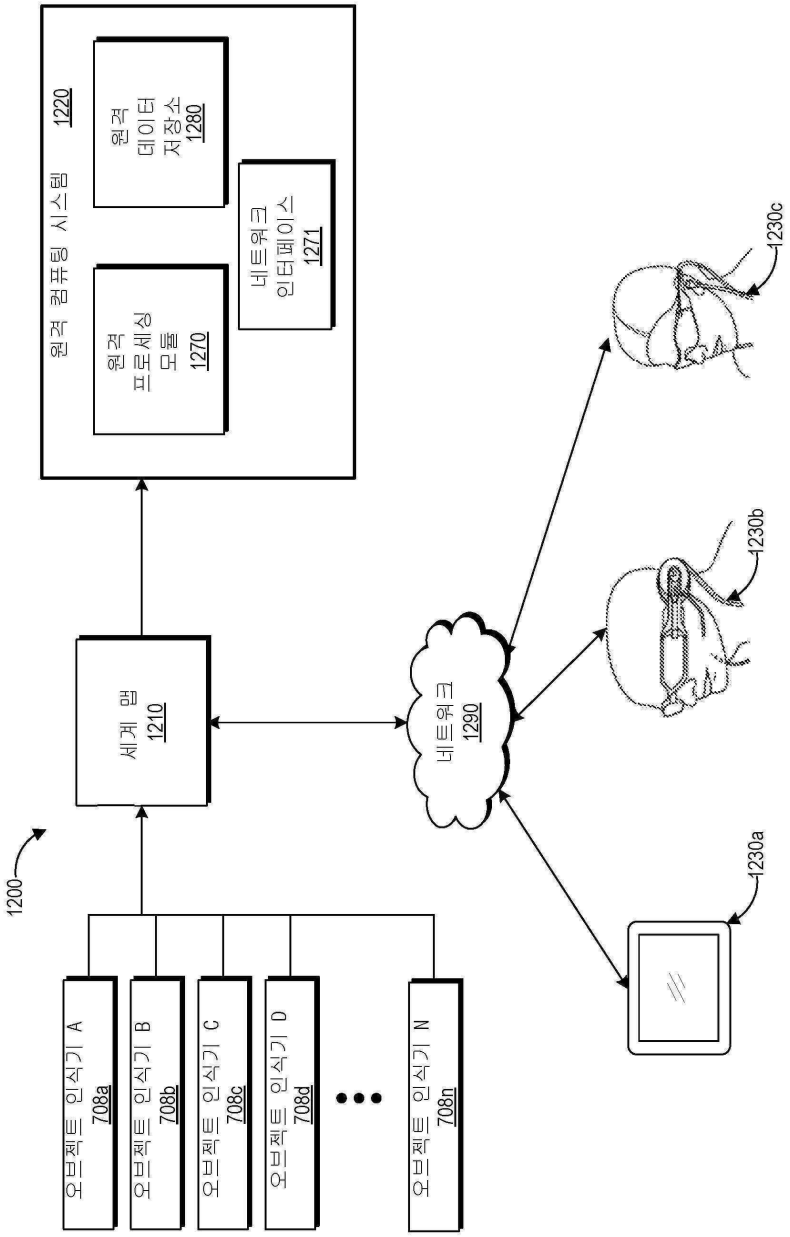
도면10


1000
토템과의 사용자의 상호작용을 검출 및 캡처
1010
선호 프레임에 대하여 토템의 포지션, 배향, 및/또는 움직임을 검출
1020
사용자의 상호작용을 토템의 포지션, 배향, 및/또는 움직임과 맵핑
1030
맵핑에 기반하여 사용자 입력을 결정
1040

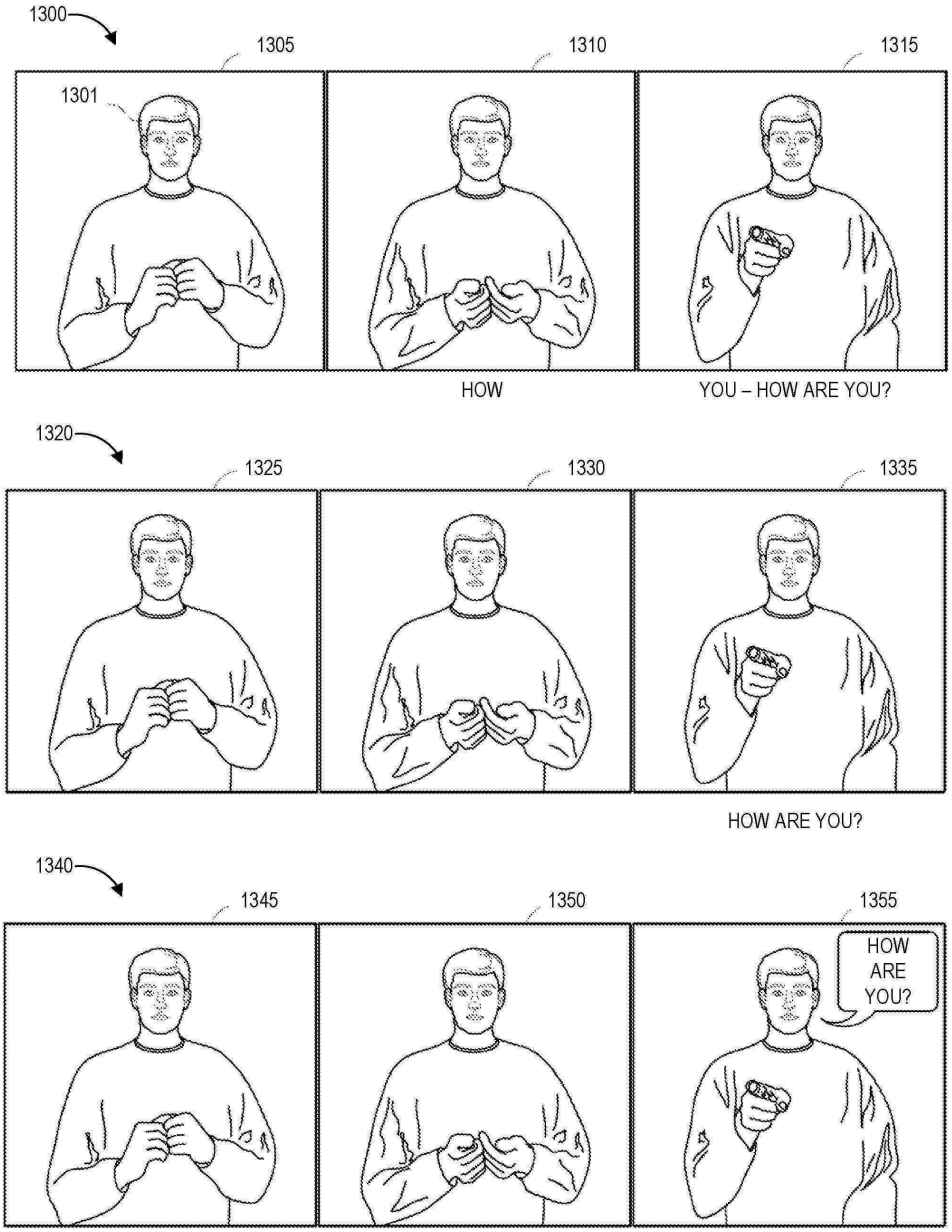
도면11


1100
1110
가상 UI를 식별
1120
가상 UI에 대한 데이터를 생성
1130
데이터를 사용자 디바이스에 전송
1140
전송되는 데이터에 기반하여 가상 UI를 디스플레이
1150
사용자로부터의 커맨드를 기다림
1160
제스처 또는 다른 커맨드를 인식?
아니오
예
1170
커맨드에 기반하여 가상 콘텐츠를 생성/액션을 수행

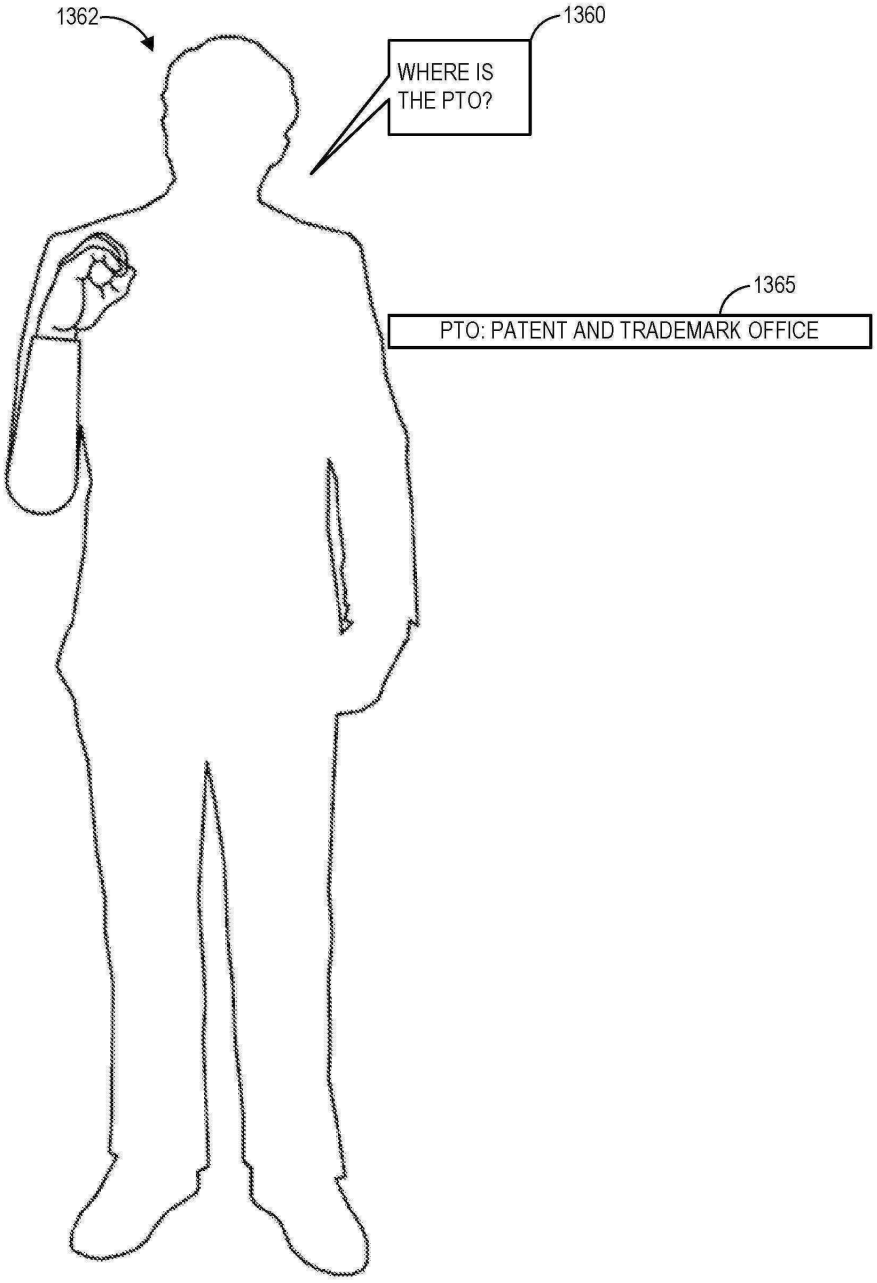
도면12


원격 컴퓨팅 시스템 1220
원격 데이터 저장소 1280
네트워크 인터페이스 1271
원격 프로세싱 모듈 1270
세계 맵 1210
네트워크 1290
1200
1230a
1230b
1230c
오브젝트 인식기 A 708a
오브젝트 인식기 B 708b
오브젝트 인식기 C 708c
오브젝트 인식기 D 708d
오브젝트 인식기 N 708n

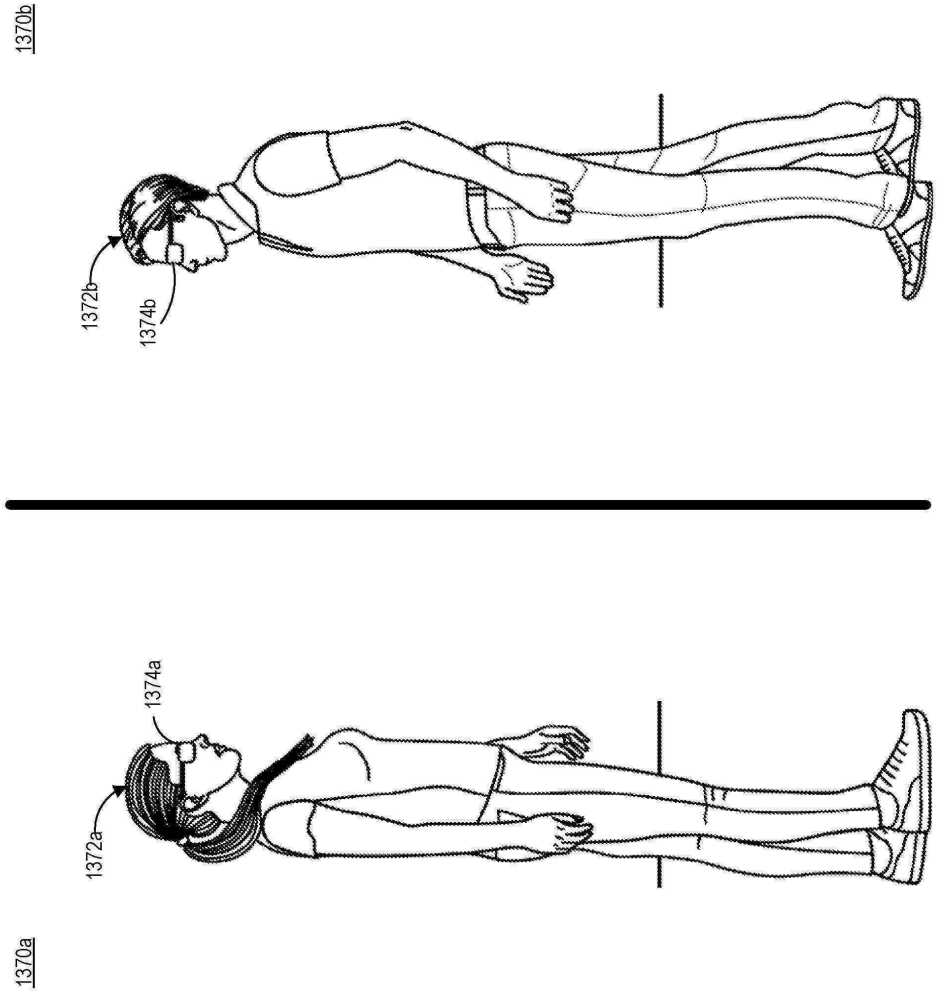
도면13a


1300
1305
1310
1315
1301
HOW
YOU – HOW ARE YOU?
1320
1325
1330
1335
HOW ARE YOU?
1340
1345
1350
1355
HOW ARE YOU?

도면13b


1362
1360
WHERE IS THE PTO?
1365
PTO: PATENT AND TRADEMARK OFFICE

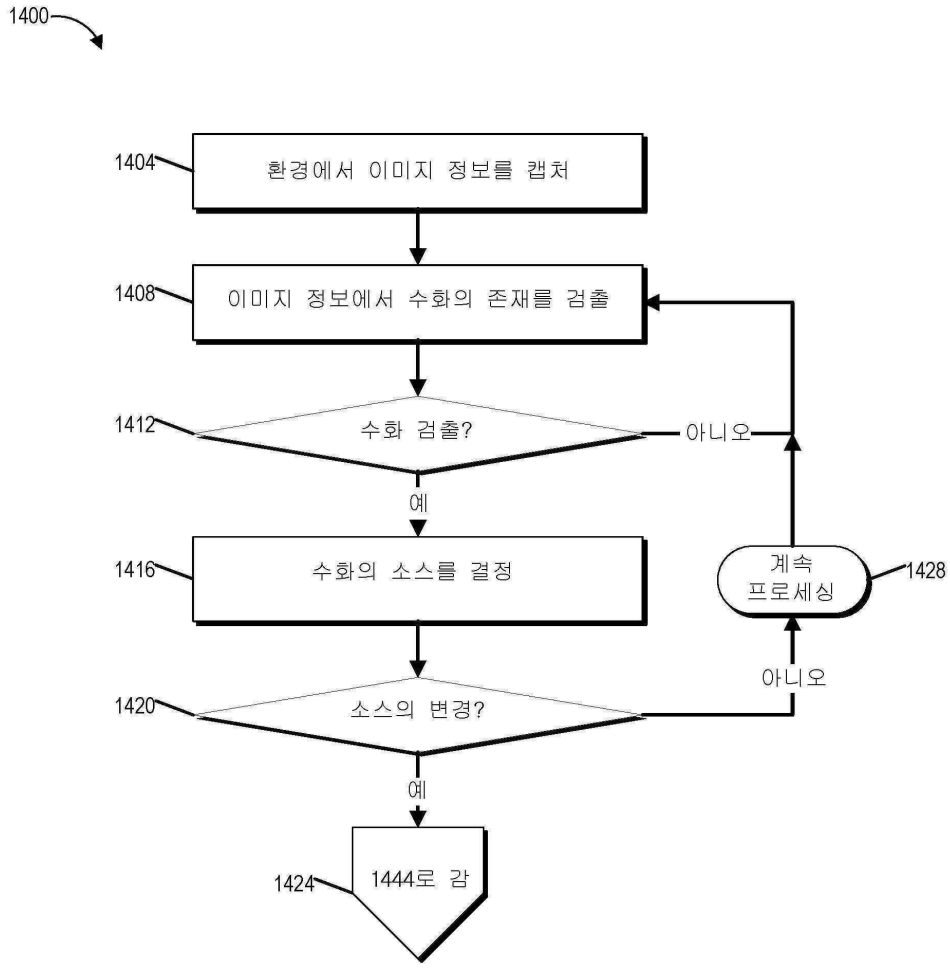
도면13c



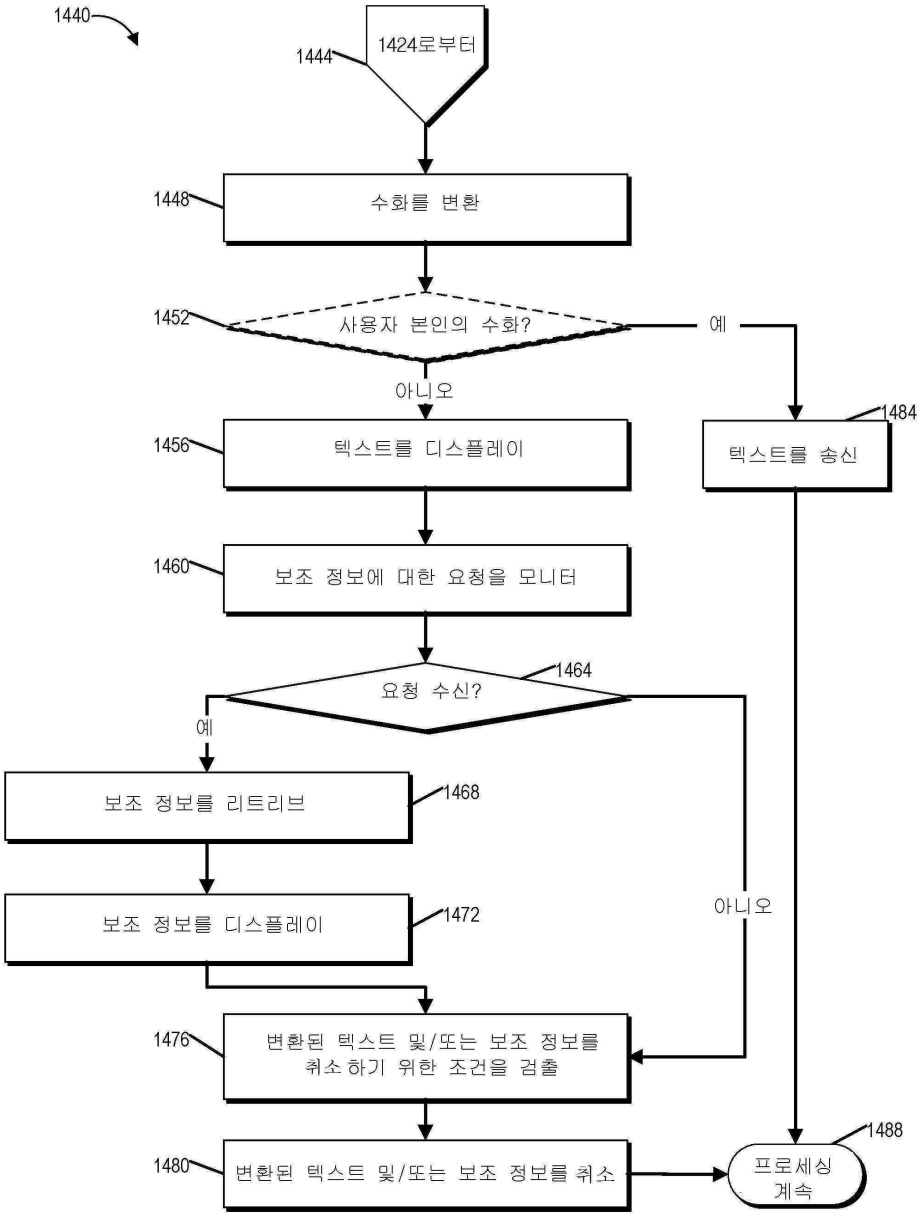
도면13d


1394
1384b
1384d
1382
1384a
1384c
1380
1392
Is there a pharmacy nearby?
I'm feeling unwell.
More Options

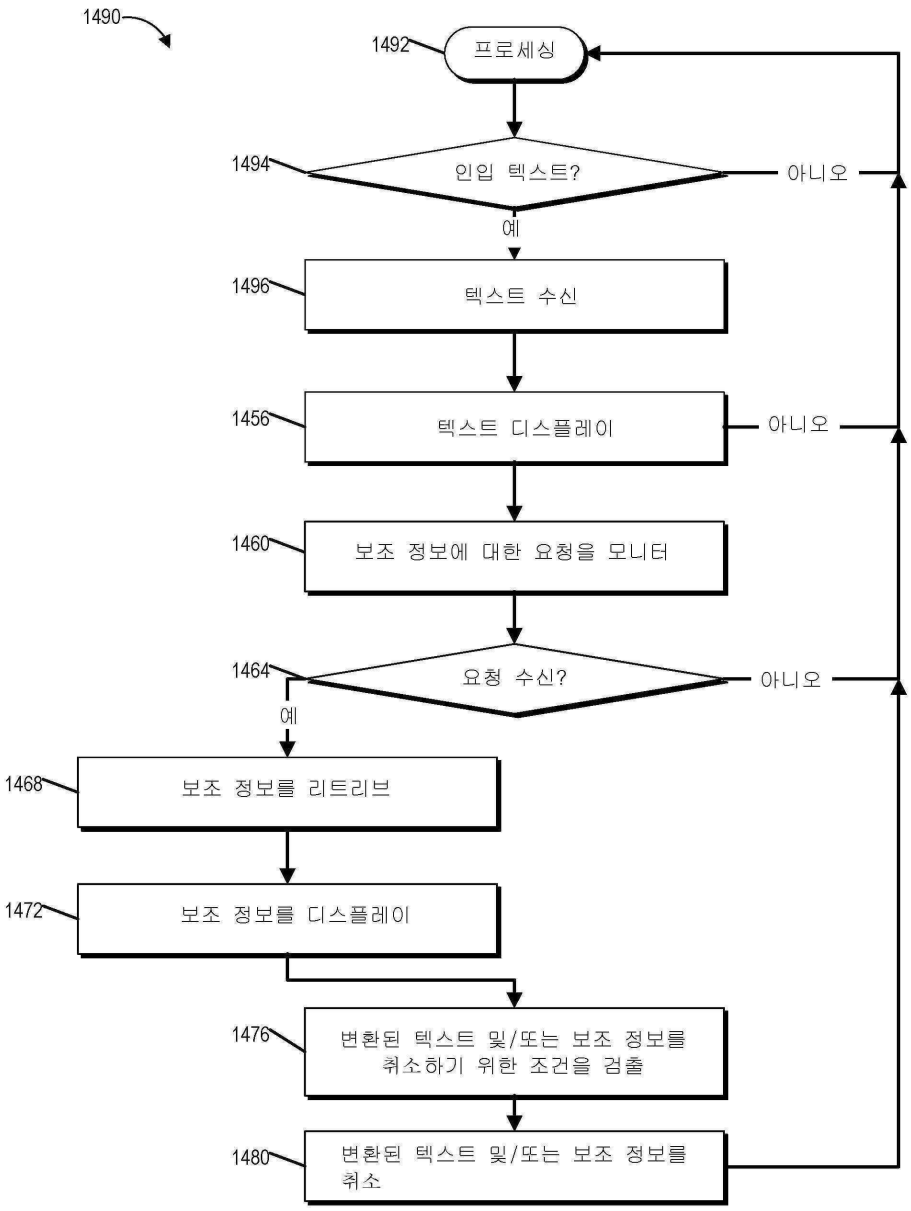
도면14a


1400
1404
환경에서 이미지 정보를 캡처
1408
이미지 정보에서 수화의 존재를 검출
1412
수화 검출?
아니오
예
1416
수화의 소스를 결정
계속 프로세싱
1428
아니오
1420
소스의 변경?
예
1424
1444로 감

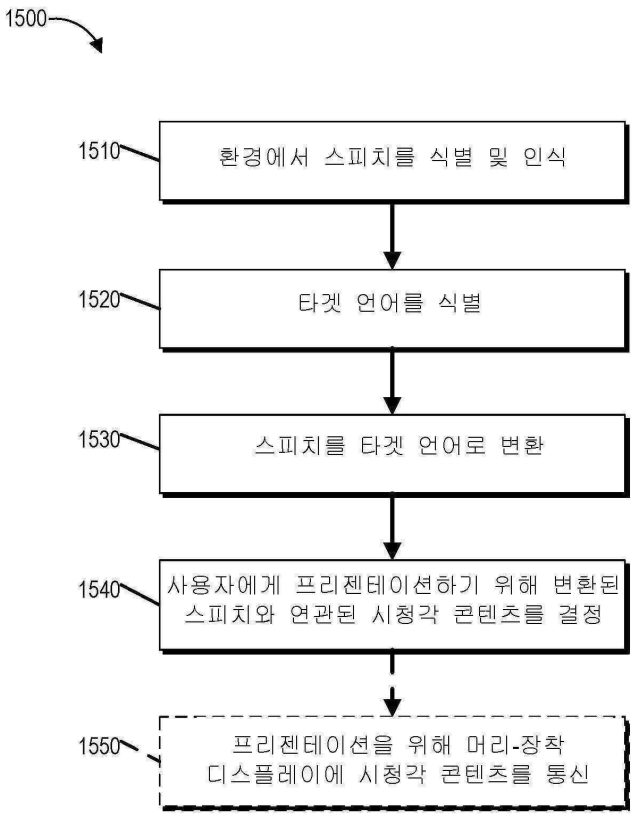
도면14b


1440
1444
1424로부터
1448
수화를 변환
1452
사용자 본인의 수화?
예
아니오
1456
텍스트를 디스플레이
1484
텍스트를 송신
1460
보조 정보에 대한 요청을 모니터
1464
요청 수신?
예
아니오
보조 정보를 리트리브
1468
보조 정보를 디스플레이
1472
1476
변환된 텍스트 및/또는 보조 정보를 취소하기 위한 조건을 검출
1480
변환된 텍스트 및/또는 보조 정보를 취소
1488
프로세싱 계속

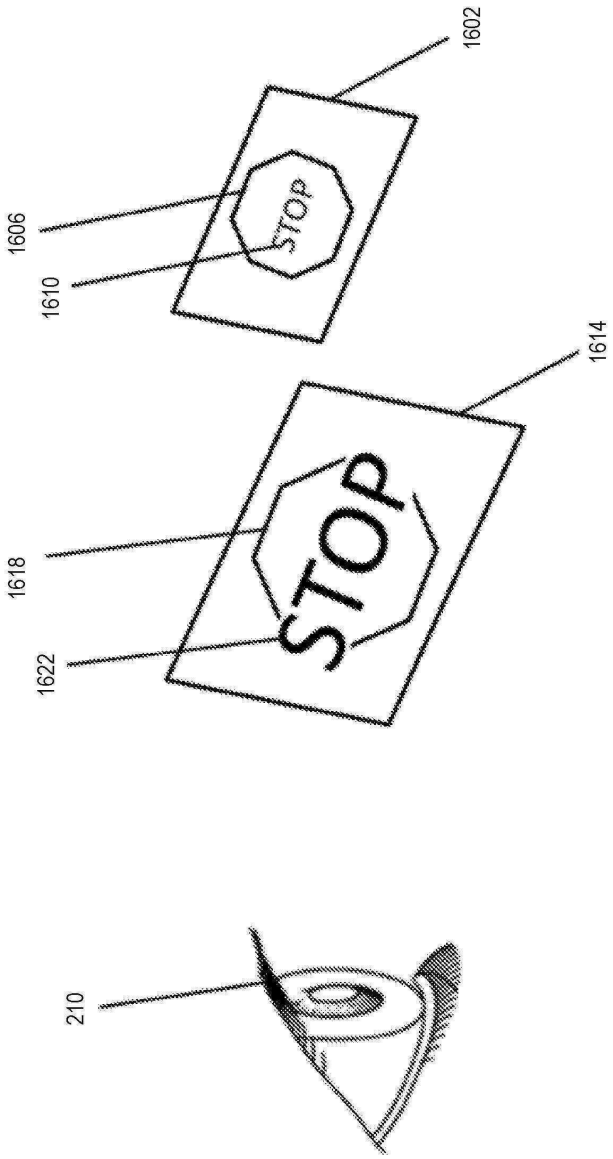
도면14c


1490
1492
프로세싱
1494
인입 텍스트?
아니오
예
1496
텍스트 수신
1456
텍스트 디스플레이
아니오
1460
보조 정보에 대한 요청을 모니터
1464
요청 수신?
아니오
예
1468
보조 정보를 리트리브
1472
보조 정보를 디스플레이
1476
변환된 텍스트 및/또는 보조 정보를 취소하기 위한 조건을 검출
1480
변환된 텍스트 및/또는 보조 정보를 취소

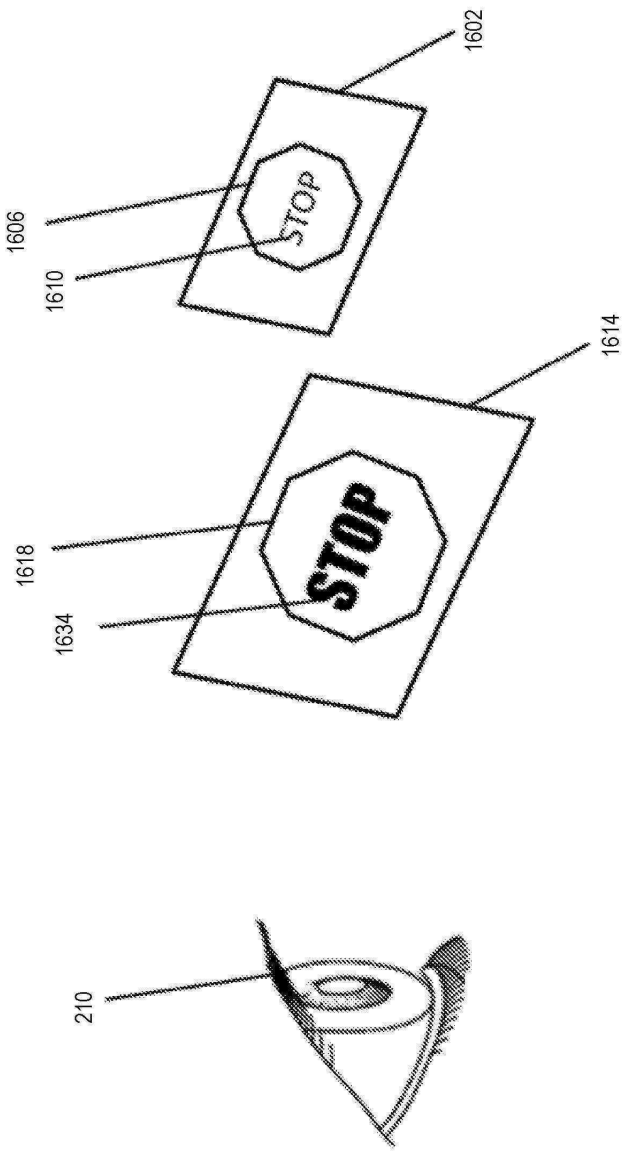
도면15


1500
1510
환경에서 스피치를 식별 및 인식
1520
타겟 언어를 식별
1530
스피치를 타겟 언어로 변환
1540
사용자에게 프리젠테이션하기 위해 변환된 스피치와 연관된 시청각 콘텐츠를 결정
1550
프리젠테이션을 위해 머리-장착 디스플레이에 시청각 콘텐츠를 통신

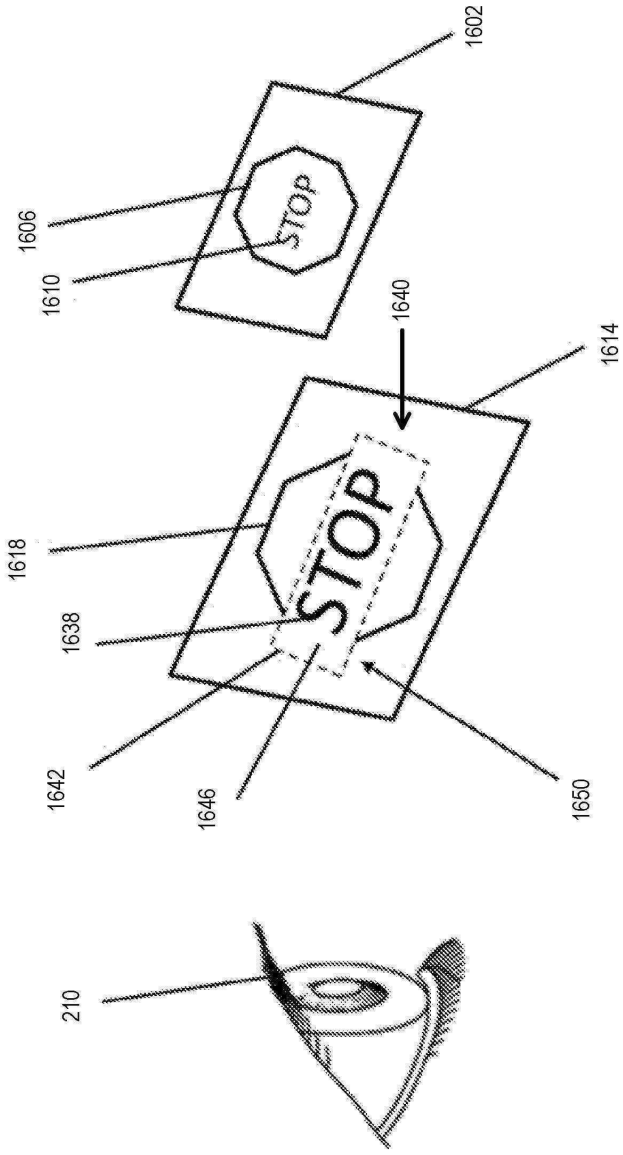
도면16a



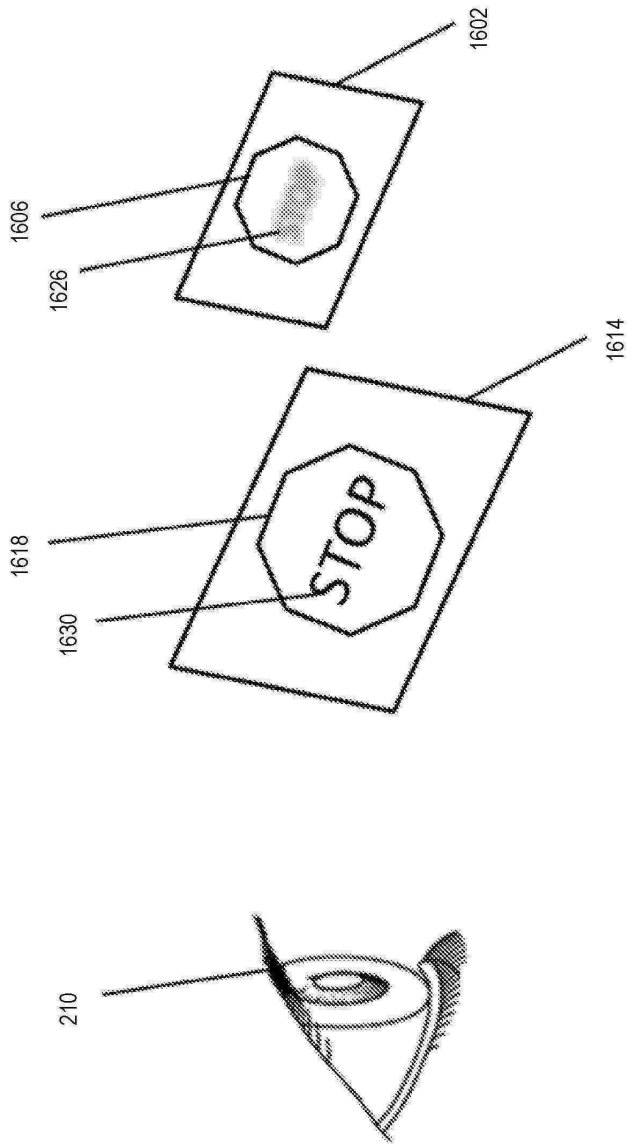
도면16b



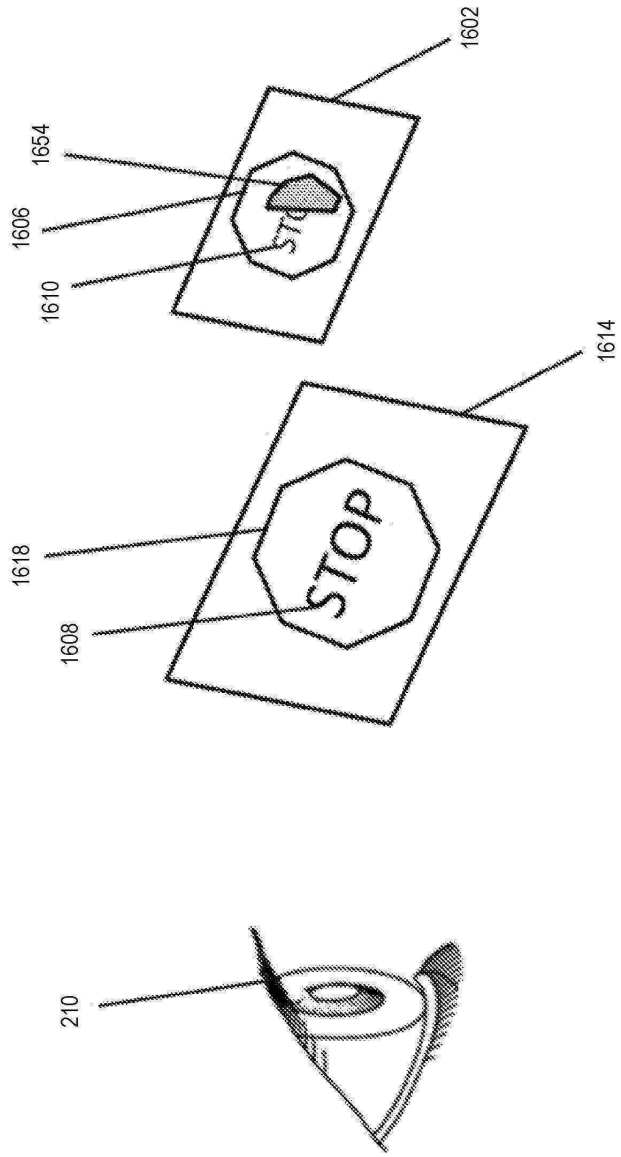
도면16c


1602
1606
1610
STOP
1640
1614
1618
1638
STOP
1642
1646
1650
210

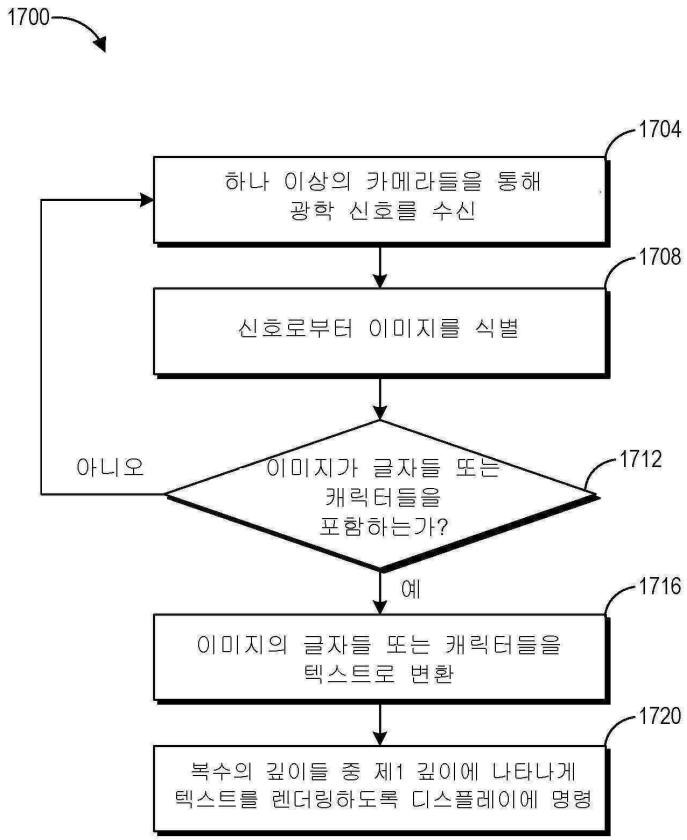
도면16d



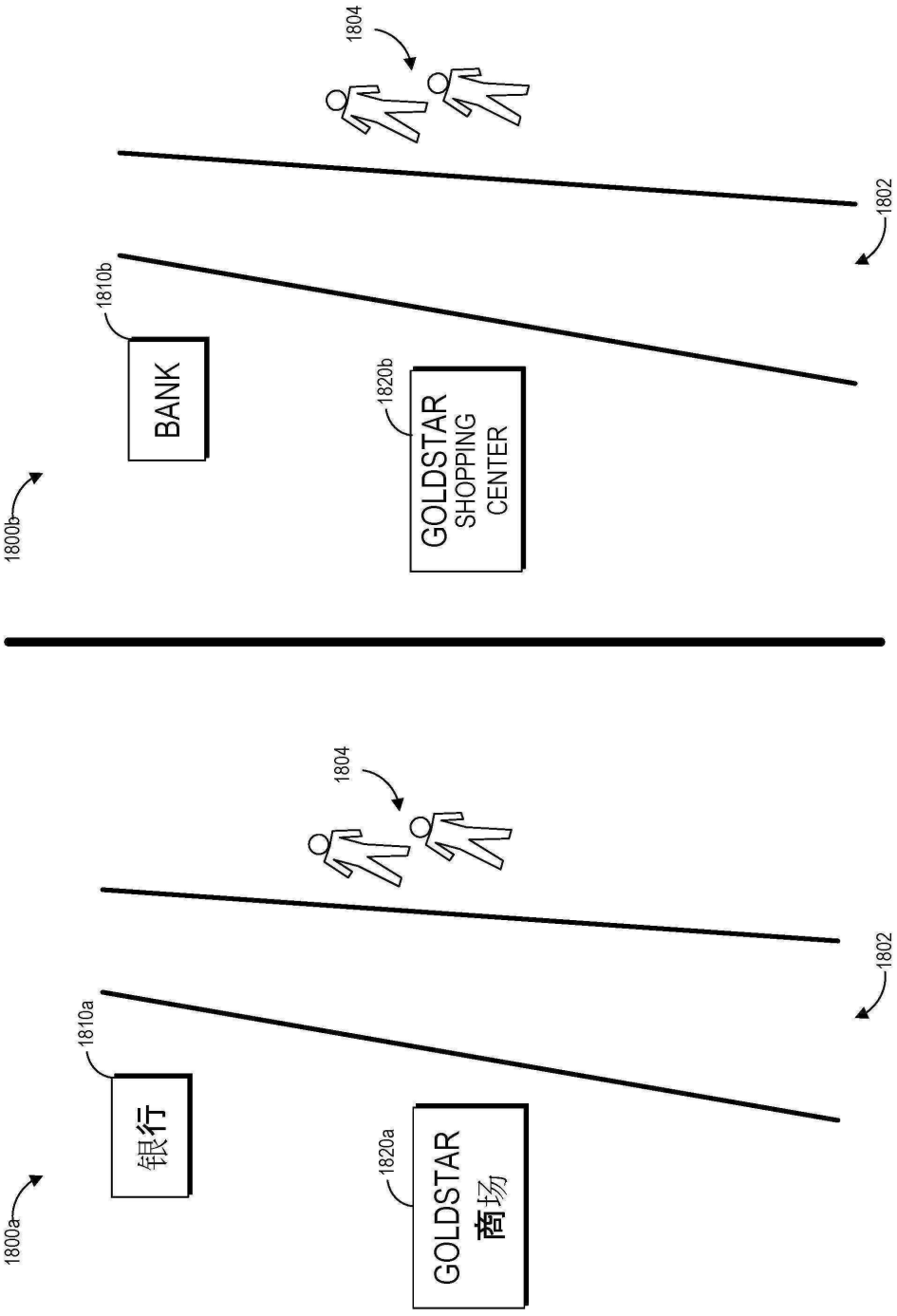
도면16e



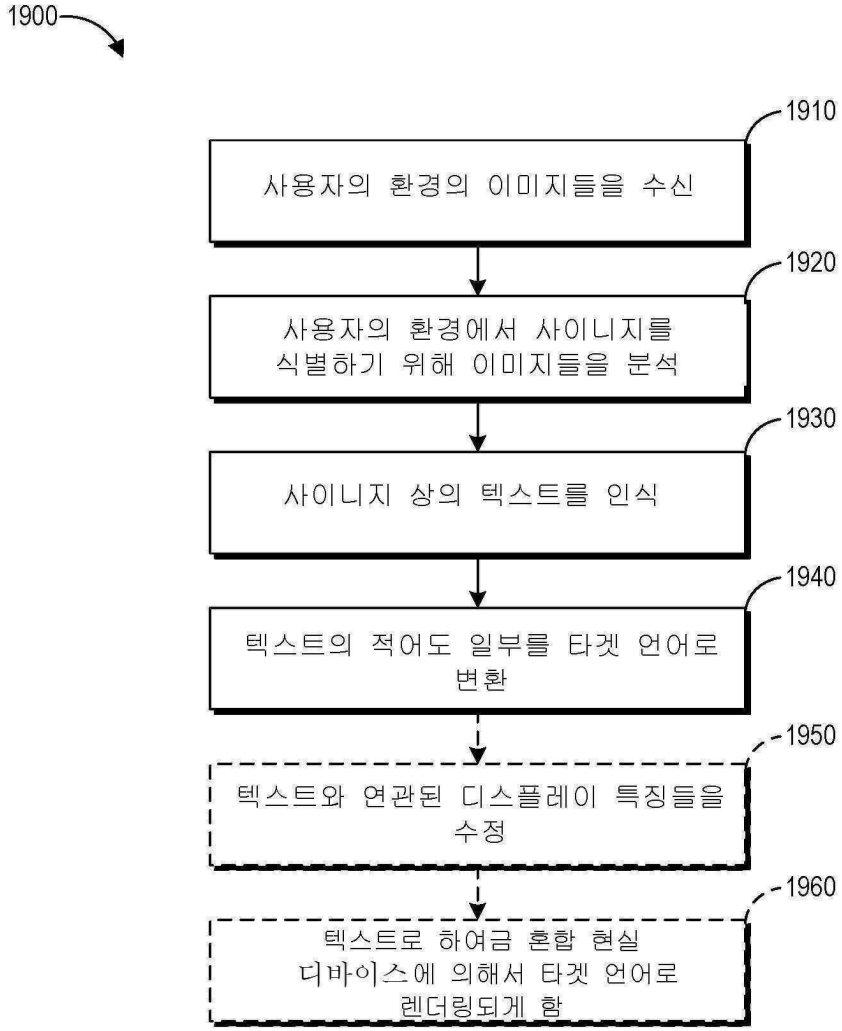
도면17


1700
1704
하나 이상의 카메라들을 통해 광학 신호를 수신
1708
신호로부터 이미지를 식별
1712
이미지가 글자들 또는 캐릭터들을 포함하는가?
아니오
예
1716
이미지의 글자들 또는 캐릭터들을 텍스트로 변환
1720
복수의 깊이들 중 제1 깊이에 나타나게 텍스트를 렌더링하도록 디스플레이에 명령

도면18


1804
1802
1810b
BANK
1820b
GOLDSTAR
SHOPPING
CENTER
1800b
1804
1802
1810a
银行
1820a
GOLDSTAR
商场
1800a

도면19


1900
1910
사용자의 환경의 이미지들을 수신
1920
사용자의 환경에서 사이니지를 식별하기 위해 이미지들을 분석
1930
사이니지 상의 텍스트를 인식
1940
텍스트의 적어도 일부를 타겟 언어로 변환
1950
텍스트와 연관된 디스플레이 특징들을 수정
1960
텍스트로 하여금 혼합 현실 디바이스에 의해서 타겟 언어로 렌더링되게 함