



등록특허 10-2227392



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2021년03월11일
(11) 등록번호 10-2227392
(24) 등록일자 2021년03월08일

- (51) 국제특허분류(Int. Cl.)
G10L 15/26 (2006.01) *G06F 40/166* (2020.01)
G06F 40/20 (2020.01) *G06F 40/40* (2020.01)
G10L 15/08 (2006.01) *G10L 15/18* (2006.01)
G10L 25/84 (2013.01) *H04N 7/14* (2006.01)
H04N 7/15 (2016.01)
- (52) CPC특허분류
G10L 15/26 (2013.01)
G06F 40/169 (2020.01)
- (21) 출원번호 10-2019-7007141
- (22) 출원일자(국제) 2017년08월10일
 심사청구일자 2020년08월10일
- (85) 번역문제출일자 2019년03월11일
- (65) 공개번호 10-2019-0038900
- (43) 공개일자 2019년04월09일
- (86) 국제출원번호 PCT/US2017/046251
- (87) 국제공개번호 WO 2018/031745
 국제공개일자 2018년02월15일
- (30) 우선권주장
 62/374,183 2016년08월12일 미국(US)
- (56) 선행기술조사문현
 KR1020160030996 A
 US20140081634 A1
 US20160093298 A1
 US20160321841 A1

- (73) 특허권자
 매직 립, 인코포레이티드
 미국 플로리다 플랜타티온 웨스트 선라이즈 블러바드 7500 (우: 33322)
- (72) 발명자
 소머즈, 제프리
 미국 33322 플로리다 플랜타티온 웨스트 선라이즈 블러바드 7500
 디바인, 제니퍼, 웹.알.
 미국 33322 플로리다 플랜타티온 웨스트 선라이즈 블러바드 7500
 (뒷면에 계속)
- (74) 대리인
 특허법인 남엔남

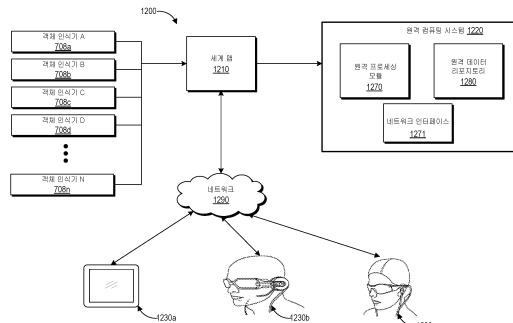
전체 청구항 수 : 총 20 항

심사관 : 노지명

(54) 발명의 명칭 단어 흐름 주석

(57) 요약

AR(augmented reality) 디바이스는 주변 오디오 데이터를 모니터링하도록 구성될 수 있다. AR 디바이스는 주변 오디오 데이터에서 스피치를 검출하고, 검출된 스피치를 텍스트로 변환하거나, 스피치에서 회귀 단어들과 같은 키워드들을 검출할 수 있다. 회귀 단어가 검출될 때, AR 디바이스는 공개 또는 개인 소스로부터 회귀 단어와 판 (뒷면에 계속)

대 표 도 - 도12

련된 보조 정보(예컨대, 정의)를 리트리브할 수 있다. AR 디바이스는 사용자가 스피치를 더 잘 이해하는 것을 돋기 위해 사용자에게 보조 정보를 디스플레이할 수 있다. AR 디바이스는 외국어 스피치의 번역을 수행할 수 있고, 사용자에게 스피커의 스피치의 텍스트(또는 번역)를 디스플레이할 수 있거나, 또는 스피치와 연관된 통계 또는 다른 정보를 디스플레이할 수 있다.

(52) CPC특허분류

G06F 40/242 (2020.01)

G06F 40/58 (2020.01)

G10L 15/1815 (2013.01)

G10L 25/84 (2013.01)

H04N 7/142 (2013.01)

H04N 7/157 (2013.01)

G10L 2015/088 (2013.01)

(72) 발명자

석, 조셉, 웨인

미국 33322 플로리다 플랜타티온 웨스트 선라이즈

블러바드 7500

케홀리, 아드리안

미국 90027 캘리포니아 로스 앤젤레스 노스 웨스턴
애비뉴 1940

명세서

청구범위

청구항 1

AR(augmented reality) 시스템으로서,

상기 AR 시스템의 사용자에게 가상 콘텐츠를 제시하도록 구성된 AR 디스플레이;

주변 사운드들을 캡처하도록 구성된 오디오 센서; 및

상기 AR 디스플레이 및 상기 오디오 센서와 통신하는 하드웨어 프로세서를 포함하고,

상기 하드웨어 프로세서는,

상기 오디오 센서에 의해 캡처된 주변 사운드들을 수신하고;

상기 주변 사운드들에서 스피치(speech)의 존재를 검출하고;

상기 검출된 스피치를 텍스트로 변환하고;

하나 이상의 단어들을 인식하기 위해 상기 텍스트를 파싱(parsing)하고;

단어 빈도 리스트 상에서 상기 하나 이상의 단어들의 위치를 결정하고 – 상기 단어 빈도 리스트는 상기 오디오 센서에 의한 단어들의 검출 빈도에 적어도 부분적으로 기초하여 랭크되는 단어들의 리스트를 포함함 –;

사용자 입력에 기초하여 상기 단어 빈도 리스트 상의 임계 위치를 결정하고;

상기 하나 이상의 단어들의 위치가 상기 단어 빈도 리스트 상의 임계 위치를 넘는지(pass) 여부에 기초하여 상기 하나 이상의 단어들을 희귀 단어로서 분류하고;

상기 희귀 단어와 연관된 쇠퇴 팩터(decay factor)를 결정하고 – 상기 쇠퇴 팩터는 상기 희귀 단어와 연관된 보조 정보의 디스플레이를 금지하는 시간 기간을 포함함 –;

상기 쇠퇴 팩터에 기초하여 상기 희귀 단어와 연관된 상기 보조 정보를 리트리브(retrieve)하고; 그리고

상기 AR 디스플레이가 상기 리트리브된 보조 정보를 렌더링하게 하도록 프로그래밍되는,

AR 시스템.

청구항 2

제 1 항에 있어서,

상기 희귀 단어는 공통 단어 사전에서 제외된 단어인,

AR 시스템.

청구항 3

제 1 항에 있어서,

상기 하드웨어 프로세서는 추가로,

상기 보조 정보의 디스플레이를 해제(dismissing)하기 위한 조건을 검출하고; 그리고

상기 검출된 조건에 대한 응답으로, 상기 AR 디스플레이가 상기 보조 정보의 디스플레이를 해제하게 하도록 프로그래밍되는,

AR 시스템.

청구항 4

제 3 항에 있어서,

상기 디스플레이를 해제하기 위한 조건을 검출하기 위해, 상기 하드웨어 프로세서는, 임계 시간 기간이 경과되었음을 결정하는 것, 다른 희귀 단어를 검출하는 것, 사용자의 포즈를 검출하는 것, 또는 사용자 입력 디바이스로부터 입력을 수신하는 것 중 적어도 하나를 수행하도록 프로그래밍되는,

AR 시스템.

청구항 5

제 1 항에 있어서,

상기 하드웨어 프로세서는 추가로,

상기 검출된 스피치의 소스를 결정하고; 그리고

상기 희귀 단어를 검출하고, 상기 검출된 스피치의 소스가 상기 AR 디스플레이의 사용자 이외의 다른 스피커와 연관된다고 결정 시에 상기 보조 정보를 리트리브 및 디스플레이하도록 프로그래밍되는,

AR 시스템.

청구항 6

제 5 항에 있어서,

상기 검출된 스피치의 소스는 하나 이상의 지향성 마이크로폰들을 포함하는 오디오 센서로부터 수집된 오디오 정보 또는 상기 AR 시스템의 외향 이미징 시스템에 의해 획득된 이미지를 중 적어도 하나에 기초하여 결정되는,

AR 시스템.

청구항 7

제 5 항에 있어서,

상기 검출된 스피치의 소스를 결정하기 위해, 상기 하드웨어 프로세서는 상기 스피커의 아이덴티티를 결정하기 위해 상기 검출된 스피치 상에서 음성 인식을 수행하도록 프로그래밍되는,

AR 시스템.

청구항 8

제 5 항에 있어서,

상기 검출된 스피치의 소스는, 상기 사용자의 환경의 다른 컴퓨팅, 상기 스피커와 연관된 다른 AR 디바이스 또는 상기 사용자의 환경의 사람 중 적어도 하나를 포함하는,

AR 시스템.

청구항 9

제 8 항에 있어서,

상기 사용자는 상기 스피커와의 텔레프레전스 세션(telepresence session)에 있고, 상기 검출된 스피치의 소스는 상기 스피커와 연관된 다른 사용자 디바이스이고,

상기 하드웨어 프로세서는 추가로, 상기 AR 디스플레이가 상기 스피커의 가상 아바타를 렌더링하게 하고, 다른 사용자의 가상 아바타 근처에 상기 보조 정보를 렌더링하게 하도록 프로그래밍되는,

AR 시스템.

청구항 10

제 5 항에 있어서,

상기 AR 디스플레이가 상기 리트리브된 보조 정보를 렌더링하게 하기 위해, 상기 하드웨어 프로세서는,

상기 소스의 위치를 결정하고; 그리고

상기 AR 디스플레이를 통해 상기 소스의 뷰(view)를 모호하게 하지 않으면서 상기 소스에 가까운 3D 공간 내의 포지션에 상기 보조 정보를 렌더링하도록 프로그래밍되는,

AR 시스템.

청구항 11

제 1 항에 있어서,

상기 보조 정보는 상기 희귀 단어의 설명을 포함하는,

AR 시스템.

청구항 12

제 1 항에 있어서,

상기 하드웨어 프로세서는 추가로,

상기 보조 정보에 주석을 달기 위해 사용자로부터 표시를 수신하고;

상기 보조 정보의 주석의 저장을 개시하고; 그리고

상기 희귀 단어의 다른 존재를 검출하는 것에 대한 응답으로, 상기 AR 디스플레이가 상기 주석 및 상기 보조 정보를 렌더링하게 하도록 프로그래밍되는,

AR 시스템.

청구항 13

제 1 항에 있어서,

상기 단어 빈도 리스트는 상기 사용자가 상기 사용자의 과거 경험에서 대면한 빈도에 의해 정렬되는 단어들의 리스트를 포함하는,

AR 시스템.

청구항 14

방법으로서,

하드웨어 프로세서 및 사용자의 환경에 가상 콘텐츠를 제시하도록 구성된 AR(augmented reality) 디스플레이를 포함하는 AR 디바이스의 제어 하에,

상기 AR 디바이스의 사용자의 환경을 모니터링하는 단계;

상기 사용자 또는 상기 환경 중 적어도 하나와 연관된 맥락 정보(contextual information)에 기초하여 상기 환경에서 관심 객체의 존재를 검출하는 단계;

단어 빈도 리스트 상에서 상기 관심 객체의 위치를 결정하는 단계 – 상기 단어 빈도 리스트는 상기 AR 디바이스에 의한 단어들의 검출 빈도에 적어도 부분적으로 기초하여 랭크되는 단어들의 리스트를 포함함 –;

사용자 입력에 기초하여 상기 단어 빈도 리스트 상의 임계 위치를 결정하는 단계;

상기 관심 객체의 위치가 상기 단어 빈도 리스트 상의 임계 위치를 넘는지 여부에 기초하여 상기 관심 객체를 희귀 단어로서 분류하는 단계;

상기 관심 객체와 연관된 쇠퇴 팩터를 결정하는 단계 – 상기 쇠퇴 팩터는 상기 관심 객체와 연관된 보조 정보의 디스플레이를 금지하는 시간 기간을 포함함 –;

상기 쇠퇴 팩터에 기초하여 상기 관심 객체에 대한 상기 보조 정보를 리트리브하는 단계;

상기 사용자의 시야에서 상기 보조 정보의 디스플레이 포지션을 결정하는 단계 – 상기 시야는 정해진 시간에 사용자가 지각할 수 있는 상기 환경의 부분을 포함함 – ; 및

상기 AR 디스플레이가 상기 디스플레이 포지션에 상기 보조 정보를 렌더링하게 하는 단계를 포함하는,
방법.

청구항 15

제 14 항에 있어서,

상기 관심 객체의 존재를 검출하는 단계는,

상기 환경의 주변 사운드들로부터 스피치를 검출하는 단계; 및

희귀 단어를 식별하기 위해 상기 스피치를 파싱(parsing)하는 단계를 포함하는,
방법.

청구항 16

제 14 항에 있어서,

상기 환경을 모니터링하는 단계는,

상기 환경의 주변 사운드들을 캡처 및 분석하는 단계; 또는

상기 환경의 이미지들을 획득하는 단계 중 적어도 하나를 포함하는,

방법.

청구항 17

제 14 항에 있어서,

상기 맥락 정보는 상기 사용자의 과거 경험과 연관되고, 상기 단어 빈도 리스트는 상기 사용자가 상기 사용자의
과거 경험에서 대면한 빈도에 의해 정렬되는 단어들의 리스트를 포함하는,

방법.

청구항 18

제 14 항에 있어서,

상기 보조 정보는 상기 관심 객체의 설명 텍스트를 포함하는,

방법.

청구항 19

제 14 항에 있어서,

상기 보조 정보에 주석을 달기 위해 상기 사용자로부터 표시를 수신하는 단계;

상기 보조 정보의 주석의 저장을 개시하는 단계; 및

상기 관심 객체의 다른 존재를 검출하는 것에 대한 응답으로, 상기 AR 디스플레이가 상기 주석 및 상기 보조 정보를 렌더링하게 하는 단계를 더 포함하는,

방법.

청구항 20

제 14 항에 있어서,

상기 보조 정보의 디스플레이 포지션을 결정하는 단계는,

상기 관심 객체의 위치를 결정하는 단계; 및

상기 AR 디스플레이를 통해, 상기 관심 객체의 뷰를 모호하게 하지 않으면서 상기 관심 객체에 가까운 상기 환경 내의 디스플레이 포지션을, 포지션으로서 식별하는 단계를 포함하는,

방법.

발명의 설명

기술 분야

[0001]

[0001] 본 출원은, 2016년 8월 12일에 출원되고 발명의 명칭이 "WORD FLOW ANNOTATION"인 미국 출원 번호 제 62/374,183호를 35 U.S.C. § 119(e) 하에서 우선권으로 주장하며, 그리하여, 이 출원의 개시내용은 그 전체가 인용에 의해 본원에 포함된다.

[0002]

[0002] 본 개시내용은 가상 현실 및 증강 현실 이미징 및 시각화 시스템들에 관한 것으로, 보다 구체적으로 시스템들의 사용자에 의한 스피치(speech)의 이해를 향상시키기 위해 스피치의 시각적 주석을 생성하는 것에 관한 것이다.

배경 기술

[0003]

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

발명의 내용

[0004]

[0004] 청취자에게의 최소 레벨의 산만함(distraction) 및 경미한 레벨의 노력으로 실시간(또는 거의 실시간)으로 관심 객체들(예컨대, 청취자에게 새로운 또는 미지의 단어들)에 관한 설명 또는 정보를 제공하기 위한 혼합 현실 시스템의 다양한 실시예들이 개시된다.

[0005]

[0005] 일부 실시예들에서, AR(augmented reality) 시스템은, AR 시스템의 사용자에게 가상 콘텐츠를 제시하도록 구성된 AR 디스플레이; 주변 사운드들을 캡처하도록 구성된 오디오 센서; 및 AR 디스플레이 및 오디오 센서와 통신하는 하드웨어 프로세서를 포함할 수 있다. 하드웨어 프로세서는, 오디오 센서에 의해 캡처된 주변 사운드들을 수신하고; 주변 사운드들에서 스피치(speech)의 존재를 검출하고; 검출된 스피치를 텍스트로 변환하고; 스피치에서 희귀 단어를 검출하고; 희귀 단어와 연관된 보조 정보를 리트리브하고; 그리고 AR 디스플레이가 리트리브된 보조 정보를 렌더링하게 하도록 프로그래밍될 수 있다.

[0006]

[0006] 일부 실시예들에서, 방법이 하드웨어 프로세서 및 사용자의 환경에 가상 콘텐츠를 제시하도록 구성된 AR(augmented reality) 디스플레이를 포함하는 AR 디바이스의 제어 하에서 수행될 수 있다. 방법은 AR 디바이스의 사용자의 환경을 모니터링하는 단계; 사용자 또는 환경 중 적어도 하나와 연관된 맥락 정보(contextual information)에 기초하여 환경에서 관심 객체의 존재를 검출하는 단계; 관심 객체에 대한 보조 정보를 리트리브하는 단계; 사용자의 시야에서 보조 정보의 디스플레이 포지션을 결정하는 단계 – 시야는 정해진 시간에 사용자가 지각할 수 있는 환경의 부분을 포함함 – ; 및 AR 디스플레이가 디스플레이 포지션에 보조 정보를 렌더링하게 하는 단계를 포함한다.

[0007]

[0007] 본 명세서에서 설명되는 청구 대상의 하나 이상의 구현들의 세부사항들은, 아래의 첨부 도면들 및 설명

에서 기술된다. 다른 특징들, 양상들, 및 이점들은 설명, 도면들, 및 청구항들로부터 명백해질 것이다. 이 개요 또는 다음의 상세한 설명 어느 것도, 본 발명의 청구 대상의 범위를 한정하거나 제한하는 것으로 의도되지 않는다.

도면의 간단한 설명

[0008]

[0008] 도 1은 사람에 의해 뷰잉되는 소정의 물리적 객체들 및 소정의 가상 현실 객체들을 갖는 혼합 현실 시나리오의 예시를 도시한다.

[0009] 도 2는 단어 흐름 주석 시스템의 실시예를 구현할 수 있는 웨어러블 시스템의 예를 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, 다중-초점 볼류메트릭 디스플레이, 이미지 또는 광 필드의 생성에 사용되는, 도파관 장치, 도파관 장치로 또는 도파관 장치로부터의 광을 광학적으로 커플링하는 광학 커플러 서브시스템, 및 제어 서브시스템을 포함하는 광학 시스템을 도시하는 개략도이다.

[0014] 도 7은 웨어러블 시스템의 일 예의 블록도이다.

[0015] 도 8은 인지된 객체들과 관련하여 가상 콘텐츠를 렌더링하는 방법의 예의 프로세스 흐름도이다.

[0016] 도 9는 단어 흐름 주석 시스템의 실시예를 포함하는 웨어러블 시스템의 다른 예의 블록도이다.

[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 사용자 디바이스들을 도시하는 전체 시스템도를 개략적으로 예시한다.

[0020] 도 13a 내지 도 13g는 관심 객체와 상호작용하는 예들 및 관심 객체의 보조 정보를 예시한다.

[0021] 도 14a 및 도 14b는 스피치에 주석을 달는 예시적인 프로세스를 도시한다.

[0022] 도 15는 관심 객체와 상호작용하는 예시적인 프로세스 및 관심 객체의 보조 정보를 예시한다.

[0023] 도 16a 및 16b는 키워드들에 기초한 스레드 식별의 예를 예시한다.

[0024] 도 17은 스레드 식별의 예시적인 프로세스를 예시한다.

[0025] 도면들 전체에 걸쳐, 참조 번호들은 참조된 엘리먼트들 사이의 대응성(correspondence)을 표시하는 데 재사용될 수 있다. 도면들은 본원에서 설명된 예시적인 실시예들을 예시하기 위해 제공되며 본 개시내용의 범위를 제한하려는 의도는 아니다.

발명을 실시하기 위한 구체적인 내용

[0009]

개요

[0010]

[0026] 청취자에게 새로운 이름들 및 레퍼런스(reference)들을 대면하는 것이 현대 세계의 일상 스피치에서 일반적인데, 그 이유는, 이들이 흔치 않은 아이디어들 또는 개념들, 특정 문제 도메인으로부터의 아이디어들 또는 개념들을 참조하기 때문이다거나, 또는 이들이 청취자가 모르는 사람들 또는 사물들에 대한 적절한 레퍼런스들이기 때문이다. 청취자는 단순히 친숙함 또는 지식의 결여로 인해 소정의 단어들을 알지 못할 수 있다. 자연히, 이는 대화를 방해하고, 당혹, 혼란, 또는 오해로 이어질 수 있다.

[0011]

[0027] 청취자에게 최소 레벨의 산만함(distraction) 및 청취자에 의한 경미한 레벨의 노력으로 실시간(또는 거의 실시간)으로 새로운 또는 미지의 단어들에 관한 설명 또는 정보를 제공할 수 있는 증강 현실 시스템을 갖는 것이 바람직할 수 있다.

- [0012] [0028] 본 개시내용은 ARD(augmented reality device)의 맥락에서 그러한 바람직한 시스템들의 예들을 개시한다. 이러한 디바이스는, 정보가 일반적으로 뷰잉 가능한 실제 세계와 나란히(또는 그 위에) 동시에 뷰잉 가능한 방식으로, 컴퓨팅 디바이스에 의해 제공되는 정보를 사용자가 시각적으로 수신할 수 있게 할 수 있다. 그러한 시스템은 통상적인 컴퓨터 스크린 상에 디스플레이될 수 있는 임의의 형태의 정보, 이를테면, 문자들, 이미지 효과들, 텍스트, 그래픽들 또는 임의의 종류의 비디오를 디스플레이하는 데 사용될 수 있다. ARD는, 눈 또는 눈을 전방에 착용된 시-쓰루(see-through) 디스플레이일 수 있는 AR 디스플레이 또는 VR 디스플레이를 포함할 수 있다. ARD는 또한 사용자 또는 사용자의 환경 내의 다른 사람에 의한 스피치를 포함한 오디오 정보를 캡처할 수 있는 오디오 센서를 포함할 수 있다.
- [0013] [0029] ARD는 단어 흐름 주석 시스템을 구현할 수 있다. 단어 흐름 주석 시스템은 스피치 인식, 희귀 단어 검출 및 ARD의 디스플레이 능력을 결합하여 단어 또는 사용자에게 적절한 보조 정보를 사용자에게 제공할 수 있다. 예로서, 제 1 사용자에 대한 ARD는 제 1 사용자가 제 2 사용자와 통신하는 텔레프레전스 세션(telepresence session) 동안 네트워크를 통해 오디오 정보를 수신할 수 있다. 단어 흐름 주석 시스템은 제 2 사용자로부터의 오디오 정보에서 스피치를 검출할 수 있고, 검출된 스피치를 텍스트로 변환하고, 변환된 텍스트에서 희귀 단어를 검색하고, 검출된 희귀 단어와 연관된 보조 정보를 리트리브할 수 있다. ARD는 ARD의 AR 디스플레이 상에 보조 정보를 디스플레이할 수 있다. 또한, 시스템은 임계 시간이 경과한 이후 또는 시스템이 다른 희귀 단어를 검출할 때, 보조 정보의 디스플레이를 해제(dismiss)할 수 있으며, 이는 유리하게는, 희귀 단어에 대한 유용한 보조 정보를 여전히 제공하면서, 제 1 사용자에 대한 산만함의 레벨을 감소시킬 수 있다. 희귀 사전 단어들에 대해, 단어의 의미(예컨대, 정의)가 디스플레이될 수 있다. 희귀 고유 명사들(예컨대, 지리적 구역들, 사람의 이름들 등)에 대해, 적절한 소스(예컨대, 백과사전)로부터의 관련 정보를 디스플레이될 수 있다. 시스템이 제공할 수 있는 주석들의 종류들은 예컨대, 인터넷 상에서 이용 가능한 정보 자원들의 거대한 어레이만큼 무제한적일 수 있다. 일부 구현들에서, 시스템은 변환된 텍스트(희귀 단어에 대한 정보 외에도 또는 이에 대안적으로)를 디스플레이의 착용자에게, 예컨대, 그 텍스트를 갖는 캡션 또는 말풍선(때로는, 콜-아웃 박스(call-out box)로서 또한 지칭됨)의 형태로 스피커 근처에 디스플레이한다. 일부 구현들에서, 보조 정보의 디스플레이는 시각적(예컨대, 이전에 언급된 캡션들 또는 말풍선들)이고; 그러나, 다른 구현들에서, 시각적 디스플레이에 부가적으로 또는 대안적으로, 보조 정보는 사용자에게 청각적으로 제시될 수 있다(이는 시각 장애가 있는 사용자들에게 유리할 수 있음).
- [0014] [0030] 희귀 단어들 외에도 또는 그에 대안으로서, 유사한 기술들이 다른 키워드들에 또한 적용될 수 있지만, 이를 키워드들 중 일부는 희귀하지 않을 수 있다. ARD는 사용자의 환경에서 관심 객체(예컨대, 사용자가 통상적으로 대면하지 못하는 객체)를 검출하고 관심 객체에 관한 보조 정보를 디스플레이할 수 있다. 본원에서 설명된 키워드들 및 희귀 단어들은 단일 단어, 구문 또는 문장을 포함할 수 있다.
- [0015] [0031] 추가로, 대화는 종종 다수의 토픽들을 포함한다. ARD는 대화에서 키워드들을 식별하고 키워드들에 기초하여 대화에서 토픽들을 식별할 수 있다. ARD는 대화와 연관된 텍스트들을 상이한 맥락 스레드들로 분리할 수 있으며, 각각의 스레드는 하나 이상의 토픽들에 대응한다. 유리하게는, 이는 ARD가 상이한 맥락 스레드들을 상이하게(예컨대, AR 디스플레이의 반대 측 상에) 제시하는 사용자 인터페이스(UI)를 제공할 수 있기 때문에, 대화와 관련된 아이디어들 및 포인트들의 사용자의 이해를 가능하게 할 수 있다.
- [0016] 웨어러블 시스템의 3D 디스플레이의 예들
- [0017] [0032] 웨어러블 시스템(본원에서 AR(augmented reality) 시스템으로 또한 지칭됨)은 2D 또는 3D 가상 이미지를 사용자에게 제시하도록 구성될 수 있다. 이미지들은 정지 이미지들, 비디오의 프레임들 또는 비디오, 이들의 조합 등일 수 있다. 웨어러블 시스템의 적어도 일부는 사용자 상호작용을 위해 VR, AR 또는 MR 환경을 단독으로 또는 조합하여 제시할 수 있는 웨어러블 디바이스 상에서 구현될 수 있다. 웨어러블 디바이스는 ARD(AR device)로서 상호 교환 가능하게 사용되는 HMD(head-mounted device)일 수 있다. 또한, 본 개시내용의 목적을 위해, "AR"이란 용어는 "MR"이란 용어와 상호 교환 가능하게 사용된다.
- [0018] [0033] 도 1은 사람에 의해 뷰잉되는 소정의 물리적 객체들 및 소정의 가상 현실 객체들을 갖는 혼합 현실 시나리오의 예시를 도시한다. 도 1에서, MR 장면(100)이 도시되며, 여기서 MR 기술의 사용자는 배경에 있는 사람들, 나무들, 빌딩들, 및 콘크리트 플랫폼(120)을 피쳐링(featuring)하는 실세계 공원-형 세팅(110)을 본다. 이들 아이템들에 더하여, MR 기술의 사용자는 또한, 그가 실세계 플랫폼(120) 상에서 있는 로봇 동상(130), 및 호박벌의 의인화인 것으로 보여지는 날고 있는 만화-형 아바타 캐릭터(140)를 보는 것을 지각하더라도, 이들 엘리먼트들은 실세계에 존재하지 않는다.

- [0019] [0034] 3D 디스플레이가 진정한 깊이감(sensation of depth) 및 보다 구체적으로, 시뮬레이팅된 표면 깊이감을 생성하기 위해, 디스플레이의 시계(visual field)의 각각의 지점이 그의 가상 깊이에 대응하는 원근조절 응답을 생성하는 것이 바람직할 수 있다. 디스플레이 지점에 대한 원근조절 응답이 수렴 및 입체시(stereopsis)의 양안 깊이 단서들에 의해 결정된 바와 같은 그 지점의 가상 깊이에 대응하지 않는 경우, 인간의 눈은 원근조절 충돌을 경험할 수 있어, 불안정한 이미징, 유해한 눈의 피로, 두통들, 그리고 원근조절 정보의 부재 시에, 표면 깊이의 거의 완전한 결여를 초래할 수 있다.
- [0020] [0035] VR, AR 및 MR 경험들은 복수의 깊이 평면들에 대응하는 이미지들이 뷰어에게 제공되는 디스플레이들을 갖는 디스플레이 시스템들에 의해 제공될 수 있다. 이미지들은 각각의 깊이 평면마다 상이할 수 있고 (예컨대, 장면 또는 객체의 약간 다른 프리젠테이션들을 제공함) 뷰어의 눈들에 의해 별개로 포커싱될 수 있어서, 상이한 깊이 평면 상에 로케이팅되는 장면에 대한 상이한 이미지 피쳐들에 포커스를 맞추도록 요구되는 눈의 원근조절에 기초하여 또는 상이한 깊이 평면들 상의 상이한 이미지 피쳐들이 아웃 포커스(out of focus)되는 것을 관찰하는 것에 기초하여 깊이 단서들을 사용자에게 제공하는 것을 돋는다. 본원의 다른 곳에서 논의된 바와 같이, 이러한 깊이 단서들은 깊이의 신뢰할 수 있는 지각들을 제공한다.
- [0021] [0036] 도 2는 AR/VR/MR 장면을 제공하도록 구성될 수 있고 본원에서 설명된 단어 흐름 주석 시스템의 실시예들을 포함할 수 있는 웨어러블 시스템(200)의 예를 예시한다. 웨어러블 시스템(200)은 또한 AR 시스템(200)으로서 지칭될 수 있다. 웨어러블 시스템(200)은 디스플레이(220), 및 디스플레이(220)의 기능을 지원하기 위한 다양한 기계적 및 전자적 모듈들 및 시스템들을 포함한다. 디스플레이(220)는 사용자, 착용자 또는 뷰어(210)에 의해 착용 가능한 프레임(230)에 커플링될 수 있다. 디스플레이(220)는 사용자(210)의 눈들 앞에 포지셔닝될 수 있다. 디스플레이(220)는 AR/VR/MR 콘텐츠를 사용자에게 제시할 수 있다. 디스플레이(220)는 사용자의 머리 상에 착용되는 머리 장착 디스플레이를 포함할 수 있다. 일부 실시예들에서, 스피커(240)는 프레임(230)에 커플링되고 사용자의 외이도에 인접하게 포지셔닝된다(일부 실시예들에서, 도시되지 않은 다른 스피커가 사용자의 다른 외이도에 인접하게 포지셔닝되어 스템레오/성형 가능(shapeable) 사운드 제어를 제공함). 디스플레이(220)는 환경으로부터 오디오 스트림을 검출하고 주변 사운드를 캡처하기 위한 오디오 센서(232)(예컨대, 마이크로폰)를 포함할 수 있다. 일부 실시예들에서, 하나 이상의 다른 오디오 센서들(도시되지 않음)이 스템레오 사운드 수신을 제공하도록 포지셔닝된다. 스템레오 사운드 수신은 사운드 소스의 위치를 결정하는 데 사용될 수 있다. 웨어러블 시스템(200)은 오디오 스트림 상에서 음성 또는 스피치 인식을 수행할 수 있다.
- [0022] [0037] 웨어러블 시스템(200)은 사용자 주위의 환경의 세계를 관찰하는 외향 이미징 시스템(464)(도 4에 도시됨)을 포함할 수 있다. 웨어러블 시스템(200)은 또한 사용자의 눈 움직임들을 추적할 수 있는 내향 이미징 시스템(462)(도 4에 도시됨)을 포함할 수 있다. 내향 이미징 시스템은 어느 한쪽 눈의 움직임들 또는 양쪽 눈의 움직임들을 추적할 수 있다. 내향 이미징 시스템(462)은 프레임(230)에 부착될 수 있고, 프로세싱 모듈(260 또는 270)과 전기 통신할 수 있으며, 이 프로세싱 모듈(260 또는 270)은 예컨대, 사용자(210)의 동공 직경들 또는 눈들의 배향들, 눈 움직임들 또는 눈 포즈(pose)를 결정하도록 내향 이미징 시스템에 의해 획득된 이미지 정보를 프로세싱할 수 있다.
- [0023] [0038] 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.
- [0024] [0039] 디스플레이(220)는 이를테면, 유선 리드 또는 무선 연결성에 의해, 다양한 구성들로 장착될 수 있는, 이를테면, 프레임(230)에 고정되게 부착되거나, 사용자에 의해 착용된 헬멧 또는 모자에 고정되게 부착되거나, 헤드폰들에 내장되거나, 그렇지 않으면 사용자(210)에게 제거 가능하게 부착되는 (예컨대, 백팩(backpack)-스타일 구성으로, 벨트-커플링 스타일 구성으로) 로컬 데이터 프로세싱 모듈(260)에 동작 가능하게 커플링(250)될 수 있다.
- [0025] [0040] 로컬 프로세싱 및 데이터 모듈(260)은 하드웨어 프로세서는 물론, 디지털 메모리 이를테면, 비-휘발성 메모리(예컨대, 플래시 메모리)를 포함할 수 있고, 이 둘 모두는 데이터의 프로세싱, 캐싱(caching) 및 저장을 보조하기 위해 활용될 수 있다. 데이터는 a) 센서들(예컨대, 프레임(230)에 동작 가능하게 커플링되거나 그렇지 않으면 사용자(210)에게 부착될 수 있음), 이를테면, 이미지 캡처 디바이스들(예컨대, 내향 이미징 시스템 또는 외향 이미징 시스템의 카메라들), 오디오 센서들(예컨대, 마이크로폰들), IMU(inertial measurement unit)들, 가속도계들, 컴퍼스(compass)들, GPS(global positioning system) 유닛들, 라디오 디바이스들, 또는 자이ロ스코프들로부터 캡처되고; 또는 b) 원격 프로세싱 모듈(270) 또는 원격 데이터 리포지토리

(repository)(280)를 사용하여 획득 또는 프로세싱되는(가능하게는, 이러한 프로세싱 또는 리트리벌(retrieval) 이후 디스플레이(220)에 전달하기 위한) 데이터를 포함할 수 있다. 로컬 프로세싱 및 데이터 모듈(260)은 통신 링크들(262 또는 264)에 의해, 이를테면, 유선 또는 무선 통신 링크들을 통하여, 원격 프로세싱 모듈(270) 또는 원격 데이터 리포지토리(280)에 동작 가능하게 커플링될 수 있어서, 이들 원격 모듈들은 로컬 프로세싱 및 데이터 모듈(260)에 대한 자원들로서 이용 가능하다. 또한, 원격 프로세싱 모듈(280) 및 원격 데이터 리포지토리(280)는 서로 동작 가능하게 커플링될 수 있다.

[0026] 일부 실시예들에서, 원격 프로세싱 모듈(270)은 데이터 또는 이미지 정보를 분석 및 프로세싱하도록 구성된 하나 이상의 프로세서들을 포함할 수 있다. 일부 실시예들에서, 원격 데이터 리포지토리(280)는 "클라우드" 자원 구성에서 인터넷 또는 다른 네트워킹 구성을 통하여 이용 가능할 수 있는 디지털 데이터 저장 설비를 포함할 수 있다. 일부 실시예들에서, 모든 데이터는 저장되고 모든 컴퓨테이션들은 로컬 프로세싱 및 데이터 모듈에서 수행되어, 원격 모듈로부터 완전히 자율적인 사용을 허용한다.

[0042] 인간 시각 시스템은 복잡하고 현실적인 깊이의 지각을 제공하는 것은 난제이다. 이론에 의해 제한됨이 없이, 객체의 뷰어들은 이접운동 및 원근조절의 조합으로 인해 객체를 "3-차원"인 것으로 지각할 수 있다고 여겨진다. 서로에 대한 두 눈들의 이접운동(vergence) 움직임들(즉, 객체를 응시하기 위해 눈들의 시선들을 수렴하도록 서로를 향하는 또는 서로 멀어지는 동공들의 롤링(rolling) 움직임들)은 눈들의 렌즈들의 포커싱(또는 "원근조절")과 밀접하게 연관된다. 정상 조건들하에서, 상이한 거리에 있는 하나의 객체로부터 다른 객체로 포커스를 변화시키기 위하여, 눈들의 렌즈들의 포커스를 변화시키거나, 또는 눈들을 원근조절하는 것은 "원근조절-이접운동 반사(accommodation-vergence reflex)"로서 알려진 관계하에서, 동일한 거리에 대한 이접운동에서의 매칭하는 변화를 자동으로 유발할 것이다. 마찬가지로, 이접운동의 변화는 정상 조건들하에서, 원근조절의 매칭 변화를 트리거할 것이다. 원근조절과 이접운동 사이의 더 양호한 매칭을 제공하는 디스플레이 시스템들은 3-차원 이미저리의 더 현실적이고 편안한 시뮬레이션들을 형성할 수 있다.

[0043] 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 객체들은, 이들 객체들이 인 포커싱(in focus)되도록 눈들(302 및 304)에 의해 원근조절된다. 눈들(302 및 304)은 z-축을 따라 상이한 거리들에 있는 객체들에 포커싱을 맞추게 하는 특정 원근조절된 상태들을 취한다. 결과적으로, 특정 원근조절된 상태는 연관된 초점 거리를 갖는, 깊이 평면들(306) 중 특정한 하나의 깊이 평면과 연관되는 것으로 말해질 수 있어서, 특정 깊이 평면의 객체들 또는 객체들의 부분들은, 눈이 해당 깊이 평면에 대해 원근조절된 상태에 있을 때 인 포커싱된다. 일부 실시예들에서, 3-차원 이미저리는 눈들(302 및 304) 각각에 대해 이미지의 상이한 프리젠테이션들을 제공함으로써, 그리고 또한 깊이 평면들 각각에 대응하는 이미지의 상이한 프리젠테이션들을 제공함으로써 시뮬레이팅될 수 있다. 예시의 명확성을 위해 별개인 것으로 도시되지만, 눈들(302 및 304)의 시야들은 예컨대, z-축을 따른 거리가 증가함에 따라 겹쳐질 수 있다는 것이 인지될 것이다. 게다가, 예시의 용이함을 위해 평평한 것으로 도시되지만, 깊이 평면의 윤곽들은 물리적 공간에서 만곡될 수 있어서, 깊이 평면의 모든 피처들은 특정 원근조절된 상태에서 눈과 인 포커싱된다는 것이 인지될 것이다. 이론에 의해 제한됨이 없이, 인간 눈이 통상적으로 깊이 지각을 제공하기 위하여 유한 수의 깊이 평면들을 해석할 수 있다고 여겨진다. 결과적으로, 지각된 깊이의 매우 믿을 만한 시뮬레이션은, 눈에, 이들 제한된 수의 깊이 평면들 각각에 대응하는 이미지의 상이한 프리젠테이션들을 제공함으로써 달성될 수 있다.

도파관 스택 어셈블리

[0044] 도 4는 이미지 정보를 사용자에게 출력하기 위한 도파관 스택의 예를 예시한다. 웨어러블 시스템(400)은 복수의 도파관들(432b, 434b, 436b, 438b, 4400b)을 사용하여 3-차원 지각을 눈/뇌에 제공하기 위하여 활용될 수 있는 도파관들의 스택, 또는 스택된 도파관 어셈블리(480)를 포함한다. 일부 실시예들에서, 웨어러블 시스템(400)은 도 2의 웨어러블 시스템(200)에 대응할 수 있고, 도 4는 그 웨어러블 시스템(200)의 일부 부분들을 더 상세히 개략적으로 보여준다. 예컨대, 일부 실시예들에서, 도파관 어셈블리(480)는 도 2의 디스플레이(220)에 통합될 수 있다.

[0045] 도 4를 계속 참조하면, 도파관 어셈블리(480)는 또한 도파관들 사이에 복수의 피처들(458, 456, 454, 452)을 포함할 수 있다. 일부 실시예들에서, 피처들(458, 456, 454, 452)은 렌즈들일 수 있다. 다른 실시예들에서, 피처들(458, 456, 454, 452)은 렌즈들이 아닐 수 있다. 오히려, 이들은 단순히 스페이서들(예컨대, 에어캡들을 형성하기 위한 클래딩 층들 또는 구조들)일 수 있다.

[0046] 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 개별 도파관을 가로질러 인입 광을 분산시키도록 구성될 수 있다. 광은 이미지 주입 디바이스들(420, 422, 424, 426, 428)의 출력 표면을 나가고 도파관들(440b, 438b, 436b, 434b, 432b)의 대응하는 입력 예지에 주입된다. 일부 실시예들에서, 단일 광 빔(예컨대, 시준된 빔)은 특정 도파관과 연관된 깊이 평면에 대응하는 특정 각도들(및 발산의 양들)로 눈(410)을 향하여 지향되는 시준된 클론 빔(cloned collimated beam)들의 전체 필드를 출력하기 위하여 각각의 도파관으로 주입될 수 있다.

[0033] [0047] 일부 실시예들에서, 이미지 주입 디바이스들(420, 422, 424, 426, 428)은 각각, 대응하는 도파관(440b, 438b, 436b, 434b, 432b)으로의 주입을 위한 이미지 정보를 각각 생성하는 이산 디스플레이들이다. 일부 다른 실시예들에서, 이미지 주입 디바이스들(420, 422, 424, 426, 428)은 예컨대, 이미지 정보를 하나 이상의 광학 도파관들(예컨대, 광섬유 케이블들)을 통하여 이미지 주입 디바이스들(420, 422, 424, 426, 428) 각각에 파이프(pipe)할 수 있는 단일 멀티플렉싱된 디스플레이의 출력 단부들이다.

[0034] [0048] 제어기(460)는 스택된 도파관 어셈블리(480) 및 이미지 주입 디바이스들(420, 422, 424, 426, 428)의 동작을 제어한다. 제어기(460)는 도파관들(440b, 438b, 436b, 434b, 432b)에 대한 이미지 정보의 타이밍 및 프로비전(provision)을 레귤레이팅하는 프로그래밍(예컨대, 비-일시적 컴퓨터-판독 가능 매체의 명령들)을 포함한다. 일부 실시예들에서, 제어기(460)는 단일 통합 디바이스, 또는 유선 또는 무선 통신 채널들에 의해 연결되는 분산 시스템일 수 있다. 제어기(460)는 일부 실시예들에서, 프로세싱 모듈들(260 또는 270)(도 2에 예시됨)의 부분일 수 있다.

[0035] [0049] 도파관들(440b, 438b, 436b, 434b, 432b)은 TIR(total internal reflection)에 의해 각각의 개별 도파관 내에서 광을 전파시키도록 구성될 수 있다. 도파관들(440b, 438b, 436b, 434b, 432b)은 각각 평면형이거나 다른 형상(예컨대, 곡선)을 가질 수 있으며, 주 최상부 및 최하부 표면들 및 이들 주 최상부와 최하부 표면들 사이에서 연장되는 예지들을 갖는다. 예시된 구성에서, 도파관들(440b, 438b, 436b, 434b, 432b)은 이미지 정보를 눈(410)에 출력하기 위해 각각의 개별 도파관 내에서 전파되는 광을 도파관 밖으로 재지향시킴으로써 도파관으로부터 광을 추출하도록 구성된 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)을 각각 포함할 수 있다. 추출된 광은 아웃커플링된 광으로서 또한 지칭될 수 있고, 광 추출 광학 엘리먼트들은 또한 아웃커플링 광학 엘리먼트들로서 지칭될 수 있다. 추출된 광 빔은, 도파관 내에서 전파되는 광이 광 재지향 엘리먼트에 부딪치는 위치들에서 도파관에 의해 출력된다. 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 예컨대, 반사성 및/또는 회절성 광학 피처들일 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 중에 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.

[0036] [0050] 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 제 1 렌즈(452)(예컨대, 네거티브 렌즈)를 통과하는 시준된 광을 보내도록 구성될 수 있다. 제 1 렌즈(452)는 약간의 볼록한 파면 곡률을 생성하도록 구성될 수 있어서, 눈/뇌가 상기 다음 위의 도파관(434b)으로부터 오는 광을 광학 무한대로부터 눈(410)을 향해 안쪽으로 더 가까운 제 1 초점 평면으로부터 오는 것으로 해석한다. 유사하게, 세번째 위의 도파관(436b)은 그 출력 광을, 눈(410)에 도달하기 전에 제 1 렌즈(452) 및 제 2 렌즈(454) 둘 모두를 통과시킨다. 제 1 및 제 2 렌즈들(452 및 454)의 결합된 광학 전력은 다른 충분 양의 파면 곡률을 생성하도록 구성될 수 있어서, 눈/뇌는 제 3 도파관(436b)으로부터 오는 광을, 상기 다음 위의 도파관(434b)으로부터의 광보다는 광학 무한대로부터 사람을 향하여 안쪽으로 훨씬 더 가까운 제 2 초점 평면으로부터 오는 것으로 해석한다.

- [0037] [0051] 다른 도파관 층들(예컨대, 도파관들(438b, 440b)) 및 렌즈들(예컨대, 렌즈들(456, 458))은 유사하게 구성되는데, 스택에서 가장 높은 도파관(440b)은 자신의 출력을, 사람과 가장 가까운 초점 평면을 나타내는 어그리게이트 초점 전력에 대해 자신과 눈 사이의 렌즈들 모두를 통하여 전송한다. 스택된 도파관 어셈블리(480)의 다른 층 상에서 세계(470)로부터 오는 광을 보거나/해석할 때 렌즈들(458, 456, 454, 452)의 스택을 보상하기 위하여, 보상 렌즈 층(430)은 아래의 렌즈 스택(458, 456, 454, 452)의 어그리게이트 파워를 보상하기 위하여 스택의 최상부에 배치될 수 있다. 이러한 구성은 이용 가능한 도파관/렌즈 쌍들이 존재하는 만큼 많은 지각된 초점 평면들을 제공한다. 도파관들의 광 추출 광학 엘리먼트들 및 렌즈들의 포커싱 양상들 둘 모두는 정적(예컨대, 동적이 아니거나 전자-활성이 아님)일 수 있다. 일부 대안적인 실시예들에서, 어느 하나 또는 둘 모두는 전자-활성 피쳐들을 사용하여 동적일 수 있다.
- [0038] [0052] 도 4를 계속 참조하면, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 자신의 개별 도파관들로부터 광을 재지향하는 것은 물론, 도파관과 연관된 특정 깊이 평면에 대해 적절한 양의 발산 또는 시준으로 이 광을 출력하도록 구성될 수 있다. 결과로서, 상이한 연관된 깊이 평면들을 가진 도파관들은 상이한 구성들의 광 추출 광학 엘리먼트들을 가질 수 있고, 이러한 광 추출 광학 엘리먼트들은 연관된 깊이 평면에 따라 상이한 양의 발산으로 광을 출력한다. 일부 실시예들에서, 본원에 논의된 바와 같이, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 특정 각도들로 광을 출력하도록 구성될 수 있는 볼륨메트릭(volumetric) 또는 표면 피쳐들일 수 있다. 예컨대, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 볼륨 홀로그램들, 표면 홀로그램들, 및/또는 회절 격자들일 수 있다. 광 추출 광학 엘리먼트들, 이를테면, 회절 격자들은 2015년 6월 25일에 공개된 미국 특허 공개 번호 제 2015/0178939호에 설명되며, 이는 그 전체가 본원에 인용에 의해 포함된다.
- [0039] [0053] 일부 실시예들에서, 광 추출 광학 엘리먼트들(440a, 438a, 436a, 434a, 432a)은 회절 패턴을 형성하는 회절 피쳐들 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.
- [0040] [0054] 일부 실시예들에서, 하나 이상의 DOE들은, 그것들이 활발하게 회절시키는 "온" 상태와 그것들이 크게 회절시키지 않는 "오프" 상태 사이에서 스위칭 가능할 수 있다. 예컨대, 스위칭 가능 DOE는, 마이크로액적들이 호스트 매질에서 회절 패턴을 포함하는 폴리머 분산형 액정 층을 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 활발하게 회절시킴).
- [0041] [0055] 일부 실시예들에서, 피사계 심도 또는 깊이 평면들의 수 및 분포는 뷰어의 눈들의 동공 크기들 또는 배향들에 기초하여 동적으로 변동될 수 있다. 피사계 심도는 뷰어의 동공 크기와 반대로 변경될 수 있다. 결과적으로, 뷰어의 눈들의 동공들의 크기들이 감소함에 따라, 피사계 심도가 증가하여서, 식별 불가능한 하나의 평면의 위치가 눈의 포커스의 깊이를 넘어서기 때문에 그 평면이 식별 가능하게 되고 동공 크기의 감소를 통해 보다 인 포커스로 나타나고 피사계 심도의 증가와 상응할 수 있다. 마찬가지로, 뷰어에게 상이한 이미지들을 제시하는 데 사용되는 이격된 깊이 평면들의 수는 감소된 동공 크기에 따라 감소될 수 있다. 예컨대, 뷰어는 하나의 깊이 평면으로부터 벗어나게 그리고 다른 깊이 평면으로 눈의 원근조절을 조정하지 않고서는, 하나의 동공 크기에서 제 1 깊이 평면 및 제 2 깊이 평면 둘 모두의 세부사항들을 명확하게 지각할 수 없을 수 있다. 그러나, 이러한 2개의 깊이 평면들은 원근조절을 변경하지 않고도 다른 동공 크기에서 사용자에게 동시에 충분히 인 포커스일 수 있다.
- [0042] [0056] 일부 실시예들에서, 디스플레이 시스템은 동공 크기 또는 배향의 결정들에 또는 특정 동공 크기 또는 배향을 나타내는 전기 신호들의 수신에 기초하여 이미지 정보를 수신하는 도파관들의 수를 변동시킬 수 있다. 예컨대, 사용자의 눈들이 2개의 도파관들과 연관된 2개의 깊이 평면들을 구별할 수 없는 경우, 제어기(460)(로컬 프로세싱 및 데이터 모듈(260)의 실시예일 수 있음)는 이를 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.

- [0043] [0057] 일부 실시예들에서, 출사 범이 뷰어의 눈의 직경 미만인 직경을 갖는 조건을 충족시키는 것이 바람직할 수 있다. 그러나 이 조건을 충족시키는 것은 뷰어의 동공들의 크기의 변동성을 고려하면 난제가 될 수 있다. 일부 실시예들에서, 이 조건은 뷰어의 동공의 크기의 결정에 대한 응답으로 출사 범의 크기를 변동시킴으로써 광범위한 동공 크기들에 걸쳐 충족된다. 예컨대, 동공 크기가 감소함에 따라, 출사 범의 크기가 또한 감소할 수 있다. 일부 실시예들에서, 출사 범 크기는 가변 어페처를 사용하여 변동될 수 있다.
- [0044] [0058] 웨어러블 시스템(400)은 세계(470)의 일부를 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 뷰잉을 위해 이용 가능한 전체 구역은 FOR(field of regard)로서 지칭될 수 있다. 착용자가 자신의 신체, 머리 또는 눈들을 움직여 실질적으로 공간의 임의의 방향을 지각할 수 있기 때문에 FOR은 웨어러블 시스템(400)을 둘러싸는 4π 스테라디안(steradian)들의 입체각을 포함할 수 있다. 다른 맥락들에서, 착용자의 움직임들은 보다 억제될 수 있고, 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 객체들을 검출하는 등을 행하는 데 사용될 수 있다.
- [0045] [0059] 웨어러블 시스템(400)은 주변 사운드를 캡처하기 위한 오디오 센서(232), 예컨대, 마이크로폰을 포함할 수 있다. 위에서 설명된 바와 같이, 일부 실시예들에서, 하나 이상의 다른 오디오 센서들은 스피치 소스의 위치 결정에 유용한 스테레오 사운드 수신을 제공하도록 포지셔닝될 수 있다. 오디오 센서(232)는 다른 예로서, 오디오 소스가 로케이팅하는 곳에 대한 그러한 유용한 지향성 정보를 또한 제공할 수 있는 지향성 마이크로폰을 포함할 수 있다. 웨어러블 시스템(400)은, 스피치의 소스의 로케이팅하는 데 있어 또는 특정 시간 순간의 활성 스피커를 결정하는 등을 위해 외향 이미징 시스템(464) 및 오디오 센서(230) 둘 모두로부터의 정보를 사용할 수 있다. 예컨대, 웨어러블 시스템(400)은, 스피커의 아이덴티티를 결정하기 위해 음성 인식을 단독으로, 또는 스피커의 반사된 이미지(예컨대, 미러에서 볼 수 있음)와 함께 사용할 수 있다. 다른 예로서, 웨어러블 시스템(400)은 지향성 마이크로폰들로부터 획득된 사운드에 기초하여 환경 내의 스피커의 포지션을 결정할 수 있다. 웨어러블 시스템(400)은 스피치의 콘텐츠를 결정하기 위해 스피치 인식 알고리즘들로 스피커의 포지션으로부터 나오는 사운드를 파싱(parse)하고 스피커의 아이덴티티(예컨대, 이름 또는 다른 인구학적 정보)를 결정하기 위해 음성 인식 기술들을 사용할 수 있다.
- [0046] [0060] 웨어러블 시스템(400)은 또한 눈 움직임들 및 얼굴 움직임들과 같은 사용자의 움직임들을 관찰하는 내향 이미징 시스템(466)(예컨대, 디지털 카메라)을 포함할 수 있다. 내향 이미징 시스템(466)은 눈(304)의 동공의 크기 및/또는 배향을 결정하기 위해 눈(410)의 이미지들을 캡처하는 데 사용될 수 있다. 내향 이미징 시스템(466)은, (예컨대, 홍채 식별을 통해) 사용자의 바이오메트릭 식별(biometric identification)을 위한 이미지들 또는 사용자가 바라보는 방향(예컨대, 눈 포즈)을 결정하는 데 사용하기 위한 이미지들을 획득하는 데 사용될 수 있다. 일부 실시예들에서, 독립적으로, 각각의 눈의 동공 크기 또는 눈 포즈를 별개로 결정하고, 그리하여 각각의 눈에 대한 이미지 정보의 프리젠테이션이 그 눈에 동적으로 맞춰지도록 허용하기 위해, 각각의 눈마다 적어도 하나의 카메라가 활용될 수 있다. 일부 다른 실시예들에서, (예컨대, 한 쌍의 눈들 당 단지 단일 카메라만을 사용하여) 단지 한쪽 눈(410)의 동공 직경 또는 배향이 결정되고 사용자의 양 눈들에 대해 유사한 것으로 가정된다. 내향 이미징 시스템(466)에 의해 획득된 이미지들은 사용자의 눈 포즈 또는 분위기(mood)를 결정하도록 분석될 수 있으며, 이는 어떤 청각적 또는 시각적 콘텐츠가 사용자에게 제시되어야 하는지를 결정하기 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.
- [0047] [0061] 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, 토텟(예컨대, 가상 사용자 입력 디바이스로서 기능함) 등을 포함할 수 있다. DOF 제어기는 그 제어기의 일부 또는 모든 가능한 병진운동(translational)들(예컨대, 좌/우, 전/후 또는 위/아래) 또는 회전들(예컨대, 요(yaw), 페치 또는 롤)로 사용자 입력을 감지할 수 있다. 병진운동 움직임들을 지원하는 다중-DOF 제어기는 3DOF로서 지칭될 수 있는 반면, 병진운동들 및 회전들을

지원하는 다중-DOF 제어기는 6DOF로서 지칭될 수 있다. 일부 경우들에서, 사용자는 손가락(예컨대, 엄지 손가락)을 사용하여, 웨어러블 시스템(400)에 입력을 제공하도록(예컨대, 웨어러블 시스템(400)에 의해 제공되는 사용자 인터페이스에 사용자 입력을 제공하도록) 터치-감지 입력 디바이스를 누르거나 스와이프(swipe)할 수 있다. 사용자 입력 디바이스(466)는 웨어러블 시스템(400)의 사용 동안 사용자의 손에 의해 보유될 수 있다. 사용자 입력 디바이스(466)는 웨어러블 시스템(400)과 유선 또는 무선 통신할 수 있다.

[0048] 도 5는 도파관에 의해 출력된 출사 빔들의 예를 도시한다. 하나의 도파관이 예시되지만, 도파관 어셈블리(480) 내의 다른 도파관들이 유사하게 기능할 수 있다는 것이 인지될 것이며, 여기서 도파관 어셈블리(480)는 다수의 도파관들을 포함한다. 광(520)은 도파관(432b)의 입력 에지(432c)에서 도파관(432b)으로 주입되고 TIR에 의해 도파관(432b) 내에서 전파된다. 광(520)이 DOE(432a)에 충돌하는 지점들에서, 광의 일부는 출사 빔들(510)로서 도파관을 나간다. 출사 빔들(510)은 실질적으로 평행한 것으로 예시되지만, 이들 출사 빔들(510)은 또한 도파관(432b)과 연관된 깊이 평면에 따라, 임의의 각도로 눈(410)으로 전파되도록 재지향될 수 있다(예컨대, 별산하는 출사 빔들을 형성함). 실질적으로 평행한 출사 빔들은, 눈(410)으로부터 먼 거리(예컨대, 광학적 무한대)에 있는 깊이 평면 상에 세팅된 것으로 보이는 이미지들을 형성하도록 광을 아웃커플링하는 광 추출 광학 엘리먼트들을 갖는 도파관을 나타낼 수 있다는 것이 인지될 것이다. 다른 도파관들 또는 광 추출 광학 엘리먼트들의 다른 세트들은 더 별산하는 출사 빔 패턴을 출력할 수 있고, 이는 눈(410)이 망막 상에 포커싱을 맞추게 하기 위해 더 가까운 거리로 원근조절하는 것을 요구할 것이고 광학적 무한대보다 눈(410)에 더 가까운 거리로부터의 광으로서 뇌에 의해 해석될 것이다.

[0049] 도 6은, 다중-초점 볼류메트릭 디스플레이, 이미지 또는 광 필드의 생성에 사용되는, 도파관 장치, 도파관 장치로 또는 도파관 장치로부터의 광을 광학적으로 커플링하는 광학 커플러 서브시스템, 및 제어 서브시스템을 포함하는 광학 시스템을 도시하는 개략도이다. 광학 시스템은 도파관 장치, 도파관 장치로 또는 상기 도파관 장치로부터의 광을 광학적으로 커플링하는 광학 커플러 서브시스템, 및 제어 서브시스템을 포함할 수 있다. 광학 시스템은 다중-초점 볼류메트릭, 이미지 또는 광 필드를 생성하는 데 사용할 수 있다. 광학 시스템은 하나 이상의 주 평면 도파관들(632a)(도 6에서 단지 하나만 도시됨) 및 주 도파관들(632a) 중 적어도 일부의 주 도파관들 각각과 연관된 하나 이상의 DOE들(632b)을 포함할 수 있다. 평면 도파관들(632b)은 도 4를 참조하여 눈의된 도파관들(432b, 434b, 436b, 438b, 440b)과 유사할 수 있다. 광학 시스템은 제 1 축(도 6의 뷰에서 수직 또는 Y-축)을 따라 광을 중계하고 제 1 축(예컨대, Y-축)을 따라 광의 유효 출사동을 확장시키기 위해 분배 도파관 장치를 사용할 수 있다. 분배 도파관 장치는, 예컨대, 분배 평면 도파관(622b) 및 분배 평면 도파관(622b)과 연관된 적어도 하나의 DOE(622a)(이중 일점 쇄선으로 예시됨)를 포함할 수 있다. 분배 평면 도파관(622b)은 그와 상이한 배향을 갖는 주 평면 도파관(632b)과 적어도 일부 면들에서 유사하거나 동일할 수 있다. 마찬가지로, 적어도 하나의 DOE(622a)는 DOE(632a)와 적어도 일부 면들에서 유사하거나 동일할 수 있다. 예컨대, 분배 평면 도파관(622b) 또는 DOE(622a)는 각각, 주 평면 도파관(632b) 또는 DOE(632a)와 동일한 재료들로 구성될 수 있다. 도 6에 도시된 광학 디스플레이 시스템(600)의 실시예들은 도 2에 도시된 웨어러블 시스템(200)에 통합될 수 있다.

[0050] 중계된 그리고 출사동 확장된 광은 분배 도파관 장치로부터 하나 이상의 주 평면 도파관들(632b)로 광학적으로 커플링될 수 있다. 주 평면 도파관(632b)은 바람직하게는, 제 1 축과 직교하는 제 2 축(예컨대, 도 6의 뷰에서 수평 또는 X-축)을 따라 광을 중계할 수 있다. 특히, 제 2 축은 제 1 축에 대해 비-직교 축일 수 있다. 주 평면 도파관(632b)은 제 2 축(예컨대, X-축)을 따라 광의 유효 출사동을 확장시킨다. 예컨대, 분배 평면 도파관(622b)은 수직 또는 Y-축을 따라 광을 중계 및 확장시키고, 수평 또는 X-축을 따라 광을 중계 및 확장시킬 수 있는 주 평면 도파관(632b)으로 그 광을 전달할 수 있다.

[0051] 광학 시스템은 단일 모드 광섬유(640)의 근위 단부에 광학적으로 커플링될 수 있는 커러 광(예컨대, 적색, 녹색 및 청색 레이저 광)의 하나 이상의 소스들(610)을 포함할 수 있다. 광 섬유(640)의 원위 단부는 압전 재료의 중공 튜브(642)를 통해 수용되거나 스레딩(thread)될 수 있다. 원위 단부는 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, 외부 표면 또는 외부 주변부 또는 직경 상에 도금될 수 있다. 코어 전극(예시되지 않음)은 또한 튜브(642)의 코어, 중앙, 내부 주변부 또는 내부 직경에 로케이팅될 수 있다.

[0052] 예컨대, 와이어들(660)을 통해 전기적으로 커플링된 구동 전자 장치(650)는 2개의 축들에서 압전 튜브(642)를 독립적으로 구부리기 위해 대향하는 전극 쌍들을 구동한다. 광섬유(644)의 돌출 원위 팁은 기계적 공진 모드들을 갖는다. 공진 주파수들은 광섬유(644)의 직경, 길이 및 재료 성질들에 의존할 수 있다. 섬유 캔

틸레버(644)의 제 1 기계적 공진 모드 근처에서 압전 튜브(642)를 진동시킴으로써, 섬유 캔틸레버(644)는 진동하게 될 수 있고, 큰 편향들을 통해 스윕(sweep)할 수 있다.

[0053] [0067] 2개의 축들에서 공진 진동을 자극함으로써, 섬유 캔틸레버(644)의 팁은 영역 필링 2-차원(2D) 스캔으로 2축 방향으로(biaxially) 스캔된다. 섬유 캔틸레버(644)의 스캔과 동기하여 광 소스(들)(610)의 세기를 변조함으로써, 섬유 캔틸레버(644)로부터 나오는 광은 이미지를 형성할 수 있다. 그러한 셋업에 대한 설명들은 그 전체가 인용에 의해 본원에 포함되는 미국 특허 공보 제 2014/0003762호에서 제공된다.

[0054] [0068] 광학 커플러 시브시스템의 컴포넌트는 스캐닝 섬유 캔틸레버(644)로부터 나오는 광을 시준할 수 있다. 시준된 광은 미러 표면(mirrored surface)(648)에 의해 적어도 하나의 DOE(diffractive optical element)(622a)를 포함하는 좁은 분배 평면 도파관(622b)으로 반사될 수 있다. 시준된 광은 TIR에 의해 분배 평면 도파관(622b)을 따라(도 6의 뷰에 대해) 수직으로 전파될 수 있고, 이렇게 하여, DOE(622a)와 반복적으로 교차한다. DOE(622a)는 바람직하게는, 낮은 회절 효율을 갖는다. 이는, 광의 프랙션(fraction)(예컨대, 10%)이 DOE(622a)와의 각각의 교차 지점에서 더 큰 주 평면 도파관(632b)의 에지를 향해 회절되게 하고, 광의 프랙션이 TIR을 통해 분배 평면 도파관(622b)의 길이 아래에서 그의 원래의 궤적을 지속하게 할 수 있다.

[0055] [0069] DOE(622a)와의 각각의 교차 지점에서, 부가적인 광이 주 도파관(632b)의 입구를 향해 회절될 수 있다. 인입 광을 다수의 아웃커플링된 세트들로 분할함으로써, 광의 출사동은 분배 평면 도파관(622b)에서 DOE(622a)에 의해 수직으로 확장될 수 있다. 분배 평면 도파관(622b) 밖으로 커플링되는 이러한 수직으로 확장된 광은 주 평면 도파관(632b)의 에지에 진입할 수 있다.

[0056] [0070] 주 도파관(632b)에 진입하는 광은 TIR을 통해 주 도파관(632b)을 따라(도 6의 뷰에 대해) 수평으로 전파될 수 있다. 광이 TIR을 통해 주 도파관(632b)의 길이의 적어도 일부를 따라 수평으로 전파됨에 따라 광은 다수의 지점들에서 DOE(632a)와 교차한다. DOE(632a)는 유리하게는, 광의 편향 및 포커싱 둘 모두를 생성하도록, 선형 회절 패턴 및 방사상 대칭 회절 패턴의 합인 위상 프로파일(phase profile)을 갖도록 설계 또는 구성될 수 있다. DOE(632a)는 유리하게는, 낮은 회절 효율(예컨대, 10%)을 가질 수 있어서, 빔의 광 중 일부만이 DOE(632a)와의 각각의 교차에 의해 뷰어의 눈을 향해 편향되는 반면에, 광의 나머지는 TIR을 통해 주 도파관(632b)을 통해 계속 전파된다.

[0057] [0071] 전파되는 광과 DOE(632a) 사이의 각각의 교차 지점에서, 광의 프랙션이 주 도파관(632b)의 인접한 면을 향해 회절되어, 광이 TIR을 벗어나 주 도파관(632b)의 면으로부터 나오게 한다. 일부 실시예들에서, DOE(632a)의 방사상 대칭 회절 패턴은 부가적으로 회절된 광에 포커스 레벨을 부여하여, 개별 빔의 광 파면을 성형(예컨대, 곡률을 부여함)하는 것은 물론, 설계된 포커스 레벨과 매칭하는 각도로 빔을 조종한다.

[0058] [0072] 따라서, 이들 상이한 경로들은 상이한 각도들의, 포커스 레벨에서 다수의 DOE들(632a)에 의해, 또는 출사동에서 상이한 필 패턴(fill pattern)들의 산출에 의해 광이 주 평면 도파관(632b) 밖으로 커플링되게 할 수 있다. 출사동에서의 상이한 필 패턴들은 다수의 깊이 평면들을 갖는 광 필드 디스플레이를 생성하는데 유리하게 사용될 수 있다. 도파관 어셈블리의 각각의 층 또는 스택의 층들의 세트(예컨대, 3개의 층들)는 각자의 컬러(예컨대, 적색, 청색, 녹색)를 생성하는 데 사용될 수 있다. 따라서, 예컨대, 3개의 인접한 층들의 제 1 세트는 제 1 초점 깊이의 적색, 청색 및 녹색 광을 각각 생성하도록 사용될 수 있다. 3개의 인접한 층들의 제 2 세트는 제 2 초점 깊이의 적색, 청색 및 녹색 광을 각각 생성하도록 사용될 수 있다. 다수의 세트들이 다양한 초점 깊이들을 갖는 풀(full) 3D 또는 4D 컬러 이미지 광 필드를 생성하도록 사용될 수 있다.

웨어러블 시스템의 다른 컴포넌트

[0060] [0073] 다수의 구현들에서, 웨어러블 시스템은 위에서 설명된 웨어러블 시스템의 컴포넌트들에 부가적으로 또는 대안적으로 다른 컴포넌트들을 포함할 수 있다. 웨어러블 시스템은 예컨대, 하나 이상의 햅틱 디바이스들 또는 컴포넌트들을 포함할 수 있다. 햅틱 디바이스들 또는 컴포넌트들은 사용자에게 촉감을 제공하도록 동작 가능할 수 있다. 예컨대, 햅틱 디바이스들 또는 컴포넌트들은 가상 콘텐츠(예컨대, 가상 객체들, 가상 툴들, 다른 가상 구조들)를 터치할 때 압력 또는 텍스처의 촉감을 제공할 수 있다. 촉감은 가상 객체가 표현하는 물리적 객체의 느낌(feel)을 복제할 수 있거나, 또는 가상 객체가 표현하는 이미지화된 객체 또는 캐릭터(예컨대, 용)의 느낌을 복제할 수 있다. 일부 구현들에서, 햅틱 디바이스들 또는 컴포넌트들은 사용자에 의해 작용될 수 있다(예컨대, 사용자 웨어러블 글러브). 일부 구현들에서, 햅틱 디바이스들 또는 컴포넌트들은 사용자에 의해 보유될 수 있다.

[0061] [0074] 웨어러블 시스템은 예컨대, 웨어러블 시스템과의 상호작용 또는 입력을 허용하기 위해 사용자에 의해

조작 가능한 하나 이상의 물리적 객체들을 포함할 수 있다. 이러한 물리적 객체들은 본원에서 토템(totem)들로서 지칭될 수 있다. 일부 토템들은 예컨대, 금속 또는 플라스틱 조작, 벽, 테이블의 표면과 같은 무생물의 형태를 취할 수 있다. 특정 구현들에서, 토템들은 실제로, 어떠한 물리적 입력 구조들(예컨대, 키들, 트리거들, 조이스틱, 트랙볼, 로커 스위치(rocker switch))도 갖지 않을 수 있다. 대신, 토템은 단순히 물리적 표면을 제공할 수 있고 웨어러블 시스템은 사용자 인터페이스가 토템의 하나 이상의 표면들 상에 있는 것으로 사용자에게 나타나게 할 수 있다. 예컨대, 웨어러블 시스템은 컴퓨터 키보드 및 트랙패드의 이미지가 토템의 하나 이상의 표면들에 상주하는 것으로 나타나게 할 수 있다. 예컨대, 웨어러블 시스템은 가상 컴퓨터 키보드 및 가상 트랙패드가 토템으로서 역할을 하는 얇은 직사각형 알루미늄 플레이트의 표면 상에 나타나게 할 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 토템을 단독으로 또는 포즈들과 함께 사용하여 웨어러블 시스템 또는 다른 사용자들과 상호작용할 수 있다.

[0062] [0075] 본 개시내용의 웨어러블 디바이스들, HMD 및 디스플레이 시스템들과 함께 사용 가능한 햅틱 디바이스들 및 토템들의 예들은, 그 전체가 본원에 인용에 의해 포함되는 미국 특허 공보 제 2015/0016777호에서 설명된다.

웨어러블 시스템들, 환경들 및 인터페이스들의 예들

[0064] [0076] 웨어러블 시스템은 렌더링된 광 필드들에서 높은 피사계 심도를 달성하기 위해 다양한 맵핑 관련 기술들을 사용할 수 있다. 가상 세계의 맵핑 시에, 실제 세계와 관련하여 가상 객체들을 정확하게 묘사하기 위해 현실 세계의 모든 피쳐들 및 지점들을 아는 것이 유리하다. 이를 위해, 웨어러블 시스템의 사용자들로부터 캡처된 FOV 이미지들이 실제 세계의 다양한 지점들 및 피쳐들에 관한 정보를 전달하는 새로운 사진들을 포함함으로써 세계 모델에 부가될 수 있다. 예컨대, 웨어러블 시스템은 맵 지점의 세트(이를테면, 2D 지점들 또는 3D 지점들)를 수집하고 새로운 맵 지점들을 발견하여 세계 모델의 보다 정확한 버전을 렌더링할 수 있다. 제 1 사용자의 세계 모델이 (예컨대, 클라우드 네트워크와 같은 네트워크를 통해) 제 2 사용자에게 통신될 수 있어서, 제 2 사용자는 제 1 사용자 주위의 세계를 경험할 수 있다.

[0065] [0077] 도 7은 MR 환경(700)의 예의 블록도이다. MR 환경(700)은 하나 이상의 사용자 웨어러블 시스템들(예컨대, 웨어러블 시스템(200) 또는 디스플레이 시스템(220)) 또는 정적 룸(stationary room) 시스템들(예컨대, 룸 카메라들 등)로부터 입력(예컨대, 사용자의 웨어러블 시스템으로부터의 시각적 입력(702), 이를테면, 룸 카메라들로부터의 정적 입력(704), 다양한 센서들로부터의 감각 입력(706), 제스처들, 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 정적 카메라들로부터의 정보 또는 상이한 관점으로부터의 다양한 단서들로 보완될 수 있다. 카메라들(이를테면, 룸 카메라들 및/또는 외향 이미징 시스템의 카메라들)에 의해 획득된 이미지 데이터는 맵핑 지점들의 세트로 감소될 수 있다.

[0066] [0078] 하나 이상의 객체 인식기들(708)은 맵 데이터베이스(710)의 도움으로 수신된 데이터(예컨대, 지점들의 모음)를 크롤링(crawl)하고, 지점을 인식 또는 맵핑하고, 이미지를 태깅하고, 의미 정보를 객체들에 첨부할 수 있다. 맵 데이터베이스(710)는 시간에 걸쳐 수집된 다양한 지점들 및 그들의 대응하는 객체들을 포함할 수 있다. 다양한 디바이스들 및 맵 데이터베이스는 클라우드에 액세스하기 위해 네트워크(예컨대, LAN, WAN 등)를 통해 서로 연결될 수 있다.

[0067] [0079] 맵 데이터베이스의 이러한 정보 및 지점들의 모음에 기초하여, 객체 인식기들(708a 내지 708n)은 환경의 객체들을 인식할 수 있다. 예컨대, 객체 인식기들은 얼굴들, 사람들, 창문들, 벽들, 사용자 입력 디바이스들, 텔레비전들, 문서들(예컨대, 본원에서의 보안 예들에서 설명된 바와 같은 여행 티켓들, 운전 면허증, 여권), 사용자의 환경의 다른 객체들 등을 인식할 수 있다. 하나 이상의 객체 인식기들이 소정의 특성들을 갖는 객체에 대해 전문화될 수 있다. 예컨대, 객체 인식기(708a)는 얼굴들을 인식하는 데 사용될 수 있는 반면에, 다른 객체 인식기는 문서들을 인식하는 데 사용될 수 있다.

[0068] [0080] 객체 인식들은 다양한 컴퓨터 비전 기술들을 사용하여 수행될 수 있다. 예컨대, 웨어러블 시스템은 장면 재구성, 이벤트 검출, 비디오 추적, 객체 인식(예컨대, 사람들 또는 문서들), 객체 포즈 추정, (예컨대, 환

경내의 사람 또는 문서 상의 이미지로부터) 얼굴 인식, 러닝(learning), 인덱싱, 모션 추정 또는 이미지 분석(예컨대, 사진들, 서명들, 식별 정보, 여행 정보 등과 같은 문서들 내의 표시들을 식별함) 등을 수행하도록 외향 이미징 시스템(464)(도 4에 도시됨)에 의해 획득된 이미지들을 분석할 수 있다. 이러한 작업들을 수행하기 위해 하나 이상의 컴퓨터 비전 알고리즘들이 사용될 수 있다. 컴퓨터 비전 알고리즘들의 비-제한적인 예들은, SIFT(Scale-invariant feature transform), SURF(speeded up robust features), ORB(oriented FAST and rotated BRIEF), BRISK(binary robust invariant scalable keypoints), FREAK(fast retina keypoint), Viola-Jones 알고리즘, Eigenfaces 접근법, Lucas-Kanade 알고리즘, Horn-Schunck 알고리즘, Mean-shift 알고리즘, vSLAM(visual simultaneous location and mapping) 기술들, 순차적 베이지안 추정기(예컨대, 칼만 필터, 확장된 칼만 필터 등), 번들 조정, 적응형 임계화(Adaptive thresholding)(및 다른 임계화 기술들), ICP(Iterative Closest Point), SGM(Semi Global Matching), SGBM(Semi Global Block Matching), 피쳐 포인트 히스토그램(Feature Point Histogram)들, (예컨대, 지원 벡터 머신, k-최근접 이웃 알고리즘, 나이브 베이즈(Naive Bayes), 뉴럴 네트워크(콘볼루셔널 또는 딥 뉴럴 네트워크들을 포함함) 또는 다른 감독/비-감독 모델들 등과 같은) 다양한 머신 러닝 알고리즘들 등을 포함한다.

[0069] [0081] 객체 인식들은 부가적으로 또는 대안적으로, 다양한 머신 러닝 알고리즘들에 의해 수행될 수 있다. 일단 트레이닝되면, 머신 러닝 알고리즘은 HMD에 의해 저장될 수 있다. 머신 러닝 알고리즘들의 일부 예들은, (예컨대, 정규 최소 제곱화 회귀 분석(Ordinary Least Squares Regression)과 같은) 회귀 알고리즘들, (예컨대, 러닝 벡터 양자화(Learning Vector Quantization)와 같은) 인스턴스-기반 알고리즘들, (예컨대, 분류 및 회귀 트리들과 같은) 결정 트리 알고리즘들, (예컨대, 나이브 베이즈(Naive Bayes)와 같은) 베이지안 알고리즘들, (예컨대, k-평균 클러스터링과 같은) 클러스터링 알고리즘들, (예컨대, 선형 알고리즘들과 같은) 연관 규칙 러닝 알고리즘들, (예컨대, 퍼셉트론(Perceptron)과 같은) 인공 뉴럴 네트워크 알고리즘들, (예컨대, 딥 볼츠만 머신, 또는 딥 뉴럴 네트워크와 같은) 딥 러닝 알고리즘들, (예컨대, 주 성분 분석(Principal Component Analysis)과 같은) 차원성 감소 알고리즘, (예컨대, 누적된 일반화(Stacked Generalization)와 같은) 양상블 알고리즘들 및/또는 다른 머신 러닝 알고리즘들을 포함하는 감독 또는 비-감독 머신 러닝 알고리즘들을 포함할 수 있다. 일부 실시예들에서, 개별 모델들은 개별 데이터 세트들에 대해 커스터마이징될 수 있다. 예컨대, 웨어러블 디바이스는 기본 모델을 생성하거나 저장할 수 있다. 기본 모델은 데이터 유형(예컨대, 텔레프레전스(telepresence) 세션의 특정 사용자), 데이터 세트(예컨대, 텔레프레전스 세션의 사용자의 획득된 부가적인 이미지들의 세트), 조건부 상황들 또는 다른 변동들에 특정한 부가적인 모델들을 생성하기 위한 시작 지점으로서 사용될 수 있다. 일부 실시예들에서, 웨어러블 HMD는 어그리게이팅된 데이터의 분석을 위한 모델들을 생성하기 위해 복수의 기술들을 활용하도록 구성될 수 있다. 다른 기술들은 미리 정의된 임계치들 또는 데이터 값들을 사용하는 것을 포함할 수 있다.

[0070] [0082] 맵 데이터베이스의 이러한 정보 및 지점들의 모음에 기초하여, 객체 인식기들(708a 내지 708n)은 객체들을 인식하고 객체들에 생명을 부여하기 위해 의미 정보를 객체들에 보충할 수 있다. 예컨대, 객체 인식기가 지점들의 세트를 문이라고 인식하는 경우, 시스템은 일부 의미 정보를 첨부할 수 있다(예컨대, 문은 힌지(hinge)를 가지며 힌지를 중심으로 90도 움직임을 가짐). 객체 인식기가 지점들의 세트를 미러라고 인식하는 경우, 시스템은 미러는 룸 내의 객체들의 이미지들을 반사할 수 있는 반사 표면을 갖는다는 의미 정보를 첨부할 수 있다. 의미 정보는 본원에서 설명된 바와 같이 객체들의 어포던스들(affordance)을 포함할 수 있다. 예컨대, 의미 정보는 객체의 범선을 포함할 수 있다. 시스템은 방향이 객체의 범선을 표시하는 벡터를 할당할 수 있다. 시간이 지남에 따라, 시스템(로컬로 상주할 수도 있거나, 또는 무선 네트워크를 통해 액세스 가능할 수도 있음)이 세계로부터 더 많은 데이터를 누적함에 따라 맵 데이터베이스가 성장한다. 객체들이 인식되면, 정보는 하나 이상의 웨어러블 시스템들에 전송될 수 있다. 예컨대, MR 환경(700)은 캘리포니아에서 일어나는 장면에 관한 정보를 포함할 수 있다. 환경(700)은 뉴욕의 하나 이상의 사용자들에 전송될 수 있다. FOV 카메라로부터 수신된 데이터 및 다른 입력들에 기초하여, 객체 인식기들 및 다른 소프트웨어 컴포넌트들은 다양한 이미지들로부터 수집된 지점을 맵핑하고, 객체들을 인식하는 등을 수행할 수 있어서, 장면이 세계의 상이한 부분에 있을 수 있는 제 2 사용자에게 정확히 "전달될" 수 있다. 환경(700)은 또한 로컬화(localization) 목적들을 위해 토폴로지 맵(topological map)을 사용할 수 있다.

[0071] [0083] 도 8은 인지된 객체들과 관련하여 가상 콘텐츠를 렌더링하는 방법(800)의 예의 프로세스 흐름도이다. 방법(800)은 가장 장면이 웨어러블 시스템의 사용자에게 어떻게 제시될 수 있는지를 설명한다. 사용자는 장면으로부터 지리적으로 멀리 있을 수 있다. 예컨대, 사용자는 뉴욕에 있을 수 있지만, 현재 캘리포니아에서 진행되고 있는 장면을 보기를 원할 수 있거나, 또는 캘리포니아에 거주하는 친구와 함께 산책을 하려 가기를 원할

수 있다.

[0072] [0084] 블록(810)에서, 웨어러블 시스템은 사용자 및 다른 사용자들로부터, 사용자의 환경에 관한 입력을 수신 할 수 있다. 이는 다양한 입력 디바이스들 및 맵 데이터베이스에 이미 보유된 지식을 통해 달성될 수 있다. 사용자의 FOV 카메라, 센서들, GPS, 눈 추적 등은 블록(810)에서 시스템에 정보를 전달한다. 시스템은 블록(820)에서 이 정보에 기초하여 회소 지점들을 결정할 수 있다. 회소 지점들은 사용자의 주변들의 다양한 객체들의 배향 및 포지션을 디스플레이 및 이해하는 데 사용될 수 있는 포즈 데이터(예컨대, 머리 포즈, 눈 포즈, 신체 포즈 또는 손 제스처들)를 결정하는 데 사용될 수 있다. 객체 인식기들(708a-708n)은 블록(830)에서 이러한 수집된 지점들을 크롤링하고 맵 데이터베이스를 사용하여 하나 이상의 객체들을 인식할 수 있다. 그 후, 이 정보는 블록(840)에서, 사용자의 개별 웨어러블 시스템에 전달될 수 있고, 따라서, 원하는 가상 장면이 블록(850)에서, 사용자에게 디스플레이될 수 있다. 예컨대, 원하는 가상 장면(예컨대, CA의 사용자)은 뉴욕의 사용자의 다양한 객체들 및 다른 주변들과 관련하여 적절한 배향, 포지션 등에서 디스플레이될 수 있다.

[0073] [0085] 도 9는 웨어러블 시스템의 다른 예의 블록도이다. 이 예에서, 웨어러블 시스템(900)은 세계의 맵 데이터를 포함하는 맵 데이터베이스(710)를 포함할 수 있는 맵(920)을 포함한다. 맵은 부분적으로 웨어러블 시스템 상에 로컬로 상주할 수 있고, (예컨대, 클라우드 시스템에서) 유선 또는 무선 네트워크에 의해 액세스 가능한 네트워킹된 저장 위치들에 부분적으로 상주할 수 있다. 포즈 프로세스(910)는 웨어러블 컴퓨팅 아키텍처(예컨대, 프로세싱 모듈(260) 또는 제어기(460)) 상에서 실행되고 웨어러블 컴퓨팅 하드웨어 또는 사용자의 포지션 및 배향을 결정하기 위해 맵(920)으로부터 데이터를 활용할 수 있다. 포즈 데이터는 사용자가 시스템을 경험하고 세계에서 동작할 때 즉석에서 수집된 데이터로부터 컴퓨팅될 수 있다. 데이터는 이미지들, 센서들(이를테면, 일반적으로 가속도계 및 자이로스코프 컴퓨트론트들을 포함하는 관성 측정 유닛들)로부터의 데이터 및 실제 또는 가상 환경의 객체들과 관련된 표면 정보를 포함할 수 있다.

[0074] [0086] 회소 지점 표현은 동시성 로컬화 및 맵핑(예컨대, 입력이 이미지들/시각 전용인 구성을 지칭하는 SLAM 또는 vSLAM) 프로세스의 출력일 수 있다. 시스템은 다양한 컴퓨트론트들이 세계 어디에 있는지 뿐만 아니라, 세상이 무엇으로 이루어져 있는지를 파악하도록 구성될 수 있다. 포즈는 맵을 채우고 맵으로부터의 데이터를 사용하는 것을 포함해서, 다수의 목표들을 달성하는 빌딩 블록일 수 있다.

[0075] [0087] 일 실시예에서, 회소 지점 포지션은 그 자체로 완전히 충분하지 않을 수 있고, 다중 초점 AR, VR 또는 MR 경험을 생성하기 위해 추가의 정보가 필요할 수 있다. 일반적으로 깊이 맵 정보를 지칭하는 밀집된 표현들이 이 챕터를 적어도 부분적으로 채우기 위해 활용될 수 있다. 이러한 정보는 스테레오(940)로서 지칭되는 프로세스로부터 컴퓨팅될 수 있으며, 여기서 깊이 정보가 삼각측량 또는 비행-시간 감지와 같은 기술을 사용하여 결정된다. 이미지 정보 및 활성 패턴들(이를테면, 활성 프로젝터들을 사용하여 생성된 적외선 패턴들), 이미지 카메라들로부터 획득된 이미지들, 또는 손 제스처들/토템(950)은 스테레오 프로세스(940)에 대한 입력으로서 역할을 할 수 있다. 상당한 양의 깊이 맵 정보가 함께 융합될 수 있으며, 그 중 일부는 표면 표현으로 요약될 수 있다. 예컨대, 수학적으로 정의 가능한 표면들은 (예컨대, 대형 지점 클라우드에 비해) 효율적일 수 있고, 게임 엔진들과 같은 다른 프로세싱 디바이스들에 의해 가능한 입력들일 수 있다. 따라서, 스테레오 프로세스(940)의 출력(예컨대, 깊이 맵)은 융합 프로세스(930)에서 결합될 수 있다. 포즈(910)는 또한 이 융합 프로세스(930)에 대한 입력일 수 있고, 융합(930)의 출력은 맵 프로세스(920)를 채우기 위한 입력이 된다. 서브-표면들이 이를테면, 토플리지 맵핑에서 서로 연결되어 더 큰 표면들을 형성할 수 있고 맵은 지점들 및 표면들의 대형 혼합물이 된다.

[0076] [0088] 혼합 현실 프로세스(960)에서의 다양한 양상들을 해결하기 위해, 다양한 입력들이 활용될 수 있다. 예컨대, 도 9에 도시된 실시예에서, 게임 파라미터들은, 시스템의 사용자가 다양한 위치들의 하나 이상의 몬스터들, 다양한 조건들(이를테면, 사용자가 몬스터를 쏘는 경우) 하에서 죽거나 도망가는 몬스터들, 다양한 위치들의 벽들 또는 다른 객체들 등을 갖는 몬스터 전투 게임을 플레이하고 있다고 결정하기 위한 입력일 수 있다. 세계 맵은 객체들의 위치에 관한 정보 또는 객체들의 의미 정보를 포함할 수 있고 세계 맵은 혼합 현실에 대한 다른 귀중한 입력일 수 있다. 세계에 대한 포즈가 또한 입력이 되며 거의 모든 상호작용 시스템에 대해 중요한 역할을 한다.

[0077] [0089] 사용자로부터의 제어들 또는 입력들은 웨어러블 시스템(900)에 대한 다른 입력이다. 본원에서 설명된 바와 같이, 사용자 입력들은 시각적 입력, 제스처들, 토템들, 오디오 입력, 감각 입력 등을 포함할 수 있다. 예컨대, 주위를 돌아다니거나 게임을 플레이하기 위해, 사용자는 웨어러블 시스템(900)에, 자신이 하기를 원하는 것에 관해 지시할 필요가 있을 수 있다. 단지 공간에서 자신을 움직이는 것 외에도, 활용될 수 있는 다양한

형태들의 사용자 제어들이 존재한다. 일 실시예에서, 토템(예컨대, 사용자 입력 디바이스) 또는 객체, 이를테면, 장난감 총은 사용자에 의해 보유되고 시스템에 의해 추적될 수 있다. 시스템은 바람직하게는, 사용자가 아이템을 보유하고 있다는 것을 인지하고 사용자가 아이템과 어떤 종류의 상호작용을 하고 있는지를 이해하도록 구성될 것이다(예컨대, 토템 또는 객체가 총인 경우, 시스템은 위치 및 배향은 물론, IMU와 같은 센서가 장착될 수 있는 트리거 또는 다른 감지 버튼 또는 엘리먼트를 사용자가 클릭하고 있는지 여부를 이해하도록 구성될 수 있으며, 이러한 센서는 그러한 활동이, 카메라들 중 임의의 것의 시야 내에 있지 않을 때조차도 무슨 일이 일어나고 있는지를 결정하는데 도움을 줄 수 있음).

[0078]

[0090] 손 제스처 추적 또는 인식은 또한 입력 정보를 제공할 수 있다. 웨어러블 시스템(900)은 버튼 누름들에 대한 손 제스쳐들, 왼쪽 또는 오른쪽, 정지, 잡기, 홀드 등을 제스처링하는 손 제스쳐들을 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(leveraging)하도록 구성될 수 있다. 예컨대, 제스쳐들은 중지를 위한 펼쳐진 손, ok를 위한 엄지 올리기, not ok를 위한 엄지 내리기; 또는 방향성 커맨드들을 위한 우측 또는 좌측 또는 위/아래로의 손 뒤집기와 같은 단순한 정적 제스처일 수도 있다.

[0079]

[0091] 눈 추적(예컨대, 특정 깊이 또는 범위에서 렌더링하도록 디스플레이 기술을 제어하기 위해 사용자가 바라보는 곳을 추적함)은 다른 입력이다. 일 실시예에서, 눈들의 이접운동은 삼각측량을 사용하여 결정될 수 있고, 그 후, 그 특정 사람에 대해 발현되는 이접운동/원근조절 모델을 사용하여, 원근조절이 결정될 수 있다. 눈 시선(예컨대, 한쪽 또는 양쪽 눈들의 방향 또는 배향)을 결정하기 위해 눈 카메라(들)에 의해 눈 추적이 수행될 수 있다. 예컨대, 눈(들) 근처에 배치된 전극들에 의한 전위들의 측정(예컨대, 전기 안구도 기록(electrooculography))과 같은 눈 추적을 위한 다른 기술들이 사용될 수 있다.

[0080]

[0092] 스피치 추적은 단독으로 또는 다른 입력들(예컨대, 토템 추적, 눈 추적, 제스처 추적 등)과 함께 사용될 수 있는 다른 입력일 수 있다. 스피치 추적은 스피치 인식, 음성 인식을 단독으로 또는 조합하여 포함할 수 있다. 시스템(900)은 환경으로부터 오디오 스트림을 수신하는 오디오 센서(예컨대, 마이크로폰)를 포함할 수 있다. 시스템(900)은 누가 말하고 있는지(예컨대, 스피치가 ARD의 착용자 또는 다른 사람 또는 음성(예컨대, 환경 내의 라우드스피커에 의해 송신된 레코딩된 음성)로부터 왔는지)를 결정하기 위한 음성 인식 기술은 물론, 무엇을 말하고 있는지를 결정하기 위한 스피치 인식 기술을 포함할 수 있다. 로컬 데이터 및 프로세싱 모듈(260) 또는 원격 프로세싱 모듈(270)은 예컨대, 허든 마르코프 모델(hidden Markov model)들, DTW(dynamic time warping)-기반 스피치 인식들, 뉴럴 네트워크들, 딥 러닝 알고리즘들, 이를테면, 딥 피드포워드 및 회귀 뉴럴 네트워크(recurrent neural network)들, 단-대-단 자동 스피치 인식들, (도 7을 참조하여 설명된) 머신 러닝 알고리즘들, 또는 음향 모델링 또는 언어 모델링을 사용하는 다른 알고리즘들 등과 같은 다양한 스피치 인식 알고리즘들을 적용함으로써 스피치의 콘텐츠를 식별하기 위해 마이크로폰으로부터의 오디오 데이터(또는, 예컨대, 사용자에 의해 시청되는 비디오 스트림과 같은 다른 스트림의 오디오 데이터)를 프로세싱할 수 있다. 일부 경우들에서, 스피치는 다수의 소스들로부터, 예컨대, 사용자 부근의 다른 사람으로부터, 사람 부근에서 플레이 중인 텔레비전 상의 아나운서로부터, 그리고 스피커(240)를 통해 ARD의 사용자에게 플레이되고 있는 스피치 콘텐츠로부터 나올 것이다. 아래에서 추가로 설명되는 바와 같이, 이들 상이한 스피치 소스들(예컨대, 사람, 텔레비전 아나운서 및 이 예의 오디오 스트림)은 콘텐츠 분석될 수 있고, 상이한 토픽들은 ARD의 사용자 인터페이스에 의해 사용자에게 상이하게 제시될 수 있다(예컨대, 상이한 스크린들로 구성된 상이한 토픽들, 상이한 스크린들로 구성된 상이한 스피커들에 의한 스피치 또는 이들의 조합).

[0081]

[0093] 로컬 데이터 및 프로세싱 모듈(260) 또는 원격 프로세싱 모듈(270)은 또한, 스피커의 아이덴티티, 이를테면, 스피커가 웨어러블 시스템(900)의 사용자(210)인지 또는 사용자가 대화중인 다른 사람인지를 식별할 수 있는 음성 인식 알고리즘들을 적용할 수 있다. 일부 예시적인 음성 인식 알고리즘들은 주파수 추정, 숨겨진 마르코프 모델들, 가우스 혼합 모델, 패턴 매칭 알고리즘들, 뉴럴 네트워크들, 매트릭스 표현, 벡터 양자화, 스피커 분리(speaker diarisation), 결정 트리들 및 DTW(dynamic time warping) 기술을 포함할 수 있다. 음성 인식 기술들은 또한, 코호트 모델들 및 세계 모델들과 같은 반-스피커 기술들을 포함할 수 있다. 스펙트럼 피쳐들은 스피커 특성들을 나타내는 데 사용될 수 있다. 로컬 데이터 및 프로세싱 모듈 또는 원격 데이터 프로세싱 모듈(270)은 음성 인식을 수행하기 위해 도 7을 참조하여 설명된 다양한 머신 러닝 알고리즘들을 사용할 수 있다.

[0082]

[0094] 단어 흐름 주석 시스템(970)의 구현은 UI(user interface)를 통해 이들 사용자 제어들 또는 입력들을

사용할 수 있다. UI 엘리먼트들(예컨대, 제어들, 팝업 창들, 버블들, 데이터 입력 필드들 등)은 예컨대, 보조 정보의 디스플레이를 해제하거나 공통 단어 사전에 단어를 부가하는 데 사용할 수 있다. 이러한 구현들의 예들 및 이를 용도들은 아래에서 추가로 설명된다.

[0083] [0095] 카메라 시스템들과 관련하여, 도 9에 도시된 예시적인 웨어러블 시스템(900)은 3쌍의 카메라들, 즉 사용자의 얼굴의 측면들에 대해 배열되는 비교적 넓은 FOV 또는 수동 SLAM 쌍의 카메라들, 스테레오 이미징 프로세스(940)를 처리하기 위해 그리고 또한, 손 제스처들을 캡처하고 사용자의 얼굴 앞의 토템/객체 추적을 위해 사용자 앞에 배향된 상이한 쌍의 카메라들을 포함할 수 있다. 스테레오 프로세스(940)를 위한 FOV 카메라들 및 카메라들의 쌍은 외향 이미징 시스템(464)(도 4에 도시됨)의 부분일 수 있다. 웨어러블 시스템(900)은 눈 백터들 및 다른 정보를 삼각 측량하기 위해 사용자의 눈들을 향해 배향되는 눈 추적 카메라들(도 4에 도시된 내향 이미징 시스템(462)의 부분일 수 있음)을 포함할 수 있다. 웨어러블 시스템(900)은 또한 장면에 텍스처를 주입하기 위해 하나 이상의 텍스처링된 광 프로젝터들(이를테면, 적외선(IR) 프로젝터들)을 포함할 수 있다.

[0084] [0096] 도 10은 웨어러블 시스템에 대한 사용자 입력을 결정하는 방법(1000)의 예의 프로세스 흐름도이다. 이 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 인식되거나 센서들(예컨대, 햅틱 글러브, 이미지 센서들, 손 추적 디바이스, 눈-추적 카메라들, 머리 포즈 센서들 등)을 통해 검출될 수 있다.

[0085] [0097] 검출된 제스처, 눈 포즈, 머리 포즈 또는 토템을 통한 입력에 적어도 부분적으로 기초하여, 웨어러블 시스템은 블록(1020)에서, 기준 프레임에 대한 토템(또는 사용자의 눈들 또는 머리 또는 제스처들)의 포지션, 배향 또는 움직임을 검출한다. 기준 프레임은 맵 지점들의 세트일 수 있으며, 이에 기초하여, 웨어러블 시스템이 토템(또는 사용자)의 움직임을 액션 또는 커맨드로 전환한다. 블록(1030)에서, 토템과 사용자의 상호작용이 맵핑된다. 기준 프레임(1020)에 대한 사용자 상호작용의 맵핑에 기초하여, 시스템은 블록(1040)에서 사용자 입력을 결정한다.

[0086] [0098] 예컨대, 사용자는 가상 페이지를 넘기고(turning) 다음 페이지로 이동하거나 하나의 UI(User Interface) 디스플레이 스크린으로부터 다른 UI 스크린으로 이동하는 것을 나타내도록 토템 또는 물리적 객체를 앞뒤로(back and forth) 이동시킬 수 있다. 다른 예로서, 사용자는 사용자의 FOR에서 상이한 실제 또는 가상 객체들을 보기 위해 자신의 머리 또는 눈을 이동시킬 수 있다. 특정 실제 또는 가상 객체에 대한 사용자의 시선이 임계 시간보다 긴 경우, 실제 또는 가상 객체가 사용자 입력으로서 선택될 수 있다. 일부 구현들에서, 사용자의 눈들의 이접운동이 추적될 수 있고, 사용자가 포커싱하는 깊이 평면에 관한 정보를 제공하는 사용자의 눈들의 원근조절 상태를 결정하기 위해 원근조절/이접운동 모델이 사용될 수 있다. 일부 구현들에서, 웨어러블 시스템은 레이 캐스팅(ray casting) 기술들을 사용하여 어떤 실제 또는 가상 객체들이 사용자의 머리 포즈 또는 눈의 포즈의 방향을 따르는지를 결정할 수 있다. 다양한 구현들에서, 레이 캐스팅 기술들은 실질적으로 작은 횡단 폭을 갖는 얇은 펜슬 광선들(pencil rays)을 캐스팅하거나 상당한 횡단 폭을 갖는 광선들(예컨대, 원뿔들 또는 절두체들)을 캐스팅하는 것을 포함할 수 있다.

[0087] [0099] 사용자 인터페이스는 본원에서 설명된 바와 같이 디스플레이 시스템(이를테면, 도 2의 디스플레이(220))에 의해 프로젝팅될 수 있다. 그것은 또한, 하나 이상의 프로젝터들과 같은 다양한 다른 기술들을 사용하여 디스플레이될 수 있다. 프로젝터들은 이미지들을 캔버스 또는 구체와 같은 물리적 객체 상에 프로젝팅할 수 있다. 사용자 인터페이스와의 상호작용들은 시스템의 부분 또는 시스템 외부의 하나 이상의 카메라들을 사용하여(이를테면, 내향 이미징 시스템(462) 또는 외향 이미징 시스템(464)을 사용하여) 추적될 수 있다.

[0088] [0100] 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 검출하기 위해 웨어러블 시스템에 의해 사용될 수 있다.

[0089] [0101] 블록(1110)에서, 웨어러블 시스템은 특정 UI를 식별할 수 있다. UI의 유형은 사용자에 의해 미리 결정될 수 있다. 웨어러블 시스템은 사용자 입력(예컨대, 제스처, 시각 데이터, 오디오 데이터, 감각 데이터, 직접 커맨드 등)에 기초하여 특정 UI가 패플레이트(populate)될 필요가 있다고 식별할 수 있다. UI는 시스템의 착용자가(예컨대, 여행 검문소에서) 문서들을 착용자에게 제시하는 사용자들을 관찰하는 보안 시나리오에 특정될 수 있다. 블록(1120)에서, 웨어러블 시스템은 가상 UI에 대한 데이터를 생성할 수 있다. 예컨대, 경계(confine)들, 일반적인 구조, UI의 형상 등과 연관된 데이터가 생성될 수 있다. 또한, 웨어러블 시스템은 사용자의 물리

적 위치의 맵 좌표들을 결정할 수 있어서 웨어러블 시스템이 사용자의 물리적 위치와 관련하여 UI를 디스플레이 할 수 있다. 예컨대, UI가 신체 중심인 경우, 웨어러블 시스템은, 링 UI가 사용자 주위에 디스플레이될 수 있거나 평면 UI가 벽 상에 또는 사용자 앞에 디스플레이될 수 있도록 사용자의 신체 스텐스, 머리 포즈 또는 눈 포즈의 좌표들을 결정할 수 있다. 본원에서 설명된 보안 맥락에서, UI가 시스템의 착용자에게 문서들을 제시하는 여행자를 둘러싼 것처럼 UI가 디스플레이될 수 있어서, 착용자는 여행자 및 여행자의 문서들을 보면서 UI를 쉽게 볼 수 있다. UI가 손 중심인 경우, 사용자의 손들의 맵 좌표들이 결정될 수 있다. 이러한 맵 지점들은 FOV 카메라들, 감각 입력 또는 임의의 다른 유형의 수집된 데이터를 통해 수신된 데이터를 통해 유도될 수 있다.

[0090]

[0102] 블록(1130)에서, 웨어러블 시스템은 클라우드로부터 디스플레이에 데이터를 전송할 수 있거나 데이터는 로컬 데이터베이스로부터 디스플레이 컴포넌트들로 전송될 수 있다. 블록(1140)에서, UI는 전송된 데이터에 기초하여 사용자에게 디스플레이된다. 예컨대, 광 필드 디스플레이에는 가상 UI를 사용자의 눈들 중 하나 또는 둘 모두에 프로젝팅할 수 있다. 일단 가상 UI가 생성되면, 웨어러블 시스템은 블록(1150)에서, 가상 UI 상에 보다 많은 가상 콘텐츠를 생성하라는 사용자로부터의 커맨드를 단순히 기다릴 수 있다. 예컨대, UI는 사용자의 신체 또는 사용자의 환경 내의 사람(예컨대, 여행자)의 신체 주위의 신체 중심 링일 수 있다. 그 후, 웨어러블 시스템은 커맨드(제스처, 머리 또는 눈 움직임, 음성 커맨드, 사용자 입력 디바이스로부터의 입력 등)를 기다릴 수 있고, 커맨드가 인식되는 경우(블록 1160), 커맨드와 연관된 가상 콘텐츠가 사용자에게 디스플레이될 수 있다 (블록 1170).

[0091]

[0103] 웨어러블 시스템들, UI들 및 사용자 경험(UX)의 부가적인 예들은 그 전체가 인용에 의해 본원에 포함되는 미국 특허 공보 제 2015/0016777호에서 설명된다.

[0092]

다수의 웨어러블 시스템들 사이의 예시적인 통신들

[0093]

[0104] 도 12는 서로 상호작용하는 다수의 사용자 디바이스들을 도시하는 전체 시스템도를 개략적으로 예시한다. 컴퓨팅 환경(1200)은 사용자 디바이스들(1230a, 1230b, 1230c)을 포함한다. 사용자 디바이스들(1230a, 1230b, 1230c)은 네트워크(1290)를 통해 서로 통신할 수 있다. 사용자 디바이스들(1230a-1230c)은 각각, 네트워크(1290)를 통해 (네트워크 인터페이스(1271)를 또한 포함할 수 있는) 원격 컴퓨팅 시스템(1220)과 통신하기 위한 네트워크 인터페이스를 포함할 수 있다. 네트워크(1290)는 LAN, WAN, 퍼어-투-퍼어 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 위치들에 클러스터링 및 로케이팅되는 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.

[0094]

[0105] 원격 컴퓨팅 시스템(1220)은 특정 사용자의 물리적 및/또는 가상 세계들에 관한 정보를 유지할 수 있는 원격 데이터 리포지토리(1280)를 포함할 수 있다. 데이터 저장소(1280)는 공통 단어 사전, 보조 정보 소스 등과 같은 단어 흐름 주석에 유용한 정보를 포함할 수 있다. 원격 데이터 리포지토리는 도 2에 도시된 원격 데이터 리포지토리(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 2에 도시된 원격 프로세싱 모듈(270)의 실시예일 수 있다. 원격 프로세싱 모듈(1270)은 사용자 디바이스들(1230a, 1230b, 1230c) 및 원격 데이터 리포지토리(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 저장의 적어도 일부는 (도 2에 도시된 바와 같이) 로컬 프로세싱 및 데이터 모듈(260)에 의해 제공될 수 있다. 원격 컴퓨팅 시스템(1220)은 정해진 사용자가, 특정 사용자 자신의 물리적 및/또는 가상 세계들에 관한 정보를 다른 사용자와 공유하는 것을 가능하게 할 수 있다.

[0095]

[0106] 사용자 디바이스는 웨어러블 디바이스(예컨대, HMD 또는 ARD), 컴퓨터, 모바일 디바이스 또는 임의의 다른 디바이스들 단독 또는 이들의 조합일 수 있다. 예컨대, 사용자 디바이스들(1230b 및 1230c)은 AR/VR/MR 콘텐츠를 제시하도록 구성될 수 있는, 도 2에 도시된 웨어러블 시스템(200)(또는 도 4에 도시된 웨어러블 시스템(400))의 실시예일 수 있다.

[0096]

[0107] 사용자 디바이스들 중 하나 이상은 도 4에 도시된 사용자 입력 디바이스(466)와 함께 사용될 수 있다. 사용자 디바이스는 (예컨대, 4에 도시된 외향 이미징 시스템(464)을 사용하여) 사용자 및 사용자의 환경에 관한 정보를 획득할 수 있다. 사용자 디바이스 및/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(raw) 정보를 프로세싱하고, 추가의 프로세싱을 위해 프로세싱된 정보를 원격 컴퓨팅 시스템(1220)에 전송할 수 있다. 사용자 디바이스는 또한, 프로세싱을 위해 원격 컴퓨팅 시스템(1220)에 원시 정보를 전송할 수 있다. 사용자 디바이스는 원격 컴퓨팅 시스템(1220)으로부터 프로세싱된 정보를 수신하고 사용자에게 프로젝팅하기 전에 최종 프로세싱을 제공할 수 있다. 사용자 디바이스는 또한 획득된 정보를 프로세싱하고 프로세싱된 정보를 다른 사용자 디바이스들에 전달할 수 있다. 사용자 디바이스는 획득된 정보를 프로세싱하는 동안 원격 데이터 리포지토리(1280)와 통신할 수 있다. 다수의 사용자 디바이스들 및/또는 다수의 서버 컴퓨터 시스템들은 획득된 이미지들의 구성 및/또는 프로세싱에 참여할 수 있다.

[0097] [0108] 물리적 세계들에 대한 정보는 시간이 지남에 따라 발전될 수 있고, 상이한 사용자 디바이스들에 의해 수집된 정보에 기초할 수 있다. 가장 세계들의 모델들은 또한, 시간이 지남에 따라 발전될 수 있고 상이한 사용자들의 입력들에 기초할 수 있다. 이러한 정보 및 모델들은 때로는, 본원에서 세계 맵 또는 세계 모델로서 지칭될 수 있다. 도 7 및 도 9를 참조하여 설명된 바와 같이, 사용자 디바이스들에 의해 획득된 정보는 세계 맵(1210)을 구성하는 데 사용될 수 있다. 세계 맵(1210)은 도 9에서 설명된 맵(920)의 적어도 일부를 포함할 수 있다. 다양한 객체 인식기들(예컨대, 708a, 708b, 708c ... 708n)은 객체들에 의미 정보를 부착할 뿐만 아니라 객체들 및 태그 이미지들을 인식하는 데 사용될 수 있다. 이를 객체 인식기들은 또한 도 7에 설명된다.

[0098] [0109] 원격 데이터 리포지토리(1280)는 데이터를 저장하고 세계 맵(1210)의 구성을 용이하게 하는 데 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 조합으로, 월드 맵(1210)을 구성 및/또는 업데이트할 수 있다. 예컨대, 사용자 디바이스는 원격 프로세싱 모듈(1270) 및 원격 데이터 리포지토리(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 디바이스(예컨대, 1230a, 1230b, 1230c)와 통신할 수 있다. 원격 컴퓨팅 시스템(1220)은 예컨대, 사용자의 이미지를 선택적으로 자르는 것, 사용자의 배경을 수정하는 것, 사용자의 환경에 가상 객체들을 부가하는 것, 보조 정보로 사용자의 스피치에 주석을 달는 것 등과 같이, 사용자 디바이스들(1230a, 1230b, 1230c)에 의해 획득된 정보를 수정할 수 있다. 원격 컴퓨팅 시스템(1220)은 프로세싱된 정보를 동일 및/또는 상이한 사용자 디바이스들에 전송할 수 있다.

[0099] [0110] 단어 흐름 주석 시스템의 실시예들의 다양한 기능성들이 이제 설명될 것이다.

예시적인 스피치 캡처

[0101] [0111] 개시된 단어 흐름 주석 시스템의 예시적인 구현은 인-퍼슨(in-person) 통신들의 맥락에서 발생한다. 이러한 맥락에서, 검출된 스피치의 소스는 단어 흐름 주석 시스템의 물리적 부근에 있다. 시스템의 오디오 센서는 주변 오디오 정보를 검출할 수 있다. 컴퓨팅 모듈(예컨대, 단어 흐름 주석 시스템(970))은 주변 오디오 정보로부터 스피치를 검출할 수 있다. 이 모듈은 본원에서 설명된 방식들을 포함하는, (도 9를 참조하여 설명된 바와 같은) 자동 스피치 인식 알고리즘을 구현할 수 있다. 컴퓨팅 모듈은 (예컨대, 스피치를 텍스트로 변환함으로써) 사용자의 스피치를 파악하고 사용자의 스피치에서 키워드를 식별할 수 있다. 키워드는 회귀 단어 또는 맥락 키워드를 포함할 수 있다. 키워드는 단독으로 또는 조합으로, 키워드 데이터베이스에 액세스하거나 맥락 정보를 결정함으로써 식별될 수 있다. 맥락 정보는 대화 내의 사용자 또는 다른 참가자들과 연관된 정보(예컨대, 인구학적 정보, 과거 활동들/경험들, 스피치 행동들 등) 또는 환경과 연관된 정보(예컨대, 대화의 위치, 환경 내의 객체들 또는 사람들 등)를 포함할 수 있다. 키워드는 웨어러블 시스템(200)의 사용자에게 프리젠테이션을 위한 가상 콘텐츠를 결정하는 데 사용될 수 있다. 예컨대, 웨어러블 시스템은 (예컨대, 설명 텍스트들 또는 이미지들과 같은) 키워드와 연관된 보조 정보를 제시할 수 있다.

[0102] [0112] 개시된 단어 흐름 주석 시스템의 다른 예시적인 구현은 원격 통신들의 맥락에서 발생한다. 이러한 예시적인 시나리오에서, 스피커는 단어 흐름 주석 시스템(및 그의 사용자)으로부터 물리적으로 원격이다. 스피커와 연관된 오디오 정보는 다수의 채널들 중 하나를 통해 청취자에 도달할 수 있다. 예컨대, 스피커 및 청취자는 전화를 통해 또는 인터넷 오디오 또는 오디오-비디오 채팅 세션을 통해 대화중일 수 있다. 스피커 및 청취자는 도 12에 예시된 바와 같이, (예컨대, 텔레프레전스 세션에서와 같이) 네트워크를 통해 통신하는 AR 시스템들을 사용하여 대화중일 수 있다. 다른 예로서, 스피커는 라디오, 텔레비전 또는 인터넷 프로그래밍 등 상에서 제시되는 것과 같은 오디오 또는 A/V 프로그램 내의 사람일 수 있다. 스피커의 음성이 청취자의 위치의 주변

사운드에서 들릴 수 있는 경우, 단어 흐름 주석 시스템은 그것이 인-퍼슨 통신 맥락에서 행한 것과 동일한 방식으로 사운드를 캡처할 수 있다. 스피커의 음성이 청취자의 위치의 주변 사운드에서 들릴 수 없는 경우(예컨대, 청취자가 단어 흐름 주석 시스템(970)의 오디오 증폭기(232)를 통해 스피커를 듣는 경우), 시스템은 상이한 방식으로 사운드를 캡처하거나, 또는 사운드를 캡처하는 단계를 스kip할 수 있다(예컨대, 시스템은 스피치와 연관된 정보가 임베딩된 전기 신호들로부터 스피치를 직접 검출할 수 있음). 원격 통신 사용에 대한 더 많은 설명이 아래의 텔레프레전스란 명칭의 하위섹션에서 이루어진다.

[0103] 스피커 음성 거절 및 소스 로컬화

[0104] [0113] AR 시스템은 사운드를, 그의 소스 또는 그의 소스의 위치와 연관시키는 능력을 통해 부가적인 유용한 애플리케이션들을 수행할 수 있다. 여러 개의 이러한 애플리케이션들이 아래에 설명된다.

[0105] [0114] 예시적인 구현은 검출된 스피치가 사용자 자신의 스피치인지를 결정한다. 이는 복수의 오디오 센서, 예컨대, 2개 이상의 마이크로폰들의 사용을 통해, 또는 지향성 오디오 센서, 예컨대, 지향성 마이크로폰의 사용을 통해 달성될 수 있다. 이러한 오디오 센서로, 시스템은 사운드의 소스의 위치를 식별할 수 있다. 따라서, 시스템은 사용자 자신의 스피치를, 다른 사람들의 것들로부터 구별할 수 있다. 시스템은 희귀 단어 검출을 위해 사용자 자신의 스피치를 프로세싱하지 않도록 구성될 수 있다.

[0106] [0115] AR 시스템의 구현은, 예컨대, 카메라들(464 및 466) 또는 정적 카메라(예컨대, 정적 입력(704))를 통해 공급된 정보로 사용자의 FOR에 대한 소스의 포지션뿐만 아니라 사운드 소스의 위치를 식별할 수 있다. 시스템은 희귀 단어 검출을 위해 사용자의 FOV 외부의 사람으로부터의 스피치를 프로세싱하지 않도록 구성될 수 있다.

[0107] [0116] 스피치를 그의 소스와 연관시키는 구현의 능력은 또한 맥락-의존적 보조 정보를 제공하는 데 있어 유용할 수 있다. 소정의 스피커들은 특정 영역 또는 맥락에서 단어들을 발화(utter)할 가능성이 매우 높다. 예컨대, 카메라 상점 직원들이 "SD"라고 말할 때, San Diego 도시를 말하기 보다는, 보안 디지털 메모리 카드에 대해 말할 가능성이 매우 높다. 따라서, 카메라 상점 직원들로서 스피커를 식별할 수 있는 구현은 맥락상 올바른 보조 정보를 사용자에게 제공할 가능성이 매우 높을 수 있다. 스피커를 식별하는 능력은 다수의 방식들로 구현될 수 있다. 예컨대, 스피치(예컨대, 대담(dialogue) 또는 독백)에서 참가자들의 이름들은 종종 스피치 초반에 도입부에서 언급된다. 또는, 사용자는 UI 상호작용을 통해 스피커의 이름을 시스템에 입력할 수 있다. 외향 카메라가 장착된 시스템은 스피커의 이미지를 캡처하고 이미지 인식을 통해 스피커를 식별할 수 있다. 시스템은 또한, 도 9에 설명된 다양한 음성 인식 알고리즘들을 사용하여 스피커의 아이덴티티를 식별하기 위해 음성 인식을 수행할 수 있다.

[0108] 스피치로부터 텍스트로의 예시적인 변환

[0109] [0117] 키워드를 식별하기 위해, 웨어러블 시스템(200)은 (예컨대, 단어 흐름 주석 시스템(970)에 의해) 오디오 스트림을 텍스트로 변환할 수 있다. 텍스트로의 스피치의 변환은 도 9를 참조하여 설명되는 딥 러닝(딥 뉴럴 네트워크들을 활용할 수 있음) 또는 다른 스피치 인식 알고리즘들과 같은 알고리즘들을 사용하여 수행될 수 있다. 웨어러블 시스템(200)은 (예컨대, 로컬 프로세싱 모듈(260)에) 스피치-텍스트 기능성을 포함할 수 있거나, 또는 (예컨대, 원격 프로세싱 모듈(270)을 통해) 스피치-텍스트 기능성에 원격으로 액세스할 수 있다. 웨어러블 시스템(200)은 (예컨대, API(application programming interface)를 통해) 상업적 스피치-텍스트 서비스들에 대한 무선 연결들을 활용할 수 있다. 단어 흐름 주석 기술은 임의의 그러한 스피치-텍스트 알고리즘을 활용하지만, 통상적이지 않은 방식으로, 즉, 사용자로부터 발생되지 않은 오디오에 스피치-텍스트 변환을 적용할 수 있다.

[0110] [0118] 단어 흐름 주석 구현은 예컨대, 위치 프로세싱 및 데이터 모듈(260)을 사용하여 웨어러블 디바이스 상에서 또는 (예컨대, 원격 컴퓨팅 시스템(1220)을 포함하는) 원격 서버 상에서 텍스트로의 스피치의 변환을 로컬로 또는 원격으로 수행할 수 있다. 변환이 로컬로 또는 원격으로 수행되든지 간에, 키워드 검출 및 보조 정보 리트리벌과 같은 다른 프로세싱 단계들은 변환이 수행되는 곳과 독립적으로 로컬로 또는 원격으로 행해질 수 있다. 예컨대, 스피치-텍스트 변환이 원격으로 행해지고 키워드 검출이 로컬로 수행되는 경우, 예컨대, 스피치-텍스트 변환이 원격으로 행해지고 키워드 검출이 로컬로 수행되는 경우, 캡처된 오디오 스트림은 네트워크를 통해, 변환이 행해지는 원격 프로세싱 모듈(270) 또는 원격 서버로 전송될 수 있고; 변환된 텍스트 스트링들은 시스템의 로컬 컴포넌트(예컨대, 로컬 프로세싱 및 데이터 모듈(260))로 리턴된다. 다른 예로서, 스피치-텍스트 변환, 키워드 검출 및 보조 정보 리트리벌이 원격으로 행해지는 경우, 캡처된 오디오 스트림은 네트워크를 통해 원격 프로세싱 모듈(270) 또는 원격 서버에 전송될 수 있고; 리트리브된 보조 정보는 시스템의 로컬 컴포넌트에

리턴될 수 있다. 로컬/원격 프로세싱의 다른 조합들이 실행 가능하다.

[0111] [0119] 본원에서 설명된 예들에서, AR 시스템(200)은 스피치-텍스트 변환을 사용하여 키워드를 검출할 수 있지만, 일부 구현들에서, AR 시스템(200)은 오디오 데이터를 사용하여 키워드를 직접 검출할 수 있다. 예컨대, AR 시스템(200)은 스피치를 텍스트로 변환할 필요 없이 키워드의 사람/사용자의 발음에 기초하여 키워드를 검색(look up)할 수 있다.

회귀 단어의 예시적인 검출

[0113] [0120] AR 시스템은 회귀 단어들과 같은 키워드들을 변환된 텍스트에서 검색할 수 있다. 회귀 단어들의 예는 AR 시스템의 사용자에 대해 드물거나 흔치 않은 단어들을 포함할 수 있다. 회귀 단어들은 흔치 않게 사용된 단어들, 다른 언어로부터의 단어들(예컨대, 영어 문장에서 사용되는 중국어 단어 guanxi), 기술 단어들, AR 시스템의 사용자가 일반적으로 종사하는 맥락 외부에 있는 전문화된 영역들로부터의 단어들 등을 포함(그러나 이에 제한되지 않음)할 수 있다. 예컨대, AR 시스템은, 단어가 AR 시스템에 의해 검출된 스피치에서 드물게 사용되는 경우 회귀한 것으로 단어를 식별할 수 있다. 그러한 회귀 단어들에 대한 정의, 번역 또는 맥락 정보를 제공함으로써, AR 시스템은 유리하게는, 사용자가 듣는 스피치를 더 잘 이해하도록 사용자를 보조할 수 있다.

[0114] [0121] 회귀 단어들은 공통 단어들의 알려진 리스트와 비교하여 검출될 수 있다. 예컨대, 단어는 그것이 공통 단어 사전 또는 데이터베이스에서 발견되지 않는 경우 회귀한 것으로 간주될 수 있다. 공통 단어 사전 또는 데이터베이스는 (예컨대, 도 2의 데이터 모듈(71)에) 로컬로 저장되거나 (예컨대, 원격 데이터 리포지토리(74)에) 원격으로 저장될 수 있다. 공통 단어 사전 또는 데이터베이스는 제 3 자 소스들로부터 이용 가능한데, 예컨대, 인터넷 상에서 이용 가능할 수 있다.

[0115] [0122] 회귀 단어는 언어의 발화의 그의 빈도 또는 단어 빈도 리스트에서 연관된 랭크를 통해 검출될 수 있다. 지프(Zipf)의 법칙은 단어 빈도 리스트에서의 랭크에 발화의 빈도를 관련시키는 예인데, 즉 단어의 발화의 빈도는 단어 빈도 리스트에서 그의 랭크에 반비례한다. 회귀 단어는 스피치에서 임계 빈도 미만의 그의 발화에 기초하여 AR 시스템에 의해 식별될 수 있다. 단어 흐름 시스템은 사용자가, 언어에 대한 사용자의 지식에 기초하여 단어들을 검출하도록 시스템을 구성할 수 있게 할 수 있다. 예컨대, 사용자는 5,000 또는 7,000(또는 다른 값)의 가장 공통적인 단어들의 리스트에 있지 않은 단어들을 검출하도록 시스템을 구성할 수 있다. 다른 예로서, 사용자는 발생 가능성이 0.01 %, 0.003 %, 0.001 % 또는 언어에서의 다른 값과 같은 임계치 미만인 단어들을 검출하도록 시스템을 구성할 수 있다. AR 시스템은 사용자-구성 가능한 디폴트 임계치(예컨대, 0.005 %)를 가질 수 있다. 통상적으로, 임계치가 높을수록, 더 많은 회귀 단어들이 플래깅(flagged)될 것이고, 임계치가 낮을수록, 더 적은 회귀 단어들이 플래깅될 것이다. 사용자는 회귀 단어에 대해 제시된 보조 정보가 지나치게 산만하게 하지 않고 사용자에게 유용한 것으로 간주되는 값으로 임계치를 세팅할 수 있다. AR 시스템은 사용자가 (예컨대, 회귀 단어들을 플래깅하는 것을 중단하도록) 임계치를 0으로 일시적으로 세팅하게 할 수 있으며, 이는 이를테면, 친숙하지 않은 환경에서 차량을 운전중인 사용자, 사용자가 산만해지길 원치 않는 종교적 또는 문화적 이벤트에 참석중인 사용자, 또는 단어들 중 상당 프랙션(fraction)이 사용자에게 회귀 단어들일 수 있는 이벤트에 참석중인 사용자(예컨대, 스피커들이 특수 의료 단어들 및 구문들을 일상적으로 말하는 의료 프레젠테이션에서 누워있는 사용자)와 같이 소정의 상황들에서 유용할 수 있다.

[0116] [0123] 일부 구현들에서, 회귀 단어는 사용 통계와 사용 임계치의 비교에 의한 단어의 드문 사용에 의해 결정되고, 단어가 흔하게 사용되는 단어들의 사용 임계치보다 덜 빈번하게 발생하는 경우, 단어는 회귀 단어로서 플래깅될 수 있다. 회귀 단어들 및/또는 흔하게 사용되는 단어들은 시스템에 의해 저장될 수 있고, 회귀 및/또는 공통 단어들의 리스트들은 사용에 기초하여 업데이트될 수 있다.

[0117] [0124] 회귀 단어는 사용자의 알려진 언어로 빈번하게 사용되는 단어일 수 있다. 예컨대 "ossify"라는 단어는 다수의 영어 스피커들에게 회귀 단어일 수 있다. 회귀 단어는 사용자에게 알려진 언어와 상이한 언어의 단어일 수 있다. 예컨대, 도 13a에 도시된 중국어 구문 "guanxi"는 다수의 영어 스피커들에게 회귀 단어일 수 있다. 회귀 단어는 전문 분야의 용어, 예컨대, 약제의 이름, 크로켓과 같이 흔치 않은 스포츠에서 사용되는 용어 등일 수 있다. 회귀 단어는 약어, 비즈니스 또는 비영리 엔티티의 이름, 주식 기호일 수 있다. 회귀 단어는 단일 단어일 필요는 없고, 단어들 그룹(예컨대, "free electron laser"), 구문(예컨대, "bring your own device" 또는 그의 약어 형태 "BYOD") 등일 수 있다. 실제로, 단어 흐름 주석 시스템은 사용자가 주석에 대한 후보들로서 단어들의 소정의 카테고리들(이를테면, 위의 예들 중 일부)을 선택하게 할 수 있도록 구성될 수 있다.

[0118] [0125] 회귀 단어들 외에도 또는 그에 대안적으로, 키워드들은 맥락 키워드들일 수 있다. 맥락 키워드들은 대

화의 토픽과 관련된 단어들 또는 대화에서 빈번하게 사용되는 단어들을 포함할 수 있다. 예로서, AR 시스템(200)의 사용자는 자신의 친구와 아프리카로의 여행에 관해 이야기중일 수 있다. "Africa"라는 단어가 영어 언어에서 희귀 단어가 아닐 수 있지만, AR 시스템(200)은 대화가 아프리카와 관련되기 때문에 그것을 맥락 키워드로서 플래깅할 수 있다. 다른 예로서, 사용자가 아프리카로의 자신의 여행에 관해 이야기하는 동안, 사용자는 아프리카에서의 자신의 서핑 경험을 언급한다. 따라서, 사용자는 자신의 스피치에서 "surf"이라는 단어를 반복적으로 사용할 수 있다. AR 시스템(200)은 "surf"이라는 단어를 맥락 키워드로서 플래깅할 수 있는데, 그 이유는 문장에서 그것이 빈번하게 사용되기 때문이다.

[0119] [0126] 일부 단어들은 맥락 키워드 및 희귀 단어 둘 모두일 수 있다. 예컨대, 어떤 사람이 법리들에 관한 강의를 하는 경우 "res judicata"라는 구문이 여러 번 나타날 수 있다. "res judicata"라는 구문은, 이것이 라틴어 구문이고 일상 스피치에서 흔하게 사용되지 않기 때문에 희귀 단어로서 간주될 수 있지만, 이 구문은 또한 법 전문가의 스피치에서 그의 빈번한 발생으로 인해 맥락 키워드로서 간주될 수 있다.

[0120] [0127] AR 시스템은 또한 단어들 및 구문들 간의 관계를 결정하고 관계에 기초하여 키워드를 검출할 수 있다. 예컨대, 사용자의 친구가 "I surfed in Africa"라고 말할 수 있다. AR 시스템은 "surf"라는 단어 및 "Africa"라는 단어 둘 모두가 키워드들이 가능성이 높다고 결정할 수 있다. 그러나 사용자가 자신의 아프리카로의 여행에 관해 이야기했기 때문에, AR 시스템은 "surf"이라는 단어보다는, "Africa"라는 단어를 키워드로서 식별할 수 있다. 그러나 사용자가 서핑에 관해 많이 알지 못하는 경우, AR 시스템은 "Africa"라는 단어 보다는, "surf"이라는 단어를 키워드로서 식별할 수 있다.

[0121] [0128] 키워드는 미리 지정될 수 있다. AR 시스템은 사용자 또는 다른 당사자(예컨대, 애플리케이션 개발자, 쇼핑 웹사이트 등)에 의해 지정된 맥락 키워드들의 데이터베이스를 포함할 수 있다. 키워드들의 데이터베이스는 예컨대, 키워드들(또는 희귀 단어들)을 부가하거나 제거하기 위해 사용자에 의해 편집 가능할 수 있다. 예로서, AR 시스템(200)의 사용자는 최근 Linda라는 이름의 아기를 낳았을 수 있다. 사용자는 맥락 키워드로 "my daughter Linda"라는 구문을 세팅할 수 있다. 따라서, 사용자가 "my daughter Linda"라고 말한 것을 AR 시스템이 검출할 때, AR 시스템은 예컨대, 아기가 자고 있는지 또는 배고픈지와 같은 사용자의 아기에 대한 보조 정보를 자동으로 제시할 수 있다. 다른 예로서, AR 시스템은 "utility bill"이라는 구문을 키워드로서 태깅할 수 있다. 따라서, AR 시스템은 스피치에서 "utility bill"라는 구문의 검출 시에 만기일 및 금액을 포함하는 사용자의 유털리티 요금을 자동으로 제시할 수 있다.

[0122] [0129] 키워드들은 사용자, 환경 또는 대화의 참가자와 연관된 맥락 정보에 기초하여 결정될 수 있다. 예시적인 단어 흐름 주석 시스템은 적어도 부분적으로, 키워드 검출 시에 사용자 행동을 활용할 수 있다. 예컨대, 사용자는 사용자가 이해하지 못하는 단어를 들으면 소정의 방향을 일시적으로 응시할 수 있다. 이러한 예시적인 시스템은 예컨대, 응시 이전에 발생하는 스피치에서 유망한 희귀 단어를 검색하기 위해 내향 카메라(466)를 사용하여 이러한 특정 응시를 검출할 수 있다. 다른 예로서, 사용자는 단어를 키워드로서 프로세싱하도록 시스템에 촉구하기 위해 다른 사람에 의해 말해진 단어를 별개로 반복할 수 있다. 시스템은 키워드들의 그의 검출을 돋기 위해 다른 사용자 행동을 사용할 수 있다. 실제로 구현은, 사용자가 키워드 검출 시에 구현에 의해 통합된 행동들을 커스터마이징하게 할 수 있을 수 있다. 예컨대, AR 시스템은, 공통 단어들 또는 희귀 단어들의 리스트에 사용자가 단어들을 부가하거나 제거할 수 있게 하는 사용자 인터페이스를 제공할 수 있다. AR 시스템은 희귀 단어 분류를 위해 사용자 자신의 스피치를 모니터링할 수 있다. 예컨대, 암호 작성자(cryptographer)의 스피치는 AES(Advanced Encryption Standard)라는 암호가 포함할 수 있다. 암호 작성자/사용자의 AES라는 암호를 여러 번(예컨대, 3 번 또는 5 번) 또는 구성 가능한 시간 지속기간에 여러 번(예컨대, 5 분에 3 번, 한 시간에 7 번 등) 검출 시에, AR 시스템은 이 사용자에 대한 공통 단어로서 AES라는 암호를 재분류할 수 있다. AR 시스템은, 그것이 검출한 단어들의 빈도들을 모니터링하고 공통 또는 희귀 단어들의 리스트들을 업데이트할 수 있다. 예컨대, (콜레스테롤 약제에 대한) "simvastatin"이라는 단어는 다수의 사용자들에 대해 희귀 단어이지만, 건강 관리 전문가들이 사용자들에 대해선 그렇지 않을 수 있다. 건강 관리 전문가인 사용자에 대해, 건강 관리 전문가의 AR 시스템에 의해 검출된 스피치에서 이 단어가 드물게 사용되지 않는다는 것을 AR 시스템이 검출하고, 이 사용자에 대한 공통 단어로서 이 단어를 재분류할 수 있다. 재분류는 예컨대, 사용자에 대한 공통 단어 사전에 단어를 부가함으로써 달성될 수 있다. 사용자에 대한 공통 단어 사전은 도 2에 예시된 시스템(200)의 로컬 데이터 모듈(260) 또는 원격 데이터 리포지토리(280)에 저장될 수 있다.

[0123] [0130] 단어가 희귀 단어인지 또는 공통 단어인지를 분류하는 것을 참조하여 이 예가 설명되지만, 일부 실시예들에서, 분류/재분류는 일반적으로 키워드에 관한 것이다. 단어가 여러 번 사용됨을 AR 시스템이 검출할 때, 또는 희귀 단어가 대화에 나타날 때, AR 시스템은 이 단어를 키워드로서 플래깅할 수 있다. 예컨대, "petrif

y"라는 단어가 사용자의 대화에서 여러 번 사용됨을 AR 시스템이 검출할 때, AR 시스템은 "petrify"라는 단어를 키워드로서 분류할 수 있다. 그러나 AR 시스템은 나중에, 사용자가 화석들과 관련된 대화에 관여하는 것으로 결정할 수 있다. AR 시스템은 "petrify"라는 단어가 더는 키워드가 아니도록 그것을 재분류할 수 있는데, 그 이유는, 사용자가 "petrify"라는 단어의 의미를 학습(다른 방식으로, 인지함)한 것으로 AR 시스템이 결정할 수 있기 때문이다.

[0124] [0131] 일부 구현들에서, 희귀 단어는 쇠퇴 팩터(decay factor)와 연관될 수 있어서, 보조 정보가 사용자에게 빈번하게 제시된 경우, 희귀 단어의 의미를 사용자가 학습한다는 가정 하에서, 희귀 단어와 연관된 보조 정보의 프리젠테이션이 (적어도 부분적으로 쇠퇴 팩터에 기초하여) 쇠퇴(decay)하거나 심지어 중단되는 경향이 있게 된다. 쇠퇴 팩터는 시간 기간을 포함할 수 있어서, 희귀 단어가 시간 기간 동안 임계 횟수보다 많이 언급되는 경우, 보조 정보의 디스플레이가 쇠퇴하거나 중단되게 된다. 쇠퇴 팩터는 숫자 카운트를 포함할 수 있어서, 희귀 단어가 언급된 횟수가 숫자 카운트를 초과하는 경우, 보조 정보의 디스플레이가 쇠퇴거나 중단되게 된다. 쇠퇴 팩터, 시간 기간 또는 숫자 카운트(또는 연관된 임계치들)는 사용자 구성 가능할 수 있다.

[0125] [0132] 키워드들 외에도 또는 이들에 대안으로서, AR 시스템은 또한 사용자의 환경에서 (예컨대, 객체 또는 사람과 같은) 관심 객체를 검출할 수 있다. AR 시스템은 도 7에 설명된 하나 이상의 객체 인식기들(708)을 사용하여 관심 객체를 검출할 수 있다. 예로서, AR 시스템(200)의 사용자는 Florida에 거주하고 있다. 사용자는 또한 휴가로 California에 갈 수 있다. 사용자는 예컨대, AR 디스플레이(220)를 통해 California의 해변에 누워있는 바다 사자를 볼 수 있다. AR 시스템(200)은 외향 이미징 시스템(464)을 사용하여 바다 사자를 검출하고 객체 인식기(708)를 사용하여 바다 사자를 식별할 수 있다. AR 시스템(200)은, 사용자가 통상적으로 Florida에서 바다 사자와 대면하거나 보지 못하기 때문에 관심 객체로서 바다 사자를 플래깅할 수 있다.

보조 정보 예시적인 리트리벌

[0127] [0133] 단어 흐름 주석 시스템은 시스템이 검출한 관심 객체와 연관된 보조 정보를 리트리브할 수 있다. 보조 정보는 텍스트, 이미지, 또는 다른 오디오 또는 시각적 정보를 포함할 수 있다. 시스템은 시각적으로(예컨대, 디스플레이(220)를 통해), 오디오로(예컨대, 오디오 증폭기(232) 또는 스피커(240)를 통해), 촉감(예컨대, 햅틱) 피드백 등을 통해 보조 정보를 제시할 수 있다.

[0128] [0134] 단어 흐름 주석 시스템은 관심 객체들 및 이들과 연관된 보조 정보를 포함하는 정보의 데이터 리포지토리(예컨대, 데이터베이스)를 포함할 수 있다. 예컨대, 데이터 리포지토리는 공통 단어들, 희귀 단어들, 다른 맥락 키워드들, (사용자가 종종 상호작용하는) 사용자의 환경의 공통 객체들 등을 저장할 수 있다. 보조 정보는 (도 7을 참조하여 설명된 바와 같은) 의미 정보, 의미 또는 설명, 객체와 연관된 구매 정보, 상태 또는 통계, 사용자가 명확히 지각할 수 없을 수 있는 관심 객체의 부가적인 세부사항들, 또는 관심 객체에 대한 이해 및 그와의 상호작용들을 용이하게 하는 다른 정보를 포함할 수 있다. 보조 정보의 예들은 도 13a 내지 도 13e, 도 15a 및 도 15b에서 추가로 설명된다.

[0129] [0135] 이러한 정보의 데이터베이스는 예컨대, 도 2의 데이터 모듈(260)에 로컬로 저장되거나 예컨대, 원격 데이터 리포지토리(280)에 원격으로 저장될 수 있다. 단어 흐름 주석 시스템은 공개적으로 액세스 가능한 정보, 예컨대, 인터넷 상의 정보를 활용하여 희귀 단어와 연관된 보조 정보를 리트리브할 수 있다. 이러한 상황에서, 시스템은 희귀 단어에 관한 질의를 인터넷 상의 자원, 사전, 백과사전 또는 다른 유사한 자원에 전송하기 위해 네트워크에 액세스할 수 있다. 이러한 자원은 일반적(예컨대, 위키피디아와 같은 범용 백과사전)이거나, 또는 예컨대, 약제들의 인덱스, 이를테면, rxlist.com 또는 광물학 데이터베이스(예컨대, webmineral.com) 상의 것과 같이 전문화될 수 있다.

[0130] [0136] 데이터 리포지토리 내의 관심 객체들은 AR 시스템(200)에 의해 업데이트될 수 있다. 예로서, AR 시스템(200)의 단어 흐름 주석 시스템(970)은 시스템이 희귀 단어들로서 식별하는 단어들의 세트로부터 단어들을 제거하도록 구성될 수 있다. 이는 예컨대, 사용자가 전문 지식을 가진 상황(예컨대, 약사는 다수의 약제들의 이름들을 알 가능성이 높음), 사용자가 연관 또는 사용을 통해 단어와 친숙해지는 상황(예컨대, Acme Corporation의 직원은 Acme Corporation에 관한 보조 정보가 필요하지 않을 가능성성이 높음)에서 유용할 수 있다. 예시적인 단어 흐름 주석 시스템은 희귀 단어들의 세트로부터 단어를 제거할 수 있다. 이는, 공통 단어 사전에 단어를 부가함으로써 구현될 수 있다. 이는, 시스템이 희귀 단어를 검출한 후에 부가적인 검사로서 구현될 수 있다. 검출된 희귀 단어가 제거될 단어들의 세트에서 발견된 경우, 시스템은 단어와 연관된 보조 정보를 리트리브하거나 디스플레이하지 않는다.

- [0131] [0137] 사용자는 UI 상호작용을 통해 희귀 단어들의 세트로부터 특정 단어를 제거할 수 있다. 예컨대, 희귀 단어 및 그의 연관된 보조 정보가 디스플레이될 때, 사용자의 입력은 사용자 입력 장치(504)를 통해 입력될 수 있다. 예시적인 시스템은 소정의 횟수, 예컨대, 3 또는 5 번 또는 사용자에 의해 구성된 임의의 수로 단어 및 그의 연관된 보조 정보를 디스플레이한 후에, 특정 단어를 그의 공통 단어 사전에 부가하도록 구성될 수 있다.
- [0132] [0138] 예컨대, 희귀 단어를 제거하거나 부가하기 위해 UI 상호작용이 사용될 수 있다. 예컨대, 사용자는 "Add the word phoropter as a common word"라고 말할 수 있고, AR 시스템은 사용자의 스피치를 검출할 수 있고 단어를 공통 단어 리스트에 부가하거나 단어를 희귀 단어 리스트로부터 제거할 수 있다.
- [0133] [0139] 데이터 리포지토리를 업데이트하는 이러한 예들이 희귀 단어들을 참조하여 설명되지만, 유사한 기술들은 또한 데이터 리포지토리 내의 다른 관심 객체들을 업데이트하는 데 사용될 수 있다.
- [0134] 보조 정보와의 예시적인 상호작용들
- [0135] [0140] AR 시스템(200)은 보조 정보와의 다양한 사용자 상호작용들을 지원하도록 구성될 수 있다. 예컨대, AR 시스템은 보조 정보를 디스플레이, 해제 또는 리콜(recall)할 수 있다. 또한, 사용자는 보조 정보의 일부를 마킹(mark)하거나 보조 정보에 주석을 달 수 있다.
- [0136] [0141] AR 시스템은 관심 객체의 상이한 발생들에 대해 특정 관심 객체와 연관된 상이한 보조 정보의 상이한 피스(piece)들을 디스플레이하도록 구성될 수 있다. 이는 쉽게 디스플레이될 수 있는 것보다 많은 보조 정보가 이용 가능할 때 유용할 수 있다. 예컨대, 구현은, 희귀 단어의 제 1 발생 시에 웹사이트 # 1로부터 보조 정보를 리트리브하고, 그의 제 2 발생 시에 웹사이트 # 2로부터 보조 정보를 리트리브하는 식으로 구성될 수 있다. 다른 예로서, 구현은 그의 제 1 발생시에 약제의 사용을 디스플레이하고, 그의 제 2 발생 시 약제의 부작용을 디스플레이하는 식으로 구성될 수 있다. 보조 정보의 어떤 피스를 디스플레이할지 결정하는 데 있어 유용한 데이터는 로컬 데이터 모듈(260) 또는 원격 데이터 리포지토리(280)에 저장될 수 있다. 보조 정보의 상이한 조각들을 사용자에게 순차적으로 제공함으로써, AR 시스템(200)은, 디스플레이되는 보조 정보가(예컨대, 디스플레이의 FOV를 너무 많이 커버함으로써) 사용자를 산만하게 하거나, 또는(예컨대, 사용자가 모든 정보를 판독할 수 있으면에는 너무 짧은 시간 동안 보조 정보를 디스플레이함으로써) 사용자가 이해하기에 너무 어렵게 될 가능성을 감소시킬 수 있다.
- [0137] [0142] AR 시스템은 사용자의 구성에 기초하여(예컨대, 사용자의 지식 레벨에 따라) 특정 희귀 단어와 연관된 상이한 보조 정보를 디스플레이하도록 구성될 수 있다. 예컨대, 사용자(예컨대, 약사)가 전문화된 약제 정보를 위해 시스템을 구성한 경우, 시스템은 약제에 관한 전문화된 정보를 리트리브 및 디스플레이할 수 있다. 그렇지 않으면, 시스템은 약제에 관한 일반 정보를 리트리브 및 디스플레이할 수 있으며, 이는 전문화된 약제 정보를 위해 시스템을 구성하지 않은 사용자에게 적절할 가능성이 높다. 시스템은 적절한 정보를 제공하기 위해 상이한 정보 소스들을 전문화된 또는 일반 정보와 연관시킬 수 있다. 예컨대, 사전은 일반 정보의 소스일 수 있고; 약제 웹사이트는 전문화된 정보의 소스일 수 있다.
- [0138] [0143] 단어 흐름 주석 시스템은 스피치의 맥락에 기초하여 희귀 단어와 연관된 보조 정보를 디스플레이하도록 구성될 수 있다. 희귀 단어의 가까운 맥락으로부터의 부가적인 단어는, 어떤 의미가 가장 가능성이 높은지 또는 어떤 종류의 보조 정보가 가장 유용할 수 있는지를 결정하는 데 사용될 수 있다. 이는, 스피치와 관련된 정보를 사용자에게 제공하거나, 또는 희귀 단어와 연관된 모호성을 해결하는 데 유용할 수 있다. 예컨대, Acme Corporation이라는 이름이 도 13d에 예시된 바와 같이 회사 간부들의 맥락의 스피치에서 나타나는 경우, 구현은 박스(1340)에 도시된 바와 같이 Acme Corporation의 간부들의 이름들 및 직함들을 리턴할 수 있다. Acme Corporation이라는 이름이 도 13e에 예시된 바와 같이 주식 가격들의 맥락의 스피치에서 나타나는 경우, 구현은 박스(1354)에 도시된 바와 같이 Acme Corporation의 주식 정보를 리턴할 수 있다. 다른 예로서, "IMDB"라는 약어는 예컨대 internet movie database", "immigrant database", "in-memory database" 등과 같은 다양한 상이한 의미들을 가질 수 있다. 단어 흐름 주석 시스템은 올바른 의미를 결정하기 위해 맥락 정보를 결정할 수 있다. 일 예에서, 단어 흐름 주석 시스템은 스피커의 아이덴티티를 결정하고 아이덴티티는 스피커의 아이덴티티와 상관되는 의미를 갖는다. 다른 예에서, 단어 흐름 주석 시스템은 희귀 단어를 둘러싼 독특한 단어들을 사용하여 그러한 단어들과, 그 단어의 특정 해석과 연관된 보조 정보 간의 상관성(correlation)들을 컴퓨팅할 수 있다. 단어 흐름 주석 시스템은 스피치의 맥락을 결정하기 위해 가장 공통적인 N개의 단어들 사이에 있지 않은 단어들을 식별 및 해석할 수 있다. 위의 IMDB 예에서, "IMDB"라는 단어는 스피커의 스피치에서 "visa", "legislation", 및 "deportation"와 같은 단어들을 동반한다. 이러한 모든 동반 단어들은 "internet movie database" 해석이 아니라 "IMDB"의 "immigrant database" 해석과 연관될 가능성이 높다. 따라서, AR 시스템은

"immigration database"로서 "IMDB"의 보조 정보를 디스플레이하도록 자동으로 선택할 수 있다.

[0139]

[0144] 단어 흐름 주석 시스템은 사용자에 대해 리트리브 또는 디스플레이될 후보인 보조 정보와 희귀 단어의 맥락 간의 상관성을 컴퓨팅할 수 있다. 맥락 분석과 연관된 프로세싱은 예컨대, 로컬 프로세싱 모듈(71) 또는 원격 프로세싱 모듈(72)에 의해 로컬로 또는 원격적으로 수행될 수 있다. 맥락 분석은 희귀 단어를 둘러싼 독특한 단어들의 사용에 기초하여, 그러한 단어들과, 그 단어의 특정 해석과 연관된 보조 정보 간의 상관성들을 컴퓨팅할 수 있다. 독특한 단어들은 가장 공통적인 N개의 단어들 사이에 있지 않은 단어들일 수 있으며, N의 일부 값은 공통적인 단어 사전의 엔트리들의 수 또는 발언 빈도 랭크의 임계치보다 작을 수 있다. 높은 상관성은 관련 보조 정보를 나타낼 수 있다. 낮은 상관성은 무관한 보조 정보를 나타낼 수 있다. 예컨대, 희귀 단어가 "IMDB"이고 다른 동반 단어들이 "visa", "legislation", 및 "deportation"를 포함하는 경우, 단어들은 "Internet Movie Database," "in-memory database" 또는 일부 다른 해석이 아니라, "Immigrant Database" 해석에 대한 보조 정보에서 발견될 가능성이 높고, 시스템은 더 높은 상관성에 기초하여 "Immigrant Database"에 대한 보조 정보를 디스플레이하도록 구성될 수 있다. 이는 약어 모호성의 한 예이며, 맥락-의존적 보조 정보의 유형이다.

[0140]

[0145] 보조 정보의 디스플레이는 다양한 방식들로 제시될 수 있다. 예컨대, 시스템은 도 13a의 박스(1308)에 예시된 바와 같이, 단지 희귀 단어들 및 연관된 보조 정보만을 디스플레이할 수 있다. 시스템은 스피치의 롤링 전사물(rolling transcript)을 디스플레이하도록 구성될 수 있다. 이 구성에서, 예컨대, 사용자가 순간적으로 산만해진 경우에, 놓친 단어들 또는 짜증나는 문장들이 신속하게 재판독될 수 있다. 본 개시내용에서, 보조 정보라는 용어는 전사물을 포함한다.

[0141]

[0146] 대화 파트너의 스피치 트레인(speech train)은 영화의 엔딩 크레딧(end credit)의 프리젠테이션과 유사한 롤링 텍스트로서 디스플레이될 수 있다. 이용 가능한 경우, 스피커들의 이름들이 또한 디스플레이될 수 있다. 시스템이 스피커들의 이름들을 획득할 수 있는 방법의 예들이 위에서 설명된다. 구현은 말풍선들에 정보(예컨대, 스피치 또는 보조 정보의 전사물), 예컨대, 도 13에 예시된 박스(1308)와 같이 스피커를 근처에 기하학적으로 로컬화된 텍스트를 배치할 수 있다.

[0142]

[0147] 스피치의 전사물을 디스플레이하는 시스템은 일부 방식으로, 예컨대, 밀줄 글기, 채색, 또는 굵은 텍스트 등으로 희귀 단어를 강조할 수 있다. 시스템은 UI 상호작용을 통해, 사용자가 현재 또는 과거 희귀 단어를 선택하고 연관된 보조 정보를 수집하거나 다시 가져오게 할 수 있을 수 있다.

[0143]

[0148] 구현은 UI 상호작용을 통해 정보에 액세스하기 위해 사용자의 눈 움직임을 최소화하도록 보조 정보(예컨대, 말풍선 또는 롤링 전사물)를 배치할 수 있다. 이러한 방식으로, UI가 단순화되고 사용자는 자신의 주의를 스피커로부터 분산시킬 필요가 없다. 판독 동작을 대화 파트너에게 최소한으로 가시적이 되도록 보조 정보가 배치될 수 있으며, 이렇게 하면, 보조 정보에 대한 사용자의 액세스를 드러내지 않으면서, 더 적은 산만함 및 더 양호한 통신을 제공할 수 있다. 예컨대, 스피커의 위치를 결정할 수 있는 구현은 스피커 옆에 보조 정보를 배치할 수 있다. 예컨대, AR 시스템의 외향 카메라로부터의 이미지들은 예컨대, 스피커의 얼굴, 제스처 등을 모호하게 하지 않는 적절한 배치의 결정을 도울 수 있다. AR 시스템은 보조 정보 디스플레이의 배치를 결정하기 위해 도 8에 예시된 프로세스 흐름(800)을 이용할 수 있다. 예컨대, 블록(2310)에서 인식된 객체는 스피치가 주석을 위해 프로세싱되는 스피커일 수 있다.

[0144]

[0149] 사용자 또는 대화 파트너에 의해 경험되는 산만함을 감소시키는 다른 예로서, 보조 정보가 오디오로 제시되는 경우, AR 시스템은, 사용자 또는 대화 파트너 어느 누구도 말하고 있지 않을 때, 정보를 제시하고 그리고/또는 대화 파트너에게 들리지 않지만, 사용자에게 충분히 큰 볼륨으로 정보를 제시할 수 있다.

[0145]

[0150] 리트리브된 보조 정보는 AR 시스템(200)의 디스플레이에 의해 사용자에게 디스플레이될 수 있다. 디스플레이된 보조 정보는 조건이 충족될 때까지 가시적인 채로 남아 있을 수 있다. 예컨대, 디스플레이되는 정보는 다음 희귀 단어(또는 다른 관심 객체)가 디스플레이될 때까지 또는 사용자 액션에 의해 해제될 때까지 고정된 시간량 동안 가시적인 채로 남아 있을 수 있다. 사용자 액션은 수동적일 수 있는데, 예컨대, 눈 움직임들일 수 있다. 내향 카메라(예컨대, 도 4의 카메라(462))는 사용자의 눈 움직임들을 검출하거나 추적하는 데 사용될 수 있다. 디스플레이는, 보조 정보의 전체 디스플레이 영역을 추적하기 위해 시스템에 의해 사용자가 관찰된 후에 눈 추적을 통해 해제될 수 있다. 예컨대, 보조 정보가 텍스트인 경우, 시스템은 텍스트 전체를 통한 사용자의 눈 움직임들(예컨대, 좌측에서 우측 그리고 위에서 아래)을 추적할 수 있다. 디스플레이는, 보조 정보의 전체 디스플레이 영역으로부터 눈길을 돌리기 위해(또는 이를 보지 않기 위해) 시스템들에 의해 사용자가 관찰된 후에 눈 추적을 통해 해제될 수 있다. 사용자 액션은 능동적일 수 있는데, 예컨대, 도 4의 사용자 입력 디바이

스(504)를 통한 입력 액션일 수 있다. 단어 흐름 주석 시스템은 특정 사용자에 대한 사용자 인터페이스(UI) 상호작용의 커스터마이징된 세트를 지원하도록 구성될 수 있다. UI 상호작용은 손가락, 포인터 또는 일부 종류의 스타일러스로, 또는 눈들의 시선 및 버튼에의 후속 응시에 의해 작동되는 버튼과 유사한 UI 엘리먼트 또는 다른 것의 형태를 취할 수 있다. 버튼은 실제 물리적 버튼이거나 AR에 디스플레이된 버튼일 수 있다. UI 상호작용들은 예컨대, 도 4와 관련하여 위에서 설명된 바와 같이 머리 포즈의 형태를 취할 수 있다. UI 상호작용 검출의 예가 도 10과 관련하여 위에서 설명된다.

[0146] [0151] 단어 흐름 주석 시스템은 보조 정보 디스플레이의 해제를 지연시키도록 사용자를 촉구할 수 있다. 예컨대, 시스템은, 디스플레이가 해제될 것임을 사용자에게 통지하기 위해, 짧게, 예컨대, 몇 초 동안 디스플레이된 보조 정보의 컬러 배합(color scheme)을 변경하거나 밝기를 감소시킬 수 있다. 위에서 설명된 것과 같은 UI 상호작용은 해제를 연기하는 데 사용될 수 있다. 예컨대, AR 시스템이 눈 추적을 통해, 사용자가 보조 정보를 능동적으로 뷰잉하고 있음을 검출하는 경우, AR 시스템은 해제를 연기할 수 있다.

[0147] [0152] 위에서 설명된 것과 같은 UI 상호작용은 또한 해제된 보조 정보를 리콜하는 데 사용될 수 있다. 예컨대, 사용자 입력 디바이스를 통한 입력 액션은 가장 최근에 디스플레이된 보조 정보를 리콜하는 데 사용될 수 있거나, 또는 리콜을 위해 이전에 디스플레이된 보조 정보의 특정 아이템을 선택하는 데 사용될 수 있다.

[0148] [0153] 보조 정보를 디스플레이하는 동안, AR 시스템은 또한 사용자가 보조 정보에 주석을 달게 할 수 있을 수 있다. 예로서, AR 시스템은 나중의 뷰잉을 위해 사용자가 보조 정보의 부분을 마킹할 수 있게 할 수 있다. 예컨대, 사용자는 보다 심도 있게 판독하도록 희귀 단어의 설명의 부분을 강조할 수 있다. 다른 예로서, 사용자가 프로 농구 게임을 시청하고 있는 동안, 해설자들이 특정 플레이어에 대해 이야기하고 있다. AR 시스템은 보조 정보로서 그 플레이어가 홍보한 제품을 제시할 수 있다. 사용자는 AR 시스템에 의해 제시된 제품을 저장하거나 북마킹(bookmark)하고 제품을 구매할지를 나중에 결정할 수 있다.

[0149] [0154] 사용자는 또한 보조 정보와 연관된 가상 콘텐츠를 부가하거나 드로잉(drawing)함으로써 관심 객체에 주석을 달 수 있다. 예컨대, 사용자는 사용자 입력 디바이스(466) 또는 손 제스처들을 사용하여, 보조 정보에 메모들 또는 그림들을 부가하거나, 또는 (예컨대, 부분을 지우거나 줄을 그음으로써) 보조 정보의 부분을 제거할 수 있다. AR 시스템은, 사용자의 주석을 저장하고, 관심 객체가 나중에 검출될 때 보조 정보와 함께 사용자의 주석을 렌더링할 수 있다.

관심 객체와의 상호작용들을 위한 예시적인 사용자 경험들

[0151] [0155] 도 13a 내지 도 13g는 관심 객체와 상호작용하는 예들 및 관심 객체의 보조 정보를 예시한다.

키워드의 존재에 기초하여 가상 콘텐츠를 제시하는 예들

[0153] [0156] 도 13a는 키워드를 검출하고 키워드에 대한 보조 정보를 제시하는 예를 예시한다. 도 13a에서, 사용자의 AR 시스템의 단어 흐름 주석 시스템(970)은 키워드들을 검출하고 키워드들의 보조 정보를 제공하도록 구성될 수 있다. AR 시스템의 사용자에게 디스플레이되는 정보가 어수선하게 흘어지는 것을 방지하기 위해, AR 시스템은 희귀 단어들의 검출 시에만 보조 정보를 제공할 수 있다.

[0154] [0157] 도 13a에서, ARD를 착용한 사용자(도시되지 않음)는 자신의 환경에서 여성(1302)을 지각할 수 있다. 여성은 콜-아웃 박스(1304)에 도시된 문장을 말한다. 문장 내에 희귀 단어일 수 있는 "guanxi"라는 단어가 있다. ARD는 (예컨대, 오디오 센서(232)를 통해)을 여성이 말한 문장 캡처하고 문장을 텍스트로 변환하고 희귀 단어로서 "guanxi"를 검출하고 "guanxi"와 연관된 보조 정보(예컨대, 정의)를 리트리브하고, 디스플레이(예컨대, 디스플레이(220)) 상에 보조 정보를 디스플레이할 수 있다. 이 예시에서, 보조 정보는 박스(1308)에 디스플레이된다. 사용자는 AR 디스플레이(220)에 의해 제시된 가상 콘텐츠의 부분으로서 이 박스(1308)를 본다. 박스(1308)의 보조 정보는, "guanxi"라는 단어를 사용함으로써 여성이 의미하는 것을 사용자가 이해하도록 도울 수 있다.

텔레프레전스 세션에서의 예시적인 사용자 경험들

[0156] [0158] 복수의 AR 시스템들의 복수의 사용자들은 시스템들의 도움을 통해 원격으로 대화를 수행할 수 있다. 예컨대, 도 12에 예시된 바와 같이, 2개의 물리적 위치들의 (디바이스들(100a 및 100b)의) 2명의 사용자들(이에 따라, 이들은 사람-제작 디바이스(man-made device)의 도움없이, 직접 서로를 보거나 들을 수 없음) 둘 모두는, 단어 흐름 주석 시스템(970)이 장착될 수 있는 AR 장치가 장착될 수 있다. 각각의 사용자의 스피치는 사용자의 각각의 디바이스의 오디오 센서에 의해 캡처되고 네트워크(1290)를 통해 송신될 수 있다. 사용자 C의 스피치는

예컨대, 보조 정보를 더한 희귀 단어들로서 또는 전사물로서 사용자 B의 디바이스에 디스플레이될 수 있고, 그 반대로 가능하다.

[0157] 단어 흐름 주석 디바이스는 스피치를 로컬로 텍스트로 변환할 수 있고 네트워크(1290)를 통해 변환된 텍스트만을 송신할 수 있다. 다른 사용자의 디바이스는 텍스트를 디스플레이하거나 텍스트를 다시 스피치로 변환할 수 있다. 이는, 대응하는 스피치를 송신하는 것보다 텍스트를 송신하는데 더 적은 양의 데이터가 요구되기 때문에 네트워크(1290)의 대역폭이 제한되는 경우 유리할 수 있다. 단어 흐름 주석 시스템은 사용자-스피커의 음성 정보를 포함할 수 있고 그 음성 정보를 사용하여 텍스트를 스피치로 다시 변환할 수 있다. 이러한 방식으로, 재구성된 스피치는 스피커 자신의 음성처럼 들릴 수 있다.

[0158] 단어 흐름 주석 시스템은 또한 AR 디스플레이(220) 상에 제시된 이미지들을 통해 텔레프레전스 대화를 향상시킬 수 있다. 예컨대, AR 디스플레이에는 보조 정보와 함께 원격 스피커의 아바타를 제시하여 참가자의 시각뿐만 아니라 청각을 사로잡을 수 있다. 위의 4와 관련하여 설명된 바와 같이, 내향 이미징 시스템(462)이 장착된 ARD는 HMD에 의해 가려진 착용자의 얼굴의 구역을 대체하기 위한 이미지들을 캡처할 수 있으며, 이는 텔레프레전스 세션 동안 제 1 호출자가 제 2 호출자의 가려지지 않은 얼굴을 볼 수 있도록 사용될 수 있으며, 그 반대로 가능하다. 제 1 사용자와 연관된 세계 맵 정보는 단어 흐름 주석 시스템들을 포함하는 텔레프레전스 세션의 제 2 사용자에게 통신될 수 있다. 이는 ARD 착용자가 보게 될 원격 사용자의 이미지들의 생성을 통해 사용자 경험을 향상시킬 수 있다.

[0159] 텔레프레전스 애플리케이션에서, 주변 사운드들을 캡처하는 것은, 인-퍼슨 시나리오의 사용자-청취자와 연관된 디바이스 보다는, 사용자-스피커와 연관된 디바이스에 의해 수행된다. 스피치의 존재의 검출 및 스피치-텍스트의 변환은 어느 한 사용자와 연관된 디바이스에 의해 수행될 수 있다. 스피치의 소스는 사운드를 캡처하는 디바이스에 기초하여 결정될 수 있는데, 예컨대, 사용자 A의 디바이스가 사운드를 캡처할 때, 사용자 A가 말하고 있다.

[0160] 도 13b는 텔레프레전스 세션에서의 단어 흐름 주석 시스템의 다른 예시적인 애플리케이션을 도시한다. 이 예에서, 여성(1326)은 사용자들(1328 및 1320)과 마케팅 그룹 미팅중에 있다. 사용자들(1328, 1320)은 각각의 ARD들(1322b, 1322c)을 착용하고 있다. ARD들(1322b 및 1322c)은 도 2에 도시된 AR 시스템(200)의 적어도 부분을 포함할 수 있다.

[0161] 여성(1326)은 (때로는 말풍선으로서 또한 지칭되는 콜아웃 박스(1312)에 도시된 바와 같이) "Let's discuss an ad campaign to sell phoropters in Bruges"라 언급한다. ARD들(1322b, 1322c)은 이 오디오를 캡처하고 (예컨대, 단어 흐름 주석 시스템(970)을 통해) 오디오에서 키워드들을 식별할 수 있다. 일부 실시예들에서, 여성은 그녀의 부근에 오디오 디바이스(예컨대, 마이크로폰)를 가질 수 있거나 ARD를 착용하여서, 여성의 ARD 또는 오디오 디바이스가 단독으로 또는 조합으로, 오디오를 캡처하고 오디오 스트림을 프로세싱(예컨대, 그것을 텍스트로 변환하거나 또는 키워드들을 식별함)할 수 있다. 일부 실시예들에서, 오디오 파일을 텍스트로 변환하거나 키워드들을 식별하는 것과 같은 오디오 프로세싱의 적어도 일부는 원격 컴퓨팅 시스템(1220)에 의해 수행될 수 있으며, 이는 각각의 사용자의 ARD에 의한 동일한 오디오 스트림의 중복 프로세싱을 회피할 수 있다.

[0162] ARD(1322b 및 1322c)는 여성의 스피치와 연관된 보조 정보를 디스플레이할 수 있다. 이 예에서, "phoropter"(a medical device used by eye care professionals to determine eyeglass prescriptions)라는 단어가 사용자(1320)에 대한 희귀 단어인 것으로 결정되고, ARD(1322c)는 박스(1316)에 "phoropter"라는 단어의 정의를 디스플레이한다. "phoropter"라는 단어가 ARD의 사용자에 대한 희귀 단어가 아닌 경우, ARD는 사용자에 대해 박스(1316)를 디스플레이하지 않을 수 있다. 이 예에서 "phoropter"라는 단어는, 사용자가 눈 관리 부서의 엔지니어이기 때문에 사용자(1328)에 대해 희귀 단어가 아니다. 그 결과, ARD(1322b)는 (박스(1316)에 도시된) 보조 정보를 착용자(1328)에게 제시하지 않을 것이다. 이 예를 계속하면, ARD(1322b)는 "Bruges"(a city in Belgium)라는 단어가 사용자(1328)에 대한 희귀 단어라고 결정할 수 있고, 도시에 관한 정보 및 도시의 위치의 작은 맵을 제공하는 그래픽(1324)을 디스플레이할 수 있다.

언어 번역

[0164] 단어 흐름 주석 시스템은 언어 번역을 위해 사용될 수 있다. 도 13c는 단어 흐름 주석 시스템의 다른 예시적인 애플리케이션을 도시하며, 여기서 단어 흐름 주석 시스템은 제 1 언어(예컨대, 스페인어)의 구문을 제 2 언어(예컨대, 영어권 사용자의 경우, 영어)로 번역할 수 있다. 이 예에서, AR 시스템(200)의 영어권 사용자

(도시되지 않음)는 외국을 방문하고 있고, 점선 박스(1332)에 도시된 바와 같이, 스피커(1390)가 "El supermercado esta al otro lado de la calle"라고 말하는 것을 듣고 있다. AR 시스템(200)은 스피커(들)의 스피치를 사용자에게 디스플레이(반드시 그럴 필요는 없음)할 수 있다(예컨대, 박스(1332)는 통상적으로 AR 시스템의 사용자에게 디스플레이되지 않지만, 일부 구현들에서, 디스플레이될 수 있음). AR 시스템(200)은 스피커(1390)의 스피치가 스페인어로 되어 있음을 결정하고, 스피치를 영어로 번역하고("The grocery store is across the street"), 번역된 텍스트를 갖는 캡션(1336)(또는 말풍선)을 디스플레이한다.

[0165] AR 시스템은 로컬 프로세싱 및 데이터 모듈(260), 원격 프로세싱 모듈(270) 또는 원격 데이터 리포지토리(280)를 단독으로 또는 조합하여 사용하여 하나의 언어에서 다른 언어로의 번역을 지원할 수 있다. 예컨대, AR 시스템은 사용자의 지배적인 언어 이외의 다른 언어를 사용자의 지배적인 언어로 번역할 수 있는 단어 흐름 주석 시스템(970)을 구현할 수 있다. 사용자의 지배적인 언어는 사용자의 모국어 또는 대화에서 사용자의 선호 언어일 수 있다. 사용자의 지배적인 언어 이외의 다른 언어는 외국어로 간주될 수 있다. 이러한 구성에서, 사용자는 도 13c에 예시된 바와 같이 맥락 보조 정보 보다는, 인입하는 단어들의 번역들을 보도록 선택할 수 있다. 사용자는 사용자의 지배적인 언어로, 스피커의 언어로, 둘 모두의 언어들로(예컨대, 행간(interlinear) 포맷으로), 검출된 스피치의 전사물, 언어들을 혼합한 전사물(예컨대, 스피커의 언어로 된 공통 단어들 및 스피커의 언어로 또는 둘 모두의 언어들로 된 비공통 단어들) 등을 보도록 선택할 수 있다. 시스템은 공통 외국어 사전의 커스터마이징(customization)을 지원할 수 있어서, 예컨대, 사용자의 모국어의 공통 단어 사전보다 작은 크기를 허용한다. 예컨대, 사용자는 특정 언어로 1,000 또는 2,500(또는 일부 다른 수)개의 가장 많이 사용된 단어들보다 덜 공통적인 임의의 단어의 번역을 보도록 선택할 수 있다. 동일한 사용자는 7,000 또는 8,000(또는 일부 다른 수)개의 단어들의 크기를 갖도록 사용자 자신의 언어의 공통 단어 사전을 선택할 수 있다.

[0166] 단어 흐름 주석 시스템은 사용자가 외국어의 스피치를 이해하는 데 도움이 될 수 있을 뿐만 아니라 사용자가 외국어를 말하는 데도 도움이 될 수 있다. 예컨대, 사용자 자신의 스피치를 외국어로 번역하도록 시스템이 구성될 수 있다. 이러한 번역은 사용자의 전체 스피치(예컨대, 전체 문장) 또는 사용자의 스피치의 선택된 단어들에 적용될 수 있다. 예컨대, 사용자는 "What is the word for 'patent'?"라고 말함으로써(또는 다른 UI 상호작용을 통해) 번역을 제공하도록 시스템에 촉구할 수 있다. 시스템은 예컨대, 디스플레이(220) 상에 관심 외국어로 "patent"에 대한 단어를 제공함으로써 응답할 수 있다. 사용자는 외국어의 단어를 볼 수 있고 자신이 단어를 말할 수 있다. 단어 흐름 주석 시스템은 번역된 단어를 오디오로 제공하기 위해 오디오 증폭기(예컨대, 스피커(240))를 포함할 수 있다. 일 구현에서, 사용자는 예컨대, 이어폰의 형태의 오디오 증폭기만을 통해 이 오디오를 들을 수 있다. 다른 구현에서, 사용자뿐만 아니라 대화 파트너는 로드 스피커를 통해 이 오디오를 들을 수 있다.

맥락 정보에 기초하여 보조 정보를 제시하는 예들

[0168] 도 13d 및 도 13e는 맥락 정보에 기초하여 보조 정보를 제시하는 예들을 예시한다. 도 13a에 설명된 바와 같이, ARD를 착용한 사용자(도시되지 않음)는 여성(1302)을 지각할 수 있다. 여성(1302)은 사용자의 물리적 환경에 존재할 수 있다. 사용자는 여성(1302)과 비즈니스 관련 대화중에 있을 수 있다. 도 13d의 텍스트 박스(1342)에 도시된 바와 같이, 여성(1302)은 "'I wonder what will happen to Acme Corp.'s executives'"라고 말한다. 사용자의 ARD는 여성(1302)이 말한 이 문장을 캡처할 수 있다. 단어 흐름 주석 시스템(970)은 여성(1302)의 스피치를 분석하고 여성(1302)이 말한 문장에서 키워드를 검출할 수 있다. 이 예에서 사용자는 비즈니스와 관련 대화중에 있기 때문에, ARD는 "Acme Corp.'s executives"를 키워드로서 식별할 수 있는데, 그 이유는 이것이 비즈니스와 관련되기 때문이다. ARD가 키워드를 결정하면, ARD는 키워드와 연관된 보조 정보에 액세스하고 이를 리트리브할 수 있다. 예컨대, ARD는 데이터 리포지토리(예컨대, 원격 데이터 리포지토리(280))에서 검색을 수행하거나 공개 데이터베이스 검색을 수행할 수 있다. 텍스트 박스(1340)에 도시된 바와 같이, ARD는 Acme Corp.'s executives의 이름들을 리트리브한다. ARD는 박스(1340)에 도시된 바와 같이 보조 정보로서 이름을 디스플레이할 수 있다.

[0169] 도 13e에서, 여성(1302)은 "Acme Corp's stock has been performing tremendously!"라고 말한다. 사용자가 비즈니스 관련 대화중이기 때문에, ARD는 여성(1302)의 스피치에서 언급된 비즈니스 정보를 키워드로서 식별할 수 있다. 이 예에서, ARD는 ("tremendously"라는 단어 보다는) "Acme Corp.'s stock""라는 구문이 키워드임을 식별하는데, 그 이유는 이 구문이 비즈니스와 관련되기 때문이다. 따라서, ARD는 키워드와 연관된 보조 정보를 검색 또는 액세스할 수 있다. 텍스트 박스(1354)에 도시된 바와 같이, ARD는 가장 콘텐츠로서 현재 주식 가격, 증가의 퍼센티지, 등락 범위를 및 이전 마감 가격을 사용자에게 제시할 수 있다.

- [0170] [0170] 도 13f는 맥락 정보에 기초하여 관심 객체와 연관된 보조 정보를 제시하는 예를 예시한다. 사용자(도시되지 않음)는 ARD를 착용하면서 야구 게임을 시청할 수 있다. ARD는 시-쓰루(see-through) 디스플레이를 포함할 수 있기 때문에, 사용자는 (예컨대, 컴퓨터 스크린 또는 텔레비전과 같은) ARD 외부의 디스플레이 상에서 플레이어중이거나 실황의 야구 게임을 시청할 수 있다. 야구 게임은 또한 일부 상황들에서, ARD에 의해 렌더링될 수 있다. 사용자가 야구 게임을 시청하는 동안, 사용자는 플레이어 Paul(1362)이 타격에 임하고 해설자가 플레이어 Paul에 대해 이야기중임을 지각할 수 있다. 따라서, ARD는 플레이어 Paul(1362)이 관심 객체임을 결정할 수 있다. 이 예에서, ARD는 단독으로 또는 조합으로, 다양한 기술들을 사용하여 관심 객체를 결정할 수 있다. 예컨대, ARD는 해설자의 스피치에서 플레이어 Paul(1362)에 대해 빈번하게 언급한다는 것을 검출하고 그것은 플레이어 Paul이 관심 객체인 팩터임을 결정할 수 있다. 다른 예로서, ARD는 게임 진행 상황을 모니터링함으로써 어느 플레이어가 타격에 임하는지를 계속 추적하고, 관심 객체로서 타격에 임하는 플레이어를 식별할 수 있다. 또 다른 예로서, 사용자가 텔레비전 상에서 야구 게임을 시청하는 동안, ARD는 텔레비전 상에 디스플레이되는 바와 같은 야구 게임의 이미지들을 획득하기 위해 외향 이미징 시스템(464)을 사용할 수 있다. ARD는 하나 이상의 객체 인식기들(708)을 사용하여 이미지들을 파싱하고 획득된 이미지들로부터 타격에 임하는 플레이어를 식별할 수 있다.
- [0171] [0171] ARD가 플레이어 Paul(1362)이 관심 객체라고 결정하면, ARD는 플레이어 Paul(1362)과 관련된 정보를 리트리브하여 사용자에게 제시할 수 있다. 도 13f의 박스(1364)에 도시된 바와 같이, ARD는 게임 통계(예컨대, 참가한 게임들의 수 및 타격 평균), 필드의 포지션, 및 플레이어 Paul(1362)이 속한 팀과 같은 플레이어 Paul의 야구 관련 정보를 제시할 수 있다.
- [0172] [0172] 보조 정보는 예컨대, 통계, 상태, 설명들, 제품 정보, 관심 객체와 연관된 부가적인 세부 사항들/사실들 등과 같은 다양한 정보를 포함할 수 있다. 도 13g는 보조 정보가 제품 정보를 포함하는 예를 예시한다. 도 13f에서 설명된 바와 같이, ARD는 플레이어 Paul(1362)이 관심 객체임을 결정할 수 있다. (도 13g에 도시된 바와 같이) 플레이어 Paul(1362)의 통계를 디스플레이하는 것 외에도 또는 그에 대안적으로, ARD는 플레이어 Paul(1362)(또는 그의 팀)이 홍보한 제품(예컨대, 티셔츠(1340))을 결정하고, 주석 박스(1368)에 제품 정보(이를테면, 티셔츠(1340)의 이미지, 티셔츠(1340)의 가격(1342), 및 티셔츠(1340)가 구매될 수 있는 웹사이트(1344))를 디스플레이할 수 있다. 어떤 제품을 렌더링하고 제품 정보에 액세스할지 결정하기 위해, ARD는 플레이어 Paul(1362)이 나타나는 광고를 찾기 위해 공개 데이터베이스 또는 ARD와 연관된 데이터 리포지토리에서 검색을 수행할 수 있다. 따라서, ARD는 플레이어 Paul(1362)이 홍보하는 제품으로서 광고에서 식별된 제품들을 결정할 수 있다. 일부 상황들에서, 플레이어 Paul(1362)은 하나 초과의 제품을 홍보할 수 있다. ARD는 보조 정보로 제품들의 서브셋트를 고를 수 있다. 예컨대, ARD는 지난해 플레이어 Paul(1362)이 홍보한 제품들을 고를 수 있다. ARD는 또한 맥락 정보에 기초하여 제품들을 고를 수 있다. 예컨대, 사용자가 스포츠 이벤트를 시청하고 있기 때문에, ARD는 플레이어 Paul(1362)에 의해 홍보된 스포츠 용품들을 결정하여 이를 스포츠 용품들의 정보를 사용자에게 제시할 수 있다. ARD는 사용자로부터 구매 결정을 수락하도록 구성될 수 있다. 예컨대, 사용자가 티셔츠(1340)를 구매하기를 원하는 경우, 사용자는 예컨대 "buy the t-shirt"라고 말하거나, 또는 사용자 입력 디바이스를 작동시켜 웹사이트(1344)에 대한 링크를 선택할 수 있고, ARD는 사용자의 구매 선택을 수락하고 (예컨대, 웹사이트(1344)에 액세스하고 티셔츠(1340)에 대한 구매 주문을 자동으로 행함으로써) 트랜잭션을 완료하기 위한 조치들을 취할 수 있다.
- [0173] 스피치에 주석을 달기 위한 예시적인 프로세스
- [0174] [0173] 도 14a 및 도 14b는 스피치에 주석을 달는 예시적인 프로세스를 도시한다. 프로세싱 흐름(1400)은 도 2에 도시된 바와 같은 AR 시스템(200)에 의해 수행될 수 있다.
- [0175] [0174] 블록(1404)에서, 시스템의 오디오 센서는 주변 사운드들을 캡처할 수 있다. 주변 사운드들은 사용자의 환경 내의 2명 이상의 사람들의 스피치들을 포함할 수 있다. 스피치는 사용자의 스피치, 사람(다른 사용자의 스피치), 또는 오디오 시작적 콘텐츠(예컨대, 영화, 음악 등)로부터의 스피치일 수 있다. 주변 사운드들은 또한 노이즈와 같은 배경 사운드들을 포함할 수 있다.
- [0176] [0175] 블록(1408)에서, 시스템은 스피치의 존재를 검출하기 위해 주변 사운드들을 모니터링한다. 이러한 검출 프로세스는 로컬로(예컨대, 로컬 프로세싱 모듈(260)에 의해) 또는 원격으로(예컨대, 원격 프로세싱 모듈(270)에 의해) 수행될 수 있다. 시스템은 캡처된 주변 사운드들에서 단어들의 스트림을 식별함으로써 스피치의 존재를 검출할 수 있다. 시스템은 또한, 스피치의 존재를 나타내는 사용자 입력, 이를테면, 사용자의 포즈 또는 사용자 입력 디바이스(466)로부터의 입력에 대한 응답으로 검출을 개시하거나 수행할 수 있다.

- [0177] [0176] 블록(1412)에서, 시스템은 스피치가 검출되는지를 결정한다. 스피치가 검출되는 경우, 흐름은 블록(1416)으로 지속된다. 스피치가 검출되지 않는 경우, 흐름은 (도시된 바와 같이) 블록(1408) 또는 블록(1404)으로 리턴한다.
- [0178] [0177] 박스들(1404 내지 1412)에서의 동작들은, 시스템 및 오디오 센서들이 턴 온될 때, 예컨대, 단어 흐름 주석 기능이 인에이블될 때, 연속적으로 또는 주기적으로(예컨대, 샘플링 주파수로) 수행될 수 있다. 이러한 동작들은 (예컨대, 적시의 인터럽트(timed interrupt)에 의해 구동되는 배경 테스크들로서) 흐름도들(1400 및 1440) 내의 다른 블록들과 병렬로 수행될 수 있다. 이들은 예시의 목적으로 프로세싱 흐름 시퀀스에서 이산 박스들로서 도시된다. 그러나 이들은 예시된 시퀀스에 의해 제한되지 않는다. 위에서 설명된 예들 이외의 다수의 프로세싱 흐름들은 시스템 설계자의 재량에 따라 가능하다.
- [0179] [0178] 블록(1416)에서, AR 시스템은 검출된 스피치의 소스를 결정할 수 있다. 예컨대, 시스템이 사용자 자신의 스피치를 프로세싱하지 않도록 구성된 경우, 또는 시스템이 AR 시스템의 FOV 내의 사람들로부터의 스피치만을 프로세싱하도록 구성되는 경우(FOV 외부의 사람들로부터 오는 스피치는 폐기되고 추가로 프로세싱되지 않을 수 있음), 스피치의 소스는 관련이 있을 수 있다. 예컨대, AR 시스템은 사용자 전방의 사람들로부터의 스피치를 프로세싱하고 사용자 뒤의 사람들로부터의 스피치는 무시할 수 있다. 스피치의 소스는 또한, 스피치의 텍스트들 또는 보조 정보의 디스플레이 위치들을 결정하는 데 사용될 수 있다. 예컨대, AR 시스템은 말하는 사람 근처에 스피치와 연관된 보조 정보 또는 스피치의 전사물을 디스플레이할 수 있다.
- [0180] [0179] 블록(1420)에서, 시스템은 스피치 소스의 변화가 있었는지를 결정한다. 예컨대, 2명의 사용자들이 대화중일 때, 시스템은 문장이 이전 스피커 또는 다른 스피커로부터 온 것인지를 결정할 수 있다. 변화가 있었던 경우, 흐름은 블록(1424)을 통해 블록(1444)으로 지속된다. 스피치의 소스에 변화가 없었던 경우, 흐름은, 주변 사운드들을 캡처하는 것(블록 1404), 스피치의 존재를 검출하는 것(블록 1408)은 물론, 도 14b에 도시된 프로세싱 단계들을 포함할 수 있는 단어 흐름 주석 프로세싱을 지속하기 위해 블록(1428)으로 이동한다. 예컨대, 시스템이 동일한 스피커가 계속 말하고 있다고 결정하는 경우, 시스템은 사운드의 캡처 및 스피치의 검출을 지속하는 것 외에도, 블록(1456)으로부터 시작하는 기능들을 수행하도록 지속될 수 있다.
- [0181] [0180] 블록(1448)에서, 예시적인 시스템은 검출된 스피치가 사용자 자신의 스피치인지를 결정할 수 있다. 만약 그렇다면, 그리고 시스템이 사용자 자신의 스피치를 프로세싱하지 않도록 구성되는 경우, 프로세싱은 블록(1484)을 통해 블록(1404)으로 리턴될 수 있다. 검출된 스피치가 사용자 자신의 스피치가 아닌 경우, 흐름은 블록(1452)으로 지속된다. 블록(1452)에서, 시스템은 스피치가 사용자의 환경 내의 소스로부터의 것인지를 결정한다. 사용자의 환경은 예컨대, 내향 이미징 시스템(462), 외향 이미징 시스템(464) 또는 정적 카메라들에 의해 캡처된 정보를 통해 결정된 바와 같이, 사용자가 보고 있거나 대화하고 있는 사람을 포함할 수 있다. 외향 이미징 시스템(464)은 사용자에 대한 스피커의 포지션을 포함하는 정보를 제공할 수 있다. 내향 이미징 시스템은 사용자가 보고 있는 방향을 포함하는 정보를 제공할 수 있다. 둘 모두의 카메라들로부터의 정보를 사용하여, 단어 흐름 주석 시스템이 사용자가 보고 있는 사람을 결정할 수 있다.
- [0182] [0181] 사용자의 환경은, 다른 예로서, 음량(loudness) 임계치 초과의 스피치를 포함할 수 있다. 단어 흐름 주석 시스템은 구성 가능한 음량 임계치 예컨대, 일반 대화들에 대해 60-70 dBA, 속삭임들에 대해 30-40 dBA를 지원할 수 있으며, 여기서 dBA는 상이한 사운드 주파수들에 대한 인간의 귀의 다양한 감도를 고려하도록 스케일링된 A-가중 데시벨 측정이다. 음량 임계치의 또는 그 이상의 스피치가 사용자의 환경 내에 있을 수 있고, 아래에서 설명되는 바와 같이 단어 흐름 주석 시스템에 의해 프로세싱될 수 있다. 음량 임계치 미만의 스피치는 사용자의 목전의 환경 내에 있지 않을 수 있고(예컨대, 배경 대화 또는 떨어진 소스로부터의 스피치), 프로세싱되지 않을 수 있다.
- [0183] [0182] 사용자의 환경은 제 3 예로서, 특정 소스로부터의 스피치를 포함할 수 있다. 예컨대, 사용자는 사무실의 스피커폰을 통해 원격 사람과의 대화를 진행중일 수 있다. 스피커폰을 통해 대화를 진행중인 동안, 사용자가 문서 또는 컴퓨터 스크린을 볼 때와 같이 사용자가 스피커폰을 보고 있지 않을 수 있다. AR 시스템은 사용자가 사용자의 환경으로서 스피커폰을 지정할 수 있게 할 수 있어서, 오디오는, 사용자의 FOV가 스피커폰을 포함하지 않더라도 프로세싱될 것이다. 또 다른 예로서, 단어 흐름 주석 시스템의 관광객 사용자는, 관광객 사용자가 관광 사이트의 뷰를 즐기면서, 예컨대, 관광 가이드를 보지 않으면서(또는 가금씩만 봄), 관광 가이드의 스피치에 주석을 달기를 원할 수 있다. 단어 흐름 주석 시스템은 관광객 사용자가 사용자의 환경으로서 관광 가이드를 지정하게 할 수 있을 수 있다. 관광 가이드의 스피치는 관광객 사용자에 대해 외국어로 되어 있을 수 있다. 이러한 지정은 언어, 예컨대, 외국어에 기초할 수 있다.

- [0184] [0183] 단어 흐름 주석 시스템의 구현은 사용자의 변하는 환경을 용이하게 할 수 있다. 스피커폰 예에서, 시스템은 대화의 지속기간 동안 사용자의 환경으로서 스피커폰을 지정하도록 구성될 수 있고 대화가 끝난 후에 음량 임계치에 기초하여 사용자의 환경을 결정하도록 재구성될 수 있다.
- [0185] [0184] 다른 예로서, 사용자의 환경이 음량 임계치에 기초하여 결정되고 사용자의 FOV 내의 스피치 소스로 제한되는 경우, 구현은 사용자의 FOV 외부로부터의 임계치를 초과하는 캡처된 스피치의 일부 량을 버퍼링할 수 있다. 이 버퍼링된 정보는, 예컨대, 사용자가 사용자의 FOV 외부에 있던 캡처된 스피치의 소스를 보기 위해 눈을 돌린 경우 리트리브 및 프로세싱될 수 있다.
- [0186] [0185] 사용자의 환경 내의 스피치에 대해, 블록(1456)에서, 시스템은 위에서 설명된 바와 같이 스피치를 텍스트로 변환할 수 있다.
- [0187] [0186] 블록(1460)에서, 시스템은 변환된 텍스트에 대해 희귀 단어 검출을 수행할 수 있다. 블록(1464)에서, 시스템은 희귀 단어가 검출되는지를 결정한다. 이들 박스들(1460 및 1464)에서 수행되는 기능들은 선행 섹션들에서 상세히 설명된다.
- [0188] [0187] 희귀 단어가 검출되지 않는 경우, 흐름은 위에서의 블록(1428)의 설명과 유사한 방식으로 단어 흐름 주석 프로세싱을 지속하도록 블록(1488)으로 이동한다.
- [0189] [0188] 희귀 단어가 검출되는 경우, 블록(1468)에서, 시스템은 희귀 단어와 연관된 보조 정보를 리트리브할 수 있다. 블록(1472)에서, 시스템은 시스템의 AR 디스플레이 상에 리트리브된 보조 정보를 디스플레이할 수 있다. 블록(1476)에서, 시스템은 보조 정보를 해제하기 위한 조건을 검출할 수 있다. 그러한 조건이 검출되는 경우, 블록(1480)에서, 시스템은 보조 정보의 디스플레이를 해제하고 블록(1488)으로 진행될 수 있다. 이들 박스들에서 수행되는 기능들은 도 13a 내지 도 13g를 참조하여 위에서 설명되었다.
- [0190] [0189] 박스들(1404 내지 1412)에 관하여 위에서 언급된 것과 유사하게, 흐름도(1440)에서의 동작들은 다른 박스 흐름도들(1400 및 1440)과 병렬로 수행될 수 있다. 이들은 예시의 목적으로 프로세싱 흐름 시퀀스에서 이산 박스들로서 도시되지만, 이들은 예시된 시퀀스에 의해 제한되지 않는다. 예컨대, 시스템이(블록(1456)에서) 부가적인 스피치를 텍스트로 변환하고, (블록(1460)에서) 부가적인 스피치에 대해 희귀 단어 검출을 수행하거나, (블록(1468)에서) 다른 희귀 단어에 관한 보조 정보를 리트리브하는 동안, 시스템은 (블록(1472)에서) 희귀 단어에 대한 보조 정보를 디스플레이하고 있을 수 있다. 다른 예로서, 시스템은, 그것이 (블록(1468)에서) 이전에 검출된 희귀 단어에 대한 보조 정보를 리트리브하는 동안, (블록(1456)에서) 스피치를 텍스트로 변환할 수 있다. 시스템 설계자의 재량에 따라 다수의 다른 프로세싱 흐름들이 가능하다.
- [0191] 관심 객체에 주석을 다는 예시적인 프로세스
- [0192] [0190] 도 15는 관심 객체와 상호작용하는 예시적인 프로세스 및 관심 객체의 보조 정보를 예시한다. 예시적인 프로세스(1500)는 본원에서 설명된 AR 시스템에 의해 수행될 수 있다.
- [0193] [0191] 블록(1510)에서, AR 시스템은 사용자의 환경에서 관심 객체를 식별할 수 있다. 관심 객체는 사람, 객체(물리적 또는 가상), 또는 주변 사운드 등을 포함할 수 있다. AR 시스템은 외향 이미징 시스템(464), 내향 이미징 시스템(462), 또는 정적 카메라들로부터 획득된 이미지들을 사용하여 관심 객체를 식별할 수 있다. AR 시스템은 또한 마이크로폰(232)으로부터 획득되거나 다른 디바이스(예컨대, 다른 사용자의 ARD, 환경 내의 마이크로폰)로부터 수신된 오디오 정보 등을 분석함으로써 관심 객체를 식별할 수 있다. 도 13f를 참조하여 설명된 바와 같이, 오디오 정보는 관심 객체를 결정하기 위해 이미지들과 함께 사용될 수 있다. 일부 구현들에서, AR 시스템은 맥락 정보에 기초하여 관심 객체를 식별할 수 있다. 본원에서 설명된 바와 같이, 맥락 정보는 사용자, 관심 객체, 사용자의 환경, 사용자의 환경 내의 객체 또는 사람들 등과 연관된 정보를 포함할 수 있다. 다른 예로서, 사용자의 환경 내의 객체는 사용자의 경험으로 인해 사용자에게 희귀(그리고 이에 따라, 관심 객체일 수 있음)할 수 있지만, 다른 사용자에게 희귀하지 않을 수 있다. 티슈 포셉(tissue forcep)들은 사용자가 일상 생활에서 이들을 빈번하게 볼 수 없기 때문에 사용자에 대해 희귀할 수 있지만 이들은 의과의사에 대해 희귀하지 않을 수 있다.
- [0194] [0192] 블록(1520)에서, AR 시스템은 맥락 정보에 기초하여 관심 객체와 연관된 보조 정보를 결정할 수 있다. AR 시스템은 예컨대, 원격 데이터 리포지토리(280), AR 장치에 로컬인 데이터 리포지토리 또는 다른 컴퓨팅 디바이스에 의해 관리되는 데이터 저장소와 같은 하나 이상의 데이터 리포지토리들로부터의 보조 정보에 액세스할 수 있다. AR 시스템은 또한 검색 엔진으로 검색들을 수행함으로써 보조 정보를 결정할 수 있다. AR 시스템은

맥락 정보에 기초하여 보조 정보의 부분만을 제시할 수 있다. 예컨대, 사용자가 스포츠 이벤트를 시청하고 있을 때, 플레이어의 보조 정보는 다른 유형의 정보(예컨대, 플레이어의 가족 구성원에 관한 정보) 보다는, 플레이어의 스포츠 관련 정보(예컨대, 게임 통계)를 포함할 수 있다. 또한, AR 시스템은 맥락 정보에 기초하여 보조 정보에 액세스하거나 이를 검색할 수 있다. 예컨대, 사용자가 주식 거래와 관련된 대화중일 때, 사용자가 "Shop Corp."이라는 구문을 들으면, AR 시스템은 "Shop Corp."의 주식 정보에 액세스할 수 있다. 그러나 사용자가 쇼핑에 관한 대화중인 경우, AR 시스템이, AR 시스템이 "Shop Corp."이라는 구문을 검출하면, Shop Corp.에 의해 판매되는 제품들에 액세스할 수 있다.

[0195] [0193] 블록(1530)에서, AR 시스템은 보조 정보를 사용자에게 제시할 수 있다. 보조 정보는 사용자의 물리적 환경 위에 오버레이된 가장 콘텐츠로서 제시될 수 있다. 예컨대, 사용자가 물리적으로 사용자의 환경에 존재하는 다른 사용자와 대화중인 동안, AR 시스템이 다른 사용자가 말한 회귀 단어를 검출하면, AR 시스템은 다른 사용자 바로 옆에 말풍선을 보여줄 수 있다. 대화가 텔레프레전스 세션에서 발생하는 경우, AR 시스템은 다른 사용자의 가장 아바타 근처에 보조 정보를 제시할 수 있다.

[0196] [0194] 블록(1540)에서, AR 시스템은 선택적으로, 보조 정보와의 사용자 상호작용을 검출할 수 있다. 예컨대, 사용자는 보조 정보에 메모들/그림들을 부가할 수 있다. 사용자는 또한 보조 정보를 해제하거나, (예컨대, 보조 정보를 보여주는 사용자 인터페이스 엘리먼트를 확장하는 손 제스처를 사용하여) 부가적인 보조 정보를 요청하거나, 또는 보조 정보를 마킹/편집할 수 있다.

[0197] [0195] 블록(1550)에서, AR 시스템은 보조 정보와의 사용자의 상호작용에 기초하여 사용자 인터페이스 동작을 선택적으로 수행할 수 있다. 예컨대, AR 시스템은 보조 정보를 해제하기 위한 사용자 입력에 대한 응답으로 보조 정보를 숨길 수 있다. AR 시스템은 또한 블록(1540)에서 사용자 상호작용에 따라 사용자의 메모들을 디스플레이하거나 보조 정보의 부분을 강조할 수 있다.

스레드 식별의 예들

[0199] [0196] 보조 정보를 결정하는 것 외에도 또는 이에 대한 대안적으로, AR 시스템은 또한 키워드들을 사용하여 대화에서 스레드들을 식별할 수 있다. 도 16a 및 16b는 키워드들에 기초한 스레드 식별의 예를 예시한다. 이들 도면들에서, 2명의 사용자들(1328 및 1320)은 그들 각각의 ARD들(1322b 및 1322c)을 착용할 수 있다. 2명의 사용자들(1328 및 1320)은 대면 대화중일 수 있거나, 또는 텔레프레전스 세션에 있을 수 있다. 2명의 사용자들(1328 및 1320)은 또한, (이를테면, 예컨대, 텍스트 메시징 애플리케이션을 통해) 텍스트 메시징을 사용하여 서로 통신할 수 있다.

[0200] [0197] 말풍선들(1612, 1614, 및 1616, 및 1618)은 사용자들(1328 및 1320)이 말한 문장들의 시퀀스를 예시한다. 이러한 말풍선들은 사용자들(1328 및 1320) 간의 대화의 콘텐츠를 예시한다. AR 시스템은 시각적 콘텐츠로서 이들 말풍선들을 사용자들(1328 및 1320)에 제시하지 않을 수 있다. 말풍선들(1612 및 1616)의 콘텐츠는 사용자(1328)가 말한 것이고 말풍선들(1614 및 1618)의 콘텐츠는 사용자(1320)가 말한 것이다. 말풍선들(1612, 1614, 1616 및 1618)에 도시된 바와 같이, 2명의 사용자들 간의 대화는 2개의 토픽들을 포함하는데, 하나는 뉴럴 네트워크 프로그래밍과 관련된 것이고 다른 하나는 사용자(1328)의 앞으로의 아프리카로의 여행과 관련된다. 그러나 두 토픽들은 두 사람들의 스피치에서 함께 묶여 있다.

[0201] [0198] 선행 섹션들에서 설명된 바와 같이, AR 시스템(200)은 키워드들에 기초하여 2개의 별개의 토픽들이 대화에 존재함을 식별할 수 있다. 이 예에서, AR 시스템은 하나의 키워드를 "neural network"로서 식별하고 다른 키워드를 "Africa"로서 식별할 수 있다. "뉴럴 네트워크"라는 키워드는 알고리즘들 및 컴퓨터 프로그래밍과 관련된 반면, "아프리카"라는 키워드는 지리적 구역을 설명하기 때문에, AR 시스템(200)은 이들 2개의 키워드들이 서로 관련되지 않는다고 결정할 수 있고, 따라서, 각각의 키워드는 별개의 토픽과 연관될 수 있다. 일부 상황들에서, AR 시스템이 다수의 키워드들을 식별했지만, 키워드들 중 일부는 동일한 주제와 연관될 수 있다. 예컨대, AR 시스템은 하나의 키워드 "vacation" 및 다른 키워드 "beach"를 검출할 수 있다. AR 시스템은 그럼에도 불구하고, 사용자가 해변에서 휴가에 관해 이야기할 수 있기 때문에 하나의 주제를 식별할 수 있다.

[0202] [0199] 다수의 토픽들이 대화에서 검출되는 경우, AR 시스템은 주제를 분리하여 상이한 대화 스레드들을 생성할 수 있다. 도 16b에 도시된 바와 같이, 도 16a의 사용자(1328 및 1320) 사이의 대화는 2개의 스레드들(1630 및 1640)로 분리되며, 여기서 스레드(1630)는 뉴럴 네트워크 대화에 대응하고 스레드(1640)는 앞으로의 아프리카 여행에 대응한다. 말풍선들(1632 및 1642)은 말풍선(1612)에 대응하고; 말풍선들(1634 및 1644)은 말풍선(1614)에 대응하고; 말풍선들(1636 및 1646)은 말풍선(1616)에 대응하고; 말풍선들(1638 및 1648)은 말풍선

(1618)에 대응한다.

[0203] [0200] AR 시스템은 사용자(예컨대, 사용자(1320))에게 2개의 별개의 스레드들로서 대화를 제시할 수 있다. 예컨대, 스레드(1630)는 사용자의 FOV의 좌측 상에 디스플레이될 수 있는 반면, 스레드(1640)는 사용자의 FOV의 우측 상에 디스플레이될 수 있다. 텍스트에 대한 상이한 컬러들 또는 그래픽들, 말풍선들 등이 상이한 스레드들을 구별하는 데 사용될 수 있다. 유리하게는, 일부 실시예들에서, 대화를 상이한 스레드들로 분리함으로써, 사용자가 ARD를 착용하는 동안, 사용자는 각각의 토픽에 대한 대화 흐름을 지각할 수 있고, 다른 사용자들과의 사용자의 상호작용에서의 혼란들을 감소시킬 수 있다. 도 16b에 2개의 스레드들이 도시되지만, 이는 예시를 위한 것이며, 스레드들의 수는 2개보다 많을 수 있다. 일부 구현들에서, AR 시스템이, 예컨대 사용자의 눈 시선의 방향의 분석을 통해 사용자(1320)가 하나의 스레드(예컨대, 스레드(1630))를 보고 있다고 결정하는 경우, AR 시스템은 예컨대, 스레드의 밝기 또는 콘트라스트를 감소시키거나 다른 스레드를 전혀 디스플레이하지 않음으로써 다른 스레드(예컨대, 스레드(1640))의 외관을 덜 강조할 수 있다.

[0204] [0201] 일부 상황들에서, 하나(또는 다수)의 스레드는 다수의 서브-스레드들을 가질 수 있다. 각각의 서브-스레드들은 스레드로부터 유도된 토픽들을 포함할 수 있다. 예컨대, 사용자(1328)가 아프리카로의 앞으로의 여행에 대해 이야기하고 있는 동안, 사용자(1328)는 그가 아프리카에서 서핑할 예정이라고 말할 수 있다. 사용자(1320)는 서핑의 토픽을 지속하고 서핑 장비 및 위치들을 논의할 수 있다. AR 시스템은 사용자들이 서핑을 논의하고 있음을 검출하면, 서핑에 대한 스레드를 완전히 새롭게 하기보다는, 아프리카에서의 서핑에 대한 서브-스레드를 제공할 수 있다.

[0205] [0202] 도 13a 내지 도 13g 및 도 16a 내지 도 16b를 참조하여 설명된 예들이 말풍선들 및 텍스트 박스들을 사용하지만, 이를 말풍선들 및 텍스트 박스들은 하나의 가능한 사용자 인터페이스 예이다. AR 시스템은 현재의 보조 정보 및 텍스트 스레드들(예컨대, 자막들, 배너들 등)을 제공하기 위해 다양한 시각적 그래픽들을 사용할 수 있다.

스레드 식별의 예시적인 프로세스

[0207] [0203] 도 17은 스레드 식별의 예시적인 프로세스를 예시한다. 프로세스(1700)는 본원에서 설명된 AR 시스템(200)의 실시예들에 의해 수행될 수 있다.

[0208] [0204] 블록(1710)에서, AR 시스템은 (예컨대, 제 1 사람과 연관된) 제 1 오디오 스트림 및 (예컨대, 제 2 사람과 연관된) 제 2 오디오 스트림을 수신할 수 있다. 이 예에서, 제 1 사람 및 제 2 사람은 서로 대화할 수 있고, AR 시스템은 각각의 사람이 말한 오디오 스트림들을 캡처할 수 있다. 텔레프레전스의 다른 예에서, 제 1 사람은 원격 위치에 있는 제 2 사람에게 말하고 있다. 제 1 사람의 AR 시스템 상의 오디오 센서는 제 1 사람의 스피치를 검출하고 제 1 오디오 스트림을 생성할 수 있고, 제 2 사람의 AR 시스템(또는 전화) 상의 오디오 센서는 제 2 사람의 스피치를 검출하고 (예컨대, 네트워크(1290)를 통해) 제 2 오디오 스트림을 제 1 사람의 AR 시스템에 통신할 수 있다. 일부 실시예들에서, 제 1 사람 또는 제 2 사람 중 어느 하나가 AR 시스템의 사용자일 수 있다. 다른 실시예들에서, 사용자는 제 1 사람과 제 2 사람 사이의 대화를 관찰하고 있을 수 있다(예컨대, 제 1 사람과 제 2 사람 사이의 토론을 시청함). AR 시스템은 네트워크(1290)를 통해 오디오 스트림(또는 전사된 오디오 스트림)을 수신함으로써 또는 오디오 센서(232)를 사용하여 제 1 오디오 스트림 또는 제 2 오디오 스트림을 획득할 수 있다.

[0209] [0205] 블록(1720)에서, AR 시스템은 제 1 오디오 스트림 및 제 2 오디오 스트림을 파싱할 수 있다. 제 1 사람과 제 2 사람 간의 대화가 다수의 토픽들을 포함할 수 있기 때문에, AR 시스템은 제 1 토픽과 연관된 제 1 키워드 및 제 2 토픽과 연관된 제 2 키워드를 식별할 수 있다. 키워드들은 선행 섹션들에서 설명된 다양한 기술을 사용하여 식별될 수 있다. 일부 실시예들에서, 토픽은 하나 초과의 키워드를 포함할 수 있다.

[0210] [0206] 블록(1730)에서, AR 시스템은 상기 제 1 토픽과 연관된 제 1 스레드 및 상기 제 2 토픽과 연관된 제 2 스레드를 생성할 수 있다. AR 시스템은 제 1 토픽과 관련된 대화의 부분들을 제 1 스레드에 통합하고 제 2 토픽과 관련된 대화의 부분들을 제 2 스레드에 통합할 수 있다.

[0211] [0207] 블록(1740)에서, AR 시스템은 제 1 스레드 또는 제 2 스레드 중 적어도 하나가 ARD에 의해 렌더링되게 할 수 있다. 도 16b에 예시된 바와 같이, ARD는 사용자의 FOV 내의 상이한 디스플레이 위치들에 제 1 스레드 및 제 2 스레드를 렌더링할 수 있다. 일부 실시예들에서, AR 시스템은 2개의 스레드들 중 하나만을 제시할 수 있다. 예컨대, AR 시스템은 스레드를 제시할 수 있으며, 그의 연관된 토픽은 현재 제 1 사람 및 제 2 사람에 의해 논의되고 있다.

[0212] 부가적인 양상들

[0208] 제 1 양상에서, 증강 현실 시스템을 통해 희귀 단어에 관한 정보를 제공하기 위한 방법은, 오디오 센서를 포함하는 AR(augmented reality) 시스템의 제어 하에서, 오디오 센서를 통해 주변 사운드들을 캡처하는 단계; 주변 사운드들에서 스피치의 존재를 검출하는 단계; 검출된 스피치를 텍스트로 변환하는 단계; 변환된 텍스트에서 희귀 단어를 검출하는 단계; 희귀 단어와 연관된 보조 정보를 리트리브하는 단계; AR 시스템을 사용하여 보조 정보를 디스플레이하는 단계; 보조 정보의 디스플레이를 해제(dismissing)하기 위한 조건을 검출하는 단계; 및 보조 정보의 디스플레이를 해제하는 단계를 포함한다.

[0214] [0209] 제 2 양상에서, 제 1 양상의 방법에 있어서, 검출하는 단계로부터의 단계들은, 검출된 스피치가 AR 시스템의 사용자의 사용자의 환경 내에 있는 경우에 수행된다.

[0215] [0210] 제 3 양상에서, 제 1 양상 또는 제 3 양상 중 어느 한 양상의 방법에 있어서, 희귀 단어는 공통 단어 사전에서 제외된 단어이다.

[0216] [0211] 제 4 양상에서, 제 3 양상의 방법에 있어서, 공통 단어 사전은 희귀 단어를 포함하도록 업데이트된다.

[0217] [0212] 제 5 양상에서, 제 1 양상 내지 제 4 양상 중 어느 한 양상의 방법에 있어서, 희귀 단어는 AR 시스템의 사용자의 관심 언어로 구성 가능한 임계치보다 덜 빈번하게 발생하는 단어이다.

[0218] [0213] 제 6 양상에서, 제 1 양상 내지 제 5 양상 중 어느 한 양상의 방법에 있어서, 희귀 단어는 AR 시스템의 사용자에게 외국어로 된 단어이고, 보조 정보는 사용자의 지배적인 언어로의 희귀 단어의 번역을 포함한다.

[0219] [0214] 제 7 양상에서, 제 1 양상 내지 제 6 양상 중 어느 한 양상의 방법에 있어서, 검출된 스피치에서 희귀 단어의 맥락을 결정하는 단계 및 맥락에 기초하여 희귀 단어와 연관된 보조 정보를 리트리브하는 단계를 더 포함한다.

[0220] [0215] 제 8 양상에서, 제 7 양상의 방법에 있어서, 맥락을 결정하는 단계는 검출된 스피치의 희귀 단어에 시 간적으로 근접하게 나타나는 단어들에 적어도 부분적으로 기초한다.

[0221] [0216] 제 9 양상에서, 제 1 양상 내지 제 8 양상 중 어느 한 양상의 방법에 있어서, 보조 정보의 디스플레이를 해제하기 위한 조건은 사용자 인터페이스 상호작용에 기초한다.

[0222] [0217] 제 10 양상에서, 제 9 양상의 방법에 있어서, 사용자 인터페이스 상호작용은 AR 시스템의 사용자의 눈 움직임들에 기초한다.

[0223] [0218] 제 11 양상에서, 제 1 양상 내지 제 10 양상 중 어느 한 양상의 방법에 있어서, 보조 정보의 디스플레이를 해제하기 위한 조건은 시간의 지속기간에 기초한다.

[0224] [0219] 제 12 양상에서, 제 1 양상 내지 제 11 양상 중 어느 한 양상의 방법에 있어서, 보조 정보의 디스플레이를 해제하기 위한 조건은 다른 희귀 단어의 검출에 적어도 부분적으로 기초한다.

[0225] [0220] 제 13 양상에서, 제 1 양상 내지 제 12 양상 중 어느 한 양상의 방법에 있어서, 보조 정보의 해제된 디스플레이를 재-디스플레이하기 위한 조건을 검출하는 단계; 및 보조 정보의 해제된 디스플레이를 재-디스플레이하는 단계를 더 포함한다.

[0226] [0221] 제 14 양상에서, 희귀 단어에 관한 정보를 제공하기 위한 AR(augmented reality) 장치는, AR 디스플레이; 오디오 센서; 컴퓨터-실행 가능 명령들 및 데이터를 저장하도록 구성된 비-일시적 데이터 저장소; 및 데이터 저장소와 통신하는 하드웨어 프로세서를 포함하며, 컴퓨터 실행 가능 명령들은, 실행될 때, 프로세서로 하여금, 오디오 센서에 의해 검출된 오디오 정보를 수신하고; 수신된 오디오 정보에서 스피치를 검출하고; 검출된 스피치에서 희귀 단어를 검출하고; 희귀 단어에 기초하여 보조 정보를 리트리브하고; 그리고 AR 디스플레이를 사용하여, 리트리브된 보조 정보를 디스플레이하게 한다.

[0227] [0222] 제 15 양상에서, 제 14 양상의 장치에 있어서, 통신 네트워크를 통해 통신하기 위한 통신 컴포넌트를 더 포함하고, 보조 정보는 통신 네트워크를 통해 통신 컴포넌트를 경유하여 리트리브된다.

[0228] [0223] 제 16 양상에서, 제 15 양상의 장치에 있어서, 프로세서는 원격 프로세서가 희귀 단어를 검출하도록 통신 네트워크를 통해 검출된 스피치를 원격 프로세서에 송신함으로써 희귀 단어를 검출한다.

- [0229] [0224] 제 17 양상에서, 제 14 양상 내지 제 16 양상 중 어느 한 양상의 장치에 있어서, 오디오 센서는 복수의 마이크로폰들 또는 지향성 마이크로폰을 포함한다.
- [0230] [0225] 제 18 양상에서, 제 14 양상 내지 제 17 양상 중 어느 한 양상의 장치에 있어서, 프로세서는 추가로, 검출된 스피치의 소스를 결정하고; 그리고 희귀 단어를 검출하고, 검출된 스피치의 소스가 AR 장치의 사용자 이외의 다른 사람인 것으로 결정 시에 보조 정보를 리트리브 및 디스플레이하도록 구성된다.
- [0231] [0226] 제 19 양상에서, 제 14 양상 내지 제 18 양상 중 어느 한 양상의 장치에 있어서, 사용자 인터페이스 상 호작용들을 수신하기 위한 사용자 입력 디바이스를 더 포함한다.
- [0232] [0227] 제 20 양상에서, 희귀 단어에 관한 정보를 제공하기 위한 AR(augmented reality) 시스템은, AR 디스플레이; 오디오 센서; 컴퓨터-실행 가능 명령들 및 데이터를 저장하도록 구성된 데이터 저장소; 및 프로세서를 포함하는 제 1 웨어러블 AR 디바이스를 포함하고, 컴퓨터 실행 가능 명령들은, 실행될 때, 프로세서로 하여금, 오디오 센서에 의해 검출된 오디오 정보를 수신하고; 수신된 오디오 정보에서 스피치를 검출하고; 검출된 스피치에서 희귀 단어를 검출하고; 희귀 단어에 기초하여 보조 정보를 리트리브하고; 그리고 AR 디스플레이 상에서, 리트리브된 보조 정보를 디스플레이하게 한다.
- [0233] [0228] 제 21 양상에서, 제 20 양상의 시스템에 있어서, 내향 이미징 시스템을 더 포함한다.
- [0234] [0229] 제 22 양상에서, 제 20 양상 또는 제 21 양상의 시스템에 있어서, 외향 이미징 시스템을 더 포함한다.
- [0235] [0230] 제 23 양상에서, 제 22 양상의 시스템에 있어서, 프로세서는 제 1 웨어러블 AR 디바이스와 연관된 사용자의 환경에 대한 스피치의 소스의 위치를 결정하기 위해 외향 이미징 시스템으로부터의 이미지 정보를 사용한다.
- [0236] [0231] 제 24 양상에서, 제 22 양상의 시스템에 있어서, 프로세서는 추가로, 스피치의 소스의 위치를 결정하기 위해 외향 이미징 시스템으로부터의 이미지 정보를 사용하고; 그리고 AR 디스플레이를 통한 소스의 뷰를 모호하게 하지 않으면서, AR 디스플레이 상에, 스피치의 소스에 공간적으로 근접하게 보조 정보의 디스플레이를 포지셔닝하도록 구성된다.
- [0237] [0232] 제 25 양상에서, 제 20 양상 내지 제 24 양상 중 어느 한 양상의 시스템에 있어서, AR 디스플레이; 오디오 센서; 및 통신 네트워크를 통해 통신하도록 구성된 통신 컴포넌트를 포함하는 제 2 웨어러블 AR 디스플레이를 더 포함하고, 제 1 웨어러블 AR 디바이스는, 통신 네트워크를 통해 통신하도록 구성된 통신 컴포넌트를 더 포함하고, 제 1 웨어러블 AR 디바이스의 오디오 센서는 오디오 정보를 검출하고, 리트리브된 보조 정보는 제 2 웨어러블 AR 디바이스의 AR 디스플레이 상에 디스플레이된다.
- [0238] [0233] 제 26 양상에서, 오디오 스트림에서 스레드를 식별하는 방법은, 제 1 오디오 스트림 및 제 2 오디오 스트림을 수신하고, 제 1 토픽과 연관된 제 1 키워드 및 제 2 토픽과 연관된 제 2 오디오 스트림을 식별하기 위해 제 1 오디오 스트림 및 제 2 오디오 스트림을 파싱하는 단계; 제 1 토픽과 연관된 제 1 스레드 및 제 2 토픽과 연관된 제 2 스레드를 생성하는 단계; 및 제 1 스레드 또는 제 2 스레드 중 적어도 하나가 증강 현실 디바이스에 의해 랜더링에 하는 단계를 포함한다.
- [0239] [0234] 제 27 양상에서, 제 26 양상의 방법에 있어서, 제 1 오디오 스트림 또는 제 2 오디오 스트림은 사람 또는 시청각 콘텐츠 중 적어도 하나로부터의 것이다.
- [0240] [0235] 제 28 양상에서, 제 26 양상 또는 제 27 양상의 방법에 있어서, 제 1 오디오 스트림은 제 1 사람으로부터의 것이고 제 2 오디오 스트림은 제 2 사람으로부터의 것이다.
- [0241] [0236] 제 29 양상에서, 제 26 양상 내지 제 28 양상 중 어느 한 양상의 방법에 있어서, 제 1 토픽은 복수의 서브-토픽들을 더 포함한다.
- [0242] [0237] 제 30 양상에서, 증강 현실 시스템은 하드웨어 프로세서 및 증강 현실 디스플레이를 포함하고, 하드웨어 프로세서는 제 26 양상 내지 제 29 양상 중 임의의 하나를 수행하도록 프로그래밍된다.
- [0243] [0238] 제 31 양상에서, AR(augmented reality) 시스템은, AR 시스템의 사용자에게 가상 콘텐츠를 제시하도록 구성된 AR 디스플레이; 주변 사운드들을 캡처하도록 구성된 오디오 센서; 및 AR 디스플레이 및 오디오 센서와 통신하는 하드웨어 프로세서를 포함하고, 하드웨어 프로세서는, 오디오 센서에 의해 캡처된 주변 사운드들을 수신하고; 주변 사운드들에서 스피치(speech)의 존재를 검출하고; 검출된 스피치를 텍스트로 변환하고; 스피치에서 희귀 단어를 검출하고; 희귀 단어와 연관된 보조 정보를 리트리브하고; 그리고 AR 디스플레이가 리트리브된

보조 정보를 렌더링하게 하도록 프로그래밍된다.

[0244] 제 32 양상에서, 제 31 양상의 AR 시스템에 있어서, 희귀 단어는 공통 단어 사전에서 제외된 단어이다.

[0245] 제 33 양상에서, 제 31 양상 또는 제 32 양상의 AR 시스템에 있어서, 하드웨어 프로세서는 추가로, 보조 정보의 디스플레이를 해제(dismissing)하기 위한 조건을 검출하고; 그리고 검출된 조건에 대한 응답으로, AR 디스플레이가 보조 정보의 디스플레이를 해제하게 하도록 프로그래밍된다.

[0246] 제 34 양상에서, 제 33 양상의 AR 시스템에 있어서, 디스플레이를 해제하기 위한 조건을 검출하기 위해, 하드웨어 프로세서는, 임계 시간 기간이 경과되었음을 결정하는 것, 다른 희귀 단어를 검출하는 것, 사용자의 포즈를 검출하는 것, 또는 사용자 입력 디바이스로부터 입력을 수신하는 것 중 적어도 하나를 수행하도록 프로그래밍된다.

[0247] 제 35 양상에서, 제 31 양상 내지 제 34 양상 중 어느 한 양상의 AR 시스템에 있어서, 하드웨어 프로세서는 추가로, 검출된 스피치의 소스를 결정하고; 그리고 희귀 단어를 검출하고, 검출된 스피치의 소스가 AR 디스플레이의 사용자 이외의 다른 스피커와 연관된다고 결정 시에 보조 정보를 리트리브 및 디스플레이하도록 프로그래밍된다.

[0248] 제 36 양상에서, 제 35 양상의 AR 시스템에 있어서, 검출된 스피치의 소스는 하나 이상의 지향성 마이크로폰들을 포함하는 오디오 센서로부터 수집된 오디오 정보 또는 AR 시스템의 외향 이미징 시스템에 의해 획득된 이미지를 중 적어도 하나에 기초하여 결정된다.

[0249] 제 37 양상에서, 제 35 양상 또는 제 36 양상의 AR 시스템에 있어서, 검출된 스피치의 소스를 결정하기 위해, 하드웨어 프로세서는 스피커의 아이덴티티를 결정하기 위해 검출된 스피치 상에서 음성 인식을 수행하도록 프로그래밍된다.

[0250] 제 38 양상에서, 제 35 양상 내지 제 37 양상 중 어느 한 양상의 AR 시스템에 있어서, 검출된 스피치의 소스는, 사용자의 환경의 다른 컴퓨팅, 스피커와 연관된 다른 AR 디바이스 또는 사용자의 환경의 사람 중 적어도 하나를 포함한다.

[0251] 제 39 양상에서, 제 38 양상의 AR 시스템에 있어서, 사용자는 스피커와의 텔레프레전스 세션(telepresence session)에 있고, 검출된 스피치의 소스는 스피커와 연관된 다른 사용자 디바이스이고, 하드웨어 프로세서는 추가로, AR 디스플레이가 스피커의 가상 아바타를 렌더링하게 하고, 다른 사용자의 가상 아바타 근처에 보조 정보를 렌더링하게 하도록 프로그래밍된다.

[0252] 제 40 양상에서, 제 35 양상 내지 제 39 양상 중 어느 한 양상의 AR 시스템에 있어서, AR 디스플레이가 리트리브된 보조 정보를 렌더링하게 하기 위해, 하드웨어 프로세서는, 소스의 위치를 결정하고; 그리고 AR 디스플레이를 통해 소스의 뷰(view)를 모호하게 하지 않으면서 소스에 가까운 3D 공간 내의 포지션에 보조 정보를 렌더링하도록 프로그래밍된다.

[0253] 제 41 양상에서, 제 31 양상 내지 제 40 양상 중 어느 한 양상의 AR 시스템에 있어서, 보조 정보는 희귀 단어의 설명을 포함한다.

[0254] 제 42 양상에서, 제 31 양상 내지 제 41 양상 중 어느 한 양상의 AR 시스템에 있어서, 하드웨어 프로세서는 추가로, 보조 정보에 주석을 달기 위해 사용자로부터 표시를 수신하고; 보조 정보의 주석의 저장을 개시하고; 그리고 희귀 단어의 다른 존재를 검출하는 것에 대한 응답으로, AR 디스플레이가 주석 및 보조 정보를 렌더링하게 하도록 프로그래밍된다.

[0255] 제 43 양상에서, 방법은, 하드웨어 프로세서 및 사용자의 환경에 가상 콘텐츠를 제시하도록 구성된 AR(augmented reality) 디스플레이를 포함하는 AR 디바이스의 제어 하에, AR 디바이스의 사용자의 환경을 모니터링하는 단계; 사용자 또는 환경 중 적어도 하나와 연관된 맥락 정보(contextual information)에 기초하여 환경에서 관심 객체의 존재를 검출하는 단계; 관심 객체에 대한 보조 정보를 리트리브하는 단계; 사용자의 시야에서 보조 정보의 디스플레이 포지션을 결정하는 단계 – 시야는 정해진 시간에 사용자가 지각할 수 있는 환경의 부분을 포함함 – ; 및 AR 디스플레이가 디스플레이 포지션에 보조 정보를 렌더링하게 하는 단계를 포함한다.

[0256] 제 44 양상에서, 제 43 양상의 방법에 있어서, 관심 객체는 키워드, 물리적 객체 또는 사람을 포함한다.

[0257] 제 45 양상에서, 제 44 양상의 방법에 있어서, 관심 객체가 키워드를 포함하는 경우, 관심 객체의 존재

를 검출하는 단계는, 환경의 주변 사운드들로부터 스피치를 검출하는 단계; 및 희귀 단어를 식별하기 위해 스피치를 파싱(parsing)하는 단계를 포함한다.

[0258] [0253] 제 46 양상에서, 제 43 양상 또는 제 45 양상의 방법에 있어서, 환경을 모니터링하는 단계는, 환경의 주변 사운드들을 캡처 및 분석하는 단계; 또는 환경의 이미지들을 획득하는 단계 중 적어도 하나를 포함한다.

[0259] [0254] 제 47 양상에서, 제 43 양상 내지 제 46 양상 중 어느 한 양상의 방법에 있어서, 맥락 정보는 사용자의 과거 경험과 연관되고, 관심 객체는 사용자가 거의 대면하지 않는 객체를 포함한다.

[0260] [0255] 제 48 양상에서, 제 43 양상 내지 제 47 양상 중 어느 한 양상의 방법에 있어서, 보조 정보는 관심 객체와 연관된 제품 정보 또는 관심 객체의 설명 텍스트 중 적어도 하나를 포함한다.

[0261] [0256] 제 49 양상에서, 제 43 양상 내지 제 48 양상 중 어느 한 양상의 방법에 있어서, 보조 정보에 주석을 달기 위해 사용자로부터 표시를 수신하는 단계; 보조 정보의 주석의 저장을 개시하는 단계; 및 관심 객체의 다른 존재를 검출하는 것에 대한 응답으로, AR 디스플레이가 주석 및 보조 정보를 렌더링하게 하는 단계를 더 포함한다.

[0262] [0257] 제 50 양상에서, 제 43 양상 내지 제 49 양상 중 어느 한 양상의 방법에 있어서, 보조 정보의 디스플레이 포지션을 결정하는 단계는, 관심 객체의 위치를 결정하는 단계; 및 AR 디스플레이를 통해, 관심 객체의 뷰를 모호하게 하지 않으면서 관심 객체에 가까운 환경 내의 디스플레이 포지션을, 포지션으로서 식별하는 단계를 포함한다.

다른 고려사항들

[0264] [0258] 본원에서 설명되고 그리고/또는 첨부 도면들에 도시되는 프로세스들, 방법들 및 알고리즘들 각각은 하나 이상의 물리적 컴퓨팅 시스템들, 하드웨어 컴퓨터 프로세서들, 애플리케이션-특정 회로 및/또는 특유 및 특정 컴퓨터 명령들을 실행하도록 구성된 전자 하드웨어에 의해 실행되는 코드 모듈들로 구현되고, 이 코드 모듈들에 의해 완전히 또는 부분적으로 자동화될 수 있다. 예컨대, 컴퓨팅 시스템들은 특정 컴퓨터 명령들로 프로그래밍된 범용 컴퓨터들(예컨대, 서버들) 또는 특수 목적 컴퓨터들, 특수 목적 회로 등을 포함할 수 있다. 코드 모듈은 실행 가능 프로그램으로 컴파일되어 링크되거나, 동적 링크 라이브러리에 설치될 수 있거나, 또는 인터프리팅된 프로그래밍 언어로 작성될 수 있다. 일부 구현들에서, 특정한 동작들 및 방법들은, 주어진 기능에 특정한 회로에 의해 수행될 수 있다.

[0265] [0259] 추가로, 본 개시내용의 기능성의 소정의 구현들은 충분히 수학적으로, 계산상으로 또는 기술적으로 복잡하여, (적절한 전문화된 실행 가능한 명령들을 활용하는) 주문형 하드웨어 또는 하나 이상의 물리적 컴퓨팅 디바이스들은 예컨대, 수반되는 계산들의 양(volume) 또는 복잡성으로 인해 또는 실질적으로 실시간으로 결과들을 제공하기 위해 그 기능성들을 수행할 필요가 있을 수 있다. 예컨대, 애니메이션들 또는 비디오는 다수의 프레임들(각각의 프레임은 수백만 개의 픽셀들을 가짐)을 포함할 수 있고, 상업적으로 합리적인 시간량 내에 원하는 이미지 프로세싱 태스크 또는 애플리케이션을 제공하기 위해, 특별히 프로그래밍된 컴퓨터 하드웨어가 비디오 데이터를 프로세싱할 필요가 있다.

[0266] [0260] 코드 모듈들 또는 임의의 유형의 데이터는, 임의의 유형의 비-일시적인 컴퓨터-판독 가능 매체, 이를테면, 하드 드라이브, 솔리드 스테이트 메모리, RAM(random access memory), ROM(read only memory), 광학 디스크, 휴발성 또는 비-휘발성 저장소, 이들의 조합들 등을 포함하는 물리적 컴퓨터 저장소 상에 저장될 수 있다. 방법들 및 모듈들(또는 데이터)은 또한, 생성된 데이터 신호들로서(예컨대, 반송파 또는 다른 아날로그 또는 디지털 전파 신호의 일부로서) 무선-기반 및 유선/케이블-기반 매체들을 포함하는 다양한 컴퓨터-판독 가능 송신 매체들 상에서 송신될 수 있고, (예컨대, 단일 또는 멀티플렉싱된 아날로그 신호의 일부로서, 또는 다수의 이산 디지털 패킷들 또는 프레임들로서) 다양한 형태들을 취할 수 있다. 개시된 프로세스들 또는 프로세스 단계들의 결과들은 임의의 유형의 비-일시적인 유형의(tangible) 컴퓨터 저장소에 지속적으로 또는 다른 방식으로 저장될 수 있거나, 또는 컴퓨터-판독 가능 송신 매체를 통해 통신될 수 있다.

[0267] [0261] 본원에서 설명되고 그리고/또는 첨부된 도면들에 도시되는 흐름도들에서의 임의의 프로세스들, 블록들, 상태들, 단계들 또는 기능성들은 프로세스의 단계들 또는 (예컨대, 논리적 또는 산술적) 특정 기능들을 구현하기 위한 하나 이상의 실행 가능 명령들을 포함하는 코드 모듈들, 세그먼트들 또는 코드 부분들을 임재적으로 나타내는 것으로 이해되어야 한다. 다양한 프로세스들, 블록들, 상태들, 단계들 또는 기능성들은 본원에서 제공된 예시적인 예들에서 조합되거나, 재배열되거나, 이들에 부가되거나, 이들로부터 제거되거나, 수정되거나, 또

는 다른 방식으로 변할 수 있다. 일부 실시예들에서, 부가적인 또는 상이한 컴퓨팅 시스템들 또는 코드 모듈들은 본원에서 설명된 기능성을 중 일부 또는 전부를 수행할 수 있다. 본원에 설명된 방법들 및 프로세스들은 또한 임의의 특정 시퀀스로 제한되지 않고, 그에 관련된 블록들, 단계들 또는 상태들은 적절한 다른 시퀀스들로, 예컨대, 직렬로, 병렬로 또는 일부 다른 방식으로 수행될 수 있다. 태스크들 또는 이벤트들은 개시된 예시적인 실시예들에 부가되거나 그로부터 제거될 수 있다. 또한, 본원에서 설명된 구현들에서의 다양한 시스템 컴포넌트들의 분리는 예시 목적들을 위한 것이며, 모든 구현들에서 이러한 분리를 요구하는 것으로 이해되어서는 안 된다. 설명된 프로그램 컴포넌트들, 방법들 및 시스템들은 일반적으로 단일 컴퓨터 제품에 함께 통합되거나 다수의 컴퓨터 제품들로 패키징될 수 있다는 것이 이해되어야 한다. 다수의 구현 변동들이 가능하다.

[0268] 프로세스들, 방법들 및 시스템들은 네트워크(또는 분산형) 컴퓨팅 환경에서 구현될 수 있다. 네트워크 환경들은, 전사적(enterprise-wide) 컴퓨터 네트워크들, 인트라넷들, LAN(Local Area Network)들, WAN(Wide Area Network)들, PAN(Personal Area Network)들, 클라우드 컴퓨팅 네트워크들, 크라우드-소스드(crowd-sourced) 컴퓨팅 네트워크들, 인터넷 및 월드 와이드 웹(World Wide Web)을 포함한다. 네트워크들은 유선 또는 무선 네트워크들 또는 임의의 다른 유형의 통신 네트워크일 수 있다.

[0269] 본 개시내용의 시스템들 및 방법들 각각은 몇몇 혁신적인 양상들을 가지며, 그 양상들 중 어떠한 단일 양상도 본원에서 개시된 바람직한 속성들을 전적으로 담당하거나 이를 위해 요구되지 않는다. 위에서 설명된 다양한 특징들 및 프로세스들은 서로 독립적으로 사용될 수 있거나, 또는 다양한 방식들로 조합될 수 있다. 모든 가능한 조합들 및 서브조합들은 본 개시내용의 범위 내에 속하는 것으로 의도된다. 본 개시내용에서 설명된 구현들에 대한 다양한 수정들은 당업자들에게 자명할 수 있으며, 본원에서 정의된 일반적인 원리들은 본 개시내용의 사상 또는 범위를 벗어나지 않으면서 다른 구현들에 적용될 수 있다. 따라서, 청구항들은 본 명세서에 도시된 구현들 또는 실시예들로 제한되도록 의도되는 것이 아니라, 본 명세서에 개시된 본 개시내용, 원리들 및 신규한 특성들과 일치하는 가장 넓은 범위에 부합할 것이다.

[0270] 별개의 구현들 또는 실시예들의 맥락에서 본 명세서에 설명된 소정의 특징들은 또한, 단일 구현 또는 실시예의 조합으로 구현될 수 있다. 대조적으로, 단일 구현 또는 실시예의 맥락에서 설명된 다양한 특징들은 또한, 별개로 다수의 구현들 또는 실시예들에서 또는 임의의 적절한 서브조합으로 구현될 수 있다. 더욱이, 특징들이 소정의 조합들로 작용하는 것으로 위에서 설명되고 심지어 초기에 이와 같이 청구될 수 있지만, 일부 경우들에서, 청구된 조합으로부터의 하나 이상의 특징들은 그 조합으로부터 제거될 수 있고, 청구된 조합은 서브조합 또는 서브조합의 변동에 관련될 수 있다. 단일 특징 또는 특징들의 그룹이 각각의 그리고 모든 각각의 실시예에 필요하거나 필수적인 것은 아니다.

[0271] 구체적으로 달리 언급되지 않거나 또는 사용된 맥락 내에서 달리 이해되지 않으면, 본원에서 사용된 조건어, 이를테면, 다른 것들 중에서도, "할 수 있다(can, could, might, may)", "예컨대" 등은 일반적으로, 소정의 실시예들이 소정의 특징들, 엘리먼트들, 및/또는 단계들을 포함하지만 다른 실시예들은 이들을 포함하지 않는다는 것을 전달하도록 의도된다. 따라서, 그러한 조건어는 일반적으로, 특징들, 엘리먼트들, 및/또는 단계들이 하나 이상의 실시예들을 위해 어떤 식으로든 요구된다는 것을, 또는 하나 이상의 실시예들이, 저자 입력 또는 프롬프팅(prompting)을 이용하거나 또는 그러한 것을 이용함이 없이, 이를 특징들, 엘리먼트들, 및/또는 단계들이 임의의 특정 실시예에 포함되는지 또는 임의의 특정 실시예들에서 수행되어야 하는지를 판단하기 위한 로직을 반드시 포함한다는 것을 암시하도록 의도되진 않는다. "포함하는(comprising, including), "갖는(having)" 등의 용어들은 동의어이며, 오픈-엔디드(open-ended) 방식으로 포괄적으로 사용되며, 부가적인 엘리먼트들, 특징들, 행동들, 동작들 등을 배제하지 않는다. 또한, "또는"이라는 용어는 (그의 배타적인 의미가 아니라) 그의 포괄적인 의미로 사용되어서, 예컨대, 리스트의 엘리먼트들을 연결하기 위해 사용될 때, "또는"이라는 용어는 리스트 내의 엘리먼트들 중 하나, 일부, 또는 전부를 의미한다. 또한, 본 명세서 및 첨부된 청구항들에서 사용된 바와 같은 단수 표현은 달리 특정되지 않는 한 "하나 이상" 또는 "적어도 하나"를 의미하는 것으로 해석될 것이다.

[0272] 본원에서 사용된 바와 같이, 리스트의 아이템들 "중 적어도 하나"를 지칭하는 어구는 단일 멤버들을 포함하여 그 아이템들의 임의의 조합을 지칭한다. 예로서, "A, B 또는 C 중 적어도 하나"는 A; B; C; A와 B; A와 C; B와 C; 그리고 A와 B와 C를 커버하는 것으로 의도된다. 특정하게 다르게 언급되지 않으면, 구문 "X, Y 또는 Z 중 적어도 하나"와 같은 접속어는, 아이템, 용어 등이 X, Y 또는 Z 중 적어도 하나일 수 있다는 것을 전달하기 위해 일반적으로 사용되는 맥락으로 달리 이해된다. 따라서, 이러한 접속어는 일반적으로, 소정의 실시예들이 X 중 적어도 하나, Y 중 적어도 하나 및 Z 중 적어도 하나가 각각 존재할 것을 요구하는 것을 암시하는 것으로

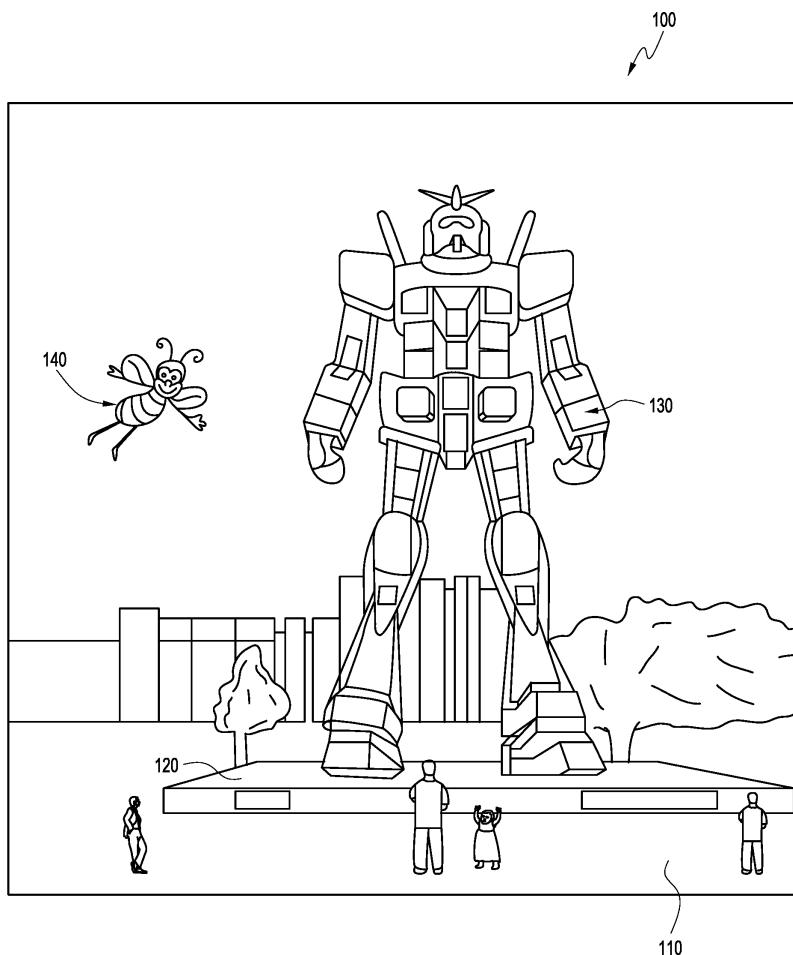
로 의도되지 않는다.

[0273]

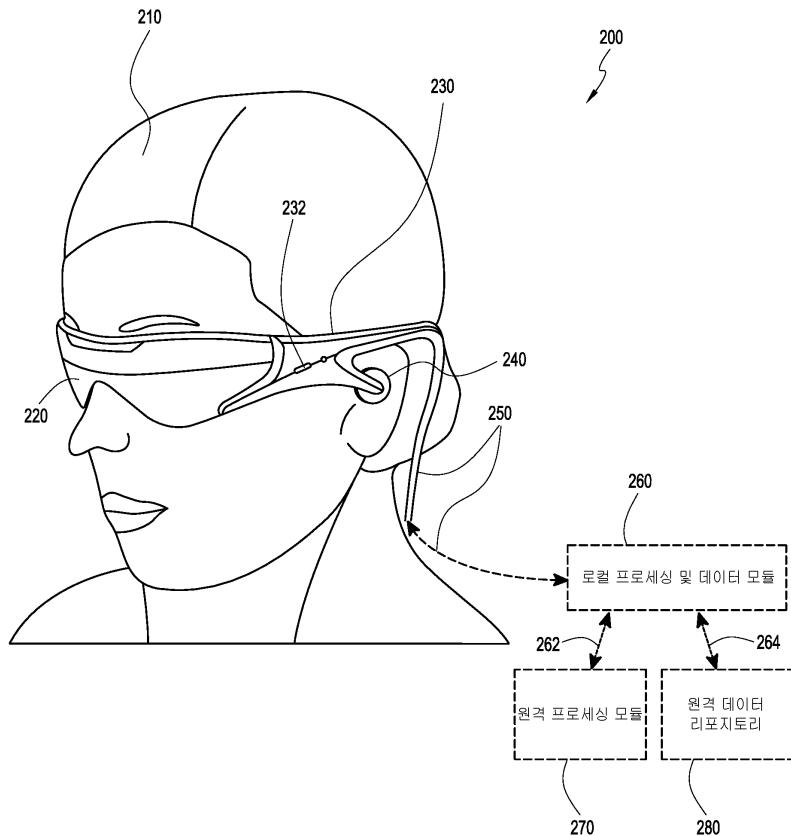
[0267] 유사하게, 동작들이 특정한 순서로 도면들에 도시될 수 있지만, 원하는 결과들을 달성하기 위해, 그러한 동작들이 도시된 특정한 순서 또는 순차적인 순서로 수행될 필요가 없거나, 모든 예시된 동작들이 수행될 필요가 없다는 것이 인지될 것이다. 추가로, 도면들은 흐름도의 형태로 둘 이상의 예시적인 프로세스들을 개략적으로 도시할 수 있다. 그러나, 도시되지 않은 다른 동작들이, 개략적으로 예시된 예시적인 방법들 및 프로세스들에 통합될 수 있다. 예컨대, 하나 이상의 부가적인 동작들은, 예시된 동작들 중 임의의 동작 이전, 이후, 그들과 동시에, 또는 그들 사이에서 수행될 수 있다. 부가적으로, 동작들은 다른 구현들에서 재배열되거나 재순서화될 수 있다. 소정의 환경들에서, 멀티태스킹 및 병렬 프로세싱이 유리할 수 있다. 또한, 위에서 설명된 구현에서의 다양한 시스템 컴포넌트들의 분리는 모든 구현들에서 그러한 분리를 요구하는 것으로서 이해되지는 않아야 하고, 그리고 설명된 프로그램 컴포넌트들 및 시스템들이 일반적으로, 단일 소프트웨어 제품으로 함께 통합되거나 다수의 소프트웨어 제품들로 패키징될 수 있다는 것이 이해되어야 한다. 부가적으로, 다른 구현들은 다음의 청구항들의 범위 내에 있다. 일부 경우들에서, 청구항들에서 열거된 액션들은, 상이한 순서로 수행될 수 있으며, 그럼에도 불구하고 원하는 결과들을 달성할 수 있다.

도면

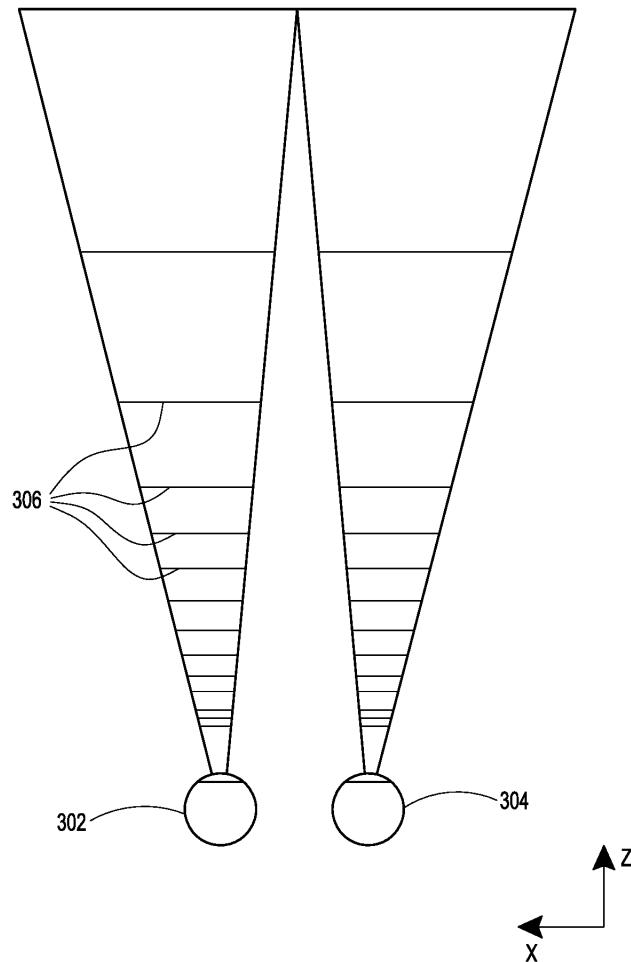
도면1



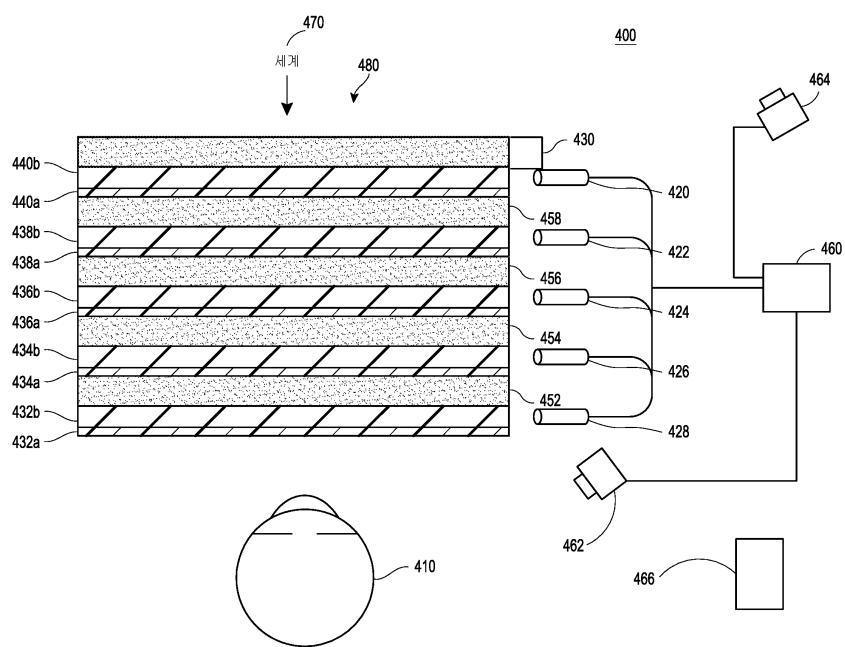
도면2



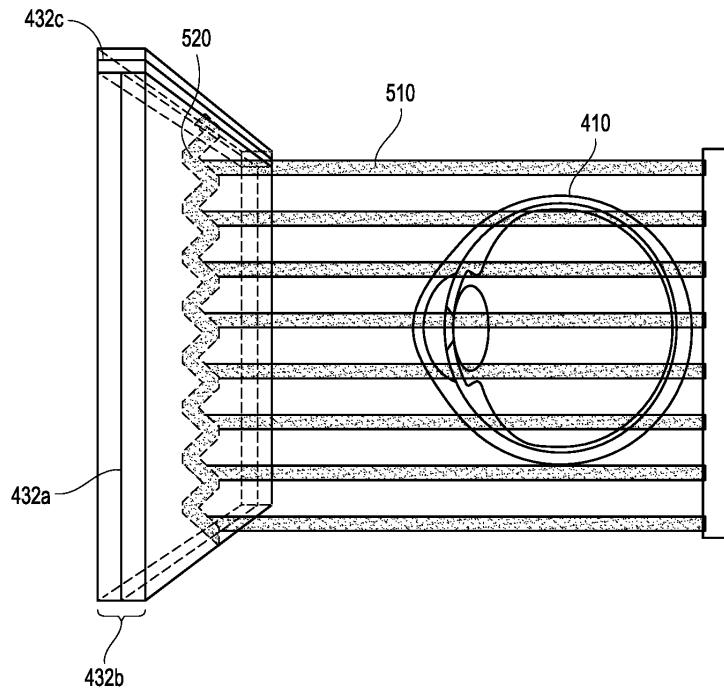
도면3



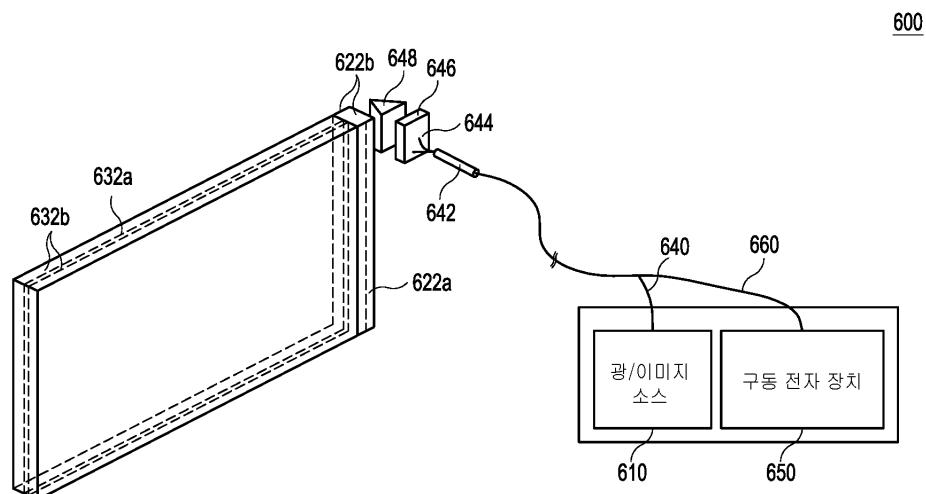
도면4



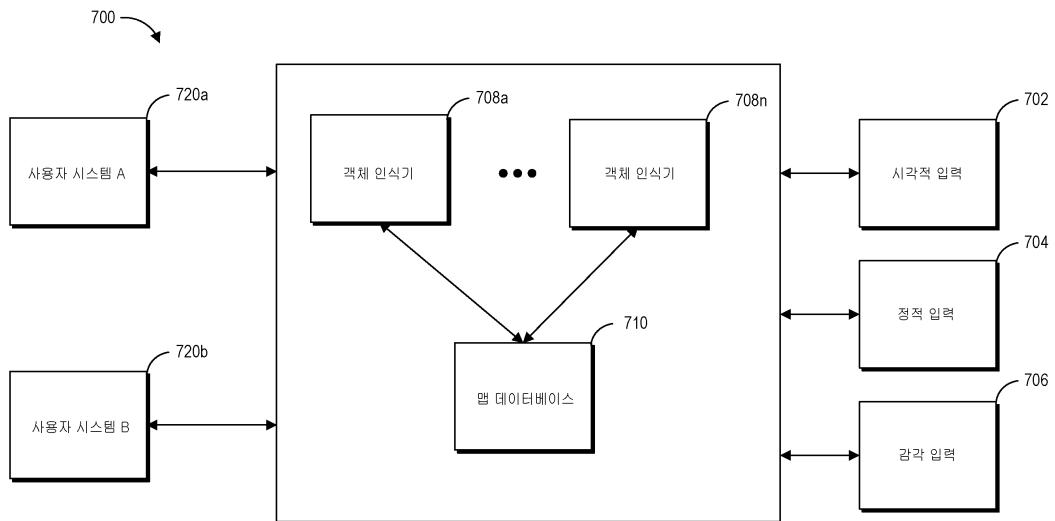
도면5



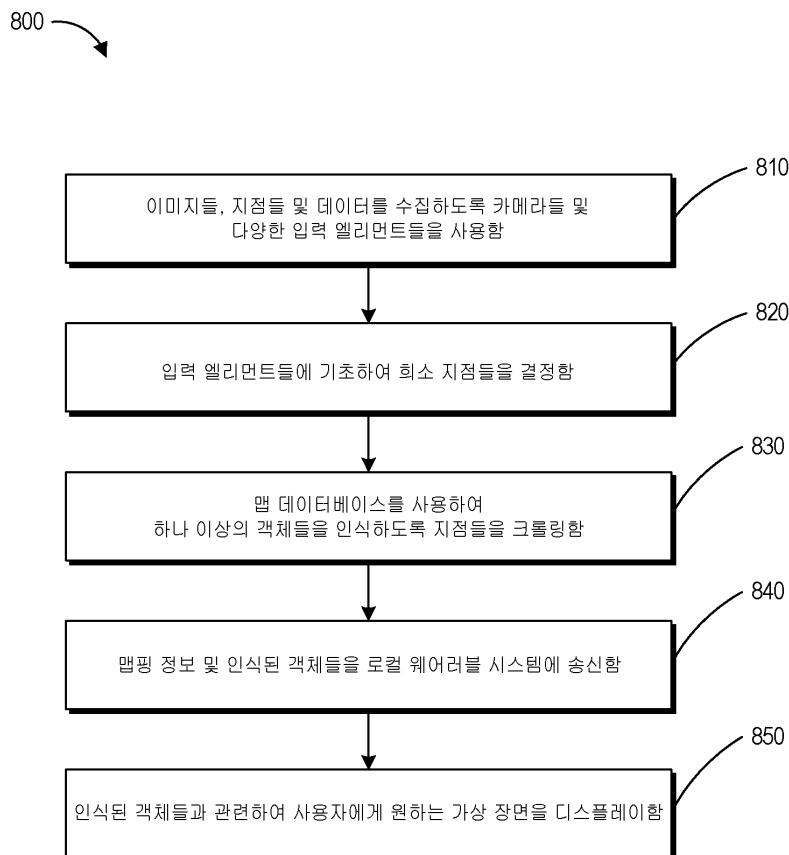
도면6



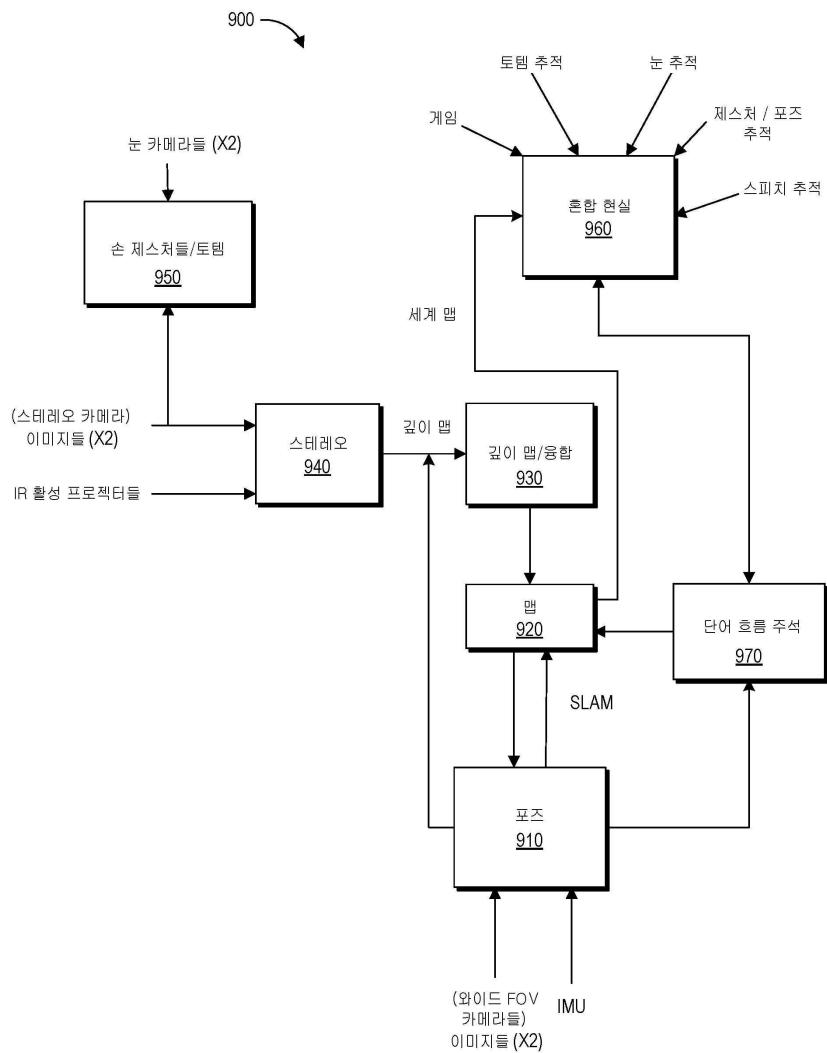
도면7



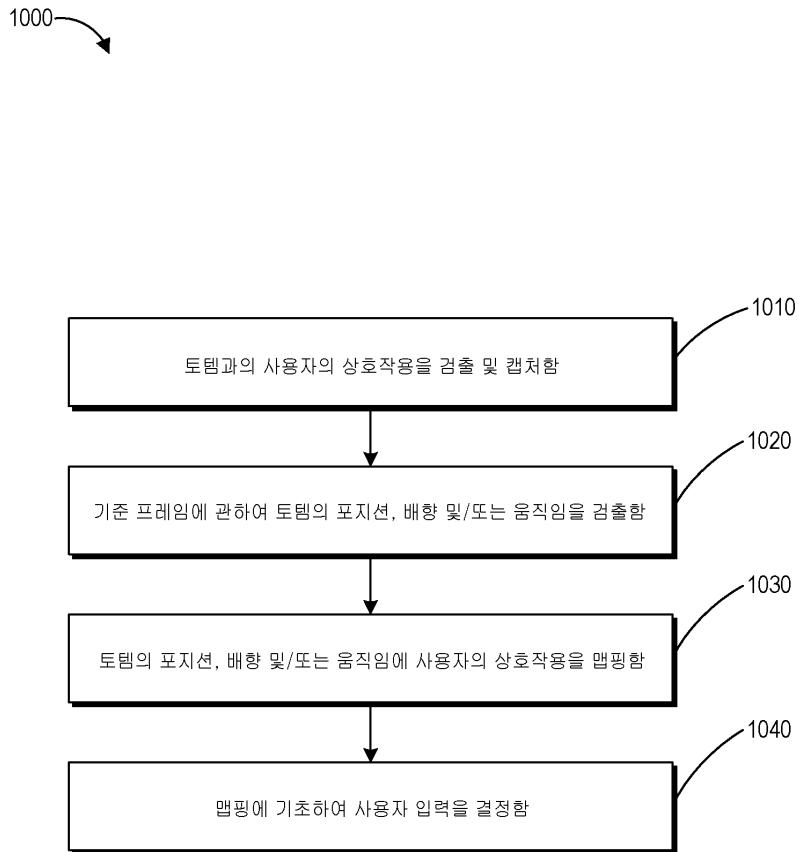
도면8



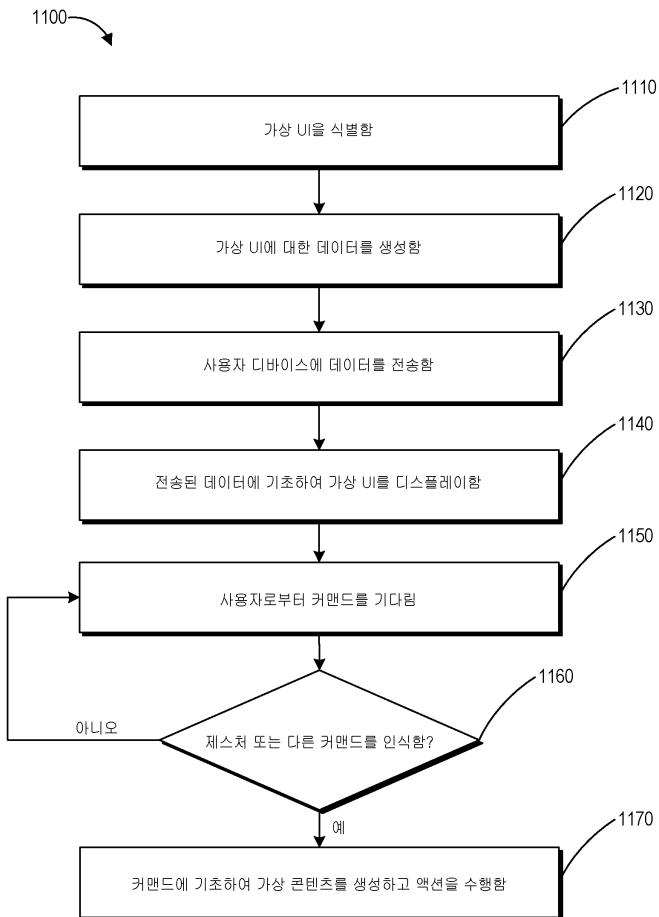
도면9



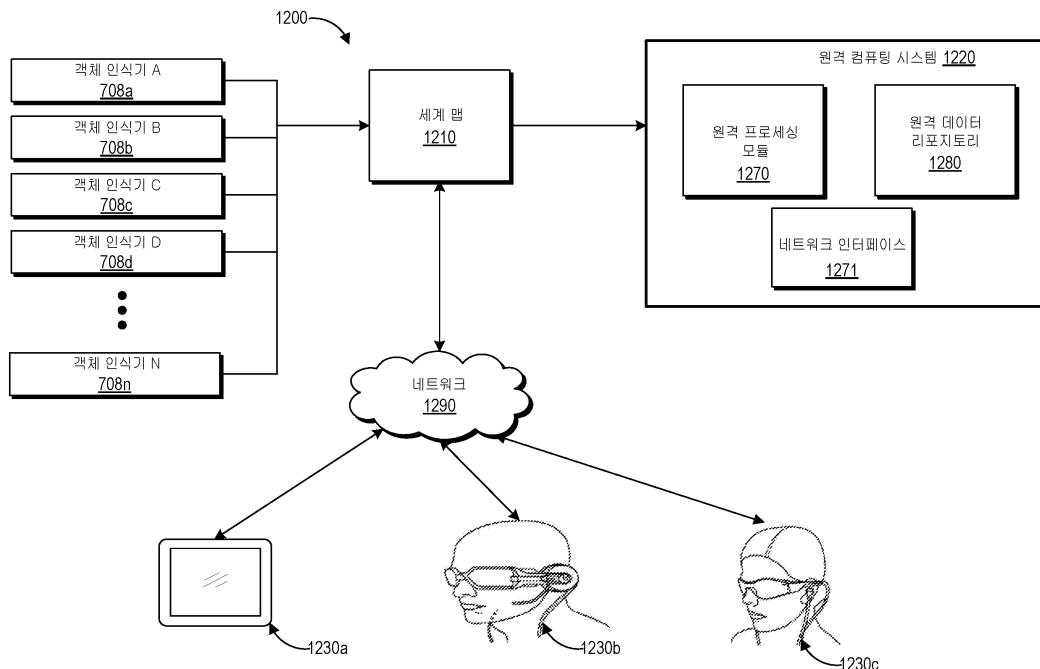
도면10



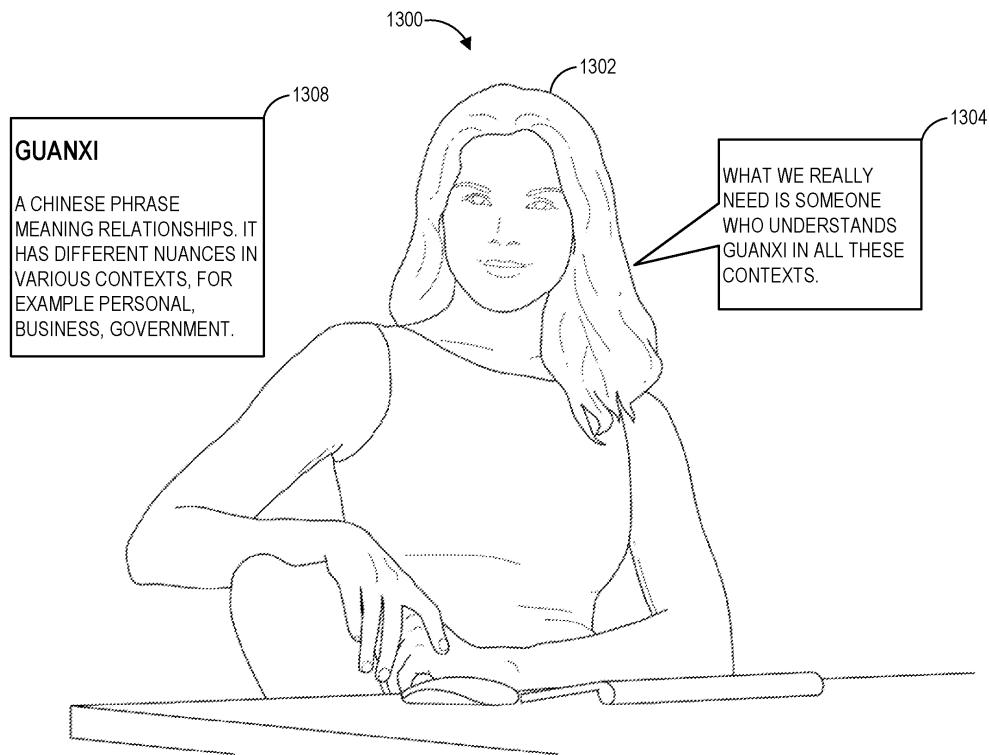
도면11



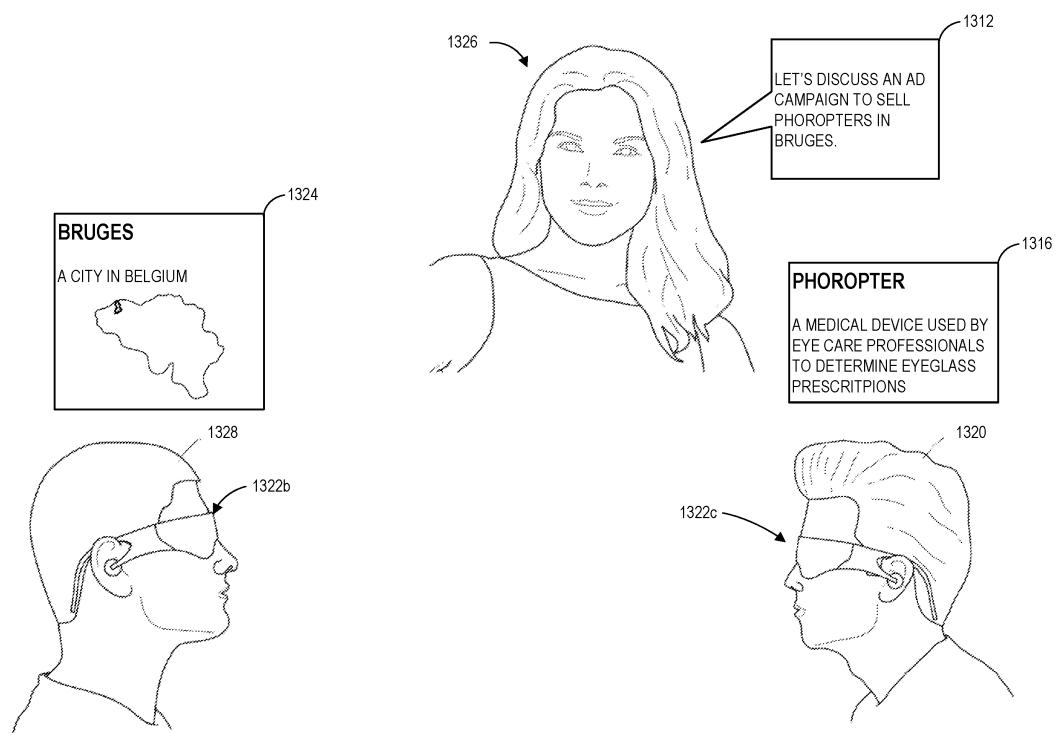
도면12



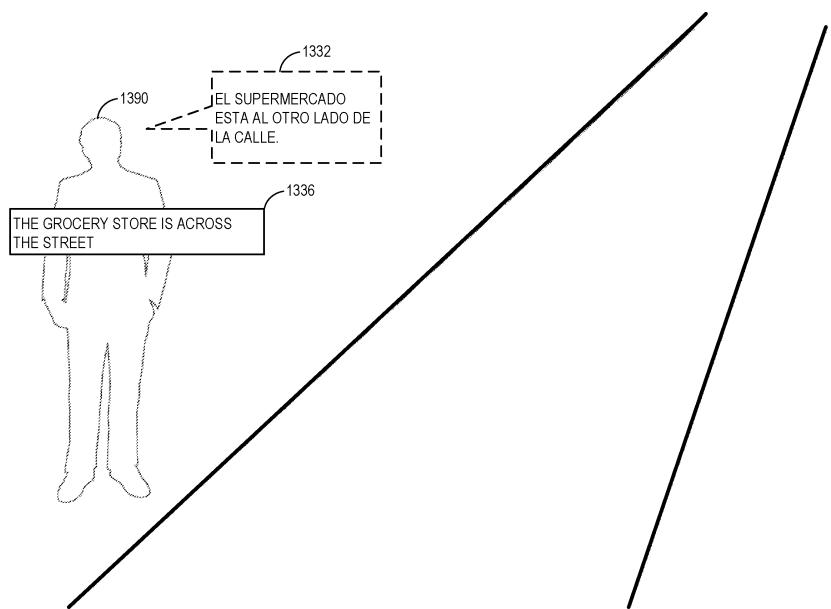
도면 13a



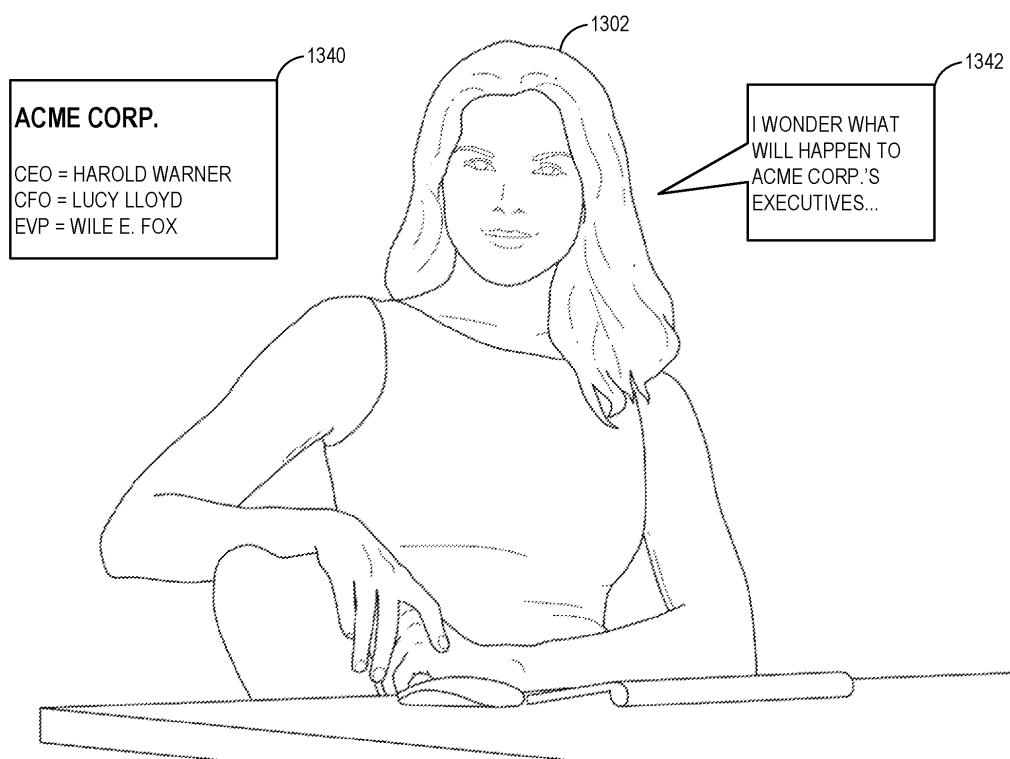
도면 13b



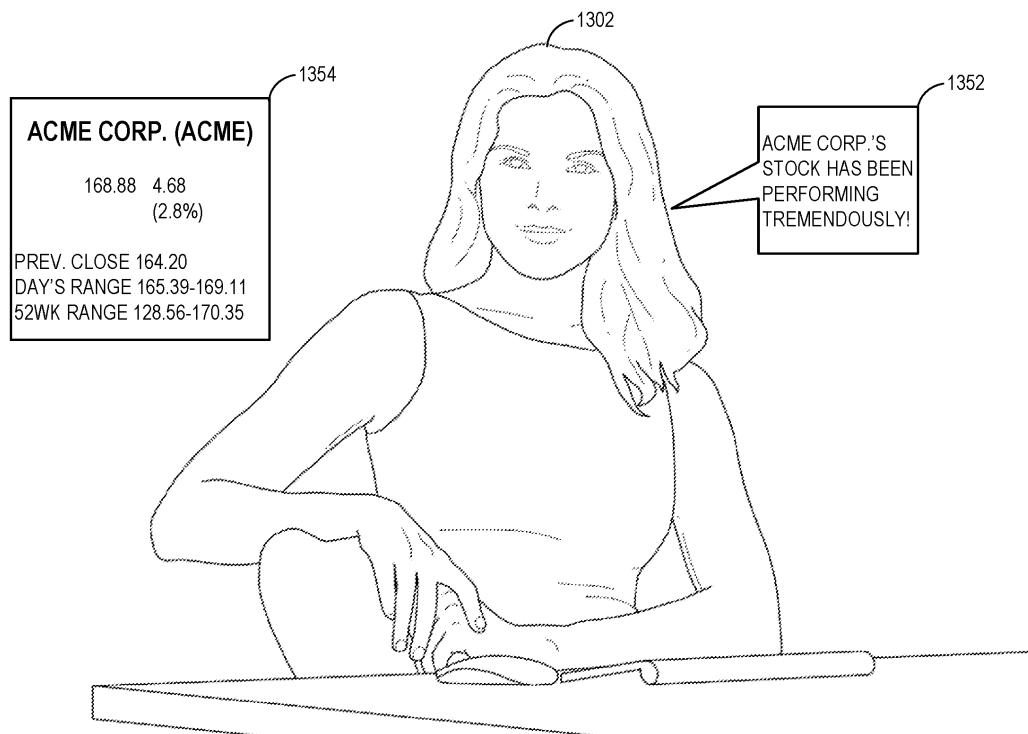
도면 13c



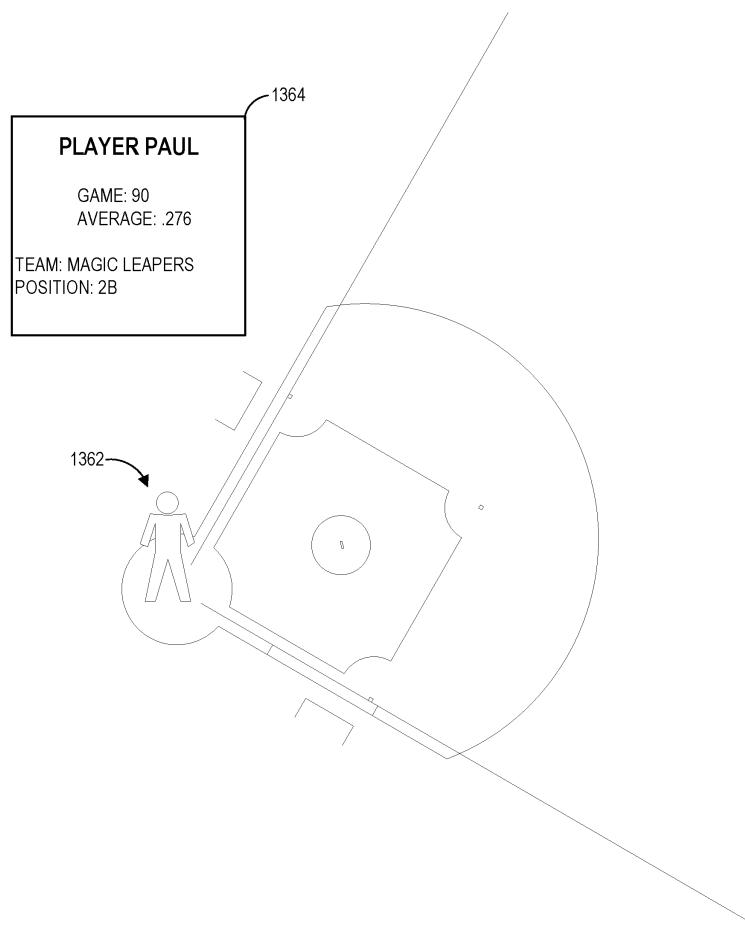
도면 13d



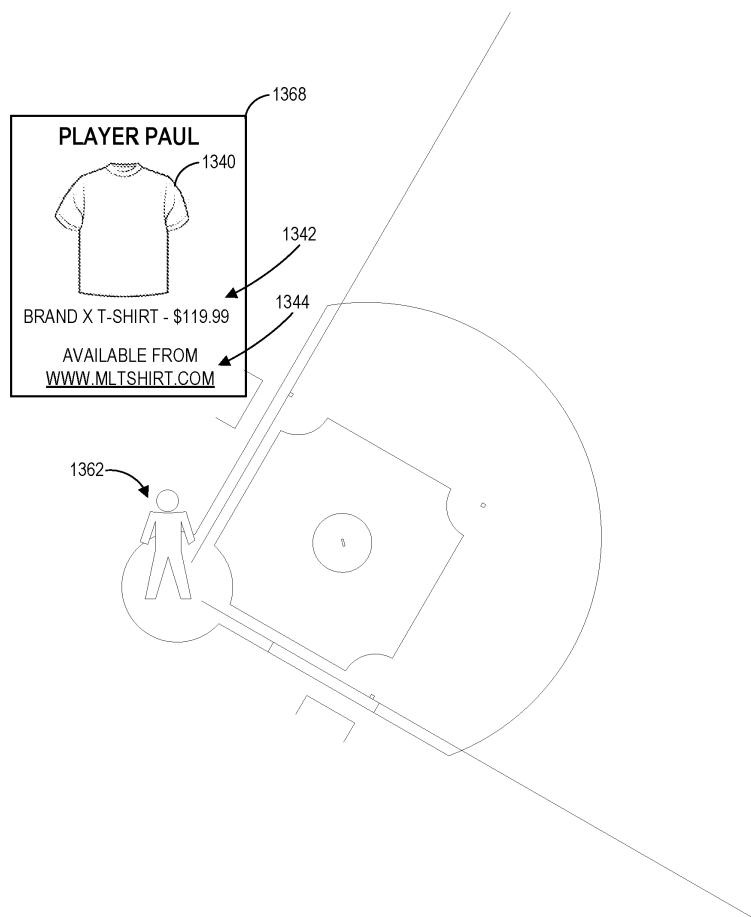
도면13e



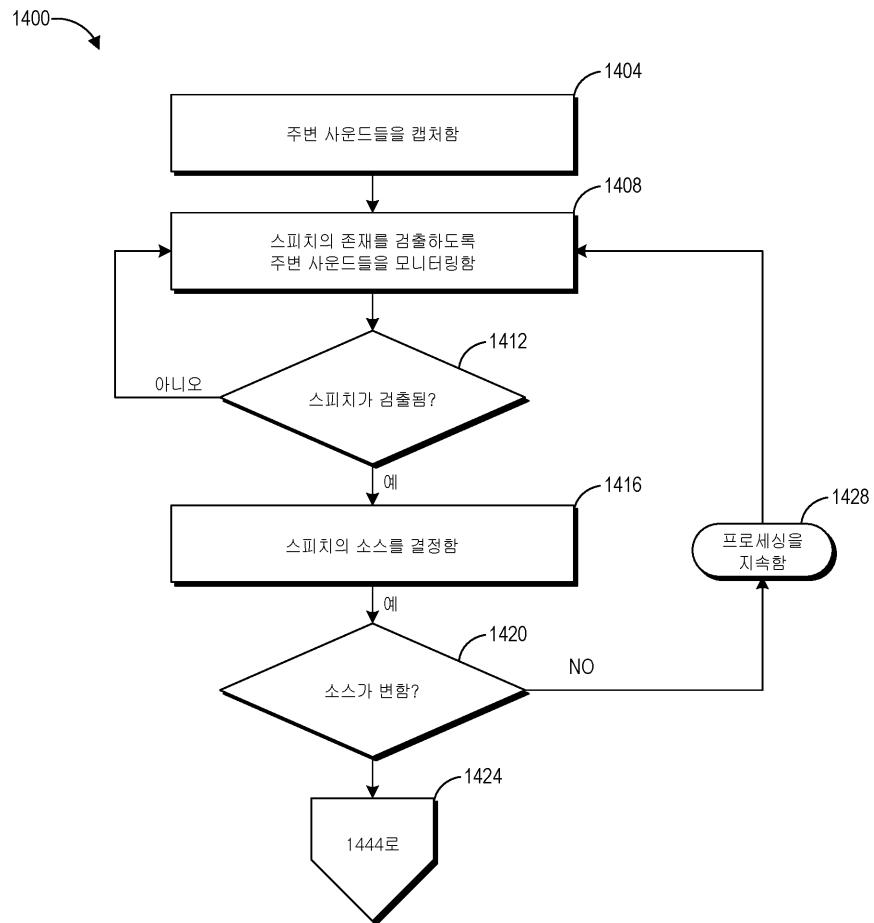
도면13f



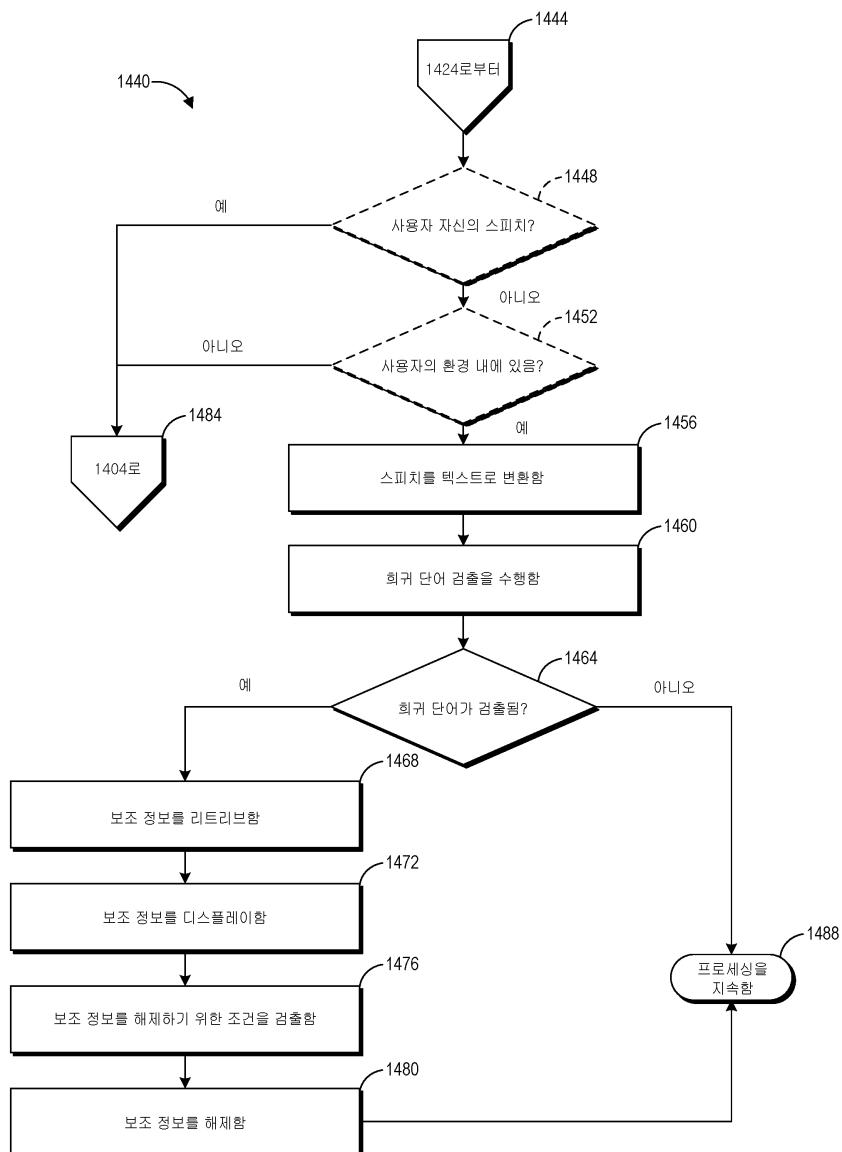
도면 13g



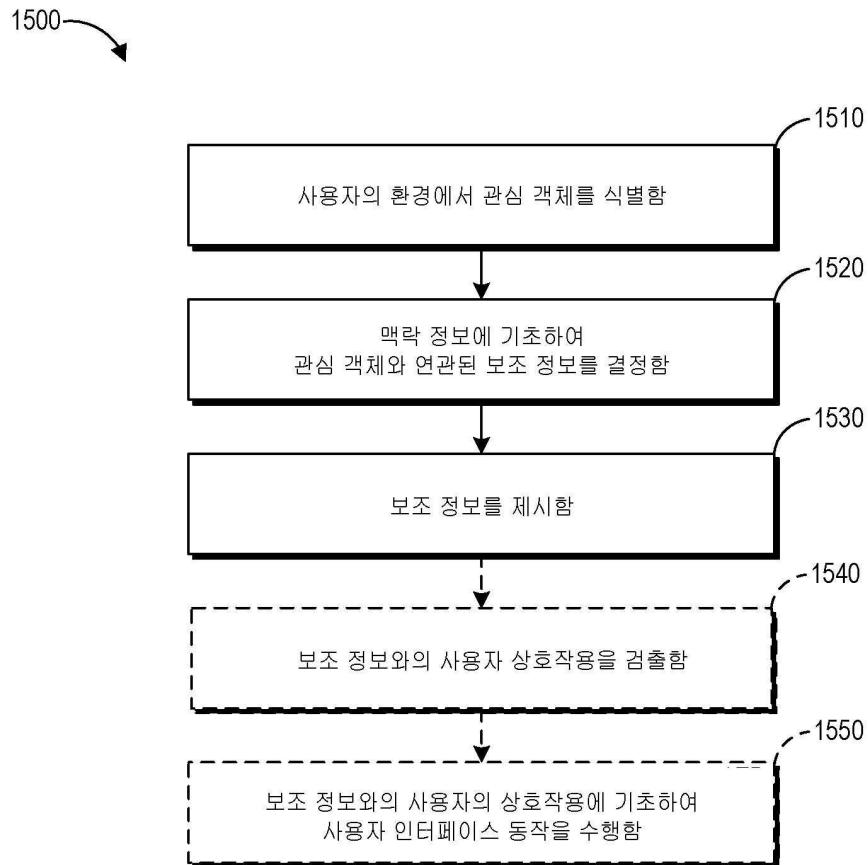
도면14a



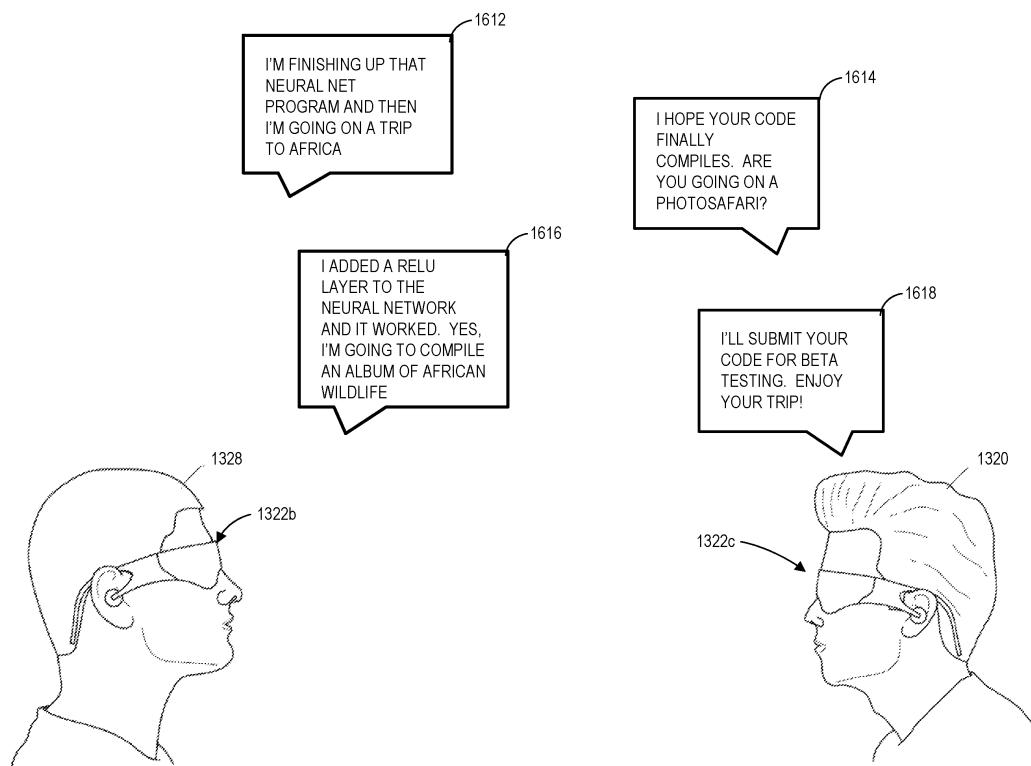
도면14b



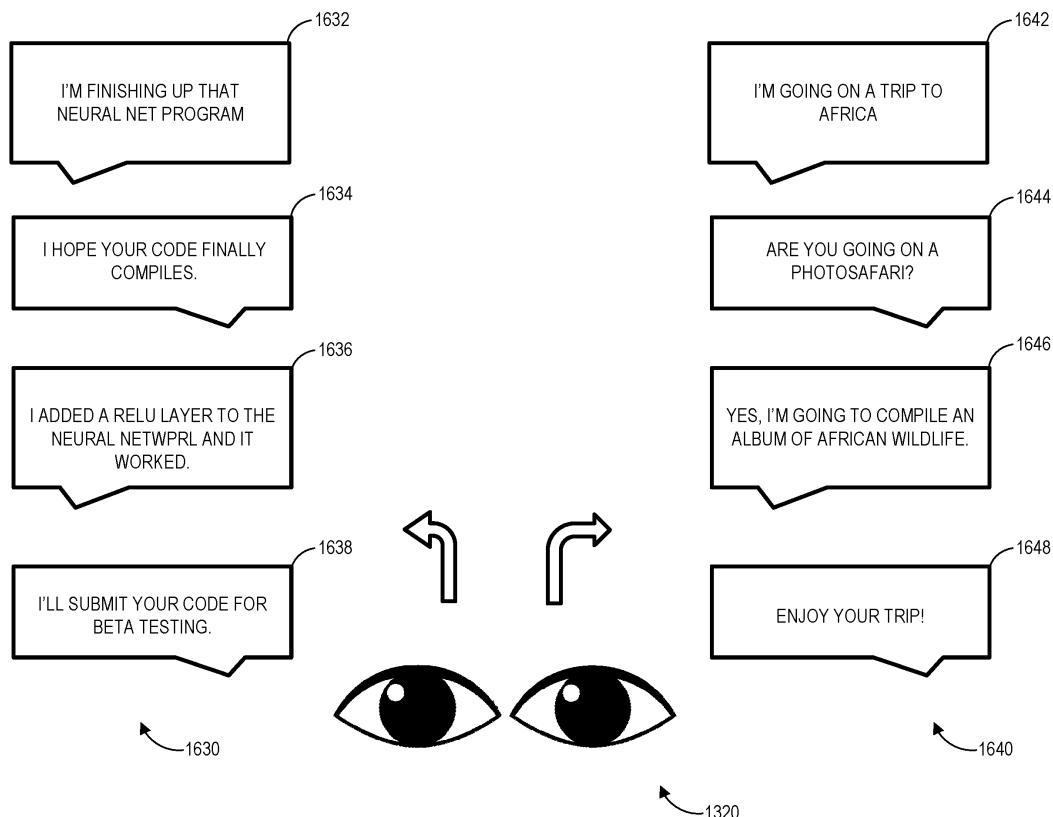
도면15



도면16a



도면16b



도면17

