



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2019년11월04일

(11) 등록번호 10-2040340

(24) 등록일자 2019년10월29일

(51) 국제특허분류(Int. Cl.)
G05D 1/12 (2006.01) G06K 9/20 (2006.01)

(21) 출원번호 10-2014-7002561

(22) 출원일자(국제) 2012년07월17일

심사청구일자 2017년06월15일

(85) 번역문제출일자 2014년01월28일

(65) 공개번호 10-2014-0053149

(43) 공개일자 2014년05월07일

(86) 국제출원번호 PCT/US2012/047089

(87) 국제공개번호 WO 2013/019402

국제공개일자 2013년02월07일

(30) 우선권주장

13/195,871 2011년08월02일 미국(US)

(56) 선행기술조사문헌

JP2006192563 A*

KR1020040008791 A

KR1020060018061 A

JP2005010244 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

마이크로소프트 테크놀로지 라이선싱, 엘엘씨

미국 워싱턴주 (우편번호 : 98052) 레드몬드 원
마이크로소프트 웨이

(72) 발명자

올리비에르 찰스 에프 III

미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 페이턴즈 마
이크로소프트 코포레이션

파우일라드 장 세바스티앙

미국 워싱턴주 98052-6399 레드몬드 원 마이크로
소프트 웨이 엘씨에이 - 인터내셔널 페이턴즈 마
이크로소프트 코포레이션

(뒷면에 계속)

(74) 대리인

제일특허법인(유)

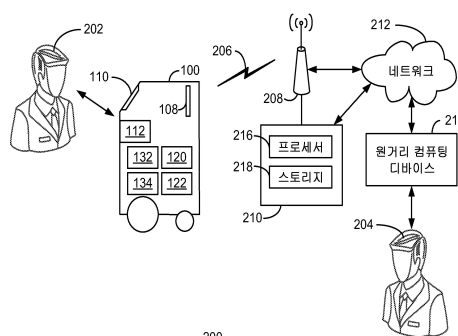
전체 청구항 수 : 총 19 항

심사관 : 신호영

(54) 발명의 명칭 피호출자 검색 기법

(57) 요약

본 발명에 따르면 로봇을 이용하여 사람과 텔레프레전스 세션을 개시하는 방법이 제공된다. 이 방법은 로봇에서 텔레프레전스 세션을 호스팅하라는 요청을 수신하는 단계와, 로봇을 통해 텔레프레전스 세션을 위한 타겟 인물에 대한 식별을 수신하는 단계를 포함한다. 그 다음 로봇은 사람에 대한 현재 위치를 검색한다. 만약 사람이 발견되면, 그 사람이 타겟 인물인지 여부와 관련된 판단이 이루어진다. 만약 발견된 사람이 타겟 인물이 아니라면, 발견된 사람이 타겟 인물에 대한 위치를 찾도록 유도된다. 로봇은 유도에 응답하여 발견된 사람에 의해 주어진 위치로 이동한다.

대표도

200

(72) 발명자

찰라비 말렉

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

클린턴 나타니엘 티

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

산체즈 러셀

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

펠론 아드리엔

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

웨일러 그레이엄

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

뷰리어넥 프랑소와

미국 워싱턴주 98052-6399 레드몬드 원 마이크로소프트 웨이 엘씨에이 - 인터내셔널 페이턴츠 마이크로소프트 코포레이션

명세서

청구범위

청구항 1

로봇을 이용하여 환경 내에 있는 타겟 인물(target person)과 텔레프레전스 세션(telepresence session)을 진행하는 시스템으로서,

상기 로봇을 상기 환경에서 이동시키도록 구성된 모터와,

카메라 및 마이크로폰을 포함하는 센서 시스템과,

로봇에 대한 모듈을 구현하도록 구성된 프로세서를 포함하되,

상기 모듈은,

상기 로봇을 상기 환경에서 이동시키도록 구성된 네비게이션 모듈과,

상기 카메라로부터의 이미지 내에 얼굴이 존재하는지를 식별하도록 구성된 얼굴 인식 모듈과,

상기 타겟 인물의 지정(designation)을 수락하고, 상기 환경 내에 있는 상기 타겟 인물의 위치를 찾기 위해서 상기 얼굴 인식 모듈 또는 상기 네비게이션 모듈, 또는 이들의 임의의 조합으로부터의 데이터를 이용하여 구성된 사람 위치추적 모듈(person location module)

을 포함하고,

상기 사람 위치추적 모듈은 발견된 인물이 상기 타겟 인물인지를 결정하고,

상기 발견된 인물이 상기 타겟 인물이 아닌 경우, 상기 사람 위치추적 모듈은

상기 타겟 인물의 위치에 대해 상기 발견된 인물에게 요청하고,

상기 요청에 응답하여 주어진 답변에서 제공된 위치로 상기 로봇을 이동시키며,

상기 발견된 인물이 상기 타겟 인물일 경우, 상기 사람 위치추적 모듈은

상기 텔레프레전스 세션에 대한 요청을 상기 타겟 인물에게 알리고,

상기 타겟 인물이 상기 텔레프레전스 세션에 대한 상기 요청을 수락하였는지 여부에 기초하여 상기 텔레프레전스 세션을 개시하는,

시스템.

청구항 2

제 1 항에 있어서,

상기 카메라에 의해 검출된 모션의 위치를 찾아내도록 구성된 모션 검출 시스템을 포함하는 시스템.

청구항 3

제 1 항에 있어서,

상기 마이크로폰에 의해 검출된 사운드의 위치를 찾아내도록 구성된 사운드 소스 위치추적 모듈을 포함하는 시스템.

청구항 4

제 3 항에 있어서,

상기 사람 위치추적 모듈은 상기 사운드 소스 위치추적 모듈에 의해 확인된 상기 사운드의 위치로 상기 로봇을

이동시킬 것을 상기 네비게이션 모듈에게 지시하도록 구성된 시스템.

청구항 5

제 1 항에 있어서,
상기 사람 위치추적 모듈은 원격 사용자에게 검색 상황을 제공하도록 구성된 시스템.

청구항 6

제 1 항에 있어서,
스피커를 더 포함하되, 상기 사람 위치추적 모듈은 상기 텔레프레전스 세션에 대한 상기 요청을 상기 스피커를 통해 알리도록 구성된 시스템.

청구항 7

제 1 항에 있어서,
상기 마이크로폰에 의해 검출된 구두 어구(spoken phrase)를 상기 로봇에 대한 커맨드 시퀀스로 변환하도록 구성된 음성 커맨드 모듈을 포함하는 시스템.

청구항 8

제 7 항에 있어서,
상기 커맨드 시퀀스는 상기 로봇을 목적지로 이동시킬 것을 상기 네비게이션 모듈에 지시하는데 사용되는 시스템.

청구항 9

제 1 항에 있어서,
상기 텔레프레전스 세션이 개시되면, 상기 텔레프레전스 세션을 수행하도록 구성된 텔레프레전스 모듈을 포함하는 시스템.

청구항 10

제 1 항에 있어서,
상기 프로세서는 상기 로봇과 통신하는 원격 컴퓨터 내에 있는 시스템.

청구항 11

제 1 항에 있어서,
상기 프로세서는 상기 로봇 내에 포함된 시스템.

청구항 12

제 1 항에 있어서,

상기 로봇과 물체 간의 거리를 결정하도록 구성된 깊이 검출 센서를 포함하는 시스템.

청구항 13

제 1 항에 있어서,
장애물 검출 모듈을 포함하는
시스템.

청구항 14

타겟 인물의 위치를 찾기 위해 자율 로봇을 이용하여 상기 타겟 인물과의 텔레프레전스 세션을 개시하는 방법으로서,

상기 로봇에서 텔레프레전스 세션을 호스팅하라는 요청을 수신하는 단계와,

상기 로봇에 의한 상기 텔레프레전스 세션을 위해 타겟 인물에 대한 식별을 수신하는 단계와,

인물에 대한 현재 위치를 검색하는 단계와,

인물이 발견되면, 상기 인물이 상기 타겟 인물인지 여부를 판정하는 단계와,

상기 발견된 인물이 상기 타겟 인물이 아닐 경우, 상기 타겟 인물의 위치에 대해 상기 발견된 인물에게 요청하고, 상기 요청에 응답하여 주어진 답변에서 제공된 위치로 상기 로봇을 이동시키는 단계와,

상기 발견된 인물이 상기 타겟 인물일 경우, 상기 텔레프레전스 세션에 대한 요청을 상기 타겟 인물에게 알리고, 상기 타겟 인물이 상기 텔레프레전스 세션에 대한 상기 요청을 수락하였는지 여부에 기초하여 상기 텔레프레전스 세션을 개시하는 단계를 포함하는,

방법.

청구항 15

제 14 항에 있어서,

얼굴 인식 소프트웨어를 이용하여 이미지 내의 인물을 식별하는 단계를 포함하는
방법.

청구항 16

제 14 항에 있어서,

상기 위치로 이동시키는 동안, 얼굴을 스캔하는 단계, 사운드를 듣는 단계, 또는 이 두 단계 모두를 포함하는
방법.

청구항 17

제 14 항에 있어서,

환경 내에서 모션의 위치를 찾는 단계와,

상기 로봇을 상기 모션의 위치로 이동시키는 단계와,

상기 모션의 위치에서 얼굴을 스캔하는 단계를 포함하는

방법.

청구항 18

제 14 항에 있어서,
사운드의 위치를 찾는 단계와,
카메라가 상기 사운드 쪽으로 향하도록 소탑(turret)을 이동시키는 단계와,
상기 사운드의 위치에서 얼굴을 스캔하는 단계를 포함하는
방법.

청구항 19

제 14 항에 있어서,
사운드가 커맨드를 포함하고 있음을 식별하는 단계와,
상기 커맨드의 위치로 이동하기 위해 움직임을 중단시키는 단계를 포함하는
방법.

발명의 설명

기술 분야

배경 기술

- [0001] 본 명세서에서 사용되는 "로봇(robot)"이라는 용어는, 로봇으로 하여금 독자적으로 기능들을 수행할 수 있게 하는 컴퓨터 하드웨어 및 소프트웨어를 포함하는 전자-기계적 디바이스를 지칭한다. 로봇의 예시로는 원격으로 조종될 필요없이 특정 타겟으로 날아가도록 구성될 수 있는 크루즈 미사일이 있다. 크루즈 미사일 상의 센서는 내부 데이터베이스에 대조되는 데이터를 출력할 수 있으며, 이는 크루즈 미사일로 하여금 의도된 타겟에 도달하는 것을 보장하도록 자신의 비행 패턴을 조정할 수 있게 한다.
- [0002] 크루즈 미사일이 군용 로봇인 반면, 소비자 레벨의 로봇들이 시장에 도입되어왔다. 예를 들어, 진공 청소기는 모터와 센서로 구성되어 진공 청소기로 하여금 독자적으로 특정 영역을 진공청소하고 충전장소로 복귀할 수 있게 한다. 다른 예시로 로봇 잔디 깎기가 도입되었으며, 로봇 잔디 깎기의 사용자는 바운더리를 정의하고 로봇 잔디 깎기는 정의된 바운더리 내에서 자동화된 방식으로 잔디를 깎도록 진행한다.
- [0003] 일부 로봇에게 원거리 위치로부터의 명령이 주어지는 것이 가능하도록 기술은 발전해왔다. 다시 말하면, 로봇은 예로서 네트워크의 방식으로 로봇으로부터 원거리에 있는 컴퓨팅 디바이스와 통신할 수 있으며, 로봇은 제공된 명령에 기초하여 동작들을 수행한다. 원거리의 사용자가 로봇이 위치된 환경에 참여하기 때문에 이를 "로보틱 텔레프레전스(robotic telepresence)"라고 지칭할 수 있다. 로보틱 텔레프레전스의 일 유형에서, 원거리의 사용자는 로봇의 환경에 있는 사람과 통신하기를 원할 수 있다.

발명의 내용

해결하려는 과제

과제의 해결 수단

- [0004] 아래에서는 본 명세서에 기술된 일부 측면들에 대한 기본적인 이해를 제공하기 위한 발명에 대한 간략한 요약물을 제시한다. 본 요약부는 청구된 청구사항에 대한 광범위한 개요가 아니다. 본 요약부는 청구된 청구사항의 기본 요소 또는 중요 요소를 식별하기 위한 것이 아니며 본 발명의 범주를 기술하고자 하는 것 또한 아니다. 본 요약부의 목적은 청구된 청구사항의 일부 개념을 아래에서 제시되는 상세한 설명에 대한 도입부로서 간략화된 형식으로 제시하고자 하는 것이다.
- [0005] 본 발명은 로봇을 포함하고 있는 환경에 사람을 위치시키고 로봇틱 텔레프레전스를 통해 통신을 개시하는 시스템 및 방법과 관련된다. 실시예는 로봇을 이용한 환경 내의 사람과의 텔레프레전스 세션(telepresence session)을 진행하는 시스템을 제공한다. 시스템은 환경에서 로봇을 이동시키도록 구성된 모터 및 카메라와 마이크로폰을 포함하는 센서 시스템을 포함한다. 프로세서는 로봇을 위한 모듈을 구현하도록 구성되며, 모듈은 환경에서 로봇을 이동시키도록 구성된 네비게이션 모듈, 환경에서의 모션(motion)을 검출하도록 구성된 모션 검출 시스템 및 마이크로폰에 의해 검출된 사운드의 위치를 찾도록 구성된 사운드 소스 위치추적 모듈을 포함한다. 얼굴 인식 모듈은 카메라로부터의 이미지 내에 얼굴이 존재하는지를 식별하도록 구성된다. 사람 위치추적 모듈은 타겟 인물의 지정(designation)을 수락하고, 환경 내의 타겟 인물의 위치를 찾기 위해서 사운드 소스 위치추적 모듈, 얼굴 인식 모듈, 네비게이션 모듈, 또는 이들의 임의의 조합으로부터의 데이터를 이용하도록 구성된다.
- [0006] 다른 실시예는 타겟 인물의 위치를 추적하기 위해 자율적인 로봇을 이용하여 타겟 인물과의 텔레프레전스 세션을 개시하는 방법을 제공한다. 이 방법은 로봇에서 텔레프레전스 세션을 호스팅하라는 요청을 수신하는 단계와 로봇에 의한 텔레프레전스 세션을 위해 타겟 인물에 대한 식별을 수신하는 단계를 포함한다. 로봇은 사람에 대한 현재 위치를 검색한다. 만약 사람이 발견되면 로봇은 그 사람이 타겟 인물인지 여부를 판정하고, 만약 아니라면 타겟 인물의 위치에 대해 그 사람에게 요청한다. 그 다음 로봇은 요청에 응답하여 주어진 답변에서 제공된 위치로 이동한다.
- [0007] 다른 실시예는 하나 이상의 프로세싱 디바이스에 의해 실행되었을 때 타겟 인물의 위치를 찾기 위한 시스템을 제공하는 컴퓨터 판독가능한 명령을 저장하는 하나 이상의 비휘발성 컴퓨터 판독가능한 저장 매체를 제공한다. 컴퓨터 판독가능한 명령은 음성 커맨드를 수신하도록 구성된 코드를 포함하되, 이러한 음성 커맨드는 로봇에게 특정 위치에서 사람을 찾도록 명령하며, 코드는 로봇을 특정 위치로 이동시키도록 구성된다.
- [0008] 본 요약부는 개념들의 선택을 간략한 형식으로 소개하고자 제공되었으며, 이러한 개념들은 아래의 상세한 설명에서 추가로 기술된다. 본 요약부는 청구된 청구사항의 중요 특성 또는 기본 특성을 식별하기 위한 것이 아니며, 청구된 청구사항의 범주를 한정하는데에 이용하기 위한 것 또한 아니다.

도면의 간단한 설명

- [0009] 도 1은 로봇틱 텔레프레전스 세션에서 사용될 수 있는 로봇의 개략도;
- 도 2는 로컬 사용자와 원거리 사용자 사이에서의 로봇틱 텔레프레전스 세션을 도시한 도면;
- 도 3은 로봇의 환경에 있는 사람의 위치를 찾아서 통신을 위한 로봇틱 텔레프레전스 세션을 확립하는 방법을 도시한 개괄적 도면;
- 도 4a 및 4b는 로봇의 환경에 있는 사람의 위치를 찾아서 통신을 위한 로봇틱 텔레프레전스 세션을 요청하는 방법의 더욱 상세한 예시를 제공하는 블록도;
- 도 5는 위치추적 기능을 수행하도록 사용될 수 있는 모듈을 나타내는 로봇의 블록도.

발명을 실시하기 위한 구체적인 내용

- [0010] 청구된 청구사항은 도면을 참조로 하여 기술되었으며, 도면들에 걸쳐 동일한 참조번호가 동일한 요소를 지칭하도록 사용되었다. 아래의 설명에서는 예시적인 목적으로 본 발명에 대한 철저한 이해를 제공하기 위한 다수의 특정한 세부사항이 기술되었다. 그러나, 청구된 청구사항이 이러한 특정한 세부사항 없이도 실시될 수 있음이 명백할 수 있다. 다른 예시에서, 잘 알려진 구조 및 디바이스가 본 발명에 대한 설명을 용이하게 하기 위해서 블록도의 형태로 도시되었다.
- [0011] 본 명세서에서 사용되는 "구성요소," "시스템," "클라이언트" 등의 용어는 하드웨어, (예로서, 실행중인) 소프트웨어,

트웨어, 및/또는 펌웨어, 또는 이들의 조합과 같은 컴퓨터 관련 엔티티를 지칭하기 위한 것이다. 예를 들어, 구성요소는 프로세서, 객체, 실행가능(executable), 프로그램, 함수, 라이브러리, 서브루틴, 및/또는 컴퓨터 상에서 구동중인 프로세스일 수 있거나 또는 소프트웨어와 하드웨어의 조합일 수 있다.

[0012] 예시로서, 서버 상에서 구동중인 애플리케이션과 서버 모두가 구성요소일 수 있다. 하나 이상의 구성요소는 프로세스 내에 존재할 수 있으며, 구성요소는 하나의 컴퓨터상에서 로컬화될 수 있고/있거나 둘 이상의 컴퓨터 사이에서 분산될 수 있다. "프로세서"라는 용어는 일반적으로 컴퓨터 시스템의 처리장치와 같은 하드웨어 구성요소를 지칭하는 것으로 이해된다.

[0013] 또한, 청구된 청구사항은 방법, 장치, 또는 개시된 청구사항을 구현하도록 컴퓨터를 제어하기 위한 소프트웨어, 펌웨어, 하드웨어, 또는 이들의 임의의 조합을 생산하기 위한 표준 프로그래밍 및/또는 엔지니어링 기술을 이용한 제조품으로서 구현될 수 있다. 본 명세서에서 사용되는 "제조품(article of manufacture)"이라는 용어는 임의의 비일시적인 컴퓨터 판독가능한 디바이스 또는 매체로부터 액세스 가능한 컴퓨터 프로그램을 포함하는 것이다.

[0014] 비일시적인 컴퓨터 판독가능한 저장 매체는 자기 저장 디바이스(예로서 하드 디스크, 플로피 디스크 및 자기 스트림 등), 광학 디스크(예로서 CD, DVD 등), 스마트 카드 및 플래시 메모리 디바이스(예로서 카드, 스틱 및 키드라이브 등)를 포함할 수 있지만, 이것으로 한정되는 것은 아니다. 반면에, 컴퓨터 판독가능한 매체는 (즉, 반드시 저장 매체인 것은 아님) 일반적으로 무선 신호 등에 대한 전송 매체와 같은 통신 매체를 추가로 포함할 수 있다.

[0015] 물론, 당업자는 청구된 청구사항의 범주 또는 사상으로 부터 벗어나지 않고 이러한 구성에 대한 다수의 수정이 이루어질 수 있음을 인식할 것이다. 또한, 본 명세서에서 사용되는 "예시적인"이라는 용어는 예시, 사례, 또는 보기로서의 역할을 함을 의미하기 위한 것이다. 본 명세서에 "예시"로서 기술된 임의의 측면 또는 설계가 반드시 다른 측면 또는 설계보다 선호된다거나 바람직한 것으로 해석되는 것은 아니다.

[0016] 본 명세서에서 사용되는 바와 같이, 로보틱 텔레프레전스 세션은 통신을 제공할 수 있으며, 이러한 통신에서 원거리 사용자는 로봇을 통해 타겟 인물과 통신할 수 있는 능력을 갖는 로봇의 비디오, 오디오 및 센서 데이터에 액세스하도록 원거리 컴퓨팅 디바이스 상의 애플리케이션을 이용한다. 본 명세서에서 사용된 바와 같이, 원거리 사용자는 예를 들어 인터넷에 의해 분리되는 로봇의 네트워크와는 다른 네트워크 상에 위치될 수 있거나, 또는 로컬 네트워크 상에 위치될 수 있다. 원거리 컴퓨팅 디바이스는 개인 컴퓨터, 모바일 디바이스, 또는 게임 콘솔 등을 포함할 수 있다. 애플리케이션은 원거리 컴퓨팅 디바이스 내의 독립형 애플리케이션일 수 있거나 또는 웹 브라우저를 통해 액세스될 수 있다.

[0017] 로보틱 텔레프레전스 세션은 원거리의 사용자에게 타겟 인물을 포함하는 로봇의 환경의 이미지를 제공할 수 있다. 이것은 통신의 리얼리티를 개선할 수 있으며, 따라서 보다 효율적이고 즐거운 통신 경험을 생성한다. 통신 도구(tool)로서, 로보틱 텔레프레전스 세션은 사업에서 중요한 기술을 제공할 뿐 아니라 우주비행사, 잠수함 승무원 및 군인과 같이 가족과 친구들로부터 멀리 떨어져 있는 사람들이 마치 집에 가까운 곳에 있는 것과 같이 느끼도록 도울 수 있다.

[0018] 그러나 타겟 인물이 로봇을 보거나, 듣거나, 또는 로봇에 액세스할 수 없기 때문에 피호출자(a called party), 또는 타겟 인물이 텔레프레전스 요청에 답할 수 없을 수도 있다. 타겟 인물이 로봇에 가까운 위치에 존재하지 않는 한, 타겟 인물을 찾기 위해 환경에서 로봇을 구동하는데에 소중한 시간이 소비될 수 있다.

[0019] 또한, 호출자가 환경에서 로봇을 구동하는 것을 허용하지 않는 것이 유용할 수도 있다. 예를 들어, 호출이 실제로 수락되기 전에는 호출자가 피호출자의 환경을 들여다보는 것을 피호출자가 원치 않을 수 있다. 또한, 피호출자가 실제로 호출을 수락할 수 있는 상황인지를 호출자가 지켜보는 것을 피호출자가 원치 않을 수 있다.

[0020] 이러한 문제를 해결하기 위해서, 본 명세서에 개시된 실시예는 환경 내에서 독자적으로 사람의 위치를 찾기 위한 접근법을 기술한다. 이러한 접근법에서, 로봇은 예를 들어 텔레프레전스 요청을 발표하고, 자신의 환경에서 (특정하거나 비특정한) 타겟 인물을 찾으며, 자신의 환경에 있는 사람들로부터의 지원을 요청함으로써, 피호출자에게 가까워지고자 시도한다. 로봇은 환경 내의 사람들을 식별하기 위해, 얼굴 식별, 얼굴 인식, 사운드 입력 및 음성 쿼리와 같은 센서 데이터와 결합하여 환경의 자동화된 네비게이션을 이용한다. 텔레프레전스 세션의 개시자(initiator)는 로봇과 지속적으로 연락해야만 하는 것은 아니며, 로봇이 해당 인물을 찾는 동안 다른 활동을 수행할 수 있다.

[0021] 로봇이 타겟 인물을 발견하였다면, 이를 개시자에게 알리고 타겟 인물에게 통신을 수락하길 원하는지를 질문할

수 있으며, 이때 통신을 개시하는 원거리 사용자가 알림을 받고 통신 채널이 개방된다. 만약 로봇이 특정 인물의 위치를 추적하라는 요청을 받지 않았다면, 로봇은 타겟 인물을 가장 먼저 발견한 사람으로 지정할 수 있으며, 그 사람에게 통신 세션을 수락하길 원하는지 질문할 수 있다. 어떠한 경우라도, 만약 그 사람이 통신을 수락하길 거부한다면 로봇은 원거리의 사용자에게 이를 알리고 로보틱 텔레프레전스 세션을 종료할 수 있다.

[0022] 도 1은 로보틱 텔레프레전스 세션에서 사용될 수 있는 로봇의 개략도이다. 로봇(100)은 헤드(102) 및 바디(104)를 포함하며, 이때 헤드(102)는 바디(104)에 대해서 이동가능하다. 로봇(100)은 바디(104)에 대해서 헤드(102)를 회전시키는 하나 이상의 모터를 포함할 수 있는 헤드 회전 모듈(106)을 포함할 수 있다. 예를 들어, 헤드 회전 모듈(106)은 바디(104)에 대해서 헤드(102)를 양쪽 각방향(angular direction)으로 45° 까지 회전시킬 수 있다. 다른 예시에서, 헤드 회전 모듈(106)은 바디(104)에 대해서 헤드(102)를 90° 회전시킬 수 있다. 다른 예시에서, 헤드 회전 모듈(106)은 바디(104)에 대해서 헤드(102)를 180° 회전시킬 수 있다.

[0023] 헤드(102)는 무선 신호를 수신 및 전송하도록 구성된 안테나(108)를 포함할 수 있다. 예를 들어, 안테나(108)는 Wi-Fi 신호, 블루투스 신호, 적외선 신호, 초음파 신호, 무선 주파수(RF) 신호, 또는 임의의 다른 적절한 신호를 수신 및 전송하도록 구성될 수 있다. 또 다른 예시에서, 안테나(108)는 데이터를 셀룰러 타워로 전송하고 셀룰러 타워로부터 수신하도록 구성될 수 있다. 로봇(100)은 안테나(108)를 통해 원거리에 위치한 컴퓨팅 디바이스와의 통신을 전송 및 수신할 수 있다.

[0024] 로봇(100)의 헤드(102)는 또한 로봇(100) 부근에 있는 개인에게 데이터를 디스플레이하기 위한 디스플레이(110)를 포함할 수 있다. 예를 들어, 디스플레이는 로봇이 위치를 찾고자 하는 사람의 이름 또는 사진을 디스플레이하도록 구성될 수 있다. 다른 예시에서, 디스플레이(110)는 로보틱 텔레프레전스 세션 동안 원거리 사용자의 이미지를 디스플레이하도록 구성될 수 있다. 디스플레이는 또한 "세션 시작" 또는 "세션 종료" 등과 같은 텔레프레전스 세션의 현재 상태를 디스플레이할 수도 있다. 실시예에서, 로봇(100)은 원거리 사용자의 이미지를 투사하도록 구성된 프로젝터(112)를 포함할 수 있으며, 이는 타겟 인물을 위한 텔레프레전스 세션의 리얼리티를 증가시킨다.

[0025] 로봇(100)의 헤드(102)는 로봇의 환경의 연속적인 이미지, 예로서 텔레프레전스 세션 동안의 타겟 인물의 연속적인 이미지를 캡처하도록 구성된 카메라(114)를 포함할 수 있다. 카메라(114)는 시스템 설계 고려사항에 따라 광범위한 카메라 유형 중 임의의 카메라일 수 있다. 사용될 수 있는 카메라의 예는 스틸 카메라, 비디오 카메라, 적외선 카메라, 입체(stereoscopic) 카메라 또는 다양한 다른 유형의 광학 센서의 어레이를 포함한다. 원거리 사용자에 대한 텔레프레전스 세션의 리얼리티를 증가시키기 위해서, 카메라(114)는 예를 들어 720p 포맷, 720i 포맷, 1080p 포맷, 1080i 포맷, 또는 다른 적절한 고해상도 비디오 포맷의 비디오 데이터의 캡처를 용이하게 하는 고해상도 비디오 카메라일 수 있다. 헤드(102)는 또한 상대적으로 낮은 해상도 데이터를 캡처하는 카메라(114)를 포함할 수도 있다. 이는 모바일폰 상에서의 비디오 디스플레이와 같이 보다 대역폭 한정된 접속 상에서의 비디오 접속을 가능하게 할 수 있다. 카메라(114)가 로봇(100)의 헤드(102) 내에 장착되기 때문에, 헤드 회전 모듈(106)은 카메라(114)가 로봇(100)의 환경의 비교적 넓은 부분의 비디오 데이터를 캡처하는 것을 가능케 할 수 있으며, 타겟 인물에 대한 검색을 더욱 효율적으로 만들 수 있다.

[0026] 로봇(100)의 헤드(102)는 로봇(100)이 자체적인 네비게이션을 수행하는 것을 도울 수 있는 센서(116)를 포함할 수 있다. 센서(116)는 임의의 적절한 유형일 수 있다. 예를 들어, 센서(116)는 로봇과 객체 사이의 거리를 결정하도록 구성되는 깊이 센서, 적외선 센서, 카메라, 로봇(100)에 가장 근접한 고도에서의 급경사(drop-off)를 검출하도록 구성된 클리프 센서(cliff sensor), GPS 센서, 가속도계, 자이로스코프, 또는 다른 센서 유형을 포함할 수 있다.

[0027] 로봇(100)의 바디(104)는 로봇(100) 내의 다른 모듈에 전력을 제공하도록 동작할 수 있는 배터리(118)를 포함할 수 있다. 배터리(118)는 예를 들어 납축전지(lead-acid battery), 젤화된 납축전지(gelled lead acid battery), 리튬 이온 배터리, 또는 니켈-카드뮴 배터리 등과 같은 충전가능한 배터리일 수 있다. 만약 배터리(118)가 충전가능하면, 로봇(100)은 배터리(118)에 전하가 비교적 쉽게 제공될 수 있도록 로봇(100)이 전원에 연결되게 하는 인터페이스를 포함할 수 있다.

[0028] 로봇(100)의 바디(104)는 스토리지(120) 및 프로세서(122)를 포함할 수 있다. 스토리지(120)는 랜덤 액세스 메모리(RAM), 판독 전용 메모리(ROM), 플래시 메모리, 하드 드라이브, 플래시 드라이브, 또는 광학 드라이브의 임의의 조합을 포함할 수 있다. 도 5와 관련하여 추가로 상세하게 기술되는 바와 같이, 스토리지(120)는 프로세서(122)에 의해 실행가능한 복수의 모듈을 저장할 수 있으며, 모듈들은 로봇(100)의 동작을 제어할 수 있다. 프로세서(122)는 단일-코어 프로세서, 멀티-코어 프로세서, 또는 예를 들어 서로 다른 기능에 대해 서로 다른

코어들이 할당된 컴퓨팅 클러스터일 수 있다. 프로세서(122)는 마더보드에 의해 호스팅되는 버스과 같은 임의의 적절한 인터페이스에 의해 로봇(100) 내의 다른 장치들과 결합될 수 있다. 프로세서(122)는 로봇(100)의 원거리 사용자가 원하는 방식으로 예를 들어 타겟 인물의 위치를 찾고 로봇(100)의 텔레프레젠텐스 세션을 수행하게 하기 위해서 로봇(100)이 로봇(100) 내의 다른 장치들뿐 아니라 원거리 컴퓨터로부터 수신된 데이터를 프로세싱하도록 이용된다. 일부 실시예에서, 로봇(100)의 텔레프레젠텐스를 구현하도록 사용되는 프로세서 및 모듈은 무선 네트워크 상에서 로봇과 접촉하는 원거리 컴퓨터 내에 위치될 수 있다.

[0029] 로봇(100)의 헤드(102) 내의 센서(116)에 추가로, 바디(104) 또한 자체적인 또는 반-자체적인 네비게이션과 결합하여 이용될 수 있는 센서(124)를 포함할 수 있다. 예를 들어, 센서(124)는 초음파 센서, 위치 센서, 적외선 센서, 카메라, 클리프 센서 등을 포함할 수 있다. 센서(116, 124)에 의해 캡처되는 데이터는 예를 들어 타겟 인물의 위치를 추적하고 로봇(100)의 텔레프레젠텐스 세션을 진행하기 위해 로봇(100)을 자체적으로 네비게이션하도록 프로세서(122)에 의해 이용될 수 있다.

[0030] 로봇(100)의 바디(104)는 로봇(100)의 하나 이상의 휠(128, 130)을 구동하도록 동작할 수 있는 드라이브 모터(126)를 포함할 수 있다. 예를 들어, 큰 휠(128)은 드라이브 휠(driving wheel)일 수 있는 반면, 작은 휠(130)은 로봇(100)의 방향을 변경할 수 있는 스티어링 휠(steering wheel)일 수 있다. 일부 실시예에서, 각각의 휠(128, 130)은 스티어링 메커니즘을 가질 수 있으며 로봇(100)의 방향을 변경하기 위해서 함께 사용될 수도 있다. 또한, 하나의 드라이브 모터(126)가 휠(128, 130) 모두를 구동하는 것으로서 도시되었지만, 다른 드라이브 모터가 휠(130) 또는 휠(128)을 구동하도록 사용될 수 있다. 센서(116, 124)로부터의 데이터에 따라, 또는 원거리 컴퓨팅 디바이스로부터의 커맨드 수신에 따라, 프로세서(122)는 바디(104)에 대한 헤드(102)의 방향을 제어하기 위해 헤드 회전 모듈(106)로 신호를 전송할 수 있으며, 로봇(100)의 방향과 위치를 제어하기 위해 드라이브 모터(126)로 신호를 전송할 수 있다.

[0031] 로봇(100)의 바디(104)는 스피커(132) 및 마이크로폰(134)을 포함할 수 있다. 마이크로폰(134)에 의해서 캡처된 오디오 데이터는 안테나(108)에 의해 원거리 컴퓨팅 디바이스에 전송될 수 있다. 따라서, 원거리 사용자는 실시간 오디오/비디오 피드를 수신할 수 있으며, 로봇(100)의 텔레프레젠텐스 세션 동안 로봇(100)의 환경을 경험할 수 있다. 스피커(132)는 로봇(100)에 가장 가까운 한 명 이상의 개인에게 오디오 데이터를 출력하도록 이용될 수 있다. 오디오 데이터는 타겟 인물의 위치 요청, 텔레프레젠텐스 세션을 개시하라는 타겟 인물에 대한 요청, 텔레프레젠텐스 세션 동안의 원거리 사용자로부터의 오디오 피드, 또는 임의의 다른 사운드일 수 있다.

[0032] 로봇(100)은 전술된 구성 및 기능적 장치로 한정되지 않으며, 다양한 다른 방식으로 구성될 수 있다. 예를 들어, 헤드 회전 모듈(106)은 로봇(100)의 헤드(102)가 바디(104)에 대한 회전 및 수직 방향에서의 기울기 모두를 포함할 수 있도록 기울기 모터(106)를 포함할 수 있다. 일부 실시예에서, 로봇(100)은 별개의 헤드(102) 및 바디(104)를 포함하지 않을 수도 있으며, 단일의 통합된 바디를 구비할 수 있다. 이러한 실시예에서, 로봇 바디는 카메라(114)를 통해 비디오 데이터를 캡처하는 것을 가능케 하도록 변경될 수 있다. 다른 실시예에서, 로봇(100)은 통합된 바디 구조를 가질 수 있지만, 카메라(114)는 카메라(114)로 하여금 환경의 상이한 모습을 획득할 수 있게 하는 관련된 모듈을 구비할 수 있다. 또한, 도 1의 바디(104)에 있는 것으로 도시된 기능적 장치들은 로봇(100)의 헤드(102)에 배치될 수도 있으며, 그 반대 경우 또한 가능하다.

[0033] 도 2는 로컬 사용자(202)와 원거리 사용자(204) 사이의 로봇(100)의 텔레프레젠텐스 세션(200)을 도시한 도면이다. 동일한 참조번호는 도 1에 대해 기술된 것과 같다. 로봇(100)은 무선으로 데이터(206)를 수신 및 전송하도록 구성된 안테나(108)를 포함할 수 있다. 예시적인 실시예에서, 로봇(100)이 파워-온 되었을 때, 로봇(100)은 액세스 포인트(208)와 자신의 프레젠텐스를 확립하도록 무선 액세스 포인트(208)와 통신할 수 있다. 액세스 포인트(208)를 통해, 로봇(100)은 예를 들어 환경 내의 로컬 컴퓨팅 디바이스(210)로 접속할 수 있다. 로봇(100)은 직접 또는 로컬 컴퓨팅 디바이스(210)를 통해서 네트워크(212)로 접속할 수 있다. 예를 들어, 네트워크(212)는 셀룰러 네트워크, 인터넷, 인트라넷 또는 가상 개인 네트워크와 같은 독점적인 네트워크, 또는 다른 적절한 네트워크일 수 있다. 로봇(100)이 자신의 모든 기능을 수행할 수 있기 때문에 통신 세션은 로컬 컴퓨팅 디바이스(210)를 이용하는 것으로 한정되지 않는다.

[0034] 원거리 사용자(204)는 원거리 컴퓨팅 디바이스(214) 상의 텔레프레젠텐스 애플리케이션을 시작할 수 있다. 보다 현실적인 경험을 위해서, 원거리 컴퓨팅 디바이스(214)는 넓은 스크린 모니터, 가상 리얼리티 디스플레이, 또는 텔레컨퍼런스 룸 등을 갖는 컴퓨터일 수 있다. 다른 실시예에서, 원거리 컴퓨팅 디바이스(214)는 데스크톱 컴퓨터, 랩톱 컴퓨터, 태블릿 컴퓨터, 게임 시스템, 또는 스마트폰 등과 같은 보다 작은 장치일 수 있다. 원거리 컴퓨팅 디바이스(214)는 네트워크(212)를 통해 로컬 컴퓨팅 디바이스(210)에 접속하거나 또는 로봇(100)에 직접

접속할 수 있다. 그 다음 원거리 컴퓨팅 디바이스(214)는 텔레프레전스 세션을 요청할 수 있다.

[0035] 텔레프레전스 세션을 제어하는 모듈은 로봇(100) 상의 스토리지(120) 내에 저장될 수 있으며, 예를 들어 텔레프레전스 세션이 요청되었을 때 로봇 상에서 프로세서(122)에 의해 로딩되어 실행될 수 있다. 일부 실시예에서, 모듈은 로컬 컴퓨팅 디바이스(210) 상의 스토리지 디바이스(218)로부터의 코드를 실행할 수 있는 로컬 컴퓨팅 디바이스(210) 내의 프로세서(216)에 의해 구동될 수 있다. 도시되지 않았지만, 원거리 컴퓨팅 디바이스(214)는 로컬 사용자(202), 예로서 텔레프레전스 세션에 대한 타겟 인물의 위치추적에 대한 상태 리포트를 제공하는 것에 추가로, 로봇(100)의 네비게이션과 관련된 원거리 사용자(204)에게 데이터를 제시할 수 있는 디스플레이 스크린을 포함할 수 있다. 예를 들어, 검색 중에 로봇(100)은 네트워크(212)를 통해서 원거리 컴퓨팅 디바이스(214)로 라이브 오디오/비디오 피드를 전송할 수 있다. 또한, 원거리 사용자(204)는 네트워크(212) 상에서 원거리 컴퓨팅 디바이스(214)에 의해 로봇(100)으로 네비게이션 및 텔레프레전스 커맨드를 전송할 수 있다.

[0036] 실시예에서, 원거리 사용자(204)는 텔레프레전스 세션에 대한 로봇(100)의 환경 내의 로컬 사용자(202)의 위치를 자체적으로 추적하도록 로봇(100)을 이용할 수 있다. 로컬 사용자(202)의 위치가 발견되어 통신이 수락되면, 텔레프레전스 세션이 시작될 수 있다. 텔레프레전스 세션에 대한 통신 채널이 핸드셰이킹, 인증 등과 다양한 동작을 통해 네트워크(212)에 의해서 원거리 컴퓨팅 디바이스(214)와 로봇(100) 사이에서 확립될 수 있다. 일 실시예에서, 비디오 전화 접속이 H.246 SVC(Scalable Video Coding) 표준과 같은 표준 프로토콜을 이용하여 확립된다. 다수의 유사한 인코딩 방안들이 실시예에서 사용될 수 있다. 원거리 사용자(204)와 로컬 사용자(202)는 그 다음 텔레프레전스 세션을 통해 통신할 수 있다.

[0037] 도 3은 로봇의 환경 내에 있는 사람의 위치를 찾고 통신을 위한 로보틱 텔레프레전스 세션을 확립하기 위한 방법(300)의 오버뷰이다. 추가적인 기능들이 포함될 수 있기 때문에 방법(300)은 도시된 기능들로 한정되지 않는다. 이 방법은 예를 들어 도 2와 관련하여 기술된 바와 같이, 원거리 사용자가 원거리 컴퓨팅 디바이스 상의 텔레프레전스 애플리케이션을 시작하는 블록(302)에서 시작한다. 블록(304)에서, 원거리 컴퓨팅 디바이스는 로봇에 대한 접속을 확립한다. 논의된 바와 같이, 로봇과 직접 접속이 확립될 수 있거나 또는 예로서 기지국으로서의 기능을 하는 로봇의 환경에 대해 로컬인 컴퓨터를 통해 접속이 확립될 수 있다. 접속이 확립되면, 로봇은 텔레프레전스 세션을 수행하기 위해 소프트웨어를 로딩할 수 있다. 블록(306)에서, 원거리 사용자는 특정 인물을 타겟으로 하는지 또는 로봇이 마주친 첫 번째 사람이 타겟 인물로서 역할을 할 것인지를 나타낸다.

[0038] 원거리 사용자가 세션을 개시한 후에, 로봇은 동작을 취하기 시작할 수 있다. 블록(308)에서, 로봇은 예를 들어 음성 알람 또는 톤(tone)을 통해서 청각적으로 또는 빛, 스크린상의 알람 등을 통해서 시각적으로 텔레프레전스 요청을 발표할 수 있다. 발표의 선택은 로컬 사용자에 의해 제어될 수 있으며, 예를 들어 개인 설정 또는 기능할 수 있는 알람을 한정하는 밤시간 설정을 설정할 수 있다. 일부 실시예에서, 로컬 사용자는 더 시끄러운 알람 또는 예를 들어 텔레프레전스 세션이 예정되어 있는지 여부와 같은 개선된 알람을 요청할 수 있다.

[0039] 블록(310)에서, 로봇은 자신의 환경 내의 사람에 대한 환경을 스캔한다. 이것은 시각적 및 청각적으로 수행될 수 있다. 예를 들어, 로봇은 이미지 내의 얼굴을 식별할 수 있거나 "이쪽으로"와 같은 음성 커맨드를 들을 수 있다. 블록(312)에서, 만약 당면한 환경에서 사람이 발견되지 않았다면, 흐름은 블록(314)으로 진행한다. 블록(314)에서, 로봇은 새로운 위치, 예로서 목적지로 이동하며, 이동하면서 사람들을 스캔하고 커맨드를 듣는다. 만약 로봇이 하나의 위치로부터 다른 위치로 이동하는 동안 객체가 로봇 앞에서 움직인다면, 로봇은 움직임을 중단하고 유효 얼굴을 스캔함으로써 예상치 못한 경로 중단이 사람 또는 객체에 의해 발생되었는지 여부를 결정할 수 있다. 만약 중단이 사람에 의해서 발생되었다면, 로봇은 그 사람의 관심을 끌 수 있다. 만약 중단이 객체에 의해서 발생되었다면, 로봇은 객체를 둘러서 이동할 수 있으며 새로운 위치로 계속해서 이동한다. 그 다음 다시 블록(310)으로 복귀하여, 로봇은 사람에 대한 새로운 위치를 스캔한다.

[0040] 만약 블록(312)에서 로봇이 당면한 환경 내의 사람을 발견하면, 프로세스 흐름은 블록(316)으로 진행한다. 블록(316)에서, 로봇은 발견한 사람이 타겟 인물인지 여부를 판정한다. 이러한 결정은 예를 들어 이미지들의 카탈로그로부터의 얼굴 인식, 그 사람이 타겟 인물이라는 확인과 같은 음성 입력, 또는 둘 모두에 의해 이루어질 수 있다. 만약 블록(316)에서 발견된 사람이 타겟 인물이 아니면, 블록(318)에서 로봇은 타겟 인물의 위치추적을 위해 발견된 인물에게 질의할 수 있다. 로봇은 응답을 디코딩할 수 있으며, 예로서 "그가 사무실에 없습니다," "그녀가 사무실 내에 있습니다," "코너를 돌아 왼쪽으로" 등과 같은 구절에 대한 오디오 스트림을 분석할 수 있다. "이곳에 없음," "사무실," "부엌" 등과 같은 보다 단순한 구절들도 사용될 수 있다. 만약 로봇이 "사무실"과 같은 위치 용어를 수신하면, 새로운 위치가 네비게이션 타겟으로서 설정될 수 있으며, 프로세스 흐름은 블록(314)으로 복귀되어 로봇은 새로운 위치 또는 목적지로 이동하며, 이동하면서 사람들을 스캔하고 커맨드

를 듣는다. 로봇은 음성 커맨드가 올바르게 해석되었는지를 보장하기 위해 입력 확인을 요청할 수 있다.

- [0041] 만약 블록(316)에서 로봇이 발견한 사람이 타겟 인물이라고 판정하거나 또는 만약 타겟이 처음 발견된 사람을 선택하도록 설정되었다면, 프로세스 흐름은 블록(320)으로 진행한다. 블록(320)에서, 로봇은 텔레프레전스 세션의 개시를 요청하며, 예를 들어 타겟 인물에게 "텔레프레전스 세션을 수락합니까?"라고 질문한다. 블록(322)에서, 로봇은 로컬 사용자가 텔레프레전스 세션을 수락하였는지를 판정한다. 예를 들어, 로컬 사용자는 "로봇 수락," "로봇 거절" 등과 같은 표현을 말할 수 있고, 로봇 상의 버튼을 누를 수 있거나, 또는 모바일폰과 같은 원격 디바이스 또는 펜던트 컨트롤을 통해 커맨드를 전송할 수 있다. 만약 블록(322)에서 로봇이 텔레프레전스 세션이 수락되었다고 판정하면, 블록(324)에서 로봇은 세션을 개시하고, 세션 참여자가 예를 들어 "로봇 종료 호출"과 같은 음성 커맨드를 주거나 버튼을 누름으로써 링크를 종료할 때까지 통신 링크를 유지한다.
- [0042] 만약 블록(322)에서 로컬 사용자가 텔레프레전스 세션을 거절하면, 프로세스 흐름은 블록(326)에서 종료되며, 로봇은 원거리 사용자에게 이를 알리고 통신 링크를 종료한다. 로봇이 세션을 종료할 때, 예를 들어 미래의 텔레프레전스 세션을 위해 디스크 상에 저장된 모듈에 의존하여 RAM으로부터 텔레프레전스 세션을 진행하기 위해 사용된 모듈을 삭제할 수 있다.
- [0043] 도 4a 및 4b는 로봇의 환경에 있는 사람의 위치를 찾아서 통신을 위한 로봇 텔레프레전스 세션을 요청하는 방법(400)의 더욱 상세한 예시를 제공하는 블록도이다. 방법(400)은 누가 기능을 수행하는지를 나타내는 세 개의 섹션들로 분할되며, 즉 도 4a는 원거리 사용자(402) 및 원거리 컴퓨팅 디바이스(404)에 의해 수행되는 기능을 도시하고, 도 4b는 로봇(406)에 의해 수행되는 기능을 도시한다.
- [0044] 블록(408)에서, 원거리 사용자(402)는 개인 컴퓨터, 태블릿 컴퓨터, 모바일폰, 또는 게임 시스템 등과 같은 원거리 컴퓨팅 디바이스(404) 상의 텔레프레전스 애플리케이션을 시작한다. 블록(410)에서, 원거리 사용자(402)는 특정 인물이 타겟화 되었는지 여부를 나타낸다. 예를 들어, 원거리 사용자(402)가 타겟 인물의 사진과 함께 있는 아이콘을 클릭할 수 있거나 또는 애플리케이션이 원거리 사용자가 특정 사람 또는 로봇(406)의 위치에 있는 임의의 사람을 타겟화 했는지를 원거리 사용자에게 질문할 수 있다.
- [0045] 블록(412)에서, 원거리 컴퓨팅 디바이스(404)는 전송된 바와 같이 인터넷, 모바일 캐리어 네트워크, 또는 임의의 다른 통신 네트워크를 통해 로봇에 접속한다. 블록(414)에서, 텔레프레전스 애플리케이션이 로딩되어 로봇(406) 또는 로컬 컴퓨터 상에서 시작될 수 있다. 블록(416)에서, 원거리 컴퓨팅 디바이스(404)는 로봇(406)이 특정한 타겟 인물을 찾아야만 하는지 또는 처음으로 마주친 사람이 텔레프레전스 요청을 수락할 수 있는지 여부를 로봇(406)에게 명시한다.
- [0046] 블록(418)에서, 로봇(406)은 스피커, 스크린, 프로젝터, 빛, 진동기 등과 같은 자신의 출력 디바이스를 이용하여 텔레프레전스 요청을 알린다. 블록(420)에서, 로봇(406)은 예를 들어 움직임을 찾거나, 얼굴을 찾거나, 또는 "로봇 수락," "로봇 거절," 또는 "이쪽으로"와 같은 특정한 음성 커맨드를 들음으로써 자신의 현재 위치로부터 자신이 있는 공간을 스캔하며 사람을 찾도록 시도한다. 유효한 사람의 얼굴은 로봇의 얼굴 인식 알고리즘에 의해 검출되는 사람 얼굴이다.
- [0047] 만약 블록(422)에서 로봇(406)이 사람들과 연관된 사운드 또는 커맨드를 식별하였다면, 블록(424)에서 로봇은 사운드를 마주하고 얼굴을 스캔하기 위해서 헤드 또는 소탑(turret)을 회전할 수 있다. 실시예에서, 로봇(406)은 사운드를 향해 이동할 수 있다. 프로세스 흐름은 블록(420)으로 복귀하여 공간을 계속해서 스캐닝한다.
- [0048] 만약 블록(422)에서 로봇(406)이 근처의 위치에서 얼굴 또는 사운드를 발견하지 못하였다면, 프로세스 흐름은 블록(426)으로 진행한다. 블록(426)에서, 로봇(406)이 공간에 대한 스캔을 완료하였는지 여부에 대한 판단이 이루어진다. 만약 그렇지 않다면, 프로세스 흐름은 블록(420)으로 진행하여 유효 얼굴 또는 음성 커맨드에 대한 스캔을 계속한다.
- [0049] 만약 블록(422)에서 로봇(406)이 유효 얼굴을 찾았다면, 예를 들어 RGB 이미지 데이터, 깊이 데이터로부터 파생되거나 또는 음성 또는 음성 커맨드의 사운드 소스 로컬화(SSL; sound source localization)로부터 파생된 위치 좌표를 사용함으로써 블록(428)에서 해당 얼굴을 향해 이동한다. 또한, SSL은 서로 다른 (예를 들어, 인접한) 공간 또는 다른 위치에 사람들이 존재한다고 결정하도록 사용될 수 있다. 피호출자에 대한 검색은 SSL에 의해 표시된 위치로 진행할 수 있다. 로봇(406)은 그 또는 그녀의 얼굴을 통해 해당 사람과 연결시킨다. 만약 블록(430)에서 로봇(406)이 특정 사람을 찾고 있지 않다고 판정되면, 즉 원거리 사용자가 로봇(406)에게 처음으로 발견된 사람이 타겟 인물로 간주된다고 명령하였다면, 프로세스 흐름은 블록(432)으로 진행한다.
- [0050] 블록(432)에서, 로봇(406)은 그 사람에게 텔레프레전스 세션을 수락할 수 있는지를 질문한다. 만약 블록(434)

에서 그 사람이 요청을 수락하였다고 결정되면, 프로세스 흐름은 블록(436)으로 진행하여 텔레프레전스 세션을 시작한다. 만약 그렇지 않다면, 프로세스 흐름은 블록(438)으로 진행하고 텔레프레전스 세션 요청이 종료된다.

[0051] 만약 블록(430)에서 로봇(406)이 특정 인물을 찾고 있다고 결정되면, 블록(440)에서 로봇(406)과 마주친 사람이 해당 인물인지 여부를 결정한다. 이것은 사람의 얼굴을 환경 내에 있을 가능성이 있는 사람들의 데이터베이스에 비교하는 얼굴 인식 알고리즘을 통해 그 사람의 신원을 확인함으로써 수행될 수 있다. 만약 타겟 인물이 식별되면, 프로세스 흐름은 전술된 바와 같이 블록(432)으로 진행한다. 만약 특정 인물이 블록(440)에서 식별되지 않았다면, 프로세스 흐름은 블록(442)으로 진행한다.

[0052] 블록(442)에서, 로봇(406)은 로봇(406)과 마주친 사람에게 타겟 인물의 위치를 질문한다. 만약 블록(444)에서 로봇(406)이 새로운 위치가 제공되지 않았다고 결정하면, 프로세스 흐름은 블록(446)으로 진행한다. 또한 만약 로봇(406)이 얼굴을 식별하지 않고 블록(426)에서 공간의 스캔을 완료하였다면 프로세스 흐름은 블록(446)으로 진행한다. 블록(446)에서, 새로운 검색 위치가 정의된다.

[0053] 새로운 검색 위치는 리스트 내에 있는 또는 근접한 다음 공간으로 이동하거나, 가장 가까운 공간으로 이동하는 것과 같은 간단한 이동 알고리즘을 이용함으로써 정의될 수 있다. 타겟 인물의 사무실로 이동하거나 시간대에 기초하여 특정 위치에서 타겟 인물을 만날 가능성이 가장 높은 위치로 이동하는 것과 같은 더욱 복잡한 알고리즘도 사용될 수 있다. 후자의 경우에서, 로봇은 위치 및 시각에 기초하여 사람들과 마주치는 빈도를 추적하는 데이터베이스를 필요로 할 수 있다. 로봇은 타겟 인물이 들렀던 마지막 위치를 판단하도록 데이터베이스 또는 다른 정보 소스를 이용할 수 있으며 해당 위치에서 검색을 시작할 수 있다.

[0054] 피호출자를 찾기 위한 다른 접근법은 텔레프레전스 세션 또는 호출이 개시되기를 기대할 수 있는 사람에 의해 로봇에게 제공되는 정보를 이용하는 것을 포함한다. 예를 들어, 누군가 로봇에게 호출하면 그/그녀가 소정의 위치에 있을 것임을 미리 명확하게 말할 수 있다. 로봇은 텔레프레전스 세션에 대한 호출 또는 요청이 수신되면 해당 위치에서 그 사람을 검색하기 시작할 수 있다. 피호출자를 찾는 또 다른 접근법은 특정 사람들을 마주칠 가능성이 높은 장소들의 히트 맵(heat map)을 이용할 수 있다. 피호출자를 어디에서 찾을 것인가에 대한 결정은 히트 맵으로부터의 데이터를 포함할 수 있다.

[0055] 블록(446)으로부터, 또는 만약 새로운 검색 위치가 블록(444)에서 제공되면, 프로세스 흐름은 블록(448)으로 진행한다. 블록(448)에서, 로봇(406)은 새로운 위치로 이동한다. 블록(450)에서 표시된 것과 같이, 로봇(406)은 새로운 위치로 이동하면서 얼굴을 스캔하고 음성 커맨드를 듣는다. 만약 유효한 얼굴 또는 음성 커맨드가 감지되면, 로봇(406)은 정지하고 프로세스 흐름은 새로운 입력을 처리하기 위해 블록(422)으로 진행한다. 블록(452)에서, 로봇은 새로운 위치에 도달하고 프로세스 흐름은 블록(418)에서 다시 시작하여 검색 절차를 반복한다. 로봇은 타겟 인물이 실제로 발견될 때까지 또는 사전설정된 시간 동안 검색 절차를 반복한다. 또한, 만약 타겟 인물이 텔레프레전스 세션을 수락하길 원치 않거나 타겟 인물이 아니라면, 로봇(406)을 다른 위치로 향하게 하거나, 다른 시간을 제안하거나, 또는 로봇(406)을 다른 사람에게 보낼 수 있다.

[0056] 도 5는 위치추적 기능을 수행하도록 사용될 수 있는 모듈을 나타내는 로봇(100)의 블록도이다. 도 1과 관련하여 기술된 바와 같이, 로봇(100)은 스토리지(120) 및 프로세서(122)를 포함한다. 도 1과 관련하여 기술된 바와 같이, 스토리지(120)는 프로세서(122)에 의해 실행가능한 모듈을 포함하는 비휘발성 스토리지 디바이스를 포함할 수 있다. 스토리지(120)는 또한 방향 그래프(directed graph)와 같은 환경의 맵(502)을 포함할 수 있으며, 여기에서 각 노드는 공간을 나타내고 각 에지는 두 공간을 연결하는 에지를 나타낸다. 맵(502)은 그리드(grid), 3차원 포인트 클라우드, 또는 임의의 수의 환경에 대한 다른 표현과 같은 측정된 정확한 정보를 포함할 수 있다. 맵(502)은 사용자에게 의해 생성될 수 있거나 또는 SLAM(simultaneous location and mapping) 기능 또는 환경의 사용자 지향 워크 네비게이션을 통한 환경의 탐사를 통해 로봇(100)에 의해 자동으로 생성될 수 있다.

[0057] 네비게이션 모듈(504)은 로봇(100)의 위치를 찾아서 로봇(100)을 새로운 위치로 향하게 하도록 맵(502)으로부터의 정보를 이용할 수 있다. 예를 들어, 타겟 위치는 사람에게 의해 제공된 음성 커맨드로부터 결정될 수 있으며 네비게이션 모듈(504)은 로봇(100)을 타겟 위치로 이동시키도록 모터를 활성화할 수 있다. 네비게이션 모듈(504)은 타겟 위치를 노드로서 생성할 수 있으며, 로봇(100)의 현재 위치로부터의 노드까지의 경로를 계산할 수 있다. 예를 들어, 맵(502)은 로봇(100)에 의해 복수의 서로 다른 노드로 해석될 수 있으며, 위치 지시 구성요소(304)는 로봇(100)의 현재 위치로부터 노드까지의 경로를 계산할 수 있고, 이러한 경로는 복수의 노드를 통과한다.

- [0058] 다른 실시예에서, 네비게이션 모듈(504)은 타겟 위치를 맵(502)과 비교할 수 있으며, 예를 들어 만약 맵(502)이 평면도 상의 좌표를 포함한다면 타겟 위치에 상응하는 좌표를 로봇(100)의 환경에 상응하는 좌표에 대해 식별할 수 있다. 그 다음 네비게이션 모듈(504)은 로봇(100)을 선택된 위치로 이동하게 할 수 있다. 예를 들어, 네비게이션 모듈(504)은 맵(502) 내의 태그되거나 태그되지 않은 위치인 타겟 위치를 포함하는 사람으로부터 음성 커맨드를 수신할 수 있다. 프로세서(122)에 의해 실행될 때, 네비게이션 모듈(504)은 로봇(100)이 환경 내의 현재 위치로부터 맵(502) 내의 선택된 위치에 해당하는 환경 내의 위치로 이동하게 할 수 있다.
- [0059] 로봇(100)이 선택된 위치를 향해서 이동할 때, 하나 이상의 장애물이 로봇(100)과 선택된 위치 사이의 경로에 존재할 수 있다. 장애물은 영구적이거나 사람이 로봇 앞의 사무실 밖으로 나가는 것과 같이 일시적인 것일 수 있다. 장애물 검출기 모듈(506)은 장애물 주변에서 로봇(100)을 조종할 수 있으며, 예를 들어 로봇(100)이 사람 또는 다른 장애물에 부딪히거나, 계단 아래로 떨어지는 등의 일을 방지한다. 스토리지(120)는 프로세서(122)에 의해 실행되었을 때 센서(116, 124)로부터 수신된 데이터를 분석하고 장애물을 검출하도록 구성되는 장애물 검출기 모듈(506)을 포함할 수 있다. 로봇(100)의 현재 위치와 선택된 위치 사이의 로봇(100)의 경로 내에서 장애물을 검출하면, 장애물 검출기 모듈(506)은 장애물이 존재한다는 표시뿐 아니라 로봇(100)의 현재 위치에 대한 장애물의 대략적인 위치를 출력할 수 있다. 예를 들어 네비게이션 모듈(504) 내의 방향 수정 구성요소는 장애물을 피하도록 로봇(100)의 코스를 변경할 수 있다. 따라서 네비게이션 모듈(504)은 예를 들어 텔레프레전스 세션에 대한 타겟 인물의 위치를 찾기 위해 환경 내에서의 위치를 찾도록 로봇(100)의 자체적인 구동과 결합하여 이용될 수 있다.
- [0060] 스토리지(120)는 또한 텔레프레전스 모듈(508)과 같은 텔레프레전스 세션을 지원하기 위해 로딩되는 다수의 모듈을 포함할 수 있다. 텔레프레전스 모듈(508)은 텔레프레전스 모듈을 요청하는 원거리 컴퓨팅 디바이스와의 통신이 개시되었을 때 로딩된다. 텔레프레전스 모듈(508)은 로봇(100)과 원거리 컴퓨팅 디바이스 사이의 통신 및 텔레프레전스 세션을 관리한다. 텔레프레전스 모듈(508)은 도 3 및 4와 관련하여 기술된 바와 같은 특정 기능을 수행하기 위해 다른 모듈을 활성화할 수 있다.
- [0061] 예를 들어, 텔레프레전스 모듈(508)은 로봇(100)의 환경 내의 사람의 위치를 찾기 위해 사람 위치추적 모듈(510)을 활성화할 수 있다. 사람 위치추적 모듈(510)은 카메라로부터의 이미지 내의 유효 얼굴을 식별하고 그 얼굴을 데이터베이스와 비교함으로써 특정 인물을 식별하기 위해 얼굴 식별 모듈(512)을 이용할 수 있다. 임의의 수의 알려진 얼굴 식별 알고리즘이 이용될 수 있으며, 이는 얼굴의 구별되는 특징을 데이터베이스에 저장된 것과 비교하는 기하학적 알고리즘 및 이미지를 나타내는 하나의 숫자 또는 연속적인 숫자들을 데이터베이스 내의 값에 비교하도록 통계적 알고리즘을 사용하는 광도계(photometric) 알고리즘을 포함한다. 실시예에서 사용될 수 있는 다양한 인식 알고리즘은 PCA(principal component analysis), 선형 식별 분석, 탄성 번치 그래프 매칭(elastic bunch graph matching), 히든 마코브모델(hidden Markov model) 및 신경망(neural network) 등을 포함한다.
- [0062] 사람 위치추적 모듈(510)은 검출된 음성 커맨드에 대한 위치 또는 방향을 식별하기 위해서 SSL(sound source location) 모듈(514)을 호출할 수 있다. 실시예에서, SSL 모듈(514)은 사운드의 위치를 삼각 측량하도록 로봇(100) 내의 다수의 마이크로폰을 이용한다. 다른 실시예에서, 마이크로폰, 스펙트럼 정보, 타이밍 분석, 상관관계 분석 및 패턴 매칭 등의 사이에서의 시간차, 레벨차 및 위상차를 포함하는 다양한 사운드 방향 알고리즘과 결합하여 두 개의 마이크로폰이 사용된다. SSL 모듈(514)은 이러한 알고리즘으로 한정되지 않으며, 임의의 수의 다른 기술들이 실시예에서 사용될 수 있다.
- [0063] 사람 위치추적 모듈(510)은 로봇에게 말한 커맨드를 식별하기 위해 음성 커맨드 모듈(516)을 호출할 수 있다. 음성 커맨드 모듈(516)은 커맨드를 식별하기 위해 임의의 수의 알려진 스피치 식별 알고리즘을 이용할 수 있다. 예를 들어, 음성 커맨드 모듈(516)에서 사용될 수 있는 하나의 알고리즘은 커맨드를 인식하기 위한 히든 마코브 알고리즘(hidden Markov algorithm)을 사용한다. 히든 마코브 모델은 예를 들어 약 10ms의 짧은 사운드의 구간 주파수 분산에 기초하여 개별적인 단어를 인식하도록 사용될 수 있다. 히든 마코브 모델의 결과는 컨텍스트 의존성 등과 같은 다양한 기술들에 의해 향상될 수 있다. 음성 커맨드 모듈(516)에 의해 스피치 인식에 사용될 수 있는 다른 알고리즘은 두 시퀀스 사이의 유사성을 측정하는 동적 타임 워핑(warping)에 기초한다.
- [0064] 임의의 수의 다른 모듈들이 전술된 기능성을 구현하거나 증강하도록 스토리지(120) 내에 존재할 수 있다. 예를 들어, 모션 검출 모듈이 포함될 수 있다. 모션 검출기 모듈은 환경 내의 모션을 식별하기 위해서 카메라 피드, 사운드 피드, 또는 임의의 다른 센서 피드를 사용할 수 있다. 이러한 경우에, 로봇은 사운드 또는 커맨드 입력을 향해 이동한 것과 동일한 방식으로 모션의 위치로 이동할 수 있다. 모션의 위치에 도달하면, 로봇은 얼굴을

스캔할 수 있다. 로봇(100)이 원거리 사용자에게 의해 직접 구동될 수 있게 하는 다른 모듈이 존재할 수도 있다.

[0065] 전술된 내용은 본 발명의 예시를 포함한다. 청구된 청구사항을 기술하기 위해 구성요소 또는 방법의 고려가능한 모든 조합을 모두 설명하는 것은 불가능하지만, 당업자는 본 발명의 다수의 다른 조합 및 치환이 가능하다는 것을 인식할 수 있을 것이다. 따라서, 청구된 청구사항은 첨부된 특허청구범위의 사상 및 범주 내에 포함되는 모든 변경, 수정 및 변형을 포함한다.

[0066] 특히 전술된 구성요소, 디바이스, 회로, 시스템 등에 의해 수행되는 다양한 기능들과 관련하여, 달리 언급되지 않는 한, 이러한 구성요소를 기술하도록 사용된 용어들은 ("수단"에 대한 지칭을 포함) 개시된 구조와 구조적으로 동일하지 않다고 해도, 본 명세서에 도시된 청구된 청구범위의 예시적인 측면들의 기능을 수행하는 (예로서, 기능적으로 동등한) 기술된 구성요소의 명시된 기능을 수행하는 임의의 구성요소에 상응하는 것이다. 이와 관련하여, 본 발명이 청구된 청구사항의 다양한 방법의 동작 및/또는 이벤트를 수행하기 위한 컴퓨터 실행가능 명령을 갖는 컴퓨터 판독가능한 저장 매체뿐 아니라 시스템을 포함한다는 것이 인식할 수 있을 것이다.

[0067] 본 발명을 구현하기 위해 예로서 적절한 API, 툴 키트, 드라이버 코드, 운영 시스템, 컨트롤, 독립형 또는 다운로드형 소프트웨어 객체 등과 같은 복수의 방식들이 존재하며, 이는 애플리케이션과 서비스가 본 명세서에 설명된 기술을 이용할 수 있게 한다. 청구된 청구사항은 API(또는 다른 소프트웨어 객체)의 관점으로부터 뿐 아니라 본 명세서에 개시된 기술에 따라 동작하는 소프트웨어 또는 하드웨어 객체로부터의 이용을 고려한다. 따라서, 본 명세서에 기술된 본 발명의 다양한 구현이 하드웨어 전체에서, 하드웨어와 소프트웨어에서 부분적으로, 그리고 소프트웨어에서의 측면들을 가질 수 있다.

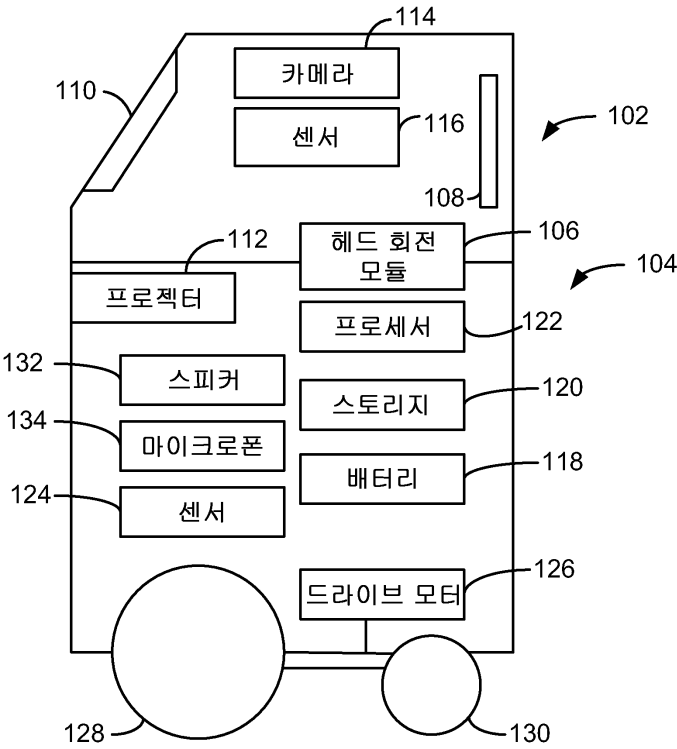
[0068] 전술된 시스템은 몇몇 구성요소들 간의 상호작용과 관련하여 기술되었다. 이러한 시스템 및 구성요소들이 언급된 구성요소들 또는 명시된 하위-구성요소들, 명시된 구성요소들 또는 하위-구성요소들의 일부, 및/또는 추가적인 구성요소를 포함할 수 있으며, 이들의 다양한 순열 및 조합에 따른다는 것이 이해될 것이다. 하위-구성요소는 또한 (계층적) 페어런트 구성요소 내에 포함되기보다는 다른 구성요소에 통신상 연결된 구성요소로서 구현될 수도 있다.

[0069] 또한, 하나 이상의 구성요소가 종합적인 기능을 제공하는 단일 구성요소로 결합될 수 있거나 또는 몇몇 개별적인 하위-구성요소로 분할될 수 있으며, 관리층과 같은 임의의 하나 이상의 중간층이 통합된 기능을 제공하기 위해서 이러한 하위-구성요소에 통신상 연결하도록 제공될 수 있음을 인지할 수 있을 것이다. 본 명세서에 기술된 임의의 구성요소가 또한 본 명세서에 특정하게 기술되지 않았지만 일반적으로 당업자에 의해 알려져 있는 하나 이상의 다른 구성요소와 상호작용할 수 있다.

[0070] 또한, 본 발명의 특정한 특성이 몇몇 구현들 중 오직 하나와 관련하여 개시되었을 수 있지만, 이러한 특성은 임의의 주어진 또는 특정한 애플리케이션에 대해 바람직하거나 유익할 수 있는 경우에 다른 구현의 하나 이상의 다른 특성들과 결합될 수 있다. 또한, "포함하는," "포괄하는," "구비하는," "내재하는"과 같은 용어들 및 이들의 변형어, 그리고 다른 유사한 용어들이 상세한 설명 또는 특허청구범위에서 사용되었으며, 이들 용어는 임의의 추가적인 또는 다른 요소를 배제하지 않는 열린 맥락의 단어로서 "포함하는"과 유사한 방식으로 포괄적으로 사용된 것이다.

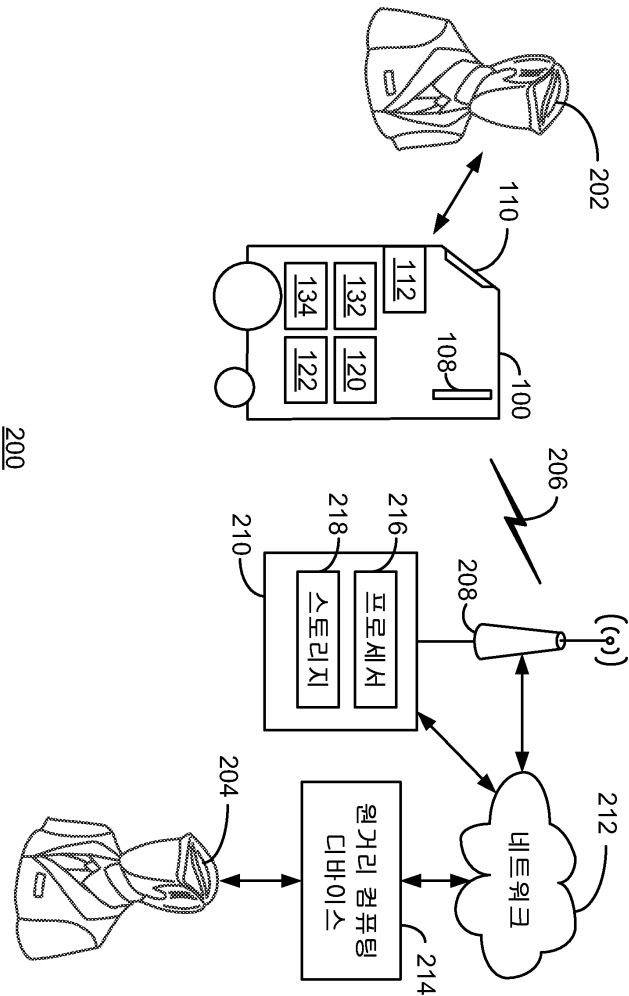
도면

도면1

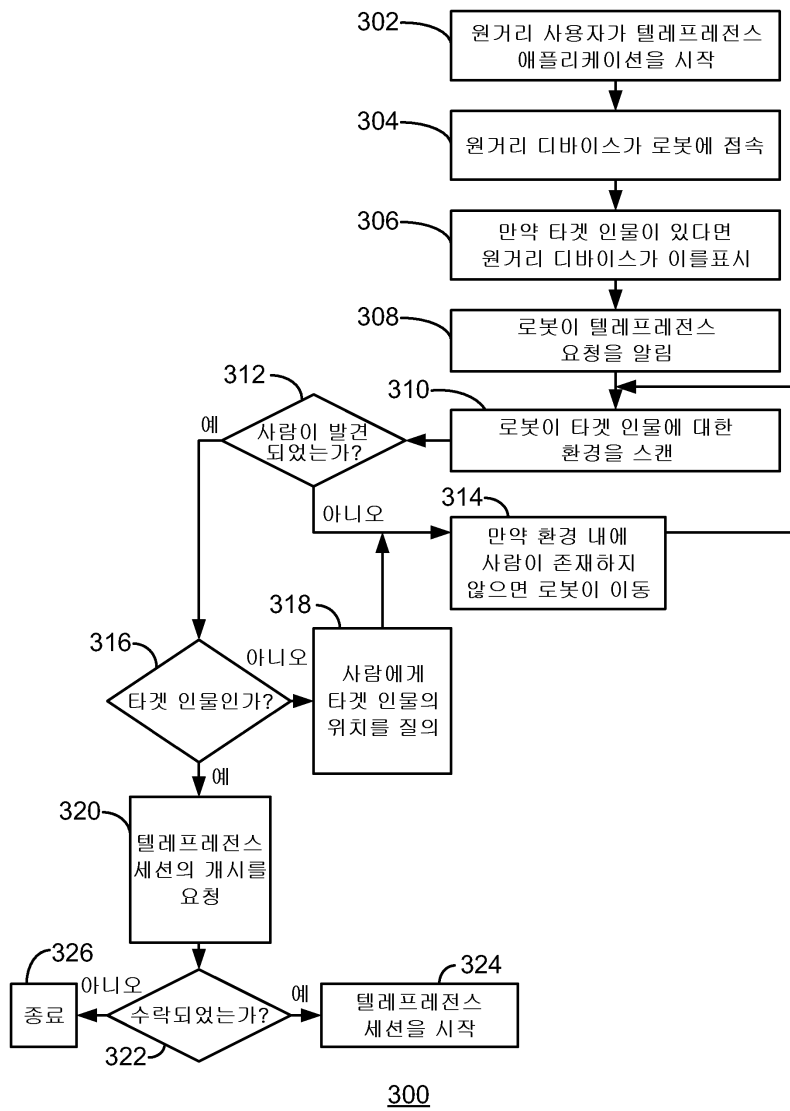


100

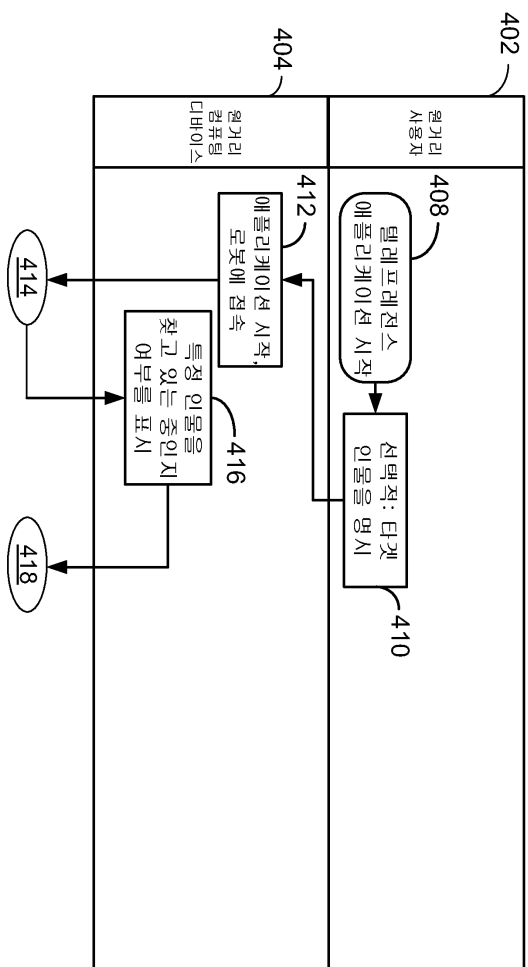
도면2



도면3

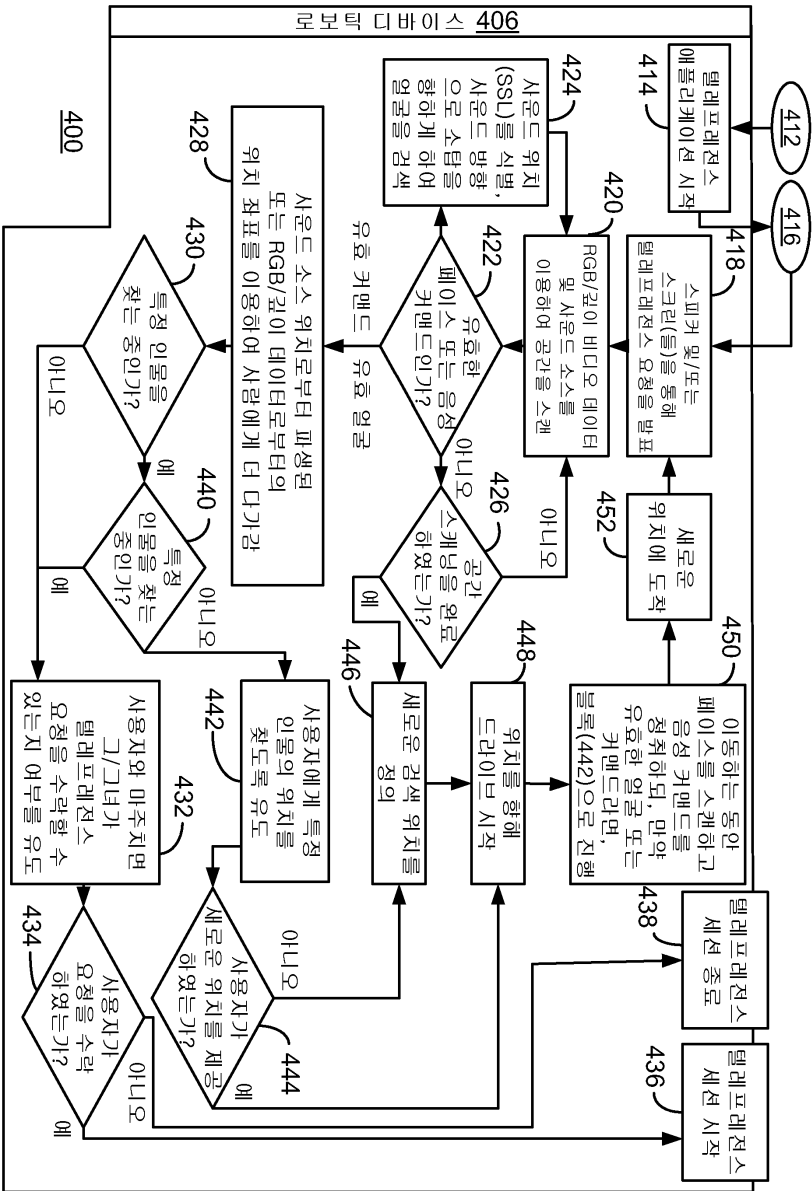


도면4a

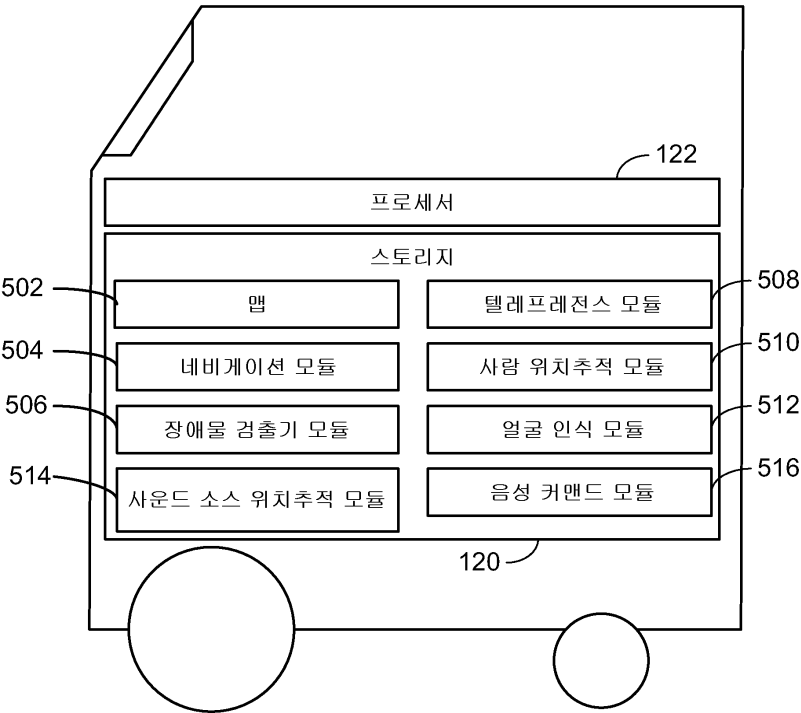


400

도면4b



도면5



100