



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2023년12월07일  
(11) 등록번호 10-2611751  
(24) 등록일자 2023년12월05일

- (51) 국제특허분류(Int. Cl.)  
G10L 15/22 (2006.01) G06F 3/16 (2018.01)  
G10L 15/08 (2006.01) G10L 17/10 (2013.01)  
G10L 17/24 (2013.01)
- (52) CPC특허분류  
G10L 15/22 (2013.01)  
G06F 3/167 (2013.01)
- (21) 출원번호 10-2022-7029524(분할)
- (22) 출원일자(국제) 2015년11월02일  
심사청구일자 2022년09월26일
- (85) 번역문제출일자 2022년08월25일
- (65) 공개번호 10-2022-0123153
- (43) 공개일자 2022년09월05일
- (62) 원출원 특허 10-2017-7015250  
원출원일자(국제) 2015년11월02일  
심사청구일자 2020년10월05일
- (86) 국제출원번호 PCT/US2015/058538
- (87) 국제공개번호 WO 2016/073321  
국제공개일자 2016년05월12일
- (30) 우선권주장  
62/074,562 2014년11월03일 미국(US)  
14/827,154 2015년08월14일 미국(US)
- (56) 선행기술조사문헌  
US20110184735 A1\*  
US20140249817 A1\*  
\*는 심사관에 의하여 인용된 문헌

- (73) 특허권자  
마이크로소프트 테크놀로지 라이선싱, 엘엘씨  
미국 워싱턴주 (우편번호 : 98052) 레드몬드 원  
마이크로소프트 웨이
- (72) 발명자  
로빈 앤드류 윌리엄  
미국 워싱턴주 98052-6399 레드몬드 원 마이크로  
소프트 웨이 마이크로소프트 테크놀로지  
라이선싱, 엘엘씨 어텐션:패턴트 그룹 도켓팅 (빌  
딩 8/1000)
- (74) 대리인  
김태홍, 김진희

전체 청구항 수 : 총 19 항

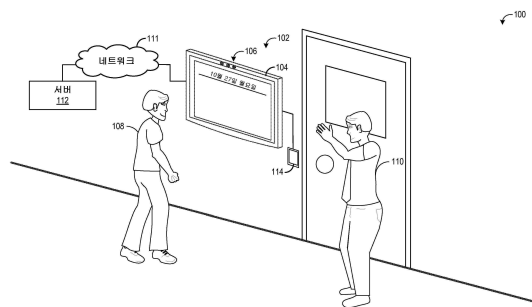
심사관 : 박금옥

(54) 발명의 명칭 키 문구 사용자 인식의 증강

(57) 요약

스피치를 통한 사용자 인식을 증강시키는 예가 제공된다. 하나의 예시적인 방법은, 컴퓨팅 디바이스 상에서, 음향 센서를 포함한 하나 이상의 센서를 통해 사용 환경을 모니터링하는 단계, 음향 센서로부터의 데이터를 통해 키 문구의 발언을 검출하는 단계, 음향 센서로부터의 선택된 데이터뿐만 아니라 음향 센서로부터의 선택된 데이(뒷면에 계속)

대표도



터 외의 상이한 시간에서 수집된 다른 환경 센서 데이터에 기초하여 키 문구가 식별된 유저에 의해 발생되었을 확률을 결정하는 단계를 포함한다. 상기 방법은 확률이 문턱 확률을 충족하거나 초과하면, 컴퓨팅 디바이스 상에 동작을 수행하는 단계를 더 포함한다.

(52) CPC특허분류

*G10L 15/08* (2013.01)

*G10L 17/10* (2013.01)

*G10L 17/24* (2013.01)

*G10L 2015/088* (2013.01)

*G10L 2015/227* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

방법으로서, 컴퓨팅 디바이스 상에서,

음향 센서에 의해 기록된 선택된 데이터를 수신하는 단계;

상기 음향 센서로부터의 상기 선택된 데이터를 통해, 키 문구(key phrase) 및 커맨드의 발언(utterance)을 스피치 인식을 통해 검출하는 단계;

식별된 사용자가 상기 키 문구를 발성했을 확률을 결정하기 위해, 상기 선택된 데이터에 대해 음성 인식을 수행하는 단계;

상기 식별된 사용자가 상기 키 문구 및 상기 커맨드의 발언 이전에 발성하고 있었는지 여부를 결정하기 위해, 상기 선택된 데이터 이전에 기록된 버퍼링된 음향 데이터에 대해 음성 인식을 수행하는 단계 - 상기 버퍼링된 음향 데이터에서 상기 식별된 사용자가 발성하고 있었다고 결정하는 것은 상기 식별된 사용자가 상기 키 문구를 발성했을 더 높은 확률에 대응함 -; 및

상기 확률이 문턱 확률을 충족하면, 상기 식별된 사용자를 인증하고 상기 컴퓨팅 디바이스 상에 상기 커맨드에 의해 지정되는 동작을 수행하는 단계

를 포함하는, 방법.

#### 청구항 2

제 1 항에 있어서,

상기 확률을 결정하는 것은 상기 키 문구를 상기 식별된 사용자의 음성 모델과 비교하는 것을 포함하는, 방법.

#### 청구항 3

제 1 항에 있어서,

이미지 데이터를 수신하는 단계 및 상기 이미지 데이터에 기초하여 사용 환경에서 한 명 이상의 사람을 식별하는 단계를 더 포함하고, 상기 확률을 결정하는 것은 상기 사용 환경에서 상기 한 명 이상의 사람의 결정된 신원(identity)에 적어도 부분적으로 기초하여 상기 확률을 결정하는 것을 포함하는 것인, 방법.

#### 청구항 4

제 1 항에 있어서,

위치 데이터를 수신하는 단계 및 상기 위치 데이터에 적어도 부분적으로 기초하여 상기 확률을 결정하는 단계를 더 포함하는, 방법.

#### 청구항 5

제 4 항에 있어서,

상기 위치 데이터는 근접 센서로부터의 근접 데이터를 포함하는 것인, 방법.

#### 청구항 6

제 5 항에 있어서,

상기 근접 센서는 액세스 카드 센서를 포함하는 것인, 방법.

#### 청구항 7

제 4 항에 있어서,

상기 위치 데이터는 상기 식별된 사용자에 대한 캘린더 정보(calendar information)를 포함하는 것인, 방법.

#### 청구항 8

제 1 항에 있어서,

사용자 행동 패턴을 검출하는 단계를 더 포함하고, 상기 확률을 결정하는 것은 상기 사용자 행동 패턴에 적어도 부분적으로 기초하여 상기 확률을 결정하는 것을 포함하는 것인, 방법.

#### 청구항 9

제 8 항에 있어서,

상기 사용자 행동 패턴은 상기 식별된 사용자가 발생하는 빈도에 관한 정보를 포함하는 것인, 방법.

#### 청구항 10

컴퓨팅 시스템에 있어서:

로직 머신; 및

상기 로직 머신에 의해 실행가능한 명령어들을 보유하는 저장 머신

을 포함하고,

상기 명령어들은,

음향 센서로부터 수신한 데이터를 통해 사용 환경을 모니터링하고;

상기 음향 센서로부터의 선택된 데이터를 통해, 커맨드가 뒤따르는 키 문구의 발언을 스피치 인식을 통해 검출하고;

상기 음향 센서로부터의 상기 선택된 데이터에 기초하여, 식별된 사용자가 상기 키 문구를 발성했을 확률의 결정을 획득하고;

상기 선택된 데이터 이전에 기록된 이력 데이터에 기초하여, 상기 키 문구 및 상기 커맨드의 발언 이전에 상기 식별된 사용자가 발성하고 있었는지 여부의 결정을 획득하고 - 상기 이력 데이터에서 상기 식별된 사용자가 발성하고 있었다는 결정의 획득은 상기 식별된 사용자가 상기 키 문구를 발성했을 더 높은 확률에 대응함 -;

상기 확률이 문턱 확률을 충족하면, 상기 식별된 사용자를 인증하고 상기 컴퓨팅 시스템 상에 상기 커맨드에 의해 지정되는 동작을 수행하도록

실행가능한 것인, 컴퓨팅 시스템.

#### 청구항 11

제 10 항에 있어서,

상기 명령어들은 상기 키 문구를 상기 식별된 사용자의 음성 모델과 비교함으로써 상기 확률의 결정을 획득하도록 실행가능한 것인, 컴퓨팅 시스템.

#### 청구항 12

제 10 항에 있어서,

상기 명령어들은 이미지 센서로부터 이미지 데이터를 수신하고 상기 이미지 데이터에 적어도 부분적으로 기초하여 상기 확률의 결정을 획득하도록 실행가능한 것인, 컴퓨팅 시스템.

#### 청구항 13

제 10 항에 있어서,

상기 명령어들은 액세스 카드 센서로부터 액세스 카드 데이터를 수신하고 상기 액세스 카드 데이터에 적어도 부

분적으로 기초하여 상기 확률의 결정을 획득하도록 실행가능한 것인, 컴퓨팅 시스템.

**청구항 14**

제 10 항에 있어서,

상기 명령어들은 캘린더 데이터를 수신하고 상기 캘린더 데이터에 적어도 부분적으로 기초하여 상기 확률의 결정을 획득하도록 실행가능한 것인, 컴퓨팅 시스템.

**청구항 15**

제 10 항에 있어서,

상기 명령어들은 센서 데이터로부터 사용자 행동 패턴을 검출하고 상기 사용자 행동 패턴에 적어도 부분적으로 기초하여 상기 확률의 결정을 획득하도록 실행가능한 것인, 컴퓨팅 시스템.

**청구항 16**

제 15 항에 있어서,

상기 사용자 행동 패턴은 상기 식별된 사용자가 발생하는 빈도에 관한 정보를 포함하는 것인, 컴퓨팅 시스템.

**청구항 17**

제 10 항에 있어서,

상기 음향 센서를 더 포함하는, 컴퓨팅 시스템.

**청구항 18**

컴퓨팅 시스템에 있어서,

음향 센서를 포함한 복수의 센서;

로직 머신; 및

상기 로직 머신에 의해 실행가능한 명령어들을 보유하는 저장 머신

을 포함하고,

상기 명령어들은,

상기 음향 센서를 포함한 상기 복수의 센서를 통해 사용 환경을 모니터링하고;

상기 음향 센서로부터의 선택된 데이터를 통해, 커맨드가 뒤따르는 키 문구의 발언을 검출하고;

상기 음향 센서로부터의 상기 선택된 데이터에 기초하여, 식별된 사용자가 상기 키 문구를 발성했을 확률을 결정하고;

상기 선택된 데이터 이전에 기록된 버퍼링된 센서 데이터에 기초하여, 상기 식별된 사용자가 상기 키 문구 및 상기 커맨드의 발언 이전에 발성하고 있었는지 여부를 결정하고 - 상기 버퍼링된 센서 데이터에서 상기 식별된 사용자가 발성하고 있었다고 결정하는 것은 상기 식별된 사용자가 상기 키 문구를 발성했을 더 높은 확률에 대응함 -;

상기 확률이 문턱 확률을 충족하면, 상기 식별된 사용자를 인증하고 상기 컴퓨팅 시스템 상에 상기 커맨드에 의해 지정되는 동작을 수행하도록

실행가능한 것인, 컴퓨팅 시스템.

**청구항 19**

제 18 항에 있어서,

상기 복수의 센서는 이미지 센서를 포함하고, 상기 버퍼링된 센서 데이터는 이미지 센서 데이터를 포함하는 것인, 컴퓨팅 시스템.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 키 문구 사용자 인식의 증강에 관한 것이다.

**배경 기술**

[0002] 컴퓨팅 시스템은 사용자 계정 정보, 컴퓨팅 시스템 상의 파일, 이메일 계정(들), 은행 계좌(들) 등과 같은 개인화된 그리고/또는 기밀의 정보로의 액세스를 허용하기 위해 다양한 사용자 식별 방법을 이용할 수 있다. 예를 들어, 사용자는 사용 세션을 시작하기 위해 사용자 이름 및 패스워드를 입력할 수 있고, 사용 세션 동안에 수행되는 동작은 그 사용자로부터 비롯되는 것으로 간주될 수 있다. 다른 방법으로는 그 사람의 음성 특성으로부터 사람을 식별하는 것을 포함한, 발성자 인식(speaker recognition)과 같은, 생체인식(biometric) 방법을 이용하는 것을 포함할 수 있다.

**발명의 내용**

[0003] 스피치(speech)를 통한 사용자 인식을 증강시키는 예가 제공된다. 하나의 예시적인 방법은, 컴퓨팅 디바이스 상에서, 음향 센서를 포함한 하나 이상의 센서를 통해 사용 환경을 모니터링하는 단계, 음향 센서로부터의 데이터를 통해 키 문구(key phrase)의 발언(utterance)을 검출하는 단계, 음향 센서로부터의 선택된 데이터뿐만 아니라 음향 센서로부터의 선택된 데이터 외의 상이한 시간에서 수집된 다른 환경 센서 데이터에 기초하여 키 문구가 식별된 유저에 의해 발생되었을 확률을 결정하는 단계를 포함한다. 상기 방법은 확률이 문턱 확률을 충족하거나 초과하면, 컴퓨팅 디바이스 상에 동작을 수행하는 단계를 더 포함한다.

[0004] 본 [발명의 내용]은 이하의 [발명을 실시하기 위한 구체적인 내용]에서 더 설명되는 단순화된 형태의 개념의 선택을 소개하기 위해 제공된다. 이 [발명의 내용]은 청구 대상의 주요 특징 또는 필수 특징을 식별하도록 의도되거나, 청구 대상의 범위를 제한하기 위해 사용되도록 의도되지 않는다. 또한, 청구 대상은 본 명세서의 임의의 부분에 언급된 임의의 또는 모든 단점을 해결하는 구현에 제한되지 않는다.

**도면의 간단한 설명**

[0005] 도 1a 및 도 1b는 키 문구 분석을 통한 사용자 인식을 위한 예시적인 사용 환경을 나타낸다.  
 도 2는 키 문구 검출을 통한 사용자 인식을 증강시키는 예시적인 방법을 설명하는 흐름도를 나타낸다.  
 도 3은 예시적인 컴퓨팅 시스템의 블록도를 나타낸다.

**발명을 실시하기 위한 구체적인 내용**

[0006] 스피치 입력에서 식별된 키 문구(key phrase)를 이용하는 음성 식별을 통한 사용자 인식과 관련된 예가 여기에 개시된다. 스피치 입력은, 커맨드가 뒤따르는, 키 문구라고도 지칭되는, 특정 단어, 단어 그룹, 소리(예를 들어, 웃음, 클릭 또는 다른 음향 사운드)의 형태를 취할 수 있다. 그러므로, 발성자(speaker) 식별을 위한 키 문구의 사용은 발성자가 식별되게 하고 동일한 발언(utterance)을 통해 컴퓨팅 시스템 기능이 호출될 수 있게 하며, 이는 직관적이고 효율적인 사용자 경험을 제공할 수 있다.

[0007] 그러나, 키 문구는 단지 몇 단어, 또는 심지어 단일 단어와 같이 상대적으로 짧을 수 있다. 키 문구에 걸친 음성 데이터의 지속시간은 짧기 때문에, 잘못된 거부(false rejection) 및/또는 잘못된 식별의 위험이 있다. 그러므로, 개시된 예는 키 문구의 발언 외에 하나 이상의 상이한 시간에서 수집된 환경 맥락 정보(environmental contextual information)로 키 문구 기반의 발성자 인식을 증강시키는 것에 관련된다. 예를 들어, 음성 식별을 수행하는 컴퓨팅 시스템은 키 문구의 발언 이전 및/또는 이후에 추가의 음향 데이터를 수집할 수 있다. 추가의 음향 데이터는 음성 식별에 의해 식별되는 발성자가 키 문구가 언급된 이전 또는 이후에 또한 발성하였는지(또는 발성하지 않았는지)를 결정하도록 분석될 수 있다. 분석이 발성자가 키 문구의 발언 이전 또는 이후에 발성하였다고 나타내면, 식별된 발성자가 키 문구를 발성했을 가능성이 더 높은 것으로 결정할 수 있다. 감지 디바이스를 통해 검출될 수 있고 키 문구를 발언한 사람의 신원을 나타낼 수 있는 임의의 환경 정보가 이용될 수 있다는 것이 이해될 것이다. 예로서, 음향 데이터, 이미지 데이터(예를 들어, 적외선 이미지 데이터, 가시 이미지 데이터, 심도 이미지 데이터 등), 위치 데이터, 모션 데이터, 및 후각 데이터를 포함하지만 그것에 한정되지 않

는다.

- [0008] 도 1a는 컴퓨팅 시스템이 키 문구 검출 및 환경 맥락 정보에 기초하여 사용자 인식을 수행할 수 있는 예시적인 사용 환경(100)을 나타낸다. 사용 환경(100)은 컴퓨팅 시스템(102)이 비어 있는 회의실 스케줄을 디스플레이하는, 회의실 외부의 복도와 같은, 사무실 환경으로서도 1에 도시된다. 예시된 사용 환경은 예시의 목적을 위해 제시되며 제한을 의도하는 것은 아니며, 개시된 예가 임의의 다른 적합한 사용 환경에서 사용될 수 있음을 이해할 것이다.
- [0009] 컴퓨팅 시스템(102)은 디스플레이 디바이스(104) 및 하나 이상의 센서(106)를 포함한다. 센서(106)는 발성자 인식 프로세스를 증강하기 위해 데이터를 감지하는 임의의 적절한 센서를 포함할 수 있다. 비제한적 예로서, 센서(106)는 하나 이상의 음향 센서(예를 들어, 마이크로폰 또는 마이크로폰 어레이) 및 하나 이상의 이미지 센서(예를 들어, 가시광 카메라, 심도 카메라, 적외선 카메라 등)를 포함할 수 있다. 또한, 사용 환경(100)은 컴퓨팅 시스템(102)으로부터 원격 위치된 추가의 센서를 포함할 수 있다. 비제한적 예로서, 사용 환경(100)은, RFID 액세스 카드와 같은, 전자적으로 판독가능한 태그의 근접을 감지하도록 구성된, RFID 센서(114)와 같은, 근접 센서를 포함할 수 있다. 컴퓨팅 시스템(102)은 또한 추가의 센서를 갖는 하나 이상의 외부 디바이스와 무선으로 통신할 수 있다. 예를 들어, 컴퓨팅 시스템(102)은 사용자(108) 및/또는 사용자(110)의 모바일 디바이스와 무선으로 통신하여 GPS 데이터, 모션 데이터 및 그 모바일 디바이스에 특정된 다른 데이터를 획득할 수 있다.
- [0010] 컴퓨팅 시스템(102)은 또한 사용자 계정 정보(예를 들어, 직원 정보)를 포함하는 서버(112) 등과 네트워크(111)를 통해 다른 컴퓨팅 시스템과 통신할 수 있다. 이러한 사용자 계정 정보는 인식된 사용자 각각에 대한 음성 패턴 데이터, 이미지 인식 데이터(예를 들어, 컬러 이미지, 심도 이미지, 적외선 이미지 등)와 같은 각각의 사용자에게 대한 식별 정보뿐만 아니라 다른 식별 정보(예를 들어, 액세스 카드 정보)를 포함할 수 있다. 컴퓨팅 시스템(102)에 관한 추가 정보는 도 3과 관련하여 이하에 제시된다
- [0011] 컴퓨팅 시스템(102)은 사용자 음성 입력에 응답하여 다양한 동작을 수행하도록 구성될 수 있다. 예를 들어, 도 1a에 도시된 바와 같이, 제 1 사용자(108)는 예를 들어, 키 문구 및 커맨드를 포함하는 스피치 입력을 수행함으로써 컴퓨팅 시스템이 제 1 사용자와 제 2 사용자 간의 미팅을 스케줄링하도록 요청할 수 있다. 특정 예로서, 제 1 사용자(108)가 키 문구 "헤이 어시스턴트(hey Assistant)"를 발성하고, 커맨드 "오늘 오후 3시에 1시간 동안 준과 나의 미팅을 예약해줘"가 뒤따를 수 있다. 컴퓨팅 시스템(102)은 센서(106)에 의해 수신된 음향 데이터로부터 키 문구를 검출하고, 키 문구를 사용하여 누구에게 방(room)을 예약할지를 결정하는 커맨드를 내리는 발성자의 신원을 인식할 수 있다. 또한, 맥락적 환경 데이터는 "존"을 제 2 사용자(110)로서 식별하는데 사용될 수도 있다.
- [0012] 그러나, 상술한 바와 같이, 키 문구의 간결성을 고려할 때, 컴퓨팅 시스템이 키 문구 데이터를 단독으로 사용하여 키 문구의 발성자를 정확하게 식별하는 것은 어려울 수 있다. 이로 인해 잘못된 긍정 또는 잘못된 부정 식별의 위험을 발생시킬 수 있다. 따라서, 발성자의 결정된 신원이 맞을 가능성을 증가시키기 위해, 컴퓨팅 시스템(102)은 키 문구의 발언 이전 및/또는 이후에 수집된 환경 맥락 데이터를 사용하여 키 문구 발성자 식별을 증강시킬 수 있다. 이러한 환경 데이터는 센서(106)를 통해 및/또는 시스템(102) 외부의 하나 이상의 컴퓨팅 시스템을 통해 수집될 수 있다. 예를 들어, 센서(106)에 의해 수집된 음향 데이터는 키 문구의 발언 이전 및/또는 이후에 센서(106) 인근에서 발성한 사람들의 신원을 결정하는데 사용될 수 있다. 이것은 키 문구가 발생되었을 때 그 환경에 있는 사용자의 신원에 관한 정보를 제공할 수 있고, 그에 따라 그 사람들 중 하나에 의해 키 문구가 발생되었을 확률을 증가시킬 수 있다. 다른 예로서, 센서(106)에 의해 수집된 이미지 데이터는 컴퓨팅 시스템(102)에 의해 사용가능한 이미지 정보를 획득하여 키 문구의 발언 이전 및/또는 이후에 사용 환경에서 존재하였던 임의의 사용자의 신원을 결정할 수 있다. 추가의 예로서, (예를 들어, 사용자의 모바일 디바이스로부터의 GPS 센서 데이터를 통해, RFID 액세스 카드 등을 통해 결정되는 바와 같은) 사용자 위치 데이터는 발성자 식별 프로세스를 증강하기 위해 사용될 수 있다.
- [0013] 환경 맥락 데이터는 임의의 적절한 방식으로 키 문구 발성자 식별을 증강하는데 사용될 수 있다. 예를 들어, 컴퓨팅 시스템(102)은 환경 맥락 데이터를 확률 결정을 위한 입력으로서 사용할 수 있다. 보다 구체적인 예로서, 센서(106)에 의해 수집된 음향 데이터는 제 1 사용자(108)가 키 문구의 발언 이전 및/또는 이후에 제 1 사용자(108)가 발생하였음을 검출하기 위해 시스템(102)에 의해 사용될 수 있고(예를 들어, 제 1 사용자(108)는 키 문구를 발언하기 전에 디스플레이 장치(104) 상의 일기 예보를 보도록 요청될 수 있다.), 이는 제 1 사용자(108)가 키 문구를 발언하였을 확률을 증가시킬 수 있다. 다른 예에서, 컴퓨팅 시스템(102)의 이미지 센서는 키 문구가 검출될 때 제 1 사용자(108) 및 제 2 사용자(110) 모두가 사용 환경(100)에 존재하였음뿐만 아니라 키 문구

가 발언되었을 때 제 2 사용자(110)가 컴퓨팅 시스템(102)을 단지 지나갔음(예를 들어, 이미지 정보는 제 2 사용자(110)가 키 문구의 발언에 이어서 사용 환경을 떠났음을 나타낼 수 있다)을 검출할 수 있다. 따라서, 제 2 사용자(110)가 키 문구를 발언하였을 확률은 이 데이터가 고려되지 않았던 경우보다 낮을 수 있다.

[0014] 도 2는 키 문구 검출을 증상시키는 방법(200)의 일례를 도시하는 흐름도이다. 방법(200)은 도 1의 컴퓨팅 시스템(102)과 같은 컴퓨팅 디바이스에 의해 수행될 수 있다. 202에서, 방법(200)은 음향 센서를 포함한 하나 이상의 센서로 사용 환경을 모니터링하는 단계를 포함한다. 상술한 바와 같이, 컴퓨팅 디바이스 상에 위치한 센서 및/또는 컴퓨팅 디바이스로부터 원격 위치한 센서와 같은 임의의 적절한 센서가 사용될 수 있다. 음향 센서에 더하여, 다른 센서의 예로는 이미지 센서, 위치 센서 및 모션 센서를 포함하지만 이에 한정되지 않는다. 일례에 있어서, 하나 이상의 센서는 사용자의 존재, 사용자와 컴퓨팅 디바이스 간의 상호작용 등에 관계없이 사용 환경을 지속적으로 모니터링할 수 있다. 다른 예에서, 하나 이상의 센서는 활동이 검출될 때까지 저전력 모드로 작동하고, 그 다음에 어떤 트리거링 활동(예를 들어, 이미지 데이터에서 사람의 모션 및/또는 존재, 음향 데이터에서 사람의 음성의 존재, 근접 센서를 통해 근접한 사람의 검출)을 검출하면 고전력 모드에서 환경 맥락 데이터를 획득할 수 있다. 하나 이상의 센서에 의해 수집된 환경 맥락 정보는 컴퓨팅 디바이스의 메모리 디바이스에 로컬 저장되고/저장되거나 원격 디바이스로 전송될 수 있다.

[0015] 204에서, 방법(200)은 선택된 음향 센서 데이터로부터 키 문구의 발언을 검출하는 단계를 포함한다. 상술한 바와 같이, 키 문구는 가능성있는 사용자들의 세트 중에서 사용자를 식별하기 위해 컴퓨팅 디바이스에 의해 사용될 수 있다. 또한, 일부 예에 있어서, 키 문구의 발언은 컴퓨팅 디바이스에 의해 수행될 동작을 호출하도록 구성된 커맨드가 뒤따를 수 있고, 여기서 동작은 식별된 사용자와 관련된다. 키 문구가 임의의 적절한 스피치 분석 파이프라인을 사용하여 검출될 수 있음이 이해될 것이다.

[0016] 계속해서, 206에서, 방법(200)은 208에서 나타난 바와 같이, 음향 센서로부터의 데이터에 기초하여 키 문구가 식별된 사용자에게 의해 발생되었을 확률을 결정하는 단계를 포함한다. 일부 예에서, 컴퓨팅 디바이스는 전적으로 키 문구에 기초하여 사용자를 식별할 수 있다. 그러나, 컴퓨팅 디바이스가 전적으로 키 문구에 기초하여 사용자를 식별할 수 없는 경우, 또는 보다 높은 확률의 식별이 요구되는 경우, 수집된 데이터의 추가 분석이 행해질 수 있다. 따라서, 확률은 또한 210에서 나타난 바와 같이, 상이한 시간에서 수집된 다른 환경 센서 데이터에 기초하여 결정될 수 있다.

[0017] 도 1과 관련하여 상술한 바와 같이, 다른 환경 센서 데이터는 키 문구의 발생 이전 및/또는 이후에 수집된 음향 데이터를 포함할 수 있다. 예를 들어, 컴퓨팅 디바이스는 음향 데이터를 연속적으로 수집하여 음향 데이터를 버퍼에 저장할 수 있다. 키 문구가 검출되면, 컴퓨팅 디바이스는 키 문구가 식별된 사용자에게 의해 발생되었을 확률과 관련된 정보에 대해 저장된 음향 데이터를 분석할 수 있다. 그러한 정보는 키 문구의 발언 이전 및/또는 이후의 시간대 내에(예를 들어, 몇 분 내에서) 식별된 사용자가 발언하였는지를 포함할 수 있지만, 이에 한정되지 않는다.

[0018] 보다 구체적인 예로서, 컴퓨팅 디바이스는, 컴퓨팅 디바이스가 키 문구를 발언할 가능성이 있는 것으로 식별하는 사용자가 키 문구의 발언 이전 및/또는 이후에 또한 발생하였는지를 결정하기 위해 추가의 음향 데이터를 분석할 수 있다. 분석이 식별된 사용자가 키 문구의 발언 이전 또는 이후에 발생하였다고 나타내면, 컴퓨팅 디바이스는 식별된 사용자에게 의해 키 문구가 발생되었을 확률을 증가시킬 수 있고, 반면에 분석이 식별된 사용자가 키 문구의 발언 이전 또는 이후에 발생하지 않았다고 나타내면, 컴퓨팅 디바이스는 식별된 사용자에게 의해 키 문구가 발생되었을 확률을 감소시킬 수 있다.

[0019] 다른 환경 센서 데이터는 또한 키 문구의 발언 검출 이전 및/또는 이후에 수집된 이미지 데이터를 포함할 수 있다. 예를 들어, 이미지 정보는 식별된 사용자가 키 문구가 발생된 이전 및/또는 이후에 컴퓨팅 디바이스와 동일한 방에 있었는지의 여부를 나타낼 수 있다. 또한, 다른 환경 센서 데이터는 컴퓨팅 디바이스로부터 원격 위치한 센서로부터 수집된 위치 데이터를 포함할 수 있다. 예를 들어, 식별된 사용자는 컴퓨팅 디바이스와 통신하는 위치 센서(예를 들어, GPS)를 갖는 휴대용 컴퓨팅 디바이스를 가질 수 있고, 위치 데이터는 키 문구가 검출되기 이전 또는 이후에 식별된 사용자가 있는 곳을 나타낼 수 있다. 이러한 위치 정보는 또한 RFID 데이터 또는 근접 센서로부터의 다른 근접 데이터, 식별된 사용자(예를 들어, 컴퓨팅 디바이스와 동일한 방에 있는 것으로 스케줄링된 사용자)의 캘린더 정보(calendar information), 및/또는 위치와 관련된 임의의 다른 적절한 데이터를 포함할 수 있다.

[0020] 보다 구체적인 예로서, 컴퓨팅 디바이스는 사용자의 캘린더 정보에 기초하여 키 문구의 발언이 검출된 시간 동안 사용자가 주어진 사용 환경(예를 들어, 회의실)에 있는 것으로 스케줄링되었는지를 결정할 수 있다. 사용자



가 사용 환경에 있는 것으로 스케줄링되었다면, 컴퓨팅 디바이스는 그 사용자에게 의해 키 문구가 발생되었을 확률을 증가시킬 수 있다. 반면에, 사용자의 캘린더 정보가 사용자가 사용 환경에 있는 것으로 스케줄링되지 않았다고 나타내면, 컴퓨팅 디바이스는 키 문구가 그 사용자에게 의해 발생되었을 확률을 낮출 수 있다.

[0021] 일부 구현 예에 있어서, 환경 데이터는 확률 결정을 위한 추가 입력으로서 사용될 수 있는 사용자 행동 패턴을 검출하기 위해 분석될 수 있다. 예를 들어, 사용자가 키 문구의 발언 이전에 발생하였는지를 결정하기 위해 음향 데이터가 분석되고, 음성 데이터의 분석은 식별된 사용자가 많이 발생하거나 조금 발생하는 경향이 있는지를 결정하기 위해 사용될 수 있다. 이 정보는 확률 결정을 위한 입력으로서 사용될 수 있고, 키워드 발언 이전 및/또는 이후에 사용자가 발생하였는지에 관한 정보가 사용자가 발생하는 경향의 빈도에 기초하여 가중될 수 있다. 사용자가 키 문구를 발언하였는지의 확률을 가중하는데 사용될 수 있는 추가의 사용자 행동 패턴 및/또는 환경 데이터는 사용자가 위치되는 환경을 포함한다. 예를 들어, 사용자는 회의실보다는 자신의 사무실에서 더 많이 발생하고, 공원 등보다는 회의실에서 더 많이 발생하는 경향이 있을 수 있다. 하나의 보다 구체적인 예에서, 사용자 행동 패턴은 환경 감지(예를 들어, 상술된 센서에 의해 수집된 데이터로부터)를 통해 검출된 이전의 사용자 행동에 기초하여 컴퓨팅 디바이스에 의해 결정될 수 있다. 사용자 행동 패턴은 사용자가 발생하는 빈도의 측정치를 포함할 수 있고, 컴퓨팅 디바이스는 식별된 사용자가 발생하는 빈도의 측정치에 기초하여 확률을 조정할 수 있다. 예를 들어, 사용자가 그 자신이 발생하는 평균 빈도보다 상대적으로 낮게 발생하면 확률은 감소될 수 있다.

[0022] 도 2에 계속해서, 방법(200)은 212에서 식별된 사용자가 키 문구를 발언하였을 결정된 확률은 문턱 확률에 비교하는 단계를 포함한다. 문턱 확률은 식별된 사용자가 실제로 키 문구를 발언하였다는 소망의 레벨의 확신을 나타내는 임의의 적절한 문턱일 수 있다. 일부 예에서, 문턱 확률은 수치(예를 들어, 0.95 또는 95%)일 수 있는 반면, 다른 예에서 문턱 확률은 모든 가능성 있는 식별된 사용자의 등급일 수 있다(예를 들어, 확률이 식별된 사용자가 모든 가능성 있는 사용자 중 최고 확률을 가졌음을 나타내는 경우, 확률은 문턱 확률을 충족 또는 초과할 수 있다). 상이한 문턱이 상이한 유형의 액세스 제한에 적용될 수 있다는 것이 이해될 것이다. 예를 들어, 발생자 식별에 기초하여 특정 사용자 프로필을 갖는 비디오 게임을 론칭하는 것보다 더 높은 문턱 확률이 발생자 식별에 기초하여 बैं킹 정보에의 액세스를 승인하는 것에 적용될 수 있다.

[0023] 확률이 문턱 확률을 충족 또는 초과하는 경우, 방법(200)은 214에서 컴퓨팅 디바이스 상에 동작을 수행하는 단계를 포함한다. 동작은 식별된 사용자가 컴퓨팅 디바이스 및/또는 컴퓨팅 디바이스를 통해 개인화된 정보에 액세스하는 것을 허용하기 위해 식별된 사용자를 인증하는 단계를 포함한다. 또한, 일부 실시예에서, 키 문구의 발언은 미팅 스케줄링과 같은 특정 동작을 수행하도록 컴퓨팅 디바이스를 위한 커맨드가 뒤따를 수 있다. 그러한 경우에, 문턱을 충족 또는 초과하는 확률에 응답하여 수행된 동작은 식별된 사용자에게 의해 커맨드된 동작을 수행하는 단계를 포함할 수 있다.

[0024] 한편, 확률이 문턱 확률을 충족 또는 초과하지 않는 경우, 방법(200)은 216에서 나타낸 바와 같이, 동작을 수행하지 않는 단계를 포함한다. 그러나, 일부 예에서, 하나보다 많은 사용자가 키 문구를 잠재적으로(potentially) 발언함으로써 식별될 수 있다. 그러므로, 218에서 방법(200)은 키 문구를 잠재적으로 발언함으로써 다른 사용자가 식별되었는지를 결정하는 단계를 포함한다. 다른 사용자가 식별되지 않았다면, 방법(200)은 종료될 수 있다. 다른 가능성 있는 사용자가 식별되었다면, 방법은 206으로 루프백(loop back)하여 하나 이상의 추가의 식별된 사용자에게 대한 확률을 결정할 수 있다.

[0025] 상술된 키 문구 발언의 검출 및 식별된 사용자 확률은 키 문구가 지시되고(directed) 있는 컴퓨터 디바이스 상에 수행되는 것으로서 설명된다. 그러나, 일부 예에서 환경 센서 데이터(음향 센서 데이터를 포함함)는 원격 디바이스로 전송될 수 있고, 키 문구 검출 및 사용자 식별이 원격 디바이스에 의해 수행된다. 다음에, 원격 디바이스는 컴퓨팅 디바이스로 식별된 사용자의 표시를 전송할 수 있다.

[0026] 상술된 키 문구 발생자 인식은 음향 데이터를 이용하지만, 키 문구 검출은 시각적 양상(modality)과 같은 대안의 또는 추가의 양상을 포함할 수 있다는 것이 이해될 것이다. 오디오 기반의 식별을 수행하기 위해 오디오 스트림의 템플릿 매칭(template matching)이 수행될 수 있거나, 오디오 스트림의 분석 및 사용자를 위한 모델에 대한 비교 또는 "사용자 아님"에 대한 비교가 수행될 수 있다. 분석 및 확률 결정은 컴퓨팅 디바이스 상에 로컬적으로 수행될 수 있거나, 예를 들어 서버 상에 원격으로 수행될 수 있다. 일부 실시예에 있어서, 입력 오디오 스트림은 키 문구 검출을 수행하기 전에 조정될 있다(예를 들어, 노이즈 억제 또는 반향 소거).

[0027] 일례에서, 컴퓨팅 디바이스는 수집된 오디오 입력을 위한 메모리 버퍼를 연속적으로 저장할 수 있고, 메모리는 로컬이거나, 클라우드에 있거나, 또는 제 2 디바이스 상에 있다. 디바이스는 키 문구를 발언하는 사용자를 검출

할 수 있다. 그 다음에, 디바이스는 키 문구를 말한 사용자를 결정한 다음, 그 이력을 검토하고 사용자가 이전에 (어떤 시간대에) 발생하였는지를 결정할 수 있다. 그렇다면, 사용자가 키 문구를 말하였을 확률이 증가된다. 그렇지 않다면 확률은 낮아진다. 또한, 이것은 이력에서 다른 사용자와 함께 수행될 수 있다, 예를 들어 (발성자 식별 시스템의) 이력에서 3명의 사용자가 있다면 3명의 사용자의 검출의 확률의 모두는, 사용자 중 어느 한 명이 계속 대화하고 있을 가능성이 많으므로, 증가할 것이다. 이력 버퍼에서의 배치(또는 제때의 키 문구에 대한 근사)는 또한 확률을 조정하기 위해 사용될 수 있다.

[0028] 상기 예에서, 사용자 대화의 확률은 이력에서 사용자의 입 움직임의 시각적 신호(cue)에 의해 증강될 수 있다. 일례에서, 이력에서의 사용자 대화와 키 문구에 대한 사용자 대화 모두의 공동 확률은 서로에 의해 증강된다. 예를 들어, 컴퓨팅 디바이스는 사용자가 이력에 걸쳐 키 문구를 대화하였던 결정된 최고의 확률을 취하고, 이 확률은 다음의 통행 상의 전체 오디오 스트림에 걸쳐 사용자 확률을 증강시키는데 사용할 수 있다. 이는 확실한 승자가 선택되거나 특정 수 반복이 지나갈 때까지 계속될 수 있다.

[0029] 다른 예에 있어서, 컴퓨팅 디바이스는 미팅이 발생하는 방 내에 위치될 수 있다. 컴퓨팅 디바이스는 키 문구를 검출한 다음에, 식별된 사용자가 방 대화에 있는지를 결정하기 위해 오디오 이력을 분석한다. 그렇지 않다면, 컴퓨팅 디바이스는 키 문구를 발언하였던 사용자의 확률을 낮출 수 있다. 그러나, 사용자에 대한 메타(meta) 정보는 사용자가 드물게 대화한다고 나타낼 수 있고, 그래서 사용자의 확률은 조정되지 않거나 소량만 조정될 수 있다. 이 정보를 학습하는 기능은 실시간으로(컴퓨팅 디바이스를 사용하여) 또는 오프라인으로 이루어질 수 있고, 미팅, 사용자 또는 시나리오 상에 투가의 메타데이터로서 저장될 수 있다. 가능성 있는 사용자들이 내부 미팅에 참가하는 외부 파트너들인 시나리오에서, 머신 러닝(machine learning)은 외부 사용자들이 많이 대화하거나 적게 대화할 가능성이 많음을 나타낼 수 있다. 그 다음에, 이 정보는 외부 사용자가 대화하였는지의 여부를 조정하기 위해 발생자 식별 결정으로 주어진다. 마찬가지로, 메타 정보는 사용자가 (그의 캘린더에 기초하여) 방에 있어야 함을 나타낼 수 있고, 확률은 그에 따라 조정될 수 있다. 일례에 있어서, 방 내의 카메라는 방 내의 사용자의 부채를 식별하거나 확인할 수 있고, 그에 따라 사용자 확률을 낮추거나 사용자를 고려사항으로부터 완전히 제거할 수 있다.

[0030] 여전히 추가의 예에서, 컴퓨팅 디바이스는 키 문구를 검출할 수 있고, 센서 이력은 사용자가 대화하고 있는 것으로 확인될 때 사용자에 대한 모델을 업데이트하기 위해 사용될 수 있다. 예를 들어, 사용자는 (예를 들어, 컴퓨팅 디바이스를 통해) 원격 위치된 사용자와 대화할 수 있고, 사용자는 키 문구를 말할 수 있다. 그 다음에, 이력 버퍼는 사용자가 (오디오 및 시각적 신호를 통해) 대화할 때의 정보를 가진다. 그 다음에, 그것은 이력에서 빼내어 사용자의 발생자 식별 모델을 업데이트하기 위해 사용된다. 그것은 사용자가 감거나 어떤 이유로 그들 음성이 상이한 상황에서 유용할 수 있다. 이 예는 새로운 사용자를 동적으로 식별하는데 사용될 수 있거나, 또는 새로운 모델을 생성하거나 새로운 디바이스 상에서 사용될 수 현재의 것으로 증강시키기 위해 구현될 수 있기 때문에 사용자가 마이크로폰을 변경하였을 때 사용될 수 있다.

[0031] 다른 예에서, 컴퓨팅 디바이스는 회의실에 위치하여 대화를 기록할 수 있다. 컴퓨팅 디바이스는 오디오 센서만을 포함할 수 있다(따라서 이미지 또는 다른 센서를 포함하지 않을 수 있다). 전체 대화는 오프라인으로 기록되고 처리될 수 있다. 디바이스는 누가 방에 있는지 알지 못하기 때문에 기록을 분석하기 위해 큰 발생자 모델 세트를 사용할 수 있다. 키 문구를 찾을 때, 디바이스는 사람의 모델이 정확한 사람 대화로서 평가된 빈도를 고려할 수 있다. 예를 들어, 사용자의 모델은 가장 높은 확률일 수 있다. 그러나, 이전 및 다음 5분의 분석은 사용자가 다시 대화하지 않는다는 것을 나타낼 수 있다. 그 다음에, 시스템은 그 키 문구에 대해 그 사용자의 확률을 낮출 수 있습니다. 그 다음에, 디바이스는 이전 및 미래의 5분을 분석하여 두번째로 높은 확률 사용자가 그 시간대 내에서 대화한다고 결정할 수 있다. 그 다음에, 디바이스는 가장 높은 확률의 사용자와는 대조적으로 두번째로 높은 확률의 사용자가 방에 있음을 높은 가능성으로 결정할 수 있다.

[0032] 컴퓨팅 디바이스는 머신 러닝을 사용할 수 있고, 이러한 환경에서, 사용자가 키 문구를 말하지만 키 문구를 발언하기 전 또는 후에 어떠한 것도 말하지 않을 가능성은 극히 낮은 확률을 갖고, 그래서 디바이스는 키 문구 중에 대화하고 있었던 사용자를 평가할 때 키 문구 이전 및 이후에 대화할 가능성이 있는 사용자만 평가할 수 있다.

[0033] 일부 실시 예에서, 여기에 설명된 방법 및 프로세스는 하나 이상의 컴퓨팅 디바이스의 컴퓨팅 시스템에 관련될 수 있다. 특히, 이러한 방법 및 프로세스는 컴퓨터 애플리케이션 프로그램 또는 서비스, 애플리케이션 프로그래밍 인터페이스(API), 라이브러리 및/또는 다른 컴퓨터 프로그램 제품으로서 구현될 수 있다.

[0034] 도 3은 상술된 방법 및 프로세스 중 하나 이상을 제정할 수 있는 컴퓨팅 시스템(300)의 비제한적인 실시예를 개

략적으로 도시한다. 컴퓨팅 시스템(300)은 단순화된 형태로 도시된다. 컴퓨팅 시스템(300)은 하나 이상의 개인용 컴퓨터, 서버 컴퓨터, 태블릿 컴퓨터, 홈 엔터테인먼트 컴퓨터, 네트워크 컴퓨팅 디바이스, 게임 디바이스, 모바일 컴퓨팅 디바이스, 모바일 통신 디바이스(예를 들어, 스마트폰) 및/또는 기타 컴퓨팅 디바이스의 형태를 취할 수 있다. 컴퓨팅 시스템(102)은 컴퓨팅 시스템(300)의 비제한적인 예일 수 있다.

[0035] 컴퓨팅 시스템(300)은 로직 머신(302) 및 저장 머신(304)을 포함한다. 컴퓨팅 시스템(300)은 도 3에 나타난 바와 같이 디스플레이 서브시스템(306), 입력 서브시스템(308), 통신 서브시스템(314) 및/또는 다른 컴포넌트를 선택적으로 포함할 수 있다.

[0036] 로직 머신(302)은 명령어들을 실행하도록 구성된 하나 이상의 물리적 디바이스들을 포함한다. 예를 들어, 로직 머신은 하나 이상의 애플리케이션, 서비스, 프로그램, 루틴, 라이브러리, 객체, 컴포넌트, 데이터 구조, 또는 다른 로직 구성의 일부인 명령어들을 실행하도록 구성될 수 있다. 이러한 명령어들은 태스크를 수행하거나, 데이터 유형을 구현하거나, 하나 이상의 컴포넌트의 상태를 변환하거나, 기술 효과를 달성하거나, 그렇지 않으면 원하는 결과에 도달하기 위해 구현될 수 있다.

[0037] 로직 머신은 소프트웨어 명령어들을 실행하도록 구성된 하나 이상의 프로세서를 포함할 수 있다. 추가적으로 또는 대안적으로, 로직 머신은 하드웨어 또는 펌웨어 명령어들을 실행하도록 구성된 하나 이상의 하드웨어 또는 펌웨어 로직 머신을 포함할 수 있다. 로직 머신의 프로세서는 단일 코어 또는 다중 코어일 수 있고, 그 위에서 실행되는 명령어들은 순차, 병렬 및/또는 분산 처리를 위해 구성될 수 있다. 로직 머신의 개별 컴포넌트는 선택적으로 조직화(coordinated) 처리를 위해 원격 위치 및/또는 구성될 수 있는 2개 이상의 분리된 디바이스들 사이에 분산될 수 있다. 로직 머신의 양상은 클라우드 컴퓨팅 구성으로 구성된 원격 액세스가능한 네트워크 컴퓨팅 디바이스에 의해 가상화되고 실행될 수 있다.

[0038] 저장 머신(304)은 여기에 설명된 방법 및 프로세스를 구현하기 위해 로직 머신에 의해 실행가능한 명령어들을 보유하도록 구성된 하나 이상의 물리적 디바이스를 포함한다. 그러한 방법 및 프로세스가 구현될 때, 저장 머신(304)의 상태는 예를 들어 상이한 데이터를 보유하도록 변환될 수 있다.

[0039] 저장 머신(304)은 제거가능한 및/또는 내장 디바이스를 포함할 수 있다. 저장 머신(304)은 그 중에서도 광학 메모리(예를 들어, CD, DVD, HD-DVD, 블루레이 디스크 등), 반도체 메모리(예를 들어, RAM, EPROM, EEPROM 등) 및/또는 자기 메모리(예를 들어, 하드 디스크 드라이브, 플로피 디스크 드라이브, 테이프 드라이브, MRAM 등)를 포함할 수 있다. 저장 머신(304)은 휘발성, 비휘발성, 동적, 정적, 읽기/쓰기, 읽기 전용, 랜덤 액세스, 순차 액세스, 위치 주소지정 가능한(addressable), 파일 주소지정 가능한, 및/또는 콘텐츠 주소지정 가능한 디바이스를 포함할 수 있다.

[0040] 저장 머신(304)은 하나 이상의 물리적 디바이스를 포함한다는 것이 인지될 것이다. 그러나, 여기서 설명된 명령어들의 양상은 대안적으로 유한한 지속시간 동안 물리적 디바이스에 의해 보유되지 않는 통신 매체(예를 들어, 전자기 신호, 광학 신호 등)에 의해 전파될 수 있다.

[0041] 로직 머신(302) 및 저장 머신(304)의 양상은 하나 이상의 하드웨어 로직 컴포넌트로 함께 통합될 수 있다. 이러한 하드웨어 로직 컴포넌트는 예를 들어 FPGA(Field-Programmable Gate Array), PASIC/ASIC(Application Specific Integrated Circuit), PSSP/ASSP(Application-Specific Standard Product), SOC(system-on-a-chip) 및 CPLD(complex programmable logic device)를 포함할 수 있다.

[0042] 디스플레이 서브시스템(306)은 저장 머신(302)에 의해 보유된 데이터의 시각적 표현을 제공하는데 사용될 수 있다. 이 시각적 표현은 그래픽 사용자 인터페이스(GUI)의 형태를 취할 수 있다. 여기서 설명된 방법 및 프로세스는 저장 머신에 의해 보유된 데이터를 변경함에 따라 저장 머신의 상태를 변환시키기 때문에, 디스플레이 서브시스템(306)의 상태는 마찬가지로 기저 데이터의 변화를 시각적으로 나타내도록 변환될 수 있다. 디스플레이 서브시스템(306)은 가상으로 사실상 모든 유형의 기술을 이용하는 하나 이상의 디스플레이 디바이스를 포함할 수 있다. 이러한 디스플레이 디바이스는 공유된 인클로저(enclosure) 내의 로직 머신(302) 및/또는 저장 머신(304)과 결합될 수 있거나, 그러한 디스플레이 디바이스는 주변 디스플레이 디바이스일 수 있다.

[0043] 입력 서브시스템(308)은 키보드, 마우스, 터치 스크린 또는 게임 컨트롤러와 같은 하나 이상의 사용자 입력 디바이스를 포함하거나 그들과 인터페이싱할 수 있다. 일부 실시 예들에서, 입력 서브시스템은 선택된 NUI(natural user input) 컴포넌트를 포함하거나 그것과 인터페이싱할 수 있다. 이러한 컴포넌트는 통합되거나 주변 장치일 수 있고, 입력 동작의 변환 및/또는 처리는 온보드 또는 오프보드로 처리될 수 있다. 예시적인 NUI 컴포넌트는 스피치 및/또는 음성 인식을 위한 마이크로폰(310)뿐만 아니라; 머신 비전 및/또는 제스처 인식

을 위한 적외선, 컬러, 입체 및/또는 심도 카메라와 같은 이미지 센서(312); 모션 검출 및/또는 의도 인식을 위한 헤드 트래커, 아이 트래커, 가속도계, 및/또는 자이로스코프; 뇌 활동을 가늠하기 위한 전계 감지 컴포넌트를 제공할 수 있다.

[0044] 통신 서브시스템(314)은 컴퓨팅 시스템(300)을 하나 이상의 다른 컴퓨팅 디바이스들과 통신가능하게 결합하도록 구성될 수 있다. 통신 서브시스템(314)은 하나 이상의 상이한 통신 프로토콜과 호환가능한 유선 및/또는 무선 통신 디바이스를 포함할 수 있다. 비제한적인 예로서, 통신 서브시스템은 무선 전화 네트워크, 또는 유선 또는 무선 로컬 또는 광역 네트워크를 통한 통신을 위해 구성될 수 있다. 일부 실시예에 있어서, 통신 서브시스템은 컴퓨팅 시스템(300)이 인터넷과 같은 네트워크를 통해 다른 장치로 및/또는 다른 장치로부터 메시지를 송신 및/또는 수신하게 할 수 있다

[0045] 또 다른 예로는 음향 센서를 포함한 하나 이상의 센서를 통해 사용 환경을 모니터링하는 단계; 음향 센서로부터 선택된 데이터를 통해 키 문구의 발음을 검출하는 단계; 음향 센서로부터의 선택된 데이터 및 음향 센서로부터의 선택된 데이터와 상이한 시간에 수집된 다른 환경 센서 데이터에 기초하여 키 문구가 식별된 사용자에게 의해 발생되었을 확률을 결정하는 단계; 확률이 문턱 확률을 충족하거나 초과하면 컴퓨팅 디바이스 상에 동작을 수행하는 단계를 포함한다. 이러한 예는 추가적으로 또는 대안적으로 다른 환경 센서 데이터가 음향 센서 데이터를 포함하는 것을 포함할 수 있다. 이러한 예에서, 다른 환경 센서 데이터는 추가적으로 또는 대안적으로 이미지 데이터를 포함할 수 있다. 그러한 예는 추가적으로 또는 대안적으로 이미지 데이터에 기초하여 사용 환경에서 한 명 이상의 사람들을 식별하는 단계를 포함할 수 있고, 확률을 결정하는 단계는 사용 환경에서 한 명 이상의 사람들의 결정된 신원에 적어도 부분적으로 기초하여 확률을 결정하는 단계를 포함한다. 이러한 예에서, 다른 환경 센서 데이터는 추가적으로 또는 대안적으로 위치 데이터를 포함할 수 있다. 이러한 예에서, 위치 데이터는 추가적으로 또는 대안적으로 근접 센서로부터의 근접 데이터를 포함할 수 있다. 그러한 예에서, 위치 데이터는 추가적으로 또는 대안적으로 식별된 사용자에게 대한 캘린더 정보를 포함할 수 있다. 이러한 예는 추가적으로 또는 대안적으로 사용자 행동 패턴을 검출하는 단계를 포함 할 수 있고, 확률을 결정하는 단계는 사용자 행동 패턴에 적어도 부분적으로 기초하여 확률을 결정하는 단계를 포함한다. 이러한 예에서, 사용자 행동 패턴 정보는 추가적으로 또는 대안적으로 식별된 사용자가 발생하는 빈도에 관한 정보를 포함할 수 있다. 상술된 예들 중 임의의 것 또는 모두가 다양한 구현예에서 임의의 적절한 방식으로 결합될 수 있다.

[0046] 컴퓨팅 시스템의 다른 예는 적어도 하나의 음향 센서를 포함한 하나 이상의 센서; 로직 머신; 로직 머신에 의해 실행 가능한 명령어들을 보유하는 저장 머신을 포함하고, 상기 명령어들은 음향 센서를 포함하는 상기 하나 이상의 센서들을 통해 사용 환경을 모니터링하고, 음향 센서로부터의 선택된 데이터를 통해 키 문구의 발음을 검출하고, 음향 센서로부터의 선택된 데이터뿐만 아니라 음향 센서로부터의 선택된 데이터 외의 상이한 시간에서 수집된 다른 환경 센서 데이터에 기초하여 키 문구가 식별된 유저에 의해 발생되었을 확률을 결정하고, 확률이 문턱 확률을 충족하거나 초과하면 컴퓨팅 디바이스 상에 동작을 수행하도록 실행가능하다. 이러한 예에서, 다른 환경 센서 데이터는 추가적으로 또는 대안적으로 이미지 데이터를 포함할 수 있고, 명령어들은 또한 이미지 데이터에 기초하여 사용 환경에서 한 명 이상의 사람을 식별하고, 사용 환경에 있는 한 명 이상의 사람들의 결정된 신원에 적어도 부분적으로 기초하여 확률을 결정하도록 실행가능할 수 있다. 이러한 예에서, 다른 환경 센서 데이터는 추가적으로 또는 대안적으로 위치 데이터를 포함할 수 있고, 위치 데이터는 근접 센서로부터의 하나 이상의 근접 데이터 및 식별된 사용자에게 대한 캘린더 정보를 포함한다. 이러한 예에서, 명령어들은 추가적으로 또는 대안적으로, 식별된 사용자가 캘린더 정보에 기초하여 키 문구의 발음이 검출된 시간 동안 사용 환경에 있는 것으로 스케줄링되었는지를 결정하고, 식별된 사용자가 사용 환경에 있는 것으로 스케줄링되었다면 식별된 사용자에게 의해 키 문구가 발생되었을 확률을 증가시키도록 실행가능할 수 있다. 이러한 예에서, 명령어들은 추가적으로 또는 대안적으로, 환경 감지를 통해 검출된 이전의 사용자 행동에 기초하여 사용자 행동 패턴 - 상기 사용자 행동 패턴은 식별된 사용자가 발생한 빈도에 관한 정보를 포함함 - 을 검출하고, 식별된 사용자가 발생하는 평균 빈도에 기초하여 확률을 결정하도록 실행가능할 수 있다. 이러한 예에서, 음향 센서로부터의 선택된 데이터 외의 상이한 시간에서 수집된 다른 환경 센서 데이터는 키 문구의 발음 이전 및/또는 이후에 수집된 추가의 음향 데이터를 포함할 수 있다. 이러한 예에서, 명령어들은 추가적으로 또는 대안적으로, 키 문구가 식별된 사용자에게 의해 발생되었을 확률을 결정하기 위해 키 문구의 발음 이전 또는 이후에 식별된 사용자가 또한 발생하였는지를 결정하도록 추가의 음향 데이터를 분석하고, 키 문구의 발음 이전 또는 이후에 식별된 사용자가 또한 발생하였다면 키 문구가 식별된 사용자에게 의해 발생되었을 확률을 증가시키도록 실행가능할 수 있다. 이러한 예에서, 명령어들은 추가적으로 또는 대안적으로 분석이 키 문구의 발음 이전 또는 이후에 식별된 사용자가 발생하지 않았다고 나타내면 식별된 사용자에게 의해 키 문구가 발생되었을 확률을 감소시키도록 실행가능할 수

있다. 상술된 예들 중 임의의 것 또는 모두는 다양한 구현예에서 임의의 적절한 방식으로 결합될 수 있다.

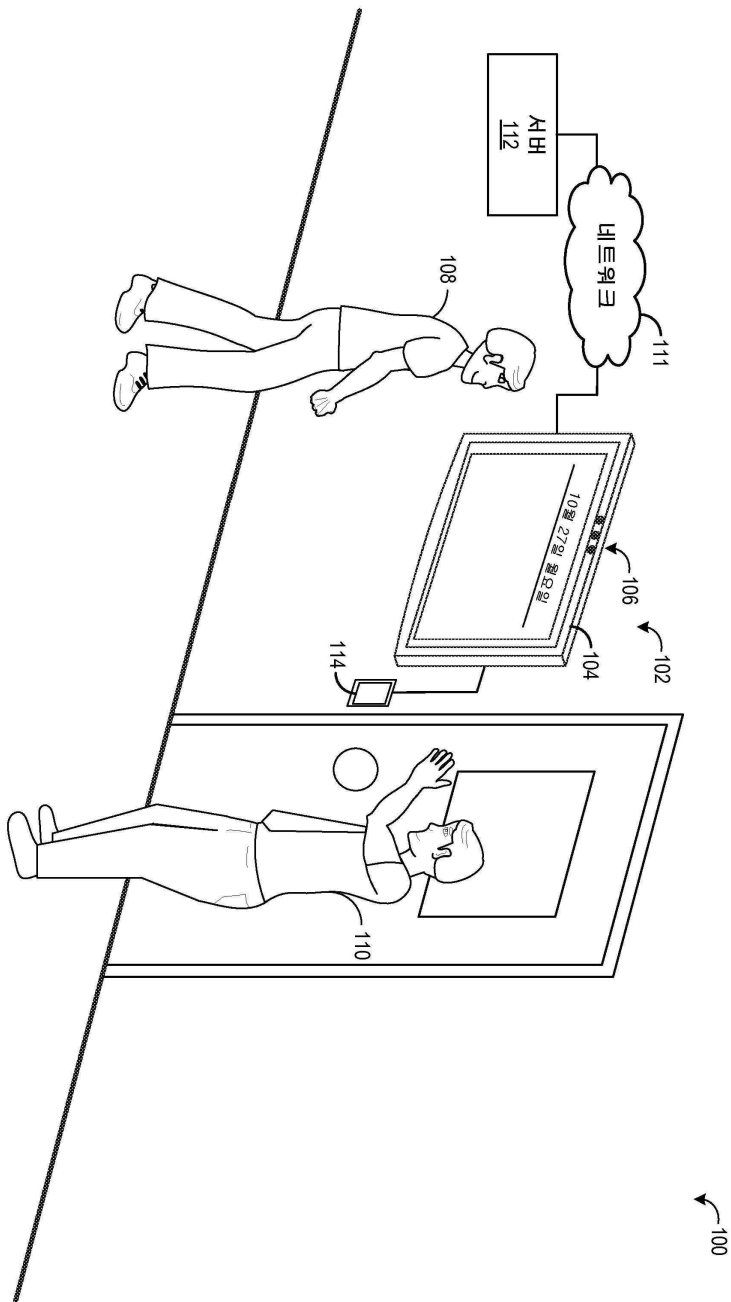
[0047] 또 다른 예는 음향 센서를 포함한 하나 이상의 센서; 컴퓨터 판독가능 명령어들을 실행하도록 구성된 프로세서; 및 명령어들을 저장하는 저장 머신을 포함하는 컴퓨팅 시스템을 제공하고, 상기 명령어들은 프로세서에 의해, 음향 센서를 포함한 하나 이상의 센서를 통해 사용 환경을 모니터링하고, 음향 센서로부터 선택된 데이터를 통해 커맨드가 뒤따르는 키 문구의 발언을 검출하고 - 상기 키 문구는 사용자를 식별하도록 구성되며, 상기 커맨드는 사용자와 관련된 컴퓨팅 시스템 동작을 호출하도록 구성됨 -, 키 문구를 발언한 사용자의 신원이 음향 센서로부터의 선택된 데이터에 기초하여 결정되면 식별된 사용자를 위한 동작을 컴퓨터 디바이스 상에 수행하고, 키 문구를 발언한 사용자의 신원이 음향 센서로부터의 선택된 데이터에 기초하여 결정되지 않으면 음향 센서로부터의 선택된 데이터 외의 상이한 시간에서 수집된 다른 환경 센서 데이터에 기초하여 키 문구가 하나 이상의 가능성 있는 사용자들에 의해 발화되었을 확률을 결정하고, 확률이 하나 이상의 다른 가능성 있는 사용자들 중 선택된 사용자에 대한 문턱 확률을 충족하거나 초과한다면 선택된 사용자를 위한 동작을 컴퓨팅 시스템 상에 수행하도록 실행가능하다. 이러한 예에서, 명령어들은 추가적으로 또는 대안적으로 환경 감지를 통해 검출된 이전의 사용자 행동에 기초하여 사용자 행동 패턴을 검출하도록 실행가능할 수 있고, 여기서 확률을 결정하는 것은 사용자 행동 패턴에 부분적으로 기초하여 확률을 결정하는 것을 포함한다. 이러한 예에서, 사용자 행동 패턴을 검출하기 위해 실행 가능한 명령어들은 식별된 사용자가 발생하는 빈도에 관한 정보를 검출하도록 실행가능한 명령어들을 포함할 수 있다. 상술된 예들 중 임의의 것 또는 모두는 다양한 구현예에서 임의의 적절한 방식으로 결합될 수 있다.

[0048] 여기에 기술된 구성 및/또는 접근법은 본질적으로 예시적인 것이며, 이들 특정 실시예 또는 예는 많은 변형이 가능하기 때문에 제한적인 의미로 고려되지 않음이 이해될 것이다. 여기에 설명된 특정 루틴 또는 방법은 임의의 수의 처리 전략 중 하나 이상을 나타낼 수 있다. 이와 같이, 예시되고/예시되거나 설명된 다양한 동작은 예시되고/예시되거나 설명된 시퀀스로, 다른 시퀀스로, 병렬로 수행되거나, 또는 생략될 수 있다. 마찬가지로, 상술된 프로세스의 순서는 변경될 수 있다.

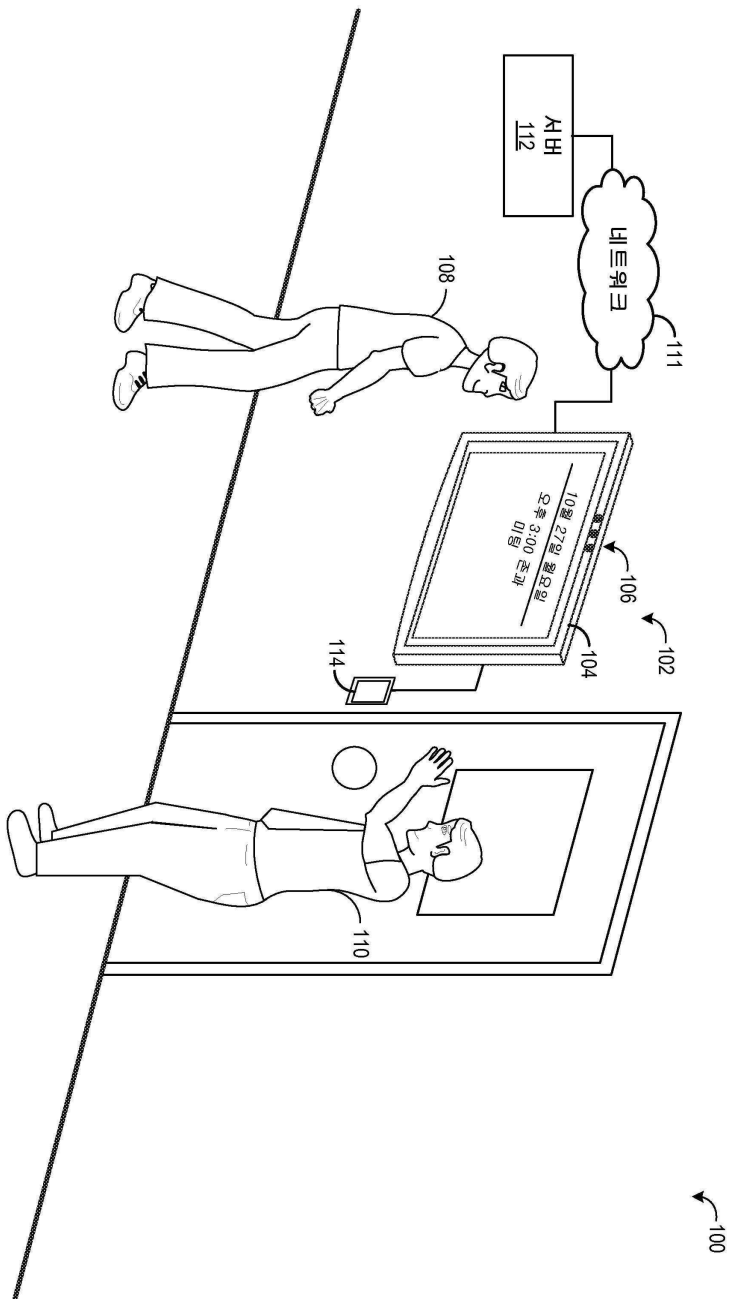
[0049] 본 명세서의 청구 대상은 여기에 개시된 다양한 프로세스, 시스템 및 구성, 및 다른 특징, 기능, 동작 및/또는 속성의 모든 신규하고 뻗하지 않은 조합 및 서브조합뿐만 아니라 임의의 및 모든 그 등가물을 포함한다.

도면

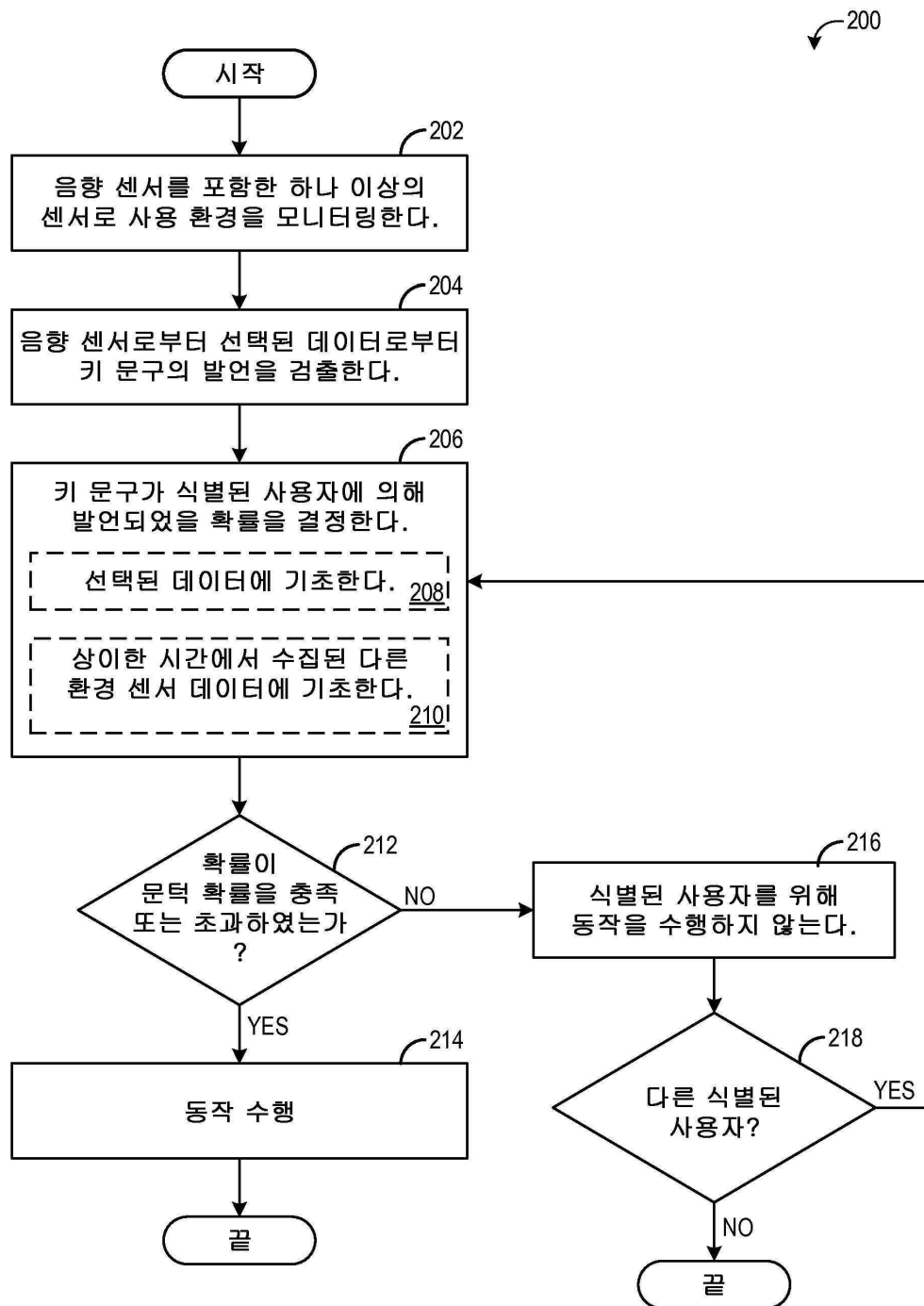
도면1a



도면1b



도면2





도면3

