



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0142775
(43) 공개일자 2023년10월11일

- (51) 국제특허분류(Int. Cl.)
G10L 15/22 (2006.01) G06F 3/16 (2018.01)
G06F 40/30 (2020.01) G06V 40/16 (2022.01)
G10L 15/18 (2006.01) G10L 15/25 (2013.01)
G10L 21/0208 (2013.01) G10L 25/51 (2013.01)
G10L 25/78 (2013.01)
- (52) CPC특허분류
G10L 15/22 (2013.01)
G06F 3/167 (2013.01)
- (21) 출원번호 10-2023-7030348
- (22) 출원일자(국제) 2022년02월24일
심사청구일자 2023년09월06일
- (85) 번역문제출일자 2023년09월06일
- (86) 국제출원번호 PCT/US2022/017652
- (87) 국제공개번호 WO 2022/191999
국제공개일자 2022년09월15일
- (30) 우선권주장
63/159,748 2021년03월11일 미국(US)
17/330,862 2021년05월26일 미국(US)

- (71) 출원인
애플 인크.
미국 캘리포니아 (우편번호 95014) 쿠퍼티노 원
애플 파크 웨이
- (72) 발명자
매디카, 스리닐
미국 94582 캘리포니아 샌 라몬 캐모마일 코트
2150
후센 압텔라지즈, 아흐메드 세라그 엘 던
미국 95014 캘리포니아 쿠퍼티노 애플 파크 웨이
1 애플 인크 내
(뒷면에 계속)
- (74) 대리인
장덕순, 백만기

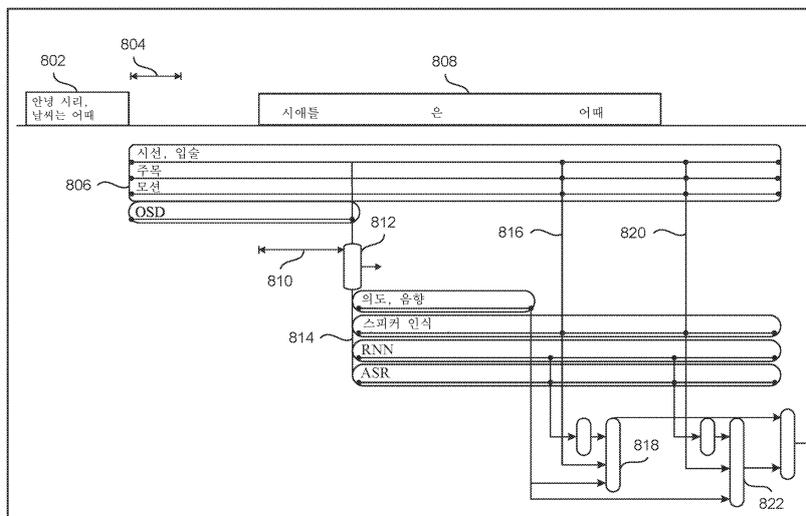
전체 청구항 수 : 총 25 항

(54) 발명의 명칭 지속적인 다이얼로그를 위한 다중 상태 디지털 어시스턴트

(57) 요약

지능형 자동화 어시스턴트를 동작시키기 위한 시스템들 및 프로세스들이 제공된다. 예를 들어, 사용자로부터 제 1 스피치 입력이 수신된다. 제 1 스피치 입력을 수신하는 것에 응답하여, 응답이 제공된다. 제 1 상태에서 디지털 어시스턴트에 대응하여 제 1 출력이 제공되고, 사용자로부터 제 2 스피치 입력이 수신된다. 제 1 복수의 값들이 획득된다. 제 1 복수의 값들에 기초하여, 제 2 스피치 입력에 대응하는 제 1 신뢰도 레벨이 획득된다. 제 1 신뢰도 레벨이 제 1 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제 2 상태에서 디지털 어시스턴트에 대응하여 제 2 출력이 제공된다. 제 2 스피치 입력은 계속해서 수신된다.

대표도 - 도8



(52) CPC특허분류

G06F 40/30 (2020.01)
G06V 40/171 (2022.01)
G10L 15/1822 (2013.01)
G10L 15/25 (2013.01)
G10L 21/0208 (2013.01)
G10L 25/51 (2013.01)
G10L 25/78 (2013.01)
G10L 2015/225 (2013.01)

(72) 발명자

만네말라, 차이타냐

미국 95014 캘리포니아 쿠퍼티노 애플 파크 웨이 1
애플 인크 내

비슈누보틀라, 스리칸트

미국 95014 캘리포니아 쿠퍼티노 애플 파크 웨이 1
애플 인크 내

바인베르그, 가렛 엘.

미국 95014 캘리포니아 쿠퍼티노 애플 파크 웨이 1
애플 인크 내

명세서

청구범위

청구항 1

컴퓨터 구현 방법으로서,

하나 이상의 프로세서들 및 메모리를 갖는 전자 디바이스에서:

사용자로부터, 제1 스피치 입력을 수신하는 단계;

상기 제1 스피치 입력을 수신하는 것에 응답하여, 상기 제1 스피치 입력에 기초하여 응답을 제공하는 단계;

제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력을 제공하는 단계;

상기 사용자로부터, 제2 스피치 입력을 수신하는 단계;

제1 복수의 값들을 획득하는 단계;

상기 제1 복수의 값들에 기초하여, 상기 제2 스피치 입력에 대응하는 제1 신뢰도 레벨을 획득하는 단계;

상기 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라:

제2 상태에서 상기 디지털 어시스턴트에 대응하는 제2 출력을 제공하는 단계; 및

상기 제2 스피치 입력을 계속해서 수신하는 단계를 포함하는, 방법.

청구항 2

제1항에 있어서,

상기 전자 디바이스가 에코 제거를 위해 구성된다는 결정에 따라:

상기 제1 스피치 입력의 검출된 종료에 응답하여 상기 제1 복수의 값들의 획득을 개시하는 단계; 및

상기 전자 디바이스가 에코 제거를 위해 구성되지 않는다는 결정에 따라:

상기 제공된 응답의 검출된 종료에 응답하여 상기 제1 복수의 값들의 획득을 개시하는 단계를 포함하는, 방법.

청구항 3

제1항 또는 제2항에 있어서,

상기 제2 스피치 입력이 최소 임계 지속기간과 연관된다는 결정에 따라:

상기 제1 복수의 값들의 각각의 값에 대해, 각자의 값이 적어도 하나의 규칙을 만족하는지의 여부를 결정하는 단계; 및

각자의 값이 적어도 하나의 규칙을 만족한다는 결정에 따라, 상기 제1 신뢰도 레벨을 증가시키는 단계를 포함하는, 방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 제1 복수의 값들을 획득하는 단계는,

상기 전자 디바이스의 디스플레이에 지향되는 사용자 시선을 검출하는 단계;

상기 사용자 시선이 디스플레이된 디지털 어시스턴트 객체에 지향되는지의 여부를 결정하는 단계; 및

상기 사용자 시선이 상기 디스플레이된 디지털 어시스턴트 객체에 지향되는지의 여부에 대한 상기 결정에 기초하여 상기 제1 복수의 값들의 각각의 값을 획득하는 단계를 포함하는, 방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 제1 복수의 값들을 획득하는 단계는,
상기 사용자와 연관된 입술 움직임을 검출하는 단계;
상기 입술 움직임을 상기 제1 스피치 입력에 대응하는지의 여부를 결정하는 단계; 및
상기 결정에 기초하여 상기 제1 복수의 값들의 각자의 값을 획득하는 단계를 포함하는, 방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 제1 복수의 값들을 획득하는 단계는,
사용자 시선과 연관된 방향을 검출하는 단계; 및
상기 결정된 방향에 기초하여 상기 제1 복수의 값들의 각자의 값을 획득하는 단계를 포함하는, 방법.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 제1 복수의 값들을 획득하는 단계는,
상기 전자 디바이스와 연관된 위치 정보를 검출하는 단계; 및
상기 위치 정보에 기초하여 상기 제1 복수의 값들의 각자의 값을 획득하는 단계를 포함하는, 방법.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 제1 복수의 값들을 획득하는 단계는,
상기 전자 디바이스에서 스피치가 검출되는지의 여부를 결정하는 단계; 및
상기 전자 디바이스에서 스피치가 검출된다는 상기 결정에 기초하여 상기 제1 복수의 값들의 각자의 값을 획득하는 단계를 포함하는, 방법.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서,
상기 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라:
제2 복수의 값들을 획득하는 단계;
상기 제1 복수의 값들 및 상기 제2 복수의 값들에 기초하여, 상기 제2 스피치 입력에 대응하는 제2 신뢰도 레벨을 획득하는 단계;
상기 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과한다는 결정에 따라:
상기 제2 스피치 입력을 계속해서 수신하는 단계; 및
상기 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과하지 않는다는 결정에 따라:
상기 제2 스피치 입력을 수신하는 것을 중지하는 단계를 포함하는, 방법.

청구항 10

제9항에 있어서, 제2 복수의 값들을 획득하는 단계는,
상기 제2 스피치 입력의 미리결정된 지속기간과 연관된 사용자 의도를 식별하는 단계;
상기 사용자 의도에 기초하여, 상기 제2 스피치 입력이 디지털 어시스턴트로 지향되는지의 여부를 결정하는 단계; 및
상기 제2 스피치 입력이 디지털 어시스턴트로 지향되는지의 여부에 대한 상기 결정에 기초하여 상기 제2 복수의 값들의 각자의 값을 획득하는 단계를 포함하는, 방법.

청구항 11

제9항 또는 제10항에 있어서, 제2 복수의 값들을 획득하는 단계는,

상기 제1 스피치 입력과 연관된 제1 스피커 프로파일을 인출하는 단계;
 상기 제2 스피치 입력과 연관된 제2 스피커 프로파일을 획득하는 단계;
 상기 제1 스피커 프로파일을 상기 제2 스피커 프로파일과 비교하는 단계; 및
 상기 비교에 기초하여 상기 제2 복수의 값들의 각자의 값을 획득하는 단계를 포함하는, 방법.

청구항 12

제9항 내지 제11항 중 어느 한 항에 있어서, 제2 복수의 값들을 획득하는 단계는,
 스피치 인식 출력에 기초하여, 격자 임베딩을 획득하는 단계;
 상기 격자 임베딩에 기초하여 사용자 의도를 결정하는 단계; 및
 상기 사용자 의도에 기초하여 상기 제2 복수의 값들의 각자의 값을 획득하는 단계를 포함하는, 방법.

청구항 13

제9항 내지 제12항 중 어느 한 항에 있어서, 제2 복수의 값들을 획득하는 단계는,
 상기 제2 스피치 입력에 기초하여 결과 후보를 식별하는 단계;
 상기 결과 후보들을 식별하는 것에 응답하여:
 업데이트된 제1 복수의 값들 및 업데이트된 제2 복수의 값들을 획득하는 단계; 및
 상기 업데이트된 제1 복수의 값들 및 상기 업데이트된 제2 복수의 값들에 기초하여, 상기 제2 스피치 입력에 대응하는 업데이트된 제2 신뢰도 레벨을 획득하는 단계를 포함하는, 방법.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력을 제공하는 단계는 제1 상태에서 디지털 어시스턴트 객체를 디스플레이하는 단계를 포함하고, 상기 방법은,
 상기 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과하지 않는다는 결정에 따라:
 상기 제1 상태에서 상기 디지털 어시스턴트 객체의 디스플레이를 유지하는 단계;
 상기 제2 스피치 입력의 미리결정된 지속기간과 연관된 제2 복수의 값들을 획득하는 단계; 및
 상기 제1 복수의 값들 및 상기 제2 복수의 값들에 기초하여, 상기 제2 스피치 입력에 대응하는 제2 신뢰도 레벨을 획득하는 단계를 포함하는, 방법.

청구항 15

제14항에 있어서,
 상기 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과한다는 결정에 따라:
 제2 상태에서 상기 디지털 어시스턴트 객체를 디스플레이하는 단계; 및
 상기 제2 상태에서 상기 디지털 어시스턴트 객체를 디스플레이하는 동안, 상기 제2 스피치 입력을 계속해서 수신하는 단계를 포함하는, 방법.

청구항 16

제14항 또는 제15항에 있어서,
 어포던스를 디스플레이하는 단계 - 상기 어포던스는 콘텍스트 정보와 연관됨 -; 및
 상기 제2 스피치 입력이 상기 콘텍스트 정보와 연관된다는 결정에 따라, 상기 제2 신뢰도 레벨을 증가시키는 단계를 포함하는, 방법.

청구항 17

제16항에 있어서, 상기 콘텍스트 정보는 제1 시맨틱 표현을 포함하고, 상기 방법은,

상기 제2 스피치 입력과 연관된 제2 시맨틱 표현을 획득하는 단계; 및

상기 제1 시맨틱 표현이 상기 제2 시맨틱 표현에 대응한다는 결정에 따라, 상기 제2 신뢰도 레벨을 증가시키는 단계를 포함하는, 방법.

청구항 18

제16항 또는 제17항에 있어서, 상기 콘텍스트 정보는 적어도 하나의 미리정의된 단어를 포함하고, 상기 방법은, 상기 제2 스피치 입력에 포함된 적어도 하나의 단어를 식별하는 단계; 및

상기 적어도 하나의 미리정의된 단어가 상기 적어도 하나의 식별된 단어에 대응한다는 결정에 따라, 상기 제2 신뢰도 레벨을 증가시키는 단계를 포함하는, 방법.

청구항 19

제1항 내지 제18항 중 어느 한 항에 있어서, 제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력을 제공하는 단계는 제1 상태에서 디지털 어시스턴트 객체를 디스플레이하는 것 및 가청 출력을 제공하는 것 중 적어도 하나를 포함하는, 방법.

청구항 20

전자 디바이스로서,

하나 이상의 프로세서들;

메모리; 및

상기 메모리에 저장되고 상기 하나 이상의 프로세서들에 의해 실행되도록 구성된 하나 이상의 프로그램들을 포함하며, 상기 하나 이상의 프로그램들은,

사용자로부터, 제1 스피치 입력을 수신하기 위한;

상기 제1 스피치 입력을 수신하는 것에 응답하여, 상기 제1 스피치 입력에 기초하여 응답을 제공하기 위한;

제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력을 제공하기 위한;

상기 사용자로부터, 제2 스피치 입력을 수신하기 위한;

제1 복수의 값들을 획득하기 위한;

상기 제1 복수의 값들에 기초하여, 상기 제2 스피치 입력에 대응하는 제1 신뢰도 레벨을 획득하기 위한;

상기 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라;

제2 상태에서 상기 디지털 어시스턴트에 대응하는 제2 출력을 제공하는 단계; 및

상기 제2 스피치 입력을 계속해서 수신하기 위한 명령어들을 포함하는, 전자 디바이스.

청구항 21

하나 이상의 프로그램들을 저장하는 비일시적 컴퓨터 판독가능 저장 매체로서, 상기 하나 이상의 프로그램들은, 제1 전자 디바이스의 하나 이상의 프로세서들에 의해 실행될 때, 상기 제1 전자 디바이스로 하여금,

사용자로부터, 제1 스피치 입력을 수신하게 하고;

상기 제1 스피치 입력을 수신하는 것에 응답하여, 상기 제1 스피치 입력에 기초하여 응답을 제공하게 하고;

제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력을 제공하게 하고;

상기 사용자로부터, 제2 스피치 입력을 수신하게 하고;

제1 복수의 값들을 획득하게 하고;

상기 제1 복수의 값들에 기초하여, 상기 제2 스피치 입력에 대응하는 제1 신뢰도 레벨을 획득하게 하고;

상기 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라:

제2 상태에서 상기 디지털 어시스턴트에 대응하는 제2 출력을 제공하게 하고; 및

상기 제2 스피치 입력을 계속해서 수신하게 하는 명령어들을 포함하는, 비밀시적 컴퓨터 판독가능 저장 매체.

청구항 22

전자 디바이스로서,

사용자로부터, 제1 스피치 입력을 수신하기 위한 수단;

상기 제1 스피치 입력을 수신하는 것에 응답하여, 상기 제1 스피치 입력에 기초하여 응답을 제공하기 위한 수단;

제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력을 제공하기 위한 수단;

상기 사용자로부터, 제2 스피치 입력을 수신하기 위한 수단;

제1 복수의 값들을 획득하기 위한 수단;

상기 제1 복수의 값들에 기초하여, 상기 제2 스피치 입력에 대응하는 제1 신뢰도 레벨을 획득하기 위한 수단;

상기 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제2 상태에서 상기 디지털 어시스턴트에 대응하는 제2 출력을 제공하기 위한 수단; 및

상기 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라, 상기 제2 스피치 입력을 계속해서 수신하기 위한 수단을 포함하는, 전자 디바이스.

청구항 23

전자 디바이스로서,

하나 이상의 프로세서들;

메모리; 및

상기 메모리에 저장되고 상기 하나 이상의 프로세서들에 의해 실행되도록 구성된 하나 이상의 프로그램들을 포함하며, 상기 하나 이상의 프로그램들은 제1항 내지 제19항 중 어느 한 항의 방법을 수행하기 위한 명령어들을 포함하는, 전자 디바이스.

청구항 24

하나 이상의 프로그램들을 저장하는 비밀시적 컴퓨터 판독가능 저장 매체로서, 상기 하나 이상의 프로그램들은, 전자 디바이스의 하나 이상의 프로세서들에 의해 실행될 때, 상기 전자 디바이스로 하여금 제1항 내지 제19항 중 어느 한 항의 방법을 수행하게 하는 명령어들을 포함하는, 비밀시적 컴퓨터 판독가능 저장 매체.

청구항 25

전자 디바이스로서,

제1항 내지 제19항 중 어느 한 항의 방법을 수행하기 위한 수단을 포함하는, 전자 디바이스.

발명의 설명

기술 분야

[0001] 관련 출원에 대한 상호 참조

[0002] 본 출원은, 2021년 3월 11일자로 출원되고 발명의 명칭이 "MULTIPLE STATE DIGITAL ASSISTANT FOR CONTINUOUS DIALOG"인 미국 가특허 출원 제63/159,748호, 및 2021년 5월 26일자로 출원되고 발명의 명칭이 "MULTIPLE STATE DIGITAL ASSISTANT FOR CONTINUOUS DIALOG"인 미국 특허 출원 제17/330,862호에 대한 우선권을 주장하며, 이들 각각의 내용들은 이로써 모든 목적들을 위하여 그들 전체가 참고로 포함된다.

[0003] 기술분야

[0004] 본 출원은 대체적으로 지능형 자동화 어시스턴트들에 관한 것으로, 보다 구체적으로는, 디지털 어시스턴트를 사용하는 지속적인 다이얼로그를 용이하게 하는 것에 관한 것이다.

배경 기술

[0005] 지능형 자동화 어시스턴트들(또는 디지털 어시스턴트들)은 인간 사용자들과 전자 디바이스들 사이의 유익한 인터페이스를 제공할 수 있다. 이러한 어시스턴트들은 사용자가 음성(spoken) 및/또는 텍스트 형태들의 자연 언어를 사용하여 디바이스들 또는 시스템들과 상호작용하도록 할 수 있다. 예를 들어, 사용자는 전자 디바이스 상에서 동작하는 디지털 어시스턴트에게 사용자 요청을 포함하는 스피치 입력을 제공할 수 있다. 디지털 어시스턴트는 스피치 입력으로부터 사용자의 의도를 해석하고 사용자의 의도를 태스크들로 조작할 수 있다. 이어서, 태스크들은 전자 디바이스의 하나 이상의 서비스들을 실행함으로써 수행될 수 있고, 사용자 요청에 응답하는 관련 출력이 사용자에게 반환될 수 있다.

[0006] 디지털 어시스턴트와의 사용자 상호작용 동안, 사용자 입력 및 디지털 어시스턴트 응답들이 교환될 수 있다. 상호작용은 사용자와 어시스턴트 사이에서 다수 차례의 교환들을 포함할 수 있다. 그러나, 전통적인 디지털 어시스턴트 시스템들은 전형적으로, 사용자와 디지털 어시스턴트 사이의 강건한 상호작용들을 용이하게 하는 능력들을 포함하지 않는다. 예를 들어, 종래의 시스템들은 대체적으로, 사용자의 후속 스피치에 적용되는 스피치 분석의 다수의 스테이지들, 예컨대 제1 세트의 값들이 분석되는 제1 스테이지 및 제2 세트의 값들이 분석되는 제2 스테이지를 포함하지 않는다. 이러한 시스템들은 또한, 필요한 경우, 사용자가 디지털 어시스턴트를 동적으로 중단하거나 정정할 수 있게 하지 않는다. 따라서, 지속적인 다이얼로그 능력들을 갖는 개선된 디지털 어시스턴트 시스템이 요구된다.

발명의 내용

[0007] 지능형 자동화 어시스턴트를 동작시키기 위한 시스템들 및 프로세스들이 제공된다. 예를 들어, 사용자로부터 제1 스피치 입력이 수신된다. 제1 스피치 입력을 수신하는 것에 응답하여, 응답이 제공된다. 제1 상태에서 디지털 어시스턴트에 대응하여 제1 출력이 제공되고, 사용자로부터 제2 스피치 입력이 수신된다. 제1 복수의 값들이 획득된다. 제1 복수의 값들에 기초하여, 제2 스피치 입력에 대응하는 제1 신뢰도 레벨이 획득된다. 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제2 상태에서 디지털 어시스턴트에 대응하여 제2 출력이 제공된다. 제2 스피치 입력은 계속해서 수신된다.

도면의 간단한 설명

- [0008] 도 1은 다양한 예들에 따른, 디지털 어시스턴트를 구현하기 위한 시스템 및 환경을 도시하는 블록도이다.
- 도 2a는 다양한 예들에 따른, 디지털 어시스턴트의 클라이언트 측 부분을 구현하는 휴대용 다기능 디바이스를 도시하는 블록도이다.
- 도 2b는 다양한 예들에 따른, 이벤트 처리를 위한 예시적인 컴포넌트들을 도시하는 블록도이다.
- 도 3은 다양한 예들에 따른, 디지털 어시스턴트의 클라이언트 측 부분을 구현하는 휴대용 다기능 디바이스를 도시한다.
- 도 4는 다양한 예들에 따른, 디스플레이 및 터치 감응형 표면을 갖는 예시적인 다기능 디바이스의 블록도이다.
- 도 5a는 다양한 예들에 따른, 휴대용 다기능 디바이스 상의 애플리케이션들의 메뉴에 대한 예시적인 사용자 인터페이스를 도시한다.
- 도 5b는 다양한 예들에 따른, 디스플레이로부터 분리된 터치 감응형 표면을 갖는 다기능 디바이스에 대한 예시적인 사용자 인터페이스를 도시한다.
- 도 6a는 다양한 예들에 따른 개인용 전자 디바이스를 도시한다.
- 도 6b는 다양한 예들에 따른 개인용 전자 디바이스를 도시하는 블록도이다.
- 도 7a는 다양한 예들에 따른 디지털 어시스턴트 시스템 또는 그의 서버 부분을 도시하는 블록도이다.
- 도 7b는 다양한 예들에 따른, 도 7a에 도시된 디지털 어시스턴트의 기능들을 도시한다.

도 7c는 다양한 예들에 따른 온톨로지(ontology)의 일부분을 도시한다.

도 8은 다양한 예들에 따른, 디지털 어시스턴트를 사용하는 지속적인 다이얼로그를 용이하게 하기 위한 프로세스를 도시한다.

도 9는 다양한 예들에 따른, 디지털 어시스턴트를 사용하는 지속적인 다이얼로그를 용이하게 하기 위한 프로세스를 도시한다.

도 10은 다양한 예들에 따른, 디지털 어시스턴트를 사용하는 지속적인 다이얼로그를 용이하게 하기 위한 프로세스를 도시한다.

도 11a 및 도 11b는 다양한 예들에 따른, 디지털 어시스턴트를 사용하는 지속적인 다이얼로그를 용이하게 하기 위한 프로세스를 도시한다.

발명을 실시하기 위한 구체적인 내용

[0009] 예들의 다음 설명에서, 첨부된 도면들이 참조되며, 실시될 수 있는 특정 예들이 도면들 내에서 예시로서 도시된다. 다양한 예들의 범주를 벗어나지 않으면서 다른 예들이 이용될 수 있고 구조적 변경이 행해질 수 있다는 것이 이해되어야 한다.

[0010] 지속적인 디지털 어시스턴트 상호작용들을 위한 종래의 기법들은 대체적으로 유효성이 결여되어 있다. 특히, 예시적인 전통적인 시스템은, 현재 콘텍스트 또는 상태에 따라 특정 값들을 고려하는 다중 상태 프레임워크와 같은, "후속" 스피치, 또는 이전 사용자 요청 및/또는 디지털 어시스턴트 응답으로 지향되는 스피치를 처리하기 위한 강건한 프레임워크를 포함하지 않는다. 대조적으로, 전통적인 시스템들은 전형적으로, 사용자 및 디바이스와 관련된 추가적인 다중 모드 입력들(예컨대, 시선, 주목(attention), 디바이스 모션, 스피커 인식 등) 대신에 스피치 관련 큐들에 초점을 맞춘다. 따라서, 이러한 시스템들은 사용자가 디바이스와 상호작용할 수 있게 하는 효과적이고 끊임 없는 수단을 제공하지 않는다.

[0011] 이하의 설명이 다양한 요소들을 기술하기 위해 "제1", "제2" 등과 같은 용어들을 사용하지만, 이러한 요소들이 그 용어들에 의해 제한되어서는 안 된다. 이러한 용어들은 하나의 요소를 다른 요소와 구별하는 데에만 사용된다. 예를 들어, 다양하게 기술된 예들의 범주로부터 벗어남이 없이, 제1 입력이 제2 입력으로 지칭될 수 있을 것이고, 이와 유사하게, 제2 입력이 제1 입력으로 지칭될 수 있을 것이다. 제1 입력 및 제2 입력은 둘 모두 입력들이고, 일부 경우들에서, 별개이고 상이한 입력들이다.

[0012] 본 명세서에서 다양하게 기술되는 예들의 설명에 사용되는 용어는 특정 예들을 기술하는 목적만을 위한 것이고, 제한하려는 의도는 아니다. 다양하게 기술된 예들의 설명 및 첨부된 청구범위에 사용되는 바와 같이, 단수의 형태("a", "an", 및 "the")는 문맥상 명백히 달리 나타내지 않는다면 복수의 형태도 마찬가지로 포함하려는 것으로 의도된다. 또한, 본 명세서에서 사용되는 바와 같은 용어 "및/또는"은 열거되는 연관된 항목들 중 하나 이상의 항목들의 임의의 그리고 모든 가능한 조합들을 나타내고 그들을 포괄하는 것임이 이해될 것이다. 용어들 "포함한다(include)", "포함하는(including)", "포함한다(comprise)", 및/또는 "포함하는(comprising)"은, 본 명세서에서 사용될 때, 언급된 특징들, 정수들, 단계들, 동작들, 요소들, 및/또는 컴포넌트들의 존재를 특정하지만, 하나 이상의 다른 특징들, 정수들, 단계들, 동작들, 요소들, 컴포넌트들, 및/또는 이들의 그룹들의 존재 또는 추가를 배제하지 않음이 추가로 이해될 것이다.

[0013] 용어 "~할 경우(if)"는 문맥에 따라 "~할 때(when)" 또는 "~할 시(upon)" 또는 "결정하는 것에 응답하여(in response to determining)" 또는 "검출하는 것에 응답하여(in response to detecting)"를 의미하는 것으로 해석될 수 있다. 유사하게, 구절 "~라고 결정되는 경우" 또는 "[언급된 조건 또는 이벤트가] 검출되는 경우"는, 문맥에 따라 "~라고 결정할 시" 또는 "~라고 결정하는 것에 응답하여" 또는 "[언급된 조건 또는 이벤트]를 검출할 시" 또는 "[언급된 조건 또는 이벤트]를 검출하는 것에 응답하여"를 의미하는 것으로 해석될 수 있다.

[0014] 1. 시스템 및 환경

[0015] 도 1은 다양한 예들에 따른 시스템(100)의 블록도를 도시한다. 일부 예들에서, 시스템(100)은 디지털 어시스턴트를 구현한다. 용어들 "디지털 어시스턴트", "가상 어시스턴트", "지능형 자동화 어시스턴트", 또는 "자동 디지털 어시스턴트"는 구어적 및/또는 문어적 형태로 입력되는 자연 언어를 해석하여 사용자 의도를 추론하고, 추론된 사용자 의도에 기초하여 동작들을 수행하는 임의의 정보 처리 시스템을 지칭한다. 예를 들어, 추론된 사용자 의도에 따라 동작하기 위해, 시스템은 다음 중 하나 이상을 수행한다: 추론된 사용자 의도를 달성하도록

설계된 단계들 및 파라미터들을 이용하여 태스크 흐름을 식별하는 것; 추론된 사용자 의도로부터의 특정 조건들을 태스크 흐름 내에 입력하는 것; 프로그램들, 방법들, 서비스들, API들 등을 호출(involve)함으로써 태스크 흐름을 실행하는 것; 및 청각적(예컨대, 스피치) 및/또는 시각적 형태로 사용자에게 대한 출력 응답들을 생성하는 것.

[0016] 구체적으로, 디지털 어시스턴트는 적어도 부분적으로 자연 언어 커맨드, 요청, 진술, 서술, 및/또는 질문의 형태로 사용자 요청을 수용할 수 있다. 전형적으로, 사용자 요청은 디지털 어시스턴트에 의해 정보제공형 답변 또는 태스크의 수행 중 어느 하나를 구한다. 사용자 요청에 대한 만족스러운 응답은 요청된 정보제공형 답변의 제공, 요청된 태스크의 수행, 또는 이 둘의 조합을 포함한다. 예를 들어, 사용자는 디지털 어시스턴트에게, "내가 지금 어디에 있지?"와 같은 질문을 한다. 사용자의 현재 위치에 기초하여, 디지털 어시스턴트는 "당신은 웨스트 게이트 근처 센트럴 파크에 있습니다"라고 답변한다. 사용자는 또한 태스크의 수행, 예를 들어 "다음주 내 여자친구의 생일 파티에 내 친구들을 초대해 줘"를 요청한다. 이에 응답하여, 디지털 어시스턴트는 "네, 바로 처리하겠습니다"라고 말함으로써 요청을 승인확인(acknowledge)할 수 있고, 이어서 사용자의 전자 주소록에 열거된 사용자의 친구들 각각에게 사용자를 대신하여 적절한 행사 예정 초대장(calendar invite)을 전송할 수 있다. 요청된 태스크의 수행 동안, 디지털 어시스턴트는 때때로 연장된 기간에 걸쳐 다수의 정보 교환을 수반하는 연속 대화로 사용자와 상호작용한다. 디지털 어시스턴트와 상호작용하여 정보 또는 다양한 태스크들의 수행을 요청하는 수많은 다른 방식이 있다. 구두 응답들을 제공하는 것 및 프로그래밍된 동작들을 취하는 것 외에도, 디지털 어시스턴트는 또한 예컨대, 텍스트, 경보들, 음악, 비디오들, 애니메이션들 등과 같은 다른 시각적 또는 청각적 형태들의 응답들을 제공한다.

[0017] 도 1에 도시된 바와 같이, 일부 예들에서, 디지털 어시스턴트는 클라이언트-서버 모델에 따라 구현된다. 디지털 어시스턴트는 사용자 디바이스(104) 상에서 실행되는 클라이언트 측 부분(102)(이하, "DA 클라이언트(102)"), 및 서버 시스템(108) 상에서 실행되는 서버 측 부분(106)(이하, "DA 서버(106)")을 포함한다. DA 클라이언트(102)는 하나 이상의 네트워크들(110)을 통해 DA 서버(106)와 통신한다. DA 클라이언트(102)는 사용자 대면 입력 및 출력 처리, 및 DA 서버(106)와의 통신과 같은 클라이언트 측 기능들을 제공한다. DA 서버(106)는 개개의 사용자 디바이스(104) 상에 각각 상주하는 임의의 수의 DA 클라이언트들(102)에 서버 측 기능들을 제공한다.

[0018] 일부 예들에서, DA 서버(106)는 클라이언트 대면 I/O 인터페이스(112), 하나 이상의 처리 모듈들(114), 데이터 및 모델들(116), 및 외부 서비스들에 대한 I/O 인터페이스(118)를 포함한다. 클라이언트 대면 I/O 인터페이스(112)는 DA 서버(106)에 대한 클라이언트 대면 입력 및 출력 처리를 용이하게 한다. 하나 이상의 처리 모듈들(114)은 데이터 및 모델들(116)을 이용하여 스피치 입력을 처리하고 자연 언어 입력에 기초하여 사용자의 의도를 결정한다. 또한, 하나 이상의 처리 모듈들(114)은 추론된 사용자 의도에 기초하여 태스크 실행을 수행한다. 일부 예들에서, DA 서버(106)는 태스크 완수 또는 정보 획득을 위해 네트워크(들)(110)를 통해 외부 서비스들(120)과 통신한다. 외부 서비스들에 대한 I/O 인터페이스(118)는 그러한 통신을 용이하게 한다.

[0019] 사용자 디바이스(104)는 임의의 적합한 전자 디바이스일 수 있다. 일부 예들에서, 사용자 디바이스(104)는 휴대용 다기능 디바이스(예컨대, 도 2a를 참조하여 후술되는 디바이스(200)), 다기능 디바이스(예컨대, 도 4를 참조하여 후술되는 디바이스(400)), 또는 개인용 전자 디바이스(예컨대, 도 6a 및 도 6b를 참조하여 후술되는 디바이스(600))이다. 휴대용 다기능 디바이스는, 예를 들어, PDA 및/또는 음악 재생기 기능들과 같은 다른 기능들을 또한 포함하는 모바일 전화기이다. 휴대용 다기능 디바이스들의 구체적인 예들은 미국 캘리포니아주 쿠파티노 소재의 애플 인크.(Apple Inc.)로부터의 애플워치(Apple Watch)®, 아이폰®, 아이팟 터치(iPod Touch)®, 및 아이패드(iPad)® 디바이스들을 포함한다. 휴대용 다기능 디바이스들의 다른 예들은 이어폰/헤드폰, 스피커, 및 랩톱 또는 태블릿 컴퓨터들을 제한 없이 포함한다. 또한, 일부 예들에서, 사용자 디바이스(104)는 비휴대용 다기능 디바이스이다. 특히, 사용자 디바이스(104)는 데스크톱 컴퓨터, 게임 콘솔, 스피커, 텔레비전, 또는 텔레비전 셋톱박스이다. 일부 예들에서, 사용자 디바이스(104)는 터치 감응형 표면(예컨대, 터치 스크린 디스플레이들 및/또는 터치패드들)을 포함한다. 또한, 사용자 디바이스(104)는 선택적으로 물리적 키보드, 마우스 및/또는 조이스틱과 같은 하나 이상의 다른 물리적 사용자 인터페이스 디바이스들을 포함한다. 다기능 디바이스들과 같은 전자 디바이스들의 다양한 예들이 더욱 상세하게 하기에 기술된다.

[0020] 통신 네트워크(들)(110)의 예들은 근거리 통신망(local area network, LAN) 및 광역 통신망(wide area network, WAN), 예컨대, 인터넷을 포함한다. 통신 네트워크(들)(110)는, 예를 들어 이더넷(Ethernet), 범용 직렬 버스(Universal Serial Bus, USB), 파이어와이어(FIREWIRE), GSM(Global System for Mobile Communications), EDGE(Enhanced Data GSM Environment), 코드 분할 다중 접속(code division multiple

access, CDMA), 시간 분할 다중 접속(time division multiple access, TDMA), 블루투스(Bluetooth), Wi-Fi, VoIP(voice over Internet Protocol), Wi-MAX, 또는 임의의 다른 적합한 통신 프로토콜과 같은 다양한 유선 또는 무선 프로토콜들을 포함하는 임의의 알려진 네트워크 프로토콜을 이용하여 구현된다.

[0021] 서버 시스템(108)은 하나 이상의 독립형 데이터 처리 장치 또는 분산형 컴퓨터 네트워크 상에 구현된다. 일부 예들에서, 서버 시스템(108)은 서버 시스템(108)의 기본 컴퓨팅 리소스들 및/또는 인프라구조 리소스들을 제공하기 위해 제3자 서비스 제공자들(예컨대, 제3자 클라우드 서비스 제공자들)의 다양한 가상 디바이스들 및/또는 서비스들을 또한 채용한다.

[0022] 일부 예들에서, 사용자 디바이스(104)는 제2 사용자 디바이스(122)를 통해 DA 서버(106)와 통신한다. 제2 사용자 디바이스(122)는 사용자 디바이스(104)와 유사 또는 동일하다. 예를 들어, 제2 사용자 디바이스(122)는 도 2a, 도 4, 및 도 6a와 도 6b를 참조하여 하기에 기술되는 디바이스들(200, 400, 또는 600)과 유사하다. 사용자 디바이스(104)는 직접 통신 접속, 예컨대, 블루투스, NFC, BTLE 등을 통해, 또는 유선 또는 무선 네트워크, 예컨대, 로컬 Wi-Fi 네트워크를 통해 제2 사용자 디바이스(122)에 통신가능하게 결합되도록 구성된다. 일부 예들에서, 제2 사용자 디바이스(122)는 사용자 디바이스(104)와 DA 서버(106) 사이의 프록시(proxy)로서 동작하도록 구성된다. 예를 들어, 사용자 디바이스(104)의 DA 클라이언트(102)는 제2 사용자 디바이스(122)를 통해 DA 서버(106)로 정보(예컨대, 사용자 디바이스(104)에서 수신된 사용자 요청)를 송신하도록 구성된다. DA 서버(106)는 정보를 처리하고 제2 사용자 디바이스(122)를 통해 사용자 디바이스(104)로 관련 데이터(예컨대, 사용자 요청에 응답하는 데이터 콘텐츠)를 반환한다.

[0023] 일부 예들에서, 사용자 디바이스(104)는 사용자 디바이스(104)로부터 송신되는 정보의 양을 감소시키기 위해 데이터에 대한 단축 요청(abbreviated request)들을 제2 사용자 디바이스(122)로 통신하도록 구성된다. 제2 사용자 디바이스(122)는 DA 서버(106)로 송신할 완전한 요청을 생성하기 위해 단축 요청에 추가하기 위한 보완 정보를 결정하도록 구성된다. 이러한 시스템 아키텍처는, 유리하게도, 제한된 통신 능력들 및/또는 제한된 배터리 전력을 갖는 사용자 디바이스(104)(예컨대, 시계 또는 유사한 소형 전자 디바이스)가, 보다 큰 통신 능력들 및/또는 배터리 전력을 갖는 제2 사용자 디바이스(122)(예컨대, 모바일 폰, 랩톱 컴퓨터, 태블릿 컴퓨터 등)를 DA 서버(106)에 대한 프록시로서 사용함으로써, DA 서버(106)에 의해 제공되는 서비스들에 액세스하도록 할 수 있다. 2개의 사용자 디바이스들(104, 122)만이 도 1에 도시되어 있지만, 일부 예들에서, 시스템(100)은 DA 서버 시스템(106)과 통신하기 위해 이러한 프록시 구성으로 구성되는 임의의 개수 및 유형의 사용자 디바이스들을 포함한다는 것이 이해되어야 한다.

[0024] 도 1에 도시된 디지털 어시스턴트가 클라이언트 측 부분(예컨대, DA 클라이언트(102)) 및 서버 측 부분(예컨대, DA 서버(106)) 둘 모두를 포함하지만, 일부 예들에서, 디지털 어시스턴트의 기능들은 사용자 디바이스 상에 설치된 독립형 애플리케이션으로서 구현된다. 게다가, 디지털 어시스턴트의 클라이언트 부분과 서버 부분 사이의 기능들의 분담은 상이한 구현예들에서 다를 수 있다. 예를 들어, 일부 예들에서, DA 클라이언트는, 오로지 사용자 대면 입력 및 출력 처리 기능들만을 제공하고 디지털 어시스턴트의 다른 모든 기능들을 백엔드 서버에 위임하는 쉘-클라이언트(thin-client)이다.

[0025] 2. 전자 디바이스

[0026] 이제, 디지털 어시스턴트의 클라이언트 측 부분을 구현하기 위한 전자 디바이스들의 실시예들에 주목한다. 도 2a는 일부 실시예들에 따른, 터치 감응형 디스플레이 시스템(212)을 갖는 휴대용 다기능 디바이스(200)를 도시하는 블록도이다. 터치 감응형 디스플레이(212)는 때때로 편의상 "터치 스크린"이라고 지칭되고, 때때로 "터치 감응형 디스플레이 시스템"으로 알려져 있거나 또는 그렇게 지칭된다. 디바이스(200)는 메모리(202)(선택적으로, 하나 이상의 컴퓨터 판독가능 저장 매체들을 포함함), 메모리 제어기(222), 하나 이상의 처리 유닛(CPU)들(220), 주변기기 인터페이스(218), RF 회로부(208), 오디오 회로부(210), 스피커(211), 마이크로폰(213), 입/출력(I/O) 서브시스템(206), 다른 입력 제어 디바이스들(216), 및 외부 포트(224)를 포함한다. 디바이스(200)는 선택적으로 하나 이상의 광 센서들(264)을 포함한다. 디바이스(200)는, 선택적으로, 디바이스(200)(예컨대, 디바이스(200)의 터치 감응형 디스플레이 시스템(212)과 같은 터치 감응형 표면) 상의 접촉들의 세기를 검출하기 위한 하나 이상의 접촉 세기 센서들(265)을 포함한다. 디바이스(200)는 선택적으로 디바이스(200) 상에 측각적 출력들을 생성하기 위한(예를 들어, 디바이스(200)의 터치 감응형 디스플레이 시스템(212) 또는 디바이스(400)의 터치패드(455)와 같은 터치 감응형 표면 상에 측각적 출력들을 생성하기 위한) 하나 이상의 측각적 출력 생성기들(267)을 포함한다. 이들 컴포넌트는 선택적으로 하나 이상의 통신 버스들 또는 신호 라인들(203)을 통해 통신한다.

[0027] 명세서 및 청구범위에서 사용되는 바와 같이, 터치 감응형 표면 상에서의 접촉의 "세기"라는 용어는 터치 감응형 표면 상에서의 접촉(예컨대, 손가락 접촉)의 힘 또는 압력(단위 면적 당 힘), 또는 터치 감응형 표면 상에서의 접촉의 힘 또는 압력에 대한 대체물(대용물(proxy))을 지칭한다. 접촉의 세기는, 적어도 4개의 구별되는 값들을 포함하고 더 전형적으로는 수백 개(예컨대, 적어도 256개)의 구별되는 값들을 포함하는 일정 범위의 값들을 갖는다. 접촉의 세기는 다양한 접근법들, 및 다양한 센서들 또는 센서들의 조합들을 이용하여, 선택적으로 결정(또는 측정)된다. 예를 들어, 터치 감응형 표면 아래의 또는 그에 인접한 하나 이상의 힘 센서들은 터치 감응형 표면 상의 다양한 지점들에서 힘을 측정하는데 선택적으로 사용된다. 일부 구현예들에서는, 다수의 힘 센서들로부터의 힘 측정치들이 접촉의 추정되는 힘을 결정하기 위해 조합(예컨대, 가중 평균)된다. 유사하게, 스타일러스의 압력 감응형 팁(tip)이 터치 감응형 표면 상의 스타일러스의 압력을 결정하는 데 선택적으로 사용된다. 대안적으로, 터치 감응형 표면 상에서 검출된 접촉 면적의 크기 및/또는 그에 대한 변화들, 접촉 부근의 터치 감응형 표면의 정전용량 및/또는 그에 대한 변화들, 및/또는 접촉 부근의 터치 감응형 표면의 저항 및/또는 그에 대한 변화들은 터치 감응형 표면 상의 접촉의 힘 또는 압력에 대한 대체물로서 선택적으로 이용된다. 일부 구현예들에서, 접촉 힘 또는 압력에 대한 대체 측정치들은 세기 임계치가 초과되었는지 여부를 결정하는데 직접 이용된다(예컨대, 세기 임계치는 대체 측정치들에 대응하는 단위로 기술된다). 일부 구현예들에서, 접촉 힘 또는 압력에 대한 대체 측정치들은 추정된 힘 또는 압력으로 변환되고, 추정된 힘 또는 압력은 세기 임계치가 초과되었는지 여부를 결정하기 위해 이용된다(예를 들어, 세기 임계치는 압력의 단위로 측정된 압력 임계치이다). 사용자 입력의 속성으로서 접촉의 세기를 사용하는 것은, 그렇지 않았으면 어포던스들을(예를 들어, 터치 감응형 디스플레이 상에) 디스플레이하고/하거나(예를 들어, 터치 감응형 디스플레이, 터치 감응형 표면, 또는 노브(knob) 또는 버튼과 같은 물리적/기계적 제어부를 통해) 사용자 입력을 수신하기 위하여 한정된 실면적을 갖는 감소된 크기의 디바이스 상에서 사용자에게 의해 액세스 가능하지 않을 수 있는 부가적인 디바이스 기능에의 사용자 액세스를 가능하게 한다.

[0028] 명세서 및 청구범위에 사용되는 바와 같이, "촉각적 출력"이라는 용어는 디바이스의 이전 위치에 대한 디바이스의 물리적 변위, 디바이스의 다른 컴포넌트(예컨대, 하우징)에 대한 디바이스의 컴포넌트(예컨대, 터치 감응형 표면)의 물리적 변위, 또는 사용자의 촉각을 이용하여 사용자에게 의해 검출된 디바이스의 질량 중심에 대한 컴포넌트의 변위를 지칭한다. 예컨대, 디바이스 또는 디바이스의 컴포넌트가 터치에 민감한 사용자의 표면(예를 들어, 사용자의 손의 손가락, 손바닥, 또는 다른 부위)과 접촉하는 상황에서, 물리적 변위에 의해 생성된 촉각적 출력은 사용자에게 의해 디바이스 또는 디바이스의 컴포넌트의 물리적 특성들의 인지된 변화에 대응하는 촉감(tactile sensation)으로서 해석될 것이다. 예컨대, 터치 감응형 표면(예컨대, 터치 감응형 디스플레이 또는 트랙패드)의 이동은, 선택적으로, 사용자에게 의해 물리적 액추에이터 버튼의 "다운 클릭" 또는 "업 클릭"으로서 해석된다. 일부 경우들에서, 사용자는 사용자의 이동에 의해 물리적으로 눌리는(예를 들어, 변위되는) 터치 감응형 표면과 연관된 물리적 액추에이터 버튼의 이동이 없는 경우에도 "다운 클릭" 또는 "업 클릭"과 같은 촉감을 느낄 것이다. 다른 예로서, 터치 감응형 표면의 이동은, 선택적으로, 터치 감응형 표면의 평활도(smoothness)에서의 변화가 존재하지 않는 경우에도, 터치 감응형 표면의 "조도(roughness)"로서 사용자에게 의해 해석 또는 감지된다. 사용자에게 의한 터치의 이러한 해석들이 사용자의 개별화된 감각 인지(sensory perception)에 영향을 받기 쉬울 것이지만, 대다수의 사용자들에게 보편적인 많은 터치 감각 인지가 있다. 따라서, 촉각적 출력이 사용자의 특정 감각 인지(예를 들어, "업 클릭", "다운 클릭", "거칠기")에 대응하는 것으로서 기술될 때, 달리 언급되지 않는다면, 생성된 촉각적 출력은 전형적인(또는 평균적인) 사용자에게 대한 기술된 감각 인지를 생성할 디바이스 또는 그의 컴포넌트의 물리적 변위에 대응한다.

[0029] 디바이스(200)는 휴대용 다기능 디바이스의 일례일 뿐이고, 디바이스(200)는, 선택적으로, 도시된 것보다 더 많거나 더 적은 컴포넌트들을 갖거나, 선택적으로, 둘 이상의 컴포넌트들을 조합하거나, 또는 선택적으로 컴포넌트들의 상이한 구성 또는 배열을 갖는다는 것이 이해되어야 한다. 도 2a에 도시된 다양한 컴포넌트들은 하나 이상의 신호 처리 및/또는 주문형 집적 회로(application-specific integrated circuit)들을 비롯한, 하드웨어, 소프트웨어, 또는 하드웨어와 소프트웨어 둘 모두의 조합으로 구현된다.

[0030] 메모리(202)는 하나 이상의 컴퓨터 판독가능 저장 매체들을 포함한다. 컴퓨터 판독가능 저장 매체들은, 예를 들어, 유형적(tangible)이고 비일시적이다. 메모리(202)는 고속 랜덤 액세스 메모리를 포함하고, 또한 하나 이상의 자기 디스크 저장 디바이스, 플래시 메모리 디바이스, 또는 다른 비휘발성 솔리드 스테이트 메모리 디바이스(non-volatile solid-state memory device)와 같은 비휘발성 메모리를 포함한다. 메모리 제어기(222)는 디바이스(200)의 다른 컴포넌트들에 의한 메모리(202)에의 액세스를 제어한다.

[0031] 일부 예들에서, 메모리(202)의 비일시적 컴퓨터 판독가능 저장 매체는, 명령어 실행 시스템, 장치, 또는 디바이

스, 예컨대, 컴퓨터 기반 시스템, 프로세서-포함(processor-containing) 시스템, 또는 명령어 실행 시스템, 장치, 또는 디바이스로부터 명령어들을 폐치하여 명령어들을 실행할 수 있는 다른 시스템에 의해 또는 그와 관련하여 사용하기 위한 (예컨대, 후술되는 프로세스들의 양태들을 수행하기 위한) 명령어들을 저장하는 데 사용된다. 다른 예들에서, (예컨대, 후술되는 프로세스들의 양태들을 수행하기 위한) 명령어들은 서버 시스템(108)의 비일시적 컴퓨터 판독가능 저장 매체(도시되지 않음) 상에 저장되거나 또는 메모리(202)의 비일시적 컴퓨터 판독가능 저장 매체와 서버 시스템(108)의 비일시적 컴퓨터 판독가능 저장 매체 사이에서 분담된다.

[0032] 주변기기 인터페이스(218)는 디바이스의 입력 및 출력 주변기기들을 CPU(220) 및 메모리(202)에 결합하는 데 사용된다. 하나 이상의 프로세서들(220)은 디바이스(200)에 대한 다양한 기능들을 수행하기 위해 그리고 데이터를 처리하기 위해 메모리(202)에 저장된 다양한 소프트웨어 프로그램들 및/또는 명령어들의 세트들을 구동 또는 실행시킨다. 일부 실시예들에서, 주변기기 인터페이스(218), CPU(220) 및 메모리 제어기(222)는 칩(204)과 같은 단일 칩 상에서 구현된다. 일부 다른 실시예들에서, 이들은 별개의 칩들 상에서 구현된다.

[0033] RF(radio frequency) 회로부(208)는 전자기 신호들이라고도 지칭되는 RF 신호들을 수신 및 전송한다. RF 회로부(208)는 전기 신호들을 전자기 신호들로/로부터 변환하고, 전자기 신호들을 통해 통신 네트워크들 및 다른 통신 디바이스들과 통신한다. RF 회로부(208)는, 선택적으로, 안테나 시스템, RF 송수신기, 하나 이상의 증폭기들, 튜너, 하나 이상의 발진기들, 디지털 신호 프로세서, CODEC 칩셋, SIM(subscriber identity module) 카드, 메모리 등을 포함하지만 이들로 제한되지 않는, 이러한 기능들을 수행하기 위한 잘 알려진 회로부를 포함한다. RF 회로부(208)는, 선택적으로, 네트워크들, 예컨대 월드 와이드 웹(WWW)으로도 지칭되는 인터넷, 인트라넷, 및/또는 무선 네트워크, 예컨대 셀룰러 전화 네트워크, 무선 LAN(local area network) 및/또는 MAN(metropolitan area network), 및 다른 디바이스들과 무선 통신에 의해 통신한다. RF 회로부(208)는, 선택적으로, 예컨대 단거리 통신 무선기기(short-range communication radio)에 의해, 근거리 통신(near field communication, NFC) 필드들을 검출하기 위한 잘 알려진 회로부를 포함한다. 무선 통신은, 선택적으로, GSM(Global System for Mobile Communications), EDGE(Enhanced Data GSM Environment), HSDPA(high-speed downlink packet access), HSUPA(high-speed uplink packet access), EV-DO(Evolution, Data-Only), HSPA, HSPA+, DC-HSPDA(Dual-Cell HSPA), LTE(long term evolution), 근거리 통신(NFC), W-CDMA(wideband code division multiple access), CDMA(code division multiple access), TDMA(time division multiple access), 블루투스(Bluetooth), BTLE(Bluetooth Low Energy), Wi-Fi(Wireless Fidelity)(예를 들어, IEEE 802.11a, IEEE 802.11b, IEEE 802.11g, IEEE 802.11n, 및/또는 IEEE 802.11ac), VoIP(voice over Internet Protocol), Wi-MAX, 이메일용 프로토콜(예를 들어, IMAP(Internet message access protocol) 및/또는 POP(post office protocol)), 인스턴트 메시징(예를 들어, XMPP(extensible messaging and presence protocol), SIMPLE(Session Initiation Protocol for Instant Messaging and Presence Leveraging Extensions), IMPS(Instant Messaging and Presence Service)), 및/또는 SMS(Short Message Service), 또는 본 문헌의 출원일 현재 아직 개발되지 않은 통신 프로토콜들을 포함한 임의의 다른 적합한 통신 프로토콜을 포함하지만 이들로 제한되지 않는 복수의 통신 표준들, 프로토콜들 및 기술들 중 임의의 것을 사용한다.

[0034] 오디오 회로부(210), 스피커(211), 및 마이크로폰(213)은 사용자와 디바이스(200) 사이에서 오디오 인터페이스를 제공한다. 오디오 회로부(210)는 주변기기 인터페이스(218)로부터 오디오 데이터를 수신하고, 그 오디오 데이터를 전기 신호로 변환하고, 그 전기 신호를 스피커(211)로 송신한다. 스피커(211)는 전기 신호를 사람이 들을 수 있는 음파로 변환한다. 오디오 회로부(210)는 또한 마이크로폰(213)에 의해 음파로부터 변환된 전기 신호들을 수신한다. 오디오 회로부(210)는 전기 신호를 오디오 데이터로 변환하고, 처리를 위해 오디오 데이터를 주변기기 인터페이스(218)로 송신한다. 오디오 데이터는 주변기기 인터페이스(218)에 의해 메모리(202) 및/또는 RF 회로부(208)로부터 인출되고/되거나 이로 송신된다. 일부 실시예들에서, 오디오 회로부(210)는 또한 헤드셋 잭(예컨대, 도 3의 312)을 포함한다. 헤드셋 잭은 출력-전용 헤드폰들, 또는 출력부(예를 들어, 한쪽 또는 양쪽 귀용 헤드폰) 및 입력부(예를 들어, 마이크로폰) 양쪽 모두를 갖는 헤드셋과 같은 착탈가능한 오디오 입/출력 주변기기들과 오디오 회로부(210) 사이의 인터페이스를 제공한다.

[0035] I/O 서브시스템(206)은 터치 스크린(212) 및 다른 입력 제어 디바이스들(216)과 같은, 디바이스(200) 상의 입/출력 주변기기들을 주변기기 인터페이스(218)에 결합시킨다. I/O 서브시스템(206)은 선택적으로 디스플레이 제어기(256), 광 센서 제어기(258), 세기 센서 제어기(259), 햅틱 피드백 제어기(261), 및 다른 입력 또는 제어 디바이스들을 위한 하나 이상의 입력 제어기들(260)을 포함한다. 하나 이상의 입력 제어기들(260)은 다른 입력 제어 디바이스들(216)로부터/로 전기 신호들을 수신/전송한다. 다른 입력 제어 디바이스들(216)은 선택적으로 물리적 버튼들(예컨대, 푸시 버튼(push button), 로커 버튼(rocker button) 등), 다이얼, 슬라이더 스위치, 조

이스틱, 클릭 휠 등을 포함한다. 일부 대안적인 실시예들에서, 입력 제어기(들)(260)는 선택적으로 키보드, 적외선 포트, USB 포트, 및 마우스와 같은 포인터 디바이스 중 임의의 것에 결합된다(또는 어떤 것에도 결합되지 않는다). 하나 이상의 버튼들(예컨대, 도 3의 308)은, 선택적으로, 스피커(211) 및/또는 마이크로폰(213)의 음량 제어를 위한 업/다운 버튼을 포함한다. 하나 이상의 버튼들은 선택적으로 푸시 버튼(예컨대, 도 3의 306)을 포함한다.

[0036] 푸시 버튼의 빠른 누르기(quick press)는 터치 스크린(212)의 잠금을 풀거나, 디바이스를 잠금해제하기 위해 터치 스크린 상의 제스처들을 사용하는 프로세스를 시작하며, 이는 2005년 12월 23일자로 출원된 미국 특허 출원 제11/322,549호(미국 특허 제7,657,849호), "Unlocking a Device by Performing Gestures on an Unlock Image"에 설명된 바와 같고, 이로써 그 출원은 그 전체가 참고로 포함된다. 푸시 버튼(예컨대, 306)의 더 긴 누르기는 디바이스(200)의 전원을 켜거나 끈다. 사용자는 버튼들 중 하나 이상의 버튼의 기능을 맞춤화할 수 있다. 터치 스크린(212)은 가상 또는 소프트 버튼들 및 하나 이상의 소프트 키보드들을 구현하는데 사용된다.

[0037] 터치 감응형 디스플레이(212)는 디바이스와 사용자 사이의 입력 인터페이스 및 출력 인터페이스를 제공한다. 디스플레이 제어기(256)는 터치 스크린(212)으로부터/으로 전기 신호들을 수신하고/하거나 전송한다. 터치 스크린(212)은 사용자에게 시각적 출력을 디스플레이한다. 시각적 출력은 그래픽, 텍스트, 아이콘들, 비디오 및 이들의 임의의 조합("그래픽"으로 총칭함)을 포함한다. 일부 실시예들에서, 시각적 출력의 일부 또는 전부가 사용자 인터페이스 객체들에 대응한다.

[0038] 터치 스크린(212)은 햅틱 및/또는 촉각적 접촉에 기초하는 사용자로부터의 입력을 수용하는 터치 감응형 표면, 센서 또는 센서들의 세트를 갖는다. 터치 스크린(212) 및 디스플레이 제어기(256)는 (메모리(202) 내의 임의의 연관된 모듈들 및/또는 명령어들의 세트들과 함께) 터치 스크린(212) 상의 접촉(및 접촉의 임의의 이동 또는 중단)을 검출하고, 검출된 접촉을 터치 스크린(212) 상에 디스플레이된 사용자 인터페이스 객체들(예컨대, 하나 이상의 소프트 키들, 아이콘들, 웹페이지들 또는 이미지들)과의 상호작용으로 변환한다. 예시적인 실시예에서, 터치 스크린(212)과 사용자 사이의 접촉 지점은 사용자의 손가락에 대응한다.

[0039] 터치 스크린(212)은, LCD(liquid crystal display) 기술, LPD(light emitting polymer display) 기술, 또는 LED(light emitting diode) 기술을 사용하지만, 다른 실시예들에서는 다른 디스플레이 기술들이 사용될 수 있다. 터치 스크린(212) 및 디스플레이 제어기(256)는, 용량성, 저항성, 적외선, 및 표면 음향파 기술들뿐만 아니라 다른 근접 센서 어레이들, 또는 터치 스크린(212)과의 하나 이상의 접촉 지점들을 결정하기 위한 다른 요소들을 포함하지만 이들로 제한되지 않는, 현재 알려져 있거나 추후에 개발되는 복수의 터치 감지 기술들 중 임의의 것을 사용하여, 접촉 및 그의 임의의 이동 또는 중단을 검출한다. 예시적인 실시예에서, 미국 캘리포니아주 쿠파티노 소재의 애플 인크.로부터의 아이폰® 및 아이패드®에서 발견되는 것과 같은 투영형 상호 정전용량 감지 기술(projected mutual capacitance sensing technology)이 이용된다.

[0040] 터치 스크린(212)의 일부 실시예들에서의 터치 감응형 디스플레이는 하기 미국 특허들: 제6,323,846호(Westerman 외), 제6,570,557호(Westerman 외), 및/또는 제6,677,932호(Westerman), 및/또는 미국 특허 공개 제2002/0015024A1호에 기재된 다중-터치 감응형 터치패드들과 유사하며, 이들 각각은 그 전체가 참고로 본 명세서에 포함된다. 그러나, 터치 스크린(212)은 디바이스(200)로부터의 시각적 출력을 디스플레이하는 반면, 터치 감응형 터치패드들은 시각적 출력을 제공하지 않는다.

[0041] 터치 스크린(212)의 일부 실시예들에서의 터치 감응형 디스플레이는 하기 출원들에 기재된 바와 같다: (1) 2006년 5월 2일자로 출원된 미국 특허 출원 제11/381,313호, "Multipoint Touch Surface Controller"; (2) 2004년 5월 6일자로 출원된 미국 특허 출원 제10/840,862호, "Multipoint Touchscreen"; (3) 2004년 7월 30일자로 출원된 미국 특허 출원 제10/903,964호, "Gestures For Touch Sensitive Input Devices"; (4) 2005년 1월 31일자로 출원된 미국 특허 출원 제11/048,264호, "Gestures For Touch Sensitive Input Devices"; (5) 2005년 1월 18일자로 출원된 미국 특허 출원 제11/038,590호, "Mode-Based Graphical User Interfaces For Touch Sensitive Input Devices"; (6) 2005년 9월 16일자로 출원된 미국 특허 출원 제11/228,758호, "Virtual Input Device Placement On A Touch Screen User Interface"; (7) 2005년 9월 16일자로 출원된 미국 특허 출원 제11/228,700호, "Operation Of A Computer With A Touch Screen Interface"; (8) 2005년 9월 16일자로 출원된 미국 특허 출원 제11/228,737호, "Activating Virtual Keys Of A Touch-Screen Virtual Keyboard"; 및 (9) 2006년 3월 3일자로 출원된 미국 특허 출원 제11/367,749호, "Multi-Functional Hand-Held Device". 이 출원들 모두는 그 전체가 본 명세서에 참고로 포함된다.

[0042] 터치 스크린(212)은, 예를 들어, 100 dpi를 초과하는 비디오 해상도를 갖는다. 일부 실시예들에서, 터치 스크

린은 대략 160 dpi의 비디오 해상도를 갖는다. 사용자는 스타일러스, 손가락 등과 같은 임의의 적합한 물체 또는 부속물을 사용하여 터치 스크린(212)과 접촉한다. 일부 실시예들에서, 사용자 인터페이스는 주로 손가락 기반 접촉들 및 제스처들을 이용하여 동작하도록 설계되는데, 이는 터치 스크린 상에서의 손가락의 더 넓은 접촉 면적으로 인해 스타일러스 기반 입력보다 덜 정밀할 수 있다. 일부 실시예들에서, 디바이스는 대략적인 손가락 기반 입력을 사용자가 원하는 동작들을 수행하기 위한 정밀한 포인터/커서 포지션 또는 커맨드로 변환한다.

[0043] 일부 실시예들에서, 터치 스크린 이외에, 디바이스(200)는 특정 기능들을 활성화 또는 비활성화하기 위한 터치패드(도시되지 않음)를 포함한다. 일부 실시예들에서, 터치패드는, 터치 스크린과는 달리, 시각적 출력을 디스플레이하지 않는 디바이스의 터치 감응형 영역이다. 터치패드는 터치 스크린(212)과는 별개인 터치 감응형 표면 또는 터치 스크린에 의해 형성되는 터치 감응형 표면의 연장부이다.

[0044] 디바이스(200)는 또한 다양한 컴포넌트들에 전력을 공급하기 위한 전력 시스템(262)을 포함한다. 전력 시스템(262)은 전력 관리 시스템, 하나 이상의 전원들(예를 들어, 배터리, 교류 전류(alternating current, AC)), 재충전 시스템, 전력 고장 검출 회로, 전력 변환기 또는 인버터, 전력 상태 표시자(예를 들어, 발광 다이오드(LED)), 및 휴대용 디바이스들 내에서의 전력의 생성, 관리 및 분배와 연관된 임의의 다른 컴포넌트들을 포함한다.

[0045] 디바이스(200)는 또한 하나 이상의 광 센서들(264)을 포함한다. 도 2a는 I/O 서브시스템(206) 내의 광 센서 제어기(258)에 결합되는 광 센서를 도시한다. 광 센서(264)는 CCD(charge-coupled device) 또는 CMOS(complementary metal-oxide semiconductor) 포토트랜지스터들을 포함한다. 광 센서(264)는 하나 이상의 렌즈들을 통해 투영되는, 주변환경으로부터의 광을 수광하고, 그 광을 이미지를 표현하는 데이터로 변환한다. 이미징 모듈(243)(카메라 모듈이라고도 지칭됨)과 함께, 광 센서(264)는 정지 이미지들 또는 비디오를 캡처한다. 일부 실시예들에서, 광 센서는 디바이스(200) 전면 상의 터치 스크린 디스플레이(212)의 반대편인 디바이스의 배면 상에 위치되어, 터치 스크린 디스플레이가 정지 및/또는 비디오 이미지 획득을 위한 뷰파인더로서 사용되게 한다. 일부 실시예들에서, 광 센서가 디바이스의 전면 상에 위치되어, 사용자가 터치 스크린 디스플레이 상에서 다른 화상 회의 참여자들을 보는 동안, 사용자의 이미지가 화상 회의를 위해 획득되게 한다. 일부 실시예들에서, 광 센서(264)의 위치는 (예를 들어, 디바이스 하우징 내의 렌즈 및 센서를 회전시킴으로써) 사용자에게 의해 변경될 수 있어, 단일 광 센서(264)가 터치 스크린 디스플레이와 함께 화상 회의와 정지 및/또는 비디오 이미지 획득 둘 모두에 사용된다.

[0046] 디바이스(200)는, 또한, 선택적으로, 하나 이상의 접촉 세기 센서들(265)을 포함한다. 도 2a는 I/O 서브시스템(206) 내의 세기 센서 제어기(259)에 결합된 접촉 세기 센서를 도시한다. 접촉 세기 센서(265)는, 선택적으로, 하나 이상의 압전 저항 스트레인 게이지, 용량성 힘 센서, 전기적 힘 센서, 압전 힘 센서, 광학적 힘 센서, 용량성 터치 감응형 표면, 또는 다른 세기 센서들(예컨대, 터치 감응형 표면 상에서의 접촉의 힘(또는 압력)을 측정하는데 사용되는 센서들)을 포함한다. 접촉 세기 센서(265)는 주변환경으로부터 접촉 세기 정보(예컨대, 압력 정보 또는 압력 정보에 대한 대응물)를 수신한다. 일부 실시예들에서, 적어도 하나의 접촉 세기 센서는 터치 감응형 표면(예컨대, 터치 감응형 디스플레이 시스템(212))과 함께 위치(collocate)되거나 그에 근접한다. 일부 실시예들에서, 적어도 하나의 접촉 세기 센서가 디바이스(200)의 전면 상에 위치한 터치 스크린 디스플레이(212)의 반대편인 디바이스(200)의 배면 상에 위치된다.

[0047] 디바이스(200)는 또한 하나 이상의 근접 센서들(266)을 포함한다. 도 2a는 주변기기 인터페이스(218)에 결합된 근접 센서(266)를 도시한다. 대안적으로, 근접 센서(266)는 I/O 서브시스템(206) 내의 입력 제어기(260)에 결합된다. 근접 센서(266)는 미국 특허 출원들 제11/241,839호, "Proximity Detector In Handheld Device"; 제11/240,788호, "Proximity Detector In Handheld Device"; 제11/620,702호, "Using Ambient Light Sensor To Augment Proximity Sensor Output"; 제11/586,862호, "Automated Response To And Sensing Of User Activity In Portable Devices"; 및 제11/638,251호, "Methods And Systems For Automatic Configuration Of Peripherals"에 기술된 바와 같이 수행되며, 이들은 이로써 그들 전체가 참고로 포함된다. 일부 실시예들에서, 근접 센서는 다가능 디바이스가 사용자의 귀 근처에 위치될 때(예컨대, 사용자가 전화 통화를 하고 있을 때) 터치 스크린(212)을 끄고 디스플레이를 끈다.

[0048] 디바이스(200)는 또한, 선택적으로, 하나 이상의 촉각적 출력 생성기들(267)을 포함한다. 도 2a는 I/O 서브시스템(206) 내의 햅틱 피드백 제어기(261)에 결합된 촉각적 출력 생성기를 도시한다. 촉각적 출력 생성기(267)는, 선택적으로, 스피커들 또는 다른 오디오 컴포넌트들과 같은 하나 이상의 전자음향 디바이스들 및/또는 모터, 솔레노이드, 전기활성 중합체, 압전 액추에이터, 정전 액추에이터, 또는 다른 촉각적 출력 생성 컴포넌트

(예컨대, 전기 신호들을 디바이스 상의 촉각적 출력들로 변환하는 컴포넌트)와 같은, 에너지를 선형 모션(linear motion)으로 변환하는 전자기계 디바이스들을 포함한다. 접촉 세기 센서(265)는 햅틱 피드백 모듈(233)로부터 촉각적 피드백 생성 명령어들을 수신하여, 디바이스(200)의 사용자에게 의해 감지될 수 있는 디바이스(200) 상의 촉각적 출력들을 생성한다. 일부 실시예들에서, 적어도 하나의 촉각적 출력 생성기는 터치 감응형 표면(예컨대, 터치 감응형 디스플레이 시스템(212))과 함께 위치되거나 그에 근접하며, 선택적으로, 터치 감응형 표면을 수직으로(예컨대, 디바이스(200)의 표면 내/외로) 또는 측방향으로(예컨대, 디바이스(200)의 표면과 동일한 평면에서 전후로) 이동시킴으로써 촉각적 출력을 생성한다. 일부 실시예들에서, 적어도 하나의 촉각적 출력 생성기 센서는 디바이스(200)의 전면 상에 위치한 터치 스크린 디스플레이(212)의 반대편인 디바이스(200)의 배면 상에 위치된다.

[0049] 디바이스(200)는 또한 하나 이상의 가속도계들(268)을 포함한다. 도 2a는 주변기기 인터페이스(218)에 결합된 가속도계(268)를 도시한다. 대안적으로, 가속도계(268)는 I/O 서브시스템(206) 내의 입력 제어기(260)에 결합된다. 가속도계(268)는, 예를 들어, 미국 특허 공개 제20050190059호, "Acceleration-based Theft Detection System for Portable Electronic Devices", 및 미국 특허 공개 제20060017692호, "Methods And Apparatuses For Operating A Portable Device Based On An Accelerometer"에 기술된 바와 같이 수행하며, 이들 둘 모두는 그 전체가 본 명세서에 참고로 포함된다. 일부 실시예들에서, 하나 이상의 가속도계들로부터 수신된 데이터의 분석에 기초하여 터치 스크린 디스플레이 상에 세로보기(portrait view) 또는 가로보기(landscape view)로 정보가 디스플레이된다. 디바이스(200)는, 선택적으로, 가속도계(들)(268) 외에도, 자력계(도시되지 않음), 및 디바이스(200)의 위치 및 배향(예컨대, 세로 또는 가로)에 관한 정보를 획득하기 위한 GPS(또는 GLONASS 또는 다른 글로벌 내비게이션 시스템) 수신기(도시되지 않음)를 포함한다.

[0050] 일부 실시예들에서, 메모리(202)에 저장된 소프트웨어 컴포넌트들은 운영 체제(226), 통신 모듈(또는 명령어들의 세트)(228), 접촉/모션 모듈(또는 명령어들의 세트)(230), 그래픽 모듈(또는 명령어들의 세트)(232), 텍스트 입력 모듈(또는 명령어들의 세트)(234), GPS 모듈(또는 명령어들의 세트)(235), 디지털 어시스턴트 클라이언트 모듈(229), 및 애플리케이션들(또는 명령어들의 세트들)(236)을 포함한다. 또한, 메모리(202)는 데이터 및 모델들, 예컨대, 사용자 데이터 및 모델들(231)을 저장한다. 게다가, 일부 실시예들에서, 메모리(도 2a의 202 또는 도 4의 470)는 도 2a 및 도 4에 도시된 바와 같이 디바이스/글로벌 내부 상태(257)를 저장한다. 디바이스/글로벌 내부 상태(257)는, 존재하는 경우, 어느 애플리케이션들이 현재 활성 상태인지를 나타내는 활성 애플리케이션 상태; 어떤 애플리케이션들, 뷰들, 또는 다른 정보가 터치 스크린 디스플레이(212)의 다양한 영역들을 점유하는지를 나타내는 디스플레이 상태; 디바이스의 다양한 센서들 및 입력 제어 디바이스들(216)로부터 획득된 정보를 포함하는 센서 상태; 및 디바이스의 위치 및/또는 자세에 관한 위치 정보 중 하나 이상을 포함한다.

[0051] 운영 체제(226)(예컨대, Darwin, RTXC, LINUX, UNIX, OS X, iOS, WINDOWS, 또는 VxWorks와 같은 임베디드 운영 체제)는 일반적인 시스템 태스크들(예컨대, 메모리 관리, 저장 디바이스 제어, 전력 관리 등)을 제어 및 관리하기 위한 다양한 소프트웨어 컴포넌트들 및/또는 드라이버들을 포함하고, 다양한 하드웨어와 소프트웨어 컴포넌트들 사이의 통신을 용이하게 한다.

[0052] 통신 모듈(228)은 하나 이상의 외부 포트들(224)을 통한 다른 디바이스들과의 통신을 가능하게 하고, 또한 RF 회로부(208) 및/또는 외부 포트(224)에 의해 수신되는 데이터를 처리하기 위한 다양한 소프트웨어 컴포넌트들을 포함한다. 외부 포트(224)(예를 들어, 범용 직렬 버스(USB), 파이어와이어(FIREWIRE) 등)는 다른 디바이스들에 직접적으로 또는 네트워크(예를 들어, 인터넷, 무선 LAN 등)를 통해 간접적으로 결합하도록 구성된다. 일부 실시예들에서, 외부 포트는 아이팟®(애플 인크.의 상표) 디바이스들에서 사용되는 30-핀 커넥터와 동일하거나 유사하고/하거나 이와 호환가능한 멀티-핀(예컨대, 30-핀) 커넥터이다.

[0053] 접촉/모션 모듈(230)은, 선택적으로, (디스플레이 제어기(256)와 함께) 터치 스크린(212), 및 다른 터치 감응형 디바이스들(예컨대, 터치패드 또는 물리적 클릭 휠)과의 접촉을 검출한다. 접촉/모션 모듈(230)은 접촉이 발생했는지를 결정하는 것(예컨대, 손가락-다운 이벤트(finger-down event)를 검출하는 것), 접촉의 세기(예컨대, 접촉의 힘 또는 압력, 또는 접촉의 힘 또는 압력에 대한 대체물)를 결정하는 것, 접촉의 이동이 있는지를 결정하고 터치 감응형 표면을 가로지르는 이동을 추적하는 것(예컨대, 하나 이상의 손가락-드래깅 이벤트(finger-dragging event)들을 검출하는 것), 및 접촉이 중지되었는지를 결정하는 것(예컨대, 손가락-업 이벤트(finger-up event) 또는 접촉 중단을 검출하는 것)과 같은, 접촉의 검출에 관련된 다양한 동작들을 수행하기 위한 다양한 소프트웨어 컴포넌트들을 포함한다. 접촉/모션 모듈(230)은 터치 감응형 표면으로부터 접촉 데이터를 수신한다. 일련의 접촉 데이터에 의해 표현되는 접촉 지점의 이동을 결정하는 것은, 선택적으로, 접촉 지점의 속력(크기), 속도(크기 및 방향), 및/또는 가속도(크기 및/또는 방향의 변화)를 결정하는 것을 포함한다. 이 동작

들은, 선택적으로, 단일 접촉들(예컨대, 한 손가락 접촉들)에 또는 다수의 동시 접촉들(예컨대, "멀티터치"/다수의 손가락 접촉들)에 적용된다. 일부 실시예들에서, 접촉/모션 모듈(230) 및 디스플레이 제어기(256)는 터치패드 상의 접촉을 검출한다.

[0054] 일부 실시예들에서, 접촉/모션 모듈(230)은 동작이 사용자에게 의해 수행되었는지 여부를 결정하는 데(예컨대, 사용자가 아이콘에 대해 "클릭"했는지 여부를 결정하는 데) 하나 이상의 세기 임계치들의 세트를 이용한다. 일부 실시예들에서, 적어도 세기 임계치들의 서브세트가 소프트웨어 파라미터들에 따라 결정된다(예컨대, 세기 임계치들은 특정 물리적 액추에이터들의 활성화 임계치들에 의해 결정되지 않으며, 디바이스(200)의 물리적 하드웨어를 변경함이 없이 조정될 수 있다). 예를 들어, 트랙패드 또는 터치 스크린 디스플레이의 마우스 "클릭" 임계치는 트랙패드 또는 터치 스크린 디스플레이 하드웨어를 변경함이 없이 넓은 범위의 미리정의된 임계 값들 중 임의의 것으로 설정될 수 있다. 추가로, 일부 구현예들에서, 디바이스의 사용자는 (예컨대, 개별 세기 임계치들을 조정함으로써 그리고/또는 복수의 세기 임계치들을 시스템 레벨 클릭 "세기" 파라미터로 한꺼번에 조정함으로써) 세기 임계치들의 세트 중 하나 이상의 세기 임계치를 조정하기 위한 소프트웨어 설정들을 제공받는다.

[0055] 접촉/모션 모듈(230)은, 선택적으로, 사용자에게 의한 제스처 입력을 검출한다. 터치 감응형 표면 상에서의 상이한 제스처들은 상이한 접촉 패턴들(예를 들어, 검출된 접촉들의 상이한 모션들, 타이밍들, 및/또는 세기들)을 갖는다. 따라서, 제스처는, 선택적으로, 특정 접촉 패턴을 검출함으로써 검출된다. 예를 들어, 손가락 탭 제스처(finger tap gesture)를 검출하는 것은 손가락-다운 이벤트를 검출한 다음에 손가락-다운 이벤트와 동일한 위치(또는 실질적으로 동일한 위치)에서(예컨대, 아이콘의 위치에서) 손가락-업(리프트오프(liftoff)) 이벤트를 검출하는 것을 포함한다. 다른 예로서, 터치 감응형 표면 상에서 손가락 스와이프 제스처(finger swipe gesture)를 검출하는 것은 손가락-다운 이벤트를 검출한 다음에 하나 이상의 손가락-드래깅 이벤트들을 검출하고, 그에 후속하여 손가락-업(리프트오프) 이벤트를 검출하는 것을 포함한다.

[0056] 그래픽 모듈(232)은, 디스플레이되는 그래픽의 시각적 효과(예컨대, 밝기, 투명도, 채도, 콘트라스트 또는 다른 시각적 속성)를 변경하기 위한 컴포넌트들을 포함하는, 터치 스크린(212) 또는 다른 디스플레이 상에서 그래픽을 렌더링 및 디스플레이하기 위한 다양한 공지된 소프트웨어 컴포넌트들을 포함한다. 본 명세서에 사용되는 바와 같이, 용어 "그래픽"은 텍스트, 웹 페이지들, 아이콘들(예를 들어, 소프트 키들을 포함한 사용자 인터페이스 객체들), 디지털 이미지들, 비디오들, 애니메이션들 등을 제한 없이 포함하는, 사용자에게 디스플레이될 수 있는 임의의 객체를 포함한다.

[0057] 일부 실시예들에서, 그래픽 모듈(232)은 사용될 그래픽을 표현하는 데이터를 저장한다. 각각의 그래픽에는, 선택적으로, 대응하는 코드가 할당된다. 그래픽 모듈(232)은, 필요한 경우 좌표 데이터 및 다른 그래픽 속성 데이터와 함께, 디스플레이될 그래픽을 특정하는 하나 이상의 코드들을 애플리케이션들 등으로부터 수신하며, 이어서 스크린 이미지 데이터를 생성하여 디스플레이 제어기(256)에 출력한다.

[0058] 햅틱 피드백 모듈(233)은 디바이스(200)와의 사용자 상호작용들에 응답하여 디바이스(200) 상의 하나 이상의 위치들에서 촉각적 출력들을 생성하기 위하여 촉각적 출력 생성기(들)(267)에 의해 이용되는 명령어들을 생성하기 위한 다양한 소프트웨어 컴포넌트들을 포함한다.

[0059] 일부 예들에서, 그래픽 모듈(232)의 컴포넌트인 텍스트 입력 모듈(234)은 다양한 애플리케이션들(예컨대, 연락처(237), 이메일(240), IM(241), 브라우저(247), 및 텍스트 입력을 필요로 하는 임의의 다른 애플리케이션)에서 텍스트를 입력하기 위한 소프트 키보드들을 제공한다.

[0060] GPS 모듈(235)은 디바이스의 위치를 결정하고, 이 정보를 다양한 애플리케이션들에서의 사용을 위해 (예컨대, 위치 기반 다이얼링에서 사용하기 위해 전화(238)에; 사진/비디오 메타데이터로서 카메라(243)에; 그리고 날씨 위젯들, 지역 옐로 페이지 위젯들 및 지도/내비게이션 위젯들과 같은 위치 기반 서비스들을 제공하는 애플리케이션들에) 제공한다.

[0061] 디지털 어시스턴트 클라이언트 모듈(229)은 디지털 어시스턴트의 클라이언트 측 기능들을 제공하기 위한 다양한 클라이언트 측 디지털 어시스턴트 명령어들을 포함한다. 예를 들어, 디지털 어시스턴트 클라이언트 모듈(229)은 휴대용 다기능 디바이스(200)의 다양한 사용자 인터페이스들(예를 들어, 마이크로폰(213), 가속도계(들)(268), 터치 감응형 디스플레이 시스템(212), 광 센서(들)(264), 다른 입력 제어 디바이스들(216) 등)을 통하여 음성 입력(예를 들어, 스피치 입력), 텍스트 입력, 터치 입력, 및/또는 제스처 입력을 수용할 수 있다. 디지털 어시스턴트 클라이언트 모듈(229)은 또한, 휴대용 다기능 디바이스(200)의 다양한 출력 인터페이스들(예컨대, 스피커(211), 터치 감응형 디스플레이 시스템(212), 촉각적 출력 생성기(들)(267) 등)을 통하여 청각적

(예컨대, 스피치 출력), 시각적, 및/또는 촉각적 형태들의 출력을 제공할 수 있다. 예를 들어, 출력은 음성, 사운드, 경보들, 텍스트 메시지들, 메뉴들, 그래픽들, 비디오들, 애니메이션들, 진동들, 및/또는 이들 중 둘 이상의 것들의 조합들로서 제공된다. 동작 동안, 디지털 어시스턴트 클라이언트 모듈(229)은 RF 회로부(208)를 사용하여 DA 서버(106)와 통신한다.

[0062] 사용자 데이터 및 모델들(231)은, 디지털 어시스턴트의 클라이언트 측 기능들을 제공하기 위해 사용자와 연관된 다양한 데이터(예컨대, 사용자-특정 어휘 데이터, 사용자 선호도 데이터, 사용자-특정 이름 발음들, 사용자의 전자 주소록으로부터의 데이터, 할 일 목록들, 쇼핑 목록들 등)를 포함한다. 또한, 사용자 데이터 및 모델들(231)은 사용자 입력을 처리하고 사용자 의도를 결정하기 위해 다양한 모델들(예컨대, 스피치 인식 모델들, 통계 언어 모델들, 자연 언어 처리 모델들, 온톨로지, 태스크 흐름 모델들, 서비스 모델들 등)을 포함한다.

[0063] 일부 예들에서, 디지털 어시스턴트 클라이언트 모듈(229)은 사용자와 연관된 콘텍스트, 현재 사용자 상호작용, 및/또는 현재 사용자 입력을 확립하기 위해 휴대용 다기능 디바이스(200)의 다양한 센서들, 서브시스템들, 및 주변기기 디바이스들을 이용하여 휴대용 다기능 디바이스(200)의 주위 환경으로부터 추가 정보를 수집한다. 일부 예들에서, 디지털 어시스턴트 클라이언트 모듈(229)은 DA 서버(106)에 사용자 입력과 함께 콘텍스트 정보 또는 그의 서브셋을 제공하여 사용자의 의도를 추론하는 것을 돕는다. 일부 예들에서, 디지털 어시스턴트는 또한 콘텍스트 정보를 이용하여 어떻게 출력들을 준비하여 사용자에게 전달할지를 결정한다. 콘텍스트 정보는 콘텍스트 데이터라고 지칭된다.

[0064] 일부 예들에서, 사용자 입력을 동반하는 콘텍스트 정보는 센서 정보, 예를 들어, 조명, 주변 소음, 주변 온도, 주위 환경의 이미지 또는 비디오 등을 포함한다. 일부 예들에서, 콘텍스트 정보는 또한 디바이스의 물리적 상태, 예를 들어, 디바이스 배향, 디바이스 위치, 디바이스 온도, 전력 레벨, 속도, 가속도, 모션 패턴, 셀룰러 신호 강도 등을 포함할 수 있다. 일부 예들에서, DA 서버(106)의 소프트웨어 상태에 관련된 정보, 예를 들어, 휴대용 다기능 디바이스(200)의 실행 중인 프로세스, 설치된 프로그램, 과거 및 현재 네트워크 활동성, 백그라운드 서비스, 에러 로그, 리소스 사용 등이 사용자 입력과 연관된 콘텍스트 정보로서 DA 서버(106)에 제공된다.

[0065] 일부 예들에서, 디지털 어시스턴트 클라이언트 모듈(229)은 DA 서버(106)로부터의 요청들에 응답하여 휴대용 다기능 디바이스(200) 상에 저장된 정보(예컨대, 사용자 데이터(231))를 선택적으로 제공한다. 일부 예들에서, 디지털 어시스턴트 클라이언트 모듈(229)은 또한 DA 서버(106)에 의한 요청에 따라 자연 언어 대화 또는 다른 사용자 인터페이스들을 통해 사용자로부터 추가 입력을 이끌어낸다. 디지털 어시스턴트 클라이언트 모듈(229)은 DA 서버(106)에 추가 입력을 전달하여, 사용자 요청에서 표현되는 사용자의 의도의 의도 추론 및/또는 이행 시에 DA 서버(106)를 돕는다.

[0066] 디지털 어시스턴트의 보다 상세한 설명은 도 7a 내지 도 7c를 참조하여 아래에 기술된다. 디지털 어시스턴트 클라이언트 모듈(229)이 하기에 기술되는 디지털 어시스턴트 모듈(726)의 임의의 개수의 하위모듈들을 포함할 수 있다는 것이 인식되어야 한다.

[0067] 애플리케이션들(236)은 다음의 모듈들(또는 명령어들의 세트들), 또는 이들의 서브세트 또는 수퍼세트(superset)를 포함한다:

[0068] • 연락처 모듈(237)(때때로 주소록 또는 연락처 목록으로 지칭됨);

[0069] • 전화 모듈(238);

[0070] • 화상 회의 모듈(239);

[0071] • 이메일 클라이언트 모듈(240);

[0072] • 인스턴트 메시징(IM) 모듈(241);

[0073] • 운동 지원 모듈(242);

[0074] • 정지 및/또는 비디오 이미지들을 위한 카메라 모듈(243);

[0075] • 이미지 관리 모듈(244);

- [0076]  비디오 재생기 모듈;
- [0077]  음악 재생기 모듈;
- [0078]  브라우저 모듈(247);
- [0079]  캘린더 모듈(248);
- [0080]  일부 예들에서, 날씨 위젯(249-1), 주식 위젯(249-2), 계산기 위젯(249-3), 알람 시계 위젯(249-4), 사전 위젯(249-5), 및 사용자에게 의해 획득되는 다른 위젯들뿐 아니라 사용자-생성 위젯들(249-6) 중 하나 이상을 포함하는 위젯 모듈들(249);
- [0081]  사용자-생성 위젯들(249-6)을 만들기 위한 위젯 생성기 모듈(250);
- [0082]  검색 모듈(251);
- [0083]  비디오 재생기 모듈 및 음악 재생기 모듈을 통합하는 비디오 및 음악 재생기 모듈(252);
- [0084]  메모 모듈(253);
- [0085]  지도 모듈(254); 및/또는
- [0086]  온라인 비디오 모듈(255).
- [0087] 메모리(202)에 저장되는 다른 애플리케이션들(236)의 예들은 다른 워드 프로세싱 애플리케이션들, 다른 이미지 편집 애플리케이션들, 그리기 애플리케이션들, 프레젠테이션 애플리케이션들, JAVA-인에이블형 애플리케이션들, 암호화, 디지털 저작권 관리, 음성 인식, 및 음성 복제를 포함한다.
- [0088] 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232) 및 텍스트 입력 모듈(234)과 함께, 연락처 모듈(237)은 주소록 또는 연락처 목록(예를 들어, 메모리(202) 또는 메모리(470) 내의 연락처 모듈(237)의 애플리케이션 내부 상태(292)에 저장됨)을 관리하는 데 사용되며, 이는 하기를 포함한다: 이름(들)을 주소록에 추가하는 것; 주소록으로부터 이름(들)을 삭제하는 것; 전화번호(들), 이메일 주소(들), 물리적 주소(들) 또는 다른 정보를 이름과 연관시키는 것; 이미지를 이름과 연관시키는 것; 이름들을 분류 및 정렬하는 것; 전화(238), 화상 회의 모듈(239), 이메일(240) 또는 IM(241)에 의한 통신을 개시하고/하거나 용이하게 하기 위해 전화번호들 또는 이메일 주소들을 제공하는 것 등.
- [0089] RF 회로부(208), 오디오 회로부(210), 스피커(211), 마이크로폰(213), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 및 텍스트 입력 모듈(234)과 함께, 전화 모듈(238)은, 전화번호에 대응하는 문자들의 시퀀스를 입력하고, 연락처 모듈(237) 내의 하나 이상의 전화번호에 액세스하며, 입력된 전화번호를 수정하고, 개개의 전화번호를 다이얼링하며, 대화를 하고, 대화가 완료된 때 접속해제하거나 끊는 데 사용된다. 전송된 바와 같이, 무선 통신은 복수의 통신 표준들, 프로토콜들 및 기술들 중 임의의 것을 이용한다.
- [0090] RF 회로부(208), 오디오 회로부(210), 스피커(211), 마이크로폰(213), 터치 스크린(212), 디스플레이 제어기(256), 광 센서(264), 광 센서 제어기(258), 접촉/모션 모듈(230), 그래픽 모듈(232), 텍스트 입력 모듈(234), 연락처 모듈(237) 및 전화 모듈(238)과 함께, 화상 회의 모듈(239)은 사용자 지시들에 따라 사용자와 한 명 이상의 다른 참여자들 사이의 화상 회의를 개시, 시행 및 종료하도록 하는 실행가능 명령어들을 포함한다.
- [0091] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232) 및 텍스트 입력 모듈(234)과 함께, 이메일 클라이언트 모듈(240)은 사용자 지시들에 응답하여 이메일을 작성, 전송, 수신, 및 관리하도록 하는 실행가능 명령어들을 포함한다. 이미지 관리 모듈(244)과 함께, 이메일 클라이언트 모듈(240)은 카메라 모듈(243)로 촬영된 정지 또는 비디오 이미지들을 갖는 이메일을 생성 및 전송하는 것을 매우 용이하게 한다.
- [0092] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232) 및 텍스트 입력 모듈(234)과 함께, 인스턴트 메시징 모듈(241)은, 인스턴트 메시지에 대응하는 문자들의 시퀀스를 입력

하고, 이전에 입력된 문자들을 수정하고, (예를 들어, 전화 기반 인스턴트 메시지들을 위한 단문자 메시지 서비스(Short Message Service, SMS) 또는 멀티미디어 메시지 서비스(Multimedia Message Service, MMS) 프로토콜을 이용하거나, 인터넷 기반 인스턴트 메시지들을 위한 XMPP, SIMPLE 또는 IMPS를 이용하여) 개개의 인스턴트 메시지를 송신하고, 인스턴트 메시지들을 수신하고, 수신된 인스턴트 메시지들을 보도록 하는 실행가능 명령어들을 포함한다. 일부 실시예들에서, 송신 및/또는 수신되는 인스턴트 메시지들은 그래픽, 사진들, 오디오 파일들, 비디오 파일들, 그리고/또는 MMS 및/또는 EMS(Enhanced Messaging Service)에서 지원되는 바와 같은 다른 첨부물들을 포함한다. 본 명세서에서 사용되는 바와 같이, "인스턴트 메시징"은 전화 기반 메시지들(예컨대, SMS 또는 MMS를 이용하여 전송되는 메시지들) 및 인터넷 기반 메시지들(예컨대, XMPP, SIMPLE 또는 IMPS를 이용하여 전송되는 메시지들) 둘 모두를 지칭한다.

- [0093] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 텍스트 입력 모듈(234), GPS 모듈(235), 지도 모듈(254), 및 음악 재생기 모듈과 함께, 운동 지원 모듈(242)은, (예컨대, 시간, 거리, 및/또는 열량 소비 목표와 함께) 운동들을 고안하고; 운동 센서들(스포츠 디바이스들)과 통신하고; 운동 센서 데이터를 수신하고; 운동을 모니터링하는데 사용되는 센서들을 교정하고; 운동을 위한 음악을 선택 및 재생하고; 운동 데이터를 디스플레이, 저장 및 송신하도록 하는 실행가능 명령어들을 포함한다.
- [0094] 터치 스크린(212), 디스플레이 제어기(256), 광 센서(들)(264), 광 센서 제어기(258), 접촉/모션 모듈(230), 그래픽 모듈(232) 및 이미지 관리 모듈(244)과 함께, 카메라 모듈(243)은, 정지 이미지들 또는 비디오(비디오 스트림을 포함함)를 캡처하고 이들을 메모리(202) 내에 저장하거나, 정지 이미지 또는 비디오의 특성을 수정하거나, 메모리(202)로부터 정지 이미지 또는 비디오를 삭제하도록 하는 실행가능 명령어들을 포함한다.
- [0095] 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 텍스트 입력 모듈(234) 및 카메라 모듈(243)과 함께, 이미지 관리 모듈(244)은 정지 및/또는 비디오 이미지들을 배열하거나, 수정(예컨대, 편집)하거나, 또는 그렇지 않으면 조작하고, 라벨링하고, 삭제하고, (예컨대, 디지털 슬라이드 쇼 또는 앨범에) 제시하고, 저장하도록 하는 실행가능 명령어들을 포함한다.
- [0096] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232) 및 텍스트 입력 모듈(234)과 함께, 브라우저 모듈(247)은, 웹 페이지들 또는 이들의 부분들뿐만 아니라 웹 페이지들에 링크된 첨부물들 및 다른 파일들을 검색하고, 그들에 링크하고, 수신하고, 그리고 디스플레이하는 것을 비롯한, 사용자 지시들에 따라 인터넷을 브라우징하도록 하는 실행가능 명령어들을 포함한다.
- [0097] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 텍스트 입력 모듈(234), 이메일 클라이언트 모듈(240), 및 브라우저 모듈(247)과 함께, 캘린더 모듈(248)은 사용자 지시들에 따라 캘린더들 및 캘린더들과 연관된 데이터(예컨대, 캘린더 엔트리들, 할 일 목록들 등)를 생성, 디스플레이, 수정, 및 저장하도록 하는 실행가능 명령어들을 포함한다.
- [0098] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 텍스트 입력 모듈(234) 및 브라우저 모듈(247)과 함께, 위젯 모듈들(249)은 사용자에게 의해 다운로드 및 사용될 수 있거나(예컨대, 날씨 위젯(249-1), 주식 위젯(249-2), 계산기 위젯(249-3), 알람 시계 위젯(249-4) 및 사진 위젯(249-5)), 또는 사용자에게 의해 생성될 수 있는(예컨대, 사용자-생성 위젯(249-6)) 미니-애플리케이션(mini-application)들이다. 일부 실시예들에서, 위젯은 HTML(Hypertext Markup Language) 파일, CSS(Cascading Style Sheets) 파일 및 자바스크립트(JavaScript) 파일을 포함한다. 일부 실시예들에서, 위젯은 XML(Extensible Markup Language) 파일 및 자바스크립트 파일(예를 들어, 야후(Yahoo)! 위젯들)을 포함한다.
- [0099] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 텍스트 입력 모듈(234) 및 브라우저 모듈(247)과 함께, 위젯 생성기 모듈(250)은 사용자에게 의해 위젯들을 생성(예를 들어, 웹 페이지의 사용자-특정 부분을 위젯으로 변경)하는 데 사용된다.
- [0100] 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232) 및 텍스트 입력 모듈(234)과 함께, 검색 모듈(251)은 사용자 지시들에 따라 하나 이상의 검색 기준들(예컨대, 하나 이상의 사용자-특정 검색어들)에 매칭되는 메모리(202) 내의 텍스트, 음악, 사운드, 이미지, 비디오, 및/또는 다른 파일들을 검색하도록 하는 실행가능 명령어들을 포함한다.
- [0101] 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 오디오 회로부(210), 스피커(211), RF 회로부(208) 및 브라우저 모듈(247)과 함께, 비디오 및 음악 재생기 모듈(252)은, 사용자가 MP3 또는 AAC 파일들과 같은 하나 이상의 파일 포맷들로 저장된 기록된 음악 및 다른 사운드 파일들을 다운로드 및

재생할 수 있도록 하는 실행가능 명령어들, 및 비디오들을 (예컨대, 터치 스크린(212) 상에서 또는 외부 포트(224)를 통해 외부의 연결된 디스플레이 상에서) 디스플레이하거나, 제시하거나, 또는 다른 방식으로 재생하도록 하는 실행가능 명령어들을 포함한다. 일부 실시예들에서, 디바이스(200)는 선택적으로 아이팟(애플 인크.의 상표)과 같은 MP3 재생기의 기능을 포함한다.

[0102] 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232) 및 텍스트 입력 모듈(234)과 함께, 메모 모듈(253)은 사용자 지시들에 따라 메모들, 할 일 목록들 등을 생성 및 관리하도록 하는 실행가능 명령어들을 포함한다.

[0103] RF 회로부(208), 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 텍스트 입력 모듈(234), GPS 모듈(235) 및 브라우저 모듈(247)과 함께, 지도 모듈(254)은 사용자 지시들에 따라 지도들 및 지도들과 연관된 데이터(예컨대, 운전 길 안내; 특정 위치에서의 또는 그 인근의 상점들 및 다른 관심 지점들에 관한 데이터; 및 다른 위치-기반 데이터)를 수신하고, 디스플레이하고, 수정하고, 저장하는 데 사용된다.

[0104] 터치 스크린(212), 디스플레이 제어기(256), 접촉/모션 모듈(230), 그래픽 모듈(232), 오디오 회로부(210), 스피커(211), RF 회로부(208), 텍스트 입력 모듈(234), 이메일 클라이언트 모듈(240) 및 브라우저 모듈(247)과 함께, 온라인 비디오 모듈(255)은 사용자가 H.264와 같은 하나 이상의 파일 포맷의 온라인 비디오들에 액세스하고, 그들을 브라우징하고, (예컨대, 스트리밍 및/또는 다운로드에 의해) 수신하고, (예컨대, 터치 스크린 상에서 또는 외부 포트(224)를 통해 외부의 연결된 디스플레이 상에서) 재생하고, 특정한 온라인 비디오로의 링크와 함께 이메일을 전송하고, 그렇지 않으면 관리하게 하는 명령어들을 포함한다. 일부 실시예들에서, 이메일 클라이언트 모듈(240)보다는 오히려 인스턴트 메시징 모듈(241)이 특정 온라인 비디오로의 링크를 전송하는 데 사용된다. 온라인 비디오 애플리케이션에 대한 추가적 설명은, 2007년 6월 20일자로 출원된 미국 가특허 출원 제60/936,562호, "Portable Multifunction Device, Method, and Graphical User Interface for Playing Online Videos" 및 2007년 12월 31일자로 출원된 미국 특허 출원 제11/968,067호, "Portable Multifunction Device, Method, and Graphical User Interface for Playing Online Videos"에서 찾아볼 수 있으며, 이들의 내용은 이로써 그 전체가 참고로 포함된다.

[0105] 앞서 식별된 모듈들 및 애플리케이션들 각각은 전술된 하나 이상의 기능들 및 본 출원에서 기술되는 방법들(예를 들어, 컴퓨터 구현 방법들 및 본 명세서에서 기술되는 다른 정보 처리 방법들)을 수행하기 위한 실행가능 명령어들의 세트에 대응한다. 이들 모듈들(예컨대, 명령어들의 세트들)은 별도의 소프트웨어 프로그램들, 절차들 또는 모듈들로서 구현될 필요는 없으며, 따라서 이들 모듈들의 다양한 서브세트들이 다양한 실시예들에서 조합되거나 달리 재배열될 수 있다. 예를 들어, 비디오 재생기 모듈은 음악 재생기 모듈과 조합되어 단일 모듈(예컨대, 도 2a의 비디오 및 음악 재생기 모듈(252))로 될 수 있다. 일부 실시예들에서, 메모리(202)는 앞서 식별된 모듈들 및 데이터 구조들의 서브세트를 저장한다. 또한, 메모리(202)는 전술되지 않은 추가의 모듈들 및 데이터 구조들을 저장한다.

[0106] 일부 실시예들에서, 디바이스(200)는 디바이스 상의 미리정의된 세트의 기능들의 동작이 터치 스크린 및/또는 터치패드를 통해 전용으로 수행되는 디바이스이다. 터치 스크린 및/또는 터치패드를 디바이스(200)의 동작을 위한 주 입력 제어 디바이스로서 사용함으로써, 디바이스(200) 상의 (푸시 버튼들, 다이얼들 등과 같은) 물리적 입력 제어 디바이스들의 수가 감소된다.

[0107] 터치 스크린 및/또는 터치 패드를 통해 전용으로 수행되는 미리정의된 세트의 기능들은, 선택적으로, 사용자 인터페이스들 간의 내비게이션을 포함한다. 일부 실시예들에서, 터치패드는, 사용자에게 의해 터치될 때, 디바이스(200)를 디바이스(200) 상에 디스플레이되는 임의의 사용자 인터페이스로부터 메인, 홈 또는 루트 메뉴로 내비게이팅한다. 이러한 실시예들에서, "메뉴 버튼"이 터치패드를 이용하여 구현된다. 일부 다른 실시예들에서, 메뉴 버튼은 터치패드 대신에 물리적 푸시 버튼 또는 다른 물리적 입력 제어 디바이스이다.

[0108] 도 2b는 일부 실시예들에 따른, 이벤트 처리를 위한 예시적인 컴포넌트들을 도시하는 블록도이다. 일부 실시예들에서, 메모리(도 2a의 202 또는 도 4의 470)는 (예컨대, 운영 체제(226)에서의) 이벤트 분류기(270) 및 개개의 애플리케이션(236-1)(예컨대, 전술된 애플리케이션들(237 내지 251, 255, 480 내지 490) 중 임의의 것)을 포함한다.

[0109] 이벤트 분류기(270)는 이벤트 정보를 수신하고, 이벤트 정보를 전달할 애플리케이션(236-1), 및 애플리케이션(236-1)의 애플리케이션 뷰(291)를 결정한다. 이벤트 분류기(270)는 이벤트 모니터(271) 및 이벤트 디스패처 모듈(274)을 포함한다. 일부 실시예들에서, 애플리케이션(236-1)은 애플리케이션이 활성화되거나 실행 중일 때

터치 감응형 디스플레이(212) 상에 디스플레이되는 현재 애플리케이션 뷰(들)를 나타내는 애플리케이션 내부 상태(292)를 포함한다. 일부 실시예들에서, 디바이스/글로벌 내부 상태(257)는 이벤트 분류기(270)에 의해 어느 애플리케이션(들)이 현재 활성화인지 결정하는 데 이용되며, 애플리케이션 내부 상태(292)는 이벤트 분류기(270)에 의해 이벤트 정보를 전달할 애플리케이션 뷰들(291)을 결정하는 데 이용된다.

- [0110] 일부 실시예들에서, 애플리케이션 내부 상태(292)는 애플리케이션(236-1)이 실행을 재개할 때 이용될 재개 정보, 애플리케이션(236-1)에 의해 디스플레이되고 있거나 디스플레이될 준비가 된 정보를 나타내는 사용자 인터페이스 상태 정보, 사용자가 애플리케이션(236-1)의 이전 상태 또는 뷰로 되돌아가는 것을 가능하게 하기 위한 상태 큐(queue), 및 사용자에게 의해 취해진 이전 동작들의 재실행(redo)/실행취소(undo) 큐 중 하나 이상과 같은 추가 정보를 포함한다.
- [0111] 이벤트 모니터(271)는 주변기기 인터페이스(218)로부터 이벤트 정보를 수신한다. 이벤트 정보는 서브이벤트(예를 들어, 다중 터치 제스처의 일부로서 터치 감응형 디스플레이(212) 상에서의 사용자 터치)에 대한 정보를 포함한다. 주변기기 인터페이스(218)는 I/O 서브시스템(206) 또는 센서, 예컨대, 근접 센서(266), 가속도계(들)(268), 및/또는 (오디오 회로부(210)를 통한) 마이크로폰(213)으로부터 수신하는 정보를 송신한다. 주변기기 인터페이스(218)가 I/O 서브시스템(206)으로부터 수신하는 정보는 터치 감응형 디스플레이(212) 또는 터치 감응형 표면으로부터의 정보를 포함한다.
- [0112] 일부 실시예들에서, 이벤트 모니터(271)는 요청들을 미리결정된 간격으로 주변기기 인터페이스(218)에 전송한다. 이에 응답하여, 주변기기 인터페이스(218)는 이벤트 정보를 송신한다. 다른 실시예들에서, 주변기기 인터페이스(218)는 중요한 이벤트(예컨대, 미리결정된 잠음 임계치를 초과하는 입력 및/또는 미리결정된 지속기간 초과 동안의 입력을 수신하는 것)가 있을 때에만 이벤트 정보를 송신한다.
- [0113] 일부 실시예들에서, 이벤트 분류기(270)는 또한 히트 뷰(hit view) 결정 모듈(272) 및/또는 활성 이벤트 인식기 결정 모듈(273)을 포함한다.
- [0114] 히트 뷰 결정 모듈(272)은 터치 감응형 디스플레이(212)가 하나 초과된 뷰를 디스플레이할 때 하나 이상의 뷰들 내에서 서브이벤트가 발생한 곳을 결정하기 위한 소프트웨어 절차들을 제공한다. 뷰들은 사용자가 디스플레이 상에서 볼 수 있는 제어부들 및 다른 요소들로 구성된다.
- [0115] 애플리케이션과 연관된 사용자 인터페이스의 다른 양태는 본 명세서에서 때때로 애플리케이션 뷰들 또는 사용자 인터페이스 창(user interface window)들로 지칭되는 한 세트의 뷰들인데, 여기서 정보가 디스플레이되고 터치 기반 제스처가 발생한다. 터치가 검출되는 (개개의 애플리케이션의) 애플리케이션 뷰들은 애플리케이션의 프로그램 또는 뷰 계층구조 내의 프로그램 레벨들에 대응한다. 예를 들어, 터치가 검출되는 최하위 레벨의 뷰는 히트 뷰로 지칭되고, 적절한 입력들로서 인식되는 이벤트들의 세트는 터치 기반 제스처를 시작하는 초기 터치의 히트 뷰에 적어도 부분적으로 기초하여 결정된다.
- [0116] 히트 뷰 결정 모듈(272)은 터치 기반 제스처의 서브이벤트들에 관련된 정보를 수신한다. 애플리케이션이 계층 구조에서 조직화된 다수의 뷰들을 갖는 경우, 히트 뷰 결정 모듈(272)은 히트 뷰를, 서브이벤트를 처리해야 하는 계층구조 내의 최하위 뷰로서 식별한다. 대부분의 상황에서, 히트 뷰는 개시되는 서브이벤트(예컨대, 이벤트 또는 잠재적 이벤트를 형성하는 서브이벤트들의 시퀀스에서의 제1 서브이벤트)가 발생하는 최하위 레벨 뷰이다. 일단 히트 뷰가 히트 뷰 결정 모듈(272)에 의해 식별되면, 히트 뷰는 전형적으로 그것이 히트 뷰로서 식별되게 한 것과 동일한 터치 또는 입력 소스에 관련된 모든 서브이벤트들을 수신한다.
- [0117] 활성 이벤트 인식기 결정 모듈(273)은 뷰 계층구조 내에서 어느 뷰 또는 뷰들이 서브이벤트들의 특정 시퀀스를 수신해야 하는지를 결정한다. 일부 실시예들에서, 활성 이벤트 인식기 결정 모듈(273)은 히트 뷰만이 서브이벤트들의 특정 시퀀스를 수신해야 하는 것으로 결정한다. 다른 실시예들에서, 활성 이벤트 인식기 결정 모듈(273)은 서브이벤트의 물리적 위치를 포함하는 모든 뷰들이 적극 참여 뷰(actively involved view)들인 것으로 결정하고, 그에 따라 모든 적극 참여 뷰들이 서브이벤트들의 특정 시퀀스를 수신해야 하는 것으로 결정한다. 다른 실시예들에서, 터치 서브이벤트들이 전적으로 하나의 특정 뷰와 연관된 영역으로 한정되었다더라도, 계층구조 내의 상위 뷰들은 여전히 적극 참여 뷰들로서 유지될 것이다.
- [0118] 이벤트 디스패처 모듈(274)은 이벤트 정보를 이벤트 인식기(예컨대, 이벤트 인식기(280))에 디스패치한다. 활성 이벤트 인식기 결정 모듈(273)을 포함하는 실시예들에서, 이벤트 디스패처 모듈(274)은 이벤트 정보를 활성 이벤트 인식기 결정 모듈(273)에 의해 결정된 이벤트 인식기에 전달한다. 일부 실시예들에서, 이벤트 디스패처 모듈(274)은 이벤트 큐 내에 이벤트 정보를 저장하는데, 이벤트 정보는 개개의 이벤트 수신기(282)에 의해 인출

된다.

- [0119] 일부 실시예들에서, 운영 체제(226)는 이벤트 분류기(270)를 포함한다. 대안적으로, 애플리케이션(236-1)은 이벤트 분류기(270)를 포함한다. 또 다른 실시예들에서, 이벤트 분류기(270)는 독립형 모듈이거나, 또는 접촉/모션 모듈(230)과 같이 메모리(202)에 저장되는 다른 모듈의 일부이다.
- [0120] 일부 실시예들에서, 애플리케이션(236-1)은 복수의 이벤트 핸들러들(290) 및 하나 이상의 애플리케이션 뷰들(291)을 포함하며, 이들의 각각은 애플리케이션의 사용자 인터페이스의 개개의 뷰 내에 발생하는 터치 이벤트들을 처리하기 위한 명령어들을 포함한다. 애플리케이션(236-1)의 각각의 애플리케이션 뷰(291)는 하나 이상의 이벤트 인식기들(280)을 포함한다. 전형적으로, 개개의 애플리케이션 뷰(291)는 복수의 이벤트 인식기들(280)을 포함한다. 다른 실시예들에서, 이벤트 인식기들(280) 중 하나 이상은 사용자 인터페이스 키트(도시되지 않음) 또는 애플리케이션(236-1)이 방법들 및 다른 속성들을 물려받는 상위 레벨 객체와 같은 별개의 모듈의 일부이다. 일부 실시예들에서, 개개의 이벤트 핸들러(290)는 데이터 업데이터(276), 객체 업데이터(277), GUI 업데이터(278), 및/또는 이벤트 분류기(270)로부터 수신된 이벤트 데이터(279) 중 하나 이상을 포함한다. 이벤트 핸들러(290)는 데이터 업데이터(276), 객체 업데이터(277) 또는 GUI 업데이터(278)를 이용하거나 호출하여 애플리케이션 내부 상태(292)를 업데이트한다. 대안적으로, 애플리케이션 뷰들(291) 중 하나 이상은 하나 이상의 개개의 이벤트 핸들러들(290)을 포함한다. 또한, 일부 실시예들에서, 데이터 업데이터(276), 객체 업데이터(277), 및 GUI 업데이터(278) 중 하나 이상은 개개의 애플리케이션 뷰(291) 내에 포함된다.
- [0121] 개개의 이벤트 인식기(280)는 이벤트 분류기(270)로부터 이벤트 정보(예컨대, 이벤트 데이터(279))를 수신하고 그 이벤트 정보로부터 이벤트를 식별한다. 이벤트 인식기(280)는 이벤트 수신기(282) 및 이벤트 비교기(284)를 포함한다. 일부 실시예들에서, 이벤트 인식기(280)는 또한 적어도 메타데이터(283) 및 이벤트 전달 명령어들(288)(서브이벤트 전달 명령어들을 포함함)의 서브세트를 포함한다.
- [0122] 이벤트 수신기(282)는 이벤트 분류기(270)로부터 이벤트 정보를 수신한다. 이벤트 정보는 서브이벤트, 예를 들어 터치 또는 터치 이동에 관한 정보를 포함한다. 서브이벤트에 따라서, 이벤트 정보는 또한 서브이벤트의 위치와 같은 추가 정보를 포함한다. 서브이벤트가 터치의 모션에 관한 것일 때, 이벤트 정보는 또한 서브이벤트의 속력 및 방향을 포함한다. 일부 실시예들에서, 이벤트들은 하나의 배향으로부터 다른 배향으로(예컨대, 세로 배향으로부터 가로 배향으로, 또는 그 반대로)의 디바이스의 회전을 포함하며, 이벤트 정보는 디바이스의 현재 배향(디바이스 자세로도 지칭됨)에 관한 대응하는 정보를 포함한다.
- [0123] 이벤트 비교기(284)는 이벤트 정보를 미리정의된 이벤트 또는 서브이벤트 정의들과 비교하고, 그 비교에 기초하여, 이벤트 또는 서브이벤트를 결정하거나, 이벤트 또는 서브이벤트의 상태를 결정 또는 업데이트한다. 일부 실시예들에서, 이벤트 비교기(284)는 이벤트 정의들(286)을 포함한다. 이벤트 정의들(286)은 이벤트들(예컨대, 서브이벤트들의 미리정의된 시퀀스들), 예를 들어 이벤트 1(287-1), 이벤트 2(287-2) 등의 정의들을 포함한다. 일부 실시예들에서, 이벤트(287) 내의 서브이벤트들은, 예를 들어, 터치 시작, 터치 종료, 터치 이동, 터치 취소, 및 다중 터치를 포함한다. 일례에서, 이벤트 1(287-1)에 대한 정의는 디스플레이된 객체 상의 더블 탭이다. 더블 탭은, 예를 들어, 미리결정된 페이즈(phase) 동안의 디스플레이된 객체 상의 제1 터치(터치 시작), 미리결정된 페이즈 동안의 제1 리프트오프(터치 종료), 미리결정된 페이즈 동안의 디스플레이된 객체 상의 제2 터치(터치 시작), 및 미리결정된 페이즈 동안의 제2 리프트오프(터치 종료)를 포함한다. 다른 예에서, 이벤트 2(287-2)에 대한 정의는 디스플레이된 객체 상의 드래깅이다. 드래깅은, 예를 들어, 미리결정된 페이즈 동안의 디스플레이된 객체 상의 터치(또는 접촉), 터치 감응형 디스플레이(212)를 가로지르는 터치의 이동, 및 터치의 리프트오프(터치 종료)를 포함한다. 일부 실시예들에서, 이벤트는 또한 하나 이상의 연관된 이벤트 핸들러들(290)에 대한 정보를 포함한다.
- [0124] 일부 실시예들에서, 이벤트 정의(287)는 개개의 사용자 인터페이스 객체에 대한 이벤트의 정의를 포함한다. 일부 실시예들에서, 이벤트 비교기(284)는 어느 사용자 인터페이스 객체가 서브이벤트와 연관되어 있는지 결정하도록 히트 테스트(hit test)를 수행한다. 예를 들어, 3개의 사용자 인터페이스 객체들이 터치 감응형 디스플레이(212) 상에 디스플레이된 애플리케이션 뷰에서, 터치 감응형 디스플레이(212) 상에서 터치가 검출되는 경우, 이벤트 비교기(284)는 3개의 사용자 인터페이스 객체들 중 어느 것이 터치(서브이벤트)와 연관되는지를 결정하도록 히트 테스트를 수행한다. 각각의 디스플레이된 객체가 개개의 이벤트 핸들러(290)와 연관되는 경우, 이벤트 비교기는 어느 이벤트 핸들러(290)가 활성화되어야 하는지 결정하는 데 히트 테스트의 결과를 이용한다. 예를 들어, 이벤트 비교기(284)는 히트 테스트를 트리거하는 객체 및 서브이벤트와 연관된 이벤트 핸들러를 선택한다.

- [0125] 일부 실시예들에서, 개개의 이벤트(287)에 대한 정의는 또한 서브이벤트들의 시퀀스가 이벤트 인식기의 이벤트 유형에 대응하는지 대응하지 않는지 여부가 결정된 후까지 이벤트 정보의 전달을 지연시키는 지연된 동작들을 포함한다.
- [0126] 개개의 이벤트 인식기(280)가 일련의 서브이벤트들이 이벤트 정의들(286) 내의 이벤트들 중 어떠한 것이라도 매칭되지 않는 것으로 결정하면, 개개의 이벤트 인식기(280)는 이벤트 불가능, 이벤트 실패, 또는 이벤트 종료 상태에 진입하고, 그 후 개개의 이벤트 인식기는 터치 기반 제스처의 후속적인 서브이벤트들을 무시한다. 이러한 상황에서, 만일 있다면, 히트 뷰에 대해 활성 상태로 유지되는 다른 이벤트 인식기들이 진행 중인 터치 기반 제스처의 서브이벤트들을 계속해서 추적 및 처리한다.
- [0127] 일부 실시예들에서, 개개의 이벤트 인식기(280)는 이벤트 전달 시스템이 어떻게 적극 참여 이벤트 인식기들에 대한 서브이벤트 전달을 수행해야 하는지를 나타내는 구성가능한 속성들, 플래그(flag)들, 및/또는 목록들을 갖는 메타데이터(283)를 포함한다. 일부 실시예들에서, 메타데이터(283)는 이벤트 인식기들이 어떻게 서로 상호작용하는지, 또는 상호작용 가능하게 되는지를 나타내는 구성가능한 속성들, 플래그들, 및/또는 목록들을 포함한다. 일부 실시예들에서, 메타데이터(283)는, 서브이벤트들이 뷰 또는 프로그램 계층구조에서의 다양한 레벨들에 전달되는지 여부를 나타내는 구성가능한 속성들, 플래그들, 및/또는 목록들을 포함한다.
- [0128] 일부 실시예들에서, 개개의 이벤트 인식기(280)는 이벤트의 하나 이상의 특정 서브이벤트들이 인식될 때 이벤트와 연관된 이벤트 핸들러(290)를 활성화시킨다. 일부 실시예들에서, 개개의 이벤트 인식기(280)는 이벤트와 연관된 이벤트 정보를 이벤트 핸들러(290)에 전달한다. 이벤트 핸들러(290)를 활성화시키는 것은 개개의 히트 뷰에 서브이벤트들을 전송(및 지연 전송)하는 것과는 별개이다. 일부 실시예들에서, 이벤트 인식기(280)는 인식된 이벤트와 연관된 플래그를 보내고, 그 플래그와 연관된 이벤트 핸들러(290)는 그 플래그를 캐치하고 미리 정의된 프로세스를 수행한다.
- [0129] 일부 실시예들에서, 이벤트 전달 명령어들(288)은 이벤트 핸들러를 활성화시키지 않으면서 서브이벤트에 관한 이벤트 정보를 전달하는 서브이벤트 전달 명령어들을 포함한다. 대신에, 서브이벤트 전달 명령어들은 일련의 서브이벤트들과 연관된 이벤트 핸들러들에 또는 적극 참여 뷰들에 이벤트 정보를 전달한다. 일련의 서브이벤트들 또는 적극 참여 뷰들과 연관된 이벤트 핸들러들은 이벤트 정보를 수신하고 미리결정된 프로세스를 수행한다.
- [0130] 일부 실시예들에서, 데이터 업데이터(276)는 애플리케이션(236-1)에서 이용되는 데이터를 생성 및 업데이트한다. 예를 들어, 데이터 업데이터(276)는 연락처 모듈(237)에서 이용되는 전화번호를 업데이트하거나, 비디오 재생기 모듈에서 이용되는 비디오 파일을 저장한다. 일부 실시예들에서, 객체 업데이터(277)는 애플리케이션(236-1)에서 이용되는 객체들을 생성 및 업데이트한다. 예를 들어, 객체 업데이터(277)는 새로운 사용자 인터페이스 객체를 생성하거나, 또는 사용자 인터페이스 객체의 위치를 업데이트한다. GUI 업데이터(278)는 GUI를 업데이트한다. 예를 들어, GUI 업데이터(278)는 터치 감응형 디스플레이 상의 디스플레이를 위해 디스플레이 정보를 준비하고 이를 그래픽 모듈(232)에 전송한다.
- [0131] 일부 실시예들에서, 이벤트 핸들러(들)(290)는 데이터 업데이터(276), 객체 업데이터(277), 및 GUI 업데이터(278)를 포함하거나 이들에 액세스한다. 일부 실시예들에서, 데이터 업데이터(276), 객체 업데이터(277), 및 GUI 업데이터(278)는 개개의 애플리케이션(236-1) 또는 애플리케이션 뷰(291)의 단일 모듈 내에 포함된다. 다른 실시예들에서, 이들은 2개 이상의 소프트웨어 모듈들 내에 포함된다.
- [0132] 터치 감응형 디스플레이 상의 사용자 터치들의 이벤트 처리에 관하여 전술한 논의는 또한 입력 디바이스들을 갖는 다기능 디바이스들(200)을 동작시키기 위한 다른 형태들의 사용자 입력들에도 적용되지만, 그 모두가 터치 스크린들 상에서 개시되는 것이 아니라는 것을 이해해야 한다. 예를 들어, 단일 또는 다수의 키보드 누르기 또는 유지(hold)와 선택적으로 조화된 마우스 이동 및 마우스 버튼 누르기; 터치패드 상에서의, 탭, 드래그, 스크롤 등과 같은 접촉 이동들; 펜 스타일러스 입력들; 디바이스의 이동; 구두 명령어들; 검출된 눈 이동들; 생체측정 입력들; 및/또는 이들의 임의의 조합은, 인식될 이벤트를 정의하는 서브이벤트들에 대응하는 입력들로서 선택적으로 이용된다.
- [0133] 도 3은 일부 실시예들에 따른, 터치 스크린(212)을 갖는 휴대용 다기능 디바이스(200)를 도시한다. 터치 스크린은, 선택적으로, 사용자 인터페이스(UI)(300) 내에서 하나 이상의 그래픽들을 디스플레이한다. 이러한 실시예는 물론 하기에 기술되는 다른 실시예들에서, 사용자는, 예를 들어, 하나 이상의 손가락들(302)(도면에서 축척대로 도시되지 않음) 또는 하나 이상의 스타일러스들(303)(도면에서 축척대로 도시되지 않음)을 이용하여 그래픽들 상에 제스처를 행함으로써 그래픽들 중 하나 이상을 선택할 수 있다. 일부 실시예들에서, 하나 이상의

그래픽들의 선택은 사용자가 하나 이상의 그래픽들과의 접촉을 중단할 때 발생한다. 일부 실시예들에서, 제스처는 선택적으로 디바이스(200)와 접촉한 손가락의 하나 이상의 탭들, (좌측에서 우측으로의, 우측에서 좌측으로의, 상측으로의 그리고/또는 하측으로의) 하나 이상의 스와이프들, 및/또는 (우측에서 좌측으로의, 좌측에서 우측으로의, 상측으로의 그리고/또는 하측으로의) 롤링을 포함한다. 일부 구현예들 또는 상황들에서, 그래픽과 부주의하여 접촉되면 그 그래픽은 선택되지 않는다. 예를 들어, 선택에 대응하는 제스처가 탭일 때, 애플리케이션 아이콘 위를 스위프(sweep)하는 스와이프 제스처는 선택적으로, 대응하는 애플리케이션을 선택하지 않는다.

[0134] 디바이스(200)는 또한 "홈" 또는 메뉴 버튼(304)과 같은 하나 이상의 물리적 버튼들을 포함한다. 이전에 기술된 바와 같이, 메뉴 버튼(304)은 디바이스(200) 상에서 실행되는 애플리케이션들의 세트 내의 임의의 애플리케이션(236)으로 내비게이팅하기 위해 사용된다. 대안적으로, 일부 실시예들에서, 메뉴 버튼은 터치 스크린(212) 상에 디스플레이된 GUI에서 소프트 키로서 구현된다.

[0135] 일 실시예에서, 디바이스(200)는 터치 스크린(212), 메뉴 버튼(304), 디바이스의 전원을 온/오프시키고 디바이스를 잠그기 위한 푸시 버튼(306), 음량 조절 버튼(들)(308), 가입자 식별 모듈(SIM) 카드 슬롯(310), 헤드셋 잭(312), 및 도킹/충전 외부 포트(224)를 포함한다. 푸시 버튼(306)은, 선택적으로, 버튼을 누르고 버튼을 미리정의된 시간 간격 동안 누른 상태로 유지함으로써 디바이스의 전원을 온/오프시키고/시키거나; 버튼을 누르고 미리정의된 시간 간격이 경과하기 전에 버튼을 누름해제함으로써 디바이스를 잠그고/잠그거나; 디바이스를 잠금 해제하거나 잠금해제 프로세스를 개시하는 데 사용된다. 대안적인 실시예에서, 디바이스(200)는 또한 마이크로폰(213)을 통해 일부 기능들의 활성화 또는 비활성화를 위한 구두 입력을 수용한다. 디바이스(200)는 또한, 선택적으로, 터치 스크린(212) 상에서의 접촉들의 세기를 검출하기 위한 하나 이상의 접촉 세기 센서들(265) 및/또는 디바이스(200)의 사용자를 위해 촉각적 출력들을 생성하기 위한 하나 이상의 촉각적 출력 생성기들(267)을 포함한다.

[0136] 도 4는 일부 실시예들에 따른, 디스플레이 및 터치 감응형 표면을 갖는 예시적인 다기능 디바이스의 블록도이다. 디바이스(400)가 휴대용일 필요는 없다. 일부 실시예들에서, 디바이스(400)는, 랩톱 컴퓨터, 데스크톱 컴퓨터, 태블릿 컴퓨터, 멀티미디어 재생기 디바이스, 내비게이션 디바이스, (어린이 학습 장난감과 같은) 교육용 디바이스, 게이밍 시스템, 또는 제어 디바이스(예컨대, 가정용 또는 산업용 제어기)이다. 디바이스(400)는 전형적으로 하나 이상의 처리 유닛(CPU)들(410), 하나 이상의 네트워크 또는 다른 통신 인터페이스들(460), 메모리(470), 및 이들 컴포넌트를 상호연결하기 위한 하나 이상의 통신 버스들(420)을 포함한다. 통신 버스들(420)은 선택적으로 시스템 컴포넌트들을 상호연결하고 이들 사이의 통신을 제어하는 회로부(때때로 칩셋이라고 지칭됨)를 포함한다. 디바이스(400)는 전형적으로 터치 스크린 디스플레이인 디스플레이(440)를 포함하는 입/출력(I/O) 인터페이스(430)를 포함한다. I/O 인터페이스(430)는 또한, 선택적으로, 키보드 및/또는 마우스(또는 다른 포인팅 디바이스)(450) 및 터치패드(455), 디바이스(400) 상에 촉각적 출력들을 생성하기 위한 촉각적 출력 생성기(457)(예컨대, 도 2a를 참조하여 기술된 촉각적 출력 생성기(들)(267)와 유사함), 센서들(459)(예컨대, 도 2a를 참조하여 기술된 접촉 세기 센서(들)(265)와 유사한 광 센서, 가속도 센서, 근접 센서, 터치 감응형 센서, 및/또는 접촉 세기 센서)을 포함한다. 메모리(470)는 DRAM, SRAM, DDR RAM 또는 다른 랜덤 액세스 솔리드 스테이트 메모리 디바이스들과 같은 고속 랜덤 액세스 메모리를 포함하며; 선택적으로 하나 이상의 자기 디스크 저장 디바이스, 광 디스크 저장 디바이스, 플래시 메모리 디바이스, 또는 다른 비휘발성 솔리드 스테이트 저장 디바이스와 같은 비휘발성 메모리를 포함한다. 메모리(470)는 선택적으로 CPU(들)(410)로부터 원격으로 위치된 하나 이상의 저장 디바이스들을 포함한다. 일부 실시예들에서, 메모리(470)는 휴대용 다기능 디바이스(200)(도 2a)의 메모리(202)에 저장된 프로그램들, 모듈들, 및 데이터 구조들과 유사한 프로그램들, 모듈들, 및 데이터 구조들, 또는 이들의 서브세트를 저장한다. 또한, 메모리(470)는, 선택적으로, 휴대용 다기능 디바이스(200)의 메모리(202) 내에 존재하지 않는 추가의 프로그램들, 모듈들 및 데이터 구조들을 저장한다. 예를 들어, 디바이스(400)의 메모리(470)는, 선택적으로, 그리기 모듈(480), 프레젠테이션 모듈(482), 워드 프로세싱 모듈(484), 웹사이트 제작 모듈(486), 디스크 저작 모듈(488), 및/또는 스프레드시트 모듈(490)을 저장하는 반면, 휴대용 다기능 디바이스(200)(도 2a)의 메모리(202)는, 선택적으로, 이들 모듈들을 저장하지 않는다.

[0137] 도 4에서의 앞서 식별된 요소들 각각은, 일부 예들에서, 이전에 언급된 메모리 디바이스들 중 하나 이상에 저장된다. 앞서 식별된 모듈들 각각은 전송된 기능을 수행하기 위한 명령어들의 세트에 대응한다. 앞서 식별된 모듈들 또는 프로그램들(예컨대, 명령어들의 세트들)은 별개의 소프트웨어 프로그램들, 절차들 또는 모듈들로서 구현될 필요는 없으며, 따라서 이러한 모듈들의 다양한 서브세트들이 다양한 실시예들에서 조합되거나 다른 방식으로 재배열된다. 일부 실시예들에서, 메모리(470)는 앞서 식별된 모듈들 및 데이터 구조들의 서브세트를 저

장한다. 또한, 메모리(470)는 전술되지 않은 추가의 모듈들 및 데이터 구조들을 저장한다.

- [0138] 이제, 예를 들어 휴대용 다기능 디바이스(200) 상에서 구현될 수 있는 사용자 인터페이스들의 실시예들에 주목한다.
- [0139] 도 5a는 일부 실시예들에 따른, 휴대용 다기능 디바이스(200) 상의 애플리케이션들의 메뉴에 대한 예시적인 사용자 인터페이스를 도시한다. 유사한 사용자 인터페이스들이 디바이스(400) 상에서 구현된다. 일부 실시예들에서, 사용자 인터페이스(500)는 하기의 요소들, 또는 그들의 서브세트나 슈퍼세트를 포함한다:
- [0140] 셀룰러 및 Wi-Fi 신호들과 같은 무선 통신(들)에 대한 신호 강도 표시자(들)(502);
- [0141]  시간(504);
- [0142]  블루투스 표시자(505);
- [0143]  배터리 상태 표시자(506);
- [0144]  다음과 같은, 빈번하게 사용되는 애플리케이션들에 대한 아이콘들을 갖는 트레이(508):
- [0145] o 부재 중 전화들 또는 음성메일 메시지들의 개수의 표시자(514)를 선택적으로 포함하는 "전화"라고 라벨링된 전화 모듈(238)에 대한 아이콘(516);
- [0146] o 읽지 않은 이메일들의 개수의 표시자(510)를 선택적으로 포함하는 "메일"이라고 라벨링된 이메일 클라이언트 모듈(240)에 대한 아이콘(518);
- [0147] o "브라우저"라고 라벨링된 브라우저 모듈(247)에 대한 아이콘(520); 및
- [0148] o 아이팟(애플 인크.의 상표) 모듈(252)로도 지칭되는, "아이팟"이라고 라벨링된 비디오 및 음악 재생기 모듈(252)에 대한 아이콘(522); 및
- [0149]  다음과 같은, 다른 애플리케이션들에 대한 아이콘들:
- [0150] o "메시지"라고 라벨링된 IM 모듈(241)에 대한 아이콘(524);
- [0151] o "캘린더"라고 라벨링된 캘린더 모듈(248)에 대한 아이콘(526);
- [0152] o "사진"이라고 라벨링된 이미지 관리 모듈(244)에 대한 아이콘(528);
- [0153] o "카메라"라고 라벨링된 카메라 모듈(243)에 대한 아이콘(530);
- [0154] o "온라인 비디오"라고 라벨링된 온라인 비디오 모듈(255)에 대한 아이콘(532);
- [0155] o "주식"이라고 라벨링된 주식 위젯(249-2)에 대한 아이콘(534);
- [0156] o "지도"라고 라벨링된 지도 모듈(254)에 대한 아이콘(536);
- [0157] o "날씨"라고 라벨링된 날씨 위젯(249-1)에 대한 아이콘(538);
- [0158] o "시계"라고 라벨링된 알람 시계 위젯(249-4)에 대한 아이콘(540);
- [0159] o "운동 지원"이라고 라벨링된 운동 지원 모듈(242)에 대한 아이콘(542);
- [0160] o "메모"라고 라벨링된 메모 모듈(253)에 대한 아이콘(544); 및
- [0161] o 디바이스(200) 및 그의 다양한 애플리케이션들(236)에 대한 설정에의 액세스를 제공하는, "설정"이라고 라벨링된, 설정 애플리케이션 또는 모듈에 대한 아이콘(546).
- [0162] 도 5a에 도시된 아이콘 라벨들은 단지 예시적인 것임에 유의해야 한다. 예를 들어, 비디오 및 음악 재생기 모듈(252)에 대한 아이콘(522)은 선택적으로 "음악" 또는 "음악 재생기"라고 라벨링된다. 기타 라벨들이 선택적으로 다양한 애플리케이션 아이콘들에 대해 사용된다. 일부 실시예들에서, 개개의 애플리케이션 아이콘에 대한 라벨은 개개의 애플리케이션 아이콘에 대응하는 애플리케이션의 이름을 포함한다. 일부 실시예들에서, 특정 애플리케이션 아이콘에 대한 라벨은 특정 애플리케이션 아이콘에 대응하는 애플리케이션의 이름과는 구별된다.
- [0163] 도 5b는 디스플레이(550)(예컨대, 터치 스크린 디스플레이(212))와는 별개인 터치 감응형 표면(551)(예컨대, 도

4의 태블릿 또는 터치패드(455)을 갖는 디바이스(예컨대, 도 4의 디바이스(400)) 상의 예시적인 사용자 인터페이스를 도시한다. 디바이스(400)는 또한, 선택적으로, 터치 감응형 표면(551) 상에서의 접촉들의 세기를 검출하기 위한 하나 이상의 접촉 세기 센서들(예컨대, 센서들(459) 중 하나 이상) 및/또는 디바이스(400)의 사용자에 대한 촉각적 출력들을 생성하기 위한 하나 이상의 촉각적 출력 생성기들(457)을 포함한다.

[0164] 후속하는 일부 예들이 (터치 감응형 표면과 디스플레이가 조합된) 터치 스크린 디스플레이(212) 상의 입력들을 참조하여 제공될 것이지만, 일부 실시예들에서, 디바이스는 도 5b에 도시된 바와 같이 디스플레이와 별개인 터치 감응형 표면 상의 입력들을 검출한다. 일부 실시예들에서, 터치 감응형 표면(예컨대, 도 5b의 551)은 디스플레이(예컨대, 550) 상의 주축(예컨대, 도 5b의 553)에 대응하는 주축(예컨대, 도 5b의 552)을 갖는다. 이 실시예들에 따르면, 디바이스는 디스플레이 상의 개개의 위치들에 대응하는 위치들(예컨대, 도 5b에서, 560은 568에 대응하고, 562는 570에 대응함)에서 터치 감응형 표면(551)과의 접촉들(예컨대, 도 5b의 560, 562)을 검출한다. 이러한 방식으로, 터치 감응형 표면(예컨대, 도 5b의 551) 상에서 디바이스에 의해 검출된 사용자 입력들(예컨대, 접촉들(560, 562) 및 그 이동들)은 터치 감응형 표면이 디스플레이와 별개일 때 디바이스에 의해 다기능 디바이스의 디스플레이(예컨대, 도 5b의 550) 상의 사용자 인터페이스를 조작하는 데 사용된다. 유사한 방법들이, 선택적으로, 본 명세서에 기술된 다른 사용자 인터페이스들에 이용된다는 것이 이해되어야 한다.

[0165] 추가적으로, 하기의 예들이 손가락 입력들(예컨대, 손가락 접촉들, 손가락 탭 제스처들, 손가락 스와이프 제스처들)을 주로 참조하여 주어지는 반면, 일부 실시예들에서, 손가락 입력들 중 하나 이상은 다른 입력 디바이스로부터의 입력(예컨대, 마우스 기반 입력 또는 스타일러스 입력)으로 대체된다는 것이 이해되어야 한다. 예컨대, 스와이프 제스처는, 선택적으로, 마우스 클릭(예컨대, 접촉 대신) 및 뒤이은 스와이프의 경로를 따른 커서의 이동(예컨대, 접촉의 이동 대신)으로 대체된다. 다른 예로서, (예컨대, 접촉의 검출에 이어 접촉을 검출하는 것을 중지하는 것 대신에) 커서가 탭 제스처의 위치 위에 위치되어 있는 동안에 탭 제스처가 선택적으로 마우스 클릭으로 대체된다. 유사하게, 다수의 사용자 입력들이 동시에 검출되는 경우, 다수의 컴퓨터 마우스들이 선택적으로 동시에 사용되거나, 또는 마우스와 손가락 접촉들이 선택적으로 동시에 사용된다는 것이 이해되어야 한다.

[0166] 도 6a는 예시적인 개인용 전자 디바이스(600)를 도시한다. 디바이스(600)는 몸체(602)를 포함한다. 일부 실시예들에서, 디바이스(600)는 디바이스들(200, 400)(예컨대, 도 2a 내지 도 4)에 대하여 기술된 특징부들의 일부 또는 전부를 포함한다. 일부 실시예들에서, 디바이스(600)는 터치 감응형 디스플레이 스크린(604)(이하, 터치 스크린(604))을 갖는다. 터치 스크린(604)에 대해 대안적으로 또는 부가적으로, 디바이스(600)는 디스플레이 및 터치 감응형 표면을 갖는다. 디바이스들(200, 400)과 마찬가지로, 일부 실시예들에서, 터치 스크린(604)(또는 터치 감응형 표면)은 가해지는 접촉들(예컨대, 터치들)의 세기를 검출하기 위한 하나 이상의 세기 센서들을 갖는다. 터치 스크린(604)(또는 터치 감응형 표면)의 하나 이상의 세기 센서들은 터치들의 세기를 표현하는 출력 데이터를 제공한다. 디바이스(600)의 사용자 인터페이스는 터치들의 세기에 기초하여 터치들에 응답하고, 이는 상이한 세기들의 터치들이 디바이스(600) 상의 상이한 사용자 인터페이스 동작들을 호출할 수 있다는 것을 의미한다.

[0167] 터치 세기를 검출 및 처리하기 위한 기법들은, 예를 들어, 관련 출원들: 2013년 5월 8일자로 출원되고 발명의 명칭이 "Device, Method, and Graphical User Interface for Displaying User Interface Objects Corresponding to an Application"인 국제 특허 출원 제PCT/US2013/040061호, 및 2013년 11월 11일자로 출원되고 발명의 명칭이 "Device, Method, and Graphical User Interface for Transitioning Between Touch Input to Display Output Relationships"인 국제 특허 출원 제PCT/US2013/069483호에서 발견되며, 이들 각각은 이로써 그 전체가 참고로 포함된다.

[0168] 일부 실시예들에서, 디바이스(600)는 하나 이상의 입력 메커니즘들(606, 608)을 갖는다. 입력 메커니즘들(606, 608)(포함되어 있는 경우)은 물리적이다. 물리적 입력 메커니즘들의 예들은 푸시 버튼들 및 회전가능 메커니즘들을 포함한다. 일부 실시예들에서, 디바이스(600)는 하나 이상의 부착 메커니즘들을 갖는다. 이러한 부착 메커니즘들(포함되어 있는 경우)은 디바이스(600)가, 예를 들어, 모자, 안경, 귀걸이, 목걸이, 셔츠, 재킷, 팔찌, 시계줄, 쇠줄(chain), 바지, 벨트, 신발, 지갑, 배낭 등에 부착될 수 있게 한다. 이 부착 메커니즘들은 디바이스(600)가 사용자에 의해 착용되도록 한다.

[0169] 도 6b는 예시적인 개인용 전자 디바이스(600)를 도시한다. 일부 실시예들에서, 디바이스(600)는 도 2a, 도 2b, 및 도 4에 대하여 기술된 컴포넌트들의 일부 또는 전부를 포함한다. 디바이스(600)는 I/O 섹션(614)을 하나 이상의 컴퓨터 프로세서들(616) 및 메모리(618)와 동작가능하게 결합하는 버스(612)를 갖는다. I/O 섹션(614)은

터치 감응형 컴포넌트(622), 및 선택적으로, 터치 세기 감응형 컴포넌트(624)를 가질 수 있는 디스플레이(604)에 연결된다. 추가로, I/O 섹션(614)은, Wi-Fi, 블루투스, 근거리 통신(NFC), 셀룰러 및/또는 다른 무선 통신 기술들을 사용하여, 애플리케이션 및 운영 체제 데이터를 수신하기 위해 통신 유닛(630)과 연결된다. 디바이스(600)는 입력 메커니즘들(606 및/또는 608)을 포함한다. 입력 메커니즘(606)은, 예를 들어, 회전가능 입력 디바이스 또는 누름가능 및 회전가능 입력 디바이스이다. 입력 메커니즘(608)은, 일부 예들에서, 버튼이다.

[0170] 입력 메커니즘(608)은, 일부 예들에서, 마이크로폰이다. 개인용 전자 디바이스(600)는, 예를 들어, GPS 센서(632), 가속도계(634), 방향 센서(640)(예컨대, 나침반), 자이로스코프(636), 모션 센서(638), 및/또는 이들의 조합과 같은, 다양한 센서들을 포함하고, 이들 모두는 I/O 섹션(614)에 동작가능하게 연결된다.

[0171] 개인용 전자 디바이스(600)의 메모리(618)는, 예를 들어, 하나 이상의 컴퓨터 프로세서들(616)에 의해 실행될 때, 컴퓨터 프로세서들로 하여금 후술되는 기술들 및 프로세스들을 수행하게 하는, 컴퓨터 실행가능 명령어들을 저장하기 위한 비밀시적 컴퓨터 판독가능 저장 매체이다. 컴퓨터 실행가능 명령어들은, 예를 들어, 명령어 실행 시스템, 장치, 또는 디바이스, 예컨대, 컴퓨터 기반 시스템, 프로세서-포함(processor-containing) 시스템, 또는 명령어 실행 시스템, 장치, 또는 디바이스로부터 명령어들을 폐지하여 명령어들을 실행할 수 있는 다른 시스템에 의해 또는 그와 관련하여 사용하기 위해 임의의 비밀시적 컴퓨터 판독가능 저장 매체 내에 또한 저장되고/되거나 전송된다. 개인용 전자 디바이스(600)는 도 6b의 컴포넌트들 및 구성으로 한정되지 않고, 다수의 구성들에서 다른 또는 추가적인 컴포넌트들을 포함할 수 있다.

[0172] 본 명세서에서 사용되는 바와 같이, "어포던스"라는 용어는 디바이스들(200, 400, 및/또는 600)(도 2a, 도 4, 및 도 6a와 도 6b)의 디스플레이 스크린 상에 디스플레이되는 사용자 상호작용(user-interactive) 그래픽 사용자 인터페이스 객체를 지칭한다. 예를 들어, 이미지(예를 들어, 아이콘), 버튼, 및 텍스트(예를 들어, 하이퍼링크)는 각각 어포던스를 구성한다.

[0173] 본 명세서에 사용되는 바와 같이, "포커스 선택자(focus selector)"라는 용어는 사용자와 상호작용하고 있는 사용자 인터페이스의 현재 부분을 나타내는 입력 요소를 지칭한다. 커서 또는 다른 위치 마커(location marker)를 포함하는 일부 구현예들에서, 커서가 특정 사용자 인터페이스 요소(예컨대, 버튼, 윈도우, 슬라이더 또는 다른 사용자 인터페이스 요소) 위에 있는 동안 터치 감응형 표면(예컨대, 도 4의 터치패드(455) 또는 도 5b의 터치 감응형 표면(551)) 상에서 입력(예컨대, 누르기 입력)이 검출될 때, 특정 사용자 인터페이스 요소가 검출된 입력에 따라 조정되도록, 커서가 "포커스 선택자"로서 기능한다. 터치 스크린 디스플레이 상의 사용자 인터페이스 요소들과의 직접적인 상호작용을 가능하게 하는 터치 스크린 디스플레이(예컨대, 도 2a의 터치 감응형 디스플레이 시스템(212) 또는 도 5a의 터치 스크린(212))을 포함하는 일부 구현예들에서, 입력(예컨대, 접촉에 의한 누르기 입력)이 특정 사용자 인터페이스 요소(예컨대, 버튼, 윈도우, 슬라이더 또는 다른 사용자 인터페이스 요소)의 위치에 있는 터치 스크린 디스플레이 상에서 검출될 때, 특정 사용자 인터페이스 요소가 검출된 입력에 따라 조정되도록, 터치 스크린 상의 검출된 접촉이 "포커스 선택자"로서 기능한다. 일부 구현예들에서, (예를 들어 포커스를 하나의 버튼으로부터 다른 버튼으로 움직이도록 탭 키 또는 화살표 키를 사용함으로써) 터치 스크린 디스플레이 상의 대응하는 커서의 이동 또는 접촉의 이동 없이 포커스가 사용자 인터페이스의 하나의 영역으로부터 사용자 인터페이스의 다른 영역으로 이동되며; 이러한 구현예들에서, 포커스 선택자는 사용자 인터페이스의 상이한 영역들 사이에서의 포커스의 이동에 따라 움직인다. 포커스 선택자가 갖는 특정 형태와 무관하게, 포커스 선택자는 일반적으로 (예를 들어, 사용자가 상호작용하려고 하는 사용자 인터페이스의 요소를 디바이스에 표시하는 것에 의해) 사용자 인터페이스와의 사용자의 의도된 상호작용을 전달하기 위해 사용자에게 의해 제어되는 사용자 인터페이스 요소(또는 터치 스크린 디스플레이 상의 접촉)이다. 예를 들어, 터치 감응형 표면(예를 들어, 터치패드 또는 터치 스크린) 상에서 누르기 입력이 검출되는 동안 각자의 버튼 위의 포커스 선택자(예를 들어, 커서, 접촉 또는 선택 박스)의 위치는 (디바이스의 디스플레이 상에 보여지는 다른 사용자 인터페이스 요소들과 달리) 사용자가 각자의 버튼을 활성화시키려고 하고 있다는 것을 나타낼 것이다.

[0174] 명세서 및 청구범위에서 사용되는 바와 같이, 접촉의 "특성 세기"라는 용어는 접촉의 하나 이상의 세기들에 기초한 접촉의 특성을 지칭한다. 일부 실시예들에서, 특성 세기는 다수의 세기 샘플들에 기초한다. 특성 세기는, 선택적으로, 미리정의된 수의 세기 샘플들, 또는 (예컨대, 접촉을 검출한 이후에, 접촉의 리프트오프를 검출하기 이전에, 접촉의 이동의 시작을 검출하기 이전 또는 이후에, 접촉의 종료를 검출하기 이전에, 접촉의 세기의 증가를 검출하기 이전 또는 이후에, 그리고/또는 접촉의 세기의 감소를 검출하기 이전 또는 이후에) 미리정의된 이벤트에 대해 미리결정된 기간(예컨대, 0.05, 0.1, 0.2, 0.5, 1, 2, 5, 10초) 동안 수집된 세기 샘플들의 세트에 기초한다. 접촉의 특성 세기는, 선택적으로, 접촉의 세기들의 최대 값, 접촉의 세기들의 중간 값(mean value), 접촉의 세기들의 평균값(average value), 접촉의 세기들의 상위 10 백분위 값(top 10

percentile value), 접촉의 세기들의 최대 값의 절반의 값, 접촉의 세기들의 최대 값의 90 퍼센트의 값 등 중 하나 이상에 기초한다. 일부 실시예들에서, 접촉의 지속기간은 (예컨대, 특성 세기가 시간의 경과에 따른 접촉의 세기의 평균일 때) 특성 세기를 결정하는데 사용된다. 일부 실시예들에서, 동작이 사용자에게 의해 수행되었는지 여부를 결정하기 위해, 특성 세기가 하나 이상의 세기 임계치들의 세트와 비교된다. 예를 들어, 하나 이상의 세기 임계치들의 세트는 제1 세기 임계치 및 제2 세기 임계치를 포함한다. 이 예에서, 제1 임계치를 초과하지 않는 특성 세기를 갖는 접촉의 결과, 제1 동작이 행해지고, 제1 세기 임계치를 초과하지만 제2 세기 임계치를 초과하지 않는 특성 세기를 갖는 접촉의 결과, 제2 동작이 행해지며, 제2 임계치 초과 특성 세기를 갖는 접촉의 결과, 제3 동작이 행해진다. 일부 실시예들에서, 특성 세기와 하나 이상의 임계치들 간의 비교는, 제1 동작을 수행할 것인지 제2 동작을 수행할 것인지 결정하기 위해 사용되기보다는, 하나 이상의 동작들을 수행할지 여부(예컨대, 개개의 동작을 수행할지 또는 개개의 동작을 수행하는 것을 보류할지 여부)를 결정하기 위해 사용된다.

[0175] 일부 실시예들에서, 특성 세기를 결정하기 위해 제스처의 일부분이 식별된다. 예를 들어, 터치 감응형 표면은 시작 위치로부터 전이하여 종료 위치(이 지점에서 접촉의 세기가 증가함)에 도달하는 연속적인 스와이프 접촉을 수신한다. 이 예에서, 종료 위치에서의 접촉의 특성 세기는 스와이프 접촉 전체가 아니라 연속적인 스와이프 접촉의 일부분에만(예컨대, 종료 위치에서의 스와이프 접촉의 부분에만) 기초한다. 일부 실시예들에서, 접촉의 특성 세기를 결정하기 전에 스와이프 접촉의 세기들에 평활화 알고리즘(smoothing algorithm)이 적용된다. 예를 들어, 평활화 알고리즘은, 선택적으로, 비가중 이동 평균(unweighted sliding-average) 평활화 알고리즘, 삼각(triangular) 평활화 알고리즘, 메디안 필터(median filter) 평활화 알고리즘, 및/또는 지수(exponential) 평활화 알고리즘 중 하나 이상을 포함한다. 일부 상황들에서, 이 평활화 알고리즘들은 특성 세기를 결정하기 위해 스와이프 접촉의 세기들에서의 좁은 급등(spike)들 또는 급감(dip)들을 제거한다.

[0176] 터치 감응형 표면 상에서의 접촉의 세기는 접촉 검출 세기 임계치, 가볍게 누르기 세기 임계치, 깊게 누르기 세기 임계치, 및/또는 하나 이상의 다른 세기 임계치와 같은, 하나 이상의 세기 임계치들에 대해 특성화된다. 일부 실시예들에서, 가볍게 누르기 세기 임계치는, 디바이스가 물리적 마우스의 버튼 또는 트랙패드를 클릭하는 것과 전형적으로 연관된 동작들을 수행하게 될 세기에 대응한다. 일부 실시예들에서, 깊게 누르기 세기 임계치는, 디바이스가 물리적 마우스의 버튼 또는 트랙패드를 클릭하는 것과 전형적으로 연관된 동작들과는 상이한 동작들을 수행하게 될 세기에 대응한다. 일부 실시예들에서, 접촉이 가볍게 누르기 세기 임계치 미만의(예컨대, 그리고 공칭 접촉 검출 세기 임계치(이 미만에서는 접촉이 더 이상 검출되지 않음) 초과) 특성 세기로 검출될 때, 디바이스는 가볍게 누르기 세기 임계치 또는 깊게 누르기 세기 임계치와 연관된 동작을 수행함이 없이 터치 감응형 표면 상의 접촉의 이동에 따라 포커스 선택자를 이동시킬 것이다. 일반적으로, 달리 언급되지 않는 한, 이 세기 임계치들은 사용자 인터페이스 도면들의 상이한 세트들 사이에서 일관성이 있다.

[0177] 가볍게 누르기 세기 임계치 미만의 세기로부터 가볍게 누르기 세기 임계치와 깊게 누르기 세기 임계치 사이의 세기로의 접촉의 특성 세기의 증가는 때때로 "가볍게 누르기" 입력으로서 지칭된다. 깊게 누르기 세기 임계치 미만의 세기로부터 깊게 누르기 세기 임계치 초과 세기로의 접촉의 특성 세기의 증가는 때때로 "깊게 누르기" 입력으로서 지칭된다. 접촉 검출 세기 임계치 미만의 세기로부터 접촉 검출 세기 임계치와 가볍게 누르기 세기 임계치 사이의 세기로의 접촉의 특성 세기의 증가는 때때로 터치 표면 상에서의 접촉을 검출하는 것으로서 지칭된다. 접촉 검출 세기 임계치 초과 세기로부터 접촉 검출 세기 임계치 미만의 세기로의 접촉의 특성 세기의 감소는 때때로 터치 표면으로부터의 접촉의 리프트오프를 검출하는 것으로서 지칭된다. 일부 실시예들에서, 접촉 검출 세기 임계치는 영(0)이다. 일부 실시예들에서, 접촉 검출 세기 임계치는 0 초과이다.

[0178] 본 명세서에 기술된 일부 실시예들에서, 하나 이상의 동작들은, 개개의 누르기 입력을 포함하는 제스처를 검출하는 것에 응답하여 또는 개개의 접촉(또는 복수의 접촉들)으로 수행되는 개개의 누르기 입력을 검출하는 것에 응답하여 수행되며, 여기서 개개의 누르기 입력은 누르기 입력 세기 임계치 초과 접촉(또는 복수의 접촉들)의 세기의 증가를 검출하는 것에 적어도 부분적으로 기초하여 검출된다. 일부 실시예들에서, 개개의 동작은, 누르기 입력 세기 임계치 초과 개개의 접촉의 세기의 증가(예컨대, 개개의 누르기 입력의 "다운 스트로크(down stroke)")를 검출하는 것에 응답하여 수행된다. 일부 실시예들에서, 누르기 입력은 누르기 입력 세기 임계치 초과 개개의 접촉의 세기의 증가 및 누르기 입력 세기 임계치 미만의 접촉의 세기의 후속하는 감소를 포함하며, 개개의 동작은 누르기 입력 임계치 미만의 개개의 접촉의 세기의 후속하는 감소(예컨대, 개개의 누르기 입력의 "업 스트로크(up stroke)")를 검출하는 것에 응답하여 수행된다.

[0179] 일부 실시예들에서, 디바이스는 때때로 "지터(jitter)"로 지칭되는 우발적인 입력들을 회피하기 위해 세기 히스 테리시스를 채용하며, 여기서 디바이스는 누르기 입력 세기 임계치에 대한 미리정의된 관계를 갖는 히스테리시

스 세기 임계치(예컨대, 히스테리시스 세기 임계치는 누르기 입력 세기 임계치보다 더 낮은 X 세기 단위이거나, 히스테리시스 세기 임계치는 누르기 입력 세기 임계치의 75%, 90% 또는 어떤 적절한 비율임)를 정의하거나 선택한다. 이와 같이, 일부 실시예들에서, 누르기 입력은 누르기 입력 세기 임계치 초과와 개개의 접촉의 세기의 증가 및 누르기 입력 세기 임계치에 대응하는 히스테리시스 세기 임계치 미만의 접촉의 세기의 후속하는 감소를 포함하며, 개개의 동작은 히스테리시스 세기 임계치 미만의 개개의 접촉의 세기의 후속하는 감소(예컨대, 개개의 누르기 입력의 "업 스트로크")를 검출하는 것에 응답하여 수행된다. 유사하게, 일부 실시예들에서, 누르기 입력은 디바이스가 히스테리시스 세기 임계치 이하에서의 세기로부터 누르기 입력 세기 임계치 이상에서의 세기로의 접촉의 세기의 증가, 및 선택적으로, 히스테리시스 세기 이하에서의 세기로의 접촉의 세기의 후속하는 감소를 검출하는 경우에만 검출되고, 개개의 동작은 누르기 입력(예컨대, 주변환경에 따른 접촉의 세기의 증가 또는 접촉의 세기의 감소)을 검출하는 것에 응답하여 수행된다.

[0180] 설명의 편의상, 누르기 입력 세기 임계치와 연관된 누르기 입력에 응답하여 또는 누르기 입력을 포함하는 제스처에 응답하여 수행되는 동작들의 설명은, 선택적으로, 누르기 입력 세기 임계치 초과와 접촉의 세기의 증가, 히스테리시스 세기 임계치 미만의 세기로부터 누르기 입력 세기 임계치 초과와 세기로의 접촉의 세기의 증가, 누르기 입력 세기 임계치 미만의 접촉의 세기의 감소, 및/또는 누르기 입력 세기 임계치에 대응하는 히스테리시스 세기 임계치 미만의 접촉의 세기의 감소 중 어느 하나를 검출하는 것에 응답하여 트리거된다. 또한, 동작이 누르기 입력 세기 임계치 미만의 접촉의 세기의 감소를 검출하는 것에 응답하여 수행되는 것으로서 기술되어 있는 예들에서, 동작은, 선택적으로, 누르기 입력 세기 임계치에 대응하고 그보다 더 낮은 히스테리시스 세기 임계치 미만의 접촉의 세기의 감소를 검출하는 것에 응답하여 수행된다.

[0181] 3. 디지털 어시스턴트 시스템

[0182] 도 7a는 다양한 예들에 따른 디지털 어시스턴트 시스템(700)의 블록도를 도시한다. 일부 예들에서, 디지털 어시스턴트 시스템(700)은 독립형 컴퓨터 시스템 상에서 구현된다. 일부 예들에서, 디지털 어시스턴트 시스템(700)은 다수의 컴퓨터들에 걸쳐 분산된다. 일부 예들에서, 디지털 어시스턴트의 모듈들 및 기능들 중 일부는 서버 부분과 클라이언트 부분으로 분담되며, 여기서 클라이언트 부분은 하나 이상의 사용자 디바이스들(예컨대, 디바이스들(104, 122, 200, 400, 또는 600)) 상에 상주하고, 예를 들어 도 1에 도시된 바와 같은 하나 이상의 네트워크들을 통해 서버 부분(예컨대, 서버 시스템(108))과 통신한다. 일부 예들에서, 디지털 어시스턴트 시스템(700)은 도 1에 도시된 서버 시스템(108)(및/또는 DA 서버(106))의 일 구현예이다. 디지털 어시스턴트 시스템(700)은 단지 디지털 어시스턴트 시스템의 일례일 뿐이라는 것, 및 디지털 어시스턴트 시스템(700)이 도시된 것보다 더 많거나 더 적은 컴포넌트들을 가질 수 있거나, 둘 이상의 컴포넌트를 조합할 수 있거나, 또는 상이한 구성 또는 배열의 컴포넌트들을 가질 수 있다는 것에 유의해야 한다. 도 7a에 도시된 다양한 컴포넌트들은 하나 이상의 신호 처리 및/또는 주문형 집적 회로를 비롯한, 하드웨어, 하나 이상의 프로세서에 의한 실행을 위한 소프트웨어 명령어들, 펌웨어, 또는 이들의 조합으로 구현된다.

[0183] 디지털 어시스턴트 시스템(700)은 메모리(702), 하나 이상의 프로세서들(704), 입/출력(I/O) 인터페이스(706), 및 네트워크 통신 인터페이스(708)를 포함한다. 이들 컴포넌트들은 하나 이상의 통신 버스들 또는 신호 라인들(710)을 통해 서로 통신할 수 있다.

[0184] 일부 예들에서, 메모리(702)는 비일시적 컴퓨터 판독가능 매체, 예컨대, 고속 랜덤 액세스 메모리 및/또는 비휘발성 컴퓨터 판독가능 저장 매체(예컨대, 하나 이상의 자기 디스크 저장 디바이스, 플래시 메모리 디바이스, 또는 다른 비휘발성 솔리드 스테이트 메모리 디바이스)를 포함한다.

[0185] 일부 예들에서, I/O 인터페이스(706)는 디스플레이들, 키보드들, 터치 스크린들, 및 마이크로폰들과 같은 디지털 어시스턴트 시스템(700)의 입/출력 디바이스들(716)을 사용자 인터페이스 모듈(722)에 결합한다. I/O 인터페이스(706)는, 사용자 인터페이스 모듈(722)과 함께, 사용자 입력들(예컨대, 음성 입력, 키보드 입력들, 터치 입력들 등)을 수신하고 그에 따라 이들을 처리한다. 일부 예들에서, 예컨대 디지털 어시스턴트가 독립형 사용자 디바이스 상에서 구현되는 경우, 디지털 어시스턴트 시스템(700)은 도 2a, 도 4, 도 6a와 도 6b 각각의 디바이스들(200, 400, 또는 600)에 대해 기술된 컴포넌트들 및 I/O 통신 인터페이스들 중 임의의 것을 포함한다. 일부 예들에서, 디지털 어시스턴트 시스템(700)은 디지털 어시스턴트 구현예의 서버 부분을 표현하고, 사용자 디바이스(예컨대, 디바이스들(104, 200, 400, 또는 600)) 상에 상주하는 클라이언트 측 부분을 통해 사용자와 상호작용할 수 있다.

[0186] 일부 예들에서, 네트워크 통신 인터페이스(708)는 유선 통신 포트(들)(712) 및/또는 무선 송신 및 수신 회로부(714)를 포함한다. 유선 통신 포트(들)는 하나 이상의 유선 인터페이스들, 예컨대 이더넷, 범용 직렬 버스

(USB), 파이어와이어 등을 통해 통신 신호들을 수신하고 전송한다. 무선 회로부(714)는 통신 네트워크들 및 다른 통신 디바이스들로부터/로 RF 신호들 및/또는 광학 신호들을 수신하고 전송한다. 무선 통신은 GSM, EDGE, CDMA, TDMA, 블루투스, Wi-Fi, VoIP, Wi-MAX, 또는 임의의 다른 적합한 통신 프로토콜과 같은 복수의 통신 표준들, 프로토콜들, 및 기술들 중 임의의 것을 이용한다. 네트워크 통신 인터페이스(708)는 인터넷, 인트라넷, 및/또는 무선 네트워크, 예컨대 셀룰러 전화 네트워크, 무선 근거리 통신망(LAN), 및/또는 도시권 통신망(MAN)과 같은 네트워크들을 이용하는 디지털 어시스턴트 시스템(700)과 다른 디바이스들 간의 통신을 가능하게 한다.

[0187] 일부 예들에서, 메모리(702) 또는 메모리(702)의 컴퓨터 판독가능 저장 매체들은, 운영 체제(718), 통신 모듈(720), 사용자 인터페이스 모듈(722), 하나 이상의 애플리케이션들(724), 및 디지털 어시스턴트 모듈(726)의 전체 또는 그들의 서브세트를 포함한, 프로그램들, 모듈들, 명령어들, 및 데이터 구조들을 저장한다. 특히, 메모리(702) 또는 메모리(702)의 컴퓨터 판독가능 저장 매체는 후술되는 프로세스들을 수행하기 위한 명령어들을 저장한다. 하나 이상의 프로세서들(704)은 이러한 프로그램들, 모듈들, 및 명령어들을 실행하고, 데이터 구조들로부터/로 판독/기록한다.

[0188] 운영 체제(718)(예를 들어, 다윈(Darwin), RTXC, LINUX, UNIX, iOS, OS X, WINDOWS, 또는 VxWorks와 같은 임베디드 운영 체제)는 일반적인 시스템 태스크들(예를 들어, 메모리 관리, 저장 디바이스 제어, 전력 관리 등)을 제어하고 관리하기 위한 다양한 소프트웨어 컴포넌트들 및/또는 드라이버들을 포함하고, 다양한 하드웨어, 펌웨어, 및 소프트웨어 컴포넌트들 간의 통신들을 용이하게 한다.

[0189] 통신 모듈(720)은 네트워크 통신 인터페이스(708)를 통해 디지털 어시스턴트 시스템(700)과 다른 디바이스들 간의 통신을 용이하게 한다. 예를 들어, 통신 모듈(720)은 도 2a, 도 4, 도 6a와 도 6b 각각에 도시된 디바이스들(200, 400, 600)과 같은 전자 디바이스들의 RF 회로부(208)와 통신한다. 통신 모듈(720)은 또한 무선 회로부(714) 및/또는 유선 통신 포트(712)에 의해 수신된 데이터를 처리하기 위한 다양한 컴포넌트들을 포함한다.

[0190] 사용자 인터페이스 모듈(722)은 I/O 인터페이스(706)를 통해 사용자로부터(예컨대, 키보드, 터치 스크린, 포인팅 디바이스, 제어기, 및/또는 마이크로폰으로부터) 커맨드들 및/또는 입력들을 수신하고, 디스플레이 상에 사용자 인터페이스 객체들을 생성한다. 사용자 인터페이스 모듈(722)은 또한 출력들(예컨대, 스피치, 사운드, 애니메이션, 텍스트, 아이콘들, 진동들, 햅틱 피드백, 조명 등)을 준비하고, I/O 인터페이스(706)를 통해(예컨대, 디스플레이들, 오디오 채널들, 스피커들, 및 터치패드들 등을 통해) 사용자에게 그들을 전달한다.

[0191] 애플리케이션들(724)은 하나 이상의 프로세서들(704)에 의해 실행되도록 구성된 프로그램들 및/또는 모듈들을 포함한다. 예를 들어, 디지털 어시스턴트 시스템이 독립형 사용자 디바이스 상에 구현되는 경우, 애플리케이션들(724)은 게임들, 캘린더 애플리케이션, 내비게이션 애플리케이션, 또는 이메일 애플리케이션과 같은 사용자 애플리케이션들을 포함한다. 디지털 어시스턴트 시스템(700)이 서버 상에서 구현되는 경우, 애플리케이션들(724)은, 예를 들어 리소스 관리 애플리케이션들, 진단 애플리케이션들, 또는 스케줄링 애플리케이션들을 포함한다.

[0192] 메모리(702)는 또한 디지털 어시스턴트 모듈(726)(또는 디지털 어시스턴트의 서버 부분)을 저장한다. 일부 예들에서, 디지털 어시스턴트 모듈(726)은 하기의 하위모듈들, 또는 그들의 서브세트 또는 슈퍼세트를 포함한다: 입출력 처리 모듈(728), 스피치-텍스트(speech-to-text, STT) 처리 모듈(730), 자연 언어 처리 모듈(732), 대화 흐름 처리 모듈(734), 태스크 흐름 처리 모듈(736), 서비스 처리 모듈(738), 및 스피치 합성 처리 모듈(740). 이러한 모듈들 각각은 디지털 어시스턴트 모듈(726)의 하기의 시스템들 또는 데이터 및 모델들, 또는 그들의 서브세트 또는 슈퍼세트 중 하나 이상에 대해 액세스한다: 온톨로지(760), 어휘 인덱스(744), 사용자 데이터(748), 태스크 흐름 모델들(754), 서비스 모델들(756), 및 ASR 시스템들(758).

[0193] 일부 예들에서, 디지털 어시스턴트 모듈(726)에서 구현되는 처리 모듈들, 데이터, 및 모델들을 사용하여, 디지털 어시스턴트는 다음 중 적어도 일부를 수행할 수 있다: 스피치 입력을 텍스트로 변환하는 것; 사용자로부터 수신된 자연 언어 입력에 표현된 사용자의 의도를 식별하는 것; 사용자의 의도를 완전히 추론하는 데 필요한 정보를(예컨대, 단어들, 게임들, 의도들 등의 중의성을 해소함(disambiguating)으로써) 능동적으로 이끌어내고 획득하는 것; 추론된 의도를 이행하기 위한 태스크 흐름을 결정하는 것; 및 태스크 흐름을 실행하여 추론된 의도를 이행하는 것.

[0194] 일부 예들에서, 도 7b에 도시된 바와 같이, I/O 처리 모듈(728)은 도 7a의 I/O 디바이스들(716)을 통해 사용자와 상호작용하거나 또는 도 7a의 네트워크 통신 인터페이스(708)를 통해 사용자 디바이스(예컨대, 디바이스들(104, 200, 400, 또는 600))와 상호작용하여, 사용자 입력(예컨대, 스피치 입력)을 획득하고 사용자 입력에 대

한 응답들을 (예컨대, 스피치 출력들로서) 제공한다. I/O 처리 모듈(728)은 사용자 입력의 수신과 함께 또는 사용자 입력의 수신 직후에, 사용자 디바이스로부터 사용자 입력과 연관된 콘텍스트 정보를 선택적으로 획득한다. 콘텍스트 정보는 사용자-특정 데이터, 어휘, 및/또는 사용자 입력과 관련된 선호도들을 포함한다. 일부 예들에서, 콘텍스트 정보는 또한 사용자 요청이 수신되는 시간에서의 사용자 디바이스의 소프트웨어 및 하드웨어 상태들, 및/또는 사용자 요청이 수신된 시간에서의 사용자의 주위 환경에 관련된 정보를 포함한다. 일부 예들에서, I/O 처리 모듈(728)은 또한 사용자 요청에 관하여 사용자에게 후속 질문들을 전송하고, 그로부터 답변들을 수신한다. 사용자 요청이 I/O 처리 모듈(728)에 의해 수신되고 사용자 요청이 스피치 입력을 포함하는 경우, I/O 처리 모듈(728)은 스피치-텍스트 변환을 위해 스피치 입력을 STT 처리 모듈(730)(또는 스피치 인식기)로 전달한다.

[0195] STT 처리 모듈(730)은 하나 이상의 ASR 시스템들(758)을 포함한다. 하나 이상의 ASR 시스템들(758)은 I/O 처리 모듈(728)을 통해 수신되는 스피치 입력을 처리하여 인식 결과를 생성할 수 있다. 각각의 ASR 시스템(758)은 프론트-엔드 스피치 프리프로세서(front-end speech pre-processor)를 포함한다. 프론트-엔드 스피치 프리프로세서는 스피치 입력으로부터 대표적인 특징을 추출한다. 예를 들어, 프론트-엔드 스피치 프리프로세서는 스피치 입력에 대해 푸리에 변환을 수행하여 대표적인 다차원 벡터의 시퀀스로서 스피치 입력을 특징짓는 스펙트럼 특징을 추출한다. 또한, 각각의 ASR 시스템(758)은 하나 이상의 스피치 인식 모델들(예컨대, 음향 모델들 및/또는 언어 모델들)을 포함하고, 하나 이상의 스피치 인식 엔진들을 구현한다. 스피치 인식 모델의 예는 은닉 마르코프 모델(hidden Markov models), 가우시안 혼합 모델(Gaussian-Mixture Models), 딥 신경망 모델(Deep Neural Network Models), n-gram 언어 모델, 및 다른 통계 모델을 포함한다. 스피치 인식 엔진의 예는 동적 시간 왜곡 기반 엔진 및 가중치 유한 상태 변환기(WFST) 기반 엔진을 포함한다. 하나 이상의 스피치 인식 모델들 및 하나 이상의 스피치 인식 엔진들은 중간 인식 결과들(예를 들어, 음소, 음소 문자열, 및 하위 단어들), 및 궁극적으로 텍스트 인식 결과들(예컨대, 단어, 단어 문자열, 또는 토큰들의 시퀀스)을 생성하기 위해 프론트-엔드 스피치 프리프로세서의 추출된 대표 특징들을 처리하는 데 사용된다. 일부 예들에서, 스피치 입력은 제3자 서비스에 의해 또는 사용자의 디바이스(예컨대, 디바이스(104, 200, 400, 또는 600)) 상에서 적어도 부분적으로 처리되어 인식 결과를 생성한다. STT 처리 모듈(730)이 텍스트 문자열(예를 들어, 단어들, 또는 단어들의 시퀀스, 또는 토큰들의 시퀀스)을 포함하는 인식 결과를 생성하면, 인식 결과는 의도 추론을 위해 자연 언어 처리 모듈(732)로 전달된다. 일부 예들에서, STT 처리 모듈(730)은 스피치 입력의 다수의 후보 텍스트 표현들을 생성한다. 각각의 후보 텍스트 표현은 스피치 입력에 대응하는 단어들 또는 토큰들의 시퀀스이다. 일부 예들에서, 각각의 후보 텍스트 표현은 스피치 인식 신뢰도 점수와 연관된다. 스피치 인식 신뢰도 점수에 기초하여, STT 처리 모듈(730)은 후보 텍스트 표현들의 순위를 정하고, 의도 추론을 위해 n-베스트(n-best)(예컨대, n번째로 높은 순위가 매겨짐) 후보 텍스트 표현(들)을 자연 언어 처리 모듈(732)에 제공하는데, 여기서 n은 0보다 큰 미리결정된 정수이다. 예를 들어, 일례에서, 가장 높은 순위(n=1) 후보 텍스트 표현만이 의도 추론을 위해 자연 언어 처리 모듈(732)에 전달된다. 다른 예에서, 5번째 순위(n=5) 후보 텍스트 표현들이 의도 추론을 위해 자연 언어 처리 모듈(732)에 전달된다.

[0196] 스피치-텍스트 처리에 대한 보다 많은 상세사항들은 2011년 9월 20일자로 출원된 "Consolidating Speech Recognition Results"에 대한 미국 실용 특허 출원 제13/236,942호에 기술되어 있으며, 그 전체 개시내용은 본 명세서에 참조로서 편입된다.

[0197] 일부 예들에서, STT 처리 모듈(730)은 음성 기호 변환 모듈(731)을 통해 인식가능한 단어들의 어휘를 포함하고/하거나 이에 액세스한다. 각각의 어휘 단어는 스피치 인식 음성 기호로 표현된 단어의 하나 이상의 후보 발음들과 연관된다. 특히, 인식가능한 단어의 어휘는 복수의 후보 발음과 연관된 단어를 포함한다. 예를 들어, 어휘는 /tə'meɪrou/ 및 /tə'matou/ 의 후보 발음들과 연관된 "tomato"라는 단어를 포함한다. 또한, 어휘 단어는 사용자로부터의 이전 스피치 입력들에 기초한 맞춤형 후보 발음들과 연관된다. 이러한 맞춤형 후보 발음들은 STT 처리 모듈(730)에 저장되고, 디바이스 상의 사용자의 프로파일을 통해 특정 사용자와 연관된다. 일부 예들에서, 단어들에 대한 후보 발음들은 단어의 철자 및 하나 이상의 언어 및/또는 음성 규칙들에 기초하여 결정된다. 일부 예들에서, 후보 발음들은 예를 들어, 알려진 정준 발음들(canonical pronunciations)에 기초하여 수동으로 생성된다.

[0198] 일부 예들에서, 후보 발음들은 후보 발음의 공통성에 기초하여 순위결정된다. 예를 들어, 후보 발음 /tə'meɪrou/ 는 /tə'matou/ 보다 높게 순위결정되는데, 그 이유는 이전 발음이 더 일반적으로 사용되는 발음이기 때문이다(예를 들어, 모든 사용자 중에서, 특정 지리적 지역에 있는 사용자의 경우, 또는 임의의 다른 적

절한 사용자들의 세트의 경우). 일부 예들에서, 후보 발음들은 후보 발음이 사용자와 연관된 맞춤형 후보 발음인지 여부에 기초하여 순위결정된다. 예를 들어, 맞춤형 후보 발음들은 정준 후보 발음들보다 높게 순위결정된다. 이는 정준 발음에서 벗어나는 고유한 발음을 가진 고유 명사를 인식하는 데 유용할 수 있다. 일부 예들에서, 후보 발음들은 지리적 기원, 국적, 또는 민족성과 같은 하나 이상의 스피치 특성들과 연관된다. 예를 들어, 후보 발음 /tə'meɪrou/는 미국과 연관되는 반면, 후보 발음 /tə'matou/는 영국과 연관된다. 또한, 후보 발음의 순위는 디바이스 상의 사용자 프로파일에 저장된 사용자의 하나 이상의 특성들(예를 들어, 지리적 기원, 국적, 민족성 등)에 기초한다. 예를 들어, 사용자가 미국과 연관되어 있다는 것이 사용자의 프로파일로부터 결정될 수 있다. 미국과 연관되어 있는 사용자에게 기초하여, 후보 발음 /tə'meɪrou/(미국과 연관됨)는 후보 발음 /tə'matou/(영국과 연관됨)보다 높게 순위결정된다. 일부 예들에서, 순위결정된 후보 발음들 중 하나는 예측된 발음(예를 들어, 가장 가능성이 있는 발음)으로서 선택된다.

[0199] 스피치 입력이 수신될 때, STT 처리 모듈(730)은 (예를 들어, 음향 모델을 사용하여) 스피치 입력에 대응하는 음소를 결정하고, 이어서 (예를 들어, 언어 모델을 사용하여) 음소에 매칭되는 단어를 결정하고자 시도하는 데 사용된다. 예를 들어, STT 처리 모듈(730)이 스피치 입력의 일부에 대응하는 음소들의 시퀀스 /tə'meɪrou/를 먼저 식별하는 경우, 그것은 이어서 어휘 인덱스(744)에 기초하여 이 시퀀스가 단어 "tomato"에 대응한다고 결정할 수 있다.

[0200] 일부 예들에서, STT 처리 모듈(730)은 근사 매칭 기법들을 사용하여 발화 중인 단어를 결정한다. 따라서, 예를 들어, STT 처리 모듈(730)은 음소들의 특정 시퀀스가 그 단어에 대한 음소들의 후보 시퀀스 중 하나가 아니더라도 음소들의 시퀀스 /tə'meɪrou/가 단어 "tomato"에 대응한다고 결정한다.

[0201] 디지털 어시스턴트의 자연 언어 처리 모듈(732)("자연 언어 프로세서")은 STT 처리 모듈(730)에 의해 생성된 n-베스트 후보 텍스트 표현(들)("단어 시퀀스(들)" 또는 "토큰 시퀀스(들)")을 취하고, 후보 텍스트 표현들의 각각을 디지털 어시스턴트에 의해 인식된 하나 이상의 "행동가능한 의도들"과 연관시키고자 시도한다. "행동가능한 의도"(또는 "사용자 의도")는 디지털 어시스턴트에 의해 수행될 수 있는 태스크를 표현하고, 태스크 흐름 모델들(754)에서 구현되는 연관 태스크 흐름을 가질 수 있다. 연관 태스크 흐름은 디지털 어시스턴트가 태스크를 수행하기 위해 취하는 일련의 프로그래밍된 동작들 및 단계들이다. 디지털 어시스턴트의 능력들의 범주는 태스크 흐름 모델들(754)에서 구현되고 저장된 태스크 흐름들의 수 및 종류, 또는 다시 말해, 디지털 어시스턴트가 인식하는 "행동가능한 의도들"의 수 및 종류에 의존한다. 그러나, 디지털 어시스턴트의 효율성은 또한 자연 언어로 표현된 사용자 요청으로부터 정확한 "행동가능한 의도(들)"를 추론하는 어시스턴트의 능력에 의존한다.

[0202] 일부 예들에서, STT 처리 모듈(730)로부터 획득된 단어들 또는 토큰들의 시퀀스에 부가하여, 자연 언어 처리 모듈(732)은 또한 예를 들어, I/O 처리 모듈(728)로부터 사용자 요청과 연관된 콘텍스트 정보를 수신한다. 자연 언어 처리 모듈(732)은 선택적으로 콘텍스트 정보를 이용하여, STT 처리 모듈(730)로부터 수신된 후보 텍스트 표현들에 포함된 정보를 명확하게 하고, 보완하고, 그리고/또는 추가로 정의한다. 콘텍스트 정보는, 예를 들어 사용자 선호도들, 사용자 디바이스의 하드웨어 및/또는 소프트웨어 상태들, 사용자 요청 전, 요청 중, 또는 요청 직후에 수집되는 센서 정보, 디지털 어시스턴트와 사용자 사이의 이전 상호작용(예컨대, 대화) 등을 포함한다. 본 명세서에 기술되는 바와 같이, 일부 예들에서, 콘텍스트 정보는 동적이고, 시간, 위치, 대화의 내용, 및 다른 인자들에 따라 변한다.

[0203] 일부 예들에서, 자연 언어 처리는, 예컨대, 온톨로지(760)에 기초한다. 온톨로지(760)는 많은 노드들을 포함하는 계층 구조(hierarchical structure)이며, 각각의 노드는 "행동가능한 의도" 또는 "행동가능한 의도들" 중 하나 이상에 관련된 "속성", 또는 다른 "속성들" 중 어느 하나를 표현한다. 전술한 바와 같이, "행동가능한 의도"는 디지털 어시스턴트가 수행할 수 있는 태스크, 즉 그것이 "행동가능"하거나 또는 영향을 미칠 수 있는 태스크를 표현한다. "속성"은 행동가능한 의도 또는 다른 속성의 하위 양태와 연관되는 파라미터를 표현한다. 온톨로지(760) 내의 행동가능한 의도 노드와 속성 노드 간의 연결성(linkage)은 속성 노드에 의해 표현되는 파라미터가 행동가능한 의도 노드에 의해 표현되는 태스크에 어떻게 관련되는지를 정의한다.

[0204] 일부 예들에서, 온톨로지(760)는 행동가능한 의도 노드들 및 속성 노드들로 구성된다. 온톨로지(760) 내에서, 각각의 행동가능한 의도 노드는 직접적으로 또는 하나 이상의 중간 속성 노드들을 통해 하나 이상의 속성 노드들에 링크된다. 유사하게, 각각의 속성 노드는 직접적으로 또는 하나 이상의 중간 속성 노드들을 통해 하나 이상의 행동가능한 의도 노드들에 링크된다. 예를 들어, 도 7c에 도시된 바와 같이, 온톨로지(760)는 "레스토랑

예약" 노드(즉, 행동가능한 의도 노드)를 포함한다. 속성 노드들 "레스토랑", (예약을 위한) "날짜/시간", 및 "인원수"는 행동가능한 의도 노드(즉, "레스토랑 예약" 노드)에 각각 직접적으로 링크된다.

[0205] 추가로, 속성 노드들 "요리", "가격대", "전화번호", 및 "위치"는 속성 노드 "레스토랑"의 하위노드들이고, 중간 속성 노드 "레스토랑"을 통해 "레스토랑 예약" 노드(즉, 행동가능한 의도 노드)에 각각 링크된다. 다른 예를 들어, 도 7c에 도시된 바와 같이, 온톨로지(760)는 또한 "리마인더 설정" 노드(즉, 다른 행동가능한 의도 노드)를 포함한다. 속성 노드들 (리마인더 설정을 위한) "날짜/시간" 및 (리마인더를 위한) "주제"는 각각 "리마인더 설정" 노드에 링크된다. 속성 "날짜/시간"이 레스토랑 예약을 하는 태스크 및 리마인더를 설정하는 태스크 둘 다에 관련되기 때문에, 속성 노드 "날짜/시간"은 온톨로지(760) 내의 "레스토랑 예약" 노드 및 "리마인더 설정" 노드 둘 다에 링크된다.

[0206] 행동가능한 의도 노드는, 그의 링크된 속성 노드들과 함께, "도메인"으로 기술된다. 본 논의에서, 각각의 도메인은 개개의 행동가능한 의도와 연관되고, 특정한 행동가능한 의도와 연관된 노드들(및 이들 사이의 관계들)의 그룹을 지칭한다. 예를 들어, 도 7c에 도시된 온톨로지(760)는 온톨로지(760) 내의 레스토랑 예약 도메인(762)의 일례 및 리마인더 도메인(764)의 일례를 포함한다. 레스토랑 예약 도메인은 행동가능한 의도 노드 "레스토랑 예약", 속성 노드들 "레스토랑", "날짜/시간", 및 "인원수", 및 하위속성 노드들 "요리", "가격대", "전화번호", 및 "위치"를 포함한다. 리마인더 도메인(764)은 행동가능한 의도 노드 "리마인더 설정", 및 속성 노드들 "주제" 및 "날짜/시간"을 포함한다. 일부 예들에서, 온톨로지(760)는 많은 도메인들로 구성된다. 각각의 도메인은 하나 이상의 속성 노드들을 하나 이상의 다른 도메인들과 공유한다. 예를 들어, "날짜/시간" 속성 노드는 레스토랑 예약 도메인(762) 및 리마인더 도메인(764) 외에도, 많은 상이한 도메인들(예컨대, 스케줄링 도메인, 여행 예약 도메인, 영화 티켓 도메인 등)과 연관된다.

[0207] 도 7c가 온톨로지(760) 내의 2개의 예시적인 도메인들을 도시하지만, 다른 도메인들은, 예를 들어 "영화를 찾다", "전화 통화를 개시한다", "길을 찾는다", "회의를 스케줄링한다", "메시지를 전송한다", "질문에 대한 답변을 제공한다", "목록을 읽는다", "내비게이션 명령어들을 제공한다", "태스크에 대한 명령어들을 제공한다" 등을 포함한다. "메시지를 전송한다" 도메인은 "메시지를 전송한다" 행동가능한 의도 노드와 연관되고, "수신자(들)", "메시지 유형", 및 "메시지 본문"과 같은 속성 노드들을 추가로 포함한다. 속성 노드 "수신자"는, 예를 들어 "수신자 이름" 및 "메시지 주소"와 같은 하위속성 노드들에 의해 추가로 정의된다.

[0208] 일부 예들에서, 온톨로지(760)는 디지털 어시스턴트가 이해할 수 있고 그에 의거하여 동작할 수 있는 모든 도메인들(및 그에 따른 행동가능한 의도들)을 포함한다. 일부 예들에서, 온톨로지(760)는, 예컨대, 전체 도메인들 또는 노드들을 추가하거나 제거함으로써 또는 온톨로지(760) 내에서 노드들 사이의 관계를 수정함으로써 수정된다.

[0209] 일부 예들에서, 다수의 관련 행동가능한 의도들과 연관된 노드들은 온톨로지(760) 내에서 "상위 도메인" 아래에 군집된다. 예를 들어, "여행" 상위 도메인은 여행에 관련된 속성 노드들 및 행동가능한 의도 노드들의 군집을 포함한다. 여행에 관련된 행동가능한 의도 노드들은 "항공 예약", "호텔 예약", "자동차 렌탈", "길찾기", "관심 지점 찾기" 등을 포함한다. 동일한 상위 도메인(예컨대, "여행" 상위 도메인) 아래의 행동가능한 의도 노드들은 많은 속성 노드들을 공동으로 갖는다. 예를 들어, "항공 예약", "호텔 예약", "자동차 렌탈", "길찾기", 및 "관심 지점 찾기"에 대한 행동가능한 의도 노드들은 속성 노드들 "시작 위치", "목적지", "출발 날짜/시간", "도착 날짜/시간", 및 "인원수" 중 하나 이상을 공유한다.

[0210] 일부 예들에서, 온톨로지(760) 내의 각각의 노드는, 노드에 의해 표현되는 속성 또는 행동가능한 의도와 관련된 단어 및/또는 구절(phrase)들의 세트와 연관된다. 각각의 노드와 연관된 단어 및/또는 구절들의 개개의 세트는 노드와 연관된 소위 "어휘"이다. 각각의 노드와 연관된 단어 및/또는 구절들의 개개의 세트는 노드에 의해 표현되는 속성 또는 행동가능한 의도와 관련하여 어휘 인덱스(744)에 저장된다. 예를 들어, 도 7b로 돌아와서, "레스토랑"의 속성에 대한 노드와 연관된 어휘는 "음식", "음료", "요리", "배고프다", "먹다", "피자", "패스트푸드", "식사" 등과 같은 단어들을 포함한다. 다른 예를 들어, "전화 통화 개시"의 행동가능한 의도에 대한 노드와 연관된 어휘는 "통화", "전화", "다이얼", "벨소리", "이 번호로 전화", "~에게 전화" 등과 같은 단어 및 구절들을 포함한다. 어휘 인덱스(744)는 상이한 언어들의 단어 및 구절들을 선택적으로 포함한다.

[0211] 자연 언어 처리 모듈(732)은 STT 처리 모듈(730)로부터 후보 텍스트 표현들(예컨대, 텍스트 문자열(들) 또는 토큰 시퀀스(들))을 수신하고, 각각의 후보 표현에 대하여, 후보 텍스트 표현 내의 단어들에 의해 어떤 노드들이 연루되는지를 결정한다. 일부 예들에서, 후보 텍스트 표현 내의 단어 또는 구절이 (어휘 인덱스(744)를 통해) 온톨로지(760) 내의 하나 이상의 노드들과 연관되는 것으로 밝혀지는 경우, 단어 또는 구절은 이들 노드들을 "

트리거" 또는 "활성화"시킨다. 활성화된 노드들의 양 및/또는 상대적 중요도에 기초하여, 자연 언어 처리 모듈(732)은 사용자가 디지털 어시스턴트로 하여금 수행하도록 의도했던 태스크로서 행동가능한 의도들 중 하나를 선택한다. 일부 예들에서, 가장 많이 "트리거된" 노드들을 갖는 도메인이 선택된다. 일부 예들에서, 가장 높은 신뢰도 값(confidence value)(예컨대, 그것의 다양한 트리거된 노드들의 상대적 중요도에 기초함)을 갖는 도메인이 선택된다. 일부 예들에서, 도메인은 트리거된 노드들의 수 및 중요도의 조합에 기초하여 선택된다. 일부 예들에서, 디지털 어시스턴트가 사용자로부터의 유사한 요청을 이전에 정확하게 해석했는지 여부와 같은 추가 인자들도 노드를 선택하는 데 있어 마찬가지로 고려된다.

[0212] 사용자 데이터(748)는 사용자-특정 정보, 예컨대, 사용자-특정 어휘, 사용자 선호도, 사용자 주소, 사용자의 초기설정 언어 및 제2 언어, 사용자의 연락처 목록, 및 각각의 사용자에 대한 다른 단기 또는 장기 정보를 포함한다. 일부 예들에서, 자연 언어 처리 모듈(732)은 사용자-특정 정보를 이용하여 사용자 입력에 포함된 정보를 보완하여 사용자 의도를 추가로 정의한다. 예를 들어, "내 생일 파티에 내 친구들을 초대하라"라는 사용자 요청에 대하여, 자연 언어 처리 모듈(732)은, "친구들"이 누구인지 그리고 "생일 파티"가 언제, 어디에서 열릴지를 결정하기 위해 사용자에게 그의/그녀의 요청에서 그러한 정보를 명확하게 제공하도록 요구하는 대신, 사용자 데이터(748)에 액세스할 수 있다.

[0213] 일부 예들에서, 자연 언어 처리 모듈(732)은 하나 이상의 기계 학습 메커니즘들(예컨대, 신경망)을 이용하여 구현됨을 인식해야 한다. 특히, 하나 이상의 기계 학습 메커니즘들은 후보 텍스트 표현 및 후보 텍스트 표현과 연관된 콘텍스트 정보를 수신하도록 구성된다. 후보 텍스트 표현 및 연관된 콘텍스트 정보에 기초하여, 하나 이상의 기계 학습 메커니즘들은 후보 행동가능한 의도들의 세트에 대하여 의도 신뢰도 점수들을 결정하도록 구성된다. 자연 언어 처리 모듈(732)은 결정된 의도 신뢰도 점수들에 기초하여 후보 행동가능한 의도들의 세트로부터 하나 이상의 후보 행동가능한 의도들을 선택할 수 있다. 일부 예들에서, 온톨로지(예컨대, 온톨로지(760))는 또한 후보 행동가능한 의도들의 세트로부터 하나 이상의 후보 행동가능한 의도들을 선택하는 데 사용된다.

[0214] 토큰 문자열에 기초하여 온톨로지를 검색하는 다른 상세 사항들이 2008년 12월 22일자로 출원된 "Method and Apparatus for Searching Using An Active Ontology"에 대한 미국 실용 특허 출원 제12/341,743호에 기술되어 있으며, 이 출원의 전체 개시내용은 본 명세서에 참조로서 편입된다.

[0215] 일부 예들에서, 일단 자연 언어 처리 모듈(732)이 사용자 요청에 기초하여 행동가능한 의도(또는 도메인)를 식별하면, 자연 언어 처리 모듈(732)은 식별된 행동가능한 의도를 표현하기 위해 구조화된 질의(structured query)를 생성한다. 일부 예들에서, 구조화된 질의는, 행동가능한 의도를 위한 도메인 내의 하나 이상의 노드들에 대한 파라미터들을 포함하고, 파라미터들 중 적어도 일부는, 사용자 요청 내에 특정되는 특정 정보 및 요건들로 채워진다. 예를 들어, 사용자는 "7시에 스시 가게에 저녁 예약을 하라"라고 말한다. 이러한 경우, 자연 언어 처리 모듈(732)은 사용자 입력에 기초하여 행동가능한 의도가 "레스토랑 예약"이라고 정확하게 식별할 수 있다. 온톨로지에 따라, "레스토랑 예약" 도메인에 대한 구조화된 질의는 {요리}, {시간}, {날짜}, {인원수} 등과 같은 파라미터들을 포함한다. 일부 예들에서, 스피치 입력, 및 STT 처리 모듈(730)을 사용하여 스피치 입력으로부터 도출된 텍스트에 기초하여, 자연 언어 처리 모듈(732)은 레스토랑 예약 도메인에 대한 부분 구조화된 질의를 생성하며, 여기서 부분 구조화된 질의는 파라미터들 {요리 = "스시"} 및 {시간 = "7pm"}을 포함한다. 그러나, 이 예에서, 사용자의 발화는 도메인과 연관된 구조화된 질의를 완성하기에 불충분한 정보를 포함한다. 따라서, {인원수} 및 {날짜}와 같은 다른 필수 파라미터들은 현재 이용가능한 정보에 기초하여 구조화된 질의 내에 특정되어 있지 않다. 일부 예들에서, 자연 언어 처리 모듈(732)은 구조화된 질의의 일부 파라미터들을, 수신된 콘텍스트 정보로 채운다. 예를 들어, 일부 예들에서, 사용자가 "내 근처"의 스시 레스토랑을 요청했다면, 자연 언어 처리 모듈(732)은 구조화된 질의의 내의 {위치} 파라미터를, 사용자 디바이스로부터의 GPS 좌표들로 채운다.

[0216] 일부 예들에서, 자연 언어 처리 모듈(732)은 STT 처리 모듈(730)로부터 수신된 각각의 후보 텍스트 표현에 대하여 다수의 후보 행동가능한 의도들을 식별한다. 또한, 일부 예들에서, 각각의 식별된 후보 행동가능한 의도에 대하여 각자의 구조화된 질의(부분 또는 완전)가 생성된다. 자연 언어 처리 모듈(732)은 각각의 후보 행동가능한 의도에 대하여 의도 신뢰도 점수를 결정하고 의도 신뢰도 점수들에 기초하여 후보 행동가능한 의도들의 순위를 정한다. 일부 예들에서, 자연 언어 처리 모듈(732)은 생성된 구조화된 질의(또는 질의들)(임의의 완성된 파라미터들을 포함함)를 태스크 흐름 처리 모듈(736)("태스크 흐름 프로세서")로 전달한다. 일부 예들에서, m-베스트(예컨대, m번째 높은 순위) 후보 행동가능한 의도들에 대한 구조화된 질의(또는 질의들)가 태스크 흐름 처리 모듈(736)에 제공되며, 여기서 m은 0보다 큰 미리결정된 정수이다. 일부 예들에서, m-베스트 후보 행동가능

한 의도들에 대한 구조화된 질의(또는 질의들)는 대응하는 후보 텍스트 표현(들)과 함께 태스크 흐름 처리 모듈(736)에 제공된다.

- [0217] 스피치 입력의 다수의 후보 텍스트 표현들로부터 결정된 다수의 후보 행동가능한 의도들에 기초하여 사용자 의도를 추론하는 다른 상세사항들은 2014년 6월 6일자로 출원된 "System and Method for Inferring User Intent From Speech Inputs"에 대한 미국 실용 특허 출원 제14/298,725호에 기재되어 있으며, 그 전체 개시 내용은 본 명세서에 참고로 포함된다.
- [0218] 태스크 흐름 처리 모듈(736)은 자연 언어 처리 모듈(732)로부터 구조화된 질의(또는 질의들)를 수신하도록, 필요하다면, 구조화된 질의를 완성하도록, 그리고 사용자의 궁극적인 요청을 "완성"하는 데 필요한 동작들을 수행하도록 구성된다. 일부 예들에서, 이들 태스크들을 완수하는 데 필요한 다양한 절차들이 태스크 흐름 모델들(754)에서 제공된다. 일부 예들에서, 태스크 흐름 모델들(754)은 사용자로부터 추가 정보를 획득하기 위한 절차들, 및 행동가능한 의도와 연관된 동작들을 수행하기 위한 태스크 흐름들을 포함한다.
- [0219] 전술한 바와 같이, 구조화된 질의를 완성하기 위해, 태스크 흐름 처리 모듈(736)은 추가 정보를 획득하고/하거나 잠재적으로 모호한 발화들의 중의성을 해소하기 위해 사용자와의 추가 대화를 개시할 필요가 있다. 그와 같은 상호작용이 필요한 경우, 태스크 흐름 처리 모듈(736)은 대화 흐름 처리 모듈(734)을 호출하여 사용자와의 대화에 참여한다. 일부 예들에서, 대화 흐름 처리 모듈(734)은 어떻게(및/또는 언제) 사용자에게 추가 정보를 물을지를 결정하고, 사용자 응답을 수신하고 처리한다. 질문들은 I/O 처리 모듈(728)을 통해 사용자들에게 제공되고 그들로부터 답변들이 수신된다. 일부 예들에서, 대화 흐름 처리 모듈(734)은 오디오 및/또는 시각적 출력을 통해 사용자에게 대화 출력을 제시하고, 음성 또는 물리적(예컨대, 클릭킹) 응답들을 통해 사용자로부터 입력을 수신한다. 상기 예에 이어서, 태스크 흐름 처리 모듈(736)이 대화 흐름 처리 모듈(734)을 호출하여 도메인 "레스토랑 예약"과 연관된 구조화된 질의를 위한 "인원수" 및 "날짜" 정보를 결정하는 경우, 대화 흐름 처리 모듈(734)은 "몇 명입니까?" 및 "언제입니까?"와 같은 질문들을 생성하여 사용자에게 전달한다. 사용자로부터 답변들이 수신되면, 이어서 대화 흐름 처리 모듈(734)은 구조화된 질의를 누락 정보로 채우거나, 또는 태스크 흐름 처리 모듈(736)에 정보를 전달하여 구조화된 질의로부터 누락 정보를 완성한다.
- [0220] 일단 태스크 흐름 처리 모듈(736)이 행동가능한 의도에 대한 구조화된 질의를 완성했다면, 태스크 흐름 처리 모듈(736)은 행동가능한 의도와 연관된 궁극적인 태스크를 수행하도록 진행한다. 따라서, 태스크 흐름 처리 모듈(736)은 구조화된 질의에 포함된 특정 파라미터들에 따라 태스크 흐름 모델에서 단계들 및 명령어들을 실행한다. 예를 들어, "레스토랑 예약"의 행동가능한 의도에 대한 태스크 흐름 모델은 레스토랑에 연락하기 위한 그리고 특정 시간의 특정 인원수에 대한 예약을 실제로 요청하기 위한 단계들 및 명령어들을 포함한다. 예를 들어, {레스토랑 예약, 레스토랑 = ABC 카페, 날짜 = 3/12/2012, 시간 = 7pm, 인원수 = 5}와 같은 구조화된 질의를 사용하여, 태스크 흐름 처리 모듈(736)은, (1) OPENTABLE®과 같은 레스토랑 예약 시스템 또는 ABC 카페의 서버에 로그인하는 단계, (2) 웹사이트 상에서 일정 서식으로 날짜, 시간, 및 인원수 정보를 입력하는 단계, (3) 그 서식을 제출하는 단계, 및 (4) 사용자의 캘린더 내에 예약을 위한 캘린더 엔트리를 만드는 단계를 수행한다.
- [0221] 일부 예들에서, 태스크 흐름 처리 모듈(736)은 사용자 입력에서 요청된 태스크를 완성하거나 또는 사용자 입력에서 요청된 정보제공형 답변을 제공하기 위해 서비스 처리 모듈(738)("서비스 처리 모듈")의 어시스턴트를 채용한다. 예를 들어, 서비스 처리 모듈(738)은 태스크 흐름 처리 모듈(736)을 대신해서 전화 통화하도록, 캘린더 엔트리를 설정하도록, 지도 검색을 호출하도록, 사용자 디바이스 상에 설치된 다른 사용자 애플리케이션들을 호출하거나 그들과 상호작용하도록, 그리고 제3자 서비스들(예를 들어, 레스토랑 예약 포털, 소셜 네트워킹 웹사이트, 은행업무 포털 등)을 호출하거나 그들과 상호작용하도록 동작한다. 일부 예들에서, 각각의 서비스에 의해 요구되는 프로토콜들 및 애플리케이션 프로그래밍 인터페이스(API)들은 서비스 모델들(756) 중 개개의 서비스 모델에 의해 특정된다. 서비스 처리 모듈(738)은 서비스를 위한 적절한 서비스 모델에 액세스하고, 서비스 모델에 따른 서비스에 의해 요구되는 프로토콜들 및 API들에 따라 서비스에 대한 요청들을 생성한다.
- [0222] 예를 들어, 레스토랑이 온라인 예약 서비스를 가능하게 했다면, 레스토랑은 예약을 하는 데 필요한 파라미터들 및 온라인 예약 서비스에 필요한 파라미터의 값들을 전달하기 위한 API들을 특정하는 서비스 모델을 제출한다. 태스크 흐름 처리 모듈(736)에 의해 요청될 때, 서비스 처리 모듈(738)은 서비스 모델에 저장된 웹 주소를 사용하여 온라인 예약 서비스와의 네트워크 연결을 확립하고, 온라인 예약 서비스의 API에 따른 포맷으로 예약의 필요한 파라미터들(예컨대, 시간, 날짜, 인원수)을 온라인 예약 인터페이스에 전송한다.
- [0223] 일부 예들에서, 자연 언어 처리 모듈(732), 대화 흐름 처리 모듈(734), 및 태스크 흐름 처리 모듈(736)은 사용

자의 의도를 추론 및 정의하도록, 사용자 의도를 더 명확히 하고 정제하게 하기 위해 정보를 획득하도록, 그리고 최종적으로 사용자의 의도를 이행하기 위해 응답(즉, 사용자에게로의 출력, 또는 태스크의 완수)을 생성하도록 총체적이고 반복적으로 사용된다. 생성된 응답은 사용자의 의도를 적어도 부분적으로 이행하는 스피치 입력에 대한 대화 응답이다. 또한, 일부 예들에서, 생성된 응답은 스피치 출력으로서 출력된다. 이들 예에서, 생성된 응답은 스피치 합성 처리 모듈(740)(예컨대, 스피치 합성기)로 전송되며, 여기서 스피치 합성 처리 모듈(740)은 스피치 형태로 대화 응답을 합성하도록 처리될 수 있다. 또 다른 예들에서, 생성된 응답은 스피치 입력에서 사용자 요청을 충족시키는 것과 관련된 데이터 콘텐츠이다.

[0224] 태스크 흐름 처리 모듈(736)이 자연 언어 처리 모듈(732)로부터 다수의 구조화된 질의들을 수신하는 예들에서, 태스크 흐름 처리 모듈(736)은 초기에 수신된 구조화된 질의들 중 제1 구조화된 질의를 처리하여 제1 구조화된 질의를 완성하고자 시도하고/하거나 제1 구조화된 질의에 의해 표현되는 하나 이상의 태스크들 또는 동작들을 실행한다. 일부 예들에서, 제1 구조화된 질의는 가장 높은 순위의 행동가능한 의도에 대응한다. 다른 예들에서, 제1 구조화된 질의는 대응하는 스피치 인식 신뢰도 점수들 및 대응하는 의도 신뢰도 점수들의 조합에 기초하여, 수신된 구조화된 질의들로부터 선택된다. 일부 예들에서, 태스크 흐름 처리 모듈(736)이 제1 구조화된 질의의 처리 시 오류를 겪게 되는 경우(예컨대, 필수 파라미터를 결정할 능력 부족으로 인해), 태스크 흐름 처리 모듈(736)은 더 낮은 순위의 행동가능한 의도에 대응하는 수신된 구조화된 질의들 중 제2 구조화된 질의를 선택 및 처리하도록 진행할 수 있다. 제2 구조화된 질의는, 예를 들어, 대응하는 후보 텍스트 표현의 스피치 인식 신뢰도 점수, 대응하는 후보 행동가능한 의도의 의도 신뢰도 점수, 제1 구조화된 질의 내의 누락된 필수 파라미터, 또는 이들의 임의의 조합에 기초하여 선택된다.

[0225] 스피치 합성 처리 모듈(740)은 사용자에게의 제시를 위한 스피치 출력을 합성하도록 구성된다. 스피치 합성 처리 모듈(740)은 디지털 어시스턴트에 의해 제공된 텍스트에 기초하여 스피치 출력을 합성한다. 예를 들어, 생성된 대화 응답은 텍스트 문자열의 형태이다. 스피치 합성 처리 모듈(740)은 텍스트 문자열을 가청 스피치 출력으로 변환한다. 스피치 합성 처리 모듈(740)은, 텍스트로부터의 스피치 출력을 생성하기 위하여 임의의 적절한 스피치 합성 기술을 사용하는데, 이는 편집 합성(concatenative synthesis), 단위 선택 합성, 다이폰 합성, 도메인-특정 합성, 포먼트 합성(formant synthesis), 조음 합성(articulatory synthesis), HMM(hidden Markov model) 기반 합성, 및 정현파 합성(sinewave synthesis)을 포함하지만 이로 한정되지 않는다. 일부 예들에서, 스피치 합성 처리 모듈(740)은 단어들에 대응하는 음소 문자열에 기초하여 개별 단어들을 합성하도록 구성된다. 예를 들어, 음소 문자열은 생성된 대화 응답의 단어와 연관된다. 음소 문자열은 단어와 연관된 메타데이터에 저장된다. 스피치 합성 처리 모듈(740)은 스피치 형태의 단어를 합성하기 위해 메타데이터 내의 음소 문자열을 직접 처리하도록 구성된다.

[0226] 일부 예들에서, 스피치 합성 처리 모듈(740)을 사용하는 대신에(또는 그에 부가하여), 스피치 합성은 원격 디바이스(예컨대, 서버 시스템(108)) 상에서 수행되고, 합성된 스피치는 사용자에게의 출력을 위해 사용자 디바이스로 전송된다. 예를 들어, 이는 디지털 어시스턴트에 대한 출력이 서버 시스템에서 생성되는 일부 구현예들에서 발생할 수 있다. 그리고 서버 시스템은 일반적으로 사용자 디바이스보다 많은 처리 능력 또는 리소스를 갖기 때문에, 클라이언트 측 합성에서 실행될 수 있는 것보다 더 높은 품질의 스피치 출력을 획득하는 것이 가능하다.

[0227] 디지털 어시스턴트에 대한 추가적인 상세 사항들은 2011년 1월 10일자로 출원되고 발명의 명칭이 "Intelligent Automated Assistant"인 미국 실용 특허 출원 제12/987,982호, 및 2011년 9월 30일자로 출원되고 발명의 명칭이 "Generating and Processing Task Items That Represent Tasks to Perform"인 미국 실용 특허 출원 제13/251,088호에서 발견될 수 있고, 이 출원들의 전체 개시내용들은 본 명세서에 참조로서 편입된다.

[0228] 4. 지속적인 다이얼로그를 위한 다중 상태 디지털 어시스턴트

[0229] 도 8 및 도 9는 전자 디바이스에서 디지털 어시스턴트를 사용하는 지속적인 다이얼로그를 위한 시스템들 및 프로세스들을 도시한다. 예를 들어, 전자 디바이스는 디바이스들(104, 200, 400, 600)(도 1, 도 2a, 도 4, 및 도 6a와 도 6b)을 포함하지만 이들로 제한되지 않는, 본 명세서에 기술된 임의의 디바이스를 포함할 수 있다. 따라서, 도 8 및 도 9와 연관된 전자 디바이스들은 전화기, 랩톱 컴퓨터, 데스크톱 컴퓨터, 태블릿, 웨어러블 디바이스(예컨대, 스마트 워치, 헤드 마운티드 디스플레이 등), 홈 스피커 등과 같은 임의의 유형의 사용자 디바이스에 대응할 수 있다는 것이 이해될 것이다. 또한, 본 명세서에 기술된 프로세스들은 디바이스로 그리고 디바이스로부터 전달된 정보를 갖는 서버에 의해 수행되거나, 디바이스 상에서 수행되거나, 또는 이들의 조합일 수 있다.

[0230] 대체적으로, 커맨드가 이어지는 트리거 구절을 포함하는 스피치 입력과 같은 제1 스피치 입력(802)이 사용자로부터 수신될 수 있다. 예를 들어, 사용자는 현재 날씨 조건들(예컨대, 현재 위치와 연관된 날씨)에 관하여 질문하기 위해 "안녕 시리, 날씨는 어때(Hey Siri, how's the weather)?"를 발화할 수 있다. 제1 스피치 입력을 수신하는 것에 응답하여, 예를 들어, 하나 이상의 날씨 데이터베이스들을 참조하여 사용자에게 "70도이고 지금은 화창합니다(It's 70 degrees and sunny right now)"와 같은 응답(804)을 제공할 수 있다. 특히, 출력의 텍스트 표현은 날씨 데이터베이스를 사용하여 획득되어, 텍스트-스피치(text-to-speech, TTS) 처리가 텍스트에 대해 수행되어 사용자에게 가청 출력을 제공하게 할 수 있다. 텍스트 표현에 대응하는 디스플레이된 출력은 또한, 가청 출력(예컨대, "70° "을 갖는 "태양" 아이콘을 포함하는 배너)과 함께 또는 그 대신에 디스플레이될 수 있다.

[0231] 대체적으로, 제1 출력이 제공되며, 제1 출력은 제1 상태에서 디지털 어시스턴트에 대응한다. 제1 출력은 디스플레이된 출력 및/또는 가청 출력을 포함할 수 있다. 예를 들어, 사용자에게 상호작용 세부사항들을 제공하기 위해 디지털 어시스턴트 객체가 다양한 상태들로 디스플레이될 수 있다. 특히, 이러한 상태들은 사용자에게, 예를 들어, 디지털 어시스턴트에 의해 제공되는 주목의 레벨에 대응하는 디지털 어시스턴트 상태, 디지털 어시스턴트에 의해 수집되는 정보의 양, 활용되는 처리 리소스들의 양 등을 사용자에게 나타낼 수 있다. 예를 들어, 일단 제공되는 응답(804)이 완료된다면(예컨대, 가청 출력이 종료됨), 디지털 어시스턴트 객체는 제1 상태에서 전자 디바이스의 디스플레이 상에 디스플레이된다. 제1 크기를 갖고 제1 애니메이션 프로파일(예컨대, 천천히 스월링(swirling)하는 조명들 및 특정 컬러 세트를 갖는 원형 구체(orb))을 갖는 디지털 어시스턴트 객체가 디스플레이될 수 있다. 제1 상태는, 예를 들어, 디지털 어시스턴트가 사용자로부터 후속 스피치 입력을 능동적으로 리스닝하고 있음을 나타낼 수 있다. 디지털 어시스턴트 객체가 디스플레이되는 동안, 사용자는 제2 스피치 입력(808)을 제공할 수 있다. 제2 스피치 입력(808)은 제1 스피치 입력(802) 및/또는 응답(804)에 다른 방식으로 관련된 후속 스피치 입력 또는 스피치 입력에 대응할 수 있다. 예를 들어, 사용자는 제1 스피치 입력(802) 및 응답(804)을 컨텍스트로서 사용하여, 시애틀 시의 날씨 상태들에 대한 암시적 요청에 대응하는 "시애틀은 어때(How about Seattle)?"를 발화할 수 있다.

[0232] 일부 예들에서, 제1 출력은 가청 출력을 포함할 수 있다. 예를 들어, 제1 출력은 디지털 어시스턴트 객체가 제1 상태(예컨대, 차임벨(chime) 또는 소프트 톤)로 전이하고 있거나 전이하였음을 나타내는 가청 신호를 포함할 수 있다. 가청 출력은 순간적일 수 있거나 또는 지속적일 수 있다. 예를 들어, 제1 가청 출력은 차임벨 또는 소프트 톤으로 시작할 수 있고, 이어서, 디지털 어시스턴트가 제1 상태에서 유지되는 지속기간(또는 더 짧은 지속기간) 동안 소프트 주변 오디오가 재생될 수 있다. 가청 출력은 전자 디바이스의 하나 이상의 스피커들에 제공될 수 있거나, 또는 2차 전자 디바이스(예컨대, 디바이스에 결합된 유선 또는 무선 헤드폰)를 통해 제공될 수 있다. 예를 들어, 사용자는 홈 스피커 시스템을 활용하고 있을 수 있거나, 또는 무선 헤드폰에 결합된 스마트폰 디바이스를 활용하고 있을 수 있다.

[0233] 추가적인 스피치가 디지털 어시스턴트로 지향되는지의 여부를 결정하기 위해 다양한 값들 및 다른 신호들이 활용될 수 있다. 특히, 일단 디지털 어시스턴트가 제1 상태로 전이하면 제1 복수의 값들(806)이 획득된다. 제1 복수의 값들(806)은 본 명세서에 기술된 바와 같이, 제2 스피치 입력(808)에 대응하는 제1 신뢰도 레벨을 결정하는 데 사용될 수 있다. 대체적으로, 제1 복수의 값들을 획득하는 것은 전자 디바이스가 에코 제거를 위해 구성되는지의 여부에 기초할 수 있다. 특히, 에코 제거를 위해 구성된 디바이스는 사용자가 요청을 제공하는 것을 완료한 직후에 제1 복수의 값들(및/또는 다른 값들)을 수집 및 분석하기 시작하는 기능을 포함할 수 있다. 이러한 기능이 인에이블되면, 사용자는 일단 응답이 제공되기 시작하면 디지털 어시스턴트 응답을 중단할 수 있다. 예를 들어, 사용자는 제1 구절 "쿠퍼티노 날씨는 어때(What is the weather in Cupertino)?"를 발화할 수 있다. 이에 응답하여, 디지털 어시스턴트는 응답 "날씨는 화창하고(It's sunny and) ..."를 출력하기 시작할 수 있다. 응답은, 예를 들어 응답이 제공되고 있는 동안 사용자가 디지털 어시스턴트를 중단할 수 있도록 텍스트-스피치(TTS) 출력을 포함할 수 있다. 특히, 사용자는 후속 발화 "미안, 샌프란시스코 말이야(Sorry, I meant San Francisco)."로 디지털 어시스턴트 응답을 중단할 수 있다. 에코 제거 인에이블형 디바이스를 사용하여, 제1 복수의 값들(806)은 일단 사용자가 제1 구절을 발화하기를 완료하면 즉시 수집되고 분석된다. 따라서, 에코 제거 인에이블형 디바이스를 사용하여, 에코 제거를 활용하기 위해 디지털 어시스턴트로부터의 임의의 가청 응답이 검출되고 처리될 수 있으며, 동시에, 디바이스는 사용자가 초기 입력을 제공하기를 완료한 후에 임의의 후속 스피치 및 관련 신호들을 즉시 분석하기 시작할 수 있다. 따라서, 전자 디바이스가 에코 제거를 위해 구성된다는 결정에 따라, 제1 복수의 값들의 획득은 제1 스피치 입력(802)의 검출된 종료에 응답하여 개시된다. 대안적으로, 전자 디바이스가 에코 제거를 위해 구성되지 않는다는 결정에 따라, 제1 복수의 값들의 획득은 제

공된 응답(804)의 검출된 종료에 응답하여 게시된다.

[0234] 제1 복수의 값들(806)은 제1 신뢰도 레벨을 결정하고/하거나 조정하기 위해 분석될 수 있다. 제1 신뢰도 레벨은 추가적인 스피치가 디지털 어시스턴트로 지향되는지의 여부에 대한 신뢰도에 대응할 수 있다. 특히, 일단 스피치의 최소 임계 지속기간(810)(예컨대, 스피치의 500 ms)이 검출되면, 본 명세서에 기술된 바와 같이, 제1 복수의 값들(806)의 각각의 값이 적어도 하나의 규칙을 만족하는지의 여부에 대한 결정이 이루어진다. 최소 임계 지속기간(810)이 검출되는 경우, 복수의 값들(806)의 각각의 값이 적어도 하나의 규칙을 만족하는지의 여부에 대한 결정이 이루어진다. 각자의 값이 적어도 하나의 규칙을 만족한다는 결정에 따라, 제1 신뢰도 레벨이 증가된다. 유사하게, 각자의 값이 적어도 하나의 규칙을 만족하지 않는다는 결정에 따라, 제1 신뢰도 레벨이 감소된다. 일부 예들에서, 각자의 값이 적어도 하나의 규칙을 만족하지 않는다는 결정에 따라, 제1 신뢰도 레벨이 조정 없이 유지된다. 다시 말해, 신뢰도 레벨은 각자의 값들 및 연관된 규칙들에 기초하여 동적으로 조정될 수 있다.

[0235] 대체적으로, 제1 복수의 값들(806)은 사용자 시선, 사용자 입술 움직임들, 사용자의 상대적 주목, 디바이스 위치 정보, 스피치 검출 등에 관련된 다양한 신호들에 기초할 수 있다. 본 명세서에서 논의된 바와 같이, 다양한 값들이 또한 사용자 프라이버시에 활용되지 않을 수 있다. 예를 들어, 사용자는 시선 또는 입술 움직임 정보를 제공하지 않는 것을 선택할 수 있거나, 또는 제한된 상황들에서만 그러한 정보를 제공할 수 있다. 일부 예들에서, (예컨대, 온-디바이스 스피치 검출기를 사용하여) 대체적인 스피치 검출이 활용된다. 구체적으로, 전자 디바이스에서 사용자 스피치가 검출되고 있음을 나타내는 값이 획득될 수 있다. 이러한 경우, 스피치 값이 스피치 규칙을 만족한다는 결정에 따라(예컨대, 사용자 스피치가 검출되고 있음), 제1 신뢰도 레벨이 증가된다. 일부 예들에서, 전자 디바이스의 디스플레이에 지향되는 사용자 시선이 검출되어, 사용자 시선이 디스플레이된 디지털 어시스턴트 객체에 지향되는지 여부에 대한 결정이 이루어지게 한다. 예를 들어, 사용자는 제2 스피치 입력(808)을 발화하기 전 및/또는 그 동안, 디스플레이된 디지털 어시스턴트 객체를 보고 있을 수 있다. 따라서, 사용자가 디지털 어시스턴트 객체를 보고 있음을 나타내는 시선 값이 획득될 수 있다. 이러한 경우, 시선 값이 시선 규칙을 만족한다는 결정에 따라(예컨대, 사용자 시선이 디지털 어시스턴트 객체에 지향되어야 함), 제1 신뢰도 레벨이 증가된다. 일부 예들에서, 사용자의 입술 움직임이 검출되고, 입술 움직임에 기초하여 대응하는 값이 획득된다. 예를 들어, 입술 움직임 값은 입술 움직임 지속기간, 입술 특정부들(예컨대, 입 외관, 안면 배향), 입술 움직임과 연관된 비짐(viseme)들 등 중 하나 이상에 대응할 수 있다. 이러한 경우, 입술 움직임 값이 입술 움직임 규칙을 만족한다는 결정에 따라, 제1 신뢰도 레벨이 증가된다. 예를 들어, 식별된 비짐은 제2 스피치 입력(808)과 연관된 사운드에 대응할 수 있고/있거나, 특정 입술 움직임 지속기간은 스피치 지속기간의 특정 부분(예컨대, 지속기간(810))에 대응할 수 있다.

[0236] 일부 예들에서, 사용자 시선과 연관된 대체적인 방향이 검출된다. 예를 들어, 사용자 시선이 대체적으로 전자 디바이스의 디스플레이에 지향된다는 결정에 기초하여 값이 획득된다. 이러한 경우, 시선 방향 값이 시선 방향 규칙을 만족한다는 결정에 따라(예컨대, 사용자 시선이 디바이스 디스플레이에 지향됨), 제1 신뢰도 레벨이 증가된다. 특히, 이러한 값은 사용자가 디지털 어시스턴트 객체를 보고 있지 않지만, 여전히 대체적으로 전자 디바이스와 결합되는 경우에 활용될 수 있다. 일부 예들에서, 위치 정보는 디바이스 위치, 디바이스 배향, 및/또는 디바이스 가속도 정보, 예컨대 디바이스 위치, 배향 및/또는 가속도에 대한 변경들과 연관될 수 있다. 예를 들어, 사용자는 제2 스피치 입력(808)을 발화하기 직전 및/또는 그 동안, 사용자의 얼굴의 대체적인 근접도로 디바이스를 들어올렸을 수 있다. 따라서, 관련 디바이스 들어올림을 나타내는 위치 값이 획득될 수 있다. 이러한 경우, 위치 값이 위치 규칙을 만족한다는 결정에 따라(예컨대, 스피치 입력 전 및/또는 그 동안의 디바이스 들어올림), 제1 신뢰도 레벨이 증가된다.

[0237] 대체적으로, 제1 신뢰도 레벨은 임계 신뢰도 레벨과 비교될 수 있다. 비교는 신뢰도 레벨이 조정됨에 따라 동적으로 발생할 수 있거나, 또는 하나 이상의 특정 시간들에서 발생할 수 있다. 예를 들어, 결정(812)은 제1 신뢰도 레벨과 제1 신뢰도 임계치의 비교를 포함할 수 있다. 결정(812)은 스피치의 최소 임계 지속기간(810)이 검출된 후에 발생할 수 있다. 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제2 상태에 있는 디지털 어시스턴트에 대응하여 제2 출력이 제공된다. 제2 상태는 디지털 어시스턴트 객체를 디스플레이하는 것 및/또는 가정 출력을 제공하는 것을 포함할 수 있다. 예를 들어, 제2 크기를 갖고 제2 애니메이션 프로파일을 갖는 디지털 어시스턴트 객체가 디스플레이될 수 있다. 제2 크기는 제1 디스플레이 상태와 연관된 제1 크기보다 더 클 수 있다. 또한, 제2 애니메이션 프로파일은 커맨드를 식별하기 위해 디바이스가 사용자의 스피치 입력을 능동적으로 샘플링하고 있다는 표시들을 포함할 수 있다. 예를 들어, 제2 상태에서의 디지털 어시스턴트 객체는 스월링 조명들(예컨대, 제1 상태에서보다 더 빠르게 스월링하는 조명들) 및 특정 컬러

세트(예컨대, 제1 상태보다 더 선명한 조명들)를 갖는 원형 구체를 포함할 수 있다. 제2 상태의 디지털 어시스턴트 객체는 또한, 검출된 스피치 진폭에 기초하여 디지털 어시스턴트 객체의 크기의 변동들을 포함할 수 있다. 디지털 어시스턴트 객체가 제2 상태에 디스플레이되는 동안, 전자 디바이스는 제2 스피치 입력을 계속해서 수신할 수 있다.

[0238] 일부 예들에서, 디지털 어시스턴트가 제2 상태로 전이하였거나 전이하고 있음을 나타내는 제2 가청 출력이 제공된다. 예를 들어, 디지털 어시스턴트 객체가 제1 상태에 있음을 나타내기 위해 제1 가청 출력(예컨대, 소프트 차임벨 또는 미세 톤)이 제공될 수 있는 반면, 디지털 어시스턴트가 제2 상태로 전이함을 나타내기 위해 제1 가청 출력과 구별되는 제2 가청 출력(예컨대, 더 큰 소리의 차임벨, 더 긴 톤 등)이 제공될 수 있다. 제1 상태 동안 제공되는 제1 지속적인 톤과는 구별되는 제2 지속적인 톤과 같은 지속적인 주변 소음(또는 미리결정된 지속기간의 잡음)이 또한 제공될 수 있다.

[0239] 도 8로 돌아가면, 단계(812)에서, 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라 제2 복수의 값들(814)이 또한 획득될 수 있다. 후속 단계에서, 제2 스피치 입력(808)에 대응하는 제2 신뢰도 레벨이 획득된다. 제2 신뢰도 레벨은, 예를 들어 단계(816)에서 결정된, 제1 복수의 값들(806) 및 제2 복수의 값들(814) 둘 모두에 기초할 수 있다. 대체적으로, 제2 복수의 값들(814)은 사용자의 스피치가 디지털 어시스턴트를 위해 의도되는지의 여부를 나타내는 추가적인 신호들에 기초할 수 있다. 특히, 이러한 신호들 및 대응하는 값들은 스피커 아이덴티티, 사용자 의도, 신경망들 등을 수반하는 결정들과 같은, 제1 복수의 신호들에 비해, 더 높고 달리 더 강건한 처리 능력들을 갖는 디바이스 프로세스들과 연관될 수 있다.

[0240] 일부 예들에서, (예컨대, 단어 "어때"의 검출에 기초하여) 질문 형태의 디지털 어시스턴트로부터의 정보에 액세스하는 것과 관련된 의도와 같은, 스피치와 연관된 사용자 의도를 식별하기 위해 미리결정된 지속기간의 스피치가 분석된다. 의도는 각자의 액션가능한 의도에 대응하는 온톨로지 노드의 식별에 기초하여 결정될 수 있다. 따라서, 제2 복수의 값들(814)에 포함된 의도 값이 결정된 의도에 대응할 수 있다. 이러한 경우, 의도 값이 의도 규칙을 만족한다는 결정에 따라(예컨대, 의도는 디지털 어시스턴트에 대한 사용자 요청과 연관되어야 함), 제2 신뢰도 레벨이 증가된다. 일부 예들에서, 스피치 인식이 수행된다. 특히, 제1 스피치 입력(802) 및 제2 스피치 입력(808)과 연관된 제1 스피커 프로파일 및 제2 스피커 프로파일이 각각 획득된다. 스피커 프로파일들은 대체적으로, 피치, 톤, 프라이(fry), 리듬, 템포 등과 같은, 스피치의 다양한 특성들을 식별할 수 있다. 스피커 프로파일들의 비교(예컨대, 하나 이상의 특성들 또는 특성들의 조합을 비교함)에 기초하여, 스피커 인식 값이 획득되어 제2 복수의 값들(814)에 포함된다. 이러한 경우, 스피커 인식 값이 스피커 인식 규칙을 만족한다는 결정에 따라(예컨대, 유사성 임계치에 기초하여 스피커 프로파일들이 매칭되거나 달리 충분히 유사함), 제2 신뢰도 레벨이 증가된다.

[0241] 일부 예들에서, 스피치 입력의 임베딩에 기초하여 사용자 의도의 표현을 획득하기 위해 순환 신경망(recurrent neural network, RNN)이 활용된다. 대체적으로, 스피치 인식 출력(예컨대, 자동 스피치 인식 컴포넌트로부터 획득됨)에 기초하여 격자 임베딩이 획득될 수 있다. 격자 임베딩은, RNN에 대한 입력으로서 스피치 인식을 사용하고 출력으로서 격자 임베딩을 획득함으로써 생성될 수 있다. 이어서, 제2 복수의 값들(814) 중의 값으로서 RNN 의도 값을 획득하기 위해 격자 임베딩에 기초하여 사용자 의도가 도출될 수 있다. 이러한 경우, RNN 의도 값이 RNN 의도 규칙을 만족한다는 결정에 따라(예컨대, 격자 임베딩으로부터 획득된 의도는 디지털 어시스턴트에 대한 사용자 요청과 연관되어야 함), 제2 신뢰도 레벨이 증가된다.

[0242] 일단 단계(816)에서 제2 신뢰도 레벨이 획득되면, 제2 신뢰도 레벨은, 예를 들어 완화 투표 프로세스(818) 동안, 제2 임계 신뢰도 레벨과 비교된다. 대체적으로, 신뢰도 레벨들을 획득하고 완화 투표 프로세스를 수행하는 것은 규칙 지향 프레임워크에 적어도 부분적으로 기초할 수 있고/있거나 데이터 학습형 모델에 적어도 부분적으로 기초할 수 있다. 예를 들어, 다양한 값들이 다른 값들보다 더 높게 가중될 수 있고, 따라서 다른 값들보다 더 많이 최종 신뢰도 값에 영향을 미칠 수 있다. 사용자 시선 값(사용자가 디지털 어시스턴트 객체를 보고 있는지의 여부와 연관됨)은, 예를 들어, 입술 움직임 값보다 더 높게 가중될 수 있다. 데이터 학습형 모델을 사용하는 접근법은 또한, 다양한 값들을 어떻게 가중할 것인지를 모델에 통지하기 위해 이전의 사용자 상호작용들로부터의 데이터를 활용할 수 있다. 예를 들어, 이전의 상호작용 데이터가, 디지털 어시스턴트에의 사용자 시선이 전형적으로, 디지털 어시스턴트로 지향되는 스피치에 대응함을 나타내는 경우, (제1 복수의 값들(806)로부터의) 각자의 사용자 시선 값의 가중치가 증가된다. 추가적인 변수들이 또한, 완화 투표 프로세스(818) 동안 고려될 수 있다. 예를 들어, 응답(804)(예컨대, 응답(804)의 시작, 응답(804)의 중간 부분, 응답(804)의 종료 등)과 제2 스피치 입력(808)의 시작 사이의 지속기간은 완화 투표 프로세스(818)의 결과에 영향을 미칠 수 있다. 특히, 응답(804)과 제2 스피치 입력(808) 사이의 더 짧은 지속기간은 제2 신뢰도 레벨에서, 증

가와 같은 조정을 초래할 수 있다.

[0243] 대체적으로, 상호작용 프로세스 전체에 걸친 다양한 값들의 변화들은 신뢰도 레벨들을 조정하거나 달리 영향을 미치는 데 사용되는 컨텍스트 정보를 제공할 수 있다. 예를 들어, 사용자가 제1 스피치 입력(802)을 제공하는 동안 제1 시선 패턴이 검출될 수 있다. 사용자가 제2 스피치 입력(808)을 제공하는 동안, 제2 시선 패턴이 검출될 수 있다. 제1 시선 패턴과 제2 시선 패턴 사이의 변화들은 다양한 신뢰도 레벨들에 영향을 미칠 수 있다. 예를 들어, 제1 시선 패턴은 사용자가 대체적으로 디바이스의 방향을 보고 있거나 고려하고 있음을 나타낼 수 있는 반면, 제2 시선 패턴은 사용자가 전체 제2 스피치 입력(808)을 말하는 동안 디지털 어시스턴트 객체를 직접 보고 있음을 나타낼 수 있다. 대체적인 주의성(attentiveness)으로부터 높은 주의성으로의 시선 패턴의 변화에 기초하여, 제2 신뢰도 값이 증가될 수 있다. 다른 변화들도 또한, 다양한 신뢰도 값들에 영향을 미칠 수 있다. 예를 들어, 사용자는 디바이스가 표면 상에 안착해 있는 동안(예컨대, 스마트폰이 테이블 상에 안착함) 초기에 디지털 어시스턴트와 상호작용할 수 있다. 사용자가 후속 스피치 입력을 발화할 때, 사용자는 디바이스를 보거나 달리 그에 더 직접적으로 말하기 위해 디바이스를 들어올렸을 수 있다. 디바이스 들어올림을 포함하는 디바이스 모션의 변화는 제2 신뢰도 레벨과 같은 다양한 신뢰도 레벨들을 증가시킬 수 있다. 다양한 음향 파라미터들, RNN 출력들, 및 다른 사용자 주의성 값들의 변화들(예컨대, 하나 이상의 별개의 입력들/출력들 사이의 변화들)은 신뢰도 값들이 각자의 변화들에 기초하여 증가되거나 감소되게 할 수 있다.

[0244] 제1 및 제2 복수의 값들은 추가로, 동적으로 업데이트되고/되거나 다양한 인자들에 기초하여 업데이트될 수 있다. 특히, 결과 후보는 제2 스피치 입력(808)의 특성들에 기초하여 식별될 수 있다. 결과 후보는 스피치 입력 내의 특정 단어(예컨대, "시애틀", "은", 또는 "어때"), 단어의 특정 부분, 스피치의 지속시간 등에 대응할 수 있다. 결과 후보를 식별하는 것에 응답하여, 예를 들어, 단계(820)에서, 업데이트된 제1 복수의 값들(806) 및 업데이트된 제2 복수의 값들(814)이 획득된다. 따라서, 단계(820)에서, 업데이트된 제1 및 제2 복수의 값들에 기초하여 업데이트된 제2 신뢰도 레벨이 획득된다. 이어서, 업데이트된 제2 신뢰도 레벨은 완화 투표 프로세스(822) 동안 제2 임계 신뢰도 레벨과 비교될 수 있다.

[0245] 대체적으로, (예컨대, 단계(818)에서) 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제2 스피치 입력은 계속해서 수신된다. 구체적으로, 사용자의 스피치가 디지털 어시스턴트로 지향된다는 충분히 높은 신뢰도가 주어지면, 디지털 어시스턴트는 제2 상태로 유지되고, 사용자의 스피치의 수신이 계속된다. 대안적으로, (예컨대, 단계(822)에서) 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과하지 않는다는 결정에 따라, 제2 스피치 입력의 수신이 중지된다. 여기서, 신뢰도 레벨은 스피치가 디지털 어시스턴트로 지향된다는 충분히 낮은 신뢰도를 반영하고, 따라서, 디지털 어시스턴트는 제2 상태를 종료한다(예컨대, 유희 또는 더 낮은 전력 상태로 복귀함).

[0246] 이제 도 9를 참조하면, 다른 예로서, 제1 신뢰도 레벨은 제1 임계 신뢰도 레벨을 초과하지 않는 것으로 결정될 수 있다. 특히, 사용자는 날씨 질의에 관한 스피치 입력(902)을 전자 디바이스에 제공할 수 있고, 스피치 입력(902)을 수신하는 것에 응답하여, 디지털 어시스턴트는 사용자에게 응답(904)(예컨대, "70도이고 지금은 화창합니다")을 제공할 수 있다. 여기서, 전자 디바이스는 사용자 시선을 검출하지 못할 수 있다(예컨대, 홈 스피커, 무선 헤드폰 등). 다른 예에서, 사용자는 사용자 시선을 검출할 수 있는 전자 디바이스에 스피치 입력(902)을 제공할 수 있지만, 사용자는 디바이스 상의 카메라들 및 다른 광 센서들이 사용자의 어떠한 얼굴 특징부들도 캡처할 수 없도록 위치될 수 있다. 따라서, 제1 복수의 값들(906)의 다양한 값들은 대응하는 규칙들을 만족하지 않는 것으로 결정될 수 있다. 특히, 시선 값은 사용자가 디지털 어시스턴트 객체를 보고 있어야 한다는 규칙을 만족하지 않을 수 있고/있거나, 시선 방향 값은 사용자가 적어도 대체적으로 전자 디바이스의 디스플레이를 보고 있어야 한다는 규칙을 만족하지 않을 수 있고/있거나, 입술 움직임 값은 사용자 입술 움직임이 사용자 스피치에 대응해야 한다는 규칙을 만족하지 않을 수 있다. 대응하는 규칙들을 만족하지 않는 각자의 값들 중 하나 이상에 기초하여, 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과하지 않는다는 결정이 이루어진다.

[0247] 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과하지 않는다는 결정에 따라, 디지털 어시스턴트 객체는 제1 상태에서 유지되는 것으로서 디스플레이될 수 있다. 예를 들어, 디지털 어시스턴트 객체는 디지털 어시스턴트가 스피치 입력을 기다리고 있지만 디지털 어시스턴트로 지향되는 스피치를 아직 인식하지 않았음을 사용자에게 나타내는 비교적 더 작은 객체(예컨대, 각자의 애니메이션 및/또는 컬러 세트를 갖는 더 작은 원형 객체)로서 계속해서 표현될 수 있다. 이러한 경우, 미리결정된 지속시간(910)과 연관된 제2 복수의 값들(908)이 결정된다. 제2 복수의 값들(908)은, 예를 들어 2초의 스피치와 같은, 미리결정된 지속시간(910)의 제2 스피치 입력(912)에 대응할 수 있다. 따라서, 제1 복수의 값들(906) 및 제2 복수의 값들(908)에 기초하여, 결정 단계(914)에서 제2 스피치 입력(912)에 대응하는 제2 신뢰도 레벨이 결정된다. 결정 단계(914)에서 제2 신뢰도 레벨이 제2 임계

신뢰도 레벨을 초과한다는 결정에 따라, 디지털 어시스턴트 객체는 제2 상태에 디스플레이되고, 제2 스피치 입력(912)은 (예컨대, 도 8과 관련하여 단계(812) 이후에 논의된 바와 같이) 계속해서 수신된다. 추가적인 변수들이 결정 단계(914) 동안 또한 고려될 수 있다. 예를 들어, 응답(904)(예컨대, 응답(904)의 시작, 응답(904)의 중간 부분, 응답(904)의 종료)과 제2 스피치 입력(912)의 시작 사이의 지속기간은 결정 단계(914)의 결과에 영향을 미칠 수 있다. 특히, 응답(904)과 제2 스피치 입력(912) 사이의 더 짧은 지속기간은 제2 신뢰도 레벨에서, 증가와 같은 조정을 초래할 수 있다.

[0248] 디스플레이된 정보와 연관된 콘텍스트는 또한, 다양한 신뢰도 레벨 조정들에 영향을 미칠 수 있다. 예를 들어, 사용자는 관심 도로들 및 지점들(예컨대, 레스토랑들, 슝들 등)과 같은 다양한 디스플레이된 어포던스들을 포함하는 지도 애플리케이션을 보고 있을 수 있다. 어포던스들과 연관된 콘텍스트 정보는 관심 지점들과 관련된 위치 정보, 경로 설정 정보, 디스플레이된 정보와 관련될 수 있는 다양한 스피치 관련 키워드들(예컨대, "가시오", "지도," "경로 설정" 등) 등을 포함할 수 있다. 콘텍스트 정보는 추가로, 사용자 시선이 다양한 어포던스들로 지향되는지의 여부를 나타낼 수 있다(예컨대, 사용자는 지도 상의 특정 관심 지점을 보고 있을 수 있음). 후속 스피치 입력이 콘텍스트 정보와 연관된다는 결정에 따라, 제2 신뢰도 레벨이 증가된다. 예를 들어, 사용자는 디스플레이된 지도 상에 묘사된 특정 레스토랑 어포던스를 응시하면서 "여기로 경로 설정해(route here)"를 발화할 수 있다. 용어 "경로 설정"의 사용이 콘텍스트 정보에 포함된 관련 키워드들 및/또는 특정 어포던스로 지향된 사용자 시선과 매칭된다면, 스피치 입력이 콘텍스트 정보와 연관된다는 결정이 이루어진다. 따라서, 각자의 신뢰도 레벨(들)은 "여기로 경로 설정해"라는 사용자의 스피치가 디지털 어시스턴트로 지향될 높은 가능성에 기초한다.

[0249] 콘텍스트 정보는 추가로, 디스플레이된 정보에 특정적인 다양한 개념들의 시맨틱 표현들을 포함할 수 있다. 예를 들어, 사용자는 메시징 애플리케이션(예컨대, 인스턴트 메시징, 이메일 등)을 보고 있을 수 있다. 메시징 애플리케이션은 대체적으로, (예컨대, 스포츠, 식사 등에 대응하는 시맨틱 공간의 상이한 영역들과는 대조적으로) 메시지들과 관련된 단어들, 구절들, 및 다른 용어에 대응하는 시맨틱 공간의 특정 영역과 연관될 수 있다. 스피치 입력을 수신하는 것에 응답하여, 입력의 시맨틱 표현이 획득될 수 있다. 사용자는, 예를 들어, "그 메시지에 대한 응답"을 발화할 수 있다. "메시지들"에 대응하는 시맨틱 표현은 단어들 "응답" 및 "메시지"의 사용에 적어도 부분적으로 기초하여 식별될 수 있다. 메시징 애플리케이션의 의미론적 표현이 스피치 입력의 시맨틱 표현에 대응한다는 결정에 따라, 스피치 입력과 연관된 각자의 신뢰도 레벨이 증가된다.

[0250] 대체적으로, 디지털 어시스턴트의 상태들은 디스플레이된 크기, 애니메이션 프로파일, 색상 스킴 등에 기초하여 달라질 수 있다. 도 10을 참조하면, 제1 디지털 어시스턴트 상태는 디지털 어시스턴트 객체(1002)와 같은, 제1 크기를 갖는 디지털 어시스턴트 객체에 대응할 수 있다. 예를 들어, 일단 디지털 어시스턴트가 제1 상태에 진입하고 디지털 어시스턴트로 지향되는 사용자 스피치(예컨대, 초기 상호작용에 관련된 후속 스피치)를 리스닝하기 시작하면, 디지털 어시스턴트 객체(1002)가 디스플레이된다. 디지털 어시스턴트 객체(1002)는 제1 속도로 이동하는 다양한 객체들을 포함할 수 있다. 일단 디지털 어시스턴트가 제2 상태로 전환되면, 디지털 어시스턴트 객체는 디지털 어시스턴트 객체(1004)를 통해 도시된 바와 같이 디스플레이될 수 있다. 특히, 디지털 어시스턴트 객체(1004)는 디지털 어시스턴트 객체(1002)보다 더 클 수 있고, 디지털 어시스턴트 객체(1004)의 크기는 수신된 스피치 진폭에 기초하여 변동할 수 있고, 디지털 어시스턴트 객체(1004) 내의 다양한 객체들은 디지털 어시스턴트 객체(1002) 내의 객체들에 비해 더 빠른 속도로 이동할 수 있다. 디지털 어시스턴트 객체(1004)의 애니메이션들 또는 색상 스킴은 또한, 수신된 스피치 진폭에 기초하여 동적으로 수정될 수 있다.

[0251] 일단 디지털 어시스턴트가 (예컨대, 종료점 검출에 기초하여) 전체 사용자 발화가 검출되었다고 결정하면, 디지털 어시스턴트는 디지털 어시스턴트 객체(1006)에 의해 표현되는 상태와 같은 제3 상태로 전환될 수 있다. 제3 상태는 처리 상태에 대응할 수 있다. 처리 상태 동안, 디지털 어시스턴트 객체(1006)는 디지털 어시스턴트 객체(1004)에 비해 크기가 수축할 수 있고, 각자의 속도로 스월링하는 다양한 객체들을 포함할 수 있다. 여기서, 객체들은 다른 상태들 동안 디지털 어시스턴트 객체 내의 객체들과는 현저하게 구별될 수 있다(예컨대, 더 작고/작거나 무색일 수 있음). 일단 디지털 어시스턴트가 사용자 스피치를 처리하기를 완료했다면, 디지털 어시스턴트는 디지털 어시스턴트 객체(1008)에 의해 표현되는 상태와 같은 제4 상태로 전환될 수 있다. 제4 상태 동안, 디지털 어시스턴트 객체(1008)는 디지털 어시스턴트 객체들(1002, 1006)의 크기와 일치하는 크기로 디스플레이될 수 있다. 또한, 디지털 어시스턴트 객체(1008)는 특정 애니메이션 프로파일 및 특정 색상 스킴(예컨대, 디지털 어시스턴트 객체(1002)에 도시된 것보다 더 조밀한 스월링 객체들의 집합)을 포함할 수 있다. 디지털 어시스턴트 객체(1008)는 디지털 어시스턴트를 표현할 수 있는 반면, 디지털 어시스턴트는 사용자 스피치 입력에 기초하여 응답(예컨대, 가청 및/또는 디스플레이된 응답)을 제공한다. 일단 디지털 어시스턴트가 응답을 제

공하기를 완료하면, 디지털 어시스턴트는 디지털 어시스턴트 객체(1002)에 의해 도시된 바와 같이 제1 상태로 복귀할 수 있다.

[0252] 일부 예들에서, 다양한 가청 출력들이 디지털 어시스턴트 상태들 사이의 전이들을 나타낼 수 있다. 특히, 도 9를 다시 참조하면, 사용자는 스마트폰에 결합된 홈 스피커 또는 무선 헤드폰과 같은 디바이스를 사용하여 디지털 어시스턴트와 상호작용할 수 있다. 디지털 어시스턴트는 일단 제1 복수의 값들(906)이 획득되기 시작하면, 제1 상태에 진입할 수 있다. 여기서, 디지털 어시스턴트가 제1 상태로 진입하는 것을 시그널링하기 위해, 제1 각자의 가청 출력이 제공된다(예컨대, 소프트 차임벨, 벨소리, 또는 유사한 톤). 지속적인 주변 오디오는 또한, 디지털 어시스턴트가 제1 상태에 있는 동안 제공될 수 있다. 일부 예들에서, 추가적인 오디오가 재생되고 있는 동안(예컨대, 사용자가 노래를 리스닝하고 있는 동안) 디지털 어시스턴트가 호출될 때, 제1 상태의 지속기간 동안 추가적인 오디오의 볼륨이 감소된다.

[0253] 또한, (예컨대, 단계(912)에서) 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과한다는 결정에 따라, 디지털 어시스턴트는 제2 상태에 진입한다. 제2 상태로의 진입은 제1 각자의 가청 출력과 구별되는 제2 각자의 가청 출력에 의해 시그널링될 수 있다(예컨대, 더 큰 소리의 또는 더 소프트한 차임벨, 벨소리, 또는 제1 출력과 유사한 톤, 2개 이상의 각자의 차임벨들, 벨소리들 등). 지속적인 주변 오디오는 또한, 디지털 어시스턴트가 제2 상태에 있는 동안 제공될 수 있는데, 이는 제1 상태 동안 제공된 임의의 지속적인 주변 오디오와는 현저하게 구별될 수 있다. 대안적으로, 지속적인 주변 오디오는 하나 이상의 상태들(예컨대, 제2 상태)에서만 제공될 수 있고, 다른 상태들(예컨대, 제1 상태)에서는 제공되지 않을 수 있다. 일부 예들에서, 사용자가 제1 상태로부터 제2 상태로의 전이 동안 미디어를 리스닝하고 있는 경우, 미디어 볼륨은 제2 상태로 전이할 때 추가로 조정될 수 있다(예컨대, 더 낮추어지거나 약간 상승됨). 예를 들어, 미디어 볼륨은 디지털 어시스턴트가 제1 상태로 전이할 때 초기에 제1 레벨로 낮추어질 수 있고, 제2 상태로 전이할 때 추가로 낮추어질 수 있다. 도 10의 디지털 어시스턴트 객체들(1006, 1008)에 대응하는 것들과 같은 다양한 추가적인 상태들은 다른 상태들과 현저하게 구별되는 각자의 오디오 출력들과 연관될 수 있다. 다양한 상태 특정적 오디오 출력들은 다양한 디지털 어시스턴트 디스플레이 상태들과 조합하여 제공될 수 있다. 예를 들어, 사용자는 무선 헤드폰을 통해 디지털 어시스턴트가 상태들을 통해 전이하는 것을 들을 수 있지만, 다양한 상태들을 통해 전이하는 디지털 어시스턴트 객체가 또한 헤드폰에 무선으로 결합된 스마트폰 상에 디스플레이될 수 있다.

[0254] 도 11a 및 도 11b는 다양한 예들에 따른, 디지털 어시스턴트를 사용하는 지속적인 다이얼로그를 위한 프로세스(1100)를 도시한다. 프로세스(1100)는, 예를 들어, 디지털 어시스턴트를 구현하는 하나 이상의 전자 디바이스들을 이용하여 수행된다. 일부 예들에서, 프로세스(1100)는 클라이언트-서버 시스템(예컨대, 시스템(100))을 사용하여 수행되고, 프로세스(1100)의 블록들은 서버(예컨대, DA 서버(106))와 클라이언트 디바이스 사이에서 임의의 방식으로 분담된다. 다른 예들에서, 프로세스(1100)의 블록들은 서버와 다수의 클라이언트 디바이스들(예컨대, 모바일 폰 및 스마트 워치) 사이에서 분담된다. 따라서, 프로세스(1100)의 일부분들이 클라이언트-서버 시스템의 특정 디바이스들에 의해 수행되는 것으로 본 명세서에 기술되지만, 프로세스(1100)는 그렇게 제한되지 않는다는 것이 이해될 것이다. 다른 예들에서, 프로세스(1100)는 단지 하나의 클라이언트 디바이스(예컨대, 사용자 디바이스(104))만을 또는 다수의 클라이언트 디바이스들만을 사용하여 수행된다. 프로세스(1100)에서, 일부 블록들은, 선택적으로, 조합되고, 일부 블록들의 순서는, 선택적으로, 변경되고, 일부 블록들은, 선택적으로, 생략된다. 일부 예들에서, 추가 단계들이 프로세스(1100)와 조합하여 수행될 수 있다.

[0255] 도 11a를 참조하면, 블록(1102)에서, 사용자로부터 제1 스피치 입력이 수신된다. 블록(1104)에서, 제1 스피치 입력을 수신하는 것에 응답하여, 제1 스피치 입력에 기초한 응답이 제공된다. 블록(1106)에서, 제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력이 제공된다. 일부 예들에서, 제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력을 제공하는 것은 제1 상태에서 디지털 어시스턴트 객체를 디스플레이하는 것 및 가청 출력을 제공하는 것 중 적어도 하나를 포함한다. 블록(1108)에서, 사용자로부터 제2 스피치 입력이 수신된다. 블록(1110)에서, 제1 복수의 값들이 획득된다. 일부 예들에서, 제1 복수의 값들을 획득하는 것은 전자 디바이스의 디스플레이에 지향된 사용자 시선을 검출하는 것, 사용자 시선이 디스플레이된 디지털 어시스턴트 객체에 지향되는지의 여부를 결정하는 것, 및 사용자 시선이 디스플레이된 디지털 어시스턴트 객체에 지향되는지의 여부에 대한 결정에 기초하여 제1 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 일부 예들에서, 제1 복수의 값들을 획득하는 것은 사용자와 연관된 입술 움직임 검출하는 것, 입술 움직임이 제1 스피치 입력에 대응하는지의 여부를 결정하는 것, 및 결정에 기초하여 제1 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 일부 예들에서, 제1 복수의 값들을 획득하는 것은 사용자 시선과 연관된 방향을 검출하는 것, 및 결정된 방향에 기초하여 제1 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 일부 예들에서, 제1 복수의 값들을 획득하는 것은

전자 디바이스와 연관된 위치 정보를 검출하는 것, 및 위치 정보에 기초하여 제1 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 일부 예들에서, 제1 복수의 값들을 획득하는 것은 전자 디바이스에서 스피치가 검출되는지의 여부를 결정하는 것, 및 전자 디바이스에서 스피치가 검출된다는 결정에 기초하여 제1 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 디바이스 및 사용자와 관련된 다양한 값들을 활용함으로써, 시스템은 사용자의 스피치가 디지털 어시스턴트로 지향되는지의 여부를 결정하기 위해 고유 정보를 활용한다. 이러한 고유 정보를 사용하는 것은 스피치 인식 정확도를 개선하고, 이어서, 디바이스를 더 효율적으로 만든다. 따라서, 이러한 특징들은 사용자가 디바이스를 더 신속하고 효율적으로 사용할 수 있게 함으로써 디바이스의 전력 사용량을 감소시키고 배터리 수명을 개선한다.

[0256] 블록(1112)에서, 제1 복수의 값들에 기초하여, 제2 스피치 입력에 대응하는 제1 신뢰도 레벨이 획득된다. 일부 예들에서, 제2 스피치 입력이 최소 임계 지속시간과 연관된다는 결정에 따라, 제1 복수의 값들의 각각의 값에 대해, 각자의 값이 적어도 하나의 규칙을 만족하는지의 여부에 대한 결정이 이루어진다. 일부 예들에서, 각자의 값이 적어도 하나의 규칙을 만족한다는 결정에 따라, 제1 신뢰도 레벨이 증가된다. 일부 예들에서, 전자 디바이스가 에코 제거를 위해 구성된다는 결정에 따라, 제1 복수의 값들을 획득하는 것은 제1 스피치 입력의 검출된 종료에 응답하여 개시된다. 일부 예들에서, 전자 디바이스가 에코 제거를 위해 구성된다는 결정에 따라, 제1 복수의 값들을 획득하는 것은 제공된 응답의 검출된 종료에 응답하여 개시된다. 에코 제거 능력들 및 스피치의 최소 임계 지속시간들을 고려함으로써, 시스템은 분석할 가장 관련있는 스피치에 초점을 맞추어 값 수집 및 신뢰도 결정들을 추가로 최적화한다. 이러한 초점은 스피치 인식 정확도를 개선하고, 이어서, 사용자 경험을 개선한다. 따라서, 이러한 특징들은, 무관한 또는 부정확한 응답들을 제거하고 초기 질의를 재표현, 조정, 또는 달리 수정할 기회를 사용자에게 추가로 제공함으로써 사용자가 디바이스를 더 효과적으로 사용하는 것을 가능하게 한다.

[0257] 도 11b를 참조하면, 블록(1114)에서, 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과하는지의 여부에 대한 결정이 이루어진다. 일부 예들에서, 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제2 복수의 값들이 획득되고, 제2 스피치 입력에 대응하는 제2 신뢰도 레벨이 제1 복수의 값들 및 제2 복수의 값들에 기초하여 획득된다. 일부 예들에서, 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제2 스피치 입력은 계속해서 수신된다. 일부 예들에서, 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과하지 않는다는 결정에 따라, 제2 스피치 입력은 수신이 중지된다. 각자의 값들로부터 도출된 신뢰도에 기초하여 스피치를 선택적으로 수신함으로써, 시스템은 디지털 어시스턴트로 지향될 가능성이 있는 관련 스피치에 초점을 더 잘 맞춘다. 관련 스피치에 초점을 맞추는 것은 무관한 스피치가 검출될 때 시스템 자원들을 보존함으로써 디바이스를 더 효율적으로 만든다. 따라서, 이러한 특징들은 사용자가 디바이스를 더 신속하고 효율적으로 사용할 수 있게 함으로써 디바이스의 전력 사용량을 감소시키고 배터리 수명을 개선한다.

[0258] 블록(1116)에서, 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과한다는 결정에 따라, 제2 상태에서 디지털 어시스턴트에 대응하는 제2 출력이 제공된다. 블록(1118)에서, 제2 스피치 입력이 계속해서 수신된다. 블록(1120)에서, 디지털 어시스턴트는 제1 상태에서 유지된다. 일부 예들에서, 제1 상태에서 디지털 어시스턴트 객체를 디스플레이하는 것을 포함하는, 제1 상태에서 디지털 어시스턴트에 대응하는 제1 출력이 제공되며, 여기서 디지털 어시스턴트 객체는 제1 신뢰도 레벨이 제1 임계 신뢰도 레벨을 초과하지 않는다는 결정에 따라 제1 상태에서 디스플레이되는 대로 유지된다. 블록(1122)에서, 제2 스피치 입력의 미리결정된 지속시간과 연관된 제2 복수의 값들이 획득된다. 일부 예들에서, 제2 복수의 값들을 획득하는 것은 제2 스피치 입력의 미리결정된 지속시간과 연관된 사용자 의도를 식별하는 것, 사용자 의도에 기초하여, 제2 스피치 입력이 디지털 어시스턴트로 지향되는지의 여부를 결정하는 것; 및 제2 스피치 입력이 디지털 어시스턴트로 지향되는지의 여부에 대한 결정에 기초하여 제2 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 일부 예들에서, 제2 복수의 값들을 획득하는 것은 제1 스피치 입력과 연관된 제1 스피커 프로파일을 인출하는 것, 제2 스피치 입력과 연관된 제2 스피커 프로파일을 획득하는 것, 제1 스피커 프로파일과 제2 스피커 프로파일과 비교하는 것, 및 비교에 기초하여 제2 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 일부 예들에서, 제2 복수의 값들을 획득하는 것은 스피치 인식 출력에 기초하여, 격자 임베딩을 획득하는 것, 격자 임베딩에 기초하여 사용자 의도를 결정하는 것, 및 사용자 의도에 기초하여 제2 복수의 값들의 각자의 값을 획득하는 것을 포함한다. 일부 예들에서, 제2 복수의 값들을 획득하는 것은 제2 스피치 입력에 기초하여 결과 후보를 식별하는 것을 포함한다. 일부 예들에서, 결과 후보를 식별하는 것에 응답하여, 업데이트된 제1 복수의 값들 및 업데이트된 제2 복수의 값들이 획득된다. 일부 예들에서, 업데이트된 제1 복수의 값들 및 업데이트된 제2 복수의 값들에 기초하여, 제2 스피치 입력에 대응하는 업데이트된 제2 신뢰도 레벨이 획득된다. 더 높은 처리 요건들과 잠재적으로 관련된 다양한 추가적인 값들을 활용함으로써, 시스템은 향상된 사용자 의도 결정을 제공하기 위해 필요할 때 이러한 선택 정보를 활용

한다. 이러한 향상된 스피치 인식은 부정확한 스피치 해석의 가능성을 감소시킴으로써 디바이스를 더 효율적으로 만든다. 따라서, 이러한 특징들은 사용자가 디바이스를 더 신속하고 효율적으로 사용할 수 있게 함으로써 디바이스의 전력 사용량을 감소시키고 배터리 수명을 개선한다.

[0259] 블록(1124)에서, 제1 복수의 값들 및 제2 복수의 값들에 기초하여, 제2 스피치 입력에 대응하는 제2 신뢰도 레벨이 획득된다. 일부 예들에서, 제2 신뢰도 레벨이 제2 임계 신뢰도 레벨을 초과한다는 결정에 따라, 디지털 어시스턴트 객체는 제2 상태에서 디스플레이되고, 제2 상태에서 디지털 어시스턴트 객체를 디스플레이하는 동안, 제2 스피치 입력은 계속해서 수신된다. 일부 예들에서, 어포던스가 디스플레이되며, 여기서 어포던스는 콘텍스트 정보와 연관된다. 일부 예들에서, 제2 스피치 입력이 콘텍스트 정보와 연관된다는 결정에 따라, 제2 신뢰도 레벨이 증가된다. 일부 예들에서, 콘텍스트 정보는 제1 시맨틱 표현을 포함한다. 일부 예들에서, 제2 스피치 입력과 연관된 제2 시맨틱 표현이 획득되고, 제1 시맨틱 표현이 제2 시맨틱 표현에 대응한다는 결정에 따라, 제2 신뢰도 레벨이 증가된다. 일부 예들에서, 콘텍스트 정보는 적어도 하나의 미리정의된 단어를 포함한다. 일부 예들에서, 제2 스피치 입력에 포함된 적어도 하나의 단어가 식별되고, 적어도 하나의 미리정의된 단어가 적어도 하나의 식별된 단어에 대응한다는 결정에 따라, 제2 신뢰도 레벨이 증가된다. 콘텍스트 정보 및 시맨틱 표현과 같은 인자들을 고려함으로써, 시스템은 수신된 사용자 스피치의 의도 결정을 추가로 향상시킨다. 이러한 향상된 스피치 인식은 부정확한 스피치 해석의 가능성을 감소시킴으로써 디바이스를 더 효율적으로 만든다. 따라서, 이러한 특징들은 사용자가 디바이스를 더 신속하고 효율적으로 사용할 수 있게 함으로써 디바이스의 전력 사용량을 감소시키고 배터리 수명을 개선한다.

[0260] 도 11a 및 도 11b를 참조하여 전술된 동작들은 선택적으로, 도 1 내지 도 4, 도 6a와 도 6b, 및 도 7a 내지 도 7c에 도시된 컴포넌트들에 의해 구현된다. 예를 들어, 프로세스(1300)의 동작들은 운영 체제(718), 애플리케이션 모듈(724), I/O 처리 모듈(728), STT 처리 모듈(730), 자연 언어 처리 모듈(732), 어휘 인덱스(744), 태스크 흐름 처리 모듈(736), 서비스 처리 모듈(738), 미디어 서비스(들)(120-1), 또는 프로세서(들)(220, 410, 704) 중 하나 이상에 의해 구현될 수 있다. 도 1 내지 도 4, 도 6a와 도 6b, 및 도 7a 내지 도 7c에 도시된 컴포넌트들에 기초하여 다른 프로세스들이 어떻게 구현되는지는 본 발명이 속한 기술분야에서 통상의 지식을 가진 자에게 자명할 것이다.

[0261] 일부 구현예들에 따르면, 전자 디바이스의 하나 이상의 프로세서들에 의한 실행을 위한 하나 이상의 프로그램들을 저장하는 컴퓨터 판독가능 저장 매체(예를 들어, 비일시적 컴퓨터 판독가능 저장 매체)가 제공되며, 여기서 하나 이상의 프로그램들은 본 명세서에 설명된 방법들 또는 프로세스들 중 임의의 것을 수행하기 위한 명령어들을 포함한다.

[0262] 일부 구현예들에 따르면, 본 명세서에 설명된 방법들 또는 프로세스들 중 임의의 것을 수행하기 위한 수단을 포함하는 전자 디바이스(예를 들어, 휴대용 전자 디바이스)가 제공된다.

[0263] 일부 구현예들에 따르면, 본 명세서에 설명된 방법들 또는 프로세스들 중 임의의 것을 수행하도록 구성된 처리 유닛을 포함하는 전자 디바이스(예를 들어, 휴대용 전자 디바이스)가 제공된다.

[0264] 일부 구현예들에 따르면, 하나 이상의 프로세서들 및 하나 이상의 프로세서들에 의한 실행을 위한 하나 이상의 프로그램들을 저장하는 메모리를 포함하는 전자 디바이스(예를 들어, 휴대용 전자 디바이스)가 제공되며, 여기서 하나 이상의 프로그램들은 본 명세서에 설명된 방법들 또는 프로세스들 중 임의의 것을 수행하기 위한 명령어들을 포함한다.

[0265] 전술한 설명은, 설명의 목적을 위해, 특정 실시예들을 참조하여 설명되었다. 그러나, 위의 예시적인 논의들은 본 발명을 개시된 정확한 형태들로 규명하거나 제한하려는 의도는 아니다. 많은 수정들 및 변형들이 위의 교시 내용들에 비추어 가능하다. 실시예들은 기술들의 원리 및 그것들의 실제적인 응용을 가장 잘 설명하기 위하여 선택되고 기술되었다. 따라서, 당업자들은 고려된 특정 사용에 적합한 바와 같이 다양한 수정을 이용하여 기술들 및 다양한 실시예들을 최상으로 활용하는 것이 가능하게 된다.

[0266] 본 개시내용 및 예들이 첨부 도면들을 참조하여 충분히 기술되었지만, 당업자들에게 다양한 변경들 및 수정들이 명백할 것이라는 것에 유의해야 한다. 그러한 변경들 및 수정들은 청구항들에 의해 정의되는 바와 같은 개시내용 및 예들의 범주 내에 포함되는 것으로 이해되어야 한다.

[0267] 전술된 바와 같이, 본 기술의 하나의 태양은 가상 어시스턴트를 사용하는 지속적인 다이얼로그를 개선하기 위해 다양한 소스들로부터 입수가능한 데이터의 수집 및 사용이다. 본 개시내용은, 일부 경우들에 있어서, 이러한 수집된 데이터가 특정 개인을 고유하게 식별하거나 또는 그와 연락하거나 그의 위치를 확인하는 데 이용될 수

있는 개인 정보 데이터를 포함할 수 있음을 고려한다. 그러한 개인 정보 데이터는 인구통계 데이터, 위치 기반 데이터, 전화번호들, 이메일 주소들, 트위터 ID들, 집 주소들, 사용자의 건강 또는 피트니스 레벨에 관한 데이터 또는 기록들(예를 들어, 바이탈 사인(vital sign) 측정치들, 약물 정보, 운동 정보), 생년월일, 또는 임의의 다른 식별 또는 개인 정보를 포함할 수 있다. 그러나, 이러한 데이터는 또한, 상기의 특징들을 구현하는 데 필수적인 것은 아니며, 제한된 경우들에 사용될 수 있거나 또는 어떠한 경우들에서도 사용되지 않을 수 있다.

[0268] 본 개시내용은 본 기술에서의 그러한 개인 정보 데이터의 이용이 사용자들에게 이득을 주기 위해 사용될 수 있음을 인식한다. 예를 들어, 눈 시선 데이터와 같은 개인 정보는 사용자가 디지털 어시스턴트 객체를 응시하고 있는지의 여부를 결정하는 데 사용될 수 있다. 게다가, 사용자에게 이득을 주는 개인 정보 데이터에 대한 다른 이용들이 또한 본 개시내용에 의해 고려된다. 예를 들어, 건강 및 피트니스 데이터는 사용자의 일반적인 웰니스(wellness)에 대한 통찰력을 제공하는 데 사용될 수 있거나, 또는 웰니스 목표를 추구하기 위한 기술을 이용하여 개인들에게 긍정적인 피드백으로서 사용될 수 있다.

[0269] 본 개시내용은 그러한 개인 정보 데이터의 수집, 분석, 공개, 전달, 저장, 또는 다른 사용을 담당하는 엔티티들이 잘 확립된 프라이버시 정책들 및/또는 프라이버시 관례들을 준수할 것임을 고려한다. 특히, 이러한 엔티티들은, 대체로 개인 정보 데이터를 사적이고 안전하게 유지시키기 위한 산업적 또는 행정적 요건들을 충족시키거나 넘어서는 것으로 인식되는 프라이버시 정책들 및 관례들을 구현하고 지속적으로 사용해야 한다. 그러한 정책들은 사용자들에 의해 쉽게 액세스가능해야 하고, 데이터의 수집 및/또는 사용이 변화됨에 따라 업데이트되어야 한다. 사용자들로부터의 개인 정보는 엔티티의 적법하며 합리적인 용도들을 위해 수집되어야 하고, 이들 적법한 용도들을 벗어나서 공유되거나 판매되어서는 안 된다. 추가로, 그러한 수집/공유는 사용자들의 통지된 동의를 수신한 후에 발생해야 한다. 부가적으로, 이러한 엔티티들은, 이러한 개인 정보 데이터에 대한 액세스를 보호하고 안전하게 하며 개인 정보 데이터에 대한 액세스를 갖는 다른 사람들이 그들의 프라이버시 정책들 및 절차들을 고수한다는 것을 보장하기 위한 임의의 필요한 조치들을 취하는 것을 고려해야 한다. 게다가, 이러한 엔티티들은 널리 인정된 프라이버시 정책들 및 관례들에 대한 그들의 고수를 증명하기 위해 제3자들에 의해 그들 자신들이 평가를 받을 수 있다. 부가적으로, 정책들 및 관례들은 수집된 그리고/또는 액세스된 특정 유형들의 개인 정보 데이터에 대해 조정되고, 관찰구역 특정 고려사항들을 포함하여 적용가능한 법률들 및 표준들로 조정되어야 한다. 예를 들어, 미국에서, 소정 건강 데이터의 수집 또는 그에 대한 액세스는 연방법 및/또는 주의 법, 예를 들어 미국 건강 보험 양도 및 책임 법령(Health Insurance Portability and Accountability Act, HIPAA)에 의해 통제될 수 있는 반면; 다른 국가들에서의 건강 데이터는 다른 규정들 및 정책들의 적용을 받을 수 있고 그에 따라 취급되어야 한다. 따라서, 상이한 프라이버시 관례들은 각각의 국가의 상이한 개인 데이터 유형들에 대해 유지되어야 한다.

[0270] 진술한 것에도 불구하고, 본 개시내용은 또한 사용자들이 개인 정보 데이터의 사용, 또는 그에 대한 액세스를 선택적으로 차단하는 실시예들을 고려한다. 즉, 본 개시내용은 그러한 개인 정보 데이터에 대한 액세스를 방지하거나 차단하기 위해 하드웨어 및/또는 소프트웨어 요소들이 제공될 수 있다는 것을 고려한다. 예를 들어, 본 기술은 사용자들이 서비스들에 대한 등록 동안에 또는 그 이후의 언제라도 개인 정보 데이터의 수집에의 참여에 "동의(opt in)" 또는 "동의하지 않음(opt out)"을 선택하게 하도록 구성될 수 있다. 다른 예에서, 사용자들은 입술 움직임 데이터를 제공하지 않을 것을 선택할 수 있다. 또 다른 예에서, 사용자들은 눈 시선 데이터와 관련하여 제공된 세부사항들을 제한할 것을 선택할 수 있다. "동의" 및 "동의하지 않음" 옵션들을 제공하는 것에 부가하여, 본 개시내용은 개인 정보의 액세스 또는 사용에 관한 통지들을 제공하는 것을 고려한다. 예를 들어, 사용자는 그들의 개인 정보 데이터가 액세스될 앱을 다운로드할 시에 통지받고, 이어서 개인 정보 데이터가 앱에 의해 액세스되기 직전에 다시 상기하게 될 수 있다.

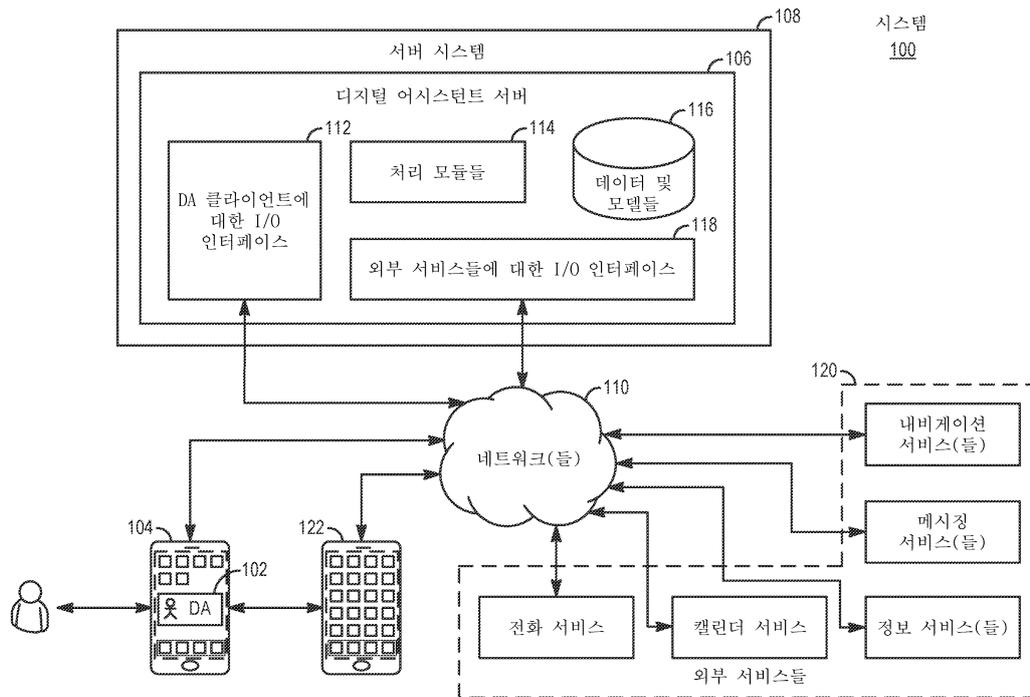
[0271] 또한, 의도하지 않은 또는 인가되지 않은 액세스 또는 이용의 위험을 최소화하는 방식으로 개인 정보 데이터가 관리되고 취급되어야 한다는 것이 본 개시내용의 의도이다. 데이터의 수집을 제한하고 데이터가 더 이상 필요하지 않게 되면 데이터를 삭제함으로써 위험이 최소화될 수 있다. 추가로, 그리고 소정의 건강 관련 애플리케이션들을 비롯하여, 적용가능할 때, 사용자의 프라이버시를 보호하기 위해 데이터 비식별화가 사용될 수 있다. 적절한 경우, 특정 식별자들(예컨대, 생년월일 등)을 제거함으로써, 저장된 데이터의 양 또는 특이성을 제어함으로써(예컨대, 주소 수준이라기보다는 오히려 도시 수준에서 위치 데이터를 수집함으로써), 데이터가 저장되는 방식을 제어함으로써(예컨대, 사용자들에 걸쳐 데이터를 집계함으로써), 그리고/또는 다른 방법들에 의해, 비식별화가 용이하게 될 수 있다.

[0272] 따라서, 본 개시내용이 하나 이상의 다양한 개시된 실시예들을 구현하기 위해 개인 정보 데이터의 사용을 광범위하게 커버하지만, 본 개시내용은 다양한 실시예들이 또한 그러한 개인 정보 데이터에 액세스할 필요 없이 구

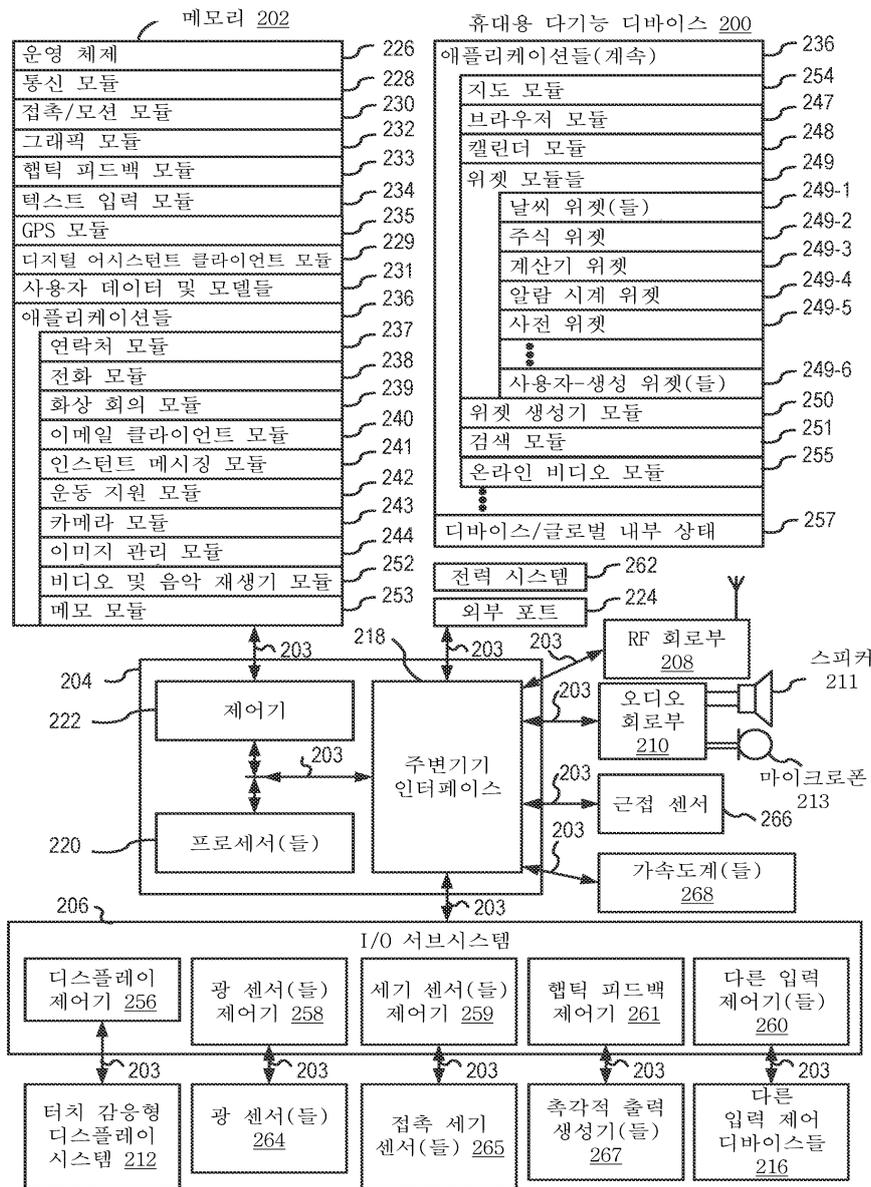
현될 수 있다는 것을 또한 고려한다. 즉, 본 기술의 다양한 실시예들은 이러한 개인 정보 데이터의 전부 또는 일부분의 결여로 인해 동작 불가능하게 되지 않는다. 예를 들어, 지속적인 다이얼로그는 익명화된 눈-시선 데이터, 지속적인 다이얼로그 시스템에 이용가능한 다른 비-개인 정보, 또는 공개적으로 입수가능한 정보와 같은 비-개인 정보 데이터 또는 가장 기본적인 최소량의 개인 정보에 기초하여 선호도들을 추론함으로써 용이해질 수 있다.

도면

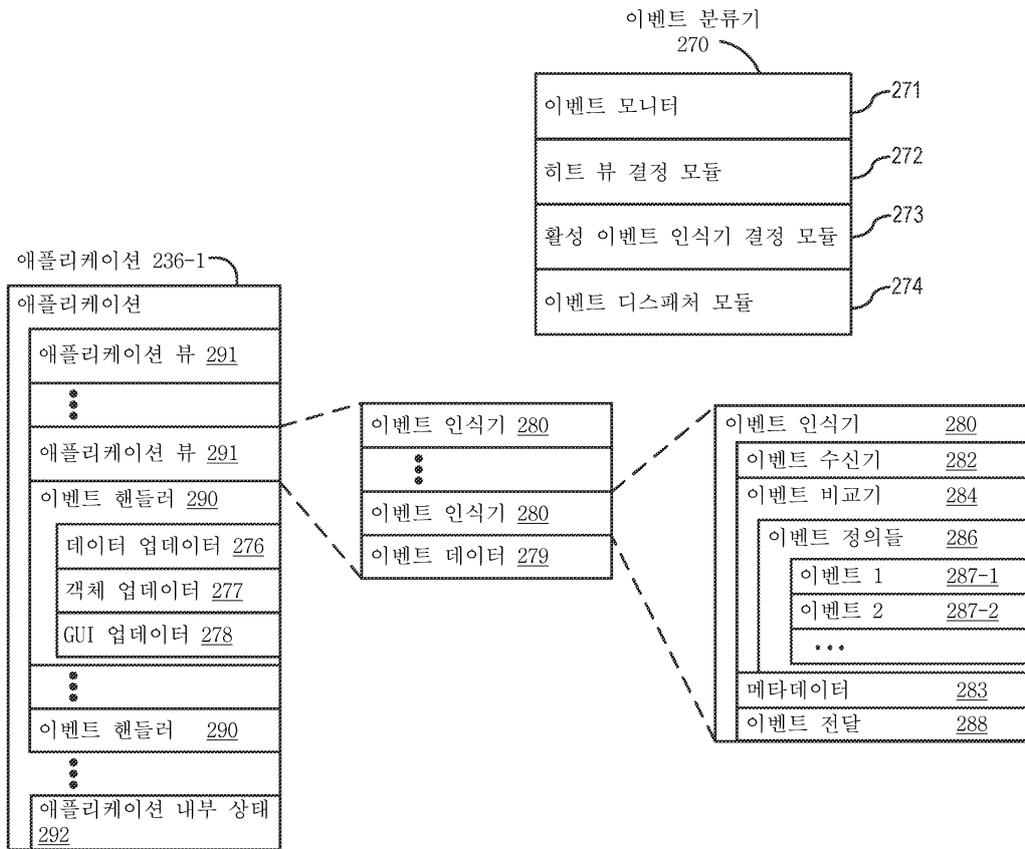
도면1



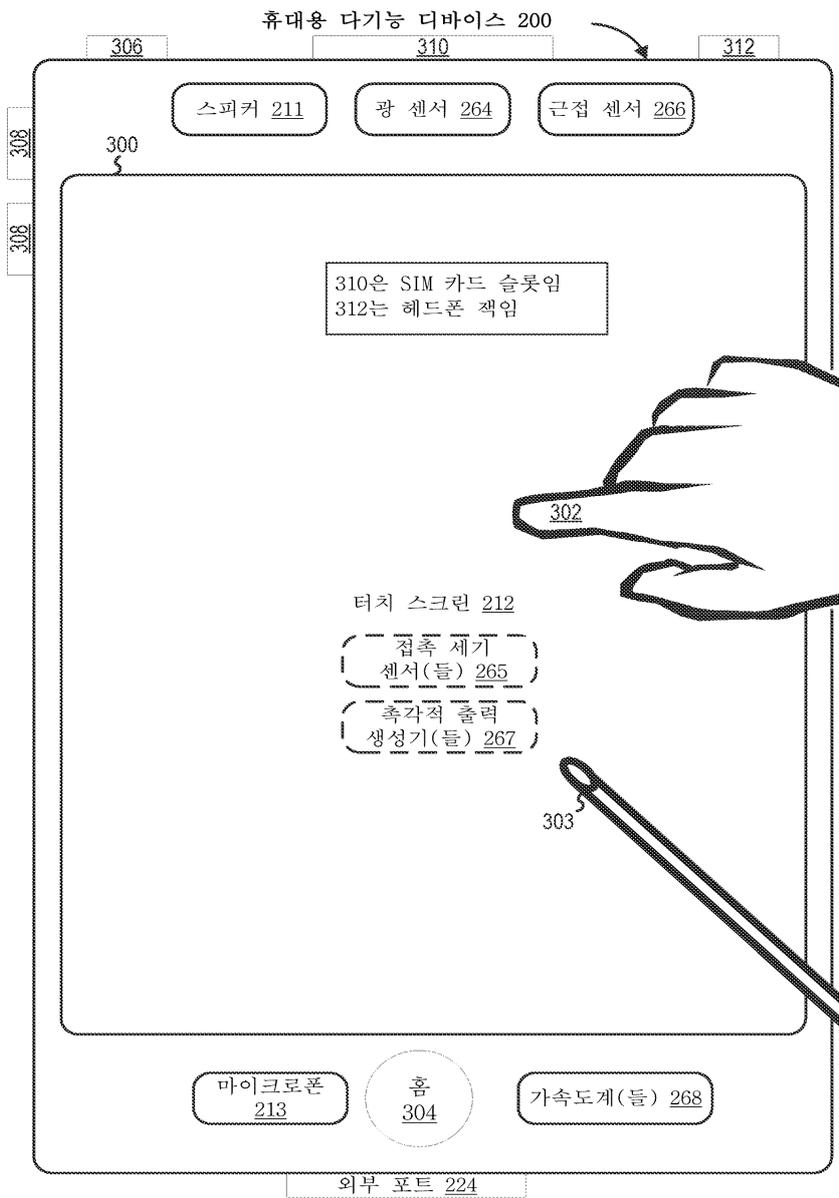
도면2a



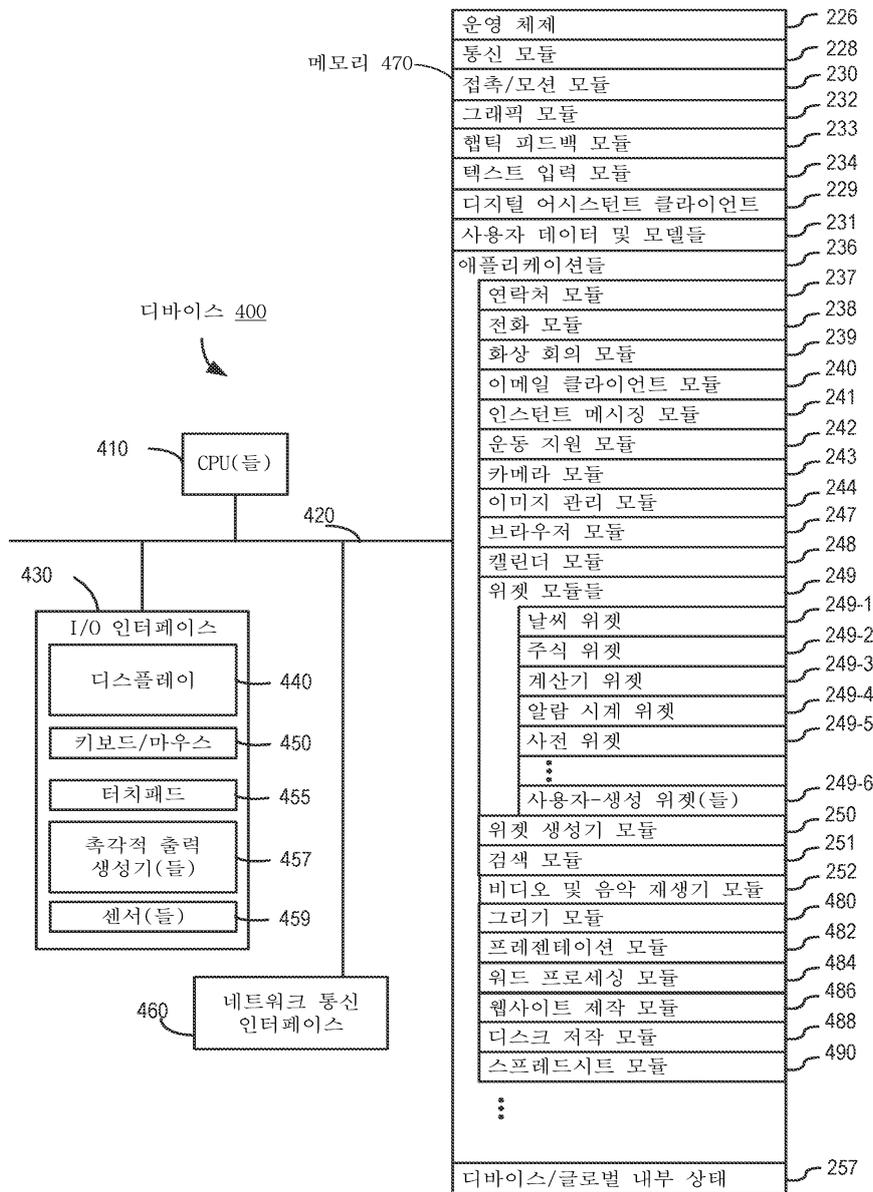
도면2b



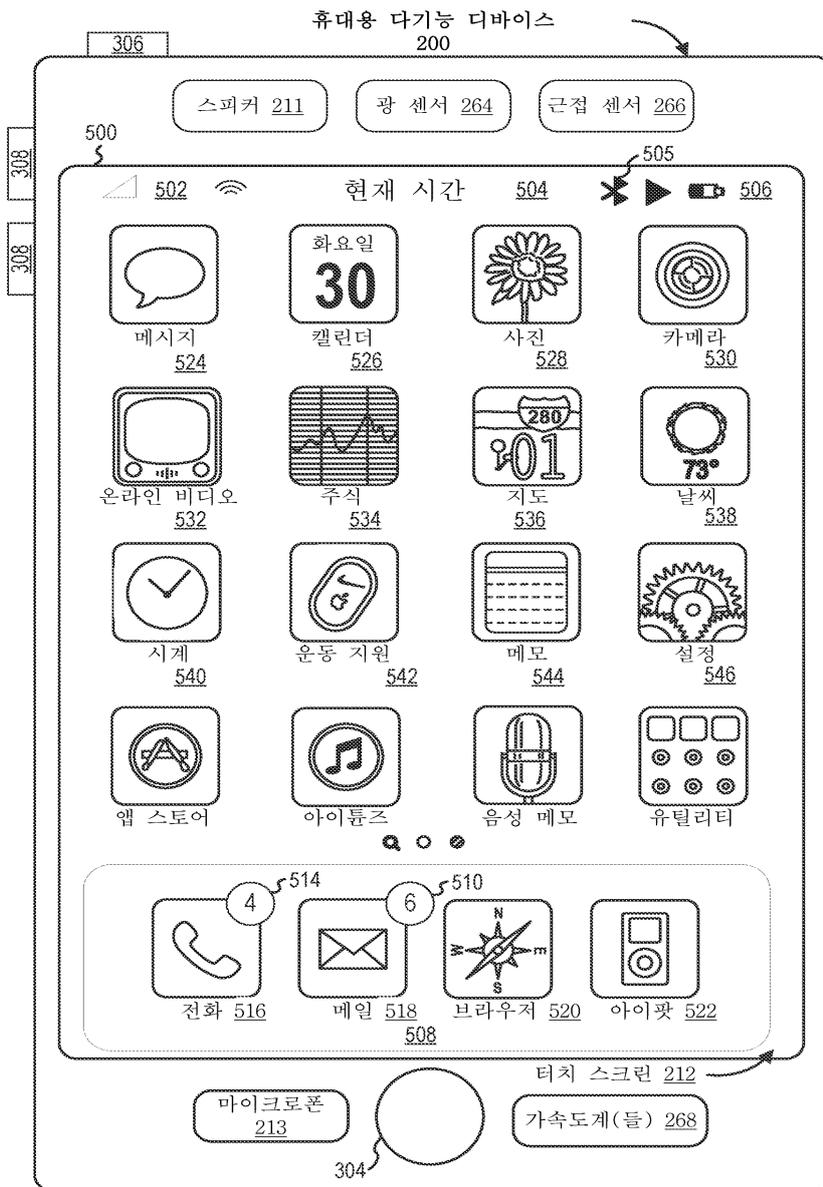
도면3



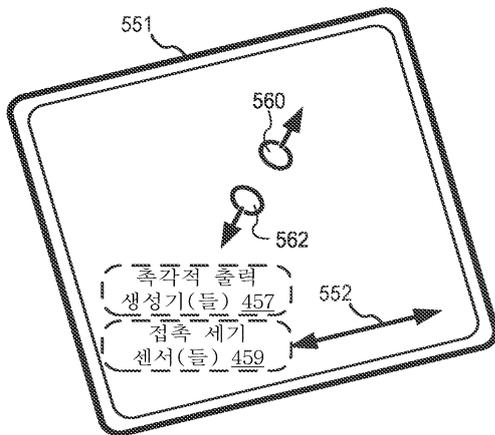
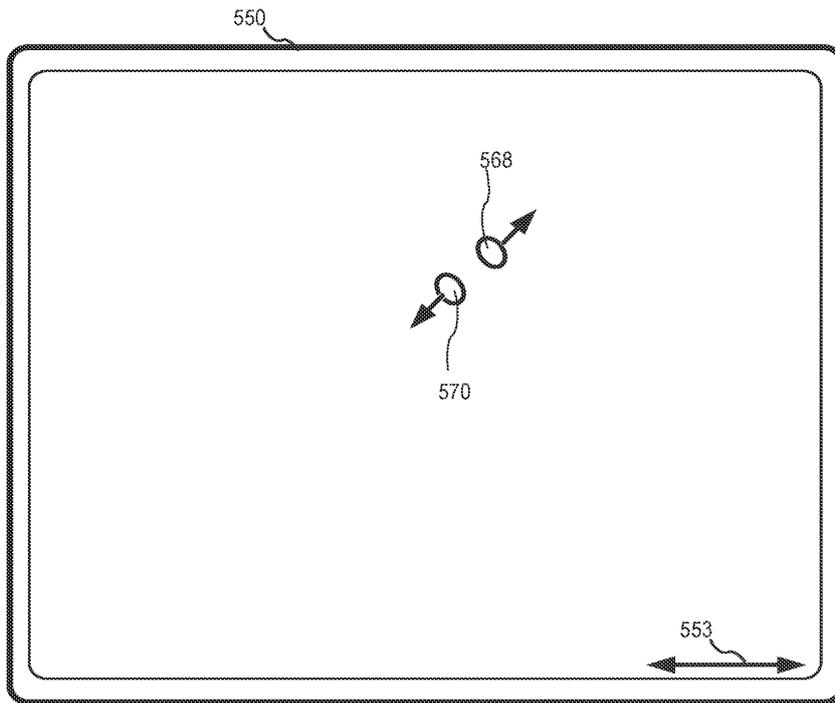
도면4



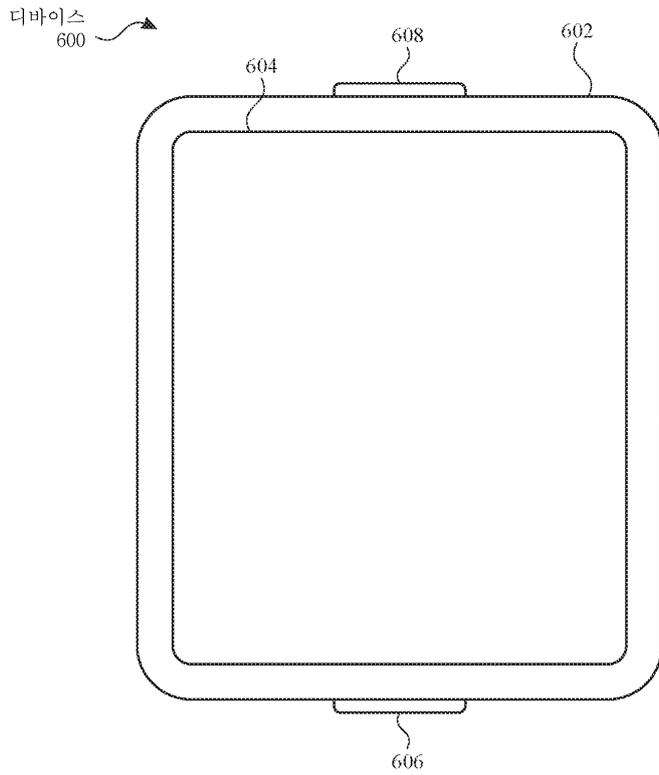
도면5a



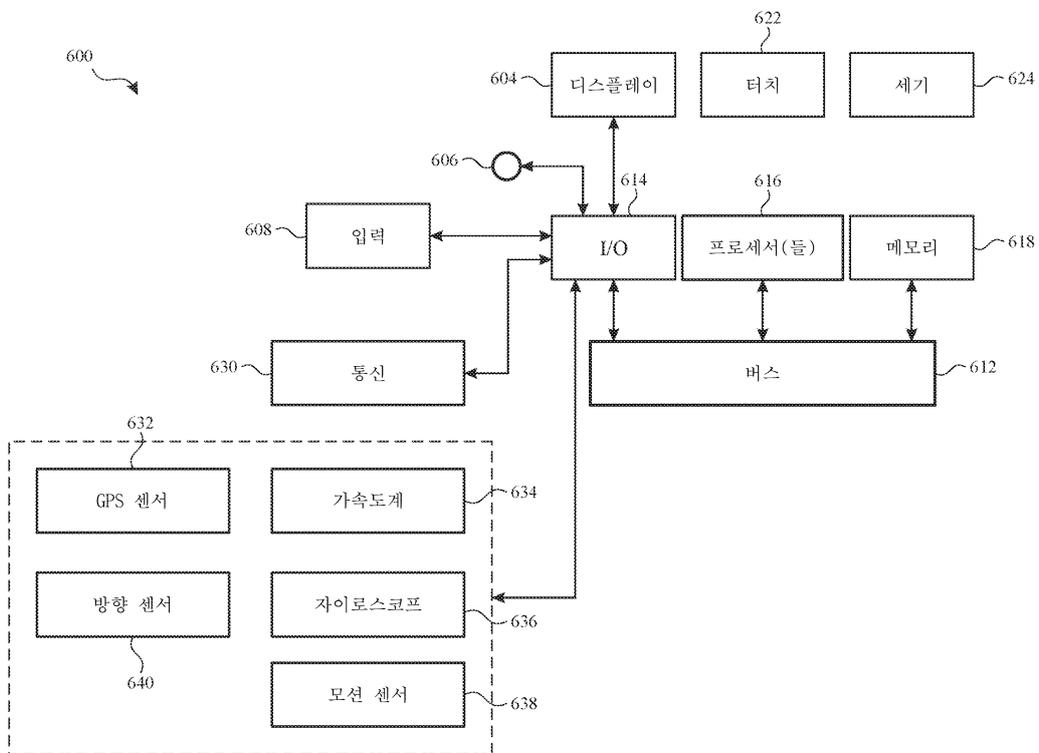
도면5b



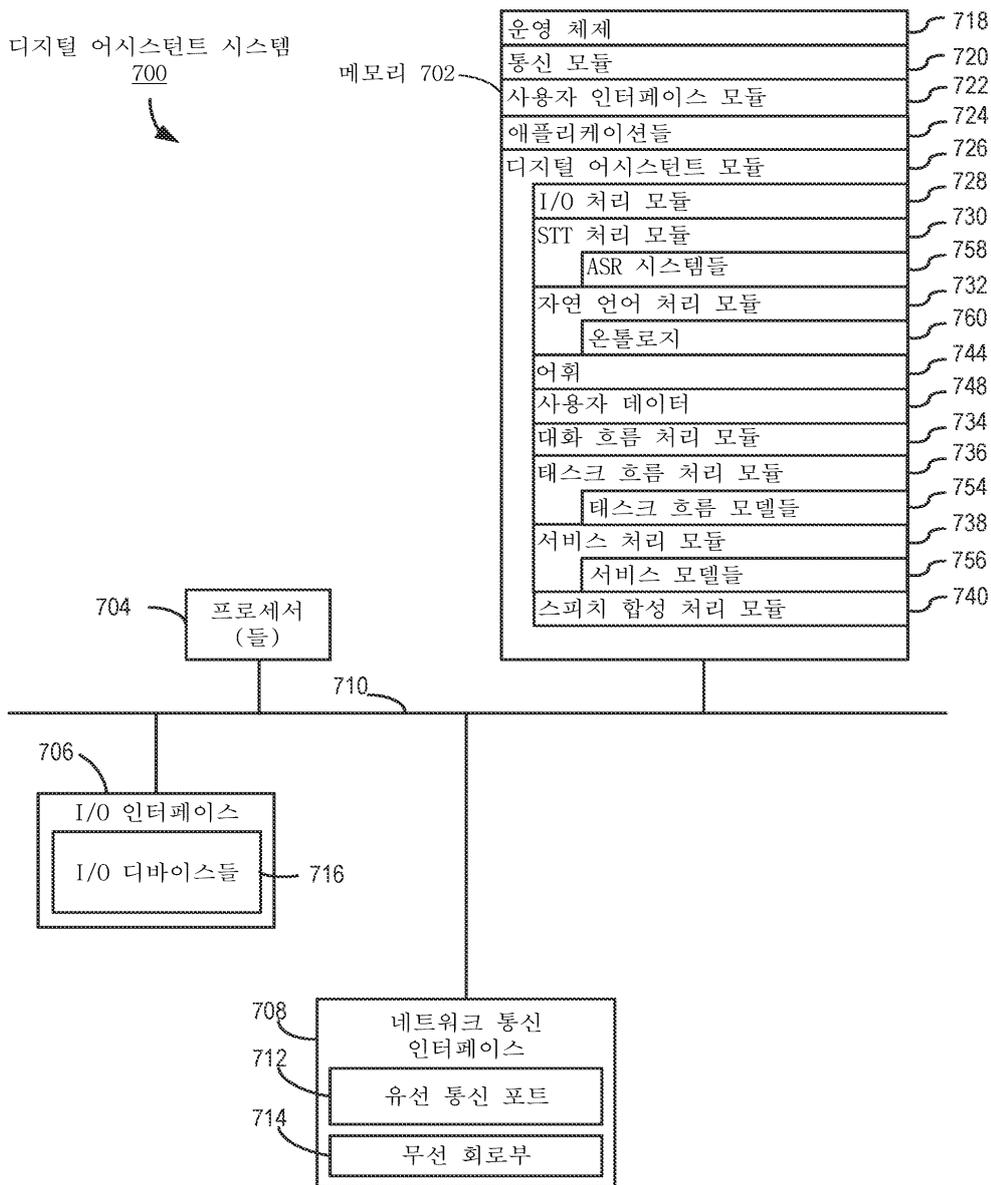
도면6a



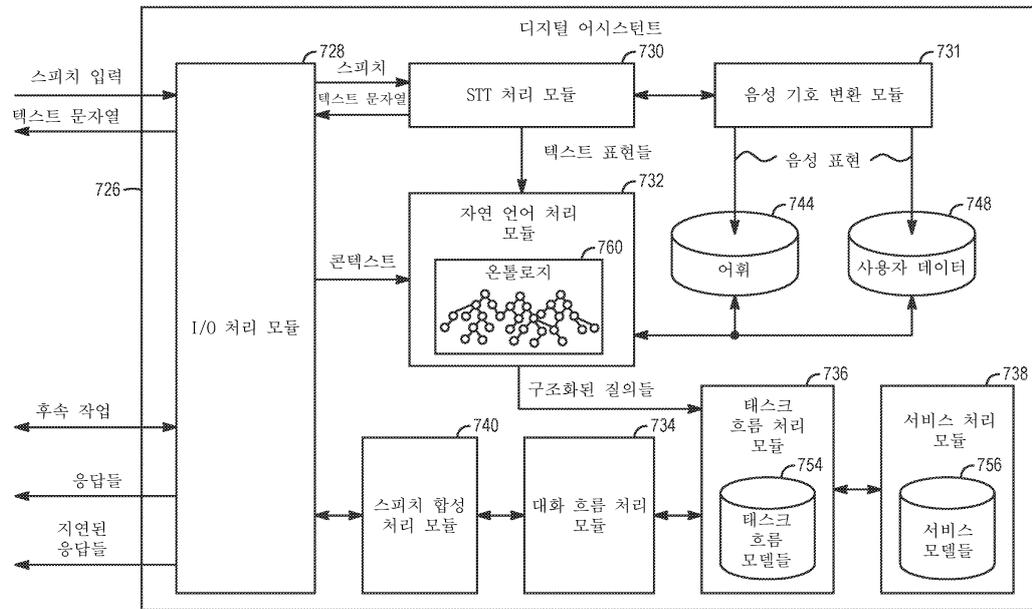
도면6b



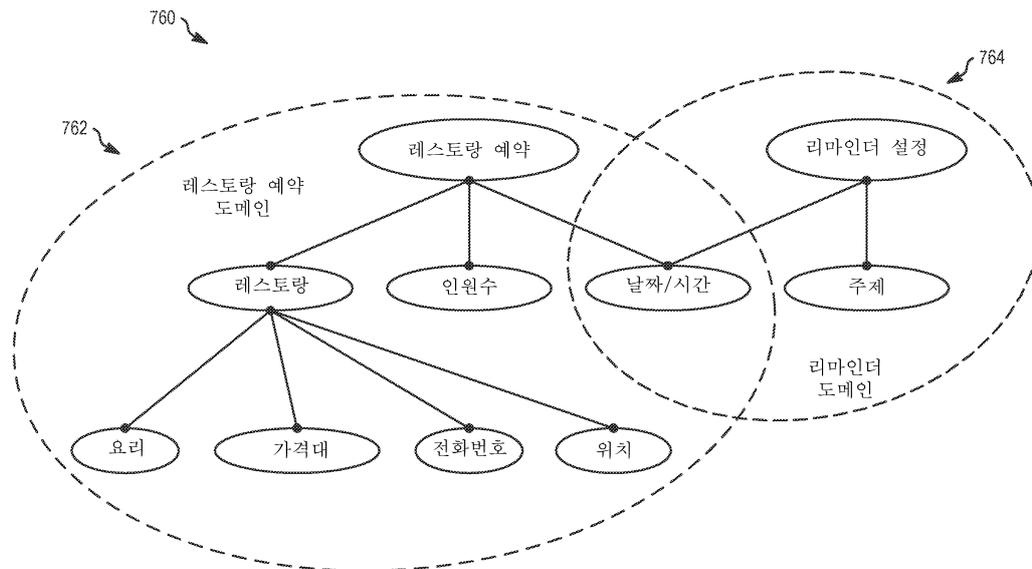
도면7a



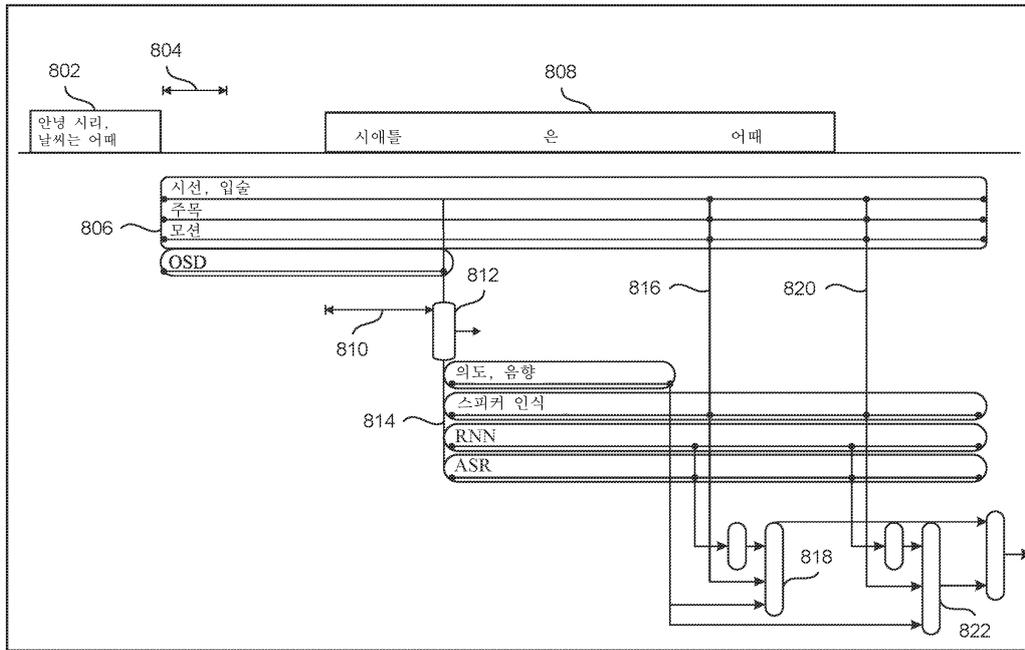
도면7b



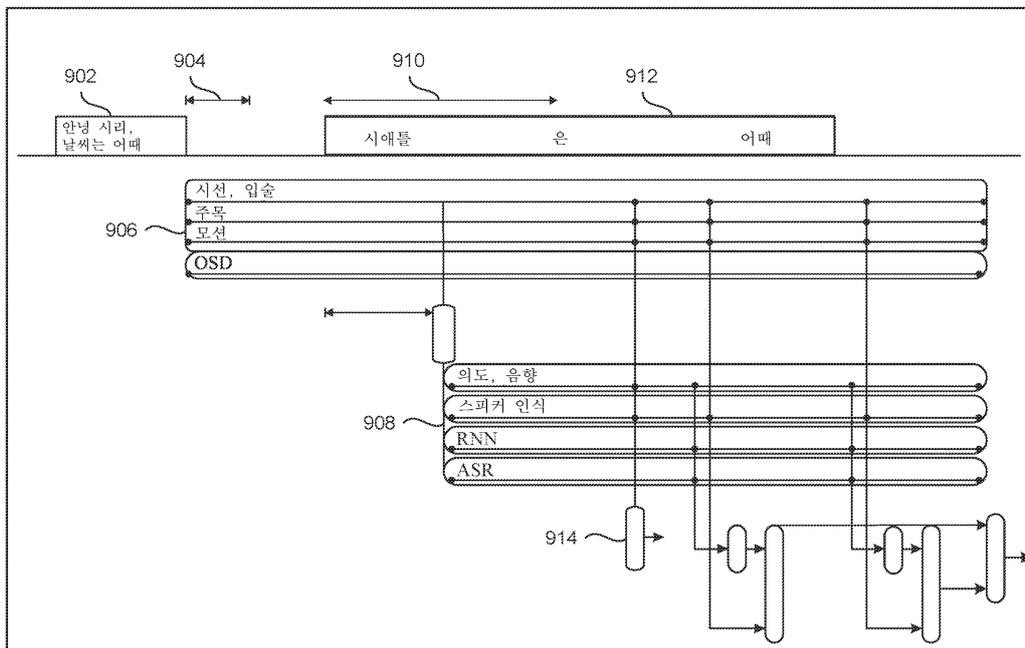
도면7c



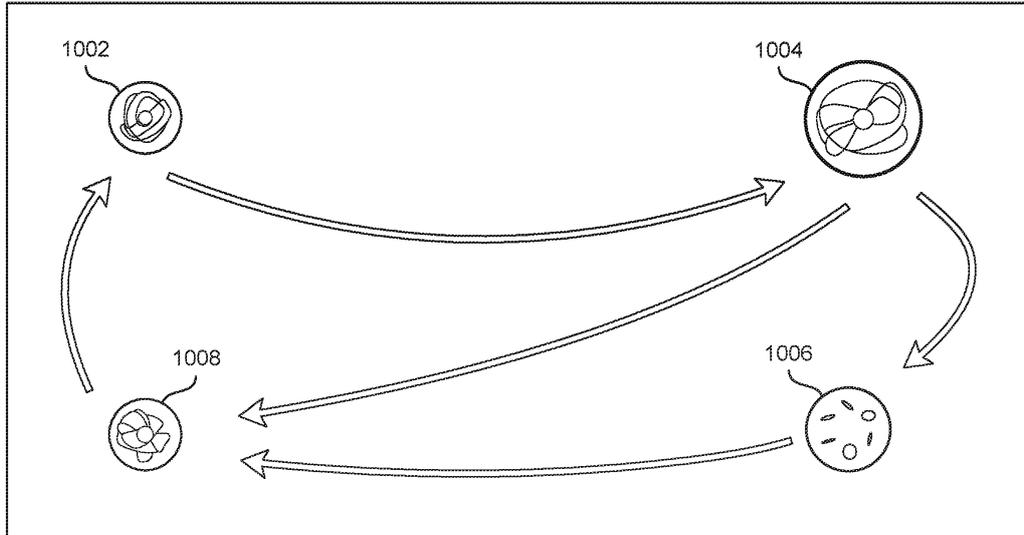
도면8



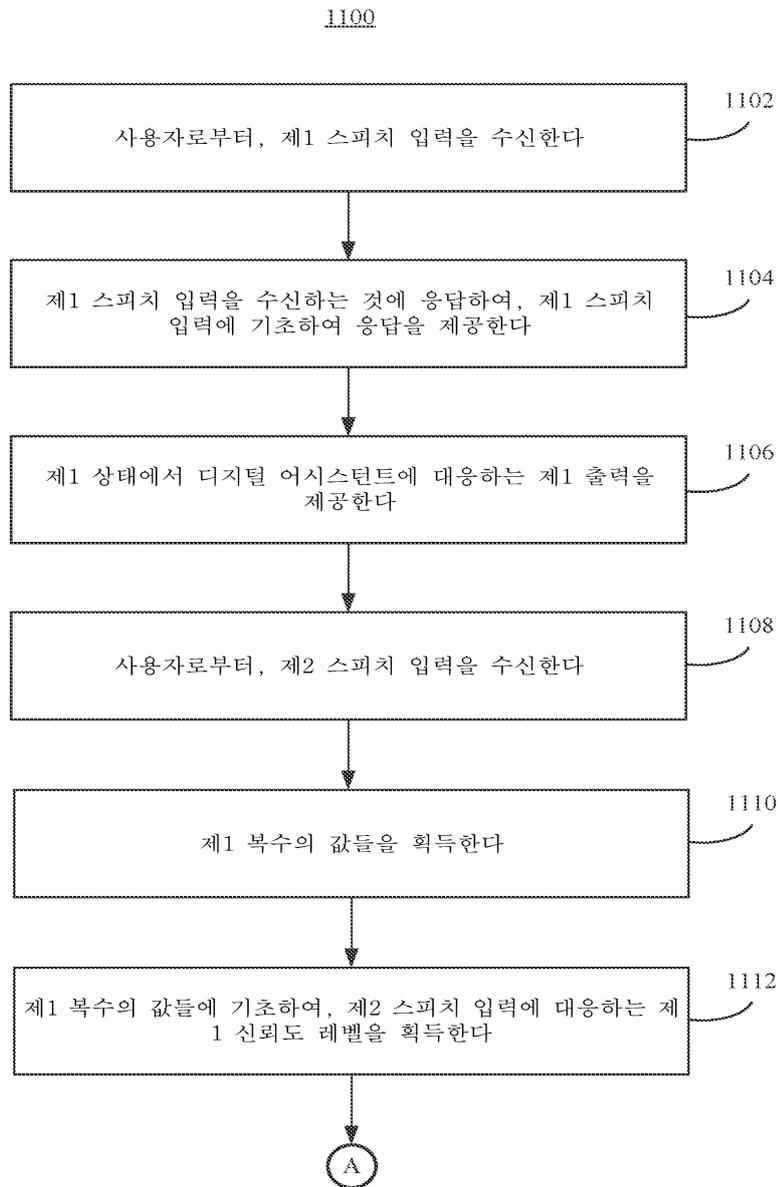
도면9



도면10



도면11a



도면11b

