



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2018년12월20일

(11) 등록번호 10-1931397

(24) 등록일자 2018년12월14일

(51) 국제특허분류(Int. Cl.)

G06F 21/31 (2013.01) G10L 17/00 (2006.01)

(21) 출원번호 10-2013-7026070

(22) 출원일자(국제) 2012년03월20일

심사청구일자 2013년10월02일

(85) 번역문제출일자 2013년10월02일

(65) 공개번호 10-2013-0133026

(43) 공개일자 2013년12월05일

(86) 국제출원번호 PCT/US2012/029810

(87) 국제공개번호 WO 2012/129231

국제공개일자 2012년09월27일

(30) 우선권주장

13/053,144 2011년03월21일 미국(US)

(56) 선행기술조사문헌

KR1020050090568 A*

KR1020020057262 A*

*는 심사관에 의하여 인용된 문헌

(73) 특허권자

애플 인크.

미국 캘리포니아 (우편번호 95014) 쿠퍼티노 원
애플 파크 웨이

(72) 발명자

체이어, 아담, 제이.

미국 95014 캘리포니아주 쿠퍼티노 메일 스탑
302-4페이지 인피니트 루프 1

(74) 대리인

권성락, 백만기, 양영준

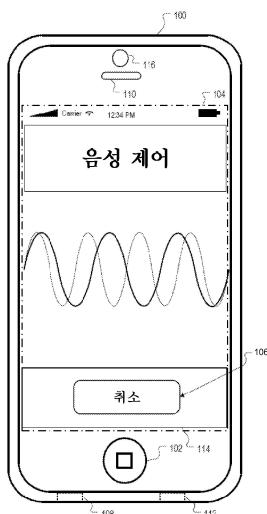
전체 청구항 수 : 총 20 항

심사관 : 서광훈

(54) 발명의 명칭 음성 인증을 사용하는 디바이스 액세스

(57) 요약

디바이스는 사용자로부터 스피치 입력을 수신하도록 구성될 수 있다. 스피치 입력은 디바이스의 제한된 특징에 액세스하기 위한 커맨드를 포함할 수 있다. 스피치 입력은 디바이스에 대해 사용자를 인증하기 위해 사용자의 음성의 성문(예를 들어, 텍스트-독립적 성문)과 비교될 수 있다. 디바이스에 대한 사용자의 성공적인 인증에 응답하여, 사용자가 다시 커맨드를 발성하거나 또는 추가적인 인증 단계들을 수행해야 할 필요 없이 제한된 특징으로의 액세스가 사용자에게 허용된다. 사용자가 디바이스에 대해 성공적으로 인증되지 않는 경우, 추가적인 인증 단계들이 디바이스에 의해 요청될 수 있다(예를 들어, 패스워드를 요청한다).

대 표 도 - 도1

명세서

청구범위

청구항 1

방법으로서,

디바이스가 잠금 상태에 있는 동안:

상기 디바이스에서 스피치 입력을 수신하는 단계 - 상기 스피치 입력은 상기 디바이스의 제한된 특징과 연관된 커맨드를 포함함 -;

상기 스피치 입력이 상기 디바이스의 허가된 사용자에 의해 발성되었는지 여부를 결정하는 단계 - 상기 결정하는 단계는 상기 허가된 사용자의 성문과 상기 스피치 입력을 비교하는 단계를 포함함 -;

상기 허가된 사용자의 성문과 상기 스피치 입력을 비교하는 동안, 상기 커맨드를 식별하기 위해 상기 스피치 입력을 프로세싱하는 단계; 및

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정되면, 상기 스피치 입력에 포함된 상기 커맨드에 따라 상기 디바이스의 상기 제한된 특징으로의 액세스를 제공하는 단계

를 포함하고,

상기 방법은 상기 디바이스의 하나 이상의 프로세서들에 의해 수행되는 방법.

청구항 2

제1항에 있어서,

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되지 않았다고 결정되면, 상기 디바이스의 상기 제한된 특징으로의 액세스를 거절하는 단계를 더 포함하는 방법.

청구항 3

제1항에 있어서,

상기 디바이스의 네트워크 인터페이스를 통해 성문 서비스로부터 상기 성문을 수신하는 단계를 더 포함하는 방법.

청구항 4

제1항에 있어서,

상기 디바이스는 핸드헬드 모바일 디바이스인 방법.

청구항 5

제1항에 있어서,

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정되면, 상기 디바이스를 잠금해제(unlocking)하는 단계;

상기 디바이스에서 후속의 스피치 입력을 수신하는 단계 - 상기 후속의 스피치 입력은 상기 디바이스의 상기 제한된 특징과 연관된 후속의 커맨드를 포함함 -; 및

상기 스피치 입력이 상기 디바이스의 상기 허가된 사용자에 의해 발성되었는지 여부를 결정하지 않고, 상기 후속의 커맨드에 따라 상기 디바이스의 상기 제한된 특징으로의 액세스를 제공하는 단계

를 더 포함하는 방법.

청구항 6

명령어들의 하나 이상의 시퀀스를 포함하는 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체로서,

상기 명령어들은, 전자 디바이스의 하나 이상의 프로세서에 의해 실행되는 경우, 상기 전자 디바이스로 하여금, 상기 전자 디바이스가 잠금 상태에 있는 동안:

스피치 입력을 수신하고 - 상기 스피치 입력은 상기 전자 디바이스의 제한된 특징과 연관된 커맨드를 포함함 -;

상기 스피치 입력이 상기 전자 디바이스의 허가된 사용자에 의해 발성되었는지 여부를 결정하고 - 상기 결정하는 것은 상기 허가된 사용자의 성문과 상기 스피치 입력을 비교하는 것을 포함함 -;

상기 허가된 사용자의 성문과 상기 스피치 입력을 비교하는 동안, 상기 커맨드를 식별하기 위해 상기 스피치 입력을 프로세싱하며;

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정되면, 상기 스피치 입력에 포함된 상기 커맨드에 따라 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 제공하게 하는, 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

청구항 7

제6항에 있어서,

상기 전자 디바이스의 하나 이상의 프로세서에 의해 실행되는 경우, 상기 전자 디바이스로 하여금, 상기 스피치 입력이 상기 허가된 사용자에 의해 발성되지 않았다고 결정되면, 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 거절하게 하는 명령어들을 더 포함하는, 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

청구항 8

제6항에 있어서,

상기 전자 디바이스의 하나 이상의 프로세서에 의해 실행되는 경우, 상기 전자 디바이스로 하여금, 상기 전자 디바이스의 네트워크 인터페이스를 통해 성문 서비스로부터 상기 성문을 수신하게 하는 명령어들을 더 포함하는, 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

청구항 9

제6항에 있어서,

상기 전자 디바이스는 핸드헬드 모바일 디바이스인, 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

청구항 10

제6항에 있어서,

상기 전자 디바이스의 하나 이상의 프로세서에 의해 실행되는 경우, 상기 전자 디바이스로 하여금,

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정되면, 상기 전자 디바이스를 잠금해제하고;

상기 전자 디바이스에서 후속의 스피치 입력을 수신하며 - 상기 후속의 스피치 입력은 상기 전자 디바이스의 상기 제한된 특징과 연관된 후속의 커맨드를 포함함 -;

상기 스피치 입력이 상기 전자 디바이스의 상기 허가된 사용자에 의해 발성되었는지 여부를 결정하지 않고, 상기 후속의 커맨드에 따라 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 제공하게 하는

명령어들을 더 포함하는, 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

청구항 11

전자 디바이스로서,

적어도 하나의 프로세서; 및

성문, 및 상기 적어도 하나의 프로세서에 의해 실행하기 위한 하나 이상의 프로그램들을 저장하는 메모리

를 포함하고,

상기 하나 이상의 프로그램은,

상기 전자 디바이스가 잠금 상태에 있는 동안:

스피치 입력을 수신하고 - 상기 스피치 입력은 상기 전자 디바이스의 제한된 특징과 연관된 커맨드를 포함함 -,

상기 스피치 입력이 상기 전자 디바이스의 허가된 사용자에 의해 발성되었는지 여부를 결정하고 - 상기 결정하는 것은 상기 허가된 사용자의 성문과 상기 스피치 입력을 비교하는 것을 포함함 -,

상기 허가된 사용자의 성문과 상기 스피치 입력을 비교하는 동안, 상기 커맨드를 식별하기 위해 상기 스피치 입력을 프로세싱하며,

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정되면, 상기 스피치 입력에 포함된 상기 커맨드에 따라 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 제공하기 위한 명령어들을 포함하는 전자 디바이스.

청구항 12

제11항에 있어서,

상기 하나 이상의 프로그램은, 상기 스피치 입력이 상기 허가된 사용자에 의해 발성되지 않았다고 결정되면, 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 거절하기 위한 명령어들을 더 포함하는 전자 디바이스.

청구항 13

제11항에 있어서,

상기 하나 이상의 프로그램은, 상기 전자 디바이스의 네트워크 인터페이스를 통해 성문 서비스로부터 상기 성문을 수신하기 위한 명령어들을 더 포함하는 전자 디바이스.

청구항 14

제11항에 있어서,

상기 전자 디바이스는 핸드헬드 모바일 디바이스인 전자 디바이스.

청구항 15

제11항에 있어서,

상기 하나 이상의 프로그램은,

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정되면, 상기 전자 디바이스를 잠금해제하고,

상기 전자 디바이스에서 후속의 스피치 입력을 수신하며 - 상기 후속의 스피치 입력은 상기 전자 디바이스의 상기 제한된 특징과 연관된 후속의 커맨드를 포함함 -,

상기 스피치 입력이 상기 전자 디바이스의 상기 허가된 사용자에 의해 발성되었는지 여부를 결정하지 않고, 상기 후속의 커맨드에 따라 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 제공하기 위한

명령어들을 더 포함하는 전자 디바이스.

청구항 16

전자 디바이스로서,

상기 전자 디바이스가 잠금 상태에 있는 동안:

스피치 입력을 수신하기 위한 수단 - 상기 스피치 입력은 상기 전자 디바이스의 제한된 특징과 연관된 커맨드를 포함함 -;

상기 전자 디바이스의 허가된 사용자의 성문과 상기 스피치 입력을 비교하기 위한 수단;

상기 비교의 결과에 기초하여, 상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정하기 위한 수단; 및

상기 커맨드에 따라 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 제공하기 위한 수단을 포함하는 전자 디바이스.

청구항 17

제16항에 있어서,

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되지 않았다고 결정되면, 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 거절하기 위한 수단을 더 포함하는 전자 디바이스.

청구항 18

제16항에 있어서,

상기 전자 디바이스의 네트워크 인터페이스를 통해 성문 서비스로부터 상기 성문을 수신하기 위한 수단을 더 포함하는 전자 디바이스.

청구항 19

제16항에 있어서,

상기 전자 디바이스는 핸드헬드 모바일 디바이스인 전자 디바이스.

청구항 20

제16항에 있어서,

상기 스피치 입력이 상기 허가된 사용자에 의해 발성되었다고 결정되면, 상기 전자 디바이스를 잠금해제하기 위한 수단;

상기 전자 디바이스에서 후속의 스피치 입력을 수신하기 위한 수단 - 상기 후속의 스피치 입력은 상기 전자 디바이스의 상기 제한된 특징과 연관된 후속의 커맨드를 포함함 -; 및

상기 스피치 입력이 상기 전자 디바이스의 상기 허가된 사용자에 의해 발성되었는지 여부를 결정하지 않고, 상기 후속의 커맨드에 따라 상기 전자 디바이스의 상기 제한된 특징으로의 액세스를 제공하기 위한 수단

을 더 포함하는 전자 디바이스.

청구항 21

삭제

청구항 22

삭제

청구항 23

삭제

청구항 24

삭제

청구항 25

삭제

청구항 26

삭제

청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

발명의 설명

기술 분야

[0001]

본 개시내용은 일반적으로 전자 디바이스의 특징들에 대한 사용자 액세스를 제어하기 위한 기법들에 관한 것이다.

배경기술

[0002]

오늘날의 컴퓨터들 및 다른 전자 디바이스들 중 다수는 사용자가 타인들에 의한 액세스로부터 컴퓨터 또는 디바이스를 잠글 수 있게 하는 특징을 제공한다. 디바이스들 중 일부는 디바이스의 그래픽 사용자 인터페이스를 통해 잠긴 디바이스를 잠금 해제하기 위한 메커니즘을 제공한다. 예를 들어, 그래픽 사용자 인터페이스는 사용자가 패스워드 또는 코드와 같은 인증 정보를 입력할 수 있게 하는 메커니즘을 제공할 수 있다.

[0003]

일부 컴퓨터들 및 다른 전자 디바이스들은 음성 커맨드 특징들을 제공할 수 있다. 예를 들어, 디바이스의 사용자는 디바이스에 연결된 마이크로폰에 대고 음성 커맨드를 발성할 수 있다. 음성 커맨드가 디바이스에 의해 수신되는 경우, 디바이스는 음성 커맨드를 인식하고 실행할 수 있다.

발명의 내용

[0004]

디바이스는 사용자로부터 스피치 입력을 수신하도록 구성될 수 있다. 스피치 입력은 디바이스의 제한된 특징에 액세스하기 위한 커맨드를 포함할 수 있다. 스피치 입력은 디바이스에 대해 사용자를 인증하기 위해 사용자의 음성의 성문(예를 들어, 텍스트-독립적 성문)과 비교될 수 있다. 디바이스에 대한 사용자의 성공적인 인증에 응답하여, 사용자가 다시 커맨드를 발성하거나 추가적인 인증 단계들을 수행해야 할 필요 없이, 제한된 특징으로의 액세스가 사용자에게 허용된다. 사용자가 디바이스에 대해 성공적으로 인증되지 않는 경우, 추가적인 인증 단계들이 디바이스에 의해 요청될 수 있다(예를 들어, 패스워드를 요청한다).

[0005]

일부 구현예들에서, 디바이스의 허가된 사용자에 대해 성문이 생성될 수 있다. 예를 들어, 사용자가 디바이스에 대고 음성 커맨드들을 발성할 시에 사용자의 음성의 하나 이상의 샘플들이 수집될 수 있다. 하나 이상의 음성 샘플들에 기초하여 성문이 생성될 수 있다. 성문은 디바이스에서 로컬로 또는 네트워크 성문 서비스(예를 들어, 네트워크 서버)에 의해 생성될 수 있다. 성문은 디바이스에 대해 사용자를 인증하기 위해 디바이스 상에서 실행하는 또는 네트워크 서비스에 의해 호스팅되는 텍스트-독립적 음성 인증 프로세스와 함께 사용될 수 있다.

[0006]

이 명세서에서 기술된 발명 대상의 특정 실시예들은 후속하는 장점들 중 하나 이상을 달성하도록 구현될 수 있다. 디바이스는 잠긴 디바이스에 액세스하기 위한 더욱 사용자-친화적인 인증 프로세스를 포함할 수 있다. 사용자의 음성이 음성 커맨드가 프로세싱되는 동시에 인증될 수 있으며, 어떠한 별도의 인증 단계도 요구되지 않는다. 디바이스는 사용자가 디바이스에 대고 음성 커맨드들을 발성하는 동안 성문을 생성할 수 있고; 어떠한 별도의 화자 인식 트레이닝 단계도 요구되지 않는다. 하기에 개시되는 음성 인증 특징들은 디바이스의 임의의/

모든 특징들로의 빠르고 안전한 음성 제어 액세스를 제공할 수 있다.

[0007] 일부 실시예들에 따라, 방법은 디바이스에서 스피치 입력을 수신하는 것을 포함한다. 스피치 입력은 디바이스의 제한된 특징과 연관된 커맨드를 포함한다. 방법은 또한 디바이스의 허가된 사용자의 성문과 스피치 입력을 비교하는 것, 및 비교의 결과에 기초하여 스피치 입력이 허가된 사용자에 의해 발성되었다고 결정하는 것을 포함한다. 방법은 커맨드에 따라 디바이스의 제한된 특징으로의 액세스를 제공하는 것을 더 포함한다. 방법은 디바이스의 하나 이상의 프로세서에 의해 수행된다.

[0008] 일부 실시예들에 따라, 방법은 디바이스에서 스피치 입력을 수신하는 것을 포함한다. 스피치 입력은 디바이스의 특징과 연관된 커맨드를 포함한다. 방법은 또한 스피치 입력에 기초하여 텍스트-독립적 성문을 생성하는 것; 및 커맨드에 따라 디바이스의 특징으로의 액세스를 제공하는 것을 포함한다. 방법은 디바이스의 하나 이상의 프로세서에 의해 수행된다.

[0009] 일부 실시예들에 따라, 방법은 디바이스에서 스피치 입력을 수신하는 것 - 스피치 입력은 디바이스의 특징과 연관된 커맨드를 포함함 - 및 스피치 입력에 기초하여 음성 샘플을 생성하는 것을 포함한다. 방법은 또한 음성 샘플에 기초하여 성문을 생성하기 위해 성문 서비스에 음성 샘플을 전송하는 것; 및 커맨드에 따라 디바이스의 특징으로의 액세스를 제공하는 것을 포함한다. 방법은 디바이스의 하나 이상의 프로세서에 의해 수행된다.

[0010] 일부 실시예들에 따라, 전자 디바이스는 하나 이상의 프로세서, 메모리, 및 하나 이상의 프로그램을 포함하고; 하나 이상의 프로그램은 메모리에 저장되고, 하나 이상의 프로세서에 의해 실행되도록 구성되고, 하나 이상의 프로그램은 전술된 방법들 중 임의의 것의 동작들을 수행하기 위한 명령들을 포함한다. 일부 실시예들에 따라, 컴퓨터 관독 가능한 저장 매체는, 전자 디바이스에 의해 실행되는 경우, 디바이스로 하여금 전술된 방법들 중 임의의 것의 동작을 실행하게 하는 명령들을 저장하고 있다. 일부 실시예들에 따라, 전자 디바이스는 전술된 방법들 중 임의의 것의 동작들을 수행하기 위한 수단을 포함한다. 일부 실시예들에 따라, 전자 디바이스에서 사용하기 위한 정보 프로세싱 장치는 전술된 방법들 중 임의의 것의 동작들을 수행하기 위한 수단을 포함한다.

[0011] 일부 실시예들에 따라, 전자 디바이스는 스피치 입력을 수신하도록 구성되는 스피치 수신 유닛을 포함한다. 스피치 입력은 전자 디바이스의 제한된 특징과 연관된 커맨드를 포함한다. 전자 디바이스는 또한 스피치 수신 유닛에 연결되는 프로세싱 유닛을 포함한다. 프로세싱 유닛은: 전자 디바이스의 허가된 사용자의 성문과 스피치 입력을 비교하고; 비교의 결과에 기초하여, 스피치 입력이 허가된 사용자에 의해 발성되었다고 결정하고; 커맨드에 따라 전자 디바이스의 제한된 특징으로의 액세스를 제공하도록 구성된다.

[0012] 일부 실시예들에 따라, 전자 디바이스는 스피치 입력을 수신하도록 구성되는 스피치 수신 유닛을 포함하고, 스피치 입력은 전자 디바이스의 특징과 연관된 커맨드를 포함한다. 전자 디바이스는 또한 스피치 수신 유닛에 연결되는 프로세싱 유닛을 포함한다. 프로세싱 유닛은: 스피치 입력에 기초하여 텍스트-독립적 성문을 생성하고; 커맨드에 따라 전자 디바이스의 특징으로의 액세스를 제공하도록 구성된다.

[0013] 일부 실시예들에 따라, 전자 디바이스는 스피치 입력을 수신하도록 구성되는 스피치 수신 유닛을 포함한다. 스피치 입력은 전자 디바이스의 특징과 연관된 커맨드를 포함한다. 전자 디바이스는 또한 스피치 수신 유닛에 연결되는 프로세싱 유닛을 포함한다. 프로세싱 유닛은 스피치 입력에 기초하여 음성 샘플을 생성하고; 음성 샘플에 기초하여 성문을 생성하기 위해 성문 서비스에 음성 샘플을 전송하고; 커맨드에 따라 전자 디바이스의 특징으로의 액세스를 제공하도록 구성된다.

[0014] 하나 이상의 구현예들의 상세항목들은 첨부 도면들 및 하기 기재에서 설명된다. 다른 특징들, 양상들 및 잠재적 장점들은 기재 및 도면들로부터, 그리고 청구항들로부터 명백할 것이다.

도면의 간단한 설명

[0015] 도 1은 음성 커맨드들을 프로세싱하도록 구성되는 예시적인 디바이스를 예시하는 도면이다.

도 2는 성문을 생성하기 위한 예시적인 프로세스의 흐름도이다.

도 3은 음성 인증을 위해 구성될 수 있는 예시적인 잠금 디바이스를 예시하는 도면이다.

도 4는 음성 인증을 위한 예시적인 프로세스의 흐름도이다.

도 5는 예시적인 네트워크 동작 환경의 블록도이다.

도 6은 도 1-4의 모바일 디바이스의 예시적인 구현예의 블록도이다.

도 7은 일부 실시예들에 따라 전자 디바이스의 기능 블록도를 예시하는 도면이다.

도 8은 일부 구현예들에 따라 전자 디바이스의 기능 블록도를 예시하는 도면이다.

도 9는 일부 실시예들에 따라 전자 디바이스의 기능 블록도를 예시하는 도면이다.

발명을 실시하기 위한 구체적인 내용

[0016] 음성 커맨드들

도 1은 음성 커맨드들을 프로세싱하도록 구성되는 예시적인 디바이스(100)를 예시한다. 예를 들어, 디바이스(100)는 모바일 디바이스, 예를 들어, 셀 폰, 스마트 폰, 전자 태블릿, 텔레비전 시스템, 개인 데이터 보조 단말(PDA), 랩톱 또는 임의의 다른 모바일 디바이스일 수 있다. 디바이스(100)는 데스크톱 컴퓨터 또는 디바이스(100)에 대해 사용자를 인증하도록 사용자에게 요구할 수 있는 임의의 다른 디바이스일 수 있다. 일부 구현예들에서, 디바이스(100)는 스피치 입력을 수신하고, 스피치 입력에 기초하여 커맨드를 결정하고, 커맨드를 실행할 수 있다. 예를 들어, 사용자는 다운 버튼(102)을 눌러서 유지함으로써 디바이스(100)의 음성 제어 특징을 활성화할 수 있다. 활성화되는 경우, 음성 제어 특징은 예를 들어, 도 1에 디스플레이되는 바와 같이, 터치 감지 디스플레이(104) 상에 음성 제어 그래픽 사용자 인터페이스를 디스플레이할 수 있다. 사용자는 메뉴 바(114)에 디스플레이된 취소 버튼(106)을 누름으로써 음성 제어 특징을 취소할 수 있다.

[0018] 일부 구현예들에서, 음성 제어 특징이 활성화되는 경우, 디바이스(100)는 마이크로폰(108)을 통해 사용자로부터의 스피치 입력을 수신할 수 있다. 일부 구현예들에서, 스피치 입력은 스피치 입력에서 발생된 단어들을 나타내는 텍스트로 변환될 수 있다. 예를 들어, 스피치 입력의 내용을 나타내는 텍스트를 생성하기 위해 스피치 입력에 대해 스피치 인식 분석 또는 모델링(예를 들어, 은닉 마코프 모델링(HMM), 동적 타임 워핑(DTW) 등)이 수행될 수 있다.

[0019] 일부 구현예들에서, 스피치 입력으로부터 생성된 텍스트는 디바이스(100)의 특징을 불러오기 위한 커맨드를 결정하기 위해 분석될 수 있다. 예를 들어, 텍스트가 단어 "호출"을 포함하는 경우, 디바이스(100)는 사용자가 전화 걸기를 원한다고 결정할 수 있고, 텔레포니 애플리케이션을 불러올 수 있다. 텍스트가 단어 "재생"을 포함하는 경우, 디바이스(100)는 사용자가 디바이스(100) 상에 저장된 미디어를 재생하기를 원한다고 결정할 수 있고, 예를 들어, 음악 또는 영화와 같은, 컨텐츠를 재생하기 위해 미디어 플레이어를 불러올 수 있다.

[0020] 일부 구현예들에서, 디바이스(100)의 음성 제어 특징은 커맨드를 결정하기 위해 디바이스(100)의 성공 또는 실패를 표시하기 위한 피드백을 사용자에게 제공할 수 있다. 예를 들어, 피드백(예를 들어, 오디오, 시각, 진동)은 어느 커맨드가 디바이스 상에서 막 실행되려고 하는지, 디바이스(100)가 스피치 입력에 기초하여 커맨드를 결정하는 데 성공하였는지 여부, 및/또는 커맨드가 디바이스(100)에 의해 성공적으로 실행되었는지의 여부를 사용자에게 표시할 수 있다. 예를 들어, 디바이스에 의해 생성된 음성은 어느 커맨드가 디바이스에 의해 막 실행되려고 하는지를 사용자에게 알릴 수 있다.

[0021] 일부 구현예들에서, 디바이스(100)의 음성 제어 특징들은 오직 디바이스가 잠금 해제 상태에 있는 경우(예를 들어, 디바이스에 액세스하는 사용자가 인증된 경우)에만 액세스될 수 있다.

[0022] 성문

도 2는 성문을 생성하기 위한 예시적인 프로세스(200)의 흐름도이다. 일부 구현예들에서, 디바이스(100)는 디바이스(100)에 의해 수신되는 스피치 입력들에 기초하여 사용자에 대한 성문을 생성하도록 구성될 수 있다. 예를 들어, 디바이스(100)는 사용자가 디바이스(100)의 음성 제어 특징들과 상호작용하는 동안 사용자의 음성의 하나 이상의 샘플들을 수집할 수 있다. 일부 구현예들에서, 디바이스(100)는 디바이스(100)에 대해 사용자를 인증하기 위해 텍스트-독립적 음성 인증 프로세스에서 성문을 사용할 수 있다.

[0024] 일부 구현예들에서, 성문을 생성하는 것은 오직 디바이스(100)가 잠금 해제 상태에 있는 경우에만 수행될 수 있다. 예를 들어, 성문을 생성하는 것은 오직 스피치 입력을 제공하는 사용자가 디바이스(100)의 소유자 또는 허가된 사용자로서 디바이스(100)에 대해 인증된 경우에만 수행되어, 허가되지 않은 사용자 또는 침입자의 음성에 기초하여 성문을 생성하는 것을 방지할 수 있다.

[0025] 단계(202)에서, 스피치 입력이 획득된다. 일부 구현예들에서, 디바이스(100)는 디바이스(100)에 연결된 마이크로폰(102)을 통해 스피치 입력을 수신하도록 구성될 수 있다. 마이크로폰(102)은 스피치 입력으로부터 오디오 데이터를 생성할 수 있다. 일부 구현예들에서, 디바이스(100)는 오디오 데이터로부터 하나 이상의 음성 샘플들

을 수집하고, 원격 성문 서비스에 음성 샘플들을 전송하도록 구성될 수 있다.

[0026] 단계(204)에서, 성문이 생성된다. 예를 들어, 하나 이상의 음성 샘플들은 사용자의 발성 패턴들의 거동 및 사용자의 성도(vocal tract)에 관한 고유한 정보에 기초하여 디바이스(100)의 허가된 사용자의 성문을 생성하도록 분석 및/또는 모델링될 수 있다. 일부 구현예들에서, 성문이 디바이스(100)에서 생성될 수 있다. 예를 들어, 오디오 데이터는 화자 인증 동안 허가된 사용자의 음성을 인식하기 위해 사용될 수 있는 성문을 생성하도록 디바이스(100)에 의해 프로세싱될 수 있다. 일부 구현예들에서, 성문이 원격 또는 네트워킹된 서비스에서 생성될 수 있다. 예를 들어, 디바이스(100)는 오디오 데이터로부터 하나 이상의 음성 샘플들을 수집하고, 도 5의 성문 서비스(508)에 음성 샘플들을 전송하도록 구성될 수 있다. 예를 들어, 음성 샘플들은 다수의 스피치 입력들로부터 시간 경과에 따라 수집될 수 있고, 음성 샘플들은 성문 서비스(508)에 여러 블록으로(in batches) 전송될 수 있다. 음성 샘플 블록들은 예를 들어, 디바이스(100)가 유휴 상태이거나 또는 낮은 자원 사용을 겪고 있는 기간들 동안 성문 서비스(508)에 전송될 수 있다. 성문 서비스(508)는 디바이스(100)로부터 수신되는 샘플들에 기초하여 성문(예를 들어, 텍스트-독립적 성문)을 생성하도록 구성될 수 있다. 성문 서비스(508)는 화자 인식 분석을 사용하여 사용자를 인증하는 경우 디바이스(100)에 의해 사용되도록 디바이스(100)에 생성된 성문을 전송할 수 있다.

[0027] 일부 구현예들에서, 디바이스(100), 또는 원격 성문 서비스(508)는 텍스트-독립적 방식으로 사람의 음성의 "시그너처(signature)" 또는 "무늬(print)"를 학습할 수 있는 성문 모듈을 포함할 수 있다. 예를 들어, 상이한 사용자의 음성들의 음성 특성들을 구별하기 위해 다양한 음소들의 사용자의 발음에 존재하는 스펙트럼 특징들의 특성들의 통계적 모델들이 구축될 수 있다. 예를 들어, 벡터 양자화(VQ) 코드북-기반 기법들이 성문을 생성하는 데 사용될 수 있다. 예를 들어, 발성, 침묵, 중단 버스트(stop burst), 비음/유음(liquid), 협착적 기식음(frication) 등과 같은 음성 특성들의 학습된 모델들을 구축하기 위해 상태들 사이의 추계적 마코프 천이들(stochastic Marchovian transitions)을 분석하는 에르고딕-HMM-기반 방법들이 성문을 생성하는 데 사용될 수 있다. 일부 구현예들에서, 먼저 명시적으로 스피치 입력으로부터의 오디오 데이터로부터 음소들 또는 음소 부류들을 결정하고, 이후 각각의 인식된 음소 카테고리에 대한 매치들의 가중된 조합에 의해 화자 검증을 수행하는 2-패스 화자 인식 방식이 사용될 수 있다.

[0028] 전술된 텍스트-독립적 화자 인증 프로세스들은 음성 인증을 위해 특정 단어 또는 특정 패스프레이즈(passphrase)를 요구하지 않고 음성 인증을 제공할 수 있다. 반면, 텍스트-종속적 화자 검증 프로세스들은 종종 화자 인식을 수행하기 위해 특정 패스프레이즈들 또는 단어 발성들을 요구하고, 따라서 종종 사용자에게 특정 단어 또는 구문을 발성할 것을 요구하는 별도의 인증 단계(예를 들어, 도전-응답 단계)를 요구한다. 텍스트-독립적 인증 프로세스들은 별도의 도전-응답 인증 단계를 요구하지 않는다.

[0029] 일부 구현예들에서, 성문이 생성되면, 성문은 디바이스(100)에 저장될 수 있다. 예를 들어, 디바이스(100)가 성문을 생성하는 경우, 성문은 디바이스(100)에 연결된 메모리 또는 비휘발성 저장소(예를 들어, 하드 드라이브)에 저장될 수 있다. 성문이 네트워크 서버에 의해(예를 들어, 성문 서비스(508)에 의해) 생성되는 경우, 디바이스(100)는 네트워크 생성된 성문을 수신하고, 성문을 메모리 또는 비휘발성 저장소에 저장할 수 있다. 네트워크 서버는 또한 그것이 생성한 성문들을 저장할 수 있다.

[0030] 단계(206)에서, 스피치 입력에 기초하여 커맨드가 결정된다. 일부 구현예들에서, 스피치 입력은 성문에 대응하는 커맨드를 결정하기 위해 프로세싱될 수 있다. 예를 들어, 스피치 입력은 스피치-대-텍스트 프로세싱을 사용하여 텍스트로 변환될 수 있고, 텍스트는 스피치 인식 프로세싱을 사용하여 커맨드를 식별하기 위해 분석될 수 있다. 예를 들어, 스피치 입력이 텍스트로 변환되면, 스피치 입력의 텍스트는, 스피치 입력 텍스트 중 임의의 것이 커맨드 텍스트에 대응(예를 들어, 매치)하는지의 여부를 결정하기 위해 디바이스(100)에 알려진 커맨드들과 연관된 텍스트와 비교될 수 있다. 스피치 입력에서 전체적으로 또는 부분적으로 텍스트적 대응성이 발견되는 경우, 디바이스(100)는 스피치 입력 텍스트에 대응하는 커맨드 텍스트에 대응하는 커맨드를 실행할 수 있다.

[0031] 일부 구현예들에서, 커맨드는 성문이 생성되는 동안 결정될 수 있다. 예를 들어, 스피치 입력이 디바이스(100)에 의해 수신되면, 스피치 입력은 성문을 생성하고 음성 커맨드를 결정하기 위해 프로세싱될 수 있다(예를 들어, 병렬로 프로세싱될 수 있다). 따라서, 단일의 스피치 입력이 성문을 생성하고 음성 커맨드를 발행하기 위해 사용될 수 있다.

[0032] 단계(208)에서, 커맨드가 실행된다. 예를 들어, 커맨드가 스피치 입력에 기초하여 결정되는 경우, 커맨드는 디바이스(400)에 의해 실행될 수 있다.

[0033] 보안 특징들

도 3은 음성 인증을 위해 구성될 수 있는 예시적인 잠긴 디바이스(100)를 예시한다. 예를 들어, 디바이스(100)는 디바이스(100) 상에 저장된 정보 또는 특징들(예를 들어, 전체 디바이스, 개별 애플리케이션들 등)로의 허가되지 않은 액세스를 방지하기 위해 (예를 들어, 사용자의 인증을 요구하는 상태로) 잠길 수 있다. 일부 구현 예들에서, 디바이스(100)의 개별 특징들이 잠길 수 있다. 예를 들어, 디바이스(100)의 개별 특징들은 디바이스(100)가 그 특징들로의 액세스를 허용하기 전에 사용자의 인증을 요구할 수 있다. 디바이스에 액세스하는 사용자가 디바이스의 허가된 사용자 또는 소유자임을 보증하기 위해 디바이스에 의해 사용자의 인증이 요구될 수 있다. 일부 구현 예들에서, 디바이스(100)는 디바이스(100)의 개별 특징들 또는 디바이스(100)로의 액세스를 승인하기 전에 사용자가 디바이스(100)의 허가된 사용자임을 인증하도록 사용자에게 요구할 수 있다. 예를 들어, 터치 감지 디스플레이(104)는 사용자가 디바이스(100)를 잠금 해제하기 위한 패스코드를 입력하게 하는 사용자 인터페이스를 디스플레이 할 수 있다. 사용자는 디바이스(100)가 잠금 해제되게 하기 위해 터치 감지 키 패드(302)를 사용하여 패스코드(예를 들어, 4 자리 숫자, 워드, 문자들의 시퀀스)를 입력할 수 있다. 다른 사용자 인증 및 디바이스 잠금 해제 메커니즘(예를 들어, 음성 인증, 얼굴 인식, 지문 인식)이 또한 가능하다.

[0035] 일부 구현 예들에서, 인증되지 않은 사용자(예를 들어, 아직 인증되지 않은 사용자)가 디바이스(100)에 입력을 제공하거나 디바이스의 특징들에 액세스하려고 시도하는 경우, 사용자의 인증이 수행될 수 있다. 예를 들어, 사용자가 패스워드 잠긴 디바이스 상에서 전화를 걸거나, 이메일 애플리케이션, 주소록 또는 캘린더에 액세스하려고 시도하는 경우, 도 3의 사용자 인터페이스가 사용자에게 제시되어 사용자로 하여금 패스워드, 코드 또는 다른 사용자 인증 입력을 입력하게 할 수 있다. 일부 구현 예들에서, 사용자가 디바이스(100)에 알려진 패스워드 또는 코드를 입력하는 경우, 사용자는 인증될 수 있고, 디바이스(100) 및/또는 디바이스(100)의 특징들은 잠금 해제될 수 있다. 사용자가 디바이스(100)에 알려지지 않은 패스워드 또는 코드를 입력하는 경우, 사용자는 인증될 수 없고, 디바이스(100) 및/또는 디바이스(100)의 특징들은 잠긴 채 유지될 수 있다. 일부 구현 예들에서, 디바이스(100)는 도 4를 참조하여 기술된 바와 같이, 사용자의 음성 인증을 수행하도록 구성될 수 있다.

[0036] 음성 인증

[0037] 도 4는 음성 인증을 위한 예시적인 프로세스(400)의 흐름도이다. 예를 들어, 사용자의 음성 인증은 잠긴 디바이스에서 스피치 입력이 수신되는 경우 스피치 입력에 대해 화자 인식 분석을 수행함으로써 수행될 수 있다. 전술된 바와 같이, 사용자의 인증은 텍스트-독립적 음성 인증 기법들을 사용하여 수행될 수 있다.

[0038] 여기서 기술된 음성 인식 특징들은 디바이스(100)의 모든 특징들 및 디바이스(100) 상에 저장된 데이터로의 빠르고 안전한 액세스를 허용할 수 있다. 예를 들어, 이들 음성 인증 특징들은 디바이스(100)의 사용자로 하여금 사용자가 디바이스(100)에 액세스하려고 시도할 때마다 패스코드를 입력해야 할 필요 없이 안전한 방식으로 디바이스(100) 상의 특징들 및 정보에 액세스하는 것을 가능하게 할 수 있다. 예를 들어, 이들 음성 인증 특징들이 없다면, 디바이스로의 사용자 액세스는 별도의 인증 단계들에 의해 느려질 수 있고, 디바이스 상에 저장된 민감한 또는 개인적 사용자 데이터가 디바이스에 대해 허가되지 않은 사용자 또는 침입자에 의해 액세스될 수 있거나, 또는 사용자가 디바이스의 음성 제어 특징들을 사용하여 액세스할 수 있는 기능성이 단지 비-개인적, 비-민감한 정보 및 커맨드들에 제한될 필요가 있을 수 있다.

[0039] 단계(402)에서, 스피치 입력이 획득된다. 예를 들어, 심지어 디바이스(100)가 잠긴 경우라도, 잠긴 디바이스(100)의 사용자는 버튼(102)을 누르고 유지하여 디바이스(100)의 음성 제어 특징들을 활성화할 수 있다. 일부 구현 예들에서, 디바이스(100)의 음성 제어 특징들이 활성화되는 경우, 디바이스(100)는 마이크로폰(108)을 통해 스피치 입력을 수신할 수 있다.

[0040] 단계(404)에서, 스피치 입력은 사용자 인증을 수행하기 위해 사용된다. 일부 구현 예들에서, 스피치 입력은 화자 인식 분석을 사용하여 디바이스(100)에 대해 사용자를 인증하기 위해 사용될 수 있다. 예를 들어, 디바이스(100)가 잠긴 경우, 스피치 입력의 음성은, 스피치 입력을 발행하는 사용자가 디바이스(100)의 허가된 사용자인지의 여부를 결정하기 위해 화자 인식 분석을 사용하여 분석될 수 있다. 예를 들어, 스피치 입력에서의 음성의 음성 특성들은 네트워크 서비스에 의해 또는 디바이스(100) 상에 저장된 허가된 사용자의 성문의 음성 특성들과 비교될 수 있다. 음성이 성문에 매치될 수 있는 경우, 사용자는 디바이스(100)의 허가된 사용자로서 인증될 수 있다. 음성이 성문에 매치될 수 없는 경우, 사용자는 디바이스(100)의 허가된 사용자로서 인증되지 않을 것이다. 사용자가 스피치 입력에 기초하여 디바이스(100)에 대해 인증될 수 없는 경우, 예러 메시지가 사용자에게 (예를 들어, 청각적으로 및/또는 시각적으로, 진동으로) 제시될 수 있다. 예를 들어, 사용자가 스피치 입력에 기초하여 인증될 수 없는 경우, 디바이스(100)는 스피커(110) 또는 라우드 스피커(112) 또는 진동 소스에 의해

제공된 진동을 통해 제시되는 사운드(예를 들어, 알람 또는 합성된 음성 메시지)를 이용하여 인증 에러를 사용자에게 통지할 수 있다. 디바이스(100)는 추가적인 인증 정보(예를 들어, 패스워드, 코드, 터치 패턴 등)을 제공하도록 사용자에게 프롬프트를 터치 인터페이스(104) 상에 제시함으로써 시각적 에러를 제시할 수 있다.

[0041] 단계(406)에서, 스피치 입력에 기초하여 커맨드가 결정될 수 있다. 전술된 바와 같이, 스피치 입력은 텍스트로 변환될 수 있고, 텍스트는 스피치 입력에 존재하는 커맨드를 결정하기 위해 프로세싱될 수 있다. 일부 구현 예들에서, 스피치 입력이 스피치 입력 내의 커맨드를 결정하기 위해 프로세싱되는 동안 사용자가 스피치 입력에 기초하여 인증될 수 있다. 즉, 사용자는 디바이스(100)에 단일 스피치 입력을 제출할 수 있고, 해당 단일 스피치 입력은 사용자를 인증하기 위해 그리고 또한 사용자가 어떤 커맨드를 디바이스가 실행하기를 원하는지를 결정하기 위해 프로세싱될 수 있다.

[0042] 단계(408)에서, 커맨드는 음성이 인증되는 경우 실행될 수 있다. 일부 구현 예들에서, 스피치 입력 내의 사용자의 음성이 허가된 사용자의 성분에 매치될 수 있는 경우, 사용자의 음성은 인증될 수 있고 디바이스는 결정된 커맨드를 실행할 수 있다. 일부 구현 예들에서, 디바이스(100)가 잠긴 동안 디바이스(100)는 결정된 커맨드를 실행할 수 있다. 예를 들어, 디바이스(100)가 커맨드를 실행하는 동안 디바이스(100)는 잠긴 채 유지될 수 있고, 따라서, 디바이스(100)에 의해 수신된 추가적인 음성(또는 비-음성) 입력은 이러한 입력을 제공하는 사용자의 인증을 요구할 것이다. 일부 구현 예들에서, 잠긴 디바이스(100)는 전술된 음성 인증 프로세스들을 사용하여 잠긴 디바이스(100)에 대해 사용자를 인증하는 것에 응답하여 잠금 해제될 수 있다. 예를 들어, 잠긴 디바이스(100)는 사용자의 음성이 디바이스(100)의 허가된 사용자에 속하는 것으로 인증되는 경우 잠금 해제될 수 있고, 따라서, 후속의 입력 또는 커맨드들은 추가적인 인증을 요구하지 않는다.

[0043] 일부 구현 예들에서, 디바이스에 대해 사용자를 인증하거나 또는 음성 인증의 결과를 확인하여 성공적인 음성 인증의 더 많은 확신을 제공하기 위해 다른 생체측정 데이터(예를 들어, 사용자의 음성 이외)가 사용될 수 있다. 예를 들어, 모바일 디바이스(100)의 전방 카메라(116)는 얼굴 인식 분석에 기초하여 디바이스의 허가된 사용자를 인식하기 위해 사용될 수 있는 사용자의 얼굴의 이미지들을 수집하기 위해 사용될 수 있다. 또 다른 예로서, 터치-감지 디스플레이(104) 또는 버튼(120)은 사용자에 대한 지문 데이터를 수집하도록 구성될 수 있고, 지문 데이터는 디바이스에 대해 사용자를 인증하기 위해 사용될 수 있다.

[0044] 일부 구현 예들에서, 다른 타입들의 생체측정 데이터를 사용하여 사용자를 인증하는 것은 수동적으로 수행될 수 있다. 예를 들어, 사용자의 인증은 사용자가 비-인증-특정 방식들로 디바이스와 상호작용하는 동안 수행될 수 있다. 예를 들어, 사용자의 지문은 음악 플레이어 오브젝트(124)와 상호작용하기 위해 사용자가 터치-감지 디스플레이를 터치하는 경우 인증될 수 있다. 전방 카메라(116)는, 예를 들어, 사용자가 디바이스(100)의 비디오 채팅 특징들과 상호작용함에 따라 사용자의 얼굴의 이미지들을 수집할 수 있다. 전방 카메라(116)는 사용자가 웹 브라우징과 같은 다른 방식들로 디바이스(100)를 조작하는 동안 얼굴 인식 분석 및 인증을 위한 이미지들을 수집할 수 있다. 수집된 이미지들은 얼굴 인식 분석을 사용하여 사용자를 인증하기 위해 사용될 수 있다. 일부 구현 예들에서, 생체측정 데이터의 조합이 수집되고 사용자가 디바이스(100)에 액세스하려고 시도하는 경우 사용자를 인증하기 위해 사용될 수 있다. 예를 들어, 화자 인식, 얼굴 인식, 지문 매칭, 또는 다른 생체측정 데이터의 조합이 디바이스(100)에 대해 사용자를 인증하기 위해 사용될 수 있다.

예시적인 네트워크 운용 환경

[0045] 도 5는 예시적인 네트워크 운용 환경(500)의 블록도이다. 도 5에서, 모바일 디바이스들(502a 및 502b)은 각각 모바일 디바이스(100)를 나타낼 수 있다. 모바일 디바이스들(502a 및 502b)은, 예를 들어, 데이터 통신시에 하나 이상의 유선 및/또는 무선 네트워크들(510)을 통해 통신할 수 있다. 예를 들어, 무선 네트워크(512), 예를 들어, 셀룰러 네트워크가, 게이트웨이(516)를 사용하여, 인터넷과 같은 광역 네트워크(WAN)(514)와 통신할 수 있다. 마찬가지로, 802.11g 무선 액세스 디바이스와 같은 액세스 디바이스(518)가 광역 네트워크(514)로의 통신 액세스를 제공할 수 있다. 일부 구현 예들에서, 음성 및 데이터 통신들 모두가 무선 네트워크(512) 및 액세스 디바이스(518) 상에서 설정될 수 있다. 예를 들어, 모바일 디바이스(502a)는, 무선 네트워크(512), 게이트웨이(516), 및 광역 네트워크(514)를 통해(예를 들어, TCP/IP 또는 UDP 프로토콜들을 사용하여) 전화 통화를 신청하고 받고(예를 들어, VoIP 프로토콜들을 사용하여), 이메일 메시지들을 송신 및 수신하고(예를 들어, POP3 프로토콜을 사용하여), 전자 문서들 및/또는 스트리밍들, 예를 들어, 웹 페이지들, 사진들, 및 비디오들을 검색할 수 있다. 마찬가지로, 일부 구현 예들에서, 모바일 디바이스(502b)는 액세스 디바이스(518) 및 광역 네트워크(514)를 통해 전화 통화를 신청하고 받고, 이메일 메시지들을 송신 및 수신하고, 전자 문서들을 검색할 수 있다. 일부 구현 예들에서, 모바일 디바이스(502a 또는 502b)는 하나 이상의 케이블들을 사용하여 액세스 디바이스(512) 및 광역 네트워크(514)와 통신할 수 있다.

이스(518)에 물리적으로 접속될 수 있고, 액세스 디바이스(518)는 개인용 컴퓨터일 수 있다. 이러한 구성에서, 모바일 디바이스(502a 또는 502b)는 "테더링된(tethered)" 디바이스라고 불릴 수 있다.

[0047] 모바일 디바이스들(502a 및 502b)은 또한 다른 수단에 의해 통신들을 설정할 수 있다. 예를 들어, 무선 디바이스(502a)는 무선 네트워크(512)를 통해 다른 무선 디바이스들, 예를 들어, 다른 모바일 디바이스들(502a 또는 502b), 셀 폰들 등과 통신할 수 있다. 마찬가지로, 모바일 디바이스들(502a 및 502b)은, Bluetooth™ 통신 디바이스들과 같은 하나 이상의 통신 서브시스템들을 사용하여, 피어-투-피어 통신들(520), 예를 들어 개인 영역 네트워크를 설정할 수 있다. 다른 통신 프로토콜들 및 토폴로지들이 또한 구현될 수 있다.

[0048] 모바일 디바이스(502a 또는 502b)는, 예를 들어, 하나 이상의 유선 및/또는 무선 네트워크들(510)을 통해 하나 이상의 서비스들(530, 540, 550, 560, 570 및 580)과 통신할 수 있다. 예를 들어, 내비게이션 서비스(530)는 내비게이션 정보, 예를 들어, 맵 정보, 위치 정보, 라우트 정보, 및 다른 정보를 모바일 디바이스(502a 또는 502b)에 제공할 수 있다. 모바일 디바이스(502b)의 사용자는 맵 기능성을 불러올 수 있고, 특정 위치에 대한 맵을 요청하고 수신할 수 있다.

[0049] 메시징 서비스(540)는, 예를 들어, 이메일 및/또는 다른 메시징 서비스들을 제공할 수 있다. 미디어 서비스(550)는 예를 들어, 노래 파일들, 오디오 북들, 영화 파일들, 비디오 클립들, 및 다른 미디어 데이터와 같은 미디어 파일들로의 액세스를 제공할 수 있다. 일부 구현예들에서, 별도의 오디오 및 비디오 서비스들(미도시)이 각각의 타입들의 미디어 파일들로의 액세스를 제공할 수 있다. 동기화 서비스(560)는 예를 들어, 동기화 서비스들(예를 들어, 동기 파일들)을 수행할 수 있다. 활성화 서비스(570)는 예를 들어, 모바일 디바이스(예를 들어, 502a 또는 502b)를 활성화하기 위한 활성화 프로세스를 수행할 수 있다.

[0050] 성문 서비스(580)는 예를 들어, 모바일 디바이스(502a 또는 502b)의 사용자들을 인증하기 위해 사용될 수 있는 성문들을 생성할 수 있다. 예를 들어, 성문 서비스(580)는 모바일 디바이스(502a 또는 502b)로부터 사용자의 음성의 샘플들을 수신하고, 음성 샘플들에 기초하여 성문을 생성할 수 있다. 모바일 디바이스(502a 또는 502b)는, 예를 들어, 사용자가 모바일 디바이스(502a 또는 502b)의 다양한 음성 특징들(예를 들어, 음성 제어, 전화, 음성 레코더 등)과 상호작용함에 따라 음성 샘플들을 수집할 수 있다. 성문 서비스(580)가 사용자에 대한 성문을 생성하면, 성문 서비스(580)는 모바일 디바이스(502a 또는 502b)에 성문을 전송할 수 있다. 성문이 모바일 디바이스(502a 또는 502b)에서 수신되면, 성문은 모바일 디바이스가 사용자의 음성에 기초하여 사용자를 인증하는 데 사용될 수 있다. 예를 들어, 성문 서비스(580)에 의해 생성된 성문들은 텍스트-독립적 성문들일 수 있다.

[0051] 모바일 디바이스(502a 또는 502b) 상의 소프트웨어에 대한 소프트웨어 업데이트들이 존재하는지의 여부를 자동으로 결정하고, 이후 모바일 디바이스(502a 또는 502b)에 소프트웨어 업데이트들을 다운로드하는 소프트웨어 업데이트 서비스를 포함하는 다른 서비스들이 제공될 수 있고, 모바일 디바이스(502a 또는 502b)에서, 소프트웨어 업데이트들은 수동적으로 또는 자동적으로 언패킹되고/되거나 설치될 수 있다.

[0052] 모바일 디바이스(502a 또는 502b)는 또한 하나 이상의 유선 및/또는 무선 네트워크들(510)을 통해 다른 데이터 및 컨텐츠에 액세스할 수 있다. 예를 들어, 뉴스 사이트, RSS 피드, 웹 사이트, 블로그, 소셜 네트워킹 사이트, 개발자 네트워크 등과 같은 컨텐츠 게시자들이 모바일 디바이스(502a 또는 502b)에 의해 액세스될 수 있다. 이러한 액세스는 모바일 디바이스(502a 또는 502b)의 웹 브라우징 기능 또는 애플리케이션(예를 들어, 브라우저)을 불러옴으로써 제공될 수 있다.

[0053] 예시적인 모바일 디바이스 아키텍처

[0054] 도 6은 도 1-4의 모바일 디바이스(100)의 예시적인 구현예의 블록도(600)이다. 모바일 디바이스(100)는 메모리 인터페이스(602), 하나 이상의 데이터 프로세서들, 이미지 프로세서들, 및/또는 중앙 처리 장치들(604), 및 주변 장치 인터페이스(606)를 포함할 수 있다. 메모리 인터페이스(602), 하나 이상의 프로세서(604) 및/또는 주변 장치 인터페이스(606)는 별도의 컴포넌트들일 수 있거나 또는 하나 이상의 집적 회로들에 통합될 수 있다. 모바일 디바이스(100) 내의 다양한 컴포넌트들은 하나 이상의 통신 버스들 또는 신호 라인들에 의해 연결될 수 있다.

[0055] 센서들, 디바이스들 및 서브시스템들이 다수의 기능들을 용이하게 하기 위해 주변 장치 인터페이스(606)에 연결될 수 있다. 예를 들어, 모션 센서(610), 광 센서(612), 및 근접 센서(614)가 배향, 조명 및 근접 기능들을 용이하게 하기 위해 주변 장치 인터페이스(606)에 연결될 수 있다. 다른 센서들(616)은 또한, 관련 기능성들을 용이하게 하기 위해, 포지셔닝 시스템(예를 들어, GPS 수신기), 온도 센서, 생체측정 센서, 또는 다른 감지 디

바이스와 같은 주변 장치 인터페이스(606)에 접속될 수 있다.

[0056] 카메라 서브시스템(620) 및 광 센서(622), 예를 들어, 전하 결합 소자(CCD) 또는 상보적 금속-산화물 반도체(CMOS) 광학 센서가 사진들 및 비디오 클립들의 기록과 같은 카메라 기능들을 용이하게 하기 위해 이용될 수 있다. 카메라 서브시스템(620) 및 광학 센서(622)는 예를 들어, 얼굴 인식 분석을 수행함으로써, 사용자의 인증 동안 사용될 사용자의 이미지들을 수집하기 위해 사용될 수 있다.

[0057] 통신 기능들은 라디오 주파수 수신기들 및 송신기들 및/또는 광학(예를 들어, 적외선) 수신기들 및 송신기들을 포함할 수 있는 하나 이상의 무선 통신 서브시스템들(624)을 통해 용이해질 수 있다. 통신 서브시스템(624)의 특정 설계 및 구현에는 모바일 디바이스(100)가 동작하도록 의도되는 통신 네트워크(들)에 의존할 수 있다. 예를 들어, 모바일 디바이스(100)는 GSM 네트워크, GPRS 네트워크, EDGE 네트워크, Wi-Fi 또는 WiMax 네트워크, 및 Bluetooth™ 네트워크를 통해 동작하도록 설계되는 통신 서브시스템들(624)을 포함할 수 있다. 특히, 무선 통신 서브시스템들(624)은 디바이스(100)가 다른 무선 디바이스들에 대한 기지국으로서 구성될 수 있도록 호스팅 프로토콜들을 포함할 수 있다.

[0058] 오디오 서브시스템(626)은 화자 인식, 음성 복제, 디지털 레코딩 및 텔레포니 기능들과 같은 음성-인에이블 기능들을 용이하게 하기 위해 스피커(628) 및 마이크로폰(630)에 연결될 수 있다. 오디오 서브시스템(626)은, 도 1-4를 참조하여 전술된 바와 같이, 음성 커맨드들, 성문 및 음성 인증의 프로세싱을 용이하게 하도록 구성될 수 있다.

[0059] I/O 서브시스템(640)은 터치 스크린 제어기(642) 및/또는 다른 입력 제어기(들)(644)를 포함할 수 있다. 터치-스크린 제어기(642)는 터치 스크린(646)에 연결될 수 있다. 터치 스크린(646) 및 터치 스크린 제어기(642)은, 예를 들어, 용량성, 저항성, 적외선, 및 표면 음파 기술들뿐만 아니라 터치 스크린(646)과의 하나 이상의 접촉 포인트들을 결정하기 위한 다른 근접 센서 어레이들 또는 다른 엘리먼트들을 포함하지만 이에 제한되지 않는 복수의 터치 감지 기술들 중 임의의 것을 사용하여 접촉 및 움직임 또는 그의 중단을 검출할 수 있다.

[0060] 다른 입력 제어기(들)(644)는 하나 이상의 버튼들, 로커(rocker) 스위치들, 셤 휠, 적외선 포트, USB 포트, 및/또는 스타일러스와 같은 포인터 디바이스와 같은 다른 입력/제어 디바이스들(648)에 연결될 수 있다. 하나 이상의 버튼들(미도시)은 스피커(628) 및/또는 마이크로폰(630)의 볼륨 제어를 위한 업/다운 버튼을 포함할 수 있다.

[0061] 일 구현예에서, 제1 지속 기간 동안 버튼을 누르는 것은 터치 스크린(646)의 잠금을 해제할 수 있고; 그리고 제1 지속 기간보다 더 긴 제2 지속 기간 동안 버튼을 누르는 것은 모바일 디바이스(100)로의 전원을 켜거나 끌 수 있다. 제3 지속 기간 동안 버튼을 누르는 것은 사용자가 마이크로폰(630)에 대고 커맨드들을 발성하여 디바이스로 하여금 발성된 커맨드를 실행하게 할 수 있는 음성 제어, 또는 음성 커맨드 모듈을 활성화할 수 있다. 사용자는 버튼들 중 하나 이상의 버튼의 기능성을 맞춤화할 수 있다. 터치 스크린(646)은 예를 들어, 또한 가상 또는 소프트 버튼들 및/또는 키보드를 구현하는 데 사용될 수 있다.

[0062] 일부 구현예들에서, 모바일 디바이스(100)는 MP3, AAC, 및 MPEG 파일들과 같은 레코딩된 오디오 및/또는 비디오 파일들을 제시할 수 있다. 일부 구현예들에서, 모바일 디바이스(100)는 iPod™와 같은 MP3 플레이어의 기능성을 포함할 수 있다. 따라서, 모바일 디바이스(100)는 iPod과 호환가능한 36-핀 커넥터를 포함할 수 있다. 다른 입력/출력 및 제어 디바이스들이 또한 사용될 수 있다.

[0063] 메모리 인터페이스(602)는 메모리(650)에 연결될 수 있다. 메모리(650)는 고속 랜덤 액세스 메모리 및/또는 비휘발성 메모리, 예컨대, 하나 이상의 자기 디스크 저장 디바이스들, 하나 이상의 광학 저장 디바이스들, 및/또는 플래시 메모리(예를 들어, NAND, NOR)를 포함할 수 있다. 메모리(650)는 예를 들어, Darwin, RTXC, LINUX, UNIX, OS X, WINDOWS와 같은 운영 체제(652), 또는 Vx Works와 같은 임베디드 운영 체제를 저장할 수 있다.

[0064] 운영 체제(652)는 기본 시스템 서비스들을 핸들링하고 하드웨어 종속적 작업들을 수행하기 위한 명령들을 포함할 수 있다. 일부 구현예들에서, 운영 체제(652)는 커널(예를 들어, UNIX 커널)일 수 있다. 일부 구현예들에서, 운영 체제(652)는 음성 인증을 수행하기 위한 명령들을 포함할 수 있다. 예를 들어, 운영 체제(652)는 도 1-4를 참조하여 기술되는 바와 같은 보안 잠금(security lockout) 및 음성 인증 특징들을 구현할 수 있다. 운영 체제(352)는 도 1-4를 참조하여 기술된 성문 및 음성 인증 특징들을 구현할 수 있다.

[0065] 메모리(650)는 또한 하나 이상의 추가적인 디바이스들, 하나 이상의 컴퓨터들 및/또는 하나 이상의 서버들과 통신하는 것을 용이하게 하기 위해 통신 명령들(654)을 저장할 수 있다. 메모리(650)는 그래픽 사용자 인터페이스 프로세싱을 용이하게 하기 위한 그래픽 사용자 인터페이스 명령들(656); 센서 관련 프로세싱 및 기능들을 용

이하게 하기 위해 센서 프로세싱 명령들(658); 전화 관련 프로세스들 및 기능들을 용이하게 하기 위한 전화 명령들(660); 전자 메시징 관련 프로세스들 및 기능들을 용이하게 하기 위한 전자 메시징 명령들(662); 웹 브라우징 관련 프로세스들 및 기능들을 용이하게 하기 위한 웹 브라우징 명령들(664); 미디어 프로세싱 관련 프로세스들 및 기능들을 용이하게 하기 위한 미디어 프로세싱 명령들(666); GPS 및 내비게이션 관련 프로세스들 및 명령들을 용이하게 하기 위한 GPS/내비게이션 명령들(668); 및/또는 카메라 관련 프로세스들 및 기능들을 용이하게 하기 위한 카메라 명령들(670)을 포함할 수 있다.

[0066] 메모리(650)는 도 1-4를 참조하여 기술된 바와 같이 보안 및/또는 인증 프로세스들과 같은 다른 프로세스들 및 기능들을 용이하게 하기 위한 다른 소프트웨어 명령들(672)을 저장할 수 있다. 예를 들어, 소프트웨어 명령들은 매 애플리케이션 또는 매 특정 기반으로 음성 인증을 수행하기 위한, 그리고 사용자가 디바이스(100) 상에 이용가능한 각각의 애플리케이션 또는 특정의 인증 요건들을 구성하게 하기 위한 명령들을 포함할 수 있다.

[0067] 메모리(650)는 또한 웹 비디오 관련 프로세스들 및 기능들을 용이하게 하기 위한 웹 비디오 명령들과 같은 다른 소프트웨어 명령들(미도시); 및/또는 웹 쇼핑 관련 프로세스들 및 기능들을 용이하게 하기 위한 웹 쇼핑 명령들을 저장할 수 있다. 일부 구현예들에서, 미디어 프로세싱 명령들(666)은 각각 오디오 프로세싱 관련 프로세스들 및 기능들 및 비디오 프로세싱 관련 프로세스들 및 기능들을 용이하게 하기 위한 오디오 프로세싱 명령들 및 비디오 프로세싱 명령들로 분할된다. 활성화 레코드 및 국제 모바일 단말기 식별 번호(International Mobile Equipment Identity, IMEI)(674) 또는 유사한 하드웨어 식별자가 또한 메모리(650)에 저장될 수 있다.

[0068] 위에서 식별된 명령들 및 애플리케이션들 각각은 전술된 하나 이상의 기능들을 수행하기 위한 명령들의 세트에 대응할 수 있다. 이들 명령들은 별도의 소프트웨어 프로그램들, 프로시저들, 또는 모듈들로서 구현될 필요가 없다. 메모리(650)는 추가적인 명령들 또는 더 적은 명령들을 포함할 수 있다. 추가로, 모바일 디바이스(100)의 다양한 기능들은, 하나 이상의 신호 프로세싱 및/또는 주문형 집적 회로들에서 구현되는 것을 포함하여, 하드웨어 및/또는 소프트웨어에서 구현될 수 있다.

[0069] 일부 실시예들에 따라, 도 7은 전술된 바와 같은 본 발명의 원리들에 따라 구성된 전자 디바이스(700)의 기능 블록도를 도시한다. 디바이스의 기능 블록들은 본 발명의 원리들을 수행하기 위해 하드웨어, 소프트웨어, 또는 하드웨어와 소프트웨어의 결합에 의해 구현될 수 있다. 도 7에 도시된 기능 블록들이 전술된 바와 같은 본 발명의 원리들을 구현하기 위해 결합되거나 서브-블록들로 분리될 수 있다는 점이 당업자에 의해 이해된다. 따라서, 여기서의 기재는 여기서 기술된 기능 블록들의 임의의 가능한 결합 또는 분리 또는 추가적인 정의를 지원할 수 있다.

[0070] 도 7에 도시된 바와 같이, 전자 디바이스(700)는 스피치 입력을 수신하도록 구성되는 스피치 수신 유닛(702)을 포함한다. 스피치 입력은 전자 디바이스(700)의 제한된 특징과 연관된 커맨드를 포함한다. 전자 디바이스(700)는 또한 스피치 수신 유닛(702)에 연결된 프로세싱 유닛(706)을 포함할 수 있다. 일부 구현예들에서, 프로세싱 유닛(706)은 비교 유닛(708), 결정 유닛(710), 액세스 제공 유닛(712), 액세스 거절 유닛(714), 스피치 프로세싱 유닛(716), 및 수신 유닛(718)을 포함한다.

[0071] 프로세싱 유닛(706)은 전자 디바이스(700)의 허가된 사용자의 성문과 스피치 입력을 비교하고(예를 들어, 비교 유닛(708)을 이용하여); 비교의 결과에 기초하여, 스피치 입력이 허가된 사용자에 의해 발성되었다고 결정하고(예를 들어, 결정 유닛(710)을 이용하여); 커맨드에 따라 전자 디바이스(700)의 제한된 특징으로의 액세스를 제공하도록(예를 들어, 액세스 제공 유닛(712)을 이용하여) 구성된다.

[0072] 일부 실시예들에서, 프로세싱 유닛(706)은 비교의 결과에 기초하여, 스피치 입력이 허가된 사용자에 의해 발성되지 않았다고 결정하고(예를 들어, 결정 유닛(710)을 이용하여); 전자 디바이스(700)의 제한된 특징으로의 액세스를 거절하도록(예를 들어, 액세스 거절 유닛(714)을 이용하여) 구성된다.

[0073] 일부 구현예들에서, 프로세싱 유닛(706)은 전자 디바이스(700)의 허가된 사용자의 성문과 스피치 입력을 비교하는 동안 커맨드를 결정하기 위해 스피치 입력을 프로세싱하도록(예를 들어, 스피치 프로세싱 유닛(716)을 이용하여) 구성된다.

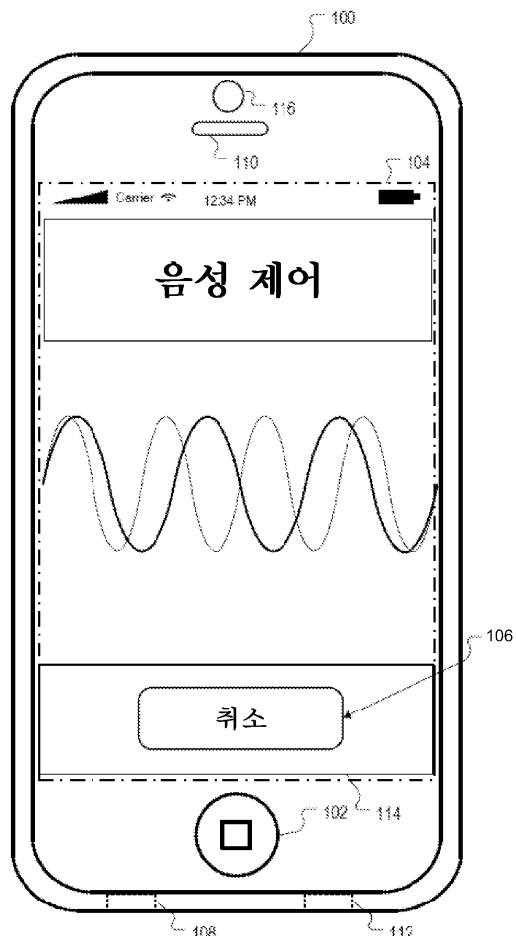
[0074] 일부 실시예들에서, 프로세싱 유닛(706)은 전자 디바이스(700)의 네트워크 인터페이스를 통해 성문 서비스로부터 성문을 수신하도록(예를 들어, 수신 유닛(718)을 이용하여) 구성된다.

[0075] 일부 실시예들에서, 전자 디바이스(700)는 모바일 디바이스이다. 일부 실시예들에서, 모바일 디바이스는 핸드 헬드 디바이스이다.

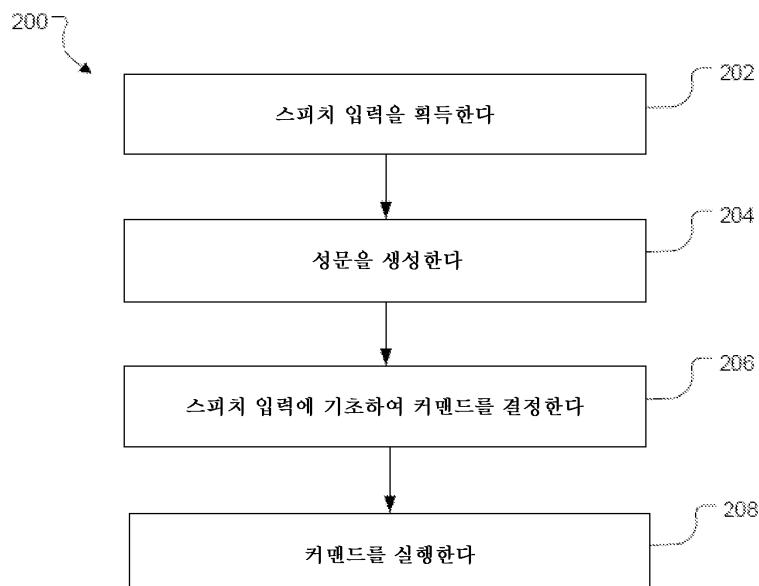
- [0076] 일부 실시예들에서, 성문은 텍스트-독립적 성문이다.
- [0077] 일부 실시예들에 따라, 도 8은 전술된 바와 같은 본 발명의 원리들에 따라 구성되는 전자 디바이스(800)의 기능 블록도를 도시한다. 디바이스의 기능 블록들은 본 발명의 원리들을 수행하기 위해 하드웨어, 소프트웨어, 또는 하드웨어와 소프트웨어의 결합에 의해 구현될 수 있다. 도 8에 설명된 기능 블록들이 전술된 바와 같은 본 발명의 원리들을 구현하기 위해 결합되거나 서브 블록들로 분리될 수 있다는 점이 당업자에 의해 이해된다. 따라서, 여기서의 기재는 여기서 기술된 기능 블록들의 임의의 가능한 결합 또는 분리 또는 추가적인 정의를 지원할 수 있다.
- [0078] 도 8에 도시된 바와 같이, 전자 디바이스(800)는 스피치 입력을 수신하도록 구성되는 스피치 수신 유닛(802)을 포함한다. 스피치 입력은 전자 디바이스(800)의 특징과 연관된 커맨드를 포함한다. 전자 디바이스(800)는 또한 스피치 수신 유닛(802)에 연결된 프로세싱 유닛(806)을 포함한다. 일부 실시예들에서, 프로세싱 유닛(806)은 생성 유닛(808), 액세스 제공 유닛(810), 스피치 프로세싱 유닛(812), 음성 샘플 생성 유닛(814), 음성 샘플 저장 유닛(816), 및 성문 생성 유닛(818)을 포함한다.
- [0079] 프로세싱 유닛(806)은 스피치 입력에 기초하여 텍스트-독립적 성문을 생성하고(예를 들어, 생성 유닛(808)을 이용하여); 커맨드에 따라 전자 디바이스(800)의 특징으로의 액세스를 제공하도록(예를 들어, 액세스 제공 유닛(810)을 이용하여) 구성된다.
- [0080] 일부 실시예들에서, 프로세싱 유닛(806)은 스피치 입력의 음성에 기초하여 텍스트-독립적 성문을 생성하는 동안 커맨드를 결정하기 위해 스피치 입력을 프로세싱하도록(예를 들어, 스피치 프로세싱 유닛(812)을 이용하여) 구성된다.
- [0081] 일부 실시예들에서, 프로세싱 유닛(806)은 스피치 입력에 기초하여 음성 샘플들을 생성하고(예를 들어, 음성 샘플 생성 유닛(814)을 이용하여); 전자 디바이스(800) 상에 음성 샘플들을 저장하고(예를 들어, 음성 샘플 저장 유닛(816)을 이용하여); 음성 샘플들에 기초하여 텍스트-독립적 성문을 생성하도록(예를 들어, 성문 생성 유닛(818)을 이용하여) 구성된다.
- [0082] 일부 실시예들에 따라, 도 9는 전술된 바와 같은 본 발명의 원리들에 따라 구성되는 전자 디바이스(900)의 기능 블록도를 도시한다. 디바이스의 기능 블록들은 본 발명의 원리들을 수행하기 위해 하드웨어, 소프트웨어, 또는 하드웨어와 소프트웨어의 결합에 의해 구현될 수 있다. 도 9에 기술된 기능 블록들이 전술된 바와 같은 본 발명의 원리들을 구현하기 위해 결합되거나 서브블록들로 분리될 수 있다는 점이 당업자에 의해 이해된다. 따라서, 여기서의 기재는 여기서 기술된 기능 블록들의 임의의 가능한 결합 또는 분리 또는 추가적인 정의를 지원할 수 있다.
- [0083] 도 9에 도시된 바와 같이, 전자 디바이스(900)는 스피치 입력을 수신하도록 구성되는 스피치 수신 유닛(902)을 포함한다. 스피치 입력은 전자 디바이스(900)의 특징과 연관된 커맨드를 포함한다. 전자 디바이스(900)는 또한 스피치 수신 유닛(902)에 연결된 프로세싱 유닛(906)을 포함한다. 일부 실시예들에서, 프로세싱 유닛(906)은 음성 샘플 생성 유닛(908), 음성 샘플 전송 유닛(910), 액세스 제공 유닛(912), 음성 샘플 저장 유닛(914), 및 수신 유닛(916)을 포함한다.
- [0084] 프로세싱 유닛은 스피치 입력에 기초하여 음성 샘플을 생성하고(예를 들어, 음성 샘플 생성 유닛(908)을 이용하여); 음성 샘플에 기초하여 성문을 생성하기 위해 성문 서비스에 음성 샘플을 전송하고(예를 들어, 음성 샘플 전송 유닛(910)을 이용하여); 커맨드에 따라 전자 디바이스(900)의 특징으로의 액세스를 제공하도록(예를 들어, 액세스 제공 유닛(912)을 이용하여) 구성된다.
- [0085] 일부 실시예들에서, 프로세싱 유닛(906)은 스피치 입력에 기초하여 음성 샘플들을 생성하고(예를 들어, 음성 샘플 생성 유닛(908)을 이용하여); 전자 디바이스(900) 상에 음성 샘플들을 저장하고(예를 들어, 음성 샘플 저장 유닛(914)을 이용하여); 음성 샘플들에 기초하여 성문을 생성하기 위해 성문 서비스에 저장된 음성 샘플들을 전송하도록(예를 들어, 음성 샘플 전송 유닛(910)을 이용하여) 구성된다.
- [0086] 일부 실시예들에서, 프로세싱 유닛(906)은 전자 디바이스(900)의 네트워크 인터페이스를 통해 성문 서비스로부터 텍스트-독립적 성문을 수신하도록(예를 들어, 수신 유닛(916)을 이용하여) 구성된다.
- [0087] 일부 실시예들에서, 음성 샘플은 스피치 입력에 기초하여 커맨드가 결정되는 동안 생성된다.

도면

도면1



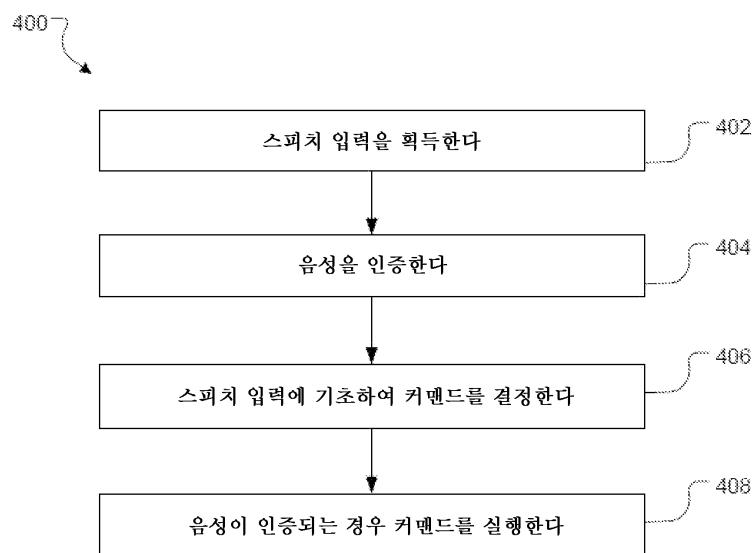
도면2



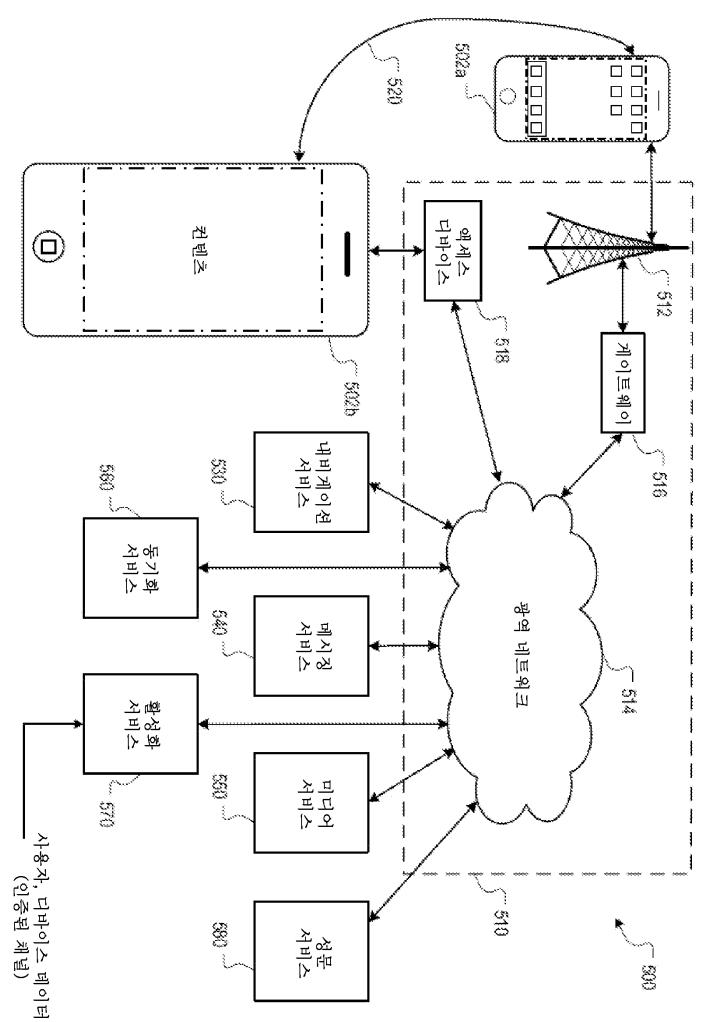
도면3



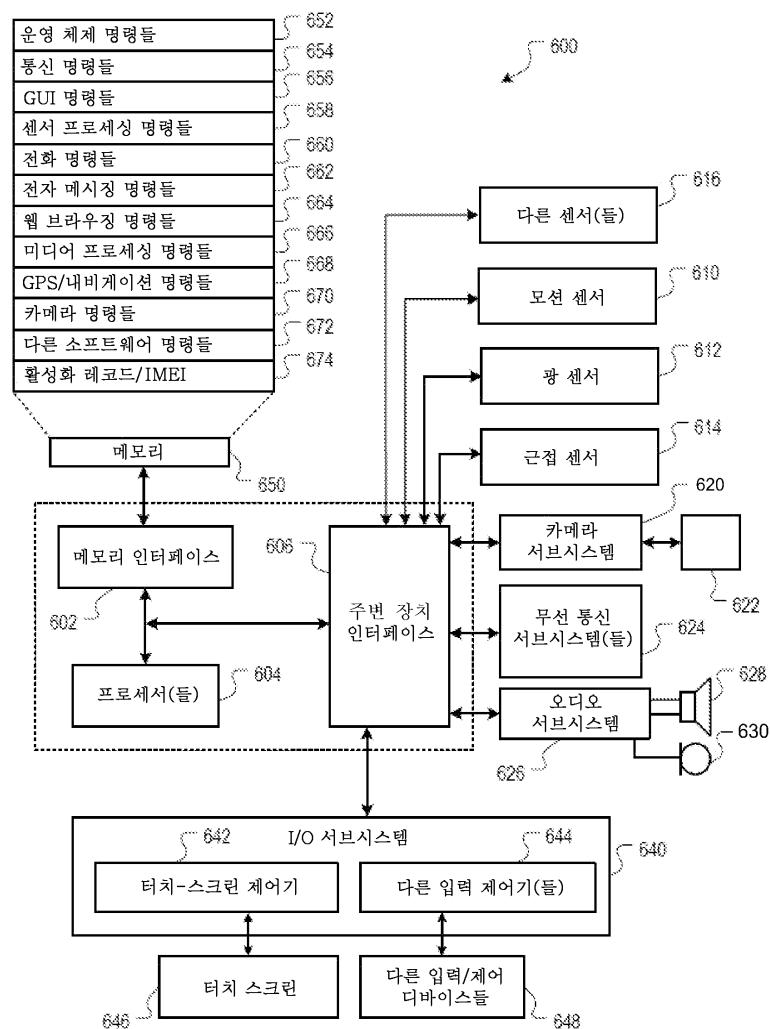
도면4



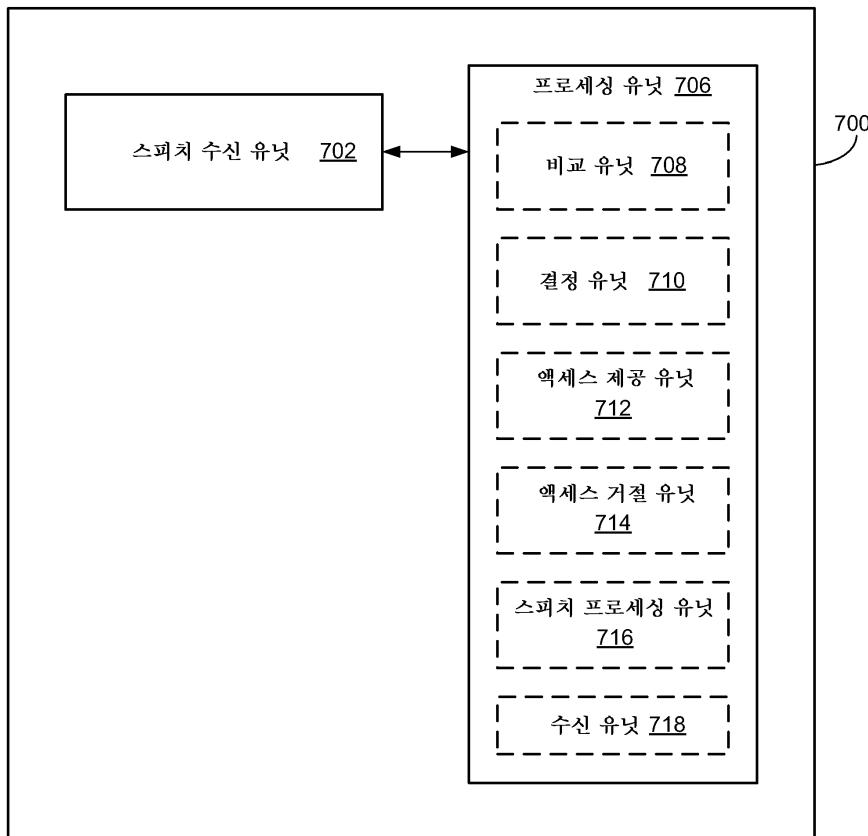
도면5



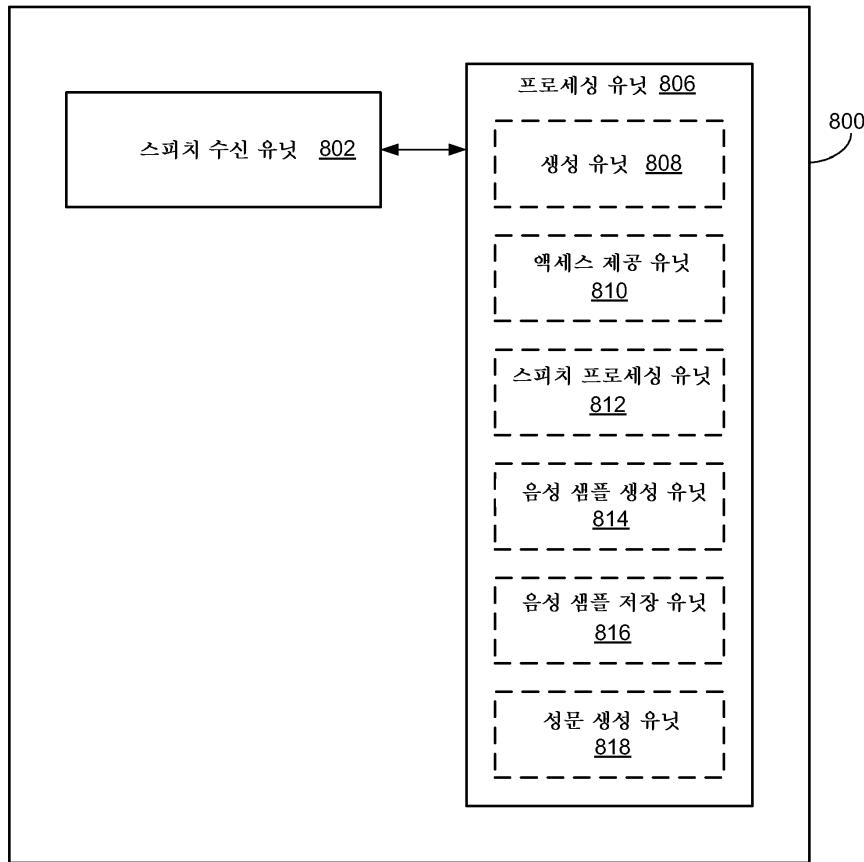
도면6



도면7



도면8



도면9

