(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0097519
(43) 공개일자 2017년08월28일

(51) 국제특허분류(Int. Cl.)
G10L 15/22 (2006.01) *G10L 15/04* (2006.01)
G10L 15/20 (2006.01) *G10L 15/28* (2006.01)
H04R 1/40 (2006.01)
(52) CPC특허분류
G10L 15/22 (2013.01)
G10L 15/04 (2013.01)
(21) 출원번호 10-2016-0019391
(22) 출원일자 2016년02월18일
심사청구일자 없음

(71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
(72) 발명자
손동일
경기도 화성시 동탄반석로 71, 솔빛마을쌍용예가아파트 446-501
김연형
경기도 용인시 기흥구 예현로 15, SK아파트 102-1902
(뒷면에 계속)
(74) 대리인
특허법인태평양

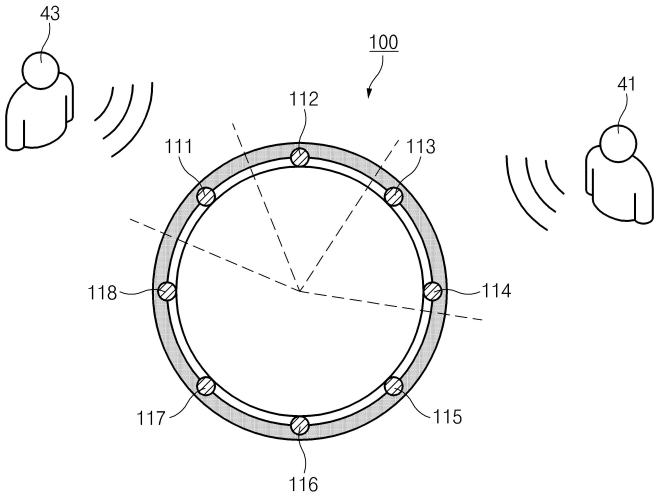
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 **음성 처리 방법 및 장치**

(57) 요 약

본 발명의 다양한 실시 예에 따른 전자 장치는, 지정된 방향을 향하도록 배치된 복수의 마이크를 포함하는 마이크 어레이(microphone array), 전자 장치 주변에 위치하는 사용자를 감지하는 센서 모듈 및 전자 장치 주변에 복수의 사용자가 존재하면 복수의 사용자 중 하나의 사용자를 선택하고, 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하고 나머지 방향으로부터 수신되는 음성을 노이즈로 처리하도록 설정된 프로세서를 포함할 수 있다. 또한, 다른 실시 예도 가능하다.

대 표 도 - 도4



(52) CPC특허분류
G10L 15/20 (2013.01)
G10L 15/28 (2013.01)
H04R 1/406 (2013.01)
(72) 발명자
윤건호
서울특별시 송파구 오금로32길 42, 현대아파트
조치현
경기도 수원시 영통구 동탄원천로915번길 36, 주공그린빌 303-603
허창룡
경기도 수원시 권선구 동수원로145번길 73, 수원아이파크시티아파트 310-804

명 세 서

청구범위

청구항 1

전자 장치에 있어서,

지정된 방향을 향하도록 배치된 복수의 마이크를 포함하는 마이크 어레이(microphone array);

상기 전자 장치 주변에 위치하는 사용자를 감지하는 센서 모듈; 및

상기 전자 장치 주변에 복수의 사용자가 존재하면 상기 복수의 사용자 중 하나의 사용자를 선택하고, 상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하고 나머지 방향으로부터 수신되는 음성을 노이즈로 처리하도록 설정된 프로세서;를 포함하는 전자 장치.

청구항 2

제1항에 있어서,

상기 프로세서는,

상기 복수의 사용자 중 지정된 명령어를 먼저 발화한 사용자를 선택하도록 설정된 전자 장치.

청구항 3

제1항에 있어서,

상기 프로세서는,

상기 복수의 사용자로부터 수신되는 음성을 이용하여 상기 복수의 사용자 각각을 식별하고, 상기 복수의 사용자 각각에 대해 우선순위를 판단하고, 상기 복수의 사용자 중 우선순위가 가장 높은 사용자를 선택하도록 설정된 전자 장치.

청구항 4

제3항에 있어서,

상기 프로세서는,

상기 복수의 사용자 중 우선순위가 가장 높은 사용자의 발화가 종료되면, 다음 우선순위를 가지는 사용자를 선택하도록 설정된 전자 장치.

청구항 5

제1항에 있어서,

상기 센서 모듈은,

사용자의 움직임에 의해 사용자의 인체를 감지하는 제1 센서; 및

지정된 방향에 위치하는 물체를 감지하는 제2 센서;를 포함하는 전자 장치.

청구항 6

제5항에 있어서,

상기 프로세서는,

상기 제1 센서가 활성화된 상태에서 상기 제1 센서에 의해 사용자가 감지되면, 상기 제1 센서를 비활성화시키고 상기 제2 센서를 활성화시키도록 설정된 전자 장치.

청구항 7

제5항에 있어서,

상기 프로세서는,

상기 제2 센서가 활성화된 상태에서 상기 제2 센서에 의해 사용자가 감지되지 않으면, 상기 제2 센서를 비활성화시키고 상기 제1 센서를 활성화시키도록 설정된 전자 장치.

청구항 8

제1항에 있어서,

상기 프로세서는,

상기 나머지 방향으로부터 수신되는 음성을 이용하여 상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성에 노이즈 캔슬링을 수행하도록 설정된 전자 장치.

청구항 9

제1항에 있어서,

디스플레이; 및

스피커;를 더 포함하고,

상기 프로세서는,

상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 인식하고, 상기 디스플레이 및 상기 스피커 중 적어도 하나를 이용하여 상기 음성에 대한 피드백을 제공하도록 설정된 전자 장치.

청구항 10

제1항에 있어서,

상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 인식하고, 상기 음성에 대응하는 기능을 수행하도록 설정된 전자 장치.

청구항 11

전자 장치의 음성 처리 방법에 있어서,

상기 전자 장치 주변에 위치하는 사용자를 감지하는 동작;

지정된 방향을 향하도록 배치된 복수의 마이크를 포함하는 마이크 어레이를 이용하여 음성을 수신하는 동작;

상기 전자 장치 주변에 복수의 사용자가 존재하면 상기 복수의 사용자 중 하나의 사용자를 선택하는 동작;

상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하는 동작; 및

나머지 방향으로부터 수신되는 음성을 노이즈로 처리하는 동작;을 포함하는 방법.

청구항 12

제11항에 있어서,

상기 복수의 사용자 중 하나의 사용자를 선택하는 동작은,

상기 복수의 사용자 중 지정된 명령어를 먼저 발화한 사용자를 선택하는 동작;을 포함하는 방법.

청구항 13

제11항에 있어서,

상기 복수의 사용자 중 하나의 사용자를 선택하는 동작은,

상기 복수의 사용자로부터 수신되는 음성을 이용하여 상기 복수의 사용자 각각을 식별하는 동작;

상기 복수의 사용자 각각에 대해 우선순위를 판단하는 동작; 및

상기 복수의 사용자 중 우선순위가 가장 높은 사용자를 선택하는 동작;을 포함하는 방법.

청구항 14

제13항에 있어서,

상기 복수의 사용자 중 하나의 사용자를 선택하는 동작은,

상기 복수의 사용자 중 우선순위가 가장 높은 사용자의 발화가 종료되면, 다음 우선순위를 가지는 사용자를 선택하는 동작;을 더 포함하는 방법.

청구항 15

제11항에 있어서,

상기 전자 장치 주변에 위치하는 사용자를 감지하는 동작은,

사용자의 움직임에 의해 사용자를 감지하는 제1 센서를 활성화시키는 동작;

상기 제1 센서에 의해 사용자가 감지되면 상기 제1 센서를 비활성화시키는 동작; 및

상기 제1 센서에 의해 사용자의 인체가 감지되면 지정된 방향에 위치하는 물체를 감지하는 제2 센서를 활성화시키는 동작;을 포함하는 방법.

청구항 16

제15항에 있어서,

상기 전자 장치 주변에 위치하는 사용자를 감지하는 동작은,

상기 제2 센서에 의해 사용자가 감지되지 않으면 상기 제2 센서를 비활성화시키는 동작; 및

상기 제2 센서에 의해 사용자가 감지되지 않으면 상기 제1 센서를 활성화시키는 동작;을 더 포함하는 방법.

청구항 17

제11항에 있어서,

상기 나머지 방향으로부터 수신되는 음성을 노이즈로 처리하는 동작은,

상기 나머지 방향으로부터 수신되는 음성을 이용하여 상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성에 노이즈 캔슬링을 수행하는 동작;을 포함하는 방법.

청구항 18

제11항에 있어서,

상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 인식하는 동작; 및

디스플레이 및 스피커 중 적어도 하나를 이용하여 상기 음성에 대한 피드백을 제공하는 동작;을 더 포함하는 방법.

청구항 19

제11항에 있어서,

상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 인식하는 동작; 및

상기 음성에 대응하는 기능을 수행하는 동작;을 더 포함하는 방법.

청구항 20

지정된 방향을 향하도록 배치된 복수의 마이크를 포함하는 마이크 어레이를 이용하여 음성을 수신하는 동작;

전자 장치 주변에 위치하는 사용자를 감지하는 동작;

상기 전자 장치 주변에 복수의 사용자가 존재하면 상기 복수의 사용자 중 하나의 사용자를 선택하는 동작;

상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하는 동작; 및

나머지 방향으로부터 수신되는 음성을 노이즈로 처리하는 동작;을 포함하는 프로그램이 기록된 컴퓨터 판독 가능 기록매체.

발명의 설명

기 술 분 야

[0001] 본 발명은 사용자로부터 수신되는 음성을 처리하는 방법 및 장치에 관한 것이다.

배 경 기 술

[0002] 전자 기술의 발달에 힘입어 다양한 유형의 전자 제품들이 개발 및 보급되고 있다. 스마트폰, 테블릿 PC 등과 같이 다양한 기능을 가지는 휴대용 전자 장치의 보급이 확대되고 있으며, 로봇과 같은 고성능의 전자 장치가 개발되고 있다.

[0003] 사용자들은 상술한 전자 장치들을 이용하여 이메일, 웹서핑, 사진 촬영, 인스턴트 메시지, 일정 관리, 비디오 재생, 오디오 재생 등 다양한 서비스를 제공받을 수 있다. 최근에는 사용자의 음성을 인식하고 인식된 사용자 음성을 이용하여 다양한 서비스를 제공하는 기술이 개발되고 있다.

발명의 내용

해결하려는 과제

[0004] 전자 장치가 마이크를 이용하여 사용자의 음성을 수신할 때 사용자의 음성뿐만 아니라 전자 장치 주변에서 발생하는 다양한 소음이 함께 수신될 수 있다. 또한, 사용자가 직접 발화한 음성뿐만 아니라 TV, 라디오 등의 장치에서 출력되는 음성에 의해 사용자 음성의 인식에 방해가되거나 잘못된 동작을 수행하는 경우가 발생할 수 있다.

[0005] 본 발명의 다양한 실시 예는, 전자 장치 주변에서 발생하는 다양한 소음을 제거하여 노이즈가 적은 사용자 음성을 획득하고, 사용자가 실제로 존재하는 상태에서 입력되는 음성만을 사용자 음성으로 처리하여 음성 인식 성능이 개선된 음성 처리 방법 및 장치를 제공할 수 있다.

과제의 해결 수단

[0006] 본 발명의 다양한 실시 예에 따른 전자 장치는, 지정된 방향을 향하도록 배치된 복수의 마이크를 포함하는 마이크 어레이(microphone array), 상기 전자 장치 주변에 위치하는 사용자를 감지하는 센서 모듈 및 상기 전자 장치 주변에 복수의 사용자가 존재하면 상기 복수의 사용자 중 하나의 사용자를 선택하고, 상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하고 나머지 방향으로부터 수신되는 음성을 노이즈로 처리하도록 설정된 프로세서를 포함할 수 있다.

[0007] 본 발명의 다양한 실시 예에 따른 전자 장치의 음성 처리 방법은, 상기 전자 장치 주변에 위치하는 사용자를 감지하는 동작, 지정된 방향을 향하도록 배치된 복수의 마이크를 포함하는 마이크 어레이를 이용하여 음성을 수신하는 동작, 상기 전자 장치 주변에 복수의 사용자가 존재하면 상기 복수의 사용자 중 하나의 사용자를 선택하는 동작, 상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하는 동작 및 나머지 방향으로부터 수신되는 음성을 노이즈로 처리하는 동작을 포함할 수 있다.

[0008] 본 발명의 다양한 실시 예에 따른 컴퓨터 판독 가능 기록매체는, 지정된 방향을 향하도록 배치된 복수의 마이크를 포함하는 마이크 어레이를 이용하여 음성을 수신하는 동작, 상기 전자 장치 주변에 위치하는 사용자를 감지하는 동작, 상기 전자 장치 주변에 복수의 사용자가 존재하면 상기 복수의 사용자 중 하나의 사용자를 선택하는 동작, 상기 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하는 동작 및 나머지 방향으로부터 수신되는 음성을 노이즈로 처리하는 동작을 포함하는 프로그램이 기록될 수 있다.

발명의 효과

[0009] 본 발명의 다양한 실시 예에 따르면, 전자 장치는 사용자가 발화한 음성 및 다른 장치에서 출력되는 음성을 정확하게 구분하여 오동작을 방지할 수 있으며, 사용자 음성에 포함된 노이즈를 제거하여 음성 인식 성능을 향상시킬 수 있다.

도면의 간단한 설명

[0010] 도 1은 본 발명의 다양한 실시 예에 따른 전자 장치의 구성을 나타내는 블럭도이다.

도 2는 본 발명의 다양한 실시 예에 따른 마이크의 배치 상태를 나타내는 도면이다.

도 3은 본 발명의 다양한 실시 예에 따른 마이크의 배치 상태를 나타내는 도면이다.

도 4는 본 발명의 다양한 실시 예에 따른 마이크의 배치 상태를 나타내는 도면이다.

도 5는 본 발명의 다양한 실시 예에 따른 사용자 인터페이스를 나타내는 도면이다.

도 6은 본 발명의 다양한 실시 예에 따른 전자 장치의 음성 처리 방법을 나타내는 흐름도이다.

도 7은 본 발명의 다양한 실시 예에 따른 전자 장치의 음성 처리 방법을 나타내는 흐름도이다.

도 8은 본 발명의 다양한 실시예에 따른 전자 장치의 예를 도시한다.

도 9는 본 발명의 다양한 실시 예에 따른 전자 장치의 구현예를 나타내는 도면이다.

도 10은 본 발명의 다양한 실시 예에 따른 네트워크 환경 내의 전자 장치를 나타내는 도면이다.

도 11은 다양한 실시 예에 따른, 전자 장치를 도시하는 블럭도이다.

도 12는 다양한 실시 예에 따른, 전자 장치를 도시하는 블럭도이다.

도 13은 본 발명의 다양한 실시예에 따른 전자 장치의 소프트웨어 블록도이다.

발명을 실시하기 위한 구체적인 내용

[0011] 이하, 본 문서의 다양한 실시예가 첨부된 도면을 참조하여 기재된다. 그러나, 이는 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 문서의 실시예의 다양한 변경(modifications), 균등물(equivalents), 및/또는 대체물(alternatives)을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다.

[0012] 본 문서에서, "가진다," "가질 수 있다," "포함한다," 또는 "포함할 수 있다" 등의 표현은 해당 특징(예: 수치, 기능, 동작, 또는 부품 등의 구성요소)의 존재를 가리키며, 추가적인 특징의 존재를 배제하지 않는다.

[0013] 본 문서에서, "A 또는 B," "A 또는/및 B 중 적어도 하나," 또는 "A 또는/및 B 중 하나 또는 그 이상"등의 표현은 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. 예를 들면, "A 또는 B," "A 및 B 중 적어도 하나," 또는 "A 또는 B 중 적어도 하나"는, (1) 적어도 하나의 A를 포함, (2) 적어도 하나의 B를 포함, 또는 (3) 적어도 하나의 A 및 적어도 하나의 B 모두를 포함하는 경우를 모두 지칭할 수 있다.

[0014] 본 문서에서 사용된 "제 1," "제 2," "첫째," 또는 "둘째,"등의 표현들은 다양한 구성요소들을, 순서 및/또는 중요도에 상관없이 수식할 수 있고, 한 구성요소를 다른 구성요소와 구분하기 위해 사용될 뿐 해당 구성요소들을 한정하지 않는다. 예를 들면, 제 1 사용자 기기와 제 2 사용자 기기는, 순서 또는 중요도와 무관하게, 서로 다른 사용자 기기를 나타낼 수 있다. 예를 들면, 본 문서에 기재된 권리 범위를 벗어나지 않으면서 제 1 구성요소는 제 2 구성요소로 명명될 수 있고, 유사하게 제 2 구성요소도 제 1 구성요소로 바꾸어 명명될 수 있다.

[0015] 어떤 구성요소(예: 제 1 구성요소)가 다른 구성요소(예: 제 2 구성요소)에 "(기능적으로 또는 통신적으로) 연결되어((operatively or communicatively) coupled with/to)" 있다거나 "접속되어(connected to)" 있다고 언급된 때에는, 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로 연결되거나, 다른 구성요소(예: 제 3 구성요소)를 통하여 연결될 수 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소(예: 제 1 구성요소)가 다른 구성요소(예: 제 2 구성요소)에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 상기 어떤 구성요소와 상기 다른 구성요소 사이에 다른 구성요소(예: 제 3 구성요소)가 존재하지 않는 것으로 이해될

수 있다.

[0016] 본 문서에서 사용된 표현 "~하도록 구성된(또는 설정된)(configured to)"은 상황에 따라, 예를 들면, "~에 적합한(suitable for)," "~하는 능력을 가지는(having the capacity to)," "~하도록 설계된(designed to)," "~하도록 변경된(adapted to)," "~하도록 만들어진(made to)," 또는 "~를 할 수 있는(capable of)"과 바꾸어 사용될 수 있다. 용어 "~하도록 구성된(또는 설정된)"은 하드웨어적으로 "특별히 설계된(specifically designed to)" 것만을 반드시 의미하지 않을 수 있다. 대신, 어떤 상황에서는, "~하도록 구성된 장치"라는 표현은, 그 장치가 다른 장치 또는 부품들과 함께 "~할 수 있는" 것을 의미할 수 있다. 예를 들면, 문구 "A, B, 및 C를 수행하도록 구성된(또는 설정된) 프로세서"는 해당 동작을 수행하기 위한 전용 프로세서(예: 임베디드 프로세서), 또는 메모리 장치에 저장된 하나 이상의 소프트웨어 프로그램들을 실행함으로써, 해당 동작들을 수행할 수 있는 범용 프로세서(generic-purpose processor)(예: CPU 또는 application processor)를 의미할 수 있다.

[0017] 본 문서에서 사용된 용어들은 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 다른 실시예의 범위를 한정하려는 의도가 아닐 수 있다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 용어들은 본 문서에 기재된 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가질 수 있다. 본 문서에 사용된 용어들 중 일반적인 사전에 정의된 용어들은, 관련 기술의 문맥상 가지는 의미와 동일 또는 유사한 의미로 해석될 수 있으며, 본 문서에서 명백하게 정의되지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다. 경우에 따라서, 본 문서에서 정의된 용어일지라도 본 문서의 실시예들을 배제하도록 해석될 수 없다.

[0018] 본 문서의 다양한 실시예들에 따른 전자 장치는, 예를 들면, 스마트폰(smartphone), 태블릿 PC(tablet personal computer), 이동 전화기(mobile phone), 영상 전화기, 전자책 리더기(e-book reader), 데스크탑 PC(desktop personal computer), 랩탑 PC(laptop personal computer), 넷북 컴퓨터(netbook computer), 워크스테이션(workstation), 서버, PDA(personal digital assistant), PMP(portable multimedia player), MP3 플레이어, 모바일 의료기기, 카메라(camera), 또는 웨어러블 장치(wearable device) 중 적어도 하나를 포함할 수 있다. 다양한 실시예에 따르면, 웨어러블 장치는 액세서리형(예: 시계, 반지, 팔찌, 발찌, 목걸이, 안경, 콘택트 렌즈, 또는 머리 착용형 장치(head-mounted-device(HMD)), 직물 또는 의류 일체형(예: 전자 의복), 신체 부착형(예: 스킨 패드(skin pad) 또는 문신), 또는 생체 이식형(예: implantable circuit) 중 적어도 하나를 포함할 수 있다.

[0019] 어떤 실시예들에서, 전자 장치는 가전 제품(home appliance)일 수 있다. 가전 제품은, 예를 들면, 텔레비전, DVD(digital video disk) 플레이어, 오디오, 냉장고, 에어컨, 청소기, 오븐, 전자레인지, 세탁기, 공기 청정기, 셋톱 박스(set-top box), 홈 오토매이션 컨트롤 패널(home automation control panel), 보안 컨트롤 패널(security control panel), TV 박스(예: 삼성 HomeSync™, 애플TV™, 또는 구글 TV™), 게임 콘솔(예: Xbox™, PlayStation™), 전자 사전, 전자 키, 캠코더(camcorder), 또는 전자 액자 중 적어도 하나를 포함할 수 있다.

[0020] 이하, 첨부 도면을 참조하여, 다양한 실시예에 따른 전자 장치가 설명된다. 본 문서에서, 사용자라는 용어는 전자 장치를 사용하는 사람 또는 전자 장치를 사용하는 장치(예: 인공지능 전자 장치)를 지칭할 수 있다.

[0021] 도 1은 본 발명의 다양한 실시 예에 따른 전자 장치의 구성을 나타내는 블럭도이다.

[0022] 도 1을 참조하면, 전자 장치(100)는 마이크 어레이(microphone array)(110), 센서 모듈(120), 통신 모듈(130), 디스플레이(140), 스피커(150), 메모리(160) 및 프로세서(170)를 포함할 수 있다.

[0023] 일 실시 예에 따르면, 마이크 어레이(110)는 지정된 방향을 향하도록 배치된 복수의 마이크를 포함할 수 있다. 일 실시 예에 따르면, 마이크 어레이(110)에 포함된 복수의 마이크는 서로 상이한 방향을 향할 수 있다. 일 실시 예에 따르면, 마이크 어레이(110)에 포함된 복수의 마이크는 사운드(예: 음성)를 수신하고, 수신된 사운드를 전기적 신호(또는, 음성 신호)로 변환할 수 있다. 일 실시 예에 따르면, 마이크 어레이(110)는 음성 신호를 프로세서(170)로 전달할 수 있다.

[0024] 일 실시 예에 따르면, 센서 모듈(120)은 전자 장치 주변에 위치하는 사용자를 감지할 수 있다. 예를 들어, 센서 모듈(120)은 PIR(passive infrared) 센서, 근접 센서, UWB(ultra wide band) 센서, 초음파 센서, 이미지 센서 또는 열감지 센서 등을 포함할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 복수의 센서 모듈(120)을 포함할 수 있다. 일 실시 예에 따르면, 복수의 센서 모듈(120) 각각은 지정된 영역에 사용자의 존재 여부, 사용자의 거리 및 사용자의 방향을 감지할 수 있다. 예를 들어, 복수의 센서 모듈(120) 각각은 마이크 어레이(110)에 포함된 복수의 마이크가 향하는 방향에 대응하여 사용자의 존재 여부를 감지할 수 있다.

[0025] 일 실시 예에 따르면, 센서 모듈(120)은 제1 센서(121) 및 제2 센서(123)를 포함할 수 있다. 일 실시 예에 따르면, 제1 센서(121)는 사용자의 인체를 감지할 수 있다. 예를 들어, 제1 센서(121)는 지정된 방향 범위에 사용자의 인체가 존재하는지 여부를 감지할 수 있다. 제1 센서(121)는, 예를 들어, PIR 센서, UWB 센서 및 열(에: 체온)감지 센서를 포함할 수 있다. 예를 들어, PIR 센서는 사용자의 인체로부터 수신되는 적외선의 변화량을 이용하여 사용자의 존재 여부를 감지할 수 있다. 일 실시 예에 따르면, 제2 센서(123)는 지정된 방향 범위에 위치하는 물체(또는, 인체)의 구체적인 방향 또는 거리를 감지할 수 있다. 제2 센서는, 예를 들어, 초음파 센서, 근접 센서 및 레이더를 포함할 수 있다. 예를 들어, 초음파 센서는 지정된 방향으로 초음파를 송신하고 물체에 반사되어 수신되는 초음파에 기초하여 물체의 구체적인 방향 또는 거리를 감지할 수 있다.

[0026] 일 실시 예에 따르면, 통신 모듈(130)은 외부 전자 장치(예: 음성 인식 서버)와 통신할 수 있다. 일 실시 예에 따르면, 통신 모듈(130)은 RF 모듈, 셀룰러 모듈, Wi-Fi(wirless-fidelity) 모듈, GNSS(global navigation satellite system) 모듈, 블루투스 모듈 또는 NFC 모듈을 포함할 수 있다. 전자 장치(100)는 상술한 모듈들 중 적어도 하나의 모듈을 통해, 예를 들면, 네트워크(예: 인터넷 망 또는 이동통신망)에 연결되어 외부 전자 장치와 통신할 수 있다.

[0027] 일 실시 예에 따르면, 디스플레이(140)는 사용자 인터페이스(또는, 컨텐츠)를 표시할 수 있다. 일 실시 예에 따르면, 디스플레이(140)는 사용자의 음성에 대응되는 피드백 정보를 표시할 수 있다. 일 실시 예에 따르면, 디스플레이(140)는 사용자 음성에 따라 사용자 인터페이스 또는 컨텐츠를 변경하여 표시할 수 있다.

[0028] 일 실시 예에 따르면, 스피커(150)는 오디오를 출력할 수 있다. 일 실시 예에 다르면, 스피커(150)는 사용자의 음성에 대응되는 음성 피드백을 출력할 수 있다.

[0029] 일 실시 예에 따르면, 메모리(160)는 사용자 음성을 인식하기 위한 데이터를 저장할 수 있다. 일 실시 예에 따르면, 메모리(160)는 사용자 음성에 대한 피드백을 제공하기 위한 데이터를 저장할 수 있다. 일 실시 예에 따르면, 메모리(160)는 사용자 정보를 저장할 수 있다. 예를 들어, 메모리(160)는 사용자의 음성을 식별하기 위한 정보를 저장할 수 잇다.

[0030] 일 실시 예에 따르면, 프로세서(170)는 전자 장치(100)의 전반적인 동작을 제어할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 마이크 어레이(110), 센서 모듈(120), 통신 모듈(130), 디스플레이(140), 스피커(150) 또는 메모리(160) 각각을 제어하여 본 발명의 다양한 실시 예에 따라 사용자의 음성을 인식하고 처리할 수 있다. 일 실시 예에 따르면, 프로세서(170)(예: 어플리케이션 프로세서)는 CPU(central processing unit), GPU(graphic processing unit), 메모리 등을 포함하는 SoC(system on chip)으로 구현될 수 있다.

[0031] 일 실시 예에 따르면, 프로세서(170)는 센서 모듈(120)로부터 수신되는 정보를 이용하여 전자 장치(100) 주변에 사용자의 존재 여부 및 사용자가 위치하는 방향을 판단할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 제1 센서(121) 및 제2 센서(123) 중 적어도 하나를 이용하여 사용자가 존재하는지 판단할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 전자 장치(100) 주변에 사용자가 감지되지 않는 상태에서는 제1 센서(121)만을 활성화시킬 수 있다. 일 실시 예에 따르면, 프로세서(170)는 제1 센서(121)가 활성화된 상태에서 제1 센서(121)에 의해 사용자의 인체가 감지되면 제2 센서(123)를 활성화시킬 수 있다. 일 실시 예에 따르면, 프로세서(170)는 제1 센서(121)에 의해 사용자의 인체가 감지되면 즉시 또는 지정된 시간이 경과한 후 제1 센서(121)를 비활성화 시킬 수 있다. 일 실시 예에 따르면, 프로세서(170)는 제2 센서(123)가 활성화된 상태에서 제2 센서(123)에 의해 사용자가 감지되지 않으면, 제1 센서(121)를 활성화시킬 수 있다. 일 실시 예에 따르면, 프로세서(170)는 제2 센서(123)가 활성화된 상태에서 제2 센서(123)에 의해 사용자가 감지되지 않으면, 즉시 또는 지정된 시간이 경과한 후 제2 센서(123)를 비활성화 시킬 수 있다.

[0032] 일 실시 예에 따르면, 프로세서(170)는 마이크 어레이(110)로부터 수신되는 음성 신호를 처리할 수 있다. 이하에서 도 2 내지 도 4를 참조하여 전자 장치 주변에 위치하는 사용자에 따른 음성 처리 방법에 대해 설명한다.

[0033] 도 2는 본 발명의 다양한 실시 예에 따른 마이크의 배치 상태를 나타내는 도면이다.

[0034] 도 2를 참조하면, 전자 장치(100)는 마이크 어레이에 포함된 복수의 마이크(111, 112, ..., 118)를 포함할 수 있다. 복수의 마이크(111, 112, ..., 118) 각각은 서로 상이한 방향을 향해 배치될 수 있다.

[0035] 일 실시 예에 따르면, 프로세서(170)는 복수의 마이크(111, 112, ..., 118)에 수신되는 음성 중 지정된 방향으로부터 수신되는 음성을 사용자 입력으로 처리하고, 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 프로세서(170)는 복수의 마이크(111, 112, ..., 118) 중 일부를 선택하고, 선택된 마이크로부

터 수신되는 음성 신호(또는, 제1 음성 신호)를 사용자 입력으로 처리하고, 선택되지 않은 마이크로부터 수신되는 음성 신호(또는, 제2 음성 신호)를 노이즈로 처리할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 제2 음성 신호를 이용하여 제1 음성 신호에 노이즈 캔슬링을 수행할 수 있다. 예를 들어, 프로세서(170)는 제2 음성 신호를 반전시켜 제2 음성 신호의 역상 신호를 생성하고, 제1 음성 신호 및 역상 신호를 합성할 수 있다.

[0036] 도 3은 본 발명의 다양한 실시 예에 따른 마이크의 배치 상태를 나타내는 도면이다.

[0037] 도 3을 참조하면, 전자 장치(100)는 마이크 어레이에 포함된 복수의 마이크(111, 112, ..., 118)를 포함할 수 있다. 복수의 마이크(111, 112, ..., 118) 각각은 서로 상이한 방향을 향해 배치될 수 있다.

[0038] 일 실시 예에 따르면, 프로세서(170)는 복수의 마이크(111, 112, ..., 118)에 수신되는 음성 중 사용자(31)가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하고, 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 프로세서(170)는 복수의 마이크(111, 112, ..., 118) 중 사용자(31)가 위치하는 방향을 향하는 제3 마이크(113) 및 제4 마이크(114)를 선택할 수 있다. 프로세서(170)는 제3 마이크(113) 및 제4 마이크(114)로부터 수신되는 음성 신호를 사용자 입력으로 처리하고, 선택되지 않은 마이크(111, 112, 115, 116, 117, 118)로부터 수신되는 음성 신호를 노이즈로 처리할 수 있다.

[0039] 일 실시 예에 따르면, 프로세서(170)는 선택되지 않은 마이크(111, 112, 115, 116, 117, 118)로부터 수신되는 음성 신호를 이용하여 제3 마이크(113) 및 제4 마이크(114)로부터 수신되는 음성 신호에 노이즈 캔슬링을 수행할 수 있다. 예를 들어, 프로세서(170)는 선택되지 않은 마이크(111, 112, 115, 116, 117, 118)로부터 수신되는 음성 신호를 반전시켜 역상 신호를 생성하고, 제3 마이크(113) 및 제4 마이크(114)로부터 수신되는 및 역상 신호를 합성할 수 있다.

[0040] 도 4는 본 발명의 다양한 실시 예에 따른 마이크의 배치 상태를 나타내는 도면이다.

[0041] 도 4를 참조하면, 전자 장치(100)는 마이크 어레이에 포함된 복수의 마이크(111, 112, ..., 118)를 포함할 수 있다. 복수의 마이크(111, 112, ..., 118) 각각은 서로 상이한 방향을 향해 배치될 수 있다.

[0042] 일 실시 예에 따르면, 프로세서(170)는 전자 장치 주변에 복수의 사용자(예: 제1 사용자(41) 및 제2 사용자(43))가 존재하면, 복수의 사용자(41, 43)가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하고, 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 프로세서(170)는 복수의 마이크(111, 112, ..., 118) 중 복수의 사용자(41, 43)가 위치하는 방향을 향하는 제1 마이크(111), 제3 마이크(113) 및 제4 마이크(114)를 선택할 수 있다. 프로세서(170)는 제1 마이크(111), 제3 마이크(113) 및 제4 마이크(114)로부터 수신되는 음성 신호를 사용자 입력으로 처리하고, 선택되지 않은 마이크(112, 115, 116, 117, 118)로부터 수신되는 음성 신호를 노이즈로 처리할 수 있다.

[0043] 일 실시 예에 따르면, 프로세서(170)는 전자 장치 주변에 복수의 사용자(예: 제1 사용자(41) 및 제2 사용자(43))가 존재하면, 복수의 사용자 중 하나를 선택할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 복수의 마이크(111, 112, ..., 118)에 수신되는 음성 중 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리하고, 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 프로세서(170)는 제1 사용자(41)가 선택되면 제1 사용자(41)가 위치하는 방향을 향하는 제3 마이크(113) 및 제4 마이크(114)로부터 수신되는 음성 신호를 사용자 입력으로 처리하고, 나머지 마이크(111, 112, 115, 116, 117, 118)로부터 수신되는 음성 신호를 노이즈로 처리할 수 있다. 다른 예를 들어, 프로세서(170)는 제2 사용자(43)가 선택되면 제2 사용자(43)가 위치하는 방향을 향하는 제1 마이크(111)로부터 수신되는 음성 신호를 사용자 입력으로 처리하고, 나머지 마이크(112, 113, 114, 115, 116, 117, 118)로부터 수신되는 음성 신호를 노이즈로 처리할 수 있다.

[0044] 일 실시 예에 따르면, 프로세서(170)는 적어도 하나의 마이크에 수신되는 음성 신호를 이용하여 사용자를 식별할 수 있다. 예를 들어, 프로세서(170)는 적어도 하나의 마이크에 수신되는 음성 신호의 특성을 분석하여 제1 사용자 및 제2 사용자를 식별할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 적어도 하나의 마이크에 수신되는 음성 신호를 메모리(160)에 저장된 음성 신호와 비교하여 사용자를 식별할 수 있다.

[0045] 일 실시 예에 따르면, 프로세서(170)는 적어도 하나의 마이크를 통해 수신되는 음성 신호를 이용하여 음성이 발화된 방향(또는, 사용자가 위치하는 방향)을 판단할 수 있다. 예를 들어, 제1 사용자가(41) 발화한 음성이 복수의 마이크 중 적어도 일부에 수신되면 적어도 하나의 마이크에 수신된 음성의 레벨(또는, 크기)에 기초하여 제1 사용자(41)의 음성이 제3 마이크(113) 및 제4 마이크(114)가 향하는 방향으로부터 발화되었다고 판단할 수 있다. 다른 예를 들어, 제2 사용자가(43) 발화한 음성이 복수의 마이크 중 적어도 일부에 수신되면 적어도 하나

의 마이크에 수신된 음성의 레벨(또는, 크기)에 기초하여 제2 사용자(43)의 음성이 제1 마이크(111)가 향하는 방향으로부터 발화되었다고 판단할 수 있다.

[0046] 일 실시 예에 따르면, 프로세서(170)는 전자 장치 주변에 복수의 사용자가 존재하면, 복수의 사용자 각각에 대해 우선순위를 판단할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 복수의 사용자 각각의 대화 이력(예: 대화 횟수, 시간, 대화 내용 등)에 기초하여 복수의 사용자 각각의 친밀도를 판단할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 복수의 사용자 각각의 친밀도에 따라 복수의 사용자 각각의 우선 순위를 판단할 수 있다.

[0047] 일 실시 예에 따르면, 프로세서(170)는 지정된 명령어가 수신되면 복수의 사용자 중 어떠한 사용자가 지정된 명령어를 발화하였는지 판단할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 전자 장치(100) 주변에 복수의 사용자(예: 제1 사용자(41) 및 제2 사용자(43))가 존재하면, 복수의 사용자 중 지정된 명령어를 먼저 발화한 사용자를 선택할 수 있다. 예를 들어, 제1 사용자가 지정된 명령어를 먼저 발화한 경우 프로세서(170)는 제1 사용자(41)가 위치하는 방향을 향하는 제3 마이크(113) 및 제4 마이크(114)로부터 수신되는 음성 신호를 사용자 입력으로 처리하고, 나머지 마이크(111, 112, 115, 116, 117, 118)로부터 수신되는 음성 신호를 노이즈로 처리할 수 있다.

[0048] 일 실시 예에 따르면, 프로세서(170)는 전자 장치(100) 주변에 복수의 사용자(예: 제1 사용자(41) 및 제2 사용자(43))가 존재하면, 복수의 사용자 중 우선순위가 높은 사용자를 선택할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 우선순위가 가장 높은 사용자의 발화가 종료되면 다음 우선순위를 가지는 사용자를 선택할 수 있다. 예를 들어, 프로세서(170)는 지정된 시간 동안 선택된 사용자로부터 음성이 발화되지 않으면 선택된 사용자의 발화가 종료되었다고 판단하고 다른 사용자를 선택할 수 있다.

[0049] 일 실시 예에 따르면, 프로세서(170)는 노이즈 캔슬링이 수행된 음성 신호를 이용하여 음성 인식을 수행할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 음성 신호를 텍스트로 변환할 수 있다. 예를 들어, 프로세서(170)는 STT(speech to text) 알고리즘을 이용하여 음성 신호를 텍스트로 변환할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 텍스트를 분석하여 사용자의 의도를 파악할 수 있다. 예를 들어, 프로세서(170)는 텍스트를 이용하여 자연어 이해(natural language understanding : NLU) 및 대화 관리(dialog management : DM)를 수행할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 인식된 음성에 포함된 사용자의 의도에 대응되는 정보(이하, 피드백 정보)를 검색 또는 생성할 수 있다. 피드백 정보는, 예를 들어, 텍스트, 오디오, 이미지 등 다양한 형태의 컨텐츠를 포함할 수 있다.

[0050] 일 실시 예에 따르면, 상술한 음성 인식 및 피드백 제공 과정 중 적어도 일부는 적어도 하나의 외부 전자 장치(예: 서버)에 의해 수행될 수 있다. 예를 들어, 프로세서(170)는 노이즈 캔슬링이 수행된 음성 신호를 외부 서버로 전송하고, 외부 서버로부터 음성 신호에 대응되는 텍스트를 수신할 수 있다. 다른 예를 들어, 프로세서(170)는 텍스트를 외부 서버로 전송하고, 외부 서버로부터 텍스트에 대응되는 피드백 정보를 수신할 수 있다.

[0051] 일 실시 예에 따르면, 프로세서(170)는 전자 장치(100) 주변에 위치하는 복수의 사용자 중 어떠한 사용자가 선택되었는지(또는, 어떠한 사용자의 음성을 인식하고 있는지) 표시할 수 있다. 예를 들어, 전자 장치(100)는 복수의 마이크(111, 112, ..., 118)가 향하는 방향에 대응되도록 배치된 복수의 발광소자(LED)를 포함할 수 있으며, 프로세서(170)는 현재 선택된 사용자가 위치하는 방향에 대응되는 발광소자를 점등시킬 수 있다.

[0052] 도 5는 본 발명의 다양한 실시 예에 따른 사용자 인터페이스를 나타내는 도면이다.

[0053] 일 실시 예에 따르면, 프로세서(170)는 디스플레이(140)에 전자 장치(100) 주변에 위치하는 복수의 사용자 중 어떠한 사용자가 선택되었는지 나타내는 사용자 인터페이스를 표시할 수 있다. 도 5를 참조하면 사용자 인터페이스는 전자 장치를 나타내는 제1 오브젝트(50), 제1 사용자를 나타내는 제2 오브젝트(51), 제2 사용자를 나타내는 제3 오브젝트(53)를 포함할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 센서 모듈(110)에 의해 제1 사용자 및 제2 사용자가 감지되면 디스플레이(140)에 감지된 사용자에 대응하는 제2 오브젝트(51) 및 제3 오브젝트(53)를 표시할 수 있다. 프로세서(170)는 사용자가 이동하는 경우 사용자의 이동에 대응되도록 제1 오브젝트 및 제3 오브젝트의 위치를 변경하여 표시할 수 있다.

[0054] 도 5를 참조하면 사용자 인터페이스는 제1 사용자의 음성을 인식할 수 있는 영역을 나타내는 제4 오브젝트(55), 제2 사용자의 음성을 인삭할 수 있는 영역을 나타내는 제5 오브젝트(57)를 포함할 수 있다. 음성 인식 가능한 영역은 사용자의 위치에 의해 결정될 수 있으며 사용자의 위치가 변경되면 음성 인식 가능한 영역도 변경될 수 있다.

[0055] 일 실시 예에 따르면, 프로세서(170)는 전자 장치(100) 주변에 위치하는 복수의 사용자 중 현재 선택된 사용자

(또는, 현재 음성 인식 중인 사용자)를 알 수 있도록 사용자 인터페이스를 표시할 수 있다. 예를 들어, 프로세서(170)는 제1 사용자가 선택중인 경우 제4 오브젝트(55)의 색상, 투명도를 제5 오브젝트와 상이하게 표시하거나 또는 제4 오브젝트(55)를 깜빡이도록 할 수 있다. 다른 예를 들어, 프로세서(170)는 현재 선택된 사용자를 지시하는 별도의 오브젝트를 표시할 수도 있다.

[0056] 일 실시 예에 따르면, 프로세서(170)는 인식된 음성에 대한 피드백을 제공할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 피드백 정보를 디스플레이(140)에 표시할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 피드백 정보를 스피커(150)를 통해 출력할 수 있다. 일 실시 예에 따르면, 프로세서(170)는 텍스트 형태의 피드백 정보가 수신되면 TTS(text to speech) 알고리즘을 이용하여 텍스트를 음성 형태로 변환하고, 음성 형태의 피드백 정보를 스피커(150)를 통해 출력할 수 있다.

[0057] 일 실시 예에 따르면, 프로세서(170)는 인식된 음성에 대응되는 기능을 수행할 수 있다. 일 실시 예에 따르면, 인식된 음성에 포함된 사용자의 의도에 대응되는 기능을 수행할 수 있다. 예를 들어, 프로세서(170)는 사용자의 의도에 따라 지정된 소프트웨어를 실행하거나 또는 디스플레이(140)에 표시된 사용자 인터페이스를 변경할 수 있다.

[0058] 도 6은 본 발명의 다양한 실시 예에 따른 전자 장치의 음성 처리 방법을 나타내는 흐름도이다.

[0059] 도 6에 도시된 흐름도는 도 1에 도시된 전자 장치(100)에서 처리되는 동작들로 구성될 수 있다. 따라서, 이하에서 생략된 내용이라 하더라도 도 1 내지 도 5를 참조하여 전자 장치(100)에 관하여 기술된 내용은 도 6에 도시된 흐름도에도 적용될 수 있다.

[0060] 일 실시 예에 따르면, 610 동작에서, 전자 장치(100)는 센서 모듈을 이용하여 전자 장치(100) 주변에 위치하는 사용자를 감지할 수 있다. 예를 들어, 전자 장치(100)는 센서 모듈을 이용하여 전자 장치(100) 주변에 사용자의 존재 여부 및 사용자가 위치하는 방향을 판단할 수 있다. 일 실시 예에 따르면, 센서 모듈은 제1 센서 및 제2 센서를 포함할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 일 실시 예에 따르면, 전자 장치(100)는 제1 센서를 이용하여 지정된 방향 범위에 사용자의 인체가 존재하는지 여부를 감지할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 제2 센서를 이용하여 지정된 방향 범위에 위치하는 물체(또는, 사용자)의 구체적인 방향 또는 거리를 감지할 수 있다. 일 실시 예에 따르면, 제2 센서(123)는 인체 여부와 관계없이 지정된 방향 범위에 위치하는 물체(또는, 사용자)를 감지할 수 있다.

[0061] 일 실시 예에 따르면, 전자 장치(100)는 제1 센서 및 제2 센서 중 적어도 하나를 이용하여 사용자가 존재하는지 판단할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 전자 장치(100) 주변에 사용자가 감지되지 않는 상태에서는 제1 센서만을 활성화시킬 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 제1 센서가 활성화된 상태에서 제1 센서에 의해 사용자의 인체가 감지되면 제2 센서를 활성화시킬 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 제1 센서에 의해 사용자의 인체가 감지되면 즉시 또는 지정된 시간이 경과한 후 제1 센서를 비활성화시킬 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 제2 센서가 활성화된 상태에서 제2 센서에 의해 사용자가 감지되지 않으면, 제1 센서를 활성화시킬 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 제2 센서가 활성화된 상태에서 제2 센서에 의해 사용자가 감지되지 않으면, 즉시 또는 지정된 시간이 경과한 후 제2 센서를 비활성화 시킬 수 있다.

[0062] 일 실시 예에 따르면, 620 동작에서, 전자 장치(100)는 마이크 어레이를 이용하여 음성을 수신할 수 있다. 일 실시 예에 따르면, 마이크 어레이는 지정된 방향을 향하도록 배치된 복수의 마이크를 포함할 수 있다. 일 실시 예에 따르면, 마이크 어레이에 포함된 복수의 마이크는 서로 상이한 방향을 향할 수 있다.

[0063] 일 실시 예에 따르면, 630 동작에서, 전자 장치(100)는 복수의 사용자가 감지되었는지 판단할 수 있다.

[0064] 일 실시 예에 따르면, 전자 장치(100)는 630 동작에서 복수의 사용자가 감지되면, 640 동작에서, 복수의 사용자 중 하나의 사용자를 선택할 수 있다.

[0065] 일 실시 예에 따르면, 전자 장치(100)는 전자 장치(100) 주변에 복수의 사용자가 존재하면, 복수의 사용자 중 지정된 명령어를 먼저 발화한 사용자를 선택할 수 있다. 예를 들어, 전자 장치(100)는 적어도 하나의 마이크에 수신되는 음성 신호를 이용하여 복수의 사용자 각각을 식별할 수 있다. 전자 장치(100)는 적어도 하나의 마이크를 통해 수신되는 음성 신호를 이용하여 음성이 발화된 방향(또는, 사용자가 위치하는 방향)을 판단할 수 있다. 전자 장치(100)는 지정된 명령어가 수신되면 복수의 사용자 중 어떠한 사용자가 지정된 명령어를 발화하였는지 판단할 수 있다.

[0066] 일 실시 예에 따르면, 전자 장치(100)는 전자 장치(100) 주변에 복수의 사용자(예: 제1 사용자(41) 및 제2 사용자(43))가 존재하면, 복수의 사용자 중 우선순위가 높은 사용자를 선택할 수 있다. 전자 장치(100)는, 예를 들어, 적어도 하나의 마이크에 수신되는 음성 신호를 이용하여 복수의 사용자 각각을 식별할 수 있다. 전자 장치(100)는, 예를 들어, 복수의 사용자 각각의 대화 이력(예: 대화 횟수, 시간, 대화 내용 등)에 기초하여 복수의 사용자 각각의 친밀도를 판단할 수 있다. 전자 장치(100)는, 예를 들어, 복수의 사용자 각각의 친밀도에 따라 복수의 사용자 각각의 우선 순위를 판단할 수 있다.

[0067] 일 실시 예에 따르면, 650 동작에서, 전자 장치(100)는 복수의 마이크에 수신되는 음성 중 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리할 수 있다.

[0068] 일 실시 예에 따르면, 670 동작에서, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 이용하여 선택된 사용자가 위치하는 방향으로부터 수신되는 음성에 노이즈 캔슬링을 수행할 수 있다.

[0069] 일 실시 예에 따르면, 전자 장치(100)는 우선순위가 가장 높은 사용자의 발화가 종료되면 다음 우선순위를 가지는 사용자를 선택할 수 있다. 예를 들어, 전자 장치(100)는 지정된 시간 동안 선택된 사용자로부터 음성이 발화되지 않으면 선택된 사용자의 발화가 종료되었다고 판단하고 다른 사용자를 선택할 수 있다.

[0070] 일 실시 예에 따르면, 전자 장치(100)는 630 동작에서 복수의 사용자가 감지되지 않으면(또는, 한명의 사용자만 감지되면), 660 동작에서, 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리할 수 있다.

[0071] 일 실시 예에 따르면, 670 동작에서, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 이용하여 상기 사용자가 위치하는 방향으로부터 수신되는 음성에 노이즈 캔슬링을 수행할 수 있다.

[0072] 일 실시 예에 따르면, 전자 장치(100)는 노이즈 캔슬링이 수행된 음성 신호를 이용하여 음성 인식을 수행할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 음성 신호를 텍스트로 변환할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 텍스트를 분석하여 사용자의 의도를 파악할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 인식된 음성에 포함된 사용자의 의도에 대응되는 정보(이하, 피드백 정보)를 검색 또는 생성할 수 있다. 피드백 정보는, 예를 들어, 텍스트, 오디오, 이미지 등 다양한 형태의 컨텐츠를 포함할 수 있다.

[0073] 일 실시 예에 따르면, 전자 장치(100)는 인식된 음성에 대한 피드백을 제공할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 피드백 정보를 디스플레이에 표시할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 피드백 정보를 스피커를 통해 출력할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 텍스트 형태의 피드백 정보가 수신되면 TTS(text to speech) 알고리즘을 이용하여 텍스트를 음성 형태로 변환하고, 음성 형태의 피드백 정보를 스피커를 통해 출력할 수 있다.

[0074] 일 실시 예에 따르면, 전자 장치(100)는 인식된 음성에 대응되는 기능을 수행할 수 있다. 일 실시 예에 따르면, 인식된 음성에 포함된 사용자의 의도에 대응되는 기능을 수행할 수 있다.

[0075] 도 7은 본 발명의 다양한 실시 예에 따른 전자 장치의 음성 처리 방법을 나타내는 흐름도이다.

[0076] 도 7에 도시된 흐름도는 도 1에 도시된 전자 장치(100)에서 처리되는 동작들로 구성될 수 있다. 따라서, 이하에서 생략된 내용이라 하더라도 도 1 내지 도 5를 참조하여 전자 장치(100)에 관하여 기술된 내용은 도 7에 도시된 흐름도에도 적용될 수 있다.

[0077] 일 실시 예에 따르면, 710 동작에서, 전자 장치(100)는 센서 모듈을 이용하여 전자 장치(100) 주변에 위치하는 사용자를 감지할 수 있다. 예를 들어, 전자 장치(100)는 센서 모듈을 이용하여 전자 장치(100) 주변에 사용자의 존재 여부 및 사용자가 위치하는 방향을 판단할 수 있다. 일 실시 예에 따르면, 센서 모듈은 제1 센서 및 제2 센서를 포함할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 일 실시 예에 따르면, 전자 장치(100)는 제1 센서를 이용하여 지정된 방향 범위에 사용자의 인체가 존재하는지 여부를 감지할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 제2 센서를 이용하여 지정된 방향 범위에 위치하는 물체(또는, 사용자)의 구체적인 방향 또는 거리를 감지할 수 있다. 일 실시 예에 따르면, 제2 센서(123)는 인체 여부와 관계없이 지정된 방향 범위에 위치하는 물체(또는, 사용자)를 감지할 수 있다.

[0078] 일 실시 예에 따르면, 720 동작에서, 전자 장치(100)는 복수의 사용자가 감지되었는지 판단할 수 있다.

[0079] 일 실시 예에 따르면, 전자 장치(100)는 720 동작에서 복수의 사용자가 감지된 후, 730 동작에서, 마이크 어레

이를 이용하여 음성을 수신할 수 있다. 일 실시 예에 따르면, 마이크 어레이는 지정된 방향을 향하도록 배치된 복수의 마이크를 포함할 수 있다. 일 실시 예에 따르면, 마이크 어레이에 포함된 복수의 마이크는 서로 상이한 방향을 향할 수 있다.

[0080] 일 실시 예에 따르면, 전자 장치(100)는, 740 동작에서, 복수의 사용자 중 하나의 사용자를 선택할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 전자 장치(100) 주변에 복수의 사용자가 존재하면, 복수의 사용자 중 지정된 명령어를 먼저 발화한 사용자를 선택할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 전자 장치(100) 주변에 복수의 사용자(예: 제1 사용자(41) 및 제2 사용자(43))가 존재하면, 복수의 사용자 중 우선순위가 높은 사용자를 선택할 수 있다.

[0081] 일 실시 예에 따르면, 750 동작에서, 전자 장치(100)는 복수의 마이크에 수신되는 음성 중 선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리할 수 있다.

[0082] 일 실시 예에 따르면, 780 동작에서, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 이용하여 선택된 사용자가 위치하는 방향으로부터 수신되는 음성에 노이즈 캔슬링을 수행할 수 있다.

[0083] 일 실시 예에 따르면, 전자 장치(100)는 우선순위가 가장 높은 사용자의 발화가 종료되면 다음 우선순위를 가지는 사용자를 선택할 수 있다. 예를 들어, 전자 장치(100)는 지정된 시간 동안 선택된 사용자로부터 음성이 발화되지 않으면 선택된 사용자의 발화가 종료되었다고 판단하고 다른 사용자를 선택할 수 있다.

[0084] 일 실시 예에 따르면, 전자 장치(100)는 720 동작에서 한명의 사용자만 감지된 후, 760 동작에서, 마이크 어레이를 이용하여 음성을 수신할 수 있다.

[0085] 일 실시 예에 따르면, 770 동작에서, 전자 장치(100)는 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리할 수 있다.

[0086] 일 실시 예에 따르면, 780 동작에서, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 노이즈로 처리할 수 있다. 예를 들어, 전자 장치(100)는 나머지 방향으로부터 수신되는 음성을 이용하여 상기 사용자가 위치하는 방향으로부터 수신되는 음성에 노이즈 캔슬링을 수행할 수 있다.

[0087] 일 실시 예에 따르면, 전자 장치(100)는 노이즈 캔슬링이 수행된 음성 신호를 이용하여 음성 인식을 수행할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 음성 신호를 텍스트로 변환할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 텍스트를 분석하여 사용자의 의도를 파악할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 인식된 음성에 포함된 사용자의 의도에 대응되는 정보(이하, 피드백 정보)를 검색 또는 생성할 수 있다. 피드백 정보는, 예를 들어, 텍스트, 오디오, 이미지 등 다양한 형태의 컨텐츠를 포함할 수 있다.

[0088] 일 실시 예에 따르면, 전자 장치(100)는 인식된 음성에 대한 피드백을 제공할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 피드백 정보를 디스플레이에 표시할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 피드백 정보를 스피커를 통해 출력할 수 있다. 일 실시 예에 따르면, 전자 장치(100)는 텍스트 형태의 피드백 정보가 수신되면 TTS(text to speech) 알고리즘을 이용하여 텍스트를 음성 형태로 변환하고, 음성 형태의 피드백 정보를 스피커를 통해 출력할 수 있다.

[0089] 일 실시 예에 따르면, 전자 장치(100)는 인식된 음성에 대응되는 기능을 수행할 수 있다. 일 실시 예에 따르면, 인식된 음성에 포함된 사용자의 의도에 대응되는 기능을 수행할 수 있다.

[0090] 도 8은 본 발명의 다양한 실시예에 따른 전자 장치의 예를 도시한다.

[0091] 도 8을 참조하면 전자 장치(예: 전자 장치(100))는 스탠드 얼론(standalone)형(801, 802, 803) 및 독킹 스테이션(docking station)형(804)으로 나뉠 수 있다. 스탠드 얼론(standalone)형 전자 장치(801, 802, 803)는 독립적으로 전자 장치의 모든 기능을 수행할 수 있다. 독킹 스테이션형 전자 장치(804)는 기능적으로 분리된 두 개 이상의 전자 장치가 하나로 결합되어 전자 장치의 모든 기능을 수행할 수 있다. 예를 들어, 독킹 스테이션형 전자 장치(804)는 본체(804a)(예: HMD(head mount display) 장치) 및 구동부(804b)를 포함하며 본체(804a)가 독킹 스테이션(구동부)에 장착되어 원하는 위치로 이동할 수 있다.

[0092] 전자 장치는 이동 여부에 따라 고정형(801) 및 이동형(802, 803, 804)으로 분류될 수 있다. 고정형 전자 장치(801)는 구동부가 없으므로 자율적으로 이동할 수 없다. 이동형 전자 장치(802, 803, 804)는 구동부를 포함하며 자율적으로 원하는 위치로 이동할 수 있다. 이동형 전자 장치(802, 803, 804)는 구동부로 휠(wheel), 캐터필러

(caterpillar) 또는 레그(leg)를 포함할 수 있다. 또한, 이동형 전자 장치(802, 803, 804)는 드론(drone)을 포함할 수 있다.

[0093] 도 9는 본 발명의 다양한 실시 예에 따른 전자 장치의 구현예를 나타내는 도면이다.

[0094] 일 실시 예에 따르면, 전자 장치(900)(예: 전자 장치(100))는 로봇의 형태로 구현될 수 있다. 일 실시 예에 따르면, 전자 장치(900)는 제1 몸체(901) 및 제2 몸체(903)를 포함할 수 있다. 제1 몸체(901)는 제2 몸체(103)의 상측에 배치될 수 있다. 일 실시 예에 따르면, 제1 몸체(901) 및 제2 몸체(903)는 각각 사람의 헤드와 바디에 각각 대응되는 형상으로 구현될 수 있다. 일 실시 예에 따르면, 전자 장치(900)는 제1 몸체(901)의 전면에 배치되는 커버(920)를 포함할 수 있다. 일 실시 예에 따르면, 커버(920)는 투명 재질 또는 반투명 재질로 구성될 수 있다. 일 실시 예에 따르면, 커버는 사용자와 인터랙션하는 방향을 나타내는 곳으로 이미지 센싱을 위한 적어도 하나의 센서, 오디오를 취득하기 위한 적어도 하나의 마이크, 오디오를 출력하기 위한 적어도 하나의 스피커, 디스플레이, 기구적인 눈 구조를 포함할 수 있으며 불빛 또는 일시적인 기구 변경을 통하여 방향을 표시할 수도 있으며, 사용자와 인터랙션 할 때 사용자 방향으로 향하는 적어도 하나 이상의 H/W 또는 기구 구조를 포함할 수 있다.

[0095] 일 실시 예에 따르면, 제1 몸체(901)는 통신 모듈(910) 및 센서 모듈(950)을 더 포함할 수도 있다. 통신 모듈(910)은 외부 전자 장치로부터 메시지를 수신할 수 있으며, 외부 전자 장치로 변환된 메시지를 송신할 수도 있다.

[0096] 카메라(940)는 전자 장치(900) 외부 환경을 촬영할 수 있다. 예를 들어, 카메라(940)는 사용자를 촬영하여 이미지를 생성할 수 있다.

[0097] 센서 모듈(950)은 외부 환경에 대한 정보를 획득할 수 있다. 예를 들어, 센서 모듈(950)은 전자 장치(900)에 사용자가 근접함을 센싱할 수 있다. 센서 모듈(950)은 근접 정보에 따라 사용자의 근접을 센싱할 수도 있으며, 또는 사용자가 이용하는 다른 전자 장치(예를 들어, 웨어러블 장치)로부터의 신호에 기초하여 사용자의 근접을 센싱할 수도 있다. 뿐만 아니라, 센서 모듈(950)은 사용자의 행동, 위치를 센싱할 수도 있다.

[0098] 구동 모듈(970)은 제1 몸체(901)를 움직이게 할 수 있는 적어도 하나의 모터를 포함할 수 있으며, 예를 들어, 제1 몸체(901)의 방향을 변경할 수 있다. 제1 몸체(901)의 방향이 변경됨에 따라, 예를 들어, 카메라(940)의 촬영 방향이 변경될 수 있다. 구동 모듈(970)의 형태는 적어도 하나 이상의 축을 중심으로 하여 상, 하 또는 좌, 우의 움직임이 가능한 형태일 수 있으며, 그 형태는 다양하게 구현될 수 있다. 전원 모듈(990)은 전자 장치(900)가 이용하는 전원을 공급할 수 있다.

[0099] 프로세서(980)는 다른 전자 장치로부터 무선으로 송신되는 메시지를 통신 모듈(910)을 통하여 획득하거나, 음성 메시지를 센서 모듈(950)을 통하여 획득할 수 있다. 일 실시 예에 따르면, 프로세서(980)는 적어도 하나의 메시지 분석 모듈을 포함할 수 있다. 일 실시 예에 따르면, 적어도 하나의 메시지 분석 모듈은 발신자가 생성한 메시지에서 수신자에게 전달 하고자 하는 주요 내용을 추출하거나 또는 내용을 분류(classify)할 수 있다.

[0100] 메모리(960)는 사용자에게 서비스를 제공하는 것과 관련된 정보를 영속적 또는 일시적으로 저장할 수 있는 저장소로, 상기 전자 장치(900)의 내부에 존재할 수 있으며, 또는 네트워크를 통하여, 클라우드 또는 다른 서버에 존재할 수 있다. 상기 메모리(960)는 전자 장치(900)에 의해 생성되거나 외부로부터 수신한 공간 정보를 저장할 수 있다. 상기 메모리(960)에는 사용자 인증을 위한 개인 정보 또는 사용자에게 서비스를 제공하는 방식에 관련된 속성 관련 정보 또는 전자 장치(900)와 인터랙션 할 수 있는 다양한 수단들 간에 관계를 파악할 수 있는 정보가 저장될 수 있다. 이때의 관계 정보는 전자 장치(900)의 사용에 따라 정보가 갱신 또는 학습되어 변경될 수도 있다. 프로세서(980)는 전자 장치(900)의 통제를 담당하며, 통신 모듈(910), 디스플레이, 스피커, 마이크, 카메라(940), 센서 모듈(950), 메모리(960), 구동 모듈(970) 및 전원 모듈(990)을 기능적으로 제어하여 사용자에게 서비스를 제공할 수 있게 한다. 일 실시 예에 따르면, 프로세서(980) 또는 메모리(960)의 적어도 일부분에 전자 장치(900)가 획득할 수 있는 정보를 판단할 수 있는 정보 판단부가 포함될 수 있으며, 이 때 정보 판단부는 센서 모듈(950) 또는 통신 모듈(910)을 통하여 획득한 정보에서 서비스를 위한 적어도 하나 이상의 데이터를 추출할 수도 있다. 한편, 전자 장치(900)가 로봇 형태로 구현된다는 것은 단순히 예시적인 것이며, 그 구현 형태에는 제한이 없다.

[0101] 도 10은 본 발명의 다양한 실시 예에 따른 네트워크 환경 내의 전자 장치를 나타내는 도면이다.

[0102] 도 10을 참조하여, 다양한 실시 예에서의, 네트워크 환경(1000) 내의 전자 장치(1001)에 대해 설명한다. 전자 장치(1001)는, 예를 들면, 도 1에 도시된 전자 장치(100)의 전체 또는 일부를 포함할 수 있다. 전자 장치(100

1)는 버스(1010), 프로세서(1020), 메모리(1030), 입출력 인터페이스(1050), 디스플레이(1060), 및 통신 인터페이스(1070)를 포함할 수 있다. 어떤 실시예에서는, 전자 장치(1001)는, 구성요소들 중 적어도 하나를 생략하거나 다른 구성요소를 추가적으로 구비할 수 있다.

[0103] 버스(1010)는, 예를 들면, 구성요소들(1010-1070)을 서로 연결하고, 구성요소들 간의 통신(예: 제어 메시지 및/또는 데이터)을 전달하는 회로를 포함할 수 있다.

[0104] 프로세서(1020)는, 중앙처리장치(central processing unit(CPU)), 어플리케이션 프로세서(application processor(AP)), 또는 커뮤니케이션 프로세서(communication processor(CP)) 중 하나 또는 그 이상을 포함할 수 있다. 프로세서(1020)는, 예를 들면, 전자 장치(1001)의 적어도 하나의 다른 구성요소들의 제어 및/또는 통신에 관한 연산이나 데이터 처리를 실행할 수 있다.

[0105] 메모리(1030)는, 휘발성 및/또는 비휘발성 메모리를 포함할 수 있다. 메모리(1030)는, 예를 들면, 전자 장치(1001)의 적어도 하나의 다른 구성요소에 관계된 명령 또는 데이터를 저장할 수 있다. 일 실시예에 따르면, 메모리(1030)는 소프트웨어 및/또는 프로그램(1040)을 저장할 수 있다.

[0106] 프로그램(1040)은, 예를 들면, 커널(1041), 미들웨어(1043), 어플리케이션 프로그래밍 인터페이스(application programming interface(API))(1045), 및/또는 어플리케이션 프로그램(또는 "어플리케이션")(1047) 등을 포함할 수 있다. 커널(1041), 미들웨어(1043), 또는 API(1045)의 적어도 일부는, 운영 시스템(operating system(OS))으로 지칭될 수 있다.

[0107] 커널(1041)은, 예를 들면, 다른 프로그램들(예: 미들웨어(1043), API(1045), 또는 어플리케이션 프로그램(1047))에 구현된 동작 또는 기능을 실행하는 데 사용되는 시스템 리소스들(예: 버스(1010), 프로세서(1020), 또는 메모리(1030) 등)을 제어 또는 관리할 수 있다. 또한, 커널(1041)은 미들웨어(1043), API(1045), 또는 어플리케이션 프로그램(1047)에서 전자 장치(1001)의 개별 구성요소에 접근함으로써, 시스템 리소스들을 제어 또는 관리할 수 있는 인터페이스를 제공할 수 있다.

[0108] 미들웨어(1043)는, 예를 들면, API(1045) 또는 어플리케이션 프로그램(1047)이 커널(1041)과 통신하여 데이터를 주고받을 수 있도록 중개 역할을 수행할 수 있다. 또한, 미들웨어(1043)는 어플리케이션 프로그램(1047)으로부터 수신된 하나 이상의 작업 요청들을 우선 순위에 따라 처리할 수 있다. 예를 들면, 미들웨어(1043)는 어플리케이션 프로그램(1047) 중 적어도 하나에 전자 장치(1001)의 시스템 리소스(예: 버스(1010), 프로세서(1020), 또는 메모리(1030) 등)를 사용할 수 있는 우선 순위를 부여하고, 상기 하나 이상의 작업 요청들을 처리할 수 있다.

[0109] API(1045)는, 예를 들면, 어플리케이션(1047)이 커널(1041) 또는 미들웨어(1043)에서 제공되는 기능을 제어하기 위한 인터페이스로, 예를 들면, 파일 제어, 창 제어, 영상 처리, 또는 문자 제어 등을 위한 적어도 하나의 인터페이스 또는 함수(예: 명령어)를 포함할 수 있다.

[0110] 입출력 인터페이스(1050)는, 예를 들면, 사용자 또는 다른 외부 기기로부터 입력된 명령 또는 데이터를 전자 장치(1001)의 다른 구성요소(들)에 전달할 수 있는 인터페이스의 역할을 할 수 있다. 또한, 입출력 인터페이스(1050)는 전자 장치(1001)의 다른 구성요소(들)로부터 수신된 명령 또는 데이터를 사용자 또는 다른 외부 기기로 출력할 수 있다.

[0111] 디스플레이(1060)는, 예를 들면, 액정 디스플레이(LCD), 발광 다이오드(LED) 디스플레이, 유기 발광 다이오드(OLED) 디스플레이, 또는 마이크로 전자기계 시스템(microelectromechanical systems(MEMS)) 디스플레이, 또는 전자종이(electronic paper) 디스플레이를 포함할 수 있다. 디스플레이(1060)는, 예를 들면, 사용자에게 각종 콘텐츠(예: 텍스트, 이미지, 비디오, 아이콘, 또는 심볼 등)을 표시할 수 있다. 디스플레이(1060)는, 터치 스크린을 포함할 수 있으며, 예를 들면, 전자 펜 또는 사용자의 신체의 일부를 이용한 터치, 제스쳐, 근접, 또는 호버링 입력을 수신할 수 있다.

[0112] 통신 인터페이스(1070)는, 예를 들면, 전자 장치(1001)와 외부 장치(예: 제 1 외부 전자 장치(1002), 제 2 외부 전자 장치(1004), 또는 서버(1006)) 간의 통신을 설정할 수 있다. 예를 들면, 통신 인터페이스(1070)는 무선 통신 또는 유선 통신을 통해서 네트워크(1062)에 연결되어 외부 장치(예: 제 2 외부 전자 장치(1004) 또는 서버(1006))와 통신할 수 있다.

[0113] 무선 통신은, 예를 들면, 셀룰러 통신 프로토콜로서, 예를 들면, LTE(long-term evolution), LTE-A(LTE Advance), CDMA(code division multiple access), WCDMA(wideband CDMA), UMTS(universal mobile

telecommunications system), WiBro(Wireless Broadband), 또는 GSM(Global System for Mobile Communications) 등 중 적어도 하나를 사용할 수 있다. 또한, 무선 통신은, 예를 들면, 근거리 통신(1064)을 포함할 수 있다. 근거리 통신(1064)은, 예를 들면, WiFi(wireless fidelity), 블루투스(Bluetooth), 블루투스 저전력(BLE), 지그비(Zigbee), NFC(near field communication), 자력 시큐어 트랜스미션(Magnetic Secure Transmission), 또는 GNSS(global navigation satellite system) 중 적어도 하나를 포함할 수 있다. GNSS는, 예를 들면, GPS(Global Positioning System), Glonass(Global Navigation Satellite System), Beidou Navigation Satellite System(이하 "Beidou") 또는 Galileo, the European global satellite-based navigation system 중 적어도 하나를 포함할 수 있다. 이하, 본 문서에서는, "GPS"는 "GNSS"와 혼용되어 사용(interchangeably used)될 수 있다.

[0114] 유선 통신은, 예를 들면, USB(universal serial bus), HDMI(high definition multimedia interface), RS-232(recommended standard232), 전력선 통신, 또는 POTS(plain old telephone service) 등 중 적어도 하나를 포함할 수 있다. 네트워크(1062)는 통신 네트워크(telecommunications network), 예를 들면, 컴퓨터 네트워크(computer network)(예: LAN 또는 WAN), 인터넷, 또는 전화 망(telephone network) 중 적어도 하나를 포함할 수 있다.

[0115] 제 1 외부 전자 장치(1002) 및 제 2 외부 전자 장치(1004) 각각은 전자 장치(1001)와 동일한 또는 다른 종류의 장치일 수 있다. 일 실시 예에 따르면, 서버(1006)는 하나 또는 그 이상의 서버들의 그룹을 포함할 수 있다. 다양한 실시 예에 따르면, 전자 장치(1001)에서 실행되는 동작들의 전부 또는 일부는 다른 하나 또는 복수의 전자 장치(예: 제1 외부 전자 장치(1002). 제2 외부 전자 장치(1004), 또는 서버(1006))에서 실행될 수 있다. 일 실시 예에 따르면, 전자 장치(1001)가 어떤 기능이나 서비스를 자동으로 또는 요청에 의하여 수행해야 할 경우에, 전자 장치(1001)는 기능 또는 서비스를 자체적으로 실행시키는 대신에 또는 추가적으로, 그와 연관된 적어도 일부 기능을 다른 장치(예: 제1 외부 전자 장치(1002). 제2 외부 전자 장치(1004) 또는 서버(1006))에게 요청할 수 있다. 다른 전자 장치(예: 제1 외부 전자 장치(1002). 제2 외부 전자 장치(1004) 또는 서버(1006))는 요청된 기능 또는 추가 기능을 실행하고, 그 결과를 전자 장치(1001)로 전달할 수 있다. 전자 장치(1001)는 수신된 결과를 그대로 또는 추가적으로 처리하여 요청된 기능이나 서비스를 제공할 수 있다. 이를 위하여, 예를 들면, 클라우드 컴퓨팅, 분산 컴퓨팅, 또는 클라이언트-서버 컴퓨팅 기술이 이용될 수 있다.

[0116] 도 11은 다양한 실시 예에 따른, 전자 장치를 도시하는 블럭도이다.

[0117] 전자 장치(1101)는, 예를 들면, 도 1에 도시된 전자 장치(100)의 전체 또는 일부를 포함할 수 있다. 전자 장치(1101)는 하나 이상의 프로세서(예: AP(application processor))(1110), 통신 모듈(1120), (가입자 식별 모듈(1124), 메모리(1130), 센서 모듈(1140), 입력 장치(1150), 디스플레이(1160), 인터페이스(1170), 오디오 모듈(1180), 카메라 모듈(1191), 전력 관리 모듈(1195), 배터리(1196), 인디케이터(1197), 및 모터(1198) 를 포함할 수 있다.

[0118] 프로세서(1110)는, 예를 들면, 운영 체제 또는 응용 프로그램을 구동하여 프로세서(1110)에 연결된 다수의 하드웨어 또는 소프트웨어 구성요소들을 제어할 수 있고, 각종 데이터 처리 및 연산을 수행할 수 있다. 프로세서(1110)는, 예를 들면, SoC(system on chip) 로 구현될 수 있다. 일 실시 예에 따르면, 프로세서(1110)는 GPU(graphic processing unit) 및/또는 이미지 신호 프로세서(image signal processor)를 더 포함할 수 있다. 프로세서(1110)는 도 11에 도시된 구성요소들 중 적어도 일부(예: 셀룰러 모듈(1121))를 포함할 수도 있다. 프로세서(1110) 는 다른 구성요소들(예: 비휘발성 메모리) 중 적어도 하나로부터 수신된 명령 또는 데이터를 휘발성 메모리에 로드(load)하여 처리하고, 다양한 데이터를 비휘발성 메모리에 저장(store)할 수 있다.

[0119] 통신 모듈(1120)은, 도 10의 통신 인터페이스(1070)와 동일 또는 유사한 구성을 가질 수 있다. 통신 모듈(1120)은, 예를 들면, 셀룰러 모듈(1121), WiFi 모듈(1122), 블루투스 모듈(1123), GNSS 모듈(1124)(예: GPS 모듈, Glonass 모듈, Beidou 모듈, 또는 Galileo 모듈), NFC 모듈(1125), MST 모듈(1126), 및 RF(radio frequency) 모듈(1127)를 포함할 수 있다.

[0120] 셀룰러 모듈(1121)은, 예를 들면, 통신망을 통해서 음성 통화, 영상 통화, 문자 서비스, 또는 인터넷 서비스 등을 제공할 수 있다. 일 실시 예에 따르면, 셀룰러 모듈(1121)은 가입자 식별 모듈(예: SIM 카드)(1129)을 이용하여 통신 네트워크 내에서 전자 장치(1101)의 구별 및 인증을 수행할 수 있다. 일 실시 예에 따르면, 셀룰러 모듈(1121)은 프로세서(1110)가 제공할 수 있는 기능 중 적어도 일부 기능을 수행할 수 있다. 일 실시 예에 따르면, 셀룰러 모듈(1121)은 커뮤니케이션 프로세서(CP: communication processor)를 포함할 수 있다.

[0121] WiFi 모듈(1122), 블루투스 모듈(1123), GNSS 모듈(1124), NFC 모듈(1125) 또는 MST 모듈(1126) 각각은, 예를 들면, 해당하는 모듈을 통해서 송수신되는 데이터를 처리하기 위한 프로세서를 포함할 수 있다. 어떤 실시예에 따르면, 셀룰러 모듈(1121), WiFi 모듈(1122), 블루투스 모듈(1123), GNSS 모듈(1124), NFC 모듈(1125) 또는 MST 모듈(1126) 중 적어도 일부(예: 두 개 이상)는 하나의 integrated chip(IC) 또는 IC 패키지 내에 포함될 수 있다.

[0122] RF 모듈(1127)은, 예를 들면, 통신 신호(예: RF 신호)를 송수신할 수 있다. RF 모듈(1127)은, 예를 들면, 트랜시버(transceiver), PAM(power amp module), 주파수 필터(frequency filter), LNA(low noise amplifier), 또는 안테나 등을 포함할 수 있다. 다른 실시예에 따르면, 셀룰러 모듈(1121), WiFi 모듈(1122), 블루투스 모듈(1123), GNSS 모듈(1124), NFC 모듈(1125) 또는 MST 모듈(1126) 중 적어도 하나는 별개의 RF 모듈을 통하여 RF 신호를 송수신할 수 있다.

[0123] 가입자 식별 모듈(1129)은, 예를 들면, 가입자 식별 모듈을 포함하는 카드 및/또는 내장 SIM(embedded SIM)을 포함할 수 있으며, 고유한 식별 정보(예: ICCID(integrated circuit card identifier)) 또는 가입자 정보(예: IMSI(international mobile subscriber identity))를 포함할 수 있다.

[0124] 메모리(1130)는, 예를 들면, 내장 메모리(1132) 또는 외장 메모리(1134)를 포함할 수 있다. 내장 메모리(1132)는, 예를 들면, 휘발성 메모리(예: DRAM(dynamic RAM), SRAM(static RAM), 또는 SDRAM(synchronous dynamic RAM) 등), 비휘발성 메모리(non-volatile Memory)(예: OTPROM(one time programmable ROM), PROM(programmable ROM), EPROM(erasable and programmable ROM), EEPROM(electrically erasable and programmable ROM), mask ROM, flash ROM, 플래시 메모리(예: NAND flash 또는 NOR flash 등), 하드 드라이브, 또는 솔리드 스테이트 드라이브(solid state drive(SSD)) 중 적어도 하나를 포함할 수 있다.

[0125] 외장 메모리(1134)는 플래시 드라이브(flash drive), 예를 들면, CF(compact flash), SD(secure digital), Micro-SD(micro secure digital), Mini-SD(mini secure digital), xD(extreme digital), MMC(multi-media card) 또는 메모리 스틱(memory stick) 등을 더 포함할 수 있다. 외장 메모리(1134)는 다양한 인터페이스를 통하여 전자 장치(1101)와 기능적으로 및/또는 물리적으로 연결될 수 있다.

[0126] 보안 모듈(1136)(또는, 보안 메모리)은 메모리(1130)보다 상대적으로 보안 레벨이 높은 저장 공간을 포함하는 모듈로써, 안전한 데이터 저장 및 보호된 실행 환경을 보장해주는 회로일 수 있다. 보안 모듈(1136)은 별도의 회로로 구현될 수 있으며, 별도의 프로세서를 포함할 수 있다. 보안 모듈(1136)은, 예를 들면, 탈착 가능한 스마트 칩, 시큐어 디지털(secure digital(SD)) 카드 내에 존재하거나, 또는 전자 장치(1101)의 고정 칩 내에 내장된 내장형 보안 요소(embedded secure element(eSE))를 포함할 수 있다. 또한, 보안 모듈(1136)은 전자 장치(1101)의 운영 체제(operating system(OS))와 다른 운영 체제로 구동될 수 있다. 예를 들면, JCOP(java card open platform) 운영 체제를 기반으로 동작할 수 있다. 센서 모듈(1140)은, 예를 들면, 물리량을 계측하거나 전자 장치(1101)의 작동 상태를 감지하여, 계측 또는 감지된 정보를 전기 신호로 변환할 수 있다. 센서 모듈(1140)은, 예를 들면, 제스처 센서(1140A), 자이로 센서(1140B), 기압 센서(1140C), 마그네틱 센서(1140D), 가속도 센서(1140E), 그립 센서(1140F), 근접 센서(1140G), 컬러(color) 센서(1140H)(예: RGB(red, green, blue) 센서), 생체 센서(1140I), 온/습도 센서(1140J), 조도 센서(1140K), 또는 UV(ultra violet) 센서(1140M) 중의 적어도 하나를 포함할 수 있다. 추가적으로 또는 대체적으로(additionally or alternatively), 센서 모듈(1140)은, 예를 들면, 후각 센서(E-nose sensor), EMG 센서(electromyography sensor), EEG 센서(electroencephalogram sensor), ECG 센서(electrocardiogram sensor), IR(infrared) 센서, 홍채 센서 및/또는 지문 센서를 포함할 수 있다. 센서 모듈(1140)은 그 안에 속한 적어도 하나 이상의 센서들을 제어하기 위한 제어 회로를 더 포함할 수 있다. 어떤 실시 예에서는, 전자 장치(1101)는 프로세서(1110)의 일부로서 또는 별도로, 센서 모듈(1140)을 제어하도록 구성된 프로세서를 더 포함하여, 프로세서(1110)가 슬립(sleep) 상태에 있는 동안, 센서 모듈(1140)을 제어할 수 있다.

[0127] 입력 장치(1150)는, 예를 들면, 터치 패널(touch panel)(1152), (디지털) 펜 센서(pen sensor)(1154), 키(key)(1156), 또는 초음파(ultrasonic) 입력 장치(1158)를 포함할 수 있다. 터치 패널(1152)은, 예를 들면, 정전식, 감압식, 적외선 방식, 또는 초음파 방식 중 적어도 하나의 방식을 사용할 수 있다. 또한, 터치 패널(1152)은 제어 회로를 더 포함할 수도 있다. 터치 패널(1152)은 택타일 레이어(tactile layer)를 더 포함하여, 사용자에게 촉각 반응을 제공할 수 있다.

[0128] (디지털) 펜 센서(1154)는, 예를 들면, 터치 패널의 일부이거나, 별도의 인식용 쉬트(sheet)를 포함할 수 있다. 키(1156)는, 예를 들면, 물리적인 버튼, 광학식 키, 또는 키패드를 포함할 수 있다. 초음파 입력 장치(1158)는

마이크(예: 마이크(1188))를 통해, 입력 도구에서 발생된 초음파를 감지하여, 상기 감지된 초음파에 대응하는 데이터를 확인할 수 있다.

[0129] 디스플레이(1160)는 패널(1162), 홀로그램 장치(1164), 또는 프로젝터(1166)를 포함할 수 있다. 패널(1162)은, 예를 들면, 유연하게(flexible), 투명하게(transparent), 또는 착용할 수 있게(wearable) 구현될 수 있다. 패널(1162)은 터치 패널(1152)과 하나의 모듈로 구성될 수도 있다. 홀로그램 장치(1164)는 빛의 간섭을 이용하여 입체 영상을 허공에 보여줄 수 있다. 프로젝터(1166)는 스크린에 빛을 투사하여 영상을 표시할 수 있다. 스크린은, 예를 들면, 전자 장치(1101)의 내부 또는 외부에 위치할 수 있다. 일 실시 예에 따르면, 디스플레이(1160)는 패널(1162), 홀로그램 장치(1164), 또는 프로젝터(1166)를 제어하기 위한 제어 회로를 더 포함할 수 있다.

[0130] 인터페이스(1170)는, 예를 들면, HDMI(high-definition multimedia interface)(1172), USB(universal serial bus)(1174), 광 인터페이스(optical interface)(1176), 또는 D-sub(D-subminiature)(1178)를 포함할 수 있다. 인터페이스(1170)는, 예를 들면, 도 10에 도시된 통신 인터페이스(1070)에 포함될 수 있다. 추가적으로 또는 대체적으로(additionally and alternatively), 인터페이스(1170)는, 예를 들면, MHL(mobile high-definition link) 인터페이스, SD(secure digital) 카드/MMC(multi-media card) 인터페이스, 또는 IrDA(infrared data association) 규격 인터페이스를 포함할 수 있다.

[0131] 오디오 모듈(1180)은, 예를 들면, 소리(sound)와 전기 신호를 쌍방향으로 변환시킬 수 있다. 오디오 모듈(1180)은, 예를 들면, 스피커(1182), 리시버(1184), 이어폰(1186), 또는 마이크(1188) 등을 통해 입력 또는 출력되는 소리 정보를 처리할 수 있다.

[0132] 카메라 모듈(1191)은, 예를 들면, 정지 영상 및 동영상을 촬영할 수 있는 장치로서, 일 실시 예에 따르면, 하나 이상의 이미지 센서(예: 전면 센서 또는 후면 센서), 렌즈, ISP(image signal processor), 또는 플래시(flash)(예: LED 또는 xenon lamp 등)를 포함할 수 있다.

[0133] 전력 관리 모듈(1195)은, 예를 들면, 전자 장치(1101)의 전력을 관리할 수 있다. 일 실시 예에 따르면, 전력 관리 모듈(1195)은 PMIC(power management integrated circuit), 충전 IC(charger integrated circuit), 또는 배터리 또는 연료 게이지(battery or fuel gauge)를 포함할 수 있다. PMIC는, 유선 및/또는 무선 충전 방식을 가질 수 있다. 무선 충전 방식은, 예를 들면, 자기공명 방식, 자기유도 방식 또는 전자기파 방식 등을 포함하며, 무선 충전을 위한 부가적인 회로, 예를 들면, 코일 루프, 공진 회로, 또는 정류기 등을 더 포함할 수 있다. 배터리 게이지는, 예를 들면, 배터리(1196)의 잔량, 충전 중 전압, 전류, 또는 온도를 측정할 수 있다. 배터리(1196)는, 예를 들면, 충전식 전지(rechargeable battery) 및/또는 태양 전지(solar battery)를 포함할 수 있다.

[0134] 인디케이터(1197)는 전자 장치(1101) 또는 그 일부(예: 프로세서(1110))의 특정 상태, 예를 들면, 부팅 상태, 메시지 상태 또는 충전 상태 등을 표시할 수 있다. 모터(1198)는 전기적 신호를 기계적 진동으로 변환할 수 있고, 진동(vibration), 또는 햅틱(haptic) 효과 등을 발생시킬 수 있다. 도시되지는 않았으나, 전자 장치(1101)는 모바일 TV 지원을 위한 처리 장치(예: GPU)를 포함할 수 있다. 모바일 TV 지원을 위한 처리 장치는, 예를 들면, DMB(digital multimedia broadcasting), DVB(digital video broadcasting), 또는 미디어플로(mediaFlo™) 등의 규격에 따른 미디어 데이터를 처리할 수 있다.

[0135] 도 12는 다양한 실시 예에 따른, 전자 장치를 도시하는 블럭도이다.

[0136] 전자 장치(1201)는, 예를 들면, 도 1에 도시된 전자 장치(100)의 전체 또는 일부를 포함할 수 있다. 도 12를 참조하면 프로세서(1210)는 영상 인식 모듈(1241)과 연결될 수 있다. 아울러, 프로세서(1210)는 행동 모듈(1244)과 연결될 수 있다. 영상 인식 모듈(1241)은 2차원 카메라(1242) 및 뎁스 카메라(1243) 중 적어도 하나를 포함할 수 있다. 영상 인식 모듈(1241)은 촬영 결과를 기초로 인식을 수행할 수 있으며, 인식 결과를 프로세서(1210)에게 전달할 수 있다. 행동 모듈(1244)는 전자 장치(1201)의 얼굴 표정을 나타내거나 얼굴이 바라보는 방향을 변경하기 위한 얼굴 표정 모터(1245), 전자 장치(1201) 바디부의 포즈, 예를 들어 팔, 다리, 또는 손가락의 위치를 변경하기 위한 바디 포즈 모터(1245), 및 전자 장치(1201)를 이동시키기 위한 이동 모터(1247) 중 적어도 하나를 포함할 수 있다. 프로세서(1210)는 얼굴 표정 모터(1245), 바디 포즈 모터(1246) 및 이동 모터(1247) 중 적어도 하나를 제어하여, 로봇 형태로 구현된 전자 장치(1201)의 움직임을 제어할 수 있다. 다양한 실시 예에 따르면, 프로세서(1210)는 외부 전자 장치로부터 수신된 모션 데이터에 기초하여 로봇 형태로 구현된 전자 장치(1201)의 표정, 헤드 또는 바디를 제어할 수 있다. 예를 들어, 전자 장치(1201)는 외부 전자 장치로부터 외부 전자 장치 사용자의 표정, 헤드 모션, 또는 바디 모션에 기초하여 생성된 모션 데이터를 수신할 수 있

다. 프로세서(1210)는 모션 데이터에 포함된 표정 데이터, 헤드 모션 데이터 또는 바디 모션 데이터 각각을 추출하고, 추출된 데이터에 기초하여 얼굴 표정 모터(1245) 또는 바디 포즈 모터(1246)를 제어할 수 있다. 전자 장치(1001)는 도 11의 구성 요소에 추가적으로 도 12의 구성 요소를 포함할 수도 있다.

[0137] 도 13은 본 발명의 다양한 실시예에 따른 전자 장치의 소프트웨어 블록도이다.

[0138] 도 13을 참조하면 전자 장치(1301)(예: 전자 장치(100))는 OS(Operating System)/시스템 소프트웨어(1301), 미들웨어(Middleware, 1110), 인텔리전트 프레임워크(Intelligent Framework, 1330)를 포함할 수 있다.

[0139] OS/시스템 소프트웨어(1301)는 전자 장치(1301)의 리소스(resource)를 분배하고 잡스케쥴링(Job scheduling) 및 프로세스 처리를 수행할 수 있다. 또한 OS/시스템 소프트웨어(1301)는 다양한 하드웨어 입력부(1309)로부터 수신된 데이터를 처리할 수 있다. 하드웨어 입력부(1309)는 뎁쓰카메라(1303), 2D카메라(1304), 센서모듈(1305), 터치센서(1306) 및 마이크 어레이(1307) 중 적어도 하나를 포함할 수 있다.

[0140] 미들웨어(1310)는 OS/시스템 소프트웨어(1301)에서 처리된 데이터를 이용하여 전자 장치(1301)의 기능을 수행할 수 있다. 일 실시예에 따르면 미들웨어(1310)는 제스처 인식 매니저(1311), 얼굴 검출/추적/인식 매니저(1312), 센서 정보처리 매니저(1313), 대화 엔진 매니저(1314), 음성합성 매니저(1315), 음원추적 매니저(1316) 및 음성 인식 매니저(1317)를 포함할 수 있다.

[0141] 한 실시예에 따르면 얼굴 검출/추적/인식 매니저(1312)는 2D 카메라(1304)에서 촬영된 이미지를 분석하여 사용자의 얼굴 위치를 검출하거나 또는 추적하고, 얼굴 인식을 통해 인증을 수행할 수 있다. 제스처 인식 매니저(511)는 2D 카메라(1304) 및 뎁쓰 카메라(1303)로 촬영된 이미지를 분석하여 사용자의 3차원 제스처를 인식할 수 있다. 음원추적 매니저(1316)는 마이크(1307)를 통해 입력된 음성을 분석하고 음원에 대한 입력 위치를 추적할 수 있다. 음성 인식 매니저(1317)는 마이크(1307)를 통해 입력된 음성을 분석하여 입력된 음성을 인식할 수 있다.

[0142] 인텔리전트 프레임워크(Intelligent Framework, 1330)는 멀티 모달 융합 모듈(1331), 사용자 패턴학습 모듈(1332) 및 행동 제어 모듈(1333)을 포함할 수 있다. 한 실시예에 따르면 멀티모달 융합 모듈(1331)은 미들웨어(1310)에서 처리된 정보를 취합하고 관리할 수 있다. 사용자 패턴 학습 모듈(1332)은 멀티모달 융합 모듈(1331)의 정보를 이용하여 사용자의 생활 패턴, 선호도 등의 유의미한 정보를 추출하고 학습할 수 있다. 행동 제어 모듈(1333)은 전자 장치(1301)가 사용자에게 피드백할 정보를 전자 장치(1301)의 움직임, 시각 정보 또는 오디오 정보로 제공할 수 있다. 즉, 행동제어블록(1333)에서는 구동부의 모터(1340)를 제어하여 전자 장치(1301)를 움직이거나 또는 디스플레이(1350)에 그래픽 객체가 표시되도록 디스플레이를 제어하거나 또는 스피커(1361, 1362)를 제어하여 오디오를 출력할 수 있다.

[0143] 사용자 모델 데이터베이스(1321)는 인텔리전트 프레임워크(1330)에서 전자 장치(1301)가 학습한 데이터를 사용자에 따라 구분하여 저장할 수 있다. 행동모델 데이터베이스(1322)는 전자 장치(1301)의 행동 제어를 위한 데이터를 저장할 수 있다. 사용자 모델 데이터베이스(1321) 및 행동모델 데이터 베이스(1322)는 전자 장치(1301)의 메모리에 저장되거나 또는 네트웍을 통해 클라우드 서버에 저장되어 다른 전자 장치(1302)에게 공유될 수도 있다.

[0144] 본 문서에서 사용된 용어 "모듈"은, 예를 들면, 하드웨어, 소프트웨어 또는 펌웨어(firmware) 중 하나 또는 둘 이상의 조합을 포함하는 단위(unit)를 의미할 수 있다. "모듈"은, 예를 들면, 유닛(unit), 로직(logic), 논리 블록(logical block), 부품(component), 또는 회로(circuit) 등의 용어와 바꾸어 사용(interchangeably use)될 수 있다. "모듈"은, 일체로 구성된 부품의 최소 단위 또는 그 일부가 될 수 있다. "모듈"은 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수도 있다. "모듈"은 기계적으로 또는 전자적으로 구현될 수 있다. 예를 들면,"모듈"은, 알려졌거나 앞으로 개발될, 어떤 동작들을 수행하는 ASIC(application-specific integrated circuit) 칩, FPGAs(field-programmable gate arrays) 또는 프로그램 가능 논리 장치(programmable-logic device) 중 적어도 하나를 포함할 수 있다.

[0145] 다양한 실시예에 따른 장치(예: 모듈들 또는 그 기능들) 또는 방법(예: 동작들)의 적어도 일부는, 예컨대, 프로그램 모듈의 형태로 컴퓨터로 읽을 수 있는 저장매체(computer-readable storage media)에 저장된 명령어로 구현될 수 있다. 상기 명령어가 프로세서(예: 도 1의 프로세서(170))에 의해 실행될 경우, 상기 하나 이상의 프로세서가 상기 명령어에 해당하는 기능을 수행할 수 있다. 컴퓨터로 읽을 수 있는 저장매체는, 예를 들면, 메모리(예: 도 1의 메모리 (160))가 될 수 있다.

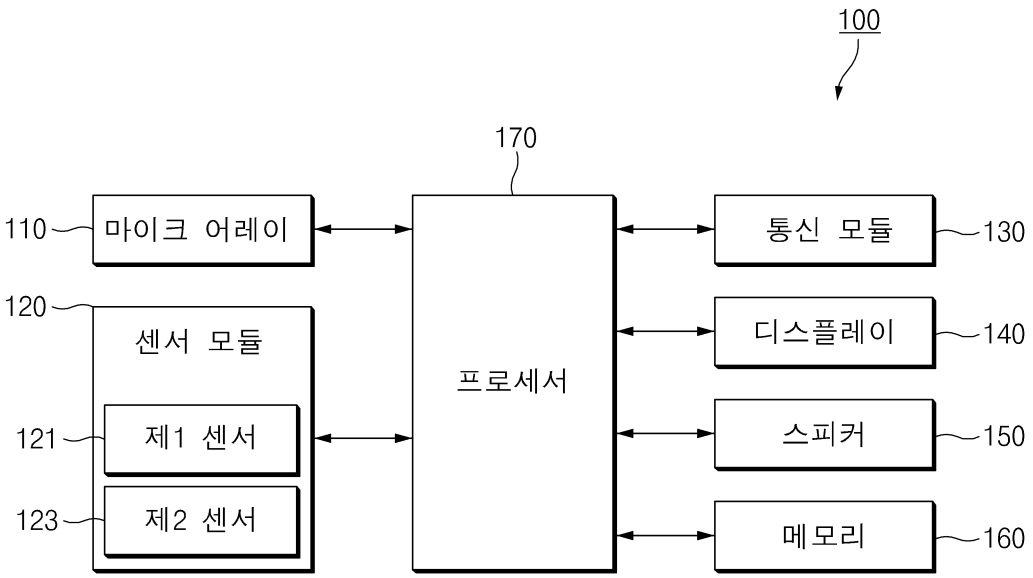
[0146] 컴퓨터로 판독 가능한 기록 매체는, 하드디스크, 플로피디스크, 마그네틱 매체(magnetic media)(예:

자기테이프), 광기록 매체(optical media)(예: CD-ROM(compact disc read only memory), DVD(digital versatile disc), 자기-광 매체(magneto-optical media)(예: 플롭티컬 디스크(floptical disk)), 하드웨어 장치(예: ROM(read only memory), RAM(random access memory), 또는 플래시 메모리 등) 등을 포함할 수 있다. 또한, 프로그램 명령에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함할 수 있다. 상술한 하드웨어 장치는 다양한 실시예의 동작을 수행하기 위해 하나 이상의 소프트웨어 모듈로서 작동하도록 구성될 수 있으며, 그 역도 마찬가지다.

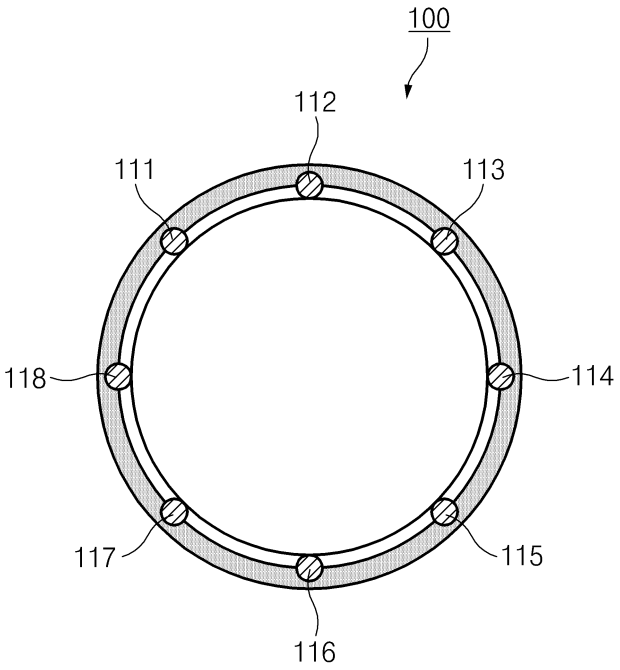
[0147] 다양한 실시예에 따른 모듈 또는 프로그램 모듈은 전술한 구성요소들 중 적어도 하나 이상을 포함하거나, 일부가 생략되거나, 또는 추가적인 다른 구성요소를 더 포함할 수 있다. 다양한 실시예에 따른 모듈, 프로그램 모듈 또는 다른 구성요소에 의해 수행되는 동작들은 순차적, 병렬적, 반복적 또는 휴리스틱(heuristic)한 방법으로 실행될 수 있다. 또한, 일부 동작은 다른 순서로 실행되거나, 생략되거나, 또는 다른 동작이 추가될 수 있다. 그리고 본 문서에 개시된 실시예는 개시된, 기술 내용의 설명 및 이해를 위해 제시된 것이며, 본 문서에서 기재된 기술의 범위를 한정하는 것은 아니다. 따라서, 본 문서의 범위는, 본 문서의 기술적 사상에 근거한 모든 변경 또는 다양한 다른 실시예를 포함하는 것으로 해석되어야 한다.

도면

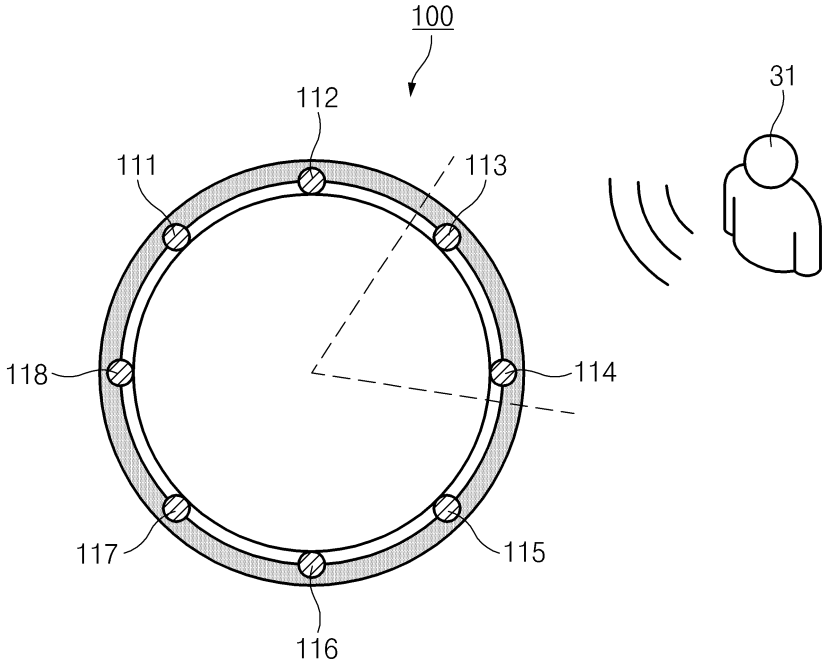
도면1



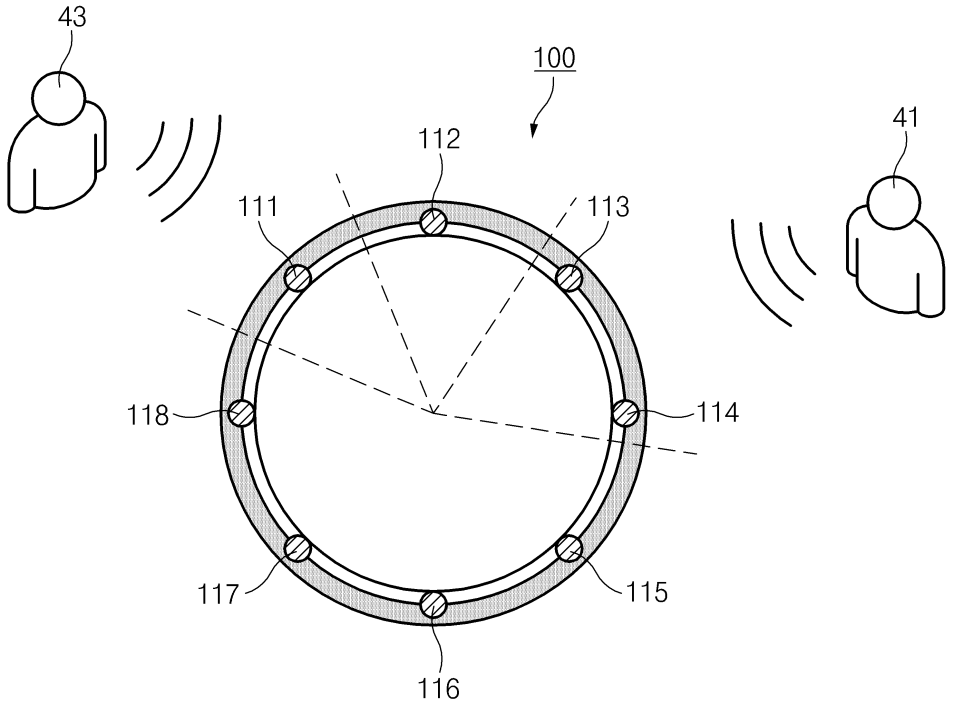
도면2


100
112
111
113
118
114
117
115
116

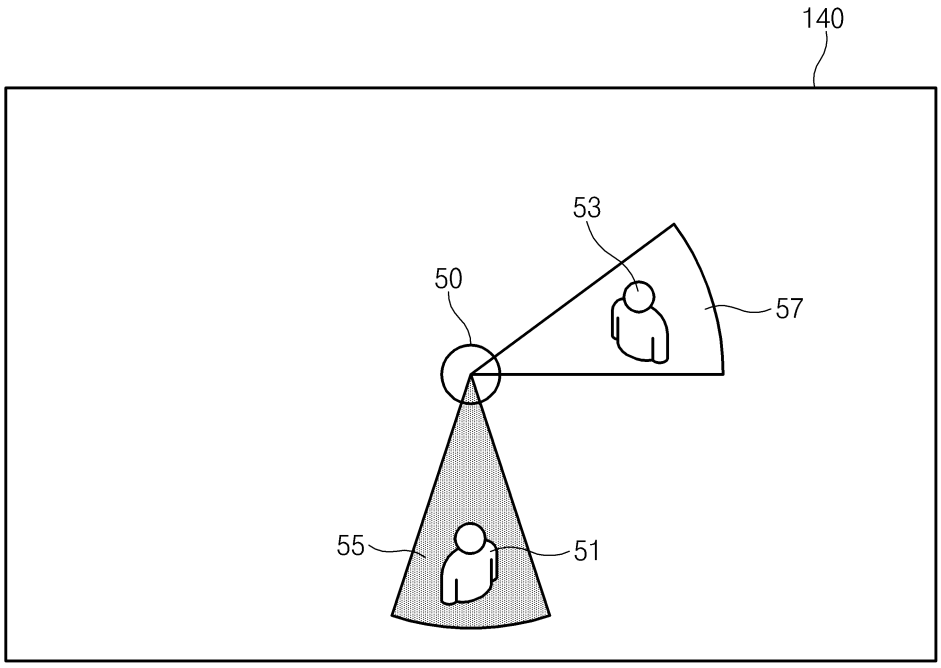
도면3


100
31
112
111
113
118
114
117
115
116

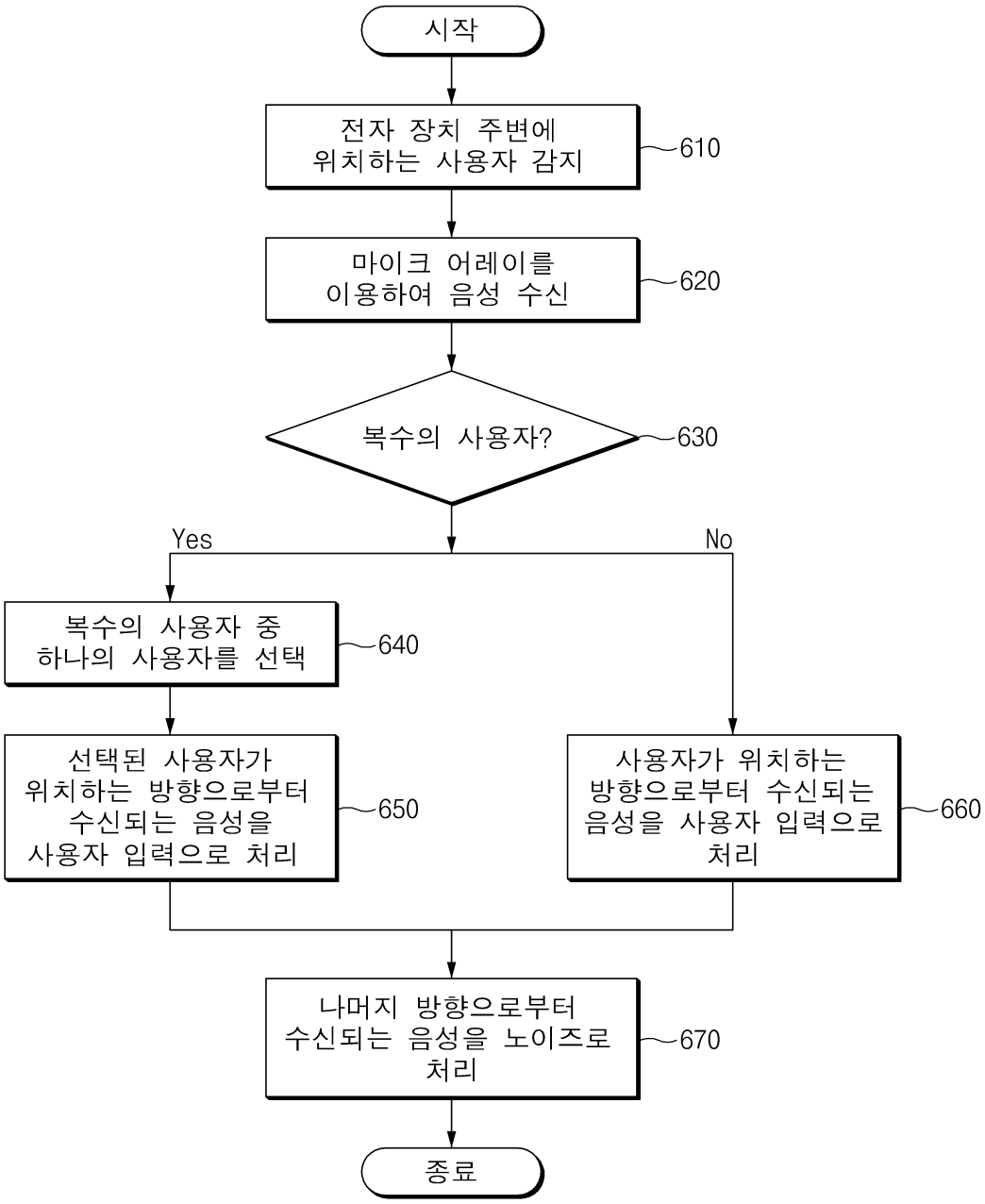
도면4



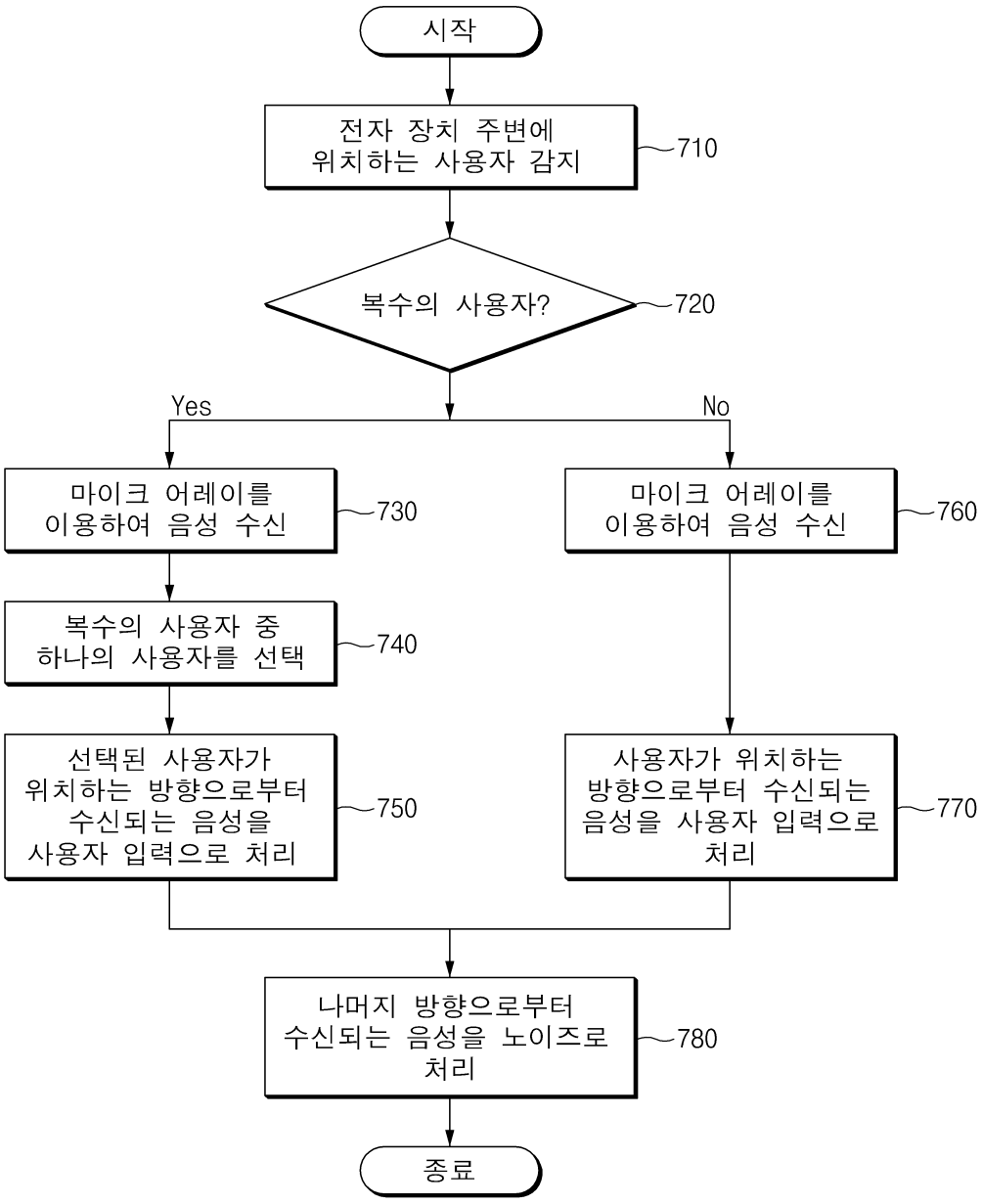
도면5



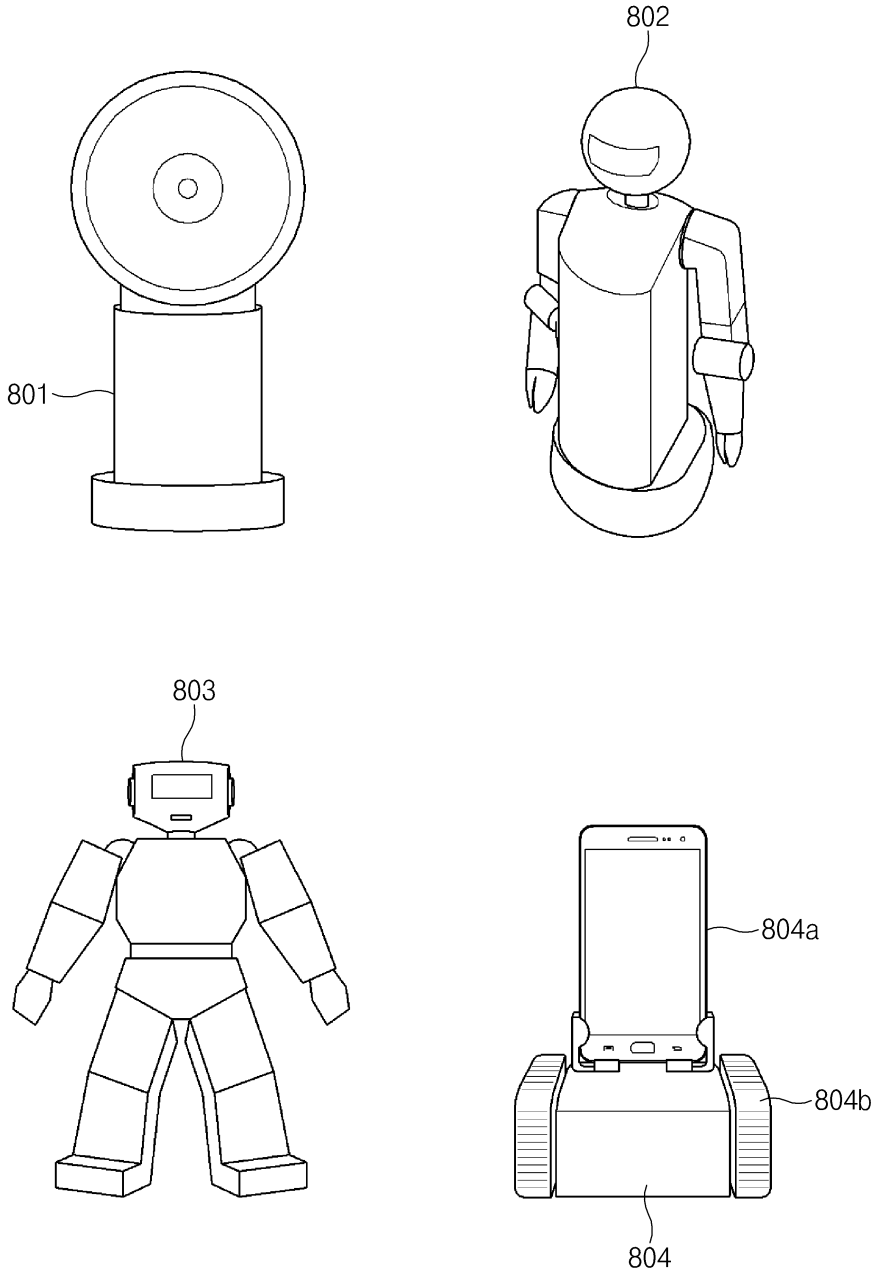
도면6


시작
전자 장치 주변에 위치하는 사용자 감지
610
마이크 어레이를 이용하여 음성 수신
620
복수의 사용자?
630
Yes
No
복수의 사용자 중 하나의 사용자를 선택
640
선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리
650
사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리
660
나머지 방향으로부터 수신되는 음성을 노이즈로 처리
670
종료

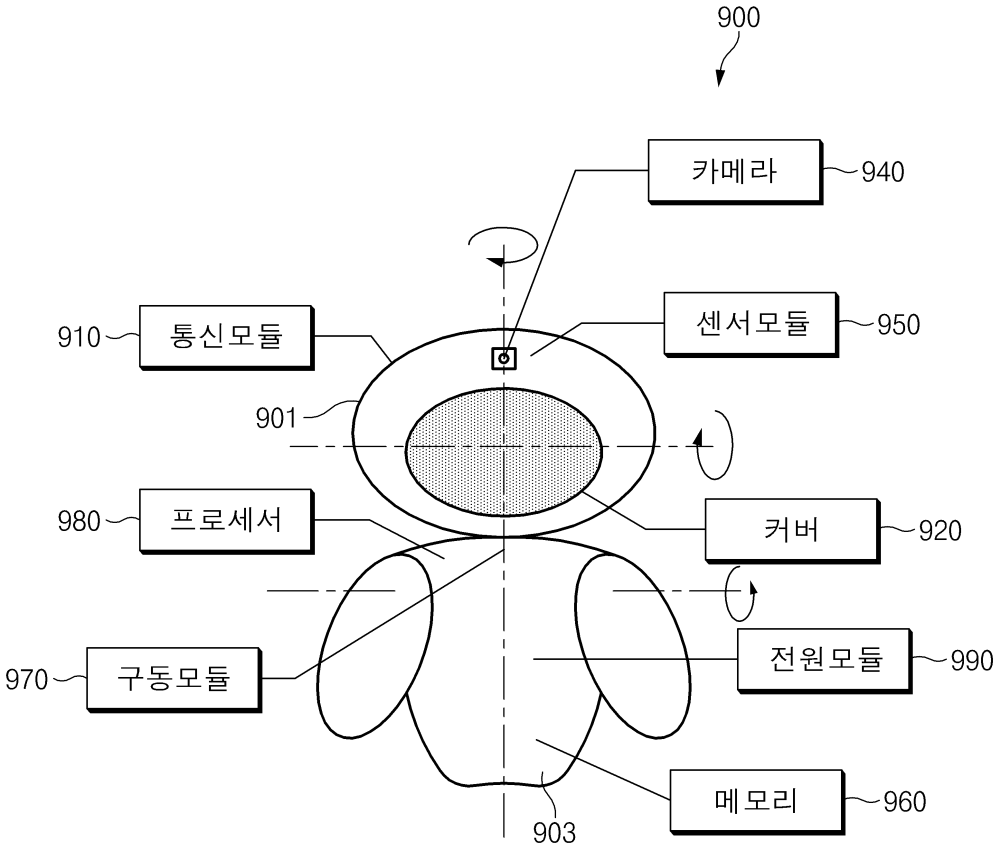
도면7


시작
전자 장치 주변에 위치하는 사용자 감지
710
복수의 사용자?
720
Yes
No
마이크 어레이를 이용하여 음성 수신
730
복수의 사용자 중 하나의 사용자를 선택
740
선택된 사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리
750
마이크 어레이를 이용하여 음성 수신
760
사용자가 위치하는 방향으로부터 수신되는 음성을 사용자 입력으로 처리
770
나머지 방향으로부터 수신되는 음성을 노이즈로 처리
780
종료

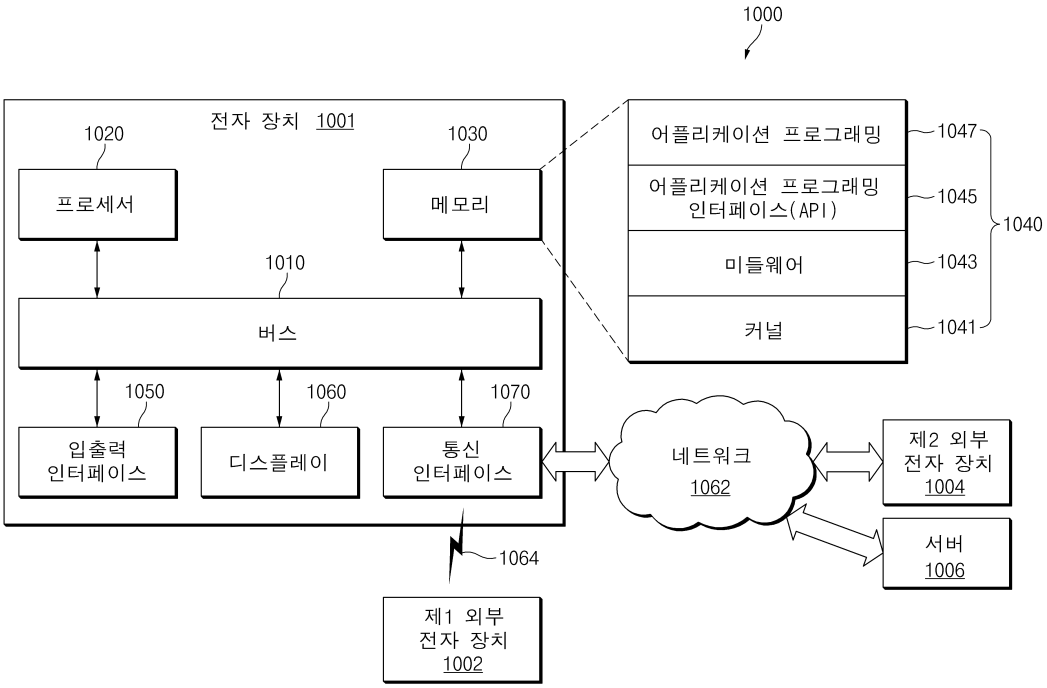
도면8



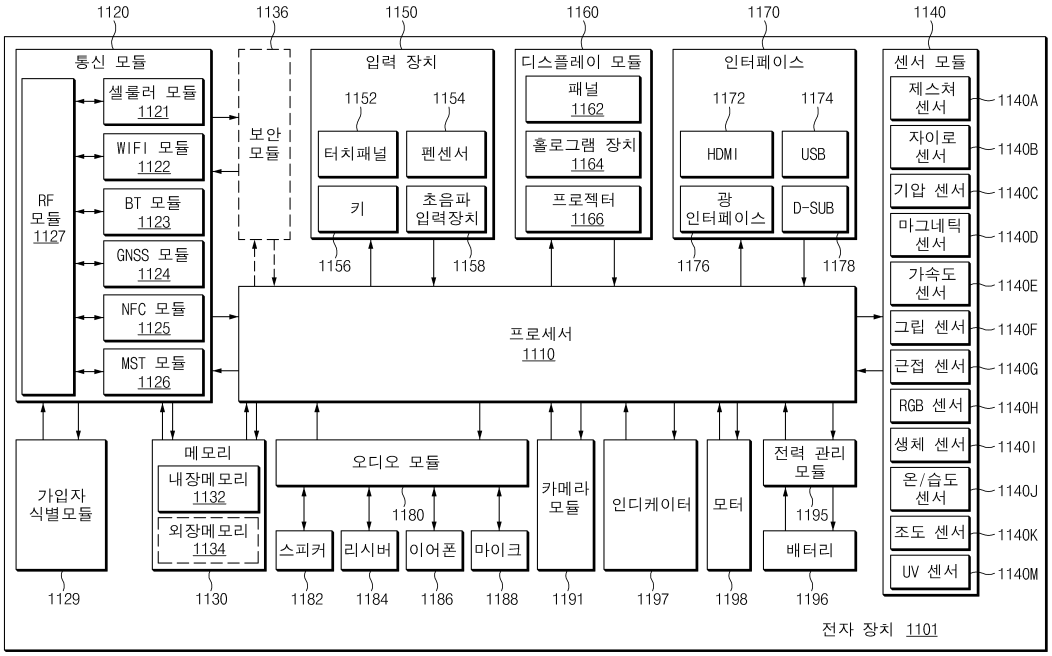
도면9


900
카메라
940
910
통신모듈
센서모듈
950
901
980
프로세서
커버
920
970
구동모듈
전원모듈
990
메모리
960
903

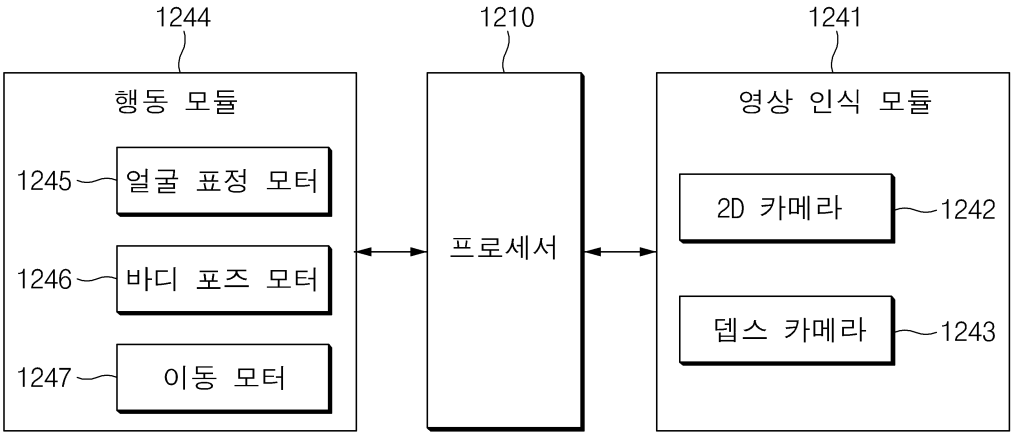
도면10


1000
전자 장치 1001
1020
프로세서
1030
메모리
1010
버스
1050
입출력 인터페이스
1060
디스플레이
1070
통신 인터페이스
1064
제1 외부 전자 장치 1002
네트워크 1062
제2 외부 전자 장치 1004
서버 1006
어플리케이션 프로그래밍
1047
어플리케이션 프로그래밍 인터페이스(API)
1045
미들웨어
1043
커널
1041
1040

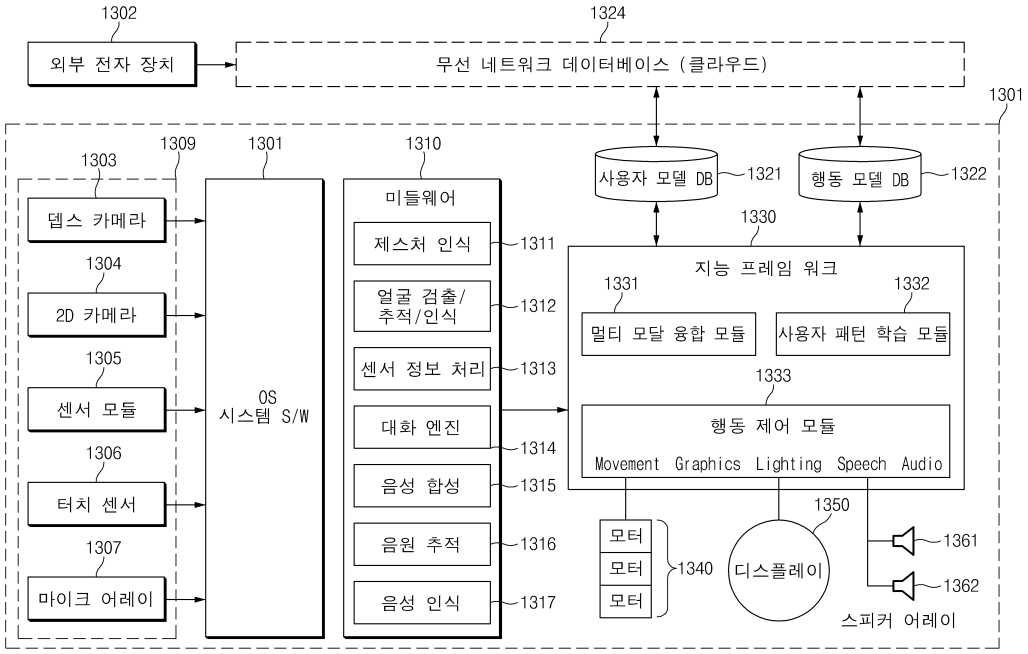
도면11


1120
1136
1150
1160
1170
1140
통신 모듈
셀룰러 모듈 1121
WIFI 모듈 1122
BT 모듈 1123
GNSS 모듈 1124
NFC 모듈 1125
MST 모듈 1126
RF 모듈 1127
보안 모듈
입력 장치
1152
1154
터치패널
펜센서
키
초음파 입력장치
디스플레이 모듈
패널 1162
홀로그램 장치 1164
프로젝터 1166
인터페이스
1172
1174
HDMI
USB
광 인터페이스
D-SUB
센서 모듈
제스쳐 센서
자이로 센서
기압 센서
마그네틱 센서
가속도 센서
그립 센서
근접 센서
RGB 센서
생체 센서
온/습도 센서
조도 센서
UV 센서
1140A
1140B
1140C
1140D
1140E
1140F
1140G
1140H
1140I
1140J
1140K
1140M
1156
1158
1176
1178
프로세서 1110
가입자 식별모듈
메모리
내장메모리 1132
외장메모리 1134
오디오 모듈
스피커
리시버
이어폰
마이크
1180
카메라 모듈
인디케이터
모터
전력 관리 모듈
배터리
1195
1129
1130
1182
1184
1186
1188
1191
1197
1198
1196
전자 장치 1101

도면12


1244
1210
1241
행동 모듈
1245
얼굴 표정 모터
1246
바디 포즈 모터
1247
이동 모터
프로세서
영상 인식 모듈
2D 카메라
1242
뎁스 카메라
1243

도면13


1302
외부 전자 장치
1324
무선 네트워크 데이터베이스 (클라우드)
1301
1309
1303
1301
1310
뎁스 카메라
2D 카메라
센서 모듈
터치 센서
마이크 어레이
1304
1305
1306
1307
OS
시스템 S/W
미들웨어
제스처 인식
얼굴 검출/
추적/인식
센서 정보 처리
대화 엔진
음성 합성
음원 추적
음성 인식
1311
1312
1313
1314
1315
1316
1317
사용자 모델 DB
1321
행동 모델 DB
1322
1330
지능 프레임 워크
1331
1332
멀티 모달 융합 모듈
사용자 패턴 학습 모듈
1333
행동 제어 모듈
Movement Graphics Lighting Speech Audio
모터
모터
모터
1340
1350
디스플레이
1361
1362
스피커 어레이