



**(19) 대한민국특허청(KR)**  
**(12) 등록특허공보(B1)**

(45) 공고일자 2023년11월28일  
(11) 등록번호 10-2606789  
(24) 등록일자 2023년11월22일

(51) 국제특허분류(Int. Cl.)  
G06F 3/16 (2018.01) G06F 3/0484 (2022.01)  
G10L 15/04 (2006.01) G10L 15/22 (2006.01)  
G10L 15/28 (2006.01)  
(52) CPC특허분류  
G06F 3/167 (2013.01)  
G06F 3/04847 (2022.01)  
(21) 출원번호 10-2018-0117192  
(22) 출원일자 2018년10월01일  
심사청구일자 2021년08월18일  
(65) 공개번호 10-2020-0037687  
(43) 공개일자 2020년04월09일  
(56) 선행기술조사문헌  
JP2008048076 A\*  
(뒷면에 계속)

(73) 특허권자  
삼성전자주식회사  
경기도 수원시 영통구 삼성로 129 (매탄동)  
(72) 발명자  
장성운  
경기도 안양시 동안구 관평로 68, 꿈마을한신아파트 708-101  
강상기  
경기도 수원시 팔달구 권광로 246, 래미안노블클래스아파트 104-301  
(뒷면에 계속)  
(74) 대리인  
특허법인태평양

전체 청구항 수 : 총 13 항

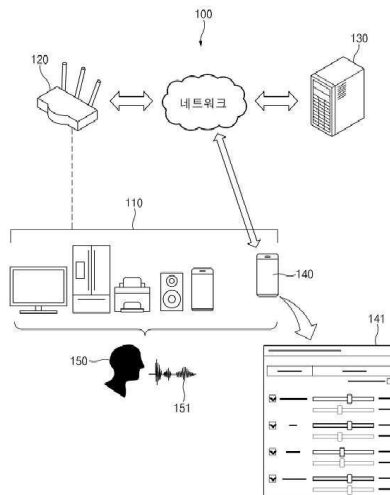
심사관 : 서광훈

(54) 발명의 명칭 복수의 음성 인식 장치들을 제어하는 방법 및 그 방법을 지원하는 전자 장치

**(57) 요약**

본 문서에 개시되는 일 실시 예에 따른 전자 장치는, 디스플레이, 마이크, 메모리, 통신 회로 및 프로세서를 포함하고, 상기 프로세서는 상기 디스플레이를 통해, 동일한 시작 발화에 대응하여 음성 인식을 시작할 수 있는 복수의 음성 인식 장치들 각각의 음성 인식 감도를 조절하기 위한 사용자 인터페이스를 표시하고, 상기 사용자 인터페이스를 통해 상기 음성 인식 감도가 변경된 경우, 상기 변경된 감도 값을 상기 복수의 음성 인식 장치들 중 적어도 일부에 전송하고, 상기 복수의 음성 인식 장치들에 사용자의 테스트 발화를 수신할 수 있도록 대기하도록 하는 신호를 송신하고, 상기 복수의 음성 인식 장치들로부터 상기 제1 발화에 대응하는 발화 정보를 수신하고, 상기 발화 정보를 기반으로 상기 사용자 인터페이스를 업데이트하도록 설정될 수 있다. 이 외에도 명세서를 통해 파악되는 다양한 실시 예가 가능하다.

**대표도 - 도1**



(52) CPC특허분류

*G10L 15/04* (2013.01)

*G10L 15/22* (2013.01)

*G10L 15/28* (2013.01)

(72) 발명자

**이남구**

경기도 수원시 영통구 권광로260번길 36, 현대힐스  
테이트아파트 108-1701

**정의석**

경기도 성남시 분당구 황새울로200번길 9-7, 판데  
온 1001호

(56) 선행기술조사문헌

KR1020160089145 A\*

JP2015122591 A

KR1020140024536 A

KR1020180085931 A

US20060100866 A1

US20160063997 A1

\*는 심사관에 의하여 인용된 문헌

## 명세서

### 청구범위

#### 청구항 1

전자 장치에 있어서,

디스플레이; 마이크; 메모리; 통신 회로; 및

프로세서;를 포함하고,

상기 프로세서는

동일한 시작 발화에 대응하여 음성 인식을 시작할 수 있는 상기 전자 장치와 별개의 복수의 음성 인식 장치들 각각의 음성 인식 감도를 조절하기 위한 사용자 인터페이스를 상기 디스플레이를 통해 표시하고,

상기 사용자 인터페이스를 통해 상기 음성 인식 감도가 변경된 경우, 변경된 감도 값을 상기 복수의 음성 인식 장치들 중 적어도 일부에 전송하고,

상기 복수의 음성 인식 장치들에 사용자의 테스트 발화를 수신하기 위해 대기하도록 하는 신호를 송신하고,

상기 복수의 음성 인식 장치들로부터 상기 테스트 발화에 대응하는 발화 정보를 수신하고,

상기 발화 정보를 기반으로 상기 복수의 음성 인식 장치들에서 상기 테스트 발화에 대한 수신 값을 표시하여 상기 사용자 인터페이스를 업데이트하도록 설정되고,

상기 발화 정보는,

상기 복수의 음성 인식 장치들의 동작 상태, 또는 상기 복수의 음성 인식 장치들 주변의 소음 정보 중 적어도 하나를 기반으로 상기 테스트 발화를 보정한 값을 포함하는 것을 특징으로 하는 전자 장치.

#### 청구항 2

제1항에 있어서, 상기 사용자 인터페이스는

상기 복수의 음성 인식 장치들 각각의 식별 정보, 및 상기 음성 인식 감도를 변경하는 제1 객체를 포함하는 전자 장치.

#### 청구항 3

제2항에 있어서, 상기 프로세서는

이동 바(bar) 또는 수치 입력 필드를 통해 상기 제1 객체를 표시하는 전자 장치.

#### 청구항 4

제2항에 있어서, 상기 사용자 인터페이스는

상기 음성 인식 감도의 기준 값, 상기 복수의 음성 인식 장치들의 상기 음성 인식 감도의 설정을 제외하는 제2 객체, 상기 음성 인식 감도를 초기화하는 제3 객체 중 적어도 하나를 포함하는 전자 장치.

#### 청구항 5

제1항에 있어서, 상기 발화 정보는

상기 복수의 음성 인식 장치들에서 수신된 상기 테스트 발화에 대한 오디오 음질(audio quality), 음압(sound pressure), 테스트 발화의 특징값 중 적어도 하나를 기반으로 보정된 값인 전자 장치.

#### 청구항 6

삭제

**청구항 7**

제1항에 있어서, 상기 프로세서는  
상기 음성 인식 감도를 조절하는 제1 객체와 상기 제1 객체와 다른 제2 객체를 오버랩하여 표시하는 전자 장치.

**청구항 8**

제7항에 있어서, 상기 프로세서는  
상기 제1 객체와 상기 제2 객체를 서로 다른 색으로 표시하는 전자 장치.

**청구항 9**

제1항에 있어서, 상기 프로세서는  
상기 발화 정보를 기반으로 상기 복수의 음성 인식 장치들 중 상기 시작 발화에 응답하는 장치를 결정하는 전자 장치.

**청구항 10**

제1항에 있어서, 상기 프로세서는  
상기 음성 인식 감도가 변경되는 경우, 상기 변경된 감도 값을 근거리 통신을 위한 네트워크 장치 또는 외부 서버를 통해, 상기 복수의 음성 인식 장치들에 전송하는 전자 장치.

**청구항 11**

제10항에 있어서, 상기 전자 장치는  
상기 네트워크 장치를 통해, 상기 복수의 음성 인식 장치들과 연결되는 전자 장치.

**청구항 12**

제11항에 있어서, 상기 프로세서는  
상기 네트워크 장치를 통해 상기 복수의 음성 인식 장치들로부터 상기 발화 정보를 수신하는 전자 장치.

**청구항 13**

제1항에 있어서, 상기 시작 발화는  
상기 테스트 발화와 서로 다르게 설정되는 전자 장치.

**청구항 14**

삭제

**청구항 15**

삭제

**청구항 16**

삭제

**청구항 17**

삭제

**청구항 18**

삭제

**청구항 19**

삭제

**청구항 20**

전자 장치에서 수행되는 음성 인식 장치의 제어 방법에 있어서,

상기 전자 장치의 디스플레이를 통해, 동일한 시작 발화에 대응하여 음성 인식 과정을 시작할 수 있는 상기 전자 장치와 별개의 복수의 음성 인식 장치들 각각의 음성 인식 감도를 조절하기 위한 사용자 인터페이스를 표시하는 동작;

상기 사용자 인터페이스를 통해 상기 음성 인식 감도가 변경된 경우, 변경된 감도 값을 복수의 음성 인식 장치들 중 적어도 일부에 전송하는 동작;

상기 복수의 음성 인식 장치들에 사용자의 테스트 발화를 수신하기 위해 대기하도록 하는 신호를 송신하는 동작;

상기 복수의 음성 인식 장치들로부터 상기 테스트 발화에 대응하는 발화 정보를 수신하는 동작; 및

상기 발화 정보를 기반으로 상기 복수의 음성 인식 장치들에서 상기 테스트 발화에 대한 수신 값을 표시하도록 상기 사용자 인터페이스를 업데이트하는 동작;을 포함하고,

상기 발화 정보는,

상기 복수의 음성 인식 장치들의 동작 상태, 또는 상기 복수의 음성 인식 장치들 주변의 소음 정보 중 적어도 하나를 기반으로 상기 테스트 발화를 보정한 값을 포함하는 것을 특징으로 하는 방법.

**발명의 설명**

**기술 분야**

[0001] 본 문서에서 개시되는 실시 예들은, 복수의 음성 인식 장치들을 제어하는 방법 및 그 방법을 지원하는 전자 장치에 관한 기술과 관련된다.

**배경 기술**

[0002] 사용자의 대화형 발화 입력(또는 음성 입력)에 대응하여 동작하는 다양한 전자 장치들이 출시되고 있다. 예를 들어, 스마트폰은 사용자의 음성 입력에 대응하여, 전화 걸기/받기, 메시지 송수신, 파일 전송, 웹 검색 등 다양한 기능을 수행할 수 있다.

[0003] 다른 예를 들어, TV, 스피커, 냉장고 등 다양한 장치들이 사용자의 음성 입력에 대응하여, 전원 온/오프, 동영상 재생, 음악 재생 등의 동작을 수행할 수 있다.

[0004] 사용자의 음성 입력에 대응하여 동작하는 전자 장치(이하, 음성 인식 장치)는 사용자의 음성 명령을 시작하는 발화 입력(wake up word)(예: Hi Bixby)(이하, 시작 발화)에 대응하여, 사용자의 발화 입력을 수신할 수 있는 대기 상태로 진입할 수 있다.

**발명의 내용**

**해결하려는 과제**

[0005] 종래 기술에 따른 전자 장치는 복수의 음성 인식 장치들이 음성 명령에 대응할 수 있는 경우, 시작 발화(예: Hi Bixby)에 대응하여 동시에 여러 장치들이 응답하는 문제점이 있다. 이를 해결하기 위해, 음성 명령을 시작하는 발화 입력을 수신한 모든 전자 장치가 반응하지 않고, 선택된 하나의 전자 장치만 사용자의 시작 발화에 응답하도록 하는 기술이 적용되고 있다. 상기 기술은, 사용 환경에 따라 소리를 발생하는 장치 (TV ON, 음악 Play)가 있는 경우, 사용자가 원하는 장치가 아닌 다른 장치가 사용자의 시작 발화에 대응하게 되는 문제점이 있다.

[0006] 또한, 사용자의 시작 발화에 대응할 음성 인식 장치를 사용자의 의도대로 세부적으로 설정할 수 없는 문제점이 있다.

**과제의 해결 수단**

[0007] 본 문서에 개시되는 일 실시 예에 따른 전자 장치는, 디스플레이, 마이크, 메모리, 통신 회로 및 프로세서를 포함하고, 상기 프로세서는 상기 디스플레이를 통해, 동일한 시작 발화에 대응하여 음성 인식을 시작할 수 있는 복수의 음성 인식 장치들 각각의 음성 인식 감도를 조절하기 위한 사용자 인터페이스를 표시하고, 상기 사용자 인터페이스를 통해 상기 음성 인식 감도가 변경된 경우, 상기 변경된 감도 값을 상기 복수의 음성 인식 장치들 중 적어도 일부에 전송하고, 상기 복수의 음성 인식 장치들에 사용자의 테스트 발화를 수신할 수 있도록 대기하도록 하는 신호를 송신하고, 상기 복수의 음성 인식 장치들로부터 상기 제1 발화에 대응하는 발화 정보를 수신하고, 상기 발화 정보를 기반으로 상기 사용자 인터페이스를 업데이트하도록 설정될 수 있다.

**발명의 효과**

[0008] 본 문서에 개시되는 실시 예들에 따른 전자 장치는 사용자가 복수의 음성 인식 장치들의 음성 입력 감도를 조절할 수 있는 사용자 인터페이스를 제공할 수 있다. 상기 사용자 인터페이스를 통해, 사용자는 테스트 음성을 입력하면서, 다양한 장치들의 음성 인식 감도를 확인할 수 있다. 사용자는 필요에 따라 특정 음성 인식 장치의 음성 인식 감도를 조절하여, 자신이 원하는 장치가 음성 명령에 대응하도록 조절할 수 있다.

[0009] 본 문서에 개시되는 실시 예들에 따른 전자 장치는 사용자의 장치 사용 환경에 적합하도록 음성 인식 감도를 자동으로 설정하여, 사용자가 원하는 장치가 음성 명령에 대응하도록 설정하는 유용한 방식을 제공할 수 있다. 이외에, 본 문서를 통해 직접적 또는 간접적으로 파악되는 다양한 효과들이 제공될 수 있다.

**도면의 간단한 설명**

- [0010] 도 1은 다양한 실시 예에 따른 음성 인식 시스템을 나타낸 도면이다.
- 도 2는 다양한 실시예에 따른 음성 인식 장치의 구성도이다.
- 도 3은 다양한 실시예에 따른 제어 전자 장치에서의 음성 인식 장치를 제어하는 방법에 관한 순서도이다.
- 도 4는 다양한 실시예에 따른 통합 UI를 나타낸다.
- 도 5는 다양한 실시예에 따른 음성 인식 감도의 수동 설정에 따른 화면 변화도이다.
- 도 6은 다양한 실시예에 따른 제어 전자 장치와 음성 인식 장치 사이의 신호 흐름도이다.
- 도 7은 다양한 실시예에 따른 자동으로 음성 인식 감도를 조절하는 순서도이다.
- 도 8은 다양한 실시예에 따른 통합 UI의 예시도이다.
- 도 9는 다양한 실시예에 따른 자동 보정에 따른 발화 정보의 변화를 나타낸다.
- 도 10은 다양한 실시 예들에 따른 네트워크 환경 내의 전자 장치의 블록도 이다.

**발명을 실시하기 위한 구체적인 내용**

[0011] 이하, 본 문서의 다양한 실시예가 첨부된 도면을 참조하여 기재된다. 다양한 실시예에 따른 음성 인식 장치(또는 전자 장치)가 설명된다. 본 문서에서, 사용자라는 용어는 음성 인식 장치를 사용하는 사람 또는 음성 인식 장치를 사용하는 장치(예: 인공지능 전자 장치)를 지칭할 수 있다.

[0012] 도 1은 다양한 실시 예에 따른 음성 인식 시스템을 나타낸 도면이다.

[0013] 도 1을 참조하면, 음성 인식 시스템(100)은 복수의 음성 인식 장치들(110), 네트워크 장치(120), 및 지능형 서버(130)를 포함할 수 있다.

[0014] 복수의 음성 인식 장치들(110) 각각은 사용자(150)의 발화 입력(151)을 수신할 수 있다. 예를 들어, 복수의 음성 인식 장치들(110) 각각은 마이크를 이용하여, 사용자(150)의 발화 입력(151)을 전기적 신호로 변환할 수 있다. 복수의 음성 인식 장치들(110)의 내부 프로세서는 장치 내부에서의 분석 또는 지능형 서버(130)을 통한 분석을 통해, 사용자(150)의 발화 입력(151)에 대응하는 동작을 수행할 수 있다.

[0015] 예를 들어, 음성 인식 장치(110)가 TV인 경우, 사용자의 대화형 발화 입력에 대응하여, 전원 턴온/턴오프, 채널 변경, 볼륨 변경, 또는 동영상 재생 등의 기능을 수행할 수 있다.

[0016] 다른 예를 들어, 음성 인식 장치(110)가 스피커인 경우, 사용자의 대화형 발화 입력에 대응하여, 전원 턴온/턴

오프, 볼륨 변경, 또는 음악 재생 등의 기능을 수행할 수 있다.

- [0017] 또 다른 예를 들어, 음성 인식 장치(110)가 냉장고인 경우, 사용자의 대화형 발화 입력에 대응하여, 온도 조절, 또는 음식 정보 검색 등의 기능을 수행할 수 있다.
- [0018] 다양한 실시예에 따르면, 복수의 음성 인식 장치들(110)은 음성 명령 서비스를 시작하는 시작 발화(예: Hi Bixby)가 서로 동일 할 수 있다. 예를 들어, 사용자(150)의 주변에 복수의 음성 인식 장치들(110)이 배치되는 상태에서, 사용자(150)가 "Hi, Bixby"를 말하는 경우, TV 스피커, 및 냉장고가 동시에 대응하는 음성 응답을 발생시킬 수 있다.
- [0019] 다양한 실시예에 따르면, 복수의 음성 인식 장치들(110) 중 적어도 하나(예: 전자 장치(140))는 복수의 음성 인식 장치들(110)의 시작 발화(151)에 대한 음성 인식 감도를 조절하기 위한 사용자 인터페이스(141)를 표시할 수 있다. 도 1에서는 복수의 음성 인식 장치들(110) 중 전자 장치(140)(이하, 제어 전자 장치)가 사용자 인터페이스(141)(이하, 통합 UI)를 표시하는 경우를 예시적으로 도시하였으나, 이에 한정되는 것은 아니다. 복수의 음성 인식 장치들(110) 중 다른 장치(예: TV, 냉장고, 또는 AI 스피커)가 통합 UI(141)를 표시할 수도 있다.
- [0020] 다양한 실시예에 따르면, 제어 전자 장치(140)의 프로세서는 통합 UI를 통해, 복수의 음성 인식 장치들(110)의 음성 인식 감도를 조절하는 입력(예: 터치 입력)을 수신할 수 있다. 제어 전자 장치(140)의 프로세서는 변경된 음성 인식 감도를 대응하는 음성 인식 장치(110)에 전송할 수 있다.
- [0021] 네트워크 장치(120)는 지정된 공간에서 근거리 무선 통신 또는 유선 통신을 통해 복수의 음성 인식 장치들(110) 중 적어도 일부와 외부 장치(예: 지능형 서버(130))의 네트워크를 연계하는 장치일 수 있다. 근거리 무선 통신은, 예를 들면, Wi-Fi(wireless fidelity), Wi-Fi Direct, Li-Fi(light fidelity), Bluetooth, BLE(Bluetooth low energy), Zigbee 등을 포함할 수 있다. 유선 통신은, 예를 들면, LAN(local area network), 전력선 통신 또는 POTS(plain old telephone service)를 포함할 수 있다. 예를 들어, 네트워크 장치(120)는 Wi-Fi 통신을 위한 AP 장치일 수 있다. 다른 예를 들어, 네트워크 장치(102)는 인터넷 연결을 위한 라우터일 수 있다.
- [0022] 도 1에서는 제어 전자 장치(140)는 별도의 통신을 통해 네트워크에 연결되고, 다른 음성 인식 장치들이 네트워크 장치(120)를 통해 네트워크에 연결되는 형태를 예시적으로 도시하였으나, 이에 한정되는 것은 아니다. 예를 들어, 다른 음성 인식 장치들과 같이, 전자 장치(140)도 동일한 네트워크 장치(120)에 연결될 수 있다. 또 다른 예를 들어, 전자 장치(140)는 제1 통신 방식(셀룰러 통신), 복수의 음성 인식 장치들(110) 중 일부(예: TV 및 냉장고)는 제2 통신 방식(LAN), 복수의 음성 인식 장치들(110) 중 다른 일부(예: 스피커 및 스마트폰)은 제3 통신 방식(Wi-Fi)를 통해 네트워크에 연결될 수도 있다.
- [0023] 지능형 서버(130)는 사용자(150)의 발화 입력(151)을 분석할 수 있다. 지능형 서버(130)는 복수의 음성 인식 장치들(110)로부터 발화 입력(151)의 적어도 일부를 수신할 수 있다. 지능형 서버(130)는 발화 입력(151)을 분석하여, 발화 입력(151)에 대응하는 복수의 음성 인식 장치들(110)의 동작에 관한 정보를 복수의 음성 인식 장치들(110)에 전송할 수 있다.
- [0024] 다양한 실시예에 따르면, 지능형 서버(130)는 제어 전자 장치(140)의 통합 UI(141)에 표시되거나, 설정되는 정보를 전달할 수 있다. 예를 들어, 지능형 서버(130)는 통합 UI(141)에서 변경된 음성 인식 감도를 수신하여, 대응하는 장치에 전송할 수 있다.
- [0025] 도 2는 다양한 실시예에 따른 음성 인식 장치의 구성도이다. 도 2는 예시적인 것으로 이에 한정되는 것은 아니다.
- [0026] 도 2를 참조하면, 음성 인식 장치(201)(예: 도 1의 복수의 음성 인식 장치(110) 또는 제어 전자 장치(140))는 프로세서(210), 메모리(220), 마이크(230), 디스플레이(240), 및 통신 회로(250)를 포함할 수 있다.
- [0027] 프로세서(210)는 음성 인식 장치(201)의 동작에 필요한 다양한 연산을 수행할 수 있다. 일 실시예에 따르면, 프로세서(210)는 통합 UI 출력부(211), 시작 발화 감지부(212), 장치 상태 인식부(213), 보상 필터부(214) 및 발화 정보 연산부(215)를 포함할 수 있다. 상기 구성은 프로세서(210)의 음성 인식과 관련된 기능에 따른 분류로서, 이에 한정되는 것은 아니다.
- [0028] 통합 UI 출력부(211)는 사용자 계정 기반 또는 동일한 근거리 통신 네트워크 장치(예: WIFI-AP)에 연결된 기기들의 음성 인식 감도를 설정할 수 있도록 하는 통합 UI를 제공할 수 있다.
- [0029] 시작 발화 감지부(212)는, 사용자가 시작 발화(예: "Hi, Bixby")를 발생키는 경우, 마이크(230)를 통해 수신

한 음성 입력의 음성 패턴을 분석하여, 시작 발화에 대응하는 음성인지를 인식할 수 있다.

- [0030] 일 실시예에 따르면, 시작 발화 감지부(212)는 DSP(Digital Signal Processor) 또는 저전력 칩을 이용한 인식 엔진을 통해 시작 발화를 인식할 수 있다. 이후, 시작 발화 감지부(212)는 음성신호가 담긴 오디오 데이터를 AP(Application Processor)의 인식 엔진으로 전달하여, 최종적으로 시작 발화가 발생했는지를 감지할 수 있다.
- [0031] 장치 상태 인식부(213)는 음성 인식 장치(201)의 동작 상태에 따른 상태 정보를 저장할 수 있다. 예를 들어, 상기 동작 상태 정보는 Media 재생 상황, 또는 볼륨 크기 등을 포함할 수 있다. 다른 예를 들어, 상기 동작 상태 정보는 전자 장치의 내부 구성에 의해 발생하는 소리(스피커의 소리, 냉장고 컴프레스의 동작 소리)에 관한 정보일 수 있다.
- [0032] 예를 들어, 상기 동작 상태 정보는 음성 인식 장치(201)이 음악을 재생 중인 지, 동영상을 재생 중인지, 볼륨이 음소거 상태인지, 세탁기가 탈수 중인지, 에어컨의 설정 온도가 몇도 인지 등에 관한 정보를 포함할 수 있다.
- [0033] 보상 필터부(214)는 음성 인식 장치(201)에서 발생하는 소음, 주변에서 발생하는 소음을 기반으로, SNR을 보정하기 위한 보상값을 이용한 연산을 수행할 수 있다. 상기 보상값은 미리 저장되거나, 실시간으로 계산되어 변경될 수 있다.
- [0034] 발화 정보 연산부(215)는 상기 보상값이 적용된 발화 정보(예: 오디오 음질(audio quality), 음압(sound pressure), 발화 입력의 특징값 등)을 계산 및 수집하여 저장할 수 있다. 상기 특징값은 발화 입력의 음성 패턴, 평균 음압, 최고음압과 최저 음압의 차이, 소리 간격, 소리 평균 주파수 중 적어도 하나일 수 있다.
- [0035] 메모리(220)는 음성 인식 장치(201)의 동작에 필요한 다양한 정보를 저장할 수 있다. 일 실시예에 따르면, 메모리(220)는 통합 UI(141)를 통해 설정된 음성 인식 감도를 저장할 수 있다.
- [0036] 마이크(230)는 사용자의 발화 입력을 수집하여, 전기적 신호로 변환할 수 있다. 디스플레이(240)는 음성 인식 장치(201)의 동작에 필요한 다양한 콘텐츠를 표시할 수 있다. 통신 회로(250)는 외부 장치와의 통신을 수행할 수 있다.
- [0037] 도 3은 다양한 실시예에 따른 제어 전자 장치에서의 음성 인식 장치를 제어하는 방법(300)에 관한 순서도이다.
- [0038] 도 3을 참조하면, 동작 310에서, 제어 전자 장치(140)의 프로세서(210)는 복수의 음성 인식 장치들(110) 각각의 음성 인식 감도를 조절하기 위한 통합 UI(141)를 표시할 수 있다.
- [0039] 상기 통합 UI(141)는 제어 가능한 복수의 음성 인식 장치들(110)의 목록을 포함할 수 있다. 복수의 음성 인식 장치들(110)은 동일한 시작 발화에 대응하여 동작하도록 설정된 장치일 수 있다.
- [0040] 다양한 실시 예에 따르면, 프로세서(210)는 지능형 서버(130) 또는 네트워크 장치(120)로부터 상기 목록을 수신하거나, 주변의 복수의 음성 인식 장치들(110)과 별도의 통신 채널을 형성하여 상기 목록을 생성할 수 있다.
- [0041] 다양한 실시 예에 따르면, 프로세서(210)는 지정된 조건에 따라 상기 목록의 순서를 변경할 수 있다. 예를 들어, 전원이 켜지고 동작 중인 장치, 사용자의 입력이 가장 최근에 발생한 장치, 머신 러닝을 통해 사용자의 장치 사용 패턴을 분석하여 해당 시간에 가장 사용 빈도가 높은 장치가 상기 목록에서 상위에 표시될 수 있다.
- [0042] 다양한 실시예에 따르면, 상기 통합 UI(141)는 복수의 음성 인식 장치들(110) 각각에 설정된 음성 인식 감도를 표시할 수 있다. 예를 들어, 음성 인식 감도는 이동 바(bar) 또는 숫자 입력 필드로 표시될 수 있다. 음성 인식 감도는 지정된 사용자 입력(예: 터치 입력)에 의해 변경될 수 있다.
- [0043] 동작 315에서, 프로세서(210)는 상기 통합 UI(141)를 통해 음성 인식 감도가 변경된 경우, 변경된 감도 값을 복수의 음성 인식 장치들(110) 중 대응하는 장치에 전송할 수 있다. 복수의 음성 인식 장치들(110)은 음성 인식 감도를 수신하는 경우, 이후에 발생하는 제1 발화(음성 인식 감도를 테스트 하기 위한 발화, 이하, 테스트 발화)를 보정하는데 이용할 수 있다.
- [0044] 동작 320에서, 프로세서(210)는 상기 통합 UI(141)의 목록에 포함된 복수의 음성 인식 장치들(110)에 사용자의 테스트 발화를 수신할 수 있도록 대기하는 신호를 송신할 수 있다.
- [0045] 다양한 실시예에 따르면, 상기 목록에 제어 전자 장치(140)가 포함되는 경우, 프로세서(210)는 마이크(230)를 동작시켜, 사용자의 테스트 발화를 수신할 수 있는 상태로 진입할 수 있다.
- [0046] 동작 330에서, 프로세서(210)는 상기 통합 UI(141)의 목록에 포함된 복수의 음성 인식 장치들(110)로부터 상기 테스트 발화에 대응하는 발화 정보를 수신할 수 있다. 예를 들어, 상기 발화 정보는 복수의 음성 인식 장치들



(110) 각각에서 수신한 테스트 발화의 SNR 값일 수 있다.

- [0047] 동작 340에서, 프로세서(210)는 발화 정보를 기반으로 상기 통합 UI(141)를 업데이트 할 수 있다. 예를 들어, 프로세서(210)는 상기 발화 정보에 대응하여, 상기 통합 UI(141)에서 복수의 음성 인식 장치들(110) 각각의 테스트 발화에 대한 SNR 값을 표시할 수 있다. 일 실시예에 따르면, 프로세서(210)는 SNR 값을 실시간으로 변화하는 그래프 형태로 표시될 수 있다.
- [0048] 일 실시예에 따르면, 상기 발화 정보는 통합 UI(141)에 포함된 복수의 음성 인식 장치들(110) 각각에 대하여 설정된 음성 인식 감도와 비교되는 방식으로 표시될 수 있다. 예를 들어, 설정된 음성 인식 감도가 제1 색(예: 회색)의 이동바(bar) 형태로 표시되는 경우, 상기 SNR 값은 상기 이동바에 오버랩되는 제2 색(예: 파란색)의 그래프의 형태로 표시될 수 있다.
- [0049] 예를 들어, 사용자가 통합 UI(141)를 통해, 복수의 음성 인식 장치들(110) 중 음악방송이 나오는 TV와 일반적인 냉장/냉동 기능을 수행하는 냉장고를 선택하여, 각각 음성 입력 감도를 변경한 경우, 프로세서(210)는 사용자에게 테스트 발화를 유도하는 가이드 메시지를 표시할 수 있다.
- [0050] 사용자가 가이드 메시지에 따라, 테스트 발화를 시작하여 종료하는 경우, 프로세서(210)는 TV와 냉장고 각각으로부터 발화 정보를 수신하여, 통합 UI(141)에 각각의 음성 인식 장치(TV, 냉장고)에서 테스트 입력을 수신한 수신 값을 그래프로 표시할 수 있다.
- [0051] 일 실시예에 따르면, 발화 정보가 반영된 통합 UI(141)는 사용자의 음성 세기(dB), 장치 동작 상태 정보 또는 주변 소음 정보 등을 기반으로, 음성 인식 상태가 좋지 않은 장치 표시(예: 아이콘, 색, 텍스트 등)를 나타낼 수 있다. 예를 들어, 음성 인식 장치(110) 각각은 설정된 음성 인식 감도, 동작 상태 정보, 또는 주변 소음 정보 등을 반영한 발화 정보를 기반으로 응답할 수 있는 기기(발화 정보에 의한 수신 값이, 음성 인식 감도의 기준값을 초과하는지), 응답을 하기에는 주변 소음/자체 소음이 커 기기의 인식률이 떨어질 장치인지 등을 판단하여, 해당 정보를 함께 제어 전자 장치(110)에 함께 전송할 수 있다.
- [0052] 통합 UI(141)는 해당 정보를 전송 받은 후 아이콘, 색깔, 텍스트 등을 통해 사용자에게 해당 음성 인식 장치(110)가 잡음 등의 영향으로 음성 인식 상태가 좋지 않음을 사용자에게 알릴 수 있다. 일 실시예에서, 통합 UI(141)는 해당 기기의 볼륨, 기기 동작 상태 등을 변경(해당 기기에서 응답을 원할 경우 일시적으로 볼륨을 줄이거나 기기 동작을 pause시키는 등)하여, 음성 인식 성능을 향상시킬 수 있음을 표시할 수 있다.
- [0053] 도 4는 다양한 실시예에 따른 통합 UI를 나타낸다. 도 4는 예시적인 것으로 이에 한정되는 것은 아니다.
- [0054] 도 4를 참조하면, 제어 전자 장치(140)의 프로세서(210)는 복수의 음성 인식 장치들(110)의 시작 발화에 대한 음성 인식 감도를 조절하기 위한 통합 UI(401)(예: 도 1의 통합 UI(141))를 표시할 수 있다.
- [0055] 통합 UI(401)는 복수의 음성 인식 장치들(110)의 식별자(예: 이름, 종류 또는 모델명)(410), 음성 인식 감도를 조절하기 위한 감도 조절 객체(420), 기본값 표시(421), 테스트 발화의 수신을 제외하도록 하는 수신 제외 객체(430), 초기화 객체(440)를 포함할 수 있다.
- [0056] 식별자(410)는 복수의 음성 인식 장치들(110)을 구분하기 위한 표시일 수 있다. 일 실시예에 따르면, 식별자(410)는 동일 종류의 장치가 없으면, 종류(TV, 스피커, 또는 냉장고)로 표시될 수 있다. 식별자(410)는, 동일 종류의 장치가 있는 경우, 모델명을 함께 표시할 수 있다.
- [0057] 감도 조절 객체(420)는 사용자 입력에 대응하여, 각 장치의 음성 인식 감도를 변경하는 객체일 수 있다. 사용자가 감도 조절 객체(420)에 지정된 입력(예: 터치 입력)을 발생시켜 이동하는 경우, 프로세서(210)는 대응하는 음성 인식 장치(110)에 사용자가 조정한 음성 입력 감도에 대응하는 감도 값을 전송할 수 있다. 예를 들어, 감도 조절 객체(420)를 좌측으로 이동하는 경우, 음성 인식 감도의 기준값이 낮아질 수 있다. 이 경우, 테스트 발화에 대한 SNR 값이 낮은 경우에도, 음성 인식 장치(110)는 시작 발화에 대응하는 응답 신호를 출력할 수 있다. 다른 예를 들어, 감도 조절 객체(420)를 우측으로 이동하는 경우, 음성 인식 감도의 기준값이 높아질 수 있다. 이 경우, 테스트 발화에 대한 SNR 값이 높은 경우에도, 음성 인식 장치(110)는 시작 발화에 대응하는 응답 신호를 출력하지 않을 수 있다.
- [0058] 기본값 표시(421)는 음성 인식 감도의 기본 값을 나타낼 수 있다. 사용자가 감도의 기준을 확인하면서, 감도 조절 객체(420)를 이동할 수 있다.
- [0059] 수신 제외 객체(430)는 대응하는 음성 인식 장치가 사용자의 시작 발화 또는 테스트 발화를 수신하지 않도록 설

정하기 위한 객체일 수 있다. 수신 제외 객체(430)가 체크된 경우, 감도 조절 객체(420)가 비활성화 될 수 있다.

- [0060] 초기화 박스(440)는 목록에 포함된 모든 감도 조절 객체(420)를 기본값으로 변경하는 객체일 수 있다. 초기화 박스(440)가 체크되는 경우, 모든 감도 조절 객체(420)가 초기화될 수 있다.
- [0061] 도 5는 다양한 실시예에 따른 음성 인식 감도의 수동 설정에 따른 화면 변화도이다.
- [0062] 도 5를 참조하면, 제어 전자 장치(140)의 프로세서(210)는 음성 인식 감도의 설정을 유도하는 제1 사용자 인터페이스(510)를 표시할 수 있다. 일 실시예에 따르면, 사용자 계정으로 등록된 복수의 음성 인식 장치들(110)을 관리하는 어플리케이션을 실행되고, 설정 메뉴에서 수신 감도를 설정하는 옵션이 선택되는 경우, 프로세서(210)는 제1 사용자 인터페이스(510)를 표시할 수 있다.
- [0063] 제1 사용자 인터페이스(510)는 가이드 메시지(511) 및 시작 버튼(512)를 포함할 수 있다. 가이드 메시지(511)는 음성 인식 감도의 수동 설정을 위해 사용자가 해야 할 동작을 설명할 수 있다. 시작 버튼(512)은 사용자의 터치 입력에 대응하여, 음성 수신 감도를 설정하는 과정을 시작하도록 할 수 있다. 일 실시예에 따르면, 사용자가 시작 버튼(512)을 터치하는 경우, 프로세서(210)는 각각의 음성 인식 장치(110)에 사용자의 테스트 발화의 수신을 대기하도록 하는 신호를 전송할 수 있다.
- [0064] 사용자가 시작 버튼(512)을 터치하는 경우, 프로세서(210)는 제2 사용자 인터페이스(520)를 표시할 수 있다. 제2 사용자 인터페이스(520)는 테스트 발화 표시(521) 및 완료 버튼(522)를 포함할 수 있다.
- [0065] 테스트 발화 표시(521)는 사용자가 발화해야 하는 테스트 발화(예: “Bixby, can you tell me the weather in Seoul”)를 표시할 수 있다. 도 5에서는 텍스트 방식으로 표시하는 경우를 예시적으로 도시하였으나, 이에 한정되는 것은 아니다. 예를 들어, 테스트 발화 표시(521)는 제어 전자 장치(140)의 스피커를 통해 음성으로 안내될 수도 있다.
- [0066] 사용자가 테스트 발화 표시(521)에 대응한 테스트 발화를 생성하고, 완료 버튼(522)을 터치하는 경우, 프로세서(210)는 제3 사용자 인터페이스(530)를 표시할 수 있다. 제3 사용자 인터페이스(530)는 통합 UI(531)를 포함할 수 있다. 통합 UI(531)는 도 4에서의 통합 UI(401)과 동일 또는 유사할 수 있다.
- [0067] 감도 조절 객체(532)는 사용자 입력에 대응하여 각 장치의 음성 인식 감도를 변경하는 객체일 수 있다. 사용자가 감도 조절 객체(420)에 지정된 입력(예: 터치 입력)을 발생시켜 이동하는 경우, 프로세서(210)는 대응하는 음성 인식 장치(110)에 사용자가 조정한 음성 입력 감도에 대응하는 감도 값을 전송할 수 있다.
- [0068] 발화 정보 그래프(533)는 각각의 음성 인식 장치에서 수신한 사용자의 테스트 발화 발화 정보에 기반한 테스트 발화의 발화 정보(예: SNR 값)를 표시할 수 있다.
- [0069] 일 실시예에 따르면, 상기 발화 정보는 통합 UI(531)에 포함된 복수의 음성 인식 장치들(110) 각각에 대하여 설정된 음성 인식 감도와 비교되는 방식으로 표시될 수 있다. 예를 들어, 설정된 음성 인식 감도가 제1 색(예: 회색)의 이동바 형태로 표시되는 경우, 상기 발화 정보는 상기 이동바에 오버랩되는 제2 색(예: 파란색)의 그래프의 형태로 표시될 수 있다.
- [0070] 다양한 실시예에 따르면, 프로세서(120)는 음성 인식 장치들(110) 각각의 인식률에 따라 감도 그래프(533)를 구분하여 표시할 수 있다. 예를 들어, 테스트 발화에 대응할 수 있는 음성 인식 장치들(110) 중 TV 또는 스피커는 자체 소음 등으로 인식률이 떨어질 수 있는 점을 반영하여, 감도 그래프(533)를 빨간색으로 표시할 수 있다. 다른 예를 들어, 테스트 발화에 대응할 수 있는 장치들 중 전등 또는 냉장고는 자체 소음이 상대적으로 작을 수 있는 점을 반영하여, 감도 그래프(533)를 녹색으로 표시할 수 있다.
- [0071] 도 6은 다양한 실시예에 따른 제어 전자 장치와 음성 인식 장치 사이의 신호 흐름도(600)이다.
- [0072] 도 6을 참조하면, 동작 610에서, 제어 전자 장치(140)는 복수의 음성 인식 장치들(110) 각각의 음성 인식 감도를 조절하기 통합 UI(141)를 표시할 수 있다. 통합 UI(141)는 복수의 음성 인식 장치들(110)의 식별자(예: 이름, 종류 또는 모델명) 및 음성 인식 감도를 조절하기 위한 객체 등을 포함할 수 있다.
- [0073] 동작 615에서, 제어 전자 장치(140)는 상기 통합 UI(141)를 통해 음성 인식 감도가 변경된 경우, 변경된 감도 값을 대응하는 음성 인식 장치(110)에 전송할 수 있다.
- [0074] 동작 620에서, 제어 전자 장치(140)는 상기 통합 UI(141)의 목록에 포함된 복수의 음성 인식 장치들(110)에 사

용자의 테스트 발화를 수신할 수 있도록 대기하는 신호를 송신할 수 있다.

- [0075] 다양한 실시예에 따르면, 동작 625에서, 음성 인식 장치(110)는 음성 인식 장치(110)의 동작 상태 정보 및 주변 소음 정보를 확인할 수 있다. 예를 들어, 상기 동작 상태 정보는 Media 재생 상황, 또는 볼륨 크기 등을 포함할 수 있다. 주변 소음 정보는 음성 인식 장치(110)의 스피커를 통해 수집된 노이즈 신호의 세기 및 패턴에 관한 정보일 수 있다.
- [0076] 다양한 실시예에 따르면, 음성 인식 장치(110) 각각은 전원 턴온/턴오프, 볼륨 변경 등의 명령어를 수신하는 경우, 미리 동작 상태 정보 및 주변 소음 정보를 계산하여 저장할 수 있다.
- [0077] 일 실시예에 따르면, 음성 인식 장치(110) 각각은(예: Intelligence Agent)는 사용자의 테스트 발화를 수신하지 않은 경우에도, 상기 정보를 미리 저장할 수 있다.
- [0078] 동작 630에서, 음성 인식 장치(110)는 스피커를 통해 테스트 발화를 수신할 수 있다. 일 실시예에 따르면, 테스트 발화는 음성 인식 장치(110)의 시작 발화와 동일하게 설정(예: "Hi, Bixby")되거나, 다르게 설정(예: 시작 발화 "Hi, Bixby", 테스트 발화 "Bixby, can you tell me the weather in Seoul")될 수도 있다.
- [0079] 동작 640에서, 음성 인식 장치(110)는 수신한 변경된 감도 값, 동작 상태 정보 및 주변 소음 정보를 기반으로 수신된 테스트 발화에 대한 보상값을 계산할 수 있다. 예를 들어, 음성 인식 장치(110)가 TV 장치이고, 동영상 을 재생 중인 경우, 음성 인식 장치(110)는 보상값을 높일 수 있다.
- [0080] 동작 650에서, 음성 인식 장치(110)는 보상 값을 기반으로 수신한 테스트 발화에 대한 발화 정보를 계산할 수 있다. 상기 발화 정보는 보상값이 적용된, 오디오 음질(audio quality), 음압(sound pressure), 테스트 발화의 음성 신호, 테스트 발화의 특징값 등을 포함할 수 있다.
- [0081] 음성 인식 장치(110)의 프로세서는 사용자가 통합 UI(510)를 통해 변경한 음성 인식 감도를 마이크 볼륨 및 SNR 값을 계산하기 위한 보상값 등의 계산 또는 결정에 이용할 수 있다.
- [0082] 음성 인식 장치(110) 각각은 제어 전자 장치(140)으로부터 수신한 음성 입력 감도/동작 상태 정보/주변 소음 정보 중 적어도 하나를 기반으로 수신한 테스트 발화에 보상값을 계산하거나, 미리 지정해 놓은 보상값들(예: 보상값에 대한 테이블) 중 하나를 선택할 수 있다.
- [0083] 예를 들어, 수신한 음성 입력 감도가 상대적으로 낮은 값으로 설정된 경우, 음성 인식 장치(110)는 보상값을 높일 수 있다. SNR이 낮은 경우에도, 음성 인식 장치(110)는 시작 발화에 대응하는 응답 신호를 출력할 수 있다. 반대로, 수신한 음성 입력 감도가 상대적으로 높은 값으로 설정된 경우, 음성 인식 장치(110)는 보상값을 낮출 수 있다. SNR이 높은 경우에도, 음성 인식 장치(110)는 시작 발화에 대응하는 응답 신호를 출력하지 않을 수 있다.
- [0084] 다른 예를 들어, 음성 인식 장치(110)의 내부 구성(예: 스피커)에 의해 소리가 발생 중인 경우, 음성 인식 장치(110)는 보상값을 높일 수 있다. 반대로, 음성 인식 장치(110)의 내부 구성(예: 스피커)에 의해 소리가 발생하지 않는 경우, 음성 인식 장치(110)는 보상값을 낮출 수 있다.
- [0085] 또 다른 예를 들어, 음성 인식 장치(110)의 주변 소음이 미리 설정된 기준값 보다 큰 경우, 음성 인식 장치(110)는 보상값을 높일 수 있다. 반대로, 음성 인식 장치(110)의 주변 소음이 미리 설정된 기준값 보다 작은 경우, 음성 인식 장치(110)는 보상값을 낮출 수 있다.
- [0086] 동작 660에서, 음성 인식 장치(110)는 발화 정보를 제어 전자 장치(140)에 전송할 수 있다.
- [0087] 일 실시 예에 따르면, 음성 인식 장치(110)는 지능형 서버(130) 또는 네트워크 장치(120)의 의해 설정되는 주소 정보(예: IP 주소)를 기반으로 발화 정보를 제어 전자 장치(140)에 전송할 수 있다.
- [0088] 다양한 실시 예에 따르면, 음성 인식을 지원하는 복수의 음성 인식 장치들(110)이 동일한 계정으로 지능형 서버(130), 또는 네트워크 장치(또는 허브 디바이스)(120)에 등록되어 있는 경우, 음성 인식 장치(110) 각각은 테스트 발화에 대한 입력 감도 및 발화 정보를 지능형 서버(130), 또는 네트워크 장치(또는 허브 디바이스)(120)에 전송할 수 있다. 지능형 서버(130) 또는 네트워크 장치(120)는 음성 인식 장치(110) 각각에서 수신 받은 발화 정보를 제어 전자 장치(140)에 전송할 수 있다.
- [0089] 동작 670에서, 제어 전자 장치(140)는 발화 정보를 기반으로 통합 UI(141)를 업데이트 할 수 있다. 제어 전자 장치(140)는 발화 정보에 대응하여, 상기 통합 UI(141)에서, 복수의 음성 인식 장치들(110) 각각의 테스트 발화

에 대한 수신 값을 표시할 수 있다. 일 실시예에 따르면, 프로세서(210)는 실시간으로 변화하는 그래프 형태로 표시될 수 있다.

- [0090] 도 7은 다양한 실시예에 따른 자동으로 음성 인식 감도를 조절하는 순서도(700)이다.
- [0091] 도 7을 참조하면, 동작 710에서, 음성 인식 장치(110)는 제어 전자 장치(140)으로부터 자동으로 음성 인식 감도를 조절하도록 하는 신호를 수신할 수 있다.
- [0092] 예를 들어, 제어 전자 장치(140)는 통합 UI(141)를 표시할 수 있다. 통합 UI(141)는 음성 인식 장치(110) 각각에서 자동으로 음성 인식 감도를 조절하도록 하기 위한 객체(예: 선택 버튼 또는 체크 박스)를 포함할 수 있다. 상기 객체에 지정된 사용자 입력(예: 터치 입력)이 발생하는 경우, 제어 전자 장치(140)는 음성 인식 장치(110) 각각에 음성 인식 감도를 자동으로 설정하도록 지시하는 신호를 송신할 수 있다.
- [0093] 동작 720에서, 음성 인식 장치(110)는 마이크를 통해, 음성 인식 장치(110)의 동작 상태 정보 또는 주변의 소음 정보 중 적어도 하나를 수집할 수 있다. 예를 들어, 동작 상태 정보는 Media 재생 상황, 또는 볼륨 크기 등을 포함할 수 있다.
- [0094] 주변 소음 정보는 음성 인식 장치(110)의 스피커를 통해 수집된 노이즈 신호의 세기 및 패턴에 관한 정보일 수 있다. 일 실시예에서, 음성 인식 장치(110)는 주기적으로 또는 지정된 동작이 실행되는 경우(예: 전원 턴온/턴오프, 또는 볼륨 조정 등), 주변의 소음 정보를 수집할 수 있다.
- [0095] 일 실시 예에 따르면, 음성 인식 장치(110) 미리 설정에 놓은 주기에 따라 주변 소음 정보를 수집할 수 있다. 다른 일 실시 예에 따르면, 음성 인식 장치(110)는 사용자의 발화 입력이 수신될 예정인 경우(예: 테스트 발화의 수신을 대기하는 신호를 수신하는 경우), 수신 중인 경우, 수신 완료된 경우, 주변 소음 정보를 수집할 수 있다.
- [0096] 동작 730에서, 음성 인식 장치(110)는 사용자의 테스트 발화를 수신할 수 있다. 일 실시예에 따르면, 테스트 발화는 음성 인식 장치(110)의 시작 발화와 동일하게 설정(예: "Hi, Bixby")되거나, 다르게 설정(예: 시작 발화 "Hi, Bixby", 테스트 발화 "Bixby, can you tell me the weather in Seoul")될 수도 있다.
- [0097] 동작 740에서, 음성 인식 장치(110)는 상기 동작 상태 정보 또는 상기 소음 정보 중 적어도 하나를 기반으로 수신한 테스트 발화를 보상하여 발화 정보를 계산할 수 있다. 상기 발화 정보는 테스트 발화에 대한 오디오 품질(audio quality), 음압(sound pressure), 발화 입력의 특징값 등을 포함할 수 있다.
- [0098] 음성 인식 장치(110)는 상기 동작 상태 정보 또는 상기 소음 정보 중 적어도 하나를 기반으로 보상값을 결정할 수 있다. 예를 들어, 음성 인식 장치(110)가 동작 상태인 경우, 보상값이 (+)일 수 있다. 주변 소음이 높은 경우, 보상값이 (+)일 수 있다.
- [0099] 음성 인식 장치(110)는 보상 값을 기반으로 수신한 테스트 발화에 대한 증폭 값 및 음성시작점, 음성 끝점, 음성 음량, 잡음 음량 등을 계산할 수 있다.
- [0100] 동작 750에서, 음성 인식 장치(110)는 상기 발화 정보를 외부 장치에 전송할 수 있다. 음성 인식 장치(110)는 상기 발화 정보를 제어 전자 장치(140) 또는 지능형 서버(130)에 전송할 수 있다. 제어 전자 장치(140)은 발화 정보를 기반으로 통합 UI(141)를 업데이트할 수 있다.
- [0101] 도 8은 다양한 실시예에 따른 음성 인식 감도의 자동 설정에 관한 통합 UI의 예시도이다. 도 8은 예시적인 것으로 이에 한정되는 것은 아니다.
- [0102] 도 8을 참조하면, 제어 전자 장치(140)의 프로세서(210)는 복수의 음성 인식 장치들(110)의 시작 발화에 대한 음성 인식 감도를 조절을 위한 통합 UI(801)(예: 도 1의 통합 UI(141))를 표시할 수 있다.
- [0103] 통합 UI(801)는 복수의 음성 인식 장치들(110)의 식별자(예: 이름, 종류 또는 모델명)(810), 음성 인식 감도를 조절하기 위한 감도 조절 객체(820), 기본값 표시(821), 테스트 발화의 수신을 제외하도록 하는 수신 제외 객체(830), 초기화 객체(840)를 포함할 수 있다.
- [0104] 다양한 실시예에 따르면, 통합 UI(801)는 복수의 음성 인식 장치들(110) 각각에서 음성 인식 감도를 자동으로 조절하도록 하는 객체(850)(이하, 자동 조절 객체)를 더 포함할 수 있다.
- [0105] 도 8에서는, 자동 조절 객체(850)가 선택 버튼으로 구현되는 경우를 예시적으로 도시하였으나, 이에 한정되는 것은 아니다. 예를 들어, 자동 조절 객체(850)는 체크 박스 형태로 구현될 수도 있다.

- [0106] 사용자 입력에 의해 자동 조절 객체(850)를 활성화 상태로 변경되는 경우, 제어 전자 장치(140)는 음성 인식 장치(110) 각각에 음성 인식 감도를 자동으로 설정하도록 지시하는 신호를 송신할 수 있다. 또한, 이 경우, 제어 전자 장치(140)는 통합 UI(801)에서, 감도 조절 객체(820)를 비활성화하여, 사용자가 조작할 수 없도록 할 수 있다.
- [0107] 도 9는 다양한 실시예에 따른 자동 보정에 따른 발화 정보의 변화를 나타낸다. 도 9는 예시적인 것으로 이에 한정되는 것은 아니다.
- [0108] 도 9를 참조하면, 테이블(901)에서, 음성 인식 장치(110)는 사용자의 테스트 발화를 수신하고, 동작 상태 정보 또는 주변의 소음 정보 중 적어도 하나를 기반으로 수신 값을 보정할 수 있다. 도 9에서는 동작 상태 정보를 기반을 보정하는 경우를 예시적으로 도시하였으나, 이에 한정되는 것은 아니다.
- [0109] 예를 들어, 휴대 전화, TV, 냉장고, AI 스피커 등의 음성 인식 장치(110)는 각각 사용자의 테스트 발화를 수신하고, SNR을 수집할 수 있다.
- [0110] 각각의 음성 인식 장치(110)는 동작 상태에 따라, SNR을 보정할 수 있다. TV가 턴온되어, 소리를 발생시키는 상태인 경우, TV는 볼륨 크기를 기반으로 테스트 발화에 대한 SNR 값을 높게 변경할 수 있다. 볼륨 크기가 클수록 추가되는 보상 값도 더 커질 수 있다. 다른 예를 들어, AI 스피커가 턴온되어, 음악을 재생 중인 경우, AI 스피커는 볼륨 크기를 기반으로 테스트 발화에 대한 SNR 값을 높게 변경할 수 있다. 볼륨 크기가 클수록 추가되는 보상 값도 더 커질 수 있다.
- [0111] 각각의 음성 인식 장치(110)는 보정된 SNR 값을 제어 전자 장치(140)로 전송할 수 있다.
- [0112] 도 10은 다양한 실시 예들에 따른 네트워크 환경(2000) 내의 전자 장치(2001)의 블록도 이다. 본 문서에 개시된 다양한 실시 예들에 따른 전자 장치는 다양한 형태의 장치가 될 수 있다. 전자 장치는, 예를 들면, 휴대용 통신 장치(예: 스마트폰), 컴퓨터 장치(예: PDA(personal digital assistant), 태블릿 PC(tablet PC), 랩탑 PC(데스크톱 PC, 워크스테이션, 또는 서버), 휴대용 멀티미디어 장치(예: 전자 책 리더기 또는 MP3 플레이어), 휴대용 의료 기기(예: 심박, 혈당, 혈압, 또는 체온 측정기), 카메라, 또는 웨어러블 장치 중 적어도 하나를 포함할 수 있다. 웨어러블 장치는 액세서리 형(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 부착 형(예: 스킨 패드 또는 문신), 또는 생체 이식 형 회로 중 적어도 하나를 포함할 수 있다. 어떤 실시 예들에서, 전자 장치는, 예를 들면, 텔레비전, DVD(digital video disk) 플레이어, 오디오 장치, 오디오 액세서리 장치(예: 스피커, 헤드폰, 또는 헤드 셋), 냉장고, 에어컨, 청소기, 오븐, 전자레인지, 세탁기, 공기 청정기, 섯톱 박스, 홈 오토메이션 컨트롤 패널, 보안 컨트롤 패널, 게임 콘솔, 전자 사진, 전자 키, 캠코더, 또는 전자 액자 중 적어도 하나를 포함할 수 있다.
- [0113] 다른 실시 예에서, 전자 장치는 네비게이션 장치, 위성 항법 시스템(GNSS(global navigation satellite system)), EDR(event data recorder)(예: 차량/선박/비행기 용 블랙박스(black box)), 자동차 인포테인먼트 장치(예: 차량용 헤드-업 디스플레이), 산업용 또는 가정용 로봇, 드론(drone), ATM(automated teller machine), POS(point of sales) 기기, 계측 기기(예: 수도, 전기, 또는 가스 계측 기기), 또는 사물 인터넷 장치(예: 전구, 스프링클러 장치, 화재 경보기, 온도 조절기, 또는 가로등) 중 적어도 하나를 포함할 수 있다. 본 문서의 실시 예에 따른 전자 장치는 전술한 기기들에 한정되지 않으며, 또한, 예를 들면, 개인의 생체 정보(예: 심박 또는 혈당)의 측정 기능이 구비된 스마트폰의 경우처럼, 복수의 장치들의 기능들을 복합적으로 제공할 수 있다. 본 문서에서, 사용자라는 용어는 전자 장치를 사용하는 사람 또는 전자 장치를 사용하는 장치(예: 인공지능 전자 장치)를 지칭할 수 있다.
- [0114] 도 10을 참조하여, 네트워크 환경(2000)에서 전자 장치(2001)(예: 도 1의 복수의 음성 인식 장치(110) 또는 제어 전자 장치(140))는 근거리 무선 통신(2098)을 통하여 전자 장치(2002)와 통신하거나, 또는 네트워크(2099)를 통하여 전자 장치(2004) 또는 서버(2008)와 통신할 수 있다. 일 실시 예에 따르면, 전자 장치(2001)는 서버(2008)를 통하여 전자 장치(2004)와 통신할 수 있다.
- [0115] 일 실시 예에 따르면, 전자 장치(2001)는 버스(2010), 프로세서(2020)(예: 도 2의 프로세서(140)), 메모리(2030)(예: 도 2의 메모리(220)), 입력 장치(2050)(예: 마이크 또는 마우스), 표시 장치(2060), 오디오 모듈(2070), 센서 모듈(2076), 인터페이스(2077), 햅틱 모듈(2079), 카메라 모듈(2080), 전력 관리 모듈(2088), 및 배터리(2089), 통신 모듈(2090)(예: 도 2의 통신 회로(250)), 및 가입자 식별 모듈(2096)을 포함할 수 있다. 어떤 실시 예에서는, 전자 장치(2001)는 구성요소들 중 적어도 하나(예: 표시 장치(2060) 또는 카메라 모듈

(2080))를 생략하거나 다른 구성요소를 추가적으로 구비할 수 있다.

- [0116] 버스(2010)는, 구성요소들(2020-2090)을 서로 연결하고, 구성요소들 간의 신호(예: 제어 메시지 또는 데이터)를 전달하는 회로를 포함할 수 있다.
- [0117] 프로세서(2020)는, 중앙처리장치(central processing unit, CPU), 어플리케이션 프로세서(application processor, AP), GPU(graphics processing unit), 카메라의 ISP(image signal processor), 또는 CP(communication processor) 중 하나 또는 그 이상을 포함할 수 있다. 일 실시 예에 따르면, 프로세서(2020)는 SoC(system on chip) 또는 SiP(system in package)로 구현될 수 있다. 프로세서(2020)는, 예를 들면, 운영 체제 또는 응용 프로그램을 구동하여 프로세서(2020)에 연결된 전자 장치(2001)의 적어도 하나의 다른 구성요소(예: 하드웨어 또는 소프트웨어 구성요소)을 제어할 수 있고, 각종 데이터 처리 및 연산을 수행할 수 있다. 프로세서(2020)는 다른 구성요소들(예: 통신 모듈(2090)) 중 적어도 하나로부터 수신된 명령 또는 데이터를 휘발성 메모리(2032)에 로드 하여 처리하고, 결과 데이터를 비 휘발성 메모리(2034)에 저장할 수 있다.
- [0118] 메모리(2030)는, 휘발성 메모리(2032) 또는 비 휘발성 메모리(2034)를 포함할 수 있다. 휘발성 메모리(2032)는, 예를 들면, RAM(random access memory)(예: DRAM, SRAM, 또는 SDRAM)로 구성될 수 있다. 비 휘발성 메모리(2034)는, 예를 들면, PROM(programmable read-only memory), OTPROM(one time PROM), EPROM(erasable PROM), EEPROM(electrically EPROM), mask ROM, flash ROM, 플래시 메모리, HDD(hard disk drive), 또는 SSD(solid state drive)로 구성될 수 있다. 또한, 비 휘발성 메모리(2034)는, 전자 장치(2001)와의 연결 형태에 따라, 그 안에 배치된 내장 메모리(2036), 또는 필요 시에만 연결하여 사용 가능한 스탠드-얼론(stand-alone) 형태의 외장 메모리(2038)로 구성될 수 있다. 외장 메모리(2038)는 플래시 드라이브(flash drive), 예를 들면, CF(compact flash), SD(secure digital), Micro-SD, Mini-SD, xD(extreme digital), MMC(multi-media card), 또는 메모리 스틱을 포함할 수 있다. 외장 메모리(2038)는 유선(예: 케이블 또는 USB(universal serial bus)) 또는 무선(예: Bluetooth)을 통하여 전자 장치(2001)와 기능적으로 또는 물리적으로 연결될 수 있다.
- [0119] 메모리(2030)는, 예를 들면, 전자 장치(2001)의 적어도 하나의 다른 소프트웨어 구성요소, 예를 들어, 프로그램(2040)에 관계된 명령 또는 데이터를 저장할 수 있다. 프로그램(2040)은, 예를 들면, 커널(kernel)(2041), 라이브러리(library)(2043), 어플리케이션 프레임워크(application framework)(2045), 또는 어플리케이션 프로그램(interchangeably "어플리케이션")(2047)을 포함할 수 있다.
- [0120] 입력 장치(2050)는, 마이크, 마우스, 또는 키보드를 포함할 수 있다. 일 실시 예에 따르면, 키보드는 물리적인 키보드로 연결되거나, 표시 장치(2060)를 통해 가상 키보드로 표시될 수 있다.
- [0121] 표시 장치(2060)는, 디스플레이, 홀로그램 장치, 또는 프로젝터 및 해당 장치를 제어하기 위한 제어 회로를 포함할 수 있다. 디스플레이는, 예를 들면, 액정 디스플레이(LCD), 발광 다이오드(LED) 디스플레이, 유기 발광 다이오드(OLED) 디스플레이, 마이크로 전자기계 시스템(MEMS) 디스플레이, 또는 전자 종이(electronic paper) 디스플레이를 포함할 수 있다. 디스플레이는, 일 실시 예에 따르면, 유연하게, 투명하게, 또는 착용할 수 있게 구현될 수 있다. 디스플레이는 사용자의 터치, 제스처, 근접, 또는 호버링(hovering) 입력을 감지할 수 터치 회로(touch circuitry) 또는 터치에 대한 압력의 세기를 측정할 수 있는 압력 센서(interchangeably "force sensor")를 포함할 수 있다. 상기 터치 회로 또는 압력 센서는 디스플레이와 일체형으로 구현되거나, 또는 디스플레이와는 별도의 하나 이상의 센서들로 구현될 수 있다. 홀로그램 장치는 빛의 간섭을 이용하여 입체 영상을 허공에 보여줄 수 있다. 프로젝터는 스크린에 빛을 투사하여 영상을 표시할 수 있다. 스크린은, 예를 들면, 전자 장치(2001)의 내부 또는 외부에 위치할 수 있다.
- [0122] 오디오 모듈(2070)은, 예를 들면, 소리와 전기 신호를 쌍방향으로 변환시킬 수 있다. 일 실시 예에 따르면, 오디오 모듈(2070)은, 입력 장치(2050)(예: 마이크)를 통해 소리를 획득하거나, 또는 전자 장치(2001)에 포함된 출력 장치(미 도시)(예: 스피커 또는 리시버), 또는 전자 장치(2001)와 연결된 외부 전자 장치(예: 전자 장치(2002)(예: 무선 스피커 또는 무선 헤드폰) 또는 전자 장치(2006)(예: 유선 스피커 또는 유선 헤드폰))를 통해 소리를 출력할 수 있다.
- [0123] 센서 모듈(2076)은, 예를 들면, 전자 장치(2001)의 내부의 작동 상태(예: 전력 또는 온도), 또는 외부의 환경 상태(예: 고도, 습도, 또는 밝기)를 측정 또는 감지하여, 그 측정 또는 감지된 상태 정보에 대응하는 전기 신호 또는 데이터 값을 생성할 수 있다. 센서 모듈(2076)은, 예를 들면, 제스처 센서, 자이로 센서, 기압 센서, 마그네틱 센서, 가속도 센서, 그립 센서, 근접 센서, 컬러(color) 센서(예: RGB(red, green, blue) 센서), IR(infrared) 센서, 생체 센서(예: 홍채 센서, 지문 센서, 또는 HRM(heartbeart rate monitoring) 센서, 후각

(electronic nose) 센서, EMG(electromyography) 센서, EEG(Electroencephalogram) 센서, ECG(Electrocardiogram) 센서), 온도 센서, 습도 센서, 조도 센서, 또는 UV(ultra violet) 센서를 포함할 수 있다. 센서 모듈(2076)은 그 안에 속한 적어도 하나 이상의 센서들을 제어하기 위한 제어 회로를 더 포함할 수 있다. 어떤 실시 예에서는, 전자 장치(2001)는 프로세서(2020) 또는 프로세서(2020)와는 별도의 프로세서(예: 센서 허브)를 이용하여, 센서 모듈(2076)을 제어할 수 있다. 별도의 프로세서(예: 센서 허브)를 이용하는 경우에, 전자 장치(2001)는 프로세서(2020)가 슬립(sleep) 상태에 있는 동안, 프로세서(2020)를 깨우지 않고 별도의 프로세서의 작동에 의하여 센서 모듈(2076)의 동작 또는 상태의 적어도 일부를 제어할 수 있다.

[0124] 인터페이스(2077)는, 일 실시 예에 따르면, HDMI(high definition multimedia interface), USB, 광 인터페이스(optical interface), RS-232(recommended standard 232), D-sub(D-subminiature), MHL(mobile high-definition link) 인터페이스, SD카드/MMC(multi-media card) 인터페이스, 또는 오디오 인터페이스를 포함할 수 있다. 연결 단자(2078)는 전자 장치(2001)와 전자 장치(2006)를 물리적으로 연결시킬 수 있다. 일 실시 예에 따르면, 연결 단자(2078)는, 예를 들면, USB 커넥터, SD 카드/MMC 커넥터, 또는 오디오 커넥터(예: 헤드폰 커넥터)를 포함할 수 있다.

[0125] 햅틱 모듈(2079)은 전기적 신호를 기계적인 자극(예: 진동 또는 움직임) 또는 전기적인 자극으로 변환할 수 있다. 예를 들면, 햅틱 모듈(2079)은 사용자에게 촉각 또는 운동 감각과 관련된 자극을 제공할 수 있다. 햅틱 모듈(2079)은 예를 들면, 모터, 압전 소자, 또는 전기 자극 장치를 포함할 수 있다.

[0126] 카메라 모듈(2080)은, 예를 들면, 정지 영상 및 동영상 촬영할 수 있다. 카메라 모듈(2080)은, 일 실시 예에 따르면, 하나 이상의 렌즈(예: 광각 렌즈 및 망원 렌즈, 또는 전면 렌즈 및 후면 렌즈), 이미지 센서, 이미지 시그널 프로세서, 또는 플래시(예: 발광 다이오드 또는 제논 램프(xenon lamp) 등)를 포함할 수 있다.

[0127] 전력 관리 모듈(2088)은 전자 장치(2001)의 전력을 관리하기 위한 모듈로서, 예를 들면, PMIC(power management integrated circuit)의 적어도 일부로서 구성될 수 있다.

[0128] 배터리(2089)는, 예를 들면, 1차 전지, 2차 전지, 또는 연료 전지를 포함하여 외부 전원에 의해 재충전되어, 상기 전자 장치(2001)의 적어도 하나의 구성 요소에 전력을 공급할 수 있다.

[0129] 통신 모듈(2090)은, 예를 들면, 전자 장치(2001)와 외부 장치(예: 제1 외부 전자 장치(2002), 제2 외부 전자 장치(2004), 또는 서버(2008)) 간의 통신 채널 수립 및 수립된 통신 채널을 통한 유선 또는 무선 통신의 수행을 지원할 수 있다. 일 실시 예에 따르면, 통신 모듈(2090)은 무선 통신 모듈(2092) 또는 유선 통신 모듈(2094)을 포함하고, 그 중 해당하는 통신 모듈을 이용하여 제1 네트워크(2098)(예: Bluetooth 또는 IrDA(infrared data association)와 같은 근거리 통신 네트워크) 또는 제2 네트워크(2099)(예: 셀룰러 네트워크와 같은 원거리 통신 네트워크)를 통하여 외부 장치와 통신할 수 있다.

[0130] 무선 통신 모듈(2092)은, 예를 들면, 셀룰러 통신, 근거리 무선 통신, 또는 GNSS 통신을 지원할 수 있다. 셀룰러 통신은, 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)을 포함할 수 있다. 근거리 무선 통신은, 예를 들면, Wi-Fi(wireless fidelity), Wi-Fi Direct, Li-Fi(light fidelity), Bluetooth, BLE(Bluetooth low energy), Zigbee, NFC(near field communication), MST(magnetic secure transmission), RF(radio frequency), 또는 BAN(body area network)을 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)을 포함할 수 있다. 본 문서에서 "GPS"는 "GNSS"와 상호 호환적으로 사용될 수 있다.

[0131] 일 실시 예에 따르면, 상기 무선 통신 모듈(2092)은, 셀룰러 통신을 지원하는 경우, 예를 들면, 가입자 식별 모듈(2096)을 이용하여 통신 네트워크 내에서 전자 장치(2001)의 구별 및 인증을 수행할 수 있다. 일 실시 예에 따르면, 무선 통신 모듈(2092)은 프로세서(2020)(예: AP)와 별개인 CP를 포함할 수 있다. 이런 경우, CP는, 예를 들면, 프로세서(2020)가 인액티브(예: 슬립) 상태에 있는 동안 프로세서(2020)를 대신하여, 또는 프로세서(2020)가 액티브 상태에 있는 동안 프로세서(2020)와 함께, 전자 장치(2001)의 구성요소들(2010-2096) 중 적어도 하나의 구성 요소와 관련된 기능들의 적어도 일부 기능을 수행할 수 있다. 일 실시 예에 따르면, 무선 통신 모듈(2092)은 셀룰러 통신 모듈, 근거리 무선 통신 모듈, 또는 GNSS 통신 모듈 중 해당하는 통신 방식만을 지원하는 복수의 통신 모듈들로 구성될 수 있다.

- [0132] 유선 통신 모듈(2094)은, 예를 들면, LAN(local area network), 전력선 통신 또는 POTS(plain old telephone service)를 포함할 수 있다.
- [0133] 제1 네트워크(2098)는, 예를 들어, 전자 장치(2001)와 제1 외부 전자 장치(2002)간의 무선으로 직접 연결을 통해 명령 또는 데이터를 송신 또는 수신 할 수 있는 Wi-Fi 다이렉트 또는 Bluetooth를 포함할 수 있다. 제2 네트워크(2099)는, 예를 들어, 전자 장치(2001)와 제2 외부 전자 장치(2004)간의 명령 또는 데이터를 송신 또는 수신할 수 있는 텔레커뮤니케이션 네트워크(예: LAN(local area network)나 WAN(wide area network)와 같은 컴퓨터 네트워크, 인터넷(internet), 또는 텔레폰(telephone) 네트워크)를 포함할 수 있다.
- [0134] 다양한 실시 예들에 따르면, 상기 명령 또는 상기 데이터는 제2 네트워크에 연결된 서버(2008)를 통해서 전자 장치(2001)와 제2 외부 전자 장치(2004)간에 송신 또는 수신될 수 있다. 제1 및 제2 외부 전자 장치(2002, 2004) 각각은 전자 장치(2001)와 동일한 또는 다른 종류의 장치일 수 있다. 다양한 실시 예들에 따르면, 전자 장치(2001)에서 실행되는 동작들의 전부 또는 일부는 다른 하나 또는 복수의 전자 장치(예: 전자 장치(2002, 2004), 또는 서버(2008)에서 실행될 수 있다. 일 실시 예에 따르면, 전자 장치(2001)가 어떤 기능이나 서비스를 자동으로 또는 요청에 의하여 수행해야 할 경우에, 전자 장치(2001)는 기능 또는 서비스를 자체적으로 실행시키는 대신에 또는 추가적으로, 그와 연관된 적어도 일부 기능을 다른 장치(예: 전자 장치(2002, 2004), 또는 서버(2008))에게 요청할 수 있다. 다른 전자 장치(예: 전자 장치(2002, 2004), 또는 서버(2008))는 요청된 기능 또는 추가 기능을 실행하고, 그 결과를 전자 장치(2001)로 전달할 수 있다. 전자 장치(2001)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.
- [0135] 다양한 실시예에 따른 전자 장치(예: 도 1의 전자 장치(140))는, 디스플레이, 마이크, 메모리, 통신 회로 및 프로세서를 포함하고, 상기 프로세서는 상기 디스플레이를 통해, 동일한 시작 발화에 대응하여 음성 인식을 시작할 수 있는 복수의 음성 인식 장치들 각각의 음성 인식 감도를 조절하기 위한 사용자 인터페이스를 표시하고, 상기 사용자 인터페이스를 통해 상기 음성 인식 감도가 변경된 경우, 상기 변경된 감도 값을 상기 복수의 음성 인식 장치들 중 적어도 일부에 전송하고, 상기 복수의 음성 인식 장치들에 사용자의 테스트 발화를 수신할 수 있도록 대기하도록 하는 신호를 송신하고, 상기 복수의 음성 인식 장치들로부터 상기 제1 발화에 대응하는 발화 정보를 수신하고, 상기 발화 정보를 기반으로 상기 사용자 인터페이스를 업데이트하도록 설정될 수 있다. 상기 프로세서의 동작은 상기 메모리에 저장되는 인스트럭션의 실행에 의해 수행될 수도 있다.
- [0136] 다양한 실시예에 따르면, 상기 사용자 인터페이스는 상기 복수의 음성 인식 장치들 각각의 식별 정보, 및 상기 음성 인식 감도를 변경하는 제1 객체를 포함할 수 있다. 상기 프로세서는 이동 바(bar) 또는 수치 입력 필드를 통해 상기 제1 객체를 표시할 수 있다. 상기 사용자 인터페이스는 상기 음성 인식 감도의 기준 값, 상기 복수의 음성 인식 장치들의 상기 음성 인식 감도의 설정을 제외하는 제2 객체, 상기 음성 인식 감도를 초기화하는 제3 객체 중 적어도 하나를 포함할 수 있다.
- [0137] 일 실시예에 따르면, 상기 발화 정보는 상기 복수의 음성 인식 장치들에서 수신된 상기 제1 발화에 대한 오디오 품질(audio quality), 음압(sound pressure), 테스트 발화의 특징값 중 적어도 하나를 기반으로 보정된 값일 수 있다. 다른 일 실시예에 따르면, 상기 발화 정보는 상기 복수의 음성 인식 장치들의 동작 상태, 상기 복수의 음성 인식 장치들 주변의 소음 정보 중 적어도 하나를 기반으로 보정된 값일 수 있다.
- [0138] 다양한 실시예에 따르면, 상기 프로세서는 상기 음성 인식 감도를 조절하는 제 1 객체와 상기 제 1객체와 다른 제 2 객체를 오버랩하여 표시할 수 있다. 상기 프로세서는 상기 제1 객체와 상기 제2 객체를 서로 다른 색으로 표시할 수 있다.
- [0139] 다양한 실시예에 따르면, 상기 프로세서는 상기 발화정보를 기반으로 상기 복수의 음성 인식 장치들 중 상기 시작 발화에 응답하는 장치를 결정할 수 있다.
- [0140] 다양한 실시예에 따르면, 상기 프로세서는 상기 음성 인식 감도가 변경되는 경우, 상기 변경된 감도 값을 근거리 통신을 위한 네트워크 장치 또는 외부 서버를 통해, 상기 복수의 음성 인식 장치들에 전송할 수 있다. 상기 전자 장치는 상기 네트워크 장치를 통해, 상기 복수의 음성 인식 장치들과 연결될 수 있다. 상기 프로세서는 상기 네트워크 장치를 통해 상기 복수의 음성 인식 장치들로부터 상기 발화 정보를 수신할 수 있다.
- [0141] 다양한 실시예에 따르면, 상기 시작 발화는 상기 제1 발화와 서로 다르게 설정될 수 있다.
- [0142] 다양한 실시예에 따른 전자 장치(예: 도 1의 전자 장치(110))는 디스플레이, 마이크, 메모리, 통신 회로 및 프



로세서를 포함하고, 상기 프로세서는 상기 전자 장치의 동작 상태 정보 또는 상기 전자 장치 주변의 소음 정보 중 적어도 하나를 수집하고, 사용자의 제1 발화를 수신하고, 상기 동작 상태 정보 또는 상기 소음 정보 중 적어도 하나를 기반으로 상기 수신한 제1 발화를 보정하여 발화 정보를 계산하고, 상기 발화 정보를 외부 장치에 전송할 수 있다.

- [0143] 다양한 실시예에 따르면, 상기 동작 상태 정보는 상기 전자 장치의 내부 구성에 의해 발생하는 소리에 관한 정보를 포함할 수 있다.
- [0144] 다양한 실시예에 따르면, 상기 프로세서는 상기 발화 정보를 근거리 통신을 위한 네트워크 장치 또는 외부 서버를 통해, 상기 외부 장치에 전송할 수 있다. 상기 전자 장치는 상기 네트워크 장치를 통해, 상기 외부 장치와 통신 채널을 형성할 수 있다.
- [0145] 다양한 실시예에 따르면, 상기 프로세서는 상기 수신된 제1 발화의 오디오 음질(audio quality), 음압(sound pressure), 제1 발화의 특징값 중 적어도 하나를 기반으로 발화 정보를 계산할 수 있다.
- [0146] 다양한 실시예에 따르면, 상기 발화 정보는 상기 수신된 제1 발화의 SNR 값을 포함할 수 있다.
- [0147] 다양한 실시예에 따른 음성 인식 장치의 제어 방법은 전자 장치에서 수행되고, 상기 전자 장치의 디스플레이를 통해, 동일한 시작 발화에 대응하여 음성 인식 과정을 시작할 수 있는 복수의 음성 인식 장치들 각각의 음성 인식 감도를 조절하기 위한 사용자 인터페이스를 표시하는 동작, 상기 사용자 인터페이스를 통해 상기 음성 인식 감도가 변경된 경우, 변경된 감도 값을 복수의 음성 인식 장치들 중 적어도 일부에 전송하는 동작, 상기 복수의 음성 인식 장치들에 사용자의 테스트 발화를 수신할 수 있도록 대기하도록 하는 신호를 송신하는 동작, 상기 복수의 음성 인식 장치들로부터 상기 테스트 발화에 대응하는 발화 정보를 수신하는 동작 및 상기 발화 정보를 기반으로 상기 사용자 인터페이스를 업데이트하는 동작을 포함할 수 있다.
- [0148] 본 문서의 다양한 실시 예들 및 이에 사용된 용어들은 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 해당 실시 예의 다양한 변경, 균등물, 및/또는 대체물을 포함하는 것으로 이해되어야 한다. 도면의 설명과 관련하여, 유사한 구성요소에 대해서는 유사한 참조 부호가 사용될 수 있다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함할 수 있다. 본 문서에서, "A 또는 B", "A 및/또는 B 중 적어도 하나", "A, B 또는 C" 또는 "A, B 및/또는 C 중 적어도 하나" 등의 표현은 함께 나열된 항목들의 모든 가능한 조합을 포함할 수 있다. "제1," "제2," "첫째," 또는 "둘째," 등의 표현들은 해당 구성요소들을, 순서 또는 중요도에 상관없이 수식할 수 있고, 한 구성요소를 다른 구성요소와 구분하기 위해 사용될 뿐 해당 구성요소들을 한정하지 않는다. 어떤(예: 제1) 구성요소가 다른(예: 제2) 구성요소에 "(기능적으로 또는 통신적으로) 연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 상기 어떤 구성요소가 상기 다른 구성요소에 직접적으로 연결되거나, 다른 구성요소(예: 제 3 구성요소)를 통하여 연결될 수 있다.
- [0149] 본 문서에서, "~하도록 설정된(adapted to or configured to)"은 상황에 따라, 예를 들면, 하드웨어적 또는 소프트웨어적으로 "~에 적합한," "~하는 능력을 가지는," "~하도록 변경된," "~하도록 만들어진," "~를 할 수 있는," 또는 "~하도록 설계된"과 상호 호환적으로(interchangeably) 사용될 수 있다. 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 설정된 (또는 구성된) 프로세서"는 해당 동작들을 수행하기 위한 전용 프로세서(예: 임베디드 프로세서), 또는 메모리 장치(예: 메모리 2030)에 저장된 하나 이상의 프로그램들을 실행함으로써, 해당 동작들을 수행할 수 있는 범용 프로세서(예: CPU 또는 AP)를 의미할 수 있다.
- [0150] 본 문서에서 사용된 용어 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어(firmware)로 구성된 유닛(unit)을 포함하며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. "모듈"은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수 있다. "모듈"은 기계적으로 또는 전자적으로 구현될 수 있으며, 예를 들면, 어떤 동작들을 수행하는, 알려졌거나 앞으로 개발될, ASIC(application-specific integrated circuit) 칩, FPGAs(field-programmable gate arrays), 또는 프로그램 가능 논리 장치를 포함할 수 있다.
- [0151] 다양한 실시 예들에 따른 장치(예: 모듈들 또는 그 기능들) 또는 방법(예: 동작들)의 적어도 일부는 프로그램 모듈의 형태로 컴퓨터로 판독 가능한 저장 매체(예: 메모리(2030))에 저장된 명령어로 구현될 수 있다. 상기 명령어가 프로세서(예: 프로세서(2020))에 의해 실행될 경우, 프로세서가 상기 명령어에 해당하는 기능을 수행할 수 있다. 컴퓨터로 판독 가능한 기록 매체는, 하드디스크, 플로피디스크, 마그네틱 매체(예: 자기테이프), 광기록 매체(예: CD-ROM, DVD, 자기-광 매체(예: 플롭티컬 디스크), 내장 메모리 등을 포함할 수 있다. 명령어는 컴

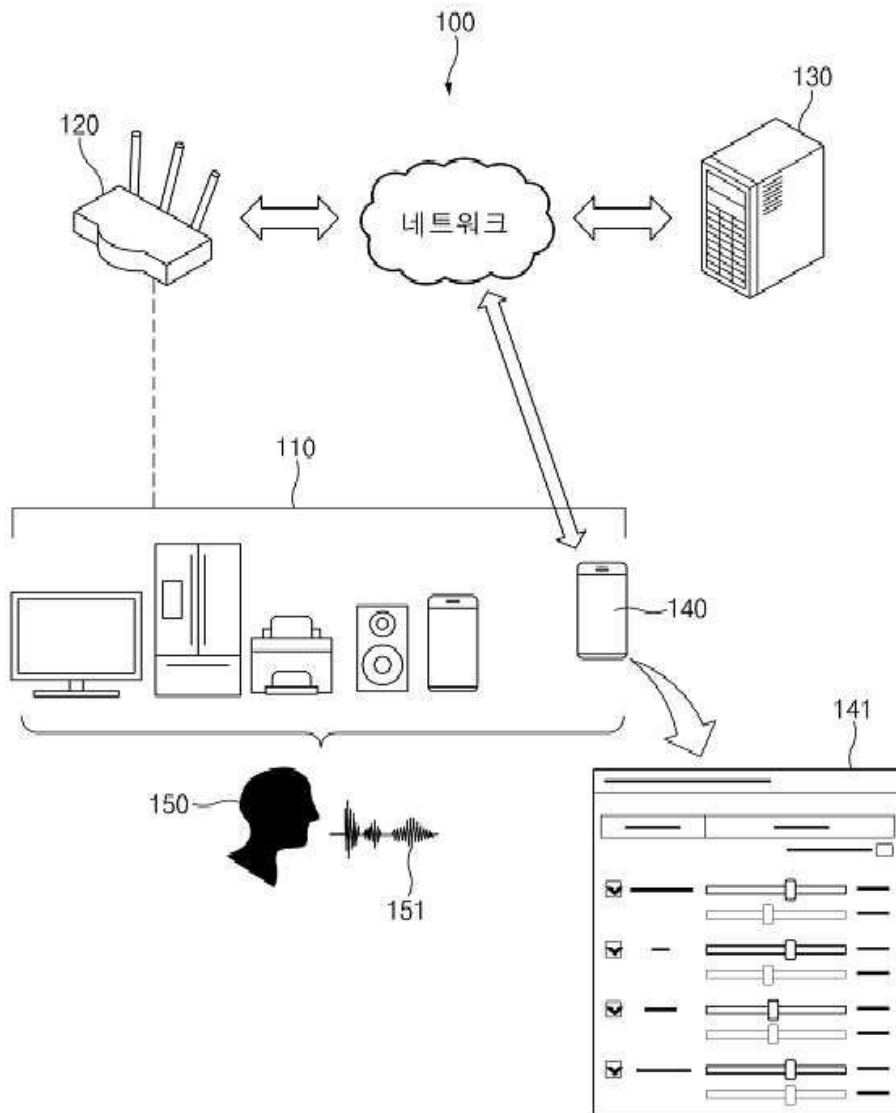
파일러에 의해 만들어지는 코드 또는 인터프리터에 의해 실행될 수 있는 코드를 포함할 수 있다.

[0152]

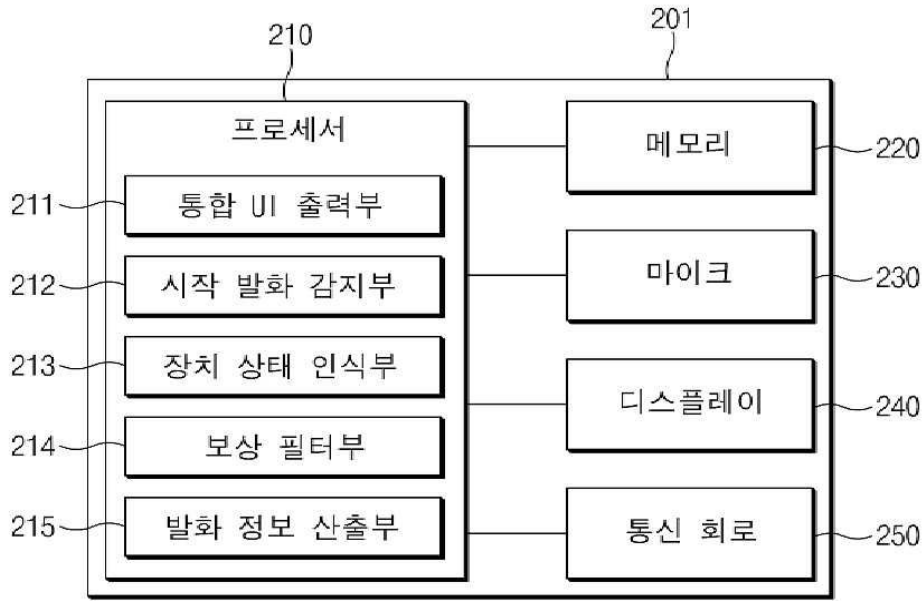
다양한 실시 예들에 따른 구성 요소(예: 모듈 또는 프로그램 모듈) 각각은 단수 또는 복수의 개체로 구성될 수 있으며, 전술한 해당 서브 구성 요소들 중 일부 서브 구성 요소가 생략되거나, 또는 다른 서브 구성 요소를 더 포함할 수 있다. 대체적으로 또는 추가적으로, 일부 구성 요소들(예: 모듈 또는 프로그램 모듈)은 하나의 개체로 통합되어, 통합되기 이전의 각각의 해당 구성 요소에 의해 수행되는 기능을 동일 또는 유사하게 수행할 수 있다. 다양한 실시 예들에 따른 모듈, 프로그램 모듈 또는 다른 구성 요소에 의해 수행되는 동작들은 순차적, 병렬적, 반복적 또는 휴리스틱(heuristic)하게 실행되거나, 적어도 일부 동작이 다른 순서로 실행되거나, 생략되거나, 또는 다른 동작이 추가될 수 있다.

**도면**

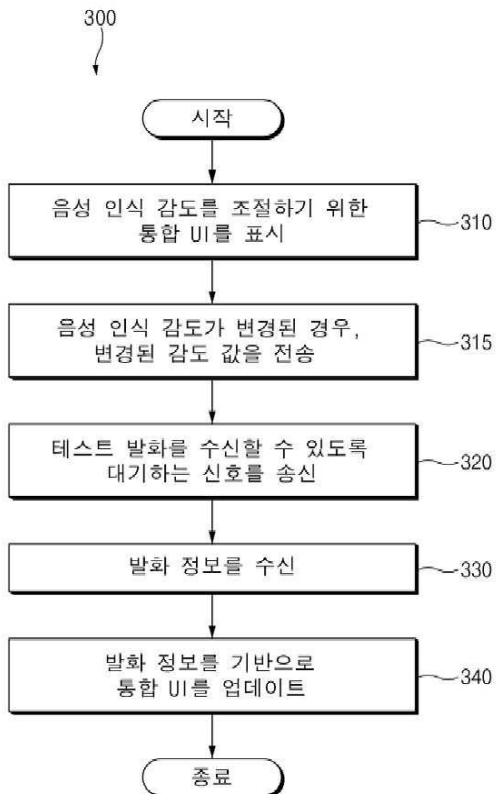
**도면1**



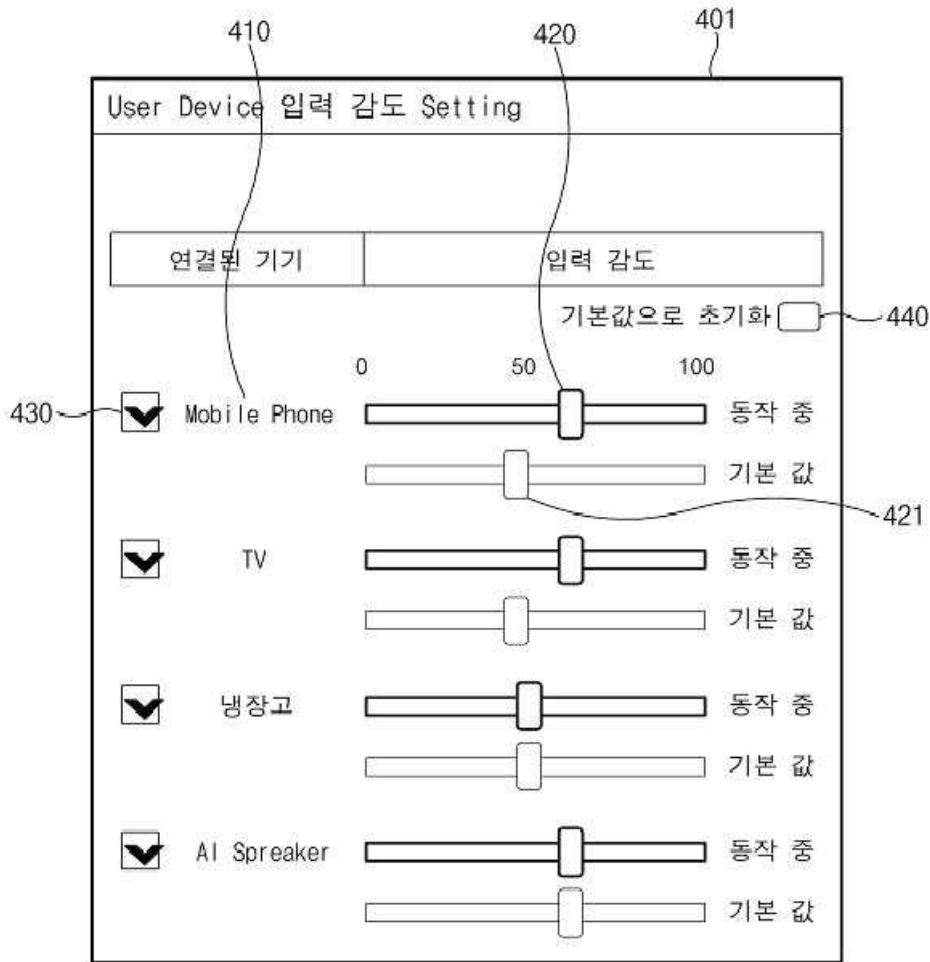
도면2



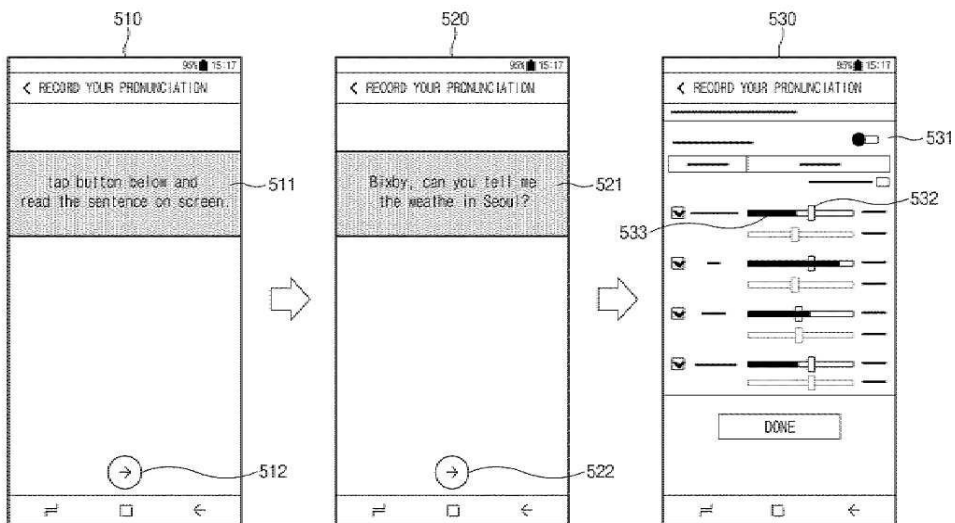
도면3



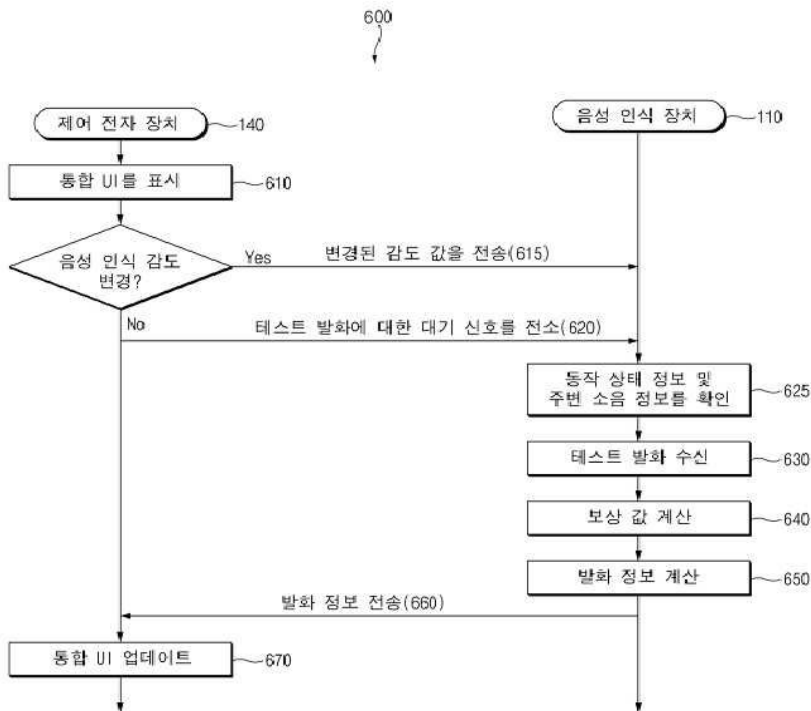
도면4



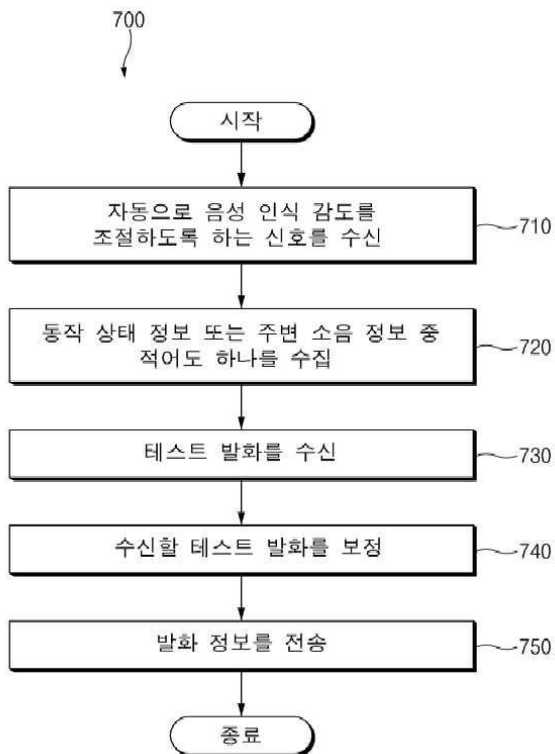
도면5



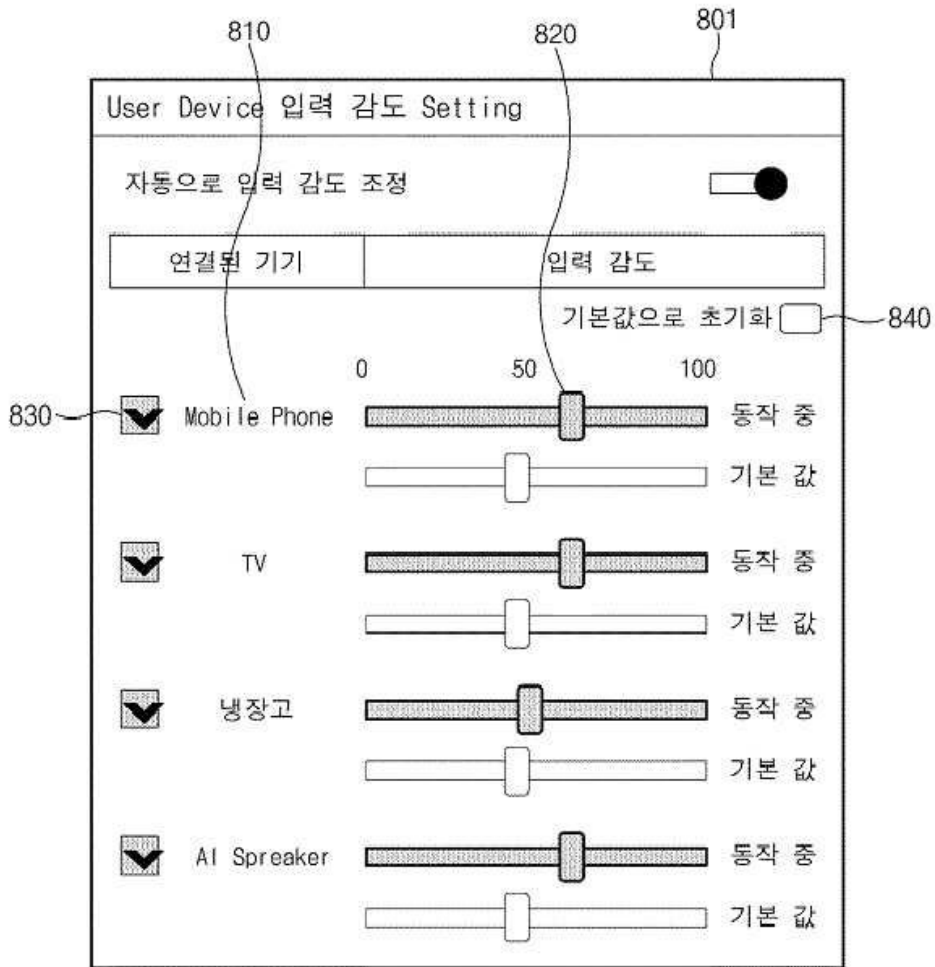
도면6



도면7



도면8



도면9

901

ID	Device	SNR	
		1st Trial (initial)	2nd Trial (SNR보정값)
A	Mobile Phone	30	30
B	TV(TV 시청중)	40	(55)
C	냉장고	50	50
D	AI Speaker (음악 Play중)	40	(45)

도면10

