



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2020-0074690
(43) 공개일자 2020년06월25일

- (51) 국제특허분류(Int. Cl.)
G10L 15/28 (2006.01) G06F 3/16 (2018.01)
G10L 15/22 (2006.01)
- (52) CPC특허분류
G10L 15/28 (2013.01)
G06F 3/011 (2013.01)
- (21) 출원번호 10-2018-0163377
- (22) 출원일자 2018년12월17일
심사청구일자 없음

- (71) 출원인
삼성전자주식회사
경기도 수원시 영통구 삼성로 129 (매탄동)
- (72) 발명자
안상원
경기도 수원시 영통구 삼성로 129(매탄동)
함성일
경기도 수원시 영통구 삼성로 129(매탄동)
(뒷면에 계속)
- (74) 대리인
정홍식, 김태현

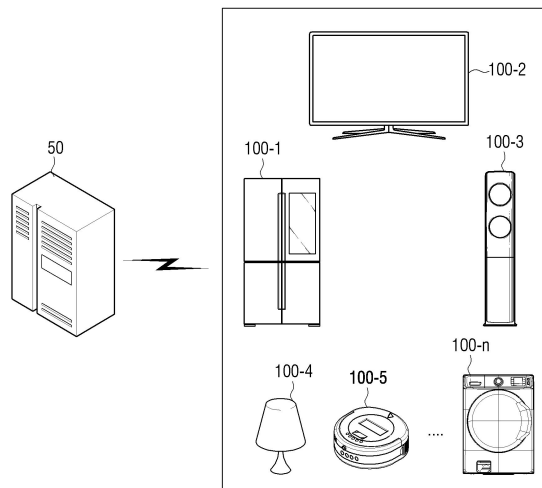
전체 청구항 수 : 총 20 항

(54) 발명의 명칭 전자 장치 및 이의 제어 방법

(57) 요약

전자 장치 및 이의 제어 방법이 제공된다. 본 전자 장치는 마이크, 통신 인터페이스, 적어도 하나의 명령을 저장하는 메모리 및 적어도 하나의 명령을 실행하는 프로세서를 포함하고, 프로세서는 마이크를 통해 수신된 사용자의 음성을 바탕으로 전자 장치 주위에 사용자가 존재하는지 여부를 판단하고, 사용자가 전자 장치 주위에 존재하면, 전자 장치와 전자 장치 주변에 존재하는 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하며, 적어도 하나의 타 전자 장치 각각으로부터 수신한 타 전자 장치에 대한 정보를 바탕으로 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정하고, 전자 장치가 허브 장치로 결정되면, 통신 인터페이스를 통해 상기 적어도 하나의 타 전자 장치 각각으로부터 사용자의 음성 데이터를 수신하여 음성 인식을 수행할 수 있다.

대표도 - 도1



(52) CPC특허분류

G06F 3/167 (2013.01)

G10L 15/22 (2013.01)

G10L 2015/225 (2013.01)

(72) 발명자

김정인

경기도 수원시 영통구 삼성로 129(매탄동)

변성호

경기도 수원시 영통구 삼성로 129(매탄동)

신재식

경기도 수원시 영통구 삼성로 129(매탄동)

정준식

경기도 수원시 영통구 삼성로 129(매탄동)

명세서

청구범위

청구항 1

전자 장치에 있어서,

마이크;

통신 인터페이스;

적어도 하나의 명령을 저장하는 메모리; 및

상기 적어도 하나의 명령을 실행하는 프로세서;를 포함하고,

상기 프로세서는,

상기 마이크를 통해 수신된 사용자의 음성을 바탕으로 상기 전자 장치 주위에 사용자가 존재하는지 여부를 판단하고,

상기 사용자가 상기 전자 장치 주위에 존재하면, 상기 전자 장치와 상기 전자 장치 주변에 존재하는 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하며,

상기 적어도 하나의 타 전자 장치 각각으로부터 수신한 상기 타 전자 장치에 대한 정보를 바탕으로 상기 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정하고,

상기 전자 장치가 허브 장치로 결정되면, 상기 통신 인터페이스를 통해 상기 적어도 하나의 타 전자 장치 각각으로부터 상기 사용자의 음성 데이터를 수신하여 음성 인식을 수행하는 전자 장치.

청구항 2

제1항에 있어서,

상기 프로세서는,

상기 사용자의 음성, 상기 전자 장치의 사용 정보, 상기 전자 장치의 사용 통계 정보, 상기 사용자의 움직임 정보 중 적어도 하나를 이용하여 상기 전자 장치 주위에 사용자가 존재하는지 여부를 판단하는 전자 장치.

청구항 3

제1항에 있어서,

상기 프로세서는,

상기 사용자가 상기 전자 장치 주위에 존재하면, 상기 통신 인터페이스를 통해 복수의 타 전자 장치로부터 상기 사용자의 음성에 대한 정보를 수신하고,

상기 복수의 타 전자 장치 중 임계값 이상의 크기를 가지는 상기 사용자의 음성을 감지한 상기 적어도 하나의 타 전자 장치를 결정하고,

상기 결정된 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하는 전자 장치.

청구항 4

제1항에 있어서,

상기 프로세서는,

상기 사용자가 상기 전자 장치 주위에 존재하면, 상기 전자 장치의 위치를 바탕으로 상기 적어도 하나의 타 전자 장치를 결정하고,

상기 결정된 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하는 전자 장치.

청구항 5

제1항에 있어서,

상기 타 전자 장치에 대한 정보는,

상기 타 전자 장치의 인터넷 연결 상태에 대한 정보, 상기 타 전자 장치의 전원 상태에 대한 정보, 상기 타 전자 장치의 메모리에 대한 정보, 상기 타 전자 장치와 사용자와의 거리에 대한 정보, 상기 타 전자 장치의 음성 인식 기능에 대한 정보 중 적어도 하나를 포함하는 전자 장치.

청구항 6

제1항에 있어서,

상기 프로세서는,

상기 적어도 하나의 타 전자 장치 각각으로부터 수신된 상기 사용자의 음성 데이터를 외부 서버에 전송하도록 상기 통신 인터페이스를 제어하고,

상기 통신 인터페이스를 통해 상기 외부 서버로부터 상기 사용자의 음성 데이터에 대한 응답 및 상기 사용자의 음성 데이터에 대응되는 제어 명령 중 적어도 하나를 수신하는 전자 장치.

청구항 7

제6항에 있어서,

상기 프로세서는,

상기 사용자의 위치를 바탕으로 상기 디바이스 그룹에 포함된 장치들 중 상기 사용자의 음성 데이터에 대한 응답을 출력할 장치를 결정하고,

상기 결정된 장치로 상기 사용자의 음성 데이터에 대한 응답을 전송하는 전자 장치.

청구항 8

제1항에 있어서,

상기 프로세서는,

상기 적어도 하나의 타 전자 장치 각각으로부터 수신된 상기 사용자의 음성 데이터를 바탕으로 상기 사용자의 음성 데이터에 대한 응답 및 상기 사용자의 음성 데이터에 대응되는 제어 명령 중 적어도 하나를 결정하는 전자 장치.

청구항 9

제1항에 있어서,

상기 프로세서는,

상기 적어도 하나의 타 전자 장치 중 하나가 허브 장치로 결정되면, 허브 장치로 결정된 타 전자 장치로 상기 마이크를 통해 수신된 사용자의 음성 데이터를 전송하도록 상기 통신 인터페이스를 제어하는 전자 장치.

청구항 10

제1항에 있어서,

상기 프로세서는,

상기 전자 장치 주위에 사용자가 감지되지 않으면, 상기 음성 인식을 종료하는 전자 장치.

청구항 11

전자 장치의 제어 방법에 있어서,

마이크를 통해 수신된 사용자의 음성을 바탕으로 상기 전자 장치 주위에 사용자가 존재하는지 여부를 판단하는

단계;

상기 사용자가 상기 전자 장치 주위에 존재하면, 상기 전자 장치와 상기 전자 장치 주변에 존재하는 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하는 단계;

상기 적어도 하나의 타 전자 장치 각각으로부터 수신한 상기 타 전자 장치에 대한 정보를 바탕으로 상기 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정하는 단계;

상기 전자 장치가 허브 장치로 결정되면, 상기 적어도 하나의 타 전자 장치 각각으로부터 상기 사용자의 음성 데이터를 수신하여 음성 인식을 수행하는 단계;를 포함하는 제어 방법.

청구항 12

제11항에 있어서,

상기 판단하는 단계는,

상기 사용자의 음성, 상기 전자 장치의 사용 정보, 상기 전자 장치의 사용 통계 정보, 상기 사용자의 움직임 정보 중 적어도 하나를 이용하여 상기 전자 장치 주위에 사용자가 존재하는지 여부를 판단하는 제어 방법.

청구항 13

제11항에 있어서,

상기 디바이스 그룹을 결정하는 단계는,

상기 사용자가 상기 전자 장치 주위에 존재하면, 복수의 타 전자 장치로부터 상기 사용자의 음성에 대한 정보를 수신하는 단계;

상기 복수의 타 전자 장치 중 임계값 이상의 크기를 가지는 상기 사용자의 음성을 감지한 상기 적어도 하나의 타 전자 장치를 결정하는 단계; 및

상기 결정된 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 생성하는 단계;를 포함하는 제어 방법.

청구항 14

제11항에 있어서,

상기 디바이스 그룹을 결정하는 단계는,

상기 사용자가 상기 전자 장치 주위에 존재하면, 상기 전자 장치의 위치를 바탕으로 상기 적어도 하나의 타 전자 장치를 결정하는 단계; 및

상기 결정된 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하는 단계;를 포함하는 제어 방법.

청구항 15

제11항에 있어서,

상기 타 전자 장치에 대한 정보는,

상기 타 전자 장치의 인터넷 연결 상태에 대한 정보, 상기 타 전자 장치의 전원 상태에 대한 정보, 상기 타 전자 장치의 메모리에 대한 정보, 상기 타 전자 장치와 사용자와의 거리에 대한 정보, 상기 타 전자 장치의 음성 인식 기능에 대한 정보 중 적어도 하나를 포함하는 제어 방법.

청구항 16

제11항에 있어서,

상기 수행하는 단계는,

상기 적어도 하나의 타 전자 장치 각각으로부터 수신된 상기 사용자의 음성 데이터를 외부 서버에 전송하는 단계; 및

상기 외부 서버로부터 상기 사용자의 음성 데이터에 대한 응답 및 상기 사용자의 음성 데이터에 대응되는 제어

명령 중 적어도 하나를 수신하는 단계;를 포함하는 제어 방법.

청구항 17

제16항에 있어서,

상기 사용자의 위치를 바탕으로 상기 디바이스 그룹에 포함된 장치들 중 상기 사용자의 음성 데이터에 대한 응답을 출력할 장치를 결정하는 단계; 및

상기 결정된 장치로 상기 사용자의 음성 데이터에 대한 응답을 전송하는 단계;를 포함하는 제어 방법.

청구항 18

제11항에 있어서,

상기 적어도 하나의 타 전자 장치 각각으로부터 수신된 상기 사용자의 음성 데이터를 바탕으로 상기 사용자의 음성 데이터에 대한 응답 및 상기 사용자의 음성 데이터에 대응되는 제어 명령 중 적어도 하나를 결정하는 단계;를 포함하는 제어 방법.

청구항 19

제11항에 있어서,

상기 적어도 하나의 타 전자 장치 중 하나가 허브 장치로 결정되면, 허브 장치로 결정된 타 전자 장치로 상기 마이크를 통해 수신된 사용자의 음성 데이터를 전송하는 단계를 포함하는 제어 방법.

청구항 20

제11항에 있어서,

상기 전자 장치 주위에 사용자가 감지되지 않으면, 상기 음성 인식을 종료하는 단계;를 포함하는 제어 방법.

발명의 설명

기술 분야

[0001] 본 개시는 전자 장치 및 이의 제어 방법에 관한 것으로, 더욱 상세하게는 텍 내에 존재하는 복수의 전자 장치 중 사용자 음성을 인식하는 전자 장치를 결정하여 대화형 서비스를 제공할 수 있는 전자 장치 및 이의 제어 방법에 관한 것이다.

배경 기술

[0002] 근래에는 인공 지능 시스템이 다양한 분야에서 이용되고 있다. 인공 지능 시스템은 기존의 룰(rule) 기반 스마트 시스템과 달리 기계가 스스로 학습하고 판단하며 똑똑해지는 시스템이다. 인공 지능 시스템은 사용할수록 인식이 향상되고 사용자 취향을 보다 정확하게 이해할 수 있게 되어, 기존 룰 기반 스마트 시스템은 점차 딥러닝 기반 인공 지능 시스템으로 대체되고 있다.

[0003] 인공 지능 기술은 기계학습(예로, 딥러닝) 및 기계학습을 활용한 요소 기술들로 구성된다.

[0004] 기계학습은 입력 데이터들의 특징을 스스로 분류/학습하는 알고리즘 기술이며, 요소기술은 딥러닝 등의 기계학습 알고리즘을 활용하여 인간 두뇌의 인지, 판단 등의 기능을 모사하는 기술로서, 언어적 이해, 시각적 이해, 추론/예측, 지식 표현, 동작 제어 등의 기술 분야로 구성된다.

[0005] 인공 지능 기술이 응용되는 다양한 분야는 다음과 같다. 언어적 이해는 인간의 언어/문자를 인식하고 응용/처리하는 기술로서, 자연어 처리, 기계 번역, 대화시스템, 질의 응답, 음성 인식/합성 등을 포함한다. 시각적 이해는 사물을 인간의 시각처럼 인식하여 처리하는 기술로서, 오브젝트 인식, 오브젝트 추적, 영상 검색, 사람 인식, 장면 이해, 공간 이해, 영상 개선 등을 포함한다. 추론 예측은 정보를 판단하여 논리적으로 추론하고 예측하는 기술로서, 지식/확률 기반 추론, 최적화 예측, 선호 기반 계획, 추천 등을 포함한다. 지식 표현은 인간의 경험정보를 지식데이터로 자동화 처리하는 기술로서, 지식 구축(데이터 생성/분류), 지식 관리(데이터 활용) 등을 포함한다. 동작 제어는 차량의 자율 주행, 로봇의 움직임 제어를 하는 기술로서, 움직임 제어(항법, 충돌,

주행), 조작 제어(행동 제어) 등을 포함한다.

[0006] 한편, 근래에는택내에 존재하는 다양한 전자 장치를 이용하여 사용자 음성을 인식하고, 사용자 음성에 대한 응답 또는 제어 명령을 제공하는 대화형 서비스가 제공되고 있다.

[0007] 다만, 종래에는 사용자 음성을 받은 모든 장치들이 기정의된 허브 장치로 사용자의 음성 데이터를 전송하고, 허브 장치가 사용자의 음성 데이터를 바탕으로 대화형 서비스를 제공하였다. 이와 같이, 정해진 허브 장치가 대화형 서비스를 제공할 경우, 사용자와 먼 곳에 존재하는 음성 데이터까지 수신하여 불필요한 잡음이 많이 포함된 음성 데이터를 이용하여 음성 인식을 수행할 수 있으며, 음성 데이터의 유실, 성능 저하 및 보안 약화와 같은 다양한 문제를 발생시킬 수 있게 된다. 뿐만 아니라, 하나의 허브 장치에 데이터가 집중되므로, 네트워크 혼잡도가 증가할 뿐만 아니라 네트워크 병목 현상이 발생하는 문제가 발생할 수 있게 된다.

발명의 내용

해결하려는 과제

[0008] 본 개시는 상술한 문제점을 해결하기 위해 안출된 것으로, 본 개시의 목적은 사용자의 위치에 따라 적응적으로 허브 장치를 결정하여 대화형 서비스를 제공할 수 있는 전자 장치 및 이의 제어 방법을 제공함에 있다.

과제의 해결 수단

[0009] 본 개시의 일 실시예에 따른, 전자 장치는, 마이크; 통신 인터페이스; 적어도 하나의 명령을 저장하는 메모리; 및 상기 적어도 하나의 명령을 실행하는 프로세서;를 포함하고, 상기 프로세서는, 상기 마이크를 통해 수신된 사용자의 음성을 바탕으로 상기 전자 장치 주위에 사용자가 존재하는지 여부를 판단하고, 상기 사용자가 상기 전자 장치 주위에 존재하면, 상기 전자 장치와 상기 전자 장치 주변에 존재하는 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하며, 상기 적어도 하나의 타 전자 장치 각각으로부터 수신한 상기 타 전자 장치에 대한 정보를 바탕으로 상기 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정하고, 상기 전자 장치가 허브 장치로 결정되면, 상기 통신 인터페이스를 통해 상기 적어도 하나의 타 전자 장치 각각으로부터 상기 사용자의 음성 데이터를 수신하여 음성 인식을 수행할 수 있다.

[0010] 본 개시의 일 실시예에 따른, 전자 장치의 제어 방법은, 마이크를 통해 수신된 사용자의 음성을 바탕으로 상기 전자 장치 주위에 사용자가 존재하는지 여부를 판단하는 단계; 상기 사용자가 상기 전자 장치 주위에 존재하면, 상기 전자 장치와 상기 전자 장치 주변에 존재하는 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하는 단계; 상기 적어도 하나의 타 전자 장치 각각으로부터 수신한 상기 타 전자 장치에 대한 정보를 바탕으로 상기 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정하는 단계; 및 상기 전자 장치가 허브 장치로 결정되면, 상기 적어도 하나의 타 전자 장치 각각으로부터 상기 사용자의 음성 데이터를 수신하여 음성 인식을 수행하는 단계;를 포함한다.

발명의 효과

[0011] 상술한 바와 같이, 사용자의 위치에 따라 적응적으로 허브 장치를 결정함으로써, 불필요한 네트워크 오버헤드를 제거할 수 있으며, 더욱 정확한 대화형 서비스를 제공할 수 있게 된다.

도면의 간단한 설명

[0012] 도 1은 본 개시의 일 실시예에 따른, 택 내의 복수의 전자 장치를 이용하여 사용자 음성에 대한 응답을 제공하는 대화 시스템을 설명하기 위한 도면,

도 2 및 도 3은 본 개시의 일 실시예에 따른, 전자 장치의 구성을 나타내는 블록도를 도시한 도면,

도 4 및 도 5는 본 개시의 다양한 실시예에 따른, 복수의 전자 장치 중 허브 장치를 결정하여 대화형 서비스를 제공하기 위한 방법을 설명하기 위한 시퀀스도들,

도 6a 내지 도 6d는 본 개시의 일 실시예에 따른, 복수의 전자 장치 중 허브 장치를 결정하여 대화형 서비스를 제공하는 실시예를 설명하기 위한 도면들,

도 7은 본 개시의 일 실시예에 따른, 전자 장치의 제어 방법을 설명하기 위한 흐름도,

도 8은 본 개시의 일 실시예에 따른, 인공지능 에이전트 시스템의 대화 시스템을 도시한 블록도이다.

발명을 실시하기 위한 구체적인 내용

- [0013] 이하, 본 문서의 다양한 실시 예가 첨부된 도면을 참조하여 기재된다. 그러나, 이는 본 문서에 기재된 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 문서의 실시 예의 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.
- [0014] 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.
- [0015] 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.
- [0016] 본 문서에서 사용된 "제1," "제2," "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.
- [0017] 어떤 구성요소(예: 제1 구성요소)가 다른 구성요소(예: 제2 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: 제3 구성요소)를 통하여 연결될 수 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소(예: 제1 구성요소)가 다른 구성요소(예: 제2 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.
- [0018] 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) 부프로세서"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.
- [0020] 본 문서에서, 사용자라는 용어는 전자 장치를 사용하는 사람 또는 전자 장치를 사용하는 장치(예: 인공지능 전자 장치)를 지칭할 수 있다.
- [0021] 이하에서는 도면을 참조하여 본 발명에 대해 상세히 설명하기로 한다. 도 1은 본 개시의 일 실시예에 따른, 맥 내의 복수의 전자 장치를 이용하여 사용자 음성에 대한 응답을 제공하는 대화 시스템을 설명하기 위한 도면이다. 대화 시스템은 복수의 전자 장치(100-1 내지 100-n) 및 외부 서버(50)를 포함할 수 있다.
- [0022] 복수의 전자 장치(100-1 내지 100-n)는 맥 내에 위치하는 전자 장치로서, 마이크를 사용하여 사용자 음성을 입력받거나 마이크가 구비된 외부 장치와 연결되어 사용자 음성 명령과 관련된 정보를 수신할 수 있는 전자 장치일 수 있다. 일 예로, 전자 장치는 텔레비전, DVD(digital video disk) 플레이어, 오디오, 냉장고, 에어컨, 청소기, 오븐, 전자레인지, 세탁기, 공기 청정기, 셋톱 박스, 홈 오토메이션 컨트롤 패널, 보안 컨트롤 패널, 미디어 박스(예: 삼성 HomeSync[™], 애플TV[™], 또는 구글 TV[™]), 게임 콘솔(예: Xbox[™], PlayStation[™]), 전자 사전, 전자 키, 캠코더, 또는 전자 액자 중 적어도 하나를 포함할 수 있다. 또 다른 실시예에 의하면, 전자 장치는, 스마트폰, 태블릿 PC, 이동 전화기, 영상 전화기, 전자책 리더기, 데스크탑 PC, 랩탑 PC, 넷북 컴퓨터, 워크스태이션, 서버, PDA, PMP(portable multimedia player), MP3 플레이어, 의료기기, 카메라, 또는 웨어러블 장치 중 적어도 하나를 포함할 수 있다. 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 또는 문신), 또는 생체 이식형 회로 중 적어도 하나를 포함할 수 있다.

- [0023] 복수의 전자 장치(100-1 내지 100-n) 각각은 마이크를 통해 사용자 음성을 입력받을 수 있다. 이때, 사용자 음성은 대화형 서비스를 수행하기 위한 트리거 음성 또는 웨이크업 음성으로서, 예로, "빅스비" 등과 같은 단어일 수 있다. 이때, 대화형 서비스는 인공지능 프로그램을 이용하여 사용자 음성에 대한 응답을 제공하고, 사용자 음성을 분석하여 제어 명령을 결정하여 전자 장치에 제공하는 서비스를 말한다.
- [0024] 복수의 전자 장치(100-1 내지 100-n) 각각은 마이크를 통해 수신된 사용자 음성을 바탕으로 현재 전자 장치 주위에 사용자가 존재하는지 여부를 판단할 수 있다. 구체적으로, 마이크를 통해 수신된 사용자 음성의 크기가 임계값 이상인 경우, 전자 장치(100)는 전자 장치(100) 주변에 사용자가 존재하는 것으로 판단할 수 있다.
- [0025] 또한, 복수의 전자 장치(100-1 내지 100-n) 각각은 사용자의 음성 크기뿐만 아니라, 전자 장치(100)의 사용 정보, 전자 장치(100)의 사용 통계 정보, 사용자의 움직임 정보 중 적어도 하나를 이용하여 전자 장치 주위에 사용자가 존재하는지 여부를 판단할 수 있다. 예로, 복수의 전자 장치(100-1 내지 100-n) 각각은 임계 시간 내에 전자 장치(100)에 구비된 입력 인터페이스(예로, 리모컨, 버튼 등)를 통해 사용자 명령을 입력하였는지 여부 등과 같은 사용 정보를 바탕으로 전자 장치(100) 주위에 사용자가 존재하는지 여부를 판단할 수 있으며, 시간대별 사용자의 전자 장치 사용 통계(또는 사용자의 위치 통계 정보) 등을 바탕으로 전자 장치(100) 주위에 사용자가 존재하는지 여부를 판단할 수 있으며, 모션 센서 등에 의해 사용자의 움직임이 감지되었는지 여부를 바탕으로 전자 장치(100) 주위에 사용자가 존재하는지 여부를 판단할 수 있다.
- [0026] 특히, 전자 장치(100)는 사용자의 음성 크기, 전자 장치(100)의 사용 정보, 전자 장치(100)의 사용 통계 정보, 사용자의 움직임 정보 중 적어도 하나를 이용하여 전자 장치(100) 주위에 사용자가 존재할 확률값을 산출할 수 있으며, 임계값 이상의 확률값이 산출되면, 전자 장치(100)는 전자 장치 주위에 사용자가 존재하는 것으로 판단할 수 있다.
- [0027] 복수의 전자 장치(100-1 내지 100-n) 중 제1 전자 장치(100-1) 주위에 사용자가 존재하는 것으로 판단되면, 제1 전자 장치(100-1)는 제1 전자 장치(100-1) 및 타 전자 장치를 포함하는 디바이스 그룹을 결정할 수 있다.
- [0028] 일 실시예로, 제1 전자 장치(100-1)는 통신 인터페이스를 통해 복수의 타 전자 장치로부터 사용자의 트리거 음성에 대한 정보를 수신하고, 복수의 타 전자 장치 중 임계값 이상의 크기를 가지는 사용자의 트리거 음성을 감지한 적어도 하나의 타 전자 장치를 결정하고, 결정된 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정할 수 있다.
- [0029] 또 다른 실시예로, 제1 전자 장치(100-1)는 제1 전자 장치(100-1)의 위치를 바탕으로 적어도 하나의 타 전자 장치를 결정하고, 결정된 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정할 수 있다. 예로, 제1 전자 장치(100-1)의 위치가 거실인 경우, 제1 전자 장치(100-1)는 거실에 위치하는 적어도 하나의 타 전자 장치를 디바이스 그룹을 결정할 수 있다. 이때, 거실에 위치하는 적어도 하나의 타 전자 장치에 대한 정보는 기 저장될 수 있다.
- [0030] 제1 전자 장치(100-1)는 디바이스 그룹 내의 적어도 하나의 타 전자 장치 각각으로부터 수신한 타 전자 장치에 대한 정보를 바탕으로 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정할 수 있다. 이때, 타 전자 장치에 대한 정보는 타 전자 장치의 인터넷 연결 상태에 대한 정보, 타 전자 장치의 전원 상태에 대한 정보, 타 전자 장치의 메모리에 대한 정보, 타 전자 장치와 사용자와의 거리에 대한 정보, 타 전자 장치의 음성 인식 기능에 대한 정보 중 적어도 하나를 포함할 수 있다. 또한, 제1 전자 장치(100-1)는 디바이스 그룹 내의 적어도 하나의 타 전자 장치 각각으로 제1 전자 장치(100-1)에 대한 정보 및 제1 전자 장치(100-1)가 수신한 사용자 음성에 대한 정보를 전송할 수 있다.
- [0031] 한편, 타 전자 장치에 대한 정보는 실시간으로 디바이스 그룹 내에서 전송될 수 있으나, 이는 일 실시예에 불과할 뿐, 타 전자 장치에 대한 정보 중 적어도 일부(예로, 타 전자 장치의 성능에 대한 정보 등)는 기 저장될 수 있다.
- [0032] 제1 전자 장치(100-1)가 허브 장치로 결정되면, 제1 전자 장치(100-1)는 제1 전자 장치(100-1) 내에 포함된 마이크를 이용하여 맥 내의 전자 장치를 제어하거나 응답을 제공받기 위한 사용자의 음성 데이터를 획득할 수 있다. 이때, 제1 전자 장치(100-1)는 디바이스 그룹 내에 포함된 타 전자 장치들로부터 맥 내의 전자 장치를 제어하거나 응답을 제공받기 위한 사용자의 음성 데이터를 입력받을 수 있다.
- [0033] 제1 전자 장치(100-1)는 제1 전자 장치(100-1) 내에서 획득한 사용자의 음성 데이터와 타 전자 장치들로부터 수

신한 사용자의 음성 데이터를 취합하여 음성 인식을 수행할 수 있다. 이때, 음성 인식은 음성 데이터를 텍스트 데이터로 변환하는 것을 포함할 수 있으나, 이는 일 실시예에 불과할 뿐, 음성 데이터를 바탕으로 변환된 텍스트 데이터를 이용으로 응답을 생성하거나 제어 명령을 생성하는 것을 포함할 수 있다. 특히, 제1 전자 장치(100-1)는 전자 장치와 사용자 사이의 거리 및 전자 장치에 구비된 마이크의 개수를 바탕으로 음성 데이터를 취합할 수 있다. 즉, 제1 전자 장치(100-1)는 사용자와 가까운 전자 장치 및 마이크 개수가 많은 전자 장치가 수신한 음성 데이터에 높은 가중치를 부가하여 음성 데이터를 취합할 수 있다.

[0034] 구체적으로, 제1 전자 장치(100-1)는 외부 서버(50)로 취합된 음성 데이터를 전송할 수 있다. 외부 서버(50)는 대화 시스템을 이용하여 사용자의 음성 데이터를 인식하고, 사용자의 음성 데이터에 대한 응답 및 제어 명령 중 적어도 하나를 결정할 수 있다. 외부 서버(50)는 획득된 응답 및 제어 명령 중 적어도 하나를 제1 전자 장치(100-1)로 전송할 수 있다. 한편, 도 1에는 외부 서버(50)가 하나인 것으로 설명하였으나, 이는 일 실시예에 불과할 뿐, 응답을 제공하기 위한 제1 외부 서버 및 제어 명령을 판단하기 위한 제2 외부 서버로 구현될 수 있다.

[0035] 제1 전자 장치(100-1)는 수신된 응답을 출력하고, 제어 명령을 수행할 장치를 결정할 수 있다. 이때, 수신된 응답을 출력하고, 제어 명령을 수행할 장치가 동일한 장치일 수 있으나, 이는 일 실시예에 불과할 뿐, 수신된 응답을 출력하는 장치와 제어 명령을 수행하는 장치는 상이한 장치일 수 있다. 예로, 제1 전자 장치(100-1)는 수신된 응답을 출력할 장치로 제2 전자 장치(100-2)로 결정할 수 있고, 제어 명령을 수행할 장치로 제3 전자 장치(100-3)를 결정할 수 있다. 이때, 제1 전자 장치(100-1)는 사용자 위치 및 스피커에 대한 정보를 바탕으로 응답을 출력할 장치로 결정할 수 있다. 구체적으로, 제1 전자 장치(100-1)는 스피커가 존재하는 전자 장치들 중 사용자와의 거리 및 스피커의 성능(예로, 출력량)을 고려하여 응답을 출력할 장치를 결정할 수 있다. 그리고, 제1 전자 장치(100)는 수신된 응답을 출력할 장치에 응답을 전송하고, 제어 명령을 수행할 장치에 제어 명령을 전송할 수 있다.

[0036] 한편, 디바이스 그룹 내에 제1 전자 장치(100-1)가 아닌 타 전자 장치가 허브 장치로 결정되면, 제1 전자 장치(100-1)는 제1 전자 장치(100-1) 내에 포함된 마이크를 이용하여 맥 내의 전자 장치를 제어하거나 응답을 제공받기 위한 사용자의 음성 데이터를 획득하고, 획득된 사용자의 음성 데이터를 허브 장치로 결정된 타 전자 장치로 전송할 수 있다.

[0037] 한편, 제1 전자 장치(100-1) 내에 대화 시스템이 저장된 경우, 제1 전자 장치(100-1)는 대화 시스템을 이용하여 취합된 음성 데이터를 인식하고, 사용자의 음성 데이터에 대한 응답 및 제어 명령 중 적어도 하나를 결정할 수 있다.

[0038] 이때, 제1 전자 장치(100-1)는 상술한 바와 같은 사용자 음성에 대한 응답을 제공하기 위하여 인공지능 에이전트(Artificial intelligence agent)를 이용할 수 있다. 이때, 인공지능 에이전트는 AI(Artificial Intelligence) 기반의 대화형 서비스(예를 들어, 음성 인식 서비스, 비서 서비스, 번역 서비스, 검색 서비스 등)를 제공하기 위한 전용 프로그램으로서, 기존의 범용 프로세서(예를 들어, CPU) 또는 별도의 AI 전용 프로세서(예를 들어, NPU(Neural Processing Unit) 또는 GPU(Graphics Processing Unit) 등)에 의해 실행될 수 있다. 특히, 인공지능 에이전트는 후술할 다양한 모듈(예로, 대화 시스템)을 제어할 수 있다.

[0039] 구체적으로, 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러진 경우, 인공지능 에이전트가 동작할 수 있다. 그리고, 인공지능 에이전트는 사용자 음성에 대한 응답을 제공할 수 있다.

[0040] 물론, 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러지면 인공지능 에이전트가 동작할 수도 있다. 또한, 인공지능 에이전트는 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러지기 이전에 기 실행된 상태일 수 있다. 이 경우, 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러진 이후에는 제1 전자 장치(100-1)의 인공지능 에이전트가 사용자 질의에 대한 응답을 제공할 수 있다. 예를 들어, 인공지능 에이전트가 AI 전용 프로세서에 의해 실행되는 경우, 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 전자 장치(100)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러지기 전에는 범용 프로세서에 의해 제1 전자 장치(100-1)의 기능이 실행되며, 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러진 이후에는 AI 전용 프로세서에 의해 제1 전자 장치(100-1)의 기능이 실행될 수 있다.

- [0041] 또한, 인공지능 에이전트는 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러지기 이전에 대기 상태일 수 있다. 여기서 대기 상태란, 인공지능 에이전트의 동작 시작을 제어하기 위해 미리 정의된 사용자 입력이 수신되는 것을 감지하는 상태이다. 인공지능 에이전트가 대기 상태인 동안 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러지면, 제1 전자 장치(100-1)는 인공지능 에이전트를 동작시키고, 동작된 인공지능 에이전트를 이용하여 사용자 음성에 대한 응답을 제공할 수 있다.
- [0042] 또한, 인공지능 에이전트는 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러지기 이전에 종료된 상태일 수 있다. 인공지능 에이전트가 종료된 상태에서 기설정된 사용자 음성(예를 들어, "빅스비" 등)가 입력되거나 제1 전자 장치(100-1)에 구비된 버튼(예를 들어, 인공지능 에이전트를 실행하기 위한 버튼)이 눌러지면, 제1 전자 장치(100-1)는 인공지능 에이전트를 실행시키고, 실행된 인공지능 에이전트를 이용하여 사용자 음성에 대한 응답을 제공할 수 있다.
- [0043] 한편, 인공지능 에이전트는 후술할 다양한 장치 또는 모듈을 제어할 수 있다. 이에 대해서는 추후 상세히 설명하기로 한다.
- [0045] 도 2는 본 개시의 일 실시예에 따른, 전자 장치의 구성을 간략히 도시한 블록도이다. 도 2에 도시된 바와 같이, 전자 장치(100)는 마이크(110), 통신 인터페이스(120), 메모리(130) 및 프로세서(140)를 포함할 수 있다. 그러나 상술한 구성에 한정되는 것은 아니며, 전자 장치의 유형에 따라 일부 구성이 추가되거나 생략될 수 있음은 물론이다.
- [0046] 마이크(110)는 사용자 음성을 입력받을 수 있다. 이때, 마이크(110)는 음성 인식의 개시를 나타내는 트리거 음성(또는 웨이크업 음성)을 수신할 수 있으며, 맥 내의 복수의 전자 장치 중 적어도 하나를 제어하거나 응답을 제공받기 위한 사용자 음성을 수신할 수 있다. 특히, 마이크(110)는 전자 장치(100) 내부에 구비될 수 있으나, 외부에 구비되어 전자 장치(100)와 전기적으로 연결될 수 있다.
- [0047] 통신 인터페이스(120)는 외부 전자 장치와 통신을 수행할 수 있다. 한편, 통신 인터페이스(120)가 외부 장치와 통신 연결되는 것은 제3 기기(예로, 중계기, 허브, 액세스 포인트, 서버 또는 게이트웨이 등)를 거쳐서 통신하는 것을 포함할 수 있다.
- [0048] 한편, 통신 인터페이스(120)는 외부 장치와 통신을 수행하기 위해 다양한 통신 모듈을 포함할 수 있다. 일 예로, 통신 인터페이스(120)는 무선 통신 모듈을 포함할 수 있으며, 예를 들면, LTE, 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 사용하는 셀룰러 통신 모듈을 포함할 수 있다. 또 다른 예로, 무선 통신 모듈은, 예를 들면, WiFi(wireless fidelity), 블루투스, 블루투스 저전력(BLE), 지그비(Zigbee), NFC(near field communication), 자력 시큐어 트랜스미션(Magnetic Secure Transmission), 라디오 프리퀀시(RF), 또는 보디 에어리어 네트워크(BAN) 중 적어도 하나를 포함할 수 있다. 또한, 통신 인터페이스(120)는 유선 통신 모듈을 포함할 수 있으며, . 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. 무선 통신 또는 유선 통신이 수행되는 네트워크는 텔레커뮤니케이션 네트워크, 예를 들면, 컴퓨터 네트워크(예: LAN 또는 WAN), 인터넷, 또는 텔레폰 네트워크 중 적어도 하나를 포함할 수 있다.
- [0049] 또한, 통신 인터페이스(120)는 주변의 타 전자 장치로부터 타 전자 장치에 대한 정보 및 타 전자 장치가 수신한 사용자의 음성에 대한 정보를 수신할 수 있다. 또한, 통신 인터페이스(120)는 주변의 타 전자 장치로 전자 장치(100)에 대한 정보 및 전자 장치가 마이크(110)를 통해 수신한 사용자의 음성에 대한 정보를 전송할 수 있다.
- [0050] 또한, 통신 인터페이스(120)는 외부 서버(50)로 취합된 사용자의 음성 데이터를 전송할 수 있으며, 외부 서버(50)로부터 사용자 음성에 대한 응답 또는 제어 명령을 수신할 수 있다.
- [0051] 한편, 통신 인터페이스(120)는 주변의 타 전자 장치와 외부 서버(50)와 상이한 통신 모듈로 통신을 수행할 수 있다. 예로, 통신 인터페이스(120)는 주변의 타 전자 장치와 근거리 통신 모듈(예로, 블루투스 통신 모듈, 지그

비 통신 모듈, 적외선 통신 모듈 등)을 이용하여 통신을 수행할 수 있으며, 외부 서버와 원거리 통신 모듈(예로, LTE 통신 모듈, 5G 통신 모듈, 와이파이 통신 모듈 등)을 이용하여 통신을 수행할 수 있다.

[0052] 메모리(130)는 전자 장치(100)의 적어도 하나의 다른 구성요소에 관계된 명령 또는 데이터를 저장할 수 있다. 특히, 메모리(130)는 비휘발성 메모리, 휘발성 메모리, 플래시메모리(flash-memory), 하드디스크 드라이브(HDD) 또는 솔리드 스테이트 드라이브(SSD) 등으로 구현될 수 있다. 메모리(130)는 프로세서(140)에 의해 액세스되며, 프로세서(140)에 의한 데이터의 독취/기록/수정/삭제/갱신 등이 수행될 수 있다. 본 개시에서 메모리라는 용어는 메모리(130), 프로세서(140) 내 롬(미도시), 램(미도시) 또는 전자 장치(100)에 장착되는 메모리 카드(미도시)(예를 들어, micro SD 카드, 메모리 스틱)를 포함할 수 있다. 또한, 메모리(130)에는 디스플레이의 디스플레이 영역에 표시될 각종 화면을 구성하기 위한 프로그램 및 데이터 등이 저장될 수 있다.

[0053] 또한, 메모리(130)는 대화 시스템을 동작하기 위한 인공지능 에이전트를 저장할 수 있다. 구체적으로, 전자 장치(100)는 사용자 발화에 대한 응답으로 자연어를 생성하기 위하여 인공지능 에이전트(Artificial intelligence agent)를 이용할 수 있다. 이때, 인공지능 에이전트는 AI(Artificial Intelligence) 기반의 서비스(예를 들어, 음성 인식 서비스, 비서 서비스, 번역 서비스, 검색 서비스 등)를 제공하기 위한 전용 프로그램이다. 특히, 인공지능 에이전트는 기존의 범용 프로세서(예를 들어, CPU) 또는 별도의 AI 전용 프로세서(예를 들어, GPU 등)에 의해 실행될 수 있다.

[0054] 또한, 전자 장치(100)가 직접 대화형 서비스를 제공할 경우, 메모리(130)는 도 8에 도시되 바와 같은 대화 시스템을 구성하는 복수의 구성(또는 모듈)을 포함할 수 있다. 이에 대해서는 도 8을 참조하여 상세히 설명하기로 한다.

[0055] 프로세서(140)는 메모리(130)와 전기적으로 연결되어 전자 장치(100)의 전반적인 동작 및 기능을 제어할 수 있다. 특히, 프로세서(140)는 마이크(110)를 통해 수신된 사용자의 음성을 바탕으로 전자 장치(100) 주위에 사용자가 존재하는지 여부를 판단할 수 있다. 사용자가 상기 전자 장치 주위에 존재하면, 프로세서(140)는 전자 장치(100)와 전자 장치(100) 주변에 존재하는 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정하며, 적어도 하나의 타 전자 장치 각각으로부터 수신한 타 전자 장치에 대한 정보를 바탕으로 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정할 수 있다. 전자 장치(100)가 허브 장치로 결정되면, 프로세서(140)는 통신 인터페이스(120)를 통해 적어도 하나의 타 전자 장치 각각으로부터 사용자의 음성 데이터를 수신하여 음성 인식을 수행할 수 있다.

[0056] 구체적으로, 전자 장치(100)는 마이크(110)를 통해 수신된 사용자의 음성을 바탕으로 전자 장치(100) 주위에 사용자가 존재하는지 여부를 판단할 수 있다. 특히, 마이크(110)를 통해 수신된 사용자 음성의 크기가 임계값 이상인 경우, 전자 장치(100)는 전자 장치(100) 주위에 사용자가 존재하는 것으로 판단할 수 있다.

[0057] 또한, 프로세서(140)는 사용자의 음성뿐만 아니라, 전자 장치(100)의 사용 정보, 전자 장치(100)의 사용 통계 정보, 사용자의 움직임 정보 중 적어도 하나를 이용하여 전자 장치 주위에 사용자가 존재하는지 여부를 판단할 수 있다.

[0058] 사용자가 전자 장치(100) 주위에 존재하는 것으로 판단되면, 프로세서(140)는 전자 장치(100)와 적어도 하나의 타 전자 장치(100)를 포함하는 디바이스 그룹을 결정할 수 있다. 일 예로, 프로세서(140)는 통신 인터페이스(110)를 통해 복수의 타 전자 장치로부터 사용자의 음성에 대한 정보를 수신하고, 복수의 타 전자 장치 중 임계값 이상의 크기를 가지는 사용자의 음성을 감지한 적어도 하나의 타 전자 장치를 결정하고, 결정된 적어도 하나의 타 전자 장치와 전자 장치(100)를 포함하는 디바이스 그룹을 결정할 수 있다. 또 다른 예로, 프로세서(140)는 기저장된 전자 장치(100)의 위치에 대한 정보를 바탕으로 적어도 하나의 타 전자 장치를 결정하고, 결정된 적어도 하나의 타 전자 장치와 전자 장치(100)를 포함하는 디바이스 그룹을 결정할 수 있다.

[0059] 디바이스 그룹이 결정되면, 프로세서(140)는 적어도 하나의 타 전자 장치 각각으로부터 수신한 타 전자 장치에 대한 정보를 바탕으로 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정할 수 있다. 이때, 타 전자 장치에 대한 정보는 타 전자 장치의 인터넷 연결 상태에 대한 정보, 타 전자 장치의 전원 상태에 대한 정보, 타 전자 장치의 메모리에 대한 정보, 타 전자 장치와 사용자와의 거리에 대한 정보, 타 전자 장치의 음성 인식 기능에 대한 정보 중 적어도 하나를 포함할 수 있다.

[0060] 특히, 전자 장치(100)가 허브 장치로 결정되면, 프로세서(140)는 마이크(110)를 통해 획득된 사용자의 음성 데이터와 적어도 하나의 타 전자 장치 각각으로부터 수신된 사용자의 음성 데이터를 취합하여 외부 서버(50)에 전송하도록 통신 인터페이스(120)를 제어하고, 통신 인터페이스(120)를 통해 외부 서버로부터 취합된 사용자의 음

성 데이터에 대한 응답 및 취합된 사용자의 음성 데이터에 대응되는 제어 명령 중 적어도 하나를 수신할 수 있다.

- [0061] 그리고, 프로세서(140)는 사용자의 위치를 바탕으로 디바이스 그룹에 포함된 장치들 중 사용자의 음성 데이터에 대한 응답을 출력할 장치를 결정하고, 결정된 장치로 사용자의 음성 데이터에 대한 응답을 전송하도록 통신 인터페이스(120)를 제어할 수 있다.
- [0062] 또한, 프로세서(140)는 적어도 하나의 타 전자 장치 각각으로부터 수신된 상기 사용자의 음성 데이터를 바탕으로 상기 사용자의 음성 데이터에 대한 응답 및 사용자의 음성 데이터에 대응되는 제어 명령 중 적어도 하나를 결정할 수 있다. 이때, 프로세서(140)는 메모리(130)에 저장된 도 8의 대화 시스템을 이용하여 응답 및 제어 명령 중 적어도 하나를 결정할 수 있다.
- [0063] 또한, 적어도 하나의 타 전자 장치 중 하나가 허브 장치로 결정되면, 프로세서(140)는 허브 장치로 결정된 타 전자 장치로 마이크(110)를 통해 수신된 사용자의 음성 데이터를 전송하도록 통신 인터페이스(120)를 제어할 수 있다.
- [0064] 그리고, 음성 인식을 수행하는 동안 전자 장치(100) 주위에 사용자가 감지되지 않으면, 프로세서(140)는 음성 인식을 종료할 수 있다.
- [0066] 도 3은 본 개시의 일 실시예에 따른, 전자 장치의 구성을 상세히 도시한 블록도이다. 도 3에 도시된 바와 같이, 전자 장치(100)는 마이크(110), 통신 인터페이스(120), 메모리(130), 디스플레이(150), 스피커(160), 입력부(170), 센서(180) 및 프로세서(140)를 포함할 수 있다. 한편, 도 3에 도시된 마이크(110), 통신 인터페이스(120), 메모리(130) 및 프로세서(140)는 도 2에서 설명하였으므로, 중복되는 설명은 생략하기로 한다.
- [0067] 디스플레이(150)는 프로세서(140)의 제어에 따라 다양한 정보를 표시할 수 있다. 특히, 디스플레이(150)는 사용자 음성에 대한 응답을 제공하기 위한 메시지를 표시할 수 있다. 디스플레이(150)는 터치 패널과 함께 터치 스크린으로 구현될 수 있다.
- [0068] 스피커(160)는 오디오 처리부에 의해 디코딩이나 증폭, 노이즈 필터링과 같은 다양한 처리 작업이 수행된 각종 오디오 데이터뿐만 아니라 각종 알람 음이나 음성 메시지를 출력하는 구성이다. 특히, 스피커(160)는 사용자 음성에 대한 응답을 자연어 형태의 음성 메시지로 출력할 수 있다. 한편, 오디오를 출력하기 위한 구성은 스피커로 구현될 수 있으나, 이는 일 실시 예에 불과할 뿐, 오디오 데이터를 출력할 수 있는 출력 단자로 구현될 수 있다.
- [0069] 입력부(170)는 전자 장치(100)를 제어하기 위한 사용자 입력을 수신할 수 있다. 특히, 입력부(170)는 사용자 손 또는 스타일러스 펜 등을 이용한 사용자 터치를 입력받기 위한 터치 패널, 사용자 조작을 입력받기 위한 버튼 등이 포함될 수 있다. 그 밖에, 입력부(170)는 다른 입력 장치(예로, 키보드, 마우스, 모션 입력부 등)로 구현될 수 있다.
- [0070] 센서(180)는 전자 장치(100)의 다양한 상태 정보를 감지할 수 있다. 예로, 센서(180)는 전자 장치(100)의 움직임 정보를 감지할 수 있는 움직임 센서(예로, 자이로 센서, 가속도 센서 등)를 포함할 수 있으며, 위치 정보를 감지할 수 있는 센서(예로, GPS(Global Positioning System) 센서), 전자 장치(100) 주위의 환경 정보를 감지할 수 있는 센서(예로, 온도 센서, 습도 센서, 기압 센서 등), 전자 장치(100)의 사용자 정보를 감지할 수 있는 센서(예로, 혈압 센서, 혈당 센서, 맥박수 센서 등), 사용자의 존재를 감지할 수 있는 센서(예로, 카메라, UWB 센서, IR 센서, 근접 센서, 광센서 등) 등을 포함할 수 있다. 그 밖에, 센서(180)는 전자 장치(100)의 외부를 촬영하기 위한 이미지 센서 등을 더 포함할 수 있다.
- [0072] 도 4는 본 개시의 일 실시예에 따른, 복수의 전자 장치 중 허브 장치를 결정하여 대화형 서비스를 제공하기 위한 방법을 설명하기 위한 시퀀스도이다. 한편, 본 개시에 따른, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2)는택 내에 존재하는 전자 장치일 수 있으며, 외부 서버(50)를 통해 대화형 서비스를 제공받을 수 있다.
- [0073] 우선, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 트리거 음성을 입력받을 수 있다(S405). 이때, 트리거 음성은 음성 인식을 수행하여 대화형 서비스를 개시하기 위한 음성으로서, 웨이크업 음성이라고도 불릴 수 있다. 한편, 본 개시의 일 실시예에서는 트리거 음성을 입력하는 것으로 설명하였으나, 이는 일 실시예에 불과

할 뿐, 트리거 음성 이외에 일반적인 사용자 음성을 입력받을 수 있다. 즉, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2)는 자연스러운 사용자의 대화를 통해 사용자의 존재 여부를 판단할 수 있다.

[0074] 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 사용자 존재를 판단할 수 있다(S410). 이때, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 수신된 트리거 음성의 크기 및 도착 시간을 바탕으로 사용자가 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각으로부터 임계 범위 내에 존재하는지 여부를 판단할 수 있다. 예로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 임계값 이상의 크기를 가지는 트리거 음성이 감지되었는지 여부 및 임계 시간 이내에 트리거 음성이 도착하였는지 여부를 판단하여 사용자가 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각으로부터 임계 범위 내에 존재하는지 여부를 판단할 수 있다. 이때, 임계 범위는 기설정된 거리(예로, 3m 이내) 또는 같은 공간 범위를 말할 수 있다. 또한, 임계값 역시 기설정된 값으로서, 일반적인 음성의 크기가 상기 기설정된 거리에서 실험적으로 감지된 크기를 말할 수 있다. 또한, 임계 시간 역시 기설정된 값으로서, 상기 사용자 음성이 상기 기설정된 거리까지 도착하는 시간을 말할 수 있다.

[0075] 또한, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 트리거 음성의 크기 이외에도 전자 장치의 사용 정보, 전자 장치의 사용 통계 정보, 사용자의 움직임 정보 중 적어도 하나를 이용하여 전자 장치 주위에 사용자가 존재하는지 여부를 판단할 수 있다. 예로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 현재 입력부(170)를 통해 현재 시간으로부터 임계 시간(예로, 30초) 내에 사용자 명령이 입력되었는지 여부를 바탕으로 전자 장치 주위에 사용자가 존재하는지 여부를 판단할 수 있으며, 시간대별 전자 장치 사용 통계 정보를 바탕으로 전자 장치 주위에 사용자가 존재하는지 여부를 판단할 수 있으며, 센서(180)를 통해 사용자가 감지되거나 사용자의 움직임이 감지되는지 여부를 판단하여 전자 장치 주위에 사용자가 존재하는지 여부를 판단할 수 있다.

[0076] 사용자가 존재하는 것으로 판단되면, 제1 전자 장치(100-1)는 제1 전자 장치(100-1)에 대한 정보 및 트리거 음성에 대한 정보(예로, 제1 전자 장치(100-1)가 수신한 트리거 음성의 크기 정보, 트리거 음성의 SNR 정보, 음성 품질 스코어에 대한 정보 등)를 제2 전자 장치(100-2)로 전송할 수 있다(S415). 이때, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 트리거 음성에 대한 정보를 바탕으로 음성 품질 스코어를 직접 산출할 수 있다.

[0077] 그리고, 제2 전자 장치(100-2)는 제1 전자 장치(100-2)에 대한 정보 및 트리거 음성에 대한 정보(예로, 제2 전자 장치(100-2)가 수신한 트리거 음성의 크기 정보 등)를 제2 전자 장치(100-2)로 전송할 수 있다(S420).

[0078] 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 수신된 트리거 음성에 대한 정보를 바탕으로 디바이스 그룹을 생성할 수 있다(S425). 구체적으로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 수신된 트리거 음성에 대한 크기 또는 SNR이 기설정된 임계값 이상인지 여부를 판단하여 디바이스 그룹을 생성할 수 있다. 예로, 제1 전자 장치(100-1)가 수신한 트리거 음성의 크기 또는 SNR이 임계값 이상이며, 제1 전자 장치(100-2)가 수신한 트리거 음성의 크기 또는 SNR이 임계값 이상이면, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 제1 전자 장치(100-1) 및 제2 전자 장치(100-2)를 디바이스 그룹으로 결정할 수 있다. 또 다른 예로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 수신된 트리거 음성에 대한 음성 품질 스코어가 기설정된 임계값 이상인지 여부를 판단하여 디바이스 그룹을 생성할 수 있다. 예로, 제1 전자 장치(100-1)가 수신한 트리거 음성의 음성 품질 스코어가 임계값 이상이며, 제1 전자 장치(100-2)가 수신한 트리거 음성의 음성 품질 스코어가 임계값 이상이면, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 제1 전자 장치(100-1) 및 제2 전자 장치(100-2)를 디바이스 그룹으로 결정할 수 있다.

[0079] 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 수신된 타 전자 장치에 대한 정보를 바탕으로 허브 장치를 결정할 수 있다(S430). 구체적으로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 전자 장치의 인터넷 연결 상태에 대한 정보, 전자 장치의 전원 상태에 대한 정보, 전자 장치의 메모리에 대한 정보, 전자 장치와 사용자와의 거리에 대한 정보, 전자 장치의 음성 인식 기능에 대한 정보 중 적어도 하나를 바탕으로 허브 장치를 결정할 수 있다.

[0080] 구체적으로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 전자 장치 내에서 인공지능 서비스를 제공할 수 있는지 여부, 전자 장치의 외부망 연결 여부, 배터리 상태, 메모리 여유 공간을 바탕으로 후보 허브 장치를 결정할 수 있다. 즉, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 전자 장치 내에서 인공지능 서비스를 제공하거나 전자 장치가 외부망에 연결되어 있으며, 배터리가 임계값 이상(예로, 10% 이상)이고, 메모리의 여유 공간이 임계값 이상(예로, 100kb)인 전자 장치를 후보 허브 장치로 결정할 수 있다. 즉, 허브 장치가 되기 위해서는, 음성인식이 가능한 인공지능 서비스를 제공하거나 외부 인터넷망에 연결되어 있어야 하며, 최소한의 배터리와 최소한의 여유 메모리가 필요할 수 있다.

- [0081] 그리고, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 전자 장치의 처리 능력(예로, CPU 속도 및 처리량), 전자 장치와 AP 사이의 통신 연결 강도, 전자 장치와 사용자 사이의 거리를 바탕으로 후보 허브 장치 중 최종 허브 장치를 결정할 수 있다. 이때, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 전자 장치의 처리 능력, 전자 장치와 AP 사이의 통신 연결 강도 및 전자 장치와 사용자 사이의 거리에 가중치를 부여하여 최종 허브 장치를 결정하기 위한 스코어를 산출할 수 있으며, 산출된 스코어를 비교하여 최종 허브 장치를 결정할 수 있다.
- [0082] 한편, 상술한 실시예에서는 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각이 최종 허브 장치를 결정하는 것으로 설명하였으나, 이는 일 실시예에 불과할 뿐, 이전에 설정된 허브 장치가 존재하는 경우, 이전에 설정된 허브 장치가 상술한 방법을 통해 최종 허브 장치를 결정할 수 있다.
- [0083] 예로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 최종 허브 장치로 제1 전자 장치(100-1)를 결정할 수 있다.
- [0084] 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 사용자 음성을 입력받을 수 있다(S435). 이때, 사용자 음성은 맥내의 전자 장치를 제어하기 위한 음성이거나 사용자 음성에 대한 응답을 제공받기 위한 사용자 음성일 수 있다. 예로, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2) 각각은 "불꺼줘"와 같은 사용자 음성을 입력받을 수 있다.
- [0085] 제2 전자 장치(100-2)는 허브 장치인 제1 전자 장치(100-1)로 사용자의 음성 데이터를 전송할 수 있다(S440).
- [0086] 제1 전자 장치(100-1)는 마이크를 통해 수신된 사용자의 음성 데이터와 제2 전자 장치(100-2)로부터 수신된 사용자의 음성 데이터를 취합할 수 있다(S445). 이때, 제1 전자 장치(100-1)는 마이크를 통해 수신된 사용자의 음성 데이터와 제2 전자 장치(100-2)로부터 수신된 사용자의 음성 데이터 각각으로부터 노이즈를 제거하고, 노이즈가 제거된 음성 데이터를 동기화하여 취합할 수 있다.
- [0087] 제1 전자 장치(100-1)는 취합된 사용자의 음성 데이터를 외부 서버(50)로 전송할 수 있다(S450). 이때, 취합된 사용자의 음성 데이터는 텍스트 데이터가 아닌 오디오 데이터일 수 있다.
- [0088] 외부 서버(50)는 수신된 사용자의 음성 데이터를 바탕으로 응답 및 제어 명령을 결정할 수 있다(S455). 이때, 외부 서버(50)는 도 8에 도시된 대화 시스템을 바탕으로 응답을 자연어 형태로 획득할 수 있다. 또한, 외부 서버(50)는 응답 및 제어 명령 모두를 결정할 수 있으나, 이는 일 실시예에 불과할 뿐, 응답 및 제어 명령 중 하나만을 결정할 수 있다.
- [0089] 외부 서버(50)는 수신된 응답 및 제어 명령을 제1 전자 장치(100)로 전송할 수 있다(S460).
- [0090] 제1 전자 장치(100-1)는 응답을 출력하고, 제어 명령을 수행할 장치를 결정할 수 있다(S465). 이때, 제1 전자 장치(100-1)는 사용자 위치 및 스피커에 대한 정보를 바탕으로 응답을 출력할 장치로 결정할 수 있다. 구체적으로, 제1 전자 장치(100-1)는 스피커가 존재하는 전자 장치들 중 사용자와의 거리 및 스피커의 성능(예로, 출력량)을 고려하여 응답을 출력할 장치를 결정할 수 있다.
- [0091] 또한, 제1 전자 장치(100-1)는 기 저장된 전자 장치들의 위치 및 기능을 바탕으로 제어 명령을 수행할 장치를 결정할 수 있다. 예로, 제1 전자 장치(100-1)는 전등이 구비된 제2 전자 장치(100-2)를 제어 명령을 수행할 장치로 결정할 수 있다. 그러나, 이는 일 실시예에 불과할 뿐, 응답이 출력될 장치와 제어 명령이 수행될 장치가 상이할 수 있다. 예로, 응답은 제1 전자 장치(100-1)가 출력할 수 있으며, 제어 명령은 제2 전자 장치(100-2)가 수행할 수 있다.
- [0092] 제1 전자 장치(100-1)는 응답 및 제어 명령을 제2 전자 장치(100-2)로 전송할 수 있으며(S470), 제2 전자 장치(100-2)는 수신된 응답을 출력하고, 제어 명령에 따라 동작을 수행할 수 있다(S475).
- [0093] 한편, 상술한 실시예에서는 제1 전자 장치(100-1)가 외부 서버를 통해 제어 명령 및 응답을 획득하는 것으로 설명하였으나, 이는 일 실시예에 불과할 뿐, 제1 전자 장치(100-1)가 취합된 음성 데이터를 인식하여 제어 명령 및 응답을 획득할 수 있다. 구체적으로, 제1 전자 장치(100-1)는 취합된 음성 데이터가 기저장된 음성 명령어 세트와 비교하여 취합된 음성 데이터에 대응되는 제어 명령 및 응답이 제1 전자 장치(100-1) 내에 저장되었는지 여부를 판단할 수 있다. 그리고, 취합된 음성 데이터에 대응되는 제어 명령 및 응답이 제1 전자 장치(100-1) 내에 저장된 경우, 제1 전자 장치(100-1)는 저장된 음성 명령어 세트를 바탕으로 제어 명령 및 응답을 획득할 수 있다. 그러나, 취합된 음성 데이터에 대응되는 제어 명령 및 응답이 제1 전자 장치(100-1)내에 저장되어 있지

많은 경우, 제1 전자 장치(100-1)는 상술한 바와 같이, 취합된 음성 데이터를 외부 서버(50)로 전송할 수 있다.

- [0095] 도 5는 본 개시의 다른 실시예에 따른, 복수의 전자 장치 중 허브 장치를 결정하여 대화형 서비스를 제공하기 위한 방법을 설명하기 위한 시퀀스도이다. 한편, 본 개시에 따른, 제1 전자 장치(100-1) 및 제2 전자 장치(100-2)는 맥 내에 존재하는 전자 장치일 수 있으며, 제1 전자 장치(100-1) 내에는 음성 인식을 수행하여 대화형 서비스를 제공할 수 있는 인공지능 프로그램이 저장될 수 있다. 한편, 도 4에서 설명한 S405 단계 내지 S445 단계는 도 5의 S505 단계 내지 S545와 동일하므로, 중복되는 설명은 생략한다.
- [0096] 제1 전자 장치(100-1)는 취합된 사용자의 음성 데이터를 바탕으로 응답 및 제어 명령을 결정할 수 있다(S550). 이때, 제1 전자 장치(100-1)는 도 8에 도시된 대화 시스템을 바탕으로 응답을 자연어 형태로 획득할 수 있다. 또한, 제1 전자 장치(100-1)는 응답 및 제어 명령 모두를 결정할 수 있으나, 이는 일 실시예에 불과할 뿐, 응답 및 제어 명령 중 하나만을 결정할 수 있다.
- [0097] 제1 전자 장치(100-1)는 응답을 출력하고, 제어 명령을 수행할 장치를 결정할 수 있다(S555). 이때, 제1 전자 장치(100-1)는 사용자 위치를 바탕으로 사용자와 가장 가까운 장치를 응답을 출력할 장치로 결정할 수 있다. 또한, 제1 전자 장치(100-1)는 기 저장된 전자 장치들의 위치 및 기능을 바탕으로 제어 명령을 수행할 장치를 결정할 수 있다. 예로, 제1 전자 장치(100-1)는 이동된 사용자와 가장 가까운 제2 전자 장치(100-2)를 응답을 출력할 장치로 결정할 수 있으며, 전등이 구비된 제2 전자 장치(100-2)를 제어 명령을 수행할 장치로 결정할 수 있다.
- [0098] 제1 전자 장치(100-1)는 응답 및 제어 명령을 제2 전자 장치(100-2)로 전송할 수 있으며(S560), 제2 전자 장치(100-2)는 수신된 응답을 출력하고, 제어 명령에 따라 동작을 수행할 수 있다(S565).
- [0100] 도 6a 내지 도 6c는 본 개시의 일 실시예에 따른, 복수의 전자 장치 중 허브 장치를 결정하여 대화형 서비스를 제공하는 실시예를 설명하기 위한 도면들이다. 맥 내에는 도 6a에 도시된 바와 같이, 제1 내지 제9 전자 장치(100-1 내지 100-9)가 포함될 수 있다. 이때, 거실에는 제1 내지 제3 전자 장치(100-1 내지 100-3)가 위치하며, 안방에는 제4 내지 제6 전자 장치(100-4 내지 100-6)가 위치하며, 화장실에는 제7 전자 장치(100-7)가 위치하며, 부엌에는 제8 및 제9 전자 장치(100-8 및 100-9)가 위치할 수 있다.
- [0101] 사용자는 도 6a에 도시된 바와 같이, 안방에서 "하이 빅스비"라는 트리거 음성을 발화할 수 있다.
- [0102] 제4 내지 제6 전자 장치(100-4 내지 100-6) 각각은 마이크를 통해 "하이 빅스비"라는 트리거 음성을 입력받을 수 있다. 제4 내지 제6 전자 장치(100-4 내지 100-6) 각각은 수신된 트리거 음성의 크기를 바탕으로 제4 내지 제6 전자 장치(100-4 내지 100-6) 주변에 사용자가 존재함을 판단할 수 있다.
- [0103] 그리고, 제4 내지 제6 전자 장치(100-4 내지 100-6) 각각은 수신된 트리거 음성의 크기를 서로 교환하여 도 6b에 도시된 바와 같이, 제4 내지 제6 전자 장치(100-4 내지 100-6)를 하나의 디바이스 그룹(600)으로 결정할 수 있다. 또는, 제4 내지 제6 전자 장치(100-4 내지 100-6) 각각은 기 저장된 전자 장치들의 위치를 바탕으로 제4 내지 제6 전자 장치(100-4 내지 100-6)를 하나의 디바이스 그룹으로 결정할 수 있다.
- [0104] 그리고, 제4 내지 제6 전자 장치(100-4 내지 100-6) 각각은 전자 장치에 대한 정보를 교환하여 허브 장치를 결정할 수 있다. 이때, 제4 내지 제6 전자 장치(100-4 내지 100-6) 각각은 상술한 바와 같은 방법으로 제5 전자 장치(100-5)를 허브 장치로 결정할 수 있다.
- [0105] 허브 장치가 결정된 후, 제4 내지 제6 전자 장치(100-4 내지 100-6) 각각은 도 6b에 도시된 바와 같이, "불켜줘"라는 사용자 음성을 획득할 수 있다. 이때, 제4 및 제6 전자 장치(100-4 및 100-6)는 마이크를 통해 수신된 음성 데이터를 제5 전자 장치(100-5)로 전송할 수 있다.
- [0106] 제5 전자 장치(100-5)는 제4 및 제6 전자 장치(100-4 및 100-6)로부터 수신된 음성 데이터와 제5 전자 장치(100-5)에 구비된 마이크를 통해 획득된 음성 데이터를 취합하여 사용자의 음성 데이터를 획득할 수 있다.
- [0107] 제5 전자 장치(100-5)는 외부 서버(50)로 취합된 음성 데이터를 전송할 수 있으며, 외부 서버(50)로부터 음성 데이터에 대한 응답인 "안방 불켜줘"와 "light. off"라는 제어 명령을 수신할 수 있다.
- [0108] 제5 전자 장치(100-5)는 수신된 응답을 출력할 장치와 제어 명령을 수신할 장치를 결정할 수 있다. 구체적으로,

제5 전자 장치(100-5)는 사용자와 가장 가까이 존재하는 제5 전자 장치(100-5)를 응답을 출력할 장치로 결정할 수 있으며, 기저장된 전자 장치들의 위치 정보 및 기능에 대한 정보를 바탕으로 안방에 존재하면서 전등을 포함하는 제4 및 제6 전자 장치(100-4 및 100-6)를 제어 명령을 수행할 장치로 결정할 수 있다.

[0109] 제5 전자 장치(100-5)는 도 6c에 도시된 바와 같이, "안방 불 끄개"라는 응답을 출력할 수 있으며, 제4 및 제6 전자 장치(100-4 및 100-6)로 "light.off"라는 제어 명령을 전송할 수 있다. 제4 및 제6 전자 장치(100-4 및 100-6)는 수신된 제어 명령을 바탕으로 전등의 전원을 오프하는 기능을 수행할 수 있다.

[0110] 또한, 제5 전자 장치(100-5)는 음성을 획득하지 않는 타 전자 장치도 제어 명령을 수신할 장치로 결정할 수 있다. 구체적으로, 제5 전자 장치(100-5)는 현재 맥 내에 불이 켜져 있는 전자 장치에 대한 정보를 획득하고, 현재 맥 내에 존재하는 전자 장치 중 불이 켜져 있는 제3, 제4, 제6, 제7 전자 장치(100-3, 100-4, 100-6, 100-7)를 제어 명령을 수행할 장치로 결정할 수 있다. 즉, 제5 전자 장치(100-5)는 사용자 음성을 수신한 제4 및 제6 전자 장치(100-4, 100-6)를 제어 명령을 수신할 전자 장치로 결정할 수 있을 뿐만 아니라, 사용자 음성을 수신하지 않은 제3 및 제7 전자 장치(100-3, 100-7)(또는 디바이스 그룹에 속하지 않은 전자 장치)를 제어 명령을 수신할 전자 장치로 결정할 수 있다.

[0111] 제5 전자 장치(100-5)는 도 6d에 도시된 바와 같이, "불 켜진 곳에 불 끄개"라는 응답을 출력할 수 있으며, 제3, 제4, 제6, 제7 전자 장치(100-3, 100-4, 100-6, 100-7)로 "light.off"라는 제어 명령을 전송할 수 있다. 제3, 제4, 제6, 제7 전자 장치(100-3, 100-4, 100-6, 100-7)는 수신된 제어 명령을 바탕으로 전등의 전원을 오프하는 기능을 수행할 수 있다.

[0112] 한편, 도 6d에서 설명한 실시예에서는 외부 서버(50)가 제5 전자 장치(100-5)를 통해 제3, 제4, 제6, 제7 전자 장치(100-3, 100-4, 100-6, 100-7)에 제어 명령을 전송하였으나, 이는 일 실시예에 불과할 뿐, 외부 서버(50)가 직접 제3, 제4, 제6, 제7 전자 장치(100-3, 100-4, 100-6, 100-7)에 제어 명령을 전송할 수 있다. 즉, 외부 서버(50)는 현재 맥 내에 불이 켜진 전자 장치로서, 제3, 제4, 제6, 제7 전자 장치(100-3, 100-4, 100-6, 100-7)를 판단하고, 판단된 제3, 제4, 제6, 제7 전자 장치(100-3, 100-4, 100-6, 100-7)에 제어 명령을 전송할 수 있다.

[0114] 도 7은 본 개시의 일 실시예에 따른, 전자 장치의 제어 방법을 설명하기 위한 흐름도이다.

[0115] 우선, 전자 장치(100)는 마이크(100)를 통해 수신된 사용자의 음성을 바탕으로 전자 장치(100) 주위에 사용자가 존재하는지 여부를 판단할 수 있다(S710). 이때, 전자 장치(100)는 사용자 음성 이외에 전자 장치(100)의 사용 정보, 전자 장치(100)의 사용 통계 정보, 사용자의 움직임 정보 중 적어도 하나를 이용하여 전자 장치(100) 주위에 사용자가 존재하는지 여부를 판단할 수 있다.

[0116] 사용자가 전자 장치(100) 주위에 존재하면, 전자 장치(100)는 전자 장치(100)와 전자 장치(100) 주변에 존재하는 적어도 하나의 타 전자 장치를 포함하는 디바이스 그룹을 결정할 수 있다(S720). 이때, 전자 장치(100)는 타 전자 장치로부터 수신된 사용자 음성의 크기를 바탕으로 디바이스 그룹을 결정할 수 있으며, 기 저장된 전자 장치들의 위치 정보를 바탕으로 디바이스 그룹을 결정할 수 있다.

[0117] 전자 장치(100)는 적어도 하나의 타 전자 장치 각각으로부터 수신한 타 전자 장치에 대한 정보를 바탕으로 디바이스 그룹 중 하나의 장치를 음성 인식을 수행하기 위한 허브 장치로 결정할 수 있다(S730). 구체적으로, 전자 장치(100)는 전자 장치들의 인터넷 연결 상태에 대한 정보, 전자 장치들의 전원 상태에 대한 정보, 전자 장치들의 메모리에 대한 정보, 전자 장치들과 사용자와의 거리에 대한 정보, 전자 장치들의 음성 인식 기능에 대한 정보 중 적어도 하나를 바탕으로 허브 장치를 결정할 수 있다.

[0118] 전자 장치(100)는 전자 장치(100)가 허브 장치로 결정되었는지 여부를 판단할 수 있다(S740).

[0119] 전자 장치(100)가 허브 장치로 결정되면(S740-Y), 전자 장치(100)는 적어도 하나의 타 전자 장치 각각으로부터 사용자의 음성 데이터를 수신하여 음성 인식을 수행할 수 있다(S760). 이때, 음성 인식은 전자 장치(100) 내에서 수행될 수 있으며, 외부 서버(50)를 통해 수행될 수 있다.

[0120] 타 전자 장치가 허브 장치로 결정되면(740-N), 전자 장치(100)는 허브 장치로 결정된 타 전자 장치로 사용자의 음성 데이터를 전송할 수 있다(S750).

[0121] 상술한 바와 같은 본 개시의 실시예에 따라, 사용자의 위치에 따라 적응적으로 허브 장치를 결정함으로써, 불필

요한 네트워크 오버헤드를 제거할 수 있으며, 더욱 정확한 대화형 서비스를 제공할 수 있게 된다.

- [0123] 도 8은 본 개시의 일 실시예에 따른, 인공지능 에이전트 시스템의 대화 시스템을 도시한 블록도이다. 도 8에 도시된 대화 시스템(800)은 가상의 인공지능 에이전트와 자연어를 통해 대화를 수행하기 위한 구성으로서, 본 개시의 일 실시예에 따르면, 대화 시스템(800)은 전자 장치(100)의 메모리(130) 내에 저장될 수 있다. 그러나, 이는 일 실시예에 불과할 뿐, 대화 시스템(400)에 포함된 적어도 하나는 외부의 적어도 하나 서버(50)에 포함될 수 있다.
- [0124] 대화 시스템(400)은 도 4에 도시된 바와 같이, 자동 음성 인식(automatic speech recognition)(ASR) 모듈(810), 자연어 이해(natural language understanding)(NLU) 모듈(820), 대화 매니저(dialogue manager)(DM) 모듈(830), 자연어 생성(natural language generator)(NLG) 모듈(840) 및 텍스트 음성 변환(text to speech)(TTS) 모듈(850)을 포함할 수 있다. 그 밖에 대화 시스템(800)은 패스 플래너(path planner) 모듈 또는 액션 플래너(action planner) 모듈을 더 포함할 수 있다.
- [0125] 자동 음성 인식(automatic speech recognition)(ASR) 모듈(810)은 전자 장치(100)로부터 수신된 사용자의 음성 데이터를 텍스트 데이터로 변환할 수 있다. 예를 들어, 자동 음성 인식 모듈(810)은 발화 인식 모듈을 포함할 수 있다. 상기 발화 인식 모듈은 음향(acoustic) 모델 및 언어(language) 모델을 포함할 수 있다. 예를 들어, 음향 모델은 발성에 관련된 정보를 포함할 수 있고, 언어 모델은 단위 음소 정보 및 단위 음소 정보의 조합에 대한 정보를 포함할 수 있다. 발화 인식 모듈은 발성에 관련된 정보 및 단위 음소 정보에 대한 정보를 이용하여 사용자 발화를 텍스트 데이터로 변환할 수 있다. 음향 모델 및 언어 모델에 대한 정보는, 예를 들어, 자동 음성 인식 데이터베이스(automatic speech recognition database)(ASR DB)(815)에 저장될 수 있다.
- [0126] 자연어 이해 모듈(820)은 문법적 분석(syntactic analyze) 또는 의미적 분석(semantic analyze)을 수행하여 사용자 의도를 파악할 수 있다. 문법적 분석은 사용자 입력을 문법적 단위(예: 단어, 구, 형태소 등)로 나누고, 나누어진 단위가 어떤 문법적인 요소를 갖는지 파악할 수 있다. 의미적 분석은 의미(semantic) 매칭, 룰(rule) 매칭, 포플러(formula) 매칭 등을 이용하여 수행할 수 있다. 이에 따라, 자연어 이해 모듈(820)은 사용자 입력이 어느 도메인(domain), 의도(intent) 또는 의도를 표현하는데 필요한 파라미터(parameter)(또는, 슬롯(slot))를 얻을 수 있다.
- [0127] 자연어 이해 모듈(820)은 도메인(domain), 의도(intend) 및 의도를 파악하는데 필요한 파라미터(parameter)(또는, 슬롯(slot))로 나누어진 매칭 규칙을 이용하여 사용자의 의도 및 파라미터를 결정할 수 있다. 예를 들어, 상기 하나의 도메인(예: 알람)은 복수의 의도(예: 알람 설정, 알람 해제 등)를 포함할 수 있고, 하나의 의도는 복수의 파라미터(예: 시간, 반복 횟수, 알람음 등)를 포함할 수 있다. 복수의 룰은, 예를 들어, 하나 이상의 필수 요소 파라미터를 포함할 수 있다. 매칭 규칙은 자연어 인식 데이터베이스(natural language understanding database)(NLU DB)(825)에 저장될 수 있다.
- [0128] 자연어 이해 모듈(820)은 형태소, 구 등의 언어적 특징(예: 문법적 요소)을 이용하여 사용자 입력으로부터 추출된 단어의 의미를 파악하고, 파악된 단어의 의미를 도메인 및 의도에 매칭시켜 사용자의 의도를 결정할 수 있다. 예를 들어, 자연어 이해 모듈(820)은 각각의 도메인 및 의도에 사용자 입력에서 추출된 단어가 얼마나 포함되어 있는지를 계산하여 사용자 의도를 결정할 수 있다. 일 실시 예에 따르면, 자연어 이해 모듈(820)은 의도를 파악하는데 기초가 된 단어를 이용하여 사용자 입력의 파라미터를 결정할 수 있다. 일 실시 예에 따르면, 자연어 이해 모듈(820)은 사용자 입력의 의도를 파악하기 위한 언어적 특징이 저장된 자연어 인식 데이터베이스(423)를 이용하여 사용자의 의도를 결정할 수 있다.
- [0129] 자연어 이해 모듈(820)은 지식 베이스(Knowledge Base)(835)를 이용하여 사용자의 의도를 결정할 수 있다. 예를 들어, 자연어 이해 모듈(820)은 사용자 정보(예: 선호 문구, 선호 콘텐츠, 연락처 리스트, 음악 리스트 등)를 이용하여 사용자의 의도를 결정할 수 있다. 일 실시 예에 따르면, 자연어 이해 모듈(820)뿐만 아니라 자동 음성 인식 모듈(810)도 지식 베이스(835)를 참고하여 사용자의 음성을 인식할 수 있다.
- [0130] 자연어 이해 모듈(820)은 사용자 입력의 의도 및 파라미터에 기초하여 패스 룰을 생성할 수 있다. 예를 들어, 자연어 이해 모듈(820)은 사용자 입력의 의도에 기초하여 실행될 앱을 선택하고, 선택된 앱에서 수행될 동작을 결정할 수 있다. 자연어 이해 모듈(820)은 결정된 동작에 대응되는 파라미터를 결정하여 패스 룰을 생성할 수 있다. 일 실시 예에 따르면, 자연어 이해 모듈(820)에 의해 생성된 패스 룰은 실행될 앱, 앱에서 실행될 동작 및 상기 동작을 실행하는데 필요한 파라미터에 대한 정보를 포함할 수 있다.

- [0131] 자연어 이해 모듈(820)은 사용자 입력의 의도 및 파라미터를 기반으로 하나의 패스 룰, 또는 복수의 패스 룰을 생성할 수 있다. 예를 들어, 자연어 이해 모듈(820)은 패스 플래너 모듈로부터 전자 장치(100)에 대응되는 패스 룰 셋을 수신하고, 사용자 입력의 의도 및 파라미터를 수신된 패스 룰 셋에 맵핑하여 패스 룰을 결정할 수 있다. 이때, 패스 룰은 앱의 기능을 수행하기 위한 동작(또는 오퍼레이션(operation))에 대한 정보 또는 동작을 실행하기 위해 필요한 파라미터에 대한 정보를 포함할 수 있다. 또한, 패스 룰은 앱의 동작 순서를 포함할 수 있다. 전자 장치는 패스 룰을 수신하고, 패스 룰에 따라 앱을 선택하고, 선택된 앱에서 패스 룰에 포함된 동작을 실행시킬 수 있다.
- [0132] 자연어 이해 모듈(820)은 사용자 입력의 의도 및 파라미터에 기초하여 실행될 앱, 앱에서 실행될 동작 및 상기 동작을 실행하는데 필요한 파라미터를 결정하여 하나의 패스 룰, 또는 복수의 패스 룰을 생성할 수 있다. 예를 들어, 자연어 이해 모듈(820)은 전자 장치(100)의 정보를 이용하여 실행될 앱 및 상기 앱에서 실행될 동작을 사용자 입력의 의도에 따라 온톨로지(ontology) 또는 그래프 모델(graph model) 형태로 배열하여 패스 룰을 생성할 수 있다. 상기 생성된 패스 룰은, 예를 들어, 패스 플래너 모듈을 통해 패스 룰 데이터베이스(path rule database)에 저장될 수 있다. 상기 생성된 패스 룰은 데이터베이스(825)의 패스 룰 셋에 추가될 수 있다.
- [0133] 자연어 이해 모듈(820)은 생성된 복수의 패스 룰 중 적어도 하나의 패스 룰을 선택할 수 있다. 예를 들어, 자연어 이해 모듈(820)은 복수의 패스 룰 최적의 패스 룰을 선택할 수 있다. 다른 예를 들어, 자연어 이해 모듈(820)은 사용자 발화에 기초하여 일부 동작만이 특정된 경우 복수의 패스 룰을 선택할 수 있다. 자연어 이해 모듈(820)은 사용자의 추가 입력에 의해 복수의 패스 룰 중 하나의 패스 룰을 결정할 수 있다.
- [0134] 대화 매니저 모듈(830)은 자연어 이해 모듈(820)에 의해 파악된 사용자의 의도가 명확한지 여부를 판단할 수 있다. 예를 들어, 대화 매니저 모듈(830)은 파라미터의 정보가 충분하지 여부에 기초하여 사용자의 의도가 명확한지 여부를 판단할 수 있다. 대화 매니저 모듈(830)은 자연어 이해 모듈(820)에서 파악된 파라미터가 태스크를 수행하는데 충분한지 여부를 판단할 수 있다. 일 실시 예에 따르면, 대화 매니저 모듈(830)은 사용자의 의도가 명확하지 않은 경우 사용자에게 필요한 정보를 요청하는 피드백을 수행할 수 있다. 예를 들어, 대화 매니저 모듈(830)은 사용자의 의도를 파악하기 위한 파라미터에 대한 정보를 요청하는 피드백을 수행할 수 있다. 또한, 대화 매니저 모듈(830)은 자연어 이해 모듈(820)에 의해 변경된 텍스트를 포함하는 사용자 질의를 확인하기 위한 메시지를 생성하여 출력할 수 있다.
- [0135] 일 실시 예에 따르면, 대화 매니저 모듈(830)은 콘텐츠 제공(content provider) 모듈을 포함할 수 있다. 콘텐츠 제공 모듈은 자연어 이해 모듈(820)에서 파악된 의도 및 파라미터에 기초하여 동작을 수행할 수 있는 경우, 사용자 입력에 대응되는 태스크를 수행한 결과를 생성할 수 있다.
- [0136] 다른 실시 예에 따르면, 대화 매니저 모듈(830)은 지식 베이스(835)를 이용하여 사용자 음성에 대한 응답을 제공할 수 있다. 이때, 지식 베이스(835)는 전자 장치(100) 내에 포함될 수 있으나, 이는 일 실시예에 불과할 뿐, 외부 서버에 포함될 수 있다.
- [0137] 자연어 생성 모듈(NLG 모듈)(840)은 지정된 정보를 텍스트 형태로 변경할 수 있다. 상기 텍스트 형태로 변경된 정보는 자연어 발화의 형태일 수 있다. 상기 지정된 정보는, 예를 들어, 추가 입력에 대한 정보, 사용자 입력에 대응되는 동작의 완료를 안내하는 정보 또는 사용자의 추가 입력을 안내하는 정보(예: 사용자 입력에 대한 피드백 정보)일 수 있다. 상기 텍스트 형태로 변경된 정보는 전자 장치(100)의 디스플레이(150)에 표시되거나, 텍스트 음성 변환 모듈(TTS 모듈)(850)에 의해 음성 형태로 변경될 수 있다.
- [0138] 텍스트 음성 변환 모듈(TTS 모듈)(850)은 텍스트 형태의 정보를 음성 형태의 정보로 변경할 수 있다. 텍스트 음성 변환 모듈(850)은 자연어 생성 모듈(840)로부터 텍스트 형태의 정보를 수신하고, 텍스트 형태의 정보를 음성 형태의 정보로 변경하여 스피커로 출력할 수 있다.
- [0139] 자연어 이해 모듈(820) 및 대화 매니저 모듈(830)은 하나의 모듈로 구현될 수 있다. 예를 들어, 자연어 이해 모듈(820) 및 대화 매니저 모듈(830)은 하나의 모듈로 구현되어 사용자의 의도 및 파라미터를 결정하고, 상기 결정된 사용자의 의도 및 파라미터에 대응되는 응답(예로, 패스 룰)을 획득할 수 있다.
- [0141] 한편, 본 개시에서 사용된 용어 "부" 또는 "모듈"은 하드웨어, 소프트웨어 또는 펌웨어로 구성된 유닛을 포함하며, 예를 들면, 로직, 논리 블록, 부품, 또는 회로 등의 용어와 상호 호환적으로 사용될 수 있다. "부" 또는 "모듈"은, 일체로 구성된 부품 또는 하나 또는 그 이상의 기능을 수행하는 최소 단위 또는 그 일부가 될 수

있다. 예를 들면, 모듈은 ASIC(application-specific integrated circuit)으로 구성될 수 있다.

[0142] 본 개시의 다양한 실시예들은 기기(machine)(예: 컴퓨터)로 읽을 수 있는 저장 매체(machine-readable storage media)에 저장된 명령어를 포함하는 소프트웨어로 구현될 수 있다. 기기는, 저장 매체로부터 저장된 명령어를 호출하고, 호출된 명령어에 따라 동작이 가능한 장치로서, 개시된 실시예들에 따른 전자 장치(예: 전자 장치(100))를 포함할 수 있다. 상기 명령이 프로세서에 의해 실행될 경우, 프로세서가 직접, 또는 상기 프로세서의 제어 하에 다른 구성요소들을 이용하여 상기 명령에 해당하는 기능을 수행할 수 있다. 명령은 컴파일러 또는 인터프리터에 의해 생성 또는 실행되는 코드를 포함할 수 있다. 기기로 읽을 수 있는 저장매체는, 비일시적(non-transitory) 저장매체의 형태로 제공될 수 있다. 여기서, '비일시적'은 저장매체가 신호(signal)를 포함하지 않으며 실재(tangible)하다는 것을 의미할 뿐 데이터가 저장매체에 반영구적 또는 임시적으로 저장됨을 구분하지 않는다.

[0143] 일시에 따르면, 본 문서에 개시된 다양한 실시예들에 따른 방법은 컴퓨터 프로그램 제품(computer program product)에 포함되어 제공될 수 있다. 컴퓨터 프로그램 제품은 상품으로서 판매자 및 구매자 간에 거래될 수 있다. 컴퓨터 프로그램 제품은 기기로 읽을 수 있는 저장 매체(예: compact disc read only memory (CD-ROM))의 형태로, 또는 어플리케이션 스토어(예: 플레이 스토어™)를 통해 온라인으로 배포될 수 있다. 온라인 배포의 경우에, 컴퓨터 프로그램 제품의 적어도 일부는 제조사의 서버, 어플리케이션 스토어의 서버, 또는 중계 서버의 메모리와 같은 저장 매체에 적어도 일시 저장되거나, 임시적으로 생성될 수 있다.

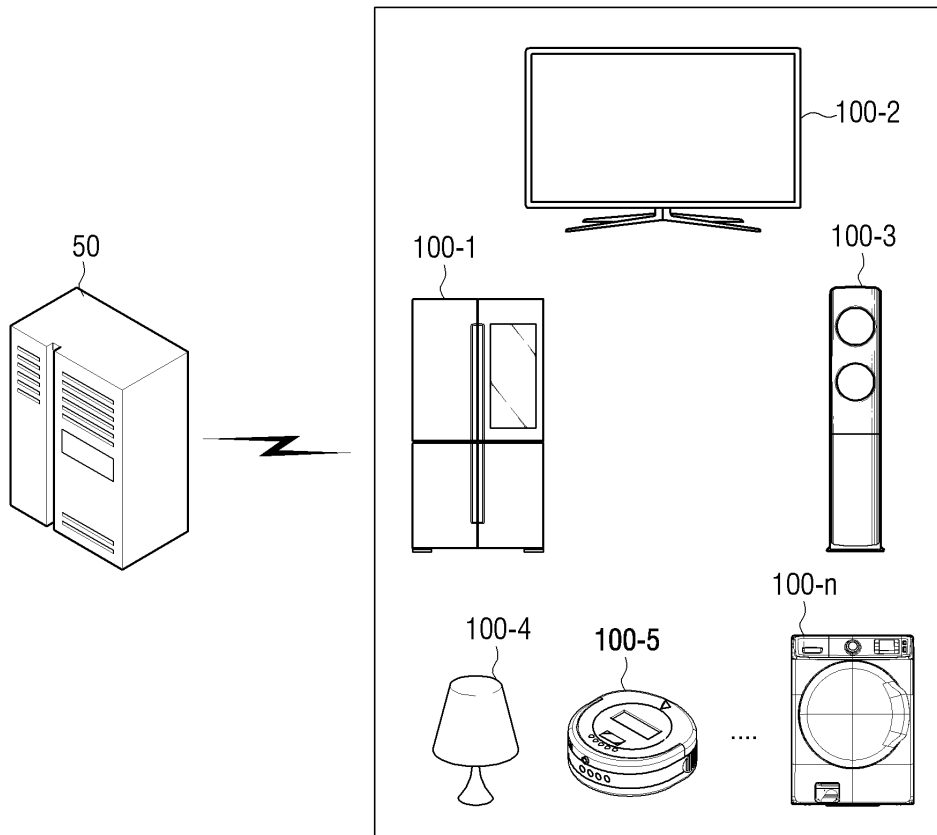
[0144] 다양한 실시예들에 따른 구성 요소(예: 모듈 또는 프로그램) 각각은 단수 또는 복수의 개체로 구성될 수 있으며, 전술한 해당 서브 구성 요소들 중 일부 서브 구성 요소가 생략되거나, 또는 다른 서브 구성 요소가 다양한 실시예에 더 포함될 수 있다. 대체적으로 또는 추가적으로, 일부 구성 요소들(예: 모듈 또는 프로그램)은 하나의 개체로 통합되어, 통합되기 이전의 각각의 해당 구성 요소에 의해 수행되는 기능을 동일 또는 유사하게 수행할 수 있다. 다양한 실시예들에 따른, 모듈, 프로그램 또는 다른 구성 요소에 의해 수행되는 동작들은 순차적, 병렬적, 반복적 또는 휴리스틱하게 실행되거나, 적어도 일부 동작이 다른 순서로 실행되거나, 생략되거나, 또는 다른 동작이 추가될 수 있다.

부호의 설명

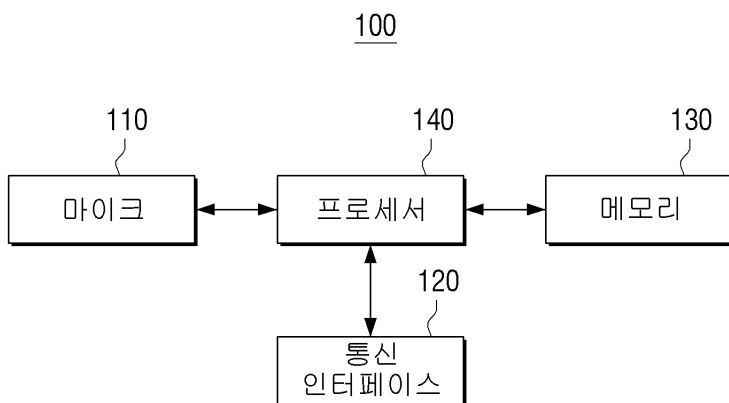
- [0145] 110: 마이크 120: 통신 인터페이스
 130: 메모리 140: 프로세서
 150: 디스플레이 160: 스피커
 170: 입력부 180: 센서

도면

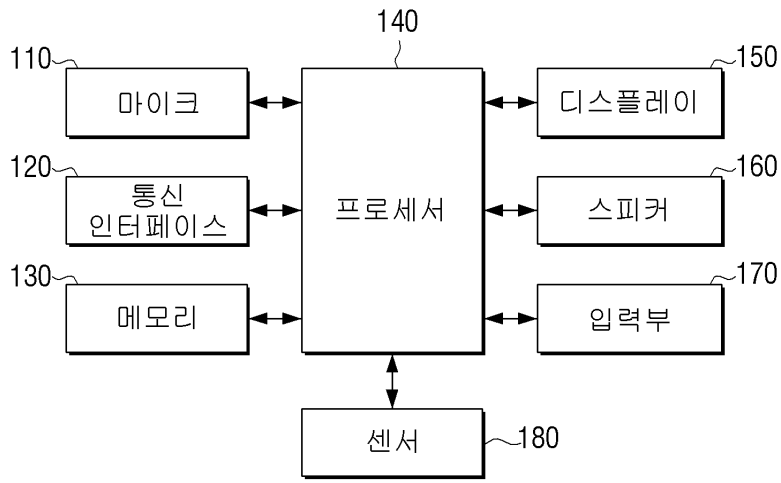
도면1



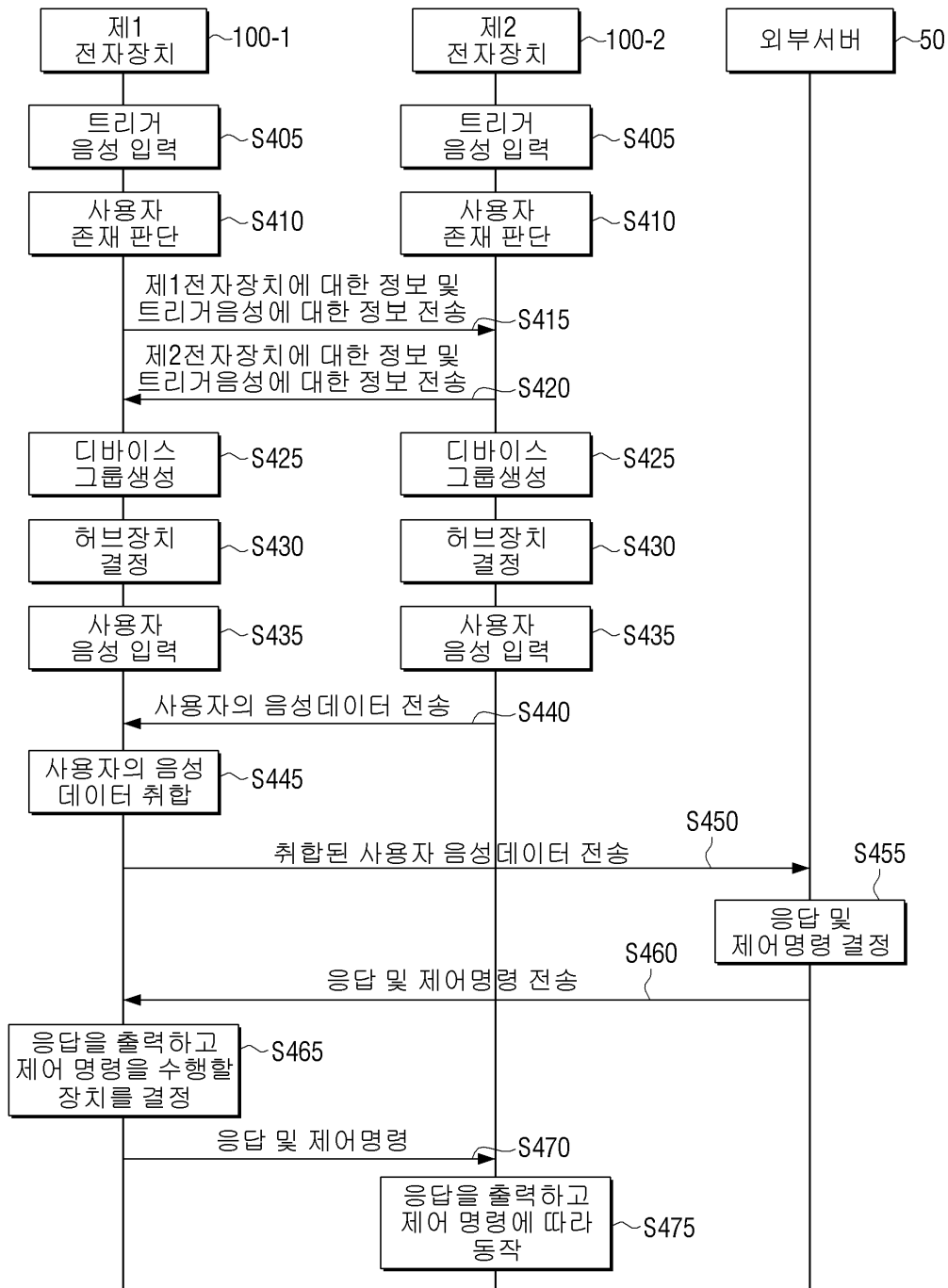
도면2



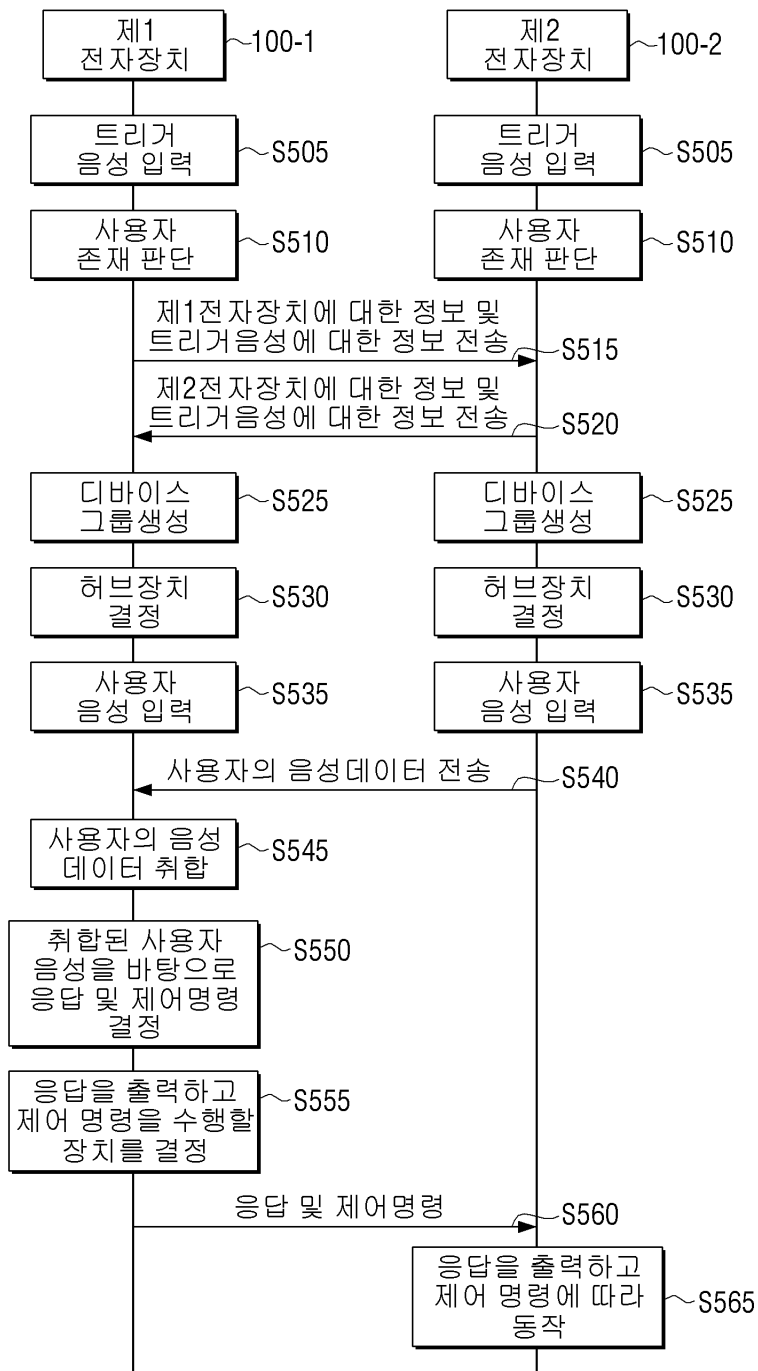
도면3



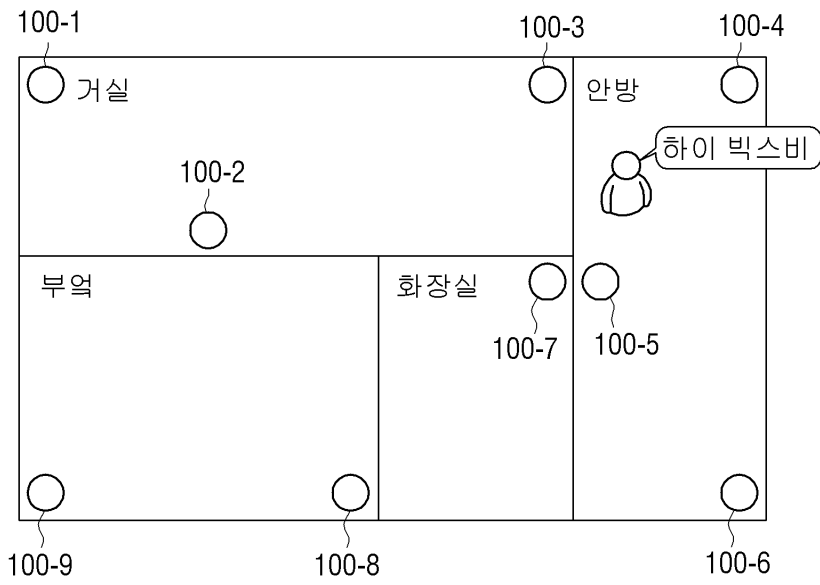
도면4



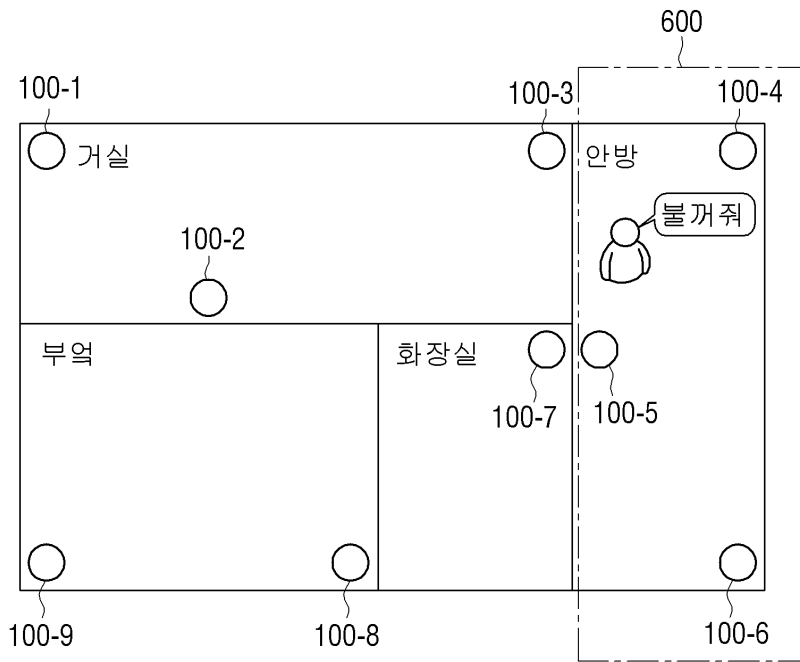
도면5



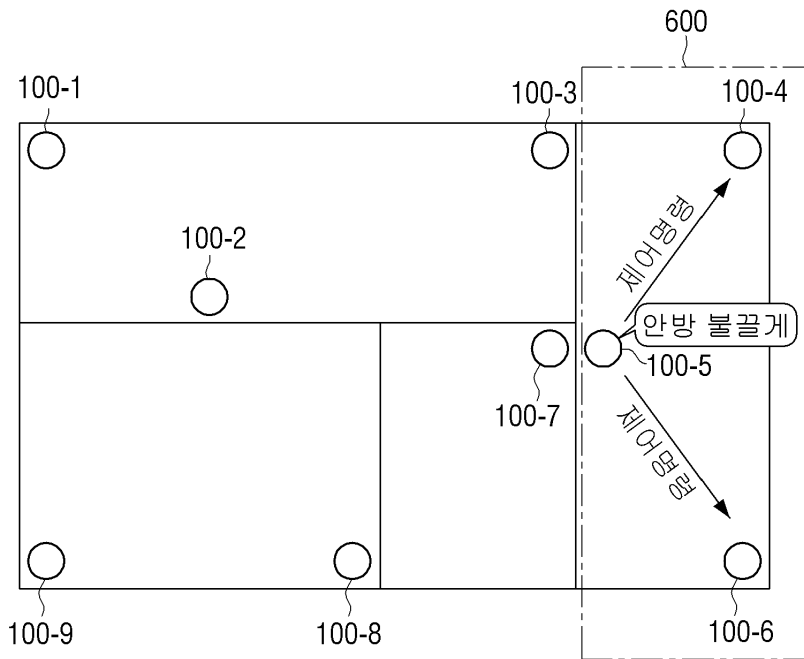
도면6a



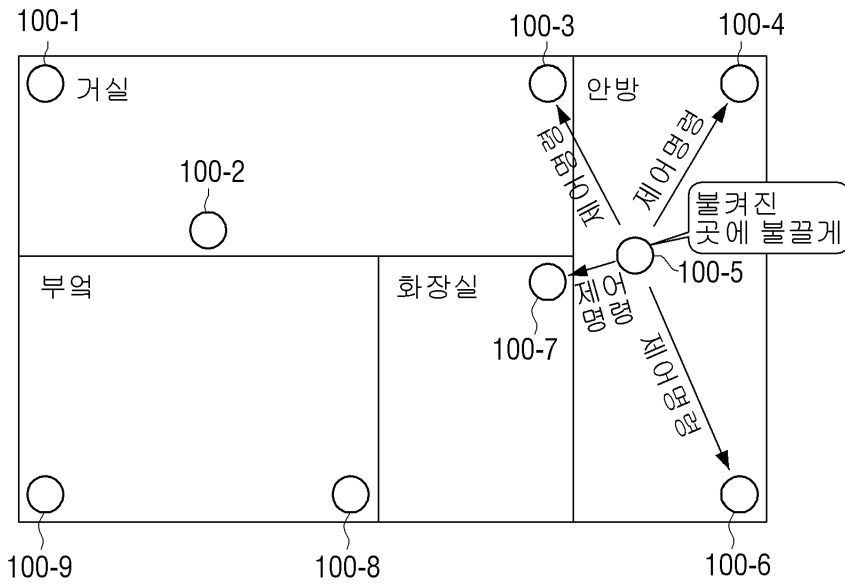
도면6b



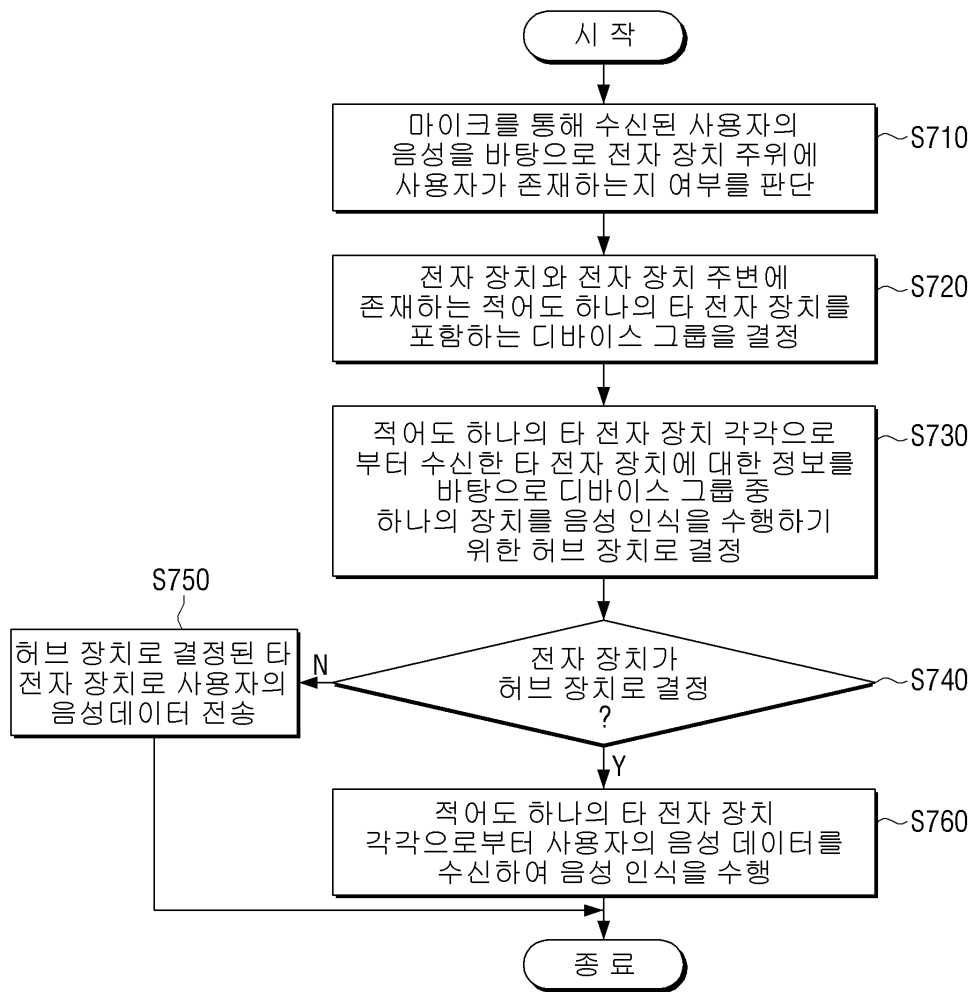
도면6c



도면6d



도면7



도면8

