



**(19) 대한민국특허청(KR)**  
**(12) 공개특허공보(A)**

(11) 공개번호 10-2019-0134366  
(43) 공개일자 2019년12월04일

(51) 국제특허분류(Int. Cl.) G10L 17/04 (2013.01) G06F 21/32 (2013.01) G10L 15/04 (2006.01) G10L 17/02 (2013.01)	(71) 출원인 삼성전자주식회사 경기도 수원시 영통구 삼성로 129 (매탄동)
(52) CPC특허분류 G10L 17/04 (2013.01) G06F 21/32 (2013.01)	(72) 발명자 최찬희 서울특별시 동작구 장승배기로4길 9, 119동 604호(상도동, 상도 더샵 아파트)
(21) 출원번호 10-2018-0059901	(74) 대리인 정홍식, 김태현
(22) 출원일자 2018년05월25일 심사청구일자 없음	

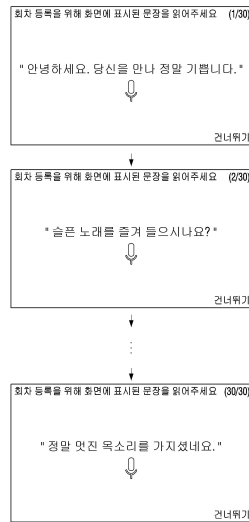
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 전자 장치, 전자 장치의 제어 방법 및 컴퓨터 판독 가능 매체

**(57) 요약**

전자 장치가 개시된다. 본 전자 장치는, 사용자 입력부, 음성 수신부, 스토리지, 상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공하고, 음성 수신부를 통해 수신된 가이드에 대응되는 음성에 기초하여 사용자 정보에 대응되는 화자 인식 모델을 생성하여 스토리지에 저장하며, 음성 수신부를 통해 수신되는 음성과 화자 인식 모델을 비교하여 업데이트된 화자 인식 모델에 기초하여, 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식하는 프로세서를 포함한다.

**대표도** - 도1



(52) CPC특허분류

*G10L 15/04* (2013.01)

*G10L 17/02* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

전자 장치에 있어서,

사용자 입력부;

음성 수신부;

스토리지; 및

상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공하고,

상기 음성 수신부를 통해 수신된 상기 가이드에 대응되는 음성에 기초하여 상기 사용자 정보에 대응되는 화자 인식 모델을 생성하여 상기 스토리지에 저장하며,

상기 음성 수신부를 통해 수신되는 음성과 상기 화자 인식 모델을 비교하여 업데이트된 화자 인식 모델에 기초하여, 상기 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식하는 프로세서;를 포함하는 전자 장치.

#### 청구항 2

제1항에 있어서,

상기 프로세서는,

상기 스토리지에 저장된 복수의 텍스트 중 일부의 텍스트를 포함하는 상기 가이드를 제공하고, 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 상기 가이드에 포함된 텍스트와 일치하는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 화자 인식 모델을 생성하는, 전자 장치.

#### 청구항 3

제1항에 있어서,

상기 프로세서는,

음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 음성이 수신되면, 상기 스토리지에 저장된 상기 화자 인식 모델과 상기 수신된 음성의 특징 값이 기설정된 임계 값 이상의 유사도를 갖는지 식별하고,

상기 기설정된 임계 값 이상의 유사도를 가지면, 상기 수신된 음성의 특징 값에 기초하여 상기 스토리지에 저장된 상기 화자 인식 모델을 업데이트하는, 전자 장치.

#### 청구항 4

제3항에 있어서,

상기 프로세서는,

상기 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 상기 스토리지에 저장된 복수의 텍스트에 포함되는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 스토리지에 저장된 화자 인식 모델을 업데이트하는, 전자 장치.

#### 청구항 5

제1항에 있어서,

상기 프로세서는,

상기 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 상기 사용자 정보에 대응되는 상기 화자 인식 모델

의 업데이트 완료 여부를 식별하고,

상기 화자 인식 모델의 업데이트가 완료되지 않은 것으로 식별되면, 상기 스토리지에 저장된 복수의 텍스트 중 사용자에게 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드를 제공하는, 전자 장치.

#### 청구항 6

제5항에 있어서,

상기 프로세서는,

상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된 음성의 특징 값에 기초하여 상기 화자 인식 모델을 업데이트하여 상기 사용자 정보에 대응되는 최종 화자 인식 모델을 생성하는, 전자 장치.

#### 청구항 7

제1항에 있어서,

상기 프로세서는,

상기 사용자 정보에 기초하여 사용자 인증이 수행되면, 상기 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스를 제공하는, 전자 장치.

#### 청구항 8

제1항에 있어서,

상기 프로세서는,

상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 상기 사용자 정보에 대응되는 화자 인식 모델이 상기 스토리지에 저장되어 있는지 식별하고,

상기 가이드는,

상기 사용자 정보에 대응되는 상기 화자 인식 모델이 상기 스토리지에 저장되어 있지 않은 경우, 상기 사용자 정보에 대응되는 상기 화자 인식 모델의 생성 여부를 선택하게 하는 유저 인터페이스를 포함하는, 전자 장치.

#### 청구항 9

제1항에 있어서,

상기 프로세서는,

상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 특정 사용자의 로그인에 최초로 수행되면, 상기 특정 사용자의 발화를 유도하기 위한 가이드를 제공하고, 상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된 음성에 기초하여 상기 특정 사용자에게 대응되는 화자 인식 모델을 생성하여 상기 스토리지에 저장하고,

상기 특정 사용자가 로그인된 상태에서 음성 인식 서비스가 제공되는 동안 상기 음성 수신부를 통해 수신되는 음성과 상기 스토리지에 저장된 상기 화자 인식 모델을 비교하여 상기 화자 인식 모델을 업데이트하는, 전자 장치.

#### 청구항 10

전자 장치의 제어 방법에 있어서,

사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공하는 단계;

음성 수신부를 통해 수신된 상기 가이드에 대응되는 음성에 기초하여 상기 사용자 정보에 대응되는 화자 인식 모델을 생성하여 저장하는 단계;

상기 음성 수신부를 통해 수신되는 음성과 상기 화자 인식 모델을 비교하여 업데이트된 화자 인식 모델에 기초하여, 상기 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식하는 단계;를 포함하는 제어 방법.

**청구항 11**

제10항에 있어서,  
 상기 가이드를 제공하는 단계는,  
 기저장된 복수의 텍스트 중 일부의 텍스트를 포함하는 상기 가이드를 제공하고,  
 상기 화자 인식 모델을 생성하여 저장하는 단계는,  
 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 상기 가이드에 포함된 텍스트와 일치하는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 화자 인식 모델을 생성하는, 제어 방법.

**청구항 12**

제10항에 있어서,  
 상기 화자 인식 모델을 업데이트하는 단계는,  
 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 음성이 수신되면, 상기 저장된 화자 인식 모델과 상기 수신된 음성의 특징 값이 기설정된 임계 값 이상의 유사도를 갖는지 식별하고,  
 상기 기설정된 임계 값 이상의 유사도를 가지면, 상기 수신된 음성의 특징 값에 기초하여 상기 저장된 화자 인식 모델을 업데이트하는, 제어 방법.

**청구항 13**

제12항에 있어서,  
 상기 화자 인식 모델을 업데이트하는 단계는,  
 상기 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 기저장된 복수의 텍스트에 포함되는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 저장된 화자 인식 모델을 업데이트하는, 제어 방법.

**청구항 14**

제10항에 있어서,  
 상기 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 상기 사용자 정보에 대응되는 상기 화자 인식 모델의 업데이트 완료 여부를 식별하는 단계; 및  
 상기 화자 인식 모델의 업데이트가 완료되지 않은 것으로 식별되면, 기저장된 복수의 텍스트 중 사용자에게 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드를 제공하는 단계;를 더 포함하는 제어 방법.

**청구항 15**

제14항에 있어서,  
 상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된 음성의 특징 값에 기초하여 상기 화자 인식 모델을 업데이트하여 상기 사용자 정보에 대응되는 최종 화자 인식 모델을 생성하는 단계;를 더 포함하는, 제어 방법.

**청구항 16**

제10항에 있어서,  
 상기 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 상기 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스를 제공하는 단계;를 더 포함하는, 제어 방법.

**청구항 17**

제10항에 있어서,

상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 상기 사용자 정보에 대응되는 화자 인식 모델이 저장되어 있는지 식별하는 단계;를 더 포함하고,

상기 가이드는,

상기 사용자 정보에 대응되는 상기 화자 인식 모델이 저장되어 있지 않은 경우, 상기 사용자 정보에 대응되는 상기 화자 인식 모델의 생성 여부를 선택하게 하는 유저 인터페이스를 포함하는, 제어 방법.

### 청구항 18

제10항에 있어서,

상기 가이드를 제공하는 단계는,

상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 특정 사용자의 로그인에 최초로 수행되면, 상기 특정 사용자의 발화를 유도하기 위한 가이드를 제공하고,

상기 화자 인식 모델을 생성하여 저장하는 단계는,

상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된 음성에 기초하여 상기 특정 사용자에 대응되는 화자 인식 모델을 생성하여 저장하고,

상기 화자 인식 모델을 업데이트하는 단계는,

상기 특정 사용자가 로그인된 상태에서 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 수신되는 음성과 상기 저장된 상기 화자 인식 모델을 비교하여 상기 화자 인식 모델을 업데이트하는, 제어 방법.

### 청구항 19

전자 장치의 프로세서에 의해 실행되어, 상기 전자 장치로 하여금,

사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공하는 단계;

음성 수신부를 통해 수신된 상기 가이드에 대응되는 음성에 기초하여 상기 사용자 정보에 대응되는 화자 인식 모델을 생성하여 저장하는 단계;

상기 음성 수신부를 통해 수신되는 음성과 상기 화자 인식 모델을 비교하여 업데이트된 화자 인식 모델에 기초하여, 상기 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식하는 단계;를 포함하는 동작을 수행하도록 하는 컴퓨터 명령어가 저장된 비일시적 컴퓨터 판독 가능 매체.

## 발명의 설명

### 기술 분야

[0001] 본 개시는 전자 장치, 전자 장치의 제어 방법 및 컴퓨터 판독 가능 매체에 관한 것으로, 보다 상세하게는 문장 독립 방식으로 화자를 인식하기 위한 시스템을 구축하기 위한 기술에 관한 것이다.

### 배경 기술

[0002] 화자 인식은 입력받은 음성 데이터를 미리 저장된 데이터와 비교하여 화자가 누구인지 식별하는 기술을 의미한다.

[0003] 화자 인식에는 사용자를 위한 화자 인식 모델을 등록하는 과정이 필요하다. 이는, 현재 입력되는 음성을 발화한 화자가 누구인지를 시스템에 알리고, 입력되는 음성에서 추출된 특징에 기초하여 입력된 음성을 발화한 사용자를 위한 화자 인식 모델을 등록하는 것이다.

[0004] 화자 인식의 방법에 따라, 화자 인식 모델을 등록하는 과정 역시 두 가지로 나뉠 수 있다. 먼저, 지정된 문장이나 단어만을 발화하게 하고 이를 기초로 사용자의 음성 특징을 추출하여 화자 인식 모델을 등록하는 문장 종속 방식이 있다. 그리고, 사용자의 자유 발화에 대해서 화자 인식을 할 수 있도록 하기 위해, 충분한 분량의 문장과 단어를 발화하게 한 뒤, 이를 기초로 사용자의 음성 특징을 추출하여 화자 인식 모델을 등록하는 문장 독립

방식도 있다.

- [0005] 문장 독립 방식은 문장 종속 방식에 비해 더욱 방대한 양의 데이터에 기반하여, 사용자가 어떤 음성을 발화하든 해당 음성에 대한 화자 인식이 가능하다는 장점이 있다.
- [0006] 다만, 종래의 문장 독립 방식은, 화자 인식 모델을 등록하는 과정에 있어 사용자가 발화해야 하는 단어 및 문장의 양이 너무 많기 때문에, 사용자의 불편을 초래할 수 있다는 문제점이 있었다.
- [0007] 특히, 종래의 문장 독립 방식은, 사용자의 발화를 녹음하는 시간이 길어질수록 사용자의 피로도, 발화 인식의 정확도, 주변의 잡음 등에 기초한 변수가 증가하여, 등록된 화자 인식 모델의 품질이 저하될 가능성이 있었다.
- [0008] 도 1은 문장 독립 방식의 화자 인식을 위한 화자 인식 모델 등록에 있어, 종래의 기술 중 일부를 설명하기 위한 도면이다. 도 1은, 화자 인식 모델을 등록하는 과정에서 디스플레이되는 이미지를 순차적으로 도시한 것으로, 각 이미지들은 사용자에게 특정 문장에 발화를 요청한다.
- [0009] 도 1을 참조하면, 종래의 화자 인식 모델 등록에 있어, 사용자는 꽤 많은 분량의 문장에 대해서 순차적으로 발화할 것을 요청받는다. 일 예로, 도 1과 같이 30 가지 텍스트에 대한 순차적 발화 요청에 따라 사용자가 발화하고 또 발화된 음성을 수신하다보면, 연속된 작업에 사용자가 불편을 느낄 수 있다.
- [0010] 또한, 지친 사용자가 발화하는 음성이 정확하지 않음으로 인해서 최종적으로 등록되는 화자 인식 모델의 정확도도 떨어질 수 있다.

## 발명의 내용

### 해결하려는 과제

- [0011] 본 개시의 목적은, 사용자의 인증에 기초하여 사용자가 발화해야 할 내용을 분할하여 요청함은 물론, 사용자가 일반적인 음성 인식 서비스를 제공받는 동안에도 사용자의 음성을 통해 화자 인식 모델을 꾸준히 업데이트하여, 문장 독립 방식의 화자 인식 모델 등록에 있어 사용자의 피로도를 줄이고 그 품질을 높이기 위한 전자 장치, 전자 장치의 제어 방법 및 컴퓨터 판독 가능 매체를 제공함에 있다.

### 과제의 해결 수단

- [0012] 본 개시의 일 실시 예에 따른 전자 장치는, 사용자 입력부, 음성 수신부, 스토리지, 상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공하고, 상기 음성 수신부를 통해 수신된 상기 가이드에 대응되는 음성에 기초하여 상기 사용자 정보에 대응되는 화자 인식 모델을 생성하여 상기 스토리지에 저장하고, 상기 음성 수신부를 통해 수신되는 음성과 상기 화자 인식 모델을 비교하여 업데이트된 상기 화자 인식 모델에 기초하여, 상기 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식하는 프로세서를 포함한다.
- [0013] 이때, 상기 프로세서는, 상기 스토리지에 저장된 복수의 텍스트 중 일부의 텍스트를 포함하는 상기 가이드를 제공하고, 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 상기 가이드에 포함된 텍스트와 일치하는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 화자 인식 모델을 생성할 수 있다.
- [0014] 그리고, 상기 프로세서는, 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 음성이 수신되면, 상기 스토리지에 저장된 상기 화자 인식 모델과 상기 수신된 음성의 특징 값이 기설정된 임계 값 이상의 유사도를 갖는지 식별하고, 상기 기설정된 임계 값 이상의 유사도를 가지면, 상기 수신된 음성의 특징 값에 기초하여 상기 스토리지에 저장된 상기 화자 인식 모델을 업데이트할 수 있다.
- [0015] 이 경우, 상기 프로세서는, 상기 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 상기 스토리지에 저장된 복수의 텍스트에 포함되는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 스토리지에 저장된 화자 인식 모델을 업데이트할 수 있다.
- [0016] 한편, 상기 프로세서는, 상기 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 상기 사용자 정보에 대응되는 상기 화자 인식 모델의 업데이트 완료 여부를 식별하고, 상기 화자 인식 모델의 업데이트가 완료되지 않은 것으로 식별되면, 상기 스토리지에 저장된 복수의 텍스트 중 사용자에게 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드를 제공할 수 있다.
- [0017] 이때, 상기 프로세서는, 상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된 음성의 특

징 값에 기초하여 상기 화자 인식 모델을 업데이트하여 상기 사용자 정보에 대응되는 최종 화자 인식 모델을 생성할 수 있다.

- [0018] 상기 프로세서는, 상기 사용자 정보에 기초하여 사용자 인증이 수행되면, 상기 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스를 제공할 수도 있다.
- [0019] 상기 프로세서는, 상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 상기 사용자 정보에 대응되는 화자 인식 모델이 상기 스토리지에 저장되어 있는지 식별하고, 상기 가이드는, 상기 사용자 정보에 대응되는 상기 화자 인식 모델이 상기 스토리지에 저장되어 있지 않은 경우, 상기 사용자 정보에 대응되는 상기 화자 인식 모델의 생성 여부를 선택하게 하는 유저 인터페이스를 포함할 수도 있다.
- [0020] 상기 프로세서는, 상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 특정 사용자의 로그인이 최초로 수행되면, 상기 특정 사용자의 발화를 유도하기 위한 가이드를 제공하고, 상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된 음성에 기초하여 상기 특정 사용자에 대응되는 화자 인식 모델을 생성하여 상기 스토리지에 저장하고, 상기 특정 사용자가 로그인된 상태에서 음성 인식 서비스가 제공되는 동안 상기 음성 수신부를 통해 수신되는 음성과 상기 스토리지에 저장된 상기 화자 인식 모델을 비교하여 상기 화자 인식 모델을 업데이트할 수도 있다.
- [0021] 한편, 본 개시의 일 실시 예에 따른 전자 장치의 제어 방법은, 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공하는 단계, 음성 수신부를 통해 수신된 상기 가이드에 대응되는 음성에 기초하여 상기 사용자 정보에 대응되는 화자 인식 모델을 생성하여 저장하는 단계, 상기 음성 수신부를 통해 수신되는 음성과 상기 화자 인식 모델을 비교하여 업데이트된 화자 인식 모델에 기초하여, 상기 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식하는 단계를 포함한다.
- [0022] 이 경우, 상기 가이드를 제공하는 단계는, 기저장된 복수의 텍스트 중 일부의 텍스트를 포함하는 상기 가이드를 제공하고, 상기 화자 인식 모델을 생성하여 저장하는 단계는, 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 상기 가이드에 포함된 텍스트와 일치하는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 화자 인식 모델을 생성할 수 있다.
- [0023] 그리고, 상기 화자 인식 모델을 업데이트하는 단계는, 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 음성이 수신되면, 상기 저장된 화자 인식 모델과 상기 수신된 음성의 특징 값이 기설정된 임계 값 이상의 유사도를 갖는지 식별하고, 상기 기설정된 임계 값 이상의 유사도를 가지면, 상기 수신된 음성의 특징 값에 기초하여 상기 저장된 화자 인식 모델을 업데이트할 수 있다.
- [0024] 이때, 상기 화자 인식 모델을 업데이트하는 단계는, 상기 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 기저장된 복수의 텍스트에 포함되는 경우, 상기 수신된 음성의 특징 값에 기초하여 상기 저장된 화자 인식 모델을 업데이트할 수 있다.
- [0025] 한편, 상기 제어 방법은, 상기 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 상기 사용자 정보에 대응되는 상기 화자 인식 모델의 업데이트 완료 여부를 식별하는 단계, 상기 화자 인식 모델의 업데이트가 완료되지 않은 것으로 식별되면, 기저장된 복수의 텍스트 중 사용자에 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드를 제공하는 단계를 더 포함할 수 있다.
- [0026] 이때, 상기 제어 방법은, 상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된 음성의 특징 값에 기초하여 상기 화자 인식 모델을 업데이트하여 상기 사용자 정보에 대응되는 최종 화자 인식 모델을 생성하는 단계를 더 포함할 수도 있다.
- [0027] 상기 제어 방법은, 상기 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 상기 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스를 제공하는 단계를 더 포함할 수도 있다.
- [0028] 상기 제어 방법은, 상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 상기 사용자 정보에 대응되는 화자 인식 모델이 저장되어 있는지 식별하는 단계를 더 포함하고, 상기 가이드는, 상기 사용자 정보에 대응되는 상기 화자 인식 모델이 저장되어 있지 않은 경우, 상기 사용자 정보에 대응되는 상기 화자 인식 모델의 생성 여부를 선택하게 하는 유저 인터페이스를 포함할 수도 있다.
- [0029] 한편, 상기 가이드를 제공하는 단계는, 상기 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 특정 사용자의 로그인이 최초로 수행되면, 상기 특정 사용자의 발화를 유도하기 위한 가이드를 제공하고, 상기 화자 인식 모델을 생성하여 저장하는 단계는, 상기 가이드에 따른 음성이 상기 음성 수신부를 통해 수신되면, 상기 수신된

음성에 기초하여 상기 특정 사용자에게 대응되는 화자 인식 모델을 생성하여 저장하고, 상기 화자 인식 모델을 업데이트하는 단계는, 상기 특정 사용자가 로그인된 상태에서 음성 인식 서비스를 제공하는 동안 상기 음성 수신부를 통해 수신되는 음성과 상기 저장된 상기 화자 인식 모델을 비교하여 상기 화자 인식 모델을 업데이트할 수 있다.

[0030] 한편, 본 개시의 일 실시 예에 따른 비밀시적 컴퓨터 판독 가능 매체에는, 전자 장치의 프로세서에 의해 실행되어, 상기 전자 장치로 하여금, 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공하는 단계, 음성 수신부를 통해 수신된 상기 가이드에 대응되는 음성에 기초하여 상기 사용자 정보에 대응되는 화자 인식 모델을 생성하여 저장하는 단계, 상기 음성 수신부를 통해 수신되는 음성과 상기 화자 인식 모델을 비교하여 업데이트된 화자 인식 모델에 기초하여, 상기 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식하는 단계를 포함하는 동작을 수행하도록 하는 컴퓨터 명령어가 저장되어 있다.

**발명의 효과**

[0031] 본 개시에 따른 전자 장치, 전자 장치의 제어 방법 및 컴퓨터 판독 가능 매체는, 문장 독립 방식의 화자 인식을 위한 화자 인식 모델을 등록함에 있어서, 기존에 많은 양의 음성을 발화해야 했던 사용자의 불편을 해소하고, 보다 정확한 화자 인식 모델을 제공할 수 있다는 효과가 있다.

[0032] 특히, 특정한 단어나 문장을 발화함으로써 특정 사용자가 인증된 상태에서 자유 발화에 대한 음성 인식에 진입하는 경우와 달리, 별도의 과정 없이 자유 발화에 대한 음성 인식이 시작되거나 또는 단순히 버튼을 누르는 동작만으로 자유 발화에 대한 음성 인식이 시작되는 경우, 아직 자유 발화의 주체가 누구인지 파악이 되지 않은 상태에서 사용자를 인식하고 인식된 사용자에게 대응되는 음성을 인식하기 위해서는, 각각의 사용자별로 많은 양의 음성이 발화되어 학습될 필요가 있으므로, 본 개시에 따른 화자 인식 모델 제공의 효과가 더욱 두드러질 수 있다.

**도면의 간단한 설명**

- [0033] 도 1은 종래의 화자 인식 모델 등록 과정의 문제점을 설명하기 위한 도면,
- 도 2는 본 개시의 일 실시 예에 따른 전자 장치의 구성을 설명하기 위한 블록도,
- 도 3은 전자 장치 및 외부 장치의 구성에 기초하여 전자 장치의 동작을 설명하기 위한 도면,
- 도 4는 사용자의 발화를 유도하는 가이드의 일부에 대한 예를 설명하기 위한 도면,
- 도 5는 화자 인식 모델을 완성하기 위해 음성 형태로 입력받아야 하는 복수의 텍스트가 있는 경우, 전자 장치의 동작을 설명하기 위한 도면,
- 도 6은 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스의 일 예를 설명하기 위한 도면,
- 도 7은 본 개시의 다양한 실시 예에 따른 전자 장치의 상세한 구성을 설명하기 위한 블록도,
- 도 8은 본 개시의 일 실시 예에 따른 전자 장치의 제어 방법을 설명하기 위한 순서도,
- 도 9는 기초가 되는 초기 화자 인식 모델을 생성하는 과정을 설명하기 위한 알고리즘,
- 도 10은 사용자에게 대한 발화 요청 없이 화자 인식 모델을 업데이트하는 과정을 설명하기 위한 알고리즘, 그리고
- 도 11은 사용자에게 대한 발화 요청을 통해 화자 인식 모델을 업데이트하는 과정을 설명하기 위한 알고리즘을 도시한다.

**발명을 실시하기 위한 구체적인 내용**

[0034] 본 개시에 대하여 구체적으로 설명하기에 앞서, 본 명세서 및 도면의 기재 방법에 대하여 설명한다.

[0035] 먼저, 본 명세서 및 청구범위에서 사용되는 용어는 본 개시의 다양한 실시 예들에서의 기능을 고려하여 일반적인 용어들을 선택하였다 하지만, 이러한 용어들은 당해 기술 분야에 종사하는 기술자의 의도나 법률적 또는 기술적 해석 및 새로운 기술의 출현 등에 따라 달라질 수 있다. 또한, 일부 용어는 출원인이 임의로 선정한 용어도 있다. 이러한 용어에 대해서는 본 명세서에서 정의된 의미로 해석될 수 있으며, 구체적인 용어 정의가 없으

면 본 명세서의 전반적인 내용 및 당해 기술 분야의 통상적인 기술 상식을 토대로 해석될 수도 있다.

- [0036] 또한, 본 명세서에 첨부된 각 도면에 기재된 동일한 참조번호 또는 부호는 실질적으로 동일한 기능을 수행하는 부품 또는 구성요소를 나타낸다. 설명 및 이해의 편의를 위해서 서로 다른 실시 예들에서도 동일한 참조번호 또는 부호를 사용하여 설명한다. 즉, 복수의 도면에서 동일한 참조 번호를 가지는 구성요소를 모두 도시되어 있다고 하더라도, 복수의 도면들이 하나의 실시 예를 의미하는 것은 아니다.
- [0037] 또한, 본 명세서 및 청구범위에서는 구성요소들 간의 구별을 위하여 "제1", "제2" 등과 같이 서수를 포함하는 용어가 사용될 수 있다. 이러한 서수는 동일 또는 유사한 구성요소들을 서로 구별하기 위하여 사용하는 것이며 이러한 서수 사용으로 인하여 용어의 의미가 한정 해석되어서는 안 된다. 일 예로, 이러한 서수와 결합된 구성 요소는 그 숫자에 의해 사용 순서나 배치 순서 등이 제한되어서는 안 된다. 필요에 따라서는, 각 서수들은 서로 교체되어 사용될 수도 있다.
- [0038] 본 명세서에서 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "구성되다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0039] 본 개시의 실시 예에서 "모듈", "유닛", "부(part)" 등과 같은 용어는 적어도 하나의 기능이나 동작을 수행하는 구성요소를 지칭하기 위한 용어이며, 이러한 구성요소는 하드웨어 또는 소프트웨어로 구현되거나 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다. 또한, 복수의 "모듈", "유닛", "부(part)" 등은 각각이 개별적인 특정한 하드웨어로 구현될 필요가 있는 경우를 제외하고는, 적어도 하나의 모듈이나 칩으로 일체화되어 적어도 하나의 프로세서로 구현될 수 있다.
- [0040] 또한, 본 개시의 실시 예에서, 어떤 부분이 다른 부분과 연결되어 있다고 할 때, 이는 직접적인 연결뿐 아니라, 다른 매체를 통한 간접적인 연결의 경우도 포함한다. 또한, 어떤 부분이 어떤 구성요소를 포함한다는 의미는, 특별히 반대되는 기재가 없는 한 다른 구성요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0041] 도 2는 전자 장치(100), 디스플레이가 구비된 외부 장치(200) 및 원격 제어 장치(300)를 도시한다. 전자 장치(100)는 TV, 스마트 TV, 셋탑박스, PC, 서버, 단말 장치, 휴대폰, 태블릿 PC 등으로 구현될 수 있다.
- [0042] 외부 장치(200)는 디스플레이를 구비한 TV, 스마트 TV, PC 등으로 구현될 수 있고, 원격 제어 장치(300)는 TV용 리모콘으로 구현되거나 또는 TV를 제어하기 위한 어플리케이션이 저장된 단말 장치의 형태로 구현될 수도 있다.
- [0043] 도 2의 (a)는, 전자 장치(100)가 디스플레이를 구비하지 않은 셋탑박스 등으로 구현된 경우를 도시한다.
- [0044] 이 경우, 전자 장치(100) 및 외부 장치(200)는 서로 신호를 주고받을 수 있다. 예로, 전자 장치(100)는 외부 장치(200)가 이미지를 디스플레이할 수 있도록 이미지 신호를 외부 장치(200)로 전송할 수 있다. 또한, 전자 장치(100)는 외부 장치(200)가 음성을 출력할 수 있도록 음성에 대한 신호를 외부 장치(200)로 전송할 수도 있다.
- [0045] 한편, 원격 제어 장치(300)는 사용자의 명령에 기초하여 전자 장치(100) 및 외부 장치(200) 중 적어도 하나를 제어하는 신호를 출력할 수 있다. 또한, 원격 제어 장치(300)는 사용자 정보를 포함하는 신호를 전자 장치(100) 및 외부 장치(200) 중 적어도 하나에 전송할 수도 있다.
- [0046] 도 2의 (b)는 전자 장치(100)가 디스플레이를 구비한 TV 또는 스마트 TV 등으로 구현된 경우를 도시한다.
- [0047] 이 경우, 전자 장치(100)는 원격 제어 장치(100)로부터 수신된 제어 신호에 기초하여 이미지를 디스플레이하거나 및/또는 음성을 출력할 수 있다. 또한, 전자 장치(100)는 원격 제어 장치(300)로부터 사용자 정보를 포함하는 신호를 수신할 수도 있다.
- [0048] 도 3은 본 개시의 일 실시 예에 따른 전자 장치(100)의 구성을 설명하기 위한 블록도이다. 도 3을 참조하면, 전자 장치(100)는 사용자 입력부(110), 음성 수신부(120), 스토리지(130) 및 프로세서(140)를 포함할 수 있다.
- [0049] 사용자 입력부(110)는 전자 장치(100)가 사용자로부터 명령 또는 정보를 받을 수 있게 하는 구성이다.
- [0050] 사용자 입력부(110)는 마이크(도시되지 않음)를 포함하는 한편 이를 이용하여 사용자가 발화한 사용자의 음성을 입력받을 수 있다. 이때, 사용자 입력부(110)는 아날로그 형태로 입력된 사용자의 음성을 디지털 형태로 변환한 뒤, 음성 수신부(120) 및 프로세서(140) 중 적어도 하나에 전송할 수 있다.

- [0051] 사용자 입력부(110)는 사용자의 명령 또는 정보를 터치로 입력받기 위해, 디스플레이(도시되지 않음)와 함께 구현된 터치 패널(도시되지 않음) 또는 별도의 터치 패드(도시되지 않음)를 포함할 수도 있다.
- [0052] 사용자 입력부(110)는 사용자가 입력한 텍스트에 대한 정보를 입력받을 수 있다. 예로, 사용자 입력부(110)는 키보드나 마우스(도시되지 않음)를 포함하는 한편 이를 이용하여 사용자로부터 특정한 텍스트를 입력받을 수 있다.
- [0053] 사용자 입력부(110)는 사용자의 음성 외에 사용자의 생체 정보에 관련된 정보를 입력받을 수 있다. 이를 위해, 사용자 입력부(110)는 사용자의 지문을 인식하기 위한 센서, 사용자의 얼굴을 인식하기 위한 카메라, 사용자의 홍채를 인식하기 위한 센서 등을 포함할 수 있다.
- [0054] 사용자 입력부(110)는, 사용자의 명령 또는 정보가 포함된 신호가 원격 제어 장치(300)로부터 전자 장치(100)의 통신부(도시되지 않음)로 전송되면, 통신부(도시되지 않음)로부터 사용자의 명령 또는 정보가 포함된 신호를 수신할 수 있다.
- [0055] 예로, 원격 제어 장치(300)는 사용자로부터 텍스트 입력을 위한 명령을 수신받으면 수신된 신호를 전자 장치(100)의 통신부(도시되지 않음)로 전송할 수 있다. 이 경우, 사용자 입력부(110)는 통신부(도시되지 않음)로부터 사용자의 명령을 포함하는 신호를 수신받는 형태로 텍스트 입력을 위한 사용자의 명령을 수신할 수 있다.
- [0056] 또는, 원격 제어 장치(300)는 지문 인식용 센서(도시되지 않음)를 이용하여 사용자로부터 지문 정보를 입력받고, 입력된 지문 정보를 포함하는 신호를 전자 장치(100)의 통신부(도시되지 않음)를 통해 사용자 입력부(110)로 전송할 수도 있다.
- [0057] 또는, 원격 제어 장치(300)는 홍채 인식용 센서(도시되지 않음)를 이용하여 사용자로부터 홍채 정보를 입력받고, 입력된 홍채 정보를 포함하는 신호를 전자 장치(100)의 통신부(도시되지 않음)를 통해 사용자 입력부(110)로 전송할 수도 있다.
- [0058] 또는, 원격 제어 장치(300)는 얼굴 인식을 위한 카메라(도시되지 않음)를 이용하여 사용자의 얼굴을 촬영하고, 촬영된 얼굴 이미지를 포함하는 신호를 전자 장치(100)의 통신부(도시되지 않음)를 통해 사용자 입력부(110)로 전송할 수도 있다.
- [0059] 음성 수신부(120)는 사용자의 음성을 수신받기 위한 구성이다.
- [0060] 음성 수신부(120)는 사용자 입력부(110)에 구비된 마이크(도시되지 않음)를 통해 입력된 사용자의 음성을 수신할 수 있다. 구체적으로는, 사용자 입력부(110)에 구비된 마이크(도시되지 않음)를 통해 아날로그 형태로 수신되었다가 디지털 형태로 변환된 사용자의 음성 신호를 수신할 수 있다.
- [0061] 음성 수신부(120)는 원격 제어 장치(300) 또는 기타 외부 장치(도시되지 않음)에 포함된 마이크(도시되지 않음)를 통해 입력된 음성 신호를, 원격 제어 장치(300) 또는 기타 외부 장치(도시되지 않음)로부터 수신할 수도 있다.
- [0062] 스토리지(130)는 전자 장치(100)의 구성요소들의 전반적인 동작을 제어하기 위한 운영체제(OS: Operating System) 및 전자 장치(100)의 구성요소와 관련된 명령 또는 데이터를 저장하기 위한 구성이다.
- [0063] 이를 위해, 스토리지(130)는 비휘발성 메모리(ex: 하드 디스크, SSD(Solid state drive), 플래시 메모리), 휘발성 메모리 등으로 구현될 수 있다.
- [0064] 스토리지(130)에는 사용자별 화자 인식 모델에 대한 데이터가 저장되어 있을 수 있다. 이때, 사용자별 화자 인식 모델이란, 각 사용자가 발화한 음성의 특징을 나타내는 하나 이상의 파라미터들을 사용자별로 모델링한 것이다.
- [0065] 음성의 특징은 Pitch, Formant, LPCC(Linear Predictive Cepstral Coefficient), MFCC(Mel-Frequency Cepstral Coefficient), PLP(Perceptual Linear Predictive) 등에 기초한 것일 수 있다.
- [0066] 스토리지(130)에는 사용자에 대한 화자 인식 모델을 생성, 업데이트 및 최종적으로 완성하기 위하여 사용자로부터 수신받아야 할 음성에 대한 정보가 저장되어 있을 수 있다.
- [0067] 구체적으로, 스토리지(130)에는 사용자별 화자 인식 모델을 생성 및/또는 업데이트하기 위해 사용자별로 수신받아야 할 음성의 시간 길이, 수신받아야 할 음성에 대응되는 텍스트의 분량에 대한 정보가 저장되어 있을 수 있다.

- [0068] 또는, 스토리지(130)에는 사용자별로 화자 인식 모델을 생성 및/또는 업데이트하기 위해 사용자별로 수신받아야 할 음성에 대응되는 복수의 텍스트들에 대한 정보가 저장되어 있을 수 있다. 즉, 스토리지(130)에는 화자 인식 모델을 생성 및/또는 업데이트하기 위해 어떤 텍스트들에 대응되는 음성을 수신받아야 하는지에 대한 정보가 저장되어 있을 수 있다.
- [0069] 예로, 스토리지(130)에는 사용자별 화자 인식 모델을 생성하기 위해 사용자별로 수신받아야 할 음성에 대응되는 기초 텍스트에 대한 정보, 사용자별 화자 인식 모델을 업데이트 내지는 완성하기 위해 사용자별로 수신받아야 할 음성에 대응되는 복수 텍스트에 대한 정보가 저장되어 있을 수 있다.
- [0070] 사용자별 화자 인식 모델을 생성 및/또는 업데이트하기 위해 사용자별로 수신받아야 할 음성에 대응되는 복수의 텍스트에 대한 설명은, 도 6에서 함께 다루어질 것이다.
- [0071] 프로세서(140)는 전자 장치(100)의 전반적인 동작을 제어할 수 있다.
- [0072] 이를 위해, 프로세서(140)는 RAM(Random Access Memory)(도시되지 않음), ROM(Read Only Memory)(도시되지 않음), CPU(central processing unit)(도시되지 않음), GPU(Graphic processing unit)(도시되지 않음) 및 시스템 버스(도시되지 않음) 등을 포함할 수 있으며, 전자 장치(100)에 포함된 다른 구성요소들의 제어에 관한 연산이나 데이터 처리를 실행할 수 있다.
- [0073] 프로세서(140)는 사용자 입력부(110)를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면 사용자 발화를 유도하기 위한 가이드를 제공할 수 있다.
- [0074] 사용자 입력부(110)를 통해 입력된 사용자 정보는 사용자의 아이디 및/또는 비밀번호, 사용자의 지문 정보, 사용자의 얼굴 인식 정보, 사용자의 홍채 정보, 전자 장치(100)의 디스플레이(도시되지 않음)를 통해 디스플레이된 패턴 이미지에 대한 사용자의 터치 정보 등을 포함할 수 있다.
- [0075] 사용자 인증은, 스토리지(130)에 기저장된 사용자 정보를 사용자 입력부(110)를 통해 입력된 사용자 정보와 비교한 결과에 기초하여 수행될 수 있다. 예로, 전자 장치(100)에 대한 사용자의 접속 내지는 로그인, 전자 장치(100)를 통해 제공되는 어플리케이션 내지는 서비스에 대한 사용자의 접속 내지는 로그인 등이 사용자 인증에 포함될 수 있다.
- [0076] 프로세서(140)는, 사용자 입력부(110)를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면, 사용자 정보에 대응되는 화자 인식 모델이 스토리지(110)에 저장되어 있는지 식별할 수 있다. 식별 결과, 아직 입력된 사용자 정보에 대응되는 화자 인식 모델이 저장되어 있지 않은 경우, 사용자 발화를 유도하기 위한 가이드에는 사용자 정보에 대응되는 화자 인식 모델의 생성 여부를 사용자가 선택하게 하는 유저 인터페이스(User Interface)가 포함될 수 있다.
- [0077] 사용자 정보에 대응되는 화자 인식 모델의 생성 여부를 선택하게 하는 유저 인터페이스는, 전자 장치(100)를 통해 디스플레이되는 이미지, 전자 장치(100)를 통해 출력되는 음성, 전자 장치(100)의 사용자 입력부(110)에 대한 사용자의 선택 명령을 포함할 수 있다.
- [0078] 화자 인식 모델을 생성하라는 명령이 유저 인터페이스에 기초하여 사용자 입력부(110)를 통해 수신되면, 프로세서(120)는, 스토리지(130)에 저장된 복수의 텍스트 중 일부의 텍스트가 포함된 형태로 사용자 발화를 유도하기 위한 가이드를 제공할 수 있다.
- [0079] 이때, 스토리지(130)에 저장된 복수의 텍스트는 사용자별 화자 인식 모델을 생성 및 업데이트하기 위해 사용자별로 수신받아야 할 음성에 대응되는 복수의 텍스트일 수 있다.
- [0080] 이 경우, 복수의 텍스트 중 일부의 텍스트는, 사용자별 화자 인식 모델을 생성하기 위해 사용자별로 수신받아야 할 음성에 대응되는 텍스트일 수 있다.
- [0081] 예로, 도 4는 사용자의 발화를 유도하는 가이드의 일부에 대한 예를 설명하기 위한 도면이다. 도 4의 화면은, 사용자의 발화를 유도하는 가이드의 일부로서 전자 장치(100)를 통해 디스플레이되거나 또는 전자 장치(100)로부터 이미지 신호를 수신한 외부 장치(200)를 통해 디스플레이될 수 있다.
- [0082] 도 4에서, "화자 인식을 위한 문장입니다."라는 텍스트는, 사용자 인증을 완료한 사용자가, 사용자 본인의 사용자 정보와 대응되는 화자 인식 모델을 최초로 생성하기 위해 발화해야 하는 텍스트이다.
- [0083] 도 4의 화면이 디스플레이된 상태에서 음성 수신부(120)에 사용자의 음성이 수신되면, 프로세서(140)는 수신된

음성에 기초하여 사용자 정보에 대응되는 화자 인식 모델을 생성하여 스토리지(130)에 저장할 수 있다.

- [0084] 구체적으로, 프로세서(140)는 수신된 음성에 대응되는 텍스트가 가이드에 포함된 텍스트(예를 들어, 도 4의 "화자 인식을 위한 문장입니다.")와 일치하는 경우, 수신된 음성의 특징 값에 기초하여 화자 인식 모델을 생성할 수 있다. 이때, 생성된 화자 인식 모델은, 입력된 사용자 정보와 대응되는 것으로 스토리지(130)에 저장될 수 있다.
- [0085] 한편, 사용자 인증이 로그인에 해당하는 경우의 예로, 프로세서(140)는 사용자 입력부(110)를 통해 입력된 사용자 정보에 기초하여 특정 사용자의 로그인이 최초로 수행되면, 특정 사용자의 발화를 유도하기 위한 가이드를 제공할 수 있다.
- [0086] 이때, 특정 사용자의 로그인은 전자 장치(100) 또는 전자 장치(100)를 통해 동작하는 어플리케이션에 대한 로그인일 수 있다.
- [0087] 그리고, 프로세서(140)는 음성 수신부(120)를 통해 수신되며 가이드에 대응되는 음성에 기초하여 사용자 정보 대응되는 화자 인식 모델을 생성 및 저장할 수 있다.
- [0088] 구체적으로, 프로세서(140)는, 가이드에 따른 음성이 음성 수신부(110)를 통해 수신되면, 수신된 음성에 기초하여 특정 사용자에게 대응되는 화자 인식 모델을 생성하여 스토리지(130)에 저장할 수 있다.
- [0089] 프로세서(140)는, 사용자 인증 수행 이후 화자 인식 모델이 생성되면, 화자 인식 모델이 어느 정도 완성되었는지 알려주기 위한 유저 인터페이스를 제공할 수도 있다. 이때, 유저 인터페이스는, 전자 장치(100) 또는 전자 장치(100)로부터 신호를 수신한 외부 장치(200)가 안내 이미지를 디스플레이하거나 또는 안내 음성을 출력하는 형태로 제공될 수 있다.
- [0090] 프로세서(140)는 음성 수신부(120)를 통해 수신되는 음성과 스토리지(130)에 저장된 화자 인식 모델을 비교하여 스토리지(130)에 저장된 상기 화자 인식 모델을 업데이트할 수 있다.
- [0091] 구체적으로, 프로세서(140)는, 음성 인식 서비스를 제공하는 동안 음성 수신부(120)를 통해 수신되는 음성을, 스토리지(130)에 저장된 화자 인식 모델과 비교하여 화자 인식 모델을 업데이트할 수 있다.
- [0092] 이때, 음성 인식 서비스는 전자 장치(100)를 통해 수행되는 어플리케이션 중 음성 인식 기능이 사용되는 어플리케이션을 통해 제공되는 서비스를 포함할 수 있다. 이 경우, 음성 인식 서비스는, 스토리지(130)에 저장된 화자 인식 모델과 대응되는 사용자 정보에 기초한 사용자 인증이 수행된 상태에서, 스토리지(130)에 저장된 어플리케이션 중 음성 인식 기능을 수행하는 어플리케이션에 의해 제공되는 서비스일 수 있다.
- [0093] 프로세서(140)는, 음성 인식 서비스를 제공하는 동안 음성 수신부(120)를 통해 음성이 수신되면, 스토리지(130)에 저장된 화자 인식 모델과 수신된 음성의 특징 값이 기설정된 임계 값 이상의 유사도를 갖는지 식별할 수 있다.
- [0094] 식별 결과 기설정된 임계 값 이상의 유사도를 가지면, 프로세서(140)는 수신된 음성의 특징 값에 기초하여 스토리지(130)에 저장된 화자 인식 모델을 업데이트할 수 있다.
- [0095] 이 경우, 프로세서(140)는 음성 인식 서비스를 제공하는 동안 음성 수신부(120)를 통해 수신되는 음성에 대응되는 텍스트가 스토리지(130)에 저장된 복수의 텍스트에 포함되는 경우에, 수신된 음성의 특징 값에 기초하여 스토리지(130)에 저장된 화자 인식 모델을 업데이트할 수도 있다. 이때, 스토리지(130)에 저장된 복수의 텍스트는 화자 인식 모델의 업데이트를 위해 수신되어야 할 음성에 대응되는 텍스트일 수 있다.
- [0096] 즉, 프로세서(140)는, 음성 인식 서비스를 제공하는 동안 음성 수신부(120)를 통해 수신되는 음성의 특징 값이 스토리지(130)에 저장된 화자 인식 모델과 기설정된 임계 값 이상의 유사도를 가지는 조건에 더하여, 수신되는 음성에 대응되는 텍스트가 스토리지(130)에 저장된 복수의 텍스트에 포함되는 조건을 함께 만족하는 경우에, 수신된 음성의 특징 값에 기초하여 스토리지(130)에 저장된 화자 인식 모델을 업데이트할 수도 있다.
- [0097] 도 5는 화자 인식 모델을 완성하기 위해 음성 형태로 입력받아야 하는 복수의 텍스트가 기설정되어 있는 경우, 전자 장치(100)의 동작을 설명하기 위한 도면이다. 이때, 복수의 텍스트는 스토리지(130)에 저장되어 있을 수 있다.
- [0098] 도 5는 화자 인식 모델을 생성 및 업데이트하기 위해 수신되어야 할 음성에 대응되는 복수의 텍스트를 부분별 블록 형태로 도시한 것이다.

- [0099] 도 5의 (a)를 참조하면, 검은색 블록들은, 복수의 텍스트 중에서 사용자의 발화 음성이 이미 수신된 등록 텍스트(510)를 의미한다. 즉, 등록 텍스트(510)에 대응되는 음성은 이미 음성 수신부(120)를 통해 수신되어 화자 인식 모델의 생성 또는 업데이트에 사용된 상황이다.
- [0100] 반면, 흰색 블록들은, 복수의 텍스트 중에서 아직 사용자의 발화 음성이 수신되지 않은 미등록 텍스트(520)를 의미한다. 즉, 미등록 텍스트(520)에 대응되는 음성은 아직 음성 수신부(120)를 통해 수신된 바 없고, 따라서 화자 인식 모델의 생성 또는 업데이트에 사용된 바도 없다.
- [0101] 이때, 음성 수신부(120)를 통해 음성이 수신되면, 프로세서(140)는 스토리지(130)에 저장된 화자 인식 모델을 수신된 음성과 비교하여 유사도를 측정하는 한편, 수신된 음성에 대응되는 것으로 인식된 텍스트가 미등록 텍스트(520) 중 적어도 일부와 일치하는지 여부를 판단할 수 있다.
- [0102] 그 결과, 도 5의 (b)는, 유사도 측정 결과 기설정된 임계 값 이상으로 측정되었고, 수신된 음성에 대응되는 것으로 인식된 텍스트가 미등록 텍스트(520) 중 일부(500-4)와 일치하는 것으로 판단된 경우를 전체로 한다.
- [0103] 도 5의 (b)를 참조하면, 도 5의 (a)에서 미등록 텍스트(520)에 포함되던 텍스트(500-4)가 등록 텍스트(510)에 포함된 것을 알 수 있다. 즉, 텍스트(500-4)에 대응되는 음성이 음성 인식부(120)를 통해 수신되어 화자 인식 모델의 업데이트에 사용된 상황임을 알 수 있다.
- [0104] 프로세서(140)는 사용자 정보에 대응되는 화자 인식 모델의 업데이트 완료 여부를 식별할 수 있다. 구체적인 예로, 프로세서(140)는 상술한 사용자 정보와 동일한 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 사용자 정보에 대응되는 화자 인식 모델의 업데이트 완료 여부를 식별할 수도 있다.
- [0105] 이 경우, 프로세서(140)는 사용자 정보에 대응되는 화자 인식 모델의 업데이트 진행 정도 내지는 완성도를 비율로서 식별할 수 있다.
- [0106] 프로세서(140)는 사용자 정보에 기초하여 사용자 인증이 수행되면, 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스를 제공할 수도 있다.
- [0107] 이때, 유저 인터페이스는, 전자 장치(100) 또는 전자 장치(100)로부터 신호를 수신한 외부 장치(200)가, 안내 이미지를 디스플레이하거나 또는 안내 음성을 출력하는 형태로 제공될 수 있다.
- [0108] 도 6은 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스의 일 예를 설명하기 위한 도면이다. 도 6은 화자 인식 모델의 업데이트 진행 정도를 나타내기 위해 디스플레이되는 안내 이미지이다.
- [0109] 도 6은 사용자 "AAA"에 대한 사용자 정보가 입력된 결과 사용자 "AAA"에 대한 사용자 인증이 수행된 상황을 가정한다.
- [0110] 도 6을 참조하면, 사용자 "AAA"의 화자 인식 모델은 업데이트 진행 정도 내지는 완성도가 80 퍼센트라는 점을 확인할 수 있다.
- [0111] 한편, 사용자 인증이 로그인에 해당하는 경우의 예로, 프로세서(140)는 특정 사용자가 로그인된 상태에서 음성 인식 서비스가 제공되는 동안 음성 수신부(120)를 통해 수신되는 음성과 스토리지(130)에 저장된 화자 인식 모델을 비교하여 화자 인식 모델을 업데이트할 수 있다.
- [0112] 이렇듯, 본 개시에 따른 전자 장치(100)는, 화자 인식 모델의 생성 및 업데이트에 필요한 모든 텍스트에 대응되는 발화를 사용자에게 무리하게 요구하지는 않으면서도, 음성 인식 서비스가 제공되는 동안 자연스럽게 수신되는 사용자의 음성을 이용함으로써, 사용자별 화자 인식 모델을 효율적으로 완성해낸다는 장점이 있다.
- [0113] 또한, 전자 장치(100)는 사용자별 화자 인식 모델의 업데이트 진행 정도를 사용자 각각의 사용자 인증 이후에 유저 인터페이스 형태로 제공할 수도 있으므로, 이를 통해 사용자는 화자 인식 모델의 등록 과정에서 지루한 동작을 반복하지 않으면서도 그 진행 과정을 명확히 인지할 수 있다.
- [0114] 프로세서(140)는 업데이트된 화자 인식 모델에 기초하여, 음성 수신부(120)를 통해 수신되는 음성에 대응되는 사용자를 인식할 수 있다.
- [0115] 구체적으로, 프로세서(140)는, 특정 사용자의 화자 인식 모델에 대한 업데이트가 계속해서 이루어진 결과 업데이트가 완료되어 특정 사용자의 최종 화자 인식 모델이 생성(완성)되면, 그 후 음성 수신부(120)를 통해 수신되는 음성이 특정 사용자의 음성인지 여부를 식별할 수 있다.

- [0116] 완성된 최종 화자 인식 모델을 이용하는 구체적인 실시 예는, 본 개시의 전자 장치(100)에 있어 화자 인식 모델의 업데이트에 대한 실시 예를 모두 설명한 뒤, 보다 자세하게 후술한다.
- [0117] 한편, 프로세서(140)는 사용자 정보에 기초하여 사용자 인증이 재차 수행되면 사용자 정보에 대응되는 화자 인식 모델의 업데이트 완료 여부를 식별하고, 화자 인식 모델의 업데이트가 완료되지 않은 것으로 식별되면, 스토리지(130)에 저장된 복수의 텍스트 중 사용자에 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드를 제공할 수도 있다.
- [0118] 이때, 사용자에 의해 발화되지 않은 텍스트란, 해당 텍스트에 대응하는 음성이 음성 수신부(120)를 통해 수신되어 해당 텍스트로 인식된 적이 아직은 없음을 의미할 수 있다. 구체적으로는, 사용자가 사용자 인증을 수행한 상태에서, 해당 텍스트에 대응하는 음성이 음성 수신부(120)를 통해 수신되어 해당 텍스트로 인식된 적이 없음을 의미할 수도 있다.
- [0119] 이 경우, 스토리지(130)에 저장된 복수의 텍스트는 화자 인식 모델을 생성 또는 업데이트하기 위해 수신되어야 할 음성에 대응되는 텍스트일 수 있다.
- [0120] 이때, 사용자에 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드는, 발화되지 않은 텍스트를 포함하는 이미지를 디스플레이하거나 또는 발화되지 않은 텍스트에 대한 음성을 출력하는 것일 수 있다.
- [0121] 일 예로, 프로세서(140)가 사용자 정보에 대응되는 화자 인식 모델의 업데이트가 90 퍼센트 정도 진행된 것으로 식별한 경우, 프로세서(140)는 남은 10 퍼센트의 업데이트를 진행하기 위해 필요한 텍스트, 즉 스토리지(130)에 저장된 복수의 텍스트 중 아직 사용자의 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드를 제공할 수 있다.
- [0122] 그리고, 프로세서(140)는 가이드에 따른 음성이 음성 수신부를 통해 수신되면, 수신된 음성의 특징 값에 기초하여 화자 인식 모델을 업데이트하여 사용자 정보에 대응되는 최종 화자 인식 모델을 생성할 수 있다.
- [0123] 구체적으로, 프로세서(140)는, 수신된 음성의 특징 값이 스토리지(130)에 저장된 화자 인식 모델의 특징 값과 기설정된 임계 값 이상 유사하고, 수신된 음성에 대응되는 텍스트가 발화되지 않은 텍스트 중 적어도 일부와 일치하면, 수신된 음성의 특징 값에 기초하여 스토리지(130)에 저장된 화자 인식 모델을 업데이트하여 사용자 정보에 대응되는 최종 화자 인식 모델을 생성할 수 있다.
- [0124] 이는, 스토리지(130)에 저장된 복수의 텍스트 중 사용자에 의해 발화되지 않은 텍스트가 별로 남지 않은 경우, 사용자에게 큰 부담을 주지 않으면서도 신속하게 사용자를 위한 화자 인식 모델을 완성하기 위한 실시 예이다.
- [0125] 상술한 실시 예들을 통해 화자 인식 모델에 대한 업데이트가 완료된 결과 최종 화자 인식 모델이 생성되어 스토리지(130)에 저장되면, 프로세서(140)는 이를 기초로 음성 수신부(120)를 통해 수신되는 음성이 누구의 발화인지를 식별할 수 있다. 또한, 프로세서(120)는 이러한 식별 결과에 기초하여 사용자 인증을 수행할 수도 있다.
- [0126] 이 경우, 사용자가 임의의 텍스트를 발화하더라도, 프로세서(140)는 최종 화자 인식 모델에 기초하여 사용자를 인식할 수 있고, 이에 기초하여 사용자의 인증이 가능하다.
- [0127] 예로, 복수의 사용자 각각의 사용자 정보에 대응되는 각각의 최종 화자 인식 모델이 스토리지(130)에 저장되어 있다면, 프로세서(140)는 음성 수신부(120)를 통해 수신되는 복수의 사용자의 음성들 각각에 대해 복수의 사용자 중 누구의 음성인지 식별하고 식별된 사용자에 대한 사용자 인증을 수행할 수 있다.
- [0128] 예로, 별도의 동작 없이 자유 발화에 대한 음성 인식이 시작되거나 또는 단순히 특정 버튼을 누르는 동작만으로 자유 발화에 대한 음성 인식이 시작되는 경우, 프로세서(140)는 발화된 임의의 음성이 누구의 음성인지 식별하여 식별된 사용자에 대한 사용자 인증을 수행한 뒤, 인증된 사용자가 자유롭게 발화한 음성에 대하여 음성 인식을 시작할 수 있다. 관련하여, 원격 제어 장치(300)에 구비된 특정 버튼을 사용자가 누른 경우, 전자 장치(100)가 원격 제어 장치(300)로부터 음성 인식을 시작하라는 요청을 포함하는 신호를 수신하고, 그 후 전자 장치(100)가 업데이트된 화자 인식 모델 및 사용자가 자유롭게 발화한 음성에 기초하여 사용자를 인식하는 한편 사용자의 자유 발화에 대한 음성 인식을 시작하는 경우를 상정할 수 있다.
- [0129] 또한, 프로세서(140)는, 사용자 입력부(110)를 통해 수신되는 사용자의 아이디/비밀번호 정보, 터치 정보, 지문 정보, 얼굴 정보 및 홍채 정보 중 적어도 하나를 포함하는 사용자 정보에 대응되는 최종 화자 인식 모델에 기초하여, 음성 수신부(120)를 통해 수신되는 음성이 사용자의 음성이 맞는지 검증하는 방법을 통해, 사용자 인증을 수행할 수도 있다.

- [0130] 일 예로, 사용자 A가 원격 제어 장치(300) 또는 전자 장치(100)에 포함된 지문 센서(도시되지 않음)에 손가락을 대는 동시에 임의의 문장을 발화하는 경우, 사용자 A가 발화한 임의의 문장에 대응되는 음성이 원격 제어 장치(300) 또는 전자 장치(100)에 포함된 마이크(도시되지 않음)를 거쳐 사용자 음성 수신부(120)로 수신될 수 있다.
- [0131] 이때, 프로세서(120)는 스토리지(130)에 저장된 최종 화자 인식 모델 중에서 수신된 음성과 신호의 특징 값이 임계 값 이상 일치하는 최종 화자 인식 모델에 대응되는 사용자 정보를, 지문 센서(도시되지 않음)를 통해 인식된 지문에 대응되는 사용자 정보와 비교한 결과, 두 사용자 정보가 모두 사용자 A에 대한 것으로서 일치된다고 판명되면, 사용자 A에 대한 사용자 인증을 수행할 수 있다.
- [0132] 한편, 별도의 사용자 인증을 거치지 않고도, 프로세서(140)는 음성 수신부(120)를 통해 수신되는 음성이 누구의 음성인지를 식별한 것에 기초하여 사용자 맞춤형 서비스를 제공할 수 있다.
- [0133] 이 경우, 사용자 맞춤형 서비스란, 연령 제한이 있는 콘텐츠의 허용 내지는 차단, 사용자별 콘텐츠 사용 이력을 고려하여 맞춤형 광고 제공, 사용자별 유저 인터페이스 사용 이력을 고려한 맞춤형 유저 인터페이스 제공 등을 포함할 수 있다.
- [0134] 도 7은 본 개시의 다양한 실시 예에 따른 전자 장치(100)의 상세한 구성을 설명하기 위한 블록도이다.
- [0135] 도 7을 참조하면, 전자 장치(100)는 사용자 입력부(110), 음성 수신부(120), 스토리지(130), 프로세서(140) 외에 입출력 포트(150), 통신부(160), 디스플레이(170), 오디오 출력부(180) 및 방송 수신부(190) 중 적어도 하나를 더 포함할 수 있다.
- [0136] 입출력 포트(150)를 통해, 전자 장치(100)는 외부로부터 이미지 및/또는 음성에 대한 신호를 수신하거나 또는 외부에 이미지 및/또는 음성 신호를 전송할 수 있다.
- [0137] 이를 위해, 입출력 포트(150)는 HDMI 포트, 디스플레이 포트, RGB 포트, DVI(Digital Visual Interface) 포트, 썬더볼트 및 컴포넌트 포트 등 유선 포트로 구현될 수 있다. 또는, 입출력 포트(150)는 와이파이나 블루투스 통신 등 무선 통신을 위한 포트도 구현될 수도 있다.
- [0138] 입출력 포트(150)는 HDMI 포트나 썬더볼트 등으로 구현되어 이미지 및 음성 신호를 함께 전송하도록 구현될 수도 있지만, 이미지 신호를 전송하는 제1 포트 및 음성 신호를 전송하는 제2 포트가 각각 별도로 구현된 것일 수도 있다.
- [0139] 예로, 전자 장치(100)가 디스플레이(170)를 포함하는 경우, 전자 장치(100)는 입출력 포트(150)를 통해 셋탑박스 등의 외부 장치(도시되지 않음)로부터 특정한 유저 인터페이스를 포함하는 이미지 및/또는 음성에 대한 신호를 수신할 수 있다.
- [0140] 다른 예로, 전자 장치(100)가 디스플레이(170)를 포함하지 않는 셋탑박스 등으로 구현된 경우, 전자 장치(100)는 특정한 유저 인터페이스를 포함하는 이미지 및/또는 음성이 외부 장치(200)를 통해 출력되도록, 입출력 포트(150)를 통해 특정한 유저 인터페이스를 포함하는 이미지 및/또는 음성을 외부 장치(200)로 전송할 수도 있다.
- [0141] 통신부(160)는 전자 장치(100)가 다양한 유형의 통신방식에 따라 다양한 유형의 외부 기기와 통신을 수행할 수 있게 하는 구성이다. 통신부(160)는 와이파이칩(도시되지 않음), 블루투스 칩(도시되지 않음) 등을 포함할 수 있다.
- [0142] 프로세서(140)는 통신부(160)를 이용하여 각종 외부 기기와 통신을 수행할 수 있다. 통신부(160)는 각종 외부 기기와 와 무선 혹은 유선으로 데이터 통신을 수행할 수 있다.
- [0143] 무선 통신 방식으로 외부 장치(도시되지 않음)와 데이터 통신을 수행할 경우, 통신부(160)는 와이파이 다이렉트(WIFI DIRECT) 통신 모듈, 블루투스(blueetooth)모듈, 적외선 통신(IrDA, infrared data association)모듈, NFC(Near Field Communication)모듈, 지그비(Zigbee) 모듈, 셀룰러 통신모듈, 3G(3세대) 이동통신 모듈, 4G(4세대) 이동통신 모듈, 4세대 LTE(Long Term Evolution) 통신 모듈 중 적어도 하나를 포함할 수 있다.
- [0144] 예로, 디스플레이(170)를 포함한 전자 장치(100)는 통신부(160)를 통해 원격 제어 장치(300)로부터 데이터를 수신할 수 있다. 구체적으로, 통신부(160)는 원격 제어 장치(300)에 입력된 사용자 정보에 대한 데이터 및/또는 신호, 원격 제어 장치(300)의 마이크(도시되지 않음)에 입력된 사용자 음성에 대한 데이터 및/또는 신호, 원격 제어 장치(300)에 입력된 사용자 명령에 대한 데이터 및/또는 신호 등을 블루투스 또는 와이파이 통신 방식으로

원격 제어 장치(300)로부터 수신할 수 있다.

- [0145] 이때, 원격 제어 장치(300)의 마이크(도시되지 않음)에 입력된 사용자 음성에 대한 데이터 및/또는 디지털 신호를 수신한 통신부(160)는, 수신된 데이터 및/또는 디지털 신호를 음성 수신부(120)로 전송할 수 있다.
- [0146] 또는, 원격 제어 장치(300)의 각종 센서(도시되지 않음)나 카메라를 통해 얻어진 홍채 인식 데이터, 얼굴 촬영 데이터, 지문 인식 데이터 등 사용자 정보 데이터를 수신한 통신부(160)는, 수신된 사용자 정보 데이터를 사용자 입력부(110) 및 프로세서(140) 중 적어도 하나로 전송할 수 있다.
- [0147] 통신부(160)는 전자 장치(100)가 수행할 수 있는 기능과 관련된 데이터를 외부 서버(도시되지 않음)와 송수신할 수도 있다.
- [0148] 한편, 통신부(160)는, 외부 장치(200)나 원격 제어 장치(300) 등과 데이터를 송수신하기 위한 제1 통신부(도시되지 않음) 및 외부 서버(도시되지 않음)와 데이터를 송수신하기 위한 제2 통신부(도시되지 않음)를 각각 별도로 포함할 수도 있다. 이때, 제1 통신부(도시되지 않음)는 이더넷 모듈, 와이파이 모듈 및 블루투스 모듈 등으로 구현되는 한편, 제2 통신부는 별도의 와이파이 모듈로 구현될 수도 있다.
- [0149] 디스플레이(170)는 프로세서(140)의 제어에 따라 영상을 표시하기 위한 구성이다. 이를 위해, 디스플레이(170)는 LCD(Liquid Crystal Display), PDP(Plasma Display Panel), OLED(Organic Light Emitting Diodes), TOLED(Transparent OLED) 등으로 구현될 수 있다. LCD로 구성되는 경우, 디스플레이(170) 내에는 a-si TFT, LTPS(low temperature poly silicon) TFT, OTFT(organic TFT) 등과 같은 형태로 구현될 수 있는 구동 회로(도시되지 않음), 백라이트 유닛(도시되지 않음) 등도 함께 포함될 수 있다.
- [0150] 한편, 디스플레이(170)는 사용자의 터치 조작을 감지할 수 있는 터치 패널이 포함된 터치스크린 형태로 구현될 수도 있으며, 이 경우 사용자 입력부(110)에 포함되는 것으로 볼 수 있다.
- [0151] 디스플레이(170)는 방송 수신부(190)를 통해 수신된 방송 신호에 포함된 이미지, 통신부(160)를 통해 수신되는 신호에 포함된 이미지, 사용자 입력부(110)를 통해 입력되는 이미지, 스토리지(130)에 저장된 이미지 파일에 포함된 이미지, 입출력 포트(150)를 통해 외부로부터 수신된 신호에 포함된 이미지 등을 출력할 수 있다.
- [0152] 오디오 출력부(180)는 프로세서(140)의 제어에 따라 특정한 음성을 출력하기 위한 구성이다. 이를 위해, 오디오 출력부는 스피커(도시되지 않음) 및 헤드폰/이어폰 출력 단자(도시되지 않음)를 포함할 수 있다.
- [0153] 오디오 출력부(180)는 방송 수신부(190)를 통해 수신된 방송 신호에 포함된 음성, 통신부(160)를 통해 수신되는 신호에 포함된 음성, 사용자 입력부(110)에 포함된 마이크(도시되지 않음)를 통해 입력되는 음성, 스토리지(130)에 저장된 오디오 파일에 포함된 음성, 입출력 포트(150)를 통해 외부로부터 수신된 신호에 포함된 음성 등을 출력할 수 있다.
- [0154] 방송 수신부(190)는 방송 콘텐츠에 대한 신호를 수신할 수 있다. 방송 콘텐츠는 영상, 오디오 및 부가 데이터(예를 들어, EPG)를 포함할 수 있으며, 방송 수신부(190)는 지상파 방송, 케이블 방송, 위성 방송, 인터넷 방송 등과 같이 다양한 소스로부터 방송 콘텐츠 신호를 수신할 수 있다.
- [0155] 방송 수신부(190)는 방송국으로부터 전송되는 방송 콘텐츠를 수신하기 위해 튜너(미도시), 복조기(미도시), 등화기(미도시) 등과 같은 구성을 포함하는 형태로 구현될 수 있다.
- [0156] 도 8은 본 개시의 일 실시 예에 따른 전자 장치의 제어 방법을 설명하기 위한 순서도이다.
- [0157] 본 제어 방법은, 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 사용자 인증이 수행되면, 사용자 발화를 유도하기 위한 가이드를 제공한다. (S810)
- [0158] 이때, 사용자 발화를 유도하기 위한 가이드는, 사용자가 발화해야 할 텍스트를 포함하는 이미지가 디스플레이되는 과정 및 사용자가 발화해야 할 텍스트에 대응되는 음성이 출력되는 과정 중 적어도 하나를 포함할 수 있다.
- [0159] 일 예로, 사용자의 발화를 유도하기 위한 가이드는, 사용자별 화자 인식 모델을 생성 및 업데이트하기 위해 수신되어야 하는 음성에 대응되는 복수의 텍스트 중 일부의 텍스트를 포함하여, 사용자가 일부의 텍스트에 대응되는 음성을 발화하도록 유도할 수 있다.
- [0160] 또한, 사용자 발화를 유도하기 위한 가이드는, 사용자 정보에 대응되는 화자 인식 모델을 생성하기 위한 발화를 수행할 것인지 여부를 사용자에게 물어보고 사용자가 선택하게 하는 유저 인터페이스를 포함할 수 있다.
- [0161] 사용자 인증이 로그인에 해당하는 경우의 일 예로, 사용자 입력부를 통해 입력된 사용자 정보에 기초하여 특정

사용자의 로그인에 최초로 수행되면, 특정 사용자의 발화를 유도하기 위한 가이드를 제공할 수 있다.

- [0162] 본 제어 방법은, 음성 수신부를 통해 수신된, 가이드에 대응되는 음성에 기초하여 사용자 정보에 대응되는 화자 인식 모델을 생성하여 저장할 수 있다. (S820)
- [0163] 만약, 사용자별 화자 인식 모델을 생성 및 업데이트하기 위해 수신되어야 하는 음성에 대응되는 복수의 텍스트 중 일부의 텍스트가 가이드에 포함된 상황에서, 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 가이드에 포함된 텍스트와 일치하는 경우, 수신된 음성의 특징에 기초하여 화자 인식 모델을 생성할 수 있다.
- [0164] 사용자 인증이 로그인에 해당하는 경우의 일 예로, 가이드에 따른 음성이 음성 수신부를 통해 수신되면, 수신된 음성에 기초하여 특정 사용자에게 대응되는 화자 인식 모델을 생성하여 저장할 수 있다.
- [0165] 상술한 S810 및 S820의 단계는, 초기 화자 인식 모델 생성 과정(900)에 포함될 수 있다.
- [0166] 초기 화자 인식 모델을 생성한 뒤, 본 제어 방법은, 음성 수신부를 통해 수신되는 음성과 화자 인식 모델을 비교하여 업데이트된 화자 인식 모델에 기초하여, 수신되는 음성에 대응되는 화자를 인식할 수 있다(S830).
- [0167] 이때, 전자 장치에 대한 사용자 인증이 유지되고 있는 상태에서, 해당 사용자에게 대한 음성 인식 서비스가 제공되는 동안 음성 수신부를 통해 수신되는 음성과 저장된 화자 인식 모델을 비교하여 저장된 화자 인식 모델을 업데이트할 수 있다.
- [0168] 이 경우, 음성 인식 서비스를 제공하는 동안 음성 수신부를 통해 음성이 수신되면, 저장된 화자 인식 모델과 수신된 음성의 특징 값이 기설정된 임계 값 이상의 유사도를 갖는지 식별하고, 식별 결과 기설정된 임계 값 이상의 유사도를 가지면, 수신된 음성의 특징 값에 기초하여 저장된 화자 인식 모델을 업데이트할 수도 있다.
- [0169] 또한, 음성 인식 서비스를 제공하는 동안 음성 수신부를 통해 수신된 음성에 대응되는 텍스트가 기저장된 복수의 텍스트에 포함되는 경우, 수신된 음성의 특징 값에 기초하여 저장된 화자 인식 모델을 업데이트할 수도 있다. 이때, 기저장된 복수의 텍스트는, 사용자별 화자 인식 모델을 생성 또는 업데이트하기 위해 음성 수신부를 통해 수신되어야 할 음성에 대응되는 텍스트일 수 있다.
- [0170] 사용자 인증이 로그인에 해당하는 경우의 일 예로, 특정 사용자가 로그인된 상태에서 음성 인식 서비스를 제공하는 동안 음성 수신부를 통해 수신되는 음성과 저장된 화자 인식 모델을 비교하여 화자 인식 모델을 업데이트할 수도 있다.
- [0171] 한편, 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 화자 인식 모델의 업데이트 진행 정도를 나타내는 유저 인터페이스를 제공할 수도 있다.
- [0172] 또한, 사용자 정보에 기초하여 사용자 인증이 재차 수행되면, 사용자 정보에 대응되는 화자 인식 모델의 업데이트 완료 여부를 식별하고, 화자 인식 모델의 업데이트가 완료되지 않은 것으로 식별되면, 기저장된 복수의 텍스트 중 사용자에게 의해 발화되지 않은 텍스트의 발화를 유도하기 위한 가이드를 제공할 수도 있다.
- [0173] 이때, 가이드에 따른 음성이 음성 수신부를 통해 수신되면, 수신된 음성의 특징 값에 기초하여 화자 인식 모델을 업데이트하여 사용자 정보에 대응되는 최종 화자 인식 모델을 생성할 수도 있다.
- [0174] 상술한 화자 인식 모델의 업데이트는, 사용자에게 대한 발화 요청 없이 화자 인식 모델을 업데이트하는 과정(1000)뿐만 아니라, 사용자에게 대한 발화 요청을 통해 화자 인식 모델을 업데이트하는 과정(1100)도 포함할 수 있다.
- [0175] 화자 인식 모델을 업데이트한 뒤, 본 제어 방법은, 업데이트된 화자 인식 모델에 기초하여, 음성 수신부를 통해 수신되는 음성에 대응되는 사용자를 인식할 수 있다.
- [0176] 구체적으로는, 업데이트가 완료되어 생성된 사용자별 최종 화자 인식 모델에 기초하여, 음성 수신부를 통해 수신되는 음성에 대응되는 사용자가 누구인지 인식할 수 있다.
- [0177] 도 9는 초기 화자 인식 모델을 생성하는 과정(900)의 일 예를 설명하기 위한 알고리즘이다.
- [0178] 먼저, 전자 장치에 사용자 정보가 입력되면(S910), 기저장된 정보와 입력된 사용자 정보를 비교하여 사용자 인증을 할 수 있다. 이때, 입력된 사용자 정보가 인증하고자 하는 사용자에게 대한 정보와 대응되지 않으면(S920 - N) 사용자 정보를 다시 입력받을 수 있다.
- [0179] 입력된 사용자 정보에 기초하여 사용자 인증이 완료되면(S920 - Y), 입력된 사용자 정보에 대응되는 화자 인식

모델이 저장되어 있는지 식별할 수 있다. (S930) 이때, 만약 입력된 사용자 정보에 대응되는 화자 인식 모델이 저장되어 있으면(S930 - Y), 곧바로 서비스에 진입할 수 있다.

- [0180] 반면, 만약 입력된 사용자 정보에 대응되는 화자 인식 모델이 저장되어 있지 않으면(S930 - N), 입력된 사용자 정보에 대응되는 화자 인식 모델을 생성할 것인지의 여부를 사용자가 선택하도록 할 수 있다(S940). 구체적으로는, 향후 사용자 본인의 음성을 인식하기 위해 필요한 화자 인식 모델을 생성하기 위한 발화 과정을 지금 수행할 것인지 여부를 사용자가 선택하도록 할 수 있다.
- [0181] 이때, 입력된 사용자 정보에 대응되는 화자 인식 모델을 생성할 것인지의 여부를 사용자에게 묻는 과정 및 사용자가 선택하는 과정은, 유저 인터페이스를 통해 진행될 수 있다.
- [0182] 이때, 유저 인터페이스는 정보를 표시하는 이미지를 디스플레이하는 과정, 정보에 대한 음성을 출력하는 과정 및 사용자의 선택을 입력받는 과정 등을 포함할 수 있다.
- [0183] 사용자가 사용자 정보에 대응되는 화자 인식 모델을 생성할 것을 선택한 경우(S940 - Y), 사용자 발화를 유도하기 위한 가이드를 제공할 수 있으며, 이때 사용자가 발화한 음성을 수신할 수 있다(S950).
- [0184] 사용자 발화를 유도하기 위한 가이드는, 사용자가 발화해야 할 텍스트를 포함하는 이미지가 디스플레이되는 과정 및 사용자가 발화해야 할 텍스트에 대응되는 음성이 출력되는 과정 중 적어도 하나를 포함할 수 있다.
- [0185] 사용자가 발화해야 할 텍스트는, 초기 화자 인식 모델을 생성하기 위해 수신받아야 하는 사용자의 음성에 대응되는 텍스트일 수 있다.
- [0186] 사용자가 발화한 음성이 수신되면, 발화 내용이 정확한 것인지 판단할 수 있다(S960). 발화 내용이 정확한 것인가에 대한 판단은, 사용자가 초기 화자 인식 모델을 생성하기 위해 수신되어야 할 음성을 정확히 발화했는가에 대한 판단을 의미한다. 이때, 발화 내용이 정확한 것인가에 대한 판단은, 사용자가 발화한 음성을 인식한 결과가 사용자가 발화해야 할 텍스트와 임계 값 이상으로 일치하는지 여부에 기초한 것일 수 있다.
- [0187] 상술한 판단 과정은, 예를 들면, 사용자 발화를 유도하기 위한 가이드가 사용자에게 "안녕하세요. 당신을 만나 정말 기쁩니다."라는 문장을 발화하기를 요청한 상태에서, 사용자가 발화한 음성이 해당 문장에 정확히 대응되는지 여부를 판단하는 과정이다.
- [0188] 판단 결과, 발화 내용이 정확하지 않으면(S960 - N) 사용자가 재발화한 음성을 수신할 필요가 있다(S950).
- [0189] 반면, 판단 결과, 발화 내용이 정확하면(S960 - Y), 수신된 음성의 특징을 추출할 수 있다(S970). 이때, 음성의 특징은 Pitch, Formant, LPCC(Linear Predictive Cepstral Coefficient), MFCC(Mel-Frequency Cepstral Coefficient), PLP(Perceptual Linear Predictive) 등에 기초한 것일 수 있다.
- [0190] 그 후, 추출된 특징에 기초하여, 입력된 사용자 정보에 대응되는 초기 화자 인식 모델을 생성하여 저장할 수 있다(S980).
- [0191] 상술한 과정들을 통해 화자 인식 모델을 완성할 수 있을 정도로 많은 양의 발화를 사용자에게 요구하는 것은 아니고, 단지 초기 화자 인식 모델을 생성할 수 있을 정도의 발화만을 요구한 채로 서비스에 진입(S990)하기 때문에, 사용자의 편의성을 증대시킬 수 있다.
- [0192] 다만, 아직 화자 인식 모델이 완성된 상태가 아니기 때문에, 생성된 화자 인식 모델을 수시로 업데이트시킬 필요가 있다.
- [0193] 도 10은 사용자에게 대한 발화 요청 없이 화자 인식 모델을 업데이트하는 과정(1000)의 일 예를 설명하기 위한 알고리즘이다.
- [0194] 본 과정(1000)은 초기 화자 인식 모델을 생성하는 과정(900)과 동일한 사용자 정보가 입력되어 동일한 사용자 인증이 수행된 결과, 동일한 사용자가 음성 인식 서비스를 이용하는 동안에 계속해서 수행될 수 있다.
- [0195] 본 과정(1000)은, 생성된 화자 인식 모델에 대응되는 사용자 정보가 입력된 결과 특정 사용자에게 대한 사용자 인증이 재차 이루어진 상태에서, 사용자가 발화한 음성을 수신할 수 있다(S1010). 이는, 인증을 완료한 사용자가 전자 장치를 통해 음성 인식 서비스를 이용하는 와중에 여러 가지 텍스트를 발화하는 상황일 수 있다.
- [0196] 음성이 수신되면, 수신된 음성에 대한 특징을 추출할 수 있다(S1020). 그리고, 추출된 특징 값을 기준에 저장되어 있던 화자 인식 모델과 비교하여 유사도를 산출할 수 있다(S1030). 즉, 입력된 사용자 정보에 대응되는 기저장된 화자 인식 모델의 특징이, 새롭게 수신된 음성의 특징과 임계 값 이상 일치하는지 판단하는 과정이다. 그

리고, 유사도가 임계 값 이상인지 판단(S1040)할 수 있다.

- [0197] 판단 결과, 유사도가 임계 값 미만이라면(S1040 - N), 수신된 음성이 입력된 사용자 정보에 대응되는 사용자의 음성이 아닐 가능성이 높으므로, 수신된 음성을 기초로 기저장된 화자 인식 모델을 업데이트할 수 없다.
- [0198] 반면, 유사도가 임계 값 이상이라면(S1040 - Y), 수신된 음성이 입력된 사용자 정보에 대응되는 사용자의 음성이 맞을 가능성이 높으므로, 수신된 음성을 기초로 기저장된 화자 인식 모델을 업데이트할 수 있다(S1050). 즉, 기저장된 화자 인식 모델의 특징에 수신된 음성의 특징을 부가하여 화자 인식 모델을 업데이트한 뒤, 업데이트된 화자 인식 모델을 다시 저장할 수 있다.
- [0199] 한편, 도 10과 달리, 저장된 화자 인식 모델의 특징과 수신된 음성의 특징을 비교한 유사도뿐만 아니라, 수신된 음성이 기저장된 텍스트에 대응되는지 여부도 함께 고려할 수 있다. 이때, 기저장된 텍스트는, 사용자별 화자 인식 모델을 업데이트함에 있어 수신되어야 하는 음성에 대응되는 텍스트일 수 있다.
- [0200] 즉, 수신된 음성의 특징이 저장된 화자 인식 모델의 특징과 임계 값 이상 유사하고, 이에 더하여, 수신된 음성이 화자 인식 모델을 업데이트하기 위해 수신되어야 하는 음성에 대응되는 텍스트에 대응되는 경우에만, 수신된 음성을 기초로 기저장된 화자 인식 모델을 업데이트할 수도 있다.
- [0201] 사용자에게 대한 발화 요청 없이 화자 인식 모델을 업데이트하는 과정(1000) 외에, 사용자에게 대한 발화 요청을 통해 화자 인식 모델을 업데이트하는 과정(1100)이 별도로 필요할 수 있다.
- [0202] 사용자에게 대한 발화 요청을 통해 화자 인식 모델을 업데이트하는 과정(1100)은, 일반적으로 화자 인식 모델에 대한 업데이트가 거의 다 완료된 시점에 수행됨으로써, 업데이트가 100 퍼센트 완료된 최종 화자 인식 모델을 생성하기 위해 수행될 수 있다.
- [0203] 사용자에게 대한 발화 요청을 통해 화자 인식 모델을 업데이트하는 과정(1100)은, 최종 화자 인식 모델을 완성하기 위해서는 기저장된 복수의 텍스트에 대응되는 음성이 모두 수신되어야 함에도, 복수의 텍스트 중 특정한 텍스트에 대응되는 음성이 아직 수신되지 않았거나 또는 장기간 수신되지 않았다면, 유용할 수 있다.
- [0204] 또한, 사용자에게 대한 발화 요청을 통해 화자 인식 모델을 업데이트하는 과정(1100)은, 최종 화자 인식 모델을 완성하기 위해 일정한 시간 또는 용량 이상의 음성이 수신되어야 하는데 아직 일정한 시간 또는 용량만큼의 음성을 수신받지 못했고, 최종 화자 인식 모델을 급하게 완성할 필요가 있는 경우, 역시 유용할 수 있다.
- [0205] 도 11은 사용자에게 대한 발화 요청을 통해 화자 인식 모델을 업데이트하는 과정(1100)의 일 예를 설명하기 위한 알고리즘이다.
- [0206] 본 과정(1100)은 사용자 정보가 입력되어(S1110) 재차 사용자 인증이 완료되면(S1120 - Y), 입력된 사용자 정보에 대응되는 화자 인식 모델의 업데이트가 완료되었는지 여부를 판단할 수 있다(S1130).
- [0207] 판단 결과, 화자 인식 모델의 업데이트가 완료되지 않은 경우(S1130 - N), 발화되지 않은 텍스트를 사용자가 발화하도록 요청하는 한편 사용자가 발화한 음성을 수신할 수 있다(S1140).
- [0208] 이때, 발화하지 않은 텍스트란, 화자 인식 모델을 완성하기 위해 복수의 텍스트에 대응되는 음성을 수신해야 하는데, 복수의 텍스트 중 일부 텍스트에 대응되는 음성만 수신되어 화자 인식 모델의 업데이트에 사용된 경우, 복수의 텍스트 중 이미 수신된 음성에 대응되지 않는 나머지 텍스트들을 의미할 수 있다.
- [0209] 이때, 발화되지 않은 텍스트를 사용자가 발화하도록 요청하는 것 외에, 화자 인식 모델을 마저 완성하는 데에 추가적으로 필요한 시간 또는 분량에 해당하는 음성을 발화하도록 사용자에게 요청할 수도 있다.
- [0210] 이 경우, 화자 인식 모델의 업데이트가 완료되지 않았을 뿐만 아니라 화자 인식 모델의 업데이트가 일정 비율 이상은 진행된 경우에만, 사용자에게 발화를 요청하고 그 음성을 수신할 수도 있다.
- [0211] 예를 들어, 화자 인식 모델의 업데이트가 20 퍼센트 정도밖에 되지 않았다면 사용자가 발화해야 하는 양이 많아 사용자에게 부담이 되므로, 화자 인식 모델의 업데이트가 90 퍼센트를 넘은 경우에만 사용자에게 대하여 발화를 요청하고 음성을 수신할 수도 있다.
- [0212] 이때, 사용자의 발화를 유도하는 가이드를 이용하여 사용자에게 발화를 요청할 수도 있다.
- [0213] 발화 요청 결과 음성이 수신되면 수신된 음성, 즉 발화 내용이 정확한지 여부를 판단할 수 있다(S1150). 이는, 사용자가, 발화되지 않은 텍스트에 대응되는 음성을 정확히 발화했는지 여부를 판단하는 것이다. 즉, 수신된 음

성을 인식한 결과가 발화되지 않은 텍스트와 대응되는지 여부를 판단하는 것이다.

- [0214] 판단 결과, 발화 내용이 정확하다면(S1150 - Y) 수신된 음성의 특징을 추출할 수 있다(S1160).
- [0215] 이 경우, 발화 내용이 정확하고, 이에 더하여, 수신된 음성의 특징이 기존의 화자 인식 모델의 특징과 일정 값 이상 유사한 경우에만, 수신된 음성의 특징을 추출할 수도 있다.
- [0216] 그리고, 추출된 특징에 기초하여 화자 인식 모델에 대한 업데이트를 최종적으로 완료할 수 있다(S1170).
- [0217] 도 8 내지 도 11을 통해 상술한 전자 장치의 제어 방법은, 도 3 및 도 7을 통해 도시 및 설명된 전자 장치(100)에 의해 수행되거나 전자 장치(100) 및 원격 제어 장치(300)를 포함하는 시스템에 의해 수행될 수도 있다.
- [0218] 또는, 상술한 제어 방법은, 디스플레이가 없는 전자 장치(100), 디스플레이가 있는 외부 장치(200) 및 원격 제어 장치(300)를 통해 수행될 수도 있다.
- [0219] 또는, 상술한 제어 방법은, 전자 장치(100) 및 전자 장치(100)와 데이터를 송수신할 수 있는 서버(도시되지 않음)를 통해 수행될 수도 있다.
- [0220] 일 예로, 사용자로부터 사용자 정보를 입력받고 사용자 음성을 수신하는 과정만 전자 장치(100)를 통해 수행되고, 입력된 사용자 정보에 기초하여 사용자 인증을 수행하고 수신된 음성을 기초로 화자 인식 모델을 생성 및 업데이트하는 과정 등은 서버(도시되지 않음)를 통해 구현될 수 있다. 이때, 서버(도시되지 않음)는 사용자 발화를 유도하기 위한 가이드를 제공하라는 명령 신호를 전자 장치(100)에 전송할 수 있다.
- [0221] 이 경우, 서버(도시되지 않음)에 의해 사용자별 최종 화자 인식 모델이 생성되면, 전자 장치(100)는 수신되는 음성에 대한 데이터를 서버로 전송할 수 있고, 서버는 수신된 음성이 누구의 음성인지 식별하여 식별 결과에 대한 데이터를 전자 장치(100)에 전송할 수도 있다.
- [0222] 다른 예로, 사용자 정보를 입력받고 사용자 음성을 수신하는 과정뿐만 아니라, 사용자 인증을 수행하는 과정 및 사용자 발화를 유도하기 위한 가이드를 제공하는 과정까지 전자 장치(100)에 의해 수행될 수 있다.
- [0223] 다만, 전자 장치(100)가 사용자 인증을 수행한 결과 사용자에게 가이드를 제공함과 동시에, 수신되는 음성은 서버(도시되지 않음)로 전송한 결과 서버(도시되지 않음)를 통해 화자 인식 모델의 생성 및 업데이트가 진행될 수도 있다.
- [0224] 이는, 수신되는 음성의 특징을 분석하여 화자 인식 모델을 생성 및 업데이트하고 그 결과를 저장하기 위해서는 일정 수준 이상의 연산속도 및 메모리 용량이 필요하다는 점을 고려한 것으로, 전자 장치(100)에 비해 비교적 연산 속도가 뛰어나고 데이터 저장량에 여유가 있는 서버(도시되지 않음)를 이용하는 실시 예이다.
- [0225] 한편, 이상에서 설명된 다양한 실시 예들은 소프트웨어(software), 하드웨어(hardware) 또는 이들의 조합된 것을 이용하여 컴퓨터 또는 이와 유사한 장치로 읽을 수 있는 기록 매체 내에서 구현될 수 있다.
- [0226] 하드웨어적인 구현에 의하면, 본 개시에서 설명되는 실시 예들은 ASICs(Application Specific Integrated Circuits), DSPs(digital signal processors), DSPDs(digital signal processing devices), PLDs(Programmable logic devices), FPGAs(field programmable gate arrays), 프로세서(processor), 제어기(controller), 마이크로 컨트롤러(micro-controllers), 마이크로 프로세서(microprocessor), 기타 기능 수행을 위한 전기적인 유닛(unit) 중 적어도 하나를 이용하여 구현될 수 있다.
- [0227] 일부의 경우에 본 명세서에서 설명되는 실시 예들이 프로세서(140) 자체로 구현될 수 있다. 소프트웨어적인 구현에 의하면 본 명세서에서 설명되는 절차 및 기능과 같은 실시 예들은 별도의 소프트웨어 모듈들로 구현될 수 있다. 상술한 소프트웨어 모듈들 각각은 본 명세서에서 설명되는 하나 이상의 기능 및 작동을 수행할 수 있다.
- [0228] 한편, 상술한 본 개시의 다양한 실시 예들에 따른 전자 장치(100)에서의 처리동작을 수행하기 위한 컴퓨터 명령어(computer instructions)는 비일시적 컴퓨터 판독 가능 매체(non-transitory computer-readable medium)에 저장될 수 있다. 이러한 비일시적 컴퓨터 판독 가능 매체에 저장된 컴퓨터 명령어는 특정 기기의 프로세서에 의해 실행되었을 때 상술한 다양한 실시 예에 따른 전자 장치(100)의 처리 동작을 상술한 특정 기기가 수행하도록 한다.
- [0229] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM

등과 같은 비밀시적 판독 가능 매체에 저장되어 제공될 수 있다.

[0230] 또한, 이상에서는 본 발명의 바람직한 실시 예에 대하여 도시하고 설명하였지만, 본 발명은 상술한 특징의 실시 예에 한정되지 아니하며, 청구범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 기술분야에서 통상의 지식을 가진 자에 의해 다양한 변형실시가 가능한 것은 물론이고, 이러한 변형실시들은 본 발명의 기술적 사상이나 전망으로부터 개별적으로 이해해서는 안 될 것이다.

**부호의 설명**


[0231] 100: 전자 장치      110: 사용자 입력부  
 120: 음성 수신부      130: 스토리지  
 140: 프로세서      150: 입출력 포트  
 160: 통신부      170: 디스플레이  
 180: 오디오 출력부      190: 방송 수신부  
 200: 외부 장치      300: 원격 제어 장치

도면

도면1

회차 등록을 위해 화면에 표시된 문장을 읽어주세요 (1/30)

"안녕하세요. 당신을 만나 정말 기쁩니다."




건너뛰기



회차 등록을 위해 화면에 표시된 문장을 읽어주세요 (2/30)

"슬픈 노래를 즐겨 들으시나요?"



건너뛰기




⋮



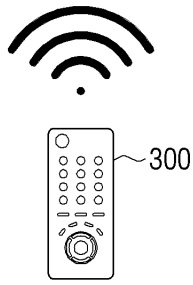
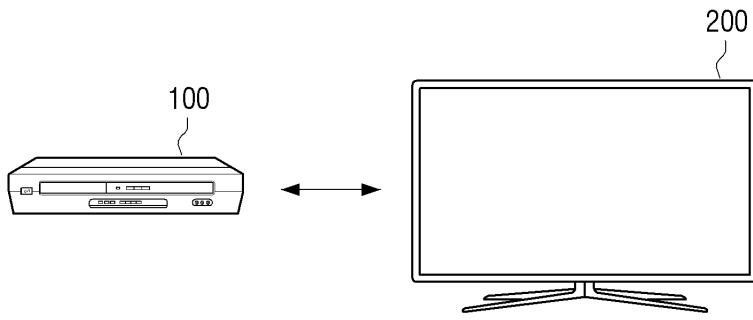
회차 등록을 위해 화면에 표시된 문장을 읽어주세요 (30/30)

"정말 멋진 목소리를 가지셨네요."

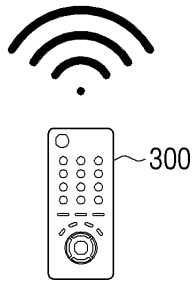
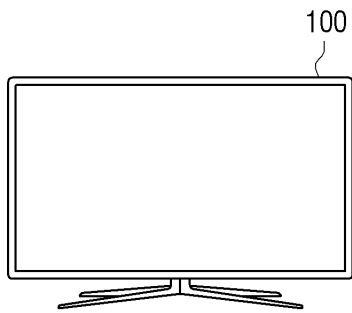


건너뛰기

도면2

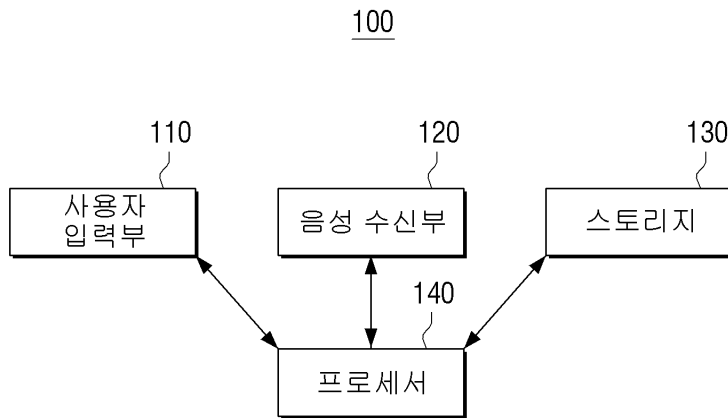


(a)

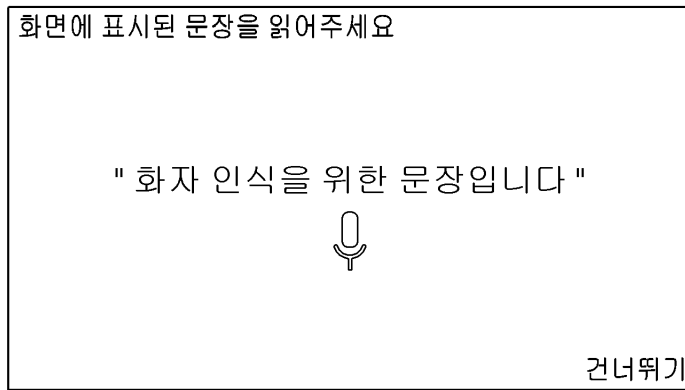


(b)

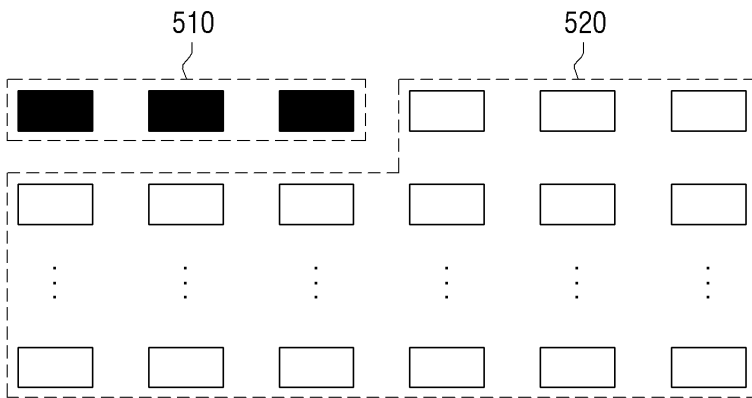
도면3



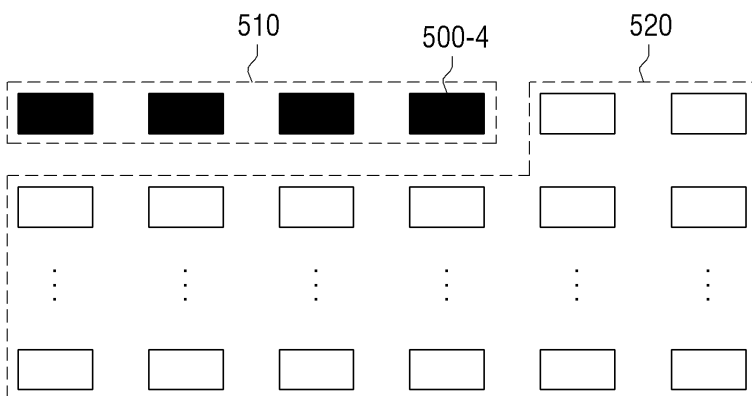
도면4



도면5

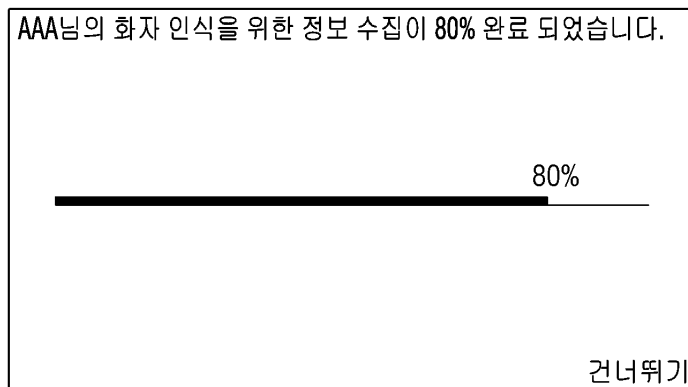


(a)

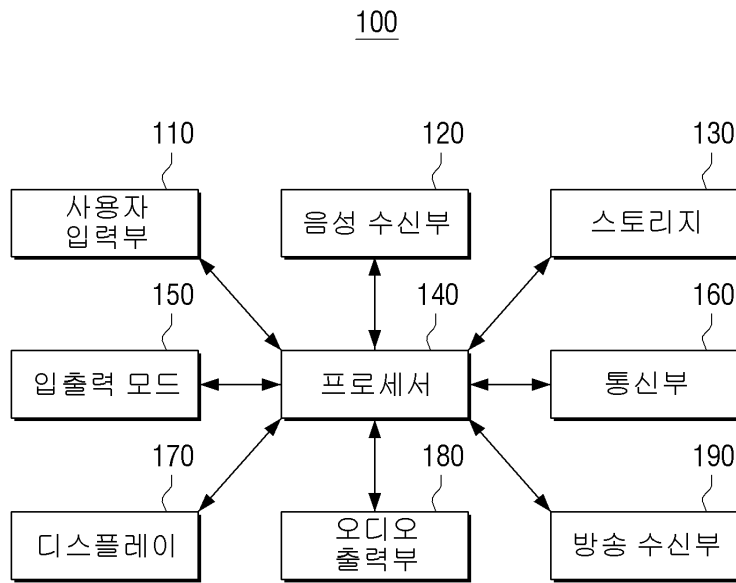


(b)

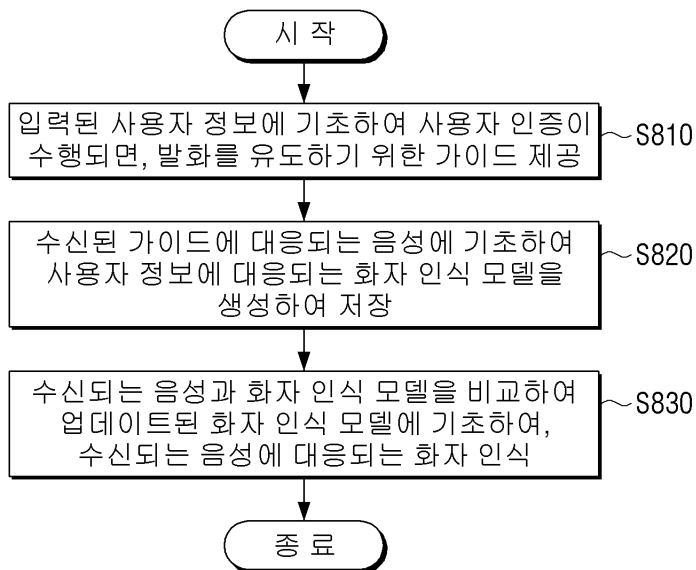
도면6



도면7

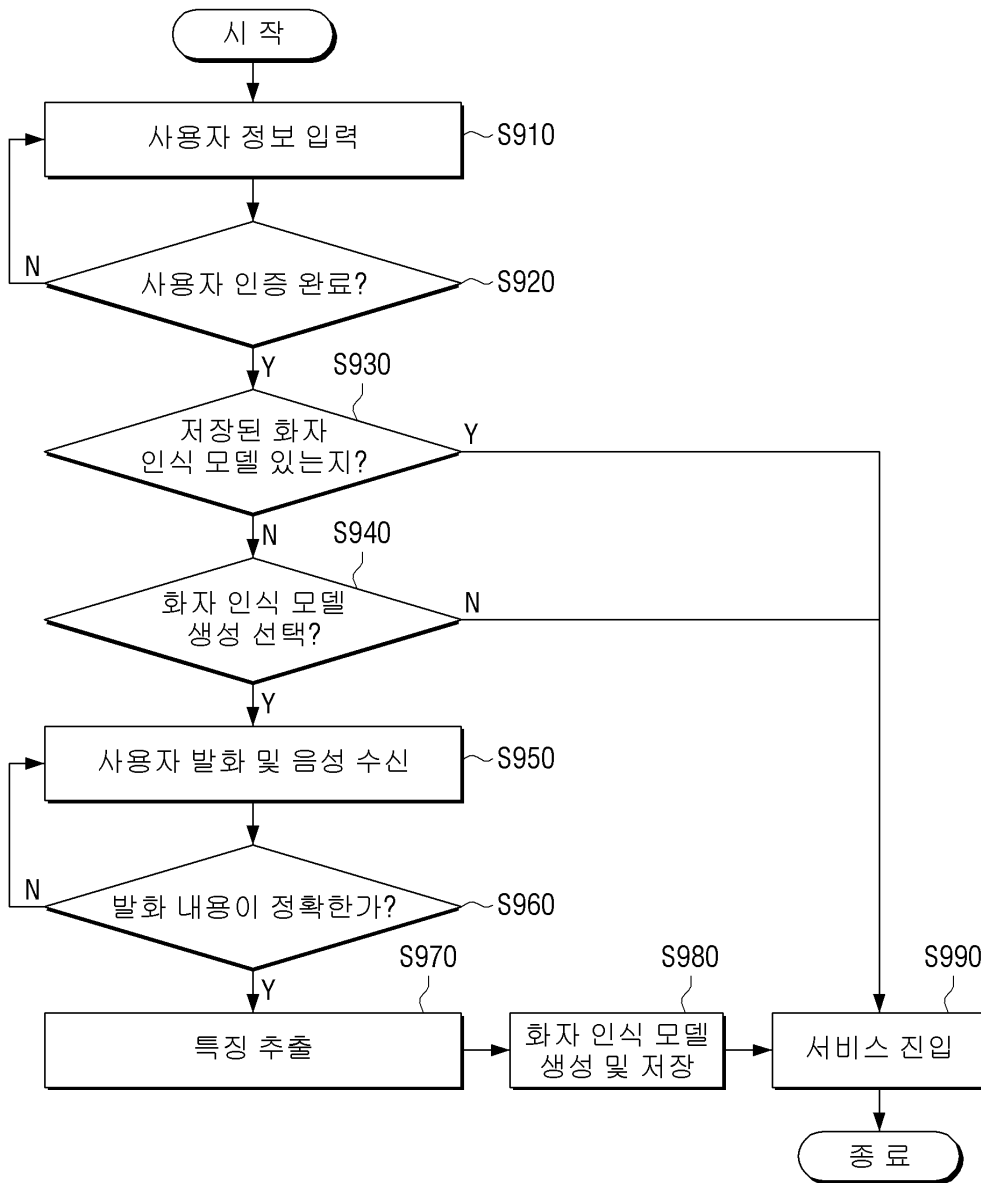


도면8

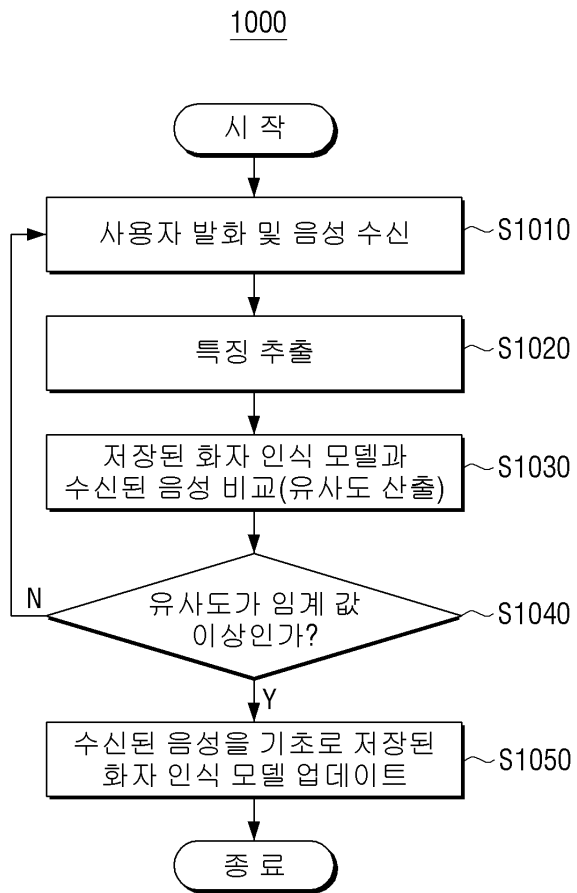


도면9

900



도면10



도면11

