



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2015년10월01일

(11) 등록번호 10-1556594

(24) 등록일자 2015년09월23일

(51) 국제특허분류(Int. Cl.)  
*G10L 15/06* (2006.01) *G10L 15/28* (2006.01)  
 (21) 출원번호 10-2009-0003026  
 (22) 출원일자 2009년01월14일  
 심사청구일자 2014년01월07일  
 (65) 공개번호 10-2010-0083572  
 (43) 공개일자 2010년07월22일  
 (56) 선행기술조사문헌  
 US06374221 B1\*  
 JP3849283 B\*  
 KR1020000032270 A\*  
 KR100657059 B1\*  
 \*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
**삼성전자 주식회사**  
 경기도 수원시 영통구 삼성로 129 (매탄동)  
 (72) 발명자  
**장종혁**  
 경기도 군포시 금산로 47, 산본2차 e-편한세상아파트 107-2305 (산본동)  
**박승권**  
 경기도 용인시 수지구 죽전로193번길 35, 반도보라아파트 108동 1802호 (죽전동)  
**이중호**  
 경기도 성남시 분당구 미금일로 22, 203동 1902호 (구미동, 까치마을)  
 (74) 대리인  
**허성원, 이동욱, 서동헌**

전체 청구항 수 : 총 16 항

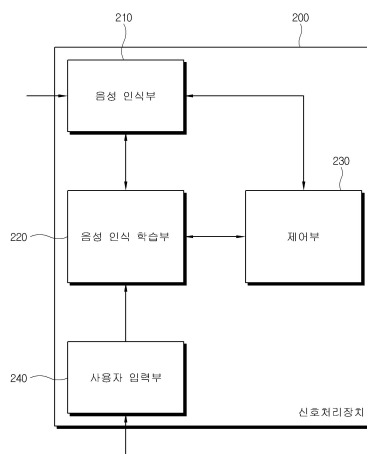
심사관 : 김응권

(54) 발명의 명칭 **신호처리장치 및 신호처리장치에서의 음성 인식 방법**

**(57) 요약**

본 발명의 일 측면에 따른 신호처리장치는, 음성을 인식하는 음성 인식부와; 음성인식의 대상인 음성을 학습하는 음성 인식 학습부와; 입력된 음성 중 상기 음성인식의 대상인 음성을 인식하도록 상기 음성 인식부를 제어하고, 상기 음성인식에 실패하는 경우에는 입력된 음성을 학습하도록 상기 음성 인식 학습부를 제어하고, 상기 음성인식에 성공하는 경우에는 인식된 음성에 대응하는 특정 동작을 수행하는 제어부를 포함하며, 이에 의할 때 의도와 달리 실패한 이전 명령들을 기억하여 최종적으로 성공한 음성 명령과 이전 명령들의 관계를 학습함으로써, 학습 이후의 시점에서는 학습 이전에 실패하였던 음성 명령도 성공적으로 활용할 수 있다.

**대표도 - 도2**



## 명세서

### 청구범위

#### 청구항 1

신호처리장치에 있어서,

사용자의 음성을 입력하는 음성 입력부와;

상기 입력된 음성과, 음성인식의 대상을 비교하여 상기 음성인식에 실패하는 경우에는 상기 사용자의 음성을 다시 입력하도록 제어하고, 상기 음성인식에 성공하는 경우에는 상기 인식된 음성에 대응하는 특정 동작을 수행하는 제어부를 포함하며,

상기 제어부는, 상기 음성인식에 실패한 음성 중에서, 상기 음성인식에 성공한 음성과의 유사도가 소정 수준 이상인 음성을 상기 음성인식의 대상에 추가하여 다음에 입력되는 음성에 대한 음성인식을 수행하는 신호처리장치.

#### 청구항 2

제1항에 있어서,

상기 음성인식의 대상인 음성을 학습하는 음성 인식 학습부를 더 포함하고,

상기 제어부는, 상기 음성인식에 성공하는 경우 상기 입력된 음성을 학습하도록 상기 음성 인식 학습부를 제어하는 신호처리장치.

#### 청구항 3

제2항에 있어서,

상기 음성 인식 학습부는, 상기 입력된 음성의 특징값 들을 상기 음성인식의 대상의 특징값과 비교하여 상기 유사도를 판단하는 신호처리장치.

#### 청구항 4

제2항에 있어서,

상기 제어부는, 상기 입력된 음성과 상기 음성인식의 대상의 유사도가 소정 수준 이상인 경우, 상기 입력된 음성의 특징을 학습하도록 상기 음성 인식 학습부를 제어하는 신호처리장치.

#### 청구항 5

제4항에 있어서,

상기 음성 인식 학습부는, acoustic 모델을 이용하여 상기 입력된 음성의 특징을 학습하는 신호처리장치.

#### 청구항 6

제2항에 있어서,

상기 제어부는, 상기 입력된 음성과 상기 음성인식의 대상의 유사도가 소정 수준 미만인 경우, 상기 입력된 음성을 상기 음성인식의 대상에 추가하도록 상기 음성 인식 학습부를 제어하는 신호처리장치.

#### 청구항 7

제6항에 있어서,

상기 음성 인식 학습부는, lexicon 모델을 이용하여 상기 입력된 음성을 상기 음성인식의 대상에 추가하는 신호처리장치.

#### 청구항 8

제1항에 있어서,

상기 사용자로부터 상기 음성인식이 실패되었는지 또는 성공되었는지 여부를 입력 받는 사용자 입력부를 더 포함하는 신호처리장치.

**청구항 9**

신호처리장치에서의 음성 인식 방법에 있어서,

사용자의 음성을 입력받는 단계와;

상기 입력된 음성과 음성인식의 대상을 비교하는 단계와;

상기 음성인식에 실패하는 경우에는 상기 사용자의 음성을 다시 입력하도록 제어하고, 상기 음성인식에 성공하는 경우에는 인식된 음성에 대응하는 특정 동작을 수행하는 단계를 포함하며,

상기 음성인식에 실패한 음성 중에서, 상기 음성인식에 성공한 음성과의 유사도가 소정 수준 이상인 음성을 상기 음성인식의 대상에 추가하여 다음에 입력되는 음성에 대한 음성인식을 수행하는 신호처리장치에서의 음성 인식 방법.

**청구항 10**

제9항에 있어서,

상기 음성인식에 성공하는 경우 상기 입력된 음성을 학습하는 신호처리장치 에서의 음성 인식 방법.

**청구항 11**

제9항에 있어서,

상기 입력된 음성의 특징값 들을 상기 음성인식의 대상의 특징값과 비교하여 상기 유사도를 판단하는 신호처리 장치에서의 음성 인식 방법.

**청구항 12**

제9항에 있어서,

상기 입력된 음성과 상기 음성인식의 대상의 유사도가 소정 수준 이상인 경우, 상기 입력된 음성의 특징을 학습 하는 신호처리장치에서의 음성 인식 방법.

**청구항 13**

제12항에 있어서,

acoustic 모델을 이용하여 상기 입력된 음성의 특징을 학습하는 신호처리장치 에서의 음성 인식 방법.

**청구항 14**

제9항에 있어서,

상기 입력된 음성과 상기 음성인식의 대상의 유사도가 소정 수준 미만인 경우, 상기 입력된 음성을 상기 음성인 식의 대상에 추가하는 신호처리장치에서의 음성 인식 방법.

**청구항 15**

제14항에 있어서,

lexicon 모델을 이용하여 상기 입력된 음성을 상기 음성인식의 대상에 추가하는 신호처리장치에서의 음성 인식 방법.

**청구항 16**

제9항에 있어서,

사용자로부터 상기 음성인식이 실패되었는지 또는 성공되었는지 여부를 입력 받는 단계를 더 포함하는 신호처리

장치에서의 음성 인식 방법.

**발명의 설명**

**발명의 상세한 설명**

**기술 분야**

[0001] 본 발명은 신호처리장치 및 신호처리장치에서의 음성 인식 방법에 관한 것으로서, 보다 상세하게는 반복적인 실수를 학습하여 이전의 실수였던 음성 명령을 성공적으로 학습할 수 있는 신호처리장치 및 신호처리장치에서의 음성 인식 방법에 관한 것이다.

**배경 기술**

[0002] 음성인식 기술에서는 입력되는 음성 명령을 더욱 정확하게 인식하는 것이 무엇보다 중요하다. 따라서 기존의 음성인식 기술들은 입력되는 음성 신호에 왜곡이 발생하는 경우에도 높은 인식률을 얻기 위해서, 다양한 입력 환경에 적응할 수 있도록 음성인식의 대상이 되는 음성을 학습한다.

[0003] 도 1은 기존의 음성인식 학습 과정을 도시한 도면이다.

[0004] 신호처리장치는 음성인식 기능을 활성화시킨다(S101).

[0005] 사용자로부터 음성을 입력받는 경우(S102), 신호처리장치는 입력된 음성을 인식한다(S103).

[0006] 이 경우, 신호처리장치는 인식한 음성을 출력하고(S104), 인식 결과에 따른 동작을 수행한다(S105). 신호처리장치는 수행되던 동작이 사용자에게 의하여 중단되었는지 판단한다(S106). 사용자가 의도하던 제어 명령에 해당하는 동작이 아닌 경우, 사용자는 신호처리장치가 수행하는 동작을 중지시킬 수 있다. 즉, 음성 인식의 결과가 사용자의 의도와 다른 경우(음성 인식이 실패할 경우), 사용자는 이를 취소하고 음성을 다시 입력할 수 있다. 따라서, 수행되던 동작이 사용자에게 의하여 중단되는 경우에는, 신호처리장치는 사용자로부터 음성 명령을 다시 입력 받는다. 반면, 수행되던 동작이 사용자에게 의하여 중단되지 않는 경우에는, 신호처리장치는 인식 결과의 신뢰도를 계산한다(S107).

[0007] 신호처리장치는 계산된 신뢰도가 일정 수준 이상인지 판단한다(S108). 일정 수준 이상인 경우, 신호처리장치는 인식한 음성을 학습한다(S109). 그러나, 신뢰도가 일정 수준 미만인 경우에는 신호처리장치는 사용자로부터 음성 명령을 다시 입력받는다.

[0008] 기존에는 음성인식 시도 시 음성인식이 성공하게 되고 그 결과가 믿을만한 수준이라고 판단되는 경우, 해당 데이터를 이용하여 성공한 음성을 재학습하였다. 예를 들어, 음성인식 결과와 음성인식 시 계산되었던 특징 벡터를 일시적으로 저장한 후, 사용자의 기기 조작 양상에 따라 인식 결과가 믿을 만하다고 판단되는 경우, 성공한 음성을 재학습에 이용하였다.

[0009] 따라서, 음성인식이 실패한 경우 또는 음성인식이 성공하였지만 신뢰도가 일정 수준 미만 이어서 음성인식 결과가 믿을만한 수준이라고 판단될 수 없는 경우에는 입력된 음성을 학습하지 않았다.

**발명의 내용**

**해결 하고자하는 과제**

[0010] 따라서, 본 발명의 목적은 의도와 달리 실패한 이전 명령들을 학습함으로써, 반복적인 실수를 학습하여 이전의 실수였던 음성 명령을 성공적으로 학습할 수 있는 신호처리장치 및 신호처리장치에서의 음성 인식 방법을 제공하는 것이다.

**과제 해결수단**

[0011] 본 발명의 일 측면에 따른 신호처리장치는, 음성을 인식하는 음성 인식부와; 음성인식의 대상인 음성을 학습하는 음성 인식 학습부와; 입력된 음성 중 상기 음성인식의 대상인 음성을 인식하도록 상기 음성 인식부를 제어하고, 상기 음성인식에 실패하는 경우에는 입력된 음성을 학습하도록 상기 음성 인식 학습부를 제어하고, 상기 음성인식에 성공하는 경우에는 인식된 음성에 대응하는 특정 동작을 수행하는 제어부를 포함할 수 있다.

[0012] 상기 신호처리장치에 있어서, 상기 제어부는, 상기 음성인식에 성공하는 경우 상기 입력된 음성을 학습하도록 상기 음성 인식 학습부를 제어할 수 있다. 상기 신호처리장치에 있어서, 상기 음성 인식 학습부는, 상기 입력된 음성의 특징값 들을 상기 음성인식의 대상인 음성의 특징값과 비교하여 유사도를 판단할 수 있다. 상기 신호처리장치에 있어서, 상기 제어부는, 상기 입력된 음성과 상기 음성인식의 대상인 음성의 유사도가 소정 수준 이상인 경우, 상기 입력된 음성의 특징을 학습하도록 상기 음성 인식 학습부를 제어할 수 있다. 상기 신호처리장치에 있어서, 상기 음성 인식 학습부는, acoustic 모델을 이용하여 상기 입력된 음성의 특징을 학습할 수 있다. 상기 신호처리장치에 있어서, 상기 제어부는, 상기 입력된 음성과 상기 음성인식의 대상인 음성의 유사도가 소정 수준 미만인 경우, 상기 입력된 음성을 상기 음성인식의 대상에 추가하도록 상기 음성 인식 학습부를 제어할 수 있다. 상기 신호처리장치에 있어서, 상기 음성 인식 학습부는, lexicon 모델을 이용하여 상기 입력된 음성을 상기 음성인식의 대상에 추가할 수 있다. 상기 신호처리장치는, 사용자로부터 상기 음성인식이 실패되었는지 또는 성공되었는지 여부를 입력받는 사용자 입력부를 더 포함할 수 있다.

[0013] 한편 본 발명의 다른 측면에 따른 신호처리장치에서의 음성 인식 방법은, 음성을 입력받는 단계와; 입력된 상기 음성 중 음성인식의 대상인 음성을 인식하는 단계와; 상기 음성인식에 실패하는 경우에는 입력된 음성을 학습하고, 상기 음성인식에 성공하는 경우에는 인식된 음성에 대응하는 특정 동작을 수행하는 단계를 포함할 수 있다.

[0014] 상기 신호처리장치에서의 음성 인식 방법에 있어서, 상기 음성인식에 성공하는 경우 상기 입력된 음성을 학습할 수 있다. 상기 신호처리장치에서의 음성 인식 방법에 있어서, 상기 입력된 음성의 특징값들을 상기 음성인식의 대상인 음성의 특징값과 비교하여 유사도를 판단할 수 있다. 상기 신호처리장치에서의 음성 인식 방법에 있어서, 상기 입력된 음성과 상기 음성인식의 대상인 음성의 유사도가 소정 수준 이상인 경우, 상기 입력된 음성의 특징을 학습할 수 있다. 상기 신호처리장치에서의 음성 인식 방법에 있어서, acoustic 모델을 이용하여 상기 입력된 음성의 특징을 학습할 수 있다. 상기 신호처리장치에서의 음성 인식 방법에 있어서, 상기 입력된 음성과 상기 음성인식의 대상인 음성의 유사도가 소정 수준 미만인 경우, 상기 입력된 음성을 상기 음성인식의 대상에 추가할 수 있다. 상기 신호처리장치에서의 음성 인식 방법에 있어서, lexicon 모델을 이용하여 상기 입력된 음성을 상기 음성인식의 대상에 추가할 수 있다. 상기 신호처리장치에서의 음성 인식 방법은, 사용자로부터 상기 음성인식이 실패되었는지 또는 성공되었는지 여부를 입력받는 단계를 더 포함할 수 있다.

**효 과**

[0015] 본 발명에 따르면 반복적인 실수를 학습하여 이전의 실수였던 음성 명령을 성공적으로 학습할 수 있다. 또한, 의도와 달리 실패한 이전 명령들을 기억하여, 최종적으로 성공한 음성 명령과 이전 명령들의 관계를 학습함으로써, 학습 이후의 시점에서는 학습 이전에 실패하였던 음성 명령도 성공적으로 활용할 수 있다.

**발명의 실시를 위한 구체적인 내용**

[0016] 이하, 첨부한 도면을 참고로 하여 본 발명의 실시예들에 대하여 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자가 용이하게 실시할 수 있도록 상세히 설명한다. 본 발명은 여러 가지 상이한 형태로 구현될 수 있으며 여기에서 설명하는 실시예들에 한정되지 않는다. 본 발명을 명확하게 설명하기 위해서 설명과 관계없는 부분은 생략하였으며, 명세서 전체를 통하여 동일 또는 유사한 구성 요소에 대해서는 동일한 참조 부호를 붙이도록 한다.

[0017] 도 2는 본 발명의 일 실시예에 따른 신호처리장치의 구성을 도시한 도면이다.

[0018] 본 발명의 일 실시예에 따른 신호처리장치(200)는 디지털 TV, 데스크 탑 컴퓨터, 노트북, 셋탑 박스, DVD(Digital Versatile Disc)/BD(Blu-ray Disc) 플레이어, DVD 레코더 등이 될 수 있다. 나아가, 음성 명령을 인식하고 이에 대응하는 동작을 수행할 수 있는 전자 디바이스라면, 본 발명의 일 실시예에 따른 신호처리장치(200)가 될 수 있다.

[0019] 본 발명의 일 실시예에 따른 신호처리장치(200)는 음성 인식부(210), 음성 인식 학습부(220), 제어부(230) 및 사용자 입력부(240)를 포함할 수 있다.

[0020] 음성 인식부(210)는 음성을 인식한다. 이 경우, 음성 인식부(210)는 입력된 음성을 양자화하여 특징값 들을 추출하고, 이를 음성인식 대상 단어의 특징값과 매칭하여 비교할 수 있다.

[0021] 음성 인식 학습부(220)는 음성인식의 대상인 음성을 학습한다. 이 경우, 음성인식 학습부(220)는 제어부(230)의 제어에 의하여, 음성인식에 실패한 입력 음성을 학습할 수 있다.

- [0022] 사용자의 의도와 다르게 음성인식이 실패하였을 경우, 사용자는 틀렸던 이전 음성 명령과 다른 방법으로 소리를 내거나(예를 들어, 음성 명령을 크게 말하거나, 혹은 또렷하게 말함) 또는 다른 변형된 음성 명령(몇 음절 또는 몇 단어가 다른 음성 명령)을 하게 된다. 따라서, 음성 인식 학습부(220)는 다음과 같이 실패한 시도를 학습하여, 반복적인 실수를 학습한다.
- [0023] 첫째, 음성인식 시 사용자 목소리의 톤, 크기, 억양, 사투리 등으로 인식 오류가 발생할 수 있다. 이러한 변수로 인하여 생긴 인식 오류를 해결하기 위하여, 음성 인식 학습부(220)는 사용자의 음성인식 시도 시의 정오 패턴 및 입력되는 음성 명령어들간의 유사도를 계산하여, 학습 이전에 틀린 것으로 분류되던 음성 명령어들과 정답인 음성 명령어간의 차이를 음성 인식의 음소 모델 재학습시 활용하여 자동화된 사용자 적응형 음성인식을 수행할 수 있다.
- [0024] 둘째, 기기 조작을 위하여 음성으로 명령어를 말하는 경우, 사용자의 습관에 따라서 혹은 명령어의 형태나 선택한 단어 또는 선택한 음절 수에 따라, 사용자는 음성인식의 대상인 명령어와 다른 명령어를 말할 수 있다. 예를 들어 TV의 볼륨을 높이는 명령어가 "볼륨 크게" 로 지정되어 있는 경우, 사용자는 무의식적으로 "볼륨 업" 또는 "소리 크게" 등으로 명령할 수 있다. 이 경우 음성 인식 오류가 발생하며, 사용자는 여러 번의 시행착오를 통해 기기에서 원하는 명령어를 하게 될 것이다. 혹은 기기에서 원하는 명령어를 말하지 못할 수도 있다. 음성 인식 학습부(220)는 이러한 경우 사용자가 명령했던 명령어들을 저장해놓은 후 유사도를 비교하여 사용자가 무의식적으로 발음한 "볼륨 업" 또는 "소리 크게" 라는 단어도 "볼륨 크게" 라는 명령어에 매칭되도록 한다.
- [0025] 이 경우, 음성 인식 학습부(220)는 Acoustic 모델 또는 Lexicon 모델에 따라 음성을 학습할 수 있다. 구체적으로, 음성 인식 학습부(220)는, 입력된 음성과 음성인식의 대상인 음성의 유사도가 소정 수준 이상인 경우에는 Acoustic 모델을 이용하여 입력된 음성의 특징을 학습할 수 있다. 또한, 음성 인식 학습부(220)는, 입력된 음성과 음성인식의 대상인 음성의 유사도가 소정 수준 미만인 경우에는 Lexicon 모델을 이용하여 입력된 음성을 음성인식의 대상에 추가할 수 있다.
- [0026] 이를 위해, 음성 인식 학습부(220)는 입력된 음성의 특징값 들을 음성인식의 대상인 음성의 특징값과 비교하여 유사도를 판단할 수 있다. 또는 음성 인식 학습부(220)는 한 음성과 음성인식 대상 단어와의 유사값을 구하여 유사도를 판단할 수도 있다.
- [0027] 한편, 음성 인식 학습부(200)는 음성 인식 결과 및 음성 인식 학습용 데이터를 저장하기 위한 메모리를 포함할 수 있다. 이 경우, 음성 인식 학습부(200)는 음성 인식 결과 및 음성 인식 학습용 데이터를 특정 형태에 맞추어 메모리 또는 디스크에 저장할 수 있다. 이에 대해서는 도 7a 내지 도 7c에서 상세히 설명한다.
- [0028] 제어부(230)는 입력된 음성 중 음성인식의 대상인 음성을 인식하도록 음성 인식부(210)를 제어할 수 있다. 또한, 제어부(230)는 음성인식에 성공하는 경우 인식된 음성에 대응하는 특정 동작을 수행할 수 있다.
- [0029] 한편, 제어부(230)는 음성인식에 실패하는 경우 입력된 음성을 학습하도록 음성 인식 학습부(220)를 제어할 수 있다. 이 경우, 제어부(230)는 입력된 음성과 음성인식의 대상인 음성의 유사도가 소정 수준 이상인 경우, 입력된 음성의 특징을 학습하도록 음성 인식 학습부(220)를 제어할 수 있다. 반면, 제어부(230)는 입력된 음성과 음성인식의 대상인 음성의 유사도가 소정 수준 미만인 경우, 입력된 음성을 음성인식의 대상에 추가하도록 음성 인식 학습부(220)를 제어할 수 있다.
- [0030] 나아가, 제어부(230)는 음성음식에 실패하는 경우뿐 아니라, 음성인식에 성공하는 경우에도 입력된 음성을 학습하도록 음성 인식 학습부(220)를 제어할 수 있다.
- [0031] 사용자 입력부(240)는 사용자로부터 음성인식이 실패되었는지 또는 성공되었는지 여부를 입력받는다. 즉, 사용자 입력부(240)는 사용자의 피드백을 입력받는다. 이 경우, 사용자의 피드백을 받는 구체적인 방법은 다양하게 구성될 수 있다. 예를 들어, 음성을 입력한 후 취소에 해당하는 명령을 일정 시간 내에 내리거나 음성 인식 결과가 틀렸음을 나타내는 입력 키를 누르는 경우 또는 일정 시간 동안 사용자의 입력이 없는 경우, 신호처리장치(200)는 음성인식이 실패한 것으로 판단할 수 있다.
- [0032] 도 3은 본 발명의 일 실시예에 따른 신호처리장치에서의 음성 인식 과정을 도시한 도면이다.
- [0033] 사용자가 음성을 입력하는 경우(S301), 신호처리장치(200)는 음성을 입력받는다. 이 경우, 신호처리장치(200)는 음성 인식 결과를 출력한다(S302).
- [0034] 신호처리장치(200)는 음성 인식 결과가 믿을 만한 수준의 결과인지 판단한다 (S303). 음성 인식 결과가 믿을 만한 수준의 결과인 경우(S303-Y), 즉 신뢰도가 일정 수준 이상인 경우, 신호처리장치(200)는 음성 인식 결과를



수행한다(S304). 구체적으로, 신호처리장치(200)는 인식된 음성에 대응하는 특정 동작을 수행할 수 있다.

[0035] 신호처리장치(200)는 사용자가 해당 동작을 취소하였는지 판단한다(S305). 만일 사용자가 해당 동작을 취소하지 않은 경우에는, 신호처리장치(200)는 음성 인식 결과를 저장한다(S306). 구체적으로, 신호처리장치(200)는 음성 인식 결과와 음성 인식 학습용 데이터를 특정 형태에 맞추어 메모리 또는 디스크에 저장할 수 있다. 이 경우, 신호처리장치(200)는 저장된 음성 인식 결과를 기초로 재학습을 수행한다(S307).

[0036] 한편, S303 단계에서 음성 인식 결과가 믿을 만한 수준의 결과가 아니라고 판단되거나(S303-N) S305 단계에서 사용자가 해당 동작을 취소한 경우에는(S305-Y), 신호처리장치(200)는 S316 단계로 이행한다. 이 경우, 신호처리장치(200)는 음성 인식 결과를 저장하고(S316), 저장된 음성 인식 결과를 기초로 재학습을 수행한다(S317). 또한, S301 단계로 되돌아가 사용자로부터 음성을 다시 입력받는다. 음성 인식 결과가 믿을 만한 수준이 아니라고 판단되거나 사용자가 해당 동작을 취소한 경우에는 음성 인식 결과를 수행하지 않기 때문에, 다시 사용자가 의도하는 제어 명령에 대응되는 음성을 입력받아야 하기 때문이다.

[0037] 이렇게 본 발명에 따르면, 기존과는 달리 음성 인식이 실패한 경우, 즉 음성 입력 결과가 일정 신뢰 수준 이하이거나 사용자가 해당 음성 인식 결과에 대한 동작을 취소한 경우에도, 해당 음성 인식 결과를 재학습한다. 즉, 의도와 달리 실패한 음성 명령들을 학습함으로써, 학습 이후의 시점에서는 학습 이전에 실패하였던 음성 명령도 성공적으로 활용할 수 있다.

[0038] 도 4는 본 발명의 일 실시예에 따른 신호처리장치에서의 음성 학습 과정을 도시한 도면이다.

[0039] 상기 음성 학습 과정에서는, 실패한 시도와 성공한 시도의 음성 특징값이 유사하다면, 사용자가 성공한 음성과 똑같은 단어를 말하였지만 억양이나, 톤, 볼륨 등으로 인하여 실패한 단어라고 가정한다. 따라서, 이 경우에는 Acoustic 모델에 의하여 실패한 시도의 음성 데이터를 학습한다.

[0040] 반면, 실패한 시도와 성공한 시도의 음성 특징값이 유사하지 않다면, 사용자가 성공한 음성과 동일한 목적으로 말하였지만 다른 단어를 말하였다고 가정한다. 따라서, 이 경우에는 Lexicon 모델에 의하여 실패한 시도의 음성 데이터를 학습한다. 구체적으로, Lexicon 모델의 음성 인식의 대상에 실패한 시도의 음성 특징값을 갖는 단어를 추가한다.

[0041] 구체적으로, 음성 인식이 단 한번에 성공한 경우에는 Acoustic 모델을 이용하여 성공한 시도의 음성 데이터를 학습하고, 두 번 이상의 시도 후에 음성 인식에 성공한 경우에는 실패한 시도의 음성 특징값과 성공한 시도의 음성 특징값과의 유사도에 따라, Acoustic 모델 또는 Lexicon 모델에 따라 음성 데이터를 학습한다.

[0042] 음성을 학습하는 경우, 가장 먼저 신호처리장치(200)는 음성인식을 시도한 총 시도회수가 1인지 판단한다(S401).

[0043] 음성인식을 시도한 총 시도회수가 1인 경우(S401-Y), 신호처리장치(200)는 Acoustic 모델을 이용하여 성공한 시도의 음성 특징값을 학습한다(S414). 반면 음성 인식을 시도한 총 시도회수가 1이 아니라고 판단되면(S401-N), 신호처리장치(200)는 음성 인식 결과값에서 실패한 시도의 음성 특징값과 성공한 시도의 음성 특징값과의 유사도를 비교한다(S402).

[0044] 신호처리장치(200)는 실패한 시도의 음성 특징값과 성공한 시도의 음성 특징값이 유사한지 판단한다(S403). 만일, 실패한 시도의 음성 특징값과 성공한 시도의 음성 특징값이 유사하면(S403-Y), 신호처리장치(200)는 Acoustic 모델을 이용하여 실패한 시도의 음성 특징값을 학습한다(S404). 반면, 실패한 시도의 음성 특징값과 성공한 시도의 음성 특징값이 유사하지 않으면(S403-N), 신호처리장치(200)는 Lexicon 모델의 음성 인식의 대상에 실패한 시도의 음성 특징값을 갖는 단어를 추가한다(S424).

[0045] 본 실시예에서는 음성 인식에 실패하는 경우에도, Acoustic 모델 또는 Lexicon 모델에 따라 음성을 학습한다. Acoustic 모델에 따라 음성을 학습하는 경우, 사용자의 발음 및 억양으로 인한 인식 오류를 알아내어 사용자의 고유 발음을 재학습할 수 있다. Lexicon 모델에 따라 음성을 학습하는 경우, 사용자가 명령어를 제대로 기억하지 못하여 같은 의미로 다른 명령어를 말하였을 경우에도 해당 단어를 인식 대상 단어에 추가하여 추후에 인식이 가능하도록 할 수 있다.

[0046] 도 5는 Acoustic 모델에 따른 음성 학습 과정을 도시한 도면이다.

[0047] 신호처리장치(200)가 디지털 TV로 구현되고, 출력되는 오디오의 볼륨을 크게 하는 명령어가 "볼륨 크게" 로 정해져 있다고 가정할 때, 음성 인식이 실패하는 첫 번째 시나리오는 다음과 같다.

- [0048] 사용자가 음성인식의 대상인 음성과 똑같은 "볼륨 크게" 라는 명령어를 말한다. 그러나, 억양이나 톤, 볼륨 등의 차이가 발생하여, 신호처리장치(200)는 사용자의 음성을 다른 단어로 오인식하거나 또는 인식에 실패할 수 있다.
- [0049] 사용자는 오인식의 경우 음성 인식 결과를 취소하고 다시 한번 "볼륨 크게" 를 말한다. 인식에 실패했을 경우에도 다시 "볼륨 크게" 를 말한다. 음성 인식에 실패한 후 다시 음성 인식을 시도하게 되면, 사용자는 일반적으로 첫 번째 시도시의 음성보다 더욱 크고 분명한 발음으로 명령어를 말하게 된다. 신호처리장치(200)는 재입력되어진 음성을 "볼륨 크게" 로 정상적으로 인식하게 되고, 상기 명령어에 대응하는 동작을 수행하게 된다.
- [0050] 본 발명에서는 상기와 같은 시나리오의 경우, Acoustic 모델에 따라 음성을 학습한다. 즉, 신호처리장치(200)는 처음으로 입력되었으나 오인식 또는 인식 실패 한 "볼륨 크게" 라는 발음과 인식에 성공한 "볼륨 크게" 와의 유사도를 비교하고, 두 발음 사이에 유사도가 일정 수준 이상이라면 사용자가 같은 단어를 말하였지만 발음이나 억양 차로 인하여 인식이 실패한 것으로 간주한다. 이 경우, 신호처리장치 (200)는 인식에 실패한 "볼륨 크게" 의 음소 모델을 분석하여 이를 학습한다.
- [0051] Acoustic 모델에 따른 음성 학습 과정에서는, 은닉 마코브 모델링(Hidden Markov Modeling: HMM) 방법을 사용할 수 있다. 은닉 마코브 모델링 방법에 의하면, 음성의 특징값을 계산하여 특정 음소가 나올 확률적인 통계값을 구함으로써 어떤 음소가 발음되었는지 구분한다. 하지만 일반적으로 말을 할 때 앞의 음소와 뒤의 음소가 현재 음소의 발음에 영향을 주게 되며, 이러한 이유로 은닉 마코브 모델링 방법에서는 tri-phonemes라는 개념을 적용하여 앞, 뒤, 현재 음소 3 가지를 조합하여 음소를 구성한다.
- [0052] 신호처리장치(200)는 Acoustic 모델 학습 자료를 구성한다(S501). 이 경우, 신호처리장치(200)는 구성된 상기 Acoustic 모델 학습 자료를 기초로, Tri-phonemes를 구성한다(S502).
- [0053] 신호처리장치(200)는 Tri-phonemes와 음소값을 매칭시킨다(S503). 즉, 신호처리장치(200)는 입력 음성의 음소와 상기 Tri-phonemes의 음소값을 비교하고, 이를 새로운 음성인식의 대상으로 학습한다.
- [0054] 이 경우, 신호처리장치(200)는 학습된 결과를 반영하여, Acoustic 모델 데이터의 해당 Tri-phonemes값을 변경한다(S504).
- [0055] 기존에는, 성공한 시도의 음성만을 학습하였다. 따라서, 사용자가 같은 단어를 말하였지만 발음이나 억양 차로 음성 인식에 실패한 경우, 입력된 음성은 학습되지 않았다. 본 발명에 의하면, 이러한 경우에도 입력된 음성을 Acoustic 모델에 의하여 학습한다. 따라서, 음성 인식의 정확도를 높일 수 있다.
- [0056] 도 6은 Lexicon 모델에 따른 음성 학습 과정을 도시한 도면이다.
- [0057] 도 5와 동일한 가정하에서, 음성 인식이 실패하는 두 번째 시나리오는 다음과 같다.
- [0058] 출력되는 오디오의 볼륨을 크게 하는 명령어가 "볼륨 크게" 로 정해져 있는 상황에서, 사용자는 "소리 크게" 라는 명령어를 말한다. 이 경우, 신호처리장치 (200)는 사용자의 음성을 다른 단어로 오인식하거나. 혹은 인식에 실패하게 된다.
- [0059] 사용자는 오인식의 경우 음성 인식 결과를 취소하고, 다시 한번 더 크고 분명한 목소리로 "소리 크게" 를 말하게 된다. 인식에 실패했을 경우에도, 다시 "소리 크게" 를 말한다. 하지만 이 명령은 다시 실패하게 되며, 이 경우 사용자는 다른 명령어의 조합도 시도할 수 있다. 예를 들어 "볼륨 업", "볼륨 크게" 등 똑같은 의미의 다른 명령어를 순차적으로 말할 수 있으며, 이때 "볼륨 크게" 라는 명령어를 말하게 되면 음성 인식은 성공한다.
- [0060] 본 발명에서는 상기와 같은 시나리오의 경우, Lexicon 모델에 따라 음성을 학습한다. 즉, 신호처리장치(200)는 처음 입력되었으나 오인식 또는 인식 실패하였던 "소리 크게" 라는 발음과 인식에 성공한 "볼륨 크게" 와의 유사도를 비교한다. 이 경우, 두 단어의 발음상의 유사도는 일정 수준을 넘지 못하기 때문에, 신호처리장치(200)는 사용자가 같은 의도의 말을 하였지만 다른 단어를 말하여 인식에 실패한 것으로 간주하고, 인식에 실패한 "소리 크게" 를 단어장에 추가한 후 "볼륨 크게" 에 대응하는 제어 동작과 연결시킨다.
- [0061] Lexicon 모델에 따라 음성을 학습하고자 하는 경우, 신호처리장치(200)는 현재 최종 인식 단어와 같은 튜플들을 Lexicon 모델 학습용 데이터에서 모두 가져온다(S601). 상기 Lexicon 모델 학습용 데이터에 대해서는 도 7c에서 후술한다.
- [0062] 신호처리장치(200)는 튜플들의 음성 특징값과 현재 입력된 음성 특징값과의 유사도를 구한다(S602). 이 경우, 현재 입력된 음성 특징값이 이전에 입력되었던 회수를 판단하여 유사도를 구할 수 있다. 예를 들어 '볼륨 크게'



라는 최종 인식 단어에 대하여 '볼륨 업' 이라고 말했다면, 이전에도 똑같이 '볼륨 업' 이라고 말한 적이 있는지 판단한다. 이 경우, 현재 입력된 음성 특징값이 이전에 입력되었던 회수가 높을수록, 유사도가 높은 것으로 판단한다.

[0063] 신호처리장치(200)는 각각의 튜플에 대하여, 유사도가 일정 수준 이상인지 판단한다(S603). 구체적으로, 최종 인식 단어가 아닌 다른 명령어를 계속 입력하는 시도가 N번 이상인지 판단할 수 있다.

[0064] 유사도가 일정 수준 이상인 경우(S603-Y), 신호처리장치(200)는 실패한 시도의 음성 특징값을 Lexicon 모델의 음성 인식의 대상에 추가한다(S604). 예를 들어, '볼륨 업' 이라는 단어의 음소(Phonemes)도 '볼륨 크게' 에 매칭되도록 음성 인식 대상 단어에 추가한다. 이 경우, 신호처리장치(200)는 lexicon 모델 학습용 데이터에서 해당 튜플을 삭제한다(S605).

[0065] 한편, S603 단계에서 유사도가 일정 수준 이상이 아니라고 판단되면(S603-N), 신호처리장치(200)는 입력된 음성 특징값을 Lexicon 모델 학습용 데이터에 추가한다(S606).

[0066] 이렇게 Lexicon 모델에 따라 음성을 학습하는 경우, 사용자가 무의식적으로 또는 습관에 의하여 '볼륨 크게' 라는 명령어를 '볼륨업' 또는 '소리 크게' 라고 발음하는 경우, 해당 단어를 Lexicon 모델에 따른 인식대상 단어에 추가함으로써 '볼륨 업', '소리 크게' 라는 명령어도 인식할 수 있다.

[0067] 도 7a는 저장된 음성 인식 결과의 일례를 도시한 도면이다.

[0068] 각각의 시도 회수, 특징값, 성공/실패가 하나의 튜플(Tuple)(710)을 이루며, 최초 음성 인식 시도부터 최초 음성 인식 성공 및 특정 동작 수행까지가 하나의 테이블로 관리된다. 성공/실패 중 성공은 음성 인식 결과가 신뢰 수준 이상이고, 사용자의 취소 요청도 없는 경우를 말한다. 도 7a에 도시된 테이블을 참조하면, 제n번째 시도에서 음성 인식이 성공할 때까지 입력된 음성의 특징값 들이 차례대로 저장되어 있는 것을 알 수 있다.

[0069] 도 7b는 저장된 Acoustic 모델 학습용 데이터의 일례를 도시한 도면이다.

[0070] 상기 Acoustic 모델 학습용 데이터는 음소 데이터의 특징값들을 저장한다. 예를 들어 ㄱ, ㄴ, ㄷ, ..., ㅏ, ㅑ, ㅓ 등 모든 음소에 대하여 음성 특징값을 매칭하여 저장할 수 있다.

[0071] 도 7b에서는 각각의 최종 인식 단어와, 최종 인식 시 음성 특징값, 학습시킬 음성 특징값이 하나의 튜플(720)을 이룬다. 이 경우, 최종 인식 시 음성 특징값과 학습시킬 음성 특징값은 일정 수준 이상의 유사도를 가진다.

[0072] 도 7c는 저장된 Lexicon 모델 학습용 데이터의 일례를 도시한 도면이다.

[0073] 상기 Lexicon 모델 학습용 데이터는 음성에서 음소를 찾아내어 특정 단어에 매칭시킨다. 구체적으로, 특정 어구 또는 단어와 음소를 매칭시켜 저장할 수 있다. 예를 들어, "맛있다" 라는 단어를 "ma-sit-da" 라는 음소 또는 "mat-it-da" 라는 음소에 매칭시켜 저장할 수 있다. 즉, Lexicon 모델 학습용 데이터는 이와 같이 특정 단어가 발음되는 형태를 정의하여 저장한다.

[0074] 도 7c에 도시된 테이블에서는 각각의 최종 인식 단어와 음성 특징값이 하나의 튜플(730)을 이루게 된다.

[0075] 비록 본 발명의 몇몇 실시예들이 도시되고 설명되었지만, 본 발명이 속하는 기술분야의 통상의 지식을 가진 당업자라면 본 발명의 원칙이나 정신에서 벗어나지 않으면서 본 실시예를 변형할 수 있음을 알 수 있을 것이다. 발명의 범위는 첨부된 청구항과 그 균등물에 의해 정해질 것이다.

**도면의 간단한 설명**

[0076] 도 1은 기존의 음성 인식 학습 과정을 도시한 도면.

[0077] 도 2는 본 발명의 일 실시예에 따른 신호처리장치의 구성을 도시한 도면.

[0078] 도 3은 본 발명의 일 실시예에 따른 신호처리장치에서의 음성 인식 과정을 도시한 도면.

[0079] 도 4는 본 발명의 일 실시예에 따른 신호처리장치에서의 음성 학습 과정을 도시한 도면.

[0080] 도 5는 Acoustic 모델에 따른 음성 학습 과정을 도시한 도면.

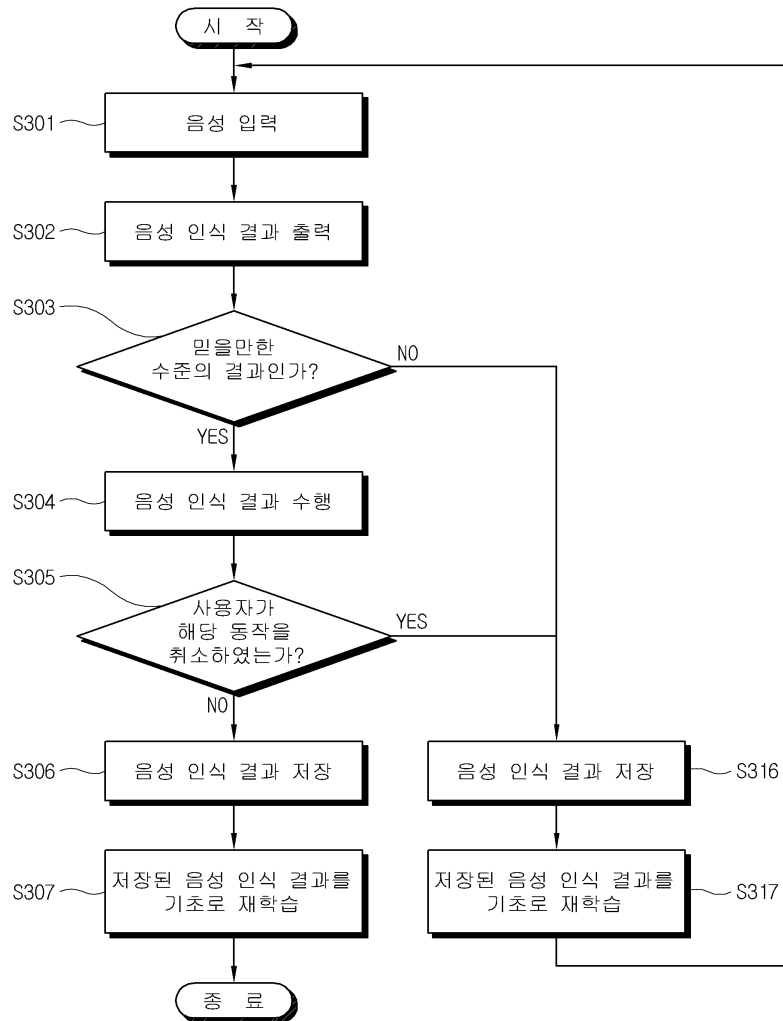
[0081] 도 6은 Lexicon 모델에 따른 음성 학습 과정을 도시한 도면.

[0082] 도 7a는 저장된 음성 인식 결과의 일례를 도시한 도면.

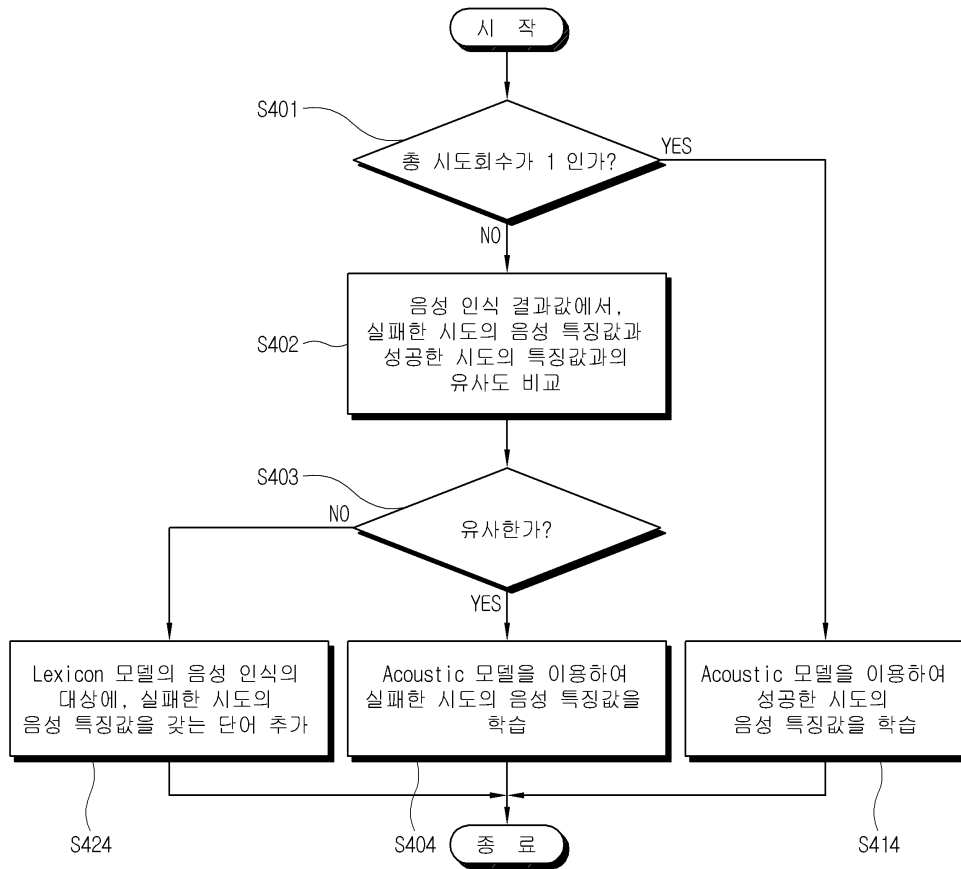




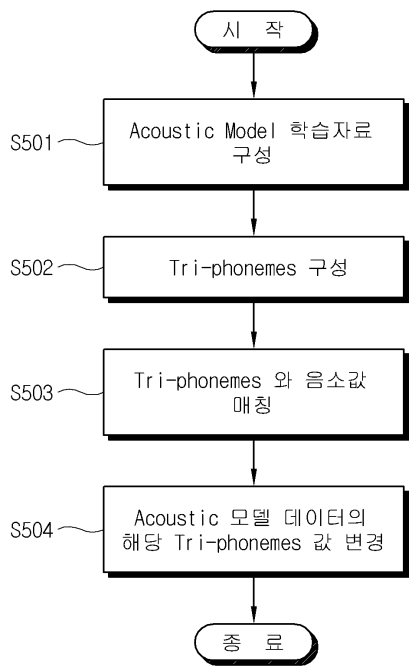
도면3



도면4

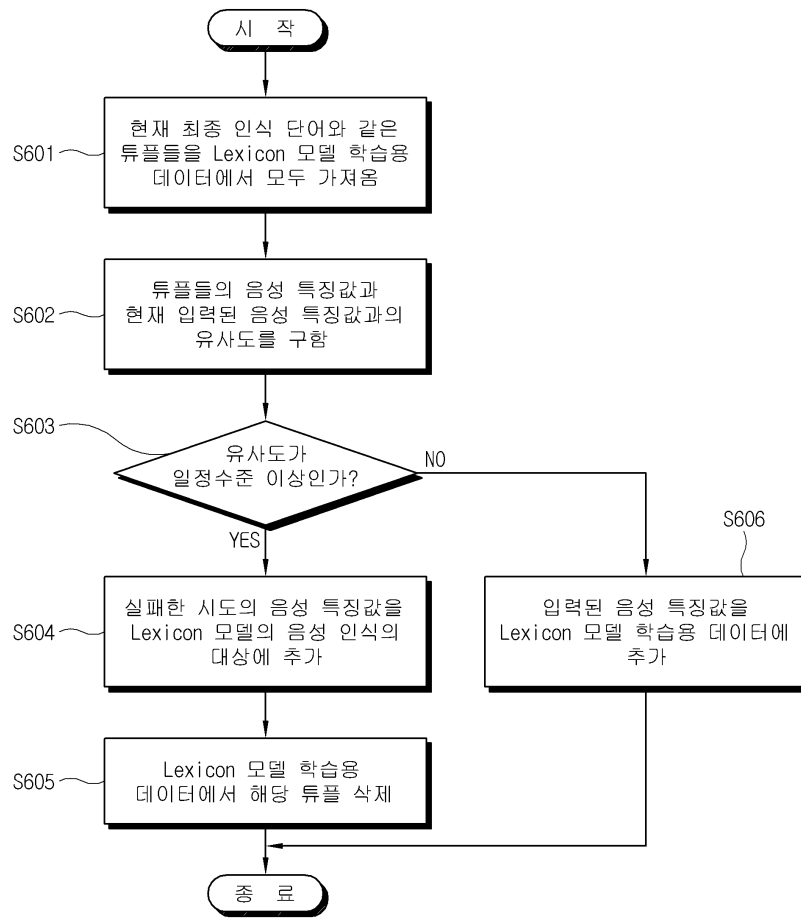


도면5





도면6



도면7a

시도 회수	특징값	성공/실패
1	<Voice Feature Data>	실패
2	<Voice Feature Data>	실패
...	...	실패
n	<Voice Feature Data>	성공

도면7b

최종 인식 단어	최종 인식시 Voice Feature Data	학습시킬 Voice Feature Data
'불룸'	<Voice Feature Data>	<Voice Feature Data>
'TV'	<Voice Feature Data>	<Voice Feature Data>
...	...	...

720

도면7c

최종 인식 단어	Voice Feature Data
'불룸'	<Voice Feature Data>
'TV'	<Voice Feature Data>
...	...

730