



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2016년01월05일

(11) 등록번호 10-1581816

(24) 등록일자 2015년12월24일

(51) 국제특허분류(Int. Cl.)

G10L 15/04 (2006.01) G09B 7/00 (2006.01)

G10L 15/22 (2006.01)

(21) 출원번호 10-2014-0138383

(22) 출원일자 2014년10월14일

심사청구일자 2014년10월14일

(56) 선행기술조사문헌

KR1020080024911 A\*

KR1020080072673 A\*

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

서강대학교산학협력단

서울특별시 마포구 백범로 35 (신수동, 서강대학교)

강원대학교산학협력단

강원도 춘천시 강원대학길 1 (효자동)

(72) 발명자

서정연

서울특별시 서초구 서초중앙로 260 B동 203호 (반포동, 삼호가든3차)

강상우

서울특별시 성북구 길음로 16 612동 1602호

(뒷면에 계속)

(74) 대리인

이지연

전체 청구항 수 : 총 5 항

심사관 : 정성윤

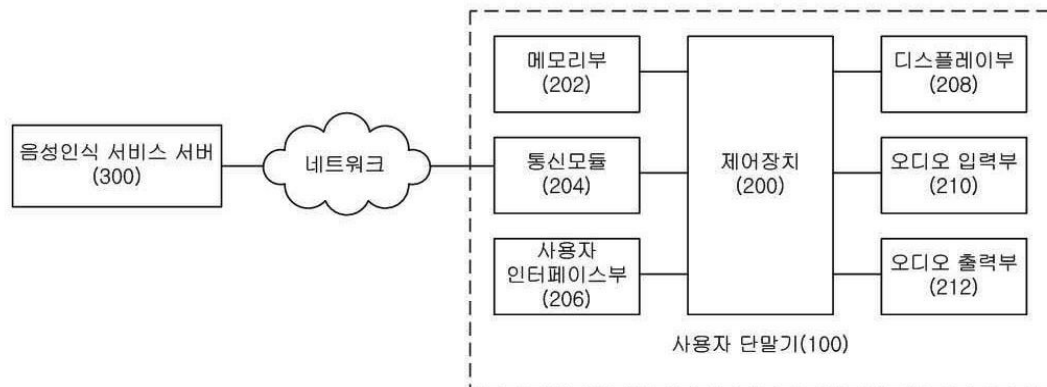
(54) 발명의 명칭 기계학습을 이용한 음성인식방법

(57) 요약

본 발명에 따르는 기계학습을 이용한 음성인식방법은, 음성신호를 입력받아 음성인식하여 다수의 후보문장정보 및 각 후보문장정보에 대한 순위를 포함하는 음성인식결과정보를 생성하는 단계; 상기 음성인식결과정보에 포함된 다수의 후보문장정보를 미리 학습된 기계학습모델에 따라 처리하여 상기 다수의 후보문장정보의 순위를 가변

(뒷면에 계속)

대표도



하여 재순위화하는 단계; 및 재순위화된 다수의 후보문장정보 중 최상위 순위의 후보문장정보를 음성인식결과로 선정하는 단계;를 포함하는 것을 특징으로 한다.

특히, 상기 기계학습모델은, 음성신호 및 정답문장정보를 입력받고, 상기 음성신호에 대해 음성인식하여 다수의 후보문장정보 및 각 후보문장정보에 대한 순위를 나타내는 문장점수를 포함하는 음성인식결과정보를 생성하고, 상기 음성인식결과정보에 상기 정답문장정보를 부가하여 정답셋정보를 생성하고, 상기 정답셋정보에 포함된 후보문장정보들과 정답문장정보의 자질정보를 추출하여 학습 데이터를 생성하고, 상기 학습 데이터를 토대로 하고 상기 후보문장정보들과 정답문장정보의 자질정보의 차이를 토대로 후보문장정보의 순위를 가변하는 기계학습을 실행하여 생성된 것임을 특징으로 한다.

(72) 발명자

**구명완**

서울특별시 양천구 목동동로 100 1312-1304 (신정동, 목동13단지아파트)

**김학수**

서울 서초구 신반포로 137, 5동 909호 (반포동, 경남아파트)

**안혁주**

강원 원주시 로아노크로 15, 101동 1403호 (단계동, 코오롱아파트)

**송영길**

강원 춘천시 모수물길 60, 105동 804호 (소양로2가, 소양로현대아파트)

**최맹식**

강원도 삼척시 미로면 상정길 43-53

이 발명을 지원한 국가연구개발사업

과제고유번호 10041678

부처명 미래창조과학부

연구관리전문기관 정보통신기술진흥센터

연구사업명 정보통신산업융합원천기술개발 (SW컴퓨팅산업원천기술개발)

연구과제명 다중영역 정보서비스를 위한 대화형 개인 비서 소프트웨어 원천 기술 개발

기 여 율 1/1

주관기관 서강대학교

연구기간 2012.06.01 ~ 2016.02.28

## 명세서

### 청구범위

#### 청구항 1

기계학습을 이용한 음성인식방법에 있어서,

음성신호를 입력받아 음성인식하여 다수의 후보문장정보 및 각 후보문장정보에 대한 순위를 포함하는 음성인식 결과정보를 생성하는 단계;

상기 음성인식결과정보에 포함된 다수의 후보문장정보를 미리 학습된 기계학습모델에 따라 처리하여 상기 다수의 후보문장정보의 순위를 가변하여 재순위화하는 단계; 및

재순위화된 다수의 후보문장정보 중 최상위 순위의 후보문장정보를 음성인식결과로 선정하는 단계;를 포함하며,

상기 기계학습모델은,

음성신호 및 정답문장정보를 입력받고,

상기 음성신호에 대해 음성인식하여 다수의 후보문장정보 및 각 후보문장정보에 대한 순위를 나타내는 문장점수를 포함하는 음성인식결과정보를 생성하고,

상기 음성인식결과정보에 상기 정답문장정보를 부가하여 정답셋정보를 생성하고,

상기 정답셋정보에 포함된 후보문장정보들과 정답문장정보의 자질정보를 추출하여 학습 데이터를 생성하고,

상기 학습 데이터를 토대로 하고 상기 후보문장정보들과 정답문장정보의 자질정보의 차이를 토대로 후보문장정보의 순위를 가변하는 기계학습을 실행하여 생성된 것임을 특징으로 하는 기계학습을 이용한 음성인식방법.

#### 청구항 2

삭제

#### 청구항 3

제1항에 있어서,

상기 기계학습은 RankSVM임을 특징으로 하는 기계학습을 이용한 음성인식방법.

#### 청구항 4

제1항에 있어서,

상기 자질정보는, 음성인식 순위와 최상위 순위의 후보문장의 문장점수와 형태소 바이그램과 품사 바이그램과 도메인사전 미등록어 개수와 도메인사전 미등록어 형태소/품사와 제너럴사전 미등록어 개수와 제너럴사전 미등록어 형태소/품사로 구성됨을 특징으로 하는 기계학습을 이용한 음성인식방법.

#### 청구항 5

제1항에 있어서,

상기 정답셋정보는 다수의 후보문장정보 중 일부만 포함되며,

상기 일부는 상기 음성인식결과정보에 포함된 후보문장정보들 중 미리 정해진 문장점수 이상의 후보문장정보들을 임을 특징으로 하는 기계학습을 이용한 음성인식방법.

#### 청구항 6

제1항에 있어서,

상기 음성인식결과정보는 미리 정해진 외부의 음성인식 서비스 서버로부터 제공됨을 특징으로 하는 기계학습을

이용한 음성인식방법.

## 발명의 설명

### 기술 분야

[0001] 본 발명은 음성인식기술에 관한 것으로, 더욱 상세하게는 사용자의 음성을 인식한 음성인식결과로 검출된 다수 개의 후보문장을 재순위화하여 최적의 후보문장을 음성인식결과로 선정함으로써, 음성 채팅 시스템의 성능을 향상시킬 수 있는 기계학습을 이용한 음성인식방법에 관한 것이다.

### 배경 기술

[0002] 음성은 인간이 사용하는 가장 보편적이고 편리한 정보전달의 수단이다. 음성에 의해 표현되는 말은 인간과 인간 사이의 의사소통의 수단으로서 뿐만 아니라 인간의 음성을 이용하여 다양한 장치를 동작시키는 수단으로서도 중요한 역할을 수행한다.

[0003] 그리고 음성 인식이란 인간과 장치를 서로 인터페이스하는 기술로서, 최근 컴퓨터 성능의 발전, 다양한 미디어의 개발, 신호 및 정보 처리 기술의 발전으로 그 필요성이 더 대두되고 있다.

[0004] 상기 음성 인식은 입력된 음성 신호의 파형 패턴이 주어졌을 때 이를 기준 패턴과 비교하여 가장 유사한 패턴을 검출하는 것으로, 입력된 음성 신호의 파형 패턴과 가장 유사한 기준 패턴을 검출하는 작업은 기준 패턴을 생성하는 학습 단계와, 상기 학습 단계에서 생성된 기준 패턴을 이용하여 입력된 음성 신호를 인식하는 단계로 요약될 수 있다.

[0005] 이러한 음성 인식 기술로는 대한민국 특허청에 대화형 언어 학습 장치를 명칭으로 하여 특허공개된 10-2009-0119043호가 있으며, 이는 제1 대화 레벨에서 상기 입력된 사용자 음성 신호를 분석하여 상기 입력된 사용자 음성 신호에 존재하는 어절 수를 카운트하는 어절 인식부; 제1 대화 레벨에서 상기 카운트한 어절 수에 매칭되어 있는 정답 문장이 존재하는지 검색하는 문장 검색부; 및 상기 카운트한 어절 수에 매칭되어 있는 정답 문장이 검색되는 경우, 상기 검색된 정답 문장에 매칭되어 있는 제2 대화 레벨의 질문 문장을 출력하도록 제어하는 제어부를 포함하는 것을 특징으로 하는 대화형 언어 학습 장치를 개시하고 있다.

[0006] 그리고 대한민국 특허청에 대화식 학습 보조 장치 및 그의 대화분석 방법을 명칭으로 하여 특허공개된 제10-2000-0032056호가 있으며, 이는 형태소 및 구문 분석과, 의미 및 담화분석에 필요한 데이터를 제공하기 위한 사전저장 장치와, 대화분석에 필요한 데이터를 제공하기 위한 지식베이스 저장 장치와, 마이크를 통해 입력된 음성을 동일한 내용의 텍스트로 변환하기 위한 음성/텍스트 변환 장치와, 상기 음성/텍스트 변환 장치로부터 출력되는 데이터를 상기 사전저장 장치의 데이터와 비교하여 순차적으로 형태소 분석, 구문 분석 및 의미 분석을 하기 위한 형태소 해석기, 구문 분석기 및 의미 해석기와, 상기 의미 해석기로부터 출력되는 데이터를 상기 사전저장 장치의 데이터를 참조하여 생략 및 대용어 해결을 하기 위한 담화 분석기와, 상기 담화 분석기로부터 출력되는 데이터를 상기 지식베이스 저장 장치의 데이터와 비교하여 용언의 하위 범주화 정보로 변환하고, 이것으로부터 화행, 영역내 키워드, 영역내 적합성 판단, 데이터베이스 질의어 생성 및 검색 등을 수행하기 위한 대화관리기와, 상기 대화 관리기로부터 출력되는 데이터를 상기 음성/텍스트 변환 장치의 데이터와 비교하여 화행, 데이터베이스 검색 결과 및 영역내 적합도로부터 사용자에게 제공할 문장을 생성하기 위한 응답 생성기를 포함하여 이루어지는 대화식 학습 보조 장치를 개시하고 있다.

[0007] 상기한 바와 같이 음성 인식 기술은 디지털 언어 학습 장치나 채팅 시스템 등의 다양한 분야에 적용되고 있다.

[0008] 특히 상기한 음성 채팅 시스템은 사용자의 음성 인식 결과 중 1순위 문장을 바탕으로 패턴 매칭이나 검색 등의 방법을 이용하여 응답을 생성한다. 이에따라 음성인식된 1순위 문장이 오인식된다면 항상 잘못된 응답을 생성하게 되는 문제가 있었다. 더욱이 음성인식 성능이 개선되고 있기는 하나 항상 올바른 인식결과를 검출하는 것은 아니었다.

[0009] 따라서 종래에는 간소한 처리절차만으로도 음성 인식 결과를 향상시킬 수 있는 기술의 개발이 절실하게 요구되었다.

## 선행기술문헌

### 특허문헌

- [0010] (특허문헌 0001) 대한민국 특허공개 제10-2009-0119043호  
(특허문헌 0002) 대한민국 특허공개 제10-2000-0032056호

## 발명의 내용

### 해결하려는 과제

- [0011] 본 발명은 사용자의 음성을 인식한 음성인식결과로 검출된 다수개의 후보문장을 재순위화하여 최적의 후보문장을 음성인식결과로 선정함으로써, 음성 채팅 시스템의 성능을 향상시킬 수 있는 기계학습을 이용한 음성인식방법을 제공하는 것을 그 목적으로 한다.

### 과제의 해결 수단

- [0012] 상기한 본 발명에 따르는 기계학습을 이용한 음성인식방법은, 음성신호를 입력받아 음성인식하여 다수의 후보문장정보 및 각 후보문장정보에 대한 순위를 포함하는 음성인식결과정보를 생성하는 단계; 상기 음성인식결과정보에 포함된 다수의 후보문장정보를 미리 학습된 기계학습모델에 따라 처리하여 상기 다수의 후보문장정보의 순위를 가변하여 재순위화하는 단계; 및 재순위화된 다수의 후보문장정보 중 최상위 순위의 후보문장정보를 음성인식결과로 선정하는 단계;를 포함하는 것을 특징으로 한다.
- [0013] 특히, 상기 기계학습모델은, 음성신호 및 정답문장정보를 입력받고, 상기 음성신호에 대해 음성인식하여 다수의 후보문장정보 및 각 후보문장정보에 대한 순위를 나타내는 문장점수를 포함하는 음성인식결과정보를 생성하고, 상기 음성인식결과정보에 상기 정답문장정보를 부가하여 정답셋정보를 생성하고, 상기 정답셋정보에 포함된 후보문장정보들과 정답문장정보의 자질정보를 추출하여 학습 데이터를 생성하고, 상기 학습 데이터를 토대로 하고 상기 후보문장정보들과 정답문장정보의 자질정보의 차이를 토대로 후보문장정보의 순위를 가변하는 기계학습을 실행하여 생성된 것임을 특징으로 한다.

### 발명의 효과

- [0014] 상기한 본 발명은 사용자의 음성을 인식한 음성인식결과로 검출된 다수개의 후보문장을 재순위화하여 최적의 후보문장을 음성인식결과로 선정함으로써, 음성 채팅 시스템의 성능을 향상시킬 수 있는 효과를 야기한다.

### 도면의 간단한 설명

- [0015] 도 1은 본 발명의 바람직한 실시예에 따른 음성 채팅 시스템의 구성도.  
도 2 및 도 3은 본 발명의 바람직한 실시예에 따른 음성 인식 방법의 흐름도.

### 발명을 실시하기 위한 구체적인 내용

- [0016] 본 발명은 사용자의 음성을 인식한 음성인식결과로 검출된 다수개의 후보문장을 재순위화하여 최적의 후보문장을 음성인식결과로 선정함으로써, 음성 채팅 시스템의 성능을 향상시킬 수 있다.
- [0017] <음성 채팅 시스템의 구성>
- [0018] 이러한 본 발명에 따르는 음성 채팅 시스템의 구성을 도 1을 참조하여 설명한다.

- [0019] 상기한 음성 채팅 시스템은 사용자 단말기(100)와 음성인식 서비스 서버(300)로 구성된다.
- [0020] 상기 사용자 단말기(100)는 사용자의 음성신호를 입력받아 음성인식 서비스 서버(300)로 전송하고, 상기 음성인식 서비스 서버(300)로부터 다수개의 후보문장정보를 포함하는 음성인식결과정보가 제공되면, 그 후보문장정보들을 재순위화하여 최적의 후보문장정보를 음성인식결과로 선정한다.
- [0021] 상기 음성인식 서비스 서버(300)는 상기 사용자 단말기(100)로부터 음성신호를 제공받아 음성인식을 이행하고, 그 음성인식에 따른 다수개의 후보문장정보를 검출하고, 각 후보문장정보에 대해 문장점수 및 순위를 부여하고, 상기 다수개의 후보문장정보와 그에 대응되는 문장점수 및 순위로 구성되는 음성인식결과정보를 생성하여 상기 사용자 단말기(100)로 반환한다. 이러한 음성인식 서비스 서버(300)로는 구글음성인식기 서비스 서버 등이 될 수 있다.
- [0022] 상기 사용자 단말기(100)는 제어장치(200)와 메모리부(202)와 통신모듈(204)과 사용자 인터페이스부(206)와 디스플레이부(208)와 오디오 입력부(210)와 오디오 출력부(212)로 구성된다.
- [0023] 상기 제어장치(200)는 상기 사용자 단말기(100)의 각부를 전반적으로 제어함은 물론이며 본 발명의 바람직한 실시예에 따라 기계학습모델인 Rank SVM 모델을 생성하고, 그 기계학습모델을 적용한 음성인식기능을 수행한다.
- [0024] 상기 메모리부(202)는 상기 제어장치(200)의 제어 프로그램을 포함하는 다양한 정보를 저장하며, 특히 본 발명에 따르는 기계학습모델인 Rank SVM 모델을 저장한다.
- [0025] 상기 통신모듈(204)은 제어장치(200)와 음성인식 서비스 서버(300) 사이의 통신을 담당한다.
- [0026] 상기 사용자 인터페이스부(206)는 사용자로부터의 각종 명령 및 정보를 입력받아 상기 제어장치(200)에 제공한다. 특히 상기 사용자 인터페이스부(206)를 통해 사용자는 정답문장정보를 입력할 수 있다.
- [0027] 상기 디스플레이부(208)는 상기 제어장치(200)의 제어에 따른 정보를 표시하여 안내한다.
- [0028] 상기 오디오 입력부(210)는 사용자가 입력한 음성에 따른 음성신호를 제어장치(200)에 제공한다.
- [0029] 상기 오디오 출력부(212)는 상기 제어장치(200)의 제어에 따른 정보를 출력한다.
- [0030] <음성인식방법의 절차>
- [0031] 상기한 본 발명의 바람직한 실시예에 따른 음성 채팅 시스템에 적용가능한 음성인식방법을 도 2 및 도 3을 참조하여 상세히 설명한다.
- [0032] <학습과정>
- [0033] 상기 도 2는 본 발명의 바람직한 실시예에 따른 음성인식결과정보 중 최적 후보문장정보를 검출하기 위한 기계학습모델(RankSVM)을 생성하는 학습과정을 도시한 도면이다.
- [0034] 상기 사용자 단말기(100)는 오디오 입력부(210) 및 사용자 인터페이스부(206)를 통해 음성 및 정답문장정보를 입력받는다(400단계).
- [0035] 이후 상기 사용자 단말기(100)는 상기 음성신호를 외부의 음성인식 서비스 서버(300)로 전송하여 음성인식을 요청하고, 그 요청에 따른 음성인식결과정보를 반환받거나, 음성정보를 입력받아 음성인식을 수행하고 그에 따른 음성인식결과정보를 생성한다(402단계). 여기서 상기 외부의 음성인식 서비스 서버(300)로는 구글음성인식기 서비스 서버 등이 될 수 있다.
- [0036] 그리고 상기 음성인식결과정보는 다수의 후보문장정보와 각 후보문장정보에 대한 문장점수가 포함된다.

표 1

| <음성인식결과정보> |        |      |
|------------|--------|------|
| 후보문장정보     | 문장점수   | 순위   |
| 제1후보문장정보   | 제1문장점수 | 제1순위 |
| 제2후보문장정보   | 제2문장점수 | 제2순위 |
| 제3후보문장정보   | 제3문장점수 | 제3순위 |

|     |     |     |
|-----|-----|-----|
| ... | ... | ... |
|-----|-----|-----|

[0038] 상기 표 1은 음성인식결과정보의 구성을 나타낸 것으로, 상기 문장점수 및 순위는 구글 음성인식기 자체에서 추출된 문장점수와 순위로서, 본 발명에서는 문장점수가 0.5 이상이면 정답으로 고려할만한 후보문장정보로 판단한다.

[0039] 상기 사용자 단말기(100)는 다수의 후보문장정보로 구성된 음성인식결과정보 중 문장점수가 미리 정해진 값 이상인 후보문장을 정답으로 추정한다(404단계).

[0040] 상기 사용자 단말기(100)는 정답으로 추정한 후보문장정보들에 대해 정답문장정보를 부가하여 정답셋정보를 생성한다(406단계). 상기 정답셋정보는 표 2에 도시한 바와 같다.

표 2

[0041]

| <정답셋정보>            |        |      |        |
|--------------------|--------|------|--------|
| 정답으로 추정된<br>후보문장정보 | 문장점수   | 순위   | 정답문장정보 |
| 제1후보문장정보           | 제1문장점수 | 제1순위 | 정답문장정보 |
| 제2후보문장정보           | 제2문장점수 | 제2순위 | 정답문장정보 |
| 제3후보문장정보           | 제3문장점수 | 제3순위 | 정답문장정보 |

[0042] 상기 정답셋정보가 생성되면, 상기 사용자 단말기(100)는 정답셋정보에 포함된 후보문장정보들과 정답문장정보들의 자질정보를 추출하여 기계학습(RankSVM)을 실행할 수 있는 학습 데이터를 생성한다(408단계). 그리고 상기 기계학습(RankSVM)을 실행하기 위한 자질정보로는 1)ASR(음성인식기) 순위와 2)ASR 최상위 후보문장의 문장점수와 3)형태소 바이그램과 4)품사 바이그램과 5)도메인사전 미등록어 개수와 6)도메인사전 미등록어 형태소/품사와 7)제너럴사전 미등록어 개수와 8)제너럴사전 미등록어 형태소/품사로 구성된다.

[0043] 상기 기계학습(RankSVM)을 실행하기 위한 자질정보에 대해 설명하기에 앞서 도메인 사전과 제너럴 사전에 대해 설명한다. 상기 도메인 사전은 본 발명에서 실험 데이터로 추출한 14,000개 문장 중, 1,000개의 실험 데이터를 제외한 나머지 13,000 문장에서 체언, 용언, 미등록어를 추출한 사전이고, 상기 제너럴 사전은 세종말뭉치와 트위터데이터를 이용하여 체언, 용언, 미등록어를 추출한 사전이다.

[0044] 상기 ASR(음성인식기) 순위는 구글음성인식기 결과로 나오는 후보문장의 순위이다. 그리고 상기 ASR 최상위 후보문장의 점수는 ASR 순위에서 설명한 바와 같이 후보문장 순위가 1위인 후보문장에서 추출된 점수이다. 그리고 상기 형태소 바이그램은 후보문장을 형태소분석하여 추출된 형태소를 바이그램으로 쪼갠 것이다. 그리고 상기 품사 바이그램은 후보문장을 형태소분석하여 추출된 품사를 바이그램으로 쪼갠 것이다. 그리고 상기 도메인사전 미등록어 개수는 형태소분석한 결과로 추출된 각각의 체언, 용언을 추출하여, 사전에 없을 때 자질의 개수를 추가한 것이다. 그리고 상기 도메인사전 미등록어 형태소/품사는 형태소분석한 결과로 추출된 각각의 체언, 용언을 추출하여 사전에 없을 때 형태소/품사에 대한 자질을 추가한 것이다. 그리고 상기 제너럴사전 미등록어 개수는 형태소분석한 결과로 추출된 각각의 체언, 용언을 추출하여 사전에 없을 때 자질의 개수를 추가한 것이다. 그리고 상기 제너럴사전 미등록어 형태소/품사는 형태소분석한 결과로 추출된 각각의 체언, 용언을 추출하여 사전에 없을 때 형태소/품사에 대한 자질을 추가한 것이다.

[0045] 상기한 바와 같은 자질정보를 추출한 학습 데이터가 생성되면, 상기 사용자 단말기는 학습 데이터를 토대로 기계학습(RankSVM)을 실행하여 기계학습모델을 생성한다(410단계).

[0046] 상기 기계학습(RankSVM)은 후보문장정보들과 정답문장정보 사이의 자질정보의 차이를 검출하고, 그 차이에 따라 후보문장들의 순위를 가변하는 기계학습모델을 생성한다.

[0047] 예를들면, 표 3에 예시한 바와 같이 후보문장정보의 순위가 가변될 수 있다.

표 3

[0048]

| <순위가변정보>        |        |      |        |
|-----------------|--------|------|--------|
| 정답으로 추정된 후보문장정보 | 문장점수   | 순위   | 가변된 순위 |
| 제1후보문장정보        | 제1문장점수 | 제1순위 | 제2순위   |
| 제2후보문장정보        | 제2문장점수 | 제2순위 | 제1순위   |
| 제3후보문장정보        | 제3문장점수 | 제3순위 | 제3순위   |

[0049]

[0050] 이에따라 본 발명에 따른 기계학습모델은 사용자의 음성을 인식함에 따라 제1 내지 제3후보문장정보가 입력되면, 그 제1 내지 제3후보문장정보의 순위를 가변된 순위로 변경함으로써, 정답문장정보에 좀더 가까운 후보문장의 순위가 최상위 순위가 될 수 있게 한다.

[0051] 또한 상기 자질정보를 추출할 때에는 상기 후보문장에 대한 순위와 점수는 그대로 채용되고 후보문장 자체는 형태소 분석한 결과만을 채용한다. 상기한 순위와 점수 정보, 형태소 등의 자질정보들은 기계학습모델이 이해할 수 있도록 정제된 뒤에, 기계학습되어 문장들의 순위를 가변한다.

[0052] <적용과정>

[0053] 상기 도 3은 본 발명의 바람직한 실시예에 따른 기계학습모델을 적용하여 음성인식결과 중 최적 후보 문장을 검출하는 적용과정을 도시한 도면이다.

[0054] 상기 사용자 단말기(100)는 사용자로부터 음성신호를 입력받는다(500단계).

[0055] 이후 상기 사용자 단말기(100)는 입력된 음성신호를 외부의 음성인식 서비스 서버(300)에 제공하여 음성인식을 요청하고, 그 요청에 따라 상기 음성인식 서비스 서버(300)가 제공하는 음성인식결과정보를 반환받거나, 독립적으로 음성인식을 수행하고 그에 따른 음성인식결과정보를 생성한다(502단계). 상기 음성인식결과정보는 다수의 후보문장정보와 각 후보문장정보에 대응되는 문장점수 및 순위를 포함한다.

[0056] 이후 상기 사용자 단말기(100)는 다수의 후보문장정보를 본 발명에 의해 생성된 기계학습모델(RankSVM)에 적용하여, 상기 다수의 후보문장정보의 순위를 가변하여 최적의 후보문장정보가 최상위 순위가 되도록 재순위화하고, 음성인식결과로서 상기 최상위 순위의 후보문장정보를 선정한다(504단계).

[0057] 상기한 본 발명의 바람직한 실시예에서는 사용자 단말기(100)가 음성인식을 외부의 음성인식 서비스 서버(300)로 요청하여, 그에 따른 결과를 제공받는 것만을 개시하였으나, 사용자 단말기(100)가 음성인식을 독립적으로 수행하는 것 역시 가능하며 이는 본 발명에 의해 자명하다.

[0058] 또한 상기한 본 발명의 바람직한 실시예에서는 기계학습모델로서 RankSVM을 채용하는 것만을 예시하였으나 다양한 기계학습모델이 채용될 수도 있으며, 이는 본 발명에 의해 자명한다.

## 부호의 설명

[0059]

100 : 사용자 단말기

200 : 제어장치

202 : 메모리부

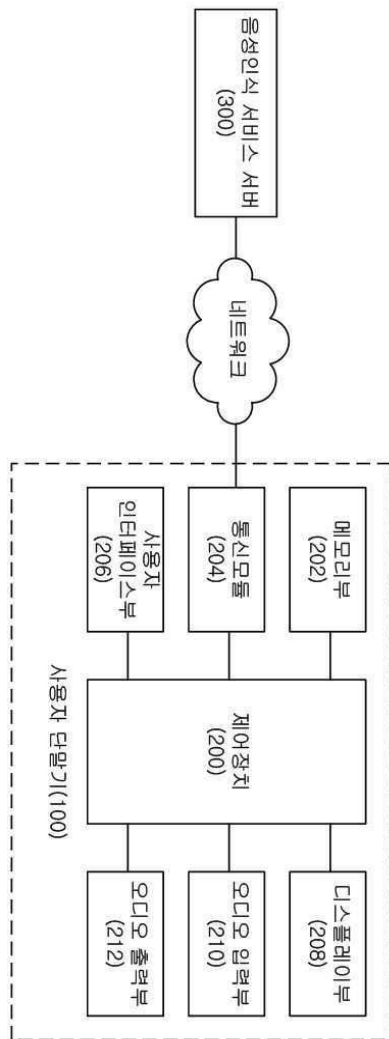
204 : 통신모듈

206 : 사용자 인터페이스부

- 208 : 디스플레이부
- 210 : 오디오 입력부
- 212 : 오디오 출력부
- 300 : 음성인식 서비스 서버

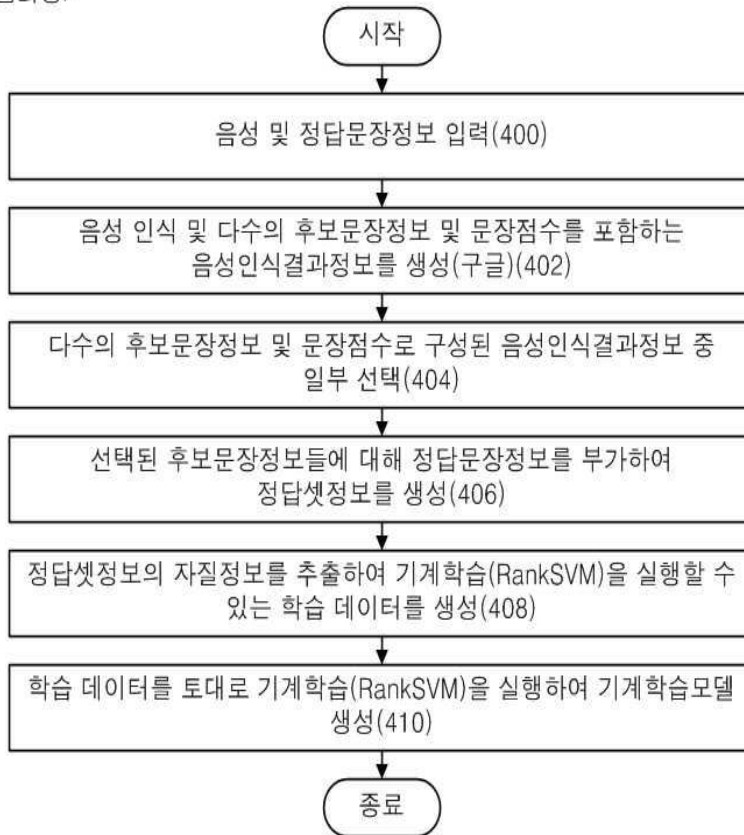
도면

도면1



도면2

<학습과정>



도면3

<적용과정>

