



이 발명을 지원한 국가연구개발사업

과제고유번호 10035252

부처명 지식경제부

연구사업명 산업원천기술개발사업(정보통신)

연구과제명 모바일 플랫폼 기반 대화모델 적용 자연어 음성 인터페이스 기술 개발

기 여 율 1/1

주관기관 한국전자통신연구원

연구기간 2010.03.01 ~ 2015.02.28

특허청구의 범위

청구항 1

인식의 대상이 되는 사용자의 음성을 수신하는 음성 수신부;

상기 사용자의 위치 정보를 파악하는 위치 정보 파악부;

상기 사용자의 위치 정보를 이용하여, 상기 사용자가 위치하고 있는 곳의 잡음 환경을 분석하고, 상기 잡음 환경에 대응하는 음향모델을 음향모델 DB(Data Base)에서 추출하는 음향모델 추출부;

상기 사용자의 위치 정보를 이용하여, 상기 사용자가 위치하고 있는 곳에 대응되는 어휘언어모델을 어휘언어모델 DB에서 추출하는 어휘언어모델 추출부; 및

상기 음향모델 및 상기 어휘언어모델을 이용하여 상기 사용자의 음성에 대한 인식을 수행하는 음성 인식부를 포함하는 것을 특징으로 하는 사용자 위치 정보를 활용한 음성 인식 장치.

명세서

기술 분야

[0001] 본 발명은 사용자 위치 정보를 활용한 음성 인식 장치 및 방법에 관한 것이다. 더욱 상세하게, 본 발명은 사용자 위치 정보를 활용하여 맞춤형 음향모델 및 언어모델을 제공함으로써 음성인식 서비스의 성능을 높일 수 있는 음성 인식 장치 및 방법에 관한 것이다.

배경 기술

[0002] 휴대폰 등의 통신 단말을 사용하여 음성통화 시에 주변 잡음이 존재하는 경우에는 좋은 통화 품질을 보장하기가 어렵다. 따라서 잡음이 존재하는 환경에서 통화 품질을 높이기 위해서는 주변 잡음 성분을 추정하여 실제 음성 신호만을 추출하는 기술이 필요하다.

[0003] 이와 더불어, 캠코더, 노트북 PC, 네비게이션, 게임기, 휴대폰 등 여러가지 단말기에서 음성을 입력받아 동작하거나 음성 데이터를 저장하는 등 음성 기반의 응용예가 증가하고 있어, 주변 잡음을 감소 또는 제거하여 좋은 품질의 음성을 추출해 내는 기술이 필요하다.

[0004] 종래에도 주변 잡음을 추정하거나 감소시키는 여러가지 방법들이 개시되어 있다. 그러나 시간에 따라 잡음의 통계적 특성이 변화하거나, 잡음의 통계적 특성을 알아내기 위한 초기 단계에서, 예측하지 못한 산발적인(sporadic) 잡음이 발생하는 경우에는 원하는 잡음 감소 또는 제거 성능을 얻지 못한다.

[0005] 관련하여, 한국특허출원 제10-2009-0085511호는 "잡음 추정 장치 및 방법과, 이를 이용한 감소 장치"에 관한 기술을 개시하고 있다.

발명의 내용

해결하려는 과제

[0006] 본 발명은 사용자 위치 정보를 활용하여 맞춤형 음향모델 및 언어모델을 제공함으로써 음성인식 서비스의 성능을 높이는 것을 목적으로 한다. 더불어, 본 발명은 누적된 사용자 입력을 바탕으로 갱신 및 관리된 데이터베이스를 이용하여 음성인식 서비스의 성능을 보다 높이는 것을 목적으로 한다.

[0007] 그리고, 본 발명은 음향모델 DB 및 어휘언어모델 DB에 미리 저장 및 분류된 음향모델 및 어휘언어모델을 이용하여 음성인식을 수행하여, 음성인식 수행 속도의 저하를 방지하는 것을 목적으로 한다.

과제의 해결 수단

[0008] 상기한 목적을 달성하기 위한 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 장치는 인식의 대상이 되는 사용자의 음성을 수신하는 음성 수신부; 상기 사용자의 위치 정보를 파악하는 위치 정보 파악부; 상기 사용자의 위치 정보를 이용하여, 상기 사용자가 위치하고 있는 곳의 잡음 환경을 분석하고, 상기 잡음 환경에 맞는 음향

모델을 음향모델 DB(Data Base)에서 추출하는 음향모델 추출부; 상기 사용자의 위치 정보를 이용하여, 상기 사용자가 위치하고 있는 곳에 대응되는 어휘언어모델을 어휘언어모델 DB에서 추출하는 어휘언어모델 추출부; 및 상기 음향모델 및 상기 어휘언어모델을 이용하여 상기 사용자의 음성에 대한 인식을 수행하는 음성 인식부를 포함하는 것을 특징으로 한다.

발명의 효과

[0009] 본 발명에 따르면, 사용자 위치 정보를 활용하여 맞춤형 음향모델 및 언어모델을 제공함으로써 음성인식 서비스의 성능을 높일 수 있다. 더불어, 본 발명은 누적된 사용자 입력을 바탕으로 갱신 및 관리된 데이터베이스를 이용하여 음성인식 서비스의 성능을 보다 높일 수 있다.

[0010] 그리고, 본 발명은 음향모델 DB 및 어휘언어모델 DB에 미리 저장 및 분류된 음향모델 및 어휘언어모델을 이용하여 음성인식을 수행하여, 음성인식 수행 속도의 저하를 방지할 수 있다.

도면의 간단한 설명

[0011] 도 1은 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 장치의 구성을 나타낸 블록도이다.

도 2는 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 방법을 설명하기 위한 플로우차트이다.

발명을 실시하기 위한 구체적인 내용

[0012] 본 발명을 첨부된 도면을 참조하여 상세히 설명하면 다음과 같다. 여기서, 반복되는 설명, 본 발명의 요지를 불필요하게 흐릴 수 있는 공지 기능, 및 구성에 대한 상세한 설명은 생략한다. 본 발명의 실시형태는 당 업계에서 평균적인 지식을 가진 자에게 본 발명을 보다 완전하게 설명하기 위해서 제공되는 것이다. 따라서, 도면에서의 요소들의 형상 및 크기 등은 보다 명확한 설명을 위해 과장될 수 있다.

[0013] 이하에서는 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 장치의 구성 및 동작에 대하여 설명하도록 한다.

[0014] 도 1은 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 장치의 구성을 나타낸 블록도이다.

[0015] 도 1을 참조하면, 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 장치(100)는 음성 수신부(110), 위치 정보 파악부(120), 잡음 환경 판단부(130), 음향모델 추출부(140), 어휘언어모델 추출부(150), 음성 인식부(160)를 포함하여 구성된다. 그리고, 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 장치(100)는 갱신부(170)를 더 포함하여 구성될 수 있다.

[0016] 음성 수신부(110)는 인식의 대상이 되는 사용자의 음성을 수신한다.

[0017] 위치 정보 파악부(120)는 음성을 발화하는 사용자의 위치 정보를 파악한다.

[0018] 잡음 환경 판단부(130)는 위치 정보 파악부(120)에서 파악된 사용자의 위치 정보를 이용하여 사용자가 위치하고 있는 곳의 잡음 환경을 분석한다. 이러한, 잡음 환경 판단부(130)는 잡음 환경에 대한 모델들이 기 정의 및 저장된 잡음 환경 DB(10)에서 잡음 환경 모델을 추출하여 잡음 환경을 분석한다. 이 때, 잡음 환경 DB(10)는 위치 및 환경에 따라 모델링된 여러 형태의 잡음 환경 모델을 저장하고 있는 데이터베이스이다. 구체적으로, 잡음 환경 DB(10)에는 자동차 내부, 지하철역, 지하철, 길거리, 음식점, 집 내부 등의 잡음 환경 모델에 대한 정의가 기 저장되어 있을 수 있다. 이 때, 잡음 환경 판단부(130)에서 결정된 잡음 환경 모델이 사용자의 입력 음성에서 얻어진 신호 특성과 크게 상이할 경우, 위치에 따른 잡음 환경 모델을 적용하지 않고, 기본적으로 제공되는 범용 잡음 환경 모델이 적용될 수 있다.

[0019] 음향모델 추출부(140)는 잡음 환경 판단부(130)에서 판단된 잡음 환경 모델에 대응되는 음향모델을 음향모델 DB(20)에서 추출한다. 이 때, 음향모델의 추출은 사용자가 음성을 입력하기 전에 백그라운드 작업을 통해 미리 수행될 수 있다. 음향모델 DB(20)는 사용자 위치 및 환경에 따른 잡음 환경 모델에 대응하여 다양한 형태로 적용된 음향모델에 대한 데이터베이스이다. 예를 들어, 음향모델 추출부(140)는 잡음 환경 판단부(130)에서 지하철역에 대응되는 잡음환경 모델이 정의된 경우, 지하철 소음 등의 노이즈 신호를 제거할 수 있는 음향모델을 추

출한다.

- [0020] 어휘언어모델 추출부(150)는 사용자의 위치 정보를 이용하여, 사용자가 위치하고 있는 곳에 대응되는 어휘언어모델을 어휘언어모델 DB(30)에서 추출한다. 이러한, 어휘언어모델 추출부(150)는 사용자의 음성에는 사용자 위치를 중심으로 주변 상호명이나 POI(Point Of Interest) 등이 포함될 가능성이 높으므로, 기존 인식 어휘에 해당 어휘를 추가하여 어휘언어모델을 구성한다. 그리고, 어휘언어모델 추출부(150)는 그 동안 사용자가 발화했던 문장에 대한 언어모델을 추가하여 어휘언어모델을 구성한다. 어휘언어모델 DB(30)는 범용의 인식 어휘 모델과 위치에 따른 인식 어휘를 저장하는 인식 어휘 DB(31), 및 그 동안 사용자가 발화했던 문장에 대한 언어 모델을 분석 및 저장하는 누적 사용자 발화 DB(32)로 구성될 수 있다. 그리고, 이러한 어휘언어모델의 추출은 사용자가 음성을 입력하기 전에 백그라운드 작업을 통해 미리 수행될 수 있다.
- [0021] 음성 인식부(160)는 음향모델 추출부(140)에서 추출된 음향모델 및 어휘언어모델 추출부(150)에서 추출된 어휘언어모델에 기반하여, 음성 수신부(110)에서 수신한 사용자의 음성을 최종 인식한다.
- [0022] 갱신부(170)는 사용자가 음성 인식을 위한 음성을 입력할 때마다 잡음환경 DB(10)의 잡음환경 모델, 음향모델 DB(20)의 음향모델, 어휘언어모델 DB(30)의 어휘언어모델을 해당 인식된 음성을 토대로 갱신한다. 즉, 갱신부(170)는 입력되는 음성 신호를 토대로 잡음환경 DB(10)의 위치에 따른 잡음환경 모델의 정보를 갱신한다. 또한, 갱신부(170)는 입력되는 음성 신호 및 잡음환경 모델을 토대로 음향모델의 정보를 갱신한다. 또한, 갱신부(170)는 사용자가 발화한 문장을 토대로 어휘언어모델 DB(30)의 누적 사용자 발화 DB(32)의 통계값을 갱신한다.
- [0023] 이하에서는 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 방법에 대하여 설명하도록 한다.
- [0024] 도 2는 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 방법을 설명하기 위한 플로우차트이다.
- [0025] 도 2를 참조하면, 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 방법은 먼저, 인식의 대상이 되는 사용자의 음성을 수신한다(S10).
- [0026] 그리고, 음성을 발화하는 사용자의 위치 정보를 파악한다(S20).
- [0027] 이 후, S20 단계에서 파악된 사용자의 위치 정보에 기반하여, 잡음환경 DB에서 사용자가 위치한 곳의 잡음 환경 모델을 분석한다(S30).
- [0028] 그리고, S30 단계에서 분석된 잡음 환경 모델에 대응되는 음향모델을 음향모델 DB에서 추출한다(S40).
- [0029] 사용자의 위치 정보를 이용하여, 사용자가 위치하고 있는 곳에 대응되는 어휘언어모델을 어휘언어모델 DB에서 추출한다(S50). 이러한, S50 단계에서는 사용자의 음성에는 사용자 위치를 중심으로 주변 상호명이나 POI(Point Of Interest) 등이 포함될 가능성이 높으므로, 기존 인식 어휘에 해당 어휘를 추가하여 어휘언어모델을 구성할 수 있다. 그리고, S50 단계에서는 그 동안 사용자가 발화했던 문장에 대한 언어모델을 추가하여 어휘언어모델을 구성할 수 있다.
- [0030] S40 단계에서 추출된 음향모델 및 S50 단계에서 추출된 어휘언어모델에 기반하여, S10 단계에서 수신된 사용자의 음성을 최종 인식한다(S60).
- [0031] 그리고, 기존의 잡음환경 DB의 잡음환경 모델, 음향모델 DB의 음향모델, 어휘언어모델 DB의 어휘언어모델을 S60 단계에서 인식된 음성을 토대로 갱신한다(S70).

[0032] 이상에서와 같이 본 발명에 따른 사용자 위치 정보를 활용한 음성 인식 장치 및 방법은 상기한 바와 같이 설명된 실시예들의 구성과 방법이 한정되게 적용될 수 있는 것이 아니라, 상기 실시예들은 다양한 변형이 이루어질 수 있도록 각 실시예들의 전부 또는 일부가 선택적으로 조합되어 구성될 수도 있다.

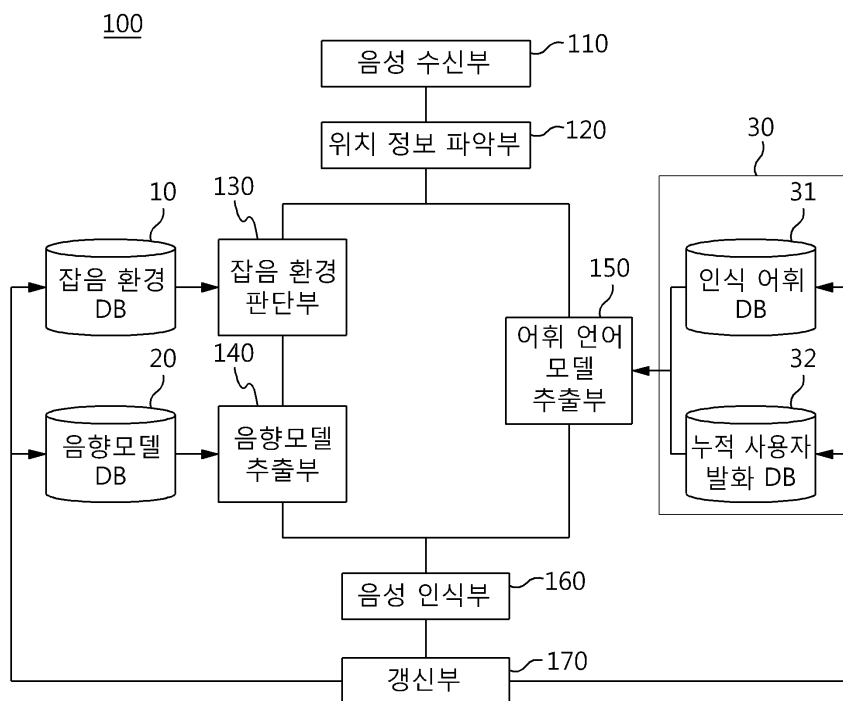
부호의 설명

- [0033] 100; 음성 인식 장치
110; 음성 수신부

- 120; 위치 정보 파악부
- 130; 잡음 환경 판단부
- 140; 음향 모델 추출부
- 150; 어휘언어모델 추출부
- 160; 음성 인식부
- 170; 갱신부
- 10; 잡음 환경 DB
- 20; 음향모델 DB
- 30; 어휘언어모델 DB
- 31; 인식 어휘 DB
- 32; 누적 사용자 발화 DB

도면

도면1



도면2

