



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2015-0031984
(43) 공개일자 2015년03월25일

(51) 국제특허분류(Int. Cl.)
G10L 15/14 (2006.01)

(21) 출원번호 10-2013-0111918
(22) 출원일자 2013년09월17일
심사청구일자 없음

(71) 출원인
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)

(72) 발명자
김동현
대전광역시 유성구 가정로 270 한국전자통신연구
원 기숙사 1동 233호

(74) 대리인
특허법인이지

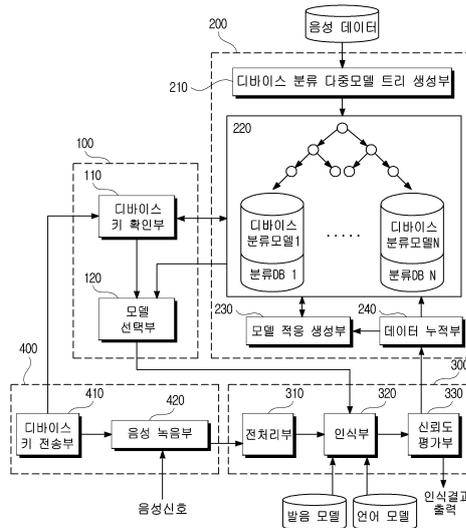
전체 청구항 수 : 총 15 항

(54) 발명의 명칭 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템 및 방법

(57) 요약

본 발명은 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템 및 방법에 관한 것으로, 상기 음성인식 시스템은, 사용자 디바이스 단말로부터 전송된 디바이스 키를 확인 및 분류하여 다중 모델 트리의 음향 모델을 선택하는 모델 선정 모듈; 음성 데이터를 사용자 디바이스 단말 별로 분류하여 다중 모델 트리를 생성하고 누적 적응하는 모델 관리 모듈; 및 상기 모델 선정 모듈로부터 선택된 음향모델을 제공받아 음성 인식을 수행하고, 신뢰도가 기 정의된 문턱값을 초과하는 데이터를 상기 모델 관리 모듈에 전달하는 음성 인식 모듈을 포함하여 구성된다.

대표도 - 도1



특허청구의 범위

청구항 1

사용자 디바이스 단말로부터 전송된 디바이스 키를 확인 및 분류하여 다중 모델 트리의 음향 모델을 선택하는 모델 선정 모듈;

음성 데이터를 사용자 디바이스 단말 별로 분류하여 다중 모델 트리를 생성하고 누적 적응하는 모델 관리 모듈; 및

상기 모델 선정 모듈로부터 선택된 음향모델을 제공받아 음성 인식을 수행하고, 신뢰도가 기 정의된 문턱값을 초과하는 데이터를 상기 모델 관리 모듈에 전달하는 음성 인식 모듈을 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 2

제 1 항에 있어서, 상기 모델 선정 모듈은,

상기 디바이스 키를 확인하여, 새로운 디바이스 키이면 상기 다중 모델 트리의 일 구성요소인 디바이스 키 트리 에 등록하고, 기존의 디바이스 키이면 상기 다중 모델 트리로부터 상기 디바이스 키 트리의 사용 가능한 음향모델 리스트를 제공받는 디바이스 키 확인부; 및

상기 사용 가능한 음향모델 리스트에서 하나의 음향 모델 또는 병렬 인식에 사용할 다중 음향 모델을 선택하고, 선택된 모델의 주소 및 상기 디바이스 키 정보를 상기 음성 인식 모듈로 전달하는 모델 선택부를 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 3

제 2 항에 있어서,

상기 음향모델 리스트는 상기 다중모델 트리에 포함된 디바이스 독립모델, 디바이스 의존모델, 세부 디바이스 의존모델 및 랜덤 ID 의존모델을 순차적으로 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 4

제 2 항에 있어서,

상기 모델 선택부는 상기 디바이스 키를 이용하여 상기 다중 모델 트리의 일 구성요소인 음향 모델 트리의 단말 노드에서 음향 모델을 검색하여 선정하고, 상기 단말 노드에 음향 모델이 없는 경우 상기 단말 노드에 연결된 상위 노드의 음향 모델을 선정하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 5

제 4 항에 있어서,

상기 모델 선택부는, 병렬 인식을 수행하는 경우, 상기 디바이스 키에 해당하는 사용자 디바이스 단말의 음향 모델과 상위 노드의 음향 모델들 중에서 기 정해진 규칙에 따라 다중 음향 모델을 선정하는 디바이스 구분 음향 모델 누적 적응을 이용한 음성인식 시스템.

청구항 6

제 1 항에 있어서, 상기 모델 관리 모듈은,

상기 음성 데이터를 사용자 디바이스 단말 별로 분류 및 특징 추출하여 상기 다중 모델 트리를 구성하는 디바이스 분류 다중모델 트리 생성부;

사용자 디바이스 단말 별로 분류되며, 계층적 음향모델, 누적 데이터 및 디바이스 키 트리를 포함하는 다중모델 트리;

상기 음성 인식 모듈로부터 상기 신뢰도가 문턱값을 초과하는 데이터를 입력받아 상기 다중 모델 트리에 음소정보와 데이터를 누적시키고, 각 노드의 데이터 누적 정도를 평가하여 기 정해진 기준값을 초과하면 모델 적응 생성부로 통보하는 데이터 누적부; 및

상기 데이터 누적부로부터 전달된 정보를 이용하여 상기 다중모델 트리에서 새로 누적된 데이터가 상기 기준값을 초과하는 노드에 해당하는 음향 모델을 누적 적응하는 모델 적응 생성부를 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 7

제 6 항에 있어서,

상기 디바이스 분류 다중모델 트리 생성부는 상기 음성 데이터를 사용자 디바이스 단말 별로 분류하여 디바이스 독립 모델에서 적용된 디바이스 의존 모델을 생성한 다중모델 트리를 생성하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 8

제 6 항에 있어서,

상기 다중모델 트리는 상기 모델 선정 모듈로부터 새로운 디바이스 키가 전달되면 상기 디바이스 키 트리에 상기 디바이스 키를 등록하고 데이터 수집을 위한 데이터 노드를 생성하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 9

제 6 항에 있어서,

상기 데이터 누적부는 상기 음성 인식 모듈로부터 전달받은 디바이스 키를 이용하여 상기 다중모델 트리에서 등록된 노드를 검색하고, 검색된 노드에서 상위 노드까지 포함하여 누적시키는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 10

제 1 항에 있어서, 상기 음성 인식 모듈은,

상기 사용자 디바이스 단말로부터 입력된 음성 신호를 전처리하는 전처리부;

상기 전처리부로부터 제공된 특징 추출된 데이터와 상기 모델 선정 모듈에 의해 선택된 음향모델을 사용하여 음성 인식을 수행하는 인식부; 및

상기 인식부로부터 전달받은 음성 인식 결과의 신뢰도를 측정하여 신뢰도가 가장 높은 결과를 인식 결과로 출력하고, 신뢰도가 상기 문턱값을 초과하는 경우 음소정보, 데이터 및 디바이스 키를 상기 모델 관리 모듈로 전달하는 신뢰도 평가부를 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 11

제 10 항에 있어서,

상기 인식부는 상기 모델 선정 모듈로부터 복수의 음향 모델을 제공받고, 상기 복수의 음향 모델에 대한 병렬 음성 인식을 수행하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 12

제 1 항에 있어서,

상기 사용자 디바이스 단말에는 음성 인식 프로그램이 설치되며,

상기 음성 인식 프로그램은,

상기 음성 인식 프로그램의 최초 설치시 디바이스 키를 생성하여 상기 모델 선정 모듈로 전달하고, 상기 음성 인식 프로그램의 사용시마다 상기 디바이스 키를 상기 모델 선정 모듈로 전달하는 디바이스 키 전송부; 및

음성 인식 수행시 입력되는 음성 데이터를 녹음하고, 녹음한 음성 데이터 또는 특징 추출된 데이터를 상기 음성 인식 모듈로 전달하는 음성 녹음부를 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 13

제 1 항에 있어서,

상기 디바이스 키는 디바이스 명, 디바이스 세부명 및 랜덤 ID를 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 14

제1 항에 있어서,

상기 다중 모델 트리는,

디바이스 기준으로 전체 데이터를 분류한 데이터 트리;

상기 데이터 트리의 노드에 있는 데이터를 적응 데이터로 활용하여 각 노드의 음향 모델을 생성한 음향모델 트리; 및

디바이스 키를 등록한 디바이스 키 트리를 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템.

청구항 15

사용자 디바이스 단말로부터 음성인식 프로그램의 설치시 전송된 디바이스 키를 디바이스 키 트리에 등록하는 단계;

상기 사용자 디바이스 단말로부터 상기 음성인식 프로그램의 실행에 따라 상기 디바이스 키가 전송되면, 상기 디바이스 키를 이용하여 사용 가능한 음향모델을 선택하는 단계;

상기 사용 가능한 음향모델 및 상기 사용자 디바이스 단말로부터 전송된 음성 데이터를 이용하여 음성 인식을 수행하는 단계;

음성 인식 결과의 신뢰도를 측정하여 상기 신뢰도가 문턱값을 초과하는 데이터는, 상기 디바이스 키에 따라 다중 모델 트리의 단말 노드 및 직계 상위 노드들에 인식 정보와 함께 데이터 누적시키는 단계; 및

상기 다중모델 트리의 각 노드의 데이터가 기준값을 초과하면, 음향모델 누적 적응을 수행하는 단계를 포함하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 방법.

명세서

기술분야

[0001] 본 발명은 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템 및 방법에 관한 것으로, 보다 상세하게는 사용자 디바이스 단말에 고유한 디바이스 키를 자동으로 생성함으로써 사용자의 관여 없이도 동일한 사용자 디바이스 단말에서 음성인식 서비스를 지속적으로 사용할수록 음성인식 성능이 점진적으로 향상될 수 있도록 하는 기술에 관한 것이다.

배경기술

[0002] 음성인식 시스템이 다양한 개인용 디바이스 단말에서 사용됨에 따라 사용자 의존적인 성능 향상을 기대하게 되었다. 사용자 의존적인 음향 모델을 만들기 위해서는 화자의 음성 데이터가 많이 필요한데, 이러한 데이터를 수집하는데 많은 시간이 필요하다. 이에 적은 양의 음성 데이터를 활용한 화자 적응 방법들이 제안되고 있다.

[0003] 그러나, 이러한 화자 적응 방법들은 데이터가 많을수록 성능 향상을 기대할 수 있기 때문에, 지속적으로 데이터를 누적해야 하는 부담감이 있다. 더 나아가, 온라인 음성인식 서비스의 경우, 개별 화자를 구분하여 데이터를 저장하기 위해 화자 인식 및 인증을 수행해야 하는 번거로움이 있다.

발명의 내용

해결하려는 과제

[0004] 따라서 본 발명은 상기와 같은 종래 기술의 문제점을 해결하기 위한 것으로, 사용자 디바이스 단말에 고유한 디바이스 키를 자동으로 생성함으로써 사용자의 관여 없이도 동일한 사용자 디바이스 단말에서 음성인식 서비스를 지속적으로 사용할수록 음성인식 성능이 점진적으로 향상될 수 있도록 하는 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템 및 방법을 제공하기 위한 것이다.

과제의 해결 수단

[0005] 상기한 목적을 달성하기 위한 본 발명의 일 실시예에 의한 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템은, 사용자 디바이스 단말로부터 전송된 디바이스 키를 확인 및 분류하여 다중 모델 트리의 음향 모델을 선택하는 모델 선정 모듈; 음성 데이터를 사용자 디바이스 단말 별로 분류하여 다중 모델 트리를 생성하고 누적 적응하는 모델 관리 모듈; 및 상기 모델 선정 모듈로부터 선택된 음향모델을 제공받아 음성 인식을 수행하고, 신뢰도가 기 정의된 문턱값을 초과하는 데이터를 상기 모델 관리 모듈에 전달하는 음성 인식 모듈을 포함한다.

[0006] 상기 모델 선정 모듈은, 상기 디바이스 키를 확인하여, 새로운 디바이스 키이면 상기 다중 모델 트리의 일 구성요소인 디바이스 키 트리에 등록하고, 기존의 디바이스 키이면 상기 다중 모델 트리로부터 상기 디바이스 키 트리의 사용 가능한 음향모델 리스트를 제공하는 디바이스 키 확인부; 및 상기 사용 가능한 음향모델 리스트에서 하나의 음향 모델 또는 병렬 인식에 사용할 다중 음향 모델을 선택하고, 선택된 모델의 주소 및 상기 디바이스 키 정보를 상기 음성 인식 모듈로 전달하는 모델 선택부를 포함할 수 있다.

[0007] 여기서, 상기 음향모델 리스트는 상기 다중모델 트리에 포함된 디바이스 독립모델, 디바이스 의존모델, 세부 디바이스 의존모델 및 랜덤 ID 의존모델을 순차적으로 포함할 수 있다.

[0008] 또한, 상기 모델 선택부는 상기 디바이스 키를 이용하여 상기 다중 모델 트리의 일 구성요소인 음향 모델 트리의 단말 노드에서 음향 모델을 검색하여 선정하고, 상기 단말 노드에 음향 모델이 없는 경우 상기 단말 노드에 연결된 상위 노드의 음향 모델을 선정할 수 있다.

[0009] 또한, 상기 모델 선택부는, 병렬 인식을 수행하는 경우, 상기 디바이스 키에 해당하는 사용자 디바이스 단말의 음향 모델과 상위 노드의 음향 모델들 중에서 기 정해진 규칙에 따라 다중 음향 모델을 선정할 수 있다.

[0010] 상기 모델 관리 모듈은, 상기 음성 데이터를 사용자 디바이스 단말 별로 분류 및 특징 추출하여 상기 다중 모델 트리를 구성하는 디바이스 분류 다중모델 트리 생성부; 사용자 디바이스 단말 별로 분류되며, 계층적 음향모델, 누적 데이터 및 디바이스 키 트리를 포함하는 다중모델 트리; 상기 음성 인식 모듈로부터 상기 신뢰도가 문턱값을 초과하는 데이터를 입력받아 상기 다중 모델 트리에 음소정보와 데이터를 누적시키고, 각 노드의 데이터 누적 정도를 평가하여 기 정해진 기준값을 초과하면 모델 적응 생성부로 통보하는 데이터 누적부; 및 상기 데이터 누적부로부터 전달된 정보를 이용하여 상기 다중모델 트리에서 새로 누적된 데이터가 상기 기준값을 초과하는 노드에 해당하는 음향 모델을 누적 적응하는 모델 적응 생성부를 포함할 수 있다.

[0011] 상기 디바이스 분류 다중모델 트리 생성부는 상기 음성 데이터를 사용자 디바이스 단말 별로 분류하여 디바이스 독립 모델에서 적용된 디바이스 의존 모델을 생성한 다중모델 트리를 생성할 수 있다.

[0012] 상기 다중모델 트리는 상기 모델 선정 모듈로부터 새로운 디바이스 키가 전달되면 상기 디바이스 키 트리에 상기 디바이스 키를 등록하고 데이터 수집을 위한 데이터 노드를 생성할 수 있다.

[0013] 상기 데이터 누적부는 상기 음성 인식 모듈로부터 전달받은 디바이스 키를 이용하여 상기 다중모델 트리에서 등록된 노드를 검색하고, 검색된 노드에서 상위 노드까지 포함하여 누적시킬 수 있다.

[0014] 상기 음성 인식 모듈은, 상기 사용자 디바이스 단말로부터 입력된 음성 신호를 전처리하는 전처리부; 상기 전처리부로부터 제공된 특징 추출된 데이터와 상기 모델 선정 모듈에 의해 선택된 음향모델을 사용하여 음성 인식을 수행하는 인식부; 및 상기 인식부로부터 전달받은 음성 인식 결과의 신뢰도를 측정하여 신뢰도가 가장 높은 결과를 인식 결과로 출력하고, 신뢰도가 상기 문턱값을 초과하는 경우 음소정보, 데이터 및 디바이스 키를 상기 모델 관리 모듈로 전달하는 신뢰도 평가부를 포함할 수 있다.

[0015] 상기 인식부는 상기 모델 선정 모듈로부터 복수의 음향 모델을 제공받고, 상기 복수의 음향 모델에 대한 병렬

음성 인식을 수행할 수 있다.

- [0016] 상기 사용자 디바이스 단말에는 음성 인식 프로그램이 설치되며, 상기 음성 인식 프로그램은, 상기 음성 인식 프로그램의 최초 설치시 디바이스 키를 생성하여 상기 모델 선정 모듈로 전달하고, 상기 음성 인식 프로그램의 사용시마다 상기 디바이스 키를 상기 모델 선정 모듈로 전달하는 디바이스 키 전송부; 및 음성 인식 수행시 입력되는 음성 데이터를 녹음하고, 녹음한 음성 데이터 또는 특징 추출된 데이터를 상기 음성 인식 모듈로 전달하는 음성 녹음부를 포함할 수 있다.
- [0017] 상기 디바이스 키는 디바이스 명, 디바이스 세부명 및 랜덤 ID를 포함할 수 있다.
- [0018] 상기 다중 모델 트리는, 디바이스 기준으로 전체 데이터를 분류한 데이터 트리; 상기 데이터 트리의 노드에 있는 데이터를 적용 데이터로 활용하여 각 노드의 음향 모델을 생성한 음향모델 트리; 및 디바이스 키를 등록한 디바이스 키 트리를 포함할 수 있다.
- [0019] 상기한 목적을 달성하기 위한 본 발명의 다른 실시예에 의한 디바이스 구분 음향모델 누적 적용을 이용한 음성 인식 방법은, 사용자 디바이스 단말로부터 음성인식 프로그램의 설치시 전송된 디바이스 키를 디바이스 키 트리에 등록하는 단계; 상기 사용자 디바이스 단말로부터 상기 음성인식 프로그램의 실행에 따라 상기 디바이스 키가 전송되면, 상기 디바이스 키를 이용하여 사용 가능한 음향모델을 선택하는 단계; 상기 사용 가능한 음향모델 및 상기 사용자 디바이스 단말로부터 전송된 음성 데이터를 이용하여 음성 인식을 수행하는 단계; 음성 인식 결과의 신뢰도를 측정하여 상기 신뢰도가 문턱값을 초과하는 데이터는, 상기 디바이스 키에 따라 다중 모델 트리의 단말 노드 및 직계 상위 노드들에 인식 정보와 함께 데이터 누적시키는 단계; 및 상기 다중모델 트리의 각 노드의 데이터가 기준값을 초과하면, 음향모델 누적 적용을 수행하는 단계를 포함한다.

발명의 효과

- [0020] 본 발명에 의하면, 모든 사용자 환경을 고려한 독립 음향 모델에서 점차 사용자 데이터를 누적하여 사용자 의존 음향 모델을 지향하는 것으로, 충분한 사용자 모델이 모여지지 않은 경우를 고려한 다단계 트리 구성의 음향 모델을 생성함으로써, 독립 음향 모델보다 사용자에게 가까운 음향 모델을 이용할 수 있는 효과가 있다.
- [0021] 또한, 공통 채널 특성을 보이는 다른 사용자들의 누적 데이터를 공동으로 활용함으로써, 채널 잡음을 줄인 음향 모델을 생성할 수 있고, 지속적인 누적 적용 프로세스에 의해 사용자 데이터를 활용한 화자 의존 음향 모델을 만들 수 있다.
- [0022] 또한, 디바이스 키를 구성함으로써, 사용자 ID 발급과 사용자 등록 및 인증을 수행함이 없이, 자동으로 사용자 등록 및 인증을 수행하는 효과가 있다.
- [0023] 더 나아가, 모델 선정에 있어서, 부가적인 인식 수행 없이도 사용자 디바이스 단말에 따라 단일 모델 또는 복수의 모델을 빠르게 선정하여 활용할 수 있다.

도면의 간단한 설명

- [0024] 도 1은 본 발명의 일 실시예에 따른 디바이스 구분 음향모델 누적 적용을 이용한 음성인식 시스템의 구조도이다.
 도 2는 본 발명의 일 실시예에 따른 디바이스 구분 음향모델 누적 적용을 이용한 음성인식 시스템에서 사용하는 디바이스 키의 구성도이다.
 도 3은 본 발명의 일 실시예에 따른 디바이스 구분 음향모델 누적 적용을 이용한 음성인식 방법의 흐름도이다.
 도 4는 본 발명의 일 실시예에 따라 디바이스 구분된 다중 모델 트리의 음향 모델과 누적 데이터 구성도이다.

발명을 실시하기 위한 구체적인 내용

- [0025] 이하에서는 도면을 참조하여 본 발명의 바람직한 실시예들을 상세히 설명한다. 이하 설명 및 첨부된 도면들에서 실질적으로 동일한 구성요소들은 각각 동일한 부호들로 나타냄으로써 중복 설명을 생략하기로 한다. 또한 본 발명을 설명함에 있어 관련된 공지기능 혹은 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그에 대한 상세한 설명은 생략하기로 한다.
- [0026] 도 1은 본 발명의 일 실시예에 따른 디바이스 구분 음향모델 누적 적용을 이용한 음성인식 시스템의

구조도이다.

- [0027] 도 1을 참조하면, 본 발명의 일 실시예에 따른 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템은 모델 선정 모듈(100), 모델 관리 모듈(200) 및 음성 인식 모듈(300)을 포함하는 음성인식 서버와, 음성 인식 프로그램(400)이 설치된 사용자 디바이스 단말을 포함하여 구성될 수 있다.
- [0028] 모델 선정 모듈(100)은 디바이스 키를 확인 및 분류하여 다중 모델 트리의 음향 모델을 선택하기 위한 것으로, 디바이스 키 확인부(110) 및 모델 선택부(120)를 포함하여 구성될 수 있다.
- [0029] 디바이스 키 확인부(110)는 사용자 디바이스 단말로부터 수신한 디바이스 키를 확인하여, 새로운 디바이스 키이면 다중모델 트리(220)의 일 구성요소인 디바이스 키 트리에 등록하고, 이미 가지고 있던 디바이스 키이면 다중 모델 트리(220)로부터 디바이스 키 트리의 사용 가능한 음향모델 리스트를 제공받는다. 여기서, 음향모델 리스트는 도 4를 참조하여 후술하는 바와 같이 다중모델 트리(220)에 포함된 디바이스 독립모델(도 4의 221), 디바이스 의존모델(도 4의 222), 세부 디바이스 의존모델(도 4의 223) 및 랜덤 ID 의존모델(도 4의 224)을 순차적으로 포함할 수 있다.
- [0030] 모델 선택부(120)는 디바이스 키 확인부(110)로부터 전달된 사용 가능한 음향모델 리스트에서 선택 옵션에 따라 최종 병렬 인식에 사용할 다중 모델을 선발하거나 단일 모델을 선택한다. 또한, 모델 선택부(120)는 다중모델 트리(220)로부터 입력된 실제 선택된 모델의 주소와 디바이스 키 정보를 음성 인식 모듈(300)의 인식부(320)로 전달한다.
- [0031] 모델 관리 모듈(200)은 음성 데이터를 사용자 디바이스 단말 별로 분류하여 다중 모델 트리를 생성하고 누적 적응하기 위한 것으로, 디바이스 분류 다중모델 트리 생성부(210), 다중모델 트리(220), 모델 적응 생성부(230) 및 데이터 누적부(240)를 포함하여 구성될 수 있다.
- [0032] 디바이스 분류 다중모델 트리 생성부(210)는 학습용 음성 데이터를 사용자 디바이스 단말 별로 분류 및 특징 추출하여 다중모델 트리를 구성한다. 구체적으로, 디바이스 분류 다중모델 트리 생성부(210)는 학습용 음성 데이터를 사용자 디바이스 단말 별로 분류하고, 각각의 데이터를 디바이스 독립 모델에 적응하여 디바이스 의존 모델을 생성한 다중모델 트리를 만든다.
- [0033] 다중모델 트리(220)는 사용자 디바이스 단말 별로 분류된 다중모델 트리로서, 계층적 음향모델, 누적 데이터 및 디바이스 키 트리를 포함한다. 상술한 디바이스 키 확인부(110)로부터 등록을 위한 새로운 디바이스 키가 전달 되면 디바이스 키 트리에 새로 등록하고 데이터를 수집할 수 있도록 데이터 노드가 생성된다.
- [0034] 모델 적응 생성부(230)는 데이터 누적부(240)로부터 전달된 정보를 이용하여 다중모델 트리(220)에서 새로 누적된 데이터가 기존 시간 이상으로 쌓인 노드에 해당하는 음향 모델을 누적적응 한다. 이 경우, 기존 모델에서 수집된 데이터를 적응 데이터로 활용하는 적응 방법으로, MAP(Maximum a Posteriori), MLLR(Maximum Likelihood Linear Regression), MAPLR 등의 음향모델 적응 기법을 활용할 수 있다.
- [0035] 데이터 누적부(240)는 음성 인식 모듈(300)의 신뢰도 평가부(330)로부터 신뢰도가 문턱 값을 초과하는 데이터를 입력받아 다중 모델 트리(220)에 음소정보와 데이터를 누적시킨다. 구체적으로, 데이터 누적부(240)는 음성 인식 모듈(300)의 신뢰도 평가부(330)로부터 전달받은 디바이스 키를 이용하여 다중모델 트리(220)에서 등록된 노드를 검색하고, 검색된 노드에서 상위 노드까지 포함하여 누적시킨다. 또한, 데이터 누적부(240)는 각 노드에 저장되는 새로 입력된 데이터의 누적 정도를 데이터 시간 기준값으로 평가한 후, 기준값을 초과하는 경우 모델 적응 생성부(230)로 알려준다.
- [0036] 음성 인식 모듈(300)은 모델 선정 모듈(100)로부터 사용자 디바이스 단말에 가장 적합한 음향모델을 제공받아 음성 인식을 수행하고, 신뢰도가 높게 인식된 데이터를 모델 관리 모듈(200)에 전달하기 위한 것으로, 전처리부(310), 인식부(320) 및 신뢰도 평가부(330)를 포함하여 구성될 수 있다.
- [0037] 전처리부(310)는 사용자 디바이스 단말로부터 입력된 음성 신호를 전처리하여 인식부(320)로 제공한다. 이 경우, 사용자 디바이스 단말로부터 음성 데이터만 수신하면, 전처리부(310)는 특징 추출을 비롯하여 CMN(Cepstral mean normalization) 및 끝점 검출, 잡음 제거 등을 수행한다. 반면, 사용자 디바이스 단말로부터 특징 추출된 데이터를 수신하면, 특징 추출을 제외한 나머지 전처리 과정을 수행할 수 있다.
- [0038] 인식부(320)는 전처리부(310)로부터 제공된 특징 추출된 데이터와 모델 선택부(120)에 의해 선발된 음향 모델을 사용하여 음성 인식을 수행한다. 이 경우, 인식부(320)는 모델 선정 모듈(100)의 모델 선택부(120)로부터 복수의 음향 모델을 제공받을 수 있고, 복수의 음향 모델에 대한 병렬 음성 인식의 수행이 가능하다. 인식부(320)는

음성 인식 결과를 신뢰도 평가부(330)로 전달한다.

- [0039] 신뢰도 평가부(330) 인식부(320)로부터 전달받은 음성 인식 결과의 신뢰도를 측정하여 신뢰도가 가장 높은 결과를 인식 결과로서 출력한다. 또한, 신뢰도 평가부(330)는 신뢰도가 문턱 값을 초과하는 경우, 해당 음소정보와 데이터, 그리고 인식부(320)로부터 전달받은 디바이스 키를 모델 관리 모듈(200)의 데이터 누적부(240)로 전달한다.
- [0040] 음성 인식 프로그램(400)은 사용자 디바이스 단말에 설치되는 음성 인식 프로그램으로, 디바이스 키 전송부(410) 및 음성 녹음부(420)를 포함하여 구성될 수 있다.
- [0041] 디바이스 키 전송부(410)는 음성 인식 프로그램(400)이 최초 설치될 때 디바이스 키를 생성하여 모델 선정 모듈(100)의 디바이스 키 확인부(110)로 전달한다. 또한, 디바이스 키 전송부(410)는 음성 인식 프로그램(400)이 사용될 때마다 음성 인식에서 사용자 디바이스 단말에 적합한 모델을 이용할 수 있도록 하기 위해 모델 선정 모듈(100)의 디바이스 키 확인부(110)로 디바이스 키를 전달한다.
- [0042] 음성 녹음부(420)는 음성 인식 프로그램(400)을 실행하여 음성 인식을 수행하는 경우 입력되는 음성 데이터를 녹음하여 음성 인식 모듈(300)의 전처리부(310)로 전달한다. 이 경우, 음성 녹음부(420)는 음성 데이터로부터 특징 추출 후 음성 데이터 대신 특징 추출된 데이터를 음성 인식 모듈(300)의 전처리부(310)로 전달할 수도 있다.
- [0043] 도 2는 본 발명의 일 실시예에 따른 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템에서 사용하는 디바이스 키의 구성도이다.
- [0044] 도 2를 참조하면, 디바이스 키(10)는 사용자 디바이스 단말에 등록되어 있는 고유 정보를 활용하여 생성되는 것으로, 사용자의 동의 없이 활용 가능하도록 구성된 것이다.
- [0045] 디바이스 키(10)는 사용자 디바이스 단말에 음성 인식 프로그램이 설치될 때 자동으로 생성되어 음성인식 서버에 등록용 정보로 제공되고, 이에 따라 다중모델 트리에 새로운 키로 등록되어 이에 적합한 음향모델 그룹이 선정된다. 이후, 음성 인식 프로그램의 사용시, 디바이스 키(10)는 음성인식 서버 인증 및 기 선정된 음향모델을 음성인식에 이용하기 위해 사용된다. 디바이스 키(10)는 디바이스 명(11), 디바이스 세부명(12) 및 랜덤 ID(13)를 포함하여 구성될 수 있다.
- [0046] 디바이스 명(11)은 통상적인 디바이스 명을 의미하는 것으로, 예를 들어 제품의 이름이 이에 해당할 수 있다. 디바이스 명(11)은 디바이스 회사 또는 프로젝트 제품 군의 이름으로, 유사한 채널 특성의 마이크로 폰과 멀티센서를 사용할 가능성이 높다. 디바이스 키(10)의 디바이스 명(11)을 이용하여 음성 데이터를 분류하고, 다중모델 트리를 형성하는데 활용할 수 있다.
- [0047] 디바이스 세부명(12)은 디바이스의 세부 명을 의미하는 것으로, 예를 들어 제품의 버전이 이에 해당할 수 있다. 세부 디바이스 별로 상이한 채널의 마이크로 폰과 멀티센서를 이용할 수 있다. 따라서, 디바이스 세부명(12)을 기준으로 음성 데이터를 분류하고, 다중모델 트리를 형성하는데 활용할 수 있다.
- [0048] 랜덤 ID(13)는 디바이스 별 고유 번호(제품 번호)에 상응하여 랜덤으로 생성된 ID를 의미하는 것이다. 랜덤 ID(13)는 디바이스 명(11)과 디바이스 세부명(12)이 동일한 경우 유일 값이 되어 등록 키로 활용한다. 동일한 디바이스 단말에서 음성인식 성능 향상을 위해 랜덤 ID(13)를 활용한 디바이스 키로 음성 데이터를 분류하고, 다중모델 트리를 형성하는데 활용할 수 있다.
- [0049] 도 3은 본 발명의 일 실시예에 따른 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 방법의 흐름도이다.
- [0050] 우선, 음성인식 프로그램이 설치된 사용자 디바이스 단말 측의 동작을 살펴본다.
- [0051] 음성 입력이 가능한 사용자 디바이스 단말에 음성인식 프로그램을 설치하고(401), 음성인식 프로그램의 최초 설치시 등록을 위한 디바이스 키를 생성하여 음성인식 서버로 전송한다(402).
- [0052] 이후, 음성인식을 위해 설치된 음성인식 프로그램을 실행하면(403), 프로그램 실행시마다 음성인식 서버로 디바이스 키를 전송한다(404).
- [0053] 이후, 음성 입력이 있는지 확인하여(405), 음성 입력이 있으면 입력된 음성 데이터를 음성인식 서버로 전송 및 음성 인식된 문자열을 수신하여(406) 결과를 출력한다(407). 이 경우, 음성 데이터의 특징 추출은 사용자 디바이스 단말에서 수행할 수도 있고, 음성인식 서버에서 수행할 수도 있다. 반면, 음성 입력이 없으면 음성인식 프

로그그램을 종료한다(408). 이후, 음성인식 프로그램을 다시 실행하는 경우 상술한 403 단계로 돌아간다.

- [0054] 다음으로, 도 1을 참조하여 상술한 바와 같이, 모델 선정 모듈, 모델 관리 모듈 및 음성 인식 모듈을 포함하는 음성인식 서버 측의 동작을 살펴본다.
- [0055] 사용자 디바이스 단말로부터 음성인식 프로그램의 설치에 따라 디바이스 키가 전송되면, 디바이스 키를 디바이스 키 트리에 등록하고(501), 다중모델 트리에 등록된 디바이스 키에 해당하는 노드를 생성함으로써(502), 데이터 저장 및 음향 모델 구성을 준비한다. 이때, 최초 등록된 디바이스 키에 대응하는 음향모델은 디바이스 키 분류에 해당하여 생성한 노드의 상위 모델이 된다.
- [0056] 이후, 사용자 디바이스 단말로부터 음성인식 프로그램의 실행에 따라 디바이스 키가 전송되면, 디바이스 키를 확인하여(503) 등록되지 않은 디바이스 키이면 상술한 501 단계로 돌아가고 등록된 디바이스 키이면 사용 가능한 음향모델 리스트를 선택하고 다중모델 트리로부터 주소 값을 전달받는다(504).
- [0057] 이후, 사용 가능한 음향모델과 사용자 디바이스 단말로부터 전송된 음성 데이터를 이용하여 음성 인식을 수행한다(505).
- [0058] 이후, 다수의 결과가 가능한 음성 인식 결과의 신뢰도를 측정하여 신뢰도가 문턱값을 초과하는 데이터를 선별하고(506), 문턱값을 초과하는 데이터는 해당 디바이스 키에 따라 다중모델 트리의 단말 노드 및 직계 상위 노드들에 인식 정보와 함께 누적시킨다(507).
- [0059] 이후, 다중모델 트리의 각 노드에 데이터가 기준값 이상으로 모이면, 음향모델 누적 적응을 수행한다(508).
- [0060] 도 4는 본 발명의 일 실시예에 따라 디바이스 구분된 다중 모델 트리의 음향 모델과 누적 데이터 구성도이다.
- [0061] 본 발명의 일 실시예에서 다중모델 트리는 크게 다음과 같이 세 가지로 구성될 수 있다.
- [0062] 첫째는 전체 데이터를 디바이스 기준으로 세부적으로 분류한 데이터 트리이고, 둘째는 데이터 트리의 노드에 있는 데이터를 적응 데이터로 활용하여 각 노드의 음향 모델을 생성한 음향모델 트리이고, 셋째는 음향 모델 선정을 위해 디바이스 키를 등록한 디바이스 키 트리이다.
- [0063] 디바이스 독립 모델(221)은 기본적으로 음성 인식 서비스에 이용되는 음향 모델로서, 음성 인식 수행시, 특별 병렬 인식시에 비교 모델로 함께 이용될 수 있다. 디바이스 독립 모델(221)은 기존의 학습용 음향 데이터를 가지고 있다.
- [0064] 디바이스 의존 모델(222)은 기존의 학습용 수집 데이터에서 디바이스 의존 데이터를 분류하고 분류된 디바이스 의존 데이터를 적응 데이터로 활용하여 디바이스 독립 모델(221)에서 음향 모델 적응하여 생성된다.
- [0065] 세부 디바이스 의존 모델(223)은 기존의 학습용 수집 데이터에서 디바이스 의존 데이터 분류를 보다 세분화하고, 세분화한 디바이스 의존 데이터를 적응 데이터로 활용하여 디바이스 의존 모델(222)에서 음향 모델 적응하여 생성된다.
- [0066] 랜덤 ID 의존 모델(224)은 다음과 같이 생성된다. 각각의 사용자 디바이스 단말에서 음성인식 프로그램을 설치할 때 전송된 디바이스 키를 이용하여 노드를 생성하고, 사용자 디바이스 단말에서 음성 인식된 데이터 중에 신뢰도가 높은 데이터를 디바이스 키 의존 데이터로 누적 보관한다. 이 경우, 직계 상위 노드들에도 데이터를 누적한다. 누적된 데이터가 기준값을 초과한 경우, 직계 상위 노드인 디바이스 의존 모델(222) 또는 세부 디바이스 의존 모델(223)에 누적 적응하여 랜덤 ID 의존모델(224)을 생성한다.
- [0067] 다음으로 도 1 내지 도 4를 참조하여, 상술한 디바이스 구분 음향모델 누적 적응을 이용한 음성인식 시스템에서, 디바이스 키 구성 및 인증 방법, 디바이스 의존 음향 모델 트리 생성 방법, 음향 모델 트리 누적 갱신 방법 및 음성 인식 모델 선정에 대해 보다 구체적으로 설명한다.
- [0068] **디바이스 키 구성 및 인증 방법**
- [0069] 사용자 디바이스 단말은 도 2에 도시된 바와 같은 디바이스 키 구성을 위해 디바이스 관련 정보를 활용할 수 있다. 예를 들어, 스마트폰의 경우, 음성인식을 위한 앱 프로그램 설치시, 디바이스 명에 대한 정보와 기타 디바이스 세부명 등의 정보를 음성인식 서버로 제공한다.
- [0070] 디바이스 명은 음성 녹음을 수행하는 사용자 디바이스 단말의 채널 특성을 크게 구분 짓기 때문에 중요한 정보이다. 예를 들어, 사용자 디바이스 단말이 태블릿인 경우와 스마트폰인 경우, 각각이 사용하는 기본 마이크 특성이 상이하므로, 녹음된 음성 데이터에 최적인 음향 모델도 상이하다. 또한, 스마트폰의 경우에도 제조사에 따

라 음성 데이터의 녹음 특성이 상이하다.

[0071] 또한, 디바이스 세부명은 사용자 디바이스 단말의 세부 구분 또는 버전 정보로서, 사용자 디바이스 단말의 새로운 버전이 출시되어 채널 특성이 명시적으로 달라진 경우 활용 가능하다.

[0072] 또한, 랜덤 ID는 개별 사용자를 구분하기 위해 랜덤 넘버를 할당하여 생성한 숫자이다. 이는, 개인 정보에 해당하는 스마트폰 전화번호나, 네트워크 장비주소 (MAC address)를 수집하는 것은 개인 정보 보호를 저해하므로, 개인과 연계성을 알 수 없도록 음성인식 프로그램의 설치시 랜덤 넘버를 생성한 것이다.

[0073] 본 발명에서는, 종래의 화자의 등록 정보를 이용하여 화자에 적합한 음성인식을 수행하는 방법에서 초기에 화자를 등록하고 사용시에 화자 인증을 거치는 불편함을 해소하기 위해, 상술한 바와 같이 구성된 디바이스 키를 활용한다.

[0074] 즉, 사용자 디바이스 단말에 음성인식 프로그램 설치시 자동 생성된 디바이스 키는 음성인식 서버에 자동 등록되고, 이후 음성인식 프로그램 실행시 디바이스 키에 의해 자동 인증함으로써 사용자의 불편함을 해소할 수 있다.

[0075] 구체적으로, 디바이스 키는 트리 형태로 분류된 디바이스 키 검색을 통하여 상위 노드인 디바이스 명과 디바이스 세부명 노드를 거쳐 단말 노드에 랜덤ID로 등록하고, 프로그램 실행시에는 디바이스 키 트리 검색을 통해 인증한 후 적합한 음향 모델을 선택하게 된다. 디바이스 키가 등록되어 디바이스 키 트리가 생성되면, 동일한 형태의 음향 모델 트리와 누적 데이터 트리도 생성된다.

[0076] **디바이스 의존 음향 모델 트리 생성 방법**

[0077] 디바이스 의존 음향 모델 트리는 기 수집된 학습용 음향 데이터를 이용하여 디바이스 독립인 음향 모델을 생성하고, 이 음향 모델을 트리 구성을 위한 전체의 루트 노드(Root Node)로 한다. 또한, 음향 모델 트리와 동일한 구조를 갖는 데이터 트리를 함께 구성하며, 이를 위해 기 수집된 학습용 데이터를 전체의 루트 노드에 위치시키고, 디바이스 키의 분류에 대응하는 데이터들을 루트 노드에서 하위 노드로 분류하여 저장한다.

[0078] 본 발명에서 제안하는 음향 모델 트리의 핵심은 채널이 상이한 사용자 디바이스 단말에서 녹음한 데이터의 분류 및 공유이다. 데이터 분류는 동일한 디바이스에 해당하는 데이터는 상위 노드에서 공유하고, 하위 노드로 갈수록 세부 디바이스와 개인 데이터로 나누어 화자 의존 데이터로 확장할 수 있도록 구성된다. 또한, 데이터 분류는 기존 수집된 데이터를 활용하여 명시적인 분류를 할 수 있고, 디바이스 키 생성 기준에 맞추어 규칙 기반으로 자동 분류할 수도 있다.

[0079] 도 4에 도시된 바와 같이, 상위 노드에 디바이스명을 기준으로 채널이 상이한 큰 그룹을 설정하고, 하위로 갈수록 디바이스 세부명으로 분류하며, 트리의 단계는 그 구성 방법에 따라 여러 단계를 거치게 하거나, 분류를 다양하게 구성할 수 있다. 단말 노드에는 디바이스 키의 랜덤 ID에 해당하는 개별 디바이스 노드가 생성되고, 개별 디바이스에서 수집한 데이터가 저장된다.

[0080] 생성한 트리의 각 노드에 존재하는 음향 데이터가 기준값 이상으로 수집된 경우, 수집된 데이터로 해당 노드의 음향 모델을 생성한다. 반면, 음향 데이터가 기준값이 미치지 않는 경우, 데이터 트리는 확장되면서 커지지만, 음향 모델 트리는 루트를 포함한 몇몇 상위 노드만 학습 모델을 생성하고, 하위 노드는 학습 모델 없이 존재하게 된다.

[0081] **음향 모델 트리 누적 갱신 방법**

[0082] 음향 모델 트리는 동일한 구성을 갖는 데이터 트리와 디바이스 키 트리를 이용하여 생성 및 관리된다. 본 발명은 온라인으로 사용자들의 데이터를 지속적으로 수집하고 모델에 반영하여 개별 사용자의 음성인식 성능을 향상시키기 위한 것으로, 이를 위해서는 개별 사용자의 데이터를 누적 시키는 방법이 필요하다.

[0083] 우선, 사용자의 음성 데이터와 인식된 전사문에 대한 신뢰도 평가를 수행하여 문턱값을 초과하는 신뢰도를 갖는 인식 결과에 대한 데이터를 수집한다. 이 경우, 신뢰도 평가는 발화 검증(Utterance Verification) 기술에서 사용하는 우도비 검증(likelihood ratio test)을 이용할 수 있다.

[0084] 문턱값을 초과하는 신뢰도를 갖는 사용자 인식 데이터는 해당 사용자의 디바이스 키를 이용하여 데이터 트리의 단말 노드에 저장되고, 각 단말 노드와 하위 노드는 연결된 상위 노드의 데이터로 활용된다. 데이터의 트리 분류는 데이터를 쪼개어 작아지는 원인이 되므로, 적은 데이터를 보유한 하위 노드는 이 데이터를 적응 데이터로 이용하고, 상위의 학습된 음향 모델을 활용하여 MAP, MLLR 등의 모델 적응 기술에 의해 하위 노드 음향 모델을

생성할 수 있다.

[0085] 음향 모델 누적 갱신은 사용자에게 대해 새로 수집한 데이터의 분량이 기준값 이상으로 많아지면 자동으로 상위 노드의 음향 모델을 활용한 모델 적응을 수행하여 갱신하는 것이다. 이에 따라 단말 노드에서 상위 노드로 음향 모델을 갱신하게 되는데, 하위 노드 음향 모델을 적응할 때 상위 노드의 음향 모델이 없는 경우에는 음향 모델이 존재하는 연결된 상위 모델을 찾아 참조할 수 있다.

[0086] 이처럼, 음향 모델 트리는 단말 노드에서 모이는 데이터를 주기적으로 점검하여 하위 노드의 음향 모델부터 음향 모델 적응을 이용하여 갱신한다. 상위 노드로 갈수록 데이터 분량이 학습 기준 분량을 넘을 경우에는 재학습 과정을 거치고, 그렇지 못하면, 새로 수집한 데이터를 기준으로 음향 모델 적응을 하여 갱신한다.

[0087] **음성 인식 모델 선정**

[0088] 사용자에게 적합한 음향 모델의 선정을 위해, 디바이스 키를 이용하여 음향 모델 트리의 단말 노드에서 모델을 찾아 선정하고, 해당 단말 노드에 음향 모델이 없는 경우에는 연결된 상위 노드의 음향 모델을 선정한다. 병렬 인식기를 사용하는 경우, 복수의 음향 모델을 사용 가능하므로, 디바이스 키에 해당하는 사용자 디바이스 단말의 음향 모델과 상위 노드의 음향 모델들 중에서 정해진 규칙에 따라 선정할 수 있다.

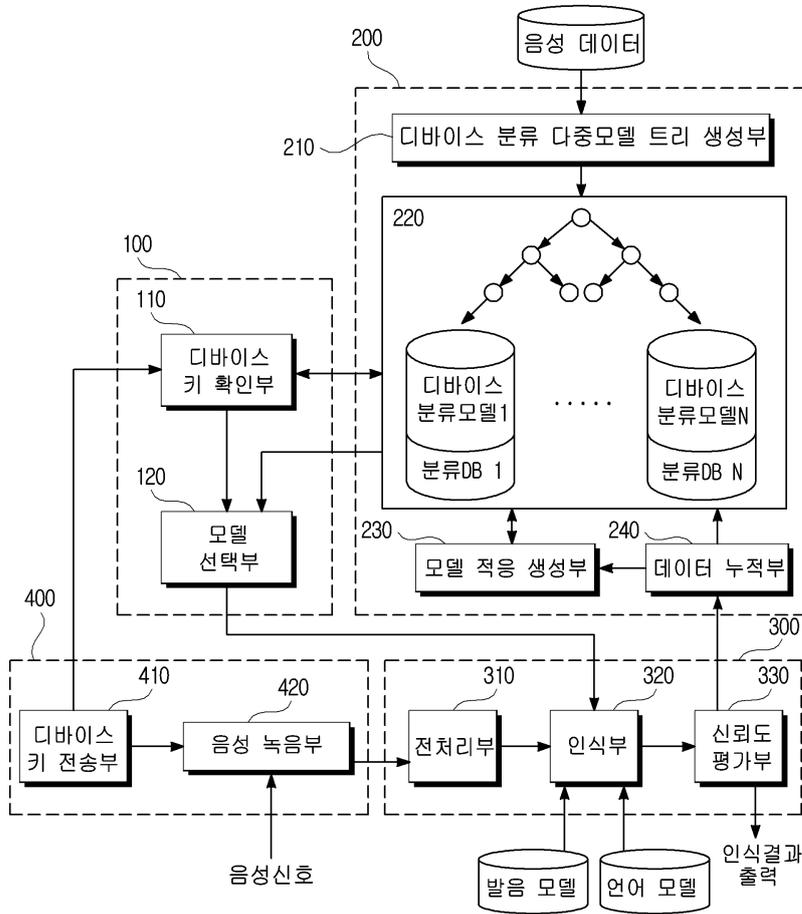
[0089] 이제까지 본 발명에 대하여 그 바람직한 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

부호의 설명

- [0090] 100: 모델 선정 모듈
- 110: 디바이스 키 확인부
- 120: 모델 선택부
- 200: 모델 관리 모듈
- 210: 디바이스 분류 다중모델 트리 생성부
- 220: 다중모델 트리
- 230: 모델 적응 생성부
- 240: 데이터 누적부
- 300: 음성 인식 모듈
- 310: 전처리부
- 320: 인식부
- 330: 신뢰도 평가부
- 400: 음성 인식 프로그램
- 410: 디바이스 키 전송부
- 420: 음성 녹음부

도면

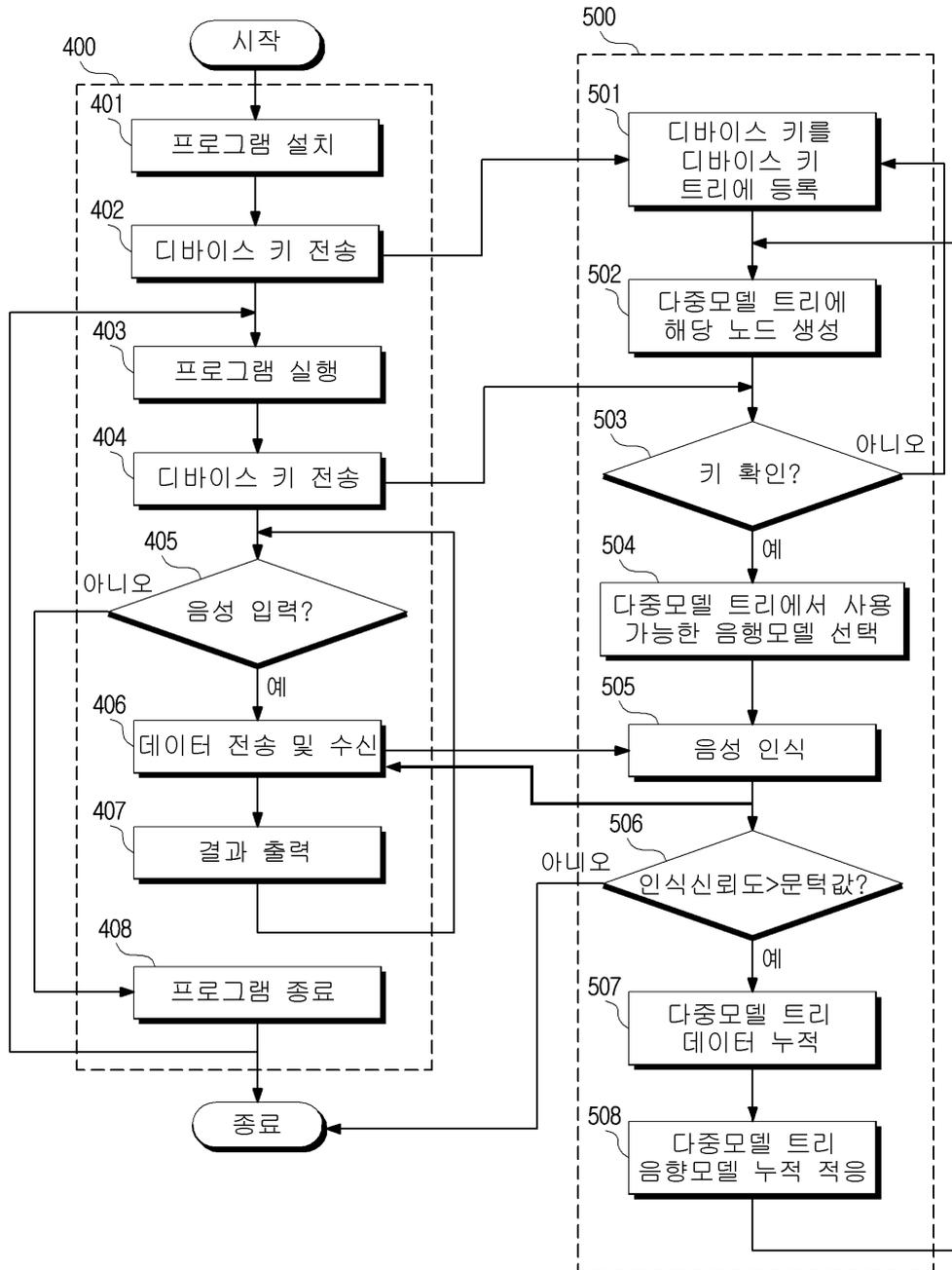
도면1



도면2



도면3



도면4

