



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2012-0066530
(43) 공개일자 2012년06월22일

(51) 국제특허분류(Int. Cl.)
G10L 15/14 (2006.01) G10L 15/10 (2006.01)
(21) 출원번호 10-2010-0127907
(22) 출원일자 2010년12월14일
심사청구일자 없음

(71) 출원인
한국전자통신연구원
대전광역시 유성구 가정로 218 (가정동)
(72) 발명자
전형배
대전광역시 서구 청사로 65, @ 109-605호 (월평동, 황실타운)
이윤근
대전광역시 서구 월평전사로 11, 무지개@ 103-1406 (월평동)
(74) 대리인
특허법인이상

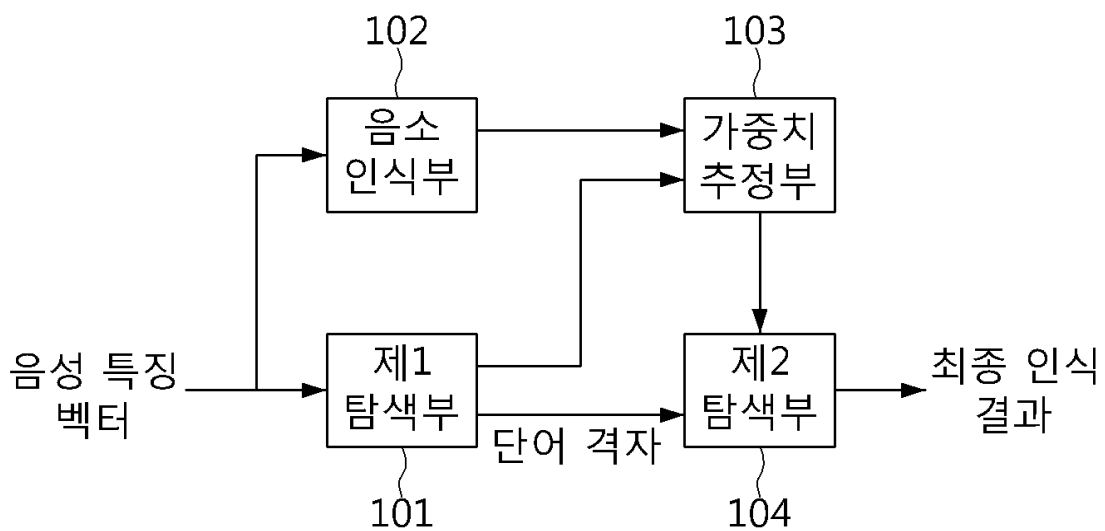
전체 청구항 수 : 총 14 항

(54) 발명의 명칭 언어 모델 가중치 추정 방법 및 이를 위한 장치

(57) 요약

연속어 음성 인식 결과에 따라 적응적으로 언어 모델 가중치를 추정하는 언어 모델 가중치 추정 방법 및 이를 위한 장치가 개시된다. 음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 음성 특징 벡터에 제 1 언어 모델을 적용하여 제 1 차 탐색을 수행하고, 단어 격자 및 단어 격자의 제 1 음향 스코어를 출력하는 제 1 탐색부, 음성 특징 벡터에 음향 모델을 적용하여 제 2 음향 스코어를 출력하는 음소 인식부, 제 1 음향 스코어 및 제 2 음향 스코어를 이용하여 상기 제 1 탐색부의 연속어 음성 인식 결과와 음소 인식부의 음소 인식 결과의 음향 스코어를 비교하여 상기 음성 인식 결과가 음소 인식 결과보다 음향 스코어가 좋은 경우 상기 제 1 언어 모델 가중치를 출력하는 가중치 추정부 및 제 1 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 적용하여 제 2 차 탐색을 수행하여 최종 인식 결과를 출력하는 제2 탐색부를 포함하여 구성될 수 있다. 따라서, 제 1차 탐색에서 정답을 찾지 못한 경우에 대해서 상대적으로 낮은 언어 모델 가중치를 적용하여, 음향 스코어의 비중을 높여 정답을 찾을 가능성을 높이게 하고, 제 2차 탐색의 성능을 높여 연속어 음성인식 시스템의 성능향상을 기대할 수 있다.

대표도



(72) 발명자

정의석

대전광역시 유성구 은구비남로 55, 705동 1301호
(지족동, 열매마을7단지)

김종진

대전광역시 서구 월평선사로 41, 백합APT 104-202
(월평동)

정훈

강원도 홍천군 홍천읍 연봉중앙로 18, 현대APT
103-1101

박전규

서울특별시 서초구 잠원동 53 잠원현대(아)
102-406

정호영

대전광역시 서구 대덕대로 415, 102동 606호 (만
년동, 상아아파트)

강병욱

충청남도 계룡시 두마면 사계로 101, 포스코 109
동 301호 (계룡 더샵 아파트)

박기영

대전광역시 유성구 지족로 362, 반석마을 아파트
302동 703호 (지족동)

이성주

대전광역시 유성구 송강로 15, 한솔@ 202동 106호
(송강동)

강점자

대전광역시 서구 둔산로 155, 크로바아파트 119동
407호 (둔산동)

송화전

대전광역시 유성구 은구비남로 56, 열매마을아파
트9단지 903-1802 (노은동)

이 발명을 지원한 국가연구개발사업

과제고유번호	10035252
부처명	지식경제부
연구사업명	정보통신산업원천기술개발사업
연구과제명	모바일 플랫폼 기반 대화모델 적용 자연어 음성 인터페이스 기술 개발
주관기관	한국전자통신연구원
연구기간	2010.03.01 ~ 2011.02.28

특허청구의 범위

청구항 1

언어 모델 가중치 추정 장치에 있어서,

음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 상기 수신된 음성 특징 벡터에 제 1 언어 모델을 적용하여 제 1 차 탐색을 수행하고, 단어 격자 및 단어 격자의 제 1 음향 스코어를 출력하는 제 1 탐색부;

상기 음성 특징 벡터에 음향 모델을 적용하여 제 2 음향 스코어를 출력하는 음소 인식부;

상기 제 1 음향 스코어 및 상기 제 2 음향 스코어를 이용하여 상기 제 1 탐색부의 연속어 음성 인식 결과와 음소 인식부의 음소 인식 결과의 음향 스코어를 비교하여 상기 음성 인식 결과가 음소 인식 결과보다 음향 스코어가 좋은 경우 상기 제 1 언어 모델 가중치를 출력하는 가중치 추정부; 및

상기 단어 격자에 상기 출력된 제 1 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 적용하여 제 2 차 탐색을 수행하여 최종 인식 결과를 출력하는 제2 탐색부를 포함하는 것을 특징으로 하는 가중치 추정 장치.

청구항 2

제 1 항에 있어서, 상기 가중치 추정부는,

상기 연속어 음성 인식 결과가 상기 음소 인식 결과보다 음향 스코어가 좋지 않을 경우 상기 제 1 언어 모델 가중치보다 낮은 가중치를 출력하는 것을 특징으로 하는 가중치 추정 장치.

청구항 3

제 2 항에 있어서, 제2 탐색부는,

상기 출력된 제 1 언어 모델 가중치보다 낮은 제 2 언어 모델 가중치를 적용하여 제 2 차 탐색을 수행하는 것을 특징으로 하는 가중치 추정 장치.

청구항 4

제 1 항에 있어서, 상기 가중치 추정부는,

상기 연속어 음성 인식 결과가 상기 음소 인식 결과보다 음향 스코어가 좋지 않고 나쁘지 않은 경우 상기 제 1 언어 모델 가중치 및 제 2 언어 모델 가중치 사이의 값에 해당하는 중간의 가중치를 출력하는 것을 특징으로 하는 가중치 추정 장치.

청구항 5

제 4 항에 있어서, 제 2 탐색부는,

상기 출력된 제 1 언어 모델 가중치 및 제 2 언어 모델 가중치 사이의 값에 해당하는 언어모델 가중치를 적용하여 제 2 차 탐색을 수행하는 것을 특징으로 하는 가중치 추정 장치.

청구항 6

제 1 항에 있어서, 상기 단어 격자는,

상기 제 1 탐색부에서 탐색된 복수개의 단어 조합을 단어간 연결 정보로 정의한 것을 특징으로 하는 가중치 추정 장치.

청구항 7

제 1 항에 있어서, 상기 단어 격자는,

상기 제 1 탐색부의 연속어 음성 인식 결과를 평가하는데 사용되는 정보를 포함하는 것을 특징으로 하는 가중치 추정 장치.

청구항 8

언어 모델 가중치 추정 방법에 있어서,

음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 상기 수신된 음성 특징 벡터에 제 1 언어 모델을 적용하여 제 1 차 탐색을 수행하고, 단어 격자 및 단어 격자의 제 1 음향 스코어를 출력하는 단계;

상기 음성 특징 벡터에 음향 모델을 적용하여 제 2 음향 스코어를 출력하는 단계;

상기 제 1 음향 스코어 및 상기 제 2 음향 스코어를 이용하여 상기 제 1 탐색부의 연속어 음성 인식 결과와 음소 인식부의 음소 인식 결과의 음향 스코어를 비교하는 단계;

상기 비교 결과 상기 음성 인식 결과가 음소 인식 결과보다 음향 스코어가 좋은 경우 상기 제 1 언어 모델 가중치를 출력하는 단계; 및

상기 단어 격자에 상기 출력된 제 1 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 적용하여 제 2 차 탐색을 수행하는 단계를 포함하는 것을 특징으로 하는 가중치 추정 방법.

청구항 9

제 8 항에 있어서, 상기 가중치를 출력하는 단계는,

상기 비교 결과 상기 음성 인식 결과가 음소 인식 결과보다 음향 스코어가 좋지 않은 경우 상기 제 1 언어 모델 가중치보다 낮은 가중치를 출력하는 것을 특징으로 하는 가중치 추정 방법.

청구항 10

제 9 항에 있어서, 제 2 차 탐색을 수행하는 단계는,

상기 제 1 언어 모델 가중치보다 낮은 제 2 언어 모델 가중치를 적용하는 것을 특징으로 하는 가중치 추정 방법.

청구항 11

제 8 항에 있어서, 상기 가중치를 출력하는 단계는,

상기 비교 결과 상기 음성 인식 결과가 음소 인식 결과보다 음향 스코어가 좋지 않고 나쁘지 않은 경우 상기 제 1 언어 모델 가중치 및 제 2 언어 모델 가중치 사이의 값에 해당하는 중간의 가중치를 출력하는 것을 특징으로 하는 가중치 추정 방법.

청구항 12

제 11 항에 있어서, 제 2 차 탐색을 수행하는 단계는,

상기 제 1 언어 모델 가중치 및 제 2 언어 모델 가중치 사이의 값에 해당하는 언어모델 가중치를 적용하여 제 2 차 탐색을 수행하는 것을 특징으로 하는 가중치 추정 방법.

청구항 13

제 8 항에 있어서, 상기 단어 격자는,

상기 제 1 탐색에서 탐색된 복수개의 단어 조합을 단어간 연결 정보로 정의한 것을 특징으로 하는 가중치 추정 방법.

청구항 14

제 8 항에 있어서, 상기 단어 격자는,

상기 제 1 탐색에서 연속어 음성 인식 결과를 평가하는데 사용되는 정보를 포함하는 것을 특징으로 하는 가중치 추정 방법.

명세서

기술 분야

[0001] 본 발명은 언어 모델 가중치 추정 방법 및 이를 위한 장치에 관한 것으로, 더욱 상세하게는 연속어 음성 인식 결과에 따라 적응적으로 언어 모델 가중치를 추정하는 언어 모델 가중치 추정 방법 및 이를 위한 장치에 관한 것이다.

배경 기술

[0002] 일반적으로, 연속 음성 인식에서는 언어 모델이 매우 중요한 역할을 하며, 음향 모델의 오류를 수정하고, 정확한 인식 결과를 나타내는데 사용되고 있다. 그러나, 현재 가장 많이 사용되고 있으며 음성 인식 산업의 표준처럼 사용되고 있는 N-gram의 경우, 학습 데이터의 영역에 따라 매우 다른 결과가 나올 수 있으며, 인식하고자 하는 음성의 영역과 언어모델 학습 데이터의 영역이 일치할 때 가장 나은 성능을 기대할 수 있다. 따라서, 방송 뉴스 인식과 같이 여러 영역을 다루는 음성 인식에서는 각 주제마다 언어모델의 영역을 적응하는 언어모델 적응 방법을 통해서 음성 인식의 정확도를 향상시킬 수 있다.

[0003] 이러한 연속 음성 인식을 위한 연속 음성 인식 엔진에서 발화에 대해 여러 인식후보인 다중인식 후보열(N-best list)을 생성하는 과정은, 단어 간의 연결을 그래프 형태로 정의한 문법 네트워크에 따라 탐색을 시도하여 단어 격자(Word lattice) 등을 중간 인식 결과를 출력하고, 단어 간 연결(Collocation) 정보, 바이그램(bigram) 및 트라이그램(trigram) 확률 등의 통계적 언어모델 등의 정보나 A-STAR 알고리즘 등을 적용하여 단어 격자를 재평가하여 N-best 리스트를 생성한다.

[0004] 이와 같이 N-best 인식 후보를 생성하는 과정에서, 언어 모델은 탐색 공간이 넓고 인식에 필요한 계산량이 많기 때문에 낮은 차수의 언어 모델을 제 1 차 탐색에 사용하여 제 1 차 탐색 결과로 단어 격자를 출력하고, 제 1차 탐색 결과인 단어 격자에 제1 차 탐색에 적용된 언어 모델보다 높은 차수의 언어 모델, 도메인 특화된 언어 모델을 적용하여 제 2 차 탐색을 수행한다.

[0005] 제 2 차 탐색은 단어 격자의 주어진 단어들의 시작과 끝점은 고정하고 음향 스코어 또는 언어 모델 스코어만을 새로 계산하기 때문에 제 1 차 탐색에 비해 상대적으로 매우 적은 계산량을 필요로 한다. 이때, 제 2 차 탐색에서 언어 모델을 적용할 때 단어 격자 상의 음향 스코어에 대한 가중치는 실험을 통해 결정된 값을 고정하여 사용한다. 하지만, 제 2 차 탐색에 사용되는 가중치를 고정하여 사용할 경우 제 1 차 탐색 결과에 따라 적응적으로 제 2 차 탐색을 수행할 수 없다는 문제점이 있다. 즉, 1 차 탐색 결과 연속어 음성 인식 결과의 스코어가 낮아 정답을 찾지 못한 경우에도 동일한 가중치를 사용하여 제 2 차 탐색을 수행하기 때문에 정답을 찾을 가능성이 낮아 진다는 문제점이 있다.

발명의 내용

해결하려는 과제

[0006] 상기와 같은 문제점을 해결하기 위한 본 발명의 제 1목적은, 연속어 음성 인식 결과에 따라 적응적으로 언어 모델 가중치를 추정하는 언어 모델 가중치 추정 장치를 제공하는데 있다.

[0007] 상기와 같은 문제점을 해결하기 위한 본 발명의 제 2목적은, 연속어 음성 인식 결과에 따라 적응적으로 언어 모델 가중치를 추정하는 언어 모델 가중치 추정 방법을 제공하는데 있다.

과제의 해결 수단

[0008] 상기한 본 발명의 제 1목적 달성을 위한 본 발명의 일 실시예에 따른 언어 모델 가중치 추정 장치는, 음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 상기 수신된 음성 특징 벡터에 제 1 언어 모델을 적용하여 제 1 차 탐색을 수행하고, 단어 격자 및 단어 격자의 제 1 음향 스코어를 출력하는 제 1 탐색부, 상기 음성특징벡터에 음향 모델을 적용하여 제 2 음향 스코어를 출력하는 음소 인식부, 상기 제 1 음향 스코어 및 상기 제 2 음향 스코어를 이용하여 상기 제 1 탐색부의 연속어 음성 인식 결과와 음소 인식부의 음소 인식 결과의 음향스코어를 비교하여 상기 연속어 음성 인식 결과가 상기 음소 인식 결과보다 음향 스코어가 좋은 경우 상기 제 1 언어 모델 가중치를 출력하는 가중치 추정부 및 상기 출력된 제 1 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 적용하여 제 2 차 탐색을 수행하여 최종 인식 결과를 출력하는 제2 탐색부를 포함하여 구성될 수 있다.

[0009] 본 발명의 제 2 목적을 달성하기 위한 본 발명의 일 실시예에 따른 언어 모델 가중치 추정 방법은, 음성 신호

로부터 변환된 음성 특징 벡터를 제 1 탐색부가 수신하고, 상기 수신된 음성 특징 벡터에 제 1 언어 모델을 적용하여 제 1 차 탐색을 수행하여 단어 격자 및 단어 격자의 제 1 음향 스코어를 가중치 추정부로 출력하는 단계, 상기 음성 특징 벡터에 음향 모델을 음소 인식부가 적용하여 제 2 음향 스코어를 계산하여 상기 가중치 추정부로 출력하는 단계, 상기 출력된 제 1 음향 스코어 및 상기 제 2 음향 스코어를 이용하여 상기 제 1 탐색부의 연속어 음성 인식 결과 음향 스코어와 음소 인식부의 음소 인식 결과의 음향 스코어를 가중치 추정부가 비교하여 상기 연속어 음성 인식 결과 음향 스코어가 상기 음소 인식 결과 음향 스코어 보다 좋다면 상기 제 1 언어 모델 가중치를 제 2 탐색부로 출력하는 단계 및 상기 출력된 제 1 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 제 2 탐색부가 적용하여 제 2 차 탐색을 수행하여 최종 인식 결과를 출력하는 단계를 포함하여 구성될 수 있다.

발명의 효과

[0010] 상기와 같은 본 발명에 따른 연속어 음성 인식 결과에 따라 적응적으로 언어 모델 가중치를 추정하는 언어 모델 가중치 추정 방법 및 이를 위한 장치를 이용할 경우에는 제 1 차 탐색에서 정답을 찾은 경우에 대해서 동일한 언어 모델 가중치를 적용하고, 제 1차 탐색에서 정답을 찾지 못한 경우에 대해서 상대적으로 낮은 언어 모델 가중치를 적용하여, 음향 스코어의 비중을 높여 정답을 찾을 가능성을 높이게 하고, 제 2차 탐색의 성능을 높여 연속어 음성인식 시스템의 성능향상을 기대할 수 있다.

도면의 간단한 설명

[0011] 도 1은 본 발명의 일 실시예에 따른 언어 모델 가중치 추정 장치의 내부 구조를 개략적으로 도시한 도면이다. 도 2는 본 발명의 일 실시예에 따른 언어 모델 가중치 추정과정을 설명하기 위한 흐름도이다.

발명을 실시하기 위한 구체적인 내용

[0012] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나, 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 각 도면을 설명하면서 유사한 참조부호를 유사한 구성요소에 대해 사용하였다.

[0013] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 본 발명의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.

[0014] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.

[0015] 본 출원에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 출원에서, "포함하다" 또는 "가지다" 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.

[0016] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 관련 기술의 문맥 상 가지는 의미와 일치하는 의미를 가지는 것으로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.

- [0017] 이하, 본 발명에 따른 바람직한 실시예를 첨부된 도면을 참조하여 상세하게 설명한다.
- [0018] 도 1은 본 발명의 일 실시예에 따른 언어 모델 가중치 추정 장치의 내부 구조를 개략적으로 도시한 도면이다.
- [0019] 도 1을 참조하면, 본 발명의 일 실시예에 따른 언어 모델 가중치 추정 장치는 제 1 탐색부(101), 음소 인식부(102), 가중치 추정부(103) 및 제 2 탐색부(104)를 포함하여 구성될 수 있다.
- [0020] 제 1 탐색부(101)는 음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 수신된 음성 특징 벡터에 제 1 언어 모델을 적용하여 제 1 차 탐색을 수행하여 단어 격자(Word lattice)를 제 2 탐색부(104)로 출력하고, 단어 격자의 제 1 음향 스코어를 가중치 추정부(103)로 출력한다. 여기서, 단어 격자는 제 1 탐색부(101)에서 탐색된 복수개의 단어 조합을 단어간 연결 정보로 정의한 것으로, 제 1 탐색부(101)의 음성 인식 결과를 평가하는데 사용되는 정보, 예를 들어 제 1 탐색부(101)로부터 실행된 제 1 차 탐색에서 가능성이 높다고 추정되는 단어들의 제 1 음향 스코어, 언어 모델 스코어, 단어의 시작점 및 끝점, 각 단어의 이전 패스에 따른 언어 모델 스코어 정보 등의 정보를 포함하고 있다.
- [0021] 음소 인식부(102)는 음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 수신된 음성 특징 벡터에 음향 모델을 적용하여 단어 격자의 제 2 음향 스코어를 출력한다. 본 발명의 일 실시예에 따르면, 음소 인식부(102)는 음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 한국어의 "자음+모음+자음"으로 구성되는 음소 탐색 네트워크에 기반하여 음소 인식을 수행하여 단어 격자의 제 2 음향 스코어를 가중치 추정부(103)로 출력한다.
- [0022] 가중치 추정부(103)는 제 1 탐색부(101)에서 출력된 제 1 음향 스코어 및 음소 인식부(102)에서 출력된 제 2 음향 스코어를 수신하고, 수신된 제 1 음향 스코어와 제 2 음향 스코어를 <수학식 1>의 변환함수를 적용하여 가중치를 추정한다.

수학식 1

[0023]
$$LLR = \text{제 2 음향 스코어} - \text{제 1 음향 스코어}$$

- [0024] <수학식 1>을 참조하면, LLR(Log-Likelihood Ratio)는 가중치 추정부(103)에서 제 1 탐색부(101)에서 출력된 제 1 음향 스코어 및 음소 인식부(102)에서 출력된 제 2 음향 스코어를 이용하여 제 2 언어 모델의 가중치를 추정하기 위해서 사용되는 추정값으로 사용되는 로그 우도 율 값이다. 제 1 음향 스코어는 제 1 탐색부(101)에서 결정된 음성 인식 결과 1best 단어열의 음향 스코어의 로그(log)값을 나타내고, 제 2 음향 스코어는 음소 인식부(102)에서 결정된 음소 인식 결과 1best 음소열의 음향 스코어의 로그값을 나타낸다.
- [0025] 먼저, 가중치 추정부(103)가 제 1 음향 스코어와 제 2 음향 스코어를 이용하여 제 1 탐색부(101)의 연속어 음성 인식 결과인 제 1 음향 스코어가 음소 인식부(102)의 음소 인식 결과인 제 2 음향 스코어 보다 상대적으로 좋다고 판단하여, 인식 결과의 신뢰가 높다고 예측함에 따라 가중치를 추정하는 경우를 설명하기로 한다. 본 발명의 일 실시예에 따르면, 가중치 추정부(103)는 제 1 탐색부(101)로부터 출력된 연속어 음성 인식 결과 및 음소 인식부(102)로부터 출력된 음소 인식 결과를 <수학식 1>에 적용하여 계산된 값과 임계값 1과 비교하여 계산된 값이 작다고 판단한 경우, 제1 탐색부(101)의 연속어 음성 인식 결과 음향 스코어가 음소 인식부(102)의 음소 인식 결과의 음향 스코어 보다 상대적으로 좋은 것이며, 인식결과 신뢰도가 높다고 예측하고, 제 1 탐색부(101)에서 사용한 언어 모델 가중치를 제 2 탐색부(104)로 출력한다.
- [0026] 둘째, 가중치 추정부(103)가 제 1 음향 스코어와 제 2 음향 스코어를 이용하여 제 1 탐색부(101)의 연속어 음성 인식 결과인 제 1 음향 스코어가 음소 인식부(102)의 음소 인식 결과인 제 2 음향 스코어 보다 상대적으로 나쁘다고 판단하여, 인식 결과의 신뢰도가 낮다고 예측함에 따라 가중치를 추정하는 경우를 설명하기로 한다. 본 발명의 일 실시예에 따르면, 가중치 추정부(103)는 제 1 탐색부(101)로부터 출력된 연속어 음성 인식 결과 및 음소 인식부(102)로부터 출력된 음소 인식 결과를 <수학식 1>에 적용하여 계산된 값과 임계값 2와 비교하여 계산된 값이 크다고 판단한 경우, 제1 탐색부(101)의 연속어 음성 인식 결과 음향스코어가 음소 인식부(102)의 음소 인식 결과 음향 스코어 보다 상대적으로 좋지 않은 것이며, 인식결과 신뢰도가 낮다고 예측하고, 제 1 탐색부(101)에서 사용한 언어 모델 가중치보다 낮은 가중치를 제 2 탐색부(104)로 출력한다.

- [0027] 셋째, 가중치 추정부(103)가 제 1 음향 스코어와 제 2 음향 스코어를 이용하여 제 1 탐색부(101)의 연속어 음성 인식 결과인 제 1 음향 스코어가 음소 인식부(102)의 음소 인식 결과인 제 2 음향 스코어 보다 좋지도 않고 나쁘지도 않다고 판단한 경우, 제 1 언어 모델의 가중치와 제 2 언어 모델의 가중치의 중간값에 해당하는 가중치를 제 2 탐색부(104)로 출력한다.
- [0028] 제 2 탐색부(104)는 가중치 추정부(103)로부터 수신된 가중치를 제 2언어 모델에 적용하여 제 2 차 탐색을 수행한다. 먼저, 제 2 탐색부(104)는 가중치 추정부(103)가 제 1 음향 스코어와 제 2 음향 스코어를 이용하여 제 1 탐색부(101)의 연속어 음성 인식 결과인 제 1 음향 스코어가 음소 인식부(102)의 음소 인식 결과인 제 2 음향 스코어 보다 상대적으로 좋다고 판단하여, 인식 결과의 신뢰가 높다고 예측함에 따라 제 1 탐색부(101)에서 사용된 언어 모델 가중치를 수신하고, 수신된 제 1 탐색부(101)에서 사용된 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 제 1 탐색부(101)로부터 출력된 단어 격자에 적용하여 제 2 차 탐색을 수행한다. 이때, 제 2 탐색부(104)는 제 1 탐색부(101)로부터 출력된 단어 격자의 음향 스코어를 고정하고, 가중치 추정부(103)로부터 수신된 제 1 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 적용한 경우의 언어 모델 스코어를 계산하여 전체 패스의 누적 스코어를 업데이트하여 가장 높은 스코어를 가진 인식 결과를 최종 인식 결과로 출력한다.
- [0029] 둘째, 제 2 탐색부(104)는 가중치 추정부(103)가 제 1 음향 스코어와 제 2 음향 스코어를 이용하여 제 1 탐색부(101)의 연속어 음성 인식 결과인 제 1 음향 스코어가 음소 인식부(102)의 음소 인식 결과인 제 2 음향 스코어 보다 상대적으로 나쁘다고 판단하여, 인식 결과의 신뢰도가 낮다고 예측함에 따라 제 1 탐색부(101)에서 사용된 언어 모델 가중치보다 낮은 가중치를 수신하고, 수신된 제 1 탐색부(101)에서 사용된 언어 모델 가중치보다 낮은 제 2 언어 모델 가중치를 제 1 탐색부(101)로부터 출력된 단어 격자에 적용하여 제 2 차 탐색을 수행한다. 이때, 제 2 탐색부(104)는 제 1 탐색부(101)로부터 출력된 단어 격자의 음향 스코어를 고정하고, 가중치 추정부(103)로부터 수신된 제 1 언어 모델 가중치보다 낮은 제 2 언어 모델 가중치를 적용한 경우의 언어 모델 스코어를 계산하여 전체 패스의 누적 스코어를 업데이트하여 가장 높은 스코어를 가진 인식 결과를 최종 인식 결과로 출력한다.
- [0030] 셋째, 제 2 탐색부(104)는 가중치 추정부(103)가 제 1 음향 스코어와 제 2 음향 스코어를 이용하여 제 1 탐색부(101)의 연속어 음성 인식 결과인 제 1 음향 스코어가 음소 인식부(102)의 음소 인식 결과인 제 2 음향 스코어 보다 좋지도 않고 나쁘지도 않다고 판단함에 따라 제 1 탐색부(101)에서 사용된 제 1언어 모델 가중치와 제 1 언어 모델 보다 낮은 가중치인 제 2 언어 모델 가중치의 중간에 해당하는 가중치를 수신하고, 수신된 언어 모델 가중치를 적용하여 제 2 탐색을 수행한다. 그러면 이하에서는, 도 2를 참조하여 본 발명의 일 실시예에 따른 언어 모델 가중치 추정 방법을 보다 구체적으로 설명하기로 한다.
- [0031] 도 2는 본 발명의 일 실시예에 따른 언어 모델 가중치 추정 방법을 설명하기 위한 흐름도이다.
- [0032] 도 2를 참조하면, 언어 모델 가중치 장치는 음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 수신된 음성 특징 벡터에 제 1 언어 모델을 적용하여 제 1 차 탐색을 수행하여 단어 격자 및 단어 격자의 제 1 음향 스코어를 출력한다(S201). 여기서, 단어 격자는 제 1 차 탐색에서 탐색된 복수개의 단어 조합을 단어간 연결 정보로 정의한 것으로, 연속어 음성 인식 결과를 평가하는데 사용되는 정보, 예를 들어 제 1 차 탐색에서 실행된 제 1 차 탐색에서 가능성이 높다고 추정되는 단어들의 제 1 음향 스코어, 언어 모델 스코어, 단어의 시작점 및 끝점, 각 단어의 이전 패스에 따른 언어 모델 스코어 정보 등의 정보를 포함하고 있다.
- [0033] 언어 모델 가중치 장치는 음성 특징 벡터에 음향 모델을 적용하여 제 2 음향 스코어를 계산하여 출력한다(S202). 본 발명의 일 실시예에 따르면, 언어 모델 가중치 장치는 음성 신호로부터 변환된 음성 특징 벡터를 수신하고, 한국어의 자음+모음+자음으로 구성되는 음소 탐색 네트워크에 기반하여 음소 인식을 수행하여 음성 특징 벡터의 제 2 음향 스코어를 출력한다.
- [0034] 언어 모델 가중치 장치는 출력된 제 1 음향 스코어 및 제 2 음향 스코어를 이용하여 연속어 음성 인식 결과의 신뢰도를 측정한다(S203). 만약, 언어 모델 가중치 장치가 측정한 신뢰도가 임계값 1보다 작다고 판단한 경우(S204), 연속어 음성 인식 결과의 음향 스코어가 음소 인식 결과 음향 스코어 보다 상대적으로 좋은 것이며, 인식결과 신뢰도가 높다고 예측함에 따라 단어 격자에 제 1 언어 모델 가중치와 동일한 제 2 언어 모델 가중치를 적용하여 제 2 차 탐색을 수행하여 최종 인식 결과를 출력(S206)한다.
- [0035] 반면, 언어 모델 가중치 장치가 측정한 신뢰도가 임계값 1보다 크다고 판단한 경우(S204), 언어 모델 가중치

도면2

