

정보검색 기법과 동적 보간 계수를 이용한 N-gram 적응

최준기, 오영환

한국과학기술원 전자전산학과 전산학부 음성인터페이스연구소

N-gram Adaptation using Information Retrieval and Dynamic Interpolation Coefficient

Joon Ki Choi, Yung-Hwan Oh

Voice Interface Laboratory, CS Div, EECS Dept, KAIST

jkchoi@speech.kaist.ac.kr, yhoh@cs.kaist.ac.kr

Abstract

연속음성인식을 위한 언어모델 적응기법은 특정 영역의 정보만을 담고 있는 적응 코퍼스를 이용해 작성한 적응 언어모델과 기본 언어모델을 병합하는 방법이다. 본 논문에서는 추가되는 자료 없이 인식 시스템이 보유하고 있는 코퍼스만을 사용하여 적응 코퍼스를 구축하기 위해 언어모델에 기반한 정보검색 기법을 사용하였다. 검색된 적응 코퍼스로 작성된 적응 언어모델과 기본 언어모델과의 병합을 위해 본 논문에서는 입력음성을 분할하여 각 구간에 최적인 동적 보간 계수를 구하는 방법을 제안하였다. 제안된 적응 코퍼스를 구하는 방법과 동적 보간 계수는 기본 언어모델 대비 절대 3.6%의 한국어 방송뉴스 인식 성능 향상을 보여주었으며 기존의 검증자료를 이용한 정적 보간 계수에 비해 상대 13.6%의 한국어 방송뉴스 인식 성능 향상을 보여주었다.

I. 서론

HMM 음향모델과 Viterbi 탐색을 기반으로 하는 연속음성인식에서 N-gram 언어모델은 성공적으로 사용되고 있다. 그러나 N-gram은 자연언어의 복잡성을 표현하기에 단순한 구조를 가지고 있으며 훈련용 코퍼스에 대한 의존성을 가지고 있기 때문에 하나의 N-gram으로 여러 영역을 동시에 표현할 수 없다는 단점을 가지고 있다[1]. 이러한 단점을 극복하기 위한 여러 방법 중 N-gram 적응은 소규모의 적응코퍼스를 이용하여 특정 영역을 강조하여 표현하는 방법으로 언어모델과 연속음성인식의 성능을 동시에 향상시킬 수 있다[2].

N-gram 적응은 크게 나누어 적응 코퍼스의 획득과 기본 언어모델과의 병합이라는 두 개의 문제로 나누어질 수 있다. 먼저 적응 코퍼스의 획득 문제는 인식 대상 음성의 영역이 알고 있는 경우 해당 영역의 코퍼스를 수집하여 사용할 수 있다. 반면에 인식 대상 음성의 영역이 미리 알려지지 않은 경우 정보검색 기법을 이용하여 적응 코퍼스를 구하는 방법이 일반적이다[3]. 본 논문에서는 기존의 여러 정보검색 기법 중 언어모델에 기반한 정보검색 기법을 사용하여 적응 코퍼스를 수집하였다. 또한 정보검색 기법에서 주로 사용되는 유니그램 질의어를 확장한 바이그램과 트라이그램의 질의어가 수집된 적응 코퍼스의 성능을 향상시키는 것을 실험적으로 확인하였다.

그리고 본 논문에서는 적응 언어모델의 또 다른 문제인 기본 언어모델과의 병합 문제해결을 위해서 음성 인식의 중간 결과를 이용하여 언어모델간의 선형 보간 병합의 보간 계수를 동적으로 구하는 방법을 제안한다. 기존의 보간 계수를 구하는 방법으로는 평가자료와 유사한 검증자료(held-out data)를 이용하여 동적 언어모델 병합 가중치를 구하거나[5][6], 고정된 정적 가중치를 사용하는 방법이 널리 사용되었으나[7], 이러한 방법은 검증자료에 의존하는 단점이 있다. 본 논문에서는 언어모델 병합의 동적 보간 계수를 정확한 추정을 위해서 인식 대상 음성을 언어모델의 변화에 대한 민감도에 따라 구간을 나누고 해당 구간의 단어 후보열을 검증자료로 사용하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2절에서는 언어모델 적응에 적합한 적응 코퍼스를 구하는 방법에 대해 서술하며 3절에서는 본 논문에서 제안한 동적 보간 계수를 구하는 방법에 대해 설명한다. 그리고 한국어 방송

뉴스인식 실험 결과와 결론을 맺는다.

II. 정보검색 기법을 이용한 적응코퍼스의 획득

인식대상 음성의 영역이 잘 알려지지 않았을 경우 N-gram 적응을 위한 적응 코퍼스는 정보검색 기법을 이용하여 외부의 새로운 데이터를 검색하여 구축할 수 있다. 그러나 본 논문에서는 순수한 언어모델의 적응 효과를 관찰하기 위하여 기본 언어모델을 작성할 때 사용한 텍스트, 즉 시스템이 미리 보유하고 있는 코퍼스만을 이용하여 적응코퍼스를 구축하였다. 또한 언어 모델 적응과정 중에 새로운 어휘의 추가도 금지하였다. 아래의 그림 1에서는 추가 코퍼스 없이 기본 코퍼스만을 이용하여 언어모델 적응을 수행하는 과정을 보여주고 있다.

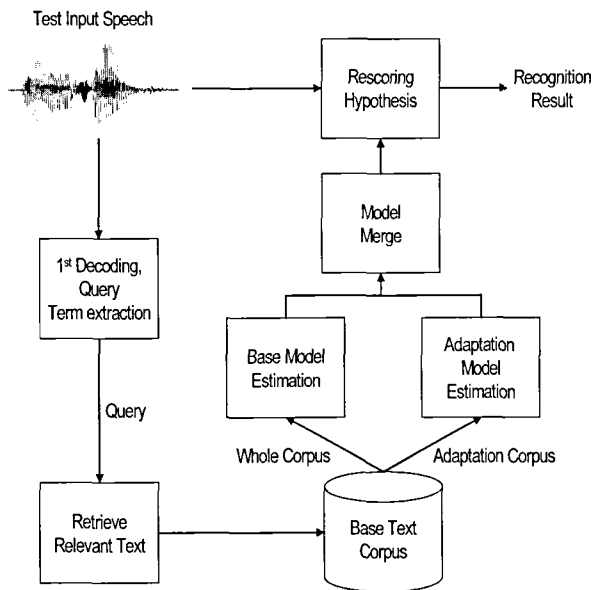


그림 1 기본 코퍼스와 정보검색을 이용한 언어모델적응

위의 그림 1에서처럼 음성인식의 중간결과로부터 추출된 정보는 정보검색의 질의어로 사용된다. 정보검색 기법을 사용하여 구축된 적응 코퍼스는 입력문장과 유사한 주제와 유사한 문형을 갖고면서도 자연언어의 특징을 잃지 않는다는 점에서 음성인식의 중간결과를 직접 적응코퍼스로 사용하는 방법에 비해 장점을 가진다.

그러나 음성인식의 중간결과를 질의어로 사용할 경우 음성인식의 오류가 대규모 코퍼스에 대한 질의어에 포함될 수 있다. 이러한 음성인식의 오류는 적응 언어 모델의 성능을 하락시키는 요인이 된다. 음성인식 오

류의 영향을 줄이기 위해서 본 논문에서는 음성인식의 중간결과에서 영역정보를 추출할 때 단어 수준의 음향학적 신뢰도를 사용하였다. 즉, 일정 수준 이상의 음향학적 신뢰도를 가지는 단어와 단어열만을 적응코퍼스를 구하기 위한 질의어로 사용하였다. 단어수준의 음향학적 신뢰도는 음향학적 사후확률을 사용하였다. 이 사후확률은 mangu가 제안한 confusion network의 링크 사후확률로부터 구해지며 식은 다음과 같다[8].

$$p(l|X) = \frac{\sum_{q_i} p(q|X)}{p(X)} \quad (1)$$

위의 식 1에서 X 는 음성의 관측벡터이며 $p(q|X)$ 는 경로 q 의 우도이다. 그리고 Q_l 은 confusion network의 링크 l 을 지난 모든 경로이다. confusion network를 예제와 적응코퍼스를 구하기 위한 질의어를 추출하는 과정은 아래의 그림 2에 도식화되어 있다.

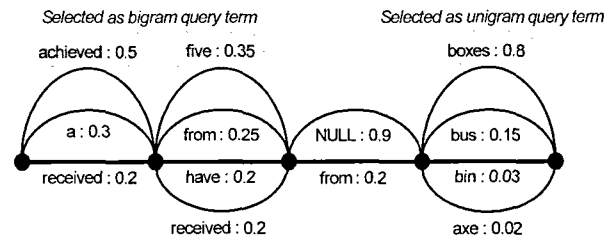


그림 2 음향학적 신뢰도를 고려한 음성인식의 중간결과 정보 추출

음향학적 신뢰도에 대한 임계값은 실험적으로 결정할 수 있다. Confusion network의 사후확률은 0부터 1까지 정규화되기 때문에 임계값을 수월하게 결정할 수 있다.

본 논문에서 사용한 정보검색 기법은 언어모델에 기반한 정보검색 기법을 사용한다[4]. 이 정보검색 기법에서는 질의어를 검색대상 문서의 언어모델에서 생성된 결과로 간주한다. 검색 결과 문서는 질의어의 언어모델과 유사한 언어모델을 가지는 문서 순으로 정렬되어 출력된다. 따라서 언어모델에 기반한 정보검색 기법은 질의어의 N-gram 분포와 유사한 문서를 검색할 수 있기 때문에 다른 정보검색 기법보다 N-gram 적응에 적합하다.

검색된 문서와 질의어의 거리는 문서 d 의 언어모델 M_d 과 질의어 q 의 언어모델 M_q 의 Kullback-Leibler 거리로 표현된다. 식은 다음과 같다.

$$Div(M_q \| M_d) = \sum_w p(w | M_q) \log \frac{p(w | M_q)}{p(w | M_d)} \quad (2)$$

식 2에서 $p(w | M_q)$ 와 $p(w | M_d)$ 는 문서와 질의어의 언어모델 값을 나타낸다. 언어모델을 구하기 위한 평탄화(smoothing) 방법으로는 Katz 평탄화 방법을 사용하였다. 여러 가지 언어모델 평탄화 기법 중 Katz 평탄화의 성능이 가장 좋았다.

또한 본 논문에서는 유니그램 질의어만을 사용하지 않고 바이그램과 트라이그램으로 확장하여 사용하였다. 질의어를 확장했을 때 질의어와 검색 대상 문서의 순서 결정 함수(ranking function)은 다음 식 3과 식 4처럼 유도될 수 있다.

$$\sum_{w_i, w_{i-1} \in d \sim q} p(w_i | M_q, w_{i-1}) \log \frac{p_s(w_i | M_d, w_{i-1})}{\alpha_d p(w_i | M_c, w_{i-1})} + \log \alpha_d \quad (3)$$

$$\sum_{w_i, w_{i-1}, w_{i-2} \in d \sim q} p(w_i | M_q, w_{i-1}, w_{i-2}) \log \frac{p_s(w_i | M_d, w_{i-1}, w_{i-2})}{\alpha_d p(w_i | M_c, w_{i-1}, w_{i-2})} + \log \alpha_d \quad (4)$$

식 3과 식 4의 α_d 는 미등록어의 언어모델 값을 부여하기 위해 문서 d 에 의해 결정되는 상수이다.

III. 선형 보간 병합의 동적 보간 계수

특정 영역에 대하여 최적화된 적응 언어모델은 위에서 서술한 방법으로 구해진 적응 코퍼스를 이용하여 구축된다. 그러나 적응 코퍼스만을 이용해서 구한 적응언어모델은 적은 학습 자료로 인한 데이터 희귀성 문제를 가지기 때문에 기본 언어모델과 병합하여 사용하는 것이 일반적이다[2]. 본 논문에서는 구현이 용이하며 효율적인 선형보간법을 이용하여 두 개의 언어모델을 병합하였다. 선형보간법의 식은 다음과 같다.

$$\hat{p}(w | h) = \lambda p_a(w | h) + (1 - \lambda) p_b(w | h) \quad (5)$$

식 5에서 P_a 는 적응 언어모델이며, P_b 는 기본 언어모델이고, λ 는 두 언어모델에 대한 보간 계수이다.

기존의 추가적인 검증자료를 이용하는 방법과는 달리 본 논문에서는 입력음성의 인식 중간결과를 검증자료로 사용하여 동적 보간 계수를 구하였다. 그러나 연속음성인식의 중간결과는 기본 언어모델과 음향모델의 변별력 부족으로 인하여 후보 단어열이 많이 생성되는 구간과 그렇지 않은 구간으로 나눌 수 있다. 후보 단

어열이 많이 생성되는 구간은 음성인식의 오류를 예상할 수 있다. 또한 자연언어의 특성상 문장에서 영역정보를 포함하고 있는 부분과 영역정보를 포함하지 않은 구간으로 나눌 수 있다. 영역정보에 민감한 부분에 대해서는 적응 언어모델의 가중치를 높여야 하지만, 그렇지 않은 구간에 대해서는 기본 언어모델의 가중치를 높이는 것이 바람직하다. 따라서 본 논문에서는 음성인식의 중간결과를 분할하여 각 구간에 적합한 보간 계수를 구하고, 해당 구간은 같은 보간 계수를 적용하여 기본 언어모델과 적응 언어모델의 가중치를 결정하고자 한다.

본 논문에서 사용한 입력음성에 대한 분할 기준은 언어모델의 변화에 따른 민감도를 사용하였다. 기본 언어모델과 적응 언어모델이 주어졌을 때 이 두 언어모델의 변화에 의해 인식 후보 단어의 순위가 변하는 구간과 변하지 않는 구간으로 분할한다. 본 논문에서 사용한 분할 기준은 음성인식의 후처리에서 널리 사용되는 신뢰도 중 하나인 음향학적 안정도(Acoustic Stability)와 유사한 의미를 가진다[9]. 입력음성을 분할하는 과정은 다음과 같다.

1. 음성인식 중간결과 재정렬

1.1 기본 언어모델로 음성인식 중간결과를 재정렬하여 N-best 리스트를 얻는다. ($N < 10$)

1.2 적응 언어모델로 음성인식 중간결과를 재정렬하여 N-best 리스트를 얻는다. ($N < 10$)

2. Levenshtein 정렬 및 구간분할

2.1 두 개의 N-best 리스트에서 각각의 순위에 맞는 쌍의 문장후보끼리 정렬한 뒤 서로 다른 부분을 찾는다.

2.2 모든 문장 후보 쌍에 대해서 항상 같은 부분을 고르고 남은 부분끼리 모아서 음성구간을 분할한다.

분할된 구간에 속하는 단어 후보열을 이용하여 언어모델 보간 계수를 결정하면 해당 구간의 영역 특성을 반영할 수 있다. 또한 Kalai[10]의 동적 보간 계수와는 달리 해당 구간의 한 언어모델의 신뢰도에 최적화된 동적 보간 계수를 제공할 수 있기 때문에 음성인식의 한 구간에서는 경쟁 관계에 있는 후보 단어열들에 대해서 단일한 적응 언어모델이 적용되는 효과를 가진다.

단어 후보열을 이용하여 동적 보간 계수를 구하는 위해 EM 알고리즘에 기반한 방법을 사용하였다[11]. 두 개의 언어모델을 HMM의 상태라고 가정하고 학습자료로 사용되는 후보 단어열을 관측벡터라고 간주하면 보간 계수는 초기 상태에서의 각 두 상태로의 천이 확률로 볼 수 있다. 따라서 보간 계수 λ 에 대한 EM

알고리즘의 갱신 식은 HMM의 천이확률 학습식과 유사하게 구할 수 있으며 다음과 같이 표현된다.

$$\lambda_{n+1}(i) = \frac{1}{M} \sum_{m=0}^{M-1} \frac{\lambda_n(i) p_j(w_{n-m} | h_{n-m})}{\sum_{j=1}^I \lambda_n(j) p_j(w_{n-m} | h_{n-m})} \quad (6)$$

IV. 실험결과

1. 한국어 방송뉴스 인식기

본 논문에서는 한국어 방송뉴스 인식기를 사용하여 제안한 적응 언어모델의 성능을 평가하였다. 트라이폰 기반의 연속 확률분포 HMM 음향모델을 학습시키기 위하여 300 시간 분량의 방송뉴스 데이터를 사용하였다. 방송뉴스는 KBS, MBC, SBS의 3 개 방송국의 저녁 뉴스 데이터를 사용하였으며 여러 가지 잡음환경이 모두 포함되어 있다. 음성처리를 위하여 13차 MFCC와 이에 해당하는 delta, delta-delta의 총 39차 특징벡터를 추출하였으며 잡음처리에 강건한 인식을 위하여 전역 CMS(Cepstrum Mean Subtraction)을 수행하였으며 VTLN(Vocal Tract Length Normalization)이 수행되었다. 훈련된 HMM은 총 3786개의 상태들로 구성되었으며 각 상태는 16개의 Gaussian mixture로 표현된다. 교차어로서의 한국어는 어근과 어미로 나누어져 많은 활용형이 가능하기 때문에 영어와 같이 단어 단위의 사전 구축이 어렵다. 따라서 본 논문에서는 의사형태소 단위의 사전과 인식단위를 사용한다. 따라서 모든 인식단위에는 품사태그가 부착되어 있다. 언어모델의 학습을 위하여 총 223M 의사형태소 분량의 신문 기사(조선일보, 동아일보)와 51M 의사형태소 분량의 방송 대본을 사용하였다. Katz 평탄화를 이용한 기본 4-gram은 SRI 언어모델 툴킷을 이용하여 구하였다 [12]. 탐색 알고리즘은 tree-trelli 알고리즘을 사용하였으며 후향 A* 알고리즘을 사용하여 래티스에 대한 재정렬을 시도하였다.

평가 자료는 학습자료와 충분한 시간 간격을 가지도록 선정되어 학습자료에 평가자료와 동일한 주제나 기사가 실리지 않도록 하였다. 학습자료로 사용된 언어모델은 1998년 1월부터 2003년 1월까지 수집된 자료이며 평가자료는 2003년 10월에 수집되었다. 평가자료는 SBS와 KBS의 두 개의 방송국에서 수집된 4일 분량의 방송 뉴스이다. 평가자료는 간단한 휴리스틱 규칙과 GMM 분류기를 이용하여 자동으로 총 39개의 꼭지기사와 459개의 문장으로 분할되었으며 각 꼭지기사는 동일한 주제를 다루고 있다. 각 꼭지기사에 대해 음성인식을 수행한 중간결과로부터 적응 코퍼스를 구하기

위한 질의를 추출하였으며 평균 305.7개의 단어가 질의어로 추출되었다. 적응 언어모델은 꼭지 기사단위로 작성되어 하나의 꼭지기사가 동일한 적응 언어모델로 재정렬 되었다.

2. 적응 코퍼스의 수집 실험

먼저 본 논문에서 제안된 적응 코퍼스 수집 방법의 유효성을 검증하기 수집된 적응 코퍼스에 대한 perplexity 실험을 수행하였다 정보 검색 기법을 사용하여 검색된 문서와 질의어의 유사도가 일정 이상인 문서들을 모아서 적응 코퍼스를 작성하게 되는데 이 유사도와 적응 언어모델의 성능간의 관계를 보기 위하여 유사도를 변형하면서 적응 언어모델의 perplexity 성능을 측정하였다. 결과는 아래의 그림과 같다.

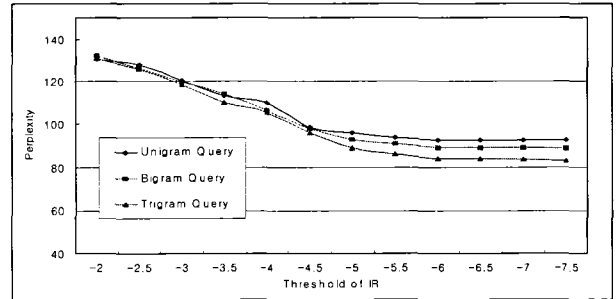


그림 3. 검색된 문서의 유사도 임계치의 변화와 적응 언어모델의 성능변화

그림 3에서 볼 수 있듯이 유사도 -6.0 근처에서 적응 언어모델의 성능이 수렴하는 것을 볼 수 있다. 또한 본 논문에서 제안한 질의어의 N-gram 확장 실험결과도 같이 표시되어있는데, 유니그램에 비해서 바이그램이나 트라이그램의 질의어가 적응 언어모델의 성능을 향상시키는 것을 알 수 있다.

다음에는 다른 정보 검색 기법과 언어모델에 기반한 정보 검색기법의 성능을 비교하는 실험을 수행하였다. 본 논문에서 사용한 방법과 벡터 모델링에 기반하여 정보검색 분야에서 널리 사용되고 있는 BM25 가중 방법[13]을 비교한 결과는 표 1과 같다.

		Perplexity	WER
기본 4-gram		136.23	16.84
적응 코퍼스 수집 방법	BM25 가중 방법	89.04	15.62
	유니그램 질의어	92.40	15.74
	바이그램 질의어	88.94	15.49
	트라이그램 질의어	84.18	15.33

표 1 기본 언어모델과 적응 코퍼스 수집 방법에 따른 적응 언어모델의 음성인식 성능 실험 결과

BM25 가중 방법의 질의어와 문서의 유사도도 역시

실험적으로 구하여 최적의 적응 코퍼스를 구하였다. 본 실험을 위해서 기본 언어모델과의 보간 계수는 검증 자료를 이용한 최적의 정적 보간 계수(0.8)를 구하여 사용하였다. 표 1에서 볼 수 있듯이 음성인식 실험에서도 N-gram 확장 질의어를 사용한 언어모델에 기반한 정보 검색 기법이 BM25 방법보다 좋은 음성인식 성능을 보여주었다.

3. 동적 보간 계수를 이용한 적응 언어모델 실험

다음으로 본 논문에서 제안된 동적 언어모델 보간 계수를 이용한 적응 언어모델의 성능 평가를 수행하였다. 먼저 언어모델의 변화에 따라 입력음성의 구간을 분할하는 방법의 유효성을 검증하기 위하여 언어모델의 변화에 민감하지 않은 구간으로 검출된 구간의 변화를 관측하였다. 그 결과 언어모델 변화에 민감하지 않은 구간은 적응 언어모델의 보간 계수를 0부터 1까지 변화시켜도 후보 단어열의 순위가 변하지 않음을 확인할 수 있었다.

그리고 본 논문에서 제안한 동적 보간 계수를 검증하기 위해 여러 가지 다른 방법으로 보간 계수를 구하고 음성인식 실험을 통해 성능을 비교하였다. 먼저 검증 자료를 이용해서 최적의 정적 보간 계수를 구하였으며(held-out static)[7], 검증 자료를 이용해서 해당 단어의 과거의 함수로 결정되는 동적 보간 계수를 구하였다(held-out history dependent)[5][6], 그리고 음성의 구간을 분할하지 않고 음성인식의 전체 후보열에 대해서 온라인 알고리즘을 통해서 구한 동적 보간 계수와 비교하였다(on-line history dependent)[10]. 실험 결과는 표 2와 같다.

보간 계수 유형	WER
held-out static	15.33
held-out history dependent	15.50
on-line history dependent	14.02
제안된 음성 구간 분할 방법	13.24

표 2 언어모델 병합의 동적 보간 계수 비교 실험

표 2에서 볼 수 있듯이 검증자료를 사용하여 구한 보간 계수에 비해서 음성인식의 중간결과를 사용하는 방법이 향상된 성능을 보여주었다. 그리고 on-line history dependent 방법에 비해서 제안된 분할된 음성 구간에 기반한 동적 보간 계수가 성능이 더 좋았으며, 이는 입력음성의 동일한 구간에 대해서는 모든 단어 후보열에 대해서 최적의 보간 계수를 구하는 방법이 각 단어 후보열에 최적인 방법보다 성능이 좋음을 확인할 수 있었다. 본 논문에서 제안한 정보검색 기법과 동적 보간 계수를 이용하여 기본 언어모델의 인식 성

능에 비해 단어 인식을 성능은 절대 3.6% 향상되었으며 기존의 정적 보간 계수를 사용하는 방법에 비해 상대적으로 13.6% 향상되었다.

V. 결론

본 논문에서는 추가적인 텍스트 자료가 없을 경우 정보검색 기법을 사용하여 기존 텍스트 코퍼스로부터 N-gram 적응에 효과적인 적응 코퍼스를 획득하는 방법을 제안하였다. 또한 질의어의 확장을 통해서 N-gram 적응을 위해서는 일반적으로 사용자가 만족하는 정보검색의 답이 아닌 N-gram 분포가 유사한 문서를 검색하는 것이 보다 효율적이라는 결론을 얻을 수 있었다.

그리고 본 논문에서는 간단한 선형 보간 언어모델 병합에서 사용될 수 있는 동적 보간 계수를 제안하였다. 본 논문에서는 입력 음성의 중간결과에서 언어모델의 변화에 민감한 구간과 그렇지 않은 구간으로 나누어 동적 보간 계수를 구간 별로 구하여 음성인식 성능의 향상을 확인하였다. 제안한 방법은 음성인식의 중간 결과에 대하여 두 개의 언어모델로 각각 신뢰도를 검사한 뒤에 특정 구간에 특정 언어모델이 우세하게 나타날 경우 해당 언어모델을 가중해 줄 수 있었다. 그리고 제안한 방법인 해당 구간에 최적의 보간 계수를 찾는 방법이 인식 후보 단어열 각각에 대해 최적의 보간 계수를 구하는 경우에 비해 성능이 좋음을 실험적으로 증명하였다.

본 논문에서는 N-gram의 적응을 위하여 동일한 N-gram을 소규모의 적응 코퍼스에서 구하여 사용하였으나 향후 청크 결합 정보나 구문 분석 결과와 같은 상위 자연언어 지식을 기본 언어모델의 성능을 저하시키지 않으면서 적응 언어모델로 사용하는 연구로 확장될 수 있다. 또한 본 논문에서는 언어모델에 대한 민감도만을 신뢰도로 사용하였으나 음향모델이나 전처리에서 제공할 수 있는 다양한 음향학적 정보를 모두 고려하여 연속음성인식의 후처리로 사용할 수 있다.

참고문헌

- [1] Rosenfeld, R. "Two Decades of Statistical Language Modeling: Where do we go from here", Proc. IEEE 88(8), 2000
- [2] Bellegarda. J. R., "An Overview of Statistical Language Model Adaptation", In Proc. ISCA Adaptation Methods in Automatic Speech

Recognition, 2001

- [3] B. Bigi, Y. Huang, R. De Mori, "Vocabulary and Language Model Adaptation using Information Retrieval", ICSLP 2004
- [4] Zhai, C. Lafferty, J., "A Study of Smoothing Methods for Language Models Applied to Adhoc Information Retrieval", In Proc. ACM SIGIR'01, 2001
- [5] Kneser. R., Steinbiss. V., "On the Dynamic Adaptation of Stochastic Language Models". Vol 2, pp 586-589, ICASSP 1993
- [6] Weintraub. M., et al, "LM95 Project Report: Fast training and portability", Technical Report 1", Center for Language and Speech Processing, Johns Hopkins University
- [7] Bulyko. I., Ostendorf. M., Stolke. A., "Getting More Mileage from Web Text Sources for Conversational Speech Language Modeling using Class Dependent Mixtures", HLT 2003
- [8] Mangu, L., "Finding Consensus in Speech Recognition", Ph. D Thesis, Johns Hopkins Univ. 2000
- [9] Wessel, et al, "Confidence Measures for Large Vocabulary Continuous Speech Recognition", IEEE TRANS. on Speech and Audio Processing, Vol. 9, No. 3, 2001
- [10] Kalai. A., Chen. S., Blum. A., Rosenfeld. R., "On-Line Algorithms for Combining Language Models", ICASSP 1999
- [11] Jelinek. F., "Self-Organized Language Modeling for Speech Recognition", in Readings in Speech Recognition", Alex Waibel and Kai-Fu Lee. Morgan Kaufmann, 1989
- [12] Stolke. A., "SRILM - An Extensible Language Modeling Toolkit", in Proc, ICSLP, Denver, 2002
- [13] Robertson. S. E., Walker. S., Beaulieu. M.M., Gatford., M., Payne. A., "Okapi at TREC-4", NIST TREC-4, 1995