

문형정보와 N-gram 단어정보를 이용한 연속음성인식 후처리

엄한용^U

황도삼

영남대학교 컴퓨터공학과

hyeom@nlp.yeungnam.ac.kr dshwang@ynucc.yeungnam.ac.kr

A post processing of continuous speech recognition using N-gram words and sentence patterns

Han-yong Eom^U

Do-sam Hwang

Dept. of Computer Engineering, Yeungnam University
and

Advanced Information Technology Research Center(AITrc)

요 약

본 논문에서는 항공편 예약이라는 제한 영역에서의 연속음성인식 시스템을 위한 후처리 방안을 제시한다. 제안하는 후처리 방안은 200 문장의 항공편 예약 텍스트 데이터를 이용하여 문형 정보를 추출한 뒤 특정 문형별로 분류하였다. 분류된 문형과 음성인식 후의 문장을 비교하여 가장 유사한 문형을 추론한다. 추론한 특정 문형에서 나올 수 있는 형태소를 형태소들간의 N-gram 정보가 수록된 데이터베이스를 이용하여 형태소를 수정하고 보완한 결과를 최종 문장으로 출력한다.

1. 서론

음성인식의 최종 목표는 불특정 화자에 대한 대어휘 연속음성인식이라 할 수 있으나, 최근까지도 불특정 화자에 대한 대어휘 연속 음성의 인식률은 매우 저조한 편이다. 그러나 음성인식 적용 대상을 특정 영역으로 제한시켜 음성인식을 행하면, 만족할 만한 인식률을 얻을 수 있다[1].

본 논문에서는 대상 영역을 항공편 예약으로 제한하고 이에 대한 후처리 방안을 제시한다. 연속음성인식 시스템에서 인식 성능에 가장 큰 영향을 미치는 요소들로는 문법, 인식단위의 모델링, 그리고 탐색법 등이 있다. 특히, 문법은 여러 개의 단어들로 구성되는 문장에서 단어와 단어간의 연결 규칙을 이용하는 것이다. 인간이 연속음성을 인식하는 과정에서도, 앞서 발음된 단어들을 사용하여 현재 인식하려는 불분명한 단어를 유추하여 인식한다. 연속음성인식 시스템에서 문법은 인식성능을 향상시키기 위해 가장 필수적인 요소이다. 그러나, 인식기에서 사용하는 문법만으로 언어의 올바른 해석을 하기에는 많은 문제점이 있다. 그래서 인식기에서는 기존 문법(CFG, N-gram)을 이용한 인식만을 한 후에, 자연 언어처리 기법을 후처리로 사용한다면 보다 나은 인식 결과를 얻을 수 있을 것이다[2, 3]. 본 논문에서는 문형

분류와 함께 특정 문형에 대한 N-gram 정보를 수록한 데이터베이스를 음성 인식기에 맞게 수정한 후처리 방안을 제시한다.

문형 분류는 후처리 단계 이전의 인식 후에 나오는 결과의 문장들과 인식기에서 사용될 수 있는 문장들에서, 문장을 가장 잘 표현할 수 있는 형태소들을 뽑아낸 다음, 이를 특정 문형으로 표현한다. 이러한 형태소를 여기에서는 '특정 형태소'라 명명하였다. 이를 N-gram 단어 정보를 사용하여 오인식된 형태소를 수정, 보완하여 문장을 재구성하게 된다. N-gram 단어 정보는 분류된 문형들에서 특정 문형에 나오는 단어들의 N-gram 정보를 담고 있는 데이터베이스이다.

2. 문형분류

2.1 대상문의 분류

하나의 문장이 일정 형태로 발화되기 위한 요인은 많다. 누가 누구에게 이야기하고 있는가, 부탁하고 있는가, 명령을 하고 있는가 등의 요인 중 하나가 선택된다. 본 논문에서는 문형을 분류하되, 문장에서 나올 수 있는 단어를 쉽게 예측할 수 있도록 인식 결과의 문장에서 조사, 접속사, 연체사와 같은 특정 형태소를 추출함으로써

기본 문형을 만들어 낸다.

항공예약 시스템의 테스트 문장들에서 추출되는 단어 들에 대해 형태소별로 좀더 자세히 분류하면, 조사(관형 격 조사, 호격 조사, 부사격 조사, 보조사, 접속 조사), 연결어미, 종결어미이며 이를 추출하여 문형 구분에 사용 한다. 특히, 문형 분류에 결정적인 역할을 할 수 있는 의 미를 가진 형태소들을 부분적으로 첨가하였다. 사용한 형태소들을 <표 2>와 같이 구분하였다.

문형 구분을 위해, 먼저 항공편 예약 시스템에서 사용 되는 문장들에 대해 형태소분석을 한 다음, 추출된 형태 소들을 기반으로 문장 트라이 구조를 만든다.

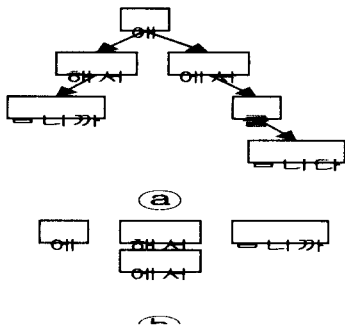
<표 2> 문형 분류에 사용되는 형태소

형태소	예
조사	에, 에서, 보다, 로, 으 로, 에는, ...
연결 어미	면, 고, 습니다만, ...
종결 어미	입니다, 어요, 습니다, 습니까, ...

2.2 트라이 구조

트라이 구조는 단어가 확정되어 있지 않은 경우의 사 전 검색에 적합한 알고리즘이지만, 여기에서는 문형 탐 색이 가능하도록 그 구조를 약간 변경하였다.

(그림 1) 에서 (a)는 트라이 구조를 (b)는 문형 탐색에 의한 그 결과를 내놓은 경우이다.



<그림 1> 트라이 구조와 탐색

<그림 1>의 ① 에서 '에' 다음에 나올 수 있는 특징 형태소로서 '해서', '에서' 가 있으나, 이 다음 나오는 형태소가 '입니다' 이라면, '에서' 다음 탐색에서는 '입니다' 가 나올 수가 없으므로 여기에서는 '해서'를 선택하게 된 다.

일반적으로, 길이 M, 후보숫자 N의 음소 격자 L[M][N]인 트라이 구조에서의 검색시간은 다음과 같다.

$$M + M * (L[0..1] \cap \text{TRIE}) + M * (L[0..2] \cap \text{TRIE}) + \dots + M * (L[0..N] \cap \text{TRIE})$$

이러한 트라이 구조의 생성은 먼저 항공편 예약 시스 템에서 사용되는 문장들을 미리 형태소분석을 하고, 문 형분류에 사용되는 특징 형태소들을 추출하여 트라이 구 조를 생성한다. 먼저, KTS(Korean Tagging System)[5] 형태소해석기를 이용하여 항공편 예약 시스템에서 사용 하는 200문장에 대해 다음과 같은 1차 결과를 얻은 다 음, 이 문장들의 문형을 분류하였다.

(KTS를 이용한 형태소 해석 결과)

1 -> 12월 3일 나고야에 가려고 합니다.

- o 12/nnn+월/mbu
- o 3/nnn+일/mbu
- x 나고야/nc+에/jca
- o 가/pv+려고/ccx
- o 하/px+됩니다/ef+/s.

(문형 분석 결과)

에 0 려고 1 됩니다 1

문형 분석 결과에서 단어들은 문형에 나올 수 있는 특 징 형태소들이고, 특징 형태소 뒤의 숫자는 문형의 종류 를 나타낸다. 특히 숫자 0은 특징 형태소가 부족하여 문 형 정보를 구성할 수 없는 경우를 나타낸다. 날짜, 이름, 고유명사 같은 경우는 트라이 구조를 생성할 때, 한 단 어로 취급하여 문형 구조에 반영시켰다. 문형구분에 결 정적인 역할을 할 수 있는 형태소(예: 편)들은 임의로 추 가하였다.

3. N-gram 단어 사전

3.1 N-gram

항공편 예약 시스템에서 사용되는 문장들에 대한 형태 소해석 결과에서 같은 문형을 갖는 문장들을 추출하고, 여기서 각 문형에서 특징 형태소들을 제외한 형태소들에 대해 N-gram 정보를 조사한다. 특징 형태소는 문형 분류 에 사용된 형태소들이며, N-gram 단어 정보는 형태소 해 석 결과에서 특징 형태소를 제외한 형태소를 바탕으로 만들어진다.

3.2 N-gram 단어 정보의 생성

형태소 해석 결과를 토대로 다음과 같은 구조를 갖는 텍스트를 생성시켰다. 텍스트를 생성할 때 날짜, 시간, 이름은 한 단어로 취급하였다.

(N-gram 단어 정보)

1. (nct+nq) 에 (pv) 려고 (px) 됩니다.

확률 -> 1

nct, 날짜(1)

nq, 고유명사(1)

pv, 가(1)

px, 하(1)

...

먼저 음소표기를 비교하여 인식된 음소와 발음 사전의 음소를 비교하여 완전히 같을 경우 1로 하고, 다를 경우

는 음소의 수와 발음 사전의 음소 중 일치하는 음소의 수를 비교하여 확률을 계산한다. 발음 사전의 구성은 다음과 같다.

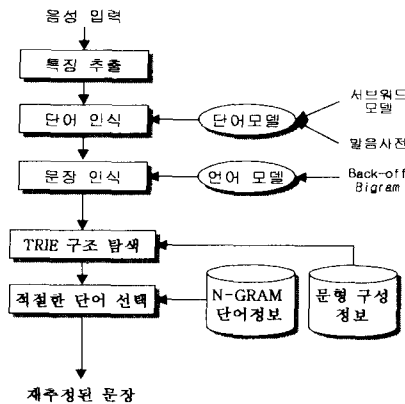
을 # u l
 대구 # d ae g~ uh
 오오사카 # ao ao s aa k aa
 예서 # eh s axr
 ...

4. 전체적인 시스템 구성과 구현

4.1 전체적인 시스템 구성

여기에서 사용되는 베이스 라인 음성인식 시스템으로는 영남대 정보통신공학과에서 개발한 한국-일본간 항공편 예약 시스템을 기반으로 한다.

<그림 2>와 같이 구성된 인식 시스템에서 먼저 문형 구성 정보를 바탕으로 트라이 구조를 만든 다음 이를 이용하여 1차 인식문에 대한 문형 구조를 탐색하게 된다.



<그림 2> 시스템 구성도

4.2 후처리에서의 동작

1차 인식 문장의 형태소분석 결과로 특징 형태소 '에', '려고', '비니다'를 추출하고 이 특징 형태소를 입력하여 트라이 구조를 검색하면 1번 문형을 출력으로 내놓게 된다. 출력 문장으로 "에 0 려고 1 비니다 1"를 추출하면, N-gram 단어정보에서 1번 문형에 수록된 단어 정보를 인식문의 수정에 사용한다는 것을 의미한다.

적절한 단어 선택은 N-gram 단어정보를 이용하여 추출된 단어들과 1차 인식되어 나온 문장에서 특징 형태소를 제외한 단어들과의 비교를 통해 가장 확률이 높은 단어를 선택하고, 그 단어가 포함된 문장을 최종 인식문으로 결정하게 된다.

예를 들어 트라이 구조 탐색결과가 "에 0 려고 1 비니다 1"이라면 N-gram 단어정보에서 1번 문형에 사용되는 단어들을 추출한다. 이를 다시 1차 인식문에서 특징 형태소를 제외한 단어들과 비교한다. 각 단어들에 대하여 그 단어가 문장 속에서 사용되는 확률을 구하고 가장 높은 확률을 가지는 문장을 선택하게 된다.

5. 실험 및 결과

본 논문에서 제시한 방법이 음성인식의 후처리로서 유효한가를 확인하기 위하여 베이스 라인 음성인식기로 항공편 예약 시스템을 선택하고, 이에 대한 후처리로 본 논문에서 제안한 방법을 적용시켜 실험을 수행하였다.

인식실험에 사용한 연속음성 및 텍스트 데이터는 13인의 사용자가 가상의 조건에서 항공편 예약에 관련된 질문과 대답에 관한 내용으로 구성되어 있었다. 13인의 음성 데이터 중 평가용 데이터는 3인의 200 문장으로 설정하고, 나머지 10 명의 음성 데이터 중 발성 상태가 양호한 8 명의 연속음성을 단어와 함께 음향 모델의 학습에 사용하였다.

이 시스템에서 후처리를 사용하지 않은 경우, 평가용 3인의 200 문장에 대한 인식 결과는 91.6%의 문장 인식을 얻었다.

본 논문에서 제안한 방법을 후처리로 적용한 결과, 오인식 문장에 대해서는 42%를 올바른 문장으로 재구성하였고, 이는 전체 문장의 인식률에 대해 2.5%의 인식률 향상을 보였다. 이상의 결과로부터 음성인식의 후처리에서 본 논문에서 제안한 방법이 유효함을 확인할 수 있었다.

6. 결론

본 연구에서는 연속음성인식 시스템의 인식률 향상을 위해 후처리로 자연언어처리를 이용하여 문장을 재구성하였고, 이를 이용하여 성능평가 실험에서 2.5%의 인식률 향상을 얻어 제안한 방법의 유효성을 확인하였다. 그러나, 아직 후처리를 하기 전의 인식 문장이 제대로 형태소 해석이 되지 않는 경우가 있었고, 이로 인해 문형 분류가 잘못되는 경우가 있었다. 향후, 음성인식기에 적합한 형태소 해석기[6]의 개발과 함께, 후처리 부분에 의미해석까지 포함된 음성인식기가 개발된다면 음성인식의 성공률을 더욱 높일 수 있다고 생각한다.

참고 문헌

[1] 오세진 외 2 인, "연속음성인식 시스템의 성능개선", 한국음향학회 하계 학술발표대회 논문집, 1997.11.
 [2] 정현열, "항공편 예약을 위한 연속음성 인식·이해 시스템", 제13회 음성통신 및 신호처리 워크샵 논문집, 1996.8.
 [3] 황철준 외 4 인, "음성인식 시스템의 성능 향상을 위한 반복학습법을 이용한 언어 모델", 한국음향학회 학술 발표 대회 논문집, 제 18 권 1(s)호, 1999.7.
 [4] 황도삼, 최기선, 김태석, "자연언어처리", 홍릉과학출판사, 1998.2.
 [5] 이상호, 서정연, 오영환, "KTS : 미등록어를 고려한 한국어 품사 태깅 시스템", 제 12 회 음성통신 및 신호처리 워크샵 논문집, 1995.6.
 [6] 이경남, 정민화, "의사 형태소 단위의 음성언어 형태소 해석", 제 10 회 한글 및 한국어 정보처리 학술대회, 1998.10.