

음절 복원 알고리즘을 이용한 핵심어 오류 보정 시스템

안 찬 식*, 오 상 엽**

Key-word Error Correction System using Syllable Restoration Algorithm

Chan-Shik Ahn*, Sang-Yeob Oh**

요 약

어휘 인식 시스템의 오류 보정 방법으로는 오류 패턴 매칭 기반 방법과 어휘 의미 패턴 기반 방법이 있으며, 이들 방법에서는 오류 보정을 위해 핵심어를 의미적으로 분석하지 못하는 문제점을 가지고 있다. 이를 개선하기 위해 본 논문에서는 음절 복원 알고리즘을 이용한 핵심어 오류 보정 시스템을 제안한다. 인식된 음소 열을 의미 분석 과정을 거쳐 음소가 갖는 의미를 파악하고 음절 복원 알고리즘을 통해 음운 변동이 적용되기 이전의 문자열로 복원하므로 핵심어를 명확히 분석하고 오인식을 줄일 수 있다. 시스템 분석을 위해 음소 유사율과 신뢰도를 이용하여 오류 보정율을 구하였으며, 어휘 인식 과정에서 오류로 판명된 어휘에 대하여 오류 보정을 수행하였다. 여러 패턴 학습을 이용한 방법과 오류 패턴 매칭 기반 방법, 어휘 의미 패턴 기반 방법의 성능 평가 결과 3.0%의 인식 향상율을 보였다.

Abstract

There are two method of error correction in vocabulary recognition system one error pattern matting base on method other vocabulary mean pattern base on method. They are a failure while semantic of key-word problem for error correction. In improving, in this paper is propose system of key-word error correction using algorithm of syllable restoration. System of key-word error correction by processing of semantic parse through recognized phoneme meaning. It's performed restore by algorithm of syllable restoration phoneme apply fluctuation before word. It's definitely parse of key-word and reduced of unrecognized. Find out error correction rate using phoneme likelihood and confidence for system parse. When vocabulary recognition perform error correction for error proved vocabulary. system performance comparison as a result of recognition improve represent 2.3% by method using error pattern learning and error pattern matting, vocabulary mean pattern base on method.

▶ Keyword : 오류 보정(error correction), 핵심어(Key-word), 음절 복원(Syllable Restoration), 음소 유사율(phoneme likelihood)

• 제1저자 : 안찬식 교신저자 : 오상엽
• 투고일 : 2010. 05. 17, 심사일 : 2010. 08. 05, 게재확정일 : 2010. 08. 17.
* 광운대학교 컴퓨터공학과 박사과정 ** 경원대학교 IT대학 컴퓨터소프트웨어
※ 본 연구는 2010년도 경원대학교 지원에 의한 결과임.

I. 서론

모바일 기기의 발전과 사용자의 증가로 인하여 어휘 기반 검색 시스템, 자동 응답 시스템 등 어휘 인식을 인터페이스로 하는 시스템들이 개발되고 있다. 어휘를 입력 인터페이스로 이용하는 모바일 단말기가 등장하면서 어휘 인식에 대한 관심이 모아지고 있으며, 어휘 인식에는 여전히 유사한 음소와 부정확한 어휘 제공에서 오류가 존재한다. 어휘의 입력에서 신호 처리를 통하여 어휘 인식 오류 보정에 관한 연구가 진행되고 있으며 사용 범위가 다양하고 화자 독립적인 시스템에서는 신호 처리만으로 어휘 인식의 효율을 높이는 것은 어려운 과제들이 따른다[1]. 따라서 어휘의 일반적인 신호 처리의 인식 결과에서 어휘 후처리를 이용한 오류 보정에 대한 연구를 진행하고 있다[2].

기존의 방법에서는 오류 패턴 매칭 기반의 오류 보정 방법과 어휘 의미 패턴 기반의 오류 보정 방법이 있다. 오류 패턴 매칭 기반의 오류 보정 방법[3]은 인식 과정에서 일정한 패턴을 갖는 오류가 존재하여 발화된 문장과 인식할 문장을 비교하여 오류 패턴을 미리 학습하여 후처리 과정에서 오류 패턴을 보정하는 방법이다. 어휘 의미 패턴 기반의 오류 보정 방법[4]은 인식할 문장을 의미적으로 분석하고 분석한 의미 정보를 특정 스트링에 포함하여 사용하는 방법이다. 오류 패턴 매칭 기반의 오류 보정 방법과 어휘 의미 패턴 기반의 오류 보정 방법에서는 오류 패턴 DB가 많이 필요하고 핵심어에만 이루어져 의미적으로 분석하기 어려운 문제점이 있다[5].

따라서 본 논문에서는 이 문제를 해결하기 위해 음절 복원 알고리즘을 이용한 핵심어 오류 보정 시스템을 제안한다. 인식된 음소 열을 의미 분석 과정을 거쳐 음소가 갖는 의미를 파악하고 음절 복원 알고리즘을 통해 음운 변동이 적용되기 이전의 문자열로 복원한다. 의미 분석 과정은 문법적인 어절과 비문법적인 어절의 의미를 파악하여 문법적인 어절은 복원 과정을 거치지 않고 인식시키고 비문법적인 어절은 복원 과정인 오류 보정을 수행하여 재인식시킨다. 의미적으로 분석하기 힘든 핵심어에만 이루어진 문장을 복원하므로 전체적인 어휘의 오인식을 줄일 수 있다. 이를 위해 음소 유사율 오류 보정 알고리즘을 개선하여 핵심어의 오인식을 줄여 문장을 복원할 수 있으며 오류 패턴 DB의 양이 감소되었다.

시스템 성능 평가는 음소 유사율을 측정하였고 신뢰도를 이용하여 평가하였다. 어휘 인식 과정에서 오류로 판명된 어휘에 대하여 오류 보정을 수행한 결과 3.0%의 인식 향상율을 보였다.

본 논문의 구성은 2장에서 음소 유사율 처리 오류 보정과 음절 복원에 대해서 살펴보고, 3장에서는 음소 유사율 오류 보정 알고리즘을 이용한 오류 보정 수행 방법에 대하여 설명한다. 4장에서는 실험 환경과 실험 결과, 5장에서는 결론을 다룬다.

II. 기존 연구

2.1 오류 패턴 매칭

오류 패턴 매칭 오류 보정은 인식과정에서의 오류는 일정한 패턴을 가지고 발생한다는 점을 전제로 하여 발화 문장과 인식 문장을 비교하여 오류 패턴을 학습하고 후처리 모듈에서 보정하는 방법으로 적은 비용과 시간으로 오류를 보정할 수 있다. 패턴 인식의 대상이 되는 패턴은 정적 패턴과 동적 패턴으로 나뉘고 고정된 영상의 패턴을 인식하는 경우와 시간에 따라 변화하는 패턴을 인식하는 경우이다. 패턴 인식은 패턴 매칭 방법에 의해 이루어지고 동적 패턴일 경우 시간상 늘이고 줄이는 신축을 허용하는 방식으로 패턴 매칭이 이루어지며 DTW(Dynamic Time Warping) 알고리즘을 이용한다[6].

길이가 다른 두 열의 인식 대상에서 한 열의 기준으로 두 열을 비교하기 위해서는 어느 한 열이 신장되거나 축소되어야 하며 직선의 의한 매핑 함수를 사용할 경우 선형 신축 매핑이고 곡선과 같은 비선형적인 경우 비선형 신축 매핑이라 한다.

선형 신축 매핑은 인식 대상열의 길이가 같은 경우 각 열의 성분끼리 일대일 대응으로 각 성분의 거리를 누적하여 총 누적 거리를 비교해 유사도를 평가한다. 하지만 인식 대상열의 길이가 다를 경우 매핑 함수를 적용하게 되며 신축 과정에서 인식 대상 열의 성분을 샘플링하고 신장 과정에서 보간하여 두 인식 대상 열의 길이를 같게 만든 후 비교하므로 인식 대상 열의 정보 손실로 인하여 패턴 매칭이 제대로 이루어지지 않아 비선형 패턴 매칭 방법을 사용하게 된다. 비선형 매핑 함수를 이용한 패턴 매칭 방법은 DTW 알고리즘이 사용된다. DTW 알고리즘은 인식 대상인 두 열의 각 성분에 대한 거리척도 값을 비용으로 설정하고 인식 대상 두 열이 이루는 격자상에서 각 열의 시작 성분에서 시작하여 끝 성분에 이르기까지 비용 테이블에 최소 비용을 순환적으로 택하여 저장하는 점화식을 이용하는 동적 계획법으로 매핑 함수를 찾아가면서 인식 대상 두 열을 비교하는 알고리즘이다. 최종적으로 끝 성분에서 비용 테이블에 저장되는 비용 값이 두 열에 대한 유사도로 표시된다. 매핑 함수의 궤적은 앞의 동적 계획법의 최

적 탐색패스를 찾는 것과 같이 탐색 과정에서 최소 비용을 택하는 경로를 별도의 경로 테이블에 매 단계마다 저장하고 끝 성분에서 최종 최소 비용을 구한 후에 역추적하여 비교 열의 궤적을 찾는다[7]. DTW 알고리즘은 비선형 시간 정규화를 갖는 패턴정합 알고리즘으로 공통적이고 균일하게 샘플간격을 갖게 음성패턴을 시간적으로 샘플화하여 인식한다.

패턴 매칭 기반의 오류 보정 방법은 인식 과정에서 인식 후보 열을 특징 추출 과정을 거쳐 기준 벡터열로 만든 후, 인식 대상 열에 대한 특징을 추출하여 시험 벡터 열을 각 후보 열과 DTW알고리즘을 이용하여 비교함으로써 후보 열의 카테고리를 인식 결과로 결정한다. 이 때 일정한 패턴을 갖는 오류가 존재하므로 오류에 대한 카테고리를 미리 작성하여 오류 패턴을 미리 학습하여 발화된 문장과 인식할 문장을 비교하여 오류로 판단되는 문장에 대해 후처리 과정에서 오류 패턴을 보정하는 방법이다[8].

2.2 어휘 의미 패턴

어휘 의미 패턴 오류 보정은 인식 문장을 의미적으로 분석하여 인식된 단어별로 의미 정보를 포함한 특정 스트링으로 대체하는 방법이다. 하지만 정보 검색 영역에서 사용되는 문장은 문장이 간결하고 사용자가 검색하고자 하는 핵심어로만 이루어진 경우가 많다. 어휘에 대한 의미를 찾기 위해선 기존의 어휘에 대한 데이터를 바탕으로 어떤 의미를 가지고 있는지 결정한다. 확률을 사용하면 기존 어휘를 바탕으로 현재 어휘와 다음 어휘를 확률적으로 예측가능하며 은닉 마코프 모델을 이용하여 구한다.

은닉 마코프 모델(HMM:Hidden Markov Model)은 숨겨져 있는 모델을 기존 모델의 확률을 이용하여 추정하는 방법이다. HMM은 초기 ($t = 0$)에 상태 i 의 확률 $\pi_i = \Pr(s_0 = i)$, 상태 i 에서 j 로의 천이 확률 $a_{ij} = \Pr(s_t = i, s_{t+1} = j)$, 상태 j 에서 심볼 k 를 관측할 확률 $b_j(k) = \Pr(x_t = k | s_t = j)$ 로 표현한다. 임의의 음성 특징벡터의 관측열 $O = (o_1, o_2, \dots, o_T)$ 이 사실임을 가정할 때 주어진 N-states HMM 모델에서의 상태열이 $q = (q_1, q_2, \dots, q_T)$ 라면 결국 관측열의 확률은 다음 식(1)과 같이 주어진다[9].

$$\begin{aligned}
 P(O|q, \lambda) &= \prod_{t=1}^T P(o_t | q_t, \lambda) \dots\dots\dots (1) \\
 &= \sum_{q \in Q} \pi_{q_1} b_{q_1}(o_1) a_{q_1 q_2} b_{q_2}(o_2) \dots a_{q_{T-1} q_T} b_{q_T}(o_T) \\
 &= \sum_{q \in Q} \prod_{t=1}^T a_{q_{t-1} q_t} b_{q_t}(o_t)
 \end{aligned}$$

초기 상태 $t = 1$ 에서 확률 $a_{s_0 s_1}(\pi_{s_1})$ 로 천이가 시작되며, 관측 O_1 는 출력 확률 $b_{s_1}(O_1)$ 로서 생성이 된다. 초기 상태 s_1 에

서 상태 s_2 로의 천이는 천이 확률 $a_{s_1 s_2}$ 로 이루어지며, 대응되는 상태 s_2 에서의 관측 O_2 를 생성될 확률은 $b_{s_2}(O_2)$ 가 된다. 이러한 과정은 상태 s_{T-1} 에서 마지막 상태 s_T 로 $a_{s_{T-1} s_T}$ 의 확률로 천이되어 기호 O_T 를 출력 확률 $b_{s_T}(O_T)$ 로 생성할 때까지 계속된다. 이러한 과정을 정의에 의하여 직접 계산하면 모든 시간 $t = 1, 2, \dots, T$ 에서는 진행 가능한 상태 수는 N 개가 되어 계산의 복잡도는 $O(N^T)$ 이 된다[10].

이러한 확률계산은 음성 구간에 따라 모델이 지수 함수적으로 증가하는 상태 열을 갖기 때문에 쉽게 계산할 수 없고 계산량이 지나치게 방대해지므로 전향, 후향 알고리즘을 이용하여 HMM 모델의 관측열의 확률을 추정한다[11].

인식 과정에서 인식 후보 열에 대해 특징 추출 과정을 수행하여 기준 벡터 열을 생성하고 인식 대상 열에 대한 특징을 추출하여 시험 벡터 열을 각 후보 열과 HMM알고리즘을 이용하여 인식 결과로 결정한다.

2.3 음소 오류 보정

오류 보정은 음소와 어휘에 대해서 결과로 나타난 인식 음소 열에 대해서 이루어진다. 음소 단위의 패턴 인식을 수행하여 얻어진 결과 리스트를 분석하여 오류 패턴을 가지고 원래의 음소로 되돌려 주는 과정이다[12]. 잡음 제거를 통하여 신호 처리에서 이루어지는 전처리 방법과 인식 결과를 분석하여 보정하는 후처리 방법으로 이루어진다. 전처리 방법의 잡음 제거는 워너 필터를 이용하여 입력 신호로부터 잡음 제거가 이루어지고 신호대 잡음비에 의해 신호가 높은 비율을 보이면 수렴하고 잡음의 비율이 높으면 제거한다. 입력된 어휘 신호의 잡음 제거는 메모리량에 따라 전체 신호에 대한 실행보다는 프레임 단위 잘라서 실행한다. 입력된 신호가 프레임 단위로 나뉘는 후에 각 프레임 별로 스펙트럼을 추정하고 스펙트럼 평활화를 수행한 후, 인식에 유용한 성분을 어휘 신호로부터 뽑아내는 특징 추출 과정을 거친다. 특징 추출은 정보의 압축과 차원의 감소 과정을 수행하며 추출된 특징의 좋고 나쁨은 인식률과 연관된다. 특징추출 과정에서는 청각 특성을 반영한 멜 필터 बैं크 분석과 프리엠퍼시스 필터를 사용한다[13].

멜 필터뱅크 분석[14]은 인간의 청각은 주파수 영역에서 음성을 비선형적으로 분석하는 특징을 가지며 1kHz 이상에서는 로그 주파수(logarithmic frequency) 영역에서 선형적으로 분석한다. 프리엠퍼시스 필터[15]는 저주파를 상대적으로 줄이고 고주파를 상대적으로 키우는 것으로 어휘 스펙트럼이 기본적으로 낮은 주파수에서 에너지가 많고 높은 주파수에서 에너지가 적은 특성을 가지고 있기 때문에 프리엠퍼시스를 통해서 스펙트럼의 경사를 평탄화한다.

음소 유사율 오류 보정은 오류로 발생한 인식 결과를 올바른 인식 결과로 보정한다. 어휘 인식 시스템에서 신호 처리만으로 보정하기 힘든 오류에 대해서 어휘 인식 결과로부터 보정하는 후처리 오류 보정을 수행한다.

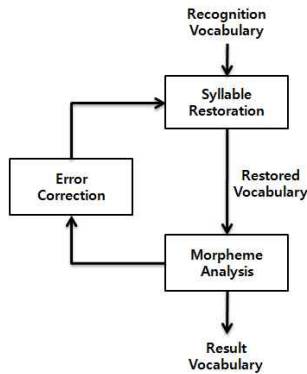


그림 1. 음소 오류 보정
Figure 1. Phoneme error restoration

그림 1은 음소 오류 보정의 구성도를 나타내고 있으며, 인식된 음소열에 대해 음절을 복원하여 음운 변동이 적용되기 이전의 문자열로 복원시킨다. 복원된 어절의 형태소 분석을 통해 문법적 어절과 비문법적 어절의 필터링을 하고 비문법적 어절에 대해 오류 보정을 수행한다[16].

III. 제안 시스템 모델

오류 보정 방법은 단순한 언어 모델이 가지는 한계점을 극복하지 못한다는 단점과 다양한 오류 패턴 DB가 필요하다. 문장이 간결하고 사용자가 검색하고자 하는 핵심어만 이루어진 경우에는 의미적으로 문장을 분석하기 힘들다. 이러한 단점을 보완하기 위해 음절 복원 알고리즘을 이용하여 핵심어 오류 보정을 위한 시스템 모델을 그림 2와 같이 구성하였다.

특징 추출을 효율적 처리하여 유사한 음소로의 인식을 보정한다. 훈련 데이터로부터 특징을 추출하고 거리 측정을 이용하여 오류율을 측정하여 거리를 계산한 후, 후처리에서 오류 보정하여 인식된 결과에 대하여 오류를 보정한다. 입력되어진 어휘로부터 핵심어 분석 처리를 통해 문장에서의 핵심어 사용 빈도 및 출연 횟수와 사용된 핵심어를 리스트로 작성하여 분석하고 오인식 처리를 통해 핵심어가 오인식된 리스트를 체크하는 과정을 거치게 된다. 오인식 처리에서 음소 유사율을 사용하여 음소 사이의 거리를 측정하고 수치로 나타내어 신뢰도 측정을 통한 인식된 결과를 확인한다. 인식 결과로부터

오류로 판명된 음소를 확인하고 오인식되어진 음소 열은 오류 보정을 수행한다.

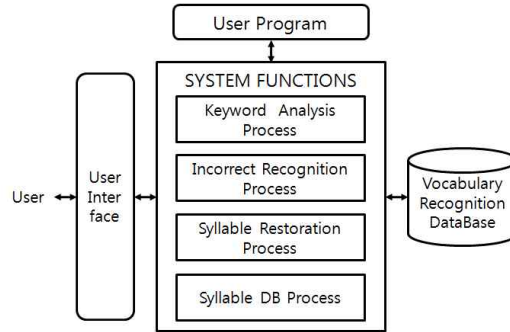


그림 2 시스템 모델
Figure 2. System Model

음소 DB 처리에서는 오인식된 어휘의 핵심어로부터 음소 열을 리스트에 추가하여 오류 음소 DB와 인식 음소 DB를 분류하여 재인식에 활용한다.

3.1 핵심어 분석

어휘 인식 시스템에서는 핵심어 모델(Keyword Model)과 필러 모델(Filler Model) 간의 구분이 명확하게 모델링되었을 경우 인식 성능이 좋게 나타나며 핵심어는 인식 대상 어휘로 사용되고 필러 모델은 그 외의 어휘들로 핵심어 모델과 유사하지 않은 모델로 사용된다. 필러 모델은 핵심어 모델과 유사하지 않으면서 얼마나 모델링이 잘되었느냐에 따라 인식 성능이 좌우된다. 핵심어 분석은 전처리부, 핵심어 검출부, 후처리부로 구성된다. 전처리부는 입력 어휘로부터 음성 구간을 검출하여 특징 파라미터를 추출하고 특징 파라미터를 12차의 멜 캡스트럼과 1차 정규화 로그 에너지, 파라미터의 델타 에너지를 사용하여 구성한다.

구성되어진 특징 파라미터는 핵심어 모델과 비핵심어인 필러 모델로 분류하여 리스트로 작성된다. 핵심어 검출을 위한 네트워크는 핵심어 모델들과 필러 모델들을 토른 패싱 알고리즘을 이용하여 결합한다[17]. 핵심어 검출 네트워크의 핵심어 모델들과 필러 모델들은 GMM(Gaussian Mixture Model)을 이용하여 어휘 모델에 대한 형태로 모델링된다. 그림 3은 핵심어 검출 네트워크에서 핵심어 모델과 필러 모델을 분류하는 과정을 나타낸다. 핵심어 모델과 필러 모델은 인식 네트워크에서 병렬로 연결되어 매 프레임마다 새로운 입력을 받아들인다.

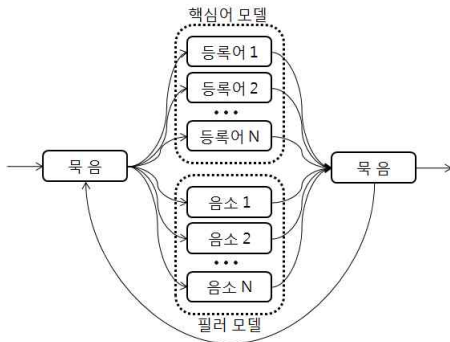


그림 3. 핵심어 모델과 필러 모델 분류 과정
Figure 3. Keyword model and filler model classification process.

후처리부는 핵심어의 오검출을 방지하기 위해 핵심어가 실제 발생되었는지를 검증하고 신뢰도를 계산한다.

3.2 오인식 처리

핵심어 분석을 통해 핵심어 모델에 의해 인식되어진 결과는 그대로 인식 결과로 사용하고 필러 모델에 의해 인식되어진 결과는 오인식 처리 과정을 거쳐 재인식을 통하여 핵심어와 필러 모델로 재분류된다. 오인식의 일정한 패턴을 찾기 위해 핵심어 모델을 이용하여 여러 번 인식을 수행하고 수행된 결과로부터 오인식 패턴을 정리하였다. 오인식 처리에 사용하기 위한 패턴은 오인식이 자주 발생하는 음소와 음소의 오인식 패턴을 쌍으로 묶어 오류 패턴 사전에 구축하였다. 인식 결과와 미리 구축된 오류 패턴 사전을 비교하여 오인식이 예상되는 부분을 검출하여 사전상의 원 단어로 치환하여 오류를 처리한다.

다음은 오인식 처리에 대한 알고리즘이다.

```

Begin
  Initialize the nonrecognition
Job1 : Get Input Vector XI
  If Vector XI is exist then next
  else goto end
  Clear NL
  Add the NL to Evaluation(n)
  Process the phone transition and
  Add the Expand(n) to NL
  Insert BP to BPframe
  Get the frame of word list BP for transition
  Process word transition
  Process the likelihood(ALL n∈CL)
  If BP=NL Job2 else Job 3
Job2 : Recognition
  Process word transition
  Create the max likelihood of BP table
Job3 : Non-Recognition
  Replace CL, NL
  goto Job1
    
```

오인식 오류 처리에 대한 알고리즘에서 입력되어진 어휘로부터 구축되어진 오류 패턴 사전과 비교하는 과정을 거쳐 음소 유사율로 표현된다. 유사율은 확률 값으로 나타나므로 음소의 신뢰도가 낮으며 음소와 음소 유사율이 높은 음소일수록 오류 보정율은 커지게 되며 음소의 신뢰도가 높거나 음소와 음소 유사율이 낮은 음소일수록 오류 보정율은 작아지는 것을 확인하였다.

3.3 음절 복원

필러 모델 인식되어 오인식으로 분류되어진 음소에 대해 음절을 복원하기 위한 방법으로 음절 복원 알고리즘을 이용하였다. 음절 복원 처리는 오류 패턴 사전과 비교하여 신뢰도와 음소 유사율을 고려하여 처리한다. 음소의 신뢰도가 높아도 음소 유사율이 높은 비슷한 음소로 대치되어 오류가 발생할 수 있으며, 신뢰도가 낮아도 제대로 인식될 수 있기 때문이다 [18]. 신뢰도가 낮으며 음소 유사율이 높은 음소를 가지고 있는 음소를 먼저 오류 보정을 수행해야 정확한 오류 보정을 할 수 있다. 식 (2)는 하나의 음절에 대한 오류 보정율을 나타낸다.

$$R_{\text{syllable}} = \frac{1}{n} \sum_{k=1}^n (1 - \alpha_k) b_k \dots\dots\dots (2)$$

n 은 한 음절의 음소 수, α_k 는 k 번째 음소의 신뢰도, b_k 는 k 번째 음소에 대한 유사율을 나타낸다. $\alpha = 1$ 일 때, 오류 보정율은 0의 값을 가지며, $\alpha \neq 1$ 일 때는 $0 < R_{\text{syllable}} < 1$ 의 값을 가진다. 음절을 구성하는 음소의 신뢰도와 음소에 대한 타 음소와의 유사율을 이용하여 각각의 음소에 대한 후보 음절을 생성한다. 생성된 후보 음절을 통하여 우선순위 어절을 선정한다.

인식 오류의 경우 인식 결과에 상위 2-3위에 정확한 인식 결과를 포함하고 있는 경우가 대부분이므로 인식 결과를 분석하여 오류를 분류하고 오류의 대한 복원을 수행한다. 음절 복원 알고리즘은 우선순위 음절을 선정하는데 사용되며 우선순위 음절은 오류 보정율의 평균으로 구하여 음절 복원 및 오류 패턴 사전과 비교하여 수행한다.

입력된 음절로부터 음소의 신뢰도와 유사율을 구하여 오류 보정율이 0의 값을 갖게 되면 보정이 이루어지지 않고 0과 1 사이의 값을 가질 때 보정이 이루어지며 1에 가까운 값이 가장 높은 보정율을 나타낸다.

음절 복원 알고리즘은 다음과 같다.

```

Begin
  Initialize the nonrecognition
Job1 : Get Input Vector XI
  If Vector XI is exist then next
  else goto end
  Clear NL
  Process the likelihood(ALL n∈CL),
  Add the NL to Evaluation(n)
  Process the phone transition and
  Add the Expand(n) to NL
  Insert BP to BPframe
  Get the frame of word list BP for transition
  Process word transition
  Replace CL, NL
  Goto Job1
Job2 : Create the word graph list having
  max likelihood of BP table
    
```

3.4 음절 DB 관리

음절 DB 관리에서는 입력되어진 음절로부터 핵심어 모델과 필러 모델을 구축하였으며 오인식의 일정한 패턴을 갖는 오류 패턴 DB를 그림 4와 같이 구축하였다. 오인식의 일정한 패턴을 바탕으로 오인식 패턴을 정리하여 사전 DB를 구축한 오인식 패턴 DB는 오인식이 자주 발생하는 음소와 음소의 오인식 패턴을 쌍으로 묶어 오류 패턴 사전 DB를 구축하였다.

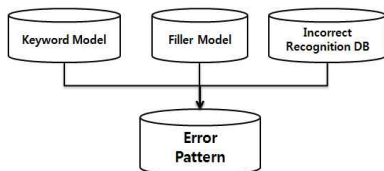


그림 4. 오류 패턴 DB
Figure 4. Error pattern DB

구축되어진 데이터는 재인식 시 오류 패턴 DB로 활용하고 입력되어진 음절로부터 핵심어 모델과 필러 모델의 유사율과 오류 보정율을 계산하기 위하여 사용된다.

IV. 실험결과 및 분석

본 논문에서 음절 복원 알고리즘을 핵심어 오류 보정 시스템을 구성하여 인식 실험을 수행하였다. 본 실험에서는 단말기에서 사용되는 단어를 선정하여 총 20개를 표 1과 같이 선별하였고 후처리 실험을 위해 5명의 사용자가 100회의 데이

터를 구축하였다. 어휘는 실내 환경과 잠음 환경에서 이동기기 등에 내장되어 있는 내장형 마이크로폰을 사용하여 16kHz Mono로 녹음 하였고, 16bit PCM 양자화를 사용하였다. 실험 어휘는 실내 10명, 실외 5명 등 총 15명의 성인 남성이 참가하였다. 녹음된 데이터는 인식기 학습을 위해 특징 추출 방법을 사용하였고 인식기는 SITEC에서 개발한 ECHOS[19]를 이용하였다. ECHOS는 각 단어별 데이터로 학습된 인식 모델을 이용하여 발화된 단어의 인식 가능한 단어들의 인식 가능 확률을 표현하고 최대값을 가지는 단어를 최종 결과로 선정한다.

표 1. 실험 데이터
Table 1. Experimentation data

No.	단어	No.	단어	No.	단어	No.	단어
1	위	2	아래	3	시작	4	다음
5	종료	6	상위	7	하위	8	정지
9	크게	10	작게	11	확대	12	축소
13	취소	14	예	15	찾기	16	도착지
17	출발지	18	경로	19	확인	20	아니오

음소 유사율의 구성에 따른 음소 유사율의 정확성을 확인하기 위하여 특징 추출 방법을 사용하여 음소 유사율을 구성하고 어휘 인식 후처리에서 오류보정 시스템 모델에 적용하였다.

표 2는 기존의 에러 패턴 학습, 에러 패턴 매칭, 어휘 의미 패턴, 제안방법의 인식률과 오류 보정율을 비교 실험한 결과이다.

표 2. 오류 보정율 비교
Table 2. Comparison of error correction

오류 보정	인식률(%)	보정율(%)
에러 패턴 학습	80.2	4.3
에러 패턴 매칭	82.6	5.5
어휘 의미 패턴	84.5	6.3
제안 방법	83.2	6.1

에러 패턴 학습을 이용한 오류 보정의 경우 4.1%, 에러 패턴 매칭의 경우 5.5%, 어휘 의미 패턴의 오류 보정의 경우 6.3%의 인식 향상율을 보였으며, 제안한 방법을 이용한 경우 6.1%의 보정율을 보였다.

V. 결론

어휘 인식 시스템에서 오류 보정을 통한 인식률 향상을 위하여 오류 패턴 매칭 방법과 어휘 의미 패턴 방법을 이용한다. 이들 방법에서는 음절을 구성하고 있는 어휘가 핵심어로 이루어져 의미적으로 분석하기 어려운 문제점과 이를 복원하기 위한 오류 패턴 DB가 많이 필요하여 본 논문에서는 음절 복원 알고리즘을 적용한 핵심어 오류 보정 시스템을 제안하였다.

인식된 음소 열을 의미 분석 과정을 거쳐 음소가 갖는 의미를 파악하고 음절 복원 알고리즘을 통해 음운 변동이 적용되기 이전의 문자열로 복원한다. 의미 분석 과정은 문법적인 어절과 비문법적인 어절의 의미를 파악하여 문법적인 어절은 복원 과정을 거치지 않고 인식시키고 비문법적인 어절은 복원 과정인 오류 보정을 수행하여 재인식시킨다. 의미적으로 분석하기 힘든 핵심어로만 이루어진 문장을 복원하므로 전체적인 어휘의 오인식을 줄일 수 있다. 이를 위해 음소 유사율 오류 보정 알고리즘을 적용하여 핵심어의 오인식을 줄여 문장을 복원할 수 있으며 오류 패턴 DB의 양이 감소되었다.

제안한 음절 복원 알고리즘을 적용한 핵심어 오류 보정 시스템으로 인하여 인식률을 향상시킬 수 있는 장점을 확인하였으며 검색 시 속도와 인식률에서 기존 에러 패턴 학습, 에러 패턴 매칭 오류 보정을 시스템보다 나은 결과를 얻을 수 있었다. 시스템 성능 평가는 음소 유사율을 측정하였고 신뢰도를 이용하여 평가하였다. 어휘 인식 과정에서 오류로 판명된 어휘에 대하여 오류 보정을 수행한 결과 보정율이 6.1%를 나타내었으며 에러 패턴 학습의 인식률의 비교 3.0%의 인식 향상율을 보였다.

참고문헌

- [1] 한동조, 최기호, “음성인식 후처리에서 음소 유사율을 이용한 오류보정에 관한 연구,” 한국TTS학회논문지, 제 6권, 제 3호, 77-86쪽, 2007년 12월.
- [2] 조시원, 이동욱, “음성 인식 후처리를 위한 연속 음절 문장의 키워드 추출 알고리즘,” 대한전기학회, 학술대회 2008 정보 및 제어 심포지엄(ICS'08) 논문집, 170-171쪽, 2008년 4월.
- [3] S. Kaki, E. Sumita, and H. Iida, “A method for correction speech recognition using the statistical features of character co-occurrence,” Proc. COLING-ACL, pp.653-657, Aug, 1998.
- [4] 김용현, 정민화, “에러패턴 학습과 후처리 모듈을 이용한 연속 음성 인식의 성능향상,” Proc. KISS Spring Semiannual Conf. 제 27권, 제 1호, 441-443쪽, 2000년 4월.
- [5] M. W. Jeong, B. C. Kim, and G. G. Lee, “Semantic-oriented error correction for spoken query processing,” Proc. IEEE Workshop on ASRU, pp.156-161, Nov, 2003.
- [6] 최용선, 오상훈, 이수영, “핵심어 검출을 위한 단일 끝점 DTW 알고리즘,” 대한전자공학회논문지, 제 41권, SP편 제 3호, 209-219쪽, 2004년 5월.
- [7] Eiichi Tanaka and Tamotsu Kasai, “Synchronization and Substitution Error-correcting codes for the Levenshtein Metric,” IEEE Trans. Information Theory, Vol.IT-22, No.2, pp.156-176, 1976.
- [8] E. K. Ringer and J. F. Allen, “A fertility channel model for post-correction of continuous speech recognition,” Proc. ICSLP, pp.897-900, Oct, 1996.
- [9] S. Kaki, E. Sumita, and H. Iida, “A method for correction speech recognition using the statistical features of character co-occurrence,” Proc. COLING-ACL, pp.653-657, Aug, 1998.
- [10] L. Rabiner and B. H. Juang, “Fundamentals of Speech Recognition”, Prentice-Hall, 1993.
- [11] 안찬식, 오상엽, “MLHF 모델을 적용한 어휘 인식 탐색 최적화 시스템,” 한국컴퓨터정보학회지, 제 14권, 제 10호, 217-223쪽, 2009년 10월.
- [12] 송원문, 김명원, “문맥 및 사용 패턴 정보를 이용한 음성 인식 후처리,” 정보처리학회논문지, 제 13-B권, 제 5호, 553-560쪽, 2006년.
- [13] 박미성, 김미진, 김계성, 최재혁, 이상조, “연속 음성 인식 후처리를 위한 음절 복원 rule-based 시스템과 형태소분석기법의 적용,” 대한전자공학회논문지, 제 36권, 제 3호, 47-57쪽, 1999년 3월.
- [14] M. F. Gales, “Model-based techniques for noise robust speech recognition,” Ph. D. dissertation, University of Cambridge, Sept, 1995.
- [15] S. Young, D. Kershaw, J. Odell, D. Ollason, Valtcher, P. Woodland, “The HTK Book,” Cambridge University Engineering Department, 2002.
- [16] 박미성, 김미진, 김계성, 김성규, 이문희, 최재혁, 이상조, “형태소 분석 기법을 이용한 음성 인식 후처리,” 대한전자공학회논문지, 제 36권, 제 4호, 65-77쪽, 1999년 4월.

- [17] 안찬식, 오상엽, “공유모델 인식 성능 향상을 위한 효율적인 연속 어휘 군집화 모델링,” 한국컴퓨터정보학회지, 제 15권, 제 1호, 177-183쪽, 2010년 1월.
- [18] T. Jitsuhiro, S. Takatoshi, and K. Aikawa, “Rejection of out-of-vocabulary words using phoneme confidence likelihood,” Proc. ICSSP, pp.217-220, May, 1998.
- [19] 음성정보기술산업지원센터, “한국어 음성인식 플랫폼 사용자 매뉴얼(ECHOS Manual),” 135-308쪽, 2006년.

저 자 소개



안 찬 식

2002: 광운대학교 컴퓨터공학과
공학석사
2004: 광운대학교 컴퓨터공학과
박사수료
관심분야: 음성인식, 분산처리,
음성/음향 신호처리



오 상 엽

1999: 광운대학교 전자계산학과
이학박사.
현 재: 경원대학교 IT대학
컴퓨터소프트웨어 교수
관심분야: 소프트웨어공학, 버전관리,
소프트웨어제사용, 형상관리,
객체지향, 음성인식, 분산
처리, 음성/음향 신호처리