
어휘 인식 시스템의 인식률 향상을 위한 어휘 유사율 처리 지원

김규호*, 오상엽**

Vocabulary Likelihood rate Process support for Recognition rate Improvement of Vocabulary Recognition System

Kyuhoo Kim*, Sang Yeob Oh**

요 약 어휘 인식 모델에서는 정확하지 않은 어휘로부터 특징을 추출하기 때문에 어휘가 실제 어휘와 유사한 어휘로 인식되거나 인식이 되지 않는 현상이 나타난다. 이를 위해 본 논문에서는 효율적인 형상 형성을 지원하는 시스템을 모델링하고 구현하였으며, 형상 형성 정보를 효율적으로 처리하고 어휘 유사율 관리를 최적화하기 위해 데이터베이스 검색에서 facet 방법을 응용하였다. 본 논문에서 제안한 시스템을 적용한 결과 시스템 성능에서 어휘 종속 인식률은 95.31%, 어휘 독립 인식률은 97.38%의 인식률을 나타내었다.

주제어 : 어휘 인식, 형상형성, 어휘 데이터, facet, 인식율

Abstract In the vocabulary recognition model, system has some problems that vocabulary is nor recognize and similar vocabulary recognition is created., because it is caused by system extract vocabulary feature from inaccurate vocabulary. To solve this problems, this paper propose the system modeling and implementation for efficient configuration thread support system, it process the configuration thread information and it apply the facet method in database retrieve for optimization of vocabulary likelihood rate. Proposed system showed 95.31% of vocabulary dependency recognition rate and 97.38% vocabulary independency recognition rate in system performance.

Key Words : Vocabulary recognition, configuration thread, vocabulary data, facet, recognition rate

1. 서 론

대부분의 어휘 인식 시스템에서는 구현되는 어휘 인식 시스템의 사용 어휘를 미리 지정하고 이에 대한 어휘들을 데이터베이스로 구축하여 사용하고, 이 데이터베이스에서 인식되는 어휘나 어휘 모델을 훈련하게 된다. 이와 같은 어휘 인식 시스템에서는 데이터베이스에 정의된 어휘들은 인식률이 높게 나타나지만 새로운 어휘를 인식할 필요가 있을 때는 다시 모델을 훈련해야 하는 불편함을 가지고 있다.[1][5] 또한, 어휘 인식에서 정보 검색을 위한 영역으로 사용되는 어휘들은 어휘가 간결하고 사용

자의 정보 검색을 위한 핵심어로 구성되어진다. 따라서 정보 검색을 위한 영역으로 사용되는 어휘들은 의미적인 분석의 어려움이 따르고 어휘의 오인식으로 인해 문장이 전체적으로 다른 의미를 가지는 단점이 있다[3].

본 논문에서 제안한 어휘 인식 시스템에서 입력 어휘는 어휘 유사율 처리를 통해 어휘소 사이의 거리를 측정하여 수치로 나타내고, 인식 결과로부터 오류로 판명되어진 어휘를 확인하여 분류하고 오인식되어진 어휘열은 문법적 어절과 비문법적 어절에 대해 필터링하여 문법적 어절은 인식시키고 비문법적 어절은 오류 보정을 수행한다. 어휘 유사율 처리는 어휘 사이의 거리를 측정하여 비

**본 논문은 2012년 가천대학교의 교내연구비 지원에 의한 결과임(GCU-2012-R202).

*울지대학교 의료 IT 마케팅학과 교수(제1저자)

**가천대학교 IT대학 인터랙티브미디어학과 교수(교신저자)

논문접수: 2012년 10월 25일, 1차 수정을 거쳐, 심사완료: 2012년 11월 20일

슷한 어휘를 찾아 관리하며, 등록되어진 음절과 미등록되어진 음절을 분류하여 음절 복원 시 등록어를 핵심어로 인식한다.

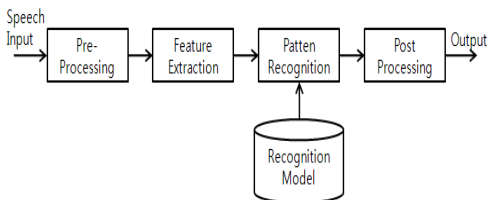
이러한 문제를 해결하고, 효율적으로 형상 형성을 지원하기 위한 시스템을 모델링하고 구현하여 훈련된 어휘 정보로부터 새로운 어휘 정보 생성을 위한 정보를 제공하여 어휘 분석의 효율성을 제공하고, 어휘 오인식 문제를 줄이도록 하였으며, 제안된 방법에서는 데이터베이스의 검색 지원을 위해 확장 facet 방법을 응용하여 설계하였다. 이로 인해 자주 액세스되는 어휘의 facet과 facet 내의 각 어휘에 대해 사용자가 빈번히 참조되는 어휘 facet에 높은 가중치를 할당하여 사용자가 특정 facet를 검색할 때, 도움이 되도록 하였다. 본 논문에서 제안한 시스템을 적용한 결과 시스템 성능에서 어휘 종속 인식률은 95.31%, 어휘 독립 인식률은 97.38%의 인식률을 나타내었다.

제 2장에서는 어휘 인식 탐색 시스템과 형상형성 제어에 대해 설명하고, 제 3장에서는 본 논문에서 제안한 어휘 유사율 처리와 형상형성 관리에 대하여 설명한다. 제 4장에서는 제안한 시스템의 실험 결과에 대하여 설명하고 제 5장에서 결론을 기술한다.

2. 관련연구

2.1 어휘 인식 시스템

어휘 인식 시스템은 화자의 음성이 마이크로폰으로 입력이 되면 화자 음성의 시작점과 끝점을 검출한다. 음성 구간이 검출되면 특징 검출을 하고, 검출된 특징을 이용하여 잡음제거 및 채널 보상을 수행한다. 전처리 과정을 거친 특징 벡터는 모델을 사용하여 인식 대상 어휘를 비교하는 인식과정을 거친다. 인식된 단어는 등록된 단어인지 아닌지를 판별하는 발화 검증 단계를 거치게 되며 일반적인 음성인식 과정은 그림 1과 같이 나타낸다[6].



[그림 1] 어휘인식 과정

음성 인식에서의 전처리 과정은 음성의 음향학적 특징을 포함하는 Formant, Pitch, Autocorrelation 등과 음성의 발생 모델에 근거한 LPC(Linear Predictive Coefficient), 청각특성의 비선형성에 기인한 MFCC(Mel Frequency Cepstrum Coefficients), PLP(Perceptual linear predictive) 등의 특징 파라미터를 추출하는 과정이다. 음성 인식의 경우 화자의 개인성이나 주위 잡음 환경에 비교적 강한 성능을 보이는 MFCC나 PLP를 주로 이용하고, 화자 인식의 경우 이러한 특성이 화자의 개인성을 손상시켜 화자간의 변별력을 약화시키는 주요한 원인이 되므로 성도 모델에 기인한 LPC나 이를 캡스트럼 변환한 LPCC를 이용한다[4].

전처리 과정에서 첫 단계가 음성구간의 추출 및 분절이며 마이크를 통해 입력된 음성에서 언어 정보 또는 화자 정보를 포함한 음성 구간 외에 묵음 및 잡음 구간 등을 찾아내고 레이블링하는 분절 알고리즘은 전체 시스템의 성능에 매우 큰 영향을 미친다.

2.2 형상형성

형상 형성의 초기 형태라고 할 수 있는 것으로는 파일 하나의 이력을 관리할 수 있는 시스템인 SCCS와 RCS가 있고, 형상 형성을 위한 보조도구로는 make 프로그램이 있다. 이 2가지 시스템을 결합하면 기본적인 형상 형성 기능을 수행할 수 있다. Make 프로그램은 본래 시스템 구성(build)에 사용되는 도구이지만, 의존 관계를 표현하는 규칙에 버전 선택 명령어를 사용하면 특정 버전이나 가장 최근의 버전을 자동으로 획득할 수 있는 방법을 제공한다. 그러나 make도구에서 파일들 간의 관계 설정에 사용되는 makefile의 관계설정 방법과 버전 선택 방법은 복잡하고, 여러 개의 makefile을 유지하는 경우 이들간의 정규적인 매개변수 전달 방법이 존재하지 않을 뿐만 아니라 단순한 버전 록(version lock)으로만 동시성 제어를 하고 있기 때문에 이 도구를 대규모 환경이나 분산 환경에 사용하는 것은 불가능하다는 단점을 갖는다.

이에 비해, Apollo를 기반으로 하는 DSEE[2]는 시스템 모형(system models)과 BCT(Bound Configuration Threads)라는 높은 수준의 시스템 형성형성 방법과 버전 선택 방법을 제공한다. Clearcase는 DSEE의 기능을 확장하고 Apollo와 같은 특정 운영체제나 하드웨어에 종속되지 않고 여러 플랫폼, 특히 UNIX상에서 작동할 수 있도록 개선하였다. 그러나 시스템 모형과 BCT를 이용한

방법은 시스템 형성정보가 개발 생산성에만 초점이 맞추어져 있다는 점과 버전을 선택하는 방법이 경직되어 있다는 점에서 형상 형성에 대한 관리 기능 측면에서 단점을 가진다. 최근 PC상에서 사용되는 버전 제어 도구로 MS사의 SourceSafe가 있다. 이 도구에서는 형상 형성과 파일 공유를 위해 레이블과 핀이라는 방법을 사용한다. CCC/Harvest는 앞에서 설명한 Clearcase나 SourceSafe와는 다른 방법으로 형상 형성을 지원한다. 다른 형상 형성 방법과는 달리 프로세스 지원을 중요하게 생각하는 철학 위에서 개발되었기 때문에 다른 시스템에서는 지원하지 않는 package와 package group 이라는 개념을 사용하고, 상태(state)를 옮겨 다니면서 작업을 하도록 지원한다. Sablime은 CCC/Harvest의 package와 유사한 개념인 MR(Modification Request)을 사용한다. MR에서는 버전 변경 요구 내용을 서술하며, 이를 이용하여 버전을 인출하고 저장한다. Sablime에서는 생명 주기 모형이 고정되어 있어서 사용자의 요구에 맞게 고칠 수 없는 문제를 가진다.

3. 시스템 모델

어휘 인식 문장이 간결하고 사용자가 검색하고자 하는 핵심어로만 이루어진 경우가 많으므로 의미적으로 분석하기 힘들다. 이러한 단점을 보완하기 위해 어휘 유사율 처리를 통해 정확한 어휘의 입력을 제공한다.

3.1 어휘 유사율 처리

어휘 유사율은 미리 학습되어진 음절로부터 생성된 어휘와 입력되어지는 음절로부터 생성된 어휘를 비교하여 유사율을 구한다. 유사율은 두 어휘 사이의 거리를 계산하여 가장 근접한 값을 취하는 형태이며, 두 어휘 집합 간의 분리도를 측정하는 통계적 수단으로 두 개의 가우시안 분포 사이의 거리를 측정한다. 계산이 단순하고 두 어휘간의 오류에 대한 경계 값을 제공함으로써 계산 시 유연성을 갖는다.

식 (1)은 어휘 집합 ω_1 과 어휘 집합 ω_2 에 대해 두 어휘 집합을 정의한다.

$$b = -\ln \int_{\Omega} [P(X|\omega_1) P(X|\omega_2)]^{\frac{1}{2}} dX \quad (1)$$

$P(X|\omega_1)$ 와 $P(X|\omega_2)$ 는 어휘 집합 ω_1 과 ω_2 의 대한 확률 밀도 함수이고, Ω 는 확률 분포상에서 정의되는 랜덤 함수 x 의 영역이다. 어휘 집합의 확률 분포를 정규 분포로 가정할 경우 식 (2)와 같이 나타낸다.

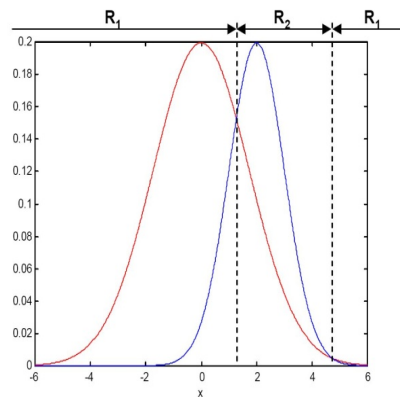
$$b = \frac{1}{8}(u_2 - u_1)^T \left[\frac{\Sigma_1 + \Sigma_2}{2} \right]^{-1} (u_2 - u_1) + \frac{1}{2} \ln \frac{|\Sigma_1 + \Sigma_2|}{|\Sigma_1| |\Sigma_2|} \quad (2)$$

u_i 는 어휘 집합 ω_i 의 평균 벡터를 나타내고 Σ_i 는 어휘 집합 ω_i 의 공분산 행렬을 나타낸다.

$$P(X|\omega_1) = \frac{1}{\sqrt{2\pi}\sqrt{3}} e^{-\frac{1x^2}{2 \cdot 3}} \quad (3)$$

$$P(X|\omega_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-2)^2} \quad (4)$$

식 (3)과 식 (4)는 두 어휘 집합에 위한 어휘 유사율을 우도비로 측정하기 위해 계산된 식이며, 두 식에 의해 어휘 집합을 측정하여 어휘 유사율을 그림 2에 표현하였으며, 식 (3)과 (4)에서 우도비 측정 계산식의 결과로 R_1 과 R_2 의 간격이 결정된다.



[그림 2] 두 어휘 집합의 측정된 어휘 유사율

식 (5)은 평균과 분산의 의한 거리를 계산하기 위한 수식으로 어휘 유사율을 측정하는데 사용된다.

$$d(i, j) = \left[\frac{1}{n} \sum_{k=1}^n \frac{(\mu_{ik} - \mu_{jk})^2}{\sigma_{ik} \sigma_{jk}} \right]^{\frac{1}{2}} \quad (5)$$

식 (6)은 식(5)의 어휘 유사율 분포를 [0, 1]사이에서 분포하도록 정규화를 위해 사용한다.

$$S_p = \frac{b_{max} - b_{xy}}{b_{max} - b_{min}} \quad (6)$$

S_p 는 두 어휘의 유사율을 측정하였으며 b_{max} 와 b_{min} , b_{xy} 는 최대값, 최소값, 두 어휘 와 의 거리를 나타낸다. 유사한 어휘일수록 1에 가깝고, 유사하지 않을수록 0에 가깝게 나타낸다.

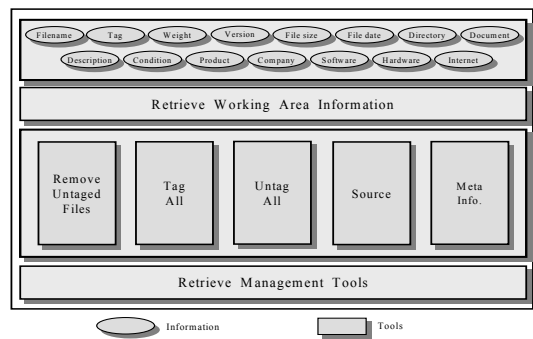
3.2 형상형성 제어

본 논문에서는 객체 지향 프로그래밍 언어인 MFC를 사용하여, 어휘인식에서 형상형성 지원을 위한 시스템과 형상 형성 정보 관리 프로그램을 구현하고, 이를 위한 사용자 인터페이스를 제공한다. 이것은 어휘의 등록과 검색을 지원하여 형상 형성 정보를 제공하기 위한 사용자 환경을 구축한다. 시스템에서 사용되는 CompVersion 클래스는 두 개의 어휘를 입력받아서 이를 비교하여 처리하기 위한 클래스이다. CopyProject 클래스는 특정 facet를 사용자가 작업 영역에 복사해 주는 작업을 처리하고 데이터베이스로부터 제공된 facet을 후에 수정하여 새로운 facet을 생성할 수 있도록 한다. VersionDetail 클래스는 facet의 생성과 삭제, facet에 대한 설명을 추가할 수 있는 기능을 지원한다. VersionView 클래스는 트리 관리를 위한 클래스로서 데이터베이스에서 사용되는 facet 구조에 저장된 어휘를 트리 구조를 이용하여 표시해 주는 역할을 수행한다. MainForm 클래스는 CompVersion 클래스와 연계되어 있으며, 실제 두 개의 버전 간의 비교 작업은 이 클래스에서 수행된다. FacetSet 클래스는 facet에 대한 가중치를 처리한다. 어휘 제어는 구현된 VocabularySearch 클래스를 이용하여 다양한 검색 기능을 수행하고, WindowApp클래스와 Dialog 클래스를 사용하여 탐색을 위한 환경을 지원한다. 버전 제어 시스템을 위한 VocabularyRetrieve 클래스는 다음과 같다.

또한, 형상형성을 지원하기 위한 형상형성 관리 도구들과 작업정보는 다음 그림 3과 같다.

```
class VocabularyRetrieve {
public:
    VocabularySearch(PTWindowsObject pParent,
                    char drive);

    virtual ~VocabularyRetrieve();
    void Retrieve();
protected:
    virtual int checkVocabulary(struct fsblk& fs) = 0;
    void print(struct fsblk& fs);
    char filePattern[13];
    int startDisk;
    char startPath[MAXPATH];
};
```



[그림 3] 형상형성 관리 도구들과 작업정보

어휘 비교 결과는 라이브러리의 해당 어휘가 속하는 facet에 저장되며, 이 정보는 그림 2에서와 같이 메타 정보에 포함되어 처리된다.

4. 실험 결과 및 분석

본 논문에서 제안한 시스템을 다른 시스템과의 최적화된 어휘 관리를 고려하여 비교 분석하였다. 3장에서 설명한 바와 같이 본 논문에서 제안한 시스템은 최적화된 어휘 관리를 지원하며 데이터베이스에서 facet 방법을 적용하여 미등록된 어휘 관리의 효율을 증진시킬 수 있다. 어휘 목록은 서울 시내의 지역명 100개, 지하철역명 50개로 구성하였다. 인식 실험에서는 실험에 참가한 화자가 어휘 목록을 5회 발음하여 총 500단어를 대상으로 실험을 수행하였다. 어휘는 실내 환경과 잡음 환경에서 이동 기기 등에 내장되어 있는 내장형 마이크로폰을 사용하여 16kHz Mono로 녹음 하였고, 16bit PCM 양자화를 사용하였다. 실험 어휘는 실내 5명의 성인 남성이 참가하였다. 기존의 인식기와 비교 실험을 위하여 캠브리지 대학

의 HTK(Hidden Markov Toolkit), 카네기멜론대학의 Sphinx를 사용하였다.[7]

[표 1]은 기존의 방식인 HTK Sphinx와 본 논문에서 제안한 방법을 이용한 실내 환경에서의 실험을 나타낸다. 실내 환경은 50~55dB에서 실험 하였다. 결과에서 보는 것과 같이 시스템 성능 평가 결과 어휘 중속 인식률은 95.31%, 어휘 독립 인식률은 97.38%의 인식률을 나타내었다.

〈표 1〉 인식률

어휘	인식률(%)		
	HTK	Sphinx	제안방법
어휘 중속	94.79	93.15	95.31
어휘 독립	96.77	95.56	97.38

소음이 50~55dB인 실내 환경 실험에서 어휘 중속 실험과 어휘 독립 실험을 한 결과 어휘 중속 실험에서는 제안 방법이 HTK 보다는 조금 낮게 나타났지만 Sphinx 보다는 2.16% 높게 나타났다. 어휘 독립 실험에서도 제안 방법이 HTK 보다는 조금 높게 나타났지만 Sphinx 보다는 1.82% 높게 나타난 것을 확인할 수 있었다.

5. 결론

본 논문에서는 본 논문에서 제안한 어휘 인식 시스템에서 입력 어휘는 어휘 유사율 처리를 통해 어휘소 사이의 거리를 측정하여 수치로 나타내고, 어휘 사이의 거리를 측정하여 비슷한 어휘를 찾아 관리하며, 등록되어진 음절과 미등록되어진 음절을 분류하여 음절 복원 시 등록어를 핵심어로 인식한다. 또한 어휘 인식 시스템을 효율적으로 관리하기 위한 형상 형성을 지원하기 위한 시스템을 모델링하고 구현하여 훈련된 어휘 정보로부터 새로운 어휘 정보 생성을 위한 정보를 제공하여 어휘 분석의 효율성을 제공하고, 어휘 오인식 문제를 줄이도록 하였으며, 제안된 방법에서는 데이터베이스의 검색 지원을 위해 확장 facet 방법을 응용하여 설계하였다. 본 논문에서 제안한 시스템을 적용한 결과 시스템 성능에서 어휘 중속 인식률은 95.31%, 어휘 독립 인식률은 97.38%의 인식률을 나타내었다.

참고 문헌

- [1] 안찬식, 오상엽 (2012). CHMM 어휘 인식에서 형상 형성 제어를 이용한 가우시안 모델 최적화, 디지털 정책연구 논문지, 10(7), 167-172.
- [2] 이태훈 (1997). 분산 소프트웨어 개발에서 관계관리 지원을 위한 형상 관리 기법, 서울대학교 계산 통계학과 박사학위 청구논문.
- [3] M. F. Gales (1995). Model-based techniques for noise robust speech recognition, Ph. D. dissertation, University of Cambridge..
- [4] T. Jitsuhiro, S. Takatoshi, and K. Aikawa(1998). Rejection of out-of-vocabulary words using phoneme confidence likelihood", Proc. ICSSP, 217-220.
- [5] Yeonja Lim, Youngjik Lee (1995). Implementation of the POW(Phonetically Optimized Words) algorithm for speech database, Proc. of ICASSP, 89-91.
- [6] L. Rabiner and B. H. Juang (1993). Fundamentals of Speech Recognition, Prentice-Hall, 1993.
- [7] S. Young, D. Kershaw, J. Odell, D. Ollason, Valtcher and P. Woodland (2002). The HTK Book, Cambridge University Engineering Department, pp.17-81, 2002.

오 상 엽



- 1991년 2월: 광운대학교 대학원 전자계산학과(이학석사)
- 1999년 2월: 광운대학교 대학원 전자계산학과(이학박사)
- 2007년 2월~현재: 가천대학교 IT대학 인터랙티브미디어학과 교수
- 관심분야: ,비전관리, 형상관리, 음성/음향 신호 처리, 차량 통신

· E-Mail: syoh@gachon.ac.kr

김 규 호



- 1991년 2월: 광운대학교 대학원 전자계산학과(이학석사)
- 1998년 8월: 광운대학교 대학원 전자계산학과(이학박사)
- 1992년 3월~현재: 을지대학교의료 IT마케팅학과 교수
- 관심분야: u-Healthcare System Network Management

· E-Mail : khkim@eulji.ac.kr