

자동 대소문자 식별을 이용한 영어 음성인식 결과의 가독성 향상

김지환(LG전자기술원)

<차 례>

- | | |
|---------------|---------------------|
| 1. 서론 | 4. 실험 방법 및 결과 측정 방법 |
| 2. 기존의 연구 | 5. 자동 대소문자 생성 결과 |
| 3. 자동 대소문자 생성 | 6. 결론 |

<Abstract>

Readability Enhancement of English Speech Recognition Output Using Automatic Capitalisation Classification

Ji-Hwan Kim

A modified speech recogniser have been proposed for automatic capitalisation generation to improve the readability of English speech recognition output. In this modified speech recogniser, every word in its vocabulary is duplicated: once in a de-capitalised form and again in the capitalised forms. In addition its language model is re-trained on mixed case texts.

In order to evaluate the performance of the proposed system, experiments of automatic capitalisation generation were performed for 3 hours of Broadcast News(BN) test data using the modified HTK BN transcription system. The proposed system produced an F-measure of 0.7317 for automatic capitalisation generation with an SER of 48.55, a precision of 0.7736 and a recall of 0.6942.

* Keywords: Capitalisation generation, Readability, Speech recognition

1. 서 론

음성인식기는 발음 사전(pronunciation dictionary)에 있는 단어들로 조합되는 단어열 중, 입력 음성에 대해 확률이 가장 높은 단어열을 음성인식 결과로 출력한다. 영어의 경우, 동일한 문자열을 가진 단어에 대해서 대문자로 표기된 경우와 소문자만으로 표기된 두 경우 같은 발음을 가지게 되므로, 입력 음성에 대해서 발성된 단어의 대/소문자화 여부를 판별할 수 없게 되며, Standard Normalised Orthographical Representation(SNOR) [1]로 알려져 있는 표준 음성인식기 출력 포맷 또한 대소문자 구분에 대한 정보를 포함하지 않고 있다. 따라서, 자동으로 대소문자가 구분 된다면 영어 음성인식 결과의 가독성은 크게 향상된다.

받아쓰기(dictation)가 수행되는 경우, 받아쓰기 시스템은 화자에게 필요한 시점에 “대문자로”, “소문자로” 등을 발성하게 함으로서 대소문자를 구분할 수 있다. 그러나 뉴스 자료에 대한 음성인식과 같이 화자가 자신의 발성이 자동으로 음성인식 되고 있다는 것을 모르는 경우에는, 이와 같은 발성된 대소문자 관련 정보는 얻을 수가 없다. 특히 음성 자료의 입력의 경우 음성인식 오류로 인하여 자동 대소문자 생성은 더욱 어려운 문제가 되게 된다.

본 논문에서는 영어 음성인식 결과의 가독성을 향상시키기 위해서 자동으로 대소문자를 생성하는 방법을 제안하며, 뉴스 음성 자료에 대해서 본 방법의 유효성을 검증한다. 대소문자의 자동 생성을 위하여, 기존의 음성인식기에서 수정된 음성인식기가 기술된다. 이 수정된 음성 인식기에서는 어휘(vocabulary) 내의 모든 단어들이 복사가 되는데, 한번은 소문자 형태로, 그리고 다른 한번은 대문자 형태로 복사가 된다. 또한 언어 모델이 대소문자 구분이 된 텍스트 상에서 재학습된다.

본 논문은 6개의 장으로 구성되어 있다. 2장에서는 기존의 연구를 소개한다. 3장에서는 기존의 음성인식기에 수정을 가하여 음성인식 결과와 함께 대소문자 구분이 가능한 방법을 제시한다. 4장에서는 뉴스자료를 이용한 실험 방법과 결과 측정 방법을 설명한다. 5장에서는 HTK 방송뉴스 전사(broadcast news transcription) 시스템을 이용해서 3시간 분량의 뉴스 데이터에 대해서 수행된 자동 대소문자 구분 결과를 보여준다. 마지막으로 6장에서는 본 논문에 대한 결론을 내린다.

2. 기존의 연구

자동 대소문자를 생성하는 방법들이 워드프로세서에서 상용화 되어 제공되고 있다. 워드프로세서의 문법 및 철자 확인기들은 단어별 대소문자에 대한 추천을 제공한다. 대표적인 예가 가장 널리 사용되는 워드프로세서인 마이크로소프트사의

워드에서 제공하는 대소문자 생성기이다. 워드의 대소문자 생성기는 미국 특허로서 [2]에 자세히 기술되어 있다. 워드에서는 현재의 단어가 문장의 처음에 위치하는지의 여부를 sentence capitalisation state machine에 의해서 식별된다. 하나의 단어는 단어에 대한 문자열들과 주변에 위치한 구두점으로 정의된다. Sentence capitalisation state machine은 state간의 천이 가능 여부를 판단하기 위해서 현재 단어의 문자열들을 사용한다. 예를 들면, state machine이 문장 끝 구두점을 통과하면, capitalisation state machine은 state를 end punctuation state로 천이한다. 단어의 문자열들이 capitalisation state machine을 통과하면서, 특정 단어가 문장의 끝에 놓였는지를 결정하게 되고, 이어서 새 문장이 시작되는 위치에 놓인 다음 단어가 대문자로서 시작해야 하는지를 결정하게 된다. 문장의 처음에서 사용되지 않은 단어들의 대소문자 구분은 사전을 이용해서 결정한다. 즉, 단어의 가장 빈도가 높은 대소문자 타입으로 대소문자 타입이 결정 된다.

대문자화가 예상되는 위치에 놓인 대문자화된 단어(예: 문장의 첫 번째 단어)는 고유명사 일수도 있고 또는 일반 단어가 문장에서의 위치 때문에 대문자화되어 있을 수 있다. 대문자화된 단어들이 고유 명사이기에 대문자화가 되었는지, 혹은 문장의 첫 번째에 위치했기 때문에 대문자화되었는지를 구분하는 방법이 [3]에서 기술되었다. 이 방법에서는 문서 전체를 스캔해서 대문자화가 예상되는 위치에 놓인 대문자화된 단어의 의미를 분석하고 고유 명사 여부를 판별했다.

많은 수의 대문자화된 단어들이 문장의 첫 번째 단어 또는 Named Entity(NE. NE는 인명, 지명, 기관명 등을 포함함) [4]이기 때문에, NE 인식과 음성인식기의 출력의 가독성 향상과는 밀접한 관련이 있다. 생성된 구두점과 대소문자는 NE 인식에 있어서 유용한 단서를 제공한다. [5]에서의 NE 인식 실험에서, 구두점과 대소문자 정보가 제공될 때 NE 인식의 성능이 크게 향상되는 것을 보여주었다.

자동 대소문자 인식에 있어서의 NE의 중요성은 [6]에서 언급되었다. 이 연구에서는 NE 태그가 포함된 언어모델을 이용해서, NE를 인식하고, 인식된 NE는 대문자로, 그 외의 단어들은 소문자로 음성인식기의 결과를 제공함으로써 대소문자를 자동으로 생성했다. 그러나 대문자화된 단어들은 크게 문장의 첫 단어와 NE로 나누어지기 때문에, NE 인식만으로는 대소문자가 충분히 구분되지 않는다. 따라서 자동으로 문장 경계를 정확히 인식 할 수 있다면, 자동 문장 경계 인식과 NE 인식을 이용하여 자동으로 대소문자를 생성하는 것도 가능하게 된다.

3. 자동 대소문자 생성

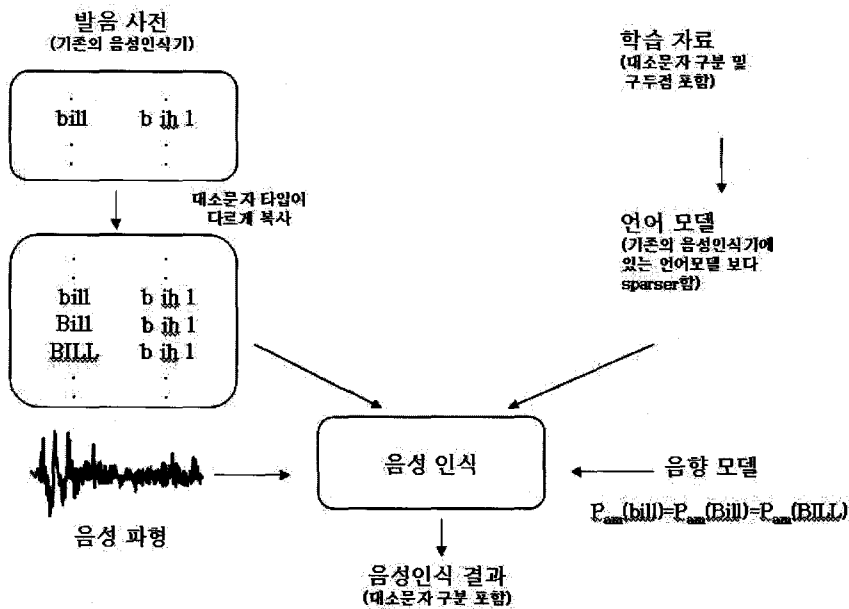
음성인식은 입력된 음성 신호에 대해서 최적의 단어 열을 찾는 것에 초점을 맞추기 때문에, 일반적으로 음성인식 시스템은 단어의 대소문자 정보를 식별하지

않는다. 따라서, 기존의 음성인식 시스템에서 어휘에 등록된 단어들과 발음사전에 등록된 단어들이 대소문자가 구분되어 등록되어 있지 않다. 이와 더불어, 기존의 음성인식 시스템에서는 언어 모델을 대소문자가 구분된 텍스트에서 학습 시키지 않는다.

본 3장에서 기술되는 자동 대소문자 생성방법은 기존의 음성인식기에 수정을 가하여 음성인식 결과와 대소문자 생성이 동시에 진행되는 방법이다. 이를 위해 다음 세 가지를 수정 한다:

1. 어휘에 있는 모든 단어들을 같은 문자열을 유지한 채, 모든 문자가 대문자인 경우(예: BILL), 모든 문자가 소문자인 경우(예: bill), 첫 번째 문자만 대문자인 경우(예: Bill)의 세 가지의 대소문자 타입으로 복사를 한다.(대소문자 타입의 자세한 사항에 대해서는 4장 및 <표 2> 참조)
2. 발음 사전에 있는 모든 단어들에 대해서 어휘의 단어들을 복사한 것과 같은 방식으로 복사를 한다. 단, 같은 문자열을 가지는 복사된 단어들은 같은 발음을 가진다.
3. 언어모델을 대소문자가 구분된 텍스트 상에서 재학습 시킨다.

<그림 1>은 본 논문에서 제안한 대소문자 생성 시스템의 개요를 보여준다. 하나의 단어로부터 복사된 단어들이 같은 발음을 가지고 있기 때문에, 복사된 단어



<그림 1> 제안한 대소문자 생성 방법의 개요도. 음성인식과 대소문자 생성이 동시에 수행됨

들은 같은 acoustic 모델 score를 가지게 된다. 그러나, 복사된 단어별로 다른 언어 모델 score를 가지기 때문에, hypothesis상에 있는 단어들은 각자 다른 대소문자 타입을 가지게 된다. 음성인식의 decoding이 수행되고, 대소문자 구분을 포함하는 최적의 hypothesis가 생성된다.

문장 경계 정보가 문장 처음에 위치한 단어를 대문자로 표기하는데 필요하기 때문에, 대소문자 생성 시스템은 구두점 생성을 하기 위하여, 발음 사전에 구두점들의 발음을 silence로 등록하고, 언어모델은 구두점을 포함하는 대소문자 구분이 된 텍스트에 대해서 훈련한다. 따라서, 제안한 방법에 의해, 음성인식 결과와 대소문자 구분, 구두점 생성이 동시에 진행된다.

4. 실험 방법 및 결과 측정 방법

뉴스 텍스트 자료와 100시간 분량의 1998 Hub-4 방송뉴스 자료가 학습 자료로서 사용되었다. 뉴스 텍스트 자료는 (이하 BNText92_97) 1992년에서 1997년 사이에 수집된 1.84억 단어로 구성된 뉴스 텍스트 자료이다. 100시간 분량의 1998 Hub-4 방송뉴스 자료는 (이하 BNAcoustic98) 뉴스에 대한 acoustic 자료와 이에 대한 전사로 구성되어 있다.

NIST 1998 Hub-4 방송뉴스 벤치마크 테스트에서 사용된 테스트 자료가 (이하 TestBNAcoustic98) 본 연구에서 테스트 자료로 사용되었다. TestBNAcoustic98은 3시간 분량의 acoustic 자료와 이에 대한 전사로 구성되어 있다. 학습 자료와 테스트 자료는 <표 1>에 정리되어 있다.

4-gram 언어모델은 BNText92_97과 BNAcoustic98에 대해서 각기 생성한 언어 모델을 perplexity 최소화 방법을 사용해서 interpolation해서 생성했다.

대소문자 타입은 모든 문자가 대문자화된 경우와 모든 문자가 다 소문자가 된 경우, 또는 단어의 첫 번째 문자만 대문자화된 경우의 세 가지로 분류했다. 이러한 분류에 대한 세부 사항은 <표 2>에 기술되어 있다. 대문자화된 한 문자로 된 단어, 예를 들면 B. B. C.에서의 initial들은 All_Cap으로 분류 되었다. <표 2>에 있는 카테고리들 중 어느 하나로 분류되지 않는 단어들이 간혹 있다. 이러한 예외 단어들은 BNAcoustic98에서 437단어, TestBNAcoustic98에서 26단어가 존재한다. 예외 단어들의 대부분은 McWethy, MacLaine와 같은 사람 이름의 성이다. BNAcoustic98과 TestBNAcoustic98에서의 예외적인 경우 모두를 수동으로 살펴보았는데, 예외단어들의 대소문자 타입을 Fst_Cap으로 처리했을 때 문제가 되는 경우가 없었다. 따라서 예외적인 모든 경우들은 Fst_Cap으로 분류되었다. <표 3>은 BNAcoustic98과 TestBNAcoustic98에서 문장에서 단어가 놓인 위치에 따라 단어의 대소문자 타입들에 대한 출현 빈도를 나타내고 있다. <표 4>는 BNAcoustic98과

TestBNAcoustic98의 통계치를 나타내고 있다.

<표 1> 학습 및 테스트 자료

자료명	내용	단어수	용도	Acoustic 자료 유무
BNtext92_97	1992_97 뉴스 텍스트	184M	학습자료	무
BNAcoustic98	100시간 분량의 1998 Hub-4 자료	774K	학습자료	유
TestBNAcoustic98	1998 벤치마크 테스트 자료	32K	테스트 자료	유

<표 2> 대소문자 타입의 분류

대소문자 타입	설명
No_Cap	단어내 모든 문자가 소문자
All_Cap	단어내 모든 문자가 대문자
Fst_Cap	첫 번째 문자만 대문자

평가 척도로는 F-measure [7]와 Slot Error Rate(SER) [7]의 척도를 이용해서 평가하게 되는데, F-measure를 구하기 위해서 필요한 Precision(P)과 Recall(R), 그리고 F-measure(F)와 SER의 정의는 다음과 같다.

$$P = \frac{\text{정확한 대소문자 타입의 단어수}}{\text{hypothesis의 단어수}} \quad (1)$$

$$R = \frac{\text{정확한 대소문자 타입의 단어수}}{\text{reference의 단어수}} \quad (2)$$

$$F = \frac{RP}{(R+P)/2} \quad (3)$$

$$SER = \frac{\text{부정확한 대소문자타입의 단어수}}{\text{reference의 단어수}} \quad (4)$$

<표 3> BNAcoustic98과 TestBNAcoustic에서 문장에서의 단어의 위치에 따른 대소문자 타입별 빈도수

Word 타입			BNAcoustic98		TestBNAcoustic98	
NE 여부	대소문자 타입	예제	문장의 첫 단어	문장의 첫 단어 아닌 경우	문장의 첫 단어	문장 첫 단어 아닌 경우
NE	No_Cap	far east	16	12,110	0	615
NE	All_Cap	B. B. C.	536	10,535	20	577
NE	Fst_Cap	Clinton	3,529	43,459	143	1,790
non-NE	No_Cap	sentence	1,587	638,477	24	26,134
non-NE	All_Cap	C. E. O.	2,842	6,887	83	141
non-NE	Fst_Cap	American	37,659	16,256	1,603	465

<표 4> BNAcoustic98과 TestBNAcoustic98의 통계

타입	빈도수	
	BNAcoustic98	TestBNAcoustic98
단어(any type)	773,893	31,595
대문자화 된 단어	121,703	4,822
NE 단어	70,230	3,149
한 문자로 된 단어(NE)	10,200	543
한 문자로 된 단어(non-NE)	2,099	78
문장	46,169	1,873

5. 자동 대소문자 생성 결과

실시간의 10배 이하에서 작동하는 (under 10 times real time. 이하 10xRT) 뉴스 자료에 대한 HTK 시스템 [8]을 자동 대소문자 생성에 이용했다. 이 시스템의 가장 첫 번째 단계는 segmentation 단계인데, 이 단계에서 연속되어지는 뉴스자료 stream 이 하나의 화자가 하나의 audio type(예: wide-band, narrow-band 등)으로 이야기 하는 segment들로 분리가 된다.

HTK 시스템에서 음성인식은 two pass로 진행이 되는데, 각 pass에서는 cross-word triphone을 기반으로 한 decision-tree state cluster된 HMM과 N-gram 언어 모델을 사용한다. 첫 번째 pass에서는 gender-independent(그러나 bandwidth에는

specific한) HMM과 trigram 언어모델을 이용해서 각 segment에 대해서 초기 전사를 생성한다. 그 후 gender-dependent HMM을 이용해서 각 segment별로 화자의 성별이 결정된다. 그 후 각 segment 별로 unsupervised Maximum Likelihood Linear Regression(MLLR) [9] 변환을 첫 번째 pass에서 만들어진 초기 전사와 gender-dependent HMM을 이용해서 수행한다. 변환된 HMM과 4-gram 언어모델을 이용해서 두 번째 pass에서 최종 음성인식 결과를 생성한다.

뉴스자료에 대한 HTK 시스템의 구현에 대한 세부사항은 [10][11]에 자세히 기술되어 있고, 10xRT HTK 시스템의 구현에 대한 세부사항은 [12]에 자세히 기술되어 있다. HTK 시스템의 속도를 높이기 위해서 10xRT 시스템에서는 단순화된 음향 모델과 단순화된 decoding 방법을 이용했다.

본 연구에서 사용한 10xRT HTK 시스템은 1998 Hub-4 방송뉴스 벤치마크 테스트에서 사용된 10xRT 시스템과 category-based 언어모델 [13] 사용 여부, 언어 모델 학습 자료 양, 어휘 크기에서 차이가 있다.

대소문자를 자동으로 생성하기 위해서 3장에서 기술된 것과 같은 방법으로 HTK 시스템의 발음 사전에 등록된 모든 단어들을 세 개의 각기 다른 대소문자 타입(All_Cap, Fst_Cap, No_Cap)으로 복사 했고, 언어모델을 대소문자 정보를 가지고 있는 BNText92_97과 BNAcoustic98에서 학습했다.

<표 5>는 수정한 HTK 시스템을 이용하여 TestBNAcoustic98에 대해서 대소문자 생성을 수행한 결과를 보여준다.

<표 5> TestBNAcoustic98에 대해 HTK 시스템을 이용한 대소문자 생성 결과(WER'은 음성인식 결과에서 대소문자 구분을 없애고, 구두점을 없앤 후 측정된 WER임)

측정지표	WER(%)	WER'(%)	Precision	Recall	F-measure	SER(%)
측정치	22.97	17.27	0.7736	0.6942	0.7317	48.55

제안한 방법은 자동 대소문자 생성에 대해 F-measure로 0.7317의 결과를 보였는데, 이때 SER은 48.55%, precision은 0.7736, 그리고 recall은 0.6942를 기록했다.

제안한 방법은 대소문자 생성이 음성인식 결과의 출력과 동시에 진행되므로 전체 시스템의 음성인식율도 전체 시스템 성능 측정에 있어서 중요한 요소가 된다. 대소문자 타입이 다른 경우 동일 문자열을 가진 단어들도 다른 단어로 구분하고, 구두점들을 각기 하나의 단어로 간주하고 시스템의 단어오류율을 측정하면 22.97%가 측정된다.

기존의 음성인식기는 구두점 및 대소문자의 구분 없이 최적의 단어열만을 출력하므로, 동일 조건에서 비교하기 위해서 제안한 시스템의 출력 결과에서 대소문

자 구분과 구두점을 없앤 후 음성인식율을 측정했다. 대소문자 구분과 구두점을 없앤 후의 단어오류율을 WER'이라 했을 때, WER'은 17.27%로 측정되었다. WER과 WER'간에는 5.7%의 차이가 존재하는데, 이는 대소문자 구분과 구두점이 포함되어 어휘 크기가 증가 했고, 이로 인해 언어 모델이 sparse 해졌기 때문이다.

HTK 10xRT 시스템을 이용해서 음성인식을 TestBNAcoustic98에 수행했을 때 단어오류율은 16.7%로 측정되었다. 동일 조건(대소문자 구분 및 구두점 없음)에 대해서 0.57%의 차이가 발생하는데, 이 차이는 문장의 첫 단어들에 의한 언어 모델의 왜곡에 의한 것으로 판단된다.

많은 경우에 있어서 문장의 첫 단어들은 NE가 아니다. 문장의 첫 단어로 사용된 단어들의 대부분은 이들 단어들이 문장의 중간에 사용되는 경우 대문자화되지 않는다. TestBNAcoustic98에는 1,873개의 문장이 있으므로, TestBNAcoustic98에 있는 문장 당 평균 단어의 수는 16.9개가 된다. 이들 문장의 첫 번째 단어들 중에서, 검사 결과 91.3%의 단어들이 NE가 아니었다. 제안한 방법에서는 같은 문자열로 구성된 단어도 대소문자 타입이 다른 경우 다른 단어로 인식이 되는데, 일반적으로 대문자화되지 않지만, 문장의 처음에 위치하는 이유로 대문자화된 단어들에 의해서 전체 단어중 약 5.4% $((1/16.9)*0.913)$ 의 단어들이 다른 단어로 인식되게 되며, 이들 단어로 인하여 언어모델이 왜곡된다.

본 논문에서 제안한 방법은 어휘의 수가 3배로 증가하므로 탐색 시간이 늘어나게 되는 단점이 있다. 그러나 복사된 단어들은 동일한 발음열을 가지게 되므로, lexical tree 상에서는 state를 공유하게 되고, 단지 final node에서만 각각의 대소문자 타입별로 별도의 단어로써 branch가 일어나게 된다. 언어모델 적용시에는 코퍼스에서 나오지 않은(대소문자 구분 기준으로) 단어들에 대해서는 언어모델 확률 값이 생성되지 않기 때문에 history에 누적되지 않게 되어 탐색 시간이 늘어나지만, 이러한 탐색에 있어서의 오버헤드가 큰 문제를 야기하지는 않는다.

6. 결 론

본 연구에서는 영어 음성인식 결과의 가독성을 향상시키기 위해서 음성인식기의 수정을 이용한 자동 대소문자를 생성하는 방법을 제안했다. 본 방법을 이용해 음성인식 결과와 이에 대한 대소문자 구분 결과를 동시에 생성하게 되어, 기존의 음성인식기가 대소문자 구분을 하지 못한 문제점에 대해 하나의 해결 방법을 제시했다.

이 수정된 음성 인식기에서는 어휘내의 모든 단어들이 복사가 되는데, 한번은 소문자 형태로, 그리고 다른 한번은 대문자 형태들로 복사가 된다. 또한 언어 모델이 대소문자 구분이 된 텍스트 상에서 재학습 된다.

제안한 시스템의 성능을 평가하기 위해서 자동 대소문자 생성의 실험이 수정된 HTK 방송뉴스 전사 시스템을 이용해서 3시간 분량의 뉴스 데이터에 대해서 수행되었다. 제안한 방법은 자동 대소문자 생성에 대해 F-measure로 0.7317의 결과를 보였는데, 이때 SER은 48.55%, precision은 0.7736, 그리고 recall은 0.6942를 기록했다.

어휘의 크기가 커짐에 따라 언어모델이 sparse 해지고, 문장 처음에 위치하는 NE가 아닌 단어에 의한 언어모델 왜곡으로 인하여 WER이 0.57% 증가하게 된다. 또한 어휘 크기의 증가로 탐색 시간 늘어나는 단점이 있다.

대문자화된 단어들은 크게 문장의 첫 단어와 NE로 나누어지기 때문에, 자동 문장 경계 인식과 NE 인식을 이용하여 자동 구두점을 생성하는 또 다른 방법을 구현할 수 있다. 이 방법은 어휘의 크기가 증가하지 않기 때문에, 어휘의 크기 증가로 인한 WER의 증가와 탐색 시간 증가 문제를 해결할 수 있으며, 위의 언어모델 왜곡 문제도 해결할 수 있을 것으로 기대된다.

추후 자동 문장 경계 인식과 NE 인식을 이용한 자동 구두점 생성 방식을 연구하여, 본 논문의 방법과 비교를 하고자 한다.

참 고 문 헌

- [1] "NIST December 1998 Hub-4e information extraction(named entities) broadcast news benchmark test results". Available at ftp://jaguar.ncsl.nist.gov/csr98/h4iene_98_official_scores_990107/index.htm.
- [2] S. Rayson, D. Hachamovitch, et. al., "Autocorrecting text typed into a word processing document", U.S. patent 5761689, 1998. Available at <http://www.delphion.com>.
- [3] A. Mikheev, "A knowledge-free method for capitalized word disambiguation", *Proc. Annual Meeting of the Association for Computational Linguistics*, pp. 159-166, 1999.
- [4] "MUC6 named entity task definition, version 2.1", *Proc. 6th Message Understanding Conference*, pp. 317-332, 1995.
- [5] F. Kubala, R. Schwartz, et. al., "Named entity extraction from speech", *Proc. Broadcast News Transcription and Understanding Workshop*, pp. 287-292, 1998.
- [6] Y. Gotoh, S. Renals, G. Williams, "Named entity tagged language models", *Proc. ICASSP*, pp. 513-516, 1999.
- [7] J. Makhoul, F. Kubala, et. al., "Performance measures for information extraction", *Proc. DARPA Broadcast News Workshop*, pp. 249-252, 1999.
- [8] P. Woodland, "The development of the HTK broadcast news transcription system: An overview", *Speech Communication*, Vol. 37, Nos. 1-2, pp. 47-67, 2002.
- [9] C. Leggetter, P. Woodland, "Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models", *Computer Speech and Language*, Vol. 9, pp. 171-185, 1995.

- [10] P. Woodland, T. Hain, et. al., "The 1997 HTK broadcast news transcription system", *Proc. Broadcast News Transcription and Understanding Workshop*, 1998.
- [11] P. Woodland, T. Hain, et. al. "The 1998 HTK broadcast news transcription system: Development and results", *Proc. DARPA Broadcast News Workshop*, pp. 265-270, 1999.
- [12] J. Odell, P. Woodland, T. Hain, "The CUHTK-Entropic 10xRT broadcast news transcription system", *Proc. DARPA Broadcast News Workshop*, pp. 271-275, 1999.
- [13] T. Niesler, E. Whittaker, P. Woodland, "Comparison of part-of-speech and automatically derived category-based language models for speech recognition", *Proc. ICASSP*, pp. 177-180, 1998.

접수일자: 2007년 2월 22일

게재결정: 2007년 3월 17일

▶ 김지환(Ji-Hwan Kim)

주소: 137-724 서울시 서초구 우면동 16번지

소속: LG전자기술원 정보기술연구소 IST 그룹

전화: 02) 526-4164

E-mail: kimjihwan@lge.com