

운율 경계강도 예측을 위한 OC1의 적용 및 CART와의 비교

The Comparison of OC1 and CART for Prosodic Boundary Index Prediction

임 동 식*, 김 진 영*, 김 선 미**

(Dong Sik Lim*, Jin Young Kim*, Sun Mi Kim**)

* 본 논문은 '97 학제간 연구의 지원에 의해 이루어진 연구 결과물 중 하나입니다.

요 약

본 논문은 연속음 인식과 합성을 위한 운율 경계강도 예측모델을 위해 최근에 널리 사용되고 있는 방법으로 분류·회귀트리라 불리는 CART(Classification And Regression Tree)와 OC1(Oblique Classifier1)을 적용한다. 운율 경계강도 수준을 4로 하고 문법적인 특징으로는 트리구조 방법으로 결정된 오른쪽 가지의 수식의 깊이(Rd)와 link grammar 방법으로 결정된 연결거리(To_Right)를 tri_gram모형과 결합하여 CART와 OC1에 적용해 각각 운율 경계강도를 예측, 비교한다. 실험을 통하여 OC1 방법이 CART 방법에 비해 더 적은 터미널 노드에 더 향상된 예측율을 보임을 확인할 수 있다.

ABSTRACT

In this paper, we apply CART(Classification And Regression tree) and OC1(Oblique Classifier1) which methods are widely used for continuous speech recognition and synthesis. We predict prosodic boundary index by applying CART and OC1, which combine right depth of tree-structured method and To_Right of link grammar method with tri_gram model. We assigned four prosodic boundary index level from 0 to 3. Experimental results show that OC1 method is superior to CART method. In other words, in spite of OC1's having fewer nodes than CART, it can make more improved prediction than CART.

I. 서 론

운율 경계강도라고 하는 것은 인간이 주어진 문장을 발음함에 있어서, 문장을 몇 개의 크고 작은 분절(도막)로 나누어 발음을 하게 되는데, 이때 결정되는 분절의 강도를 의미하는 것이다. 운율 경계강도에 따라서 휴지기의 유무와 길이, 그리고 분절음의 길이 및 피치 현상이 고유하게 나타나기 때문에 운율 경계강도를 잘 예측하는 것은 매우 중요한 일이다[1]. 종속 문법적인 해석의 한 방법으로 Hunt는 구문론적인 표현이 간단한 link grammar를 이용하였다. 이는 문법적으로 관련있는 문장성분을 연결하고 그 관계를 설명하는 레이블을 부여하였다[2]. 이 경우 link에 의해 표현된 표면-문법 관계는 문장 내 운율 경계강도를 가지고 운율구를 이룬다고 보고하고 있다. 구문론적인 정보를 보다 정확히 관찰하기 위하여 순수한 문장성분 레이

블링 방법도 시도되었다[3]. 운율 경계강도를 예측하기 위한 방법으로서 지금까지 영어권에서 많이 사용되어온 방법은 은닉마코프모델(HMM)과 신경회로망이 많이 사용되었으며, 최근 국내에서도 이러한 방법을 사용한 경계강도 예측에 관한 몇몇 보고가 있었다. 한편, 위와 같은 방법 이외에도 최근에 널리 사용되고 있는 방법이 바로 CART와 OC1 방법이다. CART는 신경회로망의 복잡한 학습 및 구조를 탈피하고자 개발되어온 방법으로서 패턴인식과 회귀분석에서 매우 유용하게 사용되어지고 있는 방법이다. 그리고, OC1방법은 CART방법보다 좀 더 다양한 분할이 가능하며 각각의 내부 노드에서 하나 또는 그 이상의 예측변수의 선형결합(linear combination)을 포함하는 결정트리를 구축하게 된다. 또한, 결정트리에서 생성된 각 노드를 관측심볼(observation sequence)로, 경계강도 지수값을 상태(state)의 개념으로 접근해 확률적 모델인 마코프(Markov)모델에 적용시킴으로써 경계강도 예측율이 향상됨을 확인할 수 있었다.

* 전남대학교 전자공학과

** 서울대학교 뉴미디어통신 공동연구소 PostDoc

접수일자: 1999년 1월 18일

II. CART 와 OCI 방법

CART는 크기의 순서가 정해져 있는 ordered 변수 외에도 categorical 변수에 대해서도 패턴인식 및 회귀분석을 적절히 수행할 수 있다. CART의 경우 변수가 연속형(continuous), 다시 말해서 실 변수(real-valued variable)인 경우이면 $x \leq c$ 형태의 질문물(x:특징변수), 이산형(discrete)인 경우에는 x가 가능한 모든 종류 부분집합에 속하는 경우를 고려하는 질문물 하게 되는 반면, OCI 방법은 각 노드에서 최적의 축 평행 분할(axis parallel split)과 사선 분할(oblique split)을 평가해서 그 둘 중 더 나은 하나를 선택하게 된다.

그리고, 본 논문의 OCI에서 국부 최소값(Local minimum)을 찾았을 때 그 국부 최소값이 전역최소값(Global minimum)에 해당하는지를 확인하는 random jump의 최대 수는 5회, search의 시작위치를 바꾸어 보는 재시작(restart)회수는 20회로 설정을 했다. OCI의 이러한 요인들은 CART에 비해 더 정확한 분류를 가능하게 한다[5].

일반적으로, OCI은 각 내부노드에서 이루어지는 질문의 형태가 선형 조합(linear combination)의 형태이며, CART는 하나의 변수 또는 OCI의 경우와 같이 선형조합(CART-LC)모두가 가능하다.

그러나, CART-LC는 OCI에 비해 약 2가지 한계점이 있는데 그것은 다음과 같다[5].

- 1) CART-LC는 deterministic한 특성때문에 split의 impurity가 시간에 따라 감소하지 않는다.
- 2) CART-LC는 local minima를 벗어나기 위한 내장 메카니즘이 없기 때문에 local minima를 벗어나기 위해 split의 impurity 증가를 종종 허용한다.

본 논문에서의 각 노드에 적용된 OCI의 알고리즘을 살펴보면 다음과 같다[5].

- 1) 특징벡터 공간으로부터 최적의 축 평행 분할 (axis-parallel split)를 찾는다. I를 어 최적 split의 불순도(impurity)라 한다.
- 2) random split H를 선택한다. 초기치로 H를 최적의 axis-parallel split로 가정하고 불순도가 더 이상 감소하지 않을 때까지 H의 계수들에 퍼터베이션 수행.
- 3) H의 불순도를 i라 할 때 만약 $i < 1$ 면 $I = 1$ 로 놓는다. 이 과정을 몇 회 반복.
- 4) I에 상당하는 split를 출력한다.

III. 실험 및 결과

1. 운율 경계강도의 측정 및 파라미터

본 연구에서는 간단한 청취실험을 통하여 운율 경계강도를 예측하였다. 음성DB로서는 녹음된 문어체 문장 150문장을 들려주고 청취자들이 0~3 사이의 값 중에서 하나

를 결정하도록 하였다. 운율 경계강도 예측을 위한 구문분석은 통계적인 문장분석을 위해 그래프의 형태가 아닌 적당한 수치로 변형되어야 한다. 본 논문에서는 문장구조를 수치적인 형태로 표현하기 위하여 다음과 같은 파라미터를 정의하여 사용하였다.

Parse : 각 어절의 품사정보, 즉 주어,동사...등.

Rd : 주어진 어절에 대해서 자기 자신 어절의 Depth와 다음 어절의 Depth의 차이에 1의 값을 더해 준 값을 의미. 예를 들어, '아름다운 꽃'이라고 한다면, '아름다운'과 '꽃'사이의 Rd는 값이 0이고 이는 아름다운이 꽃을 수식하고 있음을 의미한다.

To_Right : 연결고리 안에 나타나는 단어의 수.

즉, 문장성분 연결에서 어느 단어가 연결하는 후행하는 다른 단어 사이에 있는 단어 경계의 수. 다음의 표1은 운율 경계강도 예측을 위해 본 논문에서 적용된 파라미터 분석 예이다.

표 1. 문장 1에 대한 파라미터 분석 예
Table 1. Parameter Analysis of Sentence 1.

문장1	Word	depth	Rd	To_Right	PBI
0	생일이	3	5	6	2
1	한	7	0	0	0
2	생명이	6	0	0	1
3	태어난	5	0	0	0
4	것을	4	1	1	0
5	기념하고	4	0	0	1
6	축하하는	3	0	0	0
7	날이라는	2	0	0	0
8	것은	1	2	3	3
9	동서고급을	2	0	0	0
10	막은하고	1	1	1	2
11	모두	1	0	0	1
12	같다	0			

2. CART에 의한 운율 경계강도 예측

품사정보를 자신의 어절과 그 앞·뒤 어절을 독립변수로 적용하는 tri_gram과 구문분석 파라미터 Rd 와 To_Right를 CART에 적용하여 경계강도를 예측하였다.

위의 경우에는 전체 예러율은 24.7%이었으며, 전체 터미널 노드의 개수는 89개였다. CART분석을 통하여 얻어지는 각 파라미터의 중요성에서는 운율 경계강도를 결정하는 파라미터들은 구문론적 파라미터이었으며, 특히, 연결고리 안에 나타나는 단어의 수(To_Right)가 운율 경계강도를 예측하는 데 매우 중요하다는 사실을 알 수 있었다. 그 다음으로는 Rd 파라미터와 각 어절의 품사정보를 나타내고 있는 tri_gram 단위의 Parse 파라미터의 순서로 트리를 생성하는데 중요하게 받아들여졌다.

표 2. 특징 파라미터에 의한 운율 경계강도 분류
Table 2. Results of PBI Classification by features.

관측 1 예측	0	1	2	3
0	872	196	14	16
1	84	242	48	21
2	10	25	174	34
3	1	1	17	134

다음의 그림1은 표2 에 대해 CART를 사용한 운율 경계강도 트리의 일부분을 보여 주는 예이다.

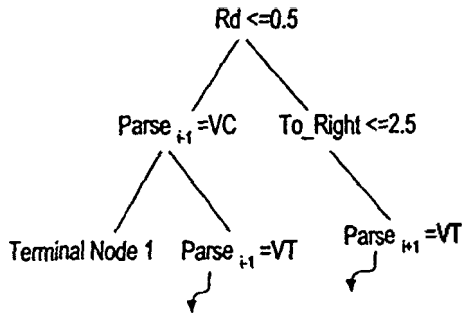


그림 1. CART에 의해 생성된 트리의 예
Fig. 1. Example of Generated tree by CART.

아래의 표3은 표2에 대해 확률적 모델인 Markov모델에 적용시킨 결과이다. 이는 학습시킨 트리에서 생성된 각 터미널 노드에서의 각각의 운율 경계강도(0,1,2,3)를 상태(states)로 보고 각 상태에서의 각 관측심볼(node)이 발생할 최적의 확률모델에 적용시킨 것이다.

$$\arg \text{Max}_{b_i, \dots, b_{N-1}} P(b_i, \dots, b_{N-1} | \text{Node}_{1, \dots, N}, \text{관측 } b_i, \dots, b_{N-1})$$

운율 경계강도는 관측 b_i, \dots, b_{N-1} 의 경계강도 열과 특징 변수인 Rd, To_Right, tri_gram 을 참조해 CART가 생성한 노드들 $\text{Node}_{1, \dots, N-1}$ 에 대하여 최대 확률을 보이는 강도들 구하는 문제이다. 이는 학습에 의해 생성된 노드와 관측치 각각의 운율 경계강도가 쌍을 이루어 운율 경계강도의 조건부 확률을 최대화 시키는 비터비 알고리즘에 적용된다[1].

표 3. Markov모델에 의한 운율 경계강도 분류
Table 3. PBI Classification by Markov Model.

관측 1 예측	0	1	2	3
0	1016	65	12	5
1	194	134	48	19
2	21	17	178	27
3	6	1	21	125

위의 경우에서 전체 에러율은 23.1%로 확률적 모델의 적용 전 보다 1.6% 더 나은 예측율을 보였다.

경계강도 0에 대한 예측율은 훨씬 향상된 반면, 1에 대한 예측율은 오히려 더 혼란스러움을 가져왔다. 이는 칭취

테스트의 과정에서 생긴 오류라고 생각 된다. 지금까지 남성화자 150문장에 대해 실험을 한 결과였고 여기에 여성화자의 낭독재 100문장을 첨가해서 학습 과 실험을 해 보았다. 남성화자 100문장, 여성화자 50문장의 150문장으로 학습을 거친 후, 남성화자 50문장, 여성화자 50문장의 100문장에 대해 실험을 해 보았다.

표 4.1. CART 에 의한 학습결과
Table 4.1. Training Results by CART.

관측 1 예측	0	1	2	3
0	1013	241	21	7
1	78	188	60	16
2	23	60	155	55
3	2	12	26	128

위 학습은 150문장 2085어절에 대한 운율 경계 강도이고 터미널 노드는 26개, 전체 에러율은 28.8%를 보였다. 이는 CART가 생성해 낼 수 있는 트리 중 가장 최적인 최소 비용(Cost)인 경우이다.

표 4.2. 학습에 의한 실험결과
Table 4.2. Test Results by Training.

관측 1 예측	0	1	2	3
0	679	166	25	5
1	81	123	51	20
2	11	37	91	56
3	3	6	30	79

학습을 거친 후, 특징벡터가 추출된 상태를 가지고 1463 어절에 관련된 운율 경계강도 실험을 해 본 결과 전체 에러율은 33.6%를 보였다. 실험 결과에 Markov 모델을 적용시켜 본 결과 전체 에러율이 29.8%로 Markov 적용 전에 비해 약 4%의 에러를 감소를 보였다.

3. OC1에 의한 운율 경계강도 예측

CART에서 실험한 결과와의 비교를 위해 똑 같은 DB와 특징변수인 Rd 와 To_Right 으로 구성된 2개의 구분분석 파라미터, tri_gram 단위의 품사정보를 적용한다. OC1에 대해 실험을 해 본 결과는 표5와 같다. 표5는 10%의 가

표 5. 특징 파라미터에 의한 운율 경계강도 분류
Table 5. PBI Classification by features.

관측 1 예측	0	1	2	3
0	1041	52	4	1
1	120	255	17	3
2	22	39	177	5
3	7	5	27	114

치기(pruning)를 한 결과이며, 여기에서 터미널 노드는 29개이다. 또한 전체 에러율은 16.0%의 에러율을 보인다.

표2의 CART 방법과 비교해 볼 때 터미널 노드의 수가 60개가 줄었으며, 8.7% 가량의 예측을 향상시키고 있다.

여기에 역시 Markov 모델을 적용한 결과는 전체 에러율이 15.5%로 Markov 적용 전에 비해 약 0.5%, CART와 비교해 보면 7.6%의 예측을 향상을 볼 수 있다.

CART에서 적용한 남·여성 혼합 문장에 대한 실험을 OCI에 적용해 본 결과는 표6과 같다.

표 6.1. OCI에 의한 학습결과
Table 6.1. Training Results by OCI.

관측 \ 예측	0	1	2	3
0	1238	31	12	1
1	173	132	36	1
2	38	44	203	8
3	5	8	49	106

위 학습에서의 터미널 노드는 5개, 전체 에러율은 19.5%를 보였다.

표 6.2. 학습에 의한 실험결과
Table 6.2. Test Results by Training.

관측 \ 예측	0	1	2	3
0	812	44	14	5
1	166	67	34	8
2	29	40	101	25
3	7	7	50	54

위 실험결과에서의 전체 에러율은 29.3%이다. 즉, OCI은 CART방법에 비해 학습에서는 9.3%, 실험에서는 4.3%의 향상된 예측을 보였다. 또한, 터미널 노드 개수 또한 21개가 줄었다. 결론적으로, OCI은 CART에 비해 터미널 노드 개수가 적으면서도 더 향상된 예측의 효과를 볼 수 있음을 알 수 있다. 그러나, 경계강도 0,1,2,3에 대해서 전반적으로 에러율이 나아진 것은 아니다. 다만, 이런 결론은 전체 에러율의 결과만을 놓고 판단한 것이다. 위에서 기술한 실험 내용을 정리해 보면 다음과 같다.

CART와 OCI에 대한 비교는 그림2와 같다.

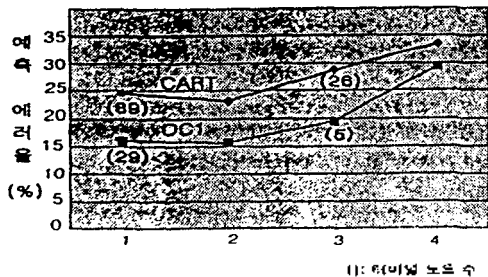


그림 2. CART 와 OCI의 비교
Fig 2. The Comparison of CART and OCI.

- 1 : tri_gram과 구문 파라미터를 적용
- 2 : 1에 대한 Markov 모델 적용
- 3 : 2명의 화자에 대한 학습
- 4 : 학습을 통한 실험결과

IV. 결 론

본 논문에서는 CART와 OCI방법을 이용하여 운율 경계강도를 예측해 보았다. 실험을 통하여 OCI방법이 CART방법에 비해 좀 더 향상된 예측을 보임을 알 수 있었다. 특징변수로는 품사정보(tri_gram단위)와 구문변수들을 이용하였다. 앞으로 인식률 향상에 좀 더 기여를 할 수 있는 특징변수 선별 등의 연구가 필요하겠고, 이를 이용하면 운율 합성부에 적용시킬 수 있으리라 생각된다.

참 고 문 헌

1. 전남대학교, 음성인식 및 합성을 위한 한국어 운율정보처리, 정보통신부 최종 연구개발 결과 보고서, 1998.
2. A.J. Hunt, Models of Prosody and Syntax and their Application to Automatic Speech Recognition, Ph.D thesis, Sydney University, 1995.
3. 김선미, 한국어의 리듬단위와 문법구조, 서울대학교 언어학과 박사학위논문, 1997.
4. L. Breiman, J.H. Friedman, R. A. Olshen, and C.J. Stone, Classification and Regression Trees, Wadsworth Statistics/Probability Series, Belmont, CA, 1984
5. Sreerama K.Murthy, On Growing Better Decision Trees from Data, the degree of Doctor of Philosophy, Baltimore, Maryland, 1995

▲ 임 동 식(Lim Dong Sik)



1998년 2월: 동신대학교 정보통신 공학과 졸업(공학사)
1998년 2월~현재: 전남대학교 대학원 전자공학과 석사 과정

※ 주관심분야: 음성인식, 합성

▲김진영(Kim Jin Young) 1962년 4월 26일생



1986년 2월: 서울대학교 전자공학과 (공학사)

1988년 2월: 서울대학교 전자공학과 (공학석사)

1994년 8월: 서울대학교 전자공학과 (공학박사)

1993년 3월~1994년 12월: 한국통신 소프트웨어연구소 전임 연구원

1995년~현재: 전남대학교 공과대학교 전자공학과 조교수

※주관심분야: 음성인식 및 음성합성, 멀티모달 MMI

▲김선미(Kim Sun Mee)



1986년 2월: 서울대학교 언어학과 (문학사)

1988년 2월: 서울대학교 언어학과 (문학석사)

1997년 2월: 서울대학교 언어학과 (문학박사)

1997년 6월~현재: 서울대학교 뉴미디어통신 공동연구소 PostDoc.

※주관심분야: 실험음성학, 한국어운음