

HMM기반 온라인 한글 인식에서의 구조적, 전역적 지식의 적용

◦조성정, 김진형
한국과학기술원 전산학과

Utilization of Structural and Global Shape Knowledge for HMM Based Korean Handwriting Recognition

◦Cho, Sung Jung , Jin H. Kim
Department of Computer Science, KAIST
E-mail: sjcho@ai.kaist.ac.kr, jkim@cs.kaist.ac.kr

요 약

본 논문은 은닉 마르코프 모델(hidden Markov model, 이하 HMM)을 이용한 온라인 한글 인식에 구조적, 전역적 지식을 적용하는 것에 대하여 논의한다. HMM 기반 인식 방법은 여러 연구에서 성공적으로 적용되었으나, 통계적 방법과 국소적 특징의 이용에 따른 한계점을 갖고 있다. 본 논문에서는 한글 기본획 단위의 구조적 분석으로 HMM의 통계적 방법의 한계를 보완하였으며, 누적각도, 자소 내 기본획별 지속 정보, 연결획의 Pen-Up 움직임의 비율 등의 전역적 특징을 이용하여 국소적 특징의 한계를 보완하였다. 이러한 구조적, 전역적 지식은 기존 HMM 인식 시스템[1,2]의 확률적 틀 안에 자소단위로 결합하였다. 실험결과 기존의 인식기[2]에 비하여 인식률이 94.8%에서 95.6%로 증가하여 14.5%의 오류를 감소할 수 있었다.

1. 개 요

온라인 문자인식은 전자펜으로 쓴 사람의 글씨를 컴퓨터로 인식하는 것을 말하며, 휴대성과 편의성의 장점을 지녀 지난 수십 년간 활발한 연구가 진행되고 있다. 이러한 연구들에 쓰인 방법론 중 특히 은닉 마르코프 모델 (HMM)이 성공적으로 적용되고 있다. HMM은 잘 정의된 수학적 이론을 배경으로 하며, 전이 확률과 관찰 확률의 이중 확률 구조로 온라인 문자 입력의 특성에 따른 획의 길이 변이와 모양 변이를 동시에 모델링 하기에 적합한 특성을 보인다[4,5]. 이런 장점으로 HMM은 일상적인 환경의 데이터에 대하여 한글과 영어 모두 80~90% 이상의 인식률을 보이고 있다 [1,2,3,6,7].

그러나, HMM 기반의 인식 방법은 통계적 방법의 한계와 국소적 특징 이용의 한계를 갖고 있다. 통계적 방법은 글자의 변이는 잘 흡수하지만, 글자간의 변별력이 약한 단점이 있다[2]. 그리고 국소적 특징은 간단한 처리로 구현이 가능하지만, 정보의 양이 부족하다는 단점이 있다. HMM의 통계적 방법을 보완하기 위한 연구로는 한글 인식에 구조적 지식을 적용한 [2]의 연구가 있다. 이 연구에서는 HMM을 기반으로 하는 한글 인식 시스템에 자소 내 필수 성분 검증, 쌍구분, 자소간 위치관계 분석을 적용하여 48%의 오류를 감소시켰다. HMM의 국소적 특징을 보완한 연구로는 [6,9]가 있다. [6]의 연구는 HMM이 생성한 알파벳을 동적 프로그래밍 매칭을 이용하여 검증하여 오류를 45%정도 줄였다. [9]의 연구에서는 점, 수평획, 하향획의 유형 등을 이용하여 영어 사건의 크기를 줄여 인식에 사용하여 오류를 줄인바 있다.

본 논문은 기존 HMM 기반 한글 인식 방법의 한계를 구조적, 전역적 지식으로 보완함을 목표로 한다. 구조적 지식으로는 자소를 이루는 기본획을 분석하여 자소 내와 자소사이의 여러 기하학적 특징을 검증하여 적용하였다. 전역적 지식으로는 자소의 전반적인 필기 경향과 특징을 나타내는 누적각도, 자소 내 기본획별 지속정보, 연결획에서의 Pen-Up 움직임 비율을 적용하였다. 이러한 지식은 HMM 기반 온라인 인식기의 확률적 틀 안에 자소단위로 결합된다.

2. 구조적 지식의 적용

한글에서 구조적 지식은 자소의 정의와 구성에 관련한 지식을 의미한다. 자소는 기본획의 조합으로 구성되며, 기본획에는 그림 1처럼

수직획, 수평획, 사선획, 역사선획 등의 직선획과 원형획이 있다. 구조적 분석은 자소에 대응하는 입력을 해당 자소를 구성하는 기본획 단위로 나눈 후, 기본획간의 기하학적 관계를 검증하여 이루어진다.

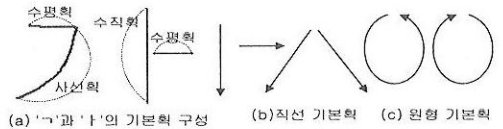


그림 1. 자소의 기본획 구성과 기본획 종류

2.1. 자소 구조 분석을 위한 기본획 추출

자소를 구성하는 기본획은 해당 자소의 훈련된 HMM 모델을 분석하여 추출할 수 있다. HMM 모델은 일반적으로 그림 2-(a)와 같은 left-to-right 모델이며, 이 경우 자체전이(self-transition) 부분이 안정적인 패턴인 기본획을 모델링한다. 그림 2-(b)는 'ㄱ' HMM 모델의 기본획 중 수평획이 상태 1의 자체전이에, 수직획이 상태 2의 자체전이에 대응됨을 보여준다 1). 방향이 일정한 직선 기본획은 하나의 자체전이에 대응이 되며, 방향이 계속 변하는 원형획은 여러 자체전이에 대응이 된다. 훈련 데이터에 대하여 인식 실험을 수행한 결과, 훈련이 충분히 되고 필순이 일정한 자소 모델들에 대하여 자체전이와 기본획이 안정적으로 대응함을 볼 수 있었다. 몇몇 자소의 기본획과 HMM의 자체전이의와의 대응관계를 정리한 예가 그림 3에 있다.

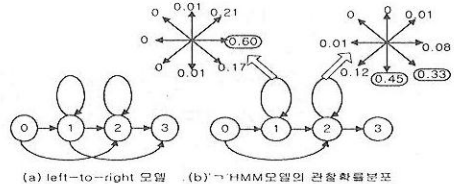


그림 2. 자소 HMM 모델구조와 예제('ㄱ'모델)

1 이러한 대응관계는 HMM 모델의 Maximum Likelihood Estimation 훈련 방법이 모델 확률을 극대화하기 위하여 유사한 필기 모양을 같은 상태(state)로 모아주는 경향이 있음에 기인한다.

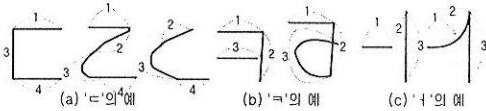


그림 3. 자소와 HMM 모델의 자체 전이와의 대응관계 예

2.2. 자소의 구조적 지식의 검증

자소를 기본획 단위로 분석한 후, 자소마다 고유하고 안정적인 기하학적 특징을 이용하여 구조적 분석을 할 수 있다. 이러한 자소의 구조적 특징으로는 자소 내에서 유지되는 기하학적 특징과 자소간에 유지되는 기하학적 특징을 들 수 있다. 자소 내 기하학적 특징으로는 자소의 기본획간 거리와 각도 관계 등이 있다. 자소간 기하학적 특징으로는 자소간 위치 관계를 들 수 있다.

자소의 기본획간 거리 관계의 예로는 'ㄱ', 'ㄴ', 'ㄷ' 등에서 유지되는 자소의 첫 기본획의 시작점과 수직획의 거리 관계를 들 수 있다. 실제로 그림 4-(b) 처럼 학습데이터에서 'ㄱ'의 시작 위치를 측정하게 되면 그림 4-(c)와 같은 위치관계를 보임을 알 수 있다.

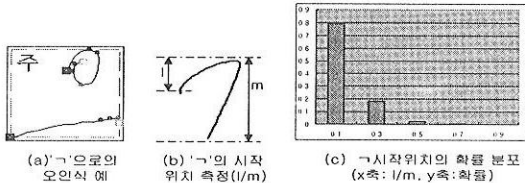


그림 4. 자소 내 기하학적 특징의 예

자소간 위치관계는 초성, 중성, 종성이 서로간에 갖는 일정한 위치 관계를 의미한다. 이러한 자소간 위치관계의 반영은 자소의 분할 오류를 막는 효과가 있다. 예를 들어 그림 5-(a) 에서 '만'을 '민'으로 오인식 한 경우에 분할된 중성의 모양은 'ㄷ'과 비슷하지만, 'ㄷ'의 시작 수평획 위치가 일반적인 경우에 비하여 오른쪽으로 치우쳐 있다. 이 경우 그림 5-(b) 처럼 중성과 중성 사이의 거리 정보를 이용하여 오류를 바로 잡을 수 있다.

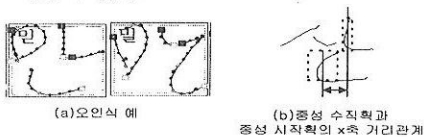


그림 5. 자소간 기하학적 특징의 예

2.3. 구조적 지식의 확률 분포

구조적 지식의 확률 분포는 자소 단위로 분할점이 표시된 학습 데이터를 이용하여 구한다. 학습 데이터에서 자소와 연결획을 분할한 후, 추출한 자소와 연결획을 해당 HMM 모델로 인식하여 기본획을 추출한다. 다음으로 기본획간에 유지되는 여러 구조적 지식을 측정하여 확률 분포를 구한다.

3. 전역적 지식의 적용

전역적 지식은 자소의 전반적인 정보를 의미하며 국소적 특징으로는 추출하기 힘든 전체적인 특징을 의미한다. 본 논문에서는 필기획의 방향 변화의 정도를 측정하는 누적각도, 기본획간 길이 비율인 자소 내 기본획별 지속 정보, 흘려 씌움의 척도인 연결획에서의 Pen-Up 움직임의 비율을 자소와 연결획 단위로 적용하였다.

3.1. 누적각도의 적용

누적각도는 그림 6-(a)에서 보듯이 자소 내에서 펜의 진행방향이 변화한 각도의 합으로 정의된다. 따라서 자소의 획의 갯수가 많고, 진

행 방향의 변화가 심할수록 값이 커지므로 자소의 복잡도를 측정하는 척도가 된다. 예를 들어 간단한 자소인 'ㄱ'의 경우 180도 정도이지만, 복잡한 자소인 'ㅁ'의 경우 1500도 정도가 되어 자소의 복잡도를 잘 반영한다. 또한 펜의 상태(매거나 불임)에 영향을 덜 받아 흘려 씌움에 안정적이라는 장점을 지닌다.

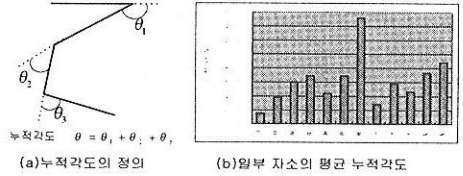


그림 6. 누적각도의 정의와 자소별 평균값

3.2. 자소 내 기본획별 지속 정보의 적용

자소를 구성하는 기본획간 길이 비율도 자소를 구별하는 특징이 된다. 이러한 기본획간 길이 비율은 각 기본획에 대응하는 자체전이에서 입력 코드가 지속한 시간에 대한 통계에서 구할 수 있다. 지속 시간을 duration modeling HMM으로 모델링하는 것은 계산량과 메모리가 많이 필요하므로[5], HMM으로 자소를 인식한 후, 후처리로 지속 시간을 계산한다[5,8,10]. 직선 기본획은 하나의 상태에 대응하므로 해당 상태에서 지속한 시간을, 원형 기본획은 여러 상태에 대응하므로 이들 상태의 지속시간의 합을 구하여 사용한다. 지속 시간을 구한 후, 평균과 표준 편차를 이용한 아래의 확률 분포 함수[2]로 지속 시간의 확률값을 계산하였다.

$$\text{확률함수 } f(x_i) = \frac{1}{\sigma} \cdot e^{-\frac{|x_i - \mu|}{\sigma}}$$

3.3. 연결획에서 Pen-Up 움직임의 비율

연결획은 자소를 서로 연결하는 획을 말한다. 연결획은 필기 습관에 따라 모양과 길이의 변이가 심하지만, 전체 이동 거리중 펜을 들고 움직인 거리의 비율이 큰 경향이 있다. 반면에 자소의 경우는 펜을 누르고 움직이는 Pen-Down 움직임 비율이 크다. 따라서 이러한 연결획의 Pen-Up 움직임 비율을 자소와 연결획을 구별하는 특징으로 이용할 수 있다.

3.4. 전역적 지식의 확률 분포

전역적 지식의 확률 분포는 학습데이터에서 자소와 연결획을 분리한 후, 자소와 연결획 단위로 통계량을 구하여 이루어진다.

4. 구조적, 전역적 지식과 HMM의 통합시스템

4.1. 기본 HMM시스템[1,2]

전처리로는 돌출점과 삐침같은 잡음을 제거하는 단계를 거친다. 입력 코딩은 입력을 일정한 단위 길이로 잘라서 재추출한 후, Pen-Down과 Pen-Up 직선에 대하여 각각 16 방향 코드를 부여하여 16방향 코드열로 변환하여 이루어진다. HMM 모델의 구조는 left-to-right 모델이며, Baum-Welch 방법으로 자소별로 학습을 시켰다[5]. 전체 인식 네트워크 구조는 자소 구성 원리와 자소의 시작, 끝 위치에 따라 구성된 한글 네트워크인 봉넷[1]으로 이루어진다. 인식은 Viterbi 알고리즘을 이용하여 수행한다.

4.2. 지식 확률의 보정(Rescoring)

구조적, 전역적 지식을 HMM 확률에 통합하려면 각 지식의 측정 확률을 보정하여야 한다. 이는 자소마다 적용되는 지식의 갯수가 다른 까닭에 더 많은 측정을 하는 자소가 1보다 작은 확률값이 계속 곱해져 손해를 보기 때문이다. 따라서 공평한 비교를 위하여 한 자소가

2 정규 확률 분포를 값이 완만하게 변화하도록 수정하였다.

측정한 확률을 확률분포의 기대값으로 나누어 보정하였다. 이는 해당 지식을 측정하지 않은 다른 자소에게는 측정 확률의 평균값을 부여함을 의미한다. 보정된 확률값은 이산확률과 연속확률의 경우 다음과 같이 구해진다.

$$\bar{p}_i = \frac{p_i}{\sum_i p_i \cdot p_i} \quad \overline{f(x)} = \frac{f(x)}{\int f(x) \cdot f(x)}$$

4.3. 지식을 추가한 네트워크 구조와 탐색

본 논문에서 구조적, 전역적 지식을 HMM 인식 네트워크에 추가하는 방법은 [2]의 연구를 응용한 것이다. 전체 인식 네트워크 구조는 한글 네트워크인 봉넷[1]에 구조적, 전역적 지식과 기존 인식기[2]의 지식을 추가한 형태이다. 각 HMM 자소모델마다 인식을 수행하여, HMM 확률값과 입력에서 자소에 대응하는 시작 시각과 끝시각을 구한다. 다음으로 자소에 대응하는 입력을 구조적, 전역적 지식으로 분석하여 지식의 확률값을 얻는다. 그리고 이 HMM 확률값과 지식 확률값을 곱하여 Viterbi 알고리즘으로 최대 확률값을 가지는 경로를 탐색한다.

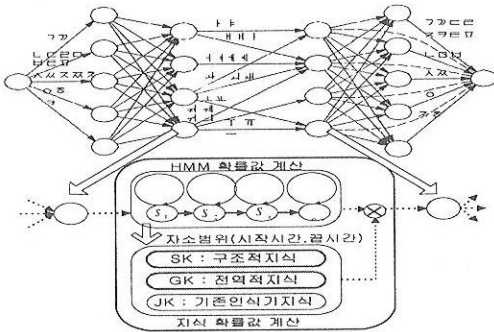


그림 7. 지식을 추가한 한글 인식 네트워크 구조

5. 실험 결과

본 논문에서 제안한 지식의 유용성을 알아보기 위하여 자소 단위의 구조적 지식이 포함된 기존 인식기[2]에 본 연구의 구조적, 전역적 지식을 추가하여 인식률을 비교하여 보았다.

5.1. 실험 환경

실험에 사용한 데이터는 KAIST 온라인 데이터로, 1992년에 WACOM SD-510C 디지털타이저에 필기에 제약없이 수집되었다. 데이터 집합은 크게 H, K, J로 구성된다. 학습에는 37명이 쓴 45,000자의 H 데이터와 21명이 쓴 7,000자의 K 데이터를 사용하였다. 실험 데이터로는 학습데이터의 필기자가 아닌 각각 10명의 H, K 데이터와, 빈도수가 가장 높은 한글 412자에 대한 39명의 필기데이터 J를 사용하였다. 데이터를 분석한 결과 K에는 또박또박 쓴 글씨가 많고, H에는 흘려 쓰고 변형이 심한 글씨가 많으며, J는 필체가 다양하였다.

5.2. 실험 결과

본 논문에서 제안한 구조적, 전역적 지식은 인식률을 평균적으로 94.8%에서 95.58%로 증가시켜서 오인식률을 14.5% 감소시켰다[표 1]. 특히 또박또박 쓴 글씨가 많은 K데이터에 대하여 효과적이어서 오인식률을 28.6% 줄였다. 반면에 흘려 쓰거나 변형이 심한 글씨에는 오인식률을 10% 정도 줄여서 비교적 덜 효과적이었다.

인식기	인식률			오류감소율		
	K(%)	H(%)	J(%)	K(%)	H(%)	J(%)
1) 기존 인식기	96.30	93.67	94.53	-	-	-
1)* 구조적, 전역적 지식	97.36	94.33	95.06	28.6	10.4	9.7

표1. 인식 실험 결과

6. 결론

본 논문에서는 기존 HMM 기반 인식기의 통계적 방법의 한계점과 국소적 특징의 한계점을 보완하기 위하여 구조적 지식과 전역적 지식을 적용하는 방법을 제안하였다. 구조적 지식을 적용하기 위하여 HMM의 자체전이와 입력을 대응시켜 기본화를 추출한 후, 기본화 간의 기하학적인 거리와 위치관계 등을 측정하여 사용하였다. 이는 기존의 인식기[2]가 자소단위의 구조적 지식을 적용한 것에 비하여 기본화 단위로 구조적 지식이 강화된 것이다. 전역적 지식으로는 자소의 전반적인 필기 경향에 대한 정보인 누적 각도, 자소 내 기본획별 지속 정보, 연결획의 Pen-Up 움직임 비율을 사용하였다. 측정 지식의 갯수가 다음에서 오는 문제는 측정 확률을 측정 확률 분포의 기대값으로 나누어 공평한 비교가 되도록 보정하여 해결하였다.

이러한 구조적, 전역적 지식의 적용 결과 인식률이 94.8%에서 95.58% 증가하였으며, 특히 또박또박 쓴 글씨에 대하여 30%에 가까운 오류를 줄일 수 있었다. 이는 실용적인 인식기에서 사용자가 오인식 때문에 여러 번 같은 글씨를 써야하는 불편을 줄일 수 있다는 점에서 사용자의 신뢰를 얻는데 중요한 장점이 된다.

향후 연구과제로는 다음 두 가지를 들 수 있다. 첫째로 구조적 분석의 신뢰를 향상하기 위하여 기본획과 HMM 상태의 대응관계를 더 강화시키고 원형획에까지 확장하는 연구가 필요하다. 다음으로 본 논문에서 사용한 다양한 지식의 신뢰도를 측정하여 가중치를 주는 연구가 필요하다.

참고 문헌

- [1] B.K.Sin and J.H.Kim, "Ligature modeling for online cursive script recognition", *IEEE PAMI*, Vol19, No.6, 1997
- [2] Jae-Ook Kwon, B.K.Sin and J.H.Kim, "Recognition of On-line Cursive Korean Characters Combining Statistical and Structural Methods", *Pattern Recognition*, Vol.30, No.8, pp.1255-1263, 1997
- [3] J.Y.Ha, *Unconstrained Handwritten Word Recognition with Interconnected Hidden Markov Models*, KAIST 박사학위 논문, 1994
- [4] X.D.Huang, Y.Ariki, and M.A.Jack, *Hidden Markov Models for Speech Recognition*, Edinburgh Information Technology Series, Edinburgh University Press, 1990
- [5] L.R.Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition", *Proceedings of IEEE*, vol.77, pp.257-285, Feb.1989.
- [6] J.Hu and M.K.Brown, "HMM based on-line handwriting recognition", *IEEE PAMI*, Vol18, No.10, 1996
- [7] H.J.Kim et al., "An HMM-based character recognition network using level-building", *Pattern Recognition*, vol.30, No.3, pp.491-502, 1997
- [8] 김진우, 초기확률분포 설정을 통한 은닉마르코프 모델의 훈련에 관한 연구: 온라인 문자인식의 적용, KAIST 석사학위 논문, 1993
- [9] 최동원, 온라인 영어 필기의 전역적 특징을 이용한 사전으로부터의 후보 단어 선택, KAIST 석사학위 논문, 1995
- [10] K.F.Lee, *Automatic speech recognition: The development of the SPHINX system*, Kluwer Academic publisher, 1989