

은닉 마르코프 모델을 이용한 온라인 혼용 필기문자 인식

이재준^{*}, 신봉기, 김진형
한국과학기술원 전산학과

On-line Mixed Character Recognition Using An HMM Network

Lee Jae Joon^{*}, Sim Bong Kee, Kim Jin Hyung
Department of Computer Science KAIST

요약

최근 펜을 이용한 필기 입력문자를 인식하는 연구가 활발히 진행 중이다 그러나 이 연구들은 한글이나 영어 등의 한가지 언어에만 제한되는 경향이 있다 본 논문에서는 펜을 이용한 한글/영어/숫자의 혼용 필기 입력에 대해 입력상의 모드 교환 없이 인식할 수 있는 혼용 필기문자 인식기를 개발하고자 한다

손으로 쓴 글씨는 출력은 물론이고 수많은 변형을 가지고 있다 본 논문에서는 시간적 제약을 받는 정보 구조와 그 다양한 변형을 모델링하는데 가치를 인정받아온 은닉 마르코프 모델을, 한글의 각 자소, 숫자, 알파벳의 필기 모델링에 이용한다 각각의 은닉 마르코프 모델을 연결해 한글 인식을 위한 봉넷 네트워크 숫자 인식을 위한 숫자 인식 네트워크 영문자 인식을 위한 영어 인식 네트워크를 구성한 뒤 이들 인식 네트워크를 구조적으로 통합하여 만든 혼용 필기문자 인식 네트워크를 제안한다 이러한 통계적 네트워크에서의 문자 인식이란 입력에 대해서 확률적인 최적 경로를 찾아내는 탐색 문제로 변환된다 이러한 경로는 비터비 알고리즘을 계층 구조의 네트워크에 확장 적용하여 효율적으로 구할 수 있다

특히 한글/영어/숫자에 대해 일관성 있는 모델 표현과 네트워크로의 확장으로서 별다른 오버헤드없이 혼용 필기문자를 인식할 수 있으며 다른 언어로의 확장이 용이하다

1 서론

온라인 (on-line) 문자 인식이란 전자펜으로 태블릿과 같은 장치에 쓴 글씨를 컴퓨터가 인식하는 것을 말한다 최근 들어 펜 컴퓨터 (pen-computer) 하드웨어 기술의 발달에 따라, 온라인 문자 인식에 대한 수요가 급격히 늘고 있으며, 이에따라 제반 연구도 활발히 진행중이다

영어를 제외한 나라에서는 자국어와 영어를 혼용해서 사용하는 경우가 많다 특히 외래어로 정착하기 전에는 정확한 뜻의 전달을 위해 영어표기 그대로 사용하는 경우도 많다 자국어와 아라비아 숫자를 혼용해 사용하는 것은 명백한 일이다 우리나라에서도 한글과 숫자 한자 영어를 혼용해서 사용하는 경우가 많다 따라서, 펜과 종이의 기능을 대체하려는 펜컴퓨터에서의 한글/영어/숫자의 혼용 허용은 필요요건 중의 하나라 할 수 있다

기존에는 필기 이전에 입력모드를 설정해 주거나 입력영역을 구분해 놓는 방법을 사용해 왔다 하지만 사람의 일상 필기형태와의 유사함과 편리함을 추구하려는 사용자 인터페이스의 구축이라는 측면에서, 이러한 제약을 없애고 입력시의 구분없이 처리하려는 혼용 필기문자 인식기의 연구가 시작되고 있다

이러한 혼용 필기문자 인식기의 구성에는 세가지 방법이 있다 첫째, 두개 이상의 각 언어의 인식기를 독립, 병렬적으로 수행하여 그 결과를 취합, 분석해 결론을 내는 방법이다 이 방법은 새로운 인식기 개발을 위한 추가 부담은 적으나, 인식기의 크기와 메모리 사용량이 두배 이상으로 증가하며, 병렬수행을 하지 못하는 경우에는 두 배 이상의 인식 시간이 걸리는 문제점이 있다 두가지 각각 인식기의 결과를 어떻게 공평하게 평가하고 판정하는 가의 어려운 점이 있다 둘째 입력 언어의 종류를 전처리를 통해 사전에 판단하여 해당 인식기만 수행하는 방법이 있다 이 방법은 사전 결정을 위한 약간의 추가 수행만이 필요하나, 각 언어의 종류를 확실하게 구분해 줄 수 있는 방법론을 찾기 어렵고, 사전 결정에 오류가 있을

경우에는 오류 수정하기가 어렵다는 위험성이 있다 셋째 하나의 방법론으로 각 인식기를 통합하는 방법이 있다 이 방법은 앞의 두 방법론이 가지고 있는 문제점을 해결해 준다 하지만 상이한 특징과 성질을 갖는 여러 언어를 한가지 방법론 안에 포용하기 어렵다는 단점을 가지고 있다 또한 새로운 인식기를 개발해야한다는 부담이 있다 이상의 방법론이 가지고 있는 어려운 점과 문제점들 때문에, 혼용 필기문자 인식기는 아직도 시도할 엄두조차 못하고 있는 경우가 많으며 초보적인 수준을 벗어나지 못하고 있다

은닉 마르코프 모델 (Hidden Markov Model 이하 HMM) 은 정형화 된 수학적 방법론으로, 순차적으로 발생하는 신호의 다양한 시간적 변형을 통계적으로 모델링하는데 뛰어나다 따라서, 시간 제약의 절대적 지배하에 있는 음성 신호 처리 및 인식에서 성공리에 활용되고 있으며, 역시 유사하게 필순, 즉 필기의 시간적 제약에 지배를 받는 온라인 필기 인식에 성공적으로 적용되고 있다 최근에 제안된 봉넷 (BongNet)[SIN 93]은 이러한 HMM을 이용해 필기의 변형을 최대한 흡수하면서, 각 자소 모델과 자소간의 연결 확률 모델을 필기순서에 알맞게 연결하여 한글 인식 네트워크를 구성함으로써 다양한 글씨체는 물론 자소간의 출력까지 허용하는 성공적인 한글 필기 인식기이다

본 논문은 HMM을 이용해 영어와 숫자 인식기를 각각 작성한 뒤 봉넷과 구조적으로 통합한, 혼용 필기문자 인식 네트워크 (Unified Recognizer UniRec)를 제안한다 이 네트워크는 한글 한 음절과 영어 알파벳 대/소문자, 숫자 모두를 하나의 커다란 틀 안에 모델링하는 필기 글자 모델이다 또한 필기에 따라 다양하게 변하는 특징을 네트워크 안의 노드 간의 확률적 전이로 표현하는 동적 정보를 가지고 있다 봉넷을 이용함으로써 한글 필기에서 생기는 여러가지 어려운 점들과 변형을 해결하며, 동일한 HMM 방법론을 사용하여 영어와 숫자 인식기를 작성하여 봉넷과 통합해 전체 네트워크를 구성함으로써, 동일한 인식 알고리즘과 처리과정을 사용하여 혼용 인식에 대한 추가부담이 적다

2 은닉 마르코프 모델

은닉 마르코프 모델 (이하 HMM)은 마르코프 체인과 각 상태에서의 관찰 심볼 확률 분포로 정의할 수 있는, 다시 말하면 확률적 마르코프 체인, 또는 마르코프 체인의 확률 함수라고 정의할 수 있다.

일반적으로 HMM은 아래와 같이 정의한다[RABI 88].

- 상태 $S = \{S_1, S_2, \dots, S_N\}$
- 관찰 심볼 $V = \{v_1, v_2, \dots, v_M\}$
- 초기 상태 확률 분포 $\Pi = \{\pi_i\}$:
 $\pi_i = P(q_0 = S_i), \quad 1 \leq i \leq N.$
- 상태 전이 확률 분포 $A = \{a_{ij}\}$:
 $a_{ij} = P(q_t = S_j | q_{t-1} = S_i), \quad 1 \leq i \leq N.$
- 관찰 심볼 확률 분포 $B = \{b_i(k)\}$:
 $b_i(k) = P(v_k | S_i), \quad 1 \leq i \leq N, 1 \leq k \leq M.$

HMM은 파라미터 $\lambda = (A, B, \Pi)$ 로 간단하게 표현하는 것이 보통이다.

본 논문에서는 시간적 제약을 구조 자체 내에 담고있는 left-to-right 형태의 HMM을 사용한다.

3 통합 인식 네트워크

3.1 한글 인식 네트워크

한글 인식 네트워크 봉넷 (BongNet)[SIN 93]은 각 자소 모델과 연결획 모델을 기반으로, 이를 한글의 글자 조합 원리를 이용해 연결함으로써 한글 필기 한 음절을 인식할 수 있도록 한 네트워크 구조이다.

한글에는 초성 19자, 중성 21자, 및 종성 27자 등, 총 67개의 자소가 있으며, 각각의 자소는 필기 습관, 필기 상태, 전체 자소의 결합형태에 따라 다양한 변형을 가지게 된다. 이러한 변형의 흡수를 위해 각각의 자소를 하나의 HMM으로 모델링 한다. 모델의 상태 수는 자소의 모양 복잡도에 따라 4 ~ 16개로 한다.

한글 한 음절을 필기하다 보면, 자소간에 '홀림'이 발생한다. 이를 모델링하기 위해 연결획 (ligature) 모델을 도입한다.[SIN 93] 연결획이란 하나의 자소에서 다음 자소로의 전이과정에서, 두 자소간의 경계 영역이 된다. 연결획은 연결획의 시작부분과 마침부분의 상대적인 위치에 따라 모델을 나누게 되며, 각각은 필기의 변형을 고려해 역시 HMM으로 모델링 한다.

한글을 필기할 때에는 초성, 중성, 종성의 순서로 필기한다. 앞서 도입한 연결획을 고려하면, 한글 한 음절은

- (초성) + (연결획) + (중성)
- (초성) + (연결획) + (중성) + (연결획) + (종성)

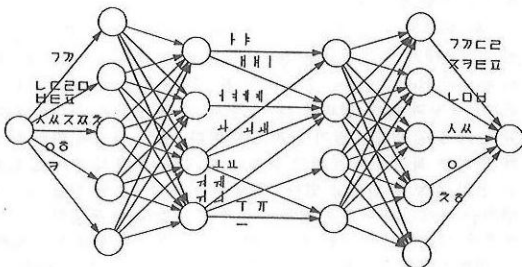


그림 1: 봉넷, 한글 인식 네트워크

과 같은 확장된 자소열로 표현할 수 있다. 이러한 방법으로 각각의 자소 모델과 연결획 모델을 연결한 것이 한글 인식 네트워크 봉넷 (BongNet)이다 (그림 1). 네트워크의 시작 노드에서 종료 노드로의 각 경로는 하나의 글자에 해당한다.

3.2 영어 인식 네트워크

영어 알파벳에는 대문자 26개, 소문자 26개 등 총 52개의 글자가 있다. 영어 알파벳은 한글 자소에 비해 곡선성분이 더 많고, 시작 부분과 끝부분에 선택적인 필기 부분이 있는 등, 다양성과 변형이 더 많다. 따라서 각각의 알파벳을 하나의 HMM으로 모델링하되, 상태의 수와 전이의 수를 늘려주어 모델의 포용성을 증대시킨다. 또한 각 상태에서 다음 상태로 관찰 심볼을 생성하지 않고 전이하는 null-arc를 모델에 추가하였다. 각 모델의 상태의 수는 알파벳의 모양 복잡도에 따라 10 ~ 20개로 한다.

각각의 알파벳에 대한 HMM을 병렬적으로 연결하면, 영문자를 인식할 수 있는 영어 인식 네트워크를 구성할 수 있다 (그림 2) [HA 92].

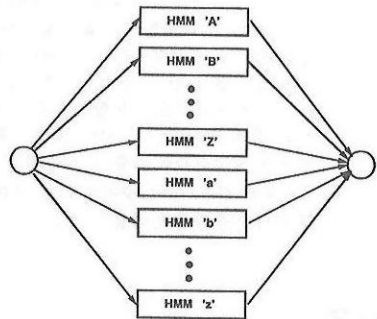


그림 2: 영어 인식 네트워크

3.3 숫자 인식 네트워크

숫자 인식기의 구성은 영어 인식기의 구성과 동일하다. '0'에서 '9'까지의 10개의 숫자에 대해 각각 HMM을 구성하며, 모델의 상태 수는 모양 복잡도에 따라 10 ~ 20개로 한다. 영어와 마찬가지로 null-arc를 모델 구조에 추가하였다.

각 숫자에 대한 HMM을 병렬적으로 연결해 숫자 인식 네트워크를 구성할 수 있다.

3.4 통합 인식 네트워크

한글/영어/숫자는 모두 HMM으로 모델링되어 있고, 유사한 구조를 가지므로, 하나의 커다란 네트워크로 통합을 시키는 것이 가능하다. 네트워크의 시작 노드와 종료 노드 사이에 한글 인식 네트워크 봉넷과 영어 인식 네트워크, 숫자 인식 네트워크를 병렬적으로 연결하여 구성된 네트워크가 바로 통합 인식 네트워크 (Unified Recognizer)이다 (그림 3).

통합 인식 네트워크는 한글 한 음절과 숫자, 영어 알파벳 모두를 모델링 하며, 다음 절에서 설명할 비터비 알고리즘을 이용해 한글/영어/숫자에 대해 인식 결과를 구할 수 있다. 네트워크의 시작 노드에서 종료 노드로의 각 경로는 하나의 글자에 해당한다. 그러므로 인식은 주어진 글씨에 대해서 그 글씨가 표현코자 하는 글자의 경로를 찾는 작업이라고 할 수 있다. 즉 통계적으로 가장 유사한 경로를 찾으면 된다.

통합 인식 네트워크의 시작 노드에서 각 한글/영어/숫자 인식 네트워크로의 경로를 차단하면, 각각의 전용 인식기로 사용할 수

통합 인식 네트워크에서 인식을 수행해 후보문자들 생성한 뒤, 후보문자중에 위의 부류에 속하는 문자가 있는 경우 쌍 구분 과정을 거쳐, 그 결과물 전체 인식 스코어에 반영하도록 하였다 쌍 구분 과정을 적용하기 이전과 이후의 인식 실험 결과는 표 1과 같다 그 결과는 그 글자에 대한 정인식률을 나타낸다

표 1 쌍 구분 실험 정인식률(%)

	적용이전	적용이후		적용이전	적용이후
O	98.2	100.0	b	91.3	97.3
o	97.4	92.2	p	80.0	94.8
0	98.8	100.0	D	69.1	87.3
6	27.5	98.8	P	83.3	92.6
h	71.3	93.9	a	99.1	95.7
n	75.9	98.2	d	65.2	88.7

표 1의 결과물 보면, o와 a'의 경우에는 인식률의 저하가 있는데, 이는 오류율을 최소로 하기 위한 통계적 방법의 불가피한 결과이다 하지만 전체적으로 보면 내부류 모두에 대해 현격한 인식률의 향상을 볼 수 있다

5.2.2 관찰 심볼 확률 분포 평활화 실험

코드화 과정을 거치면서 입력 자료는 이산화된다 이 과정에서 정보의 손실을 가져오게 되는데 코드북의 심볼 갯수를 늘려주면 정보의 손실은 줄일 수 있지만 대신 파라미터의 갯수가 증가하게 되어 많은 양의 훈련 데이터가 필요하게 된다 이 문제를 해결하기 위해, 조금 부족한 양의 훈련 데이터로 정보의 손실을 최대한 막고자 하는 방법이 관찰 심볼 확률 분포의 평활화(smoothing)이다

훈련 과정에서 구해진 HMM의 파라미터 $\lambda = (A, B, \Pi)$ 중에서 각 상태에서의 관찰 심볼 확률 분포를 나타내는 $B = \{b_i(k)\}$ 에 대한 파라미터 재추정시, 관측된 하나의 심볼에 대한 확률값만을 높여주는 것이 아니라 주변의 관찰 심볼에 대해서도 확률값을 나눠주는 방법을 말한다 HMM의 상태 i에서 k번째 관찰 심볼을 관측한 경우, 파라미터 $b_i(k)$ 의 재추정 값은,

$$b'_i(k) = \frac{w_1 b_i(k-1) + w_2 b_i(k) + w_3 b_i(k+1)}{w_1 + w_2 + w_3}$$

이다 여기서 w_1, w_2, w_3 는 가중치를 나타낸다

각 HMM의 관찰 심볼 확률 분포에 대해 평활화(smoothing)를 여러가지 비율로 행한 후의 인식 실험 결과는 표 2와 같다 각 체인 코드 심볼에 대해 양 옆의 체인 코드 심볼에만 파라미터 값을 가중치 비율에 따라 나누어 주는 방법을 사용한다 (예, 1 2 1은 좌측 코드에 25% 자신에게 50% 우측 코드에 25%를 나누어 갖는 것을 뜻한다)

표 2 평활화 실험 정인식률(%)

방법	한글	영어(대문자)	영어(소문자)	숫자
0 1 0	88.3	79.8	79.5	90.5
1 2 1	91.0	87.6	83.1	92.0
1 3 1	91.0	87.2	83.5	93.0
1 4 1	90.7	87.2	84.7	93.0
1 5 1	90.5	88.0	84.7	93.0

표 2의 결과물 보면, 훈련 데이터가 비교적 충분한 한글의 경우에는 비율 1 2 1인 경우 가장 좋은 결과를 내었다 그러나 훈련 데이터가 부족했던 영어와 숫자의 경우에는 많은 인식률의 향상을 가져왔다 특히 훈련 데이터의 양이 가장 적은 영어 대문자 경우에 가장 큰 인식률의 향상을 가져왔다

5.2.3 각 인식기에 대한 분리/통합 실험

한글/영어/숫자에 대해 각각의 독립적인 인식기를 구성한 경우와, 모두를 통합한 인식기를 구성한 경우의 인식 실험 결과는 다

표 3 각 인식기 분리/통합 실험 정인식률(%)

방법	한글	영어(대문자)	영어(소문자)	숫자
분리 인식기	91.3	90.3	88.2	97.5
통합 인식기	90.5	88.0	84.7	93.0

표 3의 결과물 보면, 한글/영어/숫자 각각의 인식기를 구성해 실험했을 때에 비해 통합한 경우의 인식률이 저하되었음을 알 수 있다 하지만 한글의 경우에는 비교적 영향이 적고, 통합에 의해 하락된 인식률도 그다지 나쁜 결과가 아닐을 알 수 있다

6 결론

본 논문에서는 정형화된 방법론인 은닉 마르코프 모델과 비터비 알고리즘을 이용하여 한글/영어/숫자의 온라인 혼용 필기 입력에 대해 입력상의 구분없이 인식할 수 있는 방법론을 제안하였다 HMM을 기반으로하여 한글의 각 자소와 연결획, 영어 알파벳, 숫자를 모델링한 뒤 이를 각각 구조적으로 연결하여 한글 인식 네트워크, 영어 인식 네트워크, 숫자 인식 네트워크를 구성하였다 다음 이들을 병렬적으로 통합하여, 동일한 전처리 과정과 인식과정을 수행해 한글/영어/숫자가 혼합된 결과물을 얻을 수 있는 혼용 필기 통합 인식 네트워크를 구성하였다 각각은 통합에 의해 인식률의 하락이 있었지만, 한글 91%, 영어 대문자 88%, 영어 소문자 85%, 숫자 93%의 인식률을 보였다

그러나, 제안된 방법론은 다음 몇가지 점에서 수정 및 보완이 필요하다 첫째 영어와 숫자의 경우에는 통합에 의한 인식률의 하락이 있다 서로 비슷한 모양의 글자에 의한 영향인데, 이를 막기 위해서는 인식과정에 새로운 특징의 반영이 필요하다 둘째, 영어와 숫자에 대해 훈련 데이터가 부족하여 만족할 만한 인식률을 얻지는 못하였다 셋째, '9', 'g' 'q' 부류 등 혼동되는 몇가지 부류에 대한 추가적인 쌍 구분 과정의 도입이 필요하다

본 통합 인식기는 하나의 방법론을 각 언어에 대해 일관성 있게 적용함으로써 다른 언어의 인식으로의 확장이 용이하다 따라서 앞으로 제스처어(gesture)와 특수 문자로의 확장을 통해, 보다 발전된 통합 인식기의 개발이 기대된다

참고 문헌

[TAPP 90] C C Tappet C Y Suen and T Wakahara 'The State of the Art in On-Line Handwriting Recognition' *IEEE Trans on PAMI* v 12 n 8, pp 787-808 August 1990

[BAUM 70] L E Baum et al 'A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains.' *Ann Math Stat* v 41 n 1 pp 164-171 1970

[RABI 88] L R Rabiner, 'A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition' *Proc of the IEEE* v 77 n 2 pp 257-286 1989

[FORN 78] G D Fooney 'The Viterbi Algorithm' *Proc IEEE* v 61 pp 268-278 Mar 1978

[SIN 93] 신봉기 김진형 "은닉 마르코프 모델을 이용한 온라인 한글 인식" 제1회 문자인식 워크샵 발표 논문집 pp 189 194 1993 5

[HA 92] 하진영 김진형, HMM을 이용한 온라인 영어인식 '제6회 한국 인지과학회 춘계 학술발표논문집 pp 264-274 1992 5

[KANG 93] 장민석등 "한글 영문 혼용 인식 구현을 위한 마스터 슬레이브 인식 시스템의 설계 기법" 제 1회 문자인식 워크샵 발표 논문집 pp 175 180 1993 5