

자소 탐색에 기반한 온라인 한글 인식

(On-line Hangul Character Recognition based on Letter-Spotting Technique)

신봉기[†] 김진형^{**}
(Bong-kee Sin) (Jin Hyung Kim)

요약 본 논문은 정상적인 필순의 한글 필기에서 가능한 모든 자소를 독립적으로 찾는 자소 탐색 방법에 의하여 글자를 인식하는 방법을 제안한다. 은닉 마르코프 모델에 기반한 탐색 모델을 각 자소별로 설계하고, 자소 탐색 방법을 기술한다. 그리고 글자 인식은 자소 탐색 결과에서 최적의 자소 조합을 찾는 격자 탐색 문제로 정의한다. 실험 결과에 의하면 매우 간단한 인식 구조에도 불구하고 자연스럽게 쓴 필기 87.78%를 인식해내었다. 본 방법의 핵심은 정확한 자소 경계를 찾는 데 있는데, 이를 위한 탐색 모델의 설계와 두 단계 탐색 알고리즘에 대하여 중점적으로 기술한다.

Abstract This paper presents an approach to on-line Hangul character recognition based on the letter spotting technique. Utilizing the regular writing order, the three types of letters are located in an input character by matching letter spotting models. Character recognition is defined as the search for the best legal combination of letters, which is performed over the lattice of letter hypotheses. Experimental results show the performance up to 87.78% for normal writing style despite the simplicity of the recognition architecture. The primary concern of the approach lies in the detection of letter boundaries which are highly nontrivial when it comes to cursive script.

1. 서론

문서 처리 및 인식 분야에서 글자 탐색이란 입력 텍스트 영상 또는 필기 데이터에서 특정 글자를 찾는 문제를 일컫는다[4,5,6]. 특정 글자의 탐색에는 그 글자의 모형 또는 참조 패턴을 탐색 키로 쓰게 되는데[7], 이를 간단히 '키 글자'라고 하며, 자소 탐색의 경우에는 '키 자소'라고 한다. 글자 탐색 기법이 온라인 필기 인식에 적용된 것은 극히 최근의 일로서, 영어 단어 인식에 응용한 [8]의 연구를 대표로 들 수 있다. 이 방법에서는 입력 단어 필기 상에서, 모든 가능한 알파베트를 중복해서 찾은 다음 알파벳 가설을 차례로 최적 연결하는 알파벳 열을 찾고자 한다. 이때 사전에 있는 알파벳 열이 단어 인식 결과가 된다.

본 논문에서는 정상적인 한글 필순을 가정하고 자소

가설 탐색 문제와 글자 인식 문제를 풀고자 한다. 통계적 모형화 도구의 하나인 은닉 마르코프 모델(HMM)을 이용하여 [1]과 같이 자소 및 연결획 단위의 필기 모형(HMM)을 만들고, 이를 기반으로 입력 필기에서 자소를 찾는다. 글자 인식은 이들 자소의 격자에서 최적 결합을 찾는 또 다른 탐색으로 이루어 진다. 전체 인식 알고리즘은 효율적인 순차 탐색 구조로 구성할 수도 있고, 최적 결과를 낼수 있지만 복잡도가 높은 island-driven 탐색[9]을 결합한 형태로 구성할 수도 있다. 본 연구의 실험에서는 자소 탐색은 심볼 수준의 순차적인 디코딩 방법을 이용하고, 글자 탐색은 자소 수준의 순차적 격자(lattice) 탐색 형태로 어떻게 구현하여 실험하였다.

이하 본 논문의 구성은 다음과 같다. 제 2절에서 HMM과 필기의 특성에 따른 필기 모형을 간단히 설명하고, 제 3절에서 자소 탐색 모형을 기술한다. 제 4절과 제 5에서는 각각 자소 탐색과 격자 탐색 방법을 기술한다. 그리고 제 6절에서는 실험 결과와 예를 보이며 제안된 방법론의 장단점을 검토한다.

† 정 회 원 . 한국통신 멀티미디어연구소

** 중 심 회 원 . 연구개발정보센터
논문접수 : 1996년 2월 4일
심사완료 : 1996년 9월 17일

2. 필기 모형

2.1 은닉 마르코프 모형

은닉 마르코프 모델(HMM)은 Markov 체인과 확률적 심볼 출력 확률 과정으로 구성된다. 확률적 Markov 체인 또는 Markov 체인의 확률 함수라고 불리기도 하는 HMM은 유한한 N 개의 상태가 연결된 방향 그래프 구조로 표현된다. V 를 출력 심볼의 집합이라 하자. HMM은 다음과 같은 세가지의 확률 파라미터로 기술한다[10].

- 상태 전이 확률 분포:

$$A = \{ a_{ij} | a_{ij} = P(q_{t+1} = j | q_t = i), 1 \leq i \leq N \},$$

$$\sum_j a_{ij} = 1$$

- 출력 심볼 확률 분포:

$$B = \{ b_j(k) | b_j(k) = P(v_k | j), 1 \leq j \leq N, v_k \in V \},$$

$$\sum_k b_j(k) = 1, 1 \leq j \leq N$$

- 초기 상태 확률 분포:

$$\pi = \{ \pi_i | \pi_i = P(q_0 = i), 1 \leq i \leq N \}, \sum_i \pi_i = 1$$

HMM은 위 세가지의 확률 파라미터를 위주로 하여 보통 $\lambda = (A, B, \pi)$ 와 같이 표시한다. HMM은 통계적 모형의 하나로서 패턴 벡터 X 의 확률 분포

$$f(X) = p(X|\lambda)$$

를 정의한다. 출력 심볼열 $X = x_1 x_2 \dots x_T$ 가 주어졌을 때, 출력 확률 $p(X|\lambda)$ 를 어떻게 계산할 것인가? 이 문제는 Baum 등의 forward/backward 알고리즘으로 해결되었다[10,11]. 그리고, HMM의 파라미터를 추정하는 문제에는 forward/backward 알고리즘을 기반으로 하여 반복적인 EM(Estimation-Maximization) 기법의 Baum-Welch 훈련 알고리즘이 나와 있다[10,11]. 한편, 심볼열 X 가 주어졌을 때, X 에 대응하는 최적 상태열 $Q = q_1 q_2 \dots q_T$ 를 찾는 세 번째 문제에 대해서는 이미 여러 분야에서 널리 쓰이고 있는 Viterbi 알고리즘이 유명하다[12]. Viterbi 알고리즘은 forward/backward 알고리즘의 forward 절차와 유사하다. 본 논문에서 기술된 자소 탐색 방법도 이 알고리즘에 기반을 두고 있다.

2.2 한글 필기 모형

복자음, 복모음을 포함하여 한글에는 초성 19자, 중성 21자, 그리고 종성 27자 등 총 67 자소가 있으며, 이들의 trigram 조합으로 글자를 구성한다. 각 글자는 특별

한 경우를 제외하고는 자소만 홀로 쓰이지 않는다. 따라서 본 논문에서는 음절 단위로 필기 모형을 구성하기로 한다. 한편 [1]에 따르면 필기는 단순히 자소의 개념적 연결이 아니라, 운동 역학적 제약과 최소의 에너지로 빨리 쓰려는 심리적 욕구에 따라 자소와 자소 사이에 특징적인 연결 패턴('연결획')을 포함한 하나의 펜 궤적으로 정의된다. 따라서 한글 필기를

$$\langle \text{글자} \rangle := \langle \text{초성} \rangle \langle \text{연결획} \rangle \langle \text{중성} \rangle \langle \text{연결획} \rangle \langle \text{종성} \rangle$$

로 정의하기로 하고, 자소 모형과 연결획 모형의 연결로 그림 1과 같은 글자 모형을 구성한다[1].

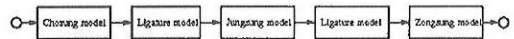
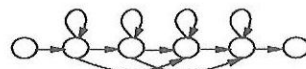
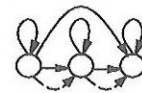


그림 1 글자 모형

한글의 구조 분석에 따르면 여섯가지의 형식 분류가 있으며, 이에 따라 글자 모형을 만드는 연구 결과를 많이 볼 수 있다. [3]에 따르면 초성은 각각 6가지의 변형, 중성은 2가지, 그리고 종성은 3가지씩의 변형으로 세분하여 표준 패턴 모형을 설계하게 된다. 하지만 일반적으로 필기에서 그러한 구분을 찾기란 매우 어렵다. 따라서 본 논문에서는 각 자소에 대해 하나씩 총 67개의 자소 모형을 설계하였다. 모든 모형은 왼쪽에서 오른쪽으로 전이하는 전진 구조(그림 2(a))를 갖도록 하였다. 온라인 필기의 시간 정보를 구조적으로 표현하고, 또 계산상의 효율을 기할 수 있기 때문이다.



(a)



(b)

그림 2 (a) 자소 마르코프 모형, (b) 연결획 마르코프 모형

인쇄체에 비하여 필기체의 가장 두드러진 특징은 흘러 쓴 영어 필기체에서와 같이 획과 획 사이를 이어주는



그림 3 탐색 오류의 예. 붉은 선은 각 자소의 영역을 표시한다: (a) 확률 비교에 의한 초성 오류, (b) 중성의 경계 오류.

합이 이루어지지 않아야 한다. 이와같은 목적의 탐색 모형 구조는 앞에서 정의한 자소, 연결획, filler HMM을 이용하여 설계할 수 있다.

자소 탐색에 있어서 가장 큰 문제는 그림 3과 같은 탐색 오류 (false alarm)가 예상 외로 빈번하게 나타난다는 것이다[8]. 이러한 탐색 오류는 종종 높은 확률을 나타내는데, 다음 단계의 글자 탐색 공간을 크게 확장시키는 요인이 된다[16]. 한편 자소의 존재를 알았다고 해도 자소와 자소가 부드러운 곡선으로 이어지는 경우가 많아 정확한 경계를 결정하기가 쉽지 않다. 올바른 글자 인식을 위해서는 자소의 경계를 정확하게 찾아내어야 한다. 글자 인식기의 성능은 자소 경계의 질 여하에 크게 좌우되기 때문이다.

위 문제점을 해결하기 위한 자소 탐색 모형은 제 2.2 절의 정의에 따라 구성된 글자 모형을 변형하여 각 자소에 대해 하나씩 구성한다. 글자 모형은 앞 절의 자소 및 연결획 HMM을 이용하여 네트워크 구조로 표현한다(그림 1). 본절에서 제안하는 각 자소별 탐색 모형의 구조는 그림 4, 5와 같다.

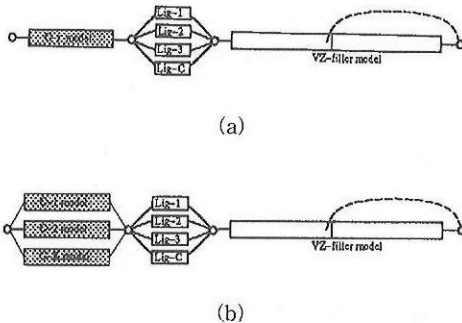


그림 4 초성 탐색 모형: 회색 모형은 키 자소 모형이고, C-1, C-2, ..., C-K는 같은 군집에 속한 자소 모형이다.

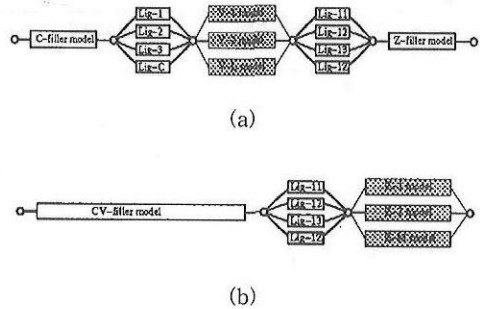


그림 5 자소 탐색 모형: (a) 중성 탐색 모형, (b) 중성 탐색 모형

초성 탐색 모형은 키 자소인 초성 HMM과 그 초성에 연결되는 연결획, 그리고 filler 모형에 연결되어 있다. 중성 모형은 이와는 달리 키 자소 중성 HMM을 가운데 두고 그림 5(b)와 같이 앞(좌측)에 filler 모형 및 연결획 모형을 붙이며, 뒤(우측)에는 연결획, filler 모형을 차례로 연결한다. 이는 한글 필기가 초성, 중성, 중성순이라는 데서 나온 것으로 앞쪽의 filler 모형은 임의의 초성을 수용하고, 뒷쪽의 filler 모형은 중성을 수용하는 것이다. 한편 받침이 없는 글자의 경우는 탐색 과정에서 해결하도록 한다. 마지막으로 중성 탐색 모형은 초성 모형과 유사하다. 중성 HMM이 탐색 네트워크에서 가장 우측에 있다는 것 외에는 다른 점이 없다.

위와 같은 구조의 모형을 이용했을 때 탐색 결과 얻을 수 있는 자소의 경계는 많은 경우 상당히 가능성을 엿볼 수 있었다. 전형적인 예로써 그림 6은 그림 3의 오류와는 달리 매우 개연성 있는 초성 탐색 결과를 보여준다.



그림 6 올바른 초성 탐색 결과-붉은 선 영역

4. 자소 탐색

4.1 탐색 알고리즘

HMM 네트워크 모형 $M = M_1 M_2 \dots M_R$ 와 시계열 입력 X 의 정합은 모형 M 에서 X 를 생성할 확률 $P(X|M)$ 을 구하는 모형 평가 문제로 정의된다. 그러나 시계열 입력을 일련의 조각(segment)으로 분할할 필요가 있을 경우

에는 일반적으로, 최적 평가를 얻을 수 있는 forward 확률 $P(X|M)$ 을 쓰지 않고 Viterbi 알고리즘에 의한 근사치인 경로-출력 결합 확률

$$P(X, Q|M) \tag{1}$$

을 사용하게 된다. 여기서 Q 는 경로, 즉 네트워크의 상태를 가리킨다. 이후 기술할 계산 과정은 모두 Viterbi 순방향 계산을 중심으로 한다. Viterbi 알고리즘의 상세한 구현 내용은 [10,12,17]을 참고하기 바란다.

HMM의 일차 마르코프 가정과 Viterbi 알고리즘의 단계별 계산은 비슷한 형식을 갖고 있다. 제 2.1절에 언급한 바와 같이 Viterbi 알고리즘은 입력 $X=x_1x_2\cdots x_T$ 와 탐색 모형 M 의 최적 상태열 \bar{Q} 의 결합 확률을 계산하는 과정, 즉, 식 (1)을 최대로 하고 그때의 \bar{Q} 를 찾는 과정을 말한다. 식 (1)은 베이스 정리에 따라

$$P(X, Q|M) = P(Q|M)P(X|Q, M) \tag{2}$$

과 같이 쓸 수 있다. 한편, 자소 탐색의 결과는

$$\max_Q P(Q|M)P(X|Q, M) \tag{3}$$

을 만족하는 상태열에서 얻을 수 있다. $Q=q_1q_2\cdots q_T$ 는 입력 X 에 배열된 모형 M 의 상태열을 나타내며, $|Q|=|X|=T$ 는 열의 길이를 가리킨다. 위에서 HMM의 기본 가정, 즉, Markov 가정과 입력 심볼간 독립 가정을 식 (3)에 적용하면

$$\max_Q \prod_{i=1}^T a_{q_{i-1}, q_i} \prod_{i=1}^T b_{q_i}(x_i) \tag{4}$$

가 된다. 여기서 $a_{q_{i-1}, q_i} = \pi_{q_i}$ 로 표시한다. 위 식을 최대로 하는 Q 를 \bar{Q} 라고 하자.

4.2 초성 탐색

초성 탐색 모형 M 은 초성 키 HMM C 와, 연결획 모형 L , filler 모형 F 의 연결로 되어 있으므로 $M=C \cdot L \cdot F$ 로 표현할 수 있다. 이 탐색 모형에 배열된 최적 상태열을

$$\bar{Q} = \bar{Q}^C \bar{Q}^{LF} = \bar{Q}^C \cdot \bar{Q}^{LF}$$

이라고 하자. 여기서 $\bar{Q}^C = q_1q_2\cdots q_r$ 는 초성 탐색 네트워크에서 C 에 속한(배열된) 상태열이고 $\bar{Q}^{LF} = q_{r+1}\cdots q_T$ 는 나머지 부분, 즉 L 과 F 에 속한 상태열이다. $0 < |\bar{Q}^C| = r < T$ 라고 하면 초성 가설은 다음과 같은 triple

로 정의된다.

$$(1, r, p_C) \tag{5}$$

이 중에서 첫째, 둘째 원소의 값은 초성의 시작과 끝 경계점을 나타내며, 마지막 원소의 값은 초성 모형의 평가

$$p_C = P(\bar{Q}^C, X_{1,1+r}|C) = \prod_{i=1}^r a_{q_{i-1}, q_i} b_{q_i}(x_i)$$

를 나타낸다.

4.3 중성 탐색

초성 탐색의 경우와 비슷하게 중성 탐색 모형을 $M=F_c \cdot L_c \cdot V \cdot L_z \cdot F_z$ 로 나타내자. V 는 중성 키 자소 HMM을 나타낸다. M 에 배열된 최적 상태열을

$$\bar{Q} = \bar{Q}^{FL_c} \bar{Q}^V \bar{Q}^{L_z F_z} = \bar{Q}^{FL_c} \cdot \bar{Q}^V \cdot \bar{Q}^{L_z F_z}$$

이라고 하자. 여기서 \bar{Q}^V 는 중성 키 모형에 속한 상태열이고 \bar{Q}^{FL_c} 과 $\bar{Q}^{L_z F_z}$ 는 전후의 나머지 부분, 즉 각각 filler 모형-연결획 모형, 연결획 모형-filler 모형에 속한 상태열이다. $|\bar{Q}^{FL_c}| = \tau_1 > 0$, $|\bar{Q}^V| = \tau > 0$, $|\bar{Q}^{L_z F_z}| = T - \tau_1 - \tau > 0$ 이라고 하면 중성 가설은 다음과 같다.

$$(\tau_1+1, \tau_1+\tau, p_V) \tag{6}$$

마지막 원소의 값은 HMM V 내의 최적 상태열과 $X_{\tau_1+1, \tau_1+\tau} = x_{\tau_1+1} \cdots x_{\tau_1+\tau}$ 와의 결합 확률

$$p_V = P(\bar{Q}^V, X_{\tau_1+1, \tau_1+\tau}|V) = \prod_{i=\tau_1+1}^{\tau_1+\tau} a_{q_{i-1}, q_i} b_{q_i}(x_i)$$

를 나타낸다.

4.4 종성 탐색

탐색 모형의 구조는 $M=F \cdot L \cdot Z$ 와 같다. 종성 Z 탐색은 키 HMM의 위치가 뒤에 있다는 점 외에는 개념적으로 초성과 별 차이가 없다. M 에 배열된 최적 상태열을

$$\bar{Q} = \bar{Q}^{FL} \bar{Q}^Z = \bar{Q}^{FL} \cdot \bar{Q}^Z$$

이라고 하자. 여기서 \bar{Q}^Z 는 종성 키 HMM Z 에 속한 상태열이고 \bar{Q}^{FL} 는 나머지 부분, 즉 filler 모형과 연결획 모형에 속한 상태열이다. $0 < |\bar{Q}^Z| = \tau < T$ 이라고 하면 종성 가설은 다음과 같다.

$$(T-\tau+1, T, p_Z) \quad (7)$$

마지막 원소의 값은 HMM Z 내의 최적 상태열과 $X_{T-\tau+1, T} = x_{T-\tau+1} \dots x_T$ 와의 결합 확률

$$p_Z = P(\overline{Q^Z}, X_{T-\tau+1, T} | Z) = \prod_{t=T-\tau+1}^T a_{\overline{q}_{t-1}} b_{\overline{q}_t}(x_t)$$

를 나타낸다.

각 자소 탐색은 서로 독립적인 소문제로서 서로간에 간섭이 없다. 따라서 모든 자소를 동시에 찾는 병렬 탐색이 가능하다. 각 자소별 탐색에는 입력 길이에 비례하는 일차 복잡도를 갖는 Viterbi 알고리즘을 사용하였다. 그러나 표준 Viterbi 알고리즘과 같이 획일적인 순방향 탐색을 할 필요는 없고 필요에 따라 편리하게 역방향 순차 탐색(중성의 경우) 또는 순방향-역방향 혼합 구조의 양방향 탐색(중성의 경우)을 구현할 수도 있다.

시간 축상에 각 자소 가설의 영역을 도시하면 그림 7과 같다. 이러한 그림을 격자라고 부른다. 그림에서 가로 빗장은 각각 자소의 영역을 나타내며 t_i 는 자소 분할점을 나타낸다.

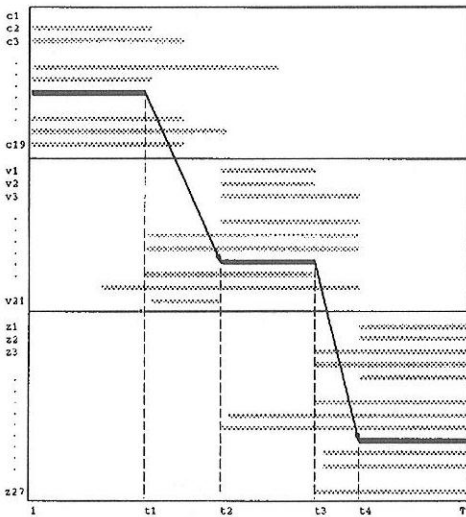


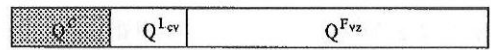
그림 7 자소 가설의 격자

5. 격자 탐색

탐색된 자소를 결합하여 하나의 완성된 글자로 인식

하는 후속 처리 과정은 격자로 표현된 자소 탐색 결과에서 초성-중성 또는 초성-중성-중성으로 적절히 연결하고 그 중에서 최적 정합값을 갖는 자소 조합을 찾는 과정을 말한다. 이와 같이 임의의 격자 내에서 진행되는 탐색을 격자 탐색이라고 한다.

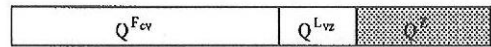
글자 가설은 연결획 영역을 사이에 끼고 자소의 가설 영역을 이어서 만든다. 이 때 자소, 연결획 영역 간에는 빈틈이나 중첩이 없어야 한다. 그리고 초성은 항상 $t=1$ 에서 시작하고, 중성은 $t=T$ 에서 끝난다. $Q^{CL\alpha F_{\alpha}}$, $Q^{F_{L\alpha}VL\alpha F_{\alpha}}$, $Q^{F_{\alpha}L\alpha Z}$ 를 각각 초성, 중성, 중성 자소 탐색 모형을 구한 경로라고 하자(그림 8).



(a)



(b)



(c)

그림 8 자소 탐색 결과 상태열 분할: (a) 초성 탐색 결과 $Q^{CL\alpha F_{\alpha}}$, (b) 중성 탐색 결과 $Q^{F_{L\alpha}VL\alpha F_{\alpha}}$, (c) 중성 탐색 결과 $Q^{F_{\alpha}L\alpha Z}$.

이들은 각각

$$\begin{aligned} Q^{CL\alpha F_{\alpha}} &= Q^C Q^{L\alpha} Q^{F_{\alpha}} \\ Q^{F_{L\alpha}VL\alpha F_{\alpha}} &= Q^{F_c} Q^{L\alpha} Q^V Q^{L\alpha} Q^{F_z} \\ Q^{F_{\alpha}L\alpha Z} &= Q^{F_{\alpha}} Q^{L\alpha} Q^Z \end{aligned}$$

와 같다. 여기서 F_c 와 F_z 는 각각 임의의 초성, 중성과 정합할 filler 모형이며, F_{cv} 와 F_{vz} 는 각각 연결획을 포함한 초성-중성, 중성-중성 부분과 정합할 filler 모형이다. 그리고 L_{cv} , L_{vz} 는 각각 초성-중성 사이, 중성-중성 사이의 연결획 모형을 나타낸다. 한편 각 경로의 길이를

$$\begin{aligned} |Q^C| &= \tau_c \\ |Q^V| &= \tau_v, |Q^{F_{L\alpha}}| = t_v - 1, \\ |Q^Z| &= \tau_z, |Q^{F_{\alpha}L\alpha}| = t_z - 1 \end{aligned}$$

라고 하자. 여기서 Q^Y 는 모형 Y 에 배열된 부분 경로를

의미하며 $|Q^k|$ 는 그 경로의 길이를 나타낸다. 그러면, 글자의 확률은

$$P(X|CVZ) = \max_{c,v,z} (\delta_{i,c}^c + \delta_{i+1,c}^{L_{cv}} + \delta_{i+1,c}^v + \delta_{i+2,c}^z + \delta_{i+2,c}^{L_{cv}} + \delta_{i+2,c}^z) \quad (8)$$

와 같이 쓸 수 있다. 모형 $K \in \{C, V, Z\}$ 에 대해서 δ_{i_1, i_2}^K 는 $[i_1, i_2]$ 구간의 부분 심볼열에 대한 K 의 확률

$$\prod_{i=i_1}^{i_2} a_{q_{i-1}} \cdot b_{q_i}(x_i)$$

를 나타낸다. 한편 $K \in \{L_{cv}, L_{cz}\}$ 일 때에는 대응 연결획 모형의 확률을 가리킨다.

모든 가능한 조합을 고려한 탐색 공간의 크기는 $|C||V||Z| = 19 \times 21 \times 27 = 11,172$ 에 달한다. 한편 식 (8)에 따르면 자소를 제외한 나머지 부분인 연결획 영역에 대하여도 계산을 해야 한다. 자소 탐색 결과 하나의 경로가 선택되면 키자소와 함께 앞뒤에 접속된 연결획 모형도 결정된다. 그림 9와 같이 초성 탐색 결과초성 모형과 filler 모형 간의 연결획 모형로 L'_{cv} 가 결정되었고, 중성 탐색의 결과 선행 filler 모형과의 사이에 L'_{cz} 가 결정되었다고 하자. 일반적으로 $L'_{cv} \neq L'_{cz}$ 이며, 각 연결획의 경계 또한 서로 차이가 난다. 따라서 자소 탐색 단계의 계산 결과는 무용지물이 된다. 하지만 일반적으로 자소의 경계는 각 자소 별로 서로 완전히 다른 것이 아니라 획의 시작과 끝부분에 집중되는 경향이 있다. 따라서 각 연결획의 경계별로 계산 결과를 기억하여 재사용함으로써 계산량을 어느 정도 줄일 수 있을 것이다.

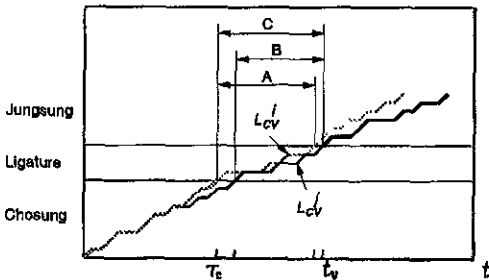


그림 9 자소 탐색 결과 연결획 경계의 차이(A, B)와 실제 경계(C); 점선은 초성 탐색 결과이고 실선은 중성 탐색 결과이다. 짧은 점선과 실선은 각각 초성의 영역, 중성의 영역을 나타낸다.

한글 완성형만을 대상으로 한 경우에는 완성형 글자에 포함 여부를 검사하는 여과 과정도 격자 탐색 과정에 엮어 넣도록 한다. 격자 탐색 자체는 동적 프로그래밍 기법 또는 A* 알고리즘 등 그래프 탐색에 쓸 수 있는 임의의 알고리즘으로 간단히 해결된다.

6. 실험 결과

6.1 필기 데이터와 모형 훈련

HMM 훈련을 위해서 자연스럽게 쓴 필기와 흘려쓴 필기 총 40명분을 사용하였다. 필기 데이터는 필순은 지카되 그외의 특별한 제약은 가하지 않고 자연스럽게 쓰도록 하였다. 글자 인식 실험에는 자연스럽게 필기 8인분(NRM)과 흘려쓴 필기 10명분(CRS), 각각 총 1,700자, 8,100자를 사용하였다.

각 글자를 글자 크기에 비례한 등간격으로 다시 샘플링한 다음, 각 점쌍을 16 방향 체인 코드로 변환하였다. 자소별로 분할한 데이터를 이용하여 각 HMM을 forward-backward 알고리즘으로 훈련하였다[11]. 훈련 반복은 100회 이내에서 끝내도록 하였다. 그리고 HMM은 상태 전이에서 심볼 출력하는 타입의 모형을 채택하였다.

6.2 글자 인식

글자 인식 결과는 표 2와 같다. 쉽게 예상할 수 있는 대로 NRM에서 보다 높은 인식을 86.71% (5개의 후보 인식률은 92.94%)를 얻었고, CRS에서는 77.61% (5개의 후보 인식률 87.94%)의 인식률을 얻었다. 이것은 흘림이 적은 경우 자소의 경계 탐색 오류가 발생할 가능성이 낮기 때문이라 판단된다.

표 2 자소-격자 탐색에 의한 인식률과 BongNet의 성능의 비교

탐색 알고리즘	데이터		비교
	NRM	CRS	
격자 탐색	86.71 %	77.61 %	기본 모델
	87.78 %	78.63 %	평활화 모델
비터비 탐색 (DP)	87.43 %	88.10 %	기본 BongNet
	93.30 %	90.41 %	평활화 BongNet

HMM 네트워크를 이용한 연구[2]에서는 심볼 출력 확률 분포를 평활화하여 인식 성능에 있어서 큰 효과가 있음을 보였다. 그 이유는 훈련 데이터가 충분치 못하여 출력 분포 파라미터가 훈련 데이터에 상당히 편향되어 있었기 때문이라고 설명하였다. 본 연구에서도 마찬가지로

로 심플 출력 분포를 평활화 하여 크지는 않지만 NRM, CRS 각각 8.05%, 4.56%의 에러 감소율을 얻을 수 있었다.

한편 본 논문의 방법에 따른 인식기의 성능을 동적 프로그래밍 탐색을 이용한 기존의 글자 네트워크 모형의 방법(표 2의 하단)과 직접 비교했을 때, NRM, CRS 각각 상당한 차이를 보여주고 있다. 평활화한 모형의 경우 각각 5.97%, 11.78%의 인식을 차이가 난다. 특히 흘림 필기 데이터 CRS에서 큰 격차를 보이는 것은, 흘려쓴 경우 자소의 경계를 찾기가 쉽지 않기 때문이라고 판단된다. 실제로, 자소 분할이 불완전하다고 해서 꼭 오인식으로 이어지는 것은 아니지만, 그림 3과 같은 자소 경계의 오류가 흘림체에서는 빈번하게 나타나게 되고 그 영향을 받는 것은 당연하다. 따라서 향후 정확한 자소 경계를 탐색하는 문제에 대해 보다 집중적으로 연구가 뒤따라야 할 것이다. 본 실험에 사용한 연결획 모형은 기존 연구 [2]의 연결획 모형을 그대로 사용하였다. 그 연결획 모형은 네트워크 구조에 이미 적용되어 있기 때문에 새로이 연결획 모형에 대한 고려가 있어야 할 것이다.

한편 글자 인식에 들어가는 계산량을 분석해보기로 하자. N 을 HMM의 상태수이라고 할 때 DP 계산량은 일반적으로 N^2T 가 된다. 실험에 사용된 연결획 모형의 상태수는 $N_L=3$ 으로 고정되어 있다. 반면 연결획을 제외한 자소 모형의 상태 수는 대상 자소의 복잡도가 다양하여 그림 10과 같이 넓은 분포를 갖는다.

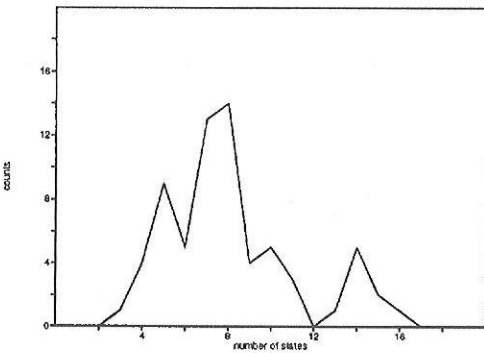


그림 10 자소 HMM의 상태수 분포

자소 모형의 평균 상태수를 \bar{N} 이라고 하자. 그리고 filler 모형의 (평균) 상태수를 N_F 라고 하자. 입력의 길이를 T 라고 하면 자소 하나를 탐색하는데 필요한 계산

량은 $T(\bar{N}^2 + lN_L^2 + N_F^2)$ 가 된다. l 은 자소 탐색 모형 하나당 평균적으로 들어가는 연결획 모형의 수이다. 총 탐색 자소의 수는 $M = |C| + |V| + |Z|$ 이므로 글자 인식에 들어가는 계산량은 대략 $MT(\bar{N}^2 + lN_L^2 + N_F^2)$ 가 된다. 참고로 BongNet[1]에서는 자소와 연결획 들의 연결에 의한 망 구조로 이미 글자 구조에 관한 추가의 정보를 활용하고 있는 셈이다. 그래서, L 을 연결획 모형의 수라고 하면, 총 계산량은 $MT\bar{N}^2 + LTN_L^2$ 로 최소화 되어 있다.

한편 일반적으로 이용되는 HMM은 우향 전진 구조에서는 계산량이 $MTk(\bar{N} + lN_L + N_F)$ 가 된다. k 는 상태당 천이의 갯수이다. $\bar{N} + lN_L + N_F$ 는 자소 탐색 모형의 크기(상대수)에 해당한다. 이를 N_{spot} 이라고 하면 $MTkN_{spot}$ 으로 쓸 수 있다. 비슷하게 BongNet에서는 $kT(M\bar{N} + LN_L)$ 이 된다. 그리고 실제로 실험에 쓰인 두 인식기 간의 계산량 비는

$$\begin{aligned} \frac{\text{자소탐색기법}}{\text{BongNet}} &= \frac{MN_{spot}}{M\bar{N} + LN_L} \\ &= \frac{3.7M\bar{N}}{1.2M\bar{N}} \\ &= 3.2 \end{aligned}$$

가 된다.

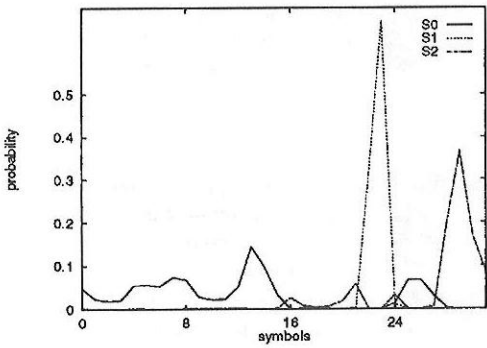
앞 실험에서는 키 자소 모형과 filler 모형 사이에 연결획 모형을 4 ~ 5개를 두었다. 연결획 모형을 하나만 두었을 때의 실험 결과는 표 3과 같다.

표 3 연결획 모형의 갯수와 훈련 여부에 따른 성능의 변화

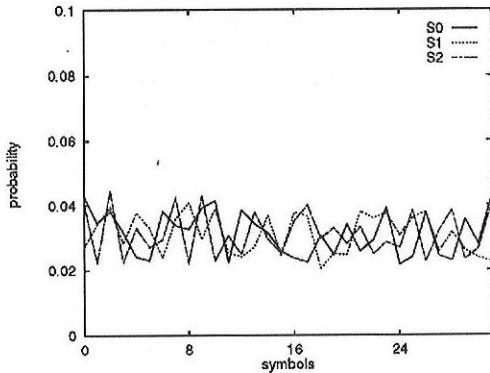
연결획 모형	NRM	CRS
문맥 종속 연결획 HMM	87.78 %	78.63 %
훈련 안된 단일 연결획 HMM	21.89 %	21.41 %
훈련된 단일 연결획 HMM	24.79 %	28.97 %

여기서 연결획 모형이 하나라는 것은 모든 초성과 모든 중성 사이의 연결 패턴을 하나의 모형에 대응 시키는 것이다. 따라서 데이터의 패턴 상으로는 변이가 매우 많기 때문에 정확한 모형화를 기대하기는 어렵다. 평활화된 자소 모형과 함께 엮은 탐색 모형을 실험한 결과, 데이터 NRM에 대하여 훈련 안된 연결획 모형과 훈련된 연결획 모형은 각각 21.89%, 24.79%의 인식률에 불과한 성능을 얻었다. 이렇게 낮은 성능은 HMM 한개로 여러 변형을 충분히 모형화할 수 없을 뿐만 아니라, 각 키자

소 별로 연결획의 모양, 방향에 대한 제약을 전혀 가할 수 없기 때문이다.



(a)



(b)

그림 11 초성-중성간 연결획 모형 각 상태의 심볼 출력 확률 분포 (상태수는 3개이며, S0은 첫번째 상태를 가리킨다): (a) 훈련된 단일 연결획 HMM, (b) 훈련 안된 단일 연결획 HMM의 경우

그림 11은 훈련 안된 연결획 모형과 훈련된 연결획 모형의 상태의 심볼 확률 분포를 도시한 것이다. 훈련된 모형(그림 11(a))의 경우에는 심볼별로 상태가 구분되어 있으며, 세개의 상태가 모든 심볼을 나누어서 모형화하고 있기 때문에 순차적으로 입력되는 심볼열 패턴에 제약이 가해진다. 이는 곧 훈련이 안된 모형과 변형 수용 능력의 측면에서 큰 차이가 있음을 의미한다. 이것은 또 키 자소의 문맥(context)에 따라 분류한 문맥 종속형 연결획 모형 (context-dependent HMM)[2]의 필요성

을 말해주고 있다.

홀려 쓴 필기에서 자소 탐색의 결과의 예는 그림 12와 같다. 각 자소의 경계점은 탐색 모형 확률 파라미터 분포에 따라서 확률적으로 결정되는데, 그림과 같이 홀려 쓴 필기에서 경우 정확한 자소 경계를 찾는 것을 거의 불가능하다고 할 때 (특히 중성-중성간의 경계의 경우), 본 방법으로 찾아낸 경계점은 개연성이 매우 높다고 판단된다.

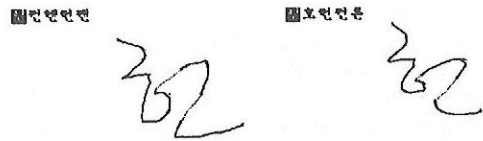


그림 12 글자 인식 및 자소 분할의 결과의 예

7. 결 론

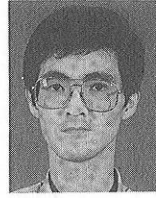
자소 위치 탐색에 의한 온라인 한글 필기 인식 방법은 기본적으로 자소 탐색의 성능에 따라서 그 성패가 결정된다고 해도 과언이 아니다. 따라서 본 논문에서는 최적의 자소 경계를 찾기 위한 자소 탐색 모형의 설계와 탐색 알고리즘의 연구에 큰 비중을 두었다. 단순한 자소와 연결획 모형의 연결이 아니라, 보다 분별력을 갖는 탐색 네트워크를 구성하기 위해서 자소의 연결 형태에 따라 문맥 종속형 연결획 모형을 사용한 점이 본 방법의 핵심이라 할 수 있다. 현재 자소 탐색시 자소 가설을 한 개씩만 생성한다. 그리고 격자 탐색에서는 자소 가설 간에 약간의 중첩이라도 허용하지 않는다. 따라서 앞으로 다중 후보를 생성토록 하고, 중첩을 허용토록 알고리즘에 수정을 가한다면 인식기에 유연성을 더할 수 있을 것이라고 판단된다.

참 고 문 헌

- [1] 신봉기, 김진형, “통계적 방법에 의한 온라인 한글 필기 인식”, 제 4회 한글 및 한국어 정보처리 학술대회 논문집, pp. 533-542, 1992. 10.
- [2] 신봉기, 김진형, “은닉 마르코프 모델 네트워크에 의한 온라인 흘림 필기 한글 인식”, 정보과학회 논문지, 제 21권, pp. 1737-1745, 1994. 9.
- [3] 이희동 외, “확장된 DP 매칭법에 의한 흘림체 한글 온라인 인식”, 대한 전자공학회 논문지, 제 15권 1호, pp. 29-37, 1989. 1.
- [4] F. R. Chen, L. D. Wilcox, and D. S. Bloomberg, “Word spotting in scanned images using hidden

Markov models," in Proc. IEEE ICASSP, pp. 1-4, 1993.

- [5] G. E. Kopec and P. A. Chow, "Document image decoding using Markov source models," in Proc. IEEE ICASSP, pp. 85-89, 1993.
- [6] S.-S. Kuo and O. E. Agazzi, "Machine vision for keyword spotting using pseudo 2D hidden Markov models," in Proc. IEEE ICASSP, pp. 81-84, 1993.
- [7] J. G. Wilpon, L. R. Rabiner, C.-H. Lee, and E. R. Goldman, "Automatic recognition of keywords in unconstrained speech using hidden Markov models," IEEE Trans. on Acoustics, Speech and Signals Processing, vol. 38, pp. 1870-1878, Nov. 1990.
- [8] S.-H. Lee and J. H. Kim, "On-line cursive script recognition by a letter spotting technique based on HMMs," in Proceedings of the 2nd Workshop on Character Recognition, (Seoul, Korea), pp. 93-104, Sep. 1994.
- [9] J. J. Wolf and W. A. Woods, "The HWIM speech understanding system," in Proc. IEEE ICASSP, (Hartford, Conn.), pp. 784-787, 1977.
- [10] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," Proc. IEEE, vol. 77, pp. 257-285, Feb. 1989.
- [11] L. E. Baum, "An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes," Inequalities, vol. 3, pp. 1-8, 1972.
- [12] A. J. Viterbi, "Error bounds for convolution codes and an asymptotically optimum decoding algorithm," IEEE Trans. on Info. Theory, vol. 13, pp. 260-269, 1967.
- [13] K. F. Lee, "Context-dependent phonetic HMMs for speaker-independent continuous speech recognition," IEEE Trans. on Acoustics, Speech and Signals Processing, vol. 38, pp. 599-609, Apr. 1990.
- [14] L. R. Bahl, P. Brown, P. V. de Souza, and R. L. Mercer, "Maximum mutual information estimation of hidden Markov model parameters for speech recognition," in Proc. IEEE ICASSP, (Tokyo), pp. 49-52, 1986.
- [15] B.-H. Juang, "Discriminative training," J. Acoust. Soc. Jpn, vol. 13, no. 6, pp. 333-339, 1992.
- [16] R. A. Sukkar and J. G. Wilpon, "A two pass classifier for utterance rejection in keyword spotting," in Proc. IEEE ICASSP, pp. 451-454, Apr. 1993.
- [17] G. D. Forney, "The viterbi algorithm," Proc. IEEE, vol. 61, pp. 268-278, Mar. 1978.



신 봉 기

1985년 서울대학교 공과대학 자원공학과 졸업. 1987년 한국과학기술원 전산학과 졸업(공학 석사). 1995년 한국과학기술원 전산학과 졸업(공학 박사). 1987년 ~ 현재 한국통신 근무. 관심분야는 패턴 인식 및 모델링, 지능형 에이전트.



김 진 형

1971년 서울대학교 공과대학 졸업. 1979년 UCLA 전산학 석사학위 취득. 1983년 UCLA 전산학 박사학위 취득. 1973년 ~ 1976년 KIST 전산실 연구원. 1976년 ~ 1977년 미국 캘리포니아 도로국 연구원. 1981년 ~ 1985년 미국 Hughes 인공지능 센터 선임 연구원. 1986년 ~ 1988년 한국정보과학회 산하 인공지능 연구회 위원장. 1989년 ~ 1990년 미국 IBM 와트슨 연구소 초빙 과학자. 1985년 ~ 현재 한국과학기술원 전산학과 교수. 1991년 ~ 현재 인공지능 연구센터 부소장. 1995년 9월 ~ 현재 연구개발정보센터(KORDIC) 소장. 관심분야는 인식 시스템, 지능형 인터페이스, Computer-Aided Education, Groupware, 지식기반 시스템.