

통계적 방법에 의한 온라인 한글 필기 인식

신 봉 기, 김 진 형
한국과학기술원 전산학과

On-line Handwritten Character Recognition with Hidden Markov Models

Sin, Bongkee, and Kim, Jin Hyung
CS Dept., KAIST

요 약

손으로 쓴 글씨는 인쇄체와 달리 많은 변형이 있다는 점이 한글 필기 인식에서 가장 큰 장애물로 통한다. 본 논문에서는 이점을 해결하면서 필기에 대한 제한을 대폭 줄인 온라인 한글 인식 방법을 제시하고자 한다.

봉넷(BongNet)은 온라인 한글 필기를 인식하기 위한 네트워크 모델이다. 글씨 인식에 들어가는 여러가지 정보를 네트워크라는 틀 안에 표현한 것인데, 기본적으로 네트워크 구조 자체가 표현하는 정적 글자 구조 정보와, 글꼴에 따라 달라지는 것으로써 노드간 확률적 이동을 나타내는 동적 정보를 포함한다.

본 모델에 따르면 한글 인식은 네트워크 안에서 최적 경로를 따라 초, 중, 종성 자소열을 찾는 문제로 변환된다. 동적 프로그래밍 기법을 이용하여 그 경로를 찾는 인식 알고리즘은 입력 데이터의 양에 정비례하는 효율성을 갖는다.

I. 서 론

한글에는 영어처럼 특별한 필기체는 없지만 글자의 종류나 필기자의 습관에 따라서 거의 반사적으로 두 획이 붙거나 또는 그외 여러가지로 변형이 생긴다. 그러한 글씨를 자소 분리형 인식기로 인식하기 위해서는 우선 어디까지가 한 자소인지를 알 필요가 있는데, 일반적으로 인식이 안된 상태에서 자소를 완벽하게 분리하는 것은 거의 불가능하다. 그리고 필기의 모양을 분석해보면 서로 다른 글자 또는 자모가 흐리는 정도에 따라 비슷한 꼴이 되는 경우가 많이 발생한다. 이것은 손운동의 물리학적 제약에 따른 필연적 흘림과, 언어 자체에 내재하는 본질적인 모호성에서 그 원인을 찾을 수 있다. 한편 음절을 기본 단위로 한 인식기는 비록 어떤 방법으로든 가능은 하겠지만 한글에는 만자가 넘는 글자가 있다는 점을 고려할 때 무리한 방법임에 틀림없다. 그러므로 완벽한 필기 글자 인식기는 불가능하다고 알려져 있다.[1]

종래의 한글 필기 인식 방법은 글자꼴 모델을 갖고서 입력 글씨와 비교하는 방법이 주를 이루고 있으며 [2, 3, 4], 최근에는 신경망을 이용한 방법도 나타나고 있다 [5, 6]. 전자의 방법은 여러가지 글자 꼴을 기억하고 있다가 그중에 가장 비슷한 것으로 판단을 한다. 그런데 사람마다 필기 습관에 따라 글씨가 천차만별이므로 글자꼴 모델의 수가 크게 늘어나게 되는 문제점이 있다. 그러나 그보다 더 큰 문제는 대부분의 인식기가 자소 별로 모델을 갖고 있기 때문에 두개의 자소를 붙여 쓰는 경우가 많은 보통의 필기를 인식하는 것은 불가능하지는

않다 하더라도 큰 무리가 따른다는 점이다. 그리고 신경망을 이용한 방법은 속도가 지나치게 느리다거나 인식 가능한 글자의 수가 제한되어 있다는 문제점을 가지고 있다.

본 논문에서는 위에서 설명한 필기 인식의 어려움을 극복하여 필기시의 제약을 가능한 최소로 줄여 줄 수 있는 한글 인식 모델을 제안하고자 한다. 봉넷(BongNet)은 온라인 한글 필기를 인식하기 위한 통계적 네트워크 모델이다. 글씨 인식에 들어가는 여러가지 정보를 네트워크라는 틀 안에 표현한 것인데, 기본적으로 네트워크 구조 자체가 표현하는 정적 글자 구조 정보와, 글꼴에 따라 달라지는 것으로서 노드간 확률적 이동을 나타내는 동적 정보를 포함한다. 전자는 한글 음절을 모델링하는 인식 네트워크 자체를 가리키며, 후자는 자소 모델이 표현하는 필기의 모양, 필순, 그리고 여러가지 변형을 말한다. 흘림에 의한 변형 모델링하는 수단으로 '연결 자소'를 제안하였고, 각 자소는 은닉 마르코프 모델로 표현하였다.

본 논문의 구성은 다음과 같다. II 장에서 마르코프 모델의 기본 개념을 소개하고, 그 다음 III 장에서 마르코프 모델로 각 자소를 표현하는 모델로 쓰일 수 있음을 보인다. 그리고 동 장에는 자소의 의미를 필기체에 초점을 맞추었을 때 변형 함수와 일관성 있는 계산 및 표현 수단으로써의 연결 자소 모델을 제안한다. 이렇게 확장된 자소 모델을 결합하여 설계한 한글 인식 네트워크 봉넷(BongNet)을 간략히 소개하고, 이어 IV 장에서 봉넷을 이용해서 손으로 쓴 글씨를 하나의 글자로 인식하는 방법을 설명한다. 다음 V 장에서는 본 네트워크 모델을 실제 데이터로 평가해본다.

II. 마르코프 모델

1. 마르코프 모델

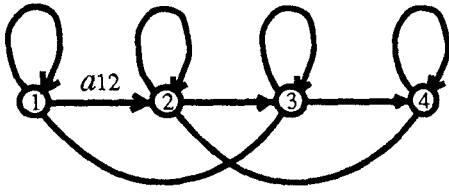
마르코프 모델은 유한개의 상태(또는 노드)와 각 상태 사이를 방향성 있게 연결하는 전이의 집합으로 된 유한한 크기의 네트워크 구조로 설명할 수 있다. 이때 확률 변수로서의 상태 X_t , $t = 0, 1, 2, \dots$ 는 다음과 같은 마르코프 성질을 만족하는 확률 시퀀스로 정의한다.

$$P(X_{t+1}=j|X_0=a, X_1=b, \dots, X_t=i) = P(X_{t+1}=j|X_t=i)$$

이것은 앞으로 상태간 전이를 반복해서 도달될 어떤 상태의 확률이 과거를 불문하고 오직 직전 상태에만 달려 있다는 것을 의미한다. 전이 확률이라고 불리우는 이 값은 $a_{ij}(t+1)$ 로 나타내기로 한다. 마르코프 모델은 또 마르코프 체인 이라고도 부르는데, 보통 그림-1과 같이 표현한다. 이 그림을 보면 노드간 전이 확률을 간단히 상태 전이 확률 행렬 $A(t) = \{a_{ij}(t)\}$ 로 표현할 수 있음을 알 수 있다. 이러한 모델에 따르면 미래의 각 상태는 이전 시간의 모델 상태에 따라서 확률적으로 예측할 수 있게 된다.

$$P_j(t+1) = \sum_i P_i(t) a_{ij}(t+1), \quad j = 1, 2, \dots, N$$

그런데 노드 S_j 의 확률 $P_j(t+1)$ 의 값이 위식과 같은 현재의 모델 상태 외에 또다른 어떤 확률적 변수에 좌우된다고 하자. 그것을 현재 상태의 어떤 확률 함수 $b_j(t+1)$ 라고 하면 각 노드의 확률은



(a)

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix}$$

(b)

그림-1. 자소 은닉 마르코프 모델의 구조

$$P_j(t+1) = \sum_i P_i(t) a_{ij}(t+1) b_j(t+1), \quad j = 1, 2, \dots, N$$

로 정의할 수 있다. 이때의 모델은 두가지의 확률 변수에 따라 좌우되는 복합 구조를 가지는 마르코프 모델이 되는데, 이것을 은닉 마르코프 모델이라 부른다. 은닉 마르코프 모델에서는 새로운 확률 변수 $b_i(t)$ 를 출력(또는 관찰) 심볼 확률이라고 한다. 그리고 $a_{ij}(t) = a_{ij}$, $b_j(t) = b_j$ 로써 시간에 대해서 무관한 경우 이를 시불변 은닉 마르코프 체인이라 하며, 특별한 구별이 없는 한 앞으로 이러한 모델을 의미한다.

2. 은닉 마르코프 모델

은닉 마르코프 모델은 Baker [7]가 처음으로 연속 음성 인식 문제에 적용할 수 있음을 보였는데, 모델 자체는 그보다 앞서 Baum 등 [8]에 의하여 효율적인 알고리즘이 출현하면서 부터 가능성을 갖고 있었다. 은닉 마르코프 모델은 앞에서 언급한 바와 같이 마르코프 체인과 그 함수로서 관찰 심볼 확률 분포로 정의할 수 있는, 다시 말하면 확률적 마르코프 체인이라고 한마디로 정의할 수 있다. 일반적으로 은닉 마르코프 모델은 문헌의 관습에 충실하자면 다음과 같이 정의한다. [9]

(1) 상태 전이 확률 분포 $A = \{a_{ij}\}$. 각 확률은

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

로 정의한다. a_{ij} 는 임의 시각에 상태 S_i 에 있을 때 다음 순간에는 S_j 로 전이할 확률을 말한다. 만약 $a_{ij} > 0$ 이면 S_i 에서 S_j 로 한번의 전이로 갈 수 있다는 것을 의미한다. 그렇지 않으면 해당 한번의 전이로 직접 갈 수 없음을 의미한다. 그리고 이들 확률 파라미터는 다음과 같은 조건을 만족한다.

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

(2) 관찰 심볼 확률 분포 $B = \{b_i(k)\}$ 로써 다음과 같이 정의된다.

$$b_i(k) = P(v_k | q_t = S_i), \quad 1 \leq i \leq N,$$

$$1 \leq k \leq M.$$

$b_i(k)$ 는 i 시각에 S_i 에서 k 번째 심볼 v_k 를 관찰할 확률을 말하며 조건

$$\sum_k b_i(k) = 1, \quad 1 \leq i \leq N$$

을 만족해야 한다.

(3) 초기 상태 전이 확률 분포 $\Pi = \{\pi_i\}$. 이것은 파라미터 A 의 특별한 경우로 볼 수도 있는데, 초기 상태 $i=0$ 일 때 각 상태 S_i 의 확률 분포를 나타낸다.

$$\begin{aligned} \pi_i &= P(q_1 = S_i), & 1 \leq i \leq N, \\ \sum_i \pi_i &= 1. \end{aligned}$$

그리고 이 밖에 모델의 기본 구조를 정의 하는 상태수 N 과 모델에서 관찰할 수 있는 M 개 심볼의 집합 $V = \{v_1, v_2, \dots, v_M\}$ 가 있다. 하지만 이중에서 위에서 설명한 동적인 특성을 기술하는 모델 파라미터만으로 다음과 같이 간단히 표현하는게 보통이다.

$$\lambda = (A, B, \Pi).$$

은닉 마르코프 모델은 시간적 제약을 받는 정보를 표현하는데 뛰어나다. 단순한 마르코프 체인으로는 기술하기 힘든 실제계의 문제에 보다 효과적으로 적용할 수 있는 모델이다. 시간적인 제약에 절대적으로 지배를 받는 음성 신호를 처리하고 인식하는데 가장 성공적인 대표적인 통계적 모델이다. 본 논문에서는 한글의 자모를 각각 은닉 마르코프 모델로 표현한다.

III. 한글 인식 모델

1. 자소 모델

다음절에서 설명될 한글 인식 네트워크는 한글 음절 일반을 표현하는 모델로 볼 수 있다. 그기본 구성 요소는 자소 하나하나에 해당하는 자소 모델이다. 한글에는 초성 19자, 중성 21자, 그리고 종성 27자가 있는데 이들의 순서적 결합에 따라 글자가 만들어진다. 총 67개의 자소 모델을 가지면 자소별로 인식은 가능할 것이다. 하지만 한글 음절을 인식하는 데는 자소별로 독립된 인식 결과만으로는 불충분하다. 한 글자는 초, 중, 종성이 대체로 네모 안에 상하좌우 등 공간적으로 적절하게 균형을 이루어 배치된 구조를 갖고 있는데 이러한 구조적 정보가 중요하다는 것은 이미 여러 곳에서[2, 3] 지적한 바와 같다.

한편 필기 습관에 따라서 이웃하는 두 획이 상호 작용으로 인해 붙거나 하는 등의 변형이 발생하게 된다. 변형은 획의 순서와 배치 관계에 따라 일정한 변형 패턴이 있는 것과, 한글 필순에 들어 맞지 않고 개인적 편리나 습관에 따른 부정형 변형이 있다. 전자의 변형은 기본적으로 손의

움직임을 최소화하려는 경제성 원리와 손의 움직임 자체가 받는 물리 법칙에서 그 원인을 찾을 수 있다. 이러한 동적인 변형은 다른 글자와 구분할 수 있도록 최소한의 특징을 갖게하는 언어적 요소로 인하여 그 정도가 제한되기도 한다. 한편 부정형 변형은 개인의 개성이라고 말할수 있는 것으로써 사람에 따라 글씨의 모양이나 필순이 달라지는 것을 말한다. 온라인 필기에서는 입력 데이터와 함께 시간적 정보, 즉 필순과 각 획의 시작, 종료점 등을 얻을 수 있는데, 이점은 오프라인 인식과 구별되는 특징으로써 온라인 필기 인식문제의 시발점이라 할 수 있다.

자소 모델을 표현하는 은닉 마르코프 모델은 이러한 시간적 제약을 받는 정보를 모델링하는데 뛰어나다. 현재 구성된 한글 인식기의 자소 모델의 구조는 그림-1.a 와 같으며, 노드수는 자소에 따라 달라 4~16 개 정도이다.

2. 연결 자소 모델

보통의 필기에는 차례로 씌어지는 두 획 사이에 원래 없었던 부분이 종종 들어가 두 획을 연결시켜 한 획으로 만든다. 이를 연결획이라고 부르자. 이것은 경제성 원리에 따라서 연필을 들고 놓는 손의 움직임을 줄여서 보다 빨리 필기하기 위한 노력의 결과이다. 이러한 연결획과 함께 정상적으로 연필을 종이에서 들어 옮긴 꺾어진 가상획을 확장된 연결획으로 정의하고, '연결 자소'로써 표현한다. 연결획과 가상획은 모양은 달라도 방향과 그 역할은 동일하다. 일반적으로 임의의 두 자소 사이의 연결획은 대체로 일정한 방향과 길이를 갖는다. 연결 자소 모델은 이러한 특징을 표현하여 한글 필기의 흘림과 그에 따르는 변형을 모델링하기 위한 것이다. 연결 자소의 또다른 역할은 평균적인 연결획 모양 - 특히 방향과 길이 - 으로 자소와 자소 사이의 공간적 배치를 표현한다는 것이다.

3. 한글 인식 네트워크

앞에서 설명한 자소 모델을 단위로 해서 완전한 한글자를 인식하기 위한 틀이 한글 인식 네트워크 봉넷(BongNet)이다. 그림-2 에서 보여주는 봉넷은 한글 제자 원리를 네트워크 모델로 구현한 것인데, 한글의 모든 가능한 음절을 인식할 수 있다는 점에서 인식기가 될 수도 있고 한글 음절 모델이라고 볼 수도 있다. 글자는

$$\begin{aligned} &<초성> + <중성> \\ &<초성> + <중성> + <종성> \end{aligned}$$

의 구조를 가진다. 그리고 자소 사이의 연결획을 고려하면

$$\begin{aligned} &<초성> + <L> + <중성> \\ &<초성> + <L> + <중성> + <L> + <종성> \end{aligned}$$

와 같은 확장된 자소열로 표현할 수 있다. 이와 같은 자소열을 묶어서 간결하게 표현한 것이 봉넷이다. 그림의 네트워크를 보면 한글자를 인식하기 위해서는 시작 노드에서 출발하여 종료 노드에 이르기 까지 여러 경로가 있다. 각 경로는 자소간이 연결 패턴이 유사한 글자들을 모델링 하는데, 각 경로상의 최적의 자소들의 연결로 인식을 하게 된다. 예를 들어 '랑'을 인식하기 위해서는

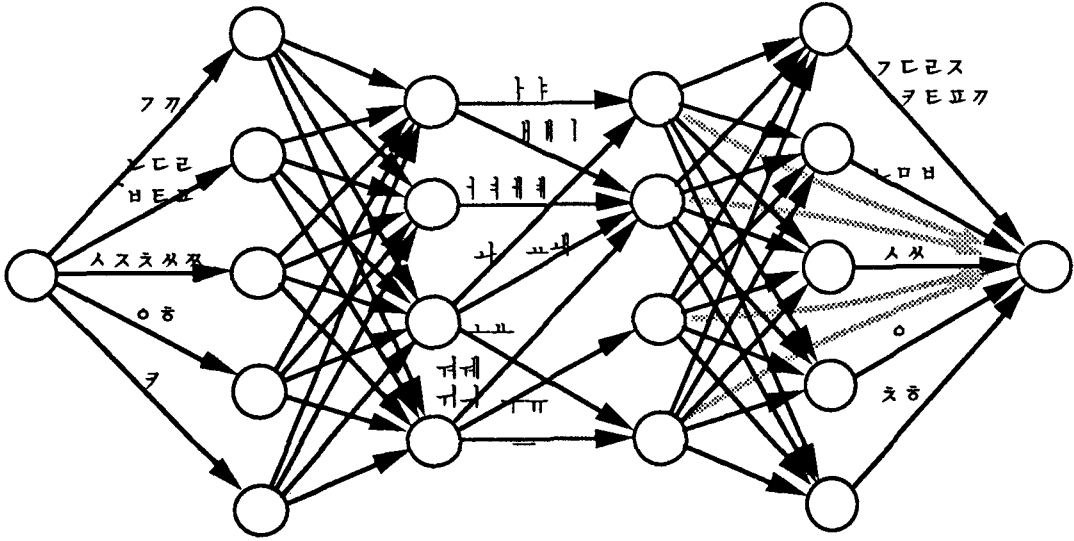


그림-2. 붕넷 - 한글 인식 네트워크

<ㄱ> + <ㄴ> + <ㅇ> + <ㅇ> + <ㅇ>

에 해당하는 경로가 있어야 하며, 그 경로상에서 각 자소(ㄱ, ㄴ, ㅇ)의 확률값이 주어진 글씨에 대해서 통계적으로 확률이 가장 높아야 한다. 그리고 그러한 경로를 찾을 수 있어야 한다.

IV. 한글 인식

1. 전처리 및 코드화 과정

전자펜으로 태블릿 위에 쓴 글씨는 일반적으로 전자 기계적인 결함으로 잡음이 포함될 소지가 있다. 그리고 부정확한 손의 움직임 때문에 나타나는 것도 있는데, 대표적으로 불필요한 점과 삐침을 들 수 있다.(그림-3.a) 이러한 잡음을 최대한 줄여서 사람이 원래 의도한 형태의 글씨로 가능한 한 가까이 변환시켜주는 과정을 전처리라고 한다. 전처리 기법으로 특별히 정형화된 이론이 있는 것은 아니며 휴리스틱 성격이 강하다.

일단 깨끗한 글씨를 얻었으면 인식기가 처리할 수 있는 데이터 형태로 변환을 해주어야 한다. 이 과정에서 가장 많은 정보를 잃기 쉽다. 만약 변환 결과 얻은 데이터로부터 원래 글씨를 거의 그대로 만들 수 있다면 정보 손실은 거의 없다고 말할 수 있다. 하지만 일반적으로 변환과정에서의 정보 손실은 어쩔 수 없고, 최소한으로 줄이는 것이 최선이다.

일반적으로 전자펜으로 필기한 데이터는, 예를 들어, 백분초 단위로 샘플링한 점좌표(x,y)의 열로 표현된다. 그런데 손으로 쓴 글씨는 획의 모양에 따라 부분별로 속도가 달라서 점들이 고른 간격으로 샘플링되지 않는다. 이런 현상에도 의미가 있긴 하지만 글씨의 모양을

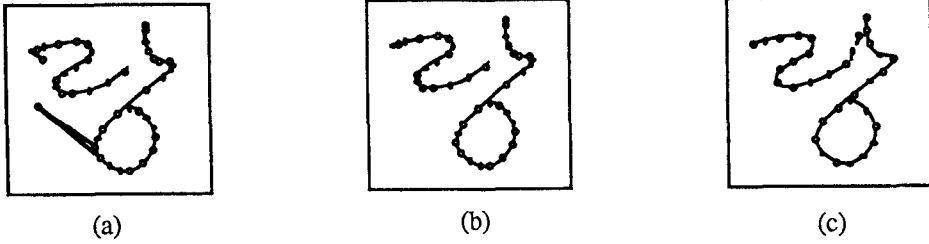


그림-3. 전처리 과정

분석하는데 중복 계산을 강요하게 된다. 그래서 형태 분석에 더 적합하도록 하기 위하여, 시간적으로 샘플링된 점열 글씨를 바탕으로 공간적 재(再)샘플링 처리를 한다. 그림-3.c와 같이 재 샘플링된 글씨는 형태상으로 왜곡이 없을 수록 좋다.

점열을 차례로 이은 선분을 코드 변환 단위로 할 때 각각은 이전의 선분 또는 코드가 무엇인지 전혀 상관하지 않는다고 가정하자. 그러면 재샘플링 점열은 곧 Freeman 방향 코드 열과 동등하다. 360도를 여덟 방향으로 나누고 차례로 매긴 번호로 점열이 이루는 각 선분을 표시한다.

한편 획과 획 사이에는 연필이 공간적으로 이동하게 되는데 이러한 연필의 이동 궤적을 가상획으로 정의하여 비슷한 방법으로 코드로 변환한다. 이때 진획과 구별을 위해서 다른 코드 번호를 체계 -- 예를 들어, 8방향 체계이면 진획의 코드값에다 8을 더하는 -- 를 사용한다. 각 가상획의 코드열은 필기한 획 순서에 따라 진획 사이에 끼워 넣는다. 이리하여 한글 음절 하나에 해당하는 글자를 쓰면 코드로 변환했을 때 하나의 코드열을 얻게 된다.

2. 인식 알고리즘

앞절에서 언급한 바와 같이 봉넷 네트워크 모델에 따라서 한글을 인식하는 것은 시작 노드에서 마지막 노드까지의 최적 경로를 찾는 문제로 정의할 수 있다. 여기서 최적의 의미는 입력된 글씨에 대해서 완전한 경로를 단위로 평가를 했을 때 가장 확률이 높은 것을 의미한다.

$$P(X, G), \quad G = g_{\langle \text{左} \rangle} g_{\langle \downarrow \rangle} g_{\langle \text{右} \rangle} g_{\langle \downarrow \rangle} g_{\langle \text{右} \rangle}$$

또는 $g_{\langle \text{左} \rangle} g_{\langle \downarrow \rangle} g_{\langle \text{右} \rangle}$

$$X = \text{입력 코드열}$$

를 경로가 G 이고 그 경로를 따라서 출력 X 가 나올 확률, 다시 말해서 입력 코드열이 경로 G 가 표현하는 글자로 인식될 확률이라고 하자. 그러면 X 를 인식하기 위한 식은 아래와 같이 쓸 수 있다.

$$P(X, G^*) = \max_G P(X, G)$$

$$= \max_G P(X|G)P(G)$$

$$= \max_G \max_D P(X_{d1}, X_{d2}, X_{d3}, X_{d4}, X_{d5} | G_1, G_2, G_3, G_4, G_5)P(G)$$

$$\begin{aligned}
&= \max_G \max_D P(X_{d1}|G_1)P(X_{d2}|G_2)...P(X_{ds}|G_s) \times P(G) \\
&= \max_G \max_D \prod_i P(X_{di}|G_i) \times P(g_{\langle i-1 \rangle})P(g_{\langle i \rangle})P(g_{\langle i+1 \rangle})P(g_{\langle i \rangle})P(g_{\langle i+1 \rangle}) \\
G^* &= \operatorname{argmax}_G \max_D \prod_i P(X_{di}|G_i) \times P(g_{\langle i-1 \rangle})P(g_{\langle i \rangle})P(g_{\langle i+1 \rangle})P(g_{\langle i \rangle})P(g_{\langle i+1 \rangle}) \\
D^* &= \max_G \operatorname{argmax}_D \prod_i P(X_{di}|G_i) \times P(g_{\langle i-1 \rangle})P(g_{\langle i \rangle})P(g_{\langle i+1 \rangle})P(g_{\langle i \rangle})P(g_{\langle i+1 \rangle})
\end{aligned}$$

여기서

X_{di} 입력 코드열을 자소의 수만큼으로 분할했을 때, 부분 코드열
($X = X_{d1}X_{d2}X_{d3}X_{d4}X_{d5}$ 의 연결(ordered concatenation))

G_i 네트워크 경로 상의 한 전이 또는 자소 모델.

$P(X_{di}|G_i)$ 는 주어진 주어진 경로 즉 자소 모델 G_i 가 입력의 일부분 X_{di} 를 모델링하는 확률.

$P(G_i)$ 는 자소의 선행 확률. G_i 가 보통의미의 한글 자소가 아닌 경우 즉 연결확이면 이것은 자소간의 관계를 의미하는 자소간 전이 확률이 된다.

윗식의 결과에 따라 최적의 경로(G^*)를 알 수 있고, 또 주어진 입력에 대해서 그 경로 위의 각 자소 모델과 대응되는 입력의 최적 분할(D^*)을 구할 수 있다.

네트워크에서 주어진 입력에 따라 최적 경로를 찾는 문제는 계산상의 복잡도가 지수 함수이기 때문에 모든 경로를 다 열거한다는 것은 불가능하다. 예를 들어 열개의 (모델수 $M=1$) 숫자로 된 수를 인식한다 할 때, 입력 심볼(코드)의 수가 T 이면 복잡도는 M^T 가 된다. 가령 $T=50$ 이면 10^{50} 이 된다.

Viterbi 알고리즘[9, 10]은 이러한 문제에서 최적해를 효율적으로 얻을 수 있는 알고리즘이다. 최적 원리(Principle of Optimality) [11]에 뿌리를 두고 있으며, 동적 프로그래밍에 의한 음성인식 방법에 널리 사용되는 알고리즘이다 [12]. $t=0$, 시작 노드에서부터 임의 노드에 도달하는 경로의 비용 또는 확률은 직전 노드에서의 비용에다가 거기서 다시 현재의 노드로 전이할 때의 추가 비용을 합하는 방식으로 구한다. 이런 계산을 반복적으로 하게 되는데, 이러한 처리 방법은 마르코프 모델이 전제로 하는 시간적 제약과 잘 맞아 떨어진다. 그리고 계산 복잡도는 M^2T 로써 입력의 길이에 비례하는 효율적인 알고리즘이다.

Viterbi 알고리즘에 따르면 각 노드의 확률(위에서는 비용)은 알고리즘 상으로 여러 가능성 중 최적 또는 최대값으로 결정되는데, 그것은 어떤 시점의 현재 노드까지 오는 경로 중 최적인 것 하나만을 기억한다는 것이다. 즉 다른 가능성은 사정없이 버린다. Viterbi 알고리즘에 따르면, 여러 경로를 끝까지 확장해 봤을 때, 현재 버려진 경로 중에서 어떤 것도 현 시점 현 노드에서의 최한 최적 경로를 끝까지 확장한 최적 경로보다 확률이 높지 못하다.

지금까지 설명한 동적 프로그래밍 기법으로 처리한 결과는 주어진 네트워크 봉넷에서의 최적 경로가 된다. 그리고 그 경로에 있는 모델의 일련의 노드(또는 상태)와 입력 코드열이 최적으로 배열된다. 다시 말하면 네트워크 파라미터에 의해서 코드열이 분할되어 각 노드

에 일대일 대응된다는 것인데, 이를 자소 모델별로 보면 자소의 경계가 자동적으로 해결된다는 것이다.

V. 실험

글자를 흘려 썼을 때에는 또박또박 쓴 자소 분리 필기의 경우 보다 훨씬 다양한 모양을 갖는다. 그러므로 이들을 모두 수용하기 위해서는 보다 많은 보통 필기를 모은 훈련 샘플을 필요로 한다. 한편 연결 자소는 그 변형의 종류에 따라 출현 빈도가 상당히 낮고, 또 희귀 자모의 결합 -- 예를 들면 'ㄷ'와 'ㅈ' 사이 -- 에 들어가는 연결 자소는 더욱 희귀하거나 거의 볼 수 없다. 이것은 자소의 훈련 문제를 매우 어렵게 한다.

자소 모델의 훈련과 인식은 '국민교육헌장'을 25 명이 필기한 글씨를 중심으로 실시하였다. 우선 자소 모델의 훈련은 글자별 필기에서 자소별로 분할하여 각 자소를 훈련 시켰다. 물론 여기서 자소란 연결 자소를 포함한 확장된 의미의 자소를 말한다. 파라미터 추정은 주어진 훈련 샘플에 대하여 반복적으로 파라미터를 재추정하는 Baum-Welch 알고리즘을 사용하였다. [8, 9] 이것은 훈련 샘플에 대해서 최적에 버금가는 maximum likelihood 파라미터를 추정하는 알고리즘이다. 훈련 샘플은 25 명중 20 명의 '국민교육헌장' 샘플에서 일부를 이용하였고, 나머지 희귀 자소와 연결 자소는 별도로 보충을 하였다.

(%)

	후보 1 개	후보 5 개
인식률	74.06	83.78

표-1. 인식률

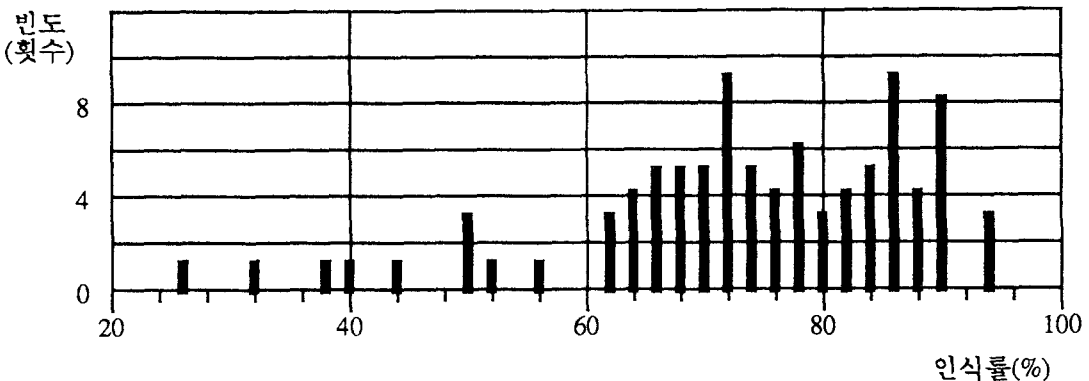


그림-4. 26명의 필기자 X 4화일에 대한 인식률 분포 - 후보 1개의 경우

인식 결과 로써 글자 후보의 갯수를 1개와 5개의 두가지로 하여 인식률을 구하였는데 표-1과 같다. 그리고 각 필기자가 쓴 문장을 네개의 화일로 분할하여 각 화일별로 인식률을 구하였는데, 화일 갯수로 본 인식률의 분포는 그림-4와 같다.

VI. 결론

본 논문에서 제시한 방법은 동적 프로그래밍 기법을 이용하여 자소의 경계를 통계적 근거에 의하여 판단한다. 자소별 인식은 본 한글 인식 네트워크 모델 봉넷(BongNet)에의하면 자소 경계 판단과 동일시할 수 있다. 자소간 경계는 자소 분석의 결과를 토대로 평가되며, 자소 인식은 가장 개연성 있는 자소 경계에서 완성되기 때문이다. 그러므로 글씨를 흘려 썼다 하더라도 자소 분리형 글씨 입력과 구별 없이 똑같은 과정으로 인식이 가능하다. 요컨대 여기서 제시한 방법은 필기의 흘림을 하나의 계산 구조로 간단히 해결할 수 있음을 보여 주었다.

참 고 문 헌

- [1] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, New York, N.Y., 1973
- [2] 이희동, "확장된 DP 매칭법에 의한 흘림체 한글 온라인 인식," *한국정보과학회논문지*, 1988
- [3] 성운재, "계층적 곡선표현 기법을 이용한 온라인 한글 필기 인식," *한국과학기술원 석사학위 논문*, 1991
- [4] 정봉 만, "온라인 한글 입력의 적응 학습과 인식에 관한 연구," *한국정보과학회논문지*, 제 16권 5호, pp. 487-498, 1991. 9.
- [5] 성태진, 방승양, "문자조합 규칙 학습에 의한 한글 온라인 필기인식기의 설계," *한국정보과학회 추계 학술발표 논문집*, pp. 223-226, 1991. 10.
- [6] 이일병, "동적 문자 인식 모형 사례 연구," *정보과학회지*, 제 9권 1호, pp. 64-78, 1991. 2.
- [7] J. K. Baker, "The DRAGON System - an Overview," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-23, no.1, pp. 24-29, Feb. 1975
- [8] L. E. Baum et al., "A Maximization Technique Occuring in the Statistical Analysis of Probabilistic Functions of Markov Chains," *Ann. Math. Statistic.*, 41., pp. 164-171, 1970
- [9] L. R. Rabiner, "Introduction to Hidden Markov Models," *IEEE ASSP Magazine*, vol. 3, no. 1, pp.4-16, Jan. 1986
- [10] A. Viterbi, "Error Bounds For Convolutional Codes And An Asymptotically Optimum Decoding Algorithm," *IEEE Trans. Info. Theory*, vol. 13, no. 2, pp. 260-269, Apr. 1967
- [11] R. E. Bellman, *Dynamic Programming*, Princeton University Press, Princeton, N.J., 1957
- [12] C. S. Meyers and L. R. Rabiner, "A level-building dynamic time warping algorithm for connected word recognition," *IEEE Trans. ASSP*, vol. ASSP-29, pp. 284-297, Apr. 1981