

필기 방향 변이를 수용하는 문자 인식 방법

이도곤^o, 김우생
광운대학교 컴퓨터공학부

Character Recognition Method Admitting a Sequence Variation of Handwritten Direction

Do Gon Lee^o, Woosaeng Kim
Dept. of Computer Engineering, Kwangwoon Univ.

요 약

사용자마다 여러 필기 방식이 존재하기 때문에 입력된 문자가 획 순서를 달리하여 필기했을 경우 오인식 발생확률이 많다고 볼 수 있다. 따라서 본 논문에서는 사용자의 서로 다른 필기 방향을 처리하는 인식 방법을 제안한다. 하나의 문자라도 필기 모양에 따라 해당 모델에서 그 문자가 발생할 확률 값이 다르지만 임계 확률 값 즉, 다양한 필기 모양에 상관없는 최소한의 발생 확률 값을 구할 수 있다. 따라서 시스템이 입력 문자를 인식할 때 어떤 모델에서의 발생 확률이 그 모델에서의 임계 확률 값보다 낮을 경우는 훈련과는 다른 필기체로 쓴 것이라고 가정할 수 있으며, 이러한 정보를 통해서 다른 필기 방향의 문자를 인식할 수가 있다.

1. 서론

전자 펜은 휴대성과 편의성의 장점을 지녀 PDA, palm PC, pocket PC, 스마트폰, Tablet PC 등의 많은 응용에 사용될 수 있기에 높은 인식률의 온라인 문자 인식 방법이 더욱 필요할 것으로 보인다[1,2]. 온라인 문자 인식 시스템은 전처리를 한 후에 정합을 하게 되는데, 이러한 정합 연구들에 쓰인 방법들 중에서 은닉 마르코프 모델 (HMM)은 음성 인식 분야뿐 아니라 문자 인식에서도 성공적으로 적용되고 있다 [3,4,5].

사람은 필체에 따라 같은 문자라도 필기 모양이 다를 수 있으며, 이러한 모양의 변이는 HMM 모델 생성 시 많은 학습데이터를 통해 흡수가 가능하다. 그러나 문자를 구성하는 획들의 필기 방향이 다른

경우에는 이러한 변이들을 하나의 모델로 흡수하기가 힘들어진다. 이러한 변이 흡수를 위해서 각 문자에 대해 서로 다른 획의 필기방향에 대응하는 여러 개의 HMM 모델들을 생성하는 방법이 있으나[6], 이것은 생성된 모델 수에 비례하여 인식시간이 길어지고 또한 모델 수가 많아지므로 오인식이 늘어날 가능성이 그만큼 많아진다.

본 논문에서는 각 문자에 대해서 하나의 HMM 모델만을 생성하면서도 서로 다른 필기 방향을 처리하는 방법을 제안한다. 하나의 문자라도 필기 모양에 따라 해당 모델에서 그 문자가 발생할 확률 값이 다르지만 임계 확률 값 즉, 다양한 필기 모양에 상관없는 최소한의 발생 확률 값을 구할 수 있다. 따라서 입력 문자를 인식할 때 어떤 모델에서의 발생 확률이 그 모델에서의 임계 확률 값보다 낮을 경우,

그 문자는 필기 방향이 다른 어떤 문자라고 가정할 수가 있으며, 이러한 정보를 통해서 다른 필기 방향의 문자를 인식할 수가 있다. 본 논문에서는 영어의 대문자만을 대상으로 제안하는 방법의 가능성을 시도해 본다.

2. 은닉 마르코프 모델(HMM)

은닉 마르코프 모델(HMM)은 유한개의 상태와 각 상태 사이를 방향성 있게 연결하는 전이의 집합으로 구성된 네트워크로 정의된다. 상태간 전이 확률은 마르코프 프로세스를 따르며 출력 확률은 시간에 독립적으로 각 상태에 지정되어 있다. 일반적으로 HMM은 $\lambda = (A, B, \pi)$ 의 기호로 표현되며 A는 상태전이 (state transition) 확률, B는 관측심볼 (observation symbol)의 출력확률, π 는 초기 상태전이 (initial state transition) 확률을 나타낸다. 한 문자에 대해 하나의 모델이 구성되고 서로 다른 n개의 문자에 대해 대응하는 각각의 모델이 $i = 1, 2, \dots, n$ 으로 표기될 때, 관측열 $O = O_1, O_2, \dots, O_T$ 에 대응하는 미지의 문자는 식(1)을 만족하는 λ_j 모델에 대응하는 문자로 인식된다.

$$\lambda_j = \max_{i=1..n} P(O | \lambda_i) \text{ ----- (1)}$$

3. 필기 방향 변이를 반영하는 문자 인식

영어 대문자의 경우 최대 4개 획의 조합으로 하나의 문자가 구성된다. 예를 들어, 문자 'C'는 사용자에게 관계없이 1 획의 조합으로 만들어진다. 그러나 많은 문자들은 사용자에게 따라 두, 세 가지 획의 조합으로 만들어질 수 있다. 예를 들어, 문자 'D'의 경우는 'l'와 'J'의 2 획의 조합으로 만들어진다. 영어 대문자 중에서 사용자에게 따라 필기 방향이 다를 수 있는 문자는 일반적으로 1획으로 만들어지는 'O'와 2획으로 만들어질 수 있는 'B', 'D', 'P', 'X'와 3획으로 만들어질 수 있는 'E', 'F' 정도이다. 예를 들어, 1획으로 만들어지는 문자 'O'의 경우는 사용자에게 따라 시계방향 또는 반시계 방향으로 쓸 수 있고, 2

획으로 만들어지는 문자 'D'의 경우는 사용자에게 따라 'l'획 다음에 'J'획으로 또는 'J'획 다음에 'l'획으로 쓸 수 있다. 본 논문에서는 사용자가 일반적으로 필기하는 방향으로 쓰여진 문자를 정필기 문자, 그렇지 않은 경우의 문자를 역필기 문자라고 정의한다. 예를 들어, 문자 'D'의 경우 'l'획 다음에 'J'획으로 쓰여진 것은 정필기 문자, 'J'획 다음에 'l'획으로 쓰여진 것은 역필기 문자이다.

본 논문에서는 역필기 문자가 특정 모델에서 인식이 될 때의 발생 확률 값이 그 모델의 임계 확률 값보다 작다는 것을 가정하고 있으며, 이것은 4장의 실험을 통해서 확인이 되었다. 따라서 시스템은 입력 문자가 정필기 문자인지 또는 역필기 문자인지를 알아내서, 역필기 문자인 경우는 정필기 문자로 바꾼 후에 인식을 시도하게 된다. 이때 역필기 문자를 정필기 문자로 바꾸는 방법은 내부적으로는 역필기 문자의 체인 코드의 순서를 바꿔주는 방법을 사용한다.

그런데 문자가 1획이나 2획으로 구성된 경우는 역필기 문자를 정필기 문자로 바꾸는 방법이 1가지 밖에 존재하지 않으나, 3획이나 4획으로 구성된 문자는 최대 4!-1 경우의 방법들이 존재할 수 있다. 따라서 본 논문에서는 사용자에게 따라 필기 방향이 다를 수 있는 'B', 'D', 'E', 'F', 'O', 'P', 'X' 들 중 최대 2획으로 구성된 'B', 'D', 'O', 'P', 'X' 들만을 대상으로 한다.

이러한 가정들을 바탕으로 본 논문에서 제안하는 알고리즘은 다음과 같다. 먼저 시스템은 입력 문자에 대한 인식을 시도할 때 3획 이상으로 구성된 문자는 일반적인 방법으로 인식을 한다. 즉, 모든 모델 중 발생확률이 가장 높은 모델에 해당하는 문자로 인식을 한다. 반면, 2획 이내로 구성된 문자의 경우는 발생 확률이 가장 높은 모델 M1의 발생 확률 값 P1을 구한다. 만약 P1이 M1의 임계 확률 값 E1보다 같거나 클 경우 인식은 성공한다. 그러나 만약 P1이 E1보다 작을 경우 시스템은 입력 문자가 역필기 문자라고 판단을 하고, 입력 문자를 정필기 문자로 바꾼 후에 다시 인식을 시도하여 특정 모델 M2의 발생 확률 값 P2를 구한다. 만약 P2가 M2의 임계 확률 값 E2보다 같거나 클 경우 인식은 성공

한다. 그러나 P2가 E2 보다 작을 경우는 P1과 P2중 확률 값이 큰 모델의 문자로 인식을 한다. 2획 이내로 구성된 문자의 경우 제안하는 알고리즘의 수도코드는 다음과 같다.

```

if 모델 M1의 발생확률값 P1 ≥ M1의 임계확률값 E1 then 인식
else 역필기 문자를 정필기 문자로 바꿈

  if 모델 M2의 발생확률값 P2 ≥ M2의 임계확률값 E2 then 인식
  else
    if P1 ≥ P2 then M1에 해당하는 문자로 인식
    else M2에 해당하는 문자로 인식
    
```

그림 1. 필기방향 변이 수용 알고리즘의 수도코드

4. HMM 구조 및 훈련

초기 본 논문에서 사용된 HMM 모델의 구조는 left-to-right 모델이며, 상태 전이의 수는 자기 전이와 다음 상태까지만 전이 할 수 있는 구조를 갖는다. 각 문자에 대응하는 HMM의 상태 수는 10개로 하였고, 각 상태에서의 각 심볼의 발생 가능성은 같게 하였으며 초기 전이 확률 분포는 균일 확률 분포로 하였다. 각 문자는 인터페이스의 박스 당 한 문자씩 필기 하게 하여 외부 분할 방식으로 분리되게 하였으며, Tablet으로부터 입력되는 데이터는 먼저 거친 점 제거와 평활화 등의 전처리 과정을 거친 후 코드화 하기 위해 일정 시간 간격으로 샘플링 된 일련의 점들을 8 방향 체인코드로 만들어 사용하였다[7]. 본 논문에서는 훈련을 위해 Baum-Welch 알고리즘을 사용하였다[3].

5. 실험 및 결과 분석

제안한 방식들의 성능을 분석하기 위해 TMC300XCi Tablet PC와 WACOM Tablet을 사용하였으며, C#과 Microsoft Tablet PC Platform SDK API로 구현하였다[11]. 각 문자는 A~Z의 26개 문자를 바탕으로 훈련과 인식을 시켰다. 본 논문에서 문자 훈련은 5명의 데이터를 사용하였고, 각 사람은 각 문자를 3

번 필기하도록 해, 각 모델은 평균적으로 15개의 해당 문자를 가지고 훈련되었다. 임계 확률 값을 구하기 위하여 각 문자에 대해서 새로운 10개의 데이터를 사용하였고, 실제의 인식 실험에서는 훈련에 참여하지 않은 5명의 데이터를 사용하였다.

5.1 임계확률 값의 설정

각 문자의 임계 확률 값을 구하기 위해 가급적 모양 변이가 큰 데이터들을 사용하였고, 올바르게 인식되면서 발생확률이 가장 낮은 값을 임계 확률 값으로 했다. 예를 들어, 표 1은 A 문자의 10개의 서로 다른 모양의 데이터들을 가지고 인식을 시도하였을 경우의 발생 확률을 보여준다.

문자	인식	발생확률	문자	인식	발생확률
A	A	0.000886	A	A	0.000033
A	A	0.000311	A	B	0.000006
A	A	0.000078	A	A	0.000012
A	A	0.000050	A	B	0.000004
A	A	0.000049	A	A	0.000010

표 1. 문자 'A'의 임계확률 값 추출과정

표 1에 나온 것처럼, 문자 A의 경우 10개의 입력 문자 중 제대로 인식이 되는 것은 8개의 경우이고, 그 중 확률 값이 가장 작은 0.00001이 임계 확률 값이 된다. 이러한 방법을 통하여 A부터 Z까지의 각 문자의 임계확률 값은 표 2와 같이 구할 수 있다.

문자	임계값	문자	임계값	문자	임계값
A	0.000010	J	0.000001	S	0.000006
B	0.000012	K	0.000001	T	0.067666
C	0.000709	L	0.000286	U	0.078252
D	0.000713	M	0.000843	V	0.000219
E	0.000680	N	0.000301	W	0.000001
F	0.000863	O	0.000129	X	0.001411
G	0.000223	P	0.000018	Y	0.009232
H	0.000139	Q	0.000004	Z	0.009417
I	0.000029	R	0.000011	-	-

표 2. 실험을 통해 추출된 문자 'A~Z'의 임계확률 값

표 3은 문자 'B', 'D', 'O', 'P', 'X'를 역필기로 인식을 시도하였을 경우의 특정 모델에서의 발생 확률과 임계 확률 값을 보여준다. 표 3에서와 같이 역필기 문자의 경우 특정 모델에서의 발생 확률 값이 임계 확률 값보다 적음을 알 수 있다.

역필기문자	모델	발생확률값	임계확률값
B	P	0.00000082	0.000018
D	P	0.00000470	0.000018
O	B	0.00000013	0.000012
P	D	0.00000002	0.000713
X	K	0.00000004	0.000001

표 3. 특정 모델에서의 발생확률값과 임계확률값 비교

5.2 제안한 방법의 성능비교

사용자는 A~Z까지의 문자를 필기하며 그 중 'B', 'D', 'O', 'P', 'X' 문자들에 대해서는 역필기를 허용 함으로써 역필기의 포함 비율에 따른 문자 인식률을 조사하였다. 본 논문에서 제안한 필기방향 변이를 수용한 방법의 경우 <그림 2>와 같이 역필기의 비율이 커질수록 제안하는 방법의 인식률이 기존 방법보다 높아지는 것을 알 수 있다.

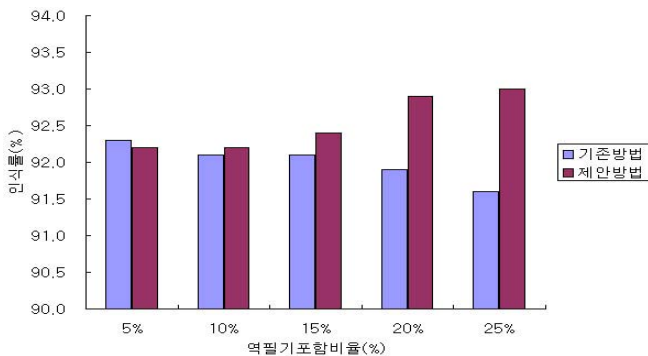


그림 2. 제안방법과 기존방법의 문자 인식률 비교

6. 결론

본 논문에서는 각 문자에 대해서 하나의 HMM 모델만을 생성하면서도 서로 다른 필기 방향을 처리하는 방법을 제안하였다. 시스템은 HMM 모델의 임

계값을 통해 문자 인식 시 정필기 문자인지 역필기 문자인지를 판별하여, 역필기 문자는 다시 정필기 문자로 바꿔 인식하는 방법을 사용하였다. 향후 과제로는 영어 대문자 외의 다른 문자와 3획 이상의 문자에 대해서도 필기 방향 변이를 수용하는 방법을 조사해 볼려고 한다.

[참고문헌]

- [1] 이성환, “온라인 필기 인식 기술의 현황”, 한국정보과학회지, 제 9권 제 1호, 1991년 2월, pp.54-63
- [2] Alexander Wolfe, “Putting Pen to Screen On Tablet PC”, IEEE Spectrum, October 2002
- [3] 이성환, “문자인식 - 이론과 실제, 흥릉과학출판사”, 1993
- [4] L. R. Labiner, “A Tutorial on HMM and Selected Applications in Speech Recognition,” Proc. IEEE, 1989
- [5] Bong-Kee Sin and Jin H. Kim, “Ligature Modeling for Online Cursive Script Recognition,” IEEE Trans. On Pattern Analysis and Machine Intelligence, Vol. 19, No. 6, June 1997, pp.623-633
- [6] 이재준, “온라인 필기 인식을 위한 데이터 기반의 HMM 구조 설계”, 한국 과학 기술원, 2000
- [7] 이수봉, 김우생 “문자 획의 특성을 이용한 전처리 기법”, 한국멀티미디어학회, 제 7권 제 1호, 2004년 5월, pp.758-761
- [8] R. Shinghal and G. T. Toussaint, “Experiments in Text Recognition with the Modified Viterbi Algorithm,” IEEE Trans. on Pattern Analysis and Machine Intelligence, 1979.
- [9] 김상운, “패턴인식 입문, 흥릉과학출판사”, 1997
- [10] 구분석, 전변환, 박명수, 김성훈, 안진모, 김재희, “On-line 문자 인식을 위한 전처리 기법”, 한국 정보과학회 인공지능 연구회 추계 학술발표 논문집, 서울, 1991년 11월, pp. 78-82
- [11] Rob Jarrett, Philip Su, “Building Tablet PC Applications”, Microsoft Press, 2003