

HMM을 이용한 수기숫자 인식에 관한 연구

조민환*

A Study on the Hand-written Number Recognition by HMM(Hidden Markov Model)

Cho Meen Hwan *

요 약

대부분의 수기 숫자 인식 시스템에서는 자모 형태를 이용한 특징 점 추출과, 형태소 적 분석기법을 많이 사용하였다. 본 연구에서는 체인코드를 사용하고, 생성된 체인코드를 미분하여 최소 값이 되는 미분코드를 만들었다. 이 미분코드는 대부분의 수기 숫자에 적용해 본 결과 숫자 변별력이 매우 뛰어난 것을 알 수 있었다. 처리 순서는 몇 개의 수기숫자를 전 처리하고, 체인코드와 미분코드를 생성 한 후, HMM 인식 네트워크를 사용하여 숫자 인식이 되었다. 처리 결과 96.1%의 수기숫자를 인식하였으며, 매우 심하게 왜곡된 숫자는 인식하지 못하였다.

Abstract

In the most of recognizing systems of hand-written numbers, extraction of feature shape by using character elements shapes and a method of morphological analysis by using then extraction of feature shapes were usually used. In this paper, however, peculiar chain-code is used, and differential code which gets minimal value by differentiating the chain-code which is generated by the peculiar chain-code is made. We found this differential code is very successful in discriminating hand-written numbers according to the result of applying to most of the hand-written numbers. Testing recognition of hand-written numbers by HMM network. From the results, we can recognize of 96.1 percentage hand-written numbers but can not recognize extremely distorted hand-written numbers.

▶ Keyword : 인식(Recognition), 방향성코드(Chain Code), HMM(Hidden Markov Model)

• 제1저자 : 조민환
• 접수일 : 2004.07.02, 심사완료일 : 2004.08.21
* 창원전문대학 컴퓨터정보처리과 조교수

I. 서론

물체를 인식하는 방법으로 물체의 형태에서 특징벡터를 추출하고, 이를 기반으로 인식알고리즘을 구현하는 방법과, 윤곽선 외에 물체 표면의 결(texture) 분석을 이용하는 방법 등이 있다. 수기 문자 인식에 대한 연구는 많은 사람들의 연구 대상이며 아직까지 연구되고 있다. 특히 수기 숫자 인식은 은행, 우체국, 라벨 인식 등 많은 영역에서 이용 가능하여 활발한 연구가 진행되고 있다. 특히 우편번호 인식은 주어진 시간 내에 많은 량의 분류작업이 요구된다. 우편번호 분류 작업은 인건비가 비싸고, 실시간으로 처리되어야 하며, 작업량의 변동률이 매우 크므로 실시간 처리에 많은 어려움이 있어 자동 분류시스템 개발이 시급히 요구되고 있다[1][2][3][8].

본 시스템은 수기된 우편 번호를 디지털 카메라로 읽은 다음 전 처리하여 잡음과 형태소 적 연결, 접합 등을 정규화 한다. 이들 영상을 세선화하고, 8방향 체인코드에서 우측과 아래 방향만 사용하여 체인코드를 생성시킨다. 사람에 따라 수기 필체는 여러 각도로 기울어지므로 기울어진 숫자의 고유 코드를 만들기 위해 미분 코드를 생성시킨다.

미분코드는 순환코드로 생각하여 최소 값이 되는 코드로 다시 변환하여 각 숫자에 대응되는 HMM(Hidden Markov Model)에 적용시켜 3자리 혹은 그 이상의 숫자를 인식하게 하였다. 다양한 모양의 18개의 6자리 숫자들을 대상으로 실험하여 결과를 분석하였다.

II. 수기숫자 구조와 전 처리

2.1 전 처리

문자인식 시스템에서 문자 데이터의 입력과 출력사이의 전 처리 과정을 블록으로 표시하면 (그림 1)과 같다.

첫째, Phase①은 인쇄된 서류를 스캐너(scanner)나 혹

은 디지털 카메라(digital camera)를 사용하여 디지털 원시 데이터를 만든다. 둘째, Phase②에서 원시 데이터를 필터링(filtering) 한 후 2진 데이터로 변환한다. 셋째, Phase③에서는 2진 데이터에서 히스토그램(histogram)을 사용하여 문자열과 문자열 사이의 공간을 찾아내어 한글의 방향을 찾아내고 동일한 히스토그램을 사용하여 각 줄에 위치한 각 글자를 발췌한다. 넷째, Phase④에서는 발췌된 글자를 구성 자소로 분리하고 분리된 각 자소의 특징을 추출한다. 다섯째, Phase⑤에서는 Phase④에서 추출된 특징을 기초로 문자인식 프로그램을 처리하여 문자 코드로 출력하게 된다.

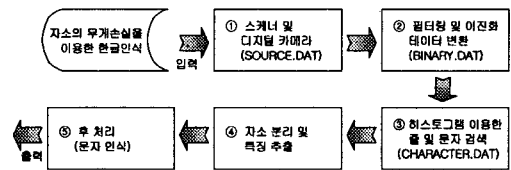


그림 1. 문자인식을 위한 전처리 과정
Fig. 1 Preprocessing phase for character recognition

2.2 문자인식을 위한 전 처리 과정

(그림 2)는 한글 문서를 디지털 카메라로 촬영하여 만든 디지털 영상 데이터를 보여준다. 이 때 각 화소의 명암도는 256이다. (그림 3)은 (그림 2)을 메디안(median)필터에 의해 잡음을 제거 후 문턱치 값을 200으로 설정하여 변환된 2진화 영상을 보여준다. (그림 2)의 경우 주변 광이 혼합되어 배경(여백)의 밝기가 매우 어렵게 나타났으나, (그림 3)에서 문턱치 값을 적절하게 선택하여 배경과 글자와의 비를 선명하게 할 수 있음을 보여준다.



그림 2. 수기 숫자 모형
Fig. 2 Handwritten number recognition model

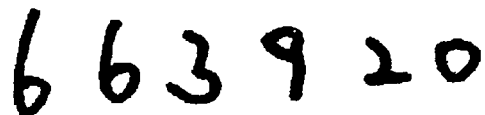


그림 3. 2진화 처리된 영상
Fig. 3 Digitized image

(그림 3)에서 문자 가장자리에는 많은 잡음이 있으므로 이 상태에서 세선화는 결과는

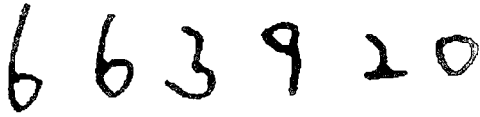


그림 4. 세선화를 위한 침식 처리된 영상
Fig. 4 Eroded image for thinning

체인 코드를 생성하기에 부적합하므로 (그림 4)와 같이 열림 처리를 한 후 세선화를 생성하였다. 그 결과 (그림 5)와 같은 결과를 얻을 수 있다.



그림 5. 세선화 처리된 영상
Fig. 5 Thinned image

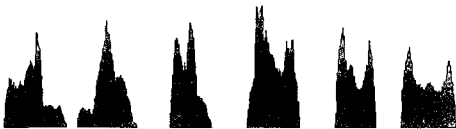


그림 6. X평면에 투사된 히스토그램
Fig. 6 Histogram for a subject projected on X plane

(그림 6)은 (그림 2)를 X 평면에 투사하여 작성된 히스토그램을 보여준다. (그림 6)에서 각 자소를 분리할 수 있는 문턱치 값을 얻을 수 있다.

III. HMM 네트워크

3.1 방향코드와 숫자모델

(그림 7)은 체인 코드 및 숫자 6에 대한 적용 경로를 보여준다. 기존 8방향코드는 진행방향 및 교점에서의 방향 설

정이 복잡하므로 상호 180도인 방향코드는 하나로 통합하여 방향설정의 혼란을 제거하였다. 숫자의 상단 끝점과 교점을 출발점으로 하여 (그림 7) 코드를 생성하며 교점을 만나게 되면 진행방향과 가장 유사한 방향을 설정하여 코드를 생성한다(9).

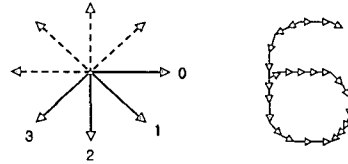


그림 7. 4 방향코드 6에 대한 체인코드 생성 경로
Fig. 7 Path of chain code generation for 4-direction code 6

(그림 7)에서 6에 대한 체인 코드는 1033222221100332211003 이 되며, HMM에 적용하기 위한 상태는 10321032103 이 된다. 이 값을 순환코드라 생각하고 최소 값을 계산하면 03103210321이 된다. 순환 체인코드의 미분 값은 3-2-13-1-1-13-1-1-1이 되며 문자 스트링으로 표현하기 위해 최저 값을 0으로 하는 초과 코드는 61262226222가 된다. 이 경우 형태 코드를 생성하면 최종 미분 코드는 12622262226이 된다(2)(4).

〈표 1〉은 0에서 9까지 수기 숫자에 대한 미분 코드를 보여준다.

표 1. 수기숫자에 대한 체인코드와 미분코드
Table 1. Chain and differential code for handwritten numbers

숫자	Chain Code	Differential Code
0	01230123	04440444
1	2	3
2	012303	044406
3	012312303	044414406
4	032	162
5	0020123	0351444
6	03103210321	12622262226
7	022	153
8	0123012303210321	044406222622444
9	01210321	22244226

3.2 HMM과 숫자 모델

은닉 마르코프 모델(HMM)은 60년대 말부터 Baum등에 의하여 연구되고 효율적인 알고리즘이 개발되면서 알려지기 시작하였으며, Baker가 처음으로 연속 음성 인식 문

제에 적용할 수 있음을 보였다. HMM은 마르코프 체인과 각 상태에서의 관찰부호 확률분포로 정의할 수 있는 확률적 마르코프 체인, 또는 마르코프 체인의 확률 함수라고 정의할 수 있다.

마르코프 모델은 유한개의 상태(state, 또는 노드)와, 각 상태 사이를 방향성 있게 연결하는 전이(transition)의 집합으로 정의한다. 각 상태에 하나의 관찰 부호(observation symbol)를 할당하고, 전이는 확률적으로 정의하며, 전이로 연결한 노드는 하나의 네트워크를 구성한다. 각 시간 $t=0,1,\dots$ 에서의 상태 q_t 에 대한 확률 변수는 현재의 상태와 이전의 상태열의 조건부 확률로 정의할 수 있으며, 일차 마르코프 성질(또는 first-order Markov assumption)을 이용하여 아래와 같은 식으로 재 정의할 수 있다.

$$\begin{aligned}
 P(q_{t+1} = S_j | q_0 = S_a, q_1 = S_b, \dots, q_t = S_i) \\
 &= P(q_{t+1} = S_j | q_t = S_i) \\
 &= S_{ij}
 \end{aligned}$$

즉, 상태 전이는 이전의 상태와 무관하고, 전이 직전의 상태만의 함수임을 나타낸다. 이 값을 전이 확률(transition probability)이라 하여, 시간에 독립적이라 가정하여 a_{ij} 로 나타낸다. 전체 네트워크에서의 전이 확률 분포는 행렬 $A = \{a_{ij}\}$ 로 표현한다.

임의의 시각 t 에서 마르코프 모델이 상태 S_j 의 확률 값은

$$\begin{aligned}
 P_t(j) &= \sum_i P_{t-1}(i) a_{ij} \\
 1 &\leq j \leq N
 \end{aligned}$$

와 같이 순환적으로 표현할 수 있다. 이를 확장하여, 각 상태에서의 관찰 부호를 확률적으로 나타냄으로써 모델링 능력을 늘릴 수 있다. 이 확률 변수를 그 상태에서의 확률 함수 $b_j(k)$ 로 나타내면 상태 S_j 의 확률 $P_t(j)$ 의 값은

$$\begin{aligned}
 P_t(j) &= \sum_i P_{t-1}(i) a_{ij} b_k(k) \\
 1 &\leq j \leq N \\
 1 &\leq k \leq M
 \end{aligned}$$

로 재정의 된다. 이러한 모델은 두 가지의 확률 변수에 따라 결정되는 복잡한 구조를 갖는 마르코프 모델이 되는데 이를 은닉 마르코프 모델 (Hidden Markov Model, HMM)

이라고 한다[5][6][7].

3.3 숫자 연결 모델

수기숫자는 숫자 3자리씩 2회 표기하여 모두 6자리로 나타낸다. 이 경우 HMM 네트워크로 표시하면 (그림 8)과 같이 나타낼 수 있다.

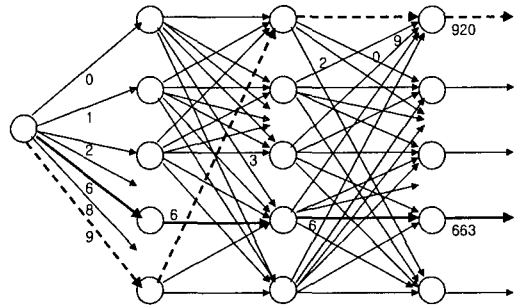


그림 8. 3자리 숫자 HMM 네트워크
Fig. 8 3-digit HMM network

(그림 8)에서 처음 3자리 숫자가 (그림 5) 세션화 된 수기 숫자의 663인 경우 굵은 실선에 표시된 경로를 취하게 되어 최종 미분 체인 코드는 12622262226035144412622262226이 됨을 알 수 있으며, 뒤 3자리 숫자의 HMM 네트워크 경로는 (그림 8)의 점선으로 표시되고, 미분 체인 코드는 01230123032103210444044404440444가 된다.

이러한 미분코드를 탐색하여 수기된 숫자를 인식할 수 있다.

IV. 숫자인식 실험 및 결과

4.1 실험 및 결과

본 연구에서 HMM 인식 네트워크를 사용하여 수기숫자를 인식실험결과 <표 2>와 같이 평균 96.1%의 인식 율로 참고문헌 [1] [5] [7] 평균 인식 율 보다 가장 우수한 결과를 나타냈다.

표 2. 수기숫자 인식의 실험결과
Table 2. The results of handwritten numbers recognition rate

부류	0	1	2	3	4	5	6	7	8	9	인식률
0	639					1	2		2		96.5%
1		660	2			1		2			97.5%
2			390	3				1	2		95.5%
3			3	920					3	2	95.1%
4		2	2		661		1			1	95.8%
5	1	2		3		625					96.9%
6	2	1				1	663				97.5%
7			1	1	3		1	672	1	1	96.1%
8	1	2			1				934	1	95.4%
9	1		1	1			1	3	2	991	95.2%
평균											96.1%

V. 결론

본 논문에서는 0에서 9의 숫자로 조합되는 수기숫자를 대상으로 한 인식시스템을 제안하였다. 기존 수기숫자의 인식 방법으로는 8방향 체인코드와 은닉 마르코프 방법을 이용한 통계처리를 많이 사용하였으나, 본 논문에서는 유사 4 방향 코드를 사용하여 8 방향에서 발생하는 처리 방향에 다른 코드 선점의 차이점을 제거하였다. 또한 각 자모의 체인코드를 미분코드로 변환하여 자소의 기울기에 관계없이 동일한 코드가 생성되게 하였다. 이러한 미분코드는 은닉 마르코프 네트워크를 이용한 경로탐색 방법으로 연속된 숫자를 인식하였다. 결과 25개의 수기숫자 중심하게 왜곡된 1개를 제외한 24개의 수기숫자를 인식하여 매우 높은 인식률을 나타내었다.

참고문헌

[1] Jay June Lee, "Data-driven Design of HMM Topology for On-line Handwriting Recognition", Ph.D. Dissertation for Department of Electrical Engineering and Computer Science / Division of Computer Science, KAIST, 2000 (in Korea).
[2] N. Merhav and Y. Ephraim, Maximun likelihood hidden Markov modeling using a dominant sequence of states. IEEE Trans. on Signal Processing, 39(9):2111-2115. Sep. 1991.

[3] J.-Y. Ha, S.-C. Oh, and Jin H. Kim. Unconstrained handwritten word recognition with interconnected hidden Markov models. In Third Int. Workshop on Frontiers in Handwriting Recognition, pp.455-460, Buffalo, New York, USA, May 1993.
[4] Michel Gilloux, Manuel Leroux, and Jean-Michel Bertille, "Strategies for Handwritten Words Recognition Using Hidden Markov Models", Proc. of 2nd Int. Cong. on Document Analysis and Recognition, Tsukuba Science City, Japan, 299-304, 1993.
[5] H. Bunke, M. Roth, and E. G. Schukat-Talamazzini, "Off-Line Cursive Handwriting Recognition Using Hidden Markov Model", Pattern Recognition, Vol.28, No.9, pp.1399-1413, 1995.
[6] Yannick Tremblay and Richard W. Prager, "Unconstrained Handwriting Recognition Using Linked Character Models", Proc. of 4th Int. Workshop on Frontiers in Handwriting Recognition, Taipei, Taiwan, pp. 449-455, 1994.
[7] Woo Sung Kim and Rae-Hong Park, "Off-line Recognition of Handwritten Korean and Alphanumeric Characters Using Hidden Markov Models", Pattern Recognition, Vol.29, No.5, pp.845-858, 1996.
[8] 조성목, "이미지 특징 추출연산자 DBAH/DBAG 와 하드웨어 실현", 한국OA학회 논문지, 제6권 제1호, 2001.
[9] 조민환, "체인코드를 이용한 훼손된 지문의 인식에 관한 연구", 한국컴퓨터정보학회 논문지, 제8권 제4호, pp63-68, 2003.

저자소개



조민환

1989년 3월 숭실대학교 산업대학원 전자계산학과(공학석사)
2001년 2월 경남대학교 대학원 컴퓨터공학과(공학박사)
1994년 3월 ~ 현재 창원전문대학 컴퓨터정보처리과 교수