

모듈러 설계 및 파이프라인 연결에 기반한 무제약 필기 숫자의 인식

Recognition of Unconstrained Handwritten Numerals
Based on Modular Design and Pipeline Connection

오일석[†], 최순만^{††}, 홍기천^{†††}, 이진선^{††††}

Il Seok Oh, Soon Man Choi, Ki Cheon Hong, Jin Seon Lee

요 약

본 논문에서는 필기체 숫자 인식 프로그램을 설계하는데 있어서의 구조적인 면의 중요성을 강조하고 두 가지 구조적 설계에 대해서 기술한다. 첫째로는 숫자 인식 프로그램에 대한 모듈러 설계를 기술하고 그에 대한 이점들을 기술한다. 이러한 구조에서 인식기는 10개의 이진 부인식기로 구성되어있으며, 각각의 부인식기는 단지 하나의 부류에 대해서만 책임을 진다. 규칙 기반 훈련과 신경망 기반 훈련을 기술한다. 둘째로는 두개 혹은 그 이상의 인식기를 파이프라인으로 연결하였다. 파이프라인에서 두 번째 인식기는 첫 번째 인식기에서 인식된 패턴을 검증하는 역할을 담당하거나, 첫 번째 인식기에서 거부된 패턴을 재인식하는 역할을 담당한다. 이제까지 얻어진 실험 결과는 제안된 구

† 전북대학교 컴퓨터과학과 조교수
Dept. of Computer Science,
Chonbuk National University
e-mail: isoh@moak.chonbuk.ac.kr

†† 전북대학교 컴퓨터과학과 박사과정
Dept. of Computer Science,
Chonbuk National University
e-mail: cilab@moak.chonbuk.ac.kr

††† 전북대학교 컴퓨터과학과 석사과정
Dept. of Computer Science,
Chonbuk National University
e-mail: cilab@moak.chonbuk.ac.kr

†††† 우석대학교 정보통신 및 컴퓨터공학부
Division of Information Communication
and Computer Engineering, WooSuk Univ.
e-mail: jslee@core.woosuk.ac.kr

조 설계의 장점을 보여주고 있다.

주제어 숫자인식, 파이프라인, 모듈러 훈련, 신경망

ABSTRACT

In this paper we emphasize the importance of architectural aspects of designing a handwritten numeral recognition program, and describe two architectural design. First, we describe the modular design of a numeral recognition program, and mention its advantages. In this design, a recognizer is composed of 10 binary subrecognizers each of which is responsible for only one class. Rule-based training and neural-based training are presented. Second, we connect two (or more) recognizers serially which we call pipelining connection. The second recognizer may act as verifier for the patterns recognized by the first recognizer, or as second chance recognizer for the patterns rejected by the first recognizer. Our experimental results obtained till now show the merits of the proposed architectural designs.

Keywords Digit recognition, Pipeline, Modular training, Neural network

1. 서론

무제약 필기체 숫자인식은 우편번호 인식 [3,4], 전표 분석 [5], 학술 논문의 표지 페이지 분석 [6], 지도 분석 [7] 등 문서 인식 작업의 자동화에 있어서 매우 중요하다 [1,2]. 숫자 인식에 관련된 연구 논문으로 구조적인 방법 [8,9,10,11,12,13,14,15], 신경망 방법 [16,17,18], 그리고 복수의 알고리즘을 이용한 다중 결합 방법 [19,20,21,22] 등이 발표되었다. 숫자 인식 문제의 가장 두드러진 특징 중의 하나는 10개의 부류만으로 구성되어 있다는

점이다.

본 논문에서는 이러한 특징을 기반으로 하여 필기체 숫자인식 프로그램을 설계하는데 있어서 구조적인 면의 중요성을 강조하고 두개의 구조적 설계를 제안한다.

첫째로는 숫자인식 프로그램을 설계하기 위해 모듈러 접근 방법을 사용하였다. 즉, 하나의 분류 문제를 10개의 이진 부문제로 나누고, 각각의 부문제를 해결하는 방법을 사용하였다. 또한 이러한 접근방법의 몇 가지 이점을 기술한다. 이러한 설계 방법에 있어서 한 개의 인식기는 10개의 이진 부인식기로 나뉘어져

있으며, 이들 각각은 한 개의 부류에만 책임을 진다. 이들 이전 부인식기는 입력 패턴이 i -부류에 속하는지 아닌 지만을 판단하여 '예' 또는 '아니오'로 답변을 한다. 규칙 기반 훈련과 신경망 기반 훈련이 소개되어지고 아울러 기존의 구조에 기반한 인식기와 모듈러 구조에 기반한 인식기 사이의 성능 비교도 기술한다. 모듈러 접근 방법과 유사한 방법을 사용한 Nadal 과 Mui의 논문을 참고하기 바란다 [21,23].

둘째로는 더 나은 성능을 얻기 위해 두개 혹은 그 이상의 인식기를 사용한다. 기존의 조합 방법은 복수의 인식기로부터 나온 인식 결과를 모은 다음 최종 인식 결과를 결정하기 위해 그 결과를 결합하는 형태였다 [19,20,21,22]. 그러나 본 논문에서는 파이프라인이라고 하는 연결방법으로 두개 혹은 그 이상의 인식기를 연속적으로 연결하였다. 파이프라인 접근 방법에서는 첫 번째 인식기의 결과를 두 번째 인식기에 넘겨준다. 이때 두 번째 인식기는 첫 번째 인식기에 의해서 인식된 패턴에 대하여 다시 검증하는 역할을 하게된다. 이러한 검증 단계는 신뢰도를 상당히 올릴 수 있다. 또다른 연결방법으로는 두 번째 인식기가 첫 번째 인식기에 의해 거부된 패턴들에 대해 이차적 기회로서의 역할을 하게 하는 방법이다. 이러한 연결방법은 인식률을 향상시킨다. 이 실험에서 첫 번째 인식기로서는 규칙 기반 인식기를

사용하였으며, 두 번째 인식기로서는 신경망 기반 인식기를 사용하였다.

본 논문의 구성은 다음과 같다. 2장에서는 모듈러 접근방법이 상세히 기술된다. 이어 3장에서는 파이프라인 연결을 기술하고, 마지막 결론을 맺는다.

2. 모듈러 구조

모듈러 접근방법의 기본적인 이점은 각각의 부류에 대해 다른 특징 공간과 다른 인식 방법을 사용할 수 있다는 것이다. 예를 들면, 7과 9 부류에 대해서는 구조적 방법을 쓰고 다른 부류에 대해서는 신경망 방법을 사용할 수 있다. 7과 9와 같이 형태상으로 작은 차이가 있는 두개의 부류를 구별하는데 가장 적합한 특징을 사용함으로써 두 부류간의 고성능 분류를 이룰수 있다. 각각의 부류에 대해 가장 적당한 특징과 인식방법을 택함으로써 가장 최고의 인식 성능을 얻을 수 있다.

또 하나의 이점은 소프트웨어 공학과 밀접한 관련이 있다는 것이다. 즉, 각각의 모듈별로 관리를 할 수 있다는 것이다. 개발 기간 동안 몇 개의 팀이 서로 독립적으로 각각의 모듈을 구현할 수 있으며, 운영 기간 동안에는 오류가 많은 모듈만 더 좋은 모듈로 대체할 수 있다.

다음은 두 개의 훈련의 예를 보여준다.

2.1 훈련 예 1

부류 i 를 위한 이진 분류기 $R_i(i=0, 1, \dots, 9)$ 는 입력 패턴이 부류 i 에 속하는지를 판별하기 위해 거부 조건들의 집합으로 다음과 같이 코딩 되어있다.

```

Ri(P) /* P : input pattern */
{
  split P into subpatterns P1', P2',
  ..., Pm', if needed;
  /* normally m=2, 3 */

  extract feature F1 from P (or P1'
  , P2', ..., Pm');
  if(veto_rule (F1) is satisfied),
    return(NO);
  /* P not belonging to class i */

  extract feature F2 from P (or
  P1', P2', ..., Pm');
  if(veto_rule(F2) is satisfied),
    return(NO);
  /* P not belonging to class i */

  :
  :

  return(YES);
  /* P belongs to the class i */
}
    
```

위의 코드에서 veto_rule(F_i)는 “F_i 의 값 $\geq THRES_i$ 와 같은 Boolean 형태를 가진다.

$THRES_i$ 는 부류 i 에 속하지 않는 패턴들을 거부하도록 설정된 임계값이다. 이러한 거부 조건을 코딩하기 위해 특징들을 설계하고 특징값을 계산하는 루틴을 작성한 다음 $THRES_i$ 값을 설정한다. 이 실험에서 이러한 과정은 수작업으로 수행되었으며 각각의 모듈은 10개에서 20개의 거부조건을 가지고 있다. 자세한 사항은 참고문헌 [24]에 잘 나타나 있다.

숫자 데이터 베이스로는 Concordia 대학의 Suen 교수에 의해 만들어진 훈련 데이터 4,000자와 테스트 데이터 2,000자가 사용되었다. 훈련 데이터와 테스트 데이터에 대한 성능은 각각 (95.95%, 0.40%, 3.65%), (92.35%, 0.85%, 6.80%)이다. (a, b, c)에서 a는 정인식률, b는 오인식률, c는 기각률이다. 자체적으로 수집한 패턴 데이터베이스도 훈련과 테스트에 사용되었다. 훈련 데이터로 8,000자와 테스트 데이터로 8,000자가 사용되었으며 각각의 인식률은 (94.35%, 0.05%, 4.70%)와 (91.30%, 1.65%, 6.25%)이다.

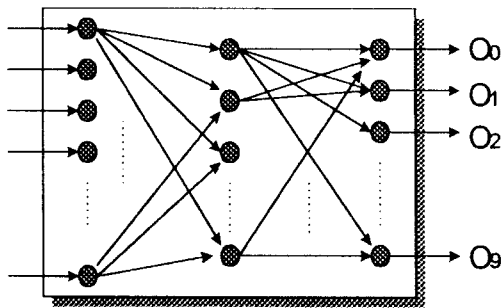
2.2 훈련 예 2

신경망 인식기는 그림 1(b)에서 보는 바와 같이 10개의 이진 네트워크 $N_i(i=0, 1, \dots, 9)$ 로 구성되어있다. 이 네트워크는 완전 연결 되어있으며, 1개의 입력층과 1개의 은닉층 그리

고 1개의 출력층으로 구성되어있다. 입력층은 21×16 의 메쉬를 수용하기 위하여 336개의 노드로 구성되어있다. 은닉층은 70개의 노드로 구성되어있다. 훈련 알고리즘으로는 역전파 알고리즘이 사용되었으며 사용된 네트워크의 특이할 점은 출력층의 노드 개수가 A(Accept Mode) 와 V(Veto Mode) 2개로 구성되어있다는 것이다. 즉, 그림 1(b)에서 보는 바와 같이 네트워크 N_i 에 대해서 부류 i 의 훈련 패턴의 기대 출력은 (1, 0)이다. 또한 나머지 9개의 부류에 대한 훈련 패턴의 기대되는 출력은 (0, 1)이다.

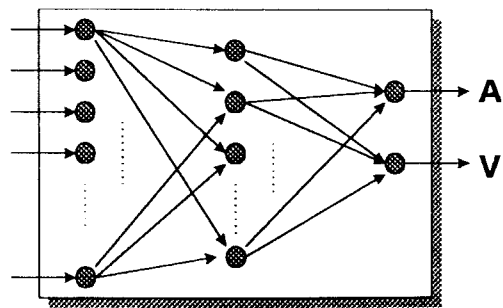
Suen 데이터베이스의 훈련 데이터와 테스트

데이터의 성능은 각각 (97.18%, 0.25%, 2.57%), (84.65%, 1.8%, 13.55%)이다. 자체 수집한 데이터베이스에 대한 성능은 (94.68%, 0.56%, 4.76%)과 (86.18%, 3.3%, 10.5%)이다. 현재 실험에서는 모든 N_i 에 대해서 21×16 메쉬 형태의 똑같은 특징을 사용하였다. 한가지 중요시 할 점은 각각의 N_i 에 대해서 서로 다른 네트워크 구조와 서로 다른 특징을 사용할 수 있다는 점이다. 이러한 접근 방법이 보다 나은 성능을 가져올 것이다. 또한, 모듈러 네트워크의 성능비교 평가를 위해 그림 1(a)에서 보는 바와 같이 10개의 출력 노드를 가진 네트워크를 훈련 시켰다. 그 결과



Training data	
input pattern	expected output
{0}	→ (1,0,...,0)
{1}	→ (0,1,...,0)
⋮	⋮
{9}	→ (0,0,...,1)

(a) 전통적인 네트워크



Training data	
input pattern	expected output
{0}	→ (0,1)
{i}	→ (1,0)
⋮	⋮
{9}	→ (0,1)

(b) 모듈러 네트워크(N_i for the class i)

<그림 1> 신경망 네트워크 구조

Suen 데이터베이스의 훈련 데이터와 테스트 데이터에 대해서 (95.35%, 1.35%, 3.3%), (84.35%, 7.8%, 7.85%)의 성능을 보였다. 이 실험 결과는 분명히 모듈러 방법의 성능이 보다 낫다는 것을 보여준다.

3. 파이프라인 다중 알고리즘

그림 2는 3단계의 파이프라인 구조를 보여준다. 각각의 단계에서 사용된 인식기는 모듈러 구조를 가진다. i 단계 인식기는 $i-1$ 번째 인식기를 보조하도록 설계하는 것이 바람직하다. i 번째 인식기($i \geq 2$)는 3.1절과 3.2절에서 설명될 검증 또는 2차 기회 역할을 한다.

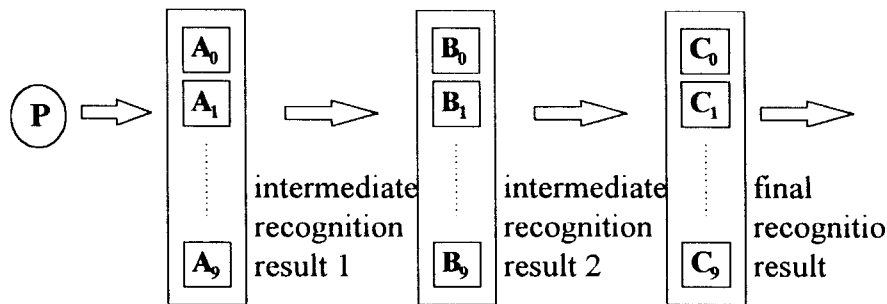
3.1 검증

만약 입력 패턴이 첫 번째 단계에서 부류(i)로 분류되었다면 두 번째 단계에서 i 번째 부인식기가 첫 번째 단계의 결과를 검증하기 위해 활성화된다. 이 부인식기가 '예'라고 했을 경우

에만 비로소 그 패턴은 부류 i 에 분류되어진다. 일반적으로 이러한 방법은 인식률은 떨어질 수 있지만 신뢰도 측면에서는 향상된 결과를 가져온다. 실험에서는 집합 R_i (2.1 절에서의 규칙 기반 인식기)는 첫 번째 인식기로 쓰여졌으며, N_i (2.2 절에서의 신경망 기반 인식기)는 두 번째 인식기로 사용되었다. 훈련 데이터와 테스트 데이터의 성능은 각각 (93.6%, 0.05%, 6.35%), (81.5%, 0.15%, 18.35%)이다. 성능 평가에서 보는 바와 같이 신뢰도는 상당히 향상되었다.

3.2 2차 기회

이 방법에서 두 번째 단계는 입력 패턴이 첫 번째 단계에 의해서 거부되었을 때만 호출된다. 두 번째 단계의 인식기는 첫 번째 단계에서 거부된 패턴들을 이용하여 훈련하는 것이 좋다. 만약 그렇게 된다면 두 번째 단계는 첫 번째 단계를 보조하도록 전문화되어진다. 이러한 방법은 인식률의 향상을 가져올 것이다.

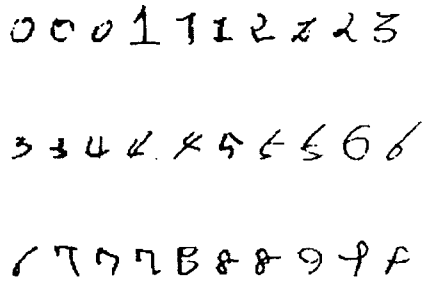


〈그림 2〉 파이프라인 구조

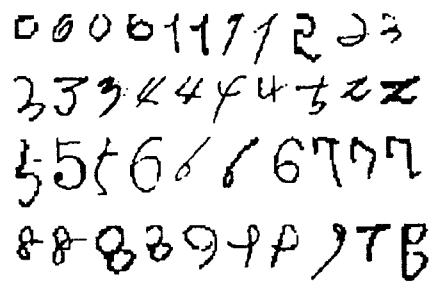
높은 신뢰도를 가지기 위해서는 처음 인식기와 두 번째 인식기 모두 신뢰도가 높아야된다. (이 실험에서는 첫 번째 단계에서 거부되는 패턴의 한정된 개수로 인하여, 두 번째 단계에서 훈련 데이터를 첫 번째 인식기의 훈련 데이터와 같은 데이터를 사용하였다.) 첫 번째 단계로 R가 사용되었고 두 번째 단계로써는 Ni가

사용되었다. 테스트 데이터에 대한 인식률은 (96.08%, 2.03%, 1.87%)으로서 인식률의 향상을 가져왔다.

그림 3은 입력 패턴에 대한 오인식(그림 3(a))과 거부된(그림 3(b)) 패턴들의 예를 보여주고 있다. 그리고 표 1은 각각의 부류에 대한 인식 결과를 보여주고 있다.



(a) 오인식된 패턴들



(b) 거부된 패턴들

〈그림 3〉오인식과 거부된 패턴 예

〈표 1〉인식 결과

	0	1	2	3	4	5	6	7	8	9	정인식	오류	거부
0	778	0	0	1	0	1	2	2	0	0	778	6	14
1	0	789	2	0	1	1	0	1	0	1	789	6	5
2	0	2	782	0	2	0	2	0	0	0	782	6	7
3	1	1	2	767	2	9	2	1	2	3	767	20	13
4	7	2	0	0	765	0	4	0	1	2	765	16	19
5	0	3	0	1	4	766	4	0	1	0	766	13	21
6	6	15	0	0	0	4	758	0	2	0	758	27	15
7	2	2	3	0	1	0	0	784	0	0	784	8	8
8	4	1	0	0	0	4	4	0	756	7	756	20	24
9	3	1	0	6	9	4	0	14	4	735	735	41	24
Σ											7,680	163	150

4. 결론

필기체 숫자 인식 프로그램을 설계하는데 있어서 구조적인 면의 중요성을 강조하고 모듈러 설계와 파이프라인 연결이라는 두 가지 구조적 설계를 제안했다. 또한 실험을 통해 이러한 구조의 장점을 입증하였다. 비록 이러한 생각은 몇몇 문헌에서 이미 사용되어졌지만, 본 논문은 필기체 숫자 인식 프로그램의 구조적인 면을 명시적으로 다루었다는데 의미가 있다.

참고문헌

1. H Bunke, P.S.P. Wang, and H.S. Baird(Edited), *Document Image Analysis*, World Scientific, 1994.
2. H.S. Baird, H. Bunke, and K. Yamamoto(Edited), *Structured Document Image Analysis*, Springer-verlag, 1992.
3. Y.L. Cun, O. Matan, B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, L.D. Jackel, and H.S. Baird, "Handwritten zip code recognition with multilayer networks," *IEEE Computer*, pp. 35-40, 1990.
4. N.W. Strathy and C.Y. Suen, "A new system for reading handwritten ZIP codes," *Proceedings of ICDAR*, pp. 74-77, 1995.
5. T.M. Ha, H. Bunke, "Model-based Analysis and Understanding of Check Forms," in *Document Image Analysis* (Edited by H. Bunke et al), *World Scientific*, pp. 1053-1080, 1994.
6. G. Nagy, S. Seth, M. Viswanathan, "A Prototype Document Image Analysis System for Technical Journals," *IEEE Computer*, Vol. 25, No. 7, pp. 10-22, 1992.
7. L. Boatto, V. Consorti, M.D. Buono et al, "An Interpretation System for land Register Maps," *IEEE Computer*, Vol. 25, No. 7, pp. 25-33, 1992.
8. W. Verschuere, B. Schaeken, Y.R. Cotert, and A. Hermanne, "Structural recognition of handwritten numerals," *Proceeding of Seventh International Conference on Pattern Recognition*,

- July, pp. 760-762, 1984.
9. M. Shridhar and A. Badreldin, "Recognition of isolated and simply connected handwritten numerals," *Pattern Recognition*, Vol. 19, No. 1, pp. 1-12., 1986.
 10. J.S. Huang, K. Chuang, "Heuristic approach to handwritten numeral recognition," *Pattern Recognition*, Vol. 19, No. 1, pp. 15-19, 1986.
 11. P. Ahmed, C.Y. Suen, "Computer recognition of totally unconstrained handwritten zip codes," *International Journal of Pattern recognition and Artificial Intelligence*, Vol. 1, No. 1, pp. 1-15, 1987.
 12. G. Baptista and M. Kulkarni, "A high accuracy algorithm for recognition of handwritten numerals," *Pattern Recognition*, Vol. 21, No. 4, pp. 287-291, 1988.
 13. A.G. Boisdulier and P. Vanhulle, "Topological modelization of handwritten digits based on a geometrical primitive decomposition," *Proceeding of 3rd USPS Advanced Technology Conference*, May, pp. 797-812, 1988.
 14. C.L. Kuan and S.N. Srihari, "A stroke-based approach to handwritten numeral recognition," *Proceeding of 3rd USPS Advanced Technology Conference*, May, pp. 1033-1041, 1988.
 15. I.S. Abuhaiba and P. Ahmed, "A fuzzy graph theoretic approach to recognize the totally unconstrained handwritten numerals," *Pattern Recognition*, Vol. 26, No. 9, pp. 1335-1350, 1993.
 16. F.A. Kananagar and M.J. Cykana, "Recognition of handwritten numerals using a multilayer neural network," *Proceeding of 3rd USPS Advanced Technology Conference*, May, pp. 768-781, 1988.
 17. Y. Lee, "Handwritten digit recognition using k nearest-neighbor, radial-basis function, and backpropagation neural network," *Neural Computation*, pp. 440-449, 1991.

18. 김영준, 이성환, "유전자 알고리즘과 결합된 다층 역전파 신경망을 이용한 무제약 필기체 숫자의 오프라인 인식," 제6회 영상처리 및 이해에 관한 워크샵, pp. 186-193, 1994.
19. P. Gader, B. Forester, M.G. Berger, A. Gillies, B. Mitchell, M. Whalen, and T. Yocum, "Recognition of handwritten digits using template and model matching," *Pattern Recognition*, Vol. 24, No. 5, pp. 421-431, 1991.
20. F. Kimura and M. Shridhar, "Handwritten numeral recognition based on multiple algorithms," *Pattern Recognition*, Vol. 24, No. 10, pp. 969-983, 1991.
21. C.Y. Suen et al, "Computer recognition of unconstrained handwritten numerals," *Proceedings of IEEE*, Vol. 80, No. 7, July, pp. 1162-1180, 1991.
22. 백종현, 정성운, 이일병, "다중결합 방법을 이용한 필기체 숫자 인식," 제1회 문자인식 워크샵, 충북대학교, pp. 149-156, 1993.
23. L. Mui, A. Agarwal, A. Gupta and P.S.P. Wang, "An adaptive modular neural network with application to unconstrained character recognition," in *Document Image Analysis* (Edited by H. Bunke et al), World Scientific, pp. 1189-1203, 1994.
24. 오일석, "부류별 방법과 거부조건에 의한 필기체 숫자의 인식," 제 2회 문자인식 워크샵, pp. 41-48, 1994.