

동적 3-D 뉴로 시스템을 이용한 오프라인 필기체 숫자 인식

김기택*, 권영철**, 이수동*

*울산대학교 컴퓨터 정보통신 공학부

**거제대학교 컴퓨터 정보계열

e-mail:kkt8023@mail.ulsan.ac.kr

Off-line Handwritten Digit Recognition Using A Dynamic 3-D Neuro System

Ki Taek Kim*, Young Chul Kwon**, Soo Dong Lee*

*Dept of Computer Engineering and Information Technology, University of Ulsan

**Dept of Computer & Information, Koje College

요 약

본 논문은 동적 3-D 뉴로 시스템(A Dynamic 3-D Neuro System)모델을 이용하여 오프라인 필기체 숫자 인식 실험을 하였다. 3-D 뉴로 시스템 모델을 사용함으로써 기존에 교육된 정보를 유지하면서 새로운 정보를 추가할 수 있는 추가학습이 가능했고, 동일한 범주의 정보에 대해서는 반복교육 횟수에 따라 교육정도가 점점 누적되는 반복교육이 가능했다. 교육과정을 통해 누적된 정보로부터 일반화된 패턴(Generalized Pattern)을 도출해 인식시 사용할 수 있었다. 패턴 인식기는 피드백루틴을 통해 미지의 입력이미지를 원형이미지로 복원한 후, 그 결과데이터를 사용하여 문자를 인식하도록 동작한다. NIST의 MNIST 데이터베이스를 사용해 실험을 하였고, 결과로 99.0%의 정인식률을 얻었다.

1. 서론

현재까지 필기체 인식과 관련된 많은 시스템 모델이 제안되었다. 대표적인 모델로는 원형 정합을 이용한 템플릿 매칭 방법, 확률 통계적인 방법, 구조적 방법, 신경회로망을 이용한 방법 등이 있다[1]. 그 중에서 신경회로망을 이용한 방법은 오프라인 문자인식에서 좋은 성능을 발휘하는 것으로 알려져 있다.

그러나 기존의 신경회로망을 이용한 방법들은 인식률을 높이기 위해 전처리 과정 중에 많은 변형을 사용함으로써 전처리 과정이 복잡해지고, 연결강도를 보정하기 위한 반복계산 등으로 인해 학습 시에 많은 시간이 걸리는 문제점이 있다. 또한 기존의 신경회로망을 이용하는 방법은 추가학습 시에 전체내용을 재학습 시켜야 하는 문제점이 있다.

본 논문에서는 필기체 인식을 위해서 Aleksander와 Stonham이 제안한 WNN(Weightless Neural Net, Ram Based Neural Network)에 3-D 기법[2]과 Lee[3]의 Feedback 기법이 통합된 인식 모델을 사용하였다. WNN[4,5,6,7]은 복잡한 전처리 과정이 필요없고, 단 한번의 학습으로 학습이 가능하며, 하드웨어적으로 구현하기가 용이하고, Bledsoe와 Browning 이 제안한 N-Tuple기법[8]을 사용한다.

실험에 사용한 데이터는 NIST의 MNIST 데이터베이스[9]

를 사용하였고, 6만개의 학습 데이터와 1만개의 테스트 데이터를 사용하여 인식률을 조사하였다. 실험결과에 의하면 99.0%의 정인식률을 얻음으로써 본 논문에서 사용한 패턴 인식모델이 패턴인식에 있어 유용함을 알 수 있었다.

본 논문의 구성은 다음과 같다. 2절에서는 관련 연구에 대해 소개한다. 3절에서는 본 논문에서 사용한 인식 모델의 학습 및 인식 알고리즘에 대해 소개하고, 4절에서는 실험 결과를 분석한다. 마지막으로 5절에서는 결론과 향후 연구방향을 제시한다.

2. 관련 연구

1. 램 기반 신경회로망(A Ram Based Neural Network)

1.1 램(A Random-Access Memory)

판별자의 기본 구성요소 인 램(RAM)의 동작과정은 다음과 같다. 학습은 입력이미지로부터 N-Tuple만큼의 픽셀을 랜덤 하게 선택하여 어드레스디코더(Address Decoder)에서 주소정보화 한 후, 해당하는 주소(Address또는 Bit)에 값(1 또는 0)을 기억시킴으로써 이루어진다. 입력노드로부터 N개의 노드를 추출하여 조합하는 방법은 퍼포먼스(Performance)를 높이기 위해 랜덤하게 추출하는 방법을 사용하고, 램의 크기는 N비트(Bit)의 주소정보가 가질 수 있는 크기인 2^N비트가 필요하다. 인식은 학습시 입력과정

과 동일한 순서로 N-Tuple의 픽셀을 선택하여 어드레스 디코더에서 주소정보화 한 후, 기억된 저장 공간의 해당 주소비트에서 값을 불러온다.[7]

예를 들면, 그림1과 같이 N이 4일 때 램의 총 비트수는 $2^4=16$ 개이고, 각 입력이 $I_1=1, I_2=0, I_3=0, I_4=1$ 일 때, 각 입력 값으로부터 만들어지는 2진 주소값, 즉 $1001_{(2)}=(9_{(10)})$ 이므로 Bit₉에 값(1 또는 0)을 기억시킨다. 인식시에는 Bit₉(또는 address 9)에 기억된 값을 출력한다.

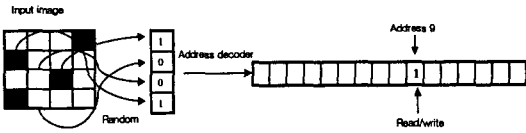


그림1. 램(A Random-Access Memory)

1.2 판별자(A Discriminator)

판별자는 그림 2와 같이 N개의 입력을 받는 K개의 램과 Summing Device(Σ)로 구성되고, 각 램의 입력라인은 입력이미지와 랜덤 매칭방법으로 각각 연결되어 있다. N개의 입력을 가지는 판별자는 $2^N(K)$ 개의 램으로 구성된다. 예를 들어, 4x4픽셀 입력이미지에 대하여 각 램 당 4개씩의 입력을 가지는 판별자를 구성하면 전체 $2^4(16)$ 개의 램으로 구성 할 수 있고, 입력이미지의 각 픽셀은 판별자에 대한 입력으로 각각 4번씩 입력패턴으로 사용된다. 즉 판별자에 대한 입력패턴의 총수는 $KN(16x4)$ 개 64개가 된다. Summing Device(Σ)는 각 램으로부터 출력을 더하여 그 합계를 출력하는 함수이다. Summing Device(Σ)의 출력값인 반응(Response) R값이 클수록 입력으로 들어온 패턴이 교육된 패턴과 유사함을 의미한다. 즉 반응값 R이 정해진 기준치 (Threshold)보다 크면 인식이 됨을 의미한다 [7].

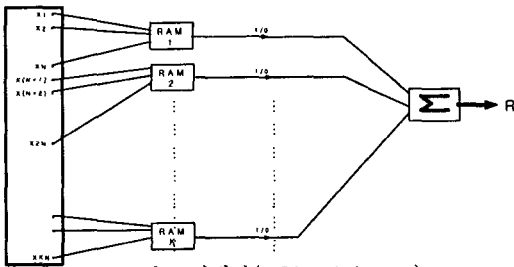


그림2. 판별자(A Discriminator)

1.3 다중 판별자(A Multi-Discriminator)

숫자 인식시스템과 같이 다양한 카테고리를 분류하는 시스템은 여러 개의 판별자를 사용한다.

이처럼 다중 판별자를 사용한 대표적 예로는 WISARD 시스템(Wilkie, Stonnen, and Aleksander's Recognition Device)[7, 10]이 있고 그림3과 같이 동작한다.

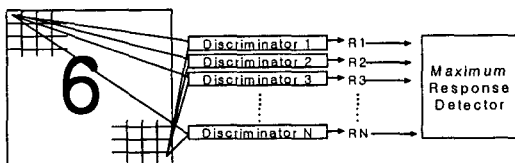


그림3. 다중 판별자(A Multi-Discriminator)

다중 판별자(Multi-discriminator)는 각 카테고리 교육 정보를 기억하고 있는 판별자와 각 판별자의 출력으로부터 가장 큰 반응(Maximum Response)을 선택하는 MRD(Maximum Response Detector)로 구성된다.

인식시 입력패턴을 각 판별자에게 보이고 MRD에서는 각 판별자들의 반응값(R)으로부터 가장 반응값(R)이 큰 것을 찾아 출력한다. 이 MRD로부터 나오는 최종 출력이 인식된 카테고리로 결정된다.

2. 3-D 뉴로 시스템(A 3-D Neuro System)

인간의 학습능력은 교육에 의해 축적된 지식이 반복교육에 의해 누적되고 또한 동일한 내용의 반복교육에 관해서는 다른 내용에 비하여 더욱 강한 반응을 나타내는 특징을 지니고 있다. 예를 들어, 단어의 암기는 동일한 단어를 반복하여 교육함으로써 암기하게 되고, 또한 여러 단어를 교육함으로써 단어의 암기수도 증가하게 된다.

3-D 뉴로 시스템[2]의 신경세포는 기존의 램 기반 신경회로망과 유사한 구조이나, 정보의 기억소자 단위를 '1' 또는 '0'의 값을 가지는 비트로 하지 않고 복수개의 램을 사용하여 교육의 반복회수(가중치)를 기억할 수 있도록 구성된다. 즉 기존의 램 기반 신경회로망에서는 하나의 램이 하나의 신경세포에 대응하는데 반해 3-D 뉴로 시스템에서는 신경세포에서 정보를 기억하는 단위소자 역할을 한다. 또한 교육정보의 누적을 위해 신경세포 내부에 연산기를 도입하여 기억된 정보를 변환하여 연산하도록 구성된다.

이러한 신경세포의 구조로 인해서 3-D 뉴로 시스템은 기존의 교육된 정보를 유지하면서 새로운 정보를 추가할 수 있는 추가교육이 가능하고, 또한 동일한 범주의 정보에 관해서는 반복교육 횟수에 따라 교육정도가 점점 누적되는 반복교육도 가능하다. 이런 반복교육에 의해 누적된 정보로부터 공통된 특징에 대해서는 주의 집중 기능과 불필요한 정보에 대해서는 망각 기능을 사용하여 일반화 패턴(Generalized Pattern)을 생성해 낼 수 있다. 즉 추가교육과 반복교육에 의해 누적된 각 램에 기억된 값 중에서 일정한 기준치(Threshold) 이상의 값을 기억하고 있는 부분만으로 재구성한 일반화 패턴을 추출할 수 있다.

본 논문에서 패턴인식기의 학습망은 3-D 뉴로 시스템을 기반으로 동작하고 인식시 교육과정 동안 누적된 교육정보로부터 일반화 패턴을 도출해 사용한다.

3. 패턴 인식기

패턴 인식기의 구조는 그림4와 같이 콘텍스트(Context)의 영향을 고려한 패턴 인식을 위해 학습망군과 인식망군의 이중망구조로 이루어진다. 학습망군과 인식망군 사이에 귀환선로를 두어 서로 대화가 가능하게 하고, 인식망의 결론이란 콘텍스트가 학습망의 프로세스에 관여하게 함으로써 진단의 결론을 점점 증폭시켜가며 안정하고 강력한 결론(Strong Decision)으로 유도 시킨다.

인간은 인식시 그때까지 경험한 지식들로부터 공통된 특징을 추출하여 하나의 대표적인 개념의 표준패턴을 생성하고, 외부의 대상을 비교해 가며 자신이 만족을 느낄 때까지 프로세스를 계속 수행한 후 비로써 결론을 내린다.

본 논문에서 사용한 패턴 인식기는 미지의 입력패턴을 인식시 학습망군에서 교육 받은 내부상태와 외부의 입력

을 피드백(Feedback)을 통해서 끊임없이 비교해 가며 변형을 제거시켜 가며 프로토타입 패턴(Prototype Pattern)을 내부적으로 만들어 간다. 이렇게 피드백 루프를 통해서 전처리된 프로토타입을 인식망군에서 인식하여 강력한 결론을 내린다.

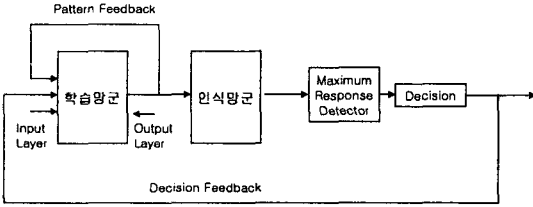


그림4. 패턴 인식기

인식 사이클은 다음의 과정을 반복 수행한다.

1. 미지의 입력패턴을 학습망의 입력과 출력에 보인다.
2. 학습망의 각 판별자들의 출력으로부터 인식망에서 결론을 내린다.
3. 인식망에서 최대 응답을 보인 학습망에서의 출력 패턴을 다시 입력으로 한다.
4. 정해진 횟수만큼 피드백(Feedback)을 하지 않았으면 프로세스 1로간다.
5. 인식망에서 최종 출력값을 구한다.
 - 1) 구해진 결과가 기준치(Threshold)보다 크면 최종 결론을 내린다.
 - 2) 주어진 기준치를 넘지 못하면 프로세스6으로 간다.
6. 처음 인식 사이클에서 인식망에서 내린 최대 응답의 바로 다음 응답을 내린 학습망의 판별자를 선택하여 그 판별자와 인식망의 조합으로 프로세스 1부터 다시 시도한다.
 - 1) 재시도 후 최종 응답이 기준치 이하이면 인식 프로세스를 종료한다.

이러한 인식 스키마(Schema)는 미지의 입력을 학습망군에서 프로토타입 패턴으로 재인식될 수 있고, 이렇게 전처리화(Preprocessing)된 패턴을 인식망군에서 인식하게 함으로써 강력한 결론(Strong Decision)을 내릴 수 있다.

또한 학습망군의 출력과 인식망군의 입력 사이에 귀환 루프(Feedback Loop)를 두어 시스템 모델이 만족할만한 결론을 얻을 때까지 프로세싱을 수행시키며, 어느 정도 기다려도 만족할 만한 결론을 내릴 수 없을 때는 억지로 결론을 내리는 대신 새로운 시도를 가지고 처음부터 다시 하여 시스템의 신뢰도를 높이고 융통성을 지니게 한다.

3.1 학습망

학습망군은 프로토타입 이미지와 필기체 이미지를 대상으로 교육하고, 교육을 통해 점진적으로 각 카테고리 이미지 정보를 교육, 흡수(Perception)할 수 있게 된다. 이렇게 생성된 정보는 차후 인식시에 일반화된 집합으로 동작시켜 소기의 목적을 달성하는데 사용된다.

학습 사이클은 다음의 과정을 반복 수행한다.

1. 기억층을 초기화한다. 처음 교육일 경우 램의 값을 '0' 으로 초기화하고, 추가교육일 경우 교육된 정보를 읽어 들여 해당 램에 세팅한다.
2. 입력과 출력에 각각 필기체 이미지와 프로토타입을 보인다. 만일 끝이면 프로세스를 종료한다.
3. 해당 범주의 판별자에 해당 프로토타입 정보를 기

역시킨다. 정보 기록시 저장할 램에 이전에 교육된 값이 존재하면 그 값에 현재 교육될 값을 더한다.

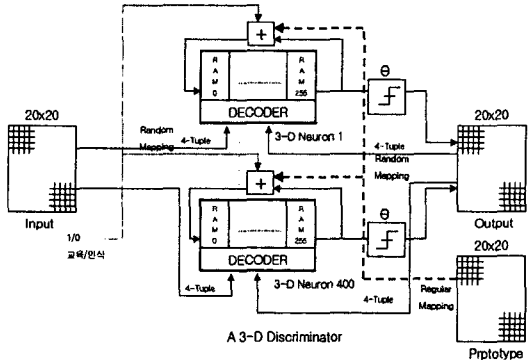
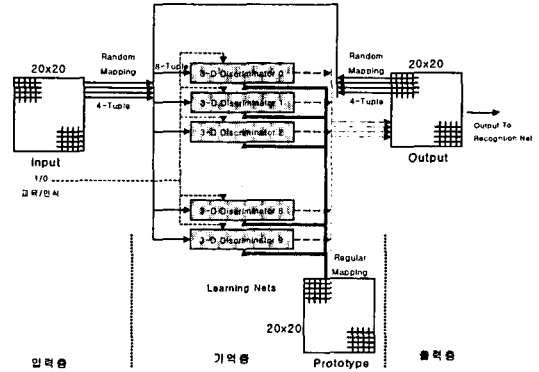


그림5. 학습망

학습망은 그림5와 같이 2차원 배열의 입력층과 여러 개의 3-D 판별자로 구성된 기억층, 그리고 입력층과 동일한 크기의 2차원 배열인 출력층으로 구성된다. 입력이미지는 20x20 크기의 0,1의 흑백이미지를 사용하였고, 8-Tuple Method를 사용해 입력층과 출력층에서 랜덤 하게 각 4개의 픽셀 중 8개의 픽셀을 조합하여 램의 입력 주소 정보로 사용하였다. 기억층의 각 3-D 판별자는 256개의 램으로 구성되는 400개의 3-D 신경세포로 구성된다. 각 신경세포에 있는 케환회로의 연산기는 케환된 램의 내부 값과 원형이미지로부터 입력된 정보('0' 또는 '1')을 가산하여 다시 램에 기억시키도록 동작하는 3-D 기법을 사용한다.[2]

3.2 인식망

인식망은 그림6과 같이 2차원 배열의 입력층과 판별자로 구성되는 기억층, 그리고 각 판별자로부터 가장 큰 반응을 구하는 인식층 MRD로 구성된다. 입력이미지는 학습망의 출력층의 입력층으로 사용되는 20x20 크기의 프로토타입 이미지를 사용하였고, 4-Tuple Method를 사용해 램의 입력 주소 정보로 사용하였다. 각 판별자는 100개의 램으로 구성된다. 인식망의 교육은 입력에 프로토타입 패턴을 보이고 해당 램에 '1' 을 가해 해당 판별자가 프로토타입을 수용하도록 교육시킨다.

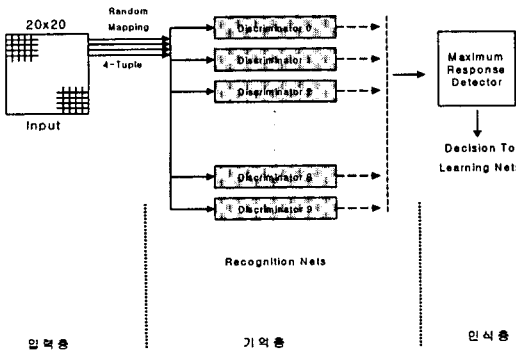


그림 6. 인식망

4. 시뮬레이션

필기체 인식시스템의 숫자 데이터는 28x28 픽셀의 고정된 크기로 정규화 되어있고, 문자정보는 20x20 크기로 중앙에 위치하도록 전처리 되어 있는, NIST의 MNIST 필기체 데이터로 6만개의 교육데이터와 1만개의 테스트 데이터를 사용하였다. 원 이미지 28x28픽셀을 20x20픽셀 크기로 축소하여 입력의 크기를 줄여 사용하였고 입력크기를 줄임으로써 판별자 램의 수를 줄일 수 있었다.

실험을 위해 사용된 시스템은 IBM-pc Celeron 533hz, 256M 이며 C언어를 사용하여 알고리즘을 구현하였다. 테스트 데이터에 대한 정인식률은 그림7에 나타내었고 x축은 각 숫자를 나타내고 y축은 인식률을 나타낸다. 전체 인식률은 99.0%의 결과를 얻었다.

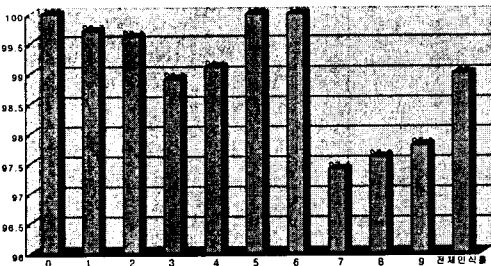


그림 7. 테스트 데이터에 대한 정인식률

그림8에는 인식과정에서 오인식된 문자들을 나타내었다. (1)은 입력1을 8로 오인식 했는데 일반적인 필기체 1이 가지는 픽셀 수보다 더 많은 픽셀정보로 인해 오인식했고 오인식 발생을 줄이기 위해서는 픽셀수가 지나치게 많은 데이터에 대해서는 크기 정규화(size normalization)와 같은 작업을 통해서 입력 픽셀 양을 줄이는 작업이 필요하다. (2),(3)과 같은 경우는 필기체 4와 9의 유사구조, 7과 9의 유사구조로 인해서 오인식이 발생하였다. (4)는 7을 1로 오인식 하였다. 이는 7이 가져야 할 픽셀정보의 부족으로 인해 오인식이 발생했다. (5)역시 8을 1로 오인식 하였다. 오인식이 발생한 경우는 육안으로도 구별하기 힘들 정도의 구조적으로 유사한 경우나 입력 정보 픽셀 양이 지나치게 많거나 너무 적은 경우에 즉 오인식이 발생함을 볼 수 있었다.

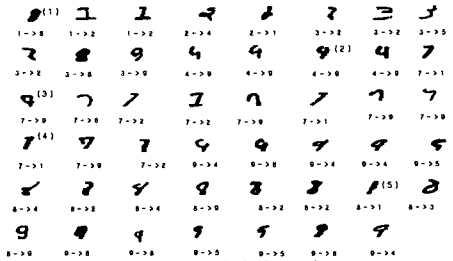


그림 8. 오인식된 문자들

5. 결론

본 논문에서는 오프라인 필기체 숫자 인식에 적합한 램 기반 신경회로망에 피드백 기법과 3-D 기법이 통합된 패턴 인식 모델(A Dynamic 3-D Neuro System)을 사용하여 필기체 인식 실험을 하였다. 학습망을 3-D 뉴로 시스템으로 구성함으로써 추가학습과 반복학습이 가능했고, 교육된 결과데이터에 기준치(Threshold)를 정해줌으로써 원형을 생성해 인식시에 사용할 수 있었다. 또한 피드백 기법을 통해 미지의 입력으로부터 해당 원형을 복원시켜 그 결과를 인식망의 입력으로 하여 강력한 결론을 내릴 수 있었다. 동적 3-D뉴로 시스템을 MNIST에 대해 실험한 결과 99.0%의 정인식률을 얻었는데 이로 본 논문에서 사용한 인식모델이 패턴인식에 있어 유용함을 알 수 있었다.

인식시에 입력이미지가 각 숫자가 지나야 할 정보량이 지나치게 많거나 적은 경우와 육안으로 구별하기 힘들 정도로 유사할 때에 오인식이 발생했는데 위와 같은 경우 오인식을 줄일 수 있는 연구가 향후 행해져야 하겠다.

참고문헌

[1] 이성환, "오프라인 필기체 문자인식 기술의 현황", 한국정보과학회지, 제11권 제5호, 1993년 10월.
 [2] 권영철 [著], "경험 유관 2진 신경회로망", 울산대학교 대학원, DHJ-00096, 1995.
 [3] Lee, S.D., "Dynamics of Feedback for Artificially Intelligent Pattern Recognition", MSc. Dissertation, Brunel Univ., England, 1981.
 [4] Aleksander. I. "Random Logic Nets: Stability and Adaptation", International Journal Man-Machine Studies, 5, 115-131, 1973.
 [5] Aleksander. I. and Stonham, T.J., "Guide to pattern recognition using random-access memories", Computers and Digital Techniques, 2, 29-40, 1979.
 [6] Aleksander. I. T. J. W. Clarke and A. P. Braga, "Weightless Neural System :A Unified Approach to Their Analysis and Design", ICONIP '94 Tutorial, 1994.
 [7] I. Aleksander and H. Morton, "An Introduction to Neural computing", Chapman & Hall, 1990.
 [8] Bledsoe. W.W. and Browning, I., "Pattern Recognition and Reading by Machines", Proceedings of the Eastern Joint Computer Conference, pp. 225-232. 1959.
 [9] Yann LeCun, NEC Research Institute, <http://yann.lecun.com/exdb/mnist/index.html>
 [10] Aleksander. I., Thomas, W.V., and Bowden, P.A., "WISARD a radical step forward in image recognition", Sensor Review, 120-124, July 1984.