

인쇄체 한글문자의 인식을 위한 계층적 신경망[†]

(A Hierarchical Organization of Neural Networks for Printed Hangul Character Recognition)

조 성 배* 김 진 형**
(Sung-Bae Cho) (Jin H Kim)

요 약

본 논문에서는 한글을 인식하고자 하는 기존 신경망 접근방법의 문제점을 분석하고, 인쇄체 한글문자를 실용적으로 인식할 수 있도록 하는 계층적 구조의 신경망을 제안한다. 이 신경망은 유형분류 신경망에 의하여 입력된 문자영상을 6가지 유형중의 하나로 분류한 후, 해당하는 문자인식 신경망을 이용하여 문자를 자소단위로 인식하는 구조로 되어 있다. 여기서는 또한 한글인식과 같은 큰 문제를 해결하는데 도움이 되는 변형을 첨가한 학습방법을 제시하고, 몇가지 실험을 통해 이와 같은 구조를 갖는 신경망 방법의 유용성을 입증한다. 상용하는 990자에 대하여 실험한 결과 변형을 첨가한 학습으로 99.28%의 인식률을 얻을 수 있었는데, 이것은 기존의 방법인 트리 Classifier와 비교하여 보다 우수한 것이었다. 또, 일반화 실험과 오인식 문자의 분석을 통해 신경망 접근 방법이 유용함을 볼 수 있었다.

ABSTRACT

In this paper, we consider the previous approaches for Hangul character recognition using neural network, and propose the hierarchical neural network which practically recognizes printed Hangul character. This system, so called Net-eye, is composed of the type classification network for grouping the characters into one of the six types by structural similarity, and six recognition networks which recognize relevant characters corresponding to each group. And the noise included learning method which is useful for practical recognition problems is proposed.

Experimental results with most frequently used 990 printed Hangul characters are as follows. By noise included learning, the recognition rate amounts up to 99.28%, which is superior to the conventional classification method, tree classifier. Neural network approach turns out to be very reasonable by means of the experiments such as analyzing the generalization capabilities and mis-classification characters.

[†] 이 논문은 시스템 공학 센터의 도움으로 Cray2s 슈퍼 컴퓨터를 사용 하였음

* 준 회 원 한국과학기술원 전산학과

** 종신회원 한국과학기술원 전산학과 교수

접수일자 1989 12 13

I. 서 론

문서에 기록된 정보를 컴퓨터에 자동으로 저장하고자 하는 문자인식은 인공지능과 패턴인식의 기본 응용 분야로 많이 연구되고 있다. 이러한 연구를 통하여 원형 비교 방법(template matching method), 통계적 방법(statistical method), 구조적 방법(syntactic method) 등이 고안되었으나 대부분 잡음에 민감한 약점을 갖고 있다. 기계에 의한 제한된 활자체로 구성되어 각 문자가 불변적인 특성을 갖는 인쇄체 문자의 경우에도 필기체 보다는 인식하기 쉽다고 하지만 실제로 입력된 문자를 처리할 수 있는 융통성 있는 방법은 아직까지도 제시되지 않고 있는 실정이다. 이것은 실제적인 입력문자를 인식하는데 기존의 방법으로는 해결하기 어려운 몇가지 문제점이 발생하기 때문이다. 인쇄체 문자라 할지라도 스캐너에서 발생하는 잡음이 추가되어 영상에 변형이 생길 뿐만 아니라, 글자체를 미적으로 보기 좋게 하고자 하여 발생하는 각 자소간의 접촉 때문에 제대로 인식하기 어렵다.

인간은 비교적 쉽게 해내는 문자의 인식을 컴퓨터가 하기 어려운 것은 기존의 계산을 주어진 입력에 대응하는 출력을 수학적인 모형에 의하여 구해내는 과정이라고 볼때, 그 사이의 연관관계를 간단한 수식으로 나타낼 수 없기 때문일 것이다. 이러한 문제를 해결하기 위한 접근 방법으로 인간의 두뇌 신경 조직을 모델로 한 신경망을 사용하려는 연구가 있다. 신경망은 프로그래머가 문제해결에 대한 정확한 지식을 갖지 못한 경우에도 그 자체의 적응능력으로 적절한 알고리즘을 생성해 낼 수 있다는 잠재력 때문에 처리 메카니즘이 분명치 않은 문자인식이나 음성인식 등의 분야에서 크게 각광을 받고 있다[18, 21, 24]. 특히, 문자인식 분야에서는 Backpropagation으로 학습시킨 신경망을 이용하는 연구가 활발히 진행되고 있다.

이와같은 이유에서 한글 문자를 인식하는데, 어느 정도의 변형이나 불완전한 입력에 대해서도 잘 작동한다는 신경망을 사용하려는 여러 시도가 있었다[1, 3, 4, 5, 6, 13]. 하지만 한글은 가능한 문자의 수가 11,172자에 이르는 큰 규모일 뿐만 아니라[10], 비교적 적은 수의 기본 자모가 서로 결합하여 문자를 만들기 때문에 문자간에 유사성이 커서 한글문자 인식은 매우 어려운 문제이다.

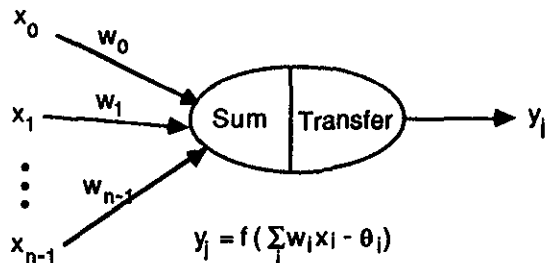
본 논문에서는 이와 같은 문제를 해결하여 인쇄체 한글문자를 실용적으로 인식할 수 있도록 하는 계층적 신경망 구조와 변형을 첨가한 학습방법을 제안하고 실험을 통해 그 유용성을 밝히고자 한다. 논문의 내용을 살펴보면 다음과 같다.

II장에서는 신경망 이론의 기초적인 아이디어를 소개하고, 신경망을 사용하여 한글을 인식하고자 하는 이제까지의 연구를 정리한다. 여기에서 기존의 신경망 접근 방법의 문제점을 분석함으로써 실용적인 인쇄체 한글인식 시스템을 신경망으로 구성할 수 있는 방법을 모색한다. 이러한 방법을 통하여 인쇄체 한글문자를 인식하기 위한 계층적 구조의 신경망, 망눈(Net-eye) 시스템을 III장에서 설계하고 구현한다. 망눈 시스템은 주어진 한글문자를 6개의 유형별로 분류한 후 해당하는 유형의 인식 신경망을 이용하여 문자를 인식한다. 이 장에서는 이와 같은 계층적 구조의 신경망이 한글문자 인식에 적절함을 보인다. IV장에서는 망눈 시스템의 인식률을 기술하고, 몇가지 실험을 통하여 이와같은 구조를 갖는 신경망의 유용성을 보인다. 즉, 기존의 패턴분류 방법인 트리 Classifier와의 비교를 통하여 신경망 분류기법의 우수성을 입증하고, 신경망을 사용한 인식의 결과가 심리학적 연구 결과와도 어느 정도 일치함을 보인다. 마지막으로 V장에서는 본 논문의 내용을 정리하고 앞으로의 연구방향을 기술한다.

II. 배 경

2.1 신경망 이론의 개요

신경망은 기본적으로 노드와 이들간의 연결로 구성된 그래프로서, 노드는 처리기(PE)라 하며, 연결은 연결 강도를 갖는다. 임의의 처리기 PE,는 (그림 2.1)과 같



(그림 2.1) 처리기의 기능

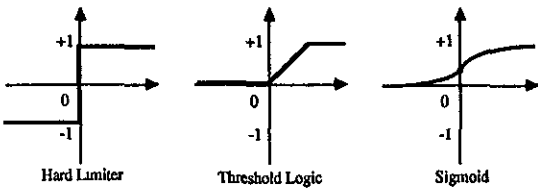
이 다른 처리기 PE_i의 출력 y_i를 그 입력 x_i로 받아 PE_j와 PE_i의 연결강도 w_{ij}에 대한 각 입력의 가중치 합 net_j를 다음과 같이 구하고,

$$net_j = \sum_i w_{ij} x_i$$

이 합을 정해진 전이함수(transfer function)에 적용시켜서 얻은 결과를 PE_j의 출력 y_j로 내보낸다

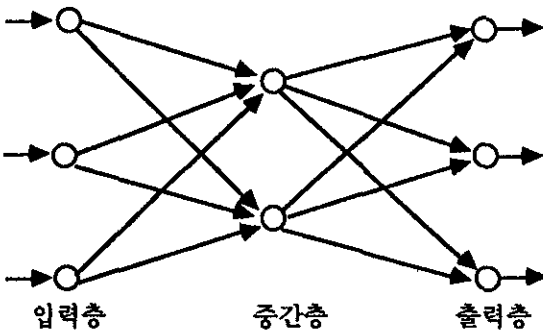
$$y_j = f(net_j + \theta_j)$$

여기에서 θ_j 는 PE_j의 고유 바이어스이다 전이함수는 (그림 2.2)에서와 같이 주로 3가지 종류의 함수, 즉 hard limit nonlinearity, threshold logic, sigmoid nonlinearity 함수 등이 쓰인다 [21] 전체의 구조는 이



(그림 2.2) 전이함수의 종류

와 같이 단순한 PE들이 (그림 2.3)과 같이 상호 연결되어 있으며, 각각의 PE는 신경망 전체의 입출력 면에서 볼때 입력 PE, 출력 PE, 내부 PE 등의 3 가지로 나뉘어 진다 이때, 신경망을 학습시킨다는 것은 신경망이 주어진 입력에 대하여 적당한 출력을 낼 수 있도록 각 노드 사이의 연결강도를 조정하는 것을 의미한다 [21, 24]



(그림 2.3) 신경망의 구조

단순한 기능의 처리기를 대규모로 연결시킨 구조를 갖는 신경망 모델은 학습 능력, 분산 지식 표현, 병렬 처리 능력, 결합 극복 능력 등의 공통적인 특성을 갖는

데, 이러한 특성으로 인하여 기존의 기호처리 중심의 인공지능 기법들이 잘 해결하지 못하는 음성 인식이나 영상 패턴인식 등의 여러 분야에서 좋은 성능을 발휘하고 있다. 신경망의 구분은 신경망의 연결 형태, 즉 망의 구조와 학습 알고리즘의 차이에 따라서 여러가지로 나눌 수 있으며, 대표적인 것으로는 Hopfield Network, Boltzmann Machine, 다층 Perceptron 등이 있다[21, 24].

2.2 신경망을 이용한 한글인식의 관련연구

한글은 그 가능한 문자의 수가 방대할 뿐만 아니라, 기본적으로 24개의 기본 자모를 규칙적으로 조합하여 문자를 이루기 때문에 각 문자 사이의 유사성이 커서 기존의 문자인식 방법으로는 인식하기 어려운 문제이다 최근에는 융통성있는 학습능력을 갖고 있는 신경망을 이용하여 한글을 인식하고자 하는 연구를 활발히 진행하고 있는데, 아직까지는 실용적인 성능의 시스템은 제시되지 않고 있다 본 절에서는 지금까지의 연구를 정리하고, 각 방법의 문제점을 분석하여 보다 유용한 시스템의 설계에 기반이 되도록 한다

현재까지의 연구중에서 성공적인 것은 대부분 기본자소를 입력으로 하거나[5, 13], 이상적인 형태를 가정하고 인위적인 잡음을 첨가하거나[6] 하는 등 간단한 영어나 숫자의 인식에 사용하는 신경망을 그대로 한글 인식에 사용해 본 것에 불과한 경우가 많다 기본자소를 입력으로 하는 전자의 경우에는 bar mask를 이용하여 입력하거나, Neocognitron 모델을 사용한 것인데 근본적으로 필기체를 인식하고자 한 것으로 문자 단위의 인식으로 확장하는 문제가 남아 있고, 최근 문자 인식에 적절하다는 RCE(Reduced Coulomb Energy) 모델을 사용하는 이 원돈등 [6]의 연구는 아직 시작단계에 있기는 하지만 이상적인 형태를 가정했을 뿐만 아니라 인식 대상의 문자 수가 너무 적어 그 유용성을 보이기에는 미흡한 점이 있다. 또한, 입력 문자에 전처리를 하여 DC 성분의 특징을 추출한 후 여의 분류만을 위해 신경망을 사용하는 경우[4]에는 특징추출 자체가 큰 부담일 뿐만 아니라, 전체 인식률이 이 단계의 성능에 크게 좌우되는 문제가 발생하여 자동적인 최적의 특징 추출이라는 신경망의 장점을 제대로 이용하지 못하는 단점이 있다

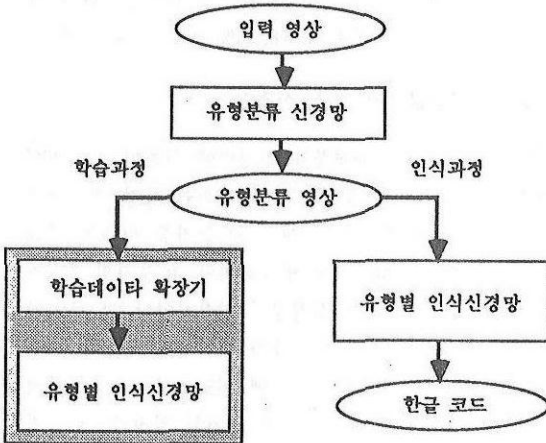
단지, 교병기[1]와 김병기[3] 등이 실제적인 입력으

로 스캐너에 의해 입력된 한글 문자를 인식하려고 시도한 바 있으나, 경쟁하는 신경망 조직을 사용한 전자의 경우 597자에 대해 부분 인식을 96.0%, 전체 인식을 83.6%의 매우 저조한 결과를 얻었고, 본 논문과 유사한 계층적 구조를 사용한 후자의 경우에도 입력영상을 20×20으로 정규화한 510자에 대해 분류율 97.67%, 전체 인식을 96.31%를 얻었다. 특히, 후자의 경우에는 유형을 불필요하게 15개씩이나 나누어 분류 신경망의 성능이 떨어져서 전체 인식을 감소시킬 수 있다.

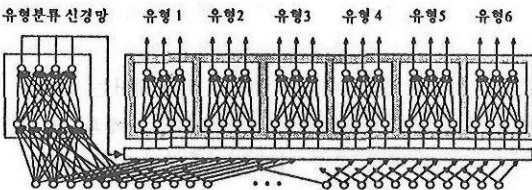
Ⅲ. 한글 인식을 위한 계층적 신경망

3.1 시스템의 구성

한글 인식을 위한 신경망 시스템은 기본적으로 입력된 문자의 유형을 분류하는 신경망과 각 유형별로 실제 문자를 인식하는 신경망으로 구성되어 있으며, 그 전체 구조와 신경망이 각각 (그림 3.1)과 (그림 3.2)에 있다.



(그림 3.1) 시스템의 전체 구성



(그림 3.2) 신경망 구성의 개관

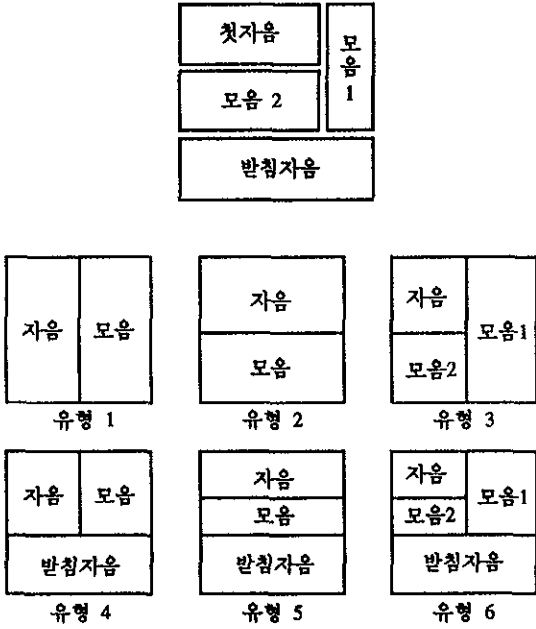
한글은 전체 가능한 클래스의 수가 11,172개에 이르고,

빈번히 상용되는 것만도 2,300여개에 이르는 복잡한 패턴분류 문제이기 때문에 영어나 숫자의 인식을 위해 사용되었던 단순한 신경망 기법을 그대로 이용하기는 어렵다. 따라서, 한글처럼 인식하여야 할 문자의 수가 많은 경우에는 단순한 신경망 방법의 적용으로는 학습이 불가능하므로 자소단위로 인식하는 것이 바람직 하다. 또, 실제 입력영상을 가정할 때 입력노드의 수가 커지는 것을 방지하기 위해 전처리 등을 통한 특징 추출을 하여 신경망의 입력으로 사용하는 방법이 있는데, 이것은 기본적으로 신경망의 장점을 최대한 이용하지 못하는 것이 된다. 이와 같은 문제를 해결하기 위하여 인식하여야 할 문자의 클래스를 몇개의 부류로 나누는 유형분류 신경망과 각 부류에 속하는 특정 문자를 인식하는 인식 신경망으로 구성된 통합 신경망(Integrated Neural Network)[15, 22, 25]의 형태로 전체 시스템을 설계하였다. 즉, 입력된 문자는 유형분류 신경망을 통하여 해당 유형이 밝혀지고, 각 유형에 속하는 문자로 학습된 해당 인식 신경망에 의하여 어떤 문자인지 결정된다.

이와 같은 계층적 구조의 신경망은 기본적으로 순수한 Backpropagation 신경망에 비하여 다음과 같은 장점을 갖는다. 첫째, 전체를 한꺼번에 학습시키는 기존의 방법에 비하여 학습에 소요되는 시간이 줄어들며, 둘째, 적은 학습 패턴으로도 높은 인식을 얻을 수 있고, 셋째, 새로운 패턴을 학습시킬 필요가 있을 때 간단히 새로운 인식 신경망을 추가하도록 하는 신경망의 구조화를 이룰 수 있다. 단지 각 부류로 나누는 기준과 방법이 다소 모호하고, 이와 같이 인식에 순서를 부여함으로써 하여 부가적인 제어 전략이 필요하다는 단점이 있다.

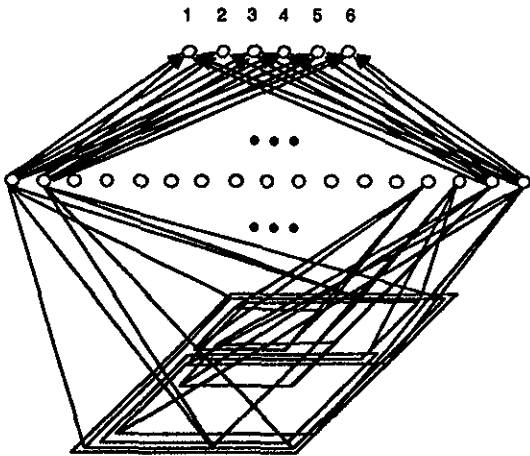
3.1.1 유형분류 신경망

전체 문자를 나누는 방법으로는 문자의 구조적 특성을 분석한 사전 지식(a priori knowledge)을 사용하는 것과 전체 문자를 일정시간 학습시킨 후 중간층의 활성화 패턴(activation pattern)을 이용하는 것이 있을 수 있다[15]. 그런데, 한글은 기본자소가 (그림 3.3)과 같이 일정한 위치에 존재하며, 그 모음의 형태와 중성의 존재 유무에 따라서 6가지의 유형으로 분류되는 구조적 특성이 있다[10]. 따라서 입력된 문자 영상을 이와 같이 6가지의 유형으로 분류할 수 있다면, 하나의 신경망



(그림 3.3) 한글의 일반적인 형태와 유형

이 인식해야 할 클래스의 수를 줄일 수 있을 뿐만 아니라, 각 유형에서 한 음절이 해석될 수 있는 가능성을 제한하게 되어 각 문자의 인식에도 큰 도움을 줄 수 있다



(그림 3.4) 유형분류 신경망

문자의 유형분류를 위한 신경망의 구조는 (그림 3.4)와 같이 1,600개의 입력 노드를 갖는 입력층과 17개의 노드를 갖는 하나의 중간층, 그리고 6개의 노드를 갖는 출력층으로 이루어져 있다[7] 학습은 Backpropaga-

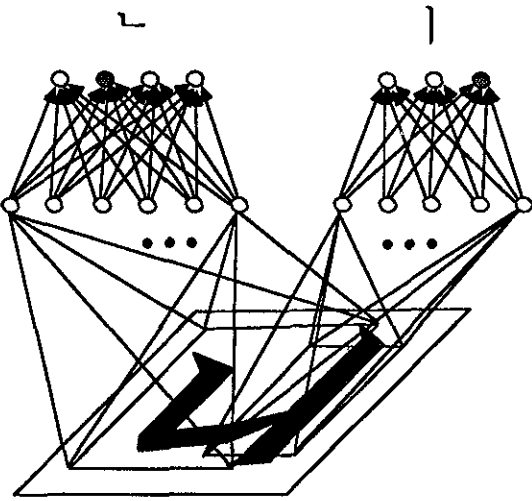
tion에 의하여 이루어 졌으며, 전체 오차가 어떤 임계치 보다 작아질때 까지 학습시켰다 이때, 중간층의 각 노드를 입력층의 모든 노드와 완전히 연결시키지 않고 각 노드가 (그림 3.3)의 각 유형의 특정 자소 부분만을 입력으로 받도록 연결시켰다. 이와 같은 구조는 각각의 소영역이 문자의 유형을 분류하는데 특정한 역할을 한다는 특성에서 나온 것으로 전체 연결의 수가 줄어들어 학습시간이 빨라질 뿐만 아니라, 각 노드의 기능을 외부적으로 부가하여 좀더 효율적인 분류 신경망이 가능하다 즉, 중간층의 각 노드가 나름대로의 국부적 결정(local decision)을 하고, 이를 토대로 출력층에서 전체 결정을 내리는 구조에서 중간층의 노드가 필요이상의 정보를 입력 받으면 오히려 혼동을 일으킬 수 있다는 점에서 특정 영역과 연결시키는 것을 정당화 할 수 있다 실제로 입력층과 출력층이 모두 연결된 구조와 비교 실험한 결과, 문제에 따른 정보를 반영하는 이와 같은 구조가 학습속도와 학습의 안정성면에서 훨씬 우수함을 알 수 있다[7] 각 노드를 모두 연결 시킨다면 27,302(40×40×17+17×6)개의 연결선이 필요한 반면, 이 신경망은 5,915개로 약 78.33% 정도 줄어들었다

3.1.2 유형별 문자 인식 신경망

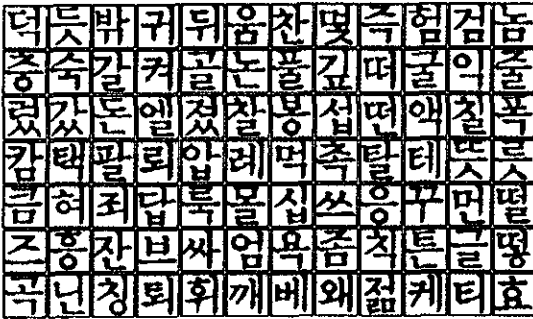
앞 절에서 설명한 유형분류 신경망에 의하여 6가지의 유형으로 나뉘어진 각 문자는 해당하는 특정 신경망으로 학습되고 인식된다 이제 전체 문자를 몇개의 클래스로 나누어 한 신경망이 인식하여야 할 문자의 클래스는 줄어들었지만, 각 신경망이 인식하여야 할 문자가 유사한 형태이기 때문에 분류하기 어렵게 되었다 그런데, 동일한 유형의 한글 문자에서는 초성, 중성, 종성 등의 각 자소가 존재하는 위치가 거의 일정하므로 각 자소별로 인식하도록 하며, 그 구조는 (그림 3.5)와 같다.

한글은 2차원적 위상구조로 각 자소가 존재하는 위치가 어느정도 일정하기는 하지만 실제로 입력된 문자영상에서는 각 자소가 접촉되어 자소별의 인식을 어렵게 하는데, 이러한 문제 때문에 기존의 방법에서는 별도의 자소를 분리하는 알고리즘을 사용하거나, 세신화 등의 전처리를 통한 구문 분석 기법을 사용하기도 하는데 그 자체가 매우 어려운 문제이다 따라서 기본자소가 있을 수 있는 위치에 적당히 중첩되는 감각 영역을 설정하고, Backpropagation 신경망을 통하여 인접 자소간의 접촉

에 의한 잡음이나 스캐너의 기계적인 특성에 의한 변형을 흡수 처리하도록 한다 (그림 3.6)에 스캐너에 의해 입력된 문자의 실제 예가 있다



(그림 3.5) 유형별 문자인식 신경망(유형 1의 경우)



(그림 3.6) 입력문자의 실제예

또, 각 유형의 문자에서 기본 자소의 위치는 대개 일정하지만, 같은 유형에 속하는 문자라도 자음이나 모음의 형태에 따라서 자소의 크기나 위치가 크게 다르다 하지만, 신경망은 단순히 결합 극부능력을 가진 원형정합 방법이 아니라 자체의 일반화 능력을 갖고 있기 때문에 이러한 문제를 해결할 수 있다 이것은 신경망을 이용한 부패된 인식에서 크기나 위치가 다르다는 이유 때문에 각기 다른 신경망을 사용하는 방법(1, 3)에 비하여 전체 시스템을 간략히 할 수 있는 장점이 있다

3.2 변형을 첨가한 학습방법

신경망을 소개하는 많은 문헌에 의하면 신경망은 변

형이 심한 입력에 대해서도 그 자체의 일반화 능력에 의하여 어느정도 올바른 인식이 가능하다고 한다(21, 24) 하지만, 실제 문제에 있어서는 변형된 입력을 잘 처리하지 못하는 경우가 있기 때문에 신경망이 이와 같은 변형을 처리하도록 하는 연구를 많이 하고 있다. 본 절에서는 이러한 연구를 정리하고, 현실적인 문제중에서 변형의 형태를 학습중에 어느정도 알 수 있는 경우에 효과적으로 인식률을 증진시킬 수 있는 학습 방법을 제안한다

신경망이 입력의 변형을 처리할 수 있도록 하기 위하여 Fukushima[17]는 Neocognitron 모델에서 전체 노드의 구성을 단순 노드(simple cell)와 복합 노드(complex cell)로 나누어 단순 노드에서는 가능한 모든 특징을 추출하고, 복합 노드에서 적당히 융통성있는 윈도우를 두어 그 범위내에서 공간적으로 이동된 특징도 발견할 수 있도록 하는 방법을 사용하여 모델 자체가 어느정도 변형된 입력도 처리하도록 한다. 그러나 이것은 생물학적으로는 바람직할지 몰라도 실제로 사용하기에는 너무 많은 저장장소를 필요로 하는 등 많은 문제를 갖고 있다

변형된 입력을 처리하는 공학적인 방법으로는 주어진 입력에서 수학적으로 변형에 무관한 특징을 추출하여 신경망의 입력으로 사용하는 것으로 Reber[23]나 Khotanzad[20] 등의 연구가 있다 이들은 주로 polar, log, discrete Fourier 변환 등을 통해 구해진 특징이나 기하학적인 특징(geometrical feature)을 신경망의 입력으로 사용하는데, 그 자체가 큰 부담일 뿐만 아니라 전체 시스템의 인식률이 전처리 과정에 크게 의존하는 단점이 있다 Widrow[26]는 불변망(invariance net)이라는 신경망을 전위(front-end)에 사용함으로써 주어진 입력을 미리 정해진 변형판에 의하여 분산시킨 후, 분류 신경망에서 다시 모아서 인식함으로 입력의 변형을 처리하고자 하였고, Waibel[25]은 시간지연신경망(Time Delay Neural Network)을 사용하였다 이것은 원래 음성인식 분야에서 각 음소의 시간적 변형을 처리하기 위한 것으로 시간에 따라서 변하는 입력에 대하여 좋은 성능을 보이고 있다

실제 스캐너를 통하여 입력된 문자는 같은 문서의 글자라도 스캐너의 작동상태 등 기계적 여건에 따라서 다르게 나타난다. 이러한 이유에서 제한된 학습 패턴으로는 현실적인 복잡한 문제에 대하여 적절한 일반화를 이

루기 어렵기 때문에 좋은 성능을 기대할 수 없다 즉, 학습하여야 할 연결선이 많은 경우에는 주어진 학습 데이터를 인식할 수 있도록 조정된 신경망이 테스트 데이터는 잘 인식하지 못하는 수가 있다 그런데, 스캐너를 통해 입력된 문자를 인식하기 위한 신경망의 일반화 능력 실험에 의하면 인식률을 저하시키는 주된 요인은 전체 문자가 상하좌우로 이동된 것임을 알 수 있다[11] 즉, 순수한 학습 패턴에 의하여 훈련된 신경망은 입력 문자가 잠깐 더 굵거나 가늘게 되는 경우나 임의의 잡음이 첨가되는 것은 어느정도 자체의 일반화 능력에 의하여 해결할 수 있지만, 전체 문자의 이동은 잘 흡수하지 못하는 경향이 있다

예상되는 입력의 변형이 비교적 크지 않은 문제에서는 신경망의 구조 자체를 복잡하게 하는 것보다는 입력에 가능한 변형이 포함된 데이터를 학습에 사용하는 것이 보다 효과적이다[15, 22] 하지만, 실제 문제에서는 학습 데이터를 획득하는 것 자체가 많은 경비와 노력을 요하므로 주어진 학습 데이터만을 이용하여 적당하게 학습패턴의 수를 늘릴 필요가 있다. 따라서, 위치나 크기의 변형이 심한 실제 문제를 신경망으로 해결하기 위한 한 방법으로 학습 데이터에 체계적으로 의미있는 변형을 첨가하여 전체 학습 패턴의 수를 확장하여 학습하는 방법을 제안한다. 가능한 입력의 변형으로 일반화하기 어려운 상하좌우로 이동시킨 새로운 패턴을 학습에 참여시킴으로써 부족한 학습 패턴으로 인한 과특수화(overspecialization) 문제를 해결하고, 이를 통하여 예상되는 변형된 입력 문자를 인식할 수 있도록 하는 응용성을 신경망에 부여한다

주어진 학습 패턴에 일종의 잡음을 첨가하는 것은 학습 패턴의 특이성을 어느정도 희석하는 효과를 냄으로써 학습 패턴의 일반적인 특성을 보다 잘 추출할 수 있도록 하여 결과적으로 인식률을 높이게 될 것이다 그런데, 변형을 첨가한 학습은 변형된 입력을 처리할 수 있도록 하기는 하지만 학습해야 할 입력관계의 수가 많아지게 되어 신경망을 안정화하는데 오랜시간이 소모되는 단점이 있기도 하므로 적당한 변형의 정도를 결정하는 것이 매우 중요하다.

IV. 실험 및 결과분석

4.1 실험 환경

실험에 사용한 시스템은 신경망의 학습속도를 감안하여 Cray2s 슈퍼 컴퓨터상에서 C 언어로 구현 하였다. 실험에 사용한 한글 문자의 집합은 한글 기계화 연구소에서 발표한 한글 문자의 짓기 순서[14]에 의해 상위 1,500자의 문자중에서 일부 문자를 제외하고 외국어 표기를 위한 문자를 추가하여 990자로 하였다[8]. 이 990자에 대한 누적 사용 빈도율은 일반적인 상용문자에 대하여 99.8%이다[14] 한글 문자의 크기는 Qnix 레이저 범 프린터의 명조체 H4 글자체를 사용하였으며, 실험은 인치당 300회소의 해상도를 갖는 광학 문서 입력 장치인 Microtek MSF300C 스캐너로 입력받은 영상을 사용하였다 실험에 사용한 문자는 모두 6쌍으로 각기 다른 문서에서 입력 받았으며, 그중에서 1쌍은 학습에, 나머지 5쌍은 인식실험에 사용하였다

실험에 사용한 한글인식 신경망은 중간층에 40개의 노드를 두고, 학습률과 관성항은 각각 0.1과 0.9를 사용한 Backpropagation 알고리즘으로 학습하였다 이때, 연결강도와 바이이스의 초기치는 -0.1과 0.1 사이의 임의의 값을 사용하였으며, 전체 학습패턴의 오차값이 특정한 임계치보다 작아질때 까지 계속 학습하였다

4.2 실험 결과의 분석

<표 4.1> 단순한 학습방법에 의한 인식률

	Group Identification			Character Recognition In each Group		Overall	
	Group	#char	#error	% correct	#error	% correct	#error
1	114	0	100.00	0	100.00	0	100.00
2	68	3	95.59	2	97.06	5	92.65
3	42	0	100.00	6	85.71	6	85.71
4	460	1	99.79	16	96.67	16	96.67
5	251	0	100.00	15	94.02	15	94.02
6	35	3	91.43	0	100.00	3	91.43
Overall	990	7	99.29	39	96.06	45	95.45

<표 4.1>은 III장에서 설명한 인체체 한글문자 인식 신경망의 인식률을 보여 준다 이 표에 의하면 실험에 사용된 990자의 문자가 유형별로 고르게 분포되지 않고 유형 4와 유형 5에 집중되어 각 유형별 신경망의 인식률이 큰 차이를 보이고 있기는 하지만, 전체 인식률이 95.45%로 경쟁하는 신경망 구조[1]를 사용한 경우의 83.6% 보다는 상당히 좋은 성능을 보인다. 따라서, 본

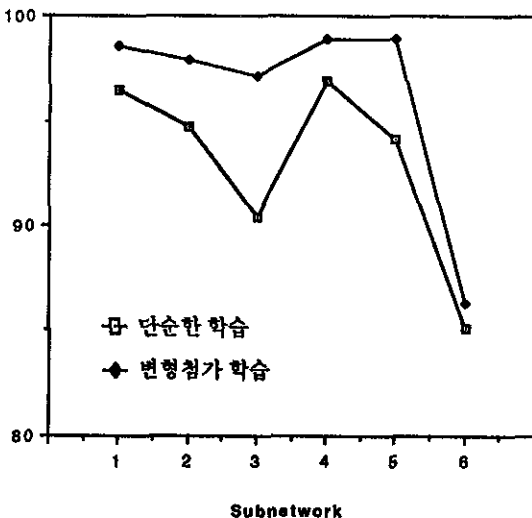
논문에서 제안한 신경망이 한글 인식을 위해 보다 적절한 구조임을 알 수 있다.

이 신경망을 학습시키기 위하여 사용한 990자의 학습 패턴을 상하좌우로 이동시킨 패턴, 즉 각각의 문자영상마다 위로, 아래로, 왼쪽으로, 오른쪽으로 각각 한 화소씩 이동시켜서 만든 5쌍의 990자를 학습에 사용하는 방법의 결과가 <표 4 2>에 있다. 변형을 첨가하여 학습 패턴의 수를 늘림으로 98.28%의 인식률을 얻었는데 이것은 단순한 학습 방법을 사용하였을때 오인식된 문자

<표 4 2> 변형을 첨가한 학습방법에 의한 인식률

Group	Group Identification			Character Recognition in each Group			Overall	
	#char	#error	% correct	#error	% correct	#error	% correct	
1	114	0	100.00	0	100.00	0	100.00	
2	68	3	95.59	0	100.00	3	95.59	
3	42	0	100.00	2	95.24	2	95.24	
4	480	1	99.79	7	98.54	7	98.54	
5	251	0	100.00	2	99.20	2	99.20	
6	35	3	91.43	0	100.00	3	91.43	
Overall	990	7	99.29	11	98.89	17	98.28	

의 거의 72% 정도를 제대로 인식한 것이다 (그림 4 1)은 각기 독립적으로 입력받은 다섯개의 테스트 문자의 집합에 대한 각 방법의 유형별 평균 인식률을 보여 준다. 이 그림에 의하면 유형 6의 오차가 전체 인식률을



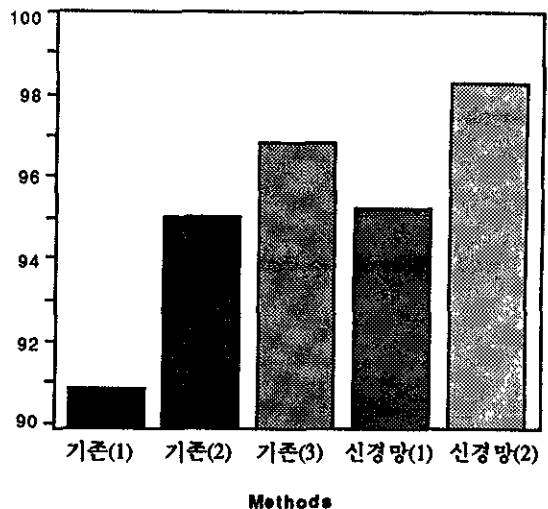
(그림 4 1) 유형별 인식률의 비교

저하시키며, 유형 6 오차의 대부분이 유형분류 부분에서 발생함을 볼 수 있다 따라서, 신경망 방법의 성능을 좀 더 향상시키기 위해서는 유형분류 신경망에서 유형 6을 나누는 부분에 특별한 처리를 하여야 할 것이다

신경망 기법을 사용한 문자인식 방법을 기존의 문자 인식 방법과 비교하기 위하여 기존의 통계적 방법으로 많이 사용하는 Nearest Neighbor 분류기법(16)을 효율적으로 구현한 트리 Classifier를 사용하였다(8). 다섯가지 테스트 문자의 집합에 대한 이들 방법의 인식률과 평균 인식률이 <표 4 3>과 <그림 4.2>에 제시되어 있다(9) 이때, 기존의 방법은 트리 Classifier 방법을 의미하는데, 기존(1)은 학습시에 하나의 패턴집합을 사용한 경우이고, 기존(2)는 5개의 패턴집합을, 그리고 기존(3)은 10개의 패턴집합을 사용한 경우이다 신경망 방법에서도 변형을 첨가하지 않고 하나의 패턴 집합만을 사용한 경우는 신경망(1)로 나타내었고, 학습패턴에

<표 4 3> 통계적 방법과의 비교

실험 화일	기존(1)	기존(2)	기존(3)	신경망(1)	신경망(2)
1	91.62	96.36	97.58	95.45	98.28
2	91.11	95.15	95.78	95.25	97.78
3	89.29	94.04	96.36	95.25	98.59
4	89.80	95.15	98.87	95.76	98.38
5	91.31	98.15	97.17	94.95	98.38
평균	90.63	95.00	97.00	95.33	98.28



(그림 4.2) 각 방법의 평균 인식률

상하좌우로 변형된 패턴을 첨가한 것은 신경망(2)로 나타내었다 이 그림에 의하면 기존의 방법은 학습 패턴을 늘려나감에 따라 어느정도 인식률이 증가되어 97.0%까지 이를 수 있지만, 변형을 첨가하여 학습시킨 신경망의 인식률 98.28%와 비교하면 다소 떨어지는 것을 볼 수 있다 여기서 주목할 사실은 단순한 신경망을 사용한 신경망(1)의 방법이 같은 조건하에서의 기존의 방법인 기존(1)에 비하여 그 인식률이 월등하다는 것이다 또, 이것은 간단한 패턴분류 문제에서 얻은 결과인, 다층구조 신경망이 일반적인 Nearest Neighbor 분류기법이나 최적의 Bayesian 기법보다 결정능력면에서 좀 더 우수할 수 있다[19]는 것을 실제 문제에서 입증한 것이다

<표 4 4> 각 유형별 자소의 혼동표

	1	2	3	4	5	6	
첫자음	ㄷ → ㅌ(1)		ㄱ → ㅋ(2)	ㄴ → ㄹ(2) ㄴ → ㄷ(1) ㄴ → ㄱ(1) ㄴ → ㅍ(1) ㄴ → ㅂ(2) ㄴ → ㅅ(1) ㄴ → ㅁ(2) ㄴ → ㅆ(2)	ㅎ → ㅈ(1)		
모음	ㅣ → ㅑ(1)		ㅣ → ㅓ(3)	ㅣ → ㅕ(1) ㅣ → ㅗ(2) ㅣ → ㅛ(7) ㅣ → ㅜ(1)	ㅡ → ㅑ(1) ㅡ → ㅓ(1) ㅡ → ㅕ(1) ㅡ → ㅗ(6) ㅡ → ㅛ(2)		
받침				ㄷ → ㅌ(1) ㅎ → ㅇ(1) ㅁ → ㅅ(1)	ㅌ → ㄷ(1) ㅓ → ㅗ(1)		

<표 4 4>는 신경망 방법을 사용하였을때 발생하는 오인식을 분석하여 각 유형별로 자소의 혼동표를 제시한 것이다[12] 이 표에 의하면 신경망에 의한 오인식 혼동이 심리학적인 연구[2]와 비교하여 어느정도 합당하다고 볼 수 있는데, 그 혼동의 원인을 분석해 보면 다음과 같다. 첫째는 자소자체의 유사성에 의한 혼동으로 'ㄷ'과 'ㅌ', 또는 'ㅌ'과 'ㄷ'의 혼동을 들 수 있다 둘째는 자소간 접촉에 의한 혼동으로 인접자소가 접촉된 문자를 자소별로 인식하고자 하여 발생한 오인식으로 'ㄴ'나 'ㄹ'와 같이 'ㄴ'이나 'ㄹ'의 마지막 획이 오른쪽으로 길게 붙어서 모음 부분과 접촉하여 'ㅣ'를 'ㄱ'로 잘못 인식하는 경우이다 셋째는 스캐너의 변형

에 의한 혼동으로 문자를 인치당 300화소로 입력하여 발생하는 것인데 첫자음이나 받침에 따라서 'ㅡ'를 'ㄱ'나 'ㅌ'로 오인식하는 경우였다.

<표 4 5> 2,350자에 대한 인식률

	Group Identification			Character Recognition In each Group		Overall	
	Group	#char	#error	% correct	#error	% correct	#error
1	149	4	97.32	4	97.32	8	94.63
2	90	2	97.78	5	94.44	7	92.22
3	109	14	87.16	48	55.96	52	52.29
4	1069	4	99.63	104	90.27	106	90.08
5	586	3	99.49	46	92.15	47	91.98
6	347	164	52.74	211	39.19	256	26.22
Overall	2350	191	91.87	418	82.21	476	79.74

<표 4 5>는 990자로 학습시킨 신경망을 한글 완성형 코드의 전체 문자인 2,350자에 대해 실험한 결과로 그 인식률이 79.74%까지 이를 수 있음을 볼 수 있다 이것은 신경망의 일반화 능력을 보여 주는 것으로 학습시키지 않은 문자도 어느정도 유추하여 인식할 수 있음을 보여준다 그런데 여기서 대부분의 오인식이 유형 3과 유형 6에서 발생하였는데, 그것은 학습에 사용한 990자 중에서 그 유형의 문자가 너무 적은 수였기 때문에 기인한 것으로 학습시에 기본 990자 이외에 유형 3과 유형 6의 문자를 좀더 추가하여 학습시키면 1,000여자의 문자만으로 학습시켜도 2,350자의 문자에 대해서 좋은 성능의 인식률을 얻을 수 있을 것이다

V. 결 론

본 논문에서는 최근에 각광 받고 있는 신경망을 이용한 실용적인 인쇄체 한글문자 인식 시스템, 망눈을 설계하고 구현하였다. 신경망은 학습기능에 의하여 주어진 입력의 연관 관계를 자동으로 생성해 낼 뿐만 아니라, 자체의 일반화 특성과 결함극복 능력으로 기존의 계산방식으로는 풀기 어려운 문제를 융통성있게 해결할 수 있는 방법으로 대두되고 있다. 하지만, 실제 문제를 해결하는데 간단한 문제에서 사용하는 것과는 같은 단순한 신경망의 적용으로는 학습시간이 너무 많이 소요될 뿐만 아니라 그 성능도 크게 떨어지는 단점이 있다 망눈 시스템의 인식률은 990자의 한글문자에 대해 단

순한 학습방법을 사용하였을때 95.45% 정도이지만, 상하좌우로 변형을 첨가하여 학습시킨 경우 전체 인식률을 98.28%까지 높일 수 있었는데, 이것은 기존의 문자 인식 시스템과 비교해 볼때 실용적인 시스템으로서의 그 가능성을 보여 주는 것이라 하겠다 또, 오인식 혼동의 분석에 의하면 이 시스템이 인간과 유사한 특성을 나타내는 것을 볼 수 있으며, 2,350자에 대한 인식 실험에 의하면 990자만으로 학습시킨 시스템이 약 80%의 인식률을 보였는데, 이로 미루어 신경망의 일반화 능력을 알 수 있었다.

앞으로 좀 더 연구하여야 할 내용은 여기서 구축한 한글문자 인식 신경망 시스템을 보다 발전시키는 연구로서 다중 글자체나 다중 크기의 글자를 인식할 수 있도록 확장하는 것이 필요하며, 보다 높은 인식률을 얻기 위해서는 기존의 기호처리 중심의 방법과 상호 보완한 통합 방법이 필수적이다. 또한 실용화를 위해서는 한글 완성형 코드의 전체 글자인 2,350자를 학습시켜야 할 것이다

감사의 글

본 논문의 연구를 위하여 Cray2s 슈퍼 컴퓨터를 사용하도록 해 주신 시스템 공학센터, 오원근 박사님의 격려에 깊이 감사드립니다

참 고 문 헌

- 1 고병기, 김진형, 양현승, “경쟁하는 신경망 조직을 이용한 인쇄체 한글 문자의 인식,” 세 1 회 영상처리 및 이해 워크샵 논문집, 포항, 1989년 1월
2. 김민식, 정찬섭, “한글의 자모 구성형태에 따른 자모 및 글자 인식,” 인지과학, Vol. 1, pp. 25-75, 1989.
3. 김병기, 김항준, “신경망 모델을 이용한 한글 문자의 형태분류와 인식,” 한국정보과학회 추계 논문발표집, 서울, pp. 303-306, 1989.
- 4 김상우, 전운호, 최중호, “신경 회로망을 이용한 한글문자의 인식,” 대한 전기학회 하계 학술대회 논문집, pp. 166-169, 1989.
5. 김태천, 김은진, 이일병, “한글 필기체 자모 인식을 위한 신경망 모형,” 한국인공지능연구회 춘계 학술발표 논문집, 서울, pp. 25-39, 1989.

- 6 이경희, 이원돈, “Coulomb Energy Network를 이용한 한글 인식 Neural Network,” 한글날 기념 학술발표 논문집, 서울, pp. 267-271, 1989년 10월.
- 7 이동현, 조원규, 양현승, 김진형, “신경망 모델을 사용한 한글 문자의 형태분류,” 한국정보과학회 춘계 학술발표 논문집, 부산, pp. 215-218, 1989년 4월.
- 8 이승호, 조창재, 김일영, 조성배, 양현승, 김진형, “한글 문서 인식 시스템 SILNOON의 개발,” 한국정보과학회 춘계 학술발표 논문집, 부산, pp. 211-214, 1989년 4월
9. 이승호, 조성배, 김진형, “문자인식에 있어서 기존의 방법과 신경망 방법의 비교,” 신경망 응용 Workshop, 서울, 1989.
- 10 이주근, “한글 문자의 인식에 관한 연구(IV),” 전자공학회지, 제 9 권 제 4 호, pp. 25-32, 1972.
- 11 조성배, “현실적인 문제를 해결하기 위한 신경망 방식과 한글문자 인식에서의 일반화 능력,” 한국과학기술원 전산학과 인공지능연구실 AIL-TR-89-2, 1989
- 12 조성배, 김진형, “인쇄체 한글문자 인식을 위한 계층적 신경망,” 한국인지과학회 연차 학술대회 논문집, 서울, 1989
13. 정상현, “역전과 학습 신경망을 이용한 한글 자소 인식에 관한 연구,” 석사학위논문, 숭실대학교, 1988
- 14 한글 기계화 연구소 : 한글 기계화 연구, 1975
15. R K. Bernhard and A.K. Wolfgang, “Design of Hierarchical Perceptron Structures and their Application to the Task of Isolated-word Recognition,” in Proceedings of the IEEE IJCNN, Vol. I, pp.243-249, 1989.
16. R.O. Duda and P.E. Hart, Pattern Classification and Scene Analysis, New York, Wiley Press, 1973.
17. K. Fukushima, “A Neural Network for Visual Pattern Recognition,” IEEE Computer, Vol. 21, No. 3, March 1988.
18. R. Hecht-Nielsen, “Neurocomputer Applications,” NATO ASI Series F. Computer and Systems Sciences, Vol. 41, pp.445-453, 1988.
19. W.Y. Huang and R P Lippmann, “Compari-

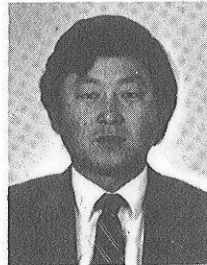
- sons between Neural Net and Conventional Classifiers,” in Proceedings of the IEEE ICNN, Vol. IV, pp.485-593, 1987.
20. A. Khotanzad and J.H. Lu, “Distortion Invariant Character Recognition by a Multi-Layer Perceptron and Backpropagation Learning,” in Proceedings of the IEEE ICNN, Vol. I, pp.625-632, 1988.
 21. R.O. Lippmann, “An Introduction to Computing with Neural Nets,” IEEE ASSP Magazine, pp.4-22, April 1987.
 22. T. Matsuoka, H. Hamada and R. Nakatsu, “Syllable Recognition Using Integrated Neural Networks,” in Proceedings of the IEEE IJCNN, Vol. I, pp.251-258, 1989.
 23. W.L. Reber, “An Artificial Neural System Design for Rotation and Scale Invariant Pattern Recognition,” in Proceedings of the IEEE ICNN, Vol. IV, pp.277-283, 1987.
 24. D.E. Rumelhart, G.E. Hinton and R.J. Williams, “Learning Internal Representations by Error Propagation,” in Parallel Distributed Processing, Vol. I, pp.318-362, Cambridge, MA:MIT Press, 1986.
 25. A. Waibel, “Connectionist Glue: Modular Design of Neural Speech Systems,” in Proceedings of the 1988 Connectionist Models Summer School, pp.417-425, 1988.
 26. B. Widrow and R.G. Winter, “Neural Nets for Adaptive Filtering and Adaptive Pattern Recognition,” IEEE Computer, pp.25-39, March 1988.



조 성 배

1988년 연세대학교 이과대학
전산학과 졸업.
1990년 한국과학기술원 전산
학과 석사학위 취득.
1990년 현재 한국과학 기술원
전산학과 박사과정 재학중.
주 관심분야는 신경망, 패턴

인식, 지능형 Man-Machine Interface 등임.



김 진 형

1971년 서울대학교 공과대학
졸업.
1979년 UCLA 전산학 석사학
위 취득.
1983년 UCLA 전산학 박사학
위 취득.
1973년~1976년 KIST 전산실

연구원.

1976년~1977년 미국 California 도로국 연구원.
1981년~1985년 미국 Hughes 인공지능센터 선임연구원.
1986년~1988년 본 학회 산하 인공지능 연구회 위원장.
1985년~현재 한국과학기술원 전산학과 부교수.
현재 국제로봇자동차시스템 논문지의 편집자.
주 관심분야는 컴퓨터시각, 전문가시스템, 지능형 Man-
Machine Interface 등임.