

역전파와 스펙트럼 분석을 이용한 이차원 영상인식

차 승 윤[†] · 김 성 락^{††}

요 약

본 논문은 스펙트럼 분석기와 신경회로망인 역전파를 이용하여 영상인식에 대한 방법을 제안하였다. 2차원 영상으로부터 추출된 외곽선 정보를 원점과 모든 경계점 사이에서 유클리드 거리로 나타내었고 영상의 전체 형태와 관련되어 있다. 고속 푸리에변환과 스펙트럼 분석기는 특징 추출과 데이터 감소의 수단으로써 사용되었으며, 3차원 특징벡터는 에너지 스펙트럼으로부터 스펙트럼 분석기에 의해서 추출되며, 이 특징벡터는 형변환, 회전 그리고 크기 변환에 무관하였다. 역전파 신경회로망은 이 스펙트럼 특징벡터와 함께 학습되고 실험된 한 개의 은닉층 모듈로 결합되었다. 본 논문에 포함된 실험은 4개의 서로 다른 형태의 자동차 영상을 대상으로 하였으며, 실험에서 잡음이 첨가된 영상에 대해서도 제시된 방법의 성능이 우수함을 보여 주고 있다.

Two Dimensional Pattern Recognition Using Spectrum Analysis And Backpropagation

Seungyoon Cha [†] · Sungnak Kim ^{††}

ABSTRACT

This paper proposes a method for pattern recognition using spectrum analyzer and Backpropagation. Contour sequences obtained from 2-D images represent the Euclidean distance between the centroid and all boundary pixels of the shape, and are related to the overall shape of images. The Fast Fourier transform of contour sequence and spectrum analyzer are used as a means of feature selection and data reduction. The three dimensional spectral feature vectors are extracted by spectrum analyzer from the energy spectrum. These Spectral feature vectors are invariant to shape translation, rotation, and scale transformations. The Backpropagation neural network which is combined with one hidden layer module is trained and tested with these spectral feature vectors. The experiments including 4 automobiles recognition process are presented to illustrate the high performance of this proposed method in the recognition problems of noisy shapes.

1. 서 론

정보화 시대를 살고 있는 우리들은 방대한 정보의

범람속에서 외부로부터 끊임없이 정보를 받아들이고 이를 처리하며 대처해 가면서 생활을 영위해 간다. 인간은 외부로부터의 정보를 시각, 청각, 후각, 미각, 촉각 등의 감각 기관을 통해 받아들이고 대뇌로 전달하여 처리한다.

그 중에서도 시각을 이용하여 처리되는 정보는 우리들의 주위에 매우 큰 비중을 차지하고 있다. 일반 사무실이나 연구실에서 정보를 문서나 책의 형태로

*이 논문은 관동대학교 연구비지원에 의하여 연구되었음.

† 종신회원:영동전문대학 전자계산과

†† 정 회 원:관동대학교 전자계산공학과

논문접수:1997년 10월 21일, 심사완료:1998년 1월 31일

처리하고 보존하며, 은행에서는 시각을 통해 화폐, 전표, 수표 등을 처리한다. 공장에서는 부품의 조립이나 설계 및 생산품에 대한 품질 검사가 대부분 시각을 통해서 이루어진다. 이와 같은 시각 정보를 인간의 시각 감각 기관인 눈을 통해서 처리하지 않고 컴퓨터가 인간처럼 감지하고 처리하게 하는 컴퓨터 시각의 실현은 자동화 시스템에서 매우 중요한 요소라고 할 수 있다[1].

컴퓨터의 시각 정보 처리는 시각 정보로 구성된 패턴을 감지하고 인식하는 패턴 인식 시스템에 의해서 이루어진다. 패턴 인식은 외부로부터 받은 정보를 감각 기관을 통하여 두뇌에서 처리하는 과정으로 패턴(pattern)은 사람이 인식하려는 정보를 물리적으로 표현한 대상으로 정의할 수 있다.

인식(recognition)은 대상물의 존재를 아는 인지(cognition)의 과정에서 한걸음 더 나아가 구체적인 의미나 정보 내용을 판정하는 것을 의미한다.

물체를 볼 때 대상물의 존재를 아는 과정이 인지라고 한다면, 구체적인 물체의 형태를 알아내는 과정은 인식이다[1][2].

패턴에는 우리가 인식 대상으로 하는 정보를 표현하여 전달하기 위하여 인간이 의도적으로 발생시킨 것과, 생물이나 자연현상, 사회현상등과 같이 자연스럽게 나타나는 경우가 있으며, 후자는 다시, 목적하는 정보를 얻기위하여 의도적으로 관측하는 경우와 자연스럽게 관측하는 경우로 나눌 수 있다.

이러한 다양한 패턴에 대하여 사람은 감각기관을 통하여 패턴을 입력받고 뇌에서 이를 인식하여 처리하는 일련의 과정을 거친다.

컴퓨터를 이용한 패턴인식에 있어서도 사람의 패턴인식과 유사한 과정을 거친다. 우선 입력장치를 통하여 데이터를 획득하고 이를 컴퓨터를 이용하여 인식을 하는 과정을 거치게 된다. 이러한 컴퓨터를 통한 인식에는 여러 가지가 있을수 있으나 많이 접근되고 사용되는 분야는 문자인식, 음성인식, 영상인식등이 있다.

이러한 컴퓨터를 이용한 다양한 패턴 인식 응용분야 중 본 논문에서는 영상인식을 대상으로 하였다. 영상 인식은 인식 대상과 인식 조건에 따라 여러 가지 형태가 있다. 일반적으로 인식 시스템은 인식 대상과 인식 주체를 고정시키고 인식 대상의 위치 및

방향을 일정하게 고정시킨다.

그러나 실제의 자연 환경에서는 인식 대상과 인식 주체가 움직이는 경우가 많으며 이 경우에 인식 대상의 영상 패턴은 위치, 크기 및 회전 변형이 나타나므로 이와 같은 변형에 무관한 패턴 인식이 이루어질 필요가 있다.

예를들어 사진기(인식의 주체)의 위치를 고정하고 자동차(인식 대상)가 사진기에 평행하게 이동을 하면 필름에 맺히는 영상은 원위치에서 이동을 하게 된다. 혹은 자동차가 사진기를 향해서 다가올 경우에는 영상은 확대되면 멀어질 경우 영상은 작아지게 된다.

또한 자동차가 회전을 하게 되면 필름에 맺히는 영상이 회전을 하게 된다. 물론 이 모든 경우에 있어서 반대로 인식 대상이 고정되고 인식 주체가 움직여도 똑같은 현상이 일어난다[3].

실제 환경에서는 사물의 위치, 크기 및 회전 변형 즉 기하학적 변형이 빈번하게 일어나며, 기계 시각 검사 장치나 자동 물체 인식기 등과 같은 영상 패턴 인식에서는 이와 같은 기하학적 변형에 무관하게 인식할 수 있는 일반적인 패턴 인식 시스템을 필요로 하고 있다.

이러한 패턴 인식 시스템은 우선 대상영상을 카메라, 스캐너 등을 이용하여 관측데이터를 획득한 후, 획득된 자료는 잡음이나 왜곡같은 인식에 불필요한 정보를 포함하고 있으므로, 컴퓨터에서 인식을 위해 필요한 자료만 얻어내는 전처리(preprocessing)를 하게 된다. 그 다음에 인식에 중요한 실마리가 되는 특징을 추출하고, 이 특징을 비교하여 패턴을 식별하는 단계를 거치게 된다[2].

이러한 형태의 영상에서 기하학적인 변형, 즉 회전, 크기변화, 위치 이동과 잡음을 첨가시에도 인식이 가능하도록 기하적인 변형에 무관한 특징 벡터를 추출하고 이를 스펙트럼 분석기를 통하여 데이터의 양을 줄여서 신경회로망인 역전파 학습 알고리즘의 입력으로 사용하여 학습을 하고 기하학적인 변형과 잡음을 포함한 영상에 대하여 실험을 하였다.

2. 패턴 인식 관련 연구

사물의 기하학적 변형에 무관한 인식을 위해 특징을 추출하는 방법에는 전체적 정보를 이용하여 특징

을 추출하는 방법과 지역적인 특징을 이용하여 특징을 추출하는 방법으로 나눌수 있다.

추출되어진 특징을 이용하여 패턴을 인식하는 과정에서는 신경망이 많이 사용되어진다.

2.1 특징 추출 방법

1) 전체 정보를 이용하는 방법

대상 영상 전체의 정보를 이용하는 방법은 영상 전체 정보로부터 추출한 특징이나 그에 상응하는 표현법에 기반을 둔 특징 추출방법으로서 사람의 눈의 망막구조를 수학적으로 모델링한 복소-대수 사상과 푸리에 기술자, 모멘트 불변자, 자기 회귀 모델, 순환 조화 전개등의 특징을 이용하여 특징을 추출한다.

이 방식은 학습에 필요한 패턴의 수가 적기 때문에 분류 인식기를 설계할 수 있는 장점이 있다. 그러나 작은 변형에 민감하여 심한 잡음이 섞이거나 손상된 패턴에 적용하기 어려운 단점이 있다[8][9].

2) 지역 정보를 이용하는 방법

지역적인 특징을 이용하여 불변 특징벡터를 추출하는 방법은 지역적 패턴으로부터 지역적 특징에 기반을 둔 추출 방법으로서 곡률함수, 호의 길이 s 와 내부각 ϕ 로 구성된 (ϕ, s) , 대상의 중심에서 외곽선까지의 거리 L 과 기준점과의 각도 ϕ 로 구성된 (L, ϕ) , 특이점, 모서리, 구멍(hole)등의 특징을 사용하는 방법이 제시된다.

이는 지역적인 특징을 이용하여 불변 특징벡터를 추출함으로써 부분 패턴으로도 분류인식을 할 수 있다는 점과, 패턴의 잡음이나 변형은 지역적으로 발생함으로 잡음이나 작은 변형에 강한 장점이 있다.

그러나 여러개의 지역적 특징 중에서 대표적인 것을 선정하는 기준이 명확하지 않고 지역적 특징에 기반을 둔 시스템이라 계산량이 많고 긴 시간이 소요되는 단점이 있다[3][4].

3) 신경망

신경망은 인간의 두뇌에 가까운 능력을 가진 정보시스템의 구현을 위하여 인간의 두뇌를 모방한 시스템으로 인간의 두뇌가 neuron이라는 기본단위들의 집합으로 이루어져 이 neuron 들이 상호 병렬 연결되어 이루어지는 정보처리 방식에 착안하여 만들어 졌

다. 즉 생물학적인 인간의 정보 처리 방식을 기계적이고 수학적인 Model로 접근한 것이다. 1943년 McCulloch와 Pitts에 의해 최초의 신경망 모델이 제시된 후 1953년 Perceptron 이 개발되어 많은 가능성을 보여 주었으나 Perceptron의 구조적인 문제점인 비선형 분리에 대한 문제로 인하여 위기에 봉착하였으나 1980년 들어서면서부터 많은 학자들에 의해 새로운 Neuron 모델 및 학습 방법등이 제시되고 이러한 발전에 따라 많은 응용분야에서 중흥기를 맞이하고 있다. 최근에는 일련의 인식분야, 제어분야, 설계분야, 의료분야등 공학 뿐만이 아닌 전분야에서 활발히 그 영역을 넓히고 있다.

이러한 신경회로망의 장점으로는 우선 신경망은 학습을 할 수 있다는 것이다. 즉 외부 환경에 맞게 자기 자신을 변화 시켜서 자기 조직화(Self Organization) 할 수 있다.

각각의 처리요소들이 자기 동작을 하면서 입력정보의 특성에 따라 다른 처리요소와의 연결을 새롭게 하거나 또는 기존의 연결을 제거하면서 학습되어질 수 있는 것이다.

두번째 장점으로는 신경망은 고도의 병렬 분산 처리가 가능하다는 것이다. 신경망은 다수의 처리요소가 동시에 작동하는 병렬처리 시스템이다. 처리요소는 극히 간단한 동작을 하는 인공의 neuron으로써 이 처리요소들이 비동기적 또는 동기적으로 병렬처리되어진다. 그래서 폰 노이만형 컴퓨터가 가지지 못한 장점들 즉 인간만이 할 수 있는 능력에 가깝게 다가가는 일이 가능한 구조이다.

세번째 장점으로는 신경망은 고장을 스스로 진단하고 수정할 수 있다. 신경회로망은 소수의 처리요소가 고장을 일으켜 작동을 하지 못할 경우 이 처리요소의 전체 시스템에 대한 영향력을 감소시키고 다른 처리요소들의 영향력을 증가시켜 보완해줄수 있다[8][9].

이러한 형태를 가지고 있는 신경망은 크게 두가지로 나누어 구분을 할 수가 있다. 그 구분은 학습하는 과정에서의 신경회로망 자체에 내장된 평가 기준에 의해 자발적으로 학습이 이루어지는 자율학습(unsupervised learning)과 입력 신호에 대해 이상적인 출력값이 유도되도록 외부에서 목표값을 부여해 주는 방법인 지도 학습(supervised learning)이 있다.

이러한 학습 방법을 따르는 신경망은 종류에는 여

러 가치가 있으나 오늘날에는 지도학습의 경우에는 MLP Multi Layer Perceptron)와 오류역전파(Error Backpropagation)알고리즘[14]이 가장 많이 사용되고 있다.

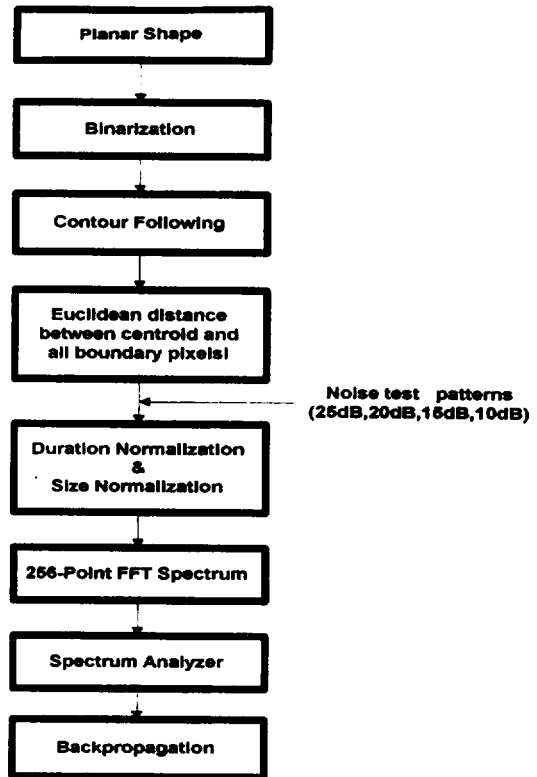
3. 불변 특징 벡터 추출

3.1 불변 특징 벡터

일반적으로 불변 특징벡터란 인식의 대상이 되는 영상을 대표하는 유클리드공간에서 정의되는 값을 말하며, 대상 영상이 기하학적인 변형인 회전(rotation), 크기변화(scaling), 위치이동(translation)되어도 크기가 변화되지 않는 벡터를 말한다[1][3].

본 논문에서는 이 지역적인 특징을 이용하여 불변 특징 벡터를 추출하는 방법에서 외곽선 정보를 이용하여 영상의 기하학적인 변형에 무관한 특징벡터를 추출하기 위해 외곽선 추적, 유클리드 거리, 정규화, 푸리에 변환, 스펙트럼 분석기를 이용 하였으며, 이를 신경회로망인 역전파학습 알고리즘의 입력으로 사용하여 학습을 하고 실험을 하였다.

이러한 패턴 인식 시스템의 전반적이 처리 과정을 (그림 1)에서 보여 주고 있다.



(그림1) Planar shape 패턴 인식 순서
(Fig. 1) Planar shape pattern recognition flow

3.2 이치화

획득되어진 영상에서 바로 인식에 필요한 특징 벡터를 추출할 수는 없다. 스캐너나 영상 캡처장치등을 이용하여 획득되어진 영상은 특정한 포맷 형태의 구조를 가지고 있으므로 이를 컴퓨터에서 입력 받아 처리가 가능한 형태인 이진 텍스트 형태로 변환을 하여야 한다[10].

이는 화상 처리의 기본 조작으로 일반적으로 이치화 처리(binarization)라고 하며 명암 도형의 농도값을 적당한 임계값 1이나 0으로 이치화 시키는 것을 말한다. 이러한 조작에 의해 화상 처리상의 대상 영역과 그 이외의 영역을 명료하게 구별할 수 있고, 영상의 경계선을 추출할 수 있다.

영상과 구분되어지는 대상화상을 이용하여 이의 경계를 구하는 경우에는 화상을 적당한 임계값(threshold)으로 이치화 한 후에 사용한다.

이러한 이치화는 가장 영상인식의 전처리과정에서 가장 기본적인 조작이다. 이 조작의 일반적인 방법은

입력영상을 $F = \{f_{ij}\}$, 출력영상은 $G = \{g_{ij}\}$ 로 하였을 때 파라미터 t 를 이용하여 다음 수식 (3-1)에 의해서 계산 되어진다[2][10].

$$\begin{aligned}
 g_{ij} &= 1 : f_{ij} \geq t \\
 &= 0 : f_{ij} < t
 \end{aligned}
 \tag{3-1}$$

3.3 외곽선 추출

이치화 처리가 된 영상을 가지고 영상의 배경으로부터 물체를 분리해 내는 경계선을 찾음으로서 외곽선 정보는 얻어진다.

외곽선이란 대상물(object)과 배경의 명암의 불연속성이 존재하는 경계면에서의 화소(pixel)들의 집합을 말한다[11].

본 논문의 인식 대상이 경계가 밀폐되어 있는 형태의 영상이므로 추출되어지는 외곽선 정보는 다음의

수식 (3-2)의 형태로 정의 할 수 있다.

$$b(N+i) = b(i) \quad (3-2)$$

대상 영상에서 외곽선을 추출하기 위하여 본 논문에서는 외곽선 추적(contour following) 알고리즘을 이용하였으며[11][12], 이는 대상물 둘레에 있는 임의의 한점을 출발점으로 하여, 본래 명암(grey-level)영상에서 단계적으로 대상물의 외곽선을 찾아가는 방법이다.

외곽선 추적(Contour following)알고리즘은 다음과 같다.

1. 외곽선의 어떤 화소에서 경계선 추적을 시작.
2. 다음 경계 화소가 정해진 영상 영역에 존재하면 left가 존재하지 않고 배경 영역에 존재하면 right
3. 새로운 화소의 위치가 영상 영역이면 시작점의 위치 수정
4. 시작점에 도착할때까지 2~3번 반복.

이러한 외곽선 추적 알고리즘을 이용하여 영상에서 추출된 외곽선 벡터(\vec{c})는 식 (3-3)과 같이 정의할 수 있다.

$$\vec{c} = [c(n)], c(n) = a(n) + jb(n), n = 0, 1, 2, \dots, N(3-3)$$

n: 외곽선 화소의 총 갯수

3.4 무게 중심과의 유클리드 거리

추출되어진 외곽선 정보를 가지고 물체의 무게 중심을 구하고 물체의 무게 중심과 각각의 외곽선 화소(pixel)들간의 유클리드 거리를 구하면 물체의 위치변동에 무관(translation invariant)한 특성을 얻어낸다.

유클리드 거리(Euclidean distance)는[11][12] 대상 영상의 형상이나 위치 관계 등의 계량적인 성질을 정의하는데 많이 사용되는 것으로 이차원상에서 정의된 두점 $x(a, b), y(a1, b1)$ 를 알고 있다면 유클리드 거리는 수식 (3-4)과 같이 정의 할 수 있다.

$$D = \sqrt{(a-a1)^2 + (b-b1)^2} \quad (3-4)$$

영상의 외곽선 화소들에서 무게 중심을 구하는 식은 (3-5), 무게 중심과 각각의 외곽선 화소와간의 유클리드 거리는 식 (3-6)에 의해서 구해진다.

$$c_{g.c} = a_{g.c} + jb_{g.c}, a_{g.c} = \frac{1}{N+1} \sum_{n=0}^N a(n)$$

$$b_{g.c} = \frac{1}{N+1} \sum_{n=0}^N b(n) \quad (3-5)$$

$$\vec{c} = [c'(n)], c'(n) = c(n) - c_{g.c}, n = 0, 1, 2, \dots, N \quad (3-6)$$

이렇게 추출되어진 벡터 \vec{c} 는 이제 위치변동에 무관한 특징벡터로 다음 전처리 과정의 입력값으로 사용되어진다.

3.5 정규화(normalization)

일반적으로 정규화 과정은 패턴이 생성될 때나 생성된 후에 여러 가지 조작에 의해 다양한 변형이 생기는데 이러한 패턴의 변형을 회복하여 기본적인 패턴으로 변환하는 과정을 말한다. 그러나 복잡한 변형을 회복하는 데에 일반적으로 유효한 방법은 없으며, 일반적으로 정규화는 개별적인 인식 방법에 포함되어진다.

정규화의 대상이 되는 것으로는 위치, 크기, 진폭 등이 있으며, 선형 변환의 범위에서 정규화 하는 것이 보통이다[1][3].

크기 변화에 무관한 특징을 얻기 위하여 본 논문에서는 정규화 과정을 다시 두가지의 과정으로 나누어 처리를 하였다.

외곽선 화소의 총갯수는 영상의 종류, 크기에 따라서 다른 값을 가지고 있다. 이러한 외곽선의 화소의 수를 고정된 크기 형태인 256개로 재추출하였다. 이 과정은 수식 (3-7)에서 보여주고 있다.

$$c(k) = b(k \cdot N/256), k = 1, 2, 3, \dots, 256 \quad (3-7)$$

N: 경계화소의 총수

과정이 끝난후 각각의 외곽선 순서 $c(k)$ 들은 유클리드 거리의 합 s 로 나누는 size 정규화 과정을 수식 (3-8)에서 보여 주고 있다.

$$d(k) = c(k)/s, k = 1, 2, 3, \dots, 256 \quad (3-8)$$

s: 유클리드 거리들의 합

이러한 정규화 과정을 마친 후 특징 벡터는 이제 영상의 크기변화에 무관한 특징 벡터로 다음 과정의 입력 값으로 사용되어진다.

3.6 고속 푸리에 변환(fast fourier transform)

영상이 주어진 경우에, 그것이 어떤 공간 주파수 성분으로 구성되어 있는가를 구하는 수단이 되는 것이 푸리에 변환이다. 일반적으로 푸리에 변환은 위치 이동에 대하여 무관한 성질을 가지고 있다.

고속 푸리에 변환(Fast Fourier Transform: FFT)은 [1][12] 다음에서 보여주는 수식 (3-9)에 의해서 구해진다.

$$F(k) = \sum_{n=0}^{255} d(n) \exp(-j2\pi kn/256) \quad (3-9)$$

여기서 $\exp(-j2\pi kn/256)$ 을 일반적으로 회전자라고 하고 일반적으로 W^{nk} 로 나타낸다. 그러므로 W 는 $\exp(-j2\pi/256)$ 로 나타낼 수 있으며 j 는 $\sqrt{-1}$ 의 값을 가지고 있는 복소사상을 나타내고 있으며 W 는 복소 평면 상에서의 단위원을 data 수 $N(256)$ 의로 분할한 $2\pi/N(256)$ 의 각도에서의 복소수값에 대응한다.

이것으로부터 W^n 는 $2\pi \times n/N$ 의 각도에서의 복소수의 값과 일치해서 W 를 곱하는 것은 복소 평면상에서 $-2\pi/N$ 만큼 시계방향으로 회전 시키는 것이 된다.

이 과정에서는 입력값으로 정규화 과정후에 추출된 벡터 $d(k)$ 를 입력으로 사용하여 변환을 시키면 이 변환 되어진 특징벡터는 영상의 회전에 무관한 특징벡터가 된다.

3.7 스펙트럼 분석기

지금 까지 위의 불변 특징 벡터 추출과정에서 추출한 데이터는 영상의 기하학적인 이동에 무관한다. 고속 푸리에 변환에서 추출된 특징은 총 256개이다. 이 256개를 신경회로망의 입력으로 사용하기에는 데이터의 양이 너무 많아 신경회로망 시스템 자체의 성능 저하와 시간 증가가 발생할 수 있다.

여기서 푸리에 변환의 출력값의 정보를 간직하고도 데이터의 양을 줄일수 있는 스펙트럼 분석기를 사용하면 데이터의 양을 줄일 수 있다[11].

스펙트럼 분석기는 Ahmed와 Bates에 의해 제안된 것으로 이 스펙트럼 분석기를 이용하여 데이터의 양을 줄일 때 데이터의 크기는 푸리에 변환의 k 에 의해서 결정되어지고 데이터의 크기는 $\log_2 k + 1$ 이다. 그러므로 본 논문에서의 데이터의 크기는 k 가 256이므로 $\log_2 2^8 + 1$ 이므로 9가 된다.

스펙트럼 분석기의 구성요소 중 첫 번째는 푸리에 변환의 첫 번째 요소를 제공해서 사용하며, 나머지 구성요소들은 수식 (3-10)에 의해서 계산되어 진다.

$$p(0) = F^2(0)$$

$$p(i) = \sum_{j=2^{i-1}}^{k-1} [F^2(j) + F^2(j+1)] \quad (3-10)$$

$i = 1, 2, 3, \dots, n, j = 2^i + 1$ 씩 증가

식 (3-10)에 의하여 나머지 요소들을 구하면 수식 다음과 같이 계산 되어진 값을 구할 수 있다.

$$P(1) = [F^2(1) + F^2(2)] + [F^2(5) + F^2(6)] + [F^2(9) + F^2(10)] + \dots + [F^2(253) + F^2(254)]$$

$$P(2) = [F^2(3) + F^2(4)] + [F^2(11) + F^2(12)] + [F^2(19) + F^2(20)] + \dots + [F^2(251) + F^2(252)]$$

$$P(3) = [F^2(7) + F^2(8)] + [F^2(23) + F^2(24)] + [F^2(39) + F^2(40)] + \dots + [F^2(246) + F^2(247)]$$

$$P(4) = [F^2(15) + F^2(16)] + [F^2(47) + F^2(48)] + [F^2(79) + F^2(80)] + \dots + [F^2(239) + F^2(240)]$$

$$P(5) = [F^2(31) + F^2(32)] + [F^2(95) + F^2(96)] + [F^2(159) + F^2(160)] + \dots + [F^2(223) + F^2(224)]$$

$$P(6) = [F^2(63) + F^2(64)] + [F^2(191) + F^2(192)]$$

$$P(7) = [F^2(127) + F^2(128)]$$

$$P(8) = [F^2(255)]$$

4. 역전파

1986년 Rumelhart에 의해 제안된 역전파 학습 알고리즘은 Widrow-Hoff 학습 규칙을 다층 네트워크와 비선형 미분 가능한 전이 함수로 확장한 것이다.

입력 벡터와 출력 벡터를 사용하여 특정한 출력벡터와 연관되는 입력벡터의 함수로 근사 시키거나, 사용자가 정의한 적절한 입력 벡터로 분류할 수 있을 때까지 네트워크를 학습한다.

시그모이드 뉴런층과 선형 출력층 및 편차를 최소화

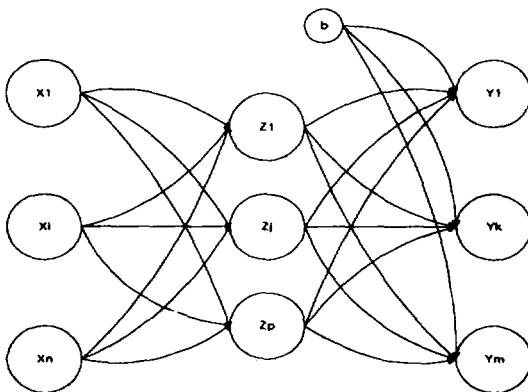
한 한 개 가지는 네트워크는 적당한 어떤 함수로 근사할 수 있다.

네트워크의 평방향 오차를 최소화 하기 위하여 역전파 학습 규칙을 사용하여 네트워크의 연결강도와 편차를 조정한다. 이것은 네트워크의 연결강도와 편차의 값을 오차에 대한 최급하강의 방향으로 계속 변화시킨다. 이를 최급하강의 절차라 한다.

각 연결강도와 편차의 변화는 네트워크의 평방향 오차에서 이들 원소의 영향에 비례한다.

훈련이 된 역전파 네트워크는 전혀 새로운 입력이 들어올 때 타당한 결과를 주고 있다. 일반적으로 새로운 입력이 들어오면 훈련에 사용된 입력 벡터에 대한 정확한 출력 벡터와 유사한 출력을 이끌어 낸다. 역전파 훈련은 전역 최소값(Global minimum)을 얻기 보다 지역 최소값(Local minimum)에 종종 빠지게 되는데 이 지역 최소값으로 만족하는 경우도 있지만 대부분의 경우에는 그렇지 못하다. 그러므로 이때는 더 많은 뉴런과 층을 갖는 네트워크를 설계해야만 좋은 결과를 얻을 수 있으나, 뉴런이나 층을 더 증가시키면 문제가 복잡해진다. 일반적인 역전파 학습 신경망의 구조는 다음의 (그림 2)와 같다[12]. 역전파 학습 신경회로망의 학습은 다음과 같은 규칙에 의해 이루어진다.

1. 신경회로망의 상태를 결정하는 가중치와 임계치를 임의의 수로 초기화



(그림 2) 역전파 학습 신경회로망
(Fig. 2) Backpropagation learning neural network

2. 학습의 반복 회수를 설정한다.
3. 학습 패턴을 설정한다.
4. 은닉층 및 출력층의 노드를 계산한다.
5. 출력층 및 은닉층 노드의 출력 오차를 계산한다.
6. 은닉층과 출력층 노드간의 가중치 및 입력층과 은닉층 노드간의 가중치를 수정한다.
7. 학습 패턴이 종료될 때 까지 3에서 6번까지의 과정을 반복 수행하고, 학습이 종료되면 단계 8로 넘어간다.
8. 모든 출력 노드의 오차가 0에 가까운 값이 되거나, 학습의 반복 횟수가 일정 횟수에 도달할 때까지 3에서 6단계 까지를 반복하고 모든 과정이 끝나면 학습을 종료한다.

5. 실험 및 결과

5.1 실험 시스템 및 실험 데이터

본 논문의 실험에 사용된 영상 입력장치는 평면영상 입력기인 'UMAX'의 칼라 스캐너를 사용하였다. 컴퓨터는 IBM pentium 기종에서 MATLAB, C언어를 사용하였으며 실험 데이터는 (그림 3)에서와 같이 서로 다른 형태의 실제 자동차 모양 4개를 획득하고 이를 전처리과정을 거쳐서 신경회로망인 역전파 신경회로망 입력으로 사용하였다.

본 논문에서는 역전파 학습 알고리즘의 성능을 향상시키기 위하여 모멘트 방법을 사용하였다. 모멘트 방법[12]은 오차 표면에서 역전파의 민감도를 조금 자세하게 감소시키므로써 네트워크가 낮은 오차의 해를 찾는 과정에서 얇은 최소점을 피하는데 도움을 주는 방법으로 오차 표면에서 네트워크가 지역적인 기울기뿐만 아니라 최근의 경향에도 응답할 수 있도록 변경한 알고리즘이다.

획득되어진 각각의 영상은 모두 300×150 pixel로 스캐닝 하였으며 PCX 형태로 입력을 받아 이를 다시 300×150 형태의 이치화된 binary 영상을 만들어 사용하였다.

이치화 영상을 가지고 영상의 기하학적 변화에 무관한 불변 특징 벡터를 추출하여 이를 역전파 신경망의 입력으로 사용하였다.

역전파 신경회로망의 학습 패턴은 각각의 원래 영상 4개와 25dB SNR 잡음을 포함한 패턴을 각각 10개

씩 만들어 총 44개(4+40)로 학습을 하는데 사용하였다.

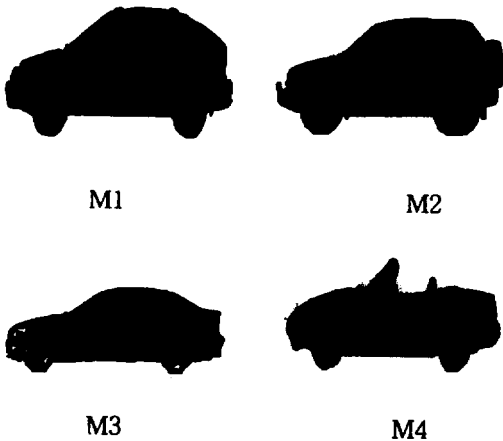
외곽선에 대한 회전, 위치이동, 크기변화 및 잡음(noise)에 대한 식별을 실험하기 위해 우선 잡음을 첨가하지 않은 noise-free 패턴 144개를 만들어 실험을 하였다. 이 noise-free 패턴의 수는 4기종×12각도×3 크기변화로 만들어진 수이다. 각도는 360도를 30도씩 12분할, 크기변형에 대해서는 원형을 100%로 하여, 80%, 60%영상을 만들어 사용하였다.

잡음을 포함한 패턴에 대해서는 식 (5-1)에 4종류의 잡음 25dB, 20dB, 15dB, 10dB를 대입하여 구한 패턴을 각각 1440개를 만들어 실험하였다.

여기서 1440개는 4기종×12각도×3크기변화×10잡음패턴수로 만들어진 수이다. 그러므로 총 테스트 패턴의 수는 5904개 이다.

$$S_n = \frac{S_s}{10^{0.1SNR}} \tag{5-1}$$

여기서 $S_s = \frac{1}{N+1} \sum_{n=0}^N |c'(n)|^2$ 은 signal power이며, $c'(n)$ 은 외곽선 벡터의 n번째 요소이다. SNR은 사전에 정의된 신호대 잡음비 (signal to noise ratio: dB) 이다[11][12].



(그림 3) 실험에 사용한 자동차의 모양
(Fig. 3) The applied car in simulation

각각의 학습 및 실험 패턴은 스펙트럼 분석기를 이용하여 만들어진 8개의 데이터 값을 이용하였다. 그

러나 이 8개의 값중에서 P(0)와 P(1), P(2), P(3)를 빼고는 거의 0에 가까운 값이 계산되어지므로 나머지는 입력값에서 제외하였다.

5.2 학습 및 실험결과

신경회로망인 역전파 학습의 입력으로 사용될 값은 스펙트럼 분석기를 통하여 나온 P(0), P(1), P(2), P(3)데이터를 이용하여 실험을 하였다. 실험에 사용된 잡음데이터는 실제 image자체에 잡음을 첨가하지 않고 실제데이터[P(0)~P(3)] 즉 수치에다 인위적인 오류값(SNR)을 집어 넣어 영상에 대한 잡음을 추가하여서 실험하였다. 실험의 결과는 다음의 표에서 보여 주고 있다.

<표 1> 자동차에 대한 인식 결과(%)
<Table 1> Result of recognition the sample data

	M1	M2	M3	M4	평균
Noise-Free	100.00	100.00	100.00	100.00	100.00
25dB SNR	100.00	100.00	100.00	100.00	100.00
20dB SNR	100.00	100.00	100.00	100.00	100.00
15dB SNR	100.00	100.00	100.00	95.43	98.86
10dB SNR	97.00	92.97	94.50	95.14	94.90

6. 결 론

본 연구에서는 2차원 영상을 인식하는데 있어서 스펙트럼 분석기와 신경회로망인 역전파를 이용하여 인식을 하였다. 대상 물체에서 획득되는 영상정보로부터 대상 물체의 종류를 인식하기 위해서는 획득되는 영상의 위치, 크기, 또는 회전에 무관하게 대상 물체를 식별해야 하며, 영상획득시 불가피하게 수반되는 잡음 속에서도 대상물체를 식별할 수 있어야 한다.

이를 위해 본 논문에서는 일련의 사전 전처리과정을 통해 영상의 위치, 크기, 회전으로 인한 영향, 그리고 잡음으로 인한 외곽선의 일부가 손상된 패턴벡터를 추출한 후, 이 패턴 벡터를 신경회로망의 입력으로 사용하였다.

본 논문에서 제안된 방법으로 실험을 하고 성능평가 위한 컴퓨터 모의 실험 대상은 4개의 실제 자동

차 데이터로 학습을 실시한 후, 각 자동차에 대해 추출된 외곽선을 확대, 축소, 회전하거나 가우시안 분포를 갖는 잡음을 다양한 신호대 잡음비로 더하여 총 5904개의 실험용 영상을 만들었다. 이 실험용 영상을 사용하여 인식 실험을 한 결과 스펙트럼 분석기와 역전파 신경회로망을 이용하였을 경우 매우 좋은 인식률을 보였다.

신경회로망에서의 학습은 상대적으로 간단하였고 빠른 학습 알고리즘과 모델트 기법을 이용하므로써 학습 패턴이 많지 않은 경우에도 좀더 나은 결과를 얻을 수 있었다.

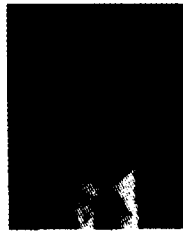
향후 연구할 내용은 본 논문에서는 외곽선 정보가 추출이 가능한 경우를 전제로 했는데, 대상의 영상을 획득하는 경우에 따라서는 외곽선 정보의 추출이 어려운 경우도 있고, 외곽선이 물체를 대표하는 특징이 되지 못하는 경우가 있다.

앞으로는 이러한 경우에 대한 연구가 요구되며 좀더 다양한 패턴에 대하여 학습을 하고 실험을 한다면 좀더 나은 결과를 얻을 수 있을 것이다.

참 고 문 헌

[1] 오영환, "패턴인식론", 정익사, 1991.
 [2] 김희승, "영상인식", 생능출판사, 1994.
 [3] 권희용, "복소-대수 사상과 수정된 이차 신경망을 이용한 회전 및 크기변형에 무관한 패턴 인식", 서울대학교 박사학위논문, 1993.
 [4] L. Gupta and M. D. Srinath, "Invariant Planar Shape Recognition Using Dynamic Alignment", *Pattern Recognition*, Vol. 21, No. 3, pp. 235-239, 1988.
 [5] J. G. Leu and I. N. Huang, "Planar shape matching based on binary tree shape representation", *Pattern Recognition*, Vol. 21, No. 6, pp. 607-622, 1988.
 [6] L. Gupta, M. R. Sayeh and R. Tammana, "A neural network approach to robust shape classification," *Pattern Recognition*, Vol. 23, No. 6, pp. 563-568, 1990.
 [7] 김상국, "신경망을 이용한 항공기의 패턴인식에 관한 연구", 경희대학교 석사학위논문, 1995.

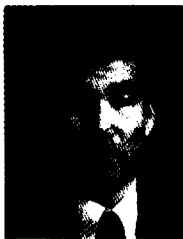
[8] 이용호, "M 변환과 환 이동 불변 신경망을 이용한 효율적인 패턴인식", 서울대학교 박사학위 논문, 1994.
 [9] R. C. Gonzalez and R. E. Woods, "*Digital Image Processing*", Addison-Wesley Publishing Company, Inc., 1992.
 [10] Anil K. Jain, "*Fundamentals of Digital Image Processing*", Prentice-Hall, 1993.
 [11] N. Ahmed and R. M. Bates, "A powerspectrum and related physical interpretation for the multidimensional BIFORE transform," *Proc. 1971 Applications of Walsh Functions Symp.*, Naval res. Lab., AD-727-000, pp. 47-50, 1971.
 [12] Laurene Fausett, "*Fundamentals of Nueral Networks*", Prentice-Hall, 1994.



차 승 운

1985년 서울산업대학교 전자계산학과(학사)
 1989년 성균관대학교 경영대학원 정보처리전공(석사)
 1996년~현재 관동대학교 전자계산공학과 박사과정
 1992년~현재 영동전문대학 전자계산과 조교수

관심분야: 패턴인식, 신경망회로, 프로그래밍언어론



김 성 략

1973년 명지대학교 전자공학과 졸업(학사)
 1979년 명지대학교 대학원 전자공학과(석사)
 1991년 단국대학교 대학원 컴퓨터전공(공학박사)
 1981년~현재 관동대학교 전자계산공학과 교수

관심분야: 전자계산기구조, 패턴인식, 신경망회로