

論文93-30B-2-8

ART와 다층 퍼셉트론을 이용한 얼굴인식 시스템의 성능분석

(Performance Analysis of Face Image Recognition System Using A R T Model and Multi-layer perceptron)

金榮日*, 安敏玉*

(Young Il Kim and Min Ock Ahn)

要約

최근 인간과 기계간의 효율적인 상호작용을 통한 쏘자동 화상인식 시스템의 필요성이 제기되고 있다. 그러나 감지기 동작에 따른 잡음이나, 형태변환 등으로 인하여 화상이 고정된 경우의 인식 시스템을 구현하는 것조차 쉽지가 않다. 본 논문은 패턴인식 분야에 매우 적합한 것으로 보고되고 있는 신경회로망 모델을 얼굴인식 문제에 적용하여 기존의 인식 시스템에서의 문제점들을 해결하고자 하였다. 카메라를 통해 얻어진 고정화상에 대해 ART 모델을 수행하였다. 새로운 얼굴이 입력되면 신경망은 후보코드를 출력하여 매칭점수를 계산한 결과, 유의수준 0.6 - 0.8에서 이 시스템은 100%의 인식율을 보였다. 가변 화상인식 시스템에서는 65개의 불변 모멘트 특징점이 추출되어, 39개는 다층 퍼셉트론을 training하기 위해 사용되었고, 26개는 testing을 위해 쓰였다. 그 결과 92.5%의 인식율을 보였다.

Abstract

Automatic image recognition system is essential for a better man-to machine interaction. Because of the noise and deformation due to the sensor operation, it is not simple to build an image recognition system even for the fixed images. In this paper, neural network which has been reported to be adequate for pattern recognition task is applied to the fixed and variational(rotation, size, position variation for the fixed image)recognition with a hope that the problems of conventional pattern recognition techniques are overcome. At fixed image recognition system, ART model is trained with face images obtained by camera. When recognizing an unknown face, this network generates a candidate code and its corresponding matching score. In the test when vigilance level 0.6 - 0.8, the system has achieved 100% correct face recognition rate. In the variational image recognition system, 65 invariant moment features sets are taken from thirteen persons. 39 data are taken to train multi-layer perceptron and other 26 data used for testing. The result shows 92.5% recognition rate.

1. 서론

*正會員, 大林專門大學 電氣科

(Dept. of Elec. Eng., Daelim Juni Univ.)

(※본 논문은 92학년도 학교법인 대림학원 학술연구 조성비에 의해 연구 되었음.)

최근 인간의 뇌가 대량의 복잡한 데이터를 효율적으로 병렬처리할 수 있을뿐만 아니라 학습능력이 있다는 사실에 근거하여 새로운 계산방식으로서 신경

회로망이 제안되었으며 여러분야에 성공적으로 응용되고 있다.^{[1] [2]}

신경회로망의 응용분야는 매우 넓지만 가장 일반적인 형태는 패턴인식이라 할 수 있다. 입력패턴이 주어지면 이 패턴이 어떤 영역에 속하는지를 알아내는 것이 목적이다.

신경회로망으로 패턴인식기를 구성할 때, 그 구조와 학습규칙에 따라 여러 종류가 있을 수 있다. 먼저 구조적으로 단층(single layer)과 다층(multi layer)으로 나누어질 수 있고, 학습규칙으로 Hebb rule을 쓰느냐 delta rule, back propagation rule을 쓰느냐에 따라 각각 성능과 특성이 결정된다. 일반적으로 다층구조가 단층구조보다는 강력하고, Hebb rule은 서로 직교성을 갖는 입력패턴만 구분할 수 있는 반면 delta rule은 선형적으로 독립인 패턴에 대해서도 적용된다. 또한 Hebb rule, delta rule, back propagation rule 등이 모두 지도학습인 반면에, 경쟁 학습은 자율학습이다.

신경회로로 되어있는 패턴인식기가 기존의 패턴인식기에 비해 우수한 점이 있다면 다음과 같다. 첫째, 구조적으로 병렬이기 때문에 모든 계산이 병렬로 이루어지므로 처리속도가 빠르다. 둘째, 입력패턴이 시험패턴(training pattern)에 비해 약간 손상된 형태 일지라도 인식할 수 있다. 셋째, 새로운 입력패턴에 대해서 가장 비슷한 범주를 찾을 수 있다. 넷째, 구조가 간단하고 복잡한 소프트웨어 부호화가 필요없다. 패턴인식은 고도의 화상해석 시스템에서 필수적인 분야이다. 전형적인 컴퓨터비전 시스템의 목적은 주어진 배경에서 화상을 해석하여 배경 내의 내용을 인식하는 것이다. 이와같은 로봇공학, remote sensing, 문서처리, 산업자동화 등 여러 분야에서 사용중이다.

이들 시스템의 대부분은 다음과 같은 4가지 블록으로 구성된다. 첫째는 화상취득으로 컴퓨터가 처리할 수 있는 배열로 변환시키는 것을 말한다. 둘째는 전처리 단계로 잡음제거 및 화질개선 등의 작업을 한다. 셋째는 특징추출 단계로 화상데이터의 차원을 줄이고, 데이터로부터 redundancy를 없애기 위해 특징점(feature)이라 불리는 숫자의 set으로 표현된다. 좋은 특징점은 같은 부류(intra-class)끼리는 차이가 적어야 하고, 다른 부류(inter-class)끼리는 차이가 커야 한다. 넷째는 분류단계로 화상인식 시스템의 마지막 단계이다.

신경회로망을 이용하여 패턴인식 시스템을 구현하는 데는 두 가지 방법이 있다. 첫째는 특징추출단계를 거치지 않고 입력패턴을 직접 신경회로망에 입력하여 패턴을 인식하는 방법이고, 또 다른 방법은 전

통적인 패턴인식 시스템의 구조에서 분류기 부분만을 신경회로망으로 대체하는 방법이다. 본 논문에서는 두 번째 방법을 이용하여 얼굴인식 시스템을 구현하고자 한다. 그러므로 특징추출을 위해서는 EDGE와 불변 모멘트를 이용하였고, EDGE와 불변 모멘트에 의해 얻어진 특징 벡터가 분류기로 사용된 신경망인 ART와 다층 퍼셉트론의 입력으로 사용된다.

본 논문은 5장으로 구성되어 있으며, 2장에서는 특징점인 EDGE와 불변 모멘트에 대해서 설명하였고, 3장에서는 분류기로 사용된 신경회로망에 대해서 설명하였으며, 4장에서는 실험결과를 보이고, 마지막 5장에서는 결론에 대해서 논하였다.

II. 특징점(Feature)

1. Edge 검출

edge 추출을 위해서는 RAMAMURTHI에 의해 제안된 아래의 알고리즘을 사용하였다.^[5] (NXN)크기의 블록에 대해 다음과 같은 방법으로 윤곽부를 검출한다.

① K, L을 0으로 SET한다.

② $g_{i,j}$ 를 (i,j)번째 pixel의 gray level이라 할 때

$$\Delta X_{ij} = g_{i+1,j} - g_{i,j}$$

$$Tx_{i,j} = (T/2)(g_{i+1,j} - g_{i,j})$$

를 계산한다.

(여기서 T는 미리 정해진 threshold이다.)

③ $K \leftarrow K+1$ if $\Delta X_{i,j} > Tx_{i,j}$

$$K \leftarrow K-1$$
 if $\Delta X_{i,j} < -Tx_{i,j}$

$$(i=1, \dots, N-1 \quad j=1, \dots, N)$$

④ $\Delta y_{i,j} = g_{i,j+1} - g_{i,j}$

$$Ty_{i,j} = (T/2)(g_{i,j+1} - g_{i,j})$$

를 계산한다.

⑤ $L \leftarrow L+1$ if $\Delta y_{i,j} > Ty_{i,j}$

$$L \leftarrow L-1$$
 if $\Delta y_{i,j} < -Ty_{i,j}$

⑥ if $K < m$ and $L < m$, then monotone

if $K < m$ and $L > m$, then edge, 0°

if $K > m$, $L > m$ and $\text{sgn}(K) = \text{sgn}(L)$, then edge, 45°

if $K > m$ and $L < m$, then edge, 90°

if $K > m$, $L > m$ and $\text{sgn}(K) = -\text{sgn}(L)$, then edge, 135°

(여기서 m은 미리 정해진 threshold이다.)

T의 선정은 윤곽부와 평탄부 영역의 구분에 영향을 준다. T는 0.1 - 0.2가 적당하다. m은 윤곽부의 강도를 정해준다. m이 크면 비교적 뚜렷한 윤곽부만 검출된다.

2. 불변 모멘트

한 화상으로부터 불변 특징점을 추출하는 것은 패턴인식에서 중요한 분야이다. 불변 특징점에 대한 수학적 배경은 대수적 invariant에 기초를 두고 있으며, 이 이론은 어느 좌표변환하에서도 변하지 않는 대수적 함수를 다루는 것이다. Hu는 대수적인 불변성을 이용하여 2차원 불변 모멘트 특징점을 제안하였다. 이와같은 불변 특징점을 사용하면 image내의 물체가 이동, 크기, 회전변화에 상관없이 동일하게 판단할 수가 있는 것이다. 일반적으로 기하학적 모멘트는 단항 $x^p y^q$ 에 입력함수 $f(x,y)$ 를 투영(projection)하는 형태로 다음과 같이 정의된다. [7]

$$M_{pq} = \iint x^p y^q f(x,y) dx dy \quad (1)$$

$p, q = 0, 1, 2, \dots$

식(1)에서의 모멘트는 각 차수마다 특성을 가지고 있는데, 영차(zero order)모멘트 M_{00} 는 입력영상 $f(x,y)$ 의 총 에너지를 의미하며, 1차 모멘트 M_{01}, M_{10} 는 영상분포의 중심위치를 표시한다. 2차 모멘트 M_{20}, M_{11}, M_{02} 는 영상의 크기 및 회전에 관한 정보를 포함하고 있다.

영차 모멘트와 1차 모멘트를 이용하면 위치에 대해 불변인 특성을 갖는 중심 모멘트(central moment)를 아래와 같이 구할 수 있다.

$$\mu_{pq} = \iint (x - x_c)^p (y - y_c)^q f(x,y) dx dy \quad (2)$$

$x_c = M_{10} / M_{00}, \quad y_c = M_{01} / M_{00}$

그리고 크기에 대해 불변인 모멘트를 얻기 위해서는 중심 모멘트를 다음과 과 같은 정규화된 모멘트로 변환을 해야한다.

$$n_{pq} = \mu_{pq} / \mu_{00}^r \quad (3)$$

$r = (p + q + 2) / 2$

즉, 정규화된 모멘트를 이용하면 영상의 크기 및 위치에 상관없이 같은것으로 판단하는 것이 가능하다.

또한 회전에 불변인 특성을 갖는 모멘트를 얻기 위해서는 정규화된 모멘트를 비선형 조합하면 식(4) - 식(10) 같은 7개의 불변 모멘트를 구할 수 있다. 그림 1은 입력 데이터로부터 불변 모멘트 특징벡터를 구하기 위한 블럭도이다.

$$\Phi_1 = n_{20} + n_{02} \quad (4)$$

$$\Phi_2 = (n_{20} - n_{02})^2 + 4n_{11}^2 \quad (5)$$

$$\Phi_3 = (n_{30} - 3n_{12})^2 + (3n_{21} - n_{03})^2 \quad (6)$$

$$\Phi_4 = (n_{30} + n_{12})^2 + (n_{21} + n_{03})^2 \quad (7)$$

$$\Phi_5 = (n_{30} - 3n_{12})(n_{30} + n_{12}) [(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] + (3n_{21} - n_{03})(n_{21} + n_{03}) [3(n_{03} + n_{12})^2 - (n_{21} + n_{03})^2] \quad (8)$$

$$\Phi_6 = (n_{20} - n_{02}) [(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] + 4n_{11} (n_{30} + n_{12})(n_{21} + n_{03}) \quad (9)$$

$$\Phi_7 = (3n_{21} - n_{03})(n_{30} + n_{12}) [(n_{30} + n_{12})^2 - 3(n_{21} + n_{03})^2] - (n_{30} - 3n_{12})(n_{21} + n_{03}) [3(n_{30} + n_{12})^2 - (n_{21} + n_{03})^2] \quad (10)$$

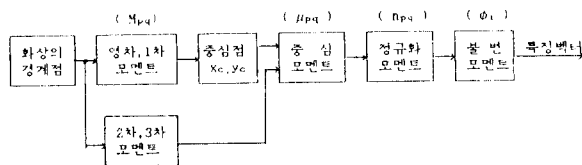


그림 1. 불변특징벡터를 구하기 위한 블럭도

Fig. 1. Block diagram for invariant feature vector.

Ⅲ. 신경회로망 분류기

1. ART 모델

ART(Adaptive Resonance Theory)모델은 이미 취득된 하향(top-down) 기대(expectation)를 이용해서 기존에 배웠던 것이 새로운 학습에 의해 지워지지 않도록 하고, 또 새로운 지식이 자동적으로 전체 지식 base에 자기일치(self-consistence)하는 방법으로 통합되도록 상향(bottom-up) 정보를 적응시키는 것이다. [3] 결과적으로 임의의 입력패턴을 받아 실시간 안에 안정된 인식코드를 스스로 구성하는 신경회로망인 것이다. 가중치와 활성화를 갖는 이 모델은 그림 2와 같은 이층구조(2-layer)로 되어 있으며 가중치를 정규화 하는 대신 입력을 정규화 하였다.

이 경쟁학습은 F1에 너무 많은 입력이 제시되거

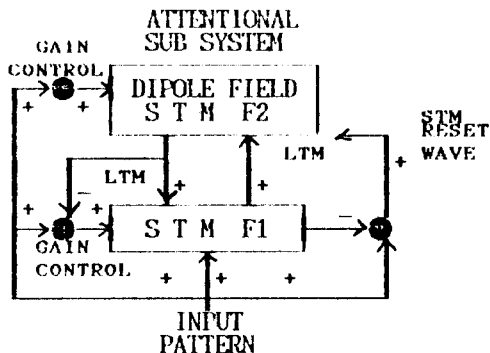


그림 2. ART 시스템

Fig. 2. ART system.

나, 또 F2의 기억 노드 수에 비해 입력패턴이 너무 많은 범주를 가지고 있는 경우를 제외하고는 인식부호의 학습은 결국 안정하게 되고 가중치는 입력환경의 구조를 반영하게 된다. 그림 2는 자신의 기억용량을 다 쓸 때까지 2진 입력패턴을 안정하게 인식부호로 배울 수 있는 ART 시스템의 구조이다. ART 시스템은 경쟁학습 모델과 마찬가지로 새로운 입력패턴을 상향 필터의 적응 가중치를 바꿈으로서 부호화된다.

2. 수정된 ART 알고리즘

본 논문에서 사용한 신경회로망에서 각 뉴우런의 상태값은 1 또는 0의 값만 가지는 2진수로서 입력과 출력도 1 또는 0의 값을 갖는다. 사용하는 비선형 함수는 하드 리미트(hard limit)함수이며, 처리중 연결강도와 임계값은 가변된다. 얼굴인식 알고리즘을 정리하면 다음과 같다.

1) 64×64의 입력노드와 20개의 출력노드를 갖는 網을 설계하고 하향 가중치와 상향 가중치를 각각 초기화 한다.

$$\begin{aligned} td_{ij} &= 1 & 0 \leq i \leq N-1 \\ bu_{ij} &= 1 / (1 + N) & 0 \leq j \leq M-1 \end{aligned}$$

$td_{ij}(t)$: 입력노드 i 와 출력노드 j 사이의 하향연결 가중치

$bu_{ij}(t)$: 입력노드 i 와 출력노드 j 사이의 상향연결 가중치

N : 입력노드의 개수

M : 출력노드의 개수

2) 첫번째의 입력패턴을 표준(exemplar)패턴으로 하고 출력단의 첫번째 노드에 기억시킨다.

3) 새로운 입력을 받아 들인다.

4) 기억된 패턴들과의 매칭점수를 계산한다.

$$u_j = \sum_{i=0}^{N-1} bu_{ij}(t) \cdot X_i \quad 0 \leq j \leq M-1$$

u_j : j 번째 출력노드의 출력값

X_i : 입력의 i 번째 값(0 또는 1)

5) 매칭점수가 가장 큰 출력노드를 선택한다.

$$u_j^* = \max(u_j)$$

j

6) 유의시험(vigilance test)을 한다.

$$|X| = \sum_{i=0}^{N-1} X_i$$

$$|TD \cdot X| = \sum_{i=0}^{N-1} td_{ij} \cdot X_i \quad \text{일 때.}$$

$|TD \cdot X| / |X| \geq \rho$ 인가를 확인한다.

만족하면 단계 ⑦로, 그렇지 않으면 단계 ⑧로 간다.

7) 최적 매칭 표본에 적응한다.

$$td_{ij}^*(t+1) = td_{ij}^*(t) \cdot X_i$$

$$bu_{ij}^*(t+1) = (td_{ij}^*(t) \cdot X_i) / (0.5 + \sum_{i=0}^{N-1} td_{ij}^*(t) \cdot X_i)$$

즉, 표준패턴과 동일한것으로 판단되어 그 패턴에 논리적 AND시키며 가중치는 자동 조절된다.

8) 새로운 패턴으로 판정하여 다음 출력노드에 저장한다.

3. 지도학습 모델

학습은 입력패턴과 출력패턴의 쌍의 집합을 가지고 수행된다. 먼저 입력패턴이 다층 퍼셉트론에 입력되어, 신경망에 의한 출력패턴이 계산된다. 만약 이 출력패턴이 목표출력패턴과 일치하면 학습은 중단된다. 그러나 이 두 값 사이에 차이가 있다면 차이를 감소시키는 방향으로 연결선의 가중치를 변경시키므로 LMS(Least Mean Square)절차라 불리기도 한다. 그리고 다층퍼셉트론의 각 노드를 활성화시키기 위한 함수로 선형함수를 사용하면 다층신경망이 등가의 단층신경망으로 변형이 가능하므로 다층구조의 장점을 살릴 수 없다. 그러므로 활성화 함수로는 비선형 함수를 사용하여야 한다. 또한 윗층의 함수는 목표패턴과 출력패턴의 차이인 항을 계산하는 과정에서 미분을 할 필요가 있으므로 미분가능한 함수이어야 한다. 이와같은 조건을 만족하는 함수를 유사선형함수(semilinear function)라고 하며 오차 역전파(error back propagation) 알고리즘에서는 식(11)의 시그모이드(sigmoid)함수를 사용한다.

$$O_{ij} = 1 / (1 + \exp(-a \cdot \theta)) \quad (11)$$

실제로 구현에 사용된 식은 식(12)와 같다.

$$\Delta W_{ij} = \eta \delta_{pj} O_{pj} + a \Delta W_{ij}(n) \quad (12)$$

식 (12)에서는 두개의 파라미터 η 와 a 가 등장하는데 이들의 역할은 다음과 같다. η 는 학습율로서 이 값이 클수록 가중치의 변화량이 큰값을 갖게된다. 학습율이 크면 일반적으로 학습을 빨리 마칠 수 있다는 장점이 있지만 한편으로는 문제의 해답을 찾지 못하고 진동을 하는 경우가 생긴다. 그러므로 학습율은 진동이 일어나지 않는 범위안에서 최대한 크게하는 것이 바람직하다. 파라미터 a 는 진동을 막으면서 학

습율을 크게 해주기 위해서 도입된 것으로 바로 전번의 반복 계산시에 구한 가중치 변화량을 다음 단계에서도 유지하도록 하는 역할을 하며 이 항을 관성항(momentum)이라고 부른다. 실험결과에 의하면 학습율이 크고 관성항이 작을수록 학습시간이 줄어드는 사실을 관찰할 수 있었다. 또한 역전과 분류기는 긴 학습시간과 적은 기억장소가 요구되며 빠른 분류 결정이 이루어진다. 그리고 이밖의 다른 분류기는 역전과 분류기와 nearest neighbor 분류기 중간에 있는 것을 알 수 있었다.

IV. 실험 및 결과

1. 화상인식 시스템 I (Edge + ART)

실험은 PC-386을 이용하여 C언어로 동작 시켰으며, 카메라(CCTV)를 이용하여 영상 데이터를 받아 FG-100-AT 영상보드에서 처리 하였다.

먼저 13명의 얼굴을 카메라로 받아서 256×256 크기의 영상 데이터로 저장한다. 영상을 겹치지 않고 연속하는 (4×4)크기의 블록으로 나누어 한 블록내에서의 값이 평균값보다 크거나 같으면 '1'로, 작으면 '0'으로 대체하여 블록을 이치화한 64×64 크기의 이진 윤곽 데이터를 얻는다. 이때 임계값을 각각 0.08, 0.09, 0.10, 0.11, 0.12로 변화시켜 윤곽정보에 변화가 생긴 총 65장의 영상 데이터를 얻는다. 표 1은 임계값이 0.12일 때 각 입력화상의 2진 데이터량을 보여주고 있다.

표 1. 2진 데이터량(임계값 T=0.12)

Table 1. Binary data(threshold T=0.12).

입력화상	1	2	3	4	5	6	7	8	9	10	
데이터량	0 1	3211 885	3205 891	3223 873	3201 895	2229 867	3244 852	3046 1095	3085 1011	3231 865	3211 885

64×64의 입력노드와 20개의 출력노드로 확장한 모델에 이 데이터를 적용한 결과 유의수준 0.6 - 0.8에서 13개의 표준패턴으로 정확히 분류 되었으나, 유의수준 0.9에서는 16개의 패턴으로 분류되었다. 이것은 윤곽정보를 얻기 위해 바꿔준 임계값이 0.08, 0.12인 경우의 데이터가 연속해서 입력되었을 때 그 차이가 커서 생긴 결과이다. 즉, 정보의 일치가 90%에 미치지 못하기 때문에 새로운 패턴으로 판단하여 저장하는 것이다. 그러나 다른 영상패턴과 일치하는 것으로 잘못 판단하는 경우는 없었으며, 이것은 얼굴

영상의 이질성 때문인 것으로 생각된다. 표 2는 임계값에 변화를 줌에 따라 동일한 입력화상의 윤곽정보량이 달라지는 것을 보여준다. 표 3은 유의수준에 따른 인식율과 유의시험을 거쳐 최종적으로 출력노드에 기억된 패턴의 수이다. 유의수준이 낮으면(0.0 - 0.5), 조금만 비슷한 패턴이 들어와도 매칭 점수가 제일 큰 노드의 기억패턴과 동일한 것으로 판단하고, 반대로 유의수준이 큰 경우(0.9 - 1.0)는 거의 똑같은 정보가 들어올 때만 동일패턴으로 판단한다. 그러므로 분야별 특성에 따라 유의수준을 조절하여 적당한 인식율로 바꾸어 사용할 수 있다.

표 2. 임계값에 따른 데이터량의 변화

Table 2. The change of data due to threshold.

임계값	0.08	0.09	0.10	0.11	0.12
윤곽데이터량	1058	1017	975	924	885

표 3. 유의수준에 따른 인식율

Table 3. The recognition rate due to vigilance level.

256×256의 원 화상을 64×64의 edge로 변환한 것을 처리한 경우 64×64의 원 화상을 64×64의 edge로 변환한 것을 처리한 경우

유의수준	0-0.5	0.6-0.89	0.9-1.0	유의수준	0-0.4	0.5-0.95	0.95-1.0
인식율	-	100%	-	인식율	-	100%	-
저장패턴수	7개 이하	13개	16개 이상	저장패턴수	10개 이하	13개	15개 이상

그림 3의 (a), (b), (c), (d)는 카메라를 통해 얻어진 얼굴화상이며, 그림 4는 각각의 얼굴화상에 대해 최종적으로 분류, 기억된 2진 윤곽 화상이다.

그림 5는 얼굴화상 (a)에 대해 정보량이 다른 5개의 윤곽 데이터가 학습이 진행되는 동안 서로 동일한 것으로 판단되어 적용되어 가는 즉, 논리적 AND 작용이 일어나는 과정 (b), (c), (d), (e), (f)를 보여준다. (f)는 최종적으로 출력노드에 기억된 윤곽선 화상이다. 최종 기억된 화상 데이터는 윤곽 데이터의 손실을 가지나, 가중치의 증가로 인하여 윤곽 데이터의 정보값은 증가한다. 그래서 다음 입력정보에 대한 반응은 더욱 강해지는 것이다.

계산량을 살펴보면 65개 입력 데이터의 처리에 걸린 총 소요시간은 20-25분 정도이며, 매칭점수를 계산하여 기억노드를 선택하고 유의시험을 거치는데는 약 15초가 걸렸다.

Competitive Learning은 Learning의 속도가 빠

르고 teaching이 필요없는 반면, 특정한 출력과 입력패턴을 관련시키기가 어렵고, 초기 가중치에 크게 의존하는 문제점도 있다. 그러나 본질적으로 Competitive Learning은 인간의 학습과 유사점을 갖고 있으므로 중요한 Learning rule중의 하나이다.

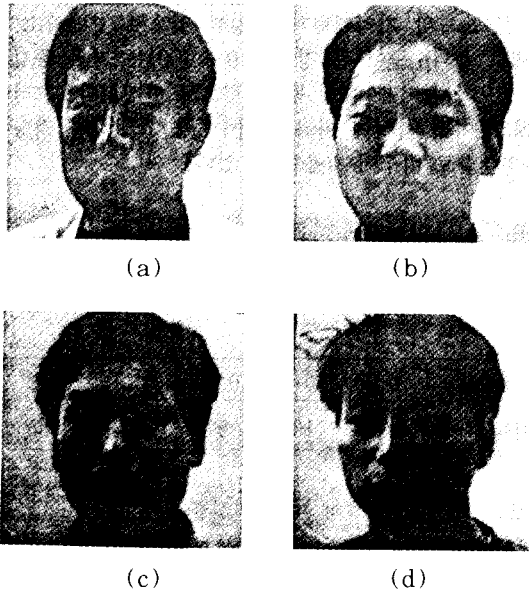


그림 3. 원 화 상
Fig. 3. Original images.

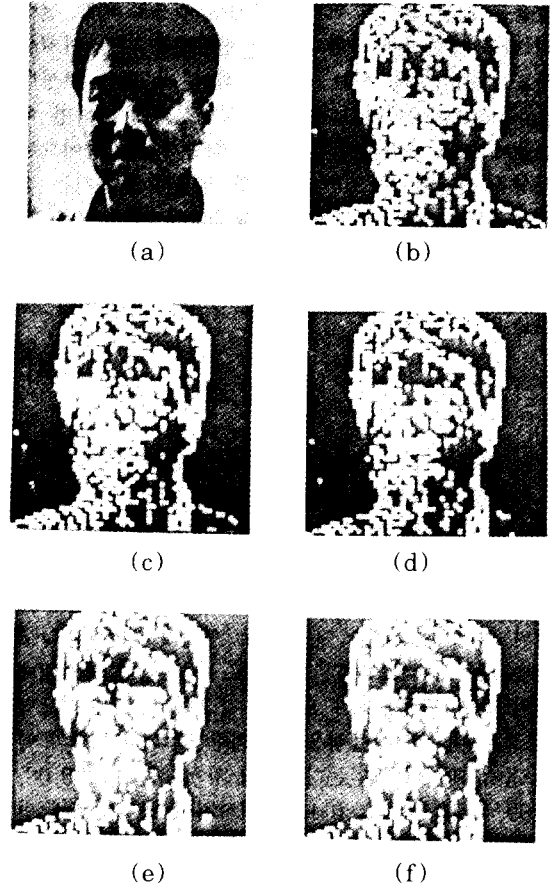


그림 5. 적응과정
Fig. 5. Adaptation procedure.

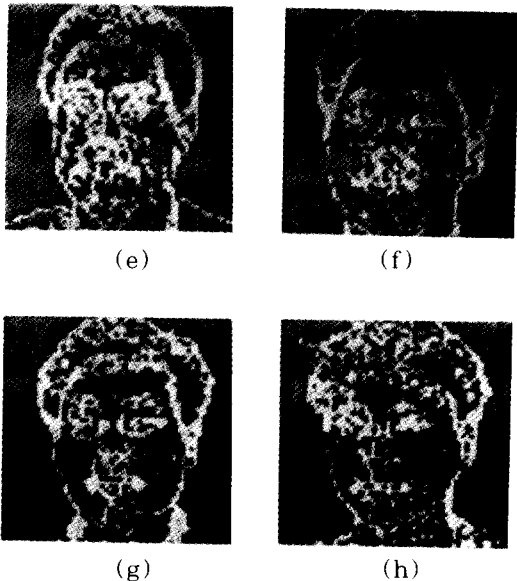


그림 4. 실험결과 화상
Fig. 4. Reconstructed images.

2. 화상인식 시스템 II(불변 모멘트+다층 퍼셉트론)
두번째 인식 시스템에서는 그림 6(a)와 같은 256 × 256 크기의 화상 데이터를 원 화상으로 저장한다. 그리고 1명당 그림 6(b) - (e)와 같이 변환된 화상들을 추가로 저장하여 총 65장의 화상 데이터를 얻는다. 그림 6(b)는 원 화상의 크기를 반으로 줄인 것이고, 그림 6(c)는 원 화상의 mirror 화상이다. 그리고 그림 6(d) - (e)는 원 화상을 각각 2° 와 20° 회전시킨 것이다.

식(4) - (10)에 주어진 7개의 불변 모멘트가 이들 화상에 대해 각각 계산되어지고, 계산된 불변 모멘트 값의 dynamic range를 줄이기 위해서 logarithm 을 취한 결과를 사용한다. 표 4에 나타난 것처럼 그림 6(b) - (e)의 결과는 원 화상에 대해 계산된 불변 모멘트값과 잘 일치한다. 그러나 데이터의 디지털적인 속성 때문에 undersampling에 의한 scale invariant error와 digitization에 의한 rotation

invariant error 그리고 양자화에 의한 error를 피할 수는 없다.^[9]

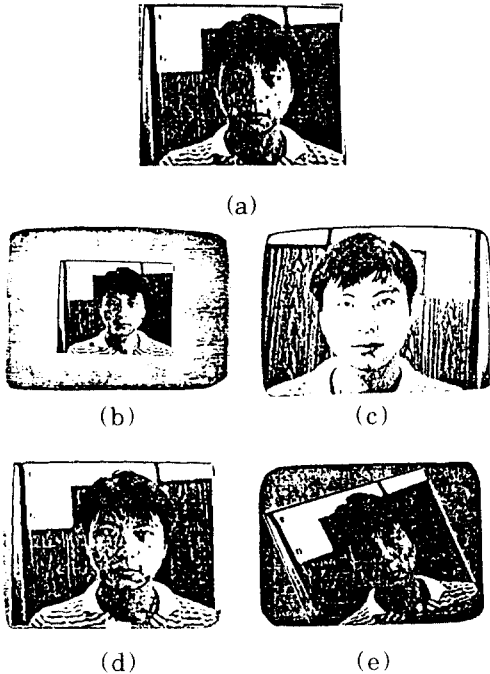


그림 6. 실험에 사용된 화상
Fig. 6. Images used in experiments.

표 4. 그림 6(a)-(e)에 대한 불변 모멘트
Table 4. Moment invariants for the images in Fig.6(a)-(e).

불변벡터 (LOG)	Original	Half Size	Mirrored	Rotated2°	Rotated20°
ϕ_1	1.7608	1.7529	1.7616	1.7584	1.7333
ϕ_2	4.8541	4.8295	4.6935	4.8050	4.5959
ϕ_3	6.2620	6.2001	6.3278	6.2189	6.2835
ϕ_4	6.4320	6.4483	6.3366	6.3578	6.3496
ϕ_5	13.8927	13.6237	13.3821	14.4169	13.6135
ϕ_6	8.9767	8.9813	8.7709	8.8787	8.7723
ϕ_7	12.7804	12.7768	12.6770	12.6463	12.6698

이렇게 하여 얻은 총 65개의 불변 모멘트 특징벡터가 같은 부류별로 분류되기 위해 분류기로 사용된 다층 퍼셉트론의 입력으로 사용된다. 즉, 다층 퍼셉트론의 입력노드수는 7이 되고, 출력노드수는 13이 되며, 가변적인 노드수를 갖는 하나의 중간층이 존재한다.

다층 퍼셉트론을 back propagation rule을 이용

하여 학습(learning) 시키기 위해서 총 65개의 특징 벡터중에서 39개가 사용되고, 나머지 26개가 testing을 위해 사용된다. training과 test 결과에 따른 인식율이 표 5와 그림 8에 나타나 있다. 그림 8을 보면 반복횟수 1500 이상과 중간층의 노드수가 30 이상에서는 인식율의 저하가 나타나는데, 이는 역전파 학습이 gradient descent method를 사용하기 때문에 종종 global minimum이 아닌 local minimum에 떨어지는 경우가 있고, 또한 이 알고리즘이 입력에 따른 출력을 보고 중간층의 가중치를 조정하므로, 중간층의 학습은 출력층에서의 가중치 조정과 비교해 볼 때 indirect하게 되기 때문으로 생각된다. 실험에서는 39개의 특징벡터를 학습 시키기 위해서 학습율과 관성율은 각각 0.7과 0.2를 사용하였다. 실험결과에 의하면 39개의 특징벡터를 학습시키기 위한 반복횟수가 1500, hidden layer의 노드수가 30일 때 가장 좋은 92.5%의 인식율을 보임을 알 수 있다.

그리고 39개의 특징벡터를 학습시키기 위해 소요된 시간은 2시간이었다.

표 5. 학습패턴과 시험패턴에 대한 인식율
Table 5. Recognition rate of learning and testing pattern.

Hidden Layer의 unit수	인식율(%) (training)	인식율(%) (testing)
0	87	70
5	90	75
10	91.6	82.5
15	91.6	80
20	95	82.5
25	95	85
30	96.7	92.5

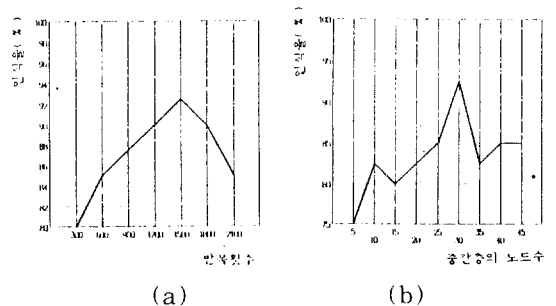


그림 8. 인식율 그래프
Fig. 8. Graph of recognition rate.

V. 결 론

첫번째 화상인식 시스템에서는 13명의 얼굴 데이터를 입력장치나 주변환경의 변화에 따른 잡음이나 형태의 변형을 대신하여 윤곽조건을 달리한 65장의 윤곽 화상 데이터를 만든 후 ART 신경회로망 모델에 적용한 결과 유의수준 0.6 - 0.8에서 13개의 패턴으로 정확히 분류되었다.

차원이 작은 화상인 경우 윤곽 정보량의 정규화를 통해 더욱 정확한 패턴분류가 가능하지만, 차원이 큰 영상의 경우 매우 작은 값이 할당되므로 계산상의 문제가 발생한다. 입력값을 정규화하지 않으면 상향 가중치를 크게(보통 1 이상) 하여야 컴퓨터 계산이 용이하지만 오히려 적응과정에서 상향 가중치가 약화되는 문제가 발생한다. 입력값을 정규화하는 대신 상향 가중치를 정규화하여 이 문제를 해결하였고, 또한 유의수준을 이용하여 패턴분류 정도를 조정할 수 있었다.

두번째 화상인식 시스템은 영상 데이터의 모멘트와 다층 퍼셉트론을 이용하여 위치, 크기, 회전에 무관한 인식 시스템으로, 불변 모멘트 특징점 자체가 위치, 크기, 회전 변화에 무관한 성질을 갖고 있기 때문에 신경회로망의 비선형적 의사결정기능과 결합하여 좋은 결과가 얻어짐을 알았다. 오인식의 원인을 보면, 영상 데이터의 디지털적인 특성으로 인한 undersampling error, digitization error, quantization error가 다층 퍼셉트론의 입력으로 사용되는 불변 모멘트 특징점에 영향을 미쳤기 때문으로 생각된다. 또한 불변 모멘트가 화상전체로부터 추출된 global feature이므로, 심한 잡음이나 화상이 부분적으로 절단된 화상에 대해서는 적용이 불가능하나 화상의 부분적 특징을 기술할 수 있는 local feature에 의한 화상인식을 통해 가능하리라 기대된다.

결론적으로 경쟁학습은 지도학습에 비해 속도가 빠르고 teaching이 필요 없으며, feature detection에는 최적이라 할 수 있는 반면, 근본적으로 특징한 출력과 입력을 관련시키기가 어렵고 초기 random weights에 크게 의존하는 등의 문제점도 있다.

參 考 文 獻

[1] D.E. Rumelhart and J. L. McClelland

(eds.), *Parallel Distributed Processing*, vol. I. MIT Press, 1986.

- [2] R. P. Lippman, "An Introduction to Computing with Neural Nets," *IEEE ASSP Magazine*, vol. 3, no. 4, pp. 4-22, April 1987.
- [3] G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine," *Computer Vision, Graphics, and Image Processing*, vol. 37, pp. 54-115, 1987.
- [4] G. A. Carpenter and S. Grossberg, "The ART of Adaptive Pattern Recognition by a Self-Organizing Neural Nets," *IEEE Computer*, pp. 77-88, Mar. 1988.
- [5] B. Ramamurthi and A. Gersho, "Nonlinear Space-Variant Processing of Block Coded Image," *IEEE ASSP-34*, no. 5, pp. 1258-1267, Oct. 1986.
- [6] A. Khotanzad and J. H. Lu, "Classification of Invariant Image Representation Using a Neural Network," *IEEE Trans. Acoustics, Speech, and Signal Processing*, vol. 38, no. 6, June 1990.
- [7] M. K. Hu, "Visual Pattern Recognition by Moment Invariants," *IRE Trans. Information Theory*, vol. 8, pp. 179-187, Feb. 1962.
- [8] R. P. Lippman, "Pattern Classification Using Neural Networks," *IEEE Communication Magazine*, pp. 47-64, Nov. 1989.
- [9] C. H. Teh and R. T. Chin, "On Digital Approximation of Moment Invariants," *Computer Vision, Graphics, and Image Processing* 33, pp. 318-326, 1986.
- [10] Yoh Han Pao, *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley, 1989.

— 著 者 紹 介 —



金 榮 日 (正會員)

1952年 10月 27日生. 1983年 송
실대 공대 전기공학과 졸업(학사)
1985년 연세대 대학원 전기공학과
졸업(석사) 1989년 동 대학원 전기
공학과 졸업(공박) 1987~89년 금
성산전 연구소 선임 연구원 1989
~1992년 동해중앙연구소장 1993~ 현재 대림전문대
학 전기과 전임강사 주관심분야는 컴퓨터비전 및 신
호처리, 자동화공학 등임.



安 敏 玉 (正會員)

1944年 8月 2日生. 1971年 2月 연
세대 공대 전기공학과 졸업 1975년
2월 연세대 공대 대학원 전기공학
과 석사학위 취득 1986년 2월 연세
대 공대 대학원 전기공학과 박사학
위 취득 1979년 2월 ~ 현재 대림
전문대학 전기과 교수 재직중 주관심분야는 제어계
측 및 자동화, 전기기기 등임.