

論文94-31B-7-14

지식에 기초한 특징추출과 역전파 알고리즘에 의한 얼굴인식

(Face Recognition Using Knowledge-Based Feature Extraction and Back-Propagation Algorithm)

李常榮*, 咸永國*, 朴來弘*

(Sang Young Lee, Young Kug Ham and Rae-Hong Park)

要約

본 논문에서는 얼굴의 특징추출과 신경회로망을 이용한 얼굴인식 알고리즘을 제안하였다. 우선 취득된 얼굴영상에서 얼굴부분이 중심에 위치하고 배경이 균일하다고 가정하여 얼굴과 배경을 분리한다. 이렇게 분리된 얼굴부분에서 수직·수평으로 투영하여 사람 얼굴에 대한 지식에 근거하여 특징을 추출하였다.

인식단계에서는 신경회로망의 오류역전파 알고리즘을 사용하여 인식하였으며, 학습시 학습속도를 줄이기 위해 학습률과 관성항을 가변시켜 사용하였다. 실험결과 본 방법은 여자 6명과 남자 14명에 대해 정확히 인식하였다.

Abstract

In this paper, we propose a method for facial feature extraction, and recognition algorithm using neural networks. First, we extract a face part from the background image based on the knowledge that it is located in the center of an input image and that the background is homogeneous. Then using vertical and horizontal projections, we extract features from the separated face image using knowledge base of human faces.

In the recognition step, we use the back propagation algorithm of the neural networks and in the learning step, to reduce the computation time, we vary learning and momentum rates. Our technique recognizes 6 women and 14 men correctly.

1. 서론

현재 많은 분야에서 컴퓨터가 사용되고 있는데, 응용 분야로서 관공서나 일반 주택의 보안 유지와 범죄 수사시 사람을 쉽게 구별할 수 있는 장치로 쓰일 수 있다. 사람을 구별하는 특징으로 목소리, 지문, '그리

고 신체 각 부분의 크기 등을 사용할 수 있으나, 사람에게 있어 서로를 구별하는 가장 정확한 방법의 하나가 얼굴의 특징을 이용하는 것이다.^[1] 얼굴을 인식하는 방법에는 측면얼굴의 윤곽 (silhouette)이나^[2-5] 정면얼굴의 특징들을 이용하거나^[6-13], 이 두가지 방법을 혼합하여 사용하는 방법 등^[14] 이 있는데, 측면얼굴을 이용할 경우 특징추출이 용이한 장점은 있으나 인식률을 높이기 위해서는 얼굴과 배경부분을 정확히 분리해야 한다는 등의 많은 어려움이 따르므로, 정면얼굴을 이용한 얼굴인식 시스템 개발이 많은

*正會員, 西江大學校 電子工學科
(Dept. of Elec. Eng., Sogang Univ.)
接受日字 : 1993年 11月 5日

접근 방법을 통하여 이루어지고 있다. 그러나 정면얼굴을 이용할 경우에도 다음과 같은 제약조건이 있어야 한다. 첫째, 안경 미착용시의 영상을 취득하여야 한다. 둘째, 눈을 감지 않은 상태에서 영상취득이 되어야 한다. 셋째, 턱수염이나 콧수염을 기르지 않은 사람의 영상이어야 한다. 넷째, 얼굴부분이 영상의 가운데 부분에 위치하도록 영상취득이 이루어져야 한다.

정면얼굴을 이용하여 얼굴을 인식하는 방법에는 2차 통계 특징 (second order statistics)을 이용하여 이진영상 (binary image)에서 정합 영역 (matching region)을 추출후 정합하는 방법^[11]과 얼굴을 구성하는 각 구성성분들의 template을 정합하는 방법^[12,13], 그리고 얼굴의 각 구성성분들 사이의 위치관계를 이용하는 방법 등^[8,9]이 있다. 2차 통계 특징을 이용한 정합방법의 경우, 정합 영역 추출시 조명에 의한 에지 변화에 민감하다는 단점이 있고, template 정합방법은 취득된 영상의 크기가 변할 경우 고정된 template에 의한 인식이 어렵다는 단점이 있다. 따라서 본 논문에서는 얼굴에 대한 기초지식을 이용하여 각 구성성분들을 추출한 후 이들 사이의 위치관계를 이용한 인식을 시도하였다. 여기서 기초지식이라 함은 얼굴영상은 대부분 좌우 대칭이고, 눈썹 아래 눈이 있고 그 아래부분에 코가 존재하며, 코 아래에 입술이 위치한다는 것 등이다. 이러한 위치관계를 이용한 기존의 논문은 얼굴의 형태 (shape)에 관련된 특징점들을 사용하여 인식을 시도하였는데, 이 경우 특징을 많이 사용하여 많은 사람을 구분할 수 있다는 장점은 있으나, 얼굴과 배경을 분리하는 단계에서 얼굴의 형태가 조명 및 각도에 의해 왜곡이 일어날 가능성이 커 오히려 인식률을 낮추는 원인이 된다. 이에 본 논문에서는 왜곡이 심하지 않은 정규화된 눈과 입의 크기 및 서로간의 위치관계에 대한 특징값 4개를 사용하여 얼굴인식 시스템을 구성하였다.

이러한 특징값들을 이용하여 인식하는 방법에는 유클리디안 거리 (Euclidean distance)를 이용한 방법과 신경회로망을 이용한 방법이 있는데, 전자는 시스템 구성이 간단한 장점은 있으나 사람수가 많아질수록 특징값들의 중첩을 해결하기 힘들기 때문에 인식률이 낮아지는 단점이 있다. 후자는 CP (counter propagation)와 BP (back propagation)가 있다. CP는 패턴분류 및 통계분석과 데이터 압축 등에 사용되고 있으며^[15], BP는 특히 패턴인식 등에 많이 사용되고 있다.^[16] 본 논문에서는 CP와 BP를 이용하여 얼굴인식을 행한 결과, CP는 학습시간이 BP에 비해 빠른 반면 인식률은 떨어지기 때문에 본 논문에서는 BP를 이용하여 얼굴을 인식하였다.^[17]

위의 방법들을 사용하여 얼굴인식 시스템을 구성한다면 다음과 같은 분야에 적용할 수 있을 것이다. 첫째, 화상전화에서 영상 전송시 얼굴인식 시스템을 이용하여 부호화한다면 적은 전송량으로도 충분한 화질을 얻을 수 있다. 둘째, 보안 시스템에서 열쇠의 분실시 따르는 불편함을 해결하기 위한 방법으로 얼굴인식 시스템이 필요하다. 셋째, 범죄수사시 범인의 몽타주를 수작업을 통하여 그리는 불편함을 해결하기 위해 데이터베이스화되어 있는 얼굴인식 시스템을 통해 진술에 상응하는 정보를 추출한다면 보다 편리하고 효과적인 결과를 얻을 수 있을 것이다.

본 논문에서는 얼굴의 각 구성성분들인 눈, 코, 그리고 입의 자동추출 방법과 신경회로망을 사용한 인식방법을 제안하였는데, 본 논문의 구성은 다음과 같다. II장에서는 특징들의 올바른 추출을 위한 전처리에 대해 기술하였고, III장에서는 특징추출 및 기운영상의 보정에 대하여 서술하였다. IV장에서는 신경회로망의 BP 알고리즘을 이용한 인식방법에 대하여 논하였으며, V장에서는 실험결과 및 분석에 대해 서술하였다. 마지막 VI장에서는 결론을 맺었다.

II. 전처리

영상취득시 나타나는 spot 잡음을 제거하기 위하여 3×3 중앙값 필터를 사용하여 제거한 후 네 방향 Sobel 에지 연산자를 사용하여 이진영상으로 만들었으며, 에지 추적 (tracking)에 의해 얼굴과 배경을 분리하였다.

1. 에지 추출

에지 추출 연산자는 얼굴의 구성성분들이 지니는 대각선 방향의 에지 성분들도 잘 검출할 수 있는 네 방향 Sobel 연산자를 사용하였으며, 문턱값 (threshold)은 Robinson이 제안한 LAT (Locally Adaptive Threshold)를 사용하여 자동으로 결정하였다.^[18] LAT는

z0			z1			z2			z3		
1	2	1	2	1	0	1	0	-1	0	-1	-2
0	0	0	1	0	-1	2	0	-2	1	0	-1
-1	-2	-1	0	-1	-2	1	0	-1	2	1	0

그림 1. Sobel 에지 연산자
Fig. 1. Sobel edge operators.

$$LAT = \frac{\text{Max}\{|zk|, k=0,1,2,3\}}{\text{Pixel}(x,y)\text{에서 low pass filter } M_0 \text{를 거친 값}} \quad (1)$$

로 주어지는데 여기서 $M_0 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$ 이고, zk 는 그림 1과 같은 Sobel 연산자를 적용하여 얻은 값이다. 위의 방법에 의해 최종 결정되는 에지점은 LAT의 값이 1보다 클 경우이다.

2. 얼굴과 배경의 분리

기존의 얼굴과 배경 분리방법을 살펴보면, Kelly^[6]는 planning 방법을 이용하여 얼굴과 배경을 분리하였고, Crow 등^[7]은 대략적인 얼굴 윤곽선인 8 8 template을 이용하여 원영상을 축소한 후 얼굴과 배경을 분리하였다. 지은미 등^[8]은 원영상을 1/4로 축소하여 이진화 영상으로 만든 후 각 방향 탐색 창(window)을 이용하여 얼굴과 배경을 분리하였다. Kato 등^[9]은 얼굴영상의 가장 긴 isodensity 라인을 얼굴과 배경의 경계선으로 결정하였다. Harasaki 등^[10]은 스테레오 영상의 초점으로 두 영상을 정합시킨 후 배경을 분리하였다. 본 논문에서는, 원영상을 이진화한 후 지식에 기초한 에지 추적 방법에 의하여 얼굴과 배경 및 어깨를 분리하였다. 또한 에지의 끊어짐이 적어야 이들을 잘 분리할 수 있으므로 이진영상에 대한 2 D 형태학의 팽창(morphological dilation)^[21]을 수행하였다. 그 방법은 식 (2)와 같다.

$$A_3 \oplus B_3 = \bigcup_{b \in B_3} (A_3 + b) \quad (2)$$

여기서 A_3 는 3×3 창이고, B_3 는 얼굴의 이진영상이다. 그리고 b 는 에지점으로서 창 A_3 를 움직이면서 이진화된 얼굴영상 (B_3)에 에지점 (b)을 하나라도 포함하면 이 창을 모두 에지점으로 결정한다. 이와 같이 결정된 에지 영상에서 얼굴과 배경을 분리하였다. 실험에 사용한 영상은 256×256 크기의 256 밝기 영상이고, 좌표의 (0, 0)점을 좌측 최상위 점으로 정하였다.

얼굴과 배경 및 어깨를 분리하는 방법은 그림 3 (a), (c)와 같이 특징 추출 단계에서 수평 수직 투영시 오류의 요인이 되는 배경과 어깨 부분이 취득된 영상의 중간 아래에 위치하기 때문에 y축의 중간에서 x축 (-x축) 방향으로 한 화소씩 증가하면서 에지의 변화가 일어나는 점을 얼굴의 좌측(우측) 최상위 점(L, R)으로 정했다. 이 점을 기준으로 하여 y축 방향으로 한 화소씩 증가시키면서 그림 2의 추적 우선순위에 따라 연속되는 에지를 얼굴과 배경 및 어깨의 경계선으로 결정한다. 이때 연속되는 에지가 존재하지 않을 경우에는 그림 2의 ③의 방향으로 10화소 옮

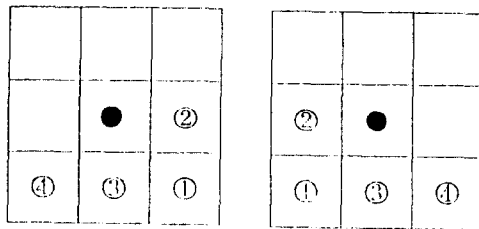


그림 2. 얼굴의 경계선을 찾는 창들 (a) 아래부분 왼쪽 창 (b) 아래부분 오른쪽 창

Fig. 2. Tracking windows of face boundaries (a) bottom-left window, (b) bottom-right window.

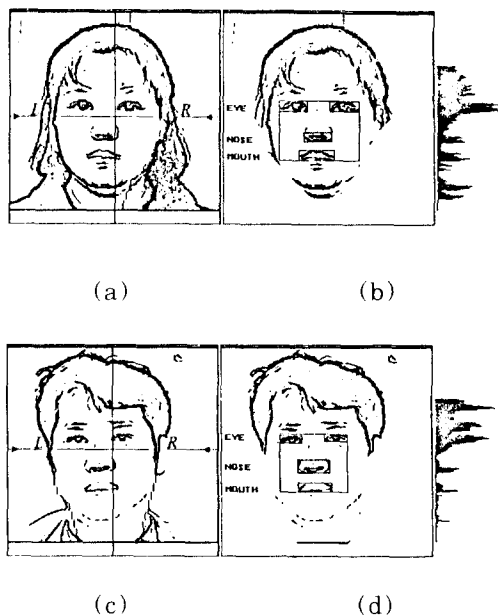


그림 3. 지식에 기초한 배경 분리과 특징추출 결과 (a) 여자영상의 배경분리 (b) (a)의 특징추출과 수평 투영값 (c) 남자영상의 배경분리 (d) (c)의 특징추출과 수평 투영값

Fig. 3. Knowledge-based background separation and feature extraction results (a) background separation of a woman image, (b) feature extraction and horizontal projection value of (a), (c) background separation of a man image, (d) feature extraction and horizontal projection value of (c).

직이며 연속되는 에지를 찾고, 그래도 연속되는 에지가 없을 경우에는 ①의 방향으로 3화소 움직인 후, 위의 과정을 되풀이한다. 이 과정은 x축의 중간지점 또는 y축의 끝지점에 도달할 때까지 계속한다. 그림 2에서 (a)는 중간 아래 왼쪽 부분을 추적하는 창이고, (b)는 중간 아래 오른쪽 부분을 추적하는 창이다. 숫자는 추적시 우선순위를 나타낸다.

Ⅲ. 특징추출 및 기운 영상의 보정

배경과 어깨 부분이 분리된 이진영상에서 각 방향 투영에 의하여 얼굴의 구성성분들에 대한 기초지식을 이용하여 인식에 사용할 특징들을 찾았다. 이때 얼굴 영상 취득시 어느 정도 기울거나 회전된 영상의 경우 오인식의 원인이 되므로 얼굴을 타원체로 가정하여 기울거나 회전된 영상의 특징값들을 보정하였다.

1. 특징추출

배경과 어깨 부분이 분리된 이진영상에서 눈, 코, 그리고 입의 끊어진 에지를 연결할 목적으로 식 (3)의 1D closing 연산을 한 후 가로와 세로 방향으로 투영하여 눈, 코, 그리고 입의 좌표 값을 찾았다.

$$C(A_i, B_i) = [A_i \oplus (-B_i)] \ominus B_i \approx \epsilon [D(A_i, B_i), -B_i] \quad (3)$$

여기서 A_i 은 이진 얼굴 영상이고, B_i 은 1×3 창이다. 그리고 ϵ 는 erosion을 의미하고, D는 dilation을 의미한다.

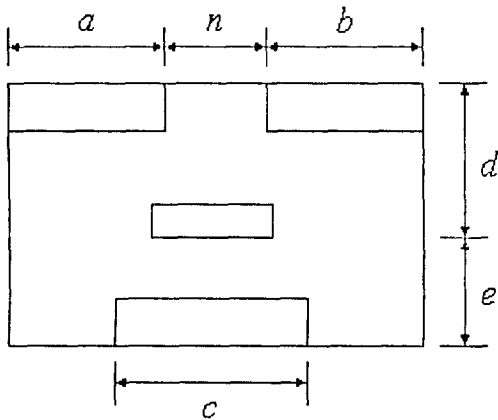


그림 4. 특징값들 $(\frac{a+b}{2n}, \frac{c}{n}, \frac{d}{n}, \frac{e}{n})$
 Fig. 4. Feature values $(\frac{a+b}{2n}, \frac{c}{n}, \frac{d}{n}, \frac{e}{n})$.

그림 4에 보인 바와 같이 정의되는 정규화된 특징값들 (눈의 폭, 입의 폭, 눈에서 코까지의 거리, 그리고 코에서 입까지의 거리)을 찾는 방법은 다음과 같다. 첫번째 단계로 모든 행을 수평방향으로 투영하여 가장 큰 투영값을 갖는 부분을 찾는다. 이 부분은 한 사람당 4장씩 20명의 실험결과 대부분 눈 또는 눈썹과 눈썹에 인접한 머리카락 부분의 y축 값에 해당한다. 이 값을 기준으로 아래쪽으로 형성되는 두개의 최대 수평 투영값을 선택한다. 이 값들이 주어진 지식과 일치하는지 판단한 후 대충의 눈, 코, 그리고 입의 y축 값을 결정한다. 즉 눈이라고 결정된 y축 값과 입이라고 결정된 y축 값 사이에서 코의 y축 값이 눈의 y축 값에 더 근접하여 있으면, 눈의 y축 값은 눈썹 또는 눈썹에 인접한 머리카락 부분의 y축 값이 되므로, 이 투영값을 삭제한 후 다시 눈, 코, 그리고 입의 대략적인 y축 값들을 결정한다. 두번째 단계로 코의 y축 값이 입의 y축 값보다 눈의 y축 값에 더 근접하여 있으면 입의 y축 값은 턱 수염 또는 턱의 값이므로 이 투영값을 삭제한 후 다시 코와 입의 대략적인 y축 값을 결정한다. 세번째 단계로 이렇게 결정된 눈의 y축 값들을 기준으로 ± 8 화소 부분에 대해서 수직 방향으로 투영한 후, 투영값의 존재 여부에 따라 그림 5와 같이 이진화하여 "1"의 부분이 가장 큰 두 부분인 x_1, x_2 부분과 x_3, x_4 부분을 눈으로 결정하였다. 그림 5에서 0과 255는 취득된 영상의 폭을 나타낸다. 이렇게 투영값에 따른 눈, 코, 그리고 입의 위치를 찾는 과정에서 사용한 투영값의 예를 그림 3에 보였고, 이 그림의 투영값에서 흐린 부분은 지식의 검증을 통해 삭제된 부분이다.

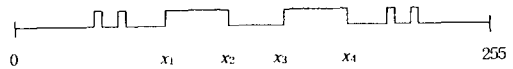


그림 5. 눈 부분 수직 투영값의 이진패턴
 Fig. 5. Binary pattern of a vertical projection across an eye.

이렇게 정한 두눈의 양끝점 (x_1 과 x_4) 사이의 아래 부분에 코와 입이 존재하므로, 대충 결정된 코와 입의 y축 값과 x_1 과 x_4 사이에서 y 8 화소 부분에 대해 수직 방향으로 투영하여 투영값이 존재하는 부분이 가장 큰 부분을 코와 입으로 결정하였다. 마지막 단계로 정확한 눈, 코, 그리고 입의 위치를 찾기 위해 각 부분의 $y \pm 10$ 화소와 x축 값 사이의 영역에서 수평투영하여 가장 에지의 연속이 많은 부분을 각 부분의 정확한 y축 값으로 결정하였다.

2. 기울거나 회전된 영상의 보정

기존의 논문에서 기울거나 영상 크기에 관계없이 인식할 수 있는 방법으로 불변 모멘트를 특징벡터로 이용하는 방법^[10]이 있는데, 이 방법은 한 사람의 취득된 영상에 대해서 여러 변환영상의 실험 셋(set)을 필요로 하는 단점이 있다. 이에 본 논문에서는 얼굴을 구성하는 눈, 코, 그리고 입의 위치관계를 이용하여 인식하므로, 변하지 않는 축을 중심으로 각도를 보정하는 방법을 제안하였다. 또한 미간 사이의 거리로 모든 특징값들을 정규화하여 영상 크기의 변화에 대처하였다. 우선 얼굴을 타원체라고 가정하였다. 즉, 그림 6(a)에서 Y축을 중심으로 회전시키더라도 \overline{AB} 의 길이는 항상 일정하다. Z축을 중심으로 회전되었을 경우 식 (4)에 의해 기울어진 각을 찾는다.

$$\text{Tilt angle} = \tan^{-1} \frac{y_2 - y_1}{x_3 - x_1} \quad (4)$$

여기서 좌표 (x_3, y_2) 는 그림 6(a)에서 보인 것처럼 오른쪽눈의 왼쪽 상단점을 의미하고, 좌표 (x_1, y_1) 는 왼쪽눈의 왼쪽 상단점을 의미한다. 두번째로 Y축을 중심으로 회전한 경우는 그림 6(b)의 입 부분 단면에 의해 입의 양끝점과 눈의 양끝점까지의 길이를 사용하여 식 (5)와 식 (6)에 의해 회전각 $\Delta\theta$ 를 찾는다. 그림 6(b)에서 실선 부분은 정확한 입의 위치이고, 점선 부분은 회전된 영상에서 입의 위치를 나타낸다.

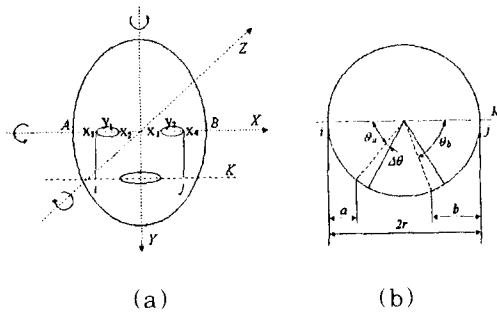


그림 6. 얼굴의 정면도와 입 부분의 단면도
(a) 얼굴의 타원체 모델 (b) 단면도
Fig. 6. Front and slice views of face
(a) elliptic model, (b) slice view.

$$a = r - r \times \cos \theta_a \quad (5)$$

$$b = r - r \times \cos \theta_b$$

여기서 a와 b는 회전된 영상에서 입의 양 끝점과 눈

의 양 끝점 사이의 거리이고, r은 입 부분 원 단면도의 반지름이다. 식 (5)로부터 θ_a 와 θ_b 를 구한 후, 원의 기본적인 성질을 이용하여 회전각 $\Delta\theta$ 를 식 (6)과 같이 구한다.

$$\Delta\theta = \frac{\theta_b - \theta_a}{2} \quad (6)$$

식 (4)와 (6)에 의해 특징값들을 보정하는데 있어서, 취득된 영상의 기운 정도가 10° 이상일 경우는 눈의 수평 투영값이 머리카락의 투영값과 겹치는 경우가 발생하여 보정이 되지않고 오인식의 원인이 된다.

IV. 오류 역전파 알고리즘을 이용한 얼굴인식

본 논문에서는 전처리 단계를 거쳐 추출된 특징에 대해 신경회로망 학습의 하나인 오류 역전파 (error back propagation) 알고리즘을 이용하여 인식하였다. 본 논문에서 사용한 인식알고리즘은 입력층과 중간층 그리고 출력층으로 이루어진 세계층의 다층퍼셉트론을 이용하였다.

1. 오류 역전파 알고리즘의 구조

신경회로망에서 중요한 구성요소는 노드라 불리우는 처리기와 이들 상호간의 연결이다. 그림 7은 오류 역전파 알고리즘의 구조를 보여주고 있다. 노드간의 연결은 방향성 연결선이고 가중치를 가지고 있다. 가중치 W_{ij} 는 노드 i가 노드 j에 미칠 수 있는 영향력의 크기를 표시한다. 양의 가중치는 흥분시키는 작용을 하고, 음의 가중치는 억제하는 작용을 한다. 가중치에 의해 회로망의 동작이 결정된다. 본 논문에서 입력층은 특징추출 단계에서 추출한 4개의 특징값을 입력층으로 사용하였다. 그리고 출력층은 20명을 인식 대상으로 하였기 때문에 20개의 출력 뉴런값을 갖도록 하였다. 그러므로 출력층에서 가장 큰 값의 출력 값을 갖는 뉴런을 인식결과로 삼았다. 또한 입력층과

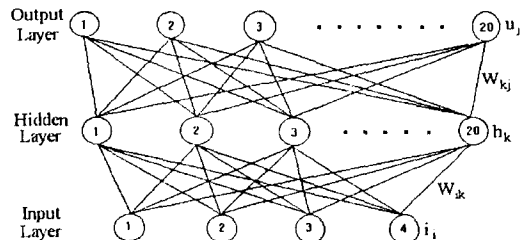


그림 7. 세계층의 신경회로망
Fig. 7. Three-layer neural networks.

출력층 사이에는 입력노드와 출력노드에 속하지 않는 중간노드가 존재하는데 이는 선형 비분리 함수와 같은 복잡한 기능을 수행하는 회로망이 필요하다.

신경회로망에 존재하는 노드는 뇌에서의 신경세포에 해당하며 활성화 값이라 불리는 출력값을 계산하는 작용을 한다. 한 노드의 활성화 값을 그 노드에 직접 연결되어 있는 노드들의 출력값과 해당 연결선의 가중치를 이용하여 계산한다. 전형적인 신경회로망의 모델에서는 다른 노드들로부터의 입력을 연결선의 가중치를 고려하여 더한 후, 이 결과가 어떤 임계치를 초과하면 그 값을 그 노드의 출력값으로 한다.

2. 학습 알고리즘 및 인식

본 연구의 학습은 입력층과 중간층 그리고 출력층으로 구성된 오류 역전파 알고리즘을 이용하였다. 오류 역전파 알고리즘은 출력의 기대치와 실제값의 오차를 감소하는 방향으로 연결 강도를 조절하고 상위층의 오차를 현재의 연결 강도를 가중치로 하여 다음 하위층에 역전파하며, 하위층에서는 이를 근거로 하여 자기층의 연결강도를 조정해 나간다. 그리고 각 노드를 활성화시키기 위한 함수로 식 (7)과 같은 비선형 시그모이드 함수를 사용함으로써 다층 신경회로망의 장점을 살릴 수 있으며, 특히 미분가능한 함수를 사용함으로써 목표 출력값과 실제 출력값의 차이를 줄일 수 있다.

$$f(x) = \frac{1}{(1+e^{-x})} \quad (7)$$

이 신경회로망을 3개층으로 구성할 경우 이론상으로는 어떠한 형태의 패턴도 형성할 수 있으나, 학습시 국소최소 (local minimum)에 빠질 우려가 있고, 학습과정이 너무 오래 걸리고, 또 기억된 패턴의 수정, 추가학습 등이 불가능하다는 단점이 있다. 그러나 이 신경회로망은 구현이 쉽고 학습방법이 용이하기 때문에 넓게 응용되고 있다.

오류 역전파 알고리즘의 학습은 입력단에서 임의의 가중치를 이용하여 학습이 시작되는데 출력층에서 j 번째 출력 뉴런 σ 는 식 (8)과 같이 정의된다.

$$\delta_j = \mu_j (1 - \mu_j) (t_j - \mu_j) \quad (8)$$

식 (8)에서 μ 는 출력단에서의 실제 출력을 나타내고, t 는 목표값을 나타낸다.

그리고 가중치 W_{kj} 는 식 (9)에 의해 변화한다.

$$\Delta W_{kj}(n) = \eta \delta_j h_k + \alpha \Delta W_{kj}(n-1) \quad (9)$$

여기서 h_k 는 중간 노드의 출력값을 나타내고, n 과 $n-1$ 는 현재의 가중치와 전단계의 가중치를 나타내는 색인 (index)이다. 이 가중치는 입력과 중간층의 연결 노드에도 사용되었다. 그리고 두 개의 파라미터 η 와 α 는 학습률과 관성항으로 학습률은 값이 클수록 가중치의 변화량이 큰 값을 갖는다. 학습률이 크면 일반적으로 학습이 빨리 이루어지지만 수렴되지 않고 진동이 일어나는 경우가 발생할 수 있으므로 각 응용에 맞는 학습률을 선택해야 한다. 그리고 관성항은 진동을 막으면서 전단계에서의 가중치 변화량을 다음 단계에서 유지하도록 하는 역할을 한다. 따라서 관성항 값의 선택도 학습에 있어 중요하다. 일반적으로 학습시 학습률은 크게 하고 관성항은 작게 할수록 학습시간이 줄어든다. 본 연구에서는 고정된 학습률과 관성항을 사용하는 기존의 방법과는 달리 학습과정에서 학습률과 관성항을 변화시키면서 학습시간과 수렴 두가지 모두 만족하도록 하였다. 그 결과 인식 결과에는 변화가 없으면서 빨리 학습할 수 있는 알고리즘을 구현할 수 있었다.

학습시 중요한 또 하나의 파라미터는 중간층의 뉴런 개수로서 뉴런의 개수가 너무 적으면 학습이 안되며, 많을 경우 학습에 소요되는 시간이 많이 걸린다. 그러므로 각 응용에 알맞는 중간층의 뉴런 수가 필요하다. 본 연구에서는 중간층의 뉴런 수를 학습률이나 관성항과 같이 자동적으로 변화시키지 않고, 중간층의 뉴런 수를 달리하면서 실험한 결과 가장 좋은 결과를 보이는 중간 뉴런 수를 본 연구의 중간 뉴런 수로 정하였다.

인식시 본 연구에서는 그림 7과 같은 세계층으로 이루어진 다층퍼셉트론 신경회로망을 이용하여 남자와 여자로 이루어진 20명에 대해 인식하였고, 입력단에서는 4개의 뉴런을 사용하였는데, 이는 사람의 얼굴에서 중요한 특징인 눈 사이의 거리, 눈과 코까지의 거리, 코에서 입까지의 거리, 그리고 입의 크기 등 4개의 특징이다. 그리고 출력단에서는 20명을 인식하기 위해 20개의 출력 뉴런을 사용하여 임계치를 넘는 출력값중에서 가장 큰 값을 갖는 뉴런을 인식결과로 선택하였다. 또한 인식과 학습에 있어 중요한 역할을 하는 중간층의 뉴런 수를 실험에 의해 20개로 정하여 인식하였다.

V. 실험결과 및 분석

실험은 PC 486을 이용하여 C언어로 구현하였으며, CCD 카메라를 이용한 256×256 영상으로 남자 14명과 여자 6명에 대해 각각 4장씩 취득한 후, 3

장은 학습에 사용하였으며, 나머지 1장은 실험 데이터로 사용하였다. 이때 실험 데이터로 사용한 얼굴영상을 그림 8에 보였다.



그림 8. 실험에 사용한 영상들
Fig. 8. Test images.

전처리 단계에서는 중앙값 필터를 이용하여 잡음을 제거하였으며, 네 방향 Sobel 연산자를 이용하여 얼굴과 배경을 분리하였다. 또한 입력시 얼굴의 기울어짐도 보정하여 특징추출에서의 오류를 줄였다. 특징추출단계에서는 수평, 수직 투영을 이용하여 얼굴부분에서 중요한 특징인 눈, 코, 입을 추출하였으며, 눈 사이의 거리, 눈과 코의 거리, 코와 입의 거리, 그리고 입의 크기를 특징으로 추출하여 인식단으로 넘겨주었다. 인식단에서는 특징추출단계에서 넘겨온 특징을 가지고 오류 역전파 알고리즘을 이용하여 학습하고 인식하였다.

학습단계에서 학습시간을 줄이기 위해 학습률과 관성항을 수렴 정도에 따라 변화시키면서 학습하였다. 즉 초기 파라미터와 변화될 파라미터 값은 사용자가 학습 초기에 입력하고, 학습시 전단계와의 수렴차이가 클 경우는 학습을 빨리하기 위해 학습항의 크기를 크게 하고 관성항의 크기는 작게 하였다. 그러나 수렴차이가 적을 경우에는 학습속도는 느리지만 국소최소를 벗어나도록 학습항의 크기는 작게 그리고 관성항의 값은 크게 선택하여 수렴속도와 국소최소 문제

를 보완하도록 하였다. 그리고 중간노드의 뉴런 수는 실험에 의해 20개로 정하였다.

학습률과 관성항을 고정시킨 후 학습한 경우와 변화시키면서 실험한 결과, 변화시키면서 학습한 경우 고정시켜 학습한 경우보다 약 20% 정도의 학습시간을 감축할 수 있었다. 또한 본 연구에서 중간 노드의 뉴런 수를 8개부터 20개까지 증가시키면서 학습시의 수렴과 인식결과를 관찰하였다. 이때 뉴런 수가 8개에서 10개까지는 수렴은 되었으나 인식률에서 100%가 되지 않았으나, 뉴런 수가 20개인 경우 수렴도 잘 되었으며, 인식률도 학습데이터 및 테스트 데이터에 대해 100%의 인식률을 보였다. 표 1은 중간노드의 뉴런 수에 따른 인식결과를 보여주고 있다.

표 1. 인식률 비교

Table 1. Comparison of recognition rate.

인식 방법	학습 데이터 인식률	테스트 데이터 인식률
유클리디안	83%	75%
CP	93%	85%
재안한 BP 방법	8	83%
	10	93%
	15	100%
	20	100%

본 논문에서는 기존의 방법과의 비교를 위해 많이 사용된 유클리디안 방법과 신경회로망의 알고리즘인 CP를 이용하여 인식을 수행하였다. 유클리디안 방법의 경우 3개 영상에 대한 특징값들의 평균벡터 $X = \{X_1, X_2, X_3, X_4\}$ 라 하고, 나머지 한 영상의 특징값 추출벡터 $Y = \{Y_1, Y_2, Y_3, Y_4\}$ 라 할 때, 유클리디안 거리는 식 (10)과 같이 정의되며, d^2 값이 가장 작은 영상을 인식 결과로 정하였다.

$$d^2 = \sum_{i=1}^4 (X_i - Y_i)^2 \tag{10}$$

이와같이 인식을 수행한 결과 20명에 대한 인식결과가 표 1에서 보여 주듯이 75% 이었다. 그러므로 유클리디안 방법을 사용할 경우 기존의 많은 논문과 같이 특징을 좀 더 많이 사용하여야 할 것이다. 그리고 CP의 경우 입력층으로는 본 논문에서 추출한 4개의 특징에 대해 정규화를 행한 특징값들을 사용하였으며, 경쟁층에서는 승자 뉴런 (winner neuron)을 구하고 출력층에서는 20명에 대한 출력값을 구하였다. CP 알고리즘의 경우도 BP와 같이 20명 각각에 대해 4장씩의 영상중에서 3개는 학습에 사용하였으며, 나

머지 1장은 실험 영상으로 사용하였다. CP 알고리즘의 경우 학습속도는 BP에 비해 빠르나 인식률은 85%로 BP 알고리즘에 비해 성능이 떨어짐을 알 수 있었다. 이는 CP 알고리즘의 경우도 유클리디안 알고리즘과 같이 입력으로 사용되는 특징의 개수를 좀 더 늘리면 좋은 결과를 얻을 수 있을 것이다.

표 2. 학습시 고정된 파라미터와 가변 파라미터 간의 경우 필요한 반복회수

Table 2. Number of iteration required with fixed and variiable parameters in learning.

파라미터 값		iteration 수
학습률 (η)	관성항 (α)	
0.9	0.1	20400
0.5	0.5	12140
0.1	0.9	10700
가변 파라미터		8200

표 2는 학습시 고정된 파라미터와 본 논문에서 사용한 가변 파라미터 간의 수렴하는데 걸리는 iteration 수를 보여주고 있다. 표 2에서 보듯이 고정된 파라미터를 사용한 경우보다 가변 파라미터를 사용한 경우가 학습시 빨리 수렴되는 것을 알 수 있다. 실험결과에서 알 수 있듯이 본 연구에서 사용한 4개의 특징으로 BP 알고리즘을 사용했을 경우 가장 좋은 성능을 나타내었으며, 다른 알고리즘에 비해 학습속도가 떨어지는 것을 보완한다면 좀 더 나은 인식 알고리즘이 될 것이다. 그리고 본 논문에서 중간층의 뉴런 개수를 실험적으로 구했으나, 학습 과정에서 가변적으로 적용시킨다면 좀 더 적합한 중간층의 뉴런 개수를 구할 수 있을 것이다.

본 논문에서는 20명에 대해 4개의 적은 특징값으로 인식이 가능하였으며, 본 연구의 알고리즘 성능을 비교하기 위해 유클리디안 알고리즘과 신경회로망의 CP 알고리즘을 이용하여 인식한 결과와 비교하였다. 또한 BP 알고리즘의 단점으로 지적되고 있는 학습시간 지연과 국부 최소값의 문제를 해결하기 위해 가변 학습률과 관성항을 이용함으로써 두가지 문제점을 보완하였다. 그러나 최적의 중간층 뉴런 개수를 얻는 문제와 좀 더 많은 사람을 인식하는 문제, 그리고 사람 수가 많아짐에 따른 좀 더 효율적인 특징추출이 필요할 것이다. 또한 데이터베이스에 저장되어 있지

않은 사람을 인식할 경우에는 데이터베이스에 저장된 다른사람으로 인식할 수도 있으므로 보안 시스템 등의 응용에서는 ID 카드 또는 지문을 이용한 인식방법 등과 병행할 필요가 있다.

VI. 결론

본 논문에서는 사람의 얼굴에서 중요한 특징이면서 외부 환경에 영향이 적은 눈, 코, 입의 위치 관계를 이용한 얼굴인식 알고리즘을 구현하였다. 전처리 단계에서는 이진화된 영상에서 눈, 코, 그리고 입의 위치를 수직과 수평투영을 이용하여 찾아내었으며, 각 부분의 크기와 위치를 4개의 특징값으로 나타내어 인식단으로 넘겨 주었다. 그리고 인식단에서는 현재 많이 사용되고 있는 신경회로망 알고리즘의 하나인 오프 역전파 알고리즘을 사용하여 인식하였다. 특히 학습시에 학습시간이 많이 걸리거나 국부 최소값에 빠진다는 단점을 보완하기 위해 학습률과 관성항을 목표값과 실제 출력값 차이에 따라 자동으로 변하는 알고리즘을 사용함으로써 고정된 학습률과 관성항을 사용한 알고리즘에 비해 약 20%의 학습시간을 단축할 수 있었다. 그리고 중간 노드의 뉴런 수는 실험에 의해 20개가 최적임을 알 수 있었다.

본 논문에서는 특징추출 단계에서 추출한 효율적인 네개의 특징으로 남자 14명과 여자 6명에 대해 오류 역전파 알고리즘으로 학습과 인식을 행한 결과 학습에 사용된 60장의 영상과 테스트에 사용된 20장의 영상에 대해 정확히 인식할 수 있었다.

앞으로의 과제는 실험에 의한 중간층의 뉴런 수 결정보다는 자동적으로 결정하도록 하는 방법과 연령의 변화와 주변환경에 덜 민감한 특징을 추출하여 다양한 연령층의 얼굴 인식과 실생활에 사용할 수 있는 안정된 얼굴인식 알고리즘을 구현하는 것이다. 또한 구성성분들인 눈과 입의 형태 등을 특징으로 사용하여 더 많은 사람을 대상으로 하는 인식 시스템을 구현하는 것이다.

參考文獻

[1] A. Samal and P. A. Iyengar, "Automatic recognition and analysis of human faces and facial expressions: A survey," *Pattern Recognition*, vol. 25, no. 1, pp. 65-77, Jan. 1992.

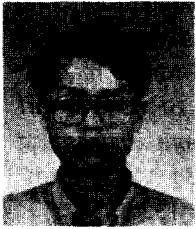
[2] G. J. Kaufman, Jr. and K. J. Breeding, "The automatic recognition

- of human faces from profile silhouettes." *IEEE Trans. Systems, Man, Cybernetics*, vol. SMC-6, no. 2, pp. 113-121, Feb. 1976.
- [3] L. D. Harmon, M. K. Khan, R. Lasch, and P. F. Ramig, "Machine identification of human faces." *Pattern Recognition*, vol. 13, no. 2, pp. 97-110, 1981.
- [4] C. J. Wu and J. S. Huang, "Human face profile recognition by computer." *Pattern Recognition*, vol. 23, no. 3, pp. 255-259, 1990.
- [5] T. Aibara, K. Ohue, and Y. Oshita, "Human face profile recognition by a P Fourier descriptor." *Optical Engineering*, vol. 32, no. 4, pp. 861-863, Apr. 1993.
- [6] M. D. Kelly, "Edge detection in pictures by computer using planning," in *Machine Intelligence*, B. Meltzer and D. Michie, Eds., Edinburgh University Press, Edinburgh, pp. 397-409, 1971.
- [7] I. Craw, H. Ellis, and J. R. Lishman, "Automatic extraction of face features," *Pattern Recognition Letters*, vol. 5, no. 2, pp. 183-187, Feb. 1987.
- [8] 지 은미, 최 형일, 이 철희, "문제영역에 관한 지식을 이용한 얼굴인식," 정보과학 회 논문지, 제 17권, 제 6호, pp. 666-675, 1990년 11월.
- [9] K. M. Lim and K. W. Oh, "A face recognition system using fuzzy logic," in *Proc. First Korean Japan Joint Conf. Computer Vision*, pp. 506-509, Seoul, Korea, Oct. 1991.
- [10] 박 용근, 김 영일, "ART와 다층 퍼셉트론을 이용한 화상인식 시스템의 성능 분석," 전자공학회 하계종합학술대회 논문집, 제 15권, 제 1호, pp. 492-496, 1992년 6월.
- [11] K. Matsuda, T. Kageyama, and T. Aibara, "Application of the second order statistics for the recognition of human faces," in *Proc. Int. Conf. Image Processing*, Singapore, pp. 543-546, Sept. 1992.
- [12] A. L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature extraction from faces using deformable templates," *Int. Journ. Computer Vision*, vol. 8, no. 2, pp. 99-112, Aug. 1992.
- [13] C. L. Huang and C. W. Chen, "Human facial feature extraction for face interpretation and recognition," *Pattern Recognition*, vol. 25, no. 12, pp. 1435-1444, Dec. 1992.
- [14] H. Agawa, G. Xu, Y. Nagashima, and F. Kishino, "Image analysis for face modeling and facial image reconstruction," in *Proc. SPIE Visual Communication and Image Processing 90*, vol. 1360, pp. 1184-1197, Lausanne, Switzerland, Oct. 1990.
- [15] R. H. Nielsen, "Counterpropagation networks," in *Pattern Recognition by Self Organizing Neural Networks*, G. A. Carpenter and S. Grossberg, Eds., Cambridge, MA: MIT Press, pp. 264-276, 1991.
- [16] D. J. Burr, "Experiments on neural net recognition of spoken and written text," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. ASSP 36, no. 7, pp. 1162-1168, July 1988.
- [17] Y. K. Ham, S. Y. Lee, and R. H. Park, "Knowledge based face recognition using neural networks," in *Proc. 1993 Int. Joint Conf. Neural Networks*, vol. 3, pp. 2125-2128, Nagoya, Japan, Oct. 1993.
- [18] G. S. Robinson, "Edge detection by compass gradient masks," *Computer Graphics Image Processing*, vol. 6, no. 5, pp. 492-501, Oct. 1977.
- [19] M. Kato, I. So, Y. Hishinuma, O. Nakamura, and T. Minami, "Description and synthesis of facial expression based on isodensity maps," in *Visual Computing*, T. L. Kunii, Ed., Springer Verlag, pp. 39-56, 1992.
- [20] H. Harasaki, M. Yano, and T.

Nishitani, "Background separation/filtering for videophone applications," in *Proc. Int. Conf. Acoustics, Speech, Signal Processing 90*, vol. 4, pp. 1981-1984, New Mexico, USA, Apr. 1990.

[21] C. R. Giardina and E. R. Dougherty, *Morphological Methods in Image and Signal Processing*, Prentice Hall, Englewood Cliffs, New Jersey, pp. 1-33, 1988.

著 者 紹 介



李 常 榮(正會員)

1968年 1月 24日生. 1991年 2月 한국과학기술원 과학기술대학 전기 및 전자공학과 졸업(공학사). 1992年 8月 ~ 현재 서강대학교 대학원 전자공학과(석사과정). 주 관심 분야는 영상처리 등임.

咸 永 國(正會員) 第 29卷 B編 第 10號 參照

1992年 서강대학교 대학원 전자공학과 졸업(공학석사). 현재 동대학원 박사과정

朴 來 弘(正會員) 第 23卷 第 6號 參照

현재 서강대학교 전자공학과 교수