

세그멘테이션에 의한 특징공간과 영상벡터를 이용한 얼굴인식

Face Recognition using the Feature Space and the Image Vector

김 선 종
(Seon-Jong Kim)

Abstract : This paper proposes a face recognition method using feature spaces and image vectors in the image plane. We obtain the 2-D feature space using the self-organizing map which has two inputs from the axis of the given image. The image vector consists of its weights and the average gray levels in the feature space. Also, we can reconstruct a normalized face by using the image vector having no connection with the size of the given face image. In the proposed method, each face is recognized with the best match of the feature spaces and the maximum match of the normally retrieval face images, respectively. For enhancing recognition rates, our method combines the two recognition methods by the feature spaces and the retrieval images. Simulations are conducted on the ORL(Olivetti Research Laboratory) images of 40 persons, in which each person has 10 facial images, and the result shows 100% recognition and 14.5% rejection rates for the 20x20 feature sizes and the 24x28 retrieval image size.

Keywords: face recognition, feature space, image vector, ORL images

I. 서론

일반적으로 얼굴과 같이 특정정보가 내포된 영상을 인식 등에 적용할 경우에는 원래의 영상을 축소하여 이용한다. 축소된 영상으로 처리하면 그 처리 시간을 줄일 수 있을 뿐만 아니라 메모리를 줄일 수 있기 때문이다. 그러나 축소된 영상을 인식에 사용함으로써 원래의 영상이 가지고 있었던 정보가 사라지거나 약해진다. 이는 영상내의 특정정보와 무관하게 전체 영상에 대한 복구에러가 최소가 되도록 축소하기 때문이다. 특히, 인식과 같은 분야에서 이러한 축소 방법을 적용하였을 경우에는 사라진 정보로 인하여 오인식을 초래할 수 있다. 즉, 영상 인식 시스템은 원래의 영상 정보를 최대한 활용하여야 인식 시스템의 성능을 좀 더 높일 수 있다. 기존의 인식분야에서 쓰이는 영상축소 방법에는 KLT(Karhunen-Loève Transform) [1]-[3], SOM(Self Organizing Map) [4], 신경망[5]-[7] 등이 사용되고 있다.

Brunelli 등[8]은 얼굴 정보에서 눈, 코, 입 및 턱에서 고유한 특징요소를 추출하여 이를 인식에 사용하였다. 이 특징추출 방법은 잡음이나 조명 등의 영향이 크며, 경우에 따라서는 특징추출이 어려운 경우도 있다. Turk 등[1]은 KLT를 이용하여 고유얼굴(eigenface)을 추출하고, 추출된 고유얼굴을 이용하여 얼굴을 인식하였으며, 조명이 일정하고 얼굴의 위치가 크게 변화되지 않은 얼굴영상에서 빠르면서도 우수한 인식성능을 가지는 시스템을 제안하였다. Lawance 등[4]은 국소(local) KLT 및 SOM을 이용하여 영상을 국소적으로 축소하고, 축소된 영상을 convolution 망[9]을 이용하여 인식하였다. 국소

적으로 축소된 영상은 전체영상에 대해 축소된 Turk 등의 고유얼굴보다 KLT를 적용하는 데에 복구 에러를 더 줄일 수 있었다. 또한 축소된 얼굴영상은 convolution 망을 이용하여 어느 정도의 이동, 크기 및 회전에 불변인식할 수 있었지만, 패턴의 수가 많아질수록 학습에 요구되는 시간이 많으며, 입력영상의 크기 및 축소된 영상의 크기가 달라지면 다른 망을 구성하여 처음부터 학습하여야 하는 문제점이 있다[10].

이와 같이 기존의 방법에서는 축소된 영상을 인식에 사용함으로써, 원래의 영상이 가지고 있는 정보를 잃어버릴 수 있으며, 이는 인식 성능에도 영향이 있다. 따라서 원래의 영상정보는 그대로 유지하면서도 축소된 영상을 얻을 수 있는 방법과 오인식을 좀 더 줄이는 방안도 강구되어야 한다.

본 논문에서는 영상내의 특정한 정보를 갖는 사람의 얼굴에 대하여 특징공간을 추출하고, 추출된 특징공간과 평균명도를 영상벡터로 하여 임의의 크기를 갖는 축소된 얼굴영상을 만들어 이를 인식하는 방법을 제안한다. 제안된 방법에서는 원래의 얼굴로부터 특징공간을 형성시킴으로써, 좀 더 정확한 정보를 갖는 특징을 얻을 수 있으며, 특징공간의 크기만 일정하다면 얼굴의 크기에 관계없이 특징의 수를 일정하게 할 수 있어서 입력 영상의 크기에 따른 인식 시스템을 변경하지 않고 그대로 사용할 수 있다. 또한 특징공간과 복구된 영상과의 성능을 조합하여 인식함으로써 인식성능을 높였으며, 인식이 확실치 않은 경우에는 인식불가(rejection)로 처리하여 오인식을 줄였다. 제안된 인식 방법의 타당성을 확인하기 위하여 40사람에 대한 영상 400개의 ORL(Olivetti Research Laboratory) 얼굴영상 데이터를 가지고 실험

하여 성능을 조사하였다.

II. 얼굴 벡터와 복구된 영상에 의한 얼굴인식 시스템의 제안

얼굴인식 방법의 기본적인 처리과정은 얼굴 영상으로부터 계산량을 줄이고, 처리를 간단하게 하기 위하여 차원을 축소하는 영상축소과정부터 시작한다. 그리고 축소된 영상에 대하여 특징점을 추출하고, 이를 학습하여 인식에 적용한다. Lawance 등[4]은 차원을 축소하는 과정에서 전체영상에 대한 정보가 아닌 국소 정보에 의하여 영상을 축소하였다. 이는 전체정보에 대한 영상을 축소하는 것보다 용이하면서도 원래의 영상으로 복구할 때의 에러를 줄일 수 있기 때문이다. 본 논문에서는 이러한 기존의 방법과 달리 원래의 영상으로부터 특징점을 추출한 후, 이로부터 축소된 영상을 얻는 방법을 제안한다. 그림 1은 기존의 인식 시스템과 제안된 인식 시스템을 비교한 것이다. 기존의 인식 방법에서는 영상을 축소한 후, 축소된 영상으로부터 특징점을 추출하지만, 제안된 방법에서는 원 영상으로부터 특징점을 추출한 후, 얻어진 특징점을 이용하여 차원을 축소시켜 축소영상을 얻는다. 따라서 제안된 인식 시스템은 축소되지 않은 원래의 영상으로부터 특징을 추출함으로써 좀 더 정확한 정보를 갖는 특징점을 얻을 수 있다.

1. 세그멘테이션을 위한 전처리

제안된 얼굴 인식 방법에서 특징공간을 추출하기 위하여, 먼저 영상을 세그멘테이션한다. 본 논문에서는 SOM을 이용하여 세그멘테이션하였다. 이는 Kohonen 학습[11]을 통하여 빠르면서도 정확하게 영상내의 특징을 추출하기 위해서이다. 따라서 입력영상 I로부터 SOM 입력 패턴을 만들기 위한 전처리를 수행하여야 한다. SOM 입력 패턴은 2차원으로 x와 y좌표로 하였다. 영상에서 한 픽셀 p(x,y,g(x,y))는 각각 x좌표, y좌표 및 (x,y) 픽셀의 명도(gray level)가 된다. 이때 x,y좌표는 Kohonen 학습을 위하여 [0,1]로 정규화(normalization)하였다. 이렇게 하면 전체 nI × mI 크기의 영상에서 nI × mI개의 입력 패턴이 발생하며, 영상의 크기가 커질수록 입력패턴의 개수가 많아져서 계산시간이 늘어난다. 그러나 얼굴영상의 크기가 커지더라도 얼굴영상내의 특징요소는 거의 일정함으로 패턴 수를 줄이기

위하여 본 논문에서는 에지영상을 찾고, 찾아진 에지영상의 크기를 양자화(quantization)하여 양자화된 에지 영상 E를 얻는다. 이렇게 하면 전체 얼굴 영상에 대한 특징을 빠르게 찾을 수 있으면서도 공간 정보를 상실하지 않는다. 그리고 nI × mI개의 입력 패턴의 개수가 에지가 있는 개수로 작아진다.

2. 특징공간 추출 및 영상 벡터

전 단계에서 얻어진 입력 패턴을 가지고 nf × mf의 2차원 특징공간을 얻기 위하여 Kohonen 학습을 수행한다. Kohonen 학습[11]은 다음과 같다.

1) 임의의 i번째 픽셀 P_i에서 정규화된 입력 s_{xi}, s_{yi}를 얻는다. 즉,

$$\begin{aligned} s_{xi} &= x_i/N_x \\ s_{yi} &= y_i/N_y \end{aligned} \tag{1}$$

2) 입력패턴 si로부터 출력유닛 k,l의 값 V_{kl}와 승리유닛의 값 V_{kl}^{*}를 각각

$$V_{kl}(t) = \sum_k \sum_l \| s_i(t) - W_{kl}(t) \| \tag{2}$$

$$V_{kl}^*(t) = \min V_{kl}(t) \tag{3}$$

에 의해 구한다.

3) 승리유닛을 중심으로 이웃함수 범위(neighborhood) 내에 있는 모든 출력 유닛에 연결된 웨이트를 다음과 같이 학습된다. 즉,

$$\Delta W_{kl}(t+1) = \alpha \beta (s_i(t) - W_{kl}(t)) \tag{4}$$

$$W_{kl}(t+1) = W_{kl}(t) + \Delta W_{kl}(t+1) \tag{5}$$

이다. 이 때 W_{kl}는 k번째 입력 유닛과 l번째 출력 유닛과 연결된 웨이트를, α는 학습률을, 그리고 β는 각 픽셀에서 양자화된 명도의 크기 g(x,y)를 각각 나타낸다.

4) 모든 입력 패턴에 대하여 1)-3)을 반복 수행한다. 전체 입력패턴에 대한 학습이 모두 끝났으면 학습범위와 학습률을 줄여 1)-4)을 반복한다.

Kohonen 학습이 끝나면, 각 출력 유닛에 연결된 웨이트 W_{ij}의 값이 특징공간이 된다. 이 특징공간은 얼굴 영상의 크기와 관계없이 일정하게 분포된다.

이 때 승리 유닛의 위치가 일정치 않고 8가지의 다른 경우로 발생하는데, 이를 일정하게 하기 위하여 초기 웨이트를 구성하였다. 이렇게 하면 출력 유닛의 위치가 크게 변하지 않을 만큼 초기 이웃함수 범위를 줄일 수 있어서 시스템의 처리시간을 줄일 수 있다.

3. 영상벡터를 이용한 임의의 크기로 얼굴 영상 복원

임의의 크기로 얼굴을 복구하기 위하여는 특징공간 W_{ij}뿐만 아니라 명도 영상 G_{ij}가 있어야 한다. 특징공간 내의 명도는 일정하다고 하고, 원 영상으로부터 특징공간 내에 위치한 픽셀 명도의 평균으로 하였다. 따라서 하나의 얼굴 영상에 대한 영상 벡터는 특징공간 W_{ij}의 nf × mf × 2(차원)개의 좌표와 명도 영상 G_{ij}에 대한 특징점인 (nf-1) × (mf-1)개의 명도로 나타난다.

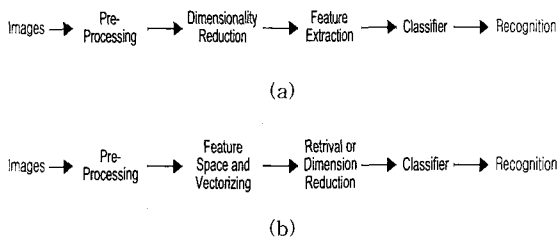


그림 1. 얼굴인식 시스템에서 (a) 기존의 방법과, (b) 제안된 방법의 블록도.
Fig. 1. Block diagram of (a) a conventional and, (b) the proposed face recognition system.

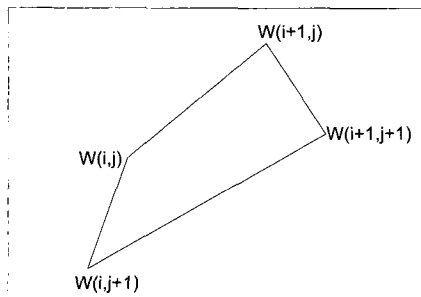


그림 2. 특징공간을 위한 세그멘테이션.
Fig. 2. Segmentation for the feature space.

따라서 특징공간인 SOM 웨이트 W_{ij} 와 명도 G_{ij} 는 영상 벡터를 만들고, 이를 이용하여 축소 및 확대가 가능하다. 즉, 고정된 크기의 영상벡터에 의해 임의의 크기에 대한 얼굴 영상 R 을 만들 수 있다. 그림 2와 같이 특징공간에서 임의의 4개의 좌표를 갖는 사각형을 조절하고 사각형 내에 있는 모든 픽셀을 동일한 명도를 갖도록 하여 영상 복원이 가능하다. 이는 복잡한 영상에 대한 세그멘테이션 결과를 SOM의 웨이트에 의해 사각형으로 근사화된 결과이다.

4. 제안된 얼굴 인식 시스템

먼저, 제안된 시스템은 모든 학습영상에 대해 영상벡터, 즉 특징공간 F 와 평균명도 영상 G , 그리고, 일정한 크기인 $N_{Rx} \times N_{Ry}$ 의 학습영상 RL 을 형성하여 데이터베이스를 만들어 놓는다. 데이터베이스를 구성한 후, 제안된 얼굴 인식 시스템의 전체적인 처리 순서는 다음과 같다.

- 1) 주어진 입력 얼굴영상 I 에 대한 에지영상을 구한다.
- 2) 에지영상을 양자화하여 E 를 구한다. 이는 특징공간을 만들기 위한 입력패턴의 수를 줄이기 위해서이다.
- 3) 구하여진 양자화 영상 E 로부터 2차원 위치벡터 si 와 양자화 값을 갖는 입력패턴 $g(x,y)$ 를 구한다.
- 4) 나열된 2차원 위치벡터와 명도를 입력패턴으로 하여 SOM 학습을 이용하여 특징공간 W_{ij} 을 구성한 후, 이를 이용하여 복구할 명도영상 G_{ij} 을 얻어 얼굴 영상벡터를 얻는다. 특징공간 추출시 사전정보나 초기값을 이용하여 처리시간을 빠르게 한다.
- 5) 모든 학습 영상에 대한 특징공간과의 유클리드 거리 Dw 를 구한다. 즉,

$$D_w^L = \sum_k \sum_l^{N_x N_y} \|W_{kl}^L - W_{kl}\| \quad (6)$$

이다. 이 때 L 은 L 번째 학습영상을 의미하며, N_{Fx} , N_{Fy} 는 각각 특징공간의 크기를 의미한다. 한다. 학습영상에 대한 Dw 중에서 최소의 값으로 나열된 학습 얼굴영상 (p_1, p_2, \dots, p_p) 을 p 개 구한다.

6) 구하여진 영상의 특징벡터를 이용하여 원하는 크기로 영상을 복원하여 RI 을 구한다.

7) 복원된 얼굴 영상에 대한 유클리드 거리 D_R 을 구한다. 즉,

$$D_R^L = \sum_i \sum_j^{N_x N_y} \|R_{ij}^L - R_{ij}^L\| \quad (7)$$

이다. N_{Rx}, N_{Ry} 는 각각 특징공간의 크기를 의미한다. 한다. 따라서 D_R 이 최소인 얼굴을 가지는 사람으로 인식된다. 즉,

$$q = \min D_R^L \quad (8)$$

8) 따라서 제안된 얼굴인식 시스템은 인식된 얼굴이 특징공간의 최소 얼굴영상 p 개 내에 포함되었으면 인식하고, 그렇지 않으면 인식불가로 처리한다. 즉,

$$q \in (p_1, p_2, \dots, p_p) \quad (9)$$

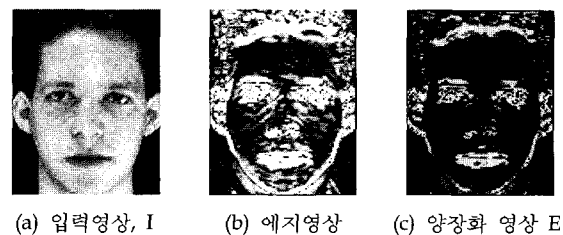
이면 인식된 것으로 가정한다.

III. 시뮬레이션 및 결과 고찰

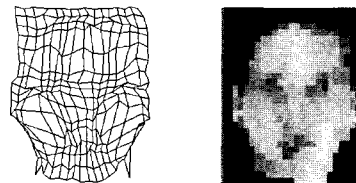
본 논문에서는 1992-94년 동안에 ORL에서 만든 얼굴영상을 사용하였다[4]. 모두 40명에 대한 얼굴영상을 가지며, 각 사람마다 10개의 영상을 가지고 있다. 동일한 사람에 대한 얼굴영상은 정면에서 20°까지 변화되었으며,



그림 3. ORL 데이터베이스의 얼굴영상의 예.
Fig. 3. Example of the ORL faces.



(a) 입력영상, I (b) 에지영상 (c) 양자화 영상 E



(d) 특징공간, F (e) 복구영상, R

그림 4. 제안된 인식 시스템의 처리결과 예.

Fig. 4. Results of the proposed processing; (a) original image, (b) edge image, (c) quantization image, (d) feature space, and (e) retrieval face image.

또한 웃는 얼굴, 안경 쓴 얼굴 등 변화가 많은 영상으로 구성되어 있다. 그리고 얼굴의 크기도 10%까지 변화가 있으며, 이들 모두는 256 명도영상을 갖는 92×112의 크기를 가지고 있다. 그 예로 한 사람에 대한 10개의 얼굴 영상을 그림 3에 나타내었다.

그림 4는 임의의 한 사람에 대한 92×112 크기의 입력영상을 제안된 인식 시스템에 따라 예지영상, 4 레벨 양자화 영상, 16×16 크기의 SOM학습 후의 특징공간을 각각 도시한 것이고, 마지막 영상은 특징공간에 의하여 N_{Rx} 와 N_{Ry} 가 20×20 크기로 축소된 영상이다. 제안된 인식 방법은 특징공간에 의해 인식 결과의 p 개의 리스트를 작성한 후, 인식 결과는 복구영상으로 인식된 결과로 인식한다. 이 인식된 결과가 특징공간에 의한 인식 리스트 p 개 내에 존재하면 인식으로 간주하고, p 개 내에 존재하지 않으면 인식불가로 간주한다. 전체 200개의 인식 대상 영상 중 복구영상 R 에 의해 인식이 성공한 경우는 특징공간에 의한 인식 리스트 p 개 내에 존재하는 경우로 제한하였고, p 개 내에 속하지 않은 경우에는 인식불가로 처리하였다. 이는 얼굴영상의 골격인 특징공간과 그 내부 명도값에 의한 복구영상의 조합으로 인식하게 하여 인식 성능을 좀 더 높일 수 있다.

그림 5는 임의의 한 입력 영상에 대하여 4 개의 각 실험에 의한 축소된 영상 R 이다. 그림 (a)와 (b)는 특징공간의 크기 N_{Fx} 와 N_{Fy} 를 12×12로, (c)와 (d)는 20×20으로 하였을 경우이고, 복구영상의 크기 N_{Rx} 와 N_{Ry} 는 (a)와 (c)는 16×16, (b)와 (d)는 24×28로 하였을 경우이다. 그림에서 특징공간의 크기가 클수록 특징벡터의 수가 많아서 정밀하게 복원됨을 알 수 있다. 즉, 제안된 인식 방법에서 사용된 축소 방법은 차원이 높은 얼굴 형태의 세그멘테이션을 단순한 2차원의 사각형 형태로 구성된 특징공간에 따라 세그멘테이션한 결과이다.

제안된 방법의 인식 성능실험은 모든 40 사람에 대한

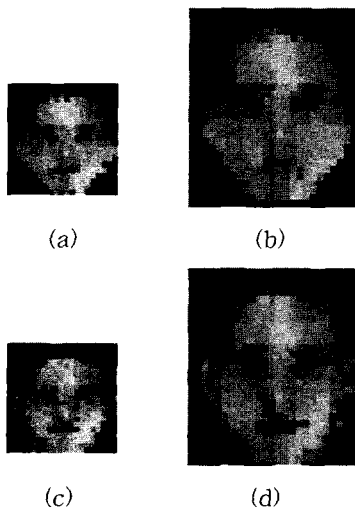


그림 5. 특징공간과 축소영상의 크기에 따른 복원영상 R 의 예.

Fig. 5. The retrieval face image vs. size of the feature space and retrieval image.

표 1. 특징공간 크기에 따른 인식률 Rr (인식불가률 Rj).

Table 1. Recognition rates, Rr vs. size of the feature space(rejection rates, Rj).

Feature space size ($N_{Fx} \times N_{Fy}$)		12×12		16×16		20×20	
		20×20 (E1)	24×28 (E2)	20×20 (E3)	24×28 (E4)	20×20 (E5)	24×28 (E6)
Method							
by D_R (%)		83.5	89.5	87.5	91.0	83.5	88.5
by D_W (%)		75.0	75.0	78.0	78.0	82.0	82.0
Proposed (D_W+D_R) (%)	$p=1$	100 (33.5)	100 (29.0)	100 (27.5)	100 (26.5)	100 (27.5)	100 (22.5)
	$p=2$	99.0 (28.5)	100 (26.0)	100 (22.0)	98.5 (19.0)	99.5 (22.5)	100 (19.5)
	$p=3$	98.0 (24.0)	99.5 (20.5)	100 (18.5)	98.0 (15.0)	97.5 (16.0)	100 (14.5)
	$p=4$	96.5 (21.5)	99.0 (19.0)	99.0 (16.0)	98.0 (13.0)	97.5 (15.5)	99.5 (13.5)
	$p=5$	95.0 (16.0)	99.0 (15.5)	99.0 (15.0)	98.0 (11.5)	96.5 (13.5)	98.5 (11.5)

여 처음 다섯 개의 얼굴영상은 학습에 사용하고, 나머지 5개의 영상에 대하여 인식성능을 테스트하는데 사용하였다. 따라서 전체 인식대상패턴 수 $N=200$ 이 된다. 인식에 성공한 패턴의 수를 N_q , 인식불가의 수를 N_r 이라 하면, 인식률 Rr , 인식불가 비율 Rj 는 각각

$$R_r = \frac{N_q}{(N - N_r)} \quad (10)$$

$$R_j = \frac{N_r}{N}$$

을 백분율로 나타낸 것이다. 각 실험에 대한 결과를 표 1에 나타내었다. 표에서 실험은 E1에서 E6까지 6가지 경우에 대하여 특징공간의 크기를 12×12에서 20×20, 그리고 복구영상의 크기를 20×20 및 24×28로 하여 각각 조사한 것이다. 복구영상에 대해서 유클리드 거리 DR 에 따른 성능은 실험 각각 83.5%에서 88.5%의 인식성능을 나타내고 있다. 그리고 특징공간에 대한 유클리드 거리 D_W 에 따른 인식성능은 각각 75%에서 82%로 복구영상으로 테스트한 결과보다 성능이 약간 떨어짐을 알 수 있다. 이는 얼굴 형태에 따른 특징공간이 복구영상보다 서로가 비슷한 형태로 나타나 오인식이 발생하기 때문이다.

그리고 제안된 방법에 의한 인식성능은 특징공간 인식 패턴의 수 p 에 따라 조사하였다. $p=1$ 일 때는 인식불가의 수가 많지만, 100%의 인식성능을 가졌다. p 가 증가함에 따라 오인식이 늘어나지만 인식불가의 수가 줄어들어 약 15%의 인식불가에서도 거의 100%의 인식률을 가짐을 알 수 있다. 가장 좋은 성능을 가지는 E6 실험에서 14.5%의 인식불가 성능으로 오인식이 전혀 없음을 알 수 있다.

그림 6은 특징벡터 및 복구영상의 크기에 따른 제안된 방법의 인식률을 도시한 것이다. p 의 값이 증가할수록 인식성능이 떨어짐을 알 수 있고, 특징공간과 크기 및 복구영상의 크기가 증가할수록 인식성능이 높게 나타난다. 전체적으로 인식성능은 95% 이상을 가지고 있으며,

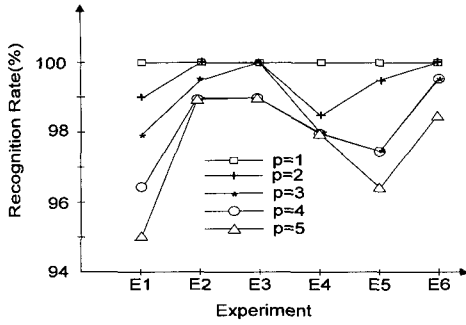


그림 6. 제안된 방법의 인식률.
Fig. 6. Recognition rates of the proposed system.

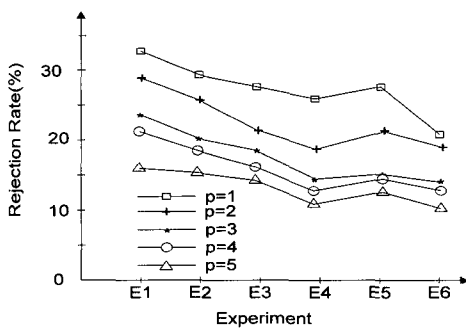


그림 7. 제안된 방법의 인식불가 비율.
Fig. 7. Rejection rates.

특징공간의 크기와 복구영상의 크기를 적당하게 조절하면 인식성능을 100%까지 높일 수 있다.

그림 7은 제안된 방법에 의한 인식불가 비율을 나타내고 있다. p 가 증가할수록 인식불가 비율이 낮아지고 있으며, 또한 특징공간의 크기 및 복구영상의 크기가 커질수록 인식불가 비율도 낮아짐을 알 수 있다.

표 2는 ORL 영상을 사용한 기존의 얼굴 인식 방법과 성능을 비교한 것이다. 이는 Lawrence 등[4]에서 나타나 있다. 기존의 방법에서는 ‘인식불가’는 고려하지 않고 있으며, 고유치를 이용한 방법은 10.5%, PCA(Principle Component Analysis)와 CN(Convolution Network)은 7.5%, 2D-HMM(Hidden Markov Model)은 5%의 오인식률을 가졌다. 가장 좋은 성능은 SOM을 이용하여 영상을 축소한 후 CN망으로 인식한 경우로 3.8%의 오인식률을 가졌다. 반면, 제안된 방법에서는 오인식이 없이 인식불가만 14.5%이다. 인식불가로 처리된 영상은 다른 정밀한 방법에 의해 2차로 인식한다면 더 나은 성능을 얻을 수 있겠다.

표 3은 기존의 방법과의 메모리의 크기를 비교한 것

표 2. 기존의 방법과 비교.
Table 2. Comparison of performance(%).

	recognition rates	rejection rates
Eigenfaces	89.5	.
PCA+CN	92.5	.
Pseudo 2D-HMM	95	.
SOM+CN	96.2	.
Proposed	100	14.5

표 3. N개의 $n \times m$ 크기의 얼굴 영상에서 특징 벡터의 메모리 크기.

Table 3. Memory size of feature in $n \times m$ size and N face images.

	메모리 크기	비고
Global KLT	$nm(1+N_e)$	N_e : 고유벡터의 수
Local KLT	$N_w(1+N_e)$	N_w : 윈도우 크기
SOM	$N_w N_s$	N_s : SOM 출력크기
Proposed	$2NN_s$	

이다. 하나의 영상이 $n \times m$ 의 크기일 경우에 대하여 전역 KLT[1,3], 국부 KLT[2,4] 및 SOM[4]과 같은 기존의 방법은 패턴의 수 N과 무관하게 나타난다. 이는 기존의 세가지 방법은 N개의 얼굴영상에 대하여 사전에 미리 학습하여 그 특징을 미리 저장하기 때문이다. 반면, 제안된 방법에서는 사전 정보없이 현재의 영상에 대해서만 처리를 한다. 따라서 하나의 얼굴영상에 대한 메모리의 크기를 조사하여 보면, 전역 KLT 방법은 원래의 영상보다 항상 크다. 국부 KLT 방법은 윈도우 크기에 직접 비례한다. 반면 제안된 방법은 SOM의 출력 크기, 즉 특징 벡터의 크기에만 의존함을 알 수 있다.

표 4는 특징공간의 크기 $N_s=8 \times 8$, 고유벡터의 수 $N_e=1$, 윈도우의 크기 $N_w=5 \times 5$ 로 하였을 때의 하나의 입력 영상에 대한 메모리의 크기를 나타내고 있다. 국부 KLT 방법이 가장 작은 메모리를 차지하고 있으며, 전역 KLT 방법이 가장 큰 메모리가 요구된다. 제안된 방법은 특징공간의 크기에 비례하여 비교적 작은 메모리를 요구

표 4. $n \times m$ 크기 하나의 얼굴 영상에서 특징 벡터의 메모리 크기.

Table 4. Memory size of feature with a $n \times m$ face image.

방법	메모리 크기
Global KLT	$2nm = 20608$
Local KLT	$2 \cdot 25 = 50$
SOM	$25 \cdot 64 = 1600$
Proposed	$2 \cdot 64 = 128$

표 5. 각 영상 축소방법에 대한 비교.

Table 5. Comparison of a performance vs. processing meth.

	Global KLT	Local KLT	SOM	Proposed
메모리	중간	작다	중간	영상 수에 비례
복구영상의 크기	고정	고정	고정	임의의 크기 가능
사전정보	필요	필요	필요	불필요
병렬화	알고리즘	알고리즘	가능	가능
특징의 수	영상크기에 비례	영상크기에 비례	영상크기에 비례	고정
디지털화	불가능	불가능	가능	가능
특징	전체 에러 최소화	부분 에러 최소화	부분 에러 최소화	물체영역 에러 최소화

한다.

표 5는 영상을 축소하는 기존의 3가지 방법과 제안된 방법을 비교한 것이다. 먼저 메모리의 크기를 살펴보면 전술한 바와 같이 제안된 방법은 사전 지식이 없는 현재 영상에 대해서만 처리함으로써 전체적으로 영상패턴 수에 비례하지만, 다른 방법은 전체 패턴을 미리 정의된 크기에 따라 학습하여, 학습된 패턴에 대해서 최소의 복구에러를 가지도록 정의한 것이다. 이들의 복구영상의 크기를 살펴보면, 기존의 방법에서는 그 크기가 입력영상에 따라 고정된다. 이는 일정한 크기를 갖는 복구영상을 얻기 위해서는 입력영상의 크기도 일정하여야 한다는 것을 의미하고 있어서, 일정한 특징패턴으로 임의의 크기를 갖는 복구영상을 만들 수 없어서 벡터화라 할 수 없다. 반면, 제안된 방법은 입력영상의 크기와 무관하게 일정한 특징공간의 크기를 얻을 수 있으며, 복구영상의 크기도 임의의 크기로 얻을 수 있어서 특징공간을 벡터화할 수 있다. 그리고 무엇보다도 기존의 방법은 백그라운드영상 및 노이즈영상까지도 포함한 전체영상에 대한 복구에러를 최소화거나 부분적으로 최소화러를 가지도록 처리되지만 제안된 방법은 물체영역, 즉 관심영역에 대해서만 특징공간을 형성하여 복구영상에서는 백그라운드영상은 사라지게 된다. 이는 특징의 수가 입력영상의 크기와 무관하게 일정하여 이후에 처리되어야 하는 인식 및 기타 특수처리를 위해 필수적으로 필요하다. 즉, 제안된 방법은 임의의 얼굴형상에 대한 벡터화 방법으로 이 영상벡터를 가지고 있으면 임의의 크기를 갖는 얼굴을 얻을 수 있으며, 크기의 정규화 등에 유용하게 사용될 수 있다.

IV. 결론

본 논문에서는 영상내의 특정한 정보를 갖는 물체, 즉 사람의 얼굴에 대한 정보만을 세그멘테이션하여 영역을 분할한 후, 이로부터 2차원의 특징공간을 추출하고 이를 영상벡터로 하여 일정한 크기를 갖는 영상으로 복원이 가능하도록 하였다. 그리고 특징공간과 복구영상의 조합으로 개개의 얼굴을 인식하는 인식 방법을 적용하였으며, 인식이 확실치 않은 경우에는 인식불가로 처리하여 오인식을 줄였다. 제안된 방법은 기존의 방법과 달리 원래의 영상으로부터 얼굴 고유의 특징점들을 추출하였으며, 복구영상을 이용한 노이즈에 대한 오인식에 문제가 없고, 얼굴의 크기가 작거나 크더라도 그 크기에 무관하게 일정한 크기로 정규화시킬 수 있어서 인식분야에 효율적으로 쓰일 수 있다. 제안된 인식 방법의 타당성을 확인하기 위하여 ORL 얼굴영상에 대한 얼굴 데이터를 가지고 실험한 결과, 특징공간의 인식 리스트 수 p 에 따라 인식 성능이 의존하였으며, $p=1$ 일 때는 인식불가의 수가 많지만, 100%의 인식성능을 가졌다. 가장 좋은 성능은 특징공간의 크기 20×20 , 복구영상의 크기 24×28 및 $p=3$ 에서 14.5%의 인식불가로 오인식이 전혀 없었다. 앞으로 SOM의 정밀학습 및 각 특징의 크기조절로 인한 인식불가의 수를 줄일 수 있는 방안과 2차 정밀 인식을

위한 연구가 동반되어야 하겠다.

참고문헌

- [1] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, pp. 71-81, 1991.
- [2] A. Pentland, B. Moghddam and T. Starner, "View-based and modular eigenspaces for face recognition," *MIT Technical Report* no. 245, 1994.
- [3] A. Pentland, R. Picard and S. Scaloff, "Photobook : Tools for content-based manipulation of image databases," *Proc. of SPIE-Storage and Retrieval for Images and Video Databases*, vol. 2185, pp. 34-47, 1994.
- [4] S. Lawrence, C. L. Giles, A. C. Tsoi and A. D. Back, "Face recognition : A convolutional neural network approach," *IEEE Trans. Neural Networks*, vol. 8, no. 1, pp. 98-113, 1997.
- [5] 김선종, "영상 세그멘테이션을 이용한 특징공간 추출 및 얼굴인식," *한국자동제어학술회의 논문집*, pp. 1141-1144, 1998. 10.
- [6] D. Valentin, H. Abdi, and A. J. O'Toole, "Categorization and identification of human face images by neural network : A review of the linear autoassociative and principal component approaches," *Journal of Biological Systems*, 2(3), pp. 413-429, 1994.
- [7] D. Valietin, H. Abdi, A. J. O'Toole, G. W. Cottrell, "Connectionist models of face processing : A survey," *Pattern Recognition*, vol. 27, pp. 1209-1230, 1994.
- [8] R. Brunelli and T. Poggio, "Face recognition : Features versus templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15(10), pp. 1042-1052, 1993.
- [9] S. P. Luttrell, "Hierarchical self-organizing network," *Proc. of the 1st IEE Conference on Artificial Neural Networks*, pp. 2-6, London, UK, 1989.
- [10] 김선종·구탁모·성효경·최홍문, "영상의 위상차를 이용한 얼굴인식," *전자공학회지* 제 35권 S편 6호, pp. 81-87, 1998.
- [11] T. Kohonen, *Self-Organizing Map*, Springer-Verlag, Berlin, 1995.



김 선 종

1965년 3월 29일생. 1989년 경북대학교 전자공학과 졸업. 동대학원 석사(1991), 박사(1996). 순천제일대학 제어계측과 전임강사(1995-1997). 1997-현재 밀양대학교 컴퓨터공학과 조교수. 관심분야는 패턴인식, 신경망, 병렬처리 등임.