

論文97-34C-10-6

# 주 시각피질에서의 단순세포 수용영역 형성에 대한 성긴 집단부호 모델을 이용한 얼굴인식

(Face recognition using a sparse population coding model for receptive field formation of the simple cells in the primary visual cortex)

金鐘圭\*, 張珠錫\*\*, 金榮一\*\*\*

(Jong-Kyu Kim, Ju-Seog Jang, and Yeong-Il Kim)

## 요 약

본 논문에서는 초기 시각피질에서의 단순세포의 수용영역 학습모델인 성긴 집단부호를 이용한 얼굴인식 방법을 제안하였다. 20명의 사람으로 부터 획득한 20개의 정면얼굴을 학습시키고, 한 사람당 10개씩의 여러 가지로 변화된 얼굴 200개를 시험하였다. 안경을 끼거나 표정이 변화된 얼굴이 포함된 정면얼굴에 대해서는 100%의 인식률을 얻었으며 회전된 얼굴을 포함한 전체 입력에 대해서는 90%의 인식률을 얻었다. 성긴 정도를 결정짓는 비선형함수의 효과를 분석하였으며, 성긴 집단부호를 이용한 인식률과 이완은 대조적인 밀집 부호의 하나인 고유벡터(얼굴)에 의한 인식률을 비교 분석하였다.

## Abstract

In this paper, we present a method that can recognize face images by use of a sparse population code that is a learning model about a receptive fields of the simple cells in the primary visual cortex. Twenty front-view facial images from twenty persons were used for the training process, and 200 varied facial images, 10 per person, were used for test. The correct recognition rate was 100% for only the front-view test facial images, which include the images either with spectacles or of various expressions, while it was 90% in average for the total input images that include rotated faces. We analyzed the effect of nonlinear function that determine the sparseness, and compared recognition rate using the sparse population code with that using eigenvectors(eigenfaces), which is compact code that makes contrast with the sparse population code.

## I. 서 론

오늘날 보안에 대한 중요성이 크게 부각되면서 얼굴

인식에 대한 관심이 점점 높아 지고 있다. 그러나 얼굴영상 인식은 다른 패턴인식과는 달리 인식하고자하는 대상이 거의 비슷하고 시간이 지남에 따라 영상이 서서히 변하기 때문에 아직까지 인식률 등에서 만족할 만한 결과를 얻지 못하고 있다.

얼굴인식 기술은 크게 얼굴의 영역분리, 특징 추출, 그리고 얼굴분류의 세 과정으로 구성된다. 이 중에서 특징추출은 인식에 있어서 가장 중요하고 근본적인 문제이다. 얼굴영상의 특징은 경계선, 등고선, 영상영역 등의 시각적 특징, 히스토그램, 모멘트 등의 통계적 특

\* 正會員, 蓮庵工業專門大 電子工學科

(Dept. of electronics Yonam College of Eng.)

\*\* 正會員, 釜慶大學校 情報通信工學科

(Dept. of Telematics Eng., Pukyong Nat'l Univ.)

\*\*\* 正會員, 慶尙大學校 電子工學科

(Dept. of electronics, Kyeongsang Nat'l Univ.)

接受日字:1997年7月21日, 수정완료일:1997年9月18日

징, 푸리에변환 등의 변환계수적 특징, Karhunen-Loeve(K-L)변환, 특이값 분해법 등의 대수적 특징으로 분류된다.<sup>[11]</sup> 본 논문에서 제안한 성긴 집단부호법에 의한 특징벡터는 대수적 특징의 범주에 속한다.

얼굴의 특징 추출은 측면얼굴을 이용하거나 정면얼굴을 이용하는데, 측면얼굴을 이용하는 경우 특징추출이 비교적 용이하다는 장점은 있으나 영상획득이 어렵고 얼굴과 배경을 정확히 분리해야하는 단점이 있다. 또한 일반적인 얼굴영상은 정면얼굴이기 때문에 대부분의 경우 정면얼굴을 이용한다. 정면얼굴에 대한 인식방법으로는 얼굴의 특징점 위치와 특징점 간의 거리나 비율을 이용하는 방법,<sup>[5]</sup> 얼굴 구성성분의 템플릿(template)을 이용하는 방법,<sup>[6-7]</sup> 그리고 영상의 대수적 특징을 이용하는 방법이 있다.<sup>[1,9-11]</sup> 얼굴 특징점 간의 위치와 거리 관계 등을 이용하는 경우는 인식을 높이기 위해서 특징점을 정확히 추출해야 하는데 정확한 특징점 추출이 어렵다는 문제점이 있다. 또 템플릿 정합법의 경우는 영상의 크기가 변할 때 고정된 템플릿에 의해 인식이 떨어진다. Yuille은 가변 템플릿에 의한 인식 방법을 개발하였으나<sup>[8]</sup> 이 방법은 계산이 너무 복잡하다. 최근에 와서는 K-L변환에 의한 고유벡터를 이용하여 얼굴인식을 하거나<sup>[9-11]</sup> 특이값 분해(singular value decomposition)법을 이용한<sup>[11]</sup> 얼굴인식이 크게 주목을 받고 있다. 이 방법은 특징벡터의 추출이 비교적 쉽고 특징 벡터가 대수적, 기하학적으로 매우 안정되어 있기 때문에 영상을 표현하고 인식하는데 대단히 유용한 수단이 된다. 그러나 인식률에서는 아직까지 만족할만한 결과를 얻지 못하고 있다. 특히 변화된 영상에 대해서는 인식이 크게 떨어진다.

본 논문에서는 포유동물의 시각체계에서의 정보처리 방법을 이용하여 얼굴을 인식하는 연구를 수행하였다. 주 시각피질에서의 단순세포의 수용영역(receptive field)은 여러 가지로 회전된 선소모양이라는 것이 오래전 실험으로 밝혀졌으며, 최근 더욱 세밀한 실험결과에 따라 길쭉한 모양의 회전된 gabor filter 형태로 모델화하고 있다. 이러한 신경세포들의 반응특성을 비교사(unsupervised) 학습 알고리즘으로 구현하기 위한 많은 연구가 있었으나<sup>[12-15]</sup> 수용영역의 특징을 완전하게 표현하지는 못했다. 최근 Olshausen과 Field는 주 시각피질에서의 성긴 집단부호 모델을 개발하였는데 그들은 자연영상에 대해 이 학습모델을 적

용시켰을 때 얻어지는 수용영역이 주 시각피질에서의 단순세포들의 것과 유사함을 보였다.<sup>[16]</sup> 주 시각피질에서 처리된 신호는 대뇌피질의 다른 시각영역(visual area)으로 전달되는데, 원숭이의 뇌연구에서 뇌의 측두(infero-temporal, IT)영역에 얼굴모양의 수용영역을 갖는 세포들이 존재한다는 사실이 밝혀져 있다.<sup>[17]</sup> 인간이나 원숭이는 IT영역에 손상이 있을 경우 얼굴을 인식하지 못한다는 것이 잘 알려져 있다.<sup>[18]</sup> Young과 Yamane는 얼굴모양에만 강하게 반응하는 세포들이 성긴 집단부호(sparse population code)화한다는 것을 뒷받침하는 원숭이 뇌실험 결과를 보고한 바 있다.<sup>[19]</sup> 그러나 IT영역의 세포가 어떻게 구체적으로 얼굴인식에 관여하는지 아직 까지 정확하게 밝혀지지 않았다.

우리는 Olshausen과 Field가 개발한 성긴 집단부호의 학습모델이 IT영역과 같은 높은 단계의 정보처리 영역에도 적용된다고 가정하고 이 학습모델을 얼굴인식에 이용하였다. 특히 성긴 집단부호를 얼굴인식에 적용하였을 때의 특성을 분석하였는데 성긴 정도를 결정하는 비선형 요소가 인식에 있어서 중요한 역할을 함을 알았다. 그리고 제안한 방식의 유용성을 알아보기 위해 성긴 집단부호와 대조적인 밀집부호의 하나인 고유벡터를 이용한 경우와 비교하였다. 남자 20명에 대해 각각 11개의 다양한 얼굴을 획득하고 이 중에서 정면얼굴 20개는 학습영상으로 사용하고 나머지 200개는 시험영상으로 사용하였다. 시험영상은 안경을 착용하거나 표정을 변화시키거나 좌로, 위로 회전된 영상, 그리고 정면영상으로 구성되어 있다.

## II. 성긴 집단부호 모델

성긴 집단부호는 밀집한 부호(compact code)와 대비되는 부호방식이다. 밀집한 부호는 데이터를 극히 낮은 차원으로 표현함으로써 입력에 있는 리던던시(redundancy)를 제거하는 부호로서 주 성분 분석법(principal component analysis)이 그 대표적인 예이다. 반면에 성긴 집단부호는 그림 1과 같이 특정한 입력에 대해 출력의 차원은 변하지 아니하고 점선으로 표시한 비활성화 세포(데이터 값 : 0)와 함께 실선으로 표시한 소수 개의 활성화된 세포(active cell, 데이터 값 : 0이 아닌 임의의 값)로 출력을 표현한다. 활성 세포의 개수는 전체 세포수에 비해 극히 적고 성기

게 분포한다. 즉 각 세포는 동일한 확률분포를 갖지만 어떤 특정 세포가 활성화될 확률은 낮고 활성 패턴에 대한 리던던시는 입력 데이터의 리던던시와 같다. 따라서 성긴 집단부호는 데이터 사이의 통계적 종속을 최소화 함으로써 세포들의 응답 히스토그램에 대한 리던던시가 최대로 되게 하는 코딩방식이라 할 수 있다<sup>[15]</sup>. 이러한 성긴 집단부호 모델은 일부 세포가 파괴되어 데이터가 유실되더라도 나머지의 활성화된 세포에 의해 입력 데이터를 효과적으로 복원할 수 있게 된다. 고등동물의 대뇌피질의 여러 영역 뿐만 아니라 곤충들의 감각기관 세포들도 이러한 형태로 코딩한다는 것이 여러 논문을 통해 많이 보고된 바 있다.<sup>[20]</sup>

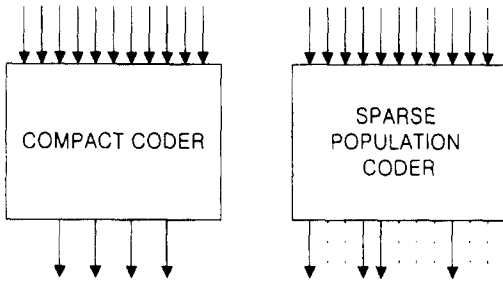


그림 1. 밀집부호와 성긴 집단부호  
Fig. 1. The compact code and the sparse population code.

입의 영상  $I(x, y)$ 에 대한 Olshausen과 Field의 성긴 집단부호 모델은 다음과 같은 비용함수(cost function)  $E$ 를 정의하고 이를 최소화하는 파라미터값을 구함으로써 얻어진다.

$$E = \sum_{x,y} [I(x, y) - \sum_{i=1}^N a_i \phi_i(x, y)]^2 + \lambda \sum_{i=1}^N S\left(\frac{a_i}{\sigma}\right) \quad (1)$$

$\phi_i(x, y)$ 는  $I(x, y)$ 를 표현하기 위한  $N$ 개의 기저함수(basis function)이고,  $a_i$ 는  $\phi_i(x, y)$ 들의 선형 결합계수이다.  $\lambda$  와  $\sigma$ 는 양수이고  $S(\cdot)$ 는 성긴 정도(sparseness)를 얻기위해 도입된 유한 최소값을 갖는 비선형함수이다.

(1)식의 우변 첫번째 항은 기저함수  $\phi_i(x, y)$ 로 영상을 얼마나 잘 표현하는지를 나타내는 척도이다. 즉 이 항의 최소값은

$$I(x, y) = \sum_{i=1}^N a_i \phi_i(x, y) \quad (2)$$

일 때이며,  $\phi_i(x, y)$ 를  $i$ 번째 신경의 수용영역함수로,  $a_i$ 를  $i$ 번째 신경의 활동세기로 해석할 수 있다.

비선형함수  $S(x)$ 는  $-e^{-x^2}, \ln(1+x^2), |x|$  등을 쓸 수 있다. Olshausen과 Field는  $I(x, y)$ 를 많은 수의 임의의 자연풍경 영상을 택하고 식 (1)을  $a_i$ 와  $\phi_i(x, y)$ 에 대해 최소화시켰을 때  $\phi_i(x, y)$ 가 주 시각 피질의 단순세포의 수용영역과 유사해 짐을 보였다.

입력영상  $I(x, y)$ 가 사람 얼굴인 경우에 대해 식 (1)을 최소화하는  $\phi_i(x, y)$ 와  $a_i$ 를 다음과 같이 구한다.  $I(x, y)$ 를  $L$ 명의 사람 얼굴로 하고  $L$ 명의 얼굴을  $F_j(x, y), j=1, 2, 3, \dots, L$  혹은 1차원 벡터로  $F_j$ 로 표현하기로 한다. 마찬가지로  $N$ 개의  $\phi_i(x, y), i=1, 2, 3, \dots, N$  역시 1차원벡터  $\phi_i$ 로 표현하기로 한다. 비선형 함수  $S(x) = \ln(1+x^2)$  를 사용하면  $j$ 번째 얼굴에 대한 비용함수  $E$ 는 다음과 같이 된다.

$$E = \|F_j - \sum_{i=1}^N a_i \phi_i\|^2 + \lambda \sum_{i=1}^N \ln\left(1 + \left(\frac{a_i}{\sigma}\right)^2\right) \quad (3)$$

모든  $F_j, j=1, 2, 3, \dots, L$ 에 대해  $E$ 를 최소화하는  $a_i$ 와  $\phi_i$ 는 다음과 같이 최대 경사강하법으로 구한다. 즉 초기 임의의  $a_i(t=0), \phi_i(t=0)$ 를 선택한 후, 먼저 특정입력  $F_j$ 에 대해 최소의  $E$ 를 주는  $a_i$ 들을 찾는다. 즉,

$$\dot{a}_i = \phi_i^T F_j - \sum_k \phi_k^T \phi_i a_k - \frac{\lambda}{\sigma} S\left(\frac{a_i}{\sigma}\right) \quad (4)$$

에서 정상상태가 되는 ( $\dot{a}_i=0$ )  $a_k$  값을 구하면 된다. (4)식에서 T는 전치를 의미한다.  $F_j, j=1, 2, 3, \dots, L$ 에 대해 각각의  $a_k$ 를 구한 다음 최소의  $E$ 를 주는  $\phi_i$ 를 구한다. 즉 마찬가지로  $\Delta \phi_i = \phi_i(t+1) - \phi_i(t) = -\frac{\eta}{2} \frac{\partial E}{\partial \phi_i}$ 가 되도록 하면

$$\Delta \phi_i = \eta \langle a_i (F_j - \sum_k a_k \phi_k) \rangle \quad (5)$$

가 된다.  $\langle \rangle$ 는  $j$ 에 대한 통계적 평균을 의미하고  $\eta$ 는 학습률이다. 각각의 얼굴,  $F_j$ 에 대한  $N$ 개의  $a_i$ 들의 값들  $a_i'$ 로 표현하자. 즉  $a_j = [a_1', a_2', \dots, a_N']^T$ 라 하면  $a_j$ 는 하나의 학습영상  $F_j$ 에 대해 유일한 값이 되고 차원도  $N \times 1$ 로 영상 자체의 차원보다 대단히 작다. 그러나 한가지 주목할 것은  $N$ 이  $L$ 보다는 클

수 있다는 것이다. 이것은 밀집코드의 일종인 고유얼굴(eigenface) 방법과는 아주 다른 성질 중의 하나이다.

성긴 집단부호를 이용한 인식의 수행과정은 다음과 같다. 위에서 기술한 학습영상  $F$ 들에 대한  $\phi_i, i = 1, 2, \dots, N$ 를 이용하여, 미지의 입력영상에 대한 벡터  $a$ 를 구하고 이것을 학습영상의  $a_i$  벡터와의 유클리드 거리를 구한다. 이 중에서 유클리드 거리가 최소가 되는 것에 대응되는 학습영상이 인식하고자 하는 영상이 된다.

### III. 실험 및 결과

#### 1. 학습영상과 시험영상

남자 20명에 대해 CCD 카메라를 통해 얻은 영상을 프레임그래버(Frame Grabber)로 디지털화 하여 180x120 크기의 다양한 얼굴 영상을 획득하였다. 즉 한 사람당 정면얼굴 3개, 옆으로 위로 10도 20도씩 돌려진 얼굴 4개, 안경낀 얼굴 1개, 눈감은 얼굴 1개, 표정이 변화된 얼굴 2개 등 모두 11개씩의 영상을 취득하였다. 이들 영상 중에서 20명의 정면얼굴 1개씩을  $\phi_i$ 를 학습시키는데 사용하였고 나머지 200개는 인식률을 얻기 위해 사용하였다. 이 때 계산량 감소를 위해 영상의 크기를 30x20으로 축소하였다. 그림 2는 한 사람에 대한 10개의 시험영상의 예이다. 그리고  $\phi_i$ 는 모두 100개이며 충분히 학습한 후의  $\phi_i$ 의 모습은 그림 3과 같다. 그림에서 보는 바와 같이  $\phi_i$ 중에는 사람얼굴 형태가 많이 존재하는데 어떠한 학습얼굴과 일치하는 것은 없다. 이것은 뇌의 IT영역에서 얼굴모양의 수용 영역을 갖는 세포들이 존재한다는 사실과 부합한다.

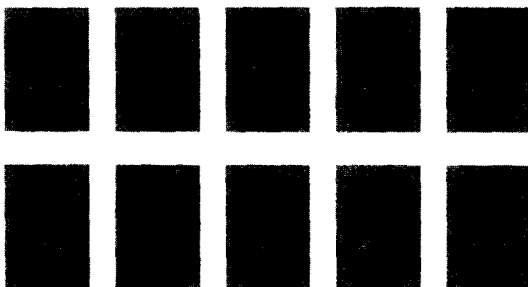


그림 2. 여러 형태의 시험영상  
Fig. 2. The various test images.

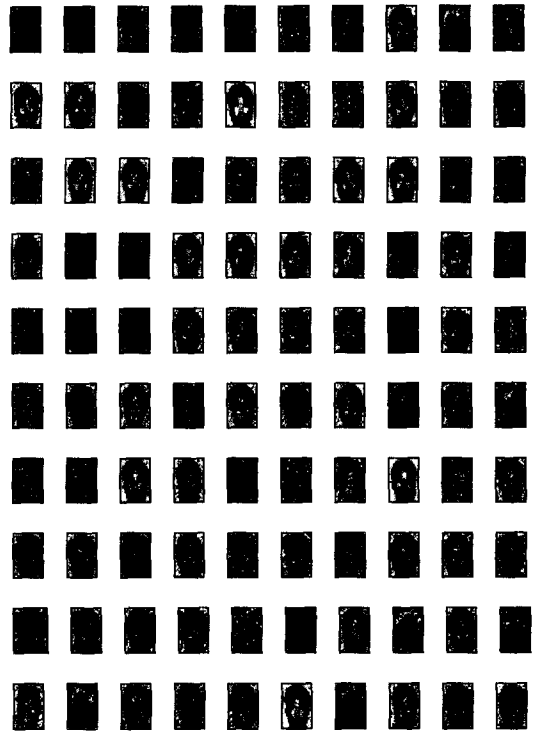


그림 3. 학습된 기저함수  $\phi_i$ 의 모습  
Fig. 3. The shape of the trained basis function  $\phi_i$ .

2. 얼굴패턴 인식에 관한 성긴 집단부호 모델의 성질 영상의 성긴 집단 부호가 얼굴영상의 인식에 어떻게 잘 작용하는지를 살펴보기 위해 학습영상과 시험영상의 벡터  $a$ 의 모습을 비교해 보았다. 여기서 벡터  $a$ 는 신경세포의 활동세기로 볼 수 있으며 신경세포의 수 즉  $\phi$ 의 개수는 100인 경우이다. 그림에서 100개의  $a_i$  값들을 10x10 평면에 늘어 놓았는데 밝은 색은 신경세포의 활동이 강한 것을 의미하고 검은색은 활동이 전연 없음을 의미하며 또한 회색 부분은 중간 정도의 활동 상태를 의미한다. 그림 4-5에서 보는 바와 같이 시험영상 m1-2, m1-3, ..., m1-11의  $a$ 모습과 학습영상 m1-1의  $a$ 의 모습은 매우 유사하고 다른 학습영상 m2-1, m3-1, m4-1, ... 등에 대한  $a$ 와는 아주 다른 모습을 보이고 있다. 따라서 m1에 대한 여러 입력모드 m1-2, m1-3, ..., m1-11의  $a$ 벡터는 학습영상 m1-1, m2-1, m3-1, ...중에서 m1-1 영상의  $a$  벡터와의 유클리드 거리가 가장 짧다. 또 입력 모드 6, 7은 옆으로 각각 10도, 20도씩 돌려진 얼굴이고 모드 9, 10은 위로 각각 10도, 20도씩 돌려진 얼굴인데 표

에서 보는 바와 같이 모드 8과 10의 경우가 유클리드 거리가 가장 크게 나타났다. 이것은 입력의 모습을 그대로 잘 표현하고 있음을 보여준다.

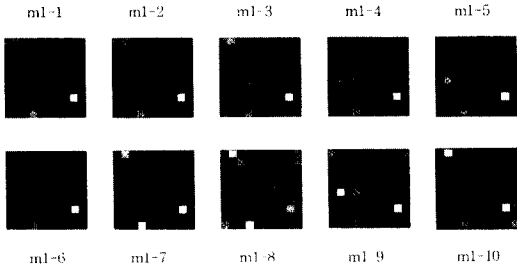


그림 4. 동일인의 시험영상에 대한  $a$  값들  
Fig. 4.  $a$  values for test images of the same person.

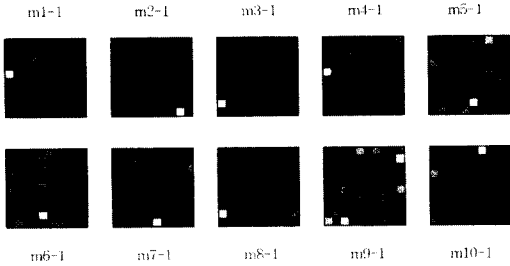


그림 5. 학습영상에 대한  $a$  값들  
Fig. 5.  $a$  values for training images.

표 1. 학습 영상의  $a$ 와 시험영상의  $a$  간의 유클리드 거리

Table 1. Euclidian distance between  $a$  of the training images and  $a$  of the test images.

training images	test images									
	m1-2	m1-3	m1-4	m1-5	m1-6	m1-7	m1-8	m1-9	m1-10	m1-11
m1-1	79	92	115	143	86	190	269	167	267	145
m2-1	313	318	319	330	308	346	379	339	390	332
m3-1	320	332	342	358	329	379	416	353	376	348
m4-1	319	323	326	340	322	352	382	321	362	326
m5-1	304	328	328	341	317	349	387	343	366	321
m6-1	377	378	384	391	371	417	436	366	401	368
m7-1	325	338	354	361	336	374	407	347	361	342
m8-1	313	295	300	312	302	326	346	303	379	317
m9-1	318	328	358	369	342	360	387	351	357	340
m10-1	325	336	351	361	337	377	409	355	376	348

\*e-4

$a$ 의 성긴 정도(sparseness)를 변화시킴에 따라 인식

에 어떤 영향이 나타나는지를 살펴보았다.  $\phi_i$ 의 개수 즉  $N=100$ 인 경우에 대해 식 (4)의  $\lambda$ 값의 변화에 따라 시험영상과 학습영상 간의 평균 유클리드 거리를 구해보면 그림 6과 같이  $\lambda=0$  즉 비선형함수에 의한 성긴 정도가 없는 경우는 학습된 자신의 얼굴과 변화된 자신의 입력얼굴과의  $a$ 벡터의 거리가 크게 나타나지만  $\lambda$ 가 증가하면 각 얼굴간의 벡터거리는 감소한다. 그러나 같은 사람의 얼굴에 대한  $a$  거리의 감소는 서로 다른 사람들의 얼굴간의  $a$ 거리 감소보다 빠르다. 즉  $a$ 에 있어서 일종의 군집화(clustering)가 발생하는 것이다. 이것은 패턴인식(혹은 분류)를 효과적으로 할 수 있게 해주는 것이다. 이러한 것을 좀 더 자세히 살펴 보기 위해서 그림 7에 특정인의 학습얼굴에 대한  $a$  벡터를 기준벡터로 했을 때 이 기준벡터와 여러 시험 얼굴에 대한  $a$ 벡터 간의 크기비와 위상차를 극좌표 형식으로 나타내었다. 기준벡터와 같은 사람에 대한  $a$ 벡터(0로 표시)는 다른 사람에 대한  $a$ 벡터(\*로 표시) 보다 기준벡터(#로 표시)에 가깝게 위치해 있고  $\lambda$  값이 증대함에 따라 기준벡터에 더욱 더 접근하면서 within class의 평균거리가 짧아지고 있음을 알 수 있다.

$i$ 번째 사람얼굴의 모드1 부터 모드11 까지를  $F_i^{(1)}, F_i^{(2)}, \dots, F_i^{(11)}$ 로 두고 이것에 대한  $a$ 벡터를 각각  $a_i^{(1)}, a_i^{(2)}, \dots, a_i^{(11)}$ 로 표기할 때 학습 얼굴  $F_i^{(1)}$ 에 대한  $a$ 벡터와 동일인의 입력얼굴  $F_i^{(k)}, k=2, 3, \dots, 11$ 에 대한  $a$ 벡터 사이의 평균거리  $d_0$ 는 다음과 같이 표현된다.

$$d_0 = \frac{1}{M} \sum_{k=2}^M \| a_i^{(1)} - a_i^{(k)} \|, M=11 \quad (6)$$

또 입력얼굴  $F_j^{(k)}, k=2, 3, \dots, 11$ 에 대한  $a$ 벡터와 다른 사람의 학습얼굴들  $F_i^{(1)}$ 에 대한  $a$ 벡터 사이의 평균거리  $d_j^k$ 는 다음 식과 같다.

$$d_j^k = \frac{1}{L} \sum_{i=2}^M \| a_i^{(1)} - a_j^{(k)} \|, k=2, 3, \dots, 11, i \neq j \quad (7)$$

여기서  $L$ 은 학습영상의 개수이다.

$\lambda$ 값이 증가하면  $d_0$ 와  $d_j^k$ 가 모두 감소되나  $d_0$ 가  $d_j^k$ 보다 훨씬 빨리 감소된다.  $\lambda$ 값의 변화에 따른 여러 형태의 입력에 대한  $d_j^k/d_0$ 의 평균값은 그림 8과 같다. 따라서  $\lambda$ 값이 증가함에 따라 인식률이 증가될 수 있음을 보여준다. 그러나  $\lambda$ 값이 너무 크면 그림 9와 같

이 E값도 커지므로 인식률은 오히려 감소하게된다.

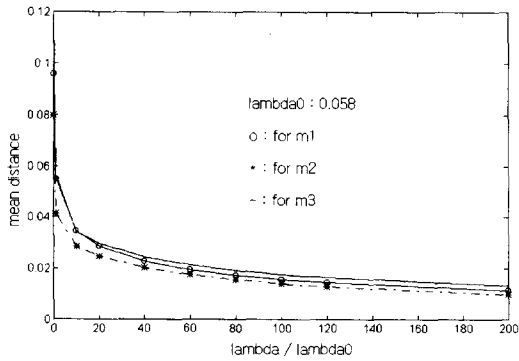


그림 6.  $\lambda$ 에 따른 학습 영상과 시험영상에 대한 벡터  $a$ 의 평균 거리  
Fig. 6. The mean distance between vector  $a$  of training images and vector  $a$  of test images as  $\lambda$  varies.

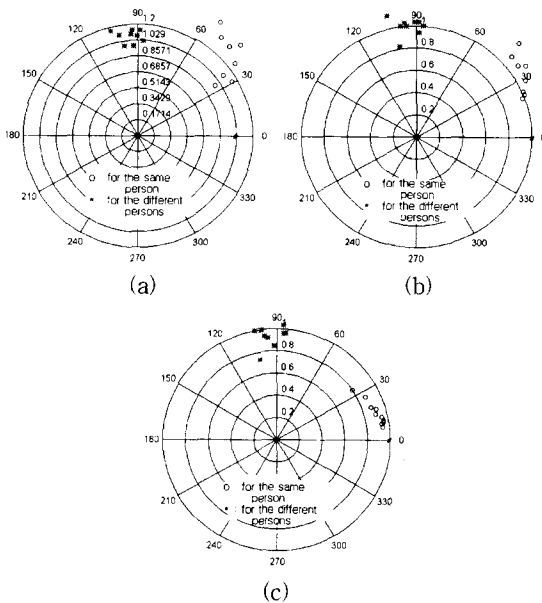


그림 7.  $\lambda$ 의 변화에 따른 벡터  $a$ 의 분포 (a)  $\lambda/\lambda_0 = 10$ , (b)  $\lambda/\lambda_0 = 100$ , (c)  $\lambda/\lambda_0 = 1000$   
Fig. 7. The distribution of vector  $a$  as  $\lambda$  varies. (a)  $\lambda/\lambda_0 = 10$ , (b)  $\lambda/\lambda_0 = 100$ , (c)  $\lambda/\lambda_0 = 1000$ .

3. 인식률

성긴 부호부호를 이용하여 여러 형태의 입력 영상에 대한 인식실험을 행하였다. 또한 동일한 입력에 대해 고유벡터(고유얼굴)을 이용한 경우와 비교하였다. 이때 입력영상의 수는 20개이고  $\lambda = 5.81$ ,  $\sigma = 0.1$ ,  $N = 100$ , 그리고 학습률  $\eta = 0.01$ 이다. 또 고유벡터에 의한 인식의 경우 고유얼굴의 수는 19개이다.

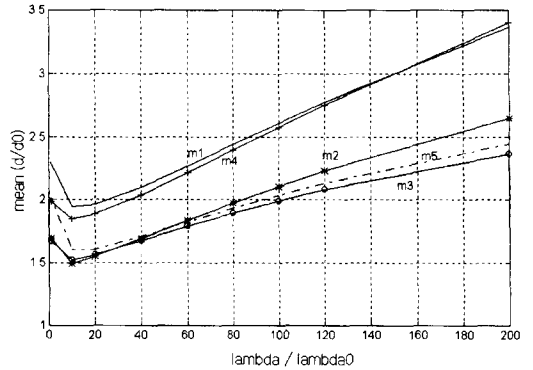


그림 8.  $\lambda$ 값의 변화에 따른 여러 형태의 입력에 대한  $d_i^*/d_0$ 의 평균  
Fig. 8. The mean distance of  $d_i^*/d_0$  for various input images as  $\lambda$  varies.

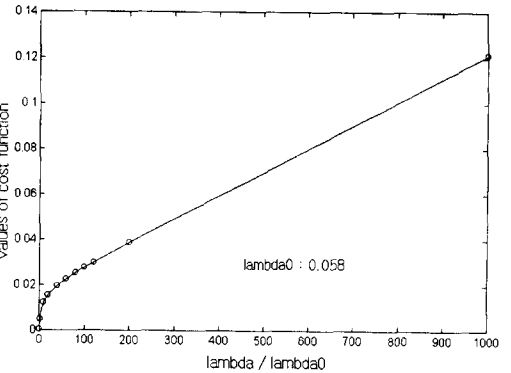


그림 9.  $\lambda$ 의 변화에 따른 비용 함수값  
Fig. 9. The values of cost function  $E$  as  $\lambda$  varies.

표 2. 인식률 비교

Table 2. Comparison of % correct recognition rate.

입력얼굴형태	스파스부호	고유벡터
정면얼굴	100%(120/120)	100%(120/120)
회전된 얼굴	75%(60/80)	72.5%(58/80)
계	90%(180/200)	89%(178/200)

표 2에서 보는 바와 같이 성긴 집단부호를 이용한 경우 안경을 끼거나, 얼굴표정이 변화된 것 등의 정면 얼굴에 대한 인식률은 100%로서 아주 만족스런 결과를 얻었다. 옆으로 또는 위로 회전된 얼굴에 대해서는 인식률이 저하되었으나 고유벡터를 이용한 경우 보다는 다소 우수하였다.

잡음에 대한 강인성을 살펴보기 위해 평균값이 0이

고 분산이 1인 가우스분포를 갖는 랜덤잡음의 정수배를 입력영상에 첨가하였다. 그림 10에서 보는 바와 같이 성긴 집단 부호를 이용한 인식 방법은 PSNR (peak signal-to-noise ratio)이 20dB 정도의 영상에 대해서도 오인식률이 0%이나 고유벡터를 이용한 경우는 PSNR이 60dB 이상이 되어야 오인식률이 5% 이내로 되었다. 따라서 성긴 집단부호는 잡음에 대단히 강한 인식 수단을 제공함을 알 수 있다.

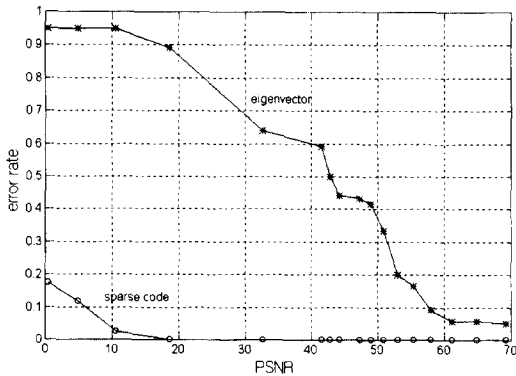


그림 10. PSNR에 따른 성긴 집단부호 방식과 고유 벡터 방식의 오인식률  
 Fig. 10. The error rate versus PSNR for sparse population code method and eigenvector method.

IV. 결 론

임의의 영상은 기저함수  $\phi_i$ 와 계수  $a_i$ 의 선형적인 조합으로 표현할 수 있는데 비선형함수에 의해  $a_i$ 의 성긴 정도를 증가시킴으로써 동일인의 학습영상과 시험 영상 간의 벡터  $a$ 에 대한 유클리드 거리가 다른 사람의 학습영상과 시험영상 간의 벡터  $a$ 에 대한 유클리드 거리 보다 급격히 감소되는 군집화 현상을 규명하였고 이러한 성질을 이용하여 얼굴인식을 행한 결과 약간의 변화된 얼굴영상 까지 포함하여 정면얼굴에 대해서 100%의 인식률을 얻었다. 또 입력영상에 잡음이 포함될 때도 아주 양호한 인식 결과를 얻었는데 동일 조건하에서 고유벡터를 이용한 얼굴인식과 비교해 볼으로써 성긴 집단부호를 이용한 얼굴인식의 유용성을 확인할 수 있었다.

그러나 제안된 방법은 비용 함수값을 최소화하는데 많은 시간이 소요된다는 단점이 있다. 따라서 이 점을

해결할 수 있는 연구가 계속되어야 할 것이다.

감사의 글

※ 이 논문은 한국과학재단의 연구과제지원에 의해 수행되었으며 이에 감사를 드린다. 아울러 영상을 수집하고 디지털화하는데 크게 수고를 해준 부경대학교 정보통신공학과 광정보처리 연구실의 신동학, 심영미, 신상호 씨에게도 감사를 드린다.

참 고 문 헌

- [ 1 ] Z. Q. Hong, "Algebraic feature extraction of images for recognition," *Pattern Recognition*, vol. 24, pp. 211-219, 1991.
- [ 2 ] R. Chellappa, C. L. Wilson and S. Sirohey, "Human and machine recognition," *Proc. IEEE*, vol. 86, pp. 705-736, 1995.
- [ 3 ] Y. Q. Cheng, K. Liu, J. Y. Yang and N. C. Gu, "Humanface recognition method based on the statistical model of small sample size," *SPIE of Intelligent Robots and Computer vision X*, vol. 1607, pp. 85-95, 1991.
- [ 4 ] C. H. Teh and R. H. Chin, "On image analysis by the method of moments," *IEEE Trans. Patt. Anal. Machine Intell.* vol. 10, pp. 496-513, 1988.
- [ 5 ] Y. K. Ham, S. Y. Lee, and R. H. Park, "Knowledge-based face recognition using neural networks," in *Proc. Int. Joint Conf. Neural Networks*, Nagoya, Japan, pp. 2125-2128, 1993.
- [ 6 ] R. J. Baron, "Mechanisms of human facial recognition," *Int. J. Man Machine Studies*, vol. 15, pp. 137-178, 1981.
- [ 7 ] D. J. Burr, "Elastic matching of line drawing," *IEEE Trans. Patt. Anal. Machine Intell.* vol. 3, pp. 1006-1015, 1981.
- [ 8 ] A. L. Yuille, "Deformable templates for face recognition," *J. Cognitive Neuroscience*, vol. 3, pp. 59-70, 1991.
- [ 9 ] M. Turk and A. Pentland, "Eigenface for recognition," *J. Cognitive Neuroscience*, vol. 3, pp. 71-86, 1991.

- [10] A. J. O'Toole, H. Abdi, K. A. Deffenbacher and D. Valentin, "Low dimensional representation of faces in high dimensions of the face space," *J. Opt. Soc. Am. A*, vol. 10, pp. 405-411, 1993.
- [11] L. Sirovich and M. Kirby, "Low dimensional procedure for the characterization of human faces," *J. Opt. Soc. Am.* vol 4, pp. 519-524, 1987.
- [12] E. Saund, "A multiple cause mixture model for unsupervised learning," *Neural Computation*, vol. 7, pp. 51-71, 1994.
- [13] H. B. Barlow, "Unsupervised learning," *Neural Computation* vol.1, pp. 295-311, 1989.
- [14] N. Intrator, "Feature extraction using an unsupervised neural network," *Neural Computation*, pp 98-107, 1992.
- [15] D. J. Field "What is the goal of sensory coding," *Neural computation*, vol 6, pp. 589-601, 1994.
- [16] A. B. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, pp. 607-609, 1996.
- [17] C. G. Gross, "Representation of visual stimuli in inferior temporal cortex," *Phil. Tran. R. Soc. Lond. B*, vol. 335, pp. 3-10, 1992.
- [18] D. Tranel, A. R. Damasio and H. Damasio, "Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity," *Neurology*, vol. 38, pp. 690-696, 1988.
- [19] M. P. Young and S. Yamane, "Sparse population coding of faces in the inferotemporal cortex," *Science*, vol. 256, pp. 1327-1331, 1992.
- [20] J. P. Miller, G. A. Jacobs and F. E. Theunissen, "Representation of sensory information in the cricket cercal sensory system. I. Response properties of the primary Interneurons," *J. of Neurophysiology*, vol. 66, pp. 1680-1689, 1991.

---

 저자 소개
 

---



金鐘圭(正會員)

1978년 2월 경북대학교 전자공학과 졸업(공학사). 1980년 2월 경북대학교 대학원 전자공학과 졸업(공학석사). 1992년 3월 ~ 현재 경상대학교 전자공학과 박사과정. 1983년 4월 ~ 1985년 2월 금성전기 기술연구소 연구원. 1985년 3월 ~ 현재 연암공업전문대 교수. 주관심분야는 음성신호처리, 패턴인식 등임



金榮一(正會員)

1979년 2월 경북대학교 전자공학과 졸업(공학사). 1981년 8월 연세대학교 대학원 전자공학과 졸업(공학석사). 1985년 2월 연세대학교 대학원 전자공학과 졸업(공학박사). 1987년 4월 ~ 현재 경상대학교 전자공학과 교수. 주관심분야는 음성신호처리, 실시간 디지털신호처리 등임

張珠錫(正會員)

1961년 6월 12일생. 1984년 2월 부산대학교 전기기계공학과 졸업(공학사). 1986년 2월 한국과학기술원 전기전자공학과 졸업(공학석사). 1989년 8월 한국과학기술원 전기전자공학과 졸업(공학박사). 1989년 11월 ~ 1991년 4월 미국 University of Colorado (Boulder 소재) 물리학과 및 Joint Institute for Laboratory Astrophysics 연구원. 1991년 8월 ~ 1993년 2월 한국 전자통신연구소 기초기술연구부 선임연구원. 1993년 3월 ~ 1995년 3월 부산수산대학교 정보통신공학과 전임강사. 1994년 7월 ~ 1995년 7월 미국 California Institute of Technology (Pasadena 소재) Computation and Neural System 객원연구원. 1995년 4월 ~ 현재 부경대학교 정보통신공학과 조교수. 주관심분야는 광정보처리, 광메모리, 광패턴인식, 신경정보처리 모델링 등임