

論文98-35C-3-10

Fourier 변환된 얼굴의 진폭스펙트럼의 Karhunen-Loeve 근사 방법에 기초한 변위불변적 얼굴인식

(Shift-invariant face recognition based on the Karhunen-
Loeve approximation of amplitude spectra of
Fourier-transformed faces)

沈 英 美 * , 張 珠 錫 * , 金 鐘 圭 **

(Young-Mi Shim, Ju-Seog Jang, and Jong-Kyu Kim)

요 약

Karhunen-Loeve 근사방법에 기초한 얼굴인식을 수행함에 있어서 얼굴영상을 직접사용하는 대신 Fourier 변환된 얼굴의 진폭스펙트럼을 사용하였다. Fourier 변환된 얼굴의 진폭스펙트럼을 얼굴인식에 이용하면 입력얼굴의 이동에 대해 불변적 성질을 가질뿐만 아니라 한 사람의 변화된 얼굴사이의 거리가 다른 사람과의 거리보다 더욱 가까워지므로 변화된 얼굴 자체에 대한 인식률이 향상됨을 알았다. 우리 방법의 유용성을 보이기 위해 남자 55명과 여자 25명에 대해 각각 11개씩의 변화된 얼굴로 컴퓨터 모의 실험을 하였다. 얼굴영상의 Fourier 변환의 진폭스펙트럼을 사용하는 것이 기존의 실제 얼굴영상을 사용하는 것에 비해 변화된 얼굴이나 이동된 얼굴에 대해서도 우수한 인식률을 가짐을 확인하였다.

Abstract

In face recognition based on the Karhunen-Loeve approximation, amplitude spectra of Fourier transformed facial images were used. We found that the use of amplitude spectra gives not only the shift-invariance property but also some improvement of recognition rates. This is because the distance between the varying faces of a person compared with that between the different persons is reduced further in general in spectral domain. To show the feasibility of our approach, we performed computer experiments on face recognition with varying facial images obtained from total 55 male and 25 females. We confirmed that the use of amplitude spectra of Fourier-transformed facial images gives better recognition rate for a variety of varying facial images including shifted ones than the use of direct facial images does.

I. 서 론

인간의 가장 놀라운 능력 중의 하나인 얼굴인식은 어릴 때부터 발달하며, 인간의 사회생활에 매우 중요한 역할을 한다. 인간수준의 얼굴인식 능력을 컴퓨터에 구현시켜 여러 응용분야에 적용하려는 연구가 활발히 이루어지고 있다. 컴퓨터에 입력되는 임의의 영상으로부터 자동적으로 얼굴을 인식하는 기술은 영상 처리나 패턴 인식, 컴퓨터 비전과 신경망 같은 다양한

* 正會員, 釜慶大學校 情報通信工學科

(Department of Telematics Engineering Pukyong National University)

** 正會員, 蓮庵工業專門大 電子工學科

(Department of Electronics Yonam College of Engineering)

接受日: 1998年1月12日, 수정완료일: 1998年2月26日

분야에 걸쳐서 활발히 연구되고 있으며, 또한 보안에 대한 중요성이 크게 부각되면서 얼굴인식에 대한 관심이 점점 높아지고 있다^[1-7].

최근 Karhunen-Loeve 근사방법^[8-10]을 얼굴의 특징을 추출하는데 사용하고 이들 특징으로부터 얼굴을 인식하는 방법이 제안되었다^[1-3]. 이 방법은 인식하고자 하는 얼굴 샘플(sample)들의 공분산 행렬의 고유벡터를 계산하고, 새롭게 입력된 얼굴을 고유벡터 공간으로 투사시켜 그 성분들을 비교함으로써 인식을 수행한다. 고유벡터는 간단히 계산할 수 있으며 고유벡터의 개수 즉 인식을 위한 비교차원이 얼굴 영상의 픽셀개수에 비해 매우 작아서 패턴인식이 대단히 간단해진다. 이 방법은 패턴인식의 단순성에 비해 비교적 높은 인식률이 얻어지는 것이 큰 장점이다. 그리고 이 방법을 다소 수정하여 인식률을 더욱 개선한 것으로 MDF (most discriminating feature) 방법이 제안되었다^[4]. 이는 샘플로 선정된 얼굴에 대해 그룹내에서의 동일한 사람의 여러 변화된 얼굴과 샘플간 거리와 서로 다른 사람의 얼굴 그룹간의 거리를 보다 분명히 함으로써 그룹간의 차별을 극대화하는 방법이다. Karhunen-Loeve 근사방법을 MDF와 대조시켜 MEF (Most Expressive Feature) 방법이라고 부를 수 있다^[3, 4].

이러한 방법의 얼굴인식에 있어서 한가지 중요한 문제는 인식하고자 하는 얼굴영상의 위치 이동에 대해 인식률이 급격히 감소한다는 것이다. 본 논문에서는 이 문제들을 해결할 수 있는 간단한 방법으로 얼굴영상을 직접 사용하는 대신 Fourier 변환된 얼굴의 진폭 스펙트럼을 사용하였다. 우리는 이러한 간단한 처리만으로 입력 얼굴영상들의 위치 이동에 무관할 뿐만 아니라 인식률도 더욱 개선할 수 있음을 처음으로 보였다.

II. Karhunen-Loeve 근사방법과 얼굴인식

1. Karhunen-Loeve 근사방법

Karhunen-Loeve 근사방법^[8-10]은 주성분 분석법 (principal component analysis)이라고도 하는데 통계적으로 변화가 있는 N 차원의 M 개 벡터를 그들의 공분산 행렬의 고유벡터로 표현하는 것이다. 즉 정규화된 직교 고유벡터를 기저(basis) 벡터로 사용하여

입의 영상을 이들 고유벡터의 가중치 조합으로 표현한다. 이때, 고유값이 큰 몇 개의 고유벡터만 사용하여도 본래의 벡터를 잘 표현할 수 있으며 사용할 고유벡터의 개수를 증가시키면 본래 영상을 더욱 정확히 나타낼 수 있다.

N 차원의 M 개 벡터를 $X_m, m = 1, 2, \dots, M$ 이라 하자. 이들의 공분산 행렬은 $C = 1/M \sum_{m=1}^M X_m X_m^T$ 이고, 이는 실수 대칭행렬이므로 서로 다른 고유값에 대해 M 개 직교 고유벡터를 얻을 수 있다. T 는 벡터의 전치치를 의미한다.

고유벡터를 정규화한 것을 U_m 이라 하고 이들의 고유값을 λ_m 이라 하자. 이때 $U_m^T U_n = \delta_{mn}$ 이고 δ_{mn} 은 Kronecker의 δ 이다. 단 index m 은 고유값이 큰 순서대로 매김되었다고 가정한다. 이때 어떤 N 차원의 임의의 입력벡터 X 는

$$X = \sum_{m=1}^M \omega_m U_m \tag{1}$$

으로 나타내어진다. 이때 ω_m 은 X 를 U_m 공간으로 투사했을 때의 성분이 되며 $\omega_m = U_m^T X$ 로 계산된다. 그리고 $M' < M$ 에 대해 $X' \approx \sum_{m=1}^{M'} \omega_m U_m$ 으로 구할 수 있고, 이는 주어진 M' 에 대해서 오차 $\|X - X'\|$ 를 최소로 한 것이다^[10].

2. 얼굴영상을 직접 사용한 인식

(1) 고유얼굴 방식

앞의 방법을 얼굴 영상에 직접 적용한 얼굴인식은 다음과 같이 수행된다. 편의상 얼굴영상의 평균(공통성분)을 제거하여 사용한다. $N_1 \times N_2 (= N)$ 개의 픽셀을 갖는 M 개의 2차원 정면 얼굴영상을 N 차원 열벡터로 표현한다. 이들 각각의 영상을 $F_m (m = 1, 2, \dots, M)$ 이라 하자. 이들의 평균 벡터는 $\bar{F} = (1/M) \sum_{m=1}^M F_m$ 이다. 평균이 제거된 이들 벡터 $F_m - \bar{F}$ 를 ψ_m 이라 표현하기로 하자. 이들의 공분산 행렬은 $C = (1/M) \sum_{m=1}^M \psi_m \psi_m^T$ 이고, C 의 차원은 $N \times N (= N_1^2 \times N_2^2)$ 으로 고유벡터와 고유값을 구하기 위한 계산량이 매우 많아진다. 따라서 다음과 같은 과정을 거쳐서 고유벡터를 구한다^[11].

먼저 행렬 $\Psi = [\psi_1 \ \psi_2 \ \dots \ \psi_M]$ 로부터 $\Psi^T \Psi$ 를 계산한다. 이 행렬의 차원은 $M \times M$ 이다. 이 행렬의 고유값과 고유벡터를 λ_i 및 V_i 라 하면

$$\Psi^T \Psi V_i = \lambda_i V_i \quad (2)$$

이고 식 (2)의 양변에 Ψ 를 곱하면

$$\Psi \Psi^T \Psi V_i = \lambda_i \Psi V_i \quad (3)$$

된다. 따라서 $\Psi \Psi^T = C$ 의 고유벡터는 ΨV_i 이고 고유값은 λ_i 이다. 즉 $U_i = \Psi V_i$ 이다. U_i 는 정규화된 것으로 간주한다. 얻어진 고유벡터를 $N_1 \times N_2$ 로 재배열하여 보면 얼굴영상과 닮아있으므로 “고유얼굴”(eigenface)라고 불린다. 그리고 평균값을 뺀 Ψ_m 을 사용하므로 고유값이 0이 아닌 고유벡터는 $M-1$ 개다.

얼굴인식은 다음과 같이 수행된다. 고유값이 큰 순서대로 $M' (< M)$ 개의 U_i 를 선택하고 U_i 를 구할 때 사용하였던 즉, 학습할 때 사용한 M 개의 Ψ_m 에 대한 특징벡터 즉,

$$\Omega_m = [U_1^T \Psi_m \ U_2^T \Psi_m \ \dots \ U_{M'}^T \Psi_m]^T, \quad m = 1, 2, \dots, M \quad (4)$$

을 저장해 둔다. 특징벡터 Ω_m 은 Ψ_m 을 U_i 공간에 투사시킬 때의 성분벡터이다. 실제 인식과정에서는 학습되지 않은 다소 변화된 입력패턴 F_{in} 에 대해서 $\Psi_{in} = F_{in} - \bar{F}$ 를 얻은 다음 식 (4)와 같이 특징을 추출하고, 즉 특징벡터 Ω_{in} 을 얻고 이를 이미 저장된 모든 M 개 Ω_m 과 비교하여 그 유클리드 거리가 최소가 되는 Ψ_m 을 찾는다. 이 방법은 M' 차원의 특징벡터공간에서 비교하므로 $N_1 \times N_2$ 화소를 갖는 얼굴영역에서 비교하는 것에 비해 계산량이 매우 줄어든다. 그리고 인식률을 높이기 위해 각 사람마다 1개 이상의 변화된 얼굴을 학습에 사용할 수도 있다^[3].

(2) Most Discriminating Feature (MDF) 방법

고유얼굴을 이용한 인식방법이 인식하고자 하는 얼굴을 최소 오류로 표현할 수 있지만^[10] 어떤 사람들의 얼굴들에 대한 차별을 가장 잘 표현한다고 볼 수는 없다. 이런 얼굴들간의 차별을 보다 뚜렷히 할 수 있는 방법으로 MDF 방법이 제안되었다^[4].

여기서는 정면 얼굴 1개만을 사용하 할 수는 없고 각 사람마다 서로 조금 변화된 얼굴을 2개 이상 사용한다. 인식하고자 하는 사람의 수를 M , 각 사람마다 서로 다른 K 개의 얼굴 영상을 사용한다고 하자. 그

러면 앞서서의 마찬가지로 총 $M \times K$ 얼굴영상에 대한 고유벡터 $U_i (i = 1, 2, \dots, MK-1)$ 를 구한다. 그리고 각각의 얼굴영상에 대해 식 (4)에서 구한 특징벡터를 $\Omega_m^k = [U_1^T \Psi_m^k \ U_2^T \Psi_m^k \ \dots \ U_{MK-1}^T \Psi_m^k]$ 라하고, m 번째 사람에 대한 Ω_m^k 의 평균을 $\bar{\Omega}_m = 1/k \sum_k \Omega_m^k$ 이라 하자. 그러면 각 사람 자신들간의 scatter matrix는

$$S_w = \sum_{m=1}^M \sum_{k=1}^K (\Omega_m^k - \bar{\Omega}_m) (\Omega_m^k - \bar{\Omega}_m)^T \quad (5)$$

로 정의된다. 그리고 한 사람과 서로 다른 사람들 간의 scatter matrix는

$$S_b = \sum_{m=1}^M (\bar{\Omega}_m - \bar{\Omega}) (\bar{\Omega}_m - \bar{\Omega})^T \quad (6)$$

로 정의된다. $\bar{\Omega}$ 는 전체 평균 $\frac{1}{M} \sum_m \bar{\Omega}_m$ 이다. 이때 $(S_w)^{-1} S_b$ 를 계산하고 이것의 고유벡터를 인식하기 위한 기저(basis)벡터 W_j 로 사용한다. W_j 의 차원은 $(MK-1) \times 1$ 이다. 그러나 Karhunen-Loeve 근사 방법에서 $MK-1$ 보다 작은 수 M' 개의 고유벡터 U_i 를 사용해도 된다. 이 경우의 W_j 의 차원은 $M' \times 1$ 이 된다.

얼굴의 인식은 다음과 같이 수행된다. 어떤 얼굴영상 F_{in} 이 입력되면 $\Psi_{in} = F_{in} - \bar{F}$ 를 계산하고 $(U_i)^T \Psi_{in}$ 을 모든 i 에 대해 계산하여 특징벡터 Ω_{in} 을 얻는다. 그리고 $W_j^T \Omega_{in}$ 을 모든 $j = 1, 2, \dots, M'$ 에 대해 계산하여 제2의 특징벡터 $Z_{in} = [W_1^T \Omega_{in} \ W_2^T \Omega_{in} \ \dots \ W_{M'}^T \Omega_{in}]^T$ 을 얻고 이것과 학습에 사용된 Ψ_m 의 Z_m 간의 유클리드 거리가 최소가 되는 것을 찾음으로써 인식을 수행한다.

그러면 MDF 방법은 마치 두 단계의 과정을 거쳐 인식을 수행하는 것 같지만 실질적으로 MDF에 대한 M' 개의 벡터 $[U_1 \ U_2 \ \dots \ U_{M'}] W_j = V_j (j=1, 2, \dots, M')$ 을 계산해 두면 앞의 고유얼굴 방식과 마찬가지로 한 번의 과정 즉 $Z_{in} = [V_1^T \Psi_{in} \ V_2^T \Psi_{in} \ \dots \ V_{M'}^T \Psi_{in}]^T$ 을 바로 계산함으로써 인식을 수행할 수 있다. 이 V_j 를, 고유얼굴 방식에서의 특징얼굴(고유벡터) U_i 에 대응되는, MDF 방식에서의 특징얼굴이라고 볼 수 있다.

3. Fourier 변환된 얼굴 영상의 진폭스펙트럼을 사용한 인식

앞 절에서, 원래의 얼굴을 직접 얼굴인식에 이용할

경우, 입력얼굴의 위치변화에 매우 민감한 특성을 가진다. 실제 얼굴인식은 카메라로 입력되는 영상을 컴퓨터에서 처리하므로 사람이 카메라의 중앙을 맞추기가 쉽지 않다. 따라서 입력 영상의 변위에 대해 둔감한 얼굴인식 시스템이 요구되므로, 본 연구에서는 입력얼굴의 광범위한 변위불변적 인식을 위해서 Fourier 변환의 진폭 스펙트럼을 사용하였다.

어떤 입력얼굴 $g(x, y)$ 의 Fourier 변환한 결과 $G(\xi, \eta)$ 는

$$G(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) \exp[-j2\pi(\xi x + \eta y)] dx dy \quad (8)$$

이다. 여기서 ξ 와 η 는 각각 x 축, y 축의 공간주파수이다. 이때 입력얼굴이 임의의 방향으로 변위된 경우 $g(x-a, y-b)$ 가 입력되면 Fourier 변환결과는

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x-a, y-b) \exp[-j2\pi(x\xi + y\eta)] dx dy \\ = G(\xi, \eta) \exp[-j2\pi(\xi a + \eta b)] \end{aligned} \quad (9)$$

이다. 그러나 식(15)의 스펙트럼의 크기만을 고려하게 되면 그 결과는

$$\mathcal{F} [v(x-a, y-b)] = |G(\xi, \eta)| \quad (10)$$

이다. 그러므로 Fourier 변환의 진폭 스펙트럼은 입력 영상의 변위에 상관없이 항상 일정한 값을 가지므로 변위불변적 성질을 가진다.

한편 분수차 Fourier 변환(Fractional Fourier Transform)의 개념이 개발된 바 있다^[11]. 분수차, 예를 들어 $1/p$ 차 Fourier 변환(fractional Fourier transform)은 이를 p 번 반복 변환하였을 때 종래의 Fourier 변환이 되는 것을 의미한다. 이러한 분수차 Fourier 변환된 결과는(분수차 Fourier 변환을 1번 수행한 것은) 입력영상의 위치 정보와 공간주파수 정보가 차수에 따라 적절히 혼합된 것으로 공간 가변적(space-variant) 성질을 지닌다. 하지만, 이 분수차 Fourier 변환의 개념을 얼굴인식에 적용시켜 변위 불변적 고유얼굴 특징을 추출하는 필터를 설계한 결과, 변환 차수 p 가 0.5 근처에서는 이동된 입력얼굴에 대해서 인식률이 다소 떨어지고 변환 차수가 0.95 이상에서만 변위 불변적 성질과 비교적 높은 인식률을 동시에 만족시켰다. 변환 차수가 0.95인 것은 위치정보

를 포함하고 있다고 보기 힘들므로 본 논문에서는 일반적인 Fourier 변환($p = 1$)의 진폭 스펙트럼을 사용하였다.

한 예로 그림 1에 본래얼굴 및 이동된 얼굴에 대해 Fast Fourier transform (FFT) 알고리즘을 이용한 Fourier 변환된 얼굴의 진폭을 나타내었다. FFT는 컴퓨터로 비교적 쉽게 계산이 된다. 진폭스펙트럼의 차원을 본래 영상 픽셀들의 차원 $N_1 \times N_2$ 와 동일하게 맞춘 후 F_m 으로 사용하여, 앞 II.2의 (1) 및 (2)와 같은 방식으로 인식실험을 수행한다.

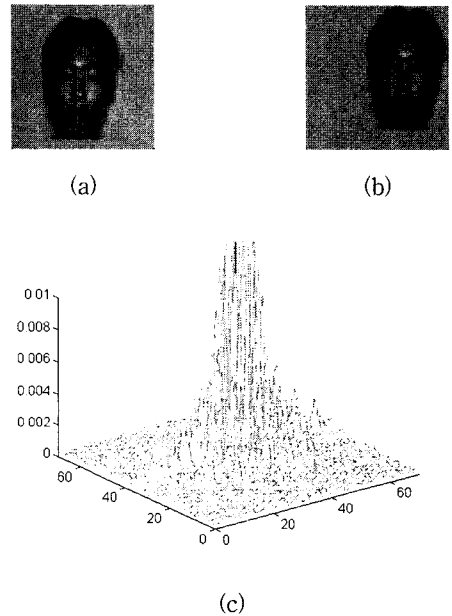


그림 1. Fourier 변환 진폭 스펙트럼의 변위 불변적 특징 (a) 좌로 10 픽셀 이동한 입력영상. (b) 상과 우로 10픽셀씩 이동한 입력영상. (c) (a)와 (b)의 Fourier 변환 진폭 스펙트럼. (가운데 dc 근처 성분은 잘랐음)

Fig. 1. The shift-invariant characteristic of amplitude spectrum of the Fourier transform. (a) An input image shifted by 10 pixels left. (b) An input image shifted by 10 pixels both right and up. (c) The amplitude spectrum of the Fourier transform of both (a) and (b) images.

III. 얼굴인식에 대한 컴퓨터 실험

1. 얼굴의 데이터베이스

얼굴인식에 대한 대부분의 연구들에 있어서, 여러가지 제약조건을 가하여 사용할 얼굴영상을 획득하고,

또한 100개 미만의 적은 데이터 집합에 대해서만 인식실험을 수행하였다. 본 연구에서는 인식률의 통계적 특성을 구하기 위해서 많은 수의 실험 얼굴을 수집하였다. 80명의 서로 다른 사람들에 대해 각각 서로 다른 11개씩의 얼굴을 CCD 카메라와 Frame Grabber를 이용하여 컴퓨터에 저장하고 이들을 0 ~ 255의 수치로 정규화하였다. 이 80명의 정면 얼굴을 그림 2에 나타내었다. 이때 얼굴영상의 차원은 모두 60×40 이다. 수집된 남자 55, 여자 25명의 얼굴을 각각 M

1~M80으로 표현하였다. 각 사람에 대한 변화된 11개의 얼굴의 예는 그림 3와 같다. 인식결과의 통계를 얻기위하여 80명의 얼굴을 20명씩 11개의 그룹으로 임의로 나누었다. 한 그룹에 속한 얼굴이 다른 그룹에도 속할 수 있다. 각 그룹마다 정면 얼굴하나와 변환된 얼굴 1개로 학습하고 나머지 변화된 얼굴로 인식을 수행한다. 그리고 전체 11개 그룹별로 독립적으로 수행된 실험결과를 평균하여 인식률을 얻었다.

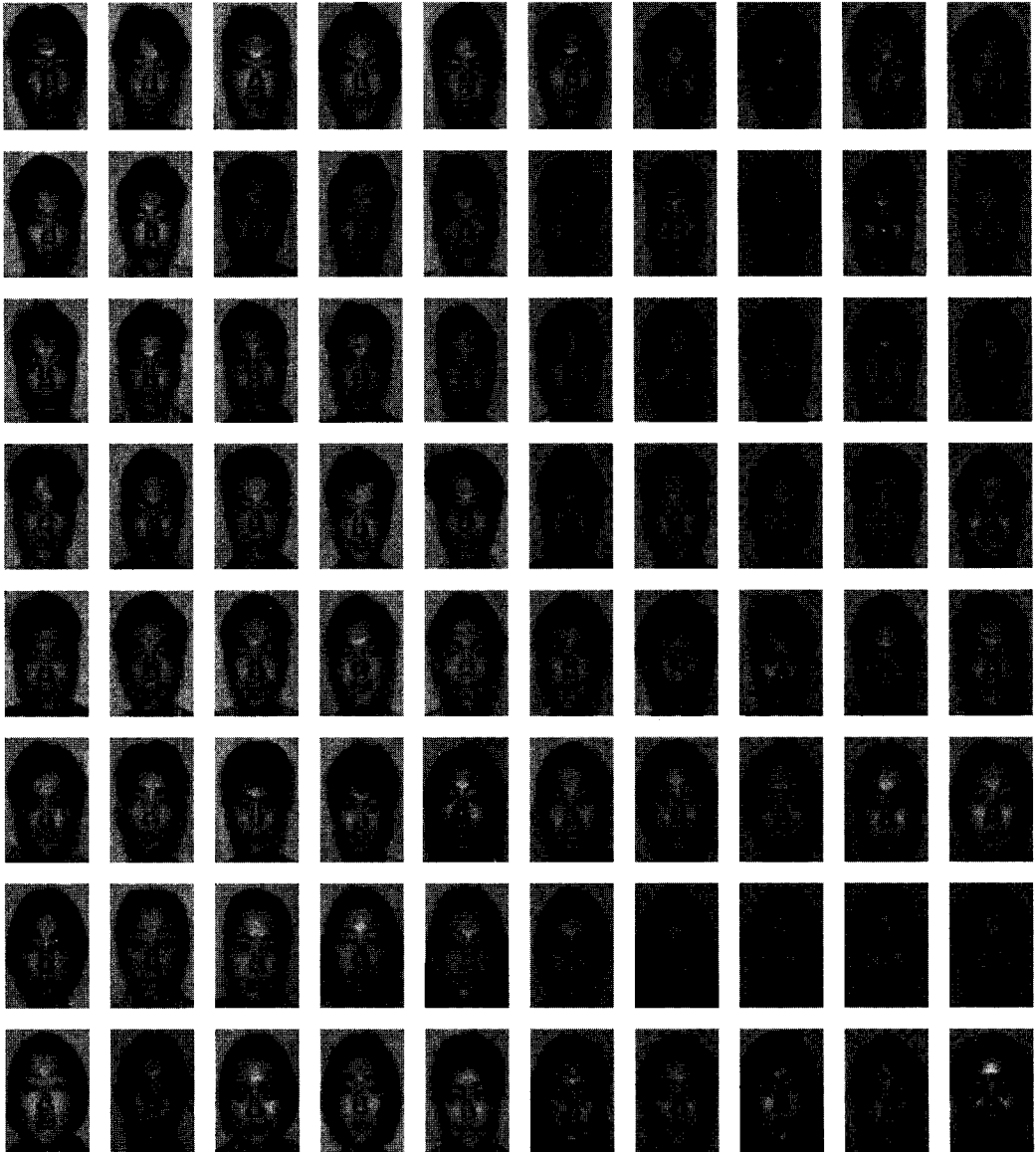


그림 2. 실험에 사용한 80명의 정면 얼굴

Fig. 2. Eighty front faces used in experiment.

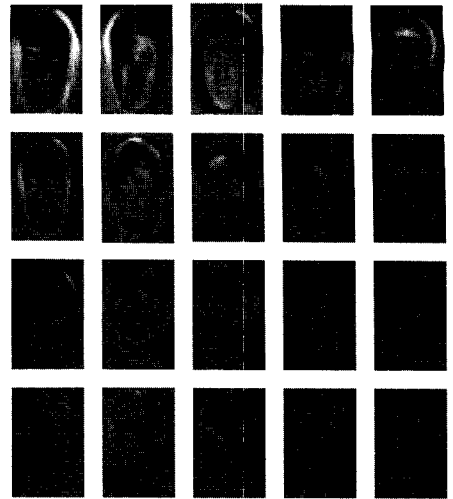
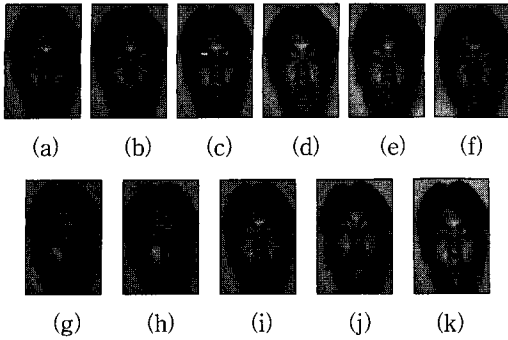


그림 3. M1의 변화된 얼굴 11가지 (a) 정면얼굴 (b) 눈을 감은 정면얼굴 (c) 안경을 낀 정면얼굴 (d) 웃는 얼굴 (e) 화난 얼굴 (f) (a)와 시간차가 나는 정면얼굴 (g) 및 (h) 오른쪽으로 회전한 얼굴 (i) 및 (j) 위로 회전한 얼굴 (k) 시간차가 나는 정면얼굴

Fig. 3. Eleven varying faces for M1. (a) Front face. (b) Front face with closed eyes. (c) Front face with glasses. (d) Front face with a smile. (e) Angry front face. (f) A front face taken at different time. (g) and (h) Rotated faces. (i) and (j) Upward faces. (k) Another front face taken at different time.

2. 인식 실험 및 인식률 비교

얼굴영상을 직접 사용한 경우와 Fourier변환 영상의 진폭스펙트럼을 사용한 경우에 대해서 각각 고유얼굴(eigenface)방법 및 MDF방법을 동일한 조건아래에서 실험하고 이들 4가지 방법을 비교하였다. 먼저 얼굴영상을 직접 사용한 경우, 얼굴영상 M1~M20이 포함된 한 그룹(그룹1)에 대해 한 사람당 2개씩 40개 얼굴을 이용하여 고유얼굴방법과 MDF 방법으로 구해지는 고유값과 고유얼굴(혹은 특징얼굴)을 그림 4에 나타내었다.

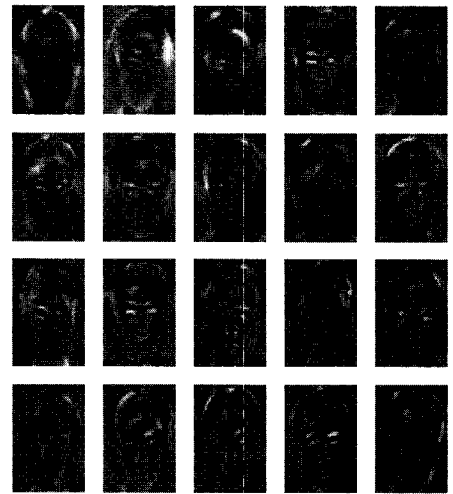
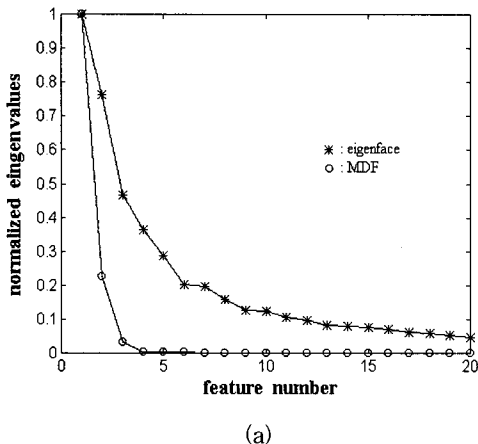


그림 4. 그룹 1에 대한 실험. (a)고유얼굴방법에서 고유벡터 U_i 의 고유값 분포 및 MDF 방법에서 W_i 의 고유값 분포. (b) 고유얼굴의 예. (c) MDF로부터 얻어진 특징얼굴의 예

Fig. 4. Experiment on group 1. (a) The distribution of eigenvalues for U_i when the eigenface method is used and those for W_i when the MDF method is used. (b) Examples of eigenfaces. (c) Examples of feature faces for the MDF method.

그림 4(a)는 고유얼굴방법과 MDF 방법으로 구해지는 고유값 중에서 고유값이 큰 20개의 정규화된 분포이며, 그림 4(b)는 고유얼굴방법으로부터 얻어지는 고유얼굴이고, 그림 4(c)는 MDF 방법으로부터 얻어

지는 특징얼굴이다. MDF특징얼굴과 고유얼굴은 뚜렷한 차이가 있으며, MDF 특징벡터의 윤곽이 더욱 뚜렷함을 볼 수 있다.

MDF 방법은 각 사람의 얼굴영상에 대해 K 개 변화된 영상들의 집합을 어떻게 설정하는가에 따라서 인식률의 차이가 난다. 본 논문에서의 실험에서는 정면얼굴과 변화된 여러 얼굴들 중에서 1개만 선택하여 $K = 2$ 인 경우에 대해 인식률을 비교해 보았다. 결과는 그림 5와 같다. 인식률이 가장 높은 집합은 그림 3에서 기준얼굴(a)와 우로 많이 돌린 얼굴(g)을 선정하였을 경우이다. 따라서 각 그룹마다 (a)와 (g) 두 개의 얼굴 영상을 하나의 집합으로 사용하였다. 동일한 조건에서 고유얼굴방법과 MDF 방법의 인식률을 조사하기 위해 고유얼굴방법에서도 MDF 방법에서 사용한 (a) 및 (g)의 얼굴들을 고유얼굴을 구하는데 사용하였다. 얼굴인식 실험결과를 그림 6에 나타내었다. 최고 인식율은 각각 85.9 %, 96.1 % 이었다. 우리의 얼굴 데이터베이스에서는 두 개씩의 얼굴로 학습시킨 고유얼굴방법이 하나의 얼굴만으로 학습시킨 고유얼굴을 이용한 것보다 평균 인식률이 약 4% 우수하였다. 하지만 MDF 방법을 이용한 결과는 2개의 얼굴을 학습시킨 고유얼굴방법의 결과보다 평균적으로 약 7% 더 우수하였다.

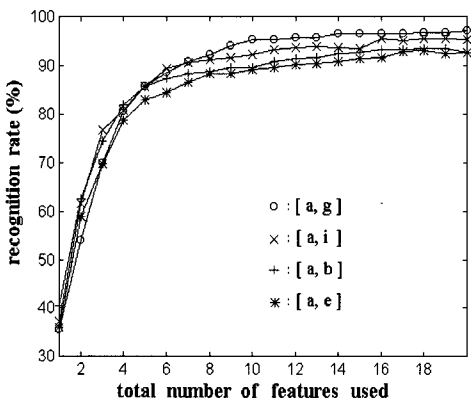


그림 5. 학습할 얼굴의 집합선정에 따른 MDF 얼굴 인식 결과

Fig. 5. Recognition rates for MDF according to two different training facial images selected.

한편 그림 6은 인식에 사용된 고유얼굴의 개수(M' 혹은 M'')에 따른 인식률을 보여주고 있다. 사용된

고유벡터에 대응되는 고유값의 총합이 전체의 90% 이상이면 인식률에 큰 변화를 가지지 않음을 알 수 있다. 이때의 고유벡터수는 평균적으로 13개정도였다. 따라서 본 논문에서는 계산량을 줄이기 위해서 4가지 인식실험에서 고유값이 90%범위인 고유얼굴을 사용하여 인식실험을 하였다.

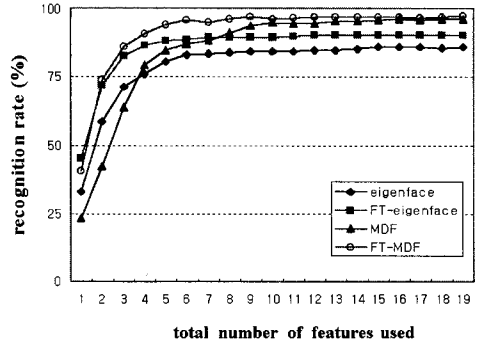


그림 6. 원래 얼굴을 사용할 때와 Fourier 변환 스펙트럼을 사용할 때의 인식률 비교

Fig. 6. Comparison of recognition rates when the original facial images and the amplitude spectra of their Fourier transform are used, respectively.

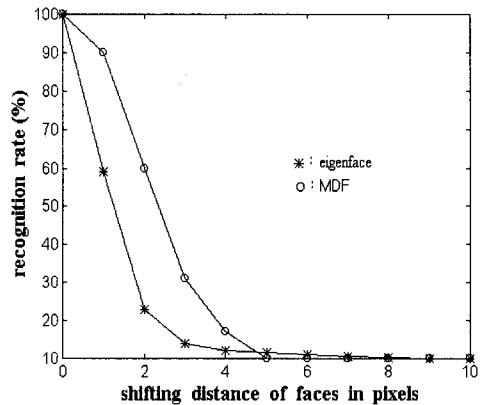


그림 7. 변위된 얼굴에 대한 인식률 감소

Fig. 7. Decrease of recognition rates for shifted faces.

일반적으로 카메라에 입력되는 얼굴은 항상 카메라의 중앙에 위치할 수 없으며, 상하 좌우의 이동을 가지기 쉽다. 따라서 고유얼굴방법과 MDF방법에 대해서 이동된 얼굴에 대해서도 인식률을 조사하였다. 그림 7에서는 앞 실험에서 학습시켰을 때 사용한 얼굴영

상의 위치이동에 대한 인식을 감소추이를 나타낸 것이다. 고유얼굴방식과 MDF방식은 좌측으로 약 2픽셀의 위치변화에도 인식이 크게 떨어짐을 알 수 있다.

표 1. 각 그룹에 대해 원래 얼굴간의 평균 거리비 및 얼굴의 진폭스펙트럼에서의 평균거리비

Table 1. Average distance ratios of original facial images and amplitude spectral images for each group.

Group	$\frac{D_{bo}}{D_{wo}}$	$\frac{D_{bf}}{D_{wf}}$
group 1	2.92	6.83
group 2	2.36	3.79
group 3	2.88	6.20
group 4	3.28	5.53
group 5	2.20	4.66
group 6	2.64	4.74
group 7	3.39	6.48
group 8	3.69	7.76
group 9	3.56	7.06
group 10	3.33	6.95
group 11	3.90	7.50

각각의 입력영상을 Fourier 변환하면 그 스펙트럼의 분포가 매우 유사하게 나타난다. 그러므로 N 차원 공간에서 학습하고자 하는 얼굴간의 Euclid 거리는 더 가까워진다. 이를 보기 위해 우리의 얼굴 데이터베이스의 각 그룹에 대해서 한 사람과 다른 사람 얼굴간의 평균 거리를 D_{bo} 로 나타내고, 또한 한 사람의 기준얼굴과 변환된 얼굴간의 평균 거리를 D_{wo} 로 나타내기로 하자. 마찬가지로 Fourier 변환된 얼굴의 진폭스펙트럼에 대해서도 그 평균 거리를 D_{bf} 및 D_{wf} 로 나타내기로 한다. 이때 11개의 그룹에 대해 D_{bf}/D_{wf} 와 D_{bo}/D_{wo} 를 측정하여 표 1에 나타내었다. 표 1에서 Fourier 변환된 얼굴의 비 D_{bf}/D_{wf} 가 원래 얼굴사이의 비 D_{bo}/D_{wo} 보다 항상 더 크다는 것을 알 수 있다. 그리고 한가지 예로 그룹 1에 포함된 20명 각각의 얼굴영상에 대해서 D_{bo}, D_{wo}, D_{bf} 그리고 D_{wf} 를 구하고 각 얼굴에 대해서 이들의 비를 그림 8에 나타내었다. 위의 결과로부터 Fourier변환의 진폭스펙트럼을 이용하는 것이 서로 다른 얼굴간의 거리가 줄어들지만 한 사람 자신의

변화된 얼굴에 비해서 상대적으로 거리가 더욱 멀어짐을 알 수 있다. 따라서 Fourier 변환의 진폭스펙트럼을 인식에 이용하는 것이 변위불변 인식외에도 오히려 인식을 높이는 효과가 있음을 예측할 수 있다.

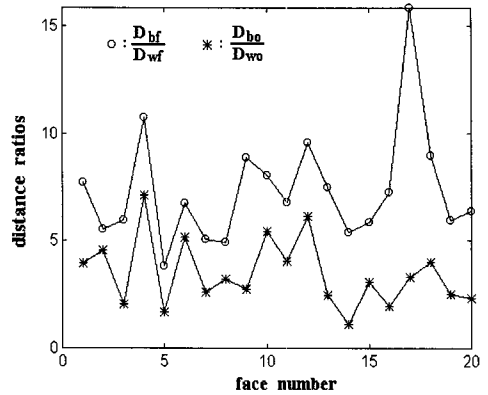


그림 8. 그룹1에서 본래 얼굴영상과 Fourier 변환된 영상의 거리비 변화

Fig. 8. Variations of distance ratios of the original facial images and the Fourier transformed faces in group1.

준비된 얼굴데이터베이스에 대해 위의 알고리즘들을 이용하여 11개의 그룹에 대해서 통계적 결과를 얻었다. 이들 역시 앞의 그림 6에 함께 나타내었다. Fourier 변환 진폭스펙트럼을 고유얼굴과 MDF 방법에 적용시킨 경우 최대 인식률은 각각 90.5%, 97.7% 였다. 따라서 Fourier 변환의 선처리한 것이 원 영상을 사용한 경우보다는 더 좋은 인식결과를 주었다. 이 실험결과는 변위된 대상에 대해서는 고려하지 않은 수치이다. Fourier 변환된 영상의 스펙트럼은 변위불변적이므로 변위되지 않은 영상에 대해서 인식하는 것은 변위된 후에도 반드시 인식을 한다. 그리고 MDF 방법에서 2개 이상의 얼굴로 구성된 집합을 학습시킬 때, 방향 변화가 있는 얼굴외에도 크기에 변화가 약간 있는 얼굴을 포함시키면, 크기변화가 다소 있는 얼굴들도 인식할 수 있을 것이다.

3. 다른 얼굴 데이터 베이스에 대한 인식실험

본 논문에서 제안하는 알고리즘에 대한 유용함을 알아보기 위해 다른 얼굴 데이터베이스에 대해서도 인식을 수행하였다. 준비된 얼굴데이터베이스는 Cambridge 대학의 Olivetti Research Laboratory(ORL) 그룹에서 제작한 40명의 얼굴과 각 사람마다 10개의

변화된 영상을 포함한 것이다^[12]. 이 얼굴 데이터는 약 2년간 그 실험실원을 대상으로 수집한 것으로서 인터넷을 통해서 다운로드 받았다. 이 ORL 얼굴 데이터는 우리 얼굴 데이터와 요구조건이 다르므로 10%의 얼굴 크기 변화가 있었고 상대적으로 변화된 정도가 심하였다. ORL 데이터의 기준 정면얼굴을 그림 9에 나타내었다. ORL 데이터에 대해서 앞에서 설명한 알고리즘을 적용시켰다. 한사람마다 10개의 얼굴 중 5개씩을 학습에 사용하고 나머지 5개로 인식률을 조사하였다. 그림 10은 이 결과를 나타내는 것이다. 이때의 인식률은 얼굴 영상을 직접 사용한 경우에 대해서는 각각 76.5%, 87%였고(위치가동이 없는 경우에 한함), Fourier 변환 진폭스펙트럼을 사용한 경우에 대해서는 각각 80%, 91%였다. ORL 그룹의 얼굴이 단순히 얼굴의 변화만을 가지고 있는 것이 아니라 얼굴의 크기가 다른 것들이 많은 관계로 우리의 얼굴영상 데이터 베이스를 사용한 경우에 비해 인식률이 다소 떨어짐을 확인하였다. 다소 크기변화가 있는 얼굴이더라도 MDF의 그룹 집합에 포함하여 학습시킴으로써, 얼굴 인식이 가능함을 알 수 있다. 이 결과에서도 변위된 얼굴은 인식율을 떨어뜨리지 않으므로 Fourier 변환 진폭스펙트럼을 이용한 것이 본래 얼굴영상을 사용하는 것보다 훨씬 더 유리하다.

IV. 결론

결론적으로 본 논문에서는 얼굴영상을 직접 사용한

경우의 인식시스템과 입력얼굴을 Fourier 변환한 후 이 스펙트럼을 이용하여 K-L 근사방법에 기초한 인식시스템을 구현하고 이들을 비교해 보았다.

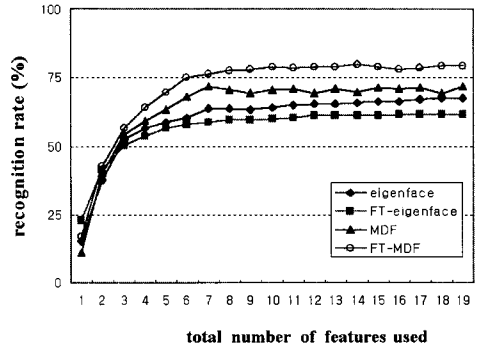


그림 10. ORL 얼굴에 대한 인식률 비교
Fig. 10. Comparison of correct recognition rates for ORL faces.

특징을 추출하는 방법으로 고유얼굴방법보다는 이것을 조금 변형시킨 MDF가 더 좋은 인식률을 가졌다. 원래 얼굴영상을 사용하면 입력영상의 위치가 조금만 변하여도 인식률이 급격히 떨어지는 것을 알 수 있었다. 우리는 단순한 FFT 알고리즘을 얼굴영상의 전처리 과정에 사용한 다음, MDF에 적용함으로써 변위불변적 인식외에도 변화된 얼굴에 대해 그 인식율을 향상시킬 수 있음을 보였다. Fourier 변환의 스펙트럼은 입력영상의 크기 변화에 다소 민감한 성질을 가지지만 이것은 MDF 방법에서 크기변화된 얼굴을 학습에 함



그림 9. ORL 얼굴집합 (정면얼굴 40개)
Fig. 9. Facial images of ORL group. (forty front faces)

게 사용함으로써 극복할 수 있음을 알았다. 본 얼굴인식에서는 처리하는 데이터량이 작아서 하나의 입력 얼굴영상을 인식하는데 기존의 펜티엄급 PC에서 FFT를 포함 약 3초가 걸렸다. 그리고 Fourier 변환을 광학적으로 렌즈하나를 이용하면 간단히 구할 수 있으므로^[13], 전체 얼굴인식을 광학적으로 구현하여 실시간 얼굴인식시스템을 실험 할 예정이다.

감사의 글

본 연구는 늘원 문화재단 및 한국과학재단 핵심전문연구비(961-0922-119-2) 지원으로 수행되었으며 지원에 감사드립니다.

참 고 논 문

[1] M. Turk and A. Pentland, "Eigenface for recognition," *J. Cognitive Neuroscience*, vol. 3, no. 1, pp. 71-86, 1991.

[2] A. J. O'Toole, H. Abdi, K. A. Deffenbacher and D. Valentin, "Low dimensional representation of faces in higher dimension of the face space," *J. Opt. Soc. Am. A*, vol. 10, no. 3, pp. 405-411, 1993.

[3] A. P. Pentland and M. A. Turk, "Face recognition using eigenfaces," in *Proc. Int. Conf. on Pattern Recognition*, pp. 586-591, 1994.

[4] D. L. Swets and J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 831-836, 1996.

[5] H. Peng and D. Zhang, "Dual eigenspace method for human face recognition," *Electron. Lett.* vol. 33, no. 4, pp. 283-284, 1997.

[6] 김종규, 장주석, 김영일, "주 시각피질에서의 단순세포 수용영역 형성에 대한 성긴 집단부호 모델을 이용한 얼굴인식," *전자공학회 논문지* 제 34권 C편 10호, pp. 43-50, 1997

[7] R. Brunelli and T. Poggio, "Face recognition: features versus templates,"

IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1042-1052, 1993.

[8] L. Sirovich and M. Kirby, "Low-dimensional procedure for the characterization of human faces," *J. Opt. Soc. Am. A*, vol. 4, no. 3, pp. 519-524, 1987.

[9] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the characterization of human faces," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 103-108, 1990.

[10] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. (Academic Press, SanDiego, 1990) chap. 9.

[11] D. Mendlovic and H. M. Ozaktas, "Fractional Fourier transforms and their optical implementation: 1," *J. Opt. Soc. Am. A*, vol. 10, pp. 1875-1881, 1993.

[12] S. Lawrence, C. L. Giles, A. C. Tsoi, "Face recognition: A convolutional neural network approach," *IEEE Trans. on Neural Networks*, vol. 8, no. 1, pp. 98-113, 1997.

[13] J. W. Goodman, *Introduction to Fourier Optics* (McGraw-Hill, SanFrancisco, 1968) chap. 5.

저 자 소 개



沈 英 美(正會員)

1974년 3월 7일생. 1996년 2월 부산
수산대학교 정보통신공학과 졸업(공
학사). 1996년 3월 ~ 현재 부경대학
교 대학원 전자공학과 석사과정. 주
관심분야는 광정보처리, 광패턴인식,
광메모리, 광연결 등임

金 鐘 圭(正會員) 第 34卷 C編 第 10號 參照

현재 연암공업전문대 교수. 주관심분
야는 음성신호처리, 패턴인식 등임

張 珠 錫(正會員) 第 34卷 B編 第 1號 參照

현재 부경대학교 정보통신공학과 조
교수. 주관심분야는 광정보처리, 광메
모리, 광패턴인식, 신경정보처리 모델
링 등임