

# Karhunen-Loeve 근사 방법과 Elastic Graph Matching을 병합한 얼굴 인식

이 형 지, 이 완 수, 정 재 호

인하대학교 전자공학과

전화 : 032-860-7420 / 핸드폰 : 019-526-4760

## Face Recognition using Karhunen-Loeve projection and Elastic Graph Matching

Hyung-Ji Lee, Wan-Su Lee, and Jae-Ho Chung

Dept. of Electronic Engineering, Inha University

E-mail : g1991177@inhavision.inha.ac.kr

### Abstract

This paper proposes a face recognition technique that effectively combines elastic graph matching (EGM) and Fisherface algorithm. EGM as one of dynamic link architecture uses not only face-shape but also the gray information of image, and Fisherface algorithm as a class specific method is robust about variations such as lighting direction and facial expression. In the proposed face recognition adopting the above two methods, the linear projection per node of an image graph reduces dimensionality of labeled graph vector and provides a feature space to be used effectively for the classification. In comparison with a conventional method, the proposed approach could obtain satisfactory results in the perspectives of recognition rates and speeds. Especially, we could get maximum recognition rate of 99.3% by leaving-one-out method for the experiments with the Yale Face Databases.

### I. 서론

정보화 시대에 맞춰 다양한 네트워크 및 정부와 기관, 개인에 대한 보안의 중요성이 대두되고 있으며, 점점 개방화되는 사회 분위기로 인하여 특정 지역의 출입통제를 위한 보안 시스템에 대한 필요성이 증대되고 있다. 이러한 이유로 사용자의 신원을 인증할 수 있는 방법이 요구되는데, 인간의 생체 특징은 그것에 대한 유일한 대안이라 할 수 있다. 특히 사용자에 게 거부감을 주지 않거나 협조적인 자세가 불가능한 경우, 얼굴을 생체 측정 수단으로 사용할 수 있다 [1].

기존의 대표적인 얼굴 인식 방법으로서 1990년대 초반 Karhunen-Loeve (K-L) 변환을 이용한 방법인 주요 성분 분석(principal component analysis, PCA) 방법이 MIT에서 제안되었으며, 이를 응용한 Fisherface 알고리즘이 등장하였다

[2][3]. 또한 신경망을 이용한 방법, 가버(Gabor) 필터 응답을 이용한 EGM (Elastic graph matching), 독립 성분 분석(independent component analysis, ICA)을 이용한 알고리즘이 소개되었다 [4][5].

본 논문은 K-L 근사 방법을 기반으로 한 Fisherface 알고리즘과 EGM 방법을 병합하여 보다 효율적인 얼굴 인식 방법을 제안하고자 한다. 제안한 알고리즘에서는 위의 두 방법을 사용하여 영상 그래프의 각 노드에 대해 Fisherface 방법을 적용함으로써 레이블 된 그래프 벡터의 차원을 줄일 뿐만 아니라 효율적으로 클래스를 구분하기 위한 특징벡터를 제공하고자 하며 실제로 이를 구현해 봄으로써 기존의 방법들과 인식을 및 인식 속도를 비교하고자 한다.

### II. Karhunen-Loeve 변환을 이용한 얼굴 인식

본 논문에서는 얼굴을 인식하기 위하여 선형 변환 이론을 사용하는데 다음과 같은 두 가지 방법이 있다. Most expressive features (MEF's)를 추출하기 위해 주요 성분 분석 또는 K-L 변환을 사용하며, most discriminating features (MDF's)를 위해 Fisherface 방법을 적용한다.

#### 2.1 주요 성분 분석 (Principal Component Analysis, PCA) 방법

주요 성분 분석 방법은 벡터를 통계학적인 성질을 이용하여 표현하는 방법이다. 즉, 식 (1)과 같이 표현되는 선형 변환의 일종이다.

$$y = A^T x \tag{1}$$

위의 식으로부터 특징 벡터  $y$ 는 원래의 벡터  $x$ 를  $A$  도메인으로 투영하여 얻어질 수 있다. 고유 성분 분석 방법은  $A$ 의 기저벡터로써  $x$ 의 공분산행렬을 사용한다. 여기에서 구해진 특징벡터들은 간단히 유클리디안 거리를 이용해서 얼굴을 인식할 수 있다.

2.2 Fisherface 방법

Fisherface 방법은 한 사람 내에서 그리고 다른 사람과의 관계를 고려한 알고리즘으로 그 방법은 다음과 같다. 우선 전체가  $N$ 개의 영상이고 사람의 수가  $c$ 이며, 한 사람당  $K$ 개의 영상을 사용한다고 가정하자. 그러면 주요 성분 분석으로부터  $N-1$ 개의 고유얼굴을 구할 수 있다. 그런 후에 클래스내 산란행렬(within-class scatter matrix,  $S_w$ )과 클래스간의 산란행렬(between-class scatter matrix,  $S_b$ )의 행렬식의 비를 최대화하기 위한  $(S_w)^{-1}S_b$ 의 고유벡터를 구한다.  $S_w$ 와  $S_b$ 는 식 (2)와 (3)처럼 정의된다.

$$S_w = \sum_{m=1}^c \sum_{k=1}^{K_m} (\mathcal{Q}_m^k - E(\mathcal{Q}_m))(\mathcal{Q}_m^k - E(\mathcal{Q}_m))^T \quad (2)$$

$$S_b = \sum_{m=1}^c (E(\mathcal{Q}_m) - E(\mathcal{Q}))(E(\mathcal{Q}_m) - E(\mathcal{Q}))^T \quad (3)$$

여기서  $M_i$ 는  $i$ 번째 클래스의 평균벡터이고  $M$ 은 전체 클래스의 평균벡터이다.  $c-1$  차원 특징 벡터는 주요 성분 분석으로부터 투영된 특징벡터를 사용하여 얻어진다. 그런 후에 미리 만들어진 그 특징벡터를 사용하여 얼굴을 인식할 수 있다 [6].

III. Elastic Graph Matching (EGM)을 이용한 얼굴 인식

그레이 레벨 정보와 물체의 모양 정보를 동시에 이용하는 접근 방법 중의 하나를 소위 dynamic link architecture (DLA)라고 한다 [4]. 이러한 패턴 인식 방법은 신경 회로망에서와 같이 자기 스스로의 발전을 기본으로 하며 훈련과정(training phase)과 적응과정(recall phase)으로 구성된다. 훈련 과정에서, 얼굴 영역은 여러 개의 노드(node)와 에지(edge)로 나타나게 되는데 이는 여러 개의 오리엔테이션(orientation)과 공간 주파수의 크기 인자를 갖는 2차원 가버 필터의 응답으로부터 얻어진다. 적응 과정에서는 미리 만들어진 사람의 모델은 새로운 사람의 얼굴 영상과 비교되며 그 과정을 통해 비용 함수(cost function)가 최소화되도록 변형되는데, 이러한 과정을 여러 얼굴 모델에 대해 수행함으로써 그 영상에 대한 비용 함수의 분포(distribution)를 비교하여 얼굴을 인식할 수 있다

3.1 가버 웨이블릿 (Gabor wavelet)

컴퓨터 비전 분야에서 2차원 가버 웨이블릿의 사용은 1980년대에 Daugman이라는 사람에 의하여 발전되었다. 최근에는 지문, 텍스처(texture) 및 얼굴 인식 등에도 많이 응용되고 있다. 가버 필터는 뇌의 시각 피질(visual cortex)을 모델로 설계되었으며 이는 복소수 값을 갖는 가우시안 포락선(Gaussian envelope) 모양의 2차원 평면파 함수로 나타낼 수 있다 [7][8].

3.2 Elastic graph matching (EGM)

EGM을 이용한 얼굴 인식은 미리 만들어진 참조 모델의 그래프를 새로운 영상의 그래프와 비교하는 방법이다. 비교하고자 하는 두 영상 그래프 사이의 관계는 정의된 유사도 함수들에 의하여 구해질 수 있다. 그래서 비용 함수  $d(G, R)$ 는 각각 노드와 에지의 유사도 함수를 독립적으로 생각하여 식 (4)와 같이 측정한다.

$$d(G, R) = \sum_{i=1}^{N_n} d_n(G_{n_i}, R_{n_i}) + \lambda \sum_{j=1}^{N_e} d_e(G_{e_j}, R_{e_j}) \quad (4)$$

식 (4)에서  $G_{n_i}$ 는 입력 영상 격자(grid)  $G$ 의  $i$ 번째 노드를 나타내며,  $R_{e_j}$ 는 참조 모델의 그래프 격자  $R$ 의  $j$ 번째 에지를 나타낸다. 그리고,  $N_n$ 과  $N_e$ 는 각각 노드와 에지의 전체 개수이며  $\lambda$ 는 그래프의 강도를 특징지어주는 가중치 인자이다. 매칭 알고리즘은 보통 적어도 다음의 두 단계로 구성된다. 먼저 그래프를 변형시키지 않고 비교될 영상의 가장 적절한 위치를 찾는데, 이를 전체 이동(Global Move)이라고 한다. 그런 다음 식 (4)를 최소화하기 위하여 그래프를 변형시키는 작업이 이어지는데 이를 지역 이동(Local Move)이라고 한다.

IV. 제안한 알고리즘

제안된 얼굴 인식 방법은 세가지 단계로 나누어진다. 그림 1에서 볼 수 있듯이 첫 번째 단계에서 얼굴의 그래프가 2차원 가버 응답으로부터 구성된다. 그 다음 두 번째 단계에서 레이블된 그래프 벡터에 대해 주요 성분 분석 및 Fisherface 방법을 적용한다. 마지막으로 세 번째 특징 벡터에 대한 매칭 알고리즘이 실행된다. 본 논문에서는 fixed graph matching (FGM)과 주요 성분 분석 및 Fisherface 방법과의 병합방법과 EGM과 주요 성분 분석 및 Fisherface 방법과의 병합방법을 제안하고자 한다. 여기에서 FGM은 지역 이동을 고려하지 않은 EGM을 말한다.

4.1. Fixed graph matching과 주요 성분 분석 및 Fisherface 방법을 병합한 알고리즘

FGM과 주요 성분 분석 방법 및 Fisherface 방법을 병합한 알고리즘은 매칭과정에서 지역 이동을 고려하지 않는 경우로써 다음과 같은 과정을 통해서 얼굴을 인식한다. 우선, 첫 번째 단계인 2차원 가버 웨이블릿 변환을 통해 그래프를 구성 및 조정을 한 후 레이블된 그래프 벡터를 얻는다. 각각의 레이블된 그래프 벡터는 영상에서  $x_i$  노드에 관한 것으로서 이는  $f(x_i)$ 로 나타낼 수 있다. 평균벡터가

$$E(f(x_i)) = (1/M) \sum_{m=1}^M f_m(x_i)$$

이면, 노드  $x_i$ 에서 정규화된 벡터  $\Psi(x_i) = f(x_i) - E(f(x_i))$ 이다. 특히, 공간 주파수 크기 및 오리엔테이션이 각각 3과 4일 경우 한 영상에 대한 레이블된 그래프 벡터  $f(x) = (f(x_1), \dots, f(x_{12 \times N}))^T$ 이며 이에 대한 평균벡터는  $E(f(x))$ 로 표기한다. 여기에서  $N$ 은 총 노드의 개수이다.  $M$ 이 전체 얼굴 영상의 개수라고 한다면, 이 때 공산 행렬  $C(x)$ 는 한 영상을 구성하는 모든 노드들에 대한 정규화된 벡터  $\Psi(x)$ 을 이용하여 만든다. 그 다음 주요 성분 분석 방법을 적용하는데,  $C(x)$ 로부터 구한  $p$ 개의 제일 큰 고유값의 순서에 따라 그와 일치되는 고유벡터  $U_1(x), \dots, U_p(x)$ 를 계산한다. 그러면 주요 성분 분석에 의해 투영된 특징벡터는 식 (5)와 같다.

$$\Omega(x) = [U_1^T(x), \dots, U_p^T(x)] \Psi(x) = P(x) \Psi(x) \quad (5)$$

여기에서  $T$ 는 전치 행렬을 구하는 연산자이며  $\Omega(x)$ 의 차원은  $p \times 1$ ,  $p \leq 12 \times N$ 이다. 그 다음 주요 성분 분석 방법으로부터 구해진 특징벡터에 Fisherface 알고리즘을 적용한다. 그래서 수식 (2)와 (3)을 이용하여 다음과 같은 최종 특징벡터를 추출한다.

$$Z = [W_1^T P(x)(J(x) - E(J(x))), \dots, W_M^T P(x)(J(x) - E(J(x)))]^T \quad (6)$$

여기에서  $W_j$ 는  $(S_w)^{-1} S_b$  행렬로부터 구해진 기저벡터이다. 얼굴을 인식하기 위하여 어떤 얼굴영상이 입력되면 그 영상에 대해 일단 가버 웨이블릿 응답을 구하고, 거기서 나온 특징벡터인 레이블된 그래프 벡터를  $F_{in}$ 이라 하면, 이에 대해  $\Psi_{in} = F_{in} - E(F)$ 를 계산하고,  $(U_i)^T \Psi_{in}$ 을 모든  $i$ 에 대해 계산하여 특징벡터  $\Omega_{in}$ 을 얻는다. 그리고  $W_j^T \Omega_{in}$ 을 모든  $j=1, 2, \dots, M-1$ 에 대해 계산하여 제 2의 특징벡터  $Z_{in} = [W_1^T \Omega_{in} \ W_2^T \Omega_{in} \ \dots \ W_{M-1}^T \Omega_{in}]^T$ 을 얻고, 이것과 학습에 사용된  $\Psi(x)$ 의  $Z$ 간의 유클리드 거리가 최소가 되는 것을 찾음으로써 인식을 수행한다.

#### 4.2. Elastic graph matching과 주요 성분 분석 및 Fisherface 방법을 병합한 알고리즘

EGM과 PCA 및 Fisherface 방법을 병합한 알고리즘은 매칭 과정에서 지역 이동을 고려하는 경우이다. 첫 번째 단계는 위에서 제안한 방법과 같다. 여기에서 공간 주파수 크기 및 오리엔테이션이 각각 3과 4일 경우  $J(x) = [J_1(x), \dots, J_{12}(x)]^T$ 이고, 주요 성분 분석 방법에 의하여  $\Omega(x)$ 를 계산한다. 여기에서  $\Omega(x)$ 의 차원은  $p \times 1$ ,  $p \leq 12$ 이다. 그 다음 Fisherface 알고리즘을 워치업 적용하여 식 (6)과 같은 특징 벡터를 노드 수만큼 계산한다. 그 다음 식 (4)에서 비용 함수를 구성하는 노드에 관한 유사도 함수  $d_n(G_n, R_n)$ 를 식 (7)과 같이 정의한다.

$$d_n(G_n, R_n) = \|Z(x)^G - Z(x)^R\| \quad (7)$$

변형은 식 (4)의 비용 함수  $d(G, R)$ 가 최소 값을 얻을 수 있을 때까지 계속된다. 결국 이러한 지역 이동 과정을 수행한 후 계산된 비용함수를 비교함으로써 얼굴을 인식한다.

### V. 실험 및 결과

제안한 알고리즘 및 기존의 방법을 비교하기 위하여 얼굴 인식 실험을 수행하였다. 알고리즘은 VC++ 6.0을 사용하여 구현하였으며, 펜티엄 프로세서 III 450MHz, 메모리 128MB의 PC에서 실험하였다. 실험에 사용된 영상은 인터넷을 통해 얻은 Yale 대학 얼굴 영상 데이터베이스를 사용하였다. 이 데이터는 총 15명의 사람이며, 각 사람 당 광원의 위치에 따른 조도의 변화, 얼굴 표정 변화, 안경 착용 여부에 따라 10장의 얼굴 영상으로 구성되어 있다. 그림 2는 Yale대학 데이터베이스의 예이며, 본 논문에서는 원래  $320 \times 243$  크기의 영상을  $128 \times 128$  크기로 바꾸어서 전체 7가지의 얼굴 인식 방법에 대해 실험하였다

#### 5.1. Leaving-one-out 방법에 의한 실험 결과

표 1은 leaving-one-out 방법으로 실험한 7가지의 얼굴 인식 알고리즘에 대한 인식률의 결과를 보여주고 있다. 입력 영상을 제외한 149개의 영상을 가지고 학습을 수행한 후 인식한 결과이다. PCA 방법의 경우 고유 벡터의 개수는 전체 고유치 값의 95%를 점유하는 개수를 사용하였으며, Fisherface 방법의 경우 특징 벡터로써의 고유 벡터 개수는 14개를 사용하였다. 또한 EGM과 FGM을 실험하는데 있어서 전체적인 이동의 경우 일정한 위치를 지정해줌으로써 그 과정을 생략하

였다. 즉, 일률적으로 150개의 얼굴 영상에 대해 좌표 값 (30, 24)의 위치를 첫 번째 노드 위치로 지정하였다. 또한 가로로 7개, 세로로 10개의 노드로 얼굴 영상 그래프의 격자를 구성했으며 노드간의 거리는 11픽셀로 지정하였다. 또한, 그래프의 강도를 특징지어주는  $\lambda$ 값은 0.003이며 2차원 가버 필터의 오리엔테이션과 공간 주파수 크기인자는 각각 4와 3인 것을 사용하였다.

표 1에서 볼 수 있듯이, 인식률에 있어서는 Fisherface 방법, EGM 알고리즘, 그리고 FGM과 Fisherface를 병합한 알고리즘이 제일 좋은 성능을 보였다. 특히, FGM과 Fisherface 방법은 99.3%의 높은 인식률과 6초의 인식 속도를 보였다. 그리고, 얼굴 영상의 위치 이동이 있다면 K-L 변환을 사용하는 방법의 경우 인식률이 급속히 감소하는 반면, EGM과 FGM의 경우 전체 이동을 통해서 이를 보완할 수 있다.

#### 5.2. Hold-out 방법에 의한 실험 결과

Hold-out 방법을 이용하여 실험한 결과는 표 2에서 볼 수 있다. 이는 각 사람마다 3장의 얼굴 영상으로 학습을 한 후 나머지 영상들을 이용하여 얼굴을 인식한 결과이다. 1, 2, 3 번째는 그림 2의 Yale 데이터베이스의 왼쪽 위부터 오른쪽 방향으로 임의로 순서를 정한 것이다.

표 2에서 볼 수 있듯이 인식률의 경우 EGM 방법과 본 논문에서 제안한 방법인 FGM과 Fisherface와의 병합 알고리즘의 순서로 인식률이 92.6%, 90.9%로 가장 높았다. 그러므로 제안한 얼굴 인식 알고리즘은 기존의 K-L 변환을 이용한 방법보다는 인식률이 개선되었으며, 특징 벡터의 차원 감소로 인하여 EGM 방법보다는 인식 속도의 성능을 개선시켰다. EGM 방법의 경우, 인식률에 있어서는 다른 방법보다 우월하지만 훈련을 하는데 있어서나, 특히 인식을 하는 경우에 많은 시간이 소요된다. 왜냐하면 가버 필터를 사용하여 영상 그래프의 격자를 구성을 해야 하며 지역 이동을 하는데 있어서 많은 시간이 소요되기 때문이다. 특히, 본 논문에서 제안하는 알고리즘은 2차원 가버 필터의 공간 주파수와 오리엔테이션의 해상도를 높일 경우, EGM 방법보다 상대적으로 인식 속도의 향상은 더욱 높아질 수 있다.

### VI. 결론

본 연구에서는 기존의 얼굴인식 방법인 EGM과 Fisherface 알고리즘의 병합 방법을 모색했으며, 그 결과 얼굴 인식 시스템의 만족할만한 인식률 및 인식 속도의 향상을 얻을 수 있었다. 특히, EGM 과정에서 Fisherface 알고리즘의 적용으로 인한 레이블된 그래프 특징 벡터의 차원 감소로 인하여 연산 처리 시간을 단축할 수 있다는 장점을 실험을 통해 알 수 있었다. 또한 본 논문에서 제안한 알고리즘의 매칭 과정 중 얼굴의 모양을 고려한 지역적 이동에 대한 효율성을 좀 더 높인다면 보다 향상된 인식률을 얻을 수 있으리라 사료된다. 향후 연구해야 할 과제로서 먼저 얼굴 위치 변화가 EGM 방법에 미치는 영향과 얼굴 영상의 효율적인 학습 모델 구성에 관한 연구가 더욱 필요하다.

#### 참고문헌(또는 Reference)

- [1] R. Chellapa and S. Sirohey, Human and Machine Recognition of Faces: A Survey, *Proceedings of the IEEE*. Vol. 83, no. 5, May 1995.
- [2] A. P. Pentland and M. A. Turk, Face Recognition Using Eigenfaces, in *Proc. the International Conference on Pattern Recognition*, pp. 586-591, 1994.

[3] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. Fisherfaces: recognition using class specific linear projection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997

[4] B. Duc, S. Fisher, and J. Bigun, "Face Authentication with Gabor Information on Deformable Graphs," *IEEE Trans. on Image Processing*, vol. 8, no. 4, April 1999.

[5] C. Kotropoulos, A. Tefas, and I. Pitas, "Face authentication using variants of elastic matching based on mathematical morphology that incorporate local discriminant coefficients," in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3729-2732, 1998.

[6] 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, August 1996.

[7] B. S. Manjunath, and W. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data" *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 18, no. 8, pp. 837-842, August 1996.

[8] M. J. Lyons, J. Budynek, and S. Akamatsu, "Automatic Classification of Single Facial Images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 12, pp. 1357-1362, December 1996.

본 연구는 정보통신부의 2000년도 대학기초연구비 지원에 의하여 진행되고 있습니다.

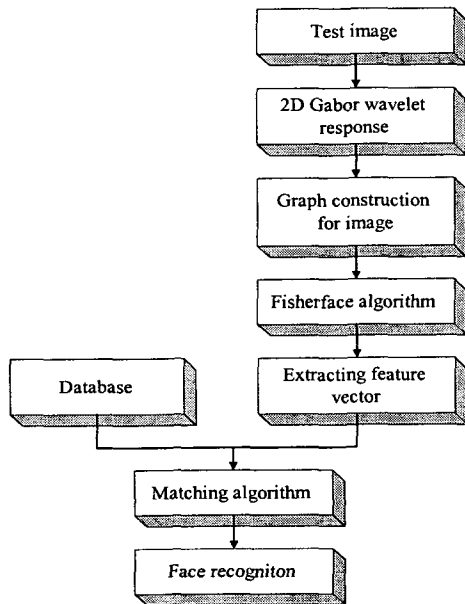


그림 1 제안하는 얼굴 인식 알고리즘

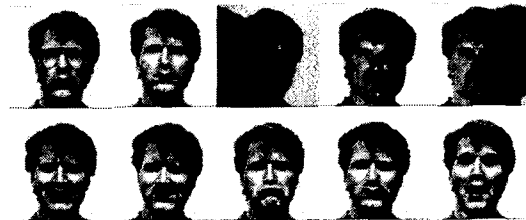


그림 2 Yale 대학 데이터베이스의 예

표 1 Leaving-one-out 방법에 의한 실험 결과

Recognition method	Recognition rate	Recognition speed
PCA	79.3%	Within 0.5 sec.
Fisherface	99.3%	Within 0.5 sec.
EGM	99.3%	Approx. 2624 sec.
FGM + PCA	88.7%	Within 6 sec.
FGM + Fisherface	99.3%	Within 6 sec.
EGM + PCA	88.0%	Within 892 sec.
EGM + Fisherface	90.0%	Within 1118 sec.

표 2 Hold-out 방법에 의한 실험 결과

Images used in training	PCA	Fisherface	EGM	FGM + PCA	FGM + Fisherface	EGM + PCA	EGM + Fisherface
1, 2, 3 <sup>rd</sup> images	81.9%	93.3%	91.4%	89.5%	94.3%	87.6%	93.9%
3, 4, 5 <sup>th</sup> images	75.2%	85.7%	94.3%	85.7%	94.3%	88.6%	88.6%
5, 6, 7 <sup>th</sup> images	82.9%	87.6%	93.3%	86.7%	93.3%	87.6%	90.5%
7, 8, 9 <sup>th</sup> images	72.4%	77.1%	93.3%	81.0%	87.6%	85.7%	83.8%
1, 9, 10 <sup>th</sup> images	70.5%	72.4%	90.5%	75.2%	84.8%	78.1%	80.0%
Recognition rate (Ave.)	76.6%	83.2%	92.6%	83.6%	90.9%	85.5%	87.2%
Recognition speed	Within 0.25 sec.	Within 0.25 sec.	Approx. 359 sec.	Within 5 sec.	Within 5 sec.	Within 235 sec.	Within 270 sec.