

論文2003-40CI-1-3

가상 데이터와 융합 분류기에 기반한 얼굴인식 (Face Recognition based on Hybrid Classifiers with Virtual Samples)

柳淵植*, 吳世泳**

(Yeon-Sik Ryu and Se-Young Oh)

요약

본 논문은 인위적으로 생성된 가상 학습 데이터와 융합 분류기를 이용한 얼굴인식 알고리즘을 제안한다. 특징공간에서의 최근접 특징 선택 방법과 연결주의 모델에 기반한 서로 다른 형태의 분류기를 융합하여 통합효과를 얻도록 하였다. 두 분류기는 모두 학습 데이터의 공간적인 분포에 따라 생성된 가상 학습데이터를 이용하여 학습되고 이용된다. 첫째로, 특징 공간에서의 각 정보(Angular Information)를 이용하는 최근접 특징각(the Nearest Feature Angle : NFA)을 이용하여 저장된 학습데이터와 가장 근접한 것을 찾고, 둘째로, 질의(Query) 얼굴 특징 정보를 정면얼굴 영상의 특징정보로 투영하여 얻은 정보에 기반한 분류기의 결과를 이용한다. 정면영상 특징정보로의 투영은 다층 신경망을 이용하여 정면 회상망(Frontal Recall Network)을 구현하였고, 이것을 여러 개 묶어 앙상블 네트워크로 구성한 Ensemble 회상망(Ensemble Recall Network)을 사용하여 일반화 성능을 향상시켰다. 끝으로, 각 분류기의 결과에 따라 융합 분류기가 최종 결과를 선택하도록 하였다. 제안된 알고리즘을 6 종류의 서로 다른 학습/시험데이터 군에 적용하여 평균 96.33%의 인식률을 얻었다. 이것은 특징라인에 기반한 방법(the Nearest Feature Line) 평균 에러율의 61.2%이며, 단일 분류기를 사용한 경우 보다 안정된 결과를 얻고 있다.

Abstract

This paper presents a novel hybrid classifier for face recognition with artificially generated virtual training samples. We utilize both the nearest neighbor approach in feature angle space and a connectionist model to obtain a synergy effect by combining the results of two heterogeneous classifiers. First, a classifier called the nearest feature angle (NFA), based on angular information, finds the most similar feature to the query from a given training set. Second, a classifier has been developed based on the recall of stored frontal projection of the query feature. It uses a frontal recall network (FRN) that finds the most similar frontal one among the stored frontal feature set. For FRN, we used an ensemble neural network consisting of multiple multiplayer perceptrons (MLPs), each of which is trained independently to enhance generalization capability. Further, both classifiers used the virtual training set generated adaptively, according to the spatial distribution of each person's training samples. Finally, the results of the two classifiers are combined to comprise the best matching class, and a corresponding similarity measure is used to make the final decision. The proposed classifier achieved an average classification rate of 96.33% against a large group of different test sets of images, and its average error rate is 61.2% that of the nearest feature line (NFL) method, and achieves a more robust classification performance.

Keywords : Feature Angle, Frontal Recall Network, Hybrid Classifier

* 正會員, LG N-Sys開發
(LG N-Sys)

** 正會員, 浦項工科大學校
(Pohang University of Science and Technology)
接受日字:2001年7月10日, 수정완료일:2002年12月30日

1. 서론

얼굴인식에 관한 연구는 다양한 응용범위와 기술적인 극복을 위해 지난 몇 년간 꾸준히 증가하고 있다. 연구 방법으로는 크게 3가지로 분류할 수 있는데, 1) 얼굴의 특징점을 추출하여 특징 점들간의 기하학적인 관계를 이용하는 것^[3, 4, 26, 27], 2) 템플릿 매칭^[1, 5, 7, 21, 31], 3) 신경망을 이용한 방법^[1, 3, 7, 19, 24, 25, 30] 등이다.

특징 점들간의 기하학적인 정보를 이용하는 방법은 간단하기는 하지만, 특징 점들을 정확하게 추출하기 위한 기술을 필요로 한다^[3, 4, 26, 27]. Gabor 필터를 이용한 특징 점들간의 그래프 매칭을 통하여 인식을 하는 방법은 연산시간이 많다^[7, 21]. 또 일반적으로 템플릿 매칭을 이용하는 방법은 영상에 직접 적용되는 방법으로, 특징 점을 추출해야 하는 번거로움이 없어서 일반적으로 많이 사용하는 방법이라 하겠다. 이러한 템플릿 매칭을 이용한 방법들 중에서 가장 많이 사용되는 것이 고유얼굴 (Eigenface)을 이용한 방법이라 하겠다^[1, 5-7, 31]. 이 방법은, PCA (Principal Component Analysis) 를 이용한 것으로 얼굴영상을 학습영상으로부터 추출한 주요 고유얼굴의 선형적합으로 표현하는 것으로, 고유얼굴 공간에 투영하여 얻은 계수 벡터들을 이용하여 분류하는 기술이다. 고유얼굴은 얼굴영상집단의 변화를 나타내는 축으로, 부 공간상(Sub Space) 에서 얼굴을 잘 표현하는 특징 공간인 반면에 분류에는 약한 면이 있으며, 이를 보완하는 접근 방법이 시도되고 있다^[22, 23].

신경망에 기반하여 얼굴을 인식하는 방법도 많이 제안되고 있다^[1, 3, 7, 19, 24, 25, 30]. 비선형 분류기로서 많이 사용되는 신경망은 많은 응용분야에 이용되고 있는데, 학습데이터군의 크기, 신경망의 구조, 다루고자 하는 문제의 복잡도 등은 일반화 문제를 어렵게 한다. 이러한 문제를 해결하기 위하여 여러 개의 신경망을 묶어서 사용하는 구조^[14-16, 18, 24, 25]와 최적의 신경망 weight를 구하는 시도^[28, 29] 등이 이루어 지고 있다. 또한 실제 환경에서 학습데이터를 많이 얻는 것이 힘든 경우가 많다. 학습데이터의 크기를 크게 만드는 작업의 일환으로, 노이즈^[10, 11], 힛트^[12], 가상 데이터^[8, 9, 13] 등의 방법을 통하여 학습데이터의 크기를 충분히 크게 만드는 작업 등이 시도되고 있다. 그러나, 신경망이 비선형 분류기로 잘 동작하기는 하나, 학습데이터의 영역에서 벗어난 시험데이터에 대해서는 여전히 제약이 있다고 하겠다.

본 논문에서는 특징공간에서의 각 정보와 가상 학습 데이터에 기반한 융합 분류기를 제안한다. 고유얼굴이 구성하는 특징공간에서 작업이 이루어지며, 전체적인 일반화 성능을 향상시키기 위하여 3 단계를 걸쳐서 분류결과를 얻는다. 첫째로, 학습 대상별(이하 클래스)로 주어진 학습데이터의 고유얼굴 공간상의 분포에 따라 생성된 가상데이터로 학습데이터의 양을 증가시킨 후, 안정된 결과를 얻기 위한 여러 개의 독립된 분류기 채널을 구성한다. 첫번째 채널은 근접 특징 각에 기반한 분류기(Classifier based on the Nearest Feature Angle : 이하 NFA) 로서, 질의 특징 벡터(query feature vector) 와 가장 유사한 것을 가상 데이터로 확장된 학습 데이터 공간에서 찾으며, 다음은, 질의 특징벡터를 정면 특징 벡터로 변환하는 정면얼굴 회상망 (Frontal Recall Network : 이하 FRN) 을 이용하는 것이다^[30]. 최종적으로, 융합 분류기(Hybrid Classifier)는 두 분류기로부터 얻은 정보에 기반 하여 최종 결정을 내린다. 이렇게 함으로써 한가지 방법만 사용하는 경우보다 각 분류기가 서로의 단점을 보완하여 보다 나은 일반화 성능을 얻을 수 있을 것으로 기대된다.

이 논문의 구성은 다음과 같다. 제 2장에서는 제안 시스템의 구조를 설명하고, 제 3장에서는 가상데이터를 이용하여 학습데이터를 확장하는 방법과 NFA 에 대하여 설명한다. 제 4장과 5장은 FRN 과 융합 분류기에 대하여 설명한다. 실험결과는 제 6장과 7장에서 설명한다.

II. 제안 시스템의 구조

얼굴 영상은 표정과 조명등의 조건에 따라 동일인 내에서 발생하는 변화의 크기가 타인과의 사이에서 측정되는 변화보다 크게 나타나는 경우가 있어서 주어진 학습데이터군이 변화 가능한 모든 경우를 대표하기 어렵다^[2]. 그 결과 설계된 분류기는 시험 데이터에 대하여 오류를 범하기 쉽다고 하겠다. 시험 데이터에 대하여 분류기가 보다 안정되게 동작하기 위해서는 입력으로부터 가능한 많은 정보를 추출하기 위한 다중경로의 처리 단계가 작업이 필요하다 하겠다. 본 논문은 <그림 1>과 같이 서로 다른 특성을 갖는 두 개의 부분류기(sub-classifier)가 병렬로 동작하는 융합 분류기를 제안한다.

효과적인 분류기를 만들기 위해서는, 학습데이터와

시험데이터의 성격을 규명하는 것이 필수적이다. 우리는 서로 다른 영역에서 잘 동작하며, 그들의 결과를 융합할 수 있는 분류기들을 제안한다. 더불어, 얼굴 학습 데이터의 부족에서 오는 문제점을 해결하기 위해서 각 클래스의 주어진 학습데이터의 특징공간상의 분포에 따라 가상 학습데이터를 생성(Augmented Training Feature Set) 하여 사용하였다. 또한, 단일 FRN을 사용하는 것보다, 사람이 다른 사람 얼굴의 여러 특징으로부터 얻은 정보를 통합하여 사용하는 것을 모방하여 다수의 FRN으로 하나의 Ensemble 회상망(Ensemble Recall Network: 이하 ERN)을 구성하여 사용하였다. 이는 다수의 신경망을 묶어서 사용함으로써 신경망 자체의 일반화 성능이 향상됨을 이용한 것이기도 하다.

<그림 1>에서, NFA는 질의 대상의 특징 벡터와 가상 학습데이터로 확장된 학습데이터 사이에서 가장 특징 각(이하 유사도)이 큰 것을 찾아서 분류하도록 하며, ERN은 질의 특징 벡터를 그것의 정면 얼굴 특징 벡터로 투사하도록 하여 학습데이터의 정면 얼굴 벡터와 유사도를 평가하여 분류한다. 두 분류기들의 역할은 다음과 같이 기대된다. <그림 2>에서 개념적으로 도시한 바와 같이, 1) NFA 나 ERN 모두 '그룹 I' 의 영역에서 잘 동작하며, 2) ERN 은 '그룹 II' 와 같이 시험 데이터가 비선형적으로 구분 가능한 영역에 존재할 때 잘 동작하며, 3) NFA 가 '그룹 III'과 같은 영역에서 잘 동작할 것을 기대한다. 융합 분류기는 두 부분류기의 결과 정보에 따라 최종 분류 결정을 내린다.

1. PCA(Principal Component Analysis)에 의한 특징 추출

본 논문에서 사용하는 특징 공간은 PCA에 의해 얻어지는 고유얼굴 공간(eigenface space)으로, PCA는 데이터 압축과 특징 추출에 널리 사용되는 통계적인 기술이다. 여기서 주어진 학습 영상들, I_1, I_2, \dots, I_N , 에 대하여 y_i 가 i 번째 학습영상의 열 벡터(Column vector) 표현이라고 하자. 매트릭스 $Y = [y_1 - m, y_2 - m, \dots, y_N - m]$ 는 $M \times N$ 매트릭스로서, M 은 얼굴영상의 화소 수이고, N 은 영상집합에 있는 영상의 수가 된다. 여기서 m 은 y_i 들의 평균벡터이다. YY^T 의 고유벡터(고유얼굴: e_1, e_2, \dots, e_M) 는 직교 베이스(orthogonal basis) 특징 공간을 형성하며, y_i 들과 임의의 얼굴 영상들은 이 고유얼굴들을 이용하여 표현이 가능하며, 복구도가

능하다. 즉, $y_i = \sum_{j=1}^M x_j e_j + m$ 으로 표현된다. 여기서 $x = (x_1, x_2, \dots, x_M)^T = W^T (y_i - m)$ 이며, W 는 고유얼굴들로 구성된 매트릭스 $[e_1, e_2, \dots, e_M]$ 이다. 일반적으로 Least Square Error 관점에서 보았을 때, 고유 값(Eigenvalue)이 큰 k 개의 고유얼굴로도 원래의 영상의 근사가 가능하며, 분류기에 사용되는 각 얼굴영상에 해당하는 특징 벡터는 $x = \{x_1, x_2, \dots, x_k\}$ 이다(PCA 방법에 대한 자세한 내용은 [7, 17, 31]을 참조하기 바람).

III. 가상 학습 데이터와 NFA 분류기

x_i 와 x_j 가 고유 얼굴 공간상의 특징 점이라고 할 때, Li 등^[1]은 동일 클래스 내의 두 특징 점, x_i 와 x_j 을 잇는 선형 특징선 $\overline{x_i x_j}$ 으로 특징 점간의 변화를 내외 보간법(interpolation, extrapolation)을 이용하여 모델링 하고, 인식할 수 있음을 제안(the Nearest Feature Line Method: 이하 NFL) 하였다. 그러나 실제로는 얼굴영상에 있어서의 변화를 모델링 하는 특징선의 변화는 매우 복잡한 양상을 띠며, 선형 특징 선으로 정확하게 모델링 하기 어렵다^[1, 20]. 더구나, <그림 3>에 나타난 바와 같이 질의 특징 점 x 와 $\overline{x_i x_j}$ 상에 직각 투사 점 x_p 간의 거리는 $\overline{x_i x_j}$ 와 $\overline{x x_p}$ 가 이루는 각(Angle)이 일정하다 하여도 x 의 위치가 어디냐에 따라 변화가 생긴다. 이러한 문제를 해결하기 위하여 각 정보

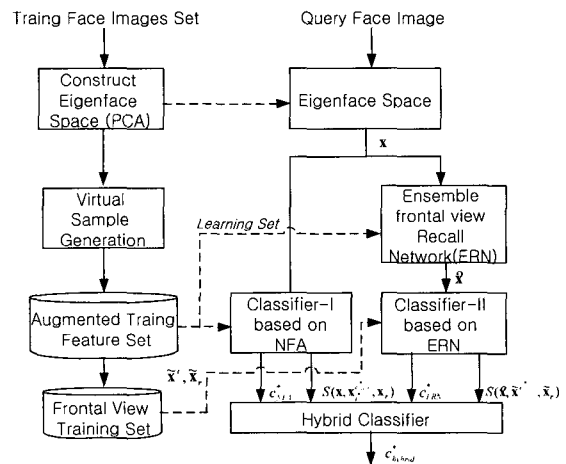


그림 1. 제안된 융합 분류기의 처리 구조
Fig. 1. The Proposed System Architecture.

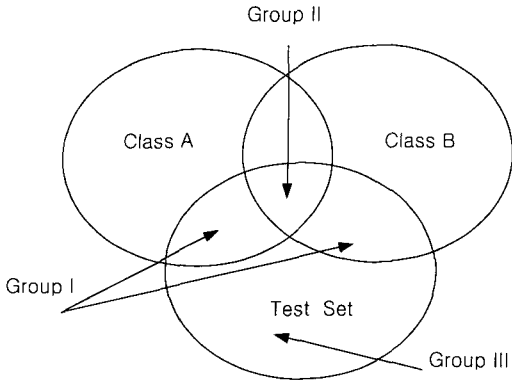


그림 2. 시험 데이터 분포의 개념적인 도식
Fig. 2. The conceptual description of test data distribution.

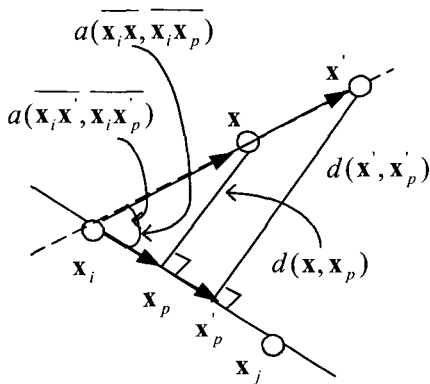


그림 3. 두 유클리디안 거리, $d(x, x_p)$, $d(x', x'_p)$ 는 각도 $a(x_i, x', x_i, x'_p)$, $a(x_i, x, x_i, x_p)$ 는 동일하다 하여도 서로 다르게 나타난다. 얼굴인식문제에 있어서 얼굴영상은 자세에 따라 영상의 변화하므로 변화를 나타내는 각 정보가 중요하다

Fig. 3. Two Euclidean distance $d(x, x_p)$ and $d(x', x'_p)$ are different, whereas the angle $a(x_i, x', x_i, x'_p)$ and $a(x_i, x, x_i, x_p)$ are the same. Angle information is also important for face recognition problems because a face image varies according to the pose.

(Angular information)에 기초한 유사도(Similarity Measure)를 정의하고, 그에 기반한 가상 학습데이터 생성방법을 제안한다.

1. 특징 각에 기반한 유사도

x_i 와 x_j 가 클래스 c 에 속한 두 특징 점으로, x_p 을 고유공간상의 기준점이라고 하자. x_i 와 x_j 사이의 특징 각 (θ)에 기반한 유사도는 다음과 같이 정의 된다.

$$S(x_i, x_j, x_p) = \cos \theta = \frac{\overline{x_i} \cdot \overline{x_j}}{\|\overline{x_i}\| \|\overline{x_j}\|} \quad (1)$$

여기서 $\overline{x_i} = x_i - x_p, \overline{x_j} = x_j - x_p, x_p = \frac{1}{N_c} \sum_{k=1}^{N_c} x_k$ 이고 N_c 는 클래스 c 의 크기이고 \cdot 는 벡터내적이며, $\|\cdot\|$ 는 L_2 -norm이다. 유사도 값 $S(\cdot)$ 는 최대 1, 최소 -1의 값을 갖는다. 유사도의 값이 클수록 두 특징 점은 서로 더욱 비슷한 것으로 평가된다.

2. 가상 학습 데이터의 생성

실제로 학습데이터는 얼굴의 모든 변화를 수용하기 어려우므로, 학습되지 않은 시험 영상을 정확하게 분류하는 것은 쉽지 않은 작업이다. 이러한 문제의 극복을 위하여, 주어진 학습 데이터에 기반하여 특징 공간상에서 가상 데이터를 생성하여 학습데이터의 수를 확장할 필요가 있다. 이렇게 확장된 학습데이터를 사용함에 따라 학습되지 않은 시험 데이터에 대하여 보다 정확한 분류를 할 수 있을 것으로 기대된다. 첫번째로, 영상 공간에서, 학습데이터를 확장하는 가장 손쉬운 방법은 미리 영상을 이용하는 것이다. 즉 얼굴 영상은 좌우 대칭이므로, 좌우를 바꾼 영상은 다른 각도에서 바라본 얼굴 영상이 된다. 둘째로, 동일 클래스에 속한 특징 점들의 분포를 평가하여 <그림 4>와 같이 두 종류의 영역, 'A' 와 'B' 로 정의 할 수 있다. 영역 'A'와 같이 두 특징 공간의 밀도가 낮은 영역, $S(x_1, x_2, x_r) \leq \alpha$ 에 대해서는, 가상 데이터 v_1 을 식 (2)와 같이 얻는다.

$$v_1 = x_1 + \mu(x_2 - x_1) \quad (2)$$

반대로 영역 'B'와 같이 특징 점간의 밀도가 높은 영역, $S(x_2, x_3, x_r) > \alpha$ 에 대해서는 가상 데이터 v_2 를 식 (3), (4)의 단계를 거쳐서 얻는다.

$$v_2 = x_4 + \mu(x_M - x_4) \quad (3)$$

$$x_M = x_2 + \mu(x_3 - x_2) \quad (4)$$

여기서 μ 는 0.5 로 사용하였으며, x_4 는 평가에 참여하지 않은 특징 점들이 되며, v_2 는 x_2, x_3, x_4 가 형성하는 공간 표면에 존재하게 된다. 클래스 c 를 위해 생성되어진 K_c 개의 가상 데이터는 원래의 학습데이터에 추가되어 클래스 c 의 학습데이터는 $X^c = \{x_1, x_2, \dots, x_{N_c}, v_1, v_2, \dots\}$

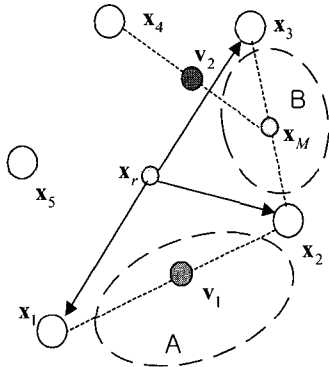


그림 4. 한 클래스의 특징 공간상의 학습 데이터 분포와 가상 학습데이터의 생성의 개념 예 x_r 은 이 클래스의 특징점들의 중심점이다

Fig. 4. Example of training set distribution of a class in feature space and virtual sample generation. The x_r is set to the average feature point of this class.

v_1, v_2, \dots, v_k)가 된다. 본 논문의 제안 방법과 NFL의 큰 구별은 1) 주어진 특징 점들의 공간적인 분포를 각 정보를 이용하여 평가하여 가상 데이터를 생성하며, 2) 가상 데이터가 세 개의 특징 점이 이루는 표면에 존재하도록 생성되어진다는 점이다.

3. NFA에 기반한 분류기

x_i^c 가 클래스 c 의 i 번째 특징 점이라고 하자. 질의 특징 점 x 와 x_i^c 사이의 유사도는 모든 클래스 C 개와 비교할 때 $N_c = \sum_{i=1}^C (N_c + K_c)$ 의 계산식을 필요로 한다. 이 유사도 값은 내림차순으로 정리되며, 각각은 클래스 구분 값(class identifier)과 특징 점 인덱스를 갖게 된다. 그 중에서 가장 큰 유사도 값을 나타내는 것을 식 (5)와 같이 찾는다.

$$S(x, x_i^{c_{NFA}}, x_r) = \max_{i'} \max_{c'} S(x, x_i^{c'}, x_r) \quad (5)$$

여기서 x_r 은 분류를 위해서는 제로 벡터로 설정하여 사용한다. 위 식으로부터 NFA의 결과로 c_{NFA}^* , i' , $S(x, x_{i'}^{c_{NFA}^*}, x_r)$ 의 값을 얻는다.

IV. 정면 회상 망의 Ensemble

사람이 영상정보로부터 사람을 구분하는 과정을 보면, 얼굴영상으로부터 얻은 많은 힌트들이 뇌의 복잡한

투영을 통하여 처리되어 이미 머리 속에 저장되어 있는 얼굴 영상을 회상하는 것임을 알 수 있다. 이러한 과정과 유사하게, 각각 서로 다른 구조를 갖는 다수의 다층신경망(MultiLayer Perceptron: 이하 MLP)을 이용한 ensemble 회상망(Ensemble Recall Network: ERN)을 제안하였다 <그림 5>.

각각의 정면 특징 회상망(FRN)은 한 개의 MLP로 구성되며, 입력 특징 점을 그에 해당하는 정면 특징 점으로 투영하는 역할을 수행한다. 즉, 인간이 누군가의 입력 얼굴영상으로부터 저장된 얼굴을 기억해 내듯이 FRN은 정면 영상에 대한 특징벡터를 기억해 낸다.

MLP는 크게 두 가지의 특징을 갖고 있다. 즉, 복잡한 입력공간에 대한 비선형 분류기로서의 역할과 학습되어지지 않은 데이터에 대한 일반화 기능이다. 조명과 자세에 따른 변화가 사람 사이의 변화보다 크게 나타나는 얼굴영상공간의 특징 공간은 매우 복잡한 비선형 공간이고^[2], 이러한 이유로 인해 MLP를 분류기로 사용하는 것은 매우 효과적이라 하겠다. 더불어 일반화 능력은 학습데이터의 부족에 따른 성능향상에 꼭 필요한 기능이라 하겠다. 또한 하나의 FRN이 아니라, 구조가 서로 다른 FRN을 여러 개 사용함으로써, 동일 입력에

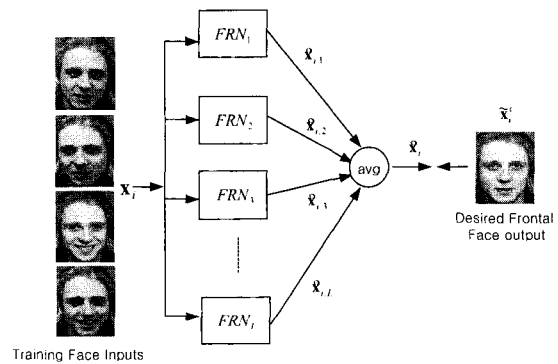


그림 5. Ensemble 회상망(Recall Network)의 구조. L 개의 서로 다른 구조로 학습된 FRN이 ERN을 구성하고 있는 경우. 좌측 영상은 FRN 학습에 사용되는 입력특징점에 해당하는 얼굴영상의 예이고, 우측 영상은 FRN의 원하는 정면얼굴특징점에 해당하는 얼굴영상의 예이다

Fig. 5. Architecture of the ensemble recall network. L FRNs trained independently constitute ensemble recall network. The 4 face example images of the left side and a face image of the right side correspond to the 4 input feature points and the desired output feature point, respectively.

대해서도 서로 다른 관점에서 정보를 처리하여 융합함으로써 보다 안정된 결과를 얻도록 하였다.

1. 정면 회상 망 (Frontal Recall Network) 의 학습
ERN 내의 각각의 FRN은 동일한 입력과 원하는 출력 데이터(desired output data)로 독립적으로 학습된다. 즉, l 번째 FRN은 클래스 X^c 의 특징 점 x_i^c 가 식 (6)과 같이 정면 특징 점 \tilde{x}^c 를 회상하도록 학습한다.

$$\tilde{x}^c \equiv \tilde{x}_{i,l} = f_l(x_i^c) \quad (6)$$

여기서 $l=1,2,\dots,L$, L 은 FRN의 개수, $i=1,2,\dots,N_c+K$.이며, $f_l(\cdot)$ 은 l 번째 FRN에 의해 구현되는 회상 함수이다. 원하는 출력 데이터로 사용되는 정면영상은 학습 데이터에서 가장 정면에 유사한 것을 선택하여 사용되며, 그 일부 예가 <그림 6>에 나타나 있다. 각각의 FRN의 출력, $\tilde{x}_{i,l}$ 은 앙상블 회상 망의 결과를 얻기 위하여 식 (7)과 같이 선형적으로 융합되어진다.

$$\tilde{x}^c \equiv \hat{x}_i = \sum_{l=1}^L \alpha_l \tilde{x}_{i,l} \quad (7)$$

여기서 선형융합의 경우 $\alpha_l = 1/L$ 로 설정되어지며, 이렇게 다수의 서로 다른 구조의 FRN의 결과를 융합함으로써 서로의 부족함을 보완하는 형태가 되어, 전체적



그림 6. 정면회상망(Frontal Recall Network)의 학습에 사용되기 위해 선택된 정면 영상의 예
Fig. 6. Some frontal face images used for training recall network.

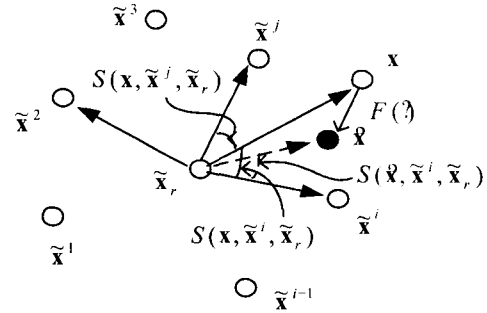


그림 7. 질의 특징 x 를 그것이 해당하는 정면특징 \hat{x} 로 변환(회상)하여 유사도 $s(\hat{x}, \tilde{x}^i, \tilde{x}^r)$ 를 얻기 위한 과정. \tilde{x}^i 는 i 클래스의 정면 특징점이고, \tilde{x}^r 는 선택된 정면특징점들의 중심점이다

Fig. 7. The process obtaining similarity measure, $s(\hat{x}, \tilde{x}^i, \tilde{x}^r)$, through transferring the query feature x to its corresponding frontal feature, \hat{x} , by ensemble recall network $f_l(\cdot)$. The \tilde{x}^i is the frontal feature of class and \tilde{x}^r is average feature point of frontal features.

으로 융합에 따른 일반화 성능 향상이 기대 된다.

2. Ensemble 회상망에 기반한 분류

\tilde{x}^c 가 클래스 c 의 정면 특징 점이라고 하자. 질의 특징 점 x 는 각각의 FRN에 의하여 정면특징 점으로 회상되어지고, ERN에 의하여 \hat{x} 가 얻어진다. 이와 각각의 클래스에 대하여 정면 특징 점으로 저장되어 있는 것과의 유사도를 평가한다 (<그림 7> 참조).

이러한 유사도의 비교에 따라서 그 결과는 내림치순으로 정리되어지면, 가장 큰 유사값을 나타내는 클래스와 유사도 값을 식 (8)로부터 얻는다.

$$S(\hat{x}, \tilde{x}^{c_{ERN}}, \tilde{x}^r) = \max_{1 \leq c \leq C} S(x, \tilde{x}^c, \tilde{x}^r) \quad (8)$$

여기서 $\tilde{x}^r = 1/C \sum_{c=1}^C \tilde{x}^c$ 로 정면 얼굴 특징 점들의 중심 점이 된다. c_{ERN} 는 가장 큰 유사도를 나타내는 클래스를 나타낸다.

V. NFA 와 ERN 의 융합 분류기

c_{ERN} 와 c_{NFA} 를 각각 ERN과 NFA으로 찾은 클래

스라 하고, $S(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}^{c_{ERN}}, \tilde{\mathbf{x}}_r)$, $S(\mathbf{x}, \mathbf{x}_p^{c_{NFI}}, \mathbf{x}_r)$ 를 그에 해당하는 유사도라고 하자. 두 분류기가 찾은 클래스와 유사도에 따라 융합 분류기는 식 (9)와 같이 판단한다.

$$c_{\text{fusion}} = \begin{cases} c_{ERN} \text{ or } c_{NFI} & \text{if } c_{NFI} = c_{ERN} \\ c_{ERN} & \text{if } S(\mathbf{x}, \mathbf{x}_p^{c_{NFI}}, \mathbf{x}_r) \geq S(\mathbf{x}, \mathbf{x}_p^{c_{ERN}}, \mathbf{x}_r) \text{ and } c_{NFI} \neq c_{ERN} \\ c_{NFI} & \text{if } S(\mathbf{x}, \mathbf{x}_p^{c_{NFI}}, \mathbf{x}_r) < S(\mathbf{x}, \mathbf{x}_p^{c_{ERN}}, \mathbf{x}_r) \text{ and } c_{NFI} \neq c_{ERN} \end{cases} \quad (9)$$

두 개의 독립된 분류기의 결과의 이용은 각 분류기가 선택한 클래스가 다른 경우 각 분류기가 찾은 유사도의 값은 클래스에 대한 확신도로 평가된다. 서로 다른 영역에 대하여 효과적인 분류기의 결과를 사용함으로써, 확신도가 높은 쪽을 선택하도록 하여, 한 개의 분류기를 사용하는 경우보다 안정된 분류기를 구성할 수 있다.

VI. 실험 결과

실험은 본 논문의 제안 방법과 Li의 NFL에 의한 방법을 평균 분류 에러 측면에서 비교하였다. 실험 데이터는 캠브리지 대학의 ORL(Olivetti Research Lab.) 얼굴 데이터 베이스를 사용하였다. 데이터 베이스는 한 사람 당 10장씩, 총 40명의 400장 영상으로 구성되어 있다. 또한, 좌우 약 20도 내외에서 자세 변화가 있으며, 표정변화와 성별, 나이 등의 분포가 다양하다 (<그림 8>에 한 사람에 대한 예를 나타내었다).

각 사람의 10장의 얼굴영상 중에서 정면영상에 가장 가까운 것을 선택하여 FRN의 원하는 출력특징 벡터의 학습영상으로 사용하였다. 나머지 영상을 Li의 방법으로 랜덤하게 분류하여, 중복 없이 학습 영상 200장, 시험 영상 200장으로 나누었다. 학습영상의 확장을 위해 전술한 바와 같이, 학습영상과 미리 영상을 함께 이용



그림 8. 한 사람의 10장 영상
Fig. 8. The set of ten images for one subject.

하여 고유얼굴을 얻었으며, 그에 기반하여 각 클래스별로 가상 학습데이터를 얻어서 학습영상을 확장하였다. 고유 얼굴은 전체 400개가 얻어지며, 이중에서 고유 값이 가장 큰 50개에 해당하는 고유얼굴을 이용하여 특징 점들을 얻었다. 각 클래스의 확장된 학습데이터의 크기는 원래 영상이 5개, 미리 영상이 5개, 생성된 가상 학습데이터 30개로 제한하여 총 40개를 사용하였다.

FRN은 50개의 입력을 갖고, 즉, 50개의 고유얼굴에 투영된 정보, 5개의 FRN을 사용하여 ERN을 구성하였다. 각 FRN의 히든 층은 각각 '50-150-150', '50-80-80', '50-200-200', '50-150-200', '50-80-100'을 갖도록 임의적으로 정하였고, 학습은 Resilient Back-propagation 방법으로 학습률 0.2로 학습하였다.

2. 서로 다른 형태의 분류기의 성능비교

6번의 랜덤하게 나는 학습/시험 데이터 군에 대하여 5 종류의 분류기에 대해 시험하였다.

<그림 9>는 9번의 시도에 대한 각 분류기의 분류 오류율을 나타낸다. <그림 10>은 평균 성능을 나타낸다. NFA분류기, ERN 분류기, NFL 분류기, 유클리디안 방법들은 각각 4.83%, 5.5%, 5.83%, 6.08%의 평균 분류 오류율을 나타내었고, 제안한 방법인 NFA와 ERN의 융합 분류기는 3.67%의 평균 분류 오류율을 나타낸다. 이것은 NFA와 ERN을 단독으로 사용한 경우보다 낮은 분류 에러율로써, 이종의 분류기가 서로 다른 영역에서 효과적으로 동작하고 있음을 보인다. 더구나, 이 에러율은 표준 NFL 방식의 61.23%에 해당한다.

2. 서로 다른 융합 분류기의 성능 비교

제안하는 방법의 각 부분류기와 NFL의 변형된 형태 (MNFL)에 대하여 융합방법에 따른 성능을 비교하였다. <그림 11>은 융합방법에 따른 평균 오류율을 나타낸다. 그림에서, NFL-I는 Li가 제안한 NFL방법의 의한 결과이고, NFL-II는 미리 영상을 포함하여 NFL을 적용한 결과이다. NFL-I 또는 NFL-II와 ERN의 융합 방법은 다음과 같은 방법으로 하였다. <그림 12>에서와 같이 $\mathbf{x}_p^{c_{NFI}}$ 를 질의 특징 점 \mathbf{x} 가 NFL에 의해 선택된 특징 선, $\overline{\mathbf{x}, \mathbf{x}_i}$, 에 수직 투사된 점이라고 하면, 그 투영 특징 점의 유사도 $S(\mathbf{x}, \mathbf{x}_p^{c_{NFI}}, \mathbf{x}_r)$ 은 식 (9)에서와 같이 ERN에 의해 얻어진 유사도와 융합되어서 최종 결론을 내도록 하였다. 여러 융합형태의 분류기에서 본 논문의 제안 방식인 NFA와 ERN의 융합방법이 NFL과

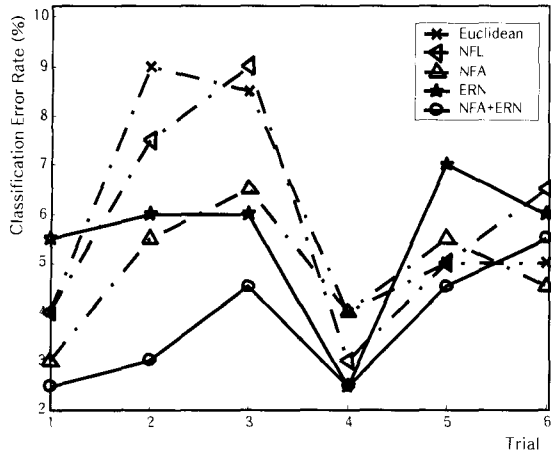


그림 9. 6 회 시도에 따른 분류 에러율 비교
Fig. 9. Comparison of classification error rate for 6 trials.

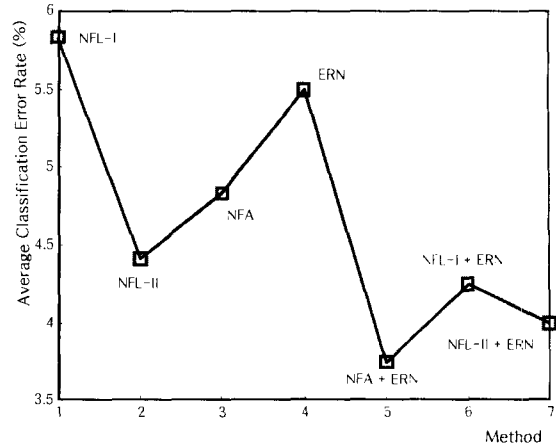


그림 11. 융합 분류 방법에 따른 평균 분류에러율의 비교
Fig. 11. Comparison of average classification error rate using hybrid methods.

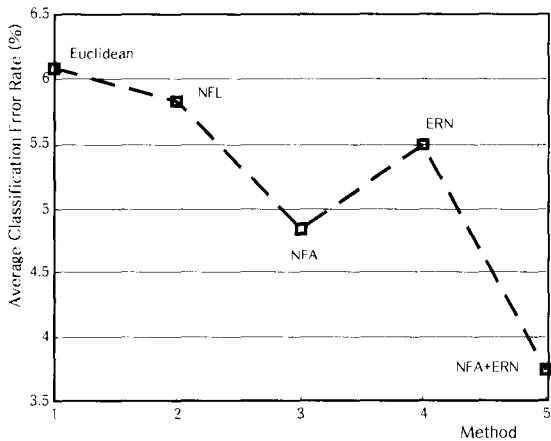


그림 10. 6 회 의 시험에 대한 평균 분류에러율 비교
Fig. 10. Comparison of average classification error rate for 6 trials.

ERN분류기에 의한 경우보다 낮은 분류 에러율을 나타내었다.

3. FRN 분류기의 수와 가상데이터 생성 임계치(α)에 따른 성능

서로 다른 구조를 갖는 FRN을 이용하는 경우, 동일한 입력과 출력에 대해서도 학습하는 경우에도 서로 다른 특징을 갖는 분류기로 동작할 것을 기대 하였고, 이러한 것이 융합되는 경우에 전체적으로 보다 나은 일반화 성능을 낼 것으로 기대했다. <그림 13>은 ERN을 구성하는 FRN의 수에 따른 평균 오류율을 보였다. <그림 14>의 결과는 5개의 FRN을 사용한 경우에 서로 다른 학습/시험 데이터 조합이 나타내는 성능의 평

균 오류율이다. 예를 들어, 2 FRN의 경우는 5 개의 FRN에서 2개를 선택하는 조합의 수인 10 가지 경우에 대한 평균오류율이다. ERN 분류기의 오류율과 그에 따른 융합 분류기의 오류율은 FRN의 수가 증가함에 따라 감소하는 것을 보여주고 있고, <그림 14>는 가상 학습데이터를 생성할 때 사용하는 임계치 값(α)에 따른 성능변화를 보인다. 최고의 성능은 $\alpha = 0.5$ 에서 96.33%를 나타내고 있다.

4. 실험 결과의 예

위의 실험을 수행하면서 NFA와 ERN 그리고 NFL 방식에 의해 찾아진 결과의 예를 보였다. <그림 15>와 <그림 16>의 시험영상은 시험영상을 고유얼굴공간에 투사하여 얻은 특징 벡터로부터 재 복구한 영상이다. <그림 15>는 NFA, ERN이 각각 정확하게 분류한 반면에, NFL 방식은 실패한 경우를 보인다. <그림 15(a)>의 두 얼굴 영상은 NFL에 의해 찾아진 두 특징 점에 해당하는 영상으로, 왼쪽 결과를 보면, 시험영상의 수염보다는 자세의 보간에 의한 영향이 크게 작용한 것으로 보이며, 오른쪽 결과는 시험 영상과 선택한 얼굴영상의 공통점이 상당히 많은 것으로 보인다. 즉, 대머리와 수염 그리고 안경을 착용한 것까지 유사성을 지니고 있다. <그림 15(b)>는 NFA에 의해 찾아진 특징 점으로 복구된 영상이며, 괄호 안의 숫자는 시험 영상과의 유사도 값을 나타낸다. 이 경우에 있어서 융합 분류기는 식 (9)에 의하여 찾아진 클래스가 동일하므로, 각 분류기의 유사도와 관계없이 찾아진 클래스를 최종

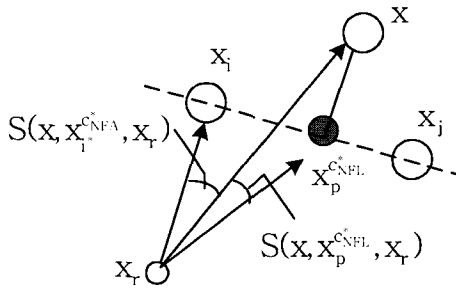


그림 12. 질의 특징점 x 로부터 $S(x, x_p^{cNFL}, x_r)$ 를 얻는 과정

Fig. 12. The diagram of obtaining $S(x, x_p^{cNFL}, x_r)$ from query feature x .

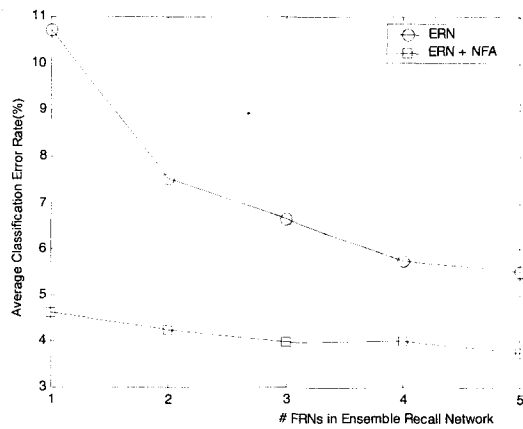


그림 13. Ensemble 회상망을 구성하는 FRN의 수에 따른 성능비교

Fig. 13. Performance comparison according to the number of FRN in ensemble recall network.

클래스로 결정한다.

<그림 16>의 결과는 NFA, NFL 방법이 실패한 경우에 대하여, ERN이 정확하게 분류한 경우의 예를 보였다. 이 경우에 융합 분류기는 식 (9)에 의해 유사도의 값이 크게 나타난 ERN의 결과를 선택하게 된다. 왼쪽 시험 영상에 대하여, NFL 분류기가 찾은 두 개의 영상은 자세의 보간에 의한 영향이 큰 것으로 보이며, <그림 16(b)>의 NFA도 얼굴 자세에 의해 영향을 받았음을 알 수 있다. 반면에 ERN은 정면얼굴로의 회상을 통해 정확한 결과를 얻고 있다.

VII. 결론

가상 학습 데이터의 생성과 각 정보(Angular Infor-

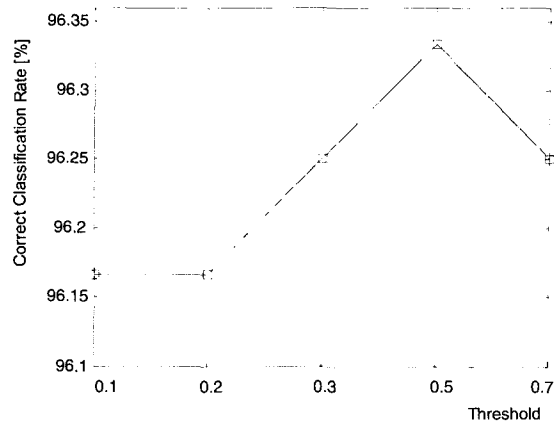


그림 14. 가상데이터 생성을 위한 임계값에 따른 성능비교

Fig. 14. Performance comparison according to the threshold.

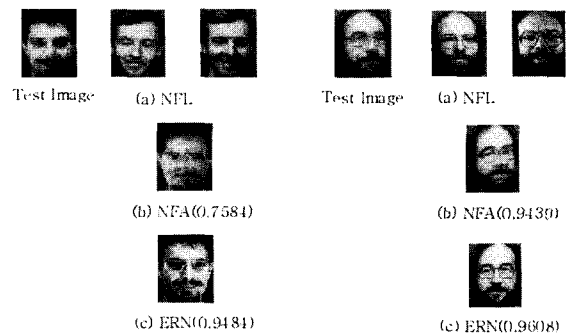


그림 15. NFL 분류기가 실패한 두가지 예

Fig. 15. Two examples of classification results in which NFL fails.

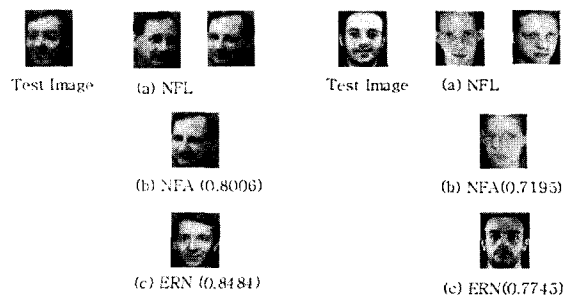


그림 16. NFL, NFA 분류기가 실패한 두가지 예

Fig. 16. Two examples of classification results in which the NFL and NFA fails.

mation)에 기반한 융합 분류기를 제안하였다. 제안 분류기의 주요 특징 점은,

1) 가상 학습 데이터가 각 클래스의 공간분포에 따라 특징 점들 중간 지점이나 특징 점들이 형성하는 표면

에 생성되며, 그에 따라 전체 학습 데이터의 크기를 확장하여 시스템의 일반화 성능향상을 도모하였다.

2) 두 개의 부분분류기가 입력정보를 서로 다르게 처리하도록 하였다. 즉, 각 정보(angular information)에 기반한 방법과 신경망의 비선형 투사기능과 일반화 기능을 이용한 정면얼굴 회상에 의한 방법을 융합하였다. 서로 다른 처리방법에 의한 결과를 융합함으로써, 두 부분분류기의 장점을 포함하는 융합 분류기를 제안하였다.

3) 두 부분분류기의 분류결과와 유사도를 최종 결과에 이용함으로써 한 개의 분류기를 사용하는 경우, 즉 NFA, ERN, NFL, 보다 안정된 결정을 내리고 있으며, 제안한 시스템은 6종류의 실험/시험 데이터 조합에 대하여 평균 96.33%의 성능을 나타내고 있다. 이것은 동일 데이터 베이스에 대하여 가장 좋은 성능으로 알려진 NFL방식의 평균 오류율에 비해 38.8% 개선된 결과이다. 앞으로의 연구는 보다 다양한 자세와 표정 변화에 대응할 수 있는 최적의 회상망(Recall Network)에 대하여 연구할 것이다.

참 고 문 헌

- [1] Stan Z. Li, Juwei Lu, Face Recognition Using the Nearest Feature Line Method, IEEE Transactions on Neural Networks, 10(2), pp. 439~443, 1999.
- [2] Y. Moses, Y. Adini, S. Ullman, Face recognition: The problem of compensating for changes in illumination direction, Proc. European Conf. Computer Vision, vol. A, pp. 286~296, 1994.
- [3] A. Samal, P. A. Iyengar, Automatic recognition and analysis of human faces and facial expression: A survey, Pattern Recognition, 25, pp. 65~77, 1992.
- [4] R. Brunelli, T. Poggio, Face Recognition: Features versus templates, IEEE Trans. Pattern Anal. Machine Intelligence, 15, pp. 1042~1052, 1993.
- [5] M. A. Turk, A. P. Pentland, Eigenfaces for recognition, J. Cognitive Neuroscience, 4(1), pp. 71~86, 1991.
- [6] A. J. O'Toole, H. Abdi, K. A. Deffenbacher, D. Valentin, Low dimensional representation of faces in higher dimensions of the face space, J. Opt. Soc. Amer. A(10), pp. 405~410, 1993.
- [7] Jun Zhang, Yong Yan, Martin Lades, Face Recognition: Eigenface, Elastic Matching, and Neural Nets, Proceeding of the IEEE, 85(9), pp. 1423~1435, 1997.
- [8] in Jang, Sungzoon Cho, Ensemble Learning Using Observational Learning Theory, International Joint Conference on Neural Network, 1999.
- [9] S. Cho, K. Cha, Evolution of neural network training set through addition of virtual samples, International Conference on Evolutionary Computation, Nagoya, Japan, pp. 685~688, 1996.
- [10] L. Holmstrom, P. Koistinen, Using additive noise in back-propagation training, IEEE transactions on Neural Networks, 3, pp. 24~38, 1992.
- [11] C. M. Bishop, Training with noise is equivalent to tikhonov regularization, Neural Computation, 7(1), pp. 108~116, 1995.
- [12] Y. Abu-Mostafa, Hints, Neural Computation, 7, pp. 639~671, 1995.
- [13] L. Breiman, Using convex pseudo-data to increase prediction accuracy, Technical report, University of California, 1999.
- [14] Roberto Battiti, Anna Maria Colla, Democracy in Neural Nets: Voting Schemes for Classification, Neural Networks, 7(4), pp. 691~707, 1994.
- [15] Yong Liu, Xin Yao, Simultaneous Training of Negatively Correlated Neural Networks in an Ensemble, IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 29(6), (1999).
- [16] D. V. Sridhar, E. B. Bartlett, R. C. Seagrave, An Information Theoretic Approach for Combining Neural Network Process Models, Neural Networks, 12, pp. 915~926, 1999.
- [17] Nikhil R. Pal, Vijay Kumar Eluri, Two Efficient Connectionist Schemes for Structure Preserving Dimensionality Reduction, IEEE Transactions on

- Neural Networks, 9(6), pp. 1142~1154, 1998.
- [18] Anil K. Jain, Robert P. W. Duin, Jianchang Mao, Statistical Pattern Recognition: A Review, IEEE Trans. Pattern Analysis and Machine Intelligence, 22(1), pp. 4~37, 2000.
- [19] S. Lawrence, C. Lee Giles, A. C. Tsoi, A. D. Back, Face recognition: A convolutional neural network approach, IEEE trans. Neural Networks, 8, pp. 98~113, 1997.
- [20] M. Bichsel, A. P. Pentland, Human face recognition and the face image set's topology, CVGIP: Image Understanding, 59, pp. 254~261, 1994.
- [21] Laurenz Wiskott, Kean-Marc Fellows, Norbert Kruger, and Christoph von der Malsburg, Face Recognition by Elastic Bunch Graph Matching, IEEE Trans. On PAMI, 19(7), pp. 775~779.
- [22] Chengjun Liu and Harry Wechsler, Learning the Face Space-Representation and Recognition, Proc. of the 15th International Conf. on Pattern Recognition, Vol.1 pp. 249~256.
- [23] Baback Moghaddam, Wasiuddin Wahid, and Alex Pentland, Beyond Eigenface: Probabilistic Matching for Face Recognition, Proc. of the 2nd International Conf. on Automatic Face and Gesture Recognition, pp. 30~35, 2000.
- [24] Srinivas Gitta and Harry Wechsler, Modular Forensic Architecture, Face Recognition from Theory to Applications, Springer, pp. 327~347, 1997.
- [25] S. Gutta, J. Huang, B. Takacs, and H. Wechsler, Face Recognition Using Ensembles of Networks, Proc. of the 13th International Conf. on Pattern Recognition, Vol. 4, 1996.
- [26] 평은섭, 김병화, 안현식, 김도현, 무게중심을 이용한 자동얼굴인식 시스템의 구현, 전자공학회 논문지, 제 33권, B편, 제 8호, pp. 114~122, 1996
- [27] 이상영, 함영국, 박래홍, 뉴로퍼지 알고리즘을 이용한 얼굴인식, 전자공학회 논문지, 제 32권, B편, 제 1권, pp. 50~63, 1995
- [28] Xin Yao, and Yong Liu, Making Use of Population Information in Evolutionary Artificial Neural Networks, IEEE Trans. on Systems, Man and Cybernetics-Part B: Cybernetics, Vol. 28, No. 3, 1998.
- [29] David E. Moriarty and Risto Miikkulainen, Hierarchical Evolution of Neural Networks, International Conference on Evolutionary Computation, pp. 428~433, 1998.
- [30] 류연식, 오세영, 고유얼굴과 신경망 융합을 이용한 정면얼굴의 Recall 및 인식, 제 12회 영상처리 및 이해에 관한 워크샵 논문지, pp. 528~533, 2000
- [31] Daniel P. Huttenlocher, Ryan H. Lillien, and Clark F. Olson View-Based Recognition using an Eigenspace Approximation to the Hausdorff Measure, IEEE Trans. on PAMI, Vol 21, No. 9, pp. 951~955, 1999.

저 자 소 개



柳 淵 植(正會員)

1988년 : 한양대학교 전기공학과 학사. 1990년 : 포항공과대학교 전자전기공학과 석사. 1990년~2001년 12월 : LG 전자. 2002년 : 포항공과대학교 전자전기공학과 박사.

현재 : LG 엔시스 개발담당 보안그룹 책임연구원. <주관심분야 : 생체인식시스템, 네트워크/호스트 보안시스템, 신경망 및 진화 연산 응용 시스템>



吳 世 泳(正會員)

1974년 : 서울대학교 전자공학과 학사. 1978년 : Case Western Reserve Univ. 전자공학과 석사. 1981년 : Case Western Reserve Univ. 전자공학과 박사. 1981년~1984년 : University of Illinois 조

교수. 1985년~1988년 : University of Florida 조교수. 현재 : 포항공과대학교 전자전기공학과 정교수. <주관심분야 : 신경망, 진화연산, 이동로봇, 지능자동차, 지능시스템>