퍼지 및 지능시스템학회 논문지 2001, Vol. 1, No. 1, pp. 39-44

영상영역 기반 퍼지 신경망을 이용한 얼굴 검출

Region Based Fuzzy Neural Networks for Face Detection

이창수* · 이정훈**
Changsu Lee* and Frank Chung-Hoon Rhee**

아이캐시
** 한양대학교 전자통신전파공학과

요 약

본 논문에서는 디지털 영상에서 얼굴 영상 검출을 위해 픽셀의 퍼지 소속도를 이용하여 신경망으로 학습하는 퍼지 신경망을 이용한 얼굴영상 검출을 제안한다. 입력 영상의 피라미드 영상에서 추출된 **20×20** 윈도우 영상 안의 각 픽셀의 소속도로 얼굴 영상 패턴을 학습하여 얼굴 영상을 검출하는 방법은 단순히 영상의 픽셀 값 하나씩만을 고려해서 각 픽셀의 소속도를 고려하여 수행하는 얼굴 영상 분할보다 얼굴 영상을 더 정확하고 인식률이 높게 검출해 낼 수 있다.

ABSTRACT

In this paper, we propose a region based fuzzy neural network method for face detection. In the proposed method, fuzzy membership degrees are assigned to preprocessed 20×20 window face and nonface image regions. These fuzzy membership degrees are then inputed into a neural network to be trained. After training, the output value of the neural network is interpreted as the degree of which a given window is a face or nonface region. If the window is determined to contain a face, post processing is then performed. Experimental results show that face images are found to be more accurately detected than using conventional neural networks.

Keywords : face detection, fuzzy membership degree, neural networks.

1. 서 론

일반적으로 영상에서 얼굴 영상을 검출하고 인식하는 알고리즘은 패턴 인식 연구에 있어서 인간과 컴퓨터의 상호작용의 연구라는 면에서 아주 중요한 문제이다. 또한 얼굴 영상 검출은 얼굴 영상 인식 연구에 아주 필수적인 부분이다. 얼굴 영상 검출과 인식은 약25여년 동안 연구되어 왔으며, 1970년대의 얼굴의 눈,코,입의 특징들에만 기초한 연구는 컴퓨터의 계산기술 부족으로 인해 소수의 테스트만이 적용되었다. 그 후의 1980년대의 잠복기를 지나 1990년대 초에 얼굴 영상 검출과 인식에 대한 연구가 되살아 났고, 근 10여년 동안 연구가 활발하게 진행되어 왔다. 그 동안 진행되어 왔던 얼굴 영상 검출에 대한 연구는 여러 가지 방법으로 연구되어 왔다. Principle component analysis(PCA)를 이용한 연구[1-3], 신경망을 이용한 연구[4-7], 패턴매칭을 이용한 방법[8-9] 등으로 연구되어 왔다.

본 논문에서는 얼굴 영상 검출을 위해 다음과 같은 알고리즘을 제시한다. 크게 전처리 과정, 퍼지 소속도 추출 과정, 신경망으로 테스트하는 과정, 후처리 과정으로 나누어진다. 입력 영상을 서브 샘플링해서 입력 영상 피라미드를 만들고 각 입력 영상 피라미드에서 20×20 픽셀 크기의 작은 윈도우 영상을 추출한다. 이 추출한 윈도우 영상은 전처리 과정을 거치고 각 픽셀의 퍼지 얼굴 영상 소속도를 구한다. 이 소속도를 신경망의 입력으로 준다. 신경망의 결과 값으로 주어진 윈도우의 영상이 얼굴에 얼마나 가까운지 알 수 있게 된다. 그리고 후처리 과정에서는 신경망의 결과 값이 임계값(threshold) 보다 높은 값이면 해당되는 입력 영상 피라미드에 얼굴의 위치를 표시한다. 그리고 각 입력 영상의 피라미드에서 잘못 찾은 영역을 없애고 확실하게 찾은 얼굴 영역을 표시한다.

2. 본 론

본 논문에서는 모든 처리 과정이 20×20 윈도우 영상 단위로 이루어진다. 따라서 얼굴의 크기가 20×20 픽셀보다 큰 얼굴을 검출해내기 위해서, 입력 영상을 1.2의 일정한 비율로 서브 샘플링해서 그림 1과 같이 각 단계별로 입력 영상 피라미드를 만든다. 각 단계의 입력 영상 피라미드에서 윈도우 영상은 한 피라미드 영상 안에서 한 픽셀 단위로 움직이면서 20×20 윈도우 영상을 추출해낸다[4].

접수일자 : 2000년 11월 18일
완료일자 : 2001년 01월 15일

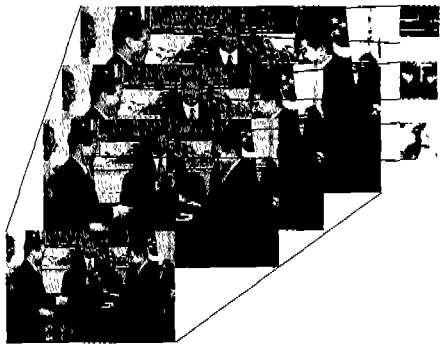

그림 1. 입력 영상 피라미드
Fig. 1. Input image pyramid

2.1. 전처리 과정

추출된 20×20 윈도우 영상은 다음의 두 가지의 전처리 과정을 거치게 된다[4],[12]. 첫째로 명암교정 과정으로 추출된 윈도우 영상의 빛에 대한 변화를 보상하기 위해서 사용된다. 윈도우 영상 전체에 걸친 픽셀 값의 변화에 대한 일차함수를 구한다(linear function fit). 그리고 원래의 추출된 윈도우 영상과 일차함수 영상의 차를 구하면 빛의 밝기에 대해 보상된 영상이 나오게 된다. 다음으로 히스토그램 균일화의 과정으로 윈도우 영상 안의 픽셀 값의 분포를 확장 시켜서 밝기에 대한 대비를 향상시킨다. 그림 2는 전처리 과정을 거치기 전의 영상과 후의 영상을 보여주고 있다.

(a) 전처리 전 영상 (b) 전처리 후 영상
(a) Images before preprocess (b) Images after preprocess
그림 2. 전처리 과정
Fig. 2. Preprocess

2.2. 얼굴영상 패턴과 배경영상 패턴 생성

일반적인 사람의 얼굴 영상에 대한 기본형(face image prototype)을 정의하고 만들기 위해서 많은 양의 영상으로부터 사람 얼굴만 추출해낸다. 이것은 얼굴 영상에 대한 학습 패턴 생성 과정으로서 정렬 알고리즘(alignment algorithm) [4],[12]을 이용한다. 정렬 알고리즘은 얼굴이 있는 영상들로부터 각 얼굴들의 특징값(눈,코,입)의 정보를 기반으로 하여 20×20 윈도우 영상에 맵핑하는 알고리즘이다[4],[12]. 본 논문에서는 CMU, World Wide Web으로부터 얻은 영상들로 정렬 알고리즘을 이용해서 한 얼굴마다 10개의 얼굴 영상들을 임의로 회전시키고, 움직이고, 크기를 변화시켜서 3,800개의 얼굴 영상들을 생성하였다. 그리고 얼굴 영상 기본형(face image prototype)은 얼굴 영상들의 평균값으로 생성하였다. 그림 3은 얼굴 영상 기본형 영상을 나타내고 그림 4는 정렬 알고리즘을 이용해서 생성한 얼굴 영상들을 보여주고 있다.

그림 3. 얼굴 영상 기본형
Fig. 3. Face image prototype

그림 4. 얼굴 영상들
Fig. 4. Face images

일반적으로 사람 얼굴의 형태는 정해져 있는데 반해서, 얼굴이 아닌 형태는 특별하게 정해져 있지 않다. 그러므로 배경 영상의 기본형(nonface image prototype)을 만들어 내는 것은 불가능한 일이다. 이러한 이유 때문에 얼굴이 없는 배경 영상들로부터 임의의 20×20 크기의 배경 영상들을 만들어 낸다. 실험에서는 임의의 배경 영상들로부터 그림 5에서 보는 것과 같은 11,112개의 얼굴 없는 배경 영상들을 생성했다.

그림 5. 배경 영상들
Fig. 5. Nonface images

2.3. 퍼지 소속도 함수 설계

생성된 얼굴 영상들로부터 얼굴 영상의 퍼지 소속도 함수를 만든다. 본 논문에서는 영상의 위치에 대한 픽셀 기반의 가우시안 함수의 퍼지 소속도 함수를 구한다. 이 방법은 히스토그램 기반의 소속도 함수[10-11]를 구하는 과정과 동일하다. 영상의 픽셀의 위치에 대한 픽셀 값을 기반으로 해서 구한 퍼지 소속도 함수는 어떤 한 영상에서 그 위치의 픽셀 값이 얼마나 얼굴 영상에 가까운지를 소속도로 나타내게 된다.

2.3.1. 정상화(normalized)된 픽셀 값 기반의 퍼지 소속도 함수

영상으로부터 주어지는 정상화(normalized)된 픽셀 값은 가능성 분포 함수(possibility distribution function)로 간주할 수 있다[10-11]. 주어지는 k 개의 특징값의 범위가 [0,1]에 오도록 정상화한 후에 그림 6과 같이 각각의 특징값들을 Q 레벨의 가우시안 퍼지 집합들로 나타낸다.

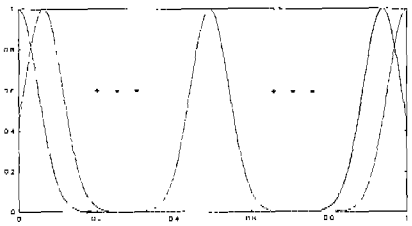

그림 6. Q 레벨의 Q개 퍼지 집합
Fig. 6. Q-level fuzzy set

여기서 u^{th} 특징값 벡터를 $x_u^j = (x_{1u}^j, \Lambda, x_{Ku}^j)$로 나타낼 때, k번째 특징값의 영역에서 정의되어지는 클래스에 대한 멤버쉽 함수 μ_k^j는 식(3.1)과 같다.

$$\mu_k^j(i) = \frac{1}{N^j} \sum_{u=1}^{N^j} f_{ku}^j(x_{ku}^j) \tag{3.1}$$

$f_k^j(0)$는 클래스 j에 대해서 x_k의 영역에 정의된 i번째 멤버쉽 함수이고, N'은 j클래스의 샘플 개수이다. 식(3.1)은 Q개의 퍼지 레벨에서의 특징값의 소속도를 준다([10-11]).

2.3.2. 퍼지 소속도 함수 설계

퍼지 소속도 함수 $m_k^j(x_k)$가 여러 개의 최고점(multiple peaks)을 가지는 형태라면, 퍼지 소속도 함수는 여러 개의 가우시안 함수의 합으로 모델화할 수 있다. 그리고 가우시안 함수를 소속도 함수로 더 유동성 있게 모델화하기 위해서 비대칭 가우시안 함수로 나타낼 수 있다. 소속도 함수 $m_k^j(x_k)$가 다중 최고점을 가지고 있다면, 소속도 함수는 식(3.2)와 같이 비대칭 가우시안 함수의 합으로써 모델화할 수 있다[10-11].

$$m_k^j(x_k) \approx \tilde{F}_k^j(x_k) = \sum_{i=1}^{L_k^j} a_{ki}^j F_{ki}^j(x_k) \qquad (3.2)$$

$$F_{ki}^j(x_k) = \begin{cases} G(x_k, c_{ki}^j, \sigma_{1ki}^j) & x_k < c_{ki}^j \\ G(x_k, c_{ki}^j, \sigma_{2ki}^j) & x_k \geq c_{ki}^j \end{cases}$$

여기서 $F_{ki}^j(x_k)$함수는 클래스 j의 특징값의 파라미터 $p_{ki}^j = (a_{ki}^j, c_{ki}^j, \sigma_{1ki}^j, \sigma_{2ki}^j)$에 해당하는 비대칭 가우시안 함수가 된다. $a_{ki}^j, c_{ki}^j, \sigma_{ki}^j$는 각각 가우시안 함수의 높이, 평균값, 표준편차 값을 나타낸다.

2.3.3. 가우시안 파라미터 학습

얼굴 영상의 소속도 함수를 앞에서 생성한 얼굴 영상들로부터 학습시킨다. 학습이 끝난 후에는 얼굴 영상의 소속도 함수의 가우시안 파라미터들을 얻게 된다. 얼굴 영상의 소속도 함수를 비대칭 가우시안 함수의 합으로 영상으로부터 학습시킬 때, 식(3.3)과 같은 목적함수(objective function)를 최소화할 수 있다.[10-11]

$$J_k^j(p_k^j) = \frac{1}{2}\left(\sum_{i=1}^{L_k^j} h_k^j(x_k, p_{ki}^j) - m_k^j(x_k)\right)^2 \qquad (3.3)$$

$h_k^j(x_k, p_{ki}^j)$는 클래스 j의 특징값 x_k에 대한 멤버쉽 함수 $m_k^j(x_k)$를 모델화하기 위해서 선택한 파라미터의 함수이다. 그리고 파라미터 벡터를 학습, 갱신시켜서 최종적인 값을 구하기 위해서 식(3.4)의 gradient descent 방법을 사용한다.

$$p_{ki}^{j(new)} = p_{ki}^{j(old)} - \rho \frac{\partial J_k^j}{\partial p_{ki}^j} \qquad (3.4)$$

실험에서 퍼지 소속도 함수로 사용되는 가우시안 함수의 높이는 1.0으로, 평균값은 특징값이 정상화된 값의 최고값으로, 표준 편차의 초기치는 0.1로 실험을 했다. 학습상수 ρ는 0.01로 초기치를 정해주었다.

2.3.4. 얼굴 영상의 퍼지 소속도 함수

제시한 방법으로 생성한 얼굴 영상의 비대칭 가우시안 퍼지 소속도 함수는 그림 7과 같다.

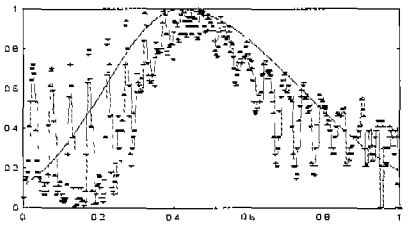

그림 7. 얼굴 영상의 가우시안 퍼지 소속도 함수, *face(i)*
Fig. 7. Gaussian fuzzy membership function of face images, *face(i)*

고려할 것은 영상 안에서 위치에 대한 픽셀 값으로 구한 퍼지 소속도 함수이다. 즉, 위치에 대한 픽셀 값의 퍼지 소속도 함수이다. 그러므로 주어지는 20×20 영상에서 한 위치의 픽셀 값이 얼굴 영상에 얼마나 가까운지를 퍼지 소속도로 나타내기 위해서 식(3.5)와 같이 얼굴 영상의 퍼지 소속도를 모델화 한다.

$$Face(i) = 1.0 - |Intensity(i) - face(i)| \qquad (3.5)$$

i : 영상의 정상화된 위치
Intensity(i) : 위치i의 정상화된 픽셀 값
Face(i) : 위치i의 얼굴영상 퍼지 소속도

2.3.5 배경 영상의 퍼지 소속도 함수

배경 영상에 대한 기본형은 정의 하기가 어렵기 때문에 본 논문에서는 배경 영상에 대한 가우시안 소속도 함수의 파라미터를 실험적으로 모델화 하였다. 그림 8의 형태를 보면 알 수 있듯이, 파라미터의 평균값을 얼굴 영상 경우와 같게 정했다. 그리고 얼굴 영상 소속도 함수의 값과 반대의 의미를 가지게 하기 위해서 그림 8과 같이 얼굴 영상 소속도의 반대의 의미를 부여하기 위해서 정상화된 위치 가 0과 1의 위치일 때 가장 높은 값을 가지도록 모델화하였다.

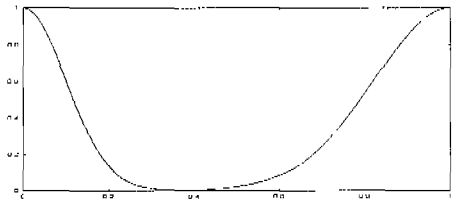

그림 8. 배경 영상의 가우시안 퍼지 소속도 함수, *nonface(i)*
Fig. 8. Gaussian fuzzy membership function of nonface images, *nonface(i)*

이 퍼지 소속도 또한 영상 안에서 위치에 대한 픽셀 값으로 구한 퍼지 소속도 함수이므로 배경 영상의 퍼지 소속도 함수는 식(3.6)과 같다.

$$Nonface(i) = 1.0 - |Intensity(i) - nonface(i)| \qquad (3.6)$$

Nonface(i):위치i의 배경영상 퍼지 소속도

결과적으로 얼굴 영상과 배경 영상의 각각에 대한 가우시안 퍼지소속도 함수의 파라미터는 표1에서 보는 것과 같다.

<표 1. 가우시안 퍼지 소속도 함수 파라미터 >
Tabel 1. Parameters of gaussian fuzzy membership function

	높이	평균	좌편차	우편차
face(i)	1.0	0.415	0.20	0.32
nonface(i)	1.0	0.415	0.18	0.10

2.3.6. 주어진 위치에 대한 픽셀의 소속도

얼굴과 배경 영상 소속도 함수를 가지고 주어진 20×20 윈도우 영상의 위치에 대한 픽셀의 최종적인 소속도는 식(3.7)과 같다.

$$Final(i) = MAX(Face(i), Nonface(i)) \tag{3.7}$$

위의 최종 소속도는 얼굴과 배경 영상의 소속도에 각각 얼마나 가까운지 구해서 두 소속도 중에 높은 값을 가지는 퍼지 소속도를 의미한다.

2.4. 신경망의 구조와 학습

신경망의 구조는 3-layer fully connected network이고, 각 레이어 뉴런의 개수는 각각 400개, 20~65개, 2개이다. 신경망의 학습 방법은 "bootstrap manner"[4] 의 방법을 채택하였다. 이 방법은 다음과 같다. 생성한 얼굴 영상들과 배경 영상들로 역 학습 방법을 통해 신경망을 학습시킨다. 그리고 이 학습된 신경망에 사람의 얼굴이 없는 배경 영상을 입력으로 주고 얼굴이라고 오검출된 영상들을 모아서 다시 이 오검출된 영상들을 배경 영상으로 신경망을 학습시키는 방법이다. 이것은 신경망이 얼굴 영상과 배경 영상에 대한 더 정확한 경계를 학습할 수 있게 해준다. 그림 9는 신경망의 구조를 보여주고 있다.

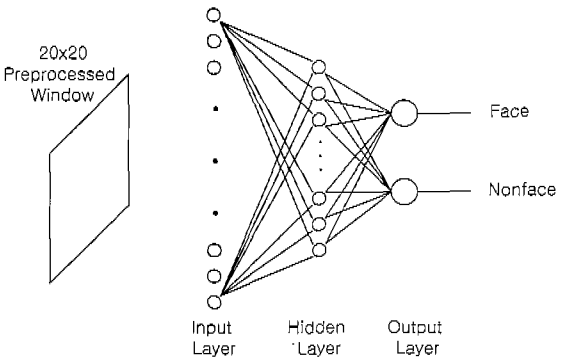

그림 9. 신경망의 구조
Fig. 9. Neural network structure

2.5. 후처리 과정

후처리 과정에서는 오검출된 영역을 제거하고, 검출된 얼굴 영역에 표시해준다[13]. 후처리 과정은 오검출 영역 제거의 방법은 다음과 같다. 우선 먼저 하나의 영상 피라미드 안에서 오검출된 영역을 제거하고 여러 다른 영상 피라미드 안에서의 오검출된 영역을 제거한다.

하나의 영상 피라미드 안에서 오검출 영역의 제거는 다음과 같다. 신경망의 값이 높게 나온 순서대로 윈도우 영상의 위치 좌표들을 정렬시키고, 각각의 윈도우 영상의 위치가 겹치는 영역이 있으면 신경망의 결과값이 낮게 나온 윈도우 위치 좌표를 제거하고 겹치는 윈도우 영상의 위치 좌표를 중심으로 겹치는 횟수를 고려한다. 이러한 방법으로 하나의 영상 피라미드 내의 겹치는 윈도우 영역을 제거한다. 그림 10은 하나의 영상 피라미드 내에서 겹치는 영역들을 고려해서 오검출된 영역을 제거를 보여주고 있다.

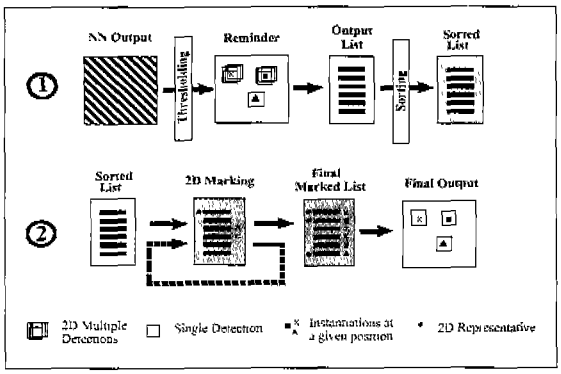

그림 10. 한 영상 피라미드 오검출 영역 제거
Fig. 10. Removal of mis-detected regions in one pyramid image

그리고 다른 여러 피라미드 영상의 오검출된 영역을 제거하는 방법은 다음과 같다. 우선 모든 피라미드 영상을 앞의 방법과 같이 각각 하나의 영상 피라미드에서 오검출된 영역을 제거한다. 그리고 각각의 다른 많은 피라미드 영상에서 겹치는 윈도우 영역의 좌표를 고려하고, 겹치는 횟수를 고려해서 여러 다른 피라미드 영상에서 오검출된 영역을 제거한다. 그림 11은 여러 피라미드 영상에서의 오검출 영역 제거를 보여주고 있다.

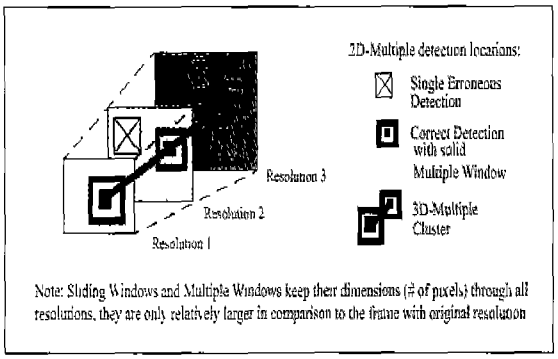

그림 11. 여러 영상 피라미드 오검출 영역 제거
Fig. 11. Removal of mis-detected regions in many pyramid images

2.6. 시스템 구성도

지금까지 설명해온 퍼지 신경망을 이용한 얼굴 영상 검출의 전체 시스템을 그림 12에서 나타내고 있다.

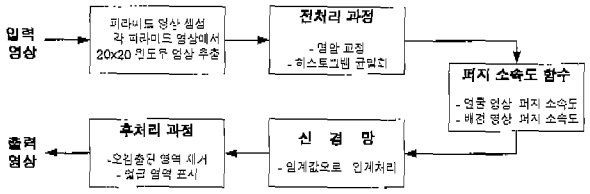

그림 12. 시스템 구성도
Fig. 12. System diagram

3. 실험 결과

본 논문에서 제시한 퍼지 소속도로 신경망을 학습하여 얼굴 영상 검출을 수행한 결과는 표2에 나타나 있다. 퍼지 소속도를 부여하지 않고 신경망을 학습시켜 수행했을 때보다 퍼지 소속도를 부여하여 신경망을 학습시킨 경우 훨씬 높은 인식률을 얻을 수 있었다. 또한 퍼지 소속도로 신경망을 학습시키면 퍼지 소속도를 사용하지 않은 일반적인 신경망보다 더 적은 수의 히든 뉴런의 개수로 동일한 인식률을 얻을 수 있었다.

이는 퍼지 소속도를 이용하여 신경망을 학습시키는 것이 얼굴 영상 검출의 방법으로서 더 향상된 방법임을 보여준다. 아래 그림 13은 실험 결과 영상들을 보여주고 있다.

<표 2. 일반 신경망과 퍼지 신경망의 비교>
Table 2. Comparison of conventional neural network and fuzzy neural network

	히든 뉴런 개수	인식률(%)
일반 신경망	20	71.2
	30	74.6
	45	80.4
	65	85.7
퍼지 신경망	20	75.2
	30	85.7
	45	90.1
	65	92.7

그림 13. 결과 영상들
Fig. 13. Result images

4. 결 론

본 논문에서는 디지털 영상에서 얼굴 영상 검출을 위해서 영상의 픽셀값의 퍼지소속도를 구해서 신경망으로 학습하는 퍼지 신경망을 이용한 얼굴 영상 검출을 제안하였다. 더 나은 검출을 위한 연구로서 더 향상된 얼굴 영상의 정의와 함께 퍼지 추정 시스템(fuzzy inference system)으로 얼굴 영상을 검출하는 시스템을 현재 연구하고 있다.

참 고 문 헌

[1] M. Turk, A. Pentland, "Eigenfaces for Recognition," Journal of Cognitive Neuroscience, vol. 3, no. 1, 1991.
[2] A. Pentland, T. Starner, N. Etcoff, A. Masoiu, O. Oliyide, and M. Turk, "Experiments with Eigenfaces," IJCAI Aug. 1993.
[3] B. Moghaddam, A. Pentland, "Probabilistic Learning for Object Detection," 5th ICCV, June 1995.
[4] A. Henry, B. Shumeet, and T. Kanade, "Neural Network-Based Face Detection," IEEE Trans. PAMI, vol. 20, no. 1, Jan. 1998.
[5] S.-H. Lin, S.-Y. Kung, L.-J. Lin, "Face Recognition /Detection by Probabilistic Decision -Based Neural Network," IEEE Trans. Neural Networks, vol. 8, no. 1, Jan. 1997.
[6] S. Lawrence, C. Giles, A. Tsoi, and A. Back, "Face Recognition: A Convolutional Neural-Network Approach," IEEE Trans. Neural Networks, vol. 8, no. 1, Jan. 1997.
[7] C. Neubauer, "Evaluation of Convolutional Neural Networks for Visual Recognition," IEEE Trans. Neural Networks, vol. 9, no. 4, July 1998.
[8] H. Wu, Q. Chen, and M. Yachida, "Face Detection From Color Images Using a Fuzzy Pattern Matching Method," IEEE Trans. PAMI, vol. 21, no. 6, June 1999.
[9] F. Weber, A. H. Hernandez, "Face Location by Template Matching with a Quadratic Discriminant Function," plettac electronic security GmbH.
[10] F. C-H. Rhee, "Fuzzy Rule Generation and Inference Methods for Pattern Recognition and Computer Vision," Ph.D. Thesis.
[11] 이창수, 이정훈, "퍼지 소속도를 이용한 얼굴 영상 분할," 한국퍼지 및 지능 시스템 학회 2000 춘계학술 대회, vol. 10, no.1, 5월 2000.
[12] K.-K. Sung, "Learning and Example Selection for Object and Pattern Detection," PhD thesis, MIT AI Lab, Jan. 1996.
[13] B. Fasel, "Fast multi-scale face detection," IDIAP-COM 98-04, IDIAP, 1998.

저 자 소 개

이 창 수 (Changsu Lee)

1995~1999년 : 한양대학교 전자공학 학사
1999~2001년 2월: 한양대학교 전자통신 전파공학 석사
현재 : ㈜아이캐시 개발본부.

관심분야 : 퍼지이론, 신경망, 패턴인식, 컴퓨터비젼, 암호학.
E-mail : cslee@icash.co.kr

이 정 훈 (Frank Chung-Hoon Rhee)

1981~1985년 : University of Southern California (USC), Dept. Electrical Eng., BSEE.
1985~1987년 : University of Missouri, Dept. of Electrical & Computer Eng., MSEE.
1990~1993년 : University of Missouri, Dept. of Electrical & Computer Eng., Ph.D.
1994~1995년 8월 : 한국전자통신연구원.
1995년 9월~현재 : 한양대학교 전자컴퓨터공학부 조교수

관심분야 : Computer Vision, Pattern Recognition, and all aspects of Computational Intelligence.
E-mail : frhee@fuzzy.hanyang.ac.kr