

다중 필터와 복합형 신경망을 이용한 얼굴 검출 기법

조일국 박현정 김호준

한동대학교 전산전자 공학부

ilgook@gmail.com, hjpark79@cmpal.com, hjkim@handong.edu

Face Detection Using Multiple Filters and Hybrid Neural Networks

Cho, Il-Gook Park, Hyun-Jung Kim, Ho-Joon

School of Computer Science and Electronic Engineering, Handong Global University

요약

본 논문에서는 방송 영상에서 조명효과와 크기변화 등에 강인한 얼굴패턴 검출기법을 제시한다. 제안된 얼굴검출 모델은 영상 전처리 과정과 얼굴패턴 검출 과정으로 이루어진다. 전처리 과정은 조명변화에 대한 보정기능과 다중필터에 의한 후보영역 선별기능으로 구분된다. 얼굴패턴 검출과정은 다단계의 특징지도 생성과정과 패턴분류 과정으로 이루어진다. 특징지도를 생성하기 위하여 가보(Gabor) 필터계층을 포함하는 CNN(Convolutional Neural Networks)모델을 도입하였다. 다양한 배경을 고려한 효과적인 학습을 위하여 본 논문에서는 억제성의 뉴런(Inhibitory neuron)을 포함하는 구조의 CNN모델을 적용한다. CNN으로부터 추출되는 특징집합은 최종 단계에서 WFMM(Weighted Fuzzy Min Max) 모델을 사용하여 분류된다. 이때 사용되는 특징집합의 크기는 분류기의 규모 및 계산량의 결정적인 역할을 준다. 이에 본 연구에서는 최종 분류 과정에 사용되는 특징의 수를 효과적으로 줄이기 위해 FMM모델을 사용하는 적응적인 특징 선별 기법을 제안한다. 또한 실제 영상을 통한 실험결과로부터 제안된 이론의 타당성을 고찰한다.

1. 서론

얼굴 검출 문제는 유비쿼터스 환경에서 인간과 컴퓨터의 대화형 시스템, 보안 시스템, 로봇 시각등의 넓은 분야에서 연구 되고 있다. 이는 시스템이 사용자를 인식 또는 인증하는 과정에서 기존의 키보드, 마우스를 통한 데이터 입력이나, 지문 또는 동공을 통한 데이터 입력에 비해 사람에게 데이터 입력의 불편함을 줄이고 좀 더 친화적인 시스템으로 구현되는데 중요한 역할을 한다.

얼굴 검출에 장애가 되는 다양한 요소들을 제거하고 정확하며 빠른 얼굴 검출을 위한 방법들은 많은 연구들을 통해 발표된 바 있으며 구현된 결과들이 많은 자료로 나타나고 있다[1-7]. 최근의 연구된 바를 보면 대부분의 연구에서 몇 가지 핵심적인 특정 기술을 사용한다[7]. 움직임 추출과 피부색 분석, 신경망 기반의 패턴 인식 모델, 특징 분석 모델 및 기하학적 모델링등을 이용한 방법 등이 제안된 바 있다. 이러한 방법들은 조명 변화, 얼굴 크기 또는 복잡한 배경에서의 얼굴 검출을 위해 복잡한 알고리즘을 적용하였고, 이로 인해 계산의 복잡도가 증가하고, 얼굴 추출 시간이 긴 경우가 대부분이다.

본 논문에서는 실제 영상의 실시간 인식문제를 대상으로 하며, 얼굴패턴 검출을 위한 신경망 기법을 소개한다. 이는 다중 필터와 복합형 신경망이 상호 결합된 형태를 가진다. 검출 과정에서 후보영역의 수를 줄이기 위해 움직임 필터와 색상 필터를 사용하였고, 조명 효과에 강인한 성능을 보이기 위하여 조명 보정을 고려하였다. 얼굴 패턴 검출 과정은 다단계의 특징지도 생성 과정과 패턴분류 과정으로 이루어진다. 특징지도 생성을 위하여 본 논문에서는 CNN을 사용하였고, 첫 단계의 특징지도 생성 단계에서 8방향의 가보 필터를 사용하여 기본 특징들을

생성하게 하였다. 또한 반대 예제에 반응하는 억제성의 뉴런을 고려하여 다양한 배경에서의 효과적인 학습을 할 수 있게 하였다. 패턴 분류 과정에서는 기존의 FMM방법에 가중치 요소를 고려한 WFMM모델을 사용하였다. 이 때 WFMM모델로 입력되는 특징집합의 크기는 신경망의 크기와 계산량의 크기를 결정하는 중요한 역할을 한다. 따라서 본 논문에서는 입력되는 특징의 수를 줄이기 위해 학습데이터에서 특징값의 발생 빈도와 분포를 고려하여 각 특징의 중요도를 평가하는 방법론을 고려 하였다. 본 논문에서는 이들에 대한 이론적 정리와 실제 실내 영상에 적용한 얼굴 검출 실험결과를 고찰한다. 본 논문의 2장에서는 대상 시스템의 구조를 간략히 설명하고, 3장에서는 다중 필터링과 조명 보정에 관하여, 4장에서는 얼굴 검출을 위한 특징 지도 생성 및 얼굴 패턴 분류와 특징의 연관도 요소를 고려한 특징 수 감소 효과 대해 소개한다. 이어서 5장에서는 실제 영상을 통한 실험 결과로부터 제안된 이론의 타당성을 고찰 한다.

2. 대상 시스템 개요

본 연구는 그림. 1과 같은 얼굴 패턴 검출 시스템을 대상으로 한다. 그림에 보인 바와 같이 시스템에서 얼굴 패턴 검출 모듈은 전처리 모듈과 패턴 검출 모듈의 두 단계로 이루어진다. 전처리 모듈에서는 다단계 필터링 과정과 조명 보정과정의 연동되어 수행된다. 이는 얼굴패턴 검출을 위한 대상의 후보 영역의 수를 줄임으로서 시스템의 실시간 응답속도를 개선한다. 이어서 두 번째 단계인 얼굴 패턴 검출 모듈에서는 두 종류의 신경망이 사용되게 되는데 이는 각각 특징 지도 생성 및 얼굴 패턴 분류 과정을 수행하게 된다. 본 연구에서는 특징 지도 생성을

위하여 기존에 제안되었던 CNN모델[1]을 변형하여 최초의 샘플링 계층에 가보 필터(Gabor Filter)를 사용하였다. 또한 패턴 분류 과정에서는 기존에 제안된 FMM 신경망[2]을 수정한 형태의 WFMM 신경망을 사용하였다. 이는 패턴 분류기능 뿐만 아니라, 학습된 신경망으로부터 특징과 하이퍼박스간의 상호 연관도 요소를 분석할 수 있게 한다. 이를 이용하여 특징을 선별적으로 추출함으로써 최종단계에서 추출되는 특징의 개수를 줄이게 한다. 이는 패턴 분류를 위한 신경망의 입력의 크기를 줄여 신경망의 규모 및 계산량을 감소시키게 된다.

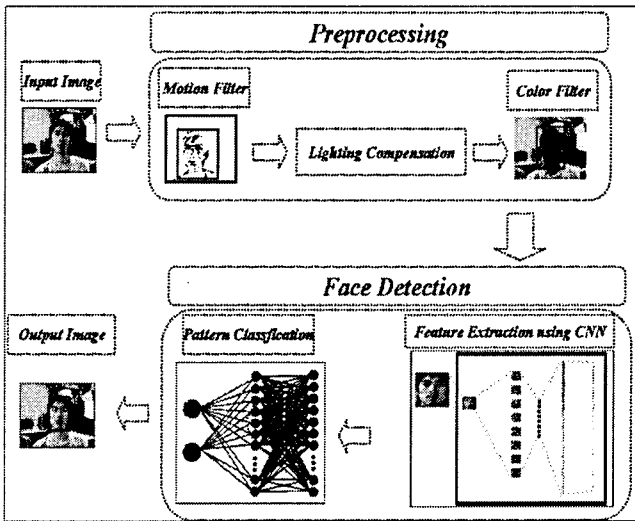


그림. 1 얼굴 검출 시스템 구성도

3. 필터링

입력 영상의 전 영역을 대상으로 하는 얼굴패턴 검출은 많은 연산 시간이 소모되고, 조명 변화에 민감하게 될 뿐만 아니라 복잡한 배경에서의 영역 검출의 정확도를 저하시킨다. 이에 본 논문에서는 얼굴패턴 검출의 전단계로서 다단계 필터와 조명 보정 기법을 제안한다. 이를 통하여 얼굴 후보 영역을 선정하여 상기의 문제점들을 개선하고자 한다.

가. 모션 필터

본 논문에서는 실내 환경에서 카메라로 입력된 연속된 영상을 입력 데이터로 고려한다. 따라서 이를 이용하면 움직임 정보를 이용할 수 있다. 단 배경과 카메라는 고정되어 있다고 가정한다. 영상 안에서의 움직임 정보는 복잡한 배경에서 움직임이 있는 영역과 움직임이 없는 영역으로 구분해 낼 수 있고, 이 정보를 이용하면 후보 영역 선정에 매우 유용한 정보가 될 수 있다. 기존의 연구에서 움직임 정보 검출을 위한 다양한 방법론들이 제안된 바 있다. 그러나 본 연구에서는 시스템의 실시간 응답특성을 고려하여 단순화된 방법을 사용하였다. 사용된 방법은 연속 영상에서의 각 픽셀 별로 색상의 변화를 산출하고 이에 임계치를 적용함으로써 움직임 정보를 생성한다.

나. 조명 보정

영상 내의 피부색은 조명 변화에 매우 민감하게 나타난다. 이에 본 연구에서는 시스템의 색상 정보 사용의 전처리 단계로 조명 보정 과정을 수행한다. 이를 위하여 "Reference White"[3] 기법을 채택하여 사용하였다. 이는 입력되는 전체 영상에서 luma(비 선형적으로 감마 보

정된 밝기 값)값의 범위를 구한 후, 이 값의 상위 일정 비율의 픽셀들을 "Reference White Pixel"로 정의한다. 이에 속한 픽셀 수가 임계치보다 많으면 조명 보정을 수행하게 되는데, 전체 픽셀에 대하여 R, G, B의 각각의 값에 대하여 Y값의 범위를 "Reference White Pixel"의 평균 값으로 나눈 값이 되도록 재조정 한다.

다. 색상 필터

본 논문에서는 실내 환경에서 획득한 다량의 데이터로부터 피부색의 특성을 분석하여 이를 색상 필터로 사용한다. 그림. 2는 실제 환경에서 얻은 영상을 대상으로 분석한 피부색의 특징분석 결과의 예이다. 이의 분석 결과를 통하여 R, G, B, 명암, Cb, Cr 등 총 9가지의 색상 영역에서의 피부색의 범위를 알아 낼 수 있다. 이 피부색의 범위는 어떤 픽셀이 가지고 있는 색상 값들이 피부색에 속하는지 속하지 않는지를 판단하게 하는 근거로 사용하게 되며, 피부색의 최종 판단은 각각의 색상 영역에 대한 퍼지 멤버십 함수의 결과를 이용한다[5].

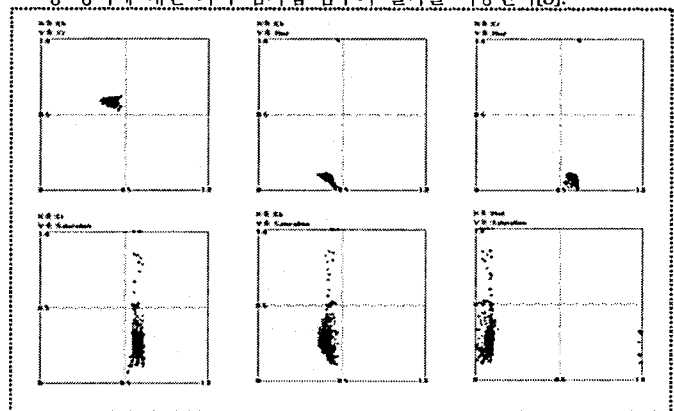


그림.2 피부색에 대한 특징 분석 결과의 예

이 때 사용되는 퍼지 멤버십 함수는 다음과 같이 정의 된다.

$$X(x : a, b, c, d) = \left\{ \begin{array}{ll} 0, & \text{if } x \leq a \text{ or } x \geq d, \\ \frac{2(x-a)^2}{(b-a)^2}, & \text{if } a < x \leq \frac{(a+b)}{2}, \\ 1 - \frac{2(x-b)^2}{(b-c)^2}, & \text{if } \frac{(a+b)}{2} < x \leq b \\ 1, & \text{if } b < x < c \\ 1 - \frac{2(x-c)^2}{(d-c)^2}, & \text{if } c \leq x < \frac{(c+d)}{2} \\ \frac{2(x-d)^2}{(d-c)^2}, & \text{if } \frac{(c+d)}{2} \leq x < d \end{array} \right. \quad (1)$$

식에서 파라미터 a, b, c 와 d는 $a < b < c < d$ 의 조건을 만족하며, 퍼지 멤버십 함수가 디자인된 의미에 맞게 조정할 수 있다. 파라미터 a와 d는 각각 특정 색상 영역에서 피부색 범위의 가장 자리를 의미한다. x값이 b와 c사이에 있으면 X(x)는 1의 결과를 나타내지만, a와 b 또는 c와 d사이에 있으면 피부색 영역이지만 상대적으로 중요도가 낮은 정보로 판단 할 수 있다. 그림. 3은 위의 퍼지 멤버십 함수를 그래프로 표현 한 것이다. 이와 같은 분석을 통하여, 적용할 특징 요소를 적용적으로 선별하여 색상 필터를 구현하게 된다.

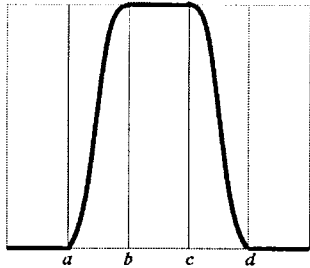


그림. 3 색상 필터에 사용한 퍼지 멤버십 함수

4. 얼굴 검출 신경망

입력된 영상에서 필터링을 거쳐 결정된 후보 영역은 CNN을 이용하여 특징 추출 및 특징 지도 생성 작업을 수행한 후 WFMM 신경망을 이용하여 얼굴 영역을 추출한다.

가. 특징 추출 및 특징 지도 생성

CNN은 다양한 형태로 필기체 인식, 얼굴 인식 등의 여러 복잡한 문제들에 적용하여 좋은 성능을 보인 바 있다[1]. CNN은 생물학적 형태를 반영한 계층적 다중 신경망으로써 목표물의 위치 이동, 크기 변화 및 왜곡에 강한 인식을 제공한다. 본 논문에서 제안하는 얼굴 검출을 위한 신경망은 그림. 4와 같은 구조를 가진다.

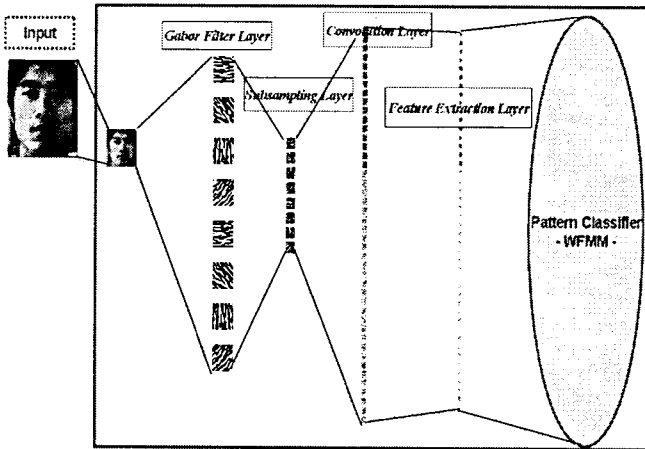


그림. 4 얼굴 검출 신경망의 구조

첫 번째 계층인 가보 필터 계층에서는 입력영상으로부터 8방향의 기본 특징을 추출한다. 연속적으로 이어지는 계층에서 확장 및 조합된 특징들을 점진적으로 추출하며 이 과정에서 뉴런간의 지역적 연결 구조와 수용영역(receptive field)의 형태로 샘플링 되어 영상에서 특징의 크기변화, 위치이동 및 회전등의 왜곡을 보완할 수 있게 한다. 상위 계층으로 갈수록 특징지도의 크기와 해상도는 작아지지만 추출되는 특징의 종류는 많아지게 된다. 본 연구에서는 대상영역의 기본 크기를 (38 X 42)로 설정 하였으며, 이 영상은 가보 필터를 거쳐 (28 X 32)의 특징지도로 생성된다. 이 특징지도는 두 번째 계층에서 (2 X 2)영역으로 표본 추출하여 (14 X 16)크기로 조정된다. 상기의 방법으로 생성된 8개의 특징 영역 중 2개의 서로 다른 조합을 선택하여 44개의 특징 지도를 생성하고, 최종적으로 (6 X 7)크기의 최종 적인 특징 지도를 생성 하게 된다. 이와 같이 생성된 특징들은 얼굴 패턴의 자극성만을 고려한 특징들로 이 특징들만 사용할 경우 반대 예제들에 의한 특징들로 검출

율이 낮아 질 수 있다. 따라서 본 연구에서는 식 (2)와 같은 방법의 억제성 뉴런을 사용하여 특징의 자극성과 억제성 작용을 모두 고려하도록 하였다.

$$z = \frac{g(\sum_{j=1}^{S_R} C_j I_j + b)}{a + f(\sum_{j=1}^{S_R} D_j I_j + d)} \quad (2)$$

위 식에서 S_R 은 입력되는 특징의 크기이고, I_j 는 j번째 입력 값, C_j 와 D_j 는 각 자극성과 억제성 입력에 대한 가중치, b와 d는 상수 바이어스 값이다.

나. 얼굴 패턴 검출 모듈

본 논문의 얼굴 검출 시스템에서 패턴 분류모듈에 사용한 신경망 모델은 가중치를 갖는 FMM 신경망이다[8]. 이는 다음 식으로 표현되는 하이퍼박스 멤버십 함수로서 그 활성화 특성이 정의 된다.

$$b_j(A_h) = \frac{1}{\sum_{i=1}^n w_{ji}} \cdot \sum_{i=1}^n w_{ji} [\max(0, 1 - \max(0, r_{ji} \min(1, a_{hi} - v_{ji}))) + \max(0, 1 - \max(0, 1 - \max(0, r_{ji} \min(1, u_{ji} - a_{hi})))) - 1.0] \quad (3)$$

식에서 $A_h = (a_{h1}, a_{h2}, \dots, a_{hm})$ 은 입력패턴으로 총 n 개의 특징 값이고, 각 특징에 대한 최소점과 최대점은 $U_j = (u_{j1}, u_{j2}, \dots, u_{jm})$, $V_j = (v_{j1}, v_{j2}, \dots, v_{jm})$ 로 표현된다. w_{ji} 는 j번째 하이퍼박스과 i번째 특징 사이의 연결 가중치를 의미하며 r 는 특징범위의 가장자리에서 퍼지 멤버십 함수의 기울기를 결정하는 매개변수로 [0,1]사이의 값을 갖는다.

다. 특징의 연관도 요소를 고려한 특징 수 감축

본 논문에서 제안된 CNN은 총 (6 X 7)크기의 특징지도 44개를 생성한다. 이들 특징지도의 개별 특징값이 모두 분류 신경망으로 입력 되려면 분류 신경망의 입력 노드수는 총 1,848개가 필요하다. 이는 신경망의 규모가 커지고 계산량이 많아질 뿐만 아니라, 다량의 불필요한 특징들을 포함하여 분류 성능을 저하 시킬 가능성이 있다. 이에 본 논문에서는 앞에서 살펴보았던 WFMM 신경망을 이용하여 특징의 중요도를 상대적으로 평가하여, 이 값을 기준으로 최적의 특징집합을 생성하는 방법을 사용하였다[8]. 학습된 신경망으로부터 특징과 클래스와의 관계를 분석하기 위하여 다음과 같은 4종류의 연관도 요소(relevance factor)를 정의한다.

- $RF1(x_i, B_j)$: 특징값 x_i 와 하이퍼박스 j 와의 연관도 요소
- $RF2(x_i, C_k)$: 특징값 x_i 와 클래스 k 와의 연관도 요소
- $RF3(X_i, C_k)$: 임의의 특징 타입과 클래스간의 연관도 요소
- $RF4(X_j)$: 주어진 패턴 분류 문제에서 임의의 특징에 대한 상대적 중요도

WFMM 신경망의 활성화 함수 및 학습기법의 특성으로부터 상기의 4

가지 척도를 다음과 같은 식으로 정의 할 수 있다. 특징값과 하이퍼 박스의 관계는 식 (4)과 같이 그들 상호간의 가중치 값으로 표현되며, 특징값과 클래스의 관계는 식 (5)와 같이 이들 가중치로부터 산출 할 수 있다.

$$RF1(x_i, B_j) = w_{ij} \quad (4)$$

$$RF2(x_i, C_k) = \left(\frac{1}{N} \sum_{B_j \in C_k} S(x_i, (u_{ji}, v_{ji})) \right) \square w_{ij} - \frac{1}{(N_B - N_k)} \sum_{B_j \in C_k} S(x_i, (u_{ji}, v_{ji})) \square w_{ij} / \sum_{B_j \in C_k} w_{ij} \quad (5)$$

이로부터 식 (6)와 같이 특징 유형과 클래스의 관계를 정의 할 수 있으며, 궁극적으로 주어진 패턴 분류 문제에서 특징의 중요도는 식(7)의 연관도 요소로 평가 할 수 있다.

$$RF3(X_i, C_k) = \frac{1}{L_i} \sum_{x_i \in X_i} RF2(x_i, C_k) \quad (6)$$

$$RF4(X_i) = \frac{1}{M} \sum_{j=1}^M RF3(X_i, C_j) \quad (7)$$

5. 실험 결과 및 고찰

본 연구는 방송 영상에서 실시간 얼굴 검출 시스템을 대상으로 한다. 그림. 5는 제안된 이론에 적용한 시스템의 출력의 예이다. 그림. 6은 특징 수 감소에 따른 성능 변화를 보인 그래프인데, 그림에 보인바와 같이 실내 영상에서 얼굴 패턴 추출기능의 성능을 유지하면서 입력 특징의 수를 최초 1848개에서 약 900개 까지 감소 시킬 수 있었다.

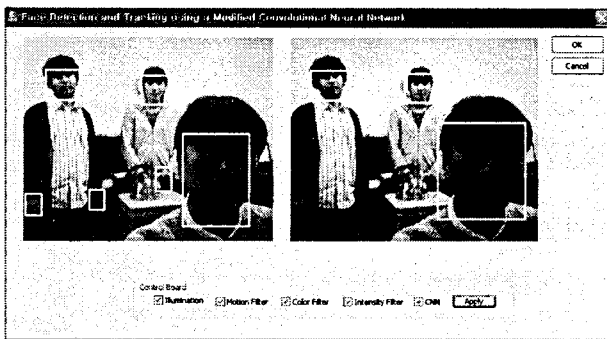


그림. 5 얼굴 검출 시스템 출력의 예

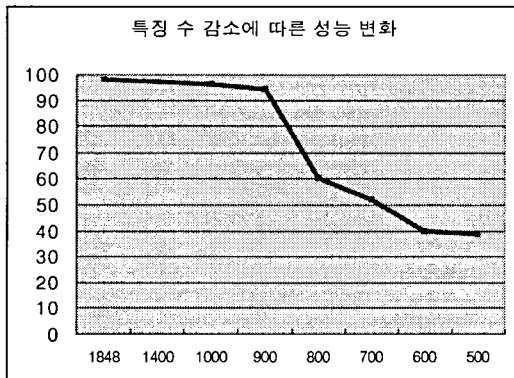


그림. 6 특징 수 감소에 따른 변화 그래프

6. 결론

본 논문에서는 얼굴 검출을 위한 방법론으로서 다단계 필터와 복합형 신경망을 이용한 모델을 제시하였다. 이는 방송 영상에서의 실시간 인식을 목표로 한다. 전처리 단계에서 조명 보정과 필터링을 통하여 후보 영역의 수를 줄임으로써 시스템의 실시간 응답 특성을 개선하였으며 시스템이 보다 강인한 성능을 보일 수 있게 하였다. 얼굴 검출 모듈에서 특징 추출을 위한 CNN은 가보 필터를 사용하여 보다 미세한 특징들을 추출하게 하였으며 반대 예제의 효과에 대해 억제성의 뉴런을 고려하여 최종적인 특징지도를 생성하게 하였다. 패턴 분류 신경망으로 사용된 WFMM은 특징집합에서 특징 값의 빈도와 분포에 따른 변화를 효과적으로 반영하여 비 정상 데이터 및 노이즈에 의한 학습 효과의 저하를 개선 할 수 있게 하였고, 또한 특징과 하이퍼박스 및 클래스 간에 연관도 요소를 평가하여 특징의 중요도가 상대적으로 높은 것들을 선별하게 하였다. 이는 패턴 분류 신경망에 입력되는 특징의 수를 감소 시켜 인식 시스템의 크기 및 계산량을 줄일 수 있게 하였다.

* 이 연구는 산업자원부 뇌과학 연구개발사업으로 수행되었음

참고문헌

- [1] Christophe Garcia, Manolis Delakis, "Convolutional Face Finder: A Neural Architecture for Fast and Robust Face Detection," IEEE Transactions on Neural Networks, Vol. 26, No. 11 2004
- [2] P. Simpson, "Fuzzy Min-Max Neural Networks Part-1: Classification," IEEE Transactions on Neural Networks, Vol.3, No.5, 1992
- [3] Rein-Lien Hsu, Mohamed Abdel-Mottaleb and Anil K. Jain, "Face Detection in Color Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, No.5, 2002
- [4] Fok Hing Chi Tivivie and Abdesselam Bouzerdoum, "Efficient Training Algorithm for a Class of Shunting Inhibitory Convolutional Neural Networks," IEEE Transactions on Neural Networks, Vol. 16, No. 3, 2005
- [5] Sanun Srisuk and Werasak Kurutach, "A New Face Detection In Color Images," Proceedings of the Fifth IEEE Conference on Automatic Face and Gesture Recognition, 2002
- [6] Raphael Feraud, Oliver J. Bernier, Jean-Emmanuel, and Michel Collobert, "A Fast and Accurate Face Detector Based on Neural Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, No.1, 2001
- [7] Ming-Hsuan Yang, Daivid J. Kreigman, Narendra Ahuja, "Detecting Faces in Images: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, No. 1, 2002
- [8] Ho-Joon Kim, Hyun-Jung Park, Yun-Seok Cho, "A Feature Saliency Measure in WFMM Neural Network-Based Pattern Classification," The 2005 International Conference on Artificial Intelligence, 2005