

## 얼굴 인식 기술 및 동향

송근원\*

### 1. 서론

최근 얼굴, 지문 및 홍채 등의 人體를 이용하여 개인의 인증, 정보 보호 및 신분 확인을 수행할 수 있는 생체 인식 기술들에 대한 연구가 활발히 진행되어, 현재 상용화 제품들이 판매되고 있다. 21C 정보화 사회는 디지털화, 인터넷, 멀티미디어 및 정보 통신의 발달로 인하여 다양한 일들을 보다 쾌적하고, 편리하고 또한 신속하게 처리할 수 있는 환경으로 급변하고 있다. 패턴 인식, 컴퓨터 비전 및 신경 회로망 기술 등을 이용한 생체 인식 기술은 친화성, 편리성, 다양성 및 쾌적성을 고려한 얼굴 로봇, 재활 공학 시스템, 가상 현실, 수화 인식, 멀티미디어 시스템 및 게임 등 다양한 산업 분야에 응용 가능하다. 이러한 생체 인식 기술의 상업적 활성화를 위하여 현재 전세계적으로 생체 인식 기술의 객관적 성능 평가에 많은 노력을 기울이고 있다. 국내에서는 2001년부터 생체 인식 기술에 대한 관심이 높아져 한국 정보 보호원에서는 2001년 2월 생체 인식 협의회의 표준 분과에서 국내 표준 제정 및 국외 표준화 활동에 대한 공동 대응 등의 역할을 수행하고 있다.

생체 인식 기술중 얼굴 인식 기술은 사용자에 게 특별한 행위를 요구할 필요가 없으며, 또한 비

접촉식으로 인식을 수행할 수 있어 편리한 장점이 있다. 현재까지 얼굴 인식 기술은 국내·외의 관련 업체와 대학을 중심으로 활발한 연구가 수행되어 왔다. 개발된 제품으로는 Identix([www.visionics.com](http://www.visionics.com))사의 FaceIt, Viisage([www.viisage.com](http://www.viisage.com))사의 VIS(Viisage's Visual Inspection System), (주)비전 인터랙티브([www.vi21.co.kr](http://www.vi21.co.kr))의 Vision Gate 등, (주)워치비전([www.watchvision.com](http://www.watchvision.com)) 및 (주)블루닉스([www.bluenics.com](http://www.bluenics.com))사의 얼굴 인식 관련 제품 등이 있다. 국내·외 대학 및 연구소에서도 얼굴 인식에 관하여 활발한 연구를 수행해 왔다. 그리고 얼굴 및 감성 인식 관련 전문 국제 학술 대회(International Conferences on Automatic Face and Gesture Recognition, International Conference on Audio and Video based Person Authentication)가 개최되어 개발한 기법들의 인식 성능에 관한 논의가 활발히 진행되고 있다. 얼굴 인식 분야는 그림 1에서와 같이 향후 많은 응용 분야에 적용될 수 있다. 또한 HCI(Human and Computer Interaction) 및 감성 공학 기술과 접목 되면 쾌적성, 사용자 편의성 및 신속성이 요구되는 분야에 널리 이용될 것으로 기대된다.

얼굴 인식에서는 주로 정지 영상 및 동영상을 이용하였다. 그리고 얼굴 인식 과정은 크게 얼굴 영역 추출, 특징 추출 및 매칭을 통한 인식 과정으

\* 위덕대학교 컴퓨터멀티미디어 공학부

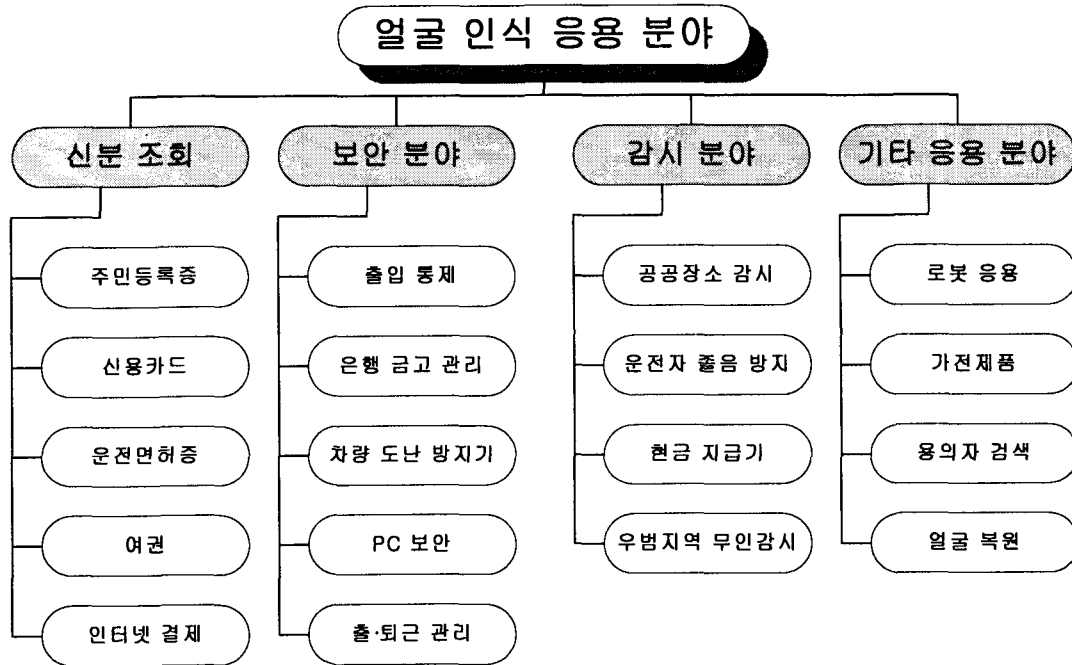


그림 1. 얼굴 인식 응용 분야

로 구성된다. 얼굴 영역이 정확히 추출되어야 높은 인식률의 얼굴 인식 시스템을 기대할 수 있으므로, 얼굴 영역 추출 과정은 얼굴 인식 기술에 있어서 기본적인 또한 중요한 과정이다. 이에 최근 정확한 얼굴 영역 추출에 관한 연구가 활발히 진행되고 있다[1-6].

본 고에서는 현재까지 연구된 얼굴 인식 기술에 대해 소개하고, 얼굴 인식 분야에서 해결해야 할 문제점에 대해 논하고자 한다.

## 2. 얼굴 인식 기술

얼굴 인식 과정은 크게 얼굴 영역 추출, 특징 추출 및 매칭을 통한 인식 과정으로 구분할 수 있다. 본 장에서는 이에 관한 기술 현황 및 동향에 대하여 알아본다.

### 2.1 얼굴 영역 추출

얼굴 인식에서 영역 추출 과정은 매우 중요하

여 현재까지 많은 연구가 수행되어 왔다. 입력 영상은 배경과 얼굴 영역 경계 부분에 존재하는 화소들의 밝기 값의 유사성, 복잡한 배경, 각 개인의 화장법 차이, 조명으로 인한 얼굴의 색상 변화, 잡음 및 주위 환경 등으로 인한 매우 다양한 특징을 나타내므로, 얼굴 영역을 정확히 추출하는 문제는 매우 어렵다. 얼굴 영역을 추출하는 기술은 크게 밝기 정보를 이용하는 방법, 움직임 정보를 이용하는 방법, 밝기 및 움직임 정보를 동시에 이용하는 방법 및 색상 정보를 이용하는 방법으로 구분된다.

얼굴 영역을 추출하는 문제는 얼굴 인식 뿐 아니라 과거 영역 기반 영상 압축 분야의 객체 분할 과정에서 많이 연구가 되어 왔으며, 근본적으로 사람이 포함된 입력 영상일 경우 얼굴 영역을 추출하는 시도는 동일하다고 볼 수 있다. 이에 밝기 정보를 이용한 기법, 움직임 정보를 이용한 기법, 밝기 및 움직임을 동시에 이용한 기법 및 칼라

정보를 이용한 얼굴 영역 추출 기법에 대해서 알아본다.

**밝기 정보를 이용한 기법**

밝기 정보를 이용하여 얼굴 영역을 추출하는 기법은 입력 영상에서 배경 영역과 얼굴 영역의 밝기가 다르다고 가정하고 영상 분할을 통하여 얼굴과 배경을 분할함으로써 얼굴 영역을 추출한다. 배경이 단순하여 얼굴 영역과 배경 영역의 경계 부근 화소들의 밝기 값의 차가 큰 경우에는 얼굴 영역을 정확히 추출할 수 있었다. 아래 그림 2에서는 이에 대한 예를 보여준다. 그러나 경계 부근 화소들의 밝기 값이 유사할 경우 또는 복잡한 배경에서는 얼굴 영역과 배경 영역이 같은 영역으로 추출될 수 있어 얼굴 영역을 정확히 추출할 수 없는 단점이 있다.

그밖에도 입력 영상에서 에지 정보를 추출한 후 이를 이용하여 얼굴 영역을 추출하는 기법들이 연구되었다. 그러나 이 기법은 밝기 정보를 이용하므로 근본적으로 위에서 언급한 밝기 정보의 한계를 극복하지 못하는 단점이 있다.

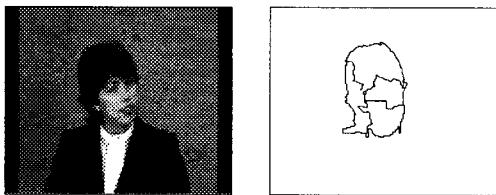


그림 2. 밝기 정보를 이용하여 얼굴 영역을 추출한 예

**움직임 정보를 이용한 기법**

얼굴 영역과 배경 영역의 경계 부근 화소들의 밝기 값이 유사하더라도 얼굴 영역에 움직임이 존재한다면 인접 프레임간의 움직임 정보를 이용하여 얼굴 영역을 정확히 추출할 수 있다. 움직임을 이용한 얼굴 영역 추출 기법은 크게 인접

프레임간의 차영상을 이용하는 경우, 움직임 추정을 통한 움직임 벡터(motion vector)를 이용하는 경우로 구분된다. 움직임 정보를 이용하는 기법들은 동영상에서 사람의 얼굴은 움직임이 있는 객체(object)라고 가정하고 인접 프레임간의 차 또는 움직임 추정을 통해 유사한 움직임 벡터 값을 나타내는 화소들을 그룹화함으로써 얼굴 영역을 추출하였다.

그림 3은 인접한 두 프레임간의 프레임 차를 이용하여 얼굴 영역을 추출한 예이다. 배경 영역이 단순하고 움직임이 없을 경우 얼굴 영역을 비교적 정확히 추출할 수 있음을 나타낸다.



그림 3. 인접 프레임간의 프레임 차를 이용하여 얼굴 영역을 추출한 예

그림 4는 배경이 단순한 영상에서 인접 프레임간의 움직임 추정을 수행한 예를 보인다. 움직임 추정을 통한 얼굴 영역 추출 기법들은 얼굴 영역의 움직임이 분명하고, 배경 영역의 움직임이 없을 경우 얼굴 영역을 정확히 추출할 수 있다. 그러나 배경 영역에서의 움직임 추정 정보가 얼굴 영역에서의 움직임 추정 정보와 유사할 경우 움직임 정보만을 이용하여 얼굴 영역을 정확히 추출할 수 없다. 그리고 일반적으로 현재까지 개발된 다양한 움직임 추정 기법들은 배경이 복잡하거나, 얼굴 영역과 배경 영역의 경계 부분에서는 정확한 움직임 정보를 추정할 수 없으므로 부정확한 얼굴 추출의 원인이 되기도 하였다. 또한 움직임을 이용할 경우 카메라의 모션 정보에 대한 고려가 필요하다.



그림 4. 인접 프레임간의 움직임 추정 예

**밝기 및 움직임 정보를 이용한 기법**

밝기 및 움직임 정보의 장점만을 결합한 형태의 분할 척도(segmentation criterion)를 이용하여 얼굴 영역을 추출하는 기법들이 연구되었다. 밝기와 움직임 정보를 동시에 고려할 경우 단순한 배경에서 분할된 많은 영역들을 효과적으로 제거하면서 얼굴 영역을 추출할 수 있는 장점이 있다. 그러나 얼굴 영역 추출시 사용되는 밝기 정보와 움직임 정보의 비(ratio)에 민감하므로 이들간의 분할 척도 비를 조절하는 어려움이 있다.

**색상 정보를 이용한 기법**

피부색은 얼굴 영역 추출시 유용한 정보가 될 수 있어 널리 이용되어 왔다. 이는 크게 HSI, YIQ 및 YCbCr 등과 같은 칼라 모델에 기반하여 피부색 구간을 설정하고 피부색 화소들을 추출함으로써 얼굴 영역을 추출하는 기법들과, 피부 영역의 색분포를 모델링 한 후 입력 화소들에 대한 피부색 및 비피부색 화소 여부를 결정하여 얼굴 영역을 추출하는 기법들로 구분된다.

피부색 구간 설정을 통한 기법으로는 HSI 및 YCbCr 칼라 모델에 기반한 기법들이 많이 연구되었다. HSI 칼라 모델에 기반한 기법은 색상(Hue) 정보에서 피부색 구간 범위를 정한 후 입력 영상 화소의 색상 값이 피부색 구간 범위에 속하면 얼굴 화소로 결정한다. YCbCr 칼라 모델을 이용할 경우 Cb 및 Cr의 피부색 구간 범위를 각각 정한 후 이에 속하는 화소들을 추출함으로써 얼굴 영역을 추출하였다. 배경 영역의 색상이 단순하고

또한 이의 색상값이 피부색 구간 범위내의 값과 차이를 나타내는 경우 얼굴 영역을 정확히 추출할 수 있다.

그림 5는 배경의 색상이 단순하고 배경 영역에 피부색 화소가 존재하지 않을 경우 색상 정보를 이용한 기법들은 비교적 얼굴 영역을 정확히 추출할 수 있음을 보여준다. 그러나 배경이 복잡하거나 배경 영역에 피부색과 유사한 색상 정보를 나타내는 화소들이 존재할 경우 얼굴 영역을 정확히 추출할 수 없다. 또한 입력 영상의 피부색은 주위 조명에 의한 영향, 각 개인의 피부색 및 화장 기법 차이 등으로 인하여 일률적인 피부색 구간 범위를 적용하면 다양한 환경 조건에서 획득된 입력 영상에서 얼굴 영역을 정확히 추출할 수 없는 단점이 있다.



그림 5. 피부색 구간 설정을 통한 얼굴 영역 추출 예(단순한 배경)

그림 6은 배경 영역에 피부색과 유사한 색상을 나타내는 화소들이 존재하고, 배경 영역이 복잡한 경우 정확한 얼굴 영역을 추출하기 어려움을 보여준다.



그림 6. 피부색 구간 설정을 통한 얼굴 영역 추출 예(복잡하고 피부색과 유사한 배경)

그림 7에서는 조명 및 화장으로 인하여 얼굴 영역의 일부가 피부색으로 추출되지 않음을 보여 준다. 이는 피부색을 이용하는 기법들은 주위 조명 변화, 각 개인의 화장 기법 차이 및 배경 영역에서 유사 피부색을 나타내는 화소들이 존재할 경우 정확한 얼굴 영역 추출이 어려움을 나타낸다. 이를 해결하기 위해서는 입력 영상에서 먼저 피부색 분포도를 분석하고 이에 기반하여 피부색 구간 범위를 설정하는 기법에 대한 연구가 필요하다.

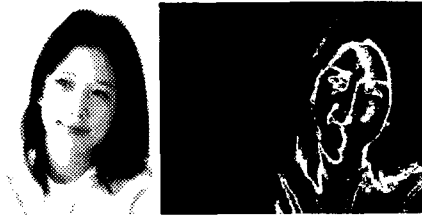


그림 7. 피부색 구간 설정을 통한 얼굴 영역 추출 예(화장 및 조명의 영향)

한편 피부 영역의 색분포를 3차원의 정규 분포로 가정하여 구한 확률 밀도 함수의 평균값과 각 화소 색조와의 차이를 임계값과 비교함으로써 각 화소를 피부 영역 화소와 비피부 영역 화소로 분류하여 얼굴 영역을 추출하는 연구가 수행되었다. 이때 피부색의 분포는 아래의 정규 분포로 모델링 된다.

$$p(x) = \frac{\|C^{-1}\|^{1/2}}{2\pi^{3/2}} \exp[-(x-m)^T C^{-1}(x-m)/2]$$

$$m = (m_Y, m_I, m_Q)^T$$

$x$ 는 각 화소의  $YIQ$  성분 벡터,  $m$ 은  $YIQ$  성분의 평균값을 나타내는 벡터이고,  $C$ 는 공분산(covariance) 행렬이다. 평균값  $m$ 과 공분산 행렬  $C$ 는 샘플 영상들의 피부 영역 화소들로부터 구할 수 있다.

$$m = \frac{1}{n} \sum_{i=1}^n x_i$$

$$C = \frac{1}{n} \sum_{i=1}^n (x_i - m)(x_i - m)^T$$

피부색 정규 분포의 평균  $m$ 과 각 화소값  $x_i$ 간의 Mahalanobis distance는 다음과 같이 정의된다.

$$D = (x_i - m)^T C^{-1}(x_i - m)$$

각 화소의 Mahalanobis distance가 임계값 이하이면 피부 영역의 화소로 간주되고 임계값보다 큰 경우는 비피부 영역 화소로 분류된다. 이 방법은 입력 영상에 대하여 피부색을 통계적으로 분석하고 피부색 화소를 추출하였으나 다양한 환경을 고려한 피부색 화소 추출에 대한 고려가 미흡하였다.

이밖에도 신경망을 이용한 기법 등이 연구되었다. 그러나 위에서 언급한 다양한 환경에서 얼굴 영역을 정확히 추출하기 위해서는 앞으로도 많은 연구가 진행되어야 할 것으로 판단된다.

## 2.2 특징 추출

특징 추출에 앞서 카메라의 줌인 및 줌아웃으로 인한 얼굴 크기 변화, 회전 변화 및 주위 환경의 영향으로 인한 밝기 변화 등을 고려하기 위하여 정규화 과정을 수행한다. 이는 보통 패턴 인식의 정규화 과정에서 널리 이용되는 어파인(affine) 변환을 사용하여 이동, 크기 및 회전 등을 정규화한다. 그리고 히스토그램 평활화를 이용하여 밝기 정규화를 수행한다.

얼굴 영역 추출 후 각 개인을 구분하는 척도가 되는 특징 추출에 관한 연구는 KL(Karhunen-Loeve) 변환을 이용하여 고유 벡터를 추출하고 이를 특징으로 하는 기법, 영상에서 방향 정보와

기울기의 크기 정보를 추출해 낼 수 있는 허프(Hough)변환을 이용하는 기법, 눈, 코 및 입간의 거리를 특징으로 이용하는 기법, 주성분 분석(PCA, Principal Component Analysis)을 이용하는 기법 및 선형 변별 해석(LDA, Linear Discriminant Analysis)등 많은 연구가 수행되어 왔다. 이중 눈, 코 및 입의 거리 정보를 이용하는 기법은 먼저 눈, 코 및 입의 정확한 추출이 필수적이다. 그러나 실제 환경에서 발생하는 안경 착용 및 머리카락의 눈가림 등의 현상이 발생할 때는 눈, 코 및 입의 정확한 추출이 어려워 대부분 상용화되는 제품에서는 PCA 및 LDA가 사용되고 있다.

Pentland 등[7]은 주성분 분석으로 물체간에 상관 관계가 없는 특징들을 추출하여 최 근접 이웃(nearest neighbor) 알고리즘으로 분류하는 방법을 제안하고 이를 얼굴 인식에 적용하였다. 주성분 분석은 입력 벡터를 공분산의 고유치와 고유벡터로 정의되는 부공간(subspace)상에 수직 투영시킨다. 이때 입력 자료의 투영 값을 나타내는 주성분들은 평균 자승 오차 측면에서 기존의 방법보다 입력 자료를 잘 표현한다.

Swet 등[8]은 주성분의 변별 성능을 높이기 위하여 주성분을 분석한 후 선형 변별하는 해석 방법을 제안하고 이를 적용하였다. 한편 Chatterjee 등[9]은 Swet 등의 선형 변별 해석 방법을 개선하여 학습 과정에서 클래스 평균을 추정하여 주성분 분석과 선형 변별 해석을 함께 할 수 있는 방안을 제시하였다. 이들 방법에는 클래스내 산포도(within-class scatter)에 대한 클래스간 산포도(between-class scatter)의 비를 변별 기준으로 이용하기 때문에 단봉형태(unimodal)로 분포된 자료 또는 여러 개의 가우스 분포를 갖는 군집으로 구성된 자료들에 대해서 변별력이 우수하다고 알려져 있다. 일반적으로 주성분 분석으로 추출한 특징으로 인식 대상 물체들을 비슷한 통계 특성을

가지는 물체끼리 개략적으로 군집화하여 분류한다. 그리고 분류된 각 소규모 군집내에서 독립성분 분석을 통해 상세히 분류하는 계층적 분류 인식 방법이 우수한 성능을 나타낸다고 알려져 있다.

### 2.3 인식 및 성능 평가

추출한 특징을 이용하여 매칭을 통해 인식하는 과정에서는 신경망(neural network), 최근접 분류기(nearest neighbor classifier) 및 SVM(support vector machines)[10] 등이 이용되어 왔다. 신경망 및 최 근접 분류기는 일반적으로 인식할 대상수가 적고, 클래스간 특징들의 유사도가 적을 경우에는 잘 분류되나 인식할 대상 수가 많아지면 클래스간 특징들의 유사도가 증가하여 인식에 상당한 어려움이 따른다. SVM은 선형 분리 가능한 샘플들에 대하여 클래스간의 거리가 최대가 되도록 학습함으로써 비교적 우수한 성능을 나타낸다고 알려져 있다. 그러나 아직까지 얼굴 영상의 복잡성 및 다양성 등의 특징을 효과적으로 고려할 수 있는 분류기에 대한 연구는 미흡하여 이에 대한 연구·개발이 요구된다. 인식은 인간이 가장 잘하므로 인간이 얼굴을 인식하는 체계를 모사하는 방식의 분류기 개발에 관한 연구가 수행되어야 할 것이다.

개발된 인식 기술에 대한 올바른 성능 평가를 위해서는 다양한 환경에서 촬영된 많은 영상들에 대하여 실험을 수행하여야 한다. 즉 표정을 달리한 영상, 카메라 및 조명을 달리한 영상 및 크기 변화된 영상들도 실험에서 고려하여야 한다. 개발한 알고리즘의 성능 평가에서는 미국의 여러 대학에서 촬영한 영상, 인터넷에서 획득된 영상 및 개발자가 임의로 촬영한 영상을 이용하기도 하지만, 대부분 얼굴 인식 FERET 평가 방법론[11]에 따

른 절차와 실험 영상을 사용하여 실험한다. FERET에는 데이터 베이스, 식별 평가 방법론, 검증 평가 방법론을 포함하고 있으며, 카메라의 줌인 및 줌아웃 영향, 조명 변화, 안경 착용 및 시간 간격 촬영 등의 다양한 조건을 고려하였다. 또한 많은 영상 데이터 베이스 구축이 되어 얼굴 인식 알고리즘의 성능 평가를 객관적으로 수행할 수 있어 최근 얼굴 인식 알고리즘의 성능 평가의 표준이 되고 있다.

### 3. 향후 연구 방향

현재까지 얼굴 영역 추출, 특징 추출 및 인식 과정을 통한 얼굴 인식 기술에 관한 연구가 활발히 수행되어 왔다. 그러나 개발된 얼굴 인식 기법들은 주위 조명 및 다양한 환경 등에 적응적으로 대처하지 못하여 여전히 해결해야 할 많은 문제점이 있다.

향후 연구 방향은 얼굴 영역 추출, 특징 추출 및 인식의 2차원적 얼굴 인식 과정에 기반한 기법 및 3차원 얼굴 모델에 기반하여 다양한 주위 환경에 우수한 인식 성능을 나타낼 수 있는 얼굴 인식 기술에 관한 연구가 필요하여 활발히 수행될 것으로 기대된다.

### 4. 결론

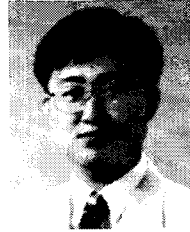
본 고에서는 얼굴 인식 기법들, 객관적 평가를 위한 실험 영상 및 방법에 관한 광범위한 내용을 기술하였다. 얼굴 인식 분야는 물체 인식 등의 다른 패턴 인식 분야보다는 인식 대상의 유사성이 크다. 그리고 복잡한 배경, 각 개인의 화장법의 차이, 조명으로 인한 얼굴의 색상 변화, 잡음 및 주위 환경 등의 영향으로 인식에 더욱 더 많은 어려움을 나타낸다. 이러한 주위 환경 요인은 인식률의

저하를 초래하므로 향후 얼굴 인식 분야는 위에서 언급한 다양한 주위 환경에 강건한 인식 알고리즘 개발이 필요하다. 이를 위해 인간이 얼굴을 인식하는 체계를 모사하는 기법을 고려한 얼굴 인식에 관한 활발한 연구가 필요하다고 사료된다.

### 참고 문헌

- [1] Kin Choong Yow and Roberto Cipolla, "Feature-based human face detection," *Image and Vision Computing*, vol.15, pp.713-735, 1997.
- [2] Froba,B. and Kublbeck,C., "Orientation template matching for face localization in complex visual scenes," *Proc. Int'I conf. Image Processing*, vol.2, pp.251-254, Jan. 2001.
- [3] V.Starovoitov and D.Samal, "Matching of faces in camera images and document photographs," *Proc. IEEE Int'I Conf. Acoustics, Speech, and Signal Processing*, vol. 4, pp.2349-2352, June 2000.
- [4] H.A. Rowley, S.Baluja, and T.Kanade, "Neural network based face detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 20, 1998.
- [5] K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *A. I. Memo 1521, MIT A.I. Lab.*, 1994.
- [6] K. Sobottka and I. Pitas, "Extraction of facial region and features using color and shape information," *In Proc, 10th Int'l Conf., Pattern Recog.*, pp.421-425, Vienna, 1996.
- [7] A. Pentland, B. Moghaddam, and T. Starner, "View-based and modula eigenspaces for face recognition," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.84-91, Seattle, Washington, June 1994.
- [8] D.L. Swet and J.J. Weng, "SHOSLIF-O: SHOSLIF for object recognition and image

- retrieval(phase II)," *Technical report cps-95-39*, Department of Computer Science, Michigan State University, Oct. 1995.
- [9] C. Chatterjee and V. P. Roychowdhury, "On self-organizing algorithms and networks for class-separability feature," *IEEE Trans. on Neural Networks*, vol. 8, no. 3, May 1997.
- [10] Guodong Guo, Stan Z. Li and Kap Luk Chan, "Support vector machines for face recognition," *Image and Vision Computing*, vol.19, pp.631-638, 2001.
- [11] P. J. Phillips, H. J. Moon, Syed A. Rizvi, and P. J. Rauss, "The FERET evaluation methodology for face-recognition algorithms," *From the Technical report NISTIR 6264*, Jan. 1999.



송 근 원

- 1993년 경북대학교 전자공학과(공학사)
- 1995년 경북대학교 전자공학과(공학석사)
- 1998년 경북대학교 전자공학과(공학박사)
- 1999년~현재 위덕대학교 조교수
- 관심분야 : 패턴 인식, 영상처리, 신호처리
- E-mail : kwsong@mail.uiduk.ac.kr