

소 특 집

얼굴인식에 관한 기술 동향
(Survey on Face Recognition)

고 재 필, 변 혜 란
연세대학교 컴퓨터과학과

I. 서 론

최근 몇 년간 인터넷을 통한 정보공유가 급속화 되면서, 공유할 수 없는 개인 정보에 대한 가치가 상승하고 있다. 더불어 개인정보 보호 문제가 대두되면서, 생체 정보를 이용한 보안기술이 주목받고 있다. 생체 정보는 개인의 생리적인 특징이나 행동적 특징을 말하며, 이는 타인에 의해 쉽게 복제될 수 없는 개인의 고유한 특징을 나타낼 수 있다. 이러한 생체 정보를 이용해 개인을 인식하는 기술에는 지문인식, 홍채인식, 서명검증, 정맥인식, 음성인식, 얼굴인식 등이 있다.

얼굴인식은 타 생체 인식 기술에 비해, 경제성과 사용자 편리성이 높다. 고가의 장비 대신 일반적인 PC 카메라를 이용할 수 있으며, 사용자는 기기에 직접 접촉을 하지 않고, 자신도 모르는 사이에 인식이 이루어질 수 있다. 이러한 이유로 얼굴 인식은 최근 몇 년간 영상 이해 분야의 가장 성공적인 응용의 하나로 주목받고 있으며, AFGR¹⁾



〈그림 1〉 생체 인식 기술

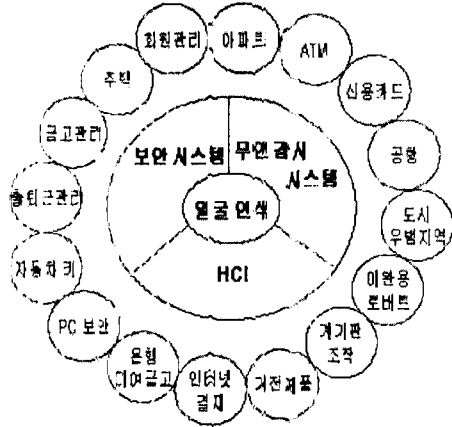
이나 AVBPA²⁾ 등과 같은 얼굴인식 분야의 전문 국제 학술 대회가 개최되고, FERET¹¹⁾과 같은 얼굴인식 평가 방법론이 정립되고 있으며, 얼굴인식 기술을 상업화에 성공한 기업들이 속속 출현하고 있는 점은 이를 잘 뒷받침 한다. 대표적인 외국기업은 <표 1>에 정리하였다.

국내의 경우, (주)위치비전, (주)프라임테크놀로지, (주)비전 인터랙티브 등이 최근 1-2년간 얼굴인식 관련 상품개발에 주력하고 있다.

〈표 1〉 대표적 얼굴인식 회사(국외)

회 사	제품명	기 술
DCS www.dcs.de	BioID	고유얼굴, 신경망
ETrue www.etrue.com	TrueFace Enging, ID	신경망
Viisage www.viisage.com	Vissage's Visual Inspection System(VIS)	고유얼굴
Visionics www.visionics.com	FaceIt Series	LFA(local feature Analysis)

1) Proceedings of the International Conferences on Automatic Face and Gesture Recognition
2) Proceedings of the International Conference on Audio and Video based Person Authentication



〈그림 2〉 얼굴인식 응용 분야



〈그림 3〉 얼굴인식 단계

검증하여 개인화인 또는 거절을 결정한다. 본 고에서는 얼굴영역 검출, 표현 및 특징추출, 인식기술 등 얼굴인식 전반을 소개하고, 얼굴인식 방법론을 개략적으로 평가할 수 있는 얼굴인식 기술 평가 방법론을 설명하고 최근 얼굴인식 연구 방향을 짚어본다.

II. 얼굴영역 검출

1. 얼굴영역 검출이란

얼굴검출은 임의의 영상에서 얼굴의 위치를 자동으로 추출하는 기술로서 얼굴관련 연구 분야 중 하나이다. 초기 얼굴검출은 자동화된 얼굴인식을 위한 전 단계 기술로서 연구가 시작되었으나 문제의 어려움과 중요성으로 인해 최근에는 하나의 독립적인 분야로 연구가 활발히 진행되고 있다. 얼굴검출은 얼굴을 포함하는 최소 영역을 찾아내는 것을 말하며, 응용에 따라 얼굴의 정확한 경계영역을 찾는 것을 포함할 수 있다. 예를 들어, 임의의 얼굴은 3차원 모델과 정합시 정확한 경계영역 검출은 필수적이다. 그러나, 일반적인 경우에는 전자의 지칭하는 경우가 대부분이다. 얼굴영역 검출은 인종, 성별, 나이별, 얼굴 움직임, 표정 변화 등의 얼굴이 가지는 변화 요인 뿐 아니라, 주변 환경 즉, 복잡한 배경, 조명 변화, 장신구 착용등에 의한 환경적인 요인이 무관하게 검출할 수 있는 기법 개발을 목표로 한다. 얼굴검출을 위해서는 얼굴 요소 검출이 수반되는 경우가 있으며, 얼굴 요소 검출이라는 또 다른 연구 분야로 구분하기도 한다. 최근에는 피부색 컬러 모델링을 통해 얼굴의 후보 영역을 찾고, 얼굴의 지형적 특징에 근거한 단계적 분석을 통해 얼굴영역을 확정짓는다.

얼굴인식은 보안 분야 뿐 아니라 그 응용분야도 다음과 같이 매우 다양하다. 크게는 보안 분야, 감시 분야, HCI(Human Computer Interface) 분야로 나눌 수 있다. 구체적인 응용 분야는 〈그림 2〉에 정리하였다.

최근 얼굴인식이 이처럼 주목받을 수 있는 데에는 상업적인 요구가 증대되었고, 이를 수용할 수 있는 지난 30년의 연구기간을 통해 쌓여진 기술력이 바탕이 되었기 때문이다. 그러나, 얼굴인식은 타 생체인식에 비해 정확도가 떨어지는 것은 여전히 극복해야 할 점으로 남아 있으며, 이는 다음과 같은 얼굴 인식의 어려움에 기인한다.

- 얼굴 자체의 변화 : 나이, 화장, 장신구
- 외부의 변화 : 배경, 조명, 포즈

얼굴인식의 일반적인 단계는 다음과 같다. 얼굴영상이 입력으로 주어지면, 복잡한 배경으로부터 얼굴영역만을 분리해내고, 분리된 얼굴영역으로부터 특징을 추출한다. 추출된 특징에 기반하여 얼굴인식을 수행한다. 마지막 단계의 얼굴인식은 다시 식별(identification)과 검증(verification)으로 나눌 수 있다. 식별은 시스템에 알려지지 않은 영상이 입력으로 주어지면 시스템을 데이터베이스에 저장된 개인 중 가장 일치되는 사람을 출력해 준다. 반면, 검증은 영상뿐 아니라 개인 식별 정보를 함께 입력하여, 시스템이 이를

2. 얼굴검출의 중요성

얼굴인식의 문제는 얼굴인식 자체의 문제가 아니라, 얼굴검출의 문제로 파악하고자 하는 견해가 많다. 정확한 얼굴검출이야말로 얼굴인식의 전처리 단계로서가 아니라 얼굴인식 관련 연구의 가장 중요한 부분 중의 하나이다. 최근 얼굴인식의 개념이 HID(Human In a Distance)로 바뀐에 따라 얼굴인식이 실제로는 얼굴검출을 의미하게 될 것이다. 얼굴검출은 얼굴인식을 위한 전 단계로의 역할 뿐 아니라 그 자체로의 응용성을 가지게 되어 얼굴감시 응용도 그 중 하나로, 공항, 범죄 우범 지역 등에서 얼굴 영상을 수집하는 부분에 활용되고 있다. 물론, 개인 사생활 침해의 소지가 크고 최근 미국 정보국의 슈퍼볼 게임 입장객을 몰래 촬영, 수집하여 범죄 용의자 검색에 활용한 사실이 밝혀지며 논란이 다분하지만 그 응용은 늘어날 것으로 사려 된다.

3. 얼굴검출 기법 분류

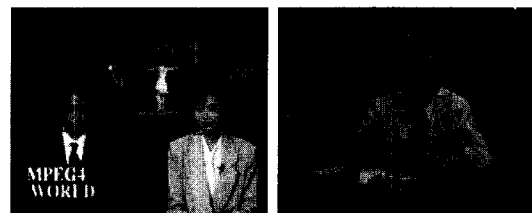
얼굴검출을 위한 수 많은 기법들이 연구되어 이들의 성능평가 또는 기술분류가 매우 난해하다. 따라서, 여기서는 활용할 수 있는 정보가 무엇이나에 따라 구분해 본다. 사용할 수 있는 정보에 따라 동일한 검출 성능을 얻기 위한 난이도가 달라지므로, 각 검출기법은 목표도 다르고, 접근 방법 또한 달리해야 할 것이다.

얼굴검출은 입력 데이터에 따라 정지영상에서의 검출과, 동영상에서의 검출로 분류할 수 있다. 동영상에서의 연속적인 검출은 얼굴추적으로 구분되나, 큰 범주로는 동영상에서의 검출로 분류한다. 동영상인 경우는 정지영상에서 사용할 수 있는 영역정보 이외에 시간정보를 추가적으로 사용할 수 있다. 이는 컬러정보와 결합하여 실시간 검출과 같은 보다 넓은 응용범위를 가진다. 그러나, 동영상은 예를 들어 카메라 움직임 보상등과 같은 또 다른 문제점을 안고 있어 실제로 시간정보가 추가된 만큼 성능 향상이 이루어 지지 않는다는 점이다.

다음은 입력영상이 가진 정보량 또는 제약조건에 따른 구분이다.

- ◎ 잘 제어된 배경을 가진 영상인 경우
이 경우는 특별히 고려할 사항이 없고 이미 알고 있는 배경에 대한 지식을 활용하여 배경을 제거함으로써 얼굴영역을 찾아 낼 수 있다.
- ◎ 컬러 정보를 사용할 수 있는 경우
전형적인 얼굴피부색을 모델링하여 얼굴을 분리해 내는 기법으로 인종에 따라, 조명 조건등에 매우 민감하다는 단점을 갖는다. 따라서, 피부색을 효과적으로 모델링할 수 있는 피부색 모델 개발과 조명에 따른 변화를 정규화하는 기법연구가 선행되어야 한다.
- ◎ 모션 정보를 사용할 수 있는 경우
살아있는 사람인 경우 얼굴이 정지해 있기가 어렵고 거의 항상 움직인다는 점을 이용하여 움직임 영역을 찾음으로써 얼굴을 찾는 기법이다. 그러나, 장면 안에 움직이는 물체가 다수 존재할 수 있기 때문에 이를 구분해야 하는 문제점이 있다. 또는 카메라 자체의 움직임으로 인한 모션을 감지해야 하는 어려움이 따른다.
- ◎ 위의 정보를 모두 사용할 수 있는 경우
각 기법의 장점들을 결합하여 얼굴을 검출하는 기법으로 효과적인 결합 방법에 대한 연구가 필요하다.

향후에는 동영상 정보를 사용할 수 있는 가능성이 커지고, 실시간 응용에만 한정되는 것이 아니라 오프라인 응용에도 확대될 것으로 추측된다. 따라서 동영상을 대상으로 하여 시간 정보를 충분히 활용할 수 있는 방안과 기존의 정지영상에서 사용하는 기법과의 적절한 결합 방안에 대한 방향으로 연구가 진행되어야 할 것이다. 보다



〈그림 4〉 복잡한 배경하에서의 얼굴영역 추출 예

현실적으로는 응용 분야를 염두해두고 기술 개발이 이루어져야 할 것으로 시켜된다. 일반적인 얼굴검출 기법으로는 Kanade 연구팀이 제안한 신형망 기반 방법^[2]과 Sung & Poggio에 의해 제안된 예제 기반 학습 방법^[3]이 대표적이다.

III. 얼굴표현 및 특징추출

인식단계에서 매칭성능을 높이기 위해서는 얼굴영상을 효과적으로 표현하고 정규화하는 과정이 반드시 필요하다. 얼굴영상을 표현하기 위해서 얼굴의 모양정보 및 텍스처 정보를 벡터화 한다. 모양정보는 얼굴 요소 즉, 눈, 코, 입 등과 같은 얼굴 요소들간의 지형적 특징으로, 거리, 비율 등의 정보가 이에 해당한다. 텍스처 정보로는 얼굴 영역내의 밝기 정보 자체로써 표현한다. 모양 정보를 이용한 얼굴인식은 먼저 눈, 코, 입과 같은 얼굴요소를 정확히 추출하는 것이 선행되어야 하는데, 이는 실제로 매우 어려운 문제로서 얼굴인식 기술에는 거의 적용되지 않고 있다. 대표적인 모양정보를 이용한 얼굴인식 기법에는 그래프 매칭 방법이 있다^[4]. 이를 제외하고는 대부분 얼굴영역 전체를 텍스처 정보로 사용하는 기법의 인식률이 높은 것으로 보고되고 있으며, 최근에는 얼굴요소 영역 기반 방식이 높은 인식률을 보고하고 있다^[5].

1. 얼굴의 벡터 표현

2차원의 얼굴영상을 1차원 벡터로 표현하는 방법은 각 픽셀 값을 차례대로 나열하는 방법이 사용된다. 그러나, 이러한 표현 방식은 위치 변환 및 포스변화에 매우 민감하며, 저장 공간이 많이 요구되어진다. 128x128 영상의 경우 2^{14} (=16,384) 바이트가 필요하며, 무엇보다도 벡터의 차원이 매우 커짐에 따라 학습을 효과적으로 수행할 수 없다. 따라서, 고차원의 얼굴영상 벡터를 효과적으로 표현하기 위한 특징추출이 반드시 요구되어지며, 주성분 분석 기법(PCA)이 널리 이용되고

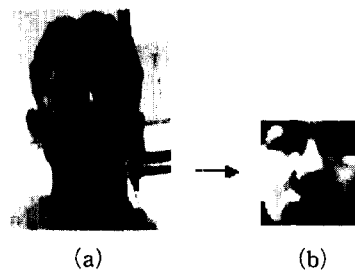
있다. 주성분 분석 기법은 뒤에서 다시 자세히 다룬다.

2. 정규화

얼굴영역의 일반적인 정규화 과정은 다음과 같다. 첫째, 어파인(affine) 변환을 수행한다. 둘째, 영상의 밝기 정보를 조정한다. 셋째, 배경 및 머리카락의 영향을 줄이기 위해 타원 마스크를 적용한다. 어파인 변환에는 이동, 크기 조정, 회전 변환을 포함한다. 두 눈을 기준으로 얼굴영상이 수직이 되도록 회전하고, 다시 일정한 크기로 변환한 후, 두 눈이 항상 일정한 위치에 오도록 이동한다. 영상의 밝기 정보는 비선형변환인 히스토그램 평활화를 적용하여 조정한다. 히스토그램 평활화를 통해 어느 정도 영상의 밝기에 무관한 얼굴영상을 얻을 수 있다. 그러나, 조명방향에 따른 그림자 효과는 제거할 수 없다. <그림 5>는 입력 영상에 대한 정규화 결과를 보여준다.

3. 특징추출

특징 추출은 모작·패턴인식 문제에 있어 중요한 부분이다. 본 고에서는 얼굴인식 기법에서 사용되고 있는 대표적인 기법 두개를 소개한다. 대부분의 특징 추출은 선형변환을 통해 이루어지며, 선형변환을 위한 기저벡터를 찾는 것이 특징추출 과정이 된다. PCA(Principal Component Analysis) 기법은 기저벡터로써 공분산 행렬의 고유벡터를 찾는 문제이고, LDA(Linear Discriminant Analysis)는 클래스내의 분산은 작



(a) Original Image (b) Normalized Image

<그림 5> 얼굴영상 정규화 예

아지고 각 클래스의 평균간 거리는 멀어지도록 하는 기저벡터를 찾는 문제이다. LDA 기법에서는 일반적으로 얼굴 벡터의 차원이 매우 크기 때문에 PCA를 적용한 후 다시 LDA를 적용한다.

PCA 기법

주성분 분석 기법은 데이터가 가지고 있는 원래의 특징을 가지고 있으면서 효율적으로 차원을 줄일 수 있도록 설계되어, 평균자승오차(mean-square error) 측면에서 최적의 선형 변환이다.

차원이 n 인 랜덤벡터 \mathbf{X} 를 평균이 0이라고 가정하면 즉 $E[\mathbf{X}]=\mathbf{0}$ 일 때, 다음과 같이 \mathbf{X} 에 대한 $n \times n$ 차원의 공분산 행렬 \mathbf{R} 은 벡터 \mathbf{X} 에 대한 외적의 기대값으로 정의된다.

$$\mathbf{R}=E[\mathbf{X}\mathbf{X}^T]$$

공분산 행렬 \mathbf{R} 의 고유값들을 $\lambda_1, \lambda_2, \dots, \lambda_n$, 대응되는 고유벡터들을 $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ 라고 하고 고유값이 큰 순서대로 대응되는 고유벡터를 나열하여 다음과 같이 쓴다.

$$\mathbf{Q}=[\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n]$$

그러면, PCA 변환은 \mathbf{X} 의 어떤 실제벡터를 \mathbf{x} 라하고 변환행렬로 \mathbf{Q} 를 사용하여 특징벡터 \mathbf{y} 를 얻는 다음 식으로 정의된다.

$$\begin{aligned} \mathbf{y} &= [y_1, y_2, \dots, y_n]^T \\ &= [\mathbf{x}^T \mathbf{q}_1, \mathbf{x}^T \mathbf{q}_2, \dots, \mathbf{x}^T \mathbf{q}_n]^T = \mathbf{Q}^T \mathbf{x} \\ y_j &= \mathbf{q}_j^T \mathbf{x} = \mathbf{x}^T \mathbf{q}_j \quad j=1, 2, \dots, n \end{aligned}$$

\mathbf{x} 를 단위벡터 \mathbf{q}_j 로 표현되는 주요방향으로 투영하여 얻은 결과를 y_j 라고 할 때 이를 주성분이라고 한다. 결국, 주성분 분석은 데이터 샘플이 표현하고 있던 분산이 최대가 되도록 새로운 축을 만들고, 이러한 축들은 서로 직교하게 됨으로써 상위 몇 개의 축에 의해서 표현되는 값은 입력 벡터가 담고 있는 대부분의 정보를 담게 된다.

LDA 기법

LDA는 데이터를 가장 잘 표현하는 공간이 아

닌, 클래스들을 가장 잘 구분할 수 있는 공간을 구성하는 기저벡터를 찾는다. 이를 위해 먼저 클래스내부 분산 행렬(within-class scatter matrix) 및 클래스간 분산 행렬(between-class scatter matrix)를 각각 아래와 같이 정의한다.

$$S_w = \sum_{j=1}^c \sum_{i=1}^{N_j} (\mathbf{x}_i^j - \mu_j) (\mathbf{x}_i^j - \mu_j)^T$$

여기서, \mathbf{x}_i^j 는 j 클래스의 i 샘플을, μ_j 는 j 클래스의 평균을, c 는 클래스 수를, 마지막으로 N_j 는 클래스 j 의 샘플수를 나타낸다.

$$S_b = \sum_{j=1}^c (\mu_j - \mu) (\mu_j - \mu)^T$$

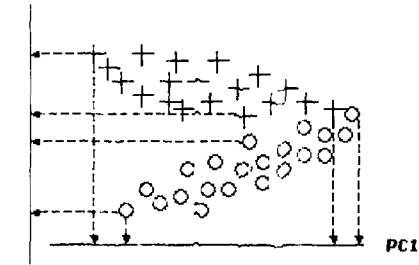
여기서, μ 는 모든 클래스들의 평균을 나타낸다.

이때, 클래스 평균간의 차이를 크게 하고, 클래스 내부의 분산을 줄일 수 있도록 하기 위해 다음과 같은 평가함수를 정의하고, 이를 최대화하는 변화행렬 \mathbf{w} 를 찾으면 된다.

$$J(\mathbf{w}) = \frac{\mathbf{w}^T S_b \mathbf{w}}{\mathbf{w}^T S_w \mathbf{w}}$$

PCA v.s. LDA

PCA는 전체 데이터에 대해 분산이 크게 되는 축을 찾게 되고, FD는 각 클래스의 분포를 보고 분류 오류를 줄일 수 있는 방향으로 축을 찾게 된다. 그러나, 일반적으로 클래스간의 분포를 알기 어렵기 때문에—이는 주어진 학습 샘플이 각 클래스의 분포를 잘 반영하고 있다고 보기 어렵다는 의미—반드시 LDA가 좋은 결과를 낼 수 있다고 말할 수는 없다. 그러나, 얼굴인식의 경우, LDA가 인식성능이 다소 우수하다고 보고되고 있다^[6]. 특히, PCA적용 후 LDA를 적용한 경우가 가장 높은 인식률을 보이는 것으로 보고되고 있다^[7]. PCA는 각 클래스의 분포에 상관없이 전체 샘플에 대해 적용될 수 있다는 것이 장점 중의 하나이며, 얼굴과 같은 고차원 데이터에 대해 효과적으로 차원을 줄여주는 기법임에는 틀림없다. <그림 6>은 PCA 및 LDA가 찾아준 첫 번째 축을 보여준다.



LD1
 <그림 6> PCA와 FA에 의해 결정된 각각의 기저 벡터

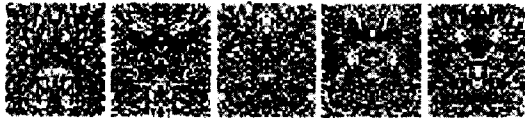


Figure 1: The first five pure LDA bases



Figure 2: The first five PCA+LDA bases



The average face and first four eigenfaces

<그림 7> FERET 데이터 베이스에 대한 LDA, PCA+LDA, LDA 기저벡터, 출처^[6]

다음 <그림 7>은 각각 LDA, PCA+LDA, PCA에 의해 구해진 기저벡터의 일부를 영상으로 표현한 것이다. PCA의 경우 이렇게 구해진 기저벡터가 얼굴과 유사하다고 하여 고유얼굴이라고 불려온다.

IV. 인식기

얼굴 인식을 위한 인식기로는 NNC(nearest neighbor classifier), NN(neural network) 등이 주로 사용되었으며, 최근에는 SVM(sup-

port vector machine)을 이용한 논문이 발표되고 있다. 본 고에서는 최근 주목받고 있는 SVM에 대하여 간략히 설명한다^[8].

SVM(Support Vector Machines)

SVM은 1995년 Vapnik이 제안한 2-클래스 분류기로서, 선형분리 가능한 샘플들에 대하여 최적의 하이퍼플레인(Optimal Hyperplane)을 만들어 줌으로서, 뛰어난 일반화 성능을 보여준다. 두 클래스에서 가장 가까운 샘플들을 Support Vector라고 하며 이들의 거리가 최대가 되도록 다음 수식에서 w 을 구하는 것으로 SVM을 학습하게 된다.

$$w \cdot x + b = 0, \quad w = \sum_{i=1}^{N_s} \alpha_i y_i s_i$$

여기서, x 는 입력벡터, b 는 상수, α_i 는 웨이트의 계수이고, N_s 는 서포트벡터의 수이며, y_i 는 목표 값이다. SVM은 선형분리가 가능하지 않은 샘플에 적용하기 위해서는 소프트마진 및 커널 함수를 도입한다. 입력 패턴을 비선형 커널 함수를 통해 다른 공간으로 맵핑함으로써 선형분리될 가능성을 높이도록 한 것으로, 다음 수식과 같다.

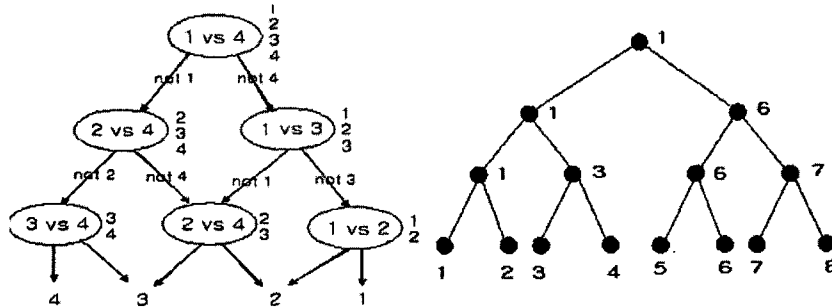
$$\sum_{i=1}^{N_s} \alpha_i y_i K(s_i, x) + b = 0$$

커널에 사용되는 대표적인 함수로는 RBF, 다항식, 퍼셉트론등이 대표적이다.

다중 SVM 분류기

SVM은 기본적으로 2 클래스 분류기 임으로 얼굴인식과 같이 다중 분류문제를 풀 수는 없다. SVM을 다중 클래스 분류기로 확장하는 방법은 기본적으로 2클래스 SVM을 결합하는 것으로서 2가지 기법이 제시되었다. 하나는 '1대나 정책(one vs all strategy)'으로 하나의 클래스와 나머지 클래스들을 구분하는 것이고, 다른 하나는 '1대1 정책(one vs one strategy)'으로 분류쌍(pair)을 구분하는 것이다^[9].

클래스가 n 개인 경우 '1대나' 정책은 각 클래스



〈그림 8〉 분류쌍 방법을 트리로 구성하는 예(왼쪽부터 각각 하향식, 상향식)

여부를 결정하는 SVM을 n 개 구성하여 출력값이 가장 높은 SVM을 승자 클래스로 결정하는 방식으로 n 번의 비교가 이루어진다. 후자의 경우는 각 클래스의 조합만큼 즉, $n(n-1)/2$ 개의 SVM을 구성하고, 트리를 이용하여 토너먼트 방식으로 승자를 결정할 경우 $n-1$ 회의 비교가 이루어진다. 전자는 SVM 구성이 간단하지만, SVM을 학습하기 위한 타 클래스 샘플 (negative sample)을 정하는 것이 모호하다는 단점이 있다. 후자는 타 클래스 샘플 수집의 모호성은 없으나 학습시켜야 할 SVM의 수가 $O(n^2)$ 다. 두 기법의 차이에도 불구하고, 얼굴인식에 있어서 분류성능의 차이는 거의 없는 것으로 보고되고 있다^[6].

V. 평가 방법론

신뢰할 수 있는 얼굴인식 시스템 개발을 지원하는데 있어서 가장 중요한 요구사항 두 가지는 큰 규모의 얼굴 영상 데이터베이스와 개발된 얼굴인식 시스템을 평가하는 테스트 절차라 할 수 있다. 본 장에서는 현재까지 소개된 평가 방법론에는 FERET과 XM2VTS가 있다. 후자는 음성정보와 더불어 멀티모달 검증을 위한 데이터베이스 및 검증 프로토콜로 구성되어 있으며, 특히, 검증 프로토콜을 위해 데이터베이스를 학습셋, 평가셋, 테스트셋으로 나누고 있는 것이 특이한 점이다. 본 장에서는 데이터베이스, 식별 평가 방법론, 검증 평가 방법론을 모두 포함하고 있는

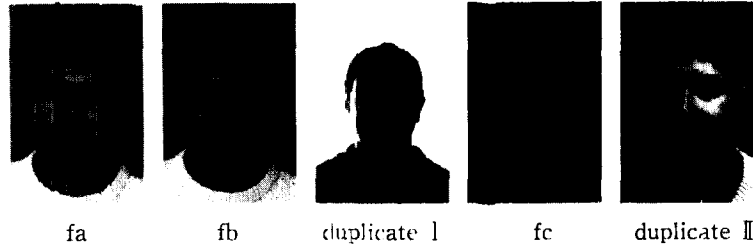
FERET 평가 방법론을 간략히 살펴보기로 한다. 보다 자세한 사항은 다음 참조문헌을 참고하기 바란다^[1].

FERET 프로토콜

FERET은 얼굴인식 알고리즘의 개발, 대규모의 얼굴영상 데이터베이스 구축, 얼굴인식 알고리즘의 성능평가 방법의 세 가지를 포함하고 있으며, 현재까지 1199명의 14,126장의 영상이 FERET 데이터베이스에 모아졌고, 1996년 9월에 FERET 얼굴인식 테스트 시리즈의 세 번째 테스트가 실시되었다.

1. 데이터베이스

FERET 데이터베이스의 얼굴 영상은 1993년 8월부터 1996년 7월 사이에 15세션으로 수집되었다. 한 번의 수집 세션은 하루 또는 이틀에 걸쳐 진행되었고, 데이터베이스의 일관성을 유지하기 위해 카메라의 위치나 물리적인 셋팅은 변화가 없도록 유의하였으나, 매 세션마다 재조립되었기 때문에, 촬영 세션이 달랐던 영상들 사이에는 약간의 변화를 볼 수 있다. 얼굴 영상은 한 사람 당 11장의 영상을 획득되었는데, 얼굴 표정을 달리한 정면 사진 2장 (fa, fb), 이중 200명에 대해서는 촬영된 2장 (fa, fb)과는 카메라 및 조명을 달리하여 세 번째 정면 사진을 획득하였다 (fc). 두 번째 촬영에서는 안경을 착용하도록 또는 머리를 뒤로 넘기도록 하고, 때때로 시간 간격을 두어 두 번째 영상을 촬영하기도 하였다. 이런 duplicate 들로 인해 데이터베이스는 얼굴에 나



〈그림 9〉 FERET 데이터베이스 영상

다단 소명, 표징, 크기 면에서 다양성을 갖게 되어 있다. 1996년 7월에는 1199명(365개의 duplicate 셋 포함)의 1564셋, 14,126의 영상을 획득되었고 이후 503개 셋은 연구자들에게 공개되었으며, 나머지는 정부에 의해 따로 관리되고 있다.

1. 평가프로토콜

FERET 얼굴인식 테스트는 다음과 같은 용어를 사용하여 FERET 데이터를 구분하여 사용한다.

- 갤러리(gallery) : 이미 알고 있는 사람들이 영상들을 포함하는 집합
- 프로브(prob) : 테스트 과정에서 인식하도록 알고리즘에 입력되는 신규 영상
- 프로브 셋(prob set) : 테스트에 사용될 갤러리 영상들의 집합
- 타겟 집합(target set) : 알고 있는 얼굴 영상들의 집합
- 쿼리 집합(query set) : 구별되어야 하는 알지 못하는 얼굴 영상들의 집합

테스트는 FB, fc, Duplicate I, Duplicate II 프로브 집합에 대하여 식별과 검증 테스트를 실시하였다. 테스트'96에서는 입력 영상만 주어지는 완전자동알고리즘과 입력영상과 두 눈의 좌표가 주어지는 반자동 알고리즘에 대한 테스트를 수행하였다.

3. 식별성능 평가

식별에서는 “프로브 영상들 인식하는데 이 알고리즘이 얼마나 좋으냐?”라는 질문에 대한 답을 주는데 있어, “갤러리에서 가장 유사도가 높은

영상이 정확한 답이 맞습니까?” 보다는 “갤러리에서 프로브 영상에 유사도가 높다고 뽑은 상위 n개 영상 중에 정의한 답이 있습니까?”로 설정하였다. 성능 평가 통계는 매치점수의 누적 그래프로 보고되고 그래프의 수평 축은 랭크 n을, 수직 축은 정확한 매칭점수를 나타낸다. 평가는 동일한 갤러리에 서로 다른 카테고리의 프로브 셋을 사용하여 알고리즘의 카테고리 별 점수를 계산하였다.

4. 검증성능 평가

검증모델에서는 프로브의 사람이 자신이 갤러리 이미지와 같은 사람이라고 주장하는 것에 대해, 알고리즘은 그 주장을 받아들이거나 또는 거부하는 것으로 결과를 준다. 검증 알고리즘의 성능은 맞는 사람에 대해 거부하는 비율(False Reject Rate)와 틀린 사람에 대해 받아들이는 비율(False Acceptance Rate)의 두 개의 성능 통계 값으로 평가되어 진다. 식별과 마찬가지로, 갤러리 영상과 프로브 영상이 있고 프로브가 갤러리 영상과 같은 사람임을 주장할 때, 두 영상 간 유사도를 사용하여 주장하는 바를 받아들일지 거부할지 결정하게 된다. 이때, 유사한 정도의 기준이 되는 상수 C는 최소값부터 최대값까지 변화시켜가면서 매 C값에 따른 FAR과 FRR의 가능한 조합을 구해 볼 수 있는데, 이것을 ROC(receiver operating characteristic or relative operating characteristic)의 그래프로 나타내 볼 수 있다. 이때, FAR과 FRR이 같아지는 점을 EER(Equal Error Rate)로 정의하고, EER이 작은 경우를 성능이 좋다고 평가할 수

있다.

5. 의의

FERET 평가 프로토콜은 서로 다른 시나리오에 대해 서로 다른 갤러리와 프로브 셋을 구성하여 알고리즘을 테스트하여 알고리즘 별 독립적인 성능 평가를 가능하게 하였고, 이 평가 프로토콜은 얼굴식별과 얼굴검증에 있어서 알고리즘들이 각기 다른 갤러리와 프로브 셋에 대해 어떤 성능을 보여주는지에 대한 직접적인 비교, 분석을 가능하게 해주기 때문에, 최근에는 얼굴인식 알고리즘의 성능 평가의 표준이 되고 있다.

VI. 최근 연구 방향

얼굴인식은 지난 30년간 패턴인식 및 영상이해, 영상처리 분야의 가장 성공한 연구분야중의 하나로 성장해 왔다. 또한 FERET과 같은 평가 방법론이 도출되면서, 얼굴인식 알고리즘 개발은 객관적인 평가 지표에 따라 최근 몇 년간 급속하게 발전되었다. 그러나, 얼굴인식은 여전히 많은 해결해야 할 문제점을 안고 있다.

특히, 조명과 포즈에 무관한 얼굴인식 알고리즘 개발을 통해 성능을 향상시키고자 하는 노력이 시도되고 있다. 얼굴은 3차원 물체이므로 조명의 작은 변화에도 2차원에서의 얼굴모습은 매우 달라지게 되므로, 이러한 조명효과를 효과적으로 제거하는 기술이 요구되어진다. 최근에 조명론 모델이 제시되어 관심을 받고 있다^[10]. 그러나, 여전히 모델을 위해 여러장의 영상이 필요하다는 단점이 지적되고 있다. 얼굴인식 자체의 문제만 볼 때, 향후 얼굴인식 연구는 조명과 포즈에 무관한 얼굴인식 알고리즘 개발에 집중될 것으로 예측된다.

실시간 감시 응용에 대한 요구 증대로, 실시간 얼굴영역 검출에 관한 연구도 활발해 질 것으로 기대된다. 동영상에서의 실시간 검출에 따른 문제점은 컬러 상수화 문제를 해결할 수 있는 컬러

피부색 모델의 개발로 집중되고 있으며, 한편으로는 적외선 카메라를 사용하는 방법도 제시되고 있다. 동영상에서 실시간 검출에 그치지 않고 인식까지 염두한다면, 포즈, 조명에 강인한 인식기술 개발은 물론, 저해상도에서의 얼굴영상 질을 향상시키기 위한 연구도 필요하다.

VII. 결 론

본 고에서는 얼굴인식의 분야의 얼굴검출, 특징추출, 인식기의 전반적인 내용을 다루었다. 또한, 객관적인 평가지침이 되는 FERET 얼굴인식 평가 방법론을 소개하였다. 얼굴검출은 응용과 사용할 수 있는 정보에 따라 구분해 보았으며, 특징추출은 얼굴인식 분야에서 널리 쓰이는 PCA 및 LDA를 다루었으며, 인식기로는 최근 주목받고 있는 SVM을 소개하였으며, 마지막으로 얼굴인식 분야에서 최근 연구 방향을 소개하였다.

향후 얼굴인식 분야의 연구는 얼굴포즈 및 조명에 무관한 얼굴인식 기술개발에 집중될 것이며, 이를위해 얼굴표현, 특징추출, 3차원 모델링, 분류기 분야 역시 지속적으로 연구되어질 것으로 전망된다.

참 고 문 헌

- [1] P.Jonathon Phillips, Hyeonjoon Moon, Syed A. Rizvi, and Patrick J. Rauss, "The FERET Evaluation Methodology for Face-Recognition Algorithms", From the Technical report NISTIR 6264, January 7, 1999.
- [2] H.A. Rowley, S.Baluja, and T. Kanade, "Neural Network Based Face Detection", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 20, 1998

- [3] K. Sung and T. Poggio, "Example-based Learning for View-based Human Face Detection", A.I. Memo 1521, MIT A.I. Lab. 1994
- [4] Wiskott, L.; Fellous, J.-M.; Kuiger, N.; von der Malsburg, C., "Face recognition by elastic bunch graph matching", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 19, pp. 775-779, 1997
- [5] Heisele, B.; Ho, P.; Poggio, T., "Face recognition with support vector machines: global versus component-based approach", IEEE International Conference on Computer Vision, Vol. 2, pp. 688-694, 2001
- [6] A.M. Martinez and A.C. Kak, "PCA versus LDA", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 29, pp. 228-233, 2001
- [7] Zhao, W.; Chellappa, R.; Krishnaswamy, A., "Discriminant analysis of principal components for face recognition", IEEE International Conference on Automatic Face and Gesture Recognition, pp. 336-341, 1998
- [8] Vladimir N. Vapnik. "Statistical Learning Theory", JOHN WILEY & SONS. INC, 1998
- [9] Gudong Guo, Stan Z. Li, and Kapluk Chan. "Face Recognition by Support Vector Machines", IEEE International Conference on Automatic Face and Gesture Recognition, pp. 196-201, 2000
- [10] A.S. Georgiades, D.J. Kriegman, P.N. Belhumeur, "Illumination Cones for Recognition Under Variable Lighting", IEEE Conference on Computer Vision and Pattern Recognition, pp. 52-58, 1998

저자 소개



高 宰 必

1996년 2월 연세대학교 학사(전산학), 1998년 8월 연세대학교 석사(컴퓨터과학), 현재: 연세대학교 박사과정, <주관심 분야: 얼굴인식, 패턴인식, 영상처리>



卞 惠 蘭

1980년 2월 연세대학교 학사(수학), 1983년 2월 연세대학교 석사(수학), 1993년 12월 Purdue Univ. 박사(전산학), 1994년 3월~1995년 2월: 한림대학교 조교수, 1995년 3월~현재: 연세대학교 부교수, <주관심 분야: 얼굴인식, 영상검색, 패턴인식, 영상처리>