

Face Recognition Grand Challenge (FRGC) 및 조명 변화에 강인한 얼굴 인식 기술 개발 동향

황원준 (삼성종합기술원), 김준모 (KAIST)

본 논문에서는 최근 얼굴 인식 평가에 많이 사용된 FRGC Ver 2.0 DB와 그 프로토콜을 간략히 소개하고 이를 이용한 다양한 얼굴 인식 방법 및 그 개발 동향에 대해서 살펴보고자 한다. FRGC는 객관적인 2D/3D 얼굴 인식 알고리즘 성능 평가를 위해서 공개되었는데, 본 논문에서는 2D 정면 얼굴 인식에 대한 내용을 위주로 기술하고자 한다. FRGC의 2D 얼굴 인식 DB는 주로 조명의 Control 유무에 따른 성능 비교를 위한 평가 프로토콜을 제안하고 있다. 이에 2004년부터 최근까지 10개 이상의 알고리즘이 발표되었고, 본 논문에서는 중요한 11개의 알고리즘을 살펴보고자 한다. 또한 이들 알고리즘에서 핵심적으로 사용되는 특징 추출 알고리즘을 살펴보고 마지막으로 각 알고리즘의 FRGC DB에서의 성능을 비교 평가하고자 한다.

I. 서론

Computer Vision 분야에서 얼굴 인식은 1990년도 초반부터 많은 연구원들이 관심을 가지는 연구 주제였다. 특히 EigenFace가 제안된 이후로 얼굴 인식 연구에서는 학습용 얼굴 영상을 모아 통계적으로 특성을 학습하는 방식이 많이 발표되었다. Independent Component Analysis (ICA), Local Feature Analysis (LFA), Locally Linear Embedding (LLE), TensorFace 등을 이용한 얼굴 인식 방법이 제안되었으나 실험적으로는 FisherFace가 가장 성공적인 얼굴 인식 성능을 보여주고 있다.

얼굴 인식을 이용한 응용 제품으로 얼굴 인식 기반 보안 솔루션, 촬영된 사진 중 동일 인물이 있는 사진을 찾아주는 서비스 등이 있다. 얼굴 인식 기반 보안 솔루션의 경우 2000년 초반에 일본의 휴대폰 제조사들이 FeliCa와 같은 RFID 스

마트카드의 보안을 높이기 위해서 앞다투어 적용하였으나 소비자들이 원하는 성능이 나오지 않아 현재는 거의 사용되고 있지 않는 상태이다. 하지만 최근 안드로이드 폰에 얼굴 인식을 이용한 잠금 화면 해제 기능이 들어가서 그 가능성을 다시 한 번 확인하고 있는 중이다. 동일 인물 검색의 경우 포털 업체들의 이미지 검색의 한 기능으로 적용되고 있으며, Apple이나 MS에서도 이러한 기능을 적극적으로 제품에 적용하려고 하고 있다.

하지만 얼굴 인식 기술은 아직까지는 모든 제품에 적용되지 못하고 있는데, 가장 큰 원인은 얼굴 인식 성능이 사람이 원하는 수준에 이르지 못했기 때문이다. 예를 들어서 현존하는 2D 영상 얼굴 인식은 얼굴의 포즈 문제를 완벽하게 풀지 못하고 있다. 몇몇 연구 기관에서는 등록 영상을 늘리거나 3D 얼굴 모델을 이용하고 있지만 소비자가 원하는 수준에 이르지 못했다. 그리고 조명, 표정 등의 변화에도 미묘한 성능 차이가 있고, 더 나가 Occlusion, 나이에 따른 얼굴 변화, 남여 및 인종 간 인식 성능 차이도 풀어야 할 문제이다.

이러한 얼굴 인식 방법의 성능을 객관적으로 평가하기 위한 시도가 많이 있었다. 미국의 경우 NIST를 중심으로 얼굴 인식 알고리즘 경쟁을 통한 인식 성능 향상을 유도하기 위해서 FERET 테스트를 진행했고, 또한 얼굴 인식 업체들의 성능을 객관적으로 평가하기 위해서 FRVT (Face Recognition Vendor Test)를 진행했다. 2000년대 중반에는 FRGC (Face Recognition Grand Challenge)를 열어서 학습/평가 DB 및 관련 프로토콜을 제공했다. 2004년경에 공개된 FRGC는 가장 많은 수의 얼굴 평가 영상을 제공하고 있었고, 이들 영상은 조명, 표정, 시간 변화를 포함하고 있었다. 마지막으로 최근에 NIST는 MBGC (Multiple Biometric Grand Challenge)를 열어 지문, 얼굴, 동공 인식을 모두 포함하는 테스트를 진행



〈그림 1〉 FRGC Ver 2.0 얼굴 영상 예제. 위쪽은 조명이 Control된 얼굴 영상의 예이고 아래 줄은 Uncontrol된 얼굴 영상의 예이다

중에 있다. 그 외 지역에서의 얼굴 인식 평가는 유럽의 경우 영국을 중심으로 BANCA DB, XM2VTS DB를 이용한 얼굴 인식 평가가 진행되고 있고, 일본의 경우 자국 업체들의 얼굴 인식 솔루션을 평가했으나 그 결과를 공개하고 있지는 않다.

II. 본 론

1. FRGC Ver 2.0 DB 및 프로토콜 소개

2004년 NIST에서는 FRGC를 통해 영상 데이터, 실험 프로토콜, 평가 툴을 FRGC 참석자들에게 제공했다. FRGC에서 제공하고 있는 고해상도 2D 정면 얼굴 영상은 Control 된 조명하에서 촬영한 Controlled Image와, 조명이 Control 되지 않은 상황에서 촬영한 Uncontrolled Image로 구성이 된다. Control 된 조명하에서 찍은 영상은 Studio 환경에서 조절된 광원을 받으며 촬영한 영상으로 주로 표정 변화와 시간에 따른 얼굴 변화가 관측이 된다. 반면 Uncontrol 된 조명하에서 촬영한 영상은 실내의 복도 조명이나 창문을 뒤로 하고 촬영된 영상들로 심한 조명 변화, 카메라 초점이 맞지 않아서 생기는 Blur, 그림자 등이 영상에서 관찰 된다.

이들 영상은 각각 4 메가 픽셀 디지털 카메라로 촬영이 되었고 1704x2272나 1200x1600의 고해상도로 촬영되었다. 2002년부터 2003년까지 촬영된 얼굴 영상으로 학습을 위해서는 222명의 12,776장의 영상을 제공해주고 있고, 이들은 Control과 Uncontrol 된 영상으로 반반 나누어져 있다. 평가를 위한 얼굴 영상은 2003년부터 2004년에 촬영된 영상으로 총 466명의 얼굴 영상으로 구성되어 있다. 이러한 얼굴 영상의 예는 〈그림 1〉을 보면 쉽게 알 수 있다.

FRGC Ver 2.0은 총 6가지 평가 실험(Exp)를 제공하는데 본 논문에서는 2D 정면 얼굴 인식 성능 평가에 가장 적합한 Exp1과 Exp4를 기준으로 기술하고자 한다. Exp1은 등록 영상(Target Set)이 Control 된 16,028장의 영상이고 평가 영상(Query Set)도 동일한 Control 영상을 사용하여 얼굴 인식 성

능을 평가하게 된다. 즉, 조명 변화가 심하지 않을 경우 시간, 표정 변화에 따른 얼굴 인식 성능을 나타내게 된다. 반면 Exp4는 등록 영상이 Control 된 영상이지만 평가 영상은 8,014장의 Uncontrol 영상을 이용하기 때문에 Exp1 보다 더 어려운 환경에서 얼굴 인식 성능을 평가하게 된다. 이러한 평가 프로토콜은 〈표 1〉에 정리되어 있다.

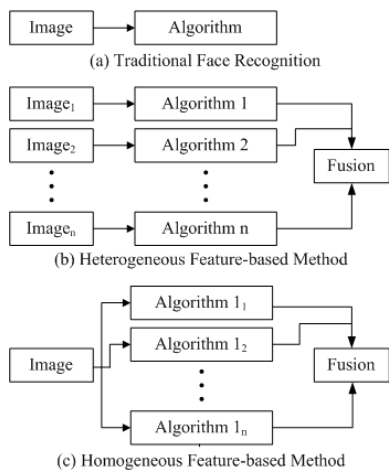
FRGC Ver 2.0에서는 얼굴 인식 성능 평가를 위한 지표로 Verification Rate (VR)과 False Acceptance Rate (FAR)을 사용하고 있다. VR은 얼굴 인식률을 나타내는 척도로 입력 영상과 선택된 등록 영상의 얼굴이 동일 인물일 때 얼굴 인식이 동일 인물로 판정할 확률을 나타내고, FAR은 두 영상이 동일 인물이 아닐 때 얼굴 인식이 동일 인물로 잘못 판단할 확률을 나타내게 된다. 그래서 어느 정도의 FAR을 허용하는지에 따라서 VR 값이 변하게 된다. FRGC의 경우 FAR값을 0.1%로 고정했을 경우의 VR값으로 성능 평가를 하게 된다. 이는 보통 FERET 평가에서 FAR값을 1%로 고정하는 것에 비해서 매우 엄격한 값을 알 수 있다. 마지막으로 평가에 사용되는 얼굴 영상의 양이 많기 때문에 영상들을 3개의 집합으로 나누어서 각각의 영상 집합에 대해 Receiver Operating Characteristic (ROC)를 구하게 된다. 이들 3개의 영상 집합을 편의상 ROC1, ROC2, ROC3라고 하면, ROC1의 경우 등록 영상과 평가 영상의 시간 차이가 학기들 안에서 촬영된 영상이고, ROC2는 1년 안이고, 마지막으로 ROC3는 학기들 사이에 촬영된 영상을 의미한다. 이를 토대로 ROC1은 영상간의 촬영 시간 차이가 적고, ROC3는 시간 차이가 큼을 알 수 있다. 하지만 실제로는 ROC3의 경우 평가에 사용되는 샘플 수가 ROC1보다 작기 때문에 실제 평가 결과를 한 가지 ROC만으로 판단하기는 힘들다.

〈표 1〉 FRGC Ver 2.0 DB 실험 프로토콜 및 영상 구성. [C]는 Controlled Image를 [U]는 Uncontrolled Image를 의미한다

Exp	Training Set	Query Set	Target Set
1	12,766장 [U/C]	16,028장 [C]	16,028장 [C]
4	12,766장 [U/C]	8,014장 [U]	16,028장 [C]

2. FRGC 向 얼굴 인식 알고리즘 리뷰

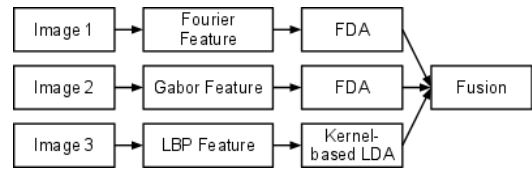
FRGC 경쟁에 참여한 알고리즘들을 <그림 2>와 같이 크게 3가지로 분류해 볼 수 있다. FRGC 초반에는 <그림 2>-(a)와 같은 전통적인 얼굴 인식 알고리즘 (Traditional Face Recognition) [2,4,6]이 많이 나왔다. 이는 한 장의 입력 얼굴 영상과 한 개의 특징만을 이용한 얼굴 인식 알고리즘이다. 하지만 이런 인식 방식은 Exp4에서 성능 개선에 한계가 있고, 이후 나온 방식은 <그림 2>-(b)와 같은 복수 입력 얼굴 영상과 복수 특징을 이용하여 복수 인식기를 만든 후에 융합하는 Heterogeneous Feature-based Method [1,3,5,7-9]가 많이 제안 되었다. 이러한 이종 특징 기반 얼굴 인식 방식은 서로 다른 특징 벡터가 상호 보완적이라는 가정하에 각자 독립적으로 학습한 뒤 이들 결과를 정합 하는 방식이다. 예를 들어서 저해상도 영상에서 Gabor 특징을 추출 하고 저해상도 영상에서 Discrete Cosine Transform (DCT)/Fourier 특징을 추출하거나 또는 Local Binary Pattern (LBP) 특징을 추출 한 뒤에 정합하는 방식을 채택하고 있다. 이런 방식은 좋은 인식 성능과 함께 조명 변화에서도 안정적인 성능을 확보할 수 있지만, 상대적으로 알고리즘의 복잡도가 증가하거나 계산 속도가 느려져 실시간 얼굴 인식기기 개발에는 사용되기 힘들다. 이에 최근에는 <그림 2>-(c)와 같이 저해상도 단일 영상에서 파라미터만 다른 동일 특징 벡터를 다수 추출 한 뒤에 독립적으로 학습하는 Homogeneous Feature-based Method [10]도 제안 되었다. 이런 방식은 안정적 성능을 확보하면서도 경량화된 알고리즘을 만들 수 있다는 이점이 있다.



<그림 2> FRGC 참여 얼굴 인식 알고리즘 분류도

가. Samsung AIT Method (2005년)

삼성종합기술원(Samsung AIT)에서는 단일 얼굴 인식기만으로는 FRGC 조명 변화 극복 및 전반적인 얼굴 인식 성능 개

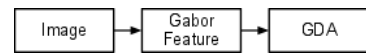


<그림 3> Samsung AIT 방식 얼굴 인식 Block-diagram

선을 유도할 수 없다고 판단하고, <그림 3>과 같은 복수 얼굴 인식이 정합 프레임 [1] 제안했다. 이는 Fourier 특징, Gabor 특징, LBP 특징 등을 독립적으로 학습하여 복수 얼굴 인식기를 만든 뒤에 이를 정합하는 방식으로, 서로 독립적인 정규화된 입력 영상을 사용하고 서로 다른 특징들을 통계적으로 융합했기 때문에 알고리즘의 복잡도는 상당히 높은 편이었다. 하지만 제안된 방식을 통해 2005년에 FRGC ver 2.0 Exp4 실험에서 약 91~92%의 높은 인식 성능을 초반에 달성할 수 있었다.

나. Liu Method (2006년)

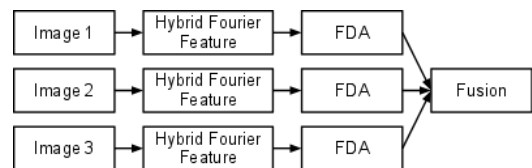
Liu [2]는 2006년도에 <그림 4>와 같은 Gabor 특징을 이용한 Kernel Fisher Analysis (KFA) 기반 얼굴 인식 프레임을 제안 했다. 이 알고리즘은 128x128 픽셀로 정규화된 얼굴 영상을 입력으로 사용하고, Gabor 필터를 적용하여 얼굴 특징을 크기(Scale) 및 각도(Orientation) 변화에 따른 얼굴 특징 벡터로 확장한다. 이렇게 확장된 특징 벡터를 커널 기반의 General Discriminant Analysis(GDA)로 학습하여 얼굴 특징의 분별도를 높이고 있다. 사용된 커널 모델은 Sigmoid 커널의 일종인 Fractional power polynomial 모델을 사용하고 있다. 이렇게 제안된 얼굴 인식 프레임은 초기 FRGC 경쟁에서 좋은 성능을 보여주었으나, 단일 인식기로는 더 좋은 성능을 내기 힘들어서 이후부터는 복수 얼굴 인식기를 정합하는 방식이 많이 나오게 된다.



<그림 4> Liu 방식 얼굴 인식 Block-diagram

다. Hwang et al. Method (2006년)

Hwang et al. [3] 은 <그림 5>와 같이 복수개의 저해상도 영상(46x56)을 이용한 Fourier 특징 기반 얼굴 인식 알고리

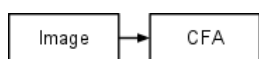


<그림 5> Hwang et al. 방식 얼굴 인식 Block-diagram

즘을 제안한다. 단일 얼굴 인식기로는 FRGC DB Ver 2.0 Exp4에서 좋은 성능을 보이지 못했고, 이를 극복하기 위해서 복수의 인식기를 생성하기 위해 입력 영상 정규화 시 두 눈 좌표 간의 거리를 바꾸어가며 복수개의 정규화 얼굴을 만들었다. 이렇게 생성된 정규화된 얼굴은 얼굴 안쪽 내부 변화에 중심을 두는 Fine 모델과 머리, 및 얼굴 라인 까지 포함하는 Coarse 모델, 그리고 중간 scale로서 위의 두 모델의 특징을 모두 포함하는 Dominant 모델로 나눌 수 있다. 각각의 얼굴 영상들을 Fourier Domain으로 변환한 뒤 주파수 특성에 따라서 저주파수, 고주파수 등 주파수를 선별적으로 선택하여 특징 벡터를 생성하고, 이 특징들은 Fisher Discriminant Analysis (FDA)를 이용 독립적으로 학습된 뒤에 가중치를 다르게 하여 정합되었다.

라. Kumar et al. Method (2006년)

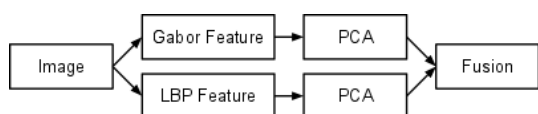
Kumar et al. ^[4]은 2006년에 Fourier 주파수 기반인 Class-dependence Feature Analysis(CFA)를 <그림 6>과 같이 제안한다. Kumar (CMU) 그룹은 과거부터 Fourier 주파수에서 Correlation 필터를 이용한 CoreFace 얼굴 인식 방법을 제안해왔다. 이전 방식은 알고리즘 복잡도가 높다는 단점이 있었으나, 새로 제안한 CFA를 이용하여 이러한 복잡도를 줄이면서 좋은 성능을 확보하게 되었다.



<그림 6> Kumar et al. 방식 얼굴 인식 Block-diagram

마. Tan and Triggs Method (2007년)

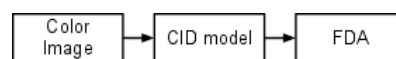
Tan과 Triggs ^[5]은 얼굴 인식 성능을 높이기 위해 다양한 특징(Heterogeneous Feature)을 추출하고 정합하는 방식을 <그림 7>과 같이 제안하였다. 이들은 이를 위해 128x128 픽셀의 고해상도 영상에서 대표적인 얼굴 인식 특징인 Gabor 특징과 Texture 인식에서 많이 사용되던 LBP 특징을 동시에 사용하였다. 각각의 특징 벡터는 Principal Component Analysis (PCA)를 통과하면서 차원을 낮춘뒤 정합하게 된다. 특히 LBP 특징의 경우 Gabor 특징과는 서로 다른 특성을 보여주어 성능 개선에 도움이 되었다.



<그림 7> Tan and Triggs 방식 얼굴 인식 Block-diagram

바. Yang and Liu Method (2008년)

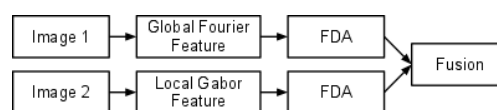
Yang과 Liu ^[6]는 <그림 8>과 같이 32x32픽셀 저해상도 영상의 컬러 정보를 활용하여 상대적으로 Exp4에서 높은 인식 성능을 2008년에 달성했다. 저자는 컬러 영상으로부터 그레이 영상을 만들 때 전통적인 가중치를 사용하지 않고 얼굴 인식 분별력을 가장 높이는 방향으로 가중치를 통계적으로 산출한 Color Image Discriminant (CID) 모델을 제안했다. 이들의 실험에 따르면 얼굴 인식에 가장 도움이 되는 컬러 정보는 적색 채널로 이는 흑인을 제외한 동양인, 백인 얼굴색의 기본이 적색이라는 점을 상기하면 타당한 실험 결과로 생각될 수 있다.



<그림 8> Yang and Liu 방식 얼굴 인식 Block-diagram

사. Su et al. Method (2009년)

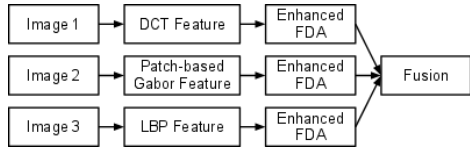
Su et al. ^[7]은 2009년에 Global 특징과 Local 특징을 결합하는 얼굴 인식 프레임워크를 <그림 9>와 같이 제안했다. Global 특징과 Local 특징을 한 계층으로 묶음으로써 양 특징들이 서로 상호 보완적인 관계가 되도록 했다. Global 특징으로는 Fourier 주파수 기반 특징을 사용하고, Local 특징 추출을 위해 얼굴의 중요 블록을 Greedy-search 알고리즘으로 찾아낸 뒤 선택된 블록에서만 Gabor 특징을 추출했다. 이렇게 추출된 약 30개의 특징들은 독립적으로 FDA에 투영되고 마지막에 인식 성능을 높이기 위해서 정합되어 조명 변화에 좋은 인식 성능을 보여주게 되었다.



<그림 9> Su et al. 방식 얼굴 인식 Block-diagram

아. Liu and Liu Method (2009년)

Liu와 Liu ^[8]는 2009년에 <그림 10>과 같이 서로 다른 특징을 사용하였고 각 특징에 특화된 새로운 컬러 영역을 제안했다. 이로써 기존의 RGB 컬러 영상과 다르게 각각 서로 다른 해상도에 얼굴 인식 성능을 높이는 방향으로 컬러 영역을 생성할 수 있었다. 특히, DCT 특징, Patch 기반 Gabor 특징, LBP 특징 등 각 특징에 맞는 컬러 영역을 만들어 독립적으로 서로 다른 특징을 추출하였고, 이렇게 추출된 특징은 개선된 FDA를 통해 얼굴 분별도가 강화되는 방향으로 투사된 후 정합되었다. FRGC Ver 2.0 Exp4에서 92%를 달성하는 좋은 성

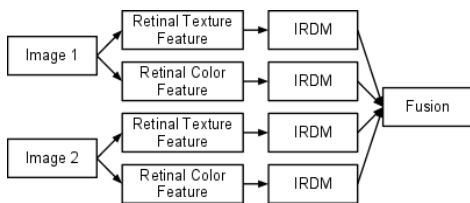


〈그림 10〉 Liu and Liu 방식 얼굴 인식 Block-diagram

능을 보여 주고 있지만, 상대적으로 128x128의 높은 해상도에서 많은 수의 Gabor 특징을 추출하고 서로 완전히 다른 특징들을 추출함으로써 알고리즘의 높은 복잡도 문제 및 재현이 힘들다는 단점은 해결하지 못했다.

자. Deng et al. Method (2010년)

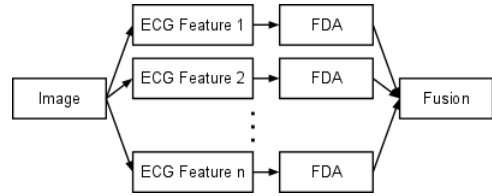
Deng et al.^[9]은 2010년에 신경 생물학과 심리학적인 관점에서 출발한 새로운 특징인 Retinal Texture Feature(RTF)와 Retinal Color Feature(RCF)를 〈그림 11〉과 같이 제안했다. RTF는 Retina 부터 Visual Cortex에 이르는 시각 정보 처리 방법에서 영감을 얻은 방식으로 정규화된 얼굴 영상에서 조명 변화를 둔감하게 만들고 Gabor 특징을 추출하게 된다. 또한 RCF는 컬러 정보를 얼굴 인식에 가장 좋은 방향으로 조합^[6]하여 새로운 컬러 정보를 만들어 내고 있다. 이렇게 생성된 특징은 FDA를 개선한 Incremental Robust Discriminant Model(IRDM)을 이용하여 최적의 특징벡터로 변환이 된다. 마지막으로 순수 얼굴만 포함한 128x128 픽셀의 정규화 영상과 전체 얼굴을 포함한 128x104 픽셀의 정규화 영상으로 나누어 얼굴의 내부/외부 정보^[3]를 모두 인식에 사용하였다. 이런 다양한 특징들을 모두 정합하여 FRGC Ver 2.0 Exp4에서 93%의 가장 높은 성능을 달성하고 있다.



〈그림 11〉 Deng et al. 방식 얼굴 인식 Block-diagram

차. Hwang et al. Method (2011년)

앞서 살펴본 것과 같이 2004년부터 지금까지 많은 학교 및 연구소의 연구를 통해 FRGC Ver 2.0 조명 변화(EXP4)에서 얼굴 인식 성능은 지속적으로 개선되어 오고 있다. 하지만 얼굴 검출 기술이 Consumer Electronics (CE) 제품 적용에 성공한 것에 비해 얼굴 인식 알고리즘은 널리 퍼지지 못하고 있다. 이는 안정적인 성능을 보여주는 얼굴 인식 알고리즘들이 상대적으로 높은 해상도(128x128 픽셀)를 사용하고 성능 향상을 위해서 많은 특징들을 추출하기 때문에 계산량이 지

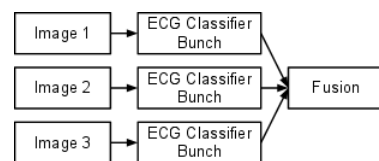


〈그림 12〉 Hwang et al. 방식 얼굴 인식 Block-diagram

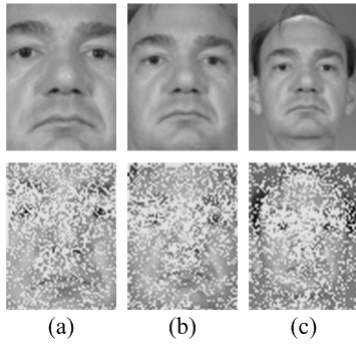
나치게 높아졌기 때문이다. 이에 2011년도에 Hwang et al.^[10]은 60x80 픽셀의 저해상도 얼굴 영상만을 사용한 경량화된 얼굴 인식 방법을 〈그림 12〉와 같이 제안했다. 이들은 Gabor 특징을 얼굴의 곡률 특징을 잘 파악할 수 있는 Curvature Gabor 특징으로 확장하고 이렇게 추출된 특징을 Boosting Theory를 사용하여 가장 좋은 특징들만을 선별하고 있다. 이렇게 선별된 특징들은 인식 성능 향상을 위해서 FDA에 투영된다. 이러한 단일 인식기를 Extended Curvature Gabor (ECG) Classifier로 명명하고 Gabor 커널 파라미터를 다르게 하여 복수의 인식기를 생성한 뒤에 정합 (ECG Classifier Bunch) 함으로써 인식기의 성능을 높이고 있다.

카. 복수 얼굴 모델 기반 ECG Classifier Bunch

본 논문에서는 [3]에서 제안된 복수 얼굴 모델을 [10]에서 제안된 ECG Classifier Bunch에 적용하여 FRGC Ver 2.0 DB에서 성능을 확인해 보고자 한다. 〈그림 13〉을 보면 제안된 방식을 쉽게 이해 할 수 있다. 입력 영상으로 60x80의 저해상도 영상을 총 3개의 Fine, Dominant, Coarse 얼굴 모델로 생성을 한 뒤에 ECG Classifier Bunch를 각각 적용하여 〈그림 14〉와 같은 Gabor 특징을 총 2400개를 Boosting 기법을 이용 선별 한다. 〈그림 14〉에서 볼 수 있는 것처럼 대부분의 특징은 눈/눈썹 및 코 주변에서 추출되는 것을 알 수 있다. 이런 실험 결과로 정면 얼굴 인식에 중요한 Gabor 특징은 눈 주변 및 코 주변임을 추정할 수 있다. 이후 추출된 Gabor 특징 벡터는 Gabor 파라미터에 따라서 여러 조합^[10]을 만들 수 있는데 본 논문에서는 앞선 논문과 다르게 4개의 인식기로 구성된 “I-Shaped ECG Classifier Bunch”를 사용하였다. 마지막으로 이들 결과를 합치는 방식은 [10]에서 제안된 LLR 기반 정합법을 이용하였다.



〈그림 13〉 복수 얼굴 모델 기반 ECG Classifier Bunch 방식 얼굴 인식 Block-diagram



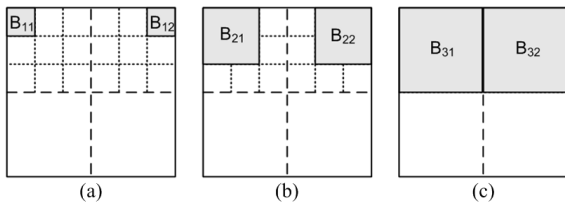
〈그림 14〉 (a) Fine 얼굴 모델, (b) Dominant 얼굴 모델, (c) Coarse 얼굴 모델. 각각의 얼굴 모델에서 총 2400개의 중요 Gabor 특징을 선별 했다

3. 얼굴 인식용 주요 특징 추출 알고리즘

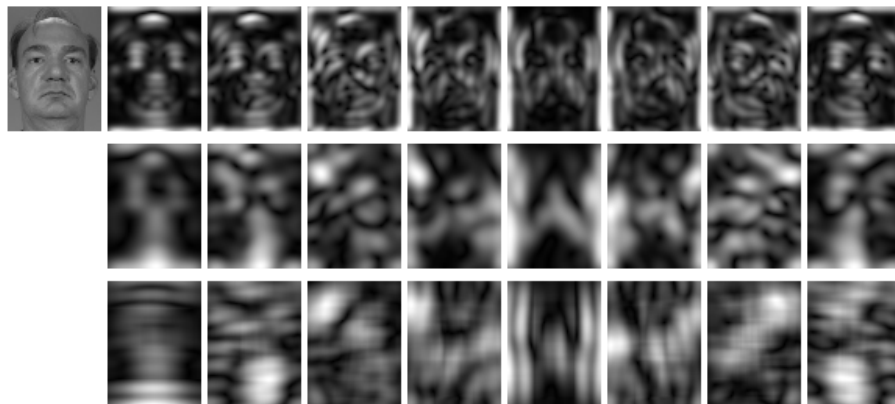
앞서 살펴본 얼굴 인식 기술들의 주요 특징 추출 알고리즘을 살펴보면 다음과 같다.

가. 2D Discrete Fourier Transform 특징 추출 기법

Fourier 변환은 동영상 압축과 같은 영상 처리 분야에서 널리 사용되고 있었다. 얼굴 인식에서 Fourier 변환의 적용은 주파수 대역별로 얼굴 특징을 분리하기 위해서 사용되었다. 예를 들어서 저주파수 대역의 특징만을 사용하고 싶을 경우 〈그림 15〉-(a)와 같이 저주파 영역만을 선택하고, 더 많은 주파수 대역을 특징으로 포함하고 싶을 경우는 〈그림 15〉-(b)



〈그림 15〉 Fourier 영역에서의 주파수 대역에 따른 얼굴 특징 선별 예



〈그림 17〉 Gabor 커널을 얼굴 영상에 적용한 뒤 나오는 Magnitude 값의 예. Gabor 커널 파라미터가 $\nu=3$ 일 때 왼쪽부터 $\sigma = \{0.5\pi, \pi, 2\pi\}$ 일 경우 나타나는 Magnitude 값의 예들이다

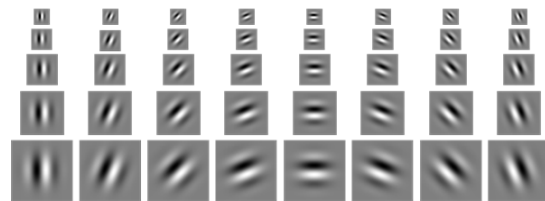
와 같은 영역만 선택하게 된다. 마지막으로 〈그림 15〉-(c)의 경우 전 주파수 대역을 나타내고 이는 즉 영상 자체를 의미하게 된다. 이러한 Fourier 변환을 이용한 특징 추출의 경우 저해상도 영상을 입력으로 사용하는 것이 유리하고, 주로 얼굴의 Global 특징을 대표하는 경우가 많다.

나. Gabor 특징 추출 기법

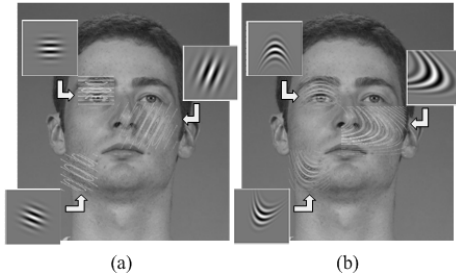
Gabor 커널은 얼굴 인식 성능 개선에 가장 많은 기여를 한 방법으로 〈그림 16〉과 같이 입력 영상을 5개의 크기 변화와 8개의 각도 변화로 분석해 주어 입력 영상 한 개에서 다양한 특징을 추출할 수 있다. 이러한 Gabor 커널의 간단한 수식은 다음과 같다.

$$\psi(x, y; \nu, \mu) = \frac{k_\nu^2}{\sigma^2} \exp\left(-\frac{k_\nu^2(x^2 + y^2)}{2\sigma^2}\right) \left[\exp(ik_\nu(x\cos\phi_\mu + y\sin\phi_\mu)) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (1)$$

이때 $k_\nu = 2^{-\frac{\nu+2}{2}}\pi$ 와 $\phi_\mu = \mu\frac{\pi}{8}$ 를 사용할 수 있고, 5개의 크기 변화 $\nu = \{0, \dots, 4\}$ 와 8개의 각도 변화 $\mu = \{0, \dots, 7\}$ 를 이용할 수 있다. Gaussian 크기는 보통 $\sigma = 2\pi$ 를 사용하는 데, 얼굴 인식 특징 확장을 위해서 〈그림 17〉과 같이 다른 값을 사용하기도 한다.



〈그림 16〉 Gabor 커널 중 Imaginary 파트 예. 사용된 Gabor 커널의 파라미터는 $\sigma = \pi$ 이고 5개의 크기 변화 와 8개의 각도 변화가 있다

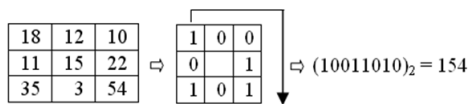


〈그림 18〉 곡률 얼굴 특징을 분석하기 위한 Curvature Gabor 커널의 특징. (a)는 일반 Gabor 커널, (b)는 Curvature Gabor 커널을 적용한 예를 보여준다

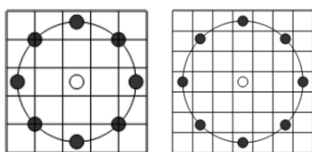
최근에는 얼굴의 곡률 특징을 추가적으로 추출하기 위해서 Curvature Gabor 특징^[10]을 활용하기도 한다. 관련된 예는 〈그림 18〉에 나타나 있다.

다. LBP 특징 추출 기법

LBP 알고리즘은 Texture Classification에서 많이 사용되던 알고리즘으로 얼굴 인식에는 인식 성능을 향상시키기 위한 부가 정보가 많이 사용이 되었다. 〈그림 19〉에서 볼 수 있듯이 중앙값을 기준으로 이보다 큰 경우는 1로 작은 값은 0으로 이진화를 한 뒤에 특정 방향으로 이진화된 값들을 정렬한다. 이후 이진화된 값을 십진수로 바꾸면 한 개의 특징 값이 나온다. 또한 〈그림 20〉에서 볼 수 있듯이 주변 값들의 거리를 다양하게 함으로써 중앙값과 주변 값의 관계를 다양하게 표현할 수 있다. 결국 LBP 특징은 중앙값과 주변 값들의 상대적인 값 차이로 특징을 표현하기 때문에 조명 변화에 강인한 성질을 보여준다. 그래서 LBP 특징은 얼굴 검출에서 많이 사용되고 있고 이전의 Haar 형 특징 보다 더 적은 수의 특징을 쓰면서 더 높은 얼굴 검출 성능을 보여주고 있다. 하지만 얼굴 인식에서의 LBP 특징은 Gabor 특징보다 현격하게 인식 성능이 떨어진다. 하지만 Gabor 특징과 LBP 특징을 서로 융



〈그림 19〉 LBP 개념도. 중앙값 15를 기준으로 주변 값들을 이진화 한 뒤, 이 값들을 다시 십진수로 바꾼다



〈그림 20〉 중앙값을 기준으로 주변 값들의 영역을 넓힌 예

합했을 경우 FRGC Ver 2.0 Exp4에서 좋은 성능 개선을 이끌어 내고 있다.

III. 실험

2004년 FRGC DB 공개 후 본문에서 소개된 다양한 알고리즘들이 얼굴 인식 성능을 높이기 위해 제안되었다. 소개된 얼굴 인식 알고리즘의 FRGC DB의 Exp1과 Exp4에서의 성능은 〈표 2〉와 〈표 3〉에 정리되었다. 이 표들은 각각의 논문에서 나온 수치를 적은 것으로, 논문에 해당 수치를 그래프로 표현해서 정확한 수치를 알 수 없거나, 값을 적지 않았을 때는 빈 칸으로 남겨 두었다. 표에 나타난 Baseline 알고리즘은 FRGC 배포 때 포함되어 있던 PCA 기반 얼굴 인식(EigenFace)이다. Baseline 알고리즘은 Exp1에서 약 66% 인식률 (FAR=0.1%)을 보였고, Exp4에서는 12~13% 인식률을 보였다. Exp1과 Exp4에서 약 50% 인식률 차이가 나는 것으로부터 우리는 Control된 환경 보다 Uncontrol된 환경이 영상 정보를 기반으로 하는 얼굴 인식에 어렵다는 사실을 다시 확인할 수 있다. FRGC DB 배포 후 나온 대부분의 얼굴 인식 알고리즘들은 Exp4의 성능을 개선하기 위한 방향으로 발전했다. 얼굴 인식 방법의 발전 경향은 앞서 기술한 것과 같이 단일 인식기 (Traditional Method)에서 복수 특징 기반 인식기 (Heterogeneous Feature-based Method)를 정합하는 형태로 바뀌었다. 즉, 알고리즘의 경량화보다는 성능 향상 위주의 복잡한 알고리즘들이 주류를 이루고 있다. 또한 과거에 Gray 영상을 이용하던 얼굴 인식 방법에서 탈피하여 컬러 영상을

〈표 2〉 FRGC Ver 2.0 DB Exp1의 성능 비교 테이블

Method	Image Size	EXP1 (VR at FAR=0.1%)		
		ROC1	ROC2	ROC3
Baseline (2004)		-	-	66%
Liu's Method (2006)	128x128	-	-	92%
Hwang's Method (2006)	46x56 (3)	95.01%	93.29%	91.51%
Kumar's Method (2006)	128x128	-	-	-
Tan's Method (2007)	128x128	-	-	-
Yang's Method (2008)	32x32	-	-	-
Su's Method (2009)	64x80/ 128x160	-	-	98%
Liu's Method (2009)	64x64/ 128x128	-	-	-
Deng's Method (2010)	128x128/ 128x104	-	-	-
Hwang's Method (2011)	60x80	98.60%	98.05%	97.49%
Proposed Method	60x80 (3)	99.02%	98.65%	98.28%

〈표 3〉 FRGC Ver 2.0 DB Exp4의 성능 비교 테이블

Method	Image Size	EXP4 (VR at FAR=0.1%)		
		ROC1	ROC2	ROC3
Baseline (2004)		13%	13%	12%
Liu's Method (2006)	128x128	-	-	76%
Hwang's Method (2006)	46x56 (3)	75.70%	75.06%	74.33%
Kumar's Method (2006)	128x128	-	-	~87.5%
Tan's Method (2007)	128x128	-	-	83.6%
Yang's Method (2008)	32x32	78.90%	78.66%	78.26%
Su's Method (2009)	64x80/ 128x160	86.9%	87.5%	88.8%
Liu's Method (2009)	64x64/ 128x128	-	-	92.43%
Deng's Method (2010)	128x128/ 128x104	93.91%	93.55%	93.12%
Hwang's Method (2011)	60x80	90.96%	90.69%	90.36%
Proposed Method	60x80 (3)	93.02%	92.91%	92.77%

적극적으로 얼굴 인식에 활용하고 있는 것 또한 주목할 만한 점이다. 과거에는 조명 변화에 따른 영향을 줄이기 위해 얼굴 인식에서는 일반적인 Gray 영상만을 사용했다. 하지만 실제로 얼굴 인식을 위해서는 일반적인 Gray 영상 변환보다는 적색을 강조하면서 한 채널로 바꾸거나 또는 얼굴 인식용 신규 컬러 채널을 사용하는 것이 Exp4에서 더 높은 성능을 보여주고 있다. 마지막으로 얼굴 인식을 위한 가장 강력한 특징은 Gabor 특징인 것을 다시 확인할 수 있다. Exp1과 Exp4에서 90% 인식률을 확보한 대부분의 얼굴 인식 알고리즘이 Gabor 특징을 기본으로 포함한다는 사실은 그 좋은 예이다. 실제 표의 수치를 보면 Exp1의 경우 Exp4보다 성능 개선 효과가 적은데, 이는 Baseline에서도 Exp1은 이미 66%를 달성했기 때문에 상대적으로 개선 효과가 적어 보였기 때문이다.

〈표 3〉의 수치를 자세히 살펴보면 Exp4의 성능의 경우 2006~2008년까지의 알고리즘은 단일 특징 기반의 알고리즘이 많았고 이들은 대부분 70~80% 인식 성능을 보여주고 있다. 2007년부터 최근까지의 알고리즘은 대부분 80~90%의 높은 성능을 보여주고 있는데, 이들은 앞선 알고리즘과는 다르게 복수 특징 기반의 얼굴 인식 알고리즘들인 것을 알 수 있다. 특히 가장 높은 성능을 보여주는 알고리즘은 Deng et al.^[9]이 제안한 방식으로 Exp4에서 약 93%의 가장 인식 성능을 보여주고 있다. 하지만 이들이 제안한 방식은 여전히 128x128의 고해상도 입력 얼굴 영상을 이용하고 대부분의 해상도에서 Gabor 특징을 뽑아내기 때문에 필연적으로 알고리즘의 복잡도가 높아진다. 최근에는 Exp4에서 90~92%를 보장하면서 입력 해상도 및 알고리즘을 경량화하는 연구도 활발히 진행 중이다.

IV. 결론

본 논문에서는 FRGC Ver 2.0 DB를 이용한 다양한 얼굴 인식 알고리즘을 비교 분석하였다. 대부분의 FRGC 얼굴 인식 논문은 조명 변화가 심한 Exp4 실험 프로토콜에서의 얼굴 인식 성능 개선에 중점을 두고 있었다. 최근 들어 Exp4에서 FAR=0.1%일 때 얼굴 인식 성능이 90% 이상 나오는 알고리즘들이 소개가 되고 있고, 이제 Exp1과 Exp4의 얼굴 인식 성능 갭은 2004년도의 50% 이상에서 5~6% 내로 떨어진 상태이다. 이러한 성능 개선의 가장 큰 이유는 Heterogeneous Feature-based Method와 같은 복수 얼굴 인식이 정합 방법 제안되었기 때문이다. 이러한 성능 개선에도 불구하고 산업 전반에 얼굴 인식이 적용되기 위해서는 알고리즘의 경량화가 필요하고 이에 대한 연구도 진행 중이다.

참고문헌

- [1] Samsung AIT, "Still Face Recognition 2D Performance," 2005 [Online]. Available: <http://www.nist.gov/itl/iad/ig/frgc-presentations.cfm>
- [2] C. Liu., "Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.28, No.5, pp.725-737, May, 2006.
- [3] W. Hwang, G. Park, J. Lee, and S. Kee, "Multiple face model of hybrid Fourier feature for large face image set," in Proc. IEEE Comput. Vis. Pattern Recognit., Jun., 2006, Vol.2, pp.1574-1581.
- [4] V. Kumar, M. Sawides, and C. Xie, "Correlation pattern recognition for face recognition," Proceedings of the IEEE, Vol.94, No.11, pp.1963-1976, 2006.
- [5] X. Tan and B. Triggs, "Fusing Gabor and LBP feature set for kernel-based face recognition," IEEE International Workshop on Analysis and Modeling of Face and Gestures, pp.235-249, 2007.
- [6] J. Yang and C. Liu, "Color image discriminant models and algorithms for face recognition," IEEE Trans. on Neural networks, Vol.19, No.12, pp.2088-2098, 2008.
- [7] Y. Su, S. Shan, X. Chen, and W. Gao, "Hierarchical ensemble of global and local classifiers for face recognition," IEEE Trans. on Image Processing, Vol.18, No.8, pp.1885-1896, Aug., 2009.
- [8] Z. Liu and C. Liu, "Robust face recognition using color information," Advances in Biometrics, Vol.5558/2009, pp.122-131, 2009.



- [9] W. Deng, J. Hu, J. Guo, W. Cai, and D. Feng, "Emulating biological strategies for uncontrolled face recognition," Pattern Recognition, Vol.43, No.6, pp.2210-2223, Jun., 2010.
- [10] W. Hwang, X. Huang, K. Roh, and J. Kim, "Face recognition system using extended curvature Gabor classifier bunch for low-resolution face image," IEEE Workshop on Biometrics, Colorado Springs, USA, Jun., 2011.



황 원 준

1999년 2월 고려대학교 전자공학과 학사.
 2001년 2월 고려대학교 전자공학과 석사.
 2001년 2월~2005년 2월 삼성종합기술원 연구원.
 2005년 3월~2008년 7월 삼성종합기술원 전문연구원.
 2008년 7월~2011년 10월 삼성전자 책임연구원.
 2011년 10월~현재 삼성종합기술원 전문연구원.
 <관심분야> 얼굴 인식, 컴퓨터 비전, 메디컬 영상 처리



김 준 모

1998년 8월 서울대학교 전기공학부 학사.
 2000년 8월 MIT EECS 석사.
 2005년 2월 MIT EECS 박사.
 2005년 5월~2009년 6월 삼성종합기술원 전문연구원.
 2009년 7월~ 현재 카이스트 전기 및 전자공학과 조교수.
 <관심분야> 확률 신호처리, 영상처리, 컴퓨터 비전, 정보 이론