

정규화된 웨이블릿 특징에 의한 얼굴 인식 (Face Recognition with Normalized Wavelet Features)

이 찬 호 [†] 박 주 철 [†] 최 형 일 ^{**}

(Chan-Ho Lee)(Ju-Chul Park)(Hyung-Il Choi)

요 약 본 논문에서는 정규화된 웨이블릿 방법에 기반한 얼굴 인식 방법을 제안한다. 추출된 얼굴 영역의 크기를 정규화하고, 배경 영역을 제거하기 위해 이진화된 가우시안 윈도우를 사용하였으며, 또한 조명의 영향을 줄이기 위해 얼굴 영역의 히스토그램을 상세화하고, 왼쪽 부분의 밝기와 오른쪽 부분의 밝기를 평균하였다. 정규화된 얼굴 영역은 극좌표계로 표현하여 사각형의 형태를 가지게 하였다. 특징으로는 가보 웨이블릿 계수를 사용한다. 가보 웨이블릿 변환은 매개 변수와 정규화된 얼굴 영역의 해상도를 바꾸어 가며 여러 차례 적용하였다. 인식에 유용한 계수들을 선택하기 위해 FD 분석을 수행하였다. 선택된 특징들은 FD 값과 함께 인식에 사용되었다. 실험 결과를 보면 제안된 방법이 매우 유망하다는 것을 알 수 있다.

Abstract This paper describes the normalized wavelet based approach for face recognition. We normalize the size of an extracted facial area and use the binarized Gaussian window to remove background areas. We also reduce illumination effects by specifying the histogram of the facial area and balancing the brightness of the left part with that of the right part of the facial area. The normalized facial area is represented in polar coordinates so that the area has the form of a rectangle. As for features, we use Gabor wavelet coefficients. The Gabor wavelet transformation is applied repeatedly by varying its parameters and the resolution of the normalized facial area. The FD(Fisher's Discriminant) analysis is applied to select useful coefficients for recognition. The resulting features together with their FD values are finally used for recognition. The experimental results show that our approach is very promising.

1. 서 론

컴퓨터, 정보 통신의 발전과 함께 정보의 가치도 더욱 커지고 있으며 이에 따라 정보의 보호가 큰 문제로 부각되고 있다. 정보를 보호하는 방법 중 가장 일반적인 방법은 패스워드(password)나 열쇠, ID 카드 등을 사용하는 것이다. 그러나 이 방법들은 분실 또는 복제의 위험이 매우 크다. 따라서, 이러한 분실, 복제 문제를 해결하기 위해 신체 일부의 특성을 이용해 신원을 확인하는

방법이 있는데, 지문, 얼굴 생김새, 홍채 또는 망막의 특징, 손등의 정맥 모양 등을 이용한 방법들과 목소리, 서명, 타이핑 습관 등의 행동학적 특징을 이용한 방법 등이 포함된다. 이 방법들은 개인의 신체적 특징이 일치하는 사람은 없고 분실, 복제의 위험이 적다는 특징을 가지고 있다.

이러한 방법들 중의 하나인 얼굴 인식 방법은 영상을 촬영하는 과정에 있어서 홍채나 정맥을 인식하는 방법에 비해 상대적으로 거부감이 적다. 홍채나 정맥을 인식하는 방법은 영상을 촬영하기 위해 눈이나 손을 입력 장치에 가까이 가져가거나 직접 접촉해야 하기 때문에 거부감을 가질 수 있지만, 얼굴 인식 방법의 경우, 일반적으로 사람들은 얼굴 사진 찍는 일에 대해 익숙하기 때문이다. 또, 적외선 카메라와 같은 특수한 입력 장치들을 필요로 하는 홍채 인식이나 정맥인식에 비해 얼굴 인식은 일반적인 CCD 카메라나 디지털 카메라만으로 구현이 가능하다.

· 본 논문은 두뇌한국 21 사업 중 핵심분야 사업(E-0075)과 첨단정보 기술연구센터를 통하여 과학재단의 일부 지원을 받았음

[†] 비 회 원 : 송실대학교 컴퓨터학과
chlee@vision.soongsil.ac.kr
jcpark@vision.soongsil.ac.kr

^{**} 종 신 회 원 : 송실대학교 컴퓨터학부 교수
hic@computing.soongsil.ac.kr

논문접수 : 1999년 10월 20일

심사완료 : 2000년 7월 25일

이러한 기존의 얼굴 인식 방법은 크게 통계적인 분석 방법, 구조적인 분석 방법 그리고 신경망을 이용한 방법으로 나눌 수 있다. 통계적인 분석 방법은 얼굴 영역에서 얼굴 인식에 적합한 특징들을 추출하여 인식에 사용하는 방법으로 Karhunen-Loeve(KL) 변환에 기반한 통계적인 방법[1,2,3]들과 가보 웨이블릿(Gabor wavelet)을 사용하여 얼굴의 특징을 추출하고 그래프 매칭 기법을 이용한 방법[4,5]들이 있으며, 구조적인 분석 방법으로는 얼굴의 구조와 얼굴 구성 요소들의 특징과 관계를 이용한 인식 방법인 기하적인 얼굴 모델에 의한 얼굴 특징 추출 방법[6]과 얼굴의 기하학적인 특징을 이용한 방법[7] 등이 있다. 그리고 신경망을 이용한 방법들로는 Brunelli[8]와 Lawrence[9]가 제안한 방법 등이 있다.

본 논문에서는 가보 웨이블릿을 사용하여 텍스처(texture) 정보를 추출하고 이를 얼굴 인식을 위한 특징으로 사용하는 얼굴 인식 방법을 제안한다. 본 논문은 크게 얼굴 추출 단계와 특징 추출 및 인식 단계로 나눌 수 있는데, 그림 1은 전체적인 구조를 보인다. 먼저 얼굴 추출 단계에서는 사전에 피부 색상에 대한 학습을 수행하여 피부 색상 모형을 구축하고, 이를 참조하여 얼굴 후보 영역을 선택한다. 얼굴 후보 영역이 결정되면 이진화와 수평, 수직 투영을 통해 두 눈을 추출하여 두 눈의 위치를 기준으로 얼굴 영역을 결정하였다. 얼굴 영상을 정규화하는 과정에서는 대부분의 얼굴 인식 방법에서 문제가 되고 있는 조명의 방향과 세기에 따른 영상의 밝기 변화를 줄이기 위해 히스토그램 상세화와 좌우 평균 얼굴을 사용하였으며, 히스토그램 상세화를 사용하여 얼굴 영상의 전체적인 밝기를 정규화하였다. 또한, 사람의 얼굴은 대부분 거의 좌우 대칭을 이루고 있는 점을 이용하여 영상의 중심을 기준으로 좌우 평균 얼굴을 만들어냄으로써 얼굴의 전체적인 모양과 특징은 변하지 않지만 조명의 방향에 따른 변화를 어느 정도 줄일 수 있는 장점을 가지도록 하였다. 특징 추출 단계에서는 가보 웨이블릿을 적용하여 계산된 가보 계수

를 특징으로 사용한다. 가보 웨이블릿은 텍스처 정보를 잘 나타낼 수 있는 방법으로 이미 여러 분야에서 텍스처 정보를 추출하기 위해 다양한 방법으로 사용되고 있다. 본 논문에서는 이 가보 웨이블릿의 가보 계수의 구분력을 높이기 위해 FD 분석을 수행하여 구분력이 큰 특징들을 선택하여 인식에 사용함으로써 인식 성능을 높이고 각각의 특징들에게 가중치를 부여함으로써 인식 정확도를 높이는 방법을 제안한다.

2. 얼굴 영역의 추출 및 정규화

2.1 얼굴 영역의 추출

CCD 카메라로 촬영된 컬러 영상에서 얼굴 영역을 찾기 위한 탐색 공간을 줄이기 위해 사전 학습을 통해 피부 색상 모형[11]을 정의하고 이를 사용하여 얼굴 후보 영역을 결정한다. 피부 색상 모형은 RGB 컬러 모형이 아닌 YIQ 컬러 모형을 사용하여 정의한다. 일반적으로 컬러 영상을 나타내기 위해 RGB 컬러 모형을 많이 사용하지만, RGB 컬러 모형은 영상을 처리할 때 3가지의 RGB 컬러 값을 모두 처리해야 하기 때문에 비효율적일 뿐 아니라, 영상의 밝기 변환에 민감하게 작용하기 때문에 동일한 영상의 동일한 영역에 대해서도 조명의 세기나 방향에 따라 컬러 값의 차이가 크게 발생할 수 있다. 반면 YIQ 컬러 모형은 밝기를 나타내는 Y 성분과 색상 정보를 가지는 I 성분과 Q 성분으로 나누어져 있어 영상 처리에서 이점을 가지고 있다[10]. 따라서 I 성분과 Q 성분을 이용하면 밝기에 독립적인 피부 색상 모형을 정의할 수 있다. 피부 색상 모형은 YIQ 컬러 모형의 I 성분과 Q 성분의 평균과 표준편차를 이용하여 정의한다.

얼굴 후보 영역의 결정은 컬러 영상의 각 화소와 피부 색상 모형 간의 마할라노비스 거리(Mahalanobis distance)[12]를 계산하여 이루어진다. 마할라노비스 거리는 평균과 표준 편차를 고려한 것으로 기하학적인 거리(Euclidean distance)가 가지는 단점을 극복할 수 있다. 피부 색상 모형은 YIQ 컬러 모형에서 색상값을 가지고 있는 I 성분과 Q 성분의 두 가지 값으로 이루어져 있으므로 거리를 측정하는 방법에 있어 기하학적인 거리보다는 마할라노비스 거리가 보다 적절한 방법이다. 따라서 피부 색상 모형의 평균에서 분산이 ± 3 의 범위 내의 색상을 피부 색상으로 간주한다. 즉, 각 화소와 피부 색상 모형 간의 마할라노비스 거리가 3 이내의 화소를 피부 색상을 가지는 화소로 결정한다.

피부 색상 모형에 의해 피부 색상에 속하는 화소들이 결정되면 레이블링(labeling)을 수행한다. 정면 얼굴 영

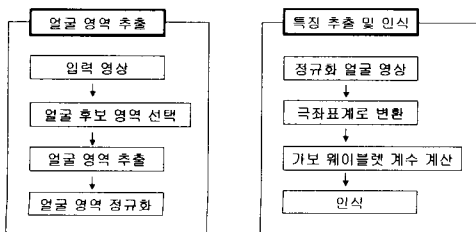


그림 1 전체적인 구조

상에서 가장 큰 영역을 차지하고 있는 것이 얼굴 영역
이므로 피부 색상을 가진 영역들 중 가장 큰 영역을 얼굴
후보 영역으로 결정한다. 그림 2는 정면 얼굴 영상과
얼굴 후보 영역 영상이다.

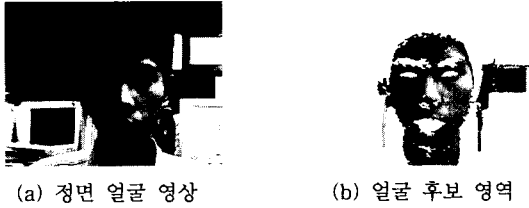


그림 2 입력 영상과 얼굴 후보 영상

얼굴 후보 영역이 결정되면 눈의 위치를 기준으로 얼굴
영역을 결정한다. 얼굴 후보 영역은 얼굴 영역 뿐만
아니라 피부 색상과 동일한 색상을 가지고 있는 귀와
목을 포함하고 있으며, 심지어 피부 색상과 유사한 배경
영역을 포함하고 있기도 하다. 따라서 귀와 목, 배경 영
역을 배제한 정확한 얼굴 영역을 찾아내기 위해서는 얼굴
의 중요 요소인 눈의 위치를 찾아내어 이를 기준으로
얼굴 영역을 추출해야 한다. 두 눈의 위치를 찾아내기
위해 얼굴 후보 영역 내부의 구멍(hole)을 포함한 얼굴
후보 영역 내의 모든 화소에 대해 YIQ 컬러 모형의 밝기
를 나타내는 Y 성분에 대해 이진화를 수행하고, 이진
화의 결과로 눈썹, 눈, 코, 입에 해당하는 화소들을 얻게
되면 이 화소들에 대해 수평, 수직 투영을 하여 두 눈의
위치를 결정한다. 우선 수평 투영을 하여 두 번째 나타
나는 봉우리(peak)를 눈 부분으로 보고, 눈 부분에 대해
다시 수직 투영을 수행하여 두 눈의 위치를 결정한다.

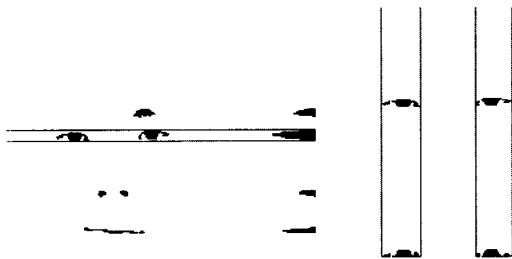
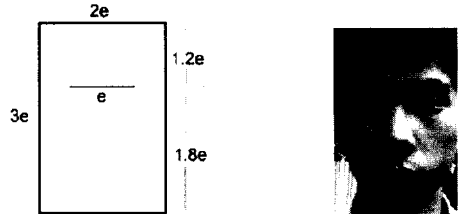


그림 3 얼굴 후보 영역 이진화 영상의 수평, 수직 투영

두 눈의 위치가 결정되면 두 눈의 위치와 눈 사이의
거리를 기준으로 얼굴 영역을 결정한다. 코나 입 같은
다른 얼굴 요소를 사용하지 않고 두 눈을 사용하는 것
은 다른 얼굴 요소와는 달리 두 개가 존재하기 때문에

눈 사이의 거리를 이용하면 대략적인 얼굴 영역을 알아
낼 수 있기 때문이다. 두 눈 사이의 거리를 e 라고 했을
때 얼굴의 폭은 $2e$ 로 하고 얼굴의 길이는 $3e$ 로 하여
직사각형 형태의 얼굴 영역을 결정한다. 그림 4는 얼굴
영역의 비율과 추출된 얼굴 영역을 보인다.



(a) 눈 사이의 거리 e 와 얼굴 (b) 추출된 얼굴 영상
영역의 비율

그림 4 얼굴 영상의 추출

2.3 얼굴 영상의 정규화

얼굴 영역이 추출되면 얼굴 영상의 크기를 정규화 한
다. 추출된 얼굴 영상은 눈 사이의 거리에 따라 그 크기
가 모두 다르기 때문에 일정한 크기의 영상으로 정규화
할 필요가 있다. 먼저 YIQ 컬러 모형의 밝기를 나타내는
Y 성분을 참조하여 추출된 얼굴 영상을 256 명암 단계
를 가지는 흑백 영상(gray image)으로 변환하고 영상의
크기를 정규화 한다. 추출된 얼굴 영상을 양선형 보간법
(bilinear interpolation)[13]을 사용하여 $n \times m$ 의 크기를
가지는 영상으로 정규화한다. 본 논문에서는 실험에 적
절한 크기를 가지는 100×150 의 크기를 가지는 영상으
로 정규화하였다.

정규화된 얼굴 영상 내부에는 얼굴 뿐 아니라 머리카
락, 귀, 목 옆 부분의 배경 영역 등이
포함되어 있다. 머리카락이나 귀는 각 개인의 고유 특성
이라기 보다는 배경이라고 볼 수 있는데, 머리카락의 경
우 머리 모양에 따라 수시로 변화하며, 유사한 머리 모
양을 가진 사람이 있을 수 있기 때문이다. 이러한 배경
영역들을 제거하기 위해 이진화된 가우시안 윈도우
(binarized Gaussian window)를 사용한다. 일반적인
가우시안 윈도우는 영상의 중앙에서 1의 값을 가지며
외곽 부분으로 갈수록 값이 작아지기 때문에 정확히 가
우시안 윈도우의 중앙에 위치한 화소만이 자기 자신의
밝기 값을 가지게 되고 나머지 화소들은 자기 자신의
값보다 작은 값을 가지게 된다. 이를 방지하기 위해 가
우시안 윈도우를 이진화하여 사용한다. 즉, 가우시안 윈
도우의 얼굴 영역에 해당하는 부분은 1의 값을 가지게
하여 이전의 밝기를 그대로 유지하고 배경 부분은 0의

값을 가지게 하여 검은 색으로 채우게 한다. 식 (1)은 2차원 가우시안 윈도우를 나타내는 식으로 여기서 (x_0, y_0) 는 영상 중심 좌표를 나타내며 σ_x 는 얼굴을 나타내기 위한 타원의 단축 길이의 1/2이며 σ_y 는 장축 길이의 1/2를 나타낸다. 상수 c 는 가우시안 함수의 최대값을 결정하기 위한 것으로 $1/c$ 가 가우시안 함수의 최대값을 나타낸다. 식 (2)는 가우시안 윈도우의 이진화를 나타내며, 식 (3)에서 $I(x, y)$ 는 (x, y) 좌표의 밝기 값을 나타내고 $G(x, y)$ 는 새로운 밝기 값을 나타낸다.

$$G(x, y) = \frac{1}{c} e^{-\left(\frac{(x-x_0)^2}{\sigma_x^2} + \frac{(y-y_0)^2}{\sigma_y^2}\right)} \quad (1)$$

$$G(x, y) = \begin{cases} 1 & \text{if } G(x, y) > 0.3 \\ 0 & \text{if } G(x, y) \leq 0.3 \end{cases} \quad (2)$$

$$I(x, y) = I(x, y) \cdot G(x, y) \quad (3)$$

본 논문에서는 배경 영역이 가장 잘 제거된다고 보여지는 $c=0.6$, $\sigma_x=35$, $\sigma_y=50$ 을 선택하여 실험에 사용하였다.

다음으로 영상의 밝기를 정규화한다. 얼굴 영상의 밝기는 같은 사람의 얼굴 영상이라도 영상의 획득 장소, 시간에 따른 차이를 가지고 있다. 심지어 같은 장소와 시간에 획득된 영상에서도 약간의 차이를 보이기도 한다. 얼굴의 특징을 추출하는 과정에서 얼굴 영상의 각 화소의 밝기 값이 그대로 사용되기 때문에 동일한 사람의 얼굴 영상의 밝기가 다르면 다른 사람으로 인식하게 된다. 따라서 얼굴 영상의 밝기 차이를 줄이기 위해서 영상 밝기의 정규화 과정이 필요하다. 정규화 방법으로는 히스토그램 상세화(Histogram Specification)[13] 방법을 사용한다. 히스토그램 상세화를 하게 되면 얼굴 영상의 평균 밝기 값이 일정하게 되어 영상의 전체적인 밝기가 정규화된다. 그림 4의 (b)는 히스토그램 상세화 후의 영상을 보여 주고 있다.

그림 4의 (b)를 보면 전체적인 밝기는 정규화가 이루어졌지만 영상 좌우의 밝기가 다른 것을 볼 수 있다. 이것은 조명의 방향에 따라 조명의 강도가 다르기 때문에 그림 4의 (b)는 오른쪽의 조명이 강한 것을 알 수 있다. 이러한 조명의 방향에 따른 얼굴 영상의 밝기 변화를 줄이기 위해 좌우 평균 얼굴을 생성한다. 사람의 얼굴은 일반적으로 거의 좌우 대칭을 이루고 있기 때문에 영상의 x 축 중간지점을 기준으로 좌우 화소의 밝기 값의 평균을 새로운 화소의 밝기 값으로 하더라도 얼굴의 전체적인 모양과 특징은 크게 변하지 않으며, 좌우 평균 얼굴을 만들어냄으로써 조명 방향에 따른 밝기 변화를 감소시킬 수 있다. 그림 5의 (c)는 좌우 평균 얼굴

영상으로 (a), (b)와 비교해 볼 때 얼굴 영상의 왼쪽과 오른쪽의 밝기 차이가 없음을 볼 수 있다.



(a) 배경 영역을 제거한 영상
(b) 히스토그램 상세화 후 영상
(c) 좌우 평균 얼굴 영상

그림 5 영상 밝기의 정규화

3. 특징의 추출과 인식

3.1 2차원 가보 필터(Gabor filter)의 적용

본 논문에서는 얼굴 인식을 위한 특징으로 텍스처 정보를 이용하며 텍스처 정보를 사용하기 위해 2차원 가보 필터 계수를 사용한다. 2차원 가보 필터 셋은 1980년도에 더그만(Dauman)[14]이 제안하였으며 2차원 가보 필터 계수는 영상에서 텍스처 정보를 추출하는 효과적인 방법으로 알려져 있다. 2차원 가보 필터는 지역적인 영상 구조의 방향과 공간 주파수에 대한 정보를 제공하며 동시에 2차원 위치에 관한 정보도 제공한다. 이러한 특성은 텍스처의 2차원 분광 특이성과 2차원 공간 위치 때문에 텍스처 분석에 매우 유용하다[15]. 극좌표계에서 사용되는 2차원 가보 필터는 식 (4)[15]와 같이 표현된다. 식 (4)는 두 개의 가우시안 함수와 하나의 복소수 함수로 구성되어 있다.

$$G(r, \theta) = e^{-i\omega(\theta-\theta_0)} e^{-(r-r_0)^2/a^2} e^{-(\theta-\theta_0)^2/\beta^2} \quad (4)$$

- i : $\sqrt{-1}$
- a : 가우시안 함수의 r 축의 폭
- β : 가우시안 함수의 θ 축의 폭
- ω : 가보 필터의 공간 주파수
- θ_0 : 가우시안 함수의 θ 축 중심
- r_0 : 가우시안 함수의 r 축 중심

가보 필터의 실수부와 허수부의 그래프를 보면 실수부는 y 축 대칭이고 허수부는 원점 대칭이다. 따라서 허수부는 조명의 밝기 값이 변한다고 해도 결국 합이 0이 된다. 그러나 실수부는 조명의 밝기 값이 변하면 값이 변하게 된다. 이와 같은 이유로 본 논문에서는 2차원 가보 필터의 허수부만을 사용하여 텍스처 정보를 추출한다. 즉 가보 필터 허수부의 가보 계수를 특징으로 사용

한다.

가보 필터를 얼굴 영상에 적용하기 전에 먼저 얼굴 영상을 극좌표계 영상으로 변환한다. 정규화된 얼굴 영상은 타원형의 얼굴 영역과 검은 색으로 채워진 배경 영역으로 나누어진다. 텍스처 특징 추출을 위한 가보 필터의 적용은 직사각형 모양의 윈도우(window) 단위로 이루어지기 때문에, 타원 내부에서는 얼굴 영상에 대해서만 가보 필터를 적용하게 되지만 타원 경계 부분에서는 얼굴 영역과 배경 영역에 필터를 적용하게 된다. 즉, 타원 경계 부분에서는 얼굴에 대한 텍스처 정보뿐 아니라 배경 영역에 대한 텍스처 정보도 가지게 된다. 배경 영역이 검은 색으로 정규화되기는 하였지만 배경 영역에 대한 텍스처 정보가 얼굴 인식에 좋지 않은 영향을 미친다는 것은 분명하다. 따라서 배경 영역에 가보 필터를 적용하지 않도록 하기 위해 방법이 필요하다. 이를 위해 xy 좌표계의 영상을 r 과 θ 의 값으로 이루어진 극좌표계로 변환하여 r 축과 θ 축 상에서 영상으로 표현한다. 변환된 영상은 직사각형의 모양을 가지게 되어 가보 필터의 적용이 용이해지고, 배경 영역에 대한 특징 값을 배제하게 된다. 반지름 r 과 각 θ 로 이루어진 극좌표계로의 변환은 배경 영역을 제외한 타원형의 얼굴 영역에 대해서만 이루어지며, 본 논문에서는 반지름 r 은 0~44의 45단계 값을 갖고, 각 θ 는 $0^\circ \sim 359^\circ$ 의 360 단계 값을 갖도록 하여 극좌표계 영상으로 변환한다.



(a)타원형의 얼굴 영상 (b) 극좌표계로 변환된 영상
그림 6 영상의 극좌표계로의 변환

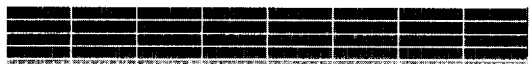
3.2 다해상도(multi-resolution) 영상

2차원 가보 필터는 텍스처 분석에 유용한 도구이지만, 적절한 텍스처를 가지고 있지 않은 영상에 대해서도 그런 것은 아니다. 영상이 적절한 텍스처를 가지고 있어야만 유용한 텍스처 분석이 이루어질 수 있는 것이다. 적절한 텍스처를 가진다는 것은 텍스처 프리미티브(texture primitive)가 적절한 크기를 가진다는 것을 의미한다. 텍스처 프리미티브가 너무 크면 텍스처라고 보기보다는 하나의 물체(object)로 간주할 수 있으며, 반면 너무 작으면 임의의 잡음(random noise)으로

도 볼 수 있기 때문이다[16].

따라서 영상이 적절한 텍스처를 가졌는지 확인하는 과정이 필요하다. 이것은 영상의 로컬 익스트리마(Local Extrema)의 밀도를 계산함으로써 이루어질 수 있다[17]. 그러나 본 논문에서는 이 과정을 처리하지 않았는데, 그것은 다해상도 영상에 대해 FD 분석 과정을 거치기 때문이다. 영상의 화소가 가지는 명암 값의 범위를 여러 단계로 나누어서, 각 영상에 대해 FD 분석을 수행하게 되므로, 적절한 텍스처를 가지지 못하는 영상은 FD 분석 과정에서 낮은 FD 값을 가지게 되고, 반대로 적절한 텍스처를 가지는 영상은 FD 분석 과정에서 높은 FD 값을 가지게 되어 얼굴 인식을 위한 특징으로 사용된다. 때문에 FD 분석 과정을 통해 영상이 적절한 텍스처를 가지고 있는가에 대한 확인을 거치지 않고도 좋은 텍스처를 가지는 특정 해상도의 영상을 선택할 수 있다.

이와 같은 이유로 각 영상은 256, 128, 64, 32, 16, 8, 4, 2 명암 단계를 가지도록 한다. 즉 한 영상에 대해 8가지 해상도를 가지도록 하여 각각의 해상도에서의 영상에 2차원 가보 필터를 적용한다. 이때 2차원 가보 필터는 영상 전체에 적용하지 않고 극좌표계 영상을 작은 크기의 셀(cell)로 분할하여 각 셀에 적용한다. 이렇게 작은 셀 단위로 처리하는 것은 영상 전체에 가보 필터를 적용하면 가보 계수를 계산하는데 상당히 많은 시간이 소모되어 인식 속도의 저하를 가져오기 때문에 가보 필터를 적용하는데 걸리는 시간 비용을 줄이기 위해서 영상 전체가 아닌 작은 셀 단위로 처리한다. 셀은 실험에 적절하다고 여겨지는 45×9 의 크기로 나누었다. 그림 7은 다단계 명암 영상의 예를 보인 것이다.



(a) 256 명암 단계 영상



(b) 2 명암 단계 영상

그림 7 다단계 명암 영상

3.3 FD 분석(Fisher's Discriminant Analysis)

2차원 가보 필터는 각 매개 변수를 변화시킴에 따라 수많은 결과를 만들어 낸다. 이 수많은 가보 계수들을 모두 사용한다면 상당한 시간이 필요하게 될 것이다. 따라서 수많은 가보 계수 중에서 보다 유용하고

결정력이 있는 매개 변수들을 선택할 필요가 있다. 본 논문에서는 가장 구분력이 큰 매개 변수들의 조합을 선택하기 위해 특징 선택 방법들 중 FD 분석[17]을 사용하여 구분력이 큰 특징들을 선택하도록 한다. FD 분석은 클래스 간의 분리성(separability)과 클래스 내의 동질성(homogeneity)을 최대화시키는 특징을 선택하는 기준으로 다음과 같이 정의된다.

$$FD_j = \frac{V_j}{\sum_i V_{ij} / (P \cdot Q)} \quad (5)$$

- V_j : 전체 훈련 데이터의 j 번째 특징의 편차
- V_{ij} : i 번째 class의 j 번째 특징의 편차
- P : class의 개수
- Q : 각 class에 소속된 훈련 데이터의 개수

2차원 가보 필터에 사용된 매개 변수의 범위는 무한히 증가할 수 있다. 따라서 모든 매개 변수에 따른 가보 계수를 계산한다는 것은 불가능하기 때문에 매개 변수의 범위를 아래와 같이 제한하여 계산 범위를 줄인다.

- α : 1~5
- β : 1~5
- ω : 1~10

3.4 인식

인식에 사용되는 특징은 FD 분석을 통해 선택된 구분력이 큰 특징을 사용한다. 각 셀에서 가장 큰 FD 값을 가지는 가보 필터의 매개 변수와 명암 단계를 선택하여 계산된 가보 계수를 특징으로 한다. 밝기 값에 의존적인 가보 필터의 실수부는 사용하지 않으며 밝기 값에 독립적인 가보 필터의 허수부에 대한 가보 계수만을 사용한다. 360×45 크기를 가지는 극좌표계 영상을 45×9 크기의 셀로 나누게 되면 모두 40개의 셀로 나누어지므로, 각 극좌표계 영상은 40개의 특징이 선택되게 된다. 특징 집합 $F = \{f_1, f_2, \dots, f_{40}\}$ 에서 f_n 은 다음과 같이 계산된다.

$$f_n = [im] \int_{\rho} \int_{\phi} e^{-i\omega(\theta_0 - \phi)} e^{-(r_0 - \rho)^2 / \alpha^2} e^{-i(\theta_0 - \phi)^2 / \beta^2} I(\rho, \phi) \rho d\rho d\phi \quad (6)$$

여기서 각 특징 f_n 은 FD 분석에서 선택된 매개 변수에 의해 계산된다.

이 때 각 셀에서 선택된 특징은 인식에 도움을 주는 것과는 도움이 되지 않는 것들이 있다. FD 값이 큰 것은 인식 성능을 높이게 되지만 FD 값이 낮은 특징은 인식에 큰 영향을 끼치지 않거나, 좋지 않은 영향을 미칠 수 있다. 따라서 각 셀에서의 가장 높은 FD 값을 가중치로 사용하여, 이 FD 값의 크기에 따라 각 특징의 인식 기여도를 조절한다. 낮은 FD 값을 가지는 특징에 대한 가중치를 낮추어, 인식에 대한 영향력

을 작게 하고, 높은 FD 값을 가지는 특징에 대한 가중치를 높여, 좋은 특징의 인식률에 대한 영향력을 크게 함으로써 인식률을 높인다. n 번째 특징에 대한 가중치는 다음과 같이 정의한다. 여기서 FD_n 은 n 번째 특징의 FD 값을 나타낸다.

$$W_n = \frac{FD_n}{\sum_{k=1}^M FD_k} \quad (7)$$

인식은 이 특징들의 차의 합이 가장 작은 값을 선택함으로써 이루어진다. 특징 집합 F 와 각 클래스간의 거리는 아래와 같이 계산된다.

$$distance = \sum_{n=1}^M (W_n \times |f_n - f_{km}|) \quad (8)$$

f_{km} 은 k class의 n 번째 특징의 대표값으로서, 본 논문에서는 FD 분석에 사용된 5장의 훈련 영상들의 n 번째 특징의 평균값을 의미한다. 즉 n 번째 셀에서 FD 분석에 의해 선택된 매개 변수에 의한 가보 계수의 평균을 의미한다.

4. 실험결과

실험에 사용된 영상은 서로 다른 개인 10명의 정면 얼굴 영상 70장을 사용하였다. 각 개인에 대해 7장의 영상을 사용하였으며, 이 중 5장의 영상은 FD 분석을 위해 사용하여 한 사람을 대표하는 특징 값을 결정하도록 하였다. 그림 8은 실험에 사용된 10명의 영상을 나타낸다.



그림 8 실험 영상

4.1 FD 분석 결과

가보 필터는 매개 변수 $\alpha, \beta, \omega, r_0, \theta_0$ 에 따라 다양한 결과 값이 계산되어지며, 또한 극좌표계 영상의 한 화소가 가질 수 있는 해상도 범위도 바꾸어 가며 가보 필터가 적용되므로 모두 6개의 매개 변수의 변화에 따른 가보 필터가 계산된다. 여기서 r_0 와 θ_0 는 각 셀의

중심으로 고정시키고 다른 매개 변수들을 변화시킴으로써 얻어지는 가보 계수들에 대한 FD 분석을 수행하였다. FD 분석에 사용된 매개 변수의 범위는 다음과 같다.

반지름 r 에 대한 가우시안 함수의 크기 α : 1~5
 각 θ 에 대한 가우시안 함수의 크기 β : 1~5
 공간 주파수 ω : 1~10
 gray level : 256,128,64,32,16,8,4,2

위의 매개 변수에 대해 각 셀이 가질 수 있는 가보 계수 값들은 모두 2000 개이다. 셀은 모두 40 개이므로 모두 80000 개의 가보 계수 값이 계산된다. 이 특징들 중에서 각 셀에 대해 FD 값이 가장 높은 매개 변수의 조합을 선택하고 가보 계수를 계산하여 인식을 위한 특징 값으로 사용한다.

4.2 인식 결과

모두 70장의 얼굴 영상이 인식 실험에 사용되었으며, 10명의 개인에 대해 각각 7장의 영상을 획득하였다. 각 개인의 영상 중 5장은 FD 분석에 사용하였고 2장은 FD 분석에 사용되지 않았다. 그림 9는 인식 결과를 나타내는 그래프이다.

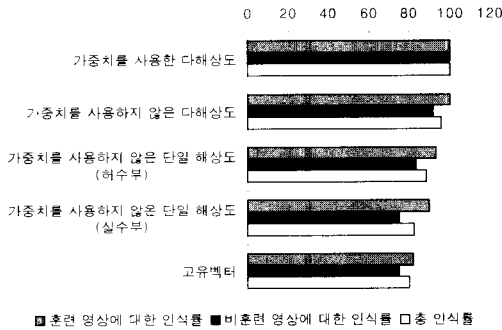


그림 9 70장의 얼굴 영상에 대한 인식 결과

그림 9는 본 논문에서 제안한 방법인 가중치를 사용한 다해상도 영상을 적용한 경우의 인식률과 가중치는 사용하지 않고 다해상도 영상만을 적용한 경우의 인식률, 가중치도 사용하지 않고 단일 해상도 영상만을 적용한 경우의 인식률을 보여준다. 고유벡터 방법은 [1]에서 제안된 방법을 사용하여 본 실험에 사용된 영상에 대한 실험한 결과이다.

실험 데이터에 대한 총 인식률을 비교해 보면 본 논문에서 제안한 가중치를 사용한 다해상도 영상을 적용한 방법의 인식률은 100%로 가장 높았으며, 가중치를 사용하지 않고 다해상도 영상만을 이용한 방법의 인식률은 97.61%를 보였다. 반면, 논문에서 제안된 방법을 사용하지 않은 경우, 다시 말해 가중치를 사용하지 않은

단일 해상도 영상을 적용하여 단순히 특징들의 차이의 합이 가장 작은 것을 선택한 경우로 가보 필터의 실수부의 특징에 대한 인식률은 83.33%, 허수부의 특징에 대한 인식률은 90.47%로 나타났다. 마지막으로 고유 벡터를 이용한 방법은 인식률 80%로 가보 필터를 이용한 방법보다 낮은 인식률을 보였다. 위 결과를 종합해 보면 가보 필터의 실수부의 특징에 대한 인식률보다 허수부의 특징에 대한 인식률이 높으며, 다해상도 영상의 사용과 가보 필터를 이용한 텍스처 특징에 가중치를 부여하면 인식률이 높아진다는 것을 알 수 있다.

5. 결론

본 논문에서는 CCD 카메라로 입력받은 영상에서 정면 얼굴 영상을 추출하여 2차원 가보 필터를 적용하여 정면 얼굴을 인식하는 방법을 제안하였다. 기존 얼굴 인식 방법에서는 고려하지 않았던 조명의 세기나, 조명의 방향에 따른 얼굴 영상의 밝기 변화를 줄이기 위해서 추출된 얼굴 영상을 정규화하는 과정에서 히스토그램 상세화와 좌우 평균 얼굴을 사용하여 얼굴 영상의 밝기 변화를 줄였으며, 얼굴 영상을 극좌표계로 변환하고 가보 필터를 적용하는 과정에서 다단계의 명암 단계를 가지게 하여 보다 좋은 텍스처를 가진 영상에서 특징을 선택할 수 있도록 하였다. 가보 필터를 이용한 기존의 얼굴 인식 방법[5]은 가보 필터를 사용함에 있어서 단순히 몇 개의 매개 변수만을 선택하여 사용하였지만 본 논문에서는 다양한 매개 변수들에 대한 가보 필터를 적용하고 FD 분석을 이용하여 보다 구분력이 큰 매개 변수들의 조합을 선택함으로써 특징들의 구분력을 높였으며, 각 특징에 가중치를 부여함으로써 인식의 정확도를 높였다. 그러나, FD 분석 과정에서 다양한 가보 필터를 수행해야 하기 때문에 많은 계산 시간이 필요하다는 것이 단점이라고 볼 수 있는데, 이것은 컴퓨터의 연산 속도가 빨라짐에 따라 해결될 수 있을 것으로 보인다. 마지막으로 실험에 필요한 데이터를 구하는 데 어려움이 있어 많은 데이터에 대하여 실험하지 못한 점이 아쉬움으로 남는다.

참고 문헌

[1] Matthew Turk and Alex Pentland, "Eigenfaces for Recognition." Journal of Cognitive Neuroscience, Vol. 3, Num. 1, 1991.
 [2] Alex Pentland, Baback Moghaddam, Thad Starner, "View-Based and Modular Eigenspaces for Face Recognition," Proc. Computer Vision and Pattern

- Recognition, Vol. 1, 1994.
- [3] Baback Moghaddam and Alex Pentland, "Probabilistic Visual Learning for Object Detection," The 5th International Conference on Computer Vision, 1995.
 - [4] B. S. Manjunath, "A Feature Based Approach to Face Recognition," CVPR92, pp.373-378, 1992.
 - [5] Laurenz Wiskott, Jean-Marc Fellous, Norbert Kruger, Christoph von der Malsburg, "Face Recognition by Elastic Bunch Graph Matching," Intelligent Biometric Technique in Fingerprint and Face Recognition, 1999.
 - [6] Shi-Hong Jeng, H. Y. M. Liao, C. C. Han, M. Y. Chern, and Y. T. Liu, Facial "Facial Feature Detection Using Geometrical Face Model: An Efficient Approach," Pattern Recognition, Vol.31, No.3, pp.273-282, 1998.
 - [7] R. Brunelli and T. Poggio, "Face Recognition through Geometrical Features," Proc. ECCV92, pp.792-800, 1992.
 - [8] R. Brunelli, T. Poggio, "HyperBF Network for Real Object Recognition," 12th International Joint Conference on Artificial Intelligence, pp.1278-1284, 1991.
 - [9] Steve Lawrence, C. Lee Giles, Ah Chung Tsoi, Andrew D. Back, "Face Recognition: A Hybrid Neural Network Approach," Technical Report UMIACS-TR-96-16 and CS-TR-3608, University of Maryland, 1996.
 - [10] Rafael C. Gonzalez, Richard E. Woods, Digital Image Processing, Addison-Wesley, 1992.
 - [11] 고일주, "컴퓨터 시각을 이용한 चेस्चि 기반 인터페이스", 숭실대학교 박사 학위 논문, 1997.
 - [12] Richard O. Duda, Peter E. Hart, Pattern Classification and Scene Analysis, A Wiley Interscience Publication, 1973.
 - [13] Randy Crane, Simplified approach to Image Processing, Prentice-Hall, 1997.
 - [14] J. G. Daugman, "Two-dimensional spectral analysis of cortical receptive field profiles," Vision Res. vol. 20, pp. 847-856, 1980.
 - [15] John G. Daugman, "High Confidence Visual Recognition of Persons by a Test of Statistical Independence," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 15, pp. 1148-1161, 1993.
 - [16] Kalle Karu, Anil K. Jain, and Ruud M. Bolle, "Is there any texture in the image?," Pattern Recognition, Vol. 29, No. 9, pp. 1437-1446, 1996.
 - [17] 고희선, "가중치를 이용한 홍채코드 생성과 인식", 숭실대학교 석사 학위 논문, 1998.



이 찬 호

1998년 군산대학교 컴퓨터과학과 학사.
2000년 숭실대학교 컴퓨터학과 석사.



박 주 철

1987년 숭실대학교 전자계산학과 학사.
1991년 숭실대학교 컴퓨터학과 석사.
1996년 ~ 현재 숭실대학교 컴퓨터학과 박사과정.



최 형 일

1979년 2월 연세대학교 전자공학과 졸업 (공학사). 1982년 6월 미시간대학교 전산공학과 졸업(공학석사). 1987년 6월 미시간대학교 전산공학과 졸업(공학박사). 현재 숭실대학교 컴퓨터학부 교수. 관심분야는 컴퓨터 비전, 퍼지 및 신경망이론, 패턴인식, 지식기반시스템 등임.