

특징정보 분석을 통한 실시간 얼굴인식

Realtime Face Recognition by Analysis of Feature Information

정재모, 배현, 김성신

부산대학교 전기공학과

Jae Mo Chung, Hyun Bae, Sung Shin Kim
School of Electrical and Computer Eng., Pusan National Univ.
(jaemochung@hanmail.net)

ABSTRACT

The statistical analysis of the feature extraction and the neural networks are proposed to recognize a human face. In the preprocessing step, the normalized skin color map with Gaussian functions is employed to extract the region of face candidate. The feature information in the region of the face candidate is used to detect the face region. In the recognition step, as a tested, the 120 images of 10 persons are trained by the backpropagation algorithm. The images of each person are obtained from the various direction, pose, and facial expression. Input variables of the neural networks are the geometrical feature information and the feature information that comes from the eigenface spaces. The simulation results of 10 persons show that the proposed method yields high recognition rates.

Keywords : Eigenface, Template matching, Distance from face space(DFFS), Distance in face space(DIFS), Principal component analysis(PCA).

I. 서론

21세기 컴퓨터와 멀티미디어 기술의 발전으로 인해 적용분야가 다양한 영상정보 분석에 대한 관심이 날로 증가하고 있다. 일반산업분야에서 비전시스템 적용은 이미 일반화되었고 무인감시시스템 및 보안시스템, 컴퓨터와 사용자간의 인터페이스 구축 등 적용범위를 넓혀가고 있다. 이러한 점에 있어 얼굴을 포함한 주요생체인식 기술의 개발에 관심이 집중되고 있다. 이중 응용분야가 넓고 정보획득이 편리한 것이 바로 얼굴인식 기술이다. 그러나 얼굴의 각도 밝기 표정 나이에 따라 계속 변하기 때문에 실용화가 쉽지 않다. 가장 중요하고 어려운 문제는 입력된 영상에서 얼굴을 추출하는 방법이다. 여기에는 피부색정보와 눈의 위치를 이용한 방법[5]과 PCA(principal component analysis)를 이용한 방법[2], 신경망을 이용한 방법[3], 2차원 가우시안 밀도함수를 이용한 방법[2], 템플릿매칭을 이용한 방법 등 다양한 방법으로 접근하고 있다. 인식단계에서는 얼굴공간(face space)상의 DIFS의 분석을 통한 인식방법과 elastic matching을 이용한 방법[4], 그리고 신경망(Neural network)에 의한 방법[4]등이 대표적이다.

본 논문에서는 피부색정보와 고유얼굴 공간상의 PCA(DIFS와 DFFS의 조합)를 이용하여 얼굴영역을 추출하는 방법[2] 그리고 템플릿 매칭을 이용한 얼굴

영역추출 방법을 비교 분석하였으면 인식 단계에서는 역전파(BP) 알고리즘을 사용하여 시뮬레이션 하였다.

II. 전체 시스템 개요

실제 얼굴인식 시스템에서는 입력영상에서 얼굴영역 추출(face detection)하는 단계와 추출한 얼굴이미지를 이용한 얼굴인식단계(face recognition)로 크게 나눌 수 있다. 얼굴영역 추출단계에서는 3가지 방법(DFFS, DFFS+ DIFS, 템플릿 매칭)을 사용하여 추출, 비교 분석하였으며 인식단계에서는 신경망을 사용하였다. 신경망은 정면이미지를 포함한 상하좌우 15도 이내의 이미지, 빛의 변화가 있는 이미지를(개인당 12개) 사용하여 학습을 시켰으며, 준비된 얼굴 샘플의 크기는 64×64로 제한하였다

III. 얼굴추출

실시간 얼굴인식 시스템에서는 얼굴추출과정이 반드시 선행되어야 하며, 얼굴 추출결과가 전체 얼굴인식 성능에 크게 영향을 미친다. 일반적으로 얼굴을 추출하는 방법에는 얼굴의 전체적인 패턴을 분석하여 얼굴을 추출하는 방법, 직관적인 상관관계로 추출하는 템플릿 매칭방법으로 나눌 수 있다. 실시간 얼굴인식시스템에서 무시할 수 없는 것이 바로 계산시간이며 상대적

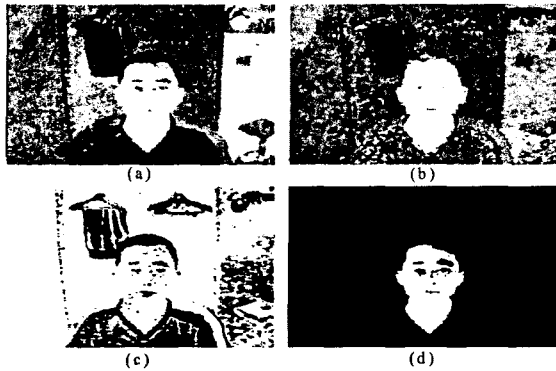


그림1. a)R-G확률이미지 b)R-B확률이미지
c)G-B확률이미지 d)얼굴후보영역.
Figure1.

a)R-G probability image b)R-B Probability image
c)G-B probability image d)face candidate region.

으로 계산량이 많은 얼굴영역 추출과정이 대부분을 차지한다. 그래서 피부색의 R,G,B색차 정보를 이용하여 얼굴후보 영역을 추출함으로써 계산시간을 줄일 수 있었다. 그리고 본 논문에서는 얼굴추출성능을 개선하기 위해 눈의 위치를 제한하여 추출성능을 개선하고자 하였다.

3.1 피부색 정보를 이용한 얼굴후보영역추출

컬러 영상은 조명변화에 아주 민감하므로 간단히 Min-Max Clustering에 의해서는 원하는 결과를 얻기가 어렵다. 실제 같은 이미지라도 조명의 변화가 있으면 카메라에서 받아들인 이미지의 R,G,B픽셀값이 현저한 차이를 보임을 알 수 있다. 이 문제를 해결하기 위해 RGB, HSV컬러공간으로의 변환을 통해 조명변화에 둔감하도록 하는 방법도 있다. 본 논문에서는 그림1의a, b, c와 같이 피부색의 R-G, R-B, G-B확률이미지의 조합으로 d)의 얼굴후보 이미지를 추출하였다. 확률함수는 식(1)의 1차원 가우시안 확률밀도함수를 사용하였다. x 는 정규화(normalize)된 R과B의(R-G, G-B) 색차값이며 m 은 평균값(mean), σ 는 분산(variance)이다. 확률이미지가 조명변화에 둔감하도록 하기 위해 빛의 세기와 방향을 다양하게 변화시키면서 추출한 색차값을 정규화(normalize) 과정을 거쳐 평균값과 분산을 구하였다. 정규화 과정을 거치지 않으면 제한된 환경을 벗어날 경우 얼굴후보 영역을 찾지 못하는 경우가 발생한다. 얼굴후보영역(d)이미지를 추출하기 위해서 침식(Erosion)과 팽창(Dilation)을 사용하여 간단히 잡음(noise)를 제거한 후 라벨링(labeling)과정을 거쳐 적당한 넓이와 가로, 세로축의 크기 비를 가지는 영역을 얼굴후보 영역으로 정의하였다. PCA 및 템플릿매칭을 이용한 얼굴추출과정은 입력영상(320*240)이 아닌 후보영역을 탐색하게된다.

$$\text{Gaussian pdf } f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-m)^2/2\sigma} \quad (1)$$

3.2 PCA를 이용한 얼굴추출

PCA(principal component analysis)의 주요 목표는

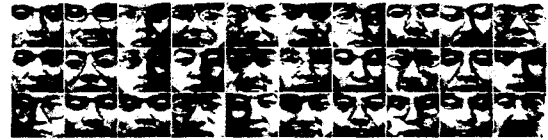


그림 2. 얼굴 샘플.
Figure 2. face samples.

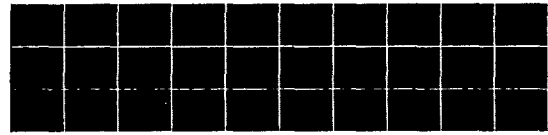


그림 3. 그림2로부터 구한 30개의 고유얼굴.
Figure 3. Thirty of the eigenfaces calculated from the Figure 2.

고차원(high dimension)의 데이터를 data set의 특징을 모두 포함하는 저차원(low dimension)의 특징공간(feature space)으로 재정의 하는 것이다. 2차원 얼굴이미지가 I*J이면 벡터 x 는 $N(I \times J)$ 차원의 고차의 벡터가 되고 PCA에 의해 벡터 x 는 저차의 특징공간으로 다시 할당하게된다.

얼굴샘플벡터 x 의 공분산행렬(covariance matrix) Σ 가 주어지면 KL(Karhunen-Loeve)변환의 기저함수를 아래 식(2)와 같이 간단히 얻을 수 있다.

$$A = P^T \Sigma P \quad (2)$$

여기서 P 는 Σ 의 고유벡터 행렬이고, Λ 는 고유값으로 구성된 대각행렬(diagonal matrix)를 나타낸다. 그림3을 보면 고유값이 가장 큰 값에 대응하는 고유벡터(고유얼굴)부터 차례대로 배열한 것으로 고유값이 클수록 얼굴이미지와 비슷함을 알 수 있다. 이 정보로부터 고유값이 큰 것부터 M 개($M \ll N$)를 선정, 이에 대응하는 고유벡터(고유얼굴)로 구성된 M 차(M dimension)의 새로운 얼굴공간을 생성한다. M 개의 고유얼굴로 구성된 행렬을 P_M 이라고 정의하고, 얼굴공간상의 특징정보(principal components) 벡터 y 는 얼굴이미지벡터 x 를 얼굴공간상에 투영(식(3))함으로써 구할 수 있다.

$$y = P_M^T (x - \bar{x}) \quad (3)$$

여기서 \bar{x} 는 평균얼굴이미지를 나타낸다. 얼굴 추출시 입력영상을 탐색하는 과정에서 탐색영역이 얼굴인지 아닌지(face or non-face)의 정보는 DFFS (distance from face space)값을 구함으로써 알 수 있으며 이 값은 M 차원의 얼굴공간상의 특징정보 벡터 y 와 P_M 를 이용해 근사화된 얼굴이미지 \hat{x} 를 복원 할 수 있다. 원래 이미지와 복원이미지의 차이 값의 norm (reconstruction error ϵ^2)이 DFFS 값이며 식(5)에 정의되어 있다.

$$\hat{x} = P_M y + \bar{x} \quad (4)$$

얼굴공간상의 특징정보 y_i 는 대응되는 고유값(eigen value)에 비례하여 변화하므로 Mahalanobis

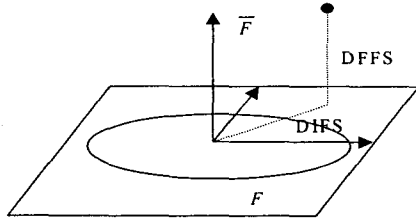


그림 4. 주 얼굴공간 F 부 얼굴공간 \bar{F}
 Figure 4. Principal face space F
 Complement face space \bar{F}



그림5. 평균 이미지.
 Figure 5. Mean Image.

$$\begin{aligned} \epsilon^2 &= \|(\mathbf{x} - \hat{\mathbf{x}})\|^2 \\ &= \|(\mathbf{x} - \hat{\mathbf{x}}) - P_M \mathbf{y}\|^2 \\ &= \|\mathbf{x} - \hat{\mathbf{x}}\|^2 - \sum_{i=1}^M y_i^2 \\ &= \sum_{i=M+1}^N y_i^2 \end{aligned} \quad (5)$$

distance의 제곱(d^2)이 얼굴공간상에 투영된 얼굴 이미지와 평균 이미지와의 거리를 나타내며 이를 DIFS(distance in face space)라고 정의한다[3]. 일반적으로 얼굴추출과정에서 DFFS를 추출기준 값으로 사용하지만 성능 향상을 위해 식(6)에 정의된 DIFS를 보완 알고리즘으로 사용하였다.

$$d^2 = \mathbf{y}^T \Lambda^{-1} \mathbf{y} = \sum_{i=1}^M \frac{y_i^2}{\lambda_i} \quad (6)$$

3.2 템플릿 매칭을 이용한 얼굴추출

템플릿 매칭의 목적은 주어진 영상 내에서 찾고자 하는 대상의 템플릿을 이동시키면서 유사도를 조사하여 그 값이 기준 값 이상일 경우 우리가 찾는 대상으로 판정한다. 이 방법은 특징정보분석이 아닌 직관적인 이미지의 비교 방법이므로 조명변화 얼굴의 각도변화 매우 민감하다. 본 논문에서는 조명변화에 둔감하도록 하기 위해 히스토그램 평활화 과정을 먼저 거치게 된다. 템플릿 이미지는 그림 5의 평균 이미지(64×64)를 사용하였으며 40×40에서 80×80사이를 11개로 나누어 각 사이즈의 이미지를 추출 후 기준 크기(64×64)로 확대 축소 한 후 템플릿 매칭과정이 이루어졌다.

IV. 기하학적 특징정보 추출

얼굴의 기하학적 특징정보는 눈, 코, 입, 눈썹의 크기 및 상대적인 위치관계, 양쪽 귀사이의 거리, 턱선의 모양 정보 등을 통칭하는 말이다. 본 논문에서 그림 6과 같이 추출된 얼굴이미지 내에서 눈, 코, 입이 존재할 후보영역을 미리 정한 후 X축과 Y축으로 투영하여 12개의 거리 특징정보를 추출하였으며, 이 정보를 바탕으로 7개의 상대적인 거리비 특징정보를

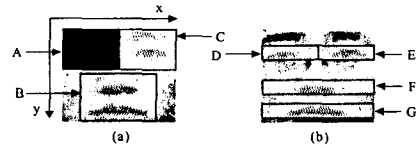


그림 6. 눈, 코, 입의 예상 후보 영역.
 Fig. 6. Expected Candidate Region of Eyes, Nose, Mouth.

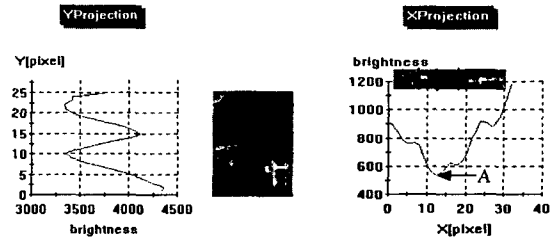


그림 7. 예상후보영역의 X,Y축에 대한 투영.
 Fig. 7. The Projection of Expected Candidate Region on to the X and Y-axis.

구하여 신경망의 입력으로 사용하였다.

V. 얼굴인식

5.1 얼굴공간선정

고유얼굴은 평균얼굴에 대한 분산을 나타내는 기저 벡터로 큰 고유값에 해당하는 고유얼굴이 큰 분산을 가진다. 그림 3.을 보면 고유값이 작아질수록 얼굴특성을 잃어갈 알 수 있다. 고유값이 작은 고유얼굴에 대한 특징정보는 얼굴에 대한 정보가 아닌 잡음으로 간주한다. 그러므로 얼굴공간 선정 시 얼굴에 대한 정보만 잘 나타낼 수 있는 공간을 선정해야 하며 얼굴이미지의 특징정보 개수는 얼굴공간의 차원(고유얼굴의 개수)과 같아진다. 일반적으로 얼굴공간 선정 시 전체 고유치의 합에 대한 사용하는 고유치의 합의 백분율이 어떤 기준치 이상이 되는 n 값을 선정하게 된다.

$$\text{백분율(\%)} = \frac{\sum_{i=1}^n \lambda_i}{\sum_{i=1}^M \lambda_i} \times 100 \quad (7)$$

λ_i : 공분산 행렬의 고유치

본 논문에서는 30(M)개의 고유치의 합의 백분율이 90%가 넘는 17(n)개의 고유얼굴로 구성된 얼굴공간을 선정하였다.

5.2 신경망에 의한 얼굴인식

일반적으로 얼굴인식 알고리즘으로는 얼굴공간상의 유클리디안 거리로 인식하는 방법과 Auto-association net에 의해 얼굴정보를 축소 후 신경망에 의해 인식하는 방법 등이 있다. 전자의 경우 비 선형성이 강한 얼굴 특징정보에 대한 인식성능이 뛰어나지 못하며 후자의 경우 얼굴이미지 전체가 Auto-association net의 입력정보이기 때문에 얼굴이미지의 크기에 제한이 따른다. 이 경우 이미지 축소과정이 필요하고 이때 얼굴특징정보를 많이 잃게 되므로 우수한 성능을 기대하기

어렵다. 이러한 단점을 극복하기 위해 얼굴공간상의 특징정보와 기하학적 특징정보를 입력으로 하는 신경망으로 역전파 알고리즘을 사용하였다.

VI. 실험결과

본 논문에서는 얼굴추출과 얼굴인식으로 나누어 연구실내 환경에서 이루어 졌다. CCD카메라에 의해 받아들여진 입력영상은 Fram Grabber에 의해 320×240해상도의 24bit RGB Color Format으로 저장되며 얼굴후보영역 추출 단계가 지나면 Gray Scale Image에서 모든 실험이 이루어졌다.

6.1 얼굴추출실험

본 논문에서는 3가지 방법으로 얼굴 추출을 하였다. 실험 시 빛의 방향 변화와 밝기 변화가 있는 2가지 환경으로 나누어 테스트하였다.

표 1. 얼굴 추출 결과1

Table 1. The First Result of Face Extraction

	DFFS	DFFS+DIFS	Template Matching	비 고
추출율(%)	48(%)	73(%)	98(%)	빛의 방향변화

표 2. 얼굴 추출 결과2

Table 2. The Second Result of Face Extraction

	DFFS	DFFS+DIFS	Template Matching	비 고
추출율(%)	66(%)	78(%)	98(%)	빛의 밝기변화

위 실험결과를 살펴보면 DFFS와 DFFS+DIFS방법은 빛의 방향과 밝기 변화 시 큰 성능 저하를 보였으며, 템플릿 매칭의 경우 빛의 변화에 크게 영향을 받지 않았으며, 추출 소요시간면에서도 우수한 결과를 보여 주었다.

6.2 얼굴인식실험

표3과 같이 입력정보가 다른 3개의 신경망 모델을 구성하고 은닉층의 개수의 변화에 따른 성능변화를 살펴 보았다. 실험 조건은 얼굴추출시와 동일하게 빛의 방향변화와 밝기변화가 있는 환경에서 각각 실험하였다.

표 3. 신경망 모델

Table 3. Neural Network Model

	입력정보	입력 개수	은닉층 개수	은닉층 node 수	학습방법	비 고
신경망 1	얼굴공간상의 특징정보	17	1-3	51	역전파방법 (Back Propagation)	
신경망 2	기하학적 특징정보	7	1-3	21		
신경망 3	얼굴공간상의 특징정보 + 기하학적 특징정보	24	1-3	72		

실험 결과를 살펴보면 신경망 모델3이 가장 우수한 성능을 나타내고 있으며, 은닉층의 개수가 증가할수록 빛의 변화에 따른 성능저하 및 분산값이 작은 것을 알 수 있다.

VII. 결론

피부색 정보를 이용한 얼굴 후보영역추출로 얼굴

표 4. 얼굴 인식 결과1

Table 4. The First Result of Face Recognition

		은닉층 개수			평균 인식율	Threshold	비 고
		1	2	3			
신경망 1	인식율	97.6(%)	82.6(%)	96.87(%)	92.36(%)	0.6/0.4	빛의 방향변화 Training Error:0.04
	학습정확도	100(%)	100(%)	100(%)			
신경망 2	인식율	40.4(%)	60.86(%)	71.87(%)	57.71(%)	0.6/0.4	
	학습정확도	98(%)	100(%)	100(%)			
신경망 3	인식율	90.47(%)	91.3(%)	90.625(%)	90.79(%)	0.6/0.4	
	학습정확도	100(%)	100(%)	100(%)			
평균 인식율		76.15(%)	78.25(%)	86.455(%)			

표 5. 얼굴 인식 결과2

Table 5. The Second Result of Face Recognition

		은닉층 개수			평균 인식율	Threshold	비 고
		1	2	3			
신경망 1	인식율	77.14(%)	85.29(%)	86.48(%)	82.97(%)	0.6/0.4	빛의 밝기변화 Training Error:0.04
	학습정확도	100(%)	100(%)	100(%)			
신경망 2	인식율	65.87(%)	64.7(%)	64.86(%)	65.14(%)	0.6/0.4	
	학습정확도	98(%)	100(%)	100(%)			
신경망 3	인식율	90.625(%)	88.23(%)	94.59(%)	91.14(%)	0.6/0.4	
	학습정확도	100(%)	100(%)	100(%)			
평균 인식율		77.87(%)	79.40(%)	81.97(%)			

추출 시 전역탐색의 비효율성을 개선 할 수 있었으며, 눈의 위치를 제한하여 얼굴 추출성능의 향상시킬 수 있었다. 3가지 얼굴추출 방법 중 템플릿 매칭이 가장 우수한 성능을 보였으며, 추출소요시간 또한 가장 짧았다. 얼굴 인식과정에서는 비 선형 맵핑능력이 좋은 신경망을 사용하였으며, 실험결과 얼굴의 기하학적 특징정보와 얼굴공간상의 특징정보를 입력으로 하는 신경망 모델3이 가장 우수한 성능을 나타내었다. 그리고 은닉층의 개수가 증가할수록 빛의 방향변화에 대하여 적은 분산값을 가지고 우수한 인식 결과를 나타내었다. 향후 과제로는 얼굴추출단계의 개선을 위해 얼굴 후보영역 추출시 추가적인 알고리즘 적용으로 보다 최적화 된 얼굴후보영역을 추출해야 할 것이며, 기하학적 특징정보를 보다 정확하고 많이 추출해 할 것이다. 그리고 빛의 방향변화에 대한 보상 알고리즘을 적용하여 적은 데이터로 우수한 인식 성능을 보이는 시스템을 구현해야 할 것이다.

VIII. 참고문헌

- [1] B. Moghaddam and A. Pentland, "Probabilistic Visual Learning for Object Representation," IEEE Transactions on , Volume: 19, Page(s): 696 -710 , July 1997
- [2] B. Menser and F Muller, "Face Detection In Color Images Using Principal Components Analysis," Seventh International Conference on vol.2, Page(s):620-624, 1999
- [3] H.A. Rowley and S. Baluja, "Rotation Invariant Neural Network-Based Face Detection," 1998 IEEE Computer Society Conference on, Page(s): 963 -963, 1998
- [4] J. Zhang and Y. Yan and M. Lades, "Face Recognition: Eigenface, Elastic Matching, and Neural Nets," Proceedings of the IEEE, Page(s): 1423 -1435, 1997
- [5] L. Zhang and P. Lenders, "Knowledge- Based Eye Detection for Human Face Recognition," Knowledge-Based Intelligent Engineering Systems and Allied Technologies, 2000. Proceedings, Page(s): 117 -120, 2000