

신경망 기반의 효율적인 얼굴 검출

강창호*, 최종무*, 김문현*

*성균관대학교 전기 전자 및 컴퓨터 공학과

e-mail : {kangch, jmchoi, mhkim}@ece.skku.ac.kr

Efficient face detection based on Neural Network

Kang, chang ho*, Choi, jong moo*, Kim, moon hyun*

*Dept. of Electrical and Computer Engineering, SungKyunKwan Univ.

요약

얼굴 영역 검출은 주어진 영상에서 얼굴의 유·무, 개수 및 위치를 검출하는 것으로 본 논문은 영상에서 얼굴을 검출하는 방법으로 신경망(Neural Network)을 적용하였다. 검출률의 향상 및 오검출률의 감소, 계산량을 최대한 줄이기 위해 후보 영역의 최적화와 얼굴의 대칭성(symmetry of face)을 이용한 좌우 평균 명암도 비교 방법, 평균 얼굴 (average face)을 이용한 템플릿 매칭을 사용하였고, 실험을 통해서 제안한 방법이 효과적으로 수행됨을 보였다.

keyword : average face, log-opponent values, symmetry of face

1. 서론

복잡한 배경을 갖는 영상으로부터의 얼굴 영역 검출은 응용분야가 넓은 기본적인면서도 중요한 문제이다. 얼굴 영역 검출을 위한 영상 처리 방법으로서 명암 영상(gray image)과 색상 영상(color image)을 이용하는 방법을 들 수 있는데, 명암 영상에 기반한 방법은 에지(edge) 검출시 잡음에 민감하고, 복잡한 배경에서 특징 추출이 어렵다는 단점이 있다. 반면 색상 영상은 명암 영상에 비해 보다 많은 정보를 가지고 있기 때문에 보다 효과적으로 특징 추출을 할 수 있는 장점을 지니고 있다[3]. 영상에서 얼굴 영역의 검출은 표정, 연령, 남·여 차이 등으로 인한 얼굴의 다양한 변화의 차이, 크기 등 시점에 따른 변화, 조명의 변화, 안경의 착용 여부, 잡음, 얼굴의 수 등의 다양한 조건들에 독립적인 강인한 방법이 필요하기 때문에 어려운 문제로 여겨져 왔다.

본 논문은 입력된 영상으로부터 얼굴이 어디에, 몇 개 위치하고 있는지를 검출하는 것이다. 검출의 회전 정도는 바로선 얼굴(upright face)로 제한되어 있지만 기존의 제안된 알고리즘에서보다 얼굴 검출의 향상 및 오검출을 줄이고 계산량을 획기적으로 줄이는 것을 목적으로 하고 있다.

본 논문의 구성은 다음과 같다. 다음 장에서는 기존의 얼굴 영역 검출을 위한 방법으로서 PCA 접근방식과 원형 기반의 접근 방식에 대해서 설명한다. 제3장에서는 제안된 알고리즘을 설명하고 제4장에서는

실험방법 및 결과에 대해서 설명한다.

2 기존의 얼굴 영역 검출 방법

임의의 영상에서 얼굴 영역을 검출하기 위한 방법으로 통계적 접근방식중 하나인 PCA와 원형 기반의 접근 방식을 살펴보기로 한다.

2. 1 PCA 기반의 접근 방식

PCA(Principal Component Analysis)는 K-L 변환을 기반으로 얼굴 영상을 영상의 크기보다 적은 수의 벡터로 표현하는 방법이다. Kirby와 Sirovich[5]는 얼굴 영상을 나타내기 위해 PCA를 처음 적용하였고, Turk와 Pentland[6]는 얼굴 영상이 존재하는 공간을 얼굴 공간(face space)라 하였고, 얼굴 공간에 존재하는 얼굴 영상을 고유 얼굴(eigenfaces)라 불리는 기저(basis)로 표현하는 방법을 제안하였다. 이 방법은 훈련 영상 집합을 $X=[X_1, X_2, \dots, X_n]$, 이 훈련 영상에 대해서 평균 영상(Ψ)을 식(1)과 같이 구한 후 각각의 훈련 영상에서 평균 영상을 뺀 차 벡터의 집합을 $A=[\phi_1, \phi_2, \dots, \phi_n]$ 라고 하면, 차 벡터에 대한 데이터 공분산(covariance)은 식(2)과 같다.

$$\Psi = \frac{1}{n} \sum_{i=1}^n X_i \quad (1)$$

$$C = \frac{1}{n-1} \sum_{i=1}^n \phi_i \phi_i^T \quad (2)$$

$$= A A^T$$

C의 고유벡터 z와 고유값 λ는 식(3)에 의하여 구할 수 있다.

$$Cz = \lambda z \quad (3)$$

$$[[C - \lambda I]] = 0$$

C의 고유값을 $\lambda_1, \lambda_2, \dots, \lambda_n$ ($\lambda_1 > \lambda_2 > \dots > \lambda_n$)라고 할 때 이에 대응하는 고유 벡터 (e_1, e_2, \dots, e_n)가 eigenfaces이다. 여기서 구해진 고유벡터 중 m개를 선택해 선형 변환 행렬(W_{pca})을 구하게 된다. 임의의 입력 영상 X는 식(4)에서와 같이 W_{pca} 에 의해 m차원으로 줄어든 벡터 공간 Y로 변환된다.

$$Y = W_{pca} X \quad (4)$$

입력 영상과 식(4)을 이용해 재구성된 영상의 차이를 이용하여 검색대상인 얼굴여부를 최종적으로 판단할 수 있다. PCA를 이용한 방법은 선형 변환된 패턴을 비교하므로 계산 속도가 빠르다는 장점이 있다. 그러나, 조명 변화, 얼굴 표정의 변화에 대해서는 선형 변환된 클래스간의 패턴 벡터의 분리가 이루어지지 않는 단점이 있다[4].

2. 2 원형(Template)기반의 접근 방식

많은 얼굴 검출 알고리즘은 원형 기반의 접근 방식을 사용한다. 이 방법에서는 얼굴의 영상을 화소의 밝기나 색으로 부호화 한 후에, 얼굴 영상 집합의 확률 모형에 의해서 처리되거나, 또는 신경망을 이용한다. 이러한 확률모형의 파라미터들은 예제 영상들로부터 자동적으로 또는 수작업에 의해 조정된다.

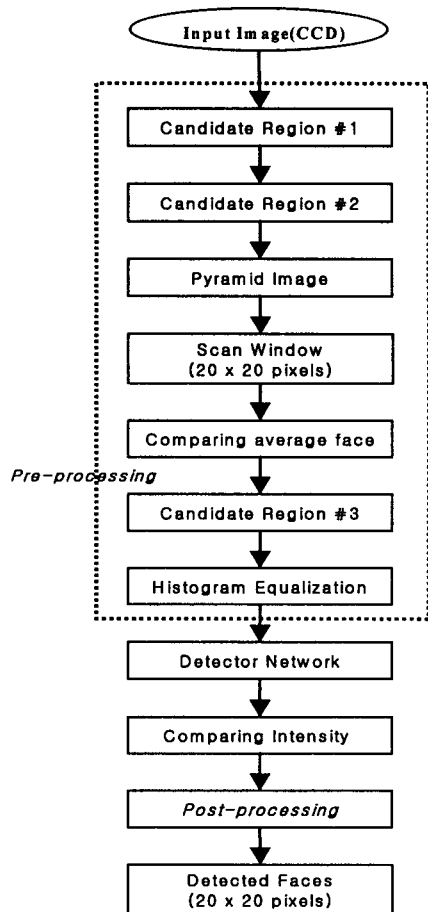
영상에서 얼굴의 위치를 찾는 간단한 방법중의 하나가 얼굴색의 정보를 이용하는 것이다. 이러한 방법에서 얼굴의 색은 가우스 분포를 갖는 정규화된 칼라 공간에서 특징 지워진다. 피부색을 이용하는 방법은 빠른 연산이 가능하고, 머리의 방향과 회전 변화에 강건한 검출을 보장한다는 장점이 있는

반면, 이 방법을 단독적으로 사용하여서는 손이나 팔의 영역과 얼굴을 구분할 수 없다는 단점이 있다.

3. 얼굴 영역 검출

본 알고리즘의 얼굴 검출 방법은 오프라인(off-line)으로 이루어지는 학습 방법과 CCD카메라로 획득한 영상으로부터 얼굴 영역을 검출하는 방법으로 이루어져 있다.

[그림1]의 얼굴 검출 흐름도에서 보여지듯이 CCD카메라로부터 획득한 영상을 전처리 하는 단계, 신명망 적용 단계, 후처리 단계를 거쳐 최종적으로 얼굴 영역을 검출하게 된다.



[그림1] 얼굴 검출 흐름도

3. 1 후보 영역 결정

얼굴 검출을 효과적으로 수행하기 위해서는 입력 영상으로부터 얼굴이 포함된 탐색영역을 적절히 제한하는 것이 필요하다. 본 논문에서는 단지 밝기의 차이만을 이용하는 것이 아니라, 밝기(brightness)

차이와 에지(edge)의 차이를 동시에 고려해 일차 후보 영역을 결정하고, 다시 색상(color)을 이용해 얼굴색(skin color)과 유사한 부분만을 다시 후보 영역으로 선택함으로써 이차적인 후보 영역을 구하게 된다. 색상을 이용한 얼굴색으로의 후보 영역 제한은 Fleck and Forsyth algorithm에 기반 하여 수정된 알고리즘을 적용했다. 이 알고리즘은 RGB color 이미지를 log-opponent (IRgBy)값으로 변환하여 얼굴색 영역을 구한다. log-opponent 값으로의 변환은 식(5)과 같다[2]. 이차적으로 결정된 후보 영역과 평균 얼굴(average face)은 템플릿 매칭 기법을 적용하여 유사성을 비교한다. 유사성의 측정은 속도를 고려하여 L0 norm을 사용하였다. 식(6)의 유사성을 평가하여 평균 얼굴과 가장 가까운 상위 5%를 최종 얼굴 검출 대상 후보 영역으로 최종 선정하게 된다. 이 방법은 효율적인 후보영역을 설정함으로써 계산량이 획기적으로 줄어들어 빠른 시간 내에 얼굴 검출이 가능하게 한다.

$$I_i = [L(R_i) + L(B_i) + L(G_i)] / 3$$

$$Rg_i = L(R_i) - L(G_i) \tag{5}$$

$$By_i = L(B_i) - [L(G_i) + L(R_i)] / 2$$

where

$$L(x) = 105 * \log_{10}(x + 1)$$

$$L0 = \sum_{i=0}^n K_i \tag{6}$$

$$K = 1 \quad \text{if } |P(i) - T(i)| > \theta_1$$

$$K = 0 \quad \text{otherwise}$$

여기에서 P(i)는 i번째 화소이고, T(i)는 템플릿의 i번째 화소이며, θ_1 는 임계치(threshold)값이다.

3. 2 얼굴 검출

전처리(pre-processing) 단계의 일부분인 후보 영역 결정 단계를 거쳐 후보 영역에 대해서만 $n \times n$ 크기(실험에서는 20×20)의 스캔 윈도우(scan window)로 후보 영역에서 이미지를 획득한 후 히스토그램 평활화(histogram equalization)를 적용시킨다[1]. 히스토그램 평활화를 거친 영상은 검출 신경망의 입력으로 들어가게 되고, 검출 신경망은 식(7)에 의해서 얼굴 또는 비얼굴로 판단한다.

$$D = \begin{cases} \text{true} & \text{if } Y \geq \theta_2 \\ \text{false} & \text{otherwise} \end{cases} \tag{7}$$

$$Y = f_2(W_2(f_1(W_1 \cdot X) + b_1) + b_2)$$

여기서 θ_2 는 임계치값이다. 오검출 비율을 줄이기 위해 검출 신경망에서 얼굴로 판단한 이미지에 대해서 식(8)의 영상의 좌우 명암도(intensity)의 평균

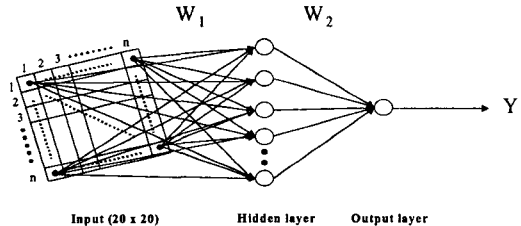
을 비교해 최종적으로 얼굴 여부를 판단하게 된다.

$$CI = \begin{cases} 1 & \text{if } |I_l - I_r| \leq \theta_3 \\ -1 & \text{otherwise} \end{cases} \tag{8}$$

$$I_l = \frac{1}{n \cdot (n/2)} \sum_{i=1}^n \sum_{j=1}^{n/2} X_{ij}$$

$$I_r = \frac{1}{n \cdot (n/2)} \sum_{i=1}^n \sum_{j=(n/2+1)}^n X_{ij}$$

여기서 θ_3 는 임계치값이다. 그리고 입력되는 영상에서 얼굴의 크기가 다를 수 있기 때문에 입력 이미지에 대해서 80%의 크기로 피라미드 이미지를 생성한 후 위의 과정을 반복 수행하게 되며 피라미드 단계는 4단계까지 적용을 하였다. 이 모든 과정을 거친 후, 중복 검출된 부분과 비얼굴을 얼굴로 검출한 부분을 해결하기 위해서 후처리 단계를 거치게 되면 최종적으로 검출된 얼굴을 얻을 수 있다. 검출 신경망의 구조는 [그림2]와 같은 2층 퍼셉트론(Perceptron)의 구조이다. 검출 신경망은 얼굴을 포함하고 있는 영상과 얼굴을 포함하고 있지 않는 영상, 두 종류의 데이터를 가지고 학습하게 된다. 얼굴을 포함하고 있지 않는 영상에 대해서는 랜덤하게 임의의 영상에서 추출하였다.

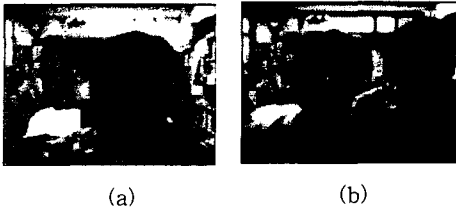


[그림2] 검출 신경망 구조

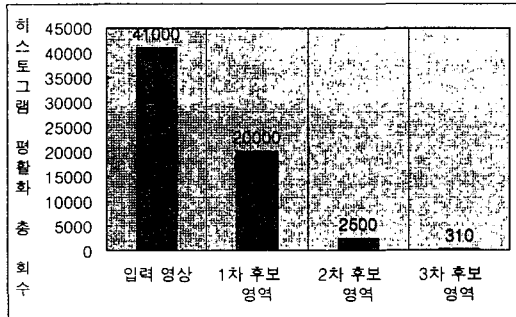
4. 실험 및 결론

영상의 획득은 CCD카메라로부터 입력받은 영상이며, 영상의 크기는 120×160 의 크기이다. [그림3]에 제안된 알고리즘으로 얼굴 검출한 결과를 보여주고 있으며, [그림4]에서는 [그림3]의 영상(a)에 대해서 후보 영역에 따른 계산량을 비교 분석하였고, [표1]은 테스트 영상 30개의 50명의 얼굴에 대해서 각 후보 영역에 따른 검출률과 에러회수를 비교 분석하였다. 제안된 방법은 얼굴의 좌우 명암도 비교를 통해 오검출을 약 46%정도 줄일 수 있었으며, 최적화된 후보 영역을 결정하여 계산량을 줄임으로써 빠른 시간 내에 검출이 가능해졌다.

제안된 알고리즘은 바로 선 얼굴(upright face)에 대해서 검출률의 향상, 계산량 및 오검출의 최소화를 목적으로 하였다. 향후 계획은 회전의 정도가 $\pm 10^\circ$ 인 범위 내에서만 얼굴을 검출할 수 있는 본 알고리즘의 방법을 개선하여 회전에 무관하게 실시간으로 얼굴을 검출할 수 있는 연구가 지속되어야 할 것이다.



[그림3] 얼굴 검출 결과 예



[그림4] 후보 영역에 따른 계산량 비교분석

	detection rate(%)	number of false(# false)
입력 영상	92	110
1차 후보 영역	92	91
2차 후보 영역	92	35
3차 후보 영역	91.3	20

[표1] 후보 영역에 따른 검출률과 에러 회수 비교 분석

참고 문헌

[1] Henry A. Rowley, Shumeet Baluja, Takeo Kanade, "Rotation Invariant Neural Network-Based Face Detection" Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1998-06-23.
 [2] Jay P.Kapur, "Face Detection in Color Images" EE499 Capstone Design Project University of

Washington Department of Electrical Engineering.
 [3] Rafael C. Gonzalez Richard E. Woods, "Digital Image Processing".
 [4] J.Yi, H.Yang, Y. Kim, "Enhanced Fisherfaces for Robust Face Recognition" First IEEE International Workshop, BMCV 2000 Seoul, Korea, May 15-17, 2000 Proceedings, pp 502-511.
 [5] M. Kirby and L. Sirovich, "Application of the Karhunen - Loeve Procedure for the Characterization of Human Faces." IEEE Trans. on PAMI, vol. 12 no 1, pp. 103-108, 1990.
 [6] M. Turk and A Pentland, "Eigenfaces for Recognition." Journal of Cognitive Neuroscience, vol. 3 no. 1, pp. 71-86, 1991.