

# 경험적 지식과 신경망을 이용한 얼굴영역 검출

서원택<sup>✉</sup>, 조범준  
조선대학교 컴퓨터공학과

## Face detection using heuristic knowledge and neural network

Wontaek Seo, Beomjoon Cho  
Dept. of Computer Science Engineering

### 요 약

본 논문은 그레이 영상에서의 사람얼굴 영역추출에 관해서 연구하였다. 얼굴영역 추출은 얼굴인식이나 사람과 컴퓨터의 인터페이스, 비디오 감시시스템을 연구하는데 있어서 반드시 거쳐야 하는 전처리 과정이라고 할 수 있다. 이러한 목적을 위해서 본 연구에서는 두 단계의 과정을 통해서 얼굴영역을 추출하였다. 첫 번째 단계는 사람얼굴에 대한 경험적 지식을 이용하여 후보영역을 획득한 다음에 두 번째 단계에서 후보영역을 웨이블릿 분해 후, 신경망을 이용하여 후보영역 중에서 얼굴영역을 검증한다. 실험결과 제안한 방법은 빠르고 정확하게 얼굴영역을 검출하였다.

### 1. 서 론

사람 얼굴영역 검출에 관한 연구는 얼굴인식이나 얼굴검증, 비디오 감시, 사람과 컴퓨터간의 상호작용을 연구하는 밀바탕 연구로서 많은 연구들이 이루어졌고, 지금도 활발하게 연구되고 있다. 이는 얼굴인식의 연구에서는 정확한 얼굴영역의 선택에 따라 인식의 정확도가 매우 큰 차이를 보이고 있기 때문이다. 지금까지의 얼굴인식에 관한 연구를 할 때, 얼굴영역 추출에 대한 정형화된 방법이 없었기 때문에 연구자들은 그들의 임의적으로 직접 얼굴영역을 선택할 수밖에 없었다. 이런 방법은 시스템에 의한 자

동얼굴인식과는 거리가 먼 방법이고, 또한 정확도를 보장할 수 없는 방법이었다. 얼굴영역 검출에 관한 연구가 중요하게 생각되는 부분도 바로 이런 이유 때문이다.

얼굴영역을 검출은 다양한 얼굴표정, 빛의 방향, 다양한 머리모양, 피부색, 얼굴의 방향, 복잡한 배경과 같은 여러 가지 요인 때문에 어려움을 겪고 있다. 그래서 제한적인 방법으로 많은 실험들이 이루어지고 있는데, 본 논문에서도 거의 정면을 향하고 있는 얼굴을 대상으로 실험을 하였다.

얼굴영역 검출에 관련된 연구는 얼굴의 각 요소의 위치정보를 이용한 방법[1][2]과 피부색 정보를 이

용한 검출방법[3][4]과 신경망을 이용한 방법[5] 등으로 구분해 볼 수 있다. 얼굴 요소의 위치정보를 이용한 방법의 단점은 주로 정면얼굴 부분만을 다룬다는 것이다. 또한 다양한 얼굴의 표정이나 액세서리, 콧수염과 같은 요인에 영향을 많이 받는다는 것이다. 피부색을 이용한 검출방법은 조명의 영향을 많이 받는다는 것과, 인종간의 차이로 선택의 폭이 넓지 않는다는 것에 있다.

본 논문에서는 두 가지 방법을 이용하여 얼굴영역을 검출하였다. 첫 번째 단계에서는 얼굴에 대한 경험적 지식을 이용하여 이미지에서 후보영역을 선택한 다음에 두 번째 단계에서 신경망을 이용하여 후보영역 중에서 얼굴영역을 선택한다. 2 장에서는 첫 번째 단계인 얼굴영역에 대한 지식적인 경험을 통해서 후보영역을 선택하는 방법을 설명하고, 3 장에서는 두 번째 단계로 웨이블릿 변환과 신경망을 이용한 얼굴영역 선택과정을 설명한다. 4장에서는 실험결과를 밝히고 5장에서 결론으로 맺는다.

## 2. 경험적 지식을 이용한 빠른 후보영역 검출

사람의 눈은 얼굴의 각 요소 중에서 가장 중요한 부분으로 감정을 잘 표현할 수 있고 얼굴영역 추출 연구에서 많이 사용되는 부분이다. 그 이유는 입과는 달리 상황에 따라 위치정보나 모양이 크게 달라지지 않기 때문이다. 이러한 이유로 본 논문에서도 눈의 특징을 이용하였다. 하지만 높은 해상도의 사진에서 눈 영역을 찾기 위해서는 시간이 오래 걸리게 된다. 따라서 본 논문에서는 원본 사진의 해상도를 낮추어서 빠르게 검색하였다.

### 2.1 경험적 지식

“사람의 얼굴에는 눈이 두 개이고, 눈과 눈 사이에는 콧등이 존재한다. 동일한 조명을 받은 얼굴에서 눈의 명도는 눈과 눈 사이의 공간의 명도보다 어둡다.”라는 경험적 지식을 바탕으로 본 논문에서는 빠른 눈 영역 추출을 시도하였다.

위의 지식을 구현하기 위해 먼저 몇몇의 얼굴 이

미지의 사이즈를 세가지 방법을 이용하여 줄였다. 그림 1은 원본이미지에 ‘nearest’ 방법을 적용하여 히스토그램 평균화를 한 영상이다.

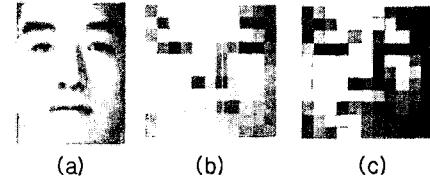


그림 1. (a) 원본이미지 (b) 11×11로 줄인이미지  
(c) 히스토그램 평활화를 적용한 이미지

사이즈 별로 살펴본 결과 11×11정도의 이미지로도 충분히 눈 영역을 쉽게 구분할 수 있다는 결론에 도달하였다.

### 2.2 후보영역추출

보통 이미지에서 얼굴영역의 크기는 상당히 다양하다. 그래서 전체적인 이미지에서 얼굴영역이 11×11정도의 틀에 들어갈 수 있도록 줄이는 비율을 결정을 해야 하는데 실형데이터를 통해서 비율을 결정하였다. 그런 다음 작은 해상도의 이미지에서 눈과 눈 사이의 거리, 눈의 높이, 눈높이와 눈 사이의 거리의 비를 조건으로 후보영역을 선택하였다.

### 2.3 두 번째 경험적 지식

사람의 얼굴에는 일반적으로 정해진 요소들이 모두 들어있고 모양도 거의 비슷하고 차지하고 있는 비율도 거의 일정하다. 때문에 “그레이 영상의 얼굴 영역을 히스토그램을 표현해보면 그 모양이 거의 일정하다”는 것을 알 수 있다. 위와 같은 경험적 지식을 바탕으로 후보영역 중에서 얼굴영역이라고 예상되는 부분을 상당부분 줄일 수 있었다.

먼저 히스토그램 영역을 정규화 시켜서 정규화 시킨 히스토그램을 비교하여 결정하였다. 이를 위해 수작업으로 30개의 얼굴이미지를 선정하여 히스토그램을 구한 후 평균을 구하여 실험 이미지의 히스토그램과 비교하였다.

아래의 그림은 명도가 다른 두 그림의 히스토그램을 비교해 보았다. 두 이미지의 히스토그램의 모양이

많이 비슷한 것을 알 수 있다.

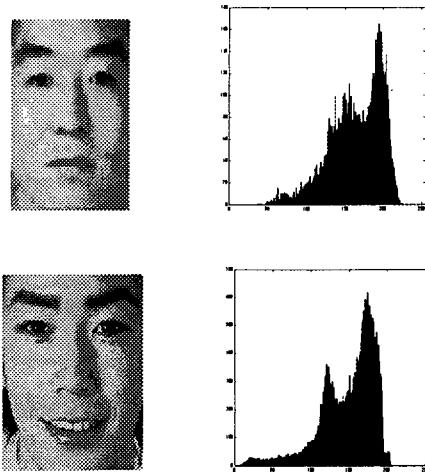


그림 2. 명암이 다른 두 이미지간의 히스토그램 비교

### 3. 신경망을 이용한 검증

일반적인 사람의 얼굴은 두 개의 눈썹, 두 개의 눈, 코, 입의 6개의 요소를 가지고 있다. 이와 같은 요소들은 이미지의 상태에 따라서 어느 특정부분은 나타나지 않거나 혹은 가려질 수 있다. 이런 점을 고려해서 다음과 같은 모델로 나타낼 수 있다.

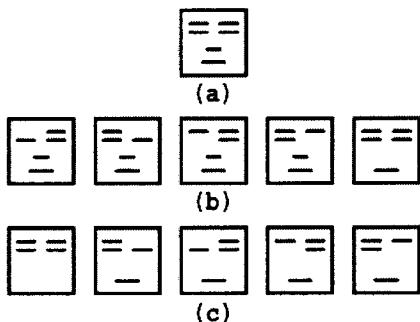


그림 3. 각 요소의 위치에 따른 얼굴 모델 (a) 6개의 요소 (b) 5개의 요소 (c) 4개의 요소

위의 모델을 이용하여 신경망을 학습시키고 후보 영역 중에서 얼굴영역을 검출해 낼 수 있었다.

### 3.1 웨이블릿 분해를 이용한 특징점 추출

웨이블릿은 S. Mallat 가 multiresolution analysis (MRA)에 대한 연구결과를 발표하면서부터 컴퓨터 비전, 이미지 프로세싱 영역에서 많이 사용되고 있다. 또한 얼굴영역에 웨이블릿을 이용하여 특징을 추출하는 방법도 현재 많이 이용되고 있다. 본 연구에서는 이 방법을 이용해서 얼굴영역에서 눈, 코, 입, 눈썹의 유무를 파악하였다.

MRA의 이론에 의하면, 이미지  $f(x,y)$ 를 MRA를 이용하여 분해한 것은 필터  $h(\cdot)$ 와  $g(\cdot)$ 를 적용하여 분해한 것과 같다는 것을 알 수 있다. 다음 그림은 MRA를 이용한 분해 과정을 나타낸 것이다.

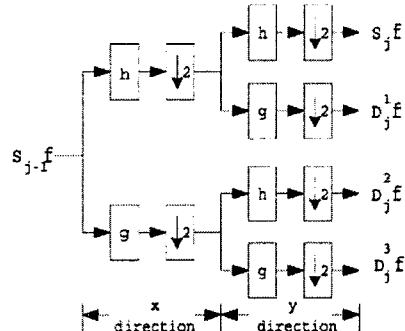


그림 4. MRA를 이용한 이미지 분해

$h(\cdot)$ 는 저주파 필터이고,  $g(\cdot)$ 는 고주파필터이다.  $\downarrow$  레벨에서의 이미지 분해과정을 다음과 같은 식으로 이루어진다.

$$\begin{aligned} S_j f &= [h * [h * s_{j-1} f]_x \downarrow 2]_y \downarrow 2, \\ d_j^1 f &= [g * [h * s_{j-1} f]_x \downarrow 2]_y \downarrow 2, \\ d_j^2 f &= [h * [g * s_{j-1} f]_x \downarrow 2]_y \downarrow 2, \\ d_j^3 f &= [g * [g * s_{j-1} f]_x \downarrow 2]_y \downarrow 2, \end{aligned}$$

위 식에서  $*$ 는 컨벌루션이고,  $\downarrow 2$ 는  $x$ 와  $y$  방향으로 샘플링하는 것이다.

위의 과정을 수행할 수 있는  $h(\cdot)$ ,  $g(\cdot)$  필터는 다음과 같은 것을 사용하였다.

$$\begin{aligned} H(z) &= 0.853 + 0.377(z + z^{-1}) - 0.111(z^2 + z^{-2}) \\ &\quad - 0.024(z^3 + z^{-3}) + 0.038(z^4 + z^{-4}) \\ G(z) &= -z^{-1}H(z^{-1}) \end{aligned}$$

위의 필터를 적용하여 윤곽선을 추출한 후에 임계값을 이용하여 이진 이미지로 변환시켰다. 이렇게 변환된 이진이미지를 오류역전파 신경망에 입력하여 얼굴영역인지를 결정하였다.

### 3.2 오류역전파 신경망

본 연구에서는 3계층 오류역전파 신경망을 사용하여 후보영역 중에서 얼굴영역을 판정하였다. 신경망을 선택한 이유는 신경망의 특성상 어느 정도의 위치적인 차이가 있어도 인식이 가능하다는 점과, 일단 훈련이 완료된 신경망은 빠른 출력을 얻을 수 있다는 장점이 있기 때문이다.

본 실험에서 사용된 신경망의 구성은 입력층에 121(11×11)개의 유닛, 중간층에 6개의 유닛, 출력층에는 1개의 유닛으로 하였다.

신경망을 훈련시키기 위해서 수작업으로 추출된 얼굴영역 이미지 50개와 얼굴이 아닌 이미지 30개를 이용하였다. 신경망의 출력은 얼굴이미지(1), 얼굴이 아닌 이미지(0)로서 두 개의 출력이었다.

## 4. 실험 결과

본 연구에서는 Pentium 4 PC를 이용하였고, Matlab 6.5 tool을 이용하였다. 실험에 사용된 이미지는 직접 촬영한 얼굴영상 20개와 인터넷에서 수집한 얼굴영상 30개, 비 얼굴영상 30개를 이용하였다. 사용된 대부분의 얼굴이미지는 동양인이었고, 사용된 이미지의 얼굴영역의 크기는 크게 차이가 나지 않는 것을 선택하였다.

실험결과는 다음표에 나타내었다.

종류	얼굴수	인식수	인식률	오인식률
단일얼굴	30	26	87%	13%
다중얼굴	48	39	81.2%	18.2%

## 5. 결론

본 논문에서는 사람얼굴에 대한 두 가지 경험적 지식을 바탕으로 얼굴영역을 추출하는 방법을 연구

하였다. 첫 번째 “사람의 얼굴에는 눈이 두 개이고, 눈과 눈 사이에는 콧등이 존재한다. 동일한 조명을 받은 얼굴에서 눈의 명도는 눈과 눈 사이의 공간의 명도보다 어둡다.”라는 지식으로 많은 후보영역을 선정할 수가 있었고, 두 번째 “그레이 영상의 얼굴영역을 히스토그램을 표현해보면 그 모양이 거의 일정하다.”라는 지식으로 후보 영역중에서 상당수를 제거할 수 있었다. 마지막으로 웨이블릿 분해를 한 이미지로 신경망을 이용해서 최종 얼굴영역을 검증하였다. 본 논문이 제안한 방법은 단일 얼굴에서는 87%의 인식률을 나타내었고, 다중얼굴영역에서는 81.2%의 인식률을 보였다. 오인식 된 얼굴영역을 살펴보면 주로 기울임이 있는 얼굴영역들이었다. 이에 본 연구를 더욱 발전시켜 기울임과 회전이 된 얼굴영역도 추출할 수 있는 강인한 방법을 연구하여야 하겠다. 또한 본 연구를 좀더 발전시켜 인식시스템에 응용할 수 있는 길을 모색해야 하겠다.

## 참고문헌

- [1] G.Z. Yang, T.S. Huang, Human face detection in a complex background, Pattern Recognition 27 (1) (1994) 43-63.
- [2] J. Miao, B. Yin, K. Wang, L. Shen, X. Chen, a hierarchical multiscale and multiangle system for human face detection in a complex background using gravity-center template, Pattern Recognition 32 (7) (1999) 1237-1248
- [3] K. Sobottka, I. Pitas, A novel method for automatic face segmentation, facial feature extraction and tracking, Signal Process. Image commun. 12 (3) (1998) 263-281
- [4] C.Garcia, G. Tziritas, Face detection using quantized skin color regions merging and wavelet packet analysis, IEEE Trans. Multimedia 1 (3) 1999) 264-277
- [5] H.A. Rowley, S. Baluja, T. Kanade, Neural network-based face detection, IEEE Trans. Pattern Anal. Mach. Intell. 20 (1) (1998) 23-38.