

## 윈도우를 사용한 얼굴영역의 추출 기법

임 성 현, 이 철 희  
연세대학교 전기전자공학과  
Email: chulhee@yonsei.ac.kr

### A Face Segmentation Algorithm Using Window

Sunghyun Lim, Chulhee Lee  
Dept. Electrical and Electronic Engineering, Yonsei University  
Email: chulhee@yonsei.ac.kr

#### Abstract

In this paper, we propose a region-based segmentation algorithm to extract human face area using a window function and neural networks. Furthermore, we apply the erosion and dilation to remove small error areas. By applying the window function, it is possible to reduce error. In particular, false segmentation of the eye and the lip can be considerably reduced. Experiments show promising results and it is expected that the proposed method can be applied to video conference and still image compression.

#### I. 서론

최근 영상에서 관심 영역 추출에 관한 연구가 여러 분야에서 진행되고 있다. 그 중에서도 인간의 얼굴 추출은 영상에서 얼굴의 중요성을 감안할 때 활용도가 매우 광범위하다. 즉, 영상에서 얼굴을 배경과 분리하는 기술은 얼굴영역의 처리, 인식, 추적에 있어 필수적인 핵심기술이다. 또한 영상 압축에서 얼굴 부분을 추출하여 보다 높은 비트를 할당한다면 제한된 비트수에서 정보를 효율적으로 전송할 수 있다.

최근 컬러의 색 정보를 이용하여 얼굴 영역을 검출하는 기법이 여러 연구자에 의해 제안되었다 [1][2][3]. 이 기법은 많은 연산량을 필요로 하는 기존의 구조체 정합 방식에 비해 빠르고 적은 메모리를 요구한다. 또한 신경망이나 또는 CrCb성분의 문턱치를 이용하여 피부 칼라 맵을 만들고 후처리과정을 통하여 원치 않는 잡음 성분을 제거 할 수 있다 [1][2]. 그러나, 얼굴의 눈이나 입 같은 정보는 피부와는 다르므로 피부 칼라 맵을 적용하면 배경으로 인식되는 경우가 있다. 이와 같이 배경 영역의 크기가 한계치를 초과하면 얼굴 추출에 있어서 상당한 문제를 발생시킬 수 있다. 이러한 문제를 해결하기 위하여 본 논문에서는 영상에 윈도우를 적용하여 눈이나 입 등에서 발생하는 잡음 성분을 줄임으로써, 후에 추출과정에서 얼굴영역에서 제외되는 결과를 방지하였다.

#### II. 신경망에 의한 얼굴 후보 추출

##### 2.1 얼굴 영역의 추출 방법

얼굴 영역을 추출하는데 있어 사용할 수 있는 특징에는 여러 가지가 있다. 얼굴의 모양이나 얼굴의 구성요소(눈 코 입 등), 그리고 얼굴의 색 정보 등을 얼굴을 구분할 수 있는 특징으로 사용할 수 있다. 얼굴의 형태를 이용하여 얼굴을 찾는 구조체 정합 방식이 오래 전부터 많이 사용되었다 [4]. 이는 얼굴의 원형을 모델

링하고 이를 주어진 영상에 적용하여 유사도를 계산하여 얼굴 영역을 추출하는 기법이다. 그러나, 실제 적용에 있어 얼굴의 크기를 알 수 없기 때문에 원형의 얼굴을 다양한 크기로 확대, 축소하여 검색하여야 하는 문제가 발생한다 [4]. 따라서 많은 처리 시간과 큰 메모리가 요구된다. 또한 눈, 코, 입 사이의 거리를 모델링하여 얼굴을 추출하는 방법이 있는데 이 방법 또한 많은 처리 시간이 요구된다 [5]. 그리고 칼라정보를 이용하는 방법이 제안되었는데 [1][2][3], 이 기법은 사람의 얼굴과 어깨로 구성되는 비디오 폰 영상과 같이 얼굴 이외의 색정보가 많이 포함하지 않는 경우에 우수한 추출 결과를 보인다. 또한 다른 방법에 비하여 처리시간이 짧아 최근에 많이 이용되는 추출 방식이며, 본 논문에서는 이 방식을 사용하였다.

## 2.2 신경망을 이용한 얼굴 후보영역 추출

사람의 얼굴 영역은 YCrCb 공간에서 일정한 패턴으로 분포되어 있다 [3]. 그림 1은 YCrCb 공간에서 사람의 얼굴의 분포도를 보여 준다.

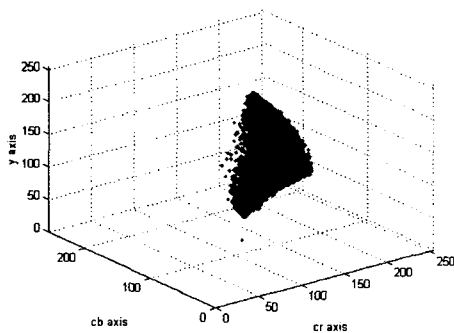


그림 1. YCrCb 공간에서의 얼굴 분포도.

그림 1에서 볼 수 있듯이, 얼굴 영역은 상당히 밀집된 분포를 가진다. 그러나, 배경은 영상에 따라 전혀 다른 분포를 가질 수 있다. 특히 배경이 얼굴과 유사한 분포를 가질 경우, 색 정보를 이용하여 얼굴 영역을 추출하는 것은 거의 불가능하다. 그러나 대부분의 영상의 경우에 배경은 YCrCb 공간에서 넓은 영역에 분포하기 때문에 얼굴영역과 겹치는 부분은 일부에 지나지 않는다. 색 정보를 이용하여 얼굴 후보 영역을 추출한 후, erosion과 dilation 과정을 통하여 중복되는 부분을 제거할 수 있다 [1].

본 논문에서는 erosion과 dilation 과정 전단계로 컬러

정보를 이용하여 얼굴 영역을 분리하기 위해 신경망을 사용하였고, 30개 영상의 얼굴영역과 배경영역을 분리한 후, 이를 학습데이터로 사용하였다. 학습 알고리즘은 backpropagation algorithm을 사용하였다. 그림 2는 학습된 신경망을 사용하여 학습에 사용하지 않은 테스트 영상을 분류한 결과를 보여준다.



그림 2. 신경망을 적용한 후의 영상.

그림 2에서 검은색 부분은 얼굴 영역으로 인식한 부분이고 흰색 부분은 배경영역으로 인식한 부분이다. 위의 그림을 보면 눈과 입 부분이 얼굴영역으로 인식되어 있지 않다. 후처리 과정으로 erosion과 dilation을 적용하면 이 부분은 얼굴 영역으로 편입될 수 있는데 만일 얼굴에서 인식이 안된 부분의 크기가 한계치를 초과하면 얼굴로 편입되지 못하고 에러가 확산되는 결과를 초래한다. 이와 같은 문제는 테스트 영상에 윈도우를 적용함으로써 얼굴에서 인식이 안된 부분의 크기를 감소시키게 되어 방지될 수 있다.

## III. 테스트 영상에 윈도우 적용

얼굴을 추출하고자하는 영상에 임의의 모양의 윈도우를 적용하게 되면 blurring 효과가 나타난다. 이는 얼굴에서 비교적 적은 영역을 갖는 눈과 입의 색이 그 주위의 피부색과 유사하게 될 수 있음을 의미한다. 따라서 신경망을 적용하여 얼굴 후보 영상을 얻었을 때 눈과 입을 얼굴 영역으로 인식하게 되고 나중에 이들이 얼굴 영역에서 제외되는 결과를 방지할 수 있다. 본 논문에서는 원형의 윈도우를 사용하였다. 윈도우의 지름을 크게 하면 눈과 입이 제외되는 에러를 제거할 수 있으나 배경 영역에서 얼굴로 인식하게 되는 확률

이 높아지고, 얼굴과 인접하는 배경영역이 얼굴영역으로 편입되는 문제가 발생한다. 실험 결과 얼굴 길이의 약 1/10정도의 윈도우를 사용하였을 때 가장 좋은 결과를 얻었다.

#### IV. 얼굴 후보로부터 얼굴 추출

##### 4.1 Erosion과 dilation 과정을 통한 밀도정규화

이 과정에서 잡음에 의해서 제외된 얼굴 영역을 복원시키고 얼굴로 오인식된 배경영역을 제외시킨다. 우선적으로 erosion과 dilation 처리를 쉽게 하기 위하여 테스트 영상을 원영상의 1/16으로 축소하였다. 즉, 4×4 픽셀을 이들의 합의 값을 갖는 1 픽셀로 전환하는데 얼굴 영역으로 인식된 픽셀에 대해서는 1을, 배경으로 인식된 픽셀에 대해서는 0을 더하여 준다. 이 과정을 수식으로 나타내면 다음과 같다.

$$M(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 O_1(4x+i, 4y+j)$$

여기서  $O_1$  은 테스트 영상을 의미하고  $M$ 은 density map을 의미한다. 예로 만일 4×4 픽셀들이 모두 얼굴로 인식되었다면  $M(x, y)$ 는 16이 된다. 따라서  $O_1$  이  $M \times N$ 의 해상도였다면  $M$ 은  $M/4 \times N/4$ 의 해상도를 갖게 된다.

이와 같이 축소된 영상에 erosion과정을 적용하여 얼굴로 오인식된 배경 영역을 제거한다. 해당 픽셀의 3×3 neighbors 에 대해서  $M(x, y)$ 가 16인 픽셀 수가 4개 이하이면  $M(x, y)$ 를 0으로 치환한다. Dilation 과정에서는  $M(x, y)$ 가 16이 아닌 점에 대해서 만약 3×3 neighbor 중에 16이 되는 점이 2개 이상이라면  $M(x, y)$  를 16으로 치환한다. 그리고 erosion 과정에서는 6개 이하로, dilation 과정에서는 5개 이상으로 threshold를 조정된 후에 같은 과정을 반복한다. 이러한 과정을 거친 후, 결과 영상을 그림 3에 나타내었다.

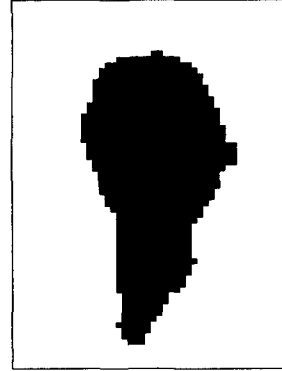


그림 3. Erosion과 Dilation 과정을 거친 후의 영상.

##### 4.2 위치상의 특성을 이용한 배경 영역 제거

이 과정에서 얼굴 영역으로 오인식된 배경을 제거하며, 그 방법은 다음과 같다.

- 1) 먼저 수평 방향에 대해서 영상을 스캔해나가면서 만약 얼굴로 인식된 점들에 대해서 해당 픽셀을 포함하여 연속하여 4개의 픽셀이 얼굴로 인식되었으면 해당 픽셀을 얼굴영역으로 포함시키고, 그렇지 않으면 배경으로 편입시킨다.
- 2) 수직 방향에 대해서도 동일한 방법을 적용한다.

만일 얼굴 전 영역이 얼굴로 인식되지 않았다면(눈이나 입 등을 모두 얼굴 영역으로 인식하지 못한 경우) 위의 과정은 에러를 확산시킬 수 있다. 그러나 본 논문에서 제안한 윈도우 영상을 적용하였을 때 크게 개선된 결과를 얻었다.

##### 4.3 윤곽선 검출을 통한 해상도의 회복

4.2에서 얻은 영상에서 얼굴영역의 경계 부분에 대해서 2.2의 신경망을 적용하여 얻은 영상의 결과를 대입함으로써 원래의 해상도를 회복할 수 있다.

## V. 실험결과

학습에 사용하지 않은 20개의 영상을 테스트 영상으로 실험한 결과 모든 경우에서 얼굴 영역을 검출하였다. 그러나 일부의 영상의 경우에 배경 영역이 얼굴 영역에 포함되거나 배경 영역이 새로운 얼굴영역으로 추출되는 문제가 발생하였다. 얼굴 영역 추출 과정에서 윈도우를 적용한 경우는 눈 영역을 정확히 검출하여 얼굴 영역의 왜곡은 없었으나 일부 영상의 경우 윈도우를 적용하는 과정에서 원치 않은 배경이 얼굴에 편입되는 결과가 발생했다.

20개의 영상을 실험한 결과 윈도우를 적용하지 않았을 때는 8개의 영상에서 눈 영역에서 에러가 발생하였으나, 윈도우를 적용한 경우에는 3개의 영상에서만 에러가 발생하였다.

## VI 결론

얼굴 영역 검출은 영상에서 얼굴 영역을 처리하고, 인식하며, 추적하는데 필수적인 기술이다. 또한, 이미지 전송에 있어서 실시간 객체 기반의 코딩을 통한 통신이 가능하기 위해서는 고속 얼굴 추출 방법이 필요하다.

얼굴 영역 추출은 유사한 개념인 얼굴 검출이나 인식과는 달리 얼굴의 경계와 내부의 모든 점을 찾아 내야 하는 난점이 있다. 그런데, 눈이나 입과 같은 부위는 컬러 정보를 이용하여 얼굴 영역을 추출할 때 제외된다. 이들로 인하여 발생하는 에러가 적은 경우 erosion과 dilation과 같은 영상 기법을 통하여 나중에 제거될 수 있으나, 에러가 큰 경우 에러가 확산되는 결과가 발생한다. 제안한 윈도우 기법을 적용한 경우 약 15% 정도만 제외하고 에러가 발생하지 않았다.

## 참고문헌

- [1] Douglas Chai, King N. Ngan. "Face Segmentation Using Skin-Color Map in Videophone Applications," 1999.
- [2] Hideaki Ishii, Minoru, Norio Akamatsu, "Face detection based on skin color information in visual scenes by neural networks," 1999.
- [3] Christophe Garcia, Georgios Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis" 1999.
- [4] Roberto Brunelli, Tomaso Poggio, "Face

Recognition: Features versus Templates" 1993.

- [5] J. Cox, Y. J. Ghosen, and P. Yianilos, "Feature-based face recognition using mixture distance," 1996