

# 수행 시간과 정확도를 고려한 얼굴 영상의 눈 영역 추출

장창혁, 박안진, 정기철  
송실대학교 IT대학 미디어학과  
{luee,anjin,kcjung}@ssu.ac.kr

## Extraction of Eye Region in Consideration for Performance and Accuracy

Changhyuk Jang, Anjin Park, Keechul Jung  
HCI Lab., School of Media,  
College of Information Technology, Soongsil University

### 요 약

얼굴 인식의 전처리 단계로써 얼굴의 특징 영역인 눈, 코, 입을 추출하는 방법들이 최근 다양하게 연구되고 있다. 얼굴 영상의 특징 영역을 추출 하는 방법에는 일반적으로 특징 점을 이용한 방법과 에지 정보를 이용한 방법이 있다. 특징 점을 이용한 방법은 높은 정확도를 보이는 반면 느린 수행시간을 보이는 문제점이 있으며, 에지 정보를 이용한 방법은 빠른 수행시간을 보이지만 정확도가 떨어지는 문제점이 있다. 본 논문에서는 정확도와 수행시간을 동시에 향상시킬 수 있는 방법을 제안한다. 빠른 수행 시간을 위해 에지 정보와 에지의 방향성 정보를 이용하여 대략적으로 영역을 추출하며, 잡음에 의해 발생된 에지나 빛에 의해 추출되지 못한 에지에서 생긴 눈 추출의 오류는 추출된 영역의 가로, 세로 비율과 각 영역의 공간 정보를 이용하여 해결한다. 실험 결과에서 85%의 정확도와 평균 0.3초의 수행 시간을 보였으며, 에지 정보를 이용한 방법의 문제점인 정확도와 특징 점을 이용한 방법의 문제점인 수행시간을 동시에 향상시킨 결과를 보였다.

### 1. 서론

얼굴 또는 표정 인식은 최근 영상 인식 분야에서 중요한 이슈로 연구되고 있다. 공상 과학 영화에서 사람의 얼굴을 인식하는 시스템을 흔히 볼 수 있는 것은 얼굴 인식 분야의 연구 진행 과정을 쉽게 알 수 있는 예라 할 수 있다. 특정 인물의 정보를 이용하여 누구인지를 확인할 수 있는 얼굴 인식 분야는 공항에서 사람의 영상 정보를 이용한 보안 절차에서 활용되고 있다. 사람의 다양한 감정을 구분하는 표정 인식은 특징적인 얼굴의 형태를 통해서 슬픔, 기쁨, 놀람, 화남 등 다양한 표정을 인식 혹은 표현하며, 최근 컴퓨터 그래픽스와 HCI(Human Computer Interaction) 분야의 발달로 3D아바타 기술에서 사람의 감정을 표현 할 때 표정 인식 기술이 활용되고 있다. 이러한 얼굴 또는 표정 인식 분야에서 얼굴의 특징적인 정보를 얻을 수 있는 영역은 눈, 코, 입이

며 가장 중요한 정보를 가지고 있는 것이 눈이다[1]. 영상에서 눈 영역을 추출하는 방법에는 영역의 특징 점을 이용하여 추출하는 방법[2-3]과 에지 정보를 이용하는 방법[4-6]등이 있다. Pantic과 Rothkrantz [2]는 룰-기반 리저닝(rule-based reasoning) 방법을 이용하여 특징 점과 영역을 추출하는 방법을 제안 하였으며, Batur과 Monson[3]는 액티브 어피어런스 모델(active appearance models)을 바탕으로 특징 점을 위한 정보를 추출하는 방법을 제안하였다. 특징 점을 이용하여 얼굴의 특징 정보를 추출하는 방법은 정확성이 매우 높은 장점이 있지만, 수행 시간이 오래 걸리는 문제점이 있다[2-3]. 빠른 수행 시간을 위해 Zuo와 Peter[4]는 에지 정보를 이용한 특징 정보 추출 방법을 제안하였으며, Gizatdinova 등[5]은 에지 검출과 에지의 기울기 벡터 정보를 이용한 영역 추출 방법을 제안 하였고. Son과 Mital[6]

는 민-필드 방법(mean-field method)을 이용한 영역 추출 방법을 제안 하였다. 하지만 영상에서 에지 정보를 통해 영역을 추출할 때 눈 주변에 약간의 노이즈가 포함될 경우와 빛에 의해 에지가 검출 되지 않을 경우 눈 영역 추출에 대해 낮은 정확도를 보였다. 이러한 문제점을 개선하기 위해 Song 등[1]은 웨이블릿 변환(wavelet transform)을 이용하여 보다 정확한 영역 추출 방법을 제안하였지만 수행 시간이 오래 걸리는 문제가 있다.

본 논문에서는 지역적인 에지 기반의 추출 방법과 전역적인 위치 기반의 추출 방법을 결합하여 수행 시간과 정확도를 향상하는 눈 영역 추출 방법을 제안한다. 수행 시간의 효율성을 위해 검출된 에지의 기울기 벡터 정보를 이용하여 대략적으로 눈 영역을 추출하며, 정확성 향상을 위해 추출된 영역의 가로, 세로 비율과 공간 정보의 관계성을 이용하여 오 추출(false)된 영역과 추출되지 않은(miss) 영역을 해결한다. 본 논문에서 제안한 방법은 평균 0.3초의 빠른 수행 시간과 평균 85%의 정확도를 보이며, 에지 정보를 이용한 방법의 문제점인 정확도와 특징 점을 이용한 방법의 문제점인 수행 시간을 동시에 향상시킨 결과를 보인다.

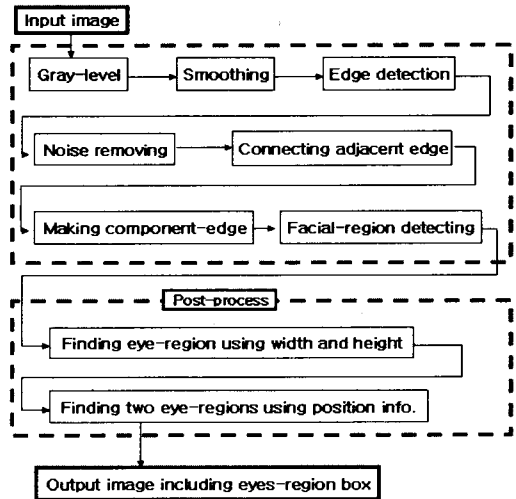
본 논문의 구성은 다음과 같다. 2장에서는 영상에서 눈 영역을 추출하기 위한 에지 기반의 방법을 설명하며, 3장에서는 위치 정보를 이용한 눈 영역 추출 방법을, 4장에서는 실험 결과를 보여 준다. 끝으로 5장은 결론 및 향후 과제를 기술한다.

2. 눈 영역 추출 구성

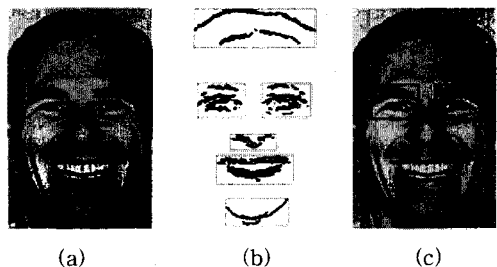
영상에서 에지 정보를 이용하여 영역을 추출하는 방법으로 Gizatdinova 등[5]은 에지의 기울기 벡터 정보를 이용한 영역 추출 방법을 제안하였다. 이 방법은 가까운 에지 점을 연결하여 하나의 구성요소(component)를 만들고, 각 구성요소의 기울기 벡터가 수평 방향에서 최대값일 때 눈, 코, 입 영역으로 추출되며, 에지의 기울기 정보만을 사용하기 때문에 가장 빠른 수행 시간을 보였다. 그러나 이 방법은 영상의 잡음 또는 조명의 차이로 노이즈 정보가 포함될 경우 정확도가 떨어지는 단점이 있다.

(그림 1)은 제안된 방법의 전체 구성도를 보여 주며 크게 두 가지 단계로 구성되어 있다. 첫째, 에지 정보를 이용하여 대략적으로 특징 영역을 추출한다. 에지를 추출하기 위한 전처리 단계로서 입력 칼라 영상을 그레이 영상으로 변환한 후 잡음 제거를 수

행 하며, 에지 방향이 수직일 경우 눈 영역 추출에서 무의미한 데이터로 간주 하여[5], 기울기가 -100에서 -80, 80에서 100사이에서 에지를 추출한다. 추출된 에지에서 대략적인 영역을 만들기 위해 가까운 에지 점을 연결하고 연결 구성요소(connected component)를 이용한다. 둘째, 위치 정보를 이용하여 눈 영역을 추출한다. 먼저 가로, 세로 비율을 이용해서 눈 영역을 추출하여 오 추출에 대한 문제점을 해결하고, 공간 정보의 관계성을 이용하여 대략적인 영역 추출 단계에서 추출되지 않은 눈 영역을 추출한다. (그림 2)는 [5]의 실험에서 사용한 영상을 적용한 것으로 에지 검출에 의한 기울기 벡터 정보와 위치 정보를 통해 두 개의 눈 영역을 추출한 결과이다.



(그림 1) 제안된 방법의 구성도.



(그림 2) 눈 영역 추출: (a) 입력 영상, (b) 기울기 정보를 이용한 영역 추출, (c) 위치 정보를 이용한 눈 영역 추출.

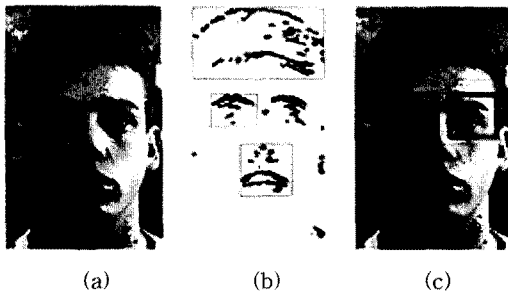
### 3. 위치 정보를 이용한 눈 영역 추출

Gizatdinova 등이 제안한 방법인 영상에서 에지 정보를 이용하여 눈 영역을 추출하는 것은 노이즈와 빛에 불필요한 데이터를 발생 시키며 정확도가 떨어지는 주요 요인이다. 노이즈에 의해 발생하는 오류는 대략적으로 추출된 영역에서 구성요소의 가로, 세로 비율을 이용하여 해결하며, 빛에 의해서 발생된 오류는 구성요소의 중심 좌표를 기반으로 한 위치 정보를 이용하여 해결한다.

실험을 통해 대략적으로 추출된 영역은 주로 이마, 턱선, 눈 영역으로 구성되어있고, 눈 영역에 대한 가로, 세로 비율( $f(x)$ )이 정사각형에 유사하며, 이마와 턱선의 에지 영역은 가로의 길이가 세로보다 긴 직사각형의 형태를 보인다.

$$f(x) = \frac{\text{height}}{\text{width}} = > 0.6 < f(x) < 1.5 \quad (1)$$

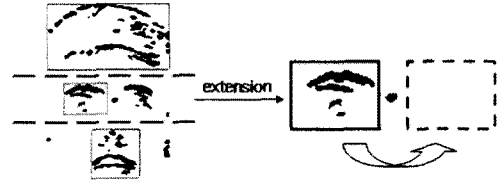
에지 영역의 가로와 세로 정보의 비율에 의해 추출된 에지 영역 중 한 개의 눈 영역을 찾은 다음 노이즈를 포함한 눈 영역 에지의 공간 정보를 통해 검출되지 않은 나머지 눈 영역을 추출 하게 된다. 이 같은 과정은 다음과 같은 두 가지의 사실에 바탕을 둔다. 첫째, 정형화된 얼굴 영상에서의 눈 영역은 동일 선상에 존재 하며, 노이즈가 포함된 에지 영역이라도 검출된 눈 영역 주위에 에지 정보가 존재 한다. 둘째, 검출된 눈 영역의 크기 정보를 통해 검출되지 않은 에지 영역의 크기 정보를 이용 할 수 있다.



(그림 3) 두 개의 눈 영역 추출을 위한 후처리 결과: (a) 입력 영상, (b) 한 개의 눈 영역 추출, (c) 후처리 과정에 의한 두 개의 눈 영역 추출.

(그림 3)의 (a)는 입력 영상이며, (b)에서 각 에지 영역에서 오른쪽 눈의 에지 정보가 검출 되지 않은 것을 볼 수 있다. 검출된 각 에지 영역에 대한 가로,

세로 비율을 이용하여 한 개의 눈 영역을 찾은 다음 동일 선상의 검출 되지 않은 에지 영역의 공간 정보를 통해 눈 영역을 추출 할 수 있다. (그림 3)의 (c)는 추출된 두 개의 눈 영역 결과를 보여 준다. (그림 4)는 추출 되지 않은 눈 영역의 추출 과정을 보여 준다. 점선에 의해 표시된 박스 영역은 검출되지 않은 눈 영역에 대한 위치 정보를 위미 한다.



(그림 4) 위치정보를 이용한 눈 영역 추출.

### 4. 실험 결과

얼굴 영역 추출을 위해 Pentium 4 CPU 3.00 GHz 768 MByte PC에서 실험을 수행 하였다. The Weizmann Institute of Science에서 제공하는 300개 [7]의 영상에서 눈 영역 추출 실험을 하였으며, <표 1>은 기존 특징 점을 이용한 방법과 본 논문에서 제안한 에지 검출과 기울기 벡터를 이용한 방법의 정확성과 수행 시간을 비교 한다. 특징 점을 이용한 방법의 정확성은 매우 높지만, 수행 시간에 문제가 있다. 이는 영상 데이터를 이용한 얼굴 영역 추출과정에서 실시간 수행에 지연이 발생 시킬 수 있다. 본 논문에서 제안한 눈 영역 추출 방법은 수행 시간과 정확성을 고려 할 때 매우 효과적이라는 실험 결과를 얻을 수 있었다. 이 방법을 이용하면, 실시간 영상 데이터에 대한 영역 검출 뿐 아니라 PDA나 모바일 기기 같은 간단한 디지털 기기에서도 실시간 영역 추출이 가능하다. 수행 결과 평균 수행 시간 0.32(sec) 와 정확성 85%의 높은 신뢰도와 효과적인 속도를 보였다.

<표 1> 정확성과 수행시간

	특징 점을 이용한 방법	Edge 검출 방법
정확성(%)	90~95	82~87
수행시간(sec)	2~8	0.2~0.45

## 5. 결론

얼굴 인식의 전처리 단계로써 특징 영역 추출을 위한 많은 연구들이 있다. 특징 점을 이용한 방법과 에지 검출을 이용한 방법의 정확성과 수행 시간을 분석해 본 결과, 특징 점을 이용한 인식 방법은 정확성이 높지만, 수행 시간이 오래 걸림을 알 수 있다. 에지에 위한 영역 추출 방법은 수행 속도 측면에서 특징 점을 이용한 방법보다 효과적이지만, 눈 영역에서 노이즈 포함 이미지에 대한 영역 추출의 정확도가 낮음을 알 수 있다. 또한 에지 정보만을 이용한 영역 추출 방법은 얼굴의 방향 정보가 정면으로 있지 않은 이미지에서 조명에 의한 어두운 영역처리가 부족한 것이 사실이다. 본 논문에서는 수행 시간의 효율성을 위해 검출된 에지의 기울기 벡터 정보를 이용하여 대략적으로 눈 영역을 추출하며, 정확성 향상을 위해 추출된 영역의 가로, 세로 비율과 공간 정보의 관계성을 이용하여 오 추출된 영역과 추출되지 않은 영역을 추출 한다. 에지 정보를 이용하여 눈 영역을 추출함으로써 매우 빠른 수행 속도를 얻었으며, 영역의 크기 비율과 각 에지 영역의 공간 정보를 통해 에지 정보를 이용한 영역 추출에서의 정확성 문제를 해결 하였다. 본 논문에서 제안한 방법은 수행 속도와 정확도를 고려 할 때 효과적인 결과를 얻을 수 있으며, 영상 이미지의 얼굴 영역에서 실시간으로 특정 영역 추출이 가능하다. 그러나 아직 제안 된 방법에서의 문제점이 남아 있다. 얼굴이미지에서 X축과 Y축으로의 기울어짐에 따른 유동적인 눈 영역 추출과 그에 따른 코, 입 영역의 추출은 해결해야 할 과제 이다.

## 참고문헌

- [1] J. Song, Z. Chi, J. Liu, "A Robust Eye Detection Method using combined Binary Edge and Intensity Information," Pattern Recognition Vol. 39, Issue 6, pp. 1110-1125, 2006
- [2] M. Pantic, J. M. Rothkrantz, "Facial Action Recognition for Facial Expression Analysis From Static Face Images," IEEE Transactions on Systems, Man and Cybernetics, Vol. 34, Issue 3, pp. 1449-1461, 2004
- [3] A. U. Batur, H. H. Monson, "Adaptive Active Appearance Models," IEEE Transactions on Image Processing, Vol. 14, Issue 11, pp. 1707-1721, 2005
- [4] F. Zuo, H. N. Peter, "Real-time Embedded Face Recognition for Smart Home," IEEE Transactions on Consumer Electronics, Vol. 51, Issue 1, pp. 183-190, 2005
- [5] Y. Gizatdinova, V. Surakka, "Feature-Based Detection of Facial Landmarks from Neutral and Expressive Facial Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, No 1, pp. 135-139, 2006
- [6] T.T. Son, S.Mita, "Face Recognition under Variable Lighting using the Mean-field Method and the Gray-level Pyramid," Proceedings of IEEE International Conference on Systems, Man and Cybernetics, Vol. 3, pp. 2107-2113, 2005
- [7] <http://www.wisdom.weizmann.ac.il/~borise/hierarchies/images/faces/faces.htm>