

# 얼굴 특징자들의 구조적 특성과 누적 히스토그램을 이용한 얼굴 표정 인식 알고리즘

김영일, 이응주  
동명정보대학교 정보통신공학과  
e-mail : ejlee@tmic.tit.ac.kr

## Face Expression Recognition Algorithm Using Geometrical Properties of Face Features and Accumulated Histogram

Young-Il Kim, Eung-Joo Lee

Dept.of Information Communication Eng, TongMyoung Univ.of Information Technology

### 요 약

본 논문에서는 얼굴의 구조적 특성과 누적 히스토그램을 이용하여 다양한 정보를 포함하고 있는 얼굴의 6가지 표정을 인식하는 알고리즘을 기술하였다. 표정 인식을 위해 특징점 추출 전처리 과정으로 입력 영상으로부터 에지 추출, 이진화, 잡음 제거, 모폴로지 기법을 이용한 팽창, 레이블링 순으로 적용한다. 본 논문은 레이블 영역의 크기를 이용해 1차 특징점 영역을 추출하고 가로 방향의 누적 히스토그램 값과 대칭성의 구조적인 관계를 이용하여 2차 특징점 추출 과정을 거쳐 정확하게 눈과 입을 찾아낸다. 또한 표정 변화를 정량적으로 측정하기 위해 추출된 특징점들의 눈과 입의 크기, 미간 사이의 거리 그리고 눈에서 입까지의 거리 정보를 이용하여 표정을 인식한다. 1, 2차 특징점 추출 과정을 거치므로 추출률이 매우 높고 특징점들의 표정에 따른 변화 거리를 이용하므로 표정 인식률이 높다. 본 논문은 안경 착용 영상과 같이 복잡한 얼굴 영상에서도 표정 인식이 가능하다.

### I. 서론

최근에 컴퓨터의 보급이 보편화되고 그래픽스 기술이 발달함에 따라 사용이 편리한 사용자 인터페이스의 개발에 대한 요구가 증가하고 있다.

표정 인식 분야에서 얼굴 표정이나 안경 착용 상태 등의 얼굴 시각 정보를 자동으로 인식하는 많은 연구가 진행되고 있다.[1][2] 이는 사람과 컴퓨터간의 더욱 지능적인 인터페이스를 추구하고 인간의 감성을 인식하여 인간 지향적인 제품을 개발하기 위한 것이다.

기존에 사용되고 있는 대표적인 얼굴 표정 인식 방법으로는 국부적인 특징(Local Feature)점들을 찾아서 특징 부분의 변화와 비교함으로써 얼굴 표정을 인식하는 방법과 전체적인 특징(Global Feature)으로 표정의 변화에 따른 얼굴 근육 움직임 정보(Optical Flow)를 이용하여 얼굴 표정을 구분하는 방법이 있다.

얼굴의 국부적인 특징을 이용하여 표정을 인식하는 방법은 안경 착용이나 수염 등의 부수적인 물질에 의해서 특징점 추출이 어렵고 인식률이 낮다는 단점을 가지고 있다. 전체적인 특징을 이용하여 표정을 인식하는 방법은 국부적인 특징을 이용한 방법에 비해 부수적인 물질의 영향은 적게 받지만 처리 시간이 많이 소요되는 단점을 가지고 있다.

본 논문은 국부적인 방법의 단점을 보완하기 위한 특징점 추출 알고리즘과 안경 착용이나 수염 등의 부수적인 물질에 대해서도 특징점 추출이 가능한 알고리즘을 제한한다. 일반적으로 얼굴 표정인식 과정은 정확한 특징점 추출을 위한 전처리 과정과 특징점 추출에 의한 얼굴표정 인식 과정으로 나누어져 실행된다.

본 논문의 구성은 먼저 2절에서는 6가지의 다양한 표정의 종류와 표정의 종류에 따라 나타나는 얼굴 특징의 분석을 설명하고 이어 3절에서는 표정인식의 특징점 추출을 위한 과정으로 기본적인 전처리 과정과 특징추출과정에서는 전처리 과정을 수행한 영상을 1차 특징점 추출 알고리즘을 사용하여 표정인식에 필요한 개체의 후보를 찾아내고 2차 특징점 추출을 이용하여 표정 인식에서 가장 중요한 정보인 눈과 입을 추출해 낸다. 다음으로 특징점인 눈, 입의 면적, 미간의 거리, 눈과 입 사이 거리를 구하여 각각의 표정을 인식하게 하였다. 4절에서는 실험을 통해 제안된 알고리즘의 타당성을 검증하고 마지막으로 5절에서 제안한 방법에 대한 평가와 앞으로의 연구방향을 중심으로 결론을 맺는다.

## II. 얼굴 표정인식을 위한 얼굴 특징 분석

표정은 특징점들의 변화에 의해 만들어지므로 각 표정에 따라 특징점이 어떻게 변화하는지를 알 필요가 있다. 웃는 표정에서는 눈썹과 눈은 반달모양으로 변하고 입의 가장자리가 올라가 얼굴의 볼 근육이 위쪽으로 올라가게 된다. 슬픈 표정은 미간이 좁아지고 눈썹, 눈, 입 가장자리가 아래로 처지게 되며, 무표정은 특징점들이 일직선상에 놓이게 된다. 놀란 표정은 눈이 크게 변하며 눈썹이 올라가고 입이 벌어져서 턱이 길어지게 된다. 윈크의 표정은 윈크하는 눈과 눈썹이 모이게 되므로 눈썹과 눈의 거리가 짧아지고 입 또한 눈 쪽으로 올라가 얼굴 근육이 윈크하는 쪽으로 모이게 된다. 졸림의 표정은 눈이 감겨지므로 눈과 눈썹의 거리가 길어지고 입은 무표정 상태와 비슷하게 일직선상에 놓이게 된다. 위와 같은 표정변화의 영상을 그림 1에 나타내었다.



그림 1. 얼굴 표정의 변화

## III. 제안한 표정 인식 알고리즘

본 논문에서 제안한 표정 인식의 단계는 크게

3단계로 전처리 단계, 특징 추출 단계, 인식의 단계로 나눈다.

### 3.1 얼굴 특징 추출을 위한 전처리 단계

본 논문에서는 입력 영상의 에지를 구하기 위하여 에지가 비교적 균일하고 특징점이 잘 구해지는 연산자를 선택하였다. 그림 2에 이 과정의 결과를 나타내었다.

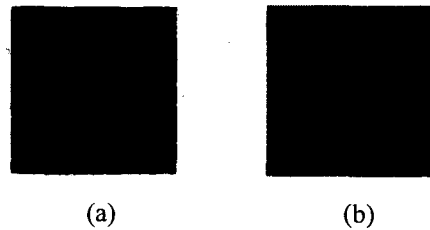


그림 2. 전처리 과정: (a) 에지추출; (b) 이진화.

안경을 착용한 영상에서는 먼저 블러링 과정을 거쳐 안경테의 고주파 성분을 줄여서 안경테의 에지 성분을 작게 하고 소벨 연산자를 이용하여 에지를 추출하였다. 이 과정으로 남아 있던 안경테의 에지는 거의 제거가 되었다. 그림 3에 이 과정의 결과를 나타내었다.

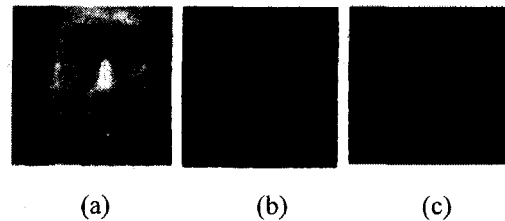


그림 3. 안경 낀 영상의 결과: (a) 슬픈 표정의 원 영상; (b) 블러링 과정; (c) 이진화.

다음으로 잡음을 제거하기 위하여 메디안 필터를 사용하였다. 메디안 필터는 특정 화소 주변 영역내의 화소 농도의 중간 값을 구하여 원하는 화소의 농도를 처리하는 것으로 영상에서 극단적인 농도의 차이가 있는 화소를 제거하므로 잡음을 제거하게 된다.

잡음을 제거한 후 끊어져 있는 특징점들의 부분을 연결하기 위해 모폴로지 기법 중에서 팽창을 사용하였다. 특징점들을 연결시키지 못하면 특징점 추출 시 특징점의 전체 영역이 추출되지 않고 일부분만이 추출이 되므로 특징점 추출이 실패하게 된다.

개체를 분리하기 위해 레이블링 알고리즘을 사용하였다. 레이블링 알고리즘은 연결되어 있는 모든 화소(연결 성분)에 동일 번호를 붙이고 다른 연결성분에는 다른 번호를 붙이는 것으로서 레이블 된 영역을 1번부터 레이블 된 개수만큼의 번호로 구분되며 영상에서는 각각의 번호가 영역의 밝기로 표현이 되어 있어서 각 연결 성분의 특징을 조사할 수가 있게 된다. 본 논문은 전처리 과정을 거치고 레이블링 알고리즘을 사용하여 약 15~20개 내외의 레이블영역을 얻을 수가 있었다. 안경을 착용한 영상에서도 동일한 알고리즘을 적용한다. 레이블링을 수행한 영상에는 눈, 입만이 아니라 원하지 않는 다른 개체가 존재하므로 그 중에서 눈, 입을 추출해야 한다. 추출의 정확성을 높이기 위해서 후보 개체의 크기 값을 이용하여 1차적으로 추출하고 히스토그램 값과 구조적인 관계를 이용하여 2차 특징 추출을 실행한다.

3.2 1차 특징점 추출

레이블링 알고리즘을 거쳐 추출된 15 ~ 20개의 레이블영역 중에서 정확히 특징점을 추출하기 위해서 레이블 영역의 면적 값을 사용하여 1차적으로 특징점을 추출하였다.

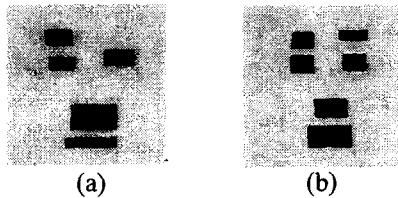


그림 4. 1차 특징점 추출 결과영상: (a) 일반영상; (b) 안경영상.

알고리즘. 1차 특징점 추출 알고리즘

```
if(Count<800 && Count>40)
    if(Area<1200 && Area>50)
        후보영역
```

특징점이 될 수 있는 후보 영역의 크기는 입력 영상의 상태에 따라 변화하므로 입력 영상에서 가장 큰 특징점의 크기 보다 작아야 하며, 가장 작은 특징점의 크기보다 커야 한다. 영상이 가장 크게 들어왔을 때와 작게 들어 왔을 때를 고려하여 화소수의 최대값은 800, 최소값은 40으로 설정하였다. Count는 레이블 영역의 화소수를

말하고 Area는 레이블영역의 면적을 의미한다. Count 와 Area는 본 논문에서 사용하는 100 x 100 영상 크기 실험 결과로 구해진 것이다. 이러한 크기 조건의 설정에 의해 처리 할 수 있는 얼굴 영상의 크기가 결정되므로 다양한 크기의 얼굴 영상을 처리하기 위해서는 충분한 크기의 설정이 요구된다.

3.3 2차 특징점 추출

1차 특징점 추출로부터 찾아낸 영역에서 정확한 특징점(눈, 입)을 찾아내기 위한 2차 방법으로 누적 히스토그램과 대칭성의 구조적 위치를 고려하여 특징점을 추출하였다. 1차 추출과정을 거치고 나면 눈썹, 코, 입, 턱선 그 외 몇 가지의 잡음이 추출되는데 그 중 눈과 입만을 찾아내는 것은 쉽지가 않으므로 많은 조건을 주어야 한다.

1차 특징점 추출과정을 거친 영상에서 가로방향으로 누적 히스토그램을 구하면 그림 5의 (a), (c)와 같다. 대칭성의 구조적인 관계를 사용하여 특징점을 추출할 수 있다. 얼굴에서 눈과 눈썹은 대칭성을 가지고 있다. 구조적으로 눈썹 아래에 눈이 존재하며 코, 입이 순서대로 존재하게 된다. 찾아낸 영역에서 대칭성을 가진 부분이 존재하면 눈썹 아래 부분은 눈으로 찾아내고 두 번째에 존재하는 히스토그램 영역을 입으로 찾아낸다.

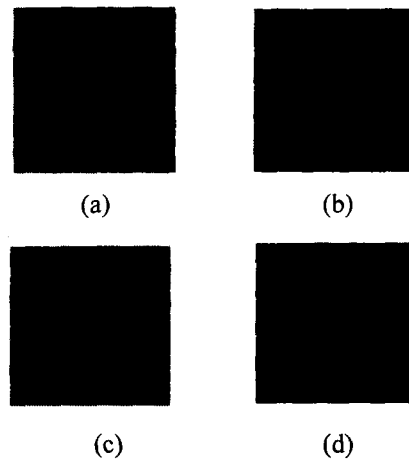


그림 5. 2차 특징점 추출: (a) 히스토그램; (b) 2차 특징점 추출; (c) 히스토그램(안경영상); (d) 2차 특징점 추출(안경영상).

윗입술과 아래 입술이 떨어져 두개의 히스토그램의 영역을 가질 때를 고려하여 히스토그램의 영역에서 7픽셀 이내의 히스토그램 영역은 같은

히스토그램 영역이라고 판정하고 추출하였다. 여기서 7픽셀은 실험영상을 통해 얻어낸 결과이다.

3.4 표정 인식 과정

영상에서 얼굴 부분을 찾아내면 인간의 표정 변화를 정량적으로 표현하는 기술 개발이 필요하다. 본 논문에서는 표정의 변화에 따라 눈 입의 면적, 미간사이의 거리, 눈에서 입까지의 거리를 가지고 표정을 인식한다. 실 데이터 13명과 예일대의 표준 얼굴영상 데이터 2명, 안경을 착용한 실 데이터 5명에 대하여 6가지 표정을 사용하여 각각의 표정에 대한 2차 모멘트까지의 평균을 로그로 환산하여 표1에 나타내었다.

표 1. 모멘트를 이용한 표정 값.

웃음		놀람		즐림	
평균	분산	평균	분산	평균	분산
1.71	4.46	1.62	3.72	1.54	4.01
슬픔		잉크		무표정	
평균	분산	평균	분산	평균	분산
1.57	4.24	1.58	4.14	1.57	4.03

또한, 모멘트 값을 사용하므로 오인하는 표정 인식률을 최대한으로 줄일 수 있었고 각 표정들마다의 특징점 변화거리를 이용함으로써 인식률을 높였다. 특징점들의 변화거리는 실험영상을 각 표정별로 나누어서 특징점 추출 후 변화거리를 구해 평균을 내어서 사용한다.

IV. 실험결과

본 논문에서 사용한 영상은 실 데이터로 안경을 쓰지 않은 영상과 안경을 착용한 영상과 예일대의 표준 얼굴 영상이며 영상 획득 시 커튼을 친 무 배경에서 자연스러운 표정을 획득하였으며 100x100사이즈의 8bit 명암도를 가지는 영상이다. 예일대의 표준 영상 또한 8bit의 100 x 100이다. 입력된 영상을 가지고 Visual C++ 6.0으로 알고리즘을 실험한 결과 보통 얼굴은 97%의 추출률을 보였으며 안경을 착용한 얼굴은 90%의 추출률을 보였다. 안경을 착용한 영상은 안경태의 종류에 따라 인식률이 달랐으며 안경이 무테이거나 은테 금테의 경우는 블러링의 과정에서 안경 에지가 제거되어 표정인식이 잘 되지만 안경테가 굵은 빨테나 크기가 작아 특징적인 눈을 가리는 안경일 경우에는 안경테를 눈 부분으로

인식하는 경우도 있다. 이런 경우에도 표정을 인식할 수 있도록 안경을 제거하거나 안경 낀 상태에서 특징점을 찾아내는 알고리즘에 관해 앞으로 연구되어야 할 것이다. 일반적인 영상에서 특징점 추출이 어려운 영상을 분석하여 보면 정면이 아닌 측면이나 기울어진 영상의 입력으로 눈 영역과 얼굴 윤곽이 연결된 경우였으며 이는 레이블링 알고리즘의 개선과 입력시의 조명에 관한 개선으로 추출율을 높일 수 있을 것으로 기대된다.

V. 결론

표정 인식은 감성, 생체 신호 및 멀티미디어 신호 연구의 기반이 되며, 얼굴 영역 분리방법의 연구, 강건하고 속도가 빠른 얼굴 인식 방법의 연구뿐만 아니라 얼굴 표정 모델링을 그래픽화하여 감정을 표현할 줄 아는 컴퓨터, 인간과 대화하는 컴퓨터를 위한 인터페이스에 있어서 많은 발전이 있을 것으로 기대된다.

본 논문은 얼굴 표정인식을 위해 먼저 구성요소 추출 알고리즘을 개발하였으며 구성요소 추출 알고리즘은 기존의 영상 처리 기법 및 얼굴 영상에 관한 지식을 효율적으로 이용하였다. 추출된 특징점의 누적 히스토그램과 구조적인 관계를 사용하여 얼굴의 다양한 표정 인식을 할 수 있었다. 현재 무표정한 표정만을 가지고 얼굴 인식을 하고 있는데 이런 표정인식에 관한 여러 연구가 활발히 진행된다면 여러 표정에서 얼굴인식을 할 수 있을 것으로 기대된다.

참고문헌

[1] R. Chellappa, C. L. Wilson and S. Sirohey, "Human and Machine Recognition of Faces, A Survey," IEEE Proc. Vol. 82, no. 5, pp. 705-740, 1995.  
 [2] Terzopoulos and K. Waters, "Analysis and Synthesis of Facial Image Using Physical and Anatomical Models," IEEE trans. Patt. Anal. Machine Intell., Vol. 15, no. 6, pp. 569-579, 1993.  
 [3] Roberto Brunelli, Tomaso Poggio, "Face Recognition : Features versus Templates," IEEE Trans. Pattern Machine Intell, Vol. 15, no. 10, 1993