

# 혼합형 특징점 추출을 이용한 얼굴 표정의 감성 인식

## Emotion Recognition of Facial Expression using the Hybrid Feature Extraction

변 광 섭\*, 박 창 현\*\*, 심 귀 보\*\*\*  
(Kwang-Sub Byun, Chang-Hyun Park, and Kwee-Bo Sim)

**Abstract** - Emotion recognition between human and human is done compositely using various features that are face, voice, gesture and etc. Among them, it is a face that emotion expression is revealed the most definitely. Human expresses and recognizes a emotion using complex and various features of the face. This paper proposes hybrid feature extraction for emotions recognition from facial expression. Hybrid feature extraction imitates emotion recognition system of human by combination of geometrical feature based extraction and color distributed histogram. That is, it can robustly perform emotion recognition by extracting many features of facial expression.

**Key Words** : 감성 인식, 얼굴 표정, 혼합형 특징점 추출, 기하학적 특징점, 컬러 분포 히스토그램

### 1. 서 론

인간의 감정 인식은 얼굴이나 음성, 몸짓 등 여러 가지 특징들을 복합적으로 이용하여 이루어진다. 그 중에서 얼굴부분에서 감정표현이 가장 확연히 드러난다. 얼굴의 여러 가지 특징(눈썹, 눈, 코, 입)의 움직임을 이용하여 감정을 표현하고 인식한다. 본 논문에서도 기존의 특징점 추출이나 형판 매칭을 이용한 감성 인식 방법을 혼합하여 인간의 감성 인식 시스템을 모방한 감성 인식을 수행한다. 즉, 눈, 코, 입, 눈썹등의 변화 특징점과 얼굴 각 영역의 컬러 히스토그램 분포를 혼합하여 얼굴에서 감정을 나타낼 수 있는 대부분의 특징들을 추출함으로써 강인한 감성 인식을 수행할 수 있다. 본 논문에서는 이러한 혼합형 특징점 추출 방식을 통해, 얼굴영역에서 감성을 인식하는 시스템을 제안한다.

인간은 여러 가지 방법을 이용하여 의사소통을 하고 서로의 감정을 표현하며 인식한다. 시각, 청각, 촉각등이 그러한 도구들의 예이다. 특히 얼굴을 이용한 감성 인식에서, 많은 채널들이 사용되기 때문에 매우 유연하고 강인한 감성인식이 가능하다. 한 채널이 실패하면 또 다른 채널에 의해 복원되고 한 채널에서의 메시지는 또 다른 채널에 의해 설명될 수도 있다.[1]

사실, 특징점 기반 정합과 형판 정합 기술은 얼굴을 인식하여 사람을 구별해내는 연구에서 사용되어져 온 것이다. 얼굴의 각부분(눈, 코, 입)의 특징을 이용한 인식에서는 각 특징의 기하학적 구조나 표현보다는 상대적인 위치나 다른 구별되는 특징을 구성하는 파라미터의 변화를 이용한다. 형판 정합의 가장 단순한 형태는 전체 얼굴을 표현하는 하나의 형판을 가지고 적절한 배열구조를 이용하여 명암도 값의 분포를 2차원으로 표현하여 구별하는 것이다 [2]. 본 논문에서는 이러한 접근법을 발전시켜 각 특징점을 나타내는 파라미터들을 찾아내고 컬러 영상으로부터 각 특징에 따른 컬러 히스토그램 분포를 혼합하여 감성인식을 위한 데이터를 추출한다. 이를 신경망의 입력으로 하여 인간의 감성 인식을 수행한다.

### 2. 감성 인식 시스템

#### 2.1. 감성 인식 시스템

인간의 얼굴 표현으로부터 감정을 인식하기 위한 과정은 크게 세 부분으로 구분된다. 첫 번째로는 감성 인식을 하는 대상이 얼굴 표정이기 때문에 얼굴 영역을 찾아야 한다. 이 과정에서는 배경이 매우 복잡하게 된 경우에 있어서도 얼굴 영역을 강인하게 찾을 수 있어야 한다. 이에 대한 많은 알고리즘이 연구되어 오고 있는데, 본 논문에서는 피부색 추출 모델을 이용하여 얼굴 영역을 검출한다. 두 번째로는 이 얼굴 영역에서 원하는 감성 정보를 얻기 위해 특징점을 추출하게 되는데, 이 과정이 매우 중요하다. 얼굴의 여러 가지 특징점, 인간이 인식하는 특징점 뿐만 아니라 인간이 인식할 수 없는 특징점까지도 감성인식을 위한 특징점 추출에 이용된다. 기

#### 저자 소개

- \* 中央大學 電子電氣工學部 碩士課程
- \*\* 中央大學 電子電氣工學部 博士課程
- \*\*\*中央大學 電子電氣工學部 正教授 · 工博

존의 특징점 추출은 기하학적 방식이나 형판 매칭, 아이겐 벡터 등 다양한 방식이 연구되어 왔는데, 본 논문에서는 이러한 알고리즘을 병렬적으로 연결한 혼합형 특징점 추출 방식을 적용한다. 마지막으로는 이 감성 정보들을 가지고 인간의 감성을 구별해내는 과정이다. 본 논문에서는 정보를 분류하는 알고리즘 가운데 가장 많이 사용되고 있는 다중 신경망 구조를 이용하여 인간의 감성을 구별해낸다. 이와 같은 얼굴 표정으로부터의 감성 인식 과정을 그림 1에 나타내었다.

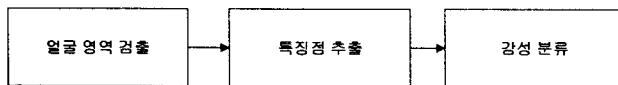


그림 1. 얼굴의 감성 인식 과정

## 2.2 얼굴 영역 검출

본 논문에서 얼굴 영역 검출은 피부색 모델을 이용하였다 [3]. 40개의 얼굴 영상으로부터 얼굴색을 추출하여 얼굴 영역을 검출한다. 학습 데이터를 위해 피부색 영역이 선택되고 RGB 형태로 저장된다. 그런데, RGB 색상은 빛의 밝기나 강도에 민감하기 때문에, RGB 색상 모델을 그대로 이용하기는 어렵다. 따라서 다음 식과 같은 강도 표준화를 통해 피부색의 편차를 감소시킬 수 있다.

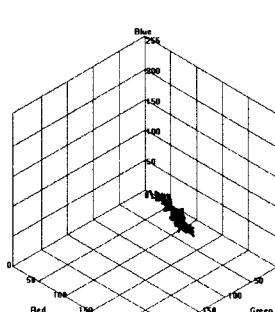
$$r = \frac{R}{R+G+B} \quad (1)$$

$$g = \frac{G}{R+G+B} \quad (2)$$

식 (1)과 (2)는 chromatic 색채 모델을 나타낸다. 먼저 얼굴 영역의 피부색이  $50 \times 20$  pixel 크기의 윈도우를 통해 RGB 형태로 추출된다. 그림 2는 얼굴 영상의 피부색 추출 모습과 추출된 RGB 색상의 분포를 나타낸다. 추출된 RGB 색상에 대해서 위 식 (1)과 (2)를 통해  $(r, g)$  색채 모델로 변환되고 가우시안 근사화를 통해 평균과 분산 값이 데이터 베이스로 저장된다. 그림 2로부터 변환된 chromatic 공간 분포와 3차원 가우시안 근사화된 모습을 그림 3에 나타내었다.

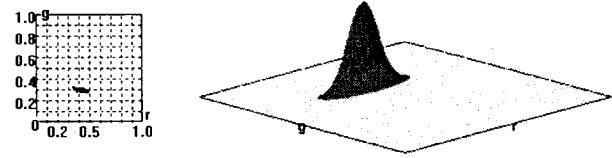


(a)



(b)

그림 2. (a) 얼굴 영상의 피부색 추출 모습 (b) 추출된 RGB 색상의 분포



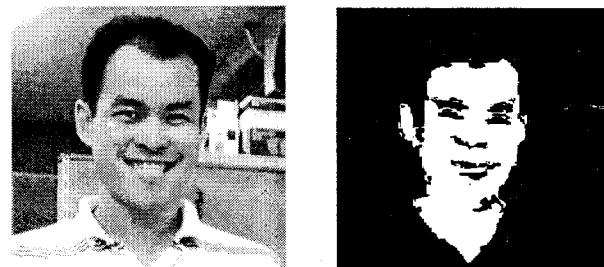
(a)

(b)

그림 3. (a) 변환된 chromatic 색채 공간 (b) 3차원 가우시안 근사화 모델

이러한 과정을 통해, 본 논문에서는 총 40 개의 영상으로부터 피부색을 추출하고 피부색 추출을 위한 표준 가우시안 모델을 얻게 된다.

새로운 입력영상은 전체적으로 chromatic 색채 공간으로 변환된 후 표준 가우시안 모델과 임계값을 적용을 이진 영상으로 변환되어 얼굴 영역을 검출하게 된다. 다음 그림 4는 원래 얼굴 영상과 얼굴 영역 검출 알고리즘에 의해 찾은 이진 얼굴 영역을 나타낸다.



(a)

(b)

그림 4. (a) 원본 얼굴 영상 (b) 얼굴 영역 검출 알고리즘에 의해 찾은 이진 얼굴 영역

## 3. 혼합형 특징점 추출과 감성 분류

### 3.1 혼합형 특징점 추출 알고리즘

다양한 특징점 추출 방식을 통합적으로 이용하기 위해, 본 논문에서는 혼합형 특징점 추출 방식을 제안한다. 그림 5는 이러한 혼합형 특징점 추출 방식의 개략도를 보여준다. 그림

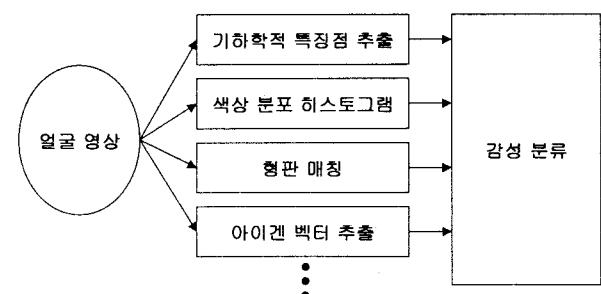


그림 5. 혼합형 특징 추출 방식의 개략도

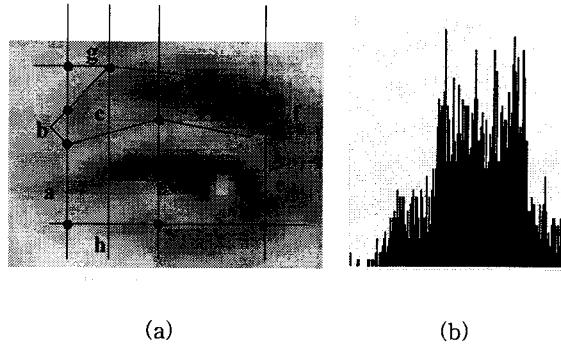


그림 6. (a) 눈썹의 특징점 추출에 사용된 파라미터 (b)  
눈 및 눈썹의 색상 분포 히스토그램

5에서 알 수 있듯이 얼굴 영상으로부터 다양한 특징점을 추출하여 감성 분류 시스템의 입력으로 이용하게 되며, 언급된 특징점 추출이외의 다른 알고리즘도 병렬적으로 추가될 수 있다. 본 연구에서는 기존에 가장 많이 사용되었던 기하학적 특징점 추출 방식과 감성 인식 분야에서는 사용되지 않았던 색상 분포 히스토그램의 두 가지에 대해 우선적으로 특징점을 추출한다.

기하학적 특징점 추출은 R. Brunelli와 T. Poggio가 제안한 방식을 사용한다 [2]. 이 알고리즘에서 사용된 특징점은 눈썹의 두께와 눈의 중심위치에서의 수직 위치, 왼쪽 눈썹의 표현(11개), 코의 수직 위치 및 폭, 입의 수직위치, 폭, 입술의 높이, 턱의 모양을 나타내는 11개의 반경, 코 위치에서의 얼굴 폭, 코 끝과 눈 사이의 중심점에서의 얼굴 폭 등 총 35개의 기하학적 특징점을 추출한다. 그림 6(a)는 이 중에서 눈썹 부분의 특징점 추출 모습을 보여준다.

또 하나의 특징점 추출 알고리즘으로 색상 분포 히스토그램을 이용하는데, 눈 및 눈썹, 코 끝 주변, 입 주변의 색상 분포 히스토그램의 모양을 추출한다. 그림 6(b)는 눈 및 눈썹 주변의 색상 분포 히스토그램 추출 모습을 보여준다.

### 3.2 병렬적 다층 신경망을 이용한 감정 분류

본 연구에서는 혼합형 특징점 추출 방식을 사용하기 때문에 특징점을 추출하는 단계의 개수만큼 다층 신경망 구조가 필요하다. 또한 우선 4개의 감정; 행복, 화남, 슬픔, 놀람등에 대하여 분류하기 때문에 신경망의 마지막 층에서는 4개의 노드를 갖게 된다. 여기에서는 기하학적 특징과 히스토그램 추출방식의 두 가지를 사용하므로 그림 7과 같은 두 개의 다층 신경망이 필요하고 마지막 단에 4개의 노드와 가중치가 추가되어 최종적인 감성 분류를 수행하게 된다. 우선 40개의 얼굴 영상으로부터 교사 학습을 통해 신경망의 노드와 가중치가 업데이트 되고, 최종적으로는 새로운 입력 영상에 대해 감성을 분류할 수 있는 시스템을 형성한다.

그림 7의 블록 안의 구조는 일반적인 다층 신경망 구조를 나타내는데, 최종 단에 연결된 4개의 노드와 가중치를 업데이트하기 위해서는 두 단계의 학습이 필요하다. 먼저 기하학적 특징점으로부터의 감성 정보를 분류하기 위한 위쪽의 다층 신경망의 학습과 히스토그램으로부터의 감성정보 분류를 위한 아래쪽의 신경망 학습이 이루어진다. 다음 이 병렬 구조

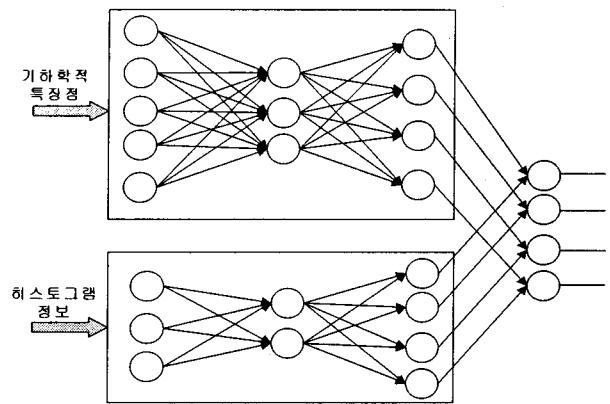


그림 7. 감정 분류를 위한 병렬적 다층 신경망 구조

를 연결하는 최종의 노드와 가중치를 위해 전체적인 학습이 다시 한번 이루어져서 전체 감성 분류 시스템을 구성한다.

### 4. 결론 및 향후 과제

얼굴 영역의 감성 인식을 위해, 인간의 복잡한 감성 인식 시스템을 기반으로 한 혼합형 특징점 추출 방식이 본 연구에서 제안되었다. 이 방식은 여러 가지 다양한 특징점 추출을 수행하고 이를 병렬적으로 분류한 후에, 최종적으로 감성을 분류한다. 혼합형 특징점 추출 방식은 다양한 감성 정보를 입력으로 하고 통합적인 감성 분류를 수행하기 때문에 강인한 감성 분류를 수행할 수 있을 것으로 기대한다.

본 논문에서는 특징점 추출로서, 기하학적 추출과 히스토그램 정보를 이용하였는데, 형판 매칭이나 아이겐벡터와 같은 다양한 특징점 추출 방식을 적용하는 것이 앞으로의 과제이다. 또한 우선적으로 4가지의 감정에 대해서만 감성 분류를 제안하였는데, 더 많은 감정을 분류하기 위해 신경망의 구조 변화와 실험이 필요하다.

### 참 고 문 헌

- [1] M. Pantic, L.J.M. Rothkrantz, "Automatic analysis of facial expressions: the state of the art", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 12, pp. 1424-1445, Dec. 2000.
- [2] R. Brunelli, T.Poggio, "Face recognition: features versus templates", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15, no. 10, pp. 1042-1052, Oct. 1993.
- [3] R. S. Feris, T. E. Campos, R. M. Cesar Junior, "Detection and Tracking of Facial Features in Video Sequences", Lecture Notes in Artificial Intelligence, Springer-Verlag press, vol. 1793, pp. 197-206, April 2000.