

## 신경회로망 기반 감성 인식 비전 시스템

이상윤, 김 성남, 주 영훈,  
군산대학교 전자정보공학부

박 창현\*, 심 귀보\*  
\* 중앙대학교 전자전기공학부

### Vision System for NN-based Emotion Recognition

Sang-Yun Lee, Sung-Nam Kim, Young-Hoon Joo, Chang-Hyun Park\*, Kwee-Bo Sim\*  
School of Electronic and Information Eng. Kunsan National University  
\* School of Electronic and Electrical Eng., Chung-Ang University

**Abstract** - In this paper, we propose the neural network based emotion recognition method for intelligently recognizing the human's emotion using vision system. In the proposed method, human's emotion is divided into four emotion (surprise, anger, happiness, sadness). Also, we use R,G,B(red, green, blue) color image data and the gray image data to get the highly trust rate of feature point extraction. For this, we propose an algorithm to extract four feature points (eyebrow, eye, nose, mouth) from the face image acquired by the color CCD camera and find some feature vectors from those. And then we apply back-propagation algorithm to the secondary feature vector(position and distance among the feature points). Finally, we show the practical application possibility of the proposed method.

(anger), 공포(fear), 즐거움(joy), 역겨움(disgust), 흥미(interest), 놀람(surprise), 경멸(contempt), 수치(shame), 슬픔(sadness), 괴로움(depress), 죄책감(guilty), 수용감(acceptance)의 12가지로 요약될 수 있다.  
본 논문에서는 상기 12가지의 감성중 분노, 즐거움, 슬픔, 놀람의 4가지 극단적인 감성을 인식하기 위하여, 비전을 이용한 신경회로망 기반 감성 인식 방법을 제안한다. 제안한 방법은 칼라 카메라로부터 획득한 칼라 얼굴 영상으로부터 특징점을 추출하고, 이를 이용해 새로운 특징벡터를 산출하는 방법을 제시한다. 얼굴 영상으로부터 신뢰성 있는 특징점을 추출하기 위하여 칼라 영상과 흑백 영상 둘 다 이용하는 방법을 소개한다. 또한, 사람마다 서로 다르게 추출된 특징벡터를 이용하여 역전파-신경회로망을 통해 얼굴 영상을 학습시키고 감성을 인식하는 방법을 소개하고 직접 본 연구실에 구축된 비전 시스템을 이용하여 본 논문에서 제안한 방법의 응용가능성을 실험한다.

#### 1. 서 론

사람의 여러 가지 감각 기능 중에서 시각이 차지하는 비중은 매우 크다고 할 수 있다. 인간이 감각에 의하여 취하는 정보의 90%를 시각이 담당하고 있다고 해도 과언이 아니다. 이렇게 사람의 시각이 중요한 정보 취득기관일 수 있는 것은, 눈의 높은 해상도와 빠른 처리 능력과 인식능력에 기인한다고 할 수 있다. 카메라 성능에 따라 다르겠지만 보통 사용되는 해상도는 640\*480 정도인 것에 비하면 사람의 눈의 해상도는 약 1,000,000\*1,000,000 정도의 높은 해상도를 가지고 있다고 한다. 최근의 컴퓨터 기술의 급속한 발전에도 불구하고 인간 시각의 정보 처리 능력에 비한다면, 아직까지도 매우 원시적이 정도의 처리 능력을 가지고 있는 정도에 불과하다. 이렇게 인간 시각의 우수한 능력에 의하여, 시각 장애인의 아닌 많은 사람들은 시각에 의존하여 생활을 영위하고 있다고 하여도 과언이 아니다. 이러한 인간 시각 기능을 모방하여 자동화 기기에 적용하려는 노력이 활발히 연구되고 있다.

인간이 누군가를 기억할 때는 그 사람의 전체 얼굴로서 기억하는 것보다 대표할 수 있는 얼굴의 특징점들의 상관관계로부터 더 많이 기억하게 된다. 얼굴영역 추출은 주로 얼굴 요소의 지형적 특징에 기반한 방법과 얼굴의 형판에 기반한 방법이 있다. 본 논문에서 적용할 얼굴의 지형적 특징에 기반한 방법은 얼굴의 주요 구성요소인 눈썹, 눈, 코, 입, 얼굴 윤곽등을 추출하여 이들의 상관관계를 사전지식으로 이용하여 얼굴 영역 및 특징을 추출하는 방법이다. 이러한 특징점들은 안경을 썼을 때, 머리모양의 변화에 따른 큰 변화에 불구하고 매우 중요하다. 따라서, 최근의 얼굴 인식 연구는 어떻게 좋은 특징점을 추출하느냐에 집중되고 있다[1].

특징점의 상관관계로부터 인간의 기본감성은 분노

#### 2. 감성 인식 비전 시스템

##### 2.1 전체 시스템의 개요

얼굴영상으로부터 감성 인식을 위한 정보를 얻는 여러 가지 연구가 진행되어 왔다 [1-4]. 칼라 카메라로부터 획득한 얼굴영상으로부터 정보를 얻는 방법은 크게 얼굴의 정면을 이용하는 방법과 측면을 이용하는 방법이 있다. 첫번째 방법은 두 번째 방법보다 얼굴의 전체적인 특징점들을 위치와 크기 판단에 있어 신뢰도가 높기 때문에 본 논문에서 얼굴의 정면을 이용하는 방법을 사용한다. IBM-PC에 호환되는 컬러 CCD 카메라와 프레임 그레버 장치를 사용하여 영상을 획득하고, 획득된 얼굴 영상으로부터 특징점을 추출하는데 사용되는 메모리 공간과 시간 지연을 줄이기 위해 320x240 해상도의 영상으로 줄인 후 실험했다.

##### 2.2 얼굴 특징점 추출 과정

그림 1은 칼라 카메라로부터 획득한 배경이 포함된 초기의 얼굴 영상을 나타낸다. 그림 2는 그림 1의 영상으로부터 배경을 제외한 얼굴의 형태와 특징점들의 형태를 얻기 위해 살색만을 추출한 영상을 나타낸다. 그림 3은 살색 정보 영상과 그레이 level의 임계값 영상의 공통부분(intersection)을 선택해 (눈썹, 눈, 코, 입, 얼굴의 윤곽)을 추출해 내고 labeling 처리 후 크기 필터링을 거쳐 각 특징 blob과 배경을 분리해 낸 그림을 도시한 것이다. 즉, 본 논문에서 사용한 얼굴 특징점 추출방법은 배경이 포함된 초기의 영상으로부터 얼굴 영상을 추출하기 위해 인간의 살색 정보를 이용하여 영상을 추출한 다음 살색 정보 영상과 그레이 영상 레벨의 임계값을 사용하여 얼굴 각 특징점들을 추출하는 방법이다.



그림 1. 원 영상



그림 4. 특징점으로부터의 특징벡터 표현



그림 2. 살색정보 영상



그림 5. 틀어진 각도의 경우

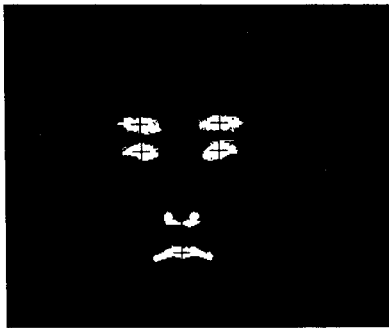


그림 3. Blob 영상

### 2.3 2차적 특징벡터 산출 방법

인간은 감성에 따라 특징적인 눈썹, 눈 또는 입의 위치나 크기에 변화가 일어난다. 본 논문에서는 이런 변화를 감성 인식에 이용하기 위해 코의 중심점을 기준으로 기하학적인 14 가지의 특징벡터를 산출하는 방법을 제안한다. 그림 4는 본 논문에서 제안한 얼굴 영상의 14 가지 특징 벡터를 추출한 예를 도시한 그림이다. 그림 4에서 추출한 14가지의 특징 벡터를 이용하여 인간의 감성을 인식하는 추론기로는 신경회로망을 이용한다. 이의 자세한 방법은 다음 절에 설명된다.

그림 5는 획득한 영상이 경사진 경우를 나타낸 그림이다. 그림 5와 같이 인간의 얼굴의 각이 틀어진 경우에는 인식에 있어 신뢰도가 낮아지므로 ⑭번 특징벡터의 x 좌표가 틀어질 경우에는  $x/r = \cos \theta$  를 이용해 각도를 구한 다음, 경사진 각 만큼 보정한 후 본 논문에서 제안한 기하학적인 특징벡터를 산출한다.

특징벡터는 감성 인식의 신뢰도 문제에 있어 중요한 요소이다. 본 논문에서 사용된 기하학적인 특징벡터 값은 다음과 같이 구한다.

① 코의 중심

$$A = \frac{(\text{point } 10) - (\text{point } 9)}{2} = (x_n' - y_n')$$

② 코의 중심과 왼쪽눈 오른쪽 끝점까지의 거리

$$(A \rightarrow 1) = \sqrt{(x_n - x_1)^2 + (y_n - y_1)^2}$$

③ 코의 중심과 오른쪽눈 왼쪽끝점까지의 거리

$$(A \rightarrow 5) = \sqrt{(x_5 - x_n)^2 + (y_n - y_5)^2}$$

④ 코의 중심과 왼쪽 눈썹왼쪽끝점까지의 거리

$$(A \rightarrow 13) = \sqrt{(x_n - x_{13})^2 + (y_n - y_{13})^2}$$

⑤ 코의 중심과 오른쪽 눈썹오른쪽끝점까지의 거리

$$(A \rightarrow 14) = \sqrt{(x_{14} - x_n)^2 + (y_n - y_{14})^2}$$

⑥ 왼쪽눈의 수직 거리

$$(4 \rightarrow 2) = (y_1 - y_2)$$

⑦ 오른쪽눈의 수직 거리

$$(8 \rightarrow 6) = (y_8 - y_6)$$

⑧ 왼쪽눈과 오른쪽눈과의 거리

$$(5 \rightarrow 1) = (x_5 - x_1)$$

⑨ 코의 중심과 입술왼쪽 끝까지의 거리

$$(11 \rightarrow A) = \sqrt{(x_n - x_{11})^2 + (y_{11} - y_n)^2}$$

⑩ 코의 중심과 입술 오른쪽 끝까지의 거리

$$(12 \rightarrow A) = \sqrt{(x_{12} - x_n)^2 + (y_{12} - y_n)^2}$$

⑪ 입술 왼쪽 끝과 오른쪽 끝까지의 거리

$$(12 \rightarrow 11) = \sqrt{(x_{12} - x_{11})^2 + (y_{12} - y_{11})^2}$$

⑫ 코의 중심과 얼굴 왼쪽면과의 거리

$$(A \rightarrow 16) = (x_n - x_{16})$$

⑬ 코의 중심과 얼굴 오른쪽면과의 거리

$$(17 \rightarrow A) = (x_{17} - x_n)$$

⑭ 코의 중심과 턱 끝까지의 거리

$$(15 \rightarrow A) = (y_{15} - x_n)$$

특징벡터 산출시 카메라의 거리에 따라 특징점들의 상관관계를 고려하여 특징 벡터 ⑦인 왼쪽 눈과 오른쪽 눈 사이의 값을 기준으로 특징 벡터의 값을 비율에 따라 산출한다. 표 1은 그림 4에서 특징 벡터를 산출한 값을 표로 나타낸 것이다.

표 1. 특징벡터의 산출 값

| 특징 벡터 요소 | 특징 벡터 값 |
|----------|---------|
| 특징 벡터 ①  | 0.462   |
| 특징 벡터 ②  | 1.404   |
| 특징 벡터 ③  | 1.442   |
| 특징 벡터 ④  | 3.135   |
| 특징 벡터 ⑤  | 2.981   |
| 특징 벡터 ⑥  | 0.788   |
| 특징 벡터 ⑦  | 0.808   |
| 특징 벡터 ⑧  | 1       |
| 특징 벡터 ⑨  | 1.635   |
| 특징 벡터 ⑩  | 1.615   |
| 특징 벡터 ⑪  | 1.770   |
| 특징 벡터 ⑫  | 2.385   |
| 특징 벡터 ⑬  | 2.365   |
| 특징 벡터 ⑭  | 2.712   |

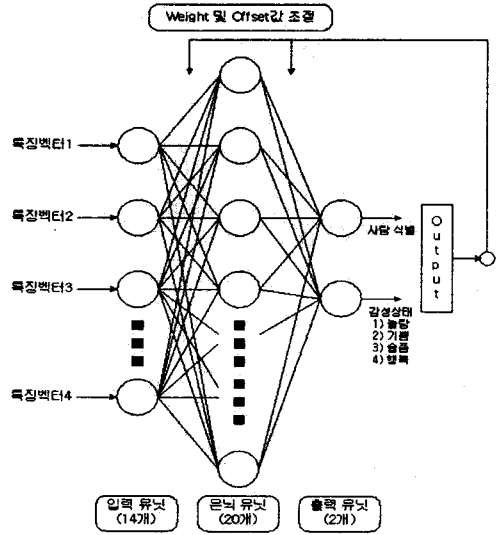


그림 6. 학습에 사용한 역전파-신경망

### 3. 신경회로망 기반 감성 인식 실험

본 논문에서는 14개의 기하학적인 특징 벡터 값을 입력으로 하는 3층 역전파 신경회로망을 이용하여 감성인식 시스템을 설계하는 방법을 사용한다. 이를 위해 그림 4에서 획득한 14개의 특징벡터를 이용하여 한명의 인간의 감성 (놀람, 슬픔, 화남, 기쁨)들에 대해 각각 50개씩 전체 200개의 샘플 데이터를 구했다. 놀람의 표정을 예를 들면, 신경회로망의 입력은 14개가 되며 입력패턴은 50개가 된다. 사람인식과 감성상태를 나타내기 위해 출력은 2가지가 된다. 표 2는 5명의 사람과 감성상태를 나타내는 목표 데이터의 한 예이다.

표 2. 목표 데이터의 예

| 출력 예      | 목표값(인간, 감성) |
|-----------|-------------|
| (인간1, 놀람) | [1, 1]      |
| (인간2, 슬픔) | [2, 2]      |
| (인간3, 화남) | [3, 3]      |
| (인간4, 기쁨) | [4, 4]      |
| (인간5, 슬픔) | [5, 2]      |

산출된 특징벡터는 그림 6의 역전파 신경망을 통해 학습 시켰다.

역전파 신경망은 학습이 수행되는 동안 가중치 정보가 향상되고 원하는 반복 회수에 도달하면 학습을 멈추고 최종 가중치 정보가 저장된다. 이 정보를 이용하여 사람의 식별과 감성인식을 수행한다. 표 3은 본 논문에서 제안한 감성 인식 시스템을 이용하여 학습에 사용되지 않은 얼굴표정을 적용한 후의 결과 값의 예이다.

표 3. 실제 출력 데이터

|        |                      |
|--------|----------------------|
| 입력 데이터 | [인간2, 기쁨 표정]의 특징벡터 값 |
| 출력 데이터 | [1.978, 4.072]       |

표 3의 결과에서 알 수 있듯이 사람의 식별 및 인식에 있어 높은 신뢰도가 있음을 알 수 있다.

### 4. 결 론

본 논문에서는 비전을 이용하여 인간의 감성 상태를 인식하기 위하여 신경회로망을 기반으로 하는 감성 인식 비전 시스템을 제안하였다. 기존의 논문에서는 얼굴 검출만을 다루었는데 본 논문에서는 사람의 식별 및 감성 상태를 동시에 인식하는 방법으로 발전시켰다. 특징 Blob 추출에 있어 그 신뢰도를 높이기 위해 살색의 컬러 정보를 이용하였고 인식 시간을 줄이기 위해 기하학적 특징벡터 산출법을 제안하였다. 제안한 방법의 적용 가능성을 실험한 결과(5명의 4가지 감성상태 실험), 사람 식별에 있어서는 모두 인식하였고 감성상태는 현재 64% 정도의 신뢰도를 얻을 수 있었다. 향후, 더 신뢰성 있는 감성 인식 시스템을 개발하기 위하여 얼굴 특징점 추출을 위한 정교한 영상 처리 알고리즘을 개발하고, 기하학적 특징벡터의 수를 줄이고 복수의 인간이 존재할 때 각각의 사람 식별과 감성 상태를 추출 할 수 있는 시스템을 개발할 예정이다.

감사의 글 : 본 연구는 산업자원부의 차세대 신기술 연구 개발 사업에 의해 지원 받았습니다.

### [참 고 문 헌]

- [1] M. Turk and A. Pentland, "Eigenfaces for Recognition", SPIE Vol.1192 IRCV VIII, 1989.
- [2] I.Craw, H.Ellis, and J.R.Lishman, "Automatic Extraction of Face Features", Pattern Recognition Letters 5, pp. 183-187, 1987.
- [3] N. Honda and S. Aida, "Analysis of Multi-variate Medical Data By Face Method", Pattern Recognition Vol. 15, No. 3, pp. 212-242, 1982.
- [4] L.D. Harmon and W.F. Hunt, "Automatic Recognition of Human Face Profiles", Computer Graphics and Image Processing, pp. 135-156, 1977.
- [5] J.L. McClelland, D.E. Rumelhart, and the PDP Research Group, "Parallel Distributed Processing Vol 1, 2, 3.", The MIT Press, Cambridge, Massachusetts, London, England, 1986.