

# 비전에 의한 감성인식

## Emotion Recognition by Vision System

이 상윤 · 오 재흥 · 주 영훈 · 심 귀보\*\*  
Sang Yun Lee, Jae heung Oh, Young Hoon Joo, Kwee Bo Sim\*

군산대학교 전자정보공학부, \*중앙대학교 전자전기공학부

**Abstract** - In this paper, we propose the neural network based emotion recognition method for intelligently recognizing the human's emotion using CCD color image. To do this, we first acquire the color image from the CCD camera, and then propose the method for recognizing the expression to be represented the structural correlation of man's feature points(eyebrows, eye, nose, mouse). It is central technology that the process of extract, separate and recognize correct data in the image. for representation is expressed by structural corelation of human's feature points In the proposed method, human's emotion is divided into four emotion (surprise, anger, happiness, sadness). Had separated complexion area using color-difference of color space by method that have separated background and human's face toughly to change such as external illumination in this paper. For this, we propose an algorithm to extract four feature points from the face image acquired by the color CCD camera and find normalization face picture and some feature vectors from those. And then we apply back-prapagation algorithm to the secondary feature vector. Finally, we show the practical application possibility of the proposed method.

### 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 취득기관 일 수 있는 것은, 눈의 높은 해상도, 빠른 처리 능력 그리고, 인식 능력에 기인한다고 할 수 있다. 컴퓨터 기술의 급속한 발전에도 불구하고 인간 시각의 정보 처리 능력에 비한다면, 아직까지도 매우 원시적인 정도의 처리 능력을 가지고 있는 정도에 불과하다. 그러나, 최근 컴퓨터 및 광학 장비의 발전과 더불어 사용자의 의지나 감성을 인지하는 보다 인간 친화적인 감성 컴퓨터 인터페이스를 구축하고자 하는 방향으로 연구들이 활발히 진행되고 있다.[6-9][11-12] 연구되어지는 여러분야중 특히 인간의 얼굴표정인식 및 감성인식의 분야가 있다.[2-4][10] 크게 두 가지로 나누어 볼 수 있다. 이는 얼굴영상에서 기하정보를 바탕으로 특징 벡터에 기반한 방법과 또 다른 접근 방법으로는 주어진 영상을 단순화하는 변형을 거쳐서 저차원의 표현형을 갖는 자료구조 인식 시스템을 들 수가 있다. 이러한 연구 분야의 접근 방법 중 본 논문에서 적용한 방법은 얼굴영상에서 기하정보를 바탕으로 한 특징벡터에 기반한 방법과 얼굴의 지형적 특징을 이용한 방법이다.[5] 이 방법은 얼굴의 주요 구성요소인 눈썹, 눈, 코, 입, 얼굴 윤곽 등을 추출하여 이들의 상관관계를 사전지식으로 이용하는 것이다. 이와 같은 얼굴 인식 연구 분야에서 집중된 문제는 정확하고 일관된 특징점 추출문제라 할 수 있다.[1]. 본 논문에서는, 비전

시스템에서 획득한 칼라 영상으로부터 얼굴의 특징들을 추출하기 위해 R, G, B 컬러 이미지 데이터를 이용하는 방법을 제안한다. 이를 위해, 먼저 컬러 CCD 카메라와 영상 보드를 통해 획득한 복합 영상으로부터 얼굴 영역을 추출하고, 두 번째로 추출된 얼굴 영상으로부터 4개의 얼굴 특징 요소(눈썹, 눈, 코, 입)의 위치와 근육 구조의 특징들을 추출하고, 세번째로 이들 데이터로부터 2차적 특징 벡터 데이터들을 찾아내는 알고리즘을 제안하고, 네 번째로 신경회로망의 역전파 알고리즘을 이용하여 각 개인의 식별 및 감정 상태를 인식하는 방법을 제안한다. 신경망을 기반으로 하는 인식 시스템의 장점은 그 유연성과 강인성에 있다고 할 것이다. 이러한 신경망의 특징은 전체 상태 공간상에서의 편향된 입력이 되더라도 네트워크의 각 함수를 거쳐 얼굴의 어느 정도의 왜곡에 대해서 일관된 결과를 얻어낼 수 있다는 데 있다. 마지막으로, 본 연구실에 구축된 비전 시스템을 이용하여 직접 실험을 통해 제안한 방법의 응용 가능성을 제시한다.

### 2. 본 론

인간의 감성은 크게 분노, 공포, 즐거움, 역겨움, 흥미, 놀람, 경멸, 수치, 슬픔, 피로움, 죄책감, 수용감등 12 가지의 감정상태로 분류할 수 있다. 인간이 누군가를 기억할 때는 그 사람의 전체 얼굴로서 기억하는 것보다 대표할 수 있는 얼굴의 특징점들의 상관관계로부터 기억을 이끌어 내게 된다. 이러한 특징의 분석과정에서 분류한다면 얼굴 요소의 지형적 특징에 기반한 방법과 얼굴의 형판에 기반한 방법으로 나누어 볼 수 있다.[9]

얼굴인식에 있어서 고려되는 중요한 요소들은 얼굴영역의 탐색, 얼굴에서 각 특징영역의 추출, 찾아진 영역에서의 특징점의 추출, 이들 정보들을 바탕으로 한 얼굴인식 등 이라고 할 수 있다. 그러한 이유 때문에 얼굴 인식 시스템에 대하여 관심이 되는 것은 얼마나 주어진 영상으로부터 정확하고 일정한 얼굴영역을 찾아내고, 영상의 여러 광학적 조건들에 대하여 정확하고 일관성 있는 데이터를 추출할 수 있는가에 달려 있기 때문이다.

**2.1 비전 시스템의 구성 및 감성인식 단계**

정면의 얼굴 영상 획득을 위해 사용된 시스템은 IBM-PC에 호환되는 컬러 CCD 카메라와 프레임 캡처 장치를 이용하여 영상을 획득하고, 획득한 영상으로부터 특징점을 추출하는데 사용되는 메모리 공간과 인식 시간을 줄이기 위해 320x240의 해상도로 축소하여 실험할 수 있도록 시스템을 구성한다. 초기에 획득한 복합 영상으로부터 영상 크기에 따른 거리별 오차를 보정하기 위해 일반화된 영상으로 표준화 시킨 영상을 생성한다.

**2.2 얼굴영역 추출**

초기에 획득한 칼라 영상으로부터 얼굴 영역의 위치를 찾기 위해 칼라 공간의 색차(chrominance)를 이용한다. 사용된 방법은 HSI(hue, saturation, intensity) 칼라 공간 모델에서 특징 blob을 잃어버리지 않도록 I를 주위 환경에 민감하지 않는 값으로 조절한다. 그리고, I의 밝기 값에 따른 H의 범위중 피부색의 범위를 찾아낸다. 그 다음, 추출된 HSI 칼라 데이터를 R, G, B 각 색상 신호로 변환시켜 피부색 범위의 이진화 영상을 추출한다. 그림 1-(a)는 상기의 방법을 이용하여 추출된 피부색 칼라 영상을 도시한 그림이다. 그림 1-(a)와 같이 이진화 표현된 피부색 영상으로부터 피부색 범위에 들어오는 다른 복합영상의 노이즈들을 제거하기 위해 각 blob들을 라벨처리한 후, size filtering을 적용하여 추출한 그림은 그림 1-(b)와 같다. 그림 2는 상기와 같은 방법으로 추출된 영역 데이터를 이용하여 생성된 복합 영상에서의 칼라 얼굴 영역을 나타낸 그림이다.

**2.3 얼굴 특징점 추출**

그림 3-(a)는 그림 2에서 획득한 얼굴 영역으로부터 계산 영역을 줄이기 위해서 영역을 분할시킨 정량화 얼굴 영역 영상이다. 그림 3-(b)는 정량화된 영상에서 2.2 절에서 제시한 피부색 추출 과정을 반복 처리하여 획득한 이진화 영상이다. 그 다음 정량화 시킨 그림 3-(a) 영상에서 얼굴의 형태와 특징점들의 형태를 얻기 위해 그림 3-(b)의 피부색 이미지와 그림 3-(c)의 최적화시킨 이진화 영상을 영역별로 계산한다. 마지막으로 추출된 영상을 라벨링처리한 후, Size filtering을 거쳐 그림 4-(b)-4-(f)와 같이 각 특징 blob을 얻는다.



(a) 피부색이미지 (b) Label 필터후 이미지  
그림 1. 복합영상에서 피부색 추출



그림 2. 얼굴영역 표출 이미지

**2.4 얼굴영상에서의 특징점 추출**

인간은 감성에 따라서 눈썹, 눈 또는 입의 위치나 크기변화가 일어난다[1]. 이런 변화를 감성 인식에 이용하기 위해 코의 중심점을 기준으로 기하학적인 11 가지의 특징벡터를 산출한다. 특징 벡터는 감성 인식의 신뢰도 문제에 있어 중요한 요소이다. 본 논문에서는 다음과 같이 기하학적인 특징벡터 값을 구한다.

- ① 코의중심  

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

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

$$(4 \rightarrow 2) = (y_4 - y_2)$$
- 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  

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

인간의 감성 변화에 민감하게 변하지 않는 8번 벡터를 기준으로 다른 모든 벡터들을 정량화시킨다.

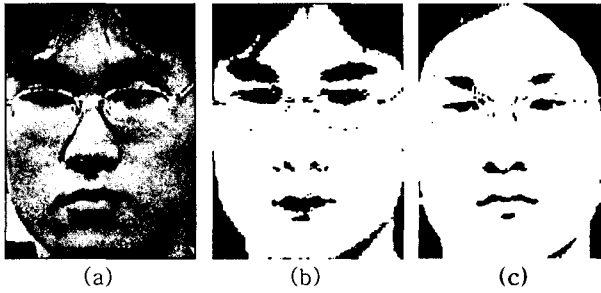


그림 3. (a) 정량화 시킨 얼굴영역 이미지  
(b) 피부색 이미지  
(c) 최적화 시킨 이진화 이미지

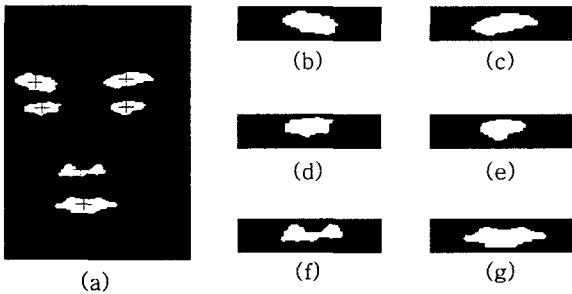


그림 4. 특징점 추출 이미지



그림 5. 특징벡터 표현 이미지

2.5 감성인식을 위한 학습 알고리즘

본 논문에서는 11개의 기하학적인 특징 벡터값을 산출한 후, 3층 역전파 신경망을 이용하여 감성인식 시스템을 설계한다(6). 획득한 11개의 특징벡터를 이용하여 인간의 감성(무표정, 놀람, 슬픔, 화남, 기쁨)들에 대해 각각 50개씩 전체 250개의 학습 이미지들에 대한 데이터를 구한다. 놀람 표정의 경우, 신경회로망의 입력은 11개의 특징벡터가 되며 입력 패턴은 50개가 된다. 출력은 사람 인식과 감성 상태를 나타내기 위해 2가지가 된다. 표 2는 5명의 사람과 감성상태를 나타내는 목표 데이터의 한 예이다. 산출된 특징벡터는 그림 11의 역전파 신경망을 통해 학습시킨다. 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 3

은 학습에 사용되지 않은 얼굴표정을 적용한 후의 결과 값의 예이다. 표 3의 결과에서 알 수 있듯이 사람의 식별 및 인식에 있어 높은 신뢰도가 있음을 알 수 있다. 그림 12는 본 연구의 실험을 위해 제작된 감성 인식 처리 화면을 나타낸 것이다. 표 4-8 은 한 사람에 대한 감성인식 실험의 결과 데이터이고 그림 6-10은 감성 인식 실험의 데이터 영상이다.

표 2. 목표 데이터의 예

출력 예	목표값(인간, 감성)
(ID 1, 무표정)	[1, 1.0]
(ID 2, 기쁨)	[2, 2.0]
(ID 3, 놀람)	[3, 3.0]
(ID 4, 화남)	[4, 4.0]
(ID 5, 슬픔)	[5, 5.0]

표 3. 실제 출력 데이터

입력 데이터	[인간1, 기쁨 표정]의 특징벡터 값
출력 데이터	[1.978, 4.072]

표 4. ID가 1인 사람의 무표정 감성상태 실험결과

FV 1	FV 2	FV 3	FV 4	FV 5	FV 6	FV 7
1.0000	0.82546	0.88435	1.42295	1.51720	0.14285	0.15584
FV 8	FV 9	FV 10	FV 11			ID
0.59210	0.39953	0.31053	0.32999			1
						EM
						5



그림 6. 무표정 감성상태 실험 영상

표 5. ID가 1인 사람의 기쁨 감성상태 실험결과

FV 1	FV 2	FV 3	FV 4	FV 5	FV 6	FV 7
1.000	0.8400	0.8199	1.3512	1.3837	0.1410	0.1410
FV 8	FV 9	FV 10	FV 11			ID
0.5443	0.5466	0.2911	0.4121			1
						EM
						4

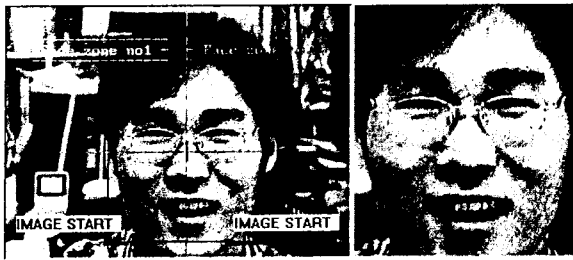


그림 7 기쁨 감성상태 실험 영상

표 6. ID가 1인 사람의 놀람 감성상태 실험결과

FV 1	FV 2	FV 3	FV 4	FV 5	FV 6	FV 7
1.000	0.7934	0.9637	1.2844	1.5286	0.1315	0.1447
FV 8	FV 9	FV 10	FV 11			ID EM
0.6233	0.5914	0.3861	0.3786			1 3



그림 8 놀람 감성상태 실험 영상

표 7. ID가 1인 사람의 화남 감성상태 실험결과

FV 1	FV 2	FV 3	FV 4	FV 5	FV 6	FV 7
1.000	0.8842	0.8792	1.4089	1.4490	0.1772	0.2151
FV 8	FV 9	FV 10	FV 11			ID EM
0.6025	0.4908	0.2951	0.3171			1 2



그림 9 화남 감성상태 실험영상

표 8. ID가 1인 사람의 슬픔 감성상태 실험결과

FV 1	FV 2	FV 3	FV 4	FV 5	FV 6	FV 7
1.000	0.8594	0.8444	1.3728	1.4197	0.1518	0.1772
FV 8	FV 9	FV 10	FV 11			ID EM
0.5443	0.5560	0.2914	0.3763			1 1



그림 10 슬픔 감성상태 실험영상

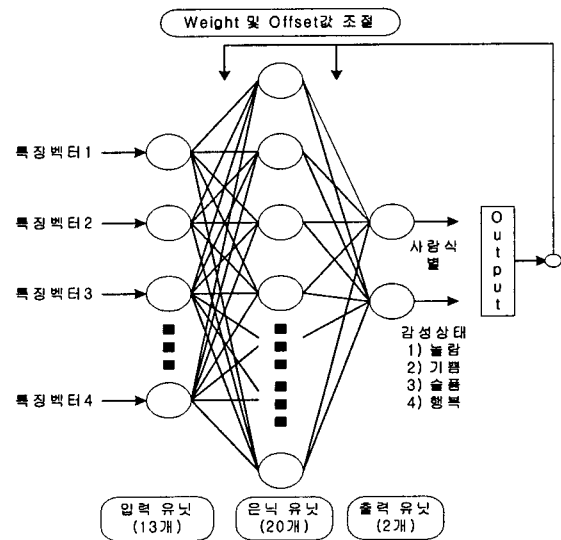


그림 11. 학습에 사용된 역전파 알고리즘

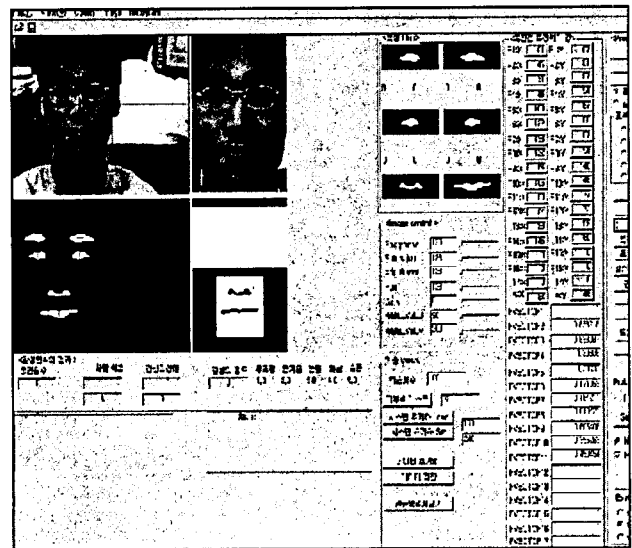


그림 12. 감성 인식 실험 예

### 3. 결 론

본 논문에서는 CCD 컬러 비전 시스템을 이용하여 인간의 감정 상태를 인식할 수 있는 감성 인식 비전 시스템을 제안하였다. 기존의 연구에서는 복합영상에서 얼굴 영역을 검출하는데 그 중요성을 다루었는데 본 논문에서는 복합영상에서 사람의 식별 및 감성상태를 동시에 인식하는 방법으로 확장 발전시켰다. 이를 위해 카메라와 사람사이의

거리에 따른 특징점의 검출 에러를 줄이기 위해서 표준화시킨 이미지를 생성하여 특징 blob 및 특징 벡터를 획득하는 과정에 사용하였다. 또한 얼굴 특징 blob 추출에 있어 그 신뢰도를 높이기 위해서 피부색 컬러 공간의 색차 정보를 이용하고, 감성 인식 시간을 줄이기 위해 간단한 기하학적 특징 벡터 산출법을 제안하였다. 마지막으로, 한사람에 대한 식별과 각각의 5가지 감성 상태를 사전에 학습시킨 신경회로망을 이용하여 실제 실험한 결과, 제안된 방법의 응용가능성을 확인하였다. 향후, 감성 인식 신뢰도를 높이기 위해, 복합 영상에서 주위 환경 변화에 따른 컬러 공간의 고유 피부색 값을 자기 동조할 수 있는 알고리즘에 관해 연구할 예정이며, 복수의 인간이 존재하고 주위 환경이 급변하는 복합 영상에서 각각의 사람 식별과 감성 상태를 인식 할 수 있는 시스템을 개발할 예정이다.

Multivariate Medical Data By Face Method," IEEE Trans. on Pattern Recognition, Vol. 15. No.3, pp. 212-242,1982.

[11] S. Morishima and H. Harashima, "Emotion Space for Analysis and Synthesis of Facial Expression," IEEE Int. Workshop on Robot and Human Communication, pp.188-193, 1993.

[12] H. Ushida, T. Takagi and T. Yamaguchi, "Recognition of Facial Expressions Using Conceptual Fuzzy Sets," IEEE Int. Conference on Fuzzy Systems, pp. 594-599, 1993.

감사의 글: 본 연구는 산업자원부의 차세대 신기술 개발사업(과제번호: N09-A08-4301-05)에 의해 지원 받았음.

#### (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] H. Wu, Q. Chen, and M. Yachida, "Face Detection from Color Images Using a Fuzzy Pattern Matching Method," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 21, No. 6, pp. 557-563, 1999.
- [6] H. Kbayashi and F. Hara, "The Recognition of Basic Facial Expressions by Neural Network," IEEE Int. Workshop on Robot and Human Communication, pp. 375-380, 1992.
- [7] H. Ebine and O. Nakamura "The Recognition of Facial Expressions Based on Production System," IEEE Int. Canadian Conference on Electrical, pp. 675-682, 1999.
- [8] I. Craw, H. Ellis and J. R. Lishman, "Automatic Extraction of Face Features," Pattern Recognition Letters, Vol. 5, pp. 183-187, 1987.
- [9] L. D. Harmon and W. F. Hunt, "Automatic Recognition of Human Face Profiles," Computer Graphics and Image Processing, Vol. 6, pp. 135-156, 1977.
- [10] N. Honda and S. Aida, "Analysis of