퍼지 및 지능시스템학회 논문지 2002, Vol. 12, No. 2, pp. 97-102

CCD 컬러영상에 의한 감성인식

Emotion Recognition by CCD Color Image

이상윤* · 주영훈* · 심귀보**

Sang-Yoon Lee*, Young Hoon Joo*, and Kwee-Bo Sim**

*군산대학교 전자정보공학부

**중앙대학교 전자전기공학부

요 약

본 논문에서는 CCD 칼라 영상을 이용하여 인간의 감성을 인식할 수 있는 방법을 제안한다. 먼저 CCD 카메라에 의해 획득한 칼라 영상으로부터 피부색 추출 방법을 이용하여 얼굴을 추출한다. 그 다음, 추출된 얼굴 영상으로부터 인간 얼굴의 특징점(눈썹, 눈, 코, 입) 들을 추출하는 방법과 각 특징점들 간의 구조적인 관계로부터 인간의 감성(놀람, 화남, 행복함, 슬픔)을 인식하는 방법을 제안한다. 본 논문에서 제안한 방법은 신경회로망을 이용하여 학습시킴으로써 인간의 감성을 인식한다. 마지막으로, 제안된 방법은 실험을 통해 그 응용 가능성을 확인한다

Abstract

In this paper, we propose the technique for recognizing the human's emotion by using the CCD color image. To do this, we first get the face image by using skin color from the original color image acquired by the CCD camera. And we propose the method for finding man's feature points(eyebrows, eye, nose, mouse) from the face image and the geometrical method for recognizing human's emotion (surprise, anger, happiness, sadness) from the structural correlation of man's feature points. The proposed method in this paper recognize the human's emotion by learning the neural network. Finally, we have proven the effectiveness of the proposed method through the experimentation.

Key words : Emotion recognition, CCD color camera, neural network, feature points, face recognition

1. 서 론

인간은 시각, 청각, 촉각, 후각, 미각 등 다섯 개의 감각을 통해 자신의 환경을 인식하게 된다. 지각과 인지 과정의 한 결과물이라 볼 수 있는 감성 역시 인간의 오감을 기초로 하여 과거의 경험과 추론 등이 더해져 만들어진다. 인간은 다른 사람들과 대화를 한다거나 자신의 감정을 전달하는 등의 일상생활에서 얼굴은 매우 중요한 요소로 작용되고 또한 그 역할이 매우 중요하다.

최근, 컴퓨터 및 광학 장비의 발전과 더불어 사용자의 의지나 감성을 인식하기 위한 연구가 활발히 진행되고 있다. 특히, 인간의 얼굴표정 인식 및 감성 인식 분야에 있어 그 관심이 높아지고 있다[1].

얼굴 영상으로부터 인간의 감정 상태를 인식하는 연구는 크게 두 가지로 나누어진다: 얼굴의 특징 점(눈썹, 눈, 코, 입)들의 위치 변화와 이들간의 기하학적인 관계로부터 유도된 특징 벡터(eigen vector)에 기반 한 방법[5-8, 10-12]과 사람 얼굴을 저 차원으로 표현할 수 있도록 하여 얼굴 영상의 기본 요소를 분석할 수 있도록 하는 형판 매칭(template matching) 방법이 있다[3-4]. 형판 매칭 방법은 획득한 얼굴 영상을 하나의 형판 영상과 비교하여 이에 따른 상관 관계를 분석함으로써 얼굴을 인식하는 것이다. 이 방법은 단순한 처리로 좋은 성능을 발휘한다는 장점이 있지만 영상의 밝기에 따라 상관도 값이 크게 변동한다는 점과 전체 얼굴에 대한 영상을 형판 데이터로 하기 때문에 메모리가 커진다는 등의 단점이 있다. 한편, 특징벡터에 기반한 방법은 빠른 인식속도와 이미지 자체를 저장하지 않고 벡터의 데이터 값을 기반으로 하기 때문에 메모리 효율이 좋다는 장점이 있다. 따라서, 특징 벡터에 기반한 연구가 활발히 진행되고 있다.

본 논문에서는, 비전 시스템에 의해 획득한 칼라 영상으로부터 인간의 감성 상태를 인식하는 방법을 제안한다. 본 논문에서는 칼라 카메라을 이용한 인간의 감성인식 방법을 제안한다. 제안한 방법은 다음과 같다: 첫째, 획득한 복합 영상으로부터 피부색을 이용하여 얼굴 영상만을 추출하기 위해 R, G, B 컬러 데이터를 이용하는 방법을 제안한다. 둘째, 추출된 얼굴 영상으로부터 얼굴의 특징 요소(눈썹, 눈, 코, 입)를 찾아내는 알고리즘을 제안한다. 셋째, 추출된 특징 요소로부터 인간의 감성 상태를 조사하기 위한 특징 벡터 산출 방법을 제안한다. 넷째, 신경회로망의 역전파 알고리즘을 이용하여 특징 벡터 데이터를 학습함으로써 개인의 식별 및 감정 상태(놀람, 분노, 즐거움, 슬픔)을 인식하는 방법을 제안한다.

마지막으로, 본 연구실에 구축된 비전 시스템을 이용하

접수일자 : 2002년 2월 15일
완료일자 : 2002년 4월 5일
본 연구는 산업자원부의 차세대 신기술개발사업(과제번호: N09-A08-4301-05)에 의해 지원 받았음.

여 실험을 통해 제안한 방법의 응용 가능성을 조사한다.

2. 감성 인식을 위한 비전 시스템

2.1 얼굴 영역 추출

초기에 획득한 칼라 영상으로부터 얼굴 영역의 위치를 찾기 위해 칼라 공간의 특정 색차 (chrominance) 정보를 이용한다. 즉, 찾고자 하는 영상은 얼굴이라는 특정 부분이기 때문에 얼굴만의 특정 색조라 할 수 있는 피부색(skin color)을 검출 기준으로 삼고 각각의 컬러 표현 방식의 공간 변환을 통해서 복합 영상으로부터 얼굴 부분을 추출하는 방법을 사용한다.

실험에 적용한 컬러 공간 변환은 두 단계를 거친다.

첫 번째는 초기의 RGB 색조 데이터로부터 피부색에 유난히 민감하게 나타나는 YIQ 컬러 공간으로 변환하는 것이다. 여기서, Y(luminance) 인자는 색의 밝기 값을 나타내고 I, Q(color image chrominance) 는 각각의 색조의 데이터가 된다.

두 번째는 초기의 RGB 색조 데이터 공간을 HSI(hue, saturation, intensity) 컬러 공간으로 변환하는 것이다. 즉, I의 밝기 값에 따른 H의 피부색 범위를 식 (1) 과 같이 각각 찾아내어 최적의 밝기 값에 의한 H의 피부색 범위를 선택하게 된다.

$$p_k=\begin{cases} aq_k & \text{if } 5\le H\le 40 \\ 0 & \text{if } H>40 \text{ or } H<5 \end{cases} \quad (1)$$

여기서, H : his 컬러 공간상의 피부색 분포 영역,

q_k: 절할 영상 (i의 값에 의한 흑백 영상)

p_k: 조절된 흑백 결과 영상

α : 조절 상수

a) 초기 영상 b) 피부색 추출영상 c) 얼굴 추출 영상

그림 1. 얼굴 추출 단계

Fig. 1. Detection step of face region

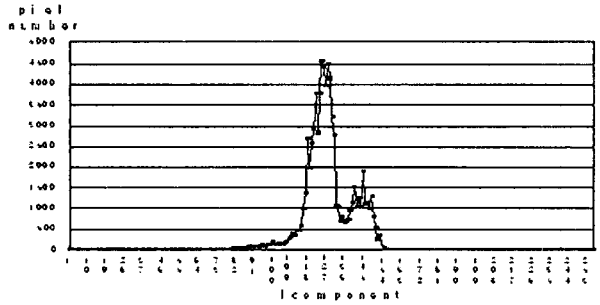


그림 2. YIQ 컬러 공간에서 I 성분의 분포

Fig. 2. Distribution of I component in YIQ color space

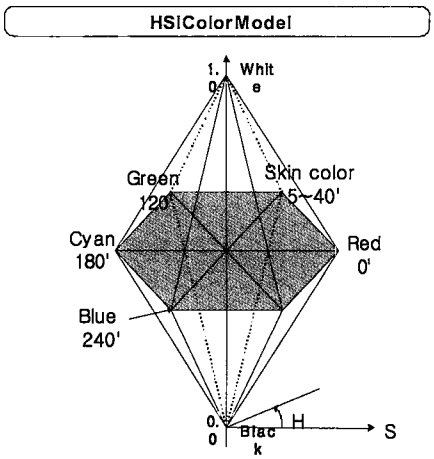


그림 3. HSI 컬러공간에서의 H의 피부색 분포 영역

Fig. 3. Distribution of H space in HSI color space

그림 1은 초기 복합 영상으로부터 얼굴 영역을 찾는 한 예를 도시한 그림이다. 그림 1-a와 같은 초기 복합 영상으로부터 YIQ 칼라 공간 변환(그림 2)과 HSI 컬러 공간 변환(그림 3)을 거친 최종 피부색 추출 영상은 그림 1 b에 나타나 있다. 그 다음 단계로, 피부색 추출 영상으로부터 얼굴 영역만을 추출하는 방법이 필요하다. 즉, 그림 1 b와 같은 영상에서 얼굴의 피부색 분포 범위 안의 물체들 중 얼굴에 해당하는 부분만을 분리해야 우리가 원하는 최종 얼굴 추출 영상을 획득할 수 있다. 본 논문에서는 얼굴 외 물체와 얼굴을 분리하기 위해 각 물체를 라벨 처리한 후, 크기 필터링(size filtering)을 적용하는 순차 연결성분 알고리즘을 통해 최종 그림 1-c와 같은 얼굴 추출 영상을 획득한다.

그림 4는 초기 복합 영상으로부터 얼굴 영역을 정규화 한 영상을 나타낸다. 이는 카메라와 대상체의 거리에 따른 영상의 크기를 일정하게 해석하기 위해 필요하다. 본 논문에서는 정규화 영상을 얻기 위해 b-스플라인 보간법을 사용한다.

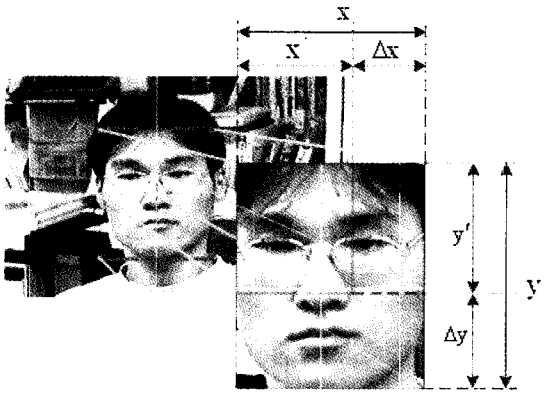


그림 4. 정규화 영상

Fig. 4. The normalized image

즉, 식 (2)와 (3)을 이용하여 x축과 y축에 매핑된 화소들의 배율을 구한다. 그리고, 식 (4)로 표현되는 b-스플라인 보간식을 이용하여 영상을 확대한다. 마지막으로 식 (5)를 이용하여 정규화 영상을 얻는다.

$$a_x = \frac{x}{\hat{x}} \quad a_x : x\text{축 화소의배율} \tag{2}$$

$$a_y = \frac{y}{\hat{y}} \quad a_y : y\text{축 화소의배율} \tag{3}$$

$$f(x) = \begin{cases} \frac{1}{2}|x|^3 - |x|^2 + \frac{2}{3} & 0 \le |x| < 1 \\ -\frac{1}{6}|x|^3 + |x|^2 - 2|x| + \frac{4}{3} & 1 \le |x| < 2 \\ 0 & 2 \le |x| \end{cases} \tag{4}$$

$$\hat{f}(x) = \max(a_x, a_y) * f(x) \tag{5}$$

여기서, $\hat{f}(x)$는 정규화한 영상이고 $f(x)$는 초기의 얼굴 영상이다.

2.2 얼굴의 특징 점 추출

2.2절에서는 초기의 복합 영상으로부터 피부색 추출 영상을 얻고, 이 영상으로부터 얼굴 부분만을 획득하는 방법과 모든 영상에서 얼굴 부분의 크기를 일정하게 하는 정규화 영상을 만들어 내는 과정을 설명하였다. 본 절에서는 정규화된 얼굴 영상으로부터 얼굴의 특징 요소인 눈썹, 눈, 코, 입등을 추출하는 방법을 설명한다.

그림 5 a는 초기의 복합 영상(그림 1 a)으로부터 획득한 정규화된 영상을 나타낸다. 그림 5 a로부터 피부색 추출을 위한 2 단계 변환과정을 거쳐 얻은 피부색 추출 영상은 그림 5-b에 나타나 있다. 그림 5 c는 그림 5 a에 나타난 영상을 이진화하여 얻은 이진화 영상을 나타낸다. 본 논문에서는 얼굴의 특징 요소들을 찾기 위해 그림 5 b와 그림 5 c의 두 영상의 중복된 화소를 이용하여 그림 6 a 와 같은 특징점 추출 영상을 만든다. 그림 6 a의 영상을 라벨 처리한 후, 크기 필터링을 거쳐 그림 6 b ~ 6 f와 같이 얼굴의 특징 요소들을 얻는다.

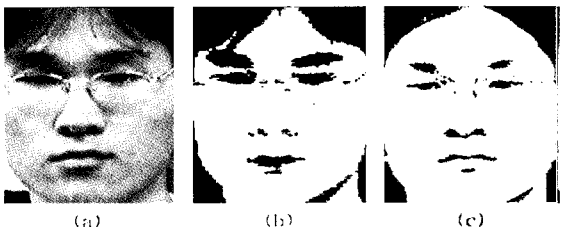


그림 5. 얼굴의 특징점 추출 과정
Fig. 5. Process for extrating the face features

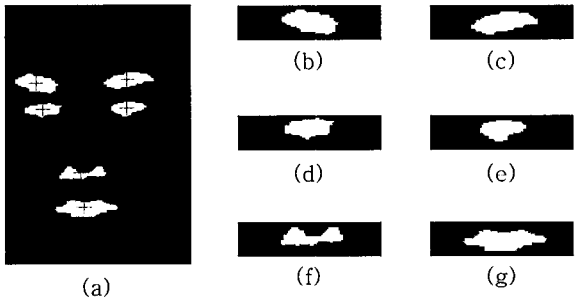


그림 6. 특징점 추출 이미지
Fig. 6. Image for feature extraction

2.3 특징벡터 추출

본 절에서는 그림 6에 의해 추출된 특징 요소들의 좌표 점들로부터 얼굴의 특징 벡터를 구하는 방법을 설명한다. 인간은 감정의 변화에 따라 눈썹, 눈, 코 또는 입의 위치나 크기상에 변화가 일어난다. 이런 변화를 감성인식에 이용하기 위해 본 논문에서는 코의 중심 좌표 점을 기준으로 기하학적인 근육구조의 14가지 특징벡터를 산출하는 방법을 제안한다. 그림 7은 본 논문에서 제안한 얼굴의 특징 벡터를 나타낸 영상이며 표 1은 14개의 특징 벡터를 산출해 내는 식을 나타낸다. 본 논문에서 제안된 방법에 의해 산출된 특징 벡터는 감성 인식의 중요한 판단 자료가 된다.

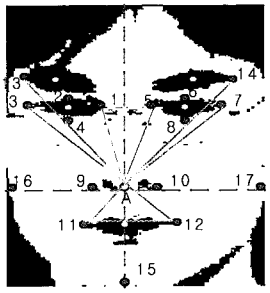


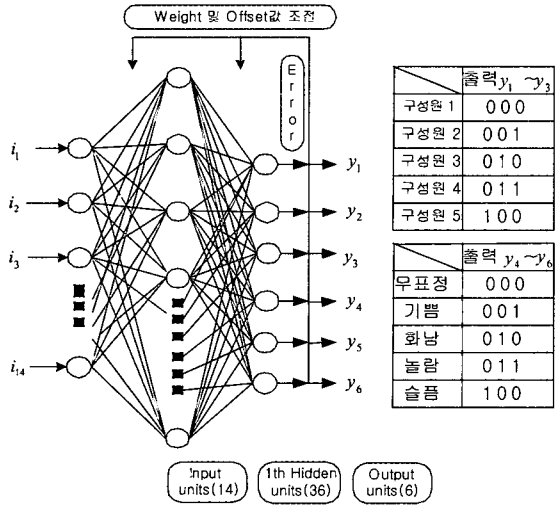
그림 7. 특징벡터
Fig. 7. Feature vector

표 1. 특징 벡터 산출 방법
Table 1. Method for extracting feature vectors

코의 중심	$A = \frac{(point10 - point9)}{2} = x_n - y_n$
코의 중심-> 왼쪽 눈의 오른쪽 끝점	$(A, 1) = \sqrt{(x_n - x_1^2) + (y_n - y_1)^2}$
코의 중심-> 오른쪽 눈의 왼쪽 끝점	$(A, 5) = \sqrt{(x_5 - x_n^2) + (y_n - y_5)^2}$
코의 중심-> 왼쪽 눈썹의 오른쪽 끝점	$(A, 13) = \sqrt{(x_n - x_{13}^2) + (y_n - y_{13})^2}$
코의 중심-> 오른쪽 눈썹의 오른쪽 끝점	$(A, 14) = \sqrt{(x_{14} - x_n^2) + (y_n - y_{14})^2}$
왼쪽 눈의 수직거리	$(4, 2) = (y_4 - y_2)$
오른쪽 눈의 수직거리	$(8, 6) = (y_8 - y_6)$
왼쪽 눈에서 오른쪽 눈까지 거리	$(5, 1) = (x_5 - x_1)$
코의 중심에서 입술 왼쪽 끝점까지 거리	$(11, A) = \sqrt{(x_n - x_{11}^2) + (y_{11} - y_n)^2}$
코의 중심에서 입술 오른쪽 끝점까지 거리	$(12, A) = \sqrt{(x_{12} - x_n^2) + (y_{12} - y_n)^2}$
입술 왼쪽 끝에서 오른쪽 끝까지 거리	$(12, 11) = \sqrt{(x_{12} - x_{11}^2) + (y_{12} - y_{11})^2}$
코의 중심에서 턱 끝까지 거리	$(15, A) = (y_{15} - x_n)$
코의 중심에서 얼굴 왼쪽 면 끝까지 거리	$(A, 16) = (x_n - x_{16})$
코의 중심에서 얼굴 오른쪽 면 끝까지 거리	$(17, A) = (x_{17} - x_n)$

2.4 감성인식을 위한 학습 알고리즘

본 논문에서는 14개의 기하학적인 특징 벡터 값을 산출한 후, 역전파 신경망의 학습을 통해 감성인식 시스템을 설계한다. 여기서, 신경망의 입력은 14개의 특징 벡터 값이며, 출력은 사람의 인식을 위해 사용한 3비트의 이진 출력 (y_1, y_2, y_3)과 감정 인식을 위해 사용한 3비트의 이진 출력 (y_4, y_5, y_6)으로 구성된다. 그림 8은 학습을 위해 사용된 신경망의 구조를 나타낸 그림이며, 학습 알고리즘은 참고문헌 [9]에 나타나 있다.



	출력 $y_1 \sim y_3$
구성원 1	000
구성원 2	001
구성원 3	010
구성원 4	011
구성원 5	100

	출력 $y_4 \sim y_6$
무표정	000
기쁨	001
화남	010
놀람	011
슬픔	100

그림 8. 역전파 알고리즘
Fig. 8. Back-propagation algorithm

3. 실험 및 결과 고찰

본 논문에서 제안한 방법의 실 적용 실험을 위해 5명의 사람을 대상으로 5가지의 감정(무표정, 놀람, 화남, 슬픔, 기쁨)에 대해 각각 50개씩의 학습 데이터 집합을 구성한다. 따라서, 학습 데이터 집합의 수는 총 2,500개가 된다. 또한, 본 실험을 위한 제약 조건은 다음과 같다: 첫째, 피부색의 색차를 이용한 컬러 데이터 변환을 전제로 하기 때문에 피부색에 가까운 의상이나 기타 얼굴을 막는 장애물은 없어야 한다. 둘째, 복수의 인간이 존재하지 않는다. 현재 복수 대상체일 경우 두 사람 중 한 사람에 대한 인식과정으로 넘어 간다.

표 2는 5명의 사람과 각 개인에 대해 5개의 감성상태를 나타내는 목표 데이터의 한 예이며, 그림 9는 본 논문에서 제안한 방법의 실험을 위해 제작된 감성 인식 실험 시스템의 초기 화면을 나타낸 것이다.

본 논문에서 제안한 감성 인식 시스템의 실험을 위해 그림 10과 같은 5개의 실험 영상을 사용하였다. 그림 10-a에서 그림 10-e까지의 영상은 각각 무표정, 기쁨, 놀람, 화남, 슬픔을 나타내는 영상이다. 표 3은 5개의 영상에 대한 실험 결과를 나타내는 입력 데이터와 출력 데이터를 나타낸 표이다. 표 3에서 알 수 있듯이 5개의 실험 영상으로부터 각각의 감성을 인식한 결과, 제안한 방법은 정확히 감성을 인식했음을 알 수 있다.

표 2. 목표 데이터의 예
Table 2. Example of target data

출력 예	목표값(인간, 감성)
(ID 1, 무표정)	[0,0,0, 0,0,0]
(ID 2, 기쁨)	[0,0,1, 0,0,1]
(ID 3, 화남)	[0,1,0, 0,1,0]
(ID 4, 놀람)	[0,1,1, 0,1,1]
(ID 5, 슬픔)	[1,0,0, 1,0,0]

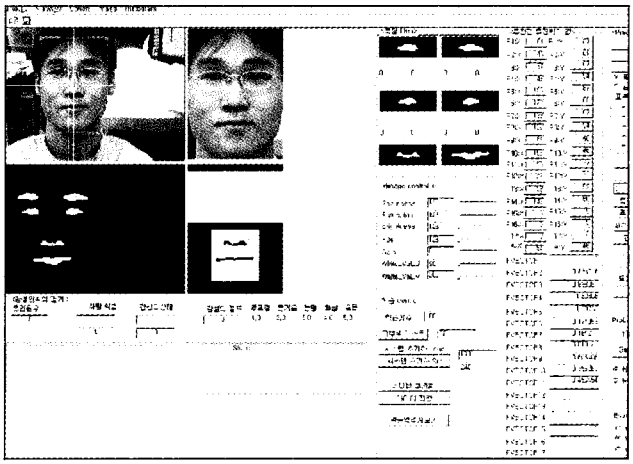


그림 9. 감성 인식 실험 시스템
Fig. 9. Experimental system for the emotion recognition

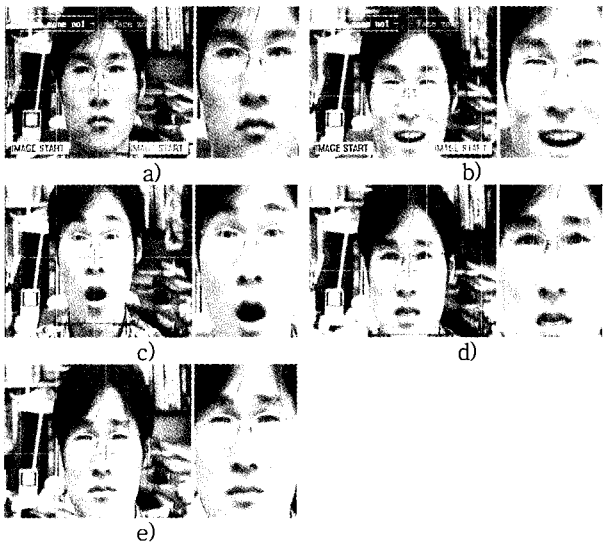


그림 10. 실험을 위한 영상
Fig. 10. Images acquired for experiment

표 3. 실험 결과 입출력 데이터
Table 3. The experimental results

	a	b	c	d	e
i_1	1.00	1.00	1.00	1.00	1.00
i_2	0.90	1.00	0.93	0.95	0.99
i_3	1.00	0.90	0.90	1.00	0.93
i_4	0.90	0.99	0.99	1.00	1.00
i_5	0.95	0.89	0.90	1.00	0.93
i_6	0.67	1.00	0.61	0.89	0.67
i_7	0.76	1.00	0.82	1.00	0.82
i_8	0.97	1.00	0.92	0.90	0.94
i_9	0.75	1.00	0.98	0.44	0.96
i_{10}	0.84	0.68	0.68	1.00	0.73
i_{11}	0.78	0.96	0.98	0.31	0.87
i_{12}	1.00	0.98	1.00	0.92	0.93
i_{13}	1.00	0.99	0.92	1.00	0.96
i_{14}	1.00	0.93	0.91	0.97	1.00
y_1	0.00	0.00	0.00	0.00	0.00
y_2	0.00	0.00	0.00	0.00	0.00
y_3	0.00	0.00	0.00	0.00	0.00
y_4	0.00	0.00	0.00	0.00	1.00
y_5	0.00	0.00	1.00	1.00	0.00
y_6	0.00	1.00	0.00	1.00	0.00

4. 결 론

본 논문에서는, 비전 시스템에 의해 획득한 칼라 영상으로부터 인간의 감성 상태를 인식하는 방법을 제안하였다. 제안한 방법은 다음과 같다: 첫째, 획득한 복합 영상으로부터 피부색을 이용하여 얼굴 영상만을 추출하기 위해 R, G, B 컬러 데이터를 이용하는 방법을 제안하였다. 둘째, 추출된 얼굴 영상으로부터 얼굴의 특징 요소(눈썹, 눈, 코, 입)를 찾아내는 알고리즘을 제안하였다. 셋째, 추출된 특징 요소로부터 인간의 감성 상태를 조사하기 위한 특징 벡터 산출 방법을 제안하였다. 넷째, 신경회로망의 역전파 알고리즘을 이용하여 특징 벡터 데이터를 학습함으로써 개인의 식별 및 감정 상태(놀람, 분노, 즐거움, 슬픔)을 인식하는 방법을 제안하였다. 마지막으로, 본 연구실에 구축된 비전 시스템을 이용하여 직접 실험을 통해 제안한 방법의 응용 가능성을 확인하였다.

참 고 문 헌

[1] 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.

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

[3] M. Turk and A. Pentland, "Face recognition Using Eigenface", Proc. of the IEEE Conference on Computer Vision, pp. 586-591, 1991.

[4] H. Ebine and O. Nakamura "The Recognition of Facial Expressions Based on Production System," IEEE Int. Canadian Conference, pp. 675-682, 1999.

[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] N. Honda and S. Aida, "Analysis of Multi-variate Medical Data By Face Method". Pattern Recognition, vol. 15. pp. 212-242, 1982.

[7] L.D. Harmon and W.F. Hunt, "Automatic Recognition of Human Face Profiles", Computer Graphics and Image Processing, pp. 135-156, 1977.

[8] I. Craw, H. Ellis, and J.R. Lishman, "Automatic Extraction of Face Features", Pattern Recognition Letters, vol. 5, pp. 183-187, 1987.

[9] 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.

[10] I. Craw, H. Ellis and J. R. Lishman, "Automatic Extraction of Face Features", Pattern Recognition Letters, vol. 5, pp. 183-187, 1987.

[11] L. D. Harmon and W. F. Hunt, "Automatic Recognition of Human Face Profiles", Computer Graphics and Image Processing, vol. 6, pp. 135-156. 1977.

[12] N. Honda and S. Aida, "Analysis of Multivariate Medical Data By Face Method", IEEE Trans. on Pattern Recognition, vol. 15. no. 3, pp. 212-242, 1982.

저 자 소 개

이상윤(Sang-Yoon Lee)
2001년 : 군산대 제어계측공학 졸업
2001년-현재 : 군산대 대학원 전자정보공학과 석사과정

관심분야 : 컴퓨터 비전시스템, 감성공학, 지능제어.

Phone : 063-469-4706
Fax : 063-469-4706
E-mail : cie93@netian.com

주영훈(Young-Hoon Joo)

1982년 : 연세대 전기공학과 졸업
1984년 : 연세대 대학원 전기공학과 졸업
1995년 : 동대학원 전기공학과 졸업 (공박)
1986～1995년 8월 : 삼성전자(주) 생산기술센타 자동화연구소(선임 연구원)
1998년 1월～1999년 2월: 미국 University of Huston 전기 및 컴퓨터공학과 (Post-doctor)
2000년～현재 : 한국 퍼지 및 지능시스템학회 편집이사
2001년～현재 : 대한전기학회 제어및시스템부문회 편집위원
1995년 9월～현재 : 군산대 공대 전자정보공학부 부교수

관심분야 : 퍼지 모델링, 유전자 알고리즘, 지능제어, 지능형 디지털 재설계. 이동로봇 시스템.

Phone : 063-469-4706
Fax : 063-469-4706
E-mail : yhjoo@kunsan.ac.kr

심귀보(Kwee-Bo Sim)

1984년 : 중앙대학교 전자공학과 공학사
1986년 : 동 대학원 전자공학과 공학석사
1990년 : The University of Tokyo 전자공학과 공학박사
2000년～현재 : 한국퍼지 및 지능시스템학회 편집이사 및 논문지 편집위원장
2001년～현재 : 대한전기학회 제어및시스템부문회 편집위원 및 학술이사
2002년～현재 : 제어자동화시스템공학회 이사
1991년～현재 : 중앙대학교 전자전기공학부 교수

관심분야 : 인공생명, 진화연산, 지능로봇시스템, 뉴로-퍼지 및 소프트 컴퓨팅, 자율분산시스템, 로봇비전, 진화하드웨어, 인공면역계 등

Phone : +82-2-820-5319
Fax : +82-2-817-0553
E-mail : kbsim@cau.ac.kr