

퍼지 추론을 기반으로 한 칼라 영상에서의 감성 인식

Fuzzy Inference-Based Human Emotion Recognition of Color Image

정 근호, 나 인호, 최 연성, 양 해권, 주 영훈
Keun-Ho Jeong, In Ho Ra, and Yeon Sung Choi, Hae Kwon Yang and Young-Hoon Joo

군산대학교 전자정보공학부

요 약

본 논문에서는 칼라 영상을 이용하여 인간의 감성을 인식할 수 있는 방법을 제안한다. 먼저 칼라 영상으로부터 피부색 추출방법을 이용하여 얼굴을 추출한다. 그 다음, 추출된 얼굴 영상으로부터 인간 얼굴의 특징 점(눈썹, 눈, 코, 입)들을 추출하는 방법과 각 특징 점들 간의 구조적인 관계로부터 인간의 감성(기쁨, 놀람, 슬픔, 분노)을 인식하는 방법을 제안한다. 본 논문에서 제안한 방법은 퍼지 추론을 기반으로 하여 인간의 감성을 인식한다. 마지막으로, 제안된 방법은 실험을 통해 그 응용 가능성을 확인한다.

1. 서 론

과학기술의 발달로 인간 생활과 사고방식이 변하고 있다. 완전 자동화되고, 일일이 손으로 기계조작을 해주는 것이 아닌 인간과의 의사소통방법을 요구하고 있다. 이 시점에서 필요로 하게된 기술 중의 하나가 사용자와의 감정교류를 위한 감성정보처리기술이다[1~3]. 이 기술은 감성정보(emotional information)를 통하여 사용자를 인식하고, 각종 표정이나 몸짓, 동작 등에서 정보 추출을 통하여 적절한 행동을 취할 수 있는 인공지능의 산물을 만들어내는데 그 목적을 둔다. 인간은 다른 사람들과 대화를 한다거나 자신의 감정을 전달하는 등의 일상 생활에서 얼굴은 매우 중요한 요소로 작용된다. 현재, 환경변화에 강인한 인간 얼굴의 인식 방법과 표정변화에 따른 인간의 감성을 알아내는 연구가 활발히 진행되고 있다 [1~8]. 얼굴영역 추출을 위한 방법은 지형적 특징추출에 기반한방법과 형판정합에 기반한 방법, 얼굴의 색상정보를 이용한 방법, 시간정보를 이용한 방법 등이 있다 [4]. 또한, 얼굴 요소의 영역추출을 위한 방법들이 있는데 이것은 투영분석 (projection analysis), 색상정보를 이용한 방법, 얼굴요소의 특성에 기반 한 추출방법과 유전자 알고리즘을 이용한 방법 등으로 나누어 볼 수 있다. 마지막으로 얼굴 영상으로부터 윤곽선을 추출하는 방법들이 있다. 이외에도 신경망, 유전자 알고리즘을 이용한 감성인식 연구 방법들이 최근 진행되고 있다[4, 6~8, 10]. 특징 점의 상관관계로부터 인간의 기본감성은 분노(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 감성 중 기쁨, 놀람, 슬픔, 분노의 4가지 극단적인 감성을 인식하기 위하여, 본 논문에서는 퍼지 추론을 기반으로 한 인간의 감정 식별 방법을 제안한다. 먼저, 칼라영상으로부터 피부색을 이용하여 얼굴영상을 얻는다. 그 다음, blob처리와 labeling을 거친 후 인간의 감성표현에 있어서 큰 움직임이 보이는 눈썹, 눈, 입의 특징 점을 정규화 한 영상 영역 안에서 추출한다. 어떠한 감성에서도 거의 움직임이 없는 양쪽 눈의 중앙의 중앙 점을 기준으로 한 각 영역으로부터 얼굴의 움직임을 검출하고, 이러한 움직임의 판단을 통합함으로써 각 표정의 관여되는 퍼지 규칙을 사용하여 인간의 감성 (기쁨, 놀람, 슬픔, 분노)을 식별한다. 마지막으로, 본 연구실에 구축된 실험시스템을 통해 제안된 방법의 응용가능성을 보인다.

2. 얼굴영역 정규화 및 감성인식

2.1 얼굴영역 추출

얼굴영역은 HSI 컬러 모델을 기반으로 하고 이진화 표현된 피부색 영상으로부터 피부색 범위에 들어오는 다른 복합영상의 노이즈들을 제거하기 위해 각 blob들을 label 처리한 후, size filtering을 적용하여 추출된다. RGB 컬러 모델은 적색(Red), 녹색(Green), 청색(Blue)으로 한 모델로써 밝기에 대한 정보도 포함되어 있으므로 이에 민감하다는

단점을 가지고 있다. 그러한 반면에 HSI 컬러모델은 색을 인식하는 속성인 색상(Hue), 채도(Saturation), 명도(Intensity)로 구성되어 있어 밝기 성분을 별도로 분할하여 사용할 수 있기에 RGB 컬러 모델보다 조명에 덜 민감하게 사용할 수 있기 때문에 본 논문에서는 얼굴영역 추출에 있어 RGB 컬러모델을 HSI 컬러 모델로 변환하여 사용하고 변환 식은 식 (1), (2), (3)과 같다.

$$H = \cos^{-1} \left(\frac{\frac{1}{2} [(R-G) + (R-B)]}{\sqrt{[(R-G)^2 + (R-B)(G-B)]}} \right) \quad (1)$$

$$S = 1 - \frac{3}{R+G+B} [\min(R, G, B)] \quad (2)$$

$$I = \frac{1}{3} (R + G + B) \quad (3)$$

색상은 0~360°로 나타내며 채도와 명도는 0~100%로 표현되어진다. 그림 1은 실험에서 사용되어진 HSI 컬러 모델 공간에서의 피부색 분포영역인 H의 범위를 나타내고 그림 2는 HSI에서의 피부색 경계 영역을 나타내며, 그림 3은 이 모델을 통하여 실험에 의해 획득한 얼굴 영역을 보여준다.

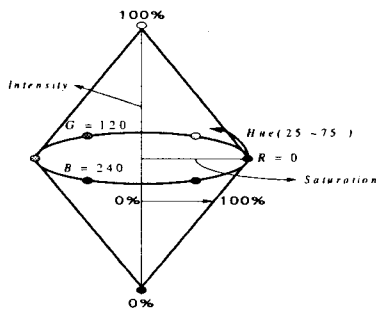


그림 1. HSI에서의 피부색 영역

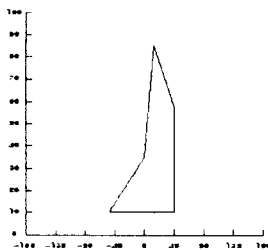


그림 2. HSI에서의 피부색 경계 영역

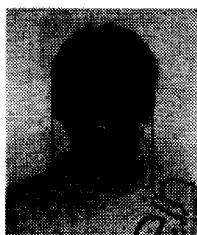


그림 3. 얼굴 영역

Labeling과 거리에 따른 size filtering을 사용하여 Noise를 제거한 영상을 이용하여 특징 점을 추출한다. 그림 4, 5 모두 놀람 감성 영상을 바탕으로 하였으며, 그림 4은 본 논문에서 제안한 얼굴의 특징 점을 나타낸 영상이다. 인간은 감성의 변화에 따라 눈썹, 눈, 코 또는 입의 위치나 크기 상에 변화가 일어난다. 이런 변화를 감성 인식에 이용하기 위해 특징 점들의 좌표들로부터 얼굴의 특징 벡터를 구하는 방법에 있어 기준 점은 인간의 얼굴에 있어 어떠한 감성에서도 변화가 거의 없는 왼쪽 눈과 오른쪽 눈 중심간의 중심점을 기준으로 하여 기하학적인 근육구조의 12가지 특징벡터를 산출하는 방법을 이용하여 특징벡터를 추출하는 방법을 제안한다. 일반적인 벡터 산출방법은 거리에 따라서 영상크기가 변화하기 때문에 이러한 문제를 극복하기 위해서 본 논문은 다른 방식의 벡터 산출방법을 제안한다. 이 방법은 비율(ratio)을 사용하는 것이다. 컬러 영상에서 검출되어진 각각의 얼굴 영역들은 특징벡터 값들이 조금씩 차이가 있다. 컬러 영상들이 거리에 민감할 경우 기하학적 근육구조를 기반으로 한 본 실험에서는 정확한 특징벡터 값을 구하기란 쉽지 않다. 이러한 문제점들을 해결하기 위하여 본 논문에서는 비율 즉, 각각의 특징 점들 중에서 기준 점과 왼쪽 눈의 중앙 점까지의 특징 벡터 값을 기준으로 하여 모든 벡터 값들을 나누어 각 영상에 대한 비율로서 이러한 문제점을 극복하여 거리에 상관없이 각 영상에 대한 정확한 특징 벡터 값을 얻는다. 그림 5은 위에서 설명한 비율을 이용한 얼굴의 특징 벡터들이 실험을 통해 얻어진 결과 영상을 나타낸 것이다. 표 1은 12개의 특징 벡터 값을 산출해 내는 식을 나타낸다. 본 논문에서 제안된 방법에 의해 산출된 특징 벡터는 퍼지 추론 알고리즘을 이용한 감성 인식에 있어서 매우 중요한 판단 자료가 된다.

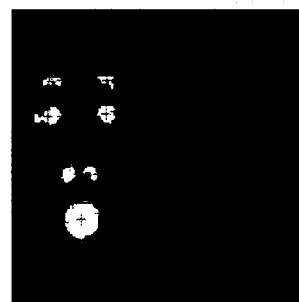


그림 4. 얼굴 영역에서의 특징 점

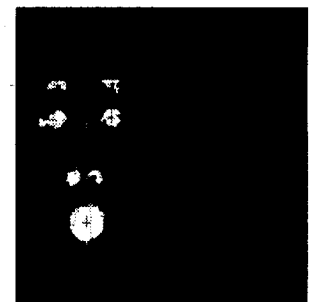


그림 5. 얼굴 영역에서의 특징 벡터

2.2 특징벡터 추출

본 절에서는 정규화 된 얼굴 영상으로부터 얼굴의 특징 요소인 눈썹, 눈, 입 등을 추출하는 방법과 이로 인해 특징 벡터를 구하는 방법에 대하여 설명한다. 본 논문에서는 좀더 정확한 특징벡터 추출을 위해 얼굴의 특징 점 추출에 있어서 정규화 된 컬러 영상을 이진화 한 후 확대하여 HSI, YCbCr, YIQ 색 모델을 재적용 시키고, blob처리를 한다.

2.3 감성 인식

본 논문에서는 12개의 기하학적인 특징 벡터 값을 산출한 후, 특징 벡터값을 입력으로 사용하여 각각 12개의 삼각형 소속함수를 사용하여 정규화 된 영상을 기준으로 하여 실험 영상들의 특징벡터들 간의 차이 값(m)으로 값을 구한다. 그 입력공간에서의 입력변수는 퍼지 근사 추론 방법 중에 하나

인 mamdani의 min-max방법을 쓴다.

표 1. 특징 벡터 산출 방법

기준점 $A(x,y)$	$A(x,y)$
기준점 ▶ 왼쪽 눈 중앙 점 $0(x_0,y_0)$	$\frac{\sqrt{(x_0-x)^2+(y_0-y)^2}}{2} = B$
기준점 ▶ 왼쪽 눈썹 왼쪽 끝점 $1(x_1,y_1)$	$\frac{\sqrt{(x_1-x)^2+(y_1-y)^2}}{B}$
기준점 ▶ 왼쪽 눈썹 오른쪽 끝점 $2(x_2,y_2)$	$\frac{\sqrt{(x_2-x)^2+(y_2-y)^2}}{B}$
기준점 ▶ 오른쪽 눈썹 왼쪽 끝점 $3(x_3,y_3)$	$\frac{\sqrt{(x_3-x)^2+(y_3-y)^2}}{B}$
기준점 ▶ 오른쪽 눈썹 오른쪽 끝점 $4(x_4,y_4)$	$\frac{\sqrt{(x_4-x)^2+(y_4-y)^2}}{B}$
기준점 ▶ 왼쪽 눈 왼쪽 끝점 $5(x_5,y_5)$	$\frac{\sqrt{(x_5-x)^2+(y_5-y)^2}}{B}$
기준점 ▶ 왼쪽 눈 오른쪽 끝점 $6(x_6,y_6)$	$\frac{\sqrt{(x_6-x)^2+(y_6-y)^2}}{B}$
기준점 ▶ 오른쪽 눈 왼쪽 끝점 $7(x_7,y_7)$	$\frac{\sqrt{(x_7-x)^2+(y_7-y)^2}}{B}$
기준점 ▶ 오른쪽 눈 오른쪽 끝점 $8(x_8,y_8)$	$\frac{\sqrt{(x_8-x)^2+(y_8-y)^2}}{B}$
기준점 ▶ 입술 왼쪽 끝점 $9(x_9,y_9)$	$\frac{\sqrt{(x_9-x)^2+(y_9-y)^2}}{B}$
기준점 ▶ 입술 오른쪽 끝점 $10(x_{10},y_{10})$	$\frac{\sqrt{(x_{10}-x)^2+(y_{10}-y)^2}}{B}$
기준점 ▶ 입술 위쪽 끝점 $11(x_{11},y_{11})$	$\frac{\sqrt{(x_{11}-x)^2+(y_{11}-y)^2}}{B}$
기준점 ▶ 입술 아래쪽 끝점 $12(x_{12},y_{12})$	$\frac{\sqrt{(x_{12}-x)^2+(y_{12}-y)^2}}{B}$

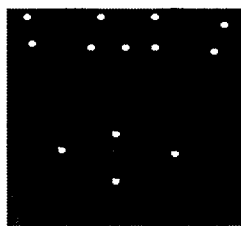


그림 6. 특징 벡터 순번

그림 6은 퍼지 규칙에서 사용되어지는 특징 벡터 순번을 나타낸다. 그림 7는 본 논문에서 사용되어진 각각의 12개 특징 벡터들의 삼각형 소속함수를 나타낸다. $E_{i,j}$ 에서 i 는 12개의 특징 벡터를 의미하고 j 는 그림 7에서 보여준 크기에 대한 4가지 (very short, short, long, very long)의 소속함수를 나타낸다. 예를 들면 $i=2$ 일 때 (그림 6에서 보여준 특징벡터 순번) 특징벡터의 차이 값이 m 일 경우 $E_{2,1}$ 는 적합도가 0.75이고, $E_{2,2}$ 는 0.2값을 가지며 $E_{2,3}$ 는 0.0, $E_{2,4}$ 는 0.0값이 구해지는 것을 볼 수가 있다. 감성 인식은 특징벡터 차이 값에 대한 각각의 12개의 소속함수를 가지고 있으며, 크게 4가

지 감성에 대한(기쁨, 놀람, 슬픔, 분노) 규칙과 사전 지식 기반으로 하여 정규화 되어진 칼라 영상 120개의 얼굴 영상 자료를 사전에 조사함으로써 구성되어 이를 통하여 감성인식을 한다. 그림 8은 인간의 감성 상태 중에서 놀람의 인식을 위해 사용된 12개의 특징벡터들의 소속함수를 기반으로 사전에 실험을 통해 정해진 규칙들을 나타낸다.

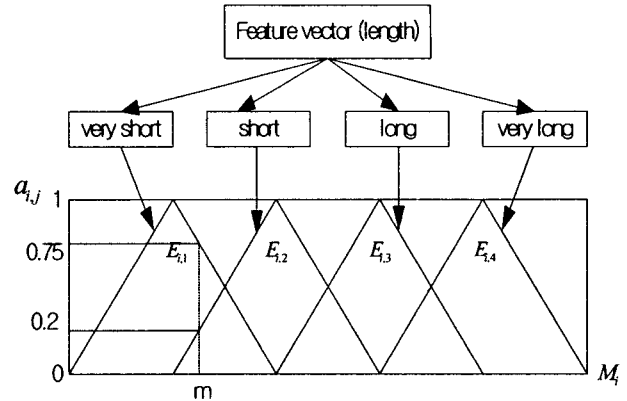


그림 7. 소속 함수

If

- $[a_{1,2}\{M_i(i=1, \text{기준점과 왼쪽눈 왼쪽 끝점에서 특징벡터의 차})\} < 0.7 \mid E_{1,2}]$
- $[a_{3,3}\{M_i(i=3, \text{기준점과 오른쪽눈 왼쪽 끝점에서 특징벡터의 차})\} < 0.1 \mid E_{3,3}]$
- $[a_{5,2}\{M_i(i=5, \text{기준점과 왼쪽눈썹 왼쪽 끝점에서 특징벡터의 차})\} < 0.8 \mid E_{5,2}]$
- $[a_{7,1}\{M_i(i=7, \text{기준점과 오른쪽눈썹 왼쪽 끝점에서 특징벡터의 차})\} > 0.4 \mid E_{7,1}]$
- $[a_{9,3}\{M_i(i=9, \text{기준점과 입술 왼쪽 끝점에서 특징벡터의 차})\} > 0.1 \mid E_{9,3}]$
- $[a_{10,4}\{M_i(i=10, \text{기준점과 입술 오른쪽 끝점에서 특징벡터의 차})\} < 0.2 \mid E_{10,4}]$
- $[a_{11,4}\{M_i(i=11, \text{기준점과 입 오른쪽 끝점에서 특징벡터의 차})\} > 0.5 \mid E_{11,4}]$
- $[a_{12,3}\{M_i(i=12, \text{기준점과 입 아래쪽 끝점에서 특징벡터의 차})\} < 0.2 \mid E_{12,3}]$
- $[a_{12,4}\{M_i(i=12, \text{기준점과 입 왼쪽 끝점에서 특징벡터의 차})\} > 0.3 \mid E_{12,4}]$

then 놀람으로 인식.

그림 8. 놀람 인식을 위한 규칙

3. 실험 및 결과고찰

본 논문은 얼굴 검출 및 감성을 인식의 알고리즘 개발을 목표로 하였으며, 본 실험에서의 제약 조건은 얼굴이 하나만 존재하여야 하고, 피부색의 색차를 이용한 컬러 데이터 변환을 전제로 하기 때문에

피부색에 가까운 의상이나 장애물을 피하였다. 제약조건을 포함하지 않는 범위 내에서 컬러 영상을 사용하였다. 제안한 방법의 실험을 위해 3명의 사람을 대상으로 4가지의 감정(기쁨, 놀람, 슬픔, 분노)에 대한 총 360개의 데이터 중에 120개를 사전 지식으로 사용하고, 실험으로는 240개의 컬러 영상을 사용하였다. 표 2를 보면 알 수 있듯이 감정 표현이 유사한 기쁨과 슬픔 감정 상태에서 서로 오인식되는 경우가 있었으며, 감정 표현이 뚜렷한 놀람이나 분노 보다 인식률이 현저히 떨어지는 것을 볼 수 있었다. 또한 슬픔 감정 상태의 경우에는 특징 점을 검출할 때 눈이 극히 작아지는 경우나 눈을 감을 경우 특징 점을 원활히 추출하지 못해 오인식의 원인이 되었다. 이러한 경우를 제외하고는 나머지 다른 감정 상태의 인식률에 있어서 제안된 시스템은 탁월한 성능을 보였으며 제안된 응용가능성을 충분히 보여주었다. 마지막으로 많은 종류의 감정과 좀더 높은 인식률을 구현하기 위해서는 잘 정의된 규칙들의 수가 추가되어지는 것이 필요하다. 표 2는 본 논문에서 제안한 감정 인식 시스템의 인식률을 나타낸 것이다. 본 논문에서 제안한 감정인식 시스템의 실험결과를 통해 그 응용 가능성을 확인하였다.

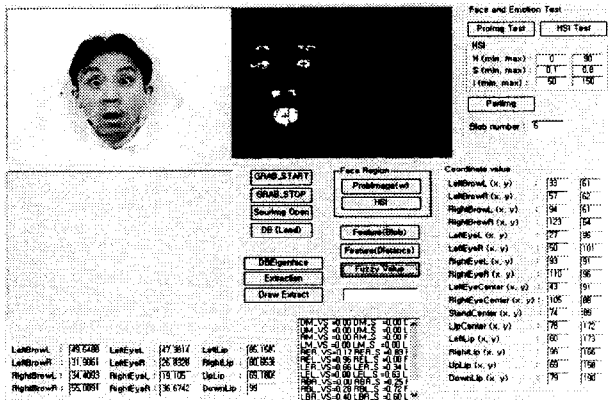


그림 9. 감정 인식 실험 시스템

표 2. 감성에 대한 인식률

	사람 1	사람 2	사람 3	총 계
기쁨	70% (14/20)	75% (15/20)	75% (15/20)	73.3%
놀람	90% (18/20)	95% (19/20)	90% (18/20)	91.7%
슬픔	75% (15/20)	75% (15/20)	70% (14/20)	73.3%
분노	80% (16/20)	85% (17/20)	75% (15/20)	80%
총 계	78.8%	82.5%	77.5%	80%

4. 결 론

본 논문에서는, 칼라 영상으로 인간의 감정 상태

를 인식하는 방법을 제안하였다. 획득한 영상으로부터 피부색을 이용하여 얼굴 영상만을 추출하기 위해 HSI 컬러 모델로 변환하여 얼굴영역을 검출하고 추출된 얼굴 영상으로부터 얼굴의 특징 요소(눈썹, 눈, 코, 입)를 찾아내어 인간 감정 상태의 특징 벡터 값을 조사하기 위한 특징 벡터 산출 방법을 제시하였으며, 특징 벡터들 간의 차이 값을 퍼지 규칙을 기반으로 한 소속함수를 이용한 퍼지화를 통해 감정상태(기쁨, 놀람, 슬픔, 분노)를 인식하는 방법을 제안하였다. 마지막으로, 본 연구실에 구축된 비전 시스템을 이용하여 직접 실험을 통해 제안한 방법의 응용 가능성을 확인하였다.

5. 참고 문헌

- [1] S. Morisima 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] L.C. De Silva, T. Miyasato, and R. Nakatsu, "Facial Emotion Recognition Using Multimodal Information", in Proc. IEEE Int. Conf. on Information, pp. 397-401, 1997.
- [3] M. Shigeo and H. Hiroshi, "Emotion Space for Analysis and Synthesis of Facial Expression", IEEE Int. Workshop on Robot and Human Communication, 1993.
- [4] R. Brunelli and T. Poggio, "Face Recognition: Features vs Templates", IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 15, 1993.
- [5] M. Turk and A. Pentland, "Face recognition Using Eigenface." Proc. of the IEEE conference on Computer Vision, pp. 586-591, 1991
- [6] M. Bichsel and A. Pentland, "Human Face Recognition and the Face Image Set's Topology", CVGIP: Image Understanding, Vol. 59, pp. 254-261, 1994.
- [7] A. Pentland, B. Moghaddam and T. Starner, "View-Based and Modular Eigenspaces for Face Recognition", Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, 1994.
- [8] 오재홍, "형판 벡터와 신경망을 이용한 감정인식", 한국퍼지 및 지능시스템학회, 12권 제 2호, pp. 325-328, 2002.
- [9] 이상운, "Eigenface를 이용한 인간의 얼굴인식과 감정인식", 군산대학교 석사학위논문, 2003.
- [10] Hideyuki EBINE and Osamu NAKAMURA, "The Recognition of Facial Expressions Based on Production System", Proceedings of the IEEE Canadian Conference, on Electrical and Computer Engineering Shaw Conference Center, pp.675-682, 1999.