

인간의 움직임 추출을 이용한 감정적인 행동 인식 시스템 개발

Emotional Human Body Recognition by Using Extraction of Human Body from Image

송민국* · 주영훈** 박진배*

Min Kook Song, Young Hoon Joo and Jin Bae Park

Abstract - Expressive face and human body gestures are among the main non-verbal communication channels in human-human interaction. Understanding human emotions through body gesture is one of the necessary skills both for humans and also for the computers to interact with their human counterparts. Gesture analysis is consisted of several processes such as detecting of hand, extracting feature, and recognizing emotions. Skin color information for tracking hand gesture is obtained from face detection region. We have revealed relationships between particular body movements and specific emotions by using HMM(Hidden Markov Model) classifier. Performance evaluation of emotional human body recognition has experimented.

Key Words :Emotion recognition, Hidden Markov Model(HMM), Pattern recognition

1. 서론

인간의 감정을 인식하는 기술은 활용가치가 높음 기술임에도 불구하고 기술개발의 어려움과 감정 표현의 난해함으로 인해 쉽게 해결되지 않고 있다. 인간의 감정을 인식하기는 방법으로 보통 영상과 음성을 이용한다.

영상을 이용한 감정인식의 방법은 이 중 특히 인간의 감정이 가장 많이 표현되는 인간의 움직임을 이용한 감정 인식 기법에 대한 연구가 활발히 진행 중이다 [3-6]. 인간의 움직임을 이용한 감정 인식은 그 인식 기법에 따라 기하학적 인식 기법, 고유공간을 이용한 인식 기법, 기타 변환을 이용한 인식 기법으로 나뉜다 [2], [7]. 이중 보다 일반적인 인간의 움직임에 대한 감정 인식을 수행하기 위해 인간의 움직임을 분석하여 인식을 수행하는 기법이 연구되고 있다. 하지만 인간의 움직임을 이용한 감정 인식은 이를 위해 다양한 선행 기술이 수정 및 개발되어야 하는 어려움이 있다. 이를 해결하기 위해 인간의 움직임 영상 분석에 대한 선행 기술들이 연구되었다 [1], [2].

본 논문에서는 기 개발된 선행 기술을 바탕으로 감정 인식 시스템을 위한 인간의 움직임 추출 기법과 은닉 마르코프 모델(HMM)을 통해 동정된 분류기를 통한 감정 인식 기법으로 제안한다. 인간의 움직임 추출은 선행 연구된 기법들로 얻어진 인간의 움직임 구성 요소에 정보를 바탕으로 감정 분류의 성능을 높이기 위한 손의 움직임을 파악하는 방법이다. 또한, 이렇게 생성된 손의 움직임 정보에 은닉 마르코프 모델을 통

해 동정된 분류기 기반 감정 인식 기법이 제안되었다. 분류기는 선행 연구된 은닉 마르코프 모델을 이용하여 동정된 분류기를 사용한다.

본 논문의 구성은 다음과 같다. 2장에서는 기 연구된 인간의 움직임 추출 기법들에 대해서 살펴본다. 3장에서는 인간의 움직임 벡터 추출 방법과 제안한 은닉 마르코프 모델을 이용한 분류기 기반 감정 인식 시스템에 대해 설명하고, 최종 개발된 감정 인식 시스템의 성능이 평가된다. 마지막으로, 4장에서 논문을 결론 맺는다.

2. 인간의 움직임 추출 방법

2.1 혼합 실루엣 추출 방법

시간에 따른 차이는 동적인 환경에 가장 적합하지만 모든 특징점 픽셀에서 추출되기 때문에 성능이 떨어진다. 배경에서의 추출은 정확한 인간의 실루엣을 추출하지만 빛과 다른 외부 요인의 영향을 많이 받게 된다. 최적의 흐름은 배경의 움직임을 가지는 이미지에서 추출하는데 사용되지만 그 계산법이 어렵고 모바일 시스템에서 실용화되기 어려운 단점이 있다. 따라서 이미지에서 움직임을 추출하는 최적의 접근법은 복합 실루엣 추출 기법이다 [6]. 복합 실루엣 추출 기법은 시간에 따른 차이에 기반한 방법이다.

먼저 연속적인 이미지들 (1)과 같이 정의한다.

$$I(x, y; t) = [I_r(x, y; t), I_g(x, y; t), I_b(x, y; t)]^T \quad (1)$$

$$= [p(t)]^T.$$

시간의 변화도는 두 이미지 사이의 시간의 차이이다. 시간과 공간의 변화도는 모서리의 정보를 가지며 다음의 (2)와 같이 정의한다.

저자 소개

- * 송민국 : 延世大學 電氣電子工學科 碩士課程
- ** 주영훈 : 群山大學 電子情報工學部 教授 · 工博
- *** 박진배 : 延世大學 電氣電子工學科 教授 · 工博

$$I_t = \frac{\partial I}{\partial t}$$

$$I_s = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right] \quad (2)$$

변화도 정보를 모서리의 핵심 값으로 구하고자 한다면, 변화도 정보는 자연수로 정의되어야 한다. 이러한 관점에서 다음의 (3)과 같이 간단히 변화도 정보를 정의한다.

$$\bar{I}_t = \| I_t \|$$

$$\bar{I}_s = \| I_s \| \quad (3)$$

모바일 시스템에서 시간의 변화도 \bar{I}_t 의 단점을 보완하기 위해 여러 방법이 도입 되었다. 카메라 자체의 문제로 인해서 우리는 시간의 변화도의 합 S_t 를 다음의 (4)와 같이 정의한다.

$$S_t = \int_{\forall p \in I} f(\bar{I}_t, \gamma_t) dp$$

$$f(\bar{I}_t, \gamma_t) = \begin{cases} 1, & \text{if } \bar{I}_t > \gamma_t \\ 0, & \text{else} \end{cases} \quad (4)$$

여기서 γ_t 는 색상 이미지를 이전 이미지로 옮기는 최소값이다. S_t 가 0이 될 때, 예전의 시간의 변화도는 현재의 시간의 변화도로 치환된다. 정확한 실루엣 정보를 얻기 위해서는 공간의 변화도와 시간의 변화도의 적절한 합이 필요하고 이는 다음 (5)와 같이 정의한다.

$$\eta \bar{I}_s(x, y) + (1 - \eta) \bar{I}_t(x, y) \quad (5)$$

여기서 η 는 불록합 합인 파라미터이다. 정확한 실루엣 정보를 얻기 위해서 η 값을 잘 정의하여야 하고, 이를 위해 본 논문에서는 움직임 영역 모델을 사용한다.

복합 실루엣은 다음의 (6)과 같이 계산된다.

$$\bar{I}_b(x, y) = \eta \bar{I}_s(x, y) + (1 - \eta) \bar{I}_t(x, y) \quad (6)$$

$$\eta = \begin{cases} \eta_R, & R(x, y) > \gamma_R \\ \eta_B, & R(x, y) < \gamma_R \end{cases}$$

여기서 η_R 는 움직임 영역의 불록합 파라미터이고, η_B 는 배경부분의 불록합 파라미터이다.

2.2 적응 뼈대 모델(Adaptive skeleton model)

움직임을 표현 하기위해 흔히 쓰이는 모델중의 하나가 뼈대 모델이다 [7]. 뼈대 모델은 가장 작은 정보만을 가지고 인간의 움직임을 표현한다.

뼈대모델은 머리의 위치와 크기를 정하는 것에서부터 출발한다. 많은 연구자들에 의해서 머리의 위치를 찾는 다양한

방법들이 연구되어왔다. 이 논문에서는 머리의 위치를 정하는 것이 목적이 아니기에 선행된 연구에서 개발한 방법을 사용하도록 하겠다.

이러한 특징점에서 스네이크 알고리즘과 유사하게 외부와 내부에 각각 해당하는 에너지 함수를 정의할 수 있다. 그리고 에너지 함수의 값을 최소화하게 나머지 특징점들의 위치를 구할 수 있다. 내부 에너지 함수는 특징점들간의 기하학적인 관계와 거리로부터 계산되고, 이때 거리는 다음의 (7)과 같이 정의한다.

$$E_i^{\text{int}} = |S_{i-} \| f_j - f_k \| | \quad (7)$$

여기서 f_j 와 f_k 는 S_i 와 연결된 추정된 특징점들이다.

외부 에너지 함수는 실루엣 이미지에서의 픽셀값들을 이용하여 정의한다. 예를 들어 머리 부분의 특징점에 대한 외부 에너지 함수는 다음의 (8)과 같다.

$$E_i^{\text{ext}} = - \sum_{x=f_{ix}-k}^{f_{ix}-k+m} \sum_{y=f_{iy}-k}^{f_{iy}-k+m} \bar{I}_b(x, y) \quad (8)$$

여기서 k 는 특징점과 실루엣 사이의 거리이고, $2m$ 은 탐색 영역의 넓이이다.

결국 에너지 함수 (9)는 불록합 파라미터 k 를 이용하여 내부에너지 함수와 외부에너지 함수의 합으로 구성된다.

$$E_i = k \bar{E}_i^{\text{ext} + (1-k) \bar{E}_i^{\text{int}}} \quad (9)$$

여기서 불록합 파라미터 k 는 인위적으로 정할 수 있다. k 의 값이 클수록 추정된 특징점들은 실루엣으로 이동하게 되지만 특징점의 모양이 어긋나게 된다.

2.3 색상 정보를 바탕으로 한 손의 움직임 추출 방법

색상 기반 손 탐색 알고리즘은 색상과 움직임 정보만을 사용하기 때문에 근처에 다른 사람의 손이 오는 경우 이를 구분하지 못하는 단점을 가진다. 따라서 본 논문에서는 이를 극복하기 위해 어깨선 추출을 이용한 보다 정확한 손 위치 파악 알고리즘을 개발하였다. 어깨선 추출 알고리즘은 목 바로 아래 어깨 점을 시작으로 원형의 탐색 공간을 각 원형으로 돌아서면서 탐색을 하는 방법으로 어깨선을 추출한다. 이때 탐색 영역안의 각 각들은 R1 과 R2 영역을 가지는데 이 영역에서 다음과 같은 에너지 함수 (11)에 의해서 에너지 값을 계산하게 된다. 최종적으로 에너지 값이 가장 큰 각으로 탐색을 해 나가는 방법을 취한다.

$$S = \left(\sum_{\forall x \in R_2} f(x) - \sum_{\forall x \in R_1} f(x) \right) / r + \sum_{\forall x \in L} f(x) \quad (11)$$

3. 모의실험

제한된 움직임 추출 알고리즘의 성능을 평가하기 위해 모의실험을 수행하였다. 모의실험은 크게 제한된 인간의 움직임 추출 알고리즘 평가와 제한된 은닉 마르코프 모델로 동정

된 감정 인식 분류기의 성능 평가로 나누어 수행되었다. 실험에 사용된 영상은 CCD 카메라로부터 얻은 320×240 크기의 24bit의 컬러 영상이 사용되었다.

먼저 인간의 움직임 추출을 위한 각 영역 데이터베이스는 소개된 적응 뼈대 모델과 색상 기반 손의 움직임 추출 알고리즘을 이용하였다. 그림 4는 얻어진 몇 개의 인간의 움직임과 손의 움직임 추출 결과이다.



그림 1. 인간의 움직임 추출
Fig 1. Extraction of human body

그림 2의 위쪽 부분은 제안된 손 영역 추출 기법으로 얻어진 손의 움직임을 분석하는 프로그램이다. 추출된 손과 팔의 크기를 정규화 시킨 후 양쪽 어깨의 특징점 위치를 (0, 0)으로 하고 손과 팔꿈치의 영상에서의 시간에 따른 위치 변화들을 나타낸 것이다. 실험하고자 하는 4가지의 감정에 대한 각각의 동작 데이터를 모으고 은닉 마르코프 모델을 이용하여 학습한다. 은닉 마르코프 모델로 학습한 모델 파라미터를 이용하여 동정된 분류기를 바탕으로 최종적으로 감정 인식 프로그램이 개발되었다. 개발된 감정 인식 프로그램의 결과는 그림 2의 아랫부분과 같다. 개발된 프로그램은 인간의 움직임 영상이 주어졌을 때, 제안한 인간의 움직임추출 기법들을 이용하여 인간의 움직임을 분석하고, 이를 이용해 손의 움직임 벡터를 생성한다. 그리고 은닉 마르코프 모델을 통해 학습된 파라미터로 동정된 분류기를 통해 최종 감정을 판별한다.

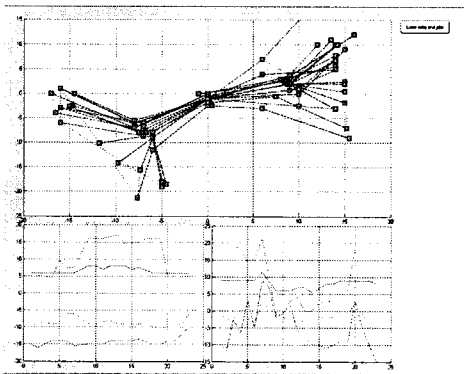


그림 2. 손의 움직임 추출
Fig 2. Extraction of human hand

표 1은 실험 결과를 나타낸다. 결과를 살펴보면 전체 인식이 71%로 매우 높음을 알 수가 있다. 이를 통해 인간의 움직임 추출과 추출된 움직임을 바탕으로 은닉 마르코프 모델을 이용해 동정된 분류기를 이용한 감정 인식이 매우 좋음을 확인 할 수 있다.

표 1 감정 인식 실험 결과

Table 1 Experiment result on emotion recognition

전체 감정 인식율	71%
각 감정 별 감정 인식율	
슬픔	73%
화남	68%
놀람	75%
역겨움	68%

4. 결론

지능형 시스템의 개발이 가속화됨에 따라 감정 인식 기술 개발의 필요성이 커지고 있다. 이에 본 논문에서는 인간의 움직임 분석을 통한 감정 인식 기법을 제안 하였다. 인간의 움직임 추출 알고리즘은 적응 뼈대 모델을 이용하여 보다 강한 환경에서 인간의 움직임을 추출하였고, 색상 기반 추출 알고리즘을 통해 보다 정확히 손의 위치를 추출하였다. 추출된 인간의 움직임에 은닉 마르코프 모델로 동정한 분류기를 이용하여 감정을 인식하였다. 최종적으로 모의실험을 통해 전체 알고리즘 수행과정을 살펴보았으며, 알고리즘의 우수함을 확인하였다.

참 고 문 헌

- [1] Haritaoglu, R. Cutler, D. Hawood and L. Davis, "Backpack: Detection of people carrying objects using silhouettes," Computer Vision and Image Understanding, pp, 385-397, No. 3, 2001
- [2] Haritaoglu, D. Harwood, and L. Davis, "A real time system for detection and tracking people" Journal of Image and Vision Computing, 1999.
- [3] Haritaoglu, D. Hawood and L. Davis, "Who? When? Where? What? A Real Time System for Detecting and Tracking People", Automatic Face and Gesture Recognition, pp. 222-227, 1998.
- [4] Blake, M. Isard, and D. Reynard, "Learning to track curves in motion of contours," Artificial Intelligence, pp.101-133, 1995
- [5] Anderson, P. Burt, and G. van der Wal., "Change detection and tracking using pyramid transformation techniques," In Proceedings of SPIE - Intelligent Robots and Computer Vision, Vol 579, pp. 72-78, 1985
- [6] 김문환, 박진배, 주영훈, 조영조, 지수영, 김해진, "지능형 로봇시스템에서 하이브리드 실루엣 추출 방법을 이용한 인간의 몸추출" 한국 퍼지 및 지능시스템학회 추계학술대회, pp. 257-260 2005
- [7] 김문환, 박진배, 주영훈, 조영조, 지수영, 김해진, "적응 뼈대 모델을 이용한 인간의 특징점 추출" 한국 퍼지 및 지능시스템학회 추계학술대회, pp. 257-260 2005