

# 인간의 제스처에 의한 감정 인식

## Emotion Recognition Based on Human Gesture

송민국\*, 박진배\*, 주영훈\*\*

Min Kook Song, Jin Bae Park, and Young Hoon Joo

\* 연세대학교 전기전자공학과

\*\* 군산대학교 전자정보공학부

### 요약

영상을 통한 감정 인식 기술은 사회의 여러 분야에서 필요성이 대두되고 있음에도 불구하고 인식 과정의 어려움으로 인해 풀리지 않는 문제로 남아 있다. 특히, 인간의 움직임에 이용한 감정 인식 기술은 많은 응용이 가능하기 때문에 개발의 필요성이 증대되고 있다. 영상을 통해 감정을 인식하는 시스템은 매우 다양한 기법들이 사용되는 복합적인 시스템이다. 따라서 이를 설계하기 위해서는 영상에서의 움직임 추출, 특징 벡터 추출 및 패턴 인식 등 다양한 기법의 연구가 필요하다. 본 논문에서는 이전에 연구된 움직임 추출 방법들을 바탕으로 한 새로운 감정 인식 시스템을 제안한다. 제안된 시스템은 은닉 마코프 모델을 통해 동정된 분류기를 이용하여 감정을 인식한다. 제안된 시스템의 성능을 평가하기 위해 평가데이터 베이스가 구축되었으며, 이를 통해 제안된 감정 인식 시스템의 성능을 확인하였다.

### Abstract

This paper is to present gesture analysis for human-robot interaction. Understanding human emotions through gesture is one of the necessary skills for the computers to interact intelligently with their human counterparts. Gesture analysis is consisted of several processes such as detecting of hand, extracting feature, and recognizing emotions. For efficient operation we used recognizing a gesture with HMM(Hidden Markov Model). We constructed a large gesture database, with which we verified our method. As a result, our method is successfully included and operated in a mobile system.

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. 감정 인식 기술은 영상을 기반으로 하는 방법과 음성을 이용하는 방법 그리고 두 가지를 모두 이용하는 방법으로 연구가 진행 중이다. 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], [8]. 이중 보다 일반적인 인간의 움직임에 대한 감정 인식을 수행하기 위해 인간의 움직임을 분석하여 인식을 수행하는 기법이 연구되고 있다. 하지만 인간의 움직임을 이용한 감정 인식은 다양한 선행 기술의 수정 및 개발되어야 하는 어려움이 있다. 인간의 움직임 영상 분석에 대한 선행 기술들은 여러 논문을 통해 연구되었다 [1-2].

본 논문에서는 기 개발된 선행 기술을 바탕으로 감정 인식시스템을 위한 인간의 움직임 추출 기법과 은닉 마르코프 모델(hidden markov model: HMM)을 통해 동정된 분류기를 통한 감정 인식 기법으로 제안한다. 인간의 움직임 추출 기법은 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 마코프 모델을 통해 동정된 분류기 기반 감정 인식 기법이 제안된다. 은닉 마코프 모델을 이용하여 설계한 감정 인식 분류기는 선형행렬부등식(linear matrix inequality) 기법을 통해 최적화된다.

본 논문의 구성은 다음과 같다. 2장에서는 기 연구된 인간의 움직임 추출 기법들에 대해서 살펴본다. 3장에서는 은닉 마르코프 모델을 설명하고, 4장에서는 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, 5장에서 논문을 결론 맺는다.

## 2. 인간의 움직임 추출 방법

### 2.1 혼합 실루엣 추출 방법

윤곽선의 특징점을 추출하는 것은 복합 영상에서 인간의 실루엣을 추출하는 것을 이용한다. 영상에서 인간의 실루엣

접수일자 : 2006년 9월 18일

완료일자 : 2007년 2월 6일

감사의 글 : 본 논문은 2005년도 정부재원(교육인적자원부 학술연구조성사업비)으로 한국학술진흥재단의 지원을 받아 수행된 연구임(KRF-2005-042-D00268)

을 추출하는 방법에는 세 가지 방법이 연구되어 왔다: 시간에 따른 차이, 배경에서의 변화, 그리고 최적의 흐름이다 [8].

시간에 따른 차이에 의한 방법은 동적인 환경에 가장 적합하지만 모든 특징점 픽셀에서 추출되기 때문에 성능이 떨어지는 단점이 있다. 배경에서의 추출 방법은 인간의 실루엣을 정확히 추출한다는 장점이 있는 반면, 빛과 다른 외부 요인의 영향을 많이 받는 단점이 있다고 알려져 있다. 최적 흐름에 의한 방법은 배경의 움직임을 갖는 이미지에서 실루엣을 추출하는데 사용되지만 그 계산방법이 어렵고 모바일 시스템에서 실용화되기 어려운 단점이 있다. 따라서, 이미지에서 움직임을 추출하는 최적의 접근 방법은 복합 실루엣 추출 기법이라고 알려져 있다[10]. 복합 실루엣 추출 기법은 시간에 따른 차이에 기반한 방법이다. 본 논문에서는 시간에 따른 차이에 의한 방법의 단점을 해결하기 위해서 다음과 같은 움직임 영역 모델과 누락되는 실루엣 정보 보상법이 제안된다.

먼저 연속적인 이미지를 (1)과 같이 정의한다.

$$I(x,y;t) = [I_x(x,y;t), I_y(x,y;t), I_b(x,y;t)]^T = [p(t)]^T \quad (1)$$

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

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

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

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  $\gamma_t$ 는 색상 이미지를 이진 이미지로 옮기는 최소값이다.  $S_t$ 가 0이 될 때, 예전의 시간의 변화도는 현재의 시간의 변화도로 치환된다.

실루엣 정보의 부족함과 정보의 왜곡을 보상하기 위해서 움직임 영역 모델을 이용한다. 지적된 문제점을 해결하기 위해서 공간의 변화도를 이용하고 오래된 시간의 변화도 정보의 연속적인 합을 이용한다. 결국 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  $\eta$ 는 볼록 합의 파라미터이다. 정확한 실루엣 정보

를 얻기 위해서  $\eta$ 값을 잘 정의하여야 하고, 이를 위해 본 논문에서는 움직임 영역 모델을 사용한다.

움직임 영역 모델은 초기 모델은 0에서 출발한다. 그리고 다음과 같은 규칙으로 갱신해 나간다.

$$R(x,y;t+1) = \begin{cases} R(x,y;t) + \gamma_i, & \bar{I}_t(x,y;t+1) > \gamma_t \\ R(x,y;t) + \gamma_d, & \bar{I}_t(x,y;t+1) < \gamma_t \end{cases}$$

$$\gamma_i = \begin{cases} \bar{\gamma}_i, & \bar{I}_s(x,y;t+1) > \gamma_s \\ \gamma_i, & \bar{I}_s(x,y;t+1) < \gamma_s \end{cases}$$

$$\gamma_d = \begin{cases} \bar{\gamma}_d, & \bar{I}_s(x,y;t+1) > \gamma_s \\ \gamma_d, & \bar{I}_s(x,y;t+1) < \gamma_s \end{cases}$$

$\gamma_s$ 는 공간의 이미지의 이진화 파라미터이다. 따라서 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  $\eta_R$ 는 움직임 영역의 볼록합 파라미터이고,  $\eta_B$ 는 배경부분의 볼록합 파라미터이다. 그림 1은 혼합 실루엣을 추출하는 전체 과정을 나타낸 순서도이다.

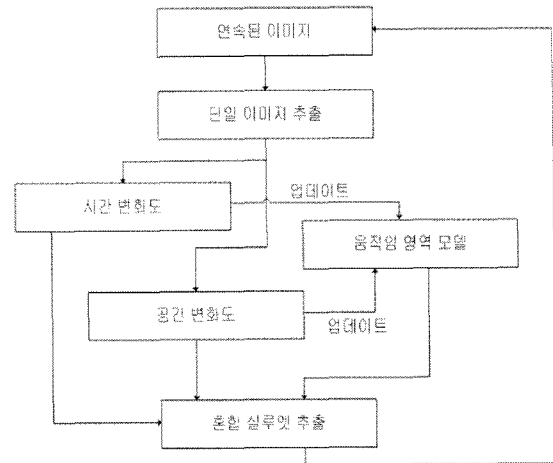


그림 1. 혼합 실루엣 추출 순서도  
Fig 1. Hybrid silhouette extraction

## 2.2 적응 뼈대 모델 (Adaptive skeleton model)

움직임을 표현 하기위해 흔히 쓰이는 모델중의 하나가 뼈대 모델이다. 뼈대 모델은 가장 작은 정보만을 가지고 인간의 움직임을 표현한다. 그림 2는 본 논문에서 사용할 뼈대 모델의 15개의 특징점 위치를 나타낸다. 점  $F_1$ 에서부터  $F_{15}$ 사이의 각각의 거리를  $S_1, \dots, S_{13}$ 으로 표현한다.

뼈대 모델은 머리의 위치와 크기를 정하는 것에서부터 출발한다. 많은 연구자들에 의해서 머리의 위치를 찾는 다양한 방법들이 연구되어왔다. 본 논문에서는 머리의 위치를 정하는 방법은 선행된 연구에서 개발한 방법을 사용한다 [9].

얼굴의 크기를 이용하여 특징점 간의 거리를 정규화 시키면 우리는 인간의 몸통 부분에 해당하는 특징점 ( $F_1, F_2, F_3, F_4, F_9$ , and  $F_{10}$ )의 대략적인 위치를 추정할 수 있다. 이러한 특징점에서 스네이크 알고리즘과 유사하게 외부와 내부에 각각 해당하는 에너지 함수를 정의할 수 있다. 그리고 고 에너지 함수의 값을 최소화하게 나머지 특징점들의 위치

를 구할 수 있다. 내부 에너지 함수는 특징점들간의 기하학적인 관계와 거리로부터 계산되고, 이때 거리는 다음의 (7)과 같이 정의한다.

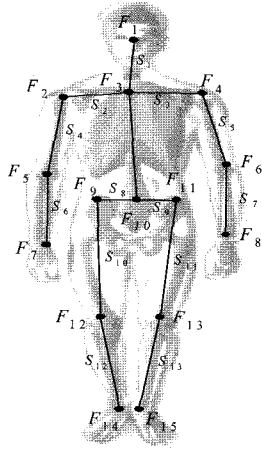


그림 2. 뼈대 모델  
Fig 2. Skeleton model

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

여기서  $f_j$ 와  $f_k$ 는  $S_i$ 와 연결된 추정된 특징점들이다. 또한 몇 개의 특징점들 사이에서는 기하학적인 관계가 정의 되어야 한다. 이러한 기하학적인 관계는 다음과 같이 표현한다.

$$g(1, f_i, f_j) = \begin{cases} 1, & f_{ix} > f_{jx} \\ 0, & f_{ix} > f_{jx} \end{cases}$$

$$g(0, f_i, f_j) = \begin{cases} 1, & f_{iy} > f_{jy} \\ 0, & f_{iy} > f_{jy} \end{cases}$$

여기서 추정된 특징점  $f_i$ 는  $f_i = [f_{ix}, f_{iy}]$ 로 정의한다.

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

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

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

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

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

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

이러한 에너지함수의 값을 작게 하기 위해서 다음과 같은 순서로 진행된다. 먼저 특징점 근처의 모든 픽셀에서의 에너지함수 값을 계산한다. 그리고 가장 작은 에너지함수 값을 가지는 픽셀을 새로운 특징점으로 정의한다. 만약 모든 픽셀에서 같은 에너지함수를 가지게 되면, 몸의 중심에서 가장 멀리 떨어진 픽셀을 새로운 특징점으로 정의한다.

### 2.3 색상 정보를 바탕으로 한 손의 움직임 추출 방법

실루엣 정보만으로 손의 특징점을 찾는 것은 충분하지 않다. 손과 팔꿈치 부분의 경우에는 움직임 동작에서 많은 변화가 나타남으로 이들의 정확한 위치를 찾는 것은 어려운 문제이다. 손의 위치를 찾기 위해서 추가적으로 색상 정보를 이용한다. 손의 위치를 찾은 이후에는 기하학적인 관계를 이용하여 팔꿈치의 위치를 찾는다. 피어(Peer)의 색상 모델은 동영상에서 색상을 찾는데 사용된다. 피어의 모델은 사람의 피부 색상과 비슷한 쓸모없는 부분도 아울러 찾기 때문에 이를 극복하기 위하여 시간의 변화도를 사용한다. 피부 색상을 가지는 이미지를  $I_c$ 라고 정의하자. 그러면 복합 피부 영역은 다음의 (10)과 같이 계산된다.

$$\bar{I}_c = I_c \bar{I} \quad (10)$$

손의 위치를 찾기 위해서는 얼굴 근처에 나타나는 피부색상 지역을 제거한다. 얼굴 영역은 적응 뼈대 모델로부터 구해진다. 얼굴 부근의 피부색상 지역을 제거하고 나서 손의 위치는 평균값 이동 탐색 알고리즘에 의해 구한다. 평균값 이동 탐색 알고리즘은 최고값을 가지는 히스토그램 값을 최소화 시키는 것이다. 자세한 알고리즘의 순서는 다음과 같다.

1 단계 :  $x$ 축과  $y$ 축의 히스토그램은 다음과 같다.

$$H_x(y) = \sum_{\forall x} \bar{I}_c(x, y), \quad H_y(x) = \sum_{\forall y} \bar{I}_c(x, y)$$

2 단계 : 평균값 이동 알고리즘은 히스토그램  $H_x(y)$ 에 적용하고  $x$ 축의 피부 영역을 구한다. 탐색 영역은 뼈대모델로 정해진다. 뼈대모델을 분석하면, 탐색영역이 추정된다.

3 단계 : 평균값 이동 알고리즘은 히스토그램  $H_y(x)$ 에 적용하고  $y$ 축의 피부 영역을 구한다. 탐색 영역은 뼈대모델과  $x$ 축의 피부 영역으로 정해진다.

4 단계 : 손의 중심부는 추출된 피부 영역으로부터 계산된다.

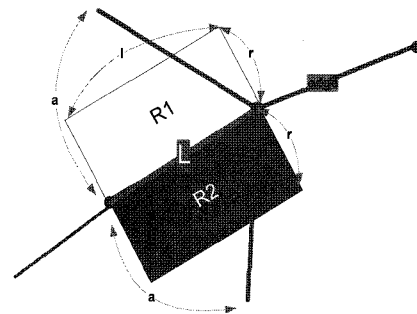
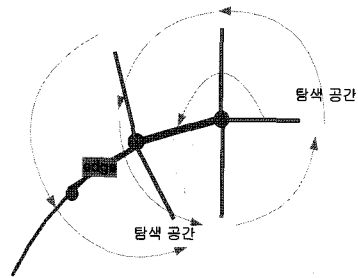


그림 3 어깨선 추출 과정  
Fig 3. Extraction of human shoulder

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 그림 3과 같이 원형의 탐색 공간을 각 원형으로 돌아서면서 탐색을 하는 방법으로 어깨선을 추출한다. 이때 탐색 영역안의 각 각들은 그림 4에서와 같은 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.

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

### 3. 은닉 마르코프 모델 (HMM)

은닉 마르코프 모델은 일반적인 마르코프 모델과 다르게 상태 변수를 직접적으로 관측할 수 없고, 단계별로 나타나는 시퀀스를 이용하여 추정할 뿐이다. 어떤 시간  $t$ 에서의 시스템의 확률 행렬은 시간  $t-1$ 에서의 확률행렬과 시간  $t$ 에서의 관측열에 의해서 결정된다. 상태 변수의 전체 수를  $N$ 으로 가정하고, 상태변수와 관측열의 시간  $t$ 에서의 초기값을 각각  $q_t, o_t$ 이라고 가정한다. 은닉 마르코프 모델에서는  $A, B, \pi$ 의 세 가지 파라미터 값을 동정해야 한다. 여기서  $A$ 는 상태 천이 확률 행렬을 의미하고,  $B$ 는 관측 심벌 확률 분포이며, 마지막으로  $\pi$ 는 초기 상태 분포를 의미한다. 만약 확률 분포  $\lambda = (A, B, \pi)$ 가 주어져 있다면 이 모델로부터 관측열  $O = \{O_1, O_2, \dots\}$ 은 다음과 같은 과정으로 생성시킬 수 있다.

1. 초기 상태 분포  $\pi$ 에 따라 초기 상태  $S_1$ 을 선택한다.
2.  $t=1$ 로 설정한다.
3. 관측 분포  $b_j(k)$ 에 따르는 관측열  $O_t$ 를 생성한다.
4. 그 상태에서 상태 천이확률 ( $A$ )에 따라 새로운 상태  $S_{t+1}$ 로 옮긴다.
5.  $t=t+1$ 로 설정하고  $t \geq T$ 까지 3-5의 과정을 반복한다.

관측열  $O = \{O_1, O_2, \dots\}$ 과 모델  $\lambda = (A, B, \pi)$ 이 주어졌을 때 우도 ( $O|\lambda$ )를 효과적으로 해결하기 위해서 전향 (Forward)과 후향(Backward) 알고리즘을 도입하여야 한다. 먼저 아래와 같이 전향 변수  $\alpha_t(i)$ 를 정의한다. 이 전향 변수는 모델  $\lambda$ 가 주어졌 있을 때, 시간  $t$ 에서의 상태  $S_t$ 와 시간  $t$ 까지의 관측열의 확률을 나타낸다. 전향 변수의 계산은 다음의 유도식 (12)로 효과적으로 계산되어진다.

$$\alpha_1(i) = \pi_i b_i(o_1), \quad 1 \leq i \leq N \quad (12)$$

$$\alpha_{t+1}(i) = b_i(o_{t+1}) \sum_{j=1}^N [\alpha_t(j) a_{ij}], \quad 1 \leq t \leq T$$

다음으로 변수  $\beta_t(i)$ 를 구하면 식 (13)과 같다.

$$\beta_T(i) = 1, \quad 1 \leq i \leq N \quad (13)$$

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \beta_{t+1}(j), \quad t = T-1, T-2, \dots, 1$$

관측열  $O = \{O_1, O_2, \dots\}$ 과 모델  $\lambda = (A, B, \pi)$ 이 주어졌 있을 때 관측열을 가장 잘 설명하는 최적 상태 열  $q = \{q_1, q_2, \dots\}$ 을 찾아야 한다. 최적 상태 열을 탐색하기 위해서는 단일한 최적의 상태 열 경로를 찾는 방법으로 본 논문에서는 비터비(Viterbi) 알고리즘을 사용한다. 비터비 알고리즘은 부분 최적인 각 상태를 통하여 최상의 단일한 상태 열을 찾는 알고리즘이다. 최적의 상태 열  $q = \{q_1, q_2, \dots\}$ 이 의미를 가지기 위해서 우리는 모델 파라미터인  $A, B, \pi$ 를 잘 정의해야 한다.

모델 파라미터의 추정을 위해서 본 논문에서는 바움-웰치(Baum-Welch) 재추정 알고리즘을 사용한다. 바움-웰치 알고리즘은 우도  $P(O|\lambda)$ 을 계산하여 관측치를 가장 잘 설명하는 모델  $\lambda$ 를 찾게 된다. 이때 모델 파라미터와 확률은 앞에서 설명한 전향변수와 후향 변수를 이용하여 다음과 같이 나타낸다. 따라서, 은닉 마르코프 모델을 이용하여 동정할 때는 정보를 추출할 때에는 숨어있는 상태 열을 잘 표현할 수 있는 초기 모델의 추정과 이를 바탕으로 상태 열을 잘 나타낼 수 있는 관측열을 잘 설정해야 한다

$$\pi = \gamma_1(i), \quad a_{ij} = \frac{\sum_{t=1}^{T-1} \zeta_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)}$$

$$b_i(k) = \frac{\sum_{t=1}^{T-1} \sum_{s.t. O_t=k} \gamma_t(i)}{\sum_{t=1}^{T-1} \gamma_t(i)}, \quad \gamma_t(i) = \frac{\alpha_t(i) \beta_t(i)}{\sum_{i=1}^N \alpha_t(i) \beta_t(i)}$$

$$\zeta_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(i)}{\sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(i)}$$

### 4. 모의 실험

제안된 움직임 추출 알고리즘의 성능을 평가하기 위해 모의 실험을 수행하였다. 모의 실험은 크게 제안된 인간의 움직임 추출 알고리즘 평가와 제안된 은닉 마르코프 모델로 동정된 감정 인식 분류기의 성능 평가로 나누어 수행되었다. 실험에 사용된 영상은 CCD 카메라로부터 얻은 320×240 크기의 24bit의 컬러 영상이 사용되었다. 감정 표현의 일반성을 위해 30명의 감정 표현 움직임을 촬영하고 촬영된 영상에서부터 움직임을 추출한다. 움직임 추출을 위해 2장에서 소개된 적응 뼈대 모델과 색상 기반 손의 움직임 추출 알고리즘을 이용한다. 제안된 시스템의 성능을 평가하기 위해 남성 20명과 여성 10명의 감정 표현 움직임 동영상으로 구성된 데이터베이스를 구축하였다. 평가에 사용된 감정은 총 4가지이며 인간의 움직임 추출과 추출된 움직임을 바탕으로 은닉 마르코프 모델을 이용해 동정된 분류기를 이용하여 성능을 평가한다

먼저 인간의 움직임 추출을 위한 각 영역 데이터베이스는 소개된 적응 뼈대 모델과 색상 기반 손의 움직임 추출 알고

리즘을 이용하였다. 그림 4는 언어진 몇 개의 인간의 움직임과 손의 움직임 추출 결과이다. 다양한 환경과 다양한 인물 그리고 다른 동작들 속에서도 본 논문이 제안한 알고리즘은 정확히 인간의 움직임 영역을 추출해내는 것을 확인하였다.

그림 5의 위쪽 부분은 제안된 손 영역 추출 기법으로 얻어진 손의 움직임을 분석하는 프로그램의 결과이다. 추출된 손과 팔의 크기를 정규화 시킨 후 양쪽 어깨의 특징점 위치를 (0, 0)으로 하고 손과 팔꿈치의 영상에서의 시간에 따른 위치 변화를 나타낸 것이다. 보다 정밀한 분석을 위해서 오른팔과 왼팔의 각각의 손의 위치와 팔꿈치의 위치변화를 같이 살펴본다. 실험하고자 하는 4가지의 감정에 대한 각각의 동작 데이터를 모으고 이를 3장에서 소개한 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 5의 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.

표 1 인간의 움직임 추출 결과  
Table 1 Experiment result on emotion recognition

인간의 움직임 추출율	98.06%
손 영역 추출율	73.84%



그림 4. 인간의 움직임 추출  
Fig 4. Extraction of human body

표 2는 실험 결과를 나타낸다. 결과를 살펴보면 전체 인식이 74%로 매우 높음을 알 수가 있다. 이를 통해 인간의 움

직임 추출과 추출된 움직임을 바탕으로 은닉 마르코프 모델을 이용해 동정된 분류기를 이용한 감정 인식이 매우 좋을 수 있다.

표 2 감정 인식 실험 결과  
Table 2 Experiment result on emotion recognition

전체 감정 인식율	74%
각 감정 별 감정 인식율	
슬픔	72%
화남	77%
놀람	71%
억제움	76%

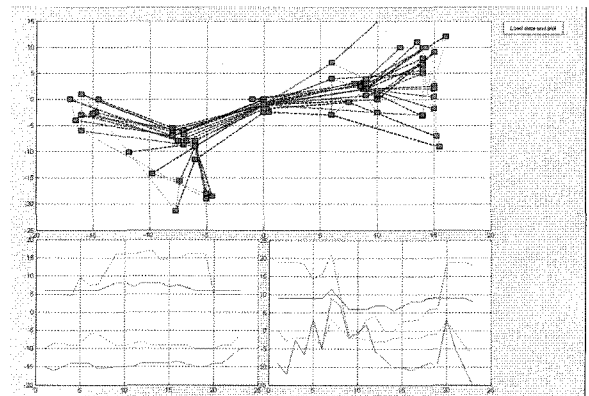


그림 5. 손의 움직임 추출 및 감정인식  
Fig 5. Extraction of human hand and emotion recognition

## 5. 결 론

감정 인식은 많은 분야에서 각광받고 있는 기술 중에 하나이다. 지능형 시스템의 개발이 가속화됨에 따라 감정 인식 기술 개발의 필요성이 커지고 있다. 이에 따라 인간의 움직임 추출을 통한 감정 인식 기술 개발이 중요해지고, 이에 본 논문에서는 인간의 움직임 분석을 통한 감정 인식 기법을 제안 하였다. 제안된 알고리즘은 인간의 움직임 추출 기법과 은닉 마르코프 모델을 통한 감정 분류 기법으로 구성 된다. 인간의 움직임 추출 알고리즘은 적응 뼈대 모델을 이용하여 보다 강인한 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, 알고리즘이 잘 동작함을 확인하였다.

## 참 고 문 헌

- [1] J. K. Aggarwal and Q. Cai., "Human Motion Analysis: A Review," Nonrigid and Articulated Motion Workshop, pp. 90-102, 1997
- [2] B. Fan, Z.-F. Wang, "Pose estimation of human

- body based on silhouette images," Int. Conf. on Information Acquisition Proceedings., pp. 296-300, June, 2004
- [3] H.R. Cutler, D. Hawood and L. Davis, "Backpack: Detection of people carrying objects using silhouettes," Seventh Int. Conf. on Computer Vision (ICCV'99), Vol 1, pp, 102-107, 1999
- [4] H. D. Harwood, and L. Davis, "A real time system for detection and tracking people" Lecture Notes in Computer Science, Vol. 1406, pp. 877, 1998
- [5] H. D. Hawood and L. Davis, "Who? When? Where? What? A Real Time System for Detecting and Tracking People", Lecture notes in computer science, Vol. 1407, pp. 877-892, 1998.
- [6] B. M. Isard, and D. Reynard, "Learning to track curves in motion of contours," Decision and Control, 1994., Pro. of the 33rd IEEE Conference, Vol 4, pp.3788-3793, 1994
- [7] A. P. Burt, and G. Wal., "Change detection and tracking using pyramid transformation techniques," In Pro. of SPIE - Intelligent Robots and Computer Vision, Vol 579, pp. 72-78, 1985
- [8] J. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," Int. J. of Computer Vision, Vol 12, No 1, pp.43-77, 1994.
- [9] 김문환, 박진배, 주영훈, 조영조, 지수영, 김혜진, "지능형 로봇시스템에서 하이브리드 실루엣 추출 방법을 이용한 인간의 몸추출" 한국 퍼지 및 지능 시스템학회 추계학술대회, Vol 15, No 2, pp. 257-260 2005
- [10] O. Masoud, N.P. Papanikolopoulos, "A method for human action recognition", Image and Vision Computing, vol. 21, No. 8, pp. 729-743, 2003

저 자 소 개

송민국(Min Kook Song)  
제 16권 6호(2006년 12월호) 참조

박진배(Jin Bae Park)  
제 16권 6호(2006년 12월호) 참조

주영훈(Young Hoon Joo)  
제 16권 6호(2006년 12월호) 참조