

# 휴먼-로봇 인터랙션을 위한 하이브리드 스켈레톤 특징점 추출

## Feature Extraction Based on Hybrid Skeleton for Human-Robot Interaction

주영훈\*, 소제윤

(Young Hoon Joo and Jea Yun So)

**Abstract :** Human motion analysis is researched as a new method for human-robot interaction (HRI) because it concerns with the key techniques of HRI such as motion tracking and pose recognition. To analysis human motion, extracting features of human body from sequential images plays an important role. After finding the silhouette of human body from the sequential images obtained by CCD color camera, the skeleton model is frequently used in order to represent the human motion. In this paper, using the silhouette of human body, we propose the feature extraction method based on hybrid skeleton for detecting human motion. Finally, we show the effectiveness and feasibility of the proposed method through some experiments.

**Keywords :** Human-robot interaction, intelligent robot, hybrid skeleton, gesture recognition

### I. 서론

오늘날 현대사회에서는 인간과 로봇과의 자연스러운 상호작용을 위하여 시각을 기반으로 한 사용자 의도 및 행위 인식에 대한 연구가 활발히 진행되고 있다. 인간 모션 추출은 시각을 기반으로 한 인식 분야에서 핵심 기술 분야로 연구 되어 왔으며 최근에는 로봇이 인간에게 자연스러운 서비스를 제공해 주거나 로봇의 동작을 제어하기 위해 연구되고 있는 분야이다. 이러한 기술로는 인간-컴퓨터 간 (Human-Computer Interaction, HCI), 인간-로봇 간(Human-Robot Interaction, HRI) 등의 상호작용 기술이 있으며 피드백이 발생하는 모든 요인을 고려한 인간 공학적이고 형태학적인 분석기술로 발전해 왔다. 그 중에서 인간 컴퓨터 상호작용 기술은 인간과 컴퓨터의 상호 작용을 위해 인간의 시각 기능을 컴퓨터에 이식하여 컴퓨터 비전에 기반 한 사용자 의도 및 행위를 인식하는 연구들이 진행되어 왔다. 입력 장비를 사용하지 않고 컴퓨터와 인터페이스 하기 위한 인간 컴퓨터 상호작용기술로서 얼굴인식 및 제스처 인식이 연구되었다[1,2].

그러나 본 논문의 기반인 인간-로봇 상호작용 기술은 로봇이 가지는 자율성, 상호작용의 양방향성, 상호작용 또는 제어 수준의 다양성 등에서 인간-컴퓨터 상호작용과는 근본적인 차이점을 가지고 있다. 인간-로봇 상호 작용기술은 다양한 의사소통 채널을 통해 인지적, 정서적 상호작용을 할 수 있도록 로봇 시스템 및 상호작용 환경을 디자인, 구현 및 평가하는 기술로서 지능형 로봇 시스템에서 핵심기술이라 말할 수 있다. 효과적인 인간-로봇 상호작용을 위해서는 인간 및 로봇 상호간의 편리성, 협력성, 친밀성을 구현하는 모듈 구조를 가진 시스템이 요구되기 때문에 로봇 동작 제어, 사용자에 대한 서비스 제공 및 가상현실 등에

응용되며 최근 몇 년 동안 휴머노이드 로봇의 동작 제어 및 지능형 로봇 시스템 형태의 가정용 로봇 등에서 활용되었다[3-7]. 인간은 일상생활에서 제스처(gesture), 표정과 같은 비 언어적인 수단을 이용하여 수많은 정보를 교환한다. 예를 들어 얼굴 표정, 손의 움직임, 시선 방향, 머리의 동작, 신체의 포즈 등이 대화중에 상대방에게 많은 정보를 전달한다. 이러한 비언어적 수단을 지능형 로봇과 인간의 대화에 사용한다면 보다 사용자 친화적인 인터페이스 실현이 가능하게 된다.

특히 근래에 들어 입는 컴퓨터(wearable computer), 인지 사용자 인터페이스(perceptual user interface), 유비쿼터스 컴퓨팅(ubiquitous computing), 감정 컴퓨팅(affective computing) 과 같이 제 4세대 정보기술의 중요성이 강조 되면서 인간의 행동에 대한 각종 인식은 컴퓨터 비전 연구자들의 많은 주목을 받고 있다. 컴퓨터를 사용하여 동작을 인식하는 것은 인체 각 부위가 시간의 흐름에 따라 어떤 모습으로 변화를 갖는지 알아내는 것을 의미한다. 인간 모션 추출을 위해서는 인간 실루엣을 추출하고 인간 관절 부위의 특징 요소를 추출해야한다.

지금까지 연구 되어온 방법들을 살펴보면 다음과 같다. 종래의 인간 몸의 특징 정보를 추출 하는 방법으로는 대부분 통계학적인 방법, 기하학적인 방법 등의 분석을 통한 방법과 다수의 카메라를 이용하여 인간 특징 정보를 추출하는 방법이 있으며 광학적 표식(marker)를 이용하여 인간 특징 정보를 추출 하는 방법이 있다[8,9].

위와 같이 인간의 모션 추출을 위해 여러 가지 방법들이 제안 되었으나 대부분의 극히 제한된 실험실 환경에서 이루어 졌거나 소수의 실험 대상자들에 대해서 인간 모션 데이터를 얻고 결과를 도출 하여 왔다. 또한 다양한 조명 환경 및 배경을 가지는 실세계 환경에서 인간 모션 추출은 매우 어렵다. 특히, 시·공간에 따른 인간 모션에 대한 다양성 및 모션 동작자에 제한을 두지 않는 상황에서 인간 모션을 추출 하기는 그리 쉬운 작업이 아니다. 따라서 다양

\* 책임저자(Corresponding Author)

논문접수 : 2007. 12. 27., 채택확정 : 2008. 1. 5.

주영훈, 소제윤 : 군산대학교 전자정보공학부

(yhjoo@kunsan.ac.kr/ifuloveme@kunsan.ac.kr)

한 환경 속에서의 카메라 영상 기반의 인간 모션 추출은 카메라로 인간의 영상을 획득하여 연속된 이미지로부터 움직임 영역의 정확한 특징 점을 찾는데 있다.

본 논문은 지능형 로봇 시스템의 단일 카메라로부터 얻어진 연속적인 영상 정보에서 단일 이미지를 획득하여 인간 실루엣을 추출하고 10명의 DB로부터 계산된 신체 비례 정보를 이용하여 구성된 가상 인간 스켈레톤 모델을 설계한다. 인간 실루엣 영상에 가상 인간 스켈레톤 모델의 초기 특징 점을 생성하고 LK 알고리즘을 이용하여 초기 특징 점을 추적한 뒤 몸통 안의 특징 점은 스네이크 알고리즘을 이용한 동정한다. 또한, 색상 기반 손 탐지 알고리즘을 이용하여 손 위치와 팔꿈치 위치를 계산함으로써 특징 점을 추출의 정확도를 높인다. 본 논문에서 구축된 시스템을 이용하여 직접 실험을 통해 제안된 방법의 응용 가능성을 제시하고 실험에 사용된 영상 정보를 바탕으로 인간 모션 추출의 정확도 여부를 판단한다.

## II. 가상 인간 스켈레톤 모델

인간의 모션을 추출하는데 있어 가장 중요한 것은 모션을 가장 잘 표현하는 정확한 특징점을 찾는 데 있다. 그리고 인간 모션을 추출하기 위해서는 인간의 신체 구조적 정보를 이용하는 것이 효율적이다. 따라서 본 논문에서는 10명 인간 신체 DB를 이용하여 인간의 신체 구조적 비례 정보를 얻는다. 추출된 신체 비례 정보를 이용하여 막대형 특징 점 추출 방식인 스켈레톤 모델 형태를 사용한 가상 인간 스켈레톤 모델을 구성한다.

그림 1은 신체 비례 정보로 구성된 가상 인간 스켈레톤 모델과 모델에 사용되는 특징 점들을 나타내고 있다.

본 논문에서는 지능형 로봇 시스템에서 인간-로봇간의 자연스러운 상호작용을 위한 정보로 인간 신체 하부는 영향이 없다고 판단되어 인간 신체 상부에만 초기 특징 점을 설계한다. 상기 가상 인간 스켈레톤 모델에서 사용된 특징 점과 뼈대의 구조로  $F_1$ 부터  $F_{11}$ 까지는 11개의 특징 점을 나타내며, 각 특징 점간의 거리는  $S_1, \dots, S_{12}$ 로 표현된다. 본 논문에서는 얼굴 위치와 가로 넓이를 기준으로 가상 인간 스켈레톤 모델 형태가 설계된다. 따라서 OpenCV에서 제공되는 얼굴 탐지 알고리즘을 이용하여 인간의 얼굴 위치와

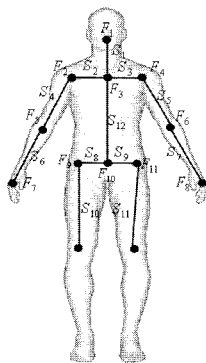


그림 1. 가상 인간 스켈레톤 모델.

Fig. 1. Skeleton model of virtual human.

얼굴 가로 길이 찾는다. 사용된 OpenCV(Open Source Computer Vision)의 얼굴 탐지 알고리즘은 인텔에서 만든 영상처리 라이브러리 공개 소스로서 많은 영상처리 분야에서 사용되고 있다[11]. 얼굴 탐지 알고리즘에서 얻어진 얼굴의 가로 넓이를 기준으로 가상 인간 신체 비율을 계산되며 계산된 신체구조비를 이용하여 가상 인간 스켈레톤 모델의 초기 특징 점을 생성한다.

가상 인간 스켈레톤 초기 특징 점  $F_i$ 는 (1)과 같이 정의된다.

$$F_i = [F_{ix}, 240 - F_{iy} + S_i] \quad (1)$$

여기서,  $S_i$ 는 10명의 DB로부터 신체 비례 정보를 이용하여 얻은 대략적인 중요 신체 관절 간의 거리, 즉 특징 점간의 거리이다. 특징 점간의 거리  $S_i$ 는 다음 (2)와 같이 계산된다.

$$S_1 = \frac{38}{27}h_w, S_2 = \frac{27}{27}h_w, \dots, S_{11} = \frac{50}{27}h_w, S_{12} = \frac{80}{27}h_w \quad (2)$$

여기서,  $h_w$ 는 OpenCV에서 제공한 얼굴 탐지 알고리즘을 이용하여 얻어진 얼굴 가로 넓이이다.

## III. 하이브리드 스켈레톤 특징점 추출

본 논문에서는 하이브리드 스켈레톤 특징 점 추출 기법을 제안한다. 가상 인간 스켈레톤 모델을 이용하여 초기 특징 점( $F_1, F_2, \dots, F_{10}, F_{11}$ )을 설계하고 LK 알고리즘을 이용하여 초기 특징 점을 추적한다. 추적에 사용된 LK 알고리즘은 1981년 Lucas-Kanade가 처음 제안한 기법으로 영상간의 움직임 변위를 계산하여 추적 하는 방식이다[12]. 연속적인 영상의 시간  $t$ 에 대해서 가상 인간 스켈레톤 모델의 신체 비례 정보에 의해 초기 특징 점 정보는 다음 (3)과 같다.

$$\begin{aligned} I(X;t) &= I(F_{ix}, F_{iy}; t + \Delta t) \\ &= I(F_{ix} + \Delta F_{ix}, F_{iy} + \Delta F_{iy}; t) \end{aligned} \quad (3)$$

상기 가상 인간 스켈레톤 초기 특징 점을 추적하기 위해 LK 알고리즘 수행한다. 시간  $t+1$ 일 때  $X$ 의 좌표의 특징 점 정보는  $I(X+d;t+1)$  정의된다. 여기서,  $X = [f_{ix}, f_{iy}]^T$ ,  $d_i = [\Delta f_{ix}, \Delta f_{iy}]^T$ 이다.

연속적인 영상에서 프레임간의  $X$  좌표의 움직임 벡터  $d$ 를

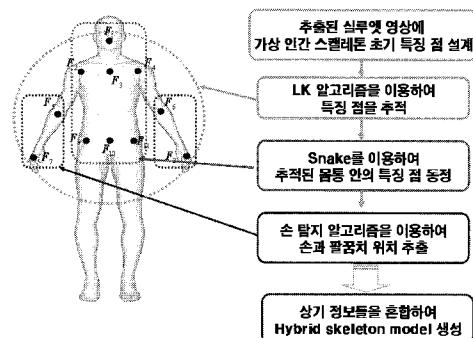


그림 2. 하이브리드 스켈레톤 특징점 추출 과정.

Fig. 2. Hybrid skeleton features extraction process.

찾기 위해  $\| \Delta d \| \leq \epsilon$  일 때까지 반복적으로 수행하며 특징 벡터들의 제곱 오차의 합 계산식은 (4)와 같다.

$$\epsilon = \sum_X [I(X+d;t) - I(X;t + \Delta t)]^2$$

$$d_i = d_i + \Delta d_i \tag{4}$$

여기서,  $\epsilon$  는 특징 벡터들의 제곱 오차 합 값,  $d_i$  는 움직임 벡터  $i = 1, \dots, 11$  이다.

LK 알고리즘을 이용하여 추적되는 몇몇의 특징 점들은 기하학적인 위치 관계를 고려해야 한다. 예를 들면 특징 점  $F_3$  와  $F_4$  는 반드시  $F_2$  의 오른쪽에 위치하여야 한다. 이러한 기하학적인 위치 관계는 함수  $g(l, f_i, f_j)$  로 표현되며, 함수  $g(l, f_i, f_j)$  는 (5)와 같이 정의된다.

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

$$g(0, f_i, f_j) = \begin{cases} 1 & f_{iy} > f_{jy} \\ 0 & f_{iy} < f_{jy} \end{cases} \tag{5}$$

최종적으로 특징 점에 대한 각 특징 벡터의 변위 값은 기하학적 관계를 바탕으로 (6)과 같이 정의 된다.

$$d_1 = |d_1 + \Delta d_1| + g(0, f_3, f_1) + g(1, f_2, f_1) + g(1, f_1, f_4)$$

$$d_2 = |d_2 + \Delta d_2| + g(1, f_2, f_3) + g(0, f_2, f_3)$$

$$d_3 = |d_3 + \Delta d_3| + g(1, f_2, f_3) + g(1, f_3, f_4) + g(0, f_1, f_3)$$

$$d_4 = |d_4 + \Delta d_4| + g(1, f_3, f_4) + g(0, f_1, f_4)$$

$$d_5 = |d_5 + \Delta d_5| + g(1, f_2, f_5) + g(1, f_5, f_9)$$

$$d_6 = |d_6 + \Delta d_6| + g(0, f_6, f_4) + g(0, f_6, f_{11})$$

$$d_7 = |d_7 + \Delta d_7|$$

$$d_8 = |d_8 + \Delta d_8|$$

$$d_9 = |d_9 + \Delta d_9| + g(1, f_9, f_{10})$$

$$d_{10} = |d_{10} + \Delta d_{10}| + g(1, f_9, f_{10}) + g(1, f_{10}, f_{11}) + g(0, f_3, f_{10})$$

$$d_{11} = |d_{11} + \Delta d_{11}| + g(1, f_{10}, f_{11}) \tag{6}$$

LK 알고리즘을 이용하여 추적된 특징 벡터에서 스네이크 알고리즘을 이용하여 신체구조비의 문제점을 극복하는 좀 더 강한 특징 벡터를 추출한다. 스네이크는 1988년 Kass 등이 제안한 알고리즘으로 영상잡음에 보다 강인한 면을 가지며, 주관적으로 요구할 수 있는 다양한 지식기반의 제한 조건들과 스네이크를 구성하는 각 점들 사이의 관계를 이용하여 경계선을 얻어 내는데 많이 이용되어 왔다.

본 논문에서는 몸통 안의  $(d_1, \dots, d_4, d_9, \dots, d_{11})$  특징 점 동정에 사용된 스네이크 알고리즘의 에너지 함수는 내부 에너지 함수와 외부 에너지 함수로 구성된다. 스네이크 알고리즘은 에너지 함수를 최소화 시키는 점을 찾는 방향으로 알고리즘이 수행된다.

기본적인 내부 에너지 함수 계산식은 (7)과 같다.

$$E_i^{Int} = |S_i - \| d_j - d_k \| | \tag{7}$$

여기서  $p_j$ 와  $p_k$ 는  $S_i$ 과 인접한 특징 점 들이다.

한 가지 예로,  $f_3$  의 내부함수는 (8)과 같이 정의 된다.

$$E_3^{Int} = |S_3 - \| d_3 - d_4 \| | \tag{8}$$

최종적으로 몸통 안의 특징 벡터  $(d_1, \dots, d_4, d_9, \dots, d_{11})$  에 대한 내부 에너지 함수는 가상 인간 스켈레톤 모델의 신체 구조비로 주어진 각 특징 점의 기하학적 관계식 (5)를 바탕으로 다음과 같이 정의된다.

$$E_1^{Int} = |S_1 - \| d_1 - d_3 \| | + g(0, d_3, d_1) + g(1, d_2, d_1) + g(1, d_1, d_4)$$

$$E_2^{Int} = |S_2 - \| d_2 - d_3 \| | + g(1, d_2, d_3) + g(0, d_1, d_2)$$

$$E_3^{Int} = |S_2 - \| d_2 - d_3 \| | + |S_3 - \| d_3 - d_4 \| | + |S_3 - \| d_3 - d_4 \| |$$

$$+ |S_{12} - \| d_3 - d_{10} \| | + g(1, f_2, f_3) + g(1, f_3, f_4) + g(0, f_1, f_3)$$

$$E_4^{Int} = |S_3 - \| d_3 - d_4 \| | + g(1, d_3, d_4) + g(0, d_1, d_4)$$

$$E_9^{Int} = |S_9 - \| d_9 - d_{10} \| | + g(1, d_9, d_{10})$$

$$E_{10}^{Int} = |S_9 - \| d_9 - d_{10} \| | + |S_{10} - \| d_{10} - d_{11} \| | +$$

$$|S_8 - \| d_3 - d_{10} \| | + g(1, f_{10}, f_{11}) + g(0, f_3, f_{10})$$

$$E_{11}^{Int} = |S_{10} - \| d_{10} - d_{11} \| | + g(1, d_{10}, d_{11}) \tag{9}$$

외부 에너지 함수는 실루엣 이미지의 픽셀 정보를 이용하여 정의된다. 기본적인 외부 에너지 함수는 (10)과 같이 정의된다.

$$E_i^{Ext} = - \sum_{x=d_x-k}^{d_x-k+m} \sum_{y=d_y-k}^{d_y-k+m} \bar{I}_b(x, y) \tag{10}$$

여기서,  $k$ 는 특징 점과 실루엣 사이의 거리를 나타내며  $m$ 은 탐색공간을 나타낸다.

인간 몸의 특징 점에 대해 (11)과 같이 외부 에너지 함수를 정의할 수 있다.

$$E_1^{Ext} = - \sum_{x=f_{1x}}^{f_{1x}+m} \sum_{y=f_{1y}}^{f_{1y}+m} \bar{I}_b(x, y)$$

$$E_2^{Ext} = - \sum_{x=f_{2x}-m}^{f_{2x}+m} \sum_{y=f_{2y}-k-m}^{f_{2y}-k+m} \bar{I}_b(x, y)$$

$$E_3^{Ext} = - \sum_{x=(f_{2x}+f_{3x})/2-m}^{(f_{2x}+f_{3x})/2+m} \sum_{y=f_{3y}-k-m}^{f_{3y}-k+m} \bar{I}_b(x, y)$$

$$- \sum_{x=(f_{3x}+f_{4x})/2-m}^{(f_{3x}+f_{4x})/2+m} \sum_{y=f_{3y}-k-m}^{f_{3y}-k+m} \bar{I}_b(x, y)$$

$$E_4^{Ext} = - \sum_{x=f_{4x}-m}^{f_{4x}+m} \sum_{y=f_{4y}-k-m}^{f_{4y}-k+m} \bar{I}_b(x, y)$$

$$E_9^{Ext} = - \sum_{x=f_{9x}-k-m}^{f_{9x}-k+m} \sum_{y=f_{9y}-m}^{f_{9y}+m} \bar{I}_b(x, y)$$

$$E_{10}^{Ext} = - \sum_{x=f_{10x}-S_3-k-m}^{f_{10x}-S_3-k+m} \sum_{y=f_{10y}-m}^{f_{10y}+m} \bar{I}_b(x, y)$$

$$- \sum_{x=f_{10x}+S_0-k+m}^{f_{10x}+S_0-k-m} \sum_{y=f_{10y}-m}^{f_{10y}+m} \bar{I}_b(x, y)$$

$$E_{11}^{Ext} = - \sum_{x=f_{11x}-k-m}^{f_{11x}-k+m} \sum_{y=f_{11y}-m}^{f_{11y}+m} \bar{I}_b(x, y) \tag{11}$$

최종적으로 에너지 함수  $E_i$ 는 내부와 외부 에너지 함수의

블록 합으로 나타내어진다.

$$E_i = K\bar{E}_i^{Ext} + (1-K)E_i^{Int} \quad (12)$$

여기서,  $K$ 는 블록 합 파라미터이며,  $\bar{E}_i^{Ext}$ 은 정규화 된 외부 에너지 함수 값으로 (13)과 같이 계산된다.

$$\bar{E}_i^{Ext} = \frac{E_i^{Ext}}{\sum E_K^{Ext}} \quad (13)$$

여기서,  $K$ 가 크면 특징 점들은 실루엣에 더 민감하게 이동할 것이고  $K$ 가 작으면 가상 인간 스켈레톤 모델의 신체 특징 비율에 민감하게 이동하게 된다.

그림 4는 초기 가상 인간 스켈레톤 특징 점에서 LK와 스네이크 알고리즘으로 추적 및 동정된 특징 벡터 예이다. 상기 그림에서와 같이 얼굴을 기준으로 총 11개의 특징 벡터가 추적 및 동정됨을 보여주고 있다. 사용된 LK 알고리즘을 이용한 특징 벡터 추적은 지능형 로봇 시스템에서 내부 진동이나 모바일 로봇과 같은 가정용 로봇의 내부 진동이나 카메라 흔들림에 상관 없는 우수한 추적 성능을 가지며 몸통 안의 특징 벡터를 스네이크 알고리즘을 이용하여 동정함으로써 신체 구조비에서 발생되는 문제점과 LK 알고리즘에서 발생하는 특징 벡터들의 손실 및 잘못 추적되는 문제점을 극복한다.

그러나 그림 5와 같이 몸통 안의 특징 점은 추적이 가능하지만 몸통 밖의 특징 점은 추적이 어려울 것 알 수 있다. 몸통 밖의 특징 점 ( $f_5, f_6, f_7, f_8$ )는 움직임이 너무 크기 때문에 LK 알고리즘을 이용하여 추적하기 어렵고 실루엣이나 움직임 정보에서 특징 정보를 얻기 힘들기 때문에 스네이크 알고리즘을 이용한 동정도 힘들다. 따라서 본 논문에서는 움직임이 많은 몸통 밖의 특징 벡터 ( $f_5, f_6, f_7, f_8$ )을 보다 정밀하고 빠른 특징 점 추출을 위한 기법이 필요하여 신체정보를 기반으로 탐색 공간을 설정하고 손 탐지 알고리즘을 사용한다.

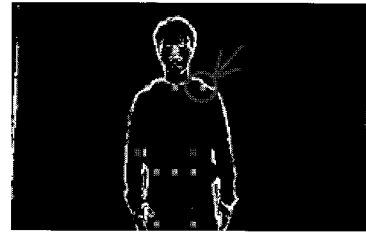
색상 기반 손 탐지 기법은 손의 피부색을 이용하는 방법으로 그 기법이 간단하고 제한된 상황에서 그 성능이 우수하기 때문에 피부색을 이용한 손 탐지 기법을 사용한다. 피부색을 탐지하기 위한 방법으로는 peer의 색상 모델을 사용한다[10,13].

Peer의 색상 모델은 조명 변화에 따라 제안된 RGB 값 범위를 선택하여 피부색 영역을 추출 하는 방법으로 간단하고 수행속도가 빨라 본 연구시스템에 적합하다. 그러나 peer의 색상 모델은 정확한 색상 영역 보다는 부정확한 영역을 포함하는 경우가 있기 때문에 움직임 정보와 색상 정보를 혼합하는 (14)와 같은 색상 정보를 사용한다.

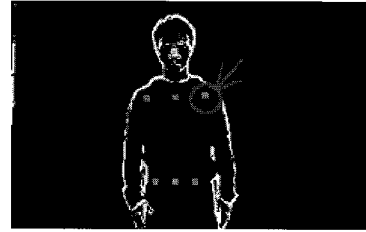
$$\bar{I}_p = I_p \bar{I}_t \quad (14)$$

여기서,  $I_p$ 는 peer의 색상 정보에서의 피부색 영역을 나타내고,  $I_t$ 는 움직임 정보이다.

손의 위치를 탐지하기 위하여 본 논문에서는 간단한 히스토그램 분석 기법을 사용한다. 연속적인 영상에서 peer의 색상 모델을 이용하여 피부색 영역을 추출한다. 추출된 피부색 영역 정보로부터 피부색 정보가 담긴  $\bar{I}_p$  영상의 히스토그램을 만들고 두 손이 겹치거나 얼굴을 중심으로 왼손과 오른손의 영역을 남



(a) LK로 추적된 특징벡터 영상



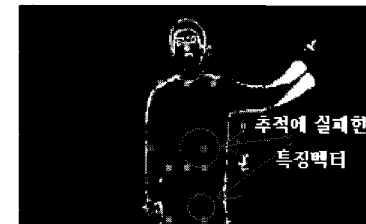
(b) 스네이크로 동정된 특징벡터 영상

그림 4. 특징 벡터 추출 예.

Fig. 4. Example of feature vector extraction.



(a) 입력 이미지



(b) 추적이 실패한 특징벡터 이미지

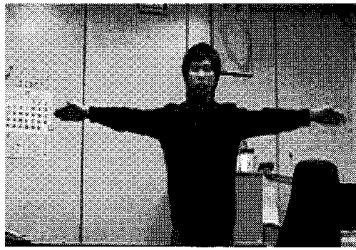
그림 5. 추적이 실패한 몸통 밖의 특징 벡터 예.

Fig. 5. Example of features vector located outside body.

지 않는다는 가정 하에 몸 중앙을 중심으로 가로 방향과 세로 방향으로 히스토그램을 만들고 평균값 이동 기법을 사용하여 얼굴의 위치를 제외한 피부색 정보가 가장 많은 부분을 손이라 판단함으로써 손의 위치를 추적 한다.

색상기반의 피부색 정보를 이용하여 손의 위치가 파악되면 어깨의 위치를 파악할 수 있다. 손 탐지 알고리즘으로 추출된 손의 위치가  $f_7$ , 상기 알고리즘으로 추정된 어깨의 위치가  $f_2$ 이다. 따라서 팔꿈치의 위치  $f_5$ 는 (15)와 같이 계산된다.

$$\begin{aligned} f_{5x} &= \frac{S_4 f_{7x} + S_6 f_{2x}}{S_4 + S_6} - \frac{1}{\sqrt{1+a^2}} \frac{2b}{d} \\ f_{5y} &= \frac{S_4 f_{7y} + S_6 f_{2y}}{S_4 + S_6} - \frac{a}{\sqrt{1+a^2}} \frac{2b}{d} \end{aligned} \quad (15)$$



(a) 입력 이미지



(b) 손 위치 추출 이미지

그림 6. 손 탐지 알고리즘을 이용한 손 위치 추출 예.

Fig. 6. Example of hand position extraction using hand searching algorithm.

여기서, (15)의 각 파라미터는 다음과 같이 계산된다.

$$a = -\frac{f_{2y} - f_{7y}}{f_{2x} - f_{7x}}$$

$$b = \sqrt{(S_4 + S_6 + d)(S_6 + d)(S_4 + d)(S_4 + S_6)}$$

$$d = \sqrt{(f_{2x} - f_{7x})^2 + (f_{2y} - f_{7y})^2}$$

또한, 팔꿈치의 위치  $f_6$ 도 상기 방법과 같은 방법으로 계산된다.

본 논문에서 최종적으로 얻고자 하는 것은 인간 관절 부위의 특징 점이다. 따라서 본 논문에서 제안한 하이브리드 스켈레톤 특징 점 추출 기법은 얼굴을 중심으로 가상의 신체 비율 정보를 통해 얻어진 초기 스켈레톤 특징 점을 LK 알고리즘으로 추적하였고, 추적된 특징 벡터에서 몸통 안의 특징 벡터를 스네이크 알고리즘을 이용하여 동정하였다. 탐색 공간을 설정 후 손 탐지 알고리즘을 이용하여 몸통 밖의 특징 점을 추출하였다. 최종적으로 상기 기법들을 블록 합하여 (16)과 같이 하이브리드 스켈레톤 모델을 생성한다.

$$F_k = (1 - \alpha - \beta)F_i + \alpha E_j + \beta C_k \quad (16)$$

$(0 < \alpha + \beta \leq 1)$

여기서,  $C_k$  는 몸통 밖의 특징 점이고,  $j = (1, \dots, 4, 9, 10, 11)$ ,  $k = (5, \dots, 8)$ 이다.

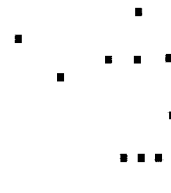
**IV. 모의실험**

본 논문에서 실험에 사용된 영상은 CCD 카메라로 부터 얻은 320×240 크기의 24 bit 컬러 영상이고 프레임 속도는 10frame/sec 이며 신체 중 상부 영상만을 사용되었다. 또한, 손을 크로스 하거나 민소매 형태는 고려하지 않았다.

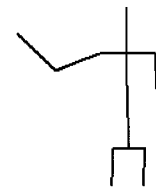
그림 7, 8은 본 논문에서 제안한 기법에 의해 생성된 하이브



(a) 입력 영상



(b) 하이브리드 특징점 추출 영상



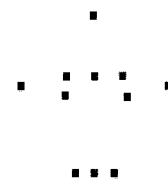
(c) 하이브리드 스켈레톤 모델 영상

그림 7. 하이브리드 스켈레톤 모델 생성 예 (1).

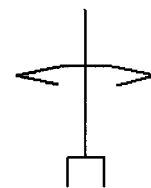
Fig. 7. Example 1 of hybrid skeleton model generation.



(a) 입력 영상



(b) 하이브리드 특징점 추출 영상



(c) 하이브리드 스켈레톤 모델 영상

그림 8. 하이브리드 스켈레톤 모델 생성 예 (2).

Fig. 8. Example 2 of hybrid skeleton model generation.

표 1. 조명변화에 따른 하이브리드 특징점 추출률.

Table 1. Extraction ratio of hybrid skeleton features for illumination changing.

구분/프레임 수	밝기 0	밝기 1	밝기 2	평균추출률
프레임 수(frame)	570	680	630	
얼굴추출 프레임 수	470	640	590	
손 추출 프레임 수	350	500	460	
팔꿈치 추출 프레임 수	260	390	360	
추출 성공률	74%	78%	76.1%	76%

리드 스켈레톤 모델의 생성 실험 결과이다. 그림에서 볼 수 있듯이 실험 결과는 정확히 특징점들을 추출함을 알 수 있다. 또한, 실험시 얼굴 추출이 가능하면 조명 변화에 크게 좌우되지 않지만 얼굴 추출이 되지 않는 조명 밝기일 경우 특징점 추출이 되지 않는 문제점이 확인 되었다. 그러나 얼굴 추출이 가능한 영상에서는 조명 변화에 의한 특징점 추출률이 크게 좌우되지 않는 결과를 얻었다. 얼굴 추출이 가능한 영상에서 조명 변화에 따른 추출률은 표 1과 같다.

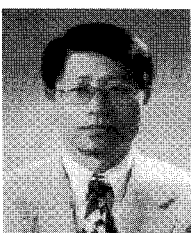
#### V. 결론

본 논문은 인간-로봇의 자연스러운 상호작용을 위한 하이브리드 스켈레톤 특징점 추출 기법을 제안하였다. 제안된 기법은 실험에 의해 만족할 만한 특징점 추출률을 보였고, 또한, 특징점 추출 결과를 스켈레톤 모델을 이용하여 인간 모션을 정확히 표현하였다. 마지막으로 제한된 환경 내에서 추출 결과로 인간-로봇 상호작용을 위한 인간 모션 추출은 본 논문에서 제안한 기법의 응용가능성을 보였다.

#### 참고문헌

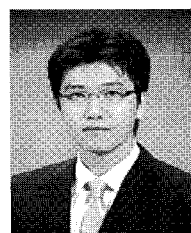
- [1] D. M. Gavrila and L. S. Davis, "Towards 3D model based tracking and recognition of human movement: a multi view approach," *Int Workshop on Face and Gesture Recognition*, vol. 162479, pp. 272-277, 1995.
- [2] V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human computer interaction: A review," *IEEE, Transaction on PAMI*, vol. 19, no. 7, pp. 677, July, 1997.

- [3] G. V. Veres, L. Gordon, J. N. Carter, and M. S. Nixon. "What image information is important in silhouette based gait recognition?," *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 776-782, 2004.
- [4] C. R. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780-785, 1997.
- [5] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 809-830, 2000.
- [6] R. Cutler and L. Davis, "Robust real-time periodic motion detection, analysis, and applications," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 781-796, 2000.
- [7] H. F. and A. J. Lipton, "Real time human motion analysis by image skeletonization," *Proceedings of The 4th Workshop of Applications of Computer Vision*, pp. 15-21, 1998.
- [8] Y. Li, A. Hilton, and J. Illingworth, "A relaxation algorithm for real-time multiple view 3D-tracking," *Journal of Image and Vision Computing*, vol. 20, pp. 841-859, 2002.
- [9] R. Hoshino, S. Yonemoto, D. Arita, and R. Taniguchi. "Real time motion capture system based-on silhouette contour analysis and inverse kinematics," *7th Korea-Japan Joint Workshop on Computer Vision*, vol. 7, pp. 157-163, 2001.
- [10] <http://opencvlibrary.sourceforge.net>
- [11] Bruce D. Lucas and T. Kanade. "An iterative image registration technique with an application to stereo vision," *International Joint Conference on Artificial Intelligence*, pp. 674-679, 1981.
- [12] C. Xu, and J. Prince, "Snakes, shapes, and gradient vector flow," *IEEE. Transactions on Image Processing*, vol. 7, no. 3, pp. 359-369, 1998.
- [13] <http://rv.fri.uni-lj.si/~peterp/publications/eurocon03.pdf>



주영훈

2007년 제 13 권 제 4 호(2007년 4월 호) 참조.



소세윤

2006년 군산대 전자정보공학부 졸업. 2006년~현재 동 대학원 석사과정. 관심분야는 지능제어, 로봇 비전, 로봇시스템.