

# HCI를 위한 컴퓨터비전 기반 사람의 2차원 움직임 정보 추출

이상환\*, 안상철\*, 김익재\*, 김형곤\*, 김재희\*\*

\*한국과학기술연구원, 영상미디어 연구센터,  
\*\*연세대학교 전기컴퓨터공학과 컴퓨터비전 연구실

## Vision based 2D Human Body Motion Extraction for HCI

SHwan Lee\*, Sang Chul Ahn\*, Ig-Jae Kim\*, Hyoung-Gon Kim\*, Jaihie Kim\*\*

\*Imaging Media Research Center, KIST

\*\*Computer Vision Lab., Dept. of Electrical and Computer Engineering, Yonsei University

### Abstract

본 논문은 특별한 마커를 사용하지 않고 연속되는 영상들에서 사람의 2차원 움직임 정보를 추출하는 알고리즘을 제안한다. 사람의 움직임 정보 추출을 위해 색상, 움직임, 윤곽선, 그리고 사용자의 개인적인 특성을 반영하는 신체적인 특성 정보를 이용한다. 본 논문에서 제안하는 방법은 사람의 움직임 정보를 이용한 가상 캐릭터의 제어와 같은 마커를 사용하지 않는 다양한 HCI 응용분야에 사용될 수 있다.

Keywords: 사람의 2차원 움직임 정보, 노 마커, 신체의 특성 정보, 색상, 윤곽선, 가상 캐릭터

### 1. 서론

하드웨어 기술의 발달과 함께 컴퓨터의 사용 증가로 인해 카메라로부터 입력된 영상을 대상으로 컴퓨터비전 방법을 이용하여 사람의 움직임 정보를 추출하고 자세를 인식하기 위한 연구들이 HCI(Human-Computer Interaction)분야에서 상당한 주목을 받고 있다[1, 2]. 사람의 움직임 정보를 이용한 사람과 컴퓨터사이의 의사소통을 위한 목적의 응용에서 사람의 움직임 정보를 추출하기 위한 목적으로 특별한 장치 또는 마커를 사용한다면, 오히려 사람의 자연스러운 움직임을 방해하게 된다. 따라서 마커를 사용하지 않고 사람의 움직임 정보를 추출하기

위한 방법들에 관한 연구들이 필요하다.

최근 들어 사람의 움직임 정보를 추출하고 자세를 인식하기 위한 목적의 연구들이 활발히 진행 중이다[1-6]. Pfinder[3]는 색상, 형태, 그리고 움직임 정보를 이용하여 사람을 여러 부분의 확률적인 모델로 모델링한다. Ghost[4]는 사람의 실루엣에 사람의 몸이 가지는 토폴로지 정보를 적용하여 다양한 자세에서 움직임 정보를 추출하고 있다. W4[5]는 사람의 움직임 정보를 얻기 위해 움직임과 형태정보를 가지는 특별한 모델과 동역학적인 템플릿 매칭을 사용하고 있다. ATR 연구소에서는 실험적인 근거를 기초로 하는 윤곽선 분석방법을 이용하고 있다[6].

그러나, 기존 대부분의 연구들이 사람의 움직임에 있어서 일부분만을 추출하거나, 몸통부분의 움직임에는 큰 비중을 두고 있지는 않다. 본 논문에서는 특별한 마커의 사용없이 사람의 2차원 움직임 정보 추출을 위해 보다 많은 특징점의 설정과 함께 몸통부분의 움직임 정보 추출에 비중을 두고 있는 알고리즘을 제시한다.

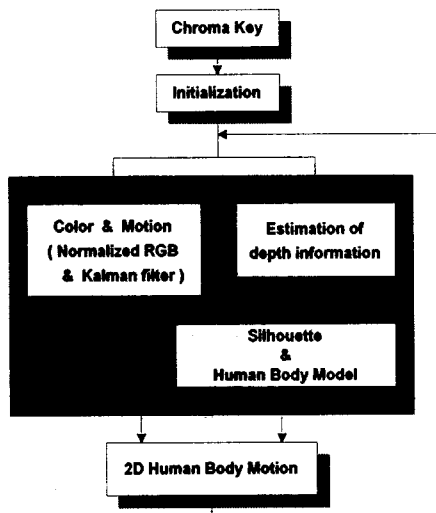


그림 1 전체 흐름도

## 2. 사람의 움직임 정보 추출

본 논문에서 제안하는 사람의 움직임 정보를 추출하는 과정은 그림 1과 같다. 제안하는 알고리즘은 최종적으로 실시간 처리를 목적으로 하고 있기 때문에 균일한 배경영역을 대상으로 한다는 가정아래 입력되는 영상에서 사람에 해당하는 영역만을 추출하는 방법으로 가장 쉽고 빠른 크로마키 방법을 사용한다.

### 2.1. 초기 과정 및 움직임을 대표하는 특징점

먼저 사용자가 카메라의 FOV(Field Of View)내로 들어오면 그림 2와 같이 미리 정의된 초기자세를 취하게 된다. 이것으로부터 사용자의 신체적인 형태에 대한 정보와 사용자가 입고 있는 상의, 하의, 그리고 사용자의 피부영역에 대한 색상정보를 추출하게 된다. 또한 움직임 추적을 위해 사용되는 칼만필터를 위한

초기값이 얻어진다. 이러한 과정이 초기과정이다.

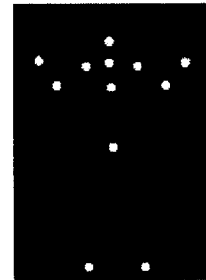


그림 2 초기자세 및 추출되는 특징점

사람의 움직임 정보를 얻기 위해서는 먼저 사람의 움직임을 잘 표현할 수 있는 특징점의 설정이 필요하다. 본 논문에서는 사람의 움직임 특성과 신체적인 특성, 그리고 기존 연구들[1-6]을 반영하여 그림 2와 같이 얼굴, 손, 발과 같이 신체의 끝부분들과 목, 어깨, 팔꿈치의 관절부분, 그리고 몸통에 해당하는 가슴중심과 골반중심을 특징점으로 설정한다.

추출되는 특징점들은 그 특성을 반영한 추출 방법에 따라 다음과 같이 크게 두 가지로 나뉜다.

#### 2.1.1. 색상을 기반으로 하는 특징점(CFP, Color-based feature point)

얼굴과 손은 피부색이라는 특별한 색상정보를 가지고 있다. 칼만필터와 함께 이러한 색상정보를 이용하여 추출되는 특징점이 CFP이다.

피부색의 색상정보는 여러 명의 사람으로부터 얻은 피부색의 자료의 분석을 통해서 정규화된 RG 색상공간상에서 하나의 가우시안 분포로 모델링된다. 초기과정에서는 이러한 모델링을 통해서 손과 얼굴영역을 얻은 다음, 미리 얻어진 상의 및 하의영역의 색상정보와의 비교를 통해서 사용자의 피부색을 가장 잘 분할할 수 있는 정규화된 RGB색상공간에서 두 가지 색상 성분과 각 성분에서 경계값을 추출하게 된다. 이러한 색상분할 과정은 상의 및 하의 영역의 분할에도 동일하게 적용된다. 또한 칼만필터의

사용에 있어서 초기값 문제를 해결하기 위해 손과 얼굴의 초기위치를 구하게 된다.

### 2.1.2. 형태를 기반으로 하는 특징점 (SFP, Structure-based feature point)

사람의 움직임 정보를 추출하기 위해 사용되는 특징점들 중 CFP를 제외한 흉부, 골반중심, 목, 어깨, 팔꿈치, 그리고 발을 형태를 기반으로 하는 특징점(SFP)으로 정의한다. 이러한 특징점들은 사람의 기본적인 형태 정보를 이용하여 추출된다.

초기과정에서는 상의의 높이, 몸통의 너비, 흉부와 둔부중심의 위치, 그리고 어깨와 목과의 거리가 얻어진다. 그리고 이러한 값들간의 관계도 얻어진다. 이것이 사용자의 형태정보이다. 이 과정에서 몸통모델이 사용자에게 맞게 설정되고 흉부와 둔부의 중심은 몸통모델로부터 얻어진다. 앞 절에서의 설명과 같이 상의와 하의의 분할을 위한 색상정보가 얻어진다.

몸통모델은 그림 3에서와 같이 간결한 상자 형태로 이루어져있다. 머리와 목을 제외한 사용자의 몸통부분을 6등분하고 있다. 위로부터 첫 번째 영역은 흉부영역이 되고 세 번째영역은 둔부영역이 된다. 그리고 각 영역의 중심이 흉부와 둔부의 중심이 된다. 초기과정에서 설정된 사용자의 형태정보, 그리고 색상정보는 매 순간 입력되는 영상에 적용된다.

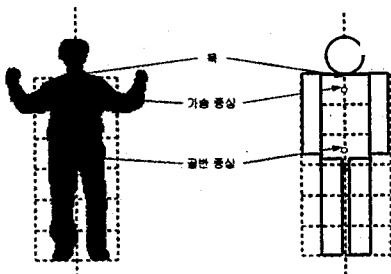


그림 3 사람의 몸통모델

## 2.2. 2차원의 움직임 정보 추출

### 2.2.1. CFP의 추적

초기과정 이후 CFP는 칼만필터를 이용하여 다음 프레임의 영상에서의 각 영역의 위치를 예측한다[7]. CFP는 짧은 시간간격동안 등가속도 운동을 한다고 가정한다.

본 논문에서는 위치값, 속도값, 가속도값을 상태변수로 가지는 상태방정식이 아니라 모든 상태변수가 위치로 표현되는 식(1)의 상태방정식과 식(2)의 측정방정식을 사용한다.

$$\begin{bmatrix} p_{k+1} \\ p_k \\ p_{k-1} \end{bmatrix} = \begin{bmatrix} 2.5 & -2 & 0.5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} p_k \\ p_{k-1} \\ p_{k-2} \end{bmatrix} + \begin{bmatrix} \omega'_{p_k} \\ \omega'_{p_{k-1}} \\ \omega'_{p_{k-2}} \end{bmatrix} \quad (1)$$

$$Z_k = [1 \ 0 \ 0] \begin{bmatrix} p_k \\ p_{k-1} \\ p_{k-2} \end{bmatrix} + v'_k \quad (2)$$

식(1)의 상태방정식은 시간과 관련된 어떠한 변수도 가지지 않는다. 즉 시간에 독립적인 특성을 가진다. 그리고 상태방정식이 가지는 상태변수와 얻어지는 측정값이 모두 위치값이므로 모든 상태변수는 예측된 다음, 보정되어지므로 보다 정확한 다음 상태의 예측이 가능해진다. 예측된 위치를 중심으로 설정된 일정한 크기의 탐색영역내에서 초기과정에서 설정된 피부색 영역의 색상정보를 만족하는 화소들을 추출한다. 추출된 영역의 위치는 CFP의 각 프레임에서의 위치값이 된다.

### 2.2.2. SFP의 추출

SFP의 추출은 CFP의 추적결과, 몸통모델, 그리고 사용자의 윤곽선정보를 사용한다. 얼굴의 중심과 사용자 영역의 중심을 잇는 선(body axis 1)과 이전 프레임에서 예측된 상의의 너비를 이용해서 현재 프레임의 영상에서 상의영역의 후보영역을 설정한다. 그리고 상의영역의 색상정보를 이용하여 상의영역을 얻는다. 그런 다

음, 상의영역의 중심과 상의의 높이를 추출하게 된다. 추출된 상의의 높이는 현재 프레임에서 사용자와 카메라사이의 거리를 정보, 즉 사람의 3차원 위치를 반영하고 있다. 따라서 상의의 높이를 이용한 정규화 과정을 통해서 초기과정에서 추출한 사용자의 형태 정보를 적용하게 된다. 얼굴의 중심과 상의영역의 중심을 잇는 선 (body axis 2)을 생성한다. 이 직선을 이용하여 목, 어깨, 흉부중심, 그리고 골반중심을 추출한다. 발은 골반의 중심을 기준으로 좌우에서 가장 아래쪽에, 그리고 가장 외곽에 위치한 윤곽선 위의 한 점으로 설정한다. 팔꿈치는 ATR연구소에서 사용한 윤곽선 분석법[6]을 변형한 방법을 이용하여 추출한다. 본 논문에서 제안하는 방법의 경우는 팔꿈치의 추출에 있어서 좌우측 최외각점의 위치와 추출된 어깨를 사용하고 있기 때문에 보다 정확한 팔꿈치의 추출이 가능

하게 된다.

### 3. 실험

CCD 카메라를 통해서 얻은 연속적인 영상데이터를 초당 10장씩 샘플링 한 영상데이터를 대상으로 실험했다. 영상의 크기는 360x243의 크기를 가진다. 영상데이터의 획득은 사용자 영역의 추출을 간결하게 하기 위해 배경이 푸른색으로 이루어진 가상 스튜디오내에서 수행되었다. 본 논문에서 제안하는 알고리즘은 C/C++을 이용하여 구현되었고, PC(Pentium III, 600MHz)에서 구동되었다. 실험결과는 그림 4에 나타나있다.

영상으로부터 추출된 사람의 움직임 정보를 가상 캐릭터에 부여하여 사람과 유사한 움직임을 갖도록 하는 작업을 수행했다. 가상 캐릭터는 Open Inventor를 사용하여 구현되어졌다. 사

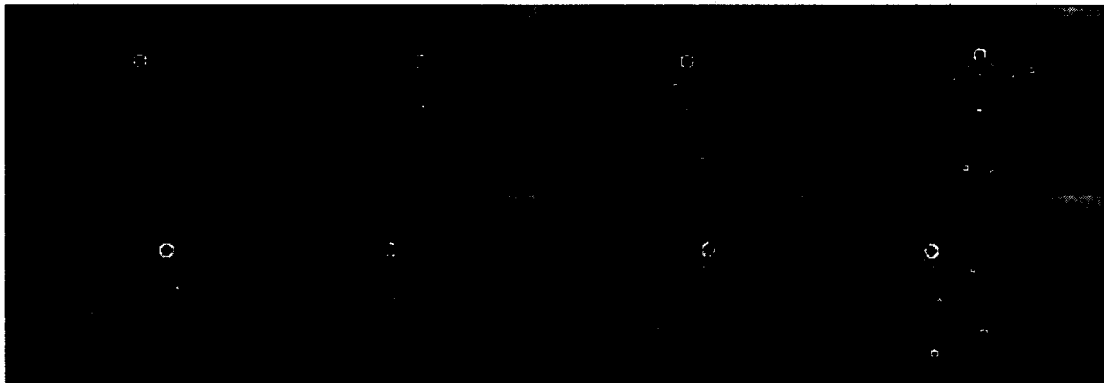


그림 4 다양한 자세와 위치를 가지는 사람의 2차원 움직임 정보 추출결과

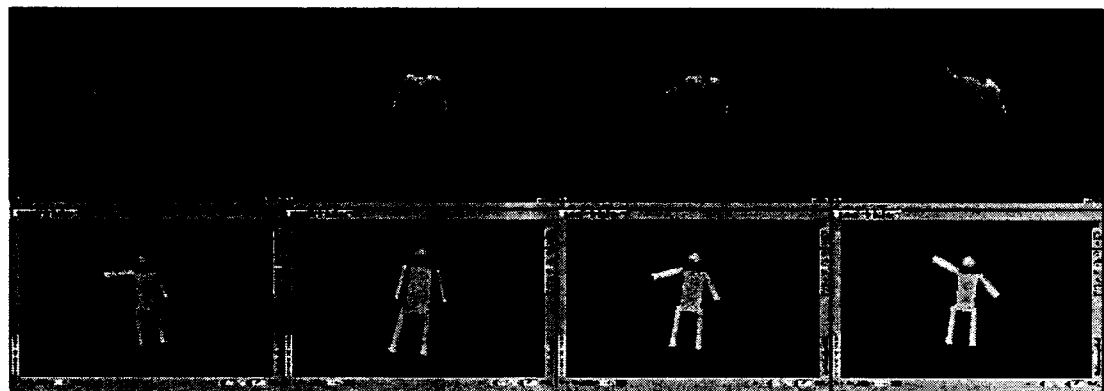


그림 5 사람의 움직임을 이용한 가상캐릭터의 제어

람의 움직임을 이용한 가상 캐릭터의 제어 결과는 그림 5이다.

#### 4. 결론

본 논문에서 제안하는 알고리즘은 어떠한 특별한 마커를 사용하지 않고 사람의 2차원 움직임 정보를 추출한다. 제안하는 알고리즘은 2차원의 영상을 통해서 사용자와 카메라사이의 거리 관계, 즉 대략적인 3차원의 거리를 예측한다. 이러한 방법은 다양한 3차원 움직임에 대해 사용자의 신체적인 정보가 반영된 흉부와 골반의 중심, 그리고 특별한 형태정보를 가지지 않는 어깨를 효율적으로 추출할 수 있게 한다.

특별한 외부적인 장치없이 사람의 움직임 정보만을 이용하여 가상캐릭터를 제어하는 작업을 수행하였다. 이러한 작업은 HCI분야에서 특별한 장치의 사용없이 사람의 움직임 정보를 이용한 보다 자연스럽게 편리한 인터페이스의 구현이 가능함을 보인다. 또한 대략적인 3차원 정보의 추출은 보다 다양한 인터페이스의 구현을 가능하게 한다.

#### 참고문헌

- [1 ] D. M. Gavrila, The Visual Analysis of Human Movement: A Survey, Computer Vision and Image Understanding, Vol. 73, No. 1, January, pp.82-98, 1999.
- [2 ] J. K. Aggarwal and Q. Cai, Human Motion Analysis: A Review, Computer Vision and Image Understanding, Vol. 73, No. 3, March, pp.428-440, 1999.
- [3 ] C. Wren, A. Azarbayejani, T. Darrell, A. Pentland, Pfinder: Real-Time Tracking of the Human Body, In Proc. Of the SPIE Conference on Integration Issues in Large Commercial Media Delivery Systems, October, 1995.
- [4 ] Ismail Haritaoglu, David Harwood and Larry S. Davis, Ghost: A Human Body

Part Labeling System Using Silhouettes, 14th International Conference on Pattern Recognition, August 16-20, 1998.

- [5 ] Ismail Haritaoglu, David Harwood and Larry S. Davis, W4: Who? When? Where? What? A Real Time System for Detecting and Tracking People, International Conference on Face and Gesture Recognition, April 14-16, 1998.
- [6 ] Kazuhiko Takahashi, Tatsumi Sakaguchi, and Jun Ohya, Real-time Estimation of Human Body Postures using Kalman Filter, RO-MAN99 8 International Workshop on Robot and Human Interaction, September 27-29, 1999.
- [7 ] Robert Grover Brown and Patrick Y. C. Hwang, Introduction to Random Signals and Applied Kalman Filtering, Addison-Wesley, 1997.