

# 컴퓨터비전을 이용한 사람의 2차원 움직임 정보 추출

이상환\*\*\*, 안상철\*, 김익재\*, 김형곤\*, 김재희\*\*  
\*한국과학기술연구원, 영상미디어 연구센터  
\*\*연세대학교 전기컴퓨터공학과 컴퓨터비전 연구실  
e-mail : \*{shlee, asc, kij, hgk}@cherry.kist.re.kr  
\*\*jhkim@bubble.yonsei.ac.kr

## Vision based 2D Human Body Motion Extraction

SHwanLee\*\*\*, 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

### 요 약

본 논문은 특별한 마커를 사용하지 않고 연속되는 영상들에서 사람의 2차원 움직임 정보를 추출하는 알고리즘을 제안한다. 사람의 움직임 정보를 추출하기 위해 색상, 움직임, 그리고 윤곽선 정보를 이용한다. 뿐만 아니라 사용자의 신체적인 차이와 특징점의 일관성을 위해 사람 몸통 모델을 사용한다. 본 논문의 알고리즘은 마커를 사용할 수 없는 HCI 응용분야에 될 수 있다..

### 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 상당히 한 주목을 받고 있다. 그리고 객체기반 통신의 기본이 되는 MPEG-4는 사람의 움직임 정보를 전달하는 응용분야를 위해 몇 가지 특징점들을 정의하고 있다[1]. 실제의 응용에서 이러한 특징점들을 추출하기 위한 목적으로 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 [2-11]. Pfunder[8]는 색상, 형태, 그리고 움직임 정보를 이용하여 사람을 여러 부분의 확률적인 모델로 모델링한다. Ghost[9]는 사람의 실루엣에 사람의 몸이 가지는 토폴로지 정보를 적용하여 다양한 자세에서 움직

임 정보를 추출하고 있다. W4[10]는 사람의 움직임 정보를 얻기 위해 움직임과 형태정보를 이용하는 특별한 모델과 동역학적인 템플릿 매칭을 사용하고 있다. W4는 일반적인 환경에서 적용 가능하나 손이 몸통내부에 위치하는 경우에는 손에 대한 어떠한 정보도 추출할 수 없다. ATR 연구소에서는 실험적인 근거를 기초로 하는 윤곽선 분석방법을 이용하고 있다[11]. 본 논문에서 제안하는 알고리즘과 가장 유사한 방법이지만, 사람의 몸통에 관한 정보의 추출에는 큰 비중을 두고 있지는 않다. 이러한 점은 앞서 언급된 연구들 역시 큰 비중을 두고 있지는 않다.

본 논문은 영상통신이나 HCI 같은 응용분야들을 위해 사람의 몸위에 특징점들을 정의하고 입력되는 영상에서 이들을 추적함으로써 해서 사람의 2차원 움직임 정보를 추출하는 알고리즘을 제안한다. 본 논문에서는 사람의 움직임에 있어서 큰 의미를 가지는 손, 얼굴, 그리고 발과 같은 부분들 뿐만 아니라 사람의 움직임에 있어서 중요한 정보가 되는 가슴과 골반부분의 특징점 정보들까지 추출하고 있다. 이러한 정보는 사용자의 신체적인 차이를 반영하여 정의하고 추출하게 된다. 그리고 이러한 움직임 정보의 추출을 위해 색상,

움직임, 윤곽선, 그리고 몸통 모델의 다양한 정보들을 사용하고 있다. Fig.1 는 본 논문에서 움직임 정보를 추출하기 위해 사용되는 특징점들이다.



Figure 1. 초기자세 및 추출되는 특징점

## 2. 사람의 움직임 정보 추출

Fig.2 는 사람의 움직임 정보를 추출하는 알고리즘의 순서도이다. 본 논문에서는 사람영역을 추출하기 위해 크로마키방법을 사용한다. 이것은 간단하고 빠르게 사람영역을 얻기 위한 목적으로 사용된다.

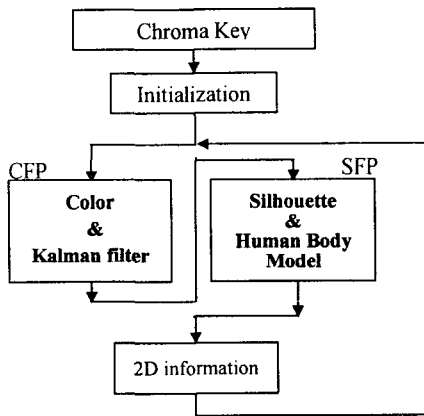


Figure 2. 사람의 움직임 정보를 추출하는 과정

### 2.1 초기 과정 및 움직임을 대표하는 특징점

사람의 움직임 정보를 추출하기 위해서는 먼저 사용자가 카메라의 FOV(Field Of View)내로 들어오면 사용자에 대한 몇 가지 정보를 추출하는 초기과정을 수행하게 된다. 사용자의 신체적인 형태에 대한 정보와 사용자가 입고 있는 상의, 하의, 그리고 사용자의 피부영역에 대한 색상정보를 추출하게 된다. 또한 움직임 추적을 위해 사용되는 칼만필터를 위한 정보가 얻어진다. Fig.1 은 초기과정을 위해 사용자가 취해야 하는 초기자세를 나타내고 있다.

사람의 움직임 정보를 표현하기 위해 움직임에 있어서 큰 의미를 가지는 부분인 얼굴, 손, 발을 특징점으로 설정한다. 뿐만 아니라 몸통의 움직임을 위해 이 부분을 크게 상체와 하체로 나누어서 고려했다. 각 영역에서 중심이 되고 형태적으로 큰 변화가 없는 부분

인 흉부와 골반영역을 정의하고 이 영역들의 중심점을 특징점으로 정의했다. 또한 팔의 움직임을 표현하는데 있어서 큰 비중을 가지는 어깨를 특징점으로 정의했다. 그리고 어깨를 추출하기 위해 보조적으로 사용되는 특징점으로 목의 측면을 정의했다. 결국 이러한 특징점의 움직임 정보가 사람의 움직임 정보를 대표할 수 있다고 가정한다.

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

피부색영역인 얼굴과 손은 칼만필터를 이용한 움직임과 정규화된 RGB 색상모델에서 색상정보를 추출한다. 이러한 특징점들이 CFP 이다. 여러 사람으로부터 얻은 피부색 정보를 이용하여 초기자세를 취하고 있는 사용자로부터 대략적인 얼굴과 손영역을 구한다. 이것으로부터 CFP 의 추출을 위해서 초기과정에서는 사용자의 피부색정보를 얻게 되고 이 색상정보는 피부색, 상의, 그리고 하의영역의 분할을 위한 임계값 설정에 사용된다. 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 특징점중 CFP 를 제외한 흉부, 골반중심, 어깨, 발, 그리고 보조 특징점인 목의 옆면을 SFP 라고 한다. 이러한 특징점들은 사람의 기본적인 형태 정보를 이용하여 추출되기 때문이다.

초기과정에서는 초기자세로부터 상의 높이, 몸통의 너비, 흉부와 둔부중심의 위치, 그리고 어깨와 목의 옆면과의 거리와 같은 사용자의 형태정보가 추출된다. 이 과정에서 몸통모델이 사용자에게 맞게 설정되고 흉부와 둔부의 중심은 몸통모델로부터 정의된다. 그리고 정규화된 RGB 색상모델에서 상의와 하의영역의 색상정보가 추출된다.

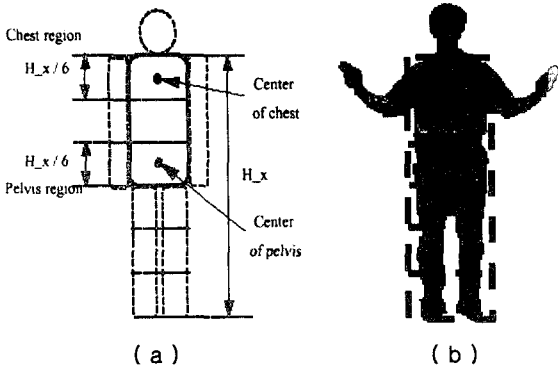
정규화된 RGB 색상공간에서 피부색, 상의, 그리고 하의영역의 색상정보는 가우시안 분포로 모델링된다. 이들 3 개영역의 분할을 위한 임계값의 설정을 위해 조건부 확률이 사용된다[12].

몸통모델은 Fig.3 에서와 같이 간결한 상자형태로 이루어져있다. 머리와 목을 제외한 사용자의 몸통부분을 6 등분하고 있다. 위로부터 첫번째 영역은 흉부영역이 되고 세번째영역은 둔부영역이 된다. 그리고 각 영역의 중심이 각각 흉부와 둔부의 중심이다. 초기과정에서 설정된 이러한 중심들은 매 프레임의 영상에서 추출된다.

## 2.2. 2 차원의 움직임 정보 추출

### 2.2.1. CFP 의 추적

CFP 는 초기과정 이후에 이전 프레임의 영상에서 얻어지는 각 영역의 위치값을 측정값으로 사용해서

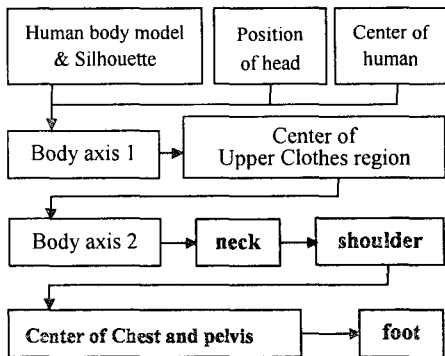


**Figure 3. 사람의 몸통 모델**  
 (a) 기본적인 사람의 몸통 모델  
 (b) 사용자에게 맞게 적용된 몸통 모델

현재 프레임의 영상에서 각 영역의 위치를 칼만필터를 이용하여 예측한다[13]. 여기서 CFP는 짧은 시간간격동안 등가속도 운동 모델을 만족하는 운동을 한다고 가정한다. 예측된 위치를 중심으로 설정된 일정한 크기의 탐색영역내에서 정규화된 RGB 색상모델을 이용하여 초기과정에서 설정된 피부색의 색상영역을 만족하는 화소들을 추출한다. 추출된 영역의 위치는 CFP의 각 위치값이 된다. 그리고 다음 프레임의 영상에서 각 영역의 위치를 예측하기 위해 칼만필터의 측정값으로 사용된다.

2.2.2. SFP의 추출

SFP의 추출은 CFP의 추적결과와 몸통모델과 사용자의 윤곽선정보를 사용하고 있다. SFP의 추출과정은 Fig.4와 같다.

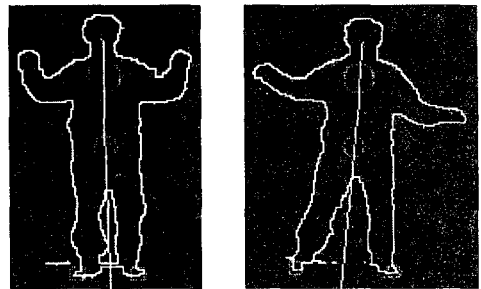


**Figure 4. SFP를 추출하는 과정**

얼굴의 중심과 사용자 영역의 중심을 잇는 선(body axis 1)과 이전 프레임에서 예측된 상의의 너비를 이용해서 현재 프레임의 영상에서 상의영역의 후보영역을 선정한다. 정규화된 RGB 색상공간에서 정의된 상의영역의 색상정보를 이용하여 상의영역을 얻는다. 이와 같은 상의영역의 추출방법은 팔부분이나 소매의 끝부

분과 같이 몸통부분과 동일한 색상을 가진 상의영역이지만 사용자의 몸통부분의 중심을 추출하는데 방해가 되는 요소들을 배제시킬 수 있다는 특성을 가진다. 여기서 상의영역의 중심과 상의의 높이를 추출하게 된다. 추출된 상의의 높이는 현재 프레임에서 사용자와 카메라사이의 거리를 정보를 반영하고 있다. 따라서 상의의 높이를 이용한 정규화 과정을 통해서 초기과정에서 추출한 사용자의 형태에 관한 정보들, 즉 흉부, 골반, 목과 어깨간의 거리등과 같은 정보들을 적용하게 된다. 이러한 방법은 사용자의 3차원 움직임에 대해 2차원의 정보만을 가진 영상에서 상의의 높이를 이용하여 사용자와 카메라사이의 거리를 관계를 예측할 수 있게 한다. 그리고 효율적으로 어깨, 흉부, 둔부의 중심과 같은 특징점을 추출할 수 있게 한다. 또한 특별한 형태를 가지지 않는 어깨를 추출할 수 있게 한다.

얼굴의 중심과 상의영역의 중심을 잇는 선(body axis 2)을 이용하여 어깨, 흉부중심, 그리고 골반중심을 추출한다. 이러한 특징점들의 추출은 Fig.5에 나타나 있다.



**Figure 5. 다양한 자세에 대한 SFP의 추출**

body axis 2를 따라서 상의의 높이를 이용한 정규화 과정을 통해서 초기 프레임에서 추출된 거리를 이용하여 흉부와 골반의 중심을 추출한다. 어깨는 상의의 높이를 이용한 정규화 과정을 통해서 얻은 초기 프레임에서의 목의 옆면과의 거리를 이용해서 추출하게 된다. 목의 옆면은 이전 프레임에서 얻은 목의 옆면의 위치정보를 이용해서 설정된 영역내에서 body axis 2를 중심으로 좌우로 거리가 가장 가까운 윤곽선 위의 한점으로 설정한다. 그리고, 골반의 중심을 기준으로 좌우에서 가장 아래쪽에, 그리고 가장 외곽에 위치한 윤곽선 위의 한 점을 발로 설정한다.

3. 실험

CCD 카메라를 통해서 얻은 연속적인 영상데이터를 초당 10장씩 샘플링 한 영상데이터를 대상으로 실험했다. 영상의 크기는 360x243의 크기를 가진다. 영상데이터를 얻는 장소는 사용자 영역의 추출을 간결하게 하기 위해 배경이 푸른색으로 이루어진 가상스튜디오내에서 수행했다. 본 논문에서 제안하는 알고리즘은 C/C++을 이용하여 구현되었고, PC(Pentium III, 600MHz)에서 구동되었다. 실험결과는 Fig. 6에 나타나

있다.

4. 결론 및 향후과제

본 논문에서 제안하는 알고리즘은 어떠한 특별한 마커를 사용하지 않고 사람의 2차원 움직임 정보를 추출한다. 제안하는 알고리즘은 2차원의 영상을 통해서 사용자와 카메라사이의 거리 관계를 예측한다. 이 예측된 거리 정보를 이용한 정규화과정을 통해서 초기과정에서 얻은 사용자에게 맞게 설정된 몸통모델과 사용자의 신체적인 정보를 적용하게 된다. 이러한 방법은 사용자의 신체적인 정보가 반영된 흉부와 골반의 중심, 그리고 특별한 형태정보를 가지지 않는 어깨를 효율적으로 추출할 수 있게 한다.

향후 과제로는 현재까지 수행된 사람의 2차원의 움직임을 추출하는 알고리즘을 3차원에 적용하는 연구와 사람의 움직임을 보다 정확하게 표현할 수 있도록 더 많은 특징점을 정의하고 추출하는 연구가 수행되어야 할 것이다.

참고문헌

[1] P.K. Doenges, T.K. Captin, F. Lavagetto, J. Ostermann, I.S. Pandzic, and E.D. Petajan, MPEG-4 : Audio/video & synthetic graphics/audio for mixed media, <http://miralabwww.unige.ch/ARTICLES/ICJ97.html>, 1997.

[2] 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.

[3] 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.

[4] Trevor Darrell, Pattie Maes, Bruce Blumberg, Alex P. Pentland, A Novel Environment for Situated Vision and Behavior, IEEE Workshop for Visual Behaviors, CVPR-94.

[5] Maylor K. Leung and Yee-Hong Yang, First Sight: A Human Body Outline Labeling System, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 4, April 1995.

[6] Shanon X. Ju, Michael J. Black, Yaser Yacoob, Cardboard People: A Parameterized Model of Articulated Image Motion, International Conference on Face and Gesture Analysis, 1996.

[7] Kazuyuki Imagawa, Shan Lu, Seiji Igi, Closed-Based Hands Tracking System for Sign Language Recognition, International Conference on Face and Gesture Recognition, April 14-16, 1998.

[8] 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.

[9] Ismail Haritaoglu, David Harwood and Larry S. Davis, Ghost: A Human Body Part LabelingSystem Using Silhouettes, 14th International Conference on Pattern Recognition, August 16-20, 1998.

[10] 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.

[11] Kazuhiko Takahashi, Tatsumi Sakaguchi, and Jun Ohya, Real-time Estimation of Human Body Postures using Kalman Filter, RO-MAN'99 8 International Workshop on Robot and Human Interaction, September 27-29, 1999.

[12] Earl Gose, Richard Johnsonbaugh, and Steve Jost, Pattern Recognition and Image Analysis, Prentice Hall, pp. 81-147, 1996.

[13] Robert Grover Brown and Patrick Y. C. Hwang, Introduction to Random Signals and Applied Kalman Filtering, Addison-Wesley, 1997.

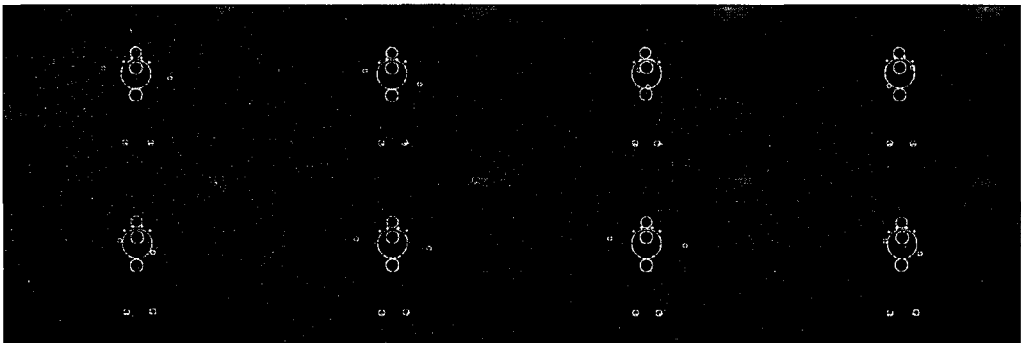


Figure 6. 다양한 자세를 취하는 사용자의 2차원 움직임 정보 추출