

# 인체 비율을 이용한 인체의 조인트 자동 검출 및 객체 추적 알고리즘

## Automatic Detecting and Tracking Algorithm of Joint of Human Body using Human Ratio

곽내정\*, 송특섭\*\*  
 목원대학교 정보통신공학과\*, 목원대학교 컴퓨터 공학과\*\*

Nae-Joung Kwak(knj0125@hanmail.net)\*, Teuk-Seob Song(teukseob@gmail.com)\*\*

### 요약

인간과 컴퓨터의 상호작용이 관심분야로 대두되면서 인체를 검출하고 추적하는 기술들에 관한 연구가 활발히 진행되고 있다. 본 논문에서는 단일카메라의 입력으로 인체의 비율을 이용하여 인체 부위를 연결하는 조인트를 자동으로 검출하고 객체를 추적하는 알고리즘을 제안한다. 제안방법은 입력영상과 배경영상의 회색조 영상과 색상 영상의 차영상을 구한 후 그 결과를 결합하여 배경과 전경을 분리하고 객체를 추출한다. 또한 얼굴길이와 인체 각 영역의 측정값을 이용하여 인체 비율을 모델링하고 추출된 객체 실루엣의 코너점과 모델링된 비율을 이용해 객체의 조인트를 자동으로 추출한다. 추출된 조인트의 움직임을 블록매칭 기법으로 객체의 움직임을 추적한다. 제안방법을 카메라로 입력되는 실험동영상에 적용한 결과 인체의 실루엣과 조인트를 자동 검출하며 추출된 조인트 또한 효율적으로 추적되었다.

■ 중심어 : | 객체검출 | 객체추적 | 실루엣 | 조인트 |

### Abstract

There have been studying many researches to detect human body and to track one with increasing interest on human and computer interaction. In this paper, we propose the algorithm that automatically extracts joints, linked points of human body, using the ratio of human body under single camera and tracks object. The proposed method gets the difference images of the grayscale images and ones of the hue images between input image and background image. Then the proposed method composes the results, splits background and foreground, and extracts objects. Also we standardize the ratio of human body using face' length and the measurement of human body and automatically extract joints of the object using the ratio and the corner points of the silhouette of object. After then, we tract the joints' movement using block-matching algorithm. The proposed method is applied to test video to be acquired through a camera and the result shows that the proposed method automatically extracts joints and effectively tracks the detected joints.

■ keyword : | Object Detection | Object Tracking | Silhouette | Joint |

## I. 서론

오늘날 인간과 컴퓨터의 상호작용이 관심으로 대두

되면서 이를 효율적으로 하기 위해 많은 연구가 되고 있다. 특히 사람과 컴퓨터의 효율적인 소통을 위한 음성인식, 수화인식, 제스처 인식 등의 분야와 사람의 행

\* 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구임  
 (No. 2010-0007549)

접수번호 : #110128-003

접수일자 : 2011년 01월 28일

심사완료일 : 2011년 04월 07일

교신저자 : 송특섭, e-mail : teukseob@mokwon.ac.kr

위를 분석하고 재구성하여 아바타 애니메이션이나 보행인식 및 인간의 행동분석, 감시시스템, 이벤트 인식 등에 적용하고자 하는 연구가 관심이 집중되고 있다. 이러한 연구의 기반 기술로서 인체를 검출하고 추적하는 기술은 비전 기술의 발전과 컴퓨터 성능의 향상으로 현재 연구가 활발히 진행되고 있다[1-3]. 기존의 연구 방법으로는 인체의 실루엣이나 윤곽선 정보 등은 관절이나 특정부위에 센서나 마커등을 부착하여 조인트를 추출하고 이를 응용하여 실루엣이나 윤곽선 및 다른 특징들을 자동으로 추출하는 방법[4-6], 동영상에서 색상을 이용하여 영상을 분할한 후 윤곽선을 분석하여 비슷한 색상으로 연결된 블랍들의 집합으로 인체를 표현하고 단일 카메라로 추적하는 방법[7][8], 네 대의 카메라에서 획득된 영상에 템플릿 매칭이 사용되거나[9], 직각으로 배치된 세대의 카메라에서 획득된 영상에 윤곽선 분석을 실행하는 방법 등이 소개되었다[10].

인체의 추적과 모델링을 위해 추출되는 특징으로는 인체의 실루엣, 윤곽선 정보, 특정 인체 부위 및 연결정보, 조인트 등이 있다. 이 특징 중 조인트는 인체의 각 부위를 연결하는 위치점으로 조인트의 위치를 이용하여 인체의 각 부분을 추정하고 그로 인하여 객체를 재구성하는 것과 추출된 조인트의 움직임을 추적하여 객체를 추적하는 것도 가능하다. 또한 조인트를 추출하여 아바타의 동작생성시 아바타의 각 관절을 표현하여 아바타 애니메이션에도 적용이 가능하며 인간의 행동 패턴을 분석할 경우에 조인트의 변화와 위치에 따라 분류하여 적용하는 것도 가능하다. 다양한 응용분야에서 조인트의 추출과 추적은 중요한 연구의 대상이 된다 [11][12]. 그러나 기존의 연구에서는 조인트 영역에 마커를 부착하거나 센서 등을 부착하여 그 위치를 획득하며 조인트의 위치를 자동으로 인식하거나 조인트의 매핑을 위한 추적하는 방법에 대한 연구는 미비한 상황이다. 마커나 센서를 부착하여 조인트를 획득할 경우 이것은 번거로운 작업이 되며 경제적인 비용도 고려된다. 또한 감시카메라 등의 응용분야에서는 마커나 센서를 사용할 수 없으며 이 경우 객체의 추적은 조인트의 추출대신 인체의 실루엣을 이용하여 객체를 추출하고 추적하는 방법을 적용하며 이때 인체의 실루엣을 추출하

기 위해 다수의 카메라가 필요하거나 스테레오 영상이 필요하다. 또한 추출된 객체는 감시카메라의 응용 분야에 적용이 가능하지만 아바타 애니메이션이나 인간의 행동 인식 분야에 적용될 때는 다른 방법을 적용해야 한다.

따라서 본 논문에서는 인간의 행동인식, 아바타 애니메이션, 감시카메라 등 다양한 응용에 적용이 가능한 인체의 조인트를 추출하고 조인트를 추적하여 객체의 움직임을 추적하는 알고리즘을 제안한다. 제안방법은 기존의 객체 추적 알고리즘에서 다수의 카메라를 사용하던 방법을 개선하여 현대의 고정 카메라를 이용하며 객체를 검출하여 실루엣을 추출하며, 또한 마커의 부착이나 센서등을 필요로하지 않고 자동으로 조인트를 추출한다. 제안방법은 먼저 배경영상 모델링 기법과 차영상 기법을 이용하여 인체의 실루엣을 추출한다. 제안방법은 추출된 실루엣을 이용하여 신체 비율 정보와 코너점 추출을 결합하여 인체의 조인트를 검출하고 검출된 조인트를 블록매칭 방법을 적용하여 추적한다.

## II. 블록매칭 알고리즘

동영상에서 객체의 움직임을 예측하는 방법은 크게 옵티컬 플로우 방법과 블록매칭 방법으로 나눌 수 있다. 이중 블록매칭 방법은 MPEG1, MPEG2, MPEG4 등에서 사용되고 있는 방법으로 현재 프레임의 특정 위치 블록이 이전 프레임의 어느 위치에서 이동한 블록인지 찾는 방법이다.

두 블록의 유사도를 평가하는 방법은 평균 절대값 차이(MAD: Mean Absolute Difference)주로 사용한다.  $t$  번째 프레임  $I_t$ 에서 좌표  $(k, l)$ 에  $M \times N$  크기의 블록과  $t-1$  번째 프레임  $I_{t-1}$ 에서 좌표  $(k+x, l+y)$ 의 블록과의 MAD는 다음과 같이 계산한다.

$$MAD_{(k,l)}(x,y) = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} |I_t(k+i, l+j) - I_{t-1}(k+x+i, l+y+j)| \quad (1)$$

위 식에서 MAD를 가장 작게 만드는 변위  $(x,y)$ 의 값을 해당블록의 움직임 벡터로 생각할 수 있다. 따라서 블록  $(k,l)$ 의 움직임 벡터를  $v(k,l)$ 로 나타내면 다음과 같다.

$$v(k,l) = \underset{(x,y)}{\operatorname{argmin}} MAD_{(k,l)}(x,y) \quad (2)$$

[그림 1]은 블록매칭의 움직임 추정과정을 그림으로 보여준다. [그림 1]의 (a)의 t번째 프레임에  $M \times N$  크기의 블록은 (b)의 t-1번째 프레임의 해당 블록 주변 검색 윈도우 안에서 해당 블록과 MAD 값이 가장 작은 위치의 블록을 찾고 그 위치 변화를 움직임 벡터로 인지하게 된다. (c)는 찾아낸 움직임 벡터의 모습을 보여준다.

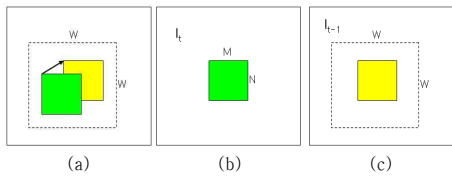


그림 1. 블록매칭 방법

블록 매칭 방법에서 특정블록 주변의 모든 변위에 대하여 MAD를 계산하는 방법을 전역 검색(full search)라 하며 가장 정확한 움직임 벡터를 찾아낼 수 있다. 그러나 검색 윈도우의 크기가 커질수록 많은 연산시간을 필요로 한다는 단점이 있다.

### III. 인체의 조인트의 자동 검출 및 추적 알고리즘

본 논문에서는 인체의 비율정보와 실루엣의 코너점을 이용하여 단일카메라에서 입력되는 영상에서 객체의 조인트(joint)를 자동으로 추출하여 조인트를 추적함으로써 객체를 추적하는 알고리즘을 제안한다. 먼저 전경과 배경을 이용하여 중심 객체의 실루엣을 추출하고 추출된 실루엣과 인체의 비율을 이용하여 조인트를 추출한다. 조인트의 자동 추출을 위해 사용되는 인체의 비율은 OpenCV에서 제공되는 얼굴 탐색기법을 이용하여

얼굴을 중심으로 얼굴의 길이와 넓이를 추출하여 실제 측정된 데이터를 고려하여 모델링하였다. 객체의 추적은 객체의 전 영역의 움직임을 추적하지 않고 추출된 조인트를 중심으로 블록매칭 기법을 적용하여 객체의 움직임을 추적한다.

#### 1. 실루엣 추출

실루엣 추출은 인체의 조인트를 추출하기 위한 전처리 단계로서 인체의 조인트 추출의 정확도를 결정하는 매우 중요한 단계로서 효율적인 인체 실루엣 추출을 위해 전경과 배경의 분리는 필수적이며 이를 이용한 객체의 분할은 매우 중요한 기술이다. 그러나 RGB 컬러 영상의 회색조 영상을 구하여 배경을 모델링한 후 차영상 기법을 적용할 경우 그림자 영역과 조명의 변화에 따라 배경과 전경의 분리가 어려워지며 정확한 객체의 검출이 불가능한 경우가 발생한다. [그림 3]의 (c)는 이와 같은 경우를 보여주는 예로 RGB 입력영상을 회색조 영상으로 변환하여 이진화 하였을 경우 [그림 3](c)에서 보듯이 조명과 그림자 등의 영향으로 부분적으로 배경과 전경의 분리가 되지 않는 영역이 발생한다.

따라서 제안방법은 입력의 회색조 영상의 차영상과 색상(hue) 영상의 차영상을 구하여 그 결과를 동시에 고려함으로써 이러한 단점을 개선한다.

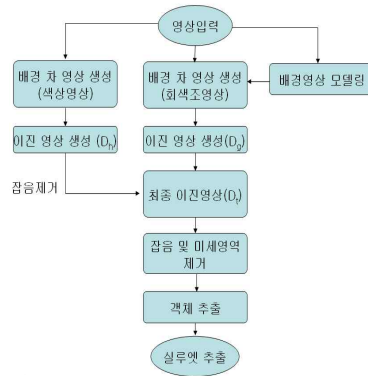


그림 2. 실루엣 추출 흐름도

[그림 2]는 제안방법의 실루엣 추출 흐름도로 먼저

카메라로 입력되는 24비트 컬러 영상  $I_C(x,y)$ 의 N 프레임의 평균으로 배경영상  $B(x,y)$ 을 모델링하며  $B(x,y)$ 와  $I_C(x,y)$ 를 256컬러의 회색조 영상으로 변환하여 회색조 영상의 차영상을 구한다. 구해진 차영상은 Otsu[13]의 방법으로 이진화하여 전경과 배경으로 구분된 이진영상 ( $D_g(x,y)$ )을 구한다. 또한 입력영상의 색상(hue) 영상을 구하여 배경영상과의 차영상을 구해 Otsu의 방법으로 이진화한다. 객체의 실루엣 추출을 위한 최종 이진화 영상은 회색조 영상과 색상 영상의 이진화 영상의 객체영역을 결합하여 얻는다. [그림 3](d)는 입력 영상의 색상 영상을 이용하여 이진화 한 결과 영상으로 객체 영역이외의 미세영역이 다수 존재한다. 따라서 최종 이진 영상을 얻기 위한 회색조 영상의 이진영상과 색상의 이진화 영상의 결합 전에 색상의 이진화 영상의 미세영역 및 잡음을 제거한다. 제안방법에서는 모폴로지컬 필터를 이용하여 미세영역 및 잡음을 제거하였다. [그림 3](e)는 제안방법을 적용하여 필터를 처리한 결과이다.

회색조의 이진영상과 색상의 이진영상의 결합은 다음의 수식을 이용하여 얻을 수 있다.

$$D_t(x,y) = \begin{cases} foreground, & D_g(x,y) = 1 \text{ 또는 } D_h(x,y) = 1 \\ background, & otherwise \end{cases} \quad (3)$$

여기서  $D_h(x,y)$ 는 색상 영상의 이진화 영상이다.

수식 (1)을 이용하여 얻은 이진영상에도 그림 3(f)에서 보여주듯이 객체 이외의 미세영역 등이 존재한다. 따라서 객체의 추출을 위하여 최종 이진화 영상에 모폴로지컬 필터 처리로 잡음을 제거하고 영역제한을 두어 중심객체 이외의 미세영역을 제거함으로 중심객체를 추출한다.

추출된 객체는 OpenCV에서 제공하는 외곽선 추적 기법을 이용하여 외곽선을 추출함으로 실루엣을 추출한다.

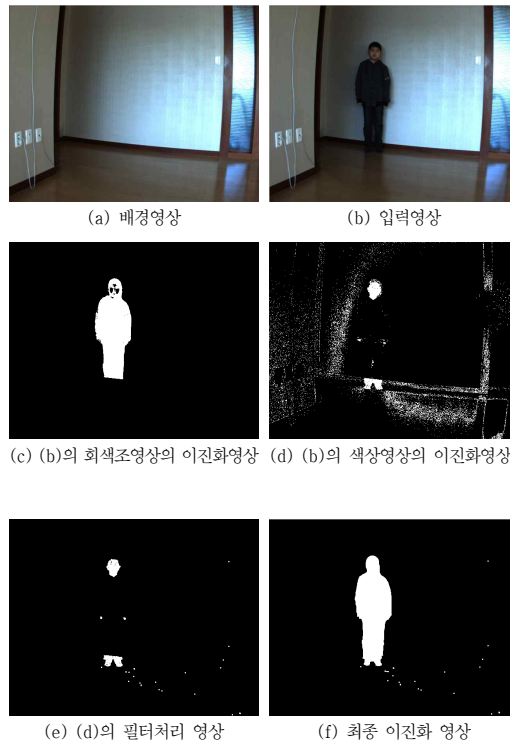


그림 3. 입력영상과 배경영상의 차영상을 이용한 이진화 결과

## 2. 인체 모델링과 조인트 검출

추출된 실루엣에서 인체의 조인트를 검출하기 위해 본 논문에서는 인체의 구조를 기반으로 [그림 4]와 같은 인체모델과 인체의 각 구조를 연결하는 16개의 조인트를 정의한다.

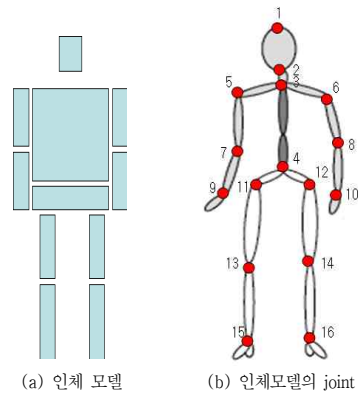


그림 4. 인체모델과 joint

각 조인트의 검출은 얼굴의 세로의 길이와 가로길이를 기준으로 각 인체 영역에 대한 비율을 구하여 이용한다. 각 비율을 도출하기 위하여 임의의 20명을 모델로 하여 각 신체의 치수를 측정하고 그 결과를 정리하였다. [표 1]은 키별 인체 영역의 평균을 구한 결과로 인체 영역의 숫자는 [그림 4]의 조인트 번호를 의미한다.

[표 1]을 고찰할 때 얼굴의 길이는 키별 데이터에 따라 변동이 없다. 따라서 비율 값을 산출하기 위한 기준으로 얼굴 길이를 사용하며 얼굴영역의 길이는 OpenCV에서 제공하는 얼굴 탐색 기법을 적용하여 구하였다. 인체 부위별 비율 값은 다음과 같이 구한다.

$$\alpha = \frac{R_m}{f_h} \quad (4)$$

여기서  $R_m$ 은 표 1의 영역별 측정치 이고  $f_h$ 는 얼굴의 길이이다.

표 1. 키별 인체영역의 치수 단위:cm

키 \ 인체영역	140-150	150-160	160-170
얼굴넓이	15	16	16
1-2(얼굴길이)	22	22	22
2-3	8	8	8
3-4	27	34	40
3-5(3-6)	17	19	21
5-7(6-8)	25	27	29
7-9(8-10)	23.5	23.5	23.5
4-11(4-12)	8	11	13
11-13(12-14)	39	41	42
13-15(14-16)	33	35	38

또한 키별 각 인체영역의 길이 비율이 다르므로 입력 영상의 객체영역 높이와 얼굴영역의 길이 및 실제 얼굴 길이를 이용하여 객체의 키를 구하여 조인트 위치 정보 계산 시 적용할 측정치를 결정한다. 다음은 객체의 실제 키를 산출하기 위한 수식이다.

$$t_h = \frac{t_r}{f_r} \times f_h \quad (5)$$

여기서  $t_h$ 는 객체의 실제 키,  $f_r$ ,  $t_r$ 은 영상에서 추출한 객체의 얼굴 높이와 객체의 높이이다.

[표 2]는  $\alpha$ 를 이용한 조인트 추출 방법으로 입력영상에서 얼굴을 인식하여 얼굴영역의 상위좌표와 하위좌

표를 기준으로 조인트를 추출한다.  $j\_N.x(y)$ 는 조인트 번호 N에서의  $x(y)$  위치이다.

인체의 치수는 각 사람마다 다르며 그 값의 평균이 [표 1]의 값이므로 조인트 추출 시 정확한 조인트의 위치를 검출하기 어렵다. 따라서 본 논문에서는 [표 1]의 측정치를 기준으로 [표 2]의 방법을 적용하여 1차적인 조인트를 검출하고 각 사람마다의 조인트 위치를 보완하기 위해 추출 객체의 인체 실루엣에서 코너점을 추출하고 추출된 코너점 중 1차 조인트와 가장 가까운 코너점을 선택하여 최종 조인트를 결정하는데 사용한다. [그림 5]는 1차 조인트3에서 가장 가까운 코너점을 추출하는 방법을 보인 예이다.

표 2. 1차 조인트 추출 방법

조인트번호	x위치	y 위치
1	얼굴영역의 상위 x좌표	얼굴영역의 상위 y좌표
2	얼굴영역의 하위 x좌표	얼굴영역의 상위 y좌표
3	$j\_2.x$	$j\_2.y + f_r \times \alpha$
4	$j\_3.x$	$j\_3.y + f_r \times \alpha$
5	$j\_3.x - f_r \times \alpha$	$j\_3.y$
6	$j\_3.x + f_r \times \alpha$	$j\_3.y$
7	$j\_5.x$	$j\_5.y + f_r \times \alpha$
8	$j\_6.x$	$j\_6.y + f_r \times \alpha$
9	$j\_7.x$	$j\_7.y + f_r \times \alpha$
10	$j\_8.x$	$j\_8.y + f_r \times \alpha$
11	$j\_4.x - \text{얼굴넓이} \times 0.5$	$j\_4.y + f_r \times \alpha$
12	$j\_4.x + \text{얼굴넓이} \times 0.5$	$j\_4.y + f_r \times \alpha$
13	$j\_11.x$	$j\_11.y + f_r \times \alpha$
14	$j\_12.x$	$j\_12.y + f_r \times \alpha$
15	$j\_13.x$	$j\_13.y + f_r \times \alpha$
16	$j\_14.x$	$j\_14.y + f_r \times \alpha$

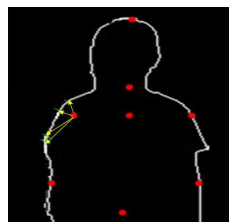


그림 5. 코너점 검출의 예

최종 조인트 검출시 조인트 1~4, 11, 12는 변경되지 않으며 조인트 5~16이 코너점을 이용하여 변경된다. 최종 조인트 추출 방법을 [표 3]에 나타냈다.  $c.x(y)$ 는 해당 조인트에서 가장 가까운 코너점의  $x(y)$  위치이다.

표 3. 코너점을 적용한 최종 조인트 추출 방법

joint번호	x위치	y 위치
5	$c.x$	$(j_{5.y}+c.y)/2$
6	$c.x$	$(j_{6.y}+c.y)/2$
7	$(j_{7.x}+c.x)/2$	$(j_{7.y}+c.y)/2$
8	$(j_{8.x}+c.x)/2$	$(j_{8.y}+c.y)/2$
9	$j_{9.x}- j_{9.x}-c.x /2$	$(j_{9.y}+c.y)/2$
10	$j_{10.x}+ j_{10.x}-c.x /2$	$(j_{10.y}+c.y)/2$
13	$(j_{4.x}+j_{13.x}+c.x)/3$	$(j_{13.y}+c.y)/2$
14	$(j_{4.x}+j_{14.x}+c.x)/3$	$(j_{14.y}+c.y)/2$
15	$(j_{15.x}+c.x)/2$	$c.y+(j_{13.y}-c.y)/2$
16	$(j_{16.x}+c.x)/2$	$c.y+(j_{14.y}-c.y)/2$

### 3. 블록매칭 기법을 이용한 조인트 추적 알고리즘

기존의 객체 추적방법은 중심객체의 추출 후 중심객체를 추적하기 위해 객체의 색상정보를 이용하거나 템플릿을 이용하는 방법 등을 사용한다. 본 논문에서는 객체의 위치를 조인트의 위치로 표현할 수 있으며 따라서 객체 전 영역에 대한 정보가 아니라 조인트를 중심으로 해서 지역정보의 움직임을 추적함으로써 객체를 추적하는 것이 가능하다. 따라서 MPEG등에서 사용되는 블록매칭 방법을 이용하여 객체를 추적하였다. 또한 카메라로 입력되는 영상에서 인체를 대상으로 하여 객체를 추적하는 경우 특별한 사건이 있을 경우를 제외하고는 인간의 행동은 급작스러운 변화가 나타나지 않으므로 인간의 신체 부위에서는 커다란 움직임의 변화가 발생하지 않는다. 따라서 본 논문에서는 처음 추출된 조인트의 위치를 중심으로 다음 프레임에서 조인트를 추적하였다. 기존의 블록매칭 방법은 객체 영역 전체를 추적하여야 하는 반면 제안 방법은 조인트를 중심으로 주변의 작은 블록만을 탐색함으로써 연산량을 줄였다.

블록은 각 조인트를 기준으로  $20 \times 20$ 의 크기로 하며 입력 영상의 탐색 윈도우는 수평/수직 방향에 대하여  $\pm 5$ 의 크기를 설정하여 블록매칭 기법을 적용한다. 이때

블록의 크기가 너무 크면 연산량이 많아지고 너무 작으면 실제적인 조인트 주변의 정보를 고려하지 못하여 잘못 추적할 수 있다. 따라서 블록의 크기를  $10 \times 10 \sim 30 \times 30$ 으로 가변하여 적용하여 연산량과 오차를 비교한 후 결정하였다. 또한 탐색윈도우는 인체의 움직임이 크지 않은 것을 고려하여 정하였다.

블록매칭 알고리즘 적용시 RGB 컬러의 입력 영상을 HSV영상으로 변환한 후 색상(hue)과 채도(saturation)의 값을 이용하여 오차가 가장 적은 점을 다음 조인트 위치로 선택한다.

[그림 6]은 처음 추출된 조인트를 중심으로 위치 정보를 이용하여 블록매칭 기법을 적용하는 것을 예로 보인 것으로 조인트를 중심으로  $20 \times 20$  크기의([그림 6]의 실선) 색상값을 중심으로  $\pm 5$ 의 윈도우([그림 6]의 점선)를 탐색하였다.

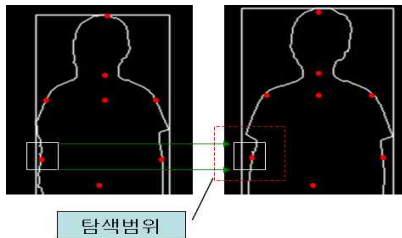


그림 6. 조인트 추적의 예

## IV. 실험 및 결과 분석

제안방법의 성능을 분석하기 위하여 실내 환경에서 실험하였으며 카메라를 이용하여 배경영상과 입력 영상을 실시간적으로 처리하였다. 제안방법은 단일 카메라의 입력으로 실루엣과 인체의 비율정보 등을 이용하여 자동으로 인체의 조인트를 추출하는 것을 목적으로 함으로 인체의 실루엣 추출을 위하여 단순한 배경에서 사람이 걸어가는 동작을 대상으로 실험하였고 시스템은 Intel cpu 2.0GHz, 1G RAM, VC++6.0과 Open CV를 이용하여 구현하였다. 입력되는 영상의 해상도는  $640 \times 480$  24bit 영상으로 초당 10프레임을 전송받는다.

1. 조인트 자동 추출 알고리즘

[그림 7]은 제안방법을 적용한 결과를 보여준다. [그림 7](c)는 제안 방법을 적용하여 그림 (b)의 인체의 실루엣을 추출한 결과이며 (d)는 16개의 조인트를 추출하여 나타낸 결과로 조인트를 추출하기 위하여 얼굴영역을 먼저 검출하므로 검출된 얼굴 영역을 나타내었다.

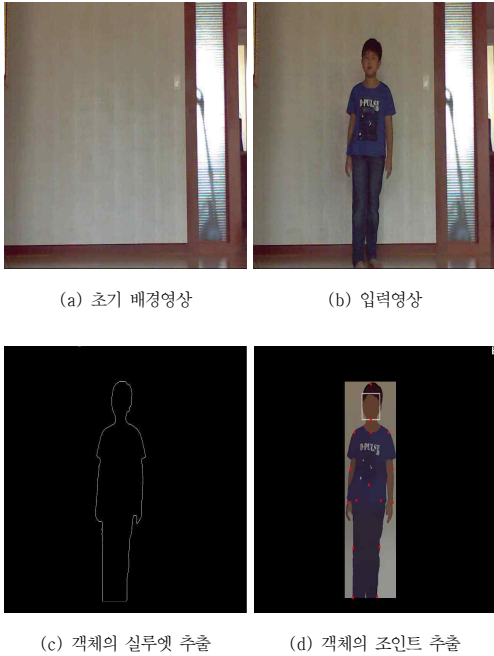
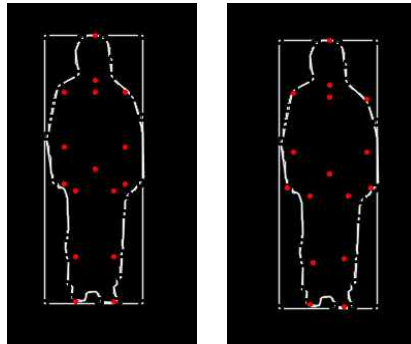


그림 7. 객체의 실루엣 추출 및 조인트 추출

[그림 8]은 [표 2]의 인체의 비율값을 적용하여 1차 조인트를 추출한 결과 영상과 [표 3]의 실루엣의 코너 점을 추출하여 최종 조인트를 추출한 결과 영상을 보여준다. 그림 (b)의 조인트 5~10과 조인트 13~16이 그림 (c)에서 변화가 있음을 알 수 있다. 이것은 실루엣 정보의 코너점을 고려함으로써 좀 더 정확한 조인트를 검출하고 있음을 보여준다.



(a) 입력영상



(b) 1차 조인트 (c) 최종 조인트

그림 8. 인체 비율과 실루엣 코너점을 이용한 조인트 추출 결과

2. 블록매칭을 이용한 조인트 추적 및 객체 추적

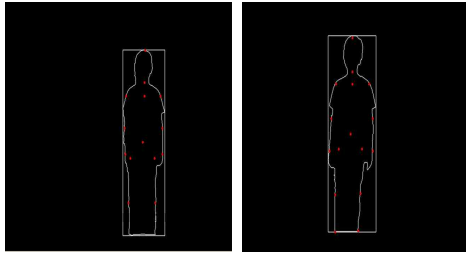
[그림 9]는 인체의 조인트 추적결과로 처음 추출된 조인트를 중심으로 다른 프레임에 블록매칭 알고리즘을 수행하여 (c)의 처음 조인트 추출된 결과에서 (d)로의 변화를 보여준다. [표 4]는 [그림 9](c)와 [그림 9](d)의 수평과 수직의 위치 변화값이다. 수평은 -17~-27, 수직은 -2~-12의 범위에서 조인트가 이동되었음을 보여준다.

표 4. 조인트의 수평/수직 위치변화값

joint	$\Delta x$	$\Delta y$	joint	$\Delta x$	$\Delta y$
1	-23	-10	9	-26	-5
2	-20	-10	10	-27	-5
3	-22	-12	11	-27	-8
4	-20	-8	12	-21	-9
5	-17	-12	13	-22	-6
6	-18	-11	14	-23	-10
7	-21	-9	15	-27	-2
8	-18	-9	16	-24	-3



(a) 추적 원본 영상 (b) 추적 대상 영상

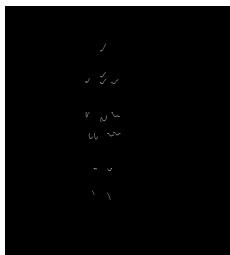


(c) (a)의 특징 추출영상 (d) (c)의 조인트 추적영상

그림 9. 객체의 조인트 추적영상



(a) 추적 원본 영상 (b) 추적 대상 영상



(c) 움직임 벡터

그림 10. 객체의 조인트 추적영상 및 움직임벡터 영상

[그림 10]은 객체의 처음 입력영상에서 24프레임 후의 영상에 대해 제안 알고리즘을 적용하여 추적한 영상

과 움직임벡터를 누적하여 얻은 결과 영상이다. 머리부분과 각 조인트의 변화를 알 수가 있다.

제안방법은 객체 추적을 위해 객체전체의 움직임을 추적하지 않고 조인트 위치의 움직임을 블록 매칭방법을 이용해 추적한다. 이것은 객체 전체에 블록 매칭방법을 적용했을 경우에 비해 연산량을 줄일 수 있다.  $M \times N$  크기의 객체를  $l \times m$ 의 블록으로 탐색 윈도우  $p \times q$  영역을 적용하여 탐색할 경우 기존의 방법은  $M \times N \times p \times q$ 의 비교횟수와 덧셈이 필요하며 제안방법은  $l \times m \times p \times q \times 16$ 의 비교횟수와 덧셈이 필요하다. 따라서 비교횟수와 덧셈에서  $\frac{M \times N}{l \times m \times 16}$ 의 연산량을 줄일 수 있다. 만약 객체의 크기가  $400 \times 200$ 인 경우  $20 \times 20$  블록을  $5 \times 5$ 의 탐색영역으로 하여 적용할 경우 기존의 블록매칭방법과 제안방법의 연산량을 비교하면 다음과 같다.

표 5. 기본블록매칭방법과 제안방법의 연산량

방법 \ 계산횟수	1프레임		10프레임	
	비교	덧셈	비교	덧셈
기본블록매칭	$32 \times 10^6 + 54$	$32 \times 10^6$	$32 \times 10^7 + 540$	$32 \times 10^7$
제안방법	$64 \times 10^4 + 16$	$64 \times 10^4$	$64 \times 10^5 + 160$	$64 \times 10^7$

[표 5]에서 보여 주듯이 제안 방법은 기존의 블록 매칭 방법에 비해 연산량을 절감하며 조인트를 추적할 수 있다.

제안방법을 다양한 동영상에 적용한 결과 실루엣 추출 및 조인트 검출 등에서 97%의 완성도를 보였으며 객체 및 조인트 추적은 95% 정도의 추적율을 얻었다.

오차의 발생 원인은 실루엣 추출 및 조인트 검출에서는 조명의 변화 및 그림자 정보등으로 인한 실루엣 추출의 실패가 주 원인이 되었으며 객체 및 조인트의 추적에서는 인체의 갑작스런 동작으로 블록매칭 알고리즘의 탐색 범위를 벗어나 조인트가 이동 하였거나 겹쳐진 영역에서 조인트를 추출하지 못하는 경우가 발생했을 때이다.



## V. 결론

본 논문에서는 단일 카메라로부터 실시간 입력되는 동영상에서 인체의 실루엣과 조인트를 자동으로 검출하고 추적하는 시스템을 제안한다. 제안방법은 마커나 센서기반 실루엣을 추출하는 다른 방법의 단점을 개선하여 객체의 실루엣 정보와 신체의 비율을 이용하여 카메라로 입력되는 동영상에서 자동으로 인체의 조인트를 추출한다. 또한 카메라의 입력을 이용하여 실루엣을 추출할 경우 다수의 카메라를 필요로하거나 주변 정보를 필요로 하는 기존의 방법과 다르게 현대의 카메라를 통해 획득한 정보로만 인체의 실루엣과 조인트를 추출한다. 또한 조인트의 자동 추출을 위해 제안방법은 얼굴길이와 넓이를 기준으로 인체의 비율을 모델링하였다. 객체의 추적시 처음 영상에서 추출된 조인트를 이용하여 지역적으로 블록매칭기법을 사용하여 조인트의 움직임 추적을 줄였고 객체 추적에서도 좋은 효율을 보였다.

제안방법은 3D 애니메이션, 인체의 행동인식, 보행인식 등의 다양한 분야에 응용이 가능하다. 그러나 조명 변화나 그림자 영역으로 인한 잡음과 검침영역의 발생에서 조인트 추출 및 추적이 강건하지 못하며 이를 개선한 연구가 필요하다.

## 참 고 문 헌

- [1] K. M. Lee and W. N. Street, "Model-based detection, segmentation and classification using on-line shape learning," *Machine vision and application*, Vol.13, No.4, pp.222-333, 2003.
- [2] G. Mori and J. Malik, "Estimating Human Body configurations using Shape Context Matching," in *Processings of ECCV*, pp.666-680, 2002.
- [3] C. Wren, A. Azarbajani, T. Darrell, and A. Pentland, "Pfinder: Real-time tracking of the human body," *IEEE trans. on PAMI*, Vol.19, No.7, pp.780-785, 1997.
- [4] L. Zhang, B. Wu, and B. Nevatia, "Detection and Tracking of Multiple Humans with Extensive Pose Articulation," *computer Vision, ICCV 2007*, pp.1-8, 2007.
- [5] T. E. de Campos and D. W. Murray, "Regression-based Hand Pose Estimation from Multiple Cameras CVPR 2006, Vol.1, pp.782-789.
- [6] Q. Delamarre and O. Faugeras, "3D articulated models and multi-view tracking with silhouettes," *Proc. ICCV*, pp.716-721, 1999(9).
- [7] G. Mori and J. Malik, "Estimating Human Body configurations using shape Context Matching," in *Proceedings of ECCV*, pp.666-680, 2002.
- [8] C. Wren, A. Azar bayejani, T. Darrell, and A. Pentland, "Pfinder: Real-time tracking of human body," *IEEE trans. on PAMI*, Vol.19, No.7, pp.780-785, 1997.
- [9] T. Horptasert, I. Haritaoglu, C. Wren, D. Harwood, L. Davis, and A. Pentland, "Real time 3D motion capture," in *Processings of Workshop on perceptual user interface*, 1998.
- [10] S. Iwasawa, J. Ohya, K. Takahashi, T. Sakaguchi, S. Kawato, K. Ebihara, and S. Morishima, "Real-time 3D estimation of human body postures from triocular images," in *Processings of Workshop on modeling people*, pp.3-10, 1999.
- [11] Andrew Hill, Chris J. Taylor, and Alan D. Brett, "A Framework for Automatic Landmark Identification Using a New Method of Nonrigid Correspondence," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.22, No.3, 2000(3).
- [12] Pengfei Zhu and Paul M.Chrlan, "On Critical Point Detection of Digital Shapes," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol.17, No.8, 1995(8).
- [13] N. Otsu, "A threshold selection method from

gray level histograms," IEEE Transactions on Systems, Man and Cybernetics, Vol.9, No.1, pp.62-66, 1979.

저 자 소 개

곽 내 정(Nae-Joung Kwak)

중신회원



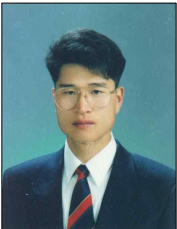
- 1993년 2월 : 충북대학교 정보통신공학과(공학사)
- 1995년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2005년 2월 : 충북대학교 정보통신공학과(공학박사)

- 2005년 3월 ~ 2006년 2월 : 목원대학교 정보통신공학과 프로그래밍 전문강사
- 2006년 3월 ~ 2009년 2월 : 목원대학교 정보통신공학과 강의전임
- 2009년 3월 ~ 현재 : 목원대학교, 한밭대학교 시간강사

<관심분야> : 멀티미디어 정보처리, 멀티미디어 통신, 컴퓨터 비전

송 특 섭(Teuk-Seob Song)

정회원



- 2001년 2월 : 연세대학교 수학과(이학박사)
- 2006년 2월 : 연세대학교 컴퓨터과학과(공학박사)
- 2006년 3월 ~ 현재 : 목원대학교 컴퓨터공학과 교수

<관심분야> : 웹환경 어노테이션, 가상환경, 사이버클래스, 스케치기반 인터페이스, 시멘틱 웹