
인체의 조인트와 움직임 정보를 이용한 인간의 행동패턴 인식

곽내정* · 송특섭**

Human Activity Pattern Recognition Using Motion Information and Joints
of Human Body

Nae-joung Kwak* · Teuk-seob Song**

이 논문은 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구임
(No. 2012-0002854)

요 약

본 논문에서는 인체의 조인트와 조인트의 움직임 정보를 이용하여 인간의 행동을 인식하는 알고리즘을 제안한다. 제안방법은 입력되는 비디오에서 객체를 추출하고 인체의 비율정보를 이용하여 조인트를 자동추출하며 각 조인트에 블록매칭 기법을 적용하여 조인트의 움직임 정보를 얻는다. 제안방법은 움직임이 있는 조인트, 조인트의 움직임의 방향벡터와 조인트의 x와 y좌표의 증가(+)와 감소(-)를 부호로 나타낸 것을 행동 인식을 위한 기본 파라미터로 사용한다. 제안된 방법은 웹카메라에서 입력되는 영상에서 8가지 행동에 대해 실험하였으며 인간의 행동 인식률에 있어 좋은 결과를 보였다.

ABSTRACT

In this paper, we propose an algorithm that recognizes human activity patterns using the human body's joints and the information of the joints. The proposed method extracts the object from inputted video, automatically extracts joints using the ratio of the human body, applies block-matching algorithm for each joint and gets the motion information of joints. The proposed method uses the joints to move, the directional vector of motions of joints, and the sign to represent the increase or decrease of x and y coordinates of joints as basic parameters for human recognition of activity. The proposed method was tested for 8 human activities of inputted video from a web camera and had the good result for the ration of recognition of the human activities.

키워드

행동 인식, 조인트, 움직임 정보

Key word

activity recognition, joint, motion information

* 정회원 : 목원대학교

** 정회원 : 목원대학교 (교신저자 teukseob@mokwon.ac.kr)

접수일자 : 2012. 02. 09

심사완료일자 : 2012. 02. 28

Open Access <http://dx.doi.org/10.6109/jkiice.2012.16.6.1179>

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

I. 서 론

최근 인간과 컴퓨터의 상호작용에 관한 관심이 대두되면서 유비쿼터스 컴퓨팅, 감정 컴퓨팅 등 인간의 환경에 지능적인 요소가 반영된 형태의 연구가 활발히 진행되고 있다. 이에 따라 제스처 인식, 수화인식, 보행패턴 인식, 행위인식 등 인간의 행동에 대한 인식은 많은 주목을 받으며 다양한 분야에 응용이 되는 중요성이 강조되고 있다. 특히 독거 노인의 증가, 슈퍼마켓이나 편의점의 범죄행위, 현금인출기 등에서의 범죄행위 및 거리에서의 폭행 등 인간의 행위 인식의 자동 인식은 다양한 범주에서 응용이 가능하다. 이러한 인간 행위의 자동인식은 영상 감시 시스템의 도입으로 가능하며 현재 지능형 감시 시스템으로 연구가 진행되고 있다[1].

영상 감시 시스템은 카메라로 입력되는 영상을 컴퓨터로 전송하여 처리하는 시스템으로 영상분석을 통해 인간의 행동 패턴을 분석해 다양한 정보를 얻어 다양한 분야에 적용이 가능하다[2,3]. 이와 같이 컴퓨터를 이용하여 영상을 분석해 인간의 행동을 인식 및 분류하는 것은 인체의 각 부위가 시간의 흐름에 따라 어떤 모습으로 변화하는지를 찾아내어 분석함으로써 가능하다. 이것은 특징을 추출하는 과정과 추출된 특징을 이용하여 행동을 분류하는 과정으로 나뉠 수 있으며 인간의 행동에 대한 특징을 추출하는 방법에 따라 인식은 인체에 센서를 부착하여 센서 정보를 분석하는 방법[4-7]과 비디오 카메라를 이용하여 입력된 영상을 이용하여 인체의 특징을 검출한 후 검출된 특징을 이용하여 행동 인식을 하는 방법[8-11]으로 크게 분류된다. 센서 정보를 이용하는 방법은 인체의 특정 부위에 센서나 마커를 부착하여 이를 응용하여 인체의 실루엣이나 윤곽선 및 다른 특징들을 추출하는 방법[6-8]으로 인체의 모델과 추적에 머리와 손 발등으로 제한되며 마커 등의 고가의 장비가 필요하다. 또한 기존의 방법들은 행위 인식 및 자세 추정 등을 위해 인체의 정보로부터 추출된 특징을 학습하고 학습된 데이터를 이용하여 입력되는 영상데이터 혹은 센서 데이터로부터 행위를 인식하거나 자세를 추정한다. 따라서 많은 량의 학습 데이터와 복잡한 학습 알고리즘이 필요하다.

본 논문에서는 입력되는 카메라 영상에서 배경과의

차영상을 이용하여 객체를 분리하고 객체의 실루엣을 추출한 후 인체의 비율정보를 이용하여 조인트를 자동 추출하고 추출된 조인트의 추적정보를 분류하여 인간의 행동패턴을 분류하는 알고리즘을 제안한다. 제안 방법은 추출된 조인트의 추적 정보를 인간의 행동 변화를 일으키는 주된 조인트를 중심으로 움직임의 방향벡터를 이용하여 행동패턴을 분류한다.

본 논문은 2장에서 객체의 실루엣 추출 알고리즘과 조인트 추출 알고리즘에 대해 기술하며 추출된 조인트의 추적으로 방향벡터에 의한 인간의 행동 패턴 분류 방법에 대하여 기술한다. 3장에서는 제안방법의 실험결과를 기술하며 4장에서는 결론 및 향후 과제에 대해 기술한다.

II. 본 론

본 논문에서는 단일 카메라에서 입력되는 입력영상과 배경정보를 이용하여 객체를 추출하고 추출된 객체의 조인트를 자동으로 추출하고 추출된 조인트 중 행동패턴 분류 관련된 조인트를 추적하고 그 움직임 벡터의 방향을 분석함으로써 인간의 행동 패턴을 분류하는 알고리즘을 제안한다. 그림 1은 제안방법의 흐름도이다.

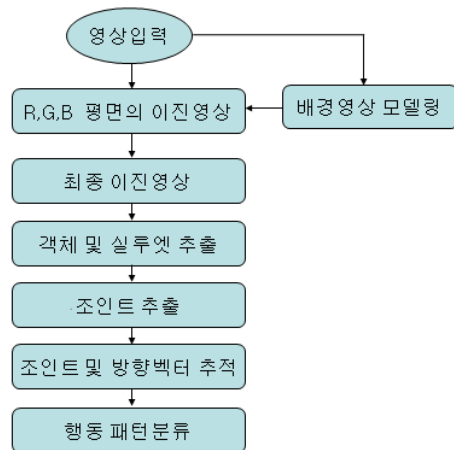


그림 1. 제안 알고리즘
Fig. 1 The proposed method

2.1. 배경모델링과 객체추출

객체 추출은 인체의 조인트를 추출하기 위한 전처리 단계로서 인체의 조인트 추출의 정확도를 결정하는 매우 중요한 단계로서 효율적인 객체 추출을 위해 객체와 배경 분리는 필수적이다.

본 논문에서는 배경과 객체의 분리를 위해 배경을 모델링 하고 임계값을 이용하여 객체와 배경을 분리한다. 배경모델링과 임계값의 설정은 다음과 같다.

1. N 프레임의 배경영상의 평균(B_{avg})을 구한다.
2. N 프레임과 평균 배경 영상의 차영상의 누적의 평균(BD_{avg})을 구한다.
3. 누적 평균의 절대값을 이용하여 배경의 상한(th_u)값과 하한값(th_l)을 정한다. 이때 th_u 와 th_l 는 다음과 같이 구한다.

$$th_u = B_{avg} + \alpha * BD_{avg} \tag{1}$$

$$th_l = B_{avg} - \beta * BD_{avg} \tag{2}$$

임계값을 이용하여 카메라로 입력되는 24비트 컬러 영상을 R, G, B 평면으로 분리 후 (1)과 식(2)의 임계값 사이의 값은 배경으로 그 외 영역은 객체로 분리한다.

$$O(x,y) = \begin{cases} object, & th_l \leq I(i,j) \leq th_u \\ background, & otherwise \end{cases} \tag{3}$$

여기서 $I(i,j)$ 는 i,j 위치에서 입력영상의 화소이고 object(객체)는 흰색(255)으로 background(배경)는 검은색(0)으로 하여 이진화한다. 이진화된 R, G, B 평면의 결과를 합하여 최종 결과 영상을 얻는다. 이때 임계값은 다수의 입력영상을 이용하여 실험에 의해 정하였다.

이진화 한 결과 영상은 객체 영역이외의 미세영역이 다수 존재한다. 따라서 모폴로지컬 필터를 이용하여 미세영역 및 잡음을 제거하고 영역제한을 두어 중심객체 이외의 미세영역을 제거함으로 중심객체를 추출한다.

2.2. 실루엣 추출과 joint 추출[12]

추출된 객체로부터 실루엣 추출 및 조인트 추출은 논문 [12]에서 적용한 방법을 이용한다. 먼저 OpenCV에서

제공하는 외곽선 추적 기법을 이용하여 외곽선을 추출함으로 실루엣을 추출하며 그림 2의 인체모델과 인체의 각 구조를 연결하는 16개의 조인트를 구한다.

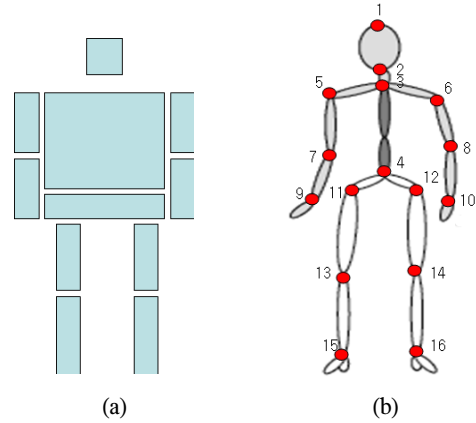


그림 2. 인체모델과 joint
 (a) 인체 모델 (b) 인체모델의 joints
 Fig. 2 Human body's model and joints
 (a) Human body's model
 (b) Joints of human body's model

각 조인트의 검출은 얼굴의 세로의 길이와 가로길이를 기준으로 [12]의 각 인체 영역에 대한 비율을 이용한다. 얼굴영역의 길이는 OpenCV에서 제공하는 얼굴 탐색 기법을 적용하여 구하였다. 인체의 치수는 각 사람마다 다르며 조인트 추출 시 정확한 조인트의 위치를 검출하기 어렵다. 따라서 인체치수의 비율을 이용하여 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.

2.3. 조인트의 움직임 정보를 이용한 행동 패턴 인식

인간의 다양한 행위 중 본 논문에서는 의자에 앉기, 한팔 들어올리기(오른쪽, 왼쪽), 양팔 들어올리기, 한다리 들어올리기(오른쪽, 왼쪽), 쪼그리기, 걷기 등 8가지의 행동을 인식한다. 16개의 조인트 중 각 동작에 따라 움직임이 발생하는 조인트는 소수이며 또한 각 조인트의 움직임 방향도 동작에 따라 다르다. 따라서 본 논문에서는 행동 인식을 위한 파라미터를 다음과 같이 정의

한다.

i) 변화 조인트

- 인간의 행동에 따라 움직임의 변화가 크게 발생하며 다른 행동과 구별이 가능한 조인트

ii) 조인트 움직임의 방향벡터(DV)

- 움직임 벡터가 변화해 가는 방향을 표현

iii) 조인트 움직임 벡터의 부호($\Delta x/\Delta y$)

- 조인트의 움직임을 현재 위치를 중심으로 x와 y좌표의 위치 증감을 +와 -로 표현한 값

변화 조인트는 걷기, 팔의 움직임, 다리의 움직임, 앉기 등 1차적으로 인체의 어느 부위에서 움직임 발생하는지를 분류할 수 있다. 움직임 벡터의 부호와 방향벡터는 정확한 동작을 구분하기 위한 것이며 움직임 방향벡터를 표현하기 위해 조인트의 움직임 영역을 그림 3과 같이 영역을 나누어 12개로 표현하였다.

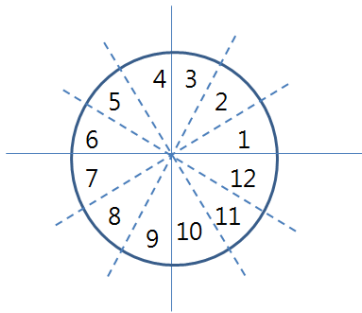


그림 3. 움직임벡터의 방향벡터
Fig. 3 Directive vector of motion vector

표 1은 행동 인식을 위해 본 논문에서 제안한 분류기준을 적용하는 유형을 정리한 것으로 ‘*’는 좌표값이 그다지 많은 변화가 있지 않고 행동 인식에 관계가 없음을 의미한다.

본 논문에서는 각 조인트의 움직임 벡터의 방향과 부호를 추출하기 위해 조인트의 위치를 중심으로 블록매칭 방법을 적용한다. 움직임 추적을 위한 탐색 블록은 각 조인트를 기준으로 20×20의 크기로 하며 입력 영상의 탐색 윈도우는 수평/수직 방향에 대하여 ±5의 크기를 설정하여 블록매칭 기법을 적용한다.

표 1. 행동인식을 위한 파라미터
Table. 1 The parameters for activity recognition

행동	조인트	Δx	Δy	DV
걷기	1	+/-	*	6 또는 7 1 또는 12
	2	+/-	*	
오른팔 들기	7	-	-	6 또는 5
	9	-	-	
왼팔 들기	8	+	-	1 또는 2
	10	+	-	
양팔 들기	7	-	-	6 또는 5
	9	-	-	
	8	+	-	1 또는 2
	10	+	-	
오른다리 들기	13	-	-	6 또는 5
	15	-	-	
왼다리 들기	14	+	-	1 또는 2
	16	+	-	
앉기	5	*	+	9 또는 10
	6	*	+	
	13	*	*	변화없음
	14	*	*	
꾸그리기	5	*	+	9 또는 10
	6	*	+	
	13	*	+	9 또는 10
	14	*	+	

III. 실험 및 결과 분석

제안방법의 성능을 분석하기 위하여 카메라를 이용하여 배경영상과 입력 영상을 실시간적으로 처리하였다. 실험영상은 한 사람당 8가지의 행동을 하고 있는 영상으로 총 4명의 사람의 영상을 이용하였으며 서로 다른 두 실내 환경에서 배경을 달리하여 구성하였다. 시스템은 Intel cpu 2.0GHz, 1G RAM, 비주얼 스튜디오 2008과 OpenCV 2.1 을 이용하여 구현하였다. 입력되는 영상의 해상도는 640×480 24bit 영상으로 초당 15 프레임을 전송받는다. 그림 4는 입력영상의 예이다.

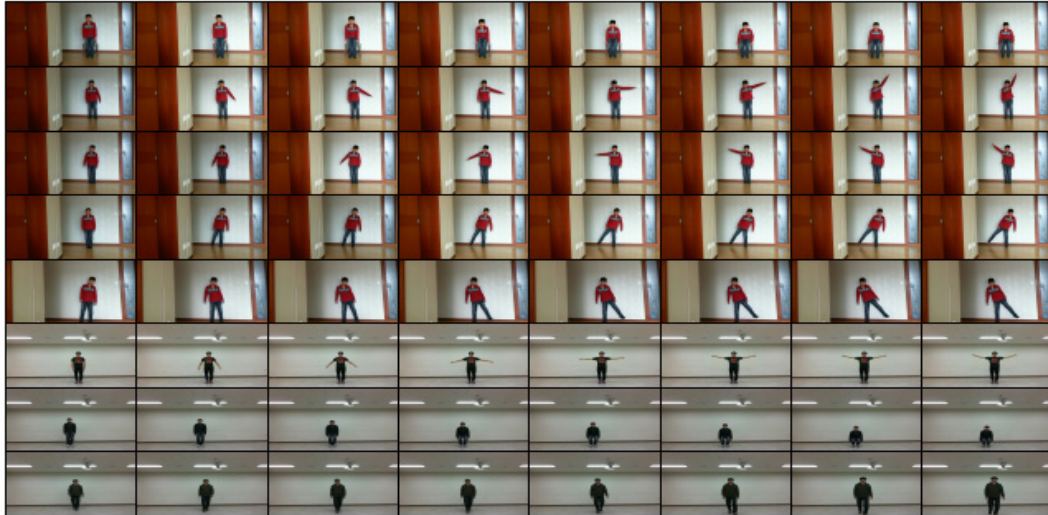


그림 4. 실험영상의 예
Fig. 4 The example of test video

그림 5는 제안방법을 실험 영상에 적용한 결과를 단계별로 나타낸 것이다. 그림 (a)의 배경을 모델링하고 (b)를 초기 입력영상으로 하여 처리한 결과를 제안 방법을 적용하여 이진영상화 한 결과를 (c)에 보인다, (d)는 (c)의 영상에서 중심객체를 추출한 후 실루엣을 추출하고 코너점을 추출한 결과, (e)는 얼굴을 검출한 결

과영상이며 (f)는 조인트를 추출한 결과영상이다. 그리고 (g)는 실험대상이 양팔을 들어올린 것을 제안방법을 이용하여 움직임을 추적하고 얼굴 영역과 움직임 영역(양 팔)의 조인트를 추적하여 표현한 영상이며, (h)는 초기 입력 영상 (b)부터 최종 조인트 추적영상 (g)까지 각 조인트의 움직임의 변화를 추적한 결과를 나타낸 영

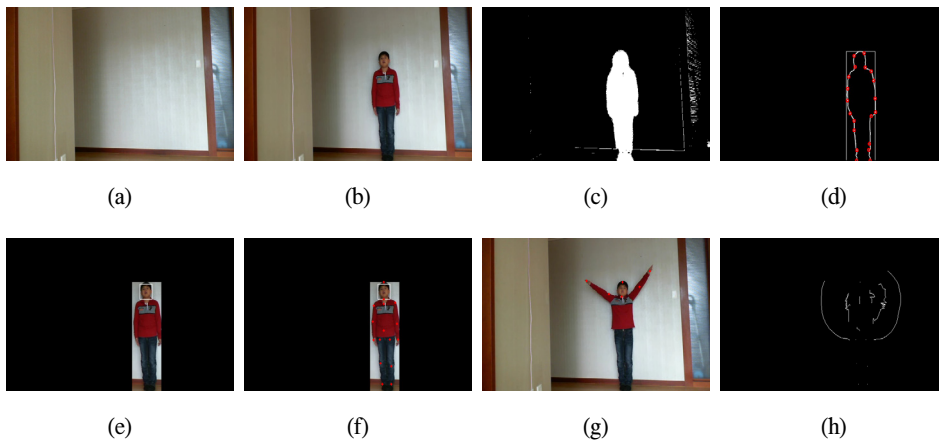


그림 5. 제안 방법을 적용한 단계별 영상. (a) 배경 (b) 초기 입력영상 (c) 이진영상 (d) 코너점 검출 (e) 얼굴 검출 (f) 조인트 추출 (g) 최종조인트영상 (h)움직임 영상
Fig. 5 Each stage image by the proposed method. (a) background (b) initial input image (c) binary image (d) corner extraction (e) face detection (f) joints extraction (g) final joint image (h) motion image

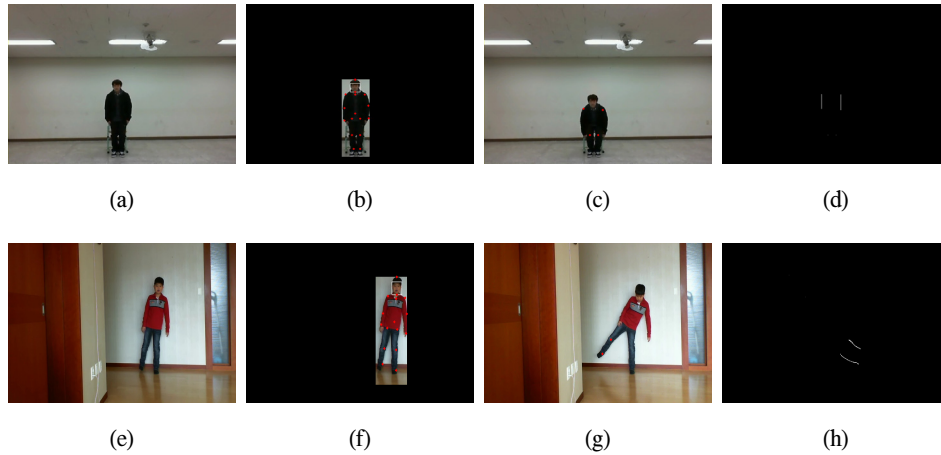


그림 6. 행동인식의 결과 (앉기와 다리들기). (a) 입력영상 (b) 조인트 영상 (c) 최종 조인트영상 (d) 움직임 영상 (e) 입력영상 (f) 조인트영상 (g) 최종조인트영상 (h)움직임 영상
 Fig. 6 The results of activity recognition (sitting and raising a leg). (a) input image (b) joint image (c) final joint image (d) motion image (e) input image (f) joint image (g) final joint image (h) motion image

상으로 양팔의 움직임이 다른 조인트에 비해 활발히 나타남이 보인다.

표 2. 입력영상에 따른 행동 인식률
 Table. 2 The ratio according to input images

행동	영상1	영상2	영상3	영상4	평균
걷기	300	300	300	300	100%
오른팔 들기	300	300	300	300	100%
왼팔 들기	300	300	300	300	100%
양팔 들기	300	300	300	300	100%
오른다리들기	300	300	300	287	98.9%
왼다리 들기	300	300	300	276	98%
앉기	300	300	300	300	100%
쭈그리기	300	300	300	300	100%

그림 6은 8가지 행동 중 앉기와 한 다리 들기에 대해 제안방법을 적용한 결과를 보여준다. 각 행동에 대해 변화 조인트를 추적하고 움직임 벡터의 움직임 변화를 잘 추적함을 보여준다.

표 2는 제안된 방법을 실험영상에 적용하여 행동을 인식한 인식률을 분석한 결과이다. 각 입력 영상은 300 프레임을 사용하였다. 모든 영상에 대해 제안방법은 100%의 인식률을 보이지만 영상4의 다리 들어올리기의 경우에는 오분류가 존재한다. 이것은 입력영상에서 다리 영역에 그림자 등으로 인하여 조인트 추출 추적시 오분류가 발생한 경우이다.

IV. 결 론

본 논문에서는 인체의 조인트와 조인트의 움직임 벡터를 이용하여 인간의 행동을 인식하는 알고리즘을 제안한다. 제안방법은 웹카메라로 입력되는 비디오에서 객체를 추출하고 추출된 객체의 조인트를 검출한다. 검출된 조인트의 움직임 벡터를 추적하여 움직임 벡터의 정보를 이용하여 인간의 행동을 인식한다. 제안된 방법은 인간의 행동 중 8가지의 정의된 행동에 대해 인식할 수 있으며 인식에 좋은 결과를 보였다.

제안 방법은 유비쿼터스 환경하에서 U-health 및 감시 카메라 등 다양한 응용에 적용이 가능하다. 그러나 조명 등에 의하여 그림자영역이 존재할 경우 인식률이 저하

되며 이를 개선할 수 있는 연구가 필요하다. 또한 복잡한 배경에서 여러 사람이 존재하는 경우 행동인식을 위한 객체 분류 등에 관련된 연구가 필요하다.

감사의 글

본 연구는 2012년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구임 (No. 2012-0002854)

참고문헌

[1] R.T. Collins, AFujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L. Wixson, "A System for Video Surveillance and Monitoring," *Technical Report CMU-RI-TR-00-12*, Carnegie Mellon University, 2000

[2] G. Gasser, N. Bird, O. Masoud, and N. Papanikolopoulos, "Human Activities Monitoring at Bus Stops," *Proceedings of the IEEE International Conference on Robotics & Automation*, '9. 90-95, 2004

[3] Ji Tao and Yap-Peng Tan, "A Probabilistic Reasoning Approach to Closed-Room People Monitoring," *IEEE ISCAS*, pp. II-185-188, 2004

[4] D.H. Wilson, A.C.Long, and C. Atkeson, "A Context-Aware Recognition Survey for Data Collection Using Ubiquitous Sensors in the Home," *In Proceedings of CHI 2005: Late Breaking Results*, pp. 1865-1868, Portland, OR, April 2005.

[5] E. Murphy-Chutorian and M. Trivedi, "Head Pose Estimation in Computer Vision: A survey," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 4, pp.607-626, April 2009.

[6] 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.

[7] Pengfei Zhu and Paul M.Chrlian, "On Critical Point Detection of Digital Shapes," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol.17, no.8, August 1995.

[8] A. Utsumi, H. Yang and J. Ohya, "Adaptive human motion tracking using non-symchrous multiple viewpoint observations," *Proc. IEEE 15th International Conf. on Pattern Recognition*, vol. 4, pp.607-610, 2000

[9] S. Iwasawa, J. Ohya, K. Takahashi, T. Sakaguchi, S. Kawato, K. Ebihara and S. Morishima, "Real-time 3D extimation of human body postures from triocular images," *in Processings of Workshop on modeling people*, pp.3-10, 1999

[10] T. E. de Campos, D. W. Murray, "Regression-based Hand Pose Estimation from Multiple Cameras," *CVPR 2006*, vol. 1, pp. 782-789

[11] Q. Delamarre and O. Faugeras, "3D articulated models and multi-view tracking with silhouettes," *Proc. ICCV*, pp. 716-721, Sep. 1999

[12] 광내정, 송특섭, "인체 비율을 이용한 인체의 조인트 자동 검출 및 객체 추적 알고리즘," 한국 콘텐츠학회 논문지, 제 11권 4호, 2011

저자소개

광내정(Nae-Joung Kwak)



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

2005년 2월 : 충북대학교 정보통신공학과 (공학박사)
2005년 3월 ~ 2006년 2월 : 목원대학교 정보통신 공학과
프로그래밍 전문강사
2006년 3월 ~ 2009년 2월 : 목원대학교 정보통신 공학과
강의전임
2009년 3월 ~ 현재 : 목원대학교, 한밭대학교
충남대학교 충북대학교 시간강사
※ 관심분야 : 멀티미디어 정보처리, 멀티미디어 통신,
컴퓨터 비전



송특섭(Teuk-Seob Song)

2001년 2월 : 연세대학교 수학과
(이학박사)

2006년 2월 : 연세대학교
컴퓨터과학과 (공학박사)

2006년 3월 ~ 현재 : 목원대학교 컴퓨터공학과 교수

※ 관심분야 : 웹환경 어노테이션, 가상환경, 사이버
클래스, 스케치기반 인터페이스, 시멘틱 웹