

모션 기반의 검색을 사용한 동적인 사람 자세 추적

정도준^{1*}, 윤정오¹
¹경운대학교 정보통신공학과

Dynamic Human Pose Tracking using Motion-based Search

Do-Joon Jung^{1*} and Jeong-Oh Yoon¹

¹Department of Information & Communication Engineering, Kyungwoon University

요약 본 논문은 단안 카메라로부터 입력된 영상에서 모션 기반의 검색을 사용한 동적인 사람 자세 추적 방법을 제안한다. 제안된 방법은 3차원 공간에서 하나의 사람 자세 후보를 생성하고, 생성된 자세 후보를 2차원 이미지 공간으로 투영하여, 투영된 사람 자세 후보와 입력 이미지와의 특징 값 유사성을 비교한다. 이 과정을 정해진 조건을 만족할 때까지 반복하여 이미지와의 유사성과, 신체 부분간 연결성이 가장 좋은 3차원 자세를 추정한다. 제안된 방법에서는 입력 이미지에 적합한 3차원 자세를 검색할 때, 2차원 영상에서 추정된 신체 각 부분들의 모션 정보를 사용해 검색 공간을 정하고 정해진 검색 공간에서 탐색하여 사람의 자세를 추정한다. 2차원 이미지 모션은 비교적 높은 제약이 있어서 검색 공간을 의미있게 줄일 수 있다. 이 방법은 모션 추정이 검색 공간을 효율적으로 할당 해주고, 자세 추적이 여러 가지 다양한 모션에 적용할 수 있다는 장점을 가진다.

Abstract This paper proposes a dynamic human pose tracking method using motion-based search strategy from an image sequence obtained from a monocular camera. The proposed method compares the image features between 3D human model projections and real input images. The method repeats the process until predefined criteria and then estimates 3D human pose that generates the best match. When searching for the best matching configuration with respect to the input image, the search region is determined from the estimated 2D image motion and then search is performed randomly for the body configuration conducted within that search region. As the 2D image motion is highly constrained, this significantly reduces the dimensionality of the feasible space. This strategy have two advantages: the motion estimation leads to an efficient allocation of the search space, and the pose estimation method is adaptive to various kinds of motion.

Key Words : Human dynamics, Pose tracking, Motion-based search, Markov network

1. 서론

사람과 컴퓨팅 기기와의 통신을 위한 자연스러운 상호작용 방법을 개발하는 것은 중요한 일이며, 상호작용의 주체인 사람의 자세와 움직임을 알아내고 분석하는 일이 다양한 방법으로 기술 개발 되고 있다[1]. 영상 기반으로 사람의 자세와 움직임을 추정하는 것은 단순히 사람의 동작을 파악하는 단계를 넘어 인간 컴퓨터 상호작용[2], 컴퓨터 애니메이션[3], 그리고 스마트룸[4] 등을 포함한

다양한 분야에서 더욱 확장된 형태의 응용을 기대할 수 있다. 하지만 기술 개발에 있어서, 사람의 신체가 관절로 연결되어 움직이기 때문에 형상에 모호성이 많고, 3차원의 사람이 2차원 영상으로 투영되어 깊이 정보가 손실되어 나타나므로 많은 어려움을 내포하고 있다. 결국, 영상 기반으로 사람의 자세를 추정하는 것은 깊이 정보가 손실된 영상에서 최적의 자세를 찾는 문제 즉, 검색 문제가 된다.

사람 자세 추적 기술은 움직임을 있는 사람의 자세 즉,

*교신저자 : 정도준(djjung@ikw.ac.kr)

접수일 10년 06월 09일

수정일 10년 06월 23일

게재확정일 10년 07월 06일

신체의 형상을 정확하게 찾아내는 것이다 [5]. 사람의 자세는 2차원 영상에서의 위치와 방향뿐만 아니라 3차원 공간에서의 위치와 방향을 포함하기 때문에, 사람의 자세를 3차원 모델을 사용하여 표현하며[6] 이 모델의 파라미터를 추정하는 것으로 사람의 자세를 추정한다. 이 모델이 고차원의 연속 공간에서 표현 될수록 사람 자세는 좀 더 사실적으로 나타나지만 검색 공간과 시간이 지수적으로 증가하여 다루기 힘들게 된다.

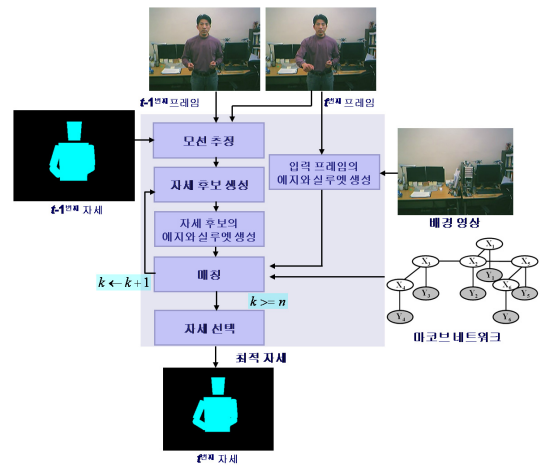
최근, 사람 자세 추적을 하기 위해 확률적인 접근법 [7,8]이 매우 널리 사용되고 있다. 이 접근법은 사람 자세 추정 및 추적을 그래프 모델(graphical model)의 추정 문제로 정형화 한다. 그래프에서 노드는 신체의 부분에 대응되고 에지는 신체 부분들의 연결에 대응된다. 사실적인 표현을 위해서 사람 신체의 회전과 위치를 3차원 연속 공간에서 표현하게 되는데[8] 그러나 이것은 자유도가 높고 따라서 사람의 자세 추정을 어렵게 한다. 또한 사람 모션에 대한 전이 사전 정보(transition prior information)를 사용한 방법[7]은 현재 프레임에서의 신체 부분이 다음 프레임에서 나타날 때, 이전 위치 근처에서 나타난다는 가정을 하고 있으므로 동적인 사람 자세 추적에 적용하기가 쉽지 않다. 이런 문제점을 해결하여 동적인 사람 자세를 추정하기 위해서는 효율적인 파라미터 추정 즉 고차원 공간에서의 효율적인 검색 방법이 필요하게 된다.

본 논문에서는 2차원 영상에서 추정된 신체 각 부분의 모션 정보를 사용해 검색 공간을 정하고 정해진 검색 공간에서 탐색하여 사람의 자세를 추정하는 방법을 제안한다. 제안된 방법은 2차원 영상에서 추정된 신체 각 부분들의 모션 근처에서 사람 자세를 추정함으로써, 속도와 크기가 서로 다른 신체 부분들의 움직임을 지닌 동적인 사람의 3차원 자세를 안정적으로 추정할 수 있다.

제안된 방법의 장점은 다음과 같다. 첫째, 2차원 영상에서 추정된 모션 정보가 3차원 공간에서 자세를 추정하는데 효율적인 정보를 제공한다. 즉, 검색공간을 신체가 움직인 곳 주변으로 한정함으로써 신체 자세 추정에 효율성이 제공된다. 둘째, 방향과 크기가 다른 신체 각 부분의 모션이 따로 추정이 되고, 추정된 모션이 자세 추정에 쓰임으로써, 임의의 움직임, 불규칙적인 움직임을 갖는 사람의 자세에도 대응하여 자세를 추정 할 수 있다. 셋째, 검색 공간을 정하기 위하여 신체 부분 검출기를 설계하지 않아도 되며, 사람의 움직임에 대한 사전 지식을 학습하지 않아도 된다. 좀 더 효율적인 2차원 혹은 3차원 공간에서의 신체 모션 추정기가 개발이 되면, 제안된 방법에 적용하여 확장할 수도 있다.

2. 제안된 방법

우리는 영상에서의 3차원 사람 자세 추정 문제를 그래프 모델에서의 추론 문제로 정형화 한다. 우리는 사람의 머리와 몸통, 팔다리가 느슨하게 연결된 신체 모델(loose-limbed body model)[8]을 사용한다. 3차원 자세와 관계된 조건부 확률은 예제로부터 학습하고, 사람자세는 입력 영상과 가장 잘 매칭 되는 신체 형상을 검색하여 추정한다. 이때 모션 벡터로부터 직접 신체 형상을 생성한다. 이 경우 각각의 신체 부분은 6차원 공간에서의 벡터로 나타나며, 이것으로 자세를 표현하게 된다. 제안된 방법의 흐름도는 그림 1과 같다.

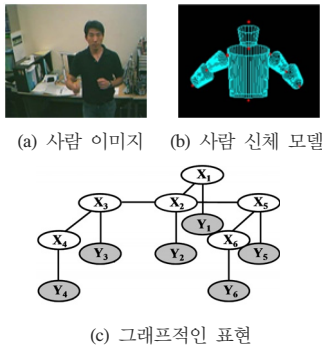


[그림 1] 제안된 방법의 흐름도

초기 자세 혹은 이전 자세와 현재 프레임과 이전 프레임이 주어지면, 제안된 방법은 신체 부분들에 대해서 각각 모션 벡터를 추정한다. 그리고 추정된 모션 벡터 근처에서 자세 후보를 생성한다. 생성된 자세 후보는 2차원 이미지 평면으로 직교 투영 되어지고 투영 평면에서 에지와 실루엣을 생성한다. 동시에 현재 프레임에서도 에지와 실루엣을 생성한다. 자세 후보의 에지와 실루엣은 현재 프레임의 에지와 실루엣과 비교하여 유사도 확률을 계산한다. 그리고 나서, 마코브 네트워크(Markov network)를 사용하여 매칭 확률을 계산한다. 매칭 확률은 마코브 네트워크에서 유사도 확률과 제약조건 확률을 사용해 계산한다. 이 과정을 미리 정한 횟수 동안 반복하고, 가장 잘 맞는 후보 자세를 선택하여 그것을 현재 프레임의 추정된 자세로 결정한다.

2.1 3차원 사람 모델

우리는 그림 2와 같은 6개의 신체 부분으로 구성된 상반신 loose-limbed 신체 모델[8]을 사용한다. 신체 각 부분은 3차원 공간에서 위치와 방향을 정의하기 위한 형상 벡터를 가진다. 전역 좌표계에서 신체 부분을 위치 시키는 것은 모션 검출기가 독립적으로 수행후 그래프 모델의 추론에 의해 그 결과를 조합할 수 있다. 그래프 모델에서 예지는 사-공간에서 인접한 신체 부분의 위치와 각도의 관련성을 나타낸다.



[그림 2] 사람 이미지, 사람 신체 모델과 그래프적인 표현

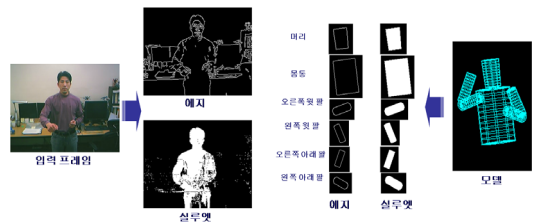
그래프 모델에 의해 신체를 표현하기 위해 각 노드의 변수를 조건부 독립으로 가정하였고, 각각의 신체 부분은 타원형의 실린더로, 6개의 정적 파라메타와 6개의 추정 파라미터를 갖도록 모델링 하였다. 정적 파라미터 $\Phi_i = [l_i, w_i^p, w_i^d, o_i^p, o_i^d]$ 는 신체 부분의 길이와 폭에 해당한다. 추정 파라미터 $X_i = [\chi_i, \Theta_i]^T$ 는 신체 부분 i 의 형상을 나타낸다. 사람 신체 자세는 X 로 표현되고 N 개의 신체 부분에 대해서는 이들의 조합인 $X = \{X_1, X_2, \dots, X_N\}$ 로 표현한다. 우리는 그림 2(c)와 같이 마코브 네트워크를 사용하여 사람 신체 모델을 표현한다. 마코브 네트워크에서 X_i 는 각 신체 부분들(머리, 몸통, 팔 다리)의 자세를 나타내고, Y_i 는 각 신체 부분들의 이미지 관측을 나타낸다. 사람의 자세를 나타내는 정적 파라미터는 미리 정해진 값으로 간주하여 사용하고, 추정 파라미터는 정해지지 않은 값으로 추정하여야 한다.

그림 2(c)에서, 신체부분 i 와 j 사이에는 각 무방향 예지는 제약조건 함수 $\psi_{ij}(X_i, X_j)$ 를 나타낸다. 또한 신체 부분의 물리적인 제약 사항을 표현하기 위해 인접한 두 신체 부분들의 대해 관절점을 정의하였으며, 물리적인 제약사항은 두 신체부분의 관절점의 유클리디언 거리

(Euclidean distance)의 정규 분포에 의해 정의하였다.

2.2 유사도 함수

우리는 $\phi_i(Y_i|X_i)$ 를 신체부분 i 의 자세가 주어진 경우의 이미지 관측의 유사도로 정의한다. 그림 3에서와 같이 자세 후보를 2차원 이미지 평면으로 직교 투영하고, 이미지에서 예지와 실루엣을 생성하여 이 둘의 유사도 확률을 계산하여 자세 추정에 사용한다. 우리는 예지와 실루엣 유사도 함수를 정의하고 이 둘을 결합하여 입력 프레임과 자세 후보와의 유사도를 계산한다.



[그림 3] 입력 프레임과 자세 후보의 예지와 실루엣

2.3 제약조건 함수

자세 추적에는 일반적으로 2가지의 사용할 수 있는 제약조건이 있다. 하나는 모션 제약조건이고 다른 하나는 자세 제약조건이다. 우리는 모션 제약조건 대신에 모션 기반의 검색방법을 사용하여 직접 자세 후보를 추정한다. $\psi_{ij}(X_i, X_j)$ 로 정의된 자세 제약조건 분포는 그래프 모델에서 인접한 신체 부분의 전위 함수(potential function)로 분해 될 수 있으며, 이 자세 제약조건 분포는 다양한 사람 자세를 나타내는 훈련 데이터로부터 학습한다. 자세 제약조건 분포는 정규 분포로 근사하고 공분산 행렬은 훈련 데이터로부터 학습한다.

2.4 자세 추정

자세 추정은 가장 좋은 매칭 확률을 갖는 사람 자세를 찾기 위해 수행된다. 우리는 마코브 네트워크를 사용해 사람 신체 모델을 표현한다. 마코브 네트워크에서 X_i 는 신체 부분들 각각의 자세를 표현하고 Y_i 는 신체 부분들 각각의 이미지 관측을 표현한다. 영상 Y 가 주어졌을 때 영상에 내제하는 사람 자세 X 를 추정하기 위해 검색한다. 마코브 네트워크에서 통계적 종속 관계를 선으로 나타내며, 노드는 통계적 독립을 나타낸다. 노드는 사람의 신체 부분을 나타내며, 이웃 하는 신체부분은 이웃 노드로 나타나며 연결 관계를 갖는다. 신체 부분은 이웃 신체 부분과 연결되고, 각 신체부분은 해당하는 이미지 관측과

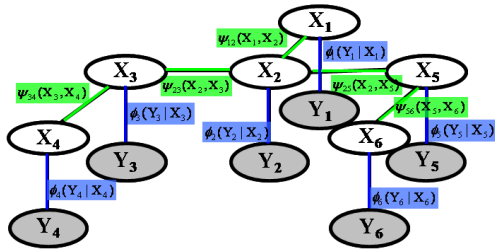
연결된다. 사람 자세를 나타내는 마코브 네트워크에서 가장 좋은 자세 \hat{X} 은 최대 우도 확률(maximum a posteriori)을 갖는 자세이다. 우리 목적은 다음 수식과 같이 마코브 네트워크에서 정의된 결합 확률을 최대화 하는 자세 파라미터를 찾는 것이다.

$$\hat{X} = \underset{X}{\operatorname{argmax}} p(X|Y) \quad (1)$$

이를 위해, 우리는 3차원 공간에서 자세 후보를 생성하고 이것을 2차원 이미지 평면으로 직교 투영한다. 그리고 나서, 마코브 네트워크에서 결합확률을 계산한다. 마코브 네트워크를 위해, 자세 X 와 이미지 투영 Y 에서의 결합 확률을 다음 수식과 같이 사용한다.

$$p(X|Y) \propto \prod_{i \in \nu} \phi_i(Y_i|X_i) \prod_{(i,j) \in \epsilon} \psi_{ij}(X_i, X_j) \quad (2)$$

ν 는 상태 노드와 그 상태 노드에 해당하는 관측 노드 사이의 모든 연결 집합이고, ϵ 는 상태 노드와 그 상태 노드와 인접한 상태 노드의 연결 집합이다. X_j 노드와 연관하여 제약조건 함수로 $\psi_{ij}(X_i, X_j)$ 함수를, 지역적 유사도 함수로 $\phi_j(Y_j|X_j)$ 함수를 사용한다. 우리는 그림 4에서 나타나듯이 마코브 네트워크에서 유사도 함수와 제약조건 함수를 사용하여 결합 확률을 계산한다.



[그림 4] 마코브 네트워크에서 유사도 함수들과 제약조건 함수들

3. 모션 기반의 검색

우리는 고차원 형상 공간에서 분절적인 사람 모션에 의해 발생하는 최적의 사람 자세를 찾고자 한다. 사람의 동적인 움직임은 비선형이며 다양한 양상이 관측된다. 따라서, 이 공간에서 사전 지식없이 방대한 검색을 하는 것은 거의 불가능한 일이다. 사람 자세 추적을 좀 더 효과적으로 수행하기 위해서는 검색 방법을 결정하는 것 뿐

만 아니라 검색 영역을 결정하는 것이 매우 중요하다. 모션 기반의 검색 함수를 사용하여 검색을 하며, 검색 영역은 3차원 위치 공간과 3차원 각도 공간에서 정의한다. 3차원 공간에서의 검색 영역은 구 형태를 띄며 이 영역은 추정된 모션 벡터에 의해서 움직이게 된다.

3.1 모션 기반의 검색 함수

모션 기반의 검색은 자세 후보를 생성하는데 사용되며, 자세 후보를 생성하기 위해 우리는 모션 정보를 근간으로 하는 중요도 함수(importance sampling function)를 정의한다. 중요도 함수를 정의함으로써 추정된 위치 근처의 상태 공간에서만 검색하게 되어 처리 시간을 줄일 수 있고, 물체의 움직임에 대한 사전 지식이 없는 등의 모든 종류의 모션에 대해서도 적응적일 수 있다. 물체의 온라인 모션 추정은 물체의 모션 수행에 대해 어떠한 가정도 하지 않으며 어떤 모션 추정 기술도 이 방법에 사용될 수 있다. 우리는 신체 부분의 모션을 이전 신체 부분 내부에 포함된 픽셀들의 모션 벡터의 평균 벡터로 사용한다. 픽셀들의 모션 벡터를 계산하기 위해 옵티컬 플로우(optical flow) 방법을 사용한다. 때때로 옵티컬 플로우의 이상값이 모션 추정 결과를 나쁘게 하기 때문에 옵티컬 플로우의 방향과 크기 분포를 이용해 이상값을 제외한다. 모션 벡터는 새로운 신체 부분의 위치와 이전 평균 상태의 차이를 알려준다.

$$\Delta X_{i,t} = [x_{i,c}(t) - x_{i,c}(t-1), y_{i,c}(t) - y_{i,c}(t-1), 0, 0, 0, 0] \quad (3)$$

$(x_{i,c}(t), y_{i,c}(t))$ 는 t 시간에 i 번째 신체 부분의 중심이고 우리는 사람 자세를 위해 다음의 정규 분포를 가정한다.

$$X_{i,t} \sim N(X_{i,t-1} + \Delta X_{i,t}, \Sigma_{i,G}) \quad (4)$$

그리고

$$\Sigma_{i,G} = \begin{pmatrix} \sigma_{i,x}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{i,y}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{i,z}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{i,rx}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{i,ry}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{i,rz}^2 \end{pmatrix} \quad (5)$$

$\sigma_{i,x}^2, \sigma_{i,y}^2, \sigma_{i,z}^2$ 는 x-y-z 방향에서의 모션 확산의 분산이고, $\sigma_{i,rx}^2, \sigma_{i,ry}^2, \sigma_{i,rz}^2$ 는 좌표계 각 축으로의 회전 확산의 분산이다. 각 자세 후보는 정규 분포에서 생성이 되고 정규분포는 6차원 공간에서 모션 벡터에 의해 정의 된다. 우리는 중요도 함수로 정규 분포에서 임의 난수 생성 방

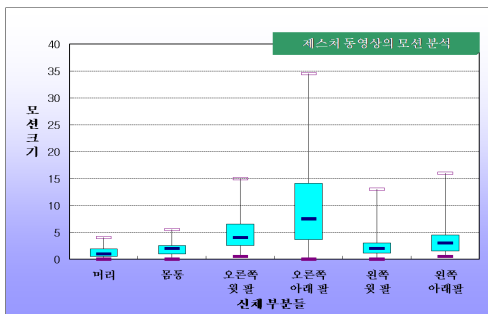
법을 사용한다.

3.2 모션 추정

우리는 모션 추정을 위해 옵티컬 플로우 방법을 사용한다. 이를 위해 신체 부분의 픽셀들에 대해 옵티컬 플로우를 계산하고 이들의 분포를 사용하여 이상값을 제거한다. 옵티컬 플로우는 방향과 크기를 가진다. 우리는 방향 이상값과 크기 이상값을 정의하여 이상값을 제거한다. 방향 이상값은 방향 평균에서 분산 보다 더 치우친 경우 ($\mu - \sigma$ 보다 작거나, 혹은 $\mu + \sigma$ 보다 큰 경우)이며, 크기 이상값도 같은 방법으로 계산하여 제거한다. 구현에 있어서 우리는 옵티컬 플로우 계산을 위해 [3]에서 설명된 OpenCV 라이브러리의 CalcOpticalFlowPyrLK 함수를 사용한다.

4. 실험 및 평가

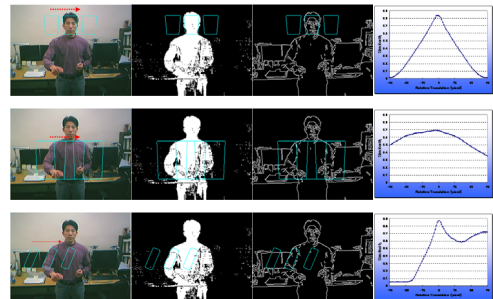
제안된 방법의 성능을 알아보기 위해 제스처 영상을 준비하여 실험하고 평가 하였다. 실험은 펜티엄 IV-2.8 GHz(512MB RAM) 시스템에서, MS Visual C++과 OpenCV[9], OpenGL[10] 라이브러리를 사용하여 수행하였다. 제스처 영상은 노이즈가 어느 정도 있는 방에서 고정된 저사양 웹카메라로부터 획득하였다. 일부 영상은 학습을 위해 사용하고 일부는 테스트를 위해 사용하였다. 영상의 해상도는 320×240이고, 초당 10 프레임의 비율로 획득하였다. 획득된 제스처 영상의 특징을 분석해 보면 그림 5와 같이 사람 머리와 몸통의 움직임은 작은 반면 팔들의 움직임과 편차는 크게 나타났다. 신체 부분의 평균 움직임은 3.2×4.6 픽셀(x-y방향)이고 분산은 12.6×42.4 픽셀이었다. 이것은 신체 각 부분의 움직임이 균일하지 않고 움직임이 많은 부분과 적은 부분이 함께 존재한다는 것을 보여준다.



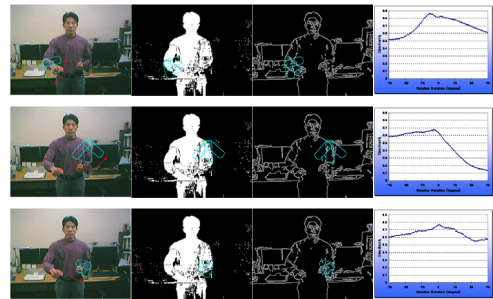
[그림 5] 제스처 동영상의 모션 분석 결과

4.1 유사도 함수의 평가

제안된 방법에 사용된 유사도 함수의 타당성을 알아보기 위해 신체 각 부분들의 상태를 임의로 변화시키면서 유사도를 계산하여 그림 6과 같이 그래프로 나타내었다. 신체 부분들의 임의의 자세를 평행 이동 하면서, 그리고 회전 하면서 이들의 변화에 대한 유사도를 계산하여 나타내었다. 그림 6에서 (a)는 평행 이동의 유사도 결과이고, (b)는 회전의 유사도 결과이며 평행 이동의 경우는 머리, 몸통, 오른쪽 위 팔에 대하여 나타내었고, 회전의 경우는 나머지 신체 부분에 대하여 나타내었다. 결과로 임의의 신체 자세가 이미지의 실제 신체 부분과 겹칠 때 가장 높은 유사도를 나타냄을 알 수 있었고 이것은 실험에서 사용한 유사도 함수가 타당함을 나타낸다.



(a) 평행 이동의 유사도 (머리, 몸통, 오른쪽 윗 팔)

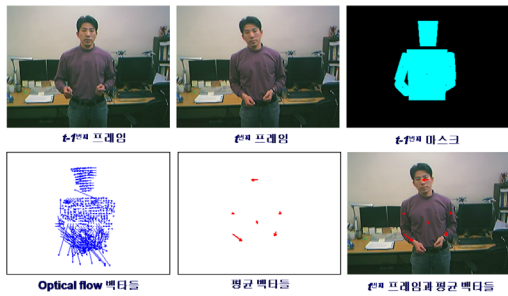


(b) 회전의 유사도(오른쪽 아래 팔, 왼쪽 위·아래 팔)

[그림 6] 유사도 함수의 평가 결과

4.2 모션 추정 결과

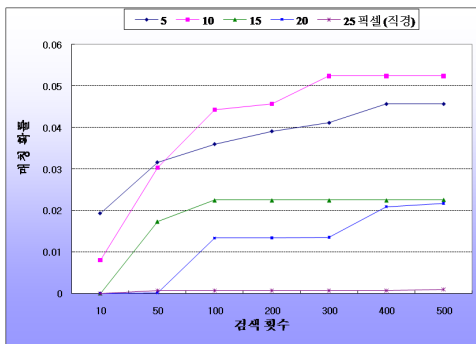
제안된 방법에서 자세 추적의 성능은 2차원 모션 추정 결과에 상당한 영향을 받는다. 그것은 사람 자세가 2차원 이미지에서 추정된 모션에서 바로 검색 영역이 결정되고 검색이 수행되어 자세가 추정되기 때문이다. 모션 추정 결과의 예는 그림 7과 같다.



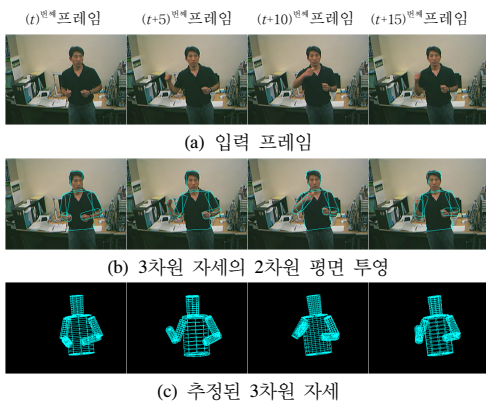
[그림 7] 모션 추정 결과의 예

4.3 사람 자세 추적 결과

가장 좋은 사람 자세를 찾기 위한 검색 영역의 크기와 검색 횟수는 제안된 방법의 성능과 수행 시간에 중요한 영향을 준다. 의미있는 검색 영역과 검색 횟수를 결정하기 위해 다양한 크기의 검색영역과 다양한 검색 횟수로 실험하였으며 그림 8과 같이 검색영역의 직경이 10 픽셀, 검색 횟수가 300번 일 때 좋은 성능을 보였다.



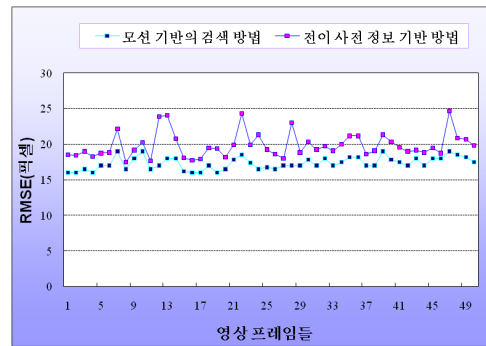
[그림 8] 검색 영역 크기와 검색 횟수에 따른 매칭 확률



[그림 9] 자세 추정 결과 예

그림 9는 사람 자세 추적 결과를 보여준다. 그림 9에서 첫 번째 열은 입력 프레임, 두 번째 열은 추정된 3차원 자세의 2차원 평면 투영, 세 번째 열은 추정된 3차원 자세를 나타낸다. 제스처 영상에는 복잡한 사람 모션이 있고, 우리는 영상에서 모션이 크지 않다는 제약 없이 실험을 수행하였다. 따라서 추정된 신체 모션은 동적인 움직임을 가진 사람의 신체 위치에 대해서 충분한 정보를 제공해 주고 안정적인 자세 추정을 가능하도록 해주었다.

우리는 각 프레임에서 RMSE(Root Mean Square Error)를 계산하고, 전체 영상에서 평균 RMSE를 계산하였다. RMSE는 평균 17.3 픽셀로 나타났다. 우리는 사람 모션에 대해 사전 지식을 사용하는 전이 사전 정보 기반의 방법과 결과를 비교하였다. 전이 사전 정보는 기본적으로 이전 영상에서 나타난 신체 부분은 다음 영상에서 가까이 나타난다는 가정을 하는 방법으로 결과는 그림 10과 같다. 결과에 따르면 제안된 방법이 제스처 영상에서 전이 사전 정보 방법에 비해 좀 더 안정적인 RMSE 결과를 보였다. 이것은 전이 사전 정보 방법에서 사용하는 전이 제약 사항이 동적인 사람의 움직임에 대해 효율적이지 않기 때문인 것으로 판단할 수 있다.



[그림 10] RMSE 비교 결과

5. 결론

본 논문에서는 모션 기반의 검색 방법을 사용한 동적인 사람 자세 추적 방법을 제안하였다. 제안된 방법은 3차원 공간에서 하나의 사람 자세 후보를 생성하고, 생성된 자세 후보를 2차원 이미지 공간으로 투영하여, 투영된 사람 자세 후보와 이미지와의 특징 값 유사성을 비교한다. 이 과정을 정해진 조건을 만족 할 때까지 반복하여 이미지와의 유사성과, 신체 부분간 연결성이 가장 좋은 3차원 자세를 추정하였다. 이미지와의 유사성 척도로는 예지와 실루엣의 유사성을, 신체 부분간의 연결성은 신체

관절 사이의 3차원 거리를 사용하였다. 신체 부분의 모양은 실린더로, 사람의 자세는 마코브 네트워크로 표현하였으며, 사람 자세의 좋은 정도는 마코브 네트워크의 결합 확률을 사용하였다. 검색 공간을 효과적으로 줄이기 위하여 2차원 이미지에서 추정된 모션 정보를 사용하여 자동으로 검색 영역을 정하였다. 복잡한 모션을 포함한 제스처 동영상에서 실험한 결과 에러율이 약 17.3 픽셀로 나타났으며 이것은 제안된 방법이 안정적인 추적 결과를 나타내는 것이었다. 제안된 방법은 다른 방법에서 사용하는 상향식 신체 부분 검출기나 전이 사전 정보를 학습하는 것을 필요로 하지 않으며, 애니메이션 콘텐츠 제작뿐만 아니라 일반인을 대상으로 하는 동작 인터페이스 분야, 상황인식 컴퓨팅 분야로의 확대 적용이 가능하다. 향후 추진 과제로는 제안된 방법을 갑자기 변하는 모션이나, 카메라 모션에 대해 실험할 것이며, 갑작스럽고 불규칙하게 변화하는 움직임에 대응할 수 있도록 연구를 진행할 계획이다.

참고문헌

[1] R. Poppe, "Vision-based human motion analysis: An overview", *Computer Vision and Image Understanding*, vol. 108, issue 1-2, pp.4-18, 2007.

[2] H. S. Park, D. J. Jung and H. J. Kim, "Vision-Based Game Interface Using Human Gesture", *Proceedings of the IEEE Pacific-Rim Symposium on Image and Video Technology, LNCS 4319*, pp. 662-671, 2006.

[3] S. Vedula, S. Baker, P. Rander, R. Collins and T. Kanade, "Three-Dimensional Scene Flow", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 475-480, 2005.

[4] X. Ren, A. Berg and J. Malik, "Recovering Human Body Configurations using Pairwise Constraints between Parts", *Proceedings of the IEEE International Conference on Computer Vision*, vol. 1, pp. 824-831, 2005.

[5] A. Agarwal and B. Triggs, "Recovering 3D Human Pose from Monocular Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 44-58, 2006.

[6] L. Sigal, M. Isard, B. H. Sigelman and M. J. Black, "Attractive people: Assembling loose-limbed models using non-parametric belief propagation", *proceedings of the Neural Information Processing Systems*, vol. 16, pp. 1539-1546, 2003.

[7] M. W. Lee and R. Nevatia, "Dynamic Human Pose Estimation using Markov Chain Monte Carlo Approach", *Proceedings of the IEEE Workshop on Motion and Video Computing*, vol. 2, pp. 168-175, 2005.

[8] L. Sigal, S. Bhatia, S. Roth, M. J. Black and M. Isard, "Tracking Loose-limbed People", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 421-428, 2004.

[9] OpenCV Reference Manual, Intel Open Source Computer Vision Library, <http://www.intel.com/technology/computing/opencv/>.

[10] OpenGL Reference Manual, Silicon Graphics Inc. Open Source 3D Graphic Library, <http://www.opengl.org/documentation/>.

정 도 준(Do-Joon Jung)

[정회원]



- 2004년 2월 : 경북대학교 컴퓨터공학과 (컴퓨터공학석사)
- 2008년 8월 : 경북대학교 컴퓨터공학과 (컴퓨터공학박사)
- 2005년 3월 ~ 현재 : 경운대학교 정보통신공학과 교수

<관심분야>
컴퓨터비전, 패턴인식, 인공지능

윤 정 오(Jeong-Oh Yoon)

[정회원]



- 1991년 2월 : 경북대학교 전자공학과 (정보통신공학석사)
- 2002년 2월 : 경북대학교 전자공학과 (정보통신공학박사)
- 1997년 3월 ~ 현재 : 경운대학교 정보통신공학과 교수

<관심분야>
멀티미디어통신, 영상처리, 정보보호