

반복 최근접점와 파티클 필터를 이용한 인간 신체 움직임 추적

(Human Body Motion Tracking Using ICP and Particle Filter)

김 대 환 [†] 김 효 정 ^{**} 김 대 진 ^{***}
(Daehwan Kim) (Hyojung Kim) (Daijin Kim)

요약 본 논문은 빠르게 움직이는 인간 신체를 추적할 수 있는 실시간 인간 신체 움직임 추적 알고리즘을 제안한다. 반복 최근접점(Iterative closest point) 알고리즘은 효율적이고 계산량이 적어 실시간 인간 신체 움직임 추적에 적합하지만 잘못된 최근접점(Closest point) 선택으로 인해 국부적 최소점(Local minimum)에 쉽게 빠지게 되어 종종 추적이 실패하게 된다. 이를 극복하기 위해, 반복 최근접점 알고리즘에 움직임 이력(Motion history) 정보를 기반으로 한 파티클 필터(Particle filter)를 결합한다. 제안하는 인간 신체 움직임 추적은 계층적 트리 구조를 사용함으로써 신체 추적 공간 크기를 줄여주며, 움직임 이력 정보를 이용한 움직임 예측 모델을 사용함으로써 빠른 인간 신체 움직임 추적을 가능하게 한다. 실험 결과는 제안하는 인간 신체 움직임 추적이 정확한 추적 성능과 빠른 수렴 속도를 가진다는 것을 보여 준다.

키워드 : 인간 신체 움직임 추적, 반복최근접점, 파티클 필터, 움직임 이력 정보, 인간 신체 모델

Abstract This paper proposes a real-time algorithm for tracking the fast moving human body. Although Iterative closest point (ICP) algorithm is suitable for real-time tracking due to its efficiency and low computational complexity, ICP often fails to converge when the human body moves fast because the closest point may be mistakenly selected and trapped in a local minimum. To overcome such limitation, we combine a particle filter based on a motion history information with the ICP. The proposed human body motion tracking algorithm reduces the search space for each limb by employing a hierarchical tree structure, and enables tracking of the fast moving human bodies by using the motion prediction based on the motion history. Experimental results show that the proposed human body motion tracking provides accurate tracking performance and fast convergence rate.

Key words : Human body motion tracking, ICP, particle filter, motion history information, human body model

· 이 연구(논문)는 산업자원부 지원으로 수행하는 21세기 프론티어 연구개발사업(인간기능 생활지원 지능 로봇 기술개발사업)의 일환으로 수행되었습니다.

· 본 연구는 한국과학재단을 통해 교육과학기술부의 WCU 사업(World Class University)으로부터 지원받아 수행되었습니다(과제번호: R33-2008-000-10103-0).

[†] 학생회원 : POSTECH 컴퓨터공학과
msoul98@postech.ac.kr

^{**} 비회원 : KT 미래기술연구소 기술기획담당 선임연구원
kimhyo_jung@kt.co.kr

^{***} 정회원 : POSTECH 컴퓨터공학과 교수
dkim@postech.ac.kr

논문접수 : 2008년 1월 25일

심사완료 : 2009년 10월 13일

Copyright©2009 한국정보과학회 : 개인 목적이거나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 얻고 비용을 지불해야 합니다.

정보과학회논문지: 소프트웨어 및 응용 제36권 제12호(2009.12)

1. 서론

비전 기반 인간 움직임 추적은 비주얼 감시, 움직임 획득과 인간 움직임 분석과 관련하여 오랜 시간 동안 연구되고 있다. 또한 인간-컴퓨터 상호 작용(Human-computer interface)이나 인간-로봇 상호 작용(Human-robot interface)으로 확장되어 사용할 수 있기 때문에 인간 움직임 추적 기술은 중요한 핵심 기술 중에 하나이다. 하지만 인간 움직임 추적은 인간 마다 신체 구성 요소들 간의 크기와 비율이 다르고 아주 복잡하고 빠른 움직임 때문에 여전히 어려운 문제로 남아 있다.

인간 움직임 추적의 어려움을 해결하고자 많은 방법들이 연구되고 있다. Sminchiseseon 등은[1] 오직 2차원 위치 정보만을 사용하여 3차원 움직임 정보를 생성하는 “Stochastic sampling” 방법을 제안했다. Lee등은[2] 인

간 각 신체 구성 요소간의 최대 유사도(Maximum likelihood)를 가지는 “Proposal map”을 제안했다. Navaratnam 등은[3] 3차원 상체 포즈를 생성하는 상향(Bottom up) 방식의 신체 구성 요소 검출과 계층적 인간 모델을 결합하였다. 하나의 카메라만을 사용하는 접근 방법들은 중첩 문제나 복잡한 환경에서의 인간만을 구분(Segmentation) 하는 문제를 해결할 수 없는 상황이다.

최근에 다중 카메라를 이용하는 방법들이 주로 연구되고 있다. Deutscher 등은[4] 다중 카메라를 이용하여 인간의 실루엣(Silhouette)과 에지(Edge)를 특징으로 어니일드 파티클 필터(Annealed particle filter) 방법을 이용하여 국부적 최소(Local minimum) 포즈에 빠지지 않도록 하는 방법을 제안했다. 하지만 이 방법은 최적의 포즈를 찾기 위해 많은 시간이 걸린다. MacCormick과 Isard[5]는 관절 대상(Articulated object)들의 효율적인 2차원 포즈 예측을 위해 분리된 샘플링(Partioned sampling) 방법을 제안했다. 그러나 이 접근 방법은 인간 신체 전체에 대한 예측으로 확장하기에는 어려움이 있다. Caranzza[6]는 기울기 하강 방법(Gradient descent method)와 그리드 탐색(Grid search)을 결합하여 인간 신체 움직임을 추적하였지만 실시간 추적에 어려움이 있고 많은 양의 텍스처 정보들이 필요하다. Demirdjian [7]은 반복 최근접점(Iterative closest point) 알고리즘을 이용하여 인간 신체 모델을 스테레오 입력 데이터에 정합(Fitting)함으로써 인간 움직임을 추적하는 방법을 제안했다. 이 방법은 계산량이 적어 실시간 움직임 추적 방법에 알맞다. 하지만 복잡하고 빠른 움직임에서는 추적 성능의 한계가 있다.

실시간 인간 움직임 추적을 위해서 효율적이고 복잡도가 낮은 반복 최근접점 알고리즘을 사용하도록 한다. 그러나 단순히 반복 최근접점 알고리즘만을 이용해서는 복잡하고 빠른 움직임을 가지는 인간 움직임을 정확하게 추적할 수 없다. 이를 극복하기 위해서 반복 최근접점 알고리즘에 3단계의 움직임 업데이트 요소를 추가한 인간 신체 움직임 추적 알고리즘을 제안하고자 한다.

첫 번째, 인간의 신체 구성 요소들이 계층적으로 연결된 구조이므로 각 신체 구성 요소들과 연결된 상위 신체 구성 요소들의 움직임에 영향을 받는다고 가정한다. 상위 신체 구성 요소의 움직임 정보를 현 신체 구성 요소에 적용하도록 한다. 이는 각 신체 구성 요소들의 움직임 검색 공간을 제한시켜 준다. 두 번째는 각 신체 구성 요소들이 정확하게 연결된 형태로 추적할 수 있도록 하는 것이다. 인간 모델 생성 시 신체 구성 요소들끼리의 연결 점을 미리 정의하여 계층적으로 신체 구성 요소들을 연결해 준다. 마지막으로 각 신체 구성 요소들의 빠른 움직임을 추적하기 위해 파티클 필터(Particle

filter)와 같은 확률 과정(Stochastic process)을 결합하도록 한다. 파티클 필터[8-10]는 확률적인 추적 특성을 가지고 있어 빨리 움직이는 물체도 잘 추적할 수 있다. 파티클 필터의 중요한 두 컴포넌트는 관측 모델(Observation model)과 상태 전이 모델(State transition model) 이다. 정확한 관측 모델을 위해서 인간 신체 모델 표면에 있는 법선 벡터(Normal vector)와 실루엣(Silhouette) 정보를 이용하도록 한다. 인간의 빠른 움직임을 반영하기 위한 상태 전이 모델은 과거의 움직임 이력(Motion history) 정보를 이용한다.

이 논문은 2장에서 반복 최근접점과 파티클 필터를 다루며, 3장에서는 인간 신체 모델에 대해서 설명하고, 4장에서는 제안하는 알고리즘에 대해 설명한다. 5장에서는 실험을 통하여 제안하는 알고리즘의 타당성을 입증한다. 6장에서 결론이 뒤따른다.

2. 이론적 배경

이 장에서는 기존의 반복 최근접점 알고리즘과 파티클 필터를 설명한다.

2.1 반복 최근접점(Iterative Closest Point)

반복 최근접점(Iterative closest point) 알고리즘은 3차원 데이터 집합들끼리 정합하는 방법이다[11,12]. 이는 주어진 두 집합의 3차원 데이터들 사이의 밀접한 대응점 집합(Closest corresponding point set)을 찾고 대응된 점 사이의 오차를 최소화하는 움직임 파라미터(Motion parameter)를 예측하는 방법이다. 비용 함수(Cost function)는 다음과 같다.

$$\min_{\Delta\theta} \sum_i |RM_i + T - D_i|^2 \tag{1}$$

여기서 M_i 는 모델 점들, D_i 는 데이터 점들을 나타낸다. 모델 점들과 데이터 점들 사이의 비용 함수를 최소화하는 회전(Rotation) 파라미터 R 과 이동(Translation) 파라미터 T 를 예측해야 한다.

반복 최근접점 알고리즘은 몇 단계의 과정을 반복하여 최적의 파라미터를 구한다.

- (1) 파라미터 R 과 T 를 초기화 한다. 반복 변수 $k=0$.
- (2) 각 M_i 에 대해서, 데이터 점 D_i 중 가장 가까운 점 (Closest point)을 계산한다.

$$Y_k = C(M, D_k) \tag{2}$$

여기서 Y_k 는 가장 가까운 점들의 쌍 집합이다.

- (3) 주어진 Y_k 에 대해서 대응점끼리의 집합 정합 방법[11]을 사용하여 변환 파라미터 R 과 T 를 계산한다.

$$(R, T, d_k) = Q(Y_k) \tag{3}$$

여기서 Q 는 정합 과정이며 d_k 는 오차값이다.

- (4) 모델 점 M_i 에 정합을 적용한다.
- (5) 오차의 변화가 적으면 반복을 멈추고, 그렇지 않으면 단계 (2)부터 다시 반복한다.

$$d_{k-1} - d_k < \tau \quad (4)$$

2.2 파티클 필터

파티클 필터는 샘플링 방법을 사용하여 베이지안 필터를 구현한 예측 기술이다. 주어진 관찰 데이터 $I(1:t) = I_1, \dots, I_t$ 로부터 알려지지 않은 상태를 예측한다.

가중치 부가 파티클 집합 $S_t = \{\theta_i^{(p)}, w_i^{(p)}\}_{p=1}^P, \sum_{p=1}^P w_i^{(p)} = 1$ 에 의해서 사후분포(Posterior distribution) $p(\theta|I_{1:t})$ 을 간략화 한다. 총 세 단계로 구성된다. 재샘플링 단계(Re-sampling), 예측 단계(Prediction), 업데이트 단계(Update)이다. 재샘플링 단계에서는 같은 가중치를 가지는 파티클 $\{\theta_{t-1}^{(p)}, 1\}_{p=1}^P$ 을 얻기 위해 파티클 S_{t-1} 을 재샘플링한다. 예측 단계에서는 파티클 $\{U_t^{(p)}\}_{p=1}^P$ 을 생성하고 상태 전이 모델 $\theta_t = F_t(\theta_{t-1}, U_t)$ 을 사용하여 $\{\theta_{t-1}^{(p)}, 1\}_{p=1}^P$ 를 전이한다. 업데이트 단계에서는 관찰 확률(Observation likelihood) $w_i^{(p)} \propto p(I_t|\theta_i^{(p)})$ 을 기반으로 각 파티클의 가중치를 업데이트한다. 표 1은 파티클 필터의 과정을 요약했다.

표 1 일반적인 파티클 필터 과정

- 초기화 : 사전 밀도 (Prior density)를 이용하여 같은 가중치를 갖는 파티클들 $\{(\theta_0^{(1)}, 1), \dots, (\theta_0^{(P)}, 1)\}$ P개를 생성한다.
- For t=1,2, ...
 - For p=1,2, ...
 - (a) 재샘플링 : 같은 가중치를 갖는 새로운 파티클 $\{\theta_{t-1}^{(p)}, 1\}_{p=1}^P$ 를 얻기 위해 파티클을 재샘플링한다.
 - (b) 예측 : 상태 전이 모델 $\theta_t^p = F_t(\theta_{t-1}^{(p)}, U_t^{(p)})$ 을 이용하여 파티클을 전이한다.
 - (c) 업데이트 : 관찰 확률 $p(I_t|\theta_i^{(p)})$ 을 이용하여 가중치 $w_i^{(p)}$ 를 업데이트한다.
- End
 - 정규화 : 가중치를 정규화 한다.

$$w_i^{(p)} = \frac{p(I_t|\theta_i^{(p)})}{\sum_{i=1}^P p(I_t|\theta_i^{(i)})}$$
 - 추측 : 최대 사후 확률(Maximum a posterior)을 이용하여 상태를 추측한다. $\hat{\theta}_t = \operatorname{argmax}_{\theta_i} w_i^{(p)}$
- End

3. 인간 신체 모델과 움직임 파라미터

인간 신체를 모델링하는 것은 인간 신체 구성 요소들의 크기와 비율이 많이 다르고 인간의 복잡한 움직임을 잘 표현할 수 있어야 되기 때문에 상당히 어려운 일이다. 일반적으로 인간을 개략적으로 가장 잘 표현할 수 있어, 다루기 쉬운 실린더 모델을 선호한다[7].

사용된 인간 신체 모델은 신체 구성 요소를 10개의 강체로 나눈다고 가정한다. 머리를 제외한 9개의 신체 구성 요소를 실린더(Cylinder)로 머리는 1개의 구(Sphere)로 구성한다. 각 실린더와 구는 인간의 계층적인 구조 아래 유니버설 조인트(Universal joint)에 의해서 연결된다. 각 실린더는 2개의 반지름 cr_x, cr_z 와 높이 ch 로 이루어지고 구는 반지름 sr_s 으로 이루어진다. 위의 파라미터를 이용하여 3차원 모델 점 집합 $X = [(x_1, y_1, z_1)^T, \dots, (x_N, y_N, z_N)^T]^T$ 을 만들 수 있다. 여기서 N 은 모델 점의 개수이다. 만약 실린더나 구 표면 위의 3차원 한 점이 $x = (x, y, z)^T$ 라 하고 그 점에 대응되는 2차원 점을 $u = (u_x, u_y)^T$ 라 한다면, 투시 투영(Perspective projection)을 사용하여 x 를 u 로 투영할 수 있다.

$$u = p(x) = \frac{f}{z} [x \ y]^T \quad (5)$$

여기서 f 는 초점 거리(Focal length)이다.

각 조인트거리 접하는 위치는 인간 구조를 고려하여 미리 결정하도록 한다. 그림 1은 인간 신체 모델을 보여준다.

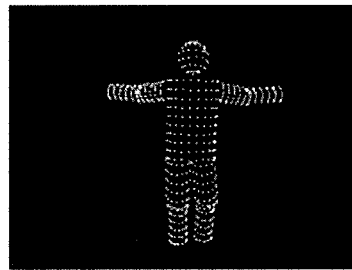


그림 1 인간 신체 모델

각 인간 신체 구성 요소의 강체 움직임 벡터(Rigid motion vector)는 3 회전 파라미터(w_x, w_y, w_z)와 3 이동 파라미터(t_x, t_y, t_z)로 구성된다. 실린더나 구의 표면 점의 3차원 좌표가 $x = (x, y, z)^T$ 라면 강체 움직임 벡터에 의해서 변환된 위치는 다음과 같이 표현할 수 있다.

$$M(x;\theta) = Rx + T \quad (6)$$

여기서 M 은 (w_x, w_y, w_z) 에 대응되는 3차원 회전 행렬 $R \in \mathbb{R}^{3 \times 3}$ 과 3차원 이동 벡터 $T \in \mathbb{R}^{3 \times 1}$ 로 표현되는 강

체 변환 함수이다. 오일러(Euler) 표현 방법에 의해서 3D 회전 행렬을 표현하는 방법은 복잡하고 어렵기 때문에 트위스트(Twist) 표현 방법[13]을 사용한다. 이에 따른 M 은 다음과 같이 간략화 할 수 있다.

$$M(x;\theta) = \begin{pmatrix} 1 & -w_z & w_y \\ w_z & 1 & -w_x \\ -w_y & w_x & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \\ t_z \end{pmatrix} \quad (7)$$

4. 인간 신체 움직임 추적

4.1 인간 신체 움직임 파라미터 업데이트

인간 신체는 단순한 개체가 아니라 여러 개의 다른 부분들이 구조적으로 연결된 개체이다. 그래서 각 신체 구성 요소들은 다른 신체 구성 요소의 움직임에 영향을 주거나 혹은 영향을 받는다. 그러나 신체 구성 요소들의 모든 유기적인 움직임을 표현하기가 어렵기 때문에 인간 신체의 구조는 계층적 트리 구조의 형태로써, 각 신체 구성 요소들의 움직임은 상위 계층의 신체 구성 요소들의 움직임에 의해서 영향을 받는다고 가정한다 [14-16]. 이 구조는 인간 신체 움직임을 표현하기에 간단하고 쉬워 3차원 인간 신체 움직임 추적에 적합하다. 일반적으로 그 트리 구조의 루트는 몸통이 되고 1 레벨 계층은 머리, 왼쪽 상부 팔, 오른쪽 상부 팔, 왼쪽 상부 다리, 오른쪽 상부 다리로 구성되고 2 레벨 계층은 왼쪽 하부 팔, 오른쪽 하부 팔, 왼쪽 하부 다리 그리고 오른쪽 하부 다리로 구성되어 진다. 그림 2는 계층적 인간 트리 구조를 보여준다.

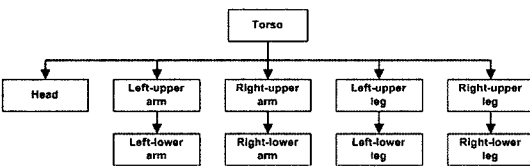


그림 2 계층적 인간 트리 구조

인간 신체 움직임을 정확하게 표현하기 위해서, 3가지의 움직임 요소들을 고려해야만 한다. 첫 번째는 각 신체 구성 요소가 계층적 구조로 이루어져 있기에 각 신체 구성 요소의 움직임은 그 신체 구성 요소의 상위 계층의 움직임에 의해서 영향을 받아야만 한다. 두 번째는 인간 신체가 유기적으로 연결된 개체기에 조인트 움직임도 고려해야만 한다. 마지막으로 각 신체 구성 요소 자신의 움직임을 반영해야만 한다. 최종적으로 인간 신체 움직임을 결정하는 t 시간에서의 i 번째 신체 구성 요소의 움직임 파라미터 업데이트(Global motion parameter updates) $\Delta\theta_{g(i),t}^*$ 는 다음과 같다.

$$\Delta\theta_{g(i),t}^* = \Delta\theta_{l(i),t}^* + \Delta\theta_{j(i),t} + \Delta\theta_{h(i),t} \quad (8)$$

여기서 $\Delta\theta_{l(i),t}^*$ 는 t 시간에서의 i 번째 신체 구성 요소 자체의 움직임 파라미터 업데이트(Local motion parameter updates)이고 $\Delta\theta_{j(i),t}$ 는 t 시간에서의 i 번째 신체 구성 요소와 상위 신체 구성 요소와의 조인트 움직임 파라미터 업데이트(Joint motion parameter updates)이다. $\Delta\theta_{h(i),t}$ 는 t 시간에 i 번째 신체 구성 요소의 상위 신체 구성 요소에 의해서 발생하는 계층적 움직임 파라미터 업데이트(Hierarchical motion parameter updates)이고, $i = 1, 2, \dots, 10, p$, p 는 i 번째 신체 구성 요소의 상위 신체 구성 요소의 인덱스이다.

계층적 움직임 파라미터 업데이트는 식 (9)로 나타낸 바와 같이 상위 신체 구성 요소의 최종 움직임 업데이트를 통해 얻을 수 있다. 몸통을 제외한 모든 신체 구성 요소들은 계층적 움직임에 영향을 받는다.

$$\Delta\theta_{h(i),t} = \Delta\theta_{g(p),t}^* \quad (9)$$

조인트 움직임 파라미터 업데이트는 식 (10)으로 나타낸 바와 같이 i 번째 신체 구성 요소와 그의 상위 신체 구성 요소와의 미리 정의된 접점 사이의 차이 벡터를 구함으로써 계산되어질 수 있다.

$$\Delta\theta_{j(i),t} = M(x_{j(p),j(i)}; \theta_{g(p),t}^*) - M(x_{j(i),j(p)}; \Delta\theta_{h(i),t} + \Delta\theta_{g(i),t-1}^*) \quad (10)$$

여기서 $\Delta\theta_{g(p),t}^*$ 는 t 시간에서의 i 번째 신체 구성 요소의 상위 최종 움직임 벡터이다. $x_{j(p),j(i)}$ 는 i 번째 신체 구성 요소의 접점과 연결된 상위 신체 구성 요소의 접점이고 $x_{j(i),j(p)}$ 는 상위 신체 구성 요소와 연결될 i 번째 신체 구성 요소의 접점이다.

마지막으로 자체의 움직임 파라미터 업데이트는 신체 구성 요소 자체의 움직임을 나타낸다. 움직임 파라미터 업데이트는 계층적 움직임 파라미터 업데이트와 조인트 움직임 파라미터 업데이트를 초기 값으로 반복 최근접점 알고리즘을 수행하여 얻을 수 있다. 계층적 트리 구조를 가지고 있는 인간의 신체 구성 요소 중 몸통을 제외한 나머지 신체 구성 요소들은 각 상위 신체 구성 요소의 연결 점을 기반으로 하여 움직임을 갖게 된다. 따라서 몸통은 기존의 비용 함수를 사용하여 움직임 파라미터를 구하면 되지만 나머지 신체 구성 요소들은 이동 파라미터 T 가 필요 없다.

4.2 파티클 필터가 내포된 반복 최근접점을 이용한 인간 신체 모델 추적

인간 신체 움직임 추적을 위해서 Demirdjian[7]에 의해 적용된 반복 최근접점 알고리즘은 각 신체 구성 요소가 작은 움직임이 발생했을 경우는 잘 동작한다. 하지만 일반적으로 인간 움직임의 순간 속도가 아주 빠르기 때문에 반복 최근접점만을 이용한 움직임 추적은 아주

제한된 성능을 보여 준다. 따라서 빠른 속도에 강인한 추적 알고리즘을 위해 반복 최근접점에 파티클 필터를 내포하도록 한다. 즉 각 시간마다 파티클 필터를 이용하여 대략적인 초기 움직임을 구한 뒤 반복 최근접점의 정합 과정을 거쳐 움직임을 추적하도록 한다. 이는 큰 움직임이 발생했을 경우 움직임이 국부적 최소점(Local minimum)에 빠지는 것을 방지해 줄뿐더러 가장 많은 시간을 소비하는 반복 최근접점의 반복 횟수를 줄여줘 추적 시간을 감소 시켜준다. 파티클 필터는 관찰 모델, 상태 전이 모델로 구성된다. 관찰 모델에서는 관찰 확률(Observation likelihood)을 간략화 하는 가중치 함수를 사용하고, 상태 전이 모델에서는 빠른 움직임을 가지는 신체 구성 요소들을 추적하기 위해서 움직임 이력 정보를 사용하도록 한다.

4.2.1 관찰 모델(Observation model)

관찰 모델은 관찰 확률을 사용하여 주어진 현재 상태에서 인간 모델 데이터와 관찰 데이터사이의 관계를 얻어내는 것이다. 일반적으로 관찰 확률을 간략화 하는 가중치 함수를 사용한다[4]. 가중치 함수를 사용하는 장점은 관찰 모델을 평가할 때 최소의 계산만을 요구한다.

가중치 함수를 생성하기 위해서 어떤 특징을 사용할 것인지 결정해야만 한다. 그 특징은 추출하기 쉬워야 하고 인간 신체 구성 요소의 움직임을 잘 반영할 수 있어야 한다. 여기서는 두 가지의 특징을 사용한다. 법선 벡터(Normal vector)와 전경 실루엣(Foreground silhouette)이다. 법선 벡터는 인간의 볼륨 정보를 반영할 수 있고 전경 실루엣 또한 인간의 형상 정보를 반영할 수 있어 움직임 추적하기에 좋은 특징들이다. 이때 제안하는 가중치 함수는 다음과 같다.

$$P(I_i|\theta_i^{(p)}) \propto \exp\left(-\frac{1}{2}\left(\frac{1}{N}\sum_{i=1}^N(1-\vec{n}_i \cdot \vec{n}(u_i)) + \frac{1}{N}\sum_{i=1}^N(1-I(u_i))\right)\right) \quad (11)$$

여기서 \vec{n} 은 인간 모델 표면 점들을 이용하여 계산된 법선 벡터, $\vec{n}(u_i)$ 는 3차원 관절 데이터의 u_i 위치에서의 법선 벡터, u_i 는 현재 상태에서 인간 모델 표면 점들에 대응하는 2차원 위치이고 $I(u_i)$ 는 미리 계산된 전경 실루엣 이미지의 u_i 위치에서의 바이너리 값이다.

가중치 함수식 (11)를 이용하여, 관찰 확률을 최대로 하는 가장 오차가 작은 파티클 $\hat{\theta}_i$ 를 선택한다.

4.2.2 상태 전이 모델(State transition model)

상태 전이 모델은 움직이는 신체 구성 요소의 동적 상태(Dynamics)를 표현한다. 일반적으로 사용하는 제로-속도 모델(Zero-velocity model)은 상태의 변화가 큰 신체 구성 요소들을 추적하기 위해서는 부적당하다. 따

라서 빠른 움직임에 적용할 수 있는 상태 전이 모델이 필요하다. 이것을 해결하기 위해서, 신체 구성 요소들의 동적 상태가 등속(Constant velocity)을 가진다고 가정하도록 한다. 이때 시간 $t-1$ 시간에서의 속도 v_{t-1} 은 다음과 같이 계산될 수 있다.

$$v_{t-1} = \theta_{t-1}^* - \theta_{t-2}^* \quad (12)$$

위의 움직임 이력 정보를 사용하여, 현재 t 시간에서의 움직임 속도는 $t-1$ 시간의 움직임 속도에 의해서 예측된다.

$$\bar{v}_t = v_{t-1} \quad (13)$$

따라서 상태 전이 모델을 이용하여 파티클 집합을 다음과 같이 생성할 수 있다.

$$\theta_t^{(p)} = \theta_{t-1}^* + \bar{v}_t + \eta_t^{(p)} \quad (14)$$

여기서 $\eta_t^{(p)}$ 는 가우시안 분포를 기반(Gaussian distribution)으로 시스템 노이즈이다.

4.3 제안하는 인간 신체 움직임 추적 알고리즘

표 2에서 제안하는 인간 신체 움직임 추적 알고리즘을 요약하였다.

5. 실험 결과 및 분석

제안하는 인간 신체 움직임 추적 알고리즘은 Microsoft Visual C++과 Open CV로 Intel Core 2 CPU와 RAM 2GHz를 가진 윈도우 PC 플랫폼 상에서 구현되어졌다. 실험 영상을 위해 포인트 그레이(Point grey)사의 범블비 2로 입력 이미지를 획득했고 3차원 관절 데이터는 범블비 카메라에 제공되는 SDK(Software development kit)를 이용하여 얻었다. 실험을 위한 여러 가지 파라미터는 표 3에 나열하였다.

제안하는 알고리즘을 평가하기 위해서 실내 환경 상에서 획득한 영상들을 이용하여 성능을 평가 하였다. 실험에 사용된 영상들의 길이는 대략 100~200 프레임 정도이며 알고리즘 동작 속도는 대략 4~5 프레임 정도였다. 카메라와 인간 사이의 거리는 대략 2m 정도이다.

5.1 빠른 속도를 가지는 팔 움직임 추적 실험

이 실험에서는 빠른 속도의 움직임이 발생했을 때 제안된 알고리즘과 기존의 알고리즘의 성능을 비교 분석하였다. 실험한 영상은 양 팔을 위 아래로 빠르게 움직이는 동작을 획득한 것이다.

1 프레임간의 팔의 각도는 대략 40~45도 정도이다. 양팔을 추적하기 위해 사용된 파티클 수는 상위 팔은 20개 하위 팔은 50개를 사용했다.

그림 3은 기존의 반복 최근접점 알고리즘과 본 논문에서 제안하는 알고리즘 추적 실험 결과 영상이다. 왼쪽 팔은 두 가지 알고리즘 다 추적에 성공했지만 상대적으

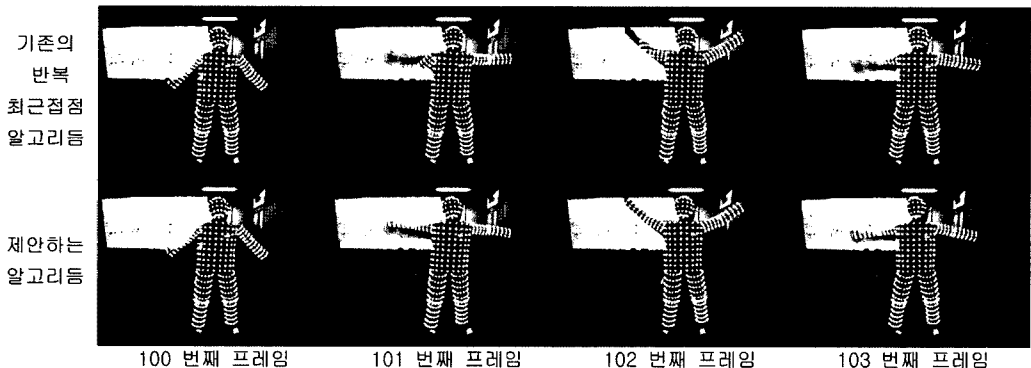
표 2 제안하는 인간 신체 움직임 추적 알고리즘

1. 인간 신체 모델을 구성하는 각 개체의 움직임 $\theta_{g(i),0}^*$ 을 초기화 과정을 통해 초기화 한다.
 2. 각 신체 구성 요소마다의 파티클 필터의 시스템 노이즈 분산 σ^2 과 파티클 수 $P(i)$ 를 설정한다. 사전 분포(Prior distribution) $p(\theta_{g(i),0}^*)$ 에 따라서 파티클 집합 $S_{g(i),0} = \left\{ \theta_{g(i),0}^{(p)}, \frac{1}{P(i)} \right\}_{p=0}^{P(i)}$ 을 생성한다.
- For $t = 1, 2, \dots$
- For $i = 1, 2, \dots, 10$
3. 식 (9)를 이용하여 계층적 움직임 파라미터 업데이트 $\Delta\theta_{h(i),t}$ 를 계산한다.
 4. 식 (10)을 이용하여 조인트 움직임 파라미터 업데이트 $\Delta\theta_{j(i),t}$ 를 계산한다.
 5. 계층적 움직임 파라미터 업데이트와 조인트 움직임 파라미터 업데이트를 업데이트 한다.
- $$\theta_{g(i),t} = \theta_{g(i),t-1}^* + \Delta\theta_{j(i),t} + \Delta\theta_{h(i),t}$$
6. 파티클 필터와 반복 최근점점 정합을 이용하여 자체 움직임 파라미터 업데이트 $\Delta\theta_{l(i),t}$ 를 얻는다.
 - i. 식 (14)를 이용하여 속도 \bar{v}_t 를 예측한다.
- For $p = 1, 2, \dots, P(i)$
- ii. 새 파티클 $\{\theta_{g(i),t}^{(p)}, 1\}$ 을 얻기 위해 $\{\theta_{g(i),t}^{(p)}, w_{g(i),t}^{(p)}\}$ 을 재생성한다.
 - iii. 식 (15)를 이용하여 예측된 위치 주위에 파티클을 생성한다.
 - iv. 식 (12)를 이용하여 관찰 확률에 따라 가중치를 업데이트 한다.
- End
- v. 각 파티클의 가중치를 정규화한다.
 - vi. 가장 큰 관찰 확률을 가지는 가장 좋은 파티클 $\hat{\theta}_{l(i),t}$ 을 선택하여 $\Delta\hat{\theta}_{l(i),t}$ 를 얻는다.
 - vii. 가장 좋은 파티클로부터 반복 최근점점을 정합을 이용하여 $\Delta\hat{\theta}_{l(i),t}$ 를 얻는다.
 - viii. 자체 움직임 파라미터 업데이트 $\Delta\theta_{l(i),t}^* = \Delta\bar{\theta}_{l(i),t} + \Delta\hat{\theta}_{l(i),t}$ 를 계산한다.
7. 자체 움직임 파라미터 업데이트 $\theta_{g(i),t}^* = \theta_{g(i),t} + \Delta\theta_{l(i),t}^*$ 를 업데이트한다.
- End
8. 입력 이미지가 더 이상 없으면 루프를 탈출한다.
- End

표 3 실험을 위한 파라미터들

항목	파라미터들	값들
카메라 (범블비2)	프레임 비율	10 fps
	이미지 사이즈	640 × 480
	초점 거리	245.130 mm
	베이스 라인	12.0023 cm
파티클	파티클 수	20~50

로 더 많은 각도를 움직인 오른쪽 팔은 제안하는 알고리즘만이 추적이 가능하였다. 기존의 알고리즘들은 움직임의 차가 너무 커 잘못된 최근점점(Closest point)을 선택해 국부적 최소점에 빠졌지만 제안하는 알고리즘은 파티클 필터의 확률적 탐색으로 잘못된 최근점점을 선택하지 않아 정확한 추적이 가능하였다. 즉 제안하는 알고리즘이



100 번째 프레임 101 번째 프레임 102 번째 프레임 103 번째 프레임

그림 3 기존 반복 최근점점 알고리즘과 제안하는 알고리즘의 추적 비교 영상

표 4 기존 반복 최근접점 알고리즘과 제안하는 알고리즘의 평균 반복 횟수와 평균 제곱 오차

	반복 최근접점 반복 횟수	평균 제곱 오차
기존의 반복 최근접점 알고리즘	14.0918	7.5217
제안하는 알고리즘	13.3513	6.8866

아주 빠른 움직임에 강인하다는 것을 보여준다.

표 4는 그림 3에 이용된 영상에서 기존의 알고리즘과 제안하는 알고리즘에서 반복 최근접점 평균 반복 횟수와 최종적으로 인간 신체 모델이 수렴되어 대응되는 3차원 관절 데이터들과의 평균 제곱 오차(Mean Squared Error)를 보여준다. 제안하는 알고리즘이 잘못된 최근접점을 선택하지 않아 반복 최근접점 알고리즘의 반복 횟수가 대략 0.7 정도 작게 되는 것을 알 수 있었고, 결과적으로 3차원 공간상에서 정합 오차를 0.7 cm 정도 낮추어 팔의 움직임 동작을 정확하게 추적함을 알 수 있었다.

5.2 일반적인 움직임을 가지는 인간 신체 움직임 추적 실험

실험에서는 일반적인 움직임이 발생했을 때 인간 신체 움직임을 얼마나 잘 추적하는가에 대한 추적 영상 결과를 보여준다. 실험 환경은 이전 실험과 같고 몸통, 팔, 다리 등을 움직이도록 한다. 그림 4는 몸통 회전과 팔의 움직임을 추적한 결과 영상이다. 이는 제안된 움직임 파라미터 업데이트가 상위 신체 구성 요소와 하위 신체 구성 요소 사이의 유기적인 움직임을 잘 표현할 수 있다는 것을 보여 준다. 그림 5는 팔로 웨이브 춤을 추고 있는 움직임을 추적한 결과 영상이다. 이 영상은 움직임 이력 정보를 이용한 파티클 필터가 위, 아래로 빠르게 움직이는 팔의 동적 상태를 잘 표현 한다는 것을 보여 준다. 마지막으로 그림 6은 Z방향으로 움직이는 팔과 다리를 추적한 결과 영상이다. 이는 제안된 알고리즘이 깊이 변화에 따른 움직임 또한 잘 추적할 수 있다는 것을 보여준다.

6. 결론

본 논문에서는 반복 최근접점(Iterative closest point) 알고리즘에 파티클 필터와 결합시켜 빠른 움직임을 가지는 신체 구성 요소들을 잘 추적할 수 있는 새로운 인간 신체 움직임 추적 알고리즘을 제안했다. 기존의 반복 최근접점 알고리즘은 실시간 인간 움직임 추적에 적합하지만 아주 빠른 움직임에 대해서는 잘못된 최근접점(Closest point)을 선택하기 때문에 극부적 최소점에 빠

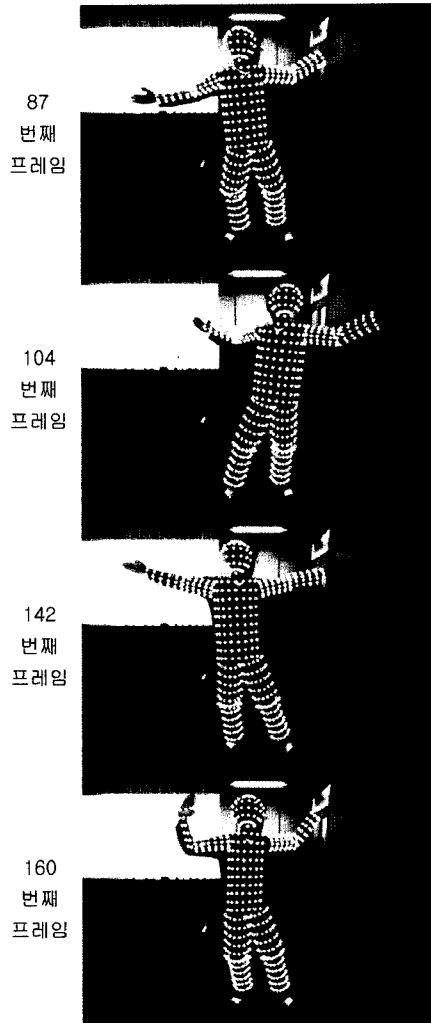


그림 4 몸통의 회전과 팔의 움직임 추적 결과

질 수 있다. 따라서 움직임 이력 정보에 이용한 파티클 필터를 반복 최근접점 알고리즘과 결합하여 인간 신체 움직임을 추적하도록 하였다. 제안하는 알고리즘은 팔이나 다리와 같이 큰 움직임을 가지는 신체 구성 요소들을 잘 추적할 수 있었다.

또한 인간의 구조를 계층적이라 가정하고 계층적 모델에 따른 계층적 움직임을 상위 신체 구성 요소들에서 구한 뒤 각 신체 구성 요소들에 적용하여 움직임을 추적 검색 공간을 줄였다.

실험 결과 제안하는 인간 신체 움직임 추적 알고리즘이 기존의 반복 최근접점 알고리즘보다 정확한 추적 성능과 빠른 수렴 속도를 가진다는 것을 보여 주었다.

제안된 알고리즘은 비전 정보만을 이용하여 인간 신체 움직임을 추적 가능할뿐더러 실시간 응용에도 적용

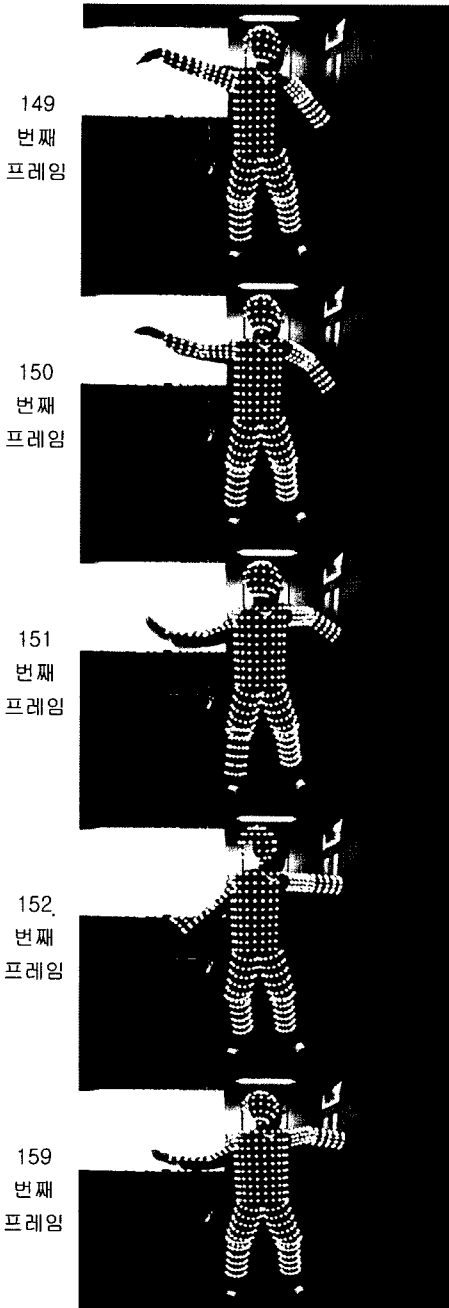


그림 5 팔로 웨이브 춤을 추고 있는 움직임 추적 결과

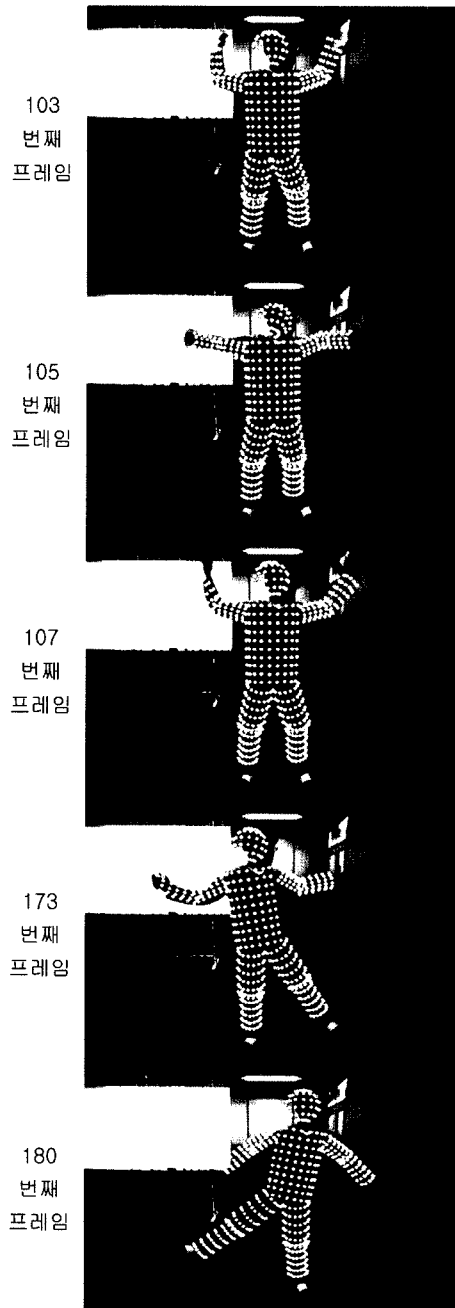


그림 6 Z축 방향으로 움직이는 팔과 다리 추적 결과

할 수 있다. 하지만 아직 복잡한 움직임이나 자체 가려짐 등의 상황에서는 정확하게 추적하는데 많은 어려움이 있다. 차후에는 좀 더 복잡하고 가려짐에도 강인한 실시간 인간 신체 움직임 추적 알고리즘으로 확장할 것이다.

참 고 문 헌 .

[1] C. Sminchisescu and B. Triggs, "Estimating articulated human motion with covariance scaled sampling," International Journal of Robotics Research, vol.22, no.6, pp.373-391, 2003.

[2] M. Lee and I. Cohen, "Proposal maps driven

- MCMC for estimating human body pose in static images," *Computer Vision and Pattern Recognition*, vol.2, pp.334-341, 2004.
- [3] R. Navaratnam, A. Thayananthan, P. Torr and R. Cipolla, "Hierarchical part-based human body pose estimation," *British Machine Vision Conference*, vol.1, pp.479-488, 2005.
- [4] J. Deutscher, A. Davision, and I. Reid, "Articulated body motion capture by annealed particle filtering," *Computer Vision and Pattern Recognition*, vol.2, pp.126-133, 2000.
- [5] J. MacComick, and M. Isard, "Partitioned sampling, articulated objects, and interface-quality hand tracking," *European Conference on Computer Vision*, vol.2 (1843), pp.3-19, 2000.
- [6] J. Caranza, C. Theobalt and M. Magnor, "Free-viewpoint video of human actors," *ACM SIGGRAPH*, pp.565-577, 2000.
- [7] D. Demirdjian, "Enforcing constraints for human body tracking," *Computer Vision and Pattern Recognition Workshop*, vol.9, pp.102-109, 2003.
- [8] K. Okuma, A. Taleghani, N. Freitas, J. Little, and D. Lowe, "A boosted particle filter: Multitarget detection and tracking," *European Conference on Computer Vision*, pp.28-39, 2004.
- [9] A. Doucet, J. Godsill and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statistics and Computing*, vol.10, no.3, pp.197-209, 2000.
- [10] S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol.50, no.2, pp.174-189, 2002.
- [11] P. Besl, and N. MacKay, "A method for registration of 3-d shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.14, no.2, pp.239-256, 1992.
- [12] Y. Chen, and G. Medioni, "Object modeling by registration of multiple range images," *Image and Vision Computing*, vol.10, no.3, pp.145-155, 1991.
- [13] C. Bregler, and J. Malik, "Tracking people with twists and exponential maps," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.8-15, 1998.
- [14] P. Felzenszwalb, and D. Huettenlocher, "Pictorial structures for object recognition," *International Journal of Computer Vision*, vol.61, no.1, pp.55-79, 2005.
- [15] M. Lee, and R. Nevatia, "Human pose tracking using multi-level structured models," *European Conference on Computer Vision*, vol.3, pp.368-381, 2006.
- [16] X. Ren, A. Berg, and J. Malik, "Recovering human body configurations using pairwise constraints between parts," *International Conference and Computer Vision*, vol.1, pp.824-831, 2005.



김 대 환

2002년 2월 동국대학교 컴퓨터 및 멀티미디어공학과(학사). 2004년 2월 POSTECH 컴퓨터공학과(석사). 2005년 3월~현재 POSTECH 컴퓨터공학과(박사과정). 관심분야는 컴퓨터 비전, 패턴 인식



김 효 정

1996년 2월 중앙대학교 전산학과(학사) 2002년 2월 서울대학교 컴퓨터공학과(석사). 2002년~현재 KT 연구소 근무. 관심분야는 컴퓨터 비전



김 대 진

1982년 2월 연세대학교 전자공학과(학사). 1984년 2월 KAIST 전기 및 전자공학과(석사). 1992년 8월 Syracuse University, Electrical and Computer Engineering(박사). 1984년 3월~1986년 12월 한국방송공사 기술연구소. 1992년 3월~1999년 6월 동아대학교 컴퓨터공학과 부교수. 1999년 7월~현재 POSTECH 컴퓨터공학과 교수. 관심분야는 지능시스템, 인간 컴퓨터 상호작용