

다중 관측 모델을 적용한 입자 필터 기반 물체 추적

Visual Object Tracking based on Particle Filters with Multiple Observation

고형승, 조용근, 강훈

중앙대학교 전자전기공학부

Hyeung Seong Ko, Yong-Gun Jo, Hoon Kang

School of Electrical and Electronic Engineering, Chung-Ang University

E-mail : hihs@sirius.cie.cau.ac.kr

ABSTRACT

본 논문에서는 CONDENSATION 알고리즘을 이용하여 입자 필터(particle filter)에 기반한 물체 추적 알고리즘을 제안한다. 입자 필터는 조건 확률 전파 모델(Conditional Density Propagation)인 베이지안(Bayesian) 추론 규칙을 적용하는 추적 구조를 갖고 있기 때문에 다른 어떤 종류의 추적 알고리즘보다 뛰어난 성능을 보인다.

논문에서는 실험 결과를 통해, 외곽(contour) 추적 입자 필터가 복잡한 환경 속에서 강인한 추적 성능을 나타냄을 증명한다.

Key words : Object Tracking, Particle Filters, Active Contour

1. 서론

물체 추적은 컴퓨터 비전에서 가장 흥미로운 분야 중 하나이다. 단지 물체의 형태나 칼라 추적하는 것뿐만 아니라 물체의 인식 또는 식별하는 것에도 관련이 있다.

최근 입자 필터(particle filter)에 의한 물체 추적 접근 방법[1][2][3]이 복잡한 환경 속을 움직이는 물체를 추적하는데 효과적이라는 것이 증명된 이래로, 컴퓨터 비전분야에서 물체 추적은 흥미로운 연구 과제가 되었다.

입자 필터는 사전분포(prior distributions)와 사후분포(posterior distributions)와 같은 베이지안 조건 확률(Bayesian conditional

probabilities)에 기반을 두고 있다.

로봇 비전에서 입자 필터 혹은 CONDENSATION(CONDitional DENSity PropagATION)알고리즘[3][4]은 로봇의 지능을 향상시킬 수 있다. 왜냐하면 추적 정보는 결국 MIQ(machine intelligent quotient)을 향상시키기 때문이다. 또한 비선형 동역학(nonlinear dynamics)이나, non-Gaussian 밀도, 가려짐(occlusion), 조명 변화와 같은 불확실성은 로봇 비전에서 입자 필터의 중요성을 더욱 부각시킨다.

이와 관련하여, 본 논문은 다중 관측 모델(multiple observation models)을 적용하여 입자 필터에 기반한 강인한 물체 추적 알고리즘을 제안한다.

비전 센서로, 웹캠(webcam)을 사용하였다. 추적 대상인 사람이 복잡한 환경 속에서 움직인다는 것을 가정하였고, 얼굴 혹은 손을 추적

본 연구는 과학기술부 인간기능 생활지원 지능로봇 기술개발 사업(과제번호 1-1) 지원으로 수행되었습니다.

하는 것이 연구의 목적이다. 이를 위해 입자 필터 즉, CONDENSATION 알고리즘을 적용하였다. 입자 필터는 확률적 접근 방법의 범주로서, 가중치가 부여된 N개의 입자는 $\{X_t^{(i)}, \pi_t^{(i)}\}_{i=1, \dots, N}$ 으로 표현되며, $X_{t-1}^{(i)}$ 는 시간 t에서 i^{th} 입자의 상태를, $\pi_t^{(i)}$ 는 가중치를 의미한다. 그러므로 분석적 해법을 유도하는 대신, 임의의 형태를 띠는 확률 분포에 기초한 몬테카를로(Monte-Carlo) [1] 시뮬레이션을 수행하였다. $X_t^{(i)}$ 의 상태로는 움직이는 물체의 위치, 속도, 회전각 등이 될 수 있다. 측정값 Z_t 는 image contrast, digital image subtraction, edge-detected silhouette, 2D 또는 3D contours, RGB 또는 HSV color 등이 될 수 있다. 추론 모델은 조건 사후 밀도(conditional posterior density)로 기술되는 베이저안 추론 규칙을 이용하며 다음과 같다.

$$\begin{aligned}
 P(X_t, Z_t) &= \kappa P(Z_t | X_t) P(X_t | Z^{t-1}) \\
 &= \kappa P Z_t, X_t \int_{X_{t-1}} P X_t, X_{t-1} P X_{t-1}, Z^{t-1} \quad (1) \\
 &= \kappa P Z_t, X_t \sum_{i=1}^N \pi_{t-1}^{(i)} P X_t, X_{t-1}^{(i)} \\
 &= \kappa P Z_t, X_t, q, X_t
 \end{aligned}$$

여기서 X_t 는 시간 t에서의 상태를 의미한다. $Z^t = \{Z_1, \dots, Z_t\}$ 는 시간 t에서의 관측도(observation)를 의미하며, κ 는 비례상수, $P(X_t | X_{t-1})$ 는 마코프 체인 운동 모델(Markov-chain motion model), $P(Z_t | X_t)$ 는 측정치 모델(measurement model), 그리고 $\pi_{t-1}^{(i)}$ 는 입자 $X_{t-1}^{(i)}$ 의 가중치를 의미한다. 여기서 $q(X_t) = \sum_{i=1}^N \pi_{t-1}^{(i)} P(X_t | X_{t-1}^{(i)})$ 는 몬테카를로 근사적분 방식으로 유도되는 N개 입자 $X_{t-1}^{(i)}$ 의 제안 분포(proposal distribution)이다. 사후 밀도 $P(X_{t-1} | Z^{t-1})$ 는 입자군 $\{X_{t-1}^{(i)}, \pi_{t-1}^{(i)}\}$ 에 의해 재귀적으로 근사 되며, 이 입자들은 $\pi_{t-1}^{(i)} = P(Z_{t-1} | X_{t-1}^{(i)})$ 에 의해 가중치를 부여 받는다.

입자 필터 알고리즘은 3가지 단계, 즉 샘플링(sampling), 예측(predicting), 측정(measuring)으로 구성된다. 알고리즘의 기본 원리는 샘플링과 예측 단계에서의 사전 밀도와 측정단계의 사후 밀도 사이의 조건 확률 전파(conditional probability propagation)이다. 따라서, 사전, 사후 조건 밀도는 재귀적 방법으로 계산된다.

입자 필터 알고리즘은 CONDENSATION

[4][7], Kalman 입자 필터(KPF)[8], 그리고 Unscented Kalman 입자 필터(UPF)[9]로 구성될 수 있다. KPF와 UPF는 Gaussian과 non-Gaussian의 상태 평가를 이용하는 예측 모델이 추가된다. 추적을 위한 입자 필터의 세가지 갱신 단계는

아래와 같이 요약될 수 있다.

- 1) 샘플링(Selection)
- 2) 예측(Particle Dynamics)
- 3) 측정(Output)

관측 모델로 피부색의 점 데이터(point cloud)와 Active contour[5]의 B-spline snake을 이용하였다.

II. 본 론

2 물체 추적을 위해 다중 관측 모델을 적용한 입자 필터.

확률적 전파의 관점에서 볼 때, CONDENSATION 알고리즘은 몇 가지 특징을 지닌 움직이는 물체를 추적하기 위한 방법을 제시한다. 만약 움직이는 물체의 특징의 개수가 K 라 하면, i^{th} 입자의 상태 혹은 contour는 $X_t^{(i)} \in R^K$ 가 된다. 기본 추적 알고리즘은 다음과 같다.

CONDENSATION Tracking Algorithm

주어진 입자 혹은 contour 군 $\{X_{t-1}^{(i)}, \pi_{t-1}^{(i)}\}$ 으로부터, 새로운 입자 혹은 contour 군 $\{X_t^{(i)}, \pi_t^{(i)}\}_{i=1, \dots, N}$ 을 구하고, 가중치 $\pi_0^{(i)} = 1/N$ 와 누적 분포(cumulative distribution) $c_0^{(i)} = i/N$ 을 초기화 한다.

- 1) 다음의 방법으로 샘플 $S_t^{(i)}$ 을 선택한다.
 - a) 균일하게 분포된 랜덤 수(random number) $r \in [0, 1]$ 을 발생시킨다.
 - b) $c_{t-1}^{(j)} \geq r$ 인 가장 작은 j 을 찾는다.
 - c) $S_t^{(i)} = X_{t-1}^{(j)}$ 을 설정한다.
- 2) $S_t^{(i)}$ 을 얻기 위해 샘플링으로부터 예측한다.

$$P(X_t | X_{t-1} = S_t^{(i)}) \quad (2)$$
- 3) 물체의 특징을 측정하고, 가중치를 갱신한다.

$$\pi_t^{(i)} = P(Z_t | X_t = \hat{S}_{t-1}^{(i)}) \quad (3)$$

$\sum_{i=1}^N \pi_t^{(i)} = 1$ 이 되도록 만들고 누적 분포 $c_t^{(i)}$ 을 계산한다. $\{X_t^{(i)}, \pi_t^{(i)}\}$ 을 갱신한다.

다중 관측 모델에서 결합 혹은 분리 연산자 AND, OR은 다중 물체 특징을 결합하기 위해 사용될 수 있다. 아래 식에 의해, 모든 입자 혹은 contour가 갱신된 후 추적된 위치를 평가한다.

$$E[g(X_t)] = \sum_{i=1}^N \pi_t^{(i)} \cdot g(\hat{S}_{t-1}^{(i)}) \quad (4)$$

상태 정보는 입자의 군집 혹은 contour가 되도록 입자의 평균값(mean value)나 역 공분산 행렬(inverse covariance matrix)의 신뢰 영역(confidence region)을 선택 한다. 예를

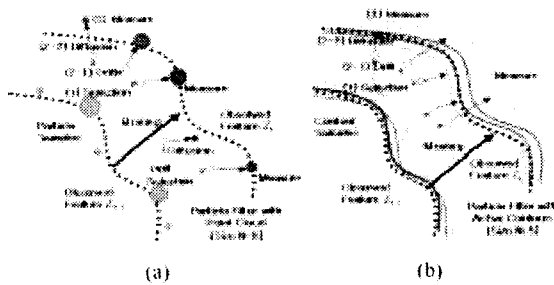


그림1. a)5개의 점 데이터 입자 샘플과 측정 단계. b)5개의 변형 가능한 외각선 입자 샘플

들어, 상태(state)의 한 요소인, 추적 물체의 평균 위치는 다음 물체의 무게 중심 추정을 위한 갱신을 위해 측정 된다.

2.1 샘플링(Selection).

샘플링 단계에서, N 개에 입자 샘플은 $[0,1]$ 의 범위에서 임의로 선택된다. 이러한 선택 방법은 진화 알고리즘(GA)의 룰렛휠 선택(roulette-wheel selection)방식과 비슷한 방법이다. 예를 들어, 대상 물체의 평균 위치는 다음 단계의 대상 물체 중심을 갱신하기 위해 평가 된다.

2.2 예측(Particle Dynamics).

예측단계에서 각 입자 혹은 contour은 다음 식에 의해 유도 된다

$$S_t^{(i)} = f(S_{t-1}^{(i)}) + \epsilon_{t-1} \quad (5)$$

보통 $f(\cdot)$ 는 서로 다른 위치에 있는 각 입자들의 비선형 drift을 표현하며, ϵ_t 는 백색 잡음(white noise)로 모델링 된 각 입자의 독립적 확산(diffusion)이다. 그러므로 잡음은 가우시안 분포(Gaussian distribution)를 나타낸다는 가정으로부터 $\hat{S}_{t-1}^{(i)}$ 을 평가 하기 위한 예측 모델(prediction model)로서, 확장된 칼만 필터(extended Kalman filter(EKF))가 사용된다. 입자나 contour는 그림 1과 같이 관측 분포(observation distribution)에서 가장 높은 가능성(likelihood) 쪽으로 유도되는 방식을 갖는 관측 모델이다.

2.3 측정(Output)

각 입자 혹은 contour $X_t^{(i)}$ 에 대해 새로운 가중치 $\pi_t^{(i)}$ 을 계산하기 위한 식은 다음과 같다.

$$\pi_t^{(i)} = P(Z_t | X_t^{(i)}) \quad (6)$$

이미지 특징 점에 대한 외곽 이미지에서 $\pi_t^{(i)} = 0$ 이다.

2.4 다중 관측 모델.

논문에서는 얼굴과 손을 추적하기 위한 2가지 특징을 선택한다. 하나는 차감(subtracted)

이미지의 contrast 정보이고, 다른 하나는 HSV 칼라 모델에서의 피부색 정보이다. 차감 이미지 ΔP_i 는 다음 식으로 측정한다.

$$\Delta P_i(x, y) = |I_t(x, y) - I_{t-\Delta}(x, y)| \quad (7)$$

여기서 $I_t(x, y)$ 는 t 시간 $[0,1]$ 의 범위를 갖는 $(x, y)^{th}$ 에서의 이미지 픽셀값 이고, $I_{t-\Delta}$ 는 시간 $t-\Delta$ 에서 나타난 이미지이다. 피부색 측정을 위한 hue 와 saturation 정보는 가우시안 밀도 함수(Gaussian density function)을 계산해서 구한다.

$$C_i(x, y) = \exp\left(\frac{-(h(x, y) - h_R(x, y))^2 - (s(x, y) - s_R(x, y))^2}{2\sigma_{hs}^2}\right) \quad (8)$$

여기서 σ_{hs} 는 HS 특징에 대한 표준 편차(standard deviation)이다. $C_i(x, y)$ 는 $(x, y)^{th}$ 에서의 피부색 특징을 의미하며, $h(x, y)$ 와 $s(x, y)$ 는 컨트롤 포인트 (x, y) 에서의 hue와 saturation 값이다. $h_R(x, y)$ 와 $s_R(x, y)$ 는 초기 입자 군에 대한, 기준 hue과 saturation 값이다. 이제 입자들과 변형 가능한 contour의 B-spline snakes를 위해 관측 모델을 고려해 보자.

· 입자로서의 점 데이터(Point Cloud):

-차감 픽셀과 피부색으로 표현되는 i^{th} 입자 점 $(x_p, y_p) \in \Delta P_i$ 에 대해 사전(prior) 가중치는 OR 연산자를 사용하여 다음과 같이 계산된다.

$$w_t^{(i)} = \sum_{(x_p, y_p) \in \Delta P_i} \{\Delta P_i(x_p, y_p) + \lambda C_i(x_p, y_p)\} \quad (9)$$

-하나의 입자에 대한 상태 벡터는 시간 t 에서 $X_t^{(i)} = [x_p(t), y_p(t), x_p(t-1), y_p(t-1)]^T$ 으로 정의되며, 입자의 운동은 다음과 같이 표현된다.

$$X_t = AX_{t-1} + Bv_t \quad (10)$$

여기서 A, B 는 동적 특성을 나타내는 행렬이고, v_t 는 확산 공분산 행렬(diffusion covariance matrix)로 표현되는 백색 잡음이다.

· 입자로서의 변형 가능한 contours:

-차감 이미지의 $(x_p, y_p) \in \Delta P_i$ 을 i^{th} contour의 컨트롤 점이라 하고, 그에 대응하는 i^{th} 피부색 컨트롤 점을 $(x_c, y_c) \in C_i$ 라 하면, 사전 가중치는 AND 연산자를 이용하여 다음과 같이 계산 된다

$$w_t^{(i)} = \sum_{(x_p, y_p) \in \Delta P_i} \Delta P_i(x_p, y_p) \sum_{(x_c, y_c) \in C_i} \lambda C_i(x_c, y_c) \quad (11)$$

-하나의 입자에 대한 상태 벡터는

$X_t^{(i)} = [\bar{x}_p(t), \bar{y}_p(t), \theta(t), \alpha(t), \bar{x}_p(t-1), \bar{y}_p(t-1), \theta(t-1), \alpha(t-1)]^T$ 로 정의되며, ΔP_i 에서 컨트롤 점의 중심이 된다. θ, α 는 회전각과 B-spline의 척도 요소

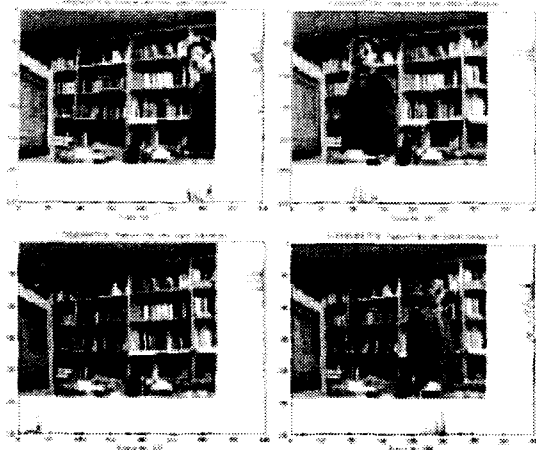


그림 2. 차감이미지와 피부색을 이용하여 입자 필터로 구현된 얼굴 추적 실험

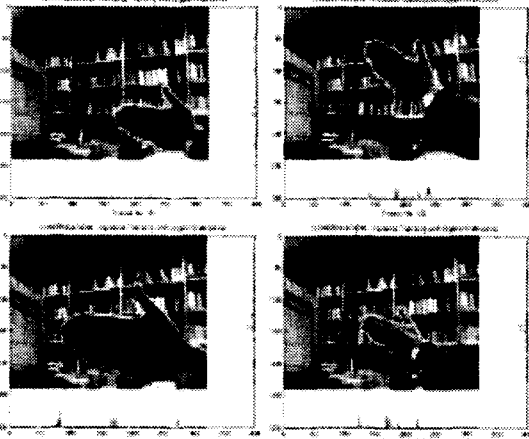


그림 3. 차감 이미지만을 이용하여 변형 가능한 외곽선으로 구현된 손 추적 실험

(scaling factor)다. λ 는 2개의 독립된 물체 특징 사이의 균형 계수(balancing coefficient)이다. 또한 사후 π_i 는 재 샘플링(resampling)과 정규화(normalization)에 의해 계산되며, 다음과 같다.

$$\pi_i = f_N[f_R(f_N(w_i))] \quad (12)$$

여기서 f_N, f_R 는 정규화 함수와 재 샘플링 함수이고, w_i 는 식(9),(11)로부터 구해진다. i^{th} 의 입자의 변형 가능한 contour는 아핀 형태 공간(affine shape-space)에서 정의되는 형태로 표현되며 다음과 같다.

$$Q' = WY_i + Q$$

$$W = \begin{bmatrix} 1 & 0 & Q_x & 0 & 0 & Q_y \\ 0 & 1 & 0 & Q_y & Q_x & 0 \end{bmatrix} \quad (13)$$

$Y_i = [\bar{x}_p, \bar{y}_p, \cos\theta - 1 + \alpha, \sin\theta, \sin\theta]^T$ 여기서 W 는 형태 행렬(shape matrix), Y_i 는



그림 4. 차감 이미지 정보와 피부색 정보를 이용하여 변형 가능한 외곽선으로 구현된 얼굴 추적 실험(N=100)

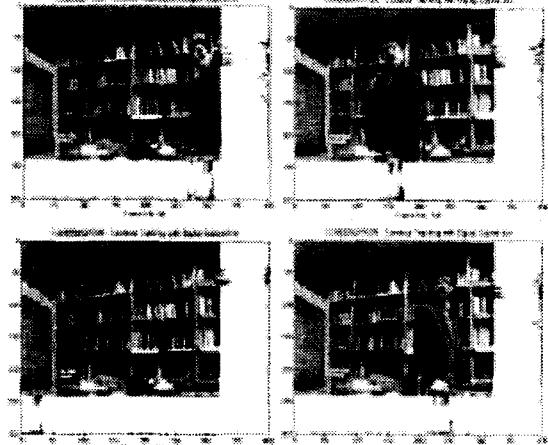


그림 5. 차감 이미지 정보와 피부색 정보를 이용하여 변형 가능한 외곽선으로 구현된 얼굴 추적 실험(N=50)

X_i 로부터 계산된 형태 행렬 벡터, 그리고 $Q, Q' = [Q'_x, Q'_y]^T$ 는 기준 컨트롤 점 벡터이다. 여기서 갱신 컨트롤 점 벡터 Q' 는 다음과 같다.

$$Q' = [x_{p,1}(t), \dots, x_{p,n}(t); y_{p,1}(t), \dots, y_{p,n}(t)]^T \quad (14)$$

여기서 n는 B-spline snake에서의 컨트롤 점의 총 개수를 나타낸다.

3 입자 필터를 이용한 물체 추적 실험 결과.

실험에서 입자 필터를 이용한 물체(얼굴, 손)를 추적하기 위해 3가지 경우의 다중 관측 모델이 적용되었다.

3.1 입자 필터에 기반한 강인한 얼굴 추적.

입자 데이터(particle cloud)에 의한 얼굴 추적 알고리즘을 시뮬레이션(simulation)하였다. Contrast 정보와 피부색 정보를 모두 사용하기

위해 입자 관측 모델(9)과 $N=50$ 인 입자 개수를 사용하였다. 여기서 $\lambda = 1/N$ 이고, 재 샘플링을 위해 $\pi^{(i)}$ 의 임계치는 $2/N$ 으로 하였다. 확산을 위한 표준 편차는 $\sigma_x = \sigma_y = 30$ 이고, hue와 saturation은 $\sigma_{hs} = 0.05$ 으로 설정하였다. 물체 특징 정보는 차감 이미지와 피부색 정보를 사용하였다. 그림 2는 점 데이터(point cloud)을 이용한 입자 필터를 사용하여 물체를 추적하는 모습이다.

3.2 변형 가능한 Active Contours을 이용한 손 추적.

Contour을 이용하여 손 추적 알고리즘을 시뮬레이션(simulation)하였다. 입자는 개수는 $n=200$ 이고, 변형 가능한 관측 모델(9)을 사용하였다. 그러나 이 경우, $\lambda = 0$ 으로 하고 contrast 특징 정보 하나만을 사용하였다. B-spline snake의 컨트롤 점의 개수는 $n=45$ 이고, 재 샘플링을 위해 $\pi^{(i)}$ 의 임계치는 $1.5/N$ 로 하였다. 확산을 위한 표준 편차는 $\sigma_x = \sigma_y = 5$ 이고, 회전각과 척도 요소는 $\sigma_\theta = 3(\text{deg.}), \sigma_\alpha = 0.05$ 로 설정하였다. 물체 특징 정보는 차감 이미지만을 사용하였다. 그림3은 변형 가능한 contour의 B-spline snake을 이용하여 손을 추적 하는 모습이다.

3.3 피부색과 변형 가능한 Active Contours을 이용한 얼굴 추적.

마지막으로 contour와 점 데이터(point clouds)을 이용하여 사람 얼굴 추적 알고리즘을 시뮬레이션(simulation)하였다. 입자의 개수는 $N=50$ 혹은 $N=100$ 이고, contrast 정보와 피부색 정보를 함께 이용하기 위해 관측 모델을 이용하였다(11). 여기서 컨트롤 점의 개수는 $n=21$ 이고, 재 샘플링을 위해 $\pi^{(i)}$ 의 임계치는 $1.5/N$ 으로 하였다. 확산을 위한 표준 편차는 $\sigma_x = \sigma_y = 5$ 이고, 회전각과 척도 요소는 $\sigma_\theta = 1(\text{deg.}), \sigma_\alpha = 0.01$ 으로 설정하였다. hue와 saturation은 $\sigma_{hs} = 0.05$ 으로 설정하였다. 그림 4,5는 이전 경우와 달리 적은 수의 입자 군을 이용하여 물체를 추적하는 모습이다.

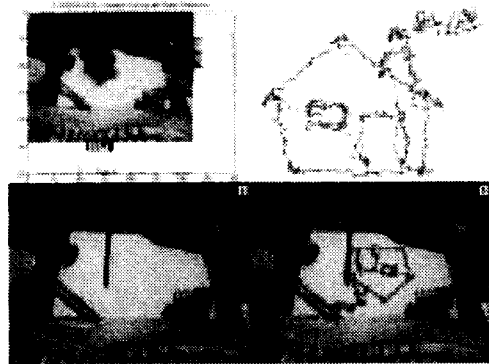


그림 6. 다중 관측 모델을 적용한 입자 필터를 이용하여 그림 그리기 추적 결과 ($N=200$)

III. 결 론

본 논문에서는 입자 데이터와 contour tracking, 차감 이미지 정보, 피부색과 같은 다중 관측 모델로 확장된 CONDENSATION 알고리즘을 이용하여 사람의 얼굴 혹은 손을 추적하였다. 사전 밀도(prior density)와 사후 확률(posterior density)로 측정된 다중 물체의 특징은 추적 대상 물체의 위치, 척도 요소, 기준 좌표에 대한 회전정보 등과 같은 상태 평가를 하기 위해 전파 된다. 실험 결과를 통해, 본 논문에서 제안하고 있는 다중 관측 모델의 입자 필터는 복잡한 환경에서의 비선형 운동과 같은 불확실한 환경 속에서 강인한 물체 추적을 위해 사용 될 수 있음을 증명하였다.

본 논문을 통해 다중 관측 모델을 적용한 입자 필터의 장점을 사람 얼굴 혹은 손 추적을 통해 입증하였다. 첫째 입자 필터를 통해 비선형 운동과 다중 관측 모델을 결합하여 물체를 추적할 수 있었고, 둘째 다중 최빈 사후 밀도(multi-modal posterior density)형태를 나타내는 복잡한 환경 속 일지라도, 강인한 추적 성능을 나타낼 수 있다. 마지막으로, 그림을 그리는 것을 추적한 결과를 보이고 있는 그림6과 같이, 입자 필터를 통해 2D 혹은 3D contour 정보를 사용하여 high-level 추론이 가능하다는 것을 증명 하였다. 향후 과제로, 다중 물체 입자 필터를 이용한 다중 물체 추적 알고리즘으로 확대 하려 한다.

IV. 참고문헌

- 1] J.M. Hammersley and K.W. Morton, "Poor man's monte carlo," J. of the Royal Statistical Society B, vol. 16, pp.23-38, 1954.

- [2] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," *IEEE Proc. F*, vol. 140, no.2, pp. 107-113, 1993.
- [3] M. Isard and A. Blake, "Contour tracking by stochastic propagation of conditional density," in *Proc. 4th European Conf. on Computer Vision (ECCV)*, pp. 343-356, April 1996.
- [4] M. Isard and A. Blake, "Condensation-conditional density propagation for visual tracking," *Int. J. Computer Vision*, vol 29, no. 1, pp. 5-28, 1998.
- [5] A. Blake and M. Isard, *Active Contours*, London: SpringerVerlag, 1998.
- [6] A. Popoulis, *Probability and Statistics*, New York: PrenticeHall, Inc., 1990.
- [7] M. Isard and A. Blake, "Icondensation: Unifying low-level and high-level tracking in a stochastic framework," in *Proc. 5th European Conf. on Computer Vision (ECCV)*, pp. 893-908, April 1998.
- [8] A. Blake, M. Isard, and D. Reynard, "Learning to track the visual motion of contours," *J. Artificial Intelligence*, vol 78, pp. 101-134, 1995.
- [9] S.J. Julier and J.K. Uhlmann, "A new extension of kalman filter to linear system," in *Proc. Aerosense: 11th Int. Symp. on Aerospace/Defense Sensing, Simulation, and Control*, Orlando, FL, pp. 182-193, 1997.
- [10] J. MacCormick and A. Blake, "Partitioned sampling, articulated objects and interface-quality hand tracking," in *Proc. 7th European Conf. on Computer Vision (ECCV)*, April 2000.