

비디오 영상내의 사람 추적을 위한 강인한 멀티-파트 추적 방법

김 태 현, 김 진 율
수원대학교 전자공학과
전화 : 031-220-2531 / 핸드폰 : 019-485-4518

A Robust Multi-part Tracking of Humans in the Video Sequence

Tae-Hyun Kim, Jin-Yul Kim
Dept. of Electronic Engineering, The University of Suwon
E-mail : thkim@mail.suwon.ac.kr

Abstract

We presents a new algorithm for tracking person in video sequence that integrates the meanshift iteration procedure into the particle filtering. Utilizing the nice property of convergence to the modes in the meanshift iteration we show that only a few sample points are sufficient, while in general the particle filtering requires a large number of sample points. Multi-parts of a person is tracked independently of each other based on the color. Then, the similarity against the reference model color and the geometric constraints between multi-parts are reflected as the sample weights. Also presented is the computer simulation results, which show successful tracking even for complex background clutter.

I. 서론

비디오 영상속의 물체 추적은 최근 HCI(human computer interface), 자동 보안 출입 시스템, 화상 회의 등을 포함하는 광범위한 응용 분야에서 그 필요성이 증가 하고 있다.

차영상을 이용한 기법[1]은 개념적으로 간단하여 널리 사용되어져 왔으며 실시간 구현이 용이하다. 그러나 카메라가 움직이는 상황에서는 주위 배경의 움직임이 고려되어야 하므로 전역 움직임 보상의 기법[2]이

수반되어야 한다. Baumberg[3]는 사람의 윤곽선(edge)을 B-spline으로 근사화하고 시간에 따른 윤곽선의 움직임을 Kalman filter로 추적하였다. 그러나 시스템의 관측모델(observation model)을 가우시안으로 단순화하였기 때문에 클러터(clutter)가 존재할 경우 많은 문제점이 발생하였다. 이를 극복하기 위해 Isard[4]는 particle filtering 기법의 일종인 condensation 알고리즘을 제시하여 non-gaussian, multi-modal의 시스템에 적용하였다. 그러나 꼭 학습과정(training)을 거쳐야 하고 관측 모델을 윤곽선에 의존하기 때문에 컬러(color) 정보를 활용하지 않고 있다. S.T. Birchfield [5]는 얼굴의 컬러와 윤곽선을 동시에 사용하는 얼굴추적 방법을 제시하였다. 그러나 일정 영역을 탐색하며 관측 모델을 추정하기 때문에 비효율적이며 윤곽선을 너무 단순화 하여 클러터 존재 시 얼굴추적에 실패하는 문제점을 갖고 있다. D. Comaniciu[6][7][8]는 기존의 방법들과는 달리 일정 범위를 탐색하지 않고 확률적 접근 방법을 통해서 몇 번의 반복 연산 수행으로 참조모델과 유사도가 높은 후보모델을 추정할 수 있는 계산량이 적은 MS(meanshift) 기법을 제시 하였다. 최근, P. Perez[9]는 particle filtering 기법에 관측확률(observation probability)로 [6]에서 제안되었던 Bhattacharyya 계수를 채용한 방법을 제시하였다.

본 논문에서는 particle filtering에 기반하여 meanshift iteration 기법을 결합시킨 새로운 추적 기법을 제안한다. 또, 추적 대상을 멀티 파트(multi-part)로 나누고 각각의 컬러 정보와 상호간의 기하학적 위치를 관측 확률로 모델링(modeling)하여 클러터가 존재하는 복잡한 배경에서도 기존의 알고리즘보다 더 강인한 추

적이 가능함을 실험을 통해 검증한다.

II. 기존 알고리즘

2.1 Particle filtering

Particle filtering[9][4]은 복잡한 클러터를 가진 환경에서 물체를 추적하기 위해 최근에 많이 사용되고 있다.

Particle filtering 알고리즘에서는 확률분포를 근사하기 위하여 가중치를 갖는 샘플들의 집합을 사용한다. 시간 $t-1$ 에서 샘플 집합이

$S_{t-1} = \{(s_{t-1}^{(n)}, \pi_{t-1}^{(n)}) \mid n = 1, \dots, N\}$ 이라 하면, 각 샘플 $s_{t-1}^{(n)}$ 은 추적하려는 물체가 있는 가능성이 높은 위치를 나타내며 이때 해당되는 관측확률은 $\pi_{t-1}^{(n)}$ 으로

근사된다. 이때 $\sum_{n=1}^N \pi_{t-1}^{(n)} = 1$ 이다. 시간 t 에서는

먼저, S_{t-1} 의 각 샘플 $s_{t-1}^{(n)}$ 로부터 관측확률값 $\pi_{t-1}^{(n)}$ 에 비례하여 그 샘플을 중복적으로 선택하되 이때 선택되는 전체의 샘플 수가 N 이 되도록 하여 새로운 샘플 집합 \tilde{S}_{t-1} 을 구성한다(Select 단계). 이렇게 얻은 각 샘플은 시스템의 dynamics model에 의하여 새로운 샘플 위치 $s_t^{(n)}$ 로 전파(propagate)되며(Propagate 단계), 전파된 각 샘플은 그 점에서 측정된 관측확률에 비례하는 가중치 $\pi_t^{(n)}$ 가 주어지며 새로운 샘플 집합 $S_t = \{(s_t^{(n)}, \pi_t^{(n)}) \mid n = 1, \dots, N\}$ 을 구성한다.

(Observe 단계). 다음 시간 단계에서는 위의 Select 단계부터 반복하게 된다. 각 시간 단계에서, 샘플 집합 S_t 로부터 mean state를 다음 식에 따라 유추할 수 있다. 즉, $E[S_t] = \sum_{i=1}^N \pi_i^{(n)} s_i^{(n)}$ 이다(Estimate 단

계). 기존의 연구에서는 particle filtering의 Observe 단계에 사람의 얼굴등의 윤곽선에 기반한 관측확률 모델[4]을 사용하거나, [9]에서 처럼 MS 알고리즘[6,7,8]에서 사용되었던 참조모델과 후보모델사이의 컬러 유사성의 정도를 나타내는 Bhattacharyya 계수에 기반한 관측확률을 사용하였다.

2.2 Meanshift 알고리즘

Meanshift iteration 기법은 주어진 후보 탐색영역의 전부를 탐색(exhaustive search)하는 대신 추적 물체가 존재하는 확률이 높은 위치(모드(mode))에 수렴하도록 하는 일련의 반복 연산을 사용하는 방법으로 참조모델과 후보모델사이의 컬러 유사성의 정도를 나타내는 Bhattacharyya 계수가 최대가 되는 위치를 추적할 수 있으며 계산량이 적어 실시간 추적에 적합하다. 그러나 이 방법은 컬러 히스토그램을 기반으로 하기 때문에 추적 대상과 비슷한 컬러의 배경 클러터가 분

포할 경우 국부 최적해에 고착(locked)될 수 있으며, 추적 대상의 사람이 여러 명이고 서로 다른 얼굴이 검출 때 그전에 추적하던 얼굴 대신 다른 얼굴을 추적하게 되는 문제점을 지니고 있다.

III. 제안하는 알고리즘

3.1 전체 구성도

본 논문에서는 particle filtering 알고리즘과 meanshift 알고리즘의 장점을 결합한 새로운 방법을 제시한다. 기존 particle filtering 알고리즘에서 추적의 정확성을 유지하기 위하여 많은 수의 샘플을 사용해야 되는데 비하여 제시된 방법에서는 효율적인 meanshift iteration의 추가적 사용에 의해 적은 수의 샘플을 사용하여도 관측확률이 높은 state에 도달할 수 있도록 하였다. 제시된 방법에서는 사람의 추적을 위하여 머리와 얼굴 등에 대하여 멀티 파트 추적을 수행하며 기존의 논문과는 달리 멀티 파트의 각 부분이 갖는 Bhattacharyya 계수의 값 이외에 각 멀티 파트들 간의 물리적 거리를 함께 고려하는 새로운 관측확률을 사용하여 추적의 정확도를 높였다.

제안하는 추적 알고리즘은 다음과 같다.

(0) 초기화 : 추적하려는 각 파트 P_j , ($j=1, 2$)에 대하여 참조 모델을 계산한다.

$$q_u^{P_j} = C \sum_{i=1}^n k(\|x_i^{P_j}\|^2) \delta[b(x_i^{P_j}) - u]$$

(1) 추적하려는 각 파트 P_j 에 대하여 다음의 단계를 독립적으로 수행한다.

(1.1) Select : 파트 P_j 에 대한 샘플 집합

$S_{t-1}^{P_j} = \{(s_{t-1}^{P_j(n)}, \pi_{t-1}^{P_j(n)}) \mid n = 1, \dots, N_{P_j}\}$ 으로부터 각 샘플로부터 관측 확률값에 비례하여 새로운 샘플 집합 $\tilde{S}_{t-1}^{P_j}$ 을 구성한다

(1.2) Propagate : 각 샘플을 시스템의 dynamics model(2차 ARP(auto-regressive processes))에 의하여 새로운 샘플 위치로 전파(propagate)한 후, 이 위치를 초기치로 하여 meanshift iteration에 의해 Bhattacharyya 계수가 최대가 되는 새로운 위치 $s_t^{P_j(n)}$ (즉, 모드)로 수렴 시킨다.

(1.3) Observe : 수렴된 모드의 위치에서 아래 3.2에 설명된 방법에 의해 관측확률을 계산하여 가중치 $\pi_t^{P_j(n)}$ 를 갱신하고 새로운 샘플 집합 $S_t^{P_j} = \{(s_t^{P_j(n)}, \pi_t^{P_j(n)}) \mid n = 1, \dots, N_{P_j}\}$ 을 구성한다.

(2) Estimate : $E[S_t^{P_1}; S_t^{P_2}] =$

$$\sum_{n=1}^{N_1} \sum_{m=1}^{N_2} \cdot \alpha_1 \pi_t^{P_1(n)} s_t^{P_1(n)} + \alpha_2 \pi_t^{P_2(m)} s_t^{P_2(m)} + \frac{1}{2} \alpha_3 \pi_t^{dst(n,m)} (s_t^{P_1(n)} + s_t^{P_2(m)});$$

여기서, $\alpha_1 + \alpha_2 + \alpha_3 = 1$ 이다.

3.2 관측확률의 계산

관측확률에는 각 멀티 파트의 컬러 정보가 참조 모델과 일치하는 정도를 나타내는 가중치 $\pi_i^{p,(n)}$ 와 멀티 파트 간의 기하학적인 거리정보를 나타내는 가중치 $\pi_i^{dist,(n,m)}$ 를 동시에 고려한다.

A. 멀티 파트 컬러 정보 가중치의 계산
 $\pi_i^{p,(n)}$ 는 다음과 같이 계산된다. $\pi_i^{p,(n)} = exp^{-\lambda(1-\rho)}$, 여기서 ρ 는 참조 모델과 후보 모델간의 Bhattacharrya 계수이다.

B. 기하학적인 거리정보를 나타내는 가중치
 본 논문에서는 멀티 파트 간의 기하학적인 거리정보를 나타내는 가중치 $\pi_i^{dist,(n,m)}$ 를 제안하여 사용하였다. 편의상 두 멀티 파트간의 거리 정보 가중치를 계산하는 예를 P_1 과 P_2 가 각각 머리와 얼굴인 경우에 대하여 설명 한다. 머리와 얼굴은 서로 일정한 간격을 가지며 보통 머리는 얼굴 위쪽의 수직선상에 존재한다. 몸을 기울이는 경우 머리와 얼굴을 잇는 직선(그림 1.의 \vec{c})은 수직선(그림 1.의 \vec{r})에 대해 일정한 각도(θ)를 가질 수 있다.

여기서 $cos\theta = \vec{c} \cdot \vec{r} / |\vec{c}| |\vec{r}|$ 이고 얼굴과 머리간의 거리 $|\vec{d}| = |\vec{c}| - l_1 - l_2$ 로 주어지고 l_1 과 l_2 는 각각 머리와 얼굴의 스케일 추정치이다. 기하학정보 θ 와 $|\vec{d}|$ 를 이용하여 다음과 같은 파라미터 ξ 를 정의한다.

$$\xi = \frac{1}{2} \left(\frac{|\vec{d}|}{(l_1+l_2)/2} + 1 - cos\theta \right)$$

그림 1의 오른쪽에 θ 와 $|\vec{d}|$ 의 값에 따른 ξ 값을 그림으로 나타냈다. 마지막으로, 파라미터 ξ 는 다음과 같은 식에 의해 기하학적인 거리정보를 나타내는 가중치로 변환된다.

$$\pi_i^{dist,(n,m)} = cos^2 \left(\frac{\pi \xi + A + 2A_0}{4(A - A_0)} \right)$$

이 변환은 ξ 가 0인 경우 최대값을 가지며, ξ 가 작은 영역에서는 느리게 감소하다가 ξ 가 커지면 빠르게 감소하는 성질을 갖고 있다.

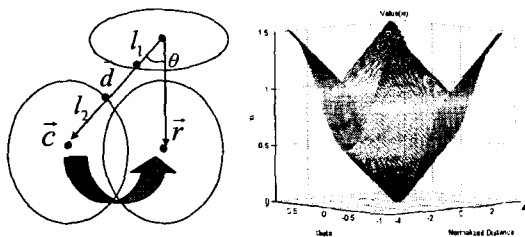


그림 1 거리의 정의와 θ 와 $|\vec{d}|$ 에 따른 ξ 의 값

3.3 기존 문제점의 극복

본 논문에서 제시된 알고리즘은 앞에서 제기한 기존 방법의 문제점을 극복할 수 있다. 먼저 국부 최적해에 고착되는 문제를 살펴보면 제시한 알고리즘은 추적 대상을 멀티 파트로 나누고 각 파트의 컬러 정보의 일치 여부와 멀티파트간의 기하학적인 거리정보를 관측확률의 계산에 반영하였으므로 기존의 알고리즘보다 배경의 클러스터에 대하여 더 강인하고 국부 최적해에 고착되는 경우가 더 적음이 실험결과 확인되었다.

제시한 알고리즘은 얼굴색의 컬러정보와 여러 다른 부분의 컬러 정보를 관측확률에 반영하기 때문에 예를 들어 서로 다른 컬러의 옷을 입은 사람은 몸통의 컬러 정보를 이용하여 구분할 수 있다.

세 번째 문제점인 샘플의 효율성을 살펴보면 제시한 알고리즘에서는 샘플들을 2nd ARP에 따라 전파 시킨 후에 meanshift iteration에 의해서 다시 한번 전파 시키는 방법을 제시하였다. 즉 여러 샘플들을 불필요한 곳으로 보내지 않고 관측확률이 높은 확률이 상대적으로 큰 좌표를 향하여 샘플들을 전파시키는 것이다. 이로써 적은 샘플을 사용하더라도 particle filtering의 확률적 접근을 허용하는 것이다. 이에 따른 기대 효과로서 계산량의 감소로 인한 실시간 추적 시스템의 구현에 한발 더 가까이 다가갈 수 있다.

IV. 실험 결과

실험에서 컬러공간(color space)으로는 HSV (hue-satuarion-value)를 사용하였으며 히스토그램은 각 컬러 요소에 대하여 $10 \times 10 \times 10$ 크기의 bin을 사용하였다. 실험에서 커널 함수로는 Epanechnikov profile[6]을 사용하였다. 그림 2는 meanshift의 국부 최적해로의 고착 문제점을 보이는 예이다. 조명변화에 의한 클러스터, 즉 창문쪽의 Bhattacharrya계수 유사도가 더 높은 경우 "MeanShift Tracker"는 국부 최적해로 발산하며 추적이 실패한다. 반면에 제시한 알고리즘은 머리와 얼굴 두 부분의 컬러와 거리정보를 관측 확률에 사용하기 때문에 얼굴의 유사도가 낮아도 머리의 컬러 유사도가 위쪽으로 발산하는 것을 막아준다. 각 파트의 추적에는 사용된 최대 meanshift iteration 회수는 5회로 제한되었다.

그림 3.에서는 회전할 때 얼굴이 사라지기 때문에 기존의 meanshift tracker는 배경의 Bhattacharrya계수 유사도가 높은 모드를 찾아서 발산하는 반면에 제시한 알고리즘은 머리와 얼굴의 기하학적인 거리 정보가 두 부분의 분리를 막고 컬러 유사도가 높은 곳으로 무조건 발산하는 것을 보상해주기 때문에 얼굴을 정확히 추적하였다.

그림 4는 머리와 얼굴의 컬러 유사도 정보와 기하학적인 위치 정보를 사용하고 부가적으로 가슴 부위의 컬러정보를 관측 확률에 반영하여 두 사람이 교차하여

진행할 때 추적 대상이 일시적으로 가려졌다가 다시 나타나면 정상적으로 원래의 대상을 지속적으로 추적할 수 있음을 보이고 있다. 반면 얼굴의 쉼표정보만을 사용하는 기존의 알고리즘에서는 이 경우 원래의 추적 대상이 아닌 전면의 위치한 사람의 얼굴을 잘못 추적하게 된다.

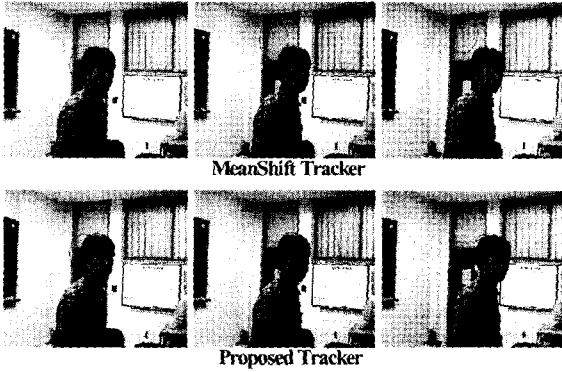


그림 2 국부 최적해를 극복하는 예1.(약 200frame)

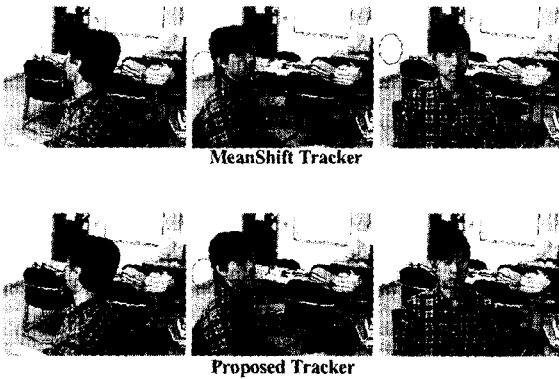


그림 3 국부 최적해를 극복하는 예2. (약 200frame)

VI. 결론 및 향후 과제

본 논문은 복잡한 배경을 갖고 카메라가 움직이는 상황에서의 영상 시퀀스에서 특정 한 사람을 추적하기 위한 particle filtering 기법 기반의 새로운 추적 방법을 제시하였다. 제시한 알고리즘은 머리와 얼굴 등에 대하여 멀티파트 추적을 수행하며 기존의 논문과는 달리 멀티파트의 각 부분이 갖는 Bhattacharrya 계수의 값 이외에 각 멀티파트들 간의 물리적 거리를 함께 고려하는 새로운 관측확률을 사용하여 추적의 정확도를 높였다. 그 결과 독복잡한 배경 클러터(clutter)가 존재하는 상황에서도 추적에 성공하였으며, 추적 대상이 일시적으로 가려져도 지속적으로 추적에 성공함을 실험을 통하여 검증하였다. 본 논문에서 제시한 추적 방법은 얼굴인식(face

recognition)의 전 처리과정, 또는 비디오 감시체제에 사용 가능하며 음성과 결합하여 HCI에 응용될 수 있다.



그림 4 일시적으로 가려지는 대상의 추적 예 (약 200frame)

References

- [1] 류성희, "무인 영상 감시를 위한 이동 물체 검출 및 추적 시스템의 실시간 구현에 관한 연구", 수원대학교 석사학위논문, 2001
- [2] 이진성, 김성대, "움직임 벡터의 신뢰도에 기반한 이동 목표물 추적 기법" 전자공학회 논문지, 2001
- [3] A. M. Baumberg, Learning deformable models for tracking human motion. PhD thesis, University of Leeds, UK, 1995
- [4] M.Isard, A.Blake, "Condensation-Conditional Density Propagation for Visual Tracking," Intern. J. Comp.Vis., 1998.
- [5] S. Bircheld, "Elliptical Head Tracking using intensity Gradients and Color Histograms," IEEE Conf. on Comp. Vis. and Pat. Rec., Santa Barbara, 1988
- [6] D. Comaniciu, P.Meer, "Mean Shift Analysis and Applications," IEEE Int'l Conf. Comp.Vis., Kerkyra, Greece,1197-1203, 1999
- [7] D. Comaniciu, V. Ramesh, P. Meer, "Real-Time Tracking of Non-Rigid Objects using Mean Shift", IEEE Conf. on Comp. Vis. and Pat. Rec. 2000
- [8] D. Comaniciu, V. Ramesh, "Robust Detection and Tracking of Human Faces with an Active Camera" IEEE Visual Surveillance 2000.
- [9] P. Perez, C. Hue, J. Vermaak, M. Gangnet, "Colr-Based Probabilistic Tracking, European Conference on Computer vision (2002) 661-675.