

영상 기반 실시간 객체 추적에서 객체 크기 추정 기법에 관한 연구

김은솔* 최유주**

*가천대학교 소프트웨어 설계·경영학과

**한독미디어대학원대학교 뉴미디어콘텐츠학과

e-mail : kesol0328@gmail.com yjchoi@kgit.ac.kr

A Study of Scaling Methods in Vision-based Real-time Object Tracking

Eun-Sol Kim* Yoo-Joo Choi**

*Dept. of Software Design and Management, Ga-Cheon University

**Dept. of Newmedia Content, Korean German Institute of Technology

요 약

본 논문에서는 커널 기반 객체 추적 방식을 이용하여 실시간으로 객체를 추적하는 기술에서 객체의 크기 추정에 대한 기존 연구를 비교 분석한다. 커널 기반 객체 추적은 추적하고자 하는 객체를 초기 프레임에서 타겟으로 설정한 후, 각 프레임마다 타겟 후보들을 제시하고 그 중 가장 객체와 유사한 후보에 초점을 맞춰가며 객체를 추적한다. 이때, 목표 객체와 객체 후보간 유사성을 기반으로 정의된 배경 영상 영상(back-projection image)을 이용하여 객체의 크기를 추정하는 방법들이 제시되고 있다. 안정적인 객체 크기 추정 방법의 설계를 위한 사전 연구로서 대표적인 객체 크기 추정 기존 연구를 비교, 분석하고자 한다.

1. 서론

최근 동적 카메라가 여러 방면에서 다양하게 사용되고 있어 객체 추적에 대한 연구가 더 중요시 되고 있다. 특히 PTZ 카메라를 이용한 지능형 영상 처리 기술의 수요가 확대되고 있는 추세이다. 그 중에서도 지능형 감시 시스템이 크게 주목을 받고 있다. 이미 여러 지자체와 산업체들은 지능형 영상 감시 보안 시스템을 개발, 사용하는 추세이다. 또 다른 예로 방송사에서는 상당수의 자동 촬영 장비들이 보급된 상태로, 여러 방면에서 사용되고 있다.

하지만 PTZ 카메라의 보급이 확대됨에 따라 비고정형 카메라를 위한 객체 추적 알고리즘에 관한 관심이 높아지고 있다. 일반적으로 고정된 카메라에서 객체를 추적하는 경우, 배경 차감 방법을 많이 적용하고 있으나, PTZ 카메라 이용 시에는 배경이 지속적으로 변화하기 때문에 이러한 배경 차감 방법을 사용할 수 없다. 그래서 관심을 모으는 추적 기법으로 커널 기반 객체 추적 방법을 들 수 있다.

커널 기반 객체 추적은 계산량이 적고, 객체의 중첩(occlusion) 등에 강건하게 동작하는 장점이 있으나 추적의 정확도를 높이기 위하여 해결하여야 할 몇가지 이슈가 있다. 첫째로, 조명 조건이 변하는 경우이다. 실외의 경우 시간 흐름에 따라, 대기 상황에 따라 조명 조건이 바뀐다. 그래서 초기에 설정한 타겟 모델의 정보가 점진적으로 갱신되어야 한다. 둘째로, 목

표 객체의 크기가 일정하지 않고 변화한다. 동적 카메라가 움직이는 경우와 객체 자체가 움직이는 경우 모두 초기에 설정한 커널의 크기와 달라진다. 특히 이러한 크기 변화가 점진적인 경우도 있지만 급격한 경우, 추적하던 객체를 놓칠 가능성이 높다. 마지막으로, 객체 추적에서는 객체의 색상 정보를 활용하기 때문에 배경 색상과 객체 색상이 유사한 경우 추적에 실패하는 경우가 있다.

본 논문에서는 커널 기반 객체 추적에서 추적 정확도를 높이기 위한 주요 이슈들 중에서 객체의 크기 추정에 초점을 맞추고자 한다. 이에 본 논문에서는 영상 기반 실시간 객체 추적에서 객체 크기 추정을 위한 강건한 알고리즘을 설계하기 위하여 이를 위한 사전 연구로서 실시간 동적 카메라 영상을 기반으로 이동하는 객체의 크기를 추정하는 기존 기법들을 분석하여 이들의 특성을 비교 분석하고자 한다.

2. 커널 기반 객체 추적

Comaniciu 방법[1]은 커널 기반 객체 추적의 기본적인 방법이다. m -bin 색상 히스토그램을 사용하여 표현한다. 타겟 모델을 식(1)로, 타겟 후보를 식(2)로 정의한다.

$$\hat{q} = \{\hat{q}_u\}_{u=1\dots m} \quad \sum_{u=1}^m \hat{q}_u = 1 \quad (1)$$

$$\hat{p}(y) = \{\hat{p}_u(y)\}_{u=1\dots m} \quad \sum_{u=1}^m \hat{p}_u = 1 \quad (2)$$

식(1)의 q 는 커널 내부의 색상 정보에 위치 가중치를 적용한 히스토그램이다. 타겟 후보 커널 내의 각 픽셀에 대한 배경 영사 가중치 w_i 는 식(3)과 같이 정의된다. 식(3)에서의 δ 는 Kronecker 델타 함수이고, x_i 는 커널 내의 i 번째 픽셀의 위치를 의미한다

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \delta[b(x_i) - u] \quad (3)$$

y 를 중심으로 하는 현재 프레임에서의 타겟 후보 커널 내의 픽셀의 개수를 n_h 라 하는 경우, 배경 영사 가중치를 기반으로 새롭게 이동되는 타겟 후보 커널의 중심 y_i 는 식(4)와 같이 정의된다.

$$\hat{y}_i = \frac{\sum_{i=1}^{n_h} x_i w_i}{\sum_{i=1}^{n_h} w_i} \quad (4)$$

3. 커널 기반 객체 추정 기법 분석

3.1 Comanicu의 방법

Comanicu는 객체 크기 추정을 위해서 이전 프레임에서의 크기를 h_{prev} 라고 정의하고, 가장 적절한 크기인 h_{opt} 를 찾기 위해 3 가지 크기를 비교하였다. 즉, 커널의 대역폭을 나타내는 h 를 각기 $h = h_{prev}$, $h = h_{prev} + \Delta h$, $h = h_{prev} - \Delta h$ 로 정의하였다. 여기서 $\Delta h = 0.1h_{prev}$ 을 의미한다. 세 개의 서로 다른 크기의 타겟 후보 커널에 대해 타겟 모델과 Bhattacharyya coefficient를 이용해 유사성을 판별하여 가장 유사성이 높은 대역폭 h 를 h_{opt} 로 정의하였다. γ 를 0.1로 정의하여 식(5)로 새로운 커널의 대역폭 h_{new} 를 정의하였다.

$$h_{new} = \gamma h_{opt} + (1 - \gamma)h_{prev} \quad (5)$$

Comanicu는 새로 찾은 객체 후보의 중심에서 이전 프레임에서의 객체 영역 크기의 $\pm 10\%$ 인 커널을 정의하고, 이중 가장 유사도가 높은 커널 크기로 객체의 크기를 추정하였다. 기존 커널 크기에서 $\pm 10\%$ 만 반영하기 때문에 급격한 크기의 변화가 일어났을 경우 추적이 어려워진다.

3.2 Ning의 방법

Ning 방법[2]은 가중치 영상(weight image)의 모멘트 특성과 밀도 분포 함수를 이용하여 타겟의 크기와 위치를 측정한다. 가중치 영상에서 모멘트를 구하는 방법은 식(6), (7), (8), (9)와 같다.

$$M_{00} = \sum_{i=1}^n w(x_i) \quad (6)$$

$$M_{10} = \sum_i w_i x_{i,1}, \quad M_{01} = \sum_{i=1}^{n_h} w_i x_{i,2} \quad (7)$$

$$M_{10} = \sum_{i=1}^{n_h} w_i x_{i,1}, \quad M_{01} = \sum_{i=1}^{n_h} w_i x_{i,2} \quad (8)$$

$$M_{20} = \sum_{i=1}^{n_h} w_i x_{i,1}^2, \quad M_{02} = \sum_{i=1}^{n_h} w_i x_{i,2}^2, \quad (9)$$

$$M_{11} = \sum_i w_i x_{i,1} x_{i,2}$$

식(9)에서의 $(x_{i,1}, x_{i,2})$ 는 후보 영역 안의 픽셀 i 의 좌표이다. 무게중심 y 는 0 차 모멘트에 대한 1 차 모멘트의 비율이다.

$$y_1 = (\bar{x}_1, \bar{x}_2) = (M_{10}/M_{00}, M_{01}/M_{00}) \quad (10)$$

식(10)의 (\bar{x}_1, \bar{x}_2) 는 타겟 후보 영역의 중심이다. 이와 같은 방식으로 구한 2 차 중심 모멘트는 객체의 크기와 방향을 묘사한다. 2 차 중심 모멘트는 새로운 y 좌표로 이동하는 식을 변환할 수 있게 한다.

$$\mu_{20} = M_{20}/M_{00} - \bar{x}_1^2 \quad (11)$$

$$\mu_{11} = M_{11}/M_{00} - \bar{x}_1 \bar{x}_2 \quad (12)$$

$$\mu_{02} = M_{02}/M_{00} - \bar{x}_2^2 \quad (13)$$

식(11), (12), (13)을 이용하여 타겟의 너비, 높이, 방향을 측정할 수 있는 행렬을 쓸 수 있다.

$$Cov = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} \quad (14)$$

이 covariance 행렬을 SVD 하면

$$Cov = U \times S \times U^T = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \times \begin{bmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{bmatrix} \times \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix}^T \quad (15)$$

인 식이 나온다. 타겟 모델을 타원인 A 라고 설정했을 때 $A = \pi ab = \pi(k\lambda_1)(k\lambda_2)$ 로 나타낼 수 있다. 그러므로 높이 a 와 너비 b 가 결정되며 다음 프레임의 타겟 후보 영역의 크기가 표현된다. Δd 는 다음 프레임에서 타겟 후보 영역의 증가 폭이다.

$$a = \sqrt{\lambda_1 A / (\pi \lambda_2)}, \quad b = \sqrt{\lambda_2 A / (\pi \lambda_1)} \quad (16)$$

$$Cov_2 = U \times \begin{bmatrix} (a + \Delta d)^2 & 0 \\ 0 & (b + \Delta d)^2 \end{bmatrix} \times U^T \quad (17)$$

타겟 후보 영역의 초기 위치는 다음의 타원 영역으로 정의된다.

$$(x - y_1) \times Cov_2^{-1} \times (x - y_1)^T \leq 1 \quad (18)$$

Ning의 방법은 타겟 모델과 타겟 후보 모두가 배경 가중치를 고려하지 않은 경우를 전제로 하여 설계되었다. 그러므로 객체 추적 정확도를 높이기 위하여 배경 가중치를 고려하는 경우에는 적용할 수 없는 방법이다.

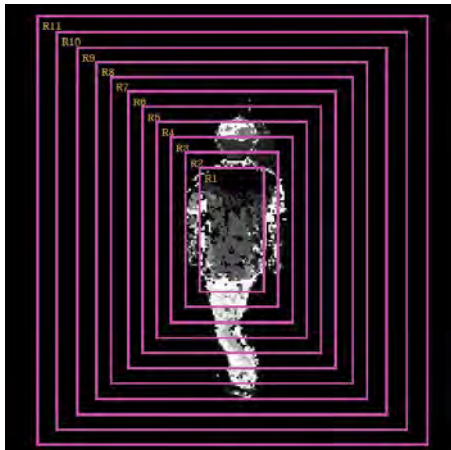
3.3 Hwang 의 방법

Hwang 의 방법[3]은 타겟 모델에만 객체의 배경 조건에 대한 가중치를 적용하여 객체 고유 색상이 강조된 가중치 영상을 이용한다. 이로 인해 객체 추적의 정확도를 높였다. 그러나 타겟 모델과 타겟 후보의 색상 히스토그램의 특성이 달라짐에 따라 타겟 모델과 타겟 후보 간의 유사성을 판별하기 위한 별도의 판별 기준이 요구되었다. 그래서 Hwang 은 초기 설정된 타겟의 커널 내의 가중치 총합을 M_{00} , 커널의 크기를 R 로 정의하고, 식(20)의 α 를 초기 타겟 모델과 타겟 후보 간의 유사도 판별 기준으로 사용하였다.

$$M_{00} = \sum_{i=1}^n w_i x_i \quad w_i x_i: \text{weight pixel value} \quad (19)$$

$$\alpha = \frac{R}{M_{00}} \quad (20)$$

즉, 타겟 모델을 정의하는 초기 프레임에서 커널의 크기와 커널 내 픽셀의 가중치 총합의 비 α 를 계산하고, 이를 각각의 커널 후보에서의 α 와 비교하여 α 가 가장 유사한 타겟 후보의 커널의 크기를 선택하였다.



(그림 1) 객체 크기 추정의 후보 영역

Hwang 은 (그림 1)에서 보이듯 현재 프레임 타겟 커널을 기준으로 너비와 높이가 3%씩 5 단계 줄이고, 5 단계 늘린 후보 커널 11 개를 정의하였다. $R_l (l = 1 \dots 11)$ 은 각 크기별 영역을 의미하며 R_l 이 포함되는 R_{l-1} 을 뺀 R_l 만의 고유 영역을 $r_l (l = 1 \dots 11)$ 로 정의한다. 그렇다면 각 크기별 가중치 총합 M_{00}^l 은 식(21)와 같이 된다. 각 크기별 객체 고유 색상 가중치 밀도 α_l 은 식(22)와 같이 된다.

$$M_{00}^l = \sum_{k=1}^l \sum_{x_i \in r_k} w_i(x_i) \quad (21)$$

$$\alpha_l = \frac{R_l}{M_{00}^l} \quad (22)$$

초기 구해진 객체 고유색상 가중치 밀도 α 와 각 크기별 가중치 밀도 α_l 의 차 식(23)이 최소가 되게 하는 크기 레벨의 값 m 을 통해서 새로운 커널의 크기를 결정한다.

$$d(l) = |\alpha - \alpha_l| \quad (l = 1 \dots 11) \quad (23)$$

$$m = \text{argmin}(d(l)) \quad (24)$$

$$A = R_m \quad (25)$$

이 방법을 사용할 경우 추적 시 크기 정확도는 높아진다. 하지만 추적하고자 하는 객체의 크기가 작은 경우에는 정의, 수행이 어렵고, 반대로 크기가 큰 경우는 계산 비용을 높이는 오버헤드가 크게 발생한다.

3.4 Jeyakar 의 방법

Jeyakar 의 방법[4]은 가중치 영상을 이용한다. 축의 중심 α 를 구하고, 각 픽셀간의 거리와 weight 에 비례하는 weight 표준편차 β 를 정의하여 크기 추정하는데 이용한다.

$$\alpha = \frac{1}{\sum_{i=0}^n w_i} \sum_{i=1}^n x_i w_i \quad (26)$$

$$\beta = \sqrt{\frac{1}{\sum_{i=0}^n w_i} \sum_{i=1}^n x_i (w_i - \alpha)^2} \quad (27)$$

초기 커널의 크기 h_{org} 와 초기 커널에서 정해진 weight 표준편차 β_{org} 를 이용해 새로운 프레임에서의 weight 표준편차 β 와 새로운 커널의 크기 h_{new} 의 비가 초기의 것들과 같다고 정의하고 식(28)를 이용해 h_{new} 를 구한다.

$$h_{new} = h_{org} \times \frac{\beta}{\beta_{org}} \quad (28)$$

하지만 이 방법은 커널의 외곽 배경 부분에 타겟 객체와 유사 색상이 존재하는 경우, 외곽 픽셀의 가중치가 높아져서 가중치 전체 평균 편차에 크게 영향을 주게 되어, 커널 크기의 확대 추정 위험이 높아지게 된다.

3.5 Dulai 의 방법

Dulai 의 방법[5]은 이전의 가우시안의 방법에서 시작되었다. 가우시안의 두 개의 타원으로 커널을 생성하여 객체를 추적했던 방법에서 계산을 단순화하였다.

$$H = \begin{bmatrix} 1 & 0 \\ h_1 & \\ 0 & 1 \\ & h_2 \end{bmatrix} \quad (29)$$

식(29)에서 H는 대역폭을 표현하는 행렬이며 h_1 과 h_2 는 커널의 대역폭을 의미하는 변수이다. 타겟의 비율이 변하지 않는다고 가정할 때, j 번째 프레임의 x와 y 방향의 대역폭을 $h_{1,j}$ 와 $h_{2,j}$ 라고 표현하면, j+1 번째 프레임에서의 대역폭인 $h_{1,j+1}$ 과 $h_{2,j+1}$ 는 식(30)과

식(31)로 표현된다.

$$h_{1,j+1} = h_{1,j}(1 + \alpha) \quad (30)$$

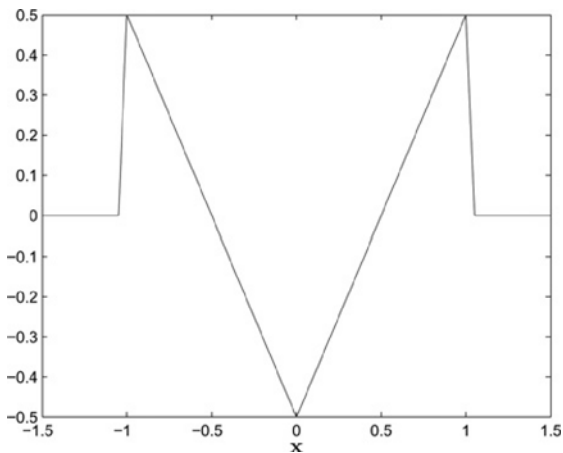
$$h_{2,j+1} = h_{2,j}(1 + \alpha) \quad (31)$$

α 는 갱신된 인자이다. α 가 음수라면 타겟의 크기가 줄어들 것이고, 양수라면 늘어날 것이다. α 를 구하는 방법은 다음과 같다.

$$\alpha = \frac{\sum_{x \in X} w(x) k_r\{\|H(y_i - x)\|\}^2}{\sum_{x \in X} w(x)} \quad (32)$$

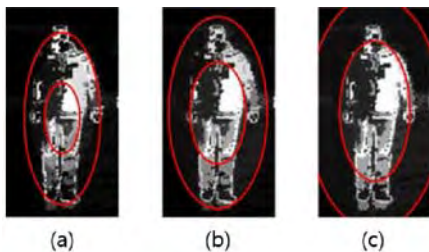
여기서 $k_r()$ 함수는 식(33)을 의미한다. (그림 2)는 커널 $k_r(x)$ 에 대한 중앙 부분 한 단면의 값을 표현하고 있다.

$$k_r\{x\} = \begin{cases} \sqrt{x} - 0.5, & x \leq 1 \\ 0, & otherwise \end{cases} \quad (33)$$



(그림 2) 커널 $k_r\{\|x\|\}^2$ 의 단면

(그림 2)에서 보듯 커널의 범위는 음수의 부분과 주변의 양수 부분까지 포함하여 -0.5~0.5 이다. 일반적인 가우시안 커널과는 다른 양수 값을 가지는 커널이 음수 값을 가지는 커널을 둘러싸고 있는 모습을 가진다. α 값은 target 의 크기에 따라 다양하게 발생할 수 있다.



(그림 3) Dulai 방법을 이용한 크기 추정

(그림 3)의 (a)는 대부분 0 이 아닌 w_i 를 가지는 픽셀이 $k_r()$ 값이 양수인 영역에 해당되기 때문에 식(32)의 α 값이 0 보다 커지게 되는 경우를 보여주고 있다.

(b)는 0 이 아닌 w_i 를 가지고 픽셀이 $k_r()$ 의 양수 영역과 음수 영역에 비슷하게 분포하기 때문에 $\alpha \approx 0$ 이 되는 경우를 보여주고 있다. (c)는 0 이 아닌 w_i 를 가지는 픽셀이 $k_r()$ 값이 음수인 영역에 주로 포함되기 때문에 식(32)의 α 는 음수가 되는 경우를 보여주고 있다.

이 방법은 mean shift 가 실행될 때마다 두 커널을 계산한다. 일반적으로 두 개의 커널에 대한 계산이 수행되어야 하는 오버헤드가 발생한다.

4. 결론

본 논문에서는 커널 기반 객체 추적에서 추적 정확도를 높이기 위한 연구 중 객체 크기 추정에 대한 연구들을 비교 분석해보았다. 비교 결과, 객체 추적의 정확도를 높이기 위해서는 배경 가중치를 적용하는 것이 필요하다[6]. 그런데 널리 알려진 방법들은 배경 가중치를 고려하지 않은 방법에서 객체 추정하는 방법을 제시하고 있다. 배경 가중치를 적용한 환경에서, 배경과 타겟 객체간 색상 유사성에 의해 발생하는 커널 외곽 픽셀의 높은 가중치의 영향을 줄일 수 있는 객체 추정 방법의 고안이 요구되고 있다.

이에 향후 연구로서, 본 분석 내용을 기반으로 커널 기반 객체 추정에 있어서 안정적으로 객체 크기를 추정할 수 있게 방법을 설계하고자 한다.

감사의 글

본 연구는 한국여성과학기술단체총연합회 “2015 년도 차세대 우수 IT 여성과학기술인 육성을 위한 고급 훈련 프로그램 운영 사업” 지원에 의하여 수행되었음.

참고문헌

- [1] D. Comaniciu, V. Ramesh, P. Meer, “Kernel-based object tracking”, IEEE Trans. of Pattern Analysis and Machine Intelligence, Vol 25, No. 5, pp. 563-577, 2003
- [2] J. Ning, L. Zhang, D. Zhang, C. Wu, “Scale and orientation adaptive mean shift tracking”, IET Comput. Vis., Vol 6, No. 1, pp. 52-61, 2012
- [3] 황만원, “배경조건을 고려한 Mean-shift 추적에서의 객체 크기 추정”, 한국지식정보기술학회 논문지, Vol. 9, No. 4, pp.519-531, Aug. 1. 2014
- [4] J. Jeyaker, R. V. Babu, K. R. Ramakrishna, “Robust object tracking with background-weighted local kernels”, Computer Vision and Image Understanding, Vol. 112, No. 3, pp. 296-309, 2008
- [5] A. Dulai, T. Sathako, “Mean shift tracking through scale and occlusion”, IET Signal Process., Vol. 6, No. 5, pp. 534-540, 2012
- [6] J. Ning, L. Zhang, D. Zhang, C. Wu, “Robust mean-shift tracking with corrected background-weighted histogram”, IET Comput. Vis., Vol. 6, No. 1, pp. 62-69, 2012