

## 척도변화와 부분적가리움에 안정한 대상추적의 한가지 방법

리금성, 하정원

선행연구[1, 2]에서는 Mean-Shift(MS)를 리용한 대상추적방법을 제기하였으나 그 방법은 척도변화와 부분적가리움에 잘 적응하지 못하는 결함이 있다.

한편 선행연구[3]에서는 척도변화와 부분적가리움문제를 논의하였으나 성능이 낮으며 일정한 조건에 제한되어있다.

본문에서는 매 화소를 전경과 배경으로 분류하고 대부분의 전경화소를 포함하는 구역을 탐색하여 이 문제를 해결하였다.

### 1. 화소분류와 적분을 리용한 대상추적방법

화소분류는 현재 프레임의 매 화소에 점수를 부여하는 과정으로서 이러한 화소분류 결과들을 적분하면 대상추적을 원만히 진행할수 있다.

화소분류와 적분을 리용한 대상추적체계를 PCI라고 한다.

PCI에서는 현재 프레임에서의 새로운 대상영역을 다음과 같이 결정한다.

$$R_f = \arg \max_R J(R) = \arg \max_R \sum_{x \in R} s(x) \quad (1)$$

여기서  $J(R)$ 는 최량화해야 할 목적함수로서  $\sum_{x \in R} s(x)$ 로 표현되는 영역  $R$ 안의 화소점수들의 적분이다. 여기로부터 PCI에서 기본은 화소점수  $s(x)$ 를 계산하는것이다.

화소점수의 분류과정을 보기로 하자.

본문에서는 화소분류를 진행하기 위한 특징  $c(x)$ 로서 화소색 YUV를 리용하고 전경 대상과 배경대상의 색히스토그램을 리용하여 분류를 진행한다.

이때 간단화를 위해 화소색  $c(x)$ 대신에 그에 대응한 히스토그램번호  $u=1, 2, \dots, m$ 에 대한  $u(x)$ 를 리용한다.

우선 전경대상이 첫 프레임에서 주어지면 그것의 히스토그램  $p_f(u)$ 를 계산한다.

다음 배경히스토그램은 항상 추적과정에 이전 프레임으로부터 만들어진다. 즉 이전 프레임의 매 방향으로  $\alpha$ 만큼 대상영역을 확장하고 그것을 둘러싸고있는 배경영역  $R_b^{t-1}$ 의 색히스토그램  $p_b^{t-1}(u)$ 를 계산한다.

여기로부터 프레임  $t$ 의 화소  $x$ 에서 화소점수의 분류는 다음과 같이 진행한다.

$$s^t(x) = \log \frac{p_f(u(x))}{p_b^{t-1}(u(x))}, \quad x \in R_s^t \quad (2)$$

여기서  $R_s^t = R_f^{t-1} \cup R_b^{t-1}$  이다.

식 (1)에 의하여 다음 프레임에서의 추적된 대상영역은 다음과 같이 계산한다.

$$R_f^t = \arg \max_R J^t(R) = \arg \max_{R \subset R_s^t} q^{t-1}(R) \sum_{x \in R} s^t(x) \quad (3)$$

여기서  $q^{t-1}(R)$  는 이전 대상위치와 현재 대상위치사이의 시간관계를 Epanechnikov핵심부를 가지고 모형화한것이다. 즉

$$q^{t-1}(R) = \max(0, 1 - d^2(R, R_f^{t-1})) \quad (4)$$

로서  $d(R, R_f^{t-1})$  은  $R$  와  $R_f^{t-1}$  의 중심들사이의 표준화된 유클리드거리이다.

우의 방법은 목적대상의 척도변화를 고려하지 않은것이다.

따라서 우리는 목적대상의 척도변화에 적응하도록 전경영역을 다음과 같이 새롭게 결정하였다.

$$R_f^t = \arg \max_R J_{scale}^t(R) = \arg \max_{R \subset R_s^t} q^{t-1}(R) \frac{1}{w \cdot h} \left[ (1 - \gamma_1 + \gamma_2) \sum_{x \in R} s^t(x) + \gamma_1 \sum_{x \in R_{\beta_1}} s^t(x) - \gamma_2 \sum_{x \in R_{\beta_2}} s^t(x) \right] \quad (5)$$

여기서  $w \cdot h$  는 영역  $R$  의 크기,  $R_{\beta_1}$  은 영역  $R$  가 약간 축소된것,  $R_{\beta_2}$  는 영역  $R$  가 약간 확장된것이다. 그리고  $\gamma_1$  과  $\gamma_2$  는 전경대상과 배경대상의 중요성을 조절하는 인자로서 첫 프레임에서 다음과 같이 결정한다.

$$\gamma_1 = \frac{\sum_{x \in R_{\beta_1}} s^0(x) - \sum_{x \in R} s^0(x)}{\sum_{x \in R} s^0(x)}, \quad \gamma_2 = \frac{\sum_{x \in R} s^0(x) - \sum_{x \in R_{\beta_2}} s^0(x)}{\sum_{x \in R} s^0(x)} \quad (6)$$

대상추적기의 또 다른 중요한 측면의 하나는 부분적가리움으로 인한 출현변화에 적응하는 능력이다.

우리는 이를 위해 전경과 배경히스토그램을 다음과 같이 결정하였다.

$$p_f(u) = m_{fl} p_{fl}(u) + m_{fs} p_{fs}(u) \quad (7)$$

$$p_b(u) = m_{bl} p_{bl}(u) + m_{bs} p_{bs}(u) \quad (8)$$

여기서  $p_f(u)$  와  $p_b(u)$  는 전경과 배경히스토그램으로서  $p_{fl}(u)$  는 전경히스토그램의 장기간구성요소를,  $p_{fs}(u)$  는 단기간구성요소를 나타낸다. 그리고  $m_{fl}$  과  $m_{fs}$  는 혼합무계로서

$$m_{fl} + m_{fs} = 1$$

이다. 유사하게  $p_{bl}(u)$  와  $p_{bs}(u)$  는 배경히스토그램의 장기간과 단기간구성요소로서 그 혼합무계도  $m_{bl} + m_{bs} = 1$  이며 이것을 결정하기 위하여 Kullback Leibler거리를 리용하는데 그 과정은 다음과 같다.

우선 전경대상에 대하여 프레임  $t$  에서의  $R_f^t$  로부터 색히스토그램  $\tilde{p}_f^t(u)$  를 계산한다.

$$d_l = \sum_u \tilde{p}_f^t(u) \log \frac{\tilde{p}_f^t(u)}{p_{fl}^{t-1}(u)}, \quad d_s = \sum_u \tilde{p}_f^t(u) \log \frac{\tilde{p}_f^t(u)}{p_{fs}^{t-1}(u)} \quad (9)$$

여기서  $d_l$  과  $d_s$  는  $\tilde{p}'_f(u)$  와 두 구성요소히스토그램사이의 Kullback Leibler거리들이다.

여기로부터 새로운 무게는 다음과 같다.

$$\tilde{m}'_{fl} = \frac{e^{-d_l/\sigma}}{e^{-d_l/\sigma} + e^{-d_s}}, \quad \tilde{m}'_{fs} = \frac{e^{-d_s}}{e^{-d_l/\sigma} + e^{-d_s}} \quad (10)$$

$$m^t_{fl} = \eta \tilde{m}'_{fl} + (1-\eta)m^{t-1}_{fl}, \quad m^t_{fs} = \eta \tilde{m}'_{fs} + (1-\eta)m^{t-1}_{fs} \quad (11)$$

여기서

$$\eta = \frac{\tilde{m}'_{fs}}{\tilde{m}'_{fl} + \tilde{m}'_{fs}} \quad (12)$$

이며  $\tilde{m}'_{fl}$  과  $\tilde{m}'_{fs}$  는 매개 구성요소의 상대적인 우량도이다. 그리고  $\sigma$  는 단기간구성요소에 대한 장기간구성요소의 사전믿음도를 반영한다.

다음 전경구성요소를 다음과 같이 갱신한다.(이 경우  $\hat{p}'_f(u)$  는  $\tilde{p}'_f(u)$  와 같다고 본다.)

$$p^t_{fl}(u) = \rho_1 \hat{p}'_f(u) + (1-\rho_1)p^{t-1}_{fl}(u) \quad (13)$$

$$p^t_{fs}(u) = \rho_2 \hat{p}'_f(u) + (1-\rho_2)p^{t-1}_{fs}(u) \quad (14)$$

여기서  $\rho_1$  과  $\rho_2$  는 갱신속도를 조종하는 파라메터이다.

류사하게 영역  $R'_b$  로부터 배경히스토그램  $\tilde{p}'_b(u)$  를 계산한다.

이때 배경구성요소는 다음과 같이 갱신된다.

$$p^t_{bl}(u) = \rho_1 \hat{p}'_b(u) + (1-\rho_1)p^{t-1}_{bl}(u) \quad (15)$$

$$p^t_{bs}(u) = \rho_2 \hat{p}'_b(u) + (1-\rho_2)p^{t-1}_{bs}(u) \quad (16)$$

여기로부터 대상추적알고리즘은 다음과 같다.

① 초기화상에서 전경영역  $R^0_f$  과 배경영역  $R^0_b$  을 설정한다.

②  $R^0_f$  과  $R^0_b$  의 색히스토그램으로  $p^0_{fl}(u)$ ,  $p^0_{fs}(u)$ ,  $p^0_{bl}(u)$ ,  $p^0_{bs}(u)$  를 초기화하고  $t=0$  으로 한다.

③  $t=t+1$  로 한다.

④ 후보영역  $R$  를 선택한다.

⑤  $R^t_f$  와  $R^t_b$  에서  $\hat{p}'_f(u)$  와  $\hat{p}'_b(u)$  를 계산한다. 그리고 식 (13)–(16)에 의해  $p^t_{fl}(u)$ ,  $p^t_{fs}(u)$ ,  $p^t_{bs}(u)$  를, 식 (7)–(12)를 리용하여  $p^t_f(u)$ ,  $p^t_b(u)$  를 계산한다.

⑥  $J(R)$  를 계산한다.

⑦ 시각  $t$  에서 식 (1)을 최소로 하는 영역  $R$  를 구한다.

⑧ 마지막 화상렬에 도달할 때까지 ③–⑥을 반복한다.

## 2. 실험결과 및 분석

실험은 좌우로 이동하면서 촬영기에 접근하는 보행자에 대하여 첫 화상에서 대상을 선택하는 방법으로 진행하였다. 보행자가 촬영기에 접근하므로 대상의 척도는 변하며 좌우로 이동하면서 부분적거리움이 나타난다.

정확성검증을 위하여 매 화상에서 주목하는 대상의 위치를 미리 선택하고 MS방법과 제안된 방법에 의한 추적결과를 비교하였다.

표에 프레임번호에 따르는 대상의 위치를 MS방법과 비교하고 그 오차를 정확한 위치와의 유클리드거리로 보여주었다.

표. 추적된 대상위치와 오차

프레임번호	정확한 위치	MS방법		제안된 방법	
		대상위치	오차	대상위치	오차
3	(302, 269)	(304, 270)	2.23	(304, 270)	2.23
7	(290, 281)	(287, 279)	3.60	(288, 280)	2.23
15	(285, 296)	(280, 293)	5.83	(283, 294)	2.82
27	(270, 310)	(276, 301)	10.81	(273, 312)	3.60
39	(291, 328)	(280, 319)	14.21	(286, 321)	8.60
57	(306, 351)	(295, 340)	17.02	(299, 348)	7.61
65	(320, 369)	(311, 361)	12.04	(316, 373)	5.65
86	(352, 390)	(341, 381)	14.21	(357, 393)	5.83

표로부터 척도변화와 부분적가리움이 있는 경우 제안된 방법이 정확하다는것을 알수 있다.

## 맺 는 말

영역의 매 화소를 그 위치와 현재 및 이전 프레임에서의 출현모형을 리용하여 전경과 배경에 속할 점수로 분류하고 그것을 적분하는 방법으로 대상추적을 진행하는 한가지 방법을 제기하였다.

## 참 고 문 헌

- [1] S. Konstantopoulos et al.; SETN, 153, 2010.
- [2] J. Jeyakar et al.; Computer Vision and Image Understanding, 112, 296, 2008.
- [3] C. Zhang et al.; The 18<sup>th</sup> International Conference on Pattern Recognition, 2006.

주제 104(2015)년 8월 5일 원고접수

## A Method for Object Tracking Robust to Scale Variation and Partial Occlusion

*Ri Kum Song, Ha Jong Won*

We proposed a method for the object tracking that is robust to different object scale and partial occlusion. We implemented the object tracking by classifying and integrating each pixel in the region into the score that belongs to the foreground and the background using the position and the appearance models of the current frame and the pre-frame and demonstrated the efficiency by tracking experiments for real video.

Key words: object tracking, mean-shift, visual tracking