A Hybrid Tracking Method Based on Active Contour and Mean Shift Algorithm

Zhu Linlin  
State Key Laboratory of Robotics  
SIA & Graduate School of the Chinese Academy of Sciences  
Shenyang, China  
zhulinlin@sia.cn

Fan Baojie  
State Key Laboratory of Robotics  
SIA & Graduate School of the Chinese Academy of Sciences  
Shenyang, China  
fanbaojie@sia.cn

Li Benjin  
Beijing Institute of Tracking and Telecommunications Technology  
Beijing, China

Tang Yandong  
State Key Laboratory of Robotics  
SIA  
Shenyang, China  
ytang@sia.cn

Abstract—Active Contour is a very accurate in target tracking and robust to variety of illumination, translation, rotation and large scale. Unfortunately, the problem that Active Contour method can not track target in real time because of huge computation is still not well resolved. On the contrast, Mean Shift is a very fast algorithm in target tracking but sensitive to the change of illumination, and can’t get the description of target’s shape. In our work, we use Mean Shift method to determine the translation motion of target and build a initial rough contour for Active Contour method. So the computation of Active Contour method in early curves evolution stage is greatly reduced and the target can be tracked accurately by minimizing the energy function in few number of interactive computation. Experimental results validate our method.

Keywords-tracking; active contour; level set; mean shift;

I. INTRODUCTION

Object tracking is an important task within the field of computer vision. In its simplest form, tracking can be defined as the problem of estimating the trajectory of an object in the image plane as it moves around a scene. Comaniciu and Meer [1] use the Mean Shift procedure for the target location in the course of tracking. The Mean Shift tracker maximizes the appearance similarity iteratively by comparing the histograms of the object and the window around the hypothesized object location. This algorithm can get the target’s translation fast. In recent years, many work based on Mean Shift has been done [2–4], make the Mean Shift tracker can handle the target’s translation, scale, rotation and occlusion, but Mean Shift can’t represent the target’s silhouette exactly which active contour can do well. Active contour model has received widely attention and research after first proposed by Kass in 1987 [5]. In tracking, active contour directly evolves the contour by minimizing the contour energy using direct minimization techniques such as gradient descent and region feature. Li, X. M [6] proposed a global translation preferred active contour, they defining a new inner product on the set of perturbations of a curve, the new inner product make the curve prefers global translation when evolve. This global translation preferred active contour is suitable for object tracking.

This paper has some similarities to work described by [7], Chang, J. S uses the region competition image segmentation approach to extract the hand. This region competition method needs a specific distribution of the target region, but both the C-V model and the geodesic model which we use in our method don’t need any prior knowledge.

In this paper, we propose a hybrid tracking method based on active contour and Mean Shift. This method can track target fast and can present target exactly. The details of our method are present in section 2, the experiments results and the conclusions are provide in section 3 and section 4.

II. THE HYBRID TRACKING METHOD

A. The overview of the tracking method

Figure 1 is the flow chat of the hybrid tracking method. As shown in the flow chat, at the beginning of the tracking we exact the silhouette contour of the target and compute the target’s feature for Mean Shift tracker. In our method, a weighted histogram computed from the region inside the target’s silhouette is used as the feature of target. On the next frame, we first compute the target’s translation by the Mean Shift procedure, then we translate the contour which we get from last frame to the right place according the translation and evolves the contour by minimizing the contour energy. In our method, C-V Active Contour model or Geodesic Active Contour Model is used. After getting the contour description target’s silhouette, we update the tracking information and go to the tracking of succeeding frame. The tracking course is shown in figure 2.

Figure 1. Flow chat of the tracking method
B. The translation tracking based on Mean Shift

The target model which center is \( x_0 \) is computed as:

\[
\hat{q}_u = C \sum_{i=1}^{n} k \left( \frac{x_i^t - x_u}{h} \right)^2 \delta \left[ b \left( x_i^t \right) - u \right] \tag{1}
\]

where \( k(x) \) is a convex and monotonic decreasing kernel silhouette, assigns smaller weights to pixel farther from the center. \( h \) is the bandwidth. \( \delta \) is the Kronecker delta function. The normalization constant \( C \) is computed as function (2):

\[
C = \frac{1}{\sum_{i=1}^{n} k \left( \frac{x_i^t - x_u}{h} \right)^2} \tag{2}
\]

Similarity the target candidates centered at \( y \) in the current frame is given by:

\[
\hat{p}_u(y) = C \sum_{i=1}^{n} k \left( \frac{x_i^t - y}{h} \right)^2 \delta \left[ b \left( x_i^t \right) - u \right] \tag{3}
\]

Then, the tracking problem turn out to find out the best \( y \) make the \( \hat{p}_u(y) \) has the biggest similarity with \( \hat{q}_u \). We chose the Bhattacharyya coefficient between \( \hat{p}_u(y) \) and \( \hat{q}_u \) as the sample estimate of similarity:

\[
\rho \left[ p(y), q \right] = \rho \left[ \hat{p}_u(y), \hat{q}_u \right] = \frac{1}{2} \sum_{s=1}^{n} \sqrt{p_s(y_q) q_s} \tag{4}
\]

Using taylor expansion around the values \( \hat{p}_u \left( y_{\text{opt}} \right) \), the linear approximation of the Bhattacharyya coefficient is obtained after some manipulations as:

\[
\rho \left[ p(y), q \right] \approx \frac{1}{2} \sum_{s=1}^{n} \sqrt{p_s(y_q) q_s} + \frac{1}{2} \sum_{s=1}^{n} p_s(y) \frac{q_s}{p_s(y_q)} \tag{5}
\]

Recalling function (1)(3) results in

\[
\rho \left[ p(y), q \right] \approx \frac{1}{2} \sum_{s=1}^{n} \sqrt{p_s(y_q) q_s} \\frac{q_s}{P_s(y_q)} \tag{6}
\]

where

\[
w_j = \sum_{i=1}^{m} \delta \left[ b(x_i) - u \right] \frac{q_s}{P_s(y_q)} \tag{7}
\]

Observe that the second term of (6) represents the density estimate computed with kernel profile \( k(x) \) at \( y \) in the current frame, with the data being weighted by \( w_i \). The mode of this density in the local neighborhood is the sought maximum which can be found employing the mean shift procedure.[1,8]. In this procedure the kernel is recursively moved from the current location to the new location according to the relation

\[
\hat{y}_{j+1} = E \left[ \frac{x_i w_i g \left( \| y_j - x_i \| ^2 \right) - 1}{h^2} \right] \tag{8}
\]

When \( \| \hat{y}_{j+1} - y_j \| \leq \varepsilon \), stop iterating. \( \hat{y}_{j+1} \) is the finally location result of current frame.

Let \( y \) be the location result of last frame, the translation of the target in current frame \( \Delta y \) can be get by the function(9).

\[
\Delta y = \hat{y}_{j+1} - y \tag{9}
\]

C. Target’s accurate presentation by evolving active contour

After get the translation of target, we use the active contour to describe the target silhouette accurately. In according to Kass et al.[5] in their classical snake approach that the problem of finding a curve that best fits the object boundary equals to the minimization over all closed planar curves \( \ell \left( q \right) \) in the plane.

1) Edge based active contour model : Geodesic model

In 1997, Castris et al.[11] proposed the Geodesic active contour model, obtained a minimal path representation of the energy.

\[
E(\ell) = \int_0^1 g \left( I(\ell(q)) \right) \xi''(q)^2 dq \tag{10}
\]

where \( g : [0, \infty] \rightarrow R \) is a strictly decreasing function, that is, \( g(r) \rightarrow 0 \) as \( r \rightarrow \infty \).

To minimize this energy function by steepest descent, consider curve \( \ell \) to be a function of time \( t \). The Euler-Lagrang equation yields the curve evolution equation
\[
\tilde{\epsilon} = \frac{1}{k} \frac{\epsilon}{k} - (\nabla g \cdot \vec{N}) \vec{N}
\]

(11)

where \( k \) is the Euclidean curvature, and \( \vec{N} \) is the unit normal vector of the surface. Eq(11) is equivalent to updating a volume map \( \phi \) with

\[
\phi_i = \frac{g}{k} |\nabla \phi| + \nabla g \cdot \nabla \phi
\]

(12)

2) Region based active contour model : C-V model

C-V active contour model is an effective model which based on Mumford – Shah functional for segmentation [10], this model can detect objects whose boundaries are not necessarily defined by gradient. In this model, the image \( u_0 \) assumed is considered as component by two regions of approximatively piecewise-constant intensities, and the boundary of these two parts is smooth. On this assumption, the energy function is defined:

\[
\inf_{\phi \in \phi} \left\{ E [c_1, c_2, \phi(u_0)] = \lambda \int (u_0 - c_1)^2 H(\phi(x,y)) dxdy + \lambda_2 \int (u_0 - c_2)^2 (1 - H(\phi(x,y))) dxdy + \mu \int \delta(\phi(x,y)) |\nabla \phi(x,y)| dxdy \right\}
\]

(13)

where the \( u_0 \) is the image, \( \phi \) is the level set function, \( \phi > 0 \) denote the area outside curve, \( \phi < 0 \) denote the area inside curve, and the constants \( c_1 \) and \( c_2 \) are the averages value of those area. \( \lambda_1, \lambda_2 \) and \( \mu \) are positive parameters. \( H(\bullet) \), \( \delta(\bullet) \) are Heaviside function and Dirac function. the first and the second term of function(13) are fitting terms, and the third term is the regularizing term. According to function (14) the level set function \( \phi \) is updating by:

\[
\frac{\partial \phi}{\partial t} = \mu \cdot \text{div} \left( \frac{\nabla \phi}{|\nabla \phi|} \right) - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2
\]

(14)

In our method, before use function (12) or (14) to updating the level set function, \( \phi \) is initialized by:

\[
\phi(t+1) = \phi_{t+1}(y - \Delta y)
\]

(15)

Where \( \phi_{t+1} \) is the level set we get from last frame.

We doing experiment with both C-V active contour model and Geodesic active contour model, the result turn out that, C-V model will get good result for the object with mean intensity, for the object with obvious boundary Geodesic model will be more adaptive. In order to reduce the computation, we set a ROI (region of interesting) for each target, don’t need to evolve contour in whole image.

III. EXPERIMENT RESULTS

The hybrid tracking method was applied to many sequences. Here we present some results on car, head and hand tracking. The tracking results are shown in figure 3-5, the car sequence (figure 3) has 200 frames of 320 * 240 pixels. Because of the complex background, we use the Mean Shift tracker with Geodesic active contour method to track the car in these frames. We implement our method on Matlab, and it takes 0.1s to handle each frame in average. From the results, we can tell that even in dynamic background, our method can track car accurately when it has translation and rotation. For the frames of human head and hand, we associate the Mean Shift tracker with the C-V active contour to get the silhouette of object. The tracking results are shown in figure 4 and 5, the iteration number of active contour is reduce to 12 and 18 from 55 and 72 after adding Mean Shift tracker.

![Car tracking in dynamic background](frame10.png) ![Head tracking](frame1.png)

![Head tracking](frame33.png) ![Head tracking](frame66.png)

![Head tracking](frame148.png) ![Head tracking](frame189.png)

![Human hand tracking](frame51.png) ![Human hand tracking](frame99.png)

![Human hand tracking](frame70.png) ![Human hand tracking](frame234.png)
IV. CONCLUSION

Based on the fact that the active contour can give accurate description for object, but need huge computation, we proposed a hybrid tracking method. Before the contour evolving, the translation of target has been get by Mean Shift procedure. So the accurate state of target will get fastly. By the way, because we use the contours to describe targets, the disturbance of background is reduced.

It should be pointed out that, for the targets don’t have obvious boundary nor the mean intensity, the active contour may not get the right curve. Finding a more effective energy model will be our future work.

REFERENCES


