Object Tracking Using an Improved Kernel Method

Yuan Chen, Shengsheng Yu, Weiping Sun, Xiaoping Chen
College of Computer Science & Technology
Huazhong University of Science & Technology
Wuhan, HuBei, China, 430074
E-mail ychen@smail.hust.edu.cn

Abstract

An improved object tracking scheme is presented based on the Kalman filter and mean-shift approach. And this scheme is robust to disturbance and occlusion of both the object and the scene. The object is selected by using FG/BG detection and represented by its center point and probability distribution. The mean-shift approach estimates the object position based on the result of the Kalman filter. A threshold of the Bhattacharyya coefficient is set to judge occlusion and when object being occluded the Kalman filter estimates the object position. Since the proposed scheme combines the space information with probability distribution, it is robust to disturbance and occlusion.

Keywords: object tracking, Kalman filter, mean shift, Bhattacharyya coefficient

1. Introduction

The purpose of computer vision is to provide computer with the ability to understand the image or video, locate and identify objects, trace their motion, interpret the relationship among them, etc. As an important element in the field of computer vision, object tracking is to generate the trajectory of the specified object over time by location its position in each frame of the video. And it is often combined with object detecting and behavior recognizing as a whole system to analyze a sequence of images which also expressed as video. There are numerous approaches for object tracking proposed before [1]. The mean-shift method is an effective one, which is based on kernel tracking. The object being tracked is represented by a primitive region such as rectangle, ellipse, and etc. A weighted histogram is computed from this region and then a mean-shift iterative procedure is performed to find the most likely region in the next frame. The mean-shift algorithm was first proposed by Fukunaga and Hostetler [2]. By it was ignored for about two decades until Cheng introduced it into the field of computer vision [3]. Comaniciu used this method in image segmentation and object tracking [4]-[6].

The Mean-shift tracking approach is to search the peak value of probability density in the appearance region and is an effective method in middle level of computer vision. It is suit for tracking objects no matter when their shape or size changing during the tracking procedure. Even sometimes the object is partially occluded by other moving objects or stationary scenes in the image. But when the whole appearance region of tracking object disappears during a serial of frames, the mean-shift method will lost the tracking object. If the object is close to another one with very similar appearance, the mean-shift tracking windows may be attracted onto the other object in the procedure of searching the peak value. So we propose an improved object tracking algorithm.

In our object tracking algorithm, the Mean-shift tracking approach is combined with the Kalman filter to improve the tracking speed and accuracy. Firstly, the object is subtracted from a video frame by using FG/BG detection method. Secondly, the specified object is represented by its pixels color probability distribution. In our test system, each frame of the video is transformed from RGB color space to HSV color space. The center point position of tracked object is also calculated for using as a state variable in the Kalman filter. Then our tracking iteration is performed by estimating the state of tracked object with the result of the Mean-shift tracking approach in a normal condition. When the object is occluded, we use the Kalman filter to estimate the object position. The Kalman filter is also used to initialize the Mean-shift tracking region in each frame.

The rest of the paper is organized as follows. In Section 2, we describe the whole flow path of our tracking system. Next, the proposed tracking method is elaborated in detail in Section 3. The tracking results of the test sequences and the performance study of our scheme are presented in Section 4. Finally, Section 5 provides conclusions and discussion.
2. Tracking system

We propose a robust object tracking method that can be put to work in many applications, such as smart video surveillance, etc. Before the object tracking procedure, we should select the moving object of interest from the image. So a background modeling method is imposed to detect foreground and background. The background pixels can be represented in the Mixtures of Guass or Hide Markov Model. Then the initial tracking region is calculated according to object contour or other appearances. After setting a proper tracking region, we will carry out our object tracking scheme in the next frames. The FG/BG detection is imposed here only for the initialization of state variables and it will be performed again if there is a necessary to adjust the tracking errors. If the background scenes are stable, the moving object can be tracked by calculating the minimum distances of foreground objects. But when camera moving, panning, tilting, or zooming, both the foreground and background are moving. Our object tracking will be effective to track object under these conditions. The whole tracking system flow path is shown in Figure 1.

3. Improved object tracking scheme

3.1. Object representation

As has mentioned in Section 2, the Kalman filter and the mean-shift approach will be imposed during object tracking. In the Kalman filter, the current state variable is estimated by the previous state variable and the last measured value. So the variables to be estimated and measured are the object’s position in the frame. We select the object’s center point as the tracking point and define \( \hat{x}_k \) as the state variable, \( y_k \) as measured value and \( \hat{y}_k \) as estimated value.

In the mean-shift approach, the tracking object is represented as a primitive region in the frame, such as rectangle, ellipse, etc. Sometimes, this region can even be selected along the object contour. Then, the object model is normalized to eliminate the difference of all kinds of objects in scale.

The primitive region of target object in the frame is represented as \( \{x_i, i=1,...,n\} \). Let \( K(x) \) be the kernel function with a convex and monotonic decreasing kernel profile. The function \( h : \mathbb{R}^2 \to \{1...m\} \) associates to the pixel at location \( x_i \). The probability of the feature \( u=1...m \) in the target object model is represented as

\[
\hat{q}_u = C \sum_{i=1}^{n} k \left( \frac{x_i - X_0}{h} \right) \delta \left( b(x_i) - u \right),
\]

where \( \delta \) is the Kronecker delta function. Since the summation of delta function for \( u=1...m \) is equal to one, the normalization constant \( C \) is derived that

\[
C = \frac{1}{\sum_{i=1}^{n} k(\|x_i\|^2)}.
\]

As to candidate object model, the same kernel function \( K(x) \) is adopted and the bandwidth is \( h \). The probability of the feature \( u=1...m \) in the candidate object model is represented as

\[
\hat{p}_u(y) = C_h \sum_{i=1}^{n} k \left( \frac{y - X_i}{h} \right) \delta \left( b(x_i) - u \right),
\]

where the normalization constant \( C_h \) is derived that
The similarity function is denoted by
\[ \hat{\rho}(y) \equiv \rho[p(y), q]. \] (5)

The object’s color probability is adopted to represent the tracked object and we transform the color image from RGB color space to HSV space. So we calculate the H component probability to represent the tracked object. By using this method, the tracking speed is increased and the tracking accuracy is not descent according to the experiment in the next section.

### 3.2. Automated object initialization

The object to be tracked should be selected from the frame before performing tracking procedure. In our scheme, we acquire the object’s region and position by combining with the FG/BG detection. The Mixtures of Guass or Hide Markov Model is adopted for FG/BG detection. The variables of Kalman filter and the mean-shift approach are initialized by the region and position of the detected object.

### 3.3. Occlusion threshold

The Bhattacharyya coefficient is calculated to evaluate the similarity of the target object model and candidate object model

\[ \hat{\rho}(y) \equiv \rho[p(y), q] = \sum_{u=1}^{m} \sqrt{\hat{p}_u(y)q_u}. \] (6)

When the tracked object occluded by other objects or background scenes, the Bhattacharyya coefficient will descent. So we define a threshold \( T_h \) to determine whether the occlusion happens. If we find that \( \rho[p(y), q] < T_h \), the tracked object may be occluded or disappear.

### 3.4. Object tracking

The tracking procedure is performed after the state variables initialized by FG/BG detection. When the tracked object is not occluded, the center point’s position is estimated by Kalman filter. And the result of Kalman filter estimation is the start position of mean-shift approach. By using this method, the iteration loops of mean-shift can be reduced. The results of mean-shift approach are set to be the tracked object’s actual state values.

\[ x_k = A_k x_{k-1} + \omega_{k-1}, \] (7)

is the system state equation and

\[ y_k = C_k x_k + \nu_k, \] (8)

is the measurement equation, where \( A_k \) is the state transfer matrix, \( C_k \) is the observation matrix. The random variables \( \omega_k \) and \( \nu_k \) are respectively represent the process and measurement noise. They are assumed to be independent, white, and with normal distribution

\[ \omega_k \sim N(0, Q), \] (9)
\[ \nu_k \sim N(0, R). \] (10)

The recurrence formula is presented to search the estimation \( \hat{x}_k \) of \( X_k \) by using the Kalman filter

\[
\hat{x}_k = A_k \hat{x}_{k-1} + H_k (y_k - C_k A_k \hat{x}_{k-1}), \quad (11)
\]
\[ H_k = P'_k C'_k (C'_k P'_k C'_k + R_k)^{-1}, \quad (12) \]
\[ P'_k = A_k P_{k-1} A_k^T + Q_{k-1}, \quad (13) \]
\[ P_k = (I - H_k C_k) P'_k. \quad (14) \]

The Kalman filter estimates the current value by the previous state variable and the last measured value and is suitable for linear system. Since the motion of the tracked object is nonlinear, we impose the mean-shift approach to track the object based on the result of the Kalman filter.

As mentioned before, the Bhattacharyya coefficient is calculated to evaluate the similarity of the target object model and candidate object model. The linear approximation of the Bhattacharyya coefficient is obtained by using Taylor expansion

\[ \rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^{m} \sqrt{p(y_0)q_u} + \frac{1}{2} \sum_{u=1}^{m} p_u(y) \sqrt{q_u}. \] (15)

Recalling (3) results in

\[ \rho[p(y), q] \approx \frac{1}{2} \sum_{u=1}^{m} \sqrt{p(y_0)q_u} + \frac{C_k}{2} \sum_{i=1}^{n} w_k \left\| y - x_i \right\|^2. \] (16)
where

\[ w_i = \sum_{u=1}^{n} \delta[b(x_i) - u] \sqrt{\frac{q_u}{P_n(y_0)}}. \]  

(17)

The new position is derived by using mean-shift approach

\[ \hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i \omega_i g\left(\frac{\hat{y}_0 - x_i}{h}\right)}{\sum_{i=1}^{n_y} \omega_i g\left(\frac{\hat{y}_0 - x_i}{h}\right)} \],

(18)

where \( g(x) = -k'(x) \).

If the tracked object is not occluded, the Kalman filter state value is updated by the mean-shift result and then tracking in the next frame. If in opposite situation, the Bhattacharyya coefficient is less than \( T_h \), the object’s position is estimated by the Kalman filter in consecutive frames until the Bhattacharyya coefficient is greater than \( T_h \).

4. Experiments

We tested our improved object tracking scheme based on some special sequences. The object to be tracked in these sequences is sometimes partially occluded and sometimes fully occluded. The scheme was tested on a PC with Intel P4 2.4G CPU and 512 M memory. As shown in Fig 2, the tracked object was fully occluded. The frame has 320×240 pixels and the frame rate is 15 fps. The tracked object was detected and tracked in frame 68 in Figure 2. From frame 84 the objects was partially occluded and in frame 97 it absolutely disappeared. In frame 110, the object was tracked when it reappeared. The Bhattacharyya coefficient during tracking is shown in Figure 3.

From Figure 4 we can see that three person coming to the camera and the two on the right side dressed in similar color. In frame 90, the tracked one was selected. The two in right side were very close to each other in frame 119. Another one was passing in front of the tracked one in frame 197.
We compared the improved object tracking scheme to the mean-shift scheme. The first sequence has 386 frames and the second has 3701 frames while we selected 712 frames for testing. It is obvious that the improved object tracking scheme consume less time than the mean-shift scheme.

Table 1 Object tracking

<table>
<thead>
<tr>
<th>sequence</th>
<th>scheme</th>
<th>frame</th>
<th>Time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Mean-Shift</td>
<td>386</td>
<td>51</td>
</tr>
<tr>
<td>1</td>
<td>ours</td>
<td>386</td>
<td>43</td>
</tr>
<tr>
<td>2</td>
<td>Mean-Shift</td>
<td>712</td>
<td>103</td>
</tr>
<tr>
<td>2</td>
<td>ours</td>
<td>712</td>
<td>87</td>
</tr>
</tbody>
</table>

5. Conclusion

We propose an improved object tracking scheme with the combine of the Kalman filter and the mean-shift approach to reduce the tracking time and this scheme is robust to disturbance and occlusion. Since the Kalman filter is adopted to estimate the start position of the mean-shift, the speed is more important than accuracy. As a result, we did not adopt nonlinear filters. With the scale changes of the tracked object, the kernel bandwidth will be adjusted adaptively. The tracking speed and accuracy can be improved even more if we simplify the object representation according to particular application.

6. References


