Research on Object Tracking Based Reliable Partial Patches Algrithm

Xueming Zhai[1, a], Mingming An[1]

[1]School of control and Computer Engineering, North China Electric Power University, Baoding 071000, China.

[a]664313357@qq.com

Abstract

Moving object tracking is an important research subject in the field of computer vision,but it is still difficult to achieve an efficient and accurate tracking result when the deformation or occlusion problem of the object occurs under the existing research conditions.In this paper, the star-shaped partial patches structure is adopted to make the representation of the object more robust. At the same time, the HSV eigenvalue combined with the Random ferns are used to characterize the patches. In addition,Markov chain is used to predict the object position at the next time,it greatly improves the accuracy of positioning the object. Finally, the occlusion problem can be solved by calculate the like hood of the foreground and background. Experiments show that using the algorithm of this paper can greatly improves the accuracy of the object tracking.

Keywords

Object tracking; partial patches; Markov chains; occlusion problem.

1. Introduction

Currently, the problem of moving object tracking in video surveillance is still a hot issue in the field of computer vision [1]. At the same time, the mismatch and loss of the object because of the deformation or occlusion problems are still the main problems.At present, the main method of tracking the object is to make a model of the target area, and then get feature matching from the image frame to determine the position of the target. However, when the target is deformed or occluded, it is easy to lose key features, resulting in target matching failure.

In recent years, many literatures have proposed corresponding methods for the problem of object deformation.Among them,the method of using the partial patches model and making the model updatin can effectively solve this problem.However, when the object is occluded,the update of the model will result in the loss of the occlusion patches,so that the purpose of accurately object tracking cannot be achieved.

At present, the mainstream methods of dealing with occlusion problem can be roughly divided into three categories: motion prediction, regional weighting and patches models. Among them, motion prediction can solve the situation where the objectt is severely occluded or completely occluded. This method is mainly based on the information of the position, direction, and speed of the object before it is occluded,and state information is used to predict where the object may appear in the next frame [2]. The shortcomings of this method are also obvious,int is only applicable to the case where the object motion has a certain regularity,so it is difficult to apply it in reality; regional weighting refers to adopting a different weight to different regions. The rules of the value of the patches to match the object is that the largeer the weight the closer to the center position.The disadvantage is that it can only reduce the influence of occlusion on the edge parts. It can not be solved when the center position is occluded; the patches model refers to the separation of the object from a whole into several partial blocks, and then these patches are determined by certain method. The features of the partial blocks are integrated as a complete object model. This approach is more flexible than region weighting and it can handle more complex occlusion situations.

For the processing of partial features,Hua Bao[3]divided the object into average rectangular blocks, while Zhiqiang Hou[4]and others adopted a random partial block model;Tu Ting[5]proposed an affine-invariant SIFT image matching algorithm,and it greatly improves the accuracy of object tracking. In dynamic scenes,Dong Wang[6]proposed a method based on background compensation to perform object tracking. Modeling the foreground and background can effectively improve the accuracy of every point in the image to the foreground and the background, which is also helpful for occlusion problem.

In order to track the object more efficiently and accurately, and solve the occlusion problem at the same time, this paper proposes an object tracking algorithm based on adaptive partial patches model, which performs trusted partitioningand and updates the paetches in real time,and deletes operations to reduce the impact of deformation on target tracking. At the same time, the occlusion problem is solved by using the similarity between the background and foreground. When the ratio between the foreground template and the background template exceeds the set threshold, it is determined that occlusion occurs, and the update of its model is stopped; when the object occlusion disappears, the model update is resumed. This method can effectively reduce the impact of occlusion on object tracking.

2. Reliable Partial Patches Model

2.1 Model Definition

In this paper, a star-shaped structure is used to characterize the object, that is, the position of the object center and each patch center form a star-shaped structure. Among them, the object position is determined by the center position of the object, and the patch center is determined by the distance between the patch center and the objectt center. The model is flexible in structure and it has a good adaptability to object deformation and apparent changes. The specific structure is shown in Fig.1.

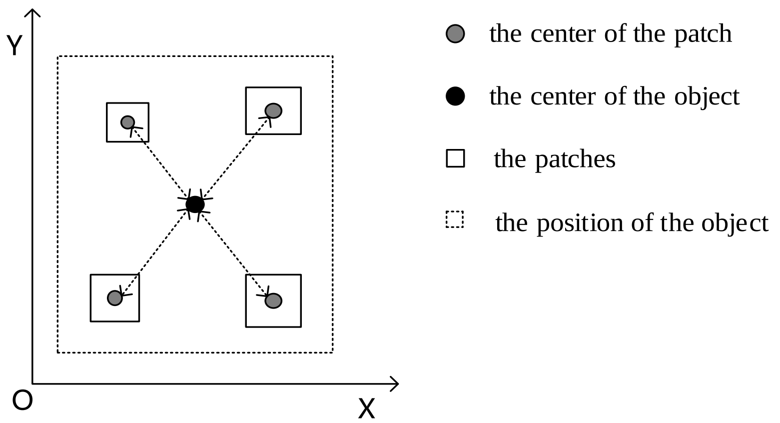

Fig 1. The structure of the patches

Xti represents the center of the patch, Xtc represents the center of the object, and Rti represents the distance from the itch patch to the object center. The position of the object center TX is defined as

$$X_t^c = \sum_{i=1}^{n} \omega_i X_t^i \tag{1}$$

ωi is the normalized weight of the itch patch.

2.2 Eigenvalue Representation

2.2.1 HSV Eigenvalue

The description of the object is mainly to describe the characteristics of the patches and the locations. This paper uses a normalized HSV color histogram to describe the color characteristics of the object, and the color histogram of the patch template can be expressed as Q={q1,q2,...,Or}, where r is the dimension of the histogram.

2.2.2 Random Ferns Eigenvalue

This paper proposes a method to characterize the patches based on the Random ferns structure. A Random ferns structural element is a block of a fixed size of m×n, which randomly generates a pairwise relationship for all pixels.Suppose the above pixel value is Pix1, and the following is Pix2, you can get a binary number:

$$B_j = \begin{cases} 1, Pix1 \geq Pix2 \\ 0, Pix1 < Pix2 \end{cases} \tag{2}$$

Assuming that the number of pixels in the itch patch is seethe bit binary is/2 can be used to characterize this patch. Considering the adaptability to the change of the moving object and the updating of the feature, the structural element is randomly divided into K groups, each group containing M structural elements, so that is/2=K×M is satisfied. The characteristics of each patch can be represented by K values, and the range of each value is 0~2M-1.

3. Matching Algorithm Based on Markov Model

3.1 Markov Chain

The Markov chain,is named by Andrey Markov (1856-1922),which is a discrete-time stochastic process with Markov properties in mathematics [7].Xt denotes the value of the random variable X at the time of discrete time t. If the probability of transition of the variable over time depends only on its current value,then $P(X_{t+1} = s_j \mid X_t = s_0, X_t = s_1, ..., X_t = s_j) = P(X_{t+1} = s_j \mid X_t = s_i)$.

That is to say, the probability of state transition depends only on the previous state. This variable is called a Markov variable, where s0, s1, ⋯, si, sj ∈ Ω are the possible states of the random variable X. This property is called Markov property, and a stochastic process with Markov property is called Markov process.

3.2 Central Location Estimatation

In this paper, the Bayesian Metropolis sampling algorithm is used to sample and predict the target center position at the next moment. Assuming that the number of generated samples is m, a Gaussian formula is used to generate a random sample.

The first sample value is related to the target center position at the previous moment, and the first sample is generated by Gaussian perturbation:

$$S_0(X_t^{c(1)}) = G(\hat{X}_{t-1}^c, \sigma^2) \tag{3}$$

The remaining m-2 sample values are related to the sample position last generated, and the generating formula is:

$$S_l(X_t^{c(l)}) = G(\hat{X}_{t-1}^c, \sigma^2) \tag{4}$$

In the process of the Metropolis sampling algorithm, a new value is generated at each time, the new value is accepted or rejected according to a certain probability. The acceptance probability formula is:

$$\alpha_l = \min\left(1, \frac{p(X_t^{c(l)} \mid X_{t-1}^c)}{p(X_{t-1}^c)}\right) \tag{5}$$

Generate a random number a from 0 to 1. When $\alpha \geq a$, accept the new sample; otherwise reject the new sample, the sample value is still taken the same value.

According to the Bayesian probability filtering formula, we set the posterior probability of the object state at time t to p(Xt|Y1:t). Here, the state value at time t is shown as Xt, and the observation value from the start to the time t is indicated asY1: t. Then the position of the object center can be predicted as follows.

$$\hat{X}_t^c = \max p(X_t^{c(l)} \mid Y_{1:t}) \qquad l=1,\ 2,\ m \tag{6}$$

After calculating the maximum posterior probability, the optimal target center estimation position at time t $\hat{X}_t^c$ can be obtained. At this time, the estimated position of each patch can be obtained by the formula7.

$$\hat{X}_t^i = \hat{X}_t^c + R_t^i \tag{7}$$

3.3 Similarity Measure

After predicting the patch position at time t, a similarity measure method is needed to evaluate the match results of the model.

In order to fully consider the spatial structure, color features and eigenvalues of the block, we set a similarity index , it can be defined as:

$$\lambda_i \propto \exp(-\| F(\hat{D}_t^i) - F(D_t^i) \|^2) \times \exp(-\| \hat{R}_t^i - R_t^i \|^2 \times \| H(\hat{D}_t^i) - H(D_t^i) \|^2) \tag{8}$$

Wherein, $\hat{D}_t^i$ represents the estimated position of the itch partial patch at time t; D_t^i represents the template of the itch partial block at time t; $F(\hat{D}_t^i) - F(D_t^i)$ represents the difference in eigenvalue between the sample and its corresponding model; $\hat{R}_t^i - R_t^i$ represents the distance offset between the sample and its corresponding model; $H(\hat{D}_t^i) - H(D_t^i)$ represents The difference between the HSV histogram mean between the sample and its corresponding model.

3.4 Matching Strategy

It can be seen from the above that the estimated center position of the itch block at time t can be obtained by the formula7.At this time, with $\hat{X}_t^i$ as the block center, select 4 times the area of the patch as the search range, and perform local poor search in this area. Calculate the similarity $\lambda$with the patch for each position, and take the position with the largest similarity as the position of the patch at time t.

After the patch position is determined, its feature value F, distance R and color feature H are updated respectively into the patch model.

4. Model Update Strategy

4.1 Match Threshold Setting

The update strategy of the model is mainly divided into three cases: when the patch matching is successful, the patch position and the patch feature value of the local model are respectively updated; when the patch matching fails, the patch information of the local model is deleted and adds a new patch to the object position according to the rule, then updates the feature value and position of the new pathch to the local model; when the patch match fails and is determined that occlusion occurs, the local model of the patches is not updated. Finally, update the center position coordinates of the local model.

In the process of tracking, this paper sets the threshold T_θ to determine whether the partial image patch belongs to the object target. When $\omega_t^i < T_\theta$, the patch model is removed and a new patch is added at the appropriate position.

When the weight of the patch is too small, the patch model should be removed. In this case, a new patch needs to be added at the target position, that is, the patch model is updated. Care must be taken when adding patches that the background cannot be added to the model, which can easily lead to the loss of the object. And don't add new patches where the patches are dense.

4.2 Update of Related Values

The object feature model established in this paper is a description of the eigenvalues of all moving targets. Real-time updating of the object model can effectively prevent the problem that the effective feature value cannot be successfully matched due to the change of the target eigenvalue, and it can improve the accuracy of the matching.

When the patch matching degree is greater than T, that means the patch matching is successful, then the feature of the patch model is updated. The update strategy at this time is: matching the feature point retention succeeded, and matching the failed feature point deletion. The new patch position that completes the matching is re-determined according to the matching feature point and the relative position of the patch. All the feature points in the patch are calculated, and the feature values are updated into the object model.

When the target match is successful, the coordinates of the object center are updated to the new center coordinate position (x', y'). At the same time, the position of the patch from the center of the object will also change, and the update method is Rti'=Xti'-Xtc'.When the patch is removed, the center, distance and other related information of the patch are also deleted; when the patch is added, its related information are updated into the model information.

4.3 Judgment of Occlusion Problem

When the object is occluded, if the object model is still updated, the occlusion information will be updated to the object template, and part of the object information will be lost, so that the object will be easily lost. Therefore, this paper adopts a certain strategy to judge the occlusion problem. When the object is judged to be occluded, the occlusion part of the patch no longer performs the model update operation.

Set the similarity between the block and the background to be b. When the ratio of the patch to the likelihood of the object and the background is less than a specific threshold θ, it is determined that occlusion occurs, and the update operation of the patch is stopped. At this time, the similarity between the patch and the background is far greater than the similarity between the patch and the object, and is determined as the occlusion of the object from the background.

5. Experimental Process and Results

5.1 Algorithm Description

The algorithm description of the whole process is shown in Table 1.

5.2 Experiment Setup

In order to verify the feasibility of the method in this paper, 20 public video sequences were selected from the OTB-50 dataset for verification and testing. These sequences contain a variety of challenging factors such as fast motion, attitude changes, deformation, scale changes, and partial occlusion.

All experiments were performed on a personal notebook with a clock speed of 2.6 GHz, a core of Intel Core i3, and a memory of 6 GB of RAM. The software simulation environment is Matlab R2016a. For comparison purposes, the code for these trackers is provided by the original author or from the official website of OTB-50.

In all experiments, we set the parameter experience to: the number of patches n = 25, the number of samples m = 50, update threshold T_θ=0.0001,occlusion threshold θ=10-20.

Table 1. Algorithm Description

Algorithm: Object tracking algorithm based on reliable partial block model	
Input:	The position of the object at time t-1
Output:	The position of the object at time t
1	For in=1 to m
2	Produce sample in;
3	Calculate αin;
4	End
5	Calculate $\hat{X}_t^c$;
6	For i=1 to n
7	$\hat{X}_t^i{}' = \hat{X}_t^c + R_t^i$;
8	$T_t^i = \left[\hat{X}_t^i + 2r_t^i\right]$;
9	End
10	In T_w^i
11	Calculate max (λ_i);
12	Calculate ω_t^i;
13	Calculate λ_b;
14	If $\omega_t^i < T_\theta$
15	Updating the model
16	Add a new patch;
17	Remove the old patch;
18	Add the new message to the model;
19	End
20	If $\lambda_b < \theta$
21	Stop updating;
22	End
23	End
24	For i=1 to r
25	$\hat{X}_t^i = \max_{\lambda_i} \hat{X}_t^i{}'$;
26	End

5.3 Tracking Results Display

Taking "davaid3" video sequence as an example, this paper used Struck algorithm, SCM algorithm and the algorithm in this paper to track and detect pedestrians in the sequence. The tracking results are shown in Fig.2. Among them, the red line box represents the tracking result of the Struck algorithm, the green line box represents the tracking result of the SCM algorithm, and the blue line box represents the algorithm used in this paper.

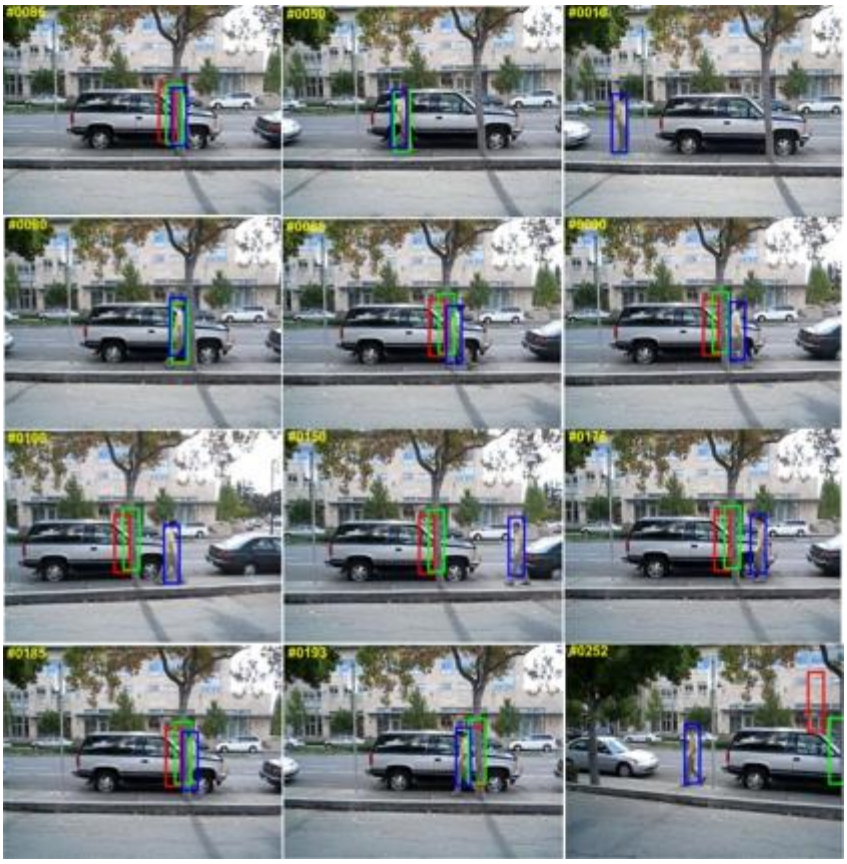

Fig 2. Comparison of three methods tracking results

According to the tracking results in the above figure, the results of the object tracking using the algorithm this paper are significantly better than the other two. When passing a large obstruction, this algorithm captures the object more accurately and does not lose the object. In addition, when the object performs irregular motion (round-trip motion), the target is not lost.

6. Conclusion

In actual scenarios, it is difficult to deal with significant apparence changes in the tracking process, such as occlusion, scale changes, deformation, fast motion, and so on. In this paper, the structural features of the object are well preserved by modeling the star structure of the obkect, and the object is represented by the reliable image patches in the local layer to improve the robustness of the model to occlusion and scale change. By predicting the object center to determine the approximate location of the patch, the workload for predicting the position of each patch is greatly reduced, thereby improving the efficiency of matching. In addition, the matching of the HSV feature and the Random ferns structural feature of the patch is used ,and it effectively reduce the mismatch caused by the single mode, so as to achieve a better matching effect.

Test results on open and challenging tracking sequences and standard visual data sets shows that the method in this chapter have significant advantages over current trackers in terms of performance, especially in terms of solving occlusion problems. In addition, we believe that the star model proposed in this paper can be used as a basic framework for target tracking. It can be integrated into different feature representation methods and sampling methods to further improve the performance of the tracking method. This is also a future research direction.

References

[1] Kaiqi Huang, Xiaoyu Chen,Yunfeng Kang, et al. Overview of Intelligent Video Surveillance Technology[J]. Chinese Journal of Computers, 2015, 20(6): 1093-1118.

[2] Xiancai Zhang, Yang Li, Yulong Xu, Jiabao Wang, Zhuang Miao.Scale-adaptive fast tracking method based on position prediction [J]. Journal of PLA University of Science and Technology: Natural Science Edition, DIO: 10.12018/j.issn.1009G3443.20160901002/2017.03.01.
[3] Zhiqiang Hou , Anqi Huang , Wangsheng Yu, et al. Visual tracking algorithm based on partial block and model updating[J].Journal of Electronics & Information Technology, 2015, 37(6):1357-1364.
[4] Ting Tu, Jixiang Lu. SIFT image matching algorithm based on block and affine invariance [J/OL]. 2018, 35(12). [2017-12-08]. http://www.arocmag.com/article /02-2018-12-065.html.
[5] Dong Wang, Hong Zhu, Kai Kang, et al. Target tracking algorithm in dynamic scene based on background compensation guidance [J]. Journal of Scientific Instrument, 2014, 35 (6): 1433-1440.
[6] Yulong Xu, Jiabao Wang, Yang Li, et al. Scale adaptive target tracking based on correlation filtering [J]. Journal of Computer Applications, 2016, 33(11): 3513-3516.
[7] Andrieu C, Doucet A, Holenstein R. Particle Markov chain Monte Carlo methods [J]. Journal of the Royal Statistical Society, 2010, 72(3):269-342.