

The Improvement of Mean - Shift Algorithm and Kalman Filter of Tracking Moving Targets in the video of Robotic Fish

Jingying Wu^{1, a}, Wei Xiang^{1, b, *}

¹ Institute of Electrical and Information Engineering, Southwest University for Nationalities, Chengdu, Sichuan, China 610041

^adream.kana@yahoo.com, ^bwei.xiang@foxmail.com, *corresponding author

Keywords: Mean shift algorithm; Kalman residuals; Kalman filter; Target tracking

Abstract. Several defects have been found in the former experiments of basic Mean shift algorithm, such as the erroneous judgments when target blocked by a large proportion of barrier. Thus, a fast target tracking algorithm based on Mean shift combined Kalman filter is proposed for the theoretical defect of basic Mean shift. The blocking issue and tracking of fast moving target are discussed in this paper and applied in the independent visual robotic fish tracking. Here are our conceptions: Firstly, the initial position of Mean shift is provided with Kalman filter in every frame, then Mean shift algorithm is used to track the position of target. Secondly, the calculation of the Kalman residuals is applied to turn on and turn off the Kalman filter under a large proportion of blocking situation, at the same time, the linear prediction of the target position is replaced by Kalman's function test. Finally, we conduct a real-time tracking experiment on the independent visual robotic fish tracking a moving ball, it is proved that the algorithm can achieve the tracking of fast moving target and robust against the barrier as well.

Introduction

Target tracking is a hot research field in computer vision. In video surveillance, target recognition, human-machine interface and other application areas, moving objects often need to be tracked effectively in complex environment. Mean Shift was first proposed by Fukunaga and it means the offset mean vector originally [1]. With the developing of Mean Shift Theory [2], the meaning of Mean Shift also has changed. The basic Mean Shift algorithm has been promoted by Cheng [3] in two aspects. Firstly, he defines a kernel function, the result of offset mean vector which is influenced by offset is variable with the distance between the sample and offset points. Secondly, he also sets a weight coefficient which leads to different importance of sample and it has expanded the scope of Mean Shift [4]. The basic thought of Mean Shift algorithm has been analyzed in this paper and the algorithm has been improved as well. Here are the following adjustments: we set the position of target center which is changed by the time as the observation value of Kalman filter and take it as the initial position of Mean Shift using Kalman filter to predict the possible position of motion target in the current frame, mean shift will find final position of target in current frame within the neighbor position. We judge whether a larger proportion of barrier happened through the analysis of Kalman filter residuals and give the corresponding processing methods to solve the issue. Then we have done a lot of tracking experiments in video. It turns out the improvement of Mean-Shift is highly real-time and can track the target in video effectively.

Basic theory of Mean shift

In the space of D dimensions, S represents a group of data. $K_h(x)$ is a multidimensional kernel function and x represents a point in this space. Then, we can get an estimate density function, density estimation is a process which we can estimate right probability density distribution from an unknown set of probability density distribution observations.

The equation of the kernel density estimation is as follows:

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i) \quad (1)$$

In this equation, $\{x_i\}_{i=1,2,\dots,n}$ is given data points in d-dimensional Euclidean space, K_h is a function which used in the threshold of density estimation. For K_h there are many different kernel functions we can choose, multidimensional kernel function is one of the most optimizing kernel functions on the basis of the least mean square.

The kernel function K_h must meet the requirement that K_h is a symmetrical probability density function : $\int_0^\infty K_h(x)dx = 1$.

The equation is as follows:

$$K_h = |H|^{-\frac{1}{2}} K(|H|^{-\frac{1}{2}} x) \quad (2)$$

In this equation is called a bandwidth matrix $d \times d$. Then we can estimate the probability density functions of the given data based on the equation (1).

Implementation improved Mean-Shift algorithm and Kalman filter[6]

Mean shift algorithm is a semi automatic tracking method which select motion target by manually assuring search window at the beginning of tracking frame. The histogram distribution of the search window is calculated by calculating the kernel function, the same method is used to calculate the histogram distribution of the current frame, with the maximum principle of two distributions' similarity, make the search window increase in the real position of moving target along the increasing maximum direction of the density.

Here is the procedure:

Target model of initial frame

Set the unit as an interval of color value, divide the feature space which is picked from pixel color values into several feature values, then in the initial frame, including target search window, the probability of u feature value is:

$$\hat{q}_u = C \sum_{i=1}^n k\left(\left\|\frac{x_0 - x_i}{h}\right\|^2\right) \delta[b(x_i - u)] \quad (3)$$

x_0 is central pixel coordinate of search window, x_i is coordinate of pixel i; $k(\|x\|^2)$ is kernel function, h is bandwidth of kernel function, generally equate half of window's width; function b and δ is to judge whether color value in x_i belong to feature value; C is a standardized constant coefficient, make all of feature values' addition of probability is 1.

Model of current frame

Similar to above, calculate feature value u's probability of search window in current frame

$$\hat{p}_u(y_0) = C_h \sum_{i=1}^{n_h} k\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right) \delta[b(x_i) - u] \quad (4)$$

y_0 is center pixel coordinate of search window in current frame, x_i is coordinate of pixel i; C_h is the same as C in (1).

Similarity function

Similarity function describe the similarity between initial frame target model and current frame target model, define as follow:

$$\hat{\rho}(y) = \rho(\hat{p}(y), \hat{q}) = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (5)$$

Mean shift vector

In order to maximize $\hat{\rho}(y)$, in the current frame, take search window position of former frame as search window position of current frame, set window center as y_0 and find the best part target

position y_1 in the neighbor of y_0 . Make (5) do Taylor expansion in $\hat{p}(y_0)$, the function can be written as follow:

$$\rho(\hat{p}(y), \hat{q}) = \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u} + \frac{C_h}{2} \sum_{i=1}^{n_h} w_i k\left(\left\|\frac{y-x_i}{h}\right\|^2\right) \quad (6)$$

$$w_i = \sum_{u=1}^m \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \delta[b(x_i - u)] \quad (7)$$

Calculate the max value through similarity function and get the Mean shift vector:

$$m_{h,G}(y) = y_1 - y_0 = \left[\frac{\sum_{i=1}^{n_h} x_i w_i g\left(\left\|\frac{\hat{y}_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{\hat{y}_0 - x_i}{h}\right\|^2\right)} \right] - y_0 \quad (8)$$

Make Mean shift algorithm do iteration repeatedly, and get the best target position y of current frame.

Kalman filter estimated initial position

We use central position of motion target as the observation value of the Kalman filter. First, we use Kalman filter to predict the target position in the next frame and take it as initial position of motion target in Mean shift, then Mean shift will find the best target position within the neighborhood and take it as observation value of Kalman filter to do calculation in the next frame.

Modeling of Kalman filter

Assume that the center of motion target is working in a straight line motion with random acceleration in X, Y axis. Acceleration a is a random variable which obey Guass distribution. Define signal vector $X(k) = [x(k) \ y(k) \ x'(k) \ y'(k)]^T$, $x(k)$ is the position component of X axis, $y(k)$ is the position component of Y axis, $x'(k)$ is the velocity of X axis, $y'(k)$ is the velocity of Y axis. Observation vector $Y(k) = [x_c(k) \ y_c(k)]^T$, $x_c(k)$ is observation value of target center in X axis, $y_c(k)$ is observation value of target center in Y axis.

Thus, Kalman filter consist of two models:

$$\text{Signal model:} \quad X(k) = A(k-1)X(k-1) + B(k)W(k) \quad (9)$$

$$\text{Observation model:} \quad Y(k) = C(k)X(k) + V(k) \quad (10)$$

$$\begin{bmatrix} x(k) \\ y(k) \\ x'(k) \\ y'(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x(k-1) \\ y(k-1) \\ x'(k-1) \\ y'(k-1) \end{bmatrix} + \begin{bmatrix} \frac{t^2}{2} \\ \frac{t^2}{2} \\ t \\ t \end{bmatrix} w(k) \quad (11)$$

$$\begin{bmatrix} x_c(k) \\ y_c(k) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x(k) \\ y(k) \\ x'(k) \\ y'(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} v(k) \quad (12)$$

Define initial value: $t=1$, $X(-1) = [x_s \ y_s \ 0 \ 0]^T$. x_s, y_s is central coordinate of initial frame target.

Treatment of block issue

Residual stands for the addition of offset square which is pointed to the estimated and the observation values of current frame. Residual of Kalman filter can be defined as follows:

$$r(k) = \sqrt{(x_c(k) - \hat{x}(k))^2 + (y_c(k) - \hat{y}(k))^2} \quad (13)$$

$\hat{x}(k)$ is the target vector of the estimated values and $\hat{x}_i(k)$ means the target vector of the observation values. Then, after some frames, Kalman filter will find its estimated value of target position and its observation value similar through the following mean shift algorithm, $r(k)$ will always maintain its value in a normal range. If being blocked in a large proportion, mean shift algorithm will begin with the initial value provided by Kalman filter, finding the “final” position, which means the position will be far from the initial one. If the residual is far from its normal value, we can conclude that the target may be blocked in a large proportion. Under this situation, we need to set a threshold value to help us judge that whether the Kalman filter has been worked in a wrong mode. According to the experiment, we set 10 as the threshold value. If $r(k) > \lambda$, then stop the Kalman filter and predict the initial value of the target in the next frame using the central position of target in the former frames, then use mean shift algorithm to find the real position of the current frame, meanwhile, calculate the value of $r(k)$ and compare with threshold value. If $r(k) < \lambda$, then restart the Kalman filter in the next frame; if $r(k) > \lambda$, keep on adjusting the value of $r(k)$ using the central position of target in the former frames until it meets the requirement. The flow chart is showed as Figure 1

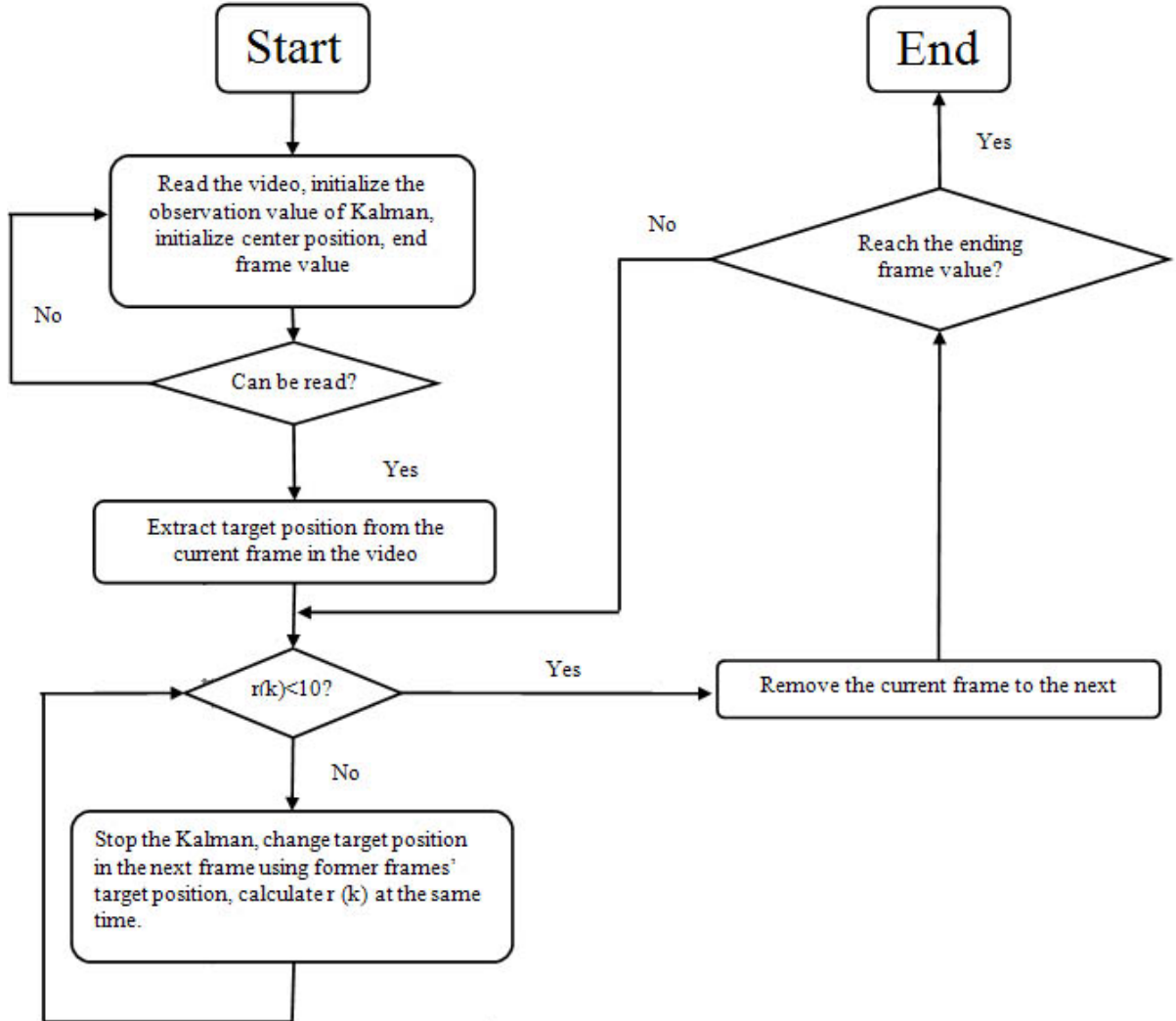


Fig.1 Flow chart of target tracking on Mean Shift algorithm

Target tracking can be automatically completed once we manually confirm the first frame.

Experiment results and analysis

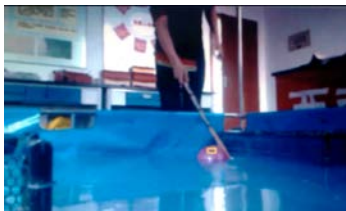
The original algorithm based on Mean Shift algorithm has been improved, to track fast moving targets more accurately, using the comparison of two domain values When selects observation value of Kalman filter. To verify the tracking principle of Mean Shift algorithm which has been improved, we'll make the experiment to track movement of robotic fish in the video files. The Mean-Shift algorithm of the first frame manually determines the undetermined tracking target in experiments which is developed under VC6.0++.

Fig.2 make use of Mean shift processing video sequence “Infinitt”, the total length of sequence is 950.

(1) The tracking results of traditional Mean-Shift algorithm

According to the figure as follow, Mean shift can track target well, as show in Fig.2, frame 8.

But Mean shift fail to track target owing to its much too fast motion, as show in Fig.2, frame 15,42.



current frame 8



current frame 15

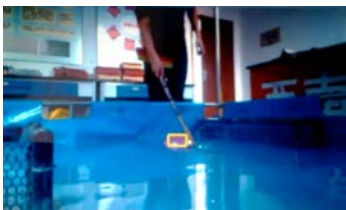


current frame 42

Fig.2 The tracking results of traditional Mean-Shift algorithm

(2) The tracking results of Mean-Shift algorithm which is improved

As you can see from figure as follow, tracking accuracy has increased significantly. “Infinitt” sequence processed by improved algorithm in this paper show that even in a high velocity, we can still track target well, as show in Fig.3



current frame 121



current frame 156



current frame 173

Fig.3 The tracking results of Mean-Shift algorithm which is improved

(3) Treatment of block issue:

We want to track the ball behind the bigger one, as show in square of frame 468. From frame 866 On, we can see the bigger ball start to block the back one, as show in frame 866,924, square of frame 866 track the ball behind the bigger one and track it again until frame 924, though there is a large portion of barrier, square of frame 924 did not recognize it wrong, which mean our algorithm succeed.



current frame 468



current frame 866



current frame 924

Fig.4 The instance about the block issue

Conclusion

Using Kalman filter combined Mean shift algorithm to solve the issue that basic Mean shift is unable to track fast moving object. Kalman filter predict possible target position of current frame according to the target center position of former frames, Mean shift find the final target position in the neighborhood. New algorithm enrich availability of known information. Judge whether a large proportion barrier happened through calculation of Kalman filter residuals, use liner prediction of former frames' target position to replace Kalman's function; meanwhile, judge the period that Kalman start to work through calculation of residuals. The experiments show that the algorithm above can achieve tracking of fast moving target and have a good robustness against barrier as well.

Acknowledgements

This work was financially supported by Innovative Research Team of the department of Sichuan Province (15TD0050) and the Fundamental Research Funds for Central University, Southwest University for Nationalities (13NZYQN04).

References

- [1] Philippe Van Kerm. Adaptive kernel density estimation[C]//9th UK States Users Meeting, Royal Statistical Society.Lodon, 2003:19-20.
- [2] Li Shijin , Ruan Xiaozhe , Feng Jun , etc. A Novel Approach to Video Transition Detection Based on Machine Learning [J]. Journal of Guang xi Normal University: Natural Science Edition, 2008, 26(1): 228-231.
- [3] Cheng Yizong . Mean Shift, mode seeking, and clustering[J]. IEEE Trans Pattern Analysis and Machine Intelligence, 1995,17(8): 790-799.
- [4] Wang Liang , Hu Weiming, Tan Tieniu. A Survey of Visual Analysis of Human Motion [J]. CHINESE J. COMPUTERS, 2002,25(3): 225-237.
- [5] Silverman B W. Density Estimation for Statistics and Data Analysis[M].London: Chapman and Hall,1986.
- [6] D. COMANICIU V. RAMESH P. MEER. Kernel-based object tracking [J]. IEEE Transactions. on Pattern Analysis and Machine Intelligence 2003 25(5) 564-577.