

Research of Moving Target Tracking Algorithm for Video

Huiying Dong*, Shengfu Chen

Shenyang Ligong University, Shenyang 110159, China

* Corresponding author's Email: huiyingdong@163.com

Abstract: In this paper, the correlation tracking algorithm based on the function of minimum average absolute difference is analyzed. In addition, a kind of correlation tracking algorithm based on particle filtering is used to avoid the target shaded by object. And last, Kalman filtering approach is used as the motion model of the multi-target tracking, and the recursive filtering method is adopted to calculate and predict the location of moving targets. The tracking results show that correlation tracking algorithm based on particle filtering can overcome the obstacle occlusion problem perfectly, which can also improve the robustness of target tracking. Besides, target tracking algorithm based on Kalman filtering works well for multi-target tracking, which has great practical significance on tracking technology for moving targets.

Keywords: particle filtering; the function of the minimum average absolute difference; Kalman filtering; target tracking; correlation tracking

1. Introduction

Tracking of moving target is to search the position of the most similar candidate target in the video by using the valid feature of the moving target and appropriate matching algorithm. In other words, it is the target location in each of the sequence images. In practical, moving target tracking is an important part between moving target detection and recognition and target analysis and understanding.

Current moving target tracking algorithms are known as the largest cross-correlation function, the minimum mean square error function, the minimum average absolute difference function, maximum matching pixel statistics, and so on. These tracking algorithms are essentially a “peak” search, that is, the “most similar” target parameter searched is the best parameter. The information exchanging between the frames only contains the “most similar” point’s information. And the other information which correlation value is less than the peak point is not used in the calculation and prediction reasonably between the frame [1]. To solve

these problems, this paper uses a kind of correlation tracking algorithm based on particle filtering to overcome the obstacle occlusion problem when it occurs by using the correlation tracking algorithm based on the function of the minimum average absolute difference. And finally target tracking algorithm based on Kalman filtering [2] algorithm extends the tracking target from single target to multi-targets.

2. The Correlation Tracking Algorithm Based on the Function of the Minimum Average Absolute Difference

The basic principle of correlation tracking is “template matching”, that is, taking certain gray values of pixels in one image, or a set of some features as a “template”, then by searching the same part or similar part of the corresponding to match the template in order to solve target motion parameters. The Correlation tracking algorithm is shown as the follows:

(1) Description of target template

The “template” used to describe the target is a set

of some features of target. Supposing the extracted template is directly formed by the gray pixels of some area of the images. This kind of template is the most intuitive because it uses the basic features of the image.

(2) Calculation of correlation value

Using the gray template of the target matches some areas of the image by using the correlation value calculation. And the zoom of correlation value indicates the higher correlation. The correlation function is chosen as the minimum average absolute difference (MAD):

$$MAD(i, j) = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N |T(m, n) - F(m+i, n+j)| \quad (1)$$

$$i = 1, 2, \dots, M, \quad j = 1, 2, \dots, N$$

(3) Motion model of the target

The ultimate purpose of target tracking is to get the description of the target state, that is, to get the motion parameters of the target. The dimensions of the motion parameters depend on how to establish the motion model of the target. Motion models include the translation model, similar model, affine model, etc. The changes of the coordinates before moving and after moving is presented by the translation model:

$$y = x + T \quad (2)$$

where x, y is two-dimensional coordinate vector which represents the translation in horizontal and vertical direction. T is a two-dimensional column vector showing the changes of translation.

(4) The search of matching value

Assume the original image size is $M \times N$, TX and TY represent the x and y direction coordinates of the target center. Then, (TX, TY) constitutes one $M \times N$ two-dimensional motion parameter space $E(i, j)$, $i = 1, 2, \dots, M$, $j = 1, 2, \dots, N$. This space can describe all possible states of the target. The tracking process is to search the best matching point in the parameter space, that is, to get the best estimate value of the motion parameters.

(5) Tracking results

By using above algorithm to track the car in video, the gray template is extracted from the target in original image. The function of minimum average absolute difference is used as matching criterion. Translation model is used as motion model. The cross symbol is used as the center of the tracking window. The tracking results of some frames in the video are shown as in Figure 1, the central pixels coordinates shown as

in Figure 3, and the values of the central coordinates shown as table 1. From the figure, we can know the tracking valid when there is no obstacle shading. To overcome the shading we use the correlation tracking algorithm based on the particle filtering.

3. Correlation Tracking Algorithm Based on Particle Filtering

3.1 Correlation tracking principle based on particle filtering

In particle filtering technique, recursive Bayesian filtering is realized by non-parametric Monte Carlo simulation method. This technique can be applied to any system with nonlinear and the non-Gaussian noise case. Its accuracy can approach optimal estimation [3]. The particle filtering theory describes a robust target tracking frame. In the tracking method based on particle filtering principle, gray template in traditional relevant tracking (template matching) method is applied as the description of the target. The sum of the particle weights is used to express the estimate value of the target parameters. The weight values of the particle are proportional to the correlation value. The specific steps are described as follows:

(1) Initialization: Assume at the initial moment K_0 , one image template $f(a, b)$ with size of $M \times N$ ($a=1, 2, \dots, M$; $b=1, 2, \dots, N$) is extracted and the initial motion parameters are $T^{init} = (TX^{init}, TY^{init})$, where TX^{init} and TY^{init} stands for the coordinate value in x and y directions respectively. The particle number is assigned as N_s , and the weight value W_t as 1, and there are two parameters for each particle:

$$T^i = (TX^i, TY^i), \quad i = 1, 2, \dots, N_s \quad (3)$$

The initial value of particle parameter is taken as:

$$TX^i = TX^{init} + b_1 \zeta, \quad TY^i = TY^{init} + b_2 \zeta \quad (4)$$

where ζ is a random number within $[-1, 1]$, b_1, b_2 are constants.

(2) State transformation in the later time K_t ($t > 0$), state transfer equation of the system is used to predict the state of each particles. Taking the first order ARP equation:

$$x_t = Ax_{t-1} + Bw_{t-1} \quad (5)$$

For particle N^i , there are:

$$\begin{aligned} TX_t^i &= A_1 TX_{t-1}^i + B_1 w_{t-1}, \\ TY_t^i &= A_2 TY_{t-1}^i + B_2 w_{t-1}, \\ &i = 1, 2, \dots, N_s \end{aligned} \quad (6)$$

where, x_t is the state of the target at t moment, A_1, A_2, B_1, B_2 are constants, w_{t-1} is the normalized noise value, usually the w_{t-1} is one random value within $[-1, 1]$.

(3) The particles can be observed after the system has observed the transmission of each particle. The minimum average absolute difference function is taken as a tool to measure the degree of similarity, that is, a similar value MAD^i ($i = 1 \dots N_s$) can be calculated for each particle.

Define the observation probability density function as:

$$p(z_k|x_k^i) = \exp\left\{-\frac{1}{2\sigma^2}MAD^i\right\} \quad (7)$$

where σ is constant. Then Gaussian modulation is worked down for the correlation value. The weights of the particles can be calculated through the recursive calculation by the following formula:

$$w_k^i = w_{k-1}^i p(z_k|x_k^i)$$

(4) Calculation of the posterior probability at k_t moment, that is the expected target parameters during target tracking (TX_t^{opt}, TY_t^{opt}), can be expressed by the sum of each weighted particle:

$$TX_t^{opt} = \sum_{i=1}^{N_s} \omega_t^i TX_t^i, \quad TY_t^{opt} = \sum_{i=1}^{N_s} \omega_t^i TY_t^i \quad (8)$$

3.2 Algorithm flow

First, the particles N_s is set and the motion model is selected. The choice of the particle number is related to the actual tracking requirements. Generally, the more particles, the more stable and higher accuracy of the tracking. But at the same time, the greater computation is required. In actual tracking situation, compromise selection or dynamic adjustment can be adopted. In the occasion that only needs target location, the motion model can be selected as translation model and system output volume is target location value. In the occasion that needs to calculate the target pose, motion model can be chosen as affine model with six parameters, which include horizontal and vertical displacement and scale, cutting scale and rotation angle. Once the motion model is selected, particle will be consistent with it, and the parameters have the same dimension.

Then the target area is determined and target template is established. At this time, the initial parameters of the template such as target location, size and angle are obtained. According to the initial target parameters, the parameters of each particle are initialized

and particle weights are set to 1 (that is, all particles equally important).

At the third frame and later stage, the algorithm switches to the iterative process of particle filter. In each frame, system state transition, systematic observation and weights calculations are executed for each particle. Then all the particles are weighted in order to output the estimation value of the target state. Finally, particle re-sampling process will be done and the next iterative process will begin.

3.3 Tracking results and analysis

Correlation tracking algorithm based on particle filtering is used to track the moving car on the road. The video is the same with that in Figure 1. In the experiment, “+” is used as the center of the tracking window, the number of particle is $N_s = 15$, translational motion model is selected. In initial center point of the target, X position coordinate is 85, and Y is 57. Figure 2 shows the 8 frames that are extracted from the tracking results of the correlation tracking algorithm based on particle filtering. From the Figure 2, we can see that the shading missed tracking by obstacle in the algorithm in section 2 is avoided successfully.

Table 1: The pixel coordinates of the correlation tracking center point based on the function of the minimum average absolute difference

Frame number	1	4	7	10	13
Coordinates	(58, 83)	(64, 93)	(69, 102)	(75, 112)	(76, 112)
Frame number	16	19	22	25	28
Coordinates	(76, 113)	(76, 113)	(76, 113)	(76, 113)	(76, 112)

Table 2: The pixel coordinates of the correlation tracking center point based on particle filtering

Frame number	1	4	7	10	13
Coordinates	(57,85)	(65,100)	(72,115)	(80,130)	(88,145)
Frame number	16	19	22	25	28
Coordinates	(95, 160)	(103, 175)	(111, 190)	(119, 205)	(126, 220)

Because the speed of the moving car changes a little within a small time interval, the car can be considered doing uniform linear motion. In each of two adjacent images, pixel distance of the target movement is approximately equal. In the process of tracking moving target, tracking windows should be able to lock target in real-time and do uniform linear motion with the target. As it can be seen from Figure 3 and Table 1, in

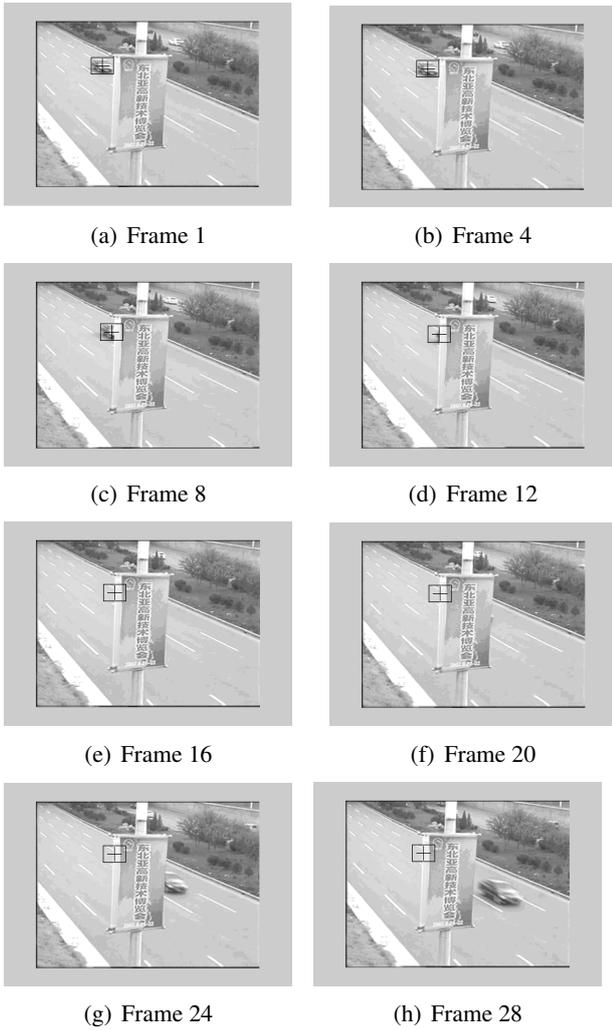


Figure 1: Tracking results of the correlation tracking algorithm based on the function of the minimum average absolute difference

the case of none obstacle blocking, the moving of the tracking windows based on the function of the minimum average absolute difference consists with the law of car uniform linear movement. But when the car blocked by obstacle billboard, tracking windows can't move with the target, which expresses the tracking is missing. As can be seen from Figure 4 and Table 2, whether the obstacles blocking or not, the moving of the tracking windows based on particle filtering does uniform linear motion with the target, and locks the target in real-time.

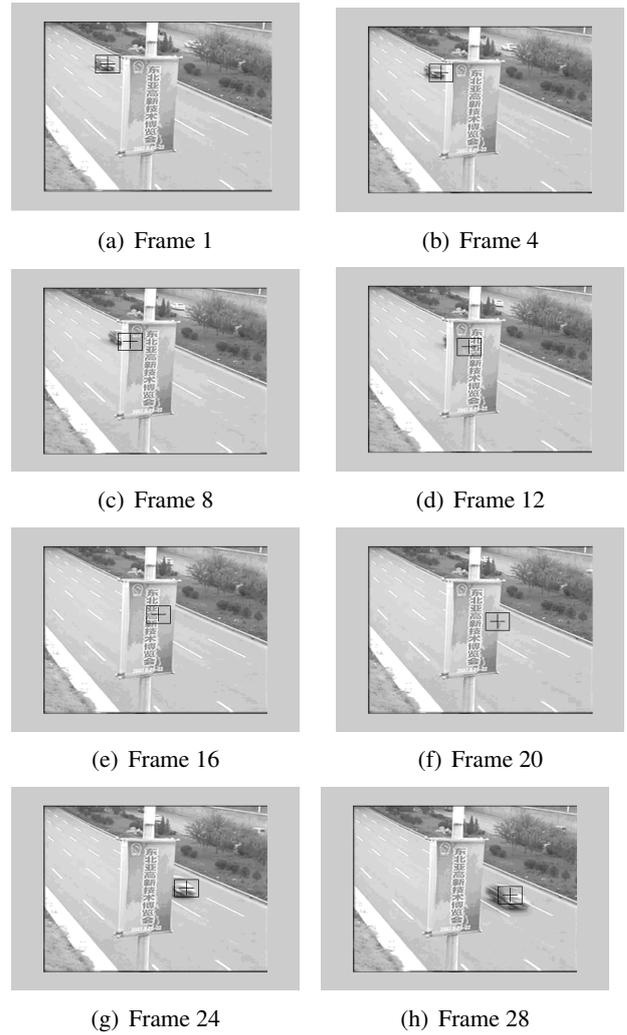


Figure 2: Tracking results of the correlation tracking algorithm based on particle filtering

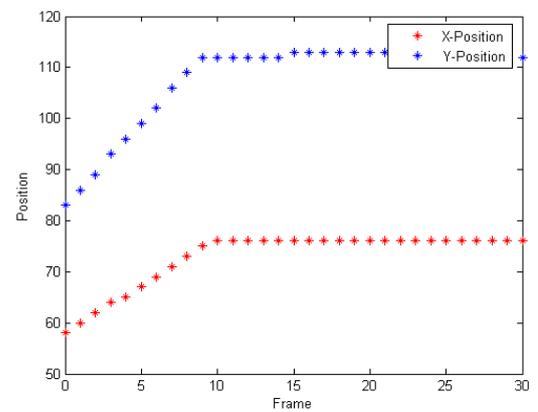


Figure 3: The center pixel coordinates of the correlation tracking windows based on the function of the minimum average absolute difference

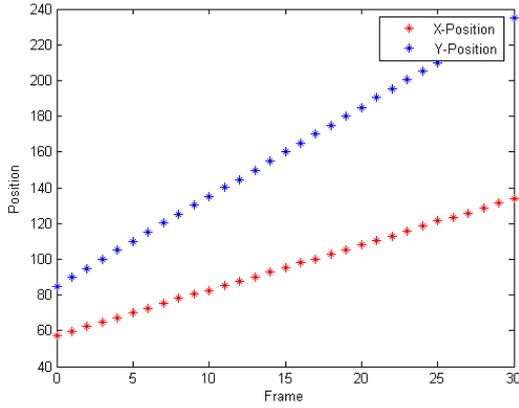


Figure 4: The center pixel coordinates of the correlation tracking windows based on particle filtering

4. Moving Targets Tracking Algorithm Based on Kalman Filtering

4.1 Kalman filter principle

Kalman filter is an efficient re-recursive filter [4],[10]. It can estimate dynamic system state from a series of incomplete included noise measurement. It can observe at any point as a starting point. Using recursive filtering method has small amount of calculation and computing features in real-time. Kalman filter uses the state equation and observation equation to describe a dynamic system. In the Kalman filter expression, we hope that the optimal linear estimate of the state vector $x(k)$ is a function of observation vector $z(1), z(2), \dots, z(k)$, where the optimal refers to the least mean-square error criterion. In addition compared with least mean square error function, the Kalman filter is optimal in a wider category of performance function.

Kalman filter reduces error covariance matrix of each time point K to the least when tracking, it consists of two steps:

1. Prediction procedures: including state prediction and error covariance prediction.
2. Modification procedures: including Kalman gain calculation and next state and error covariance modification.

The implementation process of Kalman specifically as follows:

- (1)State prediction equation:

$$\hat{X}_{k/k-1} = A\hat{X}_{k-1} + Bu_k \quad (9)$$

- (2)Covariance prediction equation:

$$\hat{P}_{k/k-1} = A\hat{P}_{k-1}A^T + Q \quad (10)$$

- (3)Kalman gain equation:

$$K_k = \hat{P}_{k/k-1}H^T(H\hat{P}_{k/k-1}H^T + R)^{-1} \quad (11)$$

- (4)State modification equation: (3)Kalman gain equation:

$$\hat{X}_k = \hat{X}_{k/k-1} + K_k(Z_k - H\hat{X}_{k/k-1}) \quad (12)$$

- (5)Covariance modification equation: (3)Kalman gain equation:

$$P_k = (\hat{P}_{k/k-1} - K_kH)\hat{P}_{k/k-1} \quad (13)$$

Kalman filter estimates motion state using the feedback control system, filter estimating the state of a certain time, and obtains the prediction value of this state. In other words, the Kalman filter equation is divided into two parts: prediction and correction. Prediction equation is responsible for using current state and error covariance estimate to obtain a priori estimate for the next time state; correction equation is responsible for feedback part and the new observation and a priori estimate are being considered together to obtain a posteriori estimate.

4.2 Target tracking based on Kalman filtering

Kalman filter predicts the location of moving targets based on the motion model of the target. When the state of the moving targets after the prediction matches the current foreground targets, image centroid and window size are used to locate precisely position of targets. First, calculating Eigen values, simultaneously using the Kalman filter to build motion model which is used for predicting the movement of target in the next frame having been extracted on the current to narrow the scope of target-matching and accelerate the speed of target matching. Second, doing the target-matching in the next frame within the specified to establish target association. Last, updating motion model and forming the target motion tracking chain to get moving target trajectory [6].

4.2.1 Calculating Kalman filter tracking Eigen values

Because time interval between two adjacent frames is small in the image sequence, the speed and direction of an object movement will not change evidently in the mutating within such a short time. So, in such a relatively small time interval, the case can be considered as moving target in the adjacent two images, of whose center of mass and tracking window size and

position has changed little, which means the continuity characteristics of the movement of the moving target. Therefore, choosing the centroid tracking and follow-up window size as the characteristic value to track targets. To track the window which traps the target image tightly and whose size is slightly larger than the target image, the target is not tracked out of the window background and noise disturbances. At the same time, due to the large amount of image data depending on the venue, in order to shorten the time, tracking window can be drawn to help reduce the size of the processing image, only processing an interesting partial image in real-time. At the same time it can also set up a few tracking window for a number of targets appearing in field of view, respectively to track moving targets. After marking good in all track windows, centroid is sought respectively for the objectives of the window. Each window centroid coordinates calculated by the following formula:

$$x = \frac{m_{10}}{m_{00}} = \frac{\sum_i \sum_j i f(i, j)}{\sum_i \sum_j f(i, j)} \quad (14)$$

$$y = \frac{m_{01}}{m_{00}} = \frac{\sum_i \sum_j j f(i, j)}{\sum_i \sum_j f(i, j)} \quad (15)$$

The $f(i, j)$ is gray value of the target image in the track windows with index (i, j) . The centroid coordinates of each tracking window is an important state parameters during tracking process of moving target, in the subsequent tracking process, the centroid coordinates becomes one of the key state variables.

4.2.2 Motion estimation model

Suppose the state vector s_{k+1} at $k+1$ moment in the model are composed of the vector s_k transfer function of time point k and noise. While the observation vector is determined by the vector s_{k+1} observation function of time point $k+1$ and noise.

Equation of state is as follows

$$s_{k+1} = A s_k + w_k \quad (16)$$

Observation equation is as follows

$$z_{k+1} = C s_{k+1} + v_{k+1} \quad (17)$$

where w_k, v_{k+1} are noise. Introduction of noise, firstly, rely on experience to determine, secondly, is obtained by studying statistics method. Suppose dynamic noise and observation noise are normal white noise with zero mean value.

s_k is the state vector, an eight-dimensional vector

$$s_k = \begin{bmatrix} x_k \\ y_k \\ \dot{x}_k \\ \dot{y}_k \\ L_{xk} \\ L_{yk} \\ \dot{L}_{xk} \\ \dot{L}_{yk} \end{bmatrix} \quad (18)$$

where x_k, y_k are respectively target cancroids coordinates;

\dot{x}_k, \dot{y}_k stands for the unit displacement of the cancroids coordinates in the x, y direction respectively;

L_{xk}, L_{yk} represent the width of tracking window in the x, y direction respectively;

$\dot{L}_{xk}, \dot{L}_{yk}$ represent the unit displacement of tracking window width in the x, y direction respectively.

z_{k+1} is the observation vector, constituted by a four-dimensional vector

$$z_{k+1} = \begin{bmatrix} x_{k+1} \\ y_{k+1} \\ L_{xk+1} \\ L_{yk+1} \end{bmatrix} \quad (19)$$

At the sampling time $t = 0.04$ seconds, time interval is very short, so it can be thought approximately that the moving target moves at a constant speed. Moreover tracking window size changes little, then the state transition matrix A is:

$$A = \begin{bmatrix} 1 & 0 & 0.04 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0.04 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0.04 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0.04 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Observation matrix C is:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

After defining the state equation and observation equation of motion model, Kalman filter is used to estimate the position of target cancroids and window size. Then cancroids can be searched in the next frame within a small area.

4.2.3 Target feature matching and model updating

For a group of motion image sequences, $F = \{f_1, f_2, \dots, f_n\}$, creating the same Cartesian coordinate system for each of these images, in the first k frame image f_k , m of targets are recorded as $R = \{r_1, r_2, \dots, r_m\}$, centroid coordinates, window width of the first i targets are recorded respectively as $x_k^i, y_k^i, L_{xy}^i, L_{yk}^i$.

First defining the mass center of the first i th objective of the first k frame and the centroid distance function of the first j objective of the first $k+1$ frame

$$D(i, j) = \frac{|c_k^i c_{k+1}^j|}{\max|c_k^i c_{k+1}^*|} \quad (20)$$

where:

$$|c_k^i c_{k+1}^j| = \sqrt{(x_k^i - x_{k+1}^j)^2 + (y_k^i - y_{k+1}^j)^2}, D(i, j) \leq 1$$

Subsequently defining area difference function, which compares windows area with the first i objective of the first k frame and the first j objective of the first $k+1$ frame

$$A(i, j) = \frac{|a_k^i - a_{k+1}^j|}{\max|a_k^i - a_{k+1}^*|} \quad (21)$$

where:

$$|a_k^i - a_{k+1}^j| = L_{xk}^i \times L_{yk}^i - L_{xk+1}^j \times L_{yk+1}^j, A(i, j) \leq 1$$

Defining similarity function

$$\Delta(i, j) = \gamma D(i, j) + \zeta A(i, j) \quad (22)$$

where γ, ζ are weighting values, and meeting the condition $\gamma > \zeta, \gamma + \zeta = 1, \Delta(i, j) \leq 1$.

The smaller $D(i, j)$, indicate the closer targets, while the smaller $A(i, j)$ indicate the more similar target shape, the smaller $\Delta(i, j)$, indicate the most probability similar of the two objectives. As a result, set the threshold T_Δ of similar function to judge that the targets are or not the same goal. If all the goals on a frame and the smallest value of the result of estimating and matching on the fore frame are over a threshold, this shows that the frame is without the same follow-up targets; If it is below this threshold, this shows that the target of the smallest value is subsequent for the follow-up frame target. While the minimum value of similar function is found, the follow-up of the same target has been found, that says, the first j objective of the first $k+1$ frame can be seen as the follow-up of the first objective of the first i frame, and the two are the same target. At this time, the characteristic value of the first j objective of the first $k+1$ frame is used as input of motion model estimating the next frame, and so on, to complete the model updates.

4.3 Experimental results and Analysis

Target tracking algorithm based on Kalman filter is adopted to track the two moving cars on the road. Experiments on

Table 3: Consecutive frames moving target centroid coordinates and window size value

Frame number	Target 1			Target 2		
	Centroid x coordinates	Centroid y coordinates	Window area	Centroid x coordinates	Centroid y coordinates	Window area
42	146	159	34×46	186	141	29×43
45	159	165	33×48	198	147	30×44
46	163	168	33×47	201	149	30×44
49	173	173	35×51	214	155	32×47
50	177	175	36×51	219	157	32×46
51	181	177	37×52	224	159	33×48
52	186	178	38×53	228	160	35×48
53	189	181	39×54	232	163	35×50

the 320×240 pixels video image sequence processing, target moving along the x direction of the right bottom of the road, take $\gamma = 0.7, \zeta = 0.3, T_\Delta = 0.6$. Figure 5 shows the 12 images extracted from the tracking experimental results respectively, Table 3 shows the centroid position of moving target and tracking window size.

It can be seen from these simulation results that the algorithm is improved for moving target tracking. In the process of target motion, each object tracking window size can do the state transfer based on the actual state, automatically adjust it's size in the process of the next frame tracking. The exist of normal white noise of mean zero has been taken into account in the process of motion state transition. Because in the process of tracking moving target, the Eigen value of the target centriod coordinates and tracking window size can be conducted the state transition in the process of the next frame tracking, such tracking algorithm can also effectively overcome the obstacle blocks and such questions in the process of tracking.

5. Conclusion

Correlation tracking algorithm based on the function of the minimum average absolute difference (MAD) uses the "peak" tracking method, which abandons all the location information less than the peak correlation value, but the algorithm is failed when faced with obstacle, so the method is not robust. Correlation tracking algorithm based on particle filtering, not only inherits the visual and practical characteristics of correlation tracking, but also reflects the "multi-peak" tracking excellence of particle filtering, which improves the robustness of tracking. Target tracking based on Kalman filtering extends the tracking target from single target to multi-targets, which captures multiple moving targets tracking effectively.

References

- [1] Chongyou Zhang, Huiying Dong, Libao Lan. "Research of Correlation Tracking Algorithm Based on

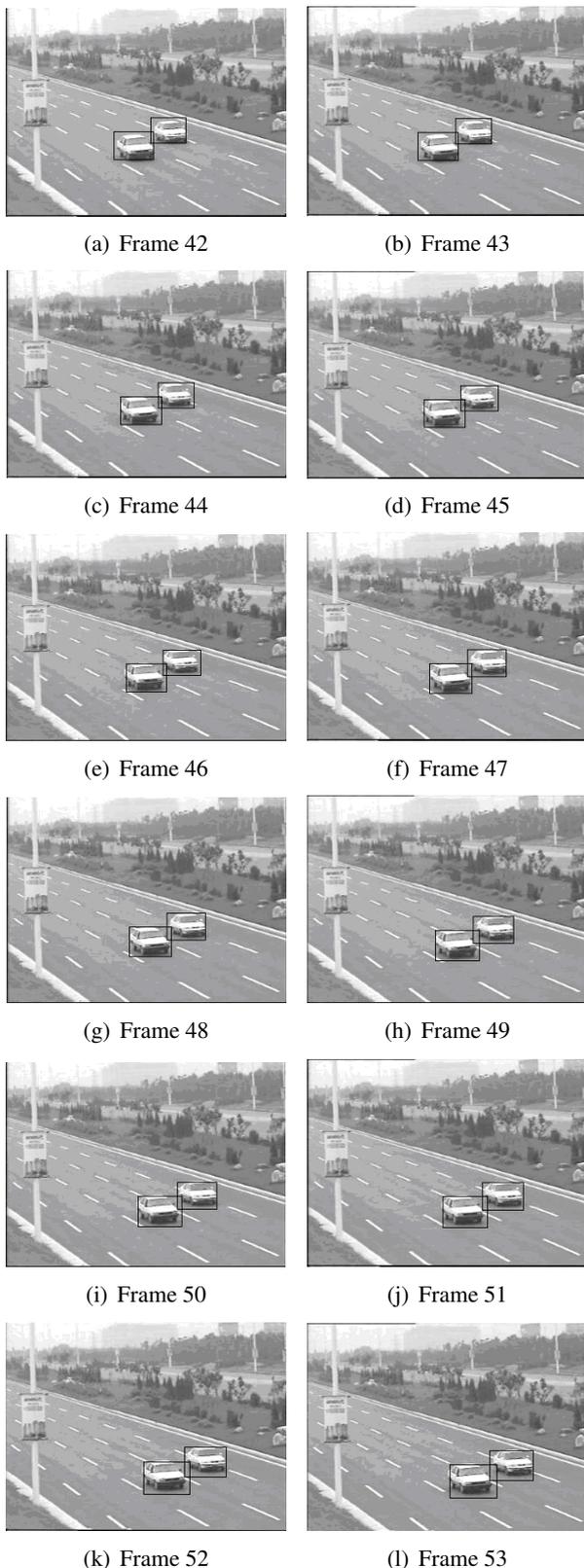


Figure 5: Target tracking result map based on Kalman filtering

- Particle Filtering”. *Transactions of Shenyang Ligong University*, 2008, 27(1): 6-9.
- [2] Huiying Dong, Shengfu Chen. “Research of Moving Targets Tracking Algorithm Based on Kalman Filtering”. *The 3rd International Conference on Intelligent Networks and Intelligent Systems*. 2010.
 - [3] Liu Bo, Li Xianli. “The Research of Visual Information Detection and Tracking Technology”[J]. *Computer Engineering*, 2000,(6): 86-89.
 - [4] Alan J. Lipton, Hironobu fujiyoshi, Raju s.Patil. “Moving target classification and tracking from real-time video”[J]. *IEEE Workshop on Applications of Computer Vision (WACV)*, Princeton NJ, 1998. 8-14.
 - [5] Wang Chunbo. *The Research of Visual Information Detection and Tracking Technology in Traffic Monitoring*[M].2000.
 - [6] Doucet A, Gordon N, Krishnamurthy V. “Particle Filters for State Estimation of Jump Markov Linear Systems”[J]. *IEEE Transactions on Signal Processing*, 2001,49: 613-624.
 - [7] Kalman, R,E and R.S.Bucy. “New results in filtering and prediction theory”, *Transactions of the ASME Journal of Basic Engineering*, Vol.83,1961.95.
 - [8] Meditch,J, S. *Stochastic Optimal Linear Estimation and Control*, McGraw-Hi-ll, 1969: 188-196.
 - [9] S. Julier, J. Uhlmann, and H.F Durrant-Whyte. “A new method for the non-linear transformation of means and covariances in filters and estimators”. *IEEE Transactions on Automatic Control*, 2000. 43(3):477-482.
 - [10] Ribaric S.G, Adrinek, Segvic S.. “Real-time active visual tracking system”. *IEEE Proceedings Mediterranean Electrotechnical Conference*. 2004, 5,vol.1:231-234.