

# Research on Mean-Shift Target Tracking Based on Image HOG Feature

Yang Xinfeng<sup>\*,1</sup> and Tian Di<sup>2</sup>

<sup>1</sup>School of Computer and Information Engineering, Nanyang Institute of Technology, Henan, Nanyang, 473004, P.R. China; <sup>2</sup>Department of Information Technology, Henan Institute of Education, Henan Zhengzhou, 450046, P.R. China

**Abstract:** In this paper, mean shift algorithm based on HOG feature has been used to track the moving object. A feature hierarchical model has been constructed, which is coarser in the entire moving object layer. It is partly based on model components featuring accurate tracking of the moving object in different layers. The moving object is tracked using nuclear tracking with two layers which is more accurate than the moving object tracking with one layer information. Partially, it is based on the whole occlusion target location that can better be forecasted. The algorithm has a good tracking performance after the moving object appears again.

**Keywords:** Hierarchical modeling, Histogram of oriented gradient, Mean shift, Moving object, Detection, Feature.

## 1. INTRODUCTION

Intelligent transportation plays an increasingly important role in modern life, and is closely related to people's daily life. From the 60s of the last century, with the rapid growth of the world economy, the rapid development of new industries such as computer technology and electronics industry, the development of these industries put forward new requirements for modern transportation. Fast, convenient and safe transportation can better serve the development of the world economy, relying solely on the increasing contradiction between urban land and other issues which are increasingly prominent.

A lot of research for target detection and tracking has been carried out in the field of intelligent transportation. Intelligent traffic can be traced back to the 1960s [1]. The 60's research focused on the radar and its related image processing, through a simple digital image, to better meet the military requirements [2]. In the 1970s, Marr, Barrow and Tenenbaum pioneered the visual computing framework, laying the foundation for the development of intelligent transportation [3, 4]. Until the 1980s, intelligent traffic has been developed by leaps and bounds. Active vision [5, 6], multi-dimensional stereo image acquisition technology [7-10], multi-sensor technology and other related technologies formed a complete foundation framework, which greatly promoted the development of intelligent traffic.

From the current research and development of intelligent transportation field, there are still many problems to be studied in depth. These problems greatly influence the development of intelligent transportation from the laboratory to practical application. Among them, "moving target detection and tracking" is one of the most challenging and difficult issues which has a wider application prospect in spatial information, electronic reconnaissance, industry and

modern defense. This issue is based on the study of intelligent traffic vehicle detection and pedestrian detection and tracking, the most fundamental is how to get higher level of characterization or semantic understanding from the bottom of the sequence image data itself. This research is closely related to this basic problem, which involves the bottom image processing technology and the description of the characteristics of the middle vision and other issues, therefore, the detection and tracking of moving objects not only have a wide range of prospects in practical applications, but also have great research value in the theoretical research.

The information of histogram of gradient image direction provides the characteristics of the object and its local characteristics. The overall structure of the object is the embodiment of the integrity of the object, but each object has its own local characteristics, the description of these local features can help correct and predict the position in the target tracking, and have good prediction effect on the occlusion, which has significance in targeting tracking.

## 2. HOG FEATURES TARGET MODELING

### 2.1. Target Description

Whether it is a pedestrian or a vehicle, the final test result is a region, which contains all the information of the target. Through the different analysis of the information within the region, we can complete the description of the target. In this paper, the feature description of the histogram of gradient direction can be expressed, which can express the target feature, and has a good effect on noise suppression and illumination. For the target area of the current frame, the feature of histogram of gradient direction can be written as Eq (1):

$$q_u = \{q_{ui}\}, \sum_{i=1}^n q_{ui} = 1 \quad (1)$$

Candidate target areas can also be similarly expressed as Eq (2):

$$p_u = \{p_{ui}\}, \sum_{i=1}^n p_{ui} = 1 \quad (2)$$

\*Address correspondence to this author at the School of Computer and Information Engineering, Nanyang Institute of Technology, Nanyang, 473004, P.R. China; Tel: 13803770071; E-mail: [ywind2005@163.com](mailto:ywind2005@163.com)

In the specific algorithm, matching may be determined by comparing the similarity of both the targets.

Because the histogram as the feature matching of the object is more common and the effect is ideal, the flexible part target modeling is adopted on the basis of the HOG feature selection. The hierarchical processing of the feature description is carried out, and the target is modeled, the detection results of the target of the overall model may be a generally rectangular area or the boundary of the object. The current target is used as the first layer of the target pyramid to construct the finer pyramid of the next level. The local characteristic of the target is described in the pyramid of the layer. If the local maximum or minimum point appears, the interval as a local feature area of the target is used for local tracking. As can be seen from the algorithm, each tracking needs to calculate one overall tracking and a number of local trackings, so its efficiency will be reduced. In order to improve the efficiency, in the absence of occlusion and loss of the target, only the overall characteristics of the target and the secondary flexible component characteristics are calculated, so that the efficiency can be improved.

### 2.2. Target Model

The target modeling is based on the need of tracking to build the model of the target location. Considering the practical application, the feature combination is set as  $\{x_i\}, i=1, \dots, n$ , and the center position of the target as  $x_0$ . Considering the flexible deformable model, some of the characteristics of the target will be described, which mainly described each partial region with feature, such as a pedestrian's head, arms and so on. Since its initial goal is obtained by target detection, the target feature is described by a rectangular box, which can be described as Eq. (3):

$$q_u = C \sum_{i=1}^n k(x_i^2) \delta(b(x_i) - u) \quad (3)$$

Among them,  $\delta$  is the Dirichlet function, which is only 1 at 0, and at 0, it is at the rest, the  $k(x)$  function is a profile function for kernel function. Considering that the characteristics in the actual tracking process are easily affected by color, light and so on, it is necessary to increase the accuracy of the target by giving a relatively large weight to the center of the distance. The final description is Eq. (4):

$$q_u = C \sum_{i=1}^n k\left(\left\|\frac{x_i - x}{h}\right\|^2\right) \delta(b(x_i) - u) \quad (4)$$

Where  $C$  is the normalized parameter, its value is Eq. (5):

$$C = \frac{1}{\sum_{i=1}^n k\left(\left\|\frac{x_i - x}{h}\right\|^2\right)} \quad (5)$$

### 2.3. Candidate Model

Candidate model is consistent with the target model, and its characteristic set is expressed as  $\{x_i\}, i=1, \dots, n_k$ , the

center coordinate point is  $y$ . In its practical application, object deformation and other problems lead to changes in the target, so there are differences between the number of its characteristics and the original target. According to the characteristics, its characteristics can be written as Eq. (6):

$$p_u = C \sum_{i=1}^{n_k} k\left(\left\|\frac{x_i - x}{h}\right\|^2\right) \delta(b(x_i) - u) \quad (6)$$

where  $C$  is also the normalized parameter, its value is Eq. (7):

$$C = \frac{1}{\sum_{i=1}^{n_k} k\left(\left\|\frac{x_i - x}{h}\right\|^2\right)} \quad (7)$$

## 3. MEAN SHIFT ALGORITHM

The mean drift algorithm is a non-kernel density estimation algorithm. The mean drift vector is closely related to the kernel density estimation. The direction of the mean drift vector is the probability density gradient direction. The problem of the kernel density estimation is obtained by the conversion of target tracking problem, and finally the kernel density estimation is completed by using the relevant theory.

### 3.1. Mean Shift Vector

Assuming that the set of samples is  $n$  points  $\{x_i\}$  of the  $d$ -dimensional space, where  $i$  is an integer of 1 to  $n$ , the mean drift vector of point  $x$  can be defined as [11] Eq. (8):

$$M_k(x) = \frac{1}{k} \sum_{x \in S_k} (x_i - x) \quad (8)$$

where  $S_k$  is a high-dimensional space, that is Eq. (9):

$$S_k = \{y | (y - x)^T (y - x) \leq r^2\} \quad (9)$$

Where  $k$  is the number of points falling into the  $S_k$  area in the  $n$  sample sets,  $x_i - x$  is the offset of the sample point relative to the center point, and the formula (1) is the geometric mean of the offset.

In order to better carry out the probability of density estimation, the introduction of kernel function can help to solve the situation that the high dimensional data estimation is consistent with the mean point.

Let  $X$  be the  $d$ -dimensional Euclidean space,  $x$  is a point of the space,  $x$  is satisfied  $\|x\|^2 = x^T x$ . Defining a function to satisfy  $K(x) = k(\|x\|^2)$ ,  $k$  is the profile function for the  $K$  function,  $k$  satisfies the properties of nonnegative, non-increasing, segmented continuous and isotropic. Such  $K$  function is called a kernel function. By defining the kernel function, it is consistent with the fact that it is important to strengthen the point near the reference point. Therefore, it can be understood that the sample is only related to the distance of the reference point.

Commonly used kernel density estimation function has unit uniform kernel function, unit Gauss kernel function and Epanechnikov kernel function, the expression is as follows:

Unit uniform kernel function is Eq. (10):

$$F(x) = \begin{cases} 1 & \|x\| < 1 \\ 0 & \text{other} \end{cases} \quad (10)$$

Unit Gaussian kernel function is Eq. (11):

$$N(x) = \exp(-\|x\|^2) \quad (11)$$

Epanechnikov kernel function is Eq. (12):

$$K_E(x) = \begin{cases} c(1-\|x\|^2) & \|x\| \leq 1 \\ 0 & \text{other} \end{cases} \quad (12)$$

From the practical application, it is necessary to consider not only the distance of the sample from the center of the mean value, but also the importance degree, so the weight coefficient  $w(x_i)$  is introduced, and then the formula (1) is extended Eq. (13):

$$M_k = \frac{\sum_{i=1}^n G_H(x_i - x) w(x_i) (x_i - x)}{\sum_{i=1}^n G_H(x_i - x) w(x_i)} \quad (13)$$

In the formula, the  $w(x_i)$  weight is a finite value and is not negative.

$G_H(x_i - x) = |H|^{-1/2} G(|H|^{-1/2}(x_i - x))$ , where  $G(x)$  is a unit kernel function,  $H$  is the bandwidth matrix; it is a positive definite square matrix. In practical application,  $H$  is set to a diagonal matrix as  $H = \text{diag}[h_1, h_2, \dots, h_d]$ , When the diagonal element is a uniform value  $h$ , the formula (13) is abbreviated as Eq. (14):

$$M_k = \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i) (x_i - x)}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i)} \quad (14)$$

The above formula is the form of the general kernel density estimation. This method is also used in this paper. When the kernel function takes the uniform kernel function and the weight is taken as 1, then the equation is degenerated into formula (1).

### 3.2. Probability Density Gradient

The probability density gradient direction is theoretically consistent with the mean drift direction, and  $M_k$  is the normalized probability density gradient. The selection of the sample point is obtained from the probability of density function  $f(x)$  by sampling. It can be seen that the sampling

point falls more in the region with large probability density, so the mean drift based on the sampling point should be the direction of the probability density gradient, which is the probability density gradient direction.

Setting the sampling point  $x_i$  of the probability density function  $f(x)$  in the  $d$ -dimensional space to satisfy the following estimate Eq. (15):

$$\hat{f}(x) = \frac{\sum_{i=1}^n K\left(\frac{x_i - x}{h}\right) w(x_i)}{h^d \sum_{i=1}^n w(x_i)} \quad (15)$$

Where  $K(x)$  is the aforementioned kernel function and the kernel function is defined as follows:

Profile function  $K(x)$  for kernel function is Eq. (16):

$$K(x) = k(\|x\|^2) \quad (16)$$

$g(x)$  is the negative derivative function,  $g(x) = -k'(x)$ , considering its probability density gradient, the gradient is estimated as follows Eq. (17):

$$\hat{\nabla} f(x) = \frac{2 \sum_{i=1}^n (x_i - x) k'\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{h^{d+2} \sum_{i=1}^n w(x_i)} \quad (17)$$

Let  $G(x)$  is the kernel function for  $g(x)$ , and then formula (17) can be written as Eq. (18):

$$\begin{aligned} \hat{\nabla} f(x) &= \frac{2 \sum_{i=1}^n (x_i - x) k'\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{h^{d+2} \sum_{i=1}^n w(x_i)} \\ &= \frac{2}{h^2} \left[ \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i)}{h^d \sum_{i=1}^n w(x_i)} \right] \left[ \frac{\sum_{i=1}^n (x_i - x) G\left(\left\|\frac{x_i - x}{h}\right\|^2\right) w(x_i)}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i)} \right] \end{aligned} \quad (18)$$

The value in the second bracket is a mean drift, the value in the first bracket is the probability density estimation of  $G(x)$  as the kernel function; the above formula can be written as Eq. (19):

$$\hat{\nabla} f(x) = \frac{2}{h^2} f_G(x) M_h \quad (19)$$

From the above formula, it can be seen that the mean drift is Eq. (20):

$$M_h = \frac{h^2 \nabla \hat{f}(x)}{2 f_G(x)} \quad (20)$$

This formula shows that the mean drift is consistent with the probability density gradient, and the mean drift obtained

by  $G(x)$  at  $x$  is proportional to the probability density gradient. In this way, obtaining the mean drift can be estimated as probability density to replace.

### 3.3. Tracking algorithm

As can be seen from the previous description, the mean drift is the amount we need to find. The solution of the amount is the iterative process of estimating the probability density. Formula (7) is modified as follows Eq. (21):

$$M_k = \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i) x_i}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i)} - x \quad (21)$$

Since the first term in the above formula is close to the point  $x$ , it can be set to  $m_k$  as in Eq. (22):

$$m_k = \frac{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i) x_i}{\sum_{i=1}^n G\left(\frac{x_i - x}{h}\right) w(x_i)} \quad (22)$$

It is calculated as follows:

Setting the initial point  $x$  and the kernel function  $G(x)$ , we can perform the following iteration to find  $m_k$ .

Step 1: Calculate  $m_k$  according to formula (22)

Step 2: Let  $x = m_k$

Step 3: Using the error limits set  $\varepsilon$ , if  $\|M_k - x\| < \varepsilon$ , stop, otherwise go to step 1 to continue.

$m_k = x + M_k$  can be calculated by iteration. From the calculation estimation algorithm it can be seen that the step size is variable. By constantly adjusting the step size, when the center near the point is reached, the step size becomes shorter, being close to the center point. So the method is estimated to be valid.

## 4. MEAN-SHIFT TARGET TRACKING BASED ON GRADIENT DIRECTION HISTOGRAM

### 4.1. Similarity Measure

Similarity measure is the realization of the detection of target tracking. Only through the similarity measure, we can better determine the location of the current target, and the candidate model provides the reference information of the current frame. There are many similarity measurement parameters, such as the Pasteur coefficient [12], histogram intersection [13], Kullback divergence parameter, and Paste coefficient is more effective in the mean shift algorithm than using other parameters [14].

Pasteur coefficient is defined as in Eq. (23):

$$\rho(y) = \rho(p(y), q) = \sum_{u=1}^m \sqrt{\hat{p}(y) \hat{q}} \quad (23)$$

The value range is  $[0, 1]$ , the closer the value is to 1, the higher the similarity of the two models. Ideally, the

probability density distribution of the two is exactly the same.

The geometric meaning is that  $p'$  is an  $m$ -dimensional vector and  $q'$  is another  $m$ -dimensional vector. Setting in Eq. (24):

$$\begin{aligned} p' &= (\sqrt{p_1}, \sqrt{p_2}, \dots, \sqrt{p_m}) \\ q' &= (\sqrt{q_1}, \sqrt{q_2}, \dots, \sqrt{q_m}) \end{aligned} \quad (24)$$

Pasteur coefficient is the angular cosine of two vectors.

According to the knowledge of the trigonometric function we can know that:

$$\cos(\theta) = \frac{p' \cdot q'}{\|p'\| \|q'\|} = \frac{\sqrt{p_1 q_1} + \sqrt{p_2 q_2} + \dots + \sqrt{p_m q_m}}{\sqrt{p_1 + p_2 + \dots + p_m} \sqrt{q_1 + q_2 + \dots + q_m}} \quad (25)$$

From its geometric meaning, the smaller the angle, the more similar the two vectors are. From the perspective of tracking target, it is the real goal of the maximum probability of the region.

### 4.2. Target Location Determination

After the current frame detects the target, the target detection is not performed in the subsequent frame, and the current target is tracked. In the determination of the target position, through the study of the location of the previous frame, the exact location is determined to achieve the purpose of tracking. The Taylor expansion of the Pasteur coefficient is shown in Eq. (26-29):

$$\rho(p(y), q) \approx \frac{1}{2} \sum_{i=1}^m \sqrt{p(y_0) q_u} + \frac{1}{2} \sum_{u=1}^m p_u(y) \sqrt{\frac{q_u}{p_u(y_0)}} \quad (26)$$

Substituting the kernel function into it:

$$\rho(p(y), q) \approx \frac{1}{2} \sum_{i=1}^m \sqrt{p(y_0) q_u} + \frac{1}{2} C_k \sum_{i=1}^m w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \quad (27)$$

From this formula, can get the following conclusions:

$$m(y) = y_1 - y_0 = \frac{\sum_{i=1}^m x_i w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^m w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)} - y_0 \quad (28)$$

The new coordinate  $y_1$  is expressed as :

$$y_1 = \frac{\sum_{i=1}^m x_i w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^m w_i g\left(\left\|\frac{y_0 - x_i}{h}\right\|^2\right)} \quad (29)$$

Then, at this time, the latest coordinate position is obtained by the similarity to achieve the purpose of tracking, Using mean drift, through a certain number of iterative operations, we can obtain a good effect of the actual coordinates.

### 4.3. Occlusion treatment

The problem of occlusion must be addressed in the tracking. Because the characteristic of flexible part structure used in this paper is local and integral, so when the object is occluded, the occlusion performed is better. When the targets are occluded, if targets are of the same type, the characteristics of the common module in a certain sense should not affect the determination of the overall character of the target. When the characteristics are inconsistent, according to the flexible structural features, the component structure can be used to restore the overall characteristics.

When the moving object reaches the back of the nonmoving object in a current scene, the object is not imaged in the video. Occlusion can be divided into partial occlusion and complete occlusion.

**Partial occlusion:** Partial region is shaded, while there are still some visible portions. At this time, the object is divided into a plurality of discrete regions, and the portion of the object is linked, which should be considered the same object. In order to determine the center of the object, this description of the identification method is used. By examining the characteristics of flexible components, we can complete the overall target recognition. For only partially visible cases, it can be determined by partial and integral contact, and when the target and the background are occluded, the situation will become more complicated, at this time, we can use occlusion between the targets to analyze and achieve the purpose of tracking.

**Full occlusion:** When the moving target moves to the background, the target disappears completely in the image and disappears after the detection frame disappears for a period of time. In this case, it is necessary to predict the position of the object. By predicting the position, the object can be rediscovered after the object occlusion to achieve the purpose of object tracking. A simple metric with reference to the target centroid will change the size of the target and cause a failure in detection. In the follow-up tracking based on the two-tier tracking system for estimation, and the overall target centroid control and flexible component feature extraction control, the current target location can provide a better prediction.

### 4.4. Algorithm Design

According to the results provided by the detection algorithm, we can initialize the basic information of the tracking algorithm. According to the test results, the central position of the rectangular box is  $y_0$ , and the HOG feature is used as the feature of the object. Therefore, it is better to choose the cloud model with HOG feature from the tracking point of view. While the target model can use the test model, in the tracking, HOG model is established with two layers. In the second layer model, the fine HOG model is constructed by sampling the first layer image, and then the flexible part model is used that dynamically constructs the sub-features of the target. In this paper, the number of dynamic models is 3-7, and the same numbers of centers  $y_i$  were constructed. In the tracking algorithm, these sub-features need to be tracked at the same time, and the tracking

algorithm is consistent with the overall target tracking algorithm.

Then its tracking characteristics are set as Eq. (30):

$$\{p(y_0), p(y_1), p(y_2), \dots, p(y_n)\} \quad (30)$$

Where  $y_i$  is the position of each feature level.

Then following the characteristics of the mean shift algorithm for tracking:

In the first step, the histogram and the Pasteur coefficient of the center coordinate point are calculated, and the histogram and Bhattacharyya coefficient of each flexible part are constructed, with the center coordinate vector and the Paste coefficient vector.

The second step seeks the weight matrix vector.

The third step is to calculate the weight vector according to the condition, and update the center coordinate matrix and the Bhattacharyya coefficient vector.

In the fourth step, the Bhattacharyya coefficient matrix is compared, and if the weighted average of the current Bhattacharyya coefficient vector is smaller than the original, the average value is taken to recalculate the Bhattacharyya coefficient vector.

The fifth step: the process ends if the error requirements are met, otherwise the second step is followed for recalculation. The specific process is shown in Fig. (1).

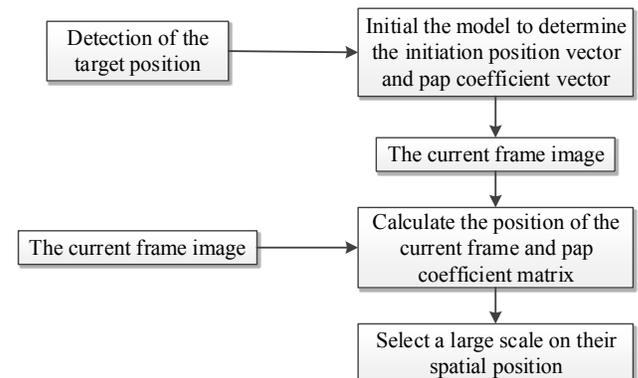


Fig. (1). Algorithm tracking process.

## 5. EXPERIMENTS AND CONCLUSION

This paper uses the highway video image and the monitoring data of one River Bridge and the indoor pedestrian data to carry out relevant experiment. The experiment is carried out from two aspects, namely, the overall goal as a tracking target and the overall goal and local information combined as a tracking target. It can be seen from the experimental results that the tracking method combined with the whole and local in this paper can track the target more accurately. The evaluation is measured using the coverage ratio, which is defined as follows in Eq. (31):

$$\Omega = 2 \frac{|R_o \cap R_i|}{|R_o| + |R_i|} \quad (31)$$

Where  $R_o$  represents the size of the actual object,  $R_r$  represents the size of the object. Because the tracking video does not provide the actual size of the target, the approximate target detection algorithm results are used as the real target in the image of the marked image, therefore, it can approximate the tracking effect to express the advantages and disadvantages of the tracking algorithm. In this paper, two algorithms are used to track, one is based on the whole target feature modeling, and another algorithm uses a combination of flexible component models and overall modeling. The experimental results are shown in Figs. (2-5) and Table 1.

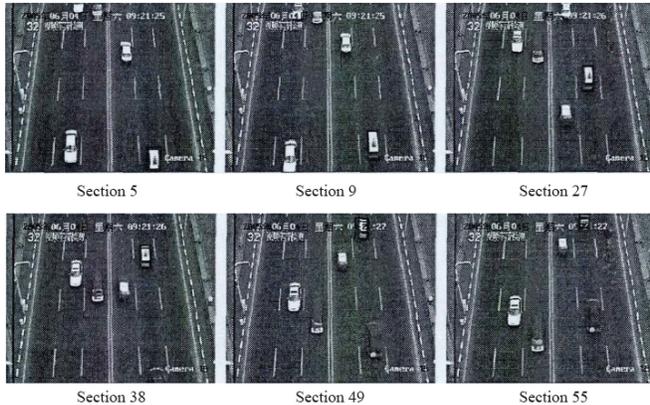


Fig. (2). One River Bridge video surveillance vehicle tracking.

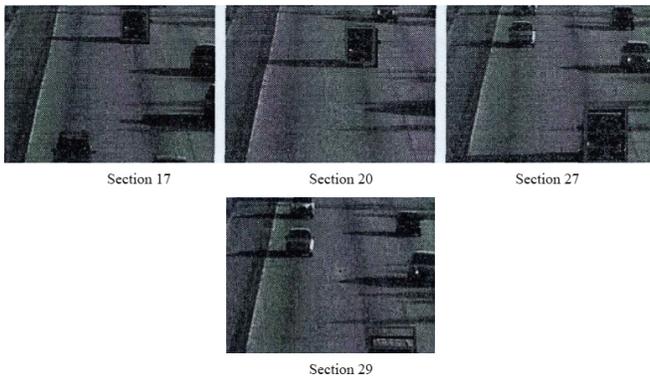


Fig. (3). Highway I Vehicle Tracking.

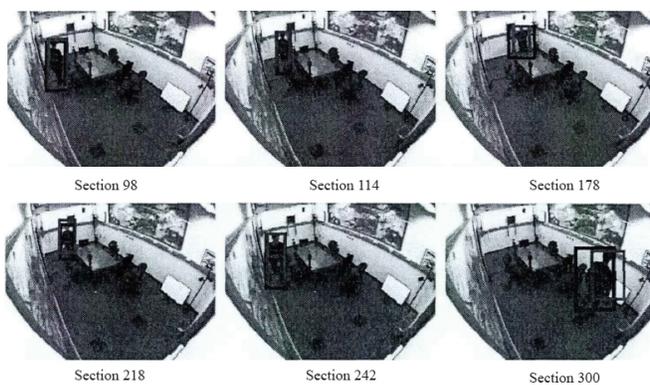


Fig. (4). Indoor pedestrian tracking.

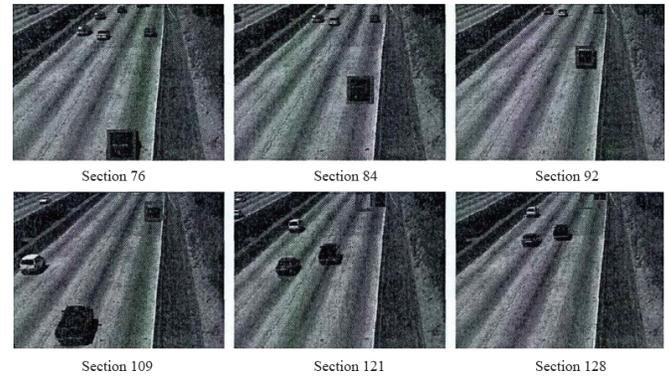


Fig. (5). Highway II track results.

Table 1. Different video segments target tracking algorithm results compare.

	Overall tracking	Overall combined with the flexible member
One River Bridge	0.903	0.913
Highway I	0.832	0.912
Highway II	0.503	0.883
Pedestrian	0.504	0.935

From the perspective of tracking, the shadow effect of (Fig. 2) video is relatively large. It can be seen from the experimental results that under the influence of shadow, this algorithm can still perform better target tracking, in relatively crowded environment, by the impact of the front and the next vehicle, causing the block between the targets in tracking but it can still track the target location. In pedestrian tracking, due to the impact of the background, the target is blocked by the background, based on considering only the mean drift algorithm, the target position deviation will appear to be relatively large after the target is blocked for some time. Experiments show that the two-layer tracking algorithm can still track the target better in the case of occlusion.

From the viewpoint of tracking time efficiency, the algorithm in this paper needs to initialize modeling when the target is detected. The difference in efficiency between the algorithm and the basic mean shift algorithm is mainly time-consuming in model construction. In practical application, such tracking speed can meet the requirements. So the algorithm is real-time and effective.

**CONCLUSION**

According to the above conclusions, the hierarchical target feature modeling combined with flexible components, using multi-layer mean drift algorithm can be better to complete the tracking. One layer of the target feature pyramid model is the target layer, and another layer is the description of the local feature layer; the combination of the two can describe the target feature more comprehensively.

Therefore, the method combined with flexible component tracking and integral tracking is used to overcome the problem of occlusion between the targets and background occlusion. Because the mean shift algorithm cannot track the goal of fast movement, therefore, for the goal of speed, the Kalman filter and particle swarm optimization algorithm can solve the problem on the basis of the same model. Because of the limited time, it cannot be studied deeply in this algorithm.

### CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

### ACKNOWLEDGEMENTS

This work was financially supported by the key scientific research project of Henan Province (15A520090) and the key science and technology project of Henan Province (122102210563 and 132102210215).

### REFERENCES

- [1] Sayrol E, Gasull A. Motion estimation using higher order statistics. *IEEE transaction on image processing*, Vol.5, No.6, 1996. pp.1077-1083.
- [2] Dubuisson M.P. and Jain A. K.. Contour extraction of moving objects in complex outdoor scenes. *International Journal of Computer Vision* Vol.14, No.1, 1995. pp.83-105.
- [3] Huang W.Y Liu Z.J., et.al, The Precise Recognition of Moving Object in Complex Background. *IEEE Third International Conference on Natural Computation*, 2007, pp.246-252.
- [4] Barrow ST. and Tenenbaum JM. *Computational vision. Processing of IEEE*, Vo1.69 No.5, 1981, pp.572-597.
- [5] Aloimonos J., Purosis and qualitative active vision the 10<sup>th</sup> international conference of Pattern Recognition, 1990 pp.246-360.
- [6] Aloimonos J., Weiss I. and Bandopadhyay A.. Active vision. *International Journal of Computer Vision*.1988.pp.333-356.
- [7] Bajcsy R. and Campos M.. Active and Exploratory Perception, *CVGIP:Image Understanding*,56(1),1992, pp.31-40.
- [8] Armen F. Aggarwai J.K. Model-based object recognition in dense-range images a review. *ACM Computing Surveys*.Vol.25, No.1, 1993, pp.5-43.
- [9] Chin R.T, and Dyer C.R. Model-based object recognition in dense-ranges a review. *ACM Computing Survey*.vol 18(1).1986. pp.67-108.
- [10] Pan L.F. Behavioral Knowledge in sensor/data fusion system. *Journal of Robotic Systems*. Vol. 7 No.3. 1990, pp.295-372.
- [11] Comaniciu D., Meer P., Mean-shift Analysis and Applications, In: *Processing of the IEEE international conference on Computer vision*.1999. pp.1197-1203.
- [12] Swain M.J. and Ballard D.H..Color Indexing, *International Journal of Computer vision*. vol.7, No.1.1991.pp.11-32.
- [13] Maggio E. and Cavallaro A..Hybrid particle Filter and Mean Shift Tracker with Adaptive Transition Model. *IEEE international conference on Acoustics, Speech and signal processing*, 2005, pp.221-224.
- [14] Comaniciu D., Meer P.. Mean Shift Analysis and Applications.h processing of the IEEE International conference on Computer Vision.1999.pp.1197-1203.

Received: May 26, 2015

Revised: July 14, 2015

Accepted: August 10, 2015

© Xinfeng and Di; Licensee *Bentham Open*.

This is an open access article licensed under the terms of the (<https://creativecommons.org/licenses/by/4.0/legalcode>), which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.