Acta Technica 62 No. 2A/2017, 39–50 © 2017 Institute of Thermomechanics CAS, v.v.i.

Study on video vehicle detection based on improved mean-shift algorithm

Geng Liwei[1], Dong Tingting[1]

Abstract. In allusion to defects of traditional Mean-Shift algorithm such as missed detection and object overlap, this paper introduces blob detection into Mean-Shift algorithm, puts forward vehicle video tracking algorithm of an improved Mean-Shift algorithm and achieves accurate and fast positioning of single object and multiple objects. The experimental result shows that Mean-Shift algorithm based on blob can not only achieve multi-object tracking, but also have more accurate tracking effect. Its tracking effect is obviously better than that of traditional Mean-Shift algorithm.

Key words. Mean-Shift algorithm, similarity, blob detection, Kalman filtering.

1. Introduction

Object tracking is a key and hotspot issue in the field of computer vision [1]. How to achieve object tracking in video sequence accurately is a difficult issue in the current research. As an important algorithm in object tracking research, Mean-Shift algorithm was first put forward by Fukunaga and Hostetler in 1975. Yizong Cheng [2] conducted extending study on it based on Mean-Shift theory and applied it to the field of computer vision. It started to arouse wide concern of domestic and foreign research scholars. Dorin Comaniciu and Peter Meer [3] boiled down nonrigid object tracking issue to optimization problem of Mean-Shift algorithm approximately. Mean-Shift algorithm has so far applied to object tracking formally. Nummiaro [4] combined Mean-Shift algorithm and particle filter and put forward multi-object tracking algorithm of an improved Mean-Shift algorithm. In allusion to infrared small object tracking problem, Chen Jiang et al. [5] combined Mean-Shift algorithm and cascade gray space and achieved good effect through application. Tracking and positioning accuracy has been greatly improved compared to Mean-Shift algorithm.

As kernel bandwidth of search window of traditional Mean-Shift algorithm remains unchanged in the tracking process, good tracking effect can be achieved for objects with minor scale change. However, stable tracking cannot be achieved and

[1]Agricultural University of Hebei, Baoding, Hebei Province, 071000, China

objects are often lost or tracking objects are missed for detection in the case of significant change of spatial scale of objects. To describe the change of object scale with time, Zhu Shengli et al. [6] used affine transformation for description, obtained transformation parameter with maximum correlation coefficient of match window in Mean-Shift and two continuous frames and calculated the change of bandwidth and the initial position of Mean-Shift. However, affine model is easily influenced by illumination intensity. Moreover, it is required to conduct correlation operation of point N in angle and scale. The calculation is complicated and real-time tracking cannot be achieved.

In allusion to defects of Mean-Shift algorithm, this paper introduces blob detection into Mean-Shift algorithm based on literature [6], puts forward vehicle video tracking algorithm of an improved Mean-Shift algorithm and achieves accurate and fast positioning of single object and multiple objects.

2. Methodology

Mean-Shift video tracking algorithm is a semi-automatic algorithm [7–8]. In starting frame, moving object is generally determined through manual or other recognition algorithms and meanwhile histogram distribution of target window is calculated. Then, histogram distribution of the corresponding window is calculated in the Nthframe of video sequence, which is used as candidated object model. The similarity of the two models is calculated and maximum similarity of the two models is regarded as the principle for moving tracking window so as to determine the position of object accurately.

2.1. Modeling

Mean-Shift algorithm generally uses probability distribution of weighting of characteristic value to determine initial object model [9]. In starting frame of video sequence, probability distribution of characteristic value at each point within target window area is counted with kernel function weighting method. To improve the efficiency of operation, characteristic space of object is transformed into multiple characteristic values, in which value q_u of the uth feature is [10–11]:

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

In formula (1), $x_0,\ x_i$, respectively, correspond to the coordinate vectors in target window center and any point in the window, $k||x||^2$ refers to the kernel function, h refers to the bandwidth of object model, $b(x_i)$ refers to the characteristic value at point x_i, C denotes the standard coefficient of characterization and the sum of probabilities of all characteristic values is 1. Candidated object model can be obtained according to feature distribution of search window located in y in frame N of computing center of video sequence [12].

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

2.2. Similarity comparison

So called $\hat{p}_u(y)$ similarity with $\hat{q}_u$ is described through the Bhattacharyya coefficient. Quantity $\hat{\rho}(\mathbf{y}) = \sum_{u=1}^{m} \sqrt{\hat{p}_u(\mathbf{y})\hat{q}_u}$ is used for measurement and maximum similarity is regarded as the selection principle. Quantity $\hat{\rho}(\mathbf{y})$ is subject to Taylor expansion at $\hat{p}_u(\mathbf{y}_0)$ of the last frame

$$\hat{\rho}(\mathbf{y}) \approx \frac{1}{2}\sum_{u=1}^{m} \sqrt{\hat{p}_u(\mathbf{y}_0)\hat{q}_u} + \frac{1}{2}\sum_{u=1}^{m} \hat{p}_u(\mathbf{y})\sqrt{\frac{\hat{q}_u}{\hat{p}_u(\mathbf{y}_0)}} . \tag{3}$$

The first item on the right-hand side of formula (3) is a constant value. Therefore, the value of the second item must be maximum in order to achieve the maximum of $\hat{\rho}(\mathbf{y})$.

2.3. Object positioning

After further concretization, formula (3) can be transcripted in the following way,

$$\hat{\rho}(\mathbf{y}) \approx \frac{1}{2}\sum_{u=1}^{m} \sqrt{\hat{p}_u(\mathbf{y}_0)\hat{q}_u} + \frac{C_h}{2}\sum_{i=1}^{n_h} \left\{ k\left(\left\|\frac{\mathbf{y} - \mathbf{x}_i}{h}\right\|^2\right) \sum_{u=1}^{m} \delta\left[b\left(\mathbf{x}_i\right) - u\right] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\mathbf{y}_0)}} \right\} . \tag{4}$$

According to Mean-Shift algorithm, the mean shift vector can be obtained through the formula

$$\mathbf{m}_{h,G}(\mathbf{x}) = \hat{\mathbf{y}}_1 - \mathbf{y}_0 = \frac{\sum_{i=1}^{n_h} \mathbf{x}_i w_i g\left(\left\|\frac{\mathbf{y}_0 - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i g\left(\left\|\frac{\mathbf{y}_0 - \mathbf{x}_i}{h}\right\|^2\right)} - \mathbf{y}_0 . \tag{5}$$

In formula (5), $w_i = \sum_{u=1}^{m} \delta\left[b\left(\mathbf{x}_i\right) - u\right] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(\mathbf{y}_0)}}$, where $g\left(\bullet\right)$ is the shadow kernel function of $k\left(\bullet\right)$. After a new position is obtained, Mean-Shift iteration is conducted with the new position as base point till the optimal central position of the object is obtained.

3. Improved Mean-Shift vehicle detection algorithms

3.1. Blob definition

Blob segmentation in target scene first came from human visual system. While observing the whole scene of an object, observers divide sceneries observed within the visual scope into several independent regions with similar features roughly according to comprehensive features of the observed object such as texture, outline and color, neglect minor details between regions, form several interested regions in vision and call each interested region as a blob.

Blob segmentation splits foreground object according to some rules under the support of such theory and makes arithmetic processing no longer pertinent to the whole foreground image and instead aim at local regions of image with comprehensive features such as color, position and size.

3.2. Blob segmentation

Blob segmentation is an important foundation of new blob detection. The exterior rectangle of a moving object is regarded as the shape of blob in the foreground. If the distance between the center of two adjacent blob rectangles, (i.e. two moving objects), is less than the sum of their width, the two moving objects represent a blob. Figure 1 shows its detection diagram.

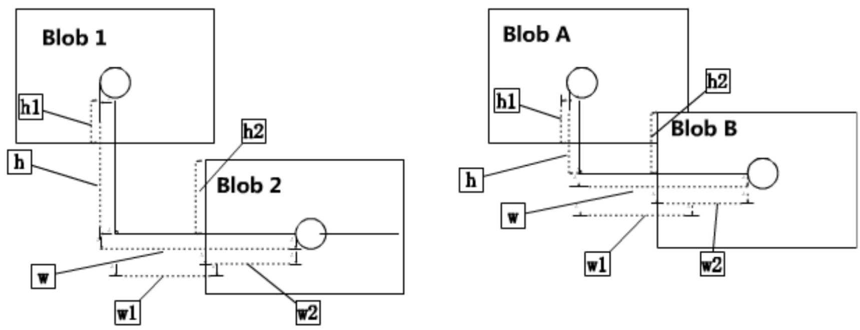

Fig. 1. Blob detection diagram

As shown in the right part of Fig. 1, if width w or height h between the center of blob 1 and blob 2 meets formula (6), blob 1 and blob 2 are two independent blobs; on the contrary, they belong to the same blob.

$$w > (w_1 + w_2)||h > (h_1 + h_2)\,. \tag{6}$$

Blob segmentation is the first step of blob detection. Based on the definition of blob, all blobs (moving vehicles) within detection region are put into a blob chain table. When a new vehicle enters the detection region, the new blob detected is added to blob chain table.

3.3. Location prediction

Kalman filtering can achieve location prediction of moving objects, reduce the calculating amount of Mean-Shift algorithm and meanwhile solve the problem of failure to extract features of target vehicle or wrong extraction caused by occlusion effectively. During actual tracking algorithm, this paper integrates location prediction of moving vehicle into Mean-Shift algorithm.

The dimension of blob is also required to set the size of Mean-Shift search box. Formulas for calculating blob dimension are as below.

$$Ratio = \frac{BlobH_{t0}}{BlobW_{t0}}\,, \tag{7}$$

$$BlobW_t = (LenTop + LenBtm) \times \frac{yPos_t - yBtm}{yTop - yBtm}\,, \tag{8}$$

$$BlobH_t = BlobW_t \times Ratio\,. \tag{9}$$

4. Multi-object tracking

Though Mean-Shift algorithm has many advantages in object tracking, it has two fatal deficiencies:In blob segmentation module, active blobs detected are put into a blob list and then the list is transmitted to tracking module. The number of blobs $BlobNum$ in the blob list is obtained before entry into tracking module. If $BlobNum > 0$, there is an object to be tracked currently. Formula (10) shows a loop computation:

$$BlobNum = \begin{cases} BlobNum - 1, \text{ if } BlobNum > 0\,, \\ 0 \text{ otherwise}\,. \end{cases} \tag{10}$$

Thread is launched for each loop computation; tracking algorithm is executed in the thread and multi-object tracking problem is achieved through multithreading.

5. Experimental simulation

5.1. Sing-object tracking

The test videos in Figs. 2 and 3 are in PETS2000 format and video object is an automobile with constantly reducing scale.

According to Figs. 2 and 3, traditional Mean-Shift algorithm can realize accurate object positioning and tracking in the process of tracking automobile with constantly reducing scale. However, when object scale changes, it is difficult to make automatic adjustment. The sequence has a total length of 299 frames. In video sequence before about frame 266, traditional Mean-Shift algorithm can track the object very well. Fig. 4 shows its tracking result, respectively the tracking result in frames 10, 19, 54, 252, 266 and 270.

According to Fig. 4, Mean-Shift algorithm has good tracking effect in the video before frame 266. The improved Mean-Shift algorithm can solve the problem of lost object tracking effectively. Fig. 6 shows its tracking result.

Fig. 2. Automobile tracking effect of traditional Mean-Shift algorithm

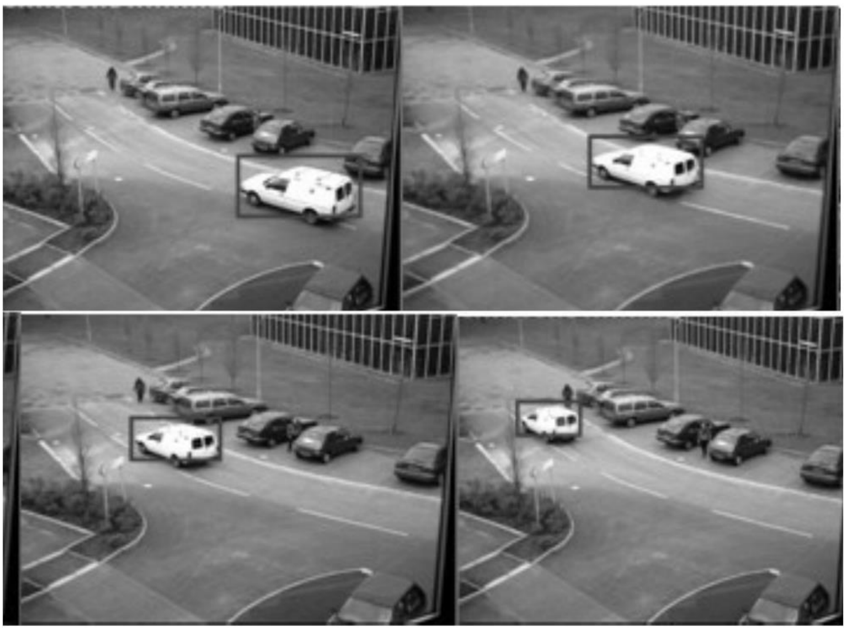

Fig. 3. Automobile tracking effect of improved Mean-Shift algorithm

Fig. 4. Video tracking result of Mean-Shift algorithm

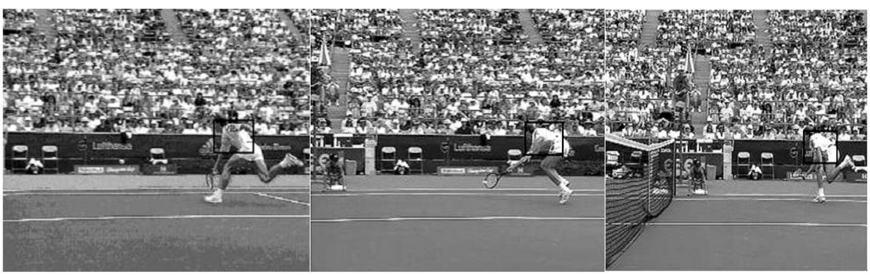

Fig. 5. Video tracking result of improved Mean-Shift algorithm

5.2. Multi-test tracking

Test videos in in Figs. 6 and 7 are in PETS2000 format and video objects are multi-object moving automobiles.

Fig. 6. Automobile tracking effect of traditional Mean-Shift algorithm

Fig. 7. Automobile tracking effect of improved Mean-Shift algorithm

According to Fig. 6, traditional Mean-Shift algorithm has some errors in multi-object tracking, which have a high frequency of occurrence. Fig. 6, left part, shows missed detection in the case of blob overlap. Fig. 6, middle part, refers to the identification of multiple vehicles into single vehicle in the case of unstable foreground and increase of search region of the algorithm. Figure 6, right part, shows object overlap when the tracking object is much larger than the actual object.

According to Fig. 7, compared to traditional Mean-Shift algorithm, problems in Fig. 6 are well solved. The improved Mean-Shift algorithm based on blob can not only achieve multi-object tracking, but also have more accurate tracking effect.

In allusion to object occlusion problem, football video sequence is used as the research object in order to verify the effectiveness of the algorithm in this paper. Figure 8 shows the tracking result of the algorithm in this paper which can solve the problem of number occlusion effectively.

According to the comparison of analysis results before and after improvement in Tables 1 and 2, the algorithm in this paper can reduce the number of iterations of Mean-shift in each frame and the total computation time of each frame is reduced relatively. The average running time of first 22 frames is 0.0185 s and the corresponding average time of Mean-Shift algorithm is 0.0284 s.

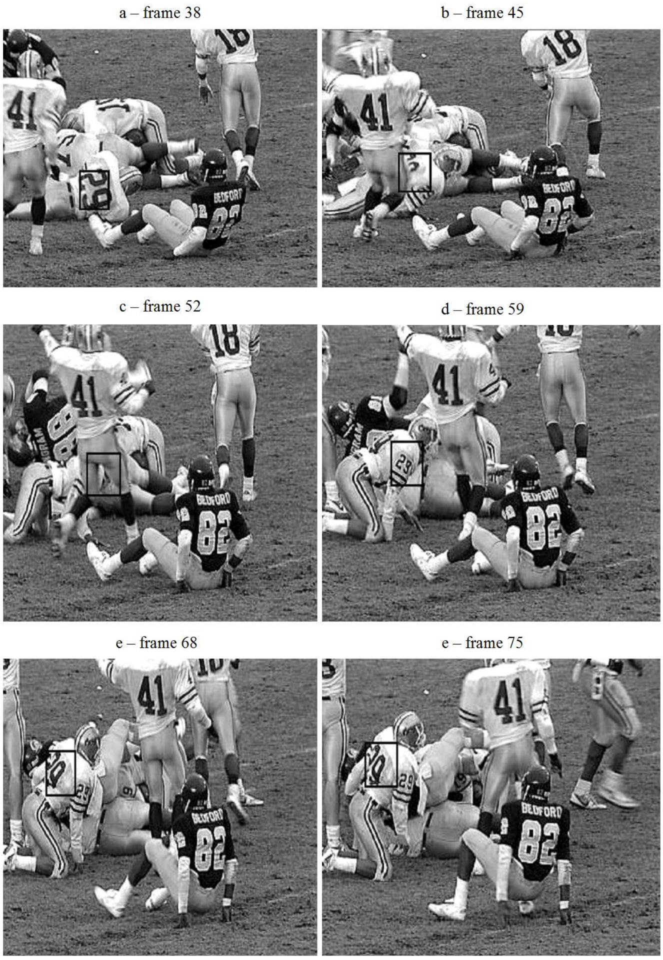

Fig. 8. Result of occlusion tracking of improved Mean-Shift algorithm

Table 1. Analysis result of improved Mean-Shift algorithm

Frame	2	3	4	5	6	7	8	9	10	11
Number of iterations	2	4	8	10	3	3	3	4	3	3
Time	0.016	0.016	0.032	0.046	0.016	0.015	0.016	0.015	0.016	0.015
Frame	12	13	14	15	16	17	18	19	20	21
Number of iterations	3	3	3	2	2	4	3	3	3	6
Time	0.016	0.015	0.016	0.015	0.016	0.017	0.016	0.016	0.016	0.022

Table 2. Analysis result of Mean-Shift algorithm

Frame	2	3	4	5	6	7	8	9	10	11
Number of iterations	3	3	18	4	3	3	6	5	3	4
Time	0.032	0.015	0.094	0.031	0.016	0.032	0.047	0.031	0.015	0.031
Frame	12	13	14	15	16	17	18	19	20	21
Number of iterations	3	3	4	3	2	5	4	4	2	3
Time	0.015	0.015	0.032	0.016	0.016	0.032	0.031	0.031	0.016	0.015

6. Conclusion

In allusion to defects of traditional Mean-Shift algorithm, this paper introduces blob detection into Mean-Shift algorithm, puts forward vehicle video tracking algorithm of an improved Mean-Shift algorithm and achieves accurate and fast positioning of single object and multiple objects. The experimental result shows that Mean-Shift algorithm based on blob can not only achieve multi-target tracking, but also have more accurate tracking effect. Its tracking effect is obviously better than that of traditional Mean-Shift algorithm.

References

[1] Z. Lv, A. Tek, F. Da Silva, C. Empereur-mot, M. Chavent, M. Baaden: *Game on, science - how video game technology may help biologists tackle visualization challenges.* PloS One *8* (2013), No. 3, e57990.

[2] W. T. Pan: *A new fruit fly optimization algorithm: Taking the financial distressmodel as an example.* Knowledge-Based Systems *26* (2012), 69–74.

[3] Y. Chen, Z. Cao: *An improved MRF-based change detection approach for multitemporal remote sensing imagery.* Signal Processing *93* (2013), No. 1, 163–175.

[4] G. F. Lu, Z. Lin, Z. Jin: *Face recognition using discriminant locality preserving projections based on maximum margin criterion.* Pattern Recognition *43* (2010), No. 10, 3572–3579.

[5] L. Ong, M. Motani: *Optimal routing for decode-forward in cooperative wireless networks.* IEEE Transactions on Communications *58* (2010), No. 8, 2345–2355.

[6] J. Yao, S. Feng, X. Zhou, Y. Liu: *Secure routing in multihop wireless ad-hoc networks with decode-and-forward relaying.* IEEE Transactions on Communications *64* (2016), No. 2, 753–764.

[7] K. Wang, K. Qiao, I. Sadooghi, X. Zhou, T. Li, M. Lang, I. Raicu: *Load-balanced and locality-aware scheduling for data-intensive workloads at extreme scales.* Concurrency and Computation Practice and Experience *28* (2016), No. 1, 70–94.

[8] Y. Lin, J. Yang, Z. Lv, W. Wei, H. Song: *A self-assessment stereo capture model applicable to the internet of things.* Sensors (Basel) *15* (2015), No. 8, 20925–20944.

[9] Y. Xie, Y. He, A. Cheng, J. Zhang: *Study on medical image enhancement based on*

IFOA improved grayscale image adaptive enhancement. Multimedia Tools and Applications *75* (2016), No. 22, 14367–14379.

[10] H. Li, H. S. Yang: *Fast and reliable image enhancement using fuzzy relaxation technique*. IEEE Transactions on Systems, Man, and Cybernetics *19*, (1989), No. 5, 1276 to 1281.

[11] H. R. Tizhoosh, G. Krell, B. Michaelis: *Knowledge-based enhancement of megavoltage images in radiation therapy using a hybrid neuro-fuzzy system*. Image and Vision Computing *19* (2001), No. 4, 217–233.

[12] S. K. Pal, R. A. King: *Image enhancement using smoothing with fuzzy sets*. IEEE Transactions on Systems, Man, and Cybernetics *11* (1981), No. 7, 494–501.

Received July 12, 2017