768

An Improved Kernel Density Estimation Approach for Moving Objects Detection

Bo Li^{1,*}, Yuhong Li² and Han Zhou³

¹College of Computer Science, South-Central University for Nationalities, Wuhan, 430074, China

²Digital Engineering and Emulation Research Center, Huazhong University of Science and Technology, Wuhan, 430074, China

³Department of Electrical and Electronic Engineering, The University of Hong Kong, Pakfulam Road, Hong Kong

Abstract: Moving object detection based on monitoring video system is often a challenging problem. Specially to monitor traffic at both day and night, in different weather and illumination conditions and with changeable background. Kernel Density Estimation (KDE) model is an effective approach to judge background and foreground, however, typical KDE uses fixed parameters, such as bandwidths, threshold, etc. This paper proposes a detection algorithm based on an Improved Kernel Density Estimation (IKDE) model. The proper bandwidths, adaptive background sample learning array, and adaptive threshold, and an improved sample updating method for sample learning array are discussed as the fundamentals of the IKDE model. Furthermore, an algorithm for restraining light field disturbance at night in video scene is proposed. Video image series are evaluated through the algorithm, and moving object detection is conducted in three different scenes. Results show that the algorithm can help to achieve a promising high accuracy and robustness for detecting moving objects.

Keywords: Adaptability, kernel density estimation, moving object detection, video analysis.

1. INTRODUCTION

Moving objects detection based on video system is considered as one of the fundamental challenging problems in the field of pattern recognition. Video-based method is a promising way due to its low cost and potential ability to collect a large amount of visual information, which can also infer higher-level information. In the past decades, many researchers devoted themselves to investigating different solutions for the intensively researched topic.

1) Related work

There are a plenty of research works from literatures on video-based approaches for object detection, such as references [1, 2]. It is well known that moving object detection is the basis for its classification. Therefore, it is necessary to detect the moving object or foreground at first. In recent years, many documents have contributed plenty to this field, see [3-5] and references therein for details. However, moving object detection based on video is a very challenging problem in applications due to the lighting changes, bad weather, etc.

Background subtraction is one of the most widely applied techniques for visual novelty detection. Many researchers detect objects based on the background difference. For example, Staufer and Grimson [6] modeled the pixels with a mixture of Gaussians. Piccardi [7] compared several background models. However, these methods are highly dependent on a perfect background model, which is very sensitive to changes in dynamic scenes. In such cases, some researchers presented some methods to detect objects through different characteristics. Robert [8] introduced a new framework to detect objects, which is based on a hierarchy of features detection and fusion. An approach was discussed in [9], based on Bayesian conditional probability density function, and [10] proposed an algorithm for the moving objects detection with the application of the color similarity analysis theory. There were few works discussing questions like the detection of objects and in particular, monitoring of traffic, both day and night, in different weather and illumination conditions, and changeable background. The drawbacks to vision-based vehicle detection include sensitivity to light and weather conditions. As the possible distributions of most background are unknown, it would be a good choice to adopt Kernel Density Estimation (KDE) [11] to establish background. The nonparametric kernel density estimation allows quick adaption to background changes.

However, typical KDE uses fixed parameters, such as bandwidths, threshold, etc. It is not flexible to deal with complex conditions.

This paper will discuss issues mentioned above and propose a detection algorithm based on an Improved Kernel Density Estimation (IKDE) model. The key questions are the



Fig. (1). The flow of moving objects detection.

effective way to establish a robust learning background sample array, and choose an adaptive threshold, etc. The algorithm implemented in this paper provides a much accurate detection when dealing with different scenes.

As a practice, those theories are verified in the Yangtze No.2 Bridge Security-Monitoring Project in Wuhan, China. Since there are about 200 thousands vehicles daily passing on the bridge, it is necessary to ensure the robustness and instantaneous nature of our video-based monitoring system.

2) Contributions

In this paper, KDE was adopted to detect the moving objects, especially the objects similar to the background. In our approach, an improved KDE model was proposed to judge the foreground and background. Initial background sample array was established by DBF (Difference Between two consecutive Frames), which could ensure the initial sample array is not contaminated by foreground. In addition, the background sample leaning array and the segmentation threshold are self-adaptively adjustable, which can make the system more robust and adaptive. Furthermore, this paper proposes an algorithm for restraining light field disturbance at night in video scene.

Since it is very difficult to detect objects at night, the disturbance light field at night is judged and removed according to some proposed conditions.

This paper is organized as follows. In Section 2, the method of object detection based on IKDE is introduced. In section 3, Disturbance light field restraining at night is discussed. In section 4, the experimental results show the accuracy achieved after the implementation of the proposed approaches. Section 5 includes conclusions and the prospect for future research.

2. MOVING OBJECT DETECTION BASED ON IM-PROVED KDE

Video-based monitoring systems rely a lot on moving object detection which is an important foundation. The basic principle of the method in this paper is the utilization of the

The Open Automation and Control Systems Journal, 2014, Volume 6 769

improved KDE algorithm. As we know, background model plays an important role in video monitoring system. However, it is very difficult to establish a robust and adaptive background model since the probability of the distribution of most background remains unknown and the background is always changeable in reality, in this study, KDE model is adopted to differentiate the background and foreground. Furthermore, the background sample learning array and segmentation threshold are self-adaptive based on the practical situation. Another difficulty is to detect the moving objects all day and night. In our approach, the disturbance light field restraining model is established to detect the foreground at night. The flow of moving objects detection can be seen in Fig. (1).

1) Kernel Density Estimation Model

Kernel density estimation model is a powerful tool which is able to classify the foreground and background without the probability distribution of the background [11]. And the fundamental idea of the classification is to capture the recent pixel information in the image sequences, and continuously update this information to get the fast changes in the background.

While the probability density distribution of pixels change quickly at any time, it is necessary to estimate the density function of this distribution with the given latest historical information if it is important to obtain sensitive detection. Suppose $x_1, x_2, ..., x_N, x_N$ are the latest values of a pixel within some historical frames, and based on these pixel values, we could get the value of xt of the pixel in probability density function at time t in the current frame with the help of nonparametric kernel Kh, it could be expressed as follows:

$$\Pr(x_{t}) = \frac{1}{N} \sum_{i=1}^{N} K_{h}(x_{t} - x_{i})$$
(1)

There are also some drawbacks of KDE approach. Although accurate, it is expensive in terms of the memory and computation time. In addition, it is always required to specify the threshold which is usually selected from some off-line experiments [1, 11]. Furthermore, a proper definition of bandwidth for the kernel function is hard to get but of vital importance, since a narrow one could lead to false positives while a wide one to false negatives [1]. Generally speaking, two bandwidth estimators could be applied to solve the problems [2, 12-14]. One is balloon estimator, of which the bandwidth is a function of the estimation point, and it depends on the sample-point estimator of the estimation point. The other is the sample point estimator changed by sample points, which means that the bandwidth is a function of sample point. The latter one is used in our study. In addition, a different kernel bandwidth is used for each pixel and for each color channel when the expected variations of pixel intensity varied over time. The kernel density function can be defined as follows:

$$\Pr(x) = \frac{1}{n} \sum_{i=1}^{n}$$
(2)

Where, x is an estimation point and x_i is a sample point. H(x_i) presents the bandwidth matrix at x_i , which determines the width and height of the kernel function. Aiming at simplicity, the $H(x_i)$ can be defined as a diagonal bandwidth matrix:

$$H = \begin{pmatrix} \sigma^{2}(c1) & 0 & 0 \\ 0 & \sigma^{2}(c2) & 0 \\ 0 & 0 & \sigma^{2}(c3) \end{pmatrix}$$
(3)

Where $\sigma^2(c)$ represents the bandwidth of a color channel. In order to reduce the computing complexity, diagonal matrix can be adopted for a m-dimension diagonal matrix H = diag(h₂₁; h₂₂; ...; h_{2m}), see [12] for details. Based on [11], the median absolute deviation over the sample for consecutive density values of the pixel is computed to estimate the kernel bandwidth for the j-th color channel of a given pixel. Namely, the median m of $|x_i - x_{i+1}|$ for each consecutive pair $(x_i - x_{i+1})$ in the sample is calculated independently for each color channel avoiding to be affected by few jumps of pixel intensities over time.

Suppose the local-in-time probability distribution to be Gaussian $N(\mu, \sigma^2)$, thus, the distribution of $(x_i - x_{i+1})$ would be $N(0, 2\sigma^2)$. Due to the symmetry of the distribution, we could get:

$$\Pr(N(0, 2\sigma^2) > m) \tag{4}$$

Hence, the standard deviation can be estimated as $= \frac{m}{m}$

$$\sigma = \frac{m}{0.68\sqrt{2}}$$

In [15], some choices are available for the kernel function. In this paper, we choose the Gaussian kernel as the kernel estimator function, Kh = N(0;H), where H represents the kernel function bandwidth.

Considering accuracy, from [16], we assume full bandwidth matrix H with a different kernel bandwidth σ_j for the j-th color channel, so the density can be estimated as:

$$\Pr(x_t) = \frac{1}{N} \sum_{i=1}^{N} \prod_{j=1}^{d} \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(x_d - x_d)^2}{2\sigma_j^2}}$$
(5)

Where, *d* refers to the number of color channels.

Based on KDE, some researchers, such as Mittal and Paragios [17], Sheikh *et al.* [18] proposed some improved models. Moreover, some other papers presented improved methodology for KDE which integrates the parametric and nonparametric classifiers and achieved the advantages of both of them [19].

Then, the pixel describing vector is presented as:

$$PDV = (f_t; x; y) \tag{6}$$

Here x and y stand for the coordinates of pixel point. Through the probability estimate on the foreground, the pixel would be judged as a foreground pixel if Pr(xt) < Th, where Th represents a global threshold of the whole frame which could be adjusted to achieve required false positives. Practically, with pre-calculated lookup tables given, the kernel function values difference $(x_i - x_i)$, the probability estimation can be calculated in a very fast way.

2) Samples learning array

When the initial selection of background samples is considered, it is very important to exclude the foreground samples. We adopt the method as follows:

Denote image series as I(x; y; i), where, x and y be the space coordinate, and i be the frame number (i = 1; 2; ...;N). And the grayscale change between two abutting frames can be denoted by DBF (Difference Between two consecutive Frames) as:

$$DBF(x, y, i) = \begin{cases} 0, if (diff < T) \\ C, if (diff \ge T) \end{cases}$$
(7)

Where, diff = |I(x, y, i+1) - I(x, y, i)|, C is a constant

number. There are some frame segments in which DBF is 0. As can be shown in Fig. (2), the average and variance for each DBF frame segment are calculated at the first step. Next, the DBF frame segments are selected to make sure the difference of average and variance among them is less than a threshold. Then, the sample pixels within will be chosen into the sample learning array for KDE calculation.

When updating the background sample learning array, it is necessary to choose background pixel into sample array, avoiding to be contaminated by foreground pixel.

This algorithm could save every pixel of the history samples array which obeys the principle of "first in first out". The length of array which is used for calculating the probability value of the follow-up pixel is set as a value, considering the calculating precision and speed. The array needs to be updated when new frame comes forth. In this paper, two updated methods are introduced [11, 13].

Selective updating method: if the new pixel is judged as background, it will be included into the sample array.

Non-selective updating method: the new pixel will be directly included into the sample array.

The first one is able to increase the sensitivity as the foreground pixels will not enter the array. When continuous pixels are judged as background the array may not be updated for a long time. Hence, the changes around pixels cannot be gathered accurately and timely.

On the other hand, the rate of foreground pixels in the array will increase by the second method, and the continuously coming forth of the foreground will lead to errors in judging the pixel of the new frame.

This paper proposed an improved sample updating method. Every new pixel is involved in the sample array whether it belongs to the background or not, but with a given weight based on its category:

$$\Pr(x) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{d} \frac{W_i}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{1}{2} \frac{(x_i - x_j)^2}{\sigma_j^2}}$$
(8)



Fig. (2). Initial sample learning array.



Fig. (3). The histogram of probability density function. (a) Initial frame. (b) Grayscale image of probability density function of (a), (c) probability histogram of (b).

Where,
$$w_i = \begin{cases} \alpha, & if(background pixel) \\ 1 - \alpha, & (otherwise) \end{cases}$$

 α is determined based on the experiments. For example, α is 0.8. Also, the length of sample array for learning is self-adaptive, based on background change. Three lines pixels can be chosen to calculate the grayscale mean, which can be regarded as the background gray mean to some extent. And the difference of gray mean between two consecutive frames is calculated. The length of sample array can be adjusted as follows:

$$LengthRatio = \begin{cases} 100\%, if (diff < Thr1) \\ 50\%, if (Thr1 < diff < Thr2) \\ 10\%, if (diff > Thr2) \end{cases}$$
(9)

Where, LengthRt stands for the length ratio in the sample array for learning accounting for the length of whole sample array.

It is indicated that the background changes suddenly when diff > Thr2, and smoothly when diff < Thr1. The adaptive sample learning array can reduce computing time when the background changes suddenly.

3) Threshold

The threshold of probability for judging foreground is determined based on experience, which is often fixed. However, it is very likely that the proportion of foreground in the whole image is often varied. Therefore, a self-adaptive threshold adjusting method is adopted in this paper.

In Fig. (**3a**) is the initial frame image, (b) is the grayscale image, in which every pixel value is 255 times the probability density function of corresponding pixels in (a), (c) describes the probability histogram of (b). As it is shown in Fig. (**3b**), the black region represents foreground, and the tint and white region is background. From Fig. (**3c**), we could get that the left part belongs to foreground, while the right part belongs to background. But it is not easy to find the exact breakpoint in the probability histogram. Therefore, in our work, the threshold Th for judging foreground or background is chosen based on real scene by human visual aid at first, and then adjusted adaptively when the light is changed suddenly.

3. DISTURBANCE LIGHT FIELD RESTRAINING AT NIGHT

The vast majority of vision-based vehicle detection papers are dedicated to daytime conditions. Nighttime conditions may be dealt with in a few ways. Various studies have trained specific models for detecting vehicles at nighttime, often by detecting the headlights and taillights of vehicles encountered on the road. At night, there are some disturbance light fields from the vehicle light as shown Fig. (4). Generally speaking, it is very difficult to remove them by medium filtering and mathematics morphology.



Fig. (4). Disturbance light.



Fig. (5). Disturbance light field processing.

Disturbance light field have some typical features, such as, high light intensity and the short-range with the moving objects. Then, it could possibly be eliminated through the points below.

If a pixel P meets these conditions below, it can be judged as disturbing light field.

f1: $P \in M(x; y)$. Where, M(x; y) is the initial foreground pixels aggregated by KDE model.

f2: Pixel feature judge. Generally speaking, the pixels in disturbance light field have certain characteristics, such as, high luminance.

f3: Location judge. The vehicle light disturbance light field is located in front of the moving vehicle.

It is important to monitor moving objects at night. In our work, a threshold of light intensity is set to judge whether it is night or not. When it is night, some preprocessing is conducted, for example, grayscale stretch. Then, the KDE model is utilized to judge the initial foreground, and remove the disturbance light field, such as vehicle light. Some follow up processes such as medium filtering and mathematic morphology operations, are done to remove noises, the flow chart is shown in Fig. (5). Some processing results at night are shown as Fig. (8).

4. EXPERIMENTAL RESULTS AND DISCUS-SION

The experimental samples came from some video monitoring scenes. One image sequences whose size is 352 * 288 pixels captured from camera on the No.2 Yangtze River Bridge in Wuhan, China, including samples in day and night, from the scene we can see that the top view of vehicles from



(a) Initial image

(b) Foreground pixel judging,

(c) Object detection

Fig. (6). Object detecting.



(a) Frame 62th

(**c**) Frame 136th





(**d**) Frame 146^{th}

Fig. (7). Detection results at daytime.

video image series is captured. Another image sequences whose size is 768 * 576 pixels are real traffic scenes from Internet. The implementation has been done in VC++6.0. In our video-based monitoring scene, these image sequences are 24bits true color, but their color information is not evident and similar to grayscale after capturing. The R, G, B channels are almost the same, that is, the frame image resembles the grayscale image. Therefore, in our work, just one of the channels is used instead, saving computation time because of color space transformations. The basic steps of objects detection is shown in Fig. (6).

Considering the calculating precision and speed, the length of array which is used for calculating the probability value of the follow-up pixel is set at 80. W_i of background point is set as 0.8~0.9, and W_i of foreground point is set as 0.2~0.1.

The results of moving objects detection of three series of samples are shown in Fig. (7), Fig. (8) and Fig. (9).

The results can be shown in three videos with different types of backgrounds and illumination condition from Fig. (6) to Fig. (9).

Fig. (6) and Fig. (7) show a scene with normal global illumination, and the main difficulty comes from the detection of dark vehicles that are like the background. There may be some dark moving objects that do not have sufficient gradient features to identify that they are vehicles in some cases. Fig. (8) shows a nighttime scene and the difficulty in the night sequence is correctly segmenting the headlight on the road with respect to the moving vehicle.

Foggy weather condition is a challenging problem as shown in Fig. (9), because many image processing algorithms are affected by the biased and low-contrast scene radiance resulting from haze weather condition.

For testing the performance of our approach, we performed two kinds of approaches, including a background



(a) Frame 197th initial image



(b) Object detection



(c) Frame 394th initail image



(d) Object detection

Fig. (8). Detection results at night.



(a) Frame 84th



(b) Foreground judge



(c) Object judge



(d) Object detection



(e) Frame 92^{th}

Li et al.



 (\mathbf{f}) Foreground judge



(g) Object judge



(h) Object detection

Fig. (9). Vehicle detection in foggy day.



(a) Initial frame



(b) Object detection using approach in this paper



(c) Object detection using background subtraction



(d) Initial frame



(e) Object detection using approach in this paper



(f) Object detection using background subtraction

Fig. (10). Comparision of results of different approaches.

Table 1. The Results of detection rate.	ts of detection rate.
---	-----------------------

Moving Vehicle	Approaches	Scene 1 (sunny day)	Scene 2 (night)	Scene 3 (foggy)
True Positive	BS	60	40	30
	This paper	63	42	33
False positive	BS	4	8	6
	This paper	3	4	3
False negative	BS	7	6	5
	This paper	4	4	2
Precision rate	BS	93.7%	83.3%	83.3%
	This paper	95.4%)	91.3%	91.7%
Recall rate	BS	89.5%	86.9%	85.7%
	This paper	94%	91.3%	94.3%

subtraction and the approach used in the paper. The results comparison can be shown in Fig. (10). From Fig. (10), it is easily found that there is low false positive in our approach. The results of detection rate can be seen in Table 1.

Where:

 $Precision \ rate = \frac{True \ positive}{True \ positive + False \ positive}$

 $Recall \ rate = \frac{True \ positive}{True \ positive + False \ negative}$

In Table 1, "BS" refers to the approach of background subtraction, and "This paper" refers to the approach used in this paper. The number of moving vehicle in our video frames adopted in the experiment is 67, 46, and 35 in the above three scenes.

It can be seen that in these videos, we obtain good vehicle precision, recall rates, which is beyond common approaches such as background subtraction, which is thus satisfactory enough to consider a practice in real system.

CONCLUSION

In this paper, we have proposed a methodology for moving object detection based on improved KDE. In our approach, an improved KDE model is introduced to judge foreground and background. Initial background sample array is established by DBF, which ensures the initial sample array is not contaminated by foreground. In addition, the background sample leaning array and the segmentation threshold are self-adaptive, which can make the system robust and adaptive. In this study, the disturbance light field in night is judged and removed based on some proposed conditions.

The experiment results show that the algorithms are feasible and the approaches are efficient. With abundant image sequences under real conditions, the experiment indicates that it is feasible to adopt this algorithm in video-based monitoring scene, especially for vehicle detection and classification to achieve accuracy. Future work can concentrate on extending the methodology to improve the selection on adaptive feature.

CONFLICT OF INTEREST

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

ACKNOWLEDGEMENTS

This investigation is supported by the Project of Natural Science Foundation of Hubei Province of China (No. 2012FFC13401).

REFERENCES

- S. Gupte, O. Masoud, R.F.K. Martin, and N.P. Papanikolopoulos, "Detection and classification of vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 37-47, 2002.
- [2] Z.H. Sun, G. Bebis, and R. Miller, "Monocular precrash vehicle detection: features and classifiers," *IEEE Transactions on Image Processing*, vol.15, no.7, pp. 2019-2034, 2006.
- [3] P. Angelov, P. Sadeghi-Tehran, and R. Ramezani, "An approach to automatic real-time novelty detection, object identification, and tracking in video streams based on recursive density estimation and evolving takagisugeno fuzzy systems," *International Journal of Intelligent Systems*, vol. 26, no. 3, pp. 189-205, 2011.
- [4] D. Szolgay, J. Benois-Pineau, R. Megret, Y. Gaestel, and J.F. Dartigues, "Detection of moving foreground objects in videos with strong camera motion," *Pattern Analysis and Applications*, vol. 14, no. 3, pp. 311-328, 2011.
- [5] C. Liu, P.C. Yuen, and G.P. Qiu, "Object motion detection using information theoretic spatio-temporal saliency," *Pattern Recognition*, vol. 42, no. 11, pp. 2897-2906, 2009.
- [6] C. Stauffer, and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 246-252, 1999.
- [7] M. Piccardi, "Background subtraction techniques: a review," *IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, pp. 3099-3104, 2004.

An Improved Kernel Density Estimation Approach

The Open Automation and Control Systems Journal, 2014, Volume 6 781

- K. Robert, "Video-based traffic monitoring at day and night [8] vehicle features detection tracking," In: 12th International IEEE Conference on Intelligent Transportation Systems, 2009, pp. 1-6.
- R.E. Bethel, B. Shapo, and C.M. Kreucher, "PDF target detection [9] and tracking", Signal Processing, vol. 90, no. 7, pp. 2164-2176, 2010.
- X. Zhang, and J. Yang, "The analysis of the color similarity prob-[10] lem in moving object detection," Signal Processing, vol. 89, no. 4, pp. 685-691, 2009.
- A. Elgammal, R. Duraiswami, D. Harwood, and L. Davis, [11] "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance," Proceedings of the IEEE, vol. 90, no. 7, pp. 1151-1163, 2002.
- A. Elgammal, R. Duraiswami, and L. Davis, "Probabilistic track-[12] ing in joint feature-spatial spaces," IEEE, vol. 1, pp. 781-788, 2003.
- [13] K. Karman, Moving object recognition using an adaptive background memory, Time Varying Image Processing and Moving Object Recognition: Elsevier Science Publishers B.V., 1990.

Received: November 26, 2014

© Li et al.; Licensee Bentham Open.

This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.

- [14] A. Monnet, A. Mittal, N. Paragios, and V. Ramesh, "Background modeling and subtraction of dynamic scenes," In: 9th IEEE International Conference on Computer Vision, 2003, pp. 1305-1312.
- B. Silverman, "Density estimation for statistics and data analysis," [15] London: Chapman & Hall/CRC, 1986, vol. 26.
- T. Duong, "ks: Kernel density estimation and kernel discriminant [16] analysis for multivariate data in r", Journal of Statistical Software, vol. 21, no. 7, pp. 1-16, 2007.
- A. Mittal, and N. Paragios, "Motion-based background subtrac-[17] tion using adaptive kernel density estimation," In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 302-309, 2004.
- [18] Y. Sheikh, and M. Shah, "Bayesian modeling of dynamic scenes for object detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 11, pp. 1778-1792, 2005.
- [19] P. Chaudhuri, A. K. Ghosh, and H. Oja, "Classification based on hybridization of parametric and nonparametric classifiers," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 7, pp. 1153-1164, 2009.

Revised: January 08, 2015

Accepted: January 21, 2015