Sensors & Transducers, Vol. 178, Issue 9, September 2014, pp. 238-243

Sensors & Transducers

Adaptive Thresholding Approach of Moving Object Detection

Man Hua, Yanling Li, Ruichun Lin
School of Computer Science, Civil Aviation Flight University of China, Guanghan, China
Tel.: +86 08385182599
E-mail: hua.man@163.com

Received: 13 May 2014 /Accepted: 29 August 2014 /Published: 30 September 2014

Abstract: Modeling background and segmenting moving objects are significant techniques for video surveillance and other video processing applications. In this paper, we proposed a novel adaptive approach modeling background and segmenting moving object with non-parametric kernel density estimation. Unlike previous approaches to object detection which detect objects by global threshold, we use a local threshold to reflect temporal persistence. With combined of global threshold and local thresholds, the proposed approach can handle scenes containing gradual illumination variations and noise and has no bootstrapping limitations. The presentation of the proposed approach is complemented by results on challenging standard test data and comparisons with other standard techniques. *Copyright © 2014 IFSA Publishing, S. L.*

Keywords: Adaptive, Thresholding, Kernel density estimation, Background, Foreground.

1. Introduction

Moving object detection and segmentation is an important research topic in the field of computer vision, which were widely used in the area of video surveillance and video compression. At present, there are 3 methods for motion detection: optical flow method, frame difference method and background subtraction method. Because of the detecting speed and effects are more ideal of Background subtraction method, it attracted a lot of researcher's interest in recent years. The main idea of Background subtraction is use the background subtraction frame subtracting background difference, and use the threshold to get the two values of difference. As a result, we get the moving target template. Therefore, the effective background modeling and the setting of threshold are the key points of background subtraction. Recently, there has been a large amount of work addressing the issues of background model representation and adaptation. A robust background modeling is to represent each pixel of the background image over time by a mixture of Gaussians [1]. This approach was first proposed by Stauffer and Grimson [2, 3], and became a standard background updating procedure for comparison. Instead of modeling the feature vectors of each pixel by a mixture of several Gaussians [4], Elgammal proposed to evaluate the probability of a background pixel using a nonparametric kernel density estimation (KDE) based on very recent historical samples in the image sequence [5]. Mittal improve this KDE based background model by introducing variable bandwidth kernels and optical flow [6].

The main limitation of most traditional statistical solutions is their need for a series of training frames absent of moving objects [7]. However, in some situations, e.g., public areas, it is difficult or impossible to control the area being monitored. In such cases it may be necessary to train the model using a sequence which contains foreground objects. Another limitation of these methods is most schemes determine the foreground threshold experimentally [8].

In this paper, we propose a method that overcomes these limitations. Our aim for such a framework is:

A reference background image that contains no moving objects may be not required.

Adaptive thresholding to make the system adaptable to scene changes and illumination variation.

This paper is organized as follows: section second introduces the adaptive global and local thresholding method based on KDE; the third section describes background update strategy; the fourth section gives the results of experimental; finally, conclusions are given.

2. Adaptive Thresholding

2.1. Modelling the Dissimilarity Measure Statistics

In this paper, we propose a novel adaptive thresholding scheme which uses two different types of adaptation. First we perform a statistics based threshold detection, then the spatial cues is used to verify the threshold and perform adjustment according to the spatial continuous of foreground.

If we estimate the density of $p(x_t)$, the possibility of each pixel belong to background, we can get the KDE graph (Fig. 1).

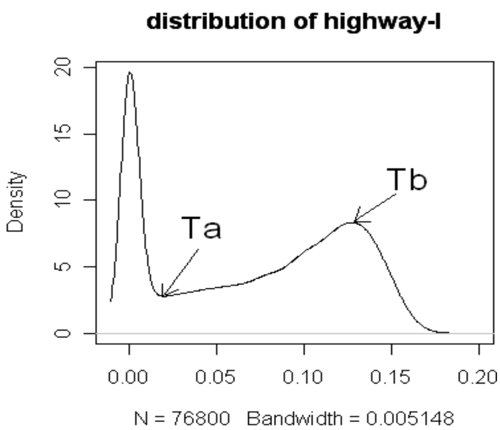


Fig. 1. Kernel density of possibility of each pixel belongs to background in an image. The training sample is 40 frames, take 50-th frame for example.

We can notice the distribution of $p(x_t)$ is usually bimodal. If the input pixel is significantly different from the corresponding background estimate, the possibility of this pixel belongs to background which represents by $p(x_t)$ has small value. While the bigger $p(x_t)$ value represents the close similarity of background. The foreground distribution usually centers around zero and has relatively small deviation and sharp crest than background, and the background deviation depends on the variations such as illumination or with animated texture. We can use the first trough from zero as the foreground threshold, as any $p(x_t)$ left to this trough can be classified as foreground safely. As mention before, the second crest is from the background distribution. Any possibility larger than this point, can be concluded as background.

Then we can determine two thresholds from the KDE graph, Ta and Tb as background and background threshold accordingly.

If we treat the KDE graph as a histogram H and the bin index (i) of histogram associated with the value of $p(x_t)$, the H(i) denotes the density of $p(x_t)$. The Ta can be found by the first match of below condition:

$$\begin{matrix} H(Ta) - H(Ta-1) < 0 \\ H(Ta) - H(Ta+1) < 0 \end{matrix} \quad \text{Ta>0}, \tag{1}$$

While Tb can be determined by finding the largest peak from Ta by:

$$\begin{matrix} Tb = \arg\max(H(Tb)), \\ H(Tb) - H(Tb-1) > 0, H(Tb) - H(Tb+1) > 0, Ta < Tb \end{matrix}, \tag{2}$$

In this manner, any $p(x_t)$ below Ta will consider as foreground and larger than Tb will take as background. The value between Ta and Tb need further information to classification.

So far, only the temporal intensity distribution is considered in background model, but the spatial clues also play an import role in foreground detection.

2.2. Robust Background Modeling

As it is usually hard to know prior knowledge of the scene, a non-parametric approach able to handle arbitrary densities is more suitable. A particular nonparametric technique that estimates the underlying density, avoids having to store the complete data, and is quite general is the kernel density estimation technique. In this technique, the underlying pdf is estimated as:

$$p(x_t) = \sum_{t=1}^{N} W(x_i) \cdot K(x_i - x_t), \tag{3}$$

where $K(x_i - x_t)$ is the kernel function, which is usually taken to be a density function, and $W(x_i)$ is the re-weighting function which can be adjusted to control the roles of different data points in the sample... N is the number of sample. If we choose our kernel function to be Gaussian, then the density can be estimated as:

$$p(x_t) = \sum_{i=1}^{N} W(x_i) \cdot \frac{1}{h\sqrt{2\pi}} \exp(-\frac{(x_i - x_t)^2}{2h^2}), \quad (4)$$

where x_t is the color/intensity feature and h is the bandwidth which controls the smooth of the estimate. Often the re-weighting function $W(x_i)$ is required to be non-negative and sum up to 1.

In most of the previous works, the uniform weights are used typically [9], same influence of each pixel assumption is made. However, assuming that the each pixel play same role in background may be flawed.

In order to obtain a reliable estimate, we use formula 5 to count continual unchanged pixel value:

$$c_i(t) = \begin{cases} c_i(t)+1, |g(t)-g(s)|<\delta \\ 0, |g(t)-g(s)|>\delta \end{cases}, 1 < s < N, \quad (5)$$

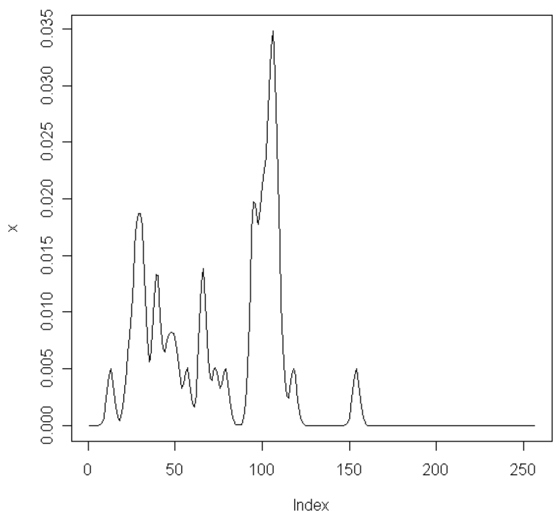
where $g(t)$ is the intensity value on time t, and δ is the small threshold used to make decision if a pixel value changed, the threshold is dependent on noise of the image, normally can use a const less than 10. Then we define the weight function $W(x_i)$ as below:

$$W(x_i) = \frac{C_i(t)}{\sum_{t=1}^{N} c_i(t)}, \quad (6)$$

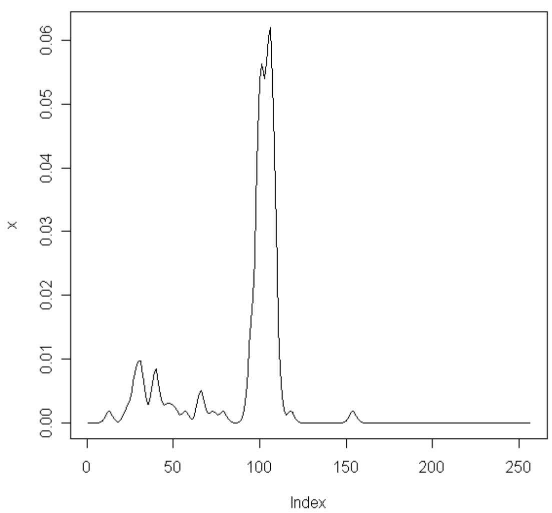
As background distribution is temporally stationary than foreground, the continual unchanged counter for a background is larger than foreground. Consequently, the more weight is assigned to background pixel in the weighted kernel density estimation.

Fig. 2 indicates the KDE generate by weighted(b) and without weighted(a), the temporal intensity distribution of the pixel without apply weight is multiple-modal, while the distribution is almost unimodal and centered at intensity 100 in weighted KDE, the interfere with foreground is suppressed.

Here the background representation is drawn by estimating the probability density function of each pixel with higher possibility belongs to background in the background model. The current pixel is declared as foreground if it is unlikely to come from this background distribution, i.e. $p(x_t)$ is smaller than some predefined threshold. It is usually not easy to determine such a threshold, a popular threshold detection scheme is based on the normalized statistics which consider the mean and the standard deviation of for all spatial locations. It can be adaptive with noise and illumination variation.



(a)



(b)

Fig. 2. The KDE generated by weighted and not weighted with pixel.

3. Background Update Strategy

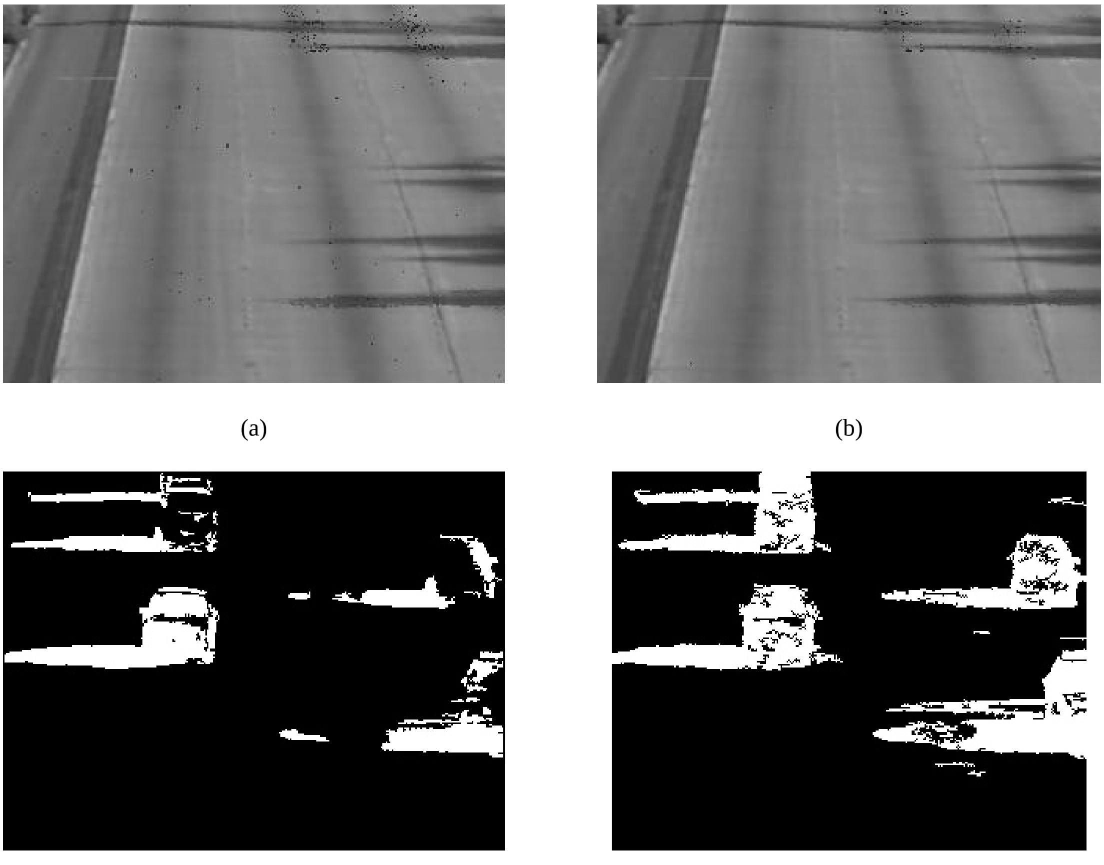
The background should be updated automatic when scene change or illumination variation abruptly or gradually. Slow-moving objects may blend into the background model if the background model adapts too fast, and it will fail to identify the portion of a foreground object that has corrupted the background model. To overcome this problem, we check if a pixel is stable enough by several continuous images to avoid the problem of blending foreground into background model. The checking is performed by:

$$| C_t(i) - C_{t-j}(i) | < \delta_c, 0 < j < L, \quad (7)$$

where $C_t(i)$ is the intensity value of pixel i at time t. δ_c is the small threshold defined to make decision if two pixel values are unchanged, and L denotes the continuous unchanged image number. If the pixel value is left unchanged in several images, it will be updated into the background. If an abrupt scene change has been detected, the background model needs re-initialization, it usually occurs large percentage (above 80 %) of foreground detected and in several continuous images [10].

4. Experimental Results

A typical traffic surveillance scenario is considered the first test sequences. The first is the challenging scene of the heavy traffic high way because moving vehicles are present in every frame of the sequence. In this example the background extraction results of weighed KDE (See Fig. 3(b)) are better than the one without adding weight (See Fig. 3(a)). As the moving vehicles occludes the background in upper location for large percentage of the total time, the normal KDE still treat the vehicles color/intensity as background, then we can observe that in the upper location there are some stain introduced by foreground. With the more weight is assigned to background pixel in the weighted kernel density estimation for temporal stable pixels, the background model can reflect the real scene. Fig. 3(c) and Fig. 3(d) are the foreground mask images for each method. Without weighted kernel shows some miss-detection in the vehicles shows, it is due to the blending of foreground which corrupts this portion of background. Our method has significant improvement compared with the typical method.



(a) (b)

(c) (d)

Fig. 3. The background extraction results of diffident ways.

We compare our algorithm's output with that of several existing techniques used for background modeling, and present compared results obtained with typical mixture of Gaussian and KDE method. Each of these methods is trained on 15 to 40 frames depending on the length of sequence. Also, the parameter for MoG method has been left unchanged in the implementation of OpenCv, and the threshold for KDE has tuned to produce the best possible results for the sequences presented in Figs. 4. Tests are performed on several sequences representative of situations which might be commonly encountered in surveillance video. Here, we describe four typical scenes:

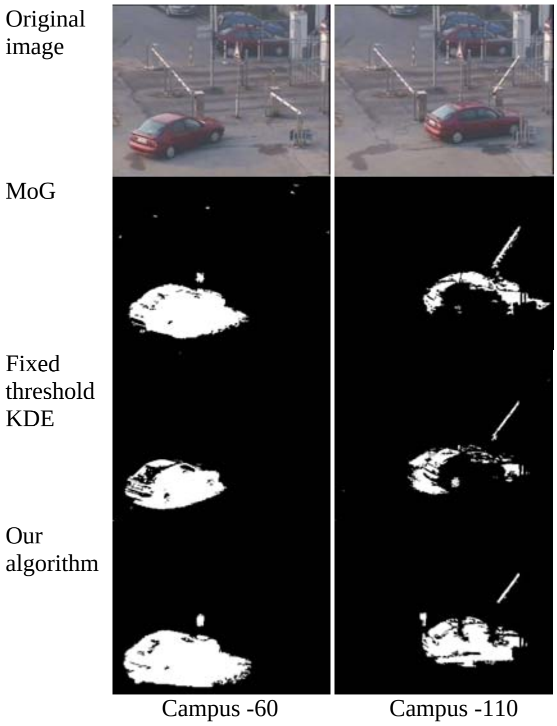


Fig. 4-1. The test results comparison of different algorithm.

Fig. 4-2. The test results comparison of different algorithm.

First column is the test result on Pets2006, as in this sequences, the moving man exists in all training image of background, it is a challenging scene as the man can blend into background easily as he moves slowly from the beginning. Both the moving person and the background details are presented in the resulting image for either the MOG or Fixed KDE, our algorithm can deal with the foreground in training phase and exclude foreground to the background model, it leads to the suppress the fake foreground output. Our method avoids the problem of blending pixel values present in many current methods. In the campus video in Fig.4-2, our approach also produces less mis-detection but the shadow issue still exists. Our method can detect the foreground nearly perfect as perform the adaptive threshold and the additional use of the special clues.

We have discovered that the adaptive thresholding methods have shown significantly better results for all the algorithms concerned.

5. Conclusions

In this paper, an adaptive target detection algorithm is proposed. The algorithm based on KDE use a method combined with global and local threshold can effectively solve the problem of threshold setting different laws of different pixels, improve the detection accuracy and avoids the phenomenon of mixing. At the same time, the algorithm based on the data distribution can adjust parameters automatically in a wide range of parameter variations. A large number of video experiments have been done and the results show that this method can correctly segment the foreground object motion and demonstrate the robustness of the method. How to eliminate the inner shadow and target "hole" will be the next step of work.

Acknowledgments

This work is supported by Joint Fund of the National Natural Science Foundation of China and the Civil Aviation Administration of China (No. U1233105).

References

[1]. N. Friedman, S. Russell, Image segmentation in video sequences: a probabilistic approach, in *Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence (UAI'97)*, Providence, USA, 19-21 May 1997, pp. 175-181.

[2]. D. Gutchess, M. Trajkonic, E. Cohen-Solal, D. Lyons, A background model initialization algorithm for video surveillance, in *Proceedings of the 8th IEEE International Conference on Computer Vision*, Vancouver, Canada, 25-27 August 2001, pp. 733-740.

[3]. Z. Q. Hou, C. Z. Han, A background reconstruction algorithm based on pixel intensity classification, *Journal of Software*, Vol. 19, Issue 9, 2005, pp. 1568-1576.

[4]. D. S. Lee, Effective Gaussian mixture learning for video background subtraction, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 27, Issue 5, 2005, pp. 827-832.

[5]. A. Elgammal, R. Duraiswami, D. Harwood, L. S. Davis, Background and foreground modeling using nonparametric kernel density estimation for visual surveillance, in *Proceedings of the IEEE International Conference on Computer Vision*, Vancouver, Canada, 25-27 August 2001, pp. 1151-1163.

[6]. A. Mittal, N. Paragios, Motion-based background subtraction using adaptive kernel density estimation, in *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, Washington, USA, 5-8 January 2004, pp. 302-309.

[7]. K. Kim, T. H. Chalidabhongse, D. Harwood, and L. S. Davis, Real-time foreground-background segmentation using codebook model, *Real-Time Imaging*, Vol. 11, Issue 3, 2005, pp. 167-256.

[8]. D. B. Xu, L. Huang, Adaptive kernel density estimation for motion detection, *Acta Automatica Sinica*, Vol. 35, Issue 4, 2009, pp. 379-385.

[9]. S. Sheather, J. Jones, A reliable data-based bandwidth selection method for kernel density estimation, *Journal of the Royal Statistical Society Series B*, Vol. 53, Issue 3, 1991, pp. 683-690.

[10]. I. Haritaoglu, D. Harwood, L. S. Davis, Real-time surveillance of people and their activities, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, Issue 8, 2000, pp. 809-830.
