Using Traffic Light Signal to Enhance Intersection Foreground Detection Based on Video Sensor Networks

Rong Ding, Shunli Wang, and Xu Liu

State Key Laboratory of Software Development Environment, School of Computer Science and Engineering, Beihang University, Beijing 100191, China

Correspondence should be addressed to Rong Ding; dingr@buaa.edu.cn

Received 4 November 2013; Revised 1 March 2014; Accepted 28 March 2014; Published 15 April 2014

Copyright © 2014 Rong Ding et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. Introduction

Intelligent video surveillance, aiming at making traffic more intelligent and decreasing the amount of vehicle accidents, is a well-studied subject area with both existing application systems and new approaches still being developed. Among this area, detecting objects at the intersection is one of the most significant focuses in typical intelligent transportation systems (ITS) applications and the basis of high-level processing. In the most real intersections, a single camera is not enough to monitor the whole scenario. Video sensor network provides a large-scale, redundant, of video streams to observe the intersections [1, 2]. Because some of video-based traffic monitoring systems include high-level description of both cars and their behaviours, continuous tracking result is significant to the high-level processing. Background subtraction is a widely used technique for foreground detection which compares an observed image with an estimated background image that does not contain any objects of interest. But before using this method, several parameters have to be determined. Among these arguments, the learning rate is more critical to the performance. If the rate is large, the slow or stopped vehicles will fuse into background quickly just as Figure 1 describes. But if we set the rate at a small value, the background will not be updated in time. In particular, in a traffic junction scene, vehicles always encounter congestion and stop-and-go when there is a red light. At this time, a reasonable learning rate becomes more significant. Usually, the majorities of current methods adjust learning rate only relying on the changes from video sequences. This causes very often the method to be unable to adapt itself with the outside world changes instantly. And when the traffic light is red, it mainly leads the tracking of vehicles to be interrupted. Once the light turns green, the cars move again and new tracking of them is constructed. In summary, it has a bad effect on the continuous tracking of foreground objects and reduces the accuracy of some high-level understanding methods, such as [3, 4]. Unlike previous methods, this paper focuses on how to adjust learning rate according to the real-time and accurate physical world signals from other sensors. And this
technique guarantees the continuous tracking of vehicles in traffic junction.

In this paper, we select the more common traffic light as the external signal to improve the results of foreground detection. Meanwhile, to divide input images into reasonable regions, road line detection result is used. Because the camera is static, road lines are just detected only once in the whole monitoring period. If the system receives a red light signal, it indicates that vehicles will slow down and stop after some frames. To avoid these interested vehicles blending in background and losing existed tracking, we decrease the learning rates of pixels in the red light region while the rates of other pixels remain unchanged. When the system receives a green light signal, it means that vehicles run through the intersection at normal speed and normal learning rates are selected. Experimental results show that this kind of environment information can greatly improve the results of the background subtraction and foreground detection.

We have done some work in improving the detection results by combining the video sequence and physical world information before, which is published in paper [5]. This paper is an expanded version of [5], which analyzes the method more in detail, sets more contrast experiments, and adds quantitative evaluation of experiments' results. The remaining of this paper is organised as follows. In Section 2, a compact review of important developments and existing improvements about background subtraction and foreground detection are presented. We propose the proposed framework and outline three classical methods and illustrate how to use our approach to make these methods perform better in Section 3. Further explanations of our method in a practical application scenario are supplied in Section 4. Meanwhile, the contrast experiment results between precious methods and our method are presented. The last section is conclusion and future work.

2. Related Research

There has been numerous works devoted to the development of background subtraction for real-time video processing. Several surveys devoted to this topic can be found in [6, 7]. The statistical tools provide a good framework to model the background of a complex traffic scene and so many methods have been developed.

During these methods, Gaussians mixture model (GMM), first presented in [8], models the distribution of the values observed at each pixel by a weighted mixture of Gaussians. GMM is able to cope with the multimodal nature of many practical situations and leads to good results when there are repetitive background motions, such as leaves shaking or water rippling. By far, GMM is the most researched and applied method. Many enhanced algorithms [9] have been proposed all along these years. Reviews of them can be seen in literature [7, 10]. The weakness of GMM lies in its strong assumption that the background is more frequently visible than the foreground and that its variance is significantly lower. Also, the initialization of the model and estimation of the parameters are problematic and uncertain.

In different real-world environments. Therefore, traditional GMM-based methods with empirical value usually are not competent to good background subtraction results. To avoid the difficulty of finding an appropriate shape for the probability density function, nonparametric methods using kernel density estimation to model background distributions have been proposed. These methods build a histogram of background values for each pixel, by collecting values sampled from the pixels recent time window [11]. In [12], a Bayesian framework which incorporates multiple types of features for modeling complex backgrounds is proposed (we abbreviate the foreground detection method in [12] as FGD for short) and solves the sudden once-off background change effectively.

For all those background subtraction methods that have been mentioned above, they still have a common problem. That is how to select an adaptive learning rate. Explicitly, background modelling methods with a global empirical learning rate are significantly penalized. Over these years, a lot of research papers discussed the adaptive learning rate and proposed various solutions for tuning the learning rate based on local intensity changes [13, 14], different level feedbacks [15, 16], and so on. Based on GMM, [17] proposed a background subtraction method using a pixel-wise adaptive learning rate for object tracking. Unlike the traditional methods that use the same experiential “learning rate,” it assigns a learning rate to each pixel relying on two parameters; one is depending on the difference of pixel intensities between the background model and the current frame and the other is depending on the duration of the pixel being classified as a background pixel. In [18], the learning rates for the mean and the variance terms are decoupled and independent so as to avoid the saturation phenomenon and degeneracy problem. They use an adaptive learning rate to update the mean and a semiparametric model for the variance. The authors of [19] use the time gap between moving and stopped objects to train the background model and get adaptive parameters for urban traffic video. Considering the slow learning problem of GMM at the beginning phase, Kaewtrakulpong and Bowden improved the update mechanism in learning step and proposed the fast-learning Gaussian mixture model [20]. In [21], the enhanced Gaussian mixture model detects still objects from moving state and adjusts learning rate to improve the performance of detecting moving object detection with intermittent stops. The authors of [22] modulated the learning rate of background model based on scene activity. In [23], an updating method with adaptive learning rate (we abbreviate this method as GMMX for short) is proposed to accurately segment the objects that move slow or stop for a while during moving.

Even though many background subtraction approaches with adaptive learning rates were proposed and indeed improved the naive GMM, as mentioned above, they still have some limits and are not proper for foreground detection at the intersection. Especially, we find that many stopped cars gradually fuse into background and cannot be traced again with the previous methods. Firstly, most of these methods [14, 15, 17, 18, 21–23] perceive sudden change only based on the image information, such as illumination changes and
Figure 1: (a) shows several stopped vehicles during red light. (b) is the corresponding detection results. As time passed, some of the foreground objects disappeared gradually.

3. Our Approach

To the best of our knowledge, there have not been any methods that utilize the traffic light signal to enhance the vision-based background subtraction at the intersection. This paper proposes to regard these similar physical world signals as the criteria to adjust background model parameters. There are several parameters in background modelling methods. Learning rate $\alpha$ which controls the updating speed of modelling is a more important parameter. The stopped foreground objects will not fuse into background by adjusting $\alpha$ according traffic light signal. Then, we use GMM [8], GMMX [23], and FGD [12] as experimental subjects and traffic light signals as external information. How these methods are enhanced by our approach is described elaborately in the following text. Note that our method can work with other existing background subtraction approaches applied in intersection scene as well.

3.1. The Proposed Framework. In Figure 2, it depicts the whole proposed framework of using traffic signal to enhance the vision-based background subtraction. Model initialization, the first step, assigns all parameters needed and initials the background model. Then, the system gets a new input image from the video sensor networks and simultaneously receives physical world signals. Next module adjusts learning rate according to the traffic light signals. The following step is background subtraction, which is the same as the previous methods. Then, it outputs the foreground detection results. Meanwhile, the system updates background model and waits for the next input image. The grey modules are newly added in the framework, which distinguish from previous methods.

As wireless sensors become widely available and their costs come down, traffic control systems integrate with wireless sensor networks (WSN) in ITS [24]. Our method is designed to obtain traffic light signals from WSN. For convenience, we set the traffic light signal manually in experimenting. The learning rate is adjusted according to traffic light signal. When it is red, a little constant is selected. And once the light turns green, a normal one is applied. Here, the two values are all determined empirically. The following content specifically describes the improved methods.

3.2. Improved GMM Modelling. In the context of a traffic surveillance system, Friedman and Russell [25] proposed
**Figure 2:** Framework of the proposed background subtraction.

![Diagram](https://example.com/diagram.png)

...to model each background pixel using a mixture of three Gaussians corresponding to road, vehicle, and shadows. The maintenance is made by using an incremental EM algorithm for real-time consideration. Stauffer and Grimson [8] generalized this idea by modelling the recent history of the color features of each pixel \( \{X_1, \ldots, X_t\} \) by a mixture of \( K \) Gaussians.

The intensity in the RGB colour space of each pixel is selected as the feature to classify. The probability of observing the current pixel value is considered given by the following formula in the multidimensional case:

\[
P(X_t) = \sum_{i=1}^{K} \omega_{ij} \eta(X_t, \mu_{ij}, \Sigma_{ij}),
\]

where the parameters are the number of Gaussian \( K \), a weight \( \omega_{ij} \) associated to the \( i \)th Gaussian at time \( t \) with mean \( \mu_{ij} \), and standard deviation \( \Sigma_{ij} \). \( \eta \) is a Gaussian probability density function

\[
\eta(X_t, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{(-1/2)(X_t - \mu)^T \Sigma^{-1} (X_t - \mu)}.
\]

For computational reasons, Stauffer and Grimson [8] assumed that the RGB colour components are independent and have the same variances. So, the covariance matrix is of the form

\[
\Sigma_{ij} = \sigma_{ij}^2 I.
\]

The \( K \) Gaussians are sorted in descending following the ratio \( \omega_{ij}/\sigma_{ij} \). The first \( B \) Gaussian distributions which exceed certain threshold \( T \) are retained for a background distribution

\[
B = \text{arg min}_b \left( \sum_{i=1}^{b} \omega_{ij} > T \right).
\]

The others are regarded as foreground distribution. When the new frame comes at time \( t+1 \), a match test is made for each pixel. And a pixel matched a Gaussian distribution if

\[
\left( \frac{1}{b} \sum_{i=1}^{b} \frac{(X_{t+1} - \mu_{ij})^T \Sigma_{ij}^{-1} (X_{t+1} - \mu_{ij})}{k \sigma_{ij}} \right)^{1/2} < k \sigma_{ij}.
\]

When a match is found with one of the \( K \) Gaussians, for the matched component, the update is done as follows:

\[
\omega_{ij,t+1} = (1 - \alpha) \omega_{ij,t} + \alpha, \\
\mu_{ij,t+1} = (1 - \rho) \mu_{ij,t} + \rho X_{t+1}, \\
\sigma_{ij,t+1}^2 = (1 - \rho) \sigma_{ij,t}^2 + \rho (X_{t+1} - \mu_{ij,t+1})(X_{t+1} - \mu_{ij,t+1})^T,
\]

where \( \alpha \) and \( \rho \) are two learning rates. For the unmatched component, only the weight is replaced by

\[
\omega_{ij,t+1} = (1 - \alpha) \omega_{ij,t}.
\]

When no match is found, the least probable distribution is replaced with a new one with initial parameters.

In our improved method, the learning rate is adaptively tuned in accordance with external physical world events. Once a new input image arrives, the system enquires traffic light to perceive environment changes and does some reasonable adjustments instantly to get the best effect. The learning rate is changed as follows:

\[
\alpha = \begin{cases} 
\alpha_{\text{red}}, & \text{when light is red} \\
\alpha_{\text{normal}}, & \text{when light is green}.
\end{cases}
\]

When light is green, \( \alpha \) is selected as \( \alpha_{\text{normal}} \), a constant used by the original methods. Once the light turns red, \( \alpha \) is correspondingly adjusted to \( \alpha_{\text{red}} \), a small constant depending on the duration time of red light. From Figure 2, we can see that only the steps of receiving signal and adjusting learning rate are added. So this idea, improving models by integrating traffic light signal, can be generalized on many other methods.

### 3.3. Improved GMMX Modelling.

The original GMM has many limitations, such as the number of Gaussians having to be predetermined, the need for good initializations, and the dependence of the results on the true distribution law which can be non-Gaussian and slow recovery from failures. To alleviate these disadvantages, numerous improvements have been proposed over the recent years. In this paper, we choose a method as the comparative experiment, which is abbreviated as GMMX [23] and has an outstanding performance on...
detecting temporarily stopped objects with adaptive learning rate.

The main contribution of paper [23] is a model number adaptive method to decrease the amount of computation and an updating method with adaptive learning rate to accurately segment the objects that move slow or stop for a while during moving (here we are only interested in the second method). The authors think the fixed learning rate causes the problem that moving objects stopping for a short time will rapidly be updated to the background model by the GMM. Thus, different learning rates should be assigned to different distributions. When a new match is found at time \( t \), the learning rate \( \alpha \) is changed as follows:

\[
\alpha_t = \max(\alpha_0, \omega_M \cdot \alpha),
\]

(9)

where \( \omega_M \) is the weight of the matched distribution. The value of \( \alpha \) should be higher than \( \alpha_s \), both of whom are constants. The reason for evaluating \( \alpha_t \) with the maximum of \( \alpha_0 \) and \( \omega_M \cdot \alpha \), rather than \( \omega_M \cdot \alpha \), is that, when \( \omega_M \) has a very low value, \( \alpha_t \) will almost be zero if being evaluated with \( \omega_M \cdot \alpha \). It will cause an object staying for a long time to be difficult to be updated into the background model.

Traffic light signal is still able to be combined with GMMX by changing \( \alpha_t \) as follows:

\[
\alpha_t = \begin{cases} 
\min(\alpha_{\text{red}}, \max(\alpha_0, \omega_M \cdot \alpha)), & \text{when light is red} \\
\max(\alpha_0, \omega_M \cdot \alpha), & \text{when light is green}, 
\end{cases}
\]

(10)

where \( \alpha_{\text{red}} \) is a little constant to prevent the objects from fusing into the background in accordance with external signals. Because the red light may last for dozens of seconds, the value of \( \alpha_t \) should not be beyond \( \alpha_{\text{red}} \). Otherwise, the method of GMMX also results in the disappearing of stopped objects, especially at a later stage when the matched component has a large \( \omega_M \). Thus, external signals are also needed to let the output be more accurate.

3.4. Improved FGD Modelling. Li et al. proposed to classify background and foreground pixels under the Bayes decision theory [12]. Let \( V_i \) be a discrete value feature vector extracted from an image sequence at the pixel \( s = (x, y) \) and time instant \( t \). According to the Bayes rule, a posterior probability of \( V_i \) from the background \( b \) or foreground \( f \) is

\[
P(C | V_i, s) = \frac{P(V_i | C, s) P(C | s)}{P(V_i | s)}, \quad C = b \text{ or } f.
\]

(11)

Using the Bayes decision rule, a pixel \( s \) is classified as background according to its feature vector \( V_i \), observed at time \( t \) if

\[
P(b | V_i, s) > P(f | V_i, s).
\]

(12)

Note that the feature vectors associated with the pixel \( s \) are either from background or from foreground objects, and it follows that

\[
P(v_i | s) = P(v_i | b, s) \cdot P(b | s) + P(v_i | f, s) \cdot P(f | s).
\]

(13)

Substituting (11) and (13) into (12), it becomes

\[
2P(v_i | b, s) \cdot P(b | s) > P(v_i | s).
\]

(14)

In this method, the colours of a pixel are chosen as the feature for stationary background, while the colour cooccurrences of interframe changes from the pixel are chosen as the feature for moving background. And a table of statistics for the possible principal features is established for each feature type at \( s \), which is denoted as

\[
S_{ij} = \begin{cases} 
P_{ij}^b = P(v_i | s) \\
P_{ij}^f = P(v_i | b, s) \\
v_i = [a_1, \ldots, a_i]^T.
\end{cases}
\]

(15)

For each feature vector \( v_i \), that is used to classify a pixel as foreground or background, the statistics of the corresponding features (colour or colour cooccurrence) is updated by

\[
\begin{align*}
p_{ij}^{b(t+1)} &= (1 - \alpha_t) p_{ij}^{b(t)} + \alpha_t M_{ij}^{s(t)}, \\
p_{ij}^{f(t+1)} &= (1 - \alpha_t) p_{ij}^{f(t)} + \alpha_t M_{ij}^{t(t)}, \\
p_{ib}^{t+1} &= (1 - \alpha_t) p_{ib}^{t} + \alpha_t (M_{ib}^{s(t)} \wedge M_{ib}^{f(t)}),
\end{align*}
\]

(16)

where \( \alpha_t \) is the learning rate which controls the speed of feature learning. \( M_{ij}^{s(t)} = 1 \) when \( s \) is labelled as the background at time \( t \); otherwise, \( M_{ij}^{s(t)} = 0 \). \( M_{ib}^{s(t)} = 1 \) when \( v_i \) of \( S_{ij}^{s(t)} \) in (15) matches \( v_i \) best and \( M_{ib}^{f(t)} = 0 \) for the remainders. A reference background image that represents the most recent appearance of the background is maintained at each time to make the background difference accurate. If \( s \) is detected as a point of insignificant change in change detection, the reference background image is updated as

\[
B(s, t + 1) = (1 - \alpha_t) B(s, t) + \alpha_t I(s, t).
\]

(17)

Also, traffic light signal is added in FGD just as GMM in Figure 2. The external signals are mainly used to adjust feature learning rate \( \alpha_2 \)

\[
\alpha_2 = \begin{cases} 
\alpha_{\text{red}}, & \text{when light is red} \\
\alpha_{\text{normal}}, & \text{when light is green}.
\end{cases}
\]

(18)

When the light is red, a little constant is set to avoid the stopped cars vanishing quickly. Once the signal changes, \( \alpha_2 \) is always set as a suitable value. This method not only ensures the robustness of foreground detection but also eliminates the missing of slow-moving or stopped vehicles in the results.

4. Experiment Results

In order to make the experiment more convincing, we record the video at a real traffic junction, where there is a traffic light. Then, we run the programs of GMM, GMMX, FGD, and their enhanced versions, collect, and analyse the executive outcomes.
Figure 3: (1)–(4) are four original frames selected randomly from the test video when the light is red. These images show us five cars that slow down and stop successively. The other images are the foreground detection results of three original methods and corresponding improved methods. (a)–(d) are the results of GMM, while (e)–(h) are the results of improved GMM. (i)–(l) are the results of GMMX, while (m)–(p) are the results of improved GMMX. (q)–(t) are the results of FGD, and the images of the last row are the results of improved FGD.
The learning rate is set to 0.005 and follows. In GMM, we choose all of the experiments. Beyond our paper, we do not perform removing shadow in the entire sequence, but removing shadow is possible. We can see that strong shadows casted by moving vehicles can be observed in the horizontal axis is the frame number, while the vertical axis is the FN value (the number of foreground pixels that are wrongly marked as background). Different colour represents different method.

4.1. Data and Qualitative Results. The test video, which consists of 2752 frames of 640 * 480 pixels and is acquired at a frequency of 25 fps (frames per seconds), is taken from a busy intersection with traffic light. From the original images, we can see that strong shadows casted by moving vehicles can be observed in the entire sequence, but removing shadow is beyond our paper. We do not perform removing shadow in all of the experiments.

The detailed selection of important parameters is as follows. In GMM, we choose \( K = 5, T = 0.6, \) and \( k = 2.5.\) The learning rate is set to 0.005 and \( \sigma \) is initialized with 30. Initial weight associated to each Gaussian is 0.05. To make the experiment more comparable, GMMX has the same values of parameters with GMM. In their enhanced versions, when light is red, the learning rate is changed to 0.0001. In FGD, we choose \( \alpha \) = 2.5 and \( \beta \) = 0 for \( i = 1, \ldots, N_2 \) and \( v_i = \{c_i, c_i\}.\) The same way, improved FGD uses the same values of parameters with FGD. Only when light is red, \( \alpha_c \) is set to 0.0001.

In Figure 3, the first row is four original frames selected randomly from the test video sequence, which are frames 2122, 2256, 2381, and 2417, respectively. These pictures show us five cars that slow down and stop successively. Then, the images below are the detection results of three initial approaches and their enhanced versions. We can see that the enhanced methods are successful to keep the stopped vehicles in the detection results, while the outputs of previous methods show that the cars in front of every image have fused into background. Our approach obviously improves the effect of detection.

4.2. Quantitative Evaluation. In the evaluation, we choose 12 frames from the segment of the test video, during which the light is red and lasts about 25 seconds. In other words, a frame is selected in every 2 seconds. We use these images to analyse the performance of our method.

Three terms are used in the quantitative evaluation: false positive (FP) is the number of background pixels that are wrongly marked as foreground; false negative (FN) is the number of foreground pixels that are wrongly marked as background; total error (TE) is the sum of FP and FN. We calculate these terms for each image according to the corresponding hand-segmented ground truth. FN, FP, and TE of every approach are the sum of four frames of FN, FP, and TE.

Figure 4 illustrates overall performance on the selected twelve frames for the three previous methods and their improvements. The total error of the improved versions is less than the previous ones. In particular, the new approach reduces FN vastly, which means that the stopped cars would not disappear before the light turns into green. So, the tracking of foreground objects will not be interrupted, which supplies a solid foundation for many high-level images processing. But it is a pity that FP increases a little and it is our next problem to be solved.

In Figure 5, FN of different methods is listed according to the frame number. The horizontal axis is the frame number, while the vertical axis is the FN value (the number of foreground pixels that are wrongly marked as background). And different colours represent different methods. From Figure 5, we can see that FN of previous methods increases rapidly, which means that the foreground objects disappear from the detection results, while FN of the improved ones stays in a low quantity. This phenomenon is more obvious in the latter period of red light.

In summary, we conclude that our method can effectively improve the accuracy and reliability of foreground segmentation. Meanwhile, various background modelling methods are able to benefit from physical world signals.

5. Conclusions

We present a novel method that utilises traffic light signal to enhance the performance of background subtraction, while...
existing methods use only image information to model and update the reference background. Then, this paper records elaborately the experimental process and results. By contrast with FN, FP, and TE of the previous methods, such as GMM, GMMX, and FGD, the responding enhanced versions obviously have a better performance. It demonstrates that background subtraction methods based on traffic light signal may have a bright future. Considering that different pixels have different characteristics of changes in colour, it is better to set different and more reasonable learning rates based on these signals for each pixel rather than some constants for all pixels, which is our future research focus. Meanwhile, in order to combine the background modelling methods with physical world signals more closely, we will think over more relations between model parameters and these signals to get a better effect.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was supported by the fund of the State Key Laboratory of Software Development Environment (Grant no. SKLSE-2012ZX-01) and the Fundamental Research Funds for the Central Universities.

References

[1] M. M. Trivedi, I. Mikic, and G. Kogut, “Distributed video networks for incident detection and management,” in Proceedings of the IEEE Intelligent Transportation Systems, pp. 155–160, October 2000.
[2] M. M. Trivedi, T. L. Gandhi, and K. S. Huang, “Distributed interactive video arrays for event capture and enhanced situational awareness,” IEEE Intelligent Systems, vol. 20, no. 5, pp. 58–65, 2005.
[3] T. A. C. Bragatto, G. I. S. Ruas, V. A. P. Benso et al., “A new approach to multiple vehicle tracking in intersections using harris corners and adaptive background subtraction,” in Proceedings of the IEEE Intelligent Vehicles Symposium (IV '08), pp. 548–553, June 2008.
[4] C. Zhang, S. C. Chen, M. L. Shyu et al., “Adaptive background learning for vehicle detection and spatio-temporal tracking,” in Proceedings of the 4th Pacific Rim Conference on Multimedia Information, Communications and Signal Processing, vol. 2, pp. 797–801, 2003.
[5] R. Ding, X. Liu, W. Cui, and Y. Wang, “Intersection foreground detection based on the cyber-physical systems,” in Proceedings of the IET International Conference on Information Science and Control Engineering, pp. 1881–1886, Shenzhen, China, 2012.
[6] M. Hedayati, W. M. D. W. Zaki, and A. Hussain, “A qualitative and quantitative comparison of real-time background subtraction algorithms for video surveillance applications,” Journal of Computational Information Systems, vol. 8, no. 2, pp. 493–505, 2012.
[7] T. Bouwmans, F. El Baf, and B. Vachon, “Background modeling using mixture of Gaussians for foreground detection—a survey;” Recent Patents on Computer Science, vol. 1, no. 3, pp. 219–237, 2008.
[8] C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR ’99), pp. 246–252, June 1999.
[9] Y.-J. X. Hou and S.-R. Gong, “Adaptive shadows detection algorithm based on Gaussian Mixture Model,” in Proceedings of the International Symposium on Information Science and Engineering (ISISE ’08), pp. 116–120, December 2008.
[10] T. Bouwmans and F. Baf, “Statistical background modeling for foreground detection: a survey,” Handbook of Pattern Recognition and Computer, vol. 4, pp. 181–199, 2010.
[11] A. Tavakkoli, M. Nicolescu, G. Babis, and M. Nicolescu, “Non-parametric statistical background modeling for efficient foreground region detection,” Machine Vision and Applications, vol. 20, no. 6, pp. 395–409, 2009.
[12] L. Li, W. Huang, I. Y.-H. Gu, and Q. Tian, “Statistical modeling of complex backgrounds for foreground object detection;” IEEE Transactions on Image Processing, vol. 13, no. 11, pp. 1459–1472, 2004.
[13] Y.-H. Li, H.-F. Tian, and Y. Zhang, “An improved Gaussian mixture background model with real-time adjustment of learning rate,” in Proceedings of the International Conference on Information, Networking and Automation (ICINA ’10), pp. V1512–V1515, October 2010.
[14] M. Shah, J. Deng, and B. J. Woodford, “Localized adaptive learning of Mixture of Gaussians models for background extraction,” in Proceedings of the 25th International Conference of Image and Vision Computing New Zealand (IVCNZ ’10), November 2010.
[15] A. Pnevmatikakis and L. Polymenakos, “2D person tracking using Kalman filtering and adaptive background learning in a feedback loop,” Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 4122, pp. 151–160, 2007.
[16] H.-H. Lin, J.-H. Chuang, and T.-L. Liu, “Regularized background adaptation: a novel learning rate control scheme for gaussian mixture modeling,” IEEE Transactions on Image Processing, vol. 20, no. 3, pp. 822–836, 2011.
[17] K. K. Ng and F. J. Delp, “Background subtraction using a pixel-wise adaptive learning rate for object tracking initialization,” in Proceedings of the Visual Information Processing and Communication II, January 2011.
[18] P. L. M. Boutefroy, A. Bouzerdoum, S. L. Phung, and A. Beghdadi, “On the analysis of background subtraction techniques using Gaussian mixture models,” in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP ’10), pp. 4042–4045, March 2010.
[19] P.-S. Hwang, K.-Y. Eom, J.-Y. Jung, and M.-H. Kim, “A statistical approach to robust background subtraction for urban traffic video,” in Proceedings of the 2nd International Workshop on Computer Science and Engineering (WCSE ’09), pp. 177–181, October 2009.
[20] P. KaewTraKulPong and R. Bowden, “An improved adaptive background mixture model for real-time tracking with shadow detection,” Video-Based Surveillance Systems, pp. 135–144, 2002.
[21] H. Fu, H. Ma, and A. Ming, “EGMM: an enhanced Gaussian mixture model for detecting moving objects with intermittent stops,” in Proceedings of the 12th IEEE International Conference on Multimedia and Expo (ICME ’11), July 2011.
[22] M. Harville, G. Gordon, and J. Woodfill, “Adaptive video background modeling using color and depth,” in Proceedings
[23] P. Suo and Y. Wang, “An improved adaptive Background modeling algorithm Based on Gaussian Mixture model,” in Proceedings of the 9th International Conference on Signal Processing (ICSP ‘08), pp. 1436–1439, October 2008.

[24] M. Tubaishat, Z. Peng, Q. Qi, and S. Yi, “Wireless sensor networks in intelligent transportation systems,” Wireless Communications and Mobile Computing, vol. 9, no. 3, pp. 287–302, 2009.

[25] N. Friedman and S. Russell, “Image segmentation in video sequences: a probabilistic approach,” in Proceedings of the 13th conference on Uncertainty in artificial intelligence (UAI ’97), pp. 175–181, San Francisco, Calif, USA, 1997.