Adaptive Background Reconstruction for Street Surveillance

N. A. Zainuddin, Y. M. Mustafah, A. A. Shafie, M. A. Rashidan, and *N. N. A. Aziz*
International Islamic University Malaysia,
Kuala Lumpur, Malaysia
e-mail: yasir@iium.edu.my

***Abstract*— In recent years, adaptive background reconstruction works have found interest in many researchers. However, the existing algorithms that have been proposed by other researchers still in the early stage of development and many aspects need to be improved. In this paper, an adaptive background reconstruction is presented. Past pixel observation is used. The proposed algorithm also has eliminated the need of the pre-training of non-moving objects in the background. The proposed algorithm is capable of reconstructing the background with moving objects in video sequence. Experimental results show that the proposed algorithms are able to reconstruct the background correctly and handle illumination and adverse weather that modifies the background.**

Keywords—adaptive background reconstruction, background subtraction, mode filtering, frame differencing

I. INTRODUCTION

Background subtraction is a common technique used by many researchers in the field of computer vision. Background reconstruction is one of the main processes in background subtraction. A good background reconstruction is essential in order to produce an effective background subtraction. The simplest method of utilizing background subtraction is by setting the reference image as the image without any moving object and it is known in advance. However, this assumption of reference image is too idealistic, especially in outdoor environments. Consider a crowded street with vehicles and pedestrians. An outdoor environment like street will experience various levels of illumination at different times of the day. It has had the tendency of experiencing adverse weather condition like fog, rain or strong wind that able to modify the reference image. In these cases, the background must be adaptive refreshed.

In this paper, background reconstruction technique is presented. In recent years, the research on the background reconstruction has found interest to many researchers. Thus, proposed methods of the background reconstruction is rising in number. There are four sections in this paper. In section two, previous works related to the adaptive background reconstruction are described. Meanwhile, in section three, design of the proposed algorithm is presented. In section four, the experimental results at various scenes are presented. Finally, the conclusion is drawn in section five.

II. RELATED WORKS

Recently, a lot of work focus on adaptive background reconstruction. Thus, several methods are introduced. The methods are temporal smoothing, pixel intensity classification, running Gaussian average, Gaussian mixture model, hidden Markov model and also kernel density estimation.

The first method is the Gaussian mixture model (GMM). Gaussian mixture model is a common method used by researchers in the detection of foreground. The main idea of the Gaussian mixture model is by having N number of Gaussian distribution function, in order to reconstruct the background pixel. The implementation of GMM has been proposed as the adaptive background reconstruction technique, especially in surveillance system since late 1990s by Stauffer et. al [1]. As it is still in the early stage, a lot of unwanted small blobs are scattered all over the frame and the main blobs aren't completely foreground. Later, many of the researchers continued to improve the proposed algorithm. Mukherjee et all [2] have incorporated the Horpreset color model in order to make the system able to detect the shadow. Mukherjee et all [2] work focus more on the shadow handling. Plus, the experimental results show that the algorithm is tested for a non-complex indoor scene with one foreground handling. Nimse et all [3] managed to reconstruct the background with shadow consideration better with a more complex scene than Mukherjee et all [2]. They use the *train station* dataset in order to have multiple object handling. They have implemented the use of window-based decision rules for the shadow and the background model is reconstructed by GMM. To sum up, many implementations of GMM are focusing on the indoor scene for the surveillance system.

Elgammal et all [4] claimed that GMM is ideal for indoor scenes only. Thus, they introduced the use of kernel density estimation (KDE) for background reconstruction. In KDE method, the background and foreground pixel is model of probability density function (pdf) and the pdf is estimated by a *kernel* function or also known as *window function*. However, the major drawback of this method is the computational cost. Later Gao et. al [5] improved the method by introducing the Marr wavelet equation in estimating the pdf. Mar wavelet is a second derivative of the Gaussian smooth function. Thus, in simpler word, they were combining the Gaussian and Kernel method together. However, from the experimental result, the background model is not reconstructed correctly as some of the background pixels are counted as a target, thus, the target is not detected correctly. Lee et. al [6] managed to reduce the need of storage. He initialize the first frame and update it at every frame by setting the learning rate. For dynamic background cases, they used threshold method. All in all, Lee

978-1-4799-7635-5/14 $31.00 © 2014 IEEE
DOI 10.1109/ICCCE.2014.73

et. al used the first frame as the background model and then, update them by learning method. This method is only ideal for the video sequence at has no target or foreground the beginning of the video sequence. For the crowd and complex scene, it is quite impossible to get such video sequence in reality.

The other method that commonly used by researchers in reconstructing the background is temporal smoothing. The implementation of this method on the traffic scene has been discovered since 1993. The main idea of temporal smoothing is combining the stored image with the new image on pixel-based. Pixels on the smoothed image will be replaced by a part from previous value combined with a new value from new image at the same position. Ridder et. al [7] have improved the method by modelling each pixel with Kalman Filter. They managed to make the system more robust the illumination changes, however, the algorithm update the background slowly. Later, in recent year, Hung et. al [8] improved the method by combining with the median filtering. By doing so, they managed to reduce the computing frequency of median operations. However, they focus on the computational performances of the algorithm with no real life situation is considered in their experiment. Motivated by [8], Asif et. al [9] use the temporal smoothing and median filtering for modelling the background. As their research more focus on the human gait, the background and foreground are non-complex. Based on their experimental result, the temporal smoothing is ideal for the non-complex scene. Thus, surveillance system that involves non-complex scene (indoor surveillance) is ideal for the implementation of temporal smoothing algorithm.

In the early of 2000s, Hou et. al [10] have introduced the pixel intensity classification (PIC) method as the background reconstruction technique. In this approach, at every frame, the difference of pixel intensity is calculated. Then, the classification is made based on the difference. The background model is assumed to be the highest frequency in the intensity value. The simulation results of Hou et. al [10] used the outdoor dataset with parking lot environment. The works on the same method is continued by Xiao et. al in 2006 and 2008 [11] [12]. Xiao et. al managed to handle more adverse weather by adapting the raining situation, however, their experimental results shown that the adaption of the algorithm to the rain is still in the early stage. Next, Cao et al. [13] have employed the improved version of PIC in reconstructing the background model for light flow traffic movement video sequence. They managed to compare their experimental results with GMM and Time averaging algorithm. The results are ideal for slow moving to medium moving foreground detection. However, for high speed foreground, for instance, the highway traffic, some of the foreground is not detected and small blobs are appeared resulted from the problem of *ghost* or blending of high speed object movement.

After analyzing these methods, the assumptions are made. In the normal condition, where there is no adverse weather, the background pixel would be the maximum frequency in the image sequence. For adverse weather condition, most of the pixels in the image frame are expected to be modified. From the assumptions, we developed a background reconstruction algorithm based on pixel intensity classification and mode filtering. However, the difference in the inter-frame pixel intensity value is not calculated as the first step, the mode filtering is done first in order to model the background.

III. ADAPTIVE MODE FILTERING

In this paper, an adaptive background reconstruction is presented. Using the assumptions stated in section 2, the pixels are classified based on the intensity value in the image sequence. If the difference between the current frame and the reference frame is greater than the threshold value, the background needed to be reconstructed again indicating it must adapt to adverse weather condition or illumination problem. The method also updates the background in a specific time in order to increase the accuracy of the reconstructed background.

A. Patches

A video sequence consists of many image frames and an image frame is composed of a very big matrix. Pixel by pixel analysis is not preferable for real-time analysis. For the analysis of the video sequence, the input video sequence will be converted into several frames first; then, each of the image frame will be divided into several patches. Then, the analysis of the patches will be done in order to reconstruct the background model. The segmentation stage of the patches is shown in following figure.

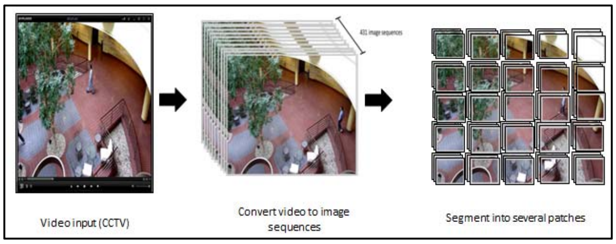


Figure 1. Segmentation to patches process

B. Background model reconstruction

After successfully segment the image to patches, the analysis is done to reconstruct the background model. The mode filtering is used as the main algorithm in this stage. In this approach, we select the most frequently occurring pixel in each of the patches as the background pixel. Example of implementation of the algorithm to a set of same location patches is shown in Figure 2.

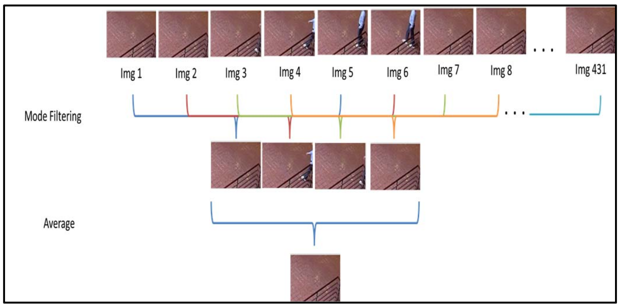


Figure 2. Background model reconstruction algorithm

The output of the each of the patches is a background patches at that location. Thus, combining all of the background patches will be produced a complete background model. Later on, based on the background model, the foreground is generated by using frame differencing technique.

IV. RESULTS AND DISCUSSION

Several video sequences are used in the experiment to evaluate the efficiency of the proposed algorithm. The proposed algorithm is tested for 3 public surveillance video sequences. There is a total of 3 video sequences that have been used which are bright, rain and busy. Bright video sequences show the public street traffic during daylight. Rain video sequences show the street during raining day, thus comparing the ability of the algorithm to handle adverse weather condition. The last video sequences show the street condition during the busy day, thus comparing the ability of handling overlapped moving objects. It is worth to note that all of the video sequences used in the experiments, are having moving objects for the entire video sequences. The results of the proposed algorithm are compared with the GMM and median filtering as both of the methods extensively mentioned in literature. The bright video sequences consist of 925 frames, the rain video sequences consist of 20236 frames and busy video sequences consist of 1170 frames.

A. Bright Video Sequence

Figure 3 shows the bright video sequence at 361th frame. The results of the background reconstruction are shown in Figure 4 (a), Figure 4(b), and Figure 4 (c) in the bright video sequence. The methods used are Median filtering, GMM and the proposed algorithm, respectively. For the bright scene, the algorithms are expected to handle shadow and edges. In Figure 4 (a) and Figure 4(b) we can see that the foreground objects produced are still intercepted with background pixels. GMM produced a better result than Median Filtering in term of shadow handling, however, some of the foreground shapes is not properly reconstructed. The car shapes in Figure 4 (b) are hardly recognized. Figure 4 (c) shows the results of our proposed algorithm. As the algorithm able to reconstruct the suitable background model, the foreground objects reconstructed is acceptable in term of shadow and edges.

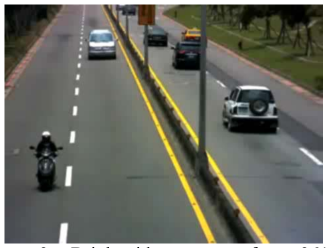
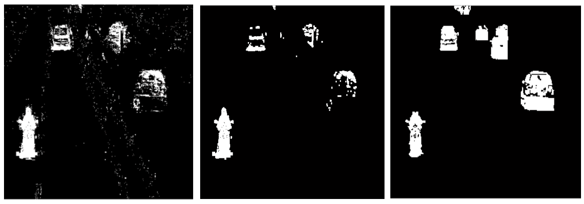


Figure 3. Bright video sequence frame 361th



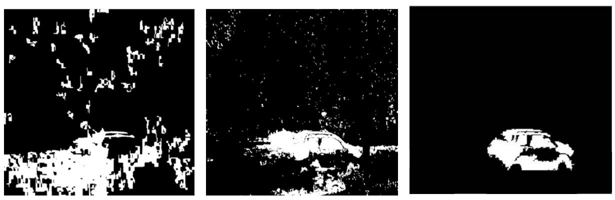
(a) (b) (c)

Figure 4. Output of proposed algorithm (bright) (a) Median Filtering (b) GMM (c) proposed method

B. Rain Video Sequence

Figure 3 shows the rain video sequence at 5751th frame. Figure 6(a), Figure 6 (b), and Figure 6 (c) are the results of the reconstruction of rain video sequence using Median filtering, GMM and the proposed algorithm, respectively. The main problem with rain video sequence is sudden lightning, strong wind, light from the car and gloomy environment. From Figure 6(a) we can see that median filtering cannot handle sudden illumination changes, thus resulting in poor result of foreground mask. From Figure 6(b) we can see that GMM perform better than median filtering, as some of the wind effect can be eliminated, however, it failed to overcome the light from the car. Figure 6(c) shows the results of our proposed algorithm. We can see that the algorithm managed to eliminate the sudden illumination and strong wind.

Figure 5. Rain video sequence frame 5751th



(a) (b) (c)

Figure 6. Output of proposed algorithm (rain) (a) Median Filtering (b) GMM (c) proposed method

C. Busy Video Sequence

Figure 7 shows the busy video sequence at 252th frame. Figure 8(a), Figure 8(b), and Figure 8(c) are the results of the reconstruction of busy video sequence using Median filtering, GMM and the proposed algorithm, respectively. From Figure 8(a), we can see that the cars overlap each other in the foreground mask due the shadow elimination of the shadow. From Figure 8(b) we can see that GMM perform better than median filtering as some of the shadows can be eliminated, however, some of the foreground pixel is classified as background pixel, thus some of the foreground is missing in the output foreground mask. Figure 8(c) shows the results of our proposed algorithm. We can see that the overlapping of the foreground is not severe. There are several outlines that can identify the border of the overlapped.

Figure 7. Busy video sequence frame 252th

(a) (b) (c)

Figure 8. Output of proposed algorithm (busy) (a) Median Filtering (b) GMM (c) proposed method

V. CONCLUSIONs

From the results of the experiments, it shows that the proposed algorithm able to reconstruct the background successfully and there is no need to have the non-moving object in the video sequence to reconstruct the background. Implementing the proposed algorithm in various condition of the scene shows that it is capable of adapting the adverse weather condition and still produces a good result for busy road. Therefore the proposed algorithm in this paper is efficient.

ACKNOWLEDGMENT

This work is supported by the Fundamental Research Grant Scheme (FRGS) Research Project FRGS13-030-0271 under the Ministry of Higher Education of Malaysia (MOHE). This work is also a collaboration between IIUM and Royal Malaysia Police (RMP).

REFERENCES

[1] C. Stauffer and W. E. L. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," in *IEEE Computer Society Conference*, Colorado, 1999.

[2] S. Mukerjee and K. Das, "An Adaptive GMM Approach to Background Subtraction for Application in Real Time Surveillance," *International Journal of Research in Engineering and Technology,* vol. 2, no. 1, pp. 25-29, 2013.

[3] M. Nimse, S. Varma and S. Patil, "Shadow Removal Using Background Subtraction and Reconstruction," *International Journal of Emerging Technology and Advanced Engineering,* vol. 4, no. 4, pp. 324-327, 2014.

[4] A. Elgammal, R. Duraiswami, D. Harwood and L. Davis, "Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance," *Proceeding of the IEEE,* vol. 90, no. 7, pp. 1151-1162, 2002.

[5] T. Gao, Z. G. Liu, W. C. Gao and J. Zhang, "A Robust Technique for Background Subtraction in Traffic Video," in *International Conference in Neuro-Information Processing*, Auckland, New Zealand, 2009.

[6] J. Lee and M. Park, "An Adaptive Background Subtraction Method Based on Kernel Density Estimation," *Journal on the Science and Technology of Sensors and Biosensors,* vol. 12, no. 9, pp. 12279-12300, 2012.

[7] C. Ridder, O. Munkelt and H. Kirchner, "Adaptive Background Estimation and Foreground Detection Using Kalman Filter," in *Proceedings of International Conference on Recent Advances in Mechatronics*, Istanbul, Turkey, 1995.

[8] M. H. Hung, J. S. Pan and H. C. Hsieh, "A Fast Algorithm of Temporal Median Filter for Background Subtraction," *Journal of Information Hiding and Multimedia Signal Processing,* vol. 5, no. 1, pp. 33-41, 2014.

[9] S. Asif, A. Javed and M. Irfan, "Human Identification On the basis of Gaits Using Time Efficient Feature Extraction and Temporal Median Background Subtraction," *International Journal Image, Graphics and Signal Processing,* vol. 3, no. 2, pp. 35-42, 2014.

[10] Z. Hou and C. Han, "A Background Reconstruction Algorithm based on Pixel Intensity Classification in Remote Video Surveillance System," *Proceedings of the Seventh International Conference on Information Fusion,* vol. 2, 2004.

[11] M. Xiao, C. Han and X. Kang, "A Background Reconstruction for Dynamic Scenes," in *IEEEE 9th International Conference on Information Fusion*, 2006.

[12] M. Xiao and L. Zhang, "A Background Reconstruction Algorithm Based on Two-Threshold Sequential," in *International Colloquium on Computing, Communication, Control, and Management*, Guangzhou, 2008.

[13] L. Cao and Y. Jiang, "An Effective Background Reconstruction Method for Video Objects Detection," in *International Conference on Networking and Distributed Computing*, 2013.