Table 4 : Estimation Model Result

	Mandatory Male			Mandatory Female			Shopping Male			Shopping Female			Leisure Male			Leisure Female		
	param	t-value		param	t-value		param	t-value		param	t-value		param	t-value		param	t-value	
EXPLANATORY VARIABLE																		
Constant	0.563	0.982		-0.611	-0.922		-0.762	-1.577		0.094	0.170		-0.714	-1.417		-0.403	-0.627	
Ojek Satisfaction Level																		
Ojek Fare	0.082	0.777		-0.199	-1.402		0.007	0.080		-0.254	-2.177	**	-0.029	-0.322		-0.203	-1.532	
Ojek Travel Time	0.044	0.409		0.062	0.606		0.049	0.555		0.125	1.414		0.106	1.125		0.053	0.528	
Ojek Safety	-0.032	-0.259		0.159	1.171		0.083	0.835		0.030	0.279		0.134	1.243		-0.055	-0.446	
Ojek Security	-0.065	-0.546		0.017	0.192		-0.136	-1.405		-0.215	-2.767	**	-0.098	-0.947		-0.147	-1.630	
Angkot Satisfaction Level																		
Angkot Fare	-0.042	-0.389		0.297	2.094	**	0.019	0.213		-0.050	-0.440		-0.024	-0.260		0.082	0.633	
Angkot Travel Time	0.057	0.557		-0.086	-0.674		0.004	0.051		0.219	2.016	**	0.041	0.468		0.113	0.928	
Angkot Safety	-0.029	-0.211		0.006	0.045		0.053	0.490		-0.070	-0.662		-0.059	-0.525		-0.091	-0.755	
Angkot Security	-0.048	-0.364		-0.067	-0.880		-0.106	-1.006		-0.065	-0.999		-0.158	-1.443		-0.043	-0.570	
Socio-Demographic Attributes																		
Young	-0.162	-0.912		-0.257	-1.546		-0.012	-0.087		0.036	0.269		0.103	0.694		0.191	1.230	
Elder	-0.503	-1.865	+	0.082	0.192		0.027	0.120		0.579	1.725	+	-0.005	-0.021		0.027	0.064	
Household Income	-0.009	-0.214		0.093	2.090	**	0.037	1.055		0.040	1.120		-0.025	-0.683		0.059	1.452	
Education Level (university)	-0.143	-0.853		0.072	0.423		0.057	0.412		0.396	2.739	**	0.170	1.158		0.147	0.901	
License Car	-0.327	-1.746	+	-0.637	-2.732	**	-0.131	-0.845		-0.509	-2.524	**	-0.208	-1.248		-0.360	-1.537	
License Motor	1.362	7.458	**	0.377	2.223	**	0.714	4.496	**	0.474	3.294	**	0.506	3.025	**	0.400	2.469	**
Residential Informations																		
Distance Home To Work	-0.014	-2.618	**	0.013	2.494	**	0.002	0.372		-0.002	-0.565		0.001	0.172		-0.007	-1.383	
Live in Jakarta	-0.253	-1.495		-0.044	-0.259		-0.008	-0.058		-0.093	-0.673		0.144	0.990		0.131	0.826	
Correlation Parameter ($\rho_{1\varepsilon}$)	0.314	5.382		0.391	7.701		0.813	19.522		0.834	19.400		0.869	19.024		0.815	15.533	
Correlation Parameter ($\rho_{2\varepsilon}$)	0.560	10.739		0.487	9.799		0.880	16.471		0.898	15.405		0.704	10.968		0.795	9.778	
Initial Log Likelihood	-442.421			-479.028			-1437.526			-766.607			-876.184			-579.632		
Converge Log Likelihood	-235.192			-261.489			-380.123			-355.339			-317.284			-250.223		
McFadden Rho Square	0.468			0.454			0.736			0.536			0.638			0.568		
Mc Fadden Rho Square Adjusted	0.425			0.414			0.722			0.512			0.616			0.536		
Number of Samples	302			253			338			339			338			339		

AN IMPROVED MIXTURE OF GAUSSIAN MODEL FOR REAL TIME VEHICLE DETECTION

Boon Kit Wong
Postgraduate Student
Department of Civil Engineering
Universiti Tunku Abdul Rahman
(UTAR)
Email: wongboonkit@yahoo.com

Oon Ee Ng
Assistant Professor
Department of Electrical Engineering
Universiti Tunku Abdul Rahman
(UTAR)
Email: ngoe@utar.edu.my

Hooi Ling Khoo
Associate Professor
Department of Civil Engineering
Universiti Tunku Abdul Rahman
(UTAR)
Email: khoohl@utar.edu.my
(Corresponding author)

ABSTRACT

This paper proposes a novel method to segment video sequences which undergoes gradual changes into foreground and background layers. The background layer contains all objects which have been stationary since the beginning of the video sequence. The foreground layer contains objects which have entered into or move within the video scene and these objects can be moving or stationary. An improved and adaptive Mixture of Gaussian (MoG) model with a feedback mechanism algorithm has been formulated. The MoG model will classify every pixel in the image as belonging either the foreground or the background layer. Every object in the foreground layer will be tracked and updated in the MoG via the feedback mechanism. This feedback avoids stationary foreground objects being updated into the MoG and thus affecting the approximation done by the MoG. This algorithm has been implemented into an Intelligent Transportation System (ITS) to detect vehicles on the road in an outdoor environment. A promising result is obtained in extracting vehicles on the road.

Keywords: Intelligent Transportation System, vehicle detection, image processing

1. INTRODUCTION

One of the important aspects in designing an Intelligent Transportation Systems (ITS) is the robustness and reliability of the automatic detection algorithm. Real time foreground object detection and segmentation from video sequences is one of the essential tasks. However, the segmentation process becomes challenging and difficult when the video scene undergoes illumination changes, weather changes and camera noise. Many methods (Porikli et al., 2003 , Atrey et al. ,2006) have been proposed for real-time foreground segmentation and detection in video sequences.

1.1 Previous Work

The most commonly used foreground segmentation technique is background subtraction. Lipton et al. (1998) proposed a two frame differencing method which subtracts the current frame from a reference frame. The pixels where the difference exceeds a predefined threshold will be classified as foreground pixels. This method is followed by morphology operations to enhance the detected region and reduce camera noise. The reference frame is then updated with the current frame. Yang & Levine (1992) proposed to model the reference frame by taking the median value of pixel color over a series of images and defining the threshold value through histogram analysis. The related temporal differencing method extracts moving region by finding pixel wise difference between two of three consecutive frames from an image sequence (Collins, et al, 2000) (Wang and Brandstein,1998), (Saptharishi et al. 2000). However, these methods fail to detect moving objects which are temporally static and do a poor job in extracting the entire relevant feature (Jagdish et al., 2011).

Another class of approaches to extract moving regions from a video sequence are the statistical methods. Statistical method approaches generate advanced background models by utilizing individual pixel or group of pixels. The background model can be updated during processing. Each pixel in the current frame will be classified as either a foreground pixel or a background pixel by comparing the statistics of the current background model. Haritaoglu et al.(2000) proposed to build a statistical model by representing each pixel with minimum intensity values, maximum intensity value and the maximum intensity difference between consecutive frames during training period. The process of rebuilding the statistical model would be periodic. However, defining the period to update and rebuilt the statistical model proved problematic since the background might change gradually or instantly. McKenna et. al. (2000) proposed an adaptive background model which combined color and gradient information. The author then performed background subtraction to cope with shadows and unreliable color information. Porikli et al. (2003) proposed variable adaptation rates to update the statistical model. The authors introduced an illumination change score where the statistical model would be updated according to the score. A variable adaptation frame rate following the pixel's activity is proposed by Magee (2004). This method runs at high adaptation rate in the less stable region of the image, this indirectly reduces the computation time. Atrey et al. (2006) proposed to apply MoG on the region of interest (ROI). The experimental result shown when applying MoG on ROI was significant reduction in processing time with a minor loss of accuracy. Zuo et al. (2007) proposed a switching based background modeling approach (MBSM). In this method, regions were classified as regions with high or low complexity. For region with high complexity, the MoG model was used. Average value was applied for region with low complexity.

1.2 Our Approach

All of the techniques discussed in the previous section were able to segment moving object from background model. However, all of the methods proved unable to segment which are initially moving and then become static. This issue commonly arises in Intelligent Transportation Systems where vehicles will not move smoothly all the time. Vehicles might become static temporally then continue to move again. Hence, this paper introduces an adaptive detection method which is able to detect moving and static vehicles on the road by segmenting the image into a foreground layer and a background layer using an adaptive Gaussian Mixture Model with feedback mechanism.

2.   METHODOLOGY

This section will illustrate the proposed method. Figure 1 is the flow chart for the proposed method. The system starts with a video or camera source as input. A frame grabber will get images from source. To reduce the processing time, the frame grabber will resize or set the region of interest (ROI) to every single image grabbed from the source. A shaking correction module is used to provide Optical Image Stabilization to correct for the vibration of the image source. The background model is then used to classify every pixel in foreground estimation module. A blob detector is implemented at the following stage to extract all the foreground objects into individual blobs. A bounding box is created around each blob. If no blobs are detected, the whole background model will be updated. However if the blobs are detected, the blob information will be sent to the tracker. The tracker will track the foreground object until the object is out of the camera or video scene. If the same foreground object is detected in the consecutive image, the tracker will send information to background model. From there, the background model will only update selected regions according to the information sent by tracker. Chintalacheruvu. N. & Muthukumar. V., 2012 proposed to detect vehicle on the road by using the Harris-Stephen corner detection. The detected corner points will be tracked from one frame to another in a video stream. Sum of Square Difference (SSD) algorithm is used to track those corner points. In this paper, centroid of the detected corner points in a region of interest will be considered as a mark or vehicle. The detected corner points might not come from a vehicle, hence it is hard to justify the centroid points of those detected points is a vehicle. Płaczek, B., 2010 proposed to construct the static background model by evaluating the local image attributes instead of particular pixels. A linguistic fuzzy logic is implemented in the image evaluation process. This algorithm is

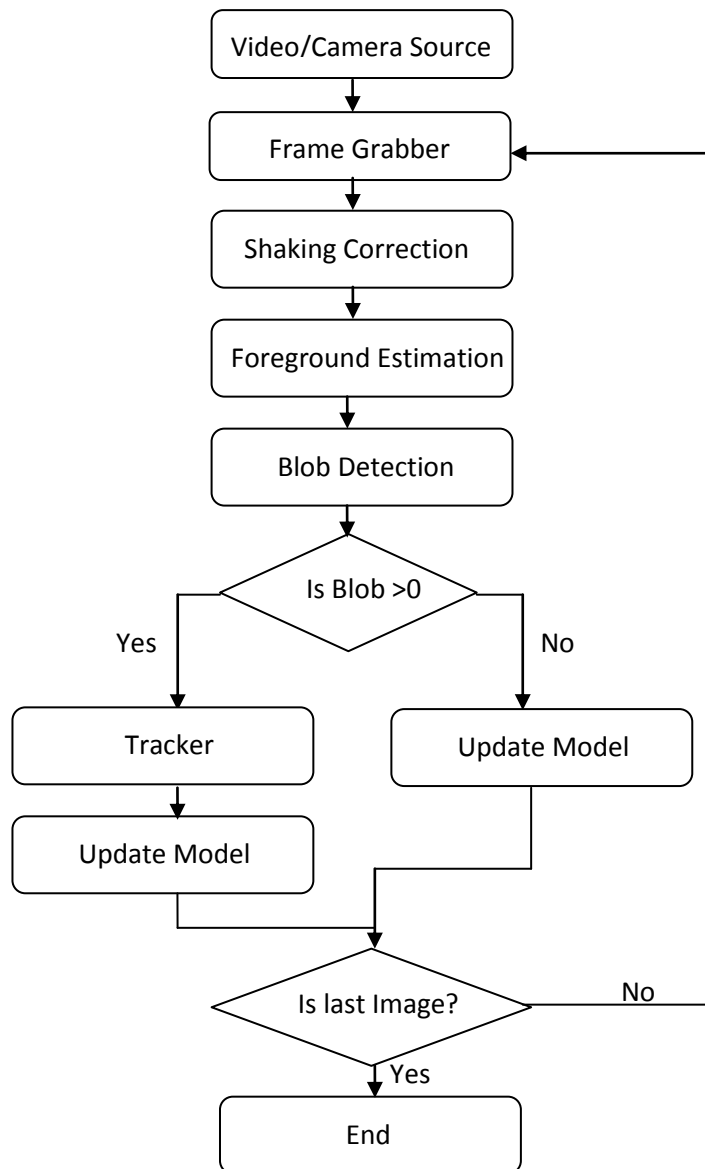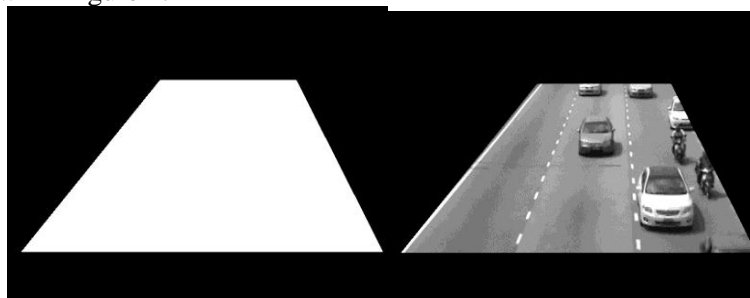unable to detect temporal static vehicle on the road.

Figure 1: Flow Chart

2.1 Frame Grabber

The frame grabber grabs images from video or camera source. In this module, image will be converted into grayscale before proceeding to the next phase. To reduce the processing time, an ROI mask is applied as shown in figure 2.

(a)   (b)

Figure 2(a) ROI mask (b) output image

2.2 Adaptive Mixture of Gaussian with feedback mechanism

According to Stauffer and Grimson(1998) method, each pixel is modeled separately by a mixture of K Gaussian where k is the number of components. This method will update the whole background model before the foreground is detected. During classification, all the pixels which fall inside those components will classify as background pixel. Hence, if the foreground object stops moving, the foreground pixels will eventually be classified as background pixels. This method also does a poor job in extracting all relevant feature pixels especially for objects with uniform color because some of the information gets updated into the background model.

In the proposed method, MoG with three components is implemented. Each component represents different layers in the image.

$$P(\mathrm{I}_t) = \sum_{i=1}^{k=3} \omega_i \eta(\mathrm{I}_t\ \mu_{i,t}\ \Sigma_{i,t}) \tag{1}$$

If the current pixel falls into any one of the components back on equation 1, the component will be updated according to equation 2. In this implementation, the standard deviation used is 2.0 (C.Stauffer and W.E.L. Grimson, 1998).

$$\omega_{i,t} = \omega_{i,t-1}$$
$$\mu_{i,t} = (1-\rho)\mu_{i,t-1} + \rho l_t \tag{2}$$
$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho(l_t - \mu_{i,t})^T(l_t - \mu_{i,t}))$$
$$\text{Where } \rho = \alpha P_\gamma(l_t\mid\mu_{i,t-1},\Sigma_{i,t}-1)$$

For components which do not match, the pixel will not be updated. If the pixel does not match any of the components, the weakest component will be replaced by a new component. In our implementation, the components are arranged in decreasing order, the strongest weight component will be the first component and the lowest weight component will be the third component. The component weight is calculated according to equation 3.

$$\omega_{i,t}/\|\Sigma_{i,t}\| \tag{3}$$

In our approach, the component with strongest weight will be classified as the background layer, unlike the method proposed in (Stauffer and Grimson, 1998). The remaining components will be classified as foreground layer. With such classification, a more complete set of relevant feature pixels can be extracted. Figure 3 shows the original image, background image and foreground image.
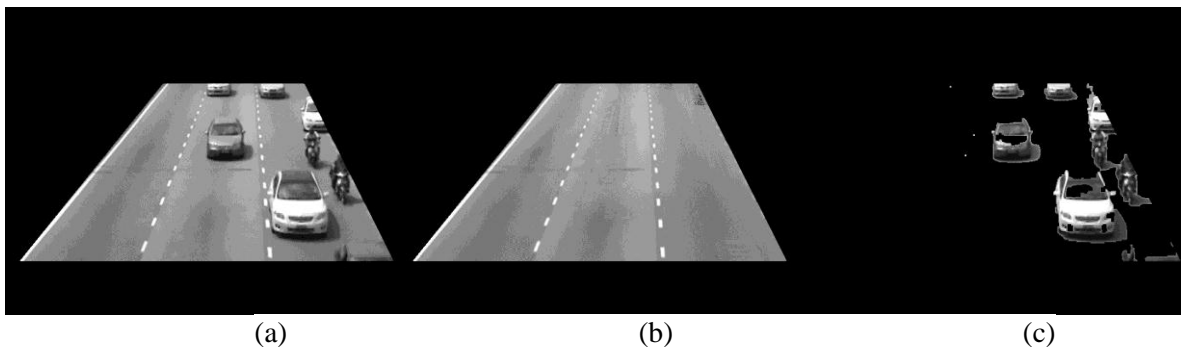
(a)   (b)   (c)

Figure 3(a) original image, 3(b) background layer, 3(c) foreground layer

According to Stauffer and Grimson(1998), the model is updated every time a new image is available and pixel classification will carried out right after that. In our approach, instead of updating the mixture model every time a new image is inserted, we carried out the foreground estimation first before updating the mixture model. The proposed method therefore separates the foreground estimation and update mixture model into two different operators. Additionally, a feedback mechanism is introduced in the update mixture module. Instead of updating the whole mixture model, the mixture model can do selective updates according to feedback mechanism. With such selective updating capability, the algorithm can define the regions which need to be updated to avoid foreground object being updated into the mixture model if the object is temporally static. The update mixture model accepts a binary mask as feedback information and does selective update according to following condition.

$$\text{if } p(x, y) == 1, \text{ then pixel will be updated}$$
$$\text{else if } p(x, y) == 0, \text{ then pixel will not be update}$$

(4)

2.3 Blob Detector

A blob detector is used to segment the foreground layer into individual blobs. Bounding boxes are created for each blob. The information of each blob is used in tracking the foreground object. All connected vehicles are considered as one blob. Figure 4(a) shows the input and (b) shows the detected blobs together with white bounding box.

(a)   (b)

Figure 4 (a) shows the foreground image and (b) shows the output image with white bounding box.

2.4 Tracker System

The tracker system will track all the vehicles entered into the camera/video scene. Each detected blob creates a tracker with blob information such as the template image, center of the blob and bounding box for the blob. If a similar blob is detected in the following scene, the tracker will updated with the new blob information. Template matching with cross correlation is implemented to find the similar blob in the scene. Normally, template matching is done using a sliding window over the entire image. The process of finding previously detected blobs might be costly if the image is large.

In this proposed method, tracker will have a searching regions represent by the circle in Figure 5. Any blobs which partially or fully overlap with the circle will be checked. The searching region is defined by parameter radius. The searching radius is currently defined as 45pixels. If the blob is similar to the tracker template image, then the tracker will stop the search and update the detected blob information. Finally, tracker system will send a feedback image which contains blobs to background model for

updating. Figure 6 shows a sequence of images with a 10 second interval. The result show that the same object is tracked and multiple objects can be tracked in single scene.
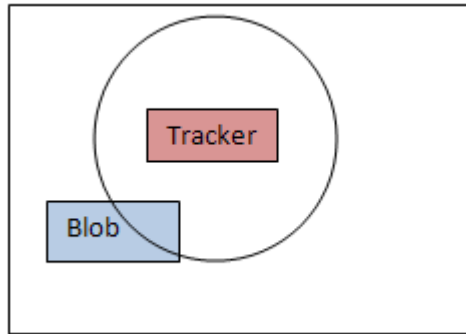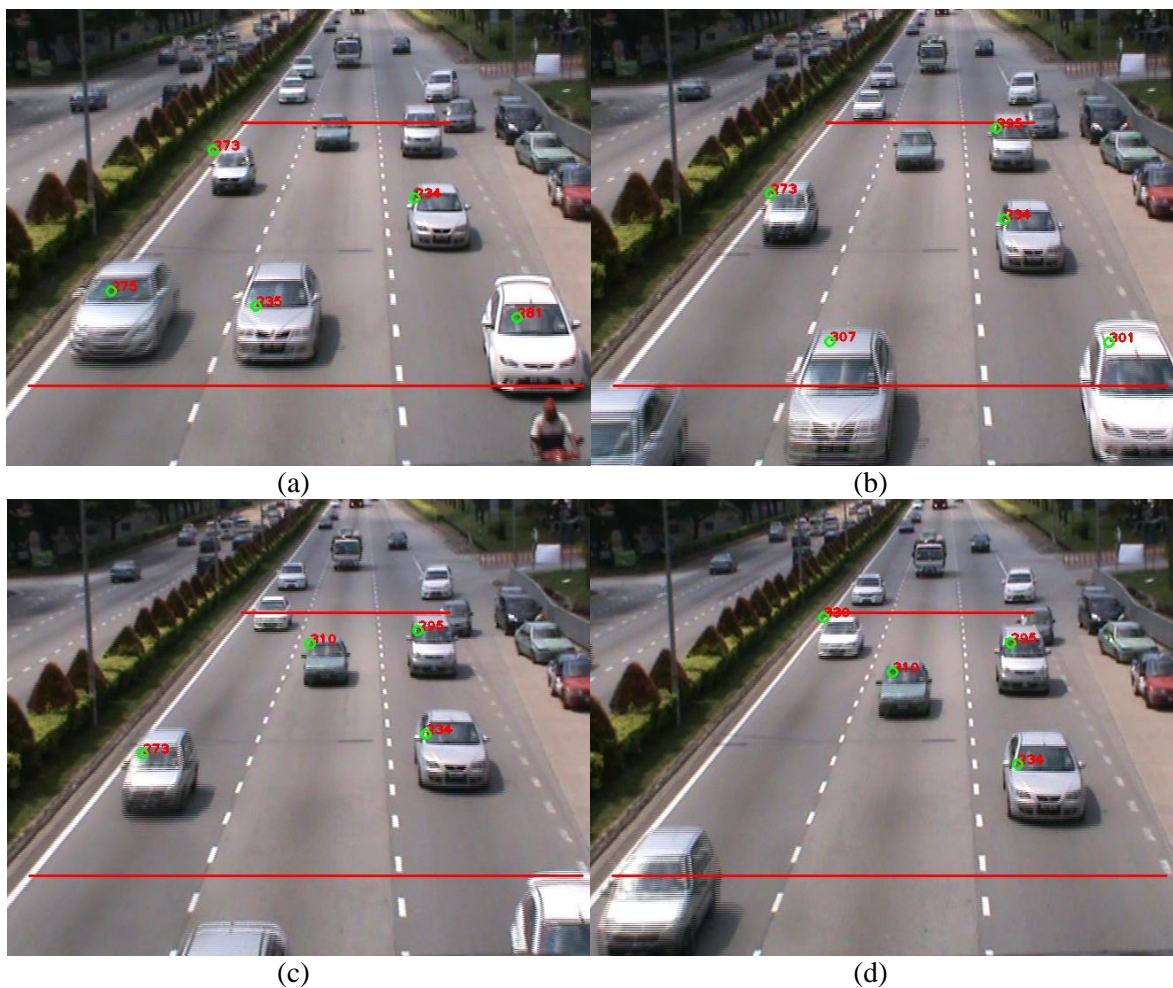

Figure 5 Tracker Searching Radius

(a)

(b)

(c)

(d)

Figure 6 Image sequence with tracked objects with interval 10 second each, the red line indicates the ROI for the proposed method.

3. EXPERIMENT RESULT

The proposed method is implemented into an intelligent vehicle system (IVS) to extract the data from the road. This method is able to segment the vehicle from the road.
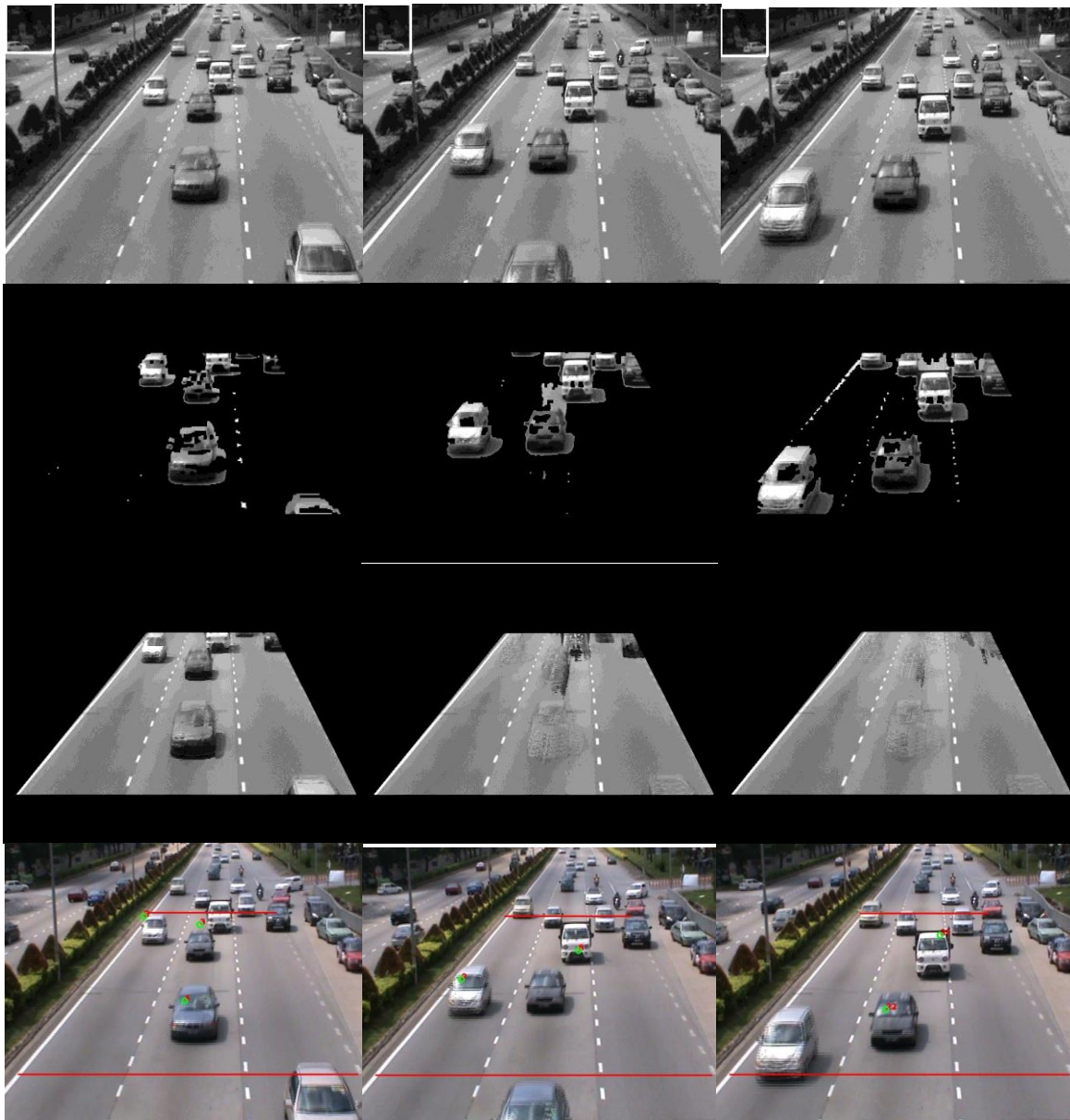
Figure 7: The top row displays the original sequence at frames 15, 35, and 55 respectively. The second row shows the foreground layer generated. The third row shows the generated background layer. The last row shows the output with tracked vehicles.

The proposed algorithm was tested on five difference roads and test conditions were complicated with slight camera vibrations, illumination changes. The shaking correction module implemented will reduce the camera vibrations. A false negative detection was registered if a vehicle was present in a detection zone and was not detected by the algorithm. A false positive detection (F1) was registered if no vehicle was present in a detection zone and detected by the algorithm. False positive detection (F2) was registered if two vehicles was closed to each other and detected as one vehicle. In this paper, the high F2 rate mainly due to blob detection implemented in this proposed method unable to segment vehicles which are closed together into individual blob.

Table 1 Experiment result on the proposed method.

Location	Accuracy (%)	False Positive (%)	False Negative(%)	
			F1	F2
	83.20	2.20	1.20	13.40
	84.60	1.75	0.90	12.75
	82.10	2.00	1.10	14.80
	84.30	3.3	1.05	11.35

4. CONCLUSION

The experimental results demonstrate that the proposed algorithm is a better way to segment images into foreground and background pixel both moving and temporal static object. This proposed method able to achieve accuracy more than 82% at medium or high traffic volume. The high false detection at medium or high traffic volume is due to vehicles which are closed together is detected as one.

5. REFERENCES

A. J. Lipton, H. Fujiyoshi, and R.S. Patil., 1998, *Moving target classification and tracking from real-time video*. In Proc. of Workshop Applications of Computer Vision, pages 129–136.

Bayona A., Juan C. S., Martinez J. M., 2009, *Comparative evaluation of stationary foreground object detection algorithms based on background subtraction techniques*

C. Anderson, P. Bert, G. Vander Wal,, 1985, *Change detection and tracking using pyramids Transformation techniques,*Proceedings of the SPIE-Intelligent Robots and Computer Vision, Vol. 579, pp. 72–78.d

Chintalacheruvu. N. & Muthukumar. V. 2012, *Video Based Vehicle Detection and Its Application in Intelligent Transportation Systems,* Journal of Transportation Technologies, 2012, 2, 305-314

C.Stauffer and W. E. L.,1999, Grimson: Adaptive *background mixture models for real-time tracking* in: Computer Vision and Pattern Recognition Fort Collins, Colorado pp. 246-252.

C. Wang, M.S. Brandstein, 1998, *A hybrid real-time face tracking system,* Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, Seattle, WA.

D.Neelima1, Gowtham Mamidisetti2,2012. *A Computer Vision Model for Vehicle Detection in Traffic*

I. Haritaoglu, D. Harwood, L.S. Davis, 2000, *Real-time surveillance of people and their activities*, IEEE Trans.Pattern Anal. Mach. Intell. 22 (8) 809–830.

JagdishLalRaheja, ChaitanyaMalireddy,Aniket Singh, L Solanki, 2011, *Electronic Computer Technology (ICECT)*, 3rd International Conference ,volume 2, page 199-203.

J.R. Bergen, et al., 1992, *A three frame algorithm for estimating two-component image motion*, IEEE Trans. Pattern Anal. Mach. Intell. 14 (9) 886–896.

Magee D., 2004, *Tracking multiple vehicles using foreground, Background and motion models*. Image and Vision Comput; 22: 143-155.

M. Saptharishi, J.B. Hampshire II, and P. Khosla., 2000, *Agent-based moving object correspondence using differential discriminative diagnosis*. In Proc. Of Computer Vision and Pattern Recognition, pages 652–658.

P. Atrey ,V. Kumar,A. Kumar ,M. Kankanhalli, 2006, *Experiential sampling based foreground/background segmentation for video surveillance*. IEEE IntConf on Multimedia and Expo (ICME), Toronto, Canada,1809-1812.

Porikli F., 2003, *Human body tracking by adaptive background models and mean-shift analysis*. IEEE Int Workshop on Performance Evaluation of Tracking and Surveillance (PETS).

Płaczek, B., 2010, *A real time vehicles detection algorithm for vision based sensors*, in: L. Bolc et al, (Eds.), Computer Vision and Graphics, Lecture Notes in Computer Science 6375, Springer-Verlag, Berlin Heidelberg, 2010, pp. 211–218

R.T. Collins, et al., 2000, *A system for video surveillance and monitoring: VSAM Anal report*, CMU-RI-TR-00-12, Technical Report, Carnegie Mellon University.

Song G.Y., Lee Y. K. & Lee J. W., 2008, V*ehicle Detection by Edge-based Candidate Generation and Appearance-based classification*

Sumer Jabri, Zoran Duric, HarryWechsler and Azriel Rosenfeld, 2000, *Tracking groups of people*, Computer Vision Image Understanding 80 (1) 42–56.

Tsukiyama and Shirai, 1985, *Detection of Movements of person from a Spare sequence of TV Images*;

Pattern Recognition 18(3/4):207-213.

W.E.L.Grimson,C.Stauffer, R. Romano and L. Lee, 1998 *Using adaptive tracking to classify and monitoractivities in a site in:* Computer Vision and Pattern Recognition Santa Barbara, California.

Y.H. Yang, M.D. Levine, 1992,*The background primal sketch: an approach for tracking moving objects Mach*. Vision Appl. 17–34.

Y.Ivanov,C.Stauffer, A. Bobick and W. E. L. Grimson, 1999 *Video surveillance of interactions in:* Second IEEE Workshop on Visual Surveillance Fort Collins, Colorado pp. 82-90.

Y. Kameda, M. Minoh, 1996, A *human motion estimation method using 3-successive video frames*, Proceedings of the International Conference on Virtual Systems and Multimedia.

Zuo J, Pan Q, Liang Y, Zhang H, Cheng Y.,2007 *Model switching based adaptive background modeling approach*. ActaAutomaticaSinica.