Vol. 5(14) Jan. 2015, PP. 1949-1962

Contents list available at IJMEC

International Journal of Mechatronics, Electrical and Computer Technology (IJMEC)

Journal Homepage: www.aeuso.org

PISSN: 2411-6173 - EISSN: 2305-0543

A Framework for Segmenting Moving Objects in Image Sequences Using Vector Quantization to Estimate the Background

Hassan Khastavaneh[1,2*], Habibollah Haron[2] and Siti Mariyam Shamsoddin[2]

[2]Faculty of Computing – UniversitiTeknologi Malaysia – Johor – Malaysia

***Corresponding Author's E-mail:** khassan3@live.utm.my

Abstract

Moving object segmentation is important in many computer vision applications. The goal of moving object segmentation is to classify pixels as foreground or background; the foreground pixels forming the moving objects. A good segmentation method should be able to do segmentation when the scene is complex as well as adaptable to changes in the environment. Many methods have been proposed for segmentation; statistical methods are the most popular ones. These methods model the background based on statistical information extracted from incoming frames. In this study, we estimate the background using vector quantization theory. The motion mask is created by subtracting incoming frames from estimated background under various conditions. The performance was measured by some metrics such as similarity and error-rate. The results have shown better accuracy of our proposed method especially when the color variation between background and foreground objects is high; and preserving the high quality background image during the segmentation process without consuming more memory which makes it suitable for real-time applications.

Keywords: *Moving Object Segmentation, Background Subtraction, Vector Quantization, Motion Detection, Video Surveillance.*

1. Introduction

Computer Vision (CV) has a broad range of applications from home to industry and medical; these applications deal with acquiring, analyzing, and understanding of images or sequence of images which their purpose is to interpret images without human involvement. Some of those applications are: diagnostic based on medical images, long-distance medical surveillance, and automatic inspection of product lines, autonomous vehicle navigation, human computer interaction, elderly fault detection, automated image and video analysis for multimedia indexing, leaving item detection, and many others. In many of CV applications like video surveillance and tracking, a sequence of images is processed and during this process we are looking for moving objects. Thus, this task is called Moving Object Segmentation (MOS) which is a primary task for mentioned applications. The ultimate goal of MOS is to distinguish moving pixels from non-moving pixels. Usually all moving pixels are not useful; it means that sometimes the environment that the object is moving on it is dynamic. Swaying tree branches and waving ocean surface are two phenomenal that make the environment dynamic and complex. Complexity itself has different context such as sudden changes in illuminations, different weather conditions and also the movement pattern of moving object. There are many methods for segmenting moving objects and also lots of challenges that affect the accuracy of segmentation. Segmentation methods are placed in three main categories namely, optical flow diagram, temporal difference, and background subtraction. Optical flow diagram is used in the cases

Article History:

Received Date: Nov. 011, 2014
Accepted Date: Jan. 13, 2015
Available Online: Feb. 16, 2015

that the camera is moving; temporal difference and background subtraction are usually used for applications with stationary camera. In temporal difference algorithms, the segmentation is done by calculating absolute difference of current frame with one or two previous frames. In such cases that we consider just one or two previous frames, the adaptability to changes in the scenes is high but, they have the problem of fragmentation. For background subtraction algorithms we need to create the Reference Background Image (RBI) then, the moving parts of current frame is obtained by subtracting the current frame from RBI($I_F = I_t - I_B$) which is called foreground image. Preserving the RBI during the segmentation process is the most important task in background subtraction algorithms. Sometimes temporal difference and background subtraction are combined together.

During the segmentation process there are some situations that make the segmentation hard. One of those situations is the scenes with complex background. As mentioned previously, complexity itself has many contexts such as dynamic background, sharp illumination changes, and also the motion pattern of moving object. So, a good segmentation algorithm should be able to ignore the movement of non-interesting objects and adaptable to sharp changes in environment like illumination. Sometimes tackling these noises reduce the accuracy of segmentation. Some segmentation methods works under a certain condition of frame rate and some of them are sensitive to coefficients and thresholds. Memory requirement and computational complexity are the factors that judge the suitability of a method for real-time applications.

In this study we proposed a framework for segmenting moving objects by a background subtraction approach. The background is modeled using Vector Quantization (VQ). In this framework pre-processing and post-processing increased the accuracy of the segmentation.

Section 2 of this article provides a review on the related studies. Section 3 goes through VQ theory. Section 4 explains the proposed approach. Section 5 will discuss the experimental results. Section 6 contains the conclusion and suggestions for future works.

2. Related Works

As mentioned previously, temporal difference, background subtraction, and optical flow diagram are three main groups of algorithms for segmenting moving objects. Because of their performance, background subtraction approaches gained more popularity among researchers. The effectiveness and performance of such approaches depends on the quality of RBI. The quality means how good the background is modeled or estimated; and how good it is adopted to changes in environment. There are many methods from simple averaging to sophisticated probabilistic and statistical ones for modeling the background.

Running Average (RA) algorithms model the background by calculating the weighted sum of incoming frames and previous background. As it is shown by Equation 1, new background is a fraction of current frame and previous background. In traditional RA it is possible to the background be polluted by the pixels belonging to the foreground. To overcome this problem selectivity RA has been proposed; as Equation 2 shows, in selectivity RA the pixels that belong to the background participate in the background update procedure [16]. In all RA algorithms selecting a proper value for adaptation rate α is a challenge. In a work done by [6], fuzzy inference was used for updating the background. In dead the update rate α, is calculated dynamically based on two input variables difference and area. RA algorithms can also be used in frequency domain [14].

$$B_{t+1}(x,y) = (1-\alpha)B_t(x,y) + \alpha F_t(x,y) \tag{1}$$

$$B_{t+1}(x,y) = \begin{cases} (1-\alpha)B_t(x,y) + \alpha F_t(x,y) \; if \; F_t(x,y) is \; background; \\ B_t(x,y) \qquad\qquad\qquad\qquad if \; F_t(x,y) is \; foreground \; pixel. \end{cases} \tag{2}$$

In statistical approaches the background is modeled by the statistics learned from incoming frames. In simple statistic for every pixel in a particular location the mean and standard deviation of that location is calculated then, foreground/background judgment is based on the position of new pixel value toward calculated mean and standard deviation. If the value is between mean and standard deviation it is background otherwise it is considered as moving pixel. In advanced cases, the Probability Density Functions (PDF) are used to model the mentioned simple concept.

Gaussian distribution is used to estimating the probability density of background. This function calculates the likelihood of a pixel to be background or foreground by determining how the pixels are far from their mean. Equation 3 shows the Gaussian function which μ is the mean value and Σ is the standard deviation. In segmentation methods based on Gaussian function the parameters should be updated by every incoming frame [12]. Another studies used Mixture of Gaussian (MoG) for storing multimodal backgrounds [19]. In this case one of the Gaussians corresponds to the background color based on its persistence and variance. As Equation 4 shows, k Gaussian is used for modeling the background; k is usually between 3 and 5; larger k means more memory consumption and more computational complexity. $w_{i,t}$ is a weight which determine persistency of every Gaussian functions [15]. Kernel Density Estimation (KDE) is another method for estimating the density. It is more general than Gaussian density estimation but, it needs more memory to model the background.

$$\eta(X_t,\mu,\Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t-\mu_t)^T\Sigma^{-1}(X_t-\mu_t)} \tag{3}$$

$$P(X_t) = \sum_{i=1}^{K} w_{i,t} * \eta\left(X_t,\mu_{i,t},\Sigma_{i,t}\right) \tag{4}$$

A work done by [9] used Sigma Delta Estimator (SDE) as a non-linear approach for modeling the background statistics. In that work, the mean and standard deviation are calculated using sign function. Multiple SDE (MSDE) was proposed to overcome the shortcomings of SDE. The principal of MSDE is to create a set of k background instead of one; each background has its own standard deviation. The long-term background has more chance to be selected as RBI in background subtraction process [10].

A probability-based background extraction was proposed by [1]; the core idea behind this method is that the probability of the background color is higher than the probability of moving object color. Bayesian network combines spatial and temporal information to do foreground/background classification. This approach is computationally expensive and needs more memory for storing statistical tables [5][3].Probabilistic Neural Network (PNN) classifies the pixels as either foreground or background. PNN plays the rule of Gaussian density estimation. In fact, it is an unsupervised classifier [2]. Competitive Neural Network (CNN) models multi modal backgrounds [7]. An effort by [8] models the background by Self Organizing Map (SOM) neural networks.

Adaptive motion histogram was used to segment moving vehicles. The idea behind this approach is that the motion information of objects which attracts human eyes is different from other objects. In this case the moving vehicle attracts human eyes more than moving tree branches. The most important parts of mentioned method are background update and motion histogram update [18]. An efficient method for moving object segmentation proposed by [4] which consists of three main modules. Those modules are background modeling (BM), alarm trigger (AT), and object extraction (OE). In mentioned method, just the changed pixels

participate in the update procedure. As mentioned, according to the literatures there are many methods for modeling the background. All these methods are different in terms of time complexity, memory consumption, and accuracy of segmentation. Usually all of these matters have effect on each other; for instance, when we use a single Gaussian for modeling the background the memory consumption, time complexity, and accuracy are low but, when a mixture of Gaussians is used more memory is used and the accuracy is increased and of course more time is needed to update Gaussian's parameters. The best method is the one, suitable for real-time applications in term of memory, time, and accuracy.

3. Vector Quantization

In this section we will discuss the Vector Quantization (VQ) which is a classical quantization technique from signal processing which allows the modeling of Probability Density Function (PDF) by the distribution of prototype vectors. In another words it maps a set of values into a finite set of representative values without losing too much data. It works by dividing a large set of vectors into small vector groups. Each group is represented by its centroid; the centroids approximate the group members. In fact it is a clustering method. The technical details of Vector Quantization theory are as below:

- Vector Space (R): A set of vectors that should be mapped (input data).
- Representative Vectors (Y): A finite set of vectors which the vectors in vector space are mapped to them; each vector in Y is called a code-word.
- Voronoi Regions (V): This regions partition the entire vector space. Each region contains vectors with smallest distance with its representative vector. Figure 1 shows voronoi regions in a two dimensional space. Voronoi regions are defined by Equation 5.

$$v_i = \{x \in R: \|x - y_i\| \leq \|x - y_j\|, for\ all\ i \neq j\} \tag{5}$$

The below pseudo code shows the procedure of VQ to model Probability Density Function.

i. Pick a vector from vector space.
ii. Specify the nearest representative by calculating the distance between candidate vector and all representative vectors.
iii. Move the nearest representative vector toward the candidate vector by a small fraction of the distance between candidate vector and nearest representative vector.
iv. Repeat the procedure.

The nearest representative vector is determined by calculating the distance between incoming vector and representative vectors. As Equation 6 shows, the Euclidean distance measure is used for calculating the distance. The closest representative vector is updated by Equation 7. In the mentioned Equations, n represents the number of scalars in every vector andx_t is the tth vector and y_t is the closest representative feature vector.

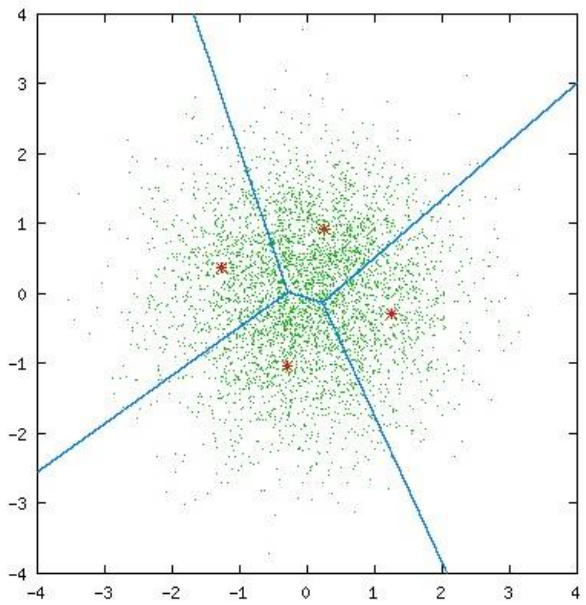


Figure 1: Voronoi regions are showed by blue lines; Input vectors are marked with green points; Code-words are marked with red points (http://www.data-compression.com/vq.shtml).

$$d = \sqrt{\sum_{i=1}^{n}(x^i - y^i)^2} \tag{6}$$

$$y_{t+1} = y_t + \alpha(x_t - y_t) \tag{7}$$

As mentioned previously, for segmentation based on background subtraction many methods try to estimate the probability density of the background. Most of these methods like Bayesian network have sophisticated procedures; this sophistication brings a huge computational and storage burden for the system. In this study VQ is used because of its straightforward procedure for modeling the background. This modeling seems to be efficient in terms of computation and storage.

4. Methodology

This section will discuss the methodology to be used and the modules that making up the system. This approach is categorized as a background subtraction method; so, background modeling and background updating are inherent characteristics of proposed method. As Figure 2 shows the block diagram of proposed method together with the process flow and sub modules that are making the system, the main modules of the proposed method are: 1) Background Modeling & Update. 2) Background Subtraction. 3) Noise Reduction. In the following sub sections every module will be discussed.

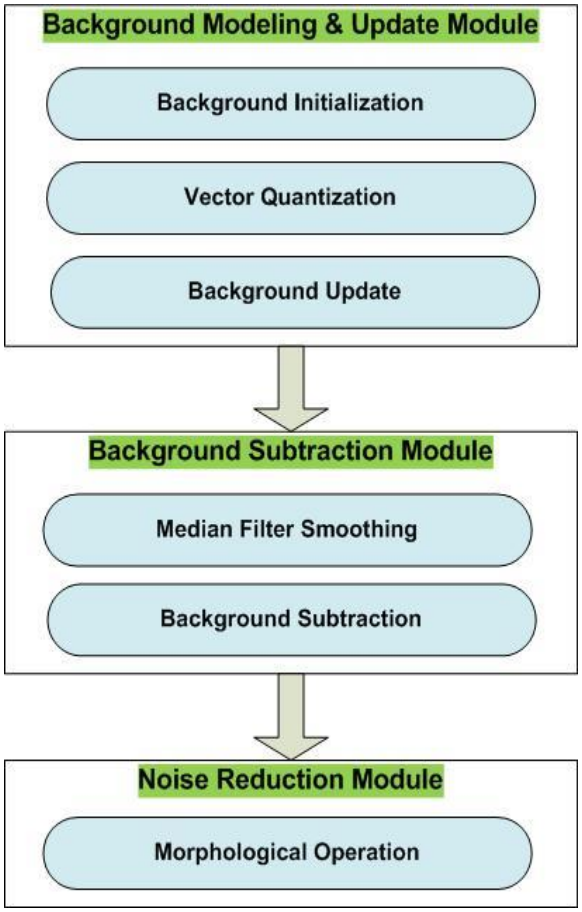


Figure 2: The block diagram of proposed Moving Object Segmentation approach.

The three modules of proposed approach are running for every incoming frame of the video. The process starts with background modeling module to create the RBI and initializing it with the first frame of the video. This reference background image will be adapted to changes in the background. Then, the background subtraction module runs to create the motion mask. Finally, the noise reduction module as our post-processing step removes the noises to enhance the motion mask. The following sections are describing the details of every module of the system.

4.1. Background Modeling & Update Module

As mentioned previously, the vital part of background subtraction algorithms is background reference image creation and also keeping it updated. The goal of this module is to create the reference background image and adopting it to new changes in the environment or background. In proposed method, RGB color components are used as features for modeling the background. Five RGB vectors store the RGB statistics for every pixel of the background. Every RGB vector keeps one aspect of the background. One of those five vectors is used to store the RBI which will be used in the subtraction module. Every of four remaining vectors are used to store different backgrounds for multi modal backgrounds. The more frequent one of the four vectors will update the RBI. Figure 3 shows the relation between the reference background vector and non-reference vectors and also the flow of update. After processing every incoming frame, one of non-reference backgrounds is updated and subsequently the reference background is updated with the data acquired from one of non-reference backgrounds. The next three sub modules will describe the details of this module.

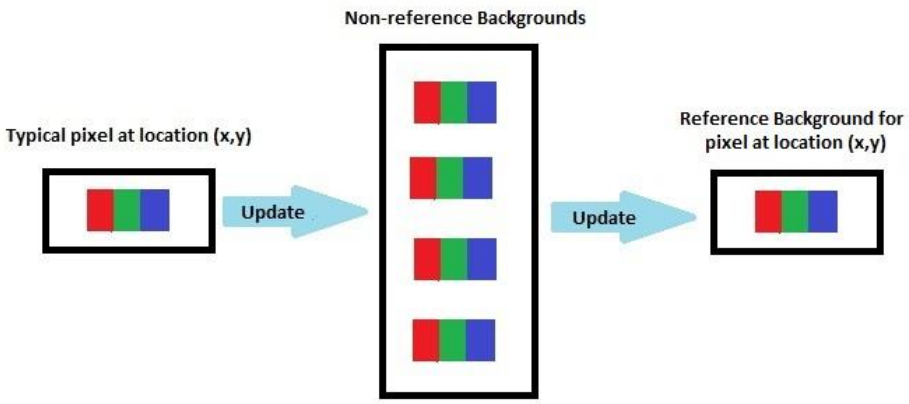


Figure 3: The relation between the reference background, non-reference backgrounds, and their update process.

4.1.1. Background Initialization: In this method we initialized the RBI and one of the non-reference backgrounds with the first frame of the video. The other non-reference backgrounds are initialized randomly. The purpose of initializing one of non-reference background images is to decrease the adaptation time. For the RBI we imperially observed the better result instead of random initialization.

4.1.2. Vector Quantization – NRBI Update: By now, we will describe how VQ is supporting the update of non-reference background images. The below pseudo code shows the procedure for updating the N-RBI:

i. For every pixel of incoming frame do.
ii. Specify the nearest N-RBI image by calculating the distance between the current pixel vector and all of N-RBI vectors.
iii. Move the nearest N-RBI vector toward the incoming pixel vector by a small fraction of the distance between incoming pixel vector and nearest N-RBI vector.
iv. Repeat the procedure for all the pixels of the frame.

In this methodology we use the VQ to form non-reference backgrounds and keeping them update. In proposed method, four representative vectors are used to store four different model of the background. And each vector keeps RGB values as features. Every incoming pixel updates its corresponding pixel in closest N-RBI vector.

4.1.3. Reference Background Image Update: After updating non-reference background images with every incoming frame. It is reference background image turns to be updated. As Equation 8 shows, the RBI is updated with the most frequent N-RBI. After every n frame of video, one of the N-RBIs that have maximum update hit is our potential N-RBI which updates the RBI. In mentioned equation, α is learning rate; based on our experiment we found the value 0.1 is the best in our cases.

$$RBI_{i+1} = (1-\alpha)RBI_i + (\alpha)N\_BRI_{potential} \quad (8)$$

4.2. Background Subtraction Module

The goal of background subtraction is to create motion mask for every incoming frame. This motion musk is applied to incoming frames for extracting moving objects. The procedure of motion mask creation starts with subtracting current frame from RBI in a pixel by pixel manner. The Equation 9 shows the mechanism of binary motion musk creation by applying a proper threshold. Using a proper value as threshold is very important and it has a direct effect on the accuracy of the motion mask.

$$M_t(x,y) = \begin{cases} 0, & |I_t(x,y) - RBI_t(x,y)| < T \\ 1, & |I_t(x,y) - RBI_t(x,y)| \geq T \end{cases} \tag{9}$$

The median filter is applied on the incoming frame and also on the RBI for reducing the noises before subtraction. The purpose of applying median filter is to reduce the noises. Camera shaking and improper setting of acquisition device can be the source of noise for incoming frames and consequently the RBI.

4.3. Noise Reduction Module

As mentioned previously, in the background subtraction approaches the most important parts of the system are modeling the background and keeping it update; and also the subtraction procedure itself is important. But, the accuracy of motion mask depends on another factor which is the noises. In our case the noise sources are acquisition noises and background noises. Acquisition noises are those created by camera shaking and complex scenes; and background noises are those created by the update procedure of the background. A good method for updating the background should have the least affect in the motion musk in term of noise. So, noise reduction is an inevitable part of such systems.

In proposed method we reduce the noises two times; before subtraction and after subtraction. The median filter is used before subtraction to smooth the RBI and current frame of the video sequences. This smoothing helps to have better results after subtraction. Technically speaking, the cooler variation in the noise area is greater than the variation between object of interest and background. So, median filter helps to smooth the noise areas and normal the variation. After subtraction usually the motion musk is polluted by salt and pepper noises. In this case the morphological operators are used to eliminate this type of noises.

5. Experimental Results

This section provides discussion and analysis of method explained in the previous sections and its corresponding results. We setup our experiment on well-known bench mark clips and sequences. After preparing test video sequences and developing proposed method, the motion mask of all video sequences has been extracted; during the extraction statistics of motion mask such as false positive, true positive, and false negative rates are extracted. Those statistics are used for performance measurement purpose. For qualitative analysis the motion musk of typical frames of sequences are used. The qualitative discussion is based on quality of this motion masks to see how much they are close to their corresponding ground truth images. For quantitative analysis some metrics like recall, precision, error rate, F, and similarity are measured. In another perspective the memory consumption and the number of frames that can be processed in one second are measured to determine whether the proposed method is suitable for real-time applications or not.

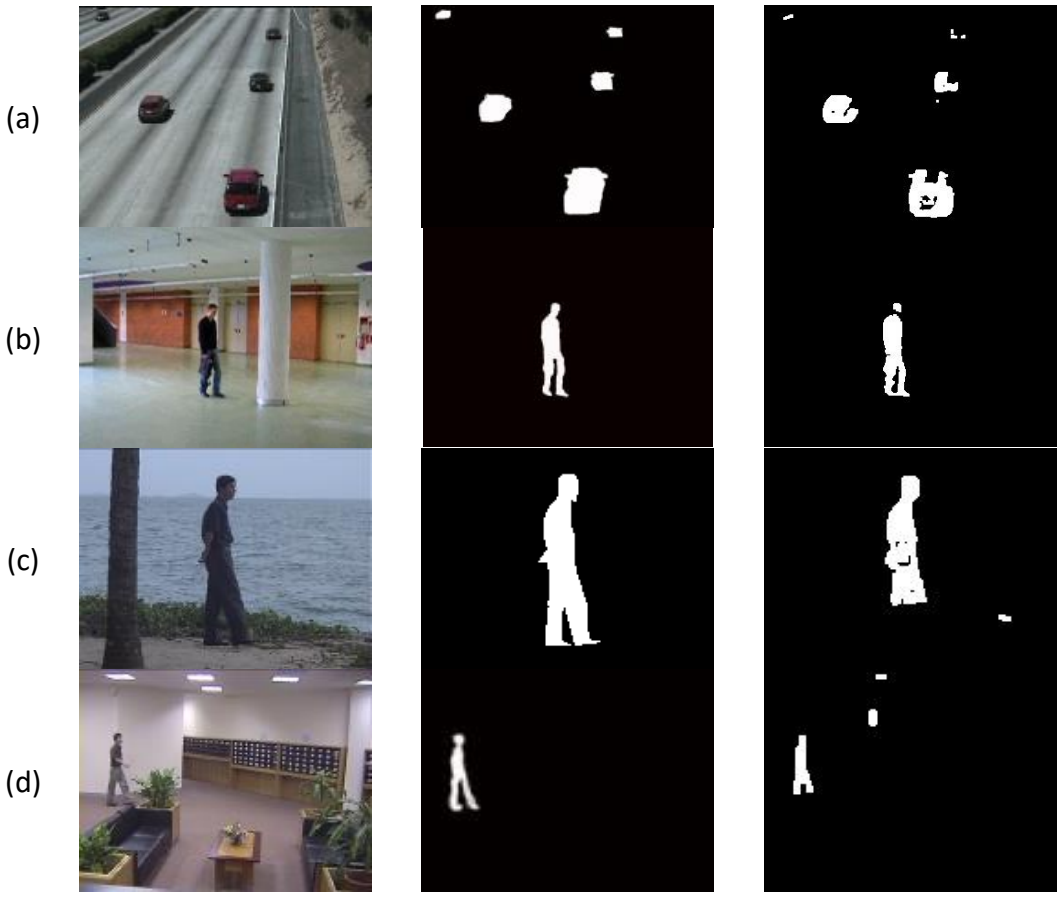
5.1. Qualitative Results

As mentioned previously, the purpose of qualitative analysis is to determine how much our motion mask is close to ground truth visually. In this analysis the frames with existence of their ground truth are used. The ground truths have been created manually for those sequences that their ground truths are not available. This section will be continued by representing the comparison between motion masks have been created by proposed method and ground truths of typical frames of different sequences.

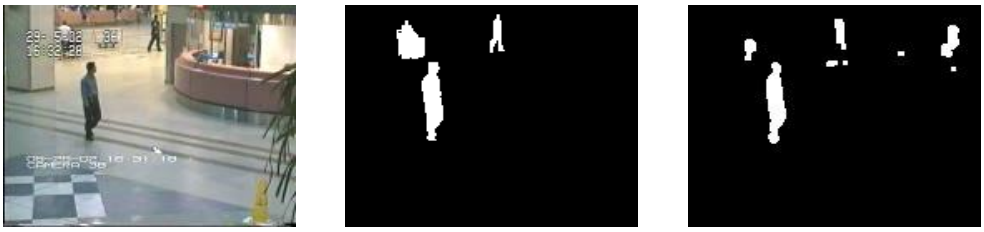
Figure 4 is showing the results of our experiment for five image sequences. As it is seen, this method can produce acceptable motion mask; they are accurate especially when the color variation between background and foreground objects is high. Figure 4(b) is a good example of such especial situations. Figure 4(c) shows the water surface sequences which from that it could be seen that the motion musk is acceptable for upper part of the person; but, the mask doesn't cover the area below the knee of the person because of low color variation.

The proposed method can perform well while the background is complex by its dynamic. Figure 4(c) is an evidence of that assertion because it performs well on tackling the noises occurred by waving water surface that makes the background complex. This performance has been achieved by having multiple Voronoi regions for storing multiple backgrounds.

In another perspective, the quality of extracted background and effect of median filter have been tested. Median filter smooth the noises in the extracted background especially in the boundary of objects. These noises are usually salt and pepper noises and they occur by camera shaking, inaccuracy of acquisition device, and dynamic backgrounds. Figure 5 shows the quality of extracted background and effect of median filter for two typical frames of two sequences.



(e)



Original Frames **Ground Truths** **Proposed Method**

Figure 4: The result of typical frames of; a) Highway II, b) MSA, c) Water Surfaces, d) Lobby, and e) Airport Hall sequences. First column, the original frames; Second Column, the ground truths; Third Column, the motion masks created by proposed method.

(a) (b) (c)

Figure 5: The results of background creation for Highway II and MSA sequences. (a) The original frames. (b) The backgrounds without median smoothing. (c) The backgrounds with median smoothing.

In this section some characteristics of proposed method such as the quality of motion mask in dynamic backgrounds and free-noise motion mask have been highlighted. Based on acquired motion mask the only shortcoming of proposed method is its inaccuracy when the color variation between foreground and background object is low. Next section will measure mentioned characteristics and analyses them quantitatively.

5.2. Quantitative Results

In the area of moving object segmentation there are three common performance measures namely memory consumption, computational time, and accuracy. In this days memory is not as important as computational time but, for real-time and embedded surveillance applications it is still important. Normally for computational time the number of frames that can be processed in one second are measured. Memory consumption means the amount of memory which is used to store all required data structures to process a frame. When we talk about the accuracy in this area, generally we want to know how much the motion mask is similar to the ground truth; and how much is the rate of error. There are different metrics for accuracy measurement which Recall, Precision, Similarity, and measure are most common ones. The Equations 10 to 13 show the way that these metrics are calculated [8]. Also Equation 14 defines the error-rate based on [1], to calculating the error rate as an important metric. Table 1 shows the value of mentioned metrics for typical frames of different sequences.

$$Recall = \frac{tp}{(tp+fn)} \tag{10}$$

$$Precision = \frac{tp}{(tp+fp)} \tag{11}$$

$$F_1 = \frac{2*R*P}{(R+P)} \tag{12}$$

$$Similarity = \frac{tp}{(tp+fp+fn)} \tag{13}$$

$$Error - Rate = \frac{Error\ Rate\ Count}{Frame\ Size} = \frac{Foreground\ Error\ Count + Background\ Error\ Count}{Frame\ Size}$$

$$= \frac{\sum_{(x,y)}[F(x,y) \oplus GT(x,y)]}{Frame\ Size} \tag{14}$$

Table 1: The accuracy measurement for sample frames of different sequences.

Sequence	Recall	Precision	F_1	Similarity	Error Rate
Highway	0.9877	0.9905	0.9891	0.9785	0.021
MSA	0.9778	0.9866	0.9822	0.9649	0.003
Water Surface	0.9717	0.9984	0.9849	0.9702	0.029
Lobby	0.9955	0.9962	0.9958	0.9917	0.008
Airport	0.9883	0.9909	0.9896	0.9794	0.020

As shown in Table 1 and Figure 6, the similarity rate between ground truth and motion mask created by proposed method is high; and also the error rate is low. All these are proving the efficiency of proposed method especially in the cases that color variation between background and foreground objects are high. In another effort the error rate for different frames of two sequences have been calculated and Figures 7 and 8 are showing their corresponding plot.

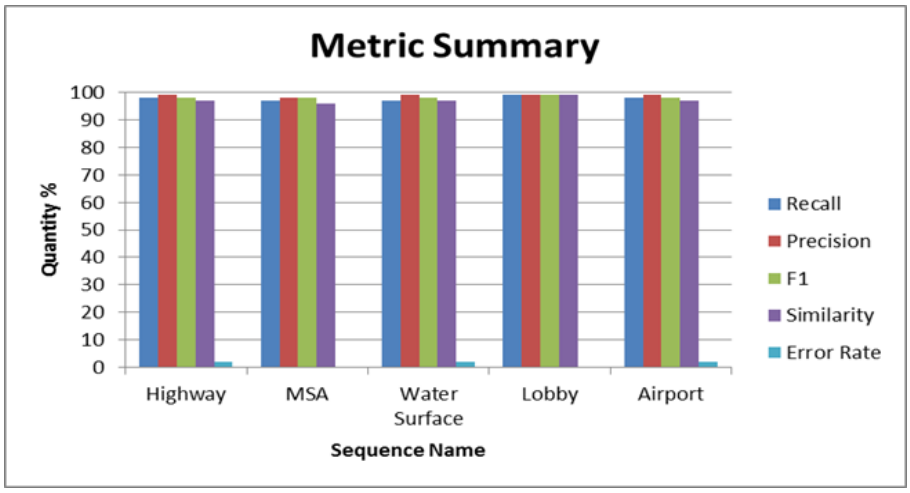


Figure 6: The accuracy measurement for sample frames of different sequences.

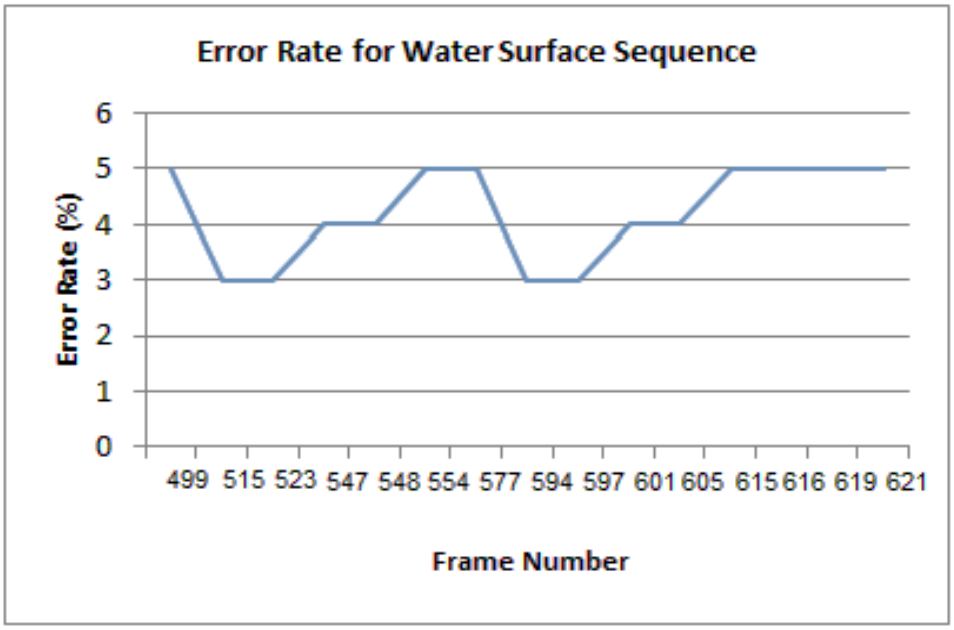


Figure 7: Error Rate for different frames of Water Surface sequence.

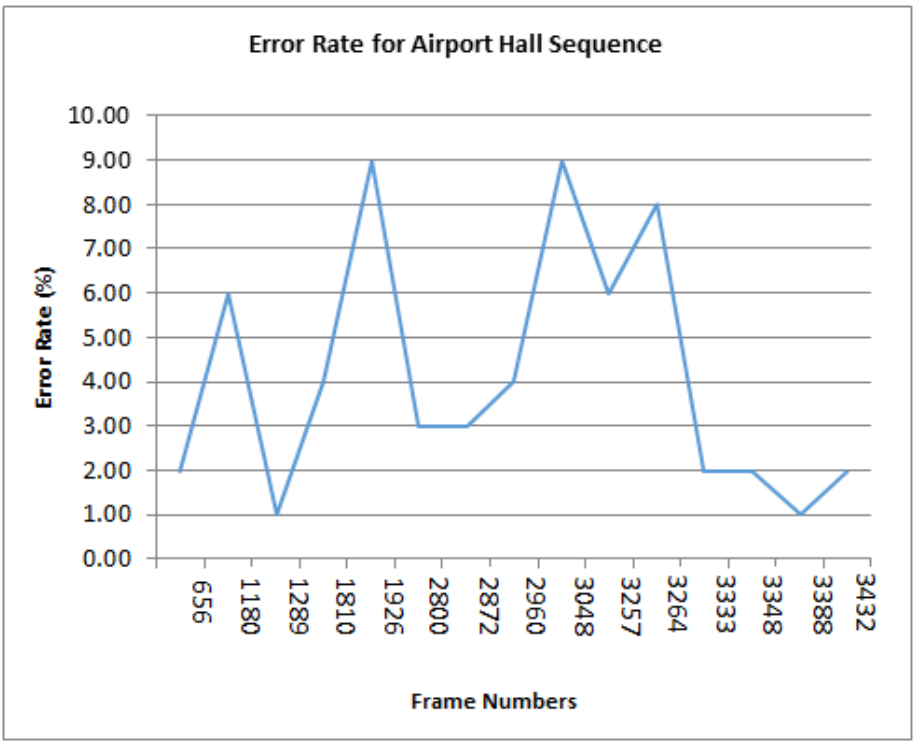


Figure 8: Error Rate for different frames of Airport Hall sequence.

As Figure 7 shows, the error rate for different frames of water surface sequence is between three and five. This values representing the durability of proposed method as one of the conditions of a good segmentation method. As Figure 8 shows, the error rate for different frames of AIRPORT HALL sequence varies from one to nine. The reason for this wide range of difference is that for frames with high error-rate the RBI is not adapted yet to sharp changes in the scene. In this case the source of change is illumination. The average error rate for both sequences is 0.04. As mentioned previously, memory consumption is important for real-time and embedded applications.

The proposed method is efficient in terms of memory consumption. Table 2 shows the evaluation of memory consumption for different frame sizes.

Table 2: Memory consumption for frames with different sizes

Frame Width(Pixel)	Frame Height(Pixel)	Memory Requirement (MByte)
320	240	12.30
176	144	4.06
160	128	3.28

As mentioned previously, the number of frames that can be processed in one second is an important factor for evaluating moving object segmentation methods. So, the proposed method runs over an environment with 2.67 GHz core 2 Duo CPU. In this effort sequences with different frame size were examined. Table 3 shows the results of this examination. In the mentioned table, the processing of average frame per second is a function of Total Frame, Total Time, and Frame Size. The implementation of proposed approach is suitable for parallel processing and hardware implementation.

Table 3: Processing time of proposed method.

Sequence Name	Total Frames	Total Time (s)	Average Frame/Second
Water Surface	633	24	26
Airport Hall	3584	169	21
Lobby	1546	49	32
MSA	528	61	9
Highway	500	55	10

Conclusion and Future Works

This paper proposed the development of moving object segmentation method based on vector quantization theory. The results of proposed method have shown the quality of motion mask is satisfactory and the motion mask is almost free of noise. The results are very good especially when the color variation between foreground and background is high. These results are based on accuracy measurement metrics like similarity and error rate and etc. This method consumes less memory in comparison with other methods especially statistical methods. Also it does not have complex computation; so, it can be a candidate for real-time applications. This method can be implemented massively parallel. We suggest implementing this method in intensity mode. In this mode it is needed to convert RGB frame to intensity. In this mode the memory consumption and computational time will be very low. And it is guessed that results will be satisfactory. Another suggestion is to use adaptive threshold. The adaptive threshold can compromise the error-rate and accurate object detection in different situations.

ACKNOWLEDGEMENTS

The authors would like to thank the Universiti Teknologi Malaysia (UTM) for its support on our Research and Development, and the Soft Computing Research Group (SCRG) for the inspiration for making this study a success.

References

[1]. Chung-Cheng Chiu, Min-Yu Ku, Li-Wey Liang,"A Robust Object Segmentation System Using a Probability-Based Background Extraction Algorithm. IEEE Transaction on Circuits and Systems for Video Technology",Vol. 20, 4 (2010), pp.518 - 528.

[2]. Culibrk D., Marques O., SocekD., Kalva H., Furht B., "Neural Network Approach to Background Modeling. IEEE Transactions on Neural networks", Vol. 18, 6 (207), pp.1614-1627.

[3]. Duong T.-T. and A.-D. Duong,"Moving Objects Segmentation in Video Sequence based on Bayesian network. Computing and Communication Technologies, Research, Innovation, and Vision for the Future (RIVF) IEEE RIVF International Conference", (2010) 1-6.

[4]. Huang, S.-C.,"An Advanced Motion Detection Algorithm With Video Quality Analysis for Video Surveillance Systems. IEEE Transactions on Circuits and Systems for Video Technology", Vol. 21, 1(2011),pp. 1 - 14.

[5]. Li L., W. Huang, Irene Yu-HuaGu, Qi Tian, "Statistical Modeling of Complex Backgrounds for Foreground Object Detection. IEEE Transactions on Image Processing", Vol. 13, 11(2004), pp.1459 - 1472.

[6]. Lijun, X.,"Moving Object Segmentation Based on Background Subtraction and Fuzzy Inference. International Conference on Mechatronic Science, Electric Engineering and Computer", (2011) 434 - 437.

[7]. Luque R. M., E. Domínguez, E. J. Palomo, J. Muñoz,"A Neural Network Approach for Video Object Segmentation in Traffic Surveillance. 5th International Conference, ICIAR", (2008) 151-158;Póvoa de Varzim, Portugal.

[8]. Maddalena, L. and A. Petrosino,"A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications. IEEE Transactions on Image Processing", Vol. 17, 7 (2008), pp. 1168-1177.

[9]. Manzanera, A. and J. C. Richefeu,"A Robust and Computationally Efficient Motion Detection Algorithm Based on Sigma-Delta Background Estimation. Fourth Indian Conference on Computer Vision, Graphics & Image Processing", (2004) 1-6; Kolkata, India.

[10]. Manzanera, A. and J. C. Richefeu,"A new motion detection algorithm based on Σ-Δ background estimation. Pattern Recognition Letters", Vol. 28, 3 (2007), pp.320-328.

[11]. Oral, M. and U. Deniz,"Centre of mass model – A novel approach to background modelling for segmentation of moving objects. Image and Vision Computing", Vol. 25, 8(2007), pp.1365 – 1376.

[12]. Stauffer, C. and W. E. L. Grimson,"Adaptive background mixture models for real-time tracking",(1999) 1-7;Massachusetts Institute of Technology,The Artificial Intelligence Laboratory.

[13]. Wang, W.-J., C.-T. Huang, et al,"VQ Applications in Steganographic Data Hiding Upon Multimedia Images. IEEE Systems Journal", Vol. 5, 4(2011), pp.528-537.

[14]. Wang W., J. Yang,"Modeling Background and Segmenting Moving Objects from Compressed Video. IEEE Transactions on Circuits and Systems for Video Technology", Vol. 18, 5(2008), pp.670 - 681.

[15]. Yan, A., J. Li, et al.,"Color Moving Object Segmentation Based on Mixture Gaussian Models, Sixth International Conference on Natural Computation (ICNC)", (2010) 1208-1211.

[16]. Yi, Z. and F. Liangzhong,"Moving Object Detection Based on Running Average Background and Temporal Difference. International Conference on Intelligent Systems and Knowledge Engineering (ISKE)", (2010) 270-272; Hangzhou.

[17]. Yu, C., X. Zheng, et al.,"Review of Intelligent Video Surveillance Technology Research. International Conference on Electronic & Mechanical Engineering and Information Technology", (2011) 230-233.

[18]. Zhang, W., Q. M. J. Wu, et al.,"Moving vehicles detection based on adaptive motion histogram. Digital Signal Processing", Vol. 20, 3(2010), pp.793–805.

[19]. Rezaee K., Mousavirad SJ., Ghezelbash M., Hadadvandia J., "Accurate Fire Detection System for Various Environments using Gaussian Mixture Model and HSV Space, Journal of Information systems and Telecommunication", Vol. 1, pp.55-62.