A Robust Method for New Object Detection in Video Surveillance Systems

Luma Akram Hamdai
Computer Engineering Department
University of Mosul
Luma.hamdi88@yahoo.com

Dr. Ahlam Fadhil Mahmood
Computer Engineering Department
University of Mosul
Ahlam.mahmood@gmail.com

Abstract

Video surveillance systems is a powerful tool for public safety and with the increasing need for more security in airports, banks, schools and other critical environments, the demand for video system is growing rapidly. Sides from the intrinsic usefulness of begin able to segment video streams into moving and background components, detecting moving blobs provide a focus of attention for recognition, classification and activity analysis, making these later processes more efficient since only "moving" pixels need to be considered. In this paper an efficient moving object detection method using modified Horprasert model for video surveillance system is present. The modified dynamic thresholds are able to detect a new object with it's shadow through different video stream in terms of light conditions. It consists of background model, distortion of brightness, color calculations and classification. Four regions are segmented depending on the thresholds foreground (moving object), background, highlight background and shadow. The proposed automatic threshold depends on background computations of brightness and is thus expected to achieved better classification performance.

Keywords: Horprasert model; background subtraction; background model; Video Surveillance.

طريقة لاكتشاف وجود أجسام جديدة في نظم المراقبة الفيديوية

د.أحلام فاضل محمود
قسم هندسة حاسبات
ahlam.mahmood@gmail.com

لمى أكرم حمدي
قسم هندسة حاسبات
luma.hamdi88@yahoo.com

لقد أصبحت أنظمة المراقبة الفيديو بلا شك أداة قوية للسلامة و الأمن العام و مع تزايد الحاجة لها في المطارات، و البنوك، و المدارس، و البيئات المزدحمة الأخرى بات من المهم متابعة تطور خوارزميتها بسرعة كبيرة. و تعتبر عملية فصل الجزء المتحرك من التتابع الفيديوي عن خلفية المشهد من أهم الخوارزميات في أنظمة المراقبة، بالأضافة لتصنيف، و تمييز و تحليل النقاط الصورية الفيديوية. في هذا البحث قدمت طريقة فعالة و مؤثرة لكشف عن وجود أجسام جديدة تدخل في صور لاحقة للتتابع الفيديوي باستخدام نموذج Horprasert. إذ أضافت الطريقة المقترحة عتبات ديناميكية تتغير حسب نوعية الدفق الفيديوي مما مكن من الكشف عن وجود كائن جديد لفيديوهات في ظروف إضافة مختلفة. يتألف التحسين المقترح من عملية تدريب لتحديد عناصر نموذج الخلفية، تشويه السطوع، و حسابات اللون و من ثم التصنيف التلقائي الى أربع مناطق أعتمادا على قيم العتبات و هي أولا: الجسم المتحرك و الخلفية و الظل. في التطوير المقترح يتم أيجاد العتبات بصورة أوتوماتيكية إذ يعتمد على عمليات حسابية لخلفية السطوع و بالتالي يتوقع أن يحقق أداء أفضل للكشف عن وجود أجسام متحركة.

1. Introduction:

Research on visual surveillance is gaining momentum due to security requirements in sensitive areas such as airport, offices, railway stations, etc. Moving object detection is the fundamental step in many visual surveillance applications like object tracking, action recognition, high level semantic description, gesture recognition, etc. Out of three major classes of moving object detection techniques, namely, image differencing, optical flow and background subtraction, the last is considered to be somewhat robust, as compared to others [1,2]. Temporal differencing method is that adjacent images are subtracted so that still objects are removed, and the motion area is preserved. This method has the advantages of low computational complexity, little need to consider the impact of background's slow changes, and insensitivity to background changes. Optical flow can be used to detect independently moving objects in the presence of camera motion; however, most optical flow computation methods are computationally complex, and cannot be applied to full-frame video streams in real-time without specialized hardware[3]. The rationale behind background subtraction is that of detecting the moving objects from the difference between the current frame and a reference frame which is often called background model[4]. There are many background subtraction techniques like non-recursive technique store in a buffer of a certain number of previous sequence frames and estimate background model as Gaussain Mixture Model[5], while recursive technique recursively update a single background model based on input frame as Sakbot(Statistic And Knowledge-Based Object detection) and RaBS(Rapid Background Subtraction from Video Sequences)[6,7]. Other techniques are pixel-based assumed that the time series of observation is independent at each pixel as SACON(Sample Consensus)and Texture based method[8,9], while region-based segment the images into regions as Rectgauss-Tex and background modeling via classification[10,11]. The technique used in proposed modified of Horprasert Model consists of building background model during a certain period of time for a static scene. After, the background model is constructed, the brightness distortion and chromaticity distortion are calculated between background model and input frame. One the basis of brightness distortion and chromaticity distortion, a brightness distortion threshold and chromaticity distortion threshold are calculated. These thresholds are used in moving objects detection. In order to delete the noises and increase detection accuracy, morphology operation is used on classified image.

The rest of this paper is organized as follows. In section 2 the brief overview of Horprasert Model is present. The modifications of Horprasert Model are proposed in section 3. Experimental results are given in section 4. Finally, conclusions are present in section 5.

2. Horprasert Model

One of the fundamental abilities of human vision is color constancy [1]. Humans tend to be able to assign a constant color to an object even under changing of illumination over time or space. The perceived color of a point in a scene depends on many factors including physical properties of the point on the surface of the object. Important physical properties of the surface in color vision are surface spectral reflectance properties, which are invariant to changes of illumination, scene composition or geometry. On Lambertain, or perfect matte surfaces, the perceived color is the product of illumination and surface spectral reflectance.

This led to our idea of designing a color model that separates these two terms; in other words, that separates the brightness from the chromaticity component. Figure 1 illustrates the proposed color model in three-dimensional RGB space[2].
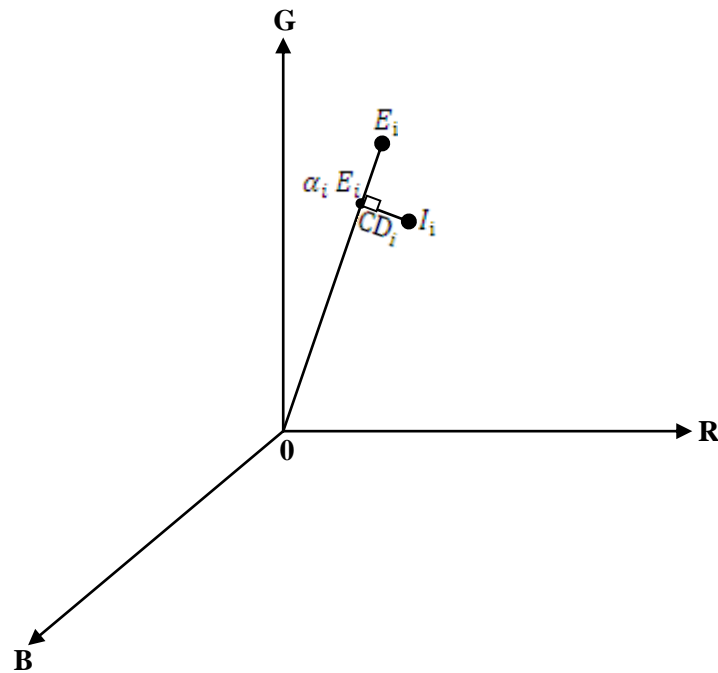
Figure 1: Horprasert color model in the RGB three-dimensional color space; the background image is statistically pixel-wise modeled. E_i represents an expected color of a given ith pixel and I_i represents the color value of the pixel in a current image. The difference between I_i and E_i is decomposed into brightness (α_i) and chromaticity (CD_i) components.

Pixel(r,c)	0,0	0,1	0,2	0,3	0,4
Ir	13	17	20	20	26
Ig	12	14	14	10	16
Ib	20	21	16	9	14
Er	15.5750	15.7250	16.45	16.25	16.6
Eg	13.3	13.4750	12.7	13.1750	12.65
Eb	15.8250	16.1250	17.35	16.6	16.7750
Sr	1.1377	1.0721	1.0235	0.6982	0.995
Sg	0.7483	0.8058	0.9539	0.7378	0.8231
Sb	1.4981	1.3818	1.6963	1.2	1.47575
A	0.0368	0.0407	0.0510	0.0357	0.0413
B	1.4270	1.2881	1.1982	1.4860	1.2645
CD	1.9987	1.5480	0.05309	0.0893	0.3417
α	0.9825	1.0461	1.0457	1.0471	0.9851
$\overline{CD}$	1.4006	1.2018	0.4431	0.0601	0.2702
$\hat{\alpha}$	-0.4753	1.1334	0.8950	1.3215	-0.3612

Consider a pixel, i, in the image; $E_i=[E_R(i); E_G(i); E_B(i)]$ represent the pixel's expected RGB color in the reference or background image. The line passing through the origin and the point Ei is called expected chromaticity line (OE). Next, let $I_i = [I_R(i); I_G(i); I_B(i)]$ denote the pixel's RGB color value in a current image that subtract from the background. Basically measuring

the distortion of I_i from E_i can be implemented by decomposing the distortion measurement into two components, brightness distortion and chromaticity distortion as defined in ref[2].

Brightness Distortion (α)

The brightness distortion(**α**) is a scalar value that brings the observed color close to the expected chromaticity line. It is obtained by minimizing

$$\Phi(\alpha)=(I_i- \alpha_i \, E_i)^2 \qquad \text{………………(1)}$$

α_i represents the pixel's strength of brightness with respect to the expected value. α_i is 1 if the brightness of the given pixel in the current image is the same as in the reference image. α_i is less than 1 if it is darker, and greater than 1 if it becomes brighter than the expected brightness.

Color Distortion (*CD*)

Color distortion is defined as the orthogonal distance between the observed color and the expected chromaticity line. The color distortion of a pixel i is given by:

$$CD_i=\|I_i-\alpha_i E_i\| \qquad \text{…………………..(2)}$$

3. Proposed Modified Horprasert Model:

Several object detection algorithms have been proposed in literature. Most of them use static thresholds to decide moving object as shown in the baseline of Horprasert model which consists of three main stages as illustrated in Figure 2.

**Figure(2) : Baseline of Horprasert Model**

The basic scheme of Horprasert background subtraction is to subtract the image from a reference image that models the background scene. Typically, the basic steps of the algorithm are as follows:

- Background modeling constructs a reference image representing the background.
- Threshold selection determines appropriate threshold values used in the subtraction operation to obtain a desired detection rate.
- Subtract operation or pixel classification classifies the type of a given pixel, i.e., the pixel in the background (including ordinary background and shaded background), or in a moving object.

This technique consists of building a reference model which represents the static background of the scene during a certain period of time. Then differences between the background image and the current image are evaluated. The differences are decomposed into brightness and chromaticity components. Applying the suitable static thresholds on the

brightness distortion(α) and chromaticity distortion(CD) of a pixel i yields an object mask M(i) which indicates the type of the pixel. This method classifies a given pixel into four categories.

But the performances of most scene change detectors decrease considerably for videos with high scene complexity and variation, so using static threshold not suitable for sudden scene configuration changes. Instead of this, a dynamically adjusted cumulative histogram is computed to find the most suitable threshold for every pixel classification in the target image. The modification in Horprasert Model consist is shown in figure(3). The key difference from previous model is, at first stage the background model is constructed by calculating four parameters $< E_i, S_i, a_i, b_i >$ for each pixel in N frames of static scene. In the second stage, the thresholds previously remain unchanged in all video sequences by determining the desired detection rate[2], now it changes dynamically to video information. At third stage the subtraction operations and classification is performed. Finally, Morphology operation is added to improve the classified images.

| Modify Background Model | → | Dynamic Threshold Selection | → | Pixel Classification | → | Morphology Operation |
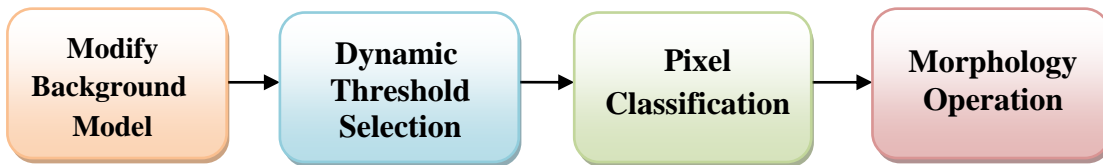
Figure(3): Block diagram of Modified Horprasert Model

3.1 Modified Background Model

In the background training process, the background model and some parameters associated with normalization are computed over a number of static background frames. For each pixel in the background frame, the mean(E_i), standard deviation(S_i), the variation of brightness (a_i), the variation of chromaticity distortion(b_i), normalized brightness distortion($\hat{\alpha}_i$) and normalized chromaticity distortion($\widehat{CD}_i$) of the ith pixel are calculated. Thus the brightness distortion in Equation(1) and chromaticity distortion in Equation(2) become:

$$\alpha_i = \min\left[\sum_{c=R,G,B}\left(\frac{I_C(i)-\alpha_i\mu_c(i)}{\sigma_C(i)}\right)^2\right]$$

$$= \frac{\left(\frac{I_R(i)\alpha_i\mu_R(i)}{\sigma_R(i)}+\frac{I_G(i)\alpha_i\mu_G(i)}{\sigma_G(i)}+\frac{I_B(i)\alpha_i\mu_B(i)}{\sigma_B(i)}\right)}{\left(\left[\frac{\mu_R(i)}{\sigma_R(i)}\right]^2+\left[\frac{\mu_G(i)}{\sigma_G(i)}\right]^2+\left[\frac{\mu_B(i)}{\sigma_B(i)}\right]^2\right)} \quad\ldots\ldots\ldots\ldots\ldots..(3)$$

$$CD_i = \sqrt{\sum_{C=R,G,B}\left(\frac{I_C(i)-\alpha i\,\mu_c(i)}{\sigma_C(i)}\right)^2} \quad\ldots\ldots\ldots\ldots\ldots..(4)$$

The variation of brightness (a_i) and chromaticity distortion(b_i) are used as normalized factors.

$$a_i = \text{RMS}(\alpha_i) = \sqrt{\frac{\sum_{i=0}^{N}(\alpha_i - 1)^2}{N}}$$

...........................(5)

$$b_i = \text{RMS}(CD_i) = \sqrt{\frac{\sum_{j=0}^{N}(CD_i)^2}{N}}$$

...........................(6)

In order to use a single threshold for all of the pixels, we need to rescale the brightness and chromaticity distortion .

$$\widehat{\alpha_i} = \frac{\alpha_i - 1}{a_i}$$

.........................(7)

$$\widehat{CD}_i = \frac{CD_i}{b_i}$$

.........................(8)

Where μ_i being the arithmetic means of each color channel for pixel i.
σ_i is the standard deviation of each color channel for pixel i.
N is a number N of images will be used.

3.2 Dynamic Threshold of Horprasert Model:

The appropriate threshold is determined by the normalization of brightness and chromaticity distortion that calculated during training period. It should be noticed the number of frame during training period is affected by choosing the threshold, increased the number of frames in background model produce better detection but requires more computation time. Therefore K is used as a scale value that changes the threshold until it becomes more appropriate from the sequence. By testing different sequences, good classification is obtained when set (k=4) for using twenty frames in training phase. The affect of number of frame in background model to the normalized brightness and normalized chromaticity is shown in figure(4) and (5).
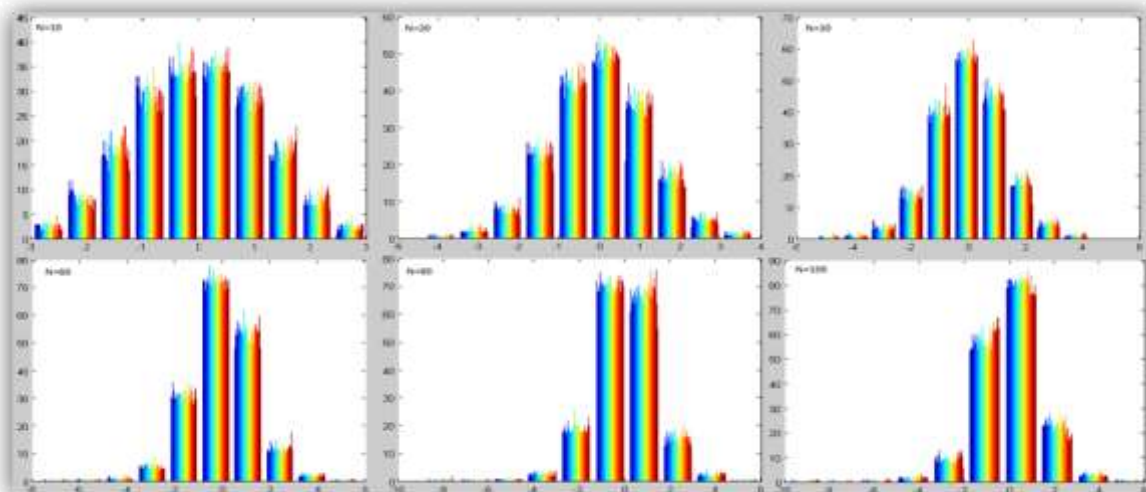
Figure(4) shows the effect of increase the number of frames in background model to normalized brightness distortion.
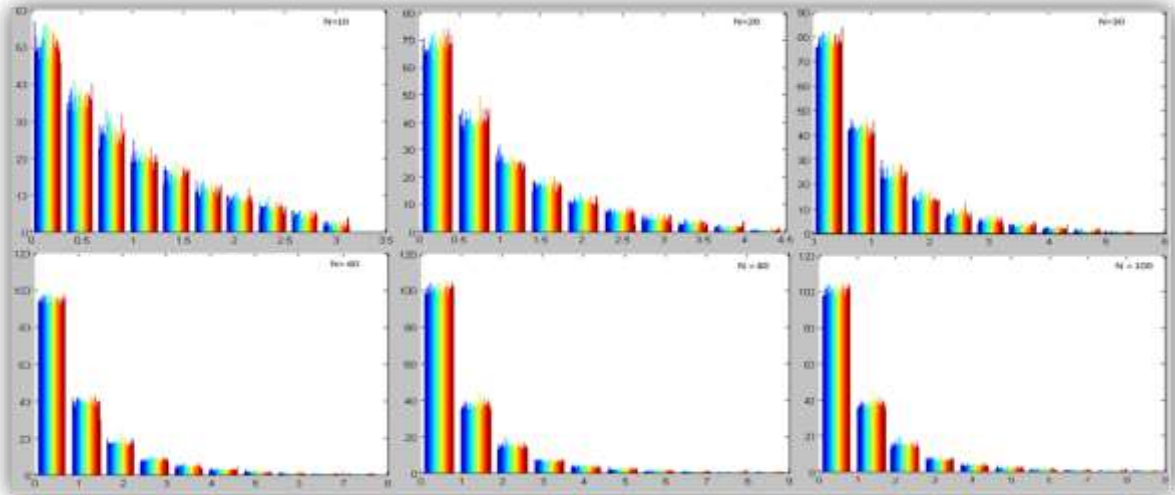
Figure(5) shows the effect of increase the number of frames in background model to normalized chromaticity distortion.

3.3 Pixel classification:-

The difference between the reference image and the current image is decomposed into brightness and chromaticity components. Applying the suitable thresholds on the normalized brightness distortion ($\widehat{\alpha_\iota}$) and the normalized chromaticity distortion ($\widehat{CD}_\iota$) of a pixel i yields an object mask M(i) which indicates the type of the pixel. The output pixel is classified into one of four categories:

- Original background (B) if it has both brightness and chromaticity similar to those of the same pixel in the background image.
- Shadow background (S) if it has similar chromaticity but lower brightness than those of the same pixel in background image.
- Highlight background (H) if it has similar chromaticity but higher brightness than the background image.
- Moving foreground object (F) if the pixel has chromaticity different from the expected values in the background image[2].

$$M(i) = \begin{cases} F : \widehat{CD}_\iota > \tau_{CD} \, or \, \widehat{\alpha_\iota} < \tau_{\alpha lo} \, , else \\ B : \widehat{\alpha_\iota} < \tau_{\alpha 1} \, and \, \widehat{\alpha_\iota} > \tau_{\alpha 2} \, , else \\ \quad S : \widehat{\alpha_\iota} < 0 , else \\ \quad H : otherwise \end{cases} \quad \dots\dots\dots\dots\dots\dots.(9)$$

$\tau_{\alpha 1}$ and $\tau_{\alpha 2}$: are thresholds that are calculated from normalized brightness distortion .
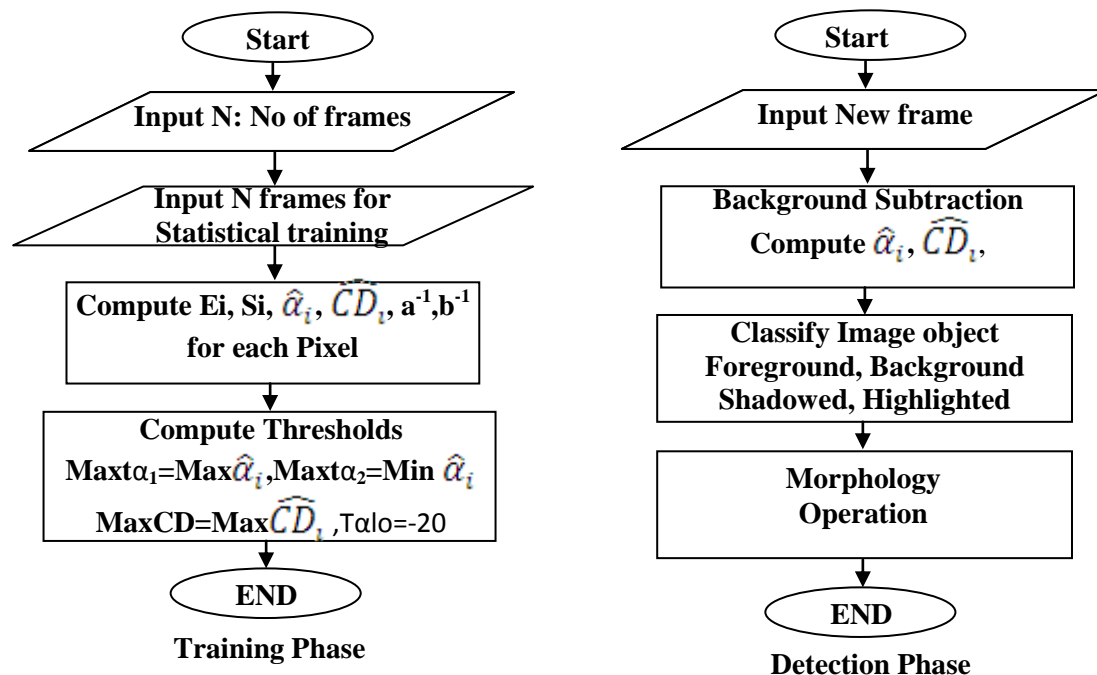
τ_{cd}: is a threshold selection that is calculated from normalized chromaticity distortion.

$\tau_{\alpha lo}$ = - 30, try and error, are considered to avoid misclassification for pixels that have darker (RGB) color. After classification is done, the shadow is detected and removed to obtain objects without distortion especially when the shadow casts to other objects. Then enters to final stage to improve the output frame.

3.4 Morphology operation:

Morphological image processing (or morphology) describes a range of image processing techniques that deal with the shape (or morphology) of features in an image Morphological to remove imperfections introduced during segmentation. It is a tool of extracting image components that are useful in the representation and description of region shape such as boundaries, skeletons etc. The goal of morphology operations are simplify image data, preserve essential shape characteristics and eliminate noise[12].

There are many morphological operations; the most basic morphological operations are dilation and erosion. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels from object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. Both dilation and erosion are produced by the interaction of a set called a structuring element with a set of pixels of interest in the image. The structuring element has both a shape of an object. In this paper, after the detection of an object, morphology operation used to eliminate the noise of the image produced as shown in figure 6.

Start

Input N: No of frames

Input N frames for Statistical training

Compute Ei, Si, $\hat{\alpha}_i$, $\widehat{CD_i}$, a^{-1},b^{-1} for each Pixel

Compute Thresholds
Maxtα$_1$=Max$\hat{\alpha}_i$,Maxtα$_2$=Min $\hat{\alpha}_i$
MaxCD=Max$\widehat{CD_i}$,Tαlo=-20

END

Training Phase

Start

Input New frame

Background Subtraction
Compute $\hat{\alpha}_i$, $\widehat{CD_i}$,

Classify Image object
Foreground, Background
Shadowed, Highlighted

Morphology Operation

END

Detection Phase

Figure(6): Flow Chart of the Modifed Horprasert algorithm

4. Experimental Result:

This section demonstrates the comparison between the Modified Horprasert Model and the Horprasert method using several video sequences for both indoor and outdoor scenes. Figure (7) shows the result of run various sequences in Matlab program, Red pixel depicts the object, black pixel depicts the background and the blue pixel for shadow, Figure (8) shows the time it takes for twenty frame to construct Background Model and Figure(9) illustrates the time it takes for classified image.

The accuracy is measured using PCC (percentage of correct classification), Figure (10) displays the percentage of errors between the Horprasert model and modification Horprasert implementations.

PCC combines all four values:

$$PCC = \frac{TP+TN}{TP+TN+FN+TN}$$ …………………….(10)

TP: which counts the number of correctly detected foreground pixels.
FP :which counts the number of background pixels incorrectly classified as foreground.
TN: which counts the number of correctly detected foreground pixels.
FN :which account for the number of foreground pixels incorrectly classified as background.

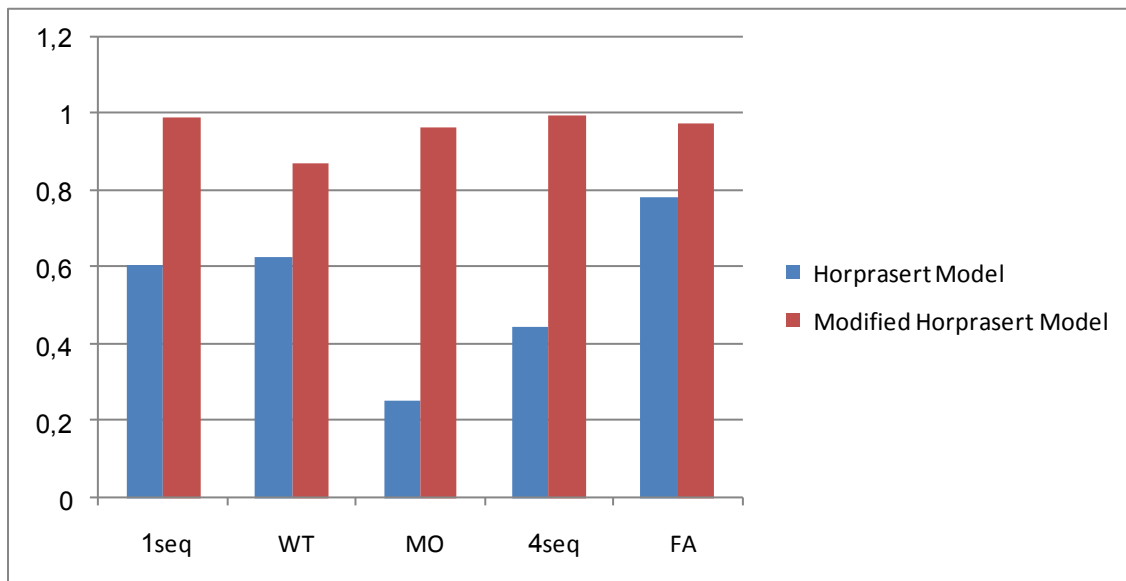|	Man Walking	Waving Trees	Moving Object	Car Parking	Foreground Aperture
A					
B					
C					
D					
E					
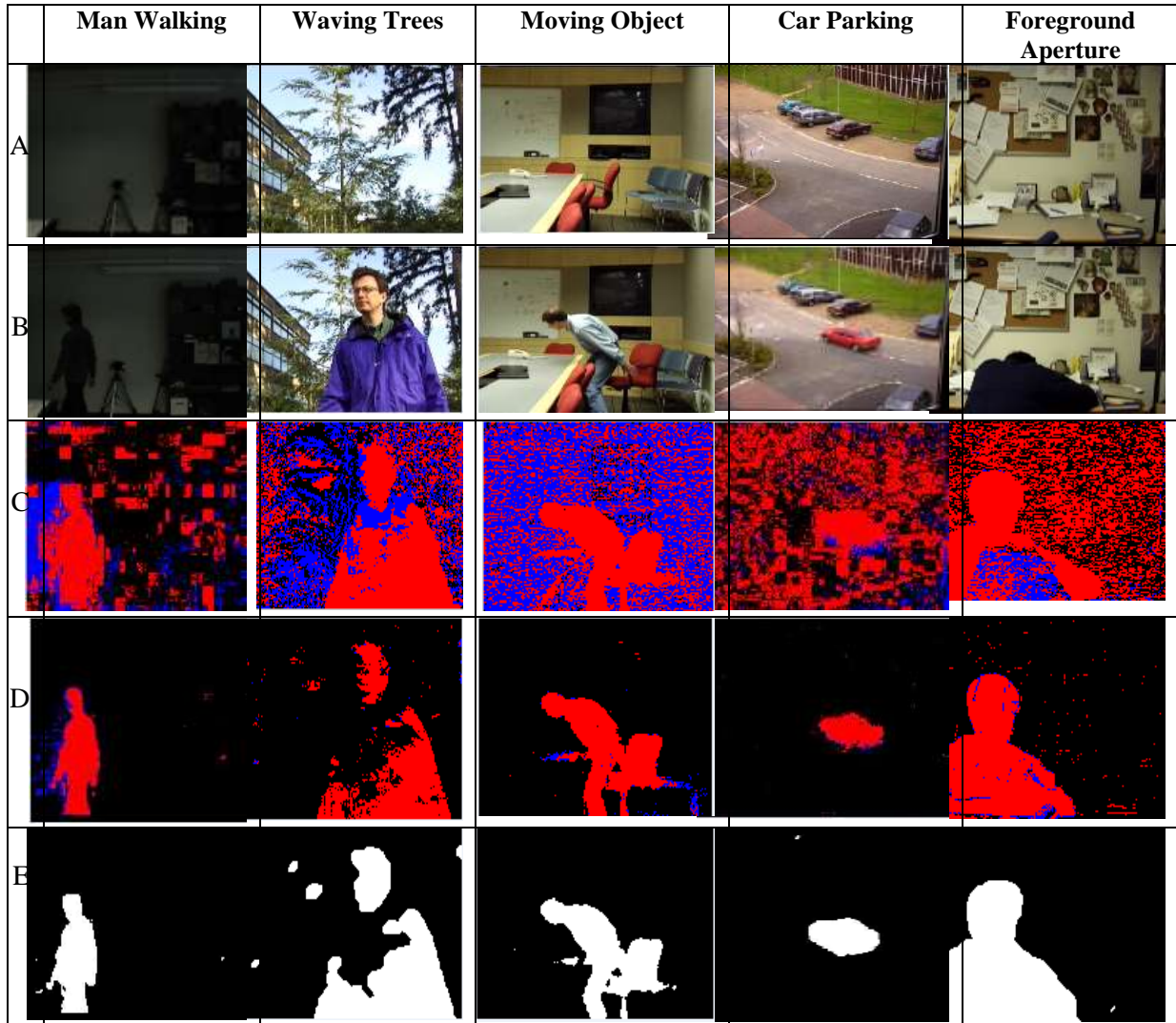
Figure(7): Show the difference from static thresholds selection and the proposed method: (A) Background model; (B) Test image ; (C) Test image in pervious run ; (D) Test image in proposed modified; (E) Proposed test image after morphology.
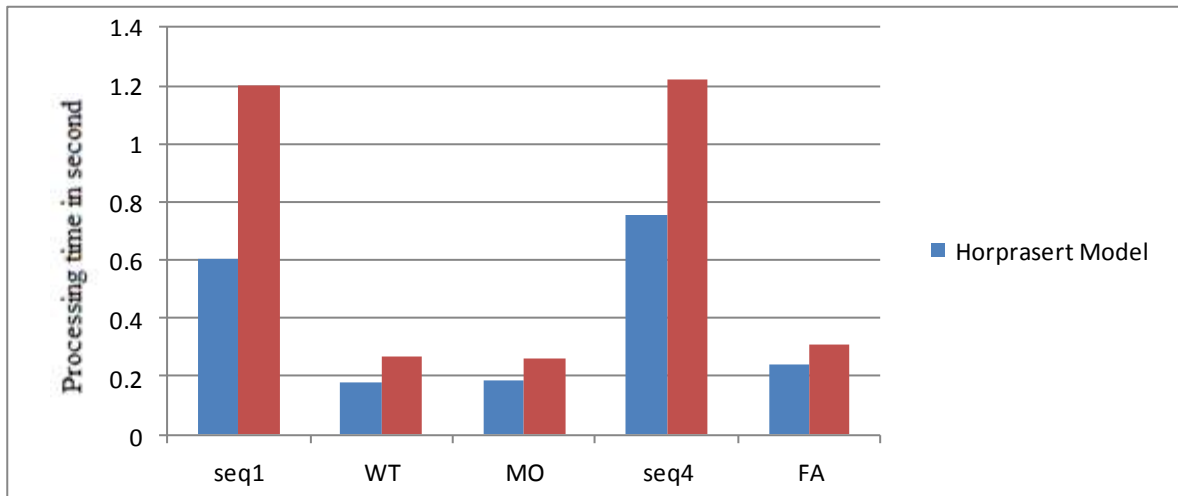
Figure(8) The time required during training background model for the previous sequences: MW- Man Walking; WT- Waving Tree; MO- Moving Object;CP-Car Parking; FA-Foreground Aperture.
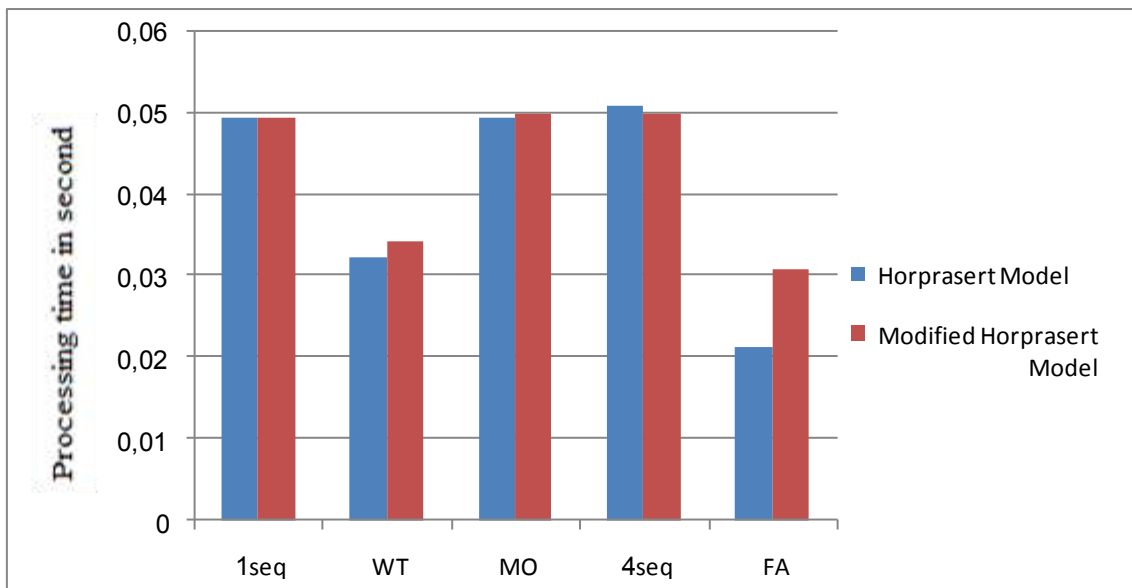
Figure(9): The time it takes to classify a test image: MW- Man Walking; WT- Waving Tree; MO- Moving Object;CP-Car Parking; FA-Foreground Aperture.

Figure(10): The comparison of (PCC) between the Modified Horprasert Model and Horprasert Model for various video sequences: MW- Man Walking; WT- Waving Tree; MO- Moving Object;CP-Car Parking; FA-Foreground Aperture.

5. Conclusion and discussion

This paper produced modification of Horprasert Model for detection new object. The modification includes the background model stage and thresholds selection stage; as well as the morphology operation. All these stages made the detection of the new objects better than previously run of Horprasert Model. The experimntal result show that the PCCs(percentage of correct classification) ranges from 87.12 % to 99.7 % for different input.

References:

[1]P. Chiranjeevi and S. Sengupta, " Robust detection of moving objects in video sequences through rough set theory framework", available in IVSL, Published in: Journal Image and Vision Computing, Vol. 30, Pages 829-842, November, 2012.

[2] T. Horprasert, D. Harwood and L. Davis, "A Statistical Approach for Real-Time Robust Background Subtraction and Shadow Detection", In Proceedings of the IEEE Frame-Rate Applications Workshop, In Proc. IEEE ICCV, Vol. 99, pp. 1-19 Key: citeulike:3878246, Kerkyra, Greece, 21–21 September 1999.

[3] Y. Su, "Surveillance Video Sequence Segmentation Based on Moving Object Detection", available in IVSL, Computer Science and Engineering, 2009. WCSE '09. Second International Workshop on, vol. 1, Page(s): 534 – 537, Qingdao, 28-30 Oct. 2009.

[4]Y. Zhu," Background Subtraction and Color Clustering Based Moving Objects Detection", available in IVSL, Information Engineering and Computer Science, 2009. ICIECS 2009. International Conference on, Wuhan, Page(s):1 – 5, 19-20 Dec. 2009.

[5] C. Stauffer and E. Grimson, "Learning patterns of activity using real time tracking", available in IVSL, IEEE Trans. Pattern Anal. Mach. Intell., Vol. 22, no. 8,pp.747-757, Aug. 2000.

[6] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams", available in ivsl, IEEE Trans. Pattern Anal. Mach. Intell.,vol. 25, no. 10, pp. 1337-1342, Oct. 2003.

[7]R. Jenifa, C.Akila and Dr.V.Kavitha, "Rapid Background subtraction from video sequence", available in IVSL, Computing, Electronics and Electrical Technologies (ICCEET), International Conference on Kumaracoil, pp. 1077 – 1086, 21-22 March 2012.

[8] H. Wang and D. Suter, "Background subtraction based on a robust consensus method", available in IVSL, Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, Hong Kong, Page(s): 223 – 226, 20-24 Aug. 2006.

[9] M. Heikkila and M. Pietikainen, "A Texture-Based Method for Modeling the Background and Detecting Moving Objects", IEEE Transactions on image processing Vol. 28 , Issue: 4, P. 657 - 662 April 2006.

[10]P. Varcheie, M. Sills-Lavoie, and G.-A. Bilodeau, "A multi scale region based motion detection and background subtraction algorithm", Sensors, Vol. 10, pp. 1041-1061, Jan. 2010.

[11] H. H. Lin, T. L. Liu, and J. c. Chuang, "Learning a scene background model via classification", IEEE Signal Process. Mag., Vol. 57, no. 5, pp. 1641-1654, May 2009.

[12] M. Goyal, " Morphological Image Processing", IJCST Vol. 2, Iss ue 4, Oct .- Dec. 2011