e-ISSN: 2458-8156

Journal Homepage: http://dergipark.gov.tr/ejens

EJENS

European Journal of Engineering and Natural Sciences

DergiPark AKADEMİK

CODEN: EJENS

Moving Object Detection Using an Adaptive Background Modeling in Dynamic Scene

M.Fatih Savas[1*],HuseyinDemirel[1],Bilgehan Erkal[1]

[1]*Karabuk University, Faculty of Engineering, Electric-Electronic Engineering Department, 78050, Karabuk, Turkey*

**CorrespondingAuthor email: muratfatihsavas@gmail.com*

Publication Info

Paper received:
29 May 2016

Revised received:
15 October 2016

Accepted:
01 March 2017

Abstract

Determination of moving foreground objects in dynamic scenes for video surveillance systems is still a problem can not be resolved exactly. In the literature; pixel-based, block-based and texture-based methods have been proposed to solve this problem. The method we propose will be block-based method which can be applied to real time in dynamic scenes. We have created non-overlapped blocks with the averages the pixels in the gray level. We used this average value to generate the background model based on a modified original KDE (Kernel Density Estimation) method. To determine the moving foreground objects and to update background model, we use an adaptive parameter which is determined according to the number of changes in the state of this pixel during the last N frames. Performance evaluation of the proposed method is tested by background methods in literature without applying post-processing techniques. Experimental results demonstrate the effectiveness and robustness of our method.

Key words

Background modeling, Moving object, Background update, Adaptive threshold

1. INTRODUCTION

Background modeling, which aims to classify each pixel as foreground and background, is one of the important processes of video surveillance system. It is a difficult process to determine moving objects in outdoor scenes such as waving trees, rippling water, moving shadow, illumination changes, camera jitters. Most of the studies conducted in recent years aim to minimize the effect of these changes which make difficult to detect efficient moving foreground. Wren et al [1]. have updated the parameters of the function regularly by creating a background model with single gaussian structure. However, this method is inadequate for dynamic background conditions. Gaussian mixture model (GMM) was proposed by Friedman and Russell [2]. and they updated the estimates by using incremental EM. In order to deal with complex scene changes, Stauffer and Grimson [3].expressed GMM with online k average approach. KaewTraKulPong and Bowden's [4].approach based on Mixture of Gaussians (MOG) improves adaptive background mixture model. To cope with background noise and illumination changes problems, Yan et al [5]. proposed a dynamic learning object determination which synthesizes the methods of background subtraction and adjacent frame difference.

Kim et al [6]. offered a structure called the codebook for real-time applications in one of the studies aimed at coping with the multimodal backgrounds. In this structure, by quantizing sample background values of each pixel into the codebook, he updated these codes in certain periods. In this structure, the codes which couldn't be reached for a long time are removed from the code table. To reduce the number of process, by using the color difference, Tu et al [7].modelled the background with boxed based codebook method. Although codebook structure offers an effective solution for dynamic background modeling, it is not fast while modelling. Maddalena et al [8]. proposed an adaptive learning rate in a neural network background modeling, a self-organizing method that does not contain a prior knowledge and model automatically background.

Elgammal et al [9]. proposed a non-parametric structure by the certain number of framers he accumulated for background modeling that expresses the color distribution of pixels with gaussian function . To eliminate the disadvantages of this structure, adaptive and less computational burden of parzen window structure is used. Ianasi, et al [10].proposed recursive density estimation with mean shift based mode to reduce the complexity of nonparametric kernel density estimation method. Tanaka et al [11]. proposed a fast PDF calculation function of structure by updating the PDF which partially estimated from the previous frame. By using HSV color values and gradient information, Park et al [12]. created the classification of background or foreground with Bayesian decision rule. Cuevas et al [16]. proposed a real-time spatial-temporal non-parametric method. To create background modeling in this method, by using both spatial and temporal data of pixels, they increased system's immunity to noise. Hoffman et al [13]. proposed the Pixel-Based Adaptive Segment (PBAS) structure by updating the background model which was created as non-parametric, with learning parameters.

Recently, Heikkila et al [10]. have proposed the texture-based background model by using local binary pattern. Although this method is effective against illumination changes, it is not strong in the uniform region. To eliminate the disadvantages of LBP method, which is inadequate in uniform regions.Yao and Odobez [14].created background modeling by RGB color feature.

2. PROPOSED METHOD

Memory burden and process time are the problems that must be solved in Real-time applications [23]-[24]-[25]. Lee and Park [15].modified the KDE model to solve this problem. In our study, we based on the modified KDE model for background model.

Moving objects such as sudden illumination change, waving trees, rippling water in the dynamic scenes create too much noise for the background model. To reduce the amount of noise and the computational burden, we connected the pixels by averaging pixels as nxn blocks. These pixel blocks are used for the background model and adaptive parameters. To build an adaptive parameter, we used the counter structure which adaptive Casares et al [16]. used in their studies. By this parameter, the foreground determination and the background update was made. While updating the model, each block is continually involved in background model. Then, normalization process was performed to equal the area to 1, which is under the model's probability function.

2.1. Background Subtraction Method

In background determination algorithms, authors mostly focused on the models such as efficient usage of storage space, reducing process time. Two problems must be solved to use the background model in real-time applications. These problems are; storage space and processing time. In real-time applications, to reduce the processing burden and to use the storage space efficiently, pixels are used as gray level and background model is updated as adaptive. But one of the disadvantages of processing at gray level is that it is open to the noise came up on stage.

A picture frame with the resolution of 160x120 size is made up of 19200 pixel. If these pixels are composed of RGB value, more time and more storage space is required. Processing on these pixels bring a lot of load on the system both in terms of time and the requirements for storage space. Therefore, to create a minimum level for processing time and storage space, the background model we propose was in gray level. However, in literature, to minimize the disadvantage of grayscale structure, kernel density estimation based model which offers an effective solution to background modeling was used.

Lee and Park [15] created the background model as pixel based. In our study, we put the average values created as 2x2 blocks in bins in background model. Instead of evaluating pixels' membership value for bins as 1 or 0, probability values are calculated according to the distance of bins from the center.

I(x,y)	*I(x,y+1)*
I(x+1,y)	*I(x+1,y+1)*

Figure 1. 2x2 block

$I(x, y)$represents the gray level value of each pixel in the image frame. Pixels in the image frame were joined by averaging as 2x2 non-overlap blocks. As computational, mean filter structure which requires less burden is used equation 1.

$$\mu(x, y) = \frac{1}{nxn}\sum_{x=1}^{n}\sum_{y=1}^{n} I(x, y) \qquad (1)$$

where μ (x, y), is the average value of the pixels in nxn block.

The following figure shows the change of the pixel at location (39,80) in the time interval 0-30 seconds and shows the gray level changes of the pixels averaged at location (39 + x,80 + y) in the time interval 0-30 seconds from the Wallflower dataset's wavingtrees test sequences. Here x,y={0,1}

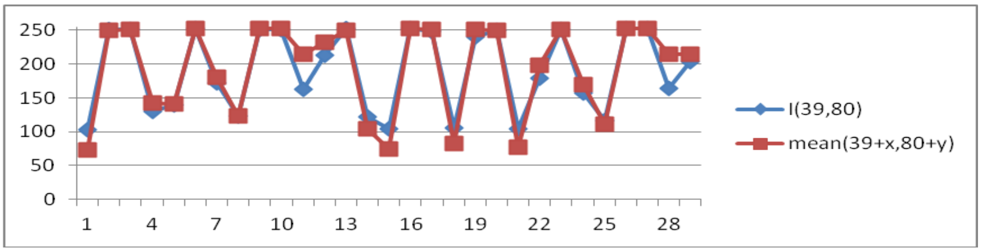


Figure 2.I(39,80) pixels and 39 + x ,80+ y block of pixels in the range of 0-50 times the change in gray level chart

The following equation 2.shows the function we used for the background model.

$$p_t(C_k) = \frac{1}{\sqrt{2\pi\left({B_d}/{2}\right)^2}}\, exp\left(-\frac{1}{2}\left(\frac{C_k-\mu_t(x,y)}{{B_d}/{2}}\right)^2\right)(2)$$

In equation 2, for each block, the histograms are formed in the width of B_d by bins.$C_k$is the central point of kth bin which belongs to nxn block pixel, $p_t(C_k)$ is the probability value of block pixel at time index t.

2.2. Adaptive Threshold Parameter

Determining the fixed threshold used for the determination of moving objects in dynamic scenes is difficult. If less values are selected as the parameter, the image may contain a lot of noise, if a large value is selected as the parameter, some values from the image may be lost. To overcome this problem, we created threshold as adaptive in our study. Adaptive threshold is created by counter structure counting the change in each pixel. This counter structure was used to reduce the computational burden.

In this structure, the parameter of $\beta_x, x \in \{1,2,3\}$ is the maximum value of a counter can count that determined by the user. $CC_n, n \in \{1,2,3\}$shows the number of counter belonging to a pixel. The value of a counter is incremented by 1 at change of each pixel's state. After counting the frame number to be counted, counter values are sequentially reset [16].The value of counters increase or decrease according to the rate of pixels' state changing. $\tau(x,y) = CC_1(x,y) + CC_2(x,y)+CC_3(x,y)$) is formed by the total value of counter of pixels. The parameter values of the pixels in moving regions take large τ value, τ values of the pixels in quasi-static area take small values. By this approach, at any time index t, we can determine the change number of pixels in past N frame.

2.3. Background Update

The following formulate is used for updating the background model. α is the fixed updating parameter. While updating , all C_k values aren't calculated, only bins interval $p_t(C_{(k-2)})$ and $p_t(C_{(k+2)})$ are updated by considering bins interval $p_t(C_{(k\pm2)})$ [15].The following equation 4.

$$p_t(C_{\boldsymbol{k}}) = \hat{p}_{t-1}(C_{\boldsymbol{k}}) + \left(\frac{1}{\alpha+100.\tau}\right)\frac{1}{\sqrt{2\pi\left({B_{\boldsymbol{d}}}/{2}\right)^2}}\, exp\left(-\frac{1}{2}\left(\frac{C_{\boldsymbol{k}}-\mu_t(x,y)}{{B_{\boldsymbol{d}}}/{2}}\right)^2\right)(4)$$

2.4. Foreground Detection

Determination of moving objects in dynamic scenes depends on a robust background modelling structure and the value of parameter to be determined. While determining the foreground, the distance between the bin center which has the largest possibility value at histogram and $\mu_t(x,y)$ is taken into consideration.Equation 5.

$$Distance|C_k - \mu_t(x,y)| \leq threshold + \tau(5)$$

If the absolute value difference between C_k, which has the largest possibility value at histogarm and the average value calculated at t time of $\mu_t(x,y)$ is bigger than threshold, foreground is determined. Equation 6.

$$I(x,y)\begin{cases} Foreground Distance|C_k - \mu_t(x,y)| \leq threshold + \tau \\ \qquad\qquad Background otherwise \end{cases}(6)$$

3. RESULTS AND DISCUSSION

We tested the performance of our method by wallflower [17]. and Li [18]. datasets. We compared our proposed method with Mixture Of Gaussian V1BGS [4], Pixel Based Adaptive Segmenter [13]. and T2FGMM_UM [22].methods from BGSLibrary which was created by Andrews Sobral [19]. Three different binary classification measurement methods were used while comparing methods. These methods are precision, recall and F-measure. When the recall and precision values are high, it shows that the performance is high. F-measure is the weighed harmonic average of recall and precision [20]-[21].Equation 7.

$$reacll = \frac{TP}{TP+FN} \qquad precision = \frac{TP}{TP+FP} \qquad F-measure = \frac{2.recall\ .precision}{recall\ +precision} \quad (7)$$

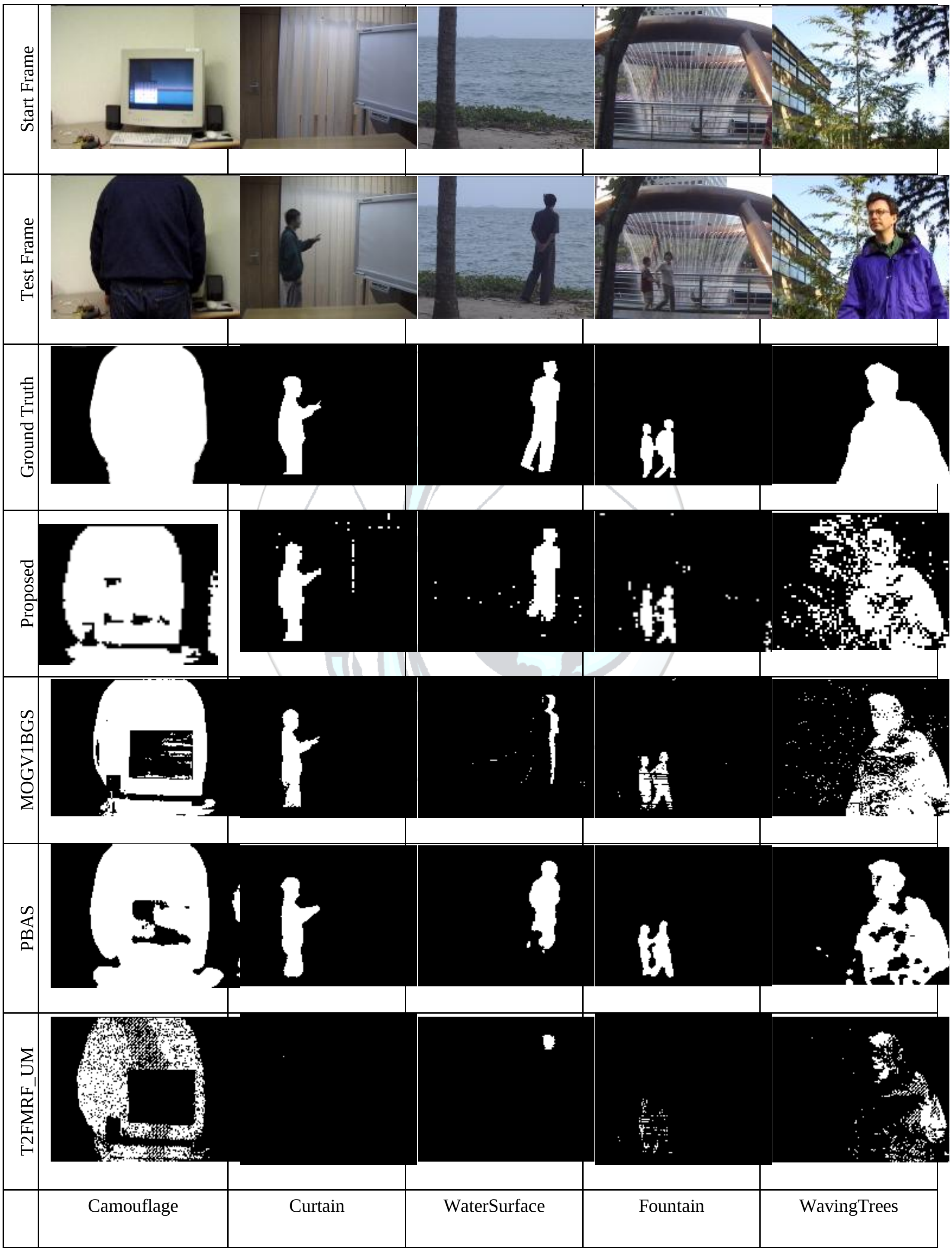
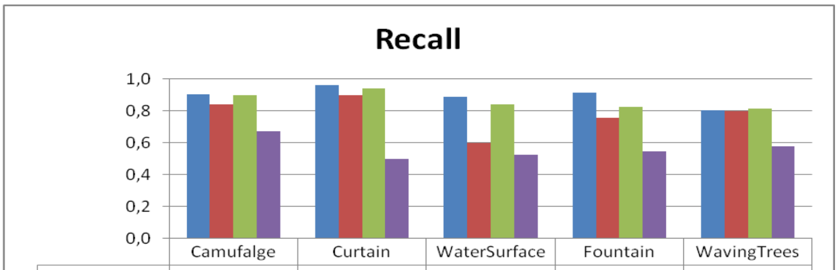
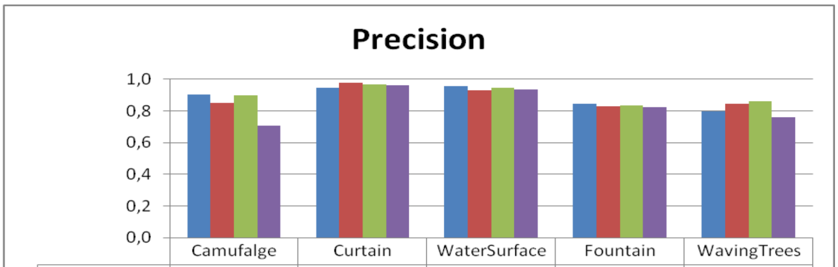


Figure 3.Experimental Results



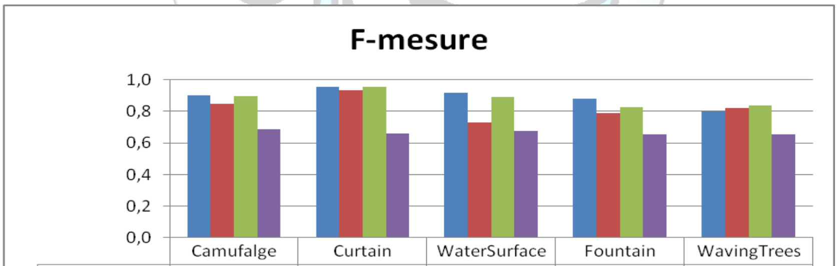
	Camufalge	Curtain	WaterSurface	Fountain	WavingTrees
Proposed	0,9031	0,9608	0,8877	0,9161	0,8043
MOGV1BGS	0,8414	0,8971	0,5988	0,7541	0,7980
PBAS	0,8966	0,9417	0,8407	0,8234	0,8142
T2FMRF_UM	0,6729	0,5003	0,5272	0,5460	0,5764

Figure 4.The recall results obtained with the proposed scheme and other methods for the Li and wallflower dataset.



	Camufalge	Curtain	WaterSurface	Fountain	WavingTrees
Proposed	0,9033	0,9452	0,9534	0,8465	0,7965
MOGV1BGS	0,8508	0,9792	0,9302	0,8274	0,8456
PBAS	0,8959	0,9672	0,9448	0,8324	0,8590
T2FMRF_UM	0,7065	0,9617	0,9338	0,8213	0,7570

Figure 5.The precision results obtained with the proposed scheme and other methods for the Li and wallflower dataset.



	Camufalge	Curtain	WaterSurface	Fountain	WavingTrees
Proposed	0,9032	0,9529	0,9194	0,8799	0,8004
MOGV1BGS	0,8461	0,9363	0,7286	0,7890	0,8211
PBAS	0,8962	0,9543	0,8897	0,8279	0,8360
T2FMRF_UM	0,6893	0,6582	0,6739	0,6559	0,6545

Figure 6.The f-measurement results obtained with the proposed scheme and other methods for the Li and wallflower dataset.

Recall, also known as detection rate, gives the percentage of detected true positives as compared to the total number of true positives in the ground truth, where TP is the total number of true positives, and FN is the total number of false negatives, which accounts for the number of foreground pixels incorrectly classified as background. Precision, also known as positive prediction, that gives the percentage of detected true positives as compared to the total number of pixels detected by the method, is generally used in conjunction with the recall. Where FP is the total number of false positives.

For five test videos, an evaluation was carried out between the models by calculating Recall, Precision and F measurement values. In test images, no post-processing technic application is performed. Test sequences are camouflage, curtain, water surface, fountain and waving trees. Some of these test sequences include continuously changing dynamic background, some of them include changing dynamic background at certain times. These test sequences have the frames with a size of 160 x 120 and 160x128.

The method we proposed in the measurements carried out with these test sequences, obtained robust results in most test data. When the waving trees scene was compared with other methods, the amount of noise was seen much more. These noises can be reduced according to the threshold parameter chosen in the process of determining foreground. On the contrary, reducing the noise can cause incorrectly classifying of some pixels belonging to foreground object. In

T2FGMM_UM method, the amount of noise was reduced, but pixels belonging to foreground object were evaluated as background. In water surface test sequence, as it is seen in the test data obtained via MOGV1BGS method, if the background model adapts to the scene very fast, the continuance of stopped objects moving at the scene on the screen will be less. The disadvantage of Pixel Based Adaptive Segmenter method which gets similar performance values as our proposed method, it contains too many parameters.

4. CONCLUSIONS

In our study, we created an effective model for real time applications by modifying KDE structure of Lee and Park [15]. By dealing with background model as gray level blocks nonoverlap, we reduced both the memory area requirement for model and the time needed for processing pixels. We determined the adaptive parameter by using counter structure of Casares et al. [16], and by determining the condition changing number of pixels in N frame that past at any t time. Thus we created an effective background modeling at grey level in dynamic scenes. The weak point of this method is, as a single band was used for background modeling it can't exactly discriminate the foreground and the background. As background model was created block-based, image belongs to foreground object is coarser than the pixel based methods. In future work, we are going to use this method for RGB band. Test data shows that, our study carried out for single color band is valid.

Acknowledgments

This work/study is supported by Karabük University graduate degrees thesis project. Project no is KBU-BAP-16/1-DR-167.

REFERENCES

[1]. C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, " Pfinder: Real-time tracking of the human body," *IEEE Trans. Pattern Anal*. 1997, vol.19, pp.780–785.

[2]. N. Friedman and S. Russell, "Image segmentation in video sequences: a probabilistic approach," *In Proceedings of the 13th Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann Publishers Inc.*, 1997, pp. 175–181.

[3]. C.Stauffer and W.E.L, Grimson, "Adaptive background mixture models for real-time tracking," *In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Santa Barbara, CA, USA, June 1998, pp. 246–252.

[4]. P,KaewTraKulPong and R,Bowden, " An improved adaptive background mixture model for real-time tracking with shadow detection," *In Proceedings of 2nd European Workshop on Advanced Video-Based Surveillance Systems(AVBS01)*, Sept. 2001.

[5]. R.Yan, X.Song and S.Yan, "Moving Object Detection Based on an Improved Gaussian Mixture Background Model", *Computing, Communication, Control, and Management*, 2009. CCCM 2009. ISECS International Colloquium on, pp. 12–15.

[6]. K.Kim, T.H,Chalidabhongse, D.Harwood and L.Davis, "Real-time foreground-background segmentation using codebook model," *Real-Time Imaging*, vol.11, pp.172–185, 2005.

[7]. Q.Tu, Y. Xu and M.Zhou, "Box-based Codebook Model for Real-time Objects Detection," *Intelligent Control and Automation, 2008*. WCICA 2008. 7th World Congress on,2008, pp. 7621-7625.

[8]. L.Maddalena and A.Petrosino,"A self-organizing approach to background subtraction for visual surveillance applications," *IEEE Trans. Image Process 2008*, pp. 1168–1177.

[9]. A.Elgammal, R.Duraiswami, D. Harwood and L.S.Davis, "Background and foreground modeling using nonparametric kernel density estimation for visual surveillance*," Proc. IEEE 2002*, vol.90, pp.1151–1163.

[10]. C.Ianasi, V. Gui, C. Toma and D.Pescaru, " A fast algorithm for background tracking in video surveillance, using nonparametric kernel density estimation*," FactaUniversitatis (NIS) Ser.: Elec. Energ*. ,vol.18, pp.127–144,2005

[11]. T.Tanaka, A.Shimada, D. Arita and R.I Taniguchi, "A Fast Algorithm for Adaptive Background Model Construction Using Parzen Density Estimation", *Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on London. 2007*, pp.528 – 533

[12]. J.G.Park and C.Lee, "Bayesian rule-based complex background modeling and foreground detection, "*Optical Engineering*, vol.49 No.2, 2010.

[13]. M. Hofmann, P. Tiefenbacher, G. Rigoll, "Background segmentation with feedback: The pixel-based adaptive segmenter," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*,2012, pp. 38–43.

[14]. J.Yao and J.M.Odobez, "Multi-Layer Background Subtraction Based on Color and Texture", *IEEE Conference on Computer Vision and Pattern Recognition, 2007.* ,pp. 1–8.

[15]. J.Lee and M.Park, "An Adaptive Background Subtraction Method Based on Kernel Density Estimation," *Sensors* , vol.12, pp.12279-12300,2012.

[16]. M.Casares, S.Velipasalar and A.Pinto, "Light-weight salient foreground detection for embedded smart cameras," *Computer Vision and Image Understanding*, pp. 1223–1237, 2010.

[17]. Wallflower dataset web site.[Online].Available: http://research.microsoft.com/en-us/um/people/jckrumm/WallFlower/TestImages.htm

[18]. Li dataset web site.[Online]. Available: http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html

[19]. Andrews Sobral web site.[Online]. Available: https://github.com/andrewssobral/bgslibrary

[20]. S.Brutzer, B.Hoferlin and G.Heidemann, "Evaluation of background subtraction techniques for video surveillance," in*: Proceedings of the CVPR, IEEE*, 2011, pp. 1937–1944.

[21]. EC Funded CAVIAR project/IST 2001 37540, web site.[Online]. Available: http://homepages.inf.ed.ac.uk/rbf/ CAVIAR/

[22]. Fida El Baf, T.Bouwmans, B.Vachon, "Type-2 Fuzzy Mixture of Gaussians Model: Application to Background Modeling," *Advances in Visual Computing*, vol. 5358, pp. 772-78,2008.

[23]. Demirel. H, *SayısalElektronik*, 2nd ed., İstanbul, BirsenYayınevi, 2015.

[24]. Topaloğlu. N, *Mikroişlemcilerve Assembly Dili, 6nd ed*., SeçkinYayıncılık, Ankara, 2015.

[25]. S.Savaş, N.Topaloğlu, B. Ciylan, "Analysis of mobile communication signals with frequency analysis method". *Gazi University Journal of Science*, Vol.25,No.2, pp.447-454,2012.