Noise Reduction using MRF and Block-Based Background Modeling in Dynamic Scenes Input

M.Fatih SAVAŞ† and Hüseyin DEMİREL††,

The Electrical/Electronic Engineering Department of Karabuk University, Turkey

Summary

Identifying the moving foreground object in dynamic scenes and making the analysis of video sequences accurate and powerful is an important process for video surveillance systems. Environmental factors such as environmental noises and sudden light changes are the main factors of the degradation of the background model. Complex algorithms are needed to create a strong background against these factors. In this study, We increased the noise immunity of the background model exposed to environmental noise by applying Markov random field (MRF) to block-based modified KDE (Kernel Density Estimation). We also reduced the storage space requirement with the KDE structure we created in blocks. Thus we have increased the applicability of this structure to a real-time structure.

Key words:

Kernel Density Estimation, Markov random field, Background modeling, Adaptive threshold parameter

1. Introduction

Identify moving objects in video sequences are the basic tasks of vision applications. [1]. There are some basic algorithms for determining moving objects. These methods are background modeling, optical flow and frame difference methods. The frame difference method is a simple method that attempts to determine motion objects by taking differences of sequential video frames.[2] However, this method is quite open to environmental noise. This method is mostly used in vision systems as a complement to other structures [3], [4]. The optical flow method attempts to determine motion in video frames by providing dense displacement fields containing motion vectors for each pixel location of the reference scene [5]. Because this method involves intensive mathematical algorithms, it is not suitable for real-time structures with limited storage space. Various optimization techniques have been applied to implement this method in visual systems [6]. The most preferred method in video vision systems is background modeling. In this method, the background is created using various frames by using video frames [7]. It is decided whether the color or the values of the pixel's intensity should be included in the model by using various updating methods. The main purpose of these methods is to try to determine effectively the moving object.

It is difficult to create a suitable background model for detecting moving objects in video sequences containing sudden light changes, dynamic backgrounds, shadows and video noises. Methods such as mean filter and median filter, Eigen-background, Mixture of Gaussian (MOG) and kernel density estimation (KDE) have been proposed in the literature to overcome these situations [8], [9]. Wren et al [10], modeling the value of a pixel's intensity as a Gaussian distribution, have proposed a parametric structure. This method is insufficient for scenes with dynamic background. Stauffer and Grimson proposed [11], [12] the Gaussian mixture model (GMM) to be able to cope with scenes in dynamic background. In order to overcome this disadvantage of parameter setting, online EM-based algorithms are proposed. Kim et al. [13] created a model structure by quantizing the pixel values in the codebook according to the frequency of pixel values in the video frames and extracting the pixel samples with reduced occurrence frequency in the model from the codebook. The disadvantage of this method is that it takes a long time to construct the model. To overcome the difficulty of determining the initial parameter value in parametric background models, Elgammal et al [14], suggested a nonparametric structure by using kernel density estimation (KDE). In this method, the background model is created by calculating the probability values of the pixels in the memory frames .To create the background model in this method, the number of frames in memory is a disadvantage. In this method, this disadvantage is solved by using recursive density estimation (RDE) methods.

The results of the methods used for background subtraction proposed in the literature are insufficient due to only the background model. Depending on the environmental noise and the weakness of the background model, the model may contain too many false positives or false negatives. Generally, these problems are solved using filtering methods or various algorithms. The MRF model, which has an important place in the literature from these methods, is widely used for noise reduction labeling and optimization problems [16] [17]. We used the MRF

Manuscript received January 5, 2017
Manuscript revised January 20, 2017

method to reduce false positives and false negatives in our work.

In this study, we concentrated on two points; block-based generation of the background model and reduction of false positives and false negatives with model adaptation of the MRF. We used the background model proposed by Lee and Park [15] to create the background model. We implemented Lee and Park's pixel-based proposed model in nxm block structure in order to reduce the system memory space requirement and processing time. nxm blocks are created in non-overlapping structure by taking the average of the pixels. We also detected a more robust foreground object using the Markov random field technique to determine moving objects in scenes with dynamic backgrounds.

2. Proposed Method

Every object detection system needs a background subtraction algorithm. The stronger the background algorithm, the greater the achievement in determining the moving object. Nonparametric model's achievement of object detection is quite high because of the ability of handling multi-modal background distribution. Lee and Park [18] proposed pixel-based adaptive kernel density estimation by adapting nonparametric structure to histogram structure.

Spatially, pixels closed to each other shows similar properties. By taking into account this feature, we have created the blocks in the frame as nxm non-overlapping blocks. So we performed block-based background adaptation to Lee and Park's [18] pixel-based histogram model. The following formula represents the probability value of the pixel-based structure of Lee and Park [18] for a pixel. $p_t^d(C_k)$ is defined as

$$p_t^d(C_k) = p_{t-1}^d(x) + \frac{1}{G_t\sqrt{2\pi}(B_d/2)^2} exp\left(-\frac{1}{2}\left(\frac{C_k - x_t^d}{B_d/2}\right)^2\right) \quad (1)$$

Here, B_d is the width of the bins of histograms created for the background model, x_t^d is the color information of the pixel at time t, C_k is the central point of each bin in the histogram, k is the number of bins in the histogram and G_t is the learning parameter and sigmoid function. d indicates the color of the pixels [18]. Here we have created the background model, using the gray level intensity value of each pixel, with a mean filter structure that does not contain too much computational cost. $\mu(x,y)$ can be defined as

$$\mu(x,y) = \frac{1}{nxm}\sum_{x=0}^{n}\sum_{y=0}^{m} I(x,y) \quad (2)$$

Here $I(x,y)$ indicates the intensity value at the gray level of the pixel at (x,y); n and m indicate the block sizes. Thus, the value of x_t in equation 1 turns into $\mu(x,y)$ value.

$$p_t(C_k) = p_{t-1}(x) + \frac{1}{\sqrt{2\pi}(B/2)^2} exp\left(-\frac{1}{2}\left(\frac{C_k - \mu_t(x,y)}{B/2}\right)^2\right) \quad (3)$$

In design, we implemented these blocks in 2x2 size.

3. Adaptive Threshold Parameter

Threshold parameter value is an important variable that determines the amount of noise included in the background model. If this parameter is a fixed number, it makes it difficult to determine moving objects in scenes with dynamic backgrounds. If the threshold parameter value is selected to a small value, the background model contains a lot of noise; if a larger value is selected, some values of the moving object are lost. In this study, in order to create the adaptive threshold parameter value, we based on the structure of the counters that Casares et al. [19] used to identify the foreground object. With this structure, we designed the threshold parameter value as a variable parameter. At the designed counter-based adaptive threshold parameter structure, there are three counters that increase the content of each pixel as foreground and background in different states. These counters perform counting up to a predefined number value. The counters are reset by sequence after a predetermined number of frames. Thus the adaptive threshold parameter value consists of the sum of the change counts of the counters for the past N frames. $\tau(x,y)$ is defined as

$$\tau(x,y) = CC_1(x,y) + CC_2(x,y) + CC_3(x,y) \quad (4)$$

Here, $CC_x(x,y)$ represents the number of state changes that each counter has. In this study, each counter holds the number of state changes for 25 frames. We adapted the threshold parameter value to the running average to avoid sudden state changes that would occur as a result of the reset of the counter values. See Eq.5 ;

$$\tau_t(x,y) = (1-\alpha)\tau_{t-1}(x,y) + \alpha\tau_t(x,y) \quad (5)$$

The total value of the counters is created the threshold parameter value. By using three separate counters for the threshold parameter value, we can easily determine the number of state changes of the pixel in past n frames at t time. We used this parameter for both the variance value in the MRF structure and the background update value.

4. Background Update

Updating the background model is done with the following formula. Here, α is a constant parameter used to provide the fit of the model. All C_k values are not calculated while updating the model, the number of bins in the range $P_t(C_{(k-2)})$ and $P_t(C_{(k+2)})$ is updated only considering the bins in the range $C_{(k\pm2)}$ [18]. $P_t(C_k)$ can be defined as

$$p_t(C_k) = \hat{p}_{t-1}(C_k) + \left(\frac{1}{(\alpha + 10*\tau(x,y))}\right)\frac{1}{\sqrt{2\pi(B_d/2)^2}}\exp\left(-\frac{1}{2}\left(\frac{C_k - \mu_t(x,y)}{B_d/2}\right)^2\right) \quad (6)$$

5. Markov Random Fields Used for Noise Reduction

Markov random fields are at the forefront of methods still popular in past day-to-day image processing algorithms. In our work we used Markov random fields to reduce noise. While performing this method for our study, Dumontier et al [22] performed the simulation based on the workings of the study.

In Markov fields, the first label field is created with the background model we created. There are two labels in the created S picture at time t. These;

$$e_s = \begin{cases} a & \text{if } s \text{ belongs to a moving object} \\ b & \text{If } s \text{ is a static background} \end{cases} \quad (7)$$

Thus, observation field can defined as

$$O_s = \left|C_k(t) - \mu(x,y)\right| \quad (8)$$

$C_k(t)$ is the center value of the bin with the highest probability value of the background model at time t. $\mu_t(x,y)$ is the spatial mean values of nxm size blocks at time t. Observation field is $o = \{O_s, s \in S\}$. Label field is $\hat{E}(t)$.

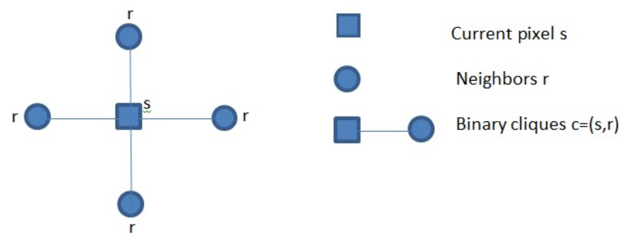

Fig. 1 Clique structure.

The most appropriate label value is found by applying iterated conditional modes which are the simplest, greedy algorithm to the $\hat{E}(t)$ label field.

5.1. Energy and Clique Function

F is used to sum the two energies. F_m is the energy of the label field. F_a is the energy of the label field with the observation field. Total energy is shown below. Equation 9,

$$F(o_s, e_s) = F_m(e_s) + F_a(o_s, e_s) \quad (9)$$

$F_m(e_s) = \sum_{c\in C} V_c(e_s, e_r)$ here c represents the cliques in the positional neighbors. C are all clicks. If we define the click potentials with the following equation;

$$V_c(e_s, e_r) = \begin{cases} -\beta & if \quad e_s = e_r \\ +\beta & if \quad e_s \neq e_r \end{cases} \quad (10)$$

The relationship between the observation field and the label field is defined by the following equation.

$$O_s = \gamma(e_s) + n \quad (11)$$

$$\gamma(e_s) = \begin{cases} 0 & if \quad e_s = b \\ C_k & if \quad e_s = a \end{cases} \quad (12)$$

Here, n is Gaussian noise with zero mean and variance σ^2. C_k was evaluated as mean value. $F_a(o_s, e_s)$ If the energy function is defined by the following equation, equation 13,

$$F_a(o_s, e_s) = \frac{1}{2\sigma^2}\sum_{s\in S}\left[O_S - \gamma(e_s)\right]^2 \quad (13)$$

Calculation of σ^2 variance value causes a computational cost on the system. Instead of this value, we used the value that we created for the threshold parameter.

$$F_m(1) + F_a(1) < F_m(0) + F_a(0) \quad (14)$$

In the case of the above equation Label 1 is evaluated. Otherwise label 0. We arrange this formulation

$$2F_m(1) < F_a(0) - F_a(1) \qquad (15)$$

Equation 13, in the case of $F_a(0)$

$$F_a(o_s, e_s) = \frac{1}{2\sigma^2}\sum_{s\in S}\left[O_s\right]^2 \qquad (16)$$

In the case of Equation $F_a(1)$;

$$F_a(o_s, e_s) = \frac{1}{2\sigma^2}\sum_{s\in S}\left[O_s - C_k(t)\right]^2 \qquad (17)$$

If we take the difference between the equations 16 and 17; equation 18,

$$\frac{C_k(t)}{2\sigma^2}(2O_S - C_k(t)) > 2F_m(1) \qquad (18)$$

$$F_a(0) - F_a(1) = \frac{C_k(t)}{2\sigma^2}(C_k(t) - 2\mu_t(x,y)) \qquad (19)$$

$$\gamma(e_s) = \begin{cases} 255 & if\ 2F_m(1) < \frac{C_k(t)}{2\sigma^2}(C_k(t) - 2\mu_t(x,y)) \\ 0 & otherwise \end{cases}$$

(20)

6. Experimental Results

We tested the performance of our method by wallflower [20]. and Li [21]. datasets. These test data are 160x120 and 160x128 sizes. We did the comparison between the method we proposed and the proposed + MRF. Thus, the MRF's noise reduction effect on the model is measured. No post processing was applied in performance measures.

Peak signal-to-noise ratio (PSNR) measurements were used in the comparison of the [proposed method] and [Markov + proposed] method. PSNR usually gives information about the relationship between the two pictures. When the PSNR is calculated, the original picture is used as the first input and the method outputs as the second input. The size of this ratio gives information about the similarity or the quality of the two images. In this calculation, A represents original images, B represents the image obtained as the output of the method, x represents the width of the image, and y represents the height of the image. The MAX value in the PSNR calculation process represents the largest value the pixel can take. For 8 bits this value is 255.

$$MSE = \sum_{i=1}^{x}\sum_{j=1}^{y}\frac{(\left|A_{ij} - B_{ij}\right|)^2}{xy} \qquad (21)$$

$$PSNR(db) = 10\log(\frac{MAX^2)}{MSE} \qquad (22)$$

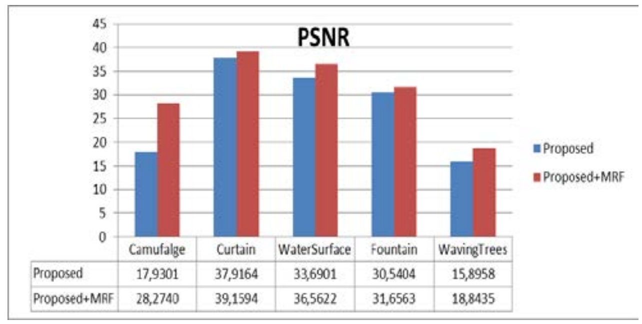

Fig. 2 Proposed and proposed +MRF the peak signal to noise ratio results for Li and wallflower data sets

When we evaluate the results, the method we proposed is adapted to the MRF and the amount of noise is reduced. Thus, a more robust model was created in dynamic scenes. Experimental results figure 3;

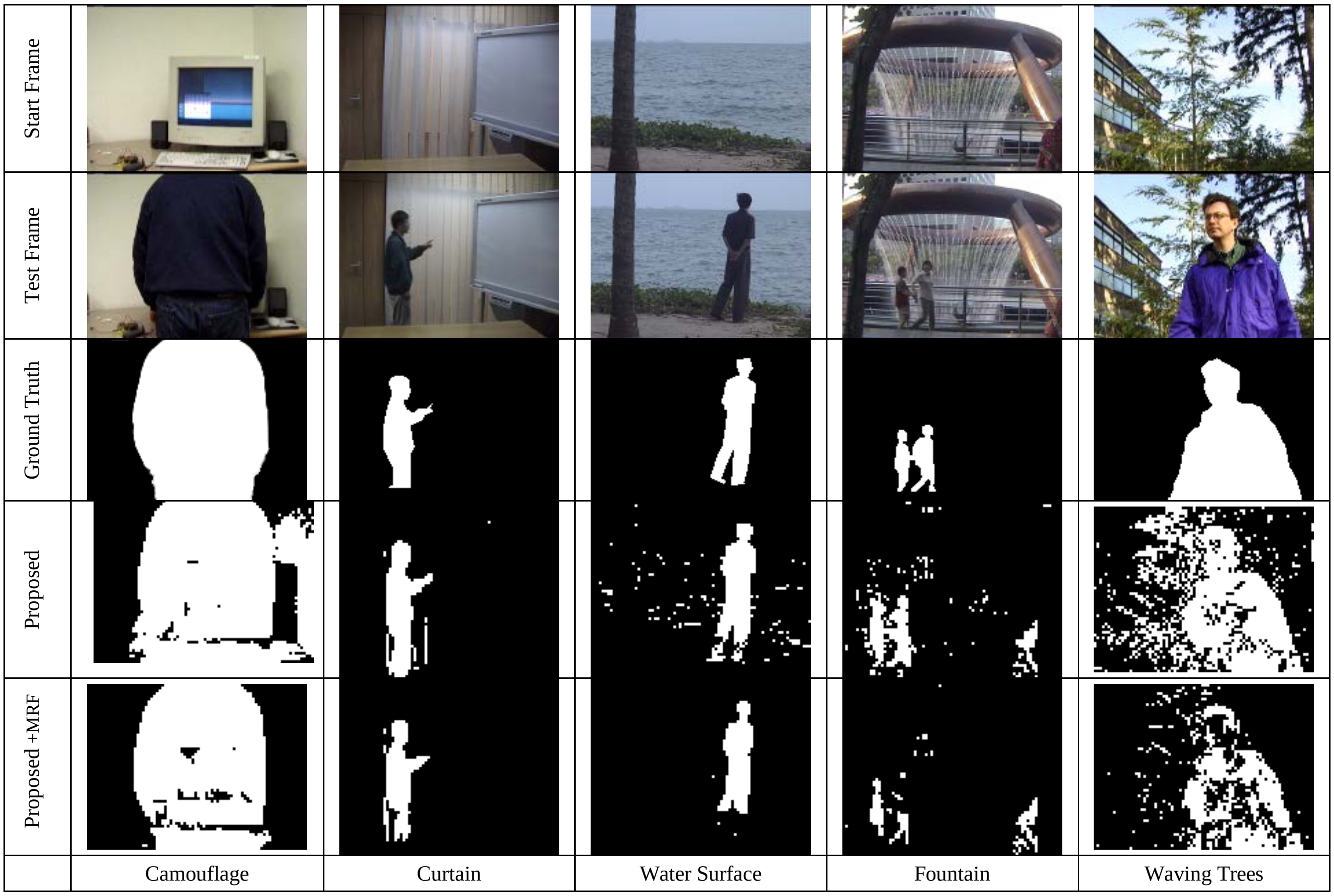

Fig.3 Experimental Results

7. Conclusions

In this study, we proposed a block-based adaptive histogram-based structure by grouping pixels at the gray level to reduce the memory space requirement of the system and shorten the processing time required to process the pixels. We have adapted the Markov random fields to our model to reduce the amount of noise of the generated background model. We used the adaptive parameter value to update both the background model and the variance value of the Markov model. Thus we reduced the computational load. Test results show that by adapting the Markov method to the background model we have modified, the amount of noise in dynamic background scenes is significantly reduced.

8. Acknowledgments

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

9. References

[1] [1] K.Yang, Z.Cai and L.Zhao, "Algorithm Research on Moving Object Detection of Surveillance Video Sequence," Optics and Photonics Journal, pp. 308-312, June 2013.

[2] H.Liu, J.Dai, R.Wang, H.Zheng and B.Zheng, "Combining background subtraction and three-frame difference to detect moving object from underwater video," IEEE Conference Publications, OCEANS 2016 - Shanghai, pp. 1-5, April 2016.

[3] J,Huang, Y.Gao, Y.Li, J.Chen and Jun Jiang, "Algorithm research on motion area segmentation based on improved frame differences," 2015 IEEE Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) , pp.1152-1155, March 2015.

[4] X.Han, Y.Gao, Z. Lu, Z.Zhang and D. Niu, "Research on Moving Object Detection Algorithm Based on Improved Three Frame Difference Method and Optical Flow," 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), pp. 580 – 584, Sept. 2015.

[5] T.Bengtssona, T. McKelveya and K.Lindströmb, "On robust optical flow estimation on image sequences with differently exposed frames using primal–dual optimization." Image and Vision Computing, vol.57, pp.78–88, Jan. 2017.

[6] T. Camus, "Real-time quantized optical flow," IEEE Conference Publications, Proceedings of Conference on Computer Architectures for Machine Perception, pp. 126 - 131 , Sept. 1995.

[7] Y.Xu, J. Dong, B.Zhang and D.Xu, "Background modeling methods in video analysis: A review and comparative evaluation," CAAI Transactions on Intelligence Technology, vol.1, pp.43-60, Jan. 2016.

[8] M.H.Sigari, N.Mozayani and H. R. Pourreza, "Fuzzy Running Average and Fuzzy Background Subtraction: Concepts and Application," IJCSNS International Journal of Computer Science and Network Security, vol. 8 no. 2 pp. 138-143, Feb.2008.

[9] K.Takahara, T.Toriu and Thi Thi Zin, "Making Background Subtraction Robust to Various Illumination Changes," IJCSNS International Journal of Computer Science and Network Security, vol. 11, pp. 241-248, March 2011.

[10] C. Wren, A. Azarhayejani, T. Darrell and A.P. Pentland, "Pfinder: Real-Time Tracking of the Human Body," IEEE Trans. Pattern Anal. Mach Intell. vol. 19 ,no.7, pp. 780–785, July 1997.

[11] C. Stauffer and E. Grimson, "Adaptive background mixture models for realtime tracking", IEEE Int. Conf. Comput. Vis. Pattern Recognit., vol. 2, pp.246-252, June 1999.

[12] C. Stauffer and E. Grimson, " Learning patterns of activity using real-time tracking, " IEEE Trans. Pattern Anal. Mach. Intell. vol.22 , no.8, pp. 747-757, Augt. 2002.

[13] K.Kim, T.H. Chalidabhongseb , D.Harwooda and Larry Davisa, "Real-time foreground–background segmentation using codebook model," Real-Time Imaging, pp.172–185, 2005.

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

[15] D.H.Parks and S.S.Fels, "Evaluation of Background Subtraction Algorithms with Post-processing," AVSS '08 Proceedings of the 2008 IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance,pp.192-199, Sept. 2008.

[16] S. Geman and D. Geman, "Stochastic relaxation, gibbs distributions and the bayesian restoration of images,"TPAMI, pp. 564–584, 1984

[17] S. Z. Li, "Markov random field modeling in image analysis," 3rd ed. Springer-Verlag New York, Inc., 2009

[18] J.Lee and M.Park, "An Adaptive Background Subtraction Method Based on Kernel Density Estimation," Sensors, vol.12, pp.12279-12300, Dec .2002.

[19] M.Casares, S.Velipasalar and A.Pinto, "Light-weight salient foreground detection for embedded smart cameras," Computer Vision and Image Understanding, vol.114, pp. 1223–1237, Nov .2010.

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

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

[22] C.Dumontier, F.Luthon and J.P.Chams, "Real time implementation of an MRF-based motion detection algorithm on a DSP board," Digital Signal Processing Workshop Proceedings, 1996., IEEE,pp.183-186, 1996.

Murat Fatih SAVAŞ, graduated from Sakarya University Electronic and Computer Education Department in 2002, received master degree from Sakarya University in 2005 and continues his education as a PhD student in Karabuk University. Now, he is a teacher in Kastamonu Vocational High School and is engaged in Background subtraction methods.

Hüseyin DEMİREL, graduated from Gazi University Electronic and Computer Education Department in 1997. After that he received MS and PhD from Gazi University in 1999 & 2010 respectively. As an assistant professor in Karabuk University Electrical and Electronic Engineering Department, he is working in digital electronics and thermoelectric modules.