2017 International Conference on Data Management, Analytics and Innovation (ICDMAI)
Zeal Education Society, Pune, India, Feb 24-26, 2017

Block-Based Quantized Histogram (BBQH) for Efficient Background Modeling and Foreground Extraction in Video

Satyabrata Maity
A.K.Choudhury School of Information Technology
University of Calcutta
Kolkata, India
Email: satyabrata.maity@gmail.com

Amlan Chakrabarti
A.K.Choudhury School of Information Technology
University of Calcutta
Kolkata, India
Email: amlanc@ieee.org

Debotosh Bhattacharjee
Department of Computer Science and Engineering
Jadavpur University
Kolkata, India
Email: debotoshb@hotmail.com

***Abstract*—This paper proposes an efficient way of background modeling and elimination for extracting foreground information from the video, applying a new block-based statistical feature extraction technique coined as Block Based Quantized Histogram ($BBQH$) for background modeling. The inclusion of contrast normalization and anisotropic smoothing in the preprocessing step, makes the feature extraction procedure more robust towards several unorthodox situations like illumination change, dynamic background, bootstrapping, noisy video and camouflaged conditions. The experimental results on the benchmark video frames clearly demonstrate that $BBQH$ has successfully extracted the foreground information despite the various irregularities. BBQH also gives the best F-measure values for most of the benchmark videos in comparison with the other state of the art methods, and hence its novelty is well justified.**

***Index Terms*—Video Application; Background modeling; Foreground extraction; Block-Based Quantized Histogram; Pixel based statistics; Global statistics; Region based statistics;**

I. Introduction

The content of any video has grossly classified in two distinct categories, specifically, (i)*Background*: the static and temporally consistent part, which is almost unchanged throughout the scene and (ii). *Foreground*: the dynamic and temporally varying part with respect to location, which defines the activity within the scene. Thus, the background information among the video frames embrace the redundant information and increase processing overhead in video analysis. Hence, the exclusion of this static information makes the analysis more efficient and purposeful. The information contents in a video are mainly the activity (the changing information due to complete an action) of the foregrounds. Thus, background modeling has a greater impact in automatic video analysis system, since a proper background modeling and elimination will provide effective foreground information for different intelligent video applications like activity recognition, object tracking, motion capture, video surveillance [10], [12], [13], [17] etc.

The density of background information (static redundant part) among the frames of a video is much greater than the same of the foreground (active part), and the dynamic property of foreground distinguishes it from the foreground. Hence, the statistical measurement of the different information in terms of intensity values can easily distinguish those two different levels of information. The third kind of things, which creates the disturbance to distinguish classify fist two kinds of information is termed as video irregularities like a dynamic background, video noise, change in illumination conditions, bootstrapping, intermittent object motion etc.

There are three main ways of computing the statistical measurement, specifically (i). *Pixel based statistics*: the statistics of intensities among the pixels of the same location in a number of consecutive frames are tested and the intensity with the highest probability is treated as the background intensity of the corresponding pixel. But, this kind of strategies are very much noise prone as the single pixel can be affected easily. (ii). *Global statistics*: the global feature of each frame are measured statistically, and background feature is computed after amalgamation. But, these types of methods can neglect the local changed in information. (iii). *Region based statistics*: the region based feature are extracted and combined to get the regions of background frame. This is robust one among all the strategies, but the region segmentation needs some sophisticated algorithms with higher costs. Thus, the proposed methods combine the advantages of above three procedures and take the block-based approach, which doesn't require any cost for segmentation, but robust towards local changes and unnecessary noises. Moreover, it can handle the critical issues like the dynamic background, foliage etc.

In the proposed approach, background modeling is performed through the extraction of the block-based quantized histogram ($BBQH$) features from the first t number of frames of any scene. Blocks are the small rectangular parts of a frame and the quantized histogram provides the number of different pixel groups present in that frame. $BBQH$ features of the corresponding block represent the number of different intensity levels present in the blocks in terms of predefined bins instead of individual intensity levels. Static information is estimated over all selected frames using the temporal density of the bins with respect to the blocks, which decides the features of the background frame of the video. Thereafter, the background information is eliminated from the current frame to detect the foreground information. Our work can handle several challenging situations like ghosting effect due to slow moving foreground, noisy conditions, bootstrapping due to the presence of foreground from the very first frame of the video, camouflaged conditions etc. The key contributions, in this research work, can be summarized as follows:

1) An efficient video preprocessing phase has been implemented, which nullifies the illumination effect to a greater extent as well as performs suitable denoising.
2) A new block based feature extraction technique coined as BBQH, which can able to model the background robustly.
3) The proposed method can handle several irregularities in the video data like ghosting effect, noisy conditions, bootstrapping, camouflaged conditions etc.

II. Related works

Information is inversely proportional to the occurrence probability. On the other hand, the probability of occurrence of the background pixels are far more than that of foreground pixels and those nearly remain unchanged between the frames throughout the scene. Background modeling is done by exploiting this fundamental property of a video. Several techniques have been proposed in [1], [2], [3], [4], [6], [7], [9], [10], [11], [12], [14], which model the background based statistical measure of the static pixels (the intensity values of the pixels remaining unchanged between the frames) in the previously observed frames. The background modeling method can be categorized into three different approaches proposed in [7] namely, (a) Pixel-level background modeling: The static part of the initial frames of any scene is modeled using statistical approaches over pixels intensity values.(b) Region-level background modeling: It is based on the evaluation of the local texture around each pixel instead of the single pixel for reducing the effects of variations in illuminations, and (c) Frame-level background modeling: It includes issues like sudden global changes of the image brightness while estimating present background. In [6], the authors used the region level background subtraction approach based on the local shape of small image regions where the region is described by the center pixel and local self-similarity descriptor. The similarity in the intensity changes among pixels is considered in [8], where the pixels are classified into several clusters based on the similarity of their intensity changes. So this strategy is a mix-up of both region and pixel level modeling. In [9] and [10], we find the general approach of background modeling using pixel history i.e. mixture of Gaussians, for modeling each pixel with some modifications. In [9], the authors tried to make the model adaptive and in [10] a modified Gaussian mixture model is employed for better results.

III. Proposed Method

This method tries to model the temporally consistent information to extract the varying information from each of the frame, which will later use for further processing. The proposed technique is divided into three main steps, namely *(i). preprocessing (ii). background modeling*, and *(iii) foreground extraction*. Fig. 1 demonstrates a complete flow diagram of the steps involved in the proposed methodology.

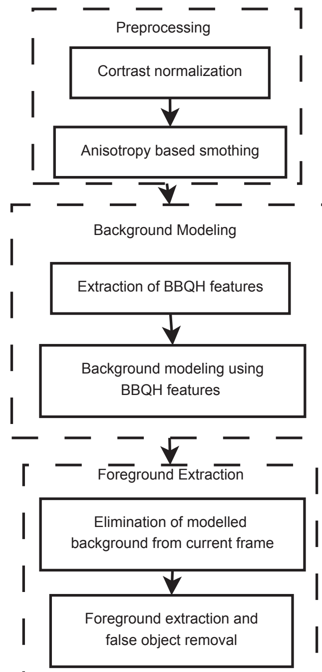

Fig. 1. Flow diagram of the proposed method

A. Preprocessing

The goal of this step is to reduce the effect of noise and illumination effect so that background modeling can be taken place more proficiently.

1) Contrast Normalization: The proposed approach extracts information using a statistical measurement of intensity values among video frames. So, the change in contrast among consecutive frames may cause bigger issues in terms of information changing. Thus, this is an important preprocessing step towards maintaining the consistency of information among frames. Black has low intensity (luminance) and white has high luminance. The contrast has been described as (L_{MAX} - L_{MIN})/(L_{MAX} + L_{MIN}), where L is the luminance value L_{MAX} and L_{MIN} are the highest and lowest luminance value respectively. So, the illumination of light has a direct impact on the contrast and the brightness of any captured frame as the images with low contrast have lower visibility as compared to a normal image as discussed in [15], [16], [18]. Thus, this step tries to reduce the effect of illumination by normalizing the contrast of the frames. ll and ul are the lower and upper limit of normalization, hv and lv are the higher and lower values of the current frame, and p_{in} and p_{out} are the input and output pixel values. The normalization is done using Eq. 1.

$$P_{out} = (P_{in} - ll)\left(\frac{ul - ll}{hv - lv}\right) + ll \tag{1}$$

2) Anisotropy in video processing: The extraction of foreground area form each frame does not require the minute details of the inner region, but the minute details in the region may cause some difficulties in order to estimate actual foreground information. Hence, the anisotropic filer is used. Unlike noise smoothing filters, generally perform smoothing operations in the whole image area (irrespective of low and high frequency region), the anisotropic filter proposed by Perona and Malik [13] provides the facility of smoothing of the low frequency regions and emphasizes the edge regions using Eq. 2.

$$\frac{\partial I(x, y, t)}{\partial x} = div\left[g\left(\parallel \nabla\perp \parallel\right)\nabla\perp\right] + a \tag{2}$$

where, $\parallel \nabla\perp \parallel$ is the gradient magnitude, and $g\left(\parallel \nabla\perp \parallel\right)$ is an "edge-stopping" function. This function is chosen to satisfy $g(x) \to 0$ when $x \to \infty$ that the diffusion is "stopped" across edges.

B. Background modeling

Histogram of an image describes the probability measurement of all intensities. On the other hand, a human can perceive a region with respect to visual similarities i.e. intensity values of a region may not be same but closure to each other. We are incorporating this fuzziness in our proposed methodology, and taking the quantized value to represent a group of pixels instead of each individual intensity values. We are dividing the whole intensity range into 16 fixed sized bins as there is no such visual difference in 16 different intensity levels of the same region. This may create problems for low contrast images, but we have performed contrast normalization before, and which stretches the intensities into the range of $0 - 255$. Besides this, the histogram of the whole image may neglect the effect of local change; we divide the frames into the uniform sized block to overcome this problem. Sometimes, it may happen that two visually different frames may have similar histograms [28]. An idea of the quantized histogram is depicted in Fig. 2, where the same blocks of three consecutive frames are shown in the first row, the general and quantized histograms are shown in the middle and last row respectively. The nature of the pictures are mostly similar for both the cases, but the complexity of later one is lesser than that of the former one.

1) Extraction of BBQH features: Let us assume that each frame containing $(l\times m)$ pixels is divided into non-overlapping blocks of size $(p \times q)$ pixels, as given in Eq. 3. There is a tradeoff between block size and method's efficiency. The union of all the blocks a frame recreate the frame itself, and we have $m1 \times n1$ number of blocks for each individual frames. If the block size is very small with respect to frame size, it becomes noise prone as it will be affected by a negligible change. On the other hand, the bigger size of blocks may overlook some important changes due to foreground movements. The processing time of the methods is inversely proportional to the size of the block. We consider the block size with respect to the resolution of the frame, but not less than 8×8.

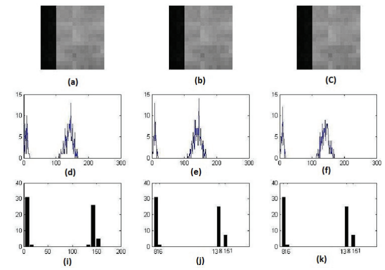

Fig. 2. Quantized histogram (a) - (c) are the same spatial block of three consecutive frames (d) - (f) are the histogram of the corresponding block (g) - (i) are the quantized histogram of corresponding block

BIN is the size of a quantized bin and B_{ij} is one of the blocks in the location of i^{th} row and j^{th} column. We have a feature vector $\mathbb{Q}_{ij}$ feature vector for the block including the representation value (weighted average), and the occurrence probability of each individual bin for the corresponding block. Eq. 4 and Eq. 5 describe the way to compute the feature vector of a block. We set a threshold of 0.03% to reduce the effect of unwanted noise, i.e. if the occurrence probability of any bin is less than the prescribed threshold, we set zero in the feature values of the corresponding bin. If β is the feature vector holding the features of a frame, it must be the integration of all $\mathbb{Q}$ in the horizontal and vertical direction as shown in Eq. 6. Continuing the same procedure in consecutive frames, we will easily get the features of all frames.

2) Background Modeling using BBQH Features: Initial t number of frames are considered to model the background feature for the scene. All the features of the corresponding frames are merged using the formula given in Eq. 7. Now MAX is the highest average occurrence probability, and the threshold for modeling the background is computed using Eq. 1. We are taking maximum probable bin divided by $\sqrt{2}$ as the most of the real world phenomenon maintain Gaussian distribution. The bins are considered as background bins if the corresponding bin greater than TH as given in Eq. 9. we are taking, t as the frame rate of the video.

After initializing the background features of a scene, we compute the feature extraction and merging in each interval of $t/2$ number of frames and merge them with the old background frame to make the procedure adaptive. If more than $2/3$ of the blocks disagree with the old values, we will start the background initialization procedure again as this case is presumed to be the change in the scene.

$$Fr = \bigcup B_{ij} \tag{3}$$

Where, $1 \leq i \leq m1(= l/p), 1 \leq j \leq n1(= n/q)$,and $B_{ij} \cap B_{xy} = \phi,\ \forall i \neq x, and\ j \neq y$

$$\mathbb{Q}_{(l,1)} = \left(\sum_{i=1}^{b} X_i\right)/b,\ \ \{X \in P \mid k = P_j/BIN\} \tag{4}$$

$$\mathbb{Q}_{(l,2)} = b/(k^2) \tag{5}$$

$$\beta(i,j).fea = \mathbb{Q} \tag{6}$$

$$\beta_{ij(l,1)} = \left(\sum_{x=1}^{t} (\mathbb{Q}_{ijx(l,1)}) \right) / t \tag{7}$$

$$TH = MAX/\sqrt{2} \tag{8}$$

$$BAK(\beta_{ijl}) = \begin{cases} 1 & If(\beta_{ij(l,1) \geqslant TH}) \\ 0 & Otherwise \end{cases} \tag{9}$$

C. Foreground Extraction

1) Elimination of Modeled Background from Current Frame: We have block based background features for a scene, which are looking like a vector of $0's$, and $1's$. Zeroes mean the corresponding bin is not the background and vice-versa. Thus, to do the segmentation to find the foreground, we are checking comparability of the intensity values of the corresponding block, and the feature values of corresponding bin. We are leveling the foreground with the values $'1's'$ and $'0's'$ as depicted by Eq. 10.

$$Frg(i,j) = \begin{cases} 1 & IF, (Frg(i,j)/BIN) = 0\, in\, BAC(\beta_{ijl} \\ 0 & otherwise \end{cases} \tag{10}$$

2) Foreground Extraction and flase Object Removal: After the extraction of a foreground, unwanted or wrongly classified areas are eliminated by removing the objects below a certain threshold. A noise removal technique is used to reduce the false foreground by removing all the foreground area below a certain threshold TH_{fg}, which is set to be 0.6% of the image size. The threshold can be scaled according to the situations and the size of the required foreground objects.

D. Complexity Analysis

The algorithmic complexity of the proposed method can be defined with the help of analyzing each of the individual modules separately as follows. Suppose, we have N numbers of pixels in a frame, for contrast normalization, we need a single iteration. But in the case of anisotropy based smoothing, we need $4-6$ iterations. So, the time complexity is linear in the case of preprocessing. For extracting $BBQH$ features from a frame, if we have n numbers of blocks and each of which containing p numbers of pixels and $N = p \times n$, for generating the histogram, we need a single iteration. Each block is represented by the feature vector of length $x << N$. The statistical measurement for modeling the background is done by exploiting these feature vectors over k consecutive frames. So the worst case complexity of the proposed method is $O(N \times k) < O(N^2)$.

IV. Experimentation

To evaluate the performance of our proposed methodology, we have experimented with some of the existing benchmark video datasets. A brief description of the datasets and the related experimental results are presented in the next part of this section.

A. Datasets and Metrics

We have used two video datasets namely changed detection (DS1) [5] and SABS (DS2)[27]. These datasets incorporate the various form of irregularities like illumination changes, camouflage, noisy video, dynamic background etc.

1) Changed detection dataset (DS1) [5]:: We have used Baseline video of changed-detection datasets for our experiment and all the given results are tested on baseline database. The following are the video files available in the datasets are a) Highway, b) Office, c) Pedestrian and d) Pets2006.

2) SABS dataset (DS2) [27]:: This dataset contains several conditions of urban traffic scenarios. The following are the video files available in the datasets are a) Basic: Basic traffic scenario, b) Camouflage, c) Darkening, d) No-camouflage, and e) Noisy nights.

The performance metrics are evaluated using the parameters like TP (True Positive), TN (True negative), FP (False Positive) and FN (False Negative), which are again extracted using the ground truth of the corresponding video frame given in the same data-sets. The metrics are:

- RE (Recall) : TP / (TP + FN)
- PR(Precision) : TP / (TP + FP)
- SP (Specificity) : TN / (TN + FP)
- FPR (False Positive Rate) : FP / (FP + TN)
- FNR (False Negative Rate) : FN / (TP + FN
- PWC (Percentage of Wrong Classifications) : 100 * (FN + FP) / (TP + FN + FP + TN)
- F-Measure : (2 * Precision * Recall) / (Precision + Recall)

TP is the response of exact foreground by applying some strategy, thus the number of TP reflects the count of accurate extraction of foreground. On the other hand, FN represents the number of mis-classified background. So, recall (RE) represents the extracted versus actual foreground of the frame. FP represents the number of mis-classified foreground. Precision (PR) represents actual extracted foreground versus total extracted foreground.

We computed all the performance metrics on data sets DS1 and DS2 and compared those with several other techniques of the related research. In the case of DS1, we compare it with respect to all the performance metrics but in the case of DS2, we only compare the results on the basis of F-measure which the only available metric for the other related works. Although the computation of F-measures needs every parameter and it is the combinations two important metrics like precision and recall.

B. Results and analysis

This subsection briefs the results of the various experiments performed for the verification of our algorithm and the analysis of the same. Fig. 3 shows the results of background modeling from first 30 frames of the corresponding video. It shows $frame1, frame5, frame10, frame20, frame30$ and the modeled background frame of the corresponding video. The background frames contain the blocks occurring most of the times in the same location among the first 30 frames of the video. In two videos shown in Fig. 3, the foreground object (human in both the cases) is eliminated while modeling the background.

We have computed the values of seven parameters (mentioned in 4.1) for our proposed approach on videos of $DS1$ and compared them with that of the other state of the art methodologies. The comparison result is shown in Table. I. The average rank of a method is the mean of the individual rank with respect to each of the parameters. The larger values in the case of parameters RE, SP, F-measure, and PR are signified better the proof of better efficiency, whereas the smaller value determines better accuracy in case of FPR, FNR, and PWC. The proposed methodology based on BBQH feature achieved the best results in the case of SP, FPR, PWC and F-measure among all the methodologies shown in Table. I. Moreover, the average rank of the BBQH is 1.333, which is the best among the list, and the second average rank is held by GMM(SG)[11], which is 3. Thus, we can infer that BBQH clearly outperforms the other state of the art methods.

In Table. II, we show the comparison results of BBQH with that of the other state of the art methods as applied on DS2 videos. Each row in Table. II provides F-measure values for the different videos having irregularities after applying the corresponding method of column-1. The best result in each of the columns is shown in boldfaced. The results show the superiority of BBQH, which gives the best F-measure values in the case of Bootstrap, noisy night and no camouflage condition. Thus, we can conclude that the proposed method performs better with respect to the other state of the art techniques irrespective of irregularities.

Contrast normalization and anisotropy based smoothing operation reduce the effect of illumination and intra-regional interference respectively. The block-based processing helps us to reduce the ambient background motion, and unwanted noise as it takes the feature of the block instead of a single pixel. The adaptive threshold selection method for background feature enables the proposed method to extract the background eminently. The capability of automatic updation of background in regular interval helps us reducing the effect of ghosting, bootstrapping, and camouflage to some extent. All these attributes of the proposed method make the procedure robust.

V. Conclusion

In this paper, we have proposed a novel methodology, named as BBQH, for background modeling of the video and subsequent foreground extraction. We have incorporated illumination normalization and noise smoothing so that it can handle changes in lighting condition and intra-regional noises. Experimental results involving benchmark dataset show that BBQH features can accurately model background of any scene and also outperforms another state of the art background modeling and elimination techniques in terms of various performance measures. In future, we wish to improvise BBQH so that it can act adaptively, and thus making it well equipped for modeling moving background.

VI. Acknowledgement

First author acknowledges the DST inspire fellowship fund from the Department of Science and Technology(DST), Government of India for supporting this research work.

References

[1] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, C. Rosenberger. "Comparative study of background subtraction algorithms", Journal of Electronic Imaging, SPIE, vol. 19, 2010

[2] W. H. Liao, J.H. Kuo, "Sleep monitoring system in real bedroom environment using texture-based background modeling approaches", J. Ambient Intelligence and Humanized Computing 4(1): 57-66 (2013)

[3] Y. Zhai, M. Shah, "A General Framework for Temporal Video Scene Segmentation", ICCV 2005: 1111-1116

[4] T. Ko , S. Soatto and D. Estrin, "Background subtraction on distributions", Proc. Eur. Conf. Comput. Vis., vol. 3, pp.276 -289 2008

[5] N. Goyette, P. M. Jodoin, F. Porikli, J. Konrad, and P. Ishwar, "changedetection.net: A new change detection benchmark dataset, in Proc. IEEE Workshop on Change Detection (CDW12) at CVPR12", Providence, RI, 16-21 Jun., 2012., http://changedetection.net/

[6] J. P. Jodoin, G. A. Bilodeau, N. Saunier, " Background subtraction based on Local Shape", arXiv:1204.6326v1

[7] O. Strauss, D. Sidib, W. Puech "Quasi-continuous histogram based motion detection",Technical Report, LE2I, 2012

[8] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection", in proc of Workshop on Advanced Video Based Surveillance Systems, 2001

[9] Y. Nonaka, A. Shimada, H.Nagahara, R. Taniguchi."Evaluation Report of Integrated Background Modeling Based on Spatio-temporal Features", in proc of IEEE Workshop on Change Detection, 2012

[10] S. Yoshinaga, A. Shimada, H. Nagahara, R. Taniguchi, "Background Model Based on Intensity Change Similarity Among Pixels", the 19th Japan-Korea Joint Workshop on Frontiers of Computer Vision, pp.276-280, 2013.01

[11] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking", inProc. Int. Conf. on Computer Vision and Pattern Recognition, Vol. 2, IEEE, Piscataway, NJ (1999)

[12] Riahi, D., St-Onge, P.L., Bilodeau, G.A. (2012). RECTGAUSS-Tex: Block-based Background Subtraction, Technical Report, cole Polytechnique de Montral, EPM-RT-2012-03

[13] Perona, Pietro, and J. Malik, "Scale-space and edge detection using anisotropic diffusion." Pattern Analysis and Machine Intelligence, IEEE Transactions on 12.7 (1990): 629-639.

[14] P. Gorur and B. Amrutur, Speeded up Gaussian mixture model algorithm for background subtraction, in Proceedings of the IEEE International Conference on Advanced Video and Signal Based Surveillance, pp. 386391, Klagenfurt, Austria, August 2011.

[15] Z. Liu, E. Zavesky, B. Shahraray, D. C. Gibbon, A. Basso, "Brief and high-interest video summary generation", evaluating the AT and T labs rushes summarizations. TVS 2008: 21-25

[16] T. Bouwmans, F. E. Baf, B. Vachon, "Background Modeling using Mixture of Gaussians for ForegroundDetection - A Survey", Recent Patents on Computer Science 1: 3. 219-237, (2008) Pages 609-617

[17] T. Ko, S. Stefano, and E. Deborah, "Background subtraction on distributions." Computer VisionECCV 2008. Springer Berlin Heidelberg, 2008. 276-289.

[18] A. Krutz, "Motion-based object segmentation using sprites and anisotropic diffusion." Image Analysis for Multimedia Interactive Services, 2007. WIAMIS'07. Eighth International Workshop on. IEEE, 2007.

[19] N. J. B. McFarlane, and C. P. Schofield. Segmentation and tracking of piglets in images. Mach. Vis. Appl. 8(3):187-193 (1995)

[20] N. M. Oliver, B. Rosario, A. P. Pentland, "A Bayesian Computer Vision System for Modeling Human Interactions " , Journal, IEEE Transactions on Pattern Analysis and Machine Intelligence archive Volume 22 Issue 8, August 2000, Page 831-843

[21] Kyungnam Kim, Chalidabhongse, T.H. ; Harwood, D. ; Davis, L.,Background modeling and subtraction by codebook construction, ICIP '04. 2004 International Conference on Image Processing, 2004. (Volume:5)

[22] Liyuan Li, Weimin Huang, Irene Y. H. Gu, Qi Tian, Foreground object detection from videos containing complex background, MULTIMEDIA '03 Proceedings of the eleventh ACM international conference on Multimedia, Pages 2-10

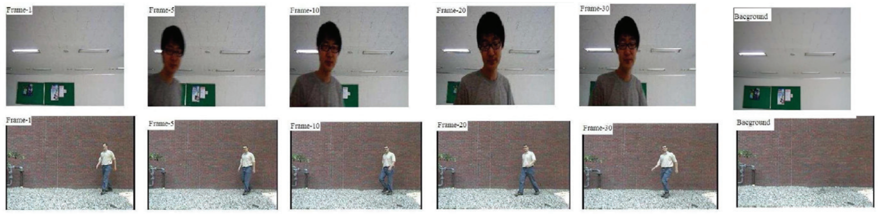

Fig. 3. Modeled Background from first k number of frames of a video scene, for each column we took 30 numbers of frames for modeling background; the figure shows frames 1,5,10,20 and,30 in the first five columns and in the last column, we present the corresponding background modeled using our proposed techniques.

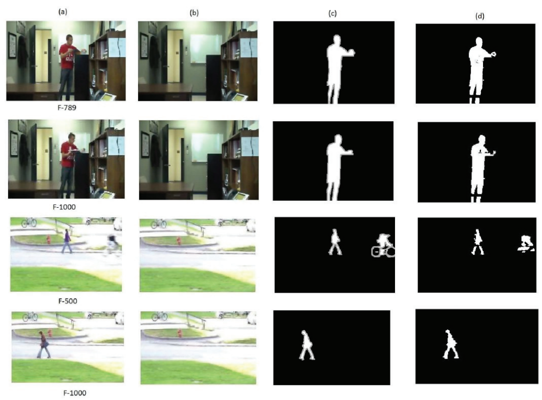

Fig. 4. Detection of foreground by eliminating modeled background, the columns in the figure represents (a)The original frame, (b) the modeled background, (c) Ground Truth and (d) our results after eliminating background

TABLE I
COMPARATIVE STUDY OF DIFFERENT PERFORMANCE METRICS BETWEEN PROPOSED APPROACH AND THE OTHER STATE OF THE ART APPROACHES ON VIDEOS OF DS1

Methods	Average RE	Average SP	Average FPR	Average FNR	Average PWC	Average F-Measure	Average Precision	Average Ranking
BBQH (Proposed)	0.8155	**0.9989**	**0.00103**	0.1846	**0.5820**	**0.8669**	0.92033	**1.333**
QCH [7]	0.7044	0.9923	0.0077	0.2956	2.2142	0.6616	0.7009	6.555
ABMM(GMM) [8]	0.5863	0.9987	0.0013	0.4137	1.9381	0.7119	**0.9532**	4.333
KDE-ISTF [9]	0.7472	0.9954	0.0046	0.2528	1.8058	0.7392	0.7998	4.444
GMM-RECTGAUSS [12]	0.6669	0.9979	0.0021	0.3331	1.5342	0.7500	0.9175	4.333
KDE-STDC [10]	0.7551	0.9940	0.0060	0.2449	1.9154	0.7554	0.7833	4.667
GMM(SG) [11]	0.8180	0.9948	0.0052	0.1820	1.5325	0.8245	0.8461	3
pROST [29]	**0.8415**	0.9937	0.0063	**0.1585**	1.15	0.8289	0.8181	3.111
RMOG [30]	0.7082	0.9981	0.0019	0.2918	1.5935	0.7848	0.9125	3.444

TABLE II
COMPARATIVE STUDY OF F-MEASURE VALUES BETWEEN PROPOSED APPROACH AND THE OTHER STATE OF THE ART APPROACHES ON VIDEOS OF DS2

Method	Bootstrap	Noisy Night	Camouflage	No Camouflage
McFarlane [19]	0.541	0.496	0.738	0.785
Stauffer [12]	0.642	0.194	0.802	0.826
Oliver [20]	–	0.213	0.802	0.824
McKenna [21]	0.301	0.098	0.624	0.656
Li [22]	0.678	0.047	0.768	0.803
Kim [23]	0.318	–	0.776	0.801
Zivkovic [24]	0.632	0.321	**0.820**	0.829
Maddalena [25]	0.495	0.263	0.793	0.811
Barnich [26]	0.685	0.271	0.741	0.799
BBQH	**0.8045**	**0.411**	0.7749	**0.8461**

[23] Stephen J. Mckenna and Sumer Jabri and Zoran Duric and Harry Wechsler and Azriel Rosenfeld, Tracking groups of people, Computer Vision and Image Understanding (2000), volume (80),pages = 42–56
[24] Zoran Zivkovic, Ferdinand van der Heijden, Efficient adaptive density estimation per image pixel for the task of background subtraction, Pattern Recognition Letters, Volume 27 Issue 7, May 2006, Pages 773-780, Elsevier Science Inc. New York, NY, USA
[25] L. Maddalena, A. Petrosino, A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications, IEEE Transactions on Image Processing, Volume 17 Issue 7, July 2008, Page 1168-1177, IEEE Press Piscataway, NJ, USA
[26] O. Barnich and M. Van Droogenbroeck,"ViBe: A powerful random technique to estimate the background in video sequences", Proc. Int. Conf. Acoust., Speech Signal Process., pp.945 -948 2009.
[27] http://www.vis.uni-stuttgart.de/index.php?id=sabs
[28] http://davidkunzman.net/photographyArticles/histograms/part2.php
[29] O. F. Seidel, C. Hage, M. Kleinsteuber "pROST - A Smoothed Lp-norm Robust Online Subspace Tracking Method for Realtime Background Subtraction in Video". In Machine Vision and Applications, 25(5):1227-1240, 2014.
[30] O. Varadarajan, S.; Miller, P.; Huiyu Zhou, "Spatial mixture of Gaussians for dynamic background modelling," Advanced Video and Signal Based Surveillance (AVSS), 2013 10th IEEE International Conference on , pp.63,68, 27-30 Aug. 2013.