JETIR.ORG
ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Statistical Background Estimation for Robust Foreground Detection

Dr. Navneet S. Ghedia[1], Dr. C. H. Vithalani[2], Dr. Ashish Kothari[3]

Principal, Sanjaybhai Rajguru College of Engineering, Rajkot, Gujarat, India[1]

Professor and Head of EC Dept., Government Engineering College, Rajkot, India [2]

Asso. Professor and Head of EC Dept., Atmiya Institute Of Technology and Science, Rajkot, India [3]

1

Abstract – Proficient categorization of the foreground and background is very critical and essential for the Visual Surveillance system. Identify the motion in the video sequences or the real time analysis, foreground classification in the basic operation hence the reliability of the surveillance system is completely depends on the segmentation approach. A large no. of algorithms and experiments are proposed and implemented for the moving object classification, recognition and detection. All approaches are usually available with the tradeoffs like computational cost, quality and accuracy of the foreground detection, environmental conditions and surveillance applications. The purpose of the paper is to present the exhaustive and comprehensive literature and study for the background estimation, background modeling and statistical foreground classification approaches. The efficacy of all the algorithm varies due to their feature properties and way of adopting the surveillance applications. In this paper we have provided the exhaustive and popular statistical pixel based parametric and non parametric approaches for the foreground detection. In all-purpose our plan is to provide the critical relative analysis of almost all the available background modeling approaches, so the researchers can effortlessly accept and modified the background model as per their requirements and applications.

***Keywords* – Background Estimation, Statistical Background Modeling, Foreground Detection, Visual Monitoring**.

I. INTRODUCTION

Visual Monitoring System requires efficient assessment of the background, background modelling and precise motion detection. Background subtraction is the easier approach to classify the background and foreground. Now a day's most of the modern smart video surveillance system requires robust detection of the foregrounds which effectively handle various environment challenges. Every visual surveillance application has different wants and requirements but moving objects for each application is remaining the same. Background subtraction, background estimation and background initialization are frequently used in video based surveillance system. Several important Visual surveillance applications are banks, department stores, airports, railways and bus stations, private and parking places, monitoring criminal activities, vehicles Patrolling and traffic management, access control, forest and fire surveillance, industrial activities, sports activities, military applications, video database management. Intelligent video surveillance system requires the exact position and activities of the moving and static foregrounds. Now a day's visual surveillance system handles diverse challenges not only pertaining to foreground or background, but it also related to the environment, cameras and system implementation. Generally, researcher try to concentrate on one of the background modeling approach, as literature proposed lots of options for the background estimation and background model. Though not a single background model technique or the algorithm capable to deal all the video sequence variations, enough estimation must determine the several scenes issues efficiently. Background model should be robust and adaptive against the scene illumination changes. Another it should not accommodate the non stationary objects such as water rippling, tree leafs weaving, snow fall, low and high intensity rain drops, weaving grass or the shadows from the stationary and moving objects. finally, the background model estimated in such a manner that, it can respond or adopt the scene changes quickly like too fast and too slow moving objects, entering or leaving objects.

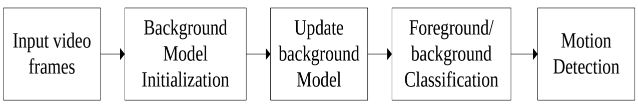

Figure 1 General Motion Detection flow

Figure 1 shows the general approach for the foreground detection. Primary step in each visual monitoring system for the foreground detection or the segmentation of the motion and background is background subtraction. Generally, it follows the background modeling, background initializations and the classifications of the background and foreground pixels along with the updating the model parameters.

Background Modeling:

Background modeling is the method to represent the initialized background. Pixel classification is the segmentation technique through it, background and foreground is confidential and over a period of time background model needs to update to handle the various challenges. Some literature classified the background modeling approaches in to Predictive and Non predictive categories, while some classified them into recursive or non recursive basis [1,5&6-2]. Previously classified those on the basis of the sliding window approach are used to approximate the backgrounds while prior used the pixels to classify the foreground and backgrounds. Frame differencing, Averaging filtering, Median filtering, Minimum and Maximum filtering, Linear predictive filter, Non parametric modeling, are some of the known Non recursive techniques. Approximated median, single Gaussian, Kalman filtering Hidden Markov Model are some of the popular recursive modeling techniques. Every model requires dealing with scenes dynamics and for first needs to initialize model. Model efficiency and accuracy depends on the background initialization, which is generates from the initial background frame. Some literature also defines it as the background generation, background extraction or the construction.

Statistical Background Model Challenges:

Background model should have robustness and adaptability to handle scene variations. In certain cases background does not available so need to train background model. Some of the useful background model challenges are as follows, [2,3 -10,11,-4,16]
Camera Location and Quality: camera can be static or PTZ and location may be fixed, aerial position or satellite images. In all cases algorithm must have to detect the motion from any video sequences. Sensing a noise or intensity variations may leads to erroneous detection and estimation. Camera quality is being varied from low to high definition quality. Low resolution will generate less no. of frames at the same time high resolution have large file size
Scene Environments: for the in indoor and outdoor environments, shadows, illuminations and occlusions are the crucial assignments sometimes dynamicity in terms of non stationary background also provides higher challenging to the designer.
Moving Foregrounds: Moving Objects have similar color as background, too much fast and slow objects, abrupt motion, object silhouette and behavior.
Tracking challenges: non-rigid object tracking, small-size object tracking, tracking a varying number of objects, complicated pose estimation, object tracking across cameras with non-overlapping views
Real Time processing: to handle the scene dynamics using low computational cost and computational time.
Literature proposed also proposed some of the other challenges such as, moving object appearance, abrupt motion, occlusions, complex backgrounds and non-rigid object deformation. The entire algorithms must focus on some of the challenges and try to provide o the optimum solution for the moving object detection efficiently.

Statistical background Model features:

Feature selection is used to distinguish the foreground and background efficiently. Every feature has its own properties to handle scene dynamics such as, illumination variations, motion variations and background variations. Sometimes algorithm may be on the basis of one feature or may be on the basis of fusion features to handle system reliability. System consistency depends on the feature selection, but somehow 100% realiability is not achieved.
Features are popularly computed in the pixel domain as the value of the pixel is directly available. The following features are commonly used [41,47]:

Intensity feature:
The scene provided by the gray level or infra red cameras should be computed using the very basic intensity features.

Color feature:
Usually, color features in the RGB color model is widely used as most of the information received from the camera are color in nature with three components either RGB or HSV. The major threat on the utilization of the color features is, it is very sensitive to illumination variations and hence the moving object shadows cannot efficiently handle. In addition to other edge or texture feature along with the color can resolve the problem.

Edge feature:
Appearance of the moving objects is varying with the changing light intensity. So, intensity and color feature cannot handle illumination changes. Hence edge feature, especially first order derivatives are used as the edge feature. Edge is removing the ghosting effect so, it is a kind of high pass filter.

Texture feature: To handle illumination variations, texture feature is the best alternatives then others. Typically, Local Binary Pattern (LBP) and Local Ternary Pattern (LTP) are the two common used texture features.

Stereo feature:
Image or frame acquisition from the stereo, 3D camera, RGB-D camera requires stereo feature for the extraction of the disparity of the scene, generally, try to extract depth from the video.

Motion feature:
Motion features is used to detect the moving object along with camera motion. It is also referred to as the optical flow approach. It is used to handle the clutter and irrelevant backgrounds.

Local Histogram feature:
It is generally computed along with the color features and used to deduce the edge features.

Local correlogram feature:
It provides spatial information of the pixel along with the color information. It also used for the tracking purposed as, it calculates the inter pixel connectivity's and relations.

Haar-like features:
It is used for the real time video surveillance system for the face detection based on the Haar wavelets.

Location features:
It is used to develop the enslavement among the pixel.
Usually, we assume that every surveillance application is considered to be in controlled environments with a clear distinguish between backgrounds and foreground. However, in real time processing it would not be possible to predict the scene dynamics for the surveillance. Every feature has its own merits and demerits, so to use multiple features for the background modeling is advisable. Some of the important and popular operators like basic, statistical or fuzzy are required for the feature aggregation to fuse features efficiently.

Statistical background Modeling:

Crucial and challenging issue in the visual surveillance system is to define the moving object under challenging environments. Set of connected pixels is used to represent the motion or moving objects is a better initiative [4,46-12]. Background modeling is required to detect the motion, a static or moving foreground objects from the background. Professional and fair approach to detect the foreground is the background modeling. The foreground is generated using the masking or thresholding techniques. Local, global, fixed – static, statistical or fuzzy threshold used to classify the pixels in the background model and the current frame.

Stauffer et.al. [5,ijcir-1] proposed a traditional mixture of Gaussian approach. They have used scene multimodality GMM for the dynamic scenes their approach fails to detect object in sudden illumination and also it will take more computational cost. Bowden et.al. [6,ijcir -6] developed the system which conquer the drawbacks of [5] learning rate. Harville et.al. [7,ijcir -11] Proposed a colour and depth based foreground background classification approach. The proposed algorithm is also cannot handle the learning rate issued efficiently, as it developed using the traditional MoG approach. Harville et.al [8,ijcir -12] proposed extended GMM approach for the detection and tracking purpose. Feedback in the algorithm resolved the learning rate issue; hence it is appropriate to handle the sudden illumination changes. Kim et.at. [9,ijcir -5] proposed vector quantization based code book approach with the help of traditional GMM approach for the motion detection. Complex and abrupt object motion able detects. LI et.al. [10,ijcir -9] explained the importance of the learning parameters and proposed Gaussian Based Mixture model to estimate the scene background. Proposed algorithm efficiently handle scene dynamics specially illumination changes. Strategies of learning parameters updation along with the Adaptive Mixture Model, algorithm able to improve the convergence speed. Katharina et.al. [11,ijcir -10] developed spatio-temporal adaptive GMM by using traditional GMM and spatial and temporal dependency. Spatio temporal feature is used to remove the shadow or the false positive or the connected region with the foregrounds, hence improved the foreground detection efficiency. Jepson et al[12,16-25] proposed W,S and L based WSL mixture model to estimate the background. Inter frame variations and stable structure deals with the dataset outliers and occlusions. Zhou et al.[13,16-26] proposed another WSL mixture model. Instead of filtering approach, algorithm uses visual features for estimating the background modeling. McKenna et al.[14,16-27] suggested Gaussian approximate distribution model for the appearance model. Gaussian model used the mixture parameters and it follows the multimodality. Han at al. [15,16-28] explained the importance of the Gaussian Mixture model. Proposed algorithm determines probability distributions automatically for the mixture parameters such as weight, mean and covariance.

Extrinsic and Intrinsic Improved Statistical Modeling:

Gaussian Mixture Model is the fundamental statistical background modeling, which able to handle most of the dataset challenges. However certain modification such as intrinsic and extrinsic makes it more robust and adaptive to handle the scene variations. Intrinsic improvements are related with Mixture model estimation and maintenance of the model parameters for the detection along with the features. Extrinsic improvements improves overall foreground detected objects by means of improving spatial and temporal object information and hence boost the foreground detection efficiency and the model robustness [16,30].

Greiffenhagen et al. [17,56] explained the importance of the mixture weight for the foreground detection. They have proposed approximation of the MAP for the matching test.
Shimada et al. [18,96] proposed the control mechanism of the no. of Gaussians. According to the scene dynamics it automatically selects the no. of Gaussians. Haque et al. [19,20,21,104-106] proposed the alternative solution for the

Threshold T. they have proposed newer S parameter for the better classification. Newer sensitive parameter is independent of the learning rate. Pnevmatikakis et al. [22,111] developed unique tracking feedback for the threshold using PPM in spatio-temporal manner. The same has been improved by Pnevmatikakis et al. [23,112] using the target covariance matrix. Fang et al. [24,145] adopts a block wise Gaussian model which consists on a vector of 3×3 neighbors of the current pixel to increase the model robustness and resolve the limitations of the MoG model.Latecki et al. [25,146] incorporated decomposes the video into spatiotemporal block so, it decreases the processing time hence increases the robustness against the noise and moving backgrounds. Bhaskar et al. [26,147] divide first the image in clusters which are generated using a color clustering mechanism of the nearest neighbor approach. In [27,67], Jain et al. incorporates the edge, intensity and Gradient feature for the sub pixel image. Fusion features provides robustness against illumination changes. Hampapur et al. [28,81] utilized texture features to deal with the sudden illumination variations. Silvestre [29,91] adopts the two camera challenge using the Time -of-Flight (TOF)

Camera. Gao et al. [30,41] propose a Spatial-Kinetic Mixture of Gaussians model (SKMGM). A fusion feature vector is taken to used to describe each pixel. Chen et al. [31,61] utilized video features for background estimation. Proposed approach ably detects the objects against the noise and illumination variations. Zhou and Zhang [32,152] suggested level set method to handle the merging splitting of the frame. It is computationally costlier but provides excellent results. Schindler et al. [33,153] present a MRF formulation for the spatial contiguity. It is very much applicable for the real time surveillance application. Sun and Yuan [34,76] proposed modified MoG model for the detection of the scene dynamics. It is used to detect the sharp changes or the sudden intensity variations of the scene. Hung et al. [35,102] propose pixel-based and block-based joint approaches for the single framework. Cristani et al [36,43] proposed an S-TAPPMOG model based on spatial sampling techniques. Parameter maintenance is done at the region level. Yang et al. [37,39] use a two-layer

Gaussian mixture model (TLGMM) approaches to deal with the dynamic background for the foreground detection. First layer deal with the small changing in the pixel while second deal on the significant variations. Ellis et al. [38,121] propose an intensity and color feature based algorithm, to deal with the sensitivity of the foreground and background. Taycher et al. [39,113] proposed algorithm based on the background formulations and maintenance. Proposed approach statistically reliable approach to incorporate the response of the high level motion. Turdu and Erdogan [40,159] proposed unique hysteresis thresholding approach for better foreground classification.

V.CONCLUSION

The purpose of this exhaustive survey is to provide the enhanced literature as well as to provide the comprehensive statistical background modeling approaches for the foreground or motion detection in various Visual Monitoring surveillance systems. This survey very well describes and incorporates the general motion detection flow and various statistical background modeling approaches in broad categories. The simple outcomes of this survey are such as, categorized the various approaches based on the simple modeling and describe the possible improvements in the model, extensively provide the study for the statistical background model challenges and the various modeling features. We have studied and provided the various possible intrinsic improvements for the better foreground classification and also provided the various possible extrinsic improvements for the improvements of overall foreground detected object appearance by means of improving spatial and temporal object information and appreciation efficiency.

VI. REFERENCES

1. Chris Stauffer and W. Eric L. Grimson, "Learning Patterns of Activity Using Real-Time Tracking,'' IEEE Transactions On cePattern Analysis and Machine Intellege, Vol. 22, No. 8, pp 747-757 August 2000.
2 C. Kim and J. Hwang (2002) Fast and automatic video object segmentation and tracking for content‐based applications. IEEE Transactions on Circuits and Systems for Video Technology. vol. 12‐2. pp. 112‐129.
3. M. J. Hossain, M. A. A. Dewan and O. Chae (2007) Moving object detection for real time video surveillance: An edge based approach. IEICE Transactions on Communications. vol. 90‐B. no. 12. pp. 3654–3664.
4. A. Yilmaz, O. Javed and M. Shah (2006) Object tracking: A survey. ACM Computing Surveys. vol. 38‐4. pp. 13.
5. D.‐Y. Lee, J.‐K. Ahn and C.‐S. Kim (2009) Fast background subtraction algorithm using two‐level sampling and silhouette detection. 16th IEEE International Conference on Image Processing. pp. 3177‐3180. Nov. 2009.
6. A. Mittal and N. Paragios, "Motion-based background subtraction using adaptive kernel density estimation," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Washington, DC, volume II, 2004, pp. 302–309.
7. A. Elgammal, D. Harwood, and L. Davis, "Nonparametric model for background subtraction," in Proc. European Conf. on Computer Vision, Dublin, Ireland, volume II, 2000, pp. 751–767.
8. C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Realtime tracking of the human body," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, pp. 780–785, 1997.
9. T. Kim and K.-H. Jo, "Generation of multiple background model by estimated camera motion using edge segments," in ICIC '08: Proceedings of the 4th international conference on Intelligent Computing. Berlin, Heidelberg: Springer-Verlag, 2008, pp. 536–543.
10. A. Elgammal, R. Duraiswami, D. Harwood, L. S. Davis, R. Duraiswami, and D. Harwood, "Background and foreground modeling using nonparametric kernel density for visual surveillance," in Proceedings of the IEEE, 2002, pp. 1151–1163.

11. N. I. Rao, H. Di, and G. Xu, “Panoramic background model under free moving camera,” in FSKD ’07: Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery. Washington, DC, USA: IEEE Computer Society, 2007, pp. 639–643.
12 C.R. Wren, A. Azarbayejani, T. Darrell, and A.P. Pentland, “Pfinder: Real-Time Tracking of the Human Body,” IEEE Trans. PAMI, vol. 19, no. 7, pp. 780-785, 1997.
13 Douglas Reynolds, “Gaussian Mixture Models”, MIT Lincoln Laboratory, 244 Wood St., Lexington, MA 02140, USA.
14.A. Yilmaz, O. Javed and M. Shah, “Object Tracking: A Survey”, ACM Journal of Computing Surveys, Vol. 38(4), , pp. 1-45,2006.
15.Greg Welch and Gary Bishop. An introduction to the kalman filter, 1995 & 2006.
16.PETS 2009 Dataset 16. S. M. Khan and M. Shah. Tracking multiple occluding people by localizing on multiple scene planes. PAMI, 31(3):505–519, 2009.
17.R.Vezzani,R.Cucchiara, "Video Surveillance Online Repository (ViSOR): an integrated framework" in Multimedia Tools and Applications, vol. 50, n. 2, Kluwer Academic Press, pp. 359-380, 2010
18. http://www.changedetection.net.