A comprehensive Survey: Background Estimation and Motion Detection Approaches

Dr. Navneet S. Ghedia[1], Dr. C.H. Vithalani[2] and Dr. Ashish M Kothari[3]
[1]Principal and Professor, Sanjaybhai Rajguru College of Engineering, Rajkot, India
[2]Professor and Head of EC Dept., Government Engineering College, Rajkot, India
[3]Asso. Professor and Head of EC Dept., Atmiya University, Rajkot, India

Abstract: *Visual Monitoring System requires efficient estimation of the background, background modelling and accurate motion detection. Background subtraction is the simplest 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. Literature shows not a single algorithm ably handle all constraints, on the basis of the applications researchers generally, focuses on the classification or segmentation techniques. The objective of this paper is to provide the rigorous and comprehensive study of the various background modeling approaches and foreground detection techniques. This survey discussed some of the popular statistical pixel based parametric and non parametric approaches, region based approached and some of the popular hybrid approaches for the background handles. The aim of the study is to handle the various datasets and real time video sequences challenges. in spite of of all the literature survey still some of the challenges like object abrupt motion, object complex silhouette, object appearance and camera motion are challenging issues for the researchers.*

Keywords: **Background Estimation, Background Modeling, Foreground Detection, Visual Monitoring.**

1. Introduction

Background subtraction, background estimation and background initialization are popularly used in various visual surveillance applications. Usually, all the video sequences are taken from the static cameras. Smart and intelligent video surveillance system requires the exact location and behaviour of the moving and static foregrounds. Now a day's visual surveillance system handles different challenges not only pertaining to foreground or background, but it also related to the environment, cameras and system implementation. Earlier before 1996, foreground and background classification was achieved on with the help of couple frame because of computer system limitations. Looking towards the present scenario the latest intelligent surveillance approaches not handles the various challenges but also deals with the different foregrounds such as moving and static foregrounds. Many of the surveillance system defines background subtraction is the primitive operation. However so many various background foreground classifications and object tracking options are available. All approaches are generally available with the tradeoffs like computational cost and the quality and accuracy of the foreground detection. Sometimes the segmentation of the dynamic region or moving object is known as background subtraction or the foreground segmentation. Real time visual monitoring system does not adopt the illumination variations, similarity in foreground and background objects, shadows and non stationary backgrounds. Figure 1 shows the general approach for the foreground detection.

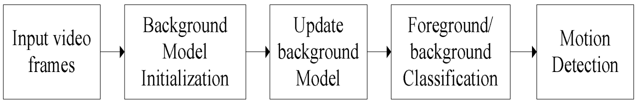

Figure 1. Fundamental Motion Detection flow

Fundamental step in every visual monitoring system for the motion detection or the segmentation of the foreground and background is background subtraction. Background subtraction generally follows the background modeling followed by background initializations and the classifications of the background and foreground pixels along with the background modeling maintenance. Background initialization is used to generate the initial background frame on which the entire background model accuracy depends. Some literature also defines it as the background generation, background extraction or the construction.

2. Background modeling

Background modeling is the technique to represent the initialized background. Pixel classification is the segmentation technique through it, background and foreground is classified and over a period of time background model needs to update to handle the various challenges. Most researchers have focuses their intention on the automated surveillance. For the automated surveillance one must aware about the behaviors and the activities of the moving and static foregrounds in a space. For the robustness and effectives of the algorithm one must have to analyze and understand the dataset, video sequences or the real time detection challenges carefully.
The prime importance in terms of implementation of the surveillance algorithm is to separate the foregrounds known as the moving objects from the backgrounds known as a static objects or the information. Background models can be classified as, Basic Background modeling, Statistical background model, background estimation using filtering or fuzzy background modeling. Figure 2 shows the different background model classifications. [1] proposed and compares the background algorithms. They have classified all the approaches in the field of pixel based and region based approaches and some are hybrid approaches. [2] proposed all the likely approaches in Row pixel, optical flow, Histogram representation, covariance representation, wavelet filtering representation, active contour representation, feature based presentation like SIFT [3] and SURF [4], segmentation based. [5] Proposed all the background modeling approach in to the broad categories of Recursive and Non Recursive modeling. [6] proposed median filtering based background modeling approach. With the exact selection of the frame rate and buffer size, algorithm is easy to adopt the dynamic environments. [7] proposed approximate median filtering approach to estimate the background. [8] Compares three background model approaches, Median filtering, Gaussian Mixture Model and Kernel Density Estimation. [9] Presented a novel feature based shape based background modeling.

Table 1 represents some of the important background modeling approaches. Usually, all the approaches are classified either pixel based or region based.

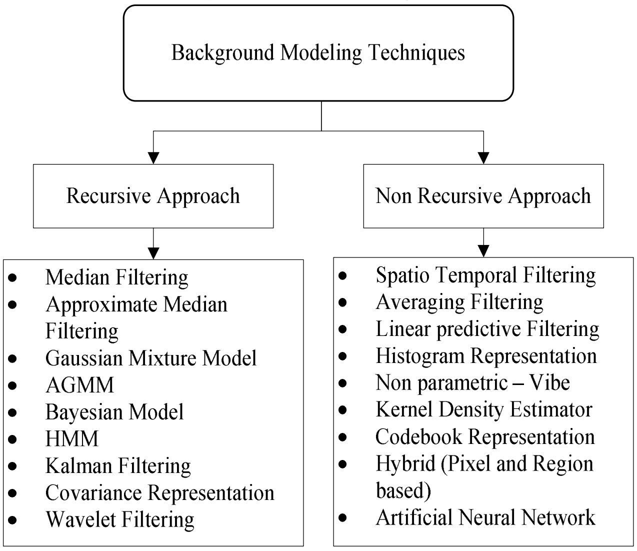

Table 1. Basic Background Modeling Techniques

3. Background model challenges

Background model must have robustness and adaptability to handle scene variations such as illuminations variations and non stationary backgrounds. In certain cases background does not available so need to train background model. Many constraints and challenges made Visual Surveillance system challenging is described as follows, [10,11]

Camera Location: 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: For every CCTV, 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. Poor quality videos due to tripod variations, weak contrast, noisy artifacts leads to increases false positives.

Environments: Generally, Motion can be detected in indoor and outdoor environments where shadows, illuminations and occlusions are the biggest constraints. However for the outdoor scenes non stationary background must be accounted.

Foregrounds: Moving Objects have similar color as background, too much fast and slow objects, abrupt motion, object silhouette and behavior.

Tracking Limitations: 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: It requires fast and low computational cost algorithm. It does not handle gradual or sudden illumination variations efficiently.

Object Appearance: illumination variations, fast camera motions, occlusion, dataset noise and outliers, non-rigid silhouette deformation, object abrupt motion.

To ensure high detection efficiency usually, all the researches have suggested the huge range of background modeling to meet the above challenges. All the various models are focused on some of the challenges and on the bases of the certain predefined applications.

4. Pixel based approach

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. Since several decades many researchers have studied and proposed algorithms for detecting the motion in the video sequences. Generally, they have proposed frame difference either first frame or previous frame approach or the background subtraction approach. With the help of detailed literature survey usually, all the background modeling approaches are classified in to the Pixel based, Region based or Hybrid approach. It can also be categorized as parametric or non-parametric based model. Initially, [12] proposed a simplest pixel based single Gaussian Model. It is a kind of parametric approach which deals with the simple scenes. Proposed single Gaussian approach does not handle scene variations. [13, 14] introduced novel pixel based Gaussian Mixture Model. They have proposed every pixel in the form of mixture components such as weight, mean and covariance. [15,16] suggested the improved version, Adaptive Gaussian Mixture Model.[17] further improved has been suggested by him. Without affecting the GMM stability, he has improved the Adaptive GMM model by updating the learning rate. [18] suggested a newer dynamic Gaussian component approach, for the improvement in the detection accuracy and computational time. [19] proposed Bayesian approach to model the background with the help of prior knowledge. [20] proposed pixel based Gaussian Components for the background modeling approach, further the segmentation is obtain with the help of threshold decision method. Above all the proposed approaches are parametric model. [21] proposed nonparametric self organization based neural network (SOBS) approach for the foreground detection. [22] proposed a vector quantization based codebook nonparametric method for the background modeling. [23] suggested a statistical computing
sample consensus (SACON) approach for the background model. This non parametric model uses the color and motion features for the detection of the motion in the video sequences. [24] proposed a unique pixel based non parametric approached known as pixel-based adaptive segmenter (PBAS). They have used pixel history to model the background pixels. [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. [26] proposed another WSL mixture model. Instead of filtering approach, algorithm uses visual features for estimating the background modeling. [27] suggested Gaussian approximate distribution model for the appearance model. Gaussian model used the mixture parameters and it follows the multimodality. [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. [29] presented a spatial color mixture Gaussian Model for the background modeling. The uniqueness of the algorithm extracts both the spatial and color information efficiently. [30] presented a prediction based non parametric pixel based Mixture of Gaussians algorithm for the background modeling which is known as PM-MOG used for classification of the foreground and background pixels. [31] presented a Bayes regulation which is in association with the background and foreground. The selected feature vector is needs to generate for the better classification of the foreground and background in the video sequences. [32] Proposed a Auto regressive Moving Average Model for pixel classification. It is textured based segmentation appearance model where Kalman is used to handle dynamicity of the scene. The unique texture based recursive filtering approach efficiently classifies the foreground and background pixels. [33] presented a motion assisted matrix restoration-based appearance background model. Low rank matrix and

sparse matrix ultimately classify the foregrounds and backgrounds. Sparse matrix is belonging to foregrounds while low rank matrix is associated with the background.

5. Region based approach

Likewise pixel based approach, region based works on the basis of the inter pixel relationship. It segments the complete image in to region and identified or extracts the foreground with the help of image region. In comparison with the other parametric or non parametric pixel based approaches region based approaches always reduces the effect of the noise. [34] proposed novel region based non parametric background modeling known as Kernel Density Estimation (KDE). [35] developed an ultimate solution for the background modeling in terms of the kernel density estimation on the bases of the color histogram. The proposed algorithm utilize the Bhattacharya distance for the calculating the pixel regions. [36] developed scale aware kernel density estimator for the background modeling. Gaussian mean shift features is used to classify the regions. [37] presented a heuristic block matching region based non parametric approach for the background modeling. They utilized the scene illumination variations to discriminate the motion from the static background. [38] extended the similar approach with the uses of fixed size regions. It is very much effective for the outdoor environments specially, in presence of the dynamic environment. [39] used image region color space for the detection for the foreground. It si also an efficient region based approach compare to color, texture or descriptor based method. [40] used the texture feature-based modeling known as local binary pattern (LBP) for the background modeling. The LBp utilized the information among the overlapping region of the foregrounds and backgrounds by means of histograms for the better classification. [41] developed binary descriptor-based background modeling to handle the outdoor environments and able to take various constraints. [42] utilized the samples of the binary descriptors to model the background.

6. Hybrid approach

The integration of the pixel and region based approaches generates the Hybrid background modeling. Hybrid approach handles the illumination variations and dynamic backgrounds for the outdoor surveillance. It also better classifies the background and represents the regions individually. [43] utilized the frame level, pixel level and region level information of the scenes and using Wiener filtering estimates the background. It can ably detect the foregrounds in presence of the sudden illumination variations and non stationary backgrounds. [44] developed integrated pixel and region-based hybrid approach for the motion estimation. It used the RGB color information and efficiently classify the foregrounds. Its computational complexity is very high. [45] proposed unique embedded hybrid approach for the foreground segmentation. [46] developed hybrid approach with the help of log like hood function and pixel classification for the generating the background model. [47] presented a Kernel based hybrid model using online SVM learning.

7. Conclusion

In this review, we have presented a survey on the foreground detection and motion detection applications. This survey also describes the general motion detection flow and various background modeling approaches in details. We have gone through and discussed the various background modeling challenges such as camera, environment, foregrounds, racking and real time processing. In this paper we have provide the important information regarding the background modeling approaches such as pixel based approaches, region based approaches and hybrid approaches for the better estimation of the background and foreground detection.

8. References

[1] Yong Xu , Jixiang Dong, Bob Zhang, Daoyun Xu, "Background modeling methods in video analysis: A review and comparative evaluation,'' CAAI Transactions on Intelligence Technology 1 **(2016)** 43e60, pp 747-757.

[2] Xi Li, Weiming Hu, Chunhua Shen, Zhongfei Zhang, Anthony Dick, Anton van den Hengel, "A Survey of Appearance Models in Visual Object Tracking", ACM Transactions on Intelligent Systems and Technology, **(2013)**

[3] Zhou, S. K., Chellappa, R., And Moghaddam, B. **(2004)**. "Visual tracking and recognition using appearance adaptive models in particle filters." IEEE Trans. on Image Processing 13, 1491–1506.

[4] Bay, H., Tuytelaars, T., And Gool, L. V. **(2006)**. "Surf: Speeded up robust features." European Conference on Computer Vision, 404–417.

[5] Saameh Golzadeh Ebrahimi, "Shadow Aware Object Detection and Vehicle Identification via License Plate Recognition" Thesis, Eastern Mediterranean University September **(2009)**.

[6] Cucchiara, Grana, Piccardi, Prati, "Detecting moving objects, ghosts and shadows in video streams", IEEE Transactions on Pattern Analysis and Machine Intelligence, (vol. 25, no. 10), **(2003)**, pp. 1337-1342.

[7] McFarlane N, Scho_eld C. "Segmentation and Tracking of Piglets in Images", Machine Vision and App., (Vol. 8, No. 3). **(1995)**, pp. 187 -193.

[8] M. Hedayati, Wan Mimi Diyana Wan Zaki_, Aini Hussain, "A Qualitative and Quantitative Comparison of Real-time Background Subtraction Algorithms for Video Surveillance Applications," Journal of Computational Information Systems 8: 2 **(2012)**, pp. 493-505.

[9] Wei-Lwun Lu, James J. Little, "*Simultaneous Tracking and Action Recognition using the PCA-HOG Descriptor"* Proceedings of the 3rd Canadian Conference on Computer and Robot Vision (CRV'06)

[10] Thierry BouwmanS, Belmar Garcia-Garcia, "Background Subtraction in Real Applications: Challenges", Computer Science Review **(2019)**.

[11] Xi Li, Weiming Hu, Chunhua Shen, Zhongfei Zhang, Anthony Dick, Anton van den Hengel, "A Survey of Appearance Models in Visual Object Tracking," ACM Transactions on Intelligent Systems and Technology, **(2013)**.

[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] C. Stauffer, E. Grimson, "Adaptive background mixture models for realtime tracking, in: Proc." IEEE Int. Conf. Comput. Vis. Pattern Recognit., vol. 2, **(1999)**, pp. 246e252.

[14] C. Stauffer, E. Grimson, IEEE Trans. Pattern Anal. Mach. Intell. 22 (8) **(2000)** 747e757.

[15] Z. Zivkovic, "Improved adaptive Gaussian mixture model for background subtraction" in: Proc. 17th Int. Conf. Pattern Recognit., vol. 2, **(2004)**, pp. 28e31.

[16] Z. Zivkovic, F. van der Heijden, Pattern Recognit. Lett. 27 **(2006)** 773e780.

[17] D.S. Lee, Pattern Anal Mach. Intell. IEEE Trans. 27 (5) **(2005)** 827e832.

[18] A. Shimada, D. Arita, R.I. Taniguchi, "Dynamic control of adaptive mixture-of-Gaussians background model" in: AVSS '06 Proceedings of the IEEE International Conference on Video and Signal Based Surveillance, IEEE Computer Society Washington, DC, USA, **(2006)**.

[19] N.M. Oliver, B. Rosario, A.P. Pentland, IEEE Trans. Pattern Anal. Mach. Zntell. 22 (8) **(2000)** 831e843.

[20] S.-Y. Chien, W.-K. Chan, Y.-H. Tseng, H.-Y. Chen, IEEE Trans. Circuits Syst. Video Technol. 23 (6) **(2013)** 921e934.

[21] L. Maddalena, A. Petrosino, IEEE Trans. Image Process. 17 (7) (2008) 1168e1177.

[22] K. Kim, T. Chalidabhongse, D. Harwood, L. Davis, Real-Time Imaging 11 (3) (2005) 172e185.

[23] H. Wang, D. Suter, Pattern Recognit. (2007) 1091e1105.

[24] M. Hofmann, P. Tiefenbacher, G. Rigoll, "Background segmentation with feedback: the pixel-based adaptive segmenter" in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, (2012), pp. 38e43.

[25] Jepson, A. D., Fleet, D. J., And El-Maraghi, T. F. **(2003)**. "Robust online appearance models for visual tracking." IEEE Trans. on Pattern Analysis and Machine Intelligence 25, 10, 1296–1311.

[26] Zhou, H., Yuan, Y., And Shi, C. **(2009)**. "Object tracking using sift features and mean shift." Computer Vision and Image Understanding, 345–352.

[27] Mckenna, S., Raja, Y., And Gong, S. **(1999)**. “Tracking colour objects using adaptive mixture models.” Image and Vision Computing 17, 223–229.
[28] Han, B. And Davis, L. **(2005)**. “On-line density-based appearance modeling for object tracking.” IEEE International Conference on Computer Vision, 1492–1499.
[29] Wang, Q., Chen, F., Xu, W., And Yang, M. **(2012)**. “Object tracking via partial least squares analysis.” IEEE Transactions on Image Processing 21, 10, 4454–4465.
[30] Y. Guo,W. Zhu, P. Jiao, and J. Chen. “Foreground detection of group-housed pigs based on the combination of mixture of gaussians using prediction mechanism and threshold segmentation.” *Biosystems Engineering*, 125:98–104, **(2014)**.
[31] L.Li, W.Huang, I.Gu, and Q.Tian, “Foreground Object Detection from VideosContaining Complex Background” , ACM **(2003)**.
[32] Jing Zhong, Stan Sclarofi , “Segmenting Foreground Objects from a Dynamic Textured Background via a Robust Kalman Filter”, in Proc. IEEE, **(2003).**
[33] Xinchen Ye, Jingyu Yang,Xin Sun, Kun Li, “Foreground Background Separation From Video Clips via Motion-Assisted Matrix Restoration”, in Proc. IEEE , **(2015)**, pp. 25772580.
[34] A. Elgammal, D. Hanvood, L.S. Davis, “Non-parametric model for background subtraction”, in: Proc. ECCV 2000, **(2000)**, pp. 751e767.
[35] Comaniciu, D., Ramesh, V., And Meer, P. **(2003)**. “Kernel-based object tracking.” IEEE Trans. on Pattern Analysis and Machine Intelligence 25, 5, 564–577.
[36] Collins, R. T., Lipton, A. J., Kanade, T., Fujiyoshi, H., Duggins, D., Tsin, Y., Tolliver, D., Enomoto, N., Hasegawa, O., Burt, P., And Wixson, L. **(2000)**. “A system for video surveillance and monitoring.” Tech. rep., VSAM final report, Carnegie Mellon University. Technical Report CMU-RI-TR-00-12.
[37] M. Seki, T. Wada, H. Fujiwara, K. Sumi, “Background subtraction based on cooccurrence of image variations”, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit, **(2003)**, pp. II-65eII-72.
[38] D. Russell, S. Gong, “A highly efficient block-based dynamic background model”, in: Proc. IEEE Conf. Adv. Video Signal Based Surveill, **(2005)**, pp. 417e422.
[39] H.-L. Eng, J. Wang, A.H.K.S. Wah, W.-Y. Yau, IEEE Trans. Image Process. 15 (6) **(2006),** 1583e1600.
[40] M. Heikkil€a, M. Pietik€ainen, IEEE Trans. Pattern Anal. Mach. Intell. 28 (4) **(2006)** 657e662.
[41] W.-C. Liu, S.-Z. Lin, M.-H. Yang, C.-R. Huang, “Real-time binary descriptor based background modeling”, in: Proc. IAPR Asian Conf. Pattern Recognit, **(2013)**, pp. 722e726.
[42] G.-H. Huang, C.-R. Huang, “Binary invariant cross color descriptor using galaxy sampling”, in: Proc. 21st Int. Conf. Pattern Recognit, **(2012)**, pp. 2610e2613.
[43] Y. Nonaka, A. Shimada, H. Nagahara, R. Taniguchi, “Evaluation report of integrated background modeling based on spatio-temporal features”, in: Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops, **(2012)**, pp. 9e14.
[44] S.-S. Huang, L.-C. Fu, P.-Y. Hsiao, IEEE Trans. Circuits Syst. Video Technol. 19 (4) **(2009)** 522e532.
[45] T.H. Tsai, C.-Y. Lin, S.-Y. Li, IEEE Trans. Circuits Syst. Video Technol. 23 (1) **(2013)** 15e29.
[46] Kelm, B. M., Pal, C., And Mccallum, A. **(2006)**. “Combining generative and discriminative methods for pixel classification with multi-conditional learning.” International Conference on Pattern Recognition, 828–832.
[47] Shen, C., Kim, J., And Wang, H. **(2010)**. “Generalized kernel-based visual tracking”. IEEE Trans. Circuits and Systems for Video Technology, 119–130.