A Static Object Detection in Image Sequences by Self Organizing Background Subtraction

Pravinkumar Yadav[1], Prof. Aditi Jahagirdar[2]

[1] PG Scholar, Department of Information Technology, MIT College of Engineering, Pune, Maharashtra, India
[2] Assistant professor, Department of Information Technology, MIT College of Engineering , Pune, Maharashtra, India

---***---

Abstract - ***In several video surveillance applications, such as the detection of abandoned/stolen objects or parked vehicles, the detection of stationary foreground objects is a critical task. In this paper the model based framework is suggested for detecting static objects. Firstly, a background subtraction based method that relies on modeling not only the background, but also the stopped foreground is implemented. Secondly, self-organizing model for image sequences which automatically adapts to scene changes is performed. Finally, we evaluate the proposed algorithm and compare results with the background segmentation algorithm using video surveillance sequences from visor datasets. Experimental results show that the proposed approach has better detection accuracy of stationary foreground regions as compared to the segmentation approach.***

Key Words: *SFS - Stopped foreground subtraction algorithm, SOBS - Self organizing background subtraction, MOG - Mixture of Gaussians*

1. INTRODUCTION

Detecting stationary foreground regions in video has recently become an active area of research in many video surveillance areas such as the detection of abandoned objects and illegally parked vehicles [15]. Video surveillance systems aim to provide automatic analysis tools that may help the supervisor personnel in order to focus his/her attention when a dangerous or strange event takes place.

There are many algorithms have been proposed that deals with the detection of stationary foreground objects, the majority of them based on background subtraction techniques. Background subtraction techniques are the most popular choice to detect stationary foreground objects [1][2][4], because they work reasonably well when the camera is stationary and the change in ambient lighting is gradual, and they represent the most popular choice to separate foreground objects from the current frame.

Many approaches have been proposed for stationary region detection in video. They can be classified based on tracking [11] or background subtraction. As tracking accuracy is significantly degraded in complex sequences, such as crowded videos, this section focuses on the second category that does not use tracking and can be applied to a wide variety of video-surveillance scenarios.

As suggested in [13], Adaptive background subtraction (ABS) has been proposed to handle photometric errors by continuously updating the background model. Combinations of fast and slow adaptation rates can be used for stationary detection [6]. However, such adaptation might decrease detection performance as static objects can be incorporated into the background before they become static [12]. Thus, slow rates are preferred that reduce the robustness to photometric errors. Moreover, background initialization is complex in crowded sequences that, if incorrect, may lead to many false positives (of foreground), which decrease stationary detection performance.

This paper combines SFS algorithm [2] along with neural network. The basic idea consists of maintaining an up-to-date model of the stopped foreground and discriminating moving objects as those that deviate from this model [1]. Neural network-based solutions are already been considered due to the fact that these methods are usually more effective and efficient than the traditional ones [3][5].

A 3-D neural model for image sequences that automatically adapts to scene changes in a self-organizing manner was targeted for modeling the background and the foreground, finalized at the detection of stopped objects. Coupled with the proposed model-based framework for stopped object detection, it enables the segmentation of stopped foreground objects against moving foreground objects [2].

The remainder of this paper is structured as follows: section 2 describes the different background subtraction based approaches, proposed system is presented in section 3, section 4 shows experimental results and section 5 closes the paper with some conclusions.

2. RELATED WORK

Two major error sources affect the performance of detection approaches based on background subtraction. The first corresponds to photometric factors (illumination changes, camouflages, shadows and reflections) whereas the second derives from sequences with high density of moving objects (multiple occlusions and algorithm initialization) [15].

The mixture of Gaussians (MOG) based BS technique proposed by Stauffer and Grimson (S&G) [10] is used most widely in real-time video analytics due to its simplicity, flexibility, and dependence on few user-tunable parameters. In this approach, observations are independently modeled at each pixel using a MOG, where each Gaussian model represents the color/intensity distribution of an environment component (e.g., sky, road, car, people etc.), and the most frequently observed models are selected as the representative of the scene background.

Thomas Sikora et. al. [8], proposes a method for the detection of static objects in crowded environments. it use two complementary background models: one model is devoted to accurately detect motion, while the other aims to achieve a representation of the empty scene. By defining some simple actions at the pixel level and at the region level, the description of the empty scene is achieved along the whole sequence and the background model can be rapidly healed upon removal of temporarily static objects, therefore improving segmentation results, especially in crowded scenarios.

The partial and total occlusion problem identified in base algorithm [4] which leads to false detection of objects (higher false positive rate) in crowded sequence is reduced by improving background subtraction and occlusion handling mechanism. A change of the background subtraction technique is an integration of a sub-sampled frame difference stage and occlusion management mechanisms have been proposed to reduce this problem. Firstly, a sub-sampling scheme based on background subtraction techniques is implemented to obtain stationary foreground regions. At the Second stage, some modifications are introduced on this base algorithm with the purpose of reducing the amount of stationary foreground detected.

A robust abandoned object detection algorithm for real-time video surveillance is proposed by Jiyan Pan, et. al. which is different from conventional approaches that mostly rely on pixel-level processing, author perform region-level analysis [9] in both background maintenance and static foreground object detection. In background maintenance, region-level information is fed back to adaptively control the learning rate. In static foreground object detection, region-level analysis double-checks the validity of candidate abandoned blobs.

An approach based on active contours [16] is proposed for discriminate detected static foreground regions between abandoned and stolen. Firstly, the static foreground object contour is extracted. Then, an active contour adjustment is performed on the current and the background frames. Finally, similarities between the initial contour and the two adjustments are studied to decide whether the object is abandoned or stolen. It provides a robust solution against non-accurate segmentations of stationary objects in the analyzed video sequence.

The proposed approach by Diego Ortego, et. al. [15] for stationary foreground detection comprises two analyses, both at pixel level on frame-by-frame basis, for foreground and motion data. Each analysis has two stages to model spatio-temporal patterns: feature extraction and history image computation. The two resulting history images are combined to get an image representing the foreground-motion variation over time, which is thresholded to get the stationary foreground mask. Finally, occlusion handling is performed to recover lost pixels due to frequent object occlusions in crowds. Background subtraction is applied to detect foreground. For extracting motion using temporal windows, they apply a median filter before and after the frame under analysis. Threshold is calculated from median filters give rise to final motion image; finally motion history image is computed.

The stationary foreground region detection algorithm described in [17], is based on the sub-sampling of the foreground-mask with the aim of detecting foreground changes at different time instants in the same pixel locations. To achieve this, authors use a background subtraction stage based on modeling each pixel with a simple Gaussian distribution. Since it is assumed that the pixels of a stationary region will remain as foreground for a period of time, a number of binary foreground mask samples are collected from the last k seconds. Then, the stationary foreground mask (S) is obtained computing the intersection of the binary foreground mask samples. Finally, each active pixel of S is determined to be part of the stationary foreground regions.

3. PROPOSED WORK

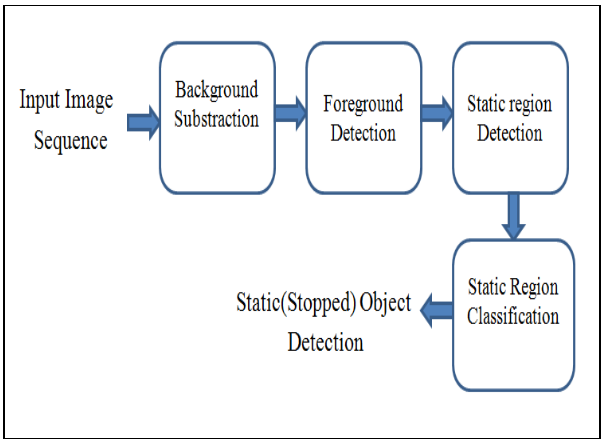


Fig -1: Proposed system architecture

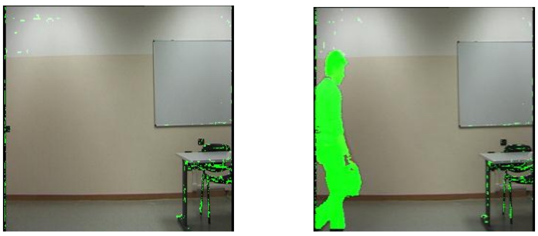
The proposed system is modular in nature. It gives classification of foreground objects into static and moving objects. Static Object Detection is done by using

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056
Volume: 02 Issue: 07 | Oct-2015 www.irjet.net p-ISSN: 2395-0072

stopped foreground subtraction (SFS) algorithm [2], where basic idea is modeling of foreground objects and classifying them as static objects those whose model maintains the same persistent features for several consecutive frames; remaining objects are classified as moving objects.

There are four major concepts defined in SFS algorithm related to pixel state for classifying objects. Those are, whether it is a background pixel, old stopped pixel, old moving pixel or new stopped pixel. For better detection accuracy neural network is used, which detect changes in pixel state by self organizing manner.

Successive image frames from video is taken as input to system. The images having RGB color space is converted into HSV color space, the conversion is taken because in HSV attributes (like hue, saturation, value) are directly corresponds to color concept, which makes it conceptually simple.

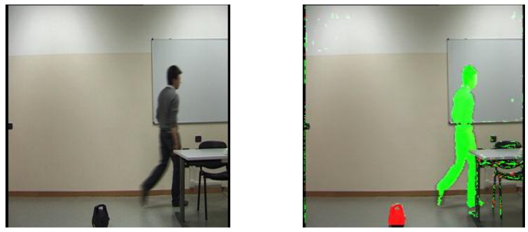


(a) (b)

Fig -2: Initialization of background and foreground in the visor dataset

After color conversions some morphological operations are performed for noise removal. Successive frame differencing gives background frame. Foreground detection then identifies pixels in the image that cannot be explained by the background. The initialization of background and foreground is shown in figure 2. Once background modeling is done only foreground pixels are observed. This background model provides a complete description of the entire background scene. Image classification is done by using neural network, where 3-D neural model for image sequences which automatically adapts to scene changes in a self-organizing manner is used [2].

Once background is identified the remaining pixels are either belongs to moving or stopped pixels from foreground. The stationary state of pixel is given by counting of consecutive co-occurrences of frame. When this value reaches to threshold the static object is detected. If suppose pixel is continuously changing it's location in successive frames it is moving foreground object which is accurately learned by using neural network. If pixel is stopped for some time and again start moving before reaching to threshold value in that case the pixel is foreground pixel which representing moving object.



(a) (b)

Fig -3: Stopped Object Detection

The self organizing approach check for every pixel between consecutive frames for changes. The 3-D neural model and 2-D model differs each other by their layered network structure and intra-layer weights update. Finally the pixel which having the same features for several consecutive frames (up to threshold) are considered as stationary pixels belongs to stopped object. The detection of static object is shown in figure 3. The threshold value for counting co-occurrence of pixel is chosen based on the total number of frames.

4. EXPERIMENTAL RESULTS

For result evaluation multiple video sequences are taken into account. Some of the videos are following same scenario of people walking with bag and leaving it on the ground. All the videos were publically available on internet [18]. These videos are nearly of same frame size. All the videos are having 130-200 image frames. The detection rate varies on the lighting chances, movement of the camera and background changes. The accuracy of SFS algorithm modeled by self organizing approach is shown below. The 10 different videos are taken and its detection parameters like TP, FP, TN and FN are calculated. This parameter leads us to the detection accuracy.

Table -1: Results of SOBS on VISOR Datasets

Image Sequence	True Positive	False Positive	True Negative	False Negative
Panini_cane	1	0	0	0
Chair	1	1	0	0
Aband 1	1	0	0	0
Aband 4	1	0	0	0
Aband 5	1	0	0	0
Aband 6	1	0	0	0
Aband 7	1	0	0	0
Aband 8	1	0	0	0
Aband 9	1	0	0	0
Aband 10	1	0	0	0
Total	10	1	0	0

From above table 1, it is found that accuracy of static object detection on SOBS approach is 90.90%. The above all sequences are also tested on Background segmentation method [11]. The detection accuracy is decreases in Background segmentation than SOBS as depicted in figure 4. The accuracy of static object detection on Background Segmentation approach is found as 71.42%.

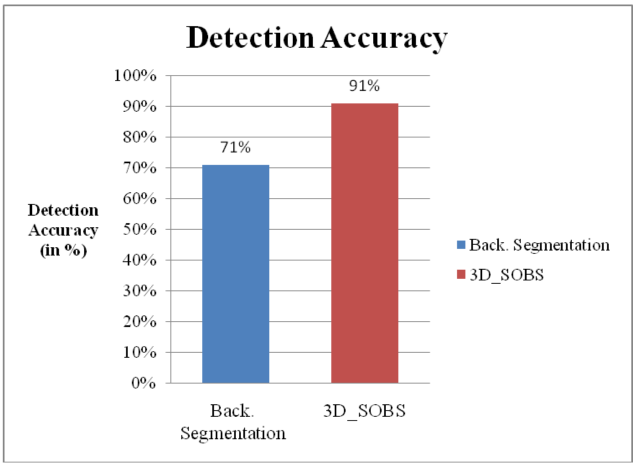


Fig -4: Chart showing accuracy comparison between SOBS and Background Segmentation

5. CONCLUSIONS AND FUTURE SCOPE

A self organizing approach for image sequences combined with Stopped foreground subtraction algorithm gives accurate detection of static objects against moving foreground objects. The accuracy of detection is largely dependent on the robust foreground analysis which separates foreground from background. The system is able to detect the static objects on different sequences with satisfactory results. Experimental results show that the detection accuracy for self organizing background subtraction method is better than the Background Segmentation. In case of low quality sequences SOBS achieves better results than Background Segmentation. The tiny objects are detected in SOBS while in Background Segmentation; the limitation of object area is present.

The static object detection by using self organizing background subtraction has high computational complexity which can be exploited later. The existing approaches have less accuracy for poor quality videos; this will be overcome by illumination invariant algorithms. In future occluded stopped objects detection is challenging task for researcher.

REFERENCES

[1] L. Maddalena and A. Petrosino, “A self-organizing approach to background subtraction for visual surveillance applications,” *IEEE Trans. Image Process., vol. 17, no. 7, pp. 1168–1177, Jul. 2008.*

[2] Lucia Maddalena, Alfredo Petrosino, “Stopped Object Detection by Learning Foreground Model in Videos”, *IEEE Transactions On Neural Networks and Learning Systems, Vol. 24, No. 5, May 2013.*

[3] D. Culibrk, O. Marques, D. Socek, H. Kalva, and B. Furht, “Neural network approach to background modeling for video object segmentation,” *IEEE Trans. Neural Netw., vol. 18, no. 6, pp. 1614–1627, Nov. 2007.*

[4] Álvaro Bayona, Juan C. SanMiguel, and José M. Martínez , “Stationary Foreground Detection Using Background Subtraction and Temporal Difference in Video Surveillance”, *7th IEEE International Conference on Image Processing 2010.*

[5] P. Wang, C. Shen, N. Barnes, and H. Zheng, “Fast and robust object detection using asymmetric totally corrective boosting,” *IEEE Trans. Neural Netw., vol. 23, no. 1, pp. 33–46, Jan. 2012.*

[6] F. Porikli, Y. Ivanov, and T. Haga, “Robust abandoned object detection using dual foregrounds”, *EURASIP J. Adv. Signal Process., Article ID 197875, 2008. 1, 2*

[7] Thi Thi Zin, Pyke Tin, Takashi Toriu and Hiromitsu Hama, “A Probability-based model for detecting abandoned objects in video surveillance systems”, *Proceedings of the World Congress on Engineering 2012 Vol II, WCE 2012, July 4 - 6, 2012, London, U.K.*

[8] Rub´en Heras Evangelio and Thomas Sikora, “Complementary Background Models for the Detection of Static and Moving Objects in Crowded Environments”, *8th IEEE International Conference on Advanced Video and Signal-Based Surveillance, 2011.*

[9] Jiyan Pan, Quanfu Fan, Sharath Pankanti, “Robust Abandoned Object Detection Using Region-Level Analysis ”, *IBM T.J. Watson Research Center, Hawthorne, NY, U.S.A.*

[10] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking”, *TPAMI, 22(8):747–757, 2000.*

[11] Hira Fatima, Syed Irtiza Ali Shah, Muqaddas Jamil, Farheen Mustafa, and Ismara Nadir “Object Recognition, Tracking and Trajectory Generation in Real-Time Video Sequence”, *International Journal of Information and Electronics Engineering, Vol. 3, No. 6, November 2013*

[12] M. Bhargava, C. Chen, M.S. Ryoo, and J.K. Aggarwal, “Detection of object abandonment using temporal logic.”, *Mach. Vision Appl., 20:271–281, 2009. 1, 2*

[13] Y. Tian, A. Senior, and M. Lu, “Robust and efficient foreground analysis in complex surveillance videos”, *Mach. Vision Appl., 23(5):967–983, 2012. 1, 2*

[14] Rajesh Kumar Tripathi, Anand Singh Jalal and Charul Bhatnagar, “A Framework for Abandoned Object Detection from Video Surveillance”, *2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 18-21 Dec. 2013.*

[15] Diego Ortego, Juan C. SanMiguel, "Stationary foreground detection for video-surveillance based on foreground and motion history images", *10th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 27-30 Aug. 2013*

[16] Luis Caro Campos, Juan Carlos SanMiguel, José M. Martínez, "Discrimination of abandoned and stolen object based on active contours", *2011 8th IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS), Aug. 30 2011-Sept. 2 2011.*

[17] Liao,H-H.; Chang,J-Y.; Chen, L-G. "A localized Approach to abandoned luggage detection with Foreground –Mask sampling", *Proc. of AVSS 2008, pp. 132-139.*

[18] VISOR – Video Surveillance online repository. It contains a large set of multimedia data. www.Openvisor.org