A Static Object Detection in Image Sequences

Pravinkumar Yadav
Department of Information Technology
MIT College of Engineering
Pune, India
pravinty@gmail.com

Aditi Jahagirdar
Department of Information Technology
MIT College of Engineering
Pune, India
Aditi.jahagirdar@mitcoe.edu.in

***Abstract*— with concerns about terrorism and global security on the rise, it has become vital to have in place efficient threat detection systems that can detect and recognize potentially dangerous situations, and alert the authorities to take appropriate action. Therefore the scientific challenge is to devise and implement automatic system able to detect static objects from scenne which are not part of background. The proposed system consists of four processing steps: (i) Background subtraction, involving separation of background and foreground,(ii) extracted foreground objects classification as stationary or moving objects,(iii) classified objects investigation for detecting static region of object (iv) final object classification for accurate modeling of background and foreground. The proposed system is able to detect static objects under different illuminations as well as in changing background environments.**

Keywords- SOD-Stopped object detection, RGB-Red, green, blue color space, HSV- Hue, saturation, value color space, SVM- Support vector machine, Background subtraction.

I. INTRODUCTION

In recent year's video surveillance system have become more popular in computer vision. Especially static object detection is the most challenging task in the video surveillance system, as there are more security threats in the areas like airport, railway station etc. The main tasks in visual surveillance systems include motion detection, object classification, tracking, activity understanding and semantic description [6]. The main concentration here is on the detection phase of a general visual surveillance system using static cameras. The detection of static objects in video streams is the relevant step of information extraction in many computer vision applications.

The process of modeling the background and determining the foreground by comparison with the frames of the sequence is often referred to as background subtraction. The usual approach to static object detection is through background subtraction that consists in maintaining an up-to date model of the background and detecting static objects as those that deviate from such a model [2]. If the state of the background is known for every frame of the sequence and there are no changes in illumination, the subtraction can be accomplished by a simple comparison between the background image and a frame of the sequence.

Background subtraction techniques can work well in the highly-crowded scenarios. The existing methods can be divided into two categories according to their use of one or more background subtraction models.

Again for each of the category, it is sub classified as:
1) Frame-to-frame (differencing) analysis and
2) Sub-sampled analysis [4].

II. STATIC OBJECT DETECTION

The detection of static object is more or less changes of the scene that stay in the same position for relatively long time, if object later start again then it is moving object [2]. For example the general overview for static object detection is shown in figure 1. The man is sitting on chair while dog is moved out from scene, but person is still present in scene for certain time period. Hence such static objects have to be detected. Therefore, the unknown objects which cannot move they are considered as stationary.

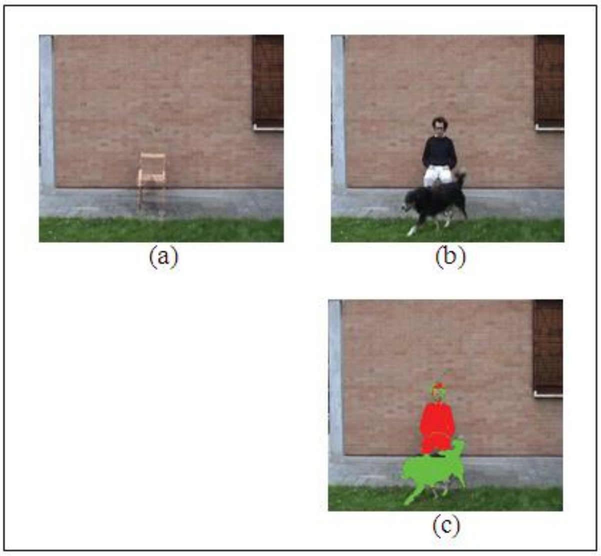

Figure 1. Static object detection general overview [2]
(a) Background of image, (b) foreground objects - man and dog in which man is static but dog is moving, (c) static object detected (in red) while dog (in green) is moving object.

A. What should be detected

Whenever an unknown object appears in the scene and remains stationary for defined amount of time period, an alarm needs to be generated.

III. RELATED WORK

One of the earliest approaches to Static object detection based on background subtraction is reported in [1]. Two processes are distinguished: pixel analysis, which determines whether a pixel is stationary, transient, or background by observing its intensity profile over time, and region analysis, which deals with the agglomeration of groups of pixels into moving and stopped regions. Stationary regions are added as a layer over the background, through a layer management process used to help solving occlusion and restart challenges.

In [3] authors present a system for the detection of static objects that, based on the detection of a dual background model, classifies pixels by means of a finite-state machine. The state machine provides the meaning for the interpretation of the results obtained from background subtraction and it can be optionally used to integrate user input. The system can thus be used both in an automatic and an interactive manner without requiring any expert knowledge from the user. The same authors also propose another system in [8], 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.

In [5], a SOD method based on double-background subtraction is reported that is robust to illumination changes and noise. The long-term background represents a clean scene; all objects located in this background are not considered static objects, but as part of the scene. The short-term background shows static objects abandoned in the scene; this image is made up of the last background image, updated with pieces of the current image. Static foreground objects are detected by thresholding the difference between the short-term and the long-term backgrounds, while dynamic foreground objects are detected by thresholding the difference between the current image and the long-term background.

Dynamic background changes, such as parked cars, left packages, or displaced chairs, are explicitly handled in [9], highlighting the importance of avoiding the mere absorption into the background model of these changes, in the light of intelligent visual surveillance. Dynamic background changes are inserted into short-term background layers, and, in order to update a long-term background model, the layer modeling technique is embedded into a codebook-based background subtraction algorithm that is robust to local and global illumination changes.

The pre-processing involves a dual-time background subtraction algorithm [10] which dynamically updates two sets of background, one after a very short interval (less than half a second) and the other after a relatively longer duration. An algorithm for tracking of abandoned objects even under occlusion is also proposed. The background segmentation is adaptive in nature and based on the Approximate Median Model. It consists of two types of reference backgrounds, Current and Buffered background, each with a different time interval. Blob analysis is done on the segmented background and a dynamic tracking algorithm is devised for tracking the blobs even under occlusion.

Other approaches are based on pixel layer-based foreground detection [7], where temporal persistent objects are introduced. To detect persistent foreground objects, the authors use a color persistence criterion, continually monitoring the color histogram of the foreground object and using correlation to decide if the color is persistent. After a user-defined time threshold, the persistent foreground object is converted to a new background layer. To reconvert persistent objects back to foreground layers when they become interesting again (restart challenge), the authors use higher level features (i.e., globally modeling the region as a whole, based on the use of region level features).

IV. MATHEMATICAL MODEL

Mathematical model is the mathematical representation of the system. The proposed system is represented using set theory. Let M be the proposed system. Input to the system is video which consist of number of consecutive frames. System M is anticipated to obtain Stopped objects from scene. There are Si numbers of states through which the whole system will undergo in order to achieve intended solution. The states interact with each other depending upon some meaningful criteria. The whole system starts and ends with two predefined states i.e. (S0: Start and S6: End).

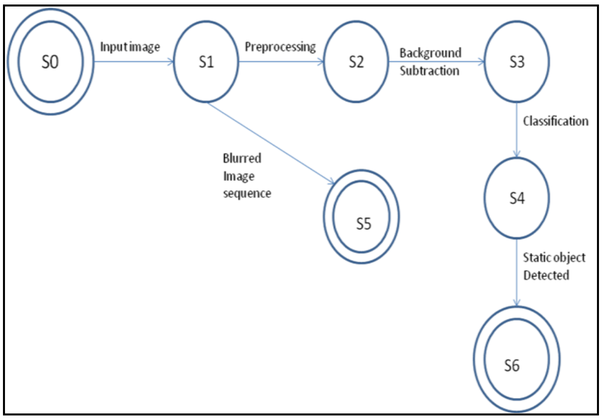

Figure 2. Mathematical model

Consider M be the proposed system,

$$M= \{I, O, F, P, BS, C\}$$

Where,

I= Input images,
O= Static Object Detection State,
F= Failure state,
P= Pre-processing state,
BS= Background subtraction state
C= Classification state

Input:

Number of images in given dataset

$$I= \{I_{src} | I_{src} \in \text{set of images } (i_1, i_2, i_3, \ldots\ldots, i_n)$$

Where,

I_{src}= Image sequence dataset
$I_{1\ldots n}$ = number of images

Output:
Static object detection

$O = \{O_d | O_d \in \text{Stopped object from scene}\}$

Where,

O_d = static object detected

Pre-Processing:

$P = \{I_{p1}, I_{p2}, \ldots, I_{pn}\}$

Where,

$I_{p1} \ldots I_{pn}$ = Number of pre-processed images

Background Subtraction:

$BS = \{(P_{i1}-P_{i2}, P_{i1}-P_{i3} \ldots, P_{i1}-P_{in} \in R_2) \| (P_{i1}-P_{i2}, P_{i2}-P_{i3} \ldots, P_{i9}-P_{in} \in R_1)\}$

Where,

$P_{i1 \ldots n}$= Set of consecutive frames

R_1, R_2= Different frames which detects objects apart from background

Static object detection:

O= {Od | (od ∈ R1) | (od ∈ R2)} ∈ St

Where,

St= threshold time,

O_d = static object detected.

Failure State:

F= {set of distorted & Blurred video}.

V. PROPOSED METHODOLOGY

The proposed system is modular in nature. It consists of different modules which are as follows:

1. Input Video
2. Read frame
3. Color space conversion
4. Noise reduction
5. Background subtraction
6. Foreground detection
7. Static region detection and classification

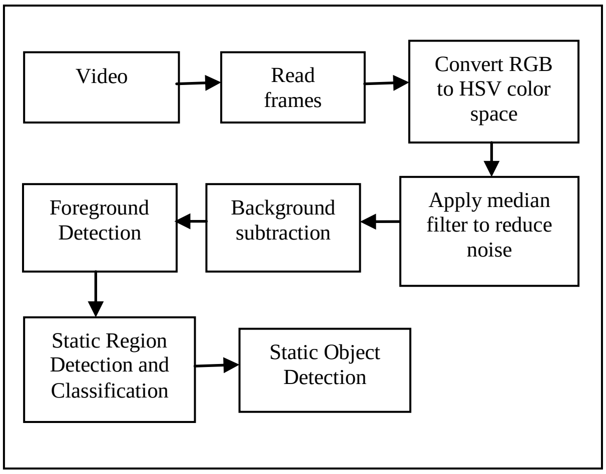

Figure 3. Proposed System Architecture

System consists of following main components:
Module 1: Input Video

In this module the video which is captured by the CCTV/Web camera is considered as an input to the system. Real time images are not taken into account, only already available videos are considered.

Module 2: Read frame
The image frames from video is obtained at defined rate.

Module 3: Color space conversion
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.

Module 4: Noise reduction
The image frames are associated with pepper noise which contains various black and white pixels throughout the image. By applying median filter, the noise is reduced with objects edges sharpening. The median filter is odd matrix filter. First it sorts the numbers in ascending order, and then places median value at consecutive positions. Due to which lowest intensity values from such image frame matrix are replaced and hence relevant noise from image is reduce. There is association of pixels which form large block of pixel this reduces the performance and affect true detection.

Module 5: Background subtraction
There are four major steps in a background subtraction algorithm are performed which are preprocessing, background modeling, foreground detection, and data validation [11]. Preprocessing consists of a collection of simple image processing tasks that change the raw input video into a format that can be processed by subsequent steps. The grayscale color conversion in preprocessing is performed.

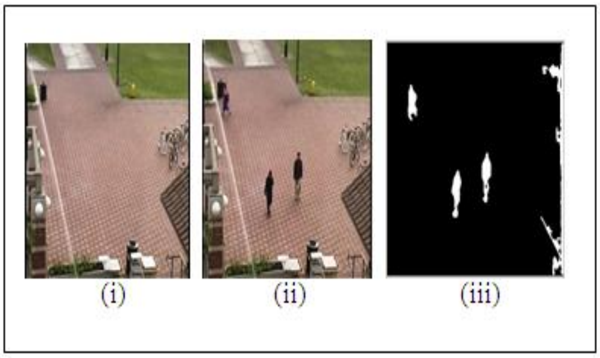

Figure 4. Preliminary Implementation results
(i) Background image, (ii) image with some foreground objects [12], and (iii) actual implementation result.

In Background modeling, it uses the new video frame to calculate and update a background model. The new image frame is subtracted from previous one. Here frame by frame subtraction is performed to differentiate background and foreground. The figure 3 shows

Preliminary implementation results for static object detection system. This background model provides a statistical description of the entire background scene. Only two frames are taken, first is background image and second is foreground image (image after some time interval). By applying bitxor function background subtraction is performed.

Module 6: Foreground detection
Foreground detection then identifies pixels in the video frame that cannot be adequately explained by the background model, and outputs them as a binary candidate foreground mask. The panini_cane video inputed now, then moving (dog) as well as static (man) objects from scene are identified. Now only differenciation between static and moving object from foreground is remaining.

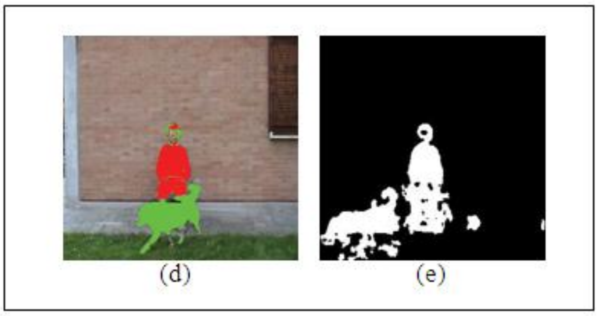

Figure 5. Background subtraction on video [2] (d) Input video sequence (frame) for background subtraction, (e) static (man) and moving (dog) object detected by implementation.

The figure 4 (e) shows different objects detected from foreground.

Module 7: Static region detection and classification
The pixels which remain Stationary for defined time period are detected as static objects. They are inserted into candidate foreground masks. Finally, data validation examines the candidate mask, eliminates those pixels that do not correspond to actual static objects, and outputs the final foreground mask. For accurate detection and reduce the effect of illuminations further classification is performed by neural/SVM.

VI. CONCLUSION AND FUTURE SCOPE

This paper proposed a system for static object detection in video surveillance system. The system is able to detect the background and foreground objects on two different sequences with satisfactory results. Further discrimination between static and moving objects from foreground is in progress.

REFERENCES

[1] R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L. Wixson, "A system for video surveillance and monitoring," *Dept. Nat. Acad. Sci., Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU-RI-TR- 00-12, 2000. J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.*

[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] Rub´en Heras Evangelio, Michael P¨atzold, and Thomas Sikora,"A System for Automatic and Interactive Detection of Static Objects", *IEEE Workshop on Person-Oriented Vision (POV), 2011.*

[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] E. Herrero-Jaraba, C. Orrite-Urunuela, and J. Senar, "Detected motion classification with a double-background and a neighborhood-based difference," *Pattern Recognit. Lett., vol. 24, no. 12, pp. 2079 2092, 2003.*

[6] 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.*

[7] K. Patwardhan, G. Sapiro, and V. Morellas, "Robust foreground detection in video using pixel layers," *IEEE Trans. Pattern Anal. Mach. Intell., vol. 30, no. 4, pp. 746–751, Apr. 2008.*

[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] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Real-time foreground-background segmentation using codebook model," *Real- Time Imag., vol. 11, no. 3, pp. 172–185, 2005.*

[10] A. Singh, S. Sawan, M. Hanmandlu, V.K. Madasu, B.C. Lovell, "An abandoned object detection system based on dual background segmentation", *Advanced Video and Signal Based Surveillance 2009.*

[11] Sen-Ching S. Cheung and Chandrika Kamath, "Robust techniques for Background subtraction In urban traffic video", *IS&T/SPIE's Symposium on Electronic Imaging San Jose, CA, United States January 18, 2004 through January 22, 2004.*

[12] VIP Base: A collection of utilities for video and image processing http://vipbase.net/doc/tool_bgsubtraction.htm