

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

## Abandoned and Removed Object Detection in a Video

Divya Maheshwari<sup>1</sup>, Dr. S. R. Gengaje<sup>2</sup>

P.G. Student, Dept. of Electronics Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India<sup>1</sup>

Head of the Dept., Dept. of Electronics Engineering, Walchand Institute of Technology, Solapur, Maharashtra, India<sup>12</sup>

**ABSTRACT:** We present a general framework for – Abandoned and removed object detection in a Video, based on background updation, background subtraction (BGS), foreground pixel detection and blob analysis. In our system, the background is initialised and updated by Approximate median method. Then, the same background model is used for BGS to detect static foreground regions without extra computation cost. Furthermore, the types of the static regions (abandoned or removed) are determined by using a method that exploits context information about the foreground masks. **We then analyse the extracted object blobs to be classified as abandoned or removed objects.** The efficiency of the proposed method is tested on several public databases, such as The Image library for intelligent detection systems (i-LIDS), IEEE Performance Evaluation of Tracking and Surveillance Workshop (PETS) 2006 datasets, actual CCTV surveillance videos and few self made test videos to depict various possible situations. The test and evaluation demonstrate our method is efficient to run in real-time, however, the work is validated using saved videos and simulated on MATLAB.

**KEYWORDS:** Abandoned object, removed object, background subtraction (BGS), foreground pixel detection, blob analysis, video surveillance.

### I. INTRODUCTION

In recent years, as a sharp increase in terrorist attacks on crowded public places, like airports, stations, subways, entrance to buildings, and other public venues, video surveillance systems have been more demanded. Many methods have been recently proposed to automatically detect abandoned objects (parked vehicles and left-luggage) in video surveillance for different applications such as traffic monitoring, public safety, retail, etc. At train/subway stations, airports, big cities, and other public spaces with high traffic flows, it becomes very challenging for security officers as well as video surveillance solutions to quickly detect objects that have been left behind.

Detecting static objects in video sequences has several applications in surveillance systems such as the detection of illegally parked vehicles in traffic monitoring or the detection of abandoned objects in public safety systems and has attracted the attention of a vast research in the field of video surveillance. In addition to security applications, video surveillance is also used to measure traffic flow, detect accidents in highways and military applications [2]. It alerts the security officers of a burglary in progress or a suspicious individual loitering in a restricted area helping to avoid threat. The automatic detection of abandoned and removed objects in real-time has recently become a topic of great interest in video surveillance domain. These items can be grouped into two main classes, dynamic suspicious behaviours (e.g., a person attempting to attack others) and static dangerous objects (e.g., luggage or bomb abandoned in public places). The scope of this paper falls into the latter category. Specifically, we investigate how to detect objects in a video as abandoned and removed. To improve the quality and the effectiveness of system various algorithms and techniques are suggested and implemented by researchers in various ways. But due to their complexity and probability issues, the implementation was not so fruitful. For security concerns 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 by raising alarm on right time. To detect abandoned and removed objects, we focus on how to detect static regions that have recently changed in the scene and how to determine whether they correspond to abandoned or removed objects.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

In general, the abandoned and removed detection is achieved by developing a system with the following analysis stages: foreground segmentation, stationary region detection, blob classification and abandoned and re-moved object discrimination [1]. The last stage of this pipeline determines the system ability to discriminate stationary foreground objects between abandoned and removed. For its implementation, the common approach is to study the similarities between features extracted from the current and background frames of the video sequence. This makes it easy for the security officials to bring the situation to notice thus preventing any security threat from occurring. The system captures the videos using static cameras.

Paper is organized as follows. Section II describes the previous research work in the concerned area, followed by section III which explains our algorithm for detection of abandoned and removed objects in a video. The flow diagram represents the various steps of the algorithm. Section IV presents experimental results showing the snapshots of the results of videos tested. Finally, Section V presents conclusion and future scope of the research work.

## II. RELATED WORK

Even though there exist numerous of object detection algorithms in the literature, most of them follow a simple flow diagram, defined by Cheung and Kamath [3], passing through four major steps, which are (1) pre-processing (simple image processing tasks that change the raw input video into a format that can be processed by subsequent steps), (2) background modelling (also known as background maintenance), (3) foreground detection (also known as background subtraction) and (4) data validation (also referred to as post-processing, used to eliminate those pixels that do not correspond to actual moving objects).

Stationary objects in multiple object tracking [4], detects the foreground objects with several moving objects and is inspired by human's visual cognition processes. It relies on tracking information to detect drop-off events. This system produced larger errors under bright lighting conditions.

Martiriggiano, Caroppo et al [5] a gradient-based method is applied to the static foreground regions to detect the type of the static regions as abandoned or removed objects. It does this by analysing the change in the amount of edge energy associated with the boundaries of the static foreground region between the current frame and the background image. They proposed a motion detection system, based on background subtraction algorithm, able to classify static foreground regions as abandoned or removed objects. A template matching procedure is applied between the edge of the foreground region and the edge detected over the segmented image.

Magno *et al* [6] employs an unobtrusive embedded platform. This method uses a wireless video sensor to detect the abandoned object. The system employs multimodal sensor integration which saves power consumption. The objective is to develop a multimodal video sensor with low power and low cost to detect abandoned objects. This uses new algorithms for energy efficient image processing without giving up the flexibility of in-field configuration. In spite of using a video sensor, the number of false positives is 13% of the total detected objects.

Singh *et al* [7] uses a dual-time background subtraction algorithm to dynamically update two sets of background. This method is dynamic, adaptive, non-probabilistic and intuitive in nature. It uses pixel color/ intensity information for background processing. The binary image is divided into a number of legitimate blobs. Once the blobs are generated, the system applies an algorithm for tracking of the abandoned objects. The system is robust to variations in lighting conditions and the number of people in the scene. The system does not classify stable objects as unattended and removed objects.

In robust detection of abandoned and removed objects in complex surveillance videos [1], the method detects abandoned and removed objects using GMM algorithm. The type of static regions is determined by a method that exploits context information. A matching method is used to detect the abandoned and removed object and it outperforms the edge based techniques. A person-detection process is integrated to differentiate static objects from

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

stationary people. The system is robust to quick-lightning changes and occlusions. The accuracy of the detection is influenced by the size of the object, light conditions, and contrast situations.

### III. OUR APPROACH

In this paper, we propose a general framework for detection of abandoned and removed objects in a video. We define an abandoned object to be a stationary object that has not been in the scene before, and a removed object to be a stationary object that has been in the scene before, but is not there anymore [1]. The overall system (see Fig. 1) is modular in nature and consists of six disparate modules with each block acting as a discreet processing unit making it easy to modify any block, provided the input and output data types remain compatible with the connecting blocks. The system includes the following modules. (i) Pre-Processing, (ii) Background Modelling, (iii) Foreground detection for Abandoned and Removed Objects, (iv) Blob Formation, (v) Feature Calculation (Blob analysis), (vi) Alarm Raising.

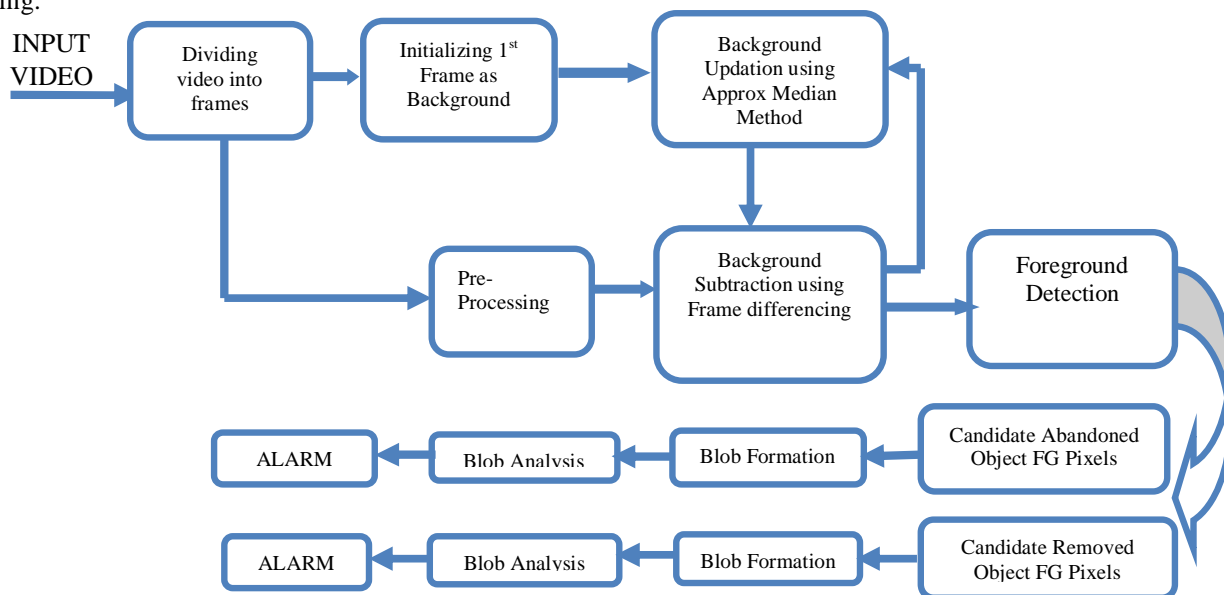


Fig.1. Flow Diagram of the proposed system

#### A. INPUT VIDEO:

For this system we are considering the recorded video as an input to the proposed system. The video stream is initially segmented into individual RGB frames. The system is compatible with any video format and resolution, however, the video resolution is scaled down to around 320\*480 to enhance the system performance. Generally the frame rate of the standard videos is 30fps or 19fps or 22fps. The proposed work will be simulated using MATLAB and Tool Boxes and will be validated using standard videos available on internet.

#### B. PRE-PROCESSING MODULE:

The input video is which is converted to frames are needed for further processing. Once the frame is grabbed, it has to go through the chain of the pre- processing. The first process is gray scaling. In image processing, a grayscale image is an image which the value of each pixel is a single sample, or we called it intensity value. This sort of image is typically composed of shades of gray, varying from black at the weakest intensity to white at the strongest intensity. The grayscale image simplifies computation drastically, compared to a color RGB image. The next process done is called as contrast enhancement which improves the perceptibility of objects in the scene by enhancing the brightness difference between objects and their backgrounds. The next step is to use median filter to filter the image. The non-linear neighbourhood operations are performed for the purpose of noise reduction that can do a better job of preserving edges than simple smoothing filters. In median filtering, the neighbouring pixels are ranked according to brightness (intensity) and the median value becomes the new value for the central pixel. After noise filtering the next

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

step is to convert the gray scale image into binary images. These filtered frames are then fed as input to the Foreground detection module.

## C. BACKGROUND MODELLING:

The algorithm is based on building the so called background of the scene and comparing each current frame with the background. The proposed algorithm to separate background and foreground in the incoming image is based on the 'Approximate Median Model' [8]. The first frame of the incoming video is initialized as 'Current Background'. Subsequently, the intensity of each pixel of this current background is incremented by one if the corresponding pixel in the current frame is greater in value or decreased by one if the current frame pixel is less in value. Each pixel in the updated background is then converged to a value for which half the updating values are greater than and half are lesser than this value, i.e. the median. In this manner the objects which are in motion during the initialization of the background will not affect the background model. This technique requires the storage of only one reference image and, hence, computationally inexpensive. The background image when finally converged, is an image of the background without any moving objects.

One problem in constructing and updating the background image is the moving objects or humans remaining motionless for longer periods of time. This way, they gradually merge into the background. Reducing the frequency of update will solve this problem because the position of objects will be random in nature and will not clutter in the same area. And before the completion of convergence and After performing a binarization process with a predefined threshold using the differencing method, we can find the target contour, and the target blob for abandoned object is obtained through the contour-filling process. Hence, an appropriate background image is obtained and, the lighting changes are updated with subsequent images.

## D. FOREGROUND DETECTION AS ABANDONED AND REMOVED OBJECTS:

Frame differencing is the simplest method for the detection of moving objects (foreground), because the background model is equal to the previous frame in image differencing. However, if the target is moving very quickly, then the blob may contain too many background pixels, whereas if the target is moving very slowly, then target information in the blob may be lost. However image differencing between current frame and a background does not suffer from the problems of differencing between successive frames and provides a useful method of segmentation [8]. The first stage of image segmentation is image differencing. Each successive image is subtracted from the background image, and then the difference image is thresholded to be converted into a binary image. Image differencing between the current frame and a reference background image gave a better segmentation result than differencing between the successive frames because it did not produce false positives where a dark shadow had removed away from an area of background. Candidate foreground pixels for abandoned object are defined as those below the abandoned threshold, i.e. the abandoned object pixels are assumed to be significantly darker than the background. The candidate foreground pixels for removed objects are defined as those above the removed threshold, i.e. the removed object pixels are assumed to be significantly brighter than the background.

## E. BLOB FORMATION:

The result of foreground detection is represented as two binary images (one for candidate abandoned object pixels and another for candidate removed object pixels) with the white portion representing foreground (blobs). In this module, we divide the binary image from the previous unit into a number of legitimate blobs. Each blob represents an object. In this step, morphological operations are performed on the binary images of the candidate foreground pixels. Morphological operations are affecting the form, structure or shape of an object, applied on binary images [9],[10]. They are used in post processing (filtering, thinning, and pruning). The two principal morphological operations done are dilation and erosion. Dilation allows objects to expand, thus potentially filling in small holes and connecting disjoint objects. Erosion shrinks objects by etching away (eroding) their boundaries.

## F. BLOB ANALYSIS AND ALARM RAISING:

Candidate foreground object pixels, which are in the binary images (one for candidate abandoned object pixels and another for candidate removed object pixels) undergo the blob analysis [7]. Blob analysis takes as an input a binary image, creates a region counter. The binary image is scanned from left to right and from top to bottom. For

# International Journal of Innovative Research in Science, Engineering and Technology

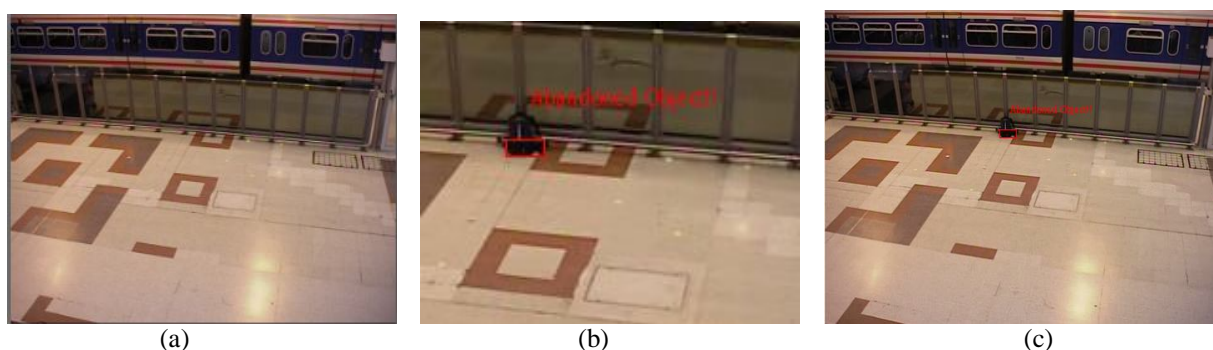
(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

every pixel the north and west pixel (4-connectivity) or the northeast, north, northwest, and west pixel (8-connectivity) is checked for a intensity value of 1 in the binary image. If none of the neighbors fit the criterion then assigning to region value of the region counter and increment region counter. If only one neighbor fits the criterion, assigning pixel to that region. If multiple neighbors match and are all members of the same region, assign pixel to their region. If multiple neighbors match and are members of different regions, then assigning pixel to one of the regions and indicating that all of these regions are the equivalent. Scan image again, assigning all equivalent regions the same region value. Thus, corresponding blob images are formed( one for abandoned blobs and one for removed blobs).

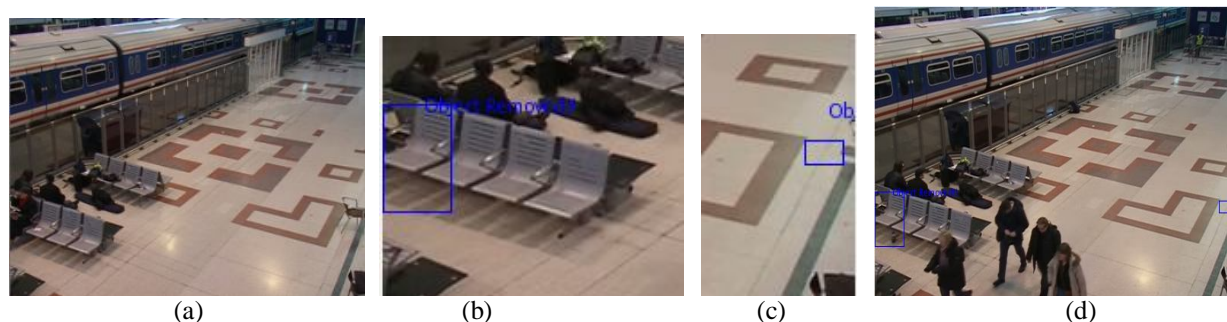
The next process in object detection is tracking the different candidate blobs so as to find which blobs correspond to abandoned objects and removed objects. The first step in this process is to create a set, Track, whose elements have two variables: hitcount\_abandoned and user defined Abandoned\_counter\_threshold( for abandoned object blobs); and hitcount\_removed and user defined removed\_counter\_threshold (for removed object blobs). The next step is to analyze the incoming binary image for all the blobs. We then run a loop through all the elements of the set. If the hitCount goes above a user defined counter\_threshold value, a visual alarm is triggered for the corresponding abandoned or removed object. This counter calculates the time for which the corresponding object stays abandoned or removed in the input video. The detection of abandoned object is shown through red rectangular block.and the detection of removed object is shown through blue rectangular block, in the input video stream.

## IV. EXPERIMENTAL RESULTS



**Fig 2.** Test sequence S1T1C3 (Courtesy of Pets), 2006, Resolution 480\*320, 30fps

(a) Frame#1: Initial frame, (b) Frame#2000: Part of frame with visual alarm for detected abandoned object, (c) Frame#3022: Last frame of sequence with alarm.



**Fig 3.** Test sequence S1T1C4(Courtesy of Pets), 2006, Resolution 480\*320, 30fps

(a) Frame#1: Initial frame, (b) Frame#900: Part of frame with visual alarm for detected removed object, (c)Frame#2900: Part of frame with visual alarm for detected second removed object, (d) Frame#3100: Last frame of sequence with alarms.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015



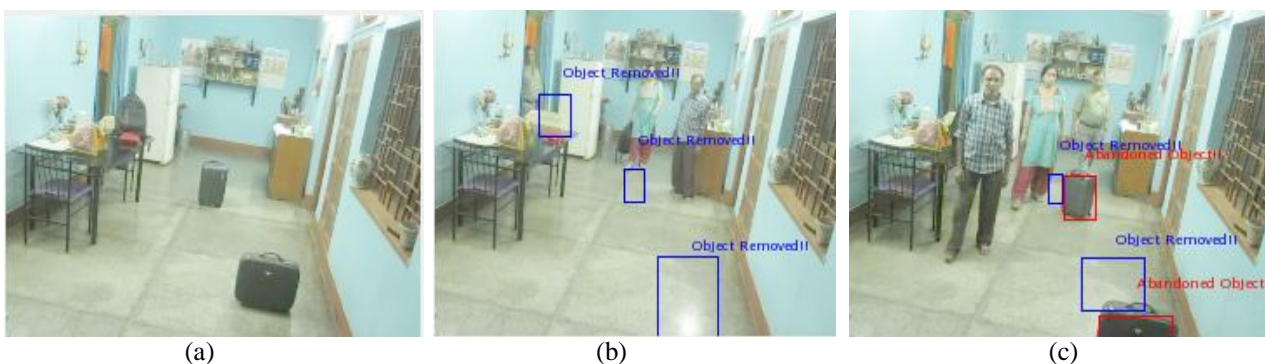
**Fig 4.** Test sequence AB-easy (Courtesy of *i-LIDS*), AVSS-2007, Resolution 480\*376, 30fps

(a) Frame#11: One of the initial frames, (b) Frame#1100: Part of frame with visual alarm for detected abandoned object, (c) Frame#2100: Last frame of sequence with alarm.



**Fig 5.** Test sequence Corridor (CCTV Footage: Courtesy WIT, Solapur), 2015, Resolution 536\*480, 20fps

(a) Frame#10: One of the initial frames, (b) Frame#1100: Frame with visual alarm for detected multiple abandoned and removed objects.



**Fig 6.** Test sequence Indoor (Self capture test video), 2015, Resolution 240\*180, 30fps

(a) Frame#10: One of the initial frames, (b) Frame#1500: Frame with visual alarm for detected removed objects, (c) Frame#2900: Last frame of sequence with alarm.

# International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

The algorithm described above is simulated using MATLAB 2014a, and is applied on standard benchmark datasets for obtaining experimental results (Fig.2 to Fig. 6). The experiments were computed on Intel Core 2 Duo Processor T7250 (2GHz) with 3 GB RAM. Five different situations involving different crowd densities and various sizes of objects were selected, which include PETS 2006 dataset, AVSS 2007 i-LIDS Abandoned Baggage Training dataset, one CCTV footage of a college corridor with numerous objects and having occlusions, and an indoor test video, snapshots of which are shown above. The results can be summarized as shown in Table. 1.

Sr. No.	DATASET	$N_{Event}$	$N_{True\ Positives}$	$N_{False\ Positives}$	$N_{False\ Negatives}$
1	PETS S1-T1-C3	1	1	0	0
2	PETS S1-T1-C4	3	2	1	0
3	AVSS i-LIDS	1	1	0	0
4	CCTV College Corridor	10	8	2	2
5	Indoor	5	5	0	0

Table 1. Summarized Results

$N_{Event}$ : Total number of events occurred in the test sequence;  $N_{True\ Positives}$ : Number of correct detections;  $N_{False\ Positives}$ : Number of false detections;  $N_{False\ Negatives}$ : Number of objects remained undetected.

## IV. CONCLUSION

This paper presented an abandoned and removed object detection system for recorded videos, based on a background segmentation scheme. The background segmentation is adaptive in nature and based on the Approximate Median Model. Blob analysis is done on the segmented background and a dynamic tracking algorithm is devised for tracking the blobs even under occlusion. Detection results show that the system can handle small variations in lighting conditions and also number of people in the scene. The system can work with any video format and resolution, however scaling is done to improve the effectiveness of the algorithm in real time. In addition, the system is simple and computationally less intensive as it avoids the use of expensive filters while achieving better detection results.

As a future work, efforts can be put in a direction to remove limitation given by static thresholds; moreover, the system can be made more robust to work under daylight or outdoor conditions.

## V. ACKNOWLEDGEMENT

I would like to seize the opportunity to express my sincere gratitude towards Head of Department and guide Dr.S.R.Gengaje, electronics department faculty members, non-teaching staff, and my fellow classmates for their invaluable co-operation and guidance.

## REFERENCES

- [1] YingLi Tian, Rogerio Schmidt Feris, , Haowei Liu, Arun Hampapur, and Ming-Ting Sun, "Robust Detection of Abandoned and Removed Objects in Complex Surveillance Videos", IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, VOL. 41, NO. 5, SEPTEMBER 2011, pp. 565-576
- [2] E. Auvinet, E. Grossmann, C. Rougier, M. Dahmane, and J. Meunier, "Left-luggage detection using homographies and simple heuristics," in *PETS*, 2006, pp. 51–58.
- [3] Cheung S-C, Kamath C. Robust techniques for background subtraction in urban traffic video. In S. Panchanathan and B. Vasudev, editors, Proc Elect Imaging: Visual Comm Image Proce 2004 (Part One) SPIE, 2004; 5308: 881-892
- [4] W. Wang and Z. Liu, "A new approach for real-time detection of abandoned and stolen objects," in IEEE Int'l Conference on Electrical and Control Engineering, 2010, pp. 128–131.
- [5] Tommaso Martiriggiano, Andrea Caroppo, Marco Leo, Paolo Spagnolo, and Tiziana D'Orazio , An Innovative Approach for Abandoned or Removed Objects Detection.

## International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 9, September 2015

- [6] M.Magno, F.Tombari, D.Brunelli, L. Di Stefano and L.Benini 2009 Multimodal abandoned/removed object detection for low power video surveillance systems in Proc AVSBS09, pp 188-193
- [7] A.Singh, S.Sawan, M.Hanmadhu, V.K.Madasu, B.C.Lovell 2009 An abandoned object detection system based on dual background segmentation in Proc. IEEE Conf. Adv. Video and Signal Based Surveillance.
- [8] N.J.B. McFarlane and C.P. Schofield, "Segmentation and tracking of piglets in images", Machine Vision and Applications, vol. 8, 1995, pp.187-193.
- BOOKS:**
- [9] Digital Image processing usinf MATLAB, By Gonsalves and woods.
- [10] A handbook of image and video processing, edited by Al Bovik
- Datasets:**
- [11] PETS 2006 Benchmark Data, <http://www.cvg.reading.ac.uk/PETS2006/data.html>
- [12] i-Lids dataset for AVSS 2007, [http://www.eecs.qmul.ac.uk/~andrea/avss2007\\_d.htm](http://www.eecs.qmul.ac.uk/~andrea/avss2007_d.htm)