CrossMark

Abandoned or removed object detection from visual surveillance: a review

Rajesh Kumar Tripathi[1] · Anand Singh Jalal[1] · Subhash Chand Agrawal[1]

Abstract

Intelligent Visual Surveillance is an important and challenging research field of image processing and computer vision. To prevent the ecological and economical losses from bomb blasting, an intelligent visual surveillance is required to keep an eye on public areas, infra-structures and discriminate an unattended object left among multiple objects at public places. An unattended object without its owner since a long time at public place is considered as an abandoned object. Identification of an abandoned object on real-time can prevent the terrorists attack through an automated video surveillance system. In recent decade, a good number of publications have been presented in the field of intelligent visual surveillance to identify the abandoned or removed objects. Furthermore, few surveys can be seen in the literature for the various human activity recognition but none of them focused deeply on abandoned or removed object detection in a review. In this paper, we present the state-of-the-art which demonstrates the overall progress of abandoned or removed object detection from the surveillance videos in the last decade. We include a brief introduction of the abandoned object detection with its issues and challenges. To acknowledge to the new researchers of this field, core technologies, and frequently used general steps to recognize abandoned or removed objects have been discussed in the literature such as foreground extraction, static object detection based on non-tracking or tracking approaches, feature extraction, classification and activity analysis to recognize abandoned object. The objective of this paper is to provide the literature review in the field of abandoned or removed object recognition from visual surveillance systems with its general framework to the researchers of this field.

1 Introduction

Usual and unusual Human Activity Recognition from Visual Surveillance is an active research area of image processing and computer vision that involves identification of human activities

✉ Rajesh Kumar Tripathi
  rajesh.tripathi@gla.ac.in; rajeshkumar.tripathi@gmail.com

[1] Department of Computer Engineering and Applications, Mathura, India

and categorization of them into normal and abnormal activities. Normal or usual activities are the daily routine life activities performed by the human beings at public places such as walking, running, jogging, hand clapping and hand waving where as Abnormal activities are the unusual or suspicious activities performed by the human beings at public places rarely such as left luggage for explosive attacks, running crowd, theft, fights and attacks, vandalism and crossing borders. Normal activities are the usual activities that are not dangerous for the human of the world but abnormal activities may be dangerous for all over the world. Among all the abnormal activities, the detection of abandoned object is a crucial activity that must have higher priority than other abnormal activities to prevent the explosive attacks. Video Surveillance captures images of moving objects in order to watch abandoned object, assault and fraud, comings and goings, prevent theft, as well as manage crowd movements and incidents [36]. Therefore, a fully automatic effective and efficient intelligent video surveillance system is to be developed. The Intelligent Surveillance System can detect un-attempted objects in different situation at public place shown in Fig. 1(a-d).

In the field of visual surveillance, less number of literature reviews has been presented to explore the progress of human activity recognition. Few papers have listed in Table 1 which shows the progress in the field of normal and abnormal human activity recognition. A few surveys have been proposed in the field of normal and abnormal activity recognition from visual surveillance but specifically a progress in the field of abandoned object detection from visual surveillance is required to be focused. The contribution of this paper is to present the progress in the field of abandoned or removed object detection. Researcher of this field can gain more knowledge about the core technologies such as single and dual foreground frame extraction, different methods of static object detection, different classification approaches to classify human and non-human objects; distinct object analyzing approaches such as finite state machine or spatial temporal analysis; that are applied to develop an intelligent visual surveillance system to detect the abandoned or removed object.

The rest of the paper is structured as follows: Section 2 discusses the issues and challenges in abandoned or removed object detection from visual surveillance. An overview of the progress in the past decade in the field of abandoned or removed object detection is discussed in section 3. The general framework for abandoned or removed object detection is discussed in Section 4. Section 5 presents the Evaluation measures and Datasets used for abandoned or removed object detection from surveillance video. Finally, the last section presents conclusion and future work.
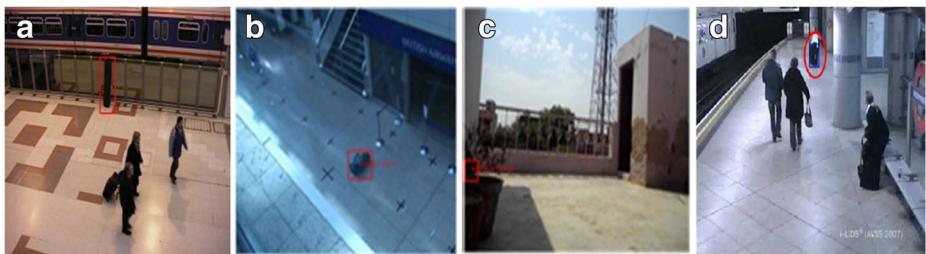

Fig. 1 (**a**) Abandoned Object Detection in Simple Video [107] (**b**) Abandoned Object Detection in Night Video Scene [107] (**c**) Abandoned Object Detection of very small object hidden in flower pot [107] (**d**) Abandoned Object Detection in complex video [105]

Table 1 Related Literature Survey

Paper	Year	References
Semantic Human Activity Recognition: A Literature Review	2015	[122]
Video-Based Abnormal Human Behavior Recognition—A Review	2012	[85]
Human Activity Analysis: A Review	2011	[1]
Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms	2010	[12]
A survey on vision-based human action recognition	2010	[86]
A Survey on Visual Surveillance of Object Motion and Behaviors	2004	[43]

2 Issues and challenges

To develop an intelligent video surveillance system for the automatic recognition of an abandoned object; there are various issues and challenges [12, 118]:

Illumination changes The foreground moving and static object detection is complex to process reliably due to dynamic changes in natural scenes such as gradual illumination changes caused by day-night change and sudden illumination change caused by weather changes or switching on the lights in the indoor videos. Various illumination effects have been shown in Fig. 2.

Shadow of objects Shadow changes the appearance of an object, which creates problem to track and detect the particular object from the video. Some of the features such as shape, motion, and background are more sensitive for a shadow. Shadow changes the shape of an object which creates problem in identification.

More crowds To detect the object from more crowded area (shown in Fig. 3(d)) is very challenging task. In such situation, abandoned object detection is very difficult.

Partial or full object occlusions In video, sometimes, objects are occluded partially or completely. This creates a problem to identify the object correctly. Partial occluded examples are shown in Fig. 4(a)-(b). In general, there are three types of occlusion which have been shown in Fig. 4(c)-(e).

Blurred objects Blurred objects segmentation is very difficult as well as finding features to identify the particular objects. Figure 3(e) shows the blurred objects in an image which is very difficult to recognize.
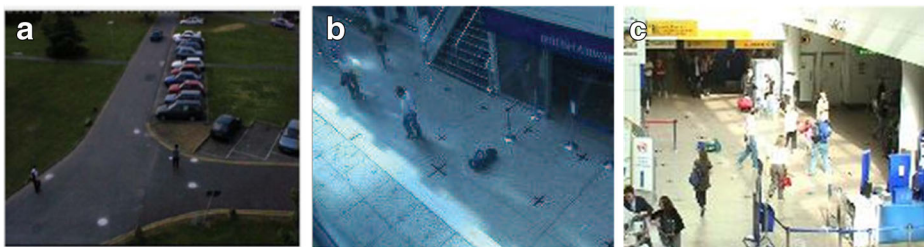

Fig. 2 (**a**) Sudden illumination change by weather changes [82] (**b**) Illumination in night due to the light effects [83] (**c**) Illumination effect in day time [83]
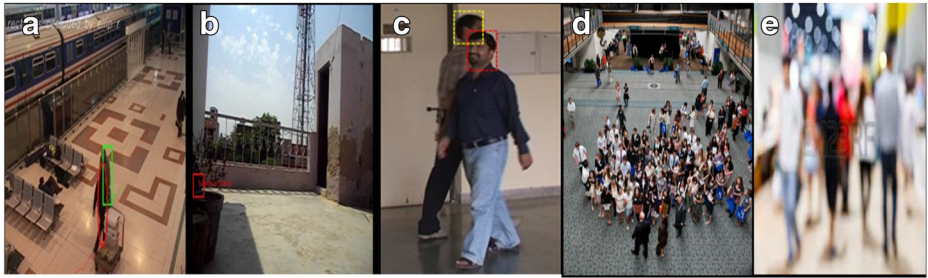
Fig. 3 (**a**) Partial occluded human beings [121] (**b**) Partial hidden abandoned object in a flower pot [108] (**c**) Occlusion with other object [47](**d**) More crowd [44] (**e**) Blurred image [45]
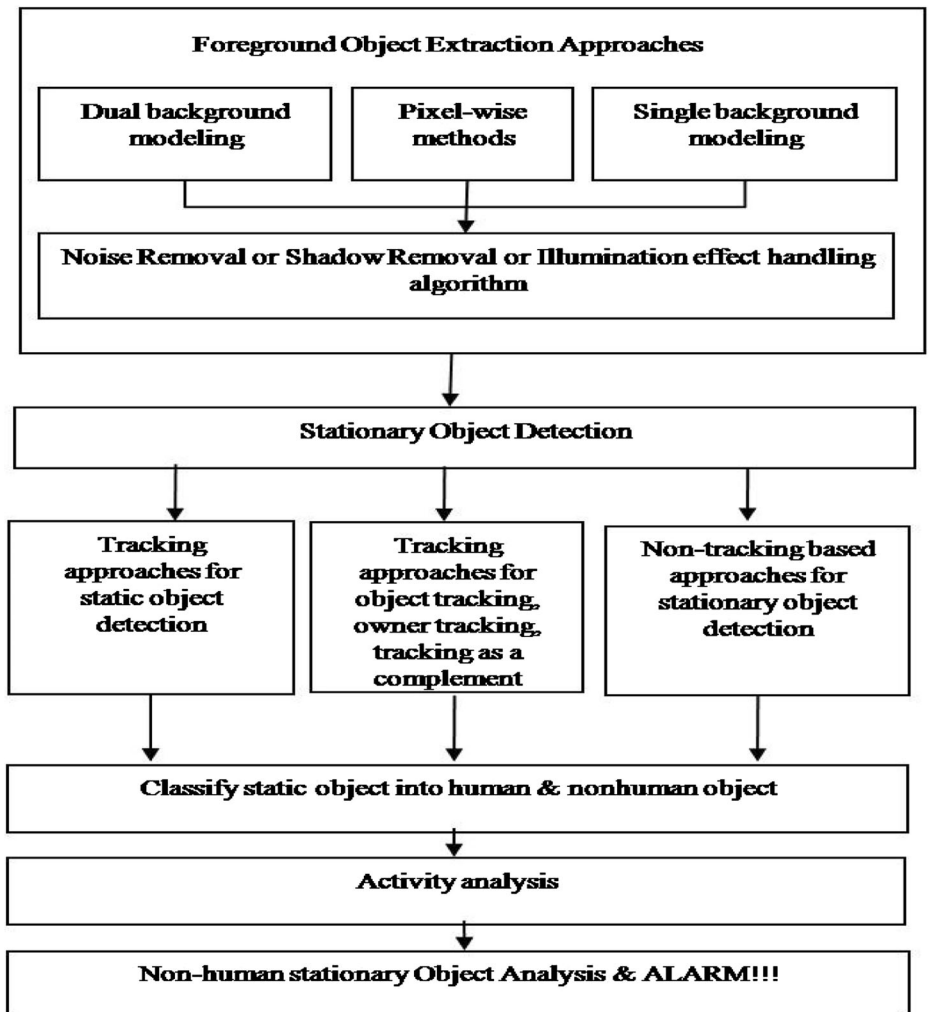

Fig. 4 General Framework for Abandoned or Stolen Object Detection from Single Static Camera

Partial or full object occlusions In video, sometimes, objects are occluded partially or completely. This creates a problem to identify the object in case of partial or full occlusion. Partial occluded examples are shown Fig. 3(a)-(c).

Poor resolution To detect the foreground objects from videos having poor resolution is very challenging task. Object boundaries identification becomes very difficult that causes incorrect object classification.

Real-time processing The more challenging task is to develop an intelligent system which works on real-time. The videos which are having complex background take more time to process it at the time of foreground object extraction and tracking of the objects. Processing time reduction is difficult and challenging for complex videos.

Static object detection In abandoned object detection, static object detection is a challenging task through the background subtraction because this method detects only the moving objects as a foreground.

Low contrast situation Low contrast situation such as black baggage with black background creates a major problem in background subtraction that causes a failure of visual surveillance system to recognize the abandoned or removed object.

3 Researches in abandoned or removed or stolen object detection from video surveillance

This section covers the progress in the field of Abandoned Object Detection from Video Surveillance till this date from single static camera, multiple static cameras and moving camera.

3.1 Research in abandoned or removed or stolen object detection from single static camera

Abandoned/stolen object detection is very difficult in case of highly crowded area, fully occluded objects, and sometimes partially occluded objects from single static cameras. Several researchers have worked to detect an abandoned or stolen object from the video surveillance to protect the people and public infrastructure from the bomb blasting performed by terrorists. There is no predefined shape and size of abandoned or removed objects. These can be in any shape and size like small and big baggage, any hidden object behind the wall or other objects, etc. Many works have been done in this field for single static cameras. This section presents the progress in the field of abandoned or removed or stolen object detection from static cameras.

In the decade of 1990s, intelligent surveillance systems have been developed in order to prevent the dangerous activities [34, 91]. The requirements of the user indicate that an intelligent visual surveillance system should be able to alert through the alarm when a dangerous situation occurs in public places. Content based retrieval technique, video-event shot detection and indexing algorithms has been suggested in [101, 102] then content based retrieval techniques and surveillance system was joined in [102]. In this work, presence of

abandoned object and video event shot including the person who left object was considered. The results were capable, but the system detects video-event shots consisting of a fixed number of frames. This problem has been resolved in [33] by introducing a new method to approximate the number of frames containing the people that was carrying the abandoned object and a novel approach that has layered content-based retrieval of video-event shots referring to potentially interesting situations. Interpretation of events is utilized for indexing criteria and defining new video-event shot detection. Interesting events refer to dangerous situations such as abandoned objects. This system provides the success rate 71% to detect abandoned object.

Later, a method to decide a person carries a baggage or not has been proposed in [39, 48]. In this decade, most of the works used intensity-level images proposed in [92] and W4 system [40]. Lately, few researchers have moved on to use color images such as the system proposed by Yang et al. [117], and Beynon et al. [6]. Bird et al. [8] also utilized color information for background segmentation. But these works present results on scenes having very few and non-occluding people.

In the decade of 2000, a good number of the researchers [8, 14, 16, 22, 27, 31, 32, 35, 49, 71, 73, 75, 78, 81, 90, 97–99, 116] presented an intelligent video surveillance system based on tracking technique to recognize drops off events, but these approaches are not suitable for the occlusion handling as well as object or human tracking in more crowded environment. For example, Spengler et al. [99] employed a Bayesian multi-people tracking module to track arbitrary number of objects. Kim et al. [98] employed Markov Chain Monte Carlo tracking to track blobs but it cannot discriminate between people and human being. In the same year, Lv et al. [69] employed blob tracker and human tracker separately to discriminate the human and abandoned object. Lu et al. [67] tried to identify objects through shape matching and object tracking. They also discriminated objects as an abandoned object and a ghost using the in-painting method which requires high computational cost. Lu et al. [68] incorporated tracking of moving objects based on shape and color features and Kalman-based filtering; and performed the classification by using Eigen features and Support Vector Machine. A package is defined as a non-human object and package ownership analysis performed using HMM-based human activity recognition. Lu et al. [66] presented a knowledge based approach to detect unattended objects that is based on accumulated knowledge about non-human and human objects from object tracking and classification. Li et al. [59] incorporated principle color representation (PCR) into a template-matching scheme to track moving objects, and also by estimating the status of a stationary object i.e. occluded or removed.

Quanfu Fan et al. [27] presented a ranking technique to detect the abandoned object with false reduction. Mahin et al. [71] exploited an approach to detect abandoned and stolen objects which incorporates the background subtraction, blob analysis, and statistics to determine static objects as abandoned or stolen objects.

The recognition of abandoned, removed or stolen object was considered when owner of the baggage is not present in the scene of the video. Owner tracking is a special and good approach that has been employed in various researches such as [7, 62–64]. Liao et al. [62] applied skin color technique and contour matching using Hough transform to track the owner. Still, these tracking approaches often fail due to appearance pattern changes and occlusions. Bhargava et al. [7] proposed a framework that utilizes spatial-temporal and contextual cues to detect baggage abandonment. In this framework, if an un-attempted object is discovered in the video, the system tracks to the previous video frames to recognize owner of that object. Lin et al. [63] presented a pixel based finite state machine to detect the stationary object and perform spatial-

temporal analysis after back tracing to the owner of the luggage. This approach performed better than the [59, 62, 72, 89, 105] on PETS 2006 dataset and [24, 25, 89, 105] on AVSS 2007 dataset. This approach handles temporary occlusions in back-tracking and it is very efficient to implement because only the foreground blobs in limited space-time windows are considered. Lately, Lin et al. [64] again proposed a method in which short term and long term modeling were used to extract foreground object where each pixel is classified in 2-bit code. This approach presented better precision and recall than [3, 27, 59, 62, 89, 105] on PETS 2006 dataset and [24, 25, 62, 89, 105] on AVSS 2007 dataset.

As we have seen above various research works based on object tracking approaches are not able to handle the occlusion problem and false positive rate could not reduced. Therefore, few researchers have employed the tracking as a complement to reduce the false positives. In [78, 93, 105], tracking has been applied to use the tracking information as a complement to reduce the false positive rate. Tian et al. [105] presented a robust and efficient framework to detect abandoned or removed objects using background subtraction with complement of tracking information to reduce false alarms. Recently, Nam [78] proposed a real time based abandoned object detection and stolen object detection using spatial-temporal features. Adaptive background modeling has been used for the removal of ghost image and stable tracking. Proposed method employed a vector matching algorithm to detect partial occluded object and also employed a tracking trajectories to reduce the false alarms.

A new approach has been introduced with the help of tracking to create feedback mechanisms for handling stopped and slow moving objects. In this aspect, Tian et al. [106] proposed a new framework to utilize two feedback mechanisms of interactions between tracking and BGS to handle stopped and slow moving objects to improve the tracking accuracy. This framework employed phong shading model to handle quick lighting changes, mixture of Gaussian methods for moving object detection, and edge energy and region growing method to classify abandoned and removed objects. The use of tracking feedback yields the percentage ground truth frames missing is 8.1% in PETS D1 C1 while 21.2% without feedback. However, this approach fails in low contrast situations where the color of the object and background is similar, e.g., black bag on a black background.

To handle the full or partial occlusion, reduction in false positive rate, improvement in the detection rate of abandoned, removed or stolen object in the crowded scene of complex video; many researchers have proposed non-tracking based approaches such as motion statistics, n number of mixture of Gaussian models, color richness methods, finite state machines, features such as area ratio, length ratio, centroids change etc.

Porikli [87] proposed a non-tracking based method that employed the mixture of Gaussian models and online update using a Bayesian update mechanism. Short-term and long-term background modeling is used to extract foreground frames. It compares the foreground maps, and updates a motion image that keeps the motion statistics. Lavee et al. [57] introduced a framework to analyze a video for suspicious event detection to reduce the demanding task of manually sorting through hours of surveillance video sequentially to determine if suspicious activity has occurred. Lavee et al. [58] introduced a framework for detecting the suspicious event from video through low-level feature extraction, event classification, and event analysis. Ferrando et al. [30] presented a multi-level object classification technique to classify stolen objects and abandoned objects. The background update was performed on low level while the feature extraction was performed on the middle level. Stolen and abandoned objects were categorized using the Bhattacharya coefficients. However, a ghost was not considered. Cheng et al. [13] developed a multi-scale parametric background modeling to extract static

foreground objects from the dynamic scene by adapting the scene change at the multiple temporal scales. However, in this method, background can absorb to the foreground object when the background training is not sufficient. Moreover, a high computational cost is required by the multi-scale multiple Gaussian mixture background modeling.

Porikli et al. [89] developed a pixel-wise method that employs dual foregrounds to find stopped objects, abandoned objects, and illegally parked vehicles. This method adapts dual background modeling using Bayesian update and a temporal consistency is achieved by evidence obtained from dual foregrounds. This has one shortcoming i.e. it also detects the static person as an abandoned object. Li et al. [60] presented a robust real-time system that employs a pixel wise static object detector and a classifier based on the color richness for detecting abandoned and removed objects. Li et al. [61] developed a video surveillance system which generated two foreground masks with the use of short-term and long-term Gaussian mixture models and also utilized Radial Reach Filter for refining foreground masks to reduce the influence of illumination changes. This system fails to detect the occluded object by a person in S4 video sequence of PETS 2006.

Caballero et al. [29] presented a work to monitor human activity by global and local finite state machines. Beyan et al. [5] fused images from optical and thermal cameras to filter out living objects and reduce the false alarm rate.

Sanmiguel and MartiNez [94] proposed a framework for single-view video event recognition with the help of hierarchical event descriptions, which takes advantages of the accuracy of probabilistic approaches as well as the descriptive capabilities of semantic-based approaches. Zin et al. [124] proposed a probability-based background subtraction algorithm based on combination of multiple background models for motion detection by removing quick lighting change adaptation and shadows. This rule based method works well in crowded environments and also can detect the very small abandoned objects in low quality videos. Sanmiguel et al. [95] developed a novel approach to discriminate between abandoned and stolen objects based on color contrast along contour of the object at pixel level from video surveillance.

Maddalena et al. [70] proposed a neural based framework for the static objects and moving object detection. Experiment was performed on i-LIDS dataset in which proposed approach detected all static objects truly excepting AB-hard video sequence. Being the good accuracy of this method, this approach is not suitable for online analysis due to having long computation time.

Tripathi et al. [108] proposed a method to detect an abandoned object from video surveillance. Contour features are used on both the consecutive frames to detect static objects in the scene. Detected static non-human and human objects are classified by using edge based matching method which is capable to generate the score for partial or full visible object. Generated score is compared with a threshold to classify whether the object is human or nonhuman. This approach has accuracy 91.6%.

Zeng et al. [120] presented a Binocular Information Reconstruction and Recognition (BIRR) approach to detect the hazardous abandoned objects from the road traffic video surveillance. Proposed method employed a static region detection algorithm for a monocular camera. To detect hazardous object, 3D object information such as height of the object and road plane equation are collected after 3-D object information reconstruction.

Szwoch [103] recently proposed a novel approach to detect stationary objects from the video surveillance. In this, author tested the stability of pixels and then clusters of pixels with stable color and brightness are extracted from the image.

3.2 Research in abandoned or removed object detection from multi-static camera

Multi-view visual surveillance is another important and emerging research area in the field of computer vision. In complex public scenes, multiple-camera systems are more commonly deployed to handle the occlusion than single-camera systems. Abandoned or Stolen Object Detection in highly crowded area is a great challenge with the full or partial occlusion handling. Few researchers have worked for the development of an intelligent surveillance system that can monitor and identify the abandoned object using the information collected from multiple cameras. This section covers the progress in the field of multi-view cameras.

The problem becomes more complicated while the scene is observed through multiple cameras. To deal with this problem, several approaches have been proposed which can be classified in two categories: calibrated and un-calibrated [3]. Calibrated cameras greatly facilitate the fusion of visual information generated by multiple cameras. Khan and Shah [51] proposed an un-calibrated approach that utilized Edges of Field of View in which advantage of the lines delimiting the field of view of each camera has taken. Similarly, Calderara et al. [11] introduced the concept of Entry Edges of Field of View to deal with false correspondences. Yue et al. [119] presented a calibrated method in which homographies were employed to resolve the occlusion problems while Mittal and Davis [76] proposed a method based on epipolar lines. Khan and Shah [52] again used a planar homography constraint to resolve occlusions and detected the locations of people on the ground plane corresponding to their feet.

Auvinet et al. [3] applied homographies and simple heuristic approaches to detect the left luggage. Rincon et al. [72] employed a modified Unscent Kalman Filter (UKF) algorithm for multi camera visual surveillance. Krahnstoever et al. [56] presented a multi view visual surveillance to detect unattended and abandoned objects. This approach tracks all the objects as well as the people in the scene and creates the spatiotemporal relationship between them. This system applied centralized tracking in a calibrated metric world coordinate system, which constrains the tracker and allows for accurate metrology regarding the distances between them and size of targets. Performance is evaluated on the PETS 2006 video data; the system detects all abandoned luggage events with only one false alarm.

Arsic et al. [2] employed a multi-camera tracking algorithm and analysis module. First of all, proposed method applied adaptive foreground segmentation based on Gaussian mixtures proposed by Stauffer and Grimson. Secondly, proposed method applied the heuristic approach to remain in real-time in all computation steps excepting high computation cost tracking methods such as UKF [111] and Condensation [46]. Homographic transformation is employed to fuse the information of the each camera view. Porikli et al. [88] proposed a pixel-based method that employs dual foregrounds for multi camera setups. This approach computed the homography matrices using multiple pairs of corresponding points, which are selected manually. This approach has one shortcoming, that it cannot discriminate the different objects, e.g. stationary human being for long duration can be identified as left object.

Guler et al. [37] identified static objects and correlated with drop-off events; then the distance of the object from the owner is utilized to declare alerts for each camera view. Finally, abandoned object detection results are obtained by fusing the information from these detectors and over multiple cameras. The left behind objects are correlated among multiple camera views using the location information and a time and object location based voting scheme.

Most of the proposed techniques for abandoned or removed object detection rely on tracking information [3, 6, 38, 56, 72, 96, 99] to detect drop-off events, while fusing

information from multiple cameras. Porikli [87, 89] stated that these methods are not well suited in complex environments like scenes involving crowds and large amounts of occlusion. In addition, they require solving a difficult problem of object tracking and detection as an intermediate step.

3.3 Research in abandoned object detection from moving video camera

This section presents a single paper for abandoned object detection from moving cameras but progress is limited in this area due to the unavailability of background subtraction method for moving cameras.

Kong et al. [55] proposed a novel framework for the detection of non-abandoned objects from moving cameras. Reference video and target videos are matched to detect the abandoned object. In this, reference video is recorded from moving camera when there is no suspicious scene in video. The target video is recorded from a camera of the same route. This system detects 21 suspicious objects out of 23 from 15 video sequences. The system fails in one video having almost at object.

4 A general framework for abandoned or removed or stolen object detection from video surveillance

In this, we have presented general framework for abandoned or removed object detection from single static camera (shown in Fig. 4). Abandoned or removed object detection includes the following important stepladders: Foreground objects detection, tracking or non-tracking based object detection, feature extraction, classification; object analysis and alarm or alert message generation. Mostly researchers follow up these steps with different algorithms or approaches to improve the recognition accuracy. Abandoned object detection from multiple cameras also includes foreground object extraction, tracking, remapping methods-camera calibration methods [109] and homographic transformation in which points of the image plane is mapped to the ground plane [3] and information is fused; then object analysis is performed to detect the abandoned object.

4.1 Foreground object extraction

Foreground object extraction from the video is the first step to extract foreground objects and remove the background details. Background subtraction is a powerful mechanism to detect the change in the sequence of frames and to extract foreground objects [74]. Foreground objects means moving objects and newly arrived objects in a video which becomes static after some time such as left luggage. However, moving objects are considered as the foreground objects while static objects are absorbed in the background image of the video after a short duration in background subtraction techniques. This concept simplifies the moving object detection from a video of static camera but difficult to detect newly arrived or removed stationary objects.

Moving object detection can be performed based on two approaches- background modeling and change detection based approaches [77]. The change detection approaches subtract the consecutive frames to recover motion and use post processing methods to recover the complete object. These methods are fast in respect to execution while lacking in accuracy. Modeling-based approaches try to model the background using some temporal and/or spatial cues. A

reasonably correct model for the background can help to separate the foreground objects much effectively compared to the previous class of methods. These methods can range from very simple to highly complex in implementation and execution.

Newly arrived stationary objects in a video can be dangerous for the human and public place. To identify such dangerous stationary objects, several authors have worked with different methods and algorithms. First step to extract the foreground object through the background subtraction is not capable to detect the stationary objects from surveillance video.

4.1.1 Moving foreground object detection

In the last decade, several researchers have worked for the moving foreground object detection from the surveillance video. These methods help in extracting the human activities from the background of the surveillance videos. Piccardi [84] presented a review on seven background subtraction methods. Wren et al. [114] introduced an independently background modeling method at each pixel location using a single Gaussian. It has low memory requirements. Lo and Velastin [65] presented temporal median filter background technique. Cucchiara et al. [19] proposed an approach based on mediod filtering, in which mediod of the pixels are computed from the buffer of image frames. Stauffer and Grimson [100] proposed a most common background model based on Mixture of Gaussians. This method handles multi-modal distributions using a mixture of various Gaussians. This approach cannot model accurately to the background having fast variations with the few (3-5) Gaussians. To solve the previous problem, Elgammal et al. [21] developed a non-parametric model to model a background which is based on Kemel Density Estimation (KDE) on the buffer of the last n background values. KDE guarantees a smoothed, continuous version of the histogram. Bouwmans [9] provided a complete survey of traditional and recent background modeling techniques to detect the foreground objects from the static cameras video. Background subtraction is a very common technique for the segmentation of foreground objects in video sequences captured by a static camera, which basically detects the moving objects from the difference between the current frame and a background model. In order to accomplish good segmentation results, the background model must be regularly updated so as to adapt to stationary changes in the scene and to the varying lighting conditions. Therefore, background subtraction techniques often do not suffice for the detection of stationary objects and are thus supplemented by an additional approach [23].

4.1.2 Stationary foreground object detection

In video surveillance, moving objects are detected easily by using several background techniques because these techniques consider only moving objects as a foreground object. Therefore, whenever a new object arrives in a video and become static; after some time, it is absorbed in the background. To detect the stationary object from surveillance video, basically different approaches have been applied to extract the static object.

Dual background approach with dual learning rate In first method, several authors used different background subtraction techniques with dual background approach with different learning rate to extract the two foreground objects for detecting the stationary objects of the video. Porikli et al. [89] proposed a system which uses dual foreground extraction from dual background modeling techniques. Therefore, a short-term and long-term background models

are created with different learning rates. Through this way, authors were able to control fastest static objects amalgamation by the background models and detect those groups of pixels classified as background by the short-term but not by the long-term background model. Dual background modeling technique has been used to detect the abandoned object by several researchers in [4, 23, 60, 61, 89, 93]. A weakness of this method is the high false alarm rate, which is typically caused by imperfect background subtraction resulting from a ghost effect stationary people, and crowded scenes. Sometimes, temporarily static objects may also get absorbed by the long-term background model after a given time based on its learning. The dual backgrounds are constructed by sampling the input video at different frame rates i.e. short-term and long-term events. This technique, however, is difficult to set appropriate parameters to sample the input video for different applications, and has no mechanism to decide whether a persistent foreground blob corresponds to an abandoned object or a removed object.

Temporal dual rate background technique Lin et al. [64] introduced a temporal dual rate background technique to estimate the static foreground object which works better than the single image based dual background models in [24, 72, 89].

Mixture of Gaussian models This approach has been applied based on n number of Mixture of Gaussian models. Mainly three Gaussian Mixture Models have been used to detect the static foreground object, moving foreground object and removed foreground object in several researches [26, 27, 105]. Tian et al. [106] used three Gaussian Mixtures of Background Model in which 1st Gaussian distribution models the persistent pixels and represents to the background pixels, static regions are updated to the 2nd Gaussian distribution and 3rd Gaussian distribution represents to the quick changing pixels. In [60], static object is detected through the pixel wise method.

Accumulating binary foreground images or tracking foreground regions to identify a static foreground Pan et al. [80] and Liao et al. [62] proposed methods to extract static foreground object based on pixels with maximum accumulated values. Liao et al. [62] extracted six foreground frames and found the intersection of these frames to find stationary object in video scene but this method fails in complex video scenes. Several authors have employed tracking approaches to track the blobs and to identify blobs as a static object based on its shape, size and position features. But, tracking technique fails in the crowded area, occlusion cases and complex environment. Tracking based approach in [116], generates the blob trajectory after the splitting of the blobs to detect the moving and stationary object in videos. Miguel et al. [75] performed stationary object detection through the analysis of trajectory graph for each blob.

4.1.3 Noise removal, shadow removal and illumination handling methods to reduce false detections

To detect the foreground objects without noise, illumination effect, and shadow is very challenging in the area of computer vision. Noise creates the problem in the identification of the object, illumination effect causes the higher false detection rate, shadow changes the appearance of the object due to that object tracking becomes very complex.

Several researchers have utilized different methods to remove the illumination effects, noise, and shadow from the video to minimize the false detections. In the abandoned object detection approaches [7, 60, 75, 90, 94], researchers used morphological operations to remove the noise from the foreground frames; Tian et al. [105] used texture information to reduce the false positives, and normalized cross correlation to remove false detection due to shadow. In [61], Radial Reach Filter has been used to reduce the false detected foreground due to the illumination changes and Gaussian smoothing to remove the small holes. In [26, 106], Phong Shading Model has been used to handle quick light changes. In [4, 8, 28, 116], 2D-convolution and color normalization to enhance the image, structure noise reduction algorithm to remove the noise, Mahalanobis distance between the source and background model pixels to handle multimodal backgrounds with moving objects and illumination changes, Gaussian blur to reduce the noise have been used respectively. In [81], Gaussian filtering, with color and gamma correction is used to remove the noise. In theft detection method [16], object size has been considered more than 50 pixels to filter out noisy regions. Fuzzy color histogram (FCH) has been used in [15] to deal with the color similarity problem. Nam [78] applied contour based approach for ghost removal. Several researchers have made a great effort to reduce the false detection rate due to the noise, shadow, and illumination effects in complex video scenes but false detection is being made through the intelligent video surveillance. Lin et al. [64] used Extended Gaussian Mixture Models instead of Codebook method because codebook generates foreground regions as a many fragmented region that causes to fail to infer the static foreground pixels. Mahin et al. [71] used morphological closing operation to fill the holes.

4.2 Localization of a static object based on tracking approaches

Object tracking is an important and challenging chore in the field of computer vision. It helps in generating the trajectory of an object over time with the tracing its position in consecutive frames of surveillance video to analyze the human behavior. Object shape representations employed for tracking are points, object contour, object silhouette, primitive geometric shapes, articulated shapes and skeletal models [118]. Sometimes, tracking of an object becomes difficult due to noise in the image, partial or full occlusion of objects, complex object shapes, illumination changes, complex object motion, and deformable objects. According to Yilmaz [118], there are three tracking categories- kernel tracking, point tracking, and silhouette tracking. Kalman filtering [50] is one of the well known and widely used methods for object tracking with its ease of use and real-time operation capability. Kalman filter assumes that the tracked object moves based on a linear dynamic system with Gaussian noise. For non-linear systems, methods based on Kalman filter are proposed, such as Extended Kalman Filter, and Unscented Kalman Filter. Kalman filter with a dynamics model of second order derivative has been used in [41]. To detect the start and the end of possible snatching events, Kalman filter has been used in [11]. Kalman filter is used when the movement is linear and to overcome this problem particle filter [54] focuses on both nonlinear and non Gaussian signals. Particle filters are an alternative to the Kalman filters due to their excellent performance in very difficult problems including signal processing, communications, navigation, and computer vision. Particle filters recently became popular in computer vision that are especially used for object detection and tracking. In the area of abandoned and removed or stolen object detection, tracking has been employed in four different ways by the eminent researchers of this field. Mostly researchers applied different tracking approaches to track the object in the video to find the position of object in each sequence of frames to decide the staticness, few of them applied tracking approach as a complement tracking to reduce the false detection, and a less number of researchers employed this tracking approach to track the owner backward.

4.2.1 Object tracking for deciding staticness

Chuang et al. [15, 16] proposed a kernel based tracking approach to track the foreground objects to analyze their transferring condition between carried object and its owner. In [35], particle filtering and blob matching techniques are used to track the objects. Miguel et al. [75] employed an object tracker to create a trajectory graph for each blob based on color, size and distance rules to identify the stationary object in the video. Yang et al. [116] employed blob tracking by using center of mass point to track the baggage and owner. Prabhakar et al. [90] applied a tracking approach to plot the trajectory of the object using centroids, height and width to find the location of object. Chitra et al. [14] also employed a trajectory system to track abandoned objects. Pavitradevi et al. [81] employed color histogram to track the colored object and a blob tracking is employed in [8]. Lu et al. [68] employed Kalman based filtering and tracked the moving objects using color and shape features. Denman et al. [20] applied condensation filter to track the people with optical flow and color. Lv et al. [69] combined a Kalman filter-based blob tracker with a shape-based human tracker to detect objects and people in motion. Spengler et al. [99] employed a Bayesian multi-people tracking module to track arbitrary number of objects. There are several works have been employed tracking approach for object or human tracking [4, 8, 41, 93, 104, 123].

4.2.2 Human tracking for owner identification

Bhargava et al. [7] performed backtracking of the owner after finding the stationary object to decide that object is abandoned or not. Liao et al. [62] used tracking algorithm based on color and human body contour to detect the owner of the abandoned object in the video. Recently, Lin et al. [63, 64] detected the static object by using pixel based finite state machine and after that back tracked to the owner to analyze the abandonment of the object. The back tracing of owner with pixels based finite state machine worked well in comparison to the pixel based FSM only.

4.2.3 Object trajectories as complement information to reduce false positives

Nam [78] employed the tracking approach only for using trajectory information as a complementary to reduce the false positives in complex video scenes. Tian et al. [105] also used tracking information as an additional cue to reduce the false detections. Sajith et al. [93] utilized tracking to reduce the false positives.

4.2.4 Interactions between background subtraction and tracking to handle slow-moving and stopped objects

Tian et al. [106] created a feedback mechanism that creates interaction between background subtraction and tracking that handles the slow moving and stopped object problem.

Most of the proposed algorithms for abandoned object detection are dependent on tracking information. These methods do not work in complex environments like scenes involving crowds and large amounts of occlusion. Several researchers have not employed the tracking based abandoned object detection due to the occlusion, complex object shapes, deformable objects and a fixed camera angle which cause erroneous tracking. Table 2 shows tracking based approaches used for object tracking to identify static objects but occlusion cannot be

handled. In spatial rule, owner is traced to check his/her presence; if owner is present in the scene of the video then static object is not considered as an abandoned object. But, in case of owner is occluded then this system will generate false alarm. Lately, few researchers started to use tracking as complement information to reduce the false positive rate and also utilized to handle the slow moving and stopped object problem but sometimes system fails when object become stationary for long duration. While tracking is being used as complement information that improves the performance of the visual surveillance system by reducing false positive rates but could not reduced false positive rate up to 0% label.

4.3 Localization of a static object based on non- tracking based approaches

The use of tracking information could not handled the occlusion, low contrast situations in a very complex video sequences. Many researchers have employed non-tracking based approaches. Selecting appropriate features plays a vital role in the recognition of abandoned or removed object from intelligent video surveillance. The goal of feature extraction is to find the most important information in the recorded video. To detect the stationary objects in a video by handling occlusion, illumination effects is very critical job. Therefore, some features of objects are extracted from video to make distinction between moving and stationary objects.

4.3.1 Dual foreground with different learning rate

Porikli et al. [87, 89] employs dual foreground with two different long-term and short-term learning rates. With these two different learning rates, two foreground masks F_L and F_S are created. If $(F_L, F_S) = (1, 0)$, then object is static. Zeng et al. [120] used dual background model based on long term and short term learning rate. In this process, region A is generated by finding the XOR operation between short term background and long term background model and region B is generated by finding the difference between short term background and long term background model. To find the static foreground object, AND operation is performed in between region A and region B. In [61], Li et al. employed two Gaussian mixture models based on long and short term models to extract two foreground masks and then applied a mathematical evaluation based on pixel value. If mask of lower learning rate $F_L(x, y)$ is 1 and mask of higher learning rate $F_S(x, y)$ is 0 then *pixel(x, y)* is the part of temporary static objects. Evangelio et al. [23] also employed dual background model to find the static object with finite state machine.

4.3.2 Pixel-wise method

Li et al. [60] proposed a pixel wise method to detect static object by generating confidence score. Pan et al. [80] employed pixel foreground score method in which pixel of previous frame is compared with pixel of the next frame and difference is computed to compare with a threshold; if it is higher than the threshold and pixel value is unchanged in both the frames, then connected component analysis is performed to generate static foreground blob. Kim et al. [53] applied region layer features to detect the static object when connected component groups foreground pixels into a region. Lin et al. [64] employed pixel based finite state machine which is found better than dual background subtraction method.

Table 2 Different tracking approaches employed for object tracking, owner tracking, tracking as complement information to reduce false positives

Purpose	Approach	References	Result Discussion with the use of trajectory of an object or owner
Object Tracking	Kernel based tracking	Chuang et al. [16], Khan and Shah [51]	Abandoned object detection rate-100% for own video dataset only.
	Particle filter and blob matching	Foucher et al. [35]	Large number of false alarms raised due to noisy tracker and track fragmented blobs.
	Trajectory graph of each blob based on color, size	Miguel et al. [75]	Detection precision for low complexity video is 97.7%, for medium complex video-93.58% and for high complex video is 76.4%.
	Blob tracking using center of mass point	Yang Rothkrantz [116]	The frame rate is 5 fps that affects to the accuracy of the tracking.
	Object tracking using center of mass, size and bounding box	Singh et al. [97].	Detection rate is over 93%, and the False Alarm Rate is under 3.6% in own video.
	Blob tracking based on height, width and centroids	Prabhakar and Ramasubramaniam [90]	Quantitative analysis is not done.
	Color object tracking based on color histogram	Pavithradevi et al. [81]	The future improvement is to improve the work of object track and improve all aspect of behavior identification in human crowd of public areas.
	Blob detection and tracking	Chitra et al. [14]	In PETS 2006 dataset, detection rate is 80% for highly complex videos, 90.3% for medium complex videos and 91.2% for low complex videos.
	Blob tracking	Bird et al. [8]	MTC8 video sequence is densely populated for that result was not good in comparison to others.
	Kalman based filtering and color, shape based tracking to moving objects	Lu et al. [68]	Quantitative analysis is not done.
	Condensation filter to track the people with optical flow and color	Denman et al. [20]	False positive for static motion is 1.18%, for shadow motion is 49.34%, for active motion is 2.33%.
	Kalman filter-based blob tracker	Lv et al. [69]	In the last sequence, out of 6 persons only 5 were detected and missed one person. Tracker missed one of them who are severely occluded by two people. The recognition algorithm of events detects all alarms and warning within an error of 9 frames (0.36 s).
	Bayesian multi-people tracking	Spengler& Schiele et al. [99]	The people tracker is able to detect and to track an arbitrary number of objects in parallel.
	Blob tracking to find static object	Joglekar et al. [49]	Quantitative analysis not discussed.
	Object tracking	Ferryman [31]	Quick lighting changes are handled with Phong shading model.

Table 2 (continued)

Purpose	Approach	References	Result Discussion with the use of trajectory of an object or owner
Owner Tracking	Backtracking	Bhargava et al. [7]	In Test1 i-LIDS, the owners entry with the luggage and the departure without the luggage are not viewable. In Test5 i-LIDS, the foreground blob of the luggage is merged with the person sitting behind it.
	Tracking based on color and human contour matching	Liao et al. [62]	In PETS 2006 video sequences 1, 2, 4, 5, 6; the luggage owner left the scene without the luggage, and successfully an alarm was raised in all these 5 cases. In video 3, owner is very close to the luggage all the time, therefore no alarm is raised.
	Spatial-temporal tracking for back- tracing the owner	Lin et al. [64]	Worked well with PETS2006 and AVSS2007 datasets. In ABODA dataset video sequence 5, 7, 8 and 11(crowded), there are 1, 1, 1 and 3 false positives respectively.
	Temporal graph is extracted by using object tracking layer	Foresti et al. [33]	Laboratory dataset- false alarm is 0.1% and miss-detection 2.0%. Genova-Rivarolo Railway station-false alarm is 1.8% and miss-detection is 3.5%. Genova-Borzoli Railway station-false alarm is 0.5% and miss-detection is 2.5%.
Object Tracking as complement to reduce false positives	Complement tracking	Nam [78]	For i-LIDS dataset, precision is 98.8%, recall is 82.28% and F-measure is 82.64%. This approach generates fewest false negative and false positives, but also leads to 18.42 and 22.71% improvement on precision and recall against Porikli2008.
	Complement tracking	Tian et al. [105]	Complement trajectory reduced approx. 25% false alarm rate i.e. reduction from 44.5 to 20.7%. When parked vehicle on road is considered as abandoned in the night time then false positive rate without tracking is 23 while with trajectory verification, false positive is only 4. In removed object detection, there are 11 true positives out of 12. Detection rate is 87.8% and average false-alarm rate is about 2.5 false alarms / minute.
	Complement tracking	Fan et al. [27]	In PETS 2006, precision and recall is 0.95 and 0.80 respectively while in AVSS-AB, precision and recall is 0.97 and 1.0 respectively.
Object Tracking to handle slow motion and stopped problem	–	Tian et al. [106]	To handle slow moving objects, smaller learning rate was used. However, this approach will fail when object remain static for long

Table 2   (continued)

Purpose	Approach	References	Result Discussion with the use of trajectory of an object or owner
			duration. It may also fail in low contrast situation where the color of the object is very similar to the background, e.g., black bag on a black background.

4.3.3 Mixture of Gaussian with k-distributions

In [106], Tian et al. used three Gaussian Mixtures of Background Model in which 1st Gaussian distribution models the persistent pixels and represents to the background pixels, static regions are updated to the 2nd Gaussian distribution and 3rd Gaussian distribution represents to the quick changing pixels. Tian et al. [105] employed three Gaussian Mixtures to model the background to detect the static region without extra computational cost. Generally, the first Gaussian distribution shows the persistent pixels and represents the background image. The repetitive variations and the relative static regions are updated to the second Gaussian distribution. The third Gaussian represents the pixels with quick changes. Femi et al. [28] utilized classic multivariate Gaussian Mixture model in which every pixel is represented by four Gaussian distribution. The pixels at each frame are classified as foreground or background by calculating the Mahalanobis distance between the source and background model pixels, and comparing this distance to a threshold. Fan et al. [26] also utilized the three Gaussian Mixtures of background model to detect static objects in [106].

4.3.4 Centroid, height and width of an object

Centroid is defined as an average of the pixels in x and y coordinates belonging to the object. If objects centroid, height and width are same in each frame, then object is found as static. These features have been used in [90].

4.3.5 Ratio histogram

Chuang et al. [15] proposed a ratio histogram based on fuzzy c-means algorithm for finding suspicious objects. In [16], novel ratio histogram has been used for finding missing colors between two pedestrians if they have interactions.

4.3.6 Probability based background modeling to detect static foreground object

Zin et al. [123, 124] developed a new background method based on probability which consists of three backgrounds. In this, first frame of a video is initialized as frequently updating background and occasionally updated background and an improved adaptive background updating method is applied by constructing two maps of pixel history. First map is a stable history map that shows the number of times a pixel is static in consecutive frames. Second map is the difference history map that represents the number of times a pixel is different from the

background in consecutive frames. Based on these two history maps, probability-based foreground (PF) is constructed for detection of stationary object.

4.3.7 Velocity of object

In [124], static and moving object are categorized on the basis of their velocity which is compared with a predefined threshold.

4.4 Classification approaches to distinguish static object into human and non-human object

Classification method plays an important role to minimize the number of false alarms in abandoned, removed or stolen object detection through an Intelligent Video Surveillance. After finding the static objects in a video, classification approach must be robust to discriminate the human and non-human stationary object so that analysis can be performed to decide whether the object is abandoned, removed or stolen. For example, a stationary human and non-human object at public place can be treated as abandoned object if there is no knowledge of the object features. Object classification methods must be highly sensitive to distinguish the static human from static abandoned object, face from skin color objects etc. If applied classifier fails to discriminate between static human and non-human objects then false detection rate increases rapidly. To classify a foreground object correctly that whether the object is removed or abandoned objects; is an important problem in background modeling, but most of the existing systems neglect it.

In general, classification methods can be categorized into three categories which are based on shape, motion, and feature. Most of the researchers have tried their best efforts to extract and combine features with different classifiers such as SVM, Multi-SVM, k-Nearest Neighbor, Cascade classifier, Neural Network, and HAR to categorize human and non-human stationary objects to make zero false positives but they could only reduced false positives up to a less number of counting.

Yang et al. [116] categorized the luggage on the basis of skin color because skin color method works well and computationally least expensive. A non-moving object without skin color has been considered as a luggage. Miguel et al. [75] applied a people detection module to identify the stationary non-human object. Bhargava et al. [7] employed k-nearest neighbor classifier to classify the foreground objects into baggage and non-baggage class where bag was considered as a solid object equivalent to the half the height of an adult person. Classification features were based on the size and shape of binary foreground blobs but bag handle created a special problem by distorting the generic shape of the luggage. For the better utilization of classifier, morphological open operations were used on the binary foreground image using cross-shaped structural elements to remove the bag handle. Prabhakar et al. [90] used blob features to assign class label and object features such as center of mass, size, and bounding box has been used to estimate a matching between objects in consecutive frames in [97]. Chitra et al. [14] used Histogram of Oriented Gradients Descriptors (HOG) and Linear Support Vector Machine (SVM) as classifier to detect the pedestrian. Pavitradevi et al. [81] employed an adjacency matrix based clustering and SVM to identify actions of people with the features extracted from frames using Histogram of gradient and Gabor algorithm. Lin et al. [64] employed a temporal consistency model with Finite State Machine to categorize static non-human object.

Li et al. [60] developed a new color richness method to discriminate the abandoned or removed object in which color is compared in between background and current frame. Li et al. [61] used a height-width ratio and a linear SVM classifier based on HOG descriptor to distinguish the non-human object. Fan et al. [26] extracted 23 features based on the structure similarity, region growing, local ternary patterns and phong shading model to train a binary classifier libSVM for foreground analysis to detect the abandoned object. Zin et al. [124] estimated the velocity to decide moving or static object and to discriminate the abandoned object and still person by using Exponent Running Average [124]. Zin et al. [123] also employed a simple rule based algorithm to distinguish between abandoned or still person. Porikli et al. [87] did not use any person classifier due to that this system is incapable to discriminate the different types of objects, e.g. a person who is static for a long duration can be detected as an abandoned object. Evangelio et al. [23] presented the use of finite state machine to classify the pixels and Collazos et al. [17] also used finite state machine for the pixel classification. Kim et al. [53] employed event-layer FSM that uses color and edge features to decide static regions as either abandoned or removed objects.

Recently, Szwoch [103] presented a classifier dependent on shape descriptors of the object contour. This work computed a vector of seven Hu moments, invariant to scale, rotation and translation for each identified stationary objects by using methods of [42], to form the final object shape descriptor. This approach applied two classifiers: the first one utilized the Support Vector Machine method [18] and the second one used the Random Forests approach [10]. The classifier has been trained using a set of shape descriptor vectors computed from the collected examples which belongs to the luggage and other object classes. Tables 3 and 4 shows several classification methods with their results and shortcomings.

4.5 Object analysis to recognize abandoned, removed or stolen object

Object analysis and decision making is a very important and challenging step for an intelligent video surveillance system to decide the abandoned, removed or stolen object correctly and raising a true alarm on real time to alert the security for the removal of abandoned object that can cause ecological and economical losses and also prevents the stolen cases at the public places of the world. To improve true positive rate and to decrease false positive rate, several eminent researchers have employed the various analysis approaches such as- Finite State Machines (FSM) [16], fusion of high gradient, low gradient and color histogram features [75], multiple spatial-temporal and contextual cues to detect a given sequence of events [7], Bayesian inference framework for event analysis [97], high level reasoning to infer the existence of abandoned luggage [116], temporal analysis [14], probabilistic event model [62], Spatial-temporal rules to backtrack owner of the luggage to identify the abandoned luggage [64], Region level analysis [80] in which after static object detection, foregroundness score is computed for each static region and if foregroundness score is found less than 0.5 then object is not considered as abandoned.

Table 3 shows the performance and shortcomings of several tracking based approaches and Table 4 shows the performance and shortcomings of non-tracking based approaches. Lin et al. [63, 64] used spatial-temporal analysis with pixel based finite state machine which worked better than [24, 27, 62, 89, 105] etc. Tian et al. [105, 106] applied edge energy and region growing method and three classifiers respectively but failed to handle low contrast problem. The Table 3 and Table 4 clearly show that there is no perfect approach to handle all the situations; therefore, improvements are required in this field to develop an intelligent visual surveillance.

Table 3 Several tracking based approaches with their classification approach, analysis approach, result discussion and shortcomings or future works

Works	Classification Approach	Object Analysis Approach	Results	Shortcoming/ future works
Nam [78]	Person, abandoned and stolen vector Matching set	Spatial temporal features based analysis	AVSS2007 Precision-98.88% Recall-82.28 F--measure-82.64%	A person is sitting on a bench is also detected as an abandoned object. People sitting or standing in static position for a long duration causes false positive detection.
Mahin [71]	Area and centroids statistics data	Blob Analysis	Able to detect objects abandoned or removed in indoor environments	No result on standard dataset
Lin [64]	back-tracing algorithm tracks the luggage owner	Spatial rule and temporal rule	PETS2006 P-1.0 R-1.0 AVSS2007 P-1.0 R-1.0	False positive in night video5 ABODA and one-one false positive in video6 and video 7 due to light switching and 3 false positives in video11 due to crowded scene.
Foggia [32]	Heat-map	Spatial temporal information	In CAVIAR dataset, abandoned object has been detected in 4 out of five where as in i-LIDS all abandoned object has been detected in AB-EASY and AB-MEDIUM video.	In CAVIAR dataset left bag behind chair was not detected and in i-LIDS AB-HARD video, one false object detected.
Chitra et al. [14]	HOG & SVM	Position of object is same for long time	Detection rate with high occlusion- 80%, Medium occlusion −90.3% and Low occlusion video-91.2%	–
Pavitrdevi et al. [81]	SVM	Adjacency matrix based clustering is used to identify the behavior	–	–
Lin et al. [63]	Pixel-based finite state machine	Spatial temporal rule	AVSS2007-Precision-1.0 Recall-1.0	Reliable tracking can be achieved for the better identification of the left-luggage by integrating an appearance based pedestrian detectors into space-time

Table 3 (continued)

Works	Classification Approach	Object Analysis Approach	Results	Shortcoming/ future works
				search-windows framework
Fan et al. [27]	High level attributes ranking (HL- RANK)	Low level temporal and spatial features	PETS 2006: Precision-0.95 Recall-0.80 AVSS 2007: Precision- 0.97 Recall-1.0	One baggage was not detected by HL-RANK in PETS2006. Missed bag is skinny and long, and similar to a person.
Sajith et al. [93]	Neural Network classifier is trained by the feature extracted from Histogram of Oriented Gradients.	Minimum and maximum object size, temporal analysis of static object	In PETS 2007, all abandoned object was detected in all four videos of S8 sequence. In PETS 2006, there are four static object and four static persons in all four videos of S3 sequence.	A stationary or motionless person was detected as an abandoned object in sequence S3 of PETS 2006 dataset.
Tian et al. [106]	Edge energy methods and region growing methods	Frame level analysis	False negative was reduced 39% while false positive was increased 2.7%.	This system fails in low contrast situations where the object and background color is very similar, e.g., black bag on a black background. Also fail to resolve occlusions correctly.
Prabhakar et al. [90]	Blob features	Classification and tracking correlation	Well worked on his own dataset	Performance was not evaluated on standard dataset.
Tian et al. [105]	Face recognition Through cascade classifier of near field people, head and shoulder detector for middle field people and 30 pixel length for far field people.	Object size and its time duration of staticness.	Detected all the abandoned objects with low- false positives rate. Some static people were identified as an abandoned object because a person classifier was not incorporated in the system. Out of total 12 removed objects, 11 were detected. Also there was a huge reduction in false-alarm rate i.e. from 44.5 to 20.7%, approx. 25% with the help of tracking.	A motionless person is recognized as an abandoned object in the video sequence S3. This problem can be resolved by incorporating a person classifier. Low contrast situations like black bag on a black background, causes missed detections.
	Skin color approach	Vertical and horizontal	Not discussed	5 fps frame rate is a good trade-off

Table 3 (continued)

Works	Classification Approach	Object Analysis Approach	Results	Shortcoming/ future works
Yang et al. [116]	worked well and motion-less split blobs without color of skin are considered as luggage.	movement, variance, mean size, and distance to nearby objects.		between system performance and accuracy.
Singh et al. [97]	Object features like size, center of mass, and bounding box	Event is analyzed by using Bayesian Inference framework.	Detection rate- 95.78% and False alarm rate-3.74	Worked only with own dataset
Chuang et al. [124]	Missed color detection with Ratio Histogram and carried object detection using GMM	Finite state machine	Abandoned object detection rate-100% on own dataset. Robbery-96%	Used 2^{12} color bins to detect the baggage by trading off between accuracy and efficiency.
Bhargavaet al. [7]	k-Nearest Neighbor (k-NN) classifier	Multiple temporal and spatial and contextual cues	Almost detects all abandoned objects of PETS2006 and i-LIDS dataset.	k- NN classifier fails to detect in Test5 i-LIDS where luggage and the person sitting close; and both are blended into the same foreground blob.
Miguel et al. [75]	Trajectory of each blob with people detection module	Fusion of three features-shape, contour and color similarity	After fusion of three features: Detection rate with low, medium and high video complexity is-99.7%, 93.58%, and 76.4% respectively.	Detection rate is low for high complexity videos.
Liao et al. [62]	Skin color for face localization and head shoulder contour	Probability score for abandonment analysis using Posteriori probability under Maximum a posteriori principle (MAP)	In video sequences 1, 2, 4, 5, 6 of PETS 2006, the owner of luggage goes outside of the scene without the luggage, system raised an alarm in all 5 video sequences successfully.	In medium and hard sequences of AVSS 2007, tracking of owner is difficult due to occlusion.
Bird et al. [8]	Blob splits, blob creation, blob merges, blob centroids, velocity features are utilized to classify by short-term logic	Percent Alarm True (PAT) scores and Per-cent Events Detected (PED) score	Three sequences having sparsely populated scenes i.e. BSHigh has good PAT and PET score, BSLow has good PED score but low PAT score due to segmentation failures caused by large potted bus in the scene, and MTC5 has also a good PED score but poor PAT score. All the three datasets	Densely populated video sequence MTC8 does not yield good result.

Table 3 (continued)

Works	Classification Approach	Object Analysis Approach	Results	Shortcoming/ future works
			have overall scores around 70%.	

5 Evaluation measures and data sets

5.1 Evaluation measures

Evaluation of the performance of an Intelligent Visual Surveillance for abandoned or removed object detection on real-time is one of the major tasks to validate the correctness and robustness. The evaluation of different abandoned or removed object detection systems have been performed in two ways; firstly qualitatively and secondly quantitatively. Qualitative evaluation approaches are performed on visual interpretation, by looking at processed image yield by an algorithm. It consists of several issues and challenges handling algorithms. Noise removal, shadow removal, illumination handling, poor resolution handling, partial or full occlusion handling, etc. improves the qualitative performance of a visual surveillance system. On the other hand, quantitative progress requires a numeric comparison of computed results with ground truth data. Due to the necessity of computing a valid ground truth data, the quantitative evaluation of an intelligent visual surveillance systems are highly challenging.

There are a number of metrics proposed in literature to quantitatively evaluate the performance of an intelligent visual surveillance system that has been discussed below.

5.1.1 Recognition accuracy

Most of the research work in abandoned or removed object detection used accuracy for measurement of evaluation. It is defined as follows:

$$Accuracy(\%) = \left(\frac{T_P + T_N}{T_P + T_N + F_P + F_N} \right) \quad (1)$$

False alarm rate determines the degree to which falsely detected objects (F_P) dominate true objects (T_P)

$$False \quad alarm \quad rate = \left(\frac{F_P}{T_P + F_P} \right) \quad (2)$$

A true positive (T_P) represents abandoned or removed object is classified as abandoned or removed by the classifier; a false negative (F_N) represents to the missed detection; a false positive (F_P) corresponds to the classification of non-abandoned object as abandoned and a true negative (T_N) stands for abandoned object classified as non-abandoned. In [26, 28, 60, 105, 106], researchers evaluated the performance by True Positive, and False Positive detection.

Table 4 Several non-tracking based approaches with their classification approach, analysis approach, result discussion and shortcomings or future works

Works	Classification Approach	Object Analysis Approach	Results	Shortcoming/ future works
Szwoch [103]	Shape descriptors of the object contour is formed by finding 7 Hu-moments then Support Vector Machine [18], and Random Forests approach [10] is pre-trained by the shape descriptor vectors.	Stationary object is unattended and there is no people around it up to a defined distance.	i-LIDS (AVSS 2007) Dataset: Precision-100% Recall-100% PETS 2006: Precision-100% Recall 85.7% Robust against short-term occlusions.	This system yields a single false negative result due to similar color object and background, so it was partially camouflaged.
Zeng et al. [120]	Binocular information reconstruction and recognition (BIRR) algorithm.	Equation of road plane and height of the abandoned objects are computed based on the 3-D information	Average detection rate-92.3% Segmentation speed = 3.108 s	To calculate the real height, the average relative error is approx. 3.38% and maximum error is lesser than 6%.
Yadav et al. [115]	Neural network	Position of pixels of an object remains unchanged in Consecutive frames.	Detection accuracy-90.90% In case of low quality sequences SOBS achieves better results than Background Segmentation and detects tiny objects also in SOBS.	Static object detection by using self organizing background sub-traction has high computational complexity.
Kim et al. [53]	SVM event-layer Finite State Machine uses edge and color features	Hierarchical FSM state transitions	Classified all static regions correctly without any false static region.	This approach can reject false static objects effectively.
Collazos et al. [17]	Finite state machine	If the people in the scene are at distance from the object larger than a predefined threshold, then it is labeled as abandoned object	In all sequences, proposed method detected all abandoned objects available in videos but proposed approach detected five false positives also.	Not evaluated on standard datasets.
Femi et al. [28]	Static region is detected by using the split and merge algorithm	Time of the presence of static object	Detected all five abandoned object and all five stolen objects.	Not evaluated on any standard dataset.
Fan et al. [26]	Region Growing and Structure Similarity Methods. Binary Classifier LibSVM trained by 23 extracted features by using LTP (Local Ternary Patterns) .	Motion for short duration isanalyzed to decide the abandonment.	In i-LIDS, not a single false positive was generated. Reduced 6% false positives on AB-L1 and and 3% on AB-L2 with only small loss of accuracy (2%).	–
Zin et al. [124]	Rule based classifier	If Exponent Running Average is greater	Detects the small abandoned objects	Quantitative analysis is not presented.

Table 4 (continued)

Works	Classification Approach	Object Analysis Approach	Results	Shortcoming/ future works
		than a predefined Threshold then Static Object is Detected as Static Person Otherwise it will be Abandoned Object.	from low quality videos.	
Evangelio et al. [23]	Finite state machine	State transition analysis	Detected all abandoned objects correctly with 5 false detection in AB MEDIUM and 6 false detection in AB HARD.	Frame rate (fps) is 15.80 for AB EASY, 13.84 for AB MEDIUM, 13.91 for AB HARD, 15.48 for PETS2006.
Li et al. [61]	HOG features are calculated when height-width ratio satisfies a threshold to input to libSVM	Analysis based on SVM result.	Detected all the abandoned baggage in PETS 2006 and PETS 2007.	Fails in dataset S4 of PETS 2006.
Li et al. [60]	Color richness in RGB color space	Static object is detected as an Abandoned object if color richness of Current frame is bigger than background frame otherwise, it is detected as a removed object	Detected all abandoned object in video of CAVIAR dataset and i- LIDS datset. 2 false positives for abandoned object and 1 false positive for removed object detection were found in AB HARD dataset.	System fails to detect a bag in video Left-Bag BehindChair.mpg because the bag was left behind the chair in CAVIAR dataset. There is no classifier to distinguish different objects.
Pan et al. [80]	Foreground score is generated	Region level analysis	Detected all the abandoned objects in i-LIDS dataset.	For ABTEA101a clip, 10 alarms are generated for non-abandoned object.
Porikli et al. [89]	–	Spatial and temporal statistics	All Abandoned object, parked vehicle are detected successfully from i-LIDS, PETS2006 and ATC dataset.	Only 1-1 false alarms are generated in AB MEDIUM, AB HARD video of i-LIDS. One false alarm in ATC-4 and two false alarms in ATC-5 are also generated. It fails to discriminate the different types of objects, e.g. a person who is stationary for long duration can be detected as a left abandoned object.
Otoom [79]	Bayesian- based classifier, Decision trees, Sequential minimal	Analysis based on classifier results.	Classification accuracy is in between 72%-79.5% and	Future work is to improve the classification performance by

Table 4 (continued)

Works	Classification Approach	Object Analysis Approach	Results	Shortcoming/ future works
	optimization [113] and Multi-BoostAB [112] based on a set of statistics of geometric primitive features		false positive rate is 6.8%-9.3%	using feature reduction methods such as Kernel Principle Component Analysis.
Porikli [87]	–	Spatial and temporal statistics	Detected all abandoned objects with low false alarm rate	System fails to discriminate the different types of objects, e.g. a person who is stationary for long duration can be detected as a left abandoned object.

5.1.2 Recall, Precision, F-measure

In [26, 31, 75, 78], researchers have utilized all these parameters to evaluate the performance of abandoned or removed object detection systems where the Precision represent the percentage of true alarms and recall represents the percentage of detected events.

$$Recall(\%) = \left(\frac{T_P}{T_P + F_N} \right) \tag{3}$$

$$Precision(\%) = \left(\frac{T_P}{T_P + F_P} \right) \tag{4}$$

$$F-measure(\%) = \left(\frac{2 \times Precision \times Recall}{Precision + Recall} \right) \times 100 \tag{5}$$

5.1.3 Frames per second (fps)

Real-Time Intelligent Video Surveillance System must have a good execution speed for processing the video frames. Several researchers [23, 70, 92] have computed execution speed of their system to decide whether the system would work on real-time or not.

5.1.4 ROC Curve

In statistics or machine learning, a receiver operating characteristic (ROC) curve is a graphical plot that reveals the performance of a binary classifier. This curve is drawn by plotting the true positive rate against the false positive rate at various threshold settings. Many researchers employed ROC analysis of the performance of the different parameters.

5.2 Data sets

Data set is one of the most important components to evaluate the performance of any system. Evaluating the proposed algorithm against a standard dataset is one of the challenging tasks in visual surveillance system. In the recent years few standard datasets are available in the field of abandoned object detection.

5.2.1 PETS 2006 datasets

To evaluate the performance of visual surveillance system to detect an abandoned or removed object; a standard dataset is required. Therefore, PETS 2006 dataset was captured by four cameras that have seven different scenarios including small object to large object like human with low complexity view to high complexity view. This dataset consists of many video sequences of real scene captured with crowd, illumination effect, and luggage left. It can be downloaded from (http://www.cvg.rdg.ac.uk/ PETS2006/data.html).

5.2.2 Pets 2007

PETS 2007 dataset designed to test loitering, theft and abandoned object detection. There are 8 video sequences captured by four cameras from different viewpoints. Last two video sequences S7 and S8 were captured for abandoned or removed object detection. These video sequences are full of bad illumination and more lighting effects. It can be downloaded from (http:// www.cvg.rdg .ac.uk/ PETS2007/data.html).

5.2.3 I-LIDS-abandoned baggage detection

i-LIDS is Imagery Library for Intelligent Detection Systems. This dataset consists of unattended bags on the platform of an underground station. There are three videos which have been categorized on the basis of scene complexity. It can be downloaded from (http://www.eecs.qmul.ac.uk/andrea /avss2007 v.html).

5.2.4 Visor [110]

Video Surveillance Online Repository provides videos for different human actions such as abandoned object, drinking water, jumping, sitting, etc. Nine videos of abandoned object are available out of forty different human action videos.

5.2.5 Candela

This dataset consists of 16 examples of abandoned objects captured inside a building lobby. Videos are around 30 s long having 352×288 resolutions. This dataset is very simple having low resolution and small size of objects present challenges for foreground segmentation. It consists of two different scenarios- first is abandoned object and second is road intersection. In first scenario, abandoned object detection can be performed over a certain period of time. The duration of time is adjustable. In several types of scenes, stationary objects should be detected. In a parking lot, a stationary object can be either a parked car or a left suitcase. It can be downloaded from (http:// www.multitel.be/va/ candela/abandon.html).

5.2.6 Cvsg

Video sequences of this dataset have been recorded using chroma based techniques to extract foreground masks in simple way. Then, these masks are composed with different backgrounds. These video sequences have many varying degrees of difficulty in terms of foreground segmentation complexity. Also video sequences contain abandoned and removed objects in the scene. It can be downloaded from (http://www.vpu.eps.uam.es/ CVSG/).

5.2.7 Cantata

This dataset consists of a total of 31 sequences of 2 min have been recorded with two cameras. Few videos include a number of people leaving abandoned objects in the scene and other videos include people removes the same objects from the scene. Videos are at standard PAL resolution and MPEG compressed. It can be downloaded from (http:// www.multitel.be/va/ cantata/LeftObject/).

5.2.8 Caviar

CAVIAR dataset consists of a number of video clips which were recorded different activities such as walking people in different lane, leaving bags, fighting, etc. It can be downloaded from (http:// www.multitel.be/va/ cantata/LeftObject/).

5.2.9 Aboda

ABODA is an abbreviation of Abandoned Objects Dataset which has been designed for reliability evaluation. This dataset comprises 11 video sequences including various real-application scenarios that are challenge for abandoned-object detection. The video scene includes changes in lighting condition, crowded scenes, night-time detection, as well as indoor and outdoor environments. It can be downloaded from (http://imp.iis.sinica.edu.tw/ABODA /index.html).

5.2.10 Vdao

VDAO refers to Video Database of Abandoned objects in a cluttered Industrial Environment. This database consists of video for single objects, single objects with extra illuminations, multiple objects and multiple objects with multiple illumination. This dataset can be downloaded from (http://www02.smt.ufrj.br/ tvdigital/database/objects/page01:html):

6 Conclusions and future work

In this review, we have discussed the various core techniques related to abandoned or removed object detection i.e. the foreground segmentation, tracking and non-tracking based approaches, feature extraction, classification and analysis. In past decades, a lot of researchers presented novel approaches with noise removal, shadow removal, illumination handling, and occlusion handling methods to improve the detection accuracy and to reduce the false positive rate. Many researchers have worked for real-time based visual surveillance system and few of them achieved processing rate very close to real time for low and middle complexity videos while

processing rate of the highly complex video frames is not as good as required. But, there is no such system that has been developed with 100% detection accuracy and 0% false detection rate for any kind of videos having complex background.

Much of the attention is required to be paid in the abandoned or removed object detection in the following areas:

Abandoned or removed object detection from single static camera: Majority of the works have been proposed for the abandoned object detection from surveillance videos captured by static cameras. Few works detected the static humans as an abandoned object. To resolve such problems, owner of the static object should be identified and system should check the presence of the owner in the scene, if owner is invisible in the scene for long duration then alarm should be raised. To resolve the problem object removal, face of the person who is picking up the static object, should match with the owner otherwise an alarm must be raised to alert the security. Future work may also resolve the low contrast situation i.e. similar color problem such as black bag and black background which lead to miss detections. Future improvements may be integration of depth and intensity cues in the form of 3D aggregation of evidence and detailed occlusion analysis. Spatial-temporal features can be extended to 3-dimensional space for the improvement of abandoned object detection method for various complex environments. Thresholding based future works can improve the performance of the surveillance system by using adaptive or hysteresis thresholding approaches.

Abandoned or removed object detection from multiple- static camera: Few works have been also proposed for abandoned object detection from the multiple views captured by multiple cameras. To incorporate these multiple views to infer the information about abandoned object can also be improved.

Abandoned or removed object detection from moving camera: Background subtraction methods are not available for moving camera due to that abandoned object detection becomes difficult. There is a large scope to detect abandoned object from videos captured by moving cameras.

References

1. Aggarwal JK, Ryoo MS (2011) Human activity analysis: a review. ACM Comput Surveys (CSUR) 43(3): 16
2. Arsic D, Hofmann M, Schuller B, Rigoll G (2007) Multi-camera person tracking and left luggage detection applying homographic transformation. In: Proceeding Tenth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, PETS
3. Auvinet E, Grossmann E, Rougier C, Dahmane M, Meunier J (2006) Left-luggage detection using homographies and simple heuristics. In: Proc. 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (PETS06), pp. 51Citeseer
4. Bangare PS, Uke NJ, Bangare SL (2012) Implementation of abandoned object detection in real time environment. Int J Comput Appl 57(12)
5. Beyan C, Yigit A, Temizel A (2011) Fusion of thermal-and visible-band video for abandoned object detection. J Electron Imaging 20(3):033,001–033,001

6. Beynon MD, Van Hook DJ, Seibert M, Peacock A, Dudgeon D (2003) Detecting abandoned packages in a multi-camera video surveillance system. In: Advanced Video and Signal Based Surveillance, 2003. Proceedings. IEEE Conference on, 221–228. IEEE

7. Bhargava M, Chen CC, Ryoo MS, Aggarwal JK (2009) Detection of object abandonment using temporal logic. Mach Vis Appl 20(5):271–281

8. Bird N, Atev S, Caramelli N, Martin R, Masoud O, Papanikolopoulos N (2006) Real time, online detection of abandoned objects in public areas. In: Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on, pp. 3775-3780.IEEE

9. Bouwmans T (2014) Traditional and recent approaches in background modeling for foreground detection: an overview. Comput Sci Rev 11:31–66

10. Breiman L (2001) Random forests. Mach Learn 45(1):5–32

11. Calderara S, Vezzani R, Prati A, Cucchiara R (2005) Entry edge of field of view for multicamera tracking in distributed video surveillance. In: Advanced Video and Signal Based Surveillance, 2005. AVSS 2005. IEEE Conference on. 93–98. IEEE

12. Candamo J, Shreve M, Goldgof DB, Sapper DB, Kasturi R (2010) Understanding transit scenes: a survey on human behavior-recognition algorithms. Intel Trans Syst, IEEE Trans 11(1):206–224

13. Cheng S, Luo X, Bhandarkar SM (2007) A multiscale parametric background model for stationary foreground object detection. In: Motion and Video Computing, 2007. WMVC'07. IEEE Workshop on. 18–18. IEEE

14. Chitra M, Geetha MK, Menaka L et al (2013) Occlusion and abandoned object detection for surveillance applications. Int J Comput Appl Technol Res 2(6):708 meta

15. Chuang CH, Hsieh JW, Tsai LW, Ju PS, Fan KC (2008) Suspicious object detection using fuzzy-color histogram. In: Circuits and Systems, 2008. ISCAS 2008. IEEE International Symposium on. 3546–3549. IEEE

16. Chuang CH, Hsieh JW, Tsai LW, Chen SY, Fan KC (2009) Carried object detection using ratio histogram and its application to suspicious event analysis. Circ Syst Video Technol, IEEE Trans 19(6):911–916

17. Collazos A, Fernandez-Lopez D, Montemayor AS, Pantrigo JJ, Delgado ML (2013) Abandoned object detection on controlled scenes using kinect. In: Natural and Artificial Computation in Engineering and Medical Applications, pp. 169–178. Springer

18. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

19. Cucchiara R, Grana C, Piccardi M, Prati A (2003) Detecting moving objects, ghosts, and shadows in video streams. Patt Anal Mach Intel, IEEE Trans 25(10):1337–1342

20. Denman S, Chandran V, Sridharan S (2008) Abandoned object detection using multi-layer motion detection. In: Proceedings of International Conference on Signal Processing and Communication Systems 2007, pp. 439-448. DSP for Communication Systems

21. Elgammal A, Harwood D, Davis L (2000) Non-parametric model for background subtraction. In: Computer VisionECCV 2000, pp. 751-767. Springer

22. Ellingsen K (2008) Salient event-detection in video surveillance scenarios. In: Proceedings of the 1st ACM workshop on Analysis and retrieval of events/actions and workflows in video streams, pp. 57-64. ACM

23. Evangelio RH, Sikora T (2010) Static object detection based on a dual background model and a finite-state machine. EURASIP J Image Video Proc 2011(1):858,502

24. Evangelio, R.H., Senst, T., Sikora, T. (2011) Detection of static objects for the task of video surveillance. In: Applications of Computer Vision (WACV), 2011 IEEE Workshop on, pp. 534-540. IEEE

25. Fan Q, Pankanti S (2011) Modeling of temporarily static objects for robust abandoned object detection in urban surveillance. In: Advanced Video and Signal-Based Surveillance (AVSS), 2011 8th IEEE International Conference on. 36–41. IEEE

26. Fan Q, Pankanti S (2012) Robust foreground and abandonment analysis for large-scale abandoned object detection in complex surveillance videos. In: Advanced Video and Signal-Based Surveillance (AVSS), 2012 IEEE Ninth International Conference on, pp. 58-63. IEEE

27. Fan, Q., Gabbur, P., Pankanti, S. (2013) Relative attributes for large-scale abandoned object detection. In: Computer Vision (ICCV), 2013 IEEE International Conference on, pp. 2736-2743. IEEE

28. Femi PS, Thaiyalnayaki K (2013) Detection of abandoned and stolen objects in videos using mixture of gaussians. International Journal of Computer Applications 70(10)

29. Fernandez-Caballero A, Castillo JC, Rodriguez-Sanchez JM (2012) Human activity monitoring by local and global finite state machines. Expert Syst Appl 39(8):6982–6993

30. Ferrando S, Gera G, Regazzoni C (2006) Classication of unattended and stolen objects in video-surveillance system. In: Video and Signal Based Surveillance, 2006. AVSS'06. IEEE International Conference on, pp. 21-21. IEEE

31. Ferryman J, Hogg D, Sochman J, Behera A, Rodriguez-Serrano JA, Worgan S, Li L, Leung V, Evans M, Cornic P et al (2013) Robust abandoned object detection integrating wide area visual surveillance and social context. Pattern Recogn Lett 34(7):789–798

32. Foggia P, Greco A, Saggese A, Vento M (2015) A method for detecting long term left baggage based on heat map

33. Foresti GL, Marcenaro L, Regazzoni CS (2002) Automatic detection and indexing of videoevent shots for surveillance applications. Multimed, IEEE Trans 4(4):459–471

34. Foresti GL, Mahonen P, Regazzoni CS (2012) Multimedia video-based surveillance systems: Requirements, Issues and Solutions, vol. 573. Springer Science & Business Media

35. Foucher S, Lalonde M, Gagnon L (2011) A system for airport surveillance: detection of people running, abandoned objects, and pointing gestures. In: SPIE Defense, Security, and Sensing, pp. 805,610-805,610. International Society for Optics and Photonics

36. Gouaillier V, Fleurant A (2009) Intelligent video surveillance: Promises and challenges. Technological and commercial intelligence report, CRIM and Techn^opole Defence and Security 456-468

37. Guler S, Farrow MK (2006) Abandoned object detection in crowded places. In: Proc. of PETS, pp. 18-23. Citeseer

38. Guler S, Silverstein J, Pushee IH et al. (2007) Stationary objects in multiple object tracking. In: Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on, pp. 248-253. IEEE

39. Haritaoglu I, Cutler R, Harwood D, Davis LS (1999) Backpack: Detection of people carrying objects using silhouettes. In: Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, vol. 1, pp. 102-107. IEEE

40. Haritaoglu I, Harwood D, Davis LS (2000) W 4: real-time surveillance of people and their activities. Patt Anal Mach Intel, IEEE Trans 22(8):809–830

41. Hoferlin M, Hoferlin B, Weiskopf D, Heidemann G (2015) Uncertainty-aware video visual analytics of tracked moving objects. J Spatial Inform Sci 2:87–117

42. Hu, M.K. (1962) Visual pattern recognition by moment invariants. information Theory, IRE Transactions on 8(2), 179-187

43. Hu W, Tan T, Wang L, Maybank S (2004) A survey on visual surveillance of object motion and behaviors. Systems, Man, and Cybernetics, Part C: Appl Rev, IEEE Trans 34(3):334–352

44. Image. URL https://www.boardofstudies.nsw.edu.au/bosstats/images/2011/2011-crowds-8-lg:jpg

45. Image (Accessed on 01 June, 2016). URL https://www.123rf.com/stockphoto/exhibitioncrowd:html

46. Isard M, Blake A (1998) Condensation conditional density propagation for visual tracking. Int J Comput Vis 29(1):5–28

47. Jalal AS, Singh V (2012) The state-of-the-art in visual object tracking. Informatica 36(3)

48. Javed O, Shah M (2002) Tracking and object classification for automated surveillance. In: Computer Vision ECCV 2002, pp. 343–357. Springer

49. Joglekar UA, Awari SB, Deshmukh SB, Kadam DM, Awari RB (2014) An abandoned object detection system using background segmentation. In: International Journal of Engineering Research and Technology, vol. 3. ESRSA Publications

50. Kalman RE (1960) A new approach to linear filtering and prediction problems. J Fluids Eng 82(1):35–45

51. Khan S, Shah M (2003) Consistent labeling of tracked objects in multiple cameras with overlapping fields of view. Patt Anal Mach Intel, IEEE Trans 25(10):1355–1360

52. Khan SM, Shah M (2006) A multiview approach to tracking people in crowded scenes using a planar homography constraint. In: Computer Vision ECCV 2006, pp. 133-146. Springer

53. Kim J, Kim D (2014) Static region classication using hierarchical finite state machine. In: Image Processing (ICIP), 2014 IEEE International Conference on, pp. 2358{2362. IEEE

54. Kitagawa G (1987) Non-gaussian statespace modeling of nonstationary time series. J Am Stat Assoc 82(400):1032–1041

55. Kong H, Audibert JY, Ponce J (2010) Detecting abandoned objects with a moving camera. Image Proc, IEEE Trans 19(8):2201–2210

56. Krahnstoever N, Tu P, Sebastian T, Perera A, Collins R (2006) Multi-view detection and tracking of travelers and luggage in mass transit environments. In: In Proc. Ninth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), vol. 258

57. Lavee G, Khan L, Thuraisingham B (2005) A framework for a video analysis tool for suspicious event detection. 79–84

58. Lavee G, Khan L, Thuraisingham B (2007) A framework for a video analysis tool for suspicious event detection. Multimed Tools Appl 35(1):109–123

59. Li L, Luo R, Ma R, Huang W, Leman K (2006) Evaluation of an ivs system for abandoned object detection on pets 2006 datasets. In: Proceedings of the 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (PETS'06), pp. 91-Citeseer

60. Li Q, Mao Y, Wang Z, Xiang W (2009) Robust real-time detection of abandoned and removed objects. In: Image and Graphics, 2009. ICIG'09. Fifth International Conference on, pp. 156-161. IEEE

61. Li X, Zhang C, Zhang D (2010) Abandoned objects detection using double illumination invariant foreground masks. In: Pattern Recognition (ICPR), 2010 20th International Conference on, pp. 436-439. IEEE
62. Liao HH, Chang JY, Chen LG (2008) A localized approach to abandoned luggage detection with foreground-mask sampling. In: Advanced Video and Signal Based Surveillance, 2008. AVSS'08. IEEE Fifth International Conference on, pp. 132-139. IEEE
63. Lin K, Chen SC, Chen CS, Lin DT, Hung YP (2014) Left-luggage detection from finite state-machine analysis in static-camera videos. In: 2014 22nd International Conference on Pattern Recognition (ICPR), pp. 4600-4605. IEEE
64. Lin K, Chen SC, Chen CS, Lin DT, Hung YP (2015) Abandoned object detection via temporal consistency modeling and back-tracing verification for visual surveillance. Info Forensics Sec, IEEE Trans 10(7): 1359–1370
65. Lo B, Velastin S (2001) Automatic congestion detection system for underground platforms. In: Intelligent Multimedia, Video and Speech Processing, 2001. Proceedings of 2001 International Symposium on, pp. 158-161. IEEE
66. Lu S, Zhang J, Feng D (2006) A knowledge-based approach for detecting unattended packages in surveillance video. In: Video and Signal Based Surveillance, 2006. AVSS'06. IEEE International Conference on, pp. 110-110. IEEE
67. Lu S, Zhang J, Feng D (2007) An efficient method for detecting ghost and left objects in surveillance video. In: Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on, pp. 540-545. IEEE
68. Lu S, Zhang J, Feng DD (2007) Detecting unattended packages through human activity recognition and object association. Pattern Recogn 40(8):2173–2184
69. Lv F, Song X, Wu B, Singh VK, Nevatia R (2006) Left luggage detection using Bayesian inference. In: Proc. of IEEE Int. Workshop on Performance Evaluation of Tracking and Surveillance, pp. 83-90. Citeseer
70. Maddalena L, Petrosino A (2013) Stopped object detection by learning foreground model in videos. Neural Netwo Learn Syst, IEEE Trans 24(5):723–735
71. Mahin FS, Islam MN, Schaefer G, Ahad MAR (2015) A simple approach for abandoned object detection. In: Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV), p. 427. The Steering Committee of The World Congress in Computer Science, Computer Engineering and Applied Computing (WorldComp)
72. Martinez-del Rincon J, Herrero-Jaraba JE, Gomez JR, Orrite-Urunuela C (2006) Automatic left luggage detection and tracking using multi-camera ukf. In: Proceedings of the 9th IEEE International Workshop on Performance Evaluation in Tracking and Surveillance (PETS'06), pp. 59-66
73. Mathew R, Yu Z, Zhang J (2005) Detecting new stable objects in surveillance video. In: Multimedia Signal Processing, 2005 IEEE 7th Workshop on, pp. 1-4. IEEE
74. McHugh JM, Konrad J, Saligrama V, Jodoin PM (2009) Foreground-adaptive background subtraction. Sign Proc Lett, IEEE 16(5):390–393
75. Miguel JCS, Martinez JM (2008) Robust unattended and stolen object detection by fusing simple algorithms. In: Advanced Video and Signal Based Surveillance, 2008. AVSS'08. IEEE Fifth International Conference on, pp. 18-25. IEEE
76. Mittal A, Davis L (2001) Unified multi-camera detection and tracking using region-matching. In: Multi-Object Tracking, 2001. Proceedings. 2001 IEEE Workshop on, pp. 3-10. IEEE
77. Mukherjee D, Wu QJ, Nguyen TM (2014) Gaussian mixture model with advanced distance measure based on support weights and histogram of gradients for background suppression. IEEE Trans Ind Inform 10(2): 1086–1096
78. Nam Y (2015) Real-time abandoned and stolen object detection based on spatio-temporal features in crowded scenes. Multimedia Tools and Applications pp. 1-26
79. Otoom AF, Gunes H, Piccardi M (2008) Feature extraction techniques for abandoned object classification in video surveillance. In: 2008 15th IEEE International Conference on Image Processing, pp. 1368-1371. IEEE
80. Pan J, Fan Q, Pankanti S (2011) Robust abandoned object detection using region-level analysis. In: Image Processing (ICIP), 2011 18th IEEE International Conference on, pp. 3597-3600. IEEE
81. Pavithradevi MK, Scholar P (2014) Detection of suspicious activities in public areas using staged matching technique. IJAICT 1
82. Pets 2001 benchmark data, (2001). URL http://www.cvg.rdg.ac.uk/PETS2001/
83. Pets 2007 benchmark data (2007). URL http://www.cvg.rdg.ac.uk/PETS2006/data.html
84. Piccardi, M. (2004) Background subtraction techniques: a review. In: Systems, man and cybernetics, 2004 IEEE international conference on, vol. 4, pp. 3099-3104. IEEE
85. Popoola OP, Wang K (2012) Video-based abnormal human behavior recognitiona review. Systems, Man, and Cybernetics, Part C: Appl Rev, IEEE Trans 42(6):865–878

86. Poppe R (2010) A survey on vision-based human action recognition. Image Vis Comput 28(6):976–990
87. Porikli F (2007) Detection of temporarily static regions by processing video at different frame rates. In: Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conferenceon, pp. 236-241. IEEE
88. Porikli, F., Yin, Z. (2007) Temporally static region detection in multi-camera systems. In: Proc. 10th IEEE Int. Workshop on Performance evaluation of tracking and surveillance: PETS, pp. 79-86
89. Porikli, F., Ivanov, Y., Haga, T. (2008) Robust abandoned object detection using dual foregrounds. EURASIP Journal on Advances in Signal Processing 2008, 30
90. Prabhakar G, Ramasubramanian B (2012) An efficient approach for real time tracking of intruder and abandoned object in video surveillance system. Int J Comput Appl 54(17):22–27
91. Regazzoni CS, Fabri G, Vernazza G (1998) Advanced video-based surveillance systems, vol. 488. Springer Science & Business Media
92. Sacchi C, Regazzoni CS (2000) A distributed surveillance system for detection of abandoned objects in unmanned railway environments. Vehicular Technol, IEEE Trans 49(5):2013–2026
93. Sajith, K., Nair, K.R. (2013) Abandoned or removed objects detection from surveillance video using codebook. In: International Journal of Engineering Research and Technology, vol. 2. ESRSA Publications
94. SanMiguel JC, Martinez JM (2012) A semantic-based probabilistic approach for real-time video event recognition. Comput Vis Image Underst 116(9):937–952
95. SanMiguel J, Caro L, Martinez J (2012) Pixel-based colour contrast for abandoned and stolen object discrimination in video surveillance. Electron Lett 48(2):86–87
96. Singh VK, Atrey PK, Kankanhalli MS (2008) Coopetitive multi-camera surveillance using model predictive control. Mach Vis Appl 19(5-6):375–393
97. Singh, R., Vishwakarma, S., Agrawal, A., Tiwari, M. (2010) Unusual activity detection for video surveillance. In: Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia, pp. 297-305. ACM
98. Smith KC, Quelhas P, Gatica-Perez D (2006) Detecting abandoned luggage items in a public space. Tech. rep., IDIAP
99. Spengler M, Schiele B (2003) Automatic detection and tracking of abandoned objects. In: Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance. Citeseer
100. Stauffer C, Grimson WEL (1999) Adaptive background mixture models for real-time tracking. In: Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on, vol. 2. IEEE
101. Stringa E, Regazzoni CS (1998) Content-based retrieval and real time detection from video sequences acquired by surveillance systems. In: Image Processing, 1998. ICIP98. Proceedings. 1998 International Conference on, pp. 138-142. IEEE
102. Stringa E, Regazzoni CS (2000) Real-time video-shot detection for scene surveillance applications. Image Proc, IEEE Trans 9(1):69–79
103. Szwoch G (2016) Extraction of stable foreground image regions for unattended luggage detection. Multimed Tools Appl 75(2):761–786
104. Tejas Naren TN, Shankar Siddharth KA, Venkat Krishnan S, Sanjeevi LR (2014) Abandoned object detection for automated video surveillance using hadoop. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering 3(3). http://www.ijareeie.com/upload/2014/apr14-specialissue3/21_R22.pdf
105. Tian Y, Feris RS, Liu H, Hampapur A, Sun MT (2011) Robust detection of abandoned and removed objects in complex surveillance videos. Systems, Man, and Cybernetics, Part C: Appl Rev, IEEE Trans 41(5):565–576
106. Tian Y, Senior A, Lu M (2012) Robust and efficient foreground analysis in complex surveillance videos. Mach Vis Appl 23(5):967–983
107. Tripathi RK, Jalal AS (2014) A framework for suspicious object detection from surveillance video. Int J Mach Intel Sensory Sign Proc 1(3):251–266
108. Tripathi RK, Jalal AS, Bhatnagar C (2013) A framework for abandoned object detection from video surveillance. In: Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG), 2013 Fourth National Conference on, pp. 1-4. IEEE
109. Tsai RY (1986) An efficient and accurate camera calibration technique for 3d machine vision. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 1986
110. Vezzani R, Cucchiara R (2010) Video surveillance online repository (visor): an integrated framework. Multimed Tools Appl 50(2):359–380
111. Wan EA, Van Der Merwe R (2000) The unscented kalman filter for nonlinear estimation. In: Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000, pp. 153-158. IEEE

112. Webb GI (2000) Multiboosting: a technique for combining boosting and wagging. Mach Learn 40(2):159–196

113. Witten I, Frank E (2000) Data mining: Practical machine learning tools with java implementations. MODELO MATEMATICO PARAMETRICO DE ESTIMACION PARA PROYECTOS DE DATA MINING IDMCOMOf

114. Wren CR, Azarbayejani A, Darrell T, Pentland AP (1997) Pfinder: real-time tracking of the human body. Patt Anal Mach Intel, IEEE Trans 19(7):780–785

115. Yadav P, Jahagirdar A (2015) A static object detection in image sequences by self organizing background subtraction

116. Yang Z, Rothkrantz L (2011) Surveillance system using abandoned object detection. In: Proceedings of the 12th International Conference on Computer Systems and Technologies, pp. 380-386. ACM

117. Yang T, Pan Q, Li SZ, Li J (2004) Multiple layer based background maintenance in complex environment. In: Image and Graphics (ICIG'04), Third International Conference on, pp. 112-115. IEEE

118. Yilmaz A, Javed O, Shah M (2006) Object tracking: a survey. A Comput Surveys (CSUR) 38(4):13

119. Yue, Z., Zhou, S.K., Chellappa, R. (2004) Robust two-camera tracking using homography. In: Acoustics, Speech, and Signal Processing, 2004. Proceedings.(ICASSP'04). IEEE International Conference on, vol. 3, pp. iii −1. IEEE

120. Zeng Y, Lan J, Ran B, Gao J, Zou J (2015) A novel abandoned object detection system based on three-dimensional image information. Sensors 15(3):6885–6904

121. Zhou Y, Benois-Pineau J, Nicolas H (2010) Multi-object particle filter tracking with automatic event analysis. In: Proceedings of the first ACM international workshop on Analysis and retrieval of tracked events and motion in imagery streams, pp. 21-26. ACM

122. Ziaeefard M, Bergevin R (2015) Semantic human activity recognition: a literature review. Pattern Recogn 8(48):2329–2345

123. Zin TT, Tin P, Toriu T, Hama H (2012) A novel probabilistic video analysis for stationary object detection in video surveillance systems. IAENG Int J Comput Sci 39(3):295–306

124. Zin TT, Tin P, Toriu T, Hama H (2012) A probability-based model for detecting abandoned objects in video surveillance systems. In: Proceedings of the World Congress on Engineering, 2

Rajesh Kumar Tripathi is working as an Assistant Professor in the Department of Computer Engineering and Applications, Institute of Engineering and Technology, GLA University, Mathura, UP, INDIA. He received his Master of Computer Application degree from the Uttar Pradesh Technical University, Lucknow, UP, INDIA. He received his M.Tech. in Computer Science and Engineering from the GLA University, Mathura, UP, INDIA. He has more than 12 years of teaching and research experience. His research interests include image processing and computer vision.

Anand Singh Jalal is working as a Professor and the Head in the Department of Computer Engineering and Applications, GLA University, Mathura, INDIA. He received his M.Tech. in Computer Science from the Devi Ahilya Vishwavidyalaya, Indore, MP, INDIA. He received his Ph.D. in the area of Computer Vision from the Indian Institute of Information Technology (IIIT), Allahabad, UP, INDIA. He has 18 years of teaching and research experience. His research interests include image processing, computer vision and pattern recognition.

Subhash Chand Agrawal is working as an Assistant Professor in the Department of Computer Engineering and Applications, Institute of Engineering and Technology, GLA University, Mathura, UP, INDIA. He received his Master of Computer Application degree from the Uttar Pradesh Technical University, Lucknow, UP, INDIA. He received his M.Tech. in Computer Science and Engineering from the GLA University, Mathura, INDIA. He has more than ten years of teaching experience. His research interests include image processing, pattern recognition and image retrieval.