Journal of Engineering Science and Technology
Vol. 8, No. 2 (2013) 199 - 216

PERFORMANCE METRICS IN VIDEO SURVEILLANCE SYSTEM

AZHAR MOHD IBRAHIM*, A. A. SHAFIE, M. M. RASHID

Department of Mechatronics Engineering, Faculty of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia
*Corresponding Author: mazary86@gmail.com

Abstract

Video surveillance is an active research topic in computer vision. One of the areas that are being actively researched is on the abilities of surveillance systems to track multiple objects over time in occluded scenes and to keep a consistent identity for each target object. These abilities enable a surveillance system to provide crucial information about moving objects behaviour and interaction. This survey reviews the recent developments in moving object detection and also different techniques and approaches in multiple objects tracking that have been developed by researchers. The algorithms and filters that can be incorporated in tracking multiples object to solve the occluded and natural busy scenes in surveillance systems are also reviewed in this paper. This survey is meant to provide researchers in the field with a summary of progress achieved up to date in multiple moving objects tracking. Despite recent progress in computer vision and other related areas, there are still major technical challenges that need to be solved before reliable automated video surveillance system can be realized.

Keywords: Surveillance systems, Object detection, Object tracking, Partial occlusion, Crowded scenes.

1. Introduction

Normally, a video surveillance system consist of a video camera, which has coverage of a constrained area defined by the fields of views of the camera. The video streams are then transmitted to a central location, displayed on one or several video monitors and recorded. The person observes the continuous video to find out if there is any activity that demands a response. Intelligent video surveillance aims to exploit the video streams via software to automatically identify specific objects, behaviours for various applications. Among the applications are:

Nomenclatures

FGobject	Foreground object
I_t	Pixel's intensity of the current image, pixel
I_{t-i}	Pixel's intensity of the image occurring i frames in the past, pixel
Th	Threshold value

Greek Symbols

Δt	Difference of two frames, pixel

Surveillance system of vehicles. Applications of this surveillance system are such as in reporting traffic congestion [1], accidents [2] and dangerous behaviour by road users [3]. Some surveillance systems, which can recognize the vehicle number plates as in [3], can be very useful to traffic departments.

Surveillance system of human. Applications on this surveillance system have the ability to analyse behaviour of the people, whether it is abnormal or normal behaviour [4]. If it is abnormal behaviour, the alarm will be automatically triggered. Besides that, the surveillance system also can be used for identification of humans using salient features of the human such as the face [5, 6]. Moreover, it can be used at public places such as supermarkets, homes [7], banks, and parking lots [8]. Figure 2 shows general flowchart of video surveillance system.

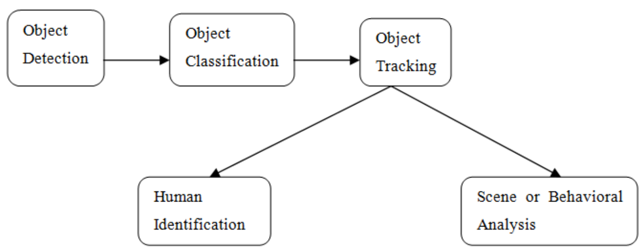


Fig. 1. General Flowchart of Video Surveillance System.

In this paper, we discussed different techniques of multiple objects tracking in a surveillance system. In section 2, we discussed object detection which is background modelling and segmentation of motion. Section 3 described several methods that have been used in moving object tracking, while section 4 discussed the algorithms and filters for tracking multiple objects. Section 5 is the conclusion of this paper.

2. Object Detection

Detection of moving objects in video streams is the first step in extracting information for many computer vision applications including video surveillance as in Fig. 2. Object detection consists of background modelling and also object segmentation.

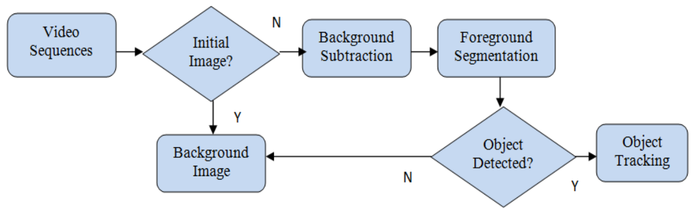


Fig. 2. General Flowchart for Moving Object Detection.

2.1. Background modelling

Background modelling is a step where an image is considered as a background in which it contains the non-moving objects in a video. Usually, the background image needs to be initialized and also maintained by updating it because of the changes in the scene. In order to achieve very good background modelling, it should also tolerate the variation in illumination.

Huwer and Niemann [9] proposed a method which combines a temporal difference method with an adaptive background model subtraction scheme so that the background model is adapted to the variation in illumination. This method performed at least 12% of recognition rates better than the method of adaptive background estimation which uses a modified Kalman filtering technique as in [10]. On the other hand, Gao et al. [11] proposed an adaptive background model estimated by using the Kalman filtering model which is based on local-region and can be used to predict the parameters of the model using a Recursive Least Square Adaptive Filter. The proposed algorithm is adaptive to the illumination changes and variation of input noise.

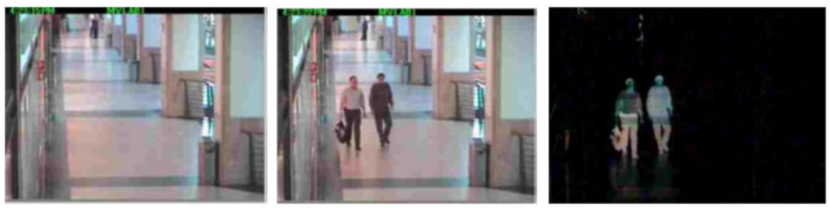
The background model should also be able to tolerate noise disturbances such as swaying tree branches; where a pixel sometimes observes the branch, changes introduced to the background image such as like objects entering the video frame and keep on being there without moving for a long period of time, and also rippling water which may reflect the sky. Grimson et al. [12] proposed a robust detector that automatically adapts to the video scene by considering each pixel as an independent statistical process, and record the observed intensity at each pixel over the previous frames. Al Najjar et al. [13] proposed a hybrid adaptive background model based on selective Gaussian modelling. The methods in [12, 13] can be used to solve the background problems such as illumination changes and also swaying of the tree branches.

2.2. Object segmentation

Object segmentation is the next step after background modelling. Object segmentation differentiates the background image with the foreground objects. In order to detect the moving object, the pixels which belong to foreground objects need to be distinguished from the background image's pixels. There are three main approaches in order to recognize foreground objects' pixels which are background subtraction, temporal differencing and optical flow.

2.2.1. Background subtraction

Background subtraction involves comparing a foreground image with a background image. The pixels where there is a considerable difference between the observed and background images show the location of the objects of interest. This can be done by subtracting the observed frame from the video scene from the background image from the video scene and thresholding the result to detect the objects of interest as shown in Fig. 3 (video scene from CAVIAR Test Set [14]).



a) Background Image b) Region of Interest c) Foreground Objects

Fig. 3. Results of Background Subtraction.

Background subtraction has better performance in extracting pixels of the foreground objects but it is very responsive to dynamic changes in the video scene such as illumination changes, noise disturbances such as swaying tree branches, and also changes introduced to the background image such as objects entering the video frame and remaining there without moving for a long period of time. Hence, adaptive background subtraction techniques that can automatically create a background image and continuously upgrade are very essential in order to achieve accurate detection of foreground objects [15].

For the person finder or Pfinder in [16], they used a Gaussian model to model the intensity of a single pixel with a single Gaussian distribution as a means to continuously update the background image. However, the method cannot be used to compensate for sudden illumination changes. Tiehan et al. [17] proposed a background subtraction algorithm to separate foreground objects from the background by compensating the input frames and compared with the given reference frame. They also replaced the widely used morphology operations by adopting a spatial filter to suppress various noises due to the swaying of the tree branches. Here, the method cannot compensate for illumination changes and large scale camera motion. Stauffer and Grimson [15] proposed adaptive background subtraction by modelling each pixel as a mixture of Gaussians and using an on-line approximation to update the background image so that can be used reliably with lighting changes, repetitive motions from clutter, and long-term scene changes. However, the technique using mixture of Gaussians involves lots of computations.

2.2.2. Temporal differencing

Temporal differencing [18] is done by comparing the pixels of two or more consecutive frames which is separated by constant times. Temporal differencing is very suitable for dynamic environments such as variation of illumination and also it does not require background initiation as in the background subtraction method. This method generally can be calculated as follows:

$$\Delta t = \max\{|I_t - I_{t-i}\} \quad (1)$$

where I_t is the pixel's intensity of the current image, and I_{t-i} is the image occurring i frames in the past. When a pixel's intensity value changes swiftly, the value of Δt increases. If the change in the intensity of the pixels at location (x, y) is more than threshold value, so, that pixel is considered as foreground object that is assigning 1 for foreground object,

$$FGobject(x,y) = \begin{cases} 1, \Delta t(x,y) > Th \\ 0, \Delta t(x,y) \leq Th \end{cases} \quad (2)$$

where Th is a threshold for classifying whether the pixel belongs to an object or background.

Rahim et al. [19] detected the moving vehicles using temporal differencing method or also known as frame differencing method. However, this method has a limitation in extracting all relevant feature pixels. Fujiyoshi et al. [20] proposed temporal differencing method using adaptive thresholding calculated from intensity changes in the past few frames in order to detect the moving objects in the outdoor environment. Recently, Zhang et al. [21] proposed motion detection technique by combining background subtraction and temporal differencing methods. Firstly, for background subtraction using adaptive background model by Gaussian model for each pixel in the image sequences, then, to update the background image is done through using three temporal differences to search for the overlapping parts of continual three frames. However, the method cannot be used when the occlusion occurs and sudden lightning changes occur.

Yoo and Park [22] proposed the temporal differencing method based on signed difference which represents the pattern of motion in a given region. In this method, a signed difference image is acquired by subtracting two consecutive frames then for each fixed blocks in the signed difference image, a motion pattern is calculated by Earth Mover's Distance (EMD). Li and Leung [23] presented combination of two motion techniques which are background subtraction and temporal differencing.

2.2.3. Optical flow

Optical flow [24] is a two-dimensional vector field that represents velocities and their directions in each point of the image. Hence, Optical flow allows moving objects to be detected based on their individual velocities. The advantages of Optical flow method are it is not sensitive to illumination changes and also noises such as shadows. Denman et al. [25] presented an algorithm that uses the motion segmentation results to update the optical flow calculations and ensure that optical flow is only calculated at pixel resolution. Tracking of the flow vectors is employed to improve performance and detect discontinuities, which can indicate the location of overlaps between objects.

Andrade et al. [26] proposed framework that detects the changes in the video scenes from optical flow and encoded by a Hidden Markov Model to classify abnormal events. Kemouche and Aouf [27] presented a motion segmentation approach based on a subtraction of background model that integrates optical flow information with colour information in the background model in which the colour background subtraction model is based on spatially global Gaussian mixture. However, moving object detection using optical flow involves expensive computations and not really suit for real-time surveillance systems. Recently, Zhang [28] presented an algorithm of detecting and tracking moving objects based on optical flow in polar-log image, as a result it increased the accuracy of optical flow calculation and also reduced the computation time.

In this survey, some of the limitations that normally faced in the object segmentation, methods that researchers used to overcome it and the results achieved are given in Table 1. Here, the system developed whether outdoor or indoor environment and the results that were given in Table 1 are as stated in their respective publications.

3. Object Tracking

When the objects is successfully detected in the video surveillance system, the next step in the video surveillance system is to successfully track the moving objects from one frame to another in following scenes. Moving object tracking is a vital task in computer vision systems for surveillance, traffic monitoring or intelligent surveillance systems. When multiple tracked objects merge into groups with various complexities of occlusion, tracking each individual object through crowds becomes a challenging task.

One of the most crucial criterions for intelligent surveillance systems is to track multiple objects over time in occluded scenes and to keep a consistent identity for each target object. This is due to its ability to provide crucial information about human behaviours, human interactions, and relationships between objects of interest [29]. In order to achieve real-time tracking of the moving objects, the exact position of moving objects must also be confirmed.

There are many challenges in tracking the moving objects such as associate target objects in consecutive video frames. This is difficult when the objects are moving fast relative to the frame rate. There are also other problems in tracking the moving objects such as when the features in the background images occluded to the features of the foreground objects. This can happen in the case where there are many people in the scene or in the occluded scenes.

According to Yilmaz et al. [30], occlusion itself can be divided into self-occlusion, interobject occlusion, and also background occlusion. Self-occlusion can occur when tracking articulated objects and it refers to the feature of an object occludes to another. Meanwhile, interobject occlusion happened when handling the tracking of multiple objects that are occluding each other. Similarly, occlusion by the background occurs when the features in the background images occluded to the features of the tracked objects.

There are many tracking methods which can be divided into five categories in general: region-based tracking, active contour or snakes based tracking, feature-based tracking, model-based tracking, and hybrid tracking. Besides that, there are many algorithms or mathematical tools which are being used by researchers for filtering and also for data association issue (to solve occluded objects) which also discussed here. The engrossed reader can use the references attached in this paper for more details on these tracking of multiple objects techniques.

Table 1. Performance Metrics for Background Modelling (O: Outdoor Environment and I: Indoor Environment).

Limitations	Solving Methods	Results	O	I	Ref.
Illumination changes	Combination of adaptive background subtraction and temporal differencing	The method performs 12% better for the static change detection and it performs 17% better result when the background representation was adapted to the image sequence.	Y	Y	[9]
	Adaptive Background Subtraction	The method gives satisfying results in updating the background model in the situation of late afternoon, a rainy day, and illumination under streetlight in the evening.	Y	N	[11]
Variant Input Noise; swaying of tree branches, rainy day, etc.	Adaptive Background Subtraction	Continuously adapt the scene in outdoor environment.	Y	Y	[12]
	Hybrid adaptive background model	The method gives precision of 0.9561, false alarm (0.2586) and detection failure (0.0439).	Y	N	[13]
Shaking Camera and input noise of swaying tree branches and shadows	Adaptive Background Subtraction	The method performs 15 fps, 2.13 of CPI (average number of clock cycles for executing an instruction) and 0.74 of MPI (average number of data memory access per instruction executed)	Y	Y	[16]
Illumination Variation, repetitive motions from clutter, long-term scene changes and rainy and snowy day	Adaptive Background Subtraction	The method performs 11-13 fps.	Y	Y	[17]
	Temporal Differencing	The method performs 30 fps. The results are given in the visual sample comparing with other method.	Y	Y	[22]
Illumination Changes, Dynamic scene changes and Background Initiation	Temporal Differencing	The results for targeted vehicles are shown in chart form and visual sample.	Y	N	[19]
	Adaptive Temporal Differencing	The method performs 11-13 fps, successfully detects 92.5% human, 93.5% human group, 92.5% bike and 100% vehicle.	Y	N	[20]
	Combination of background subtraction and temporal differencing	The method gives the detection rate of 92.6%, the rate of false negative was 7.4%, and the rate of false positive was 0%.	N	Y	[23]
Illumination changes, backgrounds disturbing, shadow, etc.	Combination of background subtraction and temporal differencing	The method gives 100% accuracy in indoor area and 94.3% accuracy in outdoor area	Y	Y	[21]

3.1. Region based tracking

The idea here is to identify a connected region in the image associated with each moving objects and then track it over time. Initialization of the process is most easily done by the background subtraction technique and motion regions are usually detected by subtracting the background image from the current image. Adaptive background subtraction techniques can be used for background subtraction or existing methods, and then tracked over time using information provided by the entire region such as motion, size, colour, shape, texture and centroid. The advantages of region based method are it does not required to predict the spatial position of blobs nor velocity of the moving objects and also it is quite simple also. Besides that, it is also computationally efficient and works well in not occluded scenes such as free-flowing traffic.

Fang et al. [31] presented multiple object tracking using region based method by separating motion of moving objects into five states: entering, leaving, merging, splitting and normal. Region corresponding is used to detect these 5 states. However, for splitting states, region based cannot be used due to the problems of distinguishing and labelling the objects. So, the adaptive colour model for each target using colour-histogram statistical characteristics is modelled. Colour-histogram statistical based tracking for splitting state is good; yet, it is very sensitive when the persons wear similar colour clothes.

Besides that, Luo and Bhandarkar [32] proposed region-based multiple object tracking using Kalman filtering and elastic matching. The Kalman filtering algorithm is used as the velocity prediction model on account of its simplicity. The elastic matching algorithm is used to measure the velocity field which is then approximated using B-spline surfaces. Since this region based tracking used Kalman filtering and also elastic matching, it does not rely on background subtraction.

3.2. Active contour based tracking or snakes

The idea is to have a representation of the bounding contour of the object and keep dynamically updating it. This method instead of tracking the whole set of pixels comprising an object, the algorithms track only the contour of the object. Chen [33] combined real-time tracking object recognition using contour-based tracking and neural fuzzy network. Contour-based model is used to extract the object's feature vector and also to track the object and a self-constructing neural fuzzy inference network is used to train and recognize moving objects.

Baumberg and Hogg [34] combined dynamic filtering with a modal based flexible shape model to track an articulated non-rigid body in motion. The active shape model used was generated automatically from real image data and incorporates variability in shape due to orientation as well as object flexibility. This method can be used to track the silhouette of a walking pedestrian in real time, but the poses and views should be well represented in the training set in order to track well.

Chen et al. [35] integrated Hidden Markov Model and unscented Kalman filter into powerful parametric contour tracking framework in a nonlinear dynamic system. This method resulted in better contour smoothness constraint based on joint probabilistic matching and robust to lighting or appearance changes. The advantage of having a contour based representation is it can reduce computational complexity but it requires accurate initialization.

3.3. Feature-based tracking

The idea here is to abandon tracking objects as a whole but instead tracks sub-features such as distinguishable points or lines on the objects. Here, the tracking algorithm is done by extracting elements, clustering them into higher level features and then matching the feature between images. The advantage of this approach is that even in the presence of partial occlusion, some of the sub-features of the moving object remain visible such as face [5, 6] in the case of human, so that visible features can be used track the moving object. However, there is a problem in grouping similar set of features that belongs to the same object such as in the case of vehicles.

In the case of vehicle tracking, Beymer et al. [36] used vehicle sub-features as a tracking feature of the vehicle to make the system robust to partial occlusion. In order to group together sub-features that come from the same vehicle, the constraint of common motion is used.

Choi et al. [37] developed an image based system that detects and tracks multiple moving vehicles from the sequences of images using quad-tree segmentation and the Scale-invariant Feature Transform (SIFT) to improve the performance which is robust to changes of the intensity, shape, and scale of object caused by movement. Feature-based methods try to extract salient features such as edges and corners in order to track the moving vehicle.

3.4. Model-based tracking

Tracking an object in a video sequence means continuously identifying its location when either the object or the camera is moving. Model-based tracking method track objects by matching projected object model which is constructed with prior knowledge, to foreground image [38]. The prior object model using 3D knowledge can come in the form of a CAD model of a scene object, a set of planar parts, or even a rough 3D model such as an ellipsoid which can be created using either automated techniques or commercially available products [39]. The human body model state consists of information about 3D location and orientation of the body and also describing the articulation of the body.

Conaire et al. [40] used adaptive appearance model to track the objects such as human and vehicles. Here the object models represented as rectangular grids of pixels, with each pixel modelled as a Gaussian distribution. Additionally, each pixel is assigned an importance depending on how often it occurs as foreground. This is used to down-weight the background pixels that should not be tracked.

3.5. Hybrid tracking

Hybrid tracking is designed as combination methods between region-based and feature-based tracking. In this method, the advantages of two methods are used by considering first the object as an entity or in the aspect of region and then by tracking its salient features. Dan and Yuan [41] presented real-time tracking in video surveillance system using both region and feature based tracking. Moving object is detected through foreground detection. Then, three features of each moving objects such as centroid, area, and average luminance are extracted. At last, the similarity function is applied to tracking. This method has good performance under dynamic scene for real-time tracking such as in different lighting and weather conditions.

Lin et al. [42] proposed a robust region and feature-based tracking algorithm with plentiful features to track objects continuously with occlusion or splitting events. Adaptive background reconstruction technique is used to handle with environmental changes and to obtain effective results of objects extraction.

Besides the methods that described in above, there are some other approaches that can be used for object tracking. Colour-based methods can track individual persons separated from a group even though they have changed motion directions before the separation. Human can be easily tracked using this method because a human is characterized by the statistics of few most significant colour features from his or her body. Li et al. [43] proposed a method for tracking persons based on the principal colours of human objects. First, an efficient human object representation method, principal colour representation (PCR), is proposed. Then, in order to track human objects, similarity measures for object matching based on PCR are introduced. These similarity measures could be used to evaluate the matching between two individuals as well as visual evident of an individual in a group. Colour based tracking is good; however, it is very sensitive when the persons wear similar colour clothes.

Li et al. [44] extend the works in [43] by presenting dominant colour histogram (DCH) for object modelling, where the dominant colours are selected by a distance measure. Here the DCH is integrated into an efficient sequential tracking algorithm for tracking multiple objects through crowds. The method is efficient and robust for tracking multiple objects (≥3) in complex occlusion for real time surveillance systems.

Recently, You et al. [6] presented a tracking algorithm by combining feature (face) and also colour histogram to track the feature. Colour histogram is used as the second feature, the tracking feature. Objects are represented as blobs and then the colour histograms in the blobs differentiate objects in a video and dynamically updated based on the subject identity.

4. Algorithms and Filters for Tracking Multiple Objects

4.1. Multiple hypotheses tracking (MHT)

There are also some other algorithm and filters that can be utilized as approaches to track the objects. The Multiple Hypotheses Tracking (MHT) is a tracking method where alternative data association hypotheses are generated whenever an occluded scene arises while tracking multiple objects. According to Yilmaz et al. [30], the Multiple Hypothesis Tracking (MHT) algorithm has several advantages such as it maintains several hypotheses for each object at each time frame. The algorithm has the ability to track objects entering the scene and terminate tracks for objects exiting the scene. It can also handle occlusions by tracking continuously even if some of the measurements from an object are missing. For each hypothesis, a prediction of each object's position in the next frame is made.

The predictions are then compared with actual measurements by evaluating a distance measure.

Polat et al. [45] presented object tracking method by combining MHT algorithm with Hausdorff matching algorithm to organize individual edges into objects given their two-dimensional models. The combined technique provides a probabilistic tracking framework which is able of tracking complex objects in messy background in video surveillance system.

4.2. Kalman filter

The Kalman filter is a set of mathematical equations that provides an efficient computational means to estimate the state of a process of past, present, and even future states. Kalman filter assumes the states are in Gaussian distribution. To use Kalman filter for object tracking, the motion of the object is assumed to be almost constant over frames. Robert [46] proposed night time vehicles tracking using Kalman filter which is associated with a reasoning module. Here, the vehicle detection consists of detecting its two headlights. Headlights are detected by looking for bright masses of pixels in the region of interest. Feature-based tracker is used here to tracks a vehicle as a line by incorporating a simple Kalman filter into the system as a prediction mechanism. The proposed method gives good results for objects tracking, but it also computationally expensive by performing minimum rate of 12 fps only.

4.3. Probabilistic data association filter

The Probabilistic Data Association Filter (PDAF) is an extension of the Kalman filter that uses a Bayesian approach to update the state when there is a single target and possibly no measurements or multiple measurements due to noise. The PDAF consists of the two main blocks which are data association and track update. Cheng and Hwang [47] proposed mechanism based on Kalman filtering and modified Probabilistic Data Association method to enhance its performance and make it more suitable for vision-based systems.

4.4. Particle filter

The particle filter is a sequential Monte Carlo algorithm which is a sampling method for approximating a distribution that makes use of its temporal structure. Particle filters are suboptimal filters. They perform sequential Monte Carlo (SMC) estimation based on a point mass representation of probability densities. Here the distribution are $P(x_t|z_{0:t})$ where x_t is the unobserved state at time t, and $z_{0:t}$ is the sequence of observations from time 0 to time t.

Particle Filter can be used for tracking single and multiple objects. It is a hypothesis tracker that approximates the filtered posterior distribution by a set of weighted particles. It weights particles based on a likelihood score and then propagates these particles according to a motion model.

Wang et al. [48] presented a multi-feature fusion model based on a particle filter for moving object tracking. Here, the authors extract multiple feature information which is colour and edge orientation information in a particle filter to achieve better tracking performance. Yang et al. [49] also used particle filter to the tracked objects using colour and edge orientation histogram features.

Xiong and Debrunner [50] proposed vehicles tracking by integrating shape and colour feature in a particle filter. Here, the particle filter perform Sequential Monte Carlo sampling method to achieve good tracking results that robust to partial occlusion, illumination change, significant clutter, target scale variations, rotations in depth and similar background colours by integrating shape and colour features.

There are also extended versions of particle filter that can be used to achieve better tracking results. Rao-Blackwellized Particle Filter is one of an extension of the particle filter. Sarkka et al. [51] presented Rao-Blackwellized (RB) particle filtering to track the multiple objects and also the births and deaths of the targets are modelled as hidden stochastic processes, which are observed through the measurements. Here, RB is used to reduce the computational cost of a multi target Monte Carlo filter.

Zhai and Yeary [52] used new algorithm for particle filter which is multiple model particle filtering which use single dynamic model, the new algorithm uses a switching state space model and a jump Markov process to approximate the true target dynamics unlike other multiple object tracking methods. Ryu and Huber [53] introduced extension to the particle filter algorithm for tracking multiple objects by presenting an approach that instantiates separate particle filters for each object and explicitly handles partial and complete occlusion for nontransparent objects that can successfully instantiate and remove filters of objects that enter or leave the scene.

4.5. Probability hypothesis density (PHD) filter

The probability hypothesis density (PHD) filter is used as a computationally efficient algorithm to propagates the first order moment of the multi-target posterior (instead of the full posterior) to reduce the growth in complexity with the number of targets from exponential to linear. The PHD filter also can be used to compensate for miss-detections and to remove noise and clutter. Nam Trung Pham et al. [54] presented a method for tracking multiple objects from video data using probability hypothesis density filter on colour measurements. Maggio et al. [55] proposed a multi-target tracking algorithm based on the Probability Hypothesis Density (PHD) filter and data association using graph matching. The proposed algorithm has the capability to remove non persistent clutter, filter miss-detections, and improves the robustness of the tracker to clutter by verifying the coherence of consecutive sets of detections, without significantly increasing the complexity of the overall algorithm.

There are some other approaches for object tracking. Yu et al. [56] presented variation of expectation maximization (EM) algorithm as a solution for multi-target tracking with an efficient data association computation. Youssef et al. [57] used discrete wavelet transform to detect and track multiple moving objects and exploited objects' colour and spatial information to recognize the moving objects.

Table 2 provides the analysis in 22 papers that were published on multiple object tracking. The survey shows the increasing research on the multiple objects tracking by using different methods especially to solve the crowded and occluded scenes as in Fig. 4. The percentage increases more than 50% from 2007-2010 compared to from 2003-2006. The researchers trying to perform the multiple objects tracking that can be used in real time surveillance system with less computation time (higher frame per second, fps) and continuously detect and track the objects (humans and vehicles) in any environment (indoor and outdoor) even in completely occluded and crowded scenes.

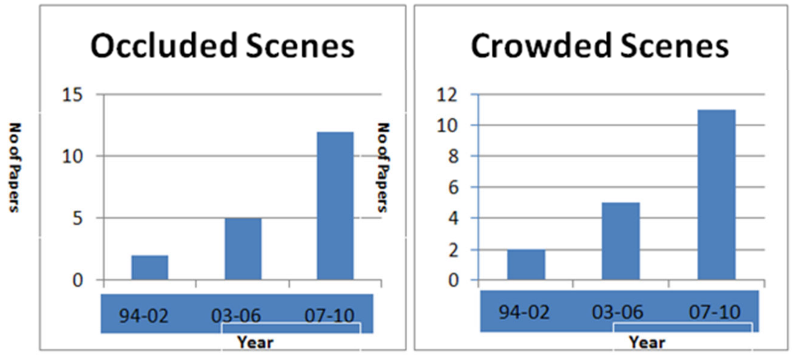


Fig. 4. Number of Papers Published on Occluded and Crowded Scenes in Multiple Object Tracking.

5. Conclusions

In this survey, an overview of the pre-processing steps for moving object detection and multiple objects tracking has been reviewed. As for the detection of moving objects, it involves background modelling and foreground segmentation. Three techniques for foreground segmentation are addressed; background subtraction, temporal differencing, and optical flow. We have discussed five general approaches to track the moving objects; region based, active-contour based, feature based, model based, and hybrid which is combination of region based and future based. We also have reviewed the mathematical tools that can be included in tracking multiple objects to solve the occluded and natural busy scenes in surveillance systems.

Most current automated video surveillance systems can process video sequence and perform object detection and segmentation, and also object tracking. However, in practice, even the best methods suffer such failures all too often, for example because the motion is too fast, a complete occlusion occurs, or simply because the target object moves momentarily out of the field of view. The adaptive background modelling and reliable method in tracking multiple moving objects is needed in order to achieve real time surveillance systems. Hence, to achieve this purpose, many problems such as object appearance changes, background clutter, partial occlusion and temporal lost should be well addressed and solved.

Table 2. Analysis in Multiple Object Tracking as Stated in the Respective Papers.

Year	Techniques	H	V	O	I	C	OC	Ref.
10	Discrete wavelet transform	Y	N	Y	Y	Y	Y	[57]
10	Feature based tracking, feature face, kernel tracker.	Y	N	N	Y	N	Y	[5]
09	Head light detection, feature based and appearance based classifier and Kalman Filter for tracking.	N	Y	Y	N	Y	Y	[46]
09	Feature based tracking, feature face, body colour histogram	Y	N	N	Y	Y	Y	[6]
08	Expectation maximization (EM) algorithm	Y	N	N	Y	Y	Y	[56]
08	Color based tracking, Dominant colour histogram, combination of directed acyclic graphs (DAGs) and depth order for tracking	Y	Y	Y	Y	Y	Y	[44]
08	Feature based tracking, Color and edge feature, Particle filter for tracking	N	Y	Y	N	Y	Y	[48]
07	Region based tracking, Color histogram matching.	Y	Y	Y	Y	Y	Y	[31]
07	Kalman filtering and modified Probabilistic Data Association	N	Y	Y	N	Y	Y	[47]
07	Probability hypothesis density (PHD), Feature Color	Y	N	Y	Y	Y	Y	[54]
07	Probability hypothesis density (PHD) and graph matching	Y	Y	Y	Y	Y	Y	[55]
07	Feature based tracking, Quad tree segmentation, Scale invariant feature transform (SIFT)	N	Y	Y	N	Y	Y	[37]
07	Region and Feature based tracking, centroid, area and average luminance feature, similarity function	Y	N	Y	Y	N	N	[41]
06	Region and Feature based tracking, region matching (1 to 1, splitting and merging matching)	Y	Y	Y	N	Y	Y	[42]
06	Model based tracking, structure preserving SUSAN noise filtering, day and night time tracking.	Y	Y	Y	N	Y	N	[40]
06	Contour based tracking, Hidden Markov Model (HMM), unscented Kalman filter, feature face	Y	N	N	Y	N	N	[35]
05	Feature based tracking, Color and edge feature, Particle filter for tracking	Y	N	Y	Y	Y	Y	[49]
05	Region based tracking, Kalman filtering, Elastic matching	Y	N	N	Y	N	Y	[32]
04	Color and edge shape feature, Monte Carlo filter based tracking	N	Y	Y	N	Y	Y	[50]
03	Color based tracking, Principal Color of human, asymmetric similarity matching, spatial distance for tracking	Y	N	N	Y	Y	Y	[43]
97	Feature based tracking, Vehicle sub feature, DSP chips	N	Y	Y	N	Y	Y	[36]
94	Contour based tracking, shape model, Kalman Filter	Y	N	Y	Y	N	N	[34]

H: Include Human tracking, V: Include Vehicle tracking, O: Dataset include outdoor environment,
I: Dataset include indoor environment, C: Dataset include crowded scenes (≥ 3 objects),
OC: Dataset include Partially Occluded Objects or Completely Occluded Objects.

References

1. Glasl, H.; Schreiber, D.; Viertl, N.; Veigl, S.; and Fernandez, G. (2008). Video based traffic congestion prediction on an embedded system. *11th International IEEE Conference on Intelligent Transportation Systems, ITSC 2008*, 950-955.
2. Chen, J.; Li, Y.; and Liu, D. (2009). The computer aided emergency management system for highway tunnels. *International Conference on Measuring Technology and Mechatronics Automation, ICMTMA 09*, 903-906.
3. Zengqiang, M.; Guosheng, G.; Wanmin, S.; and Yan, Y. (2008). Wireless monitoring system of vehicle overspeed on freeway based on GPRS. *27th Chinese Control Conference, CCC 2008,* 550-553.
4. Wu, X.; Ou, Y.; Qian, H.; and Xu, Y. (2005). A detection system for human abnormal behaviour. *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2005*, 1204- 1208.
5. Xiaoyan, D.; Ya, Z.; Wei, W.; Zhuo, W.; and Zhihua, W. (2010). A robust multiple object tracking algorithm under highly occlusion. *Seventh International Conference on CGIV*, 5-8.
6. You, W.; Jiang, H.; and Li, Z.-N. (2009). Real-time multiple object tracking in smart environments. *IEEE International Conference on ROBIO*, 818- 823.
7. Cucchiara, R.; Grana, C.; Patri, A.; Tardini, G.; and Vezzani, R. (2004). Using computer vision techniques for dangerous situation detection in domotic applications. *Proceedings IEE Workshop on Intelligent Distributed Surveillance Systems, London*, 1-5.
8. Micheloni, C.; Foresti, G.L.; and Snidaro, L. (2003). A co-operative multicamera system for video-surveillance of parking lots. *Intelligent Distributed Surveillance Systems Symposium by the IEE, London*, 21-24.
9. Huwer, S.; and Niemann, H. (2000). Adaptive change detection for real-time surveillance applications. *Proceedings Third IEEE International Workshop on Visual Surveillance*, 37-46.
10. Christof, R.; Olaf, M.; and Harald, K. (1995). Adaptive background estimation and foreground detection using Kalman-filtering. *Proceedings of the International Conference on recent Advances in Mechatronics, ICRAM 1995*, 193-199.
11. Gao, D.; Zhou, J.; and Xin, L. (2001). A novel algorithm of adaptive background estimation. *Proceedings 2001 International Conference on Image Processing*, 2, 395-398.
12. Grimson W.E.L.; Stauffer C.; Romano R.; and Lee, L. (1998). Using adaptive tracking to classify and monitor activities in a site. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 22-31.
13. Al Najjar, M.; Ghosh, S.; and Bayoumi, M. (2009) A hybrid adaptive scheme based on selective Gaussian modeling for real-time object detection. *IEEE International Symposium on Circuits and Systems*, 936-939.
14. http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/
15. Stauffer, C.; and Grimson, W.E.L. (1999). Adaptive background mixture models for real-time tracking. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2, 246-252.

16. Wren, C.; Azarbayejani, A.; Darrell, T.; and Pentland, A. (1997). Pfinder: real-time tracking of the human body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7), 780-785.
17. Tiehan Lv.; Ozer, B.; and Wolf, W. (2004). A real-time background subtraction method with camera motion compensation. *IEEE Proceedings of International Conference on Multimedia and Exhibition*, 1, 331-334.
18. Lipton, A.J.; Fujiyoshi, H.; and Patil, R.S. (1998). Moving target classification and tracking from real-time video. *Proceedings 4th IEEE Workshop on Applications of Computer Vision*, 8-14.
19. Rahim, H.A.; Sheikh, U.U.; Ahmad, R.B.; Zain, A.S.M.; and Ariffin, W.N.F.W. (2010). Vehicle speed detection using frame differencing for smart surveillance system. *10th International Conference on Information Sciences Signal Processing and their Applications (ISSPA)*, 630-633.
20. Fujiyoshi, H.; Komura, T.; Eguchi, I.; and Kayama, K. (2006). Road Observation and Information Providing System for Supporting Mobility of Pedestrian. *IEEE International Conference on Computer Vision Systems*, 37- 44.
21. Zhang, T.; Liu, Z.; Lian, X.; and Wang, X. (2010). Study on moving-objects detection technique in video surveillance system. *IEEE Conference on Chinese Control and Decision Conference (CCDC),* 2375- 2380.
22. Yoo, Y.; and Park, T.-S. (2008). A moving object detection algorithm for smart cameras. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1-8.
23. Li, L.; and Leung, M.K.H. (2001). Fusion of two different motion cues for intelligent video surveillance. *Proceedings of IEEE International Conference on Electrical and Electronic Technology*, *TENCON*, 1, 341-344.
24. Li, J.; Nikolov, S.G.; Scott-Samuel, N.E.; and Benton, C.P. (2006). Reliable real-time optical flow estimation for surveillance applications. *The Institution of Engineering and Technology Conference on Crime and Security*, 402 - 407.
25. Denman, S.; Fookes, C.; and Sridharan, S. (2009). Improved simultaneous computation of motion detection and optical flow for object tracking. *Digital Image Computing: Techniques and Applications, DICTA '09*, 175-182.
26. Andrade, E.L.; Blunsden, S.; and Fisher, R.B. (2005). Characterisation of optical flow anomalies in pedestrian traffic. *The IEE International Symposium on Imaging for Crime Detection and Prevention*, 73- 78.
27. Kemouche, M.S.; and Aouf, N. (2009). A Gaussian mixture based optical flow modeling for object detection. *3rd International Conference on Crime Detection and Prevention*, 1-6.
28. Zhang, H.-Y. (2010). Multiple moving objects detection and tracking based on Optical flow in polar-log images. *Proceedings of the Ninth International Conference on Machine Learning and Cybernetics*, 3, 1577-1582.
29. Li, L.; Huang, W.; Gu, I.Y.-H.; Luo, R.; and Tian, Q. (2008). An efficient sequential approach to tracking multiple objects through crowds for real-time intelligent CCTV systems. *IEEE Transactions on SMC*, 38(5), 1254-1269.
30. Yilmaz, A.; Javed, O.; and Shah, M. (2006). Object tracking: A survey. *ACM Computing Surveys*, 38(4), 1-45.

31. Fang, Y.; Wang, H.; Mao, S.; and Wu, X. (2007). Multi-object tracking based on region corresponding and improved color-histogram matching. *IEEE International Symposium on Signal Processing and Information Technology*, 1-4.
32. Luo, X.; and Bhandarkar, S.M. (2005). Multiple object tracking using elastic matching. *IEEE Conference on AVSS*, 123-128.
33. Chen, T. (2009). Object tracking based on active contour model by neural fuzzy network. *IITA International Conference on Control*, *Automation and Systems Engineering*, 570-574.
34. Baumberg, A.M.; and Hogg, D.C. (1994). An efficient method for contour tracking using active shape models. *IEEE Proceedings Workshop on Motion of Non-Rigid and Articulated Objects*, 194-199.
35. Chen, Y.; Rui, Y.; and Huang, T.S. (2006). Multicue HMM-UKF for real-time contour tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(9), 1525-1529.
36. Beymer, D.; McLauchlan, P.; Coifman, B.; and Malik, J. (1997). A real-time computer vision system for measuring traffic parameters. *IEEE Proceedings Computer Society Conference on Computer Vision and Pattern Recognition*, 495-501.
37. Choi, J.-Y.; Sung, K.-S.; and Yang, Y.-K. (2007). Multiple vehicles detection and tracking based on scale invariant feature transform. *IEEE Conference on Intelligent Transportation Systems*, 528-533.
38. Hu, W.; Tan, T.; Wang, L.; and Maybank S. (2004). A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 34(3), 334-352.
39. Lepetit, V.; and Fua, P. (2005). Monocular model-based 3D tracking of rigid objects: A survey. *Computer Graphics and Vision*, 1(1), 1-89.
40. Conaire, C.O.; O'Connor, N.E.; Cooke, E.; and Smeaton, A.F. (2006). Multispectral object segmentation and retrieval in surveillance video. *IEEE International Conference on Image Processing*, 2381-2384.
41. Dan, J.; and Yuan, Y. (2007). A multi-object motion-tracking method for video surveillance. *IEEE Eighth ACIS International Conference on SNPD*, 1, 402-405.
42. Lin, C.-T.; Huang, Y.-C.; Mei, T.-W.; Pu, H.-C.; and Hong, C.-T. (2006). Multi-objects tracking system using adaptive background reconstruction technique and its application to traffic parameters extraction. *IEEE International Conference on SMC*, 3, 2057-2062.
43. Li, L.; Huang, W.; Gu, I.Y.H.; Leman, K.; and Tian, Q. (2003). Principal color representation for tracking persons. *IEEE International Conference on SMC*, 1, 1007-1012.
44. Li, L.; Huang, W.; Gu, I.Y.-H.; Luo, R.; and Tian, Q. (2008). An efficient sequential approach to tracking multiple objects through crowds for real-time intelligent CCTV systems. *IEEE Transactions on SMC*, 38(5), 1254 -1269.
45. Polat, E.; Yeasin, M; and Sharma, R. (2002). Multiple complex object tracking using a combined technique. *IEEE Proceedings 16th International Conference on Pattern Recognition*, 2, 717-720.

46. Robert, K. (2009). Night-time traffic surveillance: A robust framework for multi-vehicle detection, classification and tracking. *Sixth IEEE International Conference on AVSS,* 1-6.
47. Cheng, H.Y.; and Hwang, J.N. (2007). Multiple-target tracking for crossroad traffic utilizing modified probabilistic data association. *IEEE International Conference on ICASSP*, 1, 921-924.
48. Wang, H.; Liu, C.; Xu, L.; Tang, M.; and Wu, X. (2008). Multiple feature fusion for tracking of moving objects in video surveillance. *IEEE International Conference on CIS*, 1, 554-559.
49. Yang, C.; Duraiswami, R.; and Davis, L. (2005). Fast multiple object tracking via a hierarchical particle filter. *Tenth IEEE International Conference on ICCV*, 1, 212 -219.
50. Xiong, T.; and Debrunner. C. (2004). Stochastic car tracking with line and color based features. *IEEE Transactions on Intelligent Transportation Systems*, 5(4), 324-328.
51. Sarkka, S.; Vehtari, A.; and Lampinen, J. (2007). Rao-Blackwellized particle filter for multiple target tracking. *Inf. Fusion*, 8(1), 2-15.
52. Zhai, Y.; and Yeary, M. (2007). An intelligent video surveillance system based on multiple model particle filtering. *IEEE Proceedings on IMTC 2008*, 254-258.
53. Ryu, H.R.; and Huber, M. (2007). A particle filter approach for multi-target tracking. *IEEE/RSJ International Conference on IROS*, 2753 -2760.
54. Pham, N.T.; Huang, W.; and Ong, S.H. (2007). Tracking multiple objects using probability hypothesis density filter and color measurements. *IEEE International Conference on Multimedia and Expo*, 1511-1514.
55. Maggio, E.; Piccardo, E.; Regazzoni, C.; and Cavallaro, A. (2007). Particle PHD filtering for multi-target visual tracking. *IEEE International Conference on ICASSP*, 1, 1101-1104.
56. Yu, T.; Wu, Y.; Krahnstoever, N.O.; and Tu, P.H. (2008). Distributed data association and filtering for multiple target tracking. *IEEE Conference on CVPR*, 1-8.
57. Youssef, S.M.; Hamza, M.A.; and Fayed, A.F. (2010). Detection and tracking of multiple moving objects with occlusion in smart video surveillance systems. *5th IEEE International Conference on Intelligent Systems* (*IS*), 120- 125.