Occlusion and Abandoned Object Detection for Surveillance Applications

M. Chitra
RVS college of Engineering
and Technology
Karaikal, India

M.Kalaiselvi Geetha
Annamalai University
Chidambaram, India

L.Menaka
RVS college of Engineering
and Technology
Karaikal, India

Abstract: Object detection is an important step in any video analysis. Difficulties of the object detection are finding hidden objects and finding unrecognized objects. Although many algorithms have been developed to avoid them as outliers, occlusion boundaries could potentially provide useful information about the scene's structure and composition. A novel framework for blob based occluded object detection is proposed. A technique that can be used to detect occlusion is presented. It detects and tracks the occluded objects in video sequences captured by a fixed camera in crowded environment with occlusions. Initially the background subtraction is modeled by a Mixture of Gaussians technique (MOG). Pedestrians are detected using the pedestrian detector by computing the Histogram of Oriented Gradients descriptors (HOG), using a linear Support Vector Machine (SVM) as the classifier. In this work, a recognition and tracking system is built to detect the abandoned objects in the public transportation area such as train stations, airports etc. Several experiments were conducted to demonstrate the effectiveness of the proposed approach. The results show the robustness and effectiveness of the proposed method.

Keyword: Occlusion, Histograms of Oriented Gradients descriptors, Support Vector Machine, mixture of Gaussians techniques, Blob, abandoned object.

1. INTRODUCTION

Surveillance is monitoring the behavior, activities for the purpose of influencing, managing, directing, or protecting. One of the main advantages of video surveillance is that it could be used both as a preventive mechanism and a forensic tool for crimes. Unfortunately, constant surveillance of public domains is challenging and labor-intensive. Therefore, improved automated abandoned object detection and occlusion detection systems are increasingly in demand.

Robustness of object detection is affected by occlusion in the presence of multiple objects. Occlusion is the hiding of an object by another during multiple object tracking. In video sequences Fig 1(a), occlusions create several challenges to the tracking algorithm. The task of reliable detection and tracking of multiple objects becomes highly complex for crowded scenarios. The detection of suspicious object is one of the most important task in video surveillance. Suspicious objects are generally unattended packages left in public places. This work aims to detect occlusions and abandoned object detection which captures most of the researches working in video surveillance. Object detection in monocular image sequences still suffers from a lack of robustness due to temporary occlusions, objects crossing and changing lighting conditions.

Abandoned objects are detected by difference image. An example for occluded scene and abandoned object/baggage is shown in Fig. 1(b).

(a). Occluded frame   (b). Abandoned object

Figure. 1 Occluded and abandoned frames

The rest of the paper is arranged as follows: Section 2 discusses a few approaches explored by other researchers to solve the problem. Section 3 explains the technical details of our method, Section 4. Section 5 wraps up the paper with an experimental result and conclusion of this paper.

2. RELATED WORKS

Although many algorithms have been proposed in the literature, the problem of multiple interacting objects tracking in complex scene is still open. Methods to solve the occlusion problem in multiple interacting objects tracking have been previously presented in [8, 12, 15, 16]. Sensor based occlusion tracking methods [8, 18] cannot handle the difference between moving and nonmoving obstacles. In general, solutions to both pedestrian and vehicle occlusion problems extend from spatial to temporal domains can be broadly classified into many types. Some types like active contour –based [2], stereo vision-based, region based, model based [16], feature based [12] are also seen in the literature. Object tracking in [17] overcome occlusion by fusing multiple camera inputs, but it cannot handle complete occlusion.

Several methods using color, texture and motion [13, 14, 19]. [14] Performs robust tracking that deals some instances of occlusion but it cannot handle illumination changes.

Histograms of Oriented Gradients (HOG) and Local Binary Pattern (LBP) [9] are used to detect partial occlusion but it cannot handle the articulated deformation of people. Multi-cue model widely used in tracking and detection systems [14] has achieved considerable success in object tracking because of its simplicity and robustness. However, color features cannot give good performance when an object and its background have the similar colors. Contour-based object detection and recognition method proposed in [2, 4]. [7,10] uses filter based detection and tracking.

Most of the proposed techniques for abandoned object detection rely on tracking information [24, 25] to detect drop-off events, while fusing information from multiple cameras. As stated by Porikli [26], these methods are not well suited to complex environments like scenes involving crowds and large amounts of occlusion. Stauffer and Grimson [27] present an event detection module that classifies objects, including abandoned objects, using a neural network, but is limited to detecting only one abandoned object at a time. In [28] a multi-camera video surveillance system is proposed to detect the owner of each abandoned object is determined and tracked using distance and time constraints and multiple cameras with overlapping fields of view are exploited to cope with occlusion of various types.

3. OCCLUSION AND ABANDONED OBJECT DETECTION

3.1 Occlusion Detection

Manual surveillance requires the system to be monitored by personnel and is expensive and inconvenient. Hence, efforts have been put into automatic surveillance. One of the main components of automatic video surveillance is detecting occluded objects. Fig.2 shows the proposed system architecture for occlusion detection.

Occlusion is the main cause for performance degradation in surveillance systems. Under occlusion, the objects will become overlapped and may be found moving together in a scene. Occlusion can be of three types. Self occlusion, inter-object occlusion, background occlusion. The goal of our work is to develop a general framework to detect and track objects with persistent and occasional heavy occlusion. Background subtraction is carried out by a mixture of Gaussians techniques (MOG) is in Fig.3. Blob detection for detecting occluded object is carried out in OpenCV. Multiple occluded object detection involves computing the Histogram of Oriented Gradients descriptors (HOG) and linear Support Vector Machine (SVM) as a classifier.

3.1.1 Mixture of Gaussian Algorithm(MOG)

Pixelprocesses–

$$\{X_1,\ldots\ldots,X_T\} = \{I(x_0, y_0, i): 1 \le i \le t\},$$

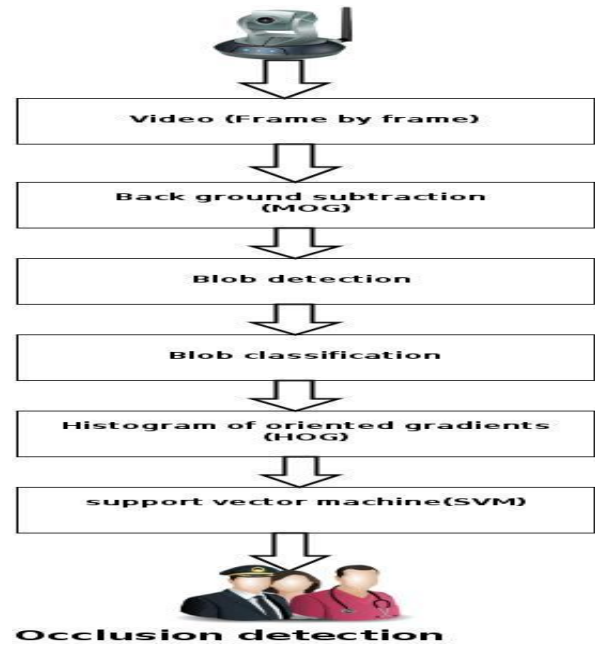Where, I- image sequence, t- time, {x0,y0}- the history of pixels.
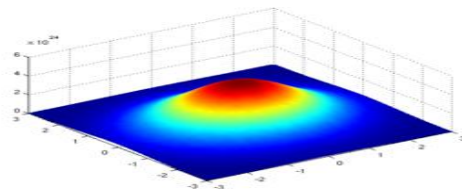

Figure. 2 Occlusion detection

Figure. 3 Mixture of Gaussians algorithms

Model the values of a particular pixel as a mixture of Gaussians. We determine which Gaussians may correspond to background colors-Based on the persistence and the variance of each of the Gaussians [23]. Pixel values that do not fit the background distributions are considered foreground until there is a Gaussian that includes them. Update the Gaussians. Pixel values that do not match one of the pixel's "background" Gaussians are grouped using connected components.

Background modeling - constructs a reference image representing the background.

Threshold selection- determines appropriate threshold values used in the subtraction operation to obtain a desired detection rate.

Pixel classification - classifies the type of a given pixel, i.e., the pixel is the part of background (including ordinary background and shaded background), or it is a moving object.
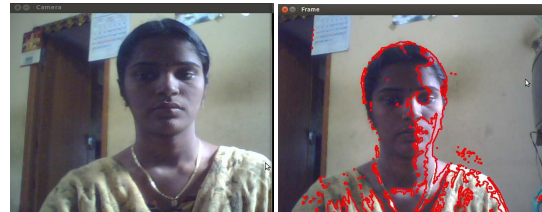

Figure. 4 Background subtraction using MOG

3.1.2 Blob Detection

Once the background subtraction is done with MOG, each foreground object is detected as a blob. Each module is applied on a bi-level image obtained though a mixture of Gaussian background subtraction and some basic filtering performed. Using OpenCV [22]. A tracked blob is considered to be occluded if its major region is covered by foreground and it should continue to be tracked if either it is occluded or its area and centroid is matched with any of the blobs. When two objects pass close to each other, they are detected as a single blob. Fig.5 shows the blob detection.

Often, one object will become occluded by the other one. One of the challenging problems is to maintain correct labeling of each object after they split again. Fig.6 demonstrates the blob merging and splitting at the time of occlusion.
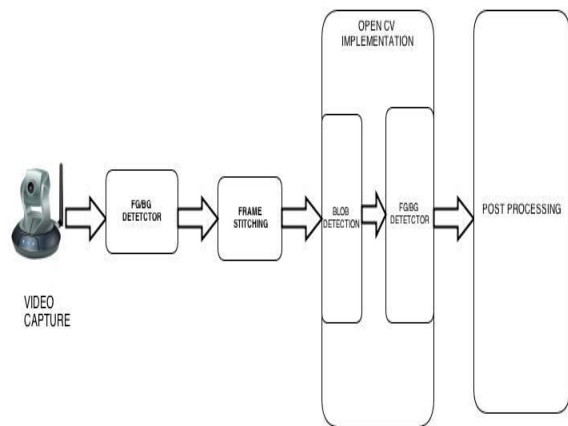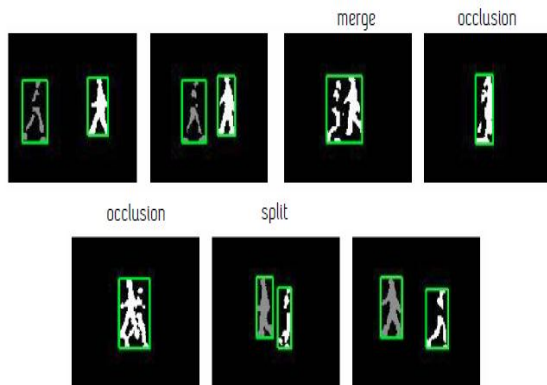

Figure. 5 Blob Detection and Tracking

Figure. 6 Blob Merging and splitting at the time of occlusion

3.1.3 Histogram of Oriented Gradients descriptors (HOG)

Histograms of Oriented Gradients (HOG) are feature descriptors used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to that of edge orientation histograms, but differs in that it is computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improved accuracy. Fig.7 shows orientation of each gradient sample rotated relative to key point orientation.

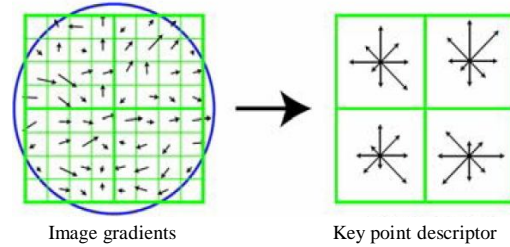Image gradients   Key point descriptor
Figure. 7 Histograms of Oriented Gradients

For normalization process all blocks are typically overlapped and rotation invariance. It is typically combined with SVM classifier.

3.1.4 SVM Classifier

The Support Vector Machine classifier is a binary classifier which looks for an optimal hyper plane as a decision function. Once trained on images containing some particular object, the SVM [20] classifier can make decisions regarding pedestrians. SVM is based on the principle of structural risk minimization. For linearly separable data, SVM finds the separating hyper plane which separates the data within the largest margin. For, linearly inseparable data, it maps the data in the input space into high dimension feature space $x \in \mathbb{R}^l \longrightarrow \Phi(x) \in \mathbb{R}^H$ With kernel function $\Phi(x)$, to find the separating hyper plane. SVM was originally developed for two class classification problems. The N class classification problem can be solved using N SVMs. Each SVM separates a single class from all the remaining classes (One-vs.-rest approach). Given a set of frames corresponding to N classed for training, N SVMs are trained. Each SVM is trained to distinguish a class and other classes in the training set. During testing, the class label y of a class x can be determined using:

$$y = \begin{cases} n, & \text{if } d_n(x) + t > 0 \\ 0, & \text{if } d_n(x) + t \leq 0 \end{cases}$$

Where, $d_n(x) = \max\{d_i(x)\}^N_{i=1}$, and $d_i(x)$ is the distance from x to the SVM hyper plane corresponding to frame i, the Classification threshold is t, and the class label y=0 stands for unknown.
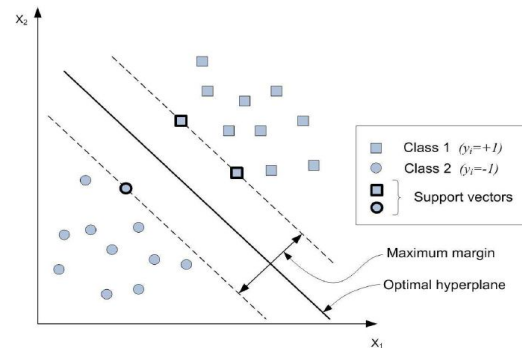

Figure. 8 Blob Optimal separating hyper plane and margin for a two dimensional feature space

3.2 Abandoned Object Detection

The detection of suspicious (dangerous) object is one of the most important task in video surveillance. An abandoned object not belonging to the background scene and remaining in the same position for a long time. Our system for abandoned object detection was designed to assist operators surveilling indoor environments, such as airports, railway, or metro stations etc. Fig 9 shows the abandoned object detection.
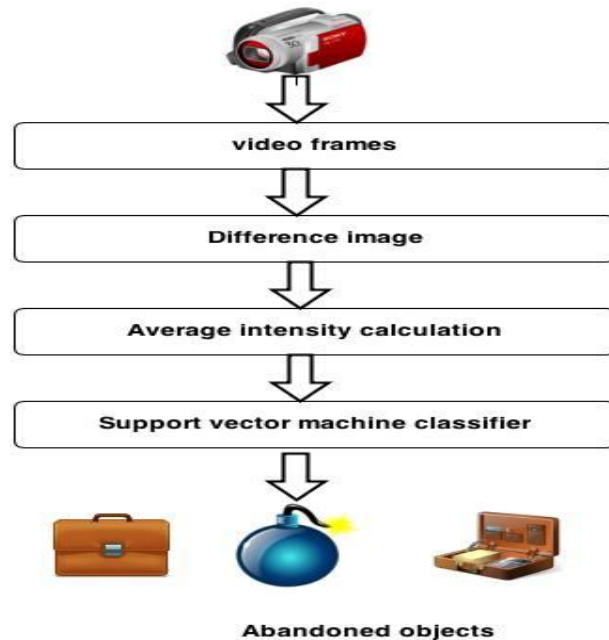

Figure. 9 Abandoned Object Detection

3.2.1 *Difference Image*

The difference image is defined as the residual image that results from subtracting the scenes at two consecutive time instants from each other. Difference images capture target motion in the scene and can be used for target tracking. Fig. 10 illustrates the basic idea behind difference images. Fig. 10(a), 10(b) show a scene with moving targets at two consecutive time instants. Fig. 10(c) shows the difference image resulting from subtracting the two image frames.
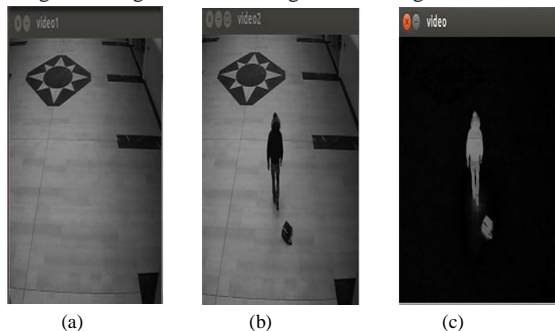

(a)   (b)   (c)

Figure. 10 (a). Image x_1 at time instance t_1 (b). Image x_2 at time instance t_2, and (c). Difference image.

4. EXPERIMENTAL RESULTS

Experiments were carried out on a PC with Intel IV cores 2.67GHz processor with C++ and the OpenCV library. The above approaches are implemented and experiments are performed on video clips freely available and PETS 2006 .The detection rate of proposed methods are shown in Table 1. The video data are processed at 20fps, and 320x240 resolution. We present some snapshots for occlusion detection and abandoned object detection in Fig.13, 14.

4.1 Occlusion detection

4.1.1 *Occlusion degree and detection rate*

The occlusion degree is calculated by the number of moving pedestrians and moving vehicles present in the frame with occlusion. For example if a video sequence contains 10 pedestrians, in which 5 pedestrians are occluded with each other, then the calculated occlusion degree is 50 %. More than 10 persons is considered as high, more than 5 as medium less than 5 as low occlusion degree. The occlusion rate is compared with the ground truth information.Fig.11 represents the graphical representation of occlusion detection.

4.1.2 *Ground truth method*

The ground-truth data allows us to do a quantitative comparison between occlusion detection method and the facts that are confirmed in an actual field. The ground truth method calculation is carried out manually.

Table 1. Detection rate of proposed method

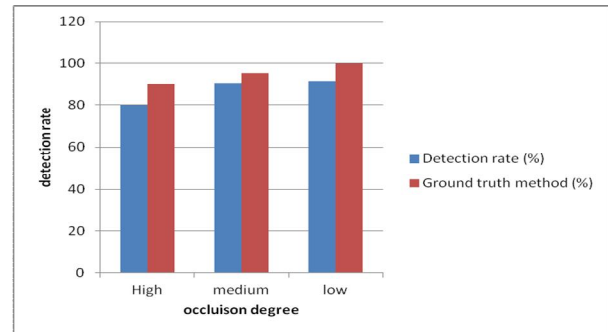| Occlusion degree (%)	High	medium	low
Detection rate	80	90.3	91.2
Ground truth method	90	95	100

Figure. 11 Graphical reprsentation of our detection rate.

4.2 Abandoned Object Detection

4.2.1 *Intensity Calculation of Difference Image*

The intensity of the difference images are taken as an input to the SVM classifier. Fig. 12 shows the graphical representation of difference images intensity values of 346 difference images from the abandoned object dataset.

Table 2. Data set for abandoned object detection

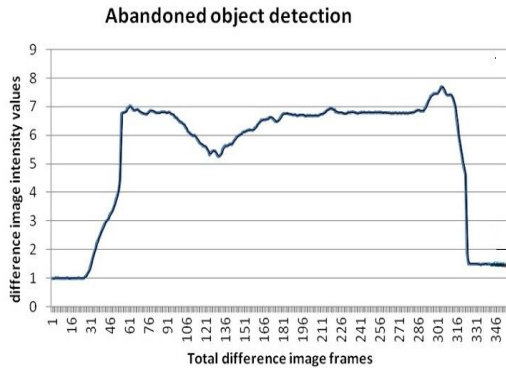| S.NO	Dataset or video name	Video length
1	Abandoned object	59
2	Video 1	30
3	PETS 2006	60

Figure. 12 Intensity values of the difference image in abandoned

object detection

4.2.2 Results with SVM Classifier

The Support Vector Machine classifier is a binary classifier which looks for an optimal hyper plane as a decision function. Once trained on images containing some particular object, the SVM classifier can make decisions regarding the presence of an object, such as a human being. Classification results using quadratic discriminant analysis were found to be quite accurate. A training set of 60 cases, where 30 of them belong to one class - bag, and the rest belong to the second class - non-bag(people) is used. Using a test set of 50 cases, gave us an accuracy of 91%. Fig. 28 shows the SVM training and testing.
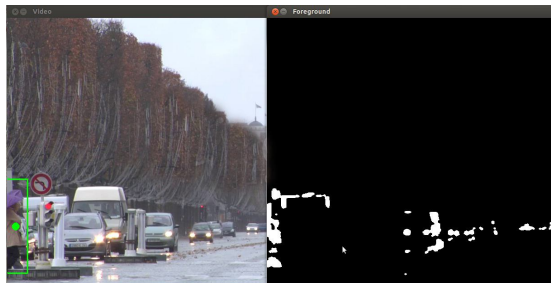

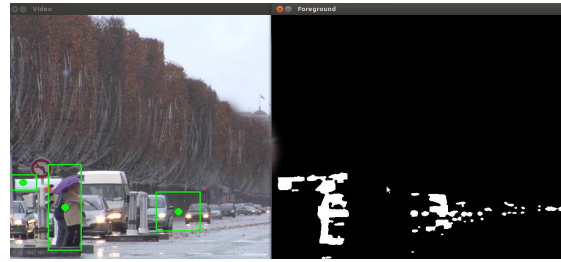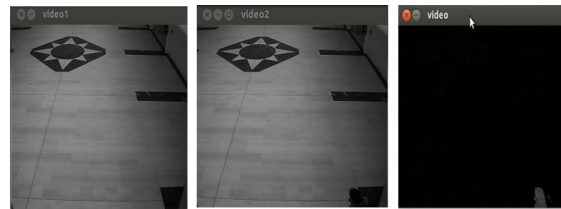Figure. 13   (a). Occlusion detetion in frame 1 at video 4

Figure. 13   (b). Occlusion detetion in frame 1 at video 10
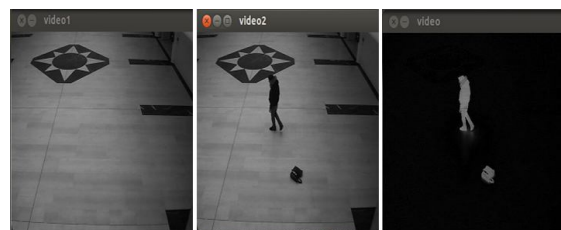
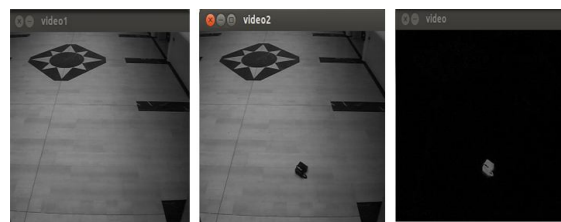Figure. 13   Occlusion detection in random frame number.

(a) Frame differencing done with first and 4th frame

(b) Frame differencing done with first and 12th frame

(c)Frame differencing done with first and 18th frame

(d) Frame differencing done with first and 26th frame.

Figure. 14 Abandoned object detection using frame differencing

at   frame 4, 12, 18, 26.

5.   CONCLUSION

This research work will lay a stepping stone for the further developments of the automatic object tracking system in a secured area. A real-time multiple occluded objects detection and abandoned object detection system is presented. Experiments on complex indoor and outdoor environments show that the system can deal with difficult situations such as ghosts and illumination changes. Moreover, it can track multiple objects with long-duration and complete occlusion. While the system is highly computationally cost effective and

accurate. The testing results which are based on different scenarios have proved that our approach can be successfully applied in real world surveillance applications. Future work includes extending multiple camera view occlusion detection and abandoned object detection.

6. REFERENCES

[1] Pallavi S. Bangare, Nilesh J. Uke, Sunil L. Bangare, " An Approach for Detecting Abandoned Object from Real Time Video", International Journal of Engineering Research and Applications (IJERA), Vol. 2, Issue 3, 2012.

[2] Geo Guo,Ting Tiang, Yizhou wang and Wen Gao, "Recovering missing Contours for occluded Object detection", IEEE Signal processing letters, Vol.19,No.8, Aug 2012.

[3] A.A.shafie,M.H Ali,Fadhlan Hafiz and Rosilizer M.Ali, "Smart video Surveillance System for Vehicle Detection and traffic Flow control", Journal of Engineering Science and technology, Vol. 6, No. 4, 2011.

[4] Tianyang Ma and Longin Jan Latecki, "From Partial Shape Matching through Local Deformation to Robust Global Shape Similarity for Object Detection", CPVR 2011.

[5] Manmohan Singh Rawat, Varun N.S, Vinotha S.R, "Occlusion Detection and Handling In Video Surveillance", 2011 International Conference on Future Information Technology, IPCSIT vol. 13 2011.

[6] Kung –Chung Lee,Anand Oka, Emmanual Pollakis and Lutz Lampe, " A Comparison between unscented Kalman filtering and Particle filtering for RSII-BasedTracking," Positioning Navigation and Communication (WPNC), 2010.

[7] Breitenstein, M.D., Reichlin, F., Leibe, B., Koller-Meier, E., Van Gool,L., " Robust tracking-by-detection using a detector confidence particle filter", IEEE International Conference on Computer Vision ,2009.

[8] Michael S. Darms, Paul E. Rybski, Christopher Baker, and Chris Urmson, "Obstacle Detection and Tracking for the Urban Challenge", IEEE Transaction on Intelligent Transportation systems,Vol. 10, No. 3, Sep 2009.

[9] Xiaoyu Wang, Tony X. Han, Shuicheng Yan, "An HOG-LBP Human Detector with Partial Occlusion Handling", Computer Vision, IEEE 12th International Conference, 2009.

[10] Junliang Xing, Haizhou Ai, Shihong Lao,"Multi-object tracking through occlusions by localtracklets filtering and global tracklets association with detection responses", Computer Vision and Pattern Recognition, 2009. CVPR 2009.

[11]   Dollar, P., Wojek, C. ., Schiele, B. , Perona, P. ,"Pedestrian detection: A benchmark", Computer Vision and Pattern Recognition, . CVPR 2009.

[12] Pang, C.C.C. ; Lam, W.W.L. ; Yung, N.H.C. , "A Method for Vehicle Count in the Presence of Multiple-Vehicle Occlusions in Traffic Images", IEEE Transactions on Intelligent Transportation Systems, Sept. 2007.

[13] Junqiu WANG and Yasushi YAGI, "Integrating Color and Shape-Texture Features for Adaptive Real-time Object Tracking", IEEE Trans on. Image Processing, 2007.

 [14] Valtteri Takala and Matti Pietik¨ainen, "Multi-Object Tracking Using Color, Texture and Motion", IEEE Trans on. Image Processing, 2007.

[15] Tao Yang, Stan Z.Li, Quan Pan, Jing Li, "Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes", Foundation of National Laboratory of Pattern Recognition and National Natural Science Foundation of China (60172037) , 2005.

[16] Clement Chun Cheong Pang, William Wai Leung Lam, Nelson Han Ching Yung, "A Novel method for Resolving vehicle occlusion in a monocular Traffic-Images Sequences", IEEE Transaction on Intelligent Transportation systems, Vol. 5, no. 3, sep 2004.

[17] Tao Zhao, Ram Nevatia "Tracking Multiple Humans in Crowded Environment", IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)

[18] MichaelBramberger, RomanPflugfelder, BernhardRinner, Helmet Schwabach and Bernhard Strobl, "Intelligent Traffic Video Sensor : Architecture and Applications", IEEE 2003.

[19] P.Perez, C.Hue, J.Vermaak, and M.Gangnet "Color-Based Probabilistic Tracking", A. Heyden et al. (Eds.): ECCV 2002, LNCS 2350, pp. 661–675, 2002.

[20] Navneet Dalal and Bill Triggs, "Histograms of Oriented Gradients for Human Detection",IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005.

[21] Asbjørn Berge, "INF 5300   Video Analysis Detecting motion", SINTEF.

[22] http://mateuszstankiewicz.eu/?p=189

[23] P. KaewTraKulPong, R. Bowden, "An Improved Adaptive Background Mixture Model for Realtime Tracking with Shadow Detection", In Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems, AVBS01. Sept 2001.

[24] S. Guler, J. A. Silverstein, and I. H. Pushee, "Stationary Objects in Multiple Object Tracking", IEEE International Conference on Advanced Video and Signal-Based Surveillance, London , UK , Sep 2007.

[25] J. M. del Rincn, J. E. Herrero-Jaraba, J. R. Gomez, and C. Orrite Urunuela, "Automatic left luggage detection and tracking using multi cameras," in PETS, 2006, pp. 59–66.

[26] F. Porikli, Y. Ivanov, and T. Haga, Robust Abandoned Object detection Using Dual Foregrounds, EURASIP Journal on Advances in Signal Processing, Vol. 2008, Hindawi Publishing Corporation.

[27] Stauffer, C., Grimson, W.E.L.: Learning patterns of activity using real-time tracking. IEEE Trans. Pattern Anal. Mach. Intell. (PAMI) 22(8), 747–757 2000

[28] M.D.Beynon, D.J.Van Hook, M.Seibert, A.Peacock, D.Dudgeon "Detecting Abandoned Packages in a Multi-camera Video Surveillance System" IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'03) July 21 - 22, 2003 Miami, Florida P. 221.