ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

A Novel Approach for Abnormality detection of Vehicle in Traffic Scenes.

S Karuppuchamy[1], and RK Selvakumar[2].

[1]Research Scholar, Manon Maniam Sundranr University, Thirunelveli, India.
[2]Professor & Head, Department of CSE, Agni College of Technology/Anna University, Chennai, India.

ABSTRACT

Video Surveillance is termed as the process of analyzing video sequences to identify the abnormality and unusual activities. Few video processing applications can be performed with Semi-autonomous video surveillance but requires significant human intervention. This paper, proposes a semantic context information based on object-specific context information and scene-specific context information to build an intelligent system for accurately detecting, tracking, and classifying normal and abnormal evens. Width distribution, paths, and entry/exist points are learned using semantic scene-specific context information. Classification is performed to efficiently improve the object detection, tracking and abnormalities. Experimental outcomes exhibit the effectiveness of semantic context capabilities for multiple actual-traffic scenes and the proposed paper has high precision accuracy with satisfactory abnormality detection and classification performance.
Keywords: Object Detection, Object Classification, Object Tracking, Gaussian Mixture Model(GMM), Linear Discriminant Analysis.

***Corresponding author**

ISSN: 0975-8585

INTRODUCTION

These days, computerized visual observation framework has been requested [7] and surveillance cameras are utilized as a part of open ranges, for example, airplane terminals, banks, shopping centers, and tram stations. Nonetheless, they were not ideally utilized because of the manual perception of the yield, which was costly and problematic. In this framework incorporates constant and proficient real time efficient computer vision calculations with a goal to help human administrators automatically. This is an aspiring objective to take care of surveillance issues of object discovery, classification, object tracking, and abnormality detection [11] throughout the years. In this paper, we endeavor to solve these issues by mining semantic context data.

Object detection is a fundamental process for further examination in video surveillance. Background modeling [3] [11] is a general technique used for stationary cameras, for extracting the moving pixels (foreground). If there are few objects in the scene, each connected component of the foreground (blob) usually corresponds to an object; this kind of blob is denoted as single-object. Several objects can be joined to form one big blob called multi-object. These multi-object looks like a foreground which makes a difficult process to create the appearance feature of every single object due to of the angle of the camera, shadow, and moving objects near each other. Therefore, it is difficult to classify and track the objects.

Various works have been proposed to take care of the crowd segmentation problem, where identifying a single human in a group is a difficult task. In [5] and [15], head discovery is utilized to help find the position of people. Moreover, they utilize dynamic programming to mitigate the inherent ambiguity. Since direction of movement are distinctive, their pictures will change, which makes these features infeasible. Also, objects in a group may have similar color, texture, and shape features. To take care of this issue, we propose a strategy method based on scene-specific context features, which reflect motion rules of objects, including direction of motion and size of object at a certain location.

Object classification process is used to categorize the detected objects into defined and predefined classes. In [8], a hybrid dynamic Bayesian network (HDBN) has been proposed a multiclass vehicle classification scheme that classifies a vehicle into one of four classes: sedan, pickup truck, SUV/minivan, and unknown. This recognition is difficult task due its diverse visual appearances Most research work in this similar area [4] used shape and motion information, such as area size, compactness, bounding box, and speed. Similarly different camera view angles also changes the object shapes drastically. In addition, the detected shapes may be noised by shadow or other actors. Another essential element is the appearance-based strategy [14] is to accomplish robust object classification in diverse camera viewing angles.

Due to its low resolution, viewing angles and shadow classification is not easy in video surveillance. Major issue in object classification is the reduction of labeling samples. Supervised learning algorithm plays a major role in combining both labeled and unlabeled data. A typical semi-supervised learning algorithm is the cot raining approach proposed by Blum and Mitchell [2]. Accurate classification is achieved by Co-training which classifies the foreground into a pedestrian or vehicle. Every generated feature is used to train a classifier, where the outputs are combined to extract the final classification results.

The other two important tasks in video surveillance is Object tracking and abnormality detection [1], [10], [11]. we can detect abnormal events, improve object tracking, and help guide vehicles [6] using the learned information,. Many approaches [12] have been proposed to learn motion patterns where few of them are based on trajectory analysis [12], [9], [13]. Spatial distance-based methods take only the pair wise similarities between trajectories.

Few drawbacks that are identified are:

- lack of probabilistic explanation for abnormality detection, Number of cluster required,
- high computational cost,
- approximate the true similarity.

ISSN: 0975-8585

Semantic scene models are generated for every cluster and Paths are identified by modeling the spatial extents of trajectory clusters [9], [13]. Entry and exit points are detected at the ends of paths based on the velocity distribution [13].

The proposed paper is organized as follows. The proposed framework is explained in Section II. Learning the scene-specific context information is detailed in Section III, classification of the foreground is elaborated in Section IV, and Implementation of the learned information for various tasks in video surveillance is discussed in Section V. Object detection and tracking are detailed in Section VI and VII. Detection of abnormalities in traffic scenes is described in details in Section VIII. Experimental results are reported and analyzed in Section IX. Finally, the analysis of the proposed scheme and conclusion is explained in Section X.

In this paper, Tsung-Lieh Lin 2010 [7] method is extended for sharing two color images S_1 and S_2 without pixel expansion. During decryption the two shares are stacked to reveal the first secret and the second secret is obtained by rotating the second secret at an angle of 180^0 and stacking with the first share. Secret color image is halftoned using Floyd and Steinberg error diffusion scheme and each color channels are processed separately for both the secrets for generating the shares using the three different processes a) Dividing process, b) Sticking Process and c) Camouflaging process.

PROPOSED TECHNIQUE

The framework of the proposed scheme is shown in Figure 1. Semantic context information includes object-specific context information and scene-specific context information. Object-specific context information contains information's like image coordinates, area in pixels, speed, and direction of motion, aspect ratio, and percentage occupancy. Scene-specific context information helps us to learned with the object-specific context information, and we consider four primary features: motion patterns of objects, width of object, paths, and sources/sinks. Then, the semantic context information is adopted to improve object detection, classification and tracking, and detect abnormal events.

For each foreground object, it is easy to obtain its object-specific context information by object detection and tracking. By trajectory analysis, GMM is adopted to learn object motion patterns and width distribution, and the graph cut algorithm is used to group similar motion patterns to get paths. Then, trajectories are further clustered by C-HMM. For each cluster of trajectories, entry/exit points and primary trajectories are learned by mean-shift-based multiple data mode-seeking algorithm. Based on the learned information, object classification, detection and tracking, and abnormal event detection has improved.

EXTRACTION OF OBJECT SPECIFIC CONTEXT INFORMATION

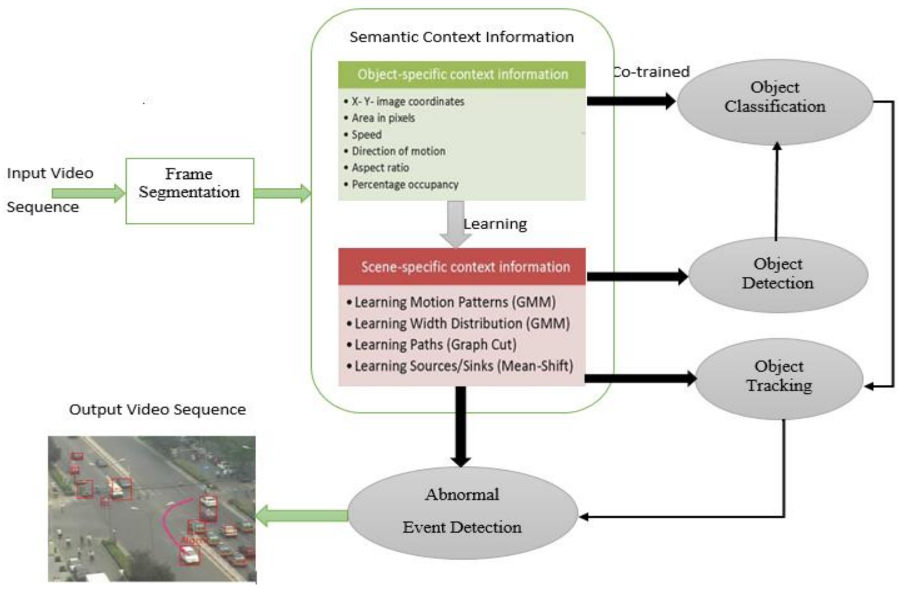

Figure 1: Proposed framework of Intelligent Video Surveillance System

Object-specific context features shows the properties of the visible objects, that can be used to distinguish objects. Object–specific context information was extracted from training features of vehicle from each frame of video, which is to be chosen to select the vehicle. The context information was extracted from each object contains-image coordinates(x and y), areas in pixel, speed, direction of motion, aspect ratio, percentage occupancy.

Co-Training Strategy

The classification process is performed as follows. Few samples are labeled in certain scenes for the purpose of training the two classifiers. The classifiers successfully classifies the unlabeled examples to identify their labels and uses these information to add those newly unlabeled examples which updates the training set continuously. This process is repeated until all the features are labeled. To ensure the accurate classification, from one training data set to another, their appearances change slowly.

The main advantages of this scheme are, first, it is a collaborative approach which uses various views of the object by which it improves each other object identification; hence a more robust classification can be achieved. Other drawback includes avoidance of mass manual labeling. Experiments were performed to demonstrate the accuracy of the classifiers. After training classifiers, performs final classification decision according to the output of the classifier with more confidence.

OBJECT CLASSIFICATION

An unsupervised learning method combined with labeled multiple features are trained using a classifier, which successfully classified the foreground into a pedestrian or vehicles, are proposed in this paper. The Classifier compares tested data and trained data, then uses the labeled data to classify the foreground.

LDA-Based Classifier

Fisher linear discriminant analysis (LDA) is used to determine the optimal direction of projection thus separating the positive and negative sample.

The projection function is given as,

$$g = w^T \quad (1)$$

Where W is product of the means of the two classes and covariance matrices. The representation of $u^t$and s^t , are as follows

$$u^t = (1/n_t)\ \sum_j^{n_t} r_j \quad (2)$$

$$S^t = (1/n_t - 1) \sum_j^{n_t} (r_j - u^r)(r_j - u^r)^T \quad (3)$$

EXTRACTION OF SCENE-SPECIFIC CONTEXT INFORMATION

Motion Patterns

Centroid of an object is used to obtain trajectory and $(x_n, y_n)$is a point in the scene image. Quadratic curve is used to describe the trajectory. For a tracked object, all points from start point to end point are collected to calculate the parameters (a, b, c) are calculated by collecting all points from start point to end point for a tracked object. The parameters are features of a trajectory. Object's moving direction is quantized into four directions. The motion patterns of every block of every trajectory, is viewed as Gaussian distributions. Scene are modeled using a mixture of Gaussian distributions for various trajectory parameters. The likelihood value of a block at time t is given as

$$P(T_t) = \sum_{i=1}^{K} w_{i,t} \times \eta\eta(T_t, u_{i,t}, \Sigma_{i,t}) \quad (4)$$

Mean:

$$u_{i,t} = (u_{i,t}^a, u_{i,t}^b, u_{i,t}^c)^T \quad (5)$$

ISSN: 0975-8585

The covariance matrices $\sum_{i,t}$ are assumed to be diagonal, and the identity matrix has proper dimensions. The K-distribution of Gaussian component are updated by computing the following equations:

$$w_{i,t} = (1-\alpha)w_{i,t-1} + \alpha(M_{i,t}) \quad (6)$$
$$u_{i,t} = (1-\rho)u_{i,t-1} + \rho T_t \quad (7)$$
$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,(t-1)}^2 + \rho(T_t - u_{i,t})^T(T_t - u_{i,t}) \quad (8)$$
$$\rho = \alpha\eta(T_t|u_{i,t}, \sigma_{i,t}) \quad (9)$$

Where $w_{i,t}$ represents the parameter weight of the i^{th} Gaussian component at time t, $\eta(T_t, u_{i,t}, \sum_{i,t})$ is the i^{th} normal distribution value of the component, $M_{i,t}$ which represents 1 for matched model or 0,α is the learning rate, and ρ is the second learning rate.

If no value matches the trajectory, then the minimum component weight can be modified using the current mean value along with its direction, high variance, and a low weight parameter. In every block, as the vehicle passes away, the motion pattern in that block determines the width of the vehicle, and this width performs the learning of the width distribution for each block.

Width Distribution for Each Block

Based on the learned patterns, the direction of movement for every block can be generated as the parameters a and b, from which the width distribution of every block can be identified.

The foreground is used to learn the width distribution w_t of block (x_0, y_0) at time 't'. In a traffic scene, the foreground can have a single-vehicle blob or a multivehicle blob. This different blobs has a significant difference in its width Therefore, the probabilistic distribution of every block width is modeled as a GMM.

The Gaussian components determines the width distribution and updates the GMM parameters using an adaptive weights in an online way just as the process of learning motion patterns of each block. The basic parameters of Gaussian component like mean w_u and variance w_σ with maximum weight are determined as the features for evaluating every block.

Paths for Scene

A block with similar motion patterns are grouped into paths. Paths can be obtained by grouping the similar patterns using a clustering algorithm. Two neighboring blocks, is expected to have similar motion patterns in common. C-HMM with Gaussian Mixture Model, is used to cluster the variable length vehicle trajectories to identify the common activity paths. Weights obtained from the Gaussian model shows the significance of the motion pattern. if the weight of the Gaussian model is greater than threshold ie weight> Th, then the Gaussian model is chosen as a primary motion pattern. Repeating the process for every block, the primary motion patterns of every block can be extracted.

Locations of the vehicles ie the entry position or exit position in the scene are called the sources and sinks. Sources and sinks should be the two ends of the path regions for each trajectory cluster. The mean-shift algorithm is used to identify these points.

OBJECT DETECTION

The intent of designing these modules is to produce a video-based surveillance system which works in real time environment. Objects can be convincingly removed from the background using subtraction process using which the blobs can be perfectly identified. Each blob, is classified using Baye's classifier and are segmented into the MV blobs and multiple SV blobs. The vehicle is the primary object which are learned using scene-specific context information to improve the detection rate. First, a classifier is adopted to classify the foreground into single-vehicle or multivehicle objects. Followed by segmenting the multivehicle blob and single-vehicle blobs are classified.

ISSN: 0975-8585

Bayes Classifier

At time 't' is the width of the scene foreground is represented as x_t. The naive Baye's classifier decides whether the foreground belongs to multivehicle (MV) or single-vehicle (SV). Bayesian decision is determined as L= $\frac{p(MV|x_t)}{p(SV|x_t)}$ (10). Generally the foreground objects that are seen are not clearly known at the initial instance. Therefore, the probability value of multi-vehicle is set as equal to the probability of single vehicle. The foreground belongs to an MV blob if probability of width of Multi-vehicle(MV) at time t is greater than a product of decision parameter (L) and probability of width of single-vehicle (SV) at time t.

To guarantee the accuracy of Bayes classifier, GMM with M components are used as means and variances to estimate the Gaussian components. In the experiment, M is set to be 2 considering that there may be two different kinds of foregrounds (MV and SV).

MV Blob Segmentation

Scene-specific context features like direction of movement, width distribution of vehicles are always stable for every fixed scene. Using such features MVs can be segmented accurately. This paper proposes a technique based on scene-specific context features to accurately predict the vehicle.

TRACKING

Tracking performs identification of objects in every frames of an video is a difficult task. This proposed scheme uses the features of objects like size, center-of-mass, bounding box, color histogram and compactness. These features are extracted to perform a comparison among objects of every consecutive frames.

Learned motion pattern is used to analyse the Motion of an object. The initial part of a motion trajectory with points are identified. The probability$P(T_0|g_m)$ of T_0under every motion pattern $g_m$is compued as$P(T_0|g_m)$ which shows the motion pattern. Motion pattern is the probability of every object is expected to move along the trajectory. The probability with highest value is selected as the most perfect value along which the object is expected to move. If the probability value is very small, then it is rejected:

$$\boldsymbol{P(T_0|g_m) = \exp(-\sum_i(a_m \times x_i^2 + b_m \times x_i + c_m - y_i)^2} \qquad (11)$$

If $g_m = (a_m, b_m, c_m)^T$ and $g_m \epsilon G_m$ then the trajectory can be fit to generate the motion pattern for every block with which the vehicle has to be passed. The motion pattern that has the maximum likelihood with trajectory can be used to improve the performance of object tracking.

As per Bayes rule, the probability is computed by $P(g_m\,|T_0) = \frac{P(T_0|g_m)P(g_m)}{\sum_m P(T_0|g_m)P(g_m)}$ (12)

Vehicle is tracked to get their trajectories for these trajectories. The proposed model predicts their future motion that is expected to be in next frames. Unusual events of moving vehicles are predicted correctly.

ABNORMAL EVENT DETECTION

Abnormal events are termed as vehicles that are violating the traffic rules by use of their trajectories are detected using the Learned motion patterns. The irregularity of a vehicle at a time period t is based on following events, that are evaluated for every individual:

- The present and earlier determined motion features.
- Features based on vehicle positions and velocities.
- The max-likelihoods of partial observations that correspond to each route modelg_m.
- Vehicle's earlier, present and expected routes.
- The maximum-likelihood thresholds $th_m^*$of each route model g_m.

ISSN: 0975-8585

Given a trajectory T, the likelyhood of T for every motion pattern $g_m$and the motion pattern which has the maximum probability with trajectory T: $m^* = argmax_m P(T|g_m)$ is computed. If the probability $P(T|g_m)$of T for that motion pattern g_m^* is less than a threshold th_m^* , then the trajectory is considered as abnormal. For every trajectories T, the probability $P(T|g_m)$ of T under motion pattern is computed. The minimum of all$P(T|g_m)$as the thresholdth_m^* is taken.

EXPERIMENTAL EVALUATION

A novel approach for video surveillance system is based on semantic context information and that can be used to improve detection, classification and tracking. Few publicly available datasets that are suitable for performing the tasks are been selected. Therefore, too many videos from real-world traffic scenes are used to evaluate the proposed methods. To prove the efficiency of motion pattern learning, the proposed algorithm is executed on various videos taken from different traffic scenes. In this paper, the experimental results are shown in for two different scenes, that include straight road and crossroad. The object–specific context information is extracted from training features of vehicle from each frame of video

Figure 2: a)Object detection b)Object Tracking c)Detection of Abnormal Event

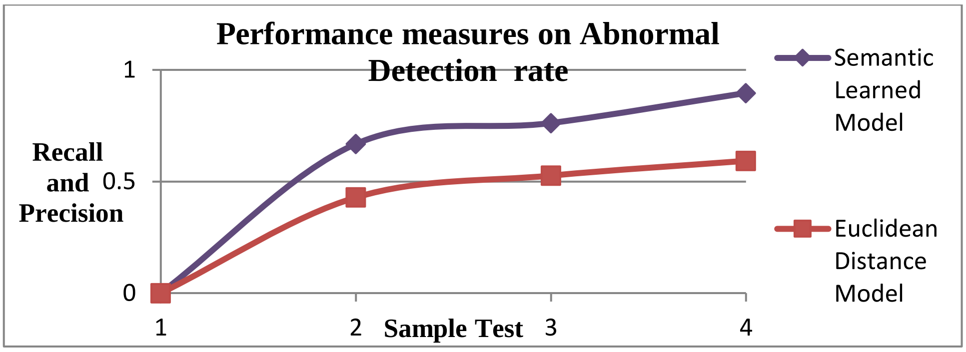

Figure 3. Abnormal Event detection Rate

Scene-specific context feature for each block are learned by adopting motion pattern and width of a vehicle, thus splitting the image into multiple blocks. As shown in Fig. 2b, a vehicle is tracked from entry point to exit point and motion pattern for each block is obtained by fitting trajectory to which the vehicle has passed. The image resolution for R and C is set to be 8 and the features are zero for blocks which the vehicle has not passed. The trajectories are clustered based on the semantic regions and their corresponding motion patterns. Figure 2a, shows that video is given as input and then vehicle is detected by bounding box based on classification with trained data. In figure 2b, green dots from start to the end point of moving vehicle shows that direction of motion and motion pattern . In figure 2c, based on the learned motion patterns, red bounding box detects abnormal events that are defined as vehicles breaking the traffic rules.

Performance is measured by Recall and Precision. Positive results are identified by Recall. Precision relates to the ability of test to identify negative results

$$Precision = \frac{True\ Positive}{(True\ Positive + False\ positive)}$$

$$Recall = \frac{True\ positive}{(\textbf{True positive} + \textbf{False Negative})}$$

Tests on i-lids parked vehicle detection dataset was done.TP is True Positive, FP is False Positive, TN is True Negative, and FN is False Negative. Scene context features of positive samples and negative samples for each scene are used to train the LDA-based classifier

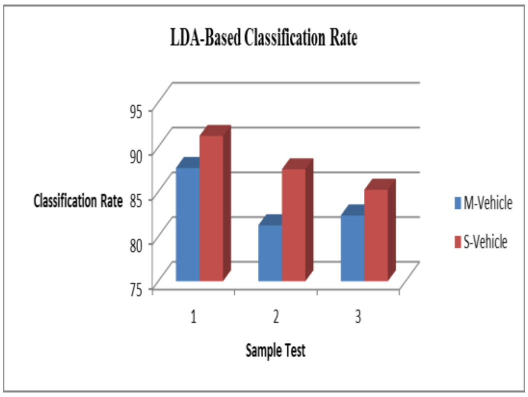

Figure 4. Performance measures on Classification rate

MV blobs are collected on i-LIDS parked vehicle detection dataset and only 91.23% blobs are correctly segmented into SV blobs.

Figure 4, shows that the performance of classification rate is measured from the graph with Sample test as x-axis and Classification accuracy as Y-axis which is measured from test 1, test2 and test3 respectively. It shows that performance of LDA-based classification of multi-vehicles (M-Vehicle) and single-vehicle (S-Vehicle) are evaluated on sample videos.

Vehicles breaking the traffic rules are detected by use of their trajectories. In figure 2c, truck turned right with high probability and this mostion is clear. Vehicle trajectories are obtained by tracking them to obtain quantitative results and model to predict their future motion in next frames. Unusual events of moving vehicles are predicted correctly.

For each vehicle class, the primary motion patterns for every block are already learned. From graph ,the system can confirm that semantic learned model has resulted in a good prediction accuracy for object behaviors. To make quantitative results that recall and precision values can be adopted to evaluate the performance from traffic scenes. These samples include 15 vehicles (class I), 11 vehicles (class II) and 8 vehicles (class III).

The semantic learned model is compared using Euclidean Distance Model. Euclidean distance model which is the baseline method used to measure the average distance between two trajectories of every cluster. This calculated distance is used as a threshold for determining the trajectories as abnormal or not. The obtained results are shown in figure 3 which demonstrates that the semantic learned method results around 90.5% accuracy. The experimental results prove that the semantic learned scheme achieves a better performance compared to other well-known methods.

CONCLUSION

This paper proposes a framework to mine semantic context information for intelligently analyzing video surveillance using traffic scenes. Learning about scene-specific context information from object-specific context information is performed initially. Then, classification of the objects is improved through combination of multiple features using a co-training framework. Information obtained using the learning process is used to improve the efficiency of object detection, tracking and abnormal events. Experimental results show the effectiveness of the object detection, classification, tracking and the abnormal effects. The co-trained framework improves the strength of various features using fusing and the samples are increased by semi supervised learning automatically. Co-trained classifier obtains better results, compared to AdaBoost classifier, and LLC classifier, has improved the identification of objects in several traffic scenes. The experimental results prove that the semantic learned scheme results a better performance when compared to other well-known methods.

REFERENCES

[1] B. Babenko, M.-H. Yang, and S. Belongie, "Robust object tracking with online multiple instance learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1619–1632, Aug. 2011.

[2] A. Blum and T. Mitchell, "Combining labeled and unlabeled data with cotraining," in *Proc. 11th Annu. Conf.Computational LearningTheory*, 1998

[3] S. Chen, J. Zhang, Y. Li, and J. Zhang, "A hierarchical model incorporating segmented regions and pixel descriptors for video background subtraction," *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 118–127, Feb.2012.

[4] R. Cutler and L. Davis, "Robust real-time periodic motion detection, analysis, and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 781–796, Aug. 2000.

[5] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 809–830, Aug. 2000.

[6] D. Herrero-Perez and H. Martinez-Barbera, "Modeling distributed transportation systems composed of flexible automated guided vehicles in flexible manufacturing systems," *IEEE Trans. Ind. Inf.*, vol. 6,no. 2, pp. 166–180, May 2010.

[7] O. Javed, K. Shafique, and M. Shah, "Automated visual surveillance in realistic scenarios," *IEEE Multimedia*, vol. 14, no. 1, pp. 30–39, Jan. –Mar. 2007

[8] M.Kafai and B. Bhanu, "Dynamic Bayesian networks for vehicle classification in video," *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 100–109,Feb. 2012.

[9] D. Makris and T. Ellis, "Automatic learning of an activity-based semantic scene model," in *Proc. IEEE Conf. Adv. Video Signal Based Surveillance*, 2003.

[10] I. Saleemi, K. Shafique, andM. Shah, "Probabilistic modeling of scene dynamics for applications in visual surveillance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 8, pp. 1472–1485, Aug. 2009

[11] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 747–757, Aug. 2000.

[12] X.Wang, K. T. Ma, G.-W. Ng, and E. Grimson, "Trajectory analysis and semantic region modelling using a nonparametric Bayesian model," in *Proc. CVPR*, 2008.

[13] X.Wang, K. Tieu, and E. Grimson, "Learning semantic scene models by trajectory analysis," in *Proc. ECCV*, 2006.

[14] L. Zhang, S. Li, X. Yuan, and S. Xiang, "Real-time object classification in video surveillance based on appearance learning," in *Proc. IEEE Int. Workshop Visual Surveillance in Conjunction With CVPR*, 2007.

[15] T. Zhao and R. Nevatia, "Bayesian human segmentation in crowded situations," in *Proc. CVPR*, 2003