2007 IEEE International Conference on
Robotics and Automation
Roma, Italy, 10-14 April 2007

Detecting Unusual Activities at Vehicular Intersections

Joaquín Salas Hugo Jiménez Joel González Juan Hurtado

Abstract— **In this work a proposal to model the activity at vehicular intersections with the aim of detecting unusual events is presented. Using the particular constraints that this kind of scenarios provide, we develop methods to detect, track, and model the activity of moving objects. Our description of activity is based on the local definition, at each pixel, of a multimodal model for the direction of motion. During operation, a particular observation is compared with the learned model. Our experiments give clear indication that the proposed scheme has a good performance in detecting such unusual events as vehicles running on red light and making forbidden turns.**

I. INTRODUCTION

According to[2]about 30% of the total road crashes occurs at vehicular intersections. Therefore the development of detection and warning technology for this kind of locations is likely to produce valuable safety benefits in terms of accident prevention. Due to the amount of information they provide, and despite some concerns about privacy[1], machine-vision systems have been seen as an appealing technology to quantify flow, measure speed, and in general to detect activity. Some algorithms based on visual info have already been developed to detect collisions at intersections [11], [3]. It has been found that detecting such events and others like bumping, passing, and jamming is quite challenging because examples of unusual events are extremely different from each other and most of the times not too frequent. Although the use of *a posteriori* video sequences may provide relevant evidence, an important aspect of the problem, as some researchers have proposed[21], is to actively use the images to detect situations that may lead to accidents.

In this work, we propose a pixel-based strategy for the detection of unusual activities. A dual layer background model is adaptively generated and updated to capture both the appearance and motion of an object. We define background modeling as the problem of separate the different elements in the scene depending on how fast they change in an image sequence. A first layer is made out from the pixel intensity level variations throughout time. To reduce the computing time we developed the method for gray scale images, although color images may offer advantages to characterize distinctive features[9] and to identify shadows[6]. In the second layer, our algorithm represents the movement orientations that may be present at each pixel using a multimodal probabilistic function. In general, an unusual activity is declared when the current direction of motion cannot be resolved under the existing model.

This work was partially supported with a grant from IPN-SIP 20061072
The authors are with CICATA-IPN Unidad Querétaro, Cerro Blanco 141, Col. Cimatario, CP 76090, Querétaro, México, `salas@ieee.org`

The detection of unusual events can be defined as a problem where one wants to classify between what is normal or common and what is not. In this sense, it has received much attention from research community. Toyama *et al.*[20] made an extensive review of the functional parts of an ideal background maintenance system while Piccardi[14] made a review of some of the main methods for background substraction.

Detecting unusual activity becomes difficult because unusual events by definition rarely occur, they may be unexpected as Zhang *et al.*[22] suggest, but at the same time they are relevant for the task. This difficulties become more significative during training.

Frequently, unusual events are modeled using Hidden Markov Models (HMMs)[5]. HMMs are perhaps the most successful framework in perceptual computing for modeling and classifying dynamic behaviors because they offer dynamic time warping, a training algorithm, and clear Bayesian semantics. Nonetheless, other possibilities that have been explored include the representation of the tracked trajectory into a binary tree structure that is used for classification[17], or the characterization of the video input as temporal templates[23].

Vehicular intersections offer a unique set of constraints, including regularity of the trajectories and predictability of the vehicular flow. Long term observation of video sequences can be used to learn the typical trajectories[10]. These trajectories can be represented with a multidimensional Gaussian distribution as in [15]. In our case, we introduce a strategy that does not require to maintain a history of all prior data points, thus making it suitable for streaming video applications. The paper is organized as follows: In §II it is presented the appearance-motion double layer used to model the scene background. Then, in §III a model to describe activity based on a probabilistic approach is introduced. Finally, in §IV we present results of the algorithm implementation on a real crossroads and conclude the paper.

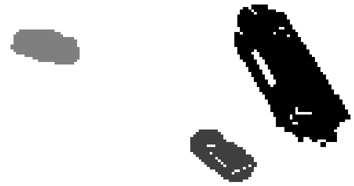
II. BACKGROUND MODEL

The background model is made out of two layers, one for appearance and another one for motion. The appearance is computed from the light intensity variations. The motion is estimated from the displacement of objects in the scene.

A. Fist layer: Appearance Model

Vehicle behavior is different inside and outside the ROI crossroads area. Inside, vehicles are always moving, while outside vehicles may be waiting for the appropriate green light. We choose the area delimited by crosswalks as our

(a) ROI



(b) Foreground objects extracted from a video frame

Fig. 1. The moving objects can be detected by subtracting the current image from the background model. The result is segmented into groups of connected pixels. This procedure is useful to both detect moving objects in a region of interest (ROI) and to update the background model considering only those regions where the variations are small.

region of interest (ROI) which can be defined as follows. Let $P = \{\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n, \mathbf{p}_{n+1}\}$ be the set of vertex points of a polygonal shaped ROI numbered counter-clockwise, where $\mathbf{p}_1 = \mathbf{p}_{n+1}$ and $\mathbf{p}_k = (x_k, y_k)$. Using these corner points, n regions can be defined such that

$$C_k(\mathbf{x}) = (y - y_{k-1})(x_k - x_{k-1}) - (x - x_{k-1})(y_k - y_{k-1}) > 0, \quad (1)$$

for $k = 2, \ldots, n+1$, is a logical predicate that divides the plane in two regions. This way, the ROI can be defined as the intersection between these regions

$$R(\mathbf{x}) = \bigcap_{k=2}^{n+1} C_k(\mathbf{x}). \quad (2)$$

$R(\mathbf{x})$ is a boolean variable that is true whenever $\mathbf{x} = (x, y)$ is inside the ROI and false otherwise.

An important processing stage includes how to obtain the initial background model[8]. The strategy that we use is computing the median of certain number of images as in [18].

Let $I(\mathbf{x}, t)$ be an image description, where $\mathbf{x}$ is a spatial position and t is a time stamp. In general, what is perceived as an image is $J(\mathbf{x}, t)$, a noisy version of $I(\mathbf{x}, t)$ given by $J(\mathbf{x}, t) = I(\mathbf{x}, t) + \delta(\mathbf{x}, t)$, where $\delta(\mathbf{x}, t)$ is assumed to be a random Gaussian variable with zero mean (i.e. Gaussian noise). We assumed that changes in illumination conditions came from smooth variations due to daylight characteristics. This assumption leaves out scenarios where illumination changes drastically from one moment to the next. In the present application, the background is supposed to be free from moving objects. Thus a single Gaussian curve can model the perceived changes in intensity. Let a Gaussian process be modeled as

$$gs(x; \mu_k, \sigma_k) = \frac{1}{\sqrt{2\pi}\sigma_k} \exp\left[-\frac{1}{2}\left(\frac{I(x) - \mu_k}{\sigma_k}\right)^2\right], \quad (3)$$

where μ_k and σ_k are respectively the mean and the standard deviation of the set of initial images. When a new observation $I(\mathbf{x}, t)$ is available, it is compared against the the parameters of the Gaussian model. If

$$\| I(x) - \mu_k \| \leq \alpha\sigma_k, \quad (4)$$

then it is assumed that the observation is likely to be produced by a perturbation of the true value similar to the one expressed by the model. Typically, α is chosen to be 3, meaning that $I(\mathbf{x})$ is within 99.73% of the cases occurring under this model. The parameters of the Gaussian are adapted as time passes by following the on-line Estimation Maximization (EM) strategy first introduced in [16]. That is,

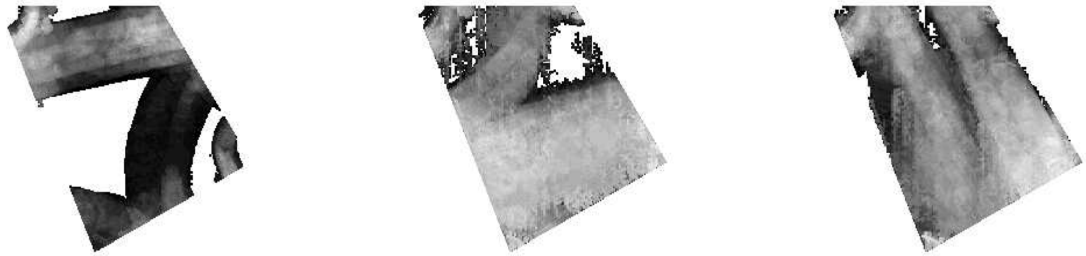
$$\mu_{t+1} = \rho\mu_t + (1 - \rho)I(\mathbf{x}, t), \quad (5)$$
$$\sigma^2_{t+1} = \rho\sigma^2_t + (1 - \rho)(I(\mathbf{x}, t) - \mu_{t+1})^2, \quad (6)$$

where $\rho \in [0, 1]$ is the learning rate.

B. Second layer: Motion Model

A second layer of the background is made out from the regular trajectories that describe moving objects in the scene. The problem of detecting where a feature A moves from one image frame to the next has many interesting facets that include objects undergoing partial or total occlusion, or being subject to complex appearance transformations. In our case, the objects are assumed to be rigid and hence, although there are some effects due to perspective and scene location, the transformations observed involves primarily rotations and translations. Furthermore, we are assuming that we can achieve a sufficiently high frame processing rate so that effectively vehicles' appearance is quite similar from frame to frame. Lucas and Kanade[12] proposed, in a milestone paper, an strategy for additive image alignment based on a Newton-Raphson type of iterative formulation. The translation of a feature between frames was computed with a steepest descend minimization strategy. In principle, a more general transformation including affine wrapping and translation could be sought. However, in practice, Shi and Tomasi[19] showed that this procedure could be numerically unstable. The procedure uses the optical flow invariant as a constraint that assumes that a gray level of an object remains equal from frame to frame. That is, let $I(\mathbf{x})$ and $G(\mathbf{x})$ be two consecutive images. It has been shown[12], [19] that the displacement $\mathbf{d}$ of a feature F can be computed using the recursive equation

$$\mathbf{d}_{k+1} = \mathbf{d}_k + Z^{-1}\mathbf{e}, \quad (7)$$



(a) (b) (c)

Fig. 2. Usual Activity Space. In (a), (b), and (c), we illustrate the number of Gaussians defined at each pixel location for the three different states in the studied scenario.

where $Z = \sum_{\mathbf{x} \in F} \begin{pmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{pmatrix}$ is the structural tensor, and $\mathbf{e} = \sum_{\mathbf{x} \in F} (I(\mathbf{x}) - G(\mathbf{x}))\, \mathbf{g}$ is a scaled version of $\mathbf{g} = (g_x, g_y)^T = \nabla I(\mathbf{x})$, the gradient. The value of Z is a good reference about how easy it is to track a feature. That is, when its eigenvalues are small the displacement is large and convergence may be difficult.

Occlusion seems to be the prime problem for robust tracking. Strategies to deal with it include the use of sub-features [4], high-level reasoning modules [21], bounding box models [3], temporal templates produced with interframe differences [13], active models [10], or multiple hypothesis[17]. In the case of this study, we do not deal explicitly with occlusion because experimentally we have made two observations. First, as it is shown in §IV, it accounts for a small portion of the problems; and second, it is common that unusual maneuvers are performed by isolated vehicles, and when it is not the case, the event is likely to be detected as an unusual activity for all the vehicles in the group.

III. Activity Space

In this work, observed activity at each pixel location is modeled with a Mixture of Gaussians (MOG) whose modes describe the main motion direction. During operation, a probabilistic measure can be assigned to a particular observation, this measure says how usual an event is. This is contrary to other approaches[10], [5] where once the trajectory of many vehicles has been accounted for, it is possible to give a unusual or usual qualification to a particular event.

A Mixture of Gaussians is calculated for each of the three possible states generated by the traffic lights. These states are next described, in one of the states (lets call it the first state) vehicles running from west to east (left to right in the images shown along the paper) and also turning to the left when driving in the same direction have the green light. In the second state the green light is for vehicles running from east to west and turning to the left when driving in the same direction. Finally the third state is when vehicles running north to south and south to north simultaneously (up-down and down-up in the images) have green light, no left turns are permitted in this state. After the third, states begin again.

A. Mixture of Gaussians

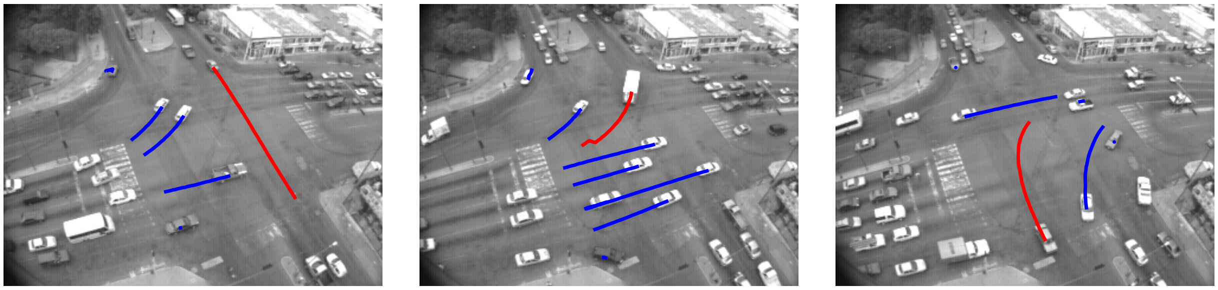
We aim to use MOG to describe the activity present at a particular pixel location as perceived from a fixed camera by a set of Gaussians. Given a set of n angular directions, $\theta_1, \ldots, \theta_n \in [0, 2\pi]$, and a family $\mathcal{F}$ of probability density functions on $\mathbb{R}$, the problem is to find the probability density $f(\theta) \in \mathcal{F}$ that is most likely to have generated the given directions. In this method, each member of the family $\mathcal{F}$ has the same general Gaussian form. Each member is distinguished by different values of a set of parameters Γ. That is[7]

$$f(\theta; \Gamma) = \sum_{k=1}^{K} p_k gs(\theta; \mu_k, \sigma_k), \tag{8}$$

where $gs(\theta; \mu_k, \sigma_k)$ is a 1-dimensional Gaussian function, as in Eq. (3), and $\Gamma = (\gamma_1, \ldots, \gamma_K) = ((p_1, \mu_1, \sigma_1),- \ldots, (p_K, \mu_K, \sigma_K))$, is a $3K$-dimensional vector containing the mixing probabilities p_k as well as the means μ_k and standard deviations σ_k of the k Gaussian functions in the mixture. When a new observation θ_t is available, it is compared against the parameters of the Gaussian models. Classification, and learning can be done as indicated in Eq. (4) through (6) respectively. After a considerable number of processed frames the MOG consists on a set of Gaussians along with the number of samples that were used to define each of them. The MOG is then pruned to eliminate Gaussians that have small support.

B. Usual Activity Space

The traffic light control at a crossroads may be seen as a deterministic machine that cycles around a number of states $S_1 \rightarrow S_2 \rightarrow \cdots \rightarrow S_n \rightarrow S_1$. At each specific state S_i, certain routes are present and others may be considered abnormal. Thus, passing on red light or making a forbidden turn may be considered abnormal because either they are happening in the wrong moment or because there were not training samples for them. Each state has a usual activity space, which is defined by an specific MOG at each pixel location. When a new state arrives the usual activity space changes in accordance. It is assumed that there is a way



(a) Running on red light. (b) Forbidden turn (too wide). (c) Forbidden turn (too wide).

Fig. 3. Some unusual events detected with our method.

to let the vision system know that a new state has arrived. This can be done by using a direct connection from the traffic light automatic controller box, for instance. Fig. 2 shows a description of the normal activity space for the three states composing the scenario studied. In Fig. 2 it is represented the number of MOG per pixel which describe the usual motion direction, the gray level indicates the number of MOG for that pixel, i.e. white zones indicate that there are no movement descriptors there, no vehicles usually pass through that image zone. A black pixel indicates a high number of MOG (movement descriptor) for that image part, i.e. a usual event. One can see that for the first state (Fig. 2, (a) vehicles running from right to left) there is a great number of MOG in the trajectory of the left turn, an appreciable but less dense number of MOG is present in the forward direction and something similar happens with right turns in the corners which are always permitted, no matter which way has green light. In the second state (Fig. 2 (b)), in which vehicles are driving from left to right, something similar happens, here the density of MOG is clearly low compared to the first state this is an indication that for example less vehicles usually take a left turn when going in this movement direction. A white area is also observable in which no vehicles passed. Finally for the third state (Fig. 2 (c)) the most interesting observation is that there is no MOG for left turns in any of the two movement directions which is coherent with the fact that no left turns are permitted by law in this state.

C. Unusual Activity Detection

Once with a model about the normal behavior of vehicles in the crossroads, it is possible to start identifying unusual events. At each pixel position, we have a MOG describing the usual directions of motions present in the training sequence. During operation, the centroid, $\mathbf{x}$, corresponding to a particular moving object give us the location to examine. The displacement computed from tracking the vehicle give us the direction of motion that is compared against the MOG.

Let $X = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\}$ be the ordered set of pixel points in the vehicle's trajectory. The probability of observing this particular trajectory is

$$\begin{aligned} p(\mathbf{x}_1, \ldots, \mathbf{x}_n) &= p(\mathbf{x}_n|\mathbf{x}_{n-1}, \ldots, \mathbf{x}_1) \\ & \quad p(\mathbf{x}_{n-1}|\mathbf{x}_{n-2}, \ldots, \mathbf{x}_1) \\ & \quad \vdots \\ & \quad p(\mathbf{x}_2|\mathbf{x}_1)p(\mathbf{x}_1). \end{aligned} \tag{9}$$

Assuming a Markovian condition, where each observation depends solely on the last one, the expression can be rewritten as

$$\begin{aligned} p(\mathbf{x}_1, \ldots, \mathbf{x}_n) = p(\mathbf{x}_n|\mathbf{x}_{n-1})p(\mathbf{x}_{n-1}|\mathbf{x}_{n-2}) \ldots \\ \ldots p(\mathbf{x}_2|\mathbf{x}_1)p(\mathbf{x}_1). \end{aligned} \tag{10}$$

Since, $\mathbf{x}_i$ and $\mathbf{x}_{i-1}$ are dependent because the new position is the previous position plus a displacement. That is, $\mathbf{x}_i = \mathbf{x}_{i-1} + a_{i-1}\mathbf{u}_{i-1}$, where a is a constant, related to the vehicle's speed, and $\mathbf{u}_{i-1}$ a unitary vector, then $p(\mathbf{x}_i|\mathbf{x}_{i-1})$ can be written as $p(\mathbf{x}_i|\mathbf{x}_{i-1}) = p(a_{i-1}\mathbf{u}_{i-1}|\mathbf{x}_{i-1})$. This way a possible measure for the likelihood of the trajectory X could be

$$\begin{aligned} L(\mathbf{x}_1, \ldots, \mathbf{x}_n) &= p(\mathbf{u}_{n-1}|\mathbf{x}_{n-1})p(\mathbf{u}_{n-2}|\mathbf{x}_{n-2}) \ldots \\ & \quad \ldots p(\mathbf{u}_1|\mathbf{x}_1) \\ &= \prod_{i=1}^{n-1} p(\mathbf{u}_i|\mathbf{x}_i). \end{aligned} \tag{11}$$

The previous condition express temporal and spatial coherence of motion and can be part of the information carried out by the blob being tracked.

IV. Experimental Results

We have programmed the algorithms to execute the method previously described using Matlab (TM). For our experiments, we used a sequence of 20,000 images, with a 320×240 resolution, the camera is located on a 28 m height tower in one of the corners of a vehicular crossroads. The traffic lights have the three states described before. The experimental image sequence comprehends 12 complete cycles through these three states. We used the first 6 cycles for training and the rest for testing. Each training cycle sequence was divided into subsequences corresponding to each one of the three different states.

TABLE I

EXPERIMENT RESULTS. **1)** SUCCESS OF VEHICLE'S TRACKING. RESULTS HAVE BEEN HAND COMPUTED AND COMPARED WITH THE ALGORITHM RESULTS. **2)** THE PERCENTAGE OF UNUSUAL EVENTS DETECTED REFERRED TO TRACKED VEHICLES ONLY; THESE HAVE BEEN DIVIDED IN TWO POSSIBLE CASES: RED LIGHT RUNNING (COLUMN 1) AND FORBIDDEN TURN (COLUMN 2); PERCENTAGES ARE SHOWN IN PARENTHESIS.

1)

State	#Vehicles	Untracked	% Error
1	262	40	15.3
2	286	33	11.5
3	176	18	10.2
Total	724	91	12.6

2)

State	#Red Light(%)	# Forbidden (%)	Total(%)
1	2(3.4)	9(0.8)	11(4.2)
2	5(1.8)	2(0.7)	7(2.5)
3	16(0)	0(0)	16(9.1)
Total	30(4.1)	4(0.6)	34(4.7)

Next, the subsequences corresponding to the same state were processed to obtain the usual event space for each particular state. So as a result of the training phase we have (a) a region of interest, (b) an initial model of the background, and (c) a description of the usual event space for each of the individual states that are part of the cycle.

The first cycle, in both the training and testing sequence, was used for background initialization. We computed the most frequent gray level for each pixel in the image. Then, a Gaussian model was used to try to adjust the gray level variations observed along the cycle. When the variations could be interpreted by the Gaussian model, the sample was used for learning and assigned to the background. Otherwise, it was assumed that a foreground object was occluding the background.

During operation, the usual event space is loaded simultaneously with the image that contains a traffic light change (change of state). The appropriate event space is then accesible and the execution continued. Next, the observed events are compared to what is considered normal for that particular state. The probabilities along the trajectory are evaluated and those with low probability value are considered unusual events.

Results are summarized in Table I. During testing, we manually counted 724 vehicles. About 87.4% of them were successfully tracked as individual vehicles. In most cases, untracked vehicles were so close together that one of them occluded the other or the moving extraction module returned them as a single connected blob. For unusual event detection that number is significant because in such a situation, as we previously noticed, vehicles tend to be isolated and were successfully tracked in all cases. The percentage of vehicle maneuvers that were classified as unusual was considerably high, about 4.7% of them all. It is interesting to notice that most of the unusual events detected are running on red light, 4.1%. Also, in this particular experiment, the third state accounts for almost half the observed unusual events.

CONCLUSION

In this paper, it was presented a reliable method for detecting such unusual events as red-light infringements and forbidden turns. Our model is based on a dual background layer. The first one deals with appearance at the intensity level. The second one with the different moving directions present on an image sequence. The model adapts to different illumination conditions and to the states caused by the traffic-light controller. The method does not require a high-level modeling of the vehicles' trajectories since the decisions are taken at pixel level. For this particular problem, the occlusion does not represent a big problem because most of the vehicles taking part in an unusual event tend to be isolated. When they are not, the statistics may be slightly affected and the type of activity (usual or unusual) that the group of vehicles develops is going to be detected. We have exploited some constraints involving the scene own conditions, including the simplicity of the background in the region of interest, the rigidity of the objects being observed, and the regularity of the usual trajectories.

Future directions of research include the detection of speedy vehicles. This could be possible done by modeling during training the different displacements present in a given pixel location, in the same way direction of motion is represented. Also, although the algorithm is highly parallel, its computing demands does not exceed standard desktop PCs capabilities for real-time implementation. Finally, the dual layer background seems suitable for dynamic background representation and extendable to other monitoring domains such as people walking on corridors.

REFERENCES

[1] Abhi Shelat Andrew J. Blumberg, Lauren S. Keeler. Automated Traffic Enforcement Which Respects "Driver Privacy". In *IEEE Conference on Intelligent Transportation Systems*, 2005.

[2] National Highway Traffic Safety Administration Arthur Carter. Intersection Collision Avoidance. http://www-nrd.nhtsa.dot.gov/departments/nrd-01/summaries/ITS_07.html, 2006.

[3] S. Atev, H. Arumugam, O. Masoud, R. Janardan, and N. P. Papanikolopoulos. A Vision-Based Approach to Collision Prediction at Traffic Intersections. *IEEE Transactions on ITS*, 6(4):416–423, December 2005.

[4] D. J. Beymer, P. McLauchlan, B. Coifman, and J. Malik. A Real Time Computer Vision System for Measuring Traffic Parameters. In *CVPR*, pages 495–501, 1997.

[5] M. T. Chan, A. Hoogs, J. Schmiederer, and M. Petersen. Detecting Rare Events in Video using Semantic Primitives with HMM. In *ICPR*, pages IV: 150–154, 2004.

[6] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati. Detecting Moving Objects, Ghosts, and Shadows in Video Streams. *IEEE Transactions on PAMI*, 25(10):1337 – 1342, 2003.

[7] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. Wiley, 2001.

[8] D. Gutchess, M. Trajkovics, E. Cohen-Solal, D. Lyons, and A. K. Jain. A Background Model Initialization Algorithm for Video Surveillance. In *ICCV*, volume 1, pages 733 – 740, 2001.

[9] T. Horprasert, D. Harwood, and L. S. Davis. A Robust Background Subtraction and Shadow Detection. In *Proceedings of the Fourth Asian Conference on Computer Vision*, 2000.

[10] N. Johnson and D. C. Hogg. Learning the Distribution of Object Trajectories for Event Recognition. *Image and Vision Computing*, 14(8):609–615, August 1996.

[11] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi. Traffic Monitoring and Accident Detection at Intersections. *IEEE Transactions on ITS*, 1(2):108–118, June 2000.

[12] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Image Understanding Workshop*, pages 121–130, 1981.

[13] G. Medioni, I. Cohen, F. Bremond, S. Hongeng, and R. Nevatia. Event detection and analysis from video streams. In *USC Computer Vision*, 2001.

[14] M. Piccardi. Background Subtraction Techniques: A Review. In *IEEE International Conference on Systems, Man, and Cybernetics*, volume 4, pages 3099 – 3104, 2004.

[15] R. Pless. Spatio-Temporal Background Models for Outdoor Surveillance. *EURASIP Journal on Applied Signal Processing*, 2005(14):2281–2291, 2005.

[16] C. Stauffer and W. E. L. Grimson. Adaptive Background Mixture Models for Real-Time Tracking. In *CVPR*, volume 2, pages 246 – 252, 1999.

[17] C. Stauffer and W. E. L. Grimson. Learning Patterns of Activity Using Real-Time Tracking. *IEEE Transactions on PAMI*, 22(8):747–757, August 2000.

[18] Jen-Chao Tai and Kai-Tai Song. Background Segmentation and its Application to Traffic Monitoring using Modified Histogram. In *Int. Conf. on Networking, Sensing and Control*, volume 1, pages 13 – 18, 2004.

[19] C. Tomasi and J. Shi. Good features to track. In *CVPR*, pages 593–600, 1994.

[20] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: Principles and Practice of Background Maintenance. In *ICCV*, volume 1, pages 255 – 261, 1999.

[21] H. Veeraraghavan, O. Masoud, and N. P. Papanikolopoulos. Computer Vision Algorithms for Intersection Monitoring. *IEEE Transactions on ITS*, 4(2):78–89, June 2003.

[22] D. Zhang, D. Gatica-Perez, S. Bengio, and I. McCowan. Semi-Supervised Adapted HMMs for Unusual Event Detection. In *CVPR*, pages I: 611–618, 2005.

[23] H. Zhong, J. Shi, and M. Visontai. Detecting Unusual Activity in Video. In *CVPR*, pages II: 819–826, 2004.