ROBUST IMAGE-BASED DETECTION OF ACTIVITY FOR TRAFFIC CONTROL

Y. Liu, P. Payeur
Vision, Imaging, Video and Audio Research Laboratory
School of Information Technology and Engineering
University of Ottawa
Ottawa, Ontario, Canada
[yliu055, ppayeur]@site.uottawa.ca

Abstract

An important application of image processing and computer vision is the development of intelligent systems for traffic monitoring, management and optimization. This paper presents a system for real-time detection of moving vehicles approaching an intersection from sequences of color images acquired by a stationary camera. The proposed approach is developed in the context of traffic lights control systems. As the system is dedicated to outdoor applications, efficient and robust vehicle detection under various weather and illumination conditions must be achieved. To deal with these ever changing conditions, vehicle detection relies on motion segmentation with dynamic background representation and on an original switching algorithm using HSV color mapping to achieve feature space segmentation. Experimental results using real outdoor sequences of images demonstrate the system's robustness under various and difficult environmental conditions.

Keywords: *Motion detection, adaptive background, traffic monitoring, outdoor image processing.*

1. INTRODUCTION

Applications of computer vision in traffic monitoring and control usually rely on motion detection, tracking and feature matching. However, achieving reliable detection of moving objects in natural outdoor environments remains a critical issue as the world under observation continuously changes. Among other things, lighting conditions, time of the day, sun position, shadows, weather conditions and reflection of surfaces all influence the content of images. Most existing vision based vehicle detection systems suffer from a lack of robustness in dealing with various natural conditions. In this paper, this aspect is addressed in the context of moving vehicles detection under changing weather and luminance conditions.

In the literature, a large body of vision research has been targeted at motion detection in outdoor scenes based on gray images. Many prototypes have been investigated and tested over the last decade in urban areas for traffic surveillance and control and helped to identify the challenges faced by automatic movement detection systems operating in outdoor environments [1, 2, 3]. Some systems have been designed to work under fair weather situations and clear lighting only [4], while others have been developed specifically to operate under precise weather conditions such as snowfall [5], heavy rain [6], or under low lighting [7]. Even though several improvements have been made on the detection of moving vehicles under various conditions, most systems do not demonstrate sufficient robustness under the vast variety of conditions under which they must operate.

Cucchiara *et al.* [8, 9] proposed to extract moving vehicles by motion detection and sharp edge closure under daylight while vehicles are identified by their headlights shape following image thresholding at nighttime. A luminance histogram is used to switch between daytime and nighttime modes. This system succeeds to detect vehicles during the day, whereas it does not consider shadow removal under bright sunny conditions. This tends to lead to important errors in the vehicle identification phase. The approach also tends to classify a significant number of static surface points as moving points when operating in nighttime mode. Detection of vehicles also reveals to be less efficient in snowfall conditions during the night. Moreover, as the switching between day and night modes only relies on the luminance parameter, the system has difficulties to cope with wet surfaces that create important reflections.

The most conventional schemes used for extraction of moving-object regions from outdoor images are the method of background subtraction [10] or spatio-temporal differentiation [11]. The main differences among the various strategies that have been proposed for background suppression are related with the way the background is updated. Recent improvements introduced in this area consist of a dynamic background updating procedure using a combination of the last few frames [12, 13]. These approaches are based on the assumption that pixels characteristics follow a Gaussian distribution and impose limitations that make them less suitable for real-time applications. Gupte *et al.* [14] presented such a moving object recognition method by adaptive background subtraction to extract vehicles from images. An important advantage of this technique is that it does not require an initial vehicle-free background image for initialisation. Moreover, as the background representation is adaptive over the time, motion can be efficiently detected. However, this approach tends to generate erroneous ghosts during the background evolution period, which affect segmentation results. Other more sophisticated background modeling schemes using kalman filters or multivalue representations have also been proposed to provide more flexibility to the presence of continuously moving objects in the background, such as swaying trees, or to temporary stopped objects, such as parked cars. These approaches demonstrate a good robustness but require more complex background models and updating procedures [15, 16].

On the other hand, techniques using color as a cue to extract information from images have also been introduced. Since color images provide more information about the objects in the scene than gray level images, researchers developed algorithms that are based on motion and color segmentation to detect moving objects [17, 18]. Motion segmentation is obtained by frame differentiation while color segmentation consists of a split-and-merge algorithm that results in an over segmentation of the image to make objects stand out.

In this paper, an approach is introduced to detect moving objects in outdoor scenes with high robustness to environmental conditions such as weather and lighting variations. The proposed algorithm concentrates on the detection of moving objects while avoiding confusion with static objects or noise by means of an adaptive background representation. The approach operates on color images and takes advantage of the hue-saturation-value (HSV) color space information to automatically adapt the feature segmentation algorithm [19]. The following sections detail the structure of the proposed detection system. Operation at various daytimes is analyzed. Experimental results using real traffic sequences are also presented to demonstrate the robustness of the proposed scheme under various weather and lighting conditions.

2. STRUCTURE OF THE PROPOSED APPROACH

A close analysis of sequences of images collected on real traffic scenes under various weather and lighting conditions allowed to put in evidence the main factors that influence the detection of moving objects. Especially the perception contrast appears to be critical as it directly influences the area of objects that is visible. For example, under sunny conditions, the visual perception contrast is high and the complete surface of moving vehicles can be extracted. However, at night, the perception contrast is low and only vehicles' headlights appear as clear features in the sequence of images. Therefore designing a single algorithm to handle all situations does not seem suitable as the relevant features are so different. On the other hand, it appears that environmental conditions share similar characteristics that can be grouped in two categories under a proper mapping of color information.

Based on these observations, the proposed approach relies on two detection schemes that work in a complementary manner to extract appropriate features from moving regions in color images. A mode selector based on HSV color mapping has been designed to automatically select the most appropriate strategy given the perception contrast achieved in the current set of images. Figure 1 shows the overall structure of the proposed detection system.

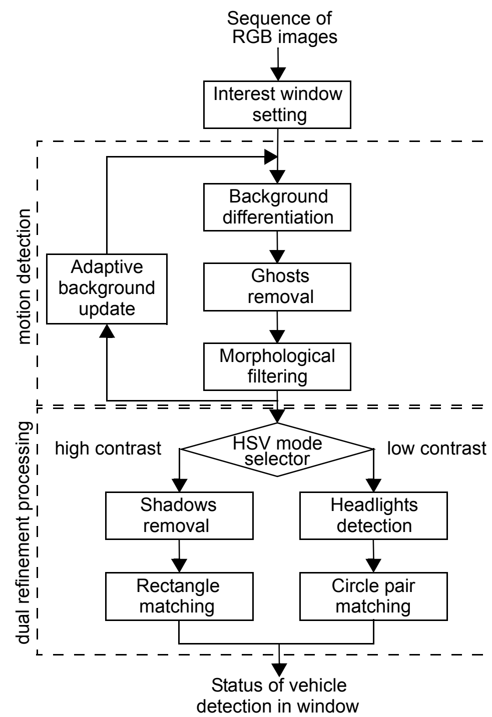


Fig. 1. Structure of the proposed robust motion detector.

When setting up the system, a search window is defined around the area of interest in the images assuming a fixed viewpoint for the camera. This window emphasizes the key observation area and significantly reduces computation time, which is critical for real-time applications. The first stage operates motion detection and segmentation on the sequence of images to detect regions of interest inside the search window. The second stage consists of a dual processing scheme that refines the detected moving areas following two different strategies that depend on the perceptual contrast available in the current set of images.

3. MOTION DETECTION AND SEGMENTATION

RGB color images provided by the camera are first processed to isolate pixels that are changing with respect to time by computing a binary mask representing the pixel-wise difference between a background representation and the current frame. Ghost effects that result from the displacement of vehicles between successive frames are then removed to improve the estimation of moving objects location. Finally the background representation is updated.

3.1 Binary Mask

The background image contains a representation of all static objects seen by the camera in the area covered by the window of interest. The foreground image to be estimated contains only the detected moving vehicles. The principle of image binarization using a brightness criterion is extended to color images segmentation as color differences appear to be less sensitive to illumination variations and inter-reflections than gray differences. The three color components of the current image, $[I_R, I_G, I_B]$, are compared with the previous background image components $[B_R, B_G, B_B]$ to compute a temporal difference image, $D(x, y, t)$, as follows:

$$D(x, y, t) = \begin{array}{l} |I_R(x, y, t) - B_R(x, y, t-1)| + \\ |I_G(x, y, t) - B_G(x, y, t-1)| + \\ |I_B(x, y, t) - B_B(x, y, t-1)| \end{array} \quad (1)$$

This difference image is thresholded at a given level, t_D, in order to extract only the pixels where a significant change in all color components has occurred. This results in a binary mask image, *M*, where pixels that correspond to moving objects have a value equal to one, while all other pixels are set to zero:

$$\begin{array}{ll} M(x, y, t) = 1 & if\ D(x, y, t) > t_D \\ M(x, y, t) = 0 & otherwise \end{array} \quad (2)$$

The threshold value, t_D, is dynamically estimated from the histogram of $D(x, y, t)$ following an adaptation of the approach proposed in [14]. Depending on environmental conditions, the histogram of $D(x, y, t)$ is more or less spread over the range of difference values as shown in figure 2. An adaptive threshold estimation technique is then used to find a dip point on the right of the peak of the histogram that has a value significantly lower than the peak value itself.

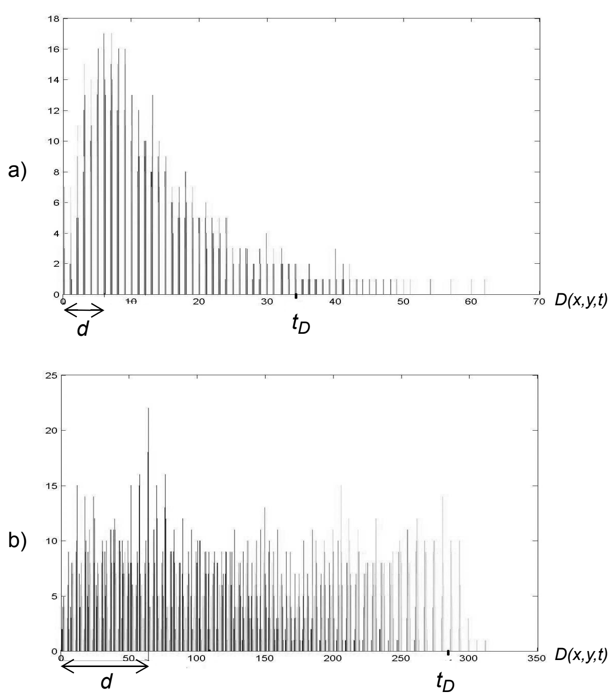


Fig. 2. Histogram of D(x,y,t): a) standard, b) under widely spread illumination changes.

The threshold estimation algorithm is defined as follows:

$$\begin{array}{ll} t_D = d \cdot s & if \quad \dfrac{max(D(x, y, t))}{d} \geq 10 \\ t_D = (d-1) \cdot \dfrac{max(D(x, y, t))}{d} & if \ 0 < \dfrac{max(D(x, y, t))}{d} < 10 \end{array} \quad (3)$$

where $max(D(x, y, t))$ is the maximum difference value computed between the current frame and the background image, and d is the distance between the peak and the origin of the graph. When the ratio $max(D(x, y, t))/d$ is larger or equal to 10, the histogram is steep and the peak tends to be closer to the origin as it corresponds to situations where changes in illumination are limited to a small number of pixels. The parameter s has been experimentally set to 7 in order to achieve good results under all weather and lighting conditions.

3.2 Ghost Detection

As the background image is progressively updated, a vehicle that starts to move or appears in the search window during the initialisation phase influences the detection for a short period of time. While the vehicle evolves inside the search window, its representation in the region where it was initially located progressively fades out but does not instantaneously disappear even though it is no longer occupying these positions. These regions of the mask image that result from the temporary memory of the background image are called "ghosts" as they do not result from real moving objects. These ghosts affect the validity of the binary mask image, as shown in figure 3b, and deteriorate the background representation. Therefore they must be eliminated.

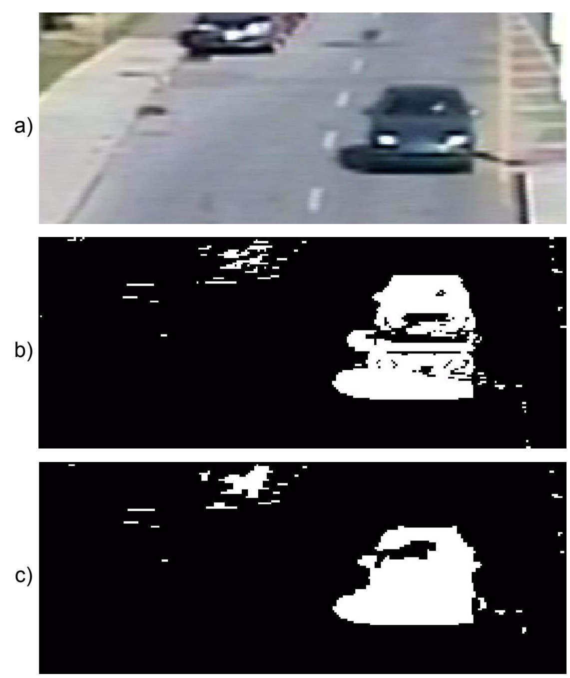


Fig. 3. Effect of ghost pixels on binary mask: a) original image, b) mask with ghost pixels, c) mask after ghost removal.

To do so, the difference between the current and the previous frames is computed over all moving pixels. As ghosts pixels have a small difference between two successive frames since their motion is only apparent and results from erroneous background pixel values, a proper correction can be applied to the binary mask. It consists in eliminating from the mask these apparently moving pixels that have a small interframe difference. The corrected binary mask is shown in figure 3c before the morphological operator is applied to filter out isolated white motion pixels from the binary mask.

3.3 Adaptive Background

In general, motion tends to generate faster changes in images than illumination and weather conditions. Representing the evolution of environmental factors over the time to allow reliable segmentation between fixed and moving components requires an adaptive background mapping to be defined and frequently updated.

To achieve this goal, a temporary background image, $TB(x, y, t)$, is computed from the binary mask image, $M(x, y, t)$ obtained previously. $TB(x, y, t)$ combines the previous background pixels, $B(x, y, t-1)$, for areas corresponding to moving objects with the current image pixels, $I(x, y, t)$, for static areas as follows:

$$TB(x, y, t) = \overline{M(x, y, t)} \wedge I(x, y, t) + M(x, y, t) \wedge B(x, y, t-1) \quad (4)$$

where $\wedge$ represents the logical AND operator.

Next, the updated background image, $B(x, y, t)$, associated with the current frame is computed as a weighted average of the temporary background, $TB(x, y, t)$, and the previous background, $B(x, y, t-1)$, as follows:

$$B(x, y, t) = (1-\alpha) \cdot B(x, y, t-1) + \alpha \cdot TB(x, y, t) \quad (5)$$

where α allows to control the background transition rate and is adaptively estimated in order to increase robustness when dealing with large illumination variations. As previously mentioned, the position of the peak in the histogram of the difference image, $D(x, y, t)$, mainly depends on illumination changes. The parameter α can therefore be defined proportionally to the ratio of the distance between the peak and the origin of the histogram of $D(x, y, t)$ and the maximum difference value computed between the current frame and the previous background image, without exceeding 0.5.

$$\alpha = min\left(0.1 + \frac{d}{max(D(x, y, t))}, 0.5\right) \quad (6)$$

In practice, this results in an increased background transition rate when illumination changes are important as it must map faster variations in the content of the scene. Figure 4 shows an example of a background image updated with the proposed approach.

Fig. 4. Adaptive background image evolution: a) initial background, b) updated background image after 6s.

4. DUAL MODE PROCESSING

Refinement of the detected moving objects is performed through a dual model processing scheme that maps vehicles as rectangular patches that contain all of their features under clear and highly contrasting daytime conditions. Under night or lighter contrasting conditions, vehicles are mapped as pairs of circular regions corresponding to their headlights. The most appropriate mode is automatically selected by a switching approach based on HSV color mapping of incoming images.

4.1 Mode Selection

Under the Hue-Saturation-Value (HSV) color mapping scheme which can be obtained by a nonlinear conversion of the original RGB images, it has been observed that the color histograms of images collected under different environmental conditions clearly differ as shown in figure 5. The cluster distribution of points in the HS histogram tends to move towards different areas of the polar representation with changes in the weather and illumination. On the other hand, color histograms of different color vehicles collected under similar conditions are rather similar one to each other.

This characteristic behavior is exploited to automatically select the motion segmentation process for different times of the day and weather conditions. Comparing the number of pixels in the image that are mapped to the upper half HS circle, *NU*, and the number of pixels that are mapped to the lower HS circle, *NL*, the most appropriate detection mode is selected.

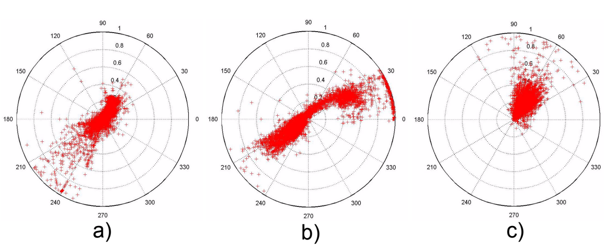


Fig. 5. HS histograms of images under different contrast conditions: a) sunny day, b) rainy day, c) night.

The ratio of highlighted pixels in the image also reveals to be critical in the selection of the detection mode. Following the HSV color mapping scheme, highlights are characterized by low Saturation and high Value values. Indeed, highlighted area appear as colorless and bright areas in the image. The ratio of highlighted pixels, H, with respect to the total number of pixels contained in the search window, W, is then computed as:

$$ratio = \frac{H}{W} \tag{7}$$

Combining the distribution of pixels on the HS histogram and the ratio of highlighted pixels, the selection process between the full feature (FF) and the headlights-based (HB) detection modes is described in pseudo-code as follows:

```
If NU < NL and ratio < φ
  Then mode=FF;
Else if NU < NL and ratio >= φ
  Then mode=HB;
Else
  Then mode=HB;
```

A proper ratio threshold level, φ, has been experimentally estimated to be 2% of pixels being highlighted. This value tends to slightly emphasize the application of the headlights-based detection mode as it provides a clearer representation of vehicles, reduces the impact of occlusion problems between vehicles and is less sensitive to shadow effects.

4.2 Shadows Detection

Shadows in images generate changes in the adaptive background and tend to be detected as moving components [20]. In the proposed approach, this effect is eliminated by discriminating shadows from moving objects taking advantage of the HSV color space mapping [21]. Given the fact that the luminance of the cast shadow is lower than that of the background while the chrominance of the shadow remains similar to that of the background, a shadow mask, $SM(x, y, t)$, is computed to extract pixels corresponding to shadow areas from the set of pixels initially tagged as moving ones in the binary mask, $M(x, y, t)$.

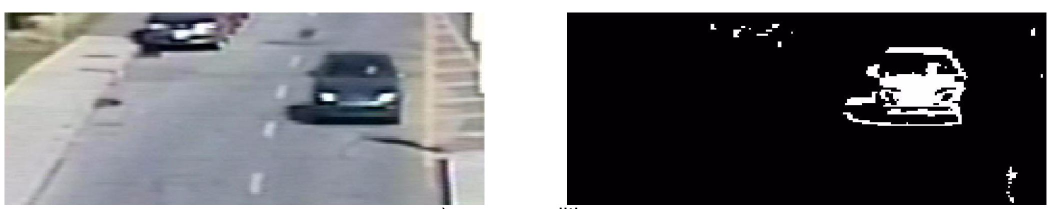
$$\begin{aligned} SM(x,y,t) = 1 \quad & if \quad \left(a < \frac{V_{I(x,y,t)}}{V_{B(x,y,t)}} < b\right) \wedge \\ & ((H_{I(x,y,t)} - H_{B(x,y,t)}) < \varepsilon_H) \wedge \\ & ((S_{I(x,y,t)} - S_{B(x,y,t)}) < \varepsilon_S) \\ SM(x,y,t) = 0 \quad & otherwise \end{aligned} \tag{8}$$

where 0<(a,b)<1 as shadow areas have lower luminance than the background, and $(\varepsilon_H, \varepsilon_S)$ are kept very small under the assumption that the chrominance of shadow and non-shadow areas is similar.

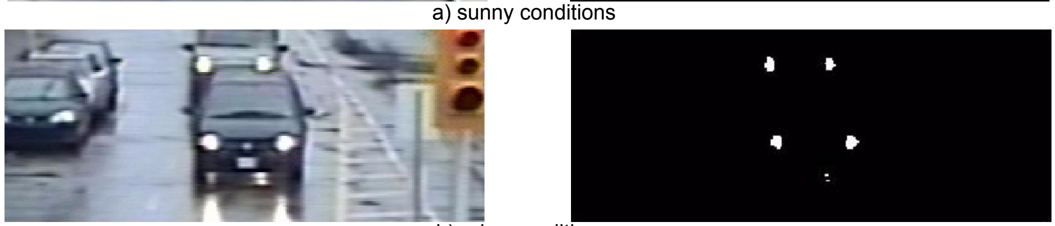
5. EXPERIMENTAL RESULTS

In the context of an application to traffic lights control, a fixed camera is installed next to a road intersection and collects image sequences of the traffic. The viewpoint is selected to cover the intersection from above. To improve system performances and to enhance the stability of the application, the field of view of the video camera is to be wide enough to observe all coming vehicles in different lanes from one direction. The goal is to use only one camera per direction of traffic.

This approach has been evaluated on sequences of real traffic scenes under various weather and illumination conditions. Original traffic images are shown on the left side of figure 6. The



a) sunny conditions



b) rainy conditions

c) at night with snow

Fig. 6. Images and corresponding binary masks under various weather and lighting conditions.

corresponding binary masks showing detected vehicles are shown on the right side of figure 6. The segmentation results using adaptive background difference and HSV color mapping were stable over a long period of time and compared advantageously with manual segmentation of images for moving objects. The approach demonstrated excellent robustness to lighting and weather changes, succeeding to detect vehicles under intense rain or snow conditions. The shadow detection and ghost removal modules also showed very good performances as rectangular or circular templates are successfully matched with vehicles while stationary objects are not detected and strong light reflection on the road is successfully eliminated.

The approach as been encoded in C++ with the help of the Intel Computer Vision Library. A frame rate of 5 images per second can currently be achieved without code optimization on a Pentium II 350 MHz processor. This appears to be sufficient for detecting cars approaching an intersection in urban areas as their speed is limited.

6. CONCLUSION

This paper introduced an approach for the detection of moving vehicles based on color image processing under various uncontrolled weather and lighting conditions. Taking advantage of the HSV color mapping, the system automatically selects the most salient features under given environmental conditions. The proposed algorithm provided excellent experimental results to detect vehicles arriving at an intersection. The automatic selection of the proper detection mode that allows to switch between a mapping of full vehicle shapes or headlights detection demonstrated very good robustness to diversified environmental conditions. The approach also efficiently handles shadow effects under strong illumination and ghosts in the adaptive background model. The resulting technology is to be extended to other applications in telesurveillance and telerobotics.

References

[1] M. Matsubara, H. Tokutome, K. Nishiyama, M. Aoki, K. Tanaka, "Development of a New Multi-Purpose Image Processing Vehicle Detector and its Implementation in the Tokyo Metropolitan Traffic Control System", in *Proc. of the Vehicle Navigation and Information Systems Conference*, pp. 95-98, 1994.

[2] E. Rowe, "The Los Angeles Automated Surveillance and Control (ATSAC) System", in *IEEE Transactions on Vehicular Technology*, vol. 40, no 1, pp. 16-20, 1991.

[3] R. M. Inigo, "Application of Machine Vision to Traffic Monitoring and Control", in *IEEE Transactions on Vehicular Technology*, vol. 38, no 3, Aug. 1989.

[4] A. Soto, A. Cipriano, "Image Processing Applied to Real-Time Measurement of Traffic Flow", in *Proc. of the IEEE Southeastern Symposium on System Theory*, pp. 312-316, 1996.

[5] H. Hase, K. Miyake, M. Yoneda, "Real-Time Snowfall Noise Elimination", in *Proc. of the IEEE Int. Conf. on Image Processing*, vol. 2, pp. 406-409, 1999.

[6] S. Kyo, T. Koga, K. Sakurai, S. Okazaki, "A Robust Vehicle Detection and Tracking System for Wet Weather Conditions Using the IMAP-VISION Image Processing Board", in *Proc. of the IEEE/IEEJ/JSAI Int. Conf. on Intelligent Transportation Systems*, pp. 423-428, 1999.

[7] R. Taktak, M. Dufaut, R. Husson, "Vehicle Detection at Night Using Image Processing and Pattern Recognition", in *Proc. of the IEEE Int. Conf. on Image Processing*, vol. 2, pp. 296-300, Nov 1994.

[8] R. Cucchiara, M. Piccardi, P. Mello, "Image Analysis and Rule-Based Reasoning for a Traffic Monitoring System", in *Proc. of the IEEE/IEEJ/JSAI/ Int. Conf. on Intelligent Transportation Systems*, pp. 758-763, 1999.

[9] R. Cucchiara, M. Piccardi, P. Mello, "Image Analysis and Rule-Based Reasoning for a Traffic Monitoring System", in *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 2, pp. 119-130, June 2000.

[10] N. Ohta, "A Statistical Approach to Background Subtraction for Surveillance Systems", in *Proc. of the IEEE Int. Conf. on Computer Vision*, vol. 2, pp. 481-486, 2001.

[11] Z. Zhu, B. Yang, G. Xu, D. Shi, "A Real-Time Vision System for Automatic Traffic Monitoring Based on 2D Spatio-Temporal Images", in *Proc. of the IEEE Workshop on Applications of Computer Vision*, pp. 162-167, Dec. 1996.

[12] M. Tsuchikawa, A. Sato, H. Koike, A. Tomono, "A Moving-Object Extraction Method Robust Against Illumination Level Changes for a Pedestrian Counting System", In *Proc. of the Int. Symposium on Computer Vision*, pp. 563-568, 1995.

[13] M. Harville, G. Gordon, J. Woodfill, "Adaptive Video Background Modeling Using Color and Depth", in *Proc. of the IEEE Int. Conf. on Image Processing*, vol. 3, pp. 90-93, 2001.

[14] S. Gupte, O. Masoud, R.F.K. Martin, N.P. Papanikolopoulos, "Detection and Classification of Vehicles", in *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 37-47, March 2002.

[15] K.P. Karmann, A. von Brandt, "Moving Object Recognition Using an Adaptive Background Memory", in V. Cappellini (ed.), *Time-Varying Image Processing and Moving Object Recognition*, 2, Elsevier, Amsterdam, The Netherlands, 1990.

[16] I. Haritaoglu, D. Harwood, L.S. Davis, "W4: Real-Time Surveillance of People and Their Activities", in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no 8, pp. 809-830, Aug. 2000.

[17] C. Zhang, M.Y. Siyal, "A New Segmentation Technique for Classification of Moving Vehicles", in *Proc. of the IEEE Vehicular Technology Conference*, vol. 1, pp. 323-326, Spring 2000.

[18] M.-P. Dubuisson, A.K. Jain, "Object Contour Extraction Using Color and Motion", in *Proc. of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 471-476, 1993.

[19] D.A. Forsyth, J. Ponce, *Computer Vision, A Modern Approach*, Prentice Hall, 2003.

[20] I. Mikic, P.C. Cosman, G.T. Kogut, M.M. Trivedi, "Moving Shadow and Object Detection in Traffic Scenes", in *Proc. of the IEEE Int. Conf. on Pattern Recognition*, vol. 1, pp. 321-324, 2000.

[21] R. Cucchiara, C. Grana, M. Piccardi, A. Pratti, S. Sirotti, "Improving shadow suppression in moving object detection with HSV color information", in *Proc. of the IEEE Int. Conf. on Intelligent Transportation Systems*, pp. 334-339, 2001.