Video Traffic Monitoring for Flow Optimization and Pollution Reduction

P. Payeur, Y. Liu

Abstract - **Vehicles are widely recognized as one the main sources of pollution in urban areas. In the context of the Kyoto protocol, means need to be identified to reduce pollution emission without compromising economic activities. Computer vision technologies can advantageously be put to contribution in developing such solutions by implementing intelligent systems for traffic monitoring, management and optimization. Improving the flow of traffic and reducing road congestion with more efficient synchronization of traffic lights can result in important reduction of carbon monoxide emissions as the duration of travels is reduced and constant speed cruise is favored. Most existing vision-based traffic monitoring systems suffer from a lack of robustness with respect to weather and lighting variations that make them unsuitable for reliable traffic light control. The present paper introduces a refined algorithm for real-time detection of moving vehicles approaching an intersection from sequences of color images. Efficient and robust vehicle detection under various weather and illumination conditions is achieved by means of an adaptive background representation and an original dual-mode feature space segmentation using HSV color mapping. Experimental results using real outdoor traffic sequences demonstrate the system's robustness under various and difficult environmental conditions.**

I. INTRODUCTION

Over the last decade, the number of vehicles circulating on our roads has been continuously increasing. The attractive cost and efficiency of transportation by trucking in comparison with massive and restricted train and naval expedition massively contributed to this increase in traffic flow on highways and urban areas. Along with the popularity of sport utility vehicles, this resulted in important increases in carbon oxides emissions and had a very negative impact on environment all around the world [1]. Unfortunately, the technology of electric vehicles is still far from being available for wide commercialisation as power storage and practical distribution of sufficient power sources still remain important issues to be addressed [2]. Difficult climates such as those found in Northern parts of Canada and Europe also preempt this technology to be widely used. Combustible fuel is therefore expected to remain the main source of energy used in transportation over the next decade. It is then necessary to develop approaches to optimize traffic management in order to protect environment. The adoption of the Kyoto protocol, some important variations in petrol cost and the increasing congestion of roads has initiated new research efforts in traffic monitoring, management and optimization.

In urban areas, important improvements on traffic flow can be achieved by an optimal programming of traffic lights according to the current traffic flow rather than time of the day. For now, this is partially achieved by means of magnetic loop detectors. Unfortunately, these loops require highly invasive installation techniques and do not allow an optimal control of traffic lights as their detection surface is limited to a few meters within the intersection area. Gray level or color image sensors are able to collect richer information about vehicles approaching an intersection as they monitor a wider area that allows to anticipate the arrival of cars at the intersection. This creates opportunities to implement higher level of intelligence in traffic light controllers such as adaptive and predictive control.

Applications of computer vision in traffic monitoring 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 with lighting and weather conditions. Most existing vision based vehicle detection systems suffer from a lack of robustness in dealing with these various natural conditions [3, 4].

In the literature, a large body of vision research has been targeted at motion detection under fair weather situations and clear lighting only [5], while specific approaches have been developed to operate under precise weather conditions such as snowfall [6], heavy rain [7], or under low lighting [8]. 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.

More sophisticated techniques propose to use different detection modes according to lighting conditions (day/night) in order to classify vehicles by types [9]. These systems generally succeed to detect vehicles both at day and during the night but neglect other aspects such as the effect of shadows. This results in misclassification of vehicles in noisy conditions such as snowfall or highly reflective wet surfaces.

The authors are with the Vision, Imaging, Video and Audio research laboratory, School of Information Technology and Engineering, University of Ottawa, Ottawa, Canada, K1N 6N5.
Email: [ppayeur, yliu055]@site.uottawa.ca

Typically, the extraction of moving-object regions from outdoor images is performed on gray scale images by background subtraction [10] or spatio-temporal differentiation [11]. Background subtraction implies a dynamic background updating procedure [12, 13, 14, 15, 16]. 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, these algorithms combine motion and color segmentation to detect moving objects [17, 18, 19]. Motion segmentation is obtained by frame differentiation while color segmentation consists of a split-and-merge algorithm that leads to an over segmentation of the image that makes objects stand out.

In this paper, an approach is introduced to detect moving objects in outdoor scenes with high robustness to environmental conditions. The proposed approach uses color image sequences and allows to simultaneously detect moving objects in various parts of the scene. By means of these multiple areas of detection, specific information about cars that are reaching or approaching an intersection as well as their location on some specific lanes (e.g. turning lanes) can be provided in real-time to the traffic light controller. The algorithm operates on images encoded in the hue-saturation-value (HSV) color space to automatically adapt the feature segmentation algorithm according to the current environmental conditions [18, 20]. The following sections detail the proposed detection system and analyze the operation at various daytimes. Experimental results on real traffic sequences are presented to demonstrate the robustness of the proposed scheme under weather and lighting conditions found in practice.

II. PROPOSED APPROACH

The perception contrast that exists in outdoor images of traffic scenes appears to be a critical factor that influences the detection of moving objects as it determines the features on which the detection needs to be based. For example, under highly contrasting sunny conditions, the complete surface of moving vehicles can be extracted. However, at night, the low contrast only allows vehicles' headlights to appear as clear features in the images. Therefore it seems suitable to develop a detection strategy that groups environmental conditions in two categories. However, such a dual processing mode requires a proper mode selector to automatically switch between different sets of features that are searched for. The proposed approach illustrated in figure 1 relies on dual detection schemes that work in a complementary manner to extract appropriate features from moving regions in color images. The mode selector uses the HSV color mapping as it is known to be less sensitive to lighting variations.

In order to provide multiple sensing and control areas along the road and to speed up processing, search windows are defined around areas of interest in the images collected from a static viewpoint. The number and position of search windows are initially selected by the operator at setting time. In the example shown in figure 2, two main areas of interest (solid lines) are defined to allow an anticipation of cars arrival at the intersection. Supplementary (dotted lines) areas can easily be added for specific direction lanes detection. Using multiple areas of detection enables programming of traffic lights to minimize idle periods for cars waiting at the intersection and to improve traffic flow thus reducing emissions.

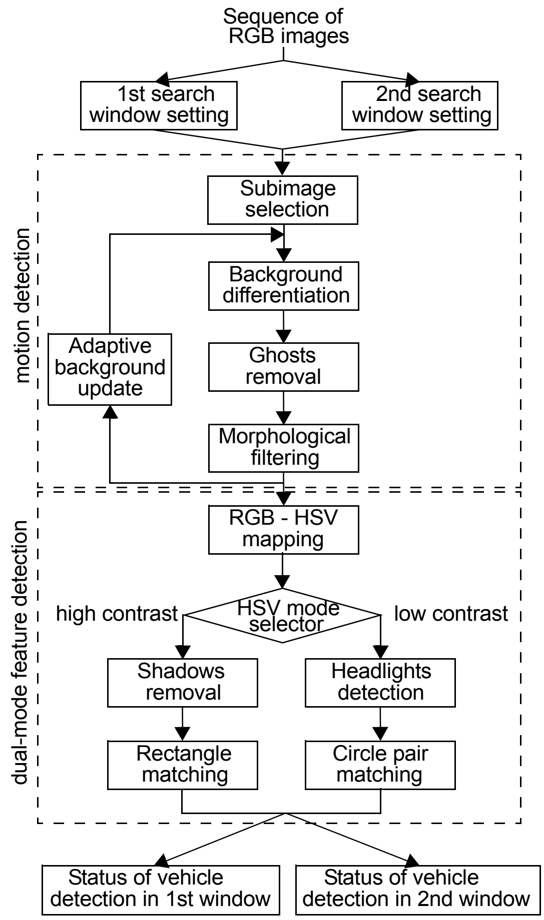


Fig. 1. Structure of the proposed robust motion detector.

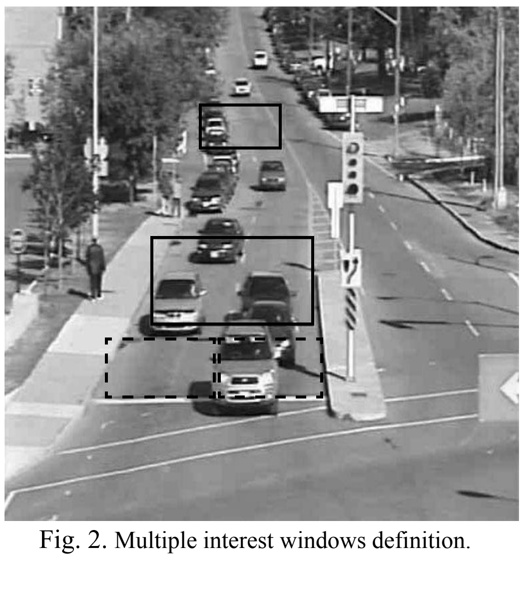


Fig. 2. Multiple interest windows definition.

III. MOTION DETECTION

Sequentially processing each search window, motion detection is achieved by computing a binary mask representing pixel-wise differences between a dynamically updated background representation and the current color frame. Morphological filtering is applied to refine the definition of moving areas before the background model is updated according to the latest binary mask to keep the track of environmental conditions. Ghost effects occurring when moving cars are present in the image at initialization or when stopped cars start to move are also eliminated from the background representation.

A. Background differentiation

A binary mask putting in evidence all moving objects is computed by comparing the current frame with an evolving background image that represents static objects and allows to model lighting and weather evolution. As color images appear to be less sensitive to illumination variations and inter-reflections than gray level ones, the principle of image binarization based on brightness level is extended to color images segmentation.

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, $D(t)$, as follows:

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

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, $M(t)$, where pixels that correspond to moving objects have a value equal to one, while all other pixels are set to zero:

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

The threshold value, t_D, is dynamically estimated from the histogram of the difference image, $D(t)$, following an adaptation of the approach proposed in [14]. Depending on environmental conditions, the spread of the histogram of $D(t)$ can vary significantly as shown in figure 3. Therefore, an adaptive threshold estimation technique has been designed to automatically find a proper value that retains only those pixels that exhibit a significant difference between the current frame and the latest version of the background image.

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

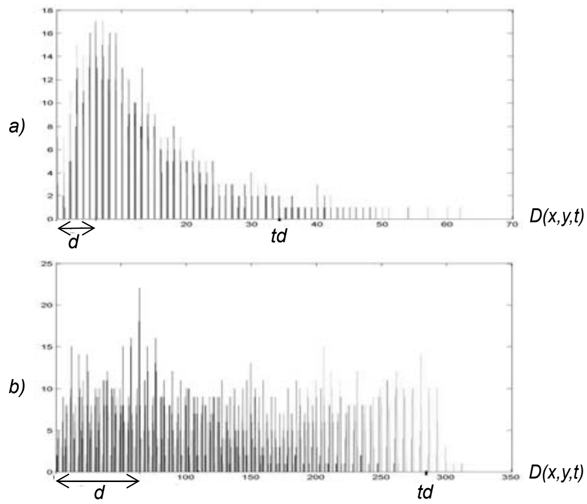


Fig. 3. Histogram of D(x,y,t): a) small number of changing pixels, b) under widely spread illumination changes.

where $max(D(x,y,t))$ is the maximum difference value 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 large, 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 (e.g. night operation). The parameter s has been experimentally set to 7 in order to optimize performances under all environmental conditions.

B. Adaptive background representation

Vehicles motion generates 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, $TB(t)$, is computed from the binary mask image, $M(t)$ obtained previously. $TB(t)$ combines the previous background pixels, $B(t-1)$, for areas corresponding to moving objects with the current image pixels, $I(t)$, for static areas as follows:

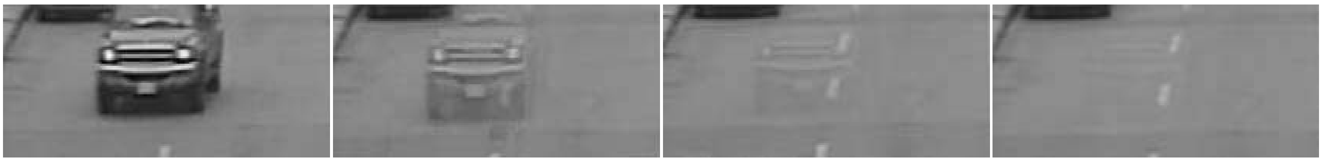
$$TB(x,y,t) = \begin{array}{l} \overline{M(x,y,t)} \wedge I(x,y,t) + \\ M(x,y,t) \wedge B(x,y,t-1) \end{array} \qquad (4)$$

where the binary mask, $M(t)$,is used as a selection factor.

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

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

The weighting parameter, α, determines the background transition rate and is adaptively estimated to increase robustness according to the magnitude of illumination variations. As the steepness of the histogram of the difference image, $D(t)$, mainly depends on illumination



a) initial background b) after 3 sec. c) after 5 sec. d) after 7 sec.

Fig. 4. Evolution of the adaptive background representation.

changes, α can be defined proportionally to the ratio of the distance between the peak and the origin of the histogram of $D(t)$ and the maximum difference value between the current frame and the previous background image. A maximum weight of 0.5 has been set experimentally to prevent abrupt transitions in the background model.

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

In practice, this results in an increased background transition rate when illumination changes are important as the background representation must then map faster variations in the static components of the scene. Figure 4 shows an example of a background image updated with the proposed approach over a period of time of 7 sec. Here α has been set to a value of 0.148 as illumination is quite constant under clear weather conditions.

C. Ghosts removal

Given the fact that the background image is progressively updated with some of the current frame information, a static object initially represented in the background that starts to move (e.g. a stopped car) requires a short period of time before it fades out from the background image. Therefore it influences movement detection over this period of time as shown in figure 5b. These regions of the binary mask image that result from the temporary memory of the background image are called "ghosts" as they correspond to virtual objects that temporarily appear in the background model. These ghosts must be eliminated to preserve coherence in the binary mask that must map only real moving objects.

A correction is applied by differentiating the current and the previous frames over all moving pixels (identified in the binary mask). Ghosts pixels have a small difference between two successive frames since their motion is virtual and results from erroneous background pixel values. Therefore a proper correction consists in eliminating these apparently moving pixels that have a small interframe difference from the binary mask. Figure 5c illustrates such a corrected binary mask.

IV. DUAL-MODE FEATURE MAPPING

Refinement of the detected moving objects is performed through a dual mode processing scheme that maps vehicles as rectangular patches that surround the complete structure of the vehicle under clear and highly contrasting daytime conditions. Under night or lighter contrasting conditions such as rain, snow or fog, vehicles are mapped as groups of circular regions corresponding to their headlights. An automatic switching approach has been developed that relies on RGB color images converted to the Hue-Saturation-Value (HSV) color mapping.

A. Mode selection

The HSV mapping of color images is a nonlinear relationship of the RGB color scheme. Under HSV mapping, it has been observed that color histograms of images collected under different environmental conditions clearly differ as shown in figure 6. The distribution of clusters of points in the HS histogram tends to move towards different areas of the polar representation with changes in weather and illumination. On the other hand, HS histograms of different color vehicles collected under similar conditions remain rather similar one to each other.

This behavior is exploited to automatically adapt the motion detection process for different environmental conditions. Comparing the number of pixels in the image, *NU*, that are mapped to the upper half of the HS polar plot, with the number of pixels, *NL*, that are mapped to the lower half of the HS polar plot, the most appropriate detection mode is selected. *NU* and *NL* are computed by an examination of the Hue angular value for each pixel in the search window to determine its location in the HS polar plot.

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, highlighted pixels are characterized by low Saturation and high Value values. Indeed, highlighted areas appear as colorless and bright regions in the image. A *ratio* of highlighted pixels, *H*, with respect to the total number of pixels contained in the search window, *W*, is then defined and computed as:

a) b) c)

Fig. 5. Effect of ghost pixels on binary mask: a) initialization image, b) mask with ghost pixels, c) mask after ghost removal.

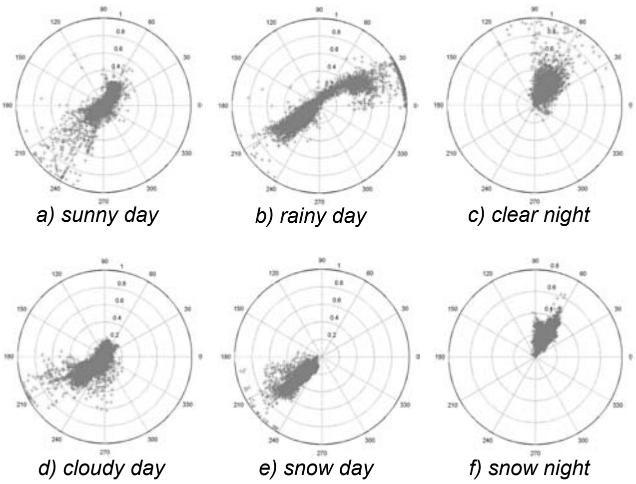


Fig. 6. Hue-Saturation histograms of images under different environmental and contrast conditions.

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

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

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

A proper ratio threshold level, τ, has been experimentally estimated to be 2% of the 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.

B. Shadows removal

Shadows in images generate unwanted changes in the adaptive background and tend to be detected as moving components [18, 21]. In the proposed approach, this effect is eliminated by discriminating shadows from moving objects taking advantage of the HSV color space mapping. 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, $SM(t)$, is computed to extract pixels corresponding to shadow areas from the set of pixels initially tagged as moving ones in the binary mask, $M(t)$.

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

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.

C. Headlights detection

Headlights detection used in the low contrast operation mode is based on the fact that pixels of interest measured under those conditions have low Saturation and high Value characteristics. Therefore, a headlight mask, $HM(t)$, of detected headlights pixels is computed by emphasizing these conditions:

$$\begin{array}{ll} HM(x,y,t) = 1 & if \quad (S_{I(x,y,t)} < c) \wedge (V_{I(x,y,t)} > d) \\ HM(x,y,t) = 0 & otherwise \end{array} \tag{9}$$

where c is set to 0.1 and d equals 0.99 in our experimental work.

V. EXPERIMENTAL RESULTS

In the context of traffic monitoring at a road intersection, a single camera is to be used per direction of traffic to provide information about as many observation windows are required. This may include different lanes associated with turning signals, anticipative detection of vehicles approaching the intersection and eventually pedestrians detection to immobilize traffic in specific directions.

The approach has been evaluated on sequences of real traffic scenes under various weather and illumination conditions for the detection of vehicles in multiple windows. The output is a count of vehicles occupying the area included in each observation window that can be directly interpreted by a high level traffic light controller.

Figure 7 shows a subset of traffic images under different environmental conditions on which the corresponding binary masks have been superimposed. 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 both during the day and at night. Shadow detection and ghost removal modules also showed very good performances.

The approach has been encoded using C++ language and Intel's CV library. A frame rate of 5 frames per second can currently be achieved without code optimization on a Pentium II 350 MHz processor. Interfacing the information provided by the proposed detection mechanism with an intelligent predictive traffic light controller is currently under investigation. It is expected that smoother traffic flow will be achieved with reduced waiting times on red lights. The resulting improvements should lead to reduction of pollutants emission and waste of energy as well as to an increased efficiency in transportation.

a) sunny day   b) cloudy day   c) rainy night   d) snowy night

Fig. 7. Images and corresponding binary masks under various weather and lighting conditions.

VI. CONCLUSION

This paper introduced an approach for the monitoring of moving vehicles based on color image processing under various uncontrolled weather and lighting conditions. The goal is to monitor traffic in various areas surrounding a road intersection from a single visual sensor to improve traffic light programming by adaptive and predictive control. Optimized traffic flow can significantly reduce pollution emission and waste of energy in transportation.

Taking advantage of the HSV color mapping scheme, the system automatically selects the most salient features under given environmental conditions. The proposed algorithm demonstrated robustness in detecting vehicles arriving at an intersection under difficult environmental conditions and in efficiently handling shadow effects under strong illumination.

VII. REFERENCES

[1] U.S. Dept. of Energy and Environmental Protection Agency, (http://www.fueleconomy.gov/feg/climate.shtml).

[2] Miller, J.M.; Emadi, A.; Rajarathnam A.V.; Ehsani, M.: "Current Status and Future Trends in More Electric Car Power Systems", in *Proc. of the IEEE Vehicular Technology Conf.*, vol. 2, pp. 1380-1384, Jul. 1999.

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

[4] Matsubara, M.; Tokutome, H.; Nishiyama, K.; Aoki, M.; Tanaka, K.: "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.

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

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

[7] Kyo, S.; Koga, T.; Sakurai, K.; Okazaki, S.: "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.

[8] Taktak, R.; Dufaut, M.; Husson, R.: "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.

[9] Cucchiara, R.; Piccardi, M.; Mello, P.: "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] Ohta, N.: "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] Zhu, Z.; Yang, B.; Xu, G.; Shi, D.: "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] Tsuchikawa, M.; Sato, A.; Koike, H.; Tomono, A.: "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] Harville, M.; Gordon, G.: Woodfill, J.: "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] Gupte, S.; Masoud, O.; Martin, R.F.K.; Papanikolopoulos, N.P.: "Detection and Classification of Vehicles", in *IEEE Transactions on Intelligent Transportation Systems*, vol. 3, no. 1, pp. 37-47, March 2002.

[15] Karmann, K.P.; von Brandt, A.: "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] Haritaoglu, I.; Harwood, D.; Davis, L.S.: "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] Zhang, C.; Siyal, M.Y.: "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] Cucchiara, R.; Grana, C.; Piccardi, M.; Pratti, A.; Sirotti, S.: "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.

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

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

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