Video-Based Automated Traffic Analysis

Pascal Bourqui[†], Fabien Claveau[‡], Amar Mitiche[†], Denis Gingras[‡]

[†] INRS-Télécommunications
16, Place du Commerce,
Verdun (Québec)
Canada, H3E 1H6
Tel : (514) 765-7782

[‡] INO
369, rue Franquet,
Sainte-Foy (Québec)
Canada, G1P 4N8
Tel : (418) 657-7006

Email: mitiche@inrs-telecom.uquebec.ca
{gingras | fclaveau | bourqui}@ino.qc.ca

keywords: real-time vision, traffic surveillance, incident detection, tracking, kalman filter

Abstract

This paper describes a real-time video-based automated traffic analysis system which can detect incidents in a highway scene. The incidents detected by the system comprise slowdowns, traffic jams and stopped vehicles. The estimation of the traffic parameters such as speed and density serve to detect incidents. Motion detection is performed using a subtractive scheme involving an estimated background image and the current image of the video input stream. Tracking of moving objects is done with Kalman filtering ([5], [6], [7], [9]). The algorithms are implemented on a Pentium PC equipped with an image processing board (C40).

1.0 Introduction

By assisting the surveillance personnel, automated traffic analysis helps reduce the response time in case of an emergency. In particular, the video-based approach can extract high level data about the scene thus making it possible to detect incidents such as traffic jams or stopped vehicles on the road. This information may be used to warn drivers or to divert traffic on alternate paths. Long term statistical data about the observed area could also be collected in order to provide a valuable input for planning traffic management solutions.

Currently, the most popular sensor technology used to analyse traffic conditions is the inductive loop. The latter makes it possible to count vehicles and measure their speed. However, the area covered by this type of sensor is limited to one vehicle and one lane. Furthermore, installing them requires burying them beneath the road surface. In view of this, and with the advent of powerful image processing platforms at affordable costs, video monitoring clearly shows to have more potential since a single video camera can monitor a multiple-lane highway in both directions over hundreds of meters.

Many approaches have been envisaged to analyse traffic using video cameras. However, most of them demand large amount of computations (3D models, occlusion reasoning, etc.) and consequently, must use very high performance hardware. Instead, we focused on making a working prototype with an open architecture (using off-the-shelf hardware), that was aimed at detecting incidents such as traffic jams and stopped vehicles. The cost of the equipment was also kept to the minimum. The system consists of one Pentium and an image processing board (C40). This resulted in a tradeoff between achieving a reasonable processing rate and the accuracy of the estimated parameters.

2.0 System Overview

The system is divided in two parts: the image processing board and the PC. The board is specialized in low-level operations: image acquisition and filtering, motion segmentation, mask labelling, mask filtering, and blob operations. The resulting information is sent to the PC, which in turn does the higher-level processing: tracking, signal processing and diagnosis. The system requires the following a priori information about the observed scene: the calibration data and the regions of interest (ROIs). The Figure 1 illustrates these components in relation with each other.

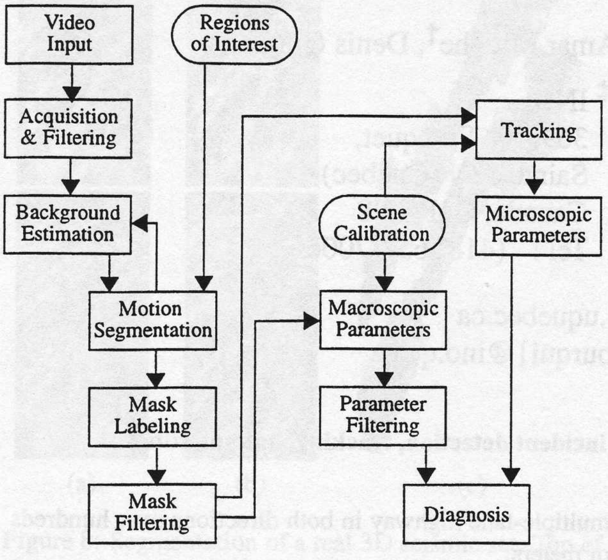


Figure 1. Block diagram of the system

As shown above, the smooth cornered rectangles indicate a priori information given to the system prior to execution. The ROIs indicate where the lanes are in the image and, for each lane, where to compute the average speed. The scene calibration is performed using Tsai's algorithm [3]. The calibration parameters are stored in a file which is used by the main program.

We will now describe the main parts of the system: motion estimation, tracking and parameter estimation.

3.0 Motion Segmentation

In order to discriminate which portion of the scene belongs to the background and to moving objects, we use a so-called *background method.* A synthesized background is generated using spatial and temporal derivatives, and is subtracted from the current image ([2], [8]). Regions where there is a large enough difference indicate moving objets in the scene. The background image is estimated using the following set of equations:

$$B(n+1,p) = B(n,p) + G(n,p)D(n,p) \tag{1}$$

$$G(n,p) = \alpha[1 - M(n,p)] + \gamma M(n,p) \tag{2}$$

$$D(n,p) = I(n,p) - B(n,p) \tag{3}$$

M is the binary movement mask, I the current image and B the background image, at time n on pixel p. The gains α and γ are based on the scene lighting and the rate of change of the background. It is crucial that the background matches closely illumination changes in the scene. The mask is made by thresholding the difference between the current image and the background:

$$M(n,p) = \begin{cases} 1 & |D(n,p)| > t(n) \\ 0 & |D(n,p)| \le t(n) \end{cases} \tag{4}$$

The threshold varies with time depending on the lighting conditions of the scene [4].

The background variance is computed differently depending on the movement mask:

$$V(n+1,p) = [1 - M(n,p)]\,T(n,p) + M(n,p)V(n,p) \tag{5}$$

$$T(n,p) = (1-\zeta)\,V(n,p) + \zeta\,(D(n,p))^2 \tag{6}$$

Where there is movement, the variance is not modified, since the pixel does not belong to the background, while in the static regions, it is a combination of the previous variance and the estimation error. This relation is controlled by the ζ factor.

4.0 Parameter Estimation

In order to make a statement about the traffic conditions, we need to estimate parameters to determine the state of each lanes and if there are stopped vehicles on the road. Two categories of parameters are estimated: *macroscopic* and *microscopic.* The ones called macroscopic are estimated for a group of vehicles and represent their global characteristics. These parameters are: average speed, average density and occupancy ratio. Their analysis for each lane will give a diagnosis about the traffic state. On the other hand, the microscopic parameter comprise the speed of each individually detected vehicle. This result is used to detect stopped vehicles.

4.1 Macroscopic parameters

The average density and the occupancy ratio are easily computed using the movement mask. Density is the number of moving object detected per lane normalized by the length of the lane, while the rate of occupancy is the ratio of moving pixels on the total area of the lane.

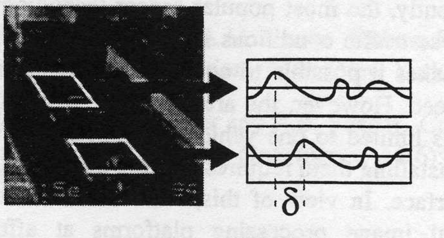


Figure 2. Illustration of average speed computation

The computation of the average speed in a given lane is performed by estimating the time delay as the vehicles cross two zones located on that lane (see Figure 2). Upon grabbing a new image, the average intensity of the pixels contained in each zone is computed. The fluctuation in time of the average intensity yields two signals, one being the delayed version of the other. By estimating the time delay, δ, between the signals and knowing the actual distance separating the two zones, the average speed is easily found.

After all the parameters have been estimated, the next step is to filter them to smooth out any sharp variation and to attenuate noise. Once this is done, the three parameters are compared to a set of rules that will determine the condition of each lane.

4.2 Microscopic parameters

The estimation of the microscopic speed of each moving object is performed by tracking the centroid of each object. Kalman filtering is used to predict the location of the objects at the next frame. Each prediction is associated with one of the new observation using a set of matching rules. If we consider a moving object at time *n*, characterized by his centroid, we can define the following model:

$$p(n+1) = p(n) + \tau\dot{p}(n) \tag{7}$$

$$\dot{p}(n+1) = \beta\dot{p}(n) + v_1(n) \tag{8}$$

$$y(n) = p(n) + v_2(n) \tag{9}$$

Where p is a point in the actual trajectory, $\dot{p}$ the velocity and y the current observation. This model is subjected to noise represented by the terms v_1 and v_2. In this model, accelerations are considered like perturbations and are contained in v_1. The constants β and τ are respectively the perspective factor and the time between frames. Thus, we can write the same model in state space:

$$x(n+1) = F(n+1, n)x(n) + v_1(n) \tag{10}$$

$$y(n) = C(n)x(n) + v_2(n) \tag{11}$$

The noise vectors v_1 and v_2 are modelled as zero-mean, white-noise gaussian processes with correlation matrix $Q_1(n)$ and $Q_2(n)$, respectively [1]. From (10) and (11) we can devise a Kalman filter which will make a prediction for each observation point if it fits into an already registered trajectory.

These predictions will be used to associate newly acquired observations with a given trajectory. When the observation is associated, we can compute an estimate for the velocity with the previous observation of the trajectory. This velocity is then converted into physical units using calibration data. We thus obtain microscopic speeds which are monitored to detect stopped vehicles.

5.0 Results

As shown in Figure 3, Figure 4 and Figure 5, the estimated background at different times is quite reliable. The outlined grey areas on the segmented images are the ROIs. We can observe that the system adapts to the ambient conditions. At night, the background estimation is harder because of headlight reflections influencing the scene. Also looking far away in the image, we can see that there is a limit on the area that can be processed by the system. The maximum distance where the analysis is still significant depends on the camera resolution.

Also, when an accident happens (Figure 6) and vehicles are immobilized, an alarm is triggered, thus drawing the operator's attention to the proper location. The side effect of the adaptive approach is that, if the segmentation is not properly performed, then the stopped vehicles are slowly absorbed in the background image as time passes. However, the absorption process would be slowed down if α and γ of (2), were functions of n and p, thus allowing different adaptation rates in the image. This would also increase the computational load.

6.0 Conclusion / Future work

Currently, the system works at 4 frames/s and can provide an accurate diagnosis in various weather conditions. The new platform will be upgraded to a C60 which will improve parameter accuracy and overall system performance. This new architecture will also allow more processing to be performed in real time. Additional software developments include shadow handling, occlusion reasoning and vehicles moving the wrong way. These shall improve the detection, tracking and diagnosis capabilities of the system.

References

[1] Simon Haykin. Adaptive filter theory, Upper Saddle River, 3e éd., Prentice Hall, 1996, 989 p.

[2] Klaus-Peter Karmann *et al.* Moving object segmentation based on adaptive reference images. *Signal Processing V: Theories and Applications*, 1990, p. 951-955.

[3] Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE journal of robotics and automation*, vol. RA-3, no. 4, august 1987, p. 323-344.

[4] Gregory W. Donohoe *et al.* Change detection for target detection and classification in video sequences. *ICASSP 88*, 1988, p. 1084-1087.

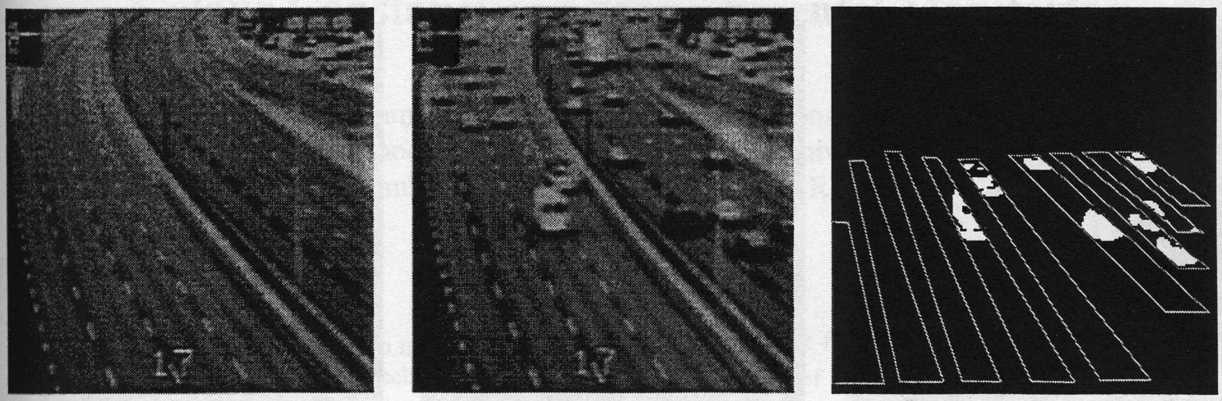
[5] D. Koller *et al.* Robust multiple car tracking with occlusion reasoning. *Technical Report UCB/ CSD-93-780*, 1994, 27 p.

[6] J. Malik *et al.* A machine vision based surveillance system for California roads. *PATH project MOU-83 Final Report*, UCB/CSD, *year unknown*, 16 p.

[7] Silvia Gil *et al.* Comparing features for target tracking in traffic scenes. *Pattern Recognition*, vol. 29, no. 8, 1996, p. 1285-1296.

[8] Amar Mitiche and Patrick Bouthemy. Computation and analysis of image motion: a synopsis of current problem and methods. *International journal of computer vision*, 1996, 65 p.

[9] N. J. Ferrier *et al.* Real-time traffic monitoring. *Workshop on the applications of computer vision*, Sarasota, FL, Dec. 1994, p. 81-88.

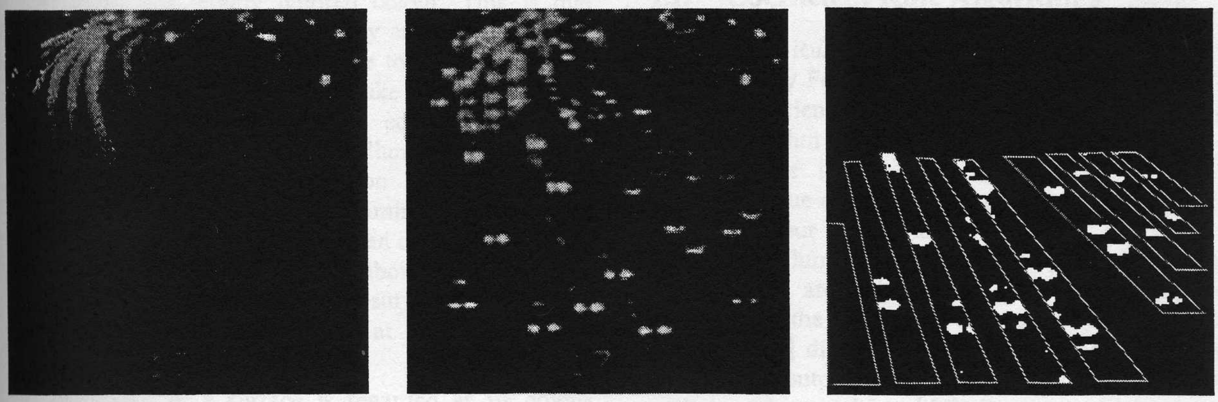


Background image Current image Segmentation image

Figure 3. Sequence at 15h05

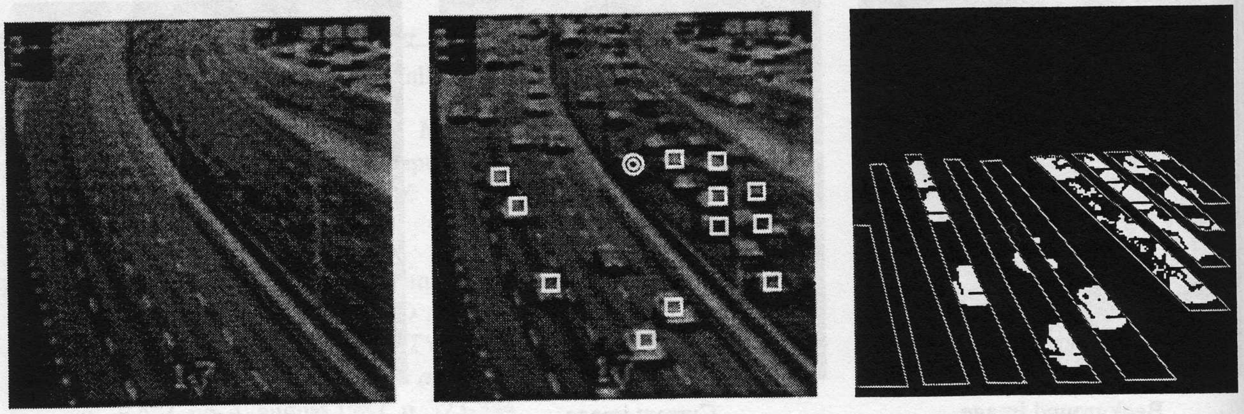
Background image Current image Segmentation image

Figure 4. Sequence at 16h05



Background image Current image Segmentation image

Figure 5. Sequence at 17h05



Background image Current image Segmentation image

Figure 6. Accident at 15h36