Proceedings of the First Young Vietnamese Scientists Meeting (YVSM '05), Nha Trang, June 12-16, 2005

Applying Computer Vision to Traffic Monitoring System in Vietnam

LE QUOC ANH, NGUYEN NGOC ANH
Union for Science-Production of New Technology, NEWSTECPRO,
17 Hoang Sam Street, Cau Giay District, Hanoi City, Vietnam
Telp: (04) 7561564 – 0912.643289
Email: quocanh@viettel.com.vn; ngocanhnguyen@voila.fr

Abstract: The purpose of this paper is to present promising results of our research on applying real-time image processing algorithms to the automatic traffic surveillance system in Vietnam.
The main functions of this system are solving problems of counting the number of vehicles passing on a road during an interval time, as well as the problems of vehicles classification and estimating the speed of the observed traffic flow from traffic scenes acquired by a camera in real-time.
The report concentrates on describing the algorithms which have been applied in automatic traffic surveillance systems in other countries, and now they are analysed and improved to reach the accuracy and to be more suitable for the traffic condition in Vietnam.
Lastly, in order to illustrate our research more precisely, a complete application of automatic traffic surveillance tested with a large number of real traffic video sequences in Hanoi city in on-line and off-line mode will be shown. The results are quite promising. For example, from 90 to 95% of vehicles are detected and counted. The speed 90 to 93% of vehicles is well estimated, depending on different situations of observed traffic flows.

1. Introduction

In order to solve the problem of traffic management, it is indispensable to have all information concerning the density as well as the speed of various kinds of transport that take part on the way. In some countries many kinds of sensor technology are applied to extract such information, for example radar, microwaves, tubes or loop detectors, etc. and more modern, the computer vision. This modern way appears to be extremely active, not necessary to interfere and is influenced by infrastructure factors like the road, the sewerage, and so on so forth.

What a pity that none of these sensor systems can be applied in the situation in Vietnam, except at the road tax station. The reason for this is the characteristics of Vietnamese traffic system: the vehicles participating on road do not move on the exact lane with a specific distance and the number of motorbikes and bikes are many times more than number of cars, bus or vans. This explains as well why some software of developed countries can not be taken in used in Vietnam.

The purpose of this paper is to propose building a vision-based traffic monitoring system for Vietnam which allows, in the first place, opening an enormous potentiality in establishing the equipment of vehicle counting and identification, and in the other place solving the problem of traffic management in Vietnam.

2. System Overview

The application of computer vision in real time into the traffic surveillance bases on the helped counting system of computer that performs the algorithms of image processing to extract the information necessary from traffic scenes acquired with cameras.

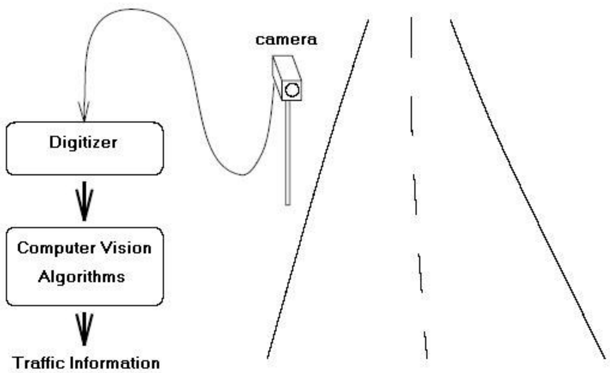

Figure 1. Model of vision-based traffic monitoring system

The information up to date extracted in real time will facilitate traffic management such as vehicle count, vehicle speed, vehicle path, vehicle density, and vehicle classification.

The general model of vision-based traffic surveillance system is illustrated in figure 1. Such traffic surveillance system will include some steps of processing as in the figure below (Figure 2):

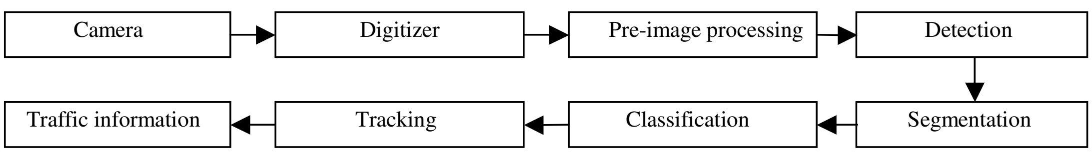

Figure 2. Block diagram of the system

The CCD camera provides live video which is digitized and fed into the computer which may well contain some special purposed hardware to cope with the extremely high data rate (~10 MBytes/s). Computer vision algorithms then perform vehicle detection, segmentation, tracking and classification. Such a single camera is able to monitor more than one lane of traffic along several hundred metres of road. A vision system could theoretically have the same powers of observation as a human observer but without the detrimental effects of tiredness or boredom. In fact, a large number of cameras are already installed on road networks for surveillance purposes in the other developed countries.

In this paper, we focus on key stages of the system, namely vehicle segmentation, vehicle classification and vehicle tracking.

3. Vehicle Detection and Segmentation

Many vision-based traffic monitoring systems rely on motion detection to segment moving regions from the image. If the regions have suitable characteristics, they are deemed to be vehicles and may then be counted or tracked as desired. There are several well established techniques for motion detection, two of which are frequently used in road monitoring systems: Frame differencing and Feature based motion detection.

In theory, there are three different kinds of efficient frame differencing algorithms to extract moving points from image sequences: *difference with background, two-frame difference* and *three-frame difference* (in other words *double-difference*). The first deals with exploiting object motion with respect to background. Whereas with the second and third algorithms, we can compute the object motion with respect to previous positions, based on the hypothesis that some object points overlap in two consecutive frames.

The feature based motion detection works by tracking prominent features from frame to frame. The first step is to identify suitable features. These should be areas of the image describing by their surroundings. Corners are a frequently used feature. After the features have been identified in all frames, the second step is to use a matching procedure to find the correspondence between these points in consecutive frames. The search for the correct correspondence is not a trivial one and iterative techniques are often used.

We have considered these algorithms. As a result we have found two algorithms realizable: *difference with background (*or method using image reference) and *three-frame difference*. Both of them are based on the fact that the difference between two frames captured at different times will reveal regions of motion.

In the *difference with background* method, a *difference image*, d(i,j) is generated by calculating the absolute difference between two frame f_1, f_2, and then thresholding the result.

$$d(i,j) = \begin{cases} 1 & \text{if } |f_1(i,j) - f_2(i,j)| > T\text{, where T is a suitable threshold} \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

There, f_1 is the incoming frame and f_2 is a *reference* or *background* frame. The reference frame is merely an image of scene with no vehicles. If the incoming frame contains no vehicles then it will be identical to the reference frame and the difference frame will contain only zeros. However, if the incoming frame does contain vehicles, then these will be shown in the difference frame.

To explain the *double-difference* algorithm, we can consider the sequence of binary $\{I_m\}$and the *difference-image* D_m is defined as:

$$D_n(i,j) = \left| I_n(i,j) - I_{n-1}(i,j) \right| \tag{2}$$

The *double-difference image* is obtained by performing a logical AND between pixels belonging to two subsequent *difference-images*, thresholded by a threshold T:

$$DD_n(i,j) = \begin{cases} 1 & \text{if } (D_{n+1}(i,j) > T) \wedge (D_n(i,j) > T) \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

The purpose of the threshold, T, is to reduce the effects of noise and changes in scene illumination.

We have realized that the method using image reference is good for extracting entire vehicle but it is difficult to initialize an image reference. Furthermore, it needs to employ a method of dynamically updating the reference frame so that it adapts to changes in scene illumination and other problems.

After testing, we have decided to choose the one based on the difference of three consecutive frames because we have found the *double-difference* method particularly robust to noise due to camera movements and changes in scene illumination. Moreover, it detects motion on the central frame I_n, where the image luminance gradient will be computed.

However, this algorithm does not detect the entire vehicle yet. Therefore, in order to improve the accuracy of vehicle detection, we proposed a small technique that combines this algorithm with the edge detection and dilation operator to find out the border of vehicles more completely and hence extract vehicles from background more exactly.

To clarify, a pixel will belong to a certain vehicle if it is achieved from *double-difference* performance. Or it appertains a vehicle if it is obtained from the edge detection and is the neighbour of double- difference pixels. Lastly, the dilation operator is performed to enlarge the boundaries of regions of moving pixels.

We have tested and achieved a very good result. Many small parts of vehicle's border can not be detected in *double-difference*, but now it is quite possible to detect them thanks to this improvement.

4. Vehicle Classification

After obtaining an initial mask for a moving object from previous steps, we may have to pre-process the mask. Normally the mask is affected by some peppercorn noises. We use median filter to remove the noises.

The purpose of this stage is to identify the vehicle to know whether it is a car, a motor-bike, a bicycle, or other kinds of means of transport. Then, it will be tracked. In this system, we have only classified them into two classes: two-wheel vehicle (motorbike, bicycle, etc.) and four-wheel vehicle (car, truck, bus, etc.). In addition we use two main features to classify vehicle: the area and the shape of unknown blobs that are extracted from the previous process. Take a vehicle as an example, if its area is less than the above threshold and bigger than the below threshold of a motorbike, then the shape of the vehicle will be taken into consideration to decide whether it is a motorbike or not.

5. Vehicle Tracking

The aim of vehicle tracking in this system is to solve two problems. The first one is that vehicle counting and tracking avoids counting one vehicle many times. The other one is vehicle speed which bases on the total time of tracking vehicles and the length of the road.

Whenever the vehicle is in the camera zone, it is detected, segmented and tracked until it disappears – get out of the zone. In order to track the object it is necessary to have a relation between the vehicles in previous frame and the vehicles in current frame. In other words these vehicles have to be determined whether they are the same object or not in the two consecutive frames.

Proposed algorithm for this problem as below:

Step 0: Create an empty database D which will contain the vehicles. Each vehicle corresponds to an entity with characteristics like position, shape, dimension, type, total tracked time, etc.

Step 1: Perform extracting the vehicles from the incoming frame.

Step 2: Vehicle classification performance.

Step 3: Each extracted vehicle will be compared to the vehicles in D: If there is not any coincidence, that means the vehicle has just entered the camera zone, so save it in D with it characteristics. If it is found out, that means it is moving in the restricted camera zone, so it is marked as flag tracked and belongs to normal move, some information of this vehicle are updated such as position, time, ... in D.

Step 4: If the objects are still in D but not updated, that means the current frame does not contain these vehicles. They can be concluded to have gotten out of the camera zone. At that time it is quite possible to give a conclusion about these vehicles as well as erase them in the D.

Step 5: Repeat the step 1 for the next frame.

6. Experiment Results

The experiment system includes a digital camera with the parameter 15 frames/s, and a computer with configurations Pentium(R) 4 CPU 2,4 GHz, RAM 128. The size of photos is 320x240 pixel. The camera is linked to the computer by the USB port.

The replaced data in some streets in Hanoi city is transferred to the computer. The figure 4 illustrates the interface of the program as well as the result screen of extracted vehicle.

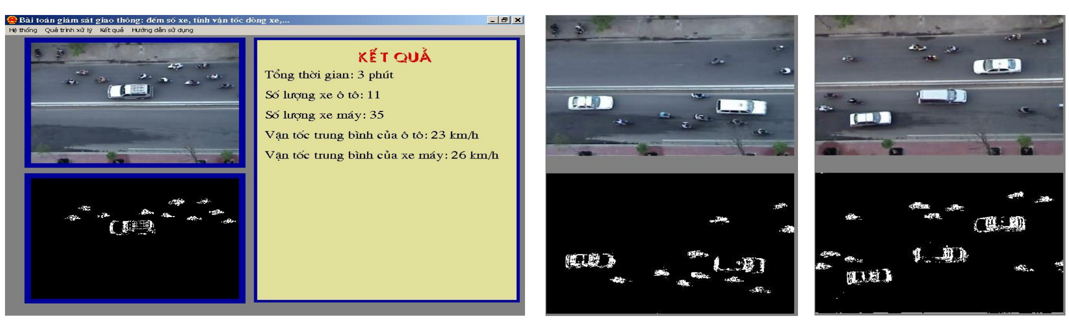

Figure 4. Interface of demo program

Comparing the result of this program after 20 minutes working with the result of manual observation, we can conclude that in good light condition, from 90% to 95% of motors and cars is detected and counted, and the accuracy of vehicle speed is 90% to 93%.

7. Conclusion

The above presents some initial results of vision-based traffic monitoring system. This research direction is quite able to be applied in Vietnamese situation. We will continue to improve this product under different conditions of road, weather, and traffic intensity in the next time.

References

[1] Belle L. Tseng, Ching-Yung Lin and John Smith, *Real-time Video Surveillance for Traffic Monitoring using Virtual Line* Analysis, Proc. IEEE International Conference on Multimedia and Expos, Switzerland, August 2002.

[2] R. Cucchiara, A. Prati, R. Vezzani, *Real Time Motion Segmentation From Moving Cameras*, in Real-Time Imaging (2004), vol. 10, n. 3, p. 127-143.

[3] R. Cucchiara R., A. Prati, R. Vezzani, *Object Segmentation in Videos From Moving Camera with MRFs on Color and Motion Features*, in Proceedings of IEEE-CS Conference on Computer Vision and Pattern Recognition (IEEE CVPR June 2003), Madison, Wisconsin, USA, vol. 1, p405-410.

[4] R. Cucchiar, C. Grana, A. Prati, *Detecting Moving Objects and their Shadows: an evaluation with the PET2002 dataset*, Proceeding of Third IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS 2002), conj. with ECCV May 2002, Copenhagen, Denmark, p18-25.

[5] Christopher John SetChell, *Applications of Computer Vision to Road-Traffic Monitoring*, PhD thesis in the Faculty of Engineering, Department of Electrical and Electronic Engineering, University of Bristol (September 1997).

[6] Cohen and G. Medioni, *Detecting and tracking moving objects in video surveillance*, IEEE Conference on Computer Vision and Pattern Recognition (1999).

[7] F.J John, *Traffic monitoring in Great Britain*, Second International Conference on Road Traffic Monitoring (1989), p. 1-4.

[8] C. Koren, J. Schvab, *Traffic survey system in Hungary*, Second International Conference on Road Traffic Monitoring (1989), p. 10-13.

[9] S.C. Lee, *Road Traffic Monitoring in Hong Kong*, Second International Conference on Road Traffic Monitoring (1989), p. 14-18.