Extracting Traffic Parameters at Intersections Through Computer Vision

K.D. Peiris and D.U.J. Sonnadara
Department of Physics, University of Colombo

ABSTRACT

A vision-based system for extracting important traffic parameters such as vehicle count, density, type and location at a three-way junction is presented. A single digital camera is used to acquire the video clip at the junction. In order to cover a wide area of the junction and to minimize the occlusion of vehicles, the camera is placed on top of a building near the junction. The accuracy of the system in vehicle counting is 91% and that of vehicle classification is 89% for small, medium and large vehicles. The errors in vehicle counting and classification occur mainly due to the occlusion of the moving vehicles. Errors also increase when the distance from the camera to the moving vehicle is high. Many of the problems in real-time implementation can be avoided to a great extent by using a multi camera system that can cover a large field of view.

1. INTRODUCTION

Information on traffic is important in the planning, maintenance, and control of any modern transport system. To resolve traffic congestion and to make good use of roads, solutions are often sought through traffic signal control systems and traffic information services.

Traffic surveillance technologies play an essential role in incident detection, traffic management, and travel time estimation. There are two basic types of traffic surveillance systems: road-based and vehicle-based. Road-based detection systems like inductive loop detectors have been a principal element of freeway surveillance and incident detection for many years. Similarly, video image detection and other roadside detection technologies have been used extensively to measure high-volume traffic conditions. However, advances in vehicle sensors and detection algorithms have given transportation authorities the opportunity to implement or enhance vehicle-based surveillance systems. Therefore, extensive efforts have been devoted to video-based measurement and analysis of traffic flow, which will greatly benefit current and future applications, including traffic control and analysis, violation detection, and vehicle identification.

Today, vision-based traffic monitoring systems are necessary to improve the traffic flow especially in heavy traffic sections. Vision-based systems are capable of providing dynamic traffic information. A video sensor is one of the most important sensors to obtain information at many road points. Vision-based camera systems are more sophisticated and powerful than those based on spot sensors, i.e., loop detectors and pneumatic sensors, since the information content associated with image sequences allows vehicle tracking and classification [1]. The application of image processing and computer vision techniques to the analysis of video sequences of traffic flow offers considerable improvements over the existing methods of traffic data collection and road traffic monitoring.

Video sensors become particularly important in traffic applications mainly due to their fast response, easy installation, operation and maintenance, and their ability to monitor wide areas. Video monitoring systems can estimate a much larger set of traffic parameters in addition to vehicle counts and speeds. These include vehicle classifications, link travel times, lane changes, rapid accelerations or decelerations, queue lengths at urban intersections, etc [2]. The parameters such as speed and density of vehicles lead to control the traffic flow by maintaining optimum time schedule for traffic signals.

This research study focuses mainly on identifying, tracking, counting and classifying the vehicles on video image sequences of traffic scenes in each lane at a three-way junction captured from a fixed digital camera during daytime, using image processing techniques.

2. METHODOLOGY

2.1 Acquiring videos from the traffic scene

At the initial stage of this project, several video sequences are acquired using a fixed digital camera. As shown in figure 1, the motion scenes are filmed with a view from above to the road surface in order to cover a large field of view and to reduce occlusions of vehicle.

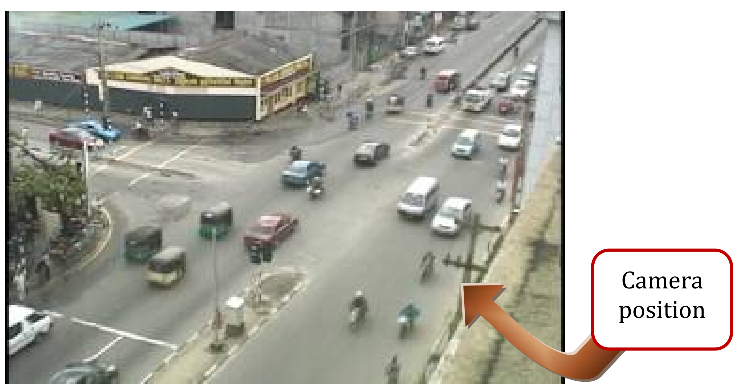

Figure 1: Traffic junction and the positioning of the camera

In most of the published works the camera is positioned in such way that the vehicles move towards the camera to avoid the side views of the movements. This is done to avoid the overlapping of the moving vehicles, which is a big obstacle in the computer vision point of view. But in this project, as shown in the figure, partial occlusion cannot be prevented because the camera should cover the three lanes simultaneously so the side views of the movements cannot be avoided and this increases the complexity of tracking vehicles significantly.

2.2 Extracting images from video sequences

A video sequence is made of basically a series of still images with a very small time interval between two consecutive images. Thus, the acquired video sequences are converted to a sequence of single frames using 'Adobe Premiere Pro 2.0' software. It is a professional, real-time, timeline based video editing software application. The images used in this work are 720×480 pixels. The video is digitized at a rate of thirty frames per second and stored in '.mat' files using the JPEG image format. These JPEG files are the sequential images required in the outer loop of the algorithm. Set of three sequential images are used in the inner loop of the developed algorithm.

2.3 Pre-processing of images

Pre-processing techniques are used to emphasize and sharpen image features for display and analysis. In this work, gray-scaling and noise filtering are used as the basic pre-processing steps.

- **Gray Scaling**

The extracted images from the video sequences are colour images. In order to apply any edge detection operator such as canny, sobel, etc. the images should be in the gray scale format. The gray scaling is also necessary to remove the colour values of the image and converts the image into a gray-scale image, which simplifies computation drastically, compared to a colour RGB image [3].

- **Noise Reduction**

Image noise is the random variation of brightness or colour information in images produced by the digital camera. In this work, median filter is used for noise reduction. Median filtering is very widely used in digital image processing because under certain conditions, it preserves edges whilst removing noise such as speckle noise and salt and pepper noise.

2.4 Vehicle identification

There are several methods for detecting moving objects in sequences taken from a stationary camera [4]. They are usually categorized into frame by frame differencing and background subtraction. Frame by frame differencing methods rely on the subtraction of subsequent frames to identify the moving regions of the image [5]. Background subtraction is a commonly used class of techniques for segmenting out objects of interest in a scene. This method is based on forming a precise background image and using it for separating moving objects from their background. The background image is specified either manually by taking an image without vehicles, or is detected in real-time by forming a mathematical or exponential average of successive images. The detection is then achieved by means of subtracting the reference image from the current image. Thresholding is performed in order to obtain presence/absence of information [6].

- **Consecutive image differencing**

This is the most direct method for making immobile objects disappear and preserving only the traces of objects in motion between successive frames. The inter-frame difference succeeds in detecting motion when temporal changes are evident [6]. Initially three consecutive images, which are already pre-processed using gray scale filter and median filter, are taken and forward and backward differences between the images are performed. In the next stage, sobel or canny edge detectors are applied to the resultant difference images.

- **Dilation**

Although the moving edges are identified using edge detectors, those edges do not make closed polygons identifiable as individual vehicles. To overcome this problem and create closed curves, a morphological operation has been used. That means, the moving edge images are enhanced by applying dilation. After the process of dilation, an image can be obtained with closed polygons where each represents a particular moving object. Here, since it is required to consider only the moving vehicles, it is required to remove the non-vehicle moving objects from the scene. For this purpose, several image processing techniques, such as 'Morphologically open binary image' method, which remove small objects, have been used.

- **Perimeter determination**

This is the final stage of moving vehicle identification process. Here, the perimeter of the dilated images is determined and the resultant convex polygons correspond to the moving vehicles. Subsequently, the bounding boxes that correspond to the identified moving objects within the region of interest can be determined i.e. because in the vehicle tracking phase it is needed to calculate the co-ordinates of the centroid of the bounding boxes which corresponds to the moving vehicles within the interested region.

2.5 Moving vehicles tracking

A vision-based traffic monitoring system needs to be able to track vehicles through the video sequence. When a physical object appears in several consecutive frames, it is necessary to identify its appearances in different frames for purposes of processing. Vehicle tracking attempts to locate, in successive frames, all vehicles that appear in the current frame. Tracking eliminates multiple counts in vehicle counting applications.

Model-based methods exploit a priori knowledge of the expected object shape. Region-based methods track connected regions that roughly correspond to the 2D shape of objects using information extracted from the entire region, such as motion, colour, and texture. Contour-based methods track the contour of the object using deformable object models (active contours). Feature-based methods track parts of the object [7].

- **Region based tracking**

In this approach, a connected region in the image, a 'blob' associated with each vehicle is identified and then tracks it over time. Blob tracking is performed by finding associations between the blobs in the current frame with those in the previous frame based on the proximity of the blobs. Euclidean distance is used for the purpose of object tracking. This approach works fairly well in free-flowing traffic. However, under congested traffic conditions, vehicles partially occlude one another instead of being spatially isolated, which makes the task of segmenting individual vehicles difficult. Such vehicles will become grouped together as one large blob in the foreground image [6]. In the case of the three-way junction, three separate regions have to be considered in order to track the vehicles in each lane simultaneously. Figure 2 represents these selected regions for vehicle tracking. There are three separate regions of interest for three lanes of the junction.

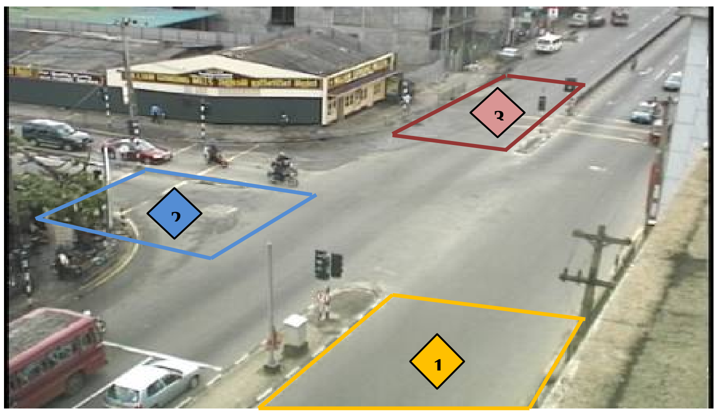

Figure 2: Selected regions for vehicle tracking

Within a particular interested region, a line of interest is defined for tracking the vehicles. As soon as a vehicle passes this line of interest, Euclidean distance measurement was started.

- **Vehicle counting and classification**

After performing the moving vehicle tracking, the next step is to count the vehicles. For a given session, it is the number of moving objects classified as vehicles. Vehicles are counted within the region of interest. Density of vehicles is the number of vehicles on a particular region in a particular time period. It is an important parameter that leads to control the traffic flow by maintaining optimum time schedule for traffic signals. In this system classification is done by classifying the vehicle as large, medium and small according to the size of the vehicles. Certain threshold values are defined for the size of the bounding boxes which enclose the tracked moving vehicles to categorize the vehicles accordingly. Table 1 shows the categorization of different types of vehicles into large, medium and small categories.

Table 1: Vehicle classification

Category	***Vehicle types***
Large	Containers, Lorries, Buses
Medium	Cars, Vans, Three wheelers
Small	Motor bikes, Bicycles

3. RESULTS AND DISCUSSION

Several video sequences are acquired using a fixed digital camera which is mounted at a high elevation on a tripod at the roadside of the three-way junction, looking down on the traffic scene. These video sequences are used for both developing and testing purposes of the system. There are three main timing cycles at the selected three-way junction. The system was developed to count the vehicles and classify them into different types in each lane of the three-way junction at each cycle. At the initial stage, the acquired video clip is sampled into multiple frames. The images used in this work are 720×480 pixels and the video is digitized at a rate of 30 frames per second.

3.1 Vehicle tracking

Tracking the detected objects is the most crucial and the hardest part of a surveillance system. For each lane of the three-way junction, separate regions of interest are defined and the vehicles are tracked within these observation zones. In this approach, a connected region in the image, a 'blob' associated with each vehicle is identified and tracked over time. Figure 3 represents a tracked vehicle within the region of interest.

Figure 3: A tracked vehicle within the observation zone

As soon as a vehicle enters in to the tracking region, the system identifies it as a vehicle and draws a bounding box that surrounds the identified blob. The centre of the bounding box is also calculated because it is needed in Euclidean distance calculations. In order to obtain robust results from counting operation, system needed to track the vehicles over time and avoid the multiple counting. For this purpose, Euclidean distance measurements have been used.

3.2 Vehicle counting

When a vehicle enters in to the tracking region the vehicle is counted by the system once. The counting procedure is continuously performed, carefully avoiding the problems due to multiple counting. The system is able to count the total number of vehicles which pass through the tracking region. Since the cycle time is a known quantity, vehicle density in each lane of the junction which is an important traffic parameter in a traffic surveillance system can be estimated. The counting and classification processes are simultaneously carried out in each lane of the three-way junction at each cycle. For the three cycles, the manually counted results and the results that are obtained from the developed system in each lane is given in Table 2.

Table 2: Vehicle counting results

	CYCLE 01			*CYCLE 02*			*CYCLE 03*		
	Path 1	Path 2	Path 3	Path 1	Path 2	Path 3	Path 1	Path 2	Path 3
Manual	12	09	09	07	-	02	56	07	35
System	11	09	08	07	-	02	43	07	20

According to the above results, the highest error of vehicle counting is in cycle 03. This is mainly due to the occlusion of vehicles. In addition, since the average distance from the camera to the moving vehicles is high for cycle 03, error is high. The average error of the system in vehicle counting is -9.5% which is a reasonable value for a real-time counting system.

3.3 Vehicle classification

The moving vehicles detected by the system are classified into three vehicular classes, Small, Medium and Large depending on the size of the bounding box. These three categories of vehicles are shown in the Figure 4. The label of the category is displayed on the top of the bounding box.

Figure 4: Vehicle classification

The vehicle classification results that are obtained manually and by the system in each lane of the three-way junction at each cycle are shown in Table 3.

Table 3: Vehicle classification results

		CYCLE 01			*CYCLE 02*			*CYCLE 03*		
		P-1	P-2	P-3	P-1	P-2	P-3	P-1	P-2	P-3
Manual	Small	01	02	02	02	-	00	18	04	09
	Medium	10	06	04	05	-	01	33	03	24
	Large	01	01	03	00	-	01	05	00	02
System	Small	01	02	01	02	-	00	12	05	05
	Medium	09	06	04	05	-	01	29	02	14
	Large	01	01	03	00	-	01	02	00	01

According to the results the average errors of classifying the vehicles into small, medium and large categories are -11.4%, -10.8% and -12.2% respectively. The highest error percentages was seen for cycle 03 and this is mainly due to the occlusion of vehicles. As in vehicle counting phase, the misclassification is increased when the distance from the camera to the moving vehicle is high. The main problem of the developed system is due to the partial occlusion of the moving vehicles. The vehicle identification is much difficult when the distance from the camera to the moving vehicle is large.

4. CONCLUSIONS

This work described a computer vision based approach to extract traffic parameters under the signalized controlled traffic conditions by analysing video sequences acquired from a monocular camera. The system is able to identify, track, count and classify vehicles into three categories: small, medium and large, in each lane of the three-way junction at each cycle.

Occlusions and pedestrian groups are the main sources of error influencing the counting and classification performance. Groups of pedestrian and unwanted moving objects must be identified and excluded from the traffic statistics. Despite these disadvantages, the system has proved to be useful in the estimation of traffic parameters under the moderate traffic conditions. The accuracy of the system in vehicle counting is 91% which is a reasonable value for a real-time vehicle counting system. For vehicle classification, the system has the accuracies of 89% small, medium and large vehicles which is also in an acceptable range.

Acknowledgements: This research is supported by the National Science Foundation (NSF) grant RG/2007/E/03.

5. REFERENCES

[1] Cucchiara R., Piccardi M. & Mello P., *Image analysis and rule-based reasoning for a traffic monitoring system*, IEEE Trasactions on Intelligent Transportation Systems, 1(2) (2002) 119-130.
[2] Coifman, B., Beymer D., McLauchlan P. & Malik J., *A real-time computer vision system for vehicle tracking and traffic surveillance*, Transportation Research Part C: Emerging Technologies, 6(4) (1998) 271-288.
[3] Chew C.L., *Moving object detection at night*, MEng Thesis, Universiti Teknologi Malaysia (2007).
[4] Radke R.J., Andra S., Al-Kofahi O. & Roysam B., *Image change detection algorithms*, IEEE Transactions on Image Processing, 14(3) (2005) 294-307.
[5] Rosin P.L., *Thresholding for change detection*, Computer Vision and Image Understanding, 86(2) (2002) 79-95.
[6] Kastrinaki V., Zervakis M. & Kalaitzakis K., *A survey of video processing techniques for traffic applications*, Image and Vision Computing, 21(4) (2003) 359-381.
[7] Messelodi S., Modena M. & Zanin M., *A vision system for the detection and classification of vehicles at urban road intersections*, Pattern Analysis and Applications, 8(1) (2005) 17-31.