Video Image Processing For Traffic Analysis

OTHMAN CHE PUAN
Jabatan Geoteknik dan Pengangkutan,
Fakulti Kejuruteraan Awam,
Universiti Teknologi Malaysia,
Sekudai,
Johor Darul Ta'zim.

ABSTRACT

In recent years the application of computer-based image processing techniques to a range of traffic data collection tasks has been successfully demonstrated. In a similar field of research carried out by the author at the University of Wales College of Cardiff, a system based on commercial image processing hardware, a 80486 IBM PC-AT and a video recorder was assembled. The main aim was to develop a system for automatic vehicle data measurement and to extend its application to the collection and analysis of pedestrian data. This paper will focus on the development of the system for vehicle detection and measurement.

A direct segmentation technique on the video images was adopted as a standard method of vehicle identification. The identification of the presence of an individual vehicle based on brightness information at relatively few sample points within the images was possible. Double threshold values were applied to the area of interest for the conversion of the area into a binary form. To compensate for the ambient lighting changes, a method of updating threshold values sequentially was introduced.

The suitability of the approach and detection algorithm was assessed by analysing a video tape containing a traffic scence for the measurement of vechicle movement. This tape is typical of tapes from which data have been extracted manually using event recorders. Although the performance of the implementation algorithms needs to be further assessed, the preliminary results have demonstrated the success of collecting data for vehicle counts, speeds and headways with reasonable accuracy.

INTRODUCTION

Traffic data, including volume, speeds, headway, intervehicle gaps, lane occupancy etc, are collected for a variety of purposes using various types of sensors and techniques. However,current traffic problems, which are more complex then ever before require comprehensive and sophisticated solution to be promoted. This increases the demand for gathering more comprehensive and better quality traffic data and statistics.

To gather such comprehensive data using conventional techniques can sometime be expensive and labour intensive. Therefore, any cost effective methods such as time-lapse photography can greatly assist in gathering the required information. The technique using 16mm film cameras was a relatively popular method of recording traffic events, but it became too expensive for general use.

Recent advances in video-cassette recording systems mean that photography can again be a useful alternative. A major disadvantage of using this method is that a considerable period of time is needed after the survey, to extract the data from the video record. Manual methods tend to be tedious and expensive , so the technique is still not particularly useful for routine surveys. However, this problem can be overcome by systems that can automatically extract the information from the video record. Research on microcomputer-based video image processing technologies applied to automatic traffic data information gathering have been widely used in Europe, Japan and United States over the last decade[1]. These systems have been reported capable, under favourable weather and lighting conditions, of detecting moving vehicles using a video camera mounted over the carriageway.

An image processing system has been developed by the author at University Wales College Cardiff for automatic traffic data collection. Its main aim was to consider the feasibility of using image processing technologies for the definition of vehicle and pedestrian interactions. This paper decribes the development of the system, implementation algorithms and the suitability of analysing video images for traffic collection.

SYSTEM ARCHITECTURE

The set up of the system is based on a collection of image processing commercial boards (Series 151 Image Processor), a host microcomputer and a video recorder. The configuration of the system is shown in a block diagram in Figure 1.

The image processing hardware sub-system forms the heart of the system. Five image processing boards each of which has a different set of image processing tasks were used. These boards are an Analog-Digital-Interface (ADI), a Frame Buffer (FB), a Histogram Feature Extraction (HF), an Arithmetic Logical Unit (ALU), and a Real Time Convolver (RTC). The image processing sub-system is connected to the microcomputer via an AT-to-VME Bus Translator.

The video bus transfers data, clock, synchronization, and control signals between sub-system modules. In this particular sub-system, it comprises five data buses. The VDI bus carries the digitised input data and timing from the interface module to the other modules. The VPI bus is a daisy-chained 16-bit data bus linking the output of each processing module to the input of the next module. It transfers 8-bit or 16 bit

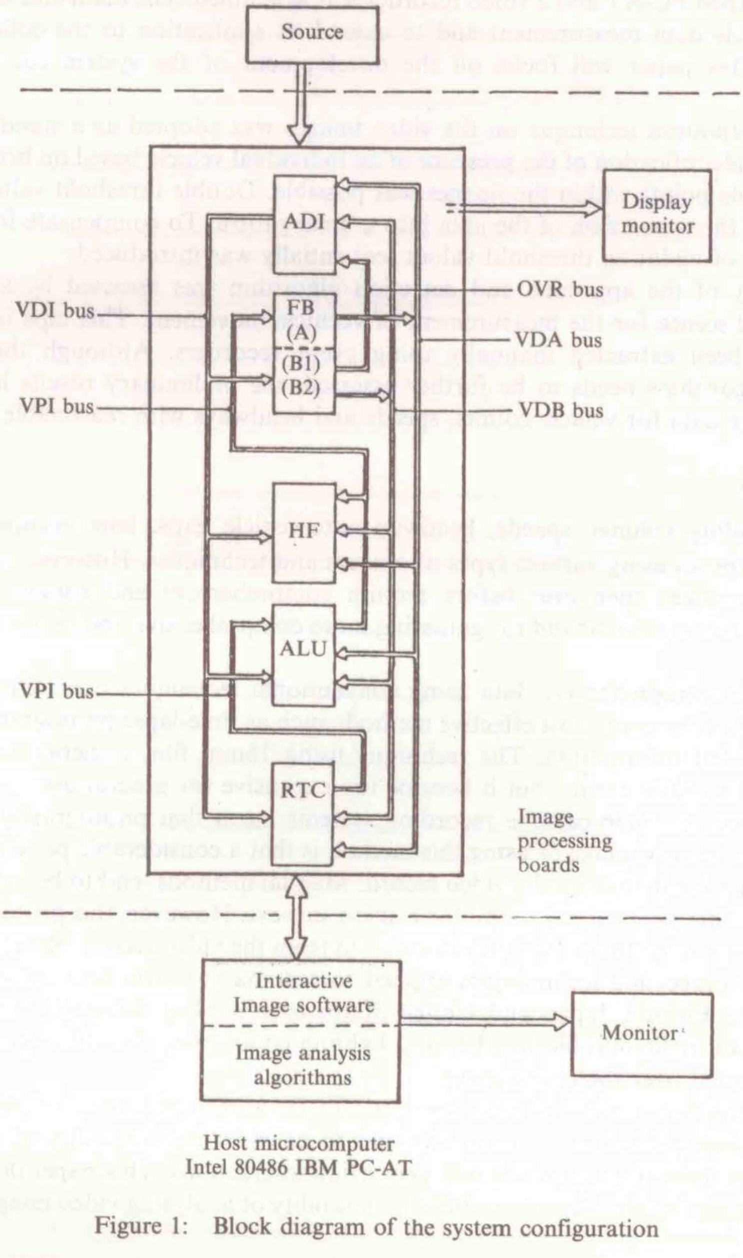


Figure 1: Block diagram of the system configuration

data using the high byte and low byte, as required. The VDA and VDB buses transfer data from the frame memory to all other modules. Finally, the OVR bus transfers optional control data to the ADI for selecting output Look-UP-Tables (LUTs) on a pixel- by-pixel basis.

The Analog-Digital-Interface is an interface module supplying data and synchronization to the other board modules. It interfaces to RS170, RS-330 and , CCIR cameras and monitors. The module is composed of an Analog-Digital-Convertor (ADC) which digitises an input analog image signal into a 8-bit or 256 grey level digital image at standard television frame rate.

The Frame Buffer contains a single 512 by 512 by 16-bit frame store (designated as frame store A), and two 512 by 512 by 8-bit frame stores (designated as B1 and B2). The 16-bit frame store functions as an accumulator holding processed images with intermediate results greater than eight bits, such as frame summation and convolutions. It receives data from VPI, and outputs data on VDA. Both frame stores B1 and B2 receive data from VDI; output data is multiplexed onto VDB. All data from frame store can be accessed directly by other modules within the sub-system including the microcomputer.

The Histogram Feature is a processor module that performs real-time analysis of pixel intensities within an image. It comprises two separate units, the Histogram unit and the Feature extraction unit. Histogram information describes the brightness, contrast, and dynamic range of an image. Histogram data accelerates intensity streching and lighting compensation. Feature information accelerates dimensional calculation and measurement. This module supports the definition of object feature of up to 16 features and records over 16000 points or more than 8000 streaks.

The Arithmetic-Logic-Unit is designed to perform most image processing functions maintaining 16-bit accuracy throughout the processing pipeline. Operations that can be performed by the ALU module include addition, subtraction, and Boolean (i.e., AND, OR and XOR) with masking.

The Real Time Convolver is a processor module performing convolutions and filtering. Convolutions on an image are performed using a 4 by 4 or 3 by 3 programmable 'kernel' in both normal and area of interest (AOI) processing modes. The board can also perform 16 by 1 Finite Impulse Response (FIR) filtering which is faster than real-time.

The Intel 80486 IBM PC-AT microcomputer running at 25 Mhz has 512 KByte RAM, 8 MByte of extended memory, 40 MByte of disk storage, and interface units. In the early stages, all the processing of an image data on brightness, thresholding, setting up area of interest, calculation of traffic parameters, and transfer of data to the attached microcomputer system were executed by the microcomputer.

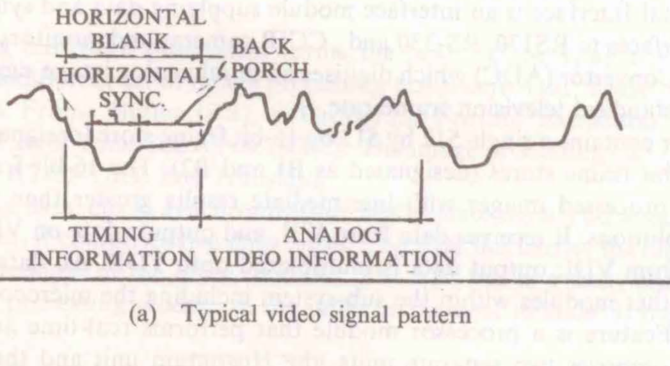
A VHS video recorder was used as a video source to the system. The related scene was recorded on a video tape using a standard television camera which comprises of 1/2 inch CCD image sensor. It produces a video output signal which conformed with of CCIR European Standard of 625 lines/frame and 25 frame/second with a normal noise ratio of 43 dB.

During the image acquisition process, the system uses a standard video format scanning technique. An analog signal containing both video and timing information, which comes alternately is sent to the ADI module for digitisation before passing to other modules. Figure 2(a) illustrates a typical analog signal pattern for the system. The analog information interval represents one horizontal line of the video image.

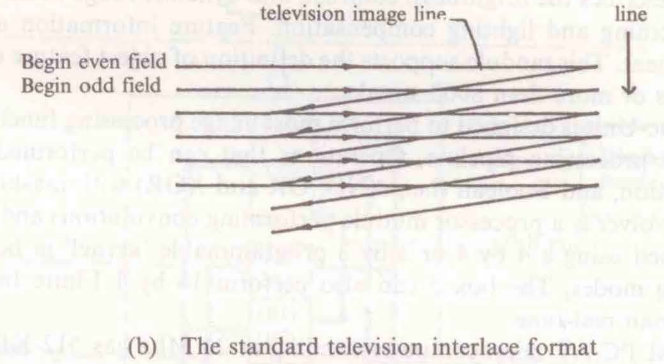
In the CCIR image, there will be 512 active lines. An analog image is read out on a line-by-line basis from left to right, top to bottom. Additionally, a technique known as interlacing is employed. Interlacing refers to the reading of all even-numbered lines, top (line zero) to bottom, followed by all odd lines. Each field of this video source is transmitted in sequence, thus creating the interlaced video image. This is shown in Figure 2(b). Because of interlacing, the television picture frame is divided into even and odd fields composed of the even-numbered lines and odd-Numbered lines, respectively. The interlacing technique is used to produce an apparent update of the entire frame in half the time that a full update actually occurs. The eyes integration of sequential fields gives the impression that the frame is updated twice as often as it really is. This results in a television monitor image with less aparent flicker.

MEASURING ALGORITHMS

Automatic measurement is possible by considering only the brightness values at relatively few sample points specified on the video images instead of dealing with all the image elements. The use of sample points



(a) Typical video signal pattern



(b) The standard television interlace format

Figure 2: Typical image digitisation format

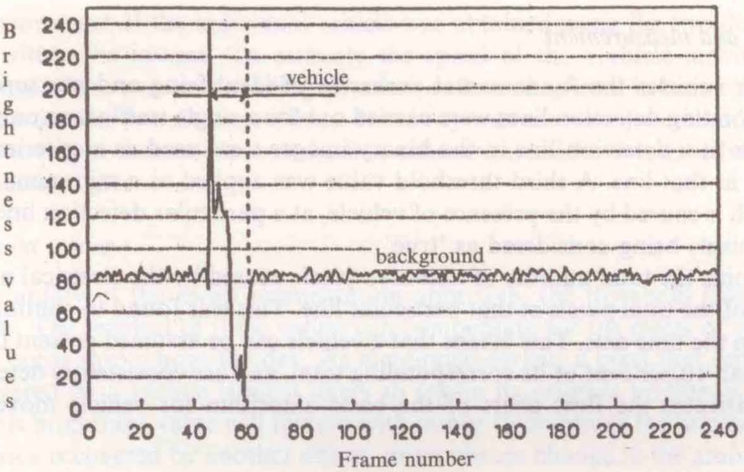
ensures that the minimum amount of image processing is required, This research uses a single line, referred to as 'detection line', as a set of sample points. This is to ensure that an object is detected at a specific reference position within an image. The detection lines were established across the vehicle's paths within the images.

The detection and measurement of vehicle movement were made by monitoring the variations of the brightness values of each point of pixel at the corresponding detection lines. Experiments showed that under favourable lighting conditions, the brightness value of a pixel in the background varied over time within a small range and with consistent amplitude. The brightness value is significantly changed when a vehicle is present at the corresponding pixel. Figure 3 illustrates this variation of brightness value of a pixel obtained from a sequence of 480 frames.

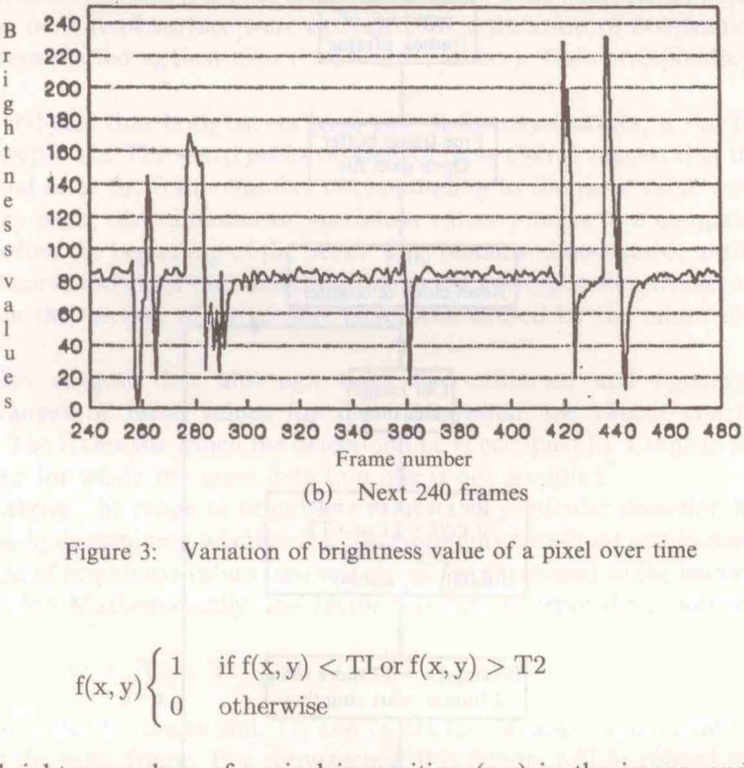
These situations are, however, true if the weather condition remains unchanged and there are no abrupt changes in the ambient lighting during the period, The level of differences in the brightness values depends on the luminance of the vehicle. It was also observed that the brightness value of a pixel representing the road surface varies along the detection line. If the detection line is located on the road surface with consistent colour, the brightness values at that line can be grouped into a short range of values.

The above criteria suggest that the objects could be identified by applying the direct segmentation line within the images. The application of this approach divides the pixel characteristics at a particular detection line into two different pixel characteristics, each having a certain uniformity. This is achieved after the appropriate threshold values for the particular image have been determined.

The minimum and maximum values of the variation in the background brightness were used as the threshold values. This means that both the minimum and maximum threshold values were established so that most values of the background brightness without vehicle presence lie between these two values. In mathematical form the image segmentation was carried out as follows:



(a) First 240 frames



(b) Next 240 frames

Figure 3: Variation of brightness value of a pixel over time

$$f(x,y)\begin{cases} 1 & \text{if } f(x,y) < TI \text{ or } f(x,y) > T2 \\ 0 & \text{otherwise} \end{cases} \qquad \text{Eqn. 1}$$

where f(x,y) is the brightness values of a pixel in position (x,y) in the image, and T1 and T2 are the minimum and maximum threshold values for a particular detection line.

Generally, if the detection line is occupied by a vehicle, its binary image will be dominated by the non-zero pixels. This, however, depends on the level of electrical noise, Which produces the points of non-zero pixels (i.e., 'false' non-zero pixels) distributed randomly in the binary image. It is possible that the detection line is also dominated by these noise points. It is therefore necessary to adopt some means of removing these noise points. In image processing procedure, there are few methods available for removing the noise points, e.g., image filtering by 'median filter', 'Fast Fourier Transform', or by applying some form of thresholding to modify the binary images.

The identification and measurement of the vehicle movement can then be carried out after the corresponding relatively 'noise free' binary images have been obtained. The procedures adopted in the processes are described in the following sub-section.

Vehicle identification and measurement

In collecting data for vehicles the fundamental processes of identifying and measuring the vehicles that pass over the corresponding detection lines were carried out for a single traffic lane only. The concentration of the non-zero pixels at a detection line in the binary images were used as a criterion in determining the presence of a vehicle at that line. A third threshold value was applied as a minimum concentration of the non-zero pixels, which is caused by the presence of vehicle, at a particular detection line. This is to eliminate the 'false' non-zero pixels being considered as 'true'

From observations, the total number of non-zero pixels caused by the electrical noise was found to be less than 10 per cent of the total pixels at that particular line. This was found to continue for not more than two or three frames in the time axis. This means that a vehicle can be assumed present if the number of non-zero pixels is more than 10 per cent of its corresponding total, and are consistently detected for more than 2 frames. Figure 4 illustrates the flow chart of the basic algorithm for vehicle movement detection and measurement.

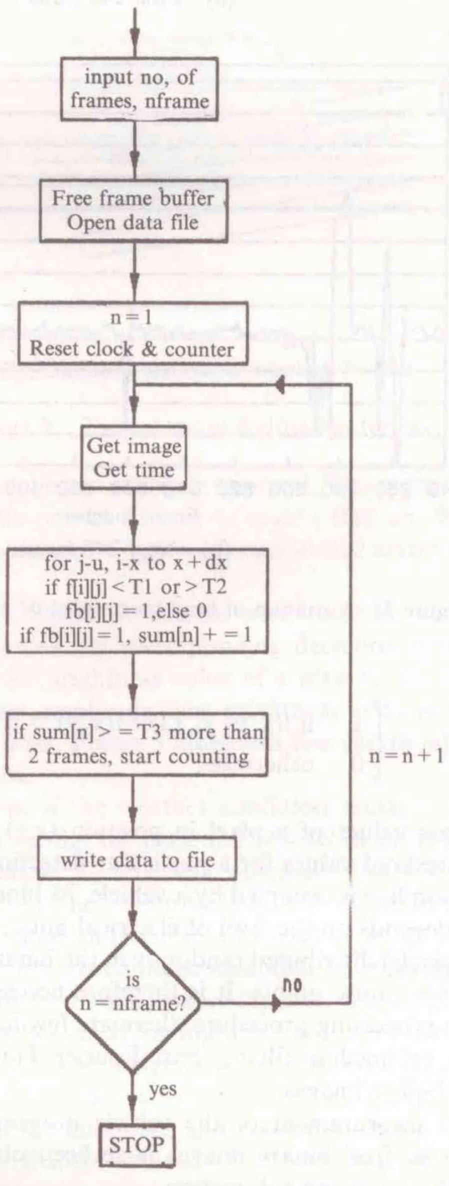


Figure 4: Basic algorithm for vehicle detection

The speed measurement of the individual vehicle was obtained using the two detection lines spaced at a known distance within the images. To estimate the speed of the vechicle moving from one point to another, the difference of the two time-readings is taken and its reciprocal in terms of hour is multiplied to the real distance between the two detection lines. Headway was measured as the interval of time between successive vehicles moving in the same lane and taken either from head to head, or from rear to rear as each vehicle passes a detection line.

Updating the threshold values

To improve the suitability of the direct segmentation approach, the threshold values are required to be updated sequentially. The renewal of these values can compensate for the efects of changes in the ambient lighting which may occur throughout the day. As mentioned earlier, a pixel that represents an area which has just been uncovered by a vechile is most likely to regain its original brightness values, possibly with minor variation. This brightness value will remain with minor fluctuations for another period of time until the corresponding area is covered by another object, or an abrupt change in the ambient lighting has taken place.

The above characteristics reflect the variations of the minimum (TI) and maximum (T2) values of the brightness at a particular detection line over time. To illustrate these characteristics, both values T1 and T2 at one detection line on a road surface were extracted for a duration of 800 frames. The corresponding values T1 and T2 were plotted against time (i.e., frame number). The corresponding curves are shown in Figure 5(a) and (b).

These Figures indicate that both curves have two distinctive features; a sharp peak and a series of variations of small amplitude. The sharp peaks on each of these curves suggest that the detection line in the few frames before and after the frame number corresponding to the peak value was being occupied by a vehicle. After a sharp peak, the minimum or maximum values returns to a comparatively different value (close to the value before the beginning of the peak). This remains almost stable until the next peak occurs.

This section of curve indicates that detection line in the corresponding frames was not occupied. The minor fluctuations in this section of curve may have been caused by the minor changes in the ambient lighting.

The curves also suggest that although both the minimum and maximum values fluctuated continuously, the ranges of these values for the background are almost consistent throughout the sequence of images. The frame for which the detection line is occupied by a vehicle shows a larger range of values than the frame for which the same detection line is not accupied.

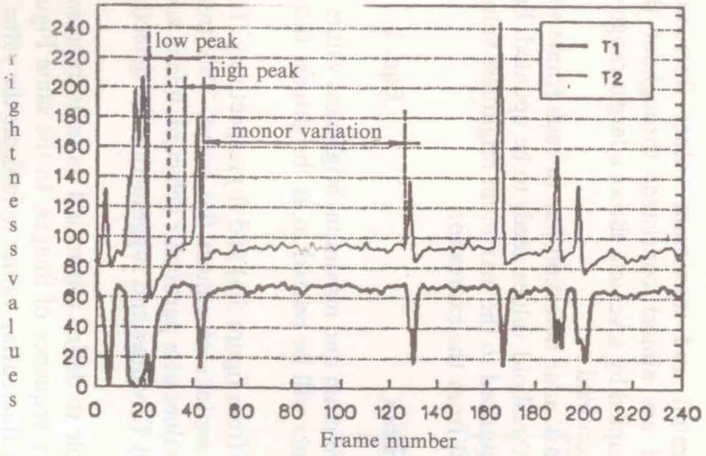
Following the above, the range of brightness values at a particular detection line in each frame was used a basic criterion in determining whether the corresponding threshold values needs to be updated. For consistency, the range of brightness values for each frame was factorised to the maximum brightness values an 8-bit image (i.e, 256). Mathematically, this factor was derived from the expression

$$w_i = (T2i - T1_i)/256 \quad i = 1, 2, 3.....n \text{ frames} \qquad \text{Eqn. 2}$$

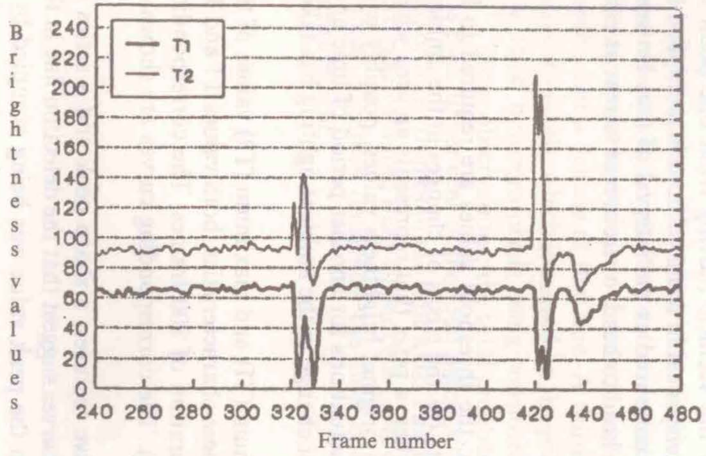
where w_i is the factor in the ith frames and, $T1_i$ and $t2_i$ are the minimum and maximum brightness values at the detection line in the same frame. For convenience, this factors will be refrred to as 'brightness factor (w)' in the remaining discussions.

Figure 5(c) and (d) show the curve of the w values obtained from figure 5(a) and (b) respectively. This curve shows a consistent pattern or trend of high peaks and sections of almost flat curve. This trend suggests better information about the changes in the brightness values at a particular detection line in each frame. Each peak corresponds to the peak in the curves of T1 and T2 values and suggests that the detection line of the corresponding frames were being occupied.

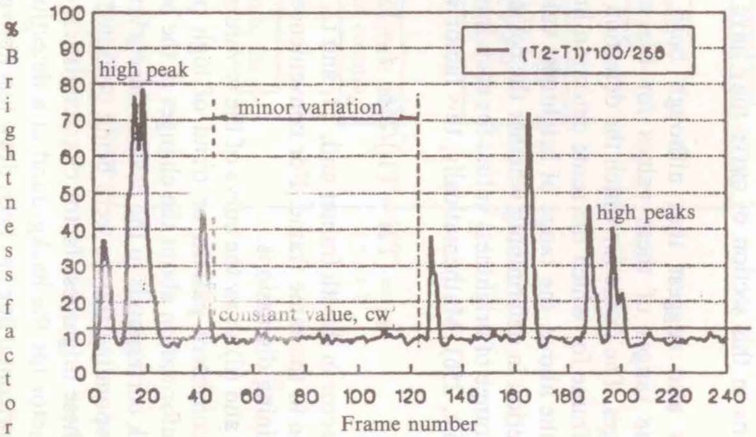
Based on these brightness factor characteristics, it is possible to define a constant value to represent the brightness factor for the background at a detection line for a sequence of images. In the same Figure 5(c) and (d), a horizontal line has been draw just above the minor fluctuating values to indicate the value of the 'constant' brightness factor (cw) which is appropriate for that image sequence. This cw value was used to identify minor and major changes in the brightness values at that detection line. The threshould updating



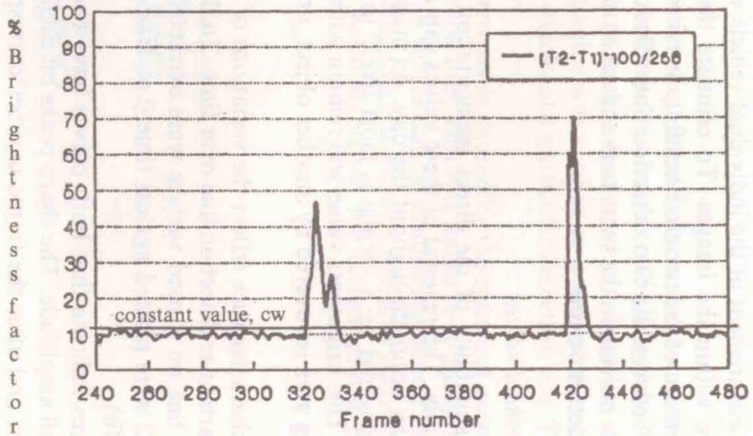
(a) First 240 frames



(b) next 240 frames



(c) Brightness factor of Image sequence (a)



(d) Brightness factor of Image sequence (b)

Figure 5 (a) & (b): Variation of maximum(T2) and minimum(T1) brightness values at a detection line over time
(c) & (d): Variation of brightness factor (cw) at a detection line over time

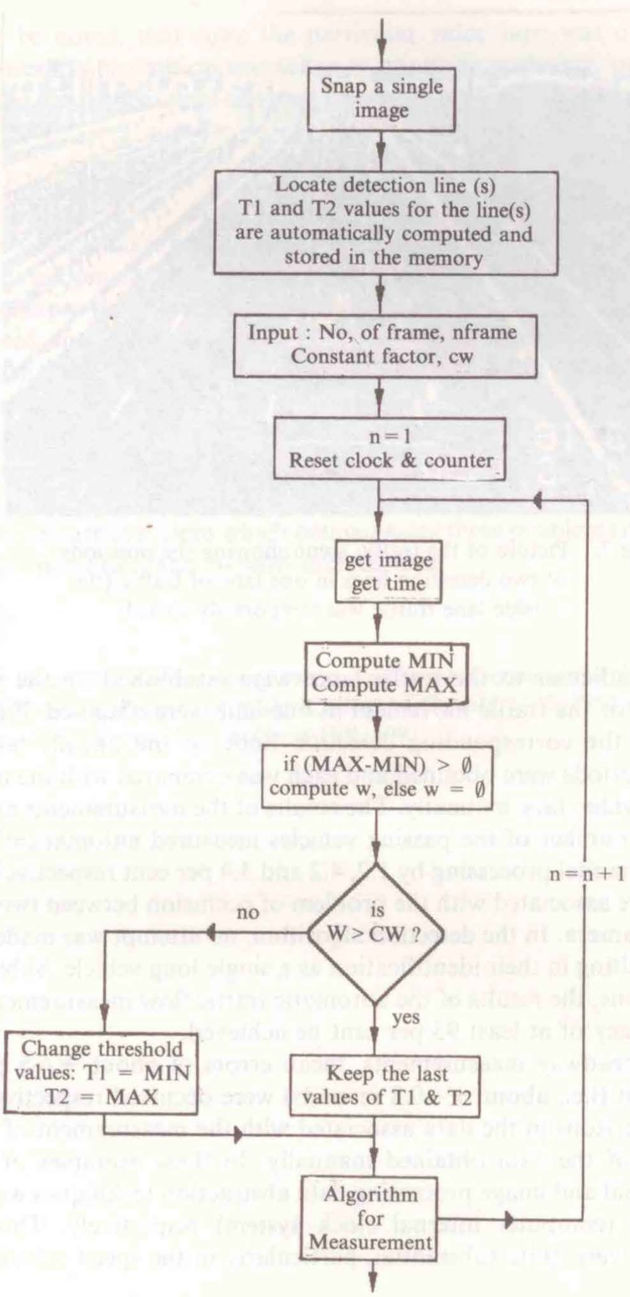


Figure 6: Flow chart of the algorithm for updating threshold values

was carried out in such a way that when the value of w is equal of lower than cw, the new threshold values that image will be defined. When w is higher than cw, the recent threshold values will not be updated. The threshold updating will not be made until the brightness factor in the next frame has indicated its closeness to the constant value. Figure 6 illustrates the improved flow chart of the detection algorithm to take into consideration the updating of threshold values.

EXPERIMENTAL RESULTS

The suitability of the adopted approach and the perfomance of the implementation algorithm in identifying and measuring vehicle movement were assessed by analysing a video tape for a traffic scene. This tape is typical of tapes from which data have been extracted manually using event recorders.

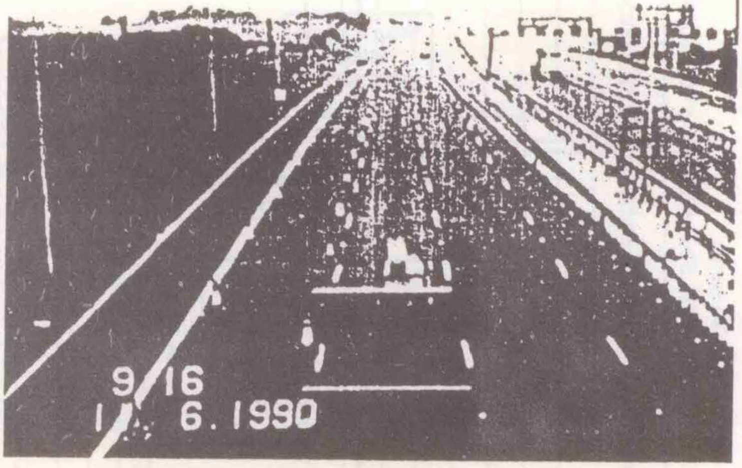


Figure 7: Picture of the traffic scene showing the positions of two detection lines in one lane of traffic (the outside lane traffic was temporarily closed)

Two detection lines perpendicular to the traffic lanes were established on the video image. As an example of measurement, data for the traffic movement in one lane were obtained. Figure 7 is view of the scene showing the position of the corresponding detection lines on the outside lane of traffic. Three measurements for five-minute periods were obtained and each was compared with the measurements which were obtained by analysing the video tape manually. The results of the measurements are shown in Table 1.

The result shows that the number of the passing vehicles measured automatically in each period of time was less then obtained by manual processing by 1.2, 4.2 and 3.4 per cent respectively. It was found that most of the missing vehicles were associated with the problem of occlusion between two or more vehicles at the detection line far from the camera. In the detection algorithm, no attempt was made to identify vehicles obscuring one another thus resulting in their identification as a single long vehicle. Subject to the frequency of the occurrence of the occlusions, the results of the automatic traffic flow measurement indicate that with the adopted approach an accuracy of at least 95 per cent be achieved.

In the vehicle speed and headway measurements, mean errors of about +/–5 per cent (i.e., about +/–5 km/h)and +– 3.5 per cent (i.e., about +– 0.8 seconds) were occurred respectively.

The accuracy of the comparisons in the data associated with the measurement of time is subjected to the limitation of the accuracy of the data obtained manually. In these examples of measurements, the accuracy of the time in the manual and image processing data abstraction techniques were 0.1 second (time-based video) and 0.01 second (computer internal clock system) respectively. This has caused some variations which in a few cases were quite substantial, particularly in the speed calculations.

Table 1: Results of measuring vehicle movement

Weather: clear/sky morning					Total
Sample No		1	2	3	
Manual count		86	72	87	244
Auto count		85	69	84	238
Count Error		–1	–2	–3	–6
	Count (%)	–1.2	–4.2	–3.4	–2.5
error	speed (kmph)	+4.8	+5.0	+5.2	+5.0
	headway (sec)	+0.08	+0.08	+0.08	–

It also must be noted, that since the particular video tape was originally for vehicles manual data abstraction purposes, no precaution was taken to minimise occlusion, or to control the visual quality, e.g., lighting control and the camera field of view. These factors are likely to adversely affect the operation of the algorithm.

CONCLUSION

Image sequence analysis is becoming an attractive field of image processing due to the new possibilities opened by the development of the computer hardware and new research problems in science and engineering. This paper devises an image processing system which is capable of extracting vehicle data automatically based on a relativity simple approach and technique. Under favourable lighting conditions, the algorithms have shown promising results in collecting vehicle data such as counts, speeds and headways within reasonable accuracy.

The effectiveness of a system, which is software-based, is still governed by the general image processing problems such as occlusion, shadows, the ambient lighting changes, and the limitation of the real-time computer processing power. This requires a great deal of research effort for complete solution. A flexible and high intelligence system which compensates these problems may be achieved by the use of faster components, large memories, and parallel architecture.

REFRENCES

1. Othman Che Puan, *The Application of Image Processing Techniques to the Collection and Analysis of Pedestrian Data*, MPhil Thesis, University of Wales College of Cardiff, U.K, 1991.
2. Don Pearson, *Image Processing*, McGraw-Hill, 1991.