APPLICATIONS OF LOW COST IMAGE PROCESSING TECHNOLOGY IN TRANSPORT

A. Rourke and M.G.H. Bell
Transport Operations Research Group
University of Newcastle upon Tyne, UK

1. INTRODUCTION

The last few years have seen the introduction of a number of image processing sub-systems for hosting by common microcomputers. Microcomputers themselves have continually been improved to provide greater data throughput. Thus, there now exists the possibility of extracting information from video images by microcomputer, in real time, at a price which is easily within the reach of any serious user.

Image processing is in effect a crude attempt at simulating a vision system. It is estimated that over 75 per cent of all the information recieved by a human is acquired through the visual process. An image processing system can be expected, therefore, to provide a very powerful and versatile means of data collection.

The application of an image processing system to traffic monitoring has a number of advantages over more conventional alternatives. There is no need to attach detectors to the road surface, for example, which should allow easier maintainence of the system hardware. In some traffic control systems today there is a duplication of equipment, with both loop detectors and video cameras (observed by a human operator) being used to monitor traffic. The use of an image processor could potentially prevent much of this duplication. Video systems can be used for a wide variety of monitioring tasks, each being addressed by writing relevant software routines. The image processor can also monitor aspects of traffic for which there are currently no other convenient or cost effective alternatives, such as pedestrian movements or area wide phenomenon like incidents,lane changing or queues.

2. THE TULIP SYSTEM

The TULIP (Traffic analysis Using Low cost Image Processing) system is currently under development at the Transport Operations Research Group at the University of Newcastle upon Tyne. The objective of this work is to develop a low cost, easy to use, image processing system which will automatically collect traffic data from a video recording in real time, using a microcomputer and a proprietary image processing sub-system. Eventually it is hoped that there will be available a suite of software capable of collecting data from videos for various different traffic situations, from simple straight forward traffic counting to complex vehicle turning movements.

The basic ingredients for such a system have recently appeared in the market place in the form of fast 16 bit microcomputers and low cost image processing sub-systems. The image processors digitise and store video images in real time, i.e. the time it takes a camera to transmit one full TV frame. Some can also perform very useful primitive processing of the image to highlight important features in the scene, also in real time, with the aid of user programmable input look-up tables (ILUTs) and a feedback loop from the frame store to the input.

Reliable video tape recorders have been available for many years at low costs, and they are now being complemented by low cost solid-state video camera/recorders.

3. RESEARCH ON IMAGE PROCESSING APPLIED TO TRAFFIC MONITORING AND CONTROL

Research into image processing for traffic applications is in progress in Europe, Japan and the USA (1). The interest in this method is due to the fact that, potentially, image processing is much more powerful and flexible than more conventional alternatives, such as buried magnetic loops.

Desirable objectives for an automatic traffic monitoring system are:-

a) vehicle count, headway, computation of speed, vehicle length and lane occupancy;

b) automatic surveillance of a length of road up to 1 Km including tracking of individual vehicles and lane changing manoeuvres;

c) automatic incident detection (based on speed changes);

d) vehicle tracking through junctions;

e) vehicle classification;

f) pedestrian traffic flow

Presently only the first of these has been successfully attained. Some of the other objectives have been attempted with varying degrees of success.

Throughout the wide field of image processing there are certain common themes, brought about by the limited speed of the current technology, and the inadequate knowledge of the problem. These are:-

Figure 1 : Camera field of view

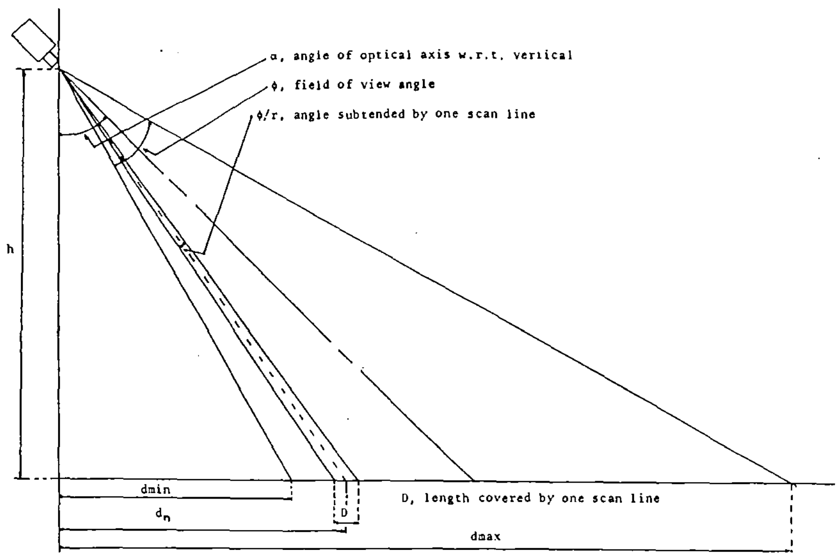
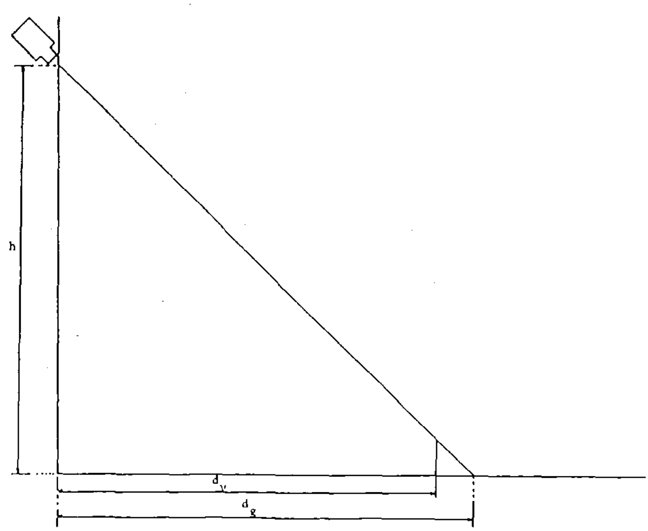


Figure 2 : Difference between detected position and actual position due to vehicle height



a) a reduction in the amount of data obtained from the camera before processing;

b) use of simple algorithms that can be efficiently implemented in real time.

A single frame, consisting of 512 by 512 picture elements or 'pixels', requires 256 Kbytes of RAM for storage, which can be found quite easily. However, to store each frame of a 5 minute video in computer memory, for example, requires 1,875 Mbytes of memory, which is not a trivial amount.

To overcome this problem various methods have been devised. Most commonly the processing is concentrated on pixels from only key regions in the image. These regions may be lines parallel or perpendicular to the roadway or small areas distributed along the lanes to be monitored. They are generally referred to as windows. This approach allows the use of general purpose microcomputers to analyse scenes in real time.

4. PRACTICALITIES OF IMAGE PROCESSING FOR TRAFFIC MONITORING

The distance over which a road may be usefully monitiored is governed by the resolution of the camera, the height of the camera above the ground and the angle of the optical axis with respect to the vertical. From Fig 1 it can be seen that

$$d_n = h * \tan [(alpha - phi/2) + (n+1/2)*(phi/r))] \quad (1)$$

$$D = h * \{\tan [\arctan (d_1 /h) + (phi/2r)] - \tan [\arctan (d_2 /h) - (phi/2r)]\} \quad (2)$$

Where

d_n	is the distance of the projected image of line n from the camera
D	is the length of roadway covered by one scan line of the image
h	is the height of the camera above the road
alpha	is the angle of the optical axis w.r.t the vertical
phi	is the field of view angle of the camera lens
r	is the number of lines in the image

Consider a camera placed 7.5m above the roadway with a field of view angle of 46 degrees with its optical axis subtending an angle of 67.5 degrees with respect to the vertical. Dickinson and Waterfall (2,3) suggest that for a reasonable probability of vehicle detection, a vehicle should cover at least three lines of the image. For vehicles 4m in length the limit to the useful field of view occurs when the scan line has a width of 1.33m. In this example this occurs at a distance of 79m from the camera.

Errors in counting vehicles can be caused by one vehicle partially obscuring or occluding another. A human operator may well be able to detect this situation, perhaps by observing another part of the image in later frames. However, if an image processing system only utilises small sections of the image, it will be unable to detect occlusions in this way. It is the responsibilty of the engineer in charge, therefore, to choose a vantage point for the camera which will minimise the chances of occlusions occuring.

Vehicle velocity measurements are made by detecting the passage of a vehicle at two separate windows a known distance apart in the scene. The number of frames it takes for the vehicle to move between the two points is directly proportional to the time, so by counting the number of frames the velocity can be determined. It is important to note the accuracy of this method and possible causes of error. One source of error is the point at which the vehicle is detected. From Fig 2 it can be seen that the distance between the two detection windows will change due to the height of the vehicle. To avoid this, the detection point should be set to the ground plane. The other major source of error is due to the temporal quantisation of the image. Frames are captured at discrete time intervals, so if two fixed points are used for detection the position of a vehicle in a frame cannot be precisely determined. There will be an associated uncertainty due to the distance a vehicle can travel in the time it takes to acquire a frame, 40 msecs. Some associated uncertainties for windows placed various distances apart can be seen in Fig 3. The distance between the points in the scene at which the vehicles are detected should be large enough to make the uncertainty as small as possible, but small enough so as to minimise the risk of missing vehicles due to lane changing manoeuvres.

The lighting of the scene is also very important for correct operation of the system. This is obviously dependent upon the weather conditions. Problems can occur if the light level changes drastically during the period of the survey or if strong shadows fall into the windows from vehicles outside them.

A further consideration when mounting the camera is the possibility of camera vibration. If the camera continually moves during the survey the system may well prove to be useless.

5. TULIP SYSTEM OPERATION

The first TULIP software routines, known as TTS_1, have been designed to perform simple analysis of traffic flowing along a straight level highway (4). The operation of TULIP is in two stages, as shown in Fig 4. Firstly a video recording of the scene to be analysed is made. This is then replayed in the laboratory and the scene analysed by the computer.

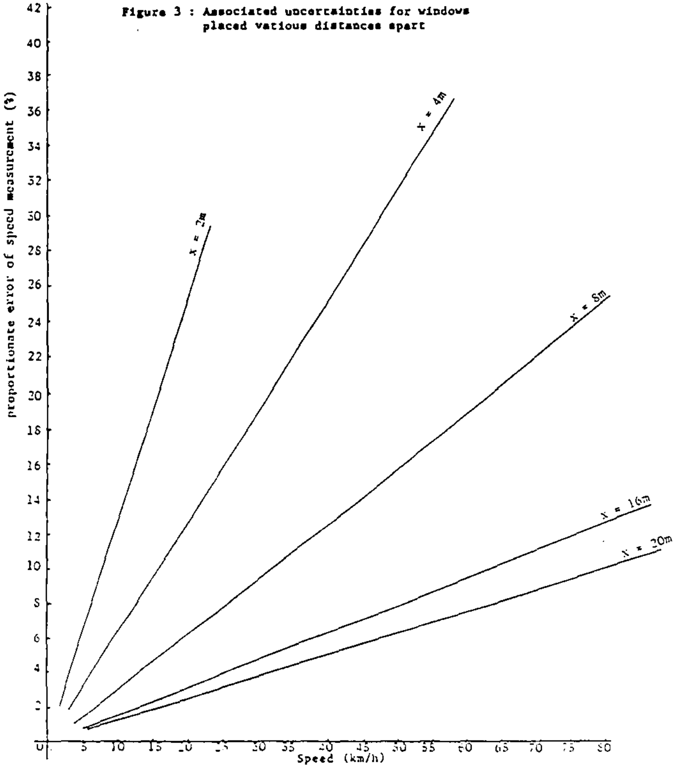


Figure 3 : Associated uncertainties for windows placed various distances apart

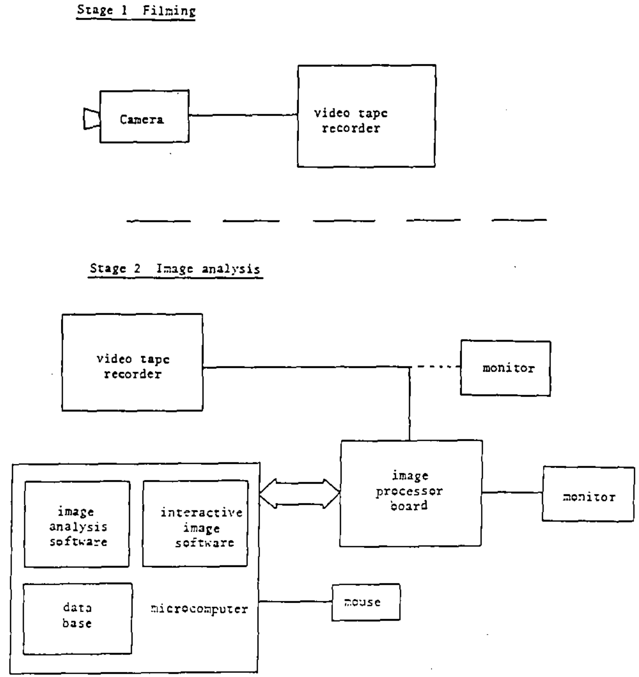


Figure 4 : Configuration of the TULIP system

The first step is being carried out in many situations at present (5). The video recordings allow engineers to study the scene repeatedly in their own time in the comfort of the laboratory. It is, however, a tedious and time consuming practice. By allowing a computer to analyse the video recording, the engineers can be engaged on more productive work.

Providing the video recordings meet certain quality criteria, they can be analysed by a microcomputer. The essential quality requirements are that there are reasonable lighting conditions (which will of course depend on the weather) and that the camera is positioned to minimise the possibility of vehicles occluding each other (an objective to be attained no matter who or what is to analyse the scene).

The microcomputer used in this system is an IBM PC-AT running at 8MHz. This acts as the host to the image processor, which is the hub of the whole system. It contains a 10MHz video ADC, which converts each line of the video picture into 512 pixels. A complete 512 line by 512 pixel image can be captured and stored in frame memory within 40 milliseconds.

Each pixel in the image has a specific grey level or intensity value representative of the illumination of its part of the scene. The image processor digitises the grey-level to a value between 0 (black) and 255 (white). These pixel intensities can be modified by passage through a user programmable input look-up table (ILUT). This enables real time segmentation of the image to be performed to highlight vehicles in the image.

There is a vast amount of information contained in a video image, so some method of data reduction must be employed. Much of the video image will be of no interest because no vehicles ever appear . The areas where vehicles appear can be designated observation areas in software and only the pixels which fall in these areas monitored by the host computer.

5.1. Automatic Traffic Counting

For vehicles to be counted automatically two observation windows per lane must be located in the scene to be analysed. This is done by the operator simply moving a cursor across the video image to define the position of these windows. The first window in each lane defined by the operator must be up stream from the second. Up to three lanes can be monitored simultaneously.

Before data collection can begin the operator must ensure that suitable values for the various segmentation parameters are determined. This is done by allowing the program to collect characteristic data from the windows in the video image. A maximum of 300 frames are monitored, but they need not be consecutive frames. Acquisition can be stopped and restarted at any time by the operator pressing any key on the keyboard. The frames monitored in this section should contain a number of vehicles passing both through the observation windows and adjacent to them so that their shadows fall in the windows if possible. The data collected is used to determine the grey scale value of the roadway and suitable threshold values to remove noise and shadows. These values are used to program the ILUT to produce a binary image. A typical grey scale distribution from such a survey is shown in Fig 5.

At this point the image processor is ready to produce binary video images suitable for processing by the data collection software. Pressing any key on the keyboard will start the data collection.

A vehicle must be counted in both observation windows before it will be included in the final data base. The data base produced by this program consists of pairs of event times and types for all vehicles detected. From this data it is a simple matter to calculate speed and headway information. A maximum of 9,362 events per window can be stored, which will allow approximately 4 hours of peak traffic to be analysed at one time.

Vehicles are classified into one of three groups, small, medium and large. Cars, motorcycles and small commercial vans will be given a small label, large commercial vans and rigid lorries will fall into the medium catagory and HGVs and buses into the large catagory. These classifications are based upon the vehicle's percieved width and length in each window.

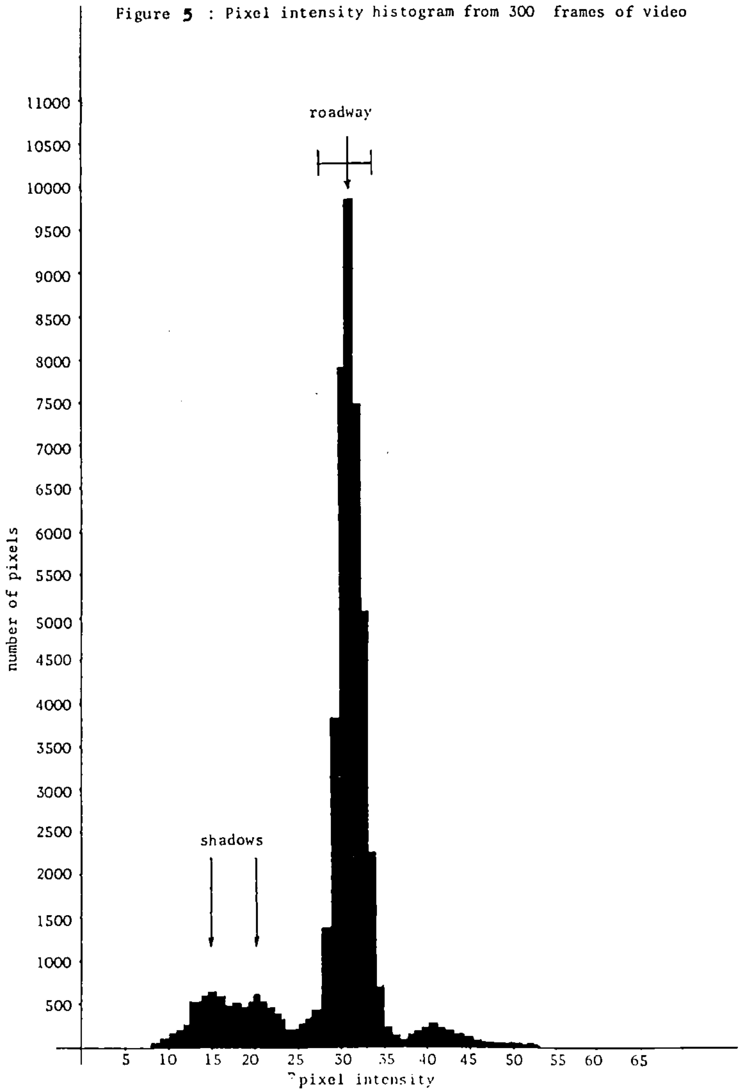
There are some instances where a lane straddling vehicle will be counted in two lanes. To reduce the occurance of this type of error a post processing function is available which will remove such vehicles. This can be done because the width of each vehicle is measured and its lateral position within the lane recorded.

The performance of TTS_1 is likely to become unreliable under heavily congested traffic situations. These situations can be detected by the system and flags set in the data base warning that the data maybe in error during certain periods. The operator can then be notified to manually inspect the video during these periods to improve the data.

5.2. Manual Traffic Counting

Presently the TULIP software is only capable of performing simple analyses as described above. To give immediate wider scope for the system a manual mode of operation is also included which is known

Figure 5 : Pixel intensity histogram from 300 frames of video



as VISTA. This algorithm was originally developed by Wootton Jeffreys Consultants (5) and has been re-written and extended for use on IBM PCs by the authors. Its inclusion enables more complex analyses to be performed, such as O-D studies, far more convienently than traditional methods allow.

In the manaual mode the operator defines a window within the scene and presses a key on his key pad each time an event of interest occurs in that window. A 16 key pad is provided, allowing up to 16 different types of event to be recorded per window. The data base produced is identical to that produced by the automatic routines and so can be passed to the same statistical analysis software without modification.

6. RESULTS

All results reported here used a pixel filtering segmentation procedure a with 3 second updating period to account for ambient light changes.

Table 1. Comparison of manual and TULIP count

Video & Duration	Light & Weather conds.	Manual Count	Computer Count	Counting Errors					
				%	i	ii	iii	iv	v
1 10 mins	Rapidly Varying	306	302	3.92	0	8	0	0	4
2 15 mins	Stable Bright	87	87	0	0	0	0	0	0
3 15 mins	Slowly varying	141	141	0	0	0	0	0	0
4 15 mins	Slowly varying	58	57	1.72	0	1	0	0	0
5 15 mins	Slowly varying	51	52	1.96	0	0	0	0	1

Errors: i) Vehicles counted as two smaller ones.
ii) Vehicles missed due to occlusion.
iii) Vehicles lost due to the segmentation and noise reduction algorithms.
iv) Vehicles missed due to light changing.
v) Extra counts due to light changing.

Video locations : 1,3,4,5 Motorway ;
2 Traffic light junction

Table 2. M6 Motorway Junction 10 7.4.1987
Video duration 89 mins

Lane	Manual Count	Computer Count	Counting Errors					
			%	i	ii	iii	iv	v
1	1529	1475	6.28	18	37	0	38	3
2	2261	2155	9.29	47	146	0	12	5
3	2293	2561	5.5	18	37	0	38	3

Table 3. M6 Motorway Junction 10 17.3.1987
Video duration 86 mins

Lane	Manual Count	Computer Count	Counting Errors					
			%	i	ii	iii	iv	v
1	1606	1552	13.8	32	62	0	73	47
2	2389	2275	13.6	81	170	0	50	25
3	2800	2548	12.1	3	152	0	144	41

Errors: i) Vehicles counted as two smaller ones.
ii) Vehicles missed due to occlusion.
iii) Vehicles lost due to the segmentation and noise reduction algorithms.
iv) Vehicles missed due to light changing or activity in adjacent lanes
v) Extra counts due to light changing or activity in adjacent lanes

7. DISCUSSION OF RESULTS

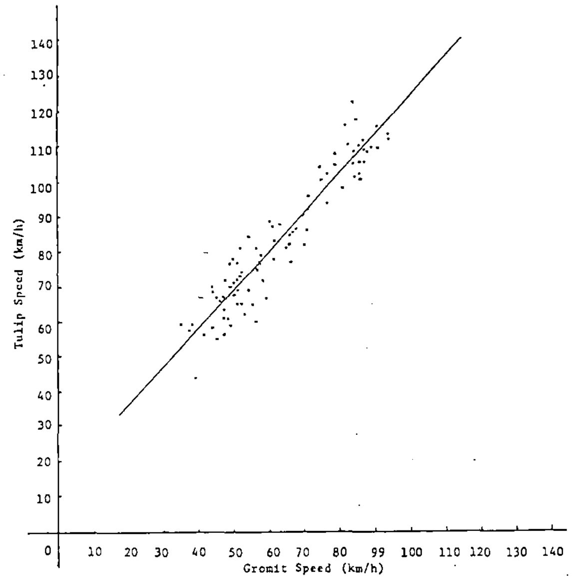
Table 1 contains results of analyses from videos taken by the author. They were all taken from footbridges approximately 10 meters above the carriageway under various lighting conditions. It can be seen that under slowly varying light levels with a favourable camera position a counting error of better than 2 per cent can be expected. This compares favourably with a manual count over a comparable period of time.

Tables 2 and 3 show the results from library videos of a motorway scene, again taken from footbridges for viewing by human operators.

In the first video the lighting was reasonably constant, it being a cloudy morning with slight rainfall. An overall counting error of 6.6 per cent was achieved, the majority of which (over half) were occlusions, which could have been reduced by a better camera angle. Also, the appearance of large bright objects in the scene, such as lorries, caused the camera auto iris to act altering the absolute pixel intensities of the image. This kind of event renders the segmentation parameters temporarily invalid and so causes system errors either by missing vehicles or making a count when no vehicles are present. This problem can be reduced by switching the auto iris off on some cameras.

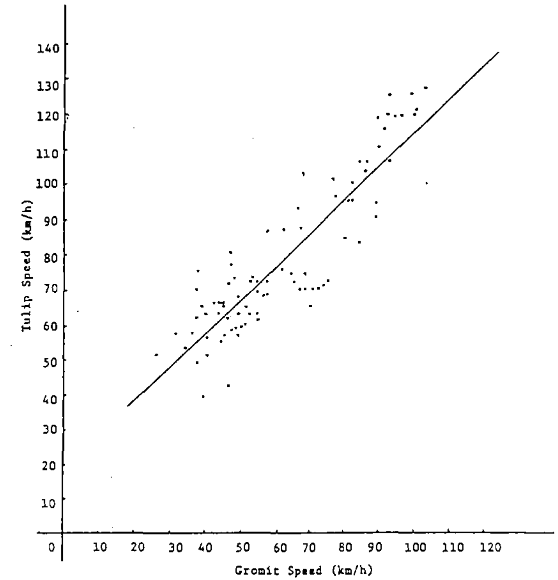
The second video suffered from the same problems as the first,and was additionally degraded in parts by the cameraman wiping the lens as rain began to fall heavily. An overall error rate of 13 per cent was obtained. Both videos were persevered with because they came with associated speed data collected from a loop detector system known as GROMIT. This speed data was in the form of one minute averaged speeds from all three lanes of the motorway. Similar data was thus extracted from the videos by the TTS_1 software. Graphs 1 and 2 show TULIP speed versus GROMIT speed for the two videos respectively.

Graph 1 shows a correlation coefficient of 0.94 with an rms error of 4.5 Kph, which represents an accuaracy of between 15 and 5 per cent over the range 30 to 100 KpH. Graph 2 shows a correlation coefficient of 0.9 with an rms error of 7 Kph which represents an accuracy of between 28 and 6.4 per cent over the range 25 to 110 Kph.



Slope	1.091914
Y intercept	14.93144
Correlation coefficient	0.9439
Standard error	4.531269

Graph 1 : M6 Motorway 7:4:1987, 1 minute averages



Slope	0.9638024
Y intercept	19.02188
Correlation coefficient	0.8977
Standard error	7.062847

Graph 2 : M6 Motorway 17:3:1987, 1 minute averages

8. FUTURE WORK

The TTS_1 software requires some further testing, principally to ascertain the reliability with which vehicle classification can be performed, but also to build up some statistical confidence of the limits of the software with regard to the lighting and weather conditions under which it can be used.

New routines are required to extend the application of TULIP into the analysis of other situations, such as turning counts at cross roads and T-junctions.

The authors are now engaged on a new project to produce a system which will analyse pedestrian traffic flow, a more complex problem than the analysis of road traffic from the point of view of the random nature of pedestrian movements. In other respects the problem is easier, for example, pedestrians do not move as quickly as vehicles and so the requirement to process every video frame can perhaps be relaxed.

In the past year or so a couple of interesting developments have occured which may be of use in this project. They appear to have the potential to increase the speed of execution of primitive image processing functions and increasing the apparent intelligence of the software. This will enable more ambiguous scenes to be analysed and better object discrimination to be performed.

The development of the Transputer seems to present the possibilty of performing primititve processing of images very quickly when used in conjuction with a frame grabber/store. These processed images, or parts thereof, can be passed to a neural network pattern recogniser for further 'high level' processing. The work on neural networks is still highly speculative, but results so far are quite promising. In certain constrained situations, where the set of objects to be identified is relatively small, these networks promise to be much more efficient and reliable than other currently available techniques, especially in the presence of noisey data.

9. AKNOWLEDGEMENTS

The authors would like to thank the Science and Engineering Research Council for providing the funds for this project and Wootton Jeffreys Consultants Ltd for donating the motorway videos.

10. REFERENCES

1. Inigo, R M (1985) Traffic Monitoring and Control using Machine Vision : A survey. IEEE Transactions on Industrial Electronics Vol IE-32, No 3, August.

2. Dickinson, K W and Waterfall, R C (1984) Image Processing Applied to Traffic 1 - A general review . Traffic Engineering and Control, 25 (1) Jan, pp 6-13.

3. Dickinson, K W and Waterfall, R C (1984) Image Processing Applied to Traffic 2 - Practical Experience. Traffic Engineering and Control, 25 (2), February

4. Rourke, A and Bell, M G H (1988) Traffic Analysis using Low Cost Image Processing - A real time system for temporary traffic censuses. TORG Research Report No.75, University of Newcastle upon Tyne.

5. Wootton, H J and Potter, R J (1981) Video Recorders, Microcomputers and New Survey Techniques. Traffic Engineering and Control, 22, pp 213-215.