For presentation at TRL Conference MOVA and Adaptive Traffic Signal Control, Birmingham, February 2005

VIDEO VEHICLE DETECTION AT SIGNALISED JUNCTIONS –
A SIMULATION-BASED STUDY

Richard Allsop*, Amer Ali, David Crosta and Benjamin Heydecker
Centre for Transport Studies
University College London

ABSTRACT

Many existing advanced methods of traffic signal control depend on information about approaching traffic provided by inductive loop detectors at particular points in the road. But analysis of images from CCTV cameras can in principle provide more comprehensive information about traffic approaching and passing through junctions, and cameras may be easier to install and maintain than loop detectors, and some systems based on video detection have already been in use for some time.

Against this background, computer simulation has been used to explore the potential of existing and immediately foreseeable capability in automatic on-line image analysis to extract information relevant to signal control from images provided by cameras mounted in acceptable positions at signal-controlled junctions. Some consequences of extracting relevant information in different ways were investigated in the context of an existing detailed simulation model of vehicular traffic moving through junctions under traffic-responsive signal control, and the development of one basic and one advanced algorithm for traffic-responsive control. The work was confined as a first step to operation of one very simple signal-controlled junction.

Two techniques for extraction of information from images were modelled - a more ambitious technique based on distinguishing most of the individual vehicles visible to the camera, and a more modest technique requiring only that the presence of vehicles in any part of the image be distinguished from the background scene. In the latter case, statistical modelling was used to estimate the number of vehicles corresponding to any single area of the image that represents vehicles rather than background.

At the simple modelled junction, each technique of extraction enabled each of the algorithms for traffic-responsive control of the signals to achieve average delays per vehicle appreciably lower than those given by System D control, and possibly competitive with those that MOVA would give, but comparison with MOVA was beyond the scope of the initial study.

These results of simulation indicate that image analysis of CCTV pictures should be able to provide sufficient information in practice for traffic-responsive control that is competitive with existing techniques. Ways in which the work could be taken further were discussed with practitioners, but have not yet been progressed.

* Centre for Transport Studies, University College London, Gower Street, London WC1E 6BT
rea@transport.ucl.ac.uk www.cts.ucl.ac.uk

1 INTRODUCTION

Traffic-responsive signal control in Britain depends on inductive loop detectors for information about approaching vehicles, as do many systems elsewhere, but one North America system, AUTOSCOPE (see *e g* Michalopoulos *et al* 1992, Panda 1999), has for more than a decade made extensive use of traffic data from analysis of CCTV images of the traffic – including applications in the Birmingham SCOOT system and in the Clyde tunnel. More recently, the French CCTV-based system MediaCity, available from the company Citilog, and based on the control algorithm CRONOS (Boillot *et al* 1992,2000) has also come into use Cameras are inherently more reliable and in many respects easier to install, modify and maintain than loop detectors, and can in principle provide more comprehensive information about traffic, pedestrian as well as vehicular, approaching and passing through the junctions being controlled.

The research reported here used computer simulation to explore the potential for analysis of CCTV images to contribute through algorithms for traffic-responsive signal control to new strategies of control. In order to keep this exploratory project within tractable bounds, it was deliberately decided to confine attention at this stage to vehicular traffic in general, and to explore what could be achieved using data from CCTV images alone and the existing or immediately foreseeable capability of image analysis. In doing so, it was recognised that resulting techniques would need to extend to pedestrians and to distinguishing particular categories of vehicle, that data from image analysis may well be used in combination with data from other forms of detection, and that this potential use of image analysis may stimulate innovation in image analysis itself.

The work was confined as a first step to operation of one very simple signal-controlled junction. Attention was also confined to gathering and use of information about traffic approaching or waiting to enter the junction, rather than including also information about vehicles crossing stoplines, because it is in respect of the former kind of information that there is the greater potential in principle for image analysis to exceed substantially the capability of loop detectors. By comparison, extraction from CCTV images of information about vehicles crossing stoplines is relatively straightforward, and the information concerned is less radically different from what can be obtained from loop detectors than is the case for traffic on approaches to junctions.

2 SCOPE OF MODELLING

An important starting point for the work was the microscopic simulation model SIGSIM (Sha’Aban 2003a.b), which is designed specifically for detailed modelling of vehicles passing through signal-controlled junctions under various forms of control. This model was adapted in straightforward ways to represent not only the positions, speeds and lengths of vehicles, but also their widths and heights, and to receive and act upon stage-change commands from control algorithms to be developed during the work. The vehicle dimensions were generated from normal distributions with means and standard deviations specific to the vehicle types: car, van and heavy goods. Alongside this model of traffic and the signals controlling it, it was necessary to develop a simulation model to generate the outputs of image analysis of CCTV images of the simulated traffic, and software to use these outputs to generate stage-change commands to control the signals.

3 GEOMETRY OF CAMERA POSITION AND APPROACHING TRAFFIC

Fundamental to the analysis and interpretation of CCTV images is the geometrical relationship between position in the two-dimensional (2D) image and position in the three-dimensional (3D) reality. This depends on the position, orientation and field of view of the camera; it is well understood (see *e g* Hartley and Zisserman 2000) and did not require new research. In the context of traffic signal control, there will in practice be many site-specific constraints on the placing of cameras to provide images of traffic on the approaches to junctions. For the purposes of this project, the camera was assumed to be placed vertically above the centre-line of a lane used by approaching traffic, and at or beyond the stopline. This assumption has two advantages: it is conservative in terms of the information contained in the image, because for a given camera height, occlusion of vehicles is maximised when the camera looks down the centreline, and it reduces the geometry from 3D to 2D, thus minimising the intricacies of the well-understood geometry. Multi-lane approaches were not considered but the methods developed extend to these, as they do to offset camera positions. Camera heights were between 5m and 25m, fields of view extending up to 400m upstream of the stopline, and camera resolutions between 512x512 and 2048x2048 pixels were considered. Particular attention was given to the occlusion of vehicles one by another in relation to their dimensions and positions, and to the number of pixels occupied vertically in the image by the visible part of the centre-line of each vehicle. The resulting geometrical model is described by Ali *et al* (1999a) and is illustrated in Figure 1, in which the stream of vehicles shown in the upper part of the diagram gives rise to occlusion as shown in the centre part, and the areas of road surface shown in the lower part obscured in the image by vehicles.

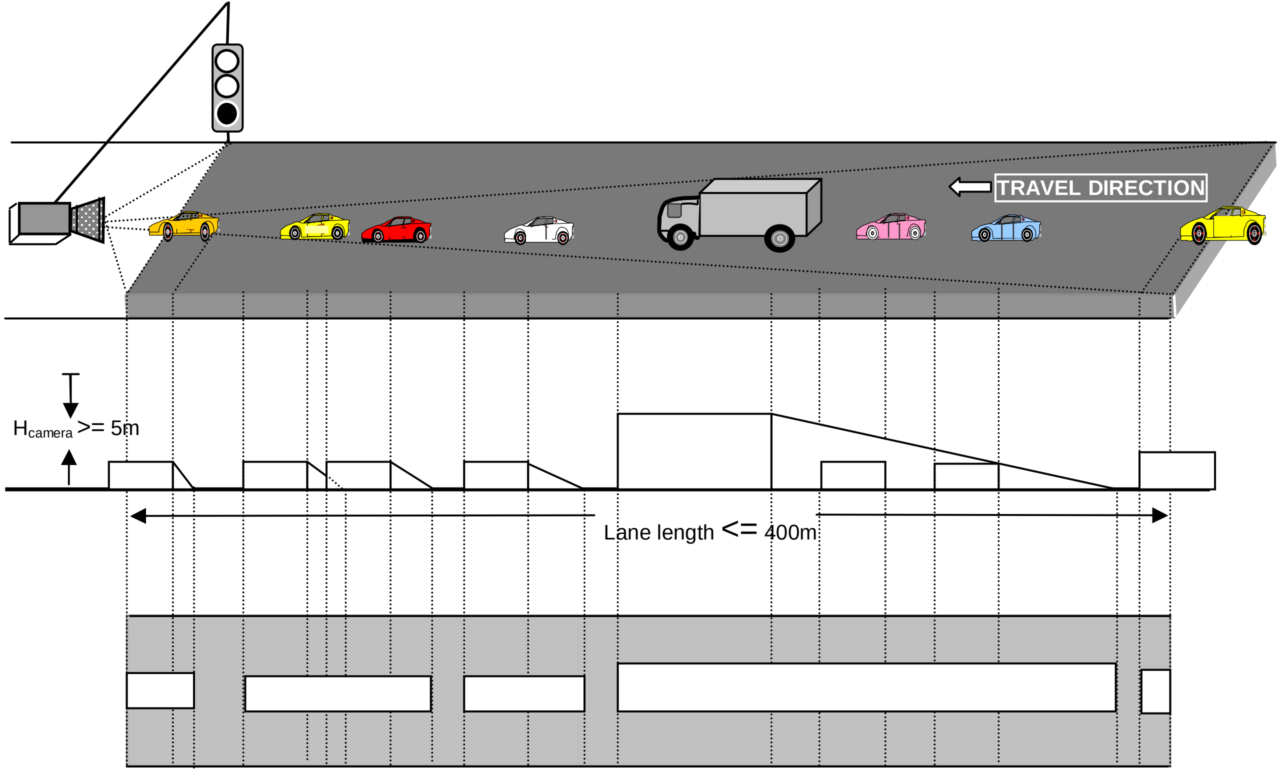

Figure 1 Illustration of camera position and some relevant aspects of the resulting image

4 ASSUMED IMAGE-ANALYSIS CAPABILITIES

This work was concerned not with image analysis itself, but with the use of information that could be obtained by that means. After updating the team's awareness of current and foreseeable image analysis capability with the help of a colleague working at the forefront of high-level image analysis research, it was decided to work with two assumptions, one ambitious and the other modest, about this capability in the context of application in the wide range of conditions to be found at signal-controlled junctions. The ambitious assumption was that, with the assumed geometry, if the visible part of a vehicle's centre-line occupied at least 5 pixels vertically in the image, that vehicle would be distinguished from its surroundings by image analysis (the 5-pixel assumption). The modest assumption was that vehicles could be distinguished sufficiently from their background (with the assumed geometry, the road surface) for the positions on the road corresponding to the front and rear of the image of a vehicle or platoon of vehicles represented by a single area in the image to be estimated, together with the speed of the front (the platoon assumption).

4.1 Distinguishing visible vehicles (DVV)

The geometrical model enables the software to calculate, for any pattern of simulated approaching vehicles in the SIGSIM model, how many vehicles on each approach would be distinguished by image analysis under the 5-pixel assumption. This is an underestimate of the number of simulated vehicles, mainly because of total or partial occlusion, and also because at lower camera resolutions and with longer fields of view, the smallest vehicles at the far end of the field of view occupy less than 5 pixels vertically even when they are completely visible. It should be possible to augment this information by means of best estimates of numbers of occluded vehicles in terms of the numbers and image-sizes of the distinguished vehicles, but this was not attempted in the time available. When vehicles are distinguished in successive images, it should also be possible to estimate their speeds, but that potential information was not used in this research. To do so, it would be necessary to make an assumption as to which of the vehicles distinguished under the 5-pixel assumption would be distinguished clearly enough in successive images for their speeds to be estimated.

4.2 Platoon estimation model (PEM)

The information available under the platoon assumption was used, together with other extractable information about areas of the image representing vehicles, to form the independent variables in the fitting of statistical models for the number of vehicles represented by a single area in the image. Independent variables additional to the position of the front and rear and the speed of the front were the speed of the rear, whether the area reached the front or rear edges of the field of view, whether the controlling signal was green or not, and time within the red or green periods at which the data were captured. The actual number of vehicles represented by each area in the image was known from the traffic simulation, and provided a corresponding value of the dependent variable for model fitting.

Three independent sets of data were generated in which the relevant variables were captured on the last 400m of the junction approach 6 times in each signal cycle (early in, midway through and late in each of the red and green periods) for about 10 minutes, giving about 400 observations in each set, each observation corresponding to an area in one of the captured images.

A wide range of models was considered, using different combinations of the independent variables, and each model was fitted separately to each of the datasets. Criteria used to assess the models were: proportion of variation explained, number and mix of variables included, number of terms whose coefficients were statistically significant, level of significance of these terms, and consistency of goodness of fit and estimated coefficients across the 3 datasets. These criteria pointed to models in which the independent variables were:

F = distance of front end of vehicle or platoon from stopline (m)

L = length of visible part of vehicle or platoon (m)

V = speed of front end of vehicle or platoon (m/s)

Several models containing these variables, their products in pairs and the product of all three were considered, and the model

$$Y = a + b\mathrm{L} + c\mathrm{FL} + d\mathrm{LV} + e\mathrm{FV} + f\,\mathrm{FLV} \qquad (*)$$

was chosen as the *platoon estimation model* (PEM) for the number of vehicles in the platoon.

The coefficients $a, b, \ldots, f$ were estimated for use in the rest of the research by fitting to all 1200 observations as a single dataset, and the fitted model with

$a = 0.6504$ $c = -0.0001088$ $e = 0.0001304$

$b = 0.08379$ $d = -0.004825$ $f = 0.000006896$

explained 82 per cent of the variation in the observed numbers of vehicles. The fit of this model is shown in Figure 2. In practical use, such a model would probably need to be calibrated for each approach where a control system using image analysis based on this type of model was installed at a junction; and the calibration would be specific to the camera position, orientation and field of view. The values of Y given by Equation (∗) are non-integer estimates of the number of vehicles represented by each relevant area in the image of traffic on the approach. The development of the PEM is described fully by Ali *et al* (1999b).

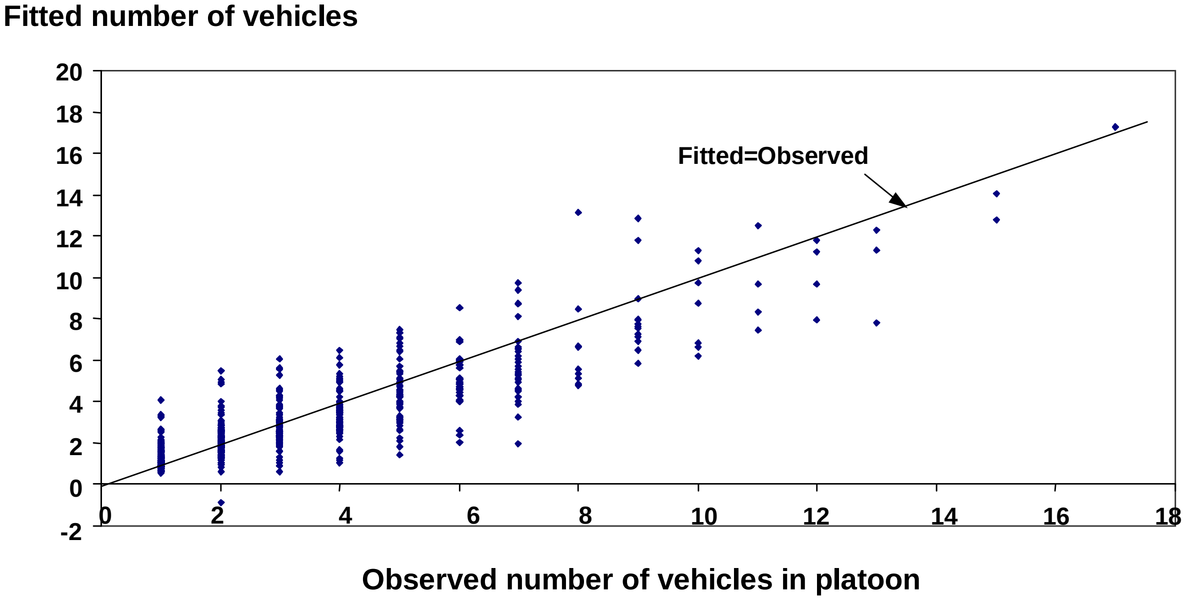

Figure 2. The estimated number of vehicles against the actual number of vehicles in the modelled platoons

5 TRAFFIC-RESPONSIVE SIGNAL CONTROL ALGORITHMS

Two signal control algorithms were used to demonstrate the use of the data that would be available from image analysis under the 5-pixel assumption or the platoon assumption. The first, designated the *preliminary algorithm* was devised by common sense as a simple technique which could be implemented quickly in order to complete the modelling process outlined in Section 2 and enable all parts of the interrelated software to be implemented in simultaneous operation. In the event, it performed well enough to be retained in use throughout the work. The second is an updated and enhanced form of the *dynamic programming* technique DYPIC (Robertson and Bretherton 1974). Both were applied in this research to two-stage control of two streams of traffic. Each can be extended in principle to junctions with more streams and more stages. For the preliminary algorithm, this would require some thought, but should not be computationally demanding. For the dynamic programming algorithm, the logic of extending it is clear, but the computational burden could become formidable. Fixed time control and traffic-responsive control by System D, each of which can be implemented within the SIGSIM traffic model, were used as baselines for comparison of the performance of the new algorithms.

5.1 Preliminary algorithm

The preliminary algorithm is based on a running comparison between the numbers N_r and N_g of vehicles estimated to be present on the approaches to the signals that are currently red and green respectively. N_r is multiplied by the time T_r since the red signal became red to give a quantity which under steady state traffic demand increases in expectation quadratically with T_r. N_g is multiplied by an adjustment factor A whose dimension is time to give a quantity which decreases in expectation as any queue present at the start of green clears, and then fluctuates about a constant mean (under steady state demand) according to the random variation in arrivals. A command to change stage is generated as soon as $T_r N_r$ exceeds AN_g. For a given demand, the average frequency of changes decreases as A increases, and the resulting average delay/vehicle at first decreases and subsequently increases as A increases, passing through a minimum which indicates the delay-minimising value of A for the given demand. Variation of delay with A is the counterpart of the variation of delay with cycle time under fixed time control. N_r and N_g can each be estimated by image analysis. The algorithm and its implementation are described by Ali et al (1999c).

5.2 Dynamic programming algorithm

After reviewing the literature of application of dynamic programming to traffic-responsive signal control, it was decided to base its application in this research upon the pioneering work of Robertson and Bretherton (1974). That work, undertaken when computing was much slower than now and the availability in practice of detailed information about vehicles approaching the signals over several hundred metres was hardly envisaged, was intended for off-line calculation. To keep the computational burden within bounds, it used a time-increment of 5s between decisions to change the signals. To realise the potential offered by data from image analysis, more frequent decisions are required, and the massive increases in available computing speed over the intervening years made it possible to envisage on-line calculations even with these shorter intervals between decisions. Against this background, a backward dynamic programming approach was developed using the time increment of 0.667s on which the SIGSIM traffic model is based. This shorter time increment enabled details of signal control and of the arrival and departure of vehicles to be represented quite closely. The new algorithm was designed to use information about all approaching vehicles already within 400m of the stopline, the longest field of view considered for the purposes of image analysis in this study. Effects of shortening the time increment (initially to 0.5s) upon the optimisation were examined and reported by Heydecker and Boardman (1999), and the algorithm as subsequently used, with data modelled as being obtained by image analysis and with a time increment of 0.667s , is described fully by Heydecker, Boardman and Addison (1999).

6 SIMULATED CONTROL USING OUTPUTS FROM MODELLED IMAGE ANALYSIS

The complete modelling procedure was used to model two-stage control of a junction with two single-lane streams of traffic, and intergreen and minimum green periods of 5s each. After a number of exploratory tests, it was decided that the principal modelling exercises would be based upon two levels of flow: 700 vehicles/h in each stream, representing a degree of saturation of 88 per cent under optimised fixed-time control, and lighter traffic at 300 vehicles/h in each stream. In each case, the traffic composition was 70 per cent cars, 20 per cent vans and 10 per cent heavy goods vehicles. Baselines for assessing performance in terms of average delay/vehicle under the preliminary and dynamic programming algorithms were provided by estimates from SIGSIM of performance under optimised fixed-time control and System D traffic-responsive control.

Control by the preliminary algorithm was simulated using image analysis data first from DVV and then from the PEM. Control by the dynamic programming algorithm was simulated using data from the PEM, but not from DVV because the latter as assumed in this research did not provide the required estimates of vehicle speed. For each algorithm, a baseline for the effectiveness of use of image analysis data in the algorithm was provided by simulating also control on the basis of perfect data - namely the positions and speeds of vehicles as they were in the SIGSIM model.

6.1 Use of outputs from distinguishing visible vehicles

At each time-increment in the SIGSIM model the corresponding numbers of vehicles in the two streams that met the 5-pixel criterion provided estimates of N_r and N_g in the preliminary algorithm. An appropriate value of A for each flow level (60 time increments or 40s for a flow of 700 vehicles/h and 40 time increments or 26.667s for a flow of 300 vehicles/h) was found by modelling for a range of values of A containing the optimum and choosing the multiple of 10 time increments that gave the lowest average delay per vehicle.

6.2 Use of outputs from the platoom estimation model

For use in the preliminary algorithm, the estimates of Y from Equation (∗) in Section 4.2 for the various image areas currently representing traffic in each stream were summed to provide values of N_r and N_g. The dynamic programming algorithm required at each time-increment an estimate of time of arrival at the stopline, if undelayed, for each vehicle estimated to be present on each approach.. To provide such estimates, the value of Y for each image area representing traffic on the approach was first rounded to an integer, upwards or downwards with probabilities such that the expectation of the rounded value was equal to the unrounded value. The estimated positions of the fronts of the vehicles were then distributed over the length of road corresponding to the image area, using a simple set of rules (Ali et al 1999a), and each vehicle was assigned the estimated speed of the front vehicle. For each approach the estimated positions and speeds of all the vehicles estimated to be present on the approach were then used to estimate the time of arrival at the stopline, if undelayed, for each vehicle using a further set of rules (Crosta 1999).

6.3 Example results

The picture given by the range of results set out fully by Ali et al (1999c) and Crosta (1999) can be summarised in the following table, in which average delays per vehicle in simulation runs representing 30 minutes of real time are given for various forms of control. At each level of flow, the same random number seed is used for each form of control. Where data from modelled image analysis are used, the modelled camera height was 10m. The field of view was 100m for input to the preliminary algorithm because the use of data from longer fields of view in that algorithm was found to be counterproductive - possibly because the algorithm does not discriminate between vehicles nearer to or further from the stopline, whereas the former are more relevant to the decision whether to change the signals now. For the dynamic programming algorithm, the field of view is 400m because this algorithm performed better with a longer field of view - probably because it considers the position and speed of each vehicle individually and can therefore benefit from the extra information provided by the larger field of view.

Average delay per vehicle over 30 minutes simulated operation

Method of control	Source of data	Arrival rate in each stream	
		700 vehicles/h	300 vehicles/h
Dynamic programming	Perfect	26.4	11.7
	PEM	26.1	11.7
Preliminary algorithm	Perfect	24.1	11.6
	PEM	23.8	11.8
	DVV	23.3	11.4
System D	Perfect	28.6	13.8
Fixed Time	Offline	30.0	15.0
Cycle time for fixed time control		50s	25s

It can be seen that the two forms of image analysis data each enable a similar level of performance to perfect data and appreciably better performance than under System D. At the heavier flow, the preliminary algorithm gave better performance than the dynamic programming algorithm. It is recognised that control by MOVA would also be expected to improve on control by System D, possibly by a comparable margin, but it was not practicable within this project to arrange for control of signals within SIGSIM by MOVA, although this would be possible in principle.

7 DISCUSSION

Although the work described here was confined to control of a single very simple signal-controlled junction, the modelling procedure established for this purpose is capable of handling more complex single junctions, and of extension to study co-ordinated control of several adjacent junctions. In particular, the platoon estimation model is relevant to the identification of platoons of vehicles moving between adjacent junctions, and estimation of their position and speed and of the numbers of vehicles they contain. As well as achieving findings in its own right, the work has therefore laid foundations for more extensive related research should a context arise in which this would be relevant. Particularly encouraging is the indication from the findings that effective use of image analysis in traffic-responsive signal control need not depend on pressing techniques of image analysis itself to the limits of foreseeable capabilities, but is likely to be achievable by using modest and robust techniques.

8 ACKNOWLEDGEMENTS

The research approach was designed in close consultation with Dr Dennis Robertson, who was in turn in close touch with the consulting firm Oscar Faber, now part of Faber Maunsell, where Peter Yendell provided helpful comments on the work. The authors are grateful to Dr Maria-Alicia Vicencio-Silva for advice about image analysis techniques and capability, and to Dr J D (Puff) Addison and Richard Boardman for their substantial input to the dynamic programming signal control algorithm developed for this study. The study as a whole was funded by the Engineering and Physical Sciences Research Council through Research Grant GR/L25462.

9 REFERENCES

ALI, A.T, D A CROSTA, R E ALLSOP, B G HEYDECKER, D I ROBERTSON and M A VICENCIO-SILVA (1999a). Improved traffic signal control using image analysis - a geometrical model. University of London Centre for Transport Studies Working Paper, University College London (unpublished).

ALI, A.T, D A CROSTA, R E ALLSOP, B G HEYDECKER, D I ROBERTSON and M A VICENCIO-SILVA (1999b). Improved traffic signal control using image analysis - a platoon estimation model. University of London Centre for Transport Studies Working Paper, University College London (unpublished).

ALI, A.T, D A CROSTA, R E ALLSOP, B G HEYDECKER, D I ROBERTSON and M A VICENCIO-SILVA (1999c). Improved traffic signal control using image analysis - a preliminary algorithm. University of London Centre for Transport Studies Working Paper, University College London (unpublished).

BOILLOT, F, J M BLOSSEVILLE, J B LESORT, V MOTYKA, M PAPAGEORGIOU and S SELLAM (1992) Optimal signal control of urban traffic networks. Sixth International Conference on Road Traffic Monitoring and Control, London, April 1992, 75-79. London: Institution of Electrical Engineers

BOILLOT, F, S MIDENET and J C PIERRELEE (2000) Real-life CRONOS evaluation. Tenth International Conference on Road Traffic Information and Control, London, April 2000, 182-186. London: Institution of Electrical Engineers

CROSTA, D A (1999). Improved traffic signal control using image analysis - dynamic programming implementation and results. University of London Centre for Transport Studies Working Paper, University College London (unpublished).

HARTLEY, R and A ZISSERMAN (2000) Multiple view geometry in computer vision. Cambridge: Cambridge University Press.

HEYDECKER, B G and R M BOARDMAN (1999). Optimisation for timing of traffic signals by dynamic programming. 31st Annual Conference of the Universities Transport Study Group, University of York, January 1999.

HEYDECKER, B G, R M BOARDMAN and J D ADDISON (1999). Improved Traffic signal control using image analysis - dynamic programming. University of London Centre for Transport Studies Working Paper, University College London (unpublished).

MICHALOPOULOS, P G, R D JACOBSON, C A ANDERSON and J C BARBARESSO (1992). Field development of Autoscope in the FAST-TRAC ATMS/ATIS programme. Traffic Engineering & Control 33 (9), 475 – 483

PANDA, D P (1999) An integrated vision sensor design for traffic management and control. In: Mastorakis, N E (ed) Recent Advances in Signal Processing and Communications. World Scientific and Engineering Press, 176-185

ROBERTSON, D I and R D BRETHERTON (1974). Optimum control of an intersection for any known sequence of vehicular arrivals. Proceedings of the 2nd IFAC-IFIP-IFORS Symposium on Traffic Control and Transportation Systems, Monte Carlo, September 1974. Amsterdam: North Holland

SHA'ABAN, J (2003a) SIGSIM User Guide Part A: SIGSIM theory (version 2.0). Centre for Transport Studies Working Paper, University College London (unpublished).

SHA'ABAN, J (2003b) SIGSIM User Guide Part B: Serial SIGSIM User Guide (version 2.0). Centre for Transport Studies Working Paper, University College London (unpublished).