The Use of Image Analysis in Determining Some Traffic Flow Characteristics

Gaetano Bosurgi
Dipartimento di Costruzioni e Tecnologie Avanzate – Università di Messina

Antonino D'Andrea
Dipartimento di Costruzioni e Tecnologie Avanzate – Università di Messina

Orazio Pellegrino
Dipartimento di Costruzioni e Tecnologie Avanzate – Università di Messina

Synopsis

The use of over-simplistic models (occasionally borrowed from the field of railway traffic research and suited, therefore, to a constrained mode of driving and rationally organized traffic) to represent driving behaviours has often proved misleading in the most critical of driving conditions.
The most restrictive scenarios have always been thought to be those relative to an isolated vehicle and to a concept of speed linked only to road geometry.
A picture of effective driving conditions has very often been built up by monitoring vehicle speeds at preordained points along the road using instruments such as speed cameras or coils but data relative to variations in speed and trajectory as a function of traffic and environmental complexity have been lost in the process.
Given that the main justification put forward for not using more complex diagnostic tools is their cost, this paper aims to illustrate a procedure designed to derive some of the most representative variables, such as speed, acceleration, distance between vehicles and trajectory using Image Analysis techniques.
In particular, video images recorded using simple cameras positioned inside the vehicle or at fixed points on the road have allowed very reliable data to be collected at very low cost.
The methodology also solves the problems associated with positioning the speed camera at a fixed point, which notoriously gives speed readings that fail to reflect situations of real ranger; it also provides easy-to-read data on interactions between vehicles.
In this paper the aim is not to propose new models of driving behaviour, already illustrated in the lately by the Authors (see the References), but suggest a technique useful to monitor the driving behaviour of unaware users belonging to the traffic flow. This method will allow in the next future to integrate and improve the existing models about with less effort and cost.
The application was tested on a provincial highway in the municipality of Messina known for its high accident rate and provided valuable information, the reliability of which can be established probabilistically.

The Use of Image Analysis in Determining Some Traffic Flow Characteristics

INTRODUCTION

Driving behaviour is a direct consequence of the stimuli that the driver receives from the road infrastructure, from the surrounding environment and from the environment inside the vehicle

While there is near-universal consensus that the former two factors certainly affect driving behaviour, the influence of the vehicle itself, and vehicle subsystems, has only recently begun to attract research attention. We only need think of the effect of radio and cell phone use, not to mention the ways in which the use of driving aids might significantly prejudice vehicle control and raise the need for more elaborate driving manoeuvres.

Driver perceptions of the road context are the result of analyses based on an unconscious selection of information received from the surroundings; the speed at which this analysis takes place will vary with driving ability and experience, the complexity of the road environment and the number of activities being performed simultaneously by the driver (Bosurgi et al., 2003, 2004, 2005).

In order to examine these issues, our study focused on driver behaviour in relation to both road environment and interaction with the vehicle. This was achieved by means of a procedure which required the development of software capable of calculating a series of characteristic variables (such as distance, speed and relative acceleration between the vehicle filming the scene and the target vehicle being pursued along the stretch of road) from recordings made using special equipment.

By manipulating these variables it is possible to derive the speed and absolute acceleration of the target vehicle on a second by second basis; this information can be used to analyse variations in these values and, consequently, to show how the car under study interacts with other vehicles.

The above algorithm was developed using MATLAB® (Gonzales & Woods, 2004) and is based on Computer Vision (CV), or artificial vision, techniques. Computer Vision is a branch of Informatics which specializes in the processing of digital images (films and photographs) in order to extrapolate properties from subjects of study (Meini, 1996) (Iannizzotto, 2004).

This study was mainly motivated by the desire to improve on more traditional approaches and propose a methodology that would make it possible to confirm or reject the validity of some interesting models used to describe driving behaviour but, above all, to study the driving behaviour of persons unaware that they are being monitored.

Such monitoring has traditionally been undertaken by placing speed cameras along the road, but this method has the following limitations:

- it only provides speed measurements at the points at which cameras are located; valuable information on acceleration-deceleration, by far the most useful in studies of road safety, is thus lost;
- it provides no information about vehicle occupation of the cross-section of road;
- it does not, therefore, allow for the influence of road context.

It must be pointed out that the many image processing applications available for use in road/driving contexts are generally employed in the design of driving aids for the so-called ‘smart car’. The methodology described in this paper, however, proposes to analyse the ‘human factor’ and attempt to identify road safety levels through observation of driver behaviour; it is hoped that this information might help suggest design solutions able to guarantee improved road safety.

CAR TRACKING

The difficulty posed by car tracking can be summarized as follows: if there is a frame and the position of the vehicle in the previous frame is known, it’s necessary to find this position in the current frame (Van Leuven et al., 2001).

Observation of a vehicle image will show it to have well-defined contours in areas of vehicle-background transition. The outline produced by these contours is not always regular but can be approximated by a rectangle (box). Identifying the position of a vehicle in a single frame is therefore the same as identifying the box that ‘contains’ it in that same frame. The size of the box will increase and decrease from frame to frame depending on whether the target vehicle is getting closer to or further away from the video camera. Since frames are captured in quick succession, it is reasonable to assume that the box will have moved little from one frame to the next. Thus, at each step the previous position of the box can be taken as the basis from which to derive a reasonably reliable estimate of the current frame’s new position (Stein et al., 2003).

To find the box that approximates the vehicle contour we used an energy minimization technique based on Bayes’ theorem. Since the box we are seeking must have very distinct contours on all sides, one of the most obvious ways to proceed is to define an objective F function that measures the extent of discontinuity of blur along a particular contour. It is clear that such a procedure will involve the use of gradient images determined from film frames. It must be remembered that the higher gradient values are typical of the areas characterized by great variations in intensity, i.e. where contours are distinct. One possible choice for F is as follows:

$$F = \frac{1}{r-l}\int_l^r G_x(t,y)\,dy + \frac{1}{r-l}\int_l^r G_x(b,y)\,dy + \frac{1}{b-t}\int_t^b G_y(x,l)\,dx + \frac{1}{b-t}\int_t^b G_y(x,r)\,dx$$

where (t, l) and (b, r) are the coordinates of the top-left and bottom-right vertices of the box respectively, while G_x e G_y are the respective vertical and horizontal gradient images. Any addendum to the second member of this formula represents an integral mean and the function to be integrated is the component of the orthogonal gradient vector at the side of the box (contour) along which the integral function is calculated (Figure 1).

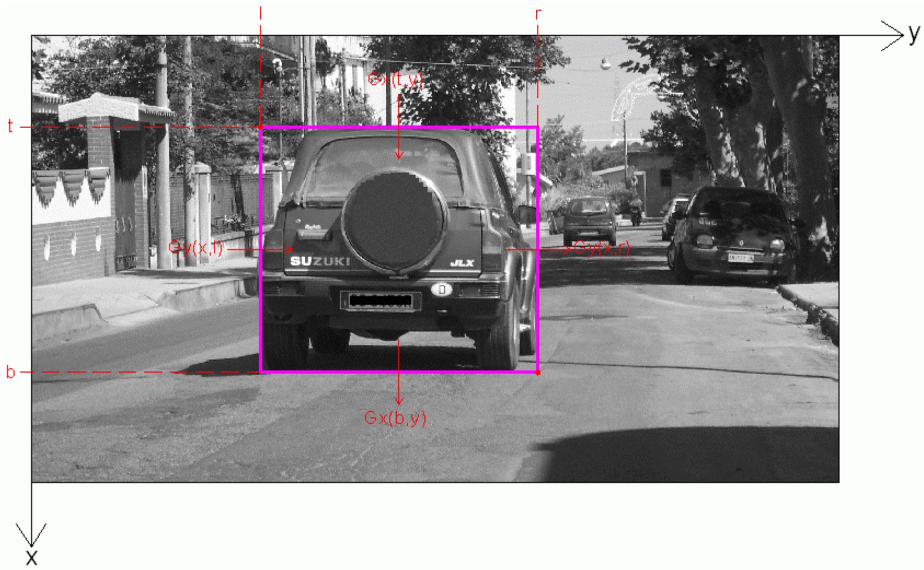

Figure 1: Segmentation of target vehicle images

The objective F function can be correlated with a Bayesian plausibility function using the Gibbs/Boltzman distribution. The latter derives from statistical mechanics and is a product of the insight that lower energy states are more likely than higher energy states. Thus, if the energy term is defined as $E = -F$ and $u = [t \quad b \quad l \quad r]^T$ indicates the four-component vector that identifies the position of the box, then the Bayesian likelihood that the box is in position u, given the information available from the image I, can be expressed as follows:

$$P(I\,|\,u) = \frac{1}{k}e^{-E(u,I)}$$

where k is a normalization factor that effectively makes the first member a probability (i.e. a number between 0 and 1).

In order to estimate *a posteriori* probability using Bayes' theorem, it is necessary at this point to introduce a term which indicates an *a priori* probability. This can be done by observing that the box which approximates the vehicle contour will not move much from its position in the previous frame and thus a simple way of defining P(u) *a priori* is to use a Gaussian probability density function based on the previous u_0 position occupied by the box:

$$P(u) = N(u_0, \Sigma) = \frac{1}{\sqrt{(2\pi)^4 \cdot \mathrm{Det}(\Sigma)}}e^{-\frac{1}{2}(u-u_0)^T \Sigma^{-1}(u-u_0)}$$

Since the Gaussian function has its peak at u_0, the probability that the multidimensional aleatory variable u will fall around u_0 is higher than the probability that u will fall within the same range but far from u_0. According to probability theory terminology, u_0 is the mean vector and Σ the covariance matrix. It must be remembered that the latter is a symmetrical matrix with variances on the main diagonal and cross-covariances or cross-variances of the aleatory variable $u = [t \quad b \quad l \quad r]^T$ outside it:

$$\Sigma = \begin{bmatrix} \sigma_t^2 & \sigma_{bt} & \sigma_{lt} & \sigma_{rt} \\ \sigma_{bt} & \sigma_b^2 & \sigma_{lb} & \sigma_{rb} \\ \sigma_{lt} & \sigma_{lb} & \sigma_l^2 & \sigma_{rl} \\ \sigma_{rt} & \sigma_{rb} & \sigma_{rl} & \sigma_r^2 \end{bmatrix}$$

In our case, the covariance matrix measures the extent of uncertainty inherent in the assumption that the box moves little from one frame to the next.
The *a posteriori* probability $P(u \mid I)$ is given by the Bayes formula:

$$P(u \mid I) = \frac{P(I \mid u) \cdot P(u)}{c}$$

where c is the normalization constant.
The position of the box in the current frame is ultimately obtained by calculating the vector u that maximizes $P(u \mid I)$ *a posteriori* and takes the name of maximum *a posteriori* (MAP):

$$u = \arg\max_u P(u \mid I)$$

By manipulating the last equations, we also have:

$$u = \arg\max_u \frac{P(I \mid u) \cdot P(u)}{c} = \arg\max_u \frac{e^{-\left[E(u,I) + \frac{1}{2}(u-u_0)^T \Sigma^{-1}(u-u_0)\right]}}{k \cdot c\sqrt{(2\pi)^4 \cdot \mathrm{Det}(\Sigma)}}$$

which is equivalent to minimizing the following sum of energy terms:

$$u = \arg\min_u \left[E(u,I) + \frac{1}{2}(u-u_0)^T \Sigma^{-1}(u-u_0)\right] =$$
$$= \arg\min_u \left[E_{veros} + E_{prior}\right] = \arg\min_u \left[E_{post}\right].$$

Briefly, to estimate the MAP it is necessary to minimize the sum E_{post} of two energy functions: $E_{veros} = E(u,I)$, which measures how probable an image I is given u, and $E_{prior} = 0.5 \cdot (u-u_0)^T \Sigma^{-1}(u-u_0)$, which codifies *a priori* knowledge about u.

Estimating the distance and relative speed of the target vehicle

The following paragraph describes a procedure designed to estimate, on a frame-to-frame basis, the distance and relative speed of a vehicle being pursued by a car equipped with a video camera. To be able to do this, it is essential to know the position and outline of the target vehicle in each of the frames; this information can be obtained using the car tracking technique illustrated in the previous paragraph. The standard scenario is: flat road surface and video camera optical axis parallel to the road surface. The scheme outlined in Figure 2 shows:

- (A) = trial vehicle equipped with videocamera V (on the rear-view mirror, for example) located at a height of H_V with respect to the road surface;
- (B) = target vehicle pursued by (A);
- Z = distance to be estimated between the video camera and rear of target vehicle (B);
- I = image plane located at focal distance f from the videocamera V;
- O = point traversed by the road horizon line on the image plane;
- P = projection onto the image plane of contact point C between rear wheel of vehicle (B) and road surface.

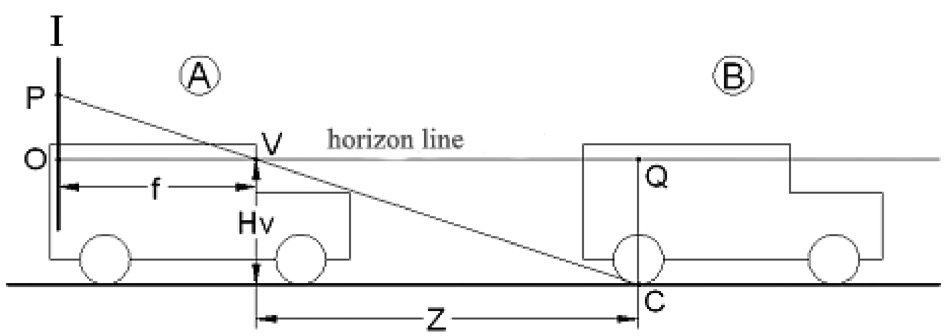

Figure 2: Scheme for evaluating distance Z between trial vehicle and target vehicle.

From the similarity of triangles VQC and VOP we can write:

$$\frac{H_V}{Z} = \frac{\overline{OP}}{f}$$

from which the distance Z is obtained:

$$Z = \frac{fH_V}{\overline{OP}}$$

with f and $\overline{OP}$ in pixels and H_V and Z in metres. The last equation shows that an increase in the $\overline{OP}$ segment, which is derived from the image after identification of the horizon line and vehicle outline, corresponds to a decrease in distance Z; the target vehicle is thus getting closer to the trial vehicle recording the scene. If an error of n pixels is made in estimating $\overline{OP}$, the consequent error Z_{err} in distance Z is equal to:

$$Z_{err} = Z - Z_n = Z - \frac{fH_V}{\overline{OP} + n} = Z - \frac{fH_V}{\frac{fH_V}{Z} + n} = \frac{nZ^2}{fH_V + nZ}$$

and since we usually have $n \cong 1e$, $fH_V >> nZ$ we ultimately obtain:

$$Z_{err} \cong \frac{n}{fH_V} Z^2$$

i.e. the error Z_{err} increases with the square of the distance.

Once the distance Z has been calculated, it is also possible to estimate the relative speed between the trial vehicle and the target vehicle. To this end, H will indicate the actual height (or W the actual width) in metres of the latter and h and h' will be the heights (or w and w'the widths) in pixels projected onto the image when the target vehicle is at distance Z and Z' respectively (Figure 3). Analogously to preceding equations, we have:

$$Z = \frac{fH}{h}; \qquad Z' = \frac{fH}{h'}$$

Relative speed can be expressed as the relationship between the variation in distance $\Delta Z = Z' - Z$ and the time interval Δt in which the variation occurs:

$$v = \frac{\Delta Z}{\Delta t} = \frac{Z' - Z}{\Delta t}$$

which, v becomes:

$$v = \frac{\frac{fH}{h'} - \frac{fH}{h}}{\Delta t} = \frac{\frac{fH(h - h')}{h'h}}{\Delta t} = \frac{h - h'}{h'} \frac{1}{\Delta t} \frac{fH}{h} = \frac{h - h'}{h'} \frac{1}{\Delta t} Z$$

and considering:

$$s = \frac{h - h'}{h'}$$

we ultimately get:

$$v = \frac{s}{\Delta t} Z$$

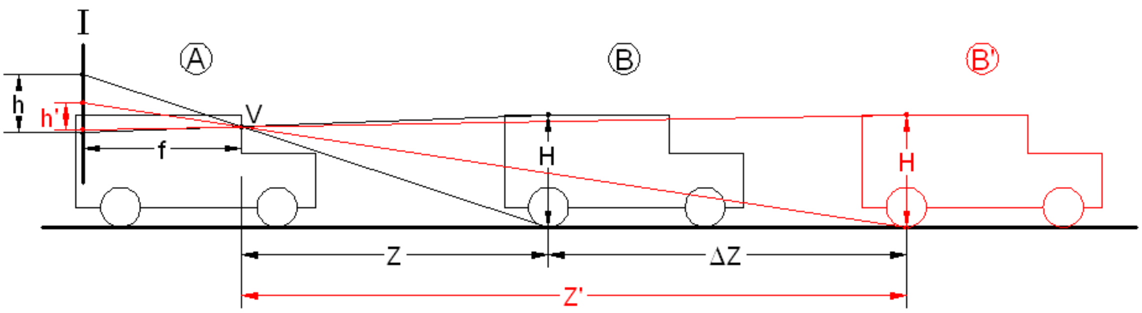

Figure 3: Scheme for estimating relative speed v between trial vehicle and target vehicle.

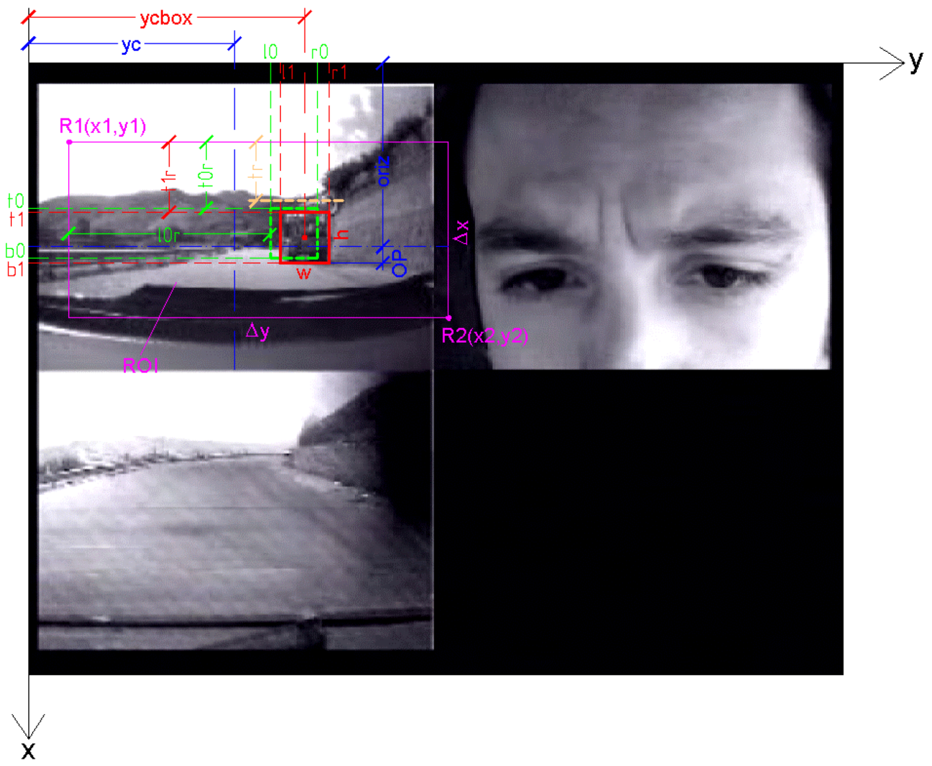

Figure 4: Graphic interface for studying the telemetry of the instrumentation-equipped vehicle.

Interpretation of results

The principal target ha been to organise a technique that permit to monitor some variables as speed, acceleration, position, etc., of the vehicles belonging to the traffic flow with the aim to evaluate the vehicle interaction.

About the experimental strategy, it is necessary to highlight that this technique has been the only way to quantify the driving behaviour of not instrumented vehicles with so low cost. Furthermore, the trial has been intentionally simple, as regards the road length and the manoeuvre implicated, to calibrate the procedure in the best way, without mistakes in the quantification of the speed or distance.

Another experimentation with a fixed camera would permit to understand the potentiality of the proposed method. Nevertheless, the analytic procedure is the same and, rather, the difficulties increase with the cameras inside the instrumented car.

This paragraph aims to show how the output data derived from our experimentation and illustrated in the figures below can be used to analyse driving behaviour. Having determined the main variables characteristic of a journey, the next step is to interpret the motion of the trial vehicle and that of the other vehicles in the three-dimensional dynamic world.

Figure 5 shows that the relative longitudinal distance Z_m between the trial vehicle (henceforth Pr) and the target vehicle (henceforth Ob) decreases steadily from one frame to the next. This means that Pr gradually gets closer to Ob.

The graph in figure 6 showing relative transversal distance Y_t, makes it possible to see which lanes the vehicles are moving in. Note that although there is limited scope for a vehicle to move sideways, the measurement of any movement in a sideways direction is of fundamental importance in interpreting its motion.

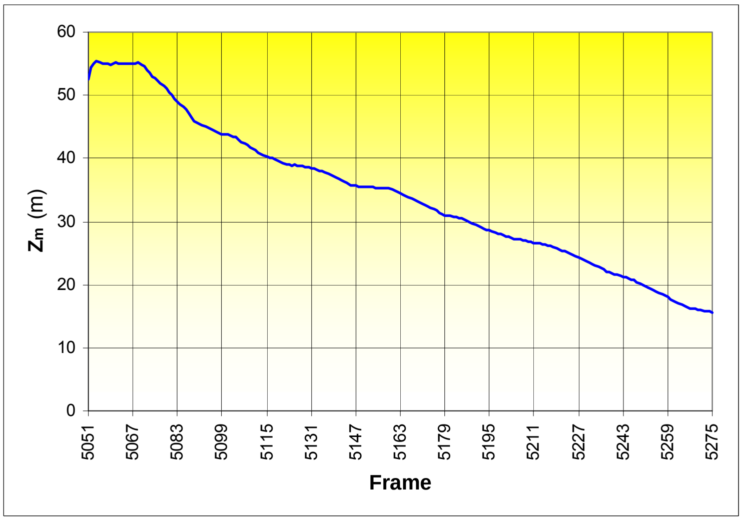

Figure 5: Relative longitudinal distance

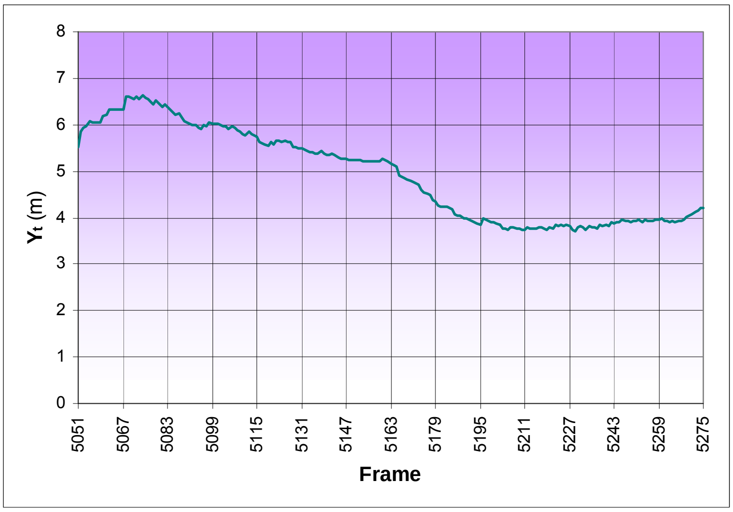

Figure 6: Relative transversal distance

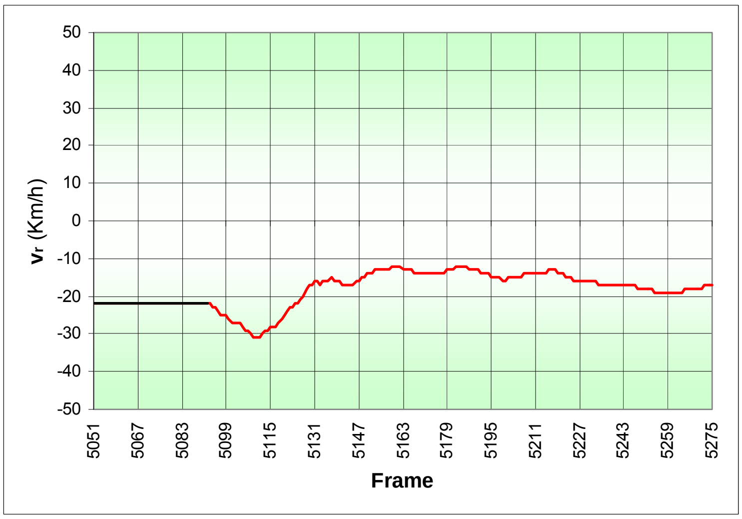

Figure 7: Relative speed

By monitoring this position over a period of time it is possible to directly identify the lane in which the vehicle is moving and, in particular, to distinguish any change in lane from the small steering actions that occur within a single lane. In our case, Y_t always assumes high values, suggesting the two vehicles to be moving in different lanes. The highest Y_t values (around 6.5÷5.0 m) are observed in the initial stage of tracking , indicating that Pr is in the fast lane at this point and that the overtaking manoeuvre starts at a bend in the road. Subsequently, Y_t decreases until it reaches values of 3.8÷4.0 m, suggesting Pr to still be in the fast

lane and to be proceeding in a straight line.
The above is confirmed by observing the section of video in which the tracking occurs. This clearly shows Pr performing an overtaking manoeuvre that starts on a bend (in particular, Ob is already negotiating the curve) and proceeding in a straight line, in line with the progress of Z_m and Y_t illustrated in Figures 6 and 7 respectively.
Analysis of the remaining graphs (Figures 7-13) will also help establish what kind of overtaking manoeuvre is involved. As is known, there are two basic scenarios for driver behaviour during an overtaking manoeuvre:

1. a car proceeding along the inside lane encounters another car ahead proceeding at slower speed and is forced to slow down and line up behind it (adjusting its own speed to that of the slower car); subsequently, as soon as the oncoming traffic lane is perceived to be free, the driver accelerates and overtakes (overtaking with acceleration);
2. a driver reaches another car proceeding at a slower speed and, seeing the oncoming traffic lane to be free, performs the overtaking manoeuvre without reduction in speed (overtaking at speed). Our study examines the second scenario. Figure 9 shows the speed v_{pr} of the trial vehicle (the vehicle performing the overtaking manoeuvre) to initially be constant (at 80 Km/h) and subsequently to decrease slightly. In addition, Figure 11 shows acceleration to remain around zero and to assume slightly negative values on the whole, in line with the trend for v_{pr}.

Let us look at the relative speed v_r graph in Figure 7. The values of v_r are always negative, which indicates Pr to be proceeding at a consistently higher speed than Ob (also compare graphs in Figures 7 and 12).

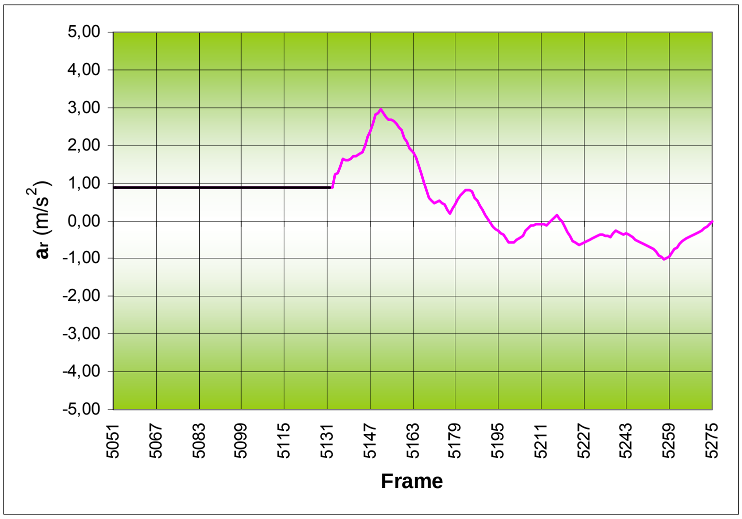

Figure 8: Relative acceleration

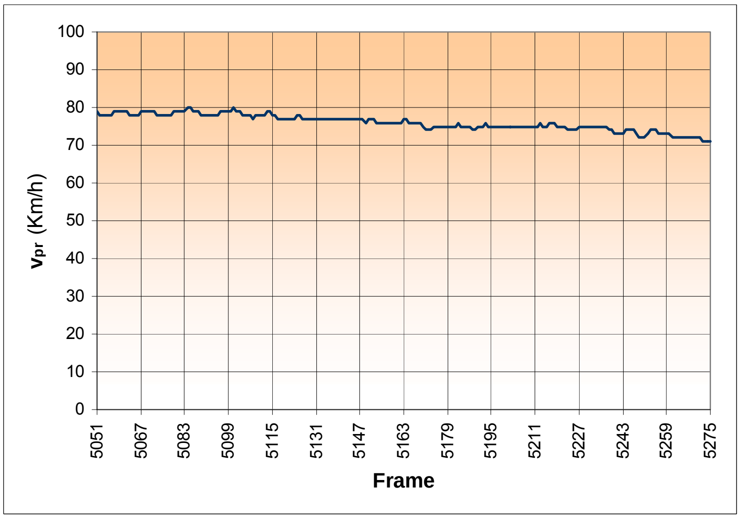

Figure 9: Speed of trial vehicle

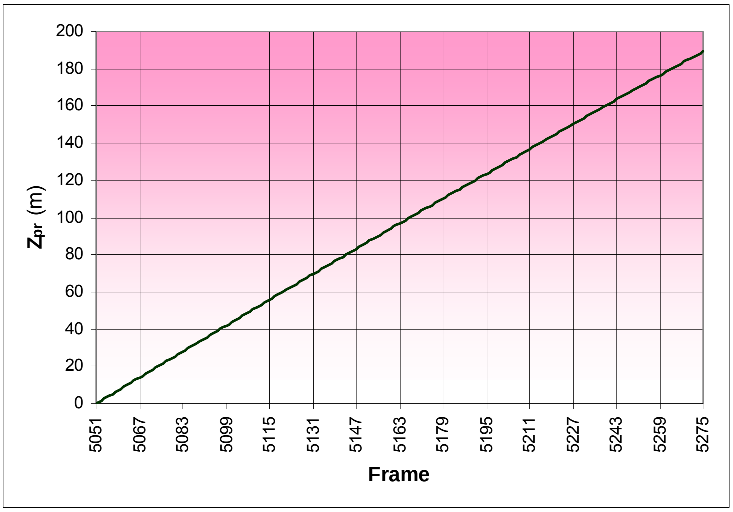

Figure 10: Space covered by the trial vehicle from the start of the sequence under analysis

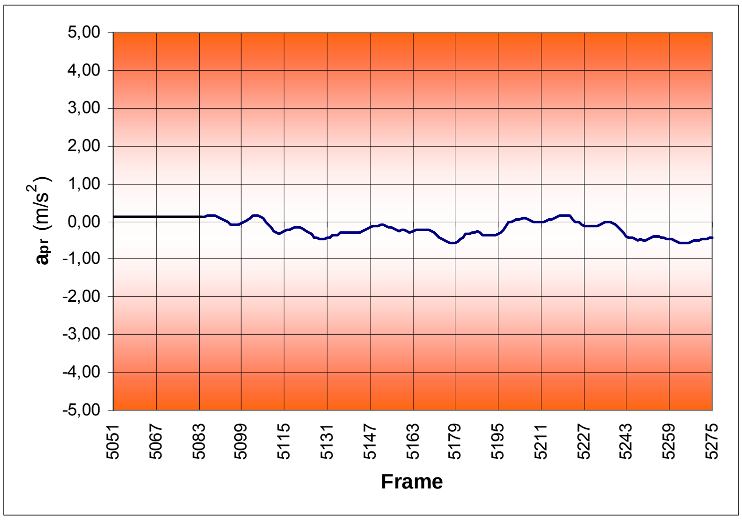

Figure 11: Trial vehicle acceleration

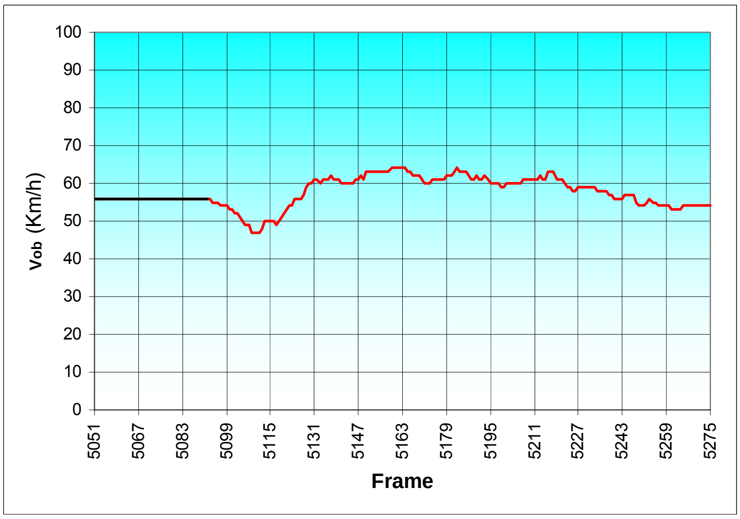

Figure 12: Target vehicle speed

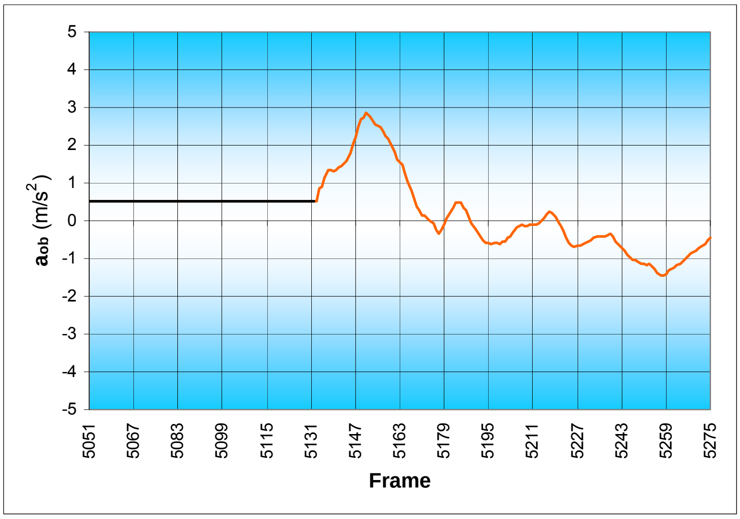

Figure 13: Target vehicle acceleration

On the stretch of road where v_r goes from around -20 Km/h to around -30 Km/h (frame 5093÷5110), i.e. where absolute relative speed values increase, Pr proceeds even faster than Ob; indeed, on this stretch v_{pr} is practically constant and v_{ob} decreases.

Unlike v_r, relative acceleration a_r presents both positive and negative values (Figure 8). The convention for the sign for a_r is similar to that adopted for v_r, i.e. positive a_r values indicate the acceleration of Ob to be greater than that of Pr.

CONCLUSIONS

This paper has illustrated the use of a tool based on Computer Vision that is able to monitor the position of vehicles within the field of vision of the mobile video camera recording the scene as well as their speed and acceleration both relative to the trial vehicle and absolute. This car tracking system, as it is known, constitutes the basis for interpreting motion and, consequently, for analysing driver behaviour in relation to the road environment and interaction with other vehicles.

One of the greatest difficulties in vehicle behaviour analysis has been identifying the lane occupied by the moving vehicle. Identifying this lane depends partly on the position of the vehicle from which the video recording is being made. We were able to identify it by measuring the distance perpendicularly to the main direction of motion. A vehicle has limited freedom to move in a sideways direction. Even when it stays in lane, its motion is accompanied by slight steering actions and it can occasionally change lanes. It is difficult to distinguish between a small steering adjustment and a (slow) lane change when analysis is based only on the relative speed of the vehicle. To make this distinction, we found it more appropriate to consider the vehicle's relative position rather than its speed.

We believe the methodology proposed to constitute a useful practical and research tool for the following reasons:

- It is indispensable in the continuous evaluation of magnitudes of movement, speed and acceleration of drivers who are unaware that they are being monitored;
- It is possible to use a knowledge of these magnitudes for numerous vehicles to evaluate interaction between them:
- It is possible to compute statistics for as large a sample of vehicles as desired;
- The limitations of speed cameras, i.e. their ability to provide measurements only at certain points in the road, are overcome;
- Models of driving behaviour proposed in literature can be verified;

REFERENCES

Gonzalez R. C., Woods R. E., Eddins S.L. (2004). *Digital Image Processing Using MATLAB®*. Prentice Hall, New Jersley, USA.

Bosurgi G., D'Andrea A., Pellegrino O. (2003). *Il VLI (Visual Load Index) quale nuovo indicatore per il controllo dei tracciati stradali*. XIII Convegno Internazionale SIIV, Padova 30-31 Ottobre 2003.

Bosurgi G., D'Andrea A., Pellegrino O. (2004). *Visual Load Index in roads*. Transportation Research Board 2004, Washington D.C., USA.

Bosurgi G., D'Andrea A., Pellegrino O. (2005). *A methodology to study driving behaviour based on the visual activity*. 83rd Transportation Research Board Annual Meeting 2005, Washington D.C., USA

Iannizzotto G. (2004) *Appunti di Calcolatori Elettronici 2*. VISILAB – Computer Vision Laboratory - Università di Messina. http://visilab.unime.it/visilab.

Meini G. (1996). *Enhancement delle immagini digitali*. Gruppo Editoriale Infomedia, http://www.giorgiomeini.it.

Van Leuven J., van Leeuwen M. B., Groen F. C. A. (2001). *Real-time vehicle tracking in image sequences*. Budapest, Hungary, May 21-23, 2001.

Dellaert F., Thorpe C. (1997). *Robust car tracking using Kalman filtering and Bayesian templates*. SPIE Conf. on Intelligent Transportation Systems, Pittsburgh, USA.

Stein G. P., Mano O., Shashua A. (2003). *Vision-based ACC with a single camera: bounds on range and range rate accuracy*. IEEE Intelligent Vehicles Symposium (IV2003), June 2003, Columbus, OH.