ESTIMATING VEHICLE VELOCITY USING RECTIFIED IMAGES

Cristina Maduro, Katherine Batista, Paulo Peixoto and Jorge Batista*
ISR-Institute of Systems and Robotics, Department of Electrical Engineering, FCTUC
University of Coimbra, Coimbra, Portugal

Keywords: Image rectification, vehicle velocity, video sequences.

Abstract: In this paper we propose a technique to estimate vehicles velocity, using rectified images that represent a top view of the highway. To rectify image sequences captured by uncalibrated cameras, this method automatically estimates two vanishing points using lines from the image plane. This approach requires two known lengths on the ground plane and can be applied to highways that are fairly straight near the surveillance camera. Once the background image is rectified it is possible to locate the stripes and boundaries of the highway lanes. This process may also be used to count vehicles, estimate their velocities and the mean velocity associated to each of the previously identified highway lanes.

1 INTRODUCTION

The incessant advances in camera technology along with the constant improvement in areas such as computer vision have lead to the development of automatic and robust methods to estimate vehicle velocity. However, this task is problematic when the image sequences do not preserve length ratios and parallelism between lines. This difficulty can be easily solved by the creation and employment of virtual images which preserve the referred characteristics, acknowledged as rectified images. These rectified images by representing a ”top view” of the observed scenario, simplify the task of estimating velocity in traffic surveillance systems. This process demands the estimation of two vanishing points using lines from the image plane. Nevertheless, it is important to state that this procedure is constrained to the precision with which the required vanishing points are estimated. Thus, a robust RANSAC (Fischler and Bolles, 1981) based algorithm is applied in the estimation of the necessary vanishing points. These are necessary to calculate the homogeneous representation of the vanishing line. This vanishing line is used in the calculation of the projective transformation that rectifies the image sequence. Once attained the required projective transformation it is then possible to rectify video sequences, and therefore estimate vehicle velocity and extract lane topology. The highway lane boundaries can be easily located once identified the position of the striped lines on the rectified background image. Seeing as the striped lines follow a periodic distribution, these can be located by applying the autocorrelation function to each line of the rectified background image. In order to estimate vehicle velocity, a Kalman filter (Kalman and Bucy, 1960) based tracking system is employed to infer future vehicle positions given a sequence of images. The estimation of an object's displacement or motion using information extracted from two consecutive images can be obtained using Lucas-Kanade's optical flow method (Lucas and Kanade, 1981). Nonetheless, we chose to employ a Kalman filter to correct the estimation provided by the previously referred method. Given the fact that the output of this procedure is represented in pixels per frame it is necessary to estimate a scale factor that relates distances on the ground plane with distances on the image plane. Once obtained the required scale factor and known the video's framerate, this procedure presents the object's velocity in the desired units. This paper is organized in four main sections. The first, named Image Rectification, focuses on the method that originates the required rectified images. The second section is referent to the procedure applied to estimate the necessary scale factor, while the third section presents the process that estimates vehicle velocity. To conclude, several results, such as, rectified images and data referent to velocity estimation are presented.

*This work was supported by BRISA, Auto-estradas de Portugal, S.A.

Maduro C., Batista K., Peixoto P. and Batista J. (2008).
ESTIMATING VEHICLE VELOCITY USING RECTIFIED IMAGES.
In *Proceedings of the Third International Conference on Computer Vision Theory and Applications*, pages 551-558
DOI: 10.5220/0001088205510558

2 IMAGE RECTIFICATION

The perspective transformation associated to image formation, distorts certain geometric properties, such as length, angle and area ratios. Due to this fact, the employment of video or image sequences in traffic surveillance is challenging, in particular for the task of vehicle velocity estimation. However, this problem can be solved by using rectified images that restore the lost geometric properties to the images of the monotorized scenario. A rectified image can be attained by estimating a homographic transformation. This estimation could be acquired by using the intrinsic and extrinsic camera parameters. Unfortunately, the surveillance cameras are uncalibrated and therefore, these parameters are unknown. Consequently, several methods have been developed in order to automatically restore geometric properties to objects moving on a ground plane. Namely, D. Dailey (Dailey and Cathey, 2005) presents a method that estimates the location of one vanishing point in order to calibrate the surveillance camera and achieve the required images. However, this method presupposes the knowledge of one of the angles of orientation of the surveillance camera, and therefore cannot be applied to all surveillance systems. On the other hand, in (Schoepflin and Dailey, 2003), associated with T. Schoepflin, D. Dailey presents a method that requires the estimation of two vanishing points from lines that are parallel and orthogonal to the road. This method estimates the camera orientation and focal length, though the height at which it is located is not automatically estimated. L. Grammatikopoulos, G. E. Karras and E. Petsa in (Lazaros Grammatikopoulos, 2002), present a method to measure vehicle speed using rectified images. This approach determines one vanishing point and requires the knowledge of one known length one the ground plane. Nevertheless, this method does not rectify images from cameras that aren't aligned accordingly to an axis parallel to the direction of motion. On the other hand, B. Bose and E. Grimson in (Bose and Grimson, 2003), present a method similar to the method employed in this study. The method proposed by Bose and Grimson achieves metric rectification of the ground plane by tracking two objects that travel with constant and possibly unequal speed. In this paper, a method presented by D. Liebowitz and A. Zisserman (Liebowitz and Zisserman, 1998) is successfully employed in the rectification of images. This technique requires the estimation of two vanishing points and the prior knowledge of two angles on the ground plane. Given the nature of a roadway structure, i.e. the large amount of parallel and perpendicular lines, these parameters can be easily obtained. In a general manner, this method estimates the projective transformation by establishing three matrices or transformations.

$$H = H_s.H_a.H_p \qquad (1)$$

where H_s represents the similarity transformation, H_a the affine and H_p the pure projective transformation. Each one of these transformations is responsible for the reinstatement of certain geometric and metric properties and can be achieved using known parameters on the image and ground planes. Namely, the pure projective transformation is responsible for restoring line parallelism and area ratios to the scenario. This transformation can be easily acquired by estimating the homogeneous representation of the vanishing line. Once known the location of two vanishing points, this representation is quite straightforward, as can be seen in the following equation:

$$l = \begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix} = v_1 \text{ x } v_2 \qquad (2)$$

where l is the homogeneous representation of the vanishing line and v_1 and v_2 the vanishing points that are represented on the upper left box in Figure 1.

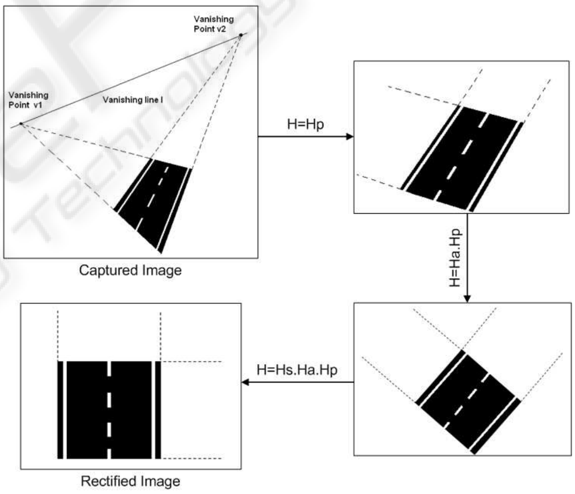

Figure 1: Stages of the rectification process.

Therefore, the pure projective transformation can be represented by the following matrix:

$$H_p = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ l_1 & l_2 & l_3 \end{bmatrix} \qquad (3)$$

where H_p represents the referred pure projective transformation. Hence, a correct estimation of this transformation relies on the accurateness of the location of the vanishing points. These are obtained by applying the Hough transform to edges extracted from the imaged highway lanes and to edges identified on the foreground image.

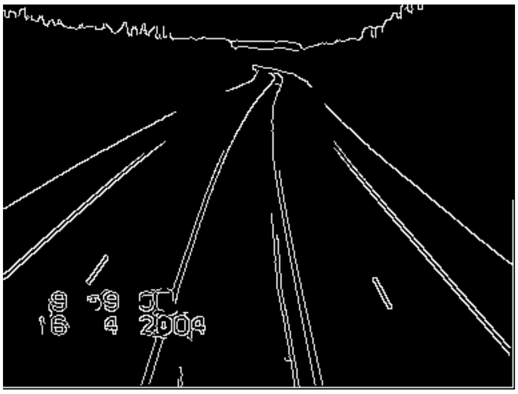

Figure 2: Edges detected by applying the Canny Edge Detector to the Background image.

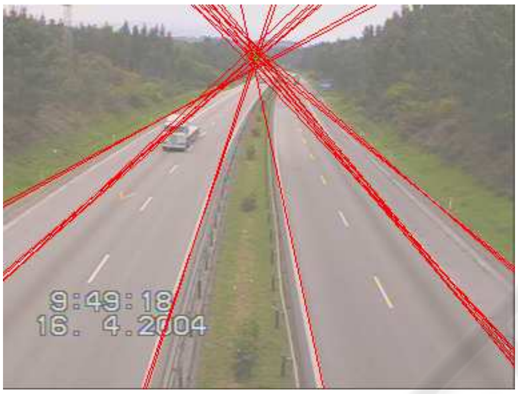

Figure 3: Lines identified using the edges represented in Figure 2.

Figure 2 represents an example of the edges detected from a background image, while Figure 3 illustrates the lines extracted from the previous image. However, edges detected from the background image may contain unnecessary edges. Due to this fact, an activity map is estimated. An activity map is an image that represents the regions that contain movement throughout the image sequence, i.e. the regions of interest. By applying an AND operator to the image shown on Figure 2 and to the activity map, results an image that contains only the sought edges. By definition a vanishing point is the intersection point of lines on the image plane that represent parallel lines on the ground plane. Nevertheless, given the fact that these lines do not intersect on an exact point, the vanishing point is situated on the point whose distance is minimal to each one of these lines. Therefore, one of the required vanishing points is obtained applying the least squares method.

To determine the second vanishing point it is necessary to identify edges that are predominately horizontal on the foreground image. This image can be attained by subtracting each frame to the background image and shows segmented vehicles. Figure 4 illustrates edges detected from the foreground, i.e. segmented vehicles, using Sobel Detector that identifies horizontal edges. The lines represented in Figure 5 where obtained from the edges represented in Figure4 and are only a few of the used in the calculation of the second vanishing point.

Figure 4: Edges detected by applying the Canny Edge Detector to the Background image.

Figure 5: Lines identified using the edges represented in Figure 4.

Given the fact that a great amount of lines obtained from the foreground image are outliers, the estimation of this vanishing point using the least squares method is erroneous. Thus, a RANSAC based algorithm is used in this estimation. However, it is important to state that the employed RANSAC algorithm must contemplate the fact that the vanishing point might be situated at infinity. In order to do so, it is necessary to adopt a 3D homogeneous representation for the extracted lines. This form of representation takes into account points at infinity. Each iteration of the RANSAC algorithm estimates a possible vanishing point using equation (4). The vanishing point with the largest number of inliers is taken as the correct vanishing point.

$$p = l_1 \text{ x } l_2, \tag{4}$$

where l_1 and l_2 represent the homogeneous coordinates of two lines. The accuracy of this estimation is crucial, and seeing as this algorithm is highly robust, though computationally heavy, it presents fine results in the estimation of the required vanishing point.
On the other hand, the affine transformation reinstates angle and length ratios of non parallel lines, and can be obtained using two known angles on the ground plane as explained in (Liebowitz and Zisserman, 1998). This approach estimates two parameters α and β using constraints on the ground plane. These parameters represent the coordinates of the circular points on the affine plane. Liebowitz and Zisserman in (Liebowitz and Zisserman, 1998) propose three types of constraints:

- A known angle on the ground plane;
- equality of two unknown angles;
- a known length ratio.

Given the orthogonal structure of the highway lanes, we chose to employ the first constraint in this algorithm, i.e. a known angle on the ground plane. Each known angle on the ground planes defines a constraint circle. This fact is quite useful seeing as α and β lie within this circle represented on a complex space defined by (α,β). Therefore, in order to obtain the required parameters one may estimate the intersection of two constraint circles obtained using two different known angles. Figure 6 illustrates two possible angles that can be used to calculate the required parameters (α and β).

$$H_a = \begin{bmatrix} \frac{1}{\beta} & \frac{-\alpha}{\beta} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

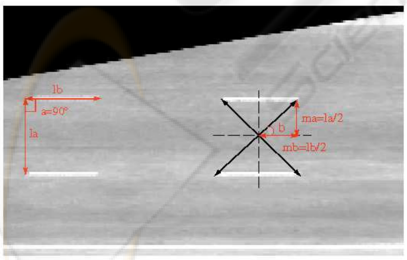

Figure 6: Representation of two possible angles obtained from two known lengths on the ground plane.

To conclude, the last transformation known as similarity transformation performs rotation, translation and isotropic scaling of the resultant image.

$$H_s = \begin{bmatrix} s.\cos\theta & -s.\sin\theta & t_x \\ s.\sin\theta & s.\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

Therefore, a rectified image can be created by applying the following transformation, along with bilinear interpolation, to each pixel of the image acquired by the surveillance camera:

$$H = H_s.H_a.H_p \tag{7}$$

3 SCALE FACTOR AND LANE PARAMETERS

Once the image sequence is rectified it is possible to measure the distance, in kilometres, travelled by the vehicle in two consecutive frames. In order to do so, one must calculate a scale factor that relates pixels in the image with the corresponding distance on the ground plane. This scale factor can be obtained by estimating the ratio between the imaged highway stripe period and the genuine stripe period on the ground plane. Hence, in order to clearly identify the striped lines on a rectified image plane, the previously attained background image is rectified. Observing the stripes presented on a rectified background image, which can be seen on Figure 7, it is possible to conclude that these follow a periodic distribution. Using a auto correlation function, represented by expression (8), the lane stripe's period can be located, since the function has peaks in the beginning and ending of each stripe, see Figure 8.

$$R(k) = \frac{E\left[(X_i - \mu).(X_{i+k} - \mu)\right]}{\sigma^2}, \tag{8}$$

where E is the expected value operator, X_i is a pixel of the straightened background image while X_{i+k} is a pixel on the same line of the referred image, but distanced k pixels from the first. μ represents the average of the pixels of each line of the straightened image, σ is the corresponding variance and k the already referred distance gap that, in this precise case, is a number of columns of the rectified background image.

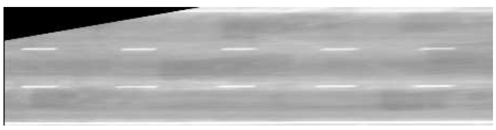

Figure 7: Example of a straightened background image.

Therefore, the autocorrelation function is applied to each line of the straightened background image, allowing the identification of the whereabouts and periods of the image stripes. Once located the image highway stripes, it is possible to situate the highway boundaries on the image plane. This information is quite useful seeing as it may be used in estimating each lane's mean vehicle velocity.

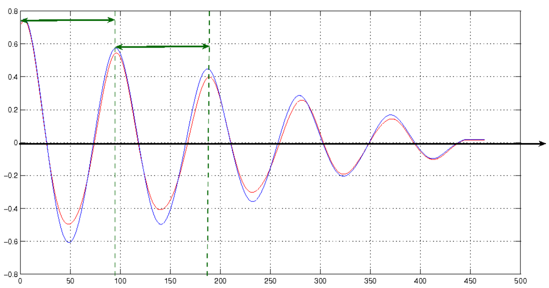

Figure 8: Autocorrelation functions of two stripe lines present on the rectified background image The green arrows represent the stripe period.

4 VEHICLE VELOCITY

To identify the location of vehicles on each frame, one must first distinguish objects from the background. This procedure is known as image segmentation and can be accomplished by subtracting each frame from the background image. Figure 9 shows the result of this operation.

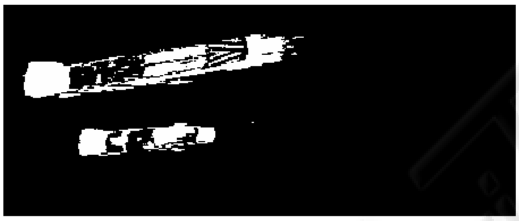

Figure 9: Example of a segmented image.

However, as can be seen in Figure 9, this process generates non-contiguous objects. This phenomenon results from the lack of information in regions distant from the camera due to the perspective distortion intrinsic in image formation. This fact is also responsible for the visible deformation of vehicle shape on rectified images. Seeing as non-contiguous objects constitute a problem to vehicle detection, certain morphological operations, such as dilation and erosion, are applied to the resulting images. Figure 10 illustrates an example of a segmented image that may be used in the identification of vehicles.

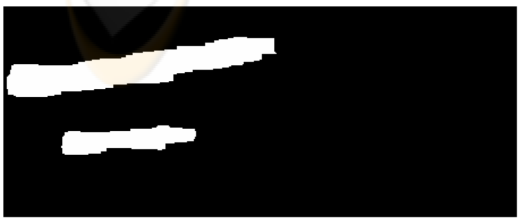

Figure 10: Example of two continuous blobs.

The velocity associated with each detected vehicle is easily attained by employing a Kalman filter (Kalman and Bucy, 1960) based tracking system. This process predicts future positions given a sequence of images, and matches these with the information provided by blobs extracted from the segmented image in order to correct the system's model. In order to do so, this process reiterates three steps for each new frame. The first step is responsible for extracting the location of each identified vehicle using the contiguous object in the segmented image, acknowledged as blob. Given two consecutive frames it is possible to estimate motion, i.e. optical flow, using Lucas-Kanade's method (Lucas and Kanade, 1981). This technique presupposes that a pixel's intensity is invariant in two successive frames and therefore, it is possible to determine motion by locating the corresponding pixel on the subsequent frame. Nevertheless, this method might be flawed when applied to traffic image sequences. Consequently, we chose to employ a Kalman filter to estimate each vehicle's position and correct the estimated velocity given by Lucas-Kanade's method. Given the fact that the imaged vehicles are travelling on a rectified image, their velocity can be considered linear and therefore, the following equations may be used to characterize a vehicle's movement:

$$\begin{aligned} x_{i+1} &= x_i + v_x.t \\ y_{i+1} &= y_i + v_y.t \end{aligned} \tag{9}$$

where v_x and v_y represent the different velocity components while x and y the vehicle's position coordinates. These expressions can be employed in the implementation of the Kalman filter that estimates each vehicle's position and corrects the estimated velocity. More precisely, each estimated state is obtained by applying the following equation:

$$\hat{x}(k) = \phi(k-1).x(k-1) = \begin{bmatrix} x, v_x, y, v_y \end{bmatrix}^T, \tag{10}$$

where ϕ is the transition matrix and can be initially defined by the following matrix:

$$\phi(0) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{11}$$

This leads to the second step of the tracking system. In order to employ a Kalman filter in this process, one must relate a given estimation with the identified position using the segmented image. This process is rather tricky due to the occasional overlapping of blobs in the segmented image or absence of detection. Thus, a failsafe sub process was implemented in order to deal with these cases.

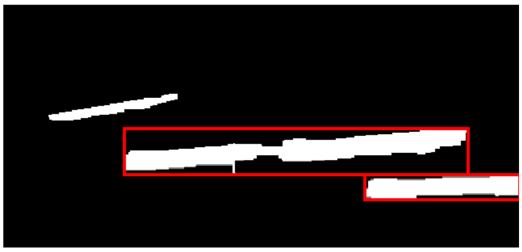

Figure 11: Representation of a blob overlying another.

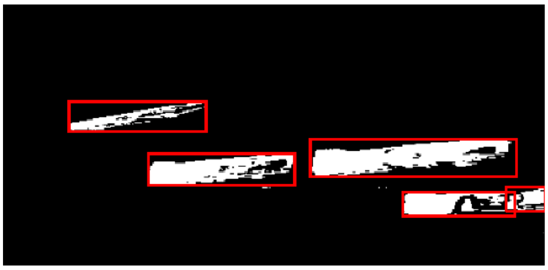

Figure 12: Result of the employment of the failsafe process to the image shown on Figure 11.

Figure 11 illustrates a segmented image where the overlapping of two blobs is visible. On the other hand, Figure 12 shows the result of the employment of failsafe process to image shown in Figure 11. As it can be quite easily seen, the failsafe process does not employ morphological operations to the segmented image and uses more permissive parameters in the detection of blobs. The final step in this process uses the estimated positions and estimated velocity provided from the Kalman filter and performs data management. More specifically, inserts new vehicles in to a linked list, removes vehicles that are no longer acknowledged on the segmented image or simply just updates data. In this context, an object's velocity is linear and therefore, once known the image sequence's framerate, estimating vehicle velocity is quite straightforward, as can be seen in the following expressions:

$$v_x = \frac{dx}{dt} \simeq 3,6.\Delta x.f.s \; [km/h]$$
$$v_y = \frac{dy}{dt} \simeq 3,6.\Delta y.f.s \; [km/h] \qquad (12)$$

where Δx and Δy are the estimated displacements in pixels between two consecutive images, s the scale factor previously obtained and f the framerate. Nonetheless, it is important to state that expression (12) estimates vehicle velocity in kilometers per hour. However, in order to obtain this estimation in another unit system, the procedure is quite similar.

This application can also count the sum of vehicles that travel on the observed scenario. To do so, an analysis is made to a control flag which indicates if a vehicle hasn't already been taken into account. The referred analysis is preformed in the middle region of the straightened image due to the fact that this region has a higher probability of including all of the highway lanes. Given the fact that all the vehicles have associated an identification of the lane where these pass through it is quite simple to calculate the mean velocity of each lane.

5 RESULTS

In this section we present several rectified images and two images that exemplify velocities obtained for several vehicles on two consecutive images. Figures 13 and 14 illustrate the results of the employment of the rectification process to images captured by two different surveillance cameras. The images on the left hand side of each figure show the captured images to which the process is applied, while the images on the right hand side are the images achieved using the rectification process.

Figure 13: Example of a captured image (on the left) and the correspondent rectified image (on the right).

Figure 14: Example of a captured image (on the left) and the correspondent rectified image (on the right).

Figure 15 illustrates the tracking system in two consecutive frames. In both images of this figure, one may see each vehicles estimated velocity, the sum of vehicles counted until that instant and each lanes mean vehicle velocity which is represented on the upper left side of each image. Labeling each lane and vehicle simplifies the task of estimating each lanes mean vehicle velocity. Therefore, each lane is identified by an incremental numerical id. For instance, the uppermost lane is labeled on the image as lane 0, the lane below as lane 1 and finally the bottommost lane as lane 2. By examining the mean vehicle velocity associated to each identified lane, one may observe that the lane identified as lane 0 has the highest mean velocity. Given the fact that this lane represents the acceleration lane, this observation was not unexpected. To conclude it is important to state that the green vectors represented on each image of Figure 15 illustrate the motion vectors for each identified vehicle, while the estimated velocities in km/h are illustrated in red above each vehicle.

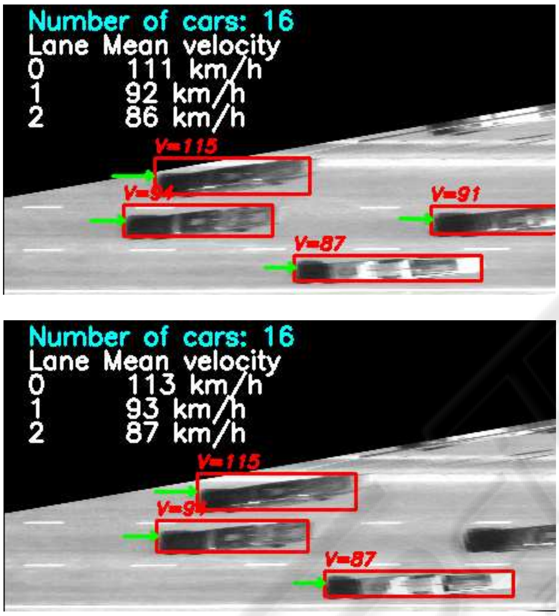

Figure 15: Estimation of vehicle velocity in two consecutive frames.

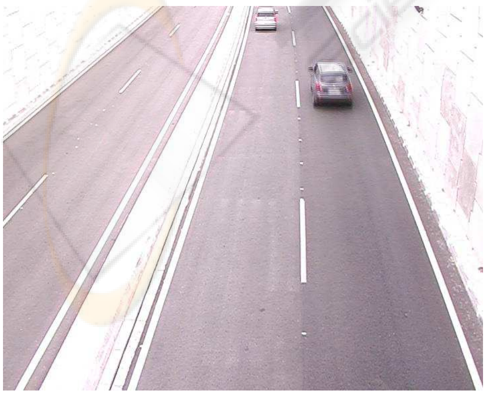

Figure 16: Captured image of the vehicle traveling at a known velocity.

In order to establish the error associated to the vehicle velocity estimation, the algorithm was applied to a video sequence containing a vehicle traveling at a known velocity. The referred vehicle is shown in Figure 16 and was traveling at 74.5 km/h according to a GPS system. To estimate this vehicle's velocity the rectification process was applied as can be seen in Figure 17. This Figure also shows the vehicle's estimated velocity between two consecutive frames.

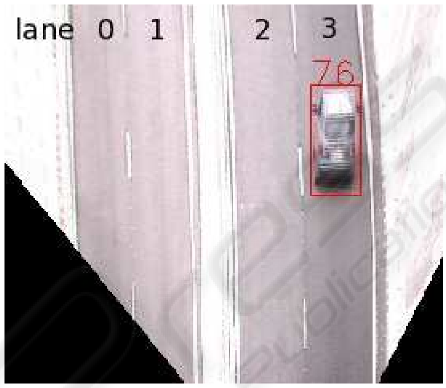

Figure 17: Representation of the vehicle traveling at a known velocity.

The result obtained applying the previously referred algorithm has an error of 2%, considering that the vehicle's real velocity was 74.5 km/h, measured by a GPS system. The graphic represented in Figure 18 illustrates estimated velocities obtained by tracking several vehicles in consecutive frames of the video sequence represented in Figure 17. Analyzing this Figure it is possible to observe that the instantaneous velocities, estimated in each frame, are influenced by noise caused by the lack of robustness of the segmentation process. Figure 18 also represents the estimated velocities of the vehicle traveling at a known velocity on the ground plane.

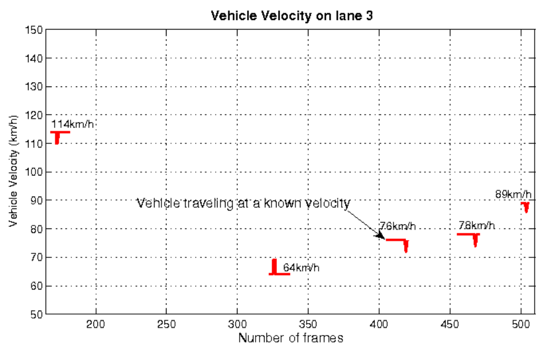

Figure 18: Graphic illustrating the velocities of several vehicles on lane 3.

6 CONCLUSIONS

This paper describes a technique to obtain a bird eye view of the ground plane in order to estimate vehicle velocity. The method requires no knowledge of camera parameters, only needs two known lengths of the highway. The rectification technique also requires that highway lanes and lane boundaries be approximately straight in the region of surveillance near the camera. This method was tested on different traffic sequences, providing fine results. To conclude, it is important to state that this procedure can be employed in other automatic traffic surveillance systems.

REFERENCES

Bose, B. and Grimson, E. (2003). Ground plane rectification by tracking moving objects. In *Proceedings of the Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*.

Dailey, D. J. and Cathey, F. W. (2005). A novel technique to dynamically measure vehicle speed using uncalibrated roadway cameras. In *Proceeding IEEE*.

Fischler, M. A. and Bolles, R. C. (1981). Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. In *Comm.Assoc. Comp. Mach.*

Kalman, R. E. and Bucy, R. S. (1960). New results in linear filtering and prediction theory. In *Transactions of the ASME, Journal of Basic Engineering (Series D)*.

Lazaros Grammatikopoulos, George Karras, E. P. (2002). Automatic estimation of vehicle speed from uncalibrated video sequences. In *International Archives of Photogrammetry and Remote Sensing*.

Liebowitz, D. and Zisserman, A. (1998). Metric rectification for perspective images of planes. In *Proceedings of Computer Vision and Pattern Recognition*.

Lucas, B. D. and Kanade, T. (1981). An iterative image registration technique with an application to stereo vision. In *Proceedings DARPA Image Understanding Workshop*.

Schoepflin, T. and Dailey, D. J. (2003). Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation. In *IEEE Transactions on Intelligent Transportation Systems*.