Vehicle Speed Estimation from CCTV Feeds

Abbas Javed, Saqib Sarfraz
Computer Vision Research Group
Department of Electrical Engineering
COMSATS Institute of Information Technology
Lahore, Pakistan
{abbasjaved, ssarfraz}@ciitlahore.edu.pk

***Abstract*—In this paper, a novel approach for vehicle speed estimation is proposed. The technique has been tested using video feeds from CCTV cameras that are already installed on roads for traffic surveillance. Vehicle speed estimation from these videos is a very challenging task because video sequence from CCTV cameras do not preserves length ratio and line parallelism. Length ratios and line parallelism are restored by projective transformation. The image is rectified for the detection of stripe lines and their length is computed. The stripes length is further used to compute the scale factor. Distance covered in consecutive frames can be related to the distance covered in ground plane by using scale factor. Detected vehicle is then tracked in consecutive frames. As the frame rate of input video is already known therefore the speed of moving vehicle in both lanes is estimated by calculating the translation between frames. The proposed framework can also be used to count vehicles in a certain direction of road. Our approach accurately calculates the speed of moving vehicle within the error of ± 3.5 km/hour of actual speed.**

***Keywords*— projective transformation, vehicle detection, vehicle tracking, distance estimation, speed estimation.**

I. Introduction

Vehicle speed estimation is an important factor in traffic monitoring. Majority of accidents occurred due to over speeding. Several different systems are used for estimating vehicle speed. i.e. use of magnetic inductive loop detector, magnetic strips, infrared sensors, radar, laser sensors etc. [1][2][3]. Magnetic inductive loop detector and strips are required to be installed on fixed location by road digging. RADAR (Radio Detection and Ranging) and LIDAR (Laser Infrared Detection and Ranging) devices are commonly used by traffic controlling authorities to monitor speed of the passing vehicles. RADAR devices transmit radio signals to the moving vehicles and receive the reflected signals from moving vehicle. Difference between the frequency of original and reflected signals is converted in to the speed of the vehicle. LIDAR devices work by calculating the time taken by LASER pulse to propagate from gun to vehicle and reflected. Based on this information distance of vehicle from gun is computed. LIDAR calculates speed with the help of numerous measurements and compares the distance covered between these measurements. RADAR and LIDAR devices based methods are active and requires operator to perform measurements.

Another approach of calculating speed is through image sequences from the camera. Cameras are installed on different roads for surveillance and traffic monitoring. These cameras are un-calibrated and require some calibration to transform pixels into real world units.

Camera motion or scene analysis are often used for camera calibration.[4][5]. Lane boundaries and stripes boundaries can be used to calibrate the camera. Many methods have been developed so far for automatic recovery of distorted projective properties. These methods require calculation of one or two vanishing point (in orthogonal and horizontal direction). Cathey et al [6] proposed a method of camera calibration using one vanishing point. This method presupposes the camera position and in order to measure speed of vehicles from a series of video images it is not necessary to completely calibrate the camera, but rather by using the vanishing point of lines in the image they establish algebraic constraints on the parameters that are sufficient to straighten the image and compute a scale factor for estimating speed. Grammatikopoulos et al [7] also proposes a method that performs image rectification with one vanishing point and one known ground length. Methods proposed in [6][7] are applicable to the scenario where camera is exactly parallel to the motion of vehicles.

Maduro et al [8] presents a method that performs image rectification using two vanishing points as presented by Liebowitz et al [9]. This approach requires two known length and angles. After image rectification this method uses Lucas Kanade Tracker [10] for tracking the vehicle and calculating speed. However method proposed by Maduro et al [8] is applicable to the three lane roads.

Rather than using vanishing points for image rectification Chorianopoulos [11] performs projective transformation with the help of 20 control points. This method accurately calculates speed with standard deviation of $\sigma S = \pm 2$ km/h.

Our proposed system is capable to compute speed from CCTV feeds and is independent of camera position and number of lanes. Similar to Chorianopoulus [11], our technique performs projective transformation with the help of control points. But instead of using 20 control points transformation is performed with the help of only 4 control points. Once image is straightened by projective transformation the stripes in the rectified image are periodically distributed. Vehicles are detected and tracked in the specified region. For vehicle tracking and detection we used same technique as proposed in [12]. With the help of known distance on ground, pixels on the image plane can be related to the world plane and hence the movement of object on image plane can be related to the displacement in the world plane. Frame rate of the video is already known so speed is calculated by dividing displacement with the time required to cover that displacement

.

The rest of this paper is organized as follows. Camera selection, calibration, vehicle detection and speed estimation is described in section II. The experimental results are provided in section III followed by the discussion and conclusion in section IV.

II. CALIBRATION AND SPEED ESTIMATION

CCTV camera is selected for image acquisition which is already installed for traffic monitoring and surveillance purposes. Our method applies some constraints on camera selection. (1) Road is fairly straight and planar in front of camera. (2) The field of view of camera should be unobstructed by large overhead signs and overpasses.

A. Camera Calibration

Camera Calibration is done in three steps i.e. (1) road boundary and lane stripe detection (2) projective transformation with control points (3) scale factor selection.

i. Road Boundary Detection

Road boundary is extracted by applying canny edge detector on input frame. In order to detect road boundary marking Hough transform is applied on the edges and boundary lines are extracted. For extracting boundary lines a threshold is selected on the basis of slope and vertical length of lines and only those lines are drawn on the image which passes this criteria. From these vertical edges, as shown in Figure 1 boundary of the road is manually selected by the user by clicking on A`,B`,C` and D` location respectively as shown in Fig 2.

Between the trapezoidal region of interest i.e. A`B`C`D` selected by the operator stripe lines are detected by calculating the slope of lines extracted by Hough transform. Only those lines are recognized as stripe lines whose slope 'm' is greater than a threshold. ~~4.~~

ii. Projective Transformation with Control Points

During perspective imaging parallel line converges to finite point i.e. rectangle in the world plane will appear as trapezoid in the image plane. This perspective distortion can be removed by projective transformation.

A planar projective transformation is a linear transformation on homogenous 3-vectors represented by a non-singular 3 x3 matrix [13]

$$\begin{bmatrix} x1' \\ x2' \\ x3' \end{bmatrix} = \begin{bmatrix} h11 & h12 & h13 \\ h21 & h22 & h23 \\ h31 & h32 & h33 \end{bmatrix} \begin{bmatrix} x1 \\ x2 \\ x3 \end{bmatrix} \qquad (1)$$

Or equation 2 can be written as

x' = Hx

where 'x'' is image plane and 'x' is the world plane.

By computing H from equation 1 and applying inverse transformation on image, perspective distortion can be restored to the world plane.

One way of computing H is by point to point correspondence of image and world plane.

Let the inhomogeneous coordinates of a pair of matching point x and x' in the world and image plane be (x,y) and (x',y') respectively. In inhomogeneous form equation 2 can be written as

$$x' = \frac{x'_1}{x'_3} = \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \quad y' = \frac{x'_2}{x'_3} = \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \qquad (2)$$

After cross multiplication equation 2 becomes

$$x'(h_{31}x + h_{32}y + h_{33}) = h_{11}x + h_{12}y + h_{13} \qquad (3)$$

$$y'(h_{31}x + h_{32}y + h_{33}) = h_{21}x + h_{22}y + h_{23} \qquad (4)$$

Eq.3 and Eq.4 are linear in the element of H. Four points selection will result in eight such equations.

Figure 1 Boundary lines using Hough Transform

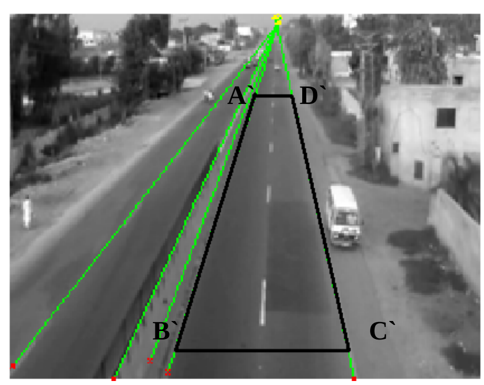
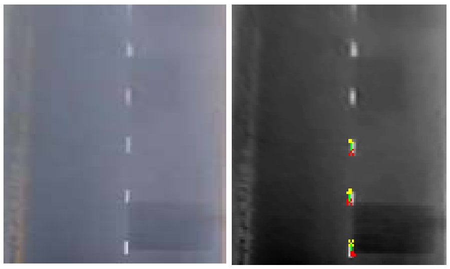


Figure 2: Manual Road Boundary Selection

The transformation H is computed by point to point correspondence between image plane as shown in fig 2 with world plane. Selected control points are A`B`C`D` in image plane and A,B,C,D in world plane. Where A,B,C. and D represents fixed location in world plane, after applying inverse transformation to fig 2 resultant straightened (rectified) image is shown in fig.3a



3(a) **3(b)**

Figure 3a : Straightened Image after Projective Transformation and

Figure 3b shows detected striped on rectified image

iii. Scale Factor Selection

Once image is rectified as shown in figure 3(a), it is possible to measure the distance travelled by vehicle on ground. For this, stripe lines are extracted by applying hough transform on edges. By comparing the period of stripes lines in image plane with the world plane scale factor ‘s’ is calculated. This scale factor ‘s’ relates the movement of vehicle in image plane with the displacement covered on ground plane in Kilometers. Scale factor ‘s’ can be mathematically shown as

$$s = \frac{stripe\ \ period\ \ on\ \ ground\ \ plane\ \ in\ feets}{stripe\ \ period\ \ in\ image\ \ plane} \qquad (5)$$

Where “stripe period in image plane” is the number of pixels between two consective stripes.

B. Background Learning

Background learning is important step in vehicle detection and stripe lane detection. Accuracy of speed estimation is dependent on correct vehicle detection and tracking of vehicle in the region between two consective stripes. Stripes cannot be detected correctly if occluded by vehicle. For this correct background learning is required which is done by exponential forgetting algorithim[14]. The background learning procedure is mathematcaly described as

$$B_n = (1-\alpha)B_{n-1} + \alpha\ I_n \qquad (6)$$

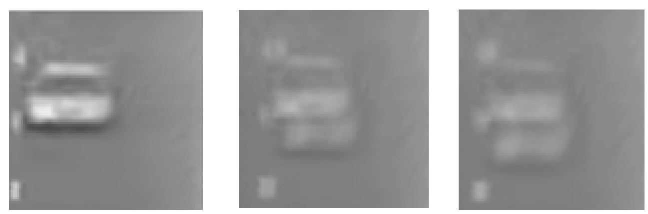
Where B_{n-1} is previous brackground image, I_n is current frame and ‘α’ is the background learning coefficient and its value is set at 0.2

The procedure for computing background is that it considers first frame of the video as background and update it by next coming frames. The updated background image computed by equation (6) is the weighted sum of previous background image and current frame. In this way possible changes in illumination due to weather, shadow and time is catered. Figure 4(a) to 4(e) shows the background learning for few consective frames.

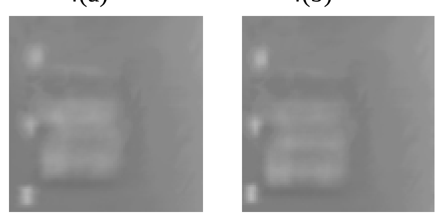
C. Vehicle Detection

The basic goal of our work is to calculate speed of the vehicle from the CCTV feed which is of low frame rate and low resolution. Due to low quality video computed edges are very weak and required some pre-processing. Foreground Objects are detected by subtracting background from the current frame. This image is converted into binary image and small holes between the blobs are filled by performing morphological operations (dilation in this case). 2D median filtering is also applied to reduce noise while preserving the edges. After these steps foreground is detected by using connected components labels. After specifying the shape (rectangle) morphological structuring element is created as shown in figure 5(a)-5(b)

So boundary of the each detected foreground object is of rectangle shape and by setting certain threshold detected foreground object is classified as vehicle. This threshold is set on the basis of width to height ratio of the rectangle. If the width to height ratio of the rectangle is below threshold the detected object is discarded and if its above the threshold foreground object is classified as moving vehicle.

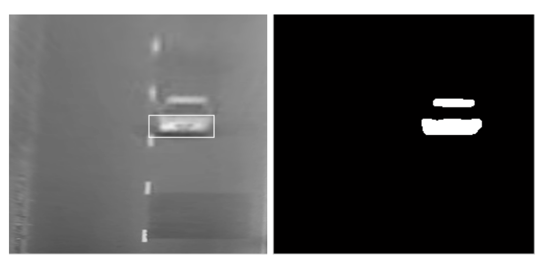


4(a) 4(b) 4(c)



4(d) 4(e)

Figure 4 (a) to Figure 4(e) shows the background for few consecutive frames



5(a) **5(b)**

Figure 5 (a) shows rectangle drawn on detected vehicle. Figure 5(b) shows detected foreground object

D. Vehicle Tracking

If there is no huge variation in the pixel intensities it is possible to track the pixels in consecutive frame. For this purpose we used the well known Lucas Kanade tracker.[10]

Detected vehicle in the video feed is passed to the lucas kanade tracker which estimates its position in the next frame by calculating the displacement using a constant acceleration dynamic model. [15]

$$I(x, y, t+\tau) = I(x-\xi, y-\eta, t) \tag{7}$$

where, 'I' is a window of pixels, 'x' and 'y' are new x and y coordinates of the object (license plate), 'ξ' and 'η' are the displacement values for previous x, y coordinates respectively

The later image taken at time t + τ can be obtained by moving every point in the current image, taken at time t, by a suitable amount. The amount of motion d = (ξ ,η) is called the displacement of the point at X = (x, y) between time instants t and t + τ , and is in general a function of x, y, t, and τ.[15]

E. Speed Estimation of Vehicle

As input video frame is rectified with control points so image scale is uniform along y axis. Each detected vehicle is tracked along y axis for sufficient number of frames. For simplicity vehicles are tracked for the frames equal to the frames captured by camera in one second. Speed of the vehicle is calculated by

$$\text{Speed} = \frac{dx}{dt} = \Delta y * s * 0.3048 * 3.6 \tag{8}$$

If frame rate is 'f' frames per second than Δy is the no.of pixels covered in 'f' frames by vehicle. 's' is the scale factor and it relates the pixels in image plane with the distance covered in feets. In order to convert speed from feet per second to kilometre per hour speed is multiplied by the appropriate factor.

III. Experiments and Results

For testing of our technique camera is installed at three different location with different angle on different road.

We have tested our framework on roads with single and double vanishing points. The advantage of this technique is that it does not require computation of vanishing point.

A. Case 1: Single Vanishing point

Camera is installed on overhead bridge on highway as shown in figure 6 and capturing video at 24 frame per seconds. Four control points A'B'C'D are selected and image of figure 6(a) is rectified as shown in figure 6(b). Length of road stripe is 4.2 m or 13.9 feets and distance between each stripe is 12 m or 39.36 feets. Stripe period in image plane is about 37 pixels so calculated scale factor 's' is 1.06. Silver Suzuki Cultus is travelling at known speed of 80 Km/hr. For testing purpose driver is instructed to maintain consatnt speed of 80Km/hr. Vehicle is detected and tracked for 24 frames. During these 24 frames vehicle covers distance of 66 pixels so estimated speed is 76.76 Km/hr.

B. Case 2: Road with Double Vanishing Point

Our technique is also tested at road with two vanishing points. Camera is installed on height on light pole situated at the right boundary of the road. Captured frame from this camera is shown in Fig 7(a). Vehicles crossed in front of the camera at known speed. i.e.

Silver Suzuki cultus at 20 Km/hr,

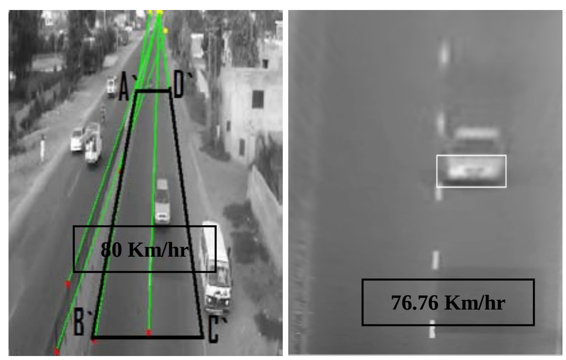
Purple Vitz at 30 Km/hr

Silver Suzuki Cultus at 40 Km/hr.

With the same procedure followed in case 1 scale factor of 0.6 is calculated. Camera frame rate is 29 frame per second and estimated vehicles speeds are shown in figure 7(b) and 7(c)

Performance of the technique is also tested at the road where camera is installed at low height on tripod stand and capturing the video at 25 frames per second. With the help of four control points A`,B`C` and D` from road boundary as shown in figure 8(a)

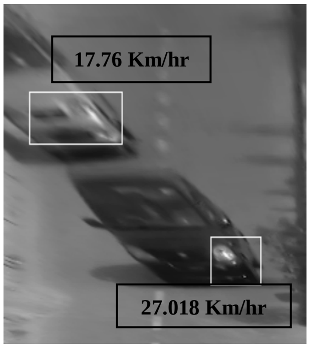
Input frame is rectified as shown in figure 8(b). During this transformation the shape of the vehicle is distorted but yet its possible to detect and track the vehicle in rectified image. Scale factor for this video is 0.7325. Silver car passed the camera at 60 Km/hr and our framework estimates its speed at 63.50 Km/hr.



6(a) 6(b)

Figure 6(a): Input image from camera, Figure 6(b) Estimated speed on rectified image

Figure 7(a) Input Image from camera



7(b)

7(c)

Figure 8a: Input Frame for Speed Estimation

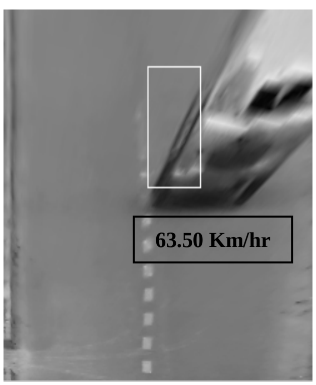


Figure 8(b): Estimated speed on rectified image

Table 1 presents the result of speed estimate of vehicle at five different known speed.

No.	Actual Speed (Km/hr)	Estimated Speed (Km/hr)	Standard Deviation σ (Km/hr)
1	20	17.76	-2.24
2	30	27.018	-2.98
3	40	37.52	-2.48
4	60	63.50	+3.5
5	80	77.52	-2.52

As shown from the Table 1 our approach gives satisfactory result of ± 3.5 Km/hr. Figure 9. Represents the estimated speed calculated by tracking vehicles in consecutive frames of video sequence shown in Figure 6(b), 7(b), 7(c) and 8(b). Figure 9 also shows the actual known speed of the vehicle. For testing purpose drivers are requested to maintain the specific speed.

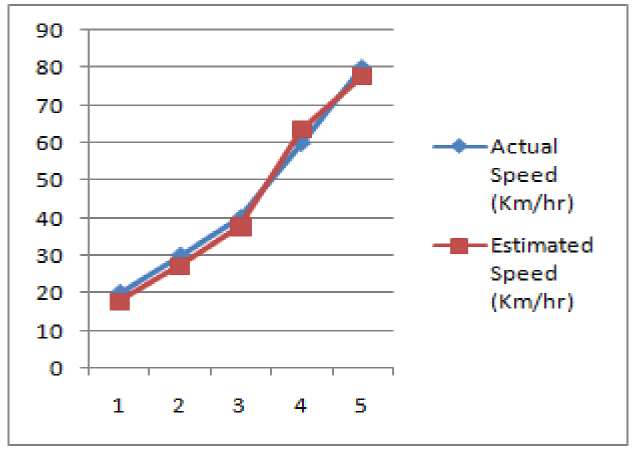


Figure 9: Shows difference of actual and estimated speed

IV. Conclusion

In this paper, we used simple approach to rectify projective distortion of the input video sequence. Our approach does not require any computation of vanishing point therefore initialization phase is very easy. It requires four control points from the boundary of road for computing transformation matrix. Accuracy of rectified image is comparable to the techniques that compute transformation matrix using single or double vanishing point.[6],[7],[8]. As only known length on ground plane is required for calculating scale factor so the primary advantage of our approach is that it can estimate speed on the roads where stripe lines are not visible. This technique can also be used for counting vehicles travelling in one direction of road as well as computing the average lane velocity. Our proposed framework can calculate speed on any road with two or three lanes and there is no restriction on camera location.

References

[1] C. Sun and S. G. Ritchie, "Individual vehicle speed estimation using single loop inductive waveforms," *Journal of Transportation Engineering-Asce*, vol. 125, pp. 531-538, 1999.

[2] Y. H. Wang and N. L. Nihan, "Freeway traffic speed estimation with single-loop outputs," *Advanced Traffic Management Systems and Automated Highway Systems 2000*, pp. 120-126,

[3] D. L. Woods, B. P. Cronin, and R. A. Hamm, "Speed Measurement with Inductance Loop Speed Traps.," Texas Transportation Institute., *Research Report FHWA/TX-95/1392-8*, 1994.

[4] O. Masoud and N. Papanikolopoulos, "Using geometric primitives to calibrate traffic scenes," in *Intelligent Robots and Systems Proceedings*, vol. 2, 2004, pp. 1878–1883.

[5] H. Kim and K. S. Hong, "Practical self-calibration of pan-tilt cameras," *IEEE Proc. on Vis. Image and Signal Process.*, vol. 148, pp. 349–355, 2001.

[6] Dailey, D. J. and Cathey, F.W. "A novel technique to dynamically measure vehicle speed using uncalibrated roadway cameras. In Intelligent Vehicles Symposium Proceedings, IEEE, pp 777-782, 2005

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

[8] Cristina Maduro, Katherine Batista, Paulo Peixoto, Jorge Batista: Estimating Vehicle Velocity Using Rectified Images. VISAPP (2) 2008: 551-558

[9] Liebowitz, D. and Zisserman, A., "Metric Rectification for Perspective Images of Planes," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 1998.

[10] Lucas, B. D., Kanade, T.: An iterative image registration technique with an application to stereo vision. Proceedings of the 7th International Joint Conference on Artificial Intelligence (1981)

[11] Chorianopoulos, P., 2001. Study of Freeway Traffic Features under Free-flow Conditions using Photogrammetric Methods. Diploma Thesis, Dept. of Surveying, NTUA, pp. 173.

[12] Shahzad M.A., Fraz M., Elahi M. A., Sarfraz M.S. Real Time Localization, Tracking and Recognition of Vehicle License Plate. In the Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP), 2011.

[13] Hartley, R., Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press.

[14] Tan, X., Li, J., and Liu, C. A video-based real-time vehicle detection method by classified background learning. World Transactions on Engineering and Technology Education 6 No. 1 (2007)

[15] Carlo Tomasi and Takeo Kanade. Detection and tracking of point features. Technical. Report CMU-CS-91-132, Carnegie Mellon University, Apr 1991.