

# Investigation on Monocular Vision based Moving Object Detection and Speed Estimation of using Low Profile Camera

Syed Nizamuddin Peeran, J. Cypto, P. Karthikeyan  
Department of production technology  
Madras Institute of Technology  
Anna University, Chennai, India

**Abstract**—Over speed of vehicles on road is major reason for several road accidents. In order to prevent the accidents, the motor vehicles rules has to be followed and especially over speed must be monitored. The estimation of moving object velocity is challenging in order to enforce law against over speed. There are several devices that have been invented to estimate the velocity. Most of these devices are either expensive or furnishing scientific evidence is difficult to file the case against violators. In this project, the low cost vision based velocity estimation is proposed to identify the over speed vehicles. The system will extract the vehicle registration number from the scene which can be used proof against the violator and their by bringing them to justice. Several algorithms were studied and tested which include foreground estimation using temporal filter and optical flow estimation using Horn-Schunck method and Lucas-Kanade method. The optimal algorithm for estimating the vehicle speed is identified as the optical flow computation using Horn-Schunck method is better based on time taken and percentage of truth of the result. Based on the results obtained, a new hybrid foreground detection algorithm which is robust to sudden lighting changes, shadow movements is created which can effectively detect the moving object velocity.

**Keywords**—Vehicle speed, moving object, vision, optical flow, foreground detection

## I. INTRODUCTION

Obtaining the speed of the moving externally without contact has been a challenge and [1] can be used in visual surveillance (outdoor and indoor environments), traffic control, behavior detection (sport activities), etc. To detect the moving objects a radiometric ratio is used for every pixel by the authors in [1] and is claimed that it is robust to sudden lighting changes. Euclidean structure of the object and the velocity estimation of the object with the help of geometric calculations in [2]. A stereo vision based system is proposed using inverse perspective map extended kalman filter [3]. There have been numerous studies on velocity estimation based on vision. This project aims at providing a system that can be used to monitor traffic and hence detect any violations of traffic rules especially speeding. For this purpose various methods to detect the velocity of an object with the help of a camera are studied and analyzed. The speeding of vehicles beyond the posted limits in both urban and rural areas is the major cause of the accidents, controlling the speed of the vehicles will hence save a lot of lives in the future. Initially, during the literature survey it has been found out that most of

the vision based systems use optical flow to solve the velocity of the object in a given scene.

The problem is to create a low cost reliable system that can be deployed to identify the speeding vehicles. In the pursuit of doing so there are several problems faced such as the selection of the appropriate hardware and software for the system, a proper algorithm to estimate the speed has to be selected from the list of algorithms available, the objects must be identified as vehicles, the location of the vehicles on the road must be identified and finally the vehicle must be identified by extracting the register number using the optical character recognition techniques.

The objective of this project is to create a camera only based system that would will be able to detect and identify the vehicle in traffic and display their speed. The system will be flexible and can be added to any added any to any existing traffic system. During the first phase the various algorithms are studied, tested and the best one is chosen the our system based on the time taken and positive results.

The overall working of the system is as follows, this system needs a video acquisition system, a real time video processing unit and a display unit to the results. There are several algorithms to detect the velocity of the objects in a given video viz., foreground extraction, optical flow and etc. they will be discussed in detail in the upcoming chapters. The objective for the first phase is to identify a perfect algorithm serving the purpose. Several works have been reported estimating the velocity of moving objects. The investigation on obtaining the velocity of the moving vehicles on the road has been done and explained in this section. It will be discussed in detail about the various steps involved in vision based velocity estimation. But over all the major steps involved in velocity estimation of traffic are perspective correction, detecting moving pixels in the frames of the video and then estimating the velocity of the moving pixels and calibrating to get the actual velocity of the vehicle.

A moving vehicle velocity estimator is presented in another paper [4] for interpreting scenes shot during bad weather conditions. An estimator based on the non-linear robust velocity estimator using the model based on contrast and brightness variation. This model has proved to be robust against falling snow patterns that have uncountable shape and size properties. This model which was not provided any knowledge about the color, ie., white is applied as vehicles and snowfall might be having the exact same white color. Instead a method is that can be used on the objective function

to minimize it in the model with following 4 variables viz., horizontal velocity and vertical velocity, contrast and brightness. Two velocity components of moving vehicles can be accurately detected as the falling snow patterns, captured in the brightness & contrast images. The method has been verified for the effectiveness with a modified cluster algorithm by comparing the recognition rate with the conventional velocity detection method.

The work by [5] is also based on a monocular camera but it is used to estimate the velocity and the structure. To estimate the structure of the object they use Lyapunov design methods. In this paper the relative velocity between the object and the camera are also estimated.

In another work by [3] in which, the position and velocity of the moving obstacle is detected using the stereo vision camera and with the help an extended Kalman filter. Normally, the position of an object is calculated with the help of triangulation or (IPM) Inverse Perspective Map, this method loses its accuracy with distance as the disparity in the image become very small. To overcome this a method is prescribed which estimating the sub pixel disparity using stripe based accurate disparity to improve the accuracy. The velocity and the position the object is then calculated with the help of the extended Kalman filter

The known object detection based algorithms are also used in the collision avoidance systems by [6]. The report highlights on various system that work using vision.

All most all the video that will be recorded will be exposed to vibration and hence video stabilization becomes key. [7] provided a solution on ROI wrapping. [2] have estimated the 3d Euclidean coordinates for a given moving feature from 2d images and hence have also estimated their velocities. Here the given camera velocity and position are already known. The technique they have used is called as the structure from motion. This can be used for deploying the system in windy areas where the pole on which the camera may be mounted is swinging in the air and hence is exposed to vibrations.

## II. LOCALIZATION OF VEHICLES IN ROAD

A work on a driver assistance system with the help of a single or a monocular camera [8] demonstrate an algorithm to track the feature points and hence establish a mathematical model of the lanes on the road. This algorithm was based on Kalman filters for dynamically tracking the target. [8] is mainly focused on the driving characteristics of the driver. The lane detection, detecting the obstacle and the data processed by an expert system would provide the driver with useful information. In the above work the lane detection systems can be used in the future works for accurate positioning of the vehicles on the road.

## III. CLASSIFICATION OF VEHICLES

In [9], the authors propose a real-time traffic monitoring system, cost-effective that can reliably perform traffic flow estimation and vehicle identification at the same time. Initially, the foreground is extracted considering pixel-wise weighting list which is used to model the dynamic background then, the foreground is formed over time based on the spatial-temporal profile image, the traffic flow was estimated based on the number of connected components in the image profile. In the

end, the vehicle type is determined according to the size of the foreground mask region. In addition, traffic velocity, flow, density and occupancy, are estimated based on the examination of the segments. Sample video is used with different fps to check the algorithm and the time take for execution is also measured.

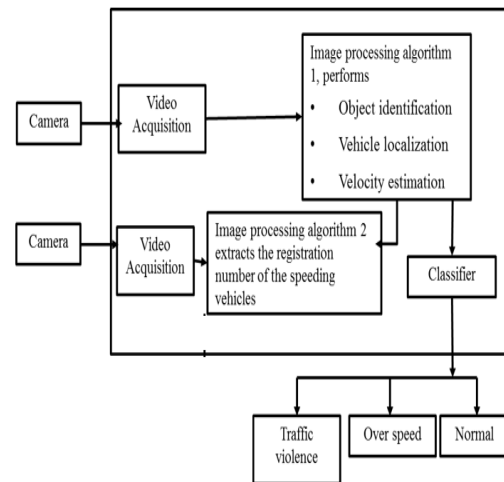


Fig. 1. Overview.

There are three methodology had been chosen for analysis, they are Back ground subtraction using centroids Optical flow Optical flow using centroids. These three methods are tested and checked against the ground truth velocity. A sample video is used with different fps to check the algorithm and the time take for execution is also measured. The very first stage in image or video processing is to acquire the image or the video from the source which is normally a camera. Video signal from the camera is analog.

### A. Camera Location (Working Distance) and Focal Length

The location of the camera 1 is chosen 15m and 3m away from the patch of the road that has to be monitored in order to avoid the problems caused but the occlusion of the vehicles the camera angle is 85°.

An adapter design has been provided to attach the dslr lenses to the low cost rpi camera this will makes camera flexible for a wide verity of applications.

To calculate the focal length of the required lenses system to be added to the camera we first need to know the distance that has to be covered by the camera sensor and the location of the camera.

The sensor size of the rpi camera is 2592 x 1.4um = 3.629mm wide.

In the figure consider L to be the size of the lense, H and W to be the height and the width of the actual object to be imaged (in our case, the road). D be the distance of the camera from the object also called as the standoff distance. We need to have a greater standoff distance to avoid the perspective distortion of the image from frame to frame.

A simplified calculation for 1/4-inch CCTV lens can be made using the following formula:

$$W = \frac{D * 3.2 \text{ mm}}{L} \quad (1)$$

Where,

W=horizontal width

D= Standoff distance

L=Focal length.

The actual width required to capture a speed of a vehicle moving at 140km/hr with the aid of a 30fps camera is determined to 40m x 7.8m.3.6.

#### IV. OPTICAL FLOW

For velocity estimation first the moving objects in the video (ie., the vehicles) will be detected. This is done with the help of optical flow. Optical flow or optic flow is the pattern of motion of surfaces, objects, and edges in a visual scene which is caused by the relative motion between observer and the object.

Sequences of ordered images enable the detection of motion as either instantaneous image velocity (discrete image displacements). The optical flow methods calculate the motion of pixels between two image frames with time tags  $t$  and  $t+\Delta t$  respectively at every voxel position. To compute the optical flow, the following constraint equation has to be solved:

$$I_x u + I_y v + I_t = 0 \quad (2)$$

Where,

$I_x$ ,  $I_y$  and  $I_t$  are the spatiotemporal image brightness derivatives along  $x$  and  $y$

$u$  is the horizontal OF

$v$  is the vertical OF

There are two major methods for determining the value of  $u$  and  $v$  they are:

1. Horn-Schunck Method
2. Lucas-Kanade Method

##### A. Horn-Schunck Method

This algorithm considers smoothness in the flow over the entire image. Their by tries to minimize distortions in flow, it only prioritizes the solutions which show more smoothness. The flow is expressed as a global energy functional which can then be minimized.

##### B. Lucas-Kanade Method

To solve the optical flow equation for  $u$  and  $v$ , this method divides the actual video frame into small sections and assumes a constant velocity in each area. Then, it makes a weighted least-square fit of the OF equation to a constant model for  $[u \ v]^T$  in each section  $\Omega$ .

##### C. Disadvantages Of Optical Flow

1. Complete vehicle body is not extracted
2. Slow moving vehicles are sometimes missed by the algorithm.
3. Computation time is high.

#### V. BACKGROUND SUBTRACTION

Foreground detection is one of the major tasks in the field of image processing which aims at detect changes in image sequences. Detecting foreground to separate these moving objects that are in the foreground from the background. It is a set of techniques that typically analyze the video sequences in real time typically recorded with the help of a stationary camera.

All detection techniques are based on the first model before the entire background of the image. That is, set the background and then see what changes occur in the background. Defining it can be challenging when it has shapes, shadows and moving objects. While defining the background is assumed to possess stationary objects that can have variations in color and intensity with time.

Scenarios where these techniques apply tend to be very diverse. Input can be a highly variable sequences of images with very different lighting, interiors, exteriors, quality their can be huge possibilities. You need a system that, in addition to process in real time, is robust to scene changes.

A very good foreground detection system should be able to get the background (estimate) whether it is a static variable. Be robust to lighting changes, repetitive movements at the bottom (leaves, waves and shadows), long-term changes. (A car comes and parks).

##### A. Temporal Average Filter

This system estimates the background model from the median of all pixels of a number of previous images. The system uses a buffer with the pixel values of the last frames to update the median for each image.

To model the background, the system examines all images in a given time period called training time. At this time we only display images and will find the median, pixel by pixel, of all the plots in the background this time.

After the training period for each new frame, each pixel value is compared with the input value of funds previously calculated. If the input pixel is within a threshold, the pixel is considered to match the background model and its value is included in the buffer. Otherwise, if the value is outside this threshold pixel is classified as foreground, and not included in the buffer.

This method cannot be considered very efficient because they do not present a rigorous statistical basis and requires a buffer that has a high computational cost.

##### B. Disadvantages of the Existing Foreground Detection Algorithm:

1. Affected by changes in lighting
2. Affected by movement in vegetation
3. Greatly affected by shadows
4. Too many false detections

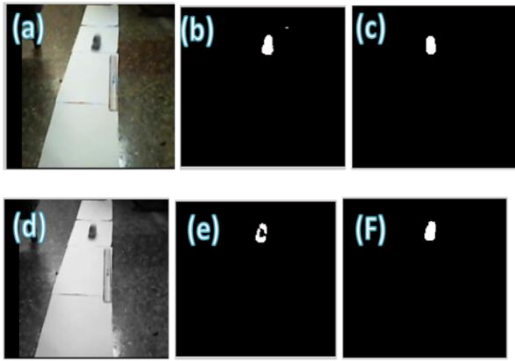


Fig. 3. Object detection results comparison for optical flow and background detection.

### VI. MODIFIED HYBRID ALGORITHM

Considering the pros and cons in both the optical flow and foreground detection a new algorithm has been created by combining both the algorithms together.

This algorithms also involves the temporal average of the first m frames to get the initial background estimation as shown in fig 4, in other words the initial background is estimated using the normal background estimator. Then the frames after the consecutive frames are used to identify the moving objects with the help of optical flow. The optical flow gives pixel velocity output thresholding the image gives the binary output. This image gives the fast moving pixels, it is compared with the previously detected background to estimate the intensity values of the missing pixels. The resulting image is an image of the stationary background.

This can be subtract from the next frame and hence the foreground object can be detected. The estimated foreground can be the used to the track the individual foreground objects from frame to frame throughout the scene. The complete foreground detection algorithm with added optical flow is shown in the fig 5. The algorithm was tested for a 30 fps, full HD video containing a single moving object (a bike). The results are acquired and have a deviation of about  $\pm 5$ km/hr. the tested videos.

### VII. RESULTS AND CONCLUSION

A set of videos with different fps and resolution are taken as test samples. As discussed in the chapters earlier the algorithms were tested for times of execution per frame for various resolutions and relation to ground truth for various algorithms.

Fig 8 shows the moving object detection performed by optical flow algorithm and its comparison to the fore ground detection algorithm. The image (a) and (d) are the input image frame to the foreground detection and optical flow algorithms respectively. The (b) is the detection result for foreground detection and (c) is the final result after correction with the help of morphological operation opening. The (e) is the detection result for optical flow and (f) is the final result after correction with the help of morphological operation closing. Both algorithms are good at identifying moving objects.

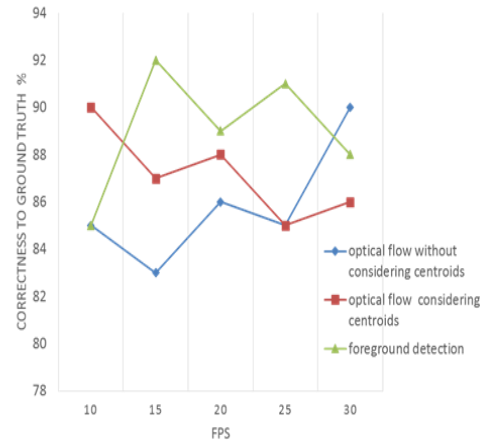


Fig. 4. Output comparison with truth.

The graph in the fig. 4 is the detection time for the carious algorithms for different resolutions. This will help identify the quickest algorithm for real time applications chosen as a problem statement for the project.

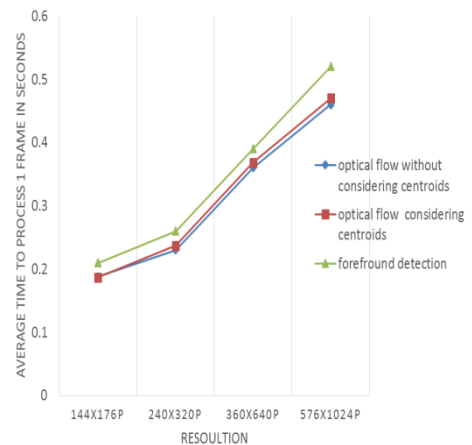


Fig. 5. Graph-processing speed.

The next figure, fig 5 is the comparison of the various algorithms with respect to the ground truth velocity. This velocity is estimated visually by looking at the time line of the video to calculate the time take for the object to travel between two known points. From the graph it is clear that the optical flow algorithm performs better in higher fps and the foreground detection performs well with higher resolution. In the above fig 6 image (a) shows the frame input, (b) shows the optical flow output, (c) is the threshold image (black and white), (d) estimated moving objects in the frame and (e) is the final estimated background.



Fig. 6. Final Algorithm Output.

The final output of the foreground is obtained by subtracting the image from the obtained background. It can be noted that the image has a hole which can be later removed with morphological operations such as filling holes in the blobs.

From the work done so far it is concluded that the optical flow algorithm is faster without considering the centroids and the foreground detection algorithm has proven to\* be more accurate for high resolution low fps images. But when looking at solving the problem in real time the chairs are turned the optical flow has lower time consumed per frame compared to any other algorithm. If accuracy is required then the foreground detection algorithm has to be used and contrarily if the higher processing speed is required then we will have to go for the optical flow algorithm. More work has to be done to optimize the code for the improvement of both accuracy and time consumed per frame. Another advantage of the optical flow algorithm is that it can function better in isolating the

vehicles in the given range of velocities alone this helps it in eliminating wanted noises from the environment, on the other hand the foreground detection algorithm detects any object that is added newly to the frame. The newly created algorithm performs better than the previous algorithms and is capable of detecting the velocity of the vehicles to 5km/hr accuracy at VGA quality for a range of 40m.

#### ACKNOWLEDGMENT

This project has been sponsored and supported by CTDI under the Young faculty members under research support scheme.

#### REFERENCES

- [1] P. Spagnolo, T. D. Orazio, M. Leo, and A. Distanto, "Moving object segmentation by background subtraction and temporal analysis," *Image and Vision Computing*, vol. 24, pp. 411-423, 5/1/ 2006.
- [2] N. Nath, D. M. Dawson, and E. Tatlicioglu, "Euclidean Position Estimation of Static Features Using a Moving Uncalibrated Camera," *Control Systems Technology, IEEE Transactions on*, vol. 20, pp. 480-485, 2012.
- [3] Y.-C. Lim, M. Lee, C.-H. Lee, S. Kwon, and J.-h. Lee, "Improvement of stereo vision-based position and velocity estimation and tracking using a stripe-based disparity estimation and inverse perspective map-based extended Kalman filter," *Optics and Lasers in Engineering*, vol. 48, pp. 859-868, 2010.
- [4] H. Sakaino, "Moving vehicle velocity estimation from obscure falling snow scenes based on brightness and contrast model," in *Image Processing. 2002. Proceedings. 2002 International Conference on*, 2002, pp. 905-908 vol.3.
- [5] V. K. Chitrakaran, D. M. Dawson, J. Chen, and H. Kannan, "Velocity and structure estimation of a moving object using a moving monocular camera," in *American Control Conference, 2006, 2006*, pp. 5159-5164.
- [6] T. Gandhi and M. M. Trivedi, "Pedestrian collision avoidance systems: a survey of computer vision based recent studies," in *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE, 2006*, pp. 976-981.
- [7] T. H. Lee, Y.-g. Lee, and B. C. Song, "Fast 3D video stabilization using ROI-based warping," *Journal of Visual Communication and Image Representation*, vol. 25, pp. 943-950, 2014.
- [8] Liao, X. Qin, X. Huang, Y. Chai, and X. Zhou, "A monocular-vision-based driver assistance system for collision avoidance," in *Intelligent Transportation Systems, 2003. Proceedings. 2003 IEEE, 2003*, pp. 463-468 vol.1.
- [9] Y. Mau-Tsuen, J. Rang-Kai, and H. Jia-Sheng, "Traffic flow estimation and vehicle-type classification using vision-based spatial-temporal profile analysis," *Computer Vision, IET*, vol. 7, pp. 394-404, 2013.