2009 International Conference on Information and Multimedia Technology

# Vision Based Relative Speed Detection of Moving Objects in a Traffic Environment

Avinash N., Dev Dutt Yadav
Mando Softtech India Pvt. Ltd.
Gurgaon – 122002, Haryana, INDIA
(avinash, devdutt)@mando.com

Murali S.
PET Research Center
PES College of Engineering
Mandya, Karnataka, INDIA
nymurali@yahoo.com

***Abstract*— Speed estimation of an object from a camera is a challenging task in the field of computer vision. Speed of the moving object from a moving camera is furthermore challenging. In this paper, a camera is used to acquire the images. The edges are detected and broken edges are merged to connect the broken components. The corners are detected to further find the optical flow of the same. The points are clustered based on the direction of optical flow of each object, thus marking the same as moving object. The object detected is used to find the displacement with a novel method of view metrology. Further, this is used to find the relative speed of the object, based on the frame rate of the camera used. The errors are computed with the known results with constant speed limits fixed initially.**

***Keywords-View Metrology, Corner Detection, Optical Flow, Moving Object Detection, Speed Calculation.***

## I. INTRODUCTION

Estimating the motion of a moving object in an unknown environment is essential for a number of applications such as autonomous vehicles and driver assistance systems. Autonomous vehicles typically use computer vision for navigation, producing a map of its environment and for detecting speed of objects. A vision-based method for estimating the position and orientation of the virtual display with respect to the environment provides flexible and cost effective solutions. Robust and efficient solutions to the problem are yet to be available.

In application of Forward Collision Warning (FCW) systems developed by TRW [1] use radar-based sensors and cameras to monitor the road ahead. They help provide object recognition and detect relative speeds between the vehicle and an approaching object. When the speed of the object is such that it represents a risk of collision, the driver receives an audible, visual or haptic warning. A forward-looking camera with object recognition capability is mounted on the windshield behind the rearview mirror. This technology provides a low-cost sensor that is able to provide advanced audible, visual or haptic warnings of potential accidents. FCW is able to provide additional functions such as Lane Departure Warning.

A single image acquired of a scene potentially contains geometrical information well enough for obtaining the depth perception. The aim of the work presented in this paper is to extract this information in a quantifiable and accurate way. This process of obtaining the 3-dimensional information requires measurements on a reference scale for which a camera is calibrated or a scene in the image is calibrated before the process. The process of measurements is traditionally an engineering task and hence it should be accurate and robust, just like any other engineering task. The measurements get affected by errors and the unavoidable noise. Thus, a proper treatment of error and its propagation through the chain of computations is necessary.

In [2][3-5], some methods were also investigated for object reconstruction from measurements in a single view both by the computer vision community and the photogrammetric communities. These methods were based on the constraints of the object to be reconstructed such as, edges, co-planarity, parallelism, perpendicularity, etc.

Wang et. al. in [6], make full use of the available scene constraints and recover more metric measurements, rather than affine entities as compared to the work by Criminisi et. al. in [7].

In this paper, we present the methodology of obtaining the three-dimensional information from the two-dimensional images based on the construction of orthographic view of objects from the perspectively distorted images. This is used to develop the theory presented here as applied to solve the speed calculation problem.

## II. SYSTEM SETUP

A forward looking camera is mounted into the car behind the windshield and just below the inside rear view mirror and camera tilt and pan is avoided. Camera frame rate is fixed 30 frames/second so that time delay between two frames can be calculated. These images passed into edge detection and broken edge merging module, this module generates an edge map for each frame. Edge map is taken as input in corner detection module, this module detects the corner and stores corner coordinates points into an array. The array is passed to the moving object detection module, it calculates optical flow between two arrays. Speed of the object is calculated using known vehicle speed. If the Relative velocity of object is equal to vehicle velocity in opposite direction then it is considered as stationary object else it is a moving object.

978-0-7695-3922-5/09 $26.00 © 2009 IEEE
DOI 10.1109/ICIMT.2009.58

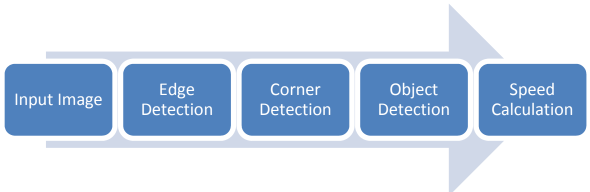



Figure1. System Flow Chart

## III. OBJECT DETECTION

Images used as input is acquired using the camera Allied GUPPY F-036C IRF. The AVT GUPPY F-036C is ultra compact, space-saving IEEE 1394 WideVGA C-mount cameras, equipped with a highly-sensitive MICRON/Aptina 1/3 progressive Scan CMOS sensor. It works up to 60 Hz frequency with full resolution and captures a high 70fps.

Canny edge detector [8] is used to detect edges in input images. Canny Edge Detection is considered to be the ideal edge detection algorithm for images. It uses a filter based on the first derivative of a Gaussian, because it is susceptible to noise present on raw unprocessed image data, so to begin with, the raw image is convolved with a Gaussian filter. The result is a slightly blurred version of the original which is not affected by a single noisy pixel to any significant degree.

After edge detection all edges are not continuous. To make them continuous, broken edge merging is required. In the edge map if the endpoint of an edge is nearly connected to another endpoint fill the gap and make them continuous. Repeat the same process for all the edges. Edge merging process produces an edge map without any broken edges. In the edge map next corner detection technique is used. Curvature of each edge has been calculated and all of the curvature local maxima are considered as corner points, then rounded corners and false corners due to boundary noise and details were eliminated. End points of line mode curve were added as corner, if they are not close to the above detected corners.

Lucas and Kanade [9] optical flow is used to match the feature points. Given point *(u, v)* in image $I_t$ the algorithm finds the point *(u, v)* in image $I_{t+1}$ that minimizes (*e*).

$$e(d_x, d_y) = \sum_{x=u-w}^{u+w} \sum_{y=v-w}^{v+w} (I(x,y) - I(x+d_x, y+d_y)) \quad (1)$$

(a) (b)

(c) (d)

Figure2. (a) input image (b) Detected edges (c) Edge map after broken edge merging process (d) Detected corner Points

All corners are assumes as feature point and nearest neighbor clustering technique used to make cluster for the feature points. The corner points which are nearly connected and similar direction are kept into same cluster. A flow $(OP)$ has 4 features $(x', y', v'_x, v'_y)$ ($x$ and $y$ are the location of the flow; $v_x$ and $v_y$are the direction of the flow). Given the two vectors of optical flows $OP_1(x', y', v'_x, v'_y)$ and $OP_2(x'', y'', v''_x, v''_y)$,Euclidean distance($dist(OP_1, OP_2)$) is defined by following.

$$dist(OP_1, OP_2) = \sqrt{(x' - x'')^2 + (y' - y'')^2 + (vx' - vx'')^2 + (vy' - vy'')^2} \quad (2)$$

If the $dist(OP_i, OP_j)$ is less than a threshold value$\theta_{dist}$, group the $OP_i$ and $OP_j$ into same group. If the case when the$OP_j$is already included in a group G', insert $OP_i$ into the group$G'$.

$$dist(OP_i, OP_j) \leq \theta_{dist} \; and \; OP_i \in G' \rightarrow OP_j \in G' \quad (3)$$

If camera is mounted in the static ground, there is no additional problem to use the optical flow and this nearest neighbor clustering. All the flows have some biased direction in some cases, because camera is also moving. Whenever there is any gap on the road, the camera has so large movement that the next image has large translation. To prevent this 'bump it out' effect, need to find the common biased vector $biased_y(vb_x, vb_y)$ if there is. Thus an optical flow $(OP_i(x, y, v_x, v_y))$ has new flow($OP'_i$) as following.

$$OP'_i = (x, y, v_x - vb_x, v_y - vb_y) \quad (4)$$

All corner points in same cluster are belongs to an object. Bounding box created by the corner point can be considered as boundary of object. Displacement of object is considered as average displacement of all the corner points. If displacement of N corner points are $x_1 + x_2 +, \dots\dots, x_n$ then

$$x_{avg} = \frac{x_1 + x_2 +, \dots\dots, x_n}{N} \quad (5)$$

### IV. DISTANCE MEASURMENT

This section presents a method of view metrology to find out the distance factors of the real world or any object from a single image. The use of perspective geometry construction from one point perspective is used in this method. The input to the method is an image having one vanishing point in it. In this method, we introduce the concept of projections in which the perspective views from in the image are orthographically projected. By converting to the orthographic projection space, the objects can be measured in true dimensions. We use the principles of construction of one point perspective in engineering graphics [10] to solve the method by performing the reverse procedure. The perspective image is represented in Euclidean space and the orthographic projections evolved from them are also represented in the same space and the calculations are performed with the use of analytical geometry. The following section explains about the basic geometry of the system followed by its representation mathematically.

#### *A. Geometry*

The approach of view metrology is performed on the principle of construction of one point perspective projection. For this method, we consider input images of type which yield one vanishing point in them.

The vanishing point 'VP' would be the key geometric cue in this method, although there are many other cues used in this method like ground line and perspective plane line. The vanishing point (VP), ground line (GL) and picture plane (PP) constructed are as depicted in figure 3.

For the measurement of the objects present in the scene, it is required to follow certain steps, viz.

1. Vanishing point determination
2. Construction of the ground line
3. Construction of the picture plane line
4. Dimension finding

NOTE: All the constructions with the layout of measurements made in our experimentation is depicted in figure 4.

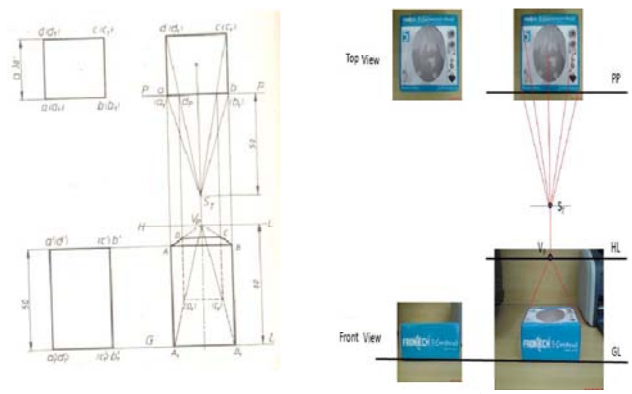
Ground line (GL) is the line parallel to the horizon line. In case of one point perspective, the horizon line passes through the vanishing point and determines the camera roll. Thus the ground line is directly constructed with either of the following two techniques.

• The slope of the ground line is fixed up based on the known camera roll angle. With the obtained slope and at the distance equal to the height of the camera mounting is constructed below the vanishing point.

• A reference parallel edge on the ground in an image which remains parallel even after perspective distortion is used as the reference parallel line to the ground line. To this reference line, a parallel line at the distance equal to the height of the camera mounting is constructed below the vanishing point. This case is for the unknown camera roll.

The picture plane line (PP) is the line parallel to the ground line (and horizon line). This is separated from the picture plane at any convenient distance from each other. Station point 'ST' lies on the perpendicular line from the vanishing point to the picture plane and at the distance of focal length away from the picture plane line. For convenience we situate the station point on the vanishing point and construct the picture plane as a parallel line to the ground line at the distance equal to the focal length of the camera and constructed above the vanishing point.

Ultimately, it is required to find out the dimensions of the objects in the scene. For any object in consideration to be measured we present a method of view metrology from one point perspective which can measure the width, height and depth of the object. Hence our consideration for experimentation is for rectangular co-ordinates system obtained from the perspective geometry. This is carried out for cases of known camera parameters, and also for unknown intrinsic camera parameters and known reference in the scene itself. These are found out by the following:




(a) (b)

Figure 3(a). Schematic of one point perspective construction and 3(b) Representation of the construction of one point perspective using images.

#### *B. Procedure to finding depth of an object*

The schematic representation of the construction for the depth measurement is shown in figure 4.

• $P_1$ being the point in the image of an object whose depth from the picture plane has to be measured. The selected points of the object should make contact on the ground as it appears in the image. Alternatively they are also considered from points which project onto the ground vertically in the image.

• Extend the line formed by joining $V_P$ and $P_1$ so that it intersect the ground line at $P_{g1}(x_{PP1}, y_{PP1})$.

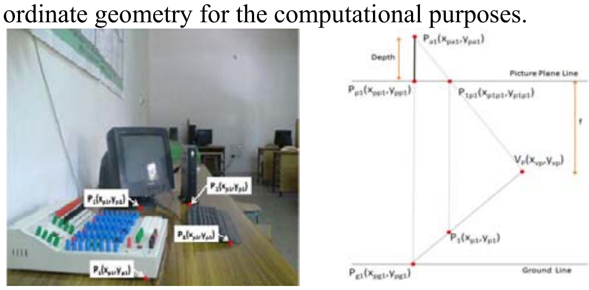
• Project $P_{g1}$ onto the picture plane line such that its foot of perpendicular is at $V_P$ $P_{P1}(x_{PP1}, y_{PP1})$on the picture plane.

• Similarly project $P_1$ onto the picture plane line such that its foot of perpendicular is at $P_{1P1}(x_{P1P1}, y_{P1P1})$ on the picture plane.

• Since from the construction, we know $S_T$ is also the same as $V_P$, Find the intersection point of the two lines formed from point $\overrightarrow{P_{g1}P_{p1}}$and $\overrightarrow{S_T P_{1p1}}$ to get$P_{a1}(x_{pa1}, y_{pa1})$.

• Measure the Euclidian distance from $P_{a1}$ to $P_{p1}$ (i.e.$|\overrightarrow{P_{a1}P_{p1}}|$) to get the actual depth of the object from the picture plane. The sum of this distance and the focal length is the depth of the object from the camera.

The entire procedure of view metrology with the help of projective geometry in one point perspective is deduced into algebraic representation with the help of Cartesian co-ordinate geometry for the computational purposes.




(a) (b)

Figure 4(a). various points of which the depth has to be measured from the image. 4(b) Corresponding projections of the points selected of which the depth is projected on to an orthographic scale.

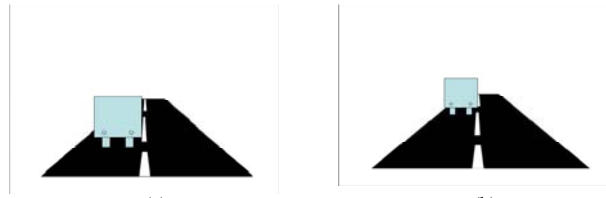
## V. SPEED CALCULATION

Speed is the rate of motion, or equivalently the rate of change of distance. Speed is a scalar quantity with dimensions length/time; the equivalent vector quantity to speed is velocity. Speed is measured in the same physical units of measurement as velocity, but does not contain the element of direction that velocity has. Speed is thus the magnitude component of velocity.

Speed is calculated by dividing the distance travelled upon time taken. To measure the speed of a moving object, we need to know the distance it has moved and the time it has taken to move that distance.

A video sequence is a series of still images acquired at various time instances. These images are regarded as frames. To know how far an object has moved in successive video frames, we need to segment the object in the scene in each frame and then using the metrology principle, calculate the depth of the segmented object in the image. Depth measurement is done on all the frames. The distance travelled by the target object with consideration from one frame to the other is calculated by subtracting the depths in these frames.

The following example of a vehicle moving on a road is schematically represented in figure 5(a), (b) and (c) and they are shown to represent the calculation of depth in each frames. Figure 5(a) is the depiction of the vehicle which is in motion and its position during the time instant '$t_1$'. Similarly figure 5(b) is the depiction of the vehicle which is in motion and its position during the time instant '$t_2$'. Figure 5(c) demonstrates the camera setup and the distance which is actually measured in each of the frame of the video.



(a) (b)

Figure 5(a). Position of the moving object at the time instant $t_1$ is depicted nearer and 5(b) position of the moving object at the time instant $t_2$ is depicted farther. Together with figure 5(a), it represents the displacement of the object

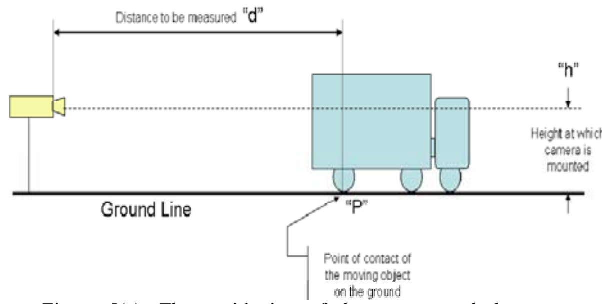



Figure 5(c). The positioning of the camera and the measurement distance is depicted.

Generally, speed is measured as distance travelled over time taken. Since the frame rate is the inverse of time, (i.e. a frame rate of 25 fps has a time between frames of 1/25 seconds) we can modify the speed formula to use the frame rate directly. This assumes the object is observed in successive frames.

*Speed = (Distance travelled in one frame) X (Frame rate)* (6)

If the distance travelled in one frame is in meters and frame rate is in frames per second (fps), then speed calculated is represented in meters per second (m/s). To calculate the distance travelled in each frame, the distance calculation formula derived in the previous chapter is used and the distance in each frame is calculated to the camera position. Thus over successive frames the difference in depth is calculated to find the distance travelled by the moving object in the video.

In kinematics, relative velocity is the vector difference between the velocities of two objects, as evaluated in terms of a single coordinate system, usually an inertial frame of reference unless specifically stated otherwise.

For example, if the velocities of particles A and B are $V_A$ and $V_B$ respectively in terms of a given inertial coordinate system, then the relative velocity of A with respect to B (also called the velocity of A relative to B, or $V_{A\,rel\,B}$) is

$$V_{A\,rel\,B} = V_A - V_B \quad (7)$$

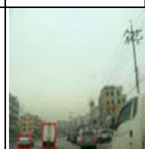
Conversely, the velocity of B relative to A is

$$V_{B\,rel\,A} = V_B - V_A \quad (8)$$

## VI. RESULTS AND DISCUSSIONS

For experimentation camera frame rate is fixed as discussed in section 2. All images are collected while source vehicle is moving with constant speed of 30 Kmph and object vehicle is moving with varying speed from 40 Kmph to 60 Kmph. As shown in Table 1 below object vehicle is a car with known speed. Actual speed of moving object is calculated.

TABLE I. INPUT IMAGE WITH VARIOUS OUTPUT IMAGES AND SPEED MEASURES

| Input Image | Detected Corner Points | Output Image | Depth (m) | Speed (m/s) |
|---|---|---|---|---|
|  |  |  | **64** | **11.1** |
|  |  |  | **71** | **12.5** |
|  |  |  | **80** | **13.3** |
|  |  |  | **91** | **14.2** |
|  |  |  | **104** | **15.3** |

We validate the method with the standard values measured from a most reliable metric tape of some of the dimensions that are also found out with our method. Thereafter any discrepancies in the values obtained in both the cases are tabulated. The results for the scene in figure 6 are computed for various images taken. This is repeated for different illumination conditions and thus a large number of images are acquired. Percentage experimental discrepancy is calculated by the equation given below.

$$\%\,Experimental\ Descripancy = \frac{|M_R - M_O|}{M_R} \times 100 \quad (9)$$

Where,

$M_R$ = Reliable Measure

$M_O$ = Experimental Measure

Error analyses performed on 500 frames and at 3 different object vehicle speeds (40, 50 and 60) and average error has been calculated by finding the median of the trial runs of each speed. Experimental setup was same as discussed above. Result is tabulated in Table 2.

TABLE II. DISCREPANCY AT VARIOUS SPEED

| Object vehicle speed | Min discrepancy (%) | Max discrepancy (%) | Median value of the Experimental discrepancy (%) |
|---|---|---|---|
| 40 | 1.1 | 9.8 | 4.8 |
| 50 | 1.3 | 10.1 | 5.2 |
| 60 | 1.5 | 10.5 | 5.5 |

## VII. CONCLUSION

In this paper, we present a method to detect the moving object on the traffic environment by using images acquired from the camera mounted on a moving vehicle. Further we present a novel method to find the distance of the moving object displacement. Thus the speed of the relative speed of the object in motion is calculated. Results are also tabulated and presented in this paper to further discuss the experimental discrepancy in the results.

## REFERENCES


[1] http://www.trw.com/sub_system/forward_collision_warning_1

[2] T. Kim, Y. Seo, K. Hong, "Physics-based 3D position analysis of a soccer ball from monocular *image sequences",* Proceedings International Conference on Computer Vision, Pg. 721–726, 1998.

[3] D. Liebowitz,A. Criminisi,A. Zisserman, *"Creating architectural models from images",* Proceedings ofEurographics,Milan, Italy, Pg. 39–50, 1999.

[4] M. Wilczkowiak, E. Boyer, P. Sturm, "Camera calibration and 3D reconstruction from single *images using parallelepipeds",* Proceedings of International Conference on Computer Vision, Vancouver, Canada, July, Vol. I, Pg. 142–148, 2001.

[5] P. Gurdjos, R. Payrissat, "About conditions for recovering the metric structure of perpendicular planes from the single ground plane to image homography", Proceedings of International Conference on Pattern Recognition, Barcelona, Spain, Sept., Vol. I, Pg. 358–361, 2000.

[6] G. Wang, Z. Hu, F. Wu, H.T. Tsui, *"Single view metrology from scene constraints",* Image and Vision Computing, Vol. 23(9), Pg. 831-840, 2005.

[7] A. Criminisi, I. Reid, A. Zisserman, *"Single view metrology",* In Proceedings 7th International Conference on Computer Vision, Kerkyra, Greece, Pg. 434-442, 1999.

[8] J. Canny, *"A Computational Approach to Edge Detection",* IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 8, Pg. 679-714, 1986.

[9] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In Proc. of the 7th IJCAI, pages 674–679, Vancouver, Canada, 1981.

[10] N.D. Bhatt, V.M. Panchal, *"Engineering Drawing"*, Charotar Publication House, 49th Edition, ISBN 81-85594-58-9, 2008.