Leader Follower Formation Control of Ground Vehicles using Dynamic Pixel Count and Inverse Perspective Mapping

Ashish Prasad[1], Gokul.M[2], Bharath.M.K[3] and S.M.Vaitheeswaran[4]
[1]Student (M.Tech, National Institute of Technology), Surathkal,Karnataka,575025
[2]Student (M.Tech, Amrita Vishwa Vidyapeetham), Coimbatore,TamilNadu,641112
[3]Project Assistant, NAL, CSIR, Bangalore, Karnataka, 560017
[4]Principal Scientist, ALD, NAL, CSIR, Bangalore, Karnataka, 560017 (corresponding and lead author)

Abstract— This paper deals with leader-follower formations of nonholonomic mobile robots, introducing a formation control strategy based on pixel counts using a commercial grade electro optics camera. Localization of the leader for motions along line of sight as well as the obliquely inclined directions are considered based on pixel variation of the images by referencing to two arbitrarily designated positions in the image frames. Based on an established relationship between the displacement of the camera movement along the viewing direction and the difference in pixel counts between reference points in the images, the range and the angle estimate between the follower camera and the leader is calculated. The Inverse Perspective Transform is used to account for non linear relationship between the height of vehicle in a forward facing image and its distance from the camera. The formulation is validated with experiments

Keywords—vision based range and angle measurement, dynamic pixel count controller, inverse mapping projection

I. INTRODUCTION

Autonomous Unmanned Systems encompassing a variety of technologies requires some form of situational awareness of the robot's environment. Generally, this is achieved through different sensors which provide additional information, but this comes at the expense of additional computational and increased costs to the system.

This paper proposes a low cost, low computation vision solution to situational awareness in a formation control convoy formation problem of autonomous ground vehicles in a leader-follower framework. The situational awareness to the agent, through vision sensor, is classified under level-4 of autonomous control levels [1-3].

The Vision system discussed within this paper uses commercial grade electro-optic camera, and differs from traditional monocular cameras in its computational method. The traditional camera generates tens of thousands of data points creating a very fine map of the environment, but does not provide any insight into the recognition of objects within the scene. The process is data intensive as sensors of this nature provide mass information in its attempt to numerically characterize the most complex of environments.

The simplified vision system presented here in the paper works by first recognizing the leader as a high contrast symbol within the camera images. The correlation problem is reduced from tens of thousands of pixels to a set of feature points and in some cases just one feature point. A vector for the relative position of the robot to the scene feature is calculated based on the dynamic pixel count in the images and is used to generate the motion commands for the follower robot. The process is repeated iteratively until the robot is correctly localized. This technology also requires an algorithm for feature point extraction, yet, this type of algorithm is heavily researched, well documented and several fast open source algorithms already exist. If the leader under measurement is not perpendicular to the optical axis, a relationship between the displacement of the camera movement and the difference in pixel counts is used to measure the photographing distance and incline angle is used.

The angle of view under which a scene is acquired and the distance of the objects from the camera, namely the perspective effect, contributes different information content to each pixel of an image. To cope with this problem a geometrical transform *Inverse Perspective Mapping*, (IPM) [4], is used in this paper. This allows removing the perspective effect from the acquired image, remapping it into a new 2-dimensional domain (the *remapped domain*).

II. PROBLEM STATEMENT

Leader-follower formations of non-holonomic ground vehicles are assumed. Each ground vehicle is equipped with a camera which provides only the bearing and position to the other robots. All computation is carried out on the leader, who receives the measurements from all followers, and transmits the control velocities to all followers.

Vehicle kinematics obeys the bicycle model. The controller's goal is to keep a desired position and orientation with respect to the leader. The problem is to find the relative position and orientation with respect to the leader so that it can be provided as feedback to the control loop and all information gathered by

the entire network can be expressed with respect to a single reference coordinate frame, the one of the leader.

III. METHODOLOGY

This section describes the algorithm used to identify a leader object in the sequence of images from a color video camera mounted on the follower vehicle to estimate the range and bearing to the target. Accurate distance measurements are crucial, as the follower robot will be planning its path based upon the data from the vision system. Finally, it is important for the hardware and software to have a fast update rate with a minimum of delay, as both can directly affect the dynamic performance of the control system for the camera.

With these requirements in mind, this work presents a system selected for their individual characteristics and complementary nature: 1) A color detector working in the RGB image space, and 2) Inverse Perspective Mapping (IPM) to create a top down view of the image 3) Angle and Distance Estimation of the leader and follower based on pixel counts in the images

Color Detection

The color detector identifies areas that match the leader's color characteristics using a number of steps shown below:

1. The image is separated into its constituent Red Green and Blue channels. The luminance $f_I(x, y)$ of the pixel (x, y) is described as [5]:

$$f_I(x, y) = 0.299R_{(x,y)} + 0.578G_{(x,y)} + 0.114B_{(x,y)} \quad (1)$$

2. Taking one of the RGB channels, the algorithm calculates the average pixel intensity from the area immediately in front of the vehicle.

3. A binarization [5] is performed using the average value from step 2 as the threshold. The pixels similar to surface (path) are thresholded to white and others are threshholded to black, where 1 signifies a road pixel, and a 0 is a non-road pixel.

4. The steps 2 and 3 are repeated for the other channels (Green and Blue).Morphological image processing operations [5], such as dilation, erosion, opening and closing are used after the process of image segmentation, in order to enable the underlying shapes to be identified and optimally reconstruct the image from their noisy precursors. Connected-components that share similar pixel intensity values are identified and then connected with each other. The connected-component labeling scans an image and groups pixels into one or more components according to pixel connectivity. Once all groups are determined, each pixel is labeled with a grey level on the basis of the component.

5. Finally, the algorithm merges the 3 new binary images into a new RGB image.

Inverse perspective mapping

IPM is a mathematical technique whereby a coordinate system may be transformed from one perspective to another. In this case we use IPM to remove the effects of perspective distortion of the road surface in the forward facing image, to an undistorted top-down view as shown in Fig. (1). As can be seen lines that appear to converge near the horizon are made parallel To get a bird eye's view image, a perspective transform relating the pixel (u,v) values seen by the camera in the original image is obtained through a transformation matrix as shown in figure.
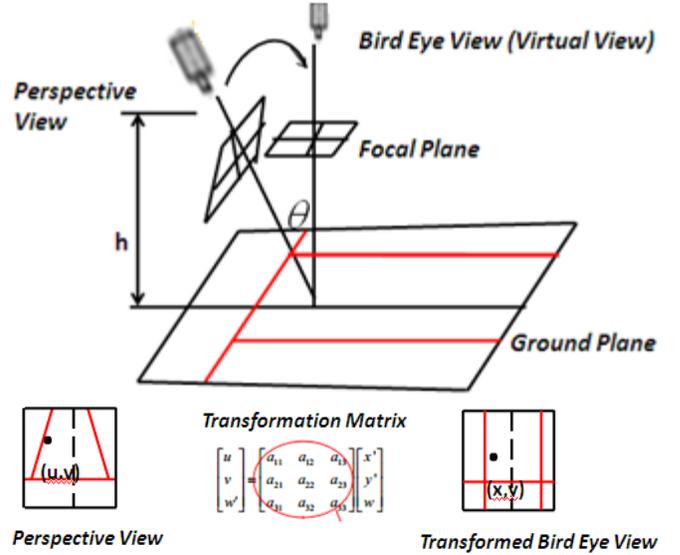

Fig.1 IPM Transformation

Transforming the image in this manner removes the non linearity of distances and the orientation. We can focus our attention on only a sub-region of the input image, which helps in reducing the run time considerably. In order to create the top-down view, we first need to find the mapping of a point on the road to its projection on the image surface. To get the IPM of the input image, we assume a flat road, and use the camera intrinsic (focal length and optical center) and extrinsic (θ ,and height above ground) parameters to perform this transformation. We start by defining a world frame $\{F_w\} = \{X_w, Y_w, Z_w\}$ centered at the camera optical center, a camera frame $\{F_c\} = \{X_c, Y_c, Z_c\}$, and an image frame $\{F_i\} = \{u, v\}$ as show in Fig.(1). We assume that the camera frame X_c axis stays in the world frame $X_w Y_w$ plane The height of the camera frame above the ground plane is h. Starting from any point in the image plane $_iP = \{u, v, 1, 1\}$, it can be shown that its projection on the road plane can be found by applying the homogeneous transformation

This mapping is expressed as [6]:

$$_iP = \{u, v, 1, 1\}= [K][T][R](x, y, z, 1) \quad (2)$$

Where R is the rotation matrix:

$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3)$$

T is the translation matrix:

$$T=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -h/\sin\theta \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(4)

And K is the camera parameter matrix:

$$K=\begin{bmatrix} f*ku & s & uo & w \\ 0 & f*kv & vo & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

(5)

Here, f is the focal length of the camera and ($ku*kv$) is the aspect ratio of the pixels, uo and vo *the camera cetre coordinates*. These transformations can be efficiently calculated in matrix form for hundreds of points.

Fig.(2) shows a typical IPM sample.The leader vehicle is extracted using object detection process and a background subtraction procedure from the transformed bird eye view images of the follower.
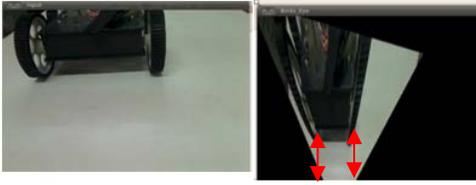
(a) Perspective (frontal)View   (b) Bird Eye View

Fig.2 IPM Sample

A pose estimation of the leader vehicle is performed in order for the follower vehicle to navigate. Only the area directly visible in front of the host vehicle is analyzed. Cropping is performed on the road segmented image, in order to get the region of interest which provides us the distance and angle of orientation of leader in term of pixels. The two red lines in Fig 2(b) indicate the distance of the leader vehicle which is being taken as two cropped images for computation as shown in Fig.(2).

Target Localization based on pixel count

In this section we outline the methodology used for range and angle estimation based on the pixel count number variation in images for the purpose of localization of the object in the images. For this we first capture an image incorporating a known-dimension trapezoid, and then the proportion relationship between the real-dimension and the image-dimension of the trapezoid are found. In this we have two cases: a. the targeted leader vehicle lying on the plane perpendicular to the optical axis of the follower vehicle camera, b. the targeted leader vehicle lies on an oblique tangent plane with respect to the optical axis of the follower vehicle camera.
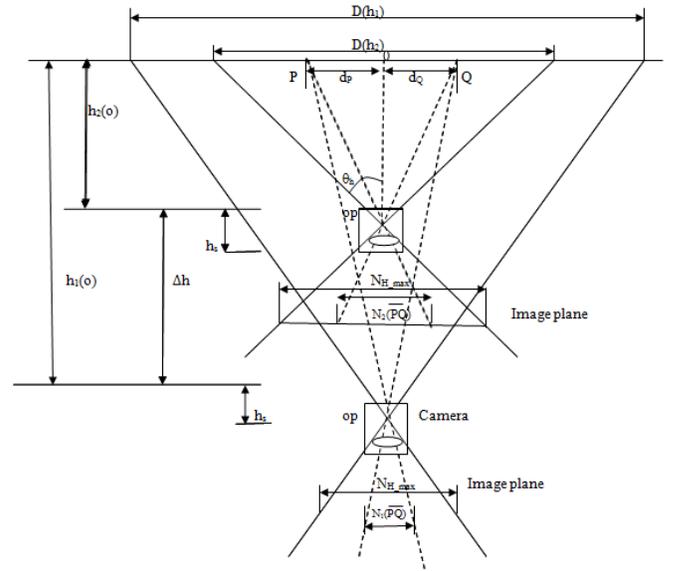*Case a*: Leader perpendicular to the optical axis of the follower camera:

Fig.3 Pixel Count Measurement Geometry using Bird Eye View Image for perpendicular Case

As shown in Fig.(3) the target plane is assumed to be perpendicular to the optical axis and the relationship between distance and variation of pixel counts at different viewing distances are obtained using Figure and geometry. The target plane formed by the leader vehicle is represented by points P, O, Q. A simple relationship between the change in leader follower distance Δ, and the difference in pixel counts between the reference points in the images at different viewing distances, are obtained with the assumption that the distance between reference points P and Q will not change. We have the following relationship between pixel counts and distances:

$$\frac{d(\overline{PQ})}{n_1(\overline{PQ})}=\frac{d(h_1)}{n_{max}}$$

(6)

$$\frac{d(\overline{PQ})}{n_2(\overline{PQ})}=\frac{d(h_2)}{n_{max}}$$

Where, points P and Q is the target plane distance. $n_1(\overline{PQ})$ and $n_2(\overline{PQ})$ are the pixel counts between points P and Q on the image plane at distances $h_1(0)$ and $h_2(0)$, respectively. Δh is the change in distance between the leader and follower positions. h_s is the distance between the optical origin (OP) and the front end of the camera on the follower. From (1) and (2), we have:

$$\frac{d(h_1)}{d(h_2)}=\frac{n_2(\overline{PQ})}{n_1(\overline{PQ})}$$

(7)

Because of the displacement, $\Delta h = h_1(0) - h_2(0)$, we have:

$$\frac{n_2(\overline{PQ})}{n_1(\overline{PQ})}=\frac{d(h_1)}{d(h_2)}=\frac{h_1(0)+h_s}{h_2(0)+h_s}$$

(8)

Substituting $h_1(0) = h_2(0) + \Delta h$ into (4), viewing distances $h_1(0)$ and $h_2(0)$ of the leader on the perpendicular plane can be obtained as:

$$h_1(O) = \frac{n_2(\overline{PQ})}{n_2(\overline{PQ}) - n_1(\overline{PQ})} * \Delta h - h_s$$

or

$$h_2(O) = \frac{n_1(\overline{PQ})}{n_2(\overline{PQ}) - n_1(\overline{PQ})} * \Delta h - h_s$$

$$(9)$$

Also, distance between points A and B can be obtained as:

$$\overline{AB} = \frac{2(h_2(0) + h_s) * \tan\theta_H}{N_{H\_max}} * N_2(\overline{AB})$$

$$= \frac{2(h_1(0) + h_s) * \tan\theta_H}{N_{H\_max}} * N_1(\overline{AB})$$

$$(10)$$

where, θ_H is the maximal horizontal view angle of the camera and N_{H_max} is the maximal pixels in a horizontal scan line of an image frame, which is fixed and known as a priori irrelevant of photographing distances. As a result of this relationship, the distance between an arbitrarily-selected point P and the origin O can also be represented as follows:

$$\overline{PQ} = \frac{2(h_2(O) + h_s) * \tan\theta_H}{n_{H\_max}} * n_2(\overline{PQ})$$

$$= \frac{2(h_1(O) + h_s) * \tan\theta_H}{n_{H\_max}} * n_1(\overline{PQ})$$

$$(11)$$

The above-mentioned relationship is only valid for leader objects or planes that are perfectly perpendicular to the optical axis.

Case b: Leader oblique to the optical axis of the follower camera:

Fig. (4)shows the configuration of the proposed image-based system for measuring the leader position on an oblique surface based on variation of pixel counts, where an incline angle θ_m exists between the oblique and perpendicular planes. Virtual planes are introduced to address the problem to measure the incline angle θ_m. Thus, distances between the camera and the virtual planes for different positions can be obtained as:

$$h_1(P) = \frac{n_2(P)}{n_2(P) - n_1(P)} * \Delta h - h_s$$

$$h_2(P) = \frac{n_1(P)}{n_2(P) - n_1(P)} * \Delta h - h_s$$
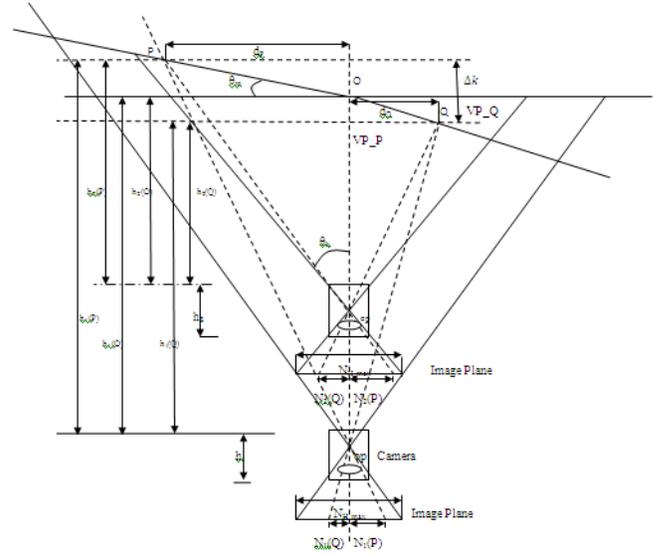
Fig.4 Pixel Count Measurement Geometry using Bird Eye View Image for Oblique Case

$$h_1(Q) = \frac{n_2(Q)}{n_2(Q) - n_1(Q)} * \Delta h - h_s$$

$$h_2(Q) = \frac{n_1(Q)}{n_2(Q) - n_1(Q)} * \Delta h - h_s \qquad (12)$$

where, $h_1(P)$ and $h_1(Q)$ are distances from the camera at position 1 to virtual plane VP_P and virtual plane VP_Q, respectively. $h_2(P)$ and $h_2(Q)$ are distances from the camera at position 2 to virtual plane VP_P and virtual plane VP_Q, respectively. $n_n(P)$ stands for the pixel counts in the image plane for point P when image is taken at position n.

From (9), (10), (11), and (12), we have:

$$h_1(P) - h_1(Q) = h_1(P) - h_1(Q) = \Delta k$$

$$\Delta k = (d_P + d_Q)\tan\theta_m$$

$$\theta_m = \tan^{-1}\frac{\Delta k}{d_P + d_Q}$$

$$(13)$$

where, Δk is the distance between the virtual planes VP_P and VP_Q, and θ_m is the incline angle. d_P and d_Q represent the distance from point P and point Q to the optical axis, respectively, which can be expressed as follows:

To this end, the distance from the optical origin at positions 1 and 2 to the oblique plane intersecting the optical axis can be expressed as:

$$h_1(O) = h_1(P) - d_P\tan\theta_m = h_1(Q) + d_Q\tan\theta_m$$

$$h_2(O) = h_2(Q) - d_P\tan\theta_m = h_2(Q) + d_Q\tan\theta_m$$

$$(14)$$

and the distance between two arbitrary points P and Q as $D(\overline{PQ})$ is derived as $D(\overline{PQ}) = \left| d_P + d_Q \right| \sec \theta_m$

Vehicle kinematics and control

The goal of the vehicle controller is to use the leader's range and bearing from the vision system to force the follower vehicle to follow the leader's path. The controller designed to follow the path the leader is delayed by an arbitrary following time, τ (typically 10 seconds or so).

If (x, y) is the global position of the follower and (x_0, y_0) is the position of the leader, the controller attempts to make $(x(t), y(t))$ track $(x_0(t - \tau), y_0(t - \tau))$ For simplicity, the delayed leader position is defined $(x_d(t), y_d(t))$.The position of the leader vehicle, (x_0, y_0), is determined using the camera vision data relative to the current follower position, as described in Section.

The controller uses a bicycle model for a kinematic model of the vehicle, which is commonly used in vehicle following applications. Referring to Fig.(5), (x, y) is the follower position, v_c is the commanded speed, and γ_c is the commanded steering. The control laws are given as

$$v_c = v_d + k_{p,e_1} e_1, \qquad k_{p,e_1} > 0 \qquad (15)$$
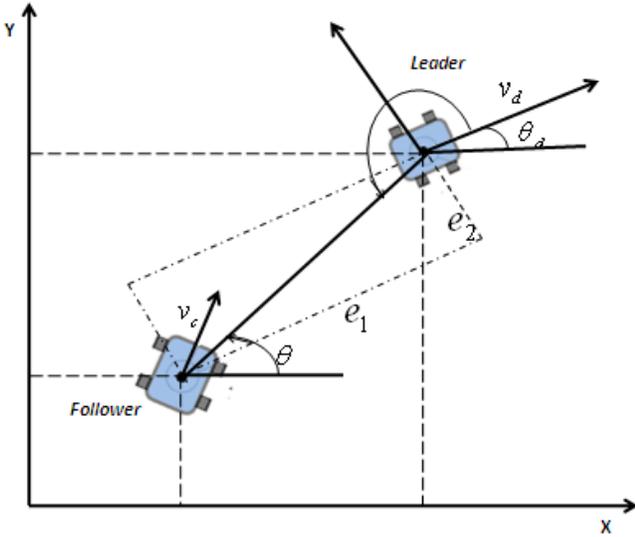
Fig.5 Leader Follower Geometry and Co-ordinate system

$$\gamma_c = k_{p,e_2} e + k_{p,e_{\theta 1}} e_\theta, \qquad k_{p,e_2} e_2, k_{p,e_{\theta 1}} e_\theta, > 0 \qquad (16)$$

$$e_1 = (x_d - x)\cos\theta + (y_d - y)\sin\theta$$

$$e_2 = -(x_d - x)\sin\theta + (y_d - y)\cos\theta$$

$$e_\theta = \theta_d - \theta \qquad (17)$$

IV. EXPERIMENTAL SET UP AND IMPLEMENTATION

In order to show the effectiveness of the algorithms, experiments were performed on a set up consisting of two ground vehicles in a convoy formation using autonomous control for the master and vision control for the slave. The robot vehicle frame and chassis is a model manufactured by Robokit. It has a differential 4-wheeled robot base with two wheels at the front and the two wheels are at the back. The electronics is developed in-house. The vehicle uses four DC motors (geared) for driving the four wheels independently. The DC motors are controlled by Atmega microcontroller and controller board is shown in figure. For this purpose, the C language, together with the OpenCV library of is used.

The follower tracks the remotely controllable leader moving in an arbitrary trajectory, while keeping a predefined inter-distance. The control system has two parts - platform inter-distance control and control of follower orientation with respect to the leader. The follower orientation control provides actions needed for follower to face the leader, which is accomplished when the leader is located in the center of image captured by the follower. The distance between the leader and the follower is obtained from the leader size in current frame. An increase in the platform inter-distance causes the decrease in the platform size in an image, and vice versa. This dependency has been experimentally obtained.

The difference between current and desired platform inter-distance causes the follower forward/backward motion, while the leader's displacement from the image center causes the follower rotation until the leader appears in the image center. Block diagram of the designed control system, containing two described subsystems, is shown in Fig.

Three different postures (angles of view) of the platform at three separation distances have been examined. For each posture and distance (9 cases) 90 frames have been processed (15 fps). Determination of the position and orientation of the leader is achieved by estimating its distance and relative orientation with regard to the follower using the procedure outlined in Section III .

Only the area directly visible in front of the host vehicle is analyzed. Cropping is performed on the road segmented image, in order to get the region of interest which provides us the distance and angle of orientation of leader in term of pixels.
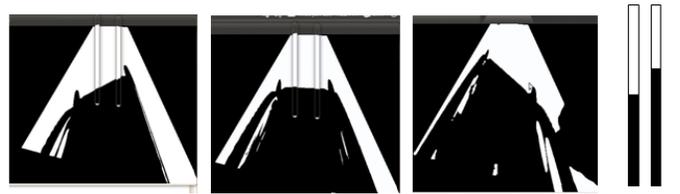
Fig.6 Distance and orientation of the leader vehicle.
(a) left orientation 35^0, (b) straight movement 0^0 (c) Right orientation 50^0 (d) cropped version used for pixel count

In case of front view (straight) as shown in fig 6(b) and for the different distances considered results show very good localization and following capabilities. For the left and right views shown in Fig (6a and 6c) the results are also in good agreement indicating the validity of the design and implementation. Fig.(7) shows typical snapshots capturing the faithful change in movements in straight line and orientation of the follower vehicle as demanded indicating the validity of the approach and method used in the paper.
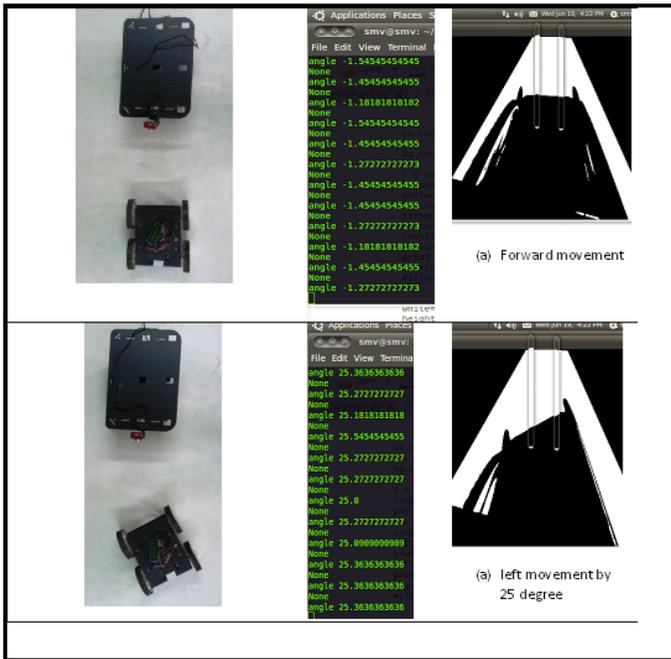
Fig. 7 Leader Follower experimental geometry ,angle measurement and Bird Eye View images showing pixel count at two points for straight and angle maneuvering

V CONCLUSIONS

The algorithms developed for vision detection and localization provide a simple but effective processing approach in real time for the leader-follower formation control approach considered in the paper. The solution is based on pixel variation of the images by referencing to two arbitrarily designated positions in the image frames. The tracking enables a follower to determine its relative position and orientation with respect to the leader in its field of view. Performance of the proposed method has been demonstrated on a laboratory setup composed of two mobile robot platforms. Experimental results demonstrate that overall accuracy of recognition is very good (in some case even 100%), thus providing accurate calculation of distance between the leader and the follower. Future work will be toward improvement of robustness of the proposed method with respect to various environment conditions (light intensity, shadows, etc) as well using multiple vehicles for validations.

REFERENCES

[1] Autonomy Levels for Unmanned Systems (ALFUS) Framework, NIST Special Publication 1011, Gaithersburg, Maryland, September 2004.

[2] U.S. DOD OUSD (AT&L) defense Systems/Land Warfare and Munitions, FY2005 Joint Robotics Program Master Plan. http://www.jointrobotics.com

[3] U. Ramachandran, K.Hong, L. Iftode, R. Jain, Rajnish Kumar, K. Rothermel, Junsuk Shin, R.Sivakumar, "Large-scale Situation Awareness with Camera Networks and Multimodal Sensing", *Proceedings of the IEEE, vol 100, No4, pp 878-892* April 2012.

[4] Davide Scaramuzza and Roland Siegwart, "Appearance-Guided Monocular Omnidirectional Visual Odometry for Outdoor Ground Vehicles", *IEEE Trans. On Robotics, Vol. 24, No. 5, pp 1015-1026* October 2008.

[5] E N Malamas , E G.M Petrakis, M Zervakis, L Petit, Jean-Didier Legat, "A survey on industrial vision systems, applications and tools*", Image and Vision Computing, Vol. 21, Issue 2, 10,* pp 171–188 February 2003.

[6] Mohamed Aly, "Real time Detection of Lane Markers in Urban Streets", 2008 IEEE Intelligent Vehicles Symposium Eindhoven University of Technology Eindhoven, The Netherlands, June 4-6, pp 7-12.