2009 IEEE International Advance Computing Conference (IACC 2009)
Patiala, India, 6–7 March 2009

# Binocular Stereo Vision Based Obstacle Avoidance Algorithm for Autonomous Mobile Robots

Saurav Kumar, Student Member IEEE and IEEE Computer Society
Computer Science Deptt. , Delhi College of Engineering, India
skumar08@ieee.org

**Abstract: Binocular Stereo vision system has been actively used for real time obstacle avoidance in Autonomous Mobile Robotics for the last century. The computation of free space is one of the essential tasks in this field. This paper describes algorithm for obstacle avoidance for mobile robots which can navigate through obstacle. While most of the paper based on stereo vision works on the disparity image but I am proposing a method based on reducing the 3D point cloud obtained from stereo camera after 3D reconstruction of the environment to build a Stochastic representation of Environment Navigation Map. The algorithm assigns each cell of the grid with a value (Free or Obstacle or Unknown) which helps the robot avoid obstacles and navigate in real time. The algorithm has been successfully tested on "Lakshya"- an UGV‡ platform in both outdoor and indoor condition.**

## I. INTRODUCTION

Localization, Map Building and obstacle avoidance are mostly performed using 2D maps (occupancy grids, line based maps) and 2.5D maps (traversability maps) for navigation but only few researchers have looked on the problem dealing with 3D range data. One of the most common precise sensors to acquire range data is a LRF, Laser Range Finder. Range data obtained from 2D LRF can not only be used for building 2D/2.5D maps because Laser is unable to detect a table/chair as a complete obstacle instead as a pairs of pole(Due to each Leg) separated by distance between the legs of the table. High Definition LIDAR e.g. Velodyne HDL-64E† is one of the commercial 3D LRF. The 3D laser data has to be clustered first for the segmentation process. An algorithm based on RBNN (Radially Bounded Nearest Neighbor graph) has been proposed [1]. While highest accurate estimate of the range information available, these Commercial 3D LRF are not suited for general research because firstly its higly expensive, secondly it requires to be fixed in one position and thirdly it is not designed for time critical application.
Since most of the mobile robots application requires it to move in a real environment, therefore recognition of surrounding objects is a big subject. The other most widely used sensor which can provide a cost effective solution to above problem faced by LRF is a stereo vision camera. The single frame stereo vision system has been used for obstacle avoidance during DARPA grand Challenge 2005 by TerraMax [2]. Processing 3D Data acquired from stereo is a challenging task. This task is majorly classified as two algorithms one based on disparity based and other based on 3D space based. Disparity based algorithm are more popular because they work directly on disparity map or depth map. The "V Disparity" [3] is well known for real time obstacle avoidance algorithm which again was used in DARPA Grand Challenge.
3D space based algorithm are mainly used for ego pose estimation.[4] [5].Obstacles are detected as clusters of image edge points reconstructed from 3D space[6].Other major research work have been done to detect Road Surface [7],[8] which concerns the field of autonomous driving and driving assisting system. The Road/Lane presented above can't be applied to autonomous robot in indoor environment.
For indoor environment, an obstacle avoidance algorithm based on disparity maps is dealt [9].The problem associated with disparity map is "Spikes" is addressed in [10]. Still the problem of disparity map is that it combines both range and height measure.
The algorithm presented in this research paper addresses the problem described above. Algorithm takes two stereo images as input and reconstructs the 3D Environment as point cloud. The point cloud is processed to remove objects like newspapers lying on the floor (which is a problem to disparity maps).The post 3D point cloud file is used to generate map of the environment which assigns each cells its occupancy value.
Section II briefly introduces the state of art involved viz The stereo vision camera, Stereo Processing to reconstruct the environment, Data structure for Navigation Map. Section III describes the Algorithm proposed. Section IV describes the experiment setup and Results. Experimental gives insight into indoor and outdoor autonomous robot developed for this research purpose. The future prospects of this paper are outlooked in Section V.

## II. STATE OF THE ART

### A. *Stereo Vision Camera*

We have used a Bumblebee Stereo Vision Camera for stereo data processing. Stereo Vision Camera have in housed two monocular SONY CMOS camera.
Stereo Image processing can be groped majorly into three points as firstly establishing correspondence in the two views viz. left view and right view, then disparity is calculated as the relative displacement between features

Figure 1. Pair of Left and Right Images

Figure 2 Pair of Rectified Left and Right Images

Coordinate in each image and finally camera geometry is used to extract the 3D information about the environment( (Refer Fgure 1,2,3,4).

Disparity for the feature A will be defined as D(A) = $x(A_{left}) - x(A_{right})$ and if the disparity of a point B is greater than disparity of the other point C then point B is more closer to point A.

From Camera geometry it has been shown that following relation holds for disparity for a point.

$$D(P) = \frac{B}{Zp} \times \frac{f}{Px} \quad \text{-----------} \ (1)$$

Where

D(P) is the Disparity Value for the point P , B is Baseline (Distance between centre of two cameras) and f is the focal length of the camera. $P_x$ is the horizontal pixel-size of the camera chip.

*B. 3D Point Cloud Model*

The 3D Point cloud model generated from the disparity map is according to the following figure.

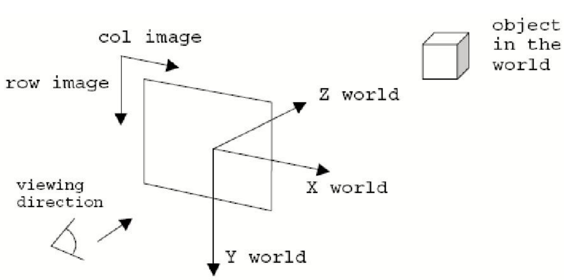



Figure 5 : Point Cloud Coordinate Model

Z represents the depth of the object from the camera. Y gives the height information. (X,Y,Z) is calculated from the disparity values using the following equation:-

Figure 3 Pair of Edges images

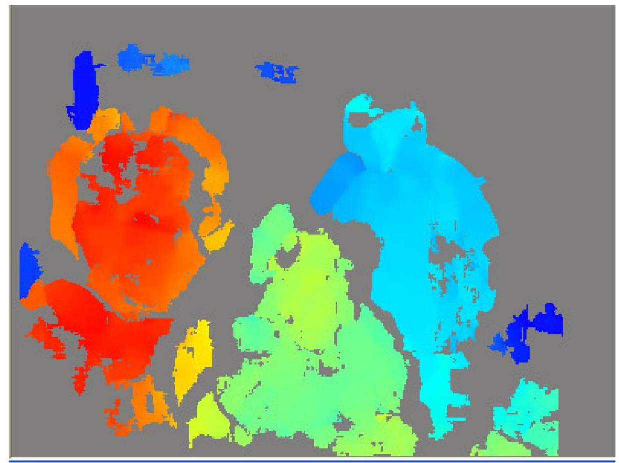


Figure 4 Disparity Image

$$\mathbf{x = P \, . \, X} \quad \text{--------------} \ (2)$$

where x is the image point , P is the 3x4 camera's Projection Matrix and X is the world Point. In general

$$\mathbf{P = K_{3x3} \, [R|t]_{3x4}} \quad \text{----------------} \ (3)$$

Where K is the intrinsic Calibration Matrix for a camera and R,t are Extrinsic Parameters.[11]

*C. Data Structure for Navigation Map*

Navigation Map is a two dimension matrix with each cell having three properties. Firstly total number of points in a cell , i.e point density for a cell and secondly the mean height of all the points in that particular cell. For e.g let us suppose in a cell, n is the point density. Points haves weights as $h_n$ then mean weight for the cell is

$$(h_1+h_2+h_3 \ldots\ldots\ldots\ldots\ldots h_n)/n$$

Third field for the cell is its type i.e whether it is occupied (0) or it is unoccupied (1) or unknown (2). Initially all the cells are assigned unknown value. On the basis of these two properties, third property is decided in the Algorithm.

## III. ALGORITHM DESCRIPTION

*A. Post Processing of 3D point Cloud*

Point cloud obtained from 3D reconstruction of the stereo Vision Camera contains around 700,000 points in our

given test environment. We start to build an elevation map it will be possible but wouldn't be computationally optimal in real time. Lot of computation power will be wasted. For the algorithm to run in real time some of the data have to be pruned before it is processed by the algorithm to build the navigation map.
The Various limits have been set and passed as an parameter to the function which is generating point cloud information of the information. The terms written above have the same correspondence as mentioned above.

A.1 *Floor Cancellation*

Prediction of the disparity for the ground has been discussed in [12].

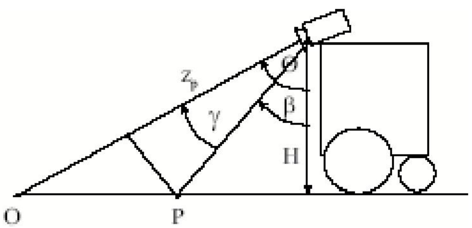



Figure 6 : Geometrical Consideration for expected disparity in an Image

Θ is the angle stereo vision camera's viewing direction is making with the normal to ground. Depth of the pixel $z_p$ can be calculated from the figure 7 as

$$Zp = \frac{H.\cos\gamma}{\cos\beta}$$

With $\gamma = \arctan((v_{p..}\, p_y)\ /\ f)$ where $p_y$ is the vertical pixel size of the camera and $vp$ is the vertical pixel coordinate in the image relative to optics center.

⇨ $\beta = \Theta + \gamma$
⇨

since in our case $\Theta = 90^0$ hence

$$Zp = -H \cot\gamma \qquad \text{-------------- (4)}$$

.

Floor cancllation is shown experimenally in figure 7. Using this depth value calculated for each γ is used to prune out the pixels near the floors. Thus a newspaper lying on the floor wouldn't be considered as obstacle which otherwise would have been a problem in case of Disparity Map based algorithm

*A.2 Ceiling Pruning*

Similarly along the Y direction, hanging obstacles above certain height (Upper Clearance) could be avoided .

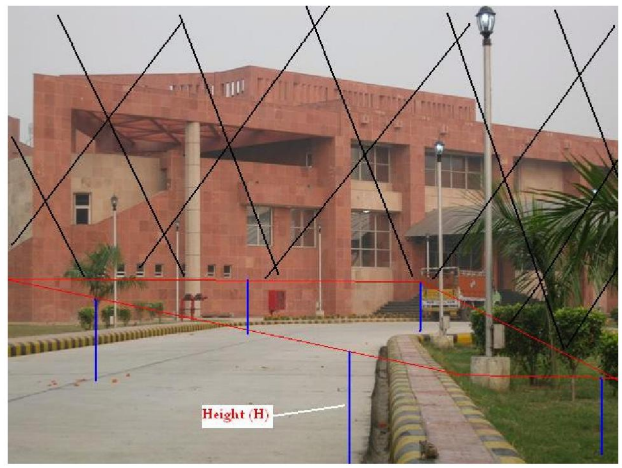
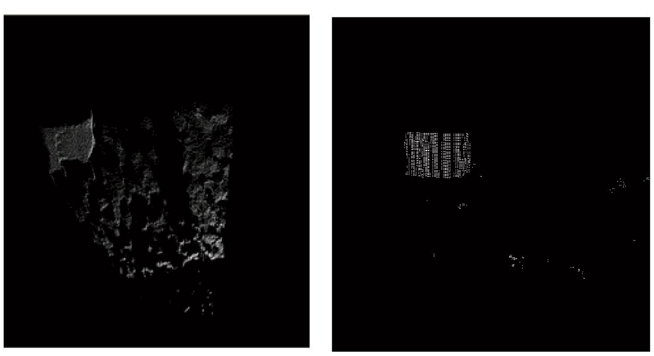



Figure 8 : Visual Imagination of Ceiling pruning

The volume above the height H (Clearance required for robot) can be avoided because if an obstacle is present in between above height H then its useless to be detected as robot will never hit it. This reduction in Height measure significantly reduces the computation required to generate 3D point Cloud. Refer Figure 8 and 9 for visualization.
Thus the computation involved in generating point cloud for trees on side roads , Clouds etc in a image can be avoided.

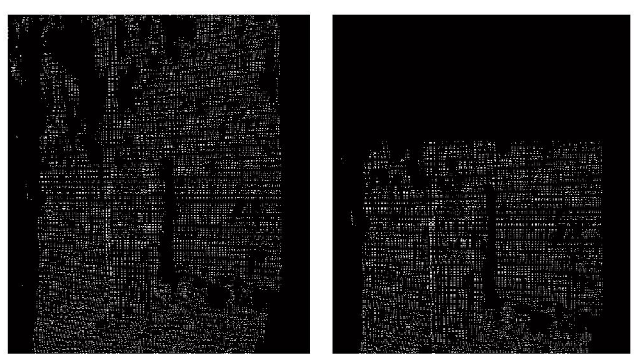
*A.3 Depth Pruning*

Keeping in mind faster computing efficiency apart from above steps it is also necessary to have a limit over field



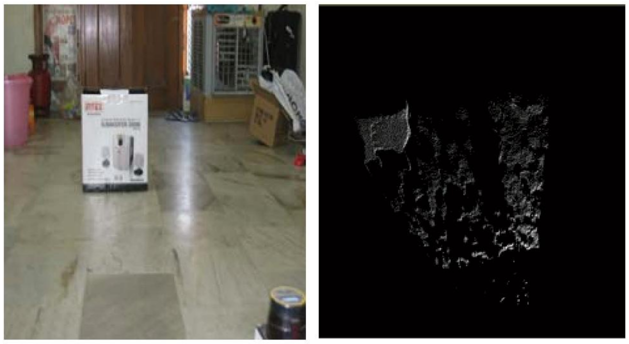
(a) (b)

Figure 7 (a) : point cloud data for a box on the floor
(b): After Floor cancellation



(a) (b)

Figure 9 (a) : Point cloud of a box against two walls
(b): Point cloud data after ceiling pruning



(a) (b)

Figure 10 (a) : Snap of room with a box on the floor (b)Point cloud data depth limited to surface of box

of view though we know that with increased depth the resolution decreases. Stereo Vision camera can look up to infinity but the depth data after a certain depth is corrupted with noise beyond our permissible level.
For e.g for outdoor purpose maximum speed we can allow our robot to move for experiment is around 7mps i.e around 25 KMPH. To apply brake even if we know obstacle 20m in advance i.e approx 3sec of time of contact, is sufficient.
Hence pruning the point cloud data after 20 m would be a wise decision to take.
While in indoor environment speed is much reduced than outdoor. So accordingly we can reduce the depth of view up to which computation is to be done.
Experimental result of depth pruning is shown in diagram 10.

*B. Raw Navigation Map Construction*

Once we have the computed the 3D point cloud information, building a Raw Navigation map is straightforward. Navigation Map is in XZ plane and I have set a space of interest of 20mX20m in front of the robot. Cell size of the map is 10cmX10cm. [Note that this can be changed upon the user's interest]

Grid size is therefore = 200X200.

A cell will contain the mean height value of all the 3D points contained in the volume above the cell. As discussed earlier a counter will be defined which basically tells us number of points contained in the particular volume i.e point density.

I have defined average point density to be as count value required for a cell to be considered for scoring in the next step. This average point density decreases as the depth of view decreases.

*C. Scoring of Map Cells*

Scoring process is defined to be process which updates the value of the cell of the Navigation Map. Initially all t


Figure 11 : Outdoor Unmanned Ground Vehicle (LAKHSYA)

the cell are assigned 'unknown'. Now if the count for a cell is greater than average point density then the particular cell is considered for scoring process.
If the mean height for a cell is greater than a minimum threshold height then the value of cell is updated to 'Obstacle (0)'.If the mean height for a cell is less than a minimum threshold height then the value of the cell is updated to 'unoccupied (1)'.

Each pixel of the navigation Map will now have a value from the set {0, 1, 2}.

*D.Obstacle Avoidance*

Navigation map is first converted to radial coordinates and fed to Vector Field Histogram [13] algorithm.

VFH finds out the density of obstacle at different angles in a sweep called as sectors. A candidate Matrix stores the value of all those sectors which have a minimum clearance for the robot to get through.

In case more than two paths are wide enough for the robot to get through then the one which is closest to its way point is selected.

## IV. Experimental Setup and results

*A Unmanned Ground Vehicle Platform*

"Lakshya"- is an outdoor unmanned ground vehicle (Figure 11) platform developed at the Innovation Lab, Delhi College of Engineering India under the Leadership of Mr. Saurav Kumar. The vehicle is driven using two Quicksilver Servo Motor. Bumblebee Stereo Vision Camera is installed for obstacle avoidance.

Indoor Platform (Figure 12) is again having the same electronics and sensor onboard as that of the Outdoor ground vehicle. An Ocean Server Heading sensor is also installed onboard which gives the vehicles heading direction w.r.t to North Pole.

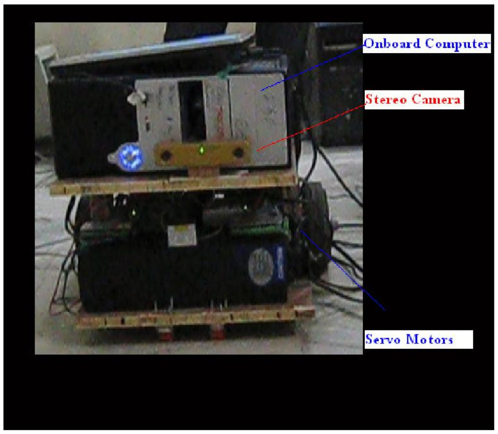



Figure 12: Indoor Platform.

Obstacle avoidance program is written in Microsoft Visual Studio 9.0 in C++ environment. It uses PGR library and Intel OpenCV Library for Navigation Map Building.

*B Results*

Different frames of the video showing the autonomous robot both moving in indoor environment is shown. Algorithm proved to very efficient and worked in real time.

**Indoor Environment**

Figure 13

Figure 14

Figure 15

Figure 16

Figure 17

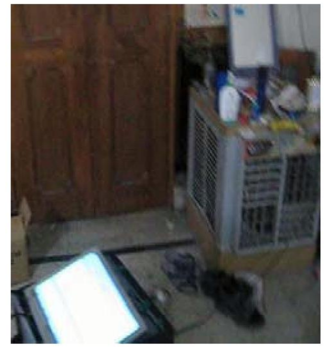


Figure 18

Figure 19

**Outdoor Environment**

Figure 20

Figure 21

Figure 22

Figure 23

Figure 24

Figure 25

Figure 26

## Acknowledgment

Innovation Lab located at Delhi College of Engineering was supported by DSIR, Govt of India and Sulabh International Social Service Organization. I would like express my acknowledgement to Puneet Kumar, Pallavi Awasthi and Gaurav Kumar .They helped me designing, and testing of the UGV platform

## References

[1] Klass Klasing, Dirk Wolherr , and Martin Buss “A clustering method for efficient segmentation of 3D laser Data ”2008IEEE International Conference on Robotics and Automation

[2] A.Broggie, C.Caraffi ,P.Paolo Porta, and P. Zani, “The Single Frame Stereo Vision System for Reliable Obstacle Detection used during the 2005 DARPA Grand Challenge on TerramMax”, in IEEE Intelligent Transportation System Conference, Toronto , Canada , September 17-20 , 2006 .

[3] R.Labayrade , D.Aubert, and j.p.. Tarel, “Real time obstacle detection on non flat road geometry trough V-disparity representation,'' in IEEE Intelligent Vehicle Symposium , Versailles , June 2002.

[4] A.Sappa , D.Geronimo, FDornaika, and A. Lipez , “Real Time Vehicle Pose Using On – Board Stereo Vision System '', Int. Conf. on Image Analysis and Recognition , LNCS, Springer Verlag ,Povoa de Varzim ,Portugal ,September 18 -20 2006.

[5] M.Cech , W. Niem, S. Abraham ,and C.Stiller , “Dynamic ego –pose estimation for driver assistance in urban environment ''. In IEEE Intelligent Vehicle Symposium , Parma , Italy,2004.

[6] S.Nendevschi , R.Dancescu, D. Frentiu, T.Martia , F.Oniga, C.Pocol, R.Schmidt , T.Garf, “High Accuracy Stereo Vision System for Far Distance Obstacle Detection,” IEEE Intelligent Vehicle Symposium ,Parma ,Italy,2004

[7] S.Nedevschi, R.Schmidt , T.Garf , R.Danscu, D.Frentiu , t. Martia , F.Oniga, C.Pocol, “3D Lane Detection System Based on Stereovision,” IEEE International Transportation System Conference (ITSC) , Washington ,USA,2004.

[8] Florin Oniga , Sergiu Nedevschi , Marc Michael Meinecke, and Thanh Binh To “Road Surface AND Obstacle Detection Based on Elevation Maps from Dense Stereo”, IEEE 2007 Conference

[9] Masko Kumano and Akihisa Ohya , “ Obstacle Avoidance of Autonomous Mobile Robot using Stereo Vision Sensor”, Shin'ichi Yuta Intelligent Robot Laboratory, University of Tsukuba, Ibaraki, Japan

[10] Don Murray and Jim Little Using real time stereo vision for mobile robot navigation,, Computer Science Dept. University of British Columbia.

[11] Richard Hartely and Andrew Zisserman, “Multiple View Geometry in Computer Vision”, Cambridge University press 2000

[12] Darius Burschka ,Stephen Lee ,”Stereo-Based Obstacle Avoidance in Indoor Environment with Active Sensor Re –Calibration “, Robotics Laboratory, Johns Hopkins University Baltimore