Journal Homepage: - www.journalijar.com
INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)
ISSN NO. 2320-5407
Article DOI: 10.21474/IJAR01/9765
DOI URL: http://dx.doi.org/10.21474/IJAR01/9765

## RESEARCH ARTICLE

# VISUAL ODOMETRY FOR AUTONOMOUS VEHICLES.

**Indushekhar Prasad Singh and Ashish Patel.**
`Research Assistant, Institute for System Research, University of Maryland, College Park, USA.

……………………………………………………………………………………………………....

**Manuscript Info**

……………………….

***Manuscript History***
Received: 12 July 2019
Final Accepted: 14 August 2019
Published: September 2019

***Key words :-***
Camera, Visual Estimation, GPS, Robotics, Visual odometry

**Abstract**

………………………………………………………………

With rapid advancements in the area of mobile robotics and industrial automation, a growing need has arisen towards accurate navigation and localization of moving objects. Camera based motion estimation is one such technique which is gaining huge popularity owing to its simplicity and use of limited resources in generating motion path. In this paper, an attempt is made to introduce this topic for covering different aspects of vision based motion estimation task.Visual odometry (VO) is the process of estimating the egomotion of an agent (e.g., vehicle, human, and robot) using only the input of a single or multiple camera attached to it. Application domains include robotics, wearable computing, augmented reality, and automotive. The advantage of VO with respect to wheel odometry is that VO is not affected by wheel slip in uneven terrain or other adverse conditions. It has been demonstrated that compared to wheel odometry, VO provides more accurate trajectory estimates, with relative position error ranging
from 0.1 to 2%. This capability makes VO an interesting supplement to wheel odometry and, additionally, to other navigation systems such as global positioning system (GPS), inertial measurement units (IMUs), and laser odometry (similar to VO, laser odometry estimates the egomotion of a vehicle by scan-matching of consecutive laser scans). In GPS-denied environments, such as underwater and aerial, VO has the utmost importance.



……………………………………………………………………………………………………....

## Introduction:-

With rising automation in different engineering fields, mobile robotics is gaining huge popularity. The unmanned vehicle is one such proliferating example that is expanding its fleet in different applications ranging from commercial to strategic use. One of the simplest mechanisms to estimate the motion of a terrestrial vehicle is to use wheel encoders. However, these have limited usage in ground vehicles and suffer from inaccuracies that occur due to wheel slippage during movement in muddy, slippery, sandy or loose terrains. The errors arising at each instant gets accumulated over time and the estimated pose drifts in proportion to the distance traveled [1]. Traditional navigation approaches such as inertial navigation systems (INS), the global positioning system (GPS), SONAR, RADAR, and LIDAR are currently in use for different applications [2]. Unavailability of GPS signals in an indoor and under-surface environment, unacceptable high drift using inertial sensors during extended GPS outages, issues of possible confusion with nearby robots for SONAR & RADAR, and the line of sight requirement for laser-based systems are some of the limitations associated with these navigation systems. One of the promising solutions lies in

**Corresponding Author:- Indushekhar Prasad Singh.**
Address:- Institute for Systems Research, University of Maryland, College Park, USA.

the art of visual odometry that helps in estimating motion information with the help of cameras mounted over the vehicle.

The onboard vision system tracks visual landmarks to estimate rotation and translation between two-time instants. The art of vision-based navigation is inspired by the behavior of a bird which relies heavily on its vision for guidance and control [3]. The initial works on estimating motion from a camera by Moravec has helped in establishing the current visual odometry (VO) pipeline [4]. Simultaneous localization and mapping (SLAM), a superset of VO, localizes and builds a map of its environment along with the trajectory of a moving object [5]. However, our discussion in this paper is limited to visual odometry, which incrementally estimates the camera pose and refines it using optimization technique. A visual odometry system consists of a specific camera arrangement, the software architecture and the hardware platform to yield camera pose at every time instant. The camera pose estimation can be either appearance or feature based. The appearancebased techniques operate on intensity values directly and matches template of sub-images over two frame or the optical flow values to estimate motion [6]. The feature-based techniques extract distinct interest points that can be tracked with the help of vectors describing the local region around the key-points. These techniques are dependent on the image texture and are generally not applicable in texture-less or low texture environments such as sandy soil, asphalt, etc. [7]. The VO technique can also be classified as geometric and learning based. The geometric VO techniques are the ones that explore camera geometry for estimating motion whereas the learning-based VO scheme trains regression model to estimate motion parameter when fed with labeled data [8-11]. The learning-based VO technique does not require the camera parameters to be known initially and can estimate trajectories with correct scale even for monocular cases [12]. The VO scheme can be implemented either with a monocular, stereo, or RGB-D camera depending on the system design. Stereo VO mimics the human vision system and can estimate the image scale immediately unlike monocular VO. However, stereo camera systems require more calibration effort and stringent camera synchronization without which the error propagates over time.

The monocular camera is preferred for inexpensive and small form factor applications such as phone, laptop, etc. where the mounting of two cameras with a specified baseline is not always feasible. Some of the approaches that aimed to recover scale information for monocular VO are the usage of IMU information [13-15], optimization approach during loop closure [16-17], and incorporating known dimensional information from walls, buildings, etc. [16-18]. An RGB-D camera provides color and depth information for each pixel in an image. The RGB-D VO starts with the 3D position of feature points which are then used to obtain transformation through iterative closest point algorithm [19-22]. The VO scheme has found its major application in the automobile industry in driver assistance and autonomous navigation [23-25]. One of the applications of visual odometry has been to estimate vehicle motion from the rearparking camera and use this information with GPS to provide accurate localization [26-28]. The task of visual servoing[29-30] (Moving the camera to a desired orientation) is very similar to the visual odometry problem requiring pose estimation for a different purpose [31-33]. These schemes are not only useful for navigation of rovers on surfaces of other planets such as Mars [34] but are also useful for tracking of satellites that needs to be repaired using a servicer [35]. Although these VO techniques have shown promising results for variety of these applications, they are sensitive to environmental changes such as lighting conditions, surrounding texture, the presence of water, snow, etc.

Some of the other conditions that lead to poor tracking data are motion blur, the presence of shadows, visual similarity, degenerate configuration, and occlusions. Along with these, some manmade errors also creep into the data during image acquisition and processing steps such as lens distortion and calibration, feature matching, triangulation, trajectory drift due to deadreckoning which lead to outliers. Therefore, the VO schemes need to be robust and have the ability to manage these issues efficiently. In order to handle the environmental conditions, different techniques have been proposed in the literature such as the usage of NIR cameras for dark environment [16-19] or usage of rank transform to handle lighting condition [20-23]. Kaess et al. handle data degeneration by dividing the image into two clusters based on disparity and computing rotation and translation with distant and nearby objects, respectively [36]. Several outlier rejection schemes have been proposed in the literature of which RANSAC and its different variants are very commonly used [37].

## Research Pipeline:-

### 1.1 Preprocessing:-

The input images are in Bayer format from which color images were recovered using the demosaic function with GBRG alignment. Then image frames were undistorted. For applications such as Visual Odometry, it is important to know the real-world location of points since the nonlinear nature of the lens distortion makes the problem challenging.

**1.2 Extraction of Camera Parameters:-**

Intrinsic matrix of the camera is calculated using the MATLAB function ReadCameraModel.m. The function gives the values of focal length ( $f_x$ and $f_y$ ) and Principal point offset ( $c_x$ and $c_y$ ).

**1.3 Feature detection and Correspondence:-**

SURF feature detector was used for detecting features. Then descriptor or feature vectors were extracted using the extractFeatures function of MATLAB computer vision toolbox. match features function was used to locate the point with matching features.

**1.4 Estimation of Fundamental matrix:-**

Estimating the fundamental matrix is the key step involved in the whole pipeline. The 8-point normalized algorithm was implemented to obtain the fundamental matrix. Using the matching points, a RANSAC function was implemented to randomly sample 8 set of matching points and checking the score of the resulting matrix with reference to other matching points. A fitness score was used to update the value of the fundamental matrix. The perpendicular distance of the matching point from the epipolar line is computed. If the distance lies within the threshold selected, we increment the score by 1. Thus, the fundamental matrix gets updated if its score is greater than the earlier score.[27-28]

**1.5 Estimation of Essential matrix:-**

Essential Matrix is computed after this by using the camera intrinsic and Fundamental matrix. It is made sure that the rank of the matrix is two by only keeping two singular values in the diagonal matrix of the svd. MATLAB code snippet is given below :

```
%% Calculating the essential matrix
E = K'* F * K;
[U,~,V] = svd(E);
E = U * [1 0 0;0 1 0;0 0 0] * V';
E = E / norm(E);
```

**1.6 Extracting camera pose:-**

For a given essential matrix E = U diag(1, 1, 0)V T , and first camera matrix P 1 = [I|0], there are four possible choices for the second camera matrix P2 namely: P 2 = [U W V T | + u3] or [U W V T | − u3] or [U W T V T | + u3] or [U W T V T | − u3]
where W is orthogonal matrix equal to [0 -1 0;1 0 0;0 0 1]. [9-11]

**1.7 Selection of correct pose:-**

Now we need to select one correct pose out of 4 poses derived from previous steps. I have used the constrained which is specific for this project to estimate the correct pose

1. Assuming that the car is moving forward only, we can filter out the transform which has positive z translation.
2. Rotation and Translation condition around y-axis: For the current scenario, rotationcan be assumed to be only around the y-axis. So that we can filter out the rotation matrix whose rotation component along other axes is approximately zero. While selecting the correct pose, the pose with minimum y translation is also one of the conditions utilized in computing the correct pose.

Finally, if none of the poses are able to satisfy all the conditions then pose of the world frame is selected as a correct pose.

**Camera pose update:-**

The last step before getting the final camera pose is updating the pose with respect to the previous frame. For updating the pose of the first frame we use the pose of the world frame and then subsequent frames use the pose of previous frames. A simple equation for that is given below:

$$R\_next = R\_current \times R\_next$$
$$t\_next = t\_current + R\_current \times t\_next$$

## Result and Discussion:-

The Plot of the position of the camera center (for each frame) based on the rotation and translation parameters between successive frames thus obtained is as shown in fig 1-fig 5 and the final plot is shown in fig 6:

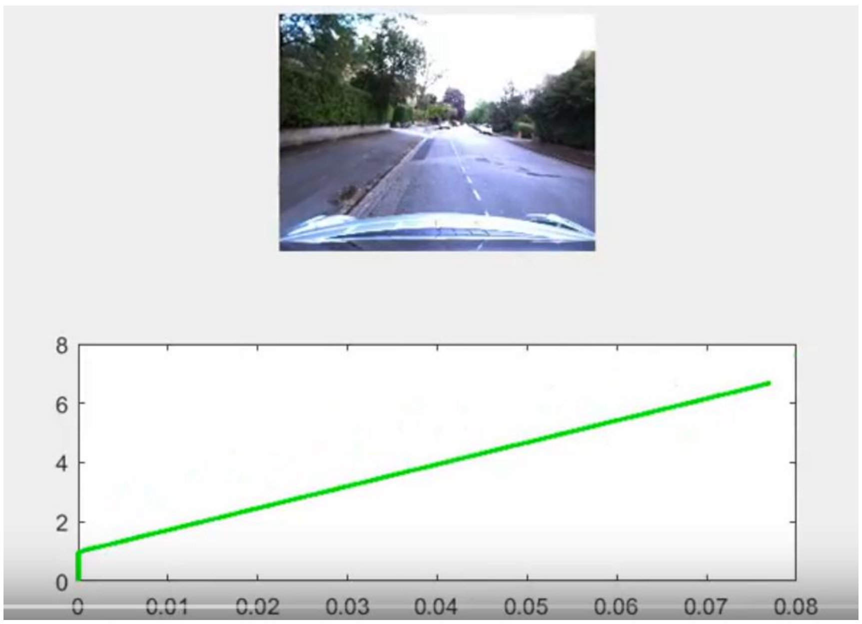


**Fig 1:-** Case 1 Plot of the camera position in x-z plane

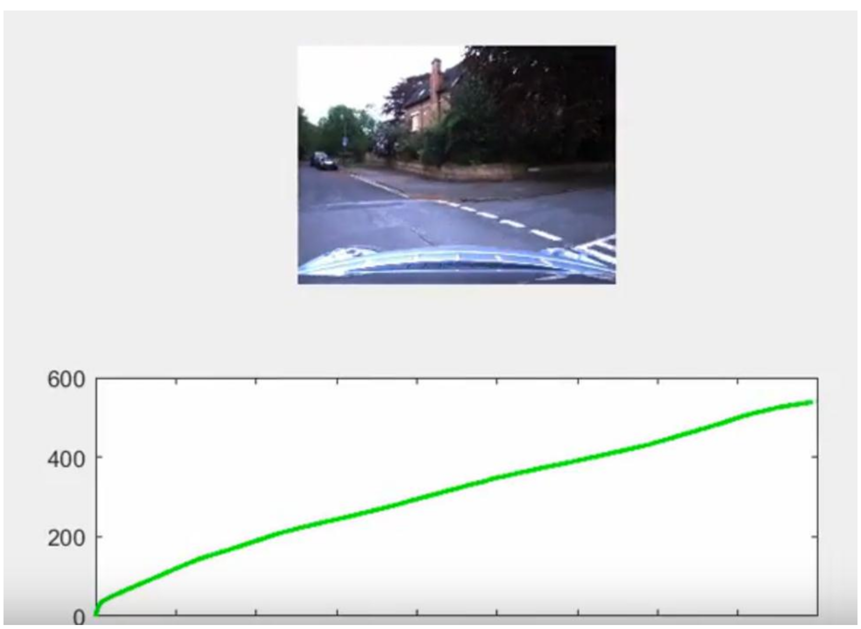


**Fig 2:-** Case 2 Plot of the camera position in x-z plane

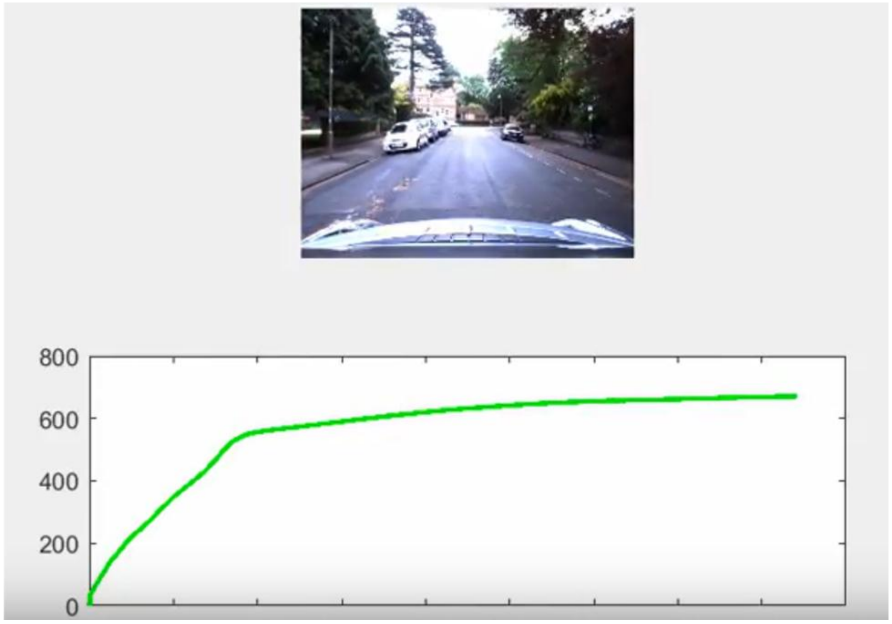


**Fig 3:-** Case 3 Plot of the camera position in x-z plane

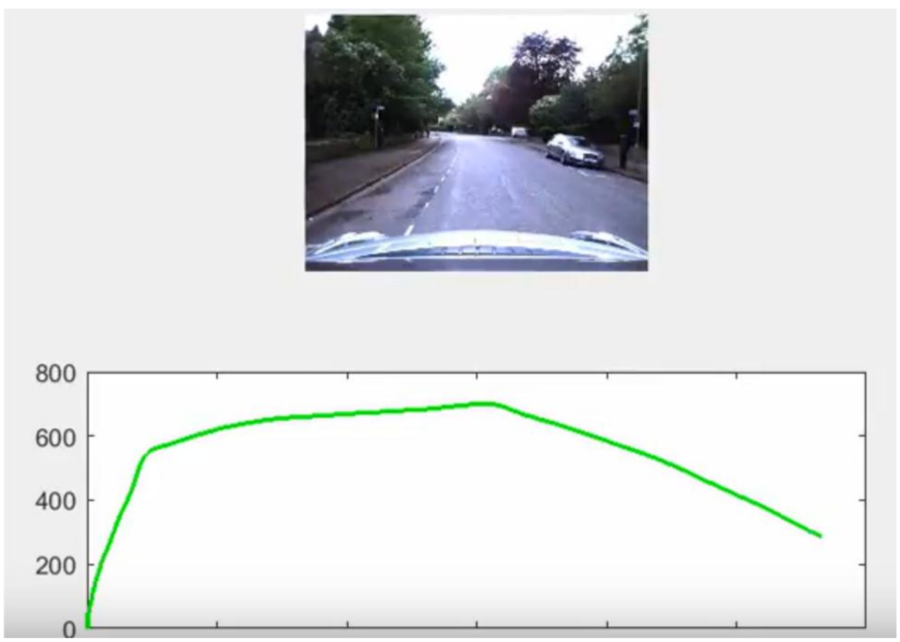


**Fig 4:-** Case 4 Plot of the camera position in x-z plane

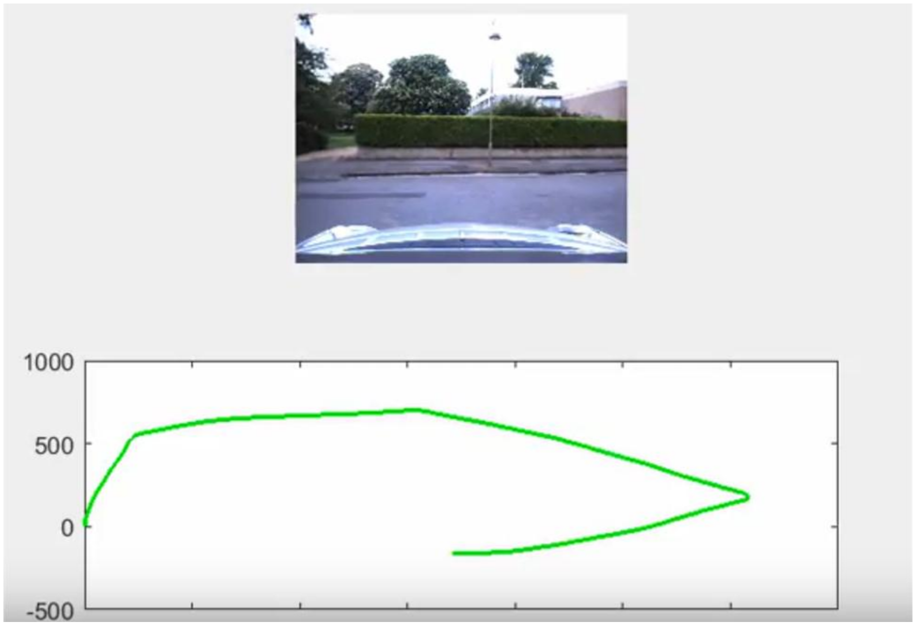

**Fig 5:-** Case 5 Plot of the camera position in x-z plane

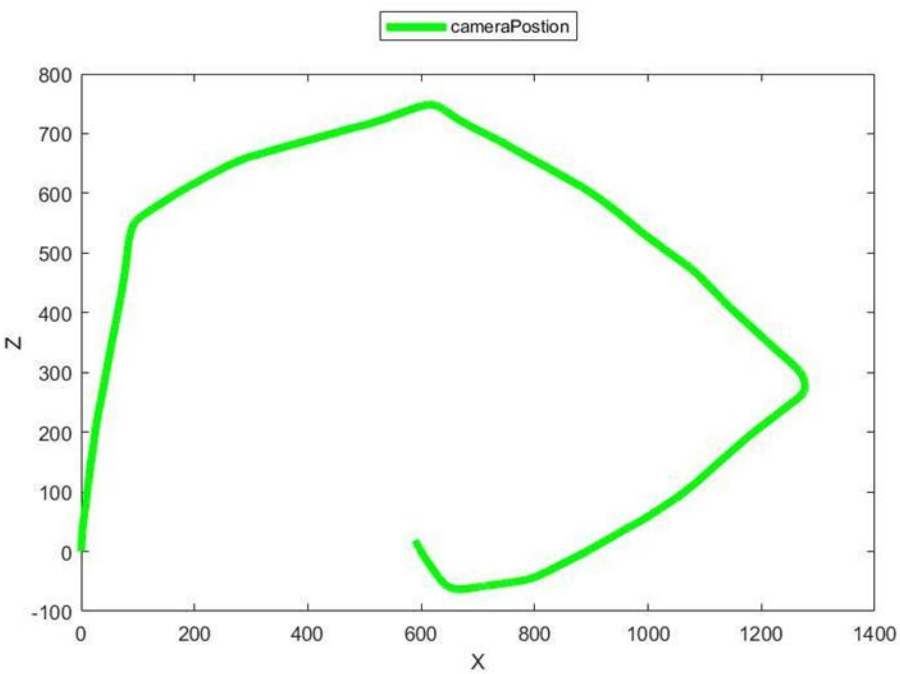

**Fig 6 :-** Final Plot of the camera position in x-z plane

The world-z-direction is forward and world-X direction is towards the right. The camera position is changing as per car movement and all the hard right turn can be easily seen in the plot, though drift is there in the value of z-axis.

**Comparison with output from MATLAB functions:-**

All the steps of the pipeline till feature detection is same in this case. Fundamental matrix is calculated using the function estimateFundamentalMatrix, which also gives index of inliers points. Only these inlier points are used to calculate the realtive pose using relativeCameraPose function. The pose obtained is updated as previous to give the final camera pose.

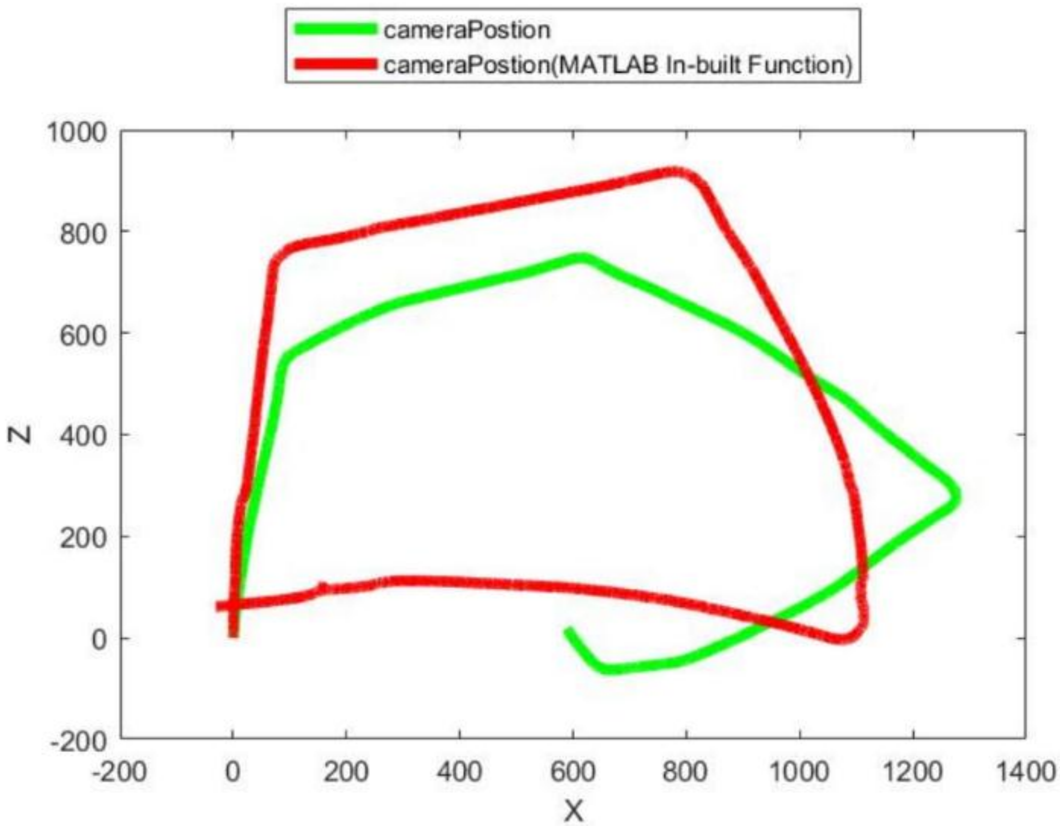


**Fig 7 :-** Final Plot of the camera position in x-z plane

The experimental video can be found here: https://www.youtube.com/watch?v=jjG6oBL6SnM&feature=youtu.be

## Conclusions:-

This research provide a holistic picture of the visual odometry technique encompassing different branches of this tree. The paper starts with an introduction to the motion estimation schemes and their wide applications in different engineering fields. Based on the experimental result on the real car. Significant improvement can been seen in the plot obtained by using MATLAB functions. Turns are sharp and movement along z axis is also greater. The car also reaches the point it started (as it is the ground truth).

## References:-


1. 1. M. O. A. Aqel, M. H. Marhaban, M. I. Saripan and N. B. Ismail, "Review of visual odometry: types, approaches, challenges, and applications," SpringerPlus, vol. 5, p. 1897, 2016.
2. R. Madison, G. Andrews, P. DeBitetto, S. Rasmussen and M. Bottkol, "Vision-aided navigation for small UAVs in GPSchallenged environments," in AIAA Infotech@ Aerospace 2007 Conference and Exhibit, 2007.
3. S. M. Ettinger, "Design and implementation of autonomous visionguided micro air vehicles," 2001.
4. H. Badino, A. Yamamoto and T. Kanade, "Visual odometry by multi-frame feature integration," in Computer Vision Workshops (ICCVW), 2013 IEEE International Conference on, 2013.
5. H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: part I," IEEE robotics \& automation magazine, vol. 13, pp. 99-110, 2006.
6. R. Gonzalez, F. Rodriguez, J. L. Guzman, C. Pradalier and R. Siegwart, "Combined visual odometry and visual compass for offroad mobile robots localization," Robotica, vol. 30, pp. 865-878, 2012.
7. D. Scaramuzza and R. Siegwart, "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles," IEEE transactions on robotics, vol. 24, pp. 1015-1026, 2008.
8. T. A. Ciarfuglia, G. Costante, P. Valigi and E. Ricci, "Evaluation of non-geometric methods for visual odometry," Robotics and Autonomous Systems, vol. 62, pp. 1717-1730, 2014.

9. Aswath S. et al. (2015) An Intelligent Rover Design Integrated with Humanoid Robot for Alien Planet Exploration. In: Kim JH., Yang W., Jo J., Sincak P., Myung H. (eds) Robot Intelligence Technology and Applications 3. Advances in Intelligent Systems and Computing, vol 345. Springer, Cham
10. Aswath Suresh, Gautam Ranjan,Sri Harsha, Kolluri Surya, Adarsh Ranjan, Nitin Ajithkumar, Sajin Sabu,Vinay Teja, Abel Varghese David, Ganesha Udupa, "Innovative Low Cost Mars Flyby Spacecraft for Safe Interplanetary Human Mission",Mars Society Convention, Marspapers,2016
11. Aswath Suresh, Sri harsha, Kolluri Surya, Debrup Laha,Dhruv Gaba,Shivam Bhardwaj,Vinay Teja,Siddhant Bhambri and Suvaansh Bhambri, "Innovative Human Mars Mission with Vertical Farming" ,Mars Society Convention,Marspapers,2017
12. V. Guizilini and F. Ramos, "Visual odometry learning for unmanned aerial vehicles," in Robotics and Automation (ICRA), 2011 IEEE International Conference on, 2011.
13. G. N{\"u}tzi, S. Weiss, D. Scaramuzza and R. Siegwart, "Fusion of IMU and vision for absolute scale estimation in monocular SLAM," Journal of intelligent \& robotic systems, vol. 61, pp. 287-299, 2011.
14. H. Strasdat, J. M. M. Montiel and A. J. Davison, "Scale drift-aware large scale monocular SLAM," Robotics: Science and Systems VI, vol. 2, 2010.
15. Aswath Suresh, Unnikrishnan VJ, Sreekuttan T Kalathil, Abin Simon, and Ganesha Udupa, " Design And Development Of An Intelligent Rover For Mars Exploration",Marspapers,2015
16. Suresh A. et al. (2017) An Advanced Spider-Like Rocker-Bogie Suspension System for Mars Exploration Rovers. In: Kim JH., Karray F., Jo J., Sincak P., Myung H. (eds) Robot Intelligence Technology and Applications 4. Advances in Intelligent Systems and Computing, vol 447. Springer, Cham
17. Aswath Suresh, Sri Harsha, Kolluri Surya, Debrup Laha, Dhruv Gaba, Siddhant Bhambri, Suvaansh Bhambri and Karthik Rangarajan, "Exploration-Probe to Jupiter Moon Europa", Mars Society Convention Marspapers,2016
18. S. Hilsenbeck, A. M{\"o}ller, R. Huitl, G. Schroth, M. Kranz and E. Steinbach, "Scale-preserving long-term visual odometry for indoor navigation," in Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on, 2012.
19. Aswath, S., Tilak, C.K., Sengar, A., Udupa, G.: Design and development of Mobile Operated Control System for Humanoid Robot. In: Advances in Computing, 3rd edn., vol. 3, pp. 50–56
20. Aswath, S., Tilak, C.K., Suresh, A., Udupa, G.: Human Gesture Recognition for Real-Time Control of Humanoid Robot. In: Proceedings of International Conference on Advances in Engineering and Technology, Singapore, March 29-30 (2014)
21. Krishnan, A.B., Aswath, S., Udupa, G.: Real Time Vision Based Humanoid Robotic Platform for Robo Soccer Competition. In: Proceedings of International Conference on Interdisciplinary Advances in Applied Computing, India, October 10-11 (2014)
22. I. Dryanovski, R. G. Valenti and J. Xiao, "Fast visual odometry and mapping from RGB-D data," in Robotics and Automation (ICRA), 2013 IEEE International Conference on, 2013.
23. M. Bertozzi, A. Broggi and A. Fascioli, "Vision-based intelligent vehicles: State of the art and perspectives," Robotics and Autonomous systems, vol. 32, pp. 1-16, 2000.
24. Suresh A., Sridhar C.P., Gaba D., Laha D., Bhambri S. (2019) Design and Development of Intelligent Self-driving Car Using ROS and Machine Vision Algorithm. In: Kim JH. et al. (eds) Robot Intelligence Technology and Applications 5. RiTA 2017. Advances in Intelligent Systems and Computing, vol 751. Springer, Cham
25. Suresh A., Laha D., Gaba D., Bhambri S. (2019) Design and Development of Innovative Pet Feeding Robot. In: Kim JH. et al. (eds) Robot Intelligence Technology and Applications 5. RiTA 2017. Advances in Intelligent Systems and Computing, vol 751. Springer, Cham
26. S. Lovegrove, A. J. Davison and J. Ibanez-Guzm{\'a}n, "Accurate visual odometry from a rear parking camera," in Intelligent Vehicles Symposium (IV), 2011 IEEE, 2011.
27. Bhardwaj S. et al. (2019) Isolated Bio-Regenerative System for Vertical Farming Through Robots in Space Explorations. In: Kim JH. et al. (eds) Robot Intelligence Technology and Applications 5. RiTA 2017. Advances in Intelligent Systems and Computing, vol 751. Springer, Cham
28. Suresh A., Gaba D., Bhambri S., Laha D. (2019) Intelligent Multi-fingered Dexterous Hand Using Virtual Reality (VR) and Robot Operating System (ROS). In: Kim JH. et al. (eds) Robot Intelligence Technology and Applications 5. RiTA 2017. Advances in Intelligent Systems and Computing, vol 751. Springer, Cham
29. M. Maimone, Y. Cheng and L. Matthies, "Two years of visual odometry on the mars exploration rovers," Journal of Field Robotics, vol. 24, pp. 169-186, 2007.
30. N. W. Oumer and G. Panin, "3D point tracking and pose estimation of a space object using stereo images," in Pattern Recognition (ICPR), 2012 21st International Conference on, 2012.

31. J. Ruppelt and G. F. Trommer, "Stereo-camera visual odometry for outdoor areas and in dark indoor environments," IEEE Aerospace and Electronic Systems Magazine, vol. 31, pp. 4-12, 2016.
32. Kumar V.S., Aswath S., Shashidhar T.S., Choudhary R.K. (2017) A Novel Design of a Full Length Prosthetic Robotic Arm for the Disabled. In: Kim JH., Karray F., Jo J., Sincak P., Myung H. (eds) Robot Intelligence Technology and Applications 4. Advances in Intelligent Systems and Computing, vol 447. Springer, Cham
33. A. I. Comport, E. Marchand, M. Pressigout and F. Chaumette, "Realtime markerless tracking for augmented reality: the virtual visual servoing framework," IEEE Transactions on visualization and computer graphics, vol. 12, pp. 615-628, 2006.
34. Suresh A., Arora C., Laha D., Gaba D., Bhambri S. (2019) Intelligent Smart Glass for Visually Impaired Using Deep Learning Machine Vision Techniques and Robot Operating System (ROS). In: Kim JH. et al. (eds) Robot Intelligence Technology and Applications 5. RiTA 2017. Advances in Intelligent Systems and Computing, vol 751. Springer, Cham
35. C. Golban, S. Istvan and S. Nedevschi, "Stereo based visual odometry in difficult traffic scenes," in Intelligent Vehicles Symposium (IV), 2012 IEEE, 2012.
36. M. Kaess, K. Ni and F. Dellaert, "Flow separation for fast and robust stereo odometry," in Robotics and Automation, 2009. ICRA'09. IEEE International Conference on, 2009.
37. R. Raguram, O. Chum, M. Pollefeys, J. Matas and J.-M. Frahm, "USAC: a universal framework for random sample consensus," IEEE transactions on pattern analysis and machine intelligence, vol. 35, pp. 2022-2038, 2013.