Stereo-Based Visual Navigation of Mobile Robots in Unknown Environments

H. Soltani, H. D. Taghirad and A.R. Norouzzadeh Ravari

Advanced Robotics and Automated Systems (ARAS),
Faculty of Electrical and Computer Engineering,
K.N. Toosi University of Technology
E-mail: taghirad@kntu.ac.ir

Abstract: *In this paper a stereo vision-based algorithm for mobile robots navigation and exploration in unknown outdoor environments is proposed. The algorithm is solely based on stereo images and implemented on a nonholonomic mobile robot. The first step for exploration in unknown environments is construction of the map of circumference in real-time. By getting disparity image from rectified stereo images and translating its data to 3D-space, point cloud model of environments is constructed. Then by projecting points to XZ plane and put local maps together based on visual odometry, global map of environment is constructed in real-time. A* algorithm is used for investigating optimal path and nonlinear back-stepping controller guides the robot to follow the identified path. Finally, the mobile robot explores for a desired object in an unknown environment through these steps. Experimental results verify the effectiveness of the proposed algorithm in real-time implementations.*

Keywords: Mobile robot, stereo vision, navigation, exploration, unknown outdoor environments.

1. Introduction

Today robots are used in various environments for different applications, from industrial fields, manu-facturing facilities to autonomous exploration of remote planets. Mobile robots have achieved great interest among scientific, industrial and military practitioners. The first step for autonomous mobile robot navigation is to find a suitable and safe path without colliding to obstacles. For this reason, different types of sensors have being used and in the last decades, visual navigation is gaining the researchers attention, since vision strategies are more simply implementable and provide a perception of environment in a single shot [1]. One of the important purpose in mobile robots research is exploration of unknown environments for map construction and particularly object detection, such as vital remarks, warning remarks etc, in an unknown environment. For this reason, an optimal path is needed to control the robot in the traversable areas without colliding with obstacles. This problem is considered in various researches and the proposed solutions are generally decomposed to two set of local and global techniques. By global technique we mean that a complete model of environment is available. These techniques are suitable in some pre-determined motion of the robot in a structured environment, and if the whole model of environment is not available, their usage is limited. Unlike global techniques, local techniques do not require access to any information of circumference from the past and robot should decide strategies from information getting from its sensors at the moment. In order to use different sensors information and to construct a map of the environment, different types of algorithms are proposed. When we use vision sensors like stereo camera also different types of algorithms are proposed for depth and disparity image computation [2]. For environmental map construction, there are various ideas proposed by researchers. Some of researches reconstruct complete 3D model of environment [3,4], whereas some others just use feature points to receive circumference perception [5,6]. used V-disparity technique are used by [7,8] to differentiate between obstacles and terrain, while in [9] this purpose is achieved by teaching circumference to robot, or fitting a plane to terrain points using RANSAC algorithm [10]. In [11] digital elevation map is constructed based on percentage of traversing points of environment. In order to determine the optimal path fuzzy techniques has been used in [12,13] , whereas some other types of path planning algorithms such as A*, D* and potential field is proposed in the literature [14]. Finally, in order to control the mobile robots motion on identified path, robot pose estimation is usually needed that it can be computed by visual odometry techniques [15] or by using existent hardware such as wheel encoders, GPS, and IMU.

In this paper, we propose an algorithm for visual navigation of mobile robots in unknown outdoor environments based on stereo vision and visual odometry. In the proposed technique, we don't access to information of environment from the past. So, we should construct

978-1-4673-1148-9/12/$31.00 ©2012 IEEE

model of circumference in real-time. After that, path planning and motion controller is used to guide the robot traverse from safe and optimal points and check for existence of desired object.

Stereo camera is the only sensing devices that we use through the whole process. This quality leads us depend only on vision information that is grabbed from environment. So that, desired goal of the robot is defined also based on vision features. This matter improves navigation process from simple pass through start to goal point to investigate desired object and exploration process. On the other hand, as the navigation process proposes in unknown outdoor environment we don't access to information of around from the past and global map of environment is updated dynamically by grabbing information of around in real-time.

2. System Overview

As shown in Fig. 1, hardware configuration that is used for system implementation is the Melon mobile robot that lies on tracked robots category. As indicated in this figure the only sensing devices attached to the Melon mobile robot is Bumblebee2 stereo camera from Point Grey Research Inc. The visual navigation system architecture proposed in this paper is presented in Fig. 2. The information provided by stereo camera is fed to the navigation module which first computes disparity image from stereo rectified images. Then disparity image is used to construct local maps. By putting local maps together based on visual odometry, global map is constructed in a real-time fashion. A* algorithm is used for finding optimal path and motion controller guides the robot to identified path by getting information about right and left robot's wheel velocities. If the robot has reached to the goal, it should be stopped, otherwise the algorithm is repeated.

Fig. 1: Melon Mobile Robot with Stereo Camera

3. Map Construction

3.1 Local Map of Circumference

Depth images are computed from rectified stereo images using semi-global matching and mutual information [16]. By accessing disparity image and considering equation (1), point cloud model of around can be computed.

$$Z = \frac{fb}{d}, X = \frac{Z\times(x-cu)}{f}, Y = \frac{Z\times(y-cv)}{f} \quad (1)$$

in which, f, b, d are focal length, base line and disparity respectively and cu, cv present image center point. Fig. 3 shows an example of point cloud models. As indicated in this figure, obstacles and around walls are identified exclusively on the point cloud model.

Building complete 3D model of environment may not be an effective way for online implementations, as we

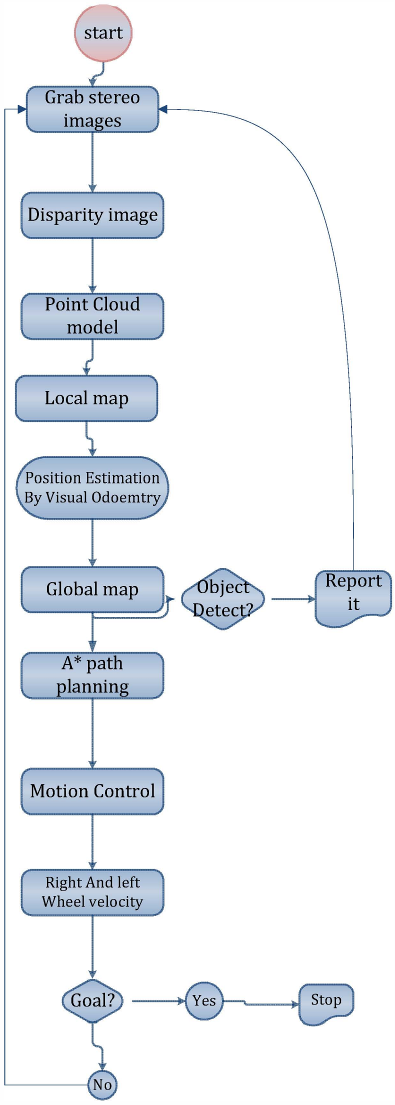

Fig. 2: The general architecture

(a)

(b)

Fig. 3: a – left frame of stereo camera, b- point cloud model

just need X and Z information of environment. Moreover, building elevation map of environment based on height and traverse criterion of points may not be implementable for Melon mobile robot, since it can traverse from low level hills in terrain. Thus, by projecting point cloud model to XZ plane, local map is constructed and environment is decomposed to traversable terrain and obstacles areas.

Depth estimation accuracy of stereo camera decreases incrementally by distance summation, therefore, we should propose a technique to get these information of environment in real-time. For this reason, we proposed using position estimation of mobile robot in order to put local maps together and construct the global map of environment.

3.2 Position Estimation

Estimation of the camera movement is an important problem in robotics and advanced driver assistance systems [17]. For such estimation and at each step, the relative orientation of the current frame with respect to the previous camera frame or a static reference frame is needed. This task is often performed using wheel speed sensors or inertial measurement units (IMUs) [18]. However, since camera systems are becoming cheaper, and more compact, and moreover, their computational power is ever increasing, the information given by images of such cameras suffices for precise motion estimation. This estimation is solely based on visual information [19], and is called visual odometry in literature [20]. Visual odometry have several advantages compared to that of other types of odometry techniques. First of all, it is more accurate compared to that of wheel speed sensors, especially in slippery terrains [21]. Other approaches that use GPS sensors or IMUs to mitigate this effect, also suffer from low accuracy and high costs.

In a first step, we extract and match corner-like image features between two consecutive stereo image pairs. For features detection, we can use different kinds of algorithms like Harris [22], SIFT [23] or SURF [24] descriptors, which are highly distinctive and thus allow robust matching. The image is divided into non-overlapping rectangles, which every of them keep a maximal number of feature points as shown in Fig. 4.

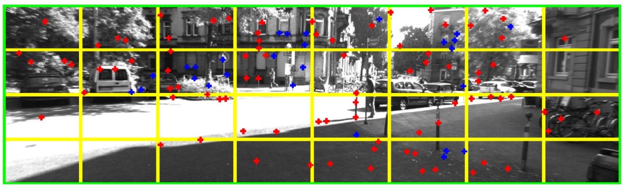

Fig. 4 : Bucketing Mechanism [17]

Based on these features matching, the movement of the robot is estimated by the trifocal tensor which relates features between three images of the same static scene. As it is shown in Fig.5, two trifocal tensor can be determined, which each of them relates the previous image pair to the current right frame or left frame by the following equations.

$$\mathcal{T}_R = \mathcal{T}(K_R, K_L, R_C, t_C, R_R, t_R, \Delta T) \quad (2)$$

$$\mathcal{T}_L = \mathcal{T}(K_R, K_L, R_C, t_C, R_L, t_L, \Delta T) \quad (3)$$

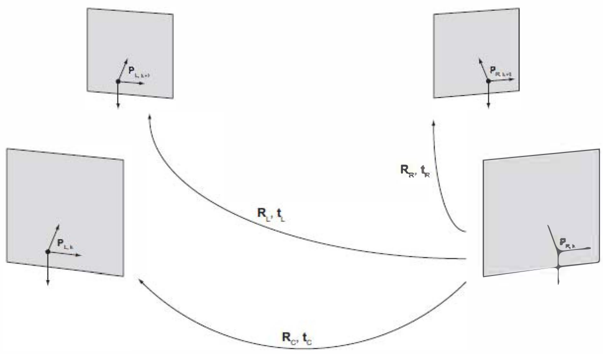

Fig. 5 : Geometric Relationship between two consecutive stereo image pairs [17]

Thus, using the trifocal tensors, a non-linear mapping of $x_{R,k} \leftrightarrow x_{L,k}$ into the current images via $x_{R,k+1} = h_R(\mathcal{T}_R, x_{R,k}, x_{L,k})$ and $x_{L,k+1} = h_L(\mathcal{T}_L, x_{R,k}, x_{L,k})$ is defined. Next, an iterated sigma point Kalman filter (ISPKF) is used to cope with the nonlinearities in the measurement equation. Outliers are detected by a random sample consensus (RANSAC) based outlier rejection scheme [25].

3.3 Global Mapping

Now, by accessing position of the camera in each moment, we can put local maps together and construct the global map. Thus, each point in local map can be transformed to global map via Equation (4).

$$\begin{bmatrix} X \\ Z \end{bmatrix} = \begin{bmatrix} cos\theta & -sin\theta \\ cos\theta & sin\theta \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} + \begin{bmatrix} t_x \\ t_z \end{bmatrix}$$
$$Y = y + t_y \quad (4)$$

in which, t_x, t_y, t_z denote movement of camera relative to reference frame and θ denotes changing rate of heading angle of robot respect to the reference frame. All these quantities are provided from visual odometry process. Fig. 6 shows an example of global map of environment. Solitary points and noise of model is removed to increase precision.

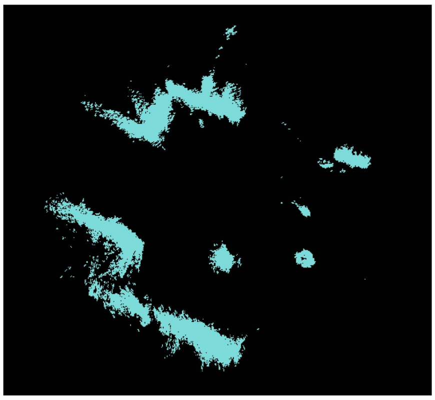

Fig. 6 : Global Map of Environment

4. Mobile Robot Navigation

4.1 Path Planning

After accessing the information of environment and building its map, we should find an optimal path for guiding the mobile robot toward its goal. For this purpose, A* algorithm is used which is developed to find an optimal path between any start and goal points [26]. Since the mobile robot shall not collide with obstacles approaching them, a distance transform algorithm is also included to keep specific distance from obstacles when the path is generated via A* algorithm. Its theory is based on distance from black color when it performs on grey scale images. In fact, whatever pixels nearer to black pixels their color is nearer to black. Therefore, by such technique a specific distance from obstacles is kept as it is shown in Fig. 7.

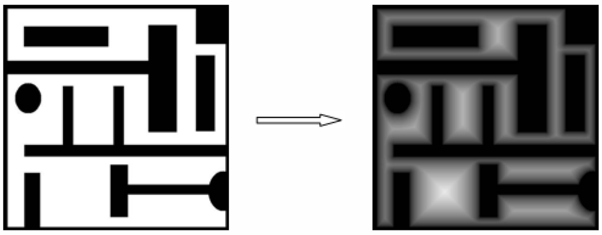

Fig. 7: Distance Transform Performance

4.2 Motion Control

Path following is achieved by non-linear back-stepping controller which commands a forward speed and a turn speed to the robot. Melon mobile robot is a non-holonomic system, since it should control two controllable output, right and left wheel's velocities, whereas position of the robot in the plane has three degrees of freedom, x, y and orientation of the robot (θ). Relations on mobile robots motion are define following:

$$\dot{x} = v\cos\theta$$
$$\dot{y} = v\sin\theta$$
$$\dot{\theta} = \omega \qquad (5)$$

In which, v, ω are linear and angular velocities, respectively. Since the objective is to follow the reference path, we should find linear and angular velocities that merge the errors to zero. Thus, error equation is defined as:

$$\begin{bmatrix} x_e \\ y_e \\ \theta_e \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_r - x \\ y_r - y \\ \theta_r - \theta \end{bmatrix} \qquad (6)$$

Moreover, error derivative equation is defined as Equation (7). Fig. 8 shows position of the mobile robot with respect to the reference pose and error posture.

$$\begin{bmatrix} \dot{x}_e \\ \dot{y}_e \\ \dot{\theta}_e \end{bmatrix} = \begin{bmatrix} y_e\omega - v + v_r\cos\theta_e \\ -x_e\omega + v_r\sin\theta_e \\ \omega_r - \omega \end{bmatrix} \qquad (7)$$

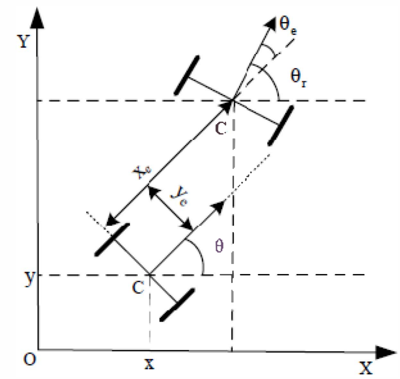

Fig. 8 : Error Posture

Finally, linear and angular velocities are selected via back-stepping control law [27,28].

$$v = k_1 x_e + v_r\cos\theta_e, \quad k_1 > 0$$
$$\omega = \omega_r + \frac{k_2 v_r y_e \sin\theta_e}{\theta_e} + k_3\theta_e, \quad k_2, k_3 > 0$$
$$k_1 = k_3 = 2\xi\sqrt{\omega_d^2(t) + \omega_n v_d^2(t)}$$
$$k_2 = \omega_n \qquad (8)$$

Equation (9) transforms linear and angular velocities to right and left wheel's velocities which are suitable to apply to the mobile robot.

$$\begin{bmatrix} \omega_r \\ \omega_l \end{bmatrix} = \frac{1}{r}\begin{bmatrix} 1 & \frac{b}{2} \\ 1 & -\frac{b}{2} \end{bmatrix}\begin{bmatrix} v \\ \omega \end{bmatrix} \qquad (9)$$

Moreover, B-Spline algorithm is proposed to the path to achieve better performance.

4.3 Object Detection

Exploration terminates by object detection process, and for this step, template matching technique is used. Our desired object has been presented to the mobile robot by only a sample image. The mobile robot will explore the desired object by its corresponding hue histogram with respect to other frames. Finally, by detecting object

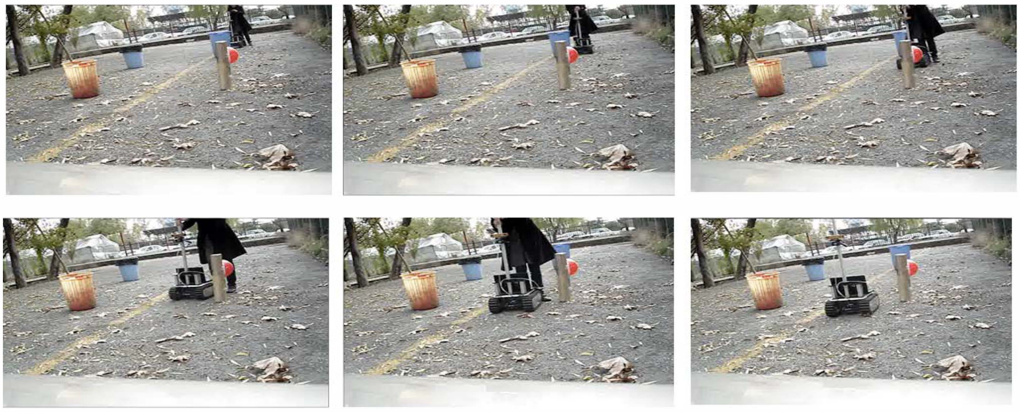

Fig. 9: Melon Mobile Robot Visual Navigation

on right and left frames location of the object is determined. Object detection techniques usually increase computational loading of the project and aren't sufficient for real-time algorithms while our technique is simple but effective approach for detecting objects in real-time.

5. Experimental Results

The navigation algorithm has been implemented and tested on Melon mobile robot which is equipped with a stereo camera as shown in Fig. 1. The algorithm has been implemented using a C++ program and OpenCV [29] and PCL [30] library. The mobile platform is exploring in an unknown outdoor environment and is continuously looking for a desired object. Therefore, stereo camera continuously grabs images from environment and disparity image is produced. Global map is constructed by putting local maps together based on visual odometry. Then, commanded wheels velocities are transmitted to mobile robot based on path planning and nonlinear motion controller. Meanwhile, the robot should check, if the desired object exists and reports it to the user. Fig. 9 shows how Melon mobile robot navigates in unknown outdoor environment. As indicated in this figure, the Melon robot traversed from safe and optimal areas without colliding with obstacles. At each moment, the robot grabs frames from stereo camera and processes them to construct disparity image (Fig. 10).

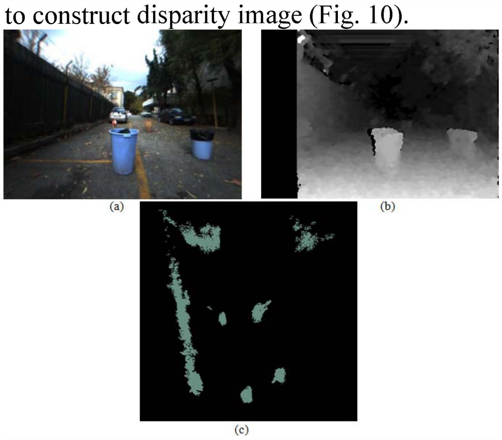

Fig. 10 : A- Left Image, B- Disparity Image, C- Global Map

Then, disparity map is used to build local map of 5 meters around the robot, as indicated in Fig. 11.In this figure, the optimal path is colored with blue that is generated by implementing A* algorithm, and distance transform algorithm, while green points indicate visual odometry results that show estimation of mobile robot's position in visual navigation process.

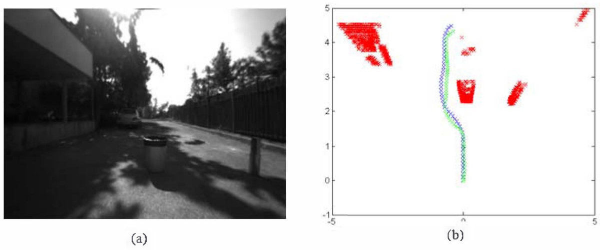

Fig. 11: A- Left Frame, B- Local Map with Path Planning and Visual Odometry Results

Visual odometry results approve that the mobile robot has well tracked the desired commanded path. Note that although shadows covered some areas of environment and half of the image becomes darker than the other half, as indicated in Fig. 11-A, the navigation algorithm performs robust and proper performance at this condition. Moreover, as the robot moves forward global map of environment is being constructed through putting local maps together and using estimated robot's pose (Fig. 10-C). At each step, the current pose of the robot is put together as start point of the motion, and frontier of visible and un-visible areas is set as final point of path planning algorithm, which is iteratively executed in an online manner. Finally, the robot is continuously looking forward to find the desired red ball as its final navigation goal (Fig. 12). As it is seen in this figure, the robot has terminated his navigation plan properly by finding its visual goal.

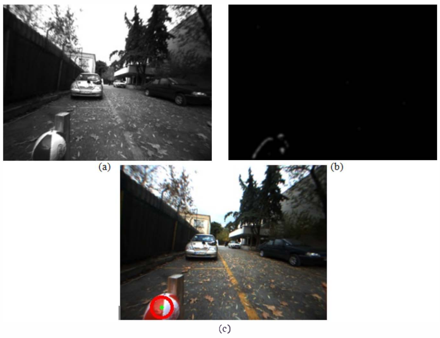

Fig. 12 : A- Right Frame of Stereo Image, B- Probability of Object Existence, C- Desired Object Detection

6. Conclusions

A stereo camera based navigation system is developed for unknown outdoor environments. The stereo camera is the only sensor and thus used for position estimation and environment modelling. The main idea is based on projecting of 3D points to XZ plane and constructing global map of environment via local maps and position estimation by visual odometry technique. Path planning and motion controller schemes are proposed to guide the mobile robot on an optimal path without colliding with obstacles. Meanwhile, the robot is continuously looking for a desired object whose image is given a priori to robot. Experimental results verify that this algorithm can suitably guide mobile robot in a real time fashion and in an unstructured outdoor environment to navigate toward traversable areas without colliding with obstacles.

References

[1] F. Bonin-Font, A. Ortiz, and G. Oliver, "Visual Navigation for Mobile Robots: A Survey," Intell Robot Syst, pp. 263-296, 2008.

[2] Aniket Murarka, M.S. Matthies, L., Shafer , “Building Safety Maps using Vision for Safe Local Mobile Robot Navigation” , S.A.: Error modeling in stereo navigation. IEEE J. Robot. Autom. 3(3), 239–248, (1987).

[3] A. Geiger, J. Ziegler and Ch. Stiller, “StereoScan: Dense 3d Reconstruction in Real-time”, Department of Measurement and Control Karlsruhe Institute of Technology.

[4] R. B. Rusu, “Leaving Flatland: Efficient Real-Time Three-Dimensional Perception and Motion Planning”, Journal of Field Robotics 26(10), 841–862 (2009) C 2009 Wiley Periodicals, Inc. Published online in Wiley InterScience (www.interscience.wiley.com) . DOI: 10.1002/rob.20313

[5] J. Zhou and B. Li, "Homography-based Ground Detection for A Mobile Robot Platform Using a Single Camera," in Jin Zhou and Baoxin Li, 2006.

[6] N. Ohnishi and A. Imiya, "Visual Navigation of Mobile Robot Using Optical Flow and Visual Potential Field," RobVis, pp. 412-426, 2008.

[7] Broggi, C. Caraffi, P. P. Porta, and P. Zani, “The single frame stereo vision system for reliable obstacle detection used during the 2005 DARPA grand challenge on TerraMax,” in Proc. IEEE Intell. Transp. Syst. Conf., Toronto, ON, Canada, Sep. 17–20, 2006, pp. 745–752.

[8] R. Labayrade and D. Aubert, “A single framework for vehicle roll, pitch, yaw estimation and obstacles detection by stereovision,” in Proc. IEEE Intell. Vehicles Symp., Columbus, OH, Jun. 2003, pp. 31–36.

[9] J. Sturm, K. Konolige, C. Stachniss, W. Burgard,” 3D Pose Estimation, Tracking and Model Learning of Articulated Objects from Dense Depth Video using Projected Texture Stereo”.

[10] K. Konolige, M. Agrawal, M. R. Blas, R. C. Bolles,” Mapping, Navigation, and Learning for Off-Road Traversal”, Journal of Field Robotics 26(1), 88–113 (2009) C _ 2008 Wiley Periodicals, Inc. Published online in Wiley InterScience (www.interscience.wiley.com).

[11] A. Chilian and H. Hirschm¨uller,” Stereo Camera Based Navigation of Mobile Robots on Rough Terrain”, The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems October 11-15, 2009 St. Louis, USA

[12] Howard A., Tunstel E., Edwards D., and C. A., "Enhancing fuzzy robot navigation systems by mimicking human visual perception of natural terrain traversability," in IFSA World Congress and 20th NAFIPS International Conference. vol. 1, 2001, pp. 7–12.

[13] M. Y. Shieh, Y. H. Chan, C. P. Cheng, and J. C. Hsieh, "Two-Stage Fuzzy Navigation Control of a Vision-Based Intelligent Shopping Service Robot," in The 30th Annual Conference of the EEE Industrial Electronics Society Busan, Korea, 2004.

[14] H. Choset, “Principles of Robot Motion—Theory, Algorithms, and Implementation”, The MIT Press Cambridge, Massachusetts London, England © 2005 Massachusetts Institute of Technology

[15] D. Nist´er ,O. Naroditsky ,J. Bergen,” Visual Odometry for Ground Vehicle Applications”, Sarnoff Corporation CN5300, Princeton NJ 08530 USA

[16] H. Hirschm¨uller ,”Stereo Processing by Semi-Global Matching and Mutual Information”, IEEE Transactions on Pattern Analysis and Machine Intelligence.

[17] B. Kitt, A. Geiger and H. Lategahn,” Visual Odometry based on Stereo Image Sequences with RANSAC-based Outlier Rejection Scheme”, Institute of Measurement and Control Systems Karlsruhe Institute of Technology

[18] Howard, “Real-time stereo visual odometry for autonomous ground vehicles,” in IEEE/RSJ International Conference on Intelligent Robots and Systems, September 2008, pp. 3946 – 3952.

[19] M. Agrawal and K. Konolige, “Real-time localization in outdoor environments using stereo vision and inexpensive gps,” in Proceedings of the 18th International Conference on Pattern Recognition, 2006, pp. 1063 – 10 068.

[20] D. Nist´er, O. Naroditsky, and J. Bergen, “Visual odometry,” in IEEE Computer Society Conference Computer Vision and Pattern Recognition,vol. 1, 2004, pp. 652 – 659.

[21] D. M. Helmick, Y. Cheng, D. S. Clouse, L. H. Matthies, and S. I. Roumeliotis, “Path following using visual odometry for a mars rover in high-slip environments,” in Proceedings of the IEEE Aerospace Conference, vol. 2, March 2004, pp. 772 – 789.

[22] C. Harris and M. Stephens, “A combined corner and edge detector,”in Proceedings of the 4th Alvey Vision Converence, 1988, pp. 147 –151.

[23] D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, vol. 60, no. 2, pp. 91– 110, 2004.

[24] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, “Speeded-up robust features (surf),” Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346 – 359, June 2008.

[25] D. Scaramuzza, F. Fraundorfer, and R. Siegwart, “Real-time monocular visual odometry for on-road vehicles with 1-point ransac,” in Proceedings of the IEEE Conference on Robotics and Automation, May 2009.

[26] S. J. Russell, P. Norvig, “Artificial Intelligence”. 2nd Edition, 2006.

[27] G. Oriolo, A. De Luca, M. Vendittelli, " WMR Control Via Dynamic Feedback Linearization: Design, Implementation, and Experimental Validation", IEEE Transactions on Control Systems Technology, Vol. 10, No. 6, November 2002.

[28] C. Samson, “Time-varying feedback stabilization of car-like wheeled mobile robots,” Int. J. Robot. Res., vol. 12, no. 1, pp. 55–64, 1993.

[29] Intel, "Open Source Computer Vision Library (OpenCV)," Intel.

[30] R. B. Rusu, S. Cousins, “3d is here : Point Cloud Library (PCL)”, IEEE International Conference on Robotics and Automation (ICRA),2011.