Stereo Vision SLAM Based Indoor Autonomous Mobile Robot Navigation

Conference Paper · December 2014

DOI: 10.1109/ROBIO.2014.7090560

4 authors:


Ebrahim Abdulla Mattar
University of Bahrain


Khalid Al-mutib
King Saud University


Mansour Alsulaiman
King Saud University


Ramdane Hedjar
King Saud University

Stereo Vision SLAM Based Indoor Autonomous Mobile Robot Navigation

Khalid N. Al-Mutib, Ebrahim A. Mattar, Mansour M. Alsulaiman, and H. Ramdane

Stereo Vision SLAM Based Indoor Autonomous Mobile Robot Navigation

Khalid N. Al-Mutib, Ebrahim A. Mattar, Mansour M. Alsulaiman, and H. Ramdane

Abstract— **This manuscript presents an autonomous navigation of a mobile robot using SLAM, while relying on an active stereo vision. We show a framework of low-level software coding which is necessary when the vision is used for multiple purposes such as obstacle discovery. The built system incorporated a number of SLAM based routines while replying on stereo vision mechanism. The system was implemented and tested on a mobile robot platform, and perform an experiment of autonomous navigation in an indoor environment.**

Keywords-component; Mobile robot, Stereo vision, Coding Hierarchy, SLAM.

I. INTRODUCTION

Stereo-vision based mobile robot navigation has always been a challenge. This is due to the complexity of integrating massive amount of information gathered by robot vision while it undergoes a motion. Visual perception of an environment, is also an important capability for mobile robots, and lots of efforts are being put by mobile robots research community. This is due to its consideration as a crucial part of mobile robot design.

There are considerable amount of efforts to let mobile robots to navigate visually within unknown, unstructured environments. In reference to Bonin-Font et. al. [1], they presented a substantial survey with a detailed summary of outstanding visual navigation studies from 1987 to late 1990s. Ghazouan et. al. [2] have also presented a new technique for model update using prior and current observations on the voxel state. Ghazouan et. al. additionally mentioned that, "Workspace is decomposed into voxels which are the smallest volume of environment. A first observation on the state of the voxels is calculated based on stereo system provided 3D points and triangulation error propagation. The proposed update function uses a credibility value that denotes how strongly a new observation shall influence the voxel state based on the age of the last observation and the homogeneity of the current observations". Guilherme and Avinash [3] have collectively published a manuscript that surveys some developments over the last 20 years within the field of vision-based mobile robot navigation. They stated that, "For indoor robots in structured environments, we have dealt separately with the cases of geometrical and topological models of space. For unstructured environments, we have discussed the cases of navigation using optical flows, using methods from the appearance-based paradigm, and by recognition of specific objects in the environment". Ulusoy [4] has presented an algorithm for active stereo vision, with depth perception for navigation. The algorithm was relying on SLAM routine for mobile navigation.

On the other hand, a work initiated and applied by Fernando et. al. [5], was related to sequential EKF feature-based SLAM algorithm. Fernando et. al. also stated that, "the features correspond to lines and corners-concave and convex- of the environment. From the SLAM architecture, a global metric map of the environment is derived. The electromyographic signals that command the robot's movements can be adapted to the patient's disabilities". The work of Elfes [6], which was associated to occupancy grid mapping, is the most widely used mobile robot mapping method. This is because of its simplicity, robustness, sufficiently flexible to be accommodated for numerous kinds of spatial sensors, in addition with ability to be adapted very well towards non-stationary environments. Stephen et. al., in their finding have provided a description of a vision-based mobile robot motion, they related that, "mobile robot localization and mapping algorithm, which uses scale-invariant image features as natural landmarks in unmodified environments. The invariance of these features to image translation, scaling and rotation makes them suitable landmarks for mobile robot localization and map building". Stephen et. al. [7], have also reported that, "to achieve SLAM, there are different types of sensor modalities such as sonar, laser range finders and vision". They added, "sonar is fast and cheap but usually very crude, whereas a laser scanning system is active, accurate but slow. Vision systems are passive and of high resolution".

A concept of on-line trajectory generation for robot motion control systems enabling instantaneous reactions to unforeseen sensor events was proposed by Kroger [8]. Kroger mentioned that "The paper presents the usage of time-variant motion constraints, such that low-level trajectory parameters can now abruptly be changed, and the system can react instantaneously within the same control cycle (typically one millisecond or less)". Further Kroger has added, "Such feature is important for instantaneous switching between state spaces and reference frames at sensor-dependent instants of time, and for the usage of the algorithm as a control sub-module in a hybrid switched robot motion control system". Further referring were also reported Murray and James [9], and the formulated in the Carnegie Mellon University by Martin and Moravec [10] as approaches to figure out paradigm for internal representation of stationary defined environment, with an evenly spaced

Khalid N. Al-Mutib is with the College of Computer & Information Sciences (CCIS), King Saud University, P. O. Box 51178, Kingdom of Saudi Arabia, (e-mail: muteb@ksu.edu.sa).

E. A. Mattar, is with the College of Engineering, University of Bahrain, P. O. Box 32038, Kingdom of Bahrain. (corresponding author e-mail: ebmattar@ieee.org).

M. M. Alsulaiman is the College of Computer & Information Sciences (CCIS), King Saud University, P. O. Box 51178, Kingdom of Saudi Arabia, (e-mail: msuliman@ksu.edu.sa).

H. Ramdane is with the College of Computer & Information Sciences (CCIS), King Saud University, P. O. Box 51178, Kingdom of Saudi Arabia, (e-mail: hedjar@ksu.edu.sa).

grids. That was based on ultrasonic range measurements. In a wider context, occupancy grids also afford data structure that permits combination of sensor data. They also provide a representation of the environment which is created with inputs from mobile sensory devices. Apart from being used directly for sensory data fusion, there also exist interesting variations of evidence grids, such a s place-centric grids as in Youngblood et. al [11]. Jia et. al. [12] , have also presented a methodology related to map buildings while relying on interactive GUI (Graphical User Interface) for an indoor service mobile robot. This is due to the sensors uncertainty. Here, the operator can modify the map compared with the real-time video from the web camera of the mobile robot. Furthermore, for improving the robot precision for self-localization, EKF was also used.

An important research work literature related to mobile vision-based navigation also to be mentioned, is the one presented by Elfes [13]. Elfes mentioned that "mobile motion were associated to the occupancy grid mapping, is the most widely used mobile robot mapping method. This is because of its simplicity, robustness, sufficiently flexible to be accommodated for numerous kinds of spatial sensors, with ability to be adapted very well towards non-stationary environments". Nevertheless, range finders capture very slight properties of the real environment where a mobile robot is to move. They cannot acquire wealthy visual information that leads to the ultimate goal of humanlike perception of the real surroundings. Range finders are also limited by their successful measurable ranges, which depend on category and shape of the range finder being used.

A. Contributed Work

The purpose of this study is to realize an autonomous mobile robot navigation while relying entirely on active vision. Our algorithm is called H-SLAM (Hybrid-SLAM). This is because we have enhanced the SLAM with some supporting routines. We have developed an active stereo vision system with a total of three degrees of freedom (pan, tilt, gaze) for each visual sensing. Blending SLAM with such active stereo vision is the main contribution of the article. The stereo visual sensing is described first. After that, framework of the implementation is proposed. Next, the structure of the navigation software is further explained. This includes the SLAM based stereo vision navigation. Finally, an indoor experimental verification results of the mobile robot is also illustrated. The implementation was achieved over a number of stages.

B. Paper Organization

The paper has been divided into six sections. In Section (I), we introduce related literature, background, and paper contribution . In Section (II), we present the overall mobile system architecture and related levels of hierarchy. Section (III) focuses on mobile robot vision and control hardware. Section (IV), provides strategy of navigation, where we indicate a number of experimentations and results, thus verifying the proposed navigation methodology. Finally, in Section (V), discusses results and conclusion remakes are concluded.

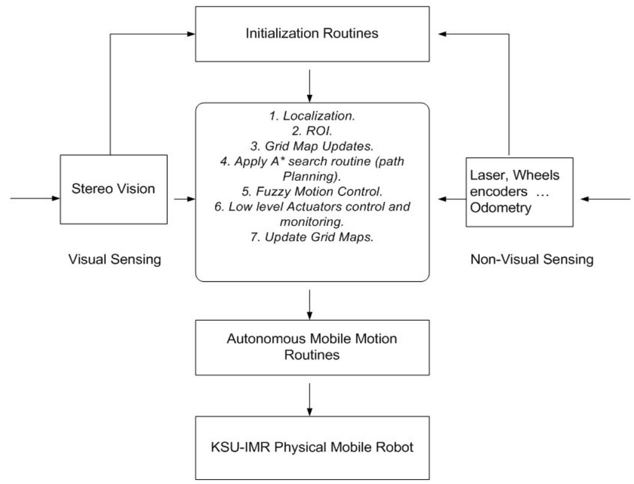

Figure 1. KSU-IMR system hierarchy.

II. MOBILE ROBOT HIERARCHY

A. Robot Hierarchy and Necessity

KSU-IMR stands for "King Saud University-Intelligent Mobile Robot". Such a mobile robot was built as a testing platform for stereo vision based navigation. Therefore, there are many tasks that are required by KSU-IMR mobile robot navigation while relying on active vision system. They are listed as follows: (*i*) Obstacle detection. (*ii*) Autonomous navigation. (*iii*) Floor landmark extraction. (*iv*) Gaining of new landmark information. In reality, the mobile robot is to perform such tasks sequentially, and in parallel once it is in motion. Therefore system hierarchy is needed to manage massive information during motion. We developed a hierarchy management for the active stereo vision system to use it commonly for multiple purposes.

B. KSU-IMR System Architecture

In practice there are many tasks which are required to the active vision system sensor. Examples of which are: (*i*) Obstacle detection. (*ii*) Landmark extraction. In reference to Fig. 1, KSU-IMR system architecture do consist of five fundamental blocks, (initialization, visual sensing, other sensing, navigation routines, and mobile body control).

C. System Hierarchy

The implementation approach is as follows: to design a mobile robotic system with stereo vision navigation capabilities. This is based on using visual data gathered from an environment. The mobile robot navigation design is to be based on a fundamental scheme; that high performance vision requires considerable computing strength. In order to fulfill this requirement, the proposed mobile robot was assembled to accommodate high performance computing power. An additional important tasks is the use of optimal path search based methods. Such a phase is achieved by creating navigation intelligence behaviors while robot is in motion. This will be based on intelligent techniques i.e. soft-computing techniques.

III. Stereo Vision And Motion System

A. Active Stereo Vision

Two cameras are fixed on a platform whose (*pan, tilt, and gaze*) are movable manually or from the computing hardware. A number of servo motors are employed for the camera head, for gaze, pan, and tilt control. The pan range of movement for each of upward and downward tilt movement ranges, is up to (±50°). Top level coding and multiple computers are used for performing CPU rigorous tasks in real time using a distributed computing architecture. This is to be capable of dealing with large number of sensing and control signals, as indicated to in Fig. 1.

B. Mobile Robot Control Hardware

This part for main computing power of POWERROB robot boards. They are used for distributing computing actions to cope with challenges from the real-time high computing power needs. Such a task also involves stereo vision and visual data processing. Other uses are, the master controller that handles all optimal path planning, controlling drive motors, and servo motors for moving camera platform. We have employed high current dual channel DC motor controller, this will give high movement toques.

C. Navigation Strategy

The in-depth structure of navigation software is shown in Fig. 2. Navigation software consists of "mobile robot navigation", and "self-localization". In reference to Fig. 2, we also show the overall coding hierarchy. It is an illustrations of the five fundamental stages of the stereo-vision navigation, including localization, ROI, grid mapping optimal path search via, and path planning.

D. Coding Interconnected Units.

Navigation routines: The main tasks performed by the onboard mobile computers are 3D depth measurement of the environment, scene analysis, optimal path search, real-time mobile path planning, and motor speed control. The mobile robot was designed in such a way to be capable of carrying high computing boards and batteries.

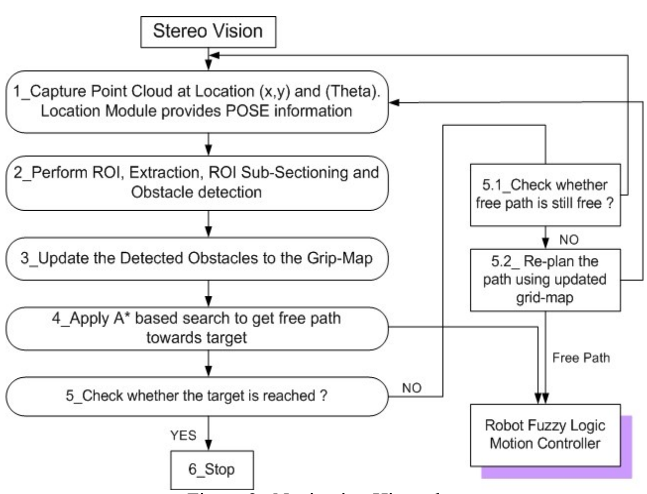

Figure 2. Navigation Hierarchy.

Another unique characteristics to be acquainted by the mobile robot, is the visual 3D perception that uses both depth from focus/defocus and the 3D binocular stereo. Building intelligence for mobile robot path planning. This is achieved by creating navigation intelligence capabilities while the robot is in motion. This is further based on intelligent path planning techniques (soft-computing methods). The four important units are (localization, map building and updating, searching optimal path planning, and motion control routine), as in Fig. 2.

E. Navigation Routines.

KSU-IMR localization at particular current time-step, the intension is the estimating robot state. This is achieved as based on knowledge of mobile initial state in addition to initial measurements up to current time. This is expressed in terms of mobile posture, the position and orientation. By definition this is a three-dimensional state vector. As stated earlier, we are integrating the vision localization techniques, hence we are laying down a heavy integration of these techniques into the developed SLAM routine.

IV. Implemenation and Results

A. Experimental Setup

The navigation strategy mentioned in previous section is implemented on an autonomous mobile robot "KSUMR", and verified by the following navigation experiment in a real environment.

B. Experimental Results

Localization Phase. As stated earlier, we are integrating the vision localization techniques, hence we are laying down a heavy integration of these techniques into the developed SLAM routine. A number of initially conducted experiments relied heavily on DEAD RECKONING localization. Errors in location estimation using DEAD RECKONING localization are accumulated in proportional to distances traveled by the robot and the ODOMETERY/GYRO inaccuracies. A number of initial experimentation have indicated that, mobile posture uncertainty increased significantly at each time step.

Monte Carlo Localization: It was chosen to implement a stereo vision Monte Carlo Localization (MCL). It will serve as a primary layer for localization parameters estimation. To enhance the robustness to localization technique, we switch to Dead Reckoning localization whenever the MCL layer failed to localize the robot due to sensor noise or unpredictable robot motion. Monte Carlo Localization was chosen as primary "localization" method. This is because it is superior in terms of its computational cost, moreover it supports multi-modal location distribution. In other words, these distributions are in fact possible locations for a robot, estimated by a motion model. Multi-modal distribution support gives us an ability to take into account multiple motion scenarios (e.g. either the robot is stationary, turning or moving straight). To validate the presented concepts, earlier presented Monte Carlo localization technique has been therefore widely tested in the laboratory ground office environment using diverse mobile robotic postures. Over a set of repeated experimentations, experimentations efforts have resulted and indicated that, such a technique is both efficient and robust. It was running comfortably in real-time. Verifying Monte Carlo localization, even under more challenging situations, the experiments described here are based on data recorded from the POWERROB mobile hardware. For localization, even though the sensory information were sensed and recorded at prior and previously defined time index, the time-stamps in the logs were used to recreate the real-time DataStream coming from the sensors. Hence, the obtained computational results are not conflicting with real data measured and collected from the mobile robot sensory instrumentation readings.

As part of the development of MCL layer, we customized the standard stereovision based Iterative Closest Point (ICP) method to suit our environment and sensor configuration. The resultant algorithm is able to localize the mobile platform with an upper bound of with respect to translational localization. Upper bound for rotational motion was stood at. Execution cost for this algorithm turns out to be very high and the robot translational speed has to be restricted at 50 (*mm/sec*). The robot's angular speed is restricted at a much slower. Reasons for such upper bounds is the high computational load on the on-board PC. The flow chart for such employed (ICP) algorithm was already presented and detailed in Fig. 3.

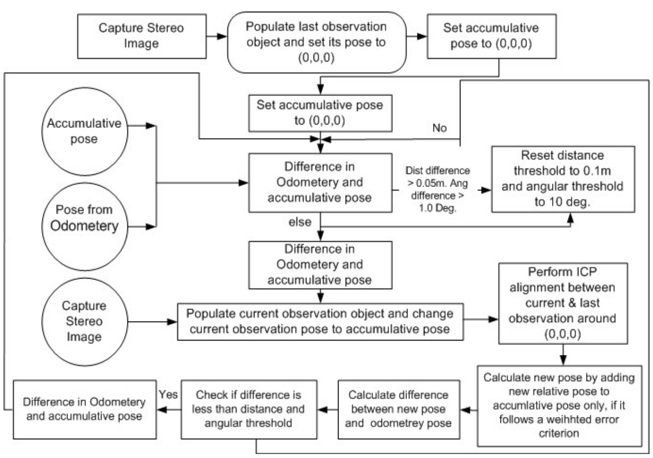

Figure 3. A flowchart of the complicated implemented ICP.

C. Localization Sensor Fusion

An effort was made into accumulating localization results from multiple vision sensors "Microsoft Kinect" and "Point Grey Bumblebee" for greater accuracy. Localization results deduced by experiments conducted on Microsoft Kinect and Point Grey Bumblebee camera suggested that Microsoft Kinect sensor has more inherent noise in depth value readings as compared to Bumble camera depth readings.

The research platform and experimental setup used in this experiment is shown in Fig. 4. A grid-based mapping approach was selected for the implementation of SLAM algorithm. Among the primary concerns about mapping, the implementation ensured that real-time execution, map-accuracy via modeling un-certainty and loop-closure detection is well-implemented within the mapping module.

In this respect, both of the results indicated to in Fig. 5 show positional error for mobile robot during localization process using each of the named sensors. Since Kinect sensor never performed better in any of the scenarios, the idea of combining localization results of both sensors was discarded. Accurate navigation and path-planning results were achieved. Resolution however, was chosen to be the consistent resolution upon which all future experiments would be conducted as it rendered the minimum cost in terms of execution time and navigation accuracy.

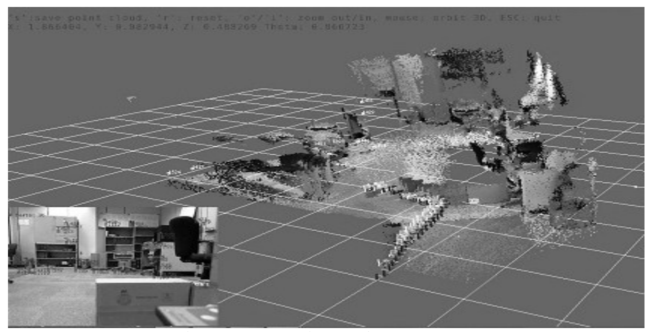

Figure 4. Experimental system setup.
Results for voxels states representation using three stereo pairs taken from different positions of the stereo cameras.

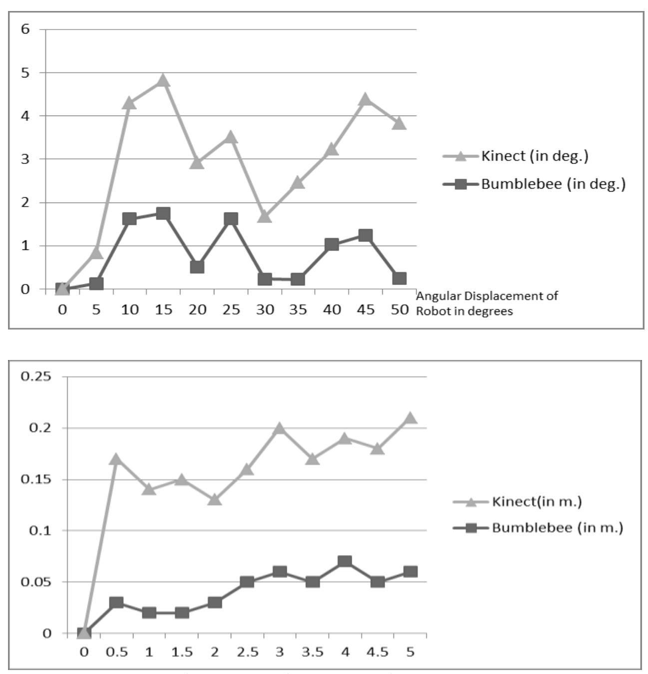

Figure 5. Robot pose and error, [14].
(*Top*) Mobile robot pose error variation in a turning scenario.
(*Bottom*) Mobile robot pose error variation in a translation scenario.

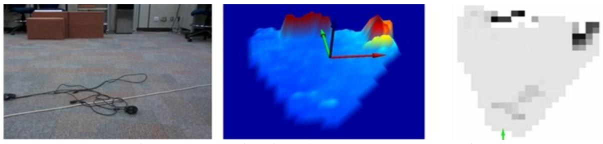

Figure 6. Path-planning over maps results.

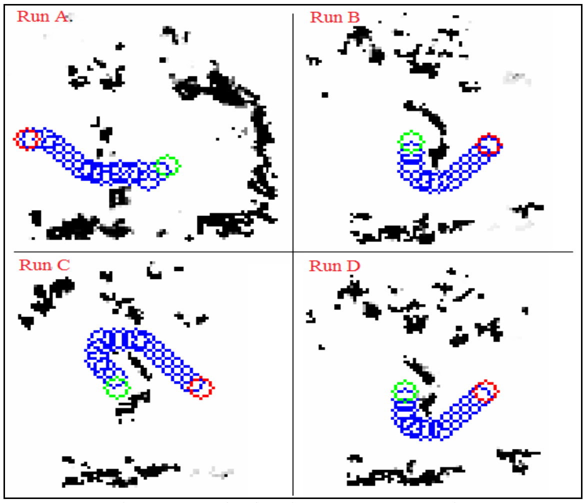

Figure 7. Path-planning over maps results.

Until now there have been no experiments designed, that require a larger area than the mentioned figure. The mapping updates and initializations are restricted to the area under observation, which roughly equals, thus any increase in map area will not trigger map-wide initialization or update. For map accuracy in terms of detection of small-sized obstacles and gradient floors, we have developed a customized algorithm and verify it over a number of experimentation trials.

The implemented mapping module can detect a loop and propagate errors in previous poses. The grid-map cells store both the height data about the environment and also the certainty measure about the existence of an obstacle. The certainty is a function of two elements i.e. the amount of time an obstacle is observed within a cell and secondly the number of points returned by the sensor that lie within each cell.

An example of a confidence map generated of obstacles in the environment (including loop detection routine) is represented in Fig. 6. Darker grid-cells represent a high probability of presence of an obstacle. Lighter shades represent an opposite hypothesis. Each observation is added to the map as a Bayesian update.

D. Map Accumulation Algorithm

The criteria for adding an observation to the built map can be defined as follows: A Bayesian update will either increase or decrease the probability of presence of an obstacle within a cell. This probability depends upon following factors: (*i*) Presence of points lying in the cell within the current observation. (*ii*) Consecutive number of observations for which a minimum number of obstacle points can be associated to a grid-cell. This mechanism handles fast moving dynamic obstacles. (*a*) All cells occluded by obstacles are not updated. For this purpose only the FOV in front of camera is considered for map updates. (*b*) The obstacle height data is only used for loop-closure detection. For obstacle detection, a combination of obstacle height and their persistence over multiple observations is employed.

E. Fast SLAM

Map-building without any knowledge of robot localization without an a priori map is unachievable. Sensor measurements need to be accurately transformed in order to be embedded within a map, for this purpose, the robot motion pose estimation needs to be accurate. The current implementation is a variant of fast SLAM. Here we use a particle filter based distribution model to update robot states and obstacle information. We use stereovision sensor to for map building and obstacle avoidance for obstacles within camera FOV. Stereovision based map is more comprehensive though not as accurate as a laser scan. This is the same reason that our navigation algorithm is far more robust to complex obstacles, such as obstacles having irregular foot print in all 3D-axes. Data association is proving to be challenge in our version of Fast SLAM. This is due to variations in illumination, specular reflection in environment and inconsistent point clouds due to variation in viewing angles. We use median and average filter along with Bayesian filters to remove noise from specular reflection.

Another layer based navigation approach was tested on the sidelines of this experimental bed. This approach presents a motion control for the autonomous robot navigation using fuzzy logic control. This requires the capability to maneuver in unstructured, dynamic and complex environment. The mobile robot uses intuitive fuzzy rules and is expected to reach a specific target or following a pre-specified trajectory while moving among unforeseen obstacles with the help of stereo-vision camera. The robot's execution depends on the choice of the task. In this approach, behavioral-based control architecture is adopted, and each local navigational task is analyzed in terms of primitive behaviors.

F. Fast SLAM Path-Planning

In reality, quite large of trails were successfully conducted experimentally within complex indoor obstacle scenarios for path-planning using the developed version of the Fast SLAM. The mobile robot has successfully reached its target location using the planned paths in an autonomous approach. Furthermore, currently we submit a set of goals to our system, hence the system plans path in sections for each of the goal. Once an obstacle scenario is significantly changed so much, in such a way that it affects the planned path, algorithm is used to re-plan the path to the goal. The planning and re-planning delays are less than a time of a

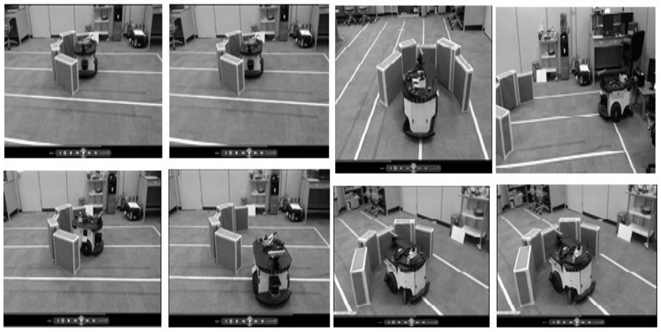

Figure 8. Video recording and shots, verifying mobile robot navigation.

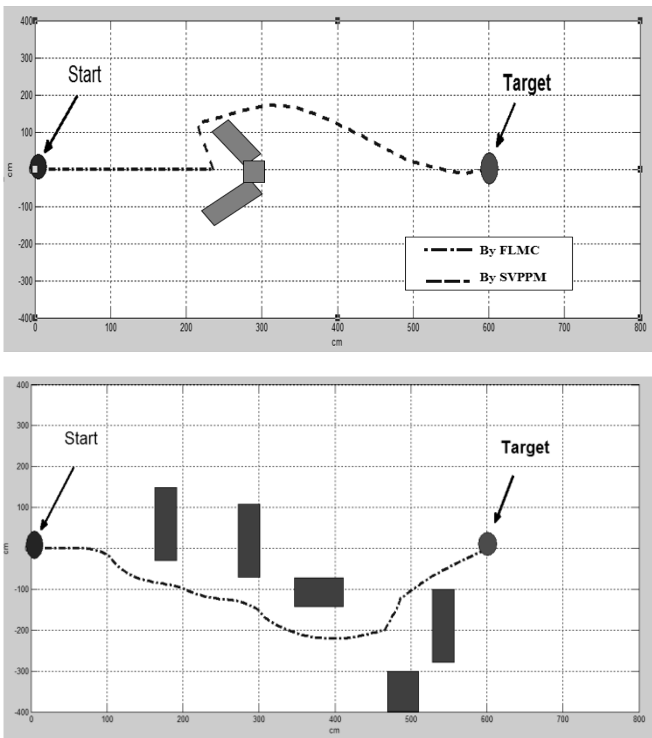

Figure 9. Graphical (*x,y*) data recording and navigation verifications. (*Top*) To target maneuvering, and motion to a posture path planning. (*Bottom*) Complicated to target maneuvering robot path planning, [14].

second long, so there exist no issues for performance degradation within path-planning module. Specifically, some of the path planning examples and runs, can be seen in Fig.7. Finally both Fig. 8 and Fig. 9 show a demonstration part of intensive experimentation and verifying the real-time mobile navigation outcomes. In Fig. 8, we show real video recording and shots, verifying mobile robot navigation without collisions. The mobile body was navigating without collisions. In addition, Fig. 9 shows the graphical (*x,y*) pose data recording and navigation verifications. For the top portion, we show a simple to target maneuvering, and motion to a posture path planning. For bottom portion, it shows mobile displacement even for complicated to target maneuvering, and robot path planning.

V. CONCLUSIONS

In this paper, we have developed the navigation strategy as based on integration of a number of important navigation routines. Hence, we also propose a framework of the vision system with the software level, which mediates the plural sensing requests and manages the vision recourses. The mobile body was experimentally tested for full navigation within an unstructured dynamic environment. A real-time stereo-vision SLAM technique was employed for motion purposes. The mobile robot has shown an excellent model of integration amount different layers and unit.

ACKNOWLEDGMENT

The research is financed by: NPST program, King Saud University. No. (08-ELE200-02). Kingdom of Saudi Arabia.

REFERENCES

[1] Bonin-Font F., Ortiz A., and Oliver G., "Visual Navigation for Mobile Robots: A Survey," Journal of Intelligent Robotics System, 53, 263–296, 2008.

[2] Ghazouan H., Tagina M., and Zapata R., "Robot Navigation Map Building Using Stereo Vision Based 3D Occupancy Grid," Journal of Artificial Intelligence: Theory and Application, 1,3, 63-72, 2010.

[3] Guilherme D. and Avinash K., "Vision for Mobile Robot Navigation: A Survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, 24, 2, 237-266, 2002.

[4] Ulusoy L., "Active Stereo Vision: Depth Perception for Navigation, Environmental MAP Formation and Object Recognition," The Middle East Technical University, Turkey, 2003.

[5] Fernando C., Natalia L., Carlos S., Fernando S., Fernando P., and Ricardo C., "SLAM algorithm applied to robotics assistance for navigation in unknown environments," Journal of Neuro Engineering and Rehabilitation, 7,10 : 1-16, 2010.

[6] Elfes A., "Using occupancy grids for mobile robot perception and navigation," IEEE Tran. of Computer Magazine, 22, 6, 46–57, 1989.

[7] Stephen S., Lowe D., and Little J., "Mobile Robot Localization and Mapping with Uncertainty using Scale-Invariant Visual Landmarks," International Journal of Robotics Research, 21, 8, 735-758, 2002.

[8] Kroger T., "On-Line Trajectory Generation: Nonconstant Motion Constraints," IEEE International Conference on Robotics and Automation River Centre, Minnesota, USA, 14-18, 2012.

[9] Murray D. and James L., "Using Real-Time Stereo Vision for Mobile Robot Navigation, Journal of Autonomous Robots, 8, 161–171, 2000.

[10] Martin M. and Moravec H., "Robot Evidence Grids," (Technical Report), CMU-RI-TR-96-06, The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 1996.

[11] Youngblood M., Holder B., and Cook J., "A Framework for Autonomous Mobile Robot Exploration and Map Learning Through the Use of Place-Centric Occupancy," ICML Workshop on Machine Learning of Spatial Knowledge, 2000.

[12] Jia S., Yang H., Li X., and Fu W., "SLAM for mobile robot based on interactive GUI," IEEE International Conference on Mechatronics and Automation (ICMA), 1308-1313, 2010.

[13] Elfes A., "Using occupancy grids for mobile robot perception and navigation," IEEE Transactions Computer Magazine, 22, 6, 46–57, 1989.

[14] "Intelligent Mobile Robotics System: Intelligent Interaction Capabilities with an Active Stereo Vision - Path Planning Navigation," King Saud University Internal Report, 2014.