MRL @Home 2013 Team Description Paper

Fariborz Mahmoudi [†], Ramin Fathzadeh[†],
Alireza Hosseini, M.J. Namazifar, Navid Nabavi, Farshid Abdollahi, Majid Iranikhah,
Salar Bayani, Farshid Abazari, Reza shahabian and Hamed Bagheri

[†] Mechatronics Research Laboratory
Qazvin Islamic Azad University
{mahmoudi,fathzadeh}@qiau.ac.ir
http://www.mrl.ir

Abstract. This paper devoted to describe MRL @Home robot. Our team has participated in RoboCup international competition since 2009. The leading improvement of MRL @Home since last year is developing Robot Operating System as the main control platform of robot. ROSARIA is recruited to control Pioneer 3-AT robot. It controls and receives position estimates, sonar and bumper, and all other current operating data sent by the robot platform. GMapping is used for navigation which is a highly efficient particle filer to learn grid maps from laser range data. Face recognition is achieved by Eigenfaces and object recognition is done by clustering algorithm. Collision free arm navigation permits the robot to grasp the objects. Speech recognition is performed by Kinect and Microsoft Speech Engine. To achieve better recognition, Speech Recognition Grammar Specification (SRGS) is used. And, a couple of improvements have been done in mechanical and electrical platform as well.

Keywords: Robina, Robot Operating System, Mobile Manipulation, Human Robot Interaction.

1 Introduction

Mechatronics Research Laboratory was established in 2003 with the support of Qazvin Islamic Azad University. MRL has several active laboratories that they are working proficiently on the robotic. Through the participation in the RoboCup competitions the teams have experienced and shared their knowledge with other research groups, and they have improved the quality of the developed technology. MRL teams are arranged to initiate several research programs on autonomous mobile robot such as: simultaneous localization and mapping, navigation strategies, motion planning, sensor fusions, scene understanding, visual odometry and exploration algorithms. With support of existence knowledge in the MRL, we have developed and improved an autonomous mobile for doing standard tests in @Home RoboCup contests. Our robot is powered by one laptop, one industrial PC and embedded controlling boards which they are connected together via network. MRL @Home lab was established in 2008 and has participated in RoboCup international competitions since 2009. Prominent experiences are obtained through our taking part in RoboCup contests. MRL @Home has recently had outstanding contributions in technical committee of @Home RoboCup league and also organizing RoboCup Iran Open competitions. This article is devoted to represent our efforts in designing and developing MRL @Home robot. Accordingly, the reset of this paper is organized as follows. Mechanical and hardware design are described in section 2 and 3 respectively. Section 4 is devoted to sketch the mobile manipulation. Human Robot Interaction (HRI) is concerned in section 5, section 6 describes task planning, and future work and team members' contributions are shown in section 7 and 8.

2 Mechanical Platform

The important tasks of MRL @Home's mechanic team are designing, improving and optimizing mechanical features and structures that are consisted of the robot's platform and its upward structure, navigation, manipulator and robot's appearance. The base of MRL @Home's robot, which is called *Robina*, is a modified Pioneer 3-AT (P3AT). It is a highly versatile all-terrain robotic platform, software compatible with all Active Media robots, chosen by many DARPA grantees and others requiring a high-performance robot with plenty of real estate for customization. We have made some changes in P3AT's motors and powering system so it can carries more weight. (See Fig.1)

Fig. 1. Robina with Pioneer 3-AT Platform

Robina have used Kinova Jaco arm that has six degrees of freedom (DOF) that shown in Fig. 2, it has a total weight of 5.7*kg* and allows a secure and esthetic fixation to mobile platform. Its maximum load is 1.5*kg* at the midrange and 1*kg* when fully extended, which is ideal for lifting objects. Carbon fiber makes it a product that is both extremely strong and light weight, powered by a 24*V* battery or a standard electrical plug. The arm has a reach of 90*cm* and is capable of moves up to 15*cm/s* in Cartesian movements and 8*rpm* in angular movements. Its hand, with its three fingers, can easily grasp objects.

Fig. 2. Kinova Jaco Arm with 6 DOF

The base platform of Robina is a four wheeled mobile robot, which two wheels of it in one side is connected be a belt to two DC motors. Using Gearbox with ratio of 1/ 4.5 makes robot's movements more powerful and accurate. The planetary gearbox is mounted in hollow space of wheel ring which is adapted on preliminary platform. It's carries out navigating, transferring electronic boards, computers, power suppliers and any other essential equipment at the same time.

Robina's Spinal material is made of carbon fiber that makes the robot's weight very light. In the head of the robot, two Dynamixels RX-64 used for pan-tilt movements and an external LCD is mounted for monitoring usages in lower part. Robot face material is made from carbon fiber that makes Lightness and shock absorption during the movement. Controlling and getting feedback data is possible with many encoders and sensors like IR sensor, sonar, laser scanners, load cell and even vision abilities of the robot if it is necessary.

3 Hardware Design

We have designed a low-cost, modular and flexible robot using industrial modules. Our goal for designing the robot is based on similarity with human body structure. The robot hardware equipped with two laptops, two Hokuyo laser scanners, two RGB-D cameras and two HD cameras. Additionally, it is equipped with lighting diodes attached behind the robot changes, its color whenever a critical event report is triggered from robot controllers. We have used several low-cost and small controlling boards which each board independently control one device and connect to other controllers and spinal cord through the RS-485 and CAN standard bus. Due to this modular network structure, each part of the robot can easily add/remove or change. So it makes the robot hardware very flexible and compatible with industrial standards. The spinal cord coordinates all of controllers and connections with computers through an Ethernet bus. It acts very similar to human body spinal cord.

The control system comprised of four main parts:

- PID controller that receives velocity commands from computer and sensors then sends suitable data to motor driver boards.
- Smart motor drivers.
- Data acquiring system that acquiring data from all sensors.
- Ethernet modem that connects spinal cord and laptop together.

4 Mobile Manipulation

A vast set of chores which could be defined for a service robot in a domestic environment contains mobile manipulation scenarios, such as grasping objects from different containers, delivering them to specific persons or placing them in different planes. So a service robot needs to integrate robust abilities to safely navigate through the environment, object detection, recognition and grasping skills [1].

4.1 Real-Time Object Detection. In order to detect objects and also acquire 3D position of them, we have equipped Robina with an “ASUS Xtion ProLive” RGB-D camera which is provides depth images in addition of RGB images. We hypothesize that most of objects in the home environment are placed on horizontal planes, so in our approach, before looking for graspable objects, we apply a Plane Model Segmentation method to find dominant planes like tables or shelves. This method uses RANSAC [2] algorithm to segment all the points within a point cloud that support a plane model. Point clouds are acquired from mentioned RGB-D camera and the OpenNi framework. Before any further processing on the acquired point clouds we down sample the points using a Voxelized Grid approach. Using this approach the underlying surface can be represented more accurately. In order to increase the accuracy of plane segmentation, we have combined surface normal estimation as an additional constraint with plane model fitting method. A surface normal needs to be estimated from the surrounding neighborhood of each point, also called K-neighborhood [3]. For finding K nearest neighbors of a specific point, FLANN-based KdTree [4][5] used as the search method of the surface normal estimation. A sample of normal estimation using this search method is shown in Fig. 3.

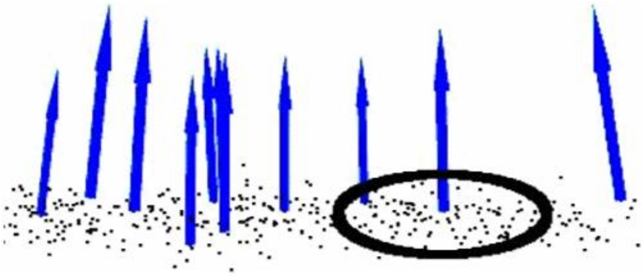

Fig. 3. Surface normal (blue arrows) estimated from a set of points. Black circle shows the defined radius of sampled point cloud dataset used for nearest-neighbor estimation.

After segmenting the plane model we construct a convex hull polygon for the set of points supported by the plane. Then, we generate a 3D polygonal prism using the planar model and a predefined height. This polygonal prism is going to be used in segmenting all points lying inside it.

The points supported by prism are considered as our region of interest (ROI). Then, Euclidean Cluster Extraction algorithm on extracted indices of mentioned prism is applied to identify individual objects. An expected result of described processes is demonstrated in Fig. 4.

(a) Example table scene (b) Depth registered point cloud (c) Segmented objects

Fig. 4. (a) Example of a dominant plane with some graspable objects resting on it. (b) Registered point cloud, aligned with RGB data. (c) Segmented objects are viewed with various colors, segmented plane with blue lines and plane convex hull with Yellow Square.

4.2 Object Recognition. In our approach at first 3D location of the segmented object in the depth image will be project into the RGB image plane which is obtained from RGB-D camera mounted on the head of the Robina. Then, a rectangular region of interest whole the pixel obtained from projecting operation will be defined. Furthermore, in order to extract key features of the object model, an object perception technique vocabulary tree [6] based Scale Invariant Feature (SIFT) [7] will be apply, on defined ROI. In this approach a previously built database from quantized SIFT features of the textured objects is used to compare extracted features of the object model in the ROI, with stored key features of different objects model. Finally, the best matches with their votes, the object model with the highest vote and higher than a predefined threshold will be considered as an identified object model.

4.3 Object Manipulation. Carrying an object from a start to the goal location requires a sequence of actions including pre-grasping, grasp acquisition and post-grasp transport [8]. In the first phase we adjust the end-effector to a pre-grasp pose. The pre-grasp pose will be calculated according to the height of the supporting plane and also pose of the object. In our approach the end-effector will be aligned parallel to the pose of the object in the vertical axis with a predefined distance from the object and 0.03*m* above the surface. Also the yaw orientation of the end-effector will be established according to the pose of the object and kinematic constraints. Once the end-effector reached to the pre-grasp pose, the end-effector will approaches to the object using a free space reaching motion and closes the finger tips. In the latter phase the object will be transported from grasp acquisition pose to the goal pose.

4.4 Navigation. Reliable obstacle avoidance depends to an accurate map, precise localization and smooth navigation. For achieving to this purpose, an accurate map will be built by GMapping algorithm [9], odometery and laser scanner's data. This algorithm is designed for long range sensor, so Robina uses a hokuyo UTM laser scanner with 30*m* scan range. It is an occupancy grid based mapping which makes a map with smaller piece of environment measurement.

In case of SLAM, Rao-Blackwellized particle filter [10][11] is used to filter some of particles to decrease complexity of SLAM procedure.

Robot's localization uses a state-of-art AMCL [12] method. AMCL is an adaptive type of MCL [9][13], which is uses probabilistic particles to estimate robot coordination in the environment.

Robot's navigation needs an algorithm to find out a way from a starting point to the destination. For this purpose DWA (dynamic window approach) [14] is proposed, which chooses the best path among the possible from beginning to destination point by evaluating each trajectory from proximity to the goal, global path, obstacle. It picks the highest scored trajectory and sends the related velocities to the mobile platform. In order to avoid from dynamic obstacles in the environment, laser scanner sensor is mounted in front of the robot in the height of 25*cm.* Detected obstacles added to the costmap will be used in further processing of robot's planner. Ultra sonic sensors mounted on the rear of robot are used to avoid from possible obstacles in the path of backward movements. In a home environment there are lots of impassable objects, which could not be perceived by 2D laser scanners. So RGB-D camera mounted on top of the robot solves this problem. It's provides 3D representation of the objects, which can be added as an obstacle to the robots costmap.

5 Human Robot Interaction

5.1 Human Tracking. One of the major fields of Human Robot Interaction is to track and then follow Human in the environment. Most of the human tracking methods are based on Kalman filter techniques.

Currently, our human tracking method is based on torso detection. For this purpose, Robina is equipped with a UBG-04LX-F01 laser range finder (LRF) at a height of 130*cm* that shown in Fig. 5. Additionally, in order to decrease the rate of false detections, laser range is limited from -45 to +45 degrees.

At the first step, a clustering algorithm will be applied on the measurements from LRF. Set of points that have been clustered properly can be considered as the next clustering candidate, if the changes are fairly low. According to the fact that the width of human torso is approximately similar, those clusters with inappropriate width would be ignored. And the structure of Human torso can be considered as oval cylindrical in Fig. 6.

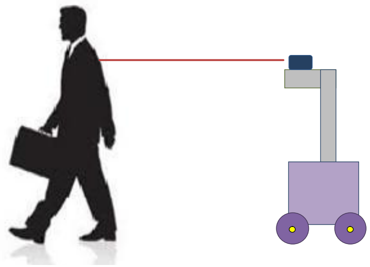

Fig. 5. The scheme of human detection using LRF

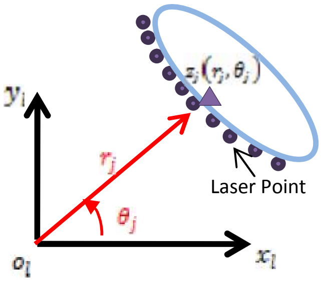

Fig. 6. The blue oval shows torso candidate obtained from laser measurements demonstrated as blue points.

After clustering, it is possible to have multiple clusters, so some of clusters might be filtered according to the following criteria:

- Humans' normal walking speed, current cluster would be rejected, if its position does not fairly close to its prior position.
- Cluster with the width of lower than 40 and higher than 60 (width of a normal person) will be rejected.

If conditions are being established, then it is selected as a candidate.

The measurement Z_j is distance from center of candidate object to robot LRF. In this measurement r_j is polar distance and θ_j is equivalent to the angle from LRF, which can be translated to a global coordination of the robot.

To observe sequential movement of each body candidate in environment, each data coming from the sensor should be validate and then clustered correctly. The position of the estimated human's torso will be update at every sampling time. The simplest way to assign a real measurement to a predicted observation is the Nearest Neighbor method [15]. Generally, according to the sensors noise and random motion mobile robot, the state and position of human aren't stable and confident. So, in order to earn optimized position state, Kalman filter method has proposed.

5.2 Speech Recognition. In the service Robots, one of the most important manners to communicate with human is the speech, so it is needed to be performed robustly. So speech recognition is supported by employing Microsoft Speech Platform to implement our speech recognition module. It is a speaker-independent with support of separated grammars and dictionaries to perform differently according to situations.

Also our speech recognition module has been developed as an ROS (Robot Operating System) package. For the better recognition Kinect sensor have used which, provides polar position finding of the talker and identifying the nearest one by the voice strength [16]. It helps the robot to make better interaction with the human by turning his head toward the speaker. Furthermore its gives us the capability of focus on the main speaker and accordingly avoids the module from capturing noises.

5.3 Natural Language Processing. The received command is a human natural language (NL) and needed to be transformed to an inner representation [17]. This quantization described in sequence of methods to be applied on the NL to achieve the claimed behavior from the Robot.

To perform Part-of-Speech tagging, Rule-Based method have been used [18]. According to the syntax all the words which have defined in the grammar of the speech Recognition module have their own predefined tags. It also helps to avoid statistical methods which are having less performance. The tagged sentences need to offer more conceptual information, So Name Entity Recognition methods are used to map verbs to the specific action and nouns to its specific entity, for example location, person or etc [19].

5.4 Face Perception. Face perception is concerned with how we recognize individual instances of a type of object. It is performed using two following stages:

Face Detection. Where images are searched to find any face, then image processing cleans up the facial image for easier recognition. For detecting faces in an image we are using Haar Cascade Face Detector of OpenCV [20] Library.

Face Recognition. Where detected and processed face is compared to a database of known faces, to decide who the person is? After generating preprocessed facial image, we perform Eigenfaces for Face Recognition [21].

5.5 Gesture Recognition. Gesture recognition is used to communicate with robot more interactive. OpenNi Skeletal tracker is employed to track human's body and the current state of body elements. Eight postures are defined for robot, such as waving to attract robot's attention, stop gesture to stopping the robot in emergency situations, clapping for demonstrating happiness, hands up for call the robot, demonstrate left and right directions by hand and stretch foots to sides to impose the robot to move in corresponding directions. In addition, sometimes the layman needs to draw attention of the robot to a specific point or location on the ground. For this purpose, robot keeps tracking the position of shoulder, hand and elbow to calculate the actual coordination of specified point.

6 Task planning

Class based system used to solve the complex tasks. An abstract action is defined, which other actions will inherit from it or other previous inherited actions. So the robot can perform a complex action by dividing it to simple actions which are the parent actions. An *action* defined as an object contains a sequence of three behaviors: *Initialization, Job, and Finalization*. Initialization separated in two parts: initialization of the parent and itself. After the action initialized, the main Job is ready to perform. The sequence of performing the main actions and its parents depends on the action requirements. These requirements might be satisfied before, after or between the sequences of processes. At last the Finalization will perform. It contains two parts: Finalization of the target action and parent actions. In our solution, additional facts for each action are stored as a global memory. These facts will be included to the initial state in the next planning [22]. By processing the global memory, it will be found out what has been done before and doesn't need to perform again.

7 Future Works

There are some ongoing tasks which are currently concerned by MRL @Home lab. In the robots mechanical structure, the spinal body of robot is going to have some major modifications. A ball screw mechanism will be replaced with current rigid carbon fiber spinal body, to add the facelift capability to the robot. Furthermore tilting laser scanner will be mounted on the chest of the robot. In case of electronic and control system, Dot Matrix LED is going to be used in showing robot facial expressions and also battery management system (BMS) to control and monitor the battery charge and discharge. BMS system has the capability of estimating the remaining charge, optimal charging and battery protection. Furthermore a new hardware architecture based on ARM processors is under development.

8 Team Members and Their Contributions

MRL @Home is comprised of following members with their contributions.
- Fariborz Mahmoudi: scientific supervisor,
- Ramin Fathzadeh: team leader,
- Alireza Hosseini: head of software group, mobile manipulation,
- Mohammad Javad Namazifar: Human Tracking, programming group,
- Farshid Abdollahi: Robot navigation, programming group,
- Farshid Abazari: Human tracking, programming group,
- Navid Nabavi: Task planning, ASR, programming group,
- Minoo Hamisi, Object Perception, programming group,
- Amirhossein Lesani: software infrastructure development, programming group,
- Bahman Ghiasvand: programming group,
- Mohammad Reza Shahabian: Head of electronic group,
- Hamed Bagheri: electronic designer,
- Bahram Jamshidi: electronic designer,
- Azin Shabani Asl: electronic designer,
- Reza Ahmadi: microcontroller programming,
- Majid Iranikhah: Head of mechanic group,
- Salar Bayani: mechanical designer,
- Reza Chegini: mechanical designer,
- Farzad Gardaneh: mechanical designer,
- Farhad Maroufkhani: mechanical designer

References

1. Jörg Stückler, Ricarda Steffens, Dirk Holz, and Sven Behnke: "Efficient 3D objects perception and grasp planning for mobile manipulation in domestic environments", 2005.
2. M. A. Fischler and R. C. Bolles: "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography", ACM, 24(6):381–395, 1981.
3. Bremner D, Demaine E, Erickson J, Iacono J, Langerman S, Morin P and Toussaint G: "Output-sensitive algorithms for computing nearest-neighbor decision boundaries". Discrete and Computational Geometry 33(4): 593–604. doi:10.1007/s00454-004-1152-0.
4. J. L. Bentley: "Multidimensional binary search trees used for associative searching", ACM, 18(9):509-517, 1975.
5. Marius Muja and David G. Lowe: "Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration", International Conference on Computer Vision Theory and Applications (VISAPP'09). 2009
6. D. Nister and H. Stewenius: "Scalable recognition with a vocabulary tree", Conference on Computer Vision and Pattern Recognition, pp. 2161–2168, Washington, 2006.
7. Lowe, David G.: "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision. 2. pp. 1150–1157.doi:10.1109/ICCV.1999.790410.
8. Lillian Y. Chang, Siddhartha S. Srinivasa and Nancy S. Pollard: "Planning pre-grasp manipulation for transport tasks", IEEE International Conference on Robotics and Automation (ICRA), May 2010.
9. Sebastian Thrun, Dieter Fox and Wolfram Burgard: "Probabilistic Robotics", 2002.
10. Giorgio Grisetti, Cyrill Stachniss and Wolfram Burgard: "Improved Techniques for Grid Mapping with Rao-Blackwellized Particle Filters", IEEE Transactions on Robotics, 2006.
11. Giorgio Grisetti, CyrillStachniss, and Wolfram Burgard: "Improving Grid-based SLAM with Rao-Blackwellized Particle Filters by Adaptive Proposals and Selective Resampling", In Proc. of the IEEE International Conference on Robotics and Automation (ICRA), 2005.
12. Lei Zhang, Rene Zapata and Pascal Lepinay: "Self-Adaptive Monte Carlo Localization for mobile robot using range sensors", IROS, 2009.
13. Sebastian Thrun, Dieter Fox, Wolfram Burgard and Frank Dellaert: "Robust Monte Carlo localization for mobile robots, Artificial Intelligence conference", 2000.
14. D. Fox, W. Burgard and S. Thrun: "The dynamic window approach to collision avoidance", Robotics & Automation Magazine, IEEE.
15. Y. Bar-Shalom and X. R. Li: "Multitarget-Multisensor Tracking: Principles and Techniques". vol. 3, 1995
16. Hari Krishna, Maganti Daniel and Gatica-Perez: "Speaker Localization for Microphone Array-Based ASR: The Effects of Accuracy on Overlapping Speech", ICMI'06, 2006.
17. Xiaoping Chen, Jianmin Ji, Jiehui Jiang and Guoqiang Jin: "A Home Robot Integrating NLP and ASP Techniques: Integrating NLP

with ASP for Intelligent Service Robots", 2009.

18. Eric Brill: "A simple Rule-based Part of Speech Tagging", HLT Proceedings of the workshop on Speech and Natural Language. Pages 152-155. 1992.
19. Maksim Tkachenko and Andrey Simanovsky: "Named Entity Recognition: Exploring Features", KONVENS 2012. pp. 118-127. 2012.
20. Rainer Lienhart and Jochen Maydt: "An Extended Set of Haar-like Features for Rapid Object Detection", IEEE ICIP, Vol. 1, pp. 900-903, Sep. 2002.
21. P. Belhumeur, J. Hespanha, and D. Kriegman: "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection", IEEE Transactions on pattern analysis and machine intelligence.
22. Thomas Keller, Patrick Eyerich, and Bernhard Nebel: "Task Planning for an Autonomous Service Robot", KI'10 Proceedings of the 33rd annual German conference on Advances in artificial. Pages 358-365, 2010.