AUT @Home 2014 Team Description Paper

Alireza Ahmadi, Edwin Babaians, Amir Badamchi, Mohsen Falahi, Shaghayegh Gharghabi, Ramin Gomari, Sara Jahangiri, Marzieh Kariminoori, Elahe Khani, Ismaiel Mehrabi, Ameneh Sheikhjafari, and Saeed Shiry Ghidary

The names are arranged in alphabetical order

Robotics Research Center, Amirkabir University of Technology (Tehran Polytechnic), Tehran, Iran
http://autathome.ir

Abstract. This paper introduces the AUT@Home robot of Amirkabir University of Technology. We describe the hardware characteristics and capabilities of this robot to perform tasks which are related to human daily life. In addition it covers our approaches for Natural Intelligence (NI) tasks such as speech and object recognition which has made our robot reliable for a wide range of tasks in home environments.
Keywords: holonomic mechanism, service robots, Human Robot Interaction (HRI), robust speech recognition, object recognition.

1 Introduction:

AUT@Home team consists of MSc and BSc students from different departments with different abilities. Our team members are previously involved in a variety of projects which cover different aspects of “at home” robots and generally “Social Robots”. These projects are advised by the faculty members of Amirkabir University of Technology. This team has been founded in December 2007, which was known as “Sourena Team”. The first participation of this team in international competition was in RoboCup2008. In that competition we had a proper demonstration and we achieved the 5th place. Additionally, we had participated in RoboCup three years consecutively. Through these years we have collected invaluable experiences which made our team stronger for future participations. At 2013 we restarted our work by building a new robot. This robot which we call that “Sepanta” participated in AUTCup2013 and achieved the 1st place in that competition.

One of the most important points of our work is that AUT@Home robot, Sepanta, is equipped with RGB-D Camera and laser scanner for 3D perception. Sepanta uses the data from these scanners for navigation, real time environment perception, object recognition, and manipulation. Besides this point, Sepanta’s actuators are highly reliable which allow us to have a perfect control on the robot manipulator and motion mechanism. Since Sepanta has two different arms, which can perform in different ways, our robot can carry out a wide range of manipulation tasks in a real world.

The remaining of this paper is organized as follows; firstly, the electrical and mechanical design of the robot is introduced. Secondly, the software platform of the robot is described. Finally, our plan for future modifications on the robot is mentioned.33

2 Hardware Design

2.1 Mechanical Part of AUT@Home Robot

Since @Home robots should be able to perform in a dynamic environment, we have designed and implemented a “Holonomic” motion mechanism for Sepanta. As a holonomic mechanism in a discoid workspace has three DOF, it can move omni-directional in the environment. To implement the holonomic mechanism, three omni-wheels are used and placed in 120° to each other. This mechanism is shown in Fig. 1.

Fig. 1. This figure shows Sepanta's motion mechanism which contains three omni-directional wheels.

In addition, based on the fact that the robot is assumed to accomplish human-like duties, we have implemented two arms for manipulation tasks. Each of these arms at least has four DOF which a shoulder has connected them to each other. Since the robot should be able to manipulate objects with different shapes, each arm has a specific gripper which enables Sepanta to use each gripper in a certain task. Fig. 2 shows robot arms and grippers.

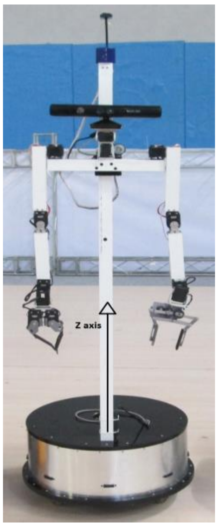

Fig. 2.AUT@Home robot, Sepanta.

Despite the fact that we have tried to design a bio-inspired mechanism for Sepanta, we have used a prismatic joint in Z axis to let the robot reach objects in all tasks and situations.

Actuators

It is clear that in robot mechanical design, the most important factor is actuators. Generally, motors can be used in three kinds of utilizations; using in revolute joints, prismatic joints, and for wheel driving.

Since Dynamixel motors [1] can satisfy all mentioned requirements, we have used them in all Sepanta's joints. These motors are manufactured by Robotis Company and introduced as "Smart Motors" which are able to work like both DC and Servo motors. It means that we can use such motors in both Sepanta's joints and motion mechanism. Table 1 shows the motors which we have used in this robot.

Actuator	Number	Location
MX-64	6	Motion Mechanism (3), Z prismatic joint (1), Elbow (2)
MX-106	2	Shoulders (2)
RX-28	4	Pan and tilt (2), right hand wrist (2)
RX-24	2	Right hand gripper (2)
AX-12+	3	Left hand gripper (3)

Table 1. this table shows the number, location, and models of actuators we have used in our robot.

It is worthy to mention that "Smart Motors" have their own embedded driver, controller, and related sensors, including encoder, temperature sensor, current meter, etc. One of the most important features of these motors is their ability to make a network via RS485 which is a reliable protocol for industrial usages.

2.2 Electronic & Sensors

In this section we will describe sensors and circuits which we have used for data acquisition and filtering. Additionally, the robot controller architecture is presented and discussed in the following text.

In this robot an AVR® microcontroller DX-IO board (DX-IO controller runs under 8bit AVR ATMega8) is used as a low level controller and Device Communication Manager (DCM). Different instruments such as actuators and sensors communicate with main processor by this board. In low-level computations on this board, we drive 3 types of sensors:

- A 3DOF magnetometer RM-G146 sensor.
- Integrated distance infrared based on GP2D120 sensors.
- Internal actuators for load, speed and absolute position sensors of Dynamixels.

In addition, we also used USBzDXL, which is a high speed USB to half-duplex serial converter, as a motor controller for Dynamixel actuators. This board enables more than 50Hz update rate for high resolution data of internal sensors. Fig.3 shows Sepanta's hardware architecture which includes low-level controller, main controller, and peripheral connected devices.

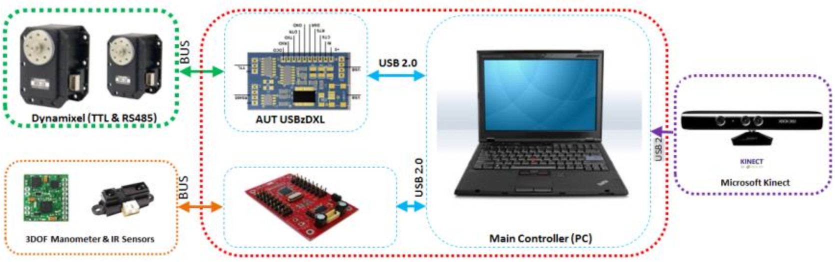

Fig. 2. DX-IO low-level controller & main PC and peripheral connected device

3 Software

The processes which run on the main controller (see Fig. 3) are known as our software part and are discussed in this section. The robot's operating system architecture is based on ROS (Robot operating system) which runs under Ubuntu. In this software there is a core node which other nodes connect to it and transfer data via this node. The main role of this node is request management between all nodes. In the following we will discuss Sepanta’s software design in detail.

3.1 Human Robot Interaction (HRI)

One of the most important subjects that should be implemented in Service Robots is Human Robot Interaction (HRI). Human detection and following is the basic concept of HRI which should be done in a real-time manner. Since the robot is not allowed to damage its surrounding environment, HRI should be safe for both human and the environment. @Home robots are designed to work in a human daily environment which is highly dynamic and these robots should be able to follow humans without any collision. In the following text we will describe human following ability of Sepanta.

As Sepanta has a Kinect sensor, depth information of a scene is available and the background of each image can be detected and removed. We have implemented this part of our algorithm using OpenNI2 and Nite2. By using these libraries background is removed and humans are detected properly. In the next step, for human following, OpenTLD is used. This library has the ability of learning and tracking a certain part of an image. It is worthy to mention that OpenTLD fails to track human in clutter backgrounds so, we introduced a novel algorithm for human following which combines the information of both Nite2 and OpenTLD which enables the robot to follow human in crowded areas. Our algorithm does not need camera calibration which this feature made it more feasible for service robots.

Another challenge which we deal with in this section is finding human who was previously in the robot’s sights and is not visible anymore. The searching algorithm starts with the robot’s neck motions. If the robot found the person searching finishes, but if it did not find that, it starts to move through its motion mechanism according to the direction that the robot saw the person in the last time.

3.2 Person Re-identification

Since a service robot has to be able to interact with humans, the human re-identification during this interaction should be real-time. It is clear that human face is one of the most significant features by which we can detect humans and distinguish them from each other. However, this feature has some deficits in cases that we cannot see the face, hence, we should not rely on only human face features. Therefore, we used other features like body features [2].

For face detection we have used Haar-cascade [3] and after that Sparse method [4] is used for face recognition. Our algorithm for human re-identification consists of the following steps;

1) Learning: in this phase we teach the robot via some training data.

2) Face detection using Haar-cascade method.

3) Applying Sparse algorithm to the detected faces.

4) Similarity calculation between detected faces and learnt faces.

After these steps, detected faces based on their similarity with learnt faces will be classified in related classes and also, if the similarity was less than a threshold the detected face will not classify in any previously known classes.

Using skeleton information provided by OpenNI2, in this classification we have employed body features including height, shoulder width, etc. besides features of the face.

3.3 Voice interaction

Verbal communication is a hallmark of human's interactions and it is obvious that a service robot should be able to have such communications with others. Sepanta is able to have verbal communication before starting a task, during its completion, and after finishing that. This ability has made our robot highly reliable in all tasks based on the fact that when Sepanta needs more information about an assigned task, she asks that information and after completing that she waits for an acknowledgment.

In this robot we have used Microsoft Speech SDK [5] for speech recognition. To avoid noisy data from the microphone Sepanta is equipped with "Sennheiser MKE400" microphone which has noise canceller. In addition, for robust sentence recognition, a grammar file is used. This auditory system is independent of speaker who talks with the robot and therefor it does not need to pass matching process during the competition. Additionally, as it is possible that different people call the robot from different directions simultaneously, Sepanta changes her direction based on the sound direction and it helps her for a better interaction with intended human.

As in a verbal communication there are two sides and Sepanta needs to talk with others, her speeches synthesize using Festival [6] software. Sepanta's verbal interface is able to recognize gender and age of the people who talks with her and she can changes her speaking rate based on her understanding of the situation.

3.4 Simultaneous-Localization and Mapping

It is clear that each mobile robot should be able to estimate its own position during its movements in the environment and create a belief (map) of its surrounding environment. If a robot be able to obtain its position accurately, with the assumption that its sensors work perfectly, the map which is created by the robot will be accurately too.

In this research for localization and motion estimation we have fused data from both a laser scanner and a Kinect sensor, although during certain tasks visual odometry using Kinect might be disabled to reduce processing. Localization using laser scanner is done using standard AMCL [7]. Visual odometry is more complex and is done using Fovis [8] algorithm but will fall back to GICP [9] if Fovis failed to provide a motion estimate.

We used a keyframe approach for SLAM and place recognition. Keyframes are defined when the robot moves a certain distance or rotates for a certain angle. If a keyframe was detected, the current image captured by the Kinect sensor would be sent to place recognition module which uses RTAB-Map [10] algorithm. In the case that the image was previously seen, the place recognition algorithm will produce a loop-closure message for the SLAM system. Our SLAM system uses g2o [11] algorithm to sparsely represent robot movements.

Although we mostly rely on laser scanner data for navigation using GMapping and ROS navigation stack, for more complex environments we can also use OctoMap [12] to produce a 3D map with Kinect point clouds, but we currently map the 3D result to two dimensions since path planning is less complex and satisfies our current needs.

3.5 Object Recognition

In this section, object recognition algorithm is described. Object recognition algorithm executes as follows; firstly, searching for planes starts and after that the detected planes will be eliminated from the point cloud. Secondly, the remaining points will be clustered based on Euclidean clustering and saved in memory as seen objects. After this step, the similarity of Viewpoint Feature Histogram (VFH) between seen and learnt objects will be calculated. Finally, if the similarity between a detected object and a learnt one was more than a threshold it would be considered as a member of that known object. It is clear that if the detected object was not as much as similar to any known object, it will be classified as unknown. The similarity calculation is so fast because of the fact that only the VFH of learnt objects are saved in a Kd-Tree structure [13].

In the next step, in mobile manipulation, object manipulation block needs the position of each detected objects to be able to reach them. Objects coordinates obtain through the Kinect sensor data and it refers to the center of the object.

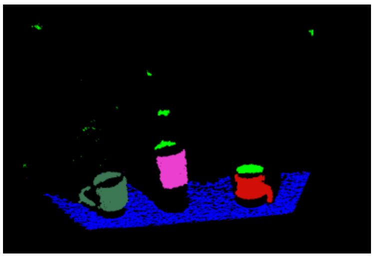

Fig. 3. this figure shows a visual output of object recognition block. The last eliminated plane in this picture is colored by blue and the objects' colors are randomly chosen. The green points refer to those parts of the picture which are classified as neither object nor plane.

3.6 Object Manipulation

After object recognition and the decision to achieve that object, the object manipulation block starts its work. The input of this block is the coordinates of tables and objects in the observed scene and its output is the trajectory which the robot should execute to reach the object.

The output trajectory of this block is influenced by two items; Obstacles which prevents the robot arm to achieve the goal position, the kinematics of the manipulator. Based on these items, during trajectory generation and execution, the position of the manipulator end-effector and joints adjusts accordingly.

As there are some well-established positions which occur frequently, we have programmed these trajectories for the robot to reach them rapidly. One of these trajectories is pre-grasp position which set the manipulator for grasping task. This position is modifies according to the height of tables and objects coordinates.

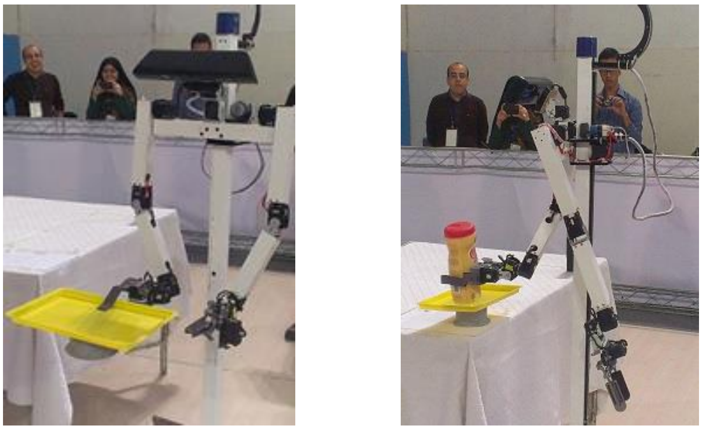

Fig. 4. Object recognition and grasping during the AUTCup 2013 in Iran

4 Conclusion

In this research we have designed and implemented a robot with seventeen degrees of freedom which enables our robot to perform easily in human daily environment. Since the robot motion platform has 3DOF, it can move omni-directional. This feature made Sepanta more maneuverable in dynamic environments. Based on the abilities of AUT@Home robot, this robot can interact with people via speech and other equipment which helps our robot to complete assigned tasks in Robocup2014. There are some novelties in both robot implementation and algorithms which are mentioned widely in this essay. These novelties are included in object recognition algorithm, human following, mechanical design and other parts. It is worthy to mention that more information about our team and Sepanta is available on our website.

Future work

There are some modifications which we have considered on our future robot to improve its performance. In the case of electronic and control system we will apply some modifications as follows:

- Using Cubieboard instead of laptop.
- Using accurate IMU for estimating movement angle of robot.
- Using 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 consuming.

In the case of mechanical design the items which are listed below will be applied:

- Modifying robot structure in order to decorate the robot based on Social Robots appearance.
- Modifying robot arm in order to increase its manipulation skills and abilities.

Additionally, in software parts, we have developed a novel approach in accurate person recognition and we want to use it in RoboCup 2014 competition.

We hope that these changes make our robot more reliable in accomplishing assigned tasks.

Team Members and their contribution

- Team leader: Ameneh Sheikhjafari,

- Person recognition: Shaghayegh Gharghabi,

- Object manipulation and object recognition: Ameneh Sheikhjafari , Marzie Karimi-noori, Ismaiel Mehrabi,

- Speech recognition: Sara Jahangiri, Amir Badamchi,

- Human tracking: Edwin Babaians, Shaghayegh Gharghabi,

- Frame work: Ramin Gomari,

- Simultaneous-Localization and Mapping:Kourosh Sartipi,

- Programming group: Elahe Khani,

- Electronic designer and mechanical designer: Alireza Sheikhjafar, Mohsen Falahi, Alireza Ahmadi.

References

1. ROBOTIS. (2013). Dynamixel Actuator. Available: http://www.robotis.com/xe/dynamixel_en. Last accessed 25th January 2014.
2. Barbosa, Igor Barros, et al. "Re-identification with rgb-d sensors." *Computer Vision–ECCV 2012. Workshops and Demonstrations*. Springer Berlin Heidelberg, 2012.
3. Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.
4. Wright, John, et al. "Robust face recognition via sparse representation." *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 31.2 (2009): 210-227.
5. Esposito, Dino. "The Microsoft Speech SDK." *Microsoft Internet Developer, Feb*(1999).
6. Black, Alan, et al. "The Festival Speech Synthesis System, Version 1.4. 2."Unpublished document available via http://www. cstr. ed. ac. uk/projects/festival. html (2001).
7. Fox, Dieter, et al. "Monte carlo localization: Efficient position estimation for mobile robots." *AAAI/IAAI* 1999 (1999): 343-349
8. Huang, Albert S., et al. "Visual odometry and mapping for autonomous flight using an RGB-D camera." *Int. Symposium on Robotics Research (ISRR),(Flagstaff, Arizona, USA)*. 2011.
9. Segal, Aleksandr, Dirk Haehnel, and Sebastian Thrun. "Generalized-ICP." *Robotics: Science and Systems*. Vol. 2. 2009.
10. Labbé, Mathieu, and François Michaud. "Appearance-Based Loop Closure Detection for Online Large-Scale and Long-Term Operation." 1-12.
11. Kuemmerle, Rainer, et al. "g2o: A general framework for graph optimization." *Proc. of the IEEE Int. Conf. on Robotics and Automation (ICRA)*. 2011.
12. Wurm, Kai M., et al. "OctoMap: A probabilistic, flexible, and compact 3D map representation for robotic systems." *Proc. of the ICRA 2010 workshop on best practice in 3D perception and modeling for mobile manipulation*. Vol. 2. 2010.

13. Rusu, Radu Bogdan, et al. "Fast 3d recognition and pose using the viewpoint feature histogram." *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010.