BahiaRT@Home 2018 Team Description Paper

Josemar R. de Souza, Marco A. C. Simões, Robson M. da Silva, Jorge Campos, Ricardo S. Matos, Samuel S. Carvalho, and Yan G. dos Anjos

Bahia State University (UNEB), ACSO Research Group, Salvador, BA, Brazil
teambahiart@gmail.com
http://www.acso.uneb.br/bill/

Abstract. Bahia Robotics Team (BahiaRT) presents Bot Intelligent Large capacity Low cost (Bill) from the Center of Computer Architecture and Operating Systems (ACSO) at the Bahia State University (UNEB). Bill is an autonomous service robot and RoboCup@Home is a perfect environment to validate its abilities and performance. The paper describes Bill's functions, such as navigation, manipulation, people and object recognition, human-robot interaction and decision making as well as the Bill's hardware and software systems. Furthermore, the paper highlights research interests and scientific contributions of BahiaRT.

Keywords: Bill, assistive robotics, RoboCup@Home.

1 Introduction

Applications in robotics advanced in the last years and evidently RoboCup has contributed for this growth. BahiaRT is a scientific cooperation group of ACSO that, since its creation (2006), have participated in the RoboCup leagues, as well as regional tournament that follows exactly the same rules, to strengthen the Brazilian engagement in this initiative and to advance its solutions for intelligent robots through competitions. Initially, the team competed in 2D Soccer Simulation League and also in the Mixed Reality demonstration competition. In Mixed Reality, BahiaRT got the third place in RoboCup 2009 and the fourth place in 2010. BahiaRT also has developed the MR-Soccer Server, the main module of MR software infrastructure. In other league named 3D Soccer Simulation, BahiaRT ranked the fourth place in 2015 and 2016 and the sixth place in 2017. In RoboCup@Home, BahiaRT got the 13th in 2015 and 21th in 2016. In At Home league of Latin-American and Brazilian Robotics Competition (LARC), BahiaRT got the second place in 2015, 2016 and third in 2017.

Fig. 1. Bill and its concept.

Bill (Fig. 1) is the proposal of BahiaRT for RoboCup@Home league. Bill born in 2014 as results of research projects in assistive robotics and its main functions are communicating with humans through natural language processing, recognizing objects, faces and gender, and navigating through unknown environments.

Section 2 describes the main advances and scientific contributions of BahiaRT to assistive robotics. Section 3 describes the architecture and main functionalities currently developed for Bill. In section 4 some experiments and results are presented. Section 5 presents the conclusions and future work. Finally, at the end of this Team Description Paper (TDP) has a brief description of each hardware and software component used in Bill's development.

2 BahiaRT's advances in innovative technology

The main research interest of BahiaRT involves robotics and artificial intelligence, specifically focusing on interaction and cooperation between human-robot, applications of artificial intelligence for services, standardization and system integration, in addition to the construction of intelligent robots with low cost components and open source. Developing autonomous robot as Bill requires knowledge of computer and mechatronics technology to develop algorithms, integrated systems and hardware solutions. Bill uses various innovative technologies in a integrated approach, such as Robot Operating System (ROS) and its packages, Arduino boards, kinect OpenNI library, TurtleBot arm, computational vision and speech algorithms, among others. The experiments and tests used to evaluate Bill's behaviors involve quantization of parameters, discretization of control tasks, details of the software and hardware, innovative technologies, and other challenger tasks. Thus, the development of new methods and technologies to provide integrated solutions and the training of practitioners to develop complex task in robotics is an important contribution of BahiaRT over the years. However, despite of the importance this type of research, there is still not several autonomous robots running in human daily lives. Furthermore, as well as in other sciences, autonomous robots can achieve different results in practice due occurrence of unpredictable events and environmental changes. Therefore, there is still a long road ahead, with much work to be done for spreading services robots in the residences around the world. To fill this gap, practitioners should provide feedback on robotic research, sharing experiences for reducing differences among countries and their laboratories. To allow replication of research and to contribute for development of robotic research, BahiaRT provides access to the codes, hardware and software list in this paper and in its website.

3 Bill's Architecture and Functionalities

The functions implemented in Bill are represented in the diagram in Figure 2 and their main features are described in the following subsections. The Bill's architecture is basically divided into levels: the high level where are the strategies functions of the robot and the low level where are located controllers and drivers that send signals and receive commands of sensors. At the end of this paper has the list of the hardware and software used.

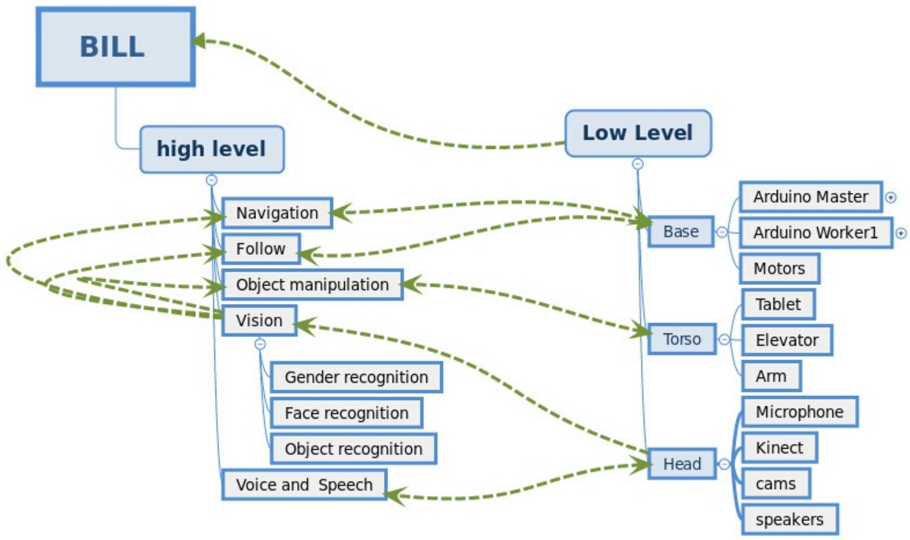

Fig. 2. Overview of the Bill's architecture and functions.

3.1 Navigation

Navigation is the keystone for efficient execution and environment interaction for robots. The components used by Bill for navigation are: encoders output, odometry, gmapping (ROS), move_base (ROS), Adaptive Monte Carlo Localization (AMCL) (ROS), map_server (ROS) and 360° laser scanner. The encoder data is used by odometry module to estimate the movements of the robot in space. Further, the odometry data is used to trace trajectory to a desired target by the move_base. Once all data are being published, the simultaneous mapping and localization using the AMCL [1] is activated integrating the 360° laser scan data. Simultaneous Localization And Mapping (SLAM) approach is used to map the environment and provide self-localization in this map. First the map is built using the incremental mapping package, Hector Mapping [2]. Then, the grid map generated is done by RPLIDAR 360 Laser Scanner sensor, which is able to get the 2D world data. The next step is creating the path planning based on the occupancy grid map that is updated based in Dynamic Voronoi ROS package approach [3]. Then the shortest path to the goal is computed by means of D* Lite algorithm [4] to ensure obstacle avoidance over incremental mapping. The motion planning in charge of getting the path planning and relating linear and angular motion is triggered, which applies the kinematics control law, and sends a message to low-level control.

3.2 People Detection and Tracking (Follow)

During Human Robot Interaction (HRI) people detection and tracking has a crucial role for service robots like Bill. The kinect OpenNI library provides position identification of key points on the human body, such as head, torso, knees, etc. This representation resembles a human skeleton, and it allows to obtain a persons' position relative to the camera on the robot. The library also assigns an ID for each person it identifies, allowing to track a specific person while he or she moves in front of the camera. That feature is used as an input for the navigation system, which will plan a path to follow a specific user in cluttered and dynamic environments.

3.3 Object Manipulation

Object manipulation plays an important role to interact with a home environment. To meet that requirement, Bill has an arm based on the TurtleBot Arm, composed of 5 Degrees of Freedom (DoF) including a gripper, which currently allows the robot to grab lightweight objects.

3.4 Vision

This module is responsible for receiving, processing and responding to external stimulus from image capture. Currently, the facial and gender recognition modules follow the proposed by [5] and have four steps: face detection, face preprocessing, training a machine-learning algorithm from collected faces and face recognition.

The first step must detect all faces in the image without worrying about recognizing it, just detecting human faces. This step is based on the Haar-Cascade classification method, which in turn is based on the cascade classification algorithm [6]. If in any step of the detection until the end of the cascade classification, one window is rejected, the classifier understands there are no faces. When this algorithm finds a face, it will separate the face from the rest of the image. To increase reliability in real-world conditions, methods of detecting facial features such as eye detection (which is being used in this project) were included, these classifiers are also based on the cascade classification [6].

After, a label is saved to identify each distinct preprocessed face, so that it can be re-identified. It is very important that faces have position variability, because saving them in single positions can make it difficult to recognize them, since people's expressions and behavior are not per-established

With preprocessed and saved images, these are trained in the Eigenfaces machine learning algorithm [7], which creates a model that allows to pre-tell who is in front of the web cam. With the algorithms of the class Face Recognizer of OpenCv library, a number responsible for measuring the reliability of being a particular person is returned.

For gender recognition is used the dataset with great variability of images of men and women assigning a label to each genre and apply two phases of facial recognition: preprocessing, for the face to match the characteristics of the dataset; and face recognition, for compaction of the proximity that an input face corresponds to the dataset.

Object recognition are based on color and shape features. The recognition is divided into four steps: (i) Acquisition of RGB-D sensor information, through the ROS [8]; (ii) Segmentation of the table in front the robot, followed by a search in the point cloud for planar structures using the Random Sample Consensus (RANSAC) method [9]; (iii) Segmentation of the objects on the table, and the points in the point cloud are converted to an OpenCV [10]; and (iv) Classification of each segmented object. Two types of features are used: (a) Color features, represented by the mean of histograms in each channel of LAB color space present in the object and (b) shape features, represented by the Hu's invariant moments [11].

3.5 Speech Recognition and Voice

The voice is the form of man machine interaction most used to give commands to the robot, either through specific commands or natural language. For recognition, Bill uses the CMU PocketSphinx, which features greater flexibility for adaptation and personalization, allowing to adapt the dictionary and acoustic models to the problem of context. After the speech recognition, the output of pocketsphinx is used to feed the state machine that applies the boost Regex library. Then, a grammar able to interpret the commands and fulfill the assigned tasks is constructed.

To speak with people, Bill use the ROS package sound play which can translate a ROS [8] topic into sounds. In the process of synthesizing, the software Festival [12] is used. It allows changing various aspects of voice, such as tone, speed of speech, among others, to ensure better understanding by the listener, allowing to generate better interaction experience.

3.6 Bill's GUI

The Bill's functions are encapsulated into a GUI module (Fig. 3) where the aforementioned functions are available in Portuguese. This GUI was designed to achieve dynamic utilization and optimal initialization of ROS nodes, using QT creator software integrated with ROS. Through the already existing message communication provided by ROS, this module was created just as a visualization tool, i.e., all the processing necessary are done in their own packages, as well as all information provided by them is sent to the interface which displays and sends requisitions to each respective package.

4 Experiments and results

Lots of experiments were made with Bill to measure and prove its abilities in different environments. For paper size limitation, here are described only the face

Fig. 3. Bill GUI

recognition results and speech abilities, however Bill responded with a similar accuracy in its other functions. Seven people were used in face recognition tests, which were divided into two steps: the former, only one person in front of the robot and the second with two people. In the first step 50 rounds were done for each person, in which the person appeared in front of the robot for 300 cycles and the robot had to recognize it. Bill attained an average of 72.64% accuracy. In the second one, 7 rounds were done combining the people in different pairs for each round. The process was made as in the first step. For every round, the robot had to recognize the people during 300 cycles. Bill scored the average result of 66.67% accuracy. The speech recognition had understanding and answering a serie of questions. Four people in two different environments and for each person the questions were asked 100 times. The indoors environment tests were done with the operator 75 cm distant from the robot. Bill obtained an average of 61.25% accuracy in this requirement. The outdoors tests, using the same settings as previous, scored an average of 56.67% accuracy in this task.

5 Conclusions and future work

This paper presented the main features of Bill the solution proposed by BahiaRT for assisting humans in their daily chores. During development and competition tests, Bill has proven to be an efficient solution with a good percentage of successful in its tasks. Market for human assistant robotics has an immense potential for growth and further works explore developing of new functionalities to make Bill more intelligent and autonomous for complex robots behaviors as well as to attend new markets. For example, an interesting research field that BahiaRT has explored is the ability of robots learning from its own experience, as well as, from other information source, or even from another robot. Future application for Bill technologies can also involve telecare robotics for patients by providing remotely located physicians with a means of medical supervision.

References

1. Frank Dellaert, Dieter Fox, Wolfram Burgard, and Sebastian Thrun. Monte carlo localization for mobile robots. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, volume 2, pages 1322–1328. IEEE, 1999.
2. Stefan Kohlbrecher, Oskar Von Stryk, Johannes Meyer, and Uwe Klingauf. A flexible and scalable slam system with full 3d motion estimation. In *Safety, Security, and Rescue Robotics (SSRR), 2011 IEEE International Symposium on*, pages 155–160. IEEE, 2011.
3. Boris Lau, Christoph Sprunk, and Wolfram Burgard. Improved updating of euclidean distance maps and voronoi diagrams. In *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, pages 281–286. IEEE, 2010.
4. James Neufeld, Michael Sokolsky, Jason Roberts, Adam Milstein, Stephen Walsh, and Michael Bowling. Autonomous geocaching: Navigation and goal finding in outdoor domains. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 1*, pages 47–54. International Foundation for Autonomous Agents and Multiagent Systems, 2008.
5. Daniel Lélis Baggio. *Mastering OpenCV with practical computer vision projects.* Packt Publishing Ltd, 2012.
6. Paul Viola and Michael Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE, 2001.
7. Peter N. Belhumeur, João P Hespanha, and David J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on pattern analysis and machine intelligence*, 19(7):711–720, 1997.
8. Morgan Quigley, Ken Conley, Brian Gerkey, Josh Faust, Tully Foote, Jeremy Leibs, Rob Wheeler, and Andrew Y Ng. Ros: an open-source robot operating system. In *ICRA workshop on open source software*, volume 3, page 5. Kobe, 2009.
9. Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
10. Gary Bradski and Adrian Kaehler. *Learning OpenCV: Computer vision with the OpenCV library.* " O'Reilly Media, Inc.", 2008.
11. Ming-Kuei Hu. Visual pattern recognition by moment invariants. *IRE transactions on information theory*, 8(2):179–187, 1962.
12. Alan Black, Paul Taylor, and Richard Caley. The festival speech synthesis system. `http://digital.cs.usu.edu/~vkulyukin/vkweb/teaching/cs6890/festival.pdf`. Accessed: 2018-01-05.

Bill Hardware and Software Description

To provide completely autonomous operation, Bill owns two main modules of control: (i) the High-level control, which includes algorithms to solve functionalities such as global task planning, navigation and tracking, recognition of objects and faces, user-interaction, among others; and (ii) a low-level to control sensors and actuators in the real world.

Bill Hardware Description

Bill has a motion base that presents higher mobility. It is a round base with 2 differential drive wheels and 2 free wheels -one in the front and other in the rear for maintaining balance. All the electronic parts were carefully checked to avoid short-circuits and increase power. The details of each hardware are described as follows:

- **Base:** Two Arduino Mega 2560; Two motors IG32P 24VDC 190 RPM Gear Motor with Encoder; One Notebook Dell Inspiron 15R-4470, intel core i7; One digital buzzer; One RPLIDAR 360°laser scaner; Three Sabertooths controlers; One LM35 linear temperature Sensor; Three batterys 11.1 volts and 2800 mAh; One digital push button;
- **Torso**: Mini actuator Firgelli Automations; One Emergency switch;
- **Arm**: five Dynamixel-ax-12A; One ArbotiX-M; Maximum load: 1kg.
- **Hea**d: One Dynamixel-ax-12A; One Microsoft Kinect sensor; Two Microsoft life Cam HD-3000; One Rode Videomic Pro.
- **Tablet**: Motorola tablet.

Bill Software Description

The low level is composed of a proportional control running on arduino boards. The communication and high level system is composed of tools developed by our team and open source applications of the Robot Operating System (ROS). The software are:

- Navigation, localization and mapping: Hector mapping.
- Face recognition: OpenCV library.
- Speech recognition: PocketSphinx library; Boost library.
- Speech generation: Festival.
- Grapichal User Interface to control and start ROS Node: QT Creator.