C.E.S.A.R-VOXAR Labs Team: RoboCup@Home 2015

Henrique Foresti, Lucas Cavalcanti, Vitor Parente, Raphael Brito, Filipe Villa Verde, Diogo Lacerda, Gabriel Finch, Tiago Barros, Veronica Teichrieb, Ermano Arruda, Daniel Queiroz, João Marcelo Teixeira, João Paulo Lima, Lucas Figueiredo, Thiago Chaves, Leonardo Lima, Wayne Ribeiro, H.D. Mabuse, Dayane Kelly

C.E.S.A.R (Centro de Estudos e Sistemas Avançados do Recife)
Rua Bione, 220, 50030-390 – Cais do Apolo, Bairro do Recife – Recife, PE – Brazil
VOXAR Labs of CIn (Centro de Informática) at UFPE (Universidade Federal de Pernambuco)
Av. Jornalista Anibal Fernandes s/n, 50740-560 – Cidade Universitária, Recife, PE – Brazil

hbf@cesar.org.br, vt@cin.ufpe.br
http://www.i-zak.org/

Abstract. This document describes the RoboCup@Home team C.E.S.A.R-VOXAR Labs from a private institute of research, Recife Center for Advanced Studies and Systems (CESAR), and VOXAR Labs, a research center of Center of Informatics of the Federal University of Pernambuco (UFPE), Brazil, for the competition to be held in Hefei in 2015. Our team will present a robot for interaction with humans, designed and constructed through Emotive Robotics concepts.

1 Introduction

The CESAR-VOXAR Labs team was formed with the objective of researching and developing robots for interaction with humans, introducing the concept of Emotive Robotics. We are submitting our work to robotics competitions as a way to validate these concepts that we are developing in a real scenario.

This team achieved the 1st place of the @home category of the CBR/LARC (Brazilian and Latino America Robotics Competition) that took place in São Carlos / Brazil, in October 2014, with i-Zak, their first produced robot. i-Zak was designed and constructed with recyclable and reused materials, and powerful processing systems that provide an excellent interaction with humans.

In order to compete in the Robocup 2015 in Hefei, in addition to the redevelopment of various subsystems of i-Zak we are developing POP-E, a new robot based on the same concepts but using a more robust technology. POP-E has a Festo Robotino platform in its base. The twist has a lifting system that allows height adjustment of the head and robot arm. The new robot is an improvement on the first one and is being developed on the same architecture (see Fig. 1).

2 Hardware and Software Description

2.1 Hardware

Our robot, i-Zak, is composed of the following hardware parts: a base platform with two driverless VEX 2.75" Omni Directional Wheel and two driving 13 cm Colson wheels; an aluminum and acrylic body made of reused materials where there is a laptop running the Robot Operating System (ROS); and a head system, composed by an Align G800 gimbal where there is a Microsoft Surface tablet mimicking its face, a Microsoft Kinect for its vision, a RODE VideoMic GO directional microphone for voice recognition and a small speaker for human interaction (see Fig.1).

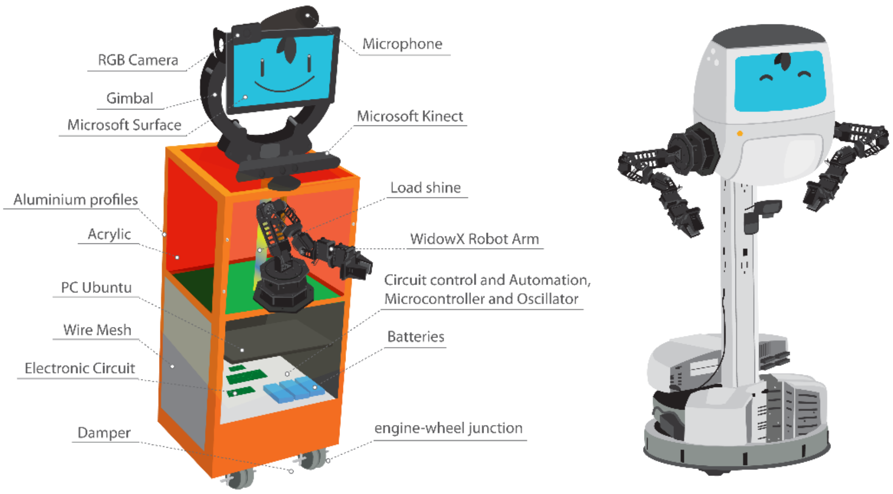

Fig. 1. i-Zak schematics (left), POP-E (right)

On the bottom of the robot there is an Arduino board connected to a PCB, by which it controls two Mabuchi DC motors and the robot's neck, as well as gathering information from the encoder circuits with infrared barrier sensors, which operate at coded discs made in a 3D printer and coupled to the robot's wheels. This same PCB is also connected to the batteries, and supplies power to several units. Connected to these batteries there are also the emergency button and the oscillator circuit responsible for the heartbeat. The Arduino receives data from ROS application through the arduinoserial lib implemented. The same notebook with ROS is also responsible for the robot audio output, delivered by one small speaker, the audio input, gathered by the RODE VideoMic GO direction microphone, and several DSP techniques implemented in the middleware layer on the PC.

On the top of the robot there is an Align G800 gimbal with three degrees of freedom (3 DOF), which emulates head movements and is controlled by the Arduino, that sends angular orientations to the servos. There is a MS Surface running the VOXAR brain, responsible for all the vision system that processes visual data from MS Kinect camera and depth sensor. The MS Surface is connected to the notebook by an ethernet cable in order to send commands based on visual processed data and receive some feedback about the robot's position and its motion via an encoder on each powered wheel for steering control.

For gripping and manipulation, we use a WidowX Robot Arm Mark II, manufactured by Trossen Robotics. It has 5 degrees of freedom (DOF) and horizontal reach of 41 cm at full stretch, holding up to 400 g for a vertical grip orientation and 29 cm reach. It has a boxed ABS frame, two MX-64 Dynamixel servos for shoulder movement and rotation, two MX-28 Dynamixel servos for elbow and wrist movement, and two AX-12 Dynamixel servos for wrist rotation and parallel gripping, opening up to 32 mm, all controlled by an ArbotiX-M Robocontroller. A new grip is being developed using AX-18 Dynamixel servos and can open up to 55 mm.

2.2 Software

Our system architecture consists of three layers: the first, written in C++, runs on MS Windows and is called VOXAR Brain. This part of the robot system processes all the visual and depth data that comes from the Kinect Camera and its depth sensor. Based on this information, the VOXAR Brain sends data to the second level of the system, packets that contain data for the robot movements and its activities. In Fig. 2 we detail the architecture.

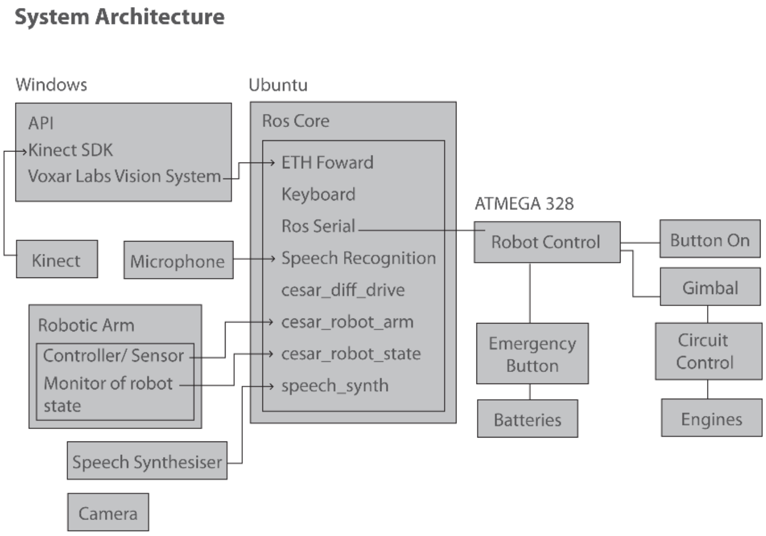

Fig. 2. i-Zak architecture

In the second layer of the robot system we have a machine with Linux running ROS (Robot Operating System). This part of our architecture has voice recognition software, which is the most important data source (along with gestures) to change the robot activities. It also sends and receives data packets from the VOXAR Brain to monitor the robot status, helping the first layer of the system choose exactly what is necessary to process. Via the vision technology processing, ROS will receive data packets with commands to control the movements.

Finally, ROS needs to pass the movements to the actuators, and it is here where our third layer enters the system. The electromechanical system is based on various controllers, including Arduino UNO and Arbotix that communicate via USB serial with ROS. The various controllers each receive a packet, and from there they compute the information that is necessary to control the actuator drives. Note that we have bidirectional communication between the VOXAR Brain and ROS, and between ROS and the hardware controllers.

ROS.

ROS (Robot Operating System) is a collaboratively developed open source framework for controlling robots. It is designed to be highly modular, consisting of small components ("nodes") which send messages to each other.

At CESAR we have expertise in customizing ROS for individual robot designs. Included in this are such areas as hardware abstraction, low level driver control, sensors/feedback and message sending.

By leveraging ready built components we are able to rapidly adapt ROS to suit every type of robot that needs to be built.

Speech Recognition.

At CESAR Robotics we are engaged in research in speech recognition. Based on the Sphinx software from Carnegie Mellon University [10], our applications are making strides in the areas of audio processing and understanding of natural language. Our current research includes measures to improve voice recognition in noisy environments, an area which is known to be problematic in robotics applications.

Speech Synthesis.

Another area of research in the Robotics department is speech synthesis. The goal is to produce speech which sounds as natural and human-like as possible. Our current line of research is based on several applications, including the "Mary" Text to Speech System developed by Saarland University in Germany [11].

Machine Learning.

In this department we have expertise in machine learning, mainly in the areas of genetic/evolutionary algorithms, sensor feedback, artificial intelligence and artificial neural networks [12].

2.3 Licenses

The robot uses ROS as the main software control. These modules are built using the BSD license. Some of the hardware controller libraries are implemented using the LGPL license.

3 Interaction

The robot was conceived to deal with functional and emotional needs of people. This is what has been referred to as "Emotive Robotics". In order to accomplish that, it is necessary for the the robot to understand its surroundings and interact according to its perceptions, creating a natural and comprehensive relationship with humans.

The robot has been designed to be useful to any person at its home. Its functionalities represent the ones of a domestic robot focusing the following concepts: companionship, productivity, health and entertainment.

Therefore, the robot's user can use it for eleven different tasks, all related to one or more of these key-concepts, namely:

- **reading** – the robot recognizes the newspaper sheet exhibited in front of it and reads it. The reading is performed accessing a local database with the corresponding mp3 files to each prerecorded page;
- **take care of the elderly** – the robot detects that the user has fallen, approaches him asking if everything is ok, and if the user doesn't respond affirmatively in the next 5 seconds the robot warns him that it is sending an alert message to a predefined cell-phone, which it does in sequence;
- **personal trainer** – the robot guides the user through a set of exercises and then detects his heartbeat, warning him if there is any abnormality;
- **nutritionist** – the user shows a known object to the robot, which exhibits nutritional information of the product on its display;
- **personal stylist** – the robot recognizes the user skeleton and dresses him with clothes and accessories selected by him;
- **objects recognition** – the robot localizes objects in a room not visible from its initial position, moves towards them and recognizes each object;
- **object grasping** – the robot recognizes an object and grasps it with its arm/hand;
- **identify people** – the robot detects the user's face and if it is a known person, greets him saying his name, and if not gives the user a fantasy name;
- **follow one person** – the robot recognizes and memorizes the person to be followed and follows this person until the end of the trajectory without bumping into anyone else;
- **musical companion** – the user performs an "air guitar" gesture and the robot plays the corresponding musical note;
- **dancing** – the robot recognizes a specific gesture performed by the user indicating that he is dancing and moves as if it was dancing together with the user.

Finally, when not performing a specific task, the robot shows a generic behavior, always looking to the closest user or the one that is talking louder. If nobody is close to the robot or no noise is present, the robot moves its head to arbitrary positions.

In order for the robot to be able to perform such tasks, it was necessary to research and develop state-of-the-art technologies in the areas of computer vision and human-robot interaction. These technologies are described in sequence.

3.1 Gesture recognition

Understand body movement performed by a person, in a way that the interaction with the robot may occur in a natural manner [1].

VOXAR Labs has been studying techniques for body gestures recognition, as well as the application of such techniques in areas such as motor rehabilitation. In 2012, the research group was awarded with two prizes for the reAIRbilitation game application: First Place of the Games For Change Award and Best Game Award in the Other Platforms Category of the 11th Indie Games Festival 2012 of the XI Brazilian Symposium on Computer Games and Digital Entertainment – SBGames 2012.

3.2 Object recognition

Make the robot capable of understanding its environment by detecting and recognizing several types of objects [2].

3.3 Object tracking

Beyond detect and recognize, it is possible to follow the movement of objects, locating them spatially in the environment [3].

VOXAR Labs has a long-term experience in 2D and 3D tracking research, being awarded in 2014 with the Third Place Prize of the Volkswagen/ISMAR Tracking Challenge 2014 - Accurate Tracking scenario.

3.4 People tracking

Follow someone's movement, tracking them in the environment.

3.5 Grasping by object pose estimation

The first problem tackled by this work is to find models that can be used for reasoning about incomplete or uncertain models of object shape. This is done by combining and extending works seen in [4][5][6][7]. Experiments are being made on several different representations for uncertainty and incompleteness. These representations are mainly based on particle filters [5] for positional uncertainty, but it is also possible to model semi-dense depth clouds with associated uncertainty values [6][7]. These models are inherently incomplete representations of shape, but policies may be applied to gather required information according to given manipulation task requirements [8], such as the ones from the Pick & Place task [4].

3.6 Robot autonomous navigation

In the system presented, the estimated localization of the robot is given by the ROS implementation of the AMCL algorithm. The navigation node used in the system is also the navigation stack that comes with the ROS installation given by the PRM algorithm [9].

4 Research and Innovation

4.1 Physical Structure

During the robot's building, we decided to use recyclable and reused materials to reduce the environmental impacts and costs.

The robot base is made from a standard computer ATX case and aluminum structure with this configuration: the two Colson driving wheels at the rear of the base and in two VEX omnidirectional wheels in front, to stabilize the structure. To power the robot driving wheels we adapted two Mabuchi DC motors used in car electric windows. We used this type of motor because of its cost and torque.

The transmission and suspension was all designed and built by our team. We first prototyped the transmission part making a 3D model and printing in a 3D printer to make some initial tests to figure out how the wheel and the motor would work together. After these tests we built the part in nylon using a lathe because nylon is stronger than the material (PLA) we used in the 3D printed prototype. The suspension structure was made with aluminum profiles because of its flexibility, so we could model it as we wanted without any special tools, and ease to obtain.

The body structure of the robot was built with aluminum profiles and acrylic. The aluminum gives the strength and the acrylic was used to fill the empty places and to give decorative look.

An oscillator circuit blinks two LED ribbons that emulate the robot's heart, and whose luminosity and oscillating frequency are dependent on the batteries' charge.

4.2 Emotive Robotics

Field research practices, adopted by user-centric design, are variations on the ethnographic research which originated in anthropology. Unlike the original academic versions, which were based on extended immersion in the field, research with focus on design for innovation is done through fast and intensive field visits.

Quantitative methods, in a general sense, help to reveal barriers and opportunities in social, economic, political and cultural dimensions, analyzing and mapping the dynamic relations between people, places, objects and institutions. It is important to highlight that quantitative methods do not have any statistical intention, so there is no objective to measure behaviors or attitudes.

Taking into consideration that we have passed from a time when the robots work exclusively isolated from humans, within production lines, and have arrived at a time of living together in domestic environments, Emotive Robotics was born of the necessity in the contemporary world for articulate robots that fulfill the needs of people. Those needs come from the functional and emotional requirements. In order to meet this complexity, the paradigm of design thinking was adopted:

“Design thinking relies on our ability to be intuitive, to recognize patterns, to construct ideas that have emotional meaning as well as functionality, to express ourselves in media other than words or symbols.” - Brown, Tim. “Change by Design.”

Therefore, emotional robots are presented as complex solutions, articulated by an interdisciplinary team to meet the functional and emotional demands of people living in a broad and sensible world.

5 Conclusion

Above we described our current robot that is a result of our team work that already achieved 1st place in the CBR (Brazilian Robotics Competition) and we want to compete in RoboCup@Home 2015 to continue our efforts and share experience with the other teams.

References

1. Alana da Gama, Veronica Teichrieb, Pascal Fallavollita, Nassir Navab. Motor rehabilitation using Kinect: a systematic review. Games for Health Journal. (accepted for publication)
2. Lucas Figueiredo, Edvar Vilar Neto, Ermano Arruda, João Marcelo Teixeira, Veronica Teichrieb. Fishtank Everywhere: Improving Viewing Experience Over 3D Content. Third International Conference, DUXU 2014, Held as Part of HCI International 2014, 2014. p. 560-571.
3. João Paulo Lima, João Marcelo Teixeira, Veronica Teichrieb. AR jigsaw puzzle with RGB-D based detection of texture-less pieces. IEEE Virtual Reality, 2014. p. 177-178.
4. J. Nunez-Varela, J. Wyatt, Where Do I Look Now? Gaze Allocation During Visually Guided Manipulation, 2012 IEEE International Conference on Robotics and Automation.
5. Thrun, S., Bugard, W., and Fox, D. 2008. Probabilistic Robotics. MIT Press Cambridge, Cambridge, MA.
6. J. Engel, T. Schöps, D. Cremers, LSD-SLAM: Large-Scale Direct Monocular SLAM, In European Conference on Computer Vision (ECCV), 2014.
7. J. Engel, Jurgen Sturm, D. Cremers, Semi-Dense Visual Odometry for a Monocular Camera, In European Conference on Computer Vision (ECCV), 2013.
8. Rajesh Rao, Decision making under uncertainty: A neural model based on POMDPs (Frontiers in Computational Neuroscience, 2010)
9. Roland Geraerts, Mark H. Overmars. A Comparative Study of Probabilistic Roadmap Planners. Workshop on the Algorithmic Foundations of Robotics, 2002. p. 43-57.
10. http://www.speech.cs.cmu.edu/
11. http://mary.dfki.de/
12. http://br.linkedin.com/pub/gabriel-finch/11/1/1b3