Emotive Robotics with I-Zak

Henrique Foresti, Gabriel Finch, Lucas Cavalcanti, Felipe Alves, Diogo Lacerda, Raphael Brito, Filipe Villa Verde, Tiago Barros, Felipe Freitas, Leonardo Lima, Wayne Ribeiro, H.D. Mabuse, and Veronica Teichrieb and João Marcelo Teixeira

C.E.S.A.R (Centro de Estudos e Sistemas Avanados do Recife),
Rua Bione, 220, 50030-390 - Cais do Apolo, Bairro do Recife - Recife, PE - Brazil
VOXAR Labs at CIn–UFPE (Universidade Federal de Pernambuco),
Av. Jornalista Anibal Fernandes s/n, 50740-560 - Cidade Universitria, Recife, PE - Brazil
{henrique.foresti,gabriel.finch,lucas.cavalcanti,felipe.alves, diogo.lacerda,raphael.brito,filipe.villa,tiago.barros, felipe.freitas,leonardo.lima,wayne.ribeiro,mabuse}@cesar.org.br
{vt,jmxnt}@cin.br
http://www.i-zak.org/

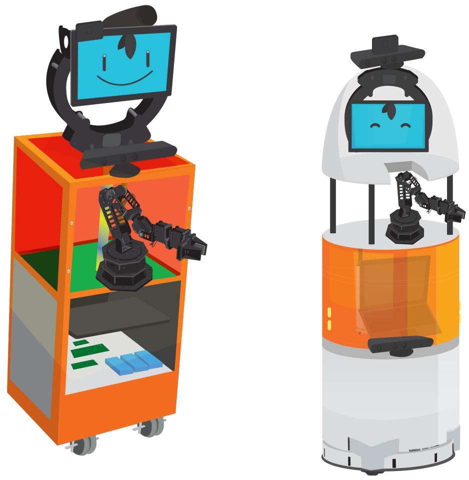
Abstract. This document describes the efforts of the RoboCup @Home team C.E.S.A.R.-VOXAR labs from a private institute of research, Recife Center for Advanced Studies and Systems (CESAR), and VOXAR Labs, a research center of Center of Informatics of the Federal University of Pernambuco (UFPE), Brazil, for the competition to be held in Leipzig in 2016. Our team will present a robot for interaction with humans, designed and constructed through Emotive Robotics concepts.

1 Introduction

The C.E.S.A.R. - 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 1st place in the @home category two times at the CBR / LARC (Brazilian and Latino America Robotics Competition), one in São Carlos / Brazil, in October 2014, with i-Zak 1.0, their first produced robot and another in Uberlândia / Brazil, in November 2015, with a new version of I-Zak.

Our first I-Zak was designed and constructed with recyclable and reused materials, and powerful processing systems that provide excellent interaction with humans. In order to compete in the Robocup 2016 in Leipzig, in addition to the redevelopment of various subsystems of the first robot we are developing a new version of I-Zak, a new robot based on the same concepts but using a more robust technology. I-Zak 2.0 has an Arlo Parallax platform in its base. The twist has a lifting system that allows height adjustment of the robot arm. The new robot is an improvement on the first one and is being developed on a similar architecture (see Fig. 1)



(a) I-Zak 1.0 (b) I-Zak 2.0

Fig. 1. I-Zak versions.

2 Hardware and Software Description

2.1 Hardware

The robot used in LARC - RoboCup Latin America 2014 was constructed using recyclable materials, to reduce the robot's cost. This year in order to have a high quality robot, we designed a robot with special constructed parts.

I-Zak appears as a cylinder with 47 centimeters of diameter, and 1.30 meters high and weighs about 25 kilograms, built upon an Arlo Robot Base. We also use acrylic parts, carbon fiber bars, WidowX robot arm, linear actuator and 3D printed parts. The Arlo Robot Base is responsible for the robot movement and is composed of 2 mabuchi motors with pneumatic rubber tires of 76 mm of diameter. The carbon fiber bars are used to support columns that are fixed on the Arlo Robot Base with specially designed 3D parts. In the middle of I-Zak, we use acrylic which was cut by a laser cutter in a circular shape, with 47cm diameter connected to the Arlo base with 3D printed parts, designed on 3D CAD software, and carbon fiber bars.

The robot head is made from a cone lampshade cut in half. Inside the lampshade is placed a gimbal to support the tablet with the robot face, and on top there is a 3D designed part to accomodate the Microsoft Kinect and microphone. The tablet simulates the face of I-Zak and the Gimbal is the neck, responsible for lateral movement of the tablet.

For gripping and manipulation, is used a WidowX Mark II Robot Arm, manufactured by Trossen Robotics. It has 5 degrees of freedom (DOF) and a horizontal reach of 41 cm fully extended, holding up to 400 g for a vertical grip orientation and 29 cm reach. It has a boxed ABS frame, two Dynamixel MX64 servos for shoulder movement and rotation, two Dynamixel MX28 servos for elbow and wrist movement, and two Dynamixel AX12 servos for wrist rotation and parallel gripping, opening up to 32 mm, all controlled by an ArbotiX-M Robocontroller. A new gripper is in development using AX18 Dynamixel servos which can open up to 55 mm.

2.2 Software

The 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 most of the visual and depth data that comes from the Kinect Camera on the top of the robot 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 activities. In Fig. 2 we detail the architecture.

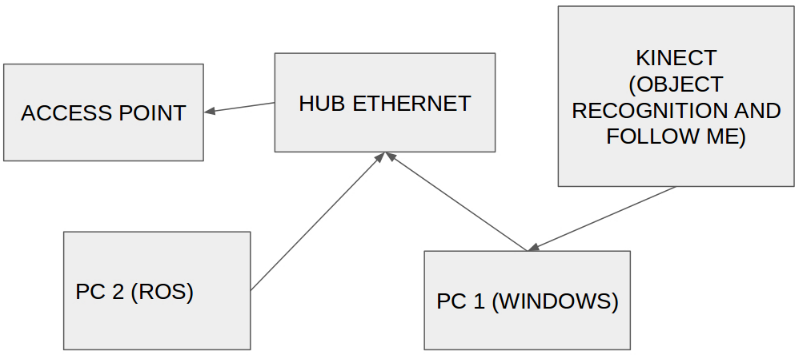


Fig. 2. PC WINDOWS - VOXAR Architecture

In the second layer of the robot system we have a machine with Linux running ROS (Robot Operating System). This part of I-Zak's 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 some features. See more on Fig. 3.

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, Parallax controler and Arbotix that communicate via USB serial with ROS. The various controllers each receive

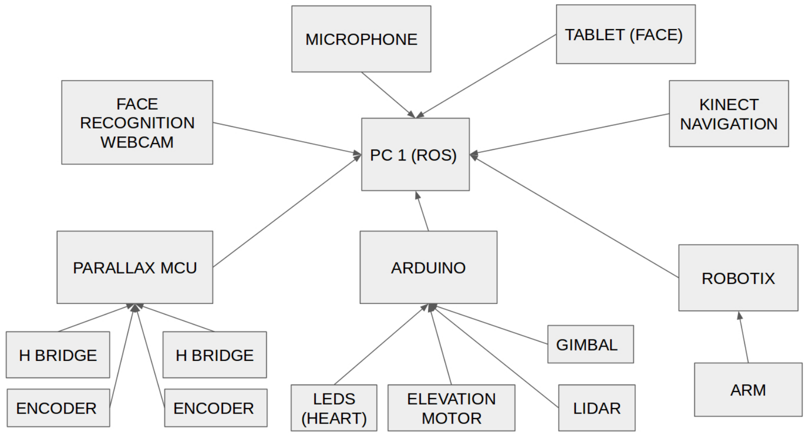


Fig. 3. PC LINUX - ROS Architecture

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. We are also developing a two level matching scheme, the first level converts sound to text, the second level then uses a phonetic representation to match the recognised text with a set of context sensitive target values. As part of our speech recognition stack, we are using IBM Watson. This is an optional element, which provides enhanced recognition capabilities for words which are not in the the robot's predefined vocabulary. Watson functions as an external computing resource: audio input is sent to the IBM BlueMix server, and if speech is recognised then the result is returned as text. The sending and receiving is carried out via a nodejs script. A node within our ROS stack is responsible for making a connection to the nodejs script and converting audio and text to and from ROS topics.

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 features represent the ones of a domestic robot focusing the following concepts: companionship, productivity, health and entertainment.

Therefore, the robots 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;
– taking 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 cellphone, 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;

- object 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 users 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 beats and starts to dance according to the music rhythm.

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 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 someones movement, tracking them in the environment. This is done via point cloud analysis to detect human forms. Some logic is implemented to re target the same individual when another passes quickly in front of the original. It is possible to track several individuals simultaneously, provided they are sufficiently separated so as to be distinguishable from each other.

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 integration of RtabMap and Ros Navigation Stack [9], which includes real time mapping and route planning, made by a point cloud generated with the Kinect mixing the camera image and its depth sensor. In addition to the Kinect sensor there is a Parallax controller that publishes the odometry and controls the wheel velocity, helping the robot.

4 Research and Innovation

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 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.

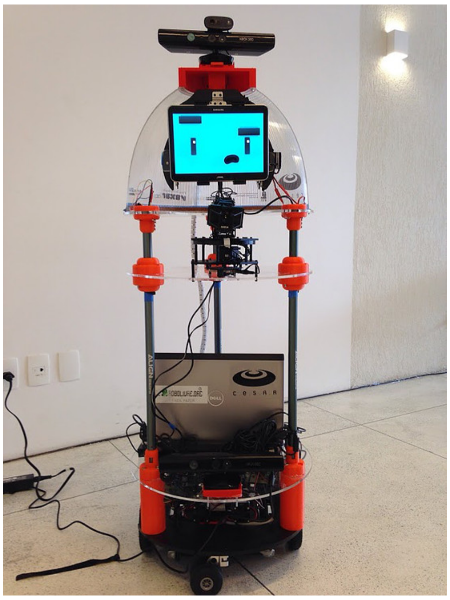
5 Conclusion

Above we described our current robot that is a result of our team work that already achieved 1st place twice in the CBR (Brazilian Robotics Competition) and we want to compete in RoboCup@Home 2016 to continue our efforts and share our experience with the other teams.

References

1. Da Gama, A., Fallavollita, P., Teichrieb, V., Navab, N.: Motor rehabilitation using Kinect: A systematic review. Games for health journal, 4(2), 123–135. (2015)
2. Lucas Figueiredo, Edvar Vilar Neto, Ermano Arruda, Joo Marcelo Teixeira, Veronica Teichrieb. Fishtank Everywhere: Improving Viewing Experience Over 3D Content. Third International Conference, DUXU 2014, Held as Part of HCI International 2014, p. 560–571. (2014)
3. Joo Paulo Lima, Joo Marcelo Teixeira, Veronica Teichrieb. AR jigsaw puzzle with RGB-D based detection of texture-less pieces. IEEE Virtual Reality, p. 177–178. (2014)
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. Thrun, S., Bugard, W., and Fox, D. Probabilistic Robotics. MIT Press Cambridge, Cambridge, MA. (2008)
5. Thrun, S., Bugard, W., and Fox, D. Probabilistic Robotics. MIT Press Cambridge, Cambridge, MA. (2008)
6. J. Engel, T. Schps, 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. p. 43-57. (2002)
10. Speech at CMU (Carnegie Mellon University), `http://www.speech.cs.cmu.edu/`
11. The MARY Text-to-Speech System (MaryTTS), `http://mary.dfki.de/`
12. Gabriel Finch (LinkedIn), `http://br.linkedin.com/pub/gabriel-finch/11/1/1b3`

– **Name of the Team:** C.E.S.A.R-VOXAR Labs Team
– **Contact Information:**
Henrique Foresti
{henrique.foresti,biorobotics-l}@cesar.org.br
– **Website:** http://www.i-zak.org/



– **Team Members:**
  - Henrique Foresti
  - Gabriel Finch
  - Lucas Cavalcanti
  - Felipe Alves
  - Diogo Lacerda
  - Raphael Brito
  - Filipe Villa Verde
  - Tiago Barros
  - Felipe Freitas
  - Leonardo Lima
  - Wayne Ribeiro
  - H.D. Mabuse
  - Veronica Teichrieb
  - João Marcelo Teixeira

– **Hardware:**
  - Kinect (2)
  - WidowX Mark II (1)
  - Arlo Robot Base (1)
  - Align G800 Gimbal (1)
  - Dynamixel MX64 (2)
  - Dynamixel MX28 (2)
  - Dynamixel AX12 (2)
  - ArbotiX-M Robocontroller (1)
  - Dynamixel AX18 (2)
  - Mabuchi DC motor (2)
  - 76mm Wheel (2)
  - Rode VideoMic Go (1)
  - Arduino UNO (1)

– **Software:**
  - VOXAR Brain
  - ROS
  - IBM Watson
  - IBM Bluemix
  - Node.js