Designing and Implementing an Interactive Social Robot from Off-the-shelf Components

Zheng-Hua Tan, Nicolai Bæk Thomsen and Xiaodong Duan

Department of Electronic Systems, Aalborg University, Denmark
e-mail: {zt,nit,xd}@es.aau.dk

Abstract. In this paper we present the design and development of the social robot called iSocioBot, which is designed to achieve long-term interaction between humans and robots in a social context. iSocioBot is 149cm tall and the mechanical body is built on top of the TurtleBot platform and designed to make people feel comfortable in its presence. All electrical components are standard off-the-shelf commercial products making a replication possible. Furthermore, the software is based on Robot Operating Software (ROS) and is made freely available. We present our experience with the design and discuss possible improvements.

Key words: Social robot, human-robot interaction, multimodal sensing, robot design

1 Introduction

Social robots have in recent years gained significant attention due to the advances in enabling technologies that make many real-world applications feasible. These applications range from social care, entertainment to education, just to name a few. According to the International Journal of Social Robotics, social robots are "robots able to interact and communicate among themselves, with humans and with the environment within the social and cultural structure in which they function" [10]. From this definition, it is obvious that the capability of interaction and communication is central for a social robot to have. Furthermore, the interaction should be natural, e.g., through speech and vision, in order to fit itself into a social environment.

Social robots are used for interacting with humans and improving humans well-being, not necessary with a goal of achieving well-defined, measurable tasks as industrial or even service robots do. With this philosophy in mind, we aim at developing robots that are suitable for various social setups with a proper height, able to move at human walking pace and rotate in place, and are a humanoid.

This work is supported by the Danish Council for Independent Research - Technology and Production Sciences under grant number: 1335-00162 (iSocioBot)

There are existing social robots, both commercial and proprietary. Nao[1] is one of the most widely used commercial social robots (e.g., [6, 1]). It has a number of different kinds of sensors and many degrees of freedom. On the other hand, it is small, 58 cm height, and moves slowly. PR2[2] robot is an advanced human size robot, which is rather expensive and thus in many ways not economically feasible or necessary. Furthermore, with a bulky body and strong arms, PR2 is more a service robot. Then there are proprietary robots that are developed by companies or research insititutes and not for sale, which, among others, include Robovie [4], Maggie [8], and Rubi [3]. There is also the iCub [7], which is a fully open-source child-like robot, however it is designed specifically for interacting with children and very expensive and unnecessarily complex for this project. Due to the lack of existing robots able to fulfil our requirements, we turned to build our own social robot called iSocioBot using off-the-shelf components and making it open source.

The primary objective of developing iSocioBot is to make a social robot being able to establish durable relationship with humans through natural interaction[3]. This relies on developing avanced capabilities to sense and express. The sensing capability is enhanced by a fusion concept we call *reinforcement fusion*, combining sensor signals in an interactive way: e.g., when a robot detects a sound direction, it turns towards the direction to see better and moves towards it to hear better. Reinforcement fusion is analogous to reinforcement learning [5], a known term in machine learning. The implication on the robot is that it should be able to move and rotate at a fast speed, e.g., that close to humans, and have a height suitable for sesing humans while they stand or sit. To maintain long-term relationship with humans, the social robot should have a model of its user and respond socially. We propose a concept of *social behaviour entrainment* to adapt behaviours to the user's. Entrainment is an interesting phonomenon found in human-human communication, e.g., speech entrainment occurs during spoken interaction where dialogue partners tend to become similar to each other in style [2]. In expressing itself, the robot will rely on both verbal and non-verbal communications. This also highlights the importance of facial expressions and overall appearance including a comfortable height. In this paper, we present the design of iSocioBot in terms of appearance, mechanics and software, and furthermore discuss some experience gained from deploying the robot in real and uncontrolled environments.

2 iSocioBot

For a robot to engange in natural and efficient interaction with persons, it needs to be able to sense what the persons are doing and be able to express itself in a similar way as they do. This requires integrating various sensors and actuators into a

[1] `https://www.aldebaran.com/en/humanoid-robot/nao-robot`
[2] `https://www.willowgarage.com/pages/pr2/overview`
[3] `http://socialrobot.dk`

single system. The physical appearance and construction along with considerations regarding choise of sensors are described in this section.

2.1 Mechanical Design

The preliminary design of the robot body is seen in Fig. 1 including specification of size. We have decided on a humanoid robot with a soft appearance because it is to interact with humans for long periods in a social context, thus it cannot resemble an industrial robot. The height of 1.49m is chosen to make users feel comfortable both while standing and sitting. In addition, it is also good for camera view on the top of the robot in terms of face detection and recognition. Figure 2 shows the actual

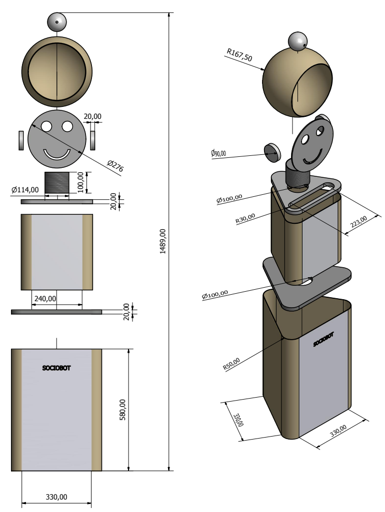

Fig. 1 The mechanical design of iSocioBot.

implementation of the design. Due to the payload capacity of the base, we needed to limit the weight of the robot body, hence plastic and light wood materials were used to build the body of the robot. Acoustic clothes was used to cover the robot to make it more friendly to users, less mechanical looking and reduce the weight. The wide crack in the body of the robot below the face in Fig. 1 was originally intended for a microphone array, however after having chosen the acoustic clothes, the microphone array was placed completely hidden inside the body with satisfactory performance. This was also the case with the loudspeaker. We built the robot head using LED array and strips, such that the robot can show different types of expressions and even animations regarding the user's varying activities.

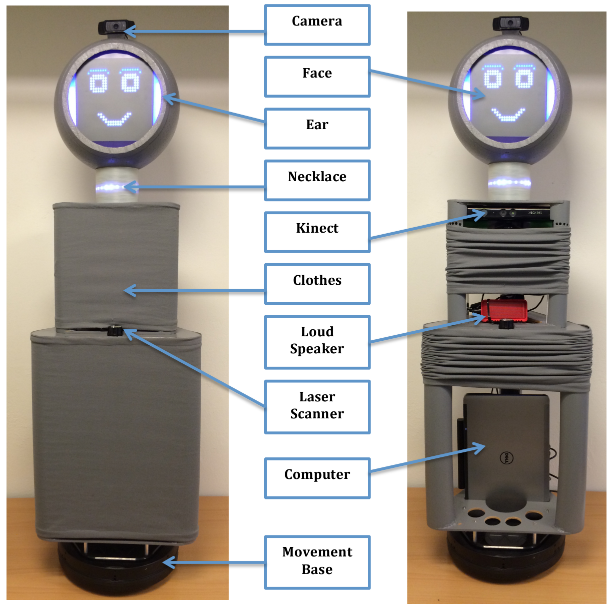

Fig. 2 The iSocioBot.

2.2 Electronic Design

iSocioBot is self-contained since all the electronic parts are inside its body and run on battery. The electronic parts are commercial off-the-shelf components making it easy to replicate the iSocioBot. The electronic system of our iSocioBot consists of computer, movement base, sensors, feedback module and power module. The details of all electronic parts can be found in Table 1.

- **Computer:** We use a DELL E6540 laptop to process the information from every sensor and control the robot. All the other electronic parts are connected to this computer through USB port.
- **Movement base:** The ability to move is achieved through a Kobuki low-cost mobile research base which is a part of the TurtleBot[4].
- **Sensors:** Four types of sensors are used to gather information for the robot; Microsoft Kinect, Singstar microphone, camera and laser scanner. The Kinect, which is hidden by the acoustic clothes due to design consideration, is only used as microphone array. Singstar microphone is a wireless handhold microphone, which is used for close-talking speech recognition. To gather visual information, we install a High-Definition camera Logitech HD Pro Webcam C920 on the top of our iSocioBot. For localization purpose, a laser scanner, Hokuyo URG-04LX-UG01, is equipped in the middle of the robot body.

[4] `http://www.turtlebot.com`

Part	Model	Note
Computer	Dell E6540	Intel Core2 i7-4800MQ Processor, 16GB RAM and 256GB SSD hard-drive
Microphone array	Microsoft Kinect	
Wireless microphone	Singstar	
Camera	Logitech C920	Full HD 1080p recording
Laser scanner	Hokuyo URG-04LX-UG01	Range (5600mm $\times$ 240°), Accuracy ($\pm$30mm)
Loudspeaker	JABRA SOLEMATE MINI	
Head	Custom	Arduino (Mega 2560) for control, one 32×32 RGB matrix LED panel for face, and three Pololu RGB LED strips
Movement base	Kobuki	Maximum translational velocity: 70cm/s, Maximum rotational velocity: 180deg/s, Payload: 5kg (hard floor), 4kg (carpet)

Table 1 Electronic components used for iSocioBot.

- **Feedback:** There are two parts for the feedback; loud speaker and head. The loud speaker is used for speech synthesis, which is also installed behind the acoustic clothes. Head is composed by face, necklace and ears, which are used to indicate the robot is listening, thinking and talking and also show different facial expressions. The hardware of head consists of Arduino (Mega 2560) for control, one 32×32 RGB matrix LED panel for face, and three Pololu RGB LED strips for two ears and one necklace. The robot can display different facial expressions based on the results from the speech recognition and person identification modules, according to predefined scripts or potentially using data-driven approaches [9]. Figure 3 shows some facial expressions that can be displayed by this system. Furthermore, by switching between two facial expressions at a low frequency, we are able to make the robot appear to be talking, listening etc. This makes the interaction with people more natural and dynamic.

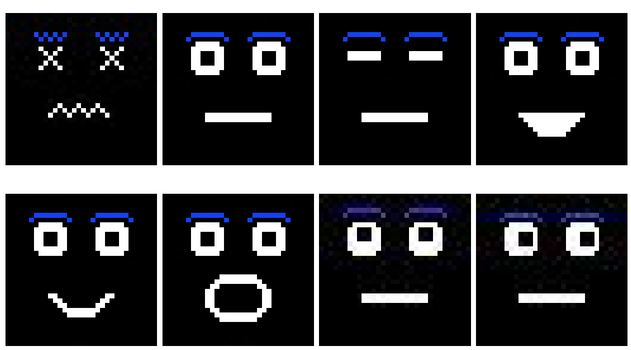

Fig. 3 Facial expression samples of the robot.

- **Power module:** The base and Kinect are powered by the battery inside the base. The head is powered by an extra battery while the speaker and laptop are powered by their built-in battery separately.

3 Software

The iSocioBot software is developed to run on The Robot Operating System (ROS) [5], which runs on top of the standard OS (e.g., Ubuntu). Having a robot which runs ROS offers many advantages such as:

- ROS offers a great framework for each module/function to communicate (e.g., fusing speaker recognition with face recognition) in real-time.
- ROS is widely used thus much high-quality software is available. This makes it easier and faster to integrate various sensors (e.g., laser scanner, camera, arm etc.) without much effort.
- ROS is open-source which means that everyone can use it for free, thus making our research as reproducible as possible .
- ROS software can be written in Python or C++, thus allowing for fast development of code using Python and computationally efficient code using C++.

The high-level structure of the iSocioBot code framework is shown in Fig. 4. All decisions and actions are made in *Decision Making* based on the input from all other modules. All modules except for *Decision Making* are written in a generic way in terms of input-output, such that only *Decision Making* needs to be re-written if a new application is desired. This also has the benefit of allowing for easy substitution of a module without having to change in other modules, e.g., if one wanted to test different face detection algorithms. It is noted that the iSocioBot code partly relies on existing code and external librabries such as OpenCV, ALIZE, and partly on our own work [12], [11].

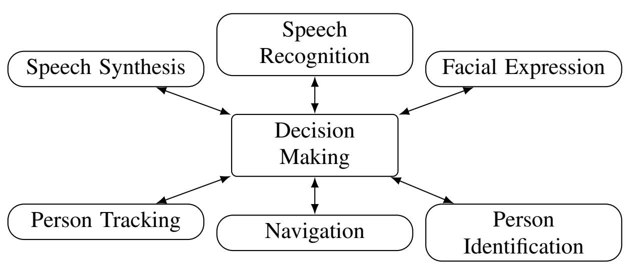

Fig. 4 Illustration of the iSocioBot software. Arrows indicate communication.

[5] http://www.ros.org/

4 Evaluation

4.1 Public Demonstrations

iSocioBot has been an attraction at three different public events where it has demonstrated different applications. At the official opening of ”Research Days of Denmark” in Aalborg and the opening of ”Safe 7” in Nibe, the robot took part in on-stage scripted dialogues with the presenters and also performed a Q/A session, where it was supposed to answer a closed set of question and turning towards the speaker. iSocioBot also participated in ”Culture Night” at the Danish Ministry of Higher Education and Science, where people were asked to take part in a dialogue with the robot and afterwards answer a questionaire about the behaviour and appearance of iSocioBot. A total of 97 persons (mainly kids at the age of 6-12 years) participated and one conclusion was, that people liked the appearance of iSocioBot compared to robots with a more mechanical appearance.

4.2 Our Own Experience

Based on the participation at the above mentioned events and running tests, we have gained some valuable experience about what should be improved for the next version of iSocioBot.

- **Choice of base:** The base of the robot has turned out to be a weak link in several ways; it can only carry a payload of 5kg which is too little when using an onboard laptop, and it is furthermore very sensitive to weight distribution resulting in unequal traction at the two active wheels which degrades the navigation and in pratice makes it unable to drive around freely, thus limiting it to rotation in-place. The weight distribution on the current design combined with the current base also forces the robot to drive over bumps etc. at very low speeds in order not to tip over.
- **Power Consumption:** When all software modules are running the laptop is only capable of running for $\approx$ 1 hour with a 9-cell 97 WH battery, which is a limiting factor for many applications. It should be noted that no effort has been made to optimizing the code or the choice of sensors with respect to power consumption.

4.3 Future Improvements

The issues regarding the base can be solved by either redistributing the weight or simply replacing it with a more powerfull base such as Pioneer 3DX[6]. The first op-

[6] `http://www.mobilerobots.com/ResearchRobots/PioneerP3DX.aspx`

tion can be achieved by making iSocioBot smaller, however this will degrade the appearance and thus the human-robot interaction, which is not desirable. The second option is more attractive from a mechanical point of view, however the price for good alternatives are considerably higher than for the current base.
One way of minimizing the power consumption is to shift all the heavy computational tasks from the on-board laptop to a server. This will also lower the weight of the robot, since only a less powerfull computer is needed to handle communication with the server, the sensors and the actuators. This solution comes at the expense of being dependent on having a WiFi available. Another drawback of this solution is the introduction of an increase in the complexity of developing software for iSocioBot.

5 Conclusions

This work has presented the first iteration of designing and implementing the social robot called iSocioBot. Based on the implementation some shortcomings have been identified and possible solutions have been presented. One direction of future work is to make a second iteration to correct these shortcomings. Other directions of research will focus more on user modelling and the application of iSocioBot.

Acknowledgements The authors would like to thank Ben Krøyer and Peter Boie Jensen for constructing the iSocioBot and Trine Skjødt Axelgaard for making the mechanical design of the iSocioBot.

References

1. Baddoura, R., Venture, G.: Social vs. useful hri: Experiencing the familiar, perceiving the robot as a sociable partner and responding to its actions. International Journal of Social Robotics **5**(4), 529–547 (2013)
2. Beňuš, Š.: Social aspects of entrainment in spoken interaction. Cognitive Computation **6**(4), 802–813 (2014)
3. Fortenberry, B., Chenu, J., Movellan, J.: Rubi: A robotic platform for real-time social interaction. In: Proceedings of the International Conference on Development and Learning (ICDL04), The Salk Institute, San Diego (2004)
4. Kanda, T., Ishiguro, H., Imai, M., Ono, T.: Development and evaluation of interactive humanoid robots. Proceedings of the IEEE **92**(11), 1839–1850 (2004)
5. Kober, J., Peters, J.: Reinforcement learning in robotics: A survey. In: Reinforcement Learning, pp. 579–610. Springer (2012)
6. Manohar, V., Crandall, J.: Programming robots to express emotions: Interaction paradigms, communication modalities, and context. Human-Machine Systems, IEEE Transactions on **44**(3), 362–373 (2014)
7. Metta, G., Sandini, G., Vernon, D., Natale, L., Nori, F.: The icub humanoid robot: an open platform for research in embodied cognition. In: Proceedings of the 8th workshop on performance metrics for intelligent systems, pp. 50–56. ACM (2008)

8. Salichs, M.A., Barber, R., Khamis, A.M., Malfaz, M., Gorostiza, J.F., Pacheco, R., Rivas, R., Corrales, A., Delgado, E., García, D.: Maggie: A robotic platform for human-robot social interaction. In: Robotics, Automation and Mechatronics, 2006 IEEE Conference on, pp. 1–7. IEEE (2006)
9. Shepstone, S.E., Tan, Z.H., Jensen, S.H.: Using audio-derived affective offset to enhance tv recommendation. IEEE Transactions on Multimedia **16**(7), 1999–2010 (2014)
10. Springer: International Journal of Social Robotics. `http://www.springer.com/about+springer/media/pressreleases?SGWID=0-11002-6-805078-0`. Accessed: 2014-12-19
11. Tan, Z.H., Lindberg, B.: Low-complexity variable frame rate analysis for speech recognition and voice activity detection. Selected Topics in Signal Processing, IEEE Journal of **4**(5), 798–807 (2010)
12. Thomsen, N.B., Tan, Z.H., Lindberg, B., Jensen, S.H.: Improving robustness against environmental sounds for directing attention of social robots. In: Proceedings of the 2nd Workshop on Multimodal Analyses Enabling Artificial Agents in Human-Machine Interaction, Singapore (2014)