A Hybrid Supervised Approach for Human Robot Interaction with Children with Autism Spectrum Disorder

Vinícius Silva, Filomena Soares, and João Sena Esteves
Center Algoritmi, University of Minho, Campus de Azurém, Guimarães, Portugal

Abstract

Individuals with Autism Spectrum Disorder (ASD) are characterized by difficulties in social communication and restricted patterns of behaviour. Technological tools such as social robots and Objects with Playware Technology (OPT) have been employed in support sessions with individuals with ASD. The present work proposes a supervised behavioural system using a hybrid approach (robot + OPT) to enable a more natural and adaptive interaction. In general, it is possible to conclude that the children understood the mechanics of the games and successfully interacted with the robot and both OPTs.

Author Keywords. Human Robot Interaction, Autism Spectrum Disorders, Playware.

1. Introduction

Emotional skills are paramount for a successful human-human communication. Individuals with Autism Spectrum Disorder (ASD) are characterized by having deficits in social communication and social interaction. Moreover, due to the diversity and specificities of symptoms, professionals have found some difficulties in developing effective support methods. Robots have been used to facilitate support processes with children with ASD, with robots acting as a mediator between the child and the game partner (Tapus et al. 2012). Research has found that interacting with the robots draws these children into a range of new social behaviours (Tapus et al. 2012; Dautenhahn and Werry 2004). Although with promising results from the children's point-of-view, the majority of the systems are controlled using the Wizard-of-Oz (WOZ) method, meaning that in reality the robot does not adapt its behaviour to the children's actions. Analogous to the use of robots, researchers have been employing Objects with Playware Technology (OPT) to interact with children with ASD. The term "playware" is suggested as a combination of intelligent play and playful experiences among users (Lund, Klitbo, and Jessen 2005).

Following this trend, the present work proposes a novel supervised behavioural system architecture using a hybrid approach (robot + OPT) to allow the detection of the child behaviours and consequently adapt the robot actions, enabling a more natural interaction and lighten up the cognitive burden on the human operator. The OPT is to be used as an add-on to the human-robot interaction with children with ASD in emotion recognition activities.

The present paper is organized as follow: Section 2 presents the materials and methods of the presented approach, the results concerning the performance of the action recognition model, and the two pilot studies conducted with children with ASD are reported in Section 3.

2. Materials and Methods

The proposed system, depicted in Figure 1, consists of a humanoid robot (ZECA – Zeno Engaging Children with Autism) capable of displaying facial cues, a computer, two OPT devices, and a 3D sensor. The PlayCube (with a 1.5-inch OLED RGB display) and PlayBrick (with a 5.0-inch touch display), shares the same internal components: an Inertial Measurement Unit (IMU), a small development board (ESP32) that already has built-in Bluetooth and Wi-Fi communication, an RGB LED ring, a Linear Resonant Actuator (LRA), and a Li-Po battery.

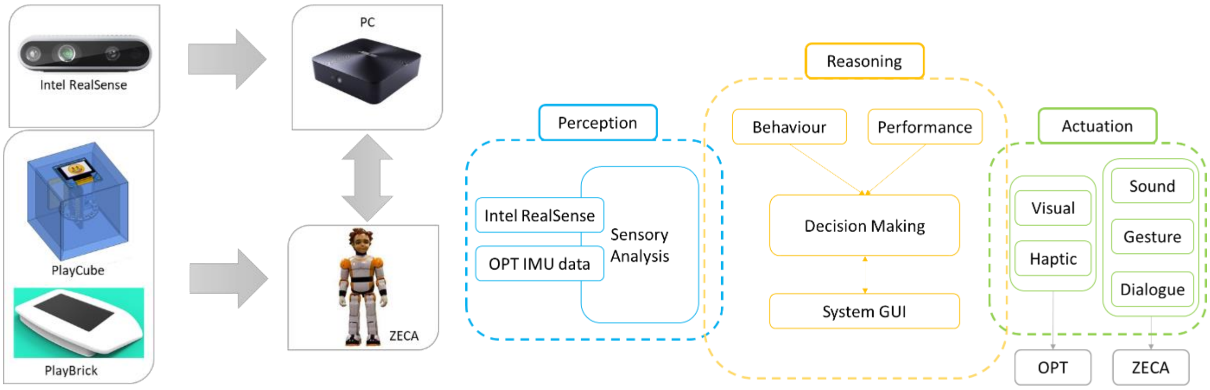

Figure 1: On the right: the proposed system. On the right, the behavioural architecture is depicted – the perception (blue), Reasoning (yellow), and Actuation (green) layers (Vinícius Silva et al. 2019).

The Intel RealSense is a USB-powered device that contains a conventional RGB full HD camera, an infrared laser projector, and a pair of depth cameras. The Intel RealSense model D435 along with the Intel RealSense SDK and Nuitrack SDK are used to track the user joints, as well as, for face tracking and head orientation.

Concerning the software architecture (Figure 1), it consists of three main layers (Vinícius Silva et al. 2019). The perception layer is responsible for sensing and processing of the data received from the sensors. The reasoning layer is influenced by the child behaviour and performance. It takes into account the outputs of the perception layer (which does face tracking for computing head motion, skeleton tracking for action recognition, and the IMU data) and the performance data in order to adapt the support session. The action recognition, Figure 2, is done by transforming the tracking data into an image representation that is used as input to a deep learning model (Vinicius Silva et al. 2020). The final output is the classification of 9 non-verbal behaviours (3 stereotypical and 6 typical behaviours), Table 1. The output of the reasoning layer will influence the dynamics of the next layer, the Action layer. Therefore, the robot actions and the OPT feedback will be influenced by the interaction flow of the session.

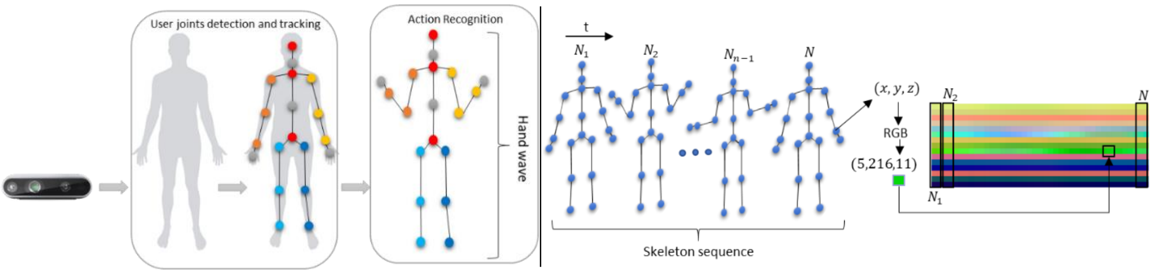

Figure 2: The overall pipeline concerning the detection of the user actions. On the right, it shows the joints encoding to colour space, where N is the number of frames.

The activities played are focused on emotion recognition. Two games were developed: Recognize, where ZECA performs a facial expression and its associated gestures, representing one of the five emotions (anger, fear, happiness, sadness, and surprised), and the child has to choose the correct facial expressions by selecting it from the OPT screen and storytelling, where the robot randomly tells a story, and the child has to identify the correct affective state of the main character. In parallel, a visual cue of the story is shown on the PlayBrick device.

3. Results and Final Remarks

Table 1 shows the average accuracy per class of the action recognitions model. The model achieved an average test accuracy of 92.5±0.5%. The model achieved a Matthews Correlation Coefficient of 91.6±0.6%. The present method is able to run in real-time at about 31 FPS on a quad-core CPU, being fast enough for most real-time applications. Additionally, the model achieved better performance when compared to other approaches on the state of the art on the detection of stereotyped behaviours – 'Hand_Wave', 'Cover_Ears', and 'Rocking' (Vinicius Silva et al. 2020).

IDLE	STANDING	CLAP	HAND_WAVE	HAND_RAISE	POINTING	COVER_EARS	TURN	ROCKING
83±2.0%	97±2.2%	92±2.6%	96±1.2%	100±1.2%	100±1.2%	88±3.1%	86±3.2%	93±2.1%

Table 1: Average accuracy per class with the standard deviation.

Two pilot studies were conducted focusing on evaluating the systems constraints with the recognize and storytelling activities – one with the PlayCube (3 children) and another with the PlayBrick (4 children) – at a school setting with children with ASD aged between 6 and 9 years old during 4 sessions. By analysing the results of the pilot studies, it is possible to conclude that the children understood the mechanics of the games and successfully interacted with both OPTs. Moreover, they reacted positively to the activity which might indicate that the developed approach allowed the children to interact in a comfortable and natural way with the system. Additionally, there was a combined attention of the child towards the robot and OPT, indicating that both components were successful in captivating the children interest. In general, the children were keen to participate in the activities. They were also attentive to the OPTs feedbacks – lights, haptic, as well as the images for correct and incorrect answers displayed on the screen. As future work, a study will be conducted with a larger sample of children with ASD, aiming to understand if and how the presented hybrid approach can be used as a valuable tool to develop skills of emotional labelling by children with ASD.

References

Dautenhahn, Kerstin, and Iain Werry. 2004. "Towards Interactive Robots in Autism Therapy: Background, Motivation and Challenges." *Pragmatics & Cognition* 12 (1): 1–35. https://doi.org/10.1075/pc.12.1.03dau.

Lund, Henrik Hautop, Thomas Klitbo, and Carsten Jessen. 2005. "Playware Technology for Physically Activating Play." *Artificial Life and Robotics* 9 (4): 165–74. https://doi.org/10.1007/s10015-005-0350-z.

Silva, Vinícius, Sandra Queirós, Filomena Soares, João Sena Esteves, and Demétrio Matos. 2019. "A Supervised Autonomous Approach for Robot Intervention with Children with Autism Spectrum Disorder." In *ICINCO 2019 - Proceedings of the 16th International Conference on Informatics in Control, Automation and Robotics*, 2:497–503. https://doi.org/10.5220/0007958004970503.

Silva, Vinicius, Filomena Soares, Joao Sena Esteves, and Gianni Vercelli. 2020. "Human Action Recognition Using an Image-Based Temporal and Spatial Representation." In *International Congress on Ultra Modern Telecommunications and Control Systems and Workshops*, 2020-Octob:41–46. https://doi.org/10.1109/ICUMT51630.2020.9222408.

Tapus, Adriana, Andreea Peca, Amir Aly, Cristina Pop, Lavinia Jisa, Sebastian Pintea, Alina S Rusu, and Daniel O David. 2012. "Children with Autism Social Engagement in Interaction with Nao, an Imitative Robot." *Interaction Studies* 13 (3): 315–47.

Acknowledgments

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020. Vinicius Silva thanks FCT for the PhD scholarship SFRH/BD/SFRH/BD/133314/2017.