CopyRobot: Interactive Mirroring Robotics Game for ASD Children

Laura Santos[1,2(✉)], Alice Geminiani[1], Ivana Olivieri[3], José Santos-Victor[2], and Alessandra Pedrocchi[1]

[1] Department of Electronics, Information and Bioengineering (DEIB) NEARLab - Neuroengineering and Medical Robotics Laboratory, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milan, Italy
laura.d.santos@tecnico.ulisboa.pt

[2] Institute for Systems and Robotics, Instituto Superior Tecnico, Lisboa, Portugal

[3] CARELab, IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy

Abstract. The family of disorders commonly known as autism is characterized by a deficit in social interaction and restricted repetitive and stereotyped patterns of behaviours, activities and interests. Motor disturbances are not part of the diagnosis of the children with autism but some studies have estimated that between 80 and 90% of children with Autism Spectrum Disorder (ASD) demonstrate some degree of motor impairments. Several therapies have been used for the improvement of motor skills, always leading to behavioural improvements as side-effects, demonstrating the importance of motor interaction and stimulation for the case of autism. Recent studies have shown that motor, imitation and social abilities are all related in people with autism. In this work, a humanoid robot is used to create a therapy that unites all these areas. The system involves a robot (NAO), a Kinect camera and Personal Computer, with the goal of facilitating the interaction between therapist and a child with ASD during a physical therapy session. To improve the imitation abilities of the child, the robot was programmed to mirror both the child and the therapist movements. After testing different tracking methodologies, the Kinect sensor was selected as the best compromise of quality and cost. Two protocols were developed, depending on who plays the role of the main actor. In the first protocol, the robot is the master and leads the interaction. It decides the exercise to execute and gives feedback to both the therapist and the child. In the second protocol, the choice of the exercise sequence is the therapist's responsibility. To promote interaction further during clinical tests, the protocol was changed to include gesture imitation. For the robot master protocol, the space theme was chosen. For the therapist master protocol, the theme of sports, that was already performed by the children in the usual therapy, was adopted. The system was tested in realistic conditions with two different autistic children. The reaction was different in each case but it demonstrated the importance of these imitation games in the treatment of this disease.

J. Henriques et al. (Eds.): MEDICON 2019, IFMBE Proceedings 76, pp. 2014–2027, 2020.
https://doi.org/10.1007/978-3-030-31635-8_239

1 Introduction

Autism Spectrum Disorder (ASD) appears on subjects in very different ways. The incidence has been growing and currently affects 1 in 160 children [1]. It is characterized by a symptomatic triad: difficulties in social interaction, communication deficit, and presence of repetitive behaviours. The cause for autism is still unknown and there is no cure. The treatment, as the Spectrum itself, is multifaceted, ranging from the standard Applied Behaviour Analysis [2], to experiments in the pool and with horses [3]. All treatments try to adapt to the needs of each child and many of them centre on the recovery of the social part. Recently, studies with ASD children proved the relation between imitation, social and motor abilities of these kids [1]. This unleashed an increasing interest in the motor and imitation recovery of ASD children. The main motor problem is related with the praxis of the movement [1], while the imitation problem (also related with this last one) centres on a difficulty of reproducing the same kinetics of another individual, although the same target can be achieved [4].

These capabilities can be developed through the use of robots, in a form of embodied mirroring (Fig. 1). This is justified by two facts. On one hand, research on mirror neurons has revealed the existence of neurons which are activated both during the execution of goal-directed movements, as well as during the observation of those movements performed by others, revealing the importance of mirroring as a form of motor training [5]. On the other hand, robots and physical mirroring are more effective than training with graphical Avatars or other forms of Virtual Reality, [6]. It was verified that the number of people that had a spontaneous imitation reactions was significantly larger when a robot was physically present, than in the situation where the robot was presented through video. Moreover, robots have been extensively used in the social therapy of children with Autism in the form of Social Assistive Robots, where they were mainly used as interactive toys, due to the connection of these children with several kinds of technology [7].

Fig. 1. Embodied mirroring

Mirroring has already been used widely in robotics requiring three implementation steps: capture of the observed human motion; mapping to robot joint angles; control the robot with the calculated angles. Most of the works have used wearable sensors and vision based without markers devices to capture the movement [8]. This was due to the low cost and high portability of both [9]. Inertial Measurement Units are the principal wearable sensors used for robot mimicry. They are constituted by a 3D gyroscope, a 3D accelerometer and a 3D magnetometer, which are responsible to measure the position and velocities of these units in relation to a global reference frame. They have been used in several works for the robotic mirroring of the upper limbs and lower limbs of a human. In terms of marker-less vision based devices, a breakthrough sensor was the Microsoft Kinect, from now on referred to only as Kinect [9]. It is able to reconstruct the 3D joint positions of a human skeleton. The first version was based on a combination of a standard camera and an infrared sensor and it was highly susceptible to noise. The second version solved this problem with a time of flight camera.

Some of the works done with Kinect in the field of embodied mirroring were [10–12]. There are two key control strategies of the robot in these works. The first is based on the end-effector trajectories, obliging the robot end effector to follow the same trajectory as the human which is controlling it. They use inverse kinematics strategies to discover the angles from the position of the end effector [8]. The second is centred on the joint space, trying to convert directly the joint angles of the human in the joint angles of the robot. The spatial vector method is one of the most used. It calculates the links between the several joints and determines the angle between them and a certain reference frame [10].

The majority of these works were mainly used for an easier control of the robot and rarely were presented in a therapeutic environment. They centre on one master to control the robot. Therefore, the main goal of this project is to develop an imitation system that can be used in a therapy for ASD children and that will allow the inclusion of a second master. In this way, child and therapist will be able to control the robot. Consequently, a setup for embodied mirroring with several exercises is implemented and tested to permit a simultaneous training of the motor, imitation and social capabilities of these children, a tool missing in the literature.

2 Methods

2.1 Mirroring System

The robot chosen for this imitation system was NAO from Aldebran Robotics. It is a small humanoid robot that was already used in some social therapies for ASD children due to their simple expressions that resemble a human without being one. It is constituted by 24 degrees of freedom but just 3 of them were used for the physical mirroring of the upper limbs: Elbow Roll, Shoulder Roll and Shoulder Pitch. Therefore, the basic arm movements allowed are the flexion and extension, abduction and adduction, horizontal abduction and adduction of the arms and the flexion and extension of the forearms. NAO has also LEDS and loudspeakers which can be used to transmit visual and audio signals, providing a bigger interaction with people.

As a tracking device the system chosen was the Kinect. Comparing to the wearable sensors, Kinect presents the big advantage of not being intrusive, which is especially important for children with autism. The 3D joint positions are obtained directly from the camera with a frame rate of 30 Hz. In total 25 key-points are extracted to form a skeleton. A median filter with a window of 5 samples is applied in order to reduce the noise. Then the Kinect reference frame is transformed on the NAO reference frame (Fig. 2) through two rotations, one around the y axis and another around the x axis, summarized by the rotation matrix presented below.

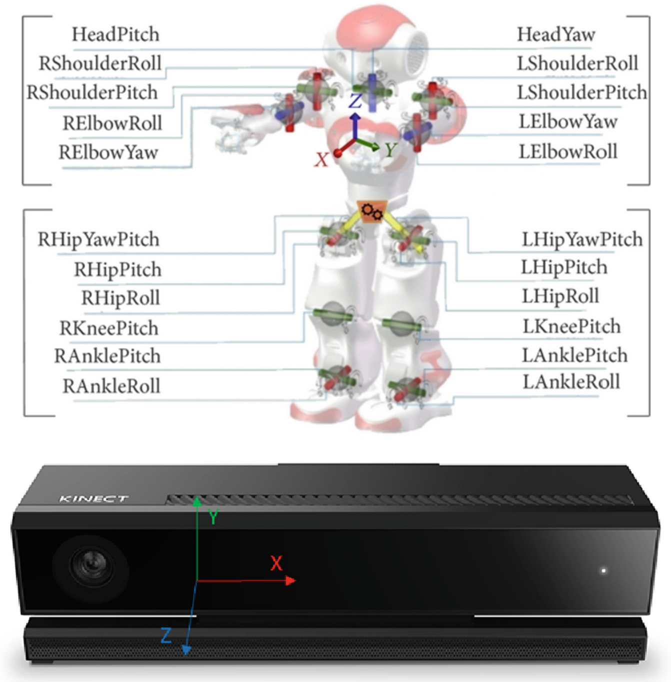


Fig. 2. NAO reference frame and Kinect reference frame

$$R_T = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

Then, vectors are created between the 3D joint positions which constitute the limbs that can be seen in Fig. 3. The angles are calculated using Eq. 2 and the vectors represented in Table 1, where **y** and **z** represent the unit vectors of the y and z axes.

$$angle = \arccos\left(\frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\,\|\mathbf{b}\|}\right) \tag{2}$$

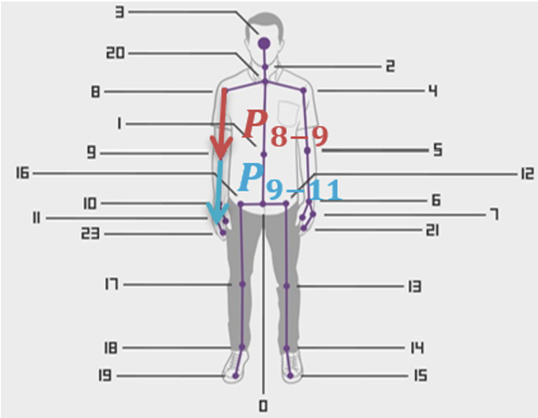


Fig. 3. 3D joints calculated by the Kinect strategy with the vectors representing the right arm ($\mathbf{P_{8-9}}$) and forearm ($\mathbf{P_{9-11}}$).

Table 1. Shoulder and Elbow angles calculated from kinect results.

Angles	**a**	**b**
Right shoulder roll	$\mathbf{P_{8-9}}$	$-\mathbf{y}$
Right shoulder pitch	$\mathbf{P_{8-9}}$	$-\mathbf{z}$
Right elbow roll	$\mathbf{P_{8-9}}$	$\mathbf{P_{9-11}}$
Left shoulder roll	$\mathbf{P_{4-5}}$	$\mathbf{y}$
Left shoulder pitch	$\mathbf{P_{4-5}}$	$-\mathbf{z}$
Left elbow roll	$\mathbf{P_{4-5}}$	$\mathbf{P_{5-7}}$

This Kinect-based system was compared with a standard measurement system of 5 inertial measurements units. The complete description of this comparison is presented on the conference paper that was written from this thesis [13]. Basically, the units were put in the chest and in the middle of both arms and forearms of a subject. Five simultaneous acquisitions of three different movements were recorded. The root mean square error between the IMU angles and the Kinect angles was calculated.

The same was done to evaluate the angles of the NAO movements. In this case the sensors were put on the arms, forearms and chest of the robot. The angles given to NAO were calculated through the movement of a subject recorded by Kinect. Then, the true angles done by the robot were measured by the 5 inertial measurement units. The root mean square error was used to compare these angles with the motion angles given to NAO.

2.2 Protocols

As this mirroring system was designed to improve the interaction between the child and the therapist, Kinect should distinguish each subject. Actually, Kinect is able to track both and the identification of each person is based on the position of the left hand. The person with the left hand more at the left of Kinect is considered the therapist. After this identification two protocols can be used: one in which the robot is the master and another where the therapist is the master. On the first protocol, NAO presents a movement. Then the therapist starts doing it. At the same time NAO is mirroring him/her. After, it is the turn of the child of executing the movement and being mirrored by NAO. If the movement is performed correctly, NAO gives a vocal feedback and its LEDs turn out to green. If the child or the therapist do not start the movement the LEDs on NAO eyes are red and it gives a negative feedback. The protocol block diagram is represented on Fig. 4a.

This protocol was adapted to be more interactive to the child, through the collaboration with IRCCS Fondazione Don Carlo Gnocchi. Actually, the movements presented by NAO are part of a story about NAO, whose goal is to train specific gestures. The story is set on the theme of the space, being NAO someone from another planet. In this way the specific gestures are to wave, move out clouds from the sky (reaching and back with two arms), take out stars (reaching and back with one arm) and finally point to the NAO planet. The feedbacks are given when a target position is achieved. This target position is related with the several parts of the movements, and it is constituted by specific target angles. The target angles were decided through the acquisition of the several movements on a group of 28 people. They were optimized in order to all the people have a positive feedback, after the correct execution of the movement.

For the second protocol, in which the therapist is the master, he/she and the child teach NAO the sports from Earth. The chosen sports were basket, bowling, swimming, tennis and skiing. In this case, first the therapist shows the movement during a certain interval, then NAO replicates it, and finally it is the turn of the child. At this moment, NAO also mirrors the child (Fig. 4b). No feedback is given by NAO because the therapist will do it.

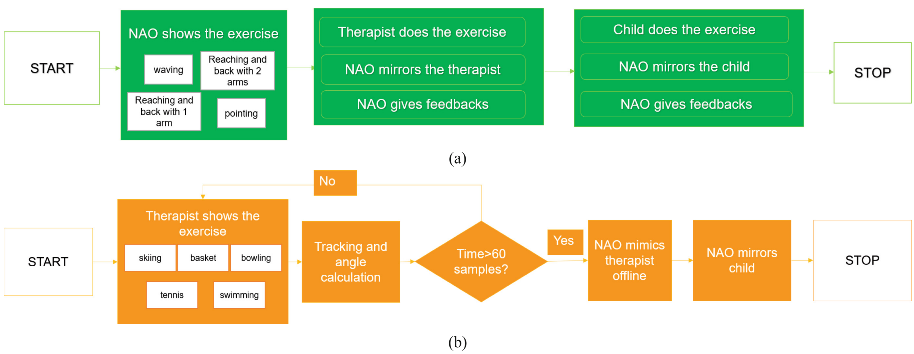


Fig. 4. Implemented protocols - Robot Master (a) and Therapist Master (b)

These protocols, in their clinic/game version were tested in two autistic children and in one healthy child. Some parameters were extracted to allow the future evaluation of the children evolution along the sessions. For the robot master protocol, it was evaluated the starting time of the movement of the children ($\mathbf{t_s}$), the error of the peak angle of the movement in relation to the target position ($\mathbf{\Delta E}$), the mean velocity on the peak angle ($\mathbf{\Delta v}$) and the maximum velocity ($\mathbf{v_i}$) and respective arrival time ($\mathbf{t_i}$). For the second protocol, the chosen parameters were calculated in relation to the whole signal since the gold standard movement is known, the movement of the therapist. In this way, the latency in relation to the beginning of the movement, the root mean square error (RMSE) between the signal of the child and the signal of the therapist and the duration of the movement were calculated. On both protocols, in order to quantify the characteristic random movement of the ASD children, it was counted the number of samples in which the change in the joints angles was larger than 0.1 rad. This parameter was called random samples.

3 Results

From Table 2, it is possible to observe that the root mean square error between the signals obtained with a Kinect and the gold standard given by the IMU sensor is generally low, as well as its variability, for both child and adult acquisitions. This means that the acquisitions done with Kinect are sufficiently similar to the ones measured with IMU. Moreover, Kinect is less invasive, it is easier to assemble the whole system and the cost is smaller in relation to the IMU system. It also allows the tracking of two people, which is impossible with the IMU, since it would require more sensors. In the same table, it is also possible to notice that the mean of the error of the adult is lower than the error of child, which can be justified by a different position of the Kinect camera, or a better performance of the person who was representing the therapist, than the child. Focusing on the angles calculated by Kinect and the angles done by the robot (Fig. 5), a difference lower than 25° can be observed. This means that the movement of the robot is sufficiently similar to the person it is mimicking, even considering that the robot has different range of motions, degrees of freedom and safety controls with respect to the human limbs. Therefore the Kinect system is a good substitute of the IMUs in a mirroring system.

The clinical acquisitions done on two boys with ASD had different outcomes. One of them had a negative reaction after the complexity of the movements requested increased and did not want to continue the session. The other boy, in the beginning was not comfortable with the presence of the robot but after he completely engaged on the therapy. While in the beginning the therapist was helping him to do the exercises, in the end he was teaching the therapist how to do correctly the movements.

The movements of this last child were compared with the movements of a healthy one to understand which parameters could be useful for the evaluation of the performance of the ASD child on future sessions. An example of the

Table 2. RMSE between Inertial Measurement Unit (IMU) and Kinect-based control signals for the adult and the child. Values are reported as mean ± SD (Standard Deviation).

Movements	RMSE (rad)	
	Adult	Child
Lateral abduction/Adduction shoulder	0.23 ± 0.12	0.41 ± 0.19
Abduction/Adduction shoulder	0.35 ± 0.07	0.41 ± 0.04
Flexion extension elbow	0.31 ± 0.10	0.30 ± 0.19

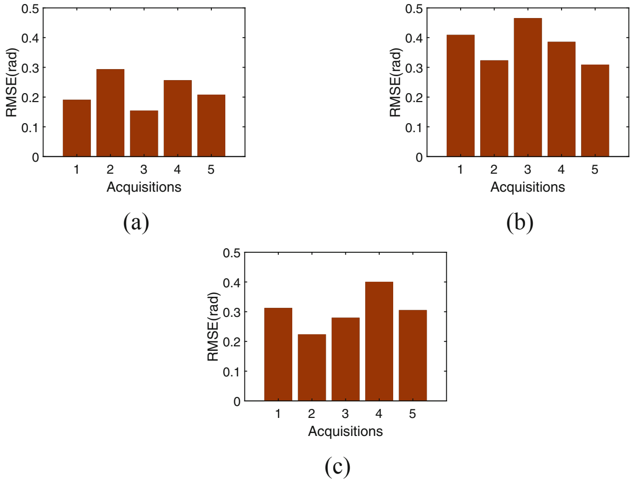


Fig. 5. Root mean square error for the 5 acquisitions done for (a) horizontal abduction and adduction, (b) abduction and adduction of the arms and (c) flexion and extension of the forearms, comparing the angles done by the robot and the angles calculated by the Kinect

Robot wave movement is shown on Fig. 6 and all parameters calculated for the several movements for the ASD child (ASD) and for the healthy child (H) are presented in Table 3. In the figure, the black straight line marks the beginning of the movement. The dashed lines represent the target angles (TA) for the internal (dark green) and external (light green) rotation and the dark green and light green stars the peak angles achieved for the internal and external rotation. From the figure, it is possible to observe the brisk movements done by the ASD child. As a consequence and also for the other movements his velocity parameters are much higher than the parameters of the healthy child. However, this reflects in a higher absolute error. In this example, it is shown for the external rotation of the arm, where one can notice the large distance between the light green star and the light green dashed line. It is also clear that the starting time is longer for the ASD child in comparison to the healthy child, reflecting his longer reaction period. The number of random samples also differs between the healthy and the ASD child being higher for the latter.

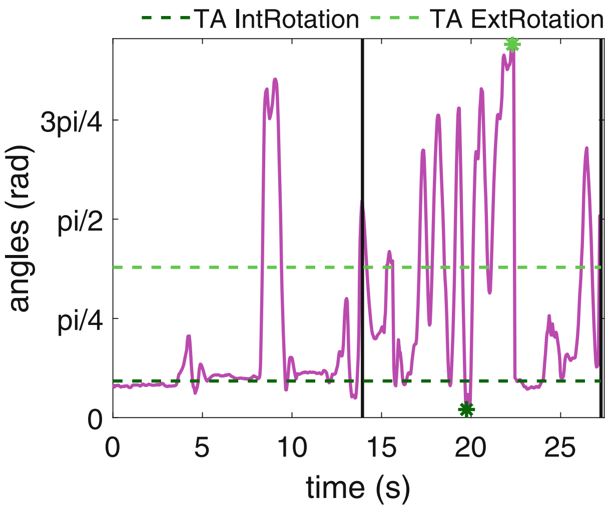
Table 3. Parameters calculated for all the movements done in the robot master protocol.

	$\mathbf{t_s}$(s)		$\mathbf{\Delta E}$(rad)		$\mathbf{\Delta v}$(rad/s)		$\mathbf{v_i}$ (rad/s)		$\mathbf{t_i}$(s)		**Random**	
	ASD	H	ASD	H	ASD	H	ASD	H	ASD	H	ASD	H
Wave	13.9	8.9	1.765	0.770	−0.197	−0.039	86	49	19.3	16.3	30	17
Clouds	9.6	8.9	2.467	0.388	−1.468	−0.133	31	6	12.1	28.0	320	49
Stars	17.4	1.1	0.391	0.271	−1.709	−1.314	13	7	26.5	13.7	14	4
Point	6.8	4.2	0.047	0.145	−3.405	−3.181	10	7	6.8	11.6	162	61

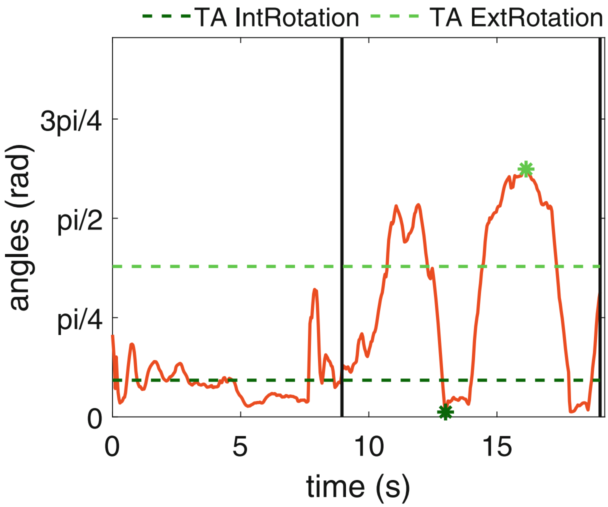
Similar conclusions could be taken from the analysis of the parameters of the therapist master protocol (Table 4). The latency tends to be smaller on the healthy child than on the ASD child, showing a larger reaction time for the latter. Observing Fig. 7, an example of the basket movement, in which both the therapist shoulder angle movement and the child movement are presented, it is clear that the shape of the signal for the ASD child differs more from the therapist signal than for the healthy child. Therefore, the root mean square error is larger in comparison with the healthy child, which happens for all the movements tested. In this case, this is due a faster execution of the movement by the ASD child. However, the duration is one of the parameters that varies between ASD child and healthy child for the several protocols. The most differentiating parameter is the number of random samples, which is always larger on the ASD child in comparison to the healthy child.

Table 4. Parameters calculated for all the movements done in the therapist master protocol.

	Latency (s)		Duration (s)		RMSE		Random	
	ASD	H	ASD	H	ASD	H	ASD	H
Basket	1.70	2.09	1.48	2.02	0.93	0.29	91	24
Swimming	1.81	0.33	0.51	2.38	0.51	0.42	101	74
Bowling	2.99	1.91	4.62	0.94	0.92	0.87	120	46
Skiing	0.90	0.40	4.33	1.26	0.99	0.40	195	45
Tennis	8.84	0.07	0.43	1.48	0.17	0.23	144	86

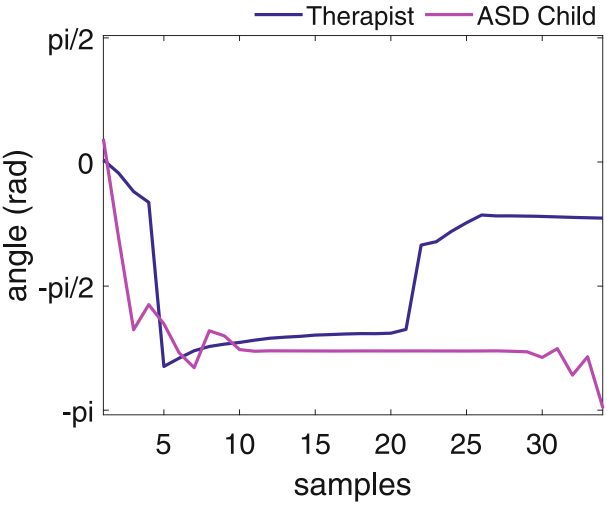


(a)

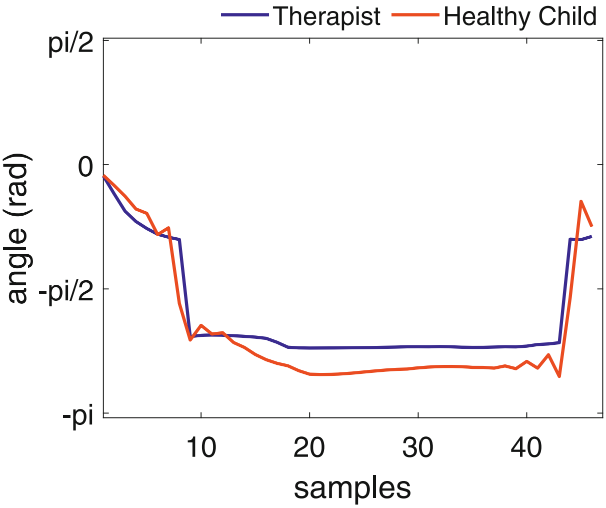


(b)

Fig. 6. Acquisitions of the robot wave movement for the ASD (a) and Healthy (b) children.



(a)



(b)

Fig. 7. Shoulder pitch angle of the basket movement done by the child and therapist where on (a) is the graph of the ASD child and on (b) the one of the healthy child.

4 Discussion and Conclusions

ASD therapies have been mostly focused on social development of these children, and only recently have included the training of motor abilities. The use of the NAO robot is a promising additional resource to stimulate motivation and participation of ASD children in the therapy. The proposed protocol promotes a global intervention to these children: (i) motor training through the execution of motor tasks in a controlled way, (ii) cognitive stimulation, induced through the different modes of mimicking games, and (iii) a social dimension, by proposing a triadic interaction between the child, the therapist and the robot. This three-fold interaction was particularly innovative in relation to existing literature, where dyadic interaction (child, robot) is the most common form of interaction.

The first part of the results has shown that the Kinect is a good substitute to the more precise IMU sensors. However the thorough validation of the Kinect still requires a larger set of movements and subjects, especially children. It has allowed to the successful implementation of an interactive mirroring system with the NAO humanoid robot, both online (Robot Master Protocol) and offline (Therapist Master Protocol). Not all movements worked equally well, namely the Clouds movement on the first protocol and the Skiing and Bowling movements on the second protocol. The main problem on the clouds movement was the inter-subject variability of the movement. The definition of the target angles for this movement was done through the use of a population of 28 adults. Probably, a similar test should be done with children, for defining target angles adapted to children. On the Skiing and Bowling movements, occlusions of the subject arms with the torso created difficulties to the computation of the subject motion angles, which occasionally blocked the robot arms behind its torso. The problem could be avoided by limiting/filtering the reference joint angles to control the NAO robot.

The whole system was tested with two ASD children, who reacted in different ways. One of them positively interacted with the robot, in a very successful and encouraging test. The other child also reacted well in the beginning, and showed interest and empathy towards the robot, but became scared by the difficulty of the movements. These different reactions are related to the different level of autism of the two children, higher in the child with less positive reactions and engagement. This case demonstrated the need of improving the system for different degrees of autism. A possibility could be to adapt the system to mirror just one person, so that the therapist could help the child to execute the movements. Alternatively, an initial step may include mirroring the child only, focusing on the imitation skills of the child. At a later stage, also the sessions with these children should continue in order to evaluate the relevance of the kinematic parameters calculated. We are currently initiating tests of the system with more children, to compare the results to the ones of the standard therapy. This will allow the complete evaluation of the efficacy and effectiveness of this type of intervention. Overall, this work shows also the importance of the clinical inputs to drive the system design (a camera and a robot) towards an attractive and impact-full therapy to autistic children.

Acknowledgment. This work was supported by Portuguese Foundation for Science and Technology (proj. UID/EEA/50009/2019) and CerebNEST, Partnering Project of the Human Brain Project.

This master thesis was awarded the top grade under a double diploma agreement between the Politecnico di Milano (Italy) and the Instituto Superior Técnico (Lisbon, Portugal), under the supervision of Professor Alessandra Pedrocchi, Eng. Alice Geminiani and Professor José Santos-Victor, to whom the author is profoundly grateful. The support of Eng. Paolo Borzatta (European House-Ambrosetti) for providing access to a NAO robot, and Doctor Ivana Olivieri (IRCCS Fondazione Don Carlo Gnocchi) for guidance in the system design towards clinical use, is highly appreciated.

A part of the work was selected to showcased in the exhibition 'Brain, wider than the sky' promoted by the Fundação Calouste Gulbenkian, and the technical work is described in a paper accepted for presentation at the International Conference of Engineering in Medicine and Biology [13].

References

1. Anzulewicz, A., Sobota, K., Delafield-Butt, J.T.: Toward the autism motor signature: gesture patterns during smart tablet gameplay identify children with autism. Sci. Rep. **6**(31107), 1–13 (2016)
2. Colebourn, J.A., Golub-Victor, A.C., Paez, A.: Developing overhand throwing skills for a child with autism with a collaborative approach in school-based therapy. Pediatr. Phys. Ther. **29**(3), 262–269 (2017)
3. Caputo, G., Ippolito, G., Mazzotta, M., Sentenza, L., Muzio, M.R., Salzano, S., Conson, M.: Effectiveness of a multisystem aquatic therapy for children with autism spectrum disorders. J. Autism Dev. Disord. **48**(6), 1945–1956 (2018)
4. Forbes, P.A.G., Pan, X., Hamilton, A.F.d.C.: Reduced mimicry to virtual reality avatars in autism spectrum disorder. J. Autism Dev. Disord. **46**(12), 3788–3797 (2016)
5. Fadiga, L., Fogassi, L., Pavesi, G., Rizzolatti, G.: Motor facilitation during action observation: a magnetic stimulation study. J. Neurophysiol. **73**(6), 2608–2611 (1995)
6. Hofree, G., Ruvolo, P., Bartlett, M.S., Winkielman, P.: Bridging the mechanical and the human mind: Spontaneous mimicry of a physically present android. PLoS ONE **9**(7), e99934 (2014)
7. Scassellati, B., Admoni, H., Matarić, M.: Robots for use in autism research. Annu. Rev. Biomed. Eng. **14**(1), 275–294 (2012)
8. Alibeigi, M., Rabiee, S., Ahmadabadi, M.N.: Inverse kinematics based human mimicking system using skeletal tracking technology. J. Intell. Robot. Syst. **85**, 27–45 (2017)
9. Spasojević, S., Ilić, T.V., Milanović, S., Potkonjak, V., Rodić, A., Santos-Victor, J.: Combined vision and wearable sensors-based system for movement analysis in rehabilitation. Methods Inf. Med. **56**, 95–111 (2017)
10. Yang, N., Duan, F., Wei, Y., Liu, C., Tan, J.T.C., Xu, B., Zhang, J.: A study of the human-robot synchronous control system based on skeletal tracking technology. In: 2013 IEEE International Conference on Robotics and Biomimetics (ROBIO), pp. 2191–2196 (2013)

11. Zannatha, J.M.I., Tamayo, A.J.M., Sánchez, A.D.G., Delgado, J.E.L., Cheu, L.E.R., Arévalo, W.A.S.: Development of a system based on 3D vision, interactive virtual environments, ergonometric signals and a humanoid for stroke rehabilitation. Comput. Methods Programs Biomed. **112**(2), 239–249 (2013)
12. Almetwally, I., Mallem, M.: Real-time tele-operation and tele-walking of humanoid robot NAO using kinect depth camera. In: 2013 10th IEEE International Conference on networking, sensing and control (ICNSC), pp. 463–466, April 2013
13. Geminiani, A., Santos, L., Casellato, C., Farabbi, A., Farella, N., Santos-Victor, J., Olivieri, I., Pedrocchi, A.: Design and validation of two embodied mirroring setups for interactive games with autistic children using the NAO humanoid robot. In: 41st IEEE Engineering in Medicine and Biology Society, Berlin, Germany (2019). Accepted with oral presentation