Design and validation of two embodied mirroring setups for interactive games with autistic children using the NAO humanoid robot

Alice Geminiani[1*], Laura Santos[1,2], Claudia Casellato[3], Andrea Farabbi[1], Nicola Farella[1],
José Santos-Victor[2], Ivana Olivieri[4] and Alessandra Pedrocchi[1]

Abstract— Socially assistive robots have shown potential benefits in therapy of child and elderly patients with social and cognitive deficits. In particular, for autistic children, humanoid robots could enhance engagement and attention, thanks to their simplified toy-like appearance and the reduced set of possible movements and expressions. The recent focus on autism-related motor impairments has increased the interest on developing new robotic tools aimed at improving not only the social capabilities but also the motor skills of autistic children. To this purpose, we have designed two embodied mirroring setups using the NAO humanoid robot. Two different tracking systems were used and compared: Inertial Measurement Units and the Microsoft Kinect, a marker-less vision based system. Both platforms were able to mirror upper limb basic movements of two healthy subjects, an adult and a child. However, despite the lower accuracy, the Kinect-based setup was chosen as the best candidate for embodied mirroring in autism treatment, thanks to the lower intrusiveness and reduced setup time. A prototype of an interactive mirroring game was developed and successfully tested with the Kinect-based platform, paving the way to the development of a versatile and powerful tool for clinical use with autistic children.

I. INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental syndrome, with a global prevalence of 1 in 160 children [1]. ASD symptoms are diverse, but a symptomatic triad is usually present: poor social interaction skills, communication deficit and presence of repetitive behaviours. Treatments tend to focus on the first symptom, trying to improve social capabilities of patients [2]. However, recent studies demonstrate that not just social skills are affected in ASD children, but also some motor problems are reported, namely in the praxis of the movement [3]. This also influences the imitative skills. Recent studies have shown that motor, imitation and social skills deficits are related, leading to an increasing interest in the recovery of motor and imitation problems in ASD patients [1]. In this scenario, robots can be useful tools for improving motor and interaction skills. The advantages of using robots are multi-fold. On one hand, autistic children are very prone to technology, on the other hand, robots have specific features that ASD children appreciate. For example, they have repetitive actions, represent a stable environment and (especially) humanoid robots have stylized human expressions, easier to be interpreted by an ASD child [2]. Robots have been already used in the treatment of autism as interactive or educational toys, raising positive reactions and increased engagement [4]. It would be important to transform them into "embodied mirrors", recruiting imitation neural circuits and thus using mirroring as a form of motor training [5]. In addition, when compared with virtual reality scenarios, robots induce a stronger effect on spontaneous imitation [6]. Therefore, robots can strongly enhance the impact of motor and imitation training.

To develop embodied mirroring platforms with humanoid robots, three steps are needed: measuring the human movements, mapping them to the robot reference frame and control the robot motion. For capturing human movements, wearable sensors and marker-less vision based devices are the main options [7]. Among wearable sensors, Inertial Measurement Units (IMUs) are the most accurate and widely used technology. For tracking human movements, they require to be attached to different body segments and to measure the rotation angles between them. Instead, marker-less vision based devices rely on cameras that are able to reconstruct the 3D position of different body joints [8]. An example of such devices available on the market is the Microsoft Kinect. In its second version, the Kinect detects 2D positions of points and then uses a time-of-flight camera to calculate their depth. In the current work, two embodied mirroring setups were developed, using the two motion capture systems, as described above, and a humanoid NAO robot. A systematic comparison between the two setups was performed to evaluate the impact of using different technologies for acquiring body motion in the context of embodied mirroring for rehabilitation, taking into account usability features of the final system. Lastly, a possible protocol for an embodied mirroring game, with the Kinect, was tested, demonstrating its potential as an interactive tool in therapies for ASD children.

II. METHODS

A. Description of the setup

A NAO robot from Aldebaran Robotics [9] was used for the embodied mirroring platform. NAO is a humanoid robot with 25 Degrees of Freedom (DOF), and endowed with sensors, LEDs and loudspeakers for interaction with the environment. We focused on movements of the upper limbs and

[1]A. Geminiani, L. Santos, A. Farabbi, N. Farella and A. Pedrocchi are with the NEARLab, Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy (*corresponding author email: alice.geminiani@polimi.it).

[2]L. Santos and J. Santos-Victor are with Institute for Systems and Robotics, Instituto Superior Técnico, Lisboa, Portugal.

[3]C. Casellato is with the Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy.

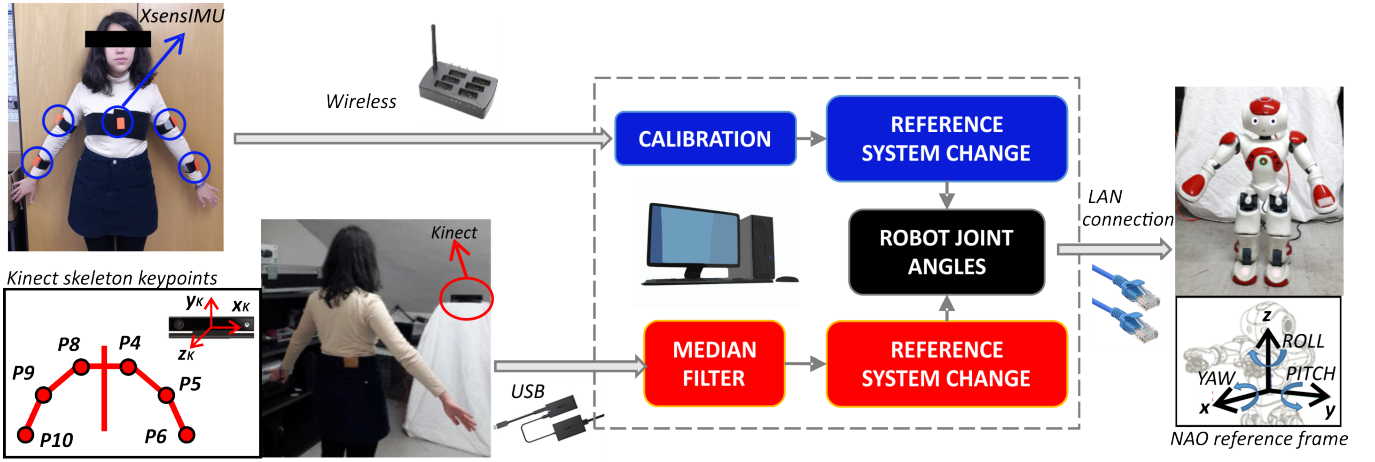[4]I. Olivieri is with the CARELab, IRCCS Fondazione Don Carlo Gnocchi, Milan, Italy.

Fig. 1: Block diagram of the embodied mirroring setups based on Inertial Measurement Unit (IMU) (top - blue) and Kinect (bottom - red) tracking systems.

chose to control 3 DOF for each limb: Shoulder Roll, Shoulder Pitch and Elbow Roll. Thus, the basic allowed movements were: flexion/extension, abduction/adduction, lateral abduction/adduction of the shoulder and flexion/extension of the elbow. To track subject movements, two systems were compared: one based on IMUs and one using the Microsoft Kinect. The subject was standing in front of NAO during the exercises. For the mirroring action, the subject's joint angles were extracted with the two tracking technologies, and provided as control signals to the NAO robot, (Fig. 1).

B. Inertial Measurement Units

For the setup based on IMUs, five motion tracking sensors from XSens Technologies were used, computing the orientation from angular velocity, acceleration and earth magnetic field through a built-in Kalman filter. IMUs were positioned on the upper limbs and chest as shown in Fig. 1. Signals from the sensors were acquired at 45 Hz. During calibration, the subject was standing, while keeping a forward flexion of both shoulders. Sensor orientation was computed with respect to the reference frame of the IMU on the chest ($^{chest}R_s$). Then, the NAO control signals were computed as the angles between limb segments, moved from the chest IMU (*chest*) to the NAO reference frame (*NAO*) (1).

$$^{NAO}R_s = {}^{NAO}R_{chest} \cdot {}^{chest}R_s \qquad (1)$$

Due to higher accuracy than the Kinect, this system was used as the *gold standard* reference for motion tracking signals.

C. Kinect

With the Microsoft Kinect, the 3D coordinates of 8 upper-limb keypoints were obtained at a frequency of 30 Hz. The signals were filtered with a median filter (5-sample window), for noise reduction. Then, two rotations were applied, one around the y-axis and one around the x-axis, to align the Kinect to the NAO coordinate frame. In the new reference frame, the shoulder joint movements were given

by the angle between the arm and the z-axis (lateral abduction/adduction) and y-axis (abduction/adduction); the elbow flexion/extension was calculated from the angle between arm and the forearm. The vectors representing the arms ($\mathbf{P_{4-5}}$ and $\mathbf{P_{8-9}}$) and forearms ($\mathbf{P_{5-7}}$ and $\mathbf{P_{9-10}}$) were obtained from the 3D coordinates. Then, the angle between two vectors was given by 2. The resulting angles (Table I) were used as control signals for the robot.

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

TABLE I: Shoulder and Elbow angles calculation with a Kinect

Angles	a	b
Right Shoulder Lateral Abduction/Adduction	$\mathbf{P_{8-9}}$	$-\mathbf{z}$
Right Shoulder Abduction/Adduction	$\mathbf{P_{8-9}}$	$-\mathbf{y}$
Right Elbow Flexion/Extension	$\mathbf{P_{8-9}}$	$\mathbf{P_{9-10}}$
Left Shoulder Lateral Abduction/Adduction	$\mathbf{P_{4-5}}$	$-\mathbf{z}$
Left Shoulder Abduction/Adduction	$\mathbf{P_{4-5}}$	$\mathbf{y}$
Left Elbow Flexion/Extension	$\mathbf{P_{4-5}}$	$\mathbf{P_{5-6}}$

D. Comparison between two tracking systems

The basic testing protocol included five acquisitions of three reference movements from a starting to a target angular position: lateral abduction/adduction (from 0 to -$\pi/2$ rad) and abduction/adduction (from $\pi/2$ to -$\pi/3$ rad) of the shoulders, flexion/extension (from 0 to $\pi/2$ rad) of the elbows. Target angles were chosen based on human and NAO expected range of motions. The selected movements allowed to test different joints and planes, and repeated three times for each acquisition. For comparison purposes, the protocol was used with two healthy subjects being tracked by the IMU and the Kinect at the same time. We chose an adult and a 11-year-old child as subjects, in view of clinical applications involving an adult therapist and a child ASD patient. The adult and the child's legal guardians gave informed consent prior to acquisitions. Due to different sampling frequencies, the Kinect signal was re-sampled and aligned to the IMU signal,

based on the maximum value of their correlation. Then, the Root Mean Square Error (RMSE) between them was calculated. Finally, other features were considered, including intrusiveness, cost and flexibility, in view of clinical use.

E. Kinect-based embodied mirroring game

Thanks to the flexibility of the Kinect-based system, an example of a turn-taking imitation game was designed and implemented, including four phases: (i) NAO demonstrating a reference movement and asking a first subject to repeat it; (ii) the first subject doing the movement while NAO is mirroring him; (iii) NAO asking a second person to repeat the movement; (iv) the second subject doing the reference movement while NAO is mirroring him.

The imitation game was tested with the same reference movements as in Section II.D. The same two subjects performed the game, positioned next to each other in front of the robot and the Kinect, to avoid occlusion problems.

III. RESULTS

The developed embodied mirroring platforms were able to successfully mimic basic movements of the upper limbs of healthy subjects. The two motion tracking systems exhibited different advantages/disadvantages in terms of accuracy and flexibility. The Kinect-based setup was chosen as the best candidate for clinical applications with ASD children, and used in the first prototype of an interactive mirroring game.

A. Validation of the IMU-based setup

The IMU-based system requires estimating the rotation matrices between each sensor and the NAO coordinate frame. We verified that the correlation between NAO control signals extracted from IMU data during two acquisitions with different sensor positioning was close to 1 (0.87 shoulder abduction/adduction, 0.92 for elbow flexion/extension), for all movements. This demonstrates the method's robustness to IMU wrong placement and trunk movements during exercise: in case of different initial positioning of the sensors, the angles calculated were the same for the same movements.

B. Comparison between two tracking systems

We took the IMU tracking system as the baseline, *gold standard*, due to its higher accuracy and lower angular errors with respect to the Kinect. Figure 2 and Table II show that the Kinect-based platform was able to mirror the subject movements with acceptable errors, compared to IMU [10]. Moreover, the cross-correlation between the Kinect and IMU signals was above 0.85 for all tested movements. In view of clinical applications with ASD children, other features of the setup were evaluated. The Kinect-based system was less intrusive, did not need any setup/calibration time and allowed recording more people at the same time, without increasing the costs. As a result, the Kinect system was chosen as the best candidate for our purposes.
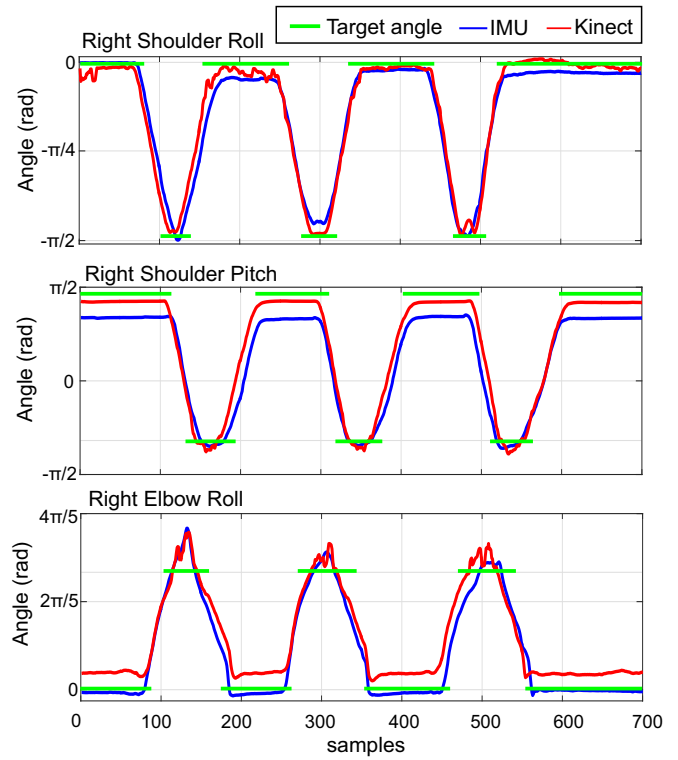

Fig. 2: Comparison between NAO control signals obtained with the two tracking systems during one example acquisition with the 3 movements of the adult left upper limb, to reach constant target angles (green lines - Par. II.D).

TABLE II: RMSE and cross-correlation between IMU and Kinect-based control signals for the adult and the child. Values are reported as mean±SD (Standard Deviation).

Movements	RMSE (rad)		Cross-Correlation	
	Adult	Child	Adult	Child
Lateral Abduction/ Adduction shoulder	0.23±0.12	0.41±0.19	0.94±0.06	0.85±0.19
Abduction/ Adduction shoulder	0.35±0.07	0.41±0.04	0.96±0.03	0.96±0.02
Flexion Extension elbow	0.31±0.10	0.30±0.19	0.94±0.04	0.91±0.13

C. Testing of the embodied mirroring game

Exploiting the capability of the Kinect-based setup to track more than one person simultaneously without increasing costs, computational load and system transparency, a first prototype of the interactive mirroring game was tested. During the session, the two subjects were successfully mirrored by the NAO, while executing upper limb movements, which were first shown by the robot (Fig. 3). The NAO mediated the interaction, guiding the turn-taking phases.

IV. DISCUSSION AND CONCLUSIONS

We developed and tested an embodied mirroring setup for therapy sessions with ASD children, to improve motor and social skills. In recent decades, robotic applications have been developed for patients with motor, social and cognitive impairments, to facilitate their participation and
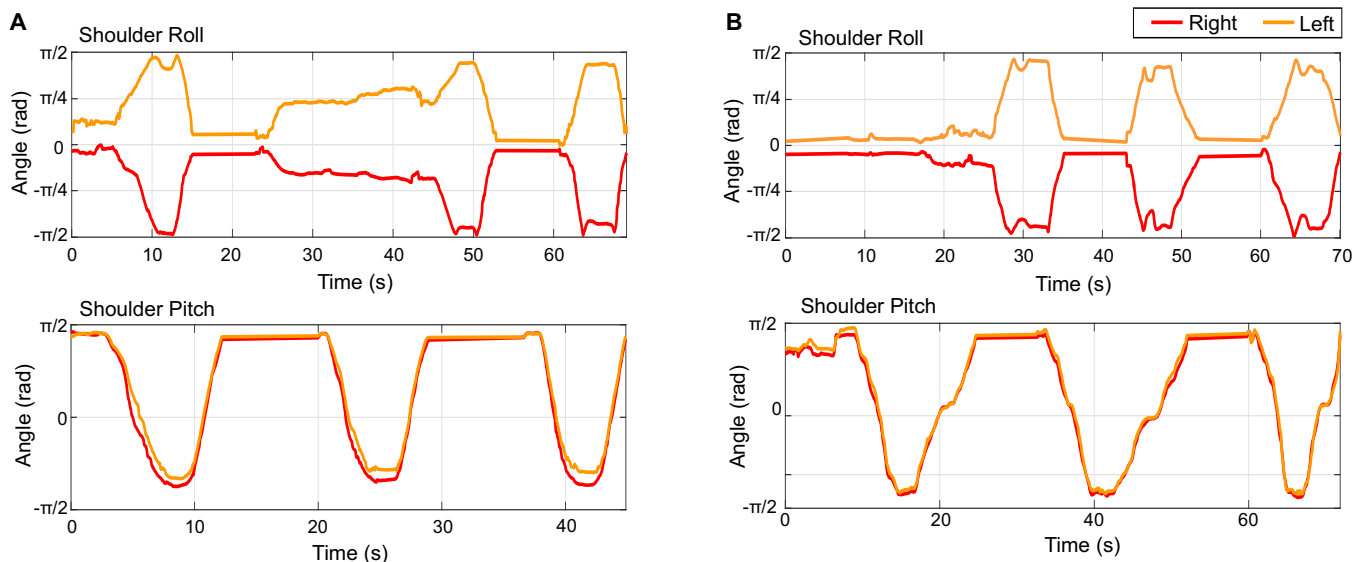
Fig. 3: Example NAO control signals in the interactive mirroring game with the two subjects, during execution of shoulder lateral abduction/adduction (top panels) and abduction/adduction (bottom panels) by the adult (A) and child (B).

engagement during therapy [11][12]. Robots have been used as coaches, teachers or assistants in daily life [13]. In case of ASD children, humanoid, child-sized and stylized robots are recommended, and non-intrusiveness and ease of setup are key [2]. Then, the link between social and motor deficits in ASD suggest that therapies should be aimed at improving both skills [3]. The platform we developed is suitable for embodied mirroring games with ASD children for several reasons: NAO robot's features perfectly match the above-cited requirements, and the Kinect tracking system successfully allowed to mirror the subject movements properly. With respect to the more accurate IMUs, the Kinect-based system has a much smaller setup time, avoiding sensor placing and calibration time, and is non intrusive. The proposed interactive game also suggests that the robotic setup has a strong potential to improve patient's social skills, mediating the interaction between two people. Future work includes the design of a feedback system, exploring NAO interactive elements, to guide movement execution and increase motivation. In order to focus also on cognitive skills, the movements will be contextualized within a semantic narrative. The final setup should be versatile and easy to use by clinicians. Motor and cognitive parameters will be used to evaluate therapy: successful execution and duration of movements will be monitored and non-functional gestures will be identified. Standard clinical scales will be used to assess impact on cognitive skills. Systematic tests with patients will be needed to eventually use the setup as therapeutic tool together with traditional exercises and games.

ACKNOWLEDGMENT

REFERENCES

[1] A. Anzulewicz, K. Sobota, and J. T. Delafield-Butt, "Toward the autism motor signature: Gesture patterns during smart tablet gameplay identify children with autism," *Sci. Rep.*, vol. 6, pp. 1–13, 2016.
[2] B. Scassellati, H. Admoni, and M. Matari, "Robots for use in autism research," *Annual Rev. Biomed. Eng.*, vol. 14, pp. 275–294, 2012.
[3] A. Cavallo, L. Romeo, C. Ansuini, J. Podda, F. Battaglia, E. Veneselli, M. Pontil, and C. Becchio, "Prospective motor control obeys to idiosyncratic strategies in autism," *Sci. Rep.*, vol. 8, 2018.
[4] C. Casellato, M. Gandolla, A. Crippa, and A. Pedrocchi, "Robotic set-up to quantify hand-eye behavior in motor execution and learning of children with autism spectrum disorder," in *IEEE ICORR*, 2017.
[5] L. Fadiga, L. Fogassi, G. Pavesi, and G. Rizzolatti, "Motor facilitation during action observation: a magnetic stimulation study," *J. Neurophys.*, vol. 73, pp. 2608–2611, 1995.
[6] P. Winkielman, E. W. Carr, G. Hofree, and L. C. Kavanagh, *The Normative Mind.* B. Broek, J. Stelmach, & . Kwiatek, 2016, ch. Imitation, Emotion, and Embodiment.
[7] M. Alibeigi, S. Rabiee, and M. N. Ahmadabadi, "Inverse kinematics based human mimicking system using skeletal tracking technology," *J. Intell. Robot. Syst.*, vol. 85, pp. 27–45, 2017.
[8] S. Spasojević, T. V. Ilić, S. Milanović, V. Potkonjak, A. Rodić, and J. Santos-Victor, "Combined vision and wearable sensors-based system for movement analysis in rehabilitation," *Methods Inf Med*, vol. 56, pp. 95–111, 2017.
[9] D. Gouaillier, V. Hugel, P. Blazevic, C. Kilner, J. Monceaux, P. Lafourcade, B. Marnier, J. Serre, and B. Maisonnier, "Mechatronic design of nao humanoid," in *IEEE Int Conf Robot Autom*, 2009, pp. 769–774.
[10] M. Huber, A. Seitz, M. Leeser, and D. Sternad, "Validity and reliability of kinect skeleton for measuring shoulder joint angles: a feasibility study," *Physiotherapy*, vol. 101, pp. 389–393, 2015.
[11] F. Lunardini, C. Casellato, A. D'Avella, T. Sanger, and A. Pedrocchi, "Robustness and reliability of synergy-based myocontrol of a multiple degree of freedom robotic arm," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, pp. 940–950, 2016.
[12] C. Casellato, A. Pedrocchi, G. Zorzi, L. Vernisse, G. Ferrigno, and N. Nardocci, "Emg-based visual-haptic biofeedback: A tool to improve motor control in children with primary dystonia," *IEEE Trans Neural Syst Rehabil Eng*, vol. 21, pp. 474–480, May 2013.
[13] E. Ambrosini, S. Ferrante, ..., and A. Pedrocchi, "A myocontrolled neuroprosthesis integrated with a passive exoskeleton to support upper limb activities," *J Electromyogr Kinesiol*, vol. 24, pp. 307 – 317, 2014.