2011 RO-MAN
20th IEEE International Symposium on Robot and
Human Interactive Communication,
July 31 - August 3, 2011, Atlanta, GA, USA

Development of a Novel Robot-Mediated Adaptive Response System for Joint Attention Task for Children with Autism

Esubalew Bekele[1,4], Uttama Lahiri[2,4], Julie Davidson[3,4], Zachary Warren[3,4], and Nilanjan Sarkar[1,2,4]
[1]*Electrical Engineering* ,[2]*Mechanical Engineering* ,[3]*Pediatrics,* [4]*AUTOS Lab,*
Vanderbilt University, Nashville, TN 37212, USA

Abstract— **With Centers for Disease Control and Prevention prevalence estimates for children with autism spectrum disorder (ASD) at 9.1 per 1,000 (1 in 110), identification and effective treatment of ASD is often characterized as a public health emergency. Emerging technology, especially robotic technology, has been shown to be appealing to these children and such interest can be harnessed to address the limitations while providing intervention services to young children with ASD. Generally the spectrum nature of autism calls for intensive, individualized intervention. However, existing robot-mediated systems tend to have limited adaptive capability that limits individualization. Our current work seeks to bridge this gap by developing a novel adaptive and individualized robot-mediated technology for children with ASD. The system is composed of a humanoid robot with its vision being augmented by several wall-mounted cameras for real-time head tracking using a distributed architecture. Based on the cues from the child's head movement, the robot intelligently adapts itself in an individualized manner to promote joint attention. The developed system is validated with two typically developing children. The validation results of the head tracker and the closed-loop nature of interaction are presented.**

I. Introduction

Autism spectrum disorder (ASD) is characterized by core deficits in social communication as well as atypical patterns of interest and activity [1]. Data from various European countries suggests a prevalence rate ranging from 0.6% to well over 1% of the population [2]. According to Centers for Disease control and prevention (CDC), an estimated 1 in 110 children in the United States has an ASD [3]. The average lifetime cost of care for individual with autism is estimated to be around $3.2 million, with average medical expenditures for individuals with ASD 4.1–6.2 times greater than for those without ASD [2]. There is a dearth of appropriate intensive intervention resources in most communities [4]. Emerging technology [5] has the potential to address these limits and may ultimately play a crucial role in providing more powerful, accessible, and individualized interventions to young children with ASD [6]. A number of recent studies investigated applications of computer technology [7], virtual reality environments [8], and robotic systems [9] to the intervention of children with ASD.

However, presently available robotic systems as applied to children with ASD are open-loop with limited adaptive capability that limits individualization of application. Preliminary results with Computer and Robot-based adaptive response technology for intervention indicates the ability of the technology to flexibly adapt to a wide variety of population while furnishing repeatable and standardized sets of stimuli [8, 9]. Thus, given the potential of intensive, individualized intervention services, the aim of our present work is to develop a novel adaptive and individualized, robot-mediated technology for ASD intervention. This will help address early intervention with children with or at-risk for ASD so as to promote intervention in core deficit areas of social communication. Our present work focuses on addressing joint attention skills, as these skills are often hypothesized to be the fundamental social communication building blocks that predict a number of neurodevelopment trajectories and outcomes related to ASD [10].

Joint attention differences are thought to have potent ramifications for development, with early joint attention difficulties predicting not only the development of ASD itself, but also difficulties in language and social outcomes for children with ASD [10]. In short, joint attention skills appear as critical ingredient for the development of social and communication skills in young children with ASD. Fundamental differences in this skill likely contribute to and/or underlie the deleterious neurodevelopmental cascade of effects associated with the disorder. To date available joint attention treatments vary widely in terms of scope and approach, from highly naturalistic techniques embedded in everyday routines to discrete trial approaches utilizing drill and practice. Literature suggests transactional approaches of joint attention that combine the advantages of developmental and discrete trial approaches via intensive graduated systems of prompts in game-like interactional frameworks, hold the most promise for improving these core skills [10]. However, these approaches are most effective when children show sustained engagement with a variety of objects (i.e., a frequent challenge for young children with ASD), can be utilized within intrinsically motivating settings, and when careful adaptation to small gains and shifts can be incorporated and utilized over longer intervals of time [11]. Given the above concerns, and the evidence that children with ASD are generally interested in robots [12 - 14], adaptive robotic technology holds great potential for creating highly motivated and controlled interactions that can be flexibly responsive to children's skills within a larger functional process. This, in turn, may promote maintenance and generalization of the skills. Specifically, robot-mediated joint attention intervention tasks may be capable of making intervention decisions based upon dynamic data analysis and subsequent adjustment in robot behavior in real-time to provide incremental shifts in reinforcement and scaffolding skill development within a transactional framework that is hard to establish in a human-centric intervention approach.

However, the robots used in earlier work [12, 14] were remotely operated and thus were not able to adapt their interactions autonomously. Subsequent findings by several investigators [1, 9, and 11] suggest that endowing a robot with the ability to automatically detect and respond to performance and attention related cues of children with ASD in an individualized manner enhances task performance as well as engagement and enjoyment of the participants themselves.

Thus the objective of this paper is to present the development of an intelligent humanoid robot-centric system capable of providing closed-loop, adaptive individualized feedback. The system is equipped with augmented vision by camera-based head tracking within a joint attention interactional framework. We present system validation results to show the operation of closed-loop nature of the interaction. A thorough study using the system with children with autism will be performed in the future. The rest of the paper is organized as follows: Section II presents the robot-mediated joint attention task architecture. Section III discusses the system validation results. Finally, Section IV summarizes the paper by drawing valuable conclusions and highlighting future directions.

II. System Architecture

The experimental setup for the robot-mediated joint attention task is shown in Fig. 1 and the system architecture is shown in Fig. 2. There is a humanoid robot that administers the joint attention task. There is one top camera mounted on the ceiling and multiple side cameras (currently three) mounted on the walls that track the head movement of the participant.

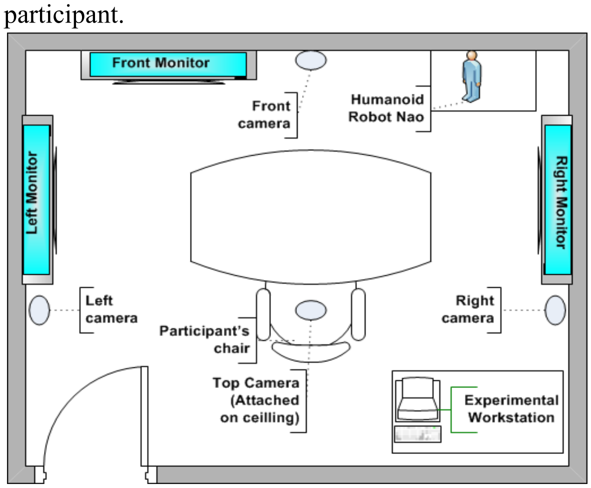

Fig. 1 The experiment room setup with the targets (LCD monitors) mounted on the walls, the top (mounted on the ceiling) and side (mounted on walls) cameras, and the humanoid robot.

The robot and each camera have their own processing modules. The system is developed using a distributed architecture with a supervisory module (i.e., a supervisor) making decision about task progression. The main sensory feedback is the head tracker, which is used for one's gaze inference.

The participant wears a hat that has an array of infrared LEDs sewn to its top and sides in straight lines. There is a supervisor that sends/receives commands and data to/from the robot as well as cameras. However, the supervisor does not centrally control the processing of the individual modules. It facilitates communication between the camera processing modules (CPM), the humanoid robot and the stimuli controllers (SC) using a network interface as depicted in Fig. 2.

In the joint attention task, the robot initiates the joint attention task (IJA) using speech as a first level of interaction. It asks the participant to look at a specific picture or a scene that is being displayed on one of the three computer monitors hung on the wall. The head tracker is triggered by the central controller as the robot issues a prompt to activate the camera processing modules for the specified trial duration to accumulate time-stamped data of the head movement to infer where the participant is looking. At the end of the trial duration, performance metrics such as effective hit counts (i.e., whether the participant actually looked at the target stimuli), duration of each hit, the latency to the first hit, and the percentage hit are computed and sent to the supervisor. The supervisor uses that information to generate rewards which are executed by the robot (such as clapping, saying "Good Job!" etc.), and/or triggers the visual stimuli themselves (e.g., the stationary picture shown on the monitor may turn into a movie clip) if the response is positive or move to the next level of prompt in case of no or inappropriate response.

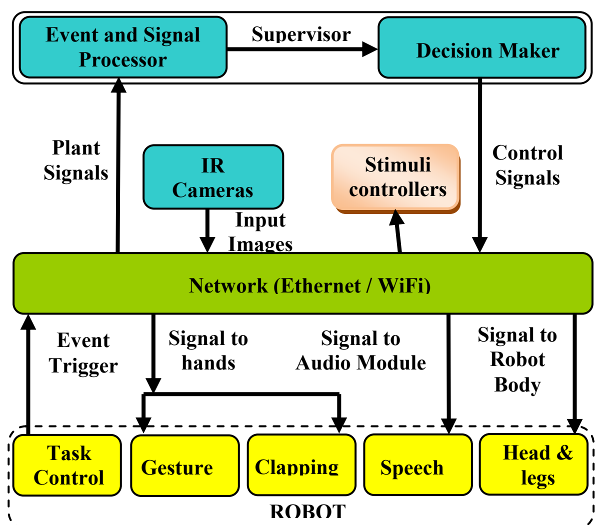

Fig. 2: System architecture showing component interaction.

A. *The Humanoid Robot, NAO*

A humanoid robot is the central component of the system architecture. It presents joint attention bids and prompts using gestures, voice commands and/or calling the participant's name. It is also responsible for presenting individualized instructions and feedback (e.g., such as applause, reward, etc.).

The humanoid robot used in this project is NAO, which is made by Aldebaran Robotics [15]. NAO is a medium childsized humanoid robot with a height of 58 cm and weight of approximately 4.3 kg (Fig. 3). Its body is made of plastic and it has 25 degrees of freedom (DOF). The robot has a variety of sensors ranging from tactile to audio microphone arrays. It also has a wide variety of actuators ranging from DC servo motors to emotion display LEDs.

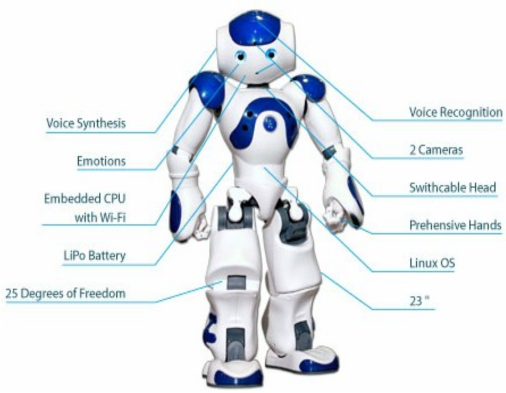

Fig. 3 The humanoid robot NAO and its capabilities

In this work, we have augmented NAO's vision by using external cameras for the head tracking for closed-loop interaction. The two CMOS vertical stereo camera sensors that NAO has are low performance CMOS sensors with frame rates of 4 - 5 frames per second (FPS) with the native resolution of 640 x 480, which was not suitable for detecting head movement in real-time.

TABLE I
LEVELS OF THE HIERARCHICAL PROTOCOL

Protocol Levels (PL)	**Robot Prompt for a child named Max**
PL 1	"Max, look" + shift (robot's) eye gaze to target
PL 2*	If NR$^{\$}$ after 5 s : "Max, look" + shift eye gaze to target
PL 3	If NR after 5s: "Max, look at that." + eye gaze shift + point to target
PL 4*	If NR after 5s: "Max, look at that." + eye gaze shift + point to target
PL 5	If NR after 5s: "Max, look at that." + eye gaze shift + point to target + audio clip sounds at target
PL 6	If NR after 5s: "Max, look at that." + eye gaze shift + point to target + audio clip sounds at target and then video onset for 30s

*PL2 and PL4 are intentionally repeated versions of PL1 and PL3, respectively. $^{\$}$NR means no or inappropriate response

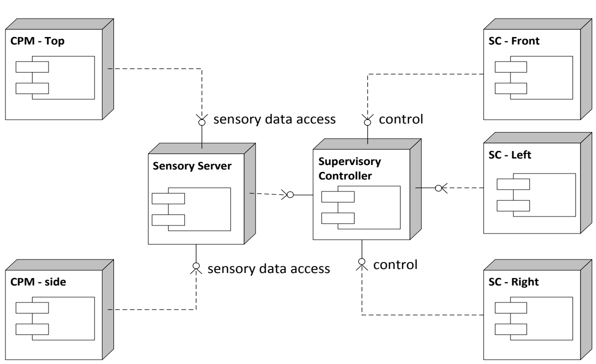

Fig. 4 Components interact via Ethernet/WiFi. Multiple Camera Processing Modules (CPM) - side are represented as one for simplicity. The stimuli controllers (SC) are displayed for each target.

The components and the algorithm used for head tracking are discussed in the following subsection. We have also developed a protocol for joint attention task-centered hierarchical prompts which is administered by the robot. Table I shows the protocol levels (PL) and their respective prompts. It gives instances of prompts for participant named Max.

The system is implemented as a component-based distributed architecture where components interact via network in real-time. Fig. 4 depicts the component dependencies and interaction. For simplicity, the robot remote module is omitted.

B. Rationale for the development of a head tracker

In JA tasks, the administrator of the task is interested in knowing whether the child looks at the desired target when prompted. Gaze is usually inferred using eye trackers, but such systems often place restrictions on head movement and detection range. Even remote non-contact commercial eye trackers that seem to avoid these limitations, are expensive, restrictive in the volume of tracking and are generally sensitive to large head movements, especially the type of head movements needed for joint attention tasks. Moreover, the eye trackers need calibration (often multiple times) with each participant and the range of the eye trackers typically requires that the participants be seated within 1-2 feet from the monitor. Given our focus on young children (i.e., 2-5 years of age) for these joint attention tasks, we wanted to develop a low-cost system that would not be limited by these factors. Joint attention tasks require large head movements and the objects of interests are typically placed between 5-10 feet from the participants. While it is necessary to know whether the child is looking at a target object in joint attention tasks, unlike typical eye tracking tasks the need to know the precise gaze coordinates is less important in JA tasks.

In order to determine whether the child is responding to JA commands, a novel near-IR light marker-based head tracker is designed using low cost off-the-shelf components for gaze inference with an assumption that gaze direction can be inferred by knowing the head direction using kinematic transformations.

C. The Head Tracker

The head tracker is composed of near-infrared cameras, arrays of IR light emitting diodes (LEDs) sewn on the top and the sides of a hat as markers, and camera processing modules (CPM). Fig. 5 below shows the experimental hat with the near-IR LEDs arrays.

Fig. 5 The hat holds the array of LEDs at side and top.

The tracker has a set of CPMs that process input images captured by the top and side cameras that are tracking the arrays of top and side LEDs on the hat, respectively. A sensory network protocol is implemented in the form of Client-Server Architecture in which each CPM is a client to a server that is monitoring them for raw time-stamped tracking data on the request from the supervisor. The server processes the raw data for the duration of a trial and produces measured performance data at the end of each trial and sends this data to a python software based client that is responsible for feedback generation and communication with NAO.

The Head Tracker has the following components:

1) IR Cameras: The IR Cameras are custom modified to see only in near infrared from cheap Logitech Pro 9000 webcams. The IR filter for the original camera is removed carefully and is replaced with a glass of similar refractive index and thickness (made from microscope slide covers and a magnetic tape of a floppy disk). The glasses are used to give the camera nearly the same focus as the original one and the magnetic disk tapes are used to suppress the visible light. Since we are interested in detecting only the IR markers on the hat worn by the participant, blocking the visible light minimizes the computational burden and reduces the possibility of false positives.

2) Camera Processing Module (CPM): These modules are equipped with light-weight contour-based image processing in the XYZ color space to detect the LEDs. Formulated based on the human eye's photo receptors (cones), the XYZ space is characterized by identifying brightness (short, medium, and long wavelength) cues better than other color spaces which is a suitable property for detection of LEDs in the IR spectrum. The tristimuli values were experimentally determined as shown below in (1) [16].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.812401 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

The detection of the LEDs is in the projective (perspective) image plane with the camera matrix and the focal length is estimated using the well-known Levenberg-Marquardt optimization algorithm.

A three dimensional (3D) approach with only top LEDs array and a stereo camera pair for 3D line reconstruction was first attempted. However, the error was found to be not acceptable for our task. A small depth difference between the two extreme points of the LEDs array creates a large angular measurement error. Subsequently, a two-view two dimensional (2D) solution (which is similar to multi-view scene reconstruction in computer graphics) with 2 side and one top LEDs arrays and multiple cameras was developed to track head movement.

The idea of a pinhole camera model and the 2D projective geometry in the image plane [17] has been extended to govern the projection of targets (flat LCD monitors in the scene which are not directly observable by the cameras). It is assumed that the perspective projection of the targets is translated and rotated on the projective image plane using the perspective projection of the coordinates from each target to the camera center.

The projections of the arrays to the top and side perspective projective image plane is shown in Fig. 6 together with the respective rays produced from one of the points with lower *y*-coordinate and extended indefinitely. When the child moves his/her head, the LEDs arrays move with the head and so do their 2D projections. Hence the ray will move around the projection plane. The intersections of rays with the projections of the targets in the top and side projection planes give the *x* and *y*-coordinates of the gaze point, respectively. Therefore, a ray-line segment intersection is performed to compute the intersection coordinate in the respective dimensions as shown in Fig. 7.

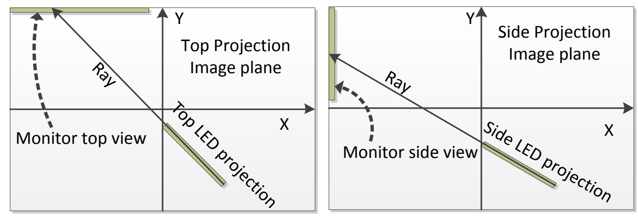

Fig. 6 Top and side perspective projections of the LED arrays and the targets (LCD monitors)

Let P_0P_1 be the projection of one of the targets on the projection plane and Q_0Q_1 be the ray formed from the LEDs projection as described above. Then, we define $\mathbf{u}$ ($u_1+\mathbf{j}u_2$), and $\mathbf{v}$ ($v_1+\mathbf{j}v_2$) to be the directional vectors for the target projection, and the LED projection ray, respectively. Vector $\mathbf{w}$ ($w_1+\mathbf{j}w_2$) is a vector from Q_0P_0. When the ray ($P(s)-Q_0$) intersects the line extended from the line segment (P_0P_1), the vector $\mathbf{v}^{\perp}$ is perpendicular to the vector $\mathbf{v}$ ($P(s)-Q_0$). This is equivalent to the perpendicular product condition [18]:

$$\mathbf{v}^{\perp}. \ (\mathbf{w}+s\mathbf{u}) = 0 \quad (2)$$

This condition can be solved to give the equation for the intersection fractions, sI, as follows.

$$sI = \frac{-v^{\perp}.w}{v^{\perp}.u} = \frac{v_2w_1 - v_1w_2}{v_1u_2 - v_2u_1} \quad (3)$$

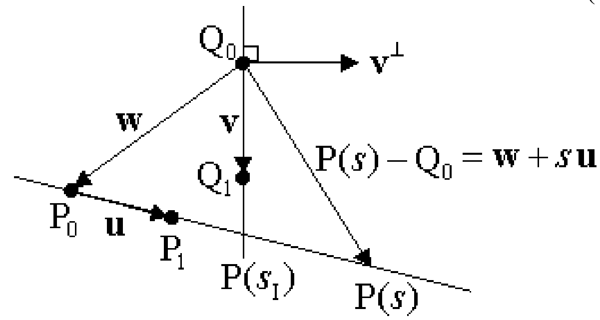

Fig. 7 Ray-line segment intersection for one target (P_0P_1) and a LEDs ray ($Q_0P(s)$)

This algorithm determines the point of intersection in each view (top and side) to get the (*x,y*) coordinates of intersection of the gaze direction vector (**v**) with the line segment (target projection to the 2D projective plane). These image projective plane coordinates are projected back to the actual lengths to each LCD monitor reference frame. This system measures the roll and the yaw angles of the head.

III. Results and Discussion

A. Validation of the head tracker

We first wanted to validate the accuracy of the developed head tracker. In order to generate ground truth data, we attached a laser pointer at the end of each IR LED array along the same straight line. The idea was to physically measure the coordinates in which the laser was pointing (ground truth) and compare them with the coordinates obtained from the head tracker. We asked a participant to look at one of the 20 points at a time distributed across a computer monitor. Each target was a 2 cm x 2 cm grid. Thus there was an inherent uncertainty of 2 cm in the ground truth coordinates.

For each target, 10 s worth of data was collected and averaged. The average error was taken as the measure of accuracy of the head tracker. The average errors for x and y coordinates were 2.6 cm and 1.5 cm, respectively. This error is tolerable for the joint attention task as the head swing is large and we are interested in capturing participant's head movement towards our targets (monitors). In our present work, we are not interested in the exact x-y coordinates of the monitor where the participant is looking at. Instead, we are interested in whether the participant hits/misses the target with certain error tolerance.

To further understand the head tracker accuracy, we analyzed the average error of all points in each row and column. Points on the same column are ideally supposed to have the same x-coordinates. However, our system was able to detect the points with an average difference of 1.2 cm. These results indicate that the system is accurate in its x coordinate measurement with an average error of 2.6 cm and its repeatability was an average of 1.2 cm.

Similarly, the y-coordinates of all points in each row should have the same value. However, our system was able to detect the points with an average difference of 2.19 cm. Thus the y-coordinate measurement (which is monitored by the side cameras) is accurate with an average error of 1.5 cm and is repeatable in measurement with 2.19 cm. The y-coordinate measurement seems more accurate but less repeatable than the x-coordinate measurement. This can be attributed to the difference in the camera geometry, optics and manual measurement error during manual initialization of the position of the targets with respect to the camera reference frames.

We also wanted to observe the effect of initialization measurement error in monitor coordinates with respect to the camera frame. We introduced an intentional position error in monitor's x-coordinate of 2.5 cm in both positive and negative x directions. We observed that the average errors in the target measurement to be 2 cm and 1.8 cm, respectively. In other words, an initialization error of 2.5 cm produced approximately 2 cm of detection error, which indicated that there was an approximately linear correlation between the initialization and measurement errors.

B. Overall system validation and performance metrics

The system is tested with a 2 and 4 years old typically developing (with no autism) participants to validate the operation of our developed system and its ability to measure the desired performance metrics. In particular, we wanted to test whether the robot could administer the JA prompts hierarchically based on real-time measurement of the head movement of the participant.

TABLE II
RESULTS OF TESTING WITH A 2 YEARS OLD TYPICALLY DEVELOPING CHILD

Trial	# of Levels	Hit Latency (ms)	Hits	Hit Duration (ms)	% of time on target
1	5	0.045	1	209.899	6.00%
2	3	0.030	1	3099.800	100.00%
3	5	0.034	1	1013.410	32.20%
4	5	0.036	1	3075.870	100.00%
5	6	1064.180	5	42.246	7.00%
				46.134	
				46.178	
				46.180	
				45.113	
6	2	0.032	1	820.112	24.56%
7	5	0.027	1	2055.180	60.35%
8	5	0.021	2	54.224	7.69%
				176.364	
9	5	0.029	2	1303.140	77.05%
				1089.280	
10	1	0.029	2	720.317	39.06%
				460.603	
11	6	N/A	0	N/A	0.00%
12	5	422.592	1	45.569	1.50%

Each participant was given 12 trials. As mentioned earlier, the robot is capable of providing 6 hierarchical prompts in each trial (Table I). From Table II, it can be seen that the 2 year old participant hit the target in 11 out of the 12 trials, i.e., a 91.7% success. However, she required 53 prompts out of the total available 72 prompts. In other words, she needed 73.6% of total available prompts. On the other hand, the 4 years old participant (refer Table III) not only hit the target in all 12 trials, i.e., a 100% success, but also required only 28 of the available 72 prompts (i.e., only 38.9% of available prompts). These results are in accord with the theory as the age increases with typically developing children, so does their joint attention skills.

TABLE III
RESULTS OF TESTING WITH A 4 YEARS OLD TYPICALLY DEVELOPING CHILD

Trial	# of Levels	Hit Latency (ms)	Hits	Hit Duration (ms)	% of time on target
1	1	0.048	2	407.759	23.44%
				309.511	
2	2	0.019	1	3068.690	100.00%
3	1	0.028	2	2114.820	80.70%
				512.879	
4	1	0.029	1	3067.100	100.00%
5	2	0.030	4	1988.070	82.46%
				43.957	
				137.151	
				361.344	
6	1	0.027	1	3073.700	100.00%
7	1	0.027	3	602.948	27.87%
				199.705	
				46.336	
8	5	200.543	1	2913.520	93.44%
9	5	629.031	2	200.764	42.37%
				1220.360	
10	5	0.024	1	3096.000	100.00%
11	1	0.021	1	1472.710	44.07%
12	3	0.018	1	3068.930	100.00%

Both participants responded promptly after they realized that they need to hit the target. This is displayed on the hit latency results. Except on a couple of trials, both participants responded within less than a second after the prompt was issued. Again, except one or two exceptional trials, they both attended the target with less in and out frequency (i.e., they looked at the target without taking their gaze off many times). The time they spent on the targets themselves measures the interest created by the context relevant video played as reinforcement after they hit the target. These are all normal responses, which can be attributed to typically developing children. We anticipate one or more of these typical behaviors to be missing in children with autism due to the nature of the disorder in impairing their joint attention skills.

IV. CONCLUSIONS

The primary contribution of the paper is to present the development of a novel closed-loop robot-mediated ASD intervention framework that has the potential to administer intervention as well as generate reinforcement strategies based on quantitative performance metrics. Robot-assisted early intervention technology with an adaptive response system will allow objective measurement of performance metrics leading to individualization of intervention while reducing the laborious tasks that are involved with intervention and diagnosis of the autism spectrum disorder. The developed head tracker and its integration with the humanoid robot provide one of the first closed-loop robot-mediated intervention platforms for autism intervention. The head tracker works at 30 fps and is capable of tracking where a subject is looking with an average error of 2.6cm and with good repeatability. The robot and the head tracker are integrated using a distributed architecture. Initial study with two typically developing children validates the system's capability in measuring several quantitative performance metrics that could be used to design new intervention strategies. We are currently recruiting children with ASD to participate in a JA study using this newly developed system.

ACKNOWLEDGEMENT

We gratefully acknowledge the Vanderbilt University Innovation and Discovery in Engineering and Science (IDEAS) grant, the National Science Foundation Grant 0967170 and the National Institute of Health Grant 1R01MH091102-01A1 that partially supported this research. However, the authors are solely responsible for the content of the article.

REFERENCES

[1] Conn K, Liu C, Sarkar N, Stone W, Warren Z: Affect-sensitive Assistive Intervention Technologies for Children with Autism: an Individual-specific Approach. In Proc. 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008), Munich, Germany, August 2008.

[2] Autism Spectrum Disorders Prevalence Rate, Autism Speaks and Center for Disease Control (CDC)- statistics, 2011

[3] Centers for Disease Control and Prevention [CDC]: Prevalence of Autism Spectrum Disorders-ADDM Network, United States, MMWR Weekly Report, 2009, 58, 1-20.

[4] Eikeseth S.: Outcome of comprehensive psycho-educational interventions for young children with autism. Research in Developmental Disabilities. 2009; 30: 158-178.

[5] Goodwin MS: Enhancing and accelerating the pace of Autism Research and Treatment: The promise of developing Innovative Technology. Focus on autism and other developmental disabilities 2008; 23:125-128.

[6] Sandall S., Hemmeter M. L., Smith B. J., & McLean M. E. (Eds.), DEC recommended practices: A comprehensive guide for practical application in early intervention/early childhood special education. Longmont, CO: Sopris West Educational Services 2005; 147–150.

[7] Liu C, Conn K, Sarkar N, Stone W: Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder. International Journal of Human-Computer Studies 2008; 66(9):662-677.

[8] Conn K, Lahiri U, Liu C, Weller R, Sarkar N, Warren Z: An Affect-Sensitive Social Interaction Paradigm Utilizing Virtual Reality Environments for Autism Intervention. HCI 2009; 703-712.

[9] Liu C, Conn K, Sarkar N, Stone W: Online Affect Detection and Adaptation in Robot Assisted Rehabilitation for Children with Autism. In Proc. 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2007), Jeju Island, Korea, August 2007.

[10] Mundy P., Stella J.: Joint attention, social orienting, and nonverbal communication in autism. In A. Wetherby & B. Prizant (Eds.). Communication and Language Intervention Series, Autism spectrum disorders: A transactional developmental perspective, Baltimore, MD: Brookes. 2000; 9: 55-77.

[11] Rani P, Liu C, Sarkar N, Vanman E: An empirical study of machine learning techniques for affect recognition in human-robot interaction. Pattern Analysis and Applications 2006; 9(1):58-69..

[12] Dautenhahn K., Chrystopher L. Nehaniv, Michael L. Walters, Robins B., Hatice Kose-Bagci, Mirza N. Assif, Mike Blow, KASPAR- a minimally expressive humanoid robot for human-robot interaction research, Applied Bionics and Biomechanics, Special Issue on Humanoid Robots, 6(3), 369–397, (2009).

[13] Kozima H., Nakagawa C., and Yasuda Y.: Interactive robots for communication-care: A case-study in autism therapy. IEEE International Workshop on Robot and Human Interactive Communication Nashville, TN. 2005; 341-346.

[14] Billard A., Robins B., Nadel J., Dautenhahn K., Building Robota, a Mini-Humanoid Robot for the Rehabilitation of Children with Autism, RESNA *Assistive Technology Journal,* 2006

[15] Aldebaran Robotics Company, URL: www.aldebaran-robotics.com

[16] Smith T.; Guild J., The C.I.E. colorimetric standards and their use, Transactions of the Optical Society, 1931

[17] Hartley R., and Zisserman A., Multiple View Geometry in Computer Vision: second edition, Cambridge University Press, 2003

[18] Brandt J., and Schneider K., Using Three-Valued Logic to Specify and Verify Algorithms of Computational Geometry, Formal Methods and Software Engineering, Lecture Notes in Computer Science, Springer, Volume 3785, pages 405-420, 2005.