

Comparing the Gaze Responses of Children with Autism and Typically Developed Individuals in Human-Robot Interaction

S. Mohammad Mavadati¹, Huanghao Feng¹, Anibal Gutierrez² and Mohammad H. Mahoor¹

Abstract—Robots are becoming a part of humans' social life as assistants, companionbots, therapists, and entertainers. One promising application of the socially assistive robots is in autism therapy, where robots are employed to enhance verbal and non-verbal skills (e.g. eye-gaze attention, facial expression mimicry) of individuals with Autism Spectrum Disorder (ASD). One important question is “How the gaze responses of individuals with ASD differ from that of Typically Developing (TD) peers when interacting with a robot?” We present the results of our recent studies for modeling and analyzing the gaze pattern of children with ASD when they interact with a robot called NAO. This paper reports the differences of gaze responses of TD and ASD group in two conversational contexts: *Speaking* versus *Listening*. We used Variable-order Markov Model (VMM) to discover the temporal gaze directional patterns of ASD and TD groups. The results reveal that the gaze responses of the TD individuals in speaking and listening contexts, can be best modeled by VMM with order zero and three, respectively. As we expected, the result show that the temporal gaze patterns of typically developed children are varying when the role in the conversational context is changed. However for the ASD individuals for both conversational contexts the VMM with order one could best fit the data. Overall, the results conclude that VMM is a powerful technique to model different gaze responses of TD and ASD individuals in speaking and listening contexts.

I. INTRODUCTION

Socially Assistive Robotics (SAR) [9] is an emerging research field that aims to enhance the humans' social responses and improve the engagement level of individuals in the community. In other words SAR uses robots, either solely or in conjunction with caregivers, to improve the social skills and wellbeing of individuals who have difficulties in their social behaviors. The social disabilities may be related to depression, stroke, or autism. Socially assistive robots hold great promise to positively impact therapies for many of these maladies, although the focus of this paper is on the use of robots in the field of autism.

Autism Spectrum Disorders (ASD) refers to a spectrum of complex developmental brain disorders causing qualitative impairments in social interaction and the presence of repetitive and stereotyped behaviors [5]. Individuals with ASD experience deficits in appropriate verbal and nonverbal communication skills including motor control, emotional facial expressions, and eye gaze attention. One critical deficit that is common among the ASD population is difficulty

¹ S. M. Mavadati, H. Feng, M.H. Mahoor are with the Department of Electrical and Computer Engineering, University of Denver {smavadat, huanghao.feng, mmahoor}@du.edu ²A. Gutierrez is with the Department of Psychology, Florida International University {anibal.gutierrez}@fiu.edu

in detecting other's eye gaze and maintaining their own eye gaze attention appropriately. These deficits often pose problems in individuals' ability to establish and maintain social relationship that may lead to anxiety, depression, and avoidance of social interactions [24].

Currently, the majority of clinical and therapeutic treatments employ human-based behavioral and educational programs to teach individuals with ASD appropriate social and behavioral skills. However, recent studies show that individuals with ASD are more comfortable using technologies (e.g. computers, iPad, robots) than interacting with humans, because they find robots to be more predictable and less complicated than human responses [25]. This evidence has encouraged several researchers to incorporate computers and robots in interventions for individuals with ASD and to investigate how ASD (and TD) individuals are performing in different social situations.

For over a decade, different robotic groups have investigated the responses of individuals with ASD to robots [21], [22], [8] and the use of robot-based interactions have shown several practical advantages for assisting individuals with ASD [14]. For instance, compared to animal assistants (e.g., service dogs), robots need less expensive training procedures. In addition, robots can provide an interface as an embodied agent, which does not exist in computer programs, iPads, or other electronic devices. In cases when a human-based therapeutic interaction is not accessible or feasible, robot interactions could be a valuable solution. However very few studies investigated the parameters and conditions that can make the robot interactions beneficial, especially for therapeutic applications.

Generally speaking, recent research suggests that children with ASD exhibit certain positive social behaviors when interacting with robots compared to their peers who do not interact with robots [18], [5], [20], [10], [25], [11]. These positive behaviors include showing emotional facial expressions (e.g., smiling), gesture imitation, and eye gaze attention. These investigations suggest that interaction with robots may be a promising intervention approach for children with ASD. For example, a social robot can be used to teach children with autism to regulate their eye gaze attention or recognize basic facial expressions.

There have also been a few studies that illustrated the effectiveness of machine-based therapy sessions compared to the conventional sessions using human therapists [18], [20]. For instance, Moore and Calvert [18] demonstrated that a computer-based program can be more effective than the human-based therapy sessions for teaching new vocabulary.

ies to children with autism, while keeping them motivated and attentive. In addition, Robins et al. [20] revealed that children with ASD have superior engagement to robot-based interaction than human interaction, and can effectively trigger positive behaviors such as joint attention and eye gaze attention.

Eye gaze and eye contact during social interactions provide communicative partners valuable information about the direction of attention and emotional and mental states of the individual they are socially engaged with. Specifically in autism research, eye gaze responses and attention levels are important cues used to identify and model ASD characteristics due to known issues with eye gaze in the population [19], [23], [3], [17]. Peirce et al. [19] showed that toddlers with ASD (as young as 14 months) spent significantly more time fixated on dynamic geometric images than the TD group. This gaze fixation signature can be used to detect the infants and toddlers at risk of autism. Shic et al. [23] utilized eye tracking tools to examine the scanning patterns of TD and ASD toddlers (2-4 years old) as they viewed static images of faces. Their study demonstrated the differential pattern of both fine attention (towards specific regions of the face) as well as gross attention (looking at the face at all) for two groups of children. Alie et al. [3] analyzed face-to-face interactions of infants (6 months of age) with their parents. They described the gaze patterns of TD and ASD infants using Markov models. Mavadati et al. [16], analyzed the gaze fixation duration and shifting frequency of young children (7-17 years old with High Functioning Autism (HFA)) while interacting with NAO. They also provided the percentage and entropy of gaze shifting of TD and ASD children during human-robot interactions.

Despite all these positive signs and evidence, research on SAR for autism therapy is in its infancy and more fundamental studies need to be carried out before the field can deploy SAR and study its effectiveness and clinical impact on autism therapy. One important un-addressed question for the autism research is “How individuals with ASD regulate their eye gaze when interacting with a robot?” and also “In what sense are the gaze responses of ASD and TD individuals are similar or distinguishable?”

This paper aims to utilize the machine learning techniques to model the gaze responses of children with ASD and compare them with their typically developing peers when interaction with NAO. In our ongoing multidisciplinary research study we employ the Variable-order Markov Model (VMM) to determine the importance of history of gaze direction (i.e. memoryless system vs. system with memory) while a child is either speaking or listening to a robot. The following sections introduce the VMM algorithm and explain the experimental settings and outcomes of our study in more detail.

II. MARKOV-BASED MODELING AND GAZE ANALYSIS

A. Markov Modeling

Markov modeling is a powerful stochastic technique that is widely used to represent and model time-sequential data. The learning ability of the Markovian models (e.g. Hidden Markov Model-HMM) has inspired several researchers to apply them for different data analysis and machine learning applications, such as automatic speech recognition [12] and human activity recognition [26].

Hidden Markov model is the most common Markovian model which contains a set of hidden states (X_i) that produce a symbol (Y_i) at each time. In the HMM algorithm, the output of the model only depends on the current state of the system so the previously observed symbols would not have any effects on the output (i.e. memoryless system). However, there are some other Markov modeling techniques that consider the current and previous states of a system to specify the output (e.g. Variable-order Markov Model).

As the gaze direction of an individual can be highly correlated with the history of gaze direction responses, this paper utilizes VMM to evaluate the effect of gaze direction history on modeling the gaze of both TD and ASD groups. Therefore, we trained two VMMs to describe the gaze responses of children. These two models are utilized in analysis of the gaze responses of each group in speaking and listening social contexts.

B. Variable-order Markov Model

In contrast to the Markov chain models where each variable in a sequence depends on a fixed number of random variables, in VMM the number of conditioning random variables may vary based on the observed sequence. This observed sequence is often called the “context”. The main advantage of VMM models is that the order may vary for different sequences, depending on their statistics. As a result, VMM produces a probability “tree” with the null space at the root and the element of the sequence as the leaves of the tree. Figure 1 illustrates a second-order VMM tree with five alphabets $\{a, b, c, d, r\}$. Considering the training sequence q_1 , this tree can represent the probability of each letter, given the current context of the system.

In order to see how the VMM algorithm is trained, let us assume Σ is a finite alphabet (e.g. $\Sigma = \{A, B, \dots, P\}$). Given a training sequence $q^n = q_1 q_2 \dots q_n$, where $q_i \in \Sigma$, VMM aims to learn a model \hat{P} , that can assign the probability of observing any sequence of symbols. Mathematically speaking [6], for any context $s \in \Sigma^*$, and a symbol $\sigma \in \Sigma$, the model generates a conditional probability distribution for $\hat{P}(\sigma|s)$. In VMM the prediction stage utilizes *average log-loss* $L(\hat{P}, x_1^T)$ of $\hat{P}(.)$ with respect to the test sequence $x^T = x_1 x_2 \dots x_T$ (T is the length of observed sequence),

$$L(\hat{P}, x_1^T) = -\frac{1}{T} \sum_{i=1}^T \log_2 \hat{P}(x_i | x_1, \dots, x_{i-1}) \quad (1)$$

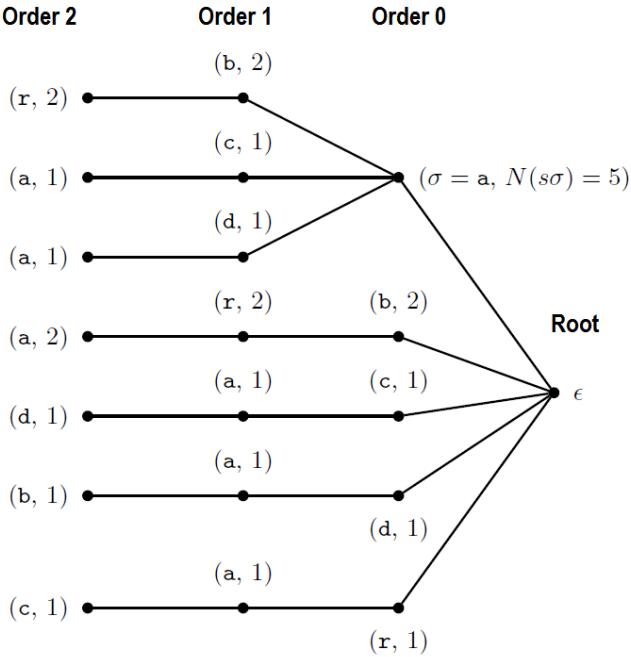


Fig. 1. The VMM tree constructed using the training sequence $q^1 = abracadabra$ by PPM approach [6].

There are several different VMM methods that can be applied for classifying a sequence of data [6]. *Prediction by Partial Match* (PPM) is an algorithm that has a robust and reliable functionality. PPM is a finite-context statistical modeling technique that can be interpreted as a mixture of several fixed-order context models of order k [7]. In other words, PPM is a combination of fixed-order context models with different values of k , ranging from zero to a pre-determined maximum (D). For each model, PPM keeps track of the length- k sequence of all characters that have been observed so far in the training sequence.

In addition, PPM can handle the zero frequency problem using *escape* mechanism [6]. In the escape mechanism, the goal is to determine the probability of unseen sub-sequence of symbols after the context s has been seen in the training sequence. In other words for each context s of length $k \leq D$, the probability of $\hat{P}_k(Esc|s)$ is allocated for all symbols that have not appeared after the context s . For instance in Figure 1, the probability $\hat{P}(d|ra) = \hat{P}(\text{escape}|ra) \cdot \hat{P}(d|a) = \frac{1}{2} \cdot \frac{1}{3+4} = 0.07$. More details about training the VMM model is described in [6].

III. GAME DESIGN AND GAZE ANNOTATION

In our study, 14 children ages 7-17 years old (Mean= 11 years) diagnosed with high-functioning autism (HFA; ASD associated with overall normal intelligence [15]) and seven TD children in the same age range were recruited as the control group. All of the parents signed the IRB informed consent forms.

We created social games to provide participants with a social context in which to interact with NAO, an autonomous, programmable humanoid robot from Aldebaran Robotics

[1]. NAO's well-designed interface (see Figure 2) and on-board multimedia system (e.g. text-to-speech synthesizer, visual data capturing and analyzing system) enabled us to specifically design social and behavioral games in which both groups of children could participate. Throughout these games, the participants were given the opportunity to engage in different social responses such as eye-gaze and joint attention, facial expression recognition and imitation. Every subject in our study participated in three sessions, and in each session s/he was spending around 30 to 45 minutes interacting with NAO.

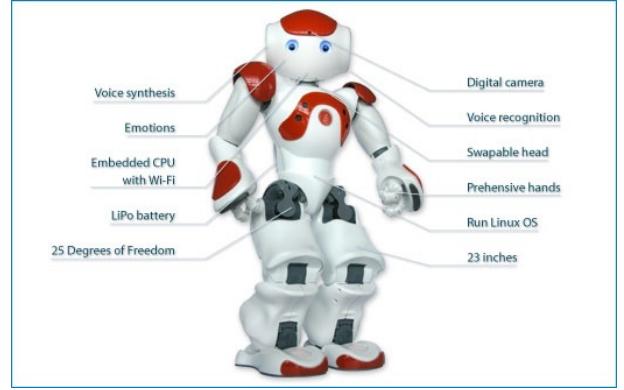


Fig. 2. NAO Platform: mechanical, Audio and visual capabilities [1].

All the sessions for this study were held in a $5 \times 5 \times 3$ meter room where four cameras recorded every interaction between the robot and the participant. These cameras captured high quality videos (HD-720P, 30 frames per second (fps)) and were connected to a recording system outside the room with an LCD screen, which allowed parents to watch the child's interaction with NAO. Figure 3 demonstrates four synchronized captured frames using four HD cameras in the room. This paper used the frontal camera to annotate the gaze responses of a child when s/he interacted with NAO.



Fig. 3. Sample images and cameras setup

The literature reveals the eye gaze is an important indicator for measuring the engagement and interest levels in conversational contexts [13]. Studies indicate that typically developing individuals look to their conversational partners nearly twice as much while listening than speaking, 75% of the time as opposed to 41%, respectively [4]. However, individuals with ASD show less interest in making eye contact and engaging in appropriate joint attention behavior with their communicative partners.

Given the importance of eye gaze direction as a non-verbal cue for regulating the flow of dyadic conversation, this paper analyzes the gaze direction of TD and ASD children while interacting with the robot. In this study, we focused on a series of conversational games in which NAO and the child verbally interacted. These games provided ideal interaction sessions for analyzing the gaze responses of children when they were listening or speaking to NAO.

To determine the gaze patterns of individuals in speaking and listening contexts, we manually labeled the videos into two categories:

- **Child Listening:** In this segment, NAO talks about different stories which came from children's story books as the child listens to NAO.
- **Child Speaking:** During this segment, NAO asks the child to share a fun story that s/he has experienced in the last week and gives enough time for the child to respond and explain their story.

After capturing the videos, both "Child Listening" and "Child Speaking" segments were extracted. Thereafter, every frame of the video that belonged to "Child Listening" (74 minutes of video in total) and "Child Speaking" (18 minutes in total) segments were manually coded off-line. By using the Continuous Measurement System [2], we labeled every video frame either as 0 (look away) or 1 (look at). Capturing data at 30 fps appeared to be beyond the gaze switching frequency; therefore, to analyze the eye gaze directions, we down-sampled the videos to 7 fps to create time units that more closely matched the observed rate of gaze switching. We utilized these coded sequences of eye gaze along with Markov-based modeling to analyze the gaze patterns.

IV. EXPERIMENTAL RESULTS

Markov models specify a probabilistic output for a given context, therefore training a model with binary letters can be a limiting factor as it can only provide two possible states for each discrete step. In order to code the gaze direction, we used a "phrase alphabet" whose "letters" represent binary sequences of a certain length. For this scheme, we defined the parameter "T" that defines the length of a letter such that each letter corresponds to one of the 2^T possible binary combinations. In a simplistic example where $T = 4$ (therefore using 16 letters), the sequence 0000 would be re-coded as A, and 1111 would be re-coded as P. Therefore for a given data string, it would be broken down into sequences of four frames which would be translated to their corresponding letters. In this way, the data string for each subject was

translated from its original form into a sequence of representative letters. In our experiments we selected 4 consecutive frames to assign $2^T = 16$ symbols (i.e. {0000, ..., 1111}) or equivalently $\Sigma = [A, B, \dots, P]$. Figure 4 demonstrates a sequence of gaze labels and the corresponding phrase alphabets for a set of consecutive video frames.

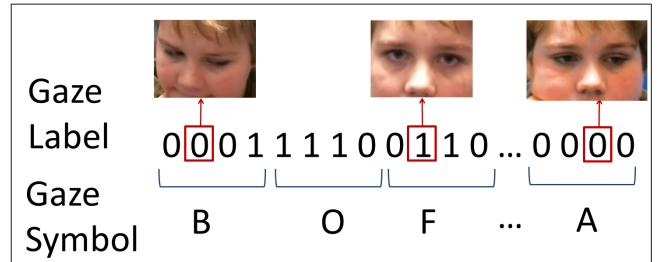


Fig. 4. Gaze labels (gaze at NAO (1) vs gaze averted (0)) and the corresponding gaze symbols representation ($n = 4$).

In order to recognize the gaze responses of the TD and ASD groups, we trained one model for the TD group and another model for the ASD group. Therefore, for any given test sample, using Equation 1 we can measure the similarity of the given test sequence to both models. The test sample is assigned to a class that has the higher similarity value. We measured the overall classification rates and reported the gaze recognition rate of VMM with different orders. Below we explain the proposed algorithm in more detail.

After defining a sequence of gaze letters, we trained the VMM to represent the gaze patterns of the TD and ASD groups. To learn the statistical probability of each alphabet in VMM, we applied the PPM algorithm, described in Section II-B. To find out how the history of gaze direction would play a role in modeling the gaze response of individuals, various VMMs with different orders have been trained and tested. We employed VMM with order zero (memoryless system) as well as the higher order Markov models (system with memory) to study the influence of gaze direction history in categorizing the gaze patterns.

We trained two separate VMM models (one for the ASD and one for the TD group) that describe the gaze patterns of individuals in the ASD group and the TD group. We used Leave One Subject Out (LOSO) cross validation technique to train a Markovian model for each group. In other words, for training the ASD and TD VMM models, we used all the data from all the subjects but one. The excluded subject was used for testing how the trained models for the ASD and TD groups can represent the gaze of each class. Hence, the similarity of the test sample is measured with respect to both the ASD and TD models and the sample is assigned to a model (class) which has the higher similarity value. The same procedure will be applied to all the participants (both ASD and TD) and we reported the average recognition rates for each group.

To evaluate the efficiency of each model, we conducted a log-likelihood similarity measurement (Equation 1), which specifies the probability of occurrence for the given sequence

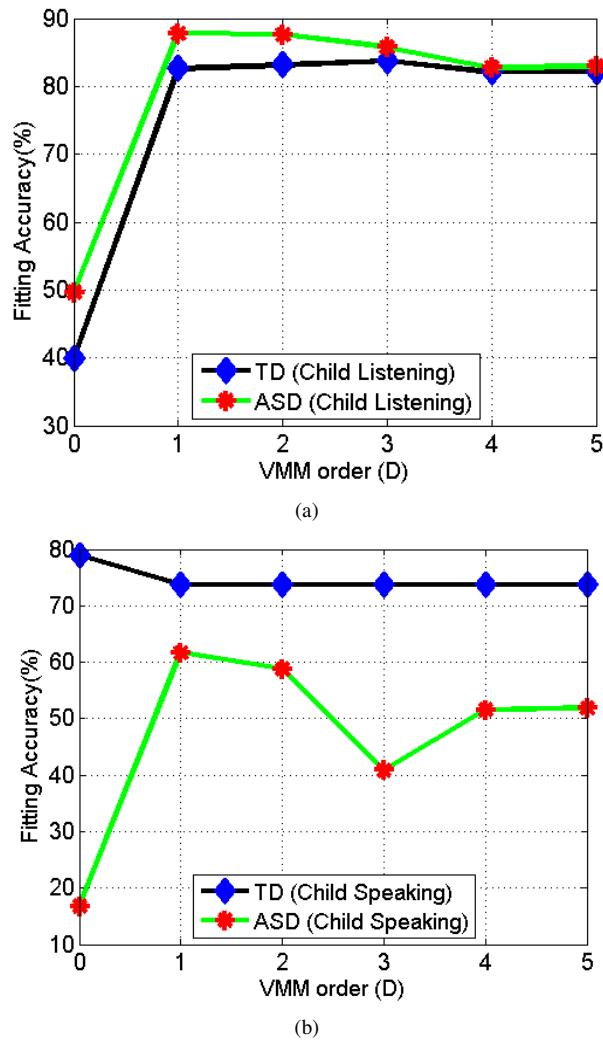


Fig. 5. Comparing Gaze Patterns of TD and ASD groups: (a) Listening Context, (b) Speaking Context.

sample for both ASD and TD models. Using the log-likelihood measure allows us to assign the test sequence to a model that has the higher similarity value. We repeated this procedure for every sequence, and the overall model fitting accuracy for both *Listening* and *Speaking* segments were reported. To measure the importance of history of gaze responses, different orders of VMM (i.e. $D = \{0, \dots, 5\}$) were studied. Figure 5 illustrates the effects of VMM order on classifying the gaze responses.

The results reported in Table I demonstrate that, VMM with order one ($D = 1$) can produce the best accuracy for the ASD group. For the TD group, in the child speaking segment the VMM with order one ($D = 1$) and for the speaking context VMM with order three ($D = 3$) provides the best results. These results can help us to discriminate between the gaze patterns of ASD and TD groups in different conversational contexts. The reported values confirm that in the “speaking context”, the memoryless model ($D = 0$) has higher accuracy for modeling the gaze responses of the TD group, although the VMM with order one ($D = 1$) has

better performance for the ASD group. In addition, for the “listening context” the gaze patterns of TD modeled well by higher order VMM ($D = 3$), but for the ASD individuals, VMM with order one ($D = 1$) can represent them accurately.

TABLE I

TYPICALLY DEVELOPED GAZE PATTERN: FITTING RATES USING VMM

Social Context of the Game	VMM Results - Fitting Rates					
	D=0	D=1	D=2	D=3	D=4	D=5
TD (Speaking)	78.95	73.68	73.68	73.68	73.68	73.68
TD (Listening)	39.89	82.58	83.15	83.71	82.02	82.02
ASD (Speaking)	16.81	61.74	58.84	40.87	51.59	51.88
ASD (Listening)	49.64	87.81	87.66	85.78	82.80	82.95

Comparing the gaze responses of ASD and TD individuals in conversational contexts confirms that, the gaze responses vary between the two groups. In other words, different orders of VMM need to be used for modeling the gaze of ASD and TD groups. More specifically, as the TD children are speaking, the memoryless model can represent their gaze patterns with high accuracy (78%). However, when the TD children are listening to NAO, higher orders of VMM ($D = 3$) would best describe the gaze responses. This reveals that as TD individuals are speaking or listening their gaze patterns vary depending on their role in the conversation. However for the ASD group, we observed that for both speaking and listening contexts, VMM with order one ($D = 1$) can model the gaze responses well. The results confirm the gaze responses of the ASD group during the speaking and listening situations is not varying significantly and the same order of VMM can model both conversational situations.

V. DISCUSSION AND CONCLUSION

This paper studied and modeled the gaze behaviors of ASD and TD children while socially interacting with a humanoid robot (NAO). In our study, we utilized Variable-order Markov Models as a powerful Markovian model to represent and temporally analyze the dynamics of eye gaze of the ASD and TD groups. The experimental results demonstrate that VMM can describe and discriminate between the gaze patterns of ASD and TD individuals in speaking and listening contexts. The results demonstrate that the TD group has different gaze patterns in speaking and listening context and their gaze responses would vary significantly depending on their role in conversational context. However for the ASD group, in listening and speaking contexts, VMM with order one ($D = 1$) can model the eye gaze of the ASD group accurately. The results validate the different characteristics of gaze patterns of ASD and TD children in two distinct social contexts (i.e. speaking and listening) and demonstrate that Markovian models can represent both the ASD and TD gaze responses reliably.

VI. ACKNOWLEDGMENT

This research work is supported by grant IIS-1450933 from the National Science Foundation.

REFERENCES

- [1] "http://www.aldebaran.com/en".
- [2] "http://measurement.psy.miami.edu/cms.phtml".
- [3] D. Alie, M. H. Mahoor, W. I. Mattson, D. R. Anderson, and D. S. Messinger. Analysis of eye gaze pattern of infants at risk of autism spectrum disorder using markov models. In *Applications of Computer Vision (WACV), 2011 IEEE Workshop on*, pages 282–287. IEEE, 2011.
- [4] Michael Argyle and Mark Cook. Gaze and mutual gaze. 1976.
- [5] S. Baron-Cohen. The cognitive neuroscience of autism. *Journal of Neurology, Neurosurgery & Psychiatry*, 75(7):945–948, 2004.
- [6] R. Begleiter, R. El-Yaniv, and G. Yona. On prediction using variable order markov models. *J. Artif. Intell. Res.(JAIR)*, 22:385–421, 2004.
- [7] J. G. Cleary and W. J. Teahan. Unbounded length contexts for ppm. *The Computer Journal*, 40(2 and 3):67–75, 1997.
- [8] Joshua J Diehl, Lauren M Schmitt, Michael Villano, and Charles R Crowell. The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Research in autism spectrum disorders*, 6(1):249–262, 2012.
- [9] David Feil-Seifer and Maja J Matarić. Defining socially assistive robotics. In *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pages 465–468. IEEE, 2005.
- [10] David Feil-Seifer and Maja J Matarić. Toward socially assistive robotics for augmenting interventions for children with autism spectrum disorders. In *Experimental robotics*, pages 201–210. Springer, 2009.
- [11] Huanghao Feng, Anibal Gutierrez, Jun Zhang, and Mohammad H Mahoor. Can nao robot improve eye-gaze attention of children with high functioning autism? In *Healthcare Informatics (ICHI), 2013 IEEE International Conference on*, pages 484–484. IEEE, 2013.
- [12] Bing Hwang Juang and Laurence R Rabiner. Hidden markov models for speech recognition. *Technometrics*, 33(3):251–272, 1991.
- [13] Adam Kendon. Some functions of gaze-direction in social interaction. *Acta psychologica*, 26:22–63, 1967.
- [14] Elizabeth S Kim, Rhea Paul, Frederick Shic, and Brian Scassellati. Bridging the research gap: Making hri useful to individuals with autism. *Journal of Human-Robot Interaction*, 1(1), 2012.
- [15] Ami Klin, FR Volkmar, SS Sparrow, DV Cicchetti, and BP Rourke. Validity and neuropsychological characterization of asperger syndrome: Convergence with nonverbal learning disabilities syndrome. *Journal of Child Psychology and Psychiatry*, 36(7):1127–1140, 1995.
- [16] S. M. Mavadati, H. Feng, Gutierrez A. Silver, S., and M. H. Mahoor. Children-robot interaction: Eye gaze analysis of children with autism during social interactions. *Int. Meeting for Autism Research*, 2014.
- [17] S Mohammad Mavadati, Mohammad H Mahoor, Kevin Bartlett, and Philip Trinh. Automatic detection of non-posed facial action units. In *Image Processing (ICIP), 2012 19th IEEE International Conference on*, pages 1817–1820. IEEE, 2012.
- [18] Monique Moore and Sandra Calvert. Brief report: Vocabulary acquisition for children with autism: Teacher or computer instruction. *Journal of autism and developmental disorders*, 30(4):359–362, 2000.
- [19] Karen Pierce, David Conant, Roxana Hazin, Richard Stoner, and Jamie Desmond. Preference for geometric patterns early in life as a risk factor for autism. *Archives of General Psychiatry*, 68(1):101–109, 2011.
- [20] B. Robins, P. Dickerson, P. Stribling, and K. Dautenhahn. Robot-mediated joint attention in children with autism. *Interaction studies*, 5(2), 2004.
- [21] Brian Scassellati. Quantitative metrics of social response for autism diagnosis. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on*, pages 585–590. IEEE, 2005.
- [22] Brian Scassellati, Henny Admoni, and Maja Matarić. Robots for use in autism research. *Annual Review of Biomedical Engineering*, 14:275–294, 2012.
- [23] F. Shic, K. Chawarska, J. Bradshaw, and B. Scassellati. Autism, eye-tracking, entropy. In *Development and Learning, 2008. ICDL 2008. 7th IEEE International Conference on*, pages 73–78. IEEE, 2008.
- [24] Lindsey Sterling, Geraldine Dawson, Annette Estes, and Jessica Greenson. Characteristics associated with presence of depressive symptoms in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 38(6):1011–1018, 2008.
- [25] A. L. Wainer and B. R. Ingersoll. The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 5(1):96–107, 2011.
- [26] J. Yamato, J. Ohya, and K. Ishii. Recognizing human action in time-sequential images using hidden markov model. In *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on*, pages 379–385. IEEE, 1992.