HHS Public Access
Author manuscript
Proc 2023 7th Int Conf Med Health Inform ICMHI 2023 (2023). Author manuscript; available in PMC 2024 May 02.

Published in final edited form as:
Proc 2023 7th Int Conf Med Health Inform ICMHI 2023 (2023). 2023 May ; 2023: 55–60. doi:10.1145/3608298.3608309.

Predicting Autism from Head Movement Patterns during Naturalistic Social Interactions

Denisa Qori McDonald,
Children's Hospital of Philadelphia Philadelphia, PA, USA

Ellis DeJardin,
Children's Hospital of Philadelphia Philadelphia, PA, USA

Evangelos Sariyanidi,
Children's Hospital of Philadelphia Philadelphia, PA, USA

John D. Herrington,
Children's Hospital of Philadelphia Philadelphia, PA, USA

University of Pennsylvania Philadelphia, PA, USA

Birkan Tunç,
Children's Hospital of Philadelphia Philadelphia, PA, USA

University of Pennsylvania Philadelphia, PA, USA

Casey J. Zampella,
Children's Hospital of Philadelphia Philadelphia, PA, USA

Robert T. Schultz
Children's Hospital of Philadelphia Philadelphia, PA, USA

University of Pennsylvania Philadelphia, PA, USA

Abstract

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterized in part by difficulties in verbal and nonverbal social communication. Evidence indicates that autistic people, compared to neurotypical peers, exhibit differences in head movements, a key form of nonverbal communication. Despite the crucial role of head movements in social communication, research on this nonverbal cue is relatively scarce compared to other forms of nonverbal communication, such as facial expressions and gestures. There is a need for scalable, reliable, and accurate instruments for measuring head movements directly within the context of social interactions. In this study, we used computer vision and machine learning to examine the head movement patterns of neurotypical and autistic individuals during naturalistic, face–to–face conversations, at both the individual (*monadic*) and interpersonal (*dyadic*) levels. Our model predicts diagnostic status using dyadic head movement data with an accuracy of 80%, highlighting the value of head movement as a marker of social communication. The monadic data pipeline had lower

mcdonalddq@chop.edu .

accuracy (69.2%) compared to the dyadic approach, emphasizing the importance of studying back-and-forth social communication cues within a true social context. The proposed classifier is not intended for diagnostic purposes, and future research should replicate our findings in larger, more representative samples.

Keywords

video analysis; bag-of-words approach; head movement patterns; dyadic features; monadic features; behavioral analysis; conversation analysis; non-verbal communication

1 INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental condition defined by difficulties in three social areas: (i) social–emotional interaction, (ii) nonverbal communication, and (iii) forming and maintaining relationships [3]. Prominent early work [25, 32] suggested that differences in nonverbal communication skills are a core trait of young children diagnosed with autism, and this finding has been substantiated by decades of further behavioral and neuroimaging research [1, 12, 27, 30, 35, 36], across several nonverbal communication modalities including eye gaze, facial expression, body actions, and head movements [26, 31, 37, 39]. There remains a pressing need to more precisely characterize nonverbal communication patterns in autism across individuals, time, and contexts, in order to better characterize its phenomenology, and continue to develop effective intervention strategies focused on enhancing nonverbal communication [2, 9, 15].

While effective nonverbal communication involves the simultaneous use of multiple cues [14], in this work, we focus on one relatively understudied nonverbal domain: use of head movements during conversations. Head movements are used in everyday communication to convey meaning, signal engagement and turn–taking, and provide structure in social interactions [6, 23]. Variations in head movement can reveal information about mental and emotional states [19, 21] and emotional dynamics between interaction partners [16, 17]. Converging evidence suggests that autistic children exhibit differences in head movement features compared to neurotypical children, with some of these differences emerging in infancy, such as head lag [7, 10]. Autistic toddlers show higher rates, acceleration, and complexity in head movements while watching movies, compared to neurotypical toddlers [20]. Older autistic children also exhibit differences in lateral head displacement and velocity while watching social stimuli (ages 2–6 years) [22], and greater and more stereotypical head movements during dyadic social interactions (ages 6–13 years) [38, 39].

Despite evidence that head movement is both critical for social communication and may be altered in autism, research is scarce compared to other nonverbal domains, such as facial expressions and gestures [17]. One barrier has been the difficulty of very precisely measuring characteristics of head movements, especially during natural social interactions. Traditionally, head movement dynamics (*e.g.,* moments of head nodding) have been studied through manual annotations by trained observers, which can be a time–consuming process, or one lacking scalability and reliability [4, 11]. Manual annotations are also unable to detect highly subtle or granular aspects of head movements that might differ between autistic and neurotypical people. Recent advances in computer vision and machine learning have enabled precise, automated analysis of human social behaviors, as they unfold over time, using computational models [8, 18, 24]. A growing body of work has demonstrated the utility of computational behavior analysis during screen–based tasks (*e.g.,* participant watching videos or looking at images) [22] and, less commonly, live social interactions [13].

In this study, we used computational behavior analysis to study head movement patterns in neurotypical and autistic individuals (ages 19–49) during face–to–face conversations with an unfamiliar adult. Our approach enables study of the progression of head movements over time during conversations, rather than just overall kinematic features (*e.g.,* speed, spatial range). We aim to understand if and how head movements alone reveal social skill differences associated with autism. We studied head movements at two levels, individual (*monadic*: analyzing the participant only) and interpersonal (*dyadic*: analyzing the back-and-forth between the two interaction partners), to determine whether measuring head movements at the dyadic level is incrementally more informative than studying the cues of a single person.

Our results show that the machine learning pipeline built to distinguish between autistic and neurotypical individuals based on head movement data from both interaction partners (*dyadic* level) achieves 80% accuracy, suggesting that head movements alone contain rich information relevant for social communication in autism.

The same pipeline performs worse (69%) when using head movement data from the participant only (*monadic* level), emphasizing the benefit of studying back-and-forth social interaction relative to more traditional, inherently *monadic* approaches. Our results motivate further study of head movement patterns in autism, and their variation across contexts and development.

2 METHODS

Sections 2.1 and 2.2 describe the study's experimental setup. Our computational approach, explained in Sections 2.3 – 2.5, uses a state–of–the–art computer vision algorithm to quantify head movements along three dimensions (yaw, pitch, roll). These signals are processed to generate *monadic* and *dyadic* features, capturing the head movement patterns used by participants alone or by both participants and conversation partners. These features are then used as input for a machine learning pipeline to predict the diagnostic status of participants (autism vs. neurotypical). This is a similar pipeline used in a previous work [24], with several modifications to better capture social dynamics in autism.

2.1 Participants and Experimental Procedure

Data collection was performed at the Center for Autism Research (CAR) at Children's Hospital of Philadelphia (CHOP). This research was approved by the institutional review board (IRB) at CHOP. Fifteen autistic and 27 neurotypical individuals matched on age, sex, and general intelligence (IQ) are included in the current study. The age range is

19.7 – 49.5 years, with a mean age of 28.2 years. Consistent with the fact that autism is disproportionately diagnosed in males, our sample includes 36 males and 6 females.

The experimental procedure consisted of a modified version of the Contextual Assessment of Social Skills (CASS) [28], a semi–structured assessment of conversational ability designed to mimic real–life first–time encounters. Participants engaged in a 3–minute face–to–face conversation with a research confederate (unaware of the dependent variables of interest). CASS confederates included undergraduate students and BA–level research assistants (all native English speakers). In order to provide opportunities for participants to initiate and develop the conversation, confederates had been trained to speak for no more than 50% of the time, and to wait 10s to initiate the conversation. If conversational pauses occurred, confederates had been trained to wait 5s before re–initiating the conversation. Otherwise, conversation partners were told to engage as they normally would.

2.2 Data Collection and Pre–Processing

Continuous audio and video of the 3–minute CASS were recorded using a specialized, custom–built "BioSensor" (Figure 1), that was placed between the participant and confederate on a floor stand. This device has two HD video cameras pointing in opposite directions, as well as two microphones, to allow for synchronous audio–video recording of the participant and the confederate as they sit facing each other. The minimal footprint and size of the device were intended to minimize the intrusiveness of the technology on the natural conversation.

For each video, the first and last 3 seconds were trimmed to remove any frames not including the CASS. This time limit was selected based on visual inspection of the videos. The videos were either recorded at a frame rate of 30 frames per second (fps), or down–sampled from 60 fps to 30 fps. Head pose data was extracted using a state–of–the–art 3D facial analysis algorithm [29]. The output consisted of time–dependent signals for the three fundamental head movement angles: *yaw*, *pitch*, and *roll* (Figure 2). In addition, for each head angle, the first discrete difference between consecutive frames (*i.e.,* velocity) was calculated, yielding another time–dependent signal for each angle.

2.3 Identifying Common Movement Patterns

We used K–Means clustering to group head movement snapshots by similarity, aiming to find general patterns. Due to the small sample size, we chose not to optimize the K parameter within cross–validation; instead, we set $K = 12$ based on results from [24]. For each participant, the temporal signal of each angle was split into overlapping windows of 4 seconds (120 frames given a 30 fps), with an overlap size of 8 time instances (roughly 0.3 seconds), in the same way as in [24]. Each window was subsequently standardized to have 0 mean and 1 standard deviation. Windowed data from all participants and all angular directions were combined and used to train a K–Means model. The resulting 12 cluster centers, each represented by an exemplar 4–second signal, captured common head movement patterns. The same process was separately repeated for the angle differential signals (*i.e.,* velocity signals).

2.4 Monadic and Dyadic Features

Monadic features were generated by counting the number of 4–second windowed samples of participants that were assigned to each cluster, yielding a K–dimensional frequency vector. This vector quantifies the number of times each head movement pattern was used by the participant. This process, also known as *bag of words* model [33, 34], was repeated for each angular direction (and for velocity signals) independently. The resulting six vectors (three angles and three velocity signals) were concatenated, yielding a final feature vector of size $6 \times K$.

In the dyadic case, we wanted to capture the interchange between the two people, since the correlation and timing of nonverbal cues is highly relevant to effective communication [5, 13]. The pattern counts were computed by counting the number of times the windowed samples of the participants and confederates were assigned to the clusters simultaneously, yielding a $K \times K$ matrix for each angle (and velocity). The cell (i, j) of the matrix included the number of times the participant made a head movement that was assigned to cluster i, while confederate simultaneously made a movement that was assigned to cluster j. Concatenating all six matrices and vectorizing the resulting tensor resulted in a $6 \times K \times K$ dimensional vector. Finally, we also considered simultaneous movements along different directions (*e.g.,* participant's movement along roll angle *vs.* confederate's movement along yaw angle). Thus the final feature vector had a size of $3 \times 6 \times K \times K$. We also prepared a third set of features by simply concatenating monadic and dyadic features.

2.5 Predicting ASD Diagnosis

Monadic, dyadic, and combined head movement features, independently, were used as the input to a binary SVM classifier with a linear kernel and $C = 1$ (*i.e.,* default settings), to predict the participants' diagnostic status: ASD *vs.* neurotypical (NT). We used 10–fold cross–validation. For each fold, part of the data was put aside for testing, and the remaining was used for training. At the end of the 10–fold cross–validation, each participant's data had been used in the testing set once. This 10–fold cross–validation was then repeated 100 times, with a different random seed used each time to shuffle the order of participants, to produce statistically robust performance metrics. In addition, in each case, leave–one–out cross–validation was also performed on the same data for consistency checks.

We also sought to elucidate the head movement features that contributed most strongly to the classification accuracy, to identify specific head movements that differ between ASD and NT groups. We extracted the feature weights computed by the SVM. Features with high absolute weights are assumed to posses a high information content. We then computed average weights for different angular directions to understand the contribution of different head movement types.

3 RESULTS AND DISCUSSION

The performance of the machine learning classifier in predicting diagnostic status is listed in Table 1. Accuracy, sensitivity, specificity, positive and negative predictive values (PPV and NPV) are listed for leave–one–out and 10 fold cross–validation. All performance metrics are significantly better than chance, suggesting the existence of informative signals in the head movement patterns of participants. Especially with the dyadic signals, the overall performance is very promising with an accuracy of 80.0%. Despite the high accuracy, it is important to note that the proposed classifier is not intended for diagnostic purposes. The low sensitivity value of 55.9% severely undermines the feasibility of using the classifier for such diagnostic purposes.

The lower accuracy of the monadic approach (69.2%) highlights the importance of studying social communication cues in a true social context, considering the behavior of both individuals in the interaction. This finding elegantly points to the fact that ASD is a condition that emerges within social contexts. The combination of monadic and dyadic features did not improve classification performance compared to using only dyadic features, potentially because concatenating the two sets of features may not be the most effective approach. The dimensionality of the dyadic features is much higher than that of the monadic features, making the contribution of the monadic ones negligible.

In all cases, the specificity values are higher than the sensitivity values, indicating that the method is better at correctly identifying neurotypical individuals (*i.e.,* true negatives) than ASD individuals (*i.e.,* true positives). This result is expected, given the small sample size and imbalanced group representation (15 autistic individuals versus 27 neurotypical individuals). In this small sample, the machine learning classifier appears to be more accurate at identifying typical movements, but less confident in detecting subtle atypicalities. Overall, it is evident that further research with larger and more balanced samples is necessary before a reliable diagnostic tool can be developed.

Figure 3 shows the importance for all types of head movement features used in classification. The weights of monadic features are displayed in Figure 3(a), and the averages across different feature types in Figure 3(c). Similarly, the weight distributions for the dyadic case are shown in Figure 3(b) and Figure 3(d). It is important to note that the interpretation of feature weights for a classifier with low accuracy may not be reliable. Therefore, here we mostly discuss feature weights for the dyadic case. In general, movement angles (i.e., magnitude) are more influential than velocity. It is possible that how an individual moves their head in response to their partner's movements (and *vice versa*) is more important in social interactions than the speed of these movements. However, it is important to note that this hypothesis may not hold for certain types of communicative cues that are primarily conveyed through the speed of the movement (*e.g.,* nodding to signal approval versus nodding to show interest).

Of the three angles (roll, pitch, and yaw), pitch and yaw have the greatest impact on this classification process. Movements along the pitch direction (*e.g.,* nodding) are often used to convey communicative cues such as agreement, attentiveness, and contemplation. Movements along the yaw direction (*e.g.,* shaking or rotating the head) may signal cues such as disagreement, attention, or orientation. While it is challenging to draw definitive and clear conclusions without further examining the signal shape of individual features (*i.e.,* how the angle changes over a 4–second window), it is clear that these types of cues play a significant role in social interactions.

As each feature in our analysis corresponds to a movement over a 4–second window, our analysis summarizes how a movement along a particular angle evolves during a conversation. It would be even more informative, and potentially more clinically relevant, to investigate how these features unfold over time in more detail. However, our analysis currently treats movements along different directions (roll, pitch, and yaw) independently, whereas real, semantically meaningful head movements involve a combination of movements along all three directions.

4 CONCLUSIONS

This work, relying on computer vision and multivariate machine learning strategies, creates computational models that distinguishing between autistic and neurotypical individuals based on their head movement patterns. Video data were collected from 3–minute, face–to–face, semi–structured conversations. Our machine learning model achieved an 80% classification accuracy when using dyadic head movement features, taking into account the coordination of head movements between both conversation partners. These results underscore the importance of studying social communication within natural social contexts to fully capture the psychological complexity of social interactions and social communication differences in autism. Computational behavior analysis, with its precision and scalability, facilitates such naturalistic studies. Future work should explore more sophisticated feature structures and machine learning models, and replicate these results with a larger, more diverse sample. Our findings provide the foundation for such explorations.

ACKNOWLEDGMENTS

This work is partially funded by the National Institutes of Health (NIH) Office of the Director (OD) and National Institute of Mental Health (NIMH) of US, under grants R01MH122599, R01MH118327, 5P50HD105354-02, and 1R01MH125958-01.

REFERENCES

[1]. Alokla Shamma. 2018. Non-verbal communication skills of children with autism spectrum disorder. (2018).

[2]. Alshurman Wael and Alsreaa Ihsani. 2015. The Efficiency of Peer Teaching of Developing Non Verbal Communication to Children with Autism Spectrum Disorder (ASD). Journal of Education and Practice 6, 29 (2015), 33–38.

[3]. American Psychiatric Association et al. 2013. Diagnostic and Statistical Manual of Mental Disorders (DSM-V)(5th ed.).

[4]. Cappella Joseph N. 1981. Mutual influence in expressive behavior: Adult–adult and infant–adult dyadic interaction. Psychological bulletin 89, 1 (1981), 101. [PubMed: 7232607]

[5]. Curto David, Clapés Albert, Selva Javier, Smeureanu Sorina, Junior Julio, Jacques CS, Gallardo-Pujol David, Guilera Georgina, Leiva David, Moeslund Thomas B, et al. 2021. Dyadformer: A multi-modal transformer for long-range modeling of dyadic interactions. In Proceedings of the IEEE/CVF International Conference on Computer Vision. 2177–2188.

[6]. Duncan Starkey. 1972. Some signals and rules for taking speaking turns in conversations. Journal of personality and social psychology 23, 2 (1972), 283.

[7]. Flanagan Joanne E, Landa Rebecca, Bhat Anjana, and Bauman Margaret. 2012. Head lag in infants at risk for autism: a preliminary study. The American Journal of Occupational Therapy 66, 5 (2012), 577–585. [PubMed: 22917124]

[8]. Georgescu Alexandra Livia, Koehler Jana Christina, Weiske Johanna, Vogeley Kai, Koutsouleris Nikolaos, and Falter-Wagner Christine. 2019. Machine learning to study social interaction difficulties in ASD. Frontiers in Robotics and AI 6 (2019), 132. [PubMed: 33501147]

[9]. Georgescu Alexandra Livia, Kuzmanovic Bojana, Roth Daniel, Bente Gary, and Vogeley Kai. 2014. The use of virtual characters to assess and train non-verbal communication in high-functioning autism. Frontiers in human neuroscience 8 (2014), 807. [PubMed: 25360098]

[10]. Gima Hirotaka, Kihara Hideki, Watanabe Hama, Nakano Hisako, Nakano Junji, Konishi Yukuo, Nakamura Tomohiko, and Taga Gentaro. 2018. Early motor signs of autism spectrum disorder in spontaneous position and movement of the head. Experimental brain research 236, 4 (2018), 1139–1148. [PubMed: 29450616]

[11]. Grammer Karl, Kruck Kirsten B, and Magnusson Magnus S. 1998. The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. Journal of Nonverbal behavior 22, 1 (1998), 3–29.

[12]. Gray Kylie Megan and Tonge Bruce John. 2001. Are there early features of autism in infants and preschool children? Journal of paediatrics and child health 37, 3 (2001), 221–226. [PubMed: 11468034]

[13]. Hale Joanna, Jamie A Ward Francesco Buccheri, Oliver Dominic, and de C Hamilton Antonia F. 2020. Are you on my wavelength? Interpersonal coordination in dyadic conversations. Journal of nonverbal behavior 44, 1 (2020), 63–83. [PubMed: 32189820]

[14]. Hall Judith A, Horgan Terrence G, and Murphy Nora A. 2019. Nonverbal communication. Annual review of psychology 70, 1 (2019), 271–294.

[15]. Hamdan Mohammed Akram. 2018. Developing a Proposed Training Program Based on Discrete Trial Training (DTT) to Improve the Non-Verbal Communication Skills in Children with Autism Spectrum Disorder (ASD). International Journal of Special Education 33, 3 (2018), 579–591.

[16]. Hammal Zakia, Cohn Jeffrey F, and George David T. 2014. Interpersonal coordination of headmotion in distressed couples. IEEE transactions on affective computing 5, 2 (2014), 155–167. [PubMed: 26167256]

[17]. Hammal Zakia, h Jeffrey F, and Messinger Daniel S. 2015. Head movement dynamics during play and perturbed mother-infant interaction. IEEE transactions on affective computing 6, 4 (2015), 361–370. [PubMed: 26640622]

[18]. Javed Hifza, Lee WonHyong, and Park Chung Hyuk. 2020. Toward an automated measure of social engagement for children with autism spectrum disorder—a personalized computational modeling approach. Frontiers in Robotics and AI (2020), 43. [PubMed: 33501211]

[19]. Kaliouby Rana el and Robinson Peter. 2005. Real-time inference of complex mental states from facial expressions and head gestures. In Real-time vision for human-computer interaction. Springer, 181–200.

[20]. Babu Pradeep Raj Krishnappa, Martino J Matias Di, Chang Zhuoqing, Perochon Sam, Aiello Rachel, Carpenter Kimberly LH, Compton Scott, Davis Naomi, Franz Lauren, Espinosa Steven, et al. 2022. Complexity analysis of head movements in autistic toddlers. Journal of Child Psychology and Psychiatry (2022).

[21]. Livingstone Steven Rand Palmer Caroline. 2016. Head movements encode emotions during speech and song. Emotion 16, 3 (2016), 365. [PubMed: 26501928]

[22]. Martin Katherine B, Hammal Zakia, Ren Gang, Cohn Jeffrey F, Cassell Justine, Ogihara Mitsunori, Britton Jennifer C, Gutierrez Anibal, and Messinger Daniel S. 2018. Objective measurement of head movement differences in children with and without autism spectrum disorder. Molecular autism 9, 1 (2018), 1–10. [PubMed: 29321841]

[23]. McClave Evelyn Z. 2000. Linguistic functions of head movements in the context of speech. Journal of pragmatics 32, 7 (2000), 855–878.

[24]. McDonald Denisa Qori, Zampella Casey J, Sariyanidi Evangelos, Manakiwala Aashvi, DeJardin Ellis, Herrington John D, Schultz Robert T, and Tunc Birkan. 2022. Head Movement Patterns during Face-to-Face Conversations Vary with Age. In Companion Publication of the 2022 International Conference on Multimodal Interaction. 185–195.

[25]. Mundy Peter, Sigman Marian, Ungerer Judy, and Sherman Tracy. 1986. Defining the social deficits of autism: The contribution of non-verbal communication measures. Journal of child psychology and psychiatry 27, 5 (1986), 657–669. [PubMed: 3771682]

[26]. Philip Ruth CM, Whalley Heather C, Stanfield Andrew C, Sprengelmeyer R, Santos Isabel M, Young Andrew W, Atkinson Anthony P, Calder AJ, Johnstone EC, Lawrie SM, et al. 2010. Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. Psychological medicine 40, 11 (2010), 1919–1929. [PubMed: 20102666]

[27]. Sara Ramos-Cabo Valentin Vulchanov, and Vulchanova Mila. 2019. Gesture and language trajectories in early development: An overview from the autism spectrum disorder perspective. Frontiers in Psychology 10 (2019), 1211. [PubMed: 31191403]

[28]. Ratto Allison B, Turner-Brown Lauren, Rupp Betty M, Mesibov Gary B, and Penn David L. 2011. Development of the contextual assessment of social skills (CASS): A role play measure of social skill for individuals with high-functioning autism. Journal of autism and developmental disorders 41, 9 (2011), 1277–1286. [PubMed: 21287253]

[29]. Sariyanidi Evangelos, Zampella Casey J, Schultz Robert T, and Tunc Birkan. 2020. Inequality-constrained and robust 3D face model fitting. In European Conference on Computer Vision. Springer, 433–449.

[30]. Schultz Robert T, Gauthier Isabel, Klin Ami, Fulbright Robert K, Anderson Adam W, Volkmar Fred, Skudlarski Pawel, Lacadie Cheryl, Cohen Donald J, and Gore John C. 2000. Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. Archives of general Psychiatry 57, 4 (2000), 331–340. [PubMed: 10768694]

[31]. Senju Atsushi and Johnson Mark H. 2009. Atypical eye contact in autism: models, mechanisms and development. Neuroscience & Biobehavioral Reviews 33, 8 (2009), 1204–1214. [PubMed: 19538990]

[32]. Sigman Marian, Mundy Peter, Sherman Tracy, and Ungerer Judy. 1986. Social interactions of autistic, mentally retarded and normal children and their caregivers. Journal of child psychology and psychiatry 27, 5 (1986), 647–656. [PubMed: 3771681]

[33]. Sivic Josef and Zisserman Andrew. 2003. Video Google: A text retrieval approach to object matching in videos. In Computer Vision, IEEE International Conference on, Vol. 3. IEEE Computer Society, 1470–1470.

[34]. Sivic Josef and Zisserman Andrew. 2008. Efficient visual search of videos cast as text retrieval. IEEE transactions on pattern analysis and machine intelligence 31, 4 (2008), 591–606.

[35]. Speer Leslie L, Cook Anne E, McMahon William M, and Clark Elaine. 2007. Face processing in children with autism: Effects of stimulus contents and type. Autism 11, 3 (2007), 265–277. [PubMed: 17478579]

[36]. Verhoeven Judith S, Cock Paul De, Lagae Lieven, and Sunaert Stefan. 2010. Neuroimaging of autism. Neuroradiology 52, 1 (2010), 3–14. [PubMed: 20033797]

[37]. Weigelt Sarah, Koldewyn Kami, and Kanwisher Nancy. 2012. Face identity recognition in autism spectrum disorders: A review of behavioral studies. Neuroscience & Biobehavioral Reviews 36, 3 (2012), 1060–1084. [PubMed: 22212588]

[38]. Zhao Zhong, Zhu Zhipeng, Zhang Xiaobin, Tang Haiming, Xing Jiayi, Hu Xinyao, Lu Jianping, Peng Qiongling, and Qu Xingda. 2021. Atypical Head Movement during Face-to-Face Interaction in Children with Autism Spectrum Disorder. Autism Research 14, 6 (2021), 1197–1208. [PubMed: 33529500]

[39]. Zhao Zhong, Zhu Zhipeng, Zhang Xiaobin, Tang Haiming, Xing Jiayi, Hu Xinyao, Lu Jianping, and Qu Xingda. 2022. Identifying autism with head movement features by implementing machine learning algorithms. Journal of Autism and Developmental Disorders 52, 7 (2022), 3038–3049. [PubMed: 34250557]

CCS CONCEPTS

• **Applied computing** → **Psychology;** • **Computing methodologies** → **Computer vision**; **Machine learning**; *Cluster analysis*; *Supervised learning by classification*; Cross-validation.

(a) (b)

Figure 1:
Experimental setup and data collection hardware. In *(a)*, the participant and confederate are having a conversation with the “BioSensor” camera placed between them. In *(b)*, the current model the of “BioSensor.”

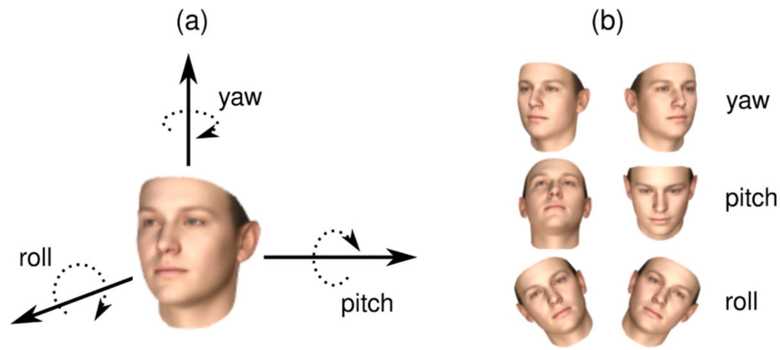
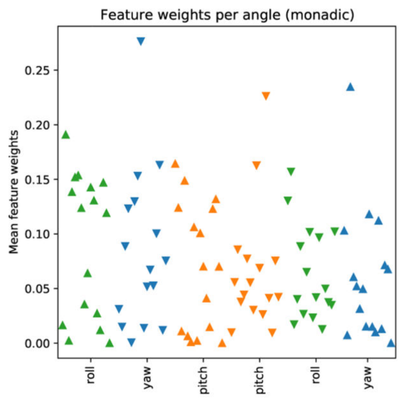
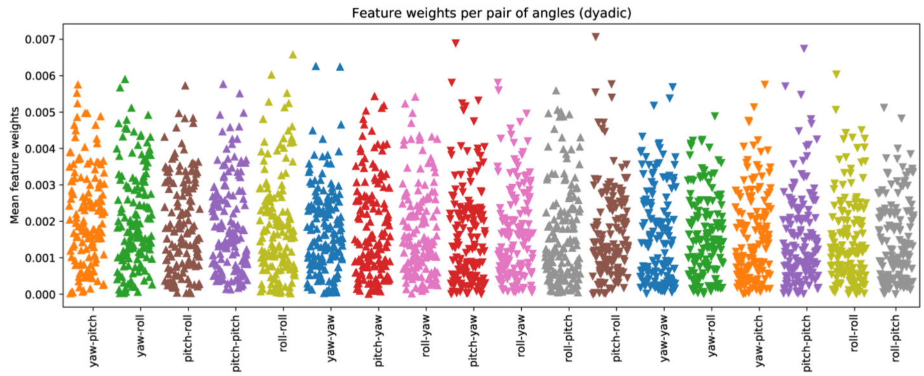


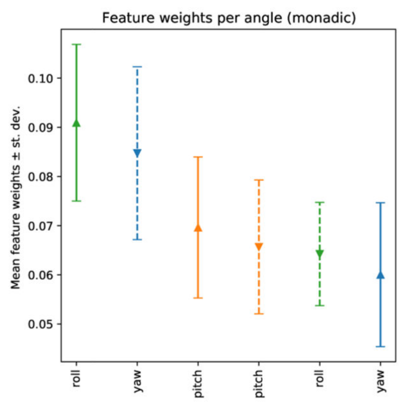
Figure 2:
Illustration of the three angles used in the study to quantify head movement.



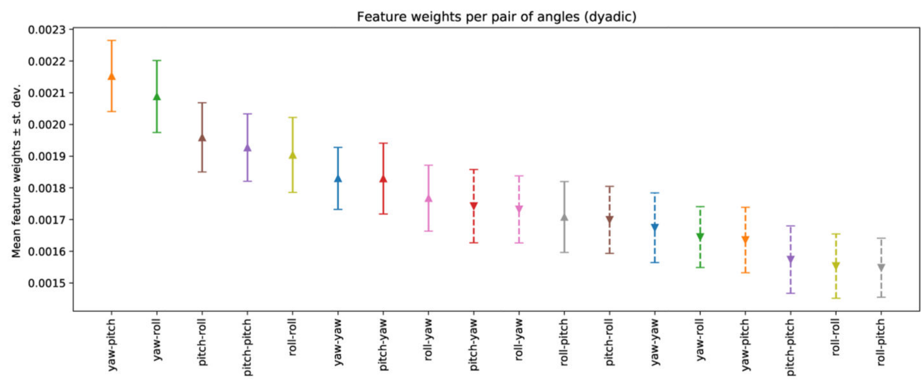
(a)



(b)



(c)



(d)

Figure 3:

The weights obtained during classification have been plotted for each feature construction scenario: *monadic* features in *(a)*, *dyadic* features in *(b)*, statistics for *monadic* features in *(c)*, and statistics for *dyadic* features in *(d)*. For each feature, the y-axis represents the weight coefficients assigned by the SVM classifier in *(a)* and *(b)*, and the mean and standard deviation of those weights in *(c)* and *(d)*. In the *monadic* case, features are the single head movement angles (upward pointing triangles) and their differentials (downward pointing triangles). The *dyadic* feature plot contains relationships between head angle pairs (upward pointing triangles) and their differentials (downward pointing triangles).

Table 1:

Classification results for both *monadic* and *dyadic* features and their combination. The reported scores are means of the 100 experimental runs. For each case, the 10–fold cross validation scores and leave-one-out cross validation (LOO) are reported.

	Accuracy	Sensitivity	Specificity	PPV	NPV
Monadic 10–fold	69.2%	54.5%	77.4%	57.5%	75.5%
Monadic LOO	66.7%	53.3%	74.1%	53.5%	74.1%
Dyadic 10–fold	80.0%	55.9%	93.3%	82.4%	79.2%
Dyadic LOO	78.6%	53.3%	92.6%	80.0%	78.1%
Monadic & Dyadic 10–fold	79.7%	56.0%	92.6%	80.0%	78.1%
Monadic & Dyadic LOO	78.6%	53.3%	92.6%	80.0%	78.1%