

Analyzing Eye-Tracking Data to Detect Joint Attention in Hexgame Experiments

Yide Yang

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Rena Kato

Kyushu Institute of Technology, 680-4 Kawazu, Iizuka-shi, Fukuoka, 820-8502, Japan

Kosei Shibata

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Hendry F. Chame

University of Lorraine, CNRS, LORIA, F-54000, France

Laurent Bougrain

University of Lorraine, CNRS, LORIA, F-54000, France

Sorbonne Université, Institut du Cerveau – Paris Brain Institute -ICM, CNRS, Inria, Inserm, AP-HP, Hopital de la Pitié Salpêtrière, 75013 Paris, France

Hiroaki Wagatsuma

Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu, 808-0196, Japan

Email: yang.yide421@mail.kyutech.jp, katorena.tyt@gmail.com, sibata.kousei927@mail.kyutech.jp, hendry.ferreira-chame@loria.fr, laurent.bougrain@loria.fr, waga@brain.kyutech.ac.jp

Abstract

This study aims to explore the mechanisms of joint attention in a strategic game experiment by analyzing Tobii eye-tracking data. In this experiment, two participants play Hexgame, during which the gaze direction is tracked and projected onto the game's board plane through perspective geometry, in order to track attention sharing correlates at particular stages of the game from behavioral data. The primary focus of this study is the analysis of eye-tracking data to identify attention coordination between players during the progression of the game. Future work will expand this framework to assess win probabilities and predict subsequent moves, providing deeper insights into strategic decision-making.

Keywords: Joint Attention, Hexgame, Eye-Tracking, Perspective Transformation

1. Introduction

Joint attention (JA), where two or more individuals coordinate their gaze towards the same object or event, is fundamental for social interaction and communication [1]. It plays a crucial role in building mutual understanding and effective cooperation, and is also essential in strategic games, as it helps to reveal players' decision-making processes [2].

Previous studies have demonstrated that coordinated gaze between participants often leads to better performance in collaborative contexts, as it enhances the sharing of information and facilitates effective coordination [3]. This ability to synchronize attention is not only important in cooperative settings but also provides insight into the dynamics of shared environments, such as understanding intentions and building shared goals [3].

Joint attention is especially relevant in the context of Autism Spectrum Disorder (ASD), where difficulties in

initiating or responding to joint attention can significantly impact social interactions and communication [4]. Research shows that deficits in joint attention hinder social skill development, but targeted interventions can help address these challenges and improve social functioning [4]. Therefore, understanding the mechanisms of joint attention has practical implications for both typical development and enhancing social abilities in individuals with ASD.

Despite extensive research on joint attention in cooperative and developmental settings, its role in competitive environments has received less attention, especially regarding how players anticipate and adapt to each other's strategies during real-time interactions. In competitive games, players must balance their focus between their own actions and their opponent's strategies, creating a dynamic environment where joint attention can become strategically advantageous [5]. Understanding this dynamic can provide deeper insights into player behavior, strategy formation, and decision-making processes.

In this study, we investigate joint attention in the context of competitive games. In particular, we focus on Hexgame, a strategy game that requires players to make rapid decisions. By utilizing Tobii eye-tracking devices, we recorded and analyzed gaze behavior between players. For identifying behavioral correlates of joint attention unfolding in critical phases of the game, we propose a methodology which consists in mapping gaze direction estimations acquired in on-board sensors to a common representation which is the game's board plane. We believe that this method can provide insights into how players adapt to and anticipate each other's moves [4] [5]. Such insights can contribute to broader research on attention dynamics in competitive environments and highlight the importance of shared visual focus in complex, adversarial interactions.

2. Methodology

This study aims to track behavioral correlates of joint attention between two participants by analyzing their eye-tracking data from an Hexgame game experiment. The following section provides a detailed explanation of the experimental setup and the data processing method, including the perspective transformation algorithm.

2.1. Data Collection

The experiment setup (Fig.1) is composed of a self-designed 8×8 Hexgame board and pieces. The board followed the standard Hexgame layout, consisting of a regular hexagonal grid with each hexagon's side-to-side distance being 60 mm. The overall board was shaped like a parallelogram, with an 8×8 active game area surrounded by boundary regions. These boundary regions, consisting of one additional row of hexagons on all sides, defined the win conditions for the red and blue sides.

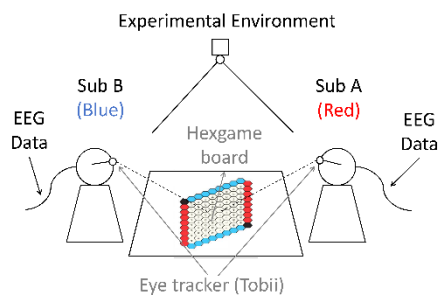


Fig.1 Diagram of Experimental Environment Setup

Hexgame is a two-player strategy game where the goal is to connect two opposite sides of the board with a chain of pieces of the player's color. Players take turns placing pieces, with the red player starting first. The game ends when a connection is made or the board is full. If no winner emerges, the game is a draw. In this experiment,

players synchronized their moves with the rhythm to ensure timing consistency in data analysis.

The experiment included 30 red and 30 blue 3D-printed game pieces. A fixed overhead camera recorded the board's movements, capturing the sequence of all piece placements. Two players competed on the board, each wearing a Tobii pro glasses 3 device for eye-tracking, with calibration completed before the experiment. EEG devices were also used, with channels 17 and 34 recording trigger signals (encoded as 2V, 4V, or 8V) for synchronization across devices.

Eight participants were divided into four groups of two, assigned to either the red or blue side. Players acted in sync with a 50 BPM rhythm, completing actions on each down beat. Each game round lasted up to 5 minutes or ended when one player won, with 1-minute breaks between rounds to reduce fatigue. This rhythmic design helps align the timing of in-game events with the EEG data. In the experiment, gaze trajectories were recorded for the participants in real-time. The dataset included images from first-person perspective, timestamp and point coordinates of eye focus on the image, and, as shown in Fig.1, we also recorded from a fixed camera perspective and EEG measures. We hope that this dataset will support future in-depth studies on joint attention mechanisms.

2.2. Data Processing

In a preprocessing stage, the recorded videos were converted into individual frames, and typical images containing the chessboard were manually selected. A training consisting of 100 images was constituted. A validation set was also selected, consisting of 20 images. The AI image annotation and dataset creation tool LabelMe was used to annotate the polygons of the chessboard for supervised training of the model.

The data analysis workflow is illustrated in Fig.2. First, gaze data, including timestamps and 2D image coordinates, is extracted from the Tobii data files and formatted into a standardized structure for subsequent processing. By using a convolutional neural network model YOLO [6], the four corners of the Hexgame board are detected in a single pass on images acquired from each participant's perspective and the board's area (i.e. the total pixel count) is calculated. A threshold of 300,000 pixels was chosen for the study, with a criterion that the area must exceed 60% of this threshold to qualify further processing.

The perspective transformation algorithm was implemented by using the OpenCV library. Thus, the `cv2.getPerspectiveTransform` method was employed to calculate a transformation matrix M from four pairs of corresponding points between the corners of the board segmented from the ceiling camera's image and the

corners detected from the images obtained from the participant's perspective. These transformation matrices are then applied respectively to each participant's perspective coordinates (using `cv2.warpPerspective`), for obtaining their projection onto the chessboard image plane in third-person perspective (as viewed from the ceiling camera). Likewise, the gaze focus location in 2D coordinates (as estimated by the Tobii eye-tracker) are transformed to the third-person perspective for each participant. Resulting from this step, the estimated gaze focus of participants are expressed in a single coordinate frame and board layout, which can facilitate the analysis and tracking of joint attention behavior correlates relative to the game's states.

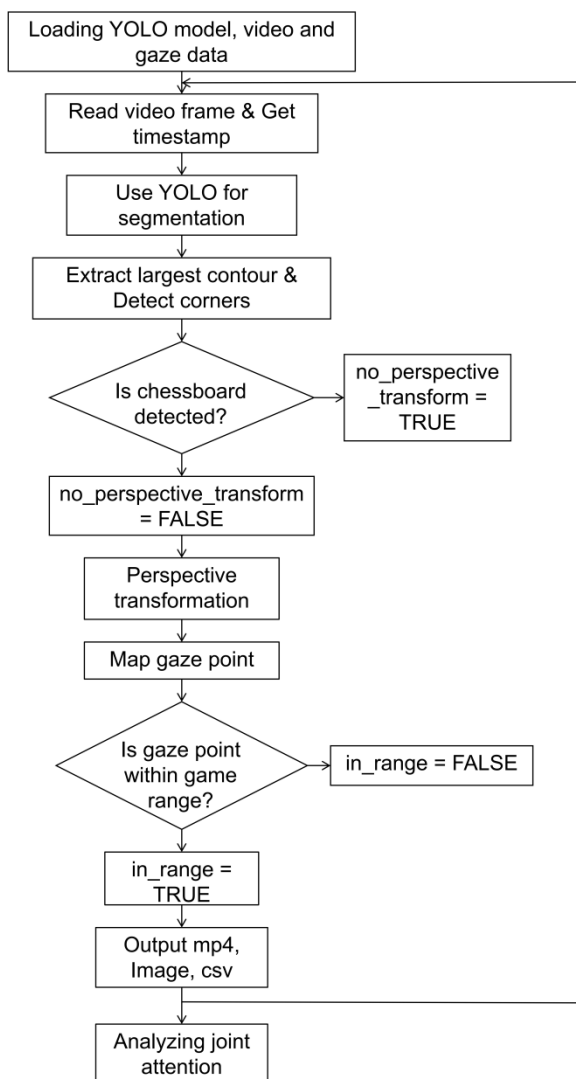


Fig. 2. Data Analysis Flowchart

In cases where the board cannot be detected or its pixel area is considered to be insufficient with respect to the established threshold, the estimation of perspective transformation is skipped for the participant; otherwise, the transformation matrix is calculated and applied to both the image and gaze points. After completing the

perspective transformation, the transformed gaze points are further evaluated to determine whether they fall within the board's boundaries.

3. Results and Discussion

As illustrated in Fig.2, our methodological proposal is structured in several stages. Next, we report on the partial results obtained at these stages.

3.1. Corner Detection Results

By inspecting the video containing the original gaze points, we observed that transformed gaze points aligned closely with positions on the board (Fig.3), which shows that a reasonably good precision can be obtained from the experiment setup when expressing estimated gaze focus of participants from first-person perspective to a board's third-person perspective.

When analyzing corner segmentation results, we observed that the detection process is influenced by various environmental factors, such as head motions away from the board and participants' hand motions obstructing the board during gameplay.

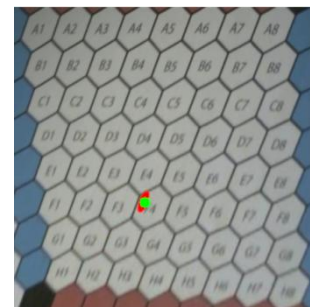
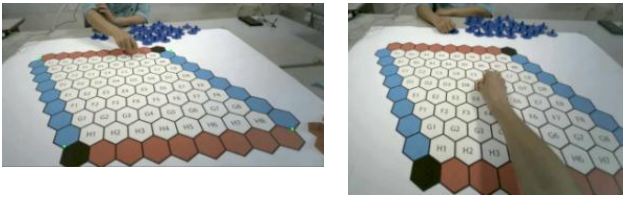


Fig. 3 Comparison of Transformed Gaze with Gaze-Overlaid Video

As shown in Fig.4, momentary occlusions happen during the game situation which compromises the board's segmentation step (Fig.4.b). In the figure it is also shown a situation where small portions of the board's corners occasionally move out of the video frame, leading to detection failures. Frames where the board was not fully detected were discarded. Finally, it is also worth mentioning that for performance reasons, the YOLO architecture was trained with an input size of 640x480 pixels, while the actual resolution of the input video was 1920x1080 pixels. Hence, the detected coordinates were remapped to the original video size which introduced approximation errors estimated to be around 20 pixels.



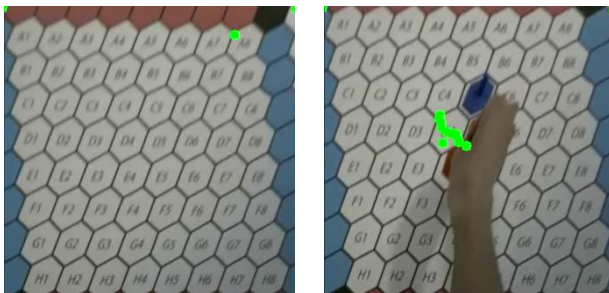
a. Reliable corners detection b. Uncertain corners detection

Fig. 4. Comparison of corners detection results

Despite the limitations addressed above, we observed that corners detection performed reliably when the board was fully contained in the image (Fig.4.a), with results aligning closely with the actual chessboard layout under most game situations.

3.2. Perspective Transformation Results

The perspective transformation stage worked well in most cases (Fig.5.a), with successful frames showing a high level of consistency between the mapped chessboard area and gaze points. Thus, most gaze points accurately fall within the chessboard boundaries. As discussed, the perspective transformation stage results are conditioned to successful corners' detection, and the latter was affected by head and arm movements. Consequently, instabilities on gaze focus estimation remapping were observed in cases where small portions of the chessboard corners temporarily moved out of the video frame compromising border detection.



a. Low uncertainty b. Uncertainty when the border is not detected

Fig. 5. Comparison of estimations from good corners detection (left) and the heuristic of reusing the last available transformation when the board is not fully contained in the field of view (right)

As shown in Fig.5b, we tested an heuristic approach for perspective transformation in cases where the segmentation algorithm failed to detect the boards' corners. This heuristic consisted in applying the last available perspective transformation, exploiting spatio-temporal statistical regularities in gaze focus location estimation. This approach showed to be limited since gaze points estimation may fluctuate which can impact

further analysis stages and data visualisation, notably, due to distinct behavior style of participants. For instance, it was observed during the game that depending on the player's behavior style, the board can leave the field of view recorded, or be obstructed by arm motions relatively frequently, resulting in fluctuations on estimations. Conversely, when movements are reduced, the heuristic is able to provide more acceptable results.

To address these issues and analyze the impact on further analysis stages and data visualisation, we implemented data preprocessing measures, including visualizing anomalous frames for review and rotating the perspective-transformed images by 180° when necessary to ensure consistent game board orientation. Additionally, we are working on optimizing the robustness of the corner detection model (e.g., using multi-frame fusion detection) and refining code logic to synchronize frame perspective transformation with gaze point transformation to further reduce the occurrence of anomalies. These improvements will enhance the stability of perspective transformation and provide more reliable data for joint attention analysis.

3.3. Future Developments

In this ongoing work, we focused on estimating and projecting estimated participants' gaze directions onto the Hexgame board plane, by segmenting the board's corners and applying perspective geometry remapping. We believe that these efforts serve as a foundational step toward the development of more precise methods for tracking joint attention correlated in board grid cells. Preliminary evaluations of our approach revealed that while good results can be achieved when the board is fully visible in each participant's field of view, challenges arise in cases where the board is partially occluded by arms or when participants look away from the board. Consequently, we are working on improving the corner detection stage of our processing pipeline.

Possible improvements include optimizing the corner detection algorithm by considering multiple frame processing and evaluating alternatives or introducing adjustments to the YOLO-based architecture model to suit our research. These improvements will provide a more robust data foundation for future research.

The current results constitute an important basis for subsequent studies. Hence, a comprehensive evaluation of the accuracy of visual focus behavioral tracking in the board plane as an estimate of joint attention correlates is being conducted. This evaluation will involve validating the current outputs, focusing on comparisons with the actual labeled points in the experimental design and conducting simulation tests under known conditions. Future developments will aim at enhancing the performance of existing models and validate them using

broader datasets to improve detection stability and robustness.

4. Conclusion

This study proposes an innovative methodological framework that integrates data from eye-tracking devices acquisitions to deep-learning techniques and perspective remapping for tracking joint attention behavior correlates unfolding in an Hexgame game interaction experience. While the partial results reported are limited, they lay a solid foundation for future improvements and applications, and constitute promising directions to study the mechanisms of joint attention in human competitive game interactions.

Acknowledgments

This work was supported in part by JSPS KAKENHI (JP17H06383, JP24K07387), and the joint research project: Collaborative project Université de Lorraine – Kyutech “Human/Robot Social Interactions: engagement and affect analysis during gaming tasks.”

References

1. Butterworth, G. (1991). "The ontogeny and phylogeny of joint visual attention." In *Natural theories of mind: Evolution, development, and simulation of everyday mindreading*, pp. 223-232.
2. Tomasello, M., Carpenter, M., & Liszkowski, U. (2007). "A new look at infant pointing." *Child Development*, 78(3), 705-722.
3. Richardson, D. C., & Dale, R. (2005). "Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension." *Cognitive Science*, 29(6), 1045-1060.
4. Mundy, P., & Sigman, M. (2006). "Joint attention, social competence, and developmental psychopathology." In *Development and Psychopathology*.
5. Cheng, J. T., Gerpott, F. H., Benson, A. J., Buckner, B., Foulsham, T., Lansu, T. A. M., Schülke, O., & Tsuchiya, K. (in press). "Eye gaze and visual attention as a window into leadership and followership: A review of empirical insights and future direction." *The Leadership Quarterly*, 34(6), 101654.
6. Redmon, J. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*.

Authors Introduction

Mr. Yide Yang



He received his Bachelor's degree in Intelligent Science and Technology in 2023 from the School of Artificial Intelligence and Data Science at Hebei University of Technology, China. He is currently pursuing his Master's degree at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, Japan.

Ms. Rena Kato



She received her associate's degree from the Department of Electrical and Electronic Systems Engineering, National Institute of Technology, Toyota College, Japan in 2024. She is currently enrolled in the Department of Systems Design and Informatics, Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology, Japan.

Mr. Kosei Shibata



He received his Master's degree in Engineering in 2023 from the Graduate School of Life Science and System Engineering, Kyushu Institute of technology (Kyutech) in Japan. He is currently a doctoral course student in Kyutech.

Dr. Hendry F. Chame



He holds degrees in Psychology and Computer Science (UCV, Venezuela), M.S. degrees in Instructional Psychology and Robotics and a Ph.D. degree in Robotics (ECN, France). His postdoctoral research includes bio-inspired (FURG, Brazil), neural (OIST, Japan), and cognitive robotics (LAAS, France). In 2021, he became an Associate Professor at the University of Lorraine, focusing on AI, robotics, and complex systems at LORIA.

Dr. Laurent Bougrain



He received his M.S. degree in artificial intelligence in 1995 from university of Paris 6, a B. degree in psychology and Ph.D. degrees in computer science in 2000 from universit  de Lorraine. He is an associate professor at the university of Lorraine (France) since 2001. He is the leader of the NeuroRhythms team of the department of the complex system, artificial intelligence and robotics at LORIA laboratory (France).

Dr. Hiroaki Wagatsuma



He received his M.S., and Ph.D. degrees from Tokyo Denki University, Japan, in 1997 and 2005, respectively. In 2009, he joined Kyushu Institute of Technology, where he is currently an Associate Professor of the Department of Human Intelligence Systems. His research interests include non-linear dynamics and robotics. He is a member of IEEE.
