Analysis of Careless Mistakes Using Gaze Information

Ryota Yabe
Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, 820-8502, Japan
Email: yabe.ryota884@mail.kyutech.jp

Noriko Takemura
Kyushu Institute of Technology, 680-4 Kawazu, Iizuka, 820-8502, Japan
Email: takemura@ai.kyutech.ac.jp

Abstract

Careless mistakes are caused by a lack of concentration, time pressure, and information overload due to multitasking. If we predict careless errors, they could be helpful in various situations, such as learning support, business efficiency improvement, and medical diagnosis support. However, it is challenging to reproduce careless error situations, and few studies have been on predicting careless errors. Therefore, we focus on Shogi (Japanese chess), where situational awareness is complex, and players must maintain long periods of concentration, making careless mistakes. In this study, we collected eye-tracking data of players during a game and annotated careless errors based on the game's contents and surveys from the players. By analyzing this data, we have examined the conditions under which careless mistakes occur).

Keywords: Careless mistakes, Shogi, Gaze analysis

1. Introduction

Careless mistakes refer to errors due to a lack of attention or insufficient confirmation despite having sufficient knowledge or ability. Careless mistakes can have a significant impact, especially in academics and work.

Several factors trigger careless mistakes. First, a lack of concentration is a significant factor. When concentration is lacking, attention becomes distracted, and one may overlook details. In particular, performing the same task for an extended period or engaging in monotonous work can lower attention levels. Second, lack of time is another important factor. When there are deadlines or time constraints, work tends to become sloppy, and verification steps may be skipped. In such situations, tasks that would normally be performed carefully can lead to a higher likelihood of careless mistakes. Furthermore, information overload due to multitasking also affects performance. By engaging in multiple tasks simultaneously, attention is divided among these tasks, making it difficult to process information accurately. As a result, when attention is required for specific tasks, it is often insufficient, leading to an increase in careless mistakes.

As mentioned above, careless mistakes are caused by various factors, such as a lack of concentration, insufficient time, and information overload due to multitasking. If we could predict careless mistakes, there would be expectations for their application in various scenarios, including learning support, work efficiency improvement, and medical diagnosis support. However, while numerous studies exist on estimating human internal states [1] [2], research on predicting careless mistakes has not yet been conducted. This lack of research is due to the difficulty in collecting data, as careless mistakes cannot be intentionally produced.

Therefore, in this study, we focus on Shogi, a game in which careless mistakes can occur, and create a dataset. Shogi requires complex situational judgment, such as defending while attacking, ignoring the opponent's attack and attacking each other, and defending from the opponent's attack. Additionally, it necessitates prolonged concentration, making it prone to fatigue and a decline in focus. In addition, because of the time limit, it is easy to feel the mental pressure of time. Given these factors, Shogi is highly susceptible to careless mistakes and is suitable for data collection.

Additionally, to predict careless errors, it is first necessary to analyze the circumstances under which careless errors occur. In this research, gaze information is utilized for that analysis. Gaze information can provide insights into what a person is focusing on, what they are struggling with, and signs of fatigue or drowsiness. This information is believed to be crucial for predicting careless mistakes. Although gaze movements are assumed to vary depending on the task being performed, it is expected that the analysis of gaze movement information during a game of Shogi in this study will provide clues for clarifying the mechanism of careless errors.

The subjects annotate the collected gaze data regarding the presence or absence of careless mistakes for each move, thereby the gaze dataset is constructed. For the analysis, various features are extracted from the gaze dataset, and F-tests and t-tests are conducted to analyze whether there is a significant difference in gaze characteristics related to careless mistakes.

2. Construction of a Gaze Dataset

2.1. *Dataset overview*

Five male subjects in their 20s compete against a shogi AI, and gaze information during the match is collected using an eye tracker. Furthermore, after the match, the subjects annotate the presence or absence of careless mistakes for each move, thereby constructing the dataset. The composition of this dataset is shown in Table 1.

Table 1. Composition of The Gaze Dataset.
("#" indicates the total count.)

ID	#Matches	#Moves	#Careless Mistake Moves
1	7	389	1
2	1	36	1
3	5	170	5
4	6	241	5
5	7	195	4
Total	26	1031	16

2.2. *Shogi GUI*

Shogi is a type of traditional Japanese two-player board game. Players move their pieces on a 9×9 square board to capture the opponent's king. The pieces consist of eight types: king, gold general, silver general, knight, lance, pawn, rook, and bishop.

In this study, five male subjects in their 20s who are familiar with shogi rules compete against the Shogi AI 'Suisho5' [3] using ShogiGUI [4] under the rule that each move must be made within 20 seconds. ShogiGUI (Fig. 1) is a graphical user interface software for Shogi that runs on Windows. In ShogiGUI, players make their moves by dragging and dropping the pieces. To adjust the strength of the opponent, the number of positions 'Suisho5' searches is adjusted by setting the node limit between 50 and 100, allowing for a match where humans can win. Before the experiments, the subjects adjust the settings by playing against the AI to ensure it is neither too strong nor weak as an opponent.

2.3. *Gaze features*

To obtain gaze information, we use the Tobii Pro X3-120 [5]. This eye tracker is a screen-based model installed at the bottom of the display with a sampling rate of 120Hz. This eye tracker can collect gaze coordinates and pupil diameter on the display; in this research, we use the gaze coordinates. The coordinates originate at the top left, with the horizontal direction as the x-axis and the vertical direction as the y-axis. The gaze heat map created from the collected data during the matches is shown in Fig. 2.

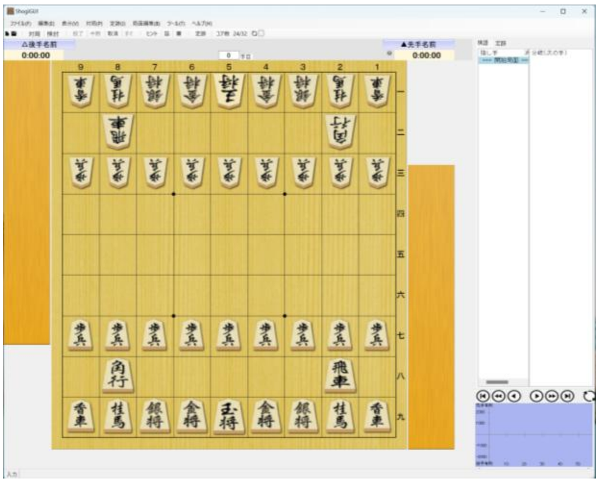

Fig.1 ShogiGUI.

Fig.2 Gaze Heat Map During the Match.

Fig. 2 shows a heat map generated from the gaze information for a single move, with a Gaussian distribution overlay on each gaze coordinate to create the heat map. From the heat map, it is possible to understand what moves the player is considering and what moves they are deliberating between.

2.4. *Annotation*

Annotations of careless mistakes are conducted. After the game, subjects annotate the game while viewing the game in ShogiGUI.

A dataset is created based on the gaze data and annotations during the match. First, the gaze data is divided into segments between the moment the opponent places a piece and when the subject places their piece. This study does not analyze the gaze data during the AI's thinking period. The moves the subject identified as having a careless mistake are annotated as such, while all other moves are annotated as without careless mistakes. Any missing values in the gaze data are entirely removed.

3. Analysis of Gaze Features

3.1. *Analysis method*

We extract features from the created dataset and conduct tests to analyze whether there is a significant difference in gaze information related to the occurrence of careless mistakes.

3.1.1. *Testing method*

In this study, F-tests and t-tests are conducted. The F-test is a method used to determine whether there is a significant difference in the variances of two data groups. In comparison, the t-test checks for a substantial difference in the means of the two data groups. First, we conduct an F-test to check whether the two data groups have equal variances. If they are equal, a Student's t-test is performed; if not, a Welch's t-test is conducted. Additionally, the significance level is set at 0.05 for both tests.

3.1.2. Gaze features

The features used for the tests are shown in Table 2. The mean and variance are calculated based on the gaze coordinates on the display. Additionally, the movement amount is determined by calculating the Euclidean distance between the gaze coordinates at the current timestamp and those from one timestamp prior, as well as by computing their sum or average. Regarding the skewness and kurtosis of the movement amount, skewness indicates the symmetry of the distribution, while kurtosis reflects the sharpness of the distribution. Based on these features, we conducted tests for each left eye, right eye, and both eyes. The data for both eyes is treated as the average of the left and right eyes.

3.2. *Analysis results*

3.2.1. *F-Test*

The results of the F-test are shown in Table 3. According to Table 3, there is a difference in the variance of both eyes' movement amount, skewness, and the kurtosis of the left eye's movement amount between the times when careless mistakes were made and when they were not.

Table 2. Features Used for the Tests.

Gaze Coordinates	Mean	x-axis
		y-axis
	Variance	x-axis
		y-axis
Movement of Gaze Coordinates	Sum	
	Mean	
	Variance	
	Skewness	
	Kurtosis	

3.2.2. *t-Test*

The results of the two-tailed tests are shown in Table 4. For the variance and skewness of the movement amount of both eyes and the kurtosis of the left eye's movement amount, which were found to have a significant difference in the F-test, Welch's t-test is performed. For the other data, a Student's t-test is conducted.

According to Table 4, there is a significant difference in the skewness and kurtosis of the right eye's movement amount between careless mistakes and not.

According to Table 3 and Table 4, several features show significant differences; however, many do not. One possible reason for this is that the sample size needs to be bigger, which may have affected the effectiveness of the tests. Therefore, increasing the number of data points in the future is necessary.

In this experiment, tests were conducted regarding gaze coordinates and gaze movement, but the treatment of missing values and the discussion of fixations and saccades are insufficient. Therefore, these factors must be considered in future analyses.

Table 3. F-test.

			left eye	right eye	both eyes
Gaze Coordinates	Mean	x-axis	0.413	0.469	0.293
		y-axis	0.346	0.297	0.353
	Variance	x-axis	0.367	0.240	0.266
		y-axis	0.236	0.300	0.238
Movement of Gaze Coordinates	Sum		0.394	0.247	0.156
	Mean		0.339	0.399	0.338
	Variance		0.062	0.098	**0.022**
	Skewness		0.083	0.116	**0.049**
	Kurtosis		**0.006**	0.079	0.131

Table 4. t-test (Two-tailed Test).

			left eye	right eye	both eyes
Gaze Coordinates	Mean	x-axis	0.137	0.161	0.435
		y-axis	0.071	0.080	0.075
	Variance	x-axis	0.745	0.892	0.309
		y-axis	0.488	0.675	0.511
Movement of Gaze Coordinates	Sum		0.279	0.224	0.156
	Mean		0.904	0.756	0.791
	Variance		0.608	0.970	0.741
	Skewness		0.798	**0.026**	0.562
	Kurtosis		0.913	**0.042**	0.709

4. Conclusion

Using gaze information during shogi matches, we employed F-tests and t-tests to test whether there is a significant difference between the gaze associated with careless mistakes and the gaze when no careless mistakes were made.

Through the tests, it was clarified that there are significant differences in features concerning variance and mean of gaze coordinates. In the future, this gaze data and the extracted features are expected to contribute to developing a model for predicting careless mistakes. Additionally, while this study utilized gaze data during shogi, where careless mistakes are likely to occur, it is essential to analyze under what circumstances careless mistakes happen beyond just shogi, broadening the discussion to other events. This remains a key challenge for the future.

References

1. S L Happy, Anirban Dasgupta, Priyadarshi Patnaik, and Aurobinda Routray. Automated alertness and emotion detection for empathic feedback during e-learning. In 2013 IEEE Fifth International Conference on Technology for Education (T4E), pp. 47–50, 2013.
2. Salma Alhagry , Aly Aly Fahmy, and Reda A. El-Khoribi. Emotion recognition based on EEG using LSTM recurrent neural network. Emotion vol. 8, no.10, pp. 355–358, 2017.
3. yaneurao, “YaneuraOu,” https://github.com/yaneurao/YaneuraOu, (accessed on December 11, 2024).
4. siganus, "Shogi GUI", http://shogigui.siganus.com/, (accessed on December 10, 2024).
5. tobii, ”Tobii Connect”, https://connect.tobii.com/s/x3-downloads?language=ja, (accessed on December 10, 2024).

Authors Introduction

Mr. Ryota Yabe

He received his Bachelor’s degree in Computer Science and Systems Engineering in 2024 from the Faculty of Computer Science and Systems Engineering, Kyushu Institute of Technology in Japan. He is currently a master student at Kyushu Institute of Technology, Japan

Dr. Noriko Takemura

She received B.S., M.S., and Ph.D. degrees in engineering from Osaka University, Japan, in 2006, 2007, and 2010, respectively. She is currently an associate professor at Graduate School of Computer Science and Systems Engineering , Kyushu Institute of Tech-nology. Her current research interests include human-centric intelligent systems