EMITTER International Journal of Engineering Technology
Vol. 10, No. 2, December 2022, pp. 338~350
DOI : 10.24003/emitter.v10i2.756

Estimation of Confidence in the Dialogue based on Eye Gaze and Head Movement Information

Dewen Cui[1], Akihiro Matsufuji[2], Yi Liu[1], Eri Sato-Shimokawara[1], Toru Yamaguchi[1]

[1]Computer Science Department, Tokyo Metropolitan University, Tokyo, Japan.
[2]Electric Works Company, Panasonic Corporation, Osaka, Japan.
Correspondence Author: cui-dewen@ed.tmu.ac.jp

Received September 30, 2022; Revised November 15, 2022; Accepted November 21, 2022

Abstract

In human-robot interaction, human mental states in dialogue have attracted attention to human-friendly robots that support educational use. Although estimating mental states using speech and visual information has been conducted, it is still challenging to estimate mental states more precisely in the educational scene. In this paper, we proposed a method to estimate human mental state based on participants' eye gaze and head movement information. Estimated participants' confidence levels in their answers to the miscellaneous knowledge question as a human mental state. The participants' non-verbal information, such as eye gaze and head movements during dialog with a robot, were collected in our experiment using an eye-tracking device. Then we collect participants' confidence levels and analyze the relationship between human mental state and non-verbal information. Furthermore, we also applied a machine learning technique to estimate participants' confidence levels from extracted features of gaze and head movement information. As a result, the performance of a machine learning technique using gaze and head movements information achieved over 80 % accuracy in estimating confidence levels. Our research provides insight into developing a human-friendly robot considering human mental states in the dialogue.

Keywords: human-robot interaction, gaze, confidence level, mental states, machine learning

1. INTRODUCTION

Biometric information, such as gaze, posture, and heartbeat obtained from intelligent sensors, is beneficial for understanding human mental states [1]. In the field of learning support, efforts to improve the learning efficiency of learners by estimating their mental states such as Emotions, conversational climate, and intention using this information have been actively studied in recent years [2,3,4]. Human physiological features can more objectively and accurately reflect a person's most natural mental state than mental state recognition based on expressions, speech, and other external features. While work based on facial expressions or speech abounds, non-verbal information recognition is less explored. Among the various observable human cognitive information, eye movements can effectively reflect changes in people's visual focus and, to some extent, the state of their mental activity. Eye movement is closely related to attention to targets, internal emotions, and task intentions. Capturing eye movement signals and analyzing the underlying eye movement information can provide more clues for mental state recognition and make eye movements an ideal input for building natural human-robot interaction systems. Therefore, eye-movement-based mental state recognition has essential research significance and application value.

Adaptive interaction with users based on human-robot interaction and providing quality services based on user needs and changes in environmental information have become the development trend for a new generation of intelligent robots. It is vital to improving robots' intelligence to sense their surroundings and understand human emotions, intentions, and service needs. As expectations grow for robots to communicate with humans in their lives, while fluency in communicating with the user is important for social robots, current dialogue-processing technologies are not yet desired. The problem with conventional dialogue technology is that robots may not correctly understand the meaning of the user's words and may thus perform unexpected responses. Therefore, to enable the correct understanding of the user's biological information and the meaning of the words, it is necessary to enable the user's speech to be correctly understood. Studies have been carried out in spoken dialogue in which confidence is presumed based on the results of speech recognition and action recognition. So, this research focuses on people's mental states and explores the certainty with which people respond to common sense questions in dialogue. People's mental states change when they are convinced and unconvinced of what they say. By estimating the degree of confidence, we can extract how people behave when they are uncertain and use the relationships to support research into robots that can interact with people in a more natural way.

2. RELATED WORKS

Eye movement technology plays a crucial role in human-computer interaction (HCI) research as the primary method of collecting visual information [5], particularly eye-tracking and eye-control techniques. For example, the development of the Tobii Pro Glasses made it possible for researchers to collect high-quality eye-tracking data and broadened the research field. Kumar et al. [6] proposed a Gaze-enhanced User Interface Design. Eye movement and mouse are used together as interaction methods, with eye movement as the primary input device and keyboard and mouse as a supplement. In the same year, they used eye control technology [7] first to examine the activities of a regular mouse and then compare the similarities and differences between eye-control methods and standard mice in terms of clicking. Eye movement technology revolutionizes the meaning of HCI by establishing a dialogue with the user's Areas of interest (AOI) and will play an essential role in the future of multi-channel HCI.

Various data, such as through speech and body movement, have been used to make presumptions about the confidence of speakers [8]. Nowadays, eye gaze information is also increasingly used in studying mental states, including stress [9], interest, confusion, and understanding, such as obtaining answers to questions through appropriate learning activities [10]. Eye gaze helps observe higher-order cognitive processing in humans. However, it is not easy to collect stable data, and data analysis methods have not been established [11]. Therefore, it is necessary to use a method based on gaze data to analyze people's problem-solving and diagnose their mental states. It is also essential to collect gaze data and understand their tendencies. A study by Ibrahim et al. examines students' visual behaviors when they tackle sequential and simultaneous problems [12]. To devise a method for examining problem-solving strategies, they studied the gaze transitions during the answering process. Yamada et al. used multiple-choice questions in their confidence estimation [13]. They proposed an application for multiple-choice answering that used an eye-tracker to analyze the user's reading and answering behavior, extract some features and estimate whether they answered confidently.

3. ORIGINALITY

Gaze tracking technology has a wide range of applications in areas such as human-computer interface interaction, educational support, and aerospace medical research. Head movements are important concomitant behaviors in the human visual gaze process and are of great importance to human vision behaviors. In previous studies, researchers have Combined head movement and gaze tracking through data fusion techniques to infer insider information such as people's conversational engagement, behavioral user state, and levels of attention [14,15,16]. This can open up new ideas for improving the reliability and practicality of gaze tracking. However, most of the research on confidence is based on systematic user interfaces that analyze people's gaze information during problem-solving and problem-processing. In this paper, we apply a new way of combining information from both and focus on the estimation of human confidence levels in human-robot dialogue. thereby providing a better estimate of the participants' confidence level. Furthermore, we chose to select features from the most representative data, the coordinate gaze data, and head movement data, and to adopt different processing methods for their respective characteristics.

4. SYSTEM DESIGN

4.1 Gaze2D data

We utilized the Tobii Pro Glasses 3 to measure the user's eye movement and obtain the gaze information. It provides a way to simultaneously collect multi-channel data, such as surroundings, ambient sound, and head movement. The device's built-in gyroscope and video camera recorded information such as the participants' head tilt, gaze video, and gaze direction. We use The Tobii Pro Glasses 3 Controller software to acquire the data. This experiment analyzes the collected gaze2D data and inertial measurement unit (IMU) data. Gaze2D data (Fig. 1) contain a 2D vector of the format: n = [x, y], where x and y are the normalized horizontal and vertical coordinates, respectively, the estimated position of the gaze concerning the calibration screen. More precisely, the point n = [0,0] lies at the top left corner of the Perspectives and n = [1,1] at the bottom right [17]. We used these data to determine the location and movement of the gaze point and to infer the confidence level of a quiz based on all changes in the user's gaze point in this response.

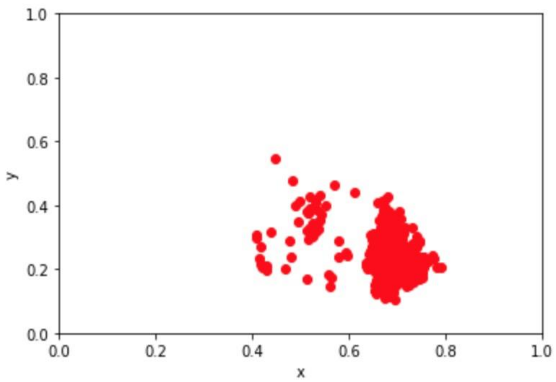

Figure 1. Example of Gaze2D data

4.2 Average Nearest Neighbor Index

Clark and Evans define the Average Nearest Neighbor Index (NNI) [18] as the distance between each feature centroid and its nearest neighbor's centroid location. It then averages all these most relative neighbor distances. If the average length is less than the average for a hypothetical random distribution, the distribution of the analyzed features is considered clustered. If the average distance exceeds a theoretical random distribution, The features are dispersed. The average nearest neighbor ratio is calculated as the following formula. where $\overline{Do}$ is the observed mean distance between each point and its nearest neighbor. and $\overline{Dr}$ is the expected mean distance for the points given in a random pattern. In the above equations, d_i equals the distance between point i and its nearest neighboring point, n corresponds to the total number of points, and A is the area of a minimum enclosing rectangle around all points, or it's a specified Area value. NNI is used in many studies to describe the spatial distribution of data. It can therefore be used for feature extraction of gaze information. Amrouche et al. propose a dynamic activity segmentation algorithm by exploiting the expressive properties of the distribution-based gaze feature NNI [19]. Mannaru et al. propose a novel way of computing NNI based on continuous hidden Markov models that model the gazes as 2D Gaussian observations [20]. Wolf et al. used NNI as one of the ocular parameters to estimate cognitive load [21]. We have used the NNI to process the gaze2D data. The spatial distribution of gaze points is described by the value of the NNI and used for feature extraction of gaze information.

$$NNI = \frac{\bar{D}_O}{\bar{D}_E} \tag{1}$$

$$\bar{D}_O = \frac{\sum_{i=1}^{n} d_i}{n} \tag{2}$$

$$\bar{D}_E = \frac{0.5}{\sqrt{n/A}} \tag{3}$$

4.3 Gaze data processing

We used Tobii Pro Glasses 3 to obtain information about the participants' gaze points while answering the questions. Next, we transformed the gaze data, which represents coordinate position information, into NNI data, representing the offset of the gaze at each period. Finally, NNI feature extraction was applied to the incoming 2D gaze coordinate stream (x, y position in normalized video coordinate [0, -1]). The output of this stage is a continuous feature signal and exhibits different signal levels at other moments [22]. Smooth low regions versus higher fluctuations regions (Fig. 2). The pattern indicates clustering if the index is less than 1. If the index is greater than 1, the trend is toward dispersion, with the peak representing the highest value of the signal at a given time. Each response of the user has been reflected in a different waveform accordingly. We base our feature extraction on these continuous signal variations.

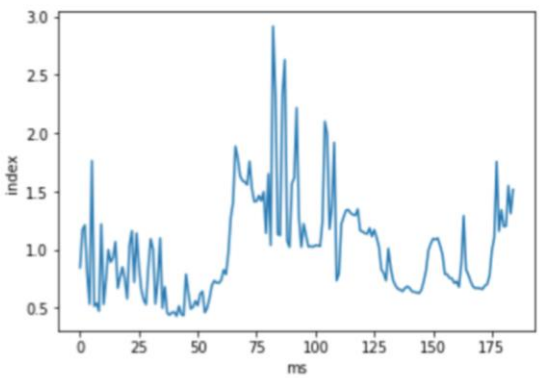

Figure 2. Example of NNI response segment

4.4 IMU data

The Head Inertial Measurement Unit (IMU) is a device that detects three-dimensional inertial motion (translational and rotational motion of three orthogonal axes). An accelerometer detects translational motion, and an angular rate sensor detects rotational motion. By combining the data from these three physical fields, the relative three-dimensional orientation of the IMU is known, and by extension, the user's head pose. Acceleration is measured along three axes in the head unit coordinate system in m/s^2. When at rest, the IMU sensor in the head unit will report an acceleration of 9.8 m/s^2 negative value in the y-component. The accelerometer is sampled at 100 Hz. The gyroscope measures rotation around three axes. A yaw movement will rotate around the Y-axis, A pitch movement will give a rotation around the X-axis, and A rolling movement will give a rotation around the Z-axis. Rotation is measured around each axis in degrees per second and is sampled at 100 Hz.

4.5 IMU data processing

Although gaze changes should be measured directly by the eye tracker, gaze shifts may be accompanied by head movements. From the study [23], it is known that the head posture data obtained from the head tracker, and in particular the amplitude information indicating the amount of change in head position and rotation angle, correlates moderately with participants' motivation to engage in dialogue and is useful for estimating attitudes to engage in dialogue. Head acceleration and rotation angle were plotted as waveforms, and their amplitude can be calculated. The greater the amplitude of this head position and rotation angle, the less positive the participant's attitude.

4.6 Feature Selection

We selected those features that effectively estimated confidence levels, as shown in Table 1. Four of these were related to gaze movements, and six were related to head movements. In addition, we selected the duration used to answer the question as a feature. In total, there were 11 feature values. When participants answered the questions, there was a significant difference in the amount of thinking time required to answer the questions in easy and difficult situations. Therefore, this was used as a feature. As generally, the user's confidence level in the answers they give does not change depending on whether they are correct. So, we did not consider the correctness of the answers. Here we use a support vector machine (SVM) to estimate confidence and use the most basic level of confidence, in which we classify the level of confidence into two categories, " Confident" and "Not Confident."

Table 1. Features for machine learning

Category	Features
Gaze	max NNI mean NNI std NNI Ratio mean-std
IMU	Amplitude of the head acceleration relative to the {x, y, z} axis Amplitude of the {roll, pitch, yaw} rotation angles
Others	duration of answer

5. EXPERIMENT AND ANALYSIS

The experiments conducted are shown in Fig. 3. We operated the robot interacting with the experimental participants in a Wizard-of-Oz (WoZ) [24] manner. A human operator observed the participants remotely, controlling the system's responses through a dedicated program. We did not disclose this to the experimental participants before or during the experiment. We conducted Dialogues in a one-question-and-answer format and set the dialogue unit to be each response of the experimental participant to the question. Specifically, it was defined as one exchange from the end of the robot discourse (i.e., the question to the experimental participant) to the end of the participants' discourse. The start of the robot discourse (when the operator Wizard presses the start button) and the end of the robot discourse is recorded by the button operation. Whenever the robot asked the participant a question during the experiment, the participant had to think and answer. After finishing responding, we used a button to move the robot to the next question. This experiment aimed to obtain data for estimating the presence or absence of confidence in answering a single question. 10 participants took participated in the experiment, including four females and six males. All of the participants are university students in their 20s (24 ± 4 years, range: 21-28 years). We use vision-adjusting lenses for Tobii Pro Glasses. to ensure that each participant's had normal or corrected-to-normal vision. In advance, we set up the 64 questions asked by the robot to be half complex and half simple, as far as possible, to collect confident/not confident data. These questions are asked by the robot, and participants answer them during human-robot communication. We asked them to fill in their level of confidence (most confident is 5 and not confident is 1) in answer to each question after the experiment and eliminated label imbalances in post-processing of the collected data.

Figure 3. Experimental View

We have created 64 question-answer lists extracted from the Japanese website (https://bonjin-ultra.com/mondai8.htm). Table 2 shows the example of questions. This collection of questions covers various issues from science, literature, and general knowledge, and it also contains the most basic questions, such as names and ages. According to participants' feedback, the question's ratio of difficulty was roughly half to half. We set a confidence label in each dialogue unit of all participants and used it to indicate their confidence level about this question. Furthermore, we manually assigned confidence labels to each dialogue unit collected from the questionnaire and gave each confidence label on a 5-point scale. One point indicated that the participant was not confident. While five points indicated that the participant was confident. The confidence level is independent of the correctness of the answer given. On the other hand, If the participant provides an incorrect answer and is still very confident in the answer, then the confidence label still counts as a 5 for this occasion. We asked participants to move from the first question to the last one and report their confidence level for each dialogue unit.

Table 2. Example of the question list

Question list
What is the population of Japan?
What is the capital of Greece?
Who is the author of Romeo and Juliet?
Who discovered gravity?
What did Madame Curie discover?
In which month does Christmas fall?
Who painted the Mona Lisa?
Who invented the telephone?
What is the smallest prime number?
What turns into dry ice when it freezes?

During the experiment, the Tobii Pro Glasses 3 captured the eye movement patterns of the observer and obtain a video recording of the participant's perspective. We intercepted the corresponding video clip during each dialogue and plotted the gaze movement (Fig. 4) and heat map (Fig. 5) of the participant's gaze points moving across the screen from the beginning to the end of each dialogue. It is easy to see that whenever the robot asks a question, most of the participant's gazes are focused on the robot's face. We also analyze it with the confidence labels to estimate whether they were confident or unconfident in the dialogue and filter out unusual gaze movements and confidence labels. The video shows that when participants are optimistic about the question, participants eye movements center on the robot's face, with a small range of motion and short reaction times. On the other hand, when participants were unconfident in their answers, the gaze point shifted significantly during the thinking process, with an extensive range of movement and a longer response time. Moreover, there are different positions and degrees of shift depending on the participants.

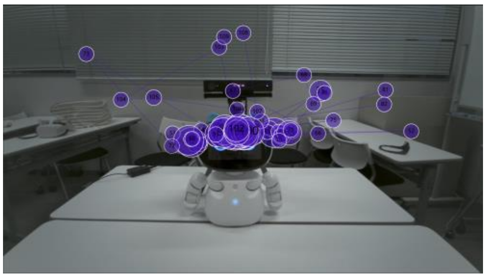

Figure 4. Example of gaze movement Trajectory during answer with confidence

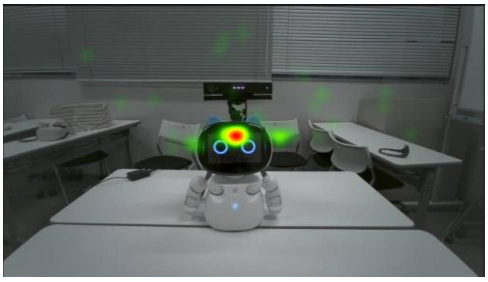

Figure 5. Example of Heat map during answer with confidence

We correlated the degree of dispersion of the gaze point with its variation, the amplitude of head position and angle, and the length of time spent answering the question with the confidence level. Machine learning (SVM: Support Vector Machine) was conducted using these parameters to verify whether they were useful for estimating confidence levels toward the dialogue. For this machine learning, conversational confidence was transformed into the two values of confident and unconfident. The confidence level was categorized as whether the participant had confidence in the dialogue. An example of the gaze dispersion of participants who responded with and without confidence was shown (Fig.6). The graphs represent the mean of a participant's NNI per dialogue unit. When participants answered confidently (dots), the values were lower. In the opposite case, the values are higher (crosses).

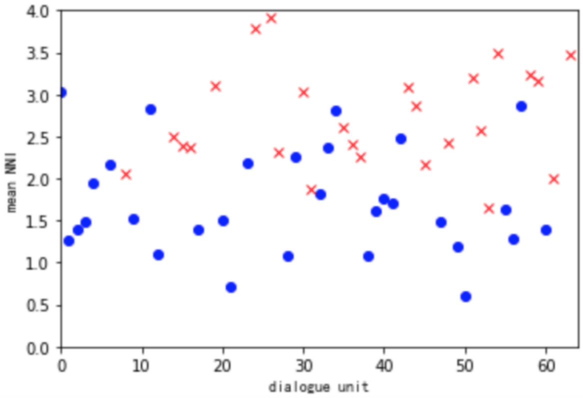

Figure 6. Gaze dispersion per dialogue unit

To demonstrate the validity of this study, we constructed several methods for comparison and examined the differences in estimation accuracy between them. The results of the research are shown in Table 3. We estimated the confidence by using: 1) only the features of the gaze or head, 2) features of the gaze or head and answering time lengths 3) the features of the gaze and head 4) all features including the eye gaze (the proposed method). First, we did not consider the two together in order to explore the effect of gaze and head features on the presumed results. The proposed methods were then compared with the other five cases. The estimated accuracy for each case was derived.. We trained machine learning models SVM with each participant's features to determine confidence and accuracy. Confidence labels 1, 2, 4, and 5 were used as confidence labels. The 30% of the data after excluding the missing data was then used as the test data and the rest as the training data. Confidence levels for binary classification were learned by dividing the data into positive and negative samples. Based on this information, we were able to make accurate predictions.

Table 3. Result for participants by the proposed method

Feature	Ave. Accuracy	Ave. Recall	Ave. Precision	Ave. F-measure
Gaze	68.6%	76.0%	74.0%	75.0%
Head	71.8%	64.9%	89.3%	75.2%
Gaze + Time	75.6%	67.5%	94.7%	78.8%
Head + Time	77.3%	79.0%	89.2%	83.8%
Gaze + Head	78.5%	81.5%	85.7%	83.5%
All Features	81.8%	84.0%	88.3%	86.1%

A total of 11 features were used, and by adding separate temporal features to the gaze and head features, were able to improve the accuracy of the predictions so the length of time was a reasonable indicator of confidence in the predictions. The results show that combining the gaze and head features with temporal features gives high accuracy results when using the selected features for machine learning. A further selection of better gaze and head features and further analysis of the correlation between gaze movement and head offset in order to better fusion of the two from the data collected. , If the range of labels we chose were reduced to 1 and 5, this would improve the accuracy of the predictions. However, due to individual differences in each individual's confidence profile when answering the questions, it also exhibited varying degrees of variation in the physiological state depending on the confidence level of the experimental participants. These results suggest that specific trait values, such as head excursion, may be valid for some situations but not others. These depend on the participants' personalities, strategies, and problem-solving habits. Therefore, further examination of different participants' characteristics is needed.

6. CONCLUSION

In this paper, we present a method to estimate the confidence of participants' responses to questions in human-robot interaction using a machine-learning model that predicts the mental state of the human psyche from the coordinates of the gaze and the tendency to move the head. In addition, it makes inferences about confident and unconfident dichotomous results. Relative to previous experiments conducted in similar environments and conditions using audio-vocal and head data [25], we achieved a improvement in the results obtained and validated. We believe that most of each user's response fragments can correspond to their own given confidence labels. However, to further improve the accuracy of the estimates, we needed to find a way to fix specific gaze and head positions through the data more accurately. The biometric differences between participants in the experiment and individual differences also deserve further analysis and discussion. In future works, we plan to combine external and internal biological information. Moreover, we plan to improve the interaction way the robot asks questions of the user and the type of questions it asks to provide a more accurate estimate of a user's confidence level.

Acknowledgments

This work was supported by JST [Moonshot R&D] [Grant Number JPMJMS2034]

REFERENCES

[1] Noroozi F., Corneanu C. A., Kamińska D., et al., **Survey on emotional body gesture recognition[J].** *IEEE transactions on affective computing,* 2018, 12(2): 505-523.

[2] Defu, Z., Matsufuji, A., Sato-Shimokawara, E., Yamaguchi, T., **Emotion Recognition based on speech data containing personal differences,** *International Symposium on Computational Intelligence and Industrial Applications,* 2018.

[3] Matsufuji, A., Shiozawa, T., Hsieh, W. F., Sato-Shimokawara, E., Yamaguchi, T., and Chen, L. -H., **The analysis of nonverbal behavior for detecting awkward situation in communication.** *In 2017*

Conference on Technologies and Applications of Artificial Intelligence (TAAI) (pp. 118-123). IEEE.

[4] Matsufuji, A., Sato-Shimokawara, E., Yamaguchi, T., **A method for estimating speaker's intention using human gestures and acoustic features in dialogue,** *Annual Conference of the Robotics Society of Japan,* 2017.

[5] Langton S. R. H., **The mutual influence of gaze and head orientation in the analysis of social attention direction,** *Quarterly Journal of Experimental Psychology,* A, 53(3), 825–845.

[6] Klyde David H., McCruer Duane T., Myers Thomas T., **Unified Pilot-Induced Oscillation Theory, Volume I: PI0 Analysis with Linear and Nonlinear Effective Vehicle Characteristics, Including Rate Limiting,** *WL - TR - 96 - 3028. Air Force Research Laboratories, Wright - Patterson AFB OH,* December 1995.

[7] Karlsson P., Allsop A., Dee-Price B. J., et al. **Eye-gaze control technology for children, adolescents and adults with cerebral palsy with significant physical disability: Findings from a systematic review[J].** *Developmental neurorehabilitation,* 2018, 21(8): 497-505.

[8] Kasano, E., Muramatsu, S., Matsufuji, A., Sato-Shimokawara E., and Yamaguchi, T., **Estimating Speaker's Confidence in Dialogue Using Speech and Motion Information,** *Conference on System Integration of the Society of Instrument and Control Engineers,* pp.729-732, 2018.

[9] Jyotsna, C., Amudha J., **Eye gaze as an indicator for stress level analysis in students.** *In 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI),* pp. 1588-1593, IEEE.

[10] Emerson A., Cloude E. B., Azevedo R., et al. **Multimodal learning analytics for game-based learning[J].** *British Journal of Educational Technology,* 2020, 51(5): 1505-1526.

[11] Graziano M. S. A., Guterstam A., Bio B. J., et al. **Toward a standard model of consciousness: Reconciling the attention schema, global workspace, higher-order thought, and illusionist theories[J].** *Cognitive Neuropsychology,* 2020, 37(3-4): 155-172.

[12] Ibrahim B., Ding L., **Sequential and simultaneous synthesis problem solving: A comparison of students' gaze transitions[J].** *Physical Review Physics Education Research,* 2021, 17(1): 010126.

[13] Kento Y., Ayano O., Hiroki F., Kensuke H., Olivier A., Koichi K., **Confidence estimation based on gaze while answering English multiple-choice questions,** *Shingaku Giho,* vol. 116, no. 461, PRMU2016-192, pp. 199-204, 2017.

[14] Chihara T., Kobayashi F., Sakamoto J., **Evaluation of mental workload during automobile driving using one-class support vector machine with eye movement data[J].** *Applied ergonomics,* 2020, 89: 103201.

[15] Stephenson L. J., Edwards S. G., Bayliss A. P., **From gaze perception to social cognition: The shared-attention system[J].** *Perspectives on Psychological Science,* 2021, 16(3): 553-576.

[16] Asteriadis S., Karpouzis K., Kollias S., **The Importance of Eye Gaze and Head Pose to Estimating Levels of Attention.** *Third International Conference on Games & Virtual Worlds for Serious Applications. IEEE,* 2011.

[17] Torrente M., Guillem, **Mobility for the severely disabled: a head-controlled wheelchair,** 2017.

[18] Sillero N., **A Simple Spatial Method for Identifying Point Clusters by Neighbourhood Relationships[J].** *Ecologies,* 2021, 2(3): 305-312.

[19] Amrouche S., Gollan B., Ferscha A., et al. **Activity segmentation and identification based on eye gaze features.** *Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference.* 2018: 75-82.

[20] Mannaru P., Balasingam B., Pattipati K., et al., **On the use of hidden Markov models for gaze pattern modelling.** *Next-Generation Analyst IV. SPIE,* 2016, 9851: 252-258.

[21] Wolf E., Martinez M., Roitberg A., et al., **Estimating mental load in passive and active tasks from pupil and gaze changes using bayesian surprise.** *Proceedings of the Workshop on Modeling Cognitive Processes from Multimodal Data.* 2018: 1-6.

[22] Sabrina A., Benedikt G., Alois F., Josef H., **Activity Segmentation and Identification based on Eye Gaze Features,** *Proceedings of the 11th Pervasive Technologies Related to Assistive Environments Conference,* June PP. 75–82, 2018

[23] Ijuin K., Jokinen K. J., Kato T., et al. **Eye-gaze in Social Robot Interactions Grounding of Information and Eye-gaze Patterns[C].** *Proceedings of the Annual Conference of JSAI 33rd (2019).* The Japanese Society for Artificial Intelligence, 2019: 3J3E402-3J3E402.

[24] Browne J T., **Wizard of oz prototyping for machine learning experiences[C]. Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems.** 2019: 1-6.

[25] Matsufuji, A., Sato-Shimokawara, E., and Yamaguchi, T., **Adaptive personalized multiple machine learning architecture for estimating human emotional states.** *Journal of Advanced Computational Intelligence and Intelligent Informatics,* 2020, 24(5): 668-675.