Reading Behavior Analysis with Gaze Tracking Data

Wen-Chung Kao, *Senior Member*, *IEEE*, Chun-Wei Shen, and Chao-Jung Wu

***Abstract*— Gaze tracking has become one of the most attractive human machine interfaces on consumer electronics. In addition to providing intelligent interactive functions on the consumer electronics, it also makes possible to collect the data of user attention. Furthermore, the data could be analyzed to predict the reading behavior and the information processing in the brain. In this paper, we propose an automatic tool for the gaze data analysis. A mathematic examination system is used for demonstrating the system performance. According to the gaze data, the experimental results show the feasibility of estimating the learner's ability, and the accuracy has achieved higher than 90%.**

I. INTRODUCTION

It is expected that integrating smarter human machine interfaces (HMI) would become the mainstream of consumer electronics and plenty of them could be achieved based on the built-in cameras. Among all of HMIs, the applications of gaze tracking [1] have prompted widespread concern [2]. That is, it makes possible to collect user behavior as well as intention [3]-[4] and provide intelligent interactive functions on the consumer electronics. Once the collection is accomplished, the gaze data will be used to analyze the reading behavior and the information processing in the brain. Thus, it is a must to develop a comprehensive data analysis tool in order to initiate this kind of intelligent HMIs on the consumer electronics.

Hence, in this paper, we propose an automatic tool for the gaze data analysis. First, all the data will be classified into fixations and saccades with a parser. Then more than 40 features will be extracted. A mathematic examination system has been used to demonstrate the performance of the proposed gaze data analysis system. In the end, the evaluation has proved that it is feasible to estimate the learner's ability with the gazedata, and the accuracy has achieved higher than 90%.

II. FEATURE EXTRACTION FROM GAZE DATA

First of all, all gaze data should be classified into fixations or saccades. Typically, human eyes are inclined to keep moving to glance through different portions of the content. The movements of the eyes are called saccades, while the eyes, which has remained still between saccades are regarded as the fixations. The eye movement trajectory, composed of a sequence of fixations and saccades, has contained a substantial amount of information of human reading behavior or the cognition process. In the proposed system, 43 features are extracted from the eye movement trajectory. The features encompass the statistics of the duration as well as the count of fixations and saccades. In addition, the system also detects the regression, which means the user has watched two regions alternatively. The regressions between the pairs of regions on the reading content indicate the cognitive process of the user reading/learning. Take a mathematical question as an example, the learner may read the text first, then the visual focus will move to the corresponding chart. Following that, the focus may move back to the text to confirm the details of the question. This process could be repeated several times for clarifying the question. It is assumed that the statistics of regression and the sequence will vary widely depending on the ability of the students. The other important features are the percentage of fixation duration for specific regions. The excellent students seem to spend more time on the key statements as along with the portions of the plots, whereas the others would have the difficulty in identifying the clues immediately. These observations confirm that the learning performance could be estimated by analyzing the gaze tracking data without asking the students to answer the questions.

III. CLASSIFICATION

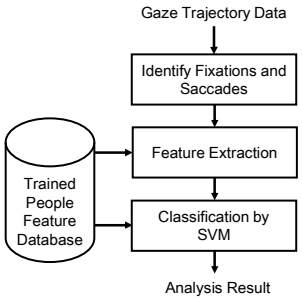

Fig. 1. Flowchart of reading behavior analysis.

As shown in Fig. 1, the proposed system first identifies the fixations and saccades from the gaze trajectory data. Then it enumerates a set of features from the detected fixations/saccades. The system adaptively selects features for different applications. The usability of extracting features appears to be quite diversified under different conditions. In the proposed data analysis system, all extracted features are arranged into a feature vector. On the other hand, a few features that are stable enough in the same class while having significant differences between the two classes will be

The authors are with National Taiwan Normal University. This research is partially supported by the "Aim for the Top University Project" and "Center of Learning Technology for Chinese" of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and Ministry of Science and Technology, Taiwan, R.O.C. under Grant no. MOST 104-2911-I-003-301 and MOST 104-2511-S-003 -016 -MY3.
Email: jungkao@ntnu.edu.tw

selected [5]. In order to realize this concept, the importance of a feature of the two classes are evaluated as

$$I = (\mu_a - \mu_b)^2/(\sigma_a^2 + \sigma_b^2) \quad (1)$$

where μ_i and σ_i denote the mean and standard deviation values of a feature type for all training samples in class i, respectively. Based on (1), the feature types are ranked based on their importance for the classification problem. The optimal dimensionality as well as feature types will be determined according to the classification result. The adopted classifier is the support vector machine (SVM), which maps the input feature vector into a high dimensional space to perform the linear classification. By feeding the classifier with a set of training data, an optimal hyperplane, which minimizes the risks for each pair of classes, can be established.

IV. Reading Behavior Analysis

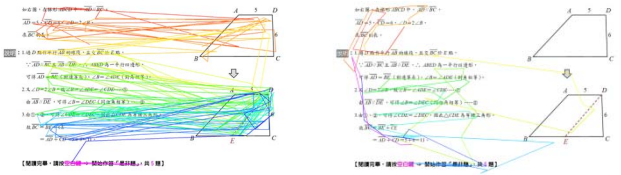

Fig.2. Two examples of gaze trajectories. Left: HA learners, Right: LA learners.

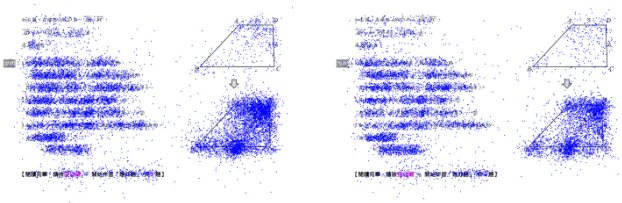

Fig. 3. Two examples of detected fixations. Left: HA learners, Right: LA learners.

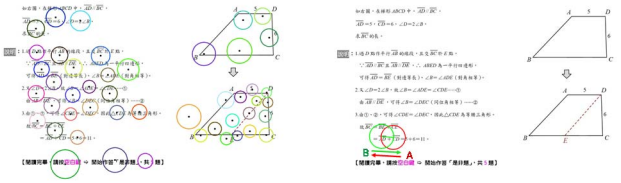

Fig. 4. Left: The fixation clusters for HA learners, Right: the definition of regression.

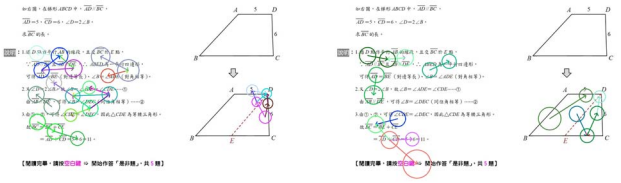

Fig. 5. The regression pairs. Left: HA learners, Right: LA learners.

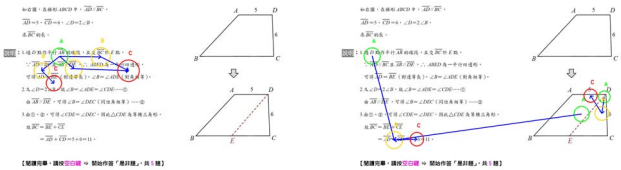

Fig. 6. The viewing sequences. Left: HA learners, Right: LA learners.

Fig. 2 shows two examples of gaze trajectories whose color indicates the viewing sequence from red to purple. These two cases are extreme ones which can be distinguished easily. By contrast, in most of the cases, the user ability cannot be identified by merely gaze trajectories. A general approach to the analysis of human reading behavior is parsing the trajectories into fixation and saccades. Fig. 3 reveals the detected fixations for the HA and LA sets, which is not easy to differentiate between the two cases. It is difficult to differentiate these two classes with the distribution of fixations points. Thus the next step is trying to make groups of fixations and ascertain their regression relationship. The left figure of Fig. 4 shows an example of fixation group and the regression group can be extracted based on the right figure of Fig. 4. If the viewing sequence of the user is region A, region B, and then region A, it will be identified as a regression. Fig. 5 shows the detected regression pairs for HA and LA people. It is quite obvious that the regression pairs for these two group people are different. Thus this kind of features can be used for judging the learning ability of the user. Fig. 6 shows the another feature which extracts the triple viewing sequences which indicate the readers view region A, region B, and then region C, continuously. Some of the most popular sequences for HA and LA people are plotted. The HA and LA people have different viewing sequences and reading behavior when they are studying the mathematic problems.

Based on the extracted features of the training samples, the system is automatically trained with SVM classifier. Currently, the proposed tool for reading behavior analysis enumerates 43 types of features, suggested by the articles of reading behavior study [2], [3]. The features contain the total/mean durations of fixation as well as saccade for each region, the regressions among fixation group, the continuous viewing sequences, and other combinations between these features. It can be validated that the HA and LA people have different features, although it is not possible to achieve the 100% correct with only the three features. With a set of experiments, the proposed system can achieve higher than 95% accuracy in identifying the ability of the readers based on their gaze trajectory data.

Reference

[1] A. Kar and P. Corcoran. "A review and analysis of eye-gaze estimation systems, algorithms and performance evaluation methods in consumer platforms," *IEEE Access*, vol. 5, pp. 16495-16519, 2017.

[2] R. Zheng, et al., "Eye-gaze tracking analysis of driver behavior while interacting with navigation systems in an urban area," *IEEE Trans. Human-Machine Systems*, vol. 46, no. 4, pp. 546 – 556, 2016.

[3] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 124, no. 3, pp. 372 - 422, 1998.

[4] M. Lai, et al., "A review of using eye-tracking technology in exploring learning from 2000 to 2012," *Educational research review*, vol. 10, pp. 90-115, 2013.

[5] W. C. Kao, M. C. Hsu, and Y. Y. Yang, "Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition," *Pattern Recognition*, vol. 43, Issue 5, pp. 1736 - 1747, May 2010.