Fixational Feature-Based Gaze Pattern Recognition using Long Short-Term Memory

Suparat Yeamkuan* and Kosin Chamnongthai†
* King Mongkut's University of Technology Thonburi, Bangkok, Thailand
E-mail: suparat.y@mail.kmutt.ac.th
† King Mongkut's University of Technology Thonburi, Bangkok, Thailand
E-mail: kosin.cha@kmutt.ac.th

***Abstract*—The pattern of eye gaze is increasingly powerful for human-computer interaction tasks. Understanding of gaze pattern can provide valuable information regarding to users' attention. Certainly, the patterns of eye gaze known as eye accessing cues are related to the cognitive processes of the human brain. In this paper we propose a method for gaze patterns recognition, where a gaze data was collected from eye tracker. Consequently, a gaze fixation feature and Long Short- Term Memory technique is employed in this work for the recognition. To evaluate the performance of the proposed method, we have an experiment with 7 examiners, in which they have to looked at 3 tasks of point, rotate and slide on screen. The experimental results show the proposed method offering favorable performance on a standard eye tracker, respectively.**

I. Introduction

The eye gaze reveals attentional focus and cognitive strategies [1]. The behavior of eye gaze is essential in developing a natural human-computer interaction (HCI) system. Generally, eye gaze has been widely used to extract the cognitive process along to attend. The knowledge of cognitive processes can intelligently assist computers to interact with humans. Therefore, gaze information is useful for achieving immersive from the user as a modality of input. The potential of eye gaze tracking and gaze-based HCI in modern consumer devices is currently an active topic for exploration, such as eye-tracker that has been widely adopted in psychology and human-computer interface research. Recently, many works have been studied [2], [3] investigating the use of gaze patterns in HCI applications. Especially, a visual fixation that provides rich information about the human cognitive process through eye gaze.

In this paper, we propose a gaze fixation feature-based structure with Long Short-Term Memory (LSTM) techniques to model recognition of human eye-gaze patterns tasks on the user interface.

The rest of this article is organized as follows. Related works and an analysis of the problem are explained in sections II and III. Consequently, the proposed method, experiments and results are introduced in sections IV and V. Finally, the paper is concluded in section VI.

II. Related Work

There has been excellent in applying gaze pattern information to study the cognitive process in the user interface which refers to psychology [4] [5].

Generally, Eye movement patterns are conducted patterns of perception. Therefore, there are many work-related to eye tracking with commercial eye-tracking devices [6]. Y.Zhang et al. [7] presented a computational model of human eye movements such as saliency map, short-time memory, saccades, Region of Interest (ROI), and retina mode which define by genetic algorithm. U. Engelke et al. [8] study the visual behaviors of observers in the subjective image, this work helps to understand of human visual behaviors in terms of individual characteristics. Recently, R.Bhattarai and M.Phothisonothai [9] proposed eye-tracking based visualizations and metrics to analyze Individual eye movement patterns, which particular use the Naive Bayes model formulate a statistical learning model. This work focuses on discrete eye movement patterns with show an excellent result that separates between saccade and fixation.

However, in terms of use eye movement in HCI. This related to the capabilities of Natural Language Processing (NLP) by cognitive-mode language processing [10] which brings the era of the user interface to become serval types of interaction behaviors such as sliding, rotation, and press button. This paper proposed a method of gaze pattern recognition using a sequence of time-series techniques as LSTM to deal with a flexible continuous sequence of saccade and fixation.

III. Problem Analysis

Generally, a saccade is the rapid eye movement between fixations to move the eye-gaze from one point to another point. Fixation is the ability to maintain the focus on a target [11]. When scanning to text or image, human eyes make saccadic movements and stop several times, and the speed of eye movement during each saccade cannot be controlled.

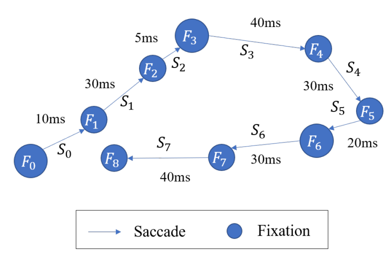

Fig. 1. An example of eye movement between saccade and fixation.

Fig. 1, shows an example of an eye movement trajectory of saccade and fixation. Let consider the sequence of saccade and fixation in term of the transitional method of Hidden Markov Models (HMM) [12], which its can address to an observation sequence by the proportion of saccade and fixation as shown in Fig. 2.

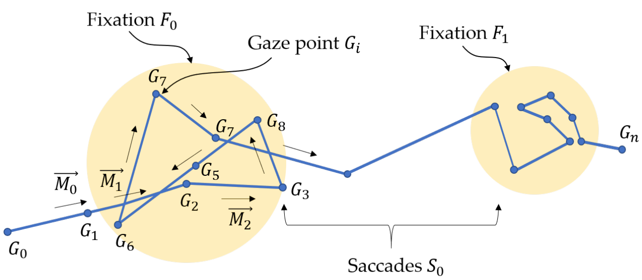

Fig. 2. An example of saccade and fixation sequence.

By mathematically, the state of variables changes over time. Fixation state $\{F\}_{i=0}^{n}$ is the model of the micro-saccades when eye gaze stationary objects. Saccade state $\{S\}_{i=0}^{n}$ is model of the rapid transit between fixation over n time. The HMM models observed fixation state in discrete time as define in (1), where $\{G\}_{i=0}^{n}$ is the state variable at time i. If we find the probability P of fixation state by Gaussian distribution where λ is HMM model parameters, the distribution of each fixation cannot be observed with a constant number of $P(G)$ and each saccade transition are deepened on user cognitive experience. This express to the uncontrol variable of the learning model that requires a flexible constant number of training data.

$$P(G_1, G_2, .., G_n|\lambda) = \sum_{i=1}^{n} P(G_1, G_2, .., G_n|\lambda) \qquad (1)$$

To improve the limitation of the traditional learning model, this has to consider the learning model that flexible on time-series sequence learning and recognition. This paper proposed LSTM [13] learning model which working on maps the sequence of past observations as fixation features input. With the ability of time sequence, this potentially solve an uncontrol fixation distribution and fixed range of data in learning along to the recognition process.

IV. Proposed Method

The proposed method consists of three main parts, as shown by the flowchart in Fig. 3, the process starts to acquire the input data from the eye tracker, following the eye-gaze data is extracted fixation to obtain feature, and finally LSTM learning model is an essential part of pattern recognition.

A. Eye Gaze Input

According to the hardware to use as an eye tracker, we use a Tobii 4C Eye Tracker for detecting the eye gaze on the screen. To obtain gaze information, we use the interaction library from Tobii [14]. The gaze data is obtained in a 2D coordinate which referring to a pixel position where the user looks on the screen. This can express the model of gaze data to $\{G\}_{i=0}^{n}$ as shown in Fig. 4.

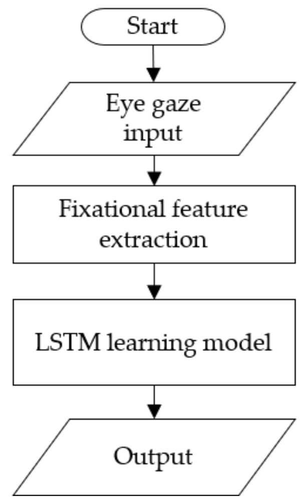

Fig. 3. Flowchart of proposed method.

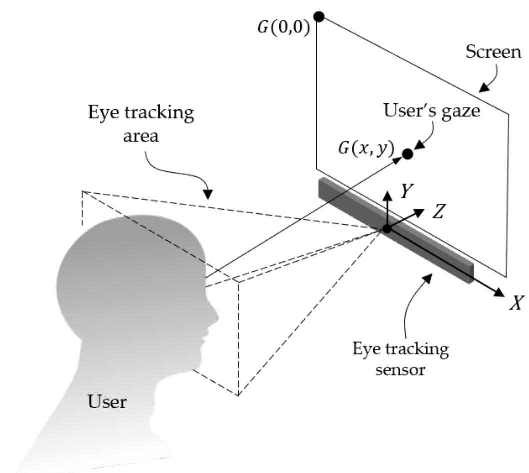

Fig. 4. System overview of eye gaze acquisition as an input.

B. Fixation Feature

By psychologically, fixations are generally referred to as saccades. Normally, when we look at an object our eyes do not travel smoothly (shown in Fig. 2) whereas the process of visual cognitive starts with a scanning path to the target. After a while, fixation spends a period of time by multipoint of gaze $\{G\}_{i=0}^{n}$ where the eye is continuously aligned direction $\{\overrightarrow{M}\}_{i=0}^{n}$ on the target, then our perception is guided by sequences of fixations and saccades over a sequence of time. This can be performing the fixation feature as follow:

$$\overrightarrow{M}_i = G(x, y)_i - G(x, y)_{i-1} \qquad (2)$$

C. LSTM Learning Model

The algorithm of recognition has two processes as show in Fig. 5: First is a training process, learning algorithm using a LSTM comprised of the fixation feature and the corresponding labels. Second is the recognition process, this input gaze data and apply the feature with the training process and feed the features to the model of learning algorithm to predict the label.

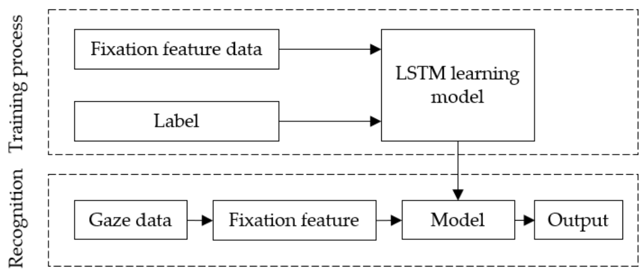

Fig. 5. LSTM learning model.

The LSTM learning model for the gaze pattern recognition, we define the model as sequential to support an input data in time series, when there is a short or long time of unknown size concerned with consecutive and current saccades. This is followed by a bidirectional LSTM and dropout layers of the model to the training data. Finally, a dense activation uses Rectified Linear Unit (ReLU) as a fully connected layer. a dense activation layer by Softmax is the final output layer used to predict the result.

Algorithm 1 Algorithm for LSTM learning model

Input: training data $\{G\}_{i=0}^{n}$ and $\{\overrightarrow{M}\}_{i=0}^{n}$

model $\leftarrow$ sequential
model $\leftarrow$ bidirectional LSTM layers($\{G\}_{i=0}^{n}$, $\{\overrightarrow{M}\}_{i=0}^{n}$)
model $\leftarrow$ dropout layers
model $\leftarrow$ dense activation layers(ReLU)
model $\leftarrow$ dense activation layers(Softmax)

V. Experiment and Results

To evaluate the proposed recognition of eye gaze pattern using LSTM, experiments with 7 examiners, 5 males, and 2 females, An eye gaze data was collecting in the experiment used eye tracker [15] to detect a gaze position. The computational of system, we use data stream as frame rates around 50 Hz which Python with Keras [16] Libary for LSTM programming in a laptop with CPU 3.4 GHz Core i3, 4GB of RAM on Windows 64 bit as the operating system. The experiment is conducted an eye gaze interaction tasks that contain 3 classes with different patterns as shown in Fig. 6.

In the experiment, each examiner is positioned in front of the eye tracker with a proper distance from the tracking area. Each examiner is allowed to move his/her head freely during the experiment. The experiment has three interaction tasks for each examiner. The software recorded each interaction task separately as shown in Fig. 7.

(a) Point (b) Rotate. (c) Slide.

Fig. 6. Experiment tasks.

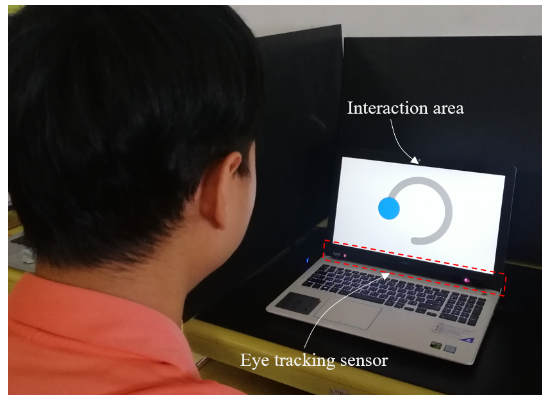

Fig. 7. Examiner during perform the experiment.

Fig. 8, shows examples of gaze patterns corresponding with the trial tasks. The behavior of the gaze position is scattered over the image of the experiment tasks. During perform the interaction of the trial task, we found the fixation shown a high density of gaze points in the "point" task by micro-saccade and a saccade path for each fixation in "rotate" and "slide" tasks.

(a) Point (b) Rotate. (c) Slide.

Fig. 8. Examples of eye gaze patterns.

For training the LSTM model, we used recorded gaze data which separate labels of tasks "point", "rotate", and "slice" where the training samples are of different sizes for each task. This depends on each examiner's experience performing each task, we found the number of training indicate the "point" task is a lower amount of number when compared with "rotate" and "slide" as shown in Fig. 9.

Experimental gaze data from each examiner was used to evaluate the performance for recognition on each interaction task and the results were measurement by the accuracy recognition. The performed experiments are fully reported in Table I. The performance overall examiners were 82.09% of the "point" task, it's lower when compared with Native Bays [9] at the same task where the result is slightly different around 8.13% in terms of accuracy. This cause by gaze information in the ”point” task is likely formal Gaussian distribution more than a sequence of time. Furthermore, the ”rotate” and ”slide” tasks reported an average accuracy of 84.86% and 85.19% respectively.

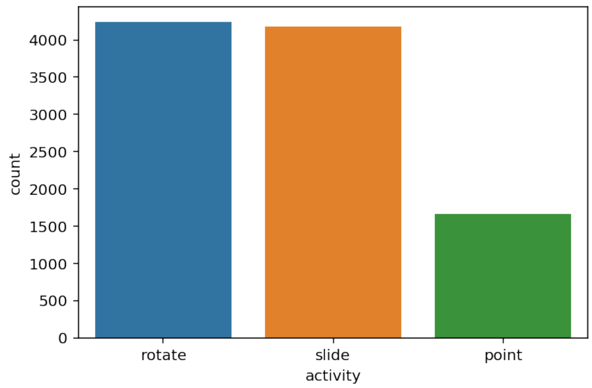

Fig. 9. Different size of training data.

TABLE I
ACCURACY OF THE EXPERIMENT

Technique	Task		
	Point	Rotate	Slide
Native Bays [9]	90.22%	-	-
Proposed method	82.09%	84.86%	85.19%

VI. CONCLUSIONS

In this paper, an eye gaze pattern recognition using gaze fixation features is presented. The method of LSTM applying with the fixation feature giving successfully recognizes three tasks regarding human interaction by eye gaze. Based on our proposed method with the fixation feature of eye gaze information that deals with the sequence of time series, the LSTM model is shown potentially use with time-sequence recognition and enables prediction of the task through eye gaze pattern. However, order to use LSTM on the short moving distance of eye gaze requires a filter process to reduce noise from previously occurred somewhere in sequence. The performance of the proposed method shows compromises the result of eye gaze pattern recognition for human-computer interaction tasks. The average accuracy of recognition of overall experiment tasks by 84.05% appropriately.

ACKNOWLEDGMENT

The research project is financially supported by Petchra Pra Jom Klao Ph.D. Research Scholarship (Grant No. 24/2561) from King Mongkut’s University of Technology Thonburi in Bangkok, Thailand.

REFERENCES

[1] Maria K. Eckstein et al, “Beyond eye gaze: What else can eyetracking reveal about cognition and cognitive development,” *Developmental Cognitive Neuroscience,* vol. 25, 2017, pp. 69-91.
[2] A. Poole and L. Ball, “Eye tracking in human-computer interaction and usability research,” *Encyclopedia of Human Computer Interaction, Publisher: Idea Group Reference,* 2006, pp. 211-219.
[3] X. Zhang et al, “Eye Tracking Based Control System for Natural Human-Computer Interaction’,’ *Computational Intelligence and Neuroscience, Hindawi,* 2017, pp. 1-9.
[4] S. Durrani and Q. Durrani, “Applying Cognitive Psychology to User Interfaces,” *International Conference on Intelligent Human Computer Interaction. Springer,* 2006, pp. 156-168.
[5] T.Stefano et al, “Developing Emotional Design: Emotions as Cognitive Processes and their Role in the Design of Interactive Technologies,” *Frontiers in Psychology,* 2017.
[6] C. T. Benjamin and L. G. Steven, “Best practices in eye tracking research,” *International Journal of Psychophysiology,* vol. 155, 2020, pp. 49-62.
[7] Y. Zhang et al, ”Eye movement data modeling using a genetic algorithm,” *2009 IEEE Congress on Evolutionary Computation,* Trondheim, 2009, pp. 1038-1044.
[8] U. Engelke, A. Maeder and H. Zepernick, ”Visual attention modelling for subjective image quality databases,” *2009 IEEE International Workshop on Multimedia Signal Processing,* Rio De Janeiro, 2009, pp. 1-6.
[9] R. Bhattarai and M. Phothisonothai, ”Eye-Tracking Based Visualizations and Metrics Analysis for Individual Eye Movement Patterns,” *2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE),* Chonburi, Thailand, 2019, pp. 381-384.
[10] A. Mishra and P.Bhattacharyya, “Cognitively Inspired Natural Language Processing:An Investigation Based on Eye-tracking,” *Cognitive Intelligence and Robotics,* Springer Singapore, 2018.
[11] D. Purves, G. J. Augustine, D. Fitzpatrick D, et al, “Neuroscience 2nd edition. Sunderland (MA): Sinauer Associates,” *Types of Eye Movements and Their Functions,* 2001.
[12] J. Kim et al, “A hidden Markov model for analyzing eye-tracking of moving objects,” *Behav Res*, vol. 52, 2020, pp. 1225–1243.
[13] A. Pulver and S. Lyu, ”LSTM with Working Memory,” *2017 International Joint Conference on Neural Networks (IJCNN),* Anchorage, AK, 2017, pp. 845-851.
[14] Tobii Interaction Library, https://developer.tobii.com/consumer-eye-trackers-interaction-library-api-reference/
[15] Tobii 4C, https://gaming.tobii.com/
[16] Keras Library, https://keras.io/api/layers/recurrent_layers/lstm/