Differences of Recognizing the Scenes and Histograms: An Eye Tracking Study

Xue Tan[1, 4], Shengfu Lu[1, 4], Mi Li[1, 2, 4*], Zhao Liu [1, 4], Tong Shang [1, 4], and Ning Zhong[1, 3, 4]

1 The International WIC Institute
Beijing University of Technology
Beijing, China
2 The School of Computer and Communication Engineering
Liaoning ShiHua University
Liaoning, China
3 The Department of Life Science and Informatics
Maebashi Institute o/Technology
Maebashi-City, Japan
4 Beijing Key Laboratory of MRI and Brain Informatics, China
* Corresponding author: Mi Li, E-mail: limi135@gmail.com

Abstract—**It is a very interesting question whether there are cognitive differences when subjects recognizing scenes and histograms. To address this issue, the study using eye tracking investigated the saccade distance, fixation duration, fixation count, mean fixation duration and pupil diameter when people recognizing the two type tasks. As is shown in the results, there exits cognitive differences in the information processing, eye movement patterns and mental overload.**

Keywords-scene; histogram; information processing; eye tracking

I.   INTRODUCTION

Eye tracking has been applied to the interface evaluation, web design and other research fields gradually. For example, there is a study, which focus on the impact on the space and location information to the interface layout design [1], and also there is a study ,which is try to understand the impact on the page layout to the visual search [2]. In addition, there are also a lot of researches on the search engines [3]. In recent years, we used eye tracking to carry out a series of studies in the page layout, page type, web search and browse and web information overload, et al. Such as the exploration of the page layout [4],the exploration of the characteristics of visual search on web pages [5,6], the exploration of whether the floating ads affect users' visual search behavior on web page on cognitive science [7], and the exploration of visual search and browsing strategies on web pages [8], as well as the exploration of strategy and processing mode of visual search under Information overload on web pages [9]. Picture is an important element of the web page, so it's important to recognize the picture, however, few studies focus on it. Therefore, this study used a real-time recording of eye movement in the eye movements when viewing the scene and histogram. By analyzing the visual behavioral indicators to explore whether there are different eye movement patterns when people were recognizing the scene and histogram, and then reveals the mental process and the law [10] when people recognized different pictures further. This study can provide a theoretical basis for the intelligent man-machine interface,

and thereby contributing to the development of artificial intelligence.

II.   EXPERIMENTS

A.   Participants

The participants were 30 undergraduates and postgraduates with the age range of 22 ~ 28 years old ($M = 25.0$, $SD = 1.4$), and 15 were female. All participants had normal or corrected-to-normal vision.

B.   Apparatus

Eye movements were recorded at the rate of 120 HZ by Tobii T120 eye-tracker, which had a 17 LCD monitor with resolution set to 1024×768 pixels and at the refresh rate of 60 HZ.

C.   Experiment Material

Two kinds of materials which were scene and histogram were black-and-white picture (See Figure 1). Scene were divided into two kinds, one was outdoor scene and the other one was indoor scene. Outdoor scene were taken from the Psychological Image Collection at Stirling (PICS; http://pics.psych.stir.ac.uk/), and indoor scene were taken form the Indoor Scene Recognition (http://web.mit.edu/torralba/www/indoor.html).The histogram were made by us according to some criteria. There were two types of charts: two-dimensional histogram and three-dimensional histogram, each type had 20 pictures. Histogram column was divided into three, four and six cases, the volume label was alphabetic and numeric characters, font type was Times New Roman, font size was 12. The gray scale, image size and resolution of all the scene and histogram were the same.

D.   Procedure

The experimental process consisted of two stages. Subjects should be familiar with the procedure in the first stage in order to understand the purpose of the experiment, and they needed to have a practice, in which four scene and histogram would be presented. The second stage was the

formal experiment, in which had four sessions to reduce the visual fatigue. There were twenty pictures in each session, half of which were scene. All the pictures were presented randomly in 2 seconds; the subjects would have a rest in two minutes and then continued to the next session. Participants needed to watch carefully in the experiment.
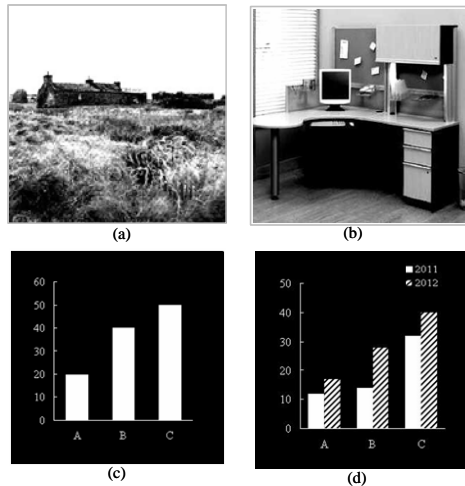

Figure 1. Examples of stimulis in the experiment: (a) Outdoor scene. (b) Indoor scene. (c) Two-dimensional histogram. (d) Three-dimensional histogram.

III. RESULTS

As shown in Figure 2, the saccade distance was significantly longer when subjects recognizing the histogram than scene, so did the pupil diameter. However, the other index, such as, fixation duration, pupil diameter and mean fixation duration was not significant.

The heat maps, which were superimposed by many fixation points, directly reflected which part was the most interesting part in the picture and the darkest color was the most interesting part. As was shown in Figure3, subjects were more concerned about central area in indoor scene, so did the outdoor scene. However, subjects were more concerned about the two-dimensional histogram's axis and column charts and were more concerned about the three-dimensional histogram's axis, column charts and the tag items.

We made an assumption that the reason was the differentiation in the amount and the layout of the information according to the result. We divided the area of interest in a 1:1 ratio for further analysis of the degree of difference; we called the central area as AOI1 and the peripheral area as AOI2. Figure4 is the results of the eye movement index in AOI1 and Figure5 is the results of the eye movement index in AOI2.

For fixation count, on the one hand, the fixation count when subjects recognizing scene (4.67) was bigger than histogram (3.86) in AOI1 and there was a significant difference between them [F (1, 58) = 13.79, P < 0.001], as shown in Figure 4(a). However, the fixation count when subjects recognizing scene (1.68) was smaller than histogram

(2.27) in AOI2 and there was a significant difference between them [F (1, 58) = 17.06, P < 0.001], as shown in Figure 5(a). The result showed that people were more concerned about the central area in scene and more concerned about the peripheral in histogram. On the other hand, the fixation count in AOI1 (4.67) was bigger than in AOI2 (1.68) for scene [F (1, 58) = 278.15, P < 0.001], and the fixation count in AOI1 (3.86) was bigger than in AOI2 (2.2) for histogram [F (1, 58) = 13.79, P < 0.001].
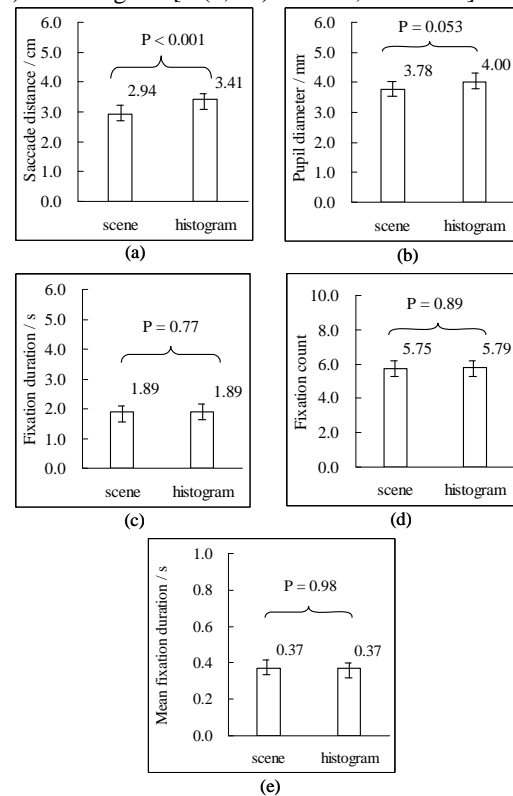

Figure 2. The results of the eye movement data on scene and histogram:(a) Saccade distance.(b) Pupil diameter.(c) Fixation duration.(d) Fixation count.(e) Mean fixation duration.

For fixation duration, on the one hand, the fixation duration when subjects recognizing scene (1.33s) was bigger than histogram (0.92s) in AOI1 and there was a significant difference between them [F (1, 58) = 30.13, P < 0.001], as shown in Figure 4(b). However, the fixation duration, when recognizing scenes (0.27s) was smaller than histograms (0.47s) in AOI2 and there was a significant difference between them [F (1, 58) = 35.15, P < 0.001], as shown in Figure 5(b). The result showed that people dealt with more information in scene' central area and dealt with more information in histogram's peripheral area. On the other hand, the fixation duration in AOI1 (1.33s) was longer than in AOI2 (0.27s) for scene [F (1, 58) = 388.01, P < 0.001], and the fixation duration in AOI1 (0.92s) was longer than in AOI2 (0.47s) for histogram [F (1, 58) = 52.31, P < 0.001].

For mean fixation duration, there was not a significant difference between scene and histogram [F (1, 58) = 0.16, P = 0.68], as shown in Figure 4(c). However, the mean fixation

duration when subjects recognizing scene (0.18s) was smaller than histogram (0.28s) in AOI2 and there was a significant difference between them [F (1, 58) = 24.02, P < 0.001], as shown in Figure 5(c). The result showed that histogram was more complex than scene, which required a longer time to identify. On the other hand, the mean fixation duration in AOI1 (0.46s) was bigger than in AOI2 (0.18s) for scene [F (1, 58) = 9.01, P < 0.001], and the mean fixation duration in AOI1 (0.41s) was bigger than in AOI2 (0.28s) for histogram [F (1, 58) = 1.98, P < 0.001].
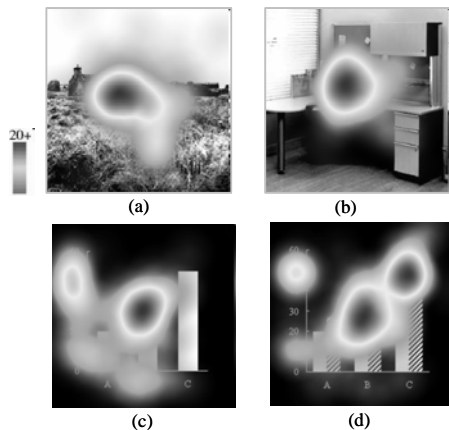

Figure 3. Examples of the heat maps when the subjects recognizing the two type tasks: (a) Outdoor scene. (b) Indoor scene. (c) Two-dimensional histogram. (d) Three-dimensional histogram.
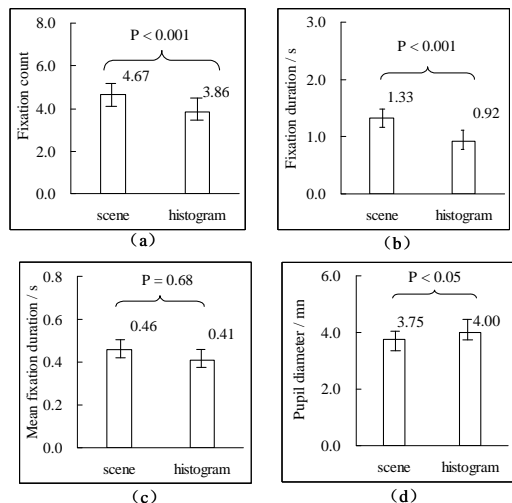

Figure 4. The results of the eye movement index in AOI1:(a) Fixation count.(b) Fixation duration.(c) Mean fixation duration.(d) Pupil diameter.

For pupil diameter, on the one hand, the pupil diameter when subjects recognizing scene (3.75mm) was smaller than histogram (4.00mm) in AOI1 and there was a significant difference between them [F (1, 58) = 5.43, P = 0.02 < 0.05], as shown in Figure 4(d). However, there was not a significant difference between scene and histogram in AOI2 [F (1, 58) = 3.71, P = 0.058 > 0.05], as shown in Figure 5(d). The result showed that there exited mental overload when

the subjects recognizing the scene and histogram. On the other hand, there was not a significant difference between AOI1 and AOI2 for scene [F (1, 58) = 0.1, P = 0.75 > 0.05], and there was also not a significant difference between AOI1 and AOI2 for histogram [F (1, 58) = 0.02, P = 0.88 > 0.05].

All these eye movement indicators confirmed the previous assumptions, which was the two types of the pictures' amount and layout of the information affected the eye movement patterns when recognizing the scene and histogram. The saccade distance and the mean fixation duration could explain the amount of the information, and fixation duration and fixation count could explain the structure of the information.

IV. DISCUSSION

There are many reasons can affect the eye movement patterns, such as, the size of the amount of the information, the layout of the information and mental overload. Our sights always focus on the main features of the pictures. The eye scan line always turns from one characteristic to another feature of the picture. The identification of the complex picture can be achieved through different levels of information processing. For familiar picture, we will treat it as a unit to identify so that we will not pay more attention to the details as we already know its main features.

In the experience, scene contained simple information so that it could be recognized easily while the histogram contained complex information so that it needed more information processing and more information integrating. This was reflected in saccade distance and mean fixation information. Saccade distance was the distance between the two gaze points [11]. The saccade occurred after processing was completed and changed the gaze point to let the new content falls in the central recess of the visual area in order to process a new content. The size of the saccade distance could reflect the size of the amount of the information obtained [12]. From the statistical data, the scene' saccade distance was significantly less than the histogram's, which meant that the information obtained from the scene was less than the information obtained from the histogram. Mean fixation duration, the speed of the watching, which meant the difference between the amount of the information at a time. The scene's mean fixation duration was longer than histogram's in the central area, while in the peripheral area the histogram's mean fixation duration was longer than scene's, there are significant differences.

Judging from the distribution of the information contained in the scene and histogram, scene is mainly distributed in the central area, while for the same size of the picture, the histogram has a different structure, and the information is distributed dispersed. This was reflected in fixation count and fixation duration. Fixation duration meant the time each gaze point lasted [13, 14], it reflected the degree of processing of the material. The scene's fixation duration was longer than histogram's in the central area, while in the peripheral area the histogram's fixation duration was longer than scene's fixation duration, which meant that the subjects were more interested in the scene's central area and more interested in the histogram's peripheral area.

Mental load is the mental activity of the individual to bear on the unit time. Many studies showed that pupil diameter was a sensitive pointer [15] which could measure the resource allocation and processing load in the cognitive processing activities. The enlarged pupil meant the greater processing load or mental effort [16] in cognitive activities. From the overall analysis, we could see that scene and histogram's difference were nearly significant, and in the central area there was a significant difference between scene and histogram, while in the periphery there was not a significant difference between them. This was because when we were looking at the coordinate, which was related to the extraction of the value of the histogram, and when we were looking at the central region of the histogram, which was involved the information integration. All those would get more mental load than directly obtained from the outside world. Scene was simple while the histogram was so complex that individuals bore more mental activity workload at a unit time when recognizing.
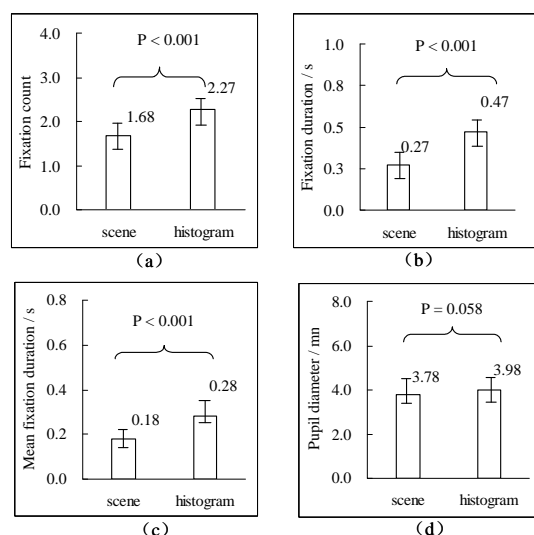

Figure 5. The results of the eye movement index in AOI 2:(a)Fixation count.(b) Fixation duration.(c) Mean fixation duration.(d) Pupil diameter.

V. CONCLUSION

Based on the analysis of the data of eye movement to recognize two types of pictures, we can make a conclusion that there exits cognitive differences in the information processing, eye movement patterns and mental overload when subject recognize scenes and histogram.

ACKNOWLEDGMENT

REFERENCES

[1] M. Terenzi, F. D. Nocera, and F. Ferlazzo, "Action, not only semantics, underlies expected location for interface elements," The 4th Italian Symposium on Human Computer Interaction, Roma, Italy, Springer, 2005, pp. 23-27.

[2] J. F. Shi, X. H. Cao, G. Wang, and B. Qu, "An eye movement study on visual search of wen page layout," Chinese journal of ergonomics, vol. 4, 2008, pp. 1-3.

[3] E. Cutrell,Z. GUAN, "What are you looking for? An eye tracking study of information usage in Web search," CHI 2007, San Jose, CA, USA, ACM Press, 2007, pp. 407-416.

[4] M. Li, Y. Y. Song, S. F. Lu, and N. Zhong, "The layout of Web pages: A study on the relation between information forms and locations using eye-tracking," LNCS, 2009, pp. 207-216.

[5] Y. Y. Song, S. F. Lu, M. Li, and N. Zhong, "The impact of the layout of web page for visual search using-tracking,"Journal of Frontiers of Computer Science and Technology, 2009, vol. 36, pp. 198-201.

[6] M. Li, J. J. Yin, S. F. Lu, and N. Zhong, "The effect of information types and floating ads for visual search on Web: an eye-tracking study," LNAI, 2009, 5819, pp. 96-105.

[7] M. Li , N. Zhong, and S. F. Lu , "A study about the characteristics of visual search on web pages," Journal of Frontiers of Computer Science and Technology, 2009, vol. 3, pp. 649-655.

[8] M. Li , N. Zhong, and S. F. Lu , "Exploring visual search and browsing strategies on web pages using the eye-tracking," Journal of Beijing University of Technology, 2011, vol. 37, pp. 773-779.

[9] M. Li, W. X. Lu, S. F. Lu ,Y. Y. Song, J. J. Yin,and N.Zhong, "Strategy and processing mode of visual search under information overload on web pages: an eye-tracking study," Journal of Beijing University of Technology, 2012, vol. 38, pp. 390-395.

[10] K. Rayner, "Eye movements in reading and information processing: 20 years of research," Psychological Bulletin, 1998, vol. 124, pp. 372-422.

[11] J. R. Anderson, "Acquisition of cognitive skill," Psychological Review, 1982, vol. 89, pp. 386-406.

[12] H. Wang, J. P. Xiong, "Characteristics of eye movement of male college students with difference sports performances when they were watching volleyball smash videos,"Journal of Physical Education, 2010, vol. 17, pp. 7-81.

[13] K. Rayner, C. M. Rotello, and A.J. Steward, "Integrating text & pictorial information: eye movements when looking at print advertisements," Journal of Experimental Psychology, 2001, vol. 7, pp. 219-226.

[14] K. Rayner, B. Miller, and M. R. Caren, "Eye movements when looking at print advertisements: the goal of the Viewer Matters," Cognitive Psychology, 2008, vol. 22, pp. 697-707.

[15] J. Beatty, "Task-evoked papillary responses, processing load, and the structure of processing resources," Psychological Bulletin,1982, vol. 91, pp. 276-292.

[16] S. P. Verney, E. Granholm, and S. P. Marshall, "Pupillary responses on the visual backward. Masking task reflect general cognitive ability," International Journal of Psychopysiology, 2004, vol. 52, pp.23-26.