Through a Web Camera Base of Eye Tracking Technology to Explore the Audience's Attention Preferences in terms of the Positions of Information and the Layout Compositions

Hui-Hui Chen, Yi-Ting Yeh, Chiao-Wen Kao, Bor-Jiunn Hwang, and Chin-Pan Huang

***Abstract*—Recording and analyzing eye movements can effectively obtain human's inner complex cognitive process of visual information and retrieve the foci of the processed information during reading. This research proposes to adopt a web camera base form for its ease to obtain, its unawareness of existence, and the unnecessary for adjustment of the participant's poses. This study meanwhile will explore the audience's attention preferences (fixation position & fixation duration), gaze areas, gaze duration, and navigation transitions regarding to the layout compositions of photos and texts and their positions to the web pages. It is to understand the attention preferences of the audience in terms of gaze areas, gaze duration, and navigation transitions with the help of the web camera base of eye tracking technology. The cross examination of the relationships between the positions of information and the layout compositions is conducted.**

***Index Terms*—Eye Gaze, Gaze Area, Navigation Transition, Attention Preferences**

I. INTRODUCTION & RELATED WORK

Along with the mature development of the digitalized information, there are plenty of learning materials and information offered via web pages. The integration of multimedia profoundly makes the presentation of information on the web in numerous ways. No matter which kind of the integration takes place on the web, it only matters to the users whether the information could be delivered effectively to them [1]. Several research studies indicated that the users' eye gazes and navigation transitions were affected not only by the users' prior reading patterns and experiences but also affected by the arrangements of the page layouts[2], [3]. Hence, the arrangements of the page layouts influence the users' choices and distribution of their attentions caught by the layouts. It was found that the users read information by scanning. Only when they found interested regions, they would just further read closely [3]. They also spent more time gazing on the information with graphics and on the preferred regions of the layouts [4].

The study [5] indicated that during human's cognitive procedure on processing the information, there was more than 80 percents of the information acquired by eyes, which were the most important sources of the sensory memory. The studies on eye tracking to explore the relationship between the eye movements and the modalities to present information were published. Although the data collected from the eye movements were important, in the past the eye movements could only be observed through the eyes of the observers, or through the interviews and the think-aloud methods [6]. However, the past way not only might be easily misconducted by the false memory, but it also might not probably reflect a person's true inner cognition. Since a German psychologist, J. Cohn, published his investigation report on the roles of the favorite colors in the 19th century, most of the researchers inherited from him using questionnaires to investigate research variables about how and what eyes navigate [7].

However, the strategy of using questionnaires might cause the similar research studies with the opposite research results due to the different research methodologies, stimuli, or equipments. By using questionnaires in investigation, the results might also be influenced by the subjective impressions of people [8], [9]. In the 20th century, researchers started to record eye movements from various aspects [6]. Recording a person's eye movement activities was recognized as a more objective and effective way to conduct research studies in how and what information eyes process [10].

Eye movements construct a series of visual tracks which reveal the paths that the eyesight of the audience navigates [2]. These visual tracks in turns may become the cues to understand the navigation patterns of the audience while reading. Recording and analyzing eye movements can effectively obtain human's inner complex cognitive process of visual information and can effective retrieve the foci of the processed information during reading[11], [12]. The final data can serve as the indicator of the outer behavior of the audience. Furthermore, through studying eye movements, that

Manuscript received December 30, 2011; revised January 17, 2012.

Hui-Hui Chen is with the Department of Computer and Communication Engineering, Ming Chuan University, Taoyuan, Taiwan. (phone: +806-3-350-7001#3749; fax: +806-3-359-3876; e-mail: huichen@ mail.mcu.edu.tw).

Yi-Ting Yeh is with the Department of Computer and Communication Engineering, Ming Chuan University, Taoyuan, Taiwan. (e-mail: 00166021@ms1.mcu.edu.tw).

Chiao-Wen Kao is with the Department of Computer Science and Information Engineering, National Central University, Taoyuan, Taiwan. (e-mail: chiaowenk@gmail.com).

Bor-Jiunn Hwang is with the Department of Computer and Communication Engineering, Ming Chuan University, Taoyuan, Taiwan. (e-mail: bjhwang@ mail.mcu.edu.tw).

Chin-Pan Huang is with the Department of Computer and Communication Engineering, Ming Chuan University, Taoyuan, Taiwan. (e-mail: hcptw@mail.mcu.edu.tw).

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

the interested gaze areas and to where the audience's attention pay may be acknowledged [5], [12], [13].

The eye tracking device, which early was used in studying reading in psychology, consists of two forms, the head mounted and the table mounted. The head mounted form is suitable for studying the motion related and the reality research of the 3D[6], but it is too heavy to be used in studying reading. Its weight may easily cause the neck of the participant uncomfortable [14] and cause the pressure and tiredness to the participant. Further, it will affect the participant's mood and make the data collected from recording eye movements inaccurate [15].

Therefore, the table mounted form is promoted generally for its ease to use and for that it affects the participant less while reading, although it can only record 2D tracks of the eye movement changes[6]. The disadvantage of the table mounted form is that it could only placed on the table or the level area. While conducting research studies on reading with either the head mounted or the table mounted, it is necessary to adjust the visual angels or to limit the movement area of the participant. The participant cannot act at the participant's great convenience nor act as naturally as possible. Besides, both forms are addition to the purchase, not general enough seen. Therefore, to record eye movements, this research proposes to adopt the camera base form for its ease to obtain, its unawareness of existence, and the unnecessary for adjustment of the participant's poses. The camera base form will be used to make the experiment closer to real condition to the participant and to construct a reliable and buildable experiment setting.

II. Research Framework

This study will explore the audience's attention preferences (fixation position & fixation duration), gaze areas, gaze duration, and navigation transitions regarding to the layout compositions of photos and texts and their positions to the web pages (as shown in Fig. 1).

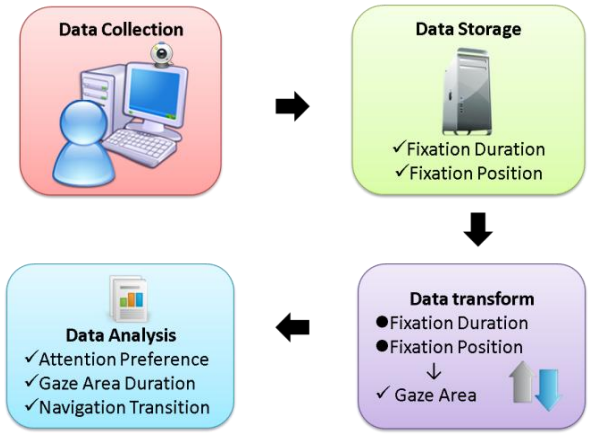

Fig. 1. Research Framework

1. Record the navigation transitions.

The web camera will be initiated to record the eye movements while the participant starts to navigate the screen pages. Five shots will be taken per second since, based on the previous research studies, the average gaze time of reading the content in Chinese takes about 220-230 millisecond.

2. Record the fixation position & fixation duration for the audience's attention preferences.
3. Transform data (fixation position & fixation duration) into gaze areas.
4. Analyze the data.

Explore the relationships between the layout compositions of photos and texts and their positions to the web pages in terms of the attention of preferences, the gaze area durations, and the navigation transitions.

III. Research Questions and Proposed Method

Regarding to the layout compositions of photos and texts and their positions to the web pages, this study comes up with three research questions, stated as followed:

A. Research Question 1: What are the audience's attention preferences in terms of fixation position & fixation duration?

Generally that in favor of eye tracking technology relies on eye movements can reflect the inner attention transformation process of the audience [11], [16]. It means that supervising the eye movements is corresponding to supervising when the cognitive process and attention initiate [17]. Despite of some research studies indicate that eye tracking records the fixation position, in fact human's visual attention could obtain information in a circumference area within 5-degree visual angel. Upon the common distance from the screen, this area is about 15 characters in width and 7 characters in height [18]. Some other studies indicate that the range of skimming information is about 4-degree visual angel while viewing photos and 2-degree visual angel while reading texts [6]. Hence, this study is to transform the fixation position and fixation duration into the gaze area and duration, and then to scratch the navigation transitions based on the gaze areas.

B. Research Question 2: Whether the audience's attention is attracted to the possessed favorite type of the layout compositions of photos and texts?

C. Research Question 3: Whether the audience's attention is attracted to the positions of information to the web pages?

Psychologically proven, within a short period of time, one's favorite tendency will drive one's eye movements. One will gaze more time on one's more favorite colors and images, and one will repeat gazing on those [17]. In addition, to properly control variables which might affect the research results, this study surveys the most commonly types seen in web portal news of the layout compositions of photos and texts. Three layout compositions of photos and texts are presented in this study: the photo on the left with the texts on the right, the photo on the right with the texts on the left, and the photo on the top with the texts on the bottom. Moreover, to reduce personal subjective causes, news and photos, such as about idols, political parties, and etc., which might arouse the audience great interests, are eliminated. Taken place are news and photos about natural issues, general knowledge, and common leisure are chosen to be integrated into the

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

designated content of the experiment layouts. Besides, the positions of the three layout compositions of photos and texts are equally shifted and arranged. Positions and layout compositions are crossly examined to answer the research questions.

Hence, this study will extend to make the null hypotheses:

Null Hypothesis $H_0$1: There is no difference of attractions between the layout compositions of photos and texts in terms of the gaze area, gaze duration and navigation transition.

Null Hypothesis $H_0$2: There is no difference between the positions of information in terms of the gaze area, gaze duration and navigation transition.

This study adapted PVS (Pattern Voting Scheme) from a recent research study [21] to record eye tracking activities and collect data for PVS requires no special user action for looking at reference points. PVS is less complex in camera calibration process and cost less. PVS finds one's fixation on the screen by determining the spatial coordinates of the eyes by distinguishing between iris and sclera with pixel gray value of pixel despite of eye colors.

With PVS, first, the eye detection is carried out by using the method based on haar-like feature [22]. After finding the eye image, take the horizontal $\overline{AB}$ from the half height (as shown in Fig. 2) and in order to make sure $\overline{AB}$ going through the iris position, $\overline{AB}$ position needs to be calibrated (as shown in Fig. 3).

The screen then is divided into N*M areas, where N along the horizontal dimension and M along the vertical dimension. Let $\overline{AB}$ be divided into K*N equal segments, where K is an odd number that can be used to set the number of polling district. And then the mean value (MV_i) of gray value in each part can be obtained, where i from 1 to KN. Compute MV_i to obtain the voting weight to indicate the iris location and estimate the visual translation direction [22].

(a) (b)

Fig. 2. Shows (a) left eye (b) right eye take the horizontal $\overline{AB}$

Fig. 3. Calibrate $\overline{AB}$ to red line

IV. Experiment and Analysis

A. Participants

15 college students (7males, 8 females), age from 20 to 25, all are able to navigate the contents of web pages and basic reading ability.

B. Experiment Materials

1. A designed page consists of three layout compositions of photos and texts (as shown in TABLE I):
   i. Composition A: the photo on the left with the texts on the right
   ii. Composition B: the photo on the right with the texts on the left
   iii. Composition C: the larger photo on the top with the texts on the bottom

TABLE I
3 Types of Layout Compositions

Composition A	Composition B	Composition C
 Photo Text	 Text Photo	 Photo Text

2. The webpage is divided into 3 blocks: Left (Block 1), Middle (Block 2), Right (Block 3). The positions of the 3 layout compositions of photos and texts are equally shifted and arranged respectively in the blocks. Therefore, there are 6 combinations (C1~C6) for the cross examination of the relationships between the positions of information and the layout compositions (as show in TABLE II).

TABLE II
Combinations for the Positions and Layout Compositions

	Block 1 (Left)	Block 2 (Middle)	Block 3 (Right)
C1	A	B	C
C2	B	C	A
C3	C	A	B
C4	A	C	B
C5	B	A	C
C6	C	B	A

3. The contents of news and photos were retrieved from several web portals during November 25th to December 5th, 2011. News and photos which might arouse the audience subjective interests are eliminated to keep the experiment objective. In consideration of the major spoken language of the participants, the experimental page contents are in Chinese.

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

C. Equipment

One web camera, one personal computer, one LCD monitor.

D. Experiment Procedure

1. Explain to the participants how to start and what to do.
2. When the participants finish navigating the presented information, they can click on the texts they gaze upon. Then, the next page will be shown with a waiting time of 2 seconds of blank page in black background to eliminate the effects caused by the previous residual image. That clicking on the texts is imitated what generally the audience will do when navigating the news on the web.
3. Until they go through all of the six pages, the experiment ends.

E. Data Analysis and Experiment Results

The web camera took five shoots per second. Data were coded with the navigation time stamp. Total of valid 4616 shoots were collected.

1. Comparisons for the respective positions of 3 layout compositions (A, B, C) on the webpage (as shown in Fig. 4):

   Data showed that while A on the Block 1, eye gazes of the participants were taken the most times with 848 shoots. The gaze area of the participants mostly remained in Block 1, which is the left block on the webpage. There were no significant difference between A on the Block 2 and Block 3. The data revealed that the attention preferences of the participants did not follow where A went.

   Meanwhile similar situations happened to B and C as well. While B on the Block 1, eye gazes of the participants were taken the most times with 981 shoots. While C on the Block 1, eye gazes of the participants were taken the most times with 946 shoots. The data revealed that the attention preferences of the participants did not follow where B or C went, either.

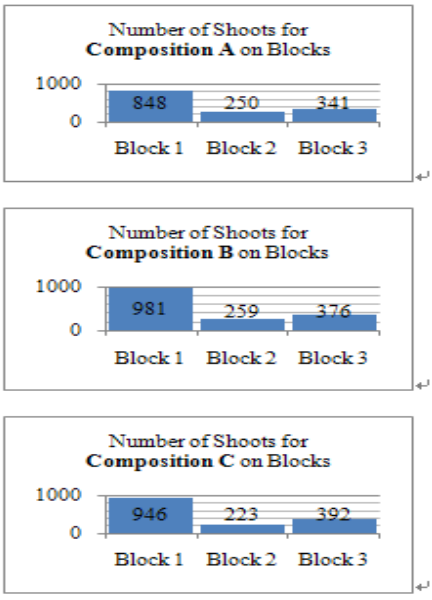

Fig. 4. Comparisons for Positions: Comparisons for the Respective

2. Comparisons for the attention preferences of 3 layout compositions.

   i. The comparisons for Composition A, B, C on the same block (as shown in Fig. 5): Data showed that the eye gazes for Composition A, B, C on the same block received close attentions contemporarily. There were no significant differences among types of layout compositions. Whereas, the attention preferences turned to be attracted to whatever type of layout composition presented on Block 1, which is the left block on the webpage.

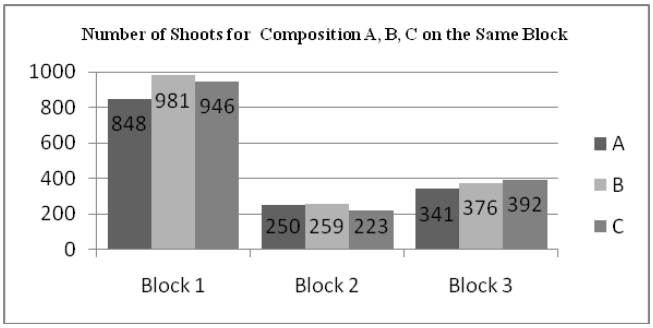

Fig. 5. Comparisons for Attention Preferences via the Eye Gazes on the Same Block

   ii. The cross comparisons for Composition A, B, C on different blocks (as shown in Fig. 6): Data showed that the cross examined eye gazes for Composition A, B, C on the different blocks received close attentions contemporarily. There were no significant differences among types of layout composition.

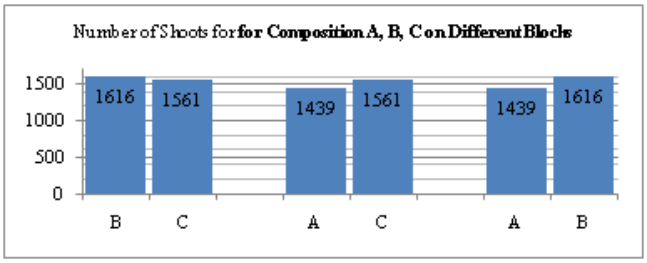

Fig. 6. Cross Comparisons for Composition A, B, C

   iii. The fixation position at start and the navigation transitions (as shown in Fig. 7): Data showed that at the first 0.4 second, the fixation position at start mostly located on Block 1, which is the left block on the webpage. Whereas, it could be noticed that even thought during only 0.4 second at start, eyes navigated across the whole webpage and gazed much less on the middle block. About at the half time of navigation, the fixation position changed slightly.

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

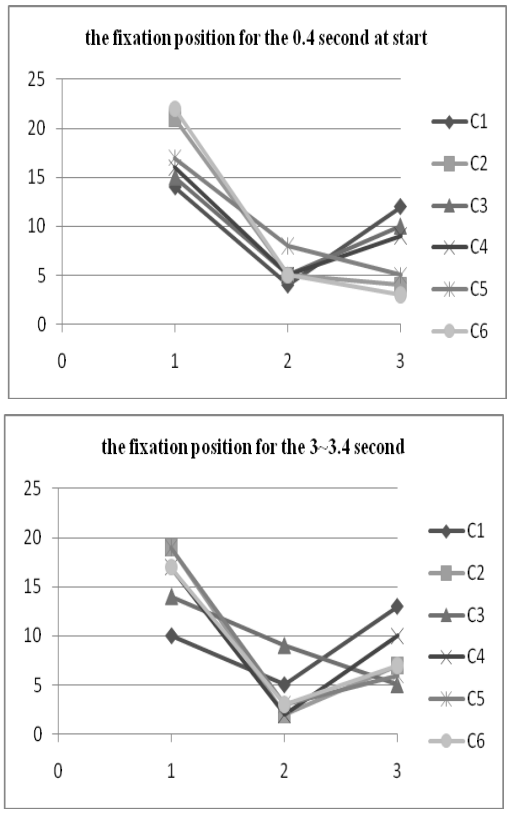

Fig. 7. Fixation Position at Start and at the Half Time

V. Conclusion

The gaze area of the participants mostly remained in the left block on the webpage. There were no significant difference of the gaze area between the middle block and right block. The result showed that the attention preferences of the participants were not affected by the layout compositions but affected by where the information was presented on the screen. In this case, the left block attracted the attention preferences of the participants. Therefore, information put on the left block of the screen can catch the audience's attention naturally.

Also, the result showed that the fixation position at start mostly located on the left block on the webpage. It also showed that eyes navigated across the whole webpage even at a glance of time but gazed much less on the middle block. Along with time, the fixation position changed slightly but data did not provide enough information to discover any pattern. In the future, it might be able to provide more insights if the experiment includes more participants.

References

[1] R. C. Li, "A Study on Icon Semantics in Human-Computer Interaction," M. S. thesis, Dept. Industrial Design, Cheng Kung Univ., Taiwan, 1993.

[2] Mario T. Garica,, Contemporary newspaper design. Englewood Cliffs, NJ: Prentice Hall, 1987.

[3] K. Holmqvist, J. Holsanova, M. Barthelson, and D. Lundqvist, "Reading or Scanning A Study of Newspaper and Net Paper Reading," in The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research, R. Radach, J. Hyona, and H. Deubel, Eds. New York: Elsevier, 2003.

[4] N. Holmberg, "Eye movement patterns and newspaper design factors. An experimental approach," M. S. thesis, Lund University Cognitive Science, 2004.

[5] M. S. Sanders and E. J McCormick, Human Factors in Engineering and Design, New York: McGraw-Hill International Editions, 1987.

[6] H. C. Chen, H. D. Lai, and F. C. Chiu. "Eye Tracking Technology for Learning and Education," *Journal of Research in Education Sciences*, vol. 55, no. 4, pp. 39-68, Dec. 2010.

[7] M. Perugini and R. Banse, "Personality, implicit self-concept and automaticity," European Journal of Personality, vol. 21, no. 3. pp. 257-261, Apr. 2007.

[8] C. Taft, "Color meaning and context: comparisons of semantic ratings of colors on samples and objects," Color Research and Application, vol. 22, no. 1, pp. 40-50, 1996.

[9] T. R. Lee, "New comparison of psychological meaning of colors in samples and objects with semantic ratings," Congress of the International Color Association, Rochester, NY, USA, Jun. 24, 2001.

[10] Y. Zhu, Experimental Psychology, 2nd ed. Beijing: Peking University Press, 2009.

[11] J. M. Henderson and A. Hollingworth, "High-level scene perception," Annual Review of Psychology, vol. 50, pp. 243-271, 1999.

[12] A. T. Duchowski, Eye tracking methodology: theory and practice, 2nd ed. New York: Springer, 2007.

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

[14] The Advanced Knowledge Provider PITOECH CO.,LTD. (2009, Apr. 12). [On-line]. Available: http://www.digitalwall.com/scripts/displaypr.asp?UID=13865

[15] H. C. Chen, S. L. Peng, C. C. Tseng, and H. W. Chiou, "An Exploratory Study of the Relation Between the Average Saccade Amplitude and Creativity Under the Eyetracker Mechanism, " Bulletin of Educational Psychology, vol. 39, pp. 127-149, 2008.

[16] J. E. Hoffman, & B. Subramaniam," The role of visual attention in saccadic eye movements," Perception and Psychophysics, vol. 57, no. 6, pp. 787-795, 1995.

[17] D. L. Tang, T. R. Lee, and C. M. Tsai. "An Exploratory Study on Relationship between Preference and Scanpath-Evidence from Color Preference Sorting Task," Chinese Journal of Psychology, vol. 47, no. 4, pp. 339-351, Dec. 2005.

[18] B. Shneiderman and C. Plaisant, Designing the User Interface: Strategies for Effective Human-Computer Interaction, 4th ed. Boston: Addison-Wesley, 2004.

[19] 2010 Taiwan WEB 100 (2010, Apr. 02). [Online]. Available: http://163.26.9.13/trip/2010%E5%8F%B0%E7%81%A3web100%E6%8E%92%E5%90%8D%E5%88%86%E9%A1%9E.htm

[20] Top 20 Websites for Different Ages in January, 2010 (2010, Mar. 03). [Online]. Available: http://blog.xuite.net/hinet_marketing/blog/32609798

[21] Takeshi Mita, Toshimitsu Kaneko , Osamu Hori ," Joint Haar-like Features for Face Detection," Proc. of the Tenth IEEE International Conference on Computer Vision ,2005.

[22] C. W. Yang, C. W. Kao, K.C. Fan, B. J. Hwang, and C. P. Huang, "Eye gaze tracking based on pattern voting scheme for mobile device," presented at The First International Conference on Instrumentation & Measurement, Computer, Communication and Control, Beijing, China, 2011.

ISBN: 978-988-19251-1-4
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)