

Scan-Path Analysis by the String-Edit Method Considering Fixation Duration

Haruhiko Takeuchi

National Institute of Advanced Industrial Science
and Technology (AIST)
1-1-1 Higashi, Tsukuba, 305-8566, Japan
takeuchi.h@aist.go.jp

Noriyuki Matsuda

University of Tsukuba

1-1-1 Tennou-dai, Tsukuba, 305-8573, Japan
mazda@sk.tsukuba.ac.jp

Abstract—Dynamic aspects of eye-tracking data are important but difficult to analyze. With string based approaches, a sequence of fixations is analyzed, however, fixation duration is not addressed. Cristino et al. recently proposed to re-code a scan-path with a long fixation by repeating the code. The modified scan path includes both fixation duration and sequence of fixations. In studying multiple records by the string-edit method enhanced with cost functions, we compared the performance of the modified coding against the ordinary one. Furthermore, we derived representative scan paths to examine the distance among the web pages used as stimuli. The usefulness of our approach is demonstrated.

Keywords- eye-tracking; scan path; data analysis; string-edit method; Levenstein method; fixation duration

I. INTRODUCTION

Eye-tracking systems are widely used in various fields in designing and evaluating human-centered systems. Among the various indices that characterize eye-tracking data [1], scan paths reflect the dynamism of eye movement in temporal sequences [2].

Despite its importance, there are very few quantitative methods for analyzing the scan path, such as a comparison of sequences. One quantitative method of scan-path analysis is the string-edit method, also known as the Levenstein method or optimal matching analysis [3, 4]. Application of the string-edit method to scan-path analysis has resulted in several important improvements. Takeuchi et al. [5] proposed a new method that incorporates the spatial distance between two fixations as substitution costs. This method has been used in various eye-tracking applications [6].

Although the string-edit method is a powerful method, it analyzed only sequence of fixations and did not address fixation duration. Cristino et al. [7] recently suggested a solution to overcome this shortcoming by modifying the coding method in combination with a string-matching algorithm originally developed for DNA sequences that normally contained repeated codes. In this paper, we propose the use of the modified scan path in the string-edit method. We compare

the the representative scan path derived from the original scan path with that derived from the modified scan path, using Web-viewing experiment data. We also demonstrate the usefulness of incorporating the modified scan path in eye-tracking data analysis.

II. METHOD

In this section, we describe the algorithm of the string-edit method, extended for eye-tracking data considering cost functions, developed by the authors [5]. We then apply this method to the modified scan path, considering fixation duration.

A. Data

The scan path is time series data of fixations, derived from eye-tracking data. The raw scan path is expressed as values of the x and y axes of a graphic display, with duration calculated from eye-tracking data. We divide the graphic area into several segments, and assign a unique letter for each segment. We then translate the scan path into a string of letters. Figure 1 presents an example of the graphical partition. In this case, the whole area is divided into nine segments, and letters “A” through “I” are assigned. An example of a string is “ABBCEHF.” How the area is divided depends on the contents of the Web pages and the purpose of the analysis. The division should be in precise order to deal with detailed contents. However, the division should be large to facilitate understanding of the general tendency of the scan path. Division by Area of Interests (AOI) can also be applied. In this case, the center position of each AOI is needed for calculation.

A	B	C
D	E	F
G	H	I

Figure 1. Example of graphic partition of 3x3 segments.

B. String-edit method with cost function

The string-edit method calculates the distance between two strings using three basic operations: deletion, addition, and substitution. Using these operations, one string is transformed into another string. Every time an operation is applied, a pre-assigned cost is accumulated. The distance between the two strings is defined as the smallest cost for matching those two strings [8].

As Josephson and Holmes [3] noted, it is reasonable to assign appropriate values for substitution cost. The authors [5] introduced cost functions for analyzing scan-path data. Substitution of spatially close fixation points costs little, whereas substitution between long distance points costs much. In this model, the substitution cost is defined by the city block distance or Euclidian distance between the central coordinates of the areas.

We define the geographical position for each letter of a string as follows. The specified position (u_1, u_2) for the letter u corresponds to the center of the assigned geographical area of the letter.

$$u = (u_1, u_2), v = (v_1, v_2) \quad (u, v \in W) \quad (1)$$

Here, W denotes the set of letters whose elements stand for the segment names. We define cost function based on Euclidian distance. The substitution cost f from u to v , and vice versa, is defined below. In the following equation, the parameter α is a normalization parameter that differs according to the magnitude of substitution cost, in contrast with the addition cost and the deletion cost, which are set to one. The substitution cost for Euclidian distance is defined using the following formula:

$$f(u, v) = \alpha \sum_{i=1}^2 \sqrt{(u_i - v_i)^2} \quad (u, v \in W) \quad (2)$$

Table I presents an example of the substitution costs based on Euclidian distance for the data in Fig. 1. Here, we assume that the effective screen size is 750 x 750 pixels and that α is 0.001.

TABLE I. EXAMPLE OF SUBSTITUTION COSTS FOR THE DATA IN FIG.1 BASED ON THE EUCLIDIAN DISTANCE MODEL

	A	B	C	D	E	F	G	H	I
A	0	0.25	0.5	0.25	0.35	0.56	0.5	0.56	0.71
B	0.25	0	0.25	0.5	0.25	0.35	0.56	0.5	0.56
C	0.5	0.25	0	0.25	0.5	0.25	0.35	0.56	0.5
D	0.25	0.5	0.25	0	0.25	0.5	0.25	0.35	0.56
E	0.35	0.25	0.5	0.25	0	0.25	0.5	0.25	0.35
F	0.56	0.35	0.25	0.5	0.25	0	0.25	0.5	0.25
G	0.5	0.56	0.35	0.25	0.5	0.25	0	0.25	0.5
H	0.56	0.5	0.56	0.35	0.25	0.5	0.25	0	0.25
I	0.71	0.56	0.5	0.56	0.35	0.25	0.5	0.25	0

C. Embedding fixation duration into scan-path data

A letter in the original scan path indicates the fixation corresponding to the segment represented by that letter. Cristino et al. [7] modified this concept. They introduced temporal binning into the string by repeating the letter in a way that was proportional to the fixation duration. In their notification, a letter in a scan path indicated 50msec fixation corresponding to the segment represented by that letter. When the fixation duration was longer than 100msec, they added extra letters until the number of letters corresponded to the fixation duration.

In this paper, we used 100msec bins (i.e., a single letter in a scan path means about 100msec fixation duration). When the fixation duration is between 100msec and 150msec, a single letter (e.g., "A") is used. When the fixation duration is between 150msec and 250msec, a double letter (e.g., "AA") is used. For 100msec bins, the number of the duplications of a letter is defined by the following formula.

$$n = [(t - 50)/100] + 1 \quad (3)$$

Here, parameter t (msec) denotes the fixation duration, and the brackets are a floor function. Thus n is the maximum integer that is no greater than the value in the brackets plus one.

D. Steps for scan-path comparison

The procedure for string-edit method with cost functions is as follows.

Step 1. Partition the screen. Divide the graphic area into segments, and assign a unique letter for each segment, using a letter code from "A" to "Z," and then from "a" to "z." AOI can also be applied instead of the grid partition. When AOI is applied, the center position of each AOI is used in step 4.

Step 2. Transform fixation data into strings. Each fixation data point is translated into a letter for each specified segment. In this stage, fixation duration data are attached to each letter. This paper refers to a string of letters without fixation duration data as an original scan path.

Step 3. Add letters for modified scan path. By checking the fixation duration attached to each letter, the same letters are duplicated. The number of duplications is calculated by formula (3). This paper refers to this string of letters as an modified scan path.

Step 4. Select the distance model. The Euclidian distance model is the first choice, and the city block distance model is the second choice. Substitution costs are automatically assigned according to the selected distance model. This substitution cost matrix is used in the next step.

Step 5. Calculate the dissimilarity. Dissimilarities among scan paths are calculated by the string-edit method. The distance is usually normalized, i.e., dividing the distance by the length of the longer string. Further analysis is performed for the derived distance matrix.

III. REAL DATA APPLICATION

A. Eye-tracking experiments

In this section, we discuss the details of our eye-tracking data, which were measured as a series of Web usability tests.

Subjects Twenty subjects were recruited for this study. Seven subjects were males, and 13 were females; their ages ranged from 19 to 48 years (average of 30 years). They were divided into two Web user groups, with eleven heavy Internet users and nine light Internet users, based on their weekly Internet usage. However, their Internet experience is not discussed in this paper.

Stimuli The stimuli presented to the subjects were ten Web pages with three different menu structures. These Web pages were chosen, because each Web page belongs to one of three distinct menu structures: menu on the left (L), menu on the right (R), and menus on both sides (B).

Apparatus A Tobii 1750 eye-tracking system with a TFT 17inch display and a maximum resolution of 1024 x 768 was used in this study. The eye tracker has a tracking rate of 50Hz.

Procedure The Web pages were randomly displayed to the subjects one at a time. Subjects were instructed to browse each Web page until the screen darked, and then to click the mouse button to proceed. The eye-tracking data for 20sec for each Web page were recorded.

B. Data preparation

This section describes the data preparation. Fixation points and their durations were extracted from the raw eye-tracking data. Here, the minimum fixation duration was set to 100msec, and the greatest radius was set to 30 pixels.

First, we transformed the fixation data into string data, as discussed in the previous section. We divided each Web page into 5x5 segments, dividing the effective screen horizontally and vertically into five partitions each. We assigned a letter code from "A" to "Z" to each segment in descending order from the first column to the fifth column in each row. Second, we translated each fixation data point into a letter that represented the corresponding segment, with strings of letters for each subject's fixation data. The original letter sequences were modified to reflect sustained. Third, we made scan paths using durations from the scan path above, by adding the same letters until each fixation corresponded to the fixation time (see step3 in the previous section). We referred to this scan path as the modified type.

Here, we present an example of the two types of scan paths. The fixation data are: "C:120, D:200, E:140, C:320, F:120, G:180, A:100, B:120, B:300, A380, K:120, L:100, L:140, F:420, G:220." The numeric value after each letter indicates fixation duration (msec).

(a) Original scan path: CDECFGABBAKLLFG

(b) Modified scan path: CDDECCCFGGABBBBAAAA
KLLFFFFGG

We treated ten Web pages with the two types of scan paths. Thus we analyzed 20 scan paths for each 10 Web pages and each scan-path type, using the string-edit method. Here, we used the Euclidian distance model. The parameter α was set to 0.001, and the effective screen size was set to 1000 x 1000. Normalization by the length of the compared scan paths was applied.

C. Results and discussion

We sought to find the representative scan path with many people viewing the same Web page. We call it representative scan path. For this purpose, the distances between scan paths were calculated by the string-edit method for each Web page. The summed distances between one scan path and others were then calculated as follows.

$$D_i = \sum_{j=1}^{20} \sqrt{(d_{ij})^2} \quad (i = 1, 2, \dots, 20) \quad (4)$$

Here, d_{ij} denotes the distance between scan path i and scan path j , calculated by the string-edit method. The representative scan path for each Web page was identified as follows. The representative scan path was defined as scan path i , such that D_i has the minimum value among all D_s .

We compared the original scan path and the modified scan path. Table II summarizes the identified representative scan paths for each Web page. For example, S16 was the scan path of subject 16. The same representative scan path was identified for three Web pages, but the representative scan paths differed for seven Web pages. Specifically, the representative scan paths were different for type L pages. For type R pages, WP5 had the same representative scan path, and the other Web pages had different representative scan paths. For type B pages, WP1 and WP3 had the same representative scan path, and other Web pages had different representative scan paths.

TABLE II. REPRESENTATIVE SCAN PATHS FOR EACH WEB PAGE

Type	Web page	Representative scan path	
		Original	Modified
L	WP2	S16	S2
	WP9	S18	S2
R	WP4	S8	S2
	WP5	S3	S3
	WP7	S5	S18
	WP10	S13	S16
B	WP1	S17	S17
	WP3	S18	S18
	WP6	S4	S19
	WP8	S2	S9

TABLE III. CORRELATION COEFFICIENTS BETWEEN ORIGINAL TYPE AND MODIFIED TYPE

Type	Web page	Correlation
L	WP2	0.78
	WP9	0.73
R	WP4	0.67
	WP5	0.91
	WP7	0.59
	WP10	0.87
B	WP1	0.87
	WP3	0.58
	WP6	0.72
	WP8	0.66

Table III presents the correlation coefficients between the original type and the modified type, calculated from value D. They varied from 0.58 to 0.91, with an average 0.74. There was some correlation between the original type and modified type, but it was not so high. WP5, which was cited from a bank company Web site, had the highest correlation (0.91). The design of WP5 was the simplest of all Web pages.

For more detailed analysis of the two types of scan paths, we selected the best and the second best representative scan paths from each Web page. We made a distance matrix of 20 representative scan paths using the string-edit method with the same cost function. We then analyzed the distance matrixes of the representative scan paths by multi-dimensional scaling (MDS) [9]. The configurations of two-dimensional results for the original type of the scan path are presented in Fig. 2, and those for the modified type of scan paths are presented in Fig. 3. Here, for example, the representative scan path of WP1 was plotted as W1. In the figures, the scan paths of the same Web page are circled.

On the configuration of the original type (Fig. 2), WP5, WP9, WP3, and WP7 were plotted far from the center position, and those Web pages were thought to be viewed uniquely. Six of the same Web pages were comparatively closely plotted, and four of them were not. In the configuration of the modified type (Fig. 3), WP5, WP9, WP1, WP6, and WP7 were plotted far from the center position, and those Web pages were thought to be viewed uniquely too. Eight of the same Web pages were comparatively closely plotted, and two of them were not. As the scan paths from the same Web page should be placed close to one another, the scan paths of the modified type had better results.

We considered reasons for the difference between the original scan path and the modified scan path. One of the most dominant reasons was the difference in string length. In the experiment above, the average number of letters was 53.3 for the original type, and 130.2 for the modified type (Table IV), making the string length of the modified type nearly twice that of the original type. In addition, the parameter setting in the

scan-path coding also affected the string length of the modified type. In this study, we used 100msec bins for one letter; however if we had used 50msec bins for one letter, the string length of the modified scan path would have increased more radically.

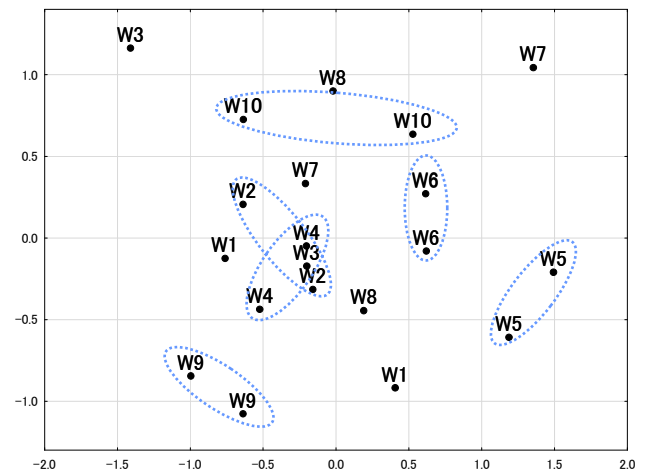


Figure 2. MDS configuration of representative scan paths of the original type

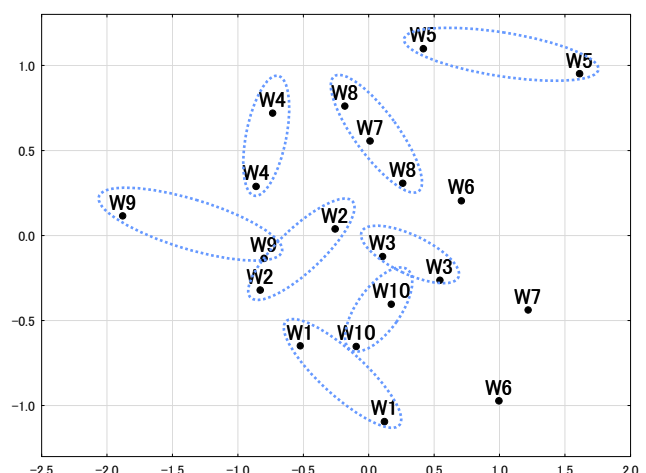


Figure 3. MDS configuration of representative scan paths of the modified type

TABLE IV. AVERAGE NUMBER OF LETTERS FOR ORIGINAL TYPE AND MODIFIED TYPE

	Number of letters	
	Original	Modified
Average	53.3	130.2
SD	10.4	31.5

The expression of the modified scan path might be an accurate expression, because the modified scan path holds the sequence, positions, and durations of fixations. However when we want to know only the sequence and positions of human Web-viewing behavior, the addition of letters may simply be additional information. Further experiments are needed to determine the merits and the demerits of the modified scan path.

IV. CONCLUSION

This paper presented an application of the string-edit method to the modified scan path, which held the sequence, positions, and durations of fixations. We compared the original scan path and the modified scan path, using Web-viewing experiment data. The average correlation coefficient was 0.74, and the extracted representative scan paths were mostly differed. We used MDS to analyze the distance matrixes of the representative scan paths derived from the original scan paths and the modified scan paths. The MDS configuration of the modified scan paths had better results.

The modified scan path differs much from the original scan path, so we need further study to select the suitable scan-path type, depending on the purpose of the analysis. In future work, we will analyze the scan path with chunks observed in eye-tracking data [10] using string-based methods. It would be interesting if we could establish a framework for analyzing scan paths with gaps. We also hope to combine network analysis [11] and the presented method in future. As Dempere-Marco [12] reported, integrated analysis is essential for eye-tracking data analysis.

REFERENCES

- [1] R.J.K. Jacob and K.S. Karn, "Eye tracking in Human-computer interaction and usability research: Ready to deliver the promises," in the *Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*, ed. by J. Hyona, R. Radach, and H. Deubel, pp. 573-605, Amsterdam, Elsevier Science, 2003.
- [2] D. Noton and L. Stark, "Scanpaths in saccadic eye movements while viewing and recognizing patterns," *Vision Research*, vol. 11, pp. 929-942, 1971.
- [3] S. Josephson and M.E. Holmes, "Attention to repeated images on the World-Wide Web: Another look at scanpath theory," *Behavior Research Methods, Instruments, & Computers*, vol. 34, pp. 539-548, 2002.
- [4] H. Hembrooke, M. Feusne and G. Gay, "Averaging scan patterns and what they can tell us," in *Proc. Symposium on Eye Tracking Research & Applications (ETRA)*, pp. 41, 2006.
- [5] H. Takeuchi and Y. Habuchi, "A Quantitative Method for Analyzing Scan Path Data Obtained by Eye Tracker," *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, pp. 283-286, Honolulu, 2007.
- [6] F. Galgani et al., "Automatic analysis of eye tracking data for medical diagnosis," *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, pp. 195-202, 2009.
- [7] F. Cristino, S. Mathôt, J. Theeuwes and I.D. Gilchrist, "ScanMatch: a novel method for comparing fixation sequences," *Behavior Research Methods*, Vol. 42(3), pp. 692-700, 2010.
- [8] D. Sankoff and J.B. Kruskal (eds.), "Time warps, string edits, and macromolecules: The Theory and Practice of Sequence Comparison," *CSLI Publications*, 1999.
- [9] J. B. Kruskal and M. Wish, *Multidimensional scaling*, Beverly Hills: Sage Publications, 1978.
- [10] N. Matsuda and H. Takeuchi, "Frequent Pattern Mining of Eye-Tracking Records decomposed into Chunks," *Abstract book of the 16th European Conference on Eye Movements (ECM)*, Marseille, 2011.
- [11] N. Matsuda and H. Takeuchi, "Joint analysis of static and dynamic importance in the eye-tracking records of Web page readers," *Journal of Eye Movement Research*, Vol. 4(1), No. 5, pp. 1-12, 2011.
- [12] L. Dempere-Marco et al., "A Novel Framework for the Analysis of Eye Movements during Visual Search for Knowledge Gathering," *Cognitive Computation*, Vol. 3, Num. 1, pp. 206-222, 2011.