Journal of Web Engineering, Vol. 14, No. 5&6 (2015) 363–385

PATTERNS IN EYETRACKING SCANPATHS AND THE AFFECTING FACTORS

SUKRU ERASLAN [1,2]

[1] *Middle East Technical University, Northern Cyprus Campus,*
99738 Kalkanli, Guzelyurt, Mersin 10, Turkey
seraslan@metu.edu.tr

[2] *School of Computer Science, University of Manchester,*
Manchester, United Kingdom
sukru.eraslan@postgrad.manchester.ac.uk

YELIZ YESILADA

Middle East Technical University, Northern Cyprus Campus,
99738 Kalkanli, Guzelyurt, Mersin 10, Turkey
yyeliz@metu.edu.tr

Web pages are typically decorated with different kinds of visual elements that help sighted people complete their tasks. Unfortunately, people accessing web pages in constrained environments, such as visually disabled and small screen device users, cannot benefit from them. In our previous work, we show that tracking the eye movements of sighted users provide good understanding of how people use these visual elements. We also show that reengineering web pages by using these visual elements can improve people's experience in constrainted environments. However, in order to reengineer web pages based on eyetracking, we first need to aggregate, analyse and understand how a group of people's eyetracking data can be combined to create a common scanpath (namely, eye movement sequence) in terms of visual elements. This paper presents an algorithm that aims to achieve this. This algorithm was developed iteratively and experimentally evaluated with an eyetracking study. This study shows that the proposed algorithm is able to identify patterns in eyetracking scanpaths and it can work well with different number of participants. We then extended our experiments to investigate the effects of the task, gender and familiarity factors on common scanpaths. The results suggest that these factors can cause some differences in common scanpaths. This study also suggests that this algorithm can be improved by considering different techniques for pre-processing the data, by addressing the drawbacks of using the hierarchical structure and by taking into account the underlying cognitive processes.

Keywords: eyetracking, scanpaths, patterns, commonality, reengineering

1 Introduction

Web pages mainly consist of different kinds of visual elements, such as menu, logo and hyperlinks. These visual elements help sighted people complete their tasks, but unfortunately small screen device users and disabled users cannot benefit from these elements. When people access web pages with small screen devices, they typically experience many difficulties [22]. For example, on small screen devices, only some parts of web pages are accessible or the complete web page is available with very small text size. Hence, they may need to scroll or zoom a lot which can be annoying. They also may need more time and effort to find their

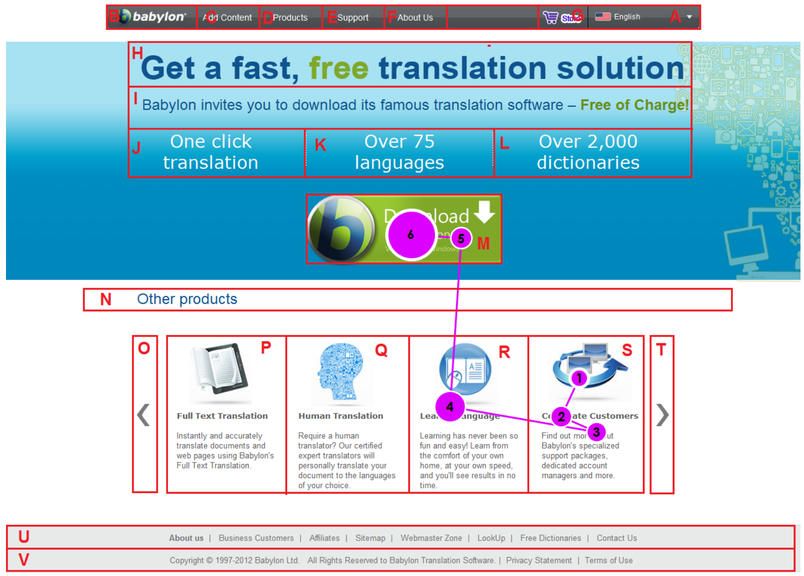

Fig. 1. A scanpath on a segmented web page.

targets. Similarly, web experience can be challenging for visually disabled users who typically use screen readers to access the web [25]. Since screen readers follow the source code of web pages, visually disabled users have to listen to unnecessary clutter to get to the main content [26].

In our previous work, we show that reengineering web pages by using the visual elements can improve the user experience in constrainted environments [4]. However, identifying visual elements and their role is the key for such reengineering process. To automatically process a web page and identify these elements, in our previous work we have extended and improved the Vision Based Page Segmentation (VIPS) algorithm [1, 2]. This extended algorithm automatically discovers visual elements and relates them to the underlying source code. It allows direct access to these visual elements via XPath. However, this algorithm does not provide any information on how these visual elements are used. In our previous work, we also show that tracking the eye movements of sighted users provide good understanding of how they are used [25]. Eyes make quick movements which are called saccades. Between saccades, eyes make fixations where they become relatively stationary. Both fixations and saccades create scanpaths which are eye movement sequences [18]. Fig. 1 shows how a web page is segmented into its visual elements and illustrates a scanpath on a segmented web page. The circles represent fixations where the larger circles represent longer fixations. The numbers in the circles show the sequence. Also, the lines between circles are saccades. This scanpath illustrates that the user first fixated a point inside of the visual element S and then fixated two other points in the same element. After that, the user moved to another point inside of the visual element R and then fixated two consecutive points in the visual element M.

In order to be able to use eyetracking data for reengineering web pages, this paper presents an algorithm called "*eMine scanpath algorithm*".[a] This algorithm analyses and aggregates a

[a] `http://emine.ncc.metu.edu.tr/`

group of people's eyetracking data to create a common scanpath in terms of visual elements of web pages (Section 3). Web pages are first automatically segmented into visual elements with the extended and improved version of the VIPS algorithm [2, 1]. Eyetracking data is then exported and related to these visual elements. This creates individual scanpaths of users in terms of visual elements. These individual scanpaths are then used by eMine scanpath algorithm to create a common scanpath. eMine scanpath algorithm was iteratively developed with the existing eyetracking data and our preliminary evaluation of this algorithm with the existing data was promising [24].

In order to experimentally evaluate the usefulness and scalability of this algorithm, we conducted a new eyetracking study with 40 participants at two different sites in two different countries: Middle East Technical University Northern Cyprus Campus and the University of Manchester in the UK (Section 4). As presented in [8], this study illustrates that eMine scanpath algorithm is able to identify a common scanpath in terms of visual elements of web pages and it can work well with different number of participants (Sections 5 and 6). We then extended our experiments in order to investigate whether or not the task, gender and familiarity can affect common scanpaths. The results show that these factors are likely to cause some differences in common scanpaths (Sections 5 and 6). This study also revealed some directions for future work, such as pre-processing eyetracking data to improve its quality and taking cognitive processes into consideration (Section 7).

2 Related Work

Eyetracking scanpaths have been analysed with different methods for different purposes. These methods typically use string representations of scanpaths which are generated using the sequence of Areas of Interest (AoIs) [21]. Different ways can be used to generate these AoIs such as using a grid layout directly [21] or the fixations' distribution over web pages [20]. However, these existing approaches typically treat a web page as an image to identify these AoIs which means these scanpaths cannot be used to process web pages. In order to address this, our previous work automatically segments a web page into its visual elements and each segment becomes an AoI [1, 2]. This allows relating AoIs with the underlying source code which is important for being able to process web pages by using the eyetracking data. Therefore, the string representation of the scanpath in Fig. 1 is generated as SSSRMM because the first three fixations are located inside of the AoI S, the next fixation is located inside of the AoI R and the last two fixations are located inside of the AoI M. We can also access to the source code of the AoIs using XPath to process them.

The Levenshtein Distance (String-Edit) algorithm has commonly been used to analyse scanpaths [14, 21]. This algorithm calculates the dissimilarity between the string representations of two scanpaths by transforming one to another with a minimum number of operations (insertion, deletion and substitution). For example, the dissimilarity between XYCZ and XYSZ is calculated as 1 (one) by the String-Edit algorithm because the substitution C with S is sufficient to transform one to another. Although the String-edit algorithm can be used to categorise scanpaths [23] and investigate differences between the behaviours of people on web pages [14], the algorithm itself is not able to identify a common scanpath for multiple scanpaths. Similar to the String-edit algorithm, Needleman and Wunsch algorithm has also been used to calculate the similarity between the string representations of two scanpaths [7].

Transition Matrix is one of the methods which use multiple scanpaths to create a matrix [23]. This matrix allows identifying the possible next and previous AoI of the particular AoI. However, when this method is considered for identifying a common scanpath, some considerable problems arise, such as what the start and end point of the common scanpath are and which probabilities should be considered.

To address these problems, some other methods can be considered. For example, the Shortest Common Supersequence method has been mentioned in literature to identify a common scanpath for multiple people but it has considerable weaknesses [19]. For example, it identifies XABCDEZ as a common scanpath for the individual scanpaths XAZ, XBZ, XCZ, XDZ and XEZ. As can be easily recognised, the common scanpath is not supported by the individual scanpaths, for instance, the common scanpath has E which is included by only one individual scanpath (XEZ). Furthermore, the common scanpath is quite longer compared to the individual scanpaths.

Some methods, such as T-Pattern [16] and eyePatterns's Discover Patterns [23], have been proposed to detect subpatterns in eyetracking scanpaths. However, eyePatterns's Discover Patterns method [23] is not tolerant of extra items in scanpaths. For instance, XYZ can be detected as a subpattern for XYZ and WXYZ but it cannot be detected for XYZ and WXUYZ because of the extra item U. This shows that this method is reductionist which means it is likely to produce unacceptable short scanpaths.

The Multiple Sequence Alignment method was proposed to identify a common scanpath but this method was not validated [11]. Moreover, the Dotplots-based algorithm was proposed to identify a common scanpath for multiple people [10]. This algorithm creates a hierarchical structure by combining a pair of scanpaths with the Dotplots algorithm. The individual scanpaths are located at leafs whereas the common scanpath is located at the root. Some statistical methods have been applied to address the reductionist approach of the Dotplots algorithm [10].

We are interested in common patterns in eyetracking data instead of individual patterns to be able to reengineer web pages. However, as can be seen above, there is not much research in identifying common scanpaths and the existing ones are likely to produce unacceptable short common scanpaths. In this paper, we present our eMine scanpath algorithm to address the limitations of these existing approaches, especially the problem of being reductionist. Since these existing methods have been designed to find a common scanpath without taking the task, gender and familiarity factors into consideration, we also present these factors that are likely to cause some differences in common scanpaths.

3 eMine Scanpath Algorithm

Algorithm 1 shows our proposed eMine scanpath algorithm [24] which takes a list of scanpaths and returns a scanpath which is common in all the given scanpaths. If there is only one scanpath, it returns that one as the common scanpath, if there is more than one scanpath, then it tries to find the most similar two scanpaths in the list by using the String-edit algorithm [14]. It then removes these two scanpaths from the list of scanpaths and introduces their common scanpath produced by the Longest Common Subsequence method (LCS) [6] to the list of scanpaths. This process is illustrated in Fig. 2 with eight scanpaths on the Babylon page (See Fig. 1). In the first iteration the most similar scanpaths are S_1 and S_7, and therefore

Algorithm 1 Find common scanpath

Input: Scanpath List
Output: Scanpath

```
if the size of Scanpath List is equal to 1 then
    return the scanpath in Scanpath List
end if
while the size of Scanpath List is not equal to 1 do
    Find the two most similar scanpaths in Scanpath List with the String-edit algorithm
    Find the common scanpath by using the Longest Common Subsequence method
    Remove the similar scanpaths from the Scanpath List
    Add the common scanpath to the Scanpath List
end while
return the scanpath in Scanpath List
```

the LCS method is applied to these two scanpaths and their common scanpath is produced. S_1 and S_7 are then removed from the list and their common scanpath is introduced to the list of scanpaths. This continues until there is only one scanpath. The abstracted version of that scanpath represents the common scanpath. In order to have abstracted version, its string representation is simplified by abstracting consecutive repetitions [3, 12]. For instance, MMPPQRSS becomes MPQRS.

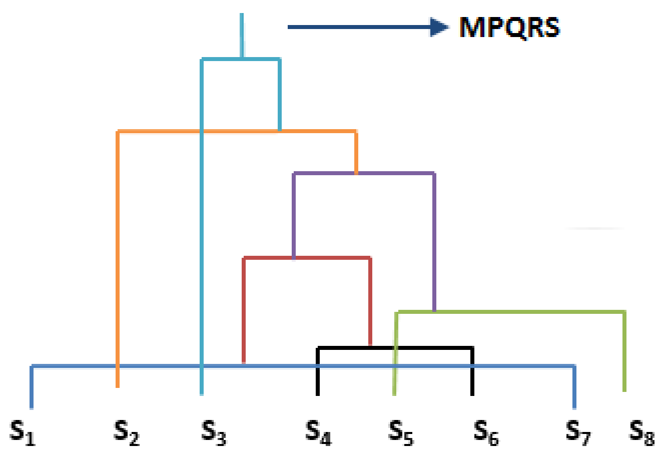

Fig. 2. eMINE scanpath algorithm applied on eight scanpaths on the Babylon page.

3.1 System Architecture and Implementation

eMine scanpath algorithm was integrated with the extended and improved version of the VIPS algorithm [2, 1]. Fig. 3 illustrates the system architecture which consists of the following parts: two input parts (web page and eyetracking data), three functional parts (web page AoI identification, an application to create string representations of scanpaths and eMine scanpath algorithm), two intermediate parts which are created as an output of one functional part and used as an input for another functional part (web page AoIs and string representations of scanpaths) and one output part (common scanpath). The functional parts are explained below.

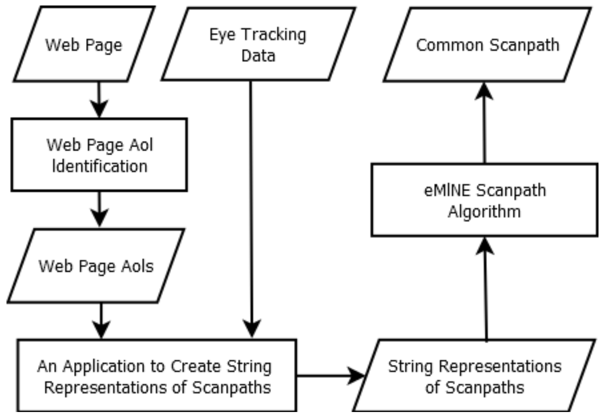

Fig. 3. System architecture

Web Page AoI Identification A web page is used as an input for the web page AoI identification part. This part creates AoIs automatically by using the extended and improved version of the VIPS algorithm [2, 1]. These AoIs represent visual elements of web pages. Even though, the extended VIPS was used, it would be easily replaced by an alternative method of AoI identification approach such as a grid-layout segmentation [21].

An Application to Create String Representations of Scanpaths The automatically generated web page AoIs and eyetracking data, provided by eyetracking software, are then used by an application to create string representations of scanpaths.

eMine Scanpath Algorithm Once the string representations are created, our scanpath algorithm is applied to them to produce a common scanpath in terms of AoIs.

eMine scanpath algorithm[b] was implemented on the Accessibility Tools Framework (ACTF)[c] which is an open-source Eclipse project.

4 An Eyetracking Study

In order to experimentally evaluate the usefulness and scalability of eMine scanpath algorithm, and investigate the effects of the task, gender and familiarity on common scanpaths identified by eMINE scanpath algorithm, we conducted an eyetracking study. This study aims to investigate the following three research questions:

1. *Usefulness:* The aim is to investigate whether or not eMine scanpath algorithm can successfully identify common scanpaths in terms of visual elements of web pages. Thus, we ask *"Can eMine scanpath algorithm identify common scanpaths in terms of visual elements of web pages?"*.

[b] `http://emine.ncc.metu.edu.tr/software.html`
[c] `http://www.eclipse.org/actf/`

2. *Scalability:* We would like to investigate whether or not eMine scanpath algorithm works well for different numbers of participants on different web pages. Hence, the research question here is *"How does the number of individual scanpaths affect common scanpaths?"*.

3. *Factors Affecting Common Scanpaths:* We also would like to investigate whether or not the task, gender and familiarity factors can affect common scanpaths identified by eMINE scanpath algorithm. Therefore, we focus on *"Can the task, gender and familiarity factors cause differences in common scanpaths?"*

4.1 Equipment

Participants sat in front of a 17" monitor with a built-in TOBII T60 eye tracker with screen resolution 1280 x 1024. The web pages were on a HP ELiteBook 8530p laptop and these web pages were shown to the participants using the eye tracker's screen. Tobii Studio eye gaze analysis software was used to record the data. Eyetracking data was also stored on that laptop, too. The collected eyetracking data were analysed on a 17" monitor with the screen resolution 1280 x 1024.

4.2 Materials

Six web pages were randomly selected from a group of pages that were used in our previous study [2, 1]. These web pages had been chosen from the top websites listed by ALEXA.com [d]. That study focused on evaluating the extended and improved version of the VIPS algorithm and to have continuity in our studies we used same set of pages. These web pages were categorised based on their visual complexity — low, medium and high [2, 1]. The ViCRAM tool had been used to determine the visual complexity of the web pages which are correlated with users' perception [17]. Two web pages were chosen randomly from each level of complexity for our study. These pages with their complexity levels are as follows: Apple (Low), Babylon (Low), AVG (Medium), Yahoo (Medium), Godaddy (High) and BBC (High).

The extended and improved VIPS algorithm segments web pages based on a granularity level where the higher levels cause segmenting web pages into smaller AoIs [2, 1]. Since the 5th segmentation granularity level was found as the most successful level with approximately 74 % user satisfaction, we decided to use the 5th level for our experiments [2, 1]. The segmented web pages can be seen in Fig. 4-9.

4.3 Procedure

The participants first read the information sheet that is about the study, what happens to the data collected and their rights. When they were happy with the information, they signed the consent form. After that, they filled in the short questionnaire which was for the purpose of collecting basic demographic information of participants — gender, age groups and education level. The participants were also asked to rank their web page usage for the six web pages with 1 (Daily), 2 (Weekly), 3 (Monthly), 4 (Less than once a month) or 5 (Never).

The participants then sat in front of the eye tracker which calibrated to their gaze. They then viewed all of the six web pages twice, one view for searching (maximum 120 seconds) and one view for browsing in a random order. For browsing tasks, the participants were given 30

[d] `http://www.alexa.com`

Apple	
Browsing	1. Can you scan the web page if you find something interesting for you?
Searching	1. Can you locate a link which allows watching the TV ads relating to iPad mini? 2. Can you locate a link labelled iPad on the main menu?
Babylon	
Browsing	1. Can you scan the web page if you find something interesting for you?
Searching	1. Can you locate a link you can download the free version of Babylon? 2. Can you find and read the names of other products of Babylon?
Yahoo	
Browsing	1. Can you scan the web page if you find something interesting for you?
Searching	1. Can you read the titles of the main headlines which have smaller images? 2. Can you read the first item under News title?
AVG	
Browsing	1. Can you scan the web page if you find something interesting for you?
Searching	1. Can you locate a link which you can download a free trial of AVG Internet Security 2013? 2. Can you locate a link which allows you to download AVG Antivirus FREE 2013?
GoDaddy	
Browsing	1. Can you scan the web page if you find something interesting for you?
Searching	1. Can you find a telephone number for technical support and read it? 2. Can you locate a text box where you can search for a new domain?
BBC	
Browsing	1. Can you scan the web page if you find something interesting for you?
Searching	1. Can you read the first item of Sport News? 2. Can you locate the table that shows market data under Business title?

Table 1. Tasks used in the eyetracking study

seconds as used in other studies [13]. The browsing and searching tasks are shown in Table 1. The researcher was responsible to check if the participants complete the tasks successfully and take notes if necessary.

At the end, the participants were asked to redraw three web pages from three different complexity levels.

4.4 User Tasks

User tasks are categorised into two groups for this study: browsing and searching. In the literature, many studies were conducted to categorise user tasks on the web [15]. G. Marchionini's Search Activities Model is one of the most popular models in this field [15]. It consists of three groups which are lookup, learn and investigate [15]. Our searching category is related to fact finding which is associated with the lookup group whereas our browsing category is related to serendipitous browsing which is associated with the investigation group. The tasks which are defined for the six web pages are listed in Table 1.

In this study, half of the participants complete searching tasks first and then complete browsing tasks. Other half completed browsing task first and then completed searching tasks. The reason is to prevent familiarity effects on eye movements which can be caused by the user tasks.

4.5 Participants

The majority of the participants comprised students, along with some academic and administrative staff at Middle East Technical University Northern Cyprus Campus and the University of Manchester. Twenty male and twenty female volunteers participated. One male participant changed his body position during the study, so the eye tracker could not record his eye movements. Another male participant had no successful eye calibration. Unfortunately, these two participants were excluded from the study. Therefore, the eyetracking data of 18 males and 20 females were used to evaluate eMine scanpath algorithm.

All of the participants use the web daily. Most of the participants (18 participants) are aged between 18 and 24 years old, then 25-34 group (14 participants) and 35-54 group (6 participants). Moreover, 14 participants completed their high/secondary schools, 6 participants have a bachelor's degree, 9 participants have a master's degree and 9 participants completed their doctorate degrees.

5 Results

In this section, we present the major findings of this study in terms of the three research questions presented in Section 4.

5.1 Usefulness

"Can eMine scanpath algorithm identify common scanpaths in terms of visual elements of web pages?"

The participants were asked to complete some searching tasks on web pages, therefore we are expecting to see that the common scanpath supports those tasks. We used eMine scanpath algorithm to identify a common scanpath for each of the six web pages. Some participants could not complete the searching tasks successfully and/or had calibration problems. These participants were defined as unsuccessful participants and excluded from the study. The success rates in completing searching tasks are as follows: Apple: 81.58 %, Babylon: 94.74 %, AVG: 94.74 %, Yahoo: 84.21 %, Godaddy: 73.68 % and BBC: 100 %. These values are calculated by dividing the number of the successful participants by the total number of the participants on the page. Table 2 shows the common scanpaths produced by eMine scanpath algorithm for the web pages where 'P' represents the number of successful participants.

Page Name	**P**	**Common Scanpath**
Apple	31	EB
Babylon	36	MPQRS
AVG	36	GI
Yahoo	32	I
Godaddy	28	OM
BBC	38	RN

Table 2. The common scanpaths produced by eMine scanpath algorithm for the web pages

On the Apple web page, 31 out of 38 participants were successful. On this page, the participants were asked to locate a link which allows watching the TV ads relating to iPad mini and then locate a main menu item iPad. EB is identified as a common scanpath for these participants. Since E is associated with the first part and B is related to the second part of the searching task, this common scanpath completely supports the searching task. Fig. 4 shows this common scanpath on the Apple web page.

On the Babylon web page, only 2 participants out of 38 were not successful. On this page, the participants were requested to locate a link which allows downloading a free version of Babylon and then read the names of other products of Babylon. The common scanpath

Fig. 4. Common scanpath on the Apple web page

for the 36 participants was identified as MPQRS shown in Fig. 5. M is related with a free version of Babylon whereas P, Q, R and S are associated with four other products of Babylon. Therefore, the common scanpath thoroughly supports the searching task.

Similar to the Babylon web page, only 2 participants were unsuccessful on the AVG web page. The searching task here was locating a link which allows downloading a free trial of AVG Internet Security 2013 and then locating a link which allows downloading AVG Antivirus FREE 2013. The common scanpath was produced as GI where G has a link to download a free trial of AVG Internet Security 2013 and I contains a link to download AVG Antivirus FREE 2013. Therefore, the common scanpath, shown in Fig. 6, entirely supports the searching task.

For the Yahoo web page, 6 participants could not be successful. The participants required to read the titles of the main headlines which have smaller images and then read the first item under News title. Since only I is produced as a common scanpath on this web page and I contains both parts of the task, the common scanpath nicely supports the searching task, too. Fig. 7 shows this common scanpath.

Since 28 out of 38 participants were successful, 10 participants were excluded for the Godaddy web page. The successful participants read the telephone number for technical support and then located a text box where they can search for a new domain. eMine scanpath algorithm produced OM as a common scanpath shown in Fig. 8. Since M contains the text box and there is no AoI in the scanpath which is related with the telephone number, the common scanpath partially supports the searching task on the Godaddy web page.

On the BBC web page, all participants completed the searching task successfully. The

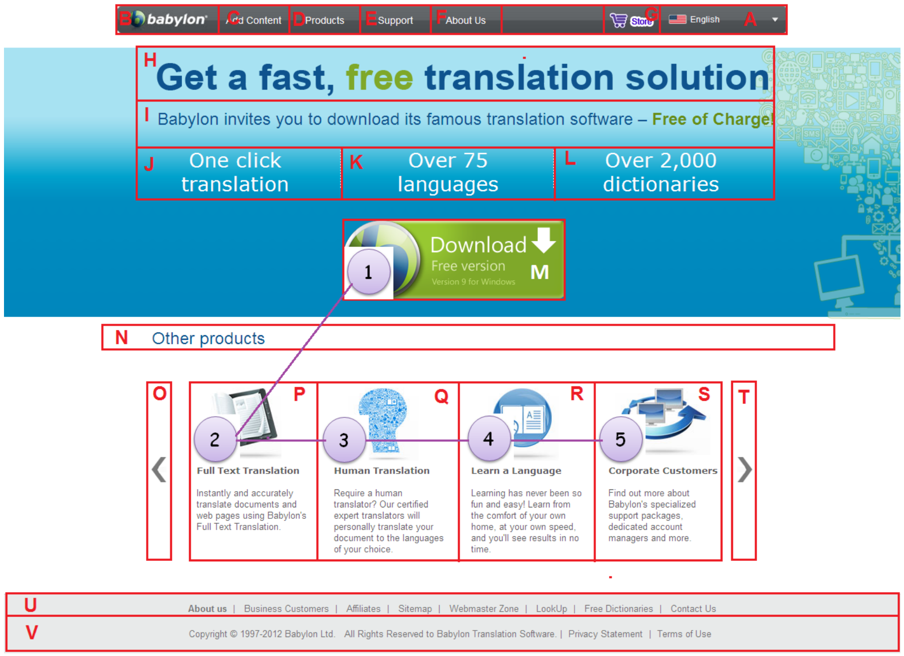

Fig. 5. Common scanpath on the Babylon web page

Fig. 6. Common scanpath on the AVG web page

Fig. 7. Common scanpath on the Yahoo web page

Fig. 8. Common scanpath on the Godaddy web page

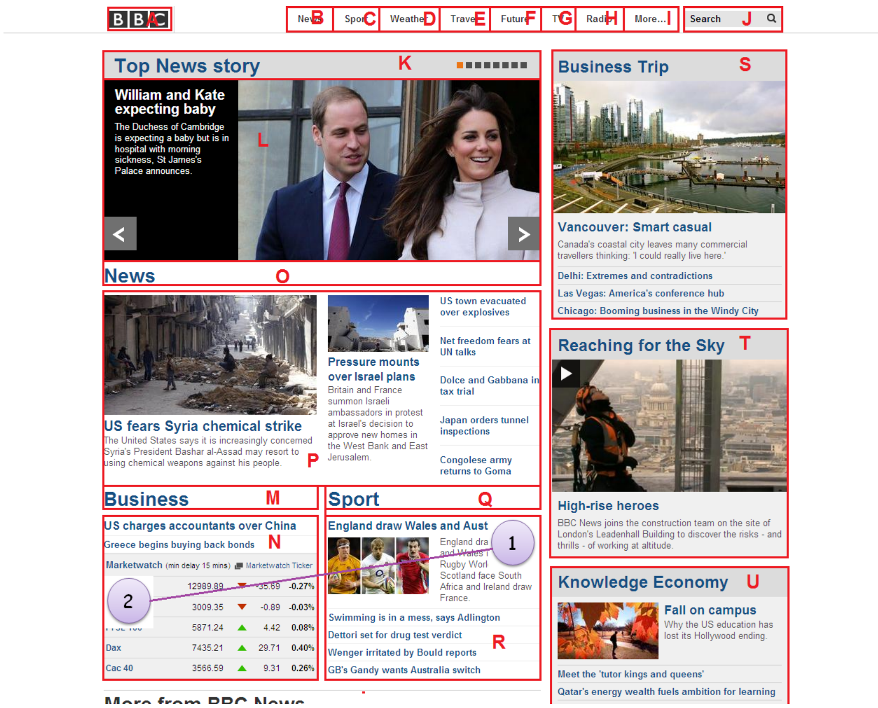

Fig. 9. Common scanpath on the BBC web page

participants were asked to read the first item of the sports news and then locate a table which shows the market data. Therefore, the participants needed to locate R and then N. As the common scanpath RN is produced, it supports the searching task very well. Fig. 9 illustrates this common scanpath on the BBC web page.

To sum up, the common scanpaths on the Apple, Babylon, AVG, Yahoo and BBC web pages completely support the searching tasks whereas the common scanpath on the Godaddy web page partially supports the searching task. Based on these results, we can suggest that our algorithm can identify common scanpaths in terms of visual elements of web pages.

5.2 Scalability

"How does the number of individual scanpaths affect common scanpaths?"

In order to test whether or not eMine scanpath algorithm works well with different numbers of individual scanpaths, we tested the algorithm with different numbers of individual participants. The participants were selected randomly from all of the successful participants. Table 3 illustrates the common scanpaths in terms of AoIs on the different web pages for 10, 20, 30 and 30+ participants while browsing and searching.

In order to see how the common scanpaths are affected when the number of participants increases, we calculated the similarities between the scanpaths which were produced for 10, 20, 30 and 30+ participants. To calculate the similarity between two common scanpaths the String-edit distance between two common scanpaths is divided by the length of the longer common scanpath to have a normalised score [9]. The purpose of a normalised score is to

Task	Page Name	P=10	P=20	P=30	P=30+
Browsing	Apple	IF	F	F	F
	Babylon	MS	M	M	M
	AVG	GIG	G	G	G
	Yahoo	IJI	I	I	I
	Godaddy	O	O	O	O
	BBC	LP	LP	P	——
Searching	Apple	EB	EB	EB	EB
	Babylon	MPQRS	MPQRS	MPQRS	MPQRS
	AVG	IGI	GI	GI	GI
	Yahoo	I	I	I	I
	Godaddy	OM	OM	-	-
	BBC	LPRN	RN	RN	RN

Table 3. The common scanpaths on the different web pages for 10, 20, 30 and 30+ participants while browsing and searching where '-' means that there was no sufficient number of successful participants and '——' means that no common scanpath was detected.

prevent any inconsistencies in similarities caused by different lengths [9, 7]. Finally, the normalised score is subtracted from 1 [9]. For example, the common scanpath for 10 participants is LPRN and the common scanpath for 20 participants is RN on the BBC web page for the searching task. The String-edit distance is calculated as 2 between two scanpaths. After that, since the length of the longer scanpath (LPRN) is equal to 4, this distance is divided by 4. As a result, the normalised score is equal to 0.5. To calculate the similarity 0.5 is subtracted from 1, so the similarity between the two common scanpaths is equal to 0.5 (50 %). Table 4 shows these similarities between the common scanpaths for the searching task on the BBC web page whereas Table 5 illustrates the similarities between the common scanpaths for the browsing task on the Yahoo web page as examples.

BBC Searching	P = 10	P = 20	P = 30	P = 30+
P = 10	—	50	50	50
P = 20	50	—	100	100
P = 30	50	100	—	100
P = 30+	50	100	100	—

Table 4. The similarities between the common scanpaths on the BBC web page for 10, 20, 30 and 30+ participants while searching

Yahoo Searching	P = 10	P = 20	P = 30	P = 30+
P = 10	—	33.33	33.33	33.33
P = 20	33.33	—	100	100
P = 30	33.33	100	—	100
P = 30+	33.33	100	100	—

Table 5. The similarities between the common scanpaths on the Yahoo web page for 10, 20, 30 and 30+ participants while browsing

For both the browsing and searching tasks, we calculated the average similarity between the common scanpaths on each web page. To calculate these average similarities we divided the sum of the similarities between the scanpaths for 10, 20, 30 and 30+ participants by the total number of the similarities. In addition, we calculated the average similarity for both the browsing and searching tasks. Since each web page typically has four scanpaths (for 10, 20, 30 and 30+ participants), we determined their weights based on the number of scanpaths. All of the pages' weights are set to 4, except the Godaddy page because of the searching task. The Godaddy page has one common scanpath for 10 participants and one common scanpath for 20 participants, therefore its weight is set to 2. When the average is calculated, we multiplied the value with its weight to find the weighted value. After that, we found the sum of the weighted value and divided it by the sum of the weights. It was found that the average similarity for the searching tasks (92.42 %) is higher than the average similarity for the browsing tasks (69.44 %). If the number of individual scanpaths does not affect common scanpaths, we should have had 100 % similarity for the browsing and searching tasks. Although we do not have 100 % similarity, we cannot conclude that the number of individual scanpaths completely affect common scanpaths because we have 92.42 % similarity for the searching tasks and 69.44 % for the browsing tasks.

Page Name	**Task**	**Average Similarity for Each Page**
Apple	Browsing	75
Babylon	Browsing	75
AVG	Browsing	66.65
Yahoo	Browsing	66.65
Godaddy	Browsing	100
BBC	Browsing	33.33
Average Similarity for the 6 Pages	Browsing	69.44
Apple	Searching	100
Babylon	Searching	100
AVG	Searching	83.3
Yahoo	Searching	100
Godaddy	Searching	100
BBC	Searching	75
Average Similarity for the 6 Pages	Searching	92.42

Table 6. The average of the similarities between the common scanpaths on each web page for 10, 20, 30 and 30+ participants

5.3 Factors Affecting Common Scanpaths

"Can the task, gender and familiarity factors cause differences in common scanpaths?"

In order to investigate the effects of the task, gender and familiarity factors, we first divided the participants into different groups and applied eMINE scanpath algorithm to each group and then compare the common scanpaths for those groups. We were then able to investigate whether or not these factors affect common scanpaths identified by eMINE scanpath algorithm. As presented below, we can suggest that these factors can cause some differences in common scanpaths.

5.3.1 Task Factor

The participants were asked to complete two kinds of tasks which were the browsing and searching tasks (See Section 4.4). For the browsing tasks, the participants freely browsed the web pages and in the searching task, they needed to locate some specific parts of web pages.

The common scanpaths of all of the successful participants on the six pages for the browsing and searching tasks can be seen in Table 7. This table also shows the similarities between the common scanpaths of the browsing and searching tasks. These similarities suggest that the task can cause some differences in common scanpaths. In particular, the common scanpath of the browsing task (EB) is completely different from the common scanpath of the searching task (F) on the Apple page. In contrast, the common scanpaths of the browsing and searching tasks are the same on the Yahoo page. As can also be seen from Table 7, eMINE scanpath algorithm typically identifies longer common scanpaths for the searching tasks — the average length for the browsing tasks is equal to 0.83 (Standard Deviation SD: 0.41) whereas the average length for the searching tasks is 2.33 (SD: 1.37).

Page Name	**Common Scanpath for the Browsing Task**	**Common Scanpath for the Searching Task**	**Similarity**
Apple	F	EB	0 %
Babylon	M	MPQRS	20 %
AVG	G	GI	50 %
Yahoo	I	I	100 %
Godaddy	O	OM	50 %
BBC	-	RN	0 %

Table 7. The common scanpaths on the different web pages for the browsing and searching tasks and the similarities between the common scanpaths that were identified for the browsing and searching tasks for the same web page

The lengths of the common scanpaths for the searching tasks varies between the web pages. For example, the length of the common scanpath on the Babylon page is 5 whereas the length of the common scanpath of the Apple page is 2. However, these lengths are related to the minimum number of the AoIs that are needed to complete the searching tasks on the web pages. For instance, the participants needed to locate at least 5 AoIs (M, P, Q, R and S) on the Babylon page and at least 2 AoIs (E and B) on the Apple page to successfully complete their searching tasks (See Table 8).

Page Name	**The minimum number of AoIs to complete the searching task successfully**	**The length of the common scanpath produced by eMINE scanpath algorithm**
Apple	2	2
Babylon	5	5
AVG	2	2
Yahoo	1	1
Godaddy	2	2
BBC	2	2

Table 8. The minimum number of AoIs to complete the searching task successfully on the pages and the length of the common scanpath produced for the web pages by eMINE scanpath algorithm

5.3.2 Gender Factor

In order to investigate the effects of the gender on common scanpaths identified by eMINE scanpath algorithm, we first split eyetracking data into a male group and a female group. We then divide the data based on the type of the tasks. After that, we used eMINE scanpath algorithm to identify common scanpaths for the male and female groups for each web page for the browsing and searching tasks. The same number of males and females were used for each web page and each task to prevent any possible effects of the number of the participants. The number differs from one web page to another because we excluded the unsuccessful participants (See Section 5.1). We also randomly excluded some successful participants to make the number of the participants equal for the male and female groups.

Task	Page Name	The number of Participants	Male Group	Female Group	Similarity
Browsing	Apple	18	F	F	100 %
	Babylon	18	MS	MR	50 %
	AVG	18	G	GIG	33.33 %
	Yahoo	18	I	IJI	33.33 %
	Godaddy	18	O	O	100 %
	BBC	18	LT	PS	0 %
				Average Similarity	**52.78 %**
Searching	Apple	15	FEB	CEB	66.67 %
	Babylon	16	MPQRS	MPQRS	100 %
	AVG	18	GI	GI	100 %
	Yahoo	14	I	I	100 %
	Godaddy	14	OM	OM	100 %
	BBC	18	LPN	RN	33.33 %
				Average Similarity	**83.33 %**

Table 9. The common scanpaths on the different web pages for the male and female groups for the browsing and searching tasks and the similarities between the common scanpaths that were identified for the male and female groups on the same web page for the same task

Table 9 shows the common scanpaths of the male and female groups on the six web pages for the browsing and searching tasks. It also illustrates the similarities between the common scanpaths of the male and female groups on the same web page for the same task. For instance, the similarity between the common scanpaths of the male and female groups for the searching task is equal to 100 % on the Babylon page. Based on the data in this table, the average similarity for the browsing tasks is equal to 52.78 % (SD: 36.54) whereas the average similarity for the searching tasks is equal to 83.33 % (SD: 25.46). These similarities suggest that the gender can cause some differences in common scanpaths, especially for the browsing tasks.

The results also suggest that eMINE scanpath algorithm is likely to produce longer common scanpaths for the female group for the browsing task — the female group's average length is equal to 2.00 (SD: 0.89) whereas the male group's average length is equal to 1.33 (SD: 0.52) for the browsing task. However, the lengths of the common scanpaths of the male and female groups on the same page for the searching task are the same, except on the BBC page. On the BBC page, the length of the male group's common scanpath is equal to 3 whereas the length of the female group's common scanpath is equal to 2.

5.3.3 Familiarity Factor

We measured the familiarity based on the user rankings. We asked the participants to rank the web pages according to their usage as 1 (Daily), 2 (Weekly), 3 (Monthly), 4 (Less than once a month) or 5 (Never) (See Section 4.3). When the participants ranked the web page with 1, 2 or 3, it means that they visit the page at least once in a month. Therefore, they are aware of the web page. However, when the participants ranked the web page with 4 or 5, it means that they might visit the web page very few times or they have not visited the web page yet. If the participants ranked the web page with 1, 2 or 3, we assumed that they are familiar with the page. Otherwise, they are not familiar. Unfortunately, we were able to use the Apple, Yahoo and BBC web pages to investigate the effects of the familiarity because there is only a few participants who were familiar with the Babylon (3 participants), AVG (3 participants) and Godaddy (1 participant) web pages.

Similar to the investigation of the gender factor, we first split eyetracking data into a familiar group and a unfamiliar group. We then divide the data based on the type of the tasks. After that, we used eMINE scanpath algorithm to identify common scanpaths for the familiar and unfamiliar groups for each web page for the browsing and searching tasks. We ensured that there were the same number of participants in the familiar and unfamiliar groups for each web page and for each task.

Task	Page Name	The number of Participants	Familiar Group	Unfamiliar Group	Similarity
Browse	Apple	11	FI	FC	50 %
	Yahoo	16	IJI	I	33.33 %
	BBC	16	LPT	LP	66.67 %
				Average Similarity	**50 %**
Search	Apple	10	FEB	EIB	33.33 %
	Yahoo	12	I	I	100 %
	BBC	16	PRN	LRN	66.67 %
				Average Similarity	**66.67 %**

Table 10. The common scanpaths on the different web pages for the familiar and unfamiliar groups for the browsing and searching tasks and the similarities between the common scanpaths that were identified for the familiar and unfamiliar groups on the same web page for the same task

Table 10 shows the common scanpaths of the familiar and unfamiliar groups on the three web pages for the browsing and searching tasks. It also shows the similarities between the common scanpaths of the familiar and unfamiliar groups on the same web page for the same task. For instance, the similarity between the common scanpaths of the familiar and unfamiliar groups for the browsing task is equal to 50 % on the Apple page. Based on the data given in this table, the average similarity for the browsing tasks is equal to 50 % (SD: 16.67) whereas the average similarity for the searching tasks is equal to 66.67 % (SD: 33.34). Hence, we can see some differences in common scanpaths, slightly more for the browsing tasks, caused by the familiarity factor.

Based on the results on these three web pages, we can also report that eMINE scanpath algorithm tends to identify longer common scanpaths for the familiar group for the browsing task — the familiar group's average length is equal to 2.67 (SD: 0.58) whereas the unfamiliar group's average length is equal to 1.67 (SD: 0.58) for the browsing task. In contrast, the lengths of the common scanpaths of the familiar and unfamiliar groups on the same page for the searching task are the same.

6 Discussion

The eMine scanpath algorithm was experimentally evaluated with an eyetracking study and this study illustrates that the algorithm is able to successfully identify common scanpaths in terms of visual elements of web pages and it is fairly scalable.

The searching tasks completed by the participants on the given pages were used to evaluate the usefulness of eMine scanpath algorithm. We expected that the common scanpaths should support these searching tasks. For instance, on the Babylon web page, the participants were asked to locate the link which allows downloading the free version of Babylon (related to AoI M) and then read the names of other products of Babylon (related to AoIs P, Q, R and S).

Therefore, we expected that the common scanpath on the Babylon web page should involve at least MPQRS for the searching tasks.

The results in Section 5.1 show that the common scanpaths produced by eMine scanpath algorithm completely support these tasks, except the common scanpath on the Godaddy page. On that page, the participants were asked to read a telephone number for technical support and locate the text box where they can search for a new domain. The common scanpath involves the AoI for the text box but does not include the AoI for the telephone number. Thus, it partially supports the searching task. There may be various reasons: (1) The participants might make a very few fixations on that AoI (2) Some participants might find the telephone number directly whereas some of them looked at many AoIs to find the telephone number. Therefore, it would be good to pre-process eyetracking data in depth to investigate the individual differences and their reasons.

Some other methods could also be used to evaluate the usefulness of eMine scanpath algorithm. One might consider calculating the similarities between the individual scanpaths and the common scanpath. Besides, the AoIs appeared in all individual scanpaths might be detected and then one part of the evaluation process could be done by using these AoIs.

The scalability of eMine scanpath algorithm was tested by using the different numbers of individual scanpaths as mentioned in Section 5.2. As expected, we can see that the algorithm is more scalable with the searching tasks because the participants were asked to complete some specific searching tasks. The average similarity is equal to 92.42 % between the common scanpaths which were produced with the different number of scanpaths for the searching tasks. However, the average similarity is equal to 69.44 % for the browsing tasks. Based on these values we can suggest that our algorithm is fairly scalable, especially in searching tasks.

There are some differences between scanpaths, such as producing LPRN for 10 participants and RN for 30+ participants on the BBC page. It is caused by using the hierarchical structure. As mentioned in Section 3, eMine scanpath algorithm uses a hierarchical structure while identifying common scanpaths. It selects the two most similar scanpaths from the list and finds their longest common subsequence. It is iteratively repeated until a single scanpath left. Because of the hierarchical structure, some information in intermediate levels can be lost because of combining two scanpaths.

Assume that there are three sequences: S1: GATACCAT S2: CTAAAGTC and S3: GCTATTGCG [6]. S1 and S2 can be aligned firstly and then S1'= - - A - A - - A - - - can be obtained [6]. Following this, S1' and S3 can be aligned and then S3'= - - - A - - - - - - - - can be obtained [6]. This example clearly illustrates that the hierarchical structure can make the method reductionist. Here, all of the three scanpaths have G and T in different locations but G and T do not exist at the end. This may cause some differences in common scanpaths. Because of this reason, eMine scanpath algorithm was not able to identify any common scanpath on the BBC page for the browsing task. When a number of individual scanpaths is increased, the different most similar scanpath pairs can be generated and this may affect common scanpaths. Although eMINE scanpath algorithm has some drawbacks because of the hierarchical structure, it still partly addresses the reductionist problem of the other existing approaches (See Section 2). To address the drawbacks of using the hierarchical structure a constraint might be created to prevent losing the AoIs appeared in all individual scanpaths in intermediate levels. Alternatively, some statistical approaches can be used to sort these AoIs and then create a common scanpath for multiple people.

We also investigated whether or not the task causes differences in common scanpaths. The results in Section 5.3 illustrates that eMINE scanpath algorithm typically provides longer common scanpaths for the searching tasks. This can be caused by the variance in the individual scanpaths in the browsing tasks. Since the participants freely browsed the web pages during their browsing tasks, it is expected to have higher variance in their scanpaths. When the variance in the individual scanpaths increases, it becomes difficult to identify a common scanpath. This is the main reason of why shorter common scanpaths are identified by eMINE scanpath algorithm for the browsing tasks.

The effects of the gender and familiarity factors on common scanpaths were also investigated. The results in Section 5.3 illustrates that the gender and familiarity factors are likely to cause some differences in common scanpaths identified by eMINE scanpath algorithm, especially for the browsing tasks. In particular, the average similarity between the common scanpaths of the male and the female groups is 52.78 % for the browsing tasks whereas it is 83.33 % for the searching tasks. Since the participants were asked to complete some specific tasks on web pages, it is expected to have higher similarities between the common scanpaths for the searching tasks. However, there is still a chance to have different lengths (See Table 9 - Searching on the BBC Page) or different AoIs (See Table 9 - Searching on the Apple Page) for the common scanpaths of the male and female groups. The effects of these factors can be considered to make web pages more usable. For example, when web designers target females, common scanpaths of females can help them to design their web pages.

Although we evaluated our eMINE scanpath algorithm and investigated the effects of the task, gender and familiarity factors on common scanpaths with an eyetracking study, this study still has some limitations. Firstly, we conducted our eyetracking study with 40 participants. Even though there were some academic and administrative staff at the universities, the majority of these participants were students (See Section 4.5). Therefore, further studies can be conducted with different user groups. Besides this, this study was based on six web pages (See Section 4.2) and therefore further studies can be conducted with a larger set of web pages.

7 Concluding Remarks and Future Work

This paper presents an algorithm and its evaluation that identifies common scanpaths in terms of visual elements of web pages. These visual elements are first automatically generated with the extended and improved version of the VIPS algorithm [2, 1]. Eyetracking data is then related to these visual elements and individual scanpaths are created in terms of these visual elements. This algorithm then uses these individual scanpaths and generates a common scanpath in terms of these visual elements. This common scanpath can be used for reengineering web pages to improve the user experience in constrainted environments. In particular, the visual element that is not included in the common scanpath can be removed and a link can be provided to access that visual element. Moreover, the visual elements can be re-ordered based on the common scanpath. Furthermore, this algorithm can be beneficial for web designers to understand how users interact with web pages.

To our knowledge, there is no work on correlating scanpaths with visual elements of web pages and the underlying source code, and this work is novel from that perspective [2, 1].

This paper also shows how the usefulness and scalability of eMine scanpath algorithm was demonstrated with an eyetracking study. The results clearly show that this algorithm is able to identify common scanpaths in terms of visual elements of web pages and it is fairly stable. This algorithm aims to address the reductionist problem that the other existing work has, but the results show that there is still room for improvement.

The eyetracking study also suggests some directions for future work. It indicates that the individual differences can affect the identification of patterns in eyetracking scanpaths. Thus, eyetracking data should be pre-processed to investigate the individual differences and their reasons. Since an eye tracker collects a large amount of data, pre-processing is also required to eliminate noisy data. It is important because noisy data are likely to decrease the commonality in scanpaths. Another benefit of pre-processing is to identify outliers which are potential to decrease the commonality, too.

As with the existing scanpath methods, eMine scanpath algorithm also tends to ignore the complexities of the underlying cognitive processes. However, when people follow a path to complete their tasks on web pages, there may be some reasons that affect their decisions. Underlying cognitive processes can be taken into account while identifying common scanpaths.

In this study, we investigated the effects of the task, gender and familiarity factors on common scanpaths, but in future studies some other factors can also be investigated. For instance, web users are also from different age groups and have different educational backgrounds. Hence, the age group and educational background factors can also be investigated to see whether or not these factors affect common scanpaths. Furthermore, the investigation of the cultural effects can also be interesting contribution.

Finally, even though there are some studies that show that reengineering web pages improve users' experience [5], as part of our future work we are still planning to conduct a user study to evaluate the effectiveness of using common scanpaths to reengineer web pages in the long run.

Acknowledgements

The project is supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) with the grant number 109E251. As such the authors would like to thank to TÜBİTAK for their continued support. We would also like to thank our participants for their time and effort. Finally, we would like to thank to Dr. Simon Harper for his contribution on the earlier version of this paper.

References

1. Elgin Akpinar and Yeliz Yesilada. Heuristic Role Detection of Visual Elements of Web Pages. In Florian Daniel, Peter Dolog, and Qing Li, editors, *Web Engineering*, volume 7977 of *Lecture Notes in Computer Science*, pages 123–131. Springer Berlin Heidelberg, 2013.
2. Elgin Akpinar and Yeliz Yesilada. Vision Based Page Segmentation Algorithm: Extended and Perceived Success. In *Sheng Q. Z. and Kjeldskov J. (eds.) Current Trends in Web Engineering - ICWE 2013 Workshops and PhD Symposium.*, volume 8295 of *Lecture Notes in Computer Science*, pages 238–252. Springer, 2013.
3. Stephan A. Brandt and Lawrence W. Stark. Spontaneous Eye Movements During Visual Imagery Reflect the Content of the Visual Scene. *J. Cognitive Neuroscience*, 9(1):27–38, January 1997.

4. Andy Brown, Caroline Jay, and Simon Harper. Audio access to calendars. In *W4A '10*, pages 1–10, New York, NY, USA, 2010. ACM.
5. Andy Brown, Caroline Jay, and Simon Harper. Using qualitative eye-tracking data to inform audio presentation of dynamic web content. *New Rev. Hypermedia Multimedia*, 16(3):281–301, December 2010.
6. Chung Han Chiang. A Genetic Algorithm for the Longest Common Subsequence of Multiple Sequences. Master's thesis, National Sun Yat-sen University, 2009.
7. Filipe Cristino, Sebastiaan Mathot, Jan Theeuwes, and Iain D. Gilchrist. Scanmatch: a novel method for comparing fixation sequences. *Behavior Research Methods*, 42(3):692–700, 2010.
8. Sukru Eraslan, Yeliz Yesilada, and Simon Harper. Identifying patterns in eyetracking scanpaths in terms of visual elements of web pages. In Sven Casteleyn, Gustavo Rossi, and Marco Winckler, editors, *Web Engineering*, volume 8541 of *Lecture Notes in Computer Science*, pages 163–180. Springer International Publishing, 2014.
9. Tom Foulsham and Geoffrey Underwood. What can Saliency Models Predict about Eye Movements? Spatial and Sequential Aspects of Fixations during Encoding and Recognition. *Journal of Vision*, 8(2):1–17, 2008.
10. Joseph H. Goldberg and Jonathan I. Helfman. Scanpath clustering and aggregation. In *ETRA '10*, pages 227–234, NY, USA, 2010. ACM.
11. Helene Hembrooke, Matt Feusner, and Geri Gay. Averaging scan patterns and what they can tell us. In *ETRA '06*, pages 41–41, New York, NY, USA, 2006. ACM.
12. Halszka Jarodzka, Kenneth Holmqvist, and Marcus Nyström. A Vector-based, Multidimensional Scanpath Similarity Measure. In *ETRA '10*, ETRA '10, pages 211–218, New York, NY, USA, 2010. ACM.
13. Caroline Jay and Andy Brown. User Review Document: Results of Initial Sighted and Visually Disabled User Investigations. Technical report, University of Manchester, 2008.
14. Sheree Josephson and Michael E. Holmes. Visual attention to repeated internet images: testing the scanpath theory on the world wide web. In *ETRA '02*, pages 43–49, NY, USA, 2002. ACM.
15. Gary Marchionini. Exploratory Search: From Finding to Understanding. *Commun. ACM*, 49(4):41–46, April 2006.
16. Marcus Mast and Michael Burmester. Exposing repetitive scanning in eye movement sequences with t-pattern detection. In *IADIS IHCI '11*, pages 137–145, Rome, Italy, 2011.
17. Eleni Michailidou. *ViCRAM: Visual Complexity Rankings and Accessibility Metrics*. PhD thesis, University of Manchester, 2010.
18. Alex Poole and Linden J. Ball. Eye tracking in human-computer interaction and usability research: Current status and future. In *Prospects, Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc*, 2005.
19. Kari-Jouko Raiha. Some applications of string algorithms in human-computer interaction. In Tapio Elomaa, Heikki Mannila, and Pekka Orponen, editors, *Algorithms and Applications*, volume 6060 of *Lecture Notes in Computer Science*, pages 196–209. Springer Berlin Heidelberg, 2010.
20. Anthony Santella and Doug DeCarlo. Robust clustering of eye movement recordings for quantification of visual interest. In *ETRA '04*, ETRA '04, pages 27–34, New York, NY, USA, 2004. ACM.
21. Haruhiko Takeuchi and Yoshiko Habuchi. A quantitative method for analyzing scan path data obtained by eye tracker. In *CIDM '07*, pages 283 –286, 1 2007-april 5 2007.
22. W3C WAI Research and Development Working Group (RDWG). Research Report on Mobile Web Accessibility. In Simon Harper, Peter Thiessen, and Yeliz Yesilada, editors, *W3C WAI Symposium on Mobile Web Accessibility*, W3C WAI Research and Development Working Group (RDWG) Notes. W3C Web Accessibility Initiative (WAI), first public working draft edition, December 2012.
23. Julia M. West, Anne R. Haake, Evelyn P. Rozanski, and Keith S. Karn. eyePatterns: software for identifying patterns and similarities across fixation sequences. In *ETRA '06*, pages 149–154, NY, USA, 2006. ACM.

24. Yeliz Yesilada, Simon Harper, and Sukru Eraslan. Experiential transcoding: An EyeTracking approach. In *W4A '13*, page 30. ACM, 2013.
25. Yeliz Yesilada, Caroline Jay, Robert Stevens, and Simon Harper. Validating the Use and Role of Visual Elements of Web Pages in Navigation with an Eye-tracking Study. In *The 17th international conference on World Wide Web*, WWW '08, pages 11–20, New York, NY, USA, 2008. ACM.
26. Yeliz Yesilada, Robert Stevens, Simon Harper, and Carole Goble. Evaluating DANTE: Semantic Transcoding for Visually Disabled Users. *ACM Trans. Comput.-Hum. Interact.*, 14(3):14, September 2007.