

# eSeeTrack –Visualizing Sequential Fixation Patterns

Hoi Ying Tsang, Melanie Tory, and Colin Swindells, University of Victoria

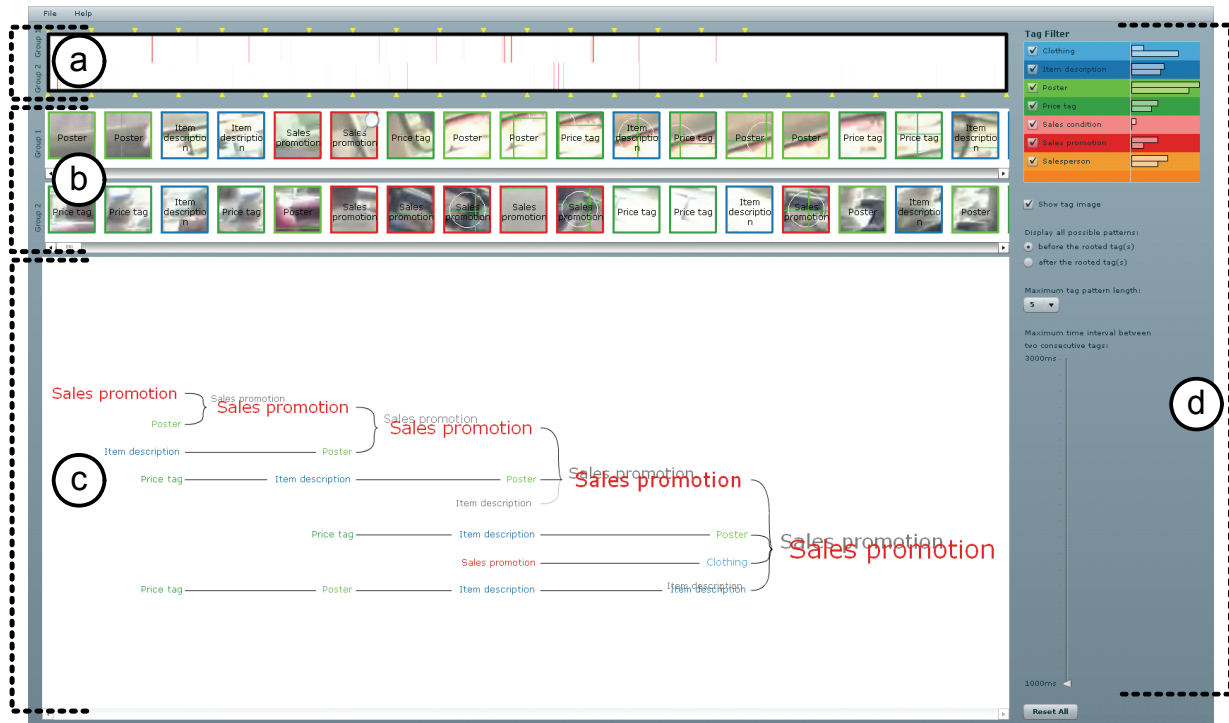


Fig. 1. Overview of eSeeTrack. a) Timeline section; b) Detailed timeline section with optional thumbnail images of fixated objects displayed; c) Tree visualization; d) Control section. Two retail stores belonging to the same chain are explored and compared; the “sales promotion” tag has been selected and the tree visualization shows the observed fixation orderings that end with “sales promotion”. In the tree visualization, labels with colors other than grey belong to store 1 and grey labels belong to store 2. Font sizes in the tree as well as the bar charts on the top of the control section reveal that sales promotion had more fixations in store 1. The most common fixation sequence (sales promotion 5 times) indicates that people tend to look at the sales promotion for an extended period of time.

**Abstract**—We introduce eSeeTrack, an eye-tracking visualization prototype that facilitates exploration and comparison of sequential gaze orderings in a static or a dynamic scene. It extends current eye-tracking data visualizations by extracting patterns of sequential gaze orderings, displaying these patterns in a way that does not depend on the number of fixations on a scene, and enabling users to compare patterns from two or more sets of eye-gaze data. Extracting such patterns was very difficult with previous visualization techniques. eSeeTrack combines a timeline and a tree-structured visual representation to embody three aspects of eye-tracking data that users are interested in: duration, frequency and orderings of fixations. We demonstrate the usefulness of eSeeTrack via two case studies on surgical simulation and retail store chain data. We found that eSeeTrack allows ordering of fixations to be rapidly queried, explored and compared. Furthermore, our tool provides an effective and efficient mechanism to determine pattern outliers. This approach can be effective for behavior analysis in a variety of domains that are described at the end of this paper.

**Index Terms**—H.5.1 [User / Machine Systems]: Human factors, H.5.2 [Information Interfaces and Presentation]: User Interfaces — GUI

## 1 INTRODUCTION

In many different domains, researchers use eye-tracking to explore where people look. Examples include cognitive science, psychology, market research, product design, human computer interaction (HCI), and sport training. Typically, their research focuses on understanding a participant’s thinking and behavior. For instance, in the study of scene perception, psychologists examine eye movements and fixation patterns in order to understand the acquisition, processing, and storage of visual information and different contexts [4]. In HCI, gaze data is studied to evaluate the usability of an interface (e.g., are users looking at too many places before finding the necessary function?) [1][4][10][15][24]. Hence, as a starting point, researchers

- Hoi Ying Tsang is with VisID Lab at University of Victoria, E-mail: hysang@uvic.ca
- Melanie Tory is with VisID Lab at University of Victoria, E-mail: mtory@cs.uvic.ca
- Colin Swindells is with VisID Lab at University of Victoria and Locarna Systems, Inc. Email: colin@locarna.com

Manuscript received 31 March 2010; accepted 1 August 2010; posted online 24 October 2010; mailed on 16 October 2010.

For information on obtaining reprints of this article, please send email to: [tcvg@computer.org](mailto:tcvg@computer.org).

are often interested in answering the following questions when they study eye movement data:

- What does an observer fixate on?
- When does an observer gaze at each object, and in what order?
- How long or how often does the observer fixate on each object?

The answers to these questions can lead to interesting implications. For example, a high number of fixations on an object likely implies viewer interest. Fixating on an object for a long period of time may imply that the viewer is puzzling over or studying the object. In addition, the order of fixations on different objects gives clues about the viewer's thinking process or strategy to complete a task. Thus, decrypting the ordering of eye movement and the gaze duration allow researchers to guess at the viewer's cognitive process [11][12][15].

Because the eye fixates on objects in different places over time, eye-tracking data varies with both space and time. However, the spatial and temporal attributes do not always need to be shown directly. Depending on the analysis goal, a user may wish to see:

- Number/duration of gazes across a spatial scene.
- Number/duration of gazes on specified regions, objects, or types of objects.
- Temporal ordering of gazes on specified regions, objects, or types of objects.

In most analyses, users want to compare eye-gaze data between groups of participants or experimental conditions. We focus on enabling such comparisons for sequential orderings of fixations, as compared to simple counts on each object type.

For static scenes (e.g., a commercial poster), fixations can be easily summarized using techniques such as heat maps or by counting fixations on user-designated areas of interest. For dynamic scenes, fixations can be tagged (i.e., labelled with the name of the object or activity); then fixations for each tag can be counted and summarized. However, in both static and dynamic cases, identifying patterns in the temporal ordering of fixations is difficult with existing tools.

Our interactive prototype, eSeeTrack, is designed to address this need. Our approach uses a timeline and a tree-structured visualization to provide a summary of serial fixations for exploration and analysis; this is shown in Figure 1 and is explained fully in section 4. It enables users to identify both frequent and infrequent orderings, and to compare orderings between different data sets (e.g., participants or conditions). While our tool is designed to visualize eye-tracking data, we envision that it could also visualize other kinds of sequential items such as event-based data and state transitions.

In this paper, we walk through the design of eSeeTrack and demonstrate its capabilities via case studies in two different applications: surgical simulation and retail store chain data. First, we discuss background information on eye-tracking and previous work on eye-tracking visualizations. We then present eSeeTrack, and discuss the features and the implementation of the system. Next, the two case studies are described and we discuss the efficiency and effectiveness of our approach. We conclude with some potential generalizations and suggestions for future work.

## 2 EYE-TRACKING BACKGROUND

### 2.1 Eye Movements and Eye Trackers

Eye movements have been studied scientifically for more than a century. In the early years, researchers discovered that the eyes move in several different ways. Two types of movement that have often interested scientists are fixations and saccades. A *fixation* happens when an observer gazes at a fixed spot in the world, defined as *point of regard*, for some duration (usually defined as at least 100 to 200 milliseconds) [15]. A *saccade* is the rapid visual transition from one fixation to another. The combination of a series of fixations and

saccades is called a *scanpath*. Due to the high velocities in saccadic movement, which can easily reach 500° per second for a large saccade [11][24], there is almost no visual input to the brain during a saccade. Hence, our visual experience consists of a set of fixations on different objects.

Eye-tracking is the process of recording the point of regard using a video-based device called an eye-tracker. The two most common types of eye trackers are head-mounted and table-mounted. In both types of tracking devices, at least one camera aims at one of the observer's eyes to record the eye motion. An additional camera is installed on the head-mounted eye tracker to film the participant's perspective of the world. Thus, a participant wearing a head-mounted tracking device is allowed to move freely around the world during the study. The table-mounted tracker is ideal when stimuli can be displayed on a computer screen. In this case, the participant must stay in front of the screen and only make minor head movements.

### 2.2 Visual Processing

Past research has confirmed that ocular motion affects tremendously what the viewer sees and links closely to the cognitive goal of the viewer [12][24]. In addition, the gaze sequence of an individual highly depends on the nature of a task performed and varies according to bottom-up or top-down processing [15][24]. In the bottom-up approach, gist or layout of the object in a scene is fixated first. In contrast, in top-down processing, eye movements are directed by prior knowledge and expectation of a scene. Furthermore, according to the mind-eye hypothesis, people are usually thinking about what they are looking at [19]. Thus, eye-gaze patterns give researchers clues about the thoughts of a participant.

## 3 PREVIOUS WORK

### 3.1 Approaches to Analyzing Eye-Tracking Data

A heat map, also known as a fixation map or a saliency map, is the most popular visualization to present eye-tracking data [20][33]. It visualizes relative frequencies or durations of fixations via a semi-transparent color map. Alternative versions of a heat map hide uninteresting parts by darkening or blurring them [25][33]. Heat maps can combine the fixations of multiple participants. Bee swarm is a variation of a heat map that uses dots to indicate the point of regard. The visualization can contain gaze data of multiple participants by using different colored dots. Bee swarm can also be integrated into a video.

Another popular approach is to define areas of interest (AOIs) and count the number of fixations in each AOI, presenting the results as summary statistics or a bar chart. This approach extends to video data by replacing spatial AOIs with tags for items of interest. Both the counting and heat map approaches neglect sequence information (i.e. the order in which items of interest were visited) and therefore only provide a summary of the content viewed.

By contrast, a gaze plot captures the sequence of eye fixations using an ordered series of circles, where circle size represents duration and number indicates order. A gaze plot can visualize eye movements of multiple participants by assigning a different color to each chain of circles. Variations of a gaze plot either arrange the fixated objects in a more abstract way [6][23][9], include additional information about each fixation [17], or miniaturize the plot [9] to enable comparison of small multiples. Gaze plots become quickly cluttered with a large number of fixations, and are limited to static scenes (e.g., a poster or interface with fixed organization). Patterns of fixation orderings are not readily apparent and comparison of such patterns is even more challenging. As a result, numerous other techniques have been developed to analyze, compare, and visualize sequences.

One such technique is the transition diagram [14], which plots AOIs as nodes in a node link graph, with a link between two AOIs if there are one or more saccades directly from one to the other. Frequencies of each transition can be encoded as line thickness or

color. Transition information can alternatively be visualized as a matrix, where cells in the matrix are colored by the frequency of transitions between each pair of AOIs [32]. However, with these methods, it is only possible to examine sequential pairs of fixations (i.e., sequences of length two).

An alternative matrix representation can be used to visualize one complete fixation sequence [10][21]; however, retrieving along sequence stored in a transition matrix is non-trivial since it requires visiting each element in the matrix. In addition, comparison of patterns cannot be performed directly, and frequency of patterns is not apparent.

Recently, string editing [22] has become a popular way to analyze and compare fixation sequences. Objects or AOIs are represented as letters in a string. Two sequences are compared by determining the number of editing operations required to transform one string into the other, usually after the sequences have been aligned. This provides a measure of the distance between two sequences. An unordered comparison is also possible, by comparing the letters but ignoring their order. Furthermore, by clustering results of an unordered comparison, AOIs can be determined automatically [22][5].

Numerous variations and improvements to string editing are described (e.g., [5][32]), and an “average” scan pattern for a group of subjects can be determined using multiple sequence alignment [13]. Alternative algorithms for determining sequence similarity have been proposed, one using dot plots [8], and one using geometric similarity [16]. A statistical test for the similarity of scan patterns has also been described [7].

Results of the above sequence analysis techniques can be visualized in a number of ways. Pairwise distances between sequences can be used as input to a multidimensional scaling (MDS) algorithm to produce a low dimensional embedding. Viewing the MDS plot enables users to see which sequences are most similar to each other [32]. Alternatively, sequences can be hierarchically clustered, with the result visualized as a cluster tree [32] or dendrogram [8].

Sequence analysis approaches analyze global sequences and transition diagrams analyze only sequential pairs. In eSeeTrack, we take an in-between approach by examining short subsequences. We assume that short subsequences may occur repeatedly within a single eye-tracking session and may be similar across users, even if the global order differs. This approach has some similarity to subsequence search in eyePatterns [32]; however, in order to conduct a search, the user must have some initial idea of what to look for. We present a visual technique to enable exploratory analysis of subsequence and their frequencies without requiring any a priori knowledge.

### 3.2 Visualizing Event Sequences

Sequences of fixations are similar to other types of time-ordered events such as state changes or sequential activities. Thus visualizations for these kinds of data may be useful. Smooth Graphs [2] and ActiviTree [29] visualize sequences of states and events using a node-link graph and a tree respectively. Each state or event is represented by a node, with transitions between them represented by edges.

However, these visualizations are insufficient for several reasons. First, ordering is difficult to discern without directional indicators such as arrow heads, which clutter the image. Second, nodes in a graph typically have several incoming and several outgoing edges, making it nearly impossible to determine which outgoing edge followed a given incoming edge without some sort of interaction. This makes it almost impossible to examine sequences of length greater than two, a problem we encountered in a previous eye-tracking analysis (in which such graphs were generated manually) [27]. Smooth graphs alleviate this problem somewhat by using smooth curves. However, individual sequences are still difficult to discern because of path overlap.

Our approach removes the above problem by replicating nodes in different paths, using a structure similar to a WordTree [31]. Unlike ActiviTree and Smooth Graphs, eSeeTrack also supports comparison of event sequences.

## 4 ESEETRACK

We designed eSeeTrack to visualize eye-tracking data based on a dynamic scene, to allow users to see all fixation patterns within a time frame without watching the whole video, and to enable users to compare fixation patterns of multiple groups of participants. As our starting point, we assume that users have automatically extracted fixations from the eye-tracking video and tagged those fixations with up to ten keywords of their choice. Since each fixation has an associated tag, in the following sections, a fixation pattern is referred as a *tag sequence*.

Tagging is used to identify categories of interest to a user (e.g., the types of objects in the scene). This task can be accomplished with commercial eye-tracking software. The level of user effort required to tag the data depends on the type of scene. For a static scene, user involvement is relatively low since the objects remain stationary over the entire eye-tracking session. The user simply needs to predefine the areas of interest (AOIs) and then tagging is automatic. For a dynamic scene, predefining the AOIs is currently impossible since the scene changes continuously. Therefore, each fixation has to be manually tagged with eye-tracking software. This process constitutes a significant amount of time and effort if the number of fixations is large. However, users typically do this tagging already to prepare for other forms of analysis, so no additional effort is required to prepare the data for eSeeTrack.

### 4.1 Overview

eSeeTrack consists of four components as seen in Figure 1: a timeline and detailed timeline on the top (panels a and b), a tree visualization at the bottom (panel c), and a control section on the right side (panel d). The tool allows users to create up to six groups of eye fixations. Each group may contain the fixations of more than one participant. Each group has a corresponding timeline and detailed timeline. The tree visualizations of each group are overlapped.

We encode tags using color so that they can be rapidly identified. We use a qualitative color map from ColorBrewer [3] and limit the number of tags to ten to ensure that the colors are easily distinguishable [30]. The tag filter, which is located at the top of the control section, labels all the tags and their associated color codes (see Figure 1 – panel d). It also serves as a filter and a summary of fixation counts. If a specific tag is unselected, the related bands in the timeline will be greyed and the associated blocks and nodes will be excluded in the detailed timeline and tree visualization. The bar chart displays the relative frequency of each tag for each participant group.

### 4.2 Timeline and Detailed Timeline

A timeline exhibits the chronological sequence of fixations. It contains a series of colored bands of different widths, each representing a fixation. Blank spaces in a timeline denote no fixation. The width of a band denotes the duration of the fixation and the color signifies the associated tag. In Figure 1 – panel a, red bands represent “sales promotion”. Detailed information about a fixation can be accessed via a tooltip, as in Figure 2.

Each timeline is segmented, with one participant per segment. Segments are the same length because the relative pattern of fixations from the start to the end of an eye-tracking session is more important than variations in the time participants took to complete the task. Tiny triangular marks on the top or bottom indicate the separation between participants within a participant group (see Figure 2). When hovering over a triangle, information about the participants and their associated video lengths are displayed.

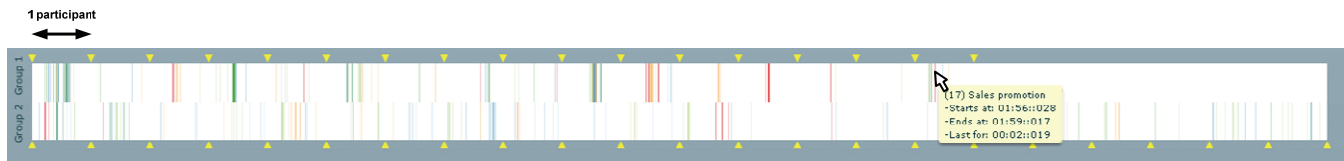


Fig. 2. Two timelines display fixations of two different groups: 16 participants in group 1 (on the top) and 22 participants in group 2 (on the bottom). A tooltip displays detailed information about the 17<sup>th</sup> fixation of group 1's last participant.

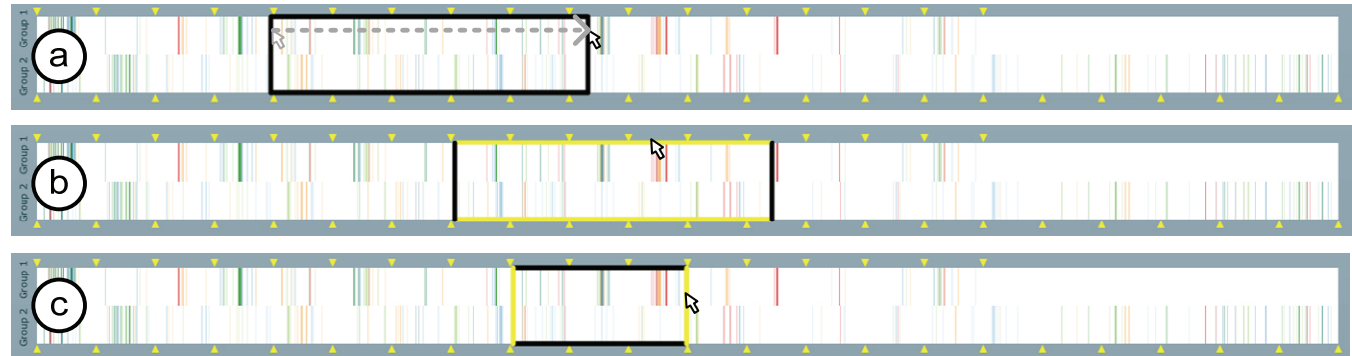


Fig. 3. A time window is used to include the fixations to be analyzed further. a) Create – dragging a box; b) Move – mouse dragging either the top or the bottom edge of the window (shown in yellow); c) Resize – mouse dragging either the left or the right edge of the window (shown in yellow).

Some fixation patterns among different participants can already be seen quickly in the timeline by observing the variation in the colors of the bands. In addition, frequently fixated objects are apparent because of the frequency of their color. In Figure 2, light green bands, which represent “poster”, are the most frequent in both group 1 and 2. The fixation data can be investigated further by creating a time window and viewing its contents in the detailed timeline and tree visualization (see Figure 3). Only fixations within the time window will be shown within the detailed timeline and counted for the bar charts and tree visualization.

A detailed timeline contains a list of square blocks labelled and colored by the fixation's tag as seen in Figure 1 – panel b. Thin bands may be lost in the timeline but are visible in this view. Detailed information about a fixation can be viewed by pressing the magnifying glass on a selected block as in Figure 4. A user can also launch the video to see the exact object that a participant was looking at. Each block can optionally display a thumbnail image of the fixated object as shown in Figure 1 – panel b.

When a block is clicked, bands corresponding to the same tag are highlighted by fading the color of other bands in the overview timeline (see Figure 1 – panel a), allowing users to see when this tag occurs over the entire timeline. In addition, selecting a tag launches a tree-structured visualization rooted at the selected tag. Thus, this tag is called the *root tag*.

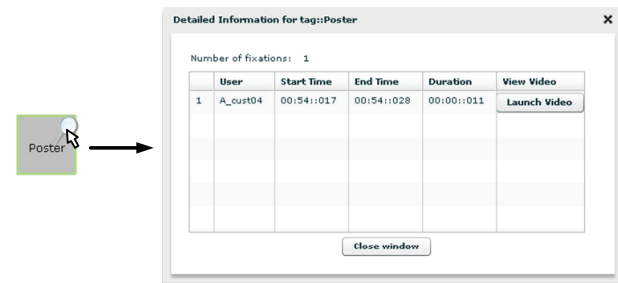


Fig. 4. Launching the information window for a particular fixation.

### 4.3 Tree Visualization

Although the ordering of fixations can be observed from colors in the timeline, it is difficult to identify ordered patterns from this view since they may be lost in a sea of lines. It is also difficult to identify which orders are the most frequent and infrequent, and to compare the patterns of multiple groups. eSeeTrack overcomes these issues by

extracting fixation patterns from the data and using a rooted-tree structure to visualize them.

Similar to a WordTree [31], our tree visualization combines the concept of tag clouds and suffix trees. Each node of the tree denotes a tag. It may include more than one label for a given node; the colored label represents the tag of group 1, the label of other groups is grey and is displayed as the shadow. The root of the tree is the tag that was selected in the detailed timeline section. The largest tree depth represents the maximum allowable length of a fixation pattern, which is adjustable in the control section. With the exception of the root, each node is connected to its parent with a curve to create a smooth visual transition. Each path from the root to a leaf node is a unique observed sequence of fixations.

The tree can be either left-rooted or right-rooted in order to display fixation patterns starting from or ending with the root tag, respectively. This allows a user to answer the questions of “What items does a user fixate on before/after looking at object X?” In Figure 1 – panel c, a right-rooted tree presents all the fixations sequences that end with the “sales promotion” category.

The size of label in a node signifies the relative frequency (or count) of sequences containing the category in its position within the sequences. The fixation count of a node is always larger than or equal to the sum of its children. To help users visually locate common tag sequences, especially for a large tree, the tree is ordered and displays the most frequent path on the top.

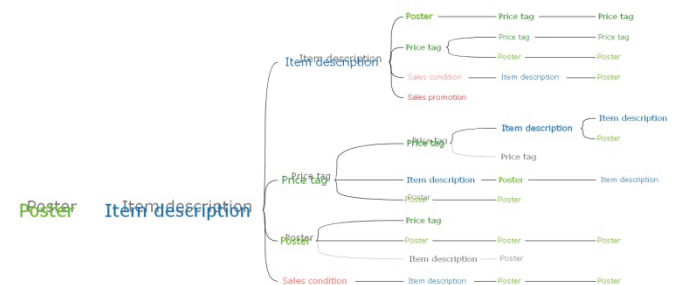


Fig. 5. Tree visualization displays the tag sequences beginning with “poster” followed by “item description”.

By highlighting nodes in the tree visualization, users can relate the selected fixation pattern to the timelines to verify the moment(s) that the participants fixated on this particular sequence of objects; the

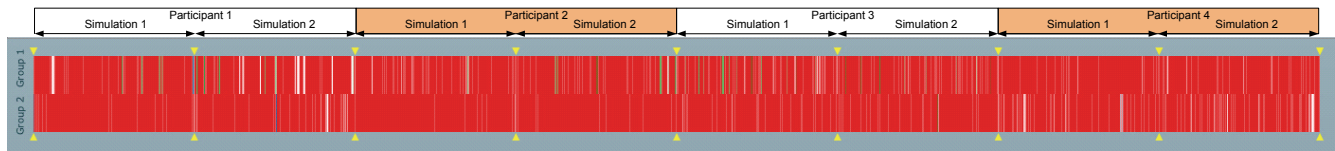


Fig. 6. Timeline section with surgical simulation fixation data; expert surgeons are in group 1 (on the top) while novice users are in group 2 (on the bottom). Bands in red– lap screen, in green – vital screen, and in blue – other.

associated bars in the timelines will be highlighted by fading the colors of other bars. In addition, the tool allows users to create a multiple-roots tree to answer questions such as “What items does a user fixate on before/after fixating on X then Y?” To accomplish this, a user clicks on a node of the tree visualization, which will add the node to the root along with any intervening nodes (see Figure 5). Similar to tags in the detailed timeline section, information about a particular category node can be accessed via the magnifying glass.

In a Word Tree, a sequence is defined as a sentence finishing with an ending punctuation mark (i.e., period, question mark, or exclamation mark). A fixation sequence in our tree visualization is built according to two user-specified criteria: 1) the maximum length of the sequence and 2) the maximum *gap duration*. The gap duration is the time span between two successive fixations. The system searches through all fixations within the time window to find all sequences that begin (or end) with the selected tag as long as the gap duration between successive fixations in the sequence is less than the threshold and the sequence length does not exceed the maximum sequence length. The threshold of the gap duration can be adjusted between 1000 to 3000 milliseconds and the sequence length can be adjusted between 2 to 5 items via widgets in the control section. Sequences of length 5 were chosen to be the upper limit of the maximum allowable length since previous research reported few consistent patterns for sequences longer than 5 [28].

#### 4.4 Implementation

eSeeTrack was developed in Adobe Flex 3.4 under Flex Builder 3 in Windows XP professional SP3. It is a web application that can be operated in browsers supporting Adobe Flash player 10 or later.

The prototype accepts two types of input files: 1) an xml file (called “fixation set file”), which contains a list of fixations made and ordered by participants of the study, and 2) Adobe Flash video file (called “video file”), which includes the scenes seen from the world perspective of a participant during the eye-tracking session plus the points of regard made by the participant. The fixation set file consists of one or more participant logs. While the fixation file is used to analyze the pattern of fixations, the video file is used to extract thumbnail images of the fixated objects and can also be played back from any fixation point.

### 5 CASE STUDIES

eSeeTrack provides a mechanism for users to initially explore eye-tracking data to find and compare possible fixation patterns that they might follow up with a statistical analysis. In order to demonstrate the capabilities of eSeeTrack, we ran two case studies on different data sets. The first data set involves eye-tracking data from a surgical simulation with novice users and expert surgeons. The second set displays eye-tracking data of customers in a chain of retail clothing stores. Both data sets were collected for real eye-tracking studies. For client confidentiality, the second data set was anonymized. The data was collected in a real retail chain that does not sell clothing; the category (tag) names were then remapped to equivalent objects in a clothing store. Since both studies involved dynamic scenes, manual tagging was required before analyzing data with eSeeTrack. It took approximately 6 seconds to tag each fixation, for total tagging times of 190 minutes for Case 1 (1889 fixations) and 130 minutes for case 2 (1303 fixations).

## 5.1 Case 1: Surgical Simulation

### 5.1.1 Background description

Laparoscopic surgery is a minimally invasive procedure in which the surgeon makes only small incisions and uses a camera and light mounted on the end of a tool (the laparoscope) to view the interior surgical site. This view is displayed on a monitor. Due to the instruments’ complicated operation and the high cost of mistakes, physicians require a lot of specialized training. As a result, surgical simulations are used to help novices become adept with the tools. It is known that experts exhibit more systematic eye-gaze patterns than novices during surgical simulation [18][26].

### 5.1.2 Data and Hypothesis

Two groups of four users participated in a surgical simulation study. A complete description of the study may be found in [26]. The first group consisted of expert surgeons and surgical residents, and the other group was made up of novice users. Each participant performed two simulations. Fixations were categorized as being on the lap screen, vital screen, or other objects. A description of these categories can be found in Table 1.

Table 1. Fixated Object Categories in the Surgical Simulation Study.

Tag	Description
Lap screen	Monitor displaying the laparoscope camera view
Vital screen	Display of a patient’s vital signs
Other	Other objects

The experimenters for this study had already completed their analysis before we began examining the data using eSeeTrack. Their analysis consisted of statistical computations as well as qualitative viewing of the eye-tracking videos. Our goal was to see whether our visualization tool could identify verifiable findings and patterns more quickly than these time-consuming forms of analysis. We were initially uninformed about the hypotheses and findings of the experimenters, except that we knew to expect different patterns among novices and experts, and thought that novice users might be more likely to focus on objects other than the lap and vital screens (e.g., their hands). We later verified our findings by comparing them to the experimenters’ hypotheses and analysis.

### 5.1.3 Analysis



Fig. 7. Tag filter and bar chart with surgical simulation data. Top bars represent fixations of experts and bottom bars represent novices.

Fixation patterns of length 5 were created for analysis. By observing the timelines and bar chart in Figures 6 and 7, and the associated summary statistics (Table 2), we can see that more than 90% of fixations are on the lap screen in both groups. The bar chart indicates that the novice users were more likely to look at “other” than expert



	Group			
	① Expert		② Novice	
	Count	%	Count	%
Lap screen	896	90.4	865	97.2
Other	3	0.3	21	2.3
Vital Screen	92	9.3	4	0.4
Total	991	100.0	890	100.0

The tree visualization showed that the most common fixation pattern for both expert and novice users was a sequence of 5 “lap

Some interesting findings emerge in the visualization when analyzing the patterns that begin or end with “vital screen” or “other” (see Figures 8 and 9). Expert surgeons switched back and forth between the “lap screen” and the “vital screen”, while novices looked at the “vital screen” for a while and then the “lap screen”, or vice versa. This implies that the expert surgeons were more conscious of the vital status of the patient and could gain that knowledge through a quick glance. These patterns can be seen in the tree visualization shown in Figure 8 where there are many more paths for experts (color) than for novices (grey shadow). The difference can be most easily observed by looking at the grey shadow sequences (novices). In all of these sequences, the novice begins looking at the vital screen and at some point switches to the lap screen, but never switches back to the vital screen.

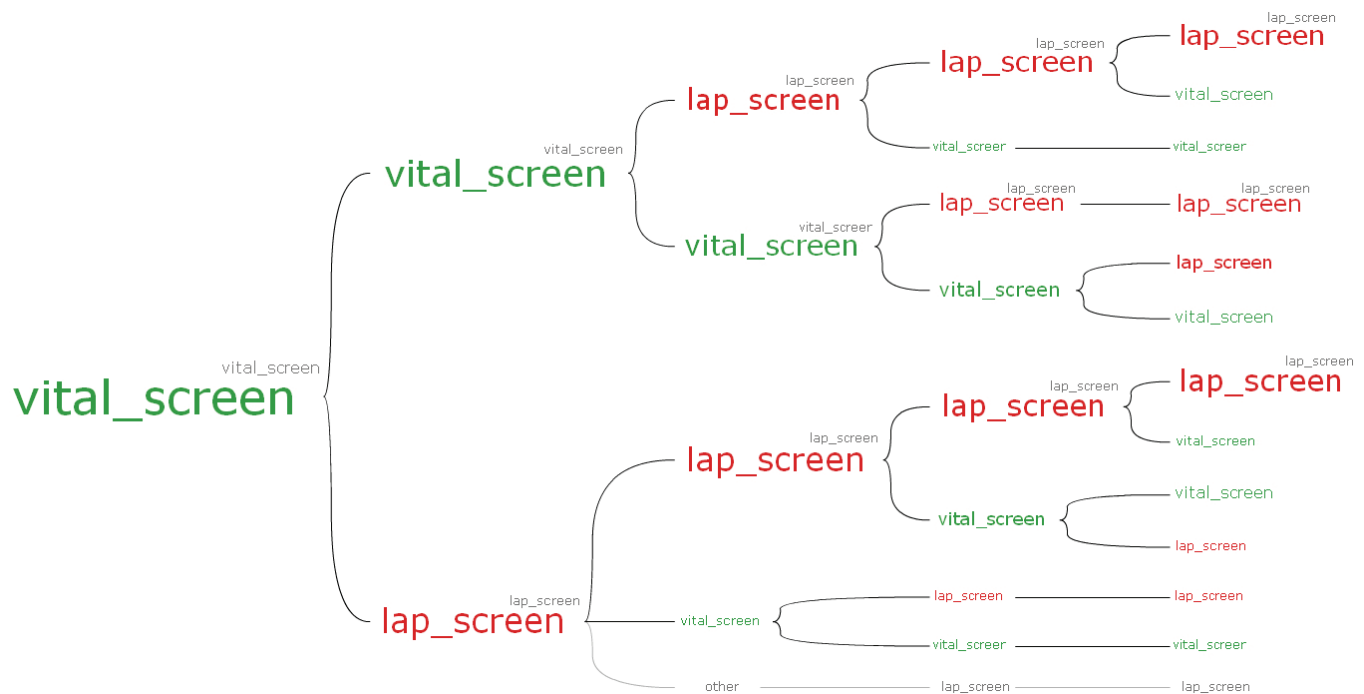


Fig. 9. Tree visualization showing the fixation patterns ending with “other”. Note that experts (shown in red and blue) have only 1 pattern, indicating that “other” was never viewed more than one time in a row. The many patterns in grey shadow belong to novices.

Only 0.3% of fixations (3 out of 991) were “other” in the expert group as compared to 2.3% (21 out of 890) for novices. Although 2.3% does not seem very large, in a surgical situation, this could have a huge impact on the patient. The tree visualization (Figure 9) reveals that once the novice users fixated on “other”, they were more likely to revisit “other” again in subsequent fixations. By contrast, the expert users never looked at “other” more than once in a row. This finding demonstrated that the experts were more task focused with distractions being shorter and less frequent. In addition, we were interested to note that “vital screen” never immediately preceded or followed “other”.

Most of our findings were hypothesized by the experimenters and identified in their analysis. However, ordered patterns of fixation tags were very difficult and time consuming for the experimenters to assess since the only way to find them was to sequentially view the videos. Thus, eSeeTrack provided more detailed information about ordered patterns than the experimenters had obtained by other means. The experimenters could only qualitatively observe that experts looked back and forth between the vital and lap screens more frequently. Specific ordered patterns could not be identified and the relative frequency of patterns could not be quantified. We therefore expect that our visualization could be very helpful for initial exploratory analysis.

## 5.2 Case 2: Retail Store Chain

### 5.2.1 Background description

One of the main tasks of marketers is to assess consumer behaviors and to develop appropriate strategies to maximize profit. They use eye-tracking systems to study customer behavior in an attempt to gain a better understanding of buyers’ decision-making processes.

In retail stores, posters are intensively used to advertise new arrivals and/or articles on sale. Marketing teams may design printed ads that intentionally minimize legally required information or details such as the terms and conditions of a sale. For example, while promotional phrases on a poster are printed in large font to attract a customer’s attention, conditions of the sale are usually printed in tiny characters or using a similar color to the background. In order to verify whether the ads direct attention as expected, marketers may use eye-tracking systems.

### 5.2.2 Data and Hypothesis

Two different retail clothing stores belonging to the same chain but with varying styles of exhibition were examined. Fixation data were collected while customers were shopping in the stores. The first group consisted of 16 customers and the second group included 22 customers. Detailed explanation of the fixation categories is found in Table 3.

Table 3. Fixated Object Categories in the Retail Store Study.

Tag	Description
<b>Clothing</b>	Clothing displayed on models or on posters
<b>Item description</b>	Information about a product such as its country of fabrication, care labels, fabric types, etc.
<b>Poster</b>	A large printed placard that advertises certain articles and/or their sales
<b>Price tag</b>	Price of a piece of merchandise shown on an attached label or on a poster
<b>Sales condition</b>	Terms and conditions clause for acquiring a sale (on a poster)
<b>Sales promotion</b>	Phases on a poster to advertise a rebate
<b>Salesperson</b>	Person who sells merchandise in a store

The first goal of this case study was to examine whether two stores carrying the same merchandise but having two different display setups led to different customer behaviors. Store managers hoped that customer fixations patterns would be similar in both stores even though the stores’ arrangements were different.

The second goal of the case study was to verify that customers did not notice condition clauses associated with a promotion. This information was intentionally deemphasized so that customers would disregard it. Therefore, we expected to see few patterns containing sales conditions.

### 5.2.3 Analysis

After analyzing the fixation data with our tool, we concluded that there were only minimal differences in fixations for the two retail stores. This can be observed in the bar chart in Figure 10 and descriptive statistics of Table 4. As shown in the figure, most categories have a similar number of fixations in both stores, with the exceptions of clothing and sales promotion categories.



Fig. 10. Tag filter and bar chart showing fixations in two retail stores.

Table 4. Fixation Count and Percentage of Fixations on Each Category in Retail Store Chain. The Top 3 Categories in Each Store Are Highlighted.

	Group			
	●Store A		●Store B	
	Count	%	Count	%
<b>Clothing</b>	41	6.1	152	24.1
<b>Item description</b>	107	15.9	96	15.2
<b>Poster</b>	221	32.9	189	30.0
<b>Price tag</b>	86	12.8	66	10.5
<b>Sales condition</b>	14	2.1	2	0.3
<b>Sales promotion</b>	85	12.6	39	6.2
<b>Salesperson</b>	118	17.6	86	13.7
<b>Total</b>	<b>672</b>	<b>100.0</b>	<b>630</b>	<b>100.0</b>

We also constructed tree visualizations rooted at each of the top three categories. Four trees were constructed (not shown) because one of the top three categories was different in each store. The trees showed nearly all possible fixation patterns, suggesting that the store design and layout did not encourage users to gaze along any particular scanpath. In addition, there was no observable difference in the trees for the two different stores. Therefore, different store arrangements did not lead to different fixation patterns, addressing the first objective of the case study.

To verify the second hypothesis, we built two tree visualizations, shown in Figure 11. The first tree showed the fixation patterns ending in “sales promotion” and the second showed patterns beginning with “sales promotion”. In both trees, the most frequent pattern was a sequence of five “sales promotion” fixations as shown by the large labels that appear first in the tree. The five instances of “sales promotion” fixations signified that participants fixated repeatedly on objects categorized as “sales promotion”. Participants were probably reading the sales promotion phrases in one poster, but they may also have examined different posters. In both cases, more than one fixation was produced. This demonstrates that the sales promotion was effective at maintaining a viewer’s attention for an extended period, as a marketer would hope. We also observed that none of the patterns in either tree contained “sales conditions”. In addition, incrementing the gap duration from 1 to 3 seconds (see Figure 12) did not show any patterns containing sales conditions. Therefore, this confirmed the hypothesis that customers did not



Fig. 11. Tree visualizations with gap duration of up to 1 second showing the fixation patterns before (on the top) and after (on the bottom) looking at “sales promotion”. Colored labels belong to store A and grey labels belong to store B. Note that “sales condition” does not appear in any of these patterns in either store.

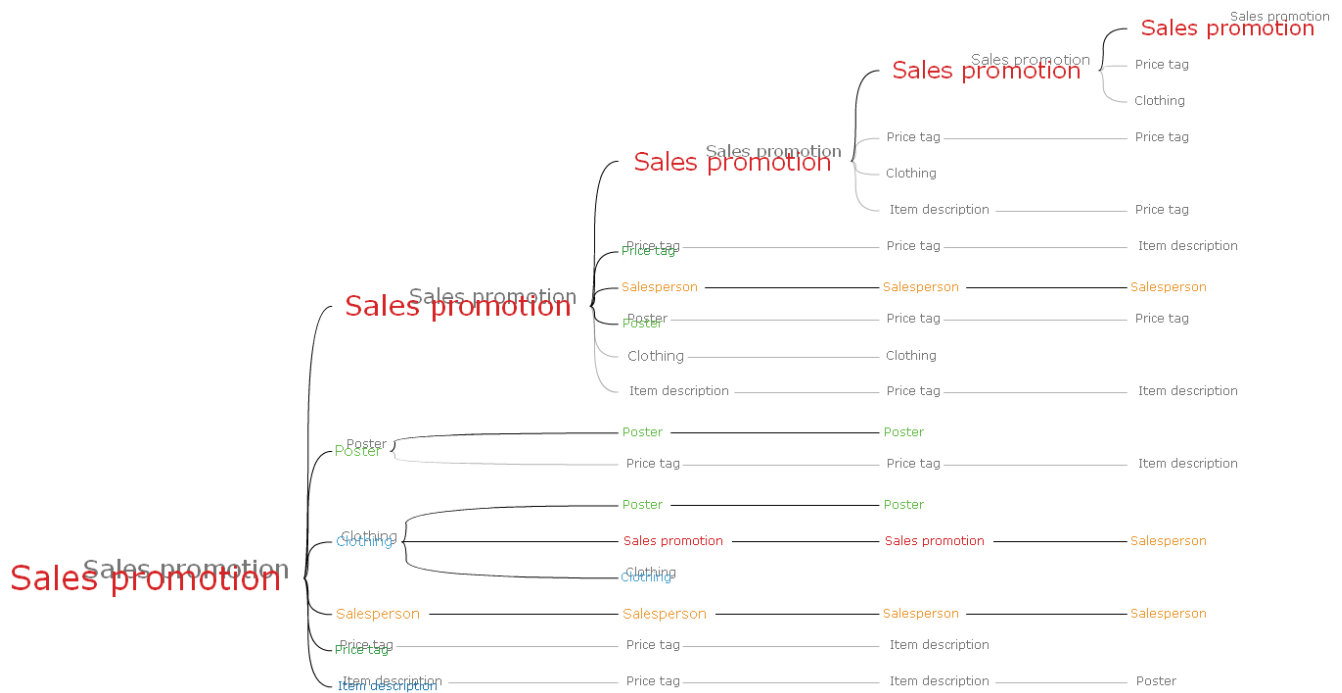


Fig. 12. Tree visualizations with gap duration of up to 3 seconds showing the fixation patterns after looking at “sales promotion”. Colored labels belong to store A and grey labels belong to store B. Note that “sales condition” does not appear in any of these patterns.

notice condition clauses associated with a promotion while they were looking at the promotion.

## 6 DISCUSSION

Development of eSeeTrack enabled us to explore and validate design ideas supporting exploration and comparison of fixation patterns. By combining a timeline and a WordTree-like visualization, our approach enables users to quickly explore short sequential patterns of

eye-gaze fixations, and compare those patterns across multiple groups. The two case studies were conducted with real eye tracking datasets to empirically test our visualization techniques. The results demonstrate that our approach can quickly reveal interesting and verifiable patterns. Frequent and infrequent tag sequences can be easily spotted in the tree visualization. The relative frequency of each fixated category can be easily obtained via the bar chart or the size of node labels. By adjusting the time window, users can quickly observe how these frequencies change across different times within a



session. Note, however, that our case studies only validated that patterns are observable with eSeeTrack; future user testing is needed to identify any usability problems and determine whether users can make sense of the information presented.

During the case studies, we discovered some shortcomings in our tool, particularly as the number of participants or fixations increases. These will be used to guide future refinements of the design. First, the detailed view can only show the ordering of fixations, not the durations or exact time. In addition, accessing the detailed information about a fixation via the magnifying glass can be cumbersome. These issues might be resolved by replacing the detailed timeline with a fisheye view or a zooming feature in the timeline and by accessing more information via a hover query. Secondly, since every parent node in the tree visualization is aligned with its middle child, when the tree becomes large, users have to scroll the panel in order to see the most frequent patterns. This problem could be solved either by rearranging the tree layout so that non-leaf nodes are always visible or by using a zooming interface. We also noticed that accessing sequence information about the group in grey shadow was more difficult than the colored group. This might be alleviated by allowing the user to interactively choose which group is shown in color, by coloring by group rather than tag name, or by using side-by-side trees for comparison instead of the drop shadow (although this would take more space). It would also be useful to align fixation patterns from different participants according to events or phases of their task (e.g., browsing vs. making a purchase in the retail scenario), and to be able to select these time windows across all participants.

The number of possible fixation categories is limited with our approach because the number of colors that can be easily distinguished is small. In practice, we have found that ten categories is sufficient for most analyses we have done so far; however, analyses requiring substantially more categories may require a new approach. One possible solution is to repeat the color set. In this case, a color may represent more than 1 fixation category. Another possible solution is to hierarchically group the tags and apply the color scheme at an intermediate level of the hierarchy.

We also imposed an upper limit of 5 on the sequence length since it was sufficient for the analyses we had done so far. This number could be easily increased to allow analysis of longer sequences. The tradeoff of increasing the length will be an increase in the time required to find sequences, and an increase in the tree size. As an alternative to leaving the choice of length up to the user, it may be possible to automatically choose an appropriate maximum length. The choice could be based on the number of tags present (interesting sequences may be longer when there are more categories) or by comparing sequences algorithmically using string editing and finding the maximum length match.

Development of eSeeTrack was motivated by our own previous challenges in examining sequences of eye-gaze fixations [27]. This type of analysis was difficult, time consuming, and very limited in scope with previous methods. Thus, by exposing time-ordered sequences of fixations, eSeeTrack supports a new type of analysis of eye-tracking data. We note that we intend eSeeTrack to support early phases of data exploration, and we anticipate that users would investigate interesting findings further with a statistical analysis.

## 7 GENERALIZATIONS

Our eye tracking example demonstrates that with some slight modifications, WordTrees can visualize more than just text data. When combined with a timeline view, this approach can effectively identify common (and uncommon) sequences of events and relate those to exact time points when they occurred. Furthermore, shadowed WordTrees can be used to compare two different sets of activities. We believe that this approach should extend beyond eye tracking, to visualize sequences of any type of time-ordered events or state transitions. Such sequences are important in any sort of behavior analysis. Examples from previous work in visualization

include observed states of animal behaviour [2] and sequences of human activities extracted from diary studies [29].

## 8 FUTURE WORK

In future work, we plan to continue refining eSeeTrack, adding new features and solving the issues stated in section 6. Some specific avenues for future work include: developing a heat map-like display for dynamic scenes, exploring ways to extend the number of tags via different visual encoding methods, developing alternate color schemes (perhaps based on participant group rather than tag), extending the allowable pattern length, and exploring alternatives to the drop-shadow method for visual tree comparison. We also wonder whether our extraction of specific ordered patterns is in fact too specific, since different numbers of repeat fixations on the same object each lead to different, but closely related, patterns. Collapsing these many similar patterns into one can be done by removing repeat elements [32], and may provide a more useful level of granularity for analysis. A more flexible user grouping strategy, based on more than one experimental variable and similar to the grouping in eyePatterns [32], may also be useful. Finally, we would like to deploy our tool for some additional case studies to further evaluate its effectiveness, and consider its extension to other types of behavior analysis beyond eye tracking.

## 9 ACKNOWLEDGEMENTS

This work was supported by Locarna Systems Inc., Mathematics of Information Technology and Complex Systems (MITACS), and the Natural Sciences and Engineering Research Council of Canada (NSERC). We would like to thank the surgery team (Dr. Stella Atkins and Geoffrey Tien of Simon Fraser University, and Dr. Bin Zheng of the Centre of Excellence for Simulation Education and Innovation at the University of British Columbia), the market research team (anonymous), and members of the VisID group at the University of Victoria for their time and valuable feedback.

## REFERENCES

- [1] A. Aula, P. Majaranta, and K.-J. Räihä. “Eye-Tracking Reveals the Personal Style for Search Result Evaluation”, *Human-Computer Interaction – INTERACT 2005*, M. F. Costabile and F. Paternò, eds., Springer, Berlin, pp. 1058-1061, 2005.
- [2] J. Blaas, C.P. Botha, R.S. Laramée. “Smooth Graphs for Visual Exploration of High-Order State Transitions”, *IEEE Trans. Visualization and Computer Graphics*, 15(6): 969-976, Nov.-Dec. 2009.
- [3] ColorBrewer. <http://colorbrewer.org>, 2010.
- [4] A.T. Duchowski. “A Breadth-First survey of eye-tracking applications”, *Behavior Research Methods, Instruments, and Computers*, 34(4): 455-470, Nov. 2002.
- [5] A.T. Duchowski, J. Driver, S. Jolaoso, B.N. Ramey, and A. Robbins. “Scanpath Comparison Revisited”, *Proc. Eye Tracking Research & Applications*, pp. 219-226, 2010.
- [6] Y. Egusa, M. Takaku, H. Terai, H. Saito, N. Kando, M. Miwa. “Visualization of User Eye Movements for Search Result Pages”, *Proc. 2nd Intl. Workshop on Evaluating Information Access*, pp. 42-46, Dec. 2008.
- [7] M. Feusner and B. Lukoff. “Testing for statistically significant differences between groups of scan patterns”, *Proc. Eye Tracking Research & Applications*, pp. 43-46, 2008.
- [8] J.H. Goldberg and J.I. Helfman. “Scanpath Clustering and Aggregation”, *Proc. Eye Tracking Research & Applications*, pp. 227-234, 2010.
- [9] J.H. Goldberg and J.I. Helfman. “Visual Scanpath Representation”, *Proc. Eye Tracking Research & Applications*, pp. 203-210, 2010.
- [10] J.H. Goldberg and X.P. Kotval. “Computer interface evaluation using eye movements: methods and constructs”, *Intl Journal of Industrial Ergonomics*, 24(6): 631-645, Oct. 1999.
- [11] R.I. Hammoud and J.B. Mulligan. “Introduction to Eye Monitoring”, *Passive Eye Monitoring*, Springer, Berlin, pp. 1-19, 2008.

- [12] M. Hayhoe and D. Ballard. "Eye movements in natural behavior", *TRENDS in Cognitive Sciences*, 9(4):188-94, Apr. 2005.
- [13] H. Hembrooke, M. Feusner, and G. Gay. "Averaging Scan Patterns and What They Can Tell Us", *Proc. Eye Tracking Research & Applications*, p.41, 2006.
- [14] K. Itoh, J.P. Hansen, and F.R. Nielsen, "Cognitive Modeling of Ship Navigation Based on Protocol and Eye-Movement Analysis", *Le Travail Humain*, 61(2): pp. 99-127, 1998.
- [15] R.J.K. Jacob and K.S. Karn, "Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises", *The Mind's Eye: Cognitive and Applied Aspect of Eye-Movement Research*, J. Hyona, R. Radach and H. Deubel, eds., Elsevier, Amsterdam, pp. 573-605, 2003.
- [16] H. Jarodzka and K. Holmqvist. "A Vector-based, Multidimensional Scanpath Similarity Measure", *Proc. Eye Tracking Research & Applications*, pp. 211-218, 2010.
- [17] C. Lankford. "GazeTracker: Software Designed to Facilitate Eye Movement Analysis", *Proc. Eye tracking Research & Applications*, pp. 51-55, 2000.
- [18] B. Law, M.S. Atkins, A.E. Kirkpatrick, A.J. Lomax. "Eye gaze patterns differentiate novice and experts in a virtual laparoscopic surgery training environment", *Proc. Eye Tracking Research & Applications*, pp.41-48, 2004.
- [19] J. Nielsen and K. Pernice. *Eyetracking Web Usability*, New Riders Press, Berkeley, p. 9-12, 2010.
- [20] M. Pomplum, H. Ritter, B. Velichkovsky. "Disambiguating Complex Visual Information: Towards Communication of Personal Views of a Scene", *Perception*, 25: 931-948, 1996.
- [21] V. Ponsoda, D. Scott and J.M. Findlay. "A probability vector and transition matrix analysis of eye movements during visual search", *ActaPsychologica*, Elsevier, Amsterdam, 88(2):167-185, 1995.
- [22] C.M. Privitera and L. Stark. "Algorithms for Defining Visual Regions of Interest: Comparison with Eye-Fixations", *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(9): 970-982, 2000.
- [23] K.-J. R  ih  , A. Aula, P. Majaranta, H. Rantala, and K. Koivunen. "Static Visualization of Temporal Eye-Tracking Data", *Human-Computer Interaction – INTERACT 2005*, M. F. Costabile and F. Patern  , eds., Springer, Berlin, pp. 946-949, 2005.
- [24] D.C. Richardson and M.J. Spivey. "Eye Tracking: Characteristics and Methods", *Encyclopedia of Biomaterials and Biomedical Engineering*, G. Wnek and G. L. Bowlin, eds., Informa Healthcare Publisher, New York, pp. 1028-1032, 2004.
- [25] O.   pakov and D. Mini  as. "Visualization of Eye Gaze Data using Heat Maps", *Electronics and Electrical Engineering*, 2:55-58, 2007.
- [26] G. Tien, M.S. Atkins, B. Zheng and C. Swindells. "Measuring Situation Awareness of Surgeons in Laparoscopic Training", *Proc. Eye Tracking Research & Applications*, pp. 149-152, 2010.
- [27] M. Tory, M.S. Atkins, A.E. Kirkpatrick, M. Nicolauo, and G.-Z. Yang. "Eyegaze Area-of-Interest Analysis of 2D and 3D Combination Displays", *IEEE Visualization*, pp. 519-526, Oct. 2005.
- [28] E. Vatikiotis-Bateson, I.-M. Eigstic, S. Yano and, K. G. Munhall. "Eye movement of perceivers during audiovisual speech perception", *Perception & Psychophysics*, 60(6): 926-940, Aug. 1998.
- [29] K. Vrotsou, J. Johansson, and M. Cooper. "ActiviTree: Interactive Visual Exploration of Sequences in Event-Based Data Using Graph Similarity", *IEEE Trans. Visualization and Computer Graphics*, 15(6): 945-952, Nov.-Dec. 2009.
- [30] C. Ware. *Information Visualization: Perception for Design (Morgan Kaufmann Integrative Technologies Series)*, 2<sup>nd</sup> edition, Morgan Kaufmann Publishers, California, pp. 123-126, 2004.
- [31] M. Wattenberg and F.B. Vi  gas. "The Word Tree, an Interactive Visual Concordance", *IEEE Trans. Visualization and Computer Graphics*, 14(6): 1221-1228, Nov.-Dec. 2008.
- [32] J.M. West, A.R. Haake, E.P. Rozanski, and K.S. Karn. "eyePatterns: Software for Identifying Patterns and Similarities Across Fixation Sequences", *Proc. Eye Tracking Research & Applications*, pp. 149-154, 2006.
- [33] D.S. Wooding. "Fixation Maps: Quantifying Eye-movement Traces", *Proc. Eye tracking Research & Applications*, pp. 31-36, 2002.