Increasing the User's Attention on the Web: Using Implicit Interaction based on Gaze Behavior to Tailor Content

Florian Alt, Alireza Sahami Shirazi, Albrecht Schmidt
University of Stuttgart,
Institute for Visualization and Interactive Systems
{firstname.lastname}@vis.uni-stuttgart.de

Julian Mennenöh
University of Duisburg-Essen,
Marketing and Retail
julian.mennenoeh@uni-duisburg-essen.de

ABSTRACT
The World Wide Web has evolved into a widely used interactive application platform, providing information, products, and services. With eye trackers we envision that gaze information as an additional input channel can be used in the future to adapt and tailor web content (e.g., news, information, ads) towards the users' attention as they implicitly interact with web pages. We present a novel approach, which allows web content to be customized on-the-fly based on the the user's gaze behavior (dwell time, duration of fixations, and number of fixations). Our system analyzes the gaze path on a page and uses this information to create adaptive content on subsequent pages. As a proof-of-concept we report on a case study with 12 participants. We presented them both randomly chosen content (baseline) as well as content chosen based on their gaze-behavior. We found a significant increase of attention towards the adapted content and evidence for changes in the user attitude based on the Elaboration Likelihood Model.

Author Keywords
Eye tracking, implicit interaction, adaptative content.

ACM Classification Keywords
H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *Interaction Styles*

General Terms
Experimentation, Human Factors

INTRODUCTION
The Internet has become a platform providing a wide variety of applications, services, and content such as news, ads, weather information, and stock trends. These content elements are often displayed alongside the main content of web pages and are an important part of many sites' business model. However, this content is often ignored by the user as a result of the banner blindness [6]. Furthermore, the success is difficult to be measured except for cases where users click.

NordiCHI'12, October 14–17, 2012, Copenhagen, Danmark.

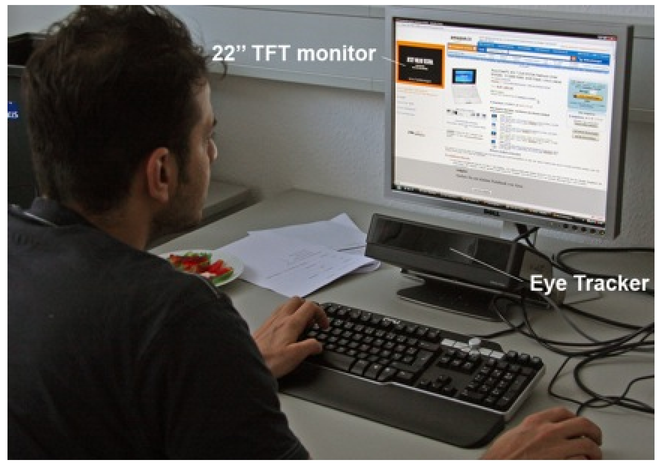

Figure 1: Adaptive WebAds: As the user interacts with the web, his gaze is implicitly monitored, fed back to the system, and the content is adapted automatically.

Eye trackers have the potential to overcome these issues. We believe that in the future many personal computers will be equipped with means for analyzing a user's gaze information for implicit interactions with web pages. Even though eye trackers are currently still expensive, prices may drop once applications for the mass market become available – similar to webcams and fingerprint readers which used to be specialized and expensive devices not too long ago.

Furthermore, eye tracking can add a new quality to the Web both for the viewer and the content provider. Since the user's gaze path and hence his attention can be measured, content can be adapted in real time and make websites more attractive. From the provider side, access to the user's gaze data is interesting as it can enhance measuring the success of web content. Currently, the most popular measures are page impressions (i.e., how often a page is served to the browser) and the clickstream (i.e., how many users clicked on content). As the gaze path is available and content users look at can be identified, a more fine-grained analysis becomes possible.

Access to the user's gaze behavior is a privacy issue and not all users may want their gaze data in the hands of the content provider. We observe that particularly digital natives are willing to share private information and communicate via public channels (e.g., Twitter). However, mechanisms need to be established for commercial use that allow users to define whom to grant access to this information.

In this paper we aim at answering two research questions. First, we are interested whether *user attention towards content* can be increased by tailoring this content based on gaze information. Second, we investigate whether there is a *cognitive effect of adaptation on the user*. In order to exploit this idea we built a system, that can adapt a web page's content elements on-the-fly. Using an eye tracker the system assesses the viewer's gaze and then feeds back the information to the browser (Figure 1). This information is used to adapt the content on subsequent pages accordingly. In a lab study with 12 participants we let the users surf the Amazon website. We adapted some of the image elements and measured whether attention would increase for these, compared to randomly chosen elements. We were able to show a significant increase of attention. Additionally, we found evidence for the peripheral route of the Elaboration Likelihood Model [28] implicating an influence on the user's attitude.

The contribution of our paper is twofold:

- We introduce the notion of gaze-based adaptation on web pages and outline how a user's gaze information can be used to create adaptive content. To show the feasibility of this approach, we present a prototype system, which allows adaptive content to be inserted into web pages on-the-fly while users implicitly interact with the page.
- We report on an experiment and show that placing adaptive content based on the user's gaze information on web pages impacts on the user's attention and attitude.

BACKGROUND AND RELATED WORK

Today, a variety of stationary high-precision eye trackers are commercially available. These are still expensive, but as a mass-market emerges we envision prices to drop significantly and the integration with public displays to become feasible. Current systems support a large freedom of head movement and allow users to behave naturally. With data rates up to 120 Hz, trackers provide real time information about the gaze point. Currently, eye tracking systems are mainly found in labs due to their high price tag. However, a lot of research has been focused on self-made *low-cost eye trackers*. Hanson et al. report on the potential of eye trackers built upon components off-the-shelf (COTS) [19]. Li et al. developed the openEyes system [24], an open hardware design with several open-source software tools for eye trackers. They showed how the performance of those low-cost trackers could be incrementally improved by replacing parts of the systems with higher quality components[1]. Current developments suggest, that eye tracking systems might be included in standard computers at little or no extra cost in the future as more applications become available. Modern laptops are already widely equipped with devices such as webcams or fingerprint readers, which were also specialized and expensive decades ago.

In general, two categories of eye tracking systems can be distinguished [12] – diagnostic and interactive systems. *Diagnostic systems* focus on the offline analysis of data gathered from user interaction to evaluate usability or user behavior.

[1]See also the homepage for the COGAIN Network of Excellence and the COGAIN Association `www.cogain.org`.

In marketing, gaze data are particularly interesting, as they provide insights into visual, cognitive, and attentive aspects of human performance, as well as how consumers disperse visual attention over difficult forms of advertising [12]. Further research on diagnostic systems includes tools focusing on web reading. Beymer and Russell present the WebGazeAnalyzer [4], a tool that supports the collection, analysis, and re-analysis of gaze data. The WebEyeMapper [34], analyzes eye tracking data and maps it onto objects on a web page, making it easy for researchers to comprehend the gaze data collected. Diagnostic systems are also used to explore user behavior on the World Wide Web (WWW). A prominent example from the field of advertising is the work of Burke [6] on the usefulness of banner ads and banner blindness.

In contrast, *interactive systems* focus on how eye trackers can be used to control applications in scenarios where people are not able to use other input devices, e.g., a surgeon during operations. Chen et al. investigate the correlation between eye and mouse movements [8]. Sibert and Jacob compare a novel gaze-based object selection technique with conventional selection by mouse [37]. Farid et al. [14] investigate how eye gaze can be used to control computer displays (e.g., navigating within large images or multiple video streams). More specific to the WWW, tracking gaze interaction on web pages requires an analysis of the intersection of eye gaze and the DOM bounding boxes of the viewed web page, as done by Reeder et al. [34] (offline) and Biedert et al. [5] (online). Whereas the previous examples are mainly used to explicitly control a system, researchers also look into how eye trackers can be used in an implicit way. The MAGIC pointing technique [41] is one of the early projects using gaze information for positioning the mouse cursor and assisting interaction. Another example is Santella et al.'s [35] application for photo cropping based on a user's gaze, which implicitly identifies regions-of-interest based on fixation data. Buscher [7] use eye tacking as a data source for realizing attention-based feedback on subdocument level. As a use case they examine personalized, context-based query expansion and ranking. Biedert et al. [5] discuss using gaze data for creating responsive text.

Furthermore, eye trackers have are used to build *attentive user interfaces* (AUIs) [38]. These UIs try to manage the attention of the users through input channels beyond conventional, explicit channels and bring the right information at the right time to the user. Computer vision and other technologies can potentially be used as input channels for AUIs, e.g., the user's presence, body posture, head direction, etc. However, the user's gaze information is a particularly rich resource. EASE, described by Wang et al. [40], uses the gaze data to assist Chinese text entry. Qvarfordt and Zhai [33] developed an interactive tourist system, which senses user interests based on eye-gaze patterns and manages data output accordingly. Drewes and Schmidt [11] enhanced the MAGIC system with a touch-sensitive mouse to ease the pointing task in graphical UIs. EyeWindows, presented by Fono et al. [15], is an attentive windowing technique that uses eye tracking for focus window selection. Selker discusses the complexity of UI designs based on simple observations of eye behavior [36].

For a framework on increasing attention through AUIs we refer to Vertegaal et al. [39]. In the context of the WWW, previous work mainly focuses on revealing user interest, e.g., [1, 2, 9]. Analyzing reading behavior provides relatively reliable results and is studied mainly in the context of implicit search queries [10, 13, 16, 20]. In contrast, inferring the relevance of images based on gaze turned out to be challenging and was only successful under controlled lab conditions [23].

Despite considerable efforts to create attentive UIs based on the user's gaze data, approaches so far mainly focus on enhancing search queries. Little is known about how gaze-behavior towards forms of content other than text can be exploited. To the best of our knowledge this project is the first approach to use real-time gaze data towards (ad) images with the aim to increase attention.

GAZE-BASED ADAPTATION ON THE WEB

When observing users in front of screens, it becomes clear that they are interacting with more than just the provided input devices (e.g., keyboard, mouse). They are often standing or sitting seemingly motionless while their eyes actively scan a section of the screen. Eye trackers could retrieve this information and allow it to be used in a variety of application domains, both offline and in real-time. The most important ones include UIs for people with disabilities, market research, advertising, and usability. For marketing research and usability engineering, eye trackers are typically used offline, meaning that experiments are carried out prior to analyzing gaze paths and behavior. In UIs for disabled persons, eye trackers can enable gaze-based, real-time interaction. Eye trackers have already been integrated into eye-typing interfaces, wheelchair steering systems, or remote controls.

Gaze data can be used for creating adaptive interfaces. Oppermann defines the term *adaptation* as a process in which a system adapts its behavior to users based on previously acquired and processed information about the user [27]. We see a large potential in adapting content to the attention or even the interest of the user.

The adaptation can be realized in two ways, by tailoring content based on (1) previously gathered knowledge, such as a profile (offline) and (2) current user behavior (real-time). The first approach (*offline*) relies upon the identification of the user and on having access to the previously assessed data. Such data is provided either explicitly by the user (e.g., a questionnaire on their interests) or implicitly by collecting data on the user's behavior (e.g., encounters at customer touch points, such as purchases). The second approach (*real-time*) assumes that the user's interest can be predicted from the currently accessed content (such as Google adSense) in addition to knowledge about the current behavior. If the user is, e.g., in an online shop looking for a specific item or if they write an email requesting information on a product, providing ads for similar products is then both feasible and sensible. However, when the user looks at more generic content, e.g., a news site, the link to a specific user's interest is less clear. For example, if the user reads about a plane crash on a news site, is it appropriate to advertise for cheap flights from the similar airline?

The advantage of the second approach is that the user's interest is determined in real-time and eliminates the need for storing personal profiles. We show that furthermore the attention towards content can be increased and new opportunities for structuring/selecting information be offered. The following scenario points out some of these opportunities.

Scenario – Adaptive Advertisements It is early evening as Paul arrives at a hotel in Stuttgart for a business meeting starting the following morning. After checking in, he prepares some slides on his notebook. An hour later he decides to quickly check his email before going out for dinner. Based on the IP address, the web-based email client detects that he is in Stuttgart, and pulls some locally relevant ads to be displayed alongside the inbox. As Paul reads his emails, his eye catches an advertisement for a Thai restaurant next to an advertisement for a musical showing in a local theatre. As he is not really up for Thai food he continues reading his email. At the same time, the eye tracker recognized that the Thai restaurant received most of Paul's attention among the ads and brings up three more restaurants, including an Italian Cucina, a German Bar, and a French Bistro. Paul is a big Pizza fan and immediately clicks the ad of the Italian restaurant to get further information on directions and the menu. He receives a coupon that saves him 3 euros on the Pizza, takes a picture of it with his mobile phone, and heads for the restaurant.

The scenario can be captured by a real-time analysis of the user's gaze input. With gaze information it becomes possible to obtain the required information on-the-fly and adapt content instantly and more appropriately. Whereas based on a click attention (or even interest) can be assumed for one page only (the page clicked), gaze input can consider any page element. The drawback of the gaze method is that the intention or motivation for the user is less clear. Clicking a content element is very likely to occur out of interest, but eye contact might be the result of subconsciously scanning the page and may not be related to the user's interest. Therefore, perception or even interest for gaze is much more difficult to determine, as (implicit) interaction only occurs for very short bursts. Hence, this paper focuses on adaptation based on the user's attention rather than on perception or interest. Mello-Thoms et al. [25] and Hauland [21] shows that the dwell time, i.e. how long a user looks, could be used to compare attention distributed between targets.

In the following we look at suitable metrics, formally present our approach, formulate the hypothesis, and introduce the Elaboration Likelihood Model (ELM) we used to evaluate and explain cognitive effects.

Metrics

To analyze the user's gaze behavior, Poole and Ball [31] identify different eye-movement metrics. In the following, we discuss those of interest to our research.

Dwell time (a.k.a. gaze or fixation cluster) The dwell time (overall gaze duration per user) can be used as a measure for attention among a set of targets [21, 25]. Yet, no conclusion on the interest towards the target can be made.

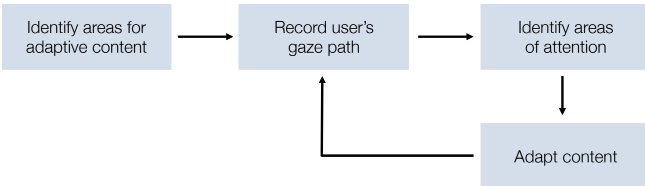

Figure 2: Cycle of creating adaptive content.

Number of fixations per target The number of fixations for a target can be used as a measure for its importance. Hence, a higher number of fixations indicates that a target (e.g., a page element) is more noticeable or important to the viewer than others [30].

Fixation duration The average duration of fixations per item can be used as an indicator for engagement or for the difficulty a user faces when extracting information [22].

Approach

As users look at different elements, their eyes are constantly moving, even if they are not consciously aware of it. With eye tracking, the user's focus and the content at this position can be determined. This information can be used to derive the attention for a given element. Our general approach for gaze-based adaptation is as follows (see also Figure 2). First, a set of content areas potentially attractive to the user needs to be identified. These content areas can be texts, images, videos, or animations and may be located in any part of the screen. Second, an eye tracking system is used to detect the user's gaze behavior towards these content areas. This information should be monitored in real-time and transferred to a server-side application. Third, the acquired gaze information is used to calculate the user's attention towards each content area. Fourth, an attention measure is used to determine elements to be adapted. Steps 2–4 are constantly repeated.

Hypotheses

Based on our approach, the following hypotheses will be answered during research:

Hypothesis 1 – Adapting content towards the users' gaze behavior increases their attention.
Today, click streams serve as both a measure for the success of web pages and as input data for adaptive content. With eye tracking we envision to find a means for an on-the-fly assessment of what attracts the user's attention and use it in real time to build attentive systems.

Hypothesis 2 – Adapting content has a cognitive impact on the user.
We assume that showing attractive content has an effect on the user engagement or results in a change in attitude.

Effect

In order to understand the effect of our approach on the user, we draw upon the widely accepted Elaboration-Likelihood-Model (ELM) [28]. The ELM tries to explain the changes in a user's attitude. The model differentiates between two routes leading to persuasion or elaboration: the central route and the peripheral route. Which route is taken depends on the user and the situation they are in. The ELM distinguishes between two factors: the motivation to process and the ability to process. If both factors are true, the central route is taken where conscious information attainment and processing takes place. As a result, the average fixation duration as a measure of engagement would increase significantly. If the peripheral route is taken, unconscious information acquisition takes place. In this case, advertisers should repeat peripheral cues to achieve persuasion [29]. Thus, a change in attitude can be accomplished through a significant increase in the number of fixations.

Through a study, we aim to find out whether gaze-based adaptation of content could support (a) the central route and thus lead to higher *engagement* (this would be the case for an increase in the average fixation duration) or (b) the peripheral route and thus lead to a *change in attitude* (this would be the case for an increase in the number of fixations). Finding any of these effects would make the approach highly interesting for advertisers as a strong influence on the user's elaboration could be assumed.

PROTOTYPE

Our central use case for adaptive content is the improvement of image-based web ads based on implicit, gaze-based interaction. Currently, web ads are among the most important means for generating revenue from web pages but many web ads are still presented in a manner that fails to anticipate the user. Advertising, however, is most effective if the right ad is provided to the right person at the right time. Many approaches have been explored to optimize this perfect match between the user (ad viewer) and the ad provided. Optimizing the match is clearly good for the advertiser, since it increases the chance that the potential customer gets aware of their product or service. Also, optimizing the match is great for the user, since the ads inform him about a product or service he is likely to be interested in. Two approaches for achieving the optimal match exist: (1) maximizing exposure, (2) adapting content based on profiling.

Maximizing exposure works by placing ads in very visible areas. This approach is simple but can be costly as well as can create a negative perception of the advertisement, if a large portion of the viewers are not interested in it and become annoyed with its omnipresence. When constantly provided irrelevant information, humans are also apt to ignore the information completely, such as with banner blindness [6], or display blindness [26]. Studies show that advertisements matching the users' interest are perceived less annoying than random advertisements. On the other hand, profiling is nowadays widely used both in the real world (e.g., shopper loyalty cards) and on the web (e.g., click stream analysis). Yet, this approach is often perceived as privacy invading as personal information needs to be stored over time.

Following our approach we suggest adapting ad content based on real time gaze behavior. This approach has the advantage that no data need to be stored and that users can opt out at anytime by simply turning off the tracking device.

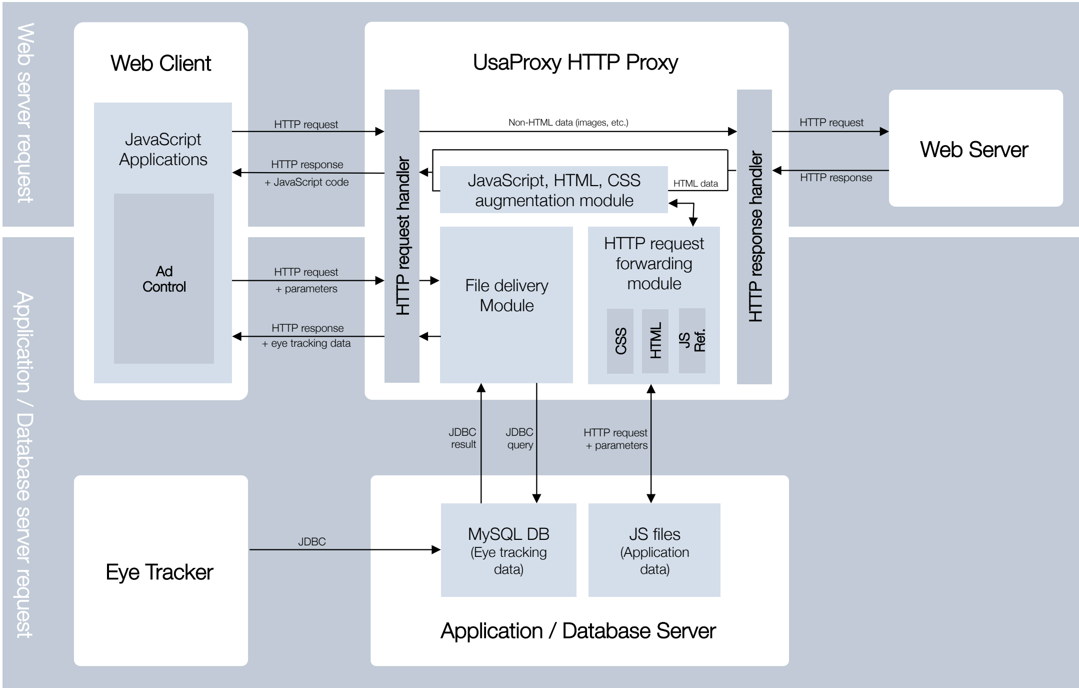

Figure 3: System Components – UsaProxy HTTP proxy, eye tracker, application/database server, web server.

The basic architecture of our system is depicted in Figure 3. It is based on a standard web architecture consisting of a web client (browser) and a web server (content provider). An eye tracker is used to implicitly track the gaze-based user interaction on a web page and store the gaze path in a database. In order to realize real-time interaction on a web page, we use an HTTP proxy to (1) insert application code into the website which handles the adaptation, and (2) to read and process the recorded gaze data to update the appropriate content and to trigger the update in the client.

Eye Tracker

In this study, we use a Tobii X120 eye tracker to extract coordinates of the eye's focal point. It is table-mounted and supports data rates of up to 120 Hz. We implemented the tracking software in such a way that the entire gaze path is time-stamped and recorded in a database at 60 Hz. This allows the gaze behavior to be monitored with regard to number of fixations, duration of fixations, and dwell time.

UsaProxy

In interactive systems that rely upon gaze data, it is necessary to process the recorded gaze data in real-time. In order to do so, we used the HTTP proxy UsaProxy [3]. It allows JavaScript code required for processing the collected data to be inserted on-the-fly on arbitrary web pages. The analysis of the data can be done either on the server or the client side.

The UsaProxy's initial task is to embed JavaScript code in any page sent from the web server to the client in response to a standard HTTP request. This makes it possible to embed content in the client instead of in the proxy. In order to add the script, UsaProxy monitors all HTTP requests that pass through it and adds a `<script>` tag inside the document's `<head>`, and its *src* attribute references the required JavaScript. As JavaScript code can be inserted into pages, data from external sources can also be loaded explicitly on to the page. Therefore, the UsaProxy was slightly modified to support XMLHttp requests to external data sources. The script, which runs inside the browser as a result of the above modifications, uses further XMLHttp requests to access the content on the external database server. Using the proxy we were able to circumvent the same origin policy [3]. We implemented a method to handle AJAX requests for downloading the data. The requests contain information about the object the user is looking at. UsaProxy retrieves and sends the current coordinates of the eye's focal point to the browser.

Apparatus

The Amazon corporate website served as the test environment, hence, allowing realistic tasks to be performed in a familiar setting, such as searching for a product (e.g., looking up the current price for an iPod touch 8GB). We used the UsaProxy to embed additional advertising elements into the web page, which then updates based on the user's gaze behavior. Our changes did not alter the look and feel or the URL of the Amazon website. In general, any website with arbitrary content elements is suitable. The sole information the system needs is the position and size of the elements to correctly associate them with the user's gaze. This setup allowed us to determine the dwell time, the number of fixations, and the duration of fixations on each page element. As suggested by Poole and Ball [31] the dwell time is used as the most appropriate means to compare attention between targets in order to decide which elements to adapt.

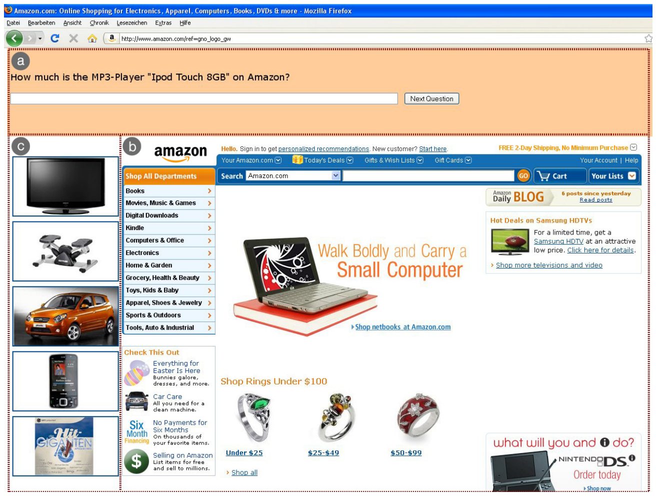

Figure 4: Layout of our test website, consisting of a task description area (a), the main content area (b), and the advertisement area (c).

Layout of the Test Web Site

Figure 4 shows the layout of the altered Amazon website with three distinct areas. We inserted a *task area* (a), allowing arbitrary tasks to be presented to the subjects. This area remained unchanged during navigation on the page. Once a task was solved, the answer was entered into the text field provided. Clicking on the 'Next Question' button triggered the system to randomly draw the next task from a database. The *main content area* (b) showed the original Amazon website. It was fully functional and users could freely navigate around the page. The *advertisement area* (c) was inserted using JavaScript. In this area, we showed different advertising elements. To maximize exposure of these elements they were inserted on the left side of the page since this area is most likely to be perceived by the users [32]. Note that this is consistent through all conditions in the study. This area was integrated into the corporate Amazon design.

Learning and Measurement Phase

To realize the adaptation of the images we defined a learning phase and a measurement phase (Figure 5) that are continuously swapped while the user is surfing the Web. In the *learning phase*, the system collects the user's gaze data hence making it possible to identify the area that received the most attention by calculating the dwell time. For the subsequent *measurement phase* the system updates the ad elements on the next page. Note, that in order to not bias the data gathered, the images are not changed while browsing a page, since this will most likely generate additional attention. Hence, image updates are seamlessly integrated upon reloading a page as a result of clicking a link.

In order to compare the effect of adapting content, the system provides two modes: in the *random mode* (Figure 5, top) the advertising images are updated randomly, in the *gaze mode* (Figure 5, bottom), the images are updated based on the collected gaze data.

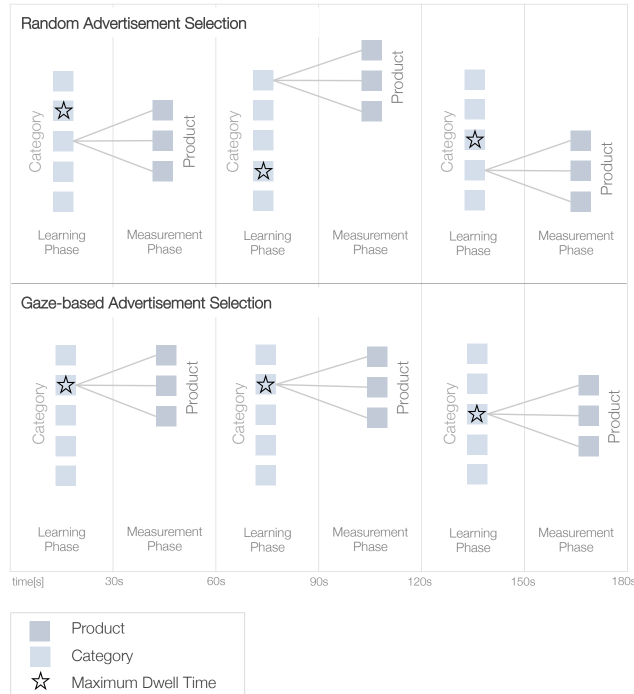

Figure 5: Adaptation of Content: (1) In random mode (random advertisement selection), product images are selected independent of the gaze. (2) In gaze-based mode (gaze-based advertisement selection), product images are adapted based on attention (dwell time).

Adaptation of Content

As adaptive content for the user study we prepared different types of advertising images: category images and product images. *Category images* represented a certain product category (e.g., cars, mobile phones, home entertainment, sport equipment, media). Those images were designed in a way such that only the type but not the brand of a product was recognizable. *Product images* represented a certain product related to one of the categories, e.g., BMW, Mercedes, and Seat for the category 'car'. In these images the brand of a product was clearly recognizable.

As we planned for a within subject design we created two image sets to avoid any learning effects. Each set consisted of five category images and three product images per category, resulting in a total of 15 product images.

In order to measure the effect of adapting content to attention, the system shows the subjects the five different category images in the learning phase and measures the dwell time for each image. After 30 seconds, the system triggers a change event so that the five category images are replaced by three product images of the same category upon reloading the page. Based on the software mode the three product images are either chosen randomly or based on the gaze data (by using the category which received the highest sum in dwell time). This procedure is repeated continuously.

EVALUATION

In the following we report on the evaluation of our approach. We first discuss the used methodology before presenting the study setup, procedure, and data analysis.

Methodology

We discussed different alternatives with regard to evaluation of the approach. A long-term field study would have allowed a large set of data to be collected and analyzed and the long-term effect on the user and his behavior to be assessed. However, the following challenges prevented us from running the study in the field. First, the eye tracker needs to be re-calibrated as the user changes their position (e.g., every time they sit down on the chair) which would have put an unbearable burden on the user and potentially led to biased data. Second, in order to collect enough data, we would first need to identify the favorite websites of the user. Then we would need to prepare and integrate the adaptable content with these websites to maintain their look and feel to avoid revealing the investigated objects.

As a consequence, only the potential of the approach in a controlled lab environment is evaluated. This allowed precise data on the gaze behavior of the user to be gathered and the effects of the approach (changes in dwell time, number of fixations, fixation duration) to be measured. As a baseline condition, randomly selected content is used as this most closely reflects the way image-based ads are currently presented in the WWW. We used a within-subject design where all users would be shown both the random and the gaze-based content (independent variables). The dependent variables were dwell time, the number of fixations, and the duration of fixations.

Furthermore, we asked the users to fill out a questionnaire after the experiment to assess recognition for the products. The questionnaire contained in total 30 product images, 15 of which were shown to the participants during the study. For each of the products the participants had to rate on a 5-Point Likert scale how sure they were to have seen the product image. We also conducted semi-structured interviews with a focus on user acceptance.

Participants

The participants were recruited via bulletin boards in the neighborhoods surrounding the university, from mailing lists, and lectures. In total, 15 participants (avg. age: 26.5 years) were selected. None of them ever participated in an eye-tracking study before and all have previously used Amazon.

Setup and Procedure

The study was conducted in an office in the lab. Users did not engage in conversations during the experiment with others around them or turn their heads away from the eye tracker. A table-mounted eye tracker in front of a 22” TFT monitor (Figure 1) allowed the users to behave more freely, as they were not restricted to a fixed (head) position.

All of the participants were briefed in the room where the eye tracker was set up and calibrated. We explained to them about the area for task descriptions on top of the website. Then, we asked them to solve the provided tasks (e.g., looking up the price for an iPad) and enter their answers into the text field for the next 20 minutes. The participants were informed about the collection of gaze data but we neither revealed the study objective nor the additional ads. The users were instructed not to leave the Amazon website. After 10 minutes, the system switched between random and gaze-based mode. To avoid any learning effects, half of the participants started with the gaze-based mode, the other half with the random mode. After the experiment the participants filled out a questionnaire and engaged in semi-structured interviews.

Data Analysis

During the study we recorded the participants' (time-stamped) gaze path, resulting in a total of 722,689 data points. During the analysis we found that for 3 participants the tracker had not recorded gaze-data properly (due to their bright eyes or glasses). Hence, we had to exclude their datasets. We used the software package ‘IBM SPSS Modeler’ to calculate the dwell time, number of fixations, and average fixation duration per ad element.

RESULTS

To identify the impact of different images on the users' attention and attitude, we used Student's t-Test because of its good performance with sample sizes below n=30 [18]. There are 24 ‘observations’ from 12 participants. Thus, we used a paired sample t-test. Since we have the directed assumption of increasing attention, the testing is one-tailed [17].

At first we compared the dwell time between random and gaze-based images. We found an increase of 191.6 ms per participant for the gaze-based adaptation, yet this is not significant (Table 1c). When digging further, however, we discovered that there is a significant increase in the number of fixations for the gaze-based images (Table 1a). This is a strong indicator that the participants consider these images to be more noticeable and hence pay more attention.

Remarkably, while the number of fixations increases significantly, the average fixation duration decreases non-significantly (Table 1b). Thus, it is likely to assume about the same fixation duration in the parent population. According to the assumptions of the ELM this indicates an unconscious information attainment in the peripheral route. An increase in the average fixation duration would indicate more elaborated and conscious information processing in the central route.

Note, that since the average dwell time increases systematically with the number of fixations and the average fixation duration, it seems that the (non significant) decreasing fixation duration lowers the increase of the dwell time (Table 1c) in this specific sample. According to the t-test, we would expect the fixation duration to be equal in the parent population.

For recognition we tested whether the participants can recognize randomly chosen better or gaze-based product images on a 5-point Likert scale (1=I definitely did not see the ad, 5=I definitely saw the ad). We found an increase for images shown in the gaze-based condition (M=3.20, SD=0.92) compared to the test condition (M=2.72, SD=0.51), yet this effect is not significant, $t(11)=1.358$, $p=0.10$ (one-tailed).

Measure (mean per user)	random study (control)	gaze-based study (test)	overall	T-test (paired sample; df=11)	Sig. (one-tailed)
(a) avg. # fixations per user	24.00	32.17	28.08	2.478	$p = 0.015$
(b) avg. duration per fixation and user	52.20	44.48	48.34	0.550	n.s.
(c) avg. dwell time per user	1238.14	1429.75	1333.94	0.511	n.s.

Table 1: Comparison of random / gaze-based image selection: The number of fixations increases significantly ($p = 0.015$).

Based on the semi-structured interview we found that 5 participants would not use the system due to privacy concerns whereas the others would be happy in case an option is provided to turn of monitoring the gaze data. Only one participant realized that the Amazon web page was customized. Additionally, none of our participants realized that we adapted content based on gaze information. The quantitative and qualitative data as well as the user feedback show that probably no or very little conscious information processing took place during the study.

DISCUSSION

The observations during the study and the analysis demonstrate that implicit gaze interaction is a powerful modality for creating new and persistent user experiences. Without additional effort for the user, content can be tailored to increase their attention, since the natural gaze movement is a rich resource for information about what they pay attention to. At the same time, the increase in the number of fixations shows that the approach has the potential to affect the users' attitude, according to the peripheral route of the Elaboration Likelihood Model. In market research and usability studies, eye tracking and offline analysis is best practice and commonly used. We demonstrated that with current technologies, it is possible to benefit from this information in real time. Our research has explored how this information can be used to adapt the user interface in real time and by these means, create an effective feedback loop. We observed that these findings have a major potential for the design and implementation of attentive UIs by creating novel and engaging ways for interacting with information systems.

So far, our research mainly focuses on image-based advertising scenarios. The results reported above show that adapting web ads based on an on-the-fly analysis of gaze behavior is feasible and effective. We demonstrated that if ads are adapted based on the dwell time extracted from the gaze behavior in real-time, it is possible to redirect the users' attention towards more specific types of ads. We showed that detailing images the users looked at could help capture the user's attention. Hence, the users' gaze behavior can help provide reactive ads, where the system takes into account the users' attention. We envision, that systems with a potentially better user experience can eventually lead to a more positive perception of advertisements.

We believe, the fact that adaptive content draws attraction independent of the interest should be highly interesting for advertisers. It implicates that new products and services could be effectively advertised if there is a link to objects users are attentive to. An example application could construct a longer chain of images leading from the user's initial attention towards an object the advertiser would like the user to look at, e.g., they looks initially at a car, then the next picture is a car in front of a house, and the following image is the house itself. We expect that such associative multi-step links may have a higher probability for the user to look at, however we have not comprehensively assessed this. Furthermore, we think that the approach described in this chapter can also be used complementary to traditional ways for targeting advertising, such as profiling.

Privacy

Such a technology poses a risk that systems acquire information about the users, which they would rather keep privately for themselves. Our approach supports a non-individualistic customization and protects the user's privacy since it does not require any data to be collected and stored about the user. We believe that this is a strong advantage over other approaches. Even though the service provider can get information about the interests of the user during their interactions with the website, no user profile is generated or stored in the process.

During the study, the participants were asked about their concerns regarding privacy when their gaze data is used to determine ads they are shown. While most of the participants were interested in this new approach, some of them stated that they would turn off this feature if this option was provided. Most of the users did not like having information about their preferences in the hands of the website owners (or in our case the UsaProxy operator). These concerns should be investigated when deploying a system relying upon the user's approval for gathering gaze data.

Opportunities for New Applications

Our study focused on the specific use case of image based ads on web pages. We chose this narrow focus as it is economically very important and by narrowing the experiments we hope to increase the reproducibility. We have so far no quantitative evidence that these findings are valid for other application areas such as images with content other than advertising or non image-based content. We expect, however, that similar effects hold for other media types and other application domains.

Additionally, we see various application domains for implicit interaction on the Web. Essentially, any web page providing content based on user profiles could benefit from the implicit interaction. One area might be a news portal, where a set of headlines is typically presented next to the main article. In this context, the gaze path could provide information about other news stories the user might be interested in and offer them alongside the current story.

CONCLUSION

In this paper we presented an eye tracking-based approach, which allows content elements on web pages to be adapted based upon a user's gaze behavior. In a case study with image-based advertisements we showed that gaze-based adaptation results in a significant increase of the users' attention, independent of his interest. Furthermore, the approach has the potential to affect the user's attitude. Our approach provides a seamless implicit interaction between users and advertisements and requires no extra effort from the users. However, privacy is a major concern. A mechanism to provide users an opportunity for opting out at anytime is essential and should be considered when deploying such systems.

We identified three potential areas of future work. First, we plan to investigate whether our approach is applicable to image-based content other than advertising. Second, we are interested in how our approach could be applied to other types of media. Third, content could be adapted based on the users' reading behavior. For example, additional content could be brought up as the users focus on a certain passage. Furthermore, text could be adapted to the user's reading abilities (e.g., text in more accessible language).

ACKNOWLEDGEMENTS

We would like to thank Dirk Kuschnerus and Stefan Legien for implementing the extension of the UsaProxy and for their help during the user study.

The research leading to these results has received funding from the European Union Seventh Framework Program (FP7/2007-2013) under grant agreement no. 244011.

REFERENCES

1. Ajanki, A., Hardoon, D., Kaski, S., Puolamaeki, K., and Shawe-Taylor, J. Can Eyes Reveal Interest? Implicit Queries from Gaze Patterns. *User Modeling and User-Adapted Interaction 19* (2009), 307–339.
2. Asteriadis, S., Tzouveli, P., Karpouzis, K., and Kollias, S. Non-Verbal Feedback on User Interest Based on Gaze Direction and Head Pose. *Semantic Media Adaptation and Personalization* (2007), 171–178.
3. Atterer, R., Wnuk, M., and Schmidt, A. Knowing the User's Every Move: User Activity Ttracking for Website Usability Evaluation and Implicit Interaction. In *Proceedings of the 15th International Conference on World Wide Web*, WWW'06, ACM (New York, NY, USA, 2006), 203–212.
4. Beymer, D., and Russell, D. M. WebGazeAnalyzer: A System for Capturing and Analyzing Web Reading Behavior Using Eye Gaze. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems*, CHI EA'05, ACM (New York, NY, USA, 2005), 1913–1916.
5. Biedert, R., Buscher, G., Schwarz, S., Hees, J., and Dengel, A. Text 2.0. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI '10 EA, ACM (New York, NY, USA, 2010), 4003–4008.
6. Burke, M., Hornof, A., Nilsen, E., and Gorman, N. High-Cost Banner Blindness: Ads Increase Perceived Workload, Hinder Visual Search, and are Forgotten. *ACM Transactions on Computer-Human Interaction (TOCHI) 12*, 4 (2005), 423–445.
7. Buscher, G., Dengel, A., and van Elst, L. Query Expansion Using Gaze-based Feedback on the Subdocument Level. In *Proceedings of the 31st International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '08, ACM (2008), 387–394.
8. Chen, M., Anderson, J., and Sohn, M. What Can a Mouse Cursor Tell Us More? Correlation of Eye/Mouse Movements on Web Browsing. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, CHI '01 EA, ACM (New York, NY, USA, 2001), 281–282.
9. Claypool, M., Le, P., Wased, M., and Brown, D. Implicit interest indicators. In *Proceedings of the 6th International Conference on Intelligent User Interfaces*, IUI'01, ACM (New York, NY, USA, 2001), 33–40.
10. Czerwinski, M., Dumais, S., Robertson, G., Dziadosz, S., Tiernan, S., and van Dantzich, M. Visualizing Implicit Queries for Information Management and Retrieval. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '99, ACM (New York, NY, USA, 1999), 560–567.
11. Drewes, H., and Schmidt, A. The MAGIC Touch: Combining MAGIC-Pointing with a Touch-Sensitive Mouse. In *Proceedings of the 12th IFIP TC 13 International Conference on Human-Computer Interaction: Part II*, INTERACT'09, Springer (2009), 415–428.
12. Duchowski, A. A breadth-first survey of eye-tracking applications. *Behavior Research Methods 34*, 4 (2002), 455–470.
13. Dumais, S., Cutrell, E., Sarin, R., and Horvitz, E. Implicit Queries (IQ) for Contextualized Search. In *Proceedings of the 27th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '04, ACM (New York, NY, USA, 2004), 594–594.
14. Farid, M., Murtagh, F., and Starck, J. Computer Display Control and Interaction Using Eye-Gaze. *Journal of the Society for Information Display 10*, 3 (2002), 289.
15. Fono, D., and Vertegaal, R. EyeWindows: Evaluation of Eye-Controlled Zooming Windows for Focus Selection. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, ACM (New York, NY, USA, 2005), 151–160.

16. Fox, S., Karnawat, K., Mydland, M., Dumais, S., and White, T. Evaluating implicit measures to improve web search. *ACM Trans. Inf. Syst. 23* (April 2005), 147–168.

17. Garson, G. D. Significance. Statnotes: Topics in Multivariate Analysis, 2011. http://faculty.chass.ncsu.edu/garson/PA765/signif.htm, last accessed 01.08.2012.

18. Garson, G. D. Student's t-test of difference of means. Statnotes: Topics in Multivariate Analysis, 2011. http://faculty.chass.ncsu.edu/garson/PA765/ttest.htm, last accessed 01.08.2012.

19. Hansen, D. W., MacKay, D. J. C., Hansen, J. P., and Nielsen, M. Eye Tracking off the Shelf. In *Proceedings of the 2004 Symposium on Eye Tracking Research & Applications*, ETRA'04, ACM (New York, NY, USA, 2004), 58–58.

20. Hardoon, D. R., Shawe-Taylor, J., Ajanki, A., PuolamŁki, K., and Kaski, S. Information Retrieval by Inferring Implicit Queries from Eye Movements. *Journal of Machine Learning Research* (2007), 179–186.

21. Hauland, G. Measuring Team Situation Awareness by Means of Eye Movement Data. In *Proceedings of HCI International 2003*, Lawrence Erlbaum Associates (2003), 230–234.

22. Just, M. A., and Carpenter, P. A. Eye fixations and cognitive processes. *Cognitive Psychology 8* (1976), 441–480.

23. Klami, A., Saunders, C., de Campos, T. E., and Kaski, S. Can relevance of images be inferred from eye movements? In *Proceeding of the 1st ACM international Conference on Multimedia Information Retrieval*, MIR'08, ACM (New York, NY, USA, 2008), 134–140.

24. Li, D., Babcock, J., and Parkhurst, D. J. openEyes: a Low-Cost Head-Mounted Eye-Tracking Solution. In *Proceedings of the 2006 Symposium on Eye Tracking Research & Applications*, ETRA'06, ACM (New York, NY, USA, 2006), 95–100.

25. Mello-Thoms, C., Nodine, C. F., and Kundel, H. L. What Attracts the Eye to the Location of Missed and Reported Breast Cancers? In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications*, ETRA '02, ACM (New York, NY, USA, 2002), 111–117.

26. Müller, J., Wilmsmann, D., Exeler, J., Buzeck, M., Schmidt, A., Jay, T., and Krüger, A. Display Blindness: The Effect of Expectations on Attention towards Digital Signage. In *Proceedings of the 7th International Conference on Pervasive Computing*, Pervasive '09, Springer-Verlag (Berlin, Heidelberg, 2009), 1–8.

27. Oppermann, R. User-Adaptive to Context-Adaptive Information Systems. *i-com 4*, 3 (2005), 4–14.

28. Petty, R., and Cacioppo, J. The Elaboration Likelihood Model of Persuasion. *Advances in Experimental Social Psychology 19*, 1 (1986), 123–205.

29. Petty, R., Cacioppo, J., and Schumann, D. Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement. *Journal of Consumer Research* (1983), 135–146.

30. Poole, A., Ball, L., and Phillips, P. In Search of Salience: A Response-Time and Eye-Movement Analysis of Bookmark Recognition. *People and Computers XVIIIDesign for Life* (2005), 363–378.

31. Poole, A., and Ball, L. J. 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).

32. Poynter Institute & Eyetools Inc. Eyetrack III: Online News Consumer Behavior in the Age of Multimedia, 2004. http://www.poynterextra.org/eyetrack2004/index.htm, last accessed 01.08.2012.

33. Qvarfordt, P., and Zhai, S. Conversing with the User Based on Eye-Gaze Patterns. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '05, ACM (New York, NY, USA, 2005), 221–230.

34. Reeder, R. W., Pirolli, P., and Card, S. K. WebEyeMapper and WebLogger: Tools for Analyzing Eye Tracking Data Collected in Web-Use Studies. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems*, CHI '01 EA, ACM (New York, NY, USA, 2001), 19–20.

35. Santella, A., Agrawala, M., DeCarlo, D., Salesin, D., and Cohen, M. Gaze-based Interaction for Semi-Automatic Photo Cropping. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI'06, ACM (New York, NY, USA, 2006), 771–780.

36. Selker, T. Visual Attentive Interfaces. *BT Technology Journal 22* (October 2004), 146–150.

37. Sibert, L. E., and Jacob, R. J. K. Evaluation of Eye Gaze Interaction. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, CHI '00, ACM (New York, NY, USA, 2000), 281–288.

38. Vertegaal, R. Introduction. *Communications of the ACM 46*, 3 (2003).

39. Vertegaal, R., Shell, J. S., Chen, D., and Mamuji, A. Designing for Augmented Attention: Towards a Framework for Attentive User Interfaces, 2006.

40. Wang, J., Zhai, S., and Su, H. Chinese input with keyboard and eye-tracking: An anatomical study. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI'01, ACM (New York, NY, USA, 2001), 349–356.

41. Zhai, S., Morimoto, C., and Ihde, S. Manual and Gaze Input Cascaded (MAGIC) Pointing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI'99, ACM (New York, NY, USA, 1999), 246–253.