Web tasks classification based on the dominant visual mechanism identification

Martinho Gonçalves
Universidade de Trás-os-Montes e Alto Douro; Instituto de Engenharia de Sistemas e Computadores do Porto
martinhofg@gmail.com

Maximino Bessa
Universidade de Trás-os-Montes e Alto Douro; Instituto de Engenharia de Sistemas e Computadores do Porto; Warwick digital lab. Institute
maxbessa@utad.pt

Telmo Adão
Universidade de Trás-os-Montes e Alto Douro; Instituto de Engenharia de Sistemas e Computadores do Porto
telmo.adao@gmail.com

Luís Magalhães
Universidade de Trás-os-Montes e Alto Douro; Instituto de Engenharia de Sistemas e Computadores do Porto
lmagalha@utad.pt

Alan Chalmers
Warwick digital lab. Institute
alan.chalmers@warwick.ac.uk

ABSTRACT

The internet contains a wide variety of websites in terms of the information that is presented, their appearance, and layouts that can range from simple to highly complex. When creating new layouts, designers always try to capture the attention of the viewer to important areas of the website. However, the important areas where users will concentrate their attention often depend on the task they are currently focussing on within the website. Knowledge as to how the human visual system (HVS) works and what activity/task each user is currently performing on the website can play a significant role in allowing developers to create more efficient, targeted web pages.

In this paper we investigate if it is possible to know what a user, at a specific time, is doing in a website. A study of a user's visual behaviour when navigating different web pages, is conducted by means of psychophysical experiments. These results are used to determine what action (namely performing a task or just exploring the website) users are performing at different points in time in a specific website. The results show that different tasks do indeed exhibit different visual patterns and it is thus possible to determine if a viewer is performing a task or simply viewing a webpage.

KEY WORDS
Eye tracking, Visual attention, Internet.

1. INTRODUCTION

The use of the Internet to access information has become essential for people around the world. Some companies devote a lot of resources to try to build the perfect layout in order to guarantee that their customers/visitors can find what they want/need and have a great experience while visiting their website. However, accomplishing this is not an easy task primarily because depending on the task a user is currently performing, when browsing websites, different areas will capture his/her attention.

William James wrote one of the most prevalent definitions of visual attention: "…is the taking possession by the mind in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought…" [James90]. Selective visual attention is the process by which one selects a portion of the visual information available to process. The visual attention is controlled by two main mechanisms, namely: Bottom-up, where one's attention can be captured involuntary through a visual stimulus, and Top-down where one's attention is captured depending on the task that is being performed and where all the long-term cognitive strategies take a real and important role [Yarbus67]. By knowing what users are doing at a given moment on the website, developers will be able to optimize the web page content (for example, by changing the layout), make the web page more usable, and predict where the advertising should be placed and/or be confident it will be seen. This will enable the most important information to be placed in these locations.

The purpose of the study presented in this paper is to verify the existence of differences in visual patterns when

surfing the Internet, depending on the tasks that are being performed. Although it is possible to state that the tasks are different, the same does not happen when it comes to identifying which kind of task the user is performing. We therefore present a study to investigate the possibility of distinguishing a user's visual behaviour when she/he is simply exploring a web page or performing a task on the same web page. In order to accomplish this, we conducted an experiment with subjects whose visual patterns, when performing different tasks, were compared with saliency maps. Saliency maps are grey scale images which specify the order in which parts of a scene are likely to be examined by the Human Visual System (HVS). Areas shown in white are attended to first [Itti00].

The paper is structured as follows: In the next section, we present the most relevant previous research and a summary of the experiment that forms the foundation of this work is given [Gonçalves11]. Section 3 details the main experiment that was conducted to investigate users' visual patterns and their visual mechanisms. Finally, in section 4, we give conclusions and discuss areas for future work.

2. BACKGROUND

The process that allows subjects to quickly move their eye position directly to the important areas in the visual field is called Selective Visual Attention. The visual attention can be captured automatically by the stimulus, using the bottom-up or top-down mechanism, which depends on the context [James90].

Bottom-up mechanisms are thought to operate on raw sensory input, where attention is rapidly and involuntarily driven to salient visual features [Itti00]. Within the bottom-up process, the visual stimulus automatically captures attention without making a conscious choice to attend to features of potential importance. A Saliency Map [Koch85] can provide information about salient, ie. visually important, areas that will attract one's attention in an involuntary way.

On the other hand, when we tell a subject to perform a specific task she/he will be driven by the top-down visual mechanism [Yarbus67]. Top-down mechanisms are related to a subject's long-term cognitive strategies, where knowledge and context play a crucial role on how one selects where to pay attention to. Yarbus showed that the way one perceives an environment depends on who he is and on the task that he is currently performing in that environment. In one of his classic experiments, Yarbus recorded eye movements, fixations and saccades (series of movements of the eye, as it changes its focus moving from one point to another [Saccades12]), of observers while inspecting Repin's painting, "The Unexpected Visitor"[Yarbus67]. Before participants had started the test, they were instructed to perform one of the seven tasks. Although all observers were presented with the same stimulus (Repin's painting), the patterns of eye movements registered were drastically different, and depended on the task they were performing.

Several studies have recently been conducted on users' visual behaviour when navigating web pages. Poole and Ball [Poole06] studied the usability and interaction between humans and computer using an eye tracking device. With the eye tracker results it is possible not only to quantify an individual's interaction with the computer, but also obtain other data, such as its meaning when relating the results with websites. Poole and Ball identified variables and factors, from evidence found in other studies, that indicated both the time that individuals take to perform certain tasks, the possibility of different task complexities, the difficulty in obtaining information, and the structuring of web pages.

Additionally, several researchers have tackled the issue of searching on the Internet using search engines, as this is one of the most performed tasks. Hsieh-Yee [Hsieh-Yee01] conducted a study on web search behaviour, through a literature review from 1995 to 2000 that included most of the studies on children and adults. The author concluded that the majority of the studies that focus on children only described their interaction with the web, while in the studies addressing the adults, the researchers' goals mainly described patterns.

What distinguishes the research on adults' web search behaviour is the use of multiple methods when collecting data collection. Research on web search behaviour requires that there is a commitment to examine users in their natural environmental conditions as well as a rigorous design and data analysis. Although these studies have an increasingly rigorous objective, they still remain poorly accepted and validated. To verify that the difference in age does influence the success level in conducting searches on the Internet, Bilal and Kirby conducted a study with children and adults [Bilal02]. In this study, they concluded that the search on the Internet is affected by the following: the users' cognitive capacity, both affective and physical states, since the user gives a different answer depending on the type of task being performed. Additionally, the study highlighted the ability of adults to reformulate the search when faced with an impasse, the type of navigation used throughout the pages, and concentration while performing a task. The cultural experience of the adults is one of the factors that most distinguishes them from children, resulting in a higher success rate for the adults.

In 2004 a study was conducted in which Granka et al. [Granka04] analysed the behaviour of individuals when they performed web searches, regarding how they interacted with the sites and the results they consequently obtained. Having as reference the time that the individuals take to select the desired link, this study has revealed the complexity of the mental process involved in the choice of the link, as well as the complexity of the information presented to the users. The order, in which they appear in the search results, is related to the time during which users pay "attention" to the links therein. Therefore, the search is automatically limited to the behaviour of the search engine and how it presents its results.

Lorigo et al. [Lorigo08] also studied online searching and presented knowledge drawn from the experiments and the difficulties encountered. By comparing different search engines, they verified that all the metrics used did not show significantly different values and thus the search in the different search engines was similar in the visual process and in the complexity of reasoning. Moreover, the individuals' preferences concerning the search engines used were different. The main difficulties identified were: interpreting the eye tracker data, analysing the interaction between individuals and the search results, and integrating eye tracker studies within usability studies. The lack of feedback and of the task's success analysis, limits the types of studies that can be done with eye trackers.

Pan et al. [Pan04] investigated the importance of visual behaviour on web pages, analysing the order in which pages were visited, as well as their complexity. They concluded that the human visual behaviour while viewing web pages results from the order in which pages are viewed, of the relationship between the site, and gender. The results of scan paths (path that the eyes follow during the viewing of web pages) showed that there is the possibility of relating the visual behaviour of different individuals with the visual structure of web pages.

Regarding the visual mechanisms, Still and Masciocchi [Still10] developed an experiment meant to verify whether a saliency model can predict or not the fixations in a web page. Taking into account the ten first fixations of each user in a web page image, they compared them with the corresponding saliency maps. They concluded that the saliency model indeed predicts the visual attention focus.

In their 2009 study Georg Buscher et al, presented an eye-tracking experiment where they first wanted to verify what areas of the page were most important for each task when a user surfed the web pages with tasks or only exploring the page [Georg09]. In a second phase they tried to develop a computational model that predicts the visual salience of the elements of a web page. They concluded that the saliency maps can predict the deployment of the users' visual attention within a web page.

3. BOTTOM UP OR TOP-DOWN APPROACH WHEN BROWSING WEB PAGES

Pervious work [Gonçalves11] showed that there are differences in the user's visual patterns when they perform various tasks. Although it is possible to state that the tasks are different, what kind of task the user is performing is unclear. Consequently, we present an experiment to investigate the possibility of distinguishing the user's visual behaviour when she/he is simply exploring a web page or performing a task on the same web page. To accomplish this we conducted another experiment with subjects whose visual patterns were recorded while performing different tasks in different web pages. These visual patterns were then compared with the computed saliency maps of each of the web pages.

Based on the different percentage levels of the visual pattern matching with the saliency map we planned to verify if it is possible to both quantify and distinguish if the users are accomplishing a task or just exploring the website. For this experiment we considered the following null hypothesis:

H0: The percentage of image covered by the visual pattern in the saliency maps, and the value of hotspots per page are the same.

3.1 PARTICIPANTS

In this test we took part 16 participants who were all University students. This set comprised 8 female and 8 male participants. The participants' average age was between 22 and 25 years old. 10 of the participants had normal vision and 6 had corrected to normal vision. All the participants use the Internet on a daily basis.

3.2 METRICS

The main metrics used in eye movement research are: fixations and "saccades", which in turn occur between anchorages. There are also several metrics deriving from these key measures, including the size of the "gaze" and the distance travelled, known as "scan path". Aspects such as the dilation of pupils and the number of times that the subject blinks his/her eyes are also possible objects of study [Salvucci00].

In this experiment we have selected the metrics corresponding to the fixations: the points where vision focuses. With these fixations, the eye tracker software automatically generates hotpots, thus creating a "heat image" as to where the user was focusing. By using this metric, we can observe and indeed verify the differences in the viewing patterns. Moreover studying the number of hotspots per page visited provides a quantitative measure regarding the dispersion of the visual field.

3.3 PROCEDURE

All experiments were conducted individually and in a controlled environment. The room was equipped with a Computer and an eye tracker (Tobii X50), in combination with Clear View 2.0.1 software [ClearView11] to collect eye movements' data. The experiments were displayed on a 17" screen with a resolution of 1024 × 768 pixels. At the end of each task, the application automatically saved all the information. All the saliency maps for each web page tested were generated using the Ezvision software [Ezvision11]. To analyse the data provided by the eye tracking system and the saliency map we used MatLab.

The experiment comprised six different tasks in three different websites namely: online journals, e-commerce and company websites. All participants performed one task in each type of website.

The assigned Task and websites were:

Task 1 (T1): Explore a newspaper web page [Expresso12];

Task 2 (T2) Find specific news on a newspaper website [Expresso12];

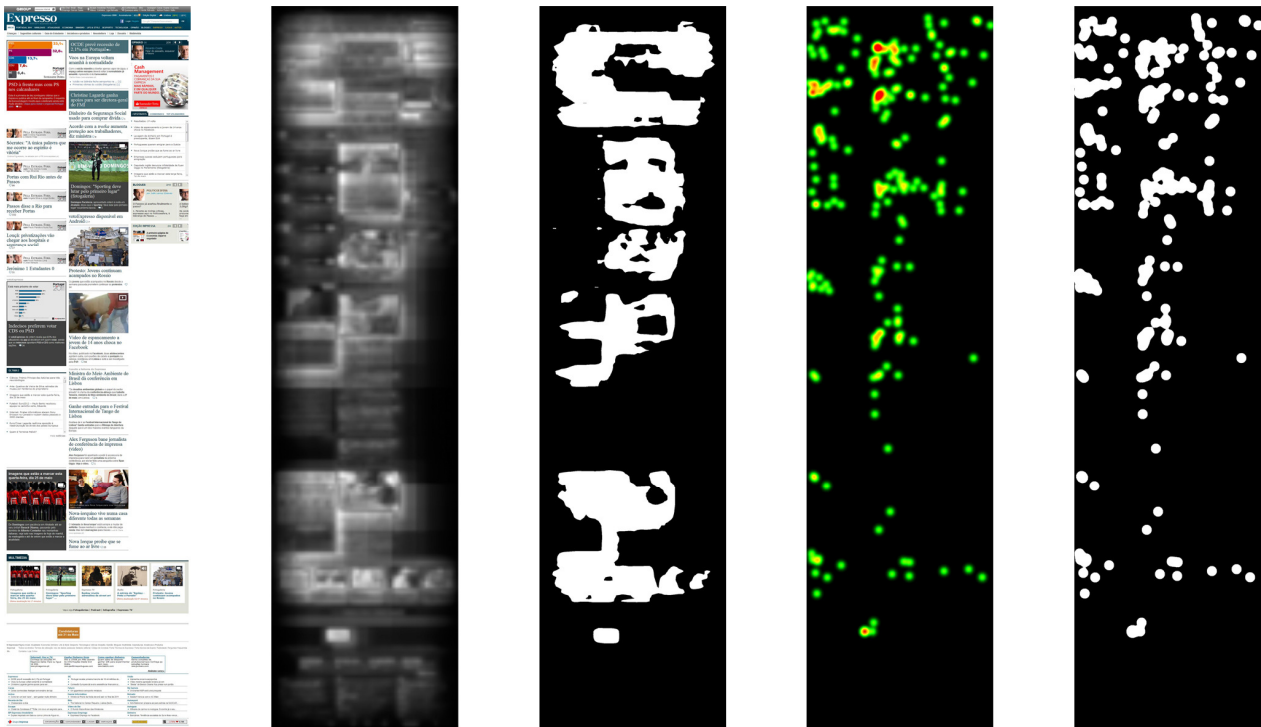Task 3 (T3): Search for specific publicity on an e-commerce web page [Miau12];

Fig. 1: Example of the images used to compare the visual pattern of subject number six with the saliency map of a web page. a) Original web page image, b) Saliency Map generated using Ezvision, c) Dichromatic image generated in MatLab from image b), d) Hotspots image from user visual patterns and e) Dichromatic image generated in MatLab from image d).

Task 4 (T4): Explore an e-commerce website Miau [Miau12];

Task 5 (T5): Explore a corporate website Microsoft [Microsoft12];

Task 6 (T6): Search for a specific web page on a corporate website [Microsoft12];

In a previous work [Gonçalves11] the task definition and web sites selection were based on the methodology used by [Pan04], but after a further analysis we observed that the mental process beneath those different tasks (ex. search for a news, buy a product, get the contact info of a company), is driven by a simple find/search task. Based on this we start to study this simple task, and later we plan to study more complex tasks. This web sites were selected to insure a representation of different web pages that can be found on the internet and all also being well-known.

After completing the tasks, the participants answered a short survey to collect their information such as age, gender, eyesight conditions and web navigation experience.

3.4 ANALYSES AND RESULTS

In order to verify the integrity of the information collected by the eye tracker, the existence of outliers in the data was investigated. Based on this, since there were some missing data from the visual patterns caused by some user's behaviour when performing the tasks that affected the eye tracker's ability to capture their data, it was decided to remove the data of subject 7 from T1, the data

relating to subject 12 from T3, the data relating to subjects 2 and 7 from T4, and the data relating to the subjects 2, 5 and 7 from T5. After this outlier removal, data was analysed to see if differences in the subjects' visual patterns compared to the Saliency Maps existed or not, H0.

In order to accomplish this we computed the saliency map, Fig.1 b of the original web page, Fig.1 a using Ezvision software [Ezvision11] and then we used MatLab to generate a dichromatic image of it, Fig.1 c. The hotspots image obtained from each user's visual patterns, Fig.1 d, was also converted to a dichromatic image with MatLab, for example Fig.1 e. After generating all images a comparison was made between the dichromatic images of the saliency maps, Fig.1 c and the hotspots, Fig.1 e.

From this comparison we considered two variables: the percentage of hotspots that matches saliency points and the percentage of hotspots area per page. This process was repeated for the six different Tasks, to test the null hypothesis, H0.

Asymp. Sig. (2-tailed)	% hotspot in salmap	% hotspot/page
Web page 1 (t1 vs t2)	0,000	0.000
Web page 2 (t3 vs t4)	0,001	0,001
Web page 3 (t5 vs t6)	0,030	0,003
Explore / Task	0,000	0,000

Table **1:** Nonparametric Test (2 Independent Samples)**.**

As can be seen in Table 1, for the comparisons between exploring a web page and completing a task on the web page 1, web pages 2 and 3 have a significance level less than 0.05, rejecting the null hypothesis. This indicates significant differences in the percentage of saliency map

covered by the users' visual patterns when performing a task or exploring a web page. In addition, when comparing the task, independently from the web page, the significance level is less than 0.05, also rejecting the null hypotheses.
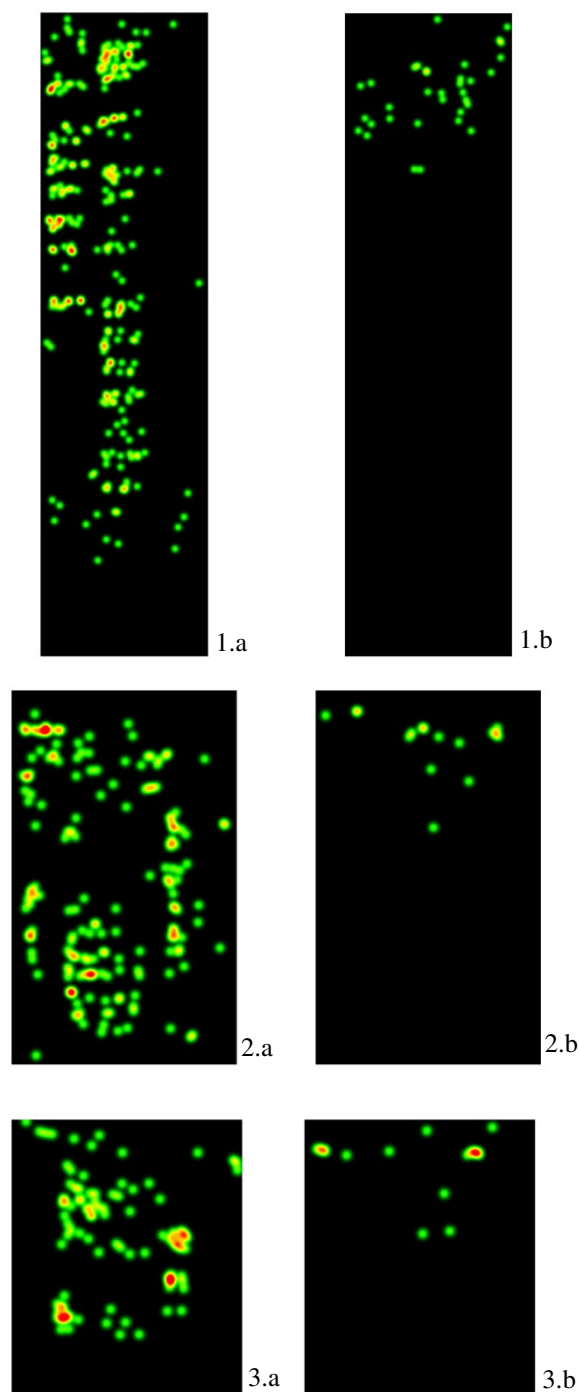
Fig. 2: 1.a visual patterns of one subject while exploring in Expresso; 1.b visual patterns of one subject while performing a task in Expresso; 2.a visual patterns of one subject while exploring in Miau; 2.b visual patterns of one subject while performing a task in Miau; 3.a visual patterns of one subject while exploring in Microsoft; 3.b visual patterns of one subject while performing a task in Microsoft. The disparity of the images height and width results from the need to represent the complete web page.

The statistical analyses have shown that there are distinct differences in the percentage of a user's visual patterns that overlaps with the web page saliency map. In addition, we also considered the hypotheses that if an image of a web page, that is salient, is related to the task that a user is performing, and, if the users' time to complete the task was very short, then the user's visual patterns were analogous to the saliency maps when exploring a web page. We therefore analysed the percentage of hotspots per page considering the following null hypothesis:

H0: The percentage of hotspot per page is the same.

The comparisons show that the percentage of hotspots per page when exploring a webpage or completing a task have a significance level lesser than 0.05. Thus the null hypothesis is rejected indicating significant differences in the percentage of hotspots per page.

Furthermore, when comparing the task, independently from the web page, the significance level is also less than 0.05. The null hypotheses is also rejected for the percentage of hotspots per page. These results suggest there are clear differences in the dominant visual mechanism (bottom-up or top-down) when the users were completing the task or exploring a web page, when considering as the value of the hotspots similar to the saliency maps and the hotpots per page.

Figure 2 shows the images of the hotspots generated by the participants' visual patterns while exploring a web page and completing a task, for each website. A similar kind of behaviour was observed in all participants. One can see clear differences in the dispersion and the number of fixations when the user is respectively exploring (Fig 2: 1.a, 2.a, 3.a) and completing a task (Fig. 2: 1.b, 2.b, 3.c). This information supports the statistical information previously presented.

4. CONCLUSIONS

The purpose of this study was to verify the existence of differences in visual patterns in the navigation of web pages, depending on the tasks that were being performed.

This study investigated the possibility of distinguishing a user's visual behaviour when she/he is simply exploring a web page or performing a task on the very same web page. This was done by considering whether the visual mechanism, bottom-up or top-down, is more dominant .The statistical analysis, Table 1, showed clear differences (significance value less than 0.05) with respect to the users' visual patterns being similar to the saliency maps when completing a task or simply exploring a web page. Additionally, there are also clear differences (significance value less than 0.05) in the number of hotspots per page. The hotspot maps (Fig. 2) generated from the participants' visual patterns support the statistical results (Table 1). We can conclude that there are differences in the visual patterns when users perform a task or simply explore a web page.

This study shows that it is possible to determine if a particular user is exploring or searching for something in a given web page by using their visual patterns and knowledge about how the visual mechanisms operates. Future work will investigate the possibility of automate this process, and identify in real time if the user is performing a task or simply exploring the web page. With this data it is possible to bring some improvements to the webpage. In real time it becomes possible, for example, to change the web page structure and rearrange the information to improve the perception of its contents so it can be more effective, to help the user finding the information that he is looking for suggesting paths or information about it or even improve a web assistant by identifying what information is the user looking for.

5. ACKNOWLEDGEMENTS

6. REFERENCES

[Bilal02] BILAL, D., AND KIRBY, J. 2002. Differences and similarities in information seeking: children and adults as Web users, Information Processing and Management 38, 649–670.

[ClearView11] Clearview 2011 software. http://www.tobii.com/eye-tracking-research/global/products/software/

[Expresso12] Expresso 2012. Journal Expresso Website. http://aeiou.expresso.pt/ accessed 2012

[Ezvision11] EZVISION 2011. software - http://ilab.usc.edu/toolkit/screenshots.shtml

[Georg09] BUSCHER, G., CUTRELL, E. AND MORRIS. M. R. 2009. What do you see when you're surfing?: using eye tracking to predict salient regions of web pages. In Proceedings of the 27th international conference on Human factors in computing systems (CHI '09). ACM, New York, NY, USA, 21-30.

[Gonçalves11] GONÇALVES, M., ROCHA, T., MAGALHÃES, L., PERES, E., BESSA, M., CHALMERS, A. 2011. "Identifying Different Visual Patterns in Web Users Behaviour", In Proceedings, Spring Conference on Computer Graphics 93-97, Viničné, Slovakia, April.

[Granka04] GRANKA, L. A., JOACHIMS, T., GAY, G. 2004. Eye-tracking analysis of user behavior in WWW search, Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval, July 25-29, Sheffield, United Kingdom

[Hsieh-Yee01] HSIEH -YEE, I. 2001. Research on Web search behavior, Library & Information Science Research 23:167–185.

[Itti00] ITTI, L. 2000. Models of Bottom-Up and Top-Down Visual Attention. PhD thesis, California Institute of Technology, Pasadena, California, Jan 2000.

[James90] JAMES, W.. 1890. The principles of psychology. Harvard. (Harvard ed. 1983).

[Koch85] KOCH, C. AND ULLMAN, S. 1985. Shifts in selective visual attention: towards the underlying neural circuitry. Hum Neurobiol, 4(4):219{227, 1985}.

[Lorigo08] LORIGO, L., HARIDASAN, M., BRYNJARSDÓTTIR, H., XIA, L., JOACHIMS, T., GAY, G., GRANKA, L., PELLACINI, F., AND PAN, B. 2008. Eye tracking and online search: Lessons learned and challenges ahead. J. Am. Soc. Inf. Sci. Technol. 59, 7 (May. 2008), 1041-1052. DOI= http://dx.doi.org/10.1002/asi.v59:7

[Miau12] Miau 2012. Miau Website. http://miau.clix.pt/home.jsp

[Microsoft12] Microsoft 2012. Microsoft Company Website. http://www.microsoft.com/pt/pt/default.aspx

[Pan04] PAN, B., HEMBROOKE, H. A., GAY, G. K., GRANKA, L. A., FEUSNER, M. K., AND NEWMAN, J. K. 2004. The determinants of web page viewing behavior: an eye-tracking study. In Proceedings of the 2004 Symposium on Eye Tracking Research & Applications (San Antonio, Texas, March 22 - 24, 2004). ETRA '04. ACM, New York, NY, 147-154. DOI= http://doi.acm.org/10.1145/968363.968391

[Poole 06] POOLE, A. & BALL, L. J. 2006. Eye Tracking in Human-Computer Interaction and Usability Research: Current Status and Future Prospects. In Ghaoui, Claude (Ed.). Encyclopedia of Human Computer Interaction. Idea Group.

[Saccades12] SACCADES 2012. Dictionary. http://dictionary.reference.com/browse/saccade.

[Salvucci00] SALVUCCI, D. D., & GOLDBERG, J. H. 2000. Identifying fixations and saccades in eye-tracking protocols. In Proceedings of the Eye Tracking Research and Applications Symposium, 71-78.

[Still10] STILL, J. D. & MASCIOCCHI, C. M. 2010. A saliency model predicts fixations in web interfaces. Paper presented at the Computer-Human Interaction Conference, 5th International Workshop on Model-Driven Development of Advanced User Interfaces, Atlanta, GA.

[Yarbus67] YARBUS, A.L. 1967. Eye movements during perception of complex objects. Eye Movements and Vision, Chapter VII, 171-196