

## Towards Fully-Accessible e-Learning Systems: Some Eye Tracking Solutions

Virginio Cantoni, Marco Porta, Alice Ravarelli

**Abstract:** *With the ever increasing diffusion of e-learning in present society, Accessibility becomes more and more an issue, prompting the need for effective technological answers. In this paper we consider the application of eye tracking to e-learning, as a way to improve the machine-learner interaction process – as well as to make up for the missing capabilities of disabled learners. Detecting the user's gaze direction can in fact be extremely useful in presence of severe disabilities, which prevent normal use of keyboard and mouse. In this work we describe some of the projects we have recently developed at the University of Pavia, aimed at improving assistive technology for e-learning through eye-controlled interfaces.*

**Key words:** *Accessibility, E-Learning, Eye Browsing, Eye Tracking, Eye Writing, Perceptive Interfaces.*

### INTRODUCTION

Machine perception can make Human-Computer Interaction more natural, imitating the ways we usually interact each other in interpersonal communication. If this is true in general, the e-learning sphere can particularly benefit from perceptive capabilities: by becoming more “human-like”, a computer able to recognize what users do or what happens around them may also improve the learning process, through the enhancement of interaction procedures [2].

According to the World Health Organization, about six hundred million persons in the world are affected by disabilities of some kind, due to several causes [14]. Even if not all of these people have disabilities impeding computer or Internet access, it is however a great part of the world population. Nevertheless, nowadays practically all existing e-learning systems can only be navigated with the mouse, and are characterized by several other accessibility lacks. In the context of Information Technology, Accessibility deals with issues like these, to reduce (or even eliminate) any gaps between healthy and disabled people in information access. On the European side, the European Commission strongly promotes several initiatives about Accessibility. For instance, *eAccessibility* – within the e-Inclusion activity – aims at ensuring people with disabilities and elderly people access ICTs on an equal basis with others. Also, the new *i2010* initiative on e-Inclusion (“To be part of the Information Society”) includes a strategy targeted on enhancing accessibility to the Information Society for all potentially disadvantaged groups [3].

In recent years, eye tracking systems [4] have greatly improved, beginning to play an important role in the assistive technology field. Eye tracking relates to the capability of some devices to detect and measure eye movements, with the aim to precisely identify the user's gaze direction (usually on a screen). The acquired data are then recorded for subsequent use, or directly exploited to provide commands to the computer. The eye tracker we use in our laboratory – the Tobii 1750, one of the most widespread eye tracking devices – combines video-oculography with infrared light reflection. The system, which looks like a common LCD screen, is provided with five NIR-LED and an infrared CCD camera, integrated in the monitor case. Infrared light generates corneal reflections whose locations are connected to gaze direction. Eye positions are recorded at a frequency of 50 Hz, and with a precision of about 0.5° of the visual field.

Eye tracking can be profitably exploited for *evaluation* purposes, allowing a great amount of data about users' behaviours to be obtained – for example, while interacting with e-learning systems. Thanks to the acquired information, users' gaze paths can be

reconstructed, and possible focuses of attention identified. Analogously, eye tracking techniques are also valuable for advertising and marketing purposes.

However, it is when eye tracking is used as a direct input source for the computer that assistive interfaces can be built. These tools can significantly improve the accessibility of software interfaces, making them also usable by disabled people. Using a keyboard and a mouse is in fact impossible in all those cases where serious motor impairments – both temporary and permanent – impede easy hand movements. In these situations, the eyes may be the only communication channel for an individual.

Several eye-controlled systems have been developed to date, but there are still some open questions connected with human-computer interaction, and with accessibility in particular. Our activity is aimed at creating more accessible and usable interfaces based on eye tracking, which could make e-learning really available to everybody.

## **SOME EYE TRACKING PROJECTS AT THE UNIVERSITY OF PAVIA**

### **Eye-S: a Full-Screen Input Modality for Pure Eye-based Communication**

*Eye-S* is a system for pure eye-based communication, allowing the input of both text and generic commands [12]. E-learning systems often require some kind of input from the user, which however cannot be provided by a severely-disabled person.

The problem of writing through the eyes has been widely considered in the past, and a number of solutions have been proposed. The simplest approach is surely that based on on-screen keyboards and dwell time: if the user looks at a certain key for more than a predefined time interval, the key is considered pressed and the corresponding letter is typed. Several studies have been carried out connected to this kind of typing and related issues (e.g. [5] and [6], focused on the efficacy of dwell time and the importance of feedback). There are also variants where the dwell time is substituted with some kind of switches, such as physical buttons, eye blinking (e.g. [13]) or facial muscle activation.

In the context of action-based gaze input (for example to perform drag and drop operations), Milekic [7] introduced the concept of “eye graffiti”, where gaze gestures are used to form a vocabulary in a way similar to the text input mechanism used in personal organizers. In these devices, natural input is obtained by “drawing” letters, or parts of them, through a pen. Our work derives basically from a similar principle.

Compared to other eye writing systems, *Eye-S* has a main advantage: it leaves the screen totally free for the display of applications, since there is no need for specific graphical input interfaces. Using an “eye gesture” approach, the user creates alphabet letters, as well as punctuation marks or specific commands definable according to the particular application to control, by means of sequences of fixations on nine (hidden) predefined areas on the screen (Fig. 1a) – “squares” (called hotspots) placed in easily identifiable positions, even if not explicitly displayed: the four vertices of the screen, the middle of each side and the center.

Normally, hotspots are not visible. However, their position is evident, and almost no effort is required to the user to remember where they are. This means that the screen is totally available for displaying any content, and no area is occupied by specific interfaces for text input.

An *eye sequence* (from which the name *Eye-S* stems) is a succession of fixations on the hotspots. When the user looks at a hotspot for more than a given threshold (e.g. 400 milliseconds), a sequence recognition process starts. If other hotspots are looked at after the initial one within configurable time intervals, and if the succession of watched hotspots pertains to a set of predefined sequences stored in a “database”, then a corresponding action is performed. If the system is being used for text input, the action will be the same as typing a key on a keyboard.

Eye sequences can be chosen arbitrarily, but in the writing context they will of course resemble the form of letters (possible eye sequences for the 'a' and 'b' characters are shown in Figure 1b).

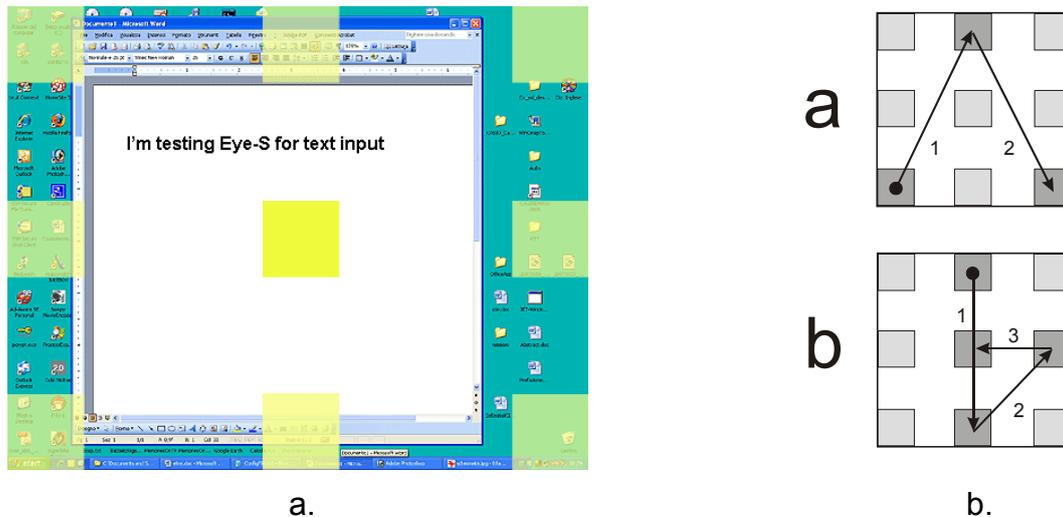


Figure 1. Eye-S: (a) Explicit display of hotspots on the screen; and (b) Examples of eye sequences for the 'a' and 'b' letters

During system use, it is generally helpful to get a feedback about the sequence composition process. To this purpose, when the user looks at the first hotspot for more than the defined dwell time, a small green square is displayed within the hotspot itself. Such square contains a '1', to indicate that this is the first hotspot of a possible sequence. If the user looks at another hotspot within a timeout, then a yellow square appears, with a '2' written inside it (and the green square is deleted). If the sequence which is being recognized is three segments long, the same happens for the third hotspot (orange square and '3' as a sequence indicator). At last, on the final hotspot of a sequence — whether it is three or four segments long — a red square is displayed which contains the character or "action" recognized, in order for the user to immediately understand that the eye gesture has been successfully detected.

### WeyeB: an Eye-Controlled Web Browser Hands-Free Navigation

WeyeB (from *Web eye Browser*) is an eye-controlled Web browser [11] enabling the two basic activities required when surfing the Web, namely page scrolling and link selection (a very common activity in e-learning systems, usually based on the Web and/or on hypermedia interaction styles).

Page scrolling can be activated by simply looking at one of two buttons, placed above and below the display area in the WeyeB interface. If the "watching time" (the dwell time) goes beyond a predefined value (e.g. 1 second), then the button is considered pressed. When this occurs, a semitransparent scroll rectangle is displayed over the page (Figure 2a). By looking within the scroll rectangle, the content can be scrolled, with a speed that is lower (or even null) when the user's gaze is in the central area of the rectangle, and increases progressively as it moves towards the upper or lower edges. The scroll rectangle is displayed as long as the user's gaze is detected within it. Once the user looks anywhere outside the rectangle for more than a certain time (2 seconds in our experiments), it disappears. This page scroll solution gives the user full control: no scrolling can occur accidentally, since the scroll rectangle must be displayed beforehand, and the explicit display of such a graphical element represents an important visual feedback.

For the selection of a link, the user must look at it for (at least) a dwell time. Subsequently, his or her gaze must be rapidly shifted upward, and then immediately downward, again (about) on the target link. This sort of "eye gesture", unlike other

techniques, does not interfere with normal reading activities.



a. b.  
 Figure 2. WeyeB: (a) Scroll rectangle; and  
 (b) Link menu displayed after the acquisition of the snapshot

The detection of a selection gesture causes a “snapshot” of the area around the link to be acquired – the precision of current eye trackers is in fact limited, and the perceived gaze may not be exactly centered on the link. The rectangular region (which has a fixed size, e.g. 200 x 100 pixels, and is centered on the fixation point) is stored in the form of a bitmap image. Such image is then analyzed with OCR techniques to extract its textual content and identify possible (textual) links present in it (through a comparison with the content of all the <a> tags in the HTML code). If the snapshot contains exactly one link, the corresponding page is loaded. If it includes more than one link, a popup menu is shown which lists all the recognized links, easily selectable by gaze (Figure 2b). Lastly, if no textual links are identified in the snapshot (e.g. because it contains an image link), the mouse cursor is simply shifted to the initial selection point and a mouse click is generated.

**e5Learning: an E-Learning Environment Based on Eye Tracking**

*e5Learning* [1][10], whose name stems from enhanced exploitation of eyes for effective eLearning, is an e-learning environment where eye tracking is exploited to allow the computer to get valuable data about users and their activities. Even if its functions are not directly related to accessibility issues, e5Learning can significantly improve the quality of distance learning, which is often the only possibility for a disabled person.

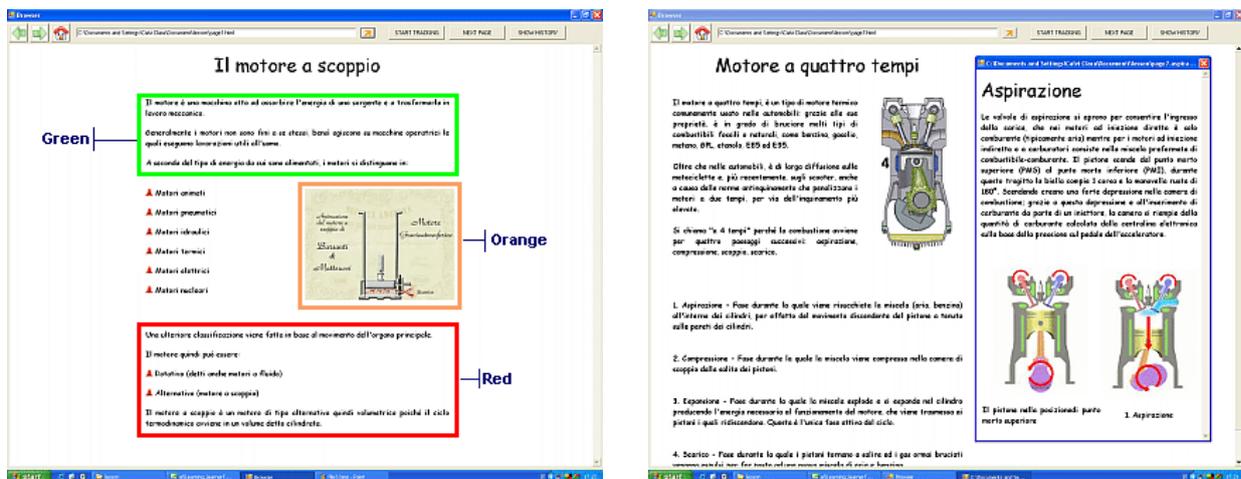
The system is characterized by three key functionalities: (1) detection of basic user activities, such as reading text and observing multimedia content, in order to maintain a "history" of user actions; (2) generation of additional content depending on the context (e.g. textual or multimedia descriptions shown when the user is reading a specific portion of text); (3) recognition of stress, high workload and tiredness states in the user, using physiological data obtainable from the eye tracker.

Thanks to the *Monitor of Accessed Screen Areas*, the author of the course can decide “how much attention” the user should pay to certain portions of content. In our prototype, a course is simply made up of web pages. We use an ad-hoc-built web browser which, along with page content, reads additional information defined by the author. Among other things, such information specifies the coordinates and sizes of screen rectangles (Regions of Interest, or Rols) corresponding to relevant portions of content, and associated data. The *History Recorder* submodule relies on the Monitor of Accessed Screen Areas and keeps track of which portions of content (Rols defined by the author) have already been accessed by the user, as well as “how much”. If, for example, in a previous session the user did not devote sufficient time to a certain area, its content might be subsequently proposed before others, independently of its position

in the logical structure of the course. Another strategy, which is the one we have actually implemented, explicitly highlights the regions which need attention. When the user presses the 'next' button to load the next page in the course, if in the current page there are ROIs that have not been fully read/observed, the system emphasizes them through colored rectangles (Fig. 3a).

Using the *Contextual Content Generator*, the creator of the course can associate new content to ROIs, and indicate the requirements for the additional information to be displayed (in the form of HTML pages appearing within a popup window, as shown in Figure 3b). A condition for the new window to be shown is that the fixation time within a ROI is higher than a threshold.

A third module composing e5Learning is the *Emotion Recognizer*. Several experiments, mainly carried out in the Psychology and Physiology fields, have demonstrated that the observation of eye behaviours can reveal much information about the user emotional state. For instance, pupil size is significantly larger after highly arousing stimuli than after neutral stimuli [9]. Other investigations suggest that the mental workload can be assessed by analyzing the fluctuation rhythm of the pupil area [8]. In particular, in our project we have considered eye factors to (potentially) identify two main user conditions: (a) high workload or non understanding, and (b) tiredness. For instance, if the average pupil size has progressively increased within a certain time interval, also user workload may have augmented. A decreased blink rate in the same period would further confirm such a supposition. When detected, these evidences could be used to dynamically modify the learning path, proposing a topic related to the main one but less complex (a sort of "break"). Or, if the user is potentially having problems in understanding something, extra information may be displayed.



a.

b.

Figure 3. *e5Learning*: (a) Coloured rectangles highlighting ROIs; and (b) Additional content displayed when the user looks at a specific ROI

## CONCLUSIONS

In this paper we have presented some eye tracking systems that can be applied to e-learning to create new advanced and accessible computer-based teaching systems (also able to better “understand” the user). Eye tracking is now an extraordinary opportunity for assistive technology, enabling new advanced solutions for human-computer interaction. Although, due to their very restricted market, prices of eye trackers are currently very high, things may totally change when these devices will start to spread, hopefully in the next few years, and current research in this field can certainly help such expectation become a reality. We strongly think that *accessible e-learning*

should become a priority in the near future: present technology is ripe to make education really available to *everybody*.

### **ACKNOWLEDGEMENTS**

This work was supported by funds from the Italian FIRB project "Software and Communication Platforms for High-Performance Collaborative Grid" (grant RBIN043TKY).

### **REFERENCES**

- [1] Calvi, C., Porta, M., Sacchi, D. e5Learning, an E-Learning Environment Based on Eye Tracking. Proceedings of the 8<sup>th</sup> IEEE International Conference on Advanced Learning Technologies (ICALT 2008), Santander, Spain, 2008.
- [2] Cantoni, V., Cellario, M., Porta, M. Perspectives and Challenges in E-Learning: Towards Natural Interaction Paradigms. Journal of Visual Languages and Computing, Elsevier Science Publishing, 15 (2004), pp. 333-345.
- [3] Communication "European i2010 initiative on e-Inclusion - to be part of the information society" – Europe's Information Society. Retrieved July 11, 2009, from [http://ec.europa.eu/information\\_society/activities/einclusion/bepartofit/overview/](http://ec.europa.eu/information_society/activities/einclusion/bepartofit/overview/).
- [4] Duchowski, A. T. Eye Tracking Methodology – Theory and Practice (2<sup>nd</sup> Ed.). London: Springer-Verlag, 2007.
- [5] Majaranta, P., Aula, A., Rähkä, K. Effects of Feed-back on Eye Typing with a Short Dwell Time. Proceedings of ETRA 2004, San Antonio, Texas, USA, March 22-24, 2004, pp. 139-146.
- [6] Majaranta, P., Mackenzie, I. S., Rähkä, K. Using motion to guide the focus of gaze during eye typing. Proceedings of 12th European Conference on Eye Movements (ECEM12), Dundee, Scotland, August, 2003.
- [7] Milekic, S. The More You Look the More You Get: Intention-Based Interface Using Gaze Tracking. Proceedings of the 7<sup>th</sup> Annual Museum and the Web Conference, Charlotte, North Carolina, USA, March 19-22, 2003.
- [8] Murata, A., Iwase, H. Evaluation of Mental Workload by Fluctuation Analysis of Pupil Area. Proc. of the 20<sup>th</sup> Int. Conf. of the IEEE Eng. in Medicine and Biology, 1998.
- [9] Partala, T., Jokiniemi, M., Surakka, V. Pupillary Responses to Emotionally Provocative Stimuli. Proc. of ETRA 2000, Palm Beach Gardens, FL, USA, 2000.
- [10] Porta, M. Implementing Eye-Based User-Aware E-Learning. Proceedings of CHI 2008 Ext. Abstr., Florence, Italy, April 5-10, 2008.
- [11] Porta, M., Ravelli, A. WeyeB, an Eye-Controlled Web Browser for Hands-Free Navigation. Proceedings of the 2<sup>nd</sup> IEEE International Conference on Human System Interaction (HSI 2009), Catania, Italy, May 21-23, 2009.
- [12] Porta, M., Turina, M. Eye-S: a Full-Screen Input Modality for Pure Eye-based Communication. Proceedings of the 5<sup>th</sup> Symposium on Eye Tracking Research & Applications (ETRA 2008), Savannah, GA, USA, March 26-28, 2008.
- [13] Rasmusson, D., Chappell, R., and Trego, M. Quick Glance: Eye Tracking Access to the Windows95 Operating Environment. Proc. of 14<sup>th</sup> International Conference on Technology and Persons with Disabilities, Los Angeles, USA, March 15-20, 1999.
- [14] World Health Organization. What is being done to improve the daily lives of people with disabilities? Retrieved July 11, 2009, from <http://www.who.int/features/qa/16/en/>, 2008.

### **ABOUT THE AUTHORS**

Virginio Cantoni, Marco Porta, Alice Ravarelli, Dipartimento di Informatica e Sistemistica, Università di Pavia, Italy. Phone: +39 0382 985 486, E-mail: {virginio.cantoni, marco.porta, alice.ravarelli}@unipv.it.