Emotional E-Learning through Eye Tracking

Marco Porta, Stefania Ricotti, Calet Jimenez Perez
Dipartimento di Informatica e Sistemistica - Università di Pavia
Via Ferrata 1 - 27100 - Pavia - Italy
marco порта@unipv.it, stefania ricotti01@universitadipavia.it, calet jimenez01@universitadipavia.it

Abstract— Besides proposing a short survey of eye tracking studies directly or indirectly connected with e-learning technology, in this paper we describe our ongoing research primarily directed to the development of an e-learning platform where eye data are exploited to obtain information about the student’s “emotional state”. The results of the preliminary tests presented here will be the starting point for more sophisticated experiments we are going to carry out.

Keywords- e-learning; human factors; eye tracking

I. INTRODUCTION

One of the main problems affecting e-learning is undoubtedly the absence of an “emotional” contact between educator and learner [1]. Apart from the psychological implications of the missing teacher in the flesh, a communication mediated exclusively by the computer lacks an important visual information stream: the one from learner to instructor. The assessment of the student’s level of comprehension can thus occur (in a more or less formal way) only a posteriori. On the contrary, in face-to-face learning the “good teacher” can catch signs of understanding difficulties, cognitive stress and tiredness by simply looking at his or her students.

Facial expressions and body language are typical examples of feedbacks that an instructor can receive. However, there is another source which can provide valuable information about learners, namely “eye behavior”. Whereas understanding eye signals may be difficult in in-person learning, because eye activities are usually characterized by micro indicators that may remain unnoticed in face-to-face communication, this ability becomes feasible using an eye tracker, i.e. a device able to follow the user’s gaze while looking at a screen [2]. Eye tracking can potentially reveal significant information about how the learning process is occurring, exploiting eye data measured in real time. Present eye trackers look almost like ordinary LCD monitors, and constrain user movements very little. Current applications of these devices are mainly in the fields of assistive technology, usability and advertising, but in the near future they are likely to spread as a new appealing technology, even incorporated into ordinary laptops [3].

The purpose of this paper is twofold. On the one hand, we propose a survey of eye tracking studies directly or indirectly related to e-learning, including those few systems which have explicitly exploited gaze input as a way for enhancing the learning experience. On the other hand, we describe our ongoing research aimed at developing an e-learning platform where eye data are used to get information about the student’s emotional state. To date, only very few e-learning systems have tried to derive emotional information from eye tracking data (e.g. AdeLE [4] and e5Learning [5]). However, such data have mostly been exploited at a very basic level and, above all, without pretending to take them as unmistakable signs of the learner’s cognitive processes. In our research we are instead seeking to build a robust e-learning platform where (at least some) eye behaviors can be really interpreted as clear indications of whether the student is having difficulty in understanding a piece of content or is getting stressed or tired. To do this, we are carrying out and planning preliminary tests in which students are asked to solve different kinds of problems, with different levels of complexity and required skills. We are in particular focusing on mathematics, which is one of the subjects that bachelor students of Electronics and Computer Engineering find more difficult. In our experiments, the student is presented with the problem on the eye tracker's screen, and his or her eye data, while solving the problem, are recorded in real time — specifically, gaze coordinates, fixation durations and pupil diameters, from which indirect but important measures, such as saccadic velocities and blink rates, can be subsequently obtained.

Several studies have been carried out to discover correlations between eye behaviors and emotional states (e.g. [6], [7] and [8]), mostly in the fields of psychology and physiology. Unlike these investigations, which have mainly been accomplished at a theoretical level only, our purpose is to actually apply “emotional hints” to a working e-learning system where real-time eye data are contextualized to the content being accessed by the student. In other words, we aim at developing a system able to reliably assess the learner’s cognitive status, at least for some categories of problems that we have properly studied in advance.

The paper is structured as follows. After presenting the basics of eye tracking technology, Section II proposes a short survey of projects and investigations connected with eye data and e-learning. The focus is both on e-learning systems explicitly based on eye tracking and on studies which have considered eye behavior as a general source of information. Section III describes the preliminary experiments we have carried out with a limited number of testers, in order to obtain indications about the feasibility of our approach. Due to lack of space, we will discuss only relevant results, highlighting the correlations between user “actions” and eye activities (especially for what concerns pupil size). Section IV, at last, draws some conclusions and provides hints for future work.

doi: 10.1109/EDUCON.2012.6201145
II. EYE TRACKING AND E-LEARNING

A. Eye Tracking Basics

Eye movements occur as sudden (almost instantaneous) saccades, followed by fixation periods of about 200-600 milliseconds. During fixations the eye is almost still. Fortunately, eye tracking has now evolved to the point where the user can move almost freely in front of the camera, within certain limits, and good accuracy (1 degree or better) is achieved throughout the whole working range. Practically all video-based remote eye trackers, which operate without contact with the user, exploit infrared or near-infrared lighting — not disturbing, since it is almost invisible. Eyes are tracked by measuring how light is reflected by the cornea and by the retina through the pupil [2].

B. E-Learning Systems Based on Eye Tracking

Among the very few e-learning systems designed to include eye tracking technology directly inside their architecture, AdeLE (Adaptive e-Learning with Eye tracking) is probably the first [4]. The main goal of the project is to dynamically capture user behavior from real-time eye-tracking data, trying to understand the learner’s cognitive processes and thus support adaptive teaching — for example, through real-time tracking of gaze duration (the time spent on an object), fixations, and blink rate, which are important factors to draw information about user interest. While the AdeLE project set itself ambitious goals, to our knowledge it has not been continued in these last years.

iDict is a translation aid designed for language courses employed to derive indications about whether and when the user needs help while reading a document [9]. Eye movements can reveal the cognitive load during text processing; for instance, long fixations and regressions may signal difficult lexical access. iDict works by tracking the reading process and identifying potentially difficult situations. When this occurs, a tooltip is displayed which shows the translation for the specific word or phrase being read, based on both gaze path data and lexical/syntactic analysis of the text.

In the system described in [10], eye movements are used both to get an indication of learner interest and focus of attention and to provide feedback to character agents. Animated agents may in fact better motivate learners, provided that they have some “intelligence” and knowledge about the user. Information about what the learner is watching can be exploited to infer both interest and current focus of attention. An “eye-aware” agent can for instance use information such as eye movements and pupil size to assess the state of the learner, and attract his or her attention when necessary.

e5Learning (enhanced exploitation of eyes for effective eLearning) is a prototype e-learning environment developed a few years ago by our research group [5]. The system allows the author of the course to decide how much time the user should spend looking at certain portions of content, be them textual or non-textual areas. Moreover, it tries to keep track of whether and how those areas have been accessed by the user. Additional content can be dynamically displayed depending on the main content being accessed at a certain moment by the user (e.g. when he or she looks at an image for more than a defined time threshold). e5Learning also includes a basic emotion recognizer module, used to detect two user states, namely “high workload/non understanding” and “tiredness”. Pupil information, fixation duration and blink rate are exploited to identify such conditions.

In [11] an e-learning environment is introduced which exploits eye tracking to assess the learner’s interest in specific subjects and to perform content adaptation. For example, a learner characterized by a strong visual memory but with weaker verbal processing capabilities will probably spend more time looking at pictures rather than at textual content. Thus, once this learning method has been recognized, the didactic content can potentially be adapted to provide mainly images and video rather than text. User behavior (including data such as the time the course has been started, visited pages and time spent on different areas) is recorded in a database.

C. General Eye Tracking Investigations on E-Learning

While the number of e-learning systems architecturally based on eye tracking is very low, several studies have also been carried out which have used eye tracking as a testing methodology.

For instance, in the experiment described in [12] an eye tracker has been employed along with an “emotion mouse”, that is a special mouse able to acquire data about heart rate, body temperature and galvanic skin response. The testers’ cognitive load and body parameters while performing different tasks (e.g. do some gym exercises, perform arithmetic calculations or play a videogame) have been estimated by means of the two devices. It has been found that body temperature decreases during calculations and increases at the end of the task; moreover, pupil size is directly correlated with the cognitive effort.

Another study [13] has considered three typical phases of e-learning activities, namely “searching”, “viewing” and “preparation”. It has for instance been found that in the searching phase pupil size is much larger than in the viewing phase (which suggests that the cognitive load is higher during a search task). During searching, the cognitive load is higher than during the other phases: saccades and blinks almost disappear and pupil size is very large. Also, saccadic rate decreases from preparation, to viewing, to searching.

The work described in [14] considers the connection between cognitive style and gaze patterns within an e-learning environment. The purpose is to adapt the learning experience as much as possible by recognizing whether the user is an “imager”, a “verbalizer” or an “intermediate”. The paper, in particular, focuses on the difference between imagers and verbalizers, confirming the existence of a diversity between the gaze patterns of the two categories.

An eye tracking study of Moodle (the open source course management system) aimed at finding how its components and teaching materials are watched by users is presented in [15]. Among other findings, the investigation discovered that the breadcrumb navigation and the “My courses” area were the most exploited navigation elements. Moreover, gaze analysis
showed that some users had significant difficulties when searching for their profile page and the logout button.

D. General Eye Tracking Studies Related to Emotion Detection

Not necessarily connected with e-learning — but potentially useful for the implementation of “user aware” e-learning platforms — some investigations have focused on eye behavior and emotions on a more general basis.

For example, the study described in [6] considers specific pupillary responses to emotionally provocative stimuli. Experiments have in particular shown that pupil size is significantly larger after highly arousing positive and negative stimuli than after neutral stimuli with medium arousal.

The work presented in [16] focuses on finding the best moments (in terms of mental workload) when to interrupt interactive tasks, in order to develop an “attention manager”. For mental workload assessment, pupil size measures have mainly been exploited. Experiments (text comprehension, mathematical calculations, search for specific items within a list and object manipulation) indicate that pupil size increases at the beginning of a task and goes back to its initial dimension at the end.

The tool briefly introduced in [17] uses an eye tracker to determine whether an image causes emotions, and of what kind. Eye parameters such as gaze direction, blink rate and pupil size are exploited to this purpose.

To develop a real-time system for cognitive load evaluation, the study described in [18] considers blink rate, pupil size and fixation data. After watching 15-second basketball videos, some players have been asked to identify attackers and defenders, as well as to remember their positions (with the task difficulty level given by the number of players to remember).

To investigate factors that may influence pupil size, in the experiment illustrated in [19] 12 testers have been asked to perform arithmetic calculations in different lighting and emotional conditions, provoked by images with different backgrounds. Studies of this kind can be useful to evaluate the user’s mental workload in real “noisy situations”, where such aspects are common and cannot be neglected.

III. Investigating Eye Behavior During Problem Solving: Our Experiments

Our final goal is to implement an e-learning platform which really takes into account the student’s emotional condition. To do this, we need to deeply understand those eye signals that more than others can indicate whether the user is going through high cognitive load or stress situations. We are thus planning and carrying out experiments aimed at both confirming outcomes of previous trials described in the literature and finding new potential signs that can be usefully exploited in our future system.

In this section we describe some preliminary (qualitative) tests we have conducted with five users. Although we will need to carry out many other experiments to confirm and refine the obtained results, they look very promising and thus worth to be presented here.

A. Participants

Five users (4 males and 1 female), aged between 27 and 50 (average of 32.8), participated in the study. All of them had normal or corrected to normal vision (two were wearing glasses).

B. Equipment and Setting

We recorded eye data using a Tobii 1750 remote eye tracker [20], characterized by an accuracy of 0.5 degrees and integrated in a 17” monitor (1280x1024). Recent studies [21][22] have demonstrated that the binocular pupillometric precision of this eye tracker is 0.15mm, and hence its resolution is enough to identify task-induced dilations. The device has a sampling rate of 50 Hz, with pupil size measured every 20 ms. Only fixations lasting at least 100 ms were considered. In addition to analytical data, also gaze plots (graphical depictions of fixations, with circles representing fixations and straight lines indicating saccades) and user audio/video recordings were exploited to examine the experiments’ outcomes.

Behind the eye tracker there was a wall painted in neutral gray and the illumination of the room was uniform and constant (in order for pupil size not to be influenced by external factors).

C. Task Description and Procedure

The task consisted in finding the right solution for three exercises concerning the Pythagorean theorem. Exercises had different difficulty levels: easy, hard or impossible. Impossible meant that the problem had no solution. Each user had to handle one exercise at a time, and all three exercises had the same structure: two slides of theory (the same for each exercise) about Pythagoras’ theorem and its consequences, a slide with the text of the exercise, a slide presenting four figures with possible graphical representations of the problem, and a slide proposing three possible ways of solving the problem. In these last two cases, users had to choose the right answer pronouncing out loud the corresponding letter (figures and solutions were identified by alphabet letters from ‘A’). Between slides, a white screen was shown for three seconds. The three exercises were presented (and possibly solved) one after the other. Users were explicitly asked to say out loud what they were thinking. Thanks to audio/video recording, this allowed us to precisely match eye behaviors to specific key phases in problem solving.

Summarizing the experiment procedure, the user’s task was composed of three subtasks (the three exercises), and each exercise was in turn divided into four sub-subtasks: reviewing the theory (already known by all users), reading the text of the exercise, looking for the correct figure which represented the problem (“figure sub-subtask” in the following), and looking for the right solution among the three proposals (“solution sub-subtask” in the following). As an example, Figure 1 shows the figure and solution sub-subtasks for the second (hard) exercise.

In order for user behavior to be as natural as possible, we
decided not to impose time limits on task duration, nor to provide any kind of suggestion on how to solve the exercises.

D. Results

Firstly, we produced graphs, for each user and sub-subtask, showing how pupil diameter changed over time. We observed correspondences between curve progress and user behavior.

In the figure sub-subtask, the moment when the user says his or her answer out loud always corresponds to a peak in the graph (Figure 2 shows some examples referred to three users). Although these are not the only crests present in the chart, most peaks match relevant user “behaviors”: for example, the moment when a solution is discarded because recognized as wrong, or when the user recognizes the right choice, just before saying it out loud. In addition, the graph for the problem without solution has a much more fluctuating trend, possibly due to the confusion it induces in the user. In such a situation, the e-learning platform could suggest a review of the theory.

Graphs for the solution sub-subtask (Figure 3) were much more complex than those for the figure sub-subtask; this may be due to both the higher difficulty of the task and the greater amount of information presented on the screen. Looking at the graph for the simpler exercise (Figure 3, bottom), in this case too we can identify peaks that correspond to the moment when the user recognized the right solution or an error. While this is also true for the impossible exercise (Figure 3, top and middle), in this case we can notice smaller pupil dilations towards the end of the task. This is known to be a sign of tiredness; when the e-learning platform finds this condition, possibly together with less eye activity and interaction, it could for instance suggest a break. The graphs for the impossible problem have many fluctuations, although we can still identify meaningful peaks. Towards the end of the task, the amount of pupil dilations tend to decrease, and the average pupil diameter is bigger than that for the other sub-subtasks.

As regards fixations, their average length was lower in the figure task than in the solution task. Although the graphs for the average fixation length do not exhibit marked similarities among users, some resemblances can nevertheless be noticed considering the charts of the three exercises for the same user.
As shown, for example, in Figure 4, graph trends are similar for the three subtasks. Further tests will hopefully confirm such tendency, possibly allowing our e-learning platform to identify specific patterns for each user. Figures 5 and 6 show example gazeplots for the figure and solution sub-subtasks. In Figure 5 it is evident that the user paid little attention to the manifestly wrong choices (B and C), to focus more on the other two depictions, which differed only in a small detail. In Figure 6, fixations are clearly more concentrated on solutions A and B, while the third (impossible) option receives little attention.

IV. CONCLUSIONS

It is now demonstrated (e.g. [16]) that pupil size acts as a function of processing load and mental effort. We intend to exploit this and other parameters — such as fixation duration and blink rate — to identify specific patterns that can provide information about the “emotional state” of the user. The e-learning platform could suggest theory revisions or clues about procedures to be followed when potential understanding problems are detected. For example, if the user of the e-learning platform is solving a problem during an assessment session and many peaks in pupil size are noticed — while the right answer has not been given yet — then it is very likely that the user is experiencing some difficulties in figuring out the problem. It is also interesting to note that ocular parameters could give information about the right moment to interrupt users. As pointed out in [23], the best time to disrupt primary tasks is when the mental workload is lowest. We are planning experiments similar to those performed in that study, but specifically applied to the mathematical field.

The preliminary tests described in this paper have mainly focused on pupil size related to user actions and thinking. However, pupil size can also depend on other factors than processing load: for example, aural stimuli, light variations, or emotions not related to the task being performed. It is thus necessary to consider other eye parameters too, such as fixation length, saccadic speed and blink rate, as well as general interaction indicators, such as mouse movement and clicks.

We are now planning more sophisticated tests, which will consider also the above quoted parameters and will be conducted with much more testers. Besides experiments conceptually similar to those described in this paper, we will also carry out more controlled trials, with precise time limits to perform the tasks and prescribed sequences of actions that the user will have to follow. This will allow testers to be better compared each other and to precisely relate user status to task completion level (with accurate quantitative data).

As said in the Introduction, unlike most investigations carried out in the past at a very general level, our final aim is to apply “emotional hints” to a working e-learning system, at least within specific contexts (such as mathematics). The
preliminary study presented in this paper is a crucial step towards such a goal.

REFERENCES


