EYE-TRACKING STUDENT'S BEHAVIOUR FOR E-LEARNING IMPROVEMENT

J. Calado, E. Marcelino-Jesus, F. Ferreira, J. Sarraipa

UNINOVA (PORTUGAL)

Abstract

Nowadays, there is a shift in society towards an increased use of technological aids, which help us in our daily life tasks. However, very little of these improvements are being used in the education system, specifically to help mitigate common academic problems that affect students, such as poor performance and disengagement issues, that can arise from identifiable symptoms like changes in behaviour and emotions. In this paper an architecture to handle eye tracker data is presented, which analyses the gaze data, captured from students in learning scenarios to predict specific behaviours and issues. A scenario where distractive stimuli was used to emulate deviations in the attention levels is presented. A methodology was defined able to detect gaze patterns for further automatization of the detection of students’ issues as lack of attention and performance. The goal is in the near future to provide Learning Management Systems with solutions able to present insightful real time cognitive data of the users and automatically detect if they are exhibiting any low performance and automatically react.

Keywords: Education; Eye-Traker; E-learning; Automatization.

1 INTRODUCTION

Nowadays, the students drop-out are one of the biggest problems existents in Higher Education Institutions (HEI) and the factors that can lead to it are diverse, namely the cultural, social and economic factors. However, physical and psychological disabilities can lead to academic marginalization, as well as, emotional and affective problems. This problematic is worsened by the shortage of educational resources that can bridge the communication gap between the faculty staff and the affective needs of these students [1].

Thus, the adaptation of technology in education has brought changes in the level of teaching, making teachers and students benefit from the existing educational technologies. However, it was necessary for teachers to learn how to use the new technologies in their classrooms, motivating and facilitating students' learning with their use. The use of these technologies in education has allowed the removal of educational boundaries, approaching this way, teachers and students and making them interact in real time [2]. The effective use of technology has the potential to transform the student-teacher relationship at the undergraduate level [3], which generates new opportunities for students. Some advantages of using technology are: technology unlocks educational boundaries; technology simplifies access to educational resources; technology motivates students; technology improves students writing and learning skills; technology makes subjects easy to learn; promotes individual learning; supports differentiated instructions; increases collaboration between teachers and students; prepares students for tomorrow technological jobs and increases students innovation and creativity [4].

In order to contribute to a solution and motivate these student's to proceed with their studies, universities are increasingly using technical aids. The technical aids or assistive technology devices for communication allow students to reach a significant level of independence [5]. These kind of technological equipment can compensate efficiently certain disabilities and, at the same time, be considered assistive technology devices [5].

The ACACIA project is a consortium of various higher education institutions from Europe and Latin America and Caribbean (LAC), united under the common goal of promoting academic integration, through the creation of 3 Centres for Professional Development and Educational Support (CADEP), in Latin America, namely Peru, Colombia and Nicaragua [1]. CADEP centres intended to provide better cooperation and innovative technological tools, to intervene in social areas like school dropout prevention, which is mandatory to establish institutional organizations [6].
This project intends not only to define a model for the LAC in order to guarantee best practices strengthening and exchange between America and Europe, in relation to the training of teachers, but also to promote the use of technologies from a diversity approach in didactic training of teachers [7].

CADEP Centres are divided to achieve specific goals, that covers wide areas that, allows to intervene more accurately and articulate the connection between CADEP centres and Universities, therefore these centres will make use of an integrated system of 5 modules that are (in Spanish and translated): 1) Empodera - Empower; 2) Innova - Innovate; 3) Cultiva - Educate; 4) Apoya – Support; 5) Convoca - Convenes described in detail in [6].

Thus, it was made some experiences in HEI with some student's namely in Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa" with "Eye Tracker" a "Kinect" and a Camera in order to study the emotions, behaviours and affective states of the student’s using this technological tools. The objective is to analyse student's behaviour's in a classroom through these three biometric acquisition devices and studying the movement of the eyes, posture of the students entering or leaving in the classroom and emotion detection through facial expressions. Consequently, the main goal is to detect and prevent future drop-outs in the faculty and find solutions to motivate students to continue their studies and improve their grades and integration. However, the authors intend to give more emphasis to the "Eye Tracker" case study.

In this paper, it is introduced in section 2 a brief state-of-the-art about biometric devices used in the scientific experience. In section 3 the application scenario is presented, detailing the biometric data acquisition and the GP3 eye tracker, gazepoint control and eye tracker data acquisition processor. In section 4 are revealed de results of this experience. Lastly, in section 5 the conclusions are presented as well as, the future work.

2 BIOMETRIC DEVICES USED IN THE SCIENTIFIC EXPERIENCE

2.1 Eye Tracking

An eye tracker is a device that measures the activity of a person’s eyes, namely the position and movement, and is usually comprised of a head-mounted or a remote (desk mounted) apparatus. The Eye trackers are commonly composed by a light source (usually infrared or near-infrared) and one or more cameras pointed at the person eyes. The use of an infrared light source is preferred due to the fact that it helps reduce the negative effects of ambient lighting conditions. The light reflection on the cornea is then used to extrapolate the direction of the gaze, as it is the brightest spot in the pupil [8]. However, other biometrics can also be measured by the eye tracker, such as:

- gaze fixation: the location where the gaze is maintained, for necessary amount of time for the eye to focus, when the visual input occurs;
- gaze duration: the amount of time for which a gaze is maintained;
- scan path: the vectors set that represent the path of gaze;
- areas of interest: refers to designated gaze locations and the collected gaze statistics for each area;
- pupil size: variations occurring in the shape and diameter of the pupil;

And various types of eye movements as well:

- saccades: the quick eye movement between two fixations;
- smooth pursuit: the eye movement while following a moving object;
- vergence: the simultaneous movement of both eyes in opposite directions to obtain or maintain single binocular vision;
- vestibulo-ocular movements: reflex eye movement, in the opposing direction of the head movement, that preserves the image in the centre of the visual field.

2.2 Kinect

Kinect is an add-on device for the Microsoft Xbox 360 gaming system that enables users to control games, movies and music with physical motion or voice commands and without the need for a
separate input controller like a joystick or keyboard [9]. This device is mainly a game controller and collection of sensors which consists of 2 RGB cameras which provide full-body 3D motion capture, facial recognition and voice recognition capabilities, stepper motor and mic array to take sound input as well [10]. Kinect can analyze the person in front of it and go through multiple “filters” to try and determine which type of body structure matches with the correct type programmed in its system [11].

The Kinect is quickly becoming a vital tool because of the 3D data it captures in visible and infrared wavelengths with very high accuracy. The Kinect is in a league of its own effectively capturing 9 million data points per second. The Kinect is an inspiring device because of its low cost and most students are already familiar with it [11].

The data is constantly being transferred between Kinect and objects in their field of view (from one side to the other) without the user having to hold anything in their hands. It is also becoming a very important tool due to its high precision, i.e., the 3D data that captures in visible and infrared wavelengths. This device can capture 9 million data points per second being nowadays, very used in day-to-day life because it is small, cheap, advanced and most students are familiar with it.

2.3 AFFECTIVA - SDK (Camera)

Affectiva Emotion Artificial Intelligence (AI) also known as Affectiva is an emotion measurement technology company, that develops software to recognize human emotions based on facial cues or physiological responses [12].

Humans use a lot of non-verbal cues, such as facial expressions, gesture, body language and tone of voice, to communicate their emotions [13]. Thus, this software brings emotional intelligence to the digital world with our emotion recognition technology that senses and analyzes facial expressions and emotions measuring unfiltered and unbiased facial expressions of emotion, using any optical sensor or just a standard webcam [14][13].

Affectiva intends to humanize technology using a Software Development Kit (SDK). SDK detects and analyse spontaneous and individual facial expressions that people show in their daily interactions for groups of 20 or more persons in real time. This software analyze recorded media in the cloud, taking just a few hours to setup, giving the ability to fully emotional enable applications, working with any optical sensor, device camera or standard webcam [15].

The computer vision algorithms identify landmark points on the face (i.e., the eyebrows corner, the tip of nose, the mouth corners, etc., then the machine learning algorithms analyze the pixels in those landmarks in order to classify facial expressions. Thus, SDK uses Facial Action Coding System (FACS) to classify facial expressions mapping them to emotions. Nowadays, Affectiva is used several different industries such as: online education, healthcare, gaming, robotics, media, and advertising, market research, automotive retail, human resources, training and coaching, video communication, experiential design, and in wearables and devices [14]. This increasing use of this technology analyzes the human expressions of emotion simply and cost-effectively getting powerful emotion data highly accurate.

3 APPLICATION SCENARIO

The proposed scenario consists in evaluating the students’ emotions, behaviours and affective states, aggregating the information, using a framework capable of assessing problematic situations. The cases studies, represented in Figure 1. comprise three biometric acquisition devices and one integration platform, for information analysis and reporting system, as a proposed solution.
3.1 Biometric Data Acquisition Case Studies

For the presented scenario three different capturing devices can be used: a Kinect, an Eye Tracker and a Camera, each of one are described in the corresponding case study.

Case #1: Gait and Posture Analysis while entering or leaving the classroom

This case study implementation makes use of a 3D motion capture device, composed by an RGB camera and an infrared depth sensor (i.e. the Kinect device), to track the student skeleton key points, in order to analyse its posture and gait, while entering and leaving the classroom.

It is expected with this observation, that a change in one student regular gait can be flagged as a signal for changes in determined emotional states.

Case #2: Eye Tracking while sitting at a desk

This case study will use a remote Eye Tracking device, which is generally composed by an infrared light source, an infrared or near-infrared camera and an RGB camera. The Eye Tracker sits in the desk, directed at the student, and tracks head movements as well as eye movements (saccades, smooth pursuit, vergence and vestibulo-ocular movements, gaze fixations and eye blinks), and pupil’s size and dilation changes (proven to be an indicator of cognitive activity).

The analysis of these biometrics can enable the recognition of the student different affective states, during the learning process, and can also be used to measure the engagement level of the student.

Case #3: Automatic Emotion Detection through Facial Expressions, while sitting at a desk

This case study will use a regular RGB camera, pointed at the student's face, in order to record the changes of the facial expressions.

An algorithm will detect and isolate the face, then extract the facial features collected, analyses the classified facial expressions, and determines the correspondent emotion or emotions.

Both eye tracking and facial expression emotion detection case studies can be deployed to a student using a computer, alongside with the sensorial device, or without the computer and just the device.

3.2 GP3 Eye Tracker, Gazepoint Control and Eye Tracker Data Acquisition Processor

The GP3 Eye Tracker, Gazepoint Control and Eye Tracker Data Acquisition Processor modules represent the data influx, collected from the student gaze, to the creation of files in the Database.

Thus, in the GP3 Eye Tracker, the Eye Tracker used was the Gazepoint GP3 Eye Tracker [17]. This eye tracker has a sample rate of 60Hz and provides the following metrics:

- Fixation POG;
- Left and right eyes POG;
• Best POG, which is the average POG data from the left and right eyes if both are available, or just either one, depending on which one is valid;
• Left and right eyes pupil coordinates and diameters, in pixels and in meters;
• Blink duration and count per minute;

It connects to a computer through Universal Serial Bus (USB) and requires two powered ports to work. Due to the basic metrics it provides, and the low accuracy results, the study and development of cognitive and affective metrics was not feasible.

The Gazepoint Control is a proprietary middleware that is required to collect the captured data from the GP3 Eye Tracker. It allows TCP client to connect, and has a set of commands it can receive, to configure the eye tracker data and control its usage. The Gazepoint Control GUI, shown in Figure 2, offers the video overview being captured by the eye tracker, an option to select the screen in use, and a calibration option, where a 5 or 9 point calibration is possible. Nevertheless, the eye gaze video, captured by the eye tracker, is not recorded.

3.2.1 Eye Tracker Data Acquisition Processor

The Eye Tracker Data Acquisition Processor is the module used to connect with the Gazepoint Control, and record the eye tracker data in the Database as XML files. This module offers the option to configure which categories of data records are received from the Gazepoint Control TCP server, to connect/disconnect from the server and to start/stop the data stream.

The TCP stream is received as a string in XML format, and contains the metrics provided by the eye tracker as well as a record counter, the time (in computer ticks and in seconds), the cursor position and optional user data. As each stream string is received, it is written in a XML file, named with the device name (“GP3”) and with the starting date and time of the Observation, which is stored in a “Records” directory representing the Database. The strings keep being written in the file until it contains all the records made during the time specified by the file creation sample rate variable. This
sample rate variable is predefined to 60 seconds but can be changed as a launch argument when the processor starts.

3.3 Profiles and Behaviours classification

Three types of Observations direct subclasses can be created: Human_Observation, Digital_Observation and Profiles.

Human_Observations represents the observations collected via the teacher or an expert input, and therefore require the link with exactly one Teacher individual.

Digital_Observations on the other hand, represents the observations collected via hardware or software devices that target the student.

The Profiles subclass and its subclasses (Attention_Disorder, Drop_Out, Sociological_Issue, Other_Profile) are intended to represent the various observation requirements necessary to trigger an alarm, flagging the situation corresponding each profile.

Off_Task: the student is not working on the activity assigned by the teacher.

The four profile classes presented, named Attention_Disorder, Sociological_Issue, Drop_Out and Other_Profile, are intended to be used as the knowledge variables, representing the conditions necessary to trigger the respective profile alert.

The Other_Profile is a class open to expand the existing profiles to future work developed using this framework.

1. Attention_Disorder

Attention deficit hyperactivity disorder is usually regarded as having three main characteristics: hyperactivity, impulsivity and inattention. The fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) defines it as “a persistent pattern of inattention and/or hyperactivity-impulsivity that interferes with functioning or development” [59]. It also states that in adults (17 years and older) five or more symptoms (of the total 18 listed) are required to be observed for at least 6 months.

2. Sociological_Issue

Oppositional defiant disorder is defined by the DSM-V as “a pattern of angry/irritable mood, argumentative/defiant behaviour, or vindictiveness” [59]. Again, at least four symptoms from any of these categories (angry and irritable mood; argumentative and defiant behaviour; or vindictiveness) must be observed for at least 6 months.

3. Drop_Out

Because there are many factors that can lead to school drop outs and there seems to be no consensus in a clear diagnose, the observations used for this profile were proposed using common sense. The same process was used for the sociological issues (which can cover a broad range of particular situations).

4. Obtained Results

The experiment consisted in a simulated test of an e-learning environment, where the participants were asked to read two excerpts of a text document, while their emotions, behaviours and affects were being recorded, by the sensors and by an expert. Each of the two excerpt readings was timed, and in one of them (the first one to half the participants, and the second one for the other half), external stimuli were introduced in the experiment, to induce distraction and disrupt the person's task, in order to record a situation resembling the conditions set for the Attention_Disorder Profile. At the end of each excerpt reading, the participant was asked to answer five questions about its content, to assess the results of their attention level, during the reading. The analysis of the measures detected by the eye tracker (Off_Task), and by the emotion detection algorithm (Anger, Contempt, Disgust, Fear, Joy, Sadness, Surprise and Engagement), are detailed in Table 1. The analysis of the observations made by the expert are detailed in Table 2, and although in some emotions the average value is always zero, it only means that the expert did not observe any of those emotions during all the experiments.
Table 1. Average Values Measured by Digital Observations

<table>
<thead>
<tr>
<th>Participant</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excerpt</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Stimuli</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Anger</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0098</td>
<td>0.0195</td>
<td>0.0001</td>
<td>0.0000</td>
</tr>
<tr>
<td>Contempt</td>
<td>0.0147</td>
<td>0.0224</td>
<td>0.0612</td>
<td>0.0226</td>
<td>0.0653</td>
<td>0.0871</td>
</tr>
<tr>
<td>Disgust</td>
<td>0.0062</td>
<td>0.0066</td>
<td>0.0044</td>
<td>0.0040</td>
<td>0.0066</td>
<td>0.0066</td>
</tr>
<tr>
<td>Fear</td>
<td>0.0012</td>
<td>0.0005</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Joy</td>
<td>0.0208</td>
<td>0.1522</td>
<td>0.0316</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1220</td>
</tr>
<tr>
<td>Sadness</td>
<td>0.0006</td>
<td>0.0007</td>
<td>0.0088</td>
<td>0.0138</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Surprise</td>
<td>0.0476</td>
<td>0.1142</td>
<td>0.0169</td>
<td>0.0018</td>
<td>0.0024</td>
<td>0.0329</td>
</tr>
</tbody>
</table>

Table 2. Average Values Observed by the Expert

<table>
<thead>
<tr>
<th>Participant</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Excerpt</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>Stimuli</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Anger</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Contempt</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Disgust</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Fear</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Joy</td>
<td>0.0000</td>
<td>0.0545</td>
<td>0.1450</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0809</td>
</tr>
<tr>
<td>Sadness</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Surprise</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Behaviour</td>
<td>Off_Task</td>
<td>0.1727</td>
<td>0.1818</td>
<td>0.0000</td>
<td>0.3909</td>
<td>0.2727</td>
</tr>
<tr>
<td></td>
<td>Engaged</td>
<td>0.7818</td>
<td>0.7818</td>
<td>1.0000</td>
<td>0.5909</td>
<td>0.7636</td>
</tr>
</tbody>
</table>

The analysis of the behaviour properties shows a good indicator for the detection of attention problems, being Off_Task, reported by the eye tracker, a suitable property to be used as a threshold for setting this Profile condition. The difference in the Off_Task average value, comparing the stimulated excerpt to the non-stimulated one, ranged between 8.8% (participant 3) and 25.9% (participant 4).

This conclusion falls in line with the assumptions made for the Attention_Disorder Profile detection, and with the Human_Observations made by the expert. However, the average Engaged values were not consistent with the Observations made by the expert, because no information is provided about this measure accuracy, which proved to be inaccurate during these tests, the Engaged property will be disregarded from the analysis.

The detection of changes in student’s behaviours is greatly facilitated with the aid of the eye tracking apparatus, which is a step forward in the prevention of drop-out intentions and academic disengagement.

5 CONCLUSIONS

The research work, presented in this paper, was developed as a response to a growing problem in HEI, concerning students’ engagement and learning problems. As society shifts to make use of more technological aids, the need to adapt the academic context, to include these benefits, is demanded. The proposed framework is one of some possible answers to mitigate these needs.

Designed to be interoperable with different types of sensorial systems, the framework proved capable of detecting the student’s behaviour, emotional and affective states, and to manage all the collected information into a knowledge-based system, able to pre-emptively and pro-actively detect situations consistent with profiles of students’ problems.
Following the framework guidelines, the implementation and deployment of a system, using an eye tracker for behaviour identification, the recognition of emotions through facial expression analysis software and recorded observations from experts in educational psychology, was successfully achieved and validated. This system capability for real-time profile warnings, is a valuable asset to assist teachers identifying problems, during the students learning process, and to help in the prevention of school drop-outs.

As a future work is expected to integrate an architecture with these three biometric acquisition devices, as well as to expand the integration with other sensors. It is also expected that the correlation, between the Digital_Observation and Human_Observation values, can be used in a machine learning implementation, that attempts to predict the detection of the profiles, further enabling the prevention of drop-out intentions and academic disengagement. It is also expected the deployment of this scientific study in other HEI, where significant data could be collected, further improving the machine learning dataset and, subsequently, the algorithm itself, for the more accurately detection of academic problematic profiles, thus fulfilling the objectives set for the Apoya module, of the ACACIA project [18].

ACKNOWLEDGEMENTS

The authors acknowledge the project ERASMUS Plus: Higher Education – International Capacity Building - ACACIA – Project reference number – 561754-EPP-1-2015-1-CO-EPKA2-CBHE-JP, (http://acacia.digital). The work has also been promoted under the project CARELINK, AAL-CALL-2016-049 funded by AAL JP, and co-funded by the European Commission and National Funding Authorities of Ireland, Belgium, Portugal and Switzerland.

REFERENCES


