

USING COMPUTER VISION FOR STUDENT-CENTRED REMOTE LAB IN ELECTRONICS

Fawzi Khattar^{1,3}, Franck Luthon^{1,2}, Benoît Larroque², Fadi Dornaika^{3,4}

¹ *University of Pau and Adour Province, Computer Science Lab LIUPPA, Anglet (FRANCE)*

² *Dept of Industrial Engineering, IUT Bayonne-Pays Basque, Anglet (FRANCE)*

³ *University of the Basque Country, UPV/EHU, San Sebastian (SPAIN)*

⁴ *IKERBASQUE, Basque Foundation for Science, Bilbao (SPAIN)*

Abstract

Technology is affecting students' life in different ways. Nowadays, students are more interested in social media, games and Internet of Things. On the other hand, most current teaching techniques appear outdated and uninteresting to students due to the fact that they do not follow the technological advances. For these reasons, students nowadays do not show much interest in education and learning. On the other hand, there is an urgent need for new learning approaches that take into account both the interests of today's students and the technological advances in several fields like telecommunication and gaming. Particularly for STEM curricula (Science, Technology, Engineering, and Math), labs play a crucial role to help students fully understand the material given during lectures. However, traditional hands-on labs are tied to space-time constraints: a student must perform the lab activity in a pre-scheduled time and in a given space. Furthermore, traditional labs are not suited for all students: some students may need more help and more motivating factors in order to take the most of the lab experience and to enhance their learning outcomes. This is not always best done in hands-on labs due to large number of students and to the presence of only one teacher that has to do lots of effort to adapt to each student's needs. Remote labs provide an alternative experience to the traditional lab, eliminating time-space constraints and introducing students to a new way of working by manipulating hardware remotely, which is used frequently in industry (SCADA: Supervisory Control and Data Acquisition). It can also be adapted to each student's needs if one is capable to infer student's mood, emotion and competence. In this paper we present the implementation of a remote lab in electronics, called LaboREM, which takes advantage of the technological advances in telecommunication and takes into consideration what current students are interested in. It is based on three parts: a remote laboratory, a learning management system and a game-like approach. The remote laboratory consists of remotely-controlled measurement devices, plus a simple robotic arm that mimics the student hand to construct electronic circuits. The robotic arm picks up electronics components equipped with magnets and places them on a breadboard. A wide angle camera is installed in order that students see what is physically happening in the lab. The electronic lab activity consists of short-time experiments (mainly tests on active filters) that last less than 5 minutes. The scenario is based on a game-like approach: a treasure hunt and a Top 10. Furthermore, a mini drone that can be controlled by the student remotely complements the static camera to add more immersion to the lab experience. The student can send specific requests to the drone in order to see some electronic instruments in a more interesting way. Moreover image processing and pattern recognition techniques are also used to infer student mood, concentration and motivation. Clues of boredom, lack of motivation and concentration (closed eyes, student not looking at the screen, student keeping moving...) are detected from the video of students' face taken from a webcam in order to adapt the scenario and the lab to each student.

Keywords: Remote lab, electronics, student motivation, industrial engineering, e-learning, game-based learning.

1 INTRODUCTION

Labs are an essential part in education. Particularly for STEM (Science, Technology, Engineering, and Math) students theories learned in class remain nice formulas and complicated concepts to most students if it's not implemented in practice as labs. Labs shed the light on the importance of the theory learned in class and introduce students to its application in real life. Thus labs are an important part in the education curriculum of STEM students. However, traditional hands-on labs suffer from time-space constraints. The labs must be scheduled in a specific time and space in a way that might not suit all different types of students. Some students might be experiencing affective states (sad, anger...) or

physiological problems at that time of the day that may inhibit the perception of information during the lab session. In addition, large number of students present in a lab session compared to a single teacher might in some cases make the learning process not the same for all of them. Some students might need more time to understand or implement a specific idea than other smart students need. Thus some students might benefit more from the lab than other students.

Remote labs are beginning to take a big part in education curriculum as technology develops due to their benefits and interesting aspects. Unlike a traditional hands-on lab, remote lab alleviates space-time constraints and gives more freedom to students to manage their time in the way that suits them. More specifically, a remote lab session can be done at anytime from anywhere given only an Internet connection and an access to the remote lab server. Furthermore remote labs introduce students to a new way of learning. Particularly for STEM students, a remote laboratory session gives the students an intuition of the importance of technology, precisely SCADA technology (supervisory control and data acquisition system) and gives them a sense of what interesting ideas they can accomplish if they put efforts into learning. It will introduce students to the possible interesting applications of the theory they learn in class and hence increase their motivation. As a result learning won't be seen as a way of getting best grades and passing exams anymore, it will be seen as an interesting topic for them. On the other hand, remote labs open the door to possible collaborations between universities. Universities can take advantage of other universities equipment and use them remotely to the benefit of its students.

Despite all these advantages that the remote lab offers, it has a major drawback. During a remote lab the student will interact with a machine, the teacher available in a traditional lab is not present during a remote lab session. This can rapidly lead to different kinds of problems. For instance, a student who has difficulties in getting the answer to a question might find himself stuck. He might end up being frustrated, confused and not motivated which will lead him to lose confidence and energy and give up on continuing the session. The problem of being confused for a short amount of time is usually a positive sign that the student is actually learning and previous studies have shown the correlation between this event and learning outcomes. However if this state persists, a need for intervention is necessary in order to help the student and prevent him from going to a frustrated state. Another problem is when the student is not concentrated on the lab. During the lab session, external events happening in the environment of the student can make him distracted for a while. He might talk with people surrounding him, or may be attracted by other events taking place. Indeed, this problem is not so crucial if it happens for a limited duration of time. Human being by nature needs some rest from time to time in order to continue concentrating and working efficiently on given task. However, if this event takes much longer time, it will affect badly the objective of learning. This problem is less present in traditional lab due to the presence of a human that can interact and unconsciously detect the affective state of the student and whether he needs help or not. The human to human communication is very strong, a good teacher can directly sense what the student is experiencing by just a few seconds of watching him in the class. This capability is normally not present in machines which do not have the intelligence required to infer the cognitive and emotional state of the student. Thus, there is an essential need to add emotional and cognitive intelligence to the machine for the remote lab to be effective. Given these challenges, we present the implementation of a remote lab in electronics. We intend to add the affective and cognitive intelligence into our remote lab by using computer vision, image processing and pattern recognition techniques. A video stream taken by the means of a webcam will serve as an input to several algorithms in order to monitor the state of the student and taking intervention in order to make the learning process more efficient. Several behavioral and emotional states can be detected and tracked over time in order to take appropriate interventions (reducing the difficulty of the lab...). In addition to that, a mini-drone equipped with a camera that can be controlled by the student is under investigation. The scenario is that by giving the student control of such an interesting modern tool will increase his motivation towards learning, add more immersion to the lab experience and make him more concentrated on the lab session. In the following section we present the implementation of the remote lab called LaboREM (for remote laboratory) and its different hardware, software and motivating aspects. After that we present a little survey on the affective state detection in learning environment. Then we present our current work towards detecting the affective and cognitive aspects of the students, and we end by a discussion, future work and perspectives.

2 LABOREM

LaboREM is a remote laboratory in electronics developed mainly for first year undergraduate students in engineering [1]. The main objective and motivation behind it is to increase the motivation of the

students. Its design is a classic client server application. The student calls for a lab session by simple URL addressing. A simple first in first out strategy is adopted to give access to the remote lab. The remote lab application is developed using NI-LabVIEW software and the simple and easy to use RFP protocol is used to pilot the remote devices. Current hardware implementation involves: (i) a robotic arm that is used to mimic the student's hand and places electronics components equipped with magnets on an electronic breadboard, (ii) a webcam with zoom control that mimics the student's eye in order that the student doesn't feel so far from what is actually happening in the lab, (iii) measurement instruments, (iv) data acquisition system (DAQ) and other components shown in Fig. 1. The platform is based on an LMS (learning management system). The global scenario of the lab is based on a game like approach in order to increase the motivation of students. This approach includes a treasure hunt and a top10 of the best measurements. The game scenario is based on four basic concepts: time spent, score obtained or mark for an activity, levels (beginner, medium, advanced) and number of lives or repetitions allowed. The learning objective of LaboREM is to enable student to wire, test remotely electronic circuits, make measurement and characterize each circuit by its time or frequency response. The electronic circuits are simple operational amplifiers (AO), active filters and oscillators.

2.1 Hardware

The remote lab consists of several connectable hardware devices that are connected together through several interfaces (USB, PCI, RS232, and GPIB). The hardware consists of an old fashioned robotic arm, a webcam, instruments (oscilloscope, multimeter, waveform generator), in addition to other equipment (data acquisition system, switching matrix...) Fig.1 shows the different hardware components of LaboREM. The robotic arm picks up electronic components from a bank of components and places them on a breadboard in order to take measurements. The electronic components are placed on a support equipped with magnets to make it easy to construct the circuit without the need for a high spatial precision or strong mechanical effort from the robot.

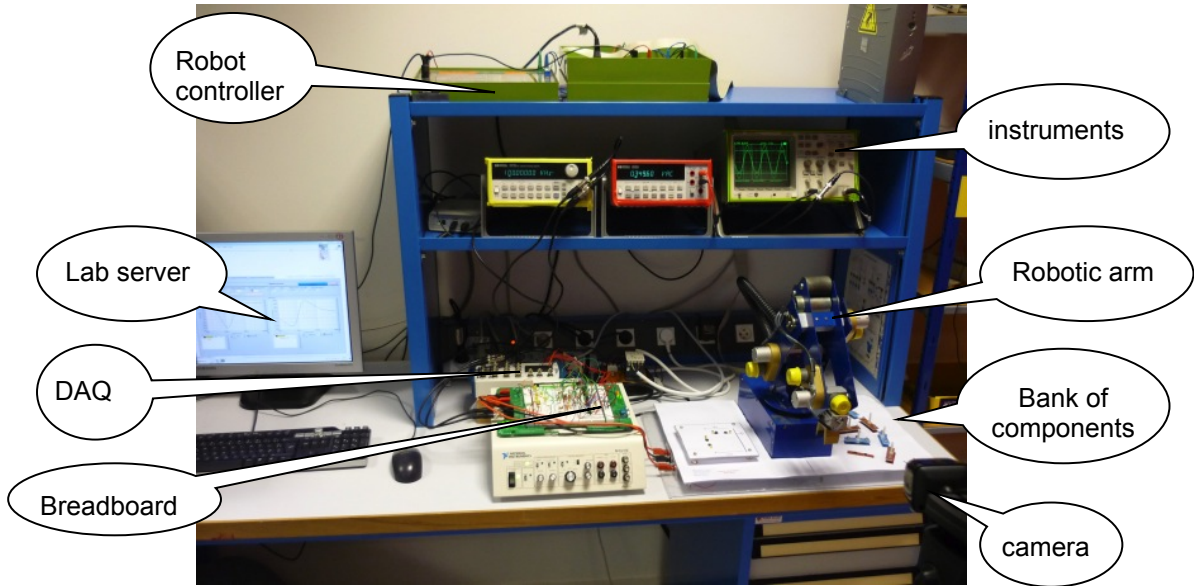


Figure 1 - LaboREM hardware components.

2.2 Network architecture

The network architecture implemented is a service-oriented architecture. It is a basic client-server architecture. Four main entities interact in order to send commands to the remote lab and retrieve measurements as shown in Fig. 2: (i) the client, (ii) the LMS, (iii) the corporate webserver, (iv) the lab server. The client sends requests via Internet. The corporate webserver receives these demands and gives access to the remote lab platform. The LMS organizes the game-based learning scenario. Finally the lab server operates the robotic arm and all other hardware instruments.

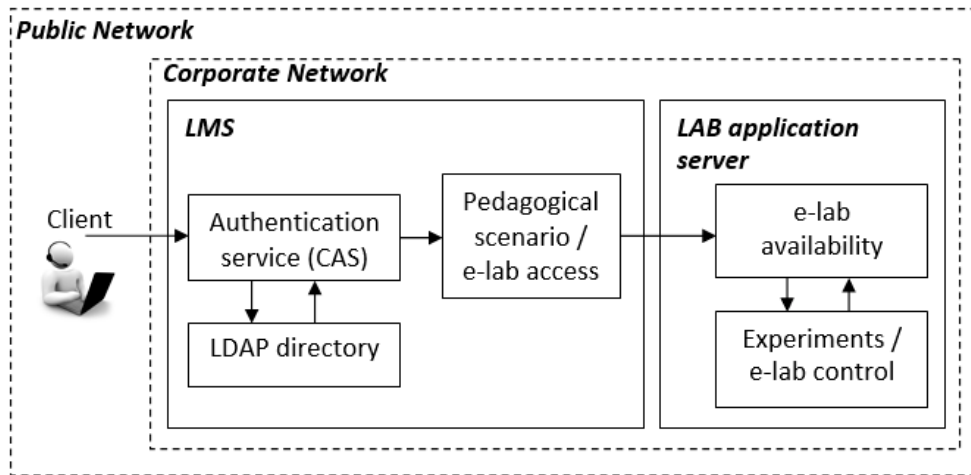


Figure 2 - Network architecture of LaboREM.

2.3 Software architecture

The lab application that is used to control the hardware components is composed of 4 main modules. Each module has specific functions to perform. The Laboratory Works module deals with the control of the robotic arm to build the requested filter. It also makes measurement, recovers data graphs and implements the top10 ranking. It also allows to select the type of experiments (currently Bode plots or signal acquisition and spectrum analyses). The Visualization module deals essentially with the video stream feedback taken from the webcam. It is also responsible for the display of the waiting queue of users and the top10 of the students. The Simulator gives numerical simulation of the experiments. Since it does not depend on the hardware, it can be used in case the robot arm encounters a failure or as a tool for comparison between the actual measurements and the simulated ones. The Virtual Manager manages the lab. Its main functionalities are storage, scheduling, initialization and security. The storage functionality stores the data that correspond to each user and to each performed experiment. Scheduling manages the waiting queue of connected clients, it is based on a simple first in first out solution. Initialization and Security reinitializes the lab to its initial state (robot arm, interface) when a new user is controlling the lab or when the experiment ends.

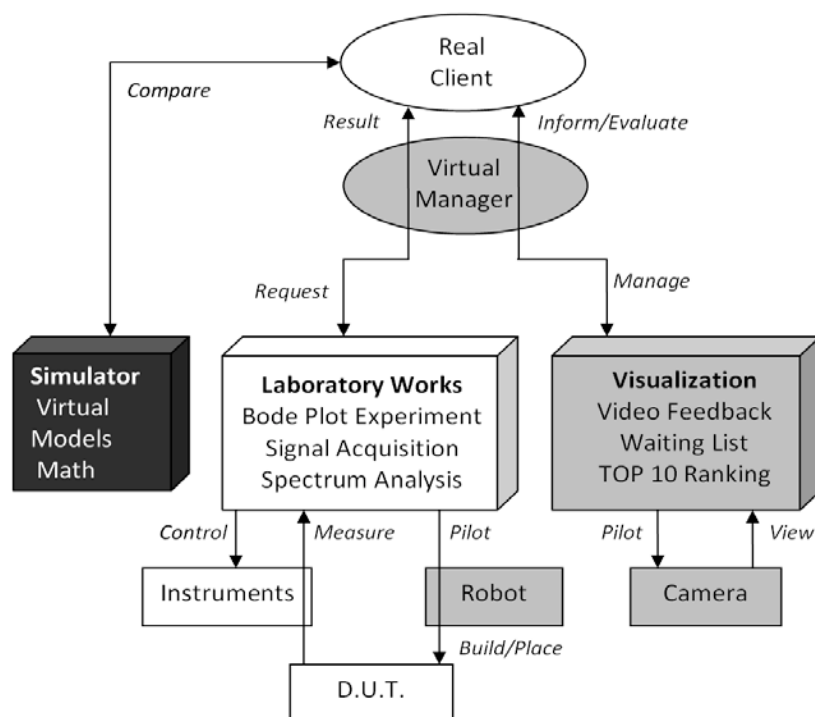


Figure 3 - Software architecture of LaboREM.

3 AFFECT AND BEHAVIOR DETECTION IN EDUCATION ENVIRONMENT

3.1 Student affect recognition

Student affect detection is a main issue that must be addressed in auto-tutoring systems. Affective state is a hidden complex process that can be observed through different modalities. Previous efforts have been put into trying to infer the human affect in order to determine the best interventions. Previous work concentrated into trying to infer the affective or cognitive state through observations of three modalities: visual channel, physiological channel and vocal channel. In the visual channel, facial expressions, head pose, gaze, posture pattern and body gestures are monitored. In the physiological channel a diversity of physiological signals (blood pressure, heart rate, muscle, nerves...) are analyzed. In the vocal channel, measures of intonation or prosody are taken. In addition some systems use context information (log data, mouse clicks, keyboard clicks, time spent on a certain task...) although it is a dependent variable on the particular auto tutor system. Given this diversity in sensing modalities, the dependence of affect interpretation on several factors (time, context, person) and the difficulty of the problem it is very difficult to make general claims on the accuracy of an affect recognition system and this is a well-known and acknowledged limitation of affect recognition technologies [2].

3.2 Visual modality

The visual modality constitutes a non-intrusive and rich modality for affect recognition. The main visual cues are head movements, body gestures and facial expressions. In automatic face expressions analysis by machines, there are mainly two streams that follow two streams in psychological research: message judgment and sign judgment. Message judgment approaches try to categorize the affect into pre-chosen categories, mainly the six basic emotion categories of Ekman directly (anger, disgust, fear, happiness, sadness and surprise). Sign judgment approaches describe first the visible behavior prior to take the high level decision of categorization. FACS (Facial action coding system) constitutes the most commonly used system to describe any facial behavior objectively. It decomposes facial expression into a combination of small facial muscle activation and head movements called action units (AU). Fig.4 shows some of the action units defined in Ekman theory. Many researchers try to find a relationship between shown behavior and the hidden affect state. In [4] authors try to identify student behaviors that correlate with student learning. They try to detect emotional cues in addition to on-task off-task cues as they explain that emotional cues might not correlate well to learning if taken independently from on-off task behaviors. Among the behavioral cues that are found interesting there are mainly head movements, talking, smiling, chair movement. They find that students in bad emotion states tend to avoid computer screen and move their head to the side which correlates well with non-learning benefits. In [5] the authors try to find a relationship between the facial expressions that are shown on a student face during interactions with an auto tutor and underlying affective or emotional state. They record the session of the students and make afterwards 4 different judgments on student affect: self-judgment, peer judgments, judgments made by two trained judges on Ekman's Facial Action Coding System (FACS) and on characteristics of dialogue. Their conclusion is that confusion is manifested by a lowered brow (AU4), the tightening of the eye lids (AU7) and a notable lack of a lip corner puller (AU12). This pattern was found consistent with previous researches [6]. Delight can be distinguished from neutral, in particular the presence of AU7 (lid lightener), AU12 (lip corner puller), AU25 (lips part), and AU 26 (jaw drop) coupled with an absence of AU 45 (blink) segregate this emotion from neutral. These patterns are generally consistent with a smile (AU7 and AU12). Frustration was not found to be so much associated with facial expression. Although some signs that are relevant to affective state in learning environment have been identified there is still a lot to be done to enable machines to infer the affective state of a student. The problem of representation of emotion theory, in a way that machines can understand and respond to it, remains an open research question.

3.3 Image processing techniques for facial expressions

Facial expression changes the appearance of a face by inducing wrinkles or bulges in specific regions of the face or/and changing the location of facial landmarks (corners of the mouth, eyebrows...). Based on this, image processing techniques try to detect these changes by either describing the appearance of certain areas of the face or analyzing movements of facial landmarks. Based on this, the process of facial expression analysis can be decomposed into mainly three main parts [7]:

- 1 Face detection: Locating the face in the image.

- 2 Facial feature extractions: detecting facial landmarks, describing the shape of facial components, or/and describe the appearance of a particular region of the face.
- 3 Analysis of extracted facial features: analysis of change in appearance, shape, or displacement of facial features.

Lot of work has been previously done in the first stage and to our date the face detector developed by Viola & Jones remains the most used due to its real time capability and high accuracy. In the second stage of the system, the first step is the detection of facial landmarks. Many types of method exist to estimate the position of these landmarks. AAM (active appearance model) is commonly used to achieve this goal. AAM method learn the shape and appearance of facial landmarks from a set of labelled facial images in the training stage and try to detect the landmarks based on this information in the testing stage. Recently regression techniques are beginning to be used. An interesting real-time facial landmark detector that uses a cascade of regression trees, taking simple image pixel differences as features, was recently proposed that can detect facial landmarks in one millisecond [8]. After the position of landmarks have been detected some approaches describe the appearance of regions in the face. Mainly Gabor filters, Haar filters and local binary patterns (LBP) descriptors are used in addition to many other image filters. Other approaches track the facial landmark and take the change in position, speed, shape, or a combination of them in time, as descriptors. In [9] the authors track landmark points using a particle filtering framework, and analyze the change in the distance between points and their displacement to code different action units and their temporal segments. They use a rigid registration transform to cope with small out of plane head rotations and change in scale. In [10] authors use the sum of movement variations of landmark points that belong to a certain facial component as input features for a neural network. Other approaches use optical flow estimation techniques applied on a face image directly and use it to detect facial movements. The third stage analyzes the features and their evolution in time in order to categorize the shown behavior in a specific category or describe the shown behavior objectively. In [9] a rule based classification method was used to encode specific shown behavior into combination of AU and their temporal segments. Many classification systems use dynamic probabilistic networks especially HMM (Hidden Markov Models) in order to model the evolution of the learning process in time and infer the hidden state of the student from the set of observable features[11]. Machine learning classification algorithms like SVM, ensemble learning techniques (AdaBoost, GentleBoost...) have also been used to perform the high level decision.































Upper Face Action Units					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
					
Inner Brow Raiser *AU 41	Outer Brow Raiser *AU 42	Brow Lowerer *AU 43	Upper Lid Raiser AU 44	Cheek Raiser AU 45	Lid Tightener AU 46
					
Lid Droop	Slit	Eyes Closed	Squint	Blink	Wink
Lower Face Action Units					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
					
Nose Wrinkler AU 15	Upper Lip Raiser AU 16	Nasolabial Deepener AU 17	Lip Corner Puller AU 18	Cheek Puffer AU 20	Dimpler AU 22
					
Lip Corner Depressor AU 23	Lower Lip Depressor AU 24	Chin Raiser *AU 25	Lip Puckerer *AU 26	Lip Stretcher *AU 27	Lip Funneler AU 28
					
Lip Tightener	Lip Pressor	Lips Part	Jaw Drop	Mouth Stretch	Lip Suck

Figure 4 - Example of AU taken from [3].

4 RESULTS

4.1 Detection of concentration signs and facial expression analysis

In order to detect the concentration and motivation of the student, several visual cues must be detected in real-time so that a proper intervention can be directly done. The face is detected by a traditional face detector made by HOG (histogram of oriented gradient) features and a linear classifier. We use the algorithm presented in [8] that can localize facial landmarks in one millisecond after the face detection step has been completed. Fig. 5 shows the landmarks detected by the algorithm. Since a real-time approach is more important than an accurate one for pose estimation, we use a simple geometric method that can estimate the coarse pose of student's head by analyzing the angles of a triangle made by connecting the eyes and the tip of the nose. The same approach is used in [10]. Due to the symmetry of the human face, a triangle formed by connecting the eyes and the tip of the nose in a frontal picture is isosceles. When the head gets far from its frontal position moving left or right this triangle will lose its isosceles property gradually. Thus a simple analysis of the angles of this triangle can give a coarse estimation of the head pose. A similar analysis can be applied to detect up and down movements of the head. Based on this, it is easy to detect if the student is looking right, left, up, down or at the screen which can give a pertinent cue for student concentration and on-off task behaviors. Fig. 6 shows the detection of multiple scenarios based on pose estimation. Furthermore a simple temporal analysis of the pose signal can detect if the student is moving his head around constantly or he is fully concentrated on the task. A simple indicator would be the variance of the signal taken on multiple frames for example. To detect facial behavior and expressions, we analyze the displacement of the facial landmark relatively to a position showing a frontal neutral face of the student. To cope with small head pose and scale changes we register each frame using a standard affine transform to the first neutral frame. The affine transform is computed from the location of three rigid points on the face (inner left eye corner, inner right eye corner and tip of the nose) [9]. Facial behavior as raising eyebrows, lowering eyebrows, talking, smiling can be easily detected. Fig. 7 shows the approach in action. However this method gives good result for near frontal views. It is robust to small variations in head pose and scale change only. In addition it requires a frame that shows a neutral expression.

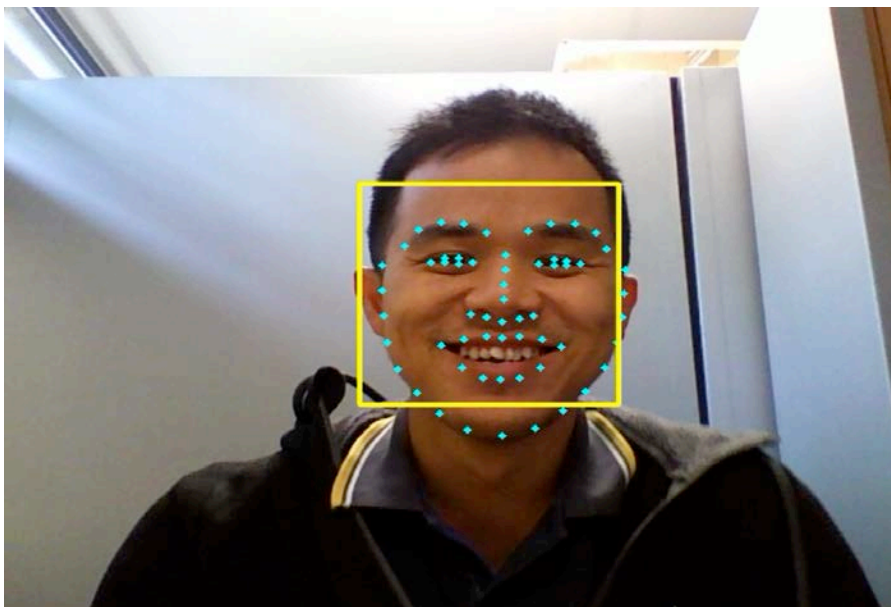


Figure 5 - Face (bounding box) and landmark (points) detection using the algorithm in [8].



Figure 6 - Different state detection. The state is indicated by the number of the green circle. In state1 the student looks left; in state 2 he looks right; in state 3 he looks up; and in state 4 he is concentrated on the given task.

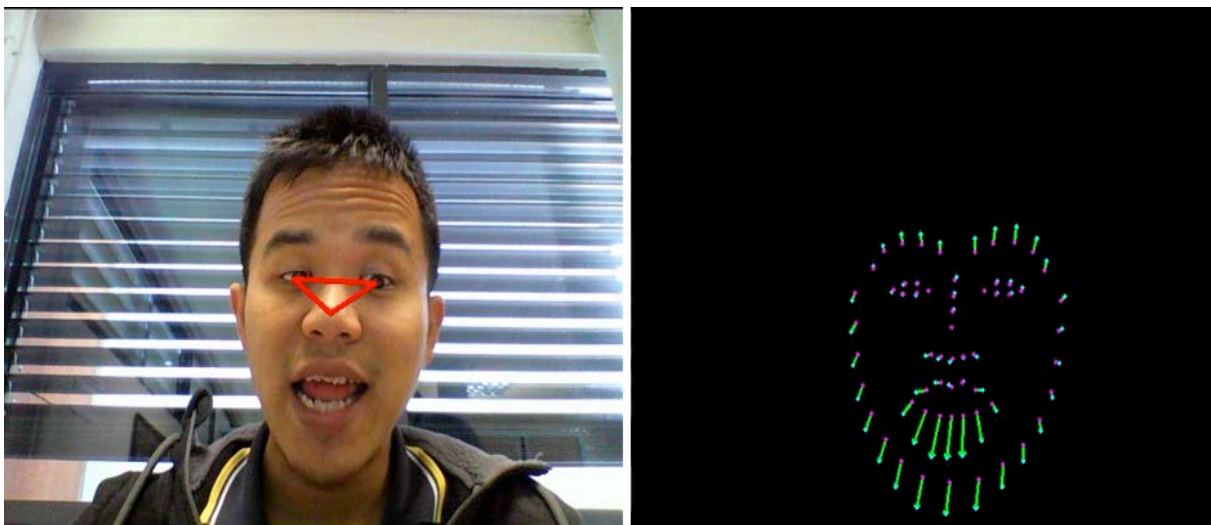


Figure 7 - Detection of eyebrows and mouth movements.

4.2 Drone for increasing the motivation of students

In order to increase the motivation of students and show the importance of the material they learn, a mini-drone equipped with a camera controlled by the student remotely is currently under investigation [12]. The scenario gives the student the possibility to control the drone in order to “walk through” the lab, examining the robot or measurement instruments. Fig. 8 shows the drone in action. Student can ask also for the drone to go and focus on a specific measurement instrument.

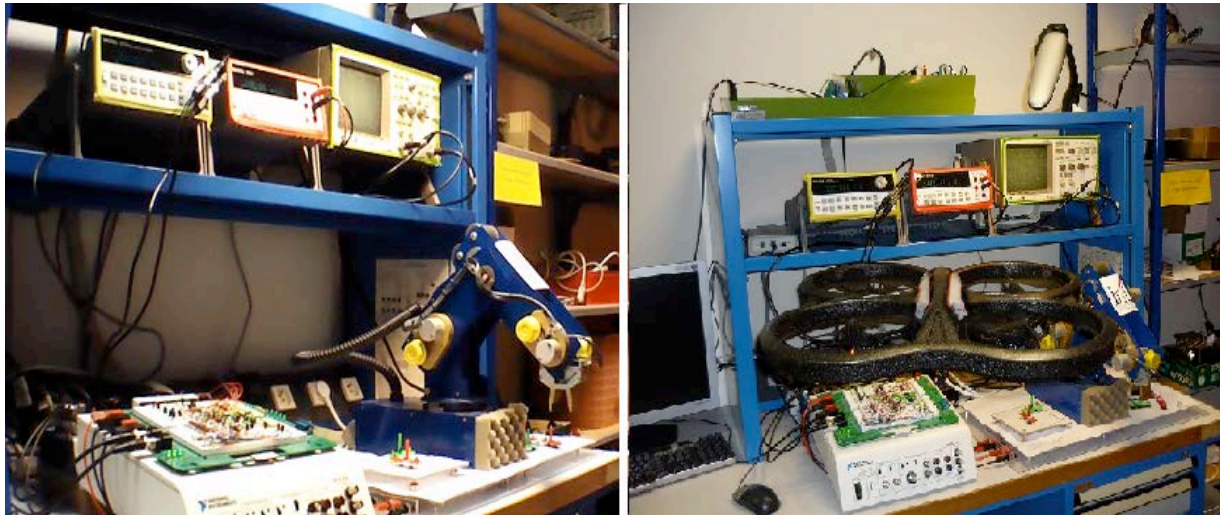


Figure 8 - Left: a picture taken from the drone camera; Right: the drone hovering around the lab hardware setup to show the measurement instruments.

5 DISCUSSION AND PERSPECTIVES

In this paper we present our implementation of a remote lab in electronics. We discuss the weak point of such a lab in dealing with the affective and cognitive state of students. We present a simple approach to detect concentration and on-task behavior based on coarse face pose estimation. We also investigate a simple approach for measuring and describing facial expressions based on the displacement of facial landmark points relatively to a neutral distribution of landmarks. However the latter approach was only found robust to small variations of scale and head pose. Future work should adopt a more robust approach. Furthermore, analysis of body movement, especially shoulder movements or posture of the student, can also help in inferring its affective state. Context information such as mouse clicks and keyboard strokes can also help. In addition, there is a need for a high level decision system that takes these evidences of student affective states into account, and performs a high level decision categorizing student affect into pre-defined states of concentration, confusion, frustration or flow. We also present the idea of using a mini-drone equipped with a camera to enhance student motivation. Future work will implement this approach to be a factor that motivates students towards learning.

REFERENCES

- [1] F. Luthon and B. Larroque, "LaboREM- A remote laboratory for game-like training in electronics," *IEEE Transactions on Learning Technologies*, vol. 8, no. 3, pp. 311–321, Jul. 2015.
- [2] M. Pantic and L. J. M. Rothkrantz, "Toward an affect-sensitive multimodal human-computer interaction," *Proc. IEEE*, vol. 91, no. 9, pp. 1370–1390, Sep. 2003.
- [3] "Facial expression recognition (Face recognition techniques) Part 1" <http://what-when-how.com/face-recognition/facial-expression-recognition-face-recognition-techniques-part-1/>
- [4] T. Dragon, I. Arroyo, B. P. Woolf, W. Burleson, R. el Kaliouby, and H. Eydgahi, "Viewing student affect and learning through classroom observation and physical Sensors," in *Intelligent Tutoring Systems*, B. P. Woolf, E. Aïmeur, R. Nkambou, and S. Lajoie, Eds. Springer Berlin Heidelberg, 2008, pp. 29–39.
- [5] B. T. McDaniel, S. D'Mello, B. G. King, P. Chipman, K. Tapp, and A. C. Graesser, "Facial features for affective state detection in learning environments," in *Proceedings of the 29th Annual Cognitive Science Society*, 2007, pp. 467–472.
- [6] S. D'Mello and A. Graesser, "Affect detection from human-computer dialogue with an intelligent tutoring system," in *Intelligent Virtual Agents*, J. Gratch, M. Young, R. Aylett, D. Ballin, and P. Olivier, Eds. Springer Berlin Heidelberg, 2006, pp. 54–67.

- [7] M. Pantic, "Machine analysis of facial behaviour: naturalistic and dynamic behaviour," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 364, no. 1535, pp. 3505–3513, Dec. 2009.
- [8] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1867–1874.
- [9] M. Pantic and I. Patras, "Detecting facial actions and their temporal segments in nearly frontal-view face image sequences," *IEEE International Conference on Systems, Man and Cybernetics*, 2005, vol. 4, pp. 3358–3363.
- [10] Y. B. Kim, S. J. Kang, S. H. Lee, J. Y. Jung, H. R. Kam, J. Lee, Y. S. Kim, J. Lee, and C. H. Kim, "Efficiently detecting outlying behavior in video-game players," *PeerJ*, vol. 3, p. e1502, Dec. 2015.
- [11] S. Afzal and P. Robinson, "Designing for automatic affect inference in learning environments," *J. Educ. Technol. Soc.*, vol. 14, no. 4, pp. 21–34, 2011.
- [12] F. Luthon and B. Larroque, "Real labworks in electronics: yes! ... but remotely controlled", *8th annual International Conference of Education, Research and Innovation (ICERI2015)*, Seville, Spain, Nov.2015, pp.8490-8500.