



International Conference on Computational Modeling and Security (CMS 2016)

Student Emotion Recognition System (SERS) for e-learning improvement based on learner concentration metric

Krithika.L.B^{a*}, Lakshmi Priya GG^b

^{a, b} School of Information Technology and Engineering, VIT University, Vellore, 632014, India

Abstract

Emotion originates from old French *smouvoir* to excite, from Latin *movēre* to disturb and *movēre* to move, which is also the theme of the paper. Emotions of a student during course engagement play a vital role in any learning environment whether it's in classrooms or in e-learning. We use excite, disturb and moving pattern of eyes and head to infer meaningful information to understand mood of the student while engaged in an e-learning environment. Emotion detection methods have been in focus of the researchers across various disciplines to understand the user involvement, effectiveness and usefulness of the system implemented or to be implemented. Our focus is on understanding and interpolating the emotional state of the learner during a learning engagement. Evaluating the emotion of a learner can progressively help in enhancing learning experience and update the learning contents. In this paper, we propose a system that can identify and monitor emotions of the student in an e-learning environment and provide a real-time feedback mechanism to enhance the e-learning aids for a better content delivery. Detection of eyes, head movement can help us understand learner concentration level. Since our metric are captured from eyes and head movement we eliminate the need of any device usage that requires physical contact to the subject under study. The proposed system helps to identify emotions and classify learner involvement and interest in the topic which are plotted as feedback to the instructor to improve learner experience.

Keywords: emotion; e-learning; concentration level

1. Introduction

Emotion plays an important role for analyzing the student's interest in class room lectures. Out of the various ways to detect emotion, the quick way is to understand the emotion symptoms through facial expression [1]. It is easy to observe the students and their reaction on a particular topic which is being taught by the instructors during an

* Corresponding author. Tel.: +91-90470 56384
E-mail address: krithika.lb@vit.ac.in

instructor led course. But the trends are changing from the instructor's teaching to self-pace e-learning with the introduction of new technologies. One of the key challenges was the un-captured emotion of the learner which plays an important role in the teaching and learning process.

Capturing the student's emotion and dynamically changing the topic to elevate the student's mood is missing in an e-learning. Most of the time in this type of environment even an important lecture or course ends up with the boredom of the student. All the effort of the instructor and their instructional aides become ineffective. In order to overcome these issues, more interactive learning technologies has been introduced in past few decades but was not able to find a better way to analyze the emotions at real-time in an e-learning environment. Thus there is a necessity for a system to analyze the mood of the student in an e-learning environment.

The consequence of this metric can also give clue to the instructor to make the topic interesting as per the student need. There is a question which can be raised in one's mind that this can be achieved by using the written feedback system but those feedbacks won't be real-time. Thus the objective of the system is to enhance the effectiveness in the e-learning environment with the help of student's emotion and also achieve real-time response as a feedback.

Based on various research, emotion plays a vital role in e-learning [2]. Similarly, the improvement of learning environment with emotion recognition has been focused by the researchers from last few decades in Computer Supported Collaborative Learning (CSCL) [3]. Experienced instructors can change their teaching style according to the feedback which they get from their students, subsequently giving extra attention. Every student could possibly require extra or reduced attention. An e-learning environment whether online or offline usually contains animated, hyper texted or text [4].

In spite of the fact that the vicinity of innovation is boundless in web-based learning environments, it is definitely not receptive to emotional responses experienced while utilizing such learning environments. The effective solutions specifically made for a classroom was not able to fulfill the requirement of the e-learner [5]. Uneven head rotation is one of the behavior that is generally found in students which is more a sign towards boredom which can be seen if the students find the content monotonous and thus future progress is affected in the relevant content.

In this paper, more focus is towards abnormal head rotation of an e-learner. Abnormal head rotation is a negative behavior and a clear sign that shows the student is not interested and bored towards the topic. Drowsiness can be detected with the help of movement of eyelid [6] and it can be detected with the help of recognition of yawning [7] which can be done by analyzing the movement of the lips. The proposed method stated can assist a developer to create a system that can detect the negative mood of the students in an e-learning environment. We continuously monitor head and eye movement to detect the boredom thereby generating a graph. Using this metrics, we can analyze whether the student felt bored or interested on a particular content of learning aid thereby providing a continuous feedback mechanism to instructor for enhancing the content and keeping the topics updated and interesting.

2. Existing Work

The ability of a person to learn cognitive procedures have been encouraged and acknowledged by many researchers [8] [9] [10]. Some research has also found the necessity of developing the learning system that can detect the emotional state of the learner [11] [12] [13]. Few researchers try to address the motivational skills with the intelligent learning system [14]. Based on the work which concluded that detecting emotion from facial expression is feasible and can predict six different emotions with 89% of accuracy [15].

In the paper [16], authors come up with more accurate findings of emotion recognition using facial expression where as it is not well examined that the facial expression can reveal the boredom and fatigue level with an individual. It has been claimed that it is necessary to detect the confidence, frustration, boredom for e-learning based environment and thus providing proper feedback [17]. The acceptance of emotions in e-learning environment has

been appealed in [18]. Although lots of work has been done on eye movement [6] [19] and yawn detection [7], drowsiness and abnormal head rotation detection has not been given much preferences in e-learning environment. Previous research has been done on emotion recognition in e-learning environment but it was not much focused on negative emotions like boredom and level of concentration.

2.1 Outline of Face Detection Algorithms

Emotion that symbolizes the inner state of an individual has captivated the researchers to define the emotion of a topic from its expression [20]. There are various face detection algorithms to extract the details of the face region. Some of the popular face detection algorithms [21] are Viola Jones, Local Binary Pattern (LBP), Ada Boost and Neural Network which are defined below:

2.1.1. Viola Jones: This is the first face or object detection algorithm framed by Viola Jones for solving the issue of face detection. It is projected in three significant ways namely through (a) an integral image (a new image) for the computation speed; (b) an efficient classifier called Ada Boost for choosing a small number of visual features from a very large set of potential features; (c) a process of cascade classifier for locating the required facial regions. High recognition accuracy and less false positive rate are its highlights. The only problem faced is it requires very long training time.

2.1.2 Local Binary Pattern: LBP is very effective to label the image features. LBP has advantages such as high speed computation and assists the wide practice in the areas of image retrieval, face recognition, etc. It is a simple approach and achieves fast computation. But the accuracy noted here is covered less.

2.1.3. AdaBoost: It is a learning algorithm which takes visual features and performs multiple iterations to generate a single composite strong learner by iteratively adding weak learners. It's simple to implement since there is need of previous information but it's sensitive to noisy data and outlier.

2.1.4. Neural Network: This system decides between several networks to improve the performance over a single network which avoids complexity for selecting the non-facial image. This type of learning network is less expensive computationally but the detection process is slow and obtains inaccurate result.

Of all the above discussed face detection algorithms, Viola Jones and LBP is considered here for the detection of the face because of its well-known characteristics.

3. Proposed System

The objective of the proposed system is to detect the concentration level of the student by continuously monitoring the head rotation and eye movement. Once the facial features are detected then it will determine whether the student is focused on the visual content (which can be an e-learning content or video lecture). The movement of the head will be then measured in terms of time duration and certain level of concentration will be analysed. **The two prominent measures of detecting the concentration level are eye movement and head rotation.**

The first criterion is to detect the eyes of the student. It is used as a distinctive characteristic to judge whether a student is interested in the topic or not and also to measure his concentration level. However, the eyes are not detected due to rotation of head or student's face turning away from screen when he is bored. All the outputs from these two components will be computed and compared to accurately judge whether the student is interested in the topic or not. The following figure 1 shows the proposed system architecture and its various stages. The first two stages deal with capturing video input and breaking into frames for processing. The next stage is the concentration level measuring module, which takes input from the sub-modules. The last stage is generation of feedback in the form of graphs which will be plotted thereby identifying the level of concentration and its duration.

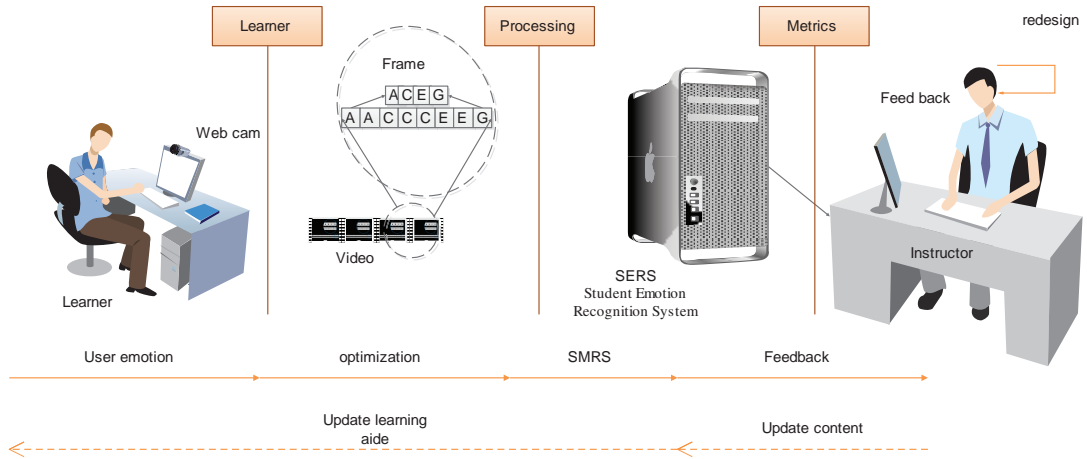


Figure 1: Proposed System Architecture-SERS

3.1 Design of the system

There are two modules in the proposed system which are eye detection module and head rotation module. The algorithm used in the proposed system is based on Viola jones algorithm and LBP.

3.1.1 Eye detection module

For detecting the status of the eyes, we need to first detect the face from the captured frame and then the eye region and hence the eyes status can be identified.

3.1.2 Head rotation module

Head rotation can be identified by detecting whether both the eyes are visible or not. If the eyes are not visible this means the head is turned around in the different direction. The graph will be plotted once the head movement and eyes are not detected with respect to the number of frames and the concentration level. The following figure 2 shows the head movement and eye detection modules.

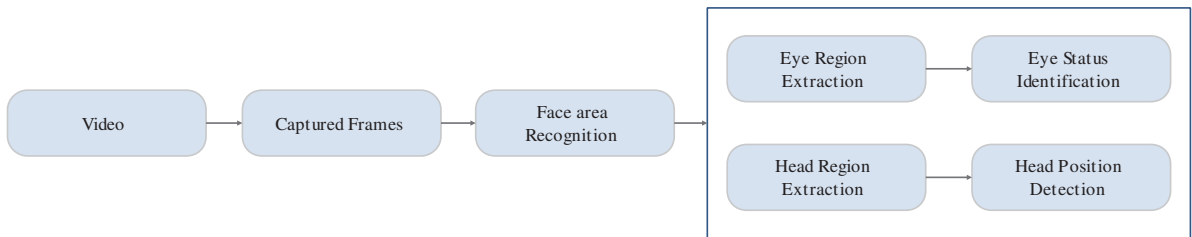


Figure 2: Head and Eye movement detection modules

4. Implementation Methodology

The implementation of SERS is done in MatLab with functions for face detection and features detection using Viola jones and LBP. Viola jones object detection framework has been used for detecting the face and other facial features. The following procedure has been proposed for detecting the three levels of concentration which are as follows:

4.1 High concentration level

For detecting the high level of concentration, it's to determine whether the eyes and the face are at right position *i.e.* whether both eyes and face are facing the camera. For determining the eyes and face status, both the eyes and the face should be facing the screen. If not, it's understood that the student is not properly facing the screen and the concentration level is given as medium or low. The number of frames for which the proper attention is not detected is monitored. If the number of frames with the medium level and low level concentration exceeds the certain limit then the student is not interested in the topic.

4.2. Medium Concentration level

For detecting the medium level of concentration, it's necessary to determine whether the eyes are at right position or not *i.e.* whether both eyes are facing the camera. For determining the eyes status, both the eyes should be facing the screen. If only one eye is visible that means the student is not properly facing the screen and the concentration level is given as medium. The number of frames for which the proper attention is not detected is monitored. If the number of frames with medium level concentration exceeds certain limit, then the student is not interested in topic. The region of interest is then found *i.e.* the eye region is detected. If the eyes are found then the counter 'E' is set zero. If eyes are not found then the random variable (for example, if 2 is assigned 'E') then it will be plotted with respect to the frames.

4.3. Low concentration level

For determining the head status, the facial features should be detected. If the facial features are not visible then the abnormal head rotation is detected. Head rotation detection can be done by measuring the period of time for which the face is detected in right position. If the face is not detected for few seconds then it is a normal head rotation. If it is not visible for long time then the student is not interested in the topic. The number of frames for which the face is not detected is monitored. If the number of frames without face detection exceeds certain limit then the concentration level is reduced.

The region of interest is then found *i.e.* the facial feature is to be detected. If the face is found, then there is no head rotation and the counter 'H' that is the head rotation counter is set zero. If the face is not detected, we can assign the counter with any variable to show the variation in the graph. If it exceeds 5 frames, then abnormal head rotation is detected else it's a normal one. The figures 3 and 4 show the flowchart of the eye and head rotation detection algorithm.

5. Results and Discussion

The system is tested on ten minutes video lecture in which a student has watched the video lecture and the following results are obtained.

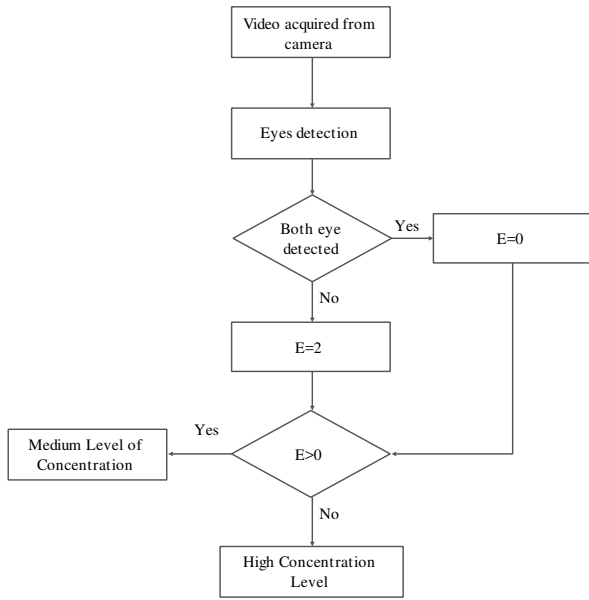


Figure3:Medium concentration level detection

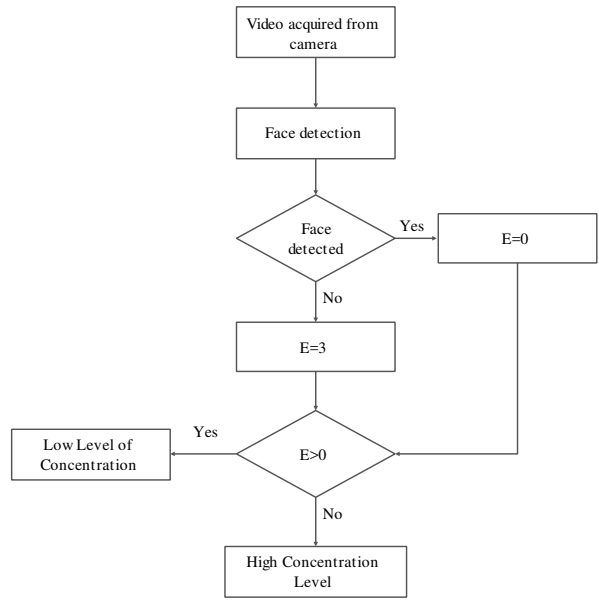


Figure4:Low concentration level detection

5.1. Head Movement

The graph is plotted with the counter value 3 with respect to the frames. The plot indicates the head is not visible in the student's video observed by the web camera, which shows that the student was not in the proper position and was not focused during the frames in which the variations are observed. In the below graph, 0 indicates that the head was visible and 3 indicates that head is not visible on a particular frame. The figure 5(a) shows the graph plotted for the head position in frame per second(fps) and Level of Concentration(LoC). So by analyzing the above graph we can conclude that there is a variation on the frame number 6, 7,12-13, 21, 23, 42-52,78-80 and 117-119. Therefore we have got the clear observation that during the mentioned frame the variation is found.

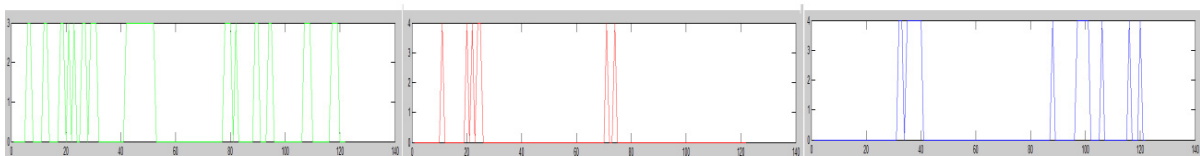


Figure5(a): Head movement graph Figure 5 (b):Left eye graphFigure 5(c):Right eye graph

5.2 Left eye

The graph is plotted with the counter value 4 with respect to the frames. The plot indicates the rotation of head towards the left which indicates that only right eye was visible in the student's video observed by the web camera, which shows that the student was not in the proper position and was not focused during the frames in which the variations are observed. In the below graph, 0 indicates that the left eye was visible and 4 indicates that left eye is not visible on a particular frame. The figure 5(b) shows the graph plotted for the left eye in frame per second(fps) and Level of Concentration(LoC). So by analyzing the below graph, we can conclude that there is a variation in the

frame numbers 11, 20, 22,24,25,71 and 74. Therefore we have got the clear observation that during the mentioned frame the variation is found.

5.3 Right eye

The graph is plotted with the counter value 4 with respect to the frames. The plot indicates the rotation of head towards the right which indicates that only left eye was visible in the student's video observed by the web camera, which shows that the student was not in the proper position and was not focused during the frames in which the variations were observed. In the below graph, 0 indicates that the left eye was visible and 4 indicates that right eye is not visible on a particular frame. The figure 5(c) shows the graph plotted for the right eye in frame per second (fps) and Level of Concentration (LoC). So by analyzing the above graph, we can conclude that there is a variation in the frame numbers 33,34,36-41,88,97-101,106,116 and 120.

5.4. Concentration level

Concentration level can be detected by doing the analysis of the entire three component's data together. The concentration level can be categorized into three levels namely high, medium and low respectively. The data value 4 is considered as medium, the data value 0 is considered as high because there is no variation and the data value 3 is considered as high level of concentration. The count value 3 is for the head and thus if the head is not visible it means the student is not at all viewing the screen and thus the concentration level is considered as low. The table 1 below represents the various level of concentration.

Table 1: Different Concentration level

Frame	Left Eye	Right Eye	Head	Concentration Level
1	4	0	0	Medium
2	0	0	3	Low
3	0	4	0	Medium
4	0	0	0	High

If there is any variation in either the left or right eye then the concentration level is considered as medium because there is a possibility that the student might be looking the screen by tilting his head either towards right or left at the same instance. If there is any variation in the head then the concentration level is considered as low because the student is not at all viewing the screen at that particular instance. If there is no variation at all it means the student is focusing on the screen and thus his concentration level is considered as high.

6. Test Results on Students

The system has been tested with 5 different students by showing them a course lecture (video) where different results have been observed which is been shown in figure 6. The different concentration levels have been observed throughout the video *i.e.* high, medium and low. Thus after analyzing, it is concluded that at some consecutive frames the concentration level is either low/medium/high. Therefore it can be said that the student was not attentive in reading the content which was displayed at that time period.

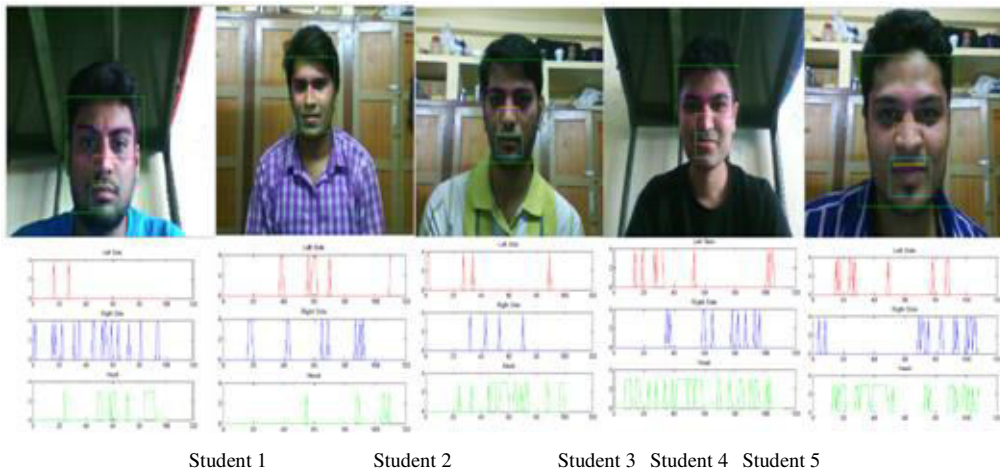


Figure 6: Test on various students and generated graphs

6.1 Discussion

The below figure 7 is score graph plotted between the score and the number of frames. The figure 7(a)(b)(c)(d)(e) are the individual graphs of the student1,2,3,4 and 5. The score are the values allotted to the level of concentration such that, for low level the score is 1, for medium level the score is 2 and for high level the score is 3. Thus the high level of concentration which is shown with score 3 and seen at the top level highlighted as green. The medium level is with score 2 is highlighted in yellow and the low level concentration level with score 1 is highlighted in red at the bottom level.

Beginning with the student 1, we can say that the concentration level is observed the least in more than 3 consecutive frames (51-54, 62-64, 86-93) which is the maximum. Therefore between these frames the student is not at all interested in the content. Next with student 2, we can say that the concentration level is minimum in more than 3 consecutive frames (89-90, 106-107). Therefore between these frames the student is not at all interested in the content. Then with student 3, the concentration level is with score 1 that can be seen as red lines at the bottom level. So for student 3, we can say that the concentration level is minimum in more than 3 or more consecutive frames which are 48-51, 54-60, 98-100. Therefore between these frames the student is not at all interested in the content.

Next with student 4, the concentration level is with score 1 that can be seen with the red lines at the bottom level. So for the student 4 we can say that the concentration level is minimum in more than 3 or more sequential frames which are 9-13, 16-18, 42-58, 50-52, 54-58 and 83-85. Therefore between these frames the student is not at all interested in the content. Finally with student 5, the concentration level is with score 1 which is highlighted in red line at the bottom level. So for the student 5 we can say that the concentration level is least in more than 3 consecutive frames (19-22, 32-37, 40-49, 90-93, 109-112). Therefore between these frames the student is not at all interested in the content.

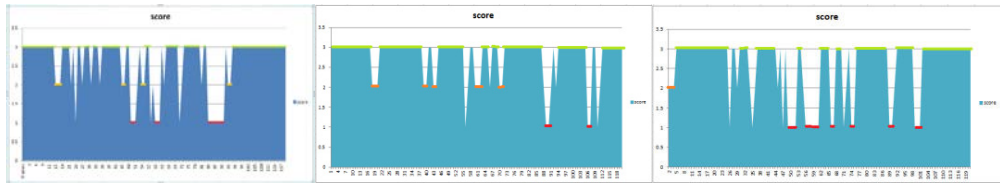


Figure7(a):Score graph of Student 1

Figure7(b):Score graph of Student 2

Figure7(c):Score graph of Student 3

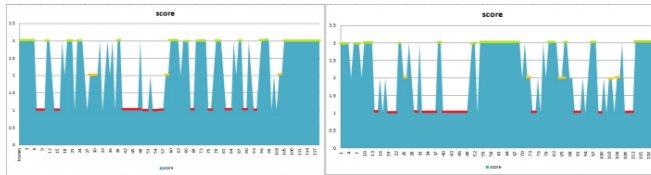


Figure7(d):Score graph of Student 4

Figure 7(e):Score graph of Student 5

Figure 7: Overall score graph of different students

6.2. Analyzing the students output

The output of each student is then analysed collectively and the outcome will be considered as a feedback to the tutor. The time frame at which the students are not interested in the topic and where the improvements are needed are done on the tutor’s side. The table 2 below shows the result of the test on each student.

Table 2:Test result on each student

Student	Frame - high concentration level	Frame - medium concentration level	Frame - low concentration level
Student 1	1-2,4-15,19-22,24,26-27,29-31,33-35,37-45,48-49,54-55,57-58,64,66-71,74-81,83-84,93,96-120	3,16-18,23,28,32,36,46-47,53,56,60,73,82,94-95	25,50-52,59,61-63,72,85-89
Student 2	1-18,21-38,41-42,45-55,58-59,67-69,72-87,92,94-105,108-109,112-120	18-20,39-40,43-44,57,60-62,66,70-71,88,91,93,111	56,89-90,106-107,110
Student 3	4-24,26-27,29-32,36-43,45-47,52-54,61-64,67-69,75-87,91-97,101-120	1-3 ,28, 33, 35, 44,54,71,90	25,34,46,48-51,55-60,65-67,70,73-74,88-89,98-100
Student 4	1-7,12-13,18,20-22,24-26,33,36,40-41,61-64,66-68,71-75,79-81,86,88-89,92,97-100,106-120	14,19,27,29-32,35-37,53,59-60,65,78,82,87,93,96,102,104-105	8-11,15-17,23,28,34,42-45,50-52,54-58,69-70,83-85,90-91,94-95,101,103
Student 5	1-4,6-8,10-13,23-25,27,31,38-39,51-52,71,78,80-83,86-87,96-88,108,113-120	5,9,16,18,26,28,50,70,72-73,76,79,84-85,88-89,93,95,101,103-104,106-107	14-15,17,20-22,29-30,32-37,40-49,53,74-75,77,90-92,94,99-100,102,105,109-112

Now, by finding the common sets we can find the frequent patterns *i.e.*frequently occurring frames in all three different concentration level of five students which are shown in table 3.

Table 3: Frequently occurring frames

	Frames with high concentration level	Frames with medium concentration level	Frames with low concentration level
Common frames	1-27,30-41,60-69,74-89,101-120	27-29,70-73	42-58,90-100

Thus with the above analysis it is clear that the maximum number of students felt bored or was not interested during the frame duration of 42-58 and 90-100. And hence the content appeared during these frames need to be revised.

7. Conclusion and Future work

In this paper, we have come forward with a constructive, iterative and a progressive method to maintain and redesign the learning aidtherebysatisfying the learner's interest withup-to-date instructional content. The quality could be achieved better based on the concentration level recognized using eye and head movement. The proposed system is efficient enough to detect the negative emotions like boredom or lack of interest of the student in e-learning environment. In future, the same methodology can be applied for other emotions as well.

References

- [1]Ekman, Paul, Wallace V. Friesen, and Phoebe Ellsworth, "Emotion in the human face: Guidelines for research and an integration of findings", Elsevier,2013.
- [2]Immordino-Yang, M.H. and Damasio, A., "We feel, therefore we learn: the relevance of affective and social neuroscience to education," *Mind, Brain, and Education*, vol. 1, pp. 3-10, 2007.
- [3]Baker, M., Andriessen, J. & Järvelä, S., "Affective learning together. Social and Emotional dimension of collaborative learning," Routledge, 2013.
- [4]A.A. Economides, "Adaptive Feedback Evaluation," *Proc. Fifth Int'l Conf. Distance Learning and Web Eng.*, pp. 134-139, 2005.
- [5]M. Martinez, "Key Design Considerations for Personalized Learning on the Web," *Educational Technology and Soc.*, vol. 4,no. 1, pp. 26-40, 2001.
- [6]Danghui Liu, Peng Sun, YanQing Xiao, Yunxia Yin, "Drowsiness Detection Based on Eyelid Movement", *Education Technology and Computer Science (ETCS), Second International Workshop, IEEE*,vol.2, pp.49-52,2010.
- [7]ShabnamAbtahi, Behnoosh Hariri, ShervinShirmohammadi, "Driver Drowsiness Monitoring Based on Yawning Detection", *Instrumentation and Measurement Technology Conference (I2MTC), IEEE* , pp.1-4,2011
- [8] A. Ortony, G.L. Clore, and A. Collins, *The Cognitive Structure of Emotions*, Cambridge Univ. Press, 1988.
- [9]Damasio, A. R., and Stuart Sutherland., "Descartes' Error: Emotion, Reason and the Human Brain." *Nature* 372.6503,1994.
- [10]Heckhausen J, HeckhausenH, "Motivation and action":pp.384-443, 2008
- [11]R. Picard, S. Papert, W. Bender, B. Blumberg, C. Breazeal, D.Cavallo, D. Machover, M. Resnick, D. Roy, and C. Strohecker, "Affective Learning—A Manifesto," *BT Technology J.*, vol. 22, no. 4,pp. 253-269, 2004.
- [12]A. de Vicente and H. Pain, "Motivation Diagnosis in Intelligent Tutoring Systems," *Proc. Fourth Int'l Conf. Intelligent Tutoring Systems*, pp. 86-95, 1998.
- [13]B. du Boulay and R. Luckin, "Modeling Human Teaching Tactics and Strategies for Tutoring Systems," *Int'l J. Artificial Intelligence in Education*, vol. 12, no. 3, pp. 235-256, 2001.
- [14]A. de Vicente and H. Pain, "Informing the Detection of the Students' Motivational State: An Empirical Study," *Proc. Sixth Int'l Conf. Intelligent Tutoring Systems*, vol.2363 of the series *Lecture Notes in Computer Science*, Springer, pp. 933-943,2002
- [15]M. J. Den Uyl and H. Van Kuilenburg, " The FaceReader: Online facial expression recognition," *Proc. Measuring Behaviour*, pp. 589-590, 2005.
- [16] S. L. Happy, A. George and A. Routray, "A real time facial expression classification system using Local Binary Patterns," in *IEEE 4th Int. Conf. on Intell. Human Computer Interaction*,pp.1-5, 2012.
- [17]Woolf, Beverly, Winslow Burelson, and IvonArroyo, "Emotional intelligence for computer tutors.", *Workshop on Modeling and Scaffolding Affective Experiences to Impact Learning at 13th International Conference on Artificial Intelligence in Education*, vol.616. 2007.
- [18]Michalis Feidakis1, Thana Daradoumis, SantiCaballé, Jordi Conesa2, "Measuring the impact of emotion awareness in e learning" , *Seventh International Conference on Complex, Intelligent, and Software Intensive Systems*,pp.391-396,2013.
- [19]H. Deubel and W. X. Schneider, "Saccade Target Selection and Object Recognition: Evidence for a Common Attentional Mechanism," *Vision Res.*, Elsevier, vol. 36, no. 12,pp.1827-1837,1995.
- [20]Konar, Amit, and ArunaChakraborty, "Emotion Recognition: A Pattern Analysis Approach", John Wiley & Sons, 2014.
- [21]Gupta, Ms Varsha, and Mr Dipesh Sharma. "A Study of Various Face Detection Methods." *METHODS* 3.5,2014.