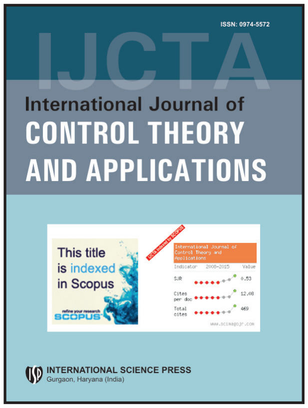

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 9 • Number 44 • 2016

Adaptation to Emotion Cognition Ability of Learner for Learner-centric Tutoring Incorporating Pedagogy Recommendation

Neelu Jyothi Ahuja[a], Ninni Singh[b] and Amit Kumar[c]

[a-c]School of Computer Science and Engineering, University of Petroleum and Energy Studies, Dehradun, Uttranchal, India, 248001. Email: [a]neelu@ddn.upes.ac.in; [b]ninnisingh1991@gmail.com; [c]amitanit007@gmail.com

Abstract: An initial tutoring strategy characterized by a learner profile and learning style is assigned to a learner, in response to an assessment conducted in the pre-tutoring phase. As tutoring commences, the learner activity is recorded, as image of the learner at periodic intervals. The performance parameters and captured image elucidating the facial expression is used to interpret the emotion, classified to determine the learner's degree of engagement and comfort level with the tutoring session. Comparison with preset threshold values triggers the decision to change the tutoring strategy for the learner. The aim is to make learning most effective, by tutoring in a manner most suitable to the learning preference of the learner. A tutoring system developed with its detailed architecture and working of the emotion recognition module of the tutoring engine, the test results obtained by testing on 20 participants have been presented in this paper.

Keyword: Emotion recognition, intelligent tutoring system, Learner style, Pedagogy recommendation.

1. INTRODUCTION

Artificial intelligence is an advanced field of research. Its use in the field of education particularly in making teaching and learning more effective has been significantly exploited in the recent past. This has resulted in development of variety of educational computer artifacts. Further developments have caused origin of the field of intelligent tutoring. This field developed at the intersection of computer science, cognitive psychology and educational research has gained immense popularity in current times particularly due to the development of Intelligent Tutoring systems (ITS) [1]. These systems termed as cognitive tutors offer tutoring in a manner that best suits the learning preference of the learner. It is a program that not only behaves like a human tutor but also follows the rules and instructions based on the progress and behavior of the learner.

ITS is an intelligent computer system which solves the learner's issue by bestowing the feedback and hints. It traces the learner's activity that predicts the learner's mastery to understand the learner's psychological mind. Cognizing psychological mind of learner makes a tutoring system an *"intelligent"* tutoring system because it solves the learner issues and offers the tutoring content in such a manner that learner can grasp easily and effectively. An ITS is distinctly different from a typical e-learning system, which is a web-based learning system, that facilitates a learner to explore a specific domain or course contents via internet. Like ITS, these systems do not adapt to the learner's learning needs and also do not offer learner-specific feedbacks and hints. ITS being adaptive in nature, has gained immense popularity in current times.

In the earlier twentieth century, the intelligent tutoring systems have grown most prevalent and effective way of learning and teaching. However, there exist some traditional comparative studies that depreciate this artificial teaching and learning techniques [2]. Resource sharing and Intelligence is one of the important key features of intelligent tutoring system (ITS) that make it different from the other traditional techniques [3] [4]. Mostly ITS intelligence is focused on the learner cognitive ability thus overlooking the real environment and original feelings. Learner cognitive ability is achieved when there is an emotional exchange between the learner and the tutoring system.

An adaptive intelligent tutoring system is the system which incorporates the learner's emotions along with the tailored personalization. For instance, in Wolcott [5], it is argued, human tutor relies on nonverbal ways like facial expressions, body language and eye contact to identify the psychological state of learners, which signify the degree of understanding and engagement. Thus, integrating an emotion recognition feature into an intelligent tutoring system will embellish its skills to administer the essential advice, effective tutoring and last but not the least make tutoring session more worthwhile.

A Tutoring system holds a knowledge or subject matter capsule which it offers to its learner. Presented below is the knowledge domain used in the current work.

The petroleum exploration is accomplished by a petroleum geologist, with primary interest to discover geological structures appropriate for hydrocarbon amassing. Seismic images are widely used by geologists and geophysicists to describe terrestrial subsurface. By utilizing their skills and knowledge human specialists manually interpret the seismic images. Nevertheless, in this whole procedure of interpretation, there is a possibility that same seismic image is differently interpreted by different seismologist. This uncertainty is because of the absence of formal thumb rules of interpretation and therefore, each geologist uses his or her own skills, expertise that he or she has established over past years [6].

Emotion Intelligence is an ability to identify or recognize the emotion of individuals and use it as a parameter to understand the psychological state of self and other individuals mind. This field has developed at the intersection of study of feeling and thinking, which are automatically generated in human mind when some situation is encountered. For example anger, smile, surprise and normal. Emotional intelligence can also be defined as an action in which someone tries to improve own cognition ability by undertaking others emotions. For example like Human tutor in a face to face class recognizes the student's emotions and tries to deliver lecture as per student cognition or state of mind and also provides a suitable problem solving feedbacks. As a result this type of tutoring prevents the student disengagement from learning. Thus ITS enhance the learner performance, improve degree of engagement and give necessary direction by means of feedbacks. These are common benefits of ITS integrated with emotional intelligence.

Emotion plays an important role during learning and teaching session as it is one of the key features for the establishment of individuality and pride. This feature benefits the tutoring system in judging the grasping potential of a learner by sensing the psychological states like happy, sad and ambiguous from his or her facial expressions. As per the requirement, the ITS changes the tutoring strategy by modifying the pedagogical style as per learner preference and grasp.

In this contribution an emotion recognition technique is projected, which is further integrated with Intelligent Tutoring System (ITS) developed to present its execution. The proposed scheme automatically partitions the captured image frame of learner, for emotion recognition it utilizes color base facial feature map, for classification it utilizes Bezier curve and some distance measure is projected and implemented.

The rest of the paper is organized under following sections: In section two related works have been discussed. In section three the proposed emotion recognition technique, its architecture algorithm, along with detailed justification have been presented. In section four implementation and experimental results have been presented. The section five presents the conclusion and future work.

2. LITERATURE REVIEW

Recent research is more focused towards emotional-adaptive intelligent tutoring system. The goal of this tutoring system is to serve the teaching material in such a manner that learner can easily grasp the concept. These systems observe the learner behavior and change the pedagogical styles as per the learner interest.

Murthy and Jadon [7] [8] has taken six emotions i.e. Sad, Happy, Surprise, Disgust, Normal and Ambiguous into consideration. For emotion recognition they utilize Eigen faces. Their main inspiration was to use the dimensionality reduction technique (Principal Component Analysis) for a larger set of data. By using this technique they achieved 83% accuracy. However they utilize PCA, which incorporates its own drawback and makes this technique more expensive because computation of the covariance matrix is performed at the expense of efficiency mainly when abundant dataset are encompassed for training purpose [9].

Lien and Colleagues [10] has taken two approaches into consideration i.e. SVD (Singular Value Decomposition) and direct matching. Firstly, they transformed images into corresponding transitional expression matrices, then they perform a direct matching operation. These two approaches impose certain drawbacks i.e. A direct matching operation provides no or little precision for computing correlation coefficients therefore facial image conversion would result in producing asymmetrical output facial images.

In Arumugam [11], for feature extraction they integrate FLD (Fisher's Linear Discriminant) and SVD (Singular Value Decomposition) and for classifier they utilize Radial Basis Function. Mainly they focused on only three types of emotions i.e. Disgust, Happy and Anger. The major drawback of this approach is that they achieve low accuracy by utilizing this combination. The computation of naïve SVD is often going outside the skill of various machines [12].

The ITSPOKE intelligent tutoring system is a dialogue system [13] that mentors a learner through a long physics qualitative question by describing each and every aspect of misconception. In order to identify the learner emotional state i.e. positive, neutral, and negative they utilize sound and prosodic features mined from learner speech. By using this technique they achieved 80.53% accuracy.

AutoTutor [14] successfully addresses more refined emotional states i.e. confusion, boredom, frustration, flow and neutral. They observed emotions from body posture, conversational cues and facial features. For affective replies from a tutoring system they utilize some animated pedagogical agent having animated facial expression, sound and speech.

Woolf [15] has taken five emotions i.e. self-confidence, frustration, boredom, motivation and fatigue into consideration. He utilized different heuristic rules for providing an effective response (changing voice and gesture, sympathetic response, graphs and hints, text messages) to learner's cognition state. He computed the degree of engagement in relation to overall impact on learner's learning and behavior.

Mao and Li [16] [17] proposed an Emotion-Sensitive ITS named "*ALICE*". Alice utilizes emotion agent that is effectively capable of recognizing the emotions of the learner through text, speech and facial expression.

They consulted human tutors to discuss all possible scenarios and developed rules. Thus 'ALICE' behaves closest to the human tutors, through the ongoing tutoring sessions.

Tian [18] proposed a framework based on the intersection of active listening and affective computing. In this framework emotion is recognized by analyzing textual interaction such as sentence typed, group discussion, chatrooms and question and answer. For providing an effective text-based response they utilize case-based reasoning.

Ekman et. al., [19] proposed a system (named as Facial Action Coding System, FACS) that elaborates distinct facial expressions. These distinct facial expression changes are based on inventory of all action units (AU). FACS determines 46 action units that are responsible for the facial expression change. For determining facial expression some rules are defined by considering action units. Table 1 shows rules generated for emotion recognition.

Table 1
Rules for Emotion Recognition

S.No.	*Emotions*	*Rules*
1.	Smile	1. Eye opening become narrowed
		2. Both the lips corners stretched or pulled obliquely
		3. Mouth is opening
2.	Sad	1. lower lip corner depress
		2. Eye is stiffen closed
3.	Surprise	1. Mouth Stretched
		2. Mouth opened
		3. Eyes are opened
		4. Uppermost eyelid Raiser

3. PROPOSED METHODOLOGY

This section presents the detailed methodology of identification of learner's emotions by recognizing the facial expression during the ongoing tutoring session of the developed intelligent tutoring system christened as SeisTutor. The working of this part of the tutor has been explained, with justification of outcomes.

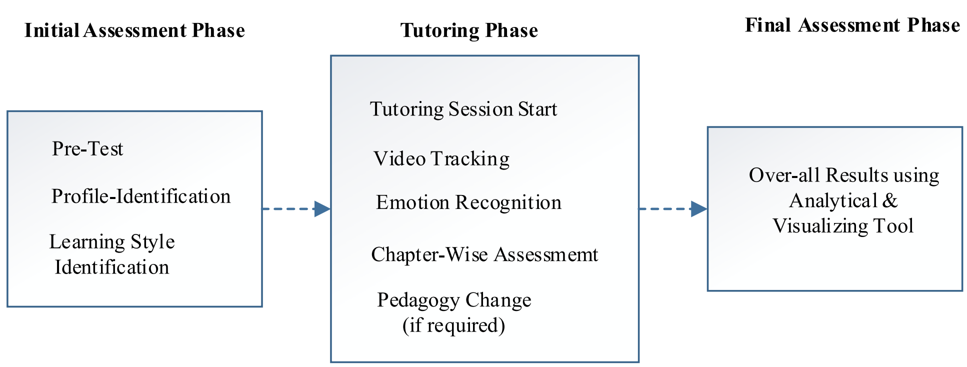

Figure 1: Intelligent Tutoring System Flow Diagram

Annexure I at the end of this paper contains the glossary of specific terms such as tutoring strategy, pedagogy, learner level, learning style etc. used in this paper.

Tutoring strategy and pedagogy style are interrelated to each other. Tutoring strategy is practically implemented in a particular pedagogy style of delivery.

For the current scope of work, three learning levels and three (Text, Visual and Verbal) learning styles have been used. Thus there are 9 tutoring strategies and the tutor maintains a given subject matter topic in nine possible variations for delivery to the learner as per adjudged tutoring strategy.

Architecture of Proposed Intelligent Tutoring System (ITS) is shown in Figure 1. The working is discussed in three phases. The first phase is Initial Assessment Phase, in which three tasks are performed. In Pre-test task set of questions are asked to learner and performance parameters such as number of correct answers (from a set of questions asked during pre-test, how many of were correct), number of hints taken (hints are available for every questions asked in pre-test. Number of hints taken indicates how many times hint was availed by the learner for a particular question), Promptness (how much time learner takes to answer a particular question), number of attempts (number of attempts taken on a question is incremented as the number of hints taken increases) and time delays (each question has pre-decided time limit to answer. Time delay is amount of time beyond this pre-set time counter) are recorded. In the Profile-Identification and the Learning Style identification task of phase one, outcomes of Pre-test are taken as input which help to decide the Learners Profile and Learning Style. The answers of the questions that were asked in pre-tutoring phase are indicative of the learning styles (Read & Write, Visual and Verbal). Thus at the end of first phase learning styles most suitable to the learner are identified. Along the same lines, the first phase also yields a learner level. The output is available as a priority list that helps to choose initial and future tutoring strategies.

The Second phase is the Tutoring Phase. In this phase, tutoring session starts for the learner as per the adjudged Learner profile and Learning style. During ongoing tutoring session, activities and emotions of the learner are captured. Several checkpoints have been incorporated to provide a provision for the ITS to change the tutoring strategy to ensure better understanding of the learner. As the tutoring session progresses at the end of every chapter, a test is administered for the learner. The performance criteria such as number of correct answers, number of hints accuracy, number of attempts and time delay are recorded. The recorded parameters and emotions form the basis for decision of change of the tutoring strategy for the learner.

Facial expression depicting emotion is of prime importance for understanding the state of mind of the learner. The scope of the paper is limited to detection and classification of emotion based on the facial expression of the learner and accordingly triggering the change of tutoring strategy. The other parameters mentioned above are equally significant, in contributing to this decision, but are not in the current scope of the paper. The ITS developed is built exactly with an analogy of a human tutor who interprets the non-verbal responses, especially the facial expressions and behavior of the learner to ascertain the learning state, engagement and comfort level, and accordingly changes his style, if required.

The third phase is the Final Assessment phase. In this phase, SeisTutor shows the chapter-wise performance of the learner by utilizing some analytical and visualizing tools (pie chart, bar chart etc.). Along with chapter wise performance, it also shows the chapter wise degree of engagement, chapter wise number of questions attempted, the performance of the learner with, initially assigned pedagogy and if, at any checkpoint the pedagogy style had been changed.

The working of this module is presented in three phases: In the first phase, the learner's image is automatically captured at regular time intervals using webcam, followed by identification of the facial region of the captured frame. In the third phase recognition of facial expression is done. The Figure 2 and Figure 3 shows the steps involved in working of emotion recognition module.

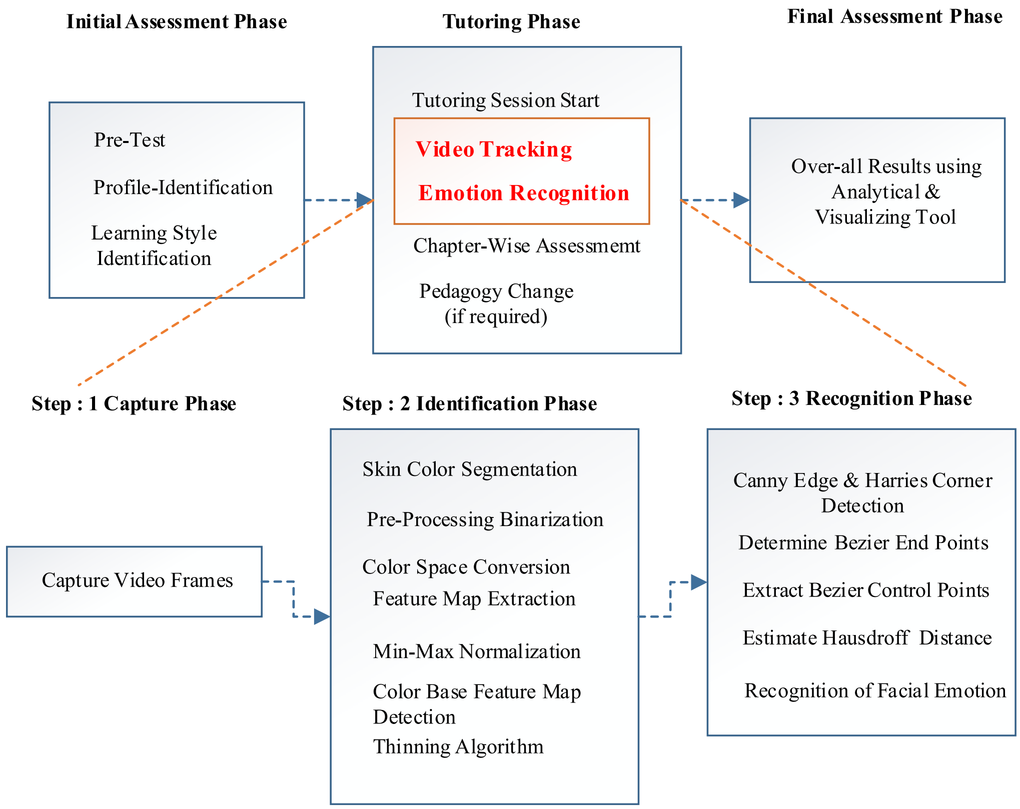

Figure 2: Emotion Recognition Module

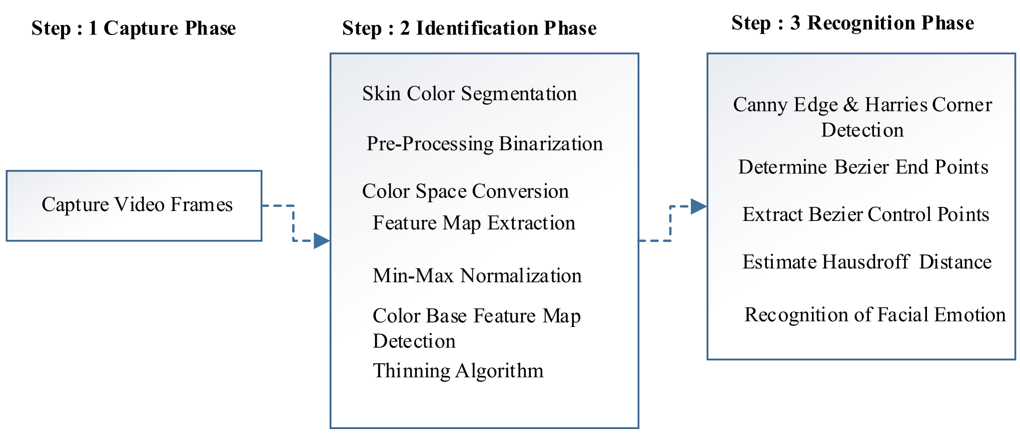

Figure 3: Flow Diagram of emotion Recognition module

As the tutoring session progresses, the webcam captures learner images periodically. The section below presents the workflow of the emotion recognition module.

In the second phase of emotion recognition, the skin color segmentation takes captured frame as an input. Skin color segmentation is used to find out the skin color of the learner. In skin color segmentation the task is to separate skin color region from a non-skin color region. In order to separate skin color from the non-skin color region, the normal RGB color segmentation is not considered to be a better solution because RGB has three components (*r* (red), *g* (green), *b* (blue)) which not only includes colors but also includes luminance. The luminance may vary from person to person due to ambient lighting on person's face. This luminance can be eliminated if RGB colored image is converted to the chromatic colored image.

$$r = \frac{R}{(R + B + G)} \tag{1}$$

$$b = \frac{B}{(R + B + G)} \tag{2}$$

$$r + g + b = 1 \tag{3}$$

Nevertheless, the result after skin color segmentation may not resemble the skin instead it identifies the region which is not having the skin color like skin and this region is eliminated from the process of emotion detection.

Now the second task is to detect the face (facial features, nose, mouth, eyebrow, and lips). In order to detect the face, the RGB snap is converted into binary image, this method is called as Pre-processing binarization. In order to convert binary image, the average of RGB of each ((r (red), g (green), b (blue))) pixels is calculated. Then the average value is compared with a threshold value, if the average is more than the threshold value then it is changed to white pixels otherwise it is changed to black pixels. After that, it detects the face and separates out the face from the captured input image.

Then to identify and locate the facial area of interest (Eyes, eyebrows, nose, lip boundaries and lower chin boundaries), feature map is used. Furthermore, this region is extracted from the snap. These snaps are mapped to eliminate the background features. The facial feature identification parameters are chosen in such a manner that it separates out the area of interest of face listed above. These boundaries are obtained using a Sobel operator. In order to find out the region of interest, i.e. for eye detection the scanning is started from the mid position of the captured image, trying to determine the consecutive black pixels and consecutive white pixels. After identifying this maximal width of the white pixels is ascertained by exploring in all directions (vertical, left and right). Then comparison is performed between the calculated widths and the previous cut (area of interest cut). If calculated width is less than the previous cut then scanning is stopped because in that case the area closer to the eyebrow may have been achieved. Now if this condition is met then the incision of the face from the beginning of forehead to a height is done.

$$\text{Height} = 1.5 \times \text{Width} \tag{4}$$

As mentioned above given incision of the face having width x then height becomes $1.5 \times x$. Now for interested feature extraction, in order to find out the mid-point between two eyes, the scanning is started from $x/4$ to $(x - x/4)$. After finding the mid position between eyes next task is to find the maximum continuous white pixels along the height between the distances (mid position of eyes). After finding this the next task is to find the upper or highest position of the eyes, and then further proceed to extract the left and right eye. For left eye the search is from Mid to $x/8$ and for right eye, the search is from Mid to $x - (x/8)$. There are some white pixels between the eyes and the eyebrow. To make the connection between the eyes and the eyebrows some black pixels are placed. These black pixels are placed in such a manner so that for the left eye, these are placed between Mid/2 and Mid/4 and for right eye these are placed between Mid + ((x + Mid)/4) to + 3 × ((x + Mid)/4). The height of the black pixels for both the eyes is (Height – (Beginning position of eyebrow/4)). Here Height indicates the Height of the snap.

Now the next task is to identify the lower boundary of the eyes. This can be accomplished by searching vertically for black pixels. For right eye, the scanning is started from Mid + ((x = Mid)/4) to Mid + 3 × ((x = Mid)/4) and for the left eye, it is started from Mid/4 to Mid – (Mid/4). After extracting all these points the last task is to identify the left side of right eye or starting pixels of the right eye then the search is made horizontally from mid to beginning point of the black pixels and vertically in between upper and lower point or positions of eyes and perform the same operation for the left eye. After finding all these points the incision of the eyes from the snap is done.

Now after eye feature extraction the next task is to extract another feature i.e. Lip Extraction. In order to perform this operation, the upper and lower points or positions of lips are to be identified. For the upper position of lip firstly the region between eyes and forehead is to be identified. After determining this, the distance to the lower position of the eyes is accumulated. Thus, by doing this, the upper position of the lip is identified. Now next job is to determine the lower position of the lip. As (1/4) position of left eye will become the upper points for lip thus for the lower position of the lip, it will be (3/4) of the right eye. After determining all these points incision of the lip from the snap is done. As soon as the boundaries are discovered, the thinning algorithm is applied for further thinning the boundaries of extracted features, i.e. lip and eyes.

To depict these boundaries in the form of curves two end points are needed. As the Bezier curve is constructed with the help of four coordinates, the two points are the end control points of the curves. Using this end control points other points are computed by Bezier interpolation method forming a three neighboring straight line.

Point X_m describes the 2D shapes curve $X_b : (p_m, q_m)\ 0 \leq m \leq n$ these points are used for further calculation in order to determine the X_b.

Whereas,

X_b defined the route between X_0 and X_n polynomial function.

$BEZ_{m,n}(b)$ Is Bernstein Polynomial [20].

$$X_b = \sum_{m=0}^{n} X_m BEZ_{m,n}(b) \tag{5}$$

$$BEZ_{m,n}(b) = (n/m) b^m (1-b)^{n-m} \tag{6}$$

$$BEZ_{m,n}(b) = (1-b) \cdot BEZ_{m,n-1}(b) + b \cdot BEZ_{m,n-1}(b) + b \cdot BEZ_{m-1,n-1}(b) \tag{7}$$

Whereas

$$BEZ_{m,m}(b) = b^m \text{ and } BEZ_{0,m}(b) = (1-b)^m$$

The points of Bezier curve are obtained by the following equations.

$$p(b) = \sum_{m=0}^{n} p_m BEZ_{m,n}(b) \tag{8}$$

$$q(b) = \sum_{m=0}^{n} q_m BEZ_{m,n}(b) \tag{9}$$

Now these straight lines are utilized to compute the Hausdroff distance [21].

$$D_H(x, y) = \max\{\min_{a \in |s,t|} \max_{b \in |u,v|} |x(a) - y(b)|\} \tag{10}$$

$$\max_{b \in |u,v|} \min_{a \in |s,t|} |x(a) - y(b)| \tag{11}$$

where, $D_H(x, y)$ between two curves $x(a)$, $a \in |s, t|$ and $y(b)$, $b \in |u, v|$.

This distance measures the line distance of spline points and offers an improved recognition rate. A reference database is made which includes the spline point for all types of emotions in normal conditions. For any captured image frames, the spline points Hausdroff distance is found and compared with the reference database. The points which are closer to the Hausdroff distance are identified and recognized to be a best match to the test frame.

4. EXPERIMENTAL RESULTS

This section presents the functionality of intelligent tutoring system through an example of ongoing tutoring session. SeisTutor was tested over 20 applicants, over a period of 3 months. Learner applicants that logged on underwent Initial assessment phase. The learner is presented with a brief test administered under two categories, one "domain knowledge test" which adjudges the learner's profile as one of the three different levels or profiles "beginner", "proficient" and "expert", and the other test, "Non-Domain Knowledge test" encompassing specific questions to adjudge the learning style of the learner. This is a key aspect of the proposed system as the tutoring system offers to deliver the learning material to learners as per their individual learning preferences. Thereby the aim is to provide effective tutoring ensuring maximum learning gain. A tutoring strategy specifically tailored for a learner is executed, to enhance his or her degree of engagement.

The observation after Initial assessment phase is shown in Table 2. It was observed that nine applicants were allotted 'beginner profile', seven of them the 'proficient profile' and rest four of them the 'Expert profile'.

Table 2
Initial Assessment Phase Observation

S.No.	*Number of Applicants*	*Learner Profile*	*Learning Style*		
			Read & Write	*Verbal*	*Visual*
1.	9	Beginner	5	–	4
2.	7	Proficient	3	1	4
3.	4	Expert	–	1	3

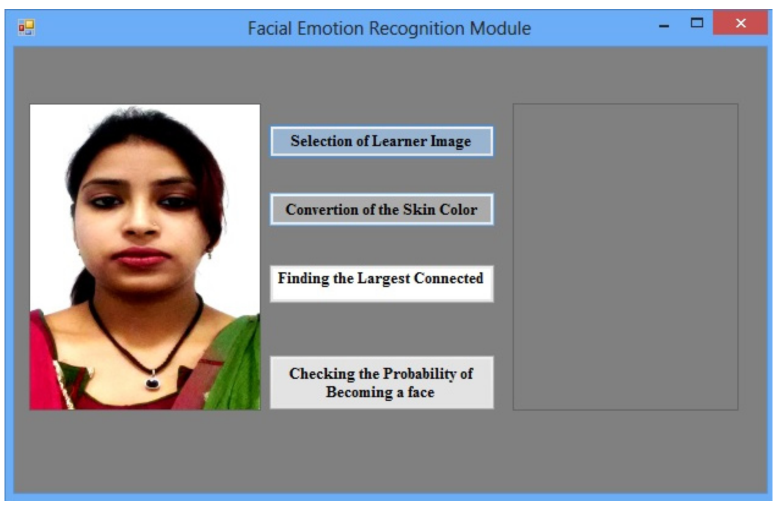

Figure 4: Selection of Learner Image

The five, beginner level learners were found to be adjudged as learning style 'read & write' oriented and four as 'visual oriented' none of them were found to be 'verbal' style oriented. Three proficient level learners were adjudged as a learning style 'read and write', one as 'verbal oriented' and four as 'visual oriented'. Out of four expert level learners one was adjudged as learning style 'verbal oriented' and three as 'visual oriented'.

After the learner profile and style identification, all the applicants underwent tutoring. Figure 11 presents the initial dashboard offered to the learner when the tutoring commences, it indicates attributes of the course/ subject matter domain being delivered such as "*Who should attend*", "*Course Overview*" and "*Course plan*" etc. Figure 12 depicts the ongoing tutoring session as per the identified tutoring strategy (a combination of learner level and style).

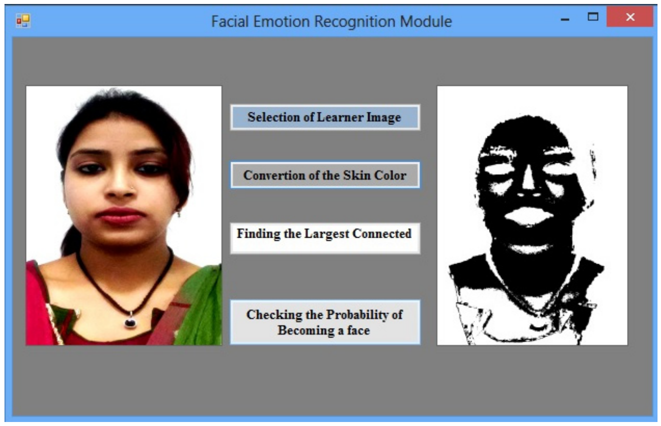

Figure 5: Conversion of the Skin Color

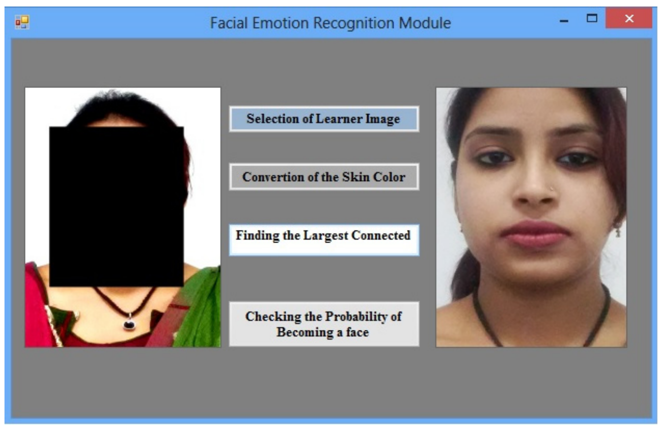

Figure 6: Finding the largest Connected Region

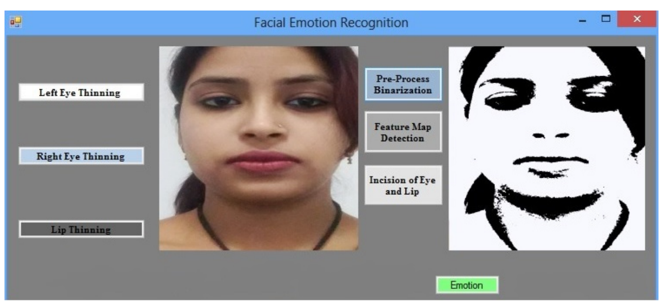

Figure 7: Pre-Processing Binarization

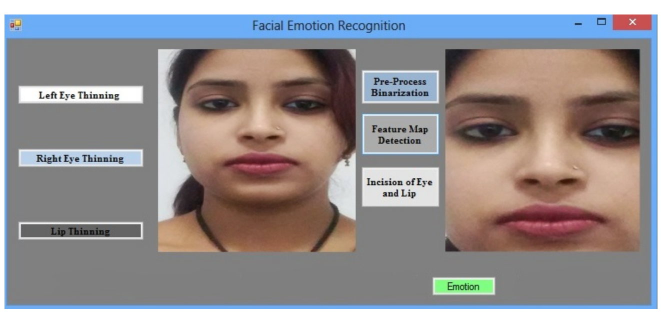

Figure 8: Feature Map Detection

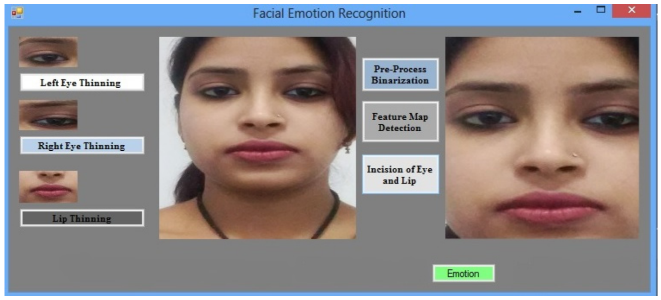

Figure 9: Incision of Eye and Lip

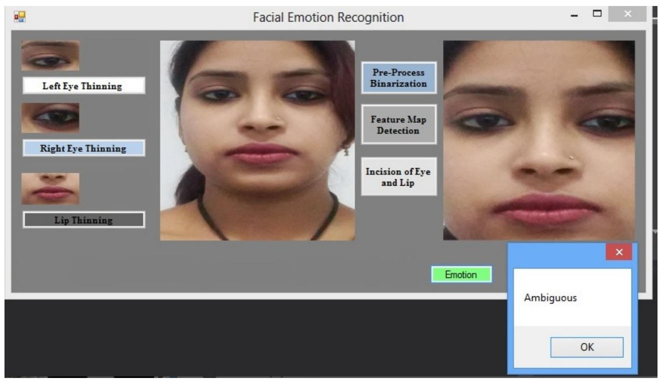

Figure 10: Emotion Detection.

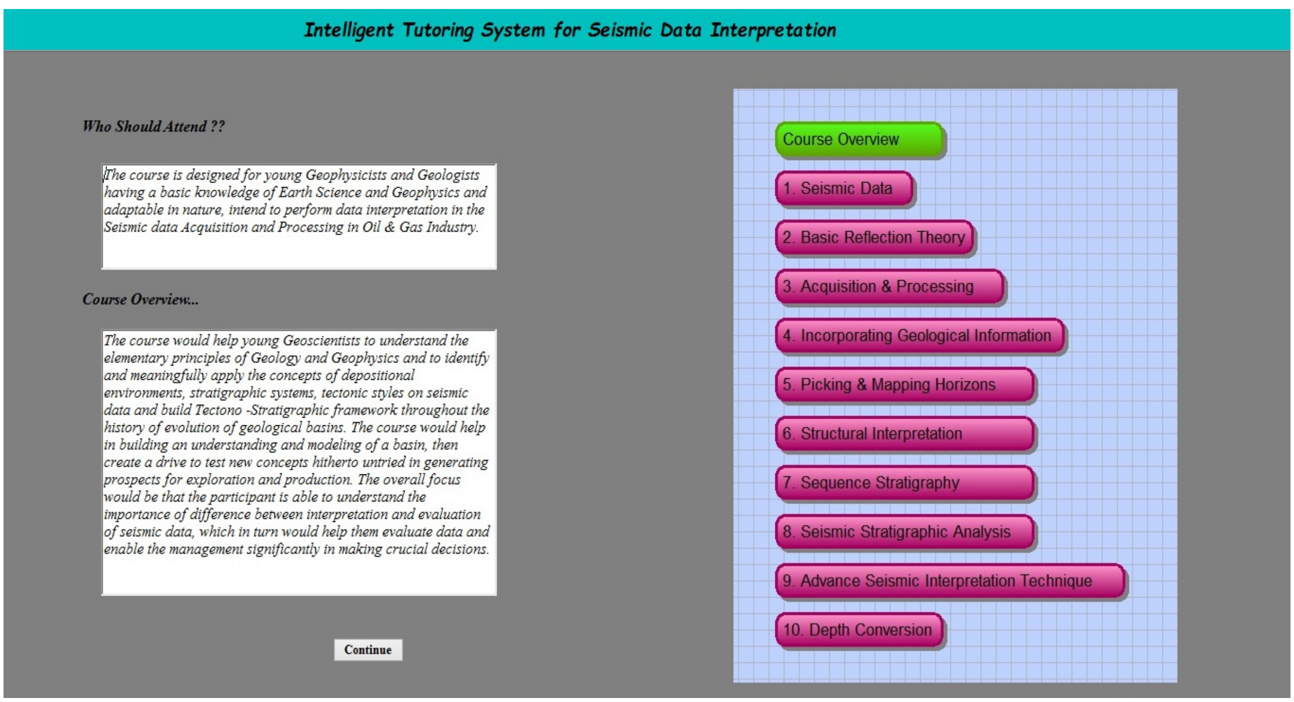

Figure. 11: Course Overview

As tutoring progresses, learner activities and performance parameters are recorded at pre-decided checkpoints. The system intelligently determines and classifies the emotion of learner. The whole process of emotion recognition is shown from Figure 4 to Figure 10.

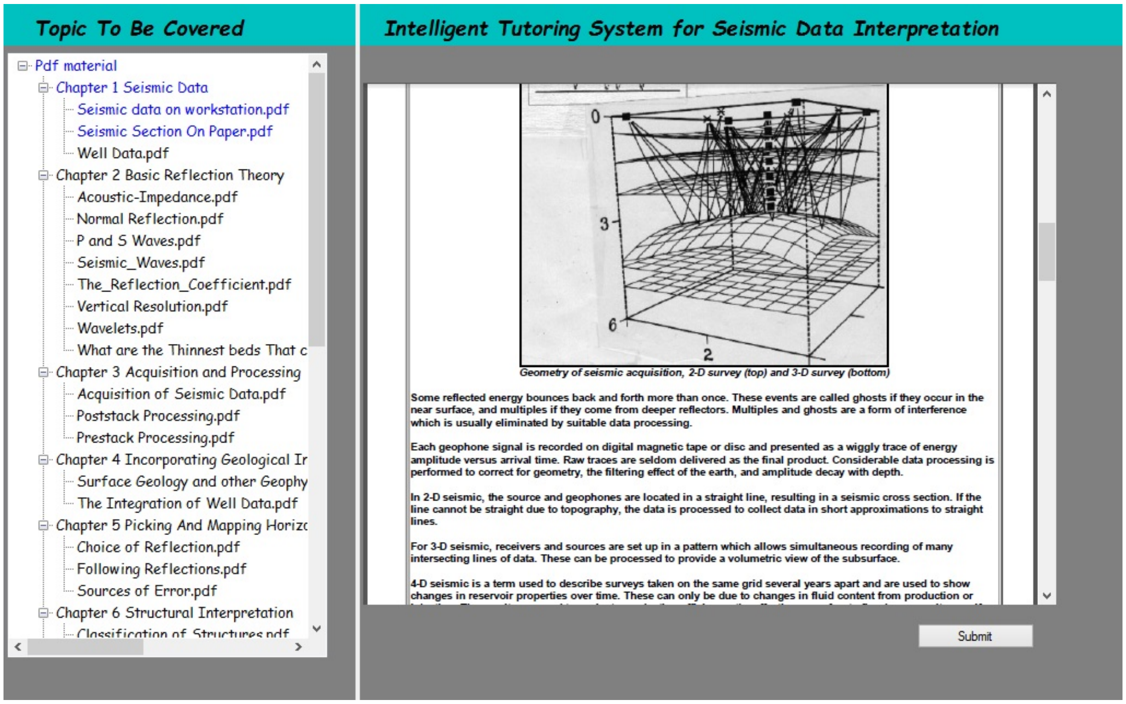

Figure 12: Tutoring session in progress

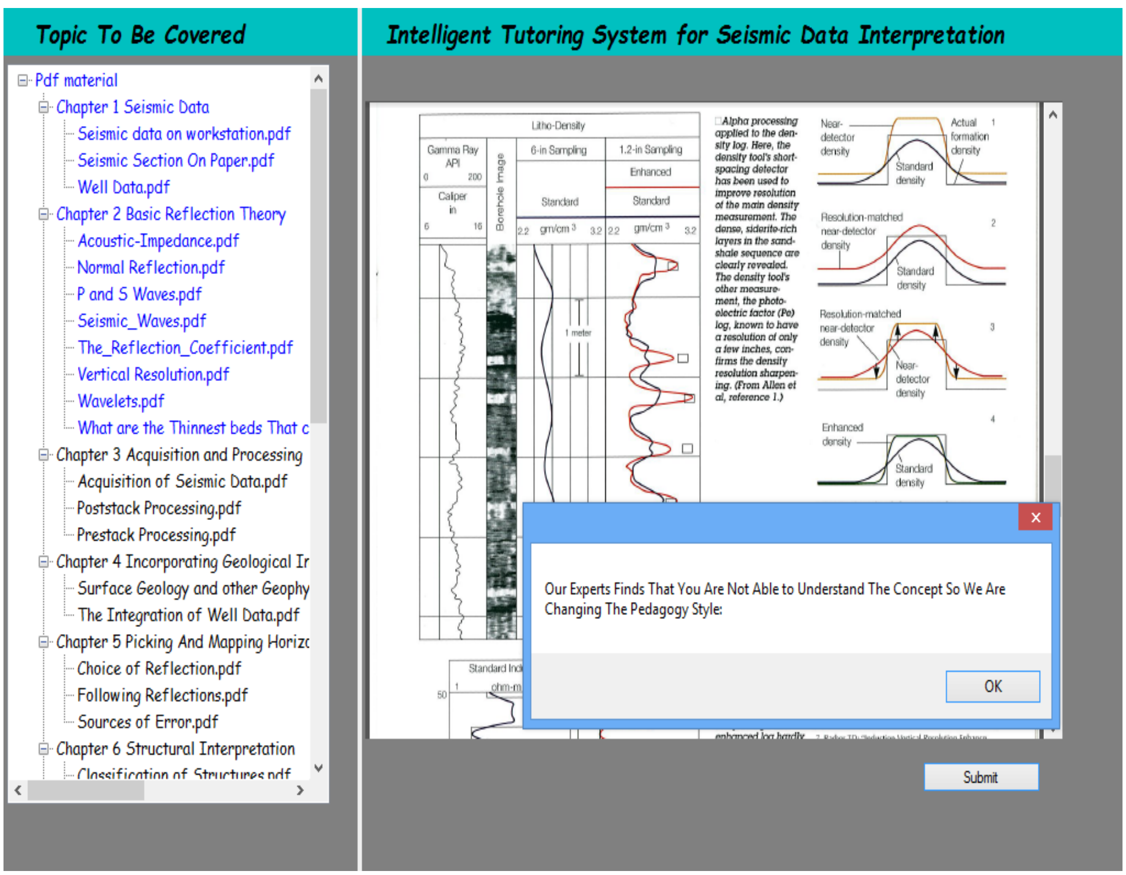

Figure 13: SeisTutor Detects Learner is not comfortable with current tutoring strategy

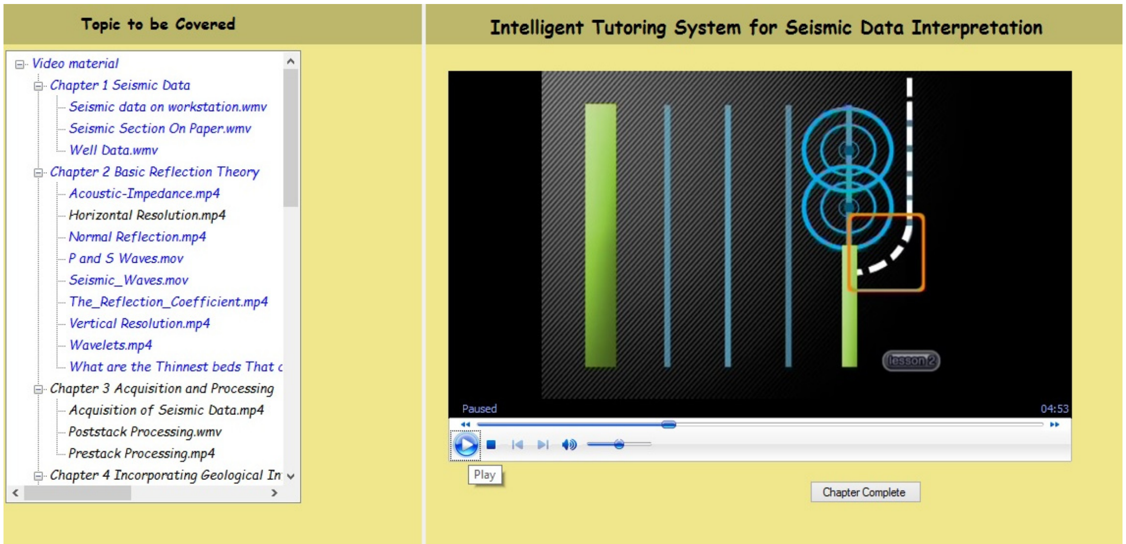

Figure 14: Pedagogy Change

For the current scope of work, five emotions have been considered i.e. sad, smile, surprise, ambiguous and normal. Reference database contains the spline coordinates of emotions. After the boundary identification of facial area of interest spline points Hausdroff distance is computed and compared with the reference database. The points which are closer to the Hausdroff distance are identified and recognized to be a best match to the image frame. By this way learner emotion is determined.

If it is identified that the learner's expressions indicate the emotions that the learner is unable to follow (Sad, Ambiguous, Surprise) the taught concept/ uncomfortable/ not sufficiently engaged, then the following message is flashed. "*Our Experts Finds That You Are Not Able to Understand the Concept So We Are Changing the Pedagogy Style:"* This causes to change the pedagogy style automatically, best suitable to the learner as shown in Figure 13. The second learning style of the priority list is now chosen as a active pedagogy style. Thus tutoring strategy also gets changed, as shown is Figure 14.

The pedagogy style is changed automatically because when a learner is in the Initial assessment phase, based on his/her response, SeisTutor makes a priority list of the learning style by using priority queue data structure. Initially, it starts tutoring by its top priority learning style. In worst case if the system detects that learner is not satisfied with the tutoring session then SeisTutor changes the pedagogy style by applying dequeuing operation in learning style priority queue and extract second top learning style.

The webcam gets initialized when tutoring session starts. After every 30 seconds SeisTutor captures the learner snap and applies tasks defined in Identification phase (as mentioned in Figure 3) and applies rules presented in Table 1. For the identification of facial expression, Table 3 presents result recorded during the tutoring session for 20 applicants. 11 applicants were comfortable with the pedagogy style that was initially intelligently identified by SeisTutor whereas 9 of them were not.

At checkpoints the tutor can trigger change of the pedagogy style. For applicants not comfortable with the initial pedagogy style. Emotion recognition and performance parameters recorded at checkpoints, form basis for decision of pedagogy change.

At checkpoints the tutor can trigger change of the pedagogy style. For applicants not comfortable with the initial pedagogy style emotion recognition and performance parameters recorded at checkpoints, from basis for decision of pedagogy change.

5. CONCLUSIONS

In a face to face class room environment, a human tutor is vigilant and observant towards the learner's behavior and adjusts the teaching style to best suit the learners learning style. Built on the same analogy, the proposed system SeisTutor, captures the learner facial expressions, to interpret emotions and changes the tutoring strategy to present pedagogy style suitable for the learner. This mechanism aims at making learning effective, ensuring enhanced learning gain, higher degree of engagement and proposes to fill the disparity between sentimental tutor and a sentimentally challenged computer. There have been efforts in past on integrating emotion detection in intelligent tutoring systems, but the novelty of the current system is to trigger and cause pedagogy changes as per the recommendations of emotion detection module, additionally supported by performance parameters, which further can lead to computation of net degree of engagement serving as an indicator of learning efficiency.

Acknowledgment

This work is being carried out at University of Petroleum and Energy Studies with the reference number SR/CSI/140/2013. The authors thankfully acknowledge the funding support received from Cognitive Science Research Initiative, Department of Science and Technology for the project. The Authors thank the management of University of Petroleum and Energy Studies, for Supporting the work and granting permission to publish it.

REFERENCES

[1] H. S. Nwana, (1990). "Intelligent tutoring systems: An overview". *Artificial Intelligence Review* **4**: 251–277.

[2] C. Qi-rong, "Research on Intelligent Tutoring System Based on Affective Model ", in *2010 2nd International Conference on Multimedia and Information Technology*, 2010, pp.7-9.

[3] M. Neji, "The integration of an emotional system in the Intelligent Tutoring System", in *3rd International Conference on Computer System and Applications*,2005, No. Figure 1.

[4] Y. Wu, W. Liu and J. Wang, "Application of Emotional Recognition in Intelligent Tutoring System", in *1st International Workshop on Knowledge Discovery and Data Mining (WKDD 2008)*, 2008, pp. 449-452.

[5] L. Wolcott, "The Distance Teacher as Reflective Practitioner", Educational Technology 1, 1995, pp. 39-43.

[6] N. J. Ahuja, P. Diwan, "An Expert System for Seismic Data Interpretation using Visual and Analytical tools ", *International Journal of Scientific & Engineering Research,* 2012, Vol. 3(4), pp. 1-13.

[7] G. R. S. Murthy and R.S. Jadon, "Recognizing Facial Expressions Using Eigen spaces*", International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007),* 2007, pp. 201-207.

[8] G.R.S. Murthy and R.S. Jadon, "Effectiveness of Eigen spaces for Facial Expressions Recognition", *International Journal of Computer Theory and Engineering,* 2009, Vol. 1(5), pp. 638-642.

[9] B. C. Becker, E. G. Ortiz and C. Flori, "Evaluation of Face R Recognition Techniques for Application to Facebook 5000 Forbes Avve University of Centr" in *8th IEEE International Conference on Automatic Face and Gesture Recognition*, 2008,pp.1-6.

[10] C. Lien, Y. Chang and C. Tien, "A Fast Facial Expression Recognition Method at Low-Resolution Images*", International Conference on Intelligent Information Hiding and Multimedia*, 2006, pp. 419-422.

[11] D.S.P. Arumugam, "Emotion Classification Using Facial Expression", *International Journal of Advance Computer Science and Applications*, Vol-2(7), 2011.

[12] H. Lu, K. N. K. Plataniotis and A.N. Venetsanopoulos, "MPCA: Mulitilinear Proncipal Component Analysis of Tensor Objects*", IEEE Transactions on Neural Networks/ a publication of the IEEE Neural Networks Council*, Vol-19(1), 2008, pp. 18-39.

[13] D. Litman and K. Forbes, "Recognizing Emotions from student speech in Tutoring Dialogues", *proceedings of the ASRU*, 2003.

[14] S. D'Mello, T. Jackson, S. Craig, B. Morgan, P. Chipman, H. White, N. Person, B. Kort, R. el. Kaliouby, R.W. Picard and A. Graesser, "AutoTutor Detects and Responds to Learners Affective and Cognitive States In", Workshop on Emotional and Cognitive Issues at the *International Conference of Intelligent Tutoring Systems*, 2008.

[15] Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). "Affect-aware tutors: Recognising and responding to student affect", *International Journal of Learning Technology, 4*(3/4), 129.

[16] Mao, X., & Li, Z. (2009). "Implementing emotion-based user-aware e-learning", In *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems - CHI EA '09* (pp. 3787-3792). New York, NY: ACM Press.

[17] Mao, X., & Li, Z. (2010). "Agent based affective tutoring systems: A Pilot study", *Computers & Education, 55*(1), 202–208.

[18] Tian, F., Gao, P., Li, L., Zhang, W., Liang, H., Qian, Y., & Zhao, R. (2014). "Recognizing and regulating e-learners' emotions based on interactive Chinese texts in e-learning systems", *Knowledge-Based Systems, 55*, 148–164.

[19] P. Ekman and W. Friesen. Facial Action Coding System (FACS): Investigator's Guide. *Conculting Psychologists Press,* 1978.

[20] M. R. Rahman, M. A. Ali, and G. Sorwar. Finding Significant Points for Parametric Curve Generation Techniques. *Journal of Advanced Computations*, 2(2), 2008.

[21] S. Pal and P. K. Biswas, "Modifiied Hausdroff Distance Transform Technique for Video Tracking," *International Conference on Vision, Graphics and Image Processing (ICVGIP-2000),* 2000.

ANNEXURE 1

Learning style: Learner's learning preference, each learner learns in his or her own distinct learning style in which his/her learning performance is improved to a reasonably degree. For the scope of this work, following three styles have been considered read-write or text, visual and verbal.

Learner profile: Depth of knowledge of a particular subject area example: Seismic data interpretation.

Accordingly, three learner levels namely 'Beginner', 'Proficient' and 'Expert' have been used in this work. 'Expert' is assumed to be holding highest degree of proficiency in a given subject matter domain.

Tutoring Strategy: It is a strategy according to which learning is to be offered to the learner. It is represented as a combination of identified learner level and learning style.

Pedagogy Style: It is a specific style in which learning is going to be delivered. It is represented as a specific sequence or pattern of aligning the subject matter and its forms while being delivered to target audience. A given tutoring strategy is manifested through one or more pedagogy styles.

Checkpoints: Checkpoints can be defined as markers depicted as milestones spaced within pre-decided time-intervals indicating points up to which tutor has recorded activities and emotions of the learner during ongoing tutoring session. This is the point where tutor is allowed to change the pedagogy style of the learner. Based on the analysis of recorded parameters the tutor may trigger style or other applicable changes.

Priority List: As a result of pre-tutoring, a list of appropriate tutoring strategies are identified for a given learner. These are ranked in the order of their appropriateness for the learner and made available in the form of priority list.

Degree of Engagement: This term is used to determine the extent of involvement of learner in tutoring session. It can also be defined as the how effectively learner is engrossed in learning and enjoys or understands the concept. This is computed based on the emotion (behavior), scores (test score after every chapter) and pedagogy change (At checkpoint if there is any change).

Promptness: How much time learner takes to respond to a given question.

Read & Write Learning Style: Read & write learning style includes information in the form of text.

Verbal Learning Style: Verbal learning style includes information in the form of sounds, spoken words, and audio clips.

Visual Learning Style: Visual learning style includes information in the form of pictures, charts, graphs etc.