A Computational System Approach To Develop Students’ Emotion Oriented System as Feedback Tool for Lecturers to Enhance Teaching and Learning

Deshinta Arrova Dewi, Zaidatol Haslinda Abdullah Sani, Premylla Jeremiah

1INTI International University, deshinta.ad@newinti.edu.my
2INTI International University, linda.sani@newinti.edu.my
3INTI International University, premylla.jeremiah@newinti.edu.my

Abstract

The quality of learning depends on aspects that integrate positive environment in the classroom. Previous works have shown that positive environment do create positive emotion and positive emotion do affect positive learning behavior. When students experience positive emotion, the learning process can be enhanced and when students experience negative emotion, learning process can be disabled. Therefore managing students learning in the classroom means managing emotions. This paper presents the efforts in recognizing student’s emotion to promote interactive teaching learning atmosphere in a Higher Education environment that is served as an alternative feedback tool for lecturer to enhance students learning experiences.

This system suggests lecturers to identify whether their teaching style has positive impact on students learning behavior by recognizing student’s emotion in the classroom. This system proposes the composition of positive class emotion as result of positive student’s emotion to encourage lecturers to retain their teaching style or to improve it.

In this paper we explore the possibilities of capturing emotions by various techniques. We compare some of the existing works and identify how emotions can be modeled in order to enhance its services and response and able to suggest meaningful information to users. We also define and discuss the appropriate capacity of meaningful information to generate analysis and recommendations according to theory of teaching and learning. We do our experiments in INTI International University, Nilai Malaysia observing approximately 200 students from various faculties.

Keywords: Emotion Oriented System, Emotion Based System, Image Processing, Computational System Approach, Interactive Teaching Learning, Feedback Tool for Lecturer

1. Introduction

Learning is defined as the process of acquiring knowledge, skills and attitude through study that can change behavior that is persistent and measurable. From the perspective of teaching and learning, we investigated parameters that affect learning to improve the way we deliver our course. We found that not only the environments, teaching methodologies and level of assessments can affect learning, but also the student’s emotion actually play an important role in learning behaviors.[1][2]

A research conducted by Sylvester[1] shows that emotion does affect learning. When students experience positive emotion, the learning process can be enhanced and when students experience negative emotion, learning process can be disabled. Dr. Lawson[2] mentioned in her research that emotion arise from memories and reactions to current events. The emotions are formed by how we think about past and present experiences. Hence, managing students learning means managing emotions.

Robert and Smith[3] from North Carolina University stressed the idea that a classroom is a site of emotion management, especially when the topics covered may expose to student’s and lecturer’s sensitivity. By having a better understanding how emotional responses are socially captured and mediated, lecturers can better facilitate the class and maximize learning.

A joint research published by four authors from US and China[4] develop a theory of emotion as a feedback system whose influence on behavior. By providing feedback and stimulating retrospective appraisal of actions, conscious emotional states can promote learning and alter guidelines for future
behavior. Another research conducted by Wright[5] said that emotions will affect behavior in both direct and indirect ways. In the same context, Dr. Farr[6] states that in order to carry out correct behavior, human develops innate drives, desire, emotions, ability to remember and learn.

The current development of Computational System has accommodated the need of understanding and exploration on how people interact with technology [7],[8],[9],[10]. Some of social practices that are fundamental part of the ways people work, live and play have been studied and channeled with the help of computing technologies.

In year 2009, the International Journal of Computer Science had published the existing research related to the consideration of emotion in system development by using persuasive computing approach. It is known as Emotion Oriented System.[10] The human emotions were captured through various properties of the human being such as voice detection, physical movement and body gesture to be interpreted in the form of XML language.

Our research attempts to use various computational system approaches to develop an Emotion Oriented System as a feedback tool for lecturers regarding their own classes. This could enhance the overall teaching learning processes in Higher Learning Institutions since not many tools are provided to help lecturers assess their own teaching at present.

This motivation comes from our previous findings that mostly lecturers know their class feedback by asking students in group or individually regarding the particular teaching session on that day. Any conclusion derived from there that relates with the continuity of lecturer’s teaching style, will be having a lack of validity and accuracy since in normal circumstances some students will respond honestly and some of them will not.

An immediate student’s feedback is very important because it reflects the impact of a lecturer’s teaching style in developing student’s emotion in learning. Hence, this research is aimed to answer the following question:

Could student’s learning be more fun and generate more positive emotion if lecturers use a variety of teaching styles according to feedback that emphasize on student’s emotion in the class?

2. Emotion Oriented System According to HUMAINE

Creating competent emotion-oriented systems is a large scale challenge. The European Network of Excellence HUMAINE (HUman-MAchine Interaction Network on Emotions) was established to prepare the scientific and technological ground for this task, with funding from the EU IST programme from 2004 to 2007. [11][12][13][14]

One fundamental of natural Human-Machine Interaction (HMI) is the ability to detect the signs of emotion emitted by the user, intentionally or unintentionally. This task is, on the one hand, heavily dependent on the emotion models used, e.g. whether the emotion is described as a category or as a region in a multidimensional space. The task is also highly dependent on the material from which to recognize emotion: for example, classifiers that work very well with acted emotional material may fail on naturalistic material [14].

Creating “emotion-oriented systems” is a key axis of future research in HCI, as we rely on emotion all the time. A user-centered system usually goes through an emotion oriented system that normally includes capturing emotion, modeling and rendering of emotions in order to understand and act according to human emotional processes.

The HUMAINE network was created in January 2004 and many articles and deliverables have been published. In order to create a common vocabulary, precise enough to be used in computer sciences in an efficient way, they published [12][13][14] a definition of emotion.

Scherer's definition of emotion as stated at [23] is the following: "Emotions are –"episodes of massive, synchronized recruitment of mental and somatic resources allowing adapting to or coping with a stimulus event subjectively appraised as being highly pertinent to the needs, goals, and values of the individuals".

The following table (Table 1) is part of concept describe in [12][13][14][23]:

542
As we can see, the HUMAINE network of excellence use generic concept of “affective states” which can be subdivided into five categories: emotions, moods, interpersonal stances, preferences/attitudes, and affect dispositions. According to the table, an emotion is characterized by:

- A high level of intensity
- A focus on the event: an emotion is caused by an environmental stimulus.
- Appraisal elicitation: a cognitive process processes the triggering event.
- A high level of synchronization: the whole body tends to react according to the emotion: facial signals, posture, physiological reactions
- A great impact on behavior
- A small duration and a high rapidity of change: emotions are a quick body response, that won’t last long and can change quickly.

### 3. Various Techniques to Capture Emotions

Picard [15] suggested several applications where it is beneficial for computers to recognize human emotions. For example, knowing the user’s emotions, the computer can become a more effective tutor.

Addressing the problem of affective communication, Bianchi-Berthouze and Lisetti as cited at [15][16][17][18] identified three key points to be considered when developing systems that capture affective information: embodiment (experiencing physical reality), dynamics (mapping experience and emotional state with its label), and adaptive interaction (conveying emotive response, responding to a recognized emotional state).

The most current researches in 2012 regarding emotion recognition have been conducted by many Research Institutes in Korea [27][28]. The researchers used 10 different emotional stimuli set to induce seven emotions, i.e., happiness, sadness, anger, fear, disgust, surprise and stress under the same conditions[27]. They identified the difference among emotions using physiological responses induced by emotional stimuli and the most optimal algorithm for emotion recognition. For this, they selected physiological signals of electrodermal activity (EDA), electrocardiogram (ECG), photoplethysmography (PPG), and skin temperature (SKT), because the signals reflect the activity of the autonomic nervous system, which plays a major role in maintaining the internal equilibrium of the body [27].

Despite these many theories, it is evident that people display these expressions to various degrees. One frequently studied task is the judgment of emotions—how well can human observers tell the emotional expressions of others, in the voice, on the face, etc? [15][16] Related questions are: Do these represent their true emotions? Can they be convincingly portrayed? How well can people conceal their emotions? In such tasks, researchers often use two different methods to describe the emotions. One approach is to label the emotions in discrete categories, i.e., human judges must choose from a prescribed list of word labels, such as joy, fear, love, surprise, sadness, etc. One problem with this approach is that the stimuli may contain blended emotions. Also, the choice of words may be too restrictive, or culturally dependent.

Another way is to have multiple dimensions or scales to describe emotions. Instead of choosing discrete labels, observers can indicate their impression of each stimulus on several continuous scales, for example, pleasant–unpleasant, attention– rejection, simple–complicated, etc. Two common scales
are valence and arousal. [22] Valence describes the pleasantness of the stimuli, with positive (or pleasant) on one end, and negative (or unpleasant) on the other. For example, happiness has a positive valence, while disgust has a negative valence. The other dimension is arousal or activation. [22]

For example, sadness has low arousal, whereas surprise has high arousal level. The different emotional labels could be plotted at various positions on a two-dimensional plane spanned by these two axes to construct a 2D emotion model. Scholsberg as cited at [17][18] suggested a three-dimensional model in which he had attention–rejection in addition to the above two.

A Bayesian approach was used to find the best match between the local observations and the learned local features model. Furthermore, a Hidden Markov Models (HMM) was employed to achieve better recognition even when the new conditions did not correspond to the conditions previously encountered during the learning phase. [17][18]

Oliver as cited in [17] used lower face tracking to extract mouth shape features and used them as inputs to an HMM based facial expression recognition system (recognizing neutral, happy, sad, and an open mouth). Chen as cited in [18] used a suite of static classifiers to recognize facial expressions, reporting on both person-dependent and person-independent results. Cohen [20] describes classification schemes for facial expression recognition in two types of settings: dynamic and static classification.

The static classifiers classify a frame in a video to one of the facial expression categories based on the tracking results of that frame. In this setting, the authors learned the structure of Bayesian networks classifiers by using as input 12 motion units given by a face tracking system. [16][18][19][20]

HUMAINE has prioritized three modalities for study – facial expression in video; speech in audio; and physiological parameters. In each of these, its first priority is to establish the reliability of alternate signal analysis algorithms for extracting basic features. Building on that, it aims to clarify the principles of effective cross-modal integration. One challenge that is immediately apparent is to deal with the different temporal structures that characterize the modalities.

In this stage, data gathered from multiple sensors need to be processed, then classified based on dynamic or discriminative models and finally combined with other sensor values and results can be matched against data sets from a database to identify matching emotional patterns. Users in a pervasive environment normally express multimodal (e.g., audio and visual) communicative signals in a complementary or redundant manner.

Besides the suggestion by HUMAINE, another way of capturing emotion of a user A is by monitoring the behavior, action of another user B interacting with A since B may possibly largely influence A’s emotional state. For example, a student generates happiness when the lecturer praises his/her work. Tracking of body movements (head, arms, torso, legs and others) to determine emotions is still a challenging area. [21]

4. System Development Methodology

Our research methodology is following the System Development Life Cycle Methodology (SDLC) that is commonly used in a software development activity. We chose SDLC because it is the most complete methodology starts from the sequence of system requirements and followed by the rest of the life cycle phases i.e., system analysis, system design, system implementation, system testing, system operational and system maintenance.

In the stage of system requirements, our software has emphasized on capturing non-linear data. The emotions and behaviors are captured by using multimedia tools such as video and camera to maintain the originality of session on that day. Later it will be analyzed by using matrix computation approach that involves image processing and pattern recognition to identify the types of emotion and behavior during learning session on that day.

Our research suggests the appropriate time frame to generate this analysis to meet the real-time requirements as a genuine feedback to improve teaching and learning processes. The proposed Emotion Oriented System is based on matrix computation analysis to cater the complexity of pervasive emotion such as feeling, expression and behavior. In its finality, the captured emotion would most likely be grouped into families (like anger, hate, rage, etc).

The databases are integrated as part of the system development. The proposed suggestions that relate to future teaching learning will be generated. As a result, the teaching learning recommendation for a lecturer is derived from the emotion analyses that were calculated on real-time basis during a particular
time. The databases also consider the major factors involved in determining emotion and the solution of effective teaching and learning such as psychology perspective, theory of teaching and learning, pervasive computing and mathematics. This integration aims to develop a solid system library as references of overall system performance.

This suggestion then is assessed from the perspective of metrics based testing to identify whether the system accuracy has been satisfied. A minimum of two sessions are required in obtaining this outcome. The first session is for capturing the emotion and provides suggestion whereas the second session is on improving the teaching methods based on suggestions given previously. The accuracy is based on this type of testing.

Our research also suggests an appropriate number of iteration to determine the success rate of a teaching learning session. A graphical conclusion of the overall teaching-learning session is also provided as self improvement for the respective lecturer.

In the stage of system testing, we impose on the system integration testing (data input, data process, database and data outcome) that should be ready for user to run. This type of testing assesses the user friendliness viewpoint on the new system’s functional operations.

The system operation and maintenance focuses on the system functionality during its operation. Some minor debugs may need to carry out to prevent major mistakes that may affect the overall system performance.

5. System Model and Implementation

Below is the model of our system prototype that consists of three major components.

![Figure 1. System Prototype](image)

The first part is the basic appearance description extraction. For each video clip, two different basic appearance descriptors were extracted. First, we compute Local Binary Patterns (LBPs) [19][22] to capture the local textural patterns from each frame of the video. The reason to apply the LBP algorithm to all the pixels in a frame rather than just to the face region is to maintain information of other visual cues (e.g., head movement and shoulder) present in the video.

The second part is data analysis to determine emotion. Our research finding shows several ways in capturing student emotion that significantly affects teaching style. One of the possible ways is by creating a Mediating Agent as suggested by [7][14][22]. The Mediating Agent must recognize a student’s emotional states to respond appropriately.

For instance, when the student is disappointed with his performance, he will probably give up the task. The agent needs to know when the student is disappointed to encourage him to keep on studying and accomplishing the task. For this reason, the Mediating Agent has a sensor component responsible for identifying student’s emotions and an affective model for storing this information.[7][14][22]

When we received several samples from the incomplete existing implementations, we found out that our research do not require the agent to encourage students to keep studying. Our research would send the agent to the database library and retrieve actions that are associated with the teaching style. For this reason, our Mediating Agent will carry out tasks to analyze the percentage of class emotion (positive and / or negative) and finding the most appropriate suggested teaching style. This type of work has been proposed by several papers [22][23][24][25].
From the perspective of the simulation model, generally a model of emotion should be able to identify events and objects involved in the interaction process. Emotion models should demonstrate the computer’s ability to identify emotions at the right intensity and appropriate time. Appraisal theories have influenced the development of computational models of our system. Essentially, our appraisal of a situation causes an emotional or affective response.

Based on the appraisal theory, the model that has been regularly used to incorporate emotion into various systems is the model that addresses the problem of representing emotions not by a set of basic emotions but by grouping emotions according to cognitive eliciting conditions. It believes that emotions arise from valenced (positive or negative) reactions to situations involving events, agents and objects [22].

Another computational model of emotions based upon the appraisal theory is the EMA (EMotion and Adaptation) [25]. The agent’s interpretation of its “agent-environment relationship” is considered to be a representation of beliefs, desires, intentions, plans and probabilities, which is referred to as causal interpretation to emphasize the importance of causal reasoning as well as the interpretative (subjective) character of the appraisal process [25].

A control-oriented method for simulation of students’ emotion recognition environment using MATLAB SIMULINK is suggested [26]. SIMULINK allows for the visual understanding how a dynamic system (a real time based) operates. Detailed dynamic models are required for performing ephemeral stability studies. Capturing students’ emotion in a classroom during teaching and learning process conducted presents a dynamic system. SIMULINK applications can be applied to represent the application.

The figures below are the prototypes of the simulation model in capturing multimedia data (pictures).

![Figure 3: Expected readings for a Grumpy mood](image)

The prototype is expected to capture the visualization of the facial emotion in the image above. The emotions captured will then display the probability percentage of each varying mood detected by the system. As an example, based on the Figure 3 above, the prototype should show a highly possible accurate emotion and the corresponding probability for the mood such as Neutral (10%), Angry (25%), Sad (5%), Grumpy (30%), Surprise (5%), Fear (5%), and Disgust (20%) while Happy and Nervous remains as 0%.

The accuracy of the prototype would very much depend on the position of the camera. The camera ought to be directly facing the students’ face in order for the accurate emotions to be detected and recognized. The prototype is also expected to provide accurate emotion detections for both Figures 4 and 5 as well.
In Figure 4 shown above, the emotions that should be detected would display a very high reading percentage of Happy (70%), Surprise (25%) and Neutral (5%) while the other moods remain at 0%.

Alternatively, for Figure 5 shown above, the prototype should provide a reading of Neutral, Angry, Happy, Surprise and Sad (all at 5%), while Disgust and Grumpy is (10%), Fear (15%), and Nervous (40%). As such, based on the readings that the prototype should be able to detect and provide the system’s ability to capture emotions ought to be satisfactory and successful.

One challenge that is hardly supported by the prototype is when the multimedia data (pictures) shows a situation where the students’ positions are not directly facing the camera (Figure 6). Our hypothesis suggests the face recognition is only being able to recognize 20%-40% of the emotion in the classroom. It was predicted the accuracy level of emotion detection most likely decreasing to up to 50%.

Alternatively, for Figure 5 shown above, the prototype should provide a reading of Neutral, Angry, Happy, Surprise and Sad (all at 5%), while Disgust and Grumpy is (10%), Fear (15%), and Nervous (40%). As such, based on the readings that the prototype should be able to detect and provide the system’s ability to capture emotions ought to be satisfactory and successful.

One challenge that is hardly supported by the prototype is when the multimedia data (pictures) shows a situation where the students’ positions are not directly facing the camera (Figure 6). Our hypothesis suggests the face recognition is only being able to recognize 20%-40% of the emotion in the classroom. It was predicted the accuracy level of emotion detection most likely decreasing to up to 50%.

Figure 6: Multimedia data that potentially generates low percentage of emotion detection.
To accommodate the challenges as above, the most recent study suggests multimodal approach to develop emotion recognition. A natural two-way interaction between the human and the computer through multiple modalities is depicted in Figure 7 suggested by [22].

![Figure 7: The multiple modalities to capture emotion](image)

In this diagram, one of the inputs to the computer is vision (video), from which gaze, posture, gestures, and facial and lip movements can be extracted. Computers may learn to recognize gestures, postures, facial expressions, eye contact, etc. Likewise, speech and voice (audio) through the microphone may convey linguistic as well as paralinguistic information.

The last part of our system is generating recommendations for lecturers regarding their teaching style in conducting class. We proposed to use knowledge based approach to generate those recommendations due to the nature of information derived from the emotion captured in the database.

The table below shows the way our system produce student’s emotion feedback for the lecturers to enhance their teaching style in order to maintain positive emotion in the classrooms.

<table>
<thead>
<tr>
<th>No</th>
<th>Suggestion</th>
<th>Emotion’s Captured (60 minutes)</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-</td>
<td>First quarter: 20% positive, 20% negative, 60% neutral</td>
<td>Not Successful, change teaching style.</td>
</tr>
</tbody>
</table>
|    |            | Second quarter: 30% positive, 40% negative, 30% neutral | Negative emotions are:  
  ➢ Boring (10%)  
  ➢ Stress (15%)  
  ➢ Sleepy (15%)  
  Suggestions:  
  ➢ Move around classes  
  ➢ More humor  
  ➢ Provide tutorial |
|    |            | Third quarter: 30% positive, 40% negative, 30% neutral |
|    |            | Fourth quarter |
Further research is required to achieve maximum utilization of multimedia and emotion data for the problem of emotion recognition, but it is clear that such methods would provide great benefit.

6. Conclusion

This research aims to promote interactive learning atmosphere between students and lecturer in a Higher Education environment by considering the student’s emotion and behavior as a pattern of generating mutual interaction that are likely to be accepted by the students during that particular session.

Our research presents the efforts in recognizing a student’s emotion and behavior through the perspective of computational system. We compared some existing works and identified how emotions and behaviors can be modeled in order to enhance the system functionalities and responses to the user by transforming the responses into meaningful information.

We also defined and discussed the appropriate capacity of meaningful information to generate analyses and recommendations of mutual interaction for the users.

References


