International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 06 | June-2018 www.irjet.net p-ISSN: 2395-0072

A BASIC REVIEW OF AFFECT DETECTION AND ITS APPLICATION IN E-LEARNING

Vidhi Ambwani[1],Shivangee Kushwah[2]

ABSTRACT:

Emotions are intrinsic behavioral qualities which help humans to interact with each other. A human being is able to understand emotions of another human being to a limited extent. As technology is all about developing an intelligent environment and reducing human efforts, it is fascinating to develop an environment which can detect human's emotional and affective states. This review paper reports about emotions and Affective states and its applications in the emerging fields. This paper also revise various methods and models on how learning systems Can be enhanced and made more effective when Merged with human's emotion

1) INTRODUCTION:

Emotions play crucial role in our lives .Perceptions differ from person to person .A person's reaction to a certain situation may be completely different from how others react in that situation. Emotions are very complex to understand and hence affective computing came in picture.
Affect detection is a technique that detects human emotions through body gestures and postures,facial expressions,voice and eye movement .With advancement in technology humans expect computers to sense their mood and mould themselves in such a way that they are able to fulfill users need. Such systems reduce the interaction gap between computer and user making it more reliable,effective and natural.
Affective computing is used in variety of fields ranging from driving a vehicle to being used in distant learning programs. System provide inputs to increase the performance of user by sensing their affective state.

The machine analysis of human affective state is valuable in various situations Applications of affect detection:

a) In order to have a better interface or to improve performance of an application emotions play an important role.Some applications whose execution is not proper are bound to lose user's attention as it causes frustration, while an attractive looking app gains more users.

b) Development is being done in the field of bots that can walk ,talk and act like human and recognize human emotions too. For the physical and mental development of child a positive environment is required and so 2 dolls namely Actimates Arthur and D.W were created having similar features to human and interacted with children to inculcate good manners in them ,motivate them and be their confidant .They aimed for the psychological growth of children by creating a happy atmosphere.

c) Sensors are needed to measure physiological signals. Truck drivers tend to fall asleep while driving and so there are sensors that detect their state report the system that alerts them to stop moving .

2) METHODS OF AFFECT DETECTION:

2.1) AFFECT DETECTION THROUGH FACIAL DETECTION:

Facial features are very expressive and it is the most efficient way to detect a person's mood.
User-independent models use facial interpretation in such a manner so that they are applicable for a large number of users.People may show different effects for similar expressions.

Some major challenges being faced in facial affect detection is:

To build a user dependent model as expressions vary from person-person and detection of spontaneous expression is tough.It is observed that user independent model has less accuracy.

Another problem faced is the model trained for a specific person can never have sufficient information as responses are obtained from a single person only and annotation is also required on obtained information Therefore we devise a system that reduces human efforts of labeling and automatically recognize facial expressions that indicate the affect state of person.

Segment based labeling is used to overcome the above anomalies.Multiple instance learning is an approach used in PADMA(Personal affect detection with minimal annotation).

Experiment was conducted where a video was given to watch that elicited certain responses(emotional state) of viewer. Facial expressions are captured from this video.These expressions are further divided into feature vectors and we get a series of these vectors. Vectors having the same characteristics undergo clustering .Expression label is assigned to each cluster,to simplify the analysis of features.Most frequently occurred gestures while affect elicitation and user's feedback(neutral,happy,interested,bored etc) after every video is watched are the key facial gestures .It effectively assign affect labels to these gestures.[3]

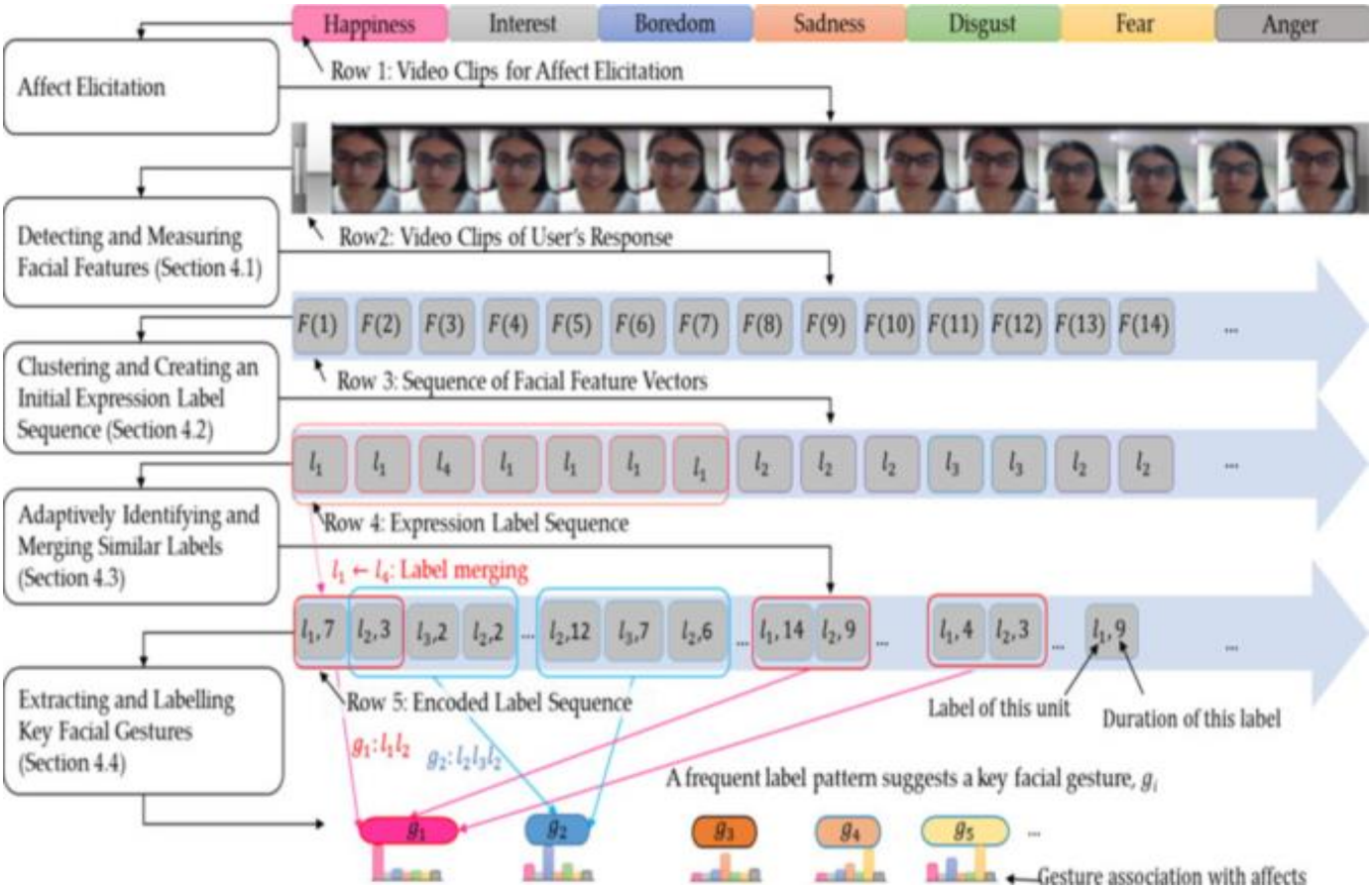


Fig 1. Personal affect detection with minimal annotation system[3]

Another experiment to detect facial expressions is by adult attachment interview(AAI).It required the participants to describe their happy and sad memories from childhood to adulthood.Observers were made to watch videos where high-resolution cameras recorded their facial expressions.To decode every facial expression the FACS[4](facial action coding system) was used and it was done using only visuals there was no audio.In most pattern recognition a training dataset is present,here there are two sets: one for emotional state and other for non-emotional where the non-emotional state is very diverse.But to detect the emotional state and distinguish them from non-emotional states one class classification is used.One of the solution to one class classification problem is novelty detection where if the probability density of an object is less than a given threshold value then it doesn't lie in the target class(emotional state).

2.2) AFFECT DETECTION THROUGH BODY GESTURES:

For different emotions humans have different head motions,these head motions are used to understand affect state of human and improve dataset of facial affect detection.Database used for study of head motion is Acted Facial expression in wild(AFEW)[5].This consists a variety of video clips that were classified under seven basic emotions i.e anger,disgust ,fear ,joy,sadness,neutral,surprise. Head posture is measured in terms of three angle of rotation(pitch,yaw,roll)and further represented as time-series of head movement.In order to categorize head motions they can be presented as the rms values for pitch,roll and yaw of head pose in terms of velocity,acceleration and angular displacement.These values must be different for different emotions.AR models are used to get the temporary dynamics for head movement data. These head movements correctly distinguish emotions and improve the efficiency of emotional recognition system.Another method for head gesture is nod and shake to recognize the affect states.With Microsoft Kinect camera head pose is determined.This pose is determined as a series of head movements in a particular order[6].Head nod is moving head in alternate up and down motion and shake is moving head in alternate left and right motion.The value of difference of angle for pitch and yaw for above movement in parallel frames helps in determining the movement.The three Hidden markov models are trained based on basic 5 head movements that are left,right,up,down

The 3 HMM'S are:

1)Shake HMM consist of left,right and still states

2)Nod HMM has up,down and still state

3)HMM for identifying other head movements other than nod and shake consist of up,down,left,right and other states.After this the value that occurs frequently for nod or shake is selected and compared with predefined values.If the predefined values are less , then gesture is head nod or shake but if not then it is other head gesture.

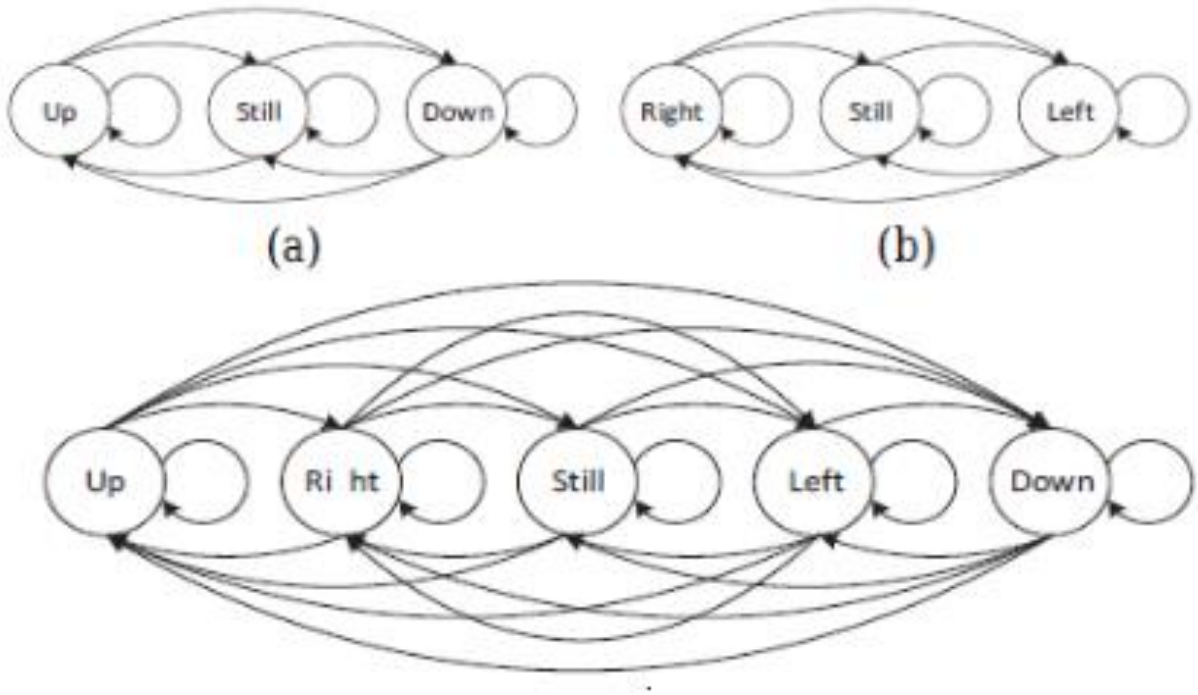


Fig 2 (a) HMM model for nod **(b)** HMM model for shake

(c)HMM model for other getsure [6]

2.3) AFFECT DETECTION THROUGH EYES

Emotions can be extracted and recognised through a person's eye movement and the increase and decrease of its pupil size while talking. They say that eyes can do all the talking that is true because while communicating ,eyes play a major role in expressing one's emotional state. Through this we can convey or detect emotion of paralysed individuals which is a breakthrough in communication between disabled individuals.
There are different and a vast number of emotions that a human goes through. But we consider only seven main emotions in this paper:-

Lying:
When people lie, their eye movements are more in number comparatively to the normal eye movements. The eye tends to move away, that is to the right or left when something is being asked. Eye contact is not maintained for long time with the other person with whom the conversation is being carried out.

Stress
When people are in stress the blink rate increases. It is noticed that the persons blink rate is higher than the normal blink rate. This happens when one is put in a challenging situation.

Disgust/Distaste
When someone is feeling the emotion disgust or disgusted then their eyes become more narrower than normal. This way one can a know a person is feeling distaste or disgusted.

Discomfort

When a person with whom you are communicating feels uncomfortable that they try blocking their eyes by looking away from you and avoid eye contact.

Happiness
Through eyes, when someone is feeling happy the eyes
are larger than normal size.The pupil size also increases when one is happy and cheerful.

Fear or Surprise
When someone is scared or surprised the eyes are wider in size than normal and is accompanied with an 'O' shaped mouth expression. This way we can say one is surprised.

Focus
When someone is focused then that person maintains eye contact for a long duration of time.The eye movements are very less and constrained.

2.4)AFFECT DETECTION THROUGH VOICE

Voice is one of the main medium through which one can understand how the other is feeling or which emotional state the person is in. human being experience different emotions for example: neutral , anger, joy, sadness, anxiety. In the modern world, with increase HCI voice is taken as an important medium through which a computer will be able to understand what a human is feeling or in which emotional state is he in. This is done by detecting and recognizing emotion for the voice signal of a person. Through this a computer is able to distinguish between different emotions that the person is going through.

[9]The recognition of voice can be divided into three categories:

1. Detection of the speaker: the speaker is detected so as to develop a model to be worked upon. To achieve this the user has to first enrol their voice further which the model will be used for
performing the next process of the system. The system can be either text-speech based or solely voice based system.
2. Detection of the speech: once the model is enrolled the process of extraction and matching is done
3. Language analysis: when considering emotion detection through voice it is essential to focus on how the model is speaking. The language used by the model and the variations in the voice are two key features to be analysed.

EMOTION RECOGNITION METHOD

In [10] to understand the way features has an impact in speech under different emotional situations data was extracted from subjects. There were three of them and their voices were recorded. They were told to express particular emotions and their voice was recorded on a device. The device that was used to record the voice was a mobile phone. The candidates were Russians and they were told to speak under different emotional state. The device that was used to record the voice was a mobile phone and kept at a distance of 15cm from the mouth. The experiments environment was an ordinary bedroom of 5mx5m.MATLAB functions were used to get the features from the recorded voice notes.
Analysis of the voice notes were done in MATLAB.

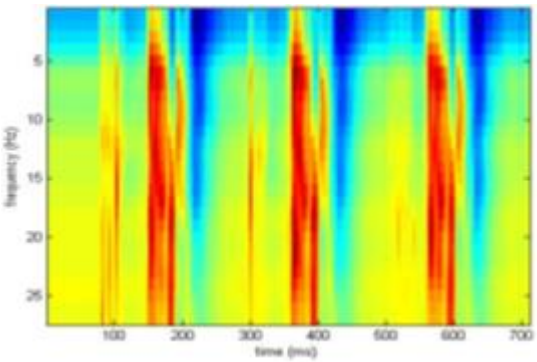
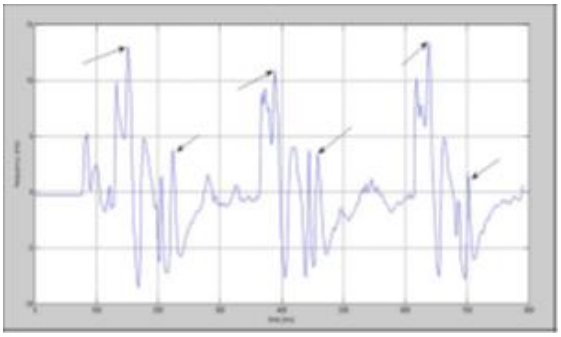


Figure 3. (a) The Power spectrum of a speech signal



(b) MFCC of a signal a candidate

Pronounced a word for three times. The distance between the indicated peaks are used to form a feature vector.
The Mel-frequency cepstral coefficients (MFCC) are widely used. It extracts the features of the signal and represents it.It takes short time spectral shape about the quality of voice and production effects for calculation of their coefficients has transfer of real log of short term spectrum of energy. Then it's performed in frequency scale. Then pre-emphasising of the speech segments are windowed. In this process the simple version of hamming window is used which is based on the

reduction of the leakage effect. The concept of windowing is based on multiplying the signal frames through the window function w(n).

To classify emotion in human a big set of data is formed. Each candidate's voice was recorded for 30 times and the recorded data was divided into 3 categories of emotional state that is neutral, anger and joy. A limited number of emotions were taken. The set of data had three fields which are feature distance, heart rate and class. Here A,B and C is denoted for neutral, anger and joy which the class attribute contains. Then a software called weka software is used to classify data. It was run nine times. The percentage varied from 10%-90%.The experiment was performed 30 times and the average value was calculated as well as the standard deviation.

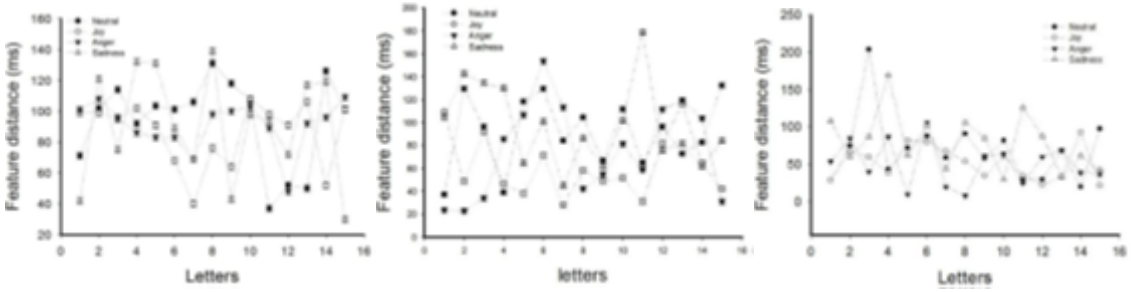


Fig-4[10].For (a) person 1, (b) person 2 and (c)person 3 is the length value between peaks for a single word.Joy has the least feature distance

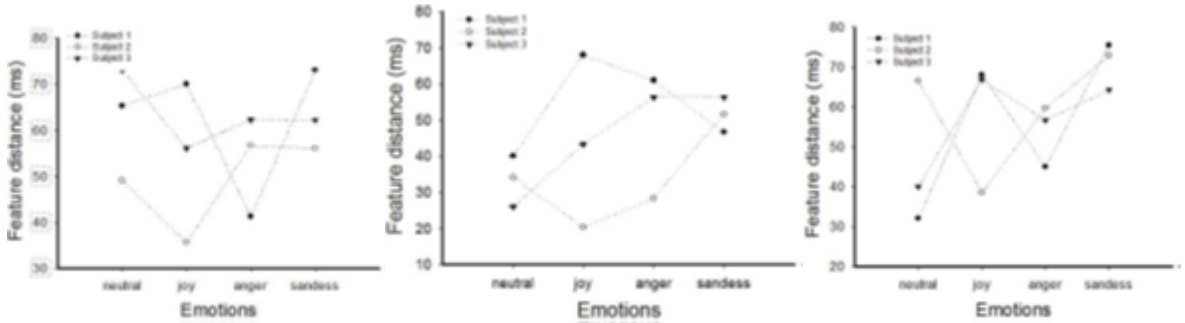


Figure 4. (b)[10] (a) Phase 1, (b) Phase 2, and (c)Phase 3 for all three candidates under emotions: neutral, joy, anger, sadness.

observing figure 2 and 3, we can infer that the vocal cord or voice signals produced by a person is influenced by the emotion of that person in a direct manner. The words being said is independent of how that person feels.

ACOUSTIC PARAMETERS

Detecting emotion through the voice of a person provides a better perspective about the emotional state of a person. This emotion detected is completely independent of the words being used by the Person. Emotion is detected by taking the person's voice as an input.

There are sources of data which are used in detecting the emotion through voice. They are:

Utterances:

This is a set of voice that has been recorded by many candidates both male and female. Each candidate has incorporated various emotions of the human nature.

Corpora of different languages:

Every language is spoken differently with each and every word being pronounced differently.There has been many situations in which the same words are pronounced differently due to the influence of the mother language over a common language like English words are pronounced differently over the world. Detecting emotion is independent by the languages that are spoken as well as the text.

Call centre corpora:
In workplaces especially call centers, customers complains and give their feedback as well as review. The conversations between the customer and the call centre employee are recorded and stored in their databases. Here customers portray combinations of emotions which gives a good outline in emotion analysis as it has various fluctuation of emotion.

TWO STREAM PROCESSING[8]

Verbal communication is very expressive and communicates the emotional state in precise manner compared to the non-verbal communication. Therefore, spoken language is comparatively more expressive. The possibility is high when is comes to capturing expressions through the language being spoken. Mainly energy and pitch is co-related to the emotional state of the speaker. We use the SHR i.e. the subharmonic to harmonic amplitude ratio for the calculation of pitch value from the utterance.
The first step is to find the location of the global maxima i.e. logf1 and taking that as a point of reference, we have to find the location of the next maxima i.e. logf2. The SHR is calculated
SHR=0.5
DA(logf1)- DA(logf2)
DA(logf1)+DA(logf2)*5
Where DA is a function called difference function which represents the difference between the odd and even samples in the log frequency scale. If threshold value is greater than SHR, f2 is a final pitch or else f1 is chosen as the final pitch. ACCORDINGLY, PITCH VALUES Are calculated for other utterances over the signal. These are the following parameter for the acoustic features.

a) Mean tone
b) Maximum tone
c) Minimum tone
d) Standard deviation of tone

These features are calculated over the first derivative of the pitch contour. These then form the 8 pitch features. The energy is then calculated. Then, following features are captured.

a) Mean energy
b) Highest energy
c) Lowest energy
d) Standard deviation of energy

Then, 4 more features of energy are generated from the first derivative of energy. Thus, there are 16 features that have been derived from pitch and energy of the signal which consists the acoustic parameters. The emotional status are captured on the contour and not on the signal. Therefore, the first derivative captured this information. The 16 dimensional feature is then calculated by the Gaussian mix model. Va the output vector is produced after the calculating the input feature vector over the aussians of different emotional categories. This what the emotional category can be recognized by calculation of acoustic parameters.

3) E-LEARNING

E- learning is a learning system based on electronic resources. It can also be referred as transferring of skills and knowledge via network based systems helping in delivering education(knowledge) on a large scale. It is useful not only in classroom projects but in corporate sector too. For example, when training programs are conducted by companies across the globe in forms of meetings and seminars.[11][12]13]

It is believed on basis of various researches that human brain can easily remember and relate to things that are visually represented. It not only grabs the attention of the mind but also the information is retained by the brain for a longer period of time. However, students may drop their focus because of no facial interaction. Due to which the focus levels may fluctuate.[13]

In e-learning systems, emotions play a vital role in order to enable students to receive knowledge anytime anywhere and build new knowledge and skills to improve performance.It is a centered more precise learning technology. Only few reports that have been explored in the field of affect detection in e-learning technology. The past research has only been focusing on the traditional classroom face to face learning.[11]

Past methodologies were centered on only classroom face to face interactions. But now due to learning in cognition process, learning can be made more simpler, interesting and effective.[11][15]

4) AFFECT DETECTION IN E-LEARNING:

Affect detection being a technique that deals with emotions and affective states of human being can be used in decision making which produces an unique environment for students and give learning experience focusing on each individual.

The relationship between learning and emotions goes hand in hand. It is difficult for a person to concentrate if he is not mentally stable. Emotions like anger, depression, sadness, anxiousness can make a person's mind wander resulting in not being able to focus on what is being taught. It is now recognized that positive (joy, satisfaction, curiosity etc)and negative(frustration, sadness etc) states trigger may lead to diversion and may result in possible hurdles while teaching and training.[14]

To overrule the disadvantages and pitfalls of the old e-learning system a Brain Control Interface known as EEG has been introduced which captures the brain waves and can be used for mapping concentration with emotion of the learner.

5) METHODOLOGY[15]

A Brain Computer-Interface is basically a system for communication where activity signals are generated using EEG.An electroencephalogram (EEG) is a device which helps in measuring the activities of brain and keeps the electrical activity of brain recorded. Special sensors in form of electrodes are attached to the head and connected by wires to the computer. BCI provides conveyance between human mind and machine without any physical contact.
To develop a strong BCI system the operations used to detect EEG signals have to be stipulated accurately.

This process has to undergo five steps. The first step is signal procurement step, second step is noise eviction step , the third step is characteristic extraction,the fourth step is producing output to the devices.
The procedure can be explained in elaborated approach by which a new system inclusive of fundamentals can be developed:

1. EEG is interfaced with software for detecting emotion signals in real time. Multi modal system is used as BCI headset and EEG signals are measured. A basic headset obtains the attention span values ranging from 0-100.

Fig. 5 an EEG sensor device

To detect visual feedback, system used kinect motion sensor which tracks the user's face and calculates the attentiveness of the user.

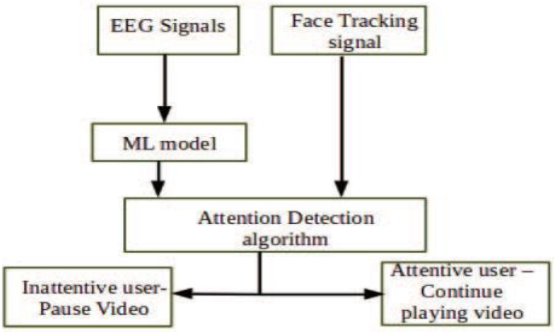


Fig 6. a flowchart explaining the attention detection algorithm

If the user's attention comes to be less or reduced than the required attention to understand the video, the video can be paused.

If the user is able to give attention as per the requirement then the video should be continued.

2. ML model-Data is obtained from BCI module wherethe signals are made free of noises anddisturbances like eye blink and duplicate values. While the video is being played two variables can be used to note the attentiveness and distraction factor of the user. For example Data obtained after processing is further divided into training and test data.Training data trains classifier and label data as(Patt) or (P!att).Stratified k fold cross validation technique is used,to test accuracy.Original training data set provides:k folds of training and test data set.

3. The third section focuses on the processing ofthe signals. The noise free signals are distinguished from each other on basis of the data which is already processed. Only selective information is processed. This feature is very critical as choice for result of appropriate classification is performed.

4. The fourth section deals with the output to the user. The software should be able to give output in which the classified signals is processed and converted into knowledge. Thus reflecting the status of the emotion and attention span of the user according to which system adapts and performs operation.[5]

Learning is a vast idea which involves incorporation of textual as well as non textual data but how the learner reacts and grasps thoroughly through these means is why these systems are evolving .The defiance is to determine the accurate support which will be able to take profits of relevant information. This will help in not alone to make the system adaptable but provide affective feedback.[13]

CONCLUSION

Advancement in technology has made it possible for computers to sense human emotions that provides better interface, make user system interaction easy and interesting. Affective computing has found its way in almost every field. Models were given and experiments were performed in order to detect the affective state.Some of the models were HMM,AAI experiments and so on.The attention span of humans is decreasing day by day and so system was developed in the field of education that helped students in performing better and made learning more effective.This system captured human signals to know the state of mind of users.Such systems were useful in driving and
many other applications and are very oftenly used now a days.

REFERENCES:

1. Emotion-sensitive Human-Computer Interaction (HCI):State of the art - Seminar paper Caroline Voeffray University of Fribourg

2. Toward an Affect-Sensitive Multimodal Human–Computer Interaction MAJA PANTIC, MEMBER, IEEE, AND LEON J. M. ROTHKRANTZ PROCEEDINGS OF THE IEEE, VOL. 91, NO. 9, SEPTEMBER 2003

2. Toward an Affect-Sensitive Multimodal Human–Computer Interaction MAJA PANTIC, MEMBER, IEEE, AND LEON J. M. ROTHKRANTZ PROCEEDINGS OF THE IEEE, VOL. 91, NO. 9, SEPTEMBER 2003

3. Identifying User-Specific Facial Affects from Spontaneous Expressions with
Minimal Annotation Michael Xuelin Huang, Grace Ngai, Kien A. Hua, Fellow, IEEE, Stephen C.F. Chan, Member, IEEE, and Hong Va Leong.Member, IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, VOL. 7, NO. 4, OCTOBER-DECEMBER 2016

4. Spontaneous Emotional Facial Expression Detection Zhihong Zeng University of Illinois at Urbana-Champaign, Urbana, USA, JOURNAL OF MULTIMEDIA, VOL. 1, NO. 5, AUGUST 2006.

5. ON THE ROLE OF HEAD MOTION IN AFFECTIVE EXPRESSION Atanu Samanta, Tanaya Guha Electrical Engineering, Indian Institute of Technology, Kanpur, India ,ICASSP 2017

6. REAL-TIME HEAD NOD AND SHAKE DETECTION FOR CONTINUOUS HUMAN AFFECT RECOGNITIONHaolin Wei1, Patricia Scanlon2, Yingbo Li1, David S. Monaghan1, Noel E. O'Connor

7. IJCSNS International Journal of Computer Science and Network Security, VOL.16 No.3, March 2016,HCI and Eye Tracking: Emotion Recognition using Hidden Markov Model,B.Rajakumari, N.SenthamaraiSelvi.

8. International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 5, May 2016, Emotion Detection Analysis through Tone of User: A Survey, Ms. Ankita Dutta Chowdhuri1, Prof. Sachin Bojewar

9. Ubiquitous computing seminar FS2014, Speech recognition for human computer interaction, Niklas Hofmann,ETH Zurich

10. Procedia Computer Science , 2015, Detection and Analysis of Emotion From Speech Signals, Assel Davletcharovaa, Sherin Sugathanb, Bibia Abrahamc, Alex Pappachen James

11. Neha G. Prabhu , Neha S. Singh , Shweta V. Singh , Professor Nilima Patil,AFFECTIVE E-LEARNING USING EMOTION DETECTION,International Journal of Technical Research and Applications e-ISSN: 2320-8163, www.ijtra.com Volume 4, Issue 2 (March-April, 2016), PP. 216-220

12.Liping Shen,Minjuan Wang and RuimimShen,Affective e-Learning:Using "Emotional" Data to improve Learning in Pervasive Learning Enviornment,Education Technology & Society,12(2),176-189.

13.Luigi Anolli,Fabrizia Mantovani-2,Massimo Balestra3,Alessia Agliati1,The Potential of Affective Computing in E-Learning

14.Emotions and Personality in Adaptive e-Learning Systems:an Affective Computing Perspective,Chapter in'Emotions and Personality in Personalized Services'.Part of the series Human-Computer Interaction series,pp 263-285

15.Namita R.Tambe,Ajitkumar Khachne,Mood Based E-learning using EEG.