Journal of Theoretical and Applied Information Technology
30th June 2016. Vol.88. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

HUMAN EMOTION DETECTION THROUGH FACIAL EXPRESSIONS

[1]KRISHNA MOHAN KUDIRI, [2]ABAS MD SAID AND [3]M YUNUS NAYAN
[1,2]Computer and Information Sciences Department, Universiti Teknologi PETRONAS, Malaysia
[3]Applied Sciences Department, Universiti Teknologi PETRONAS, Malaysia
E-mail: kmk.iiit@gmail.com, abass@petronas.com.my, yunusna@petronas.com.my

ABSTRACT

Human to human social communication in real-life is possible through different modalities like facial expressions, speech and body poses. However, facial expressions plays important role while dealing with human emotions in real-life than the other modalities. It is because facial expression provides non-verbal data towards emotions. And also gives emotion of a person towards his goal. On the other hand, speech and body poses are mostly language and culture dependent respectively which creates problem while detection emotions of a person. Thus in order to deal with the above issues, this research work focused on facial expressions instead other modalities. To improve detection performance of the system, proposed Relative Sub-Image Based features is used. Support Vector Machine with radial basis kernel is used for classification. Total six basic emotions (angry, sad, happy, disgust, boredom and surprise) are tested. From experimental results, the proposed Relative Sub-Image Based features enhanced the classification rates than the conventional features.

Keywords: *Relative Sub-Image Based (RSB), Support Vector Machine (SVM), Human Computer Interaction (HCI).*

1. INTRODUCTION

Usage of computers become so popular so that, it's become part of daily life. Now days, computer is become important object for every person. Because of that, apart from traditional computation, Human Computer Interaction (HCI) need to assist for every customer so that they can feel free to use. Due to that, HCI need to assist every common user of any age, any gender and culture. Even here users no need to be educated. He may be physically handicapped. But, still computer should assist user effectively. In order to do so, computer should not be like desktop (monitor, keyboard and mouse). According to the above scenario, HCI should not be traditional (programming) and it should be more social (touch, speech and facial expressions) to understand user emotions. Thereby, social human computer interaction comes into the picture.

Nass et al. [1] stated that, computers are social actors. In order to quote the above, an experiment was conducted between people and computer. In the experiment, people need to interact with computer in a social manner. Finally the response of people is studied. According to that, the people never felt that they are interaction with machine. They felt more social and natural. Thus, social HCI can play important role in social life of a person than traditional HCI.

Brave and Nass [2] stated that, human emotions are important to understand to make it more social. Human to human communication in social environment is a mixture of words and emotional expressions together. Emotions are influencing on human-human communication. Thus, HCI needs to estimate human emotions to make more social and comfortable to user.

Picard [3] released a book on ―effective computing that relates to, arises from or deliberately influences emotion. In the book, author discussed about possible abilities of computer as follows: computer can detect emotions, computer can express emotions and computer can behave emotionally depends on what user need. This book changed the view of computer scientists and made them to focus on social HCI.

In human-human communication, each person need to encode and decode his emotions and others emotions respectively. This may through speech, facial expressions and body moments. When two persons are communicating each other, the above processes used to be in dynamic in nature. Understanding of one modality may affect the view

Journal of Theoretical and Applied Information Technology
30th June 2016. Vol.88. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

on other modality either positively or negatively. On the other hand, understanding of one modality sometimes may create ambiguity to understand other modality. Thus, encoding and decoding of multi-sensory is error prone.

In a study about human emotions through speech and facial expressions by Mehrabian [4], it is stated that, verbal and non-verbal information of speech signal conveys 7% and 38% respectively. On the other hand author also stated that, facial expressions convey 55% emotional information than speech signal. Thus, facial expression plays higher role while dealing with emotions.

Wehrle and Kaiser [5] stated that, facial expressions of a person have the following roles to encode his emotions in human-human communication as: 1) Speech regulation: according to the listener‘s reaction or facial expression, encoder get clear idea to regulate his speech to convey him clear message. 2) Speech related signal: while talking, speaker can raise his eyebrows to support his speech to listener. So that, listener will get facial signal with verbal data. For instance, a smile while talking can indicate that, what is being said is not serious. 3) Signaling for relationship: when a couple discusses a controversial topic openly, a smile indicates no harm to their relationship. 4) Signaling for cognitive process: while thinking any hard problem, frowning often occurs. 5) Signaling for emotion: If a person smile, he may be happy inside. Otherwise, he wants to convey his message to somebody to achieve some goal. Silva et al. [6] stated as follows: 1) Speech signal is difficult to understand when noise exists. 2) Facial expression of an encoder is difficult to understand while he is speaking. 3) Dedicated emotions are easy to detect from one channel instead of all.

(a) (b)

Figure No 1: (A) Deliberate Sad Facial Expressions (B) Sad Facial Expressions During Speech

Figure 1 illustrates sad emotion in two cases namely, deliberate and during speech. In the figure, sad emotion is different for both deliberate and during speech. But, both convey same type of emotion.

According to the above discussion, it is concluded that, identification of human emotions through facial expressions are still a very challenging field that target methods to make effective human computer interaction. In order to deal with the problems, this research work introduced proposed Relative Sub-Image Based (RSB) features through facial expressions. Furthermore, Ekman et al. [7] stated that, identification of every emotion is not possible. Instead, this research work focused on list of basic emotions namely, angry, sad, disgust, fear, happy and boredom. Finally Support Vector Machine (SVM) is used for the classification of emotions through facial expressions.

The rest of the paper is organized as follows: Section 2 discusses about related work. Section 3 explains about proposed RSB based features. Section 4 presents results and discussion and finally Section 5 gives the conclusion and future work.

2. RELATED WORK

Feature extraction from facial expressions is possible in two ways namely geometric based and appearance based. Geometric based feature extraction focuses on complete face to extract emotional data. Whereas to extract emotional information, the appearance based technique focuses on facial skin changes likely, winkles and bulges.

2.1 Geometric Based Method

Lien [8] used wavelet based method for complete face regarding motion detection. PCA is used to optimize the data to process in vertical and horizontal directions. In this case, there is no guarantee that PCA always gives better accuracy. Otsuka and Ohya [9] used local regions of eyes and mouth regarding motion detection. But, the above techniques contain noise and loss of data problems.

Wang et al. [10] identified geometric displacements using manual feature points in the form of lines and dots near to eyes, eyebrows and mouth to detect emotions. However, this method failed to predict feature points automatically. Kaliuby and Robinson [25] applied color plastic dots to the users face to recognize the facial muscles clearly. This method gives better accuracy than previously unused method. However, labeling

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the points manually is not going to be useful for natural human computer interaction.

Donato et al. [11] subtracted the input image with the reference image to analyze the face. This method is only applicable to person dependent systems which are not suitable for real-time conditions. Next, the proper normalization technique is needed to normalize the training samples to get reference image. Thus, this approach is not suitable for fully automatic emotion detection systems.

2.2 Appearance Based Method

Vukadinovic and Pantic [12] used Viola-Jones algorithm to detect faces and then the face is subdivided into 20 regions. Next, each region is tested for region of interest through Gabor wavelets. Padgett and Cottrell [13] used Kernel based approach to detect emotion through facial expression. Here, Principal Component Analysis (PCA) is used for full face to detect emotions. The advantage of PCA is fast and simple. Moreover, PCA provides better accuracy for person dependent emotion detection system in natural human - computer interaction.

Hung-Hsu et al. [14] developed a novel human facial emotion detection system using angular radial transform, discrete cosine transform and Gabor filter. This system contains loss of data due to filters. Thus, the system cannot provide better accuracy for real time data. Pentland et al. [15] used Eigen faces to predict human emotions through facial expressions. The performance of the Eigen faces effects due to noise. PCA is used for dimensionality reduction. Using the PCA, the emotion detection system is able to give better accuracy. The complexity of the PCA is directly proportional to the size of the input file [16]. Furthermore, there is no guarantee that PCA always gives better accuracy. From the above two, hybrid technique came into the picture.

Ekman and Friensen [17] introduced facial action units to predict human facial expressions. According to this method, human face is divided into 44 parts and each part is marked with a point manually. Then according to the moments of these landmarks, Ekman‘s work predicts human emotions. He concluded that, compared to other methods, this method can able to predict more emotions effectively. But, it is not applicable for fully automatic systems.

From the above studies, identification of emotions through facial expression needs to be suitable for real time data. While the above mentioned methods, every method focused on facial expressions and it‘s experienced emotion in outside environment. Here no one is focused on actual meaning of emotion. According to Darwin theory [7] identification of human facial expression is helps to understand user emotions towards his goals. People use facial expressions to express their feeling to others to reach their goal. Identification of every human emotion through facial expression is not possible except basic emotions. Basic emotions relates to basic needs. Thus, identification of facial expressions for basic emotions is unique and easy to predict.

From the above discussion on 2D approach, it is concluded that no conventional method from the above approaches is supporting fully automatic system effectively. According to that, appearance based approach is used by the proposed research work in order to implement fully automatic emotion detection system for real-life conditions. To do so, the proposed research work used RSB features for feature extraction. The detailed discussion is given in the Section 3.

3. PROPOSED APPROACH

Face detection means face prediction in given input image, if there is any. This is a challenging task in real time conditions due to the following, pose and angle of the face can vary due to the camera position, target face may occluded or close up to the other objects and lightning problem etc. Furthermore, face is unique to every person and same person may look different with eye glasses, beard and moustache etc. Next, to evaluate the performance of the face detector in real time, HCI depends on the following factors: detection speed, detection accuracy, required training time and number of training samples etc. Consider the above mentioned metrics in this research work, Ada-boost based face detector by Viola and Jones [13] is used. The face detector uses cascade of sample classifier at early stage to filter out most negative samples and then uses strong classifiers to dealing with instances that look like faces.

Identification of human emotions is practically difficult, since most of the emotions are close to each other. However, basic emotions are common to every culture. Thus, basic emotions are easy to distinguish. Ekman [18] suggested 6 basic emotions

Journal of Theoretical and Applied Information Technology
30th June 2016. Vol.88. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

which can be identified easily. Now the problem is how to classify adjacent emotion which is close to each other. In this case, angry, surprise and disgust are adjacent to each other. In order to deal with the above, the proposed feature extraction is needed. Facial expressions are fast in nature. Let's consider a video file of length 8 seconds which is shown in Figure 2.

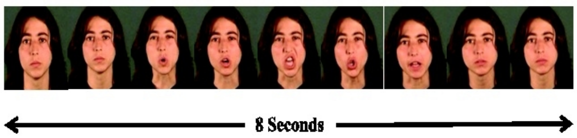


Figure No 2: Video File Of 8 Seconds Duration

This video contains both speech and facial expressions. Let's consider facial expressions in the single file. This file contains one emotion angry and neutral class which is shown in Figure 3. Here, almost every second, the emotion of that person changes from one class to another. This change is not random in nature. This change is completely depends on state diagram of Plutchik emotional circle [19].

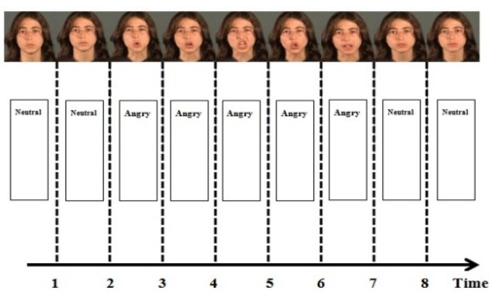


Figure No 3: Different Types Of Emotions In A Single Video File

Algorithm for Relative Sub-Image Based Features (RSB) [24]:

Step 1. Read image.
Step 2. Divide input image into sub-images.
Step 3. Calculate average of each sub-image is using pixel intensities.
Step 4. Calculate relative difference between each sub-image to its edge cent sub-images.

In order to detect human emotions when user is engaged in conversational sessions, the system needs to estimate emotions in a short amount of time. In addition, human facial expressions are fast in nature. Thus, identification of human emotions through facial expressions should be fast. From the above discussion, the proposed system used Viola-Jones algorithm to detect human face from input video. In this case, the input video is in MPEG format, and contains 30 frames per second. For each second in the video, middle frame is used for Viola-Jones algorithm to detect face which is illustrated in Figure 4. The detected face is cropped and resized for segmentation.

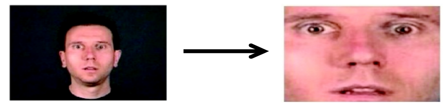


Figure No 4: Face Detection And Cropping

Next is image segmentation. In this stage, the image is segmented into 5×5. So, there are a total 25 segments. Each segment contains 100×100 size which is shown in Figure 5.

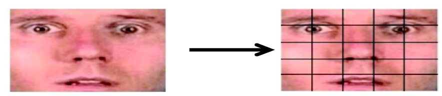


Figure No 5: Face Segmentation

This is done using an image segment function in Matlab. The way of segmenting is robust to every image. After segmenting save the segmented images separately. Next, compute the average of each segmented image. In this case, this average is done in RGB [24] separately for each sub-image which is shown in Figure 6. Thus, each sub-image creates 3 values from RGB segments. This step provides normalized data.

Figure No 6: Average Of Pixel Intensities

Finally, the computation of edge difference between every two sub-images through all possible combinations is done using Equation (1).

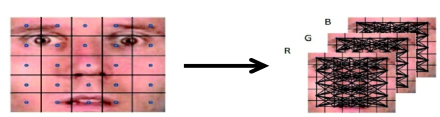


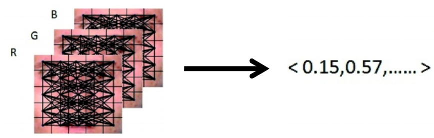
Figure No 7: Models Of RGB

$$\text{Difference} = \left[\frac{(X_2 - X_1)}{(X_1 + X_2)}\right] \quad (1)$$

Next, a RGB model is made using the edge difference between them which is shown in Figure 7. This helps to discriminate the facial expressions between them.

Journal of Theoretical and Applied Information Technology
30th June 2016. Vol.88. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



(a)

(b)

Figure No 8: (A) Relative Sub-Imaged Based Features, (B) Classification

Furthermore, the data vector contains 300 (${}^{25}C_2$) elements which is shown in Figure 8 (a). This vector goes to SVM classifier as an input or detecting emotions which is shown in Figure 8 (b). These RBS coefficients contain no loss of data compared to the PCA and Gaussian wavelet. The complexity of RSB features in the worst case is (MN). The proposed RSB uses pixel intensities to process the data which is similar to human observation in real-life conditions. It is also can able to discriminate adjacent emotions. For instance, human-human or human-computer communication is a mixture of deliberate and non-deliberate emotions. Angry and disgust are adjacent emotions. In order to predict these two emotional classes through facial expressions during speech is difficult for unimodal emotion detection systems. In this case, the proposed RSB features are suitable to deal with the problem.

Classification: Support vector machine (SVM) [20] is a popular classifier, which not only works properly with the known data, but also able to work well on unknown data. The main idea behind this SVM is to take a few class means as a typical model, and assess the contribution to split the cases as far away as possible. This is done through training. According to the statistical learning theory, larger margin always has a better ability to separate two classes in the SVM classifier. Here, the radial basis kernel is used to keep complex data in higher dimension space. Furthermore, the performance of the classifier depends on the type of kernel used in the system. Furthermore, in SVM, Gradient descent optimization algorithm is used, due to its less complex in nature. But, using this, perimeter C-value need to pick manually. Next, convergence is slow here compared with other optimization algorithms.

Database: The Database of Facial Expressions (DaFEx) [21] is a database of acted one, which is created by professional actors from Italy (4 males and 4 females). The database is suitable for unimodal and bimodal system through speech and facial expressions. It contains 1000 video samples, and each sample of 4 or 27 seconds duration. Next, the samples are divided into two parts namely, with speech and without speech. Further, the samples are divided into 6 classes of emotions, namely angry, disgust, fear, happy, sad and surprise. Every emotion is expressed in three intensities namely, high medium and lower. Here, each sample in the database labeled with the help of 80 observers. Thus, this is a suitable database for real time applications.

4. RESULTS AND DISCUSSIONS

The DaFEx database contains video samples with utterance and without utterance. Here, non-utterance one is used for training and testing. 300 facial poses for each emotional class are extracted manually and used for the training class. Next, identification of face and perimeter vector was done with the help of Viola-Jones algorithm and proposed Relative Sub-images Based feature, respectively. At this moment, two similar copies of training data were maintained for person dependent and person independent test. Furthermore, SVM with radial basis kernel is used for the classification. Here, the classification between six classes of emotions with neutral class is done using one against all approach. 10-fold cross validation is used for system validation.

Procedure of the experiment: Step 1: Initially the training set of facial expressions is used as a set of annotated images, which is created by detecting and cropping the face in the training set. This process is done with the help of Viola-Jones algorithm to create a basic model to detect facial expressions from input image.
Step 2: Create training data with the help of RSB features from facial expressions.
Step 3: Train the SVM classifier in order to detect emotions.
Step 4: Finally detect input emotions from video through facial expressions.

4.1 Emotion Detection through Deliberate Facial Expressions

Table No 1: Dafex Database Without Utterance

Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
150	150	150	150	150	150	150

Person Dependent Test:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table No 2: Comparison Between Conventional Features with proposed RSB

	PCA (%)	Gaussian Wavelet (%)	Proposed RSB (%)
Angry	85	83	**90**
Disgust	70	75	**83**
Fear	82	80	**88**
Happy	81	87	**90**
Sad	80	83	**90**
Surprise	75	73	**83**
Neutral	75	79	**90**
Average	78.28	80	**87.71**

Table 1 shows the size of the database. The database contains a total of 1050 samples, 150 samples in each class. From that 90% of samples is used for training and regaining 10% samples are used for testing. Table 2 shows the comparison between the performance of the proposed RSB features with other conventional (PCA and Gaussian Wavelet) methods through deliberate facial expressions. Identification of angry emotion is easy through facial expressions. Due to that, angry emotional class achieved better classification accuracy than other emotional classes.

On the other hand, this angry emotional class achieved 90% classification accuracy using proposed RSB features through SVM. Identification of disgust and surprise emotions is difficult through facial expressions. Facial expressions of sad and disgust look similar to each other. On the other hand, a facial expression of surprise and fear looks similar to each other. Due to that, these both disgust and surprise emotional classes achieved less accuracy than other emotional classes. But, proposed RSB achieved 83% and 83% accuracy for disgust and surprise emotions which is higher than conventional methods.

Furthermore, happy and sad emotional classes are easy to predict through facial expressions. Because, these two emotions are affect more number of facial muscles than other emotional expressions. Here also, proposed RSB achieved 90% and 90% accuracy for happy and sad emotions which is higher than the other conventional methods. On the other hand, the proposed RSB features achieved total average accuracy 87.71% which is higher. Total average accuracy of proposed RSB achieved better than conventional features which is also illustrated in Figure 9. Table 3, 4 and 5 shows confusion matrix for person dependent emotion detection through deliberate facial expressions of PCA, Gaussian wavelet and proposed RSB respectively.

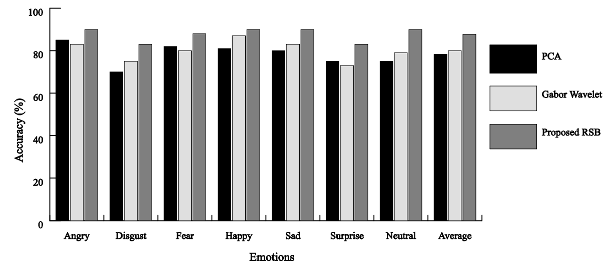


Figure No 9: Comparison Between Conventional Features With Proposed RSB

According to the above, angry emotional class confused with disgust emotion vice versa. Fear emotional class confused with happy emotion vice versa. Sad and surprise emotional class confused with fear class. However, the proposed RSB feature extraction affected less and achieved a higher recognition rate than conventional methods.

From the above confusion matrices in Table 3, 4 and 5, it is concluded that proposed feature can able to detect emotion better than the rest.

Person Independent Test:

Identification of emotions in real time scenario, input mostly used to be unknown and new. In order to handle this issue, the person independent classification is needed. Here also, the same database which is shown in Table 1 is used for training and testing. But, different actors used for training and testing.

Table 6 shows the confusion matrix for a person independent approach using RSB. Here also like Table 5, angry and disgust emotions confused each other and achieved 83% and 75% accuracy respectively. Similarly like above, fear and happy again confused and achieved 82% and 85% accuracy respectively.

In the other hand, Sad and surprise confused with happy emotion and achieved 83% and 78% accuracy respectively. Thus, the person independent classification also got similar problem like person dependent approach while classifying data. Furthermore, the accuracy of the person independent system is less compared to the person dependent system. This was done due to the unseen pattern identification problem. However, SVM is a better classifier to handle this problem. Finally, RSB performed better than conventional features in both person independent and dependent approaches.

Journal of Theoretical and Applied Information Technology
30th June 2016. Vol.88. No.3

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.2 Emotion Detection Through Facial Expressions During Speech

Table 7 shows a DaFEx database during speech. Similarly like Table 1, here also each class contains 150 samples. All these samples are confirmed to the particular emotional classes with the help of a group of observers. From that, 90% and 10% samples are used for training and testing, respectively.

Table No 7: Dafex Database During Speech

Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
150	150	150	150	150	150	150

Person Dependent Test:

Table No 8: Comparison between conventional features with proposed RSB

	PCA (%)	Gaussian Wavelet (%)	Proposed RSB (%)
Angry	63	61	**85**
Disgust	58	58	**73**
Fear	66	66	**83**
Happy	60	62	**85**
Sad	58	61	**85**
Surprise	55	50	**78**
Neutral	50	55	**88**
Average	58.57	59	**82.42**

In Table 8, neutral class achieved higher accuracy of 88% using proposed RSB features. This is done due to the problem of facial expressions with utterance. On the other hand, proposed RSB achieved better accuracy for angry, fear, happy and sad than conventional features at 85%, 83%, 85% and 85% respectively.

Finally, the proposed RSB features also achieved higher average accuracy of 82.42% than conventional features which is also illustrated in Figure 10. In Table 9, every emotion mostly confused with neutral emotion and less confused with other emotional classes.

This is due to the adjacent nature between the emotions which already discussed in Chapter 3. Furthermore, neutral emotion confused with every other class of emotions. It's due to nature of neutral class which already discussed in Chapter 3.

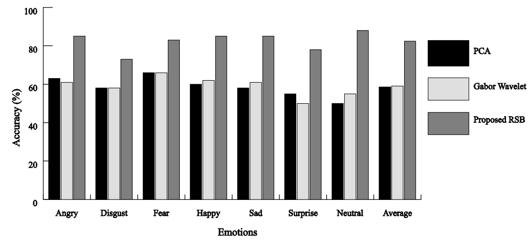


Figure No 10: Comparison Between Conventional Features With Proposed RSB

Person Independent Test:

Table 11 shows the confusion matrix for the proposed RSB. Similarly like person dependent classification, here also neutral achieved 87% higher accuracy than other emotional classes. Next disgust and surprise, emotions confused more with neutral class.

It is due to the problem with facial expressions during speech. The average accuracy of person independent test is less than person dependent test. However, proposed features still performed better than conventional features.

Furthermore, Table 12 compares both unimodal emotion detection systems through facial expression with utterance and without utterance, respectively. According to Table 12, the performance of the facial expression without utterance is higher while comparing with facial expression with utterance. Thus, the deliberate facial expression plays a major role. In order to overcome the above problem, observation of facial cures, namely; eyes, eyelids and lips helps to get improved accuracy. However, this is going to make a system semi-automatic. On the other hand, this increases system complexity. In order to deal with the above, another option is to provide a large amount of training data which can help to achieve better accuracy but, practically this is difficult.

Table 4.13 shows the comparison with proposed work with [22]. Simina et al. [22] used wavelet transform in order to estimate space frequency of input image. Daubechies wavelets are used to estimate emotions. SVM classifier is used. All the experiments were conducted using Cohn-Kanade database database. This work is not able to deal with adjacent emotions. However, proposed work achieved better accuracy than the conventional approach.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Similarly like above, Table 4.13 also shows comparison between proposed work with [26] and [27]. Classification is done with the help of SVM through radial basis kernel. Here also, this work is not able to deal with adjacent emotions. However, proposed approach showed better accuracy. According to Table 4.13, it is concluded that the proposed work performs well not only with DaFEx database, but also with other databases (Cohn-Kanade database).

5. CONCLUSION AND FUTUREWORK

The presented work provides ability to the computer to estimate user emotions on real-life conditions. It also provides fully automatic applicability through social cues. Furthermore, the proposed system predicts emotions not only for known person, but also for unknown person effectively. Future work need to focus on many issues about both unimodal and bimodal systems. The proposed RSB feature has lighting problem. The simple solution for this problem might be texture features. From the experiments, it is indicated that, performance of the system is less while dealing with real-life conditions. Regarding this issue, it might be better to have speech modality in order to deal with the problem. On the other hand, labeling of that training data is another open challenge. Furthermore, in the future it is better to increase number of emotional classes. So that, computer can understand user better.

REFERENCES

[1]. C. Nass, J. Steuer and E. R. Tauber, "Computers are social actors", ACM Conference on Human Factors in Computing Systems, Boston, USA, pp. 72-78, 1994.

[2]. S. Brave and C. Nass, "Emotion in human–computer interaction", vol. 1. New Jersey, USA: Lawrence Erlbaum Associates Inc, 2003.

[3]. R. W. Picard and R. Picard, "Affective computing", vol. 5. Cambridge, USA: MIT Press, 1997.

[4]. Mehrabian, "Communication without words", Psychological Today, vol. 2, pp. 53-55, 1968.

[5]. T. Wehrle and S. Kaiser, "Emotion and facial expression", Affective Interactions, vol. 1814, pp. 49-63, 2000.

[6]. L. C. De Silva and P. C. Ng, "Bimodal emotion recognition", IEEE 4th International Conference on Automatic Face and Gesture Recognition, Grenoble, pp. 332-335, 2000.

[7]. P. Ekman, W. V. Friensen and P. Ellsworth, "Emotion in the human face", vol. 11. Oxford, UK: Elsevier BV, 1972.

[8]. J.-J. J. Lien, Automatic Recognition of Facial Expressions Using Hidden Markov Models and Estimation of Expression Intensity, PhD Thesis, Washington University, St. Louis, 1998.

[9]. T. Otsuka and J. Ohya, "Spotting segments displaying facial expression from image sequences using HMM", IEEE 3rd International Conference on Automatic Face and Gesture Recognition, Nara, Japan, pp. 442-447, 1998.

[10]. J. Wang, L. Yin, X. Wei and Y. Sun, "3D facial expression recognition based on primitive surface feature distribution", IEEE Conference on Computer Vision and Pattern Recognition, New York, USA, pp. 1399-1406, 2006.

[11]. G. Donato, M. S. Bartlett, J. C. Hanger, P. Ekman and T. J. Sejnowski, "Classifying facial actions", Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 21, pp. 974-989, 1999.

[12]. D. Vukadinovic and M. Pantic, "Fully automatic facial feature point detection using Gabor feature based boosted classifiers", IEEE International Conference on Systems, Man and Cybernetics, Hawaii, USA, pp. 1692-1698, 2005.

[13]. C. Padgett and G. W. Cottrell, "Representing face images for emotion classification", Advances in Neural Information Processing Systems, vol. 10, pp. 894-900, 1997.

[14]. H.-H. Tsai, Y. S. Lai and Y. C. Zhang, "Using SVM to design facial expression recognition for shape and texture features", IEEE International Conference on Machine Learning and Cybernetics, Qingdao, China, pp. 2697-2704, 2010.

[15]. Pentland, B. Moghaddam, and T. Starner, "View-based and modular Eigen spaces for face recognition", IEEE Conference on Computer Vision and Pattern Recognition, Washington, USA, pp. 84-91, 1994.

[16]. P. Gosavi and S. Khot, "Facial Expression Recognition Using Principal Component Analysis", International Journal of Soft Computing and Engineering, vol. 3, pp. 2231-2307, 2013.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[17]. P. Ekman and W. V. Friesen, "Facial action coding system", vol. 1. California, USA: Consulting Psychologists Press, 1977.

[18]. Darwin, P. Ekman, and P. Prodger, "The expression of the emotions in man and animals", vol. 1. Oxford, USA: Oxford University Press, 1998.

[19]. R. Plutchik, "A general psycho evolutionary theory of emotion", Theories of Emotion. vol. 1, New York: Random House, pp. 3-33, 1980.

[20]. Punitha and M. K. Geetha, "Texture based emotion recognition from facial expressions using support vector machine", International Journal of Computer Application, vol. 80, p. 6, 2013.

[21]. Battocchi, F. Pianesi and D. Goren-Bar, "A first evaluation study of a database of kinetic facial expressions", IEEE 7th International Conference on Multimodal Interfaces, Trento, Italy, pp: 214-221, 2005.

[22]. S. Emerich, E. Lupu, and A. Apatean, "Emotions recognition by speechand facial expressions analysis", IEEE 17th European International Conference on Signal Processing Conference, Germany, pp. 1617-1621, 2009.

[23]. M. Mansoorizadeh and N. M. Charkari, "Multimodal information fusion application to human emotion recognition from face and speech", Multimedia Tools and Applications, vol. 49, pp. 277-297, 2010.

[24]. Krishna Mohan Kudiri, Abas Md Said and M. Yunus Nayan, "Emotion detection using sub-image based features through human facial expressions", International Conferece on Computer and Information Sciences, pp. 332-335, 2012.

[25]. R. E. Kaliouby and T. S. Huang, "Real-time interface of complex mental states from facial expressions and head gestures", Real-Tiime Vision for Human-Computer Interaction, vol. 25, pp. 181-200, 2005.

[26]. S. A. Caifeng, G. B Shaogang and W. M. Peter, "Facial expression recognition based on Local Binary Patterns:A comprehensive study", Image and Vision Computing vol. 27. pp. 803–816, 2009.

[27]. P. Michel and R. El Kaliouby, "Real time facial expression recognition in video using support vector machines‖, 5th International Conference on Multimodal Interfaces, Canada, pp. 258-264, 2003

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

APPENDEX

Table No 3: Confusion Matrix Of PCA

Actual /Predicted	Angry (%)	Disgust (%)	Fear (%)	Happy (%)	Sad (%)	Surprise (%)	Neutral (%)
Angry	85	6	1	0	2	0	6
Disgust	10	70	0	0	10	1	9
Fear	1	0	82	1	0	7	9
Happy	2	0	10	81	0	0	7
Sad	3	1	9	0	80	0	7
Surprise	0	2	9	1	4	75	9
Neutral	9	4	2	1	5	3	75

Table No 4: Confusion Matrix Of Gabor Wavelet

Actual /Predicted	Angry (%)	Disgust (%)	Fear (%)	Happy (%)	Sad %)	Surprise (%)	Neutral (%)
Angry	83	5	3	0	3	0	6
Disgust	17	75	2	0	1	1	4
Fear	0	0	80	3	2	8	7
Happy	3	0	6	87	0	1	3
Sad	3	2	8	0	83	0	4
Surprise	0	0	8	3	6	73	10
Neutral	5	4	5	2	3	2	79

Table No 5: Confusion Matrix Of Proposed RSB Features

Actual /Predicted	Angry (%)	Disgust (%)	Fear (%)	Happy (%)	Sad (%)	Surprise (%)	Neutral (%)
Angry	90	1	0	0	0	0	9
Disgust	3	83	0	0	0	0	14
Fear	0	0	88	0	0	2	10
Happy	0	0	0	90	0	0	10
Sad	0	1	0	0	90	1	8
Surprise	0	0	2	0	1	83	14
Neutral	2	1	1	2	1	3	90

Table No 6: Confusion Matrix For Proposed RSB

Actual /Predicted	Angry (%)	Disgust (%)	Fear (%)	Happy (%)	Sad (%)	Surprise (%)	Neutral (%)
Angry	83	2	0	0	0	0	15
Disgust	5	75	0	0	0	0	20
Fear	0	0	82	0	0	3	15
Happy	0	0	0	85	0	1	14
Sad	0	1	0	0	83	3	13
Surprise	0	0	2	0	2	78	18
Neutral	4	1	2	2	3	2	86

Table No 10: Confusion Matrix For Proposed RSB

Actual /Predicted	Angry (%)	Disgust (%)	Fear (%)	Happy (%)	Sad (%)	Surprise (%)	Neutral (%)
Angry	85	1	0	0	0	0	14
Disgust	8	73	0	0	0	0	19
Fear	0	0	83	0	0	3	14
Happy	0	0	0	85	0	0	15
Sad	0	1	0	0	85	3	11
Surprise	0	0	1	0	5	78	16
Neutral	1	2	1	4	1	3	88

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table No 11: Confusion Matrix For Proposed RSB

Actual /Predicted	Angry (%)	Disgust (%)	Fear (%)	Happy (%)	Sad (%)	Surprise (%)	Neutral (%)
Angry	80	5	0	0	0	0	15
Disgust	9	70	0	0	0	0	21
Fear	0	0	79	0	0	6	15
Happy	2	0	0	81	0	0	17
Sad	0	3	0	0	78	1	18
Surprise	0	0	2	0	2	75	21
Neutral	2	2	3	3	2	1	87

Table No 12: Comparison Between Emotion Detection Through Facial Expressions With Speech And Without Speech

	Angry	Disgust	Fear	Happy	Sad	Surprise	Neutral
Without utterance (%)	83	75	82	85	83	78	86
With utterance (%)	80	70	79	81	78	75	87

Table 4.13: Comparison with proposed work with conventional work using Cohn-Kanade database

Author	Features	Database	Classifier	Accuracy (%)
Philipp et al. [27]	Facial feature locations	Cohn-Kanad database	SVM with RBF kernel	71.8
Emerich et al. [22]	Wavelet transform	Cohn-Kanad database	SVM with RBF kernel	90.7
Caifeng et al. [26]	LBP and Template matching	Cohn-Kanad database	SVM with RBF kernel	86.12
Proposed work	RSB	Cohn-Kanad database	SVM with RBF kernel	96.2