“Emotion Detection using Facial Expression and Speech Recognition”

Imtiyaz Ahmad, Ramendra Pathak, Dr. Yaduvir Singh, Mr. Jameel Ahamad

Harcourt Butler Technical University, Kanpur (Uttar Pradesh)

Abstract

Emotion interpretation has arisen as an essential field of research that can provide some useful insight to a number of ends. People communicate their feelings through their words and facial gestures, consciously or implicitly. To interpret emotions may be used several different types of knowledge, such as voice, writing, and visual. Speech and facial expression have been the valuable tool for identifying feelings since ancient times, and have revealed numerous facets, including mentality. It is an enormous and difficult job to determine the feelings beneath these statements and facial expressions. Scientists from multiple disciplines are seeking to find an effective way to identify human emotions more effectively from different outlets, like voice and facial expressions, to tackle this issue. Computer intelligence, natural language modeling systems, etc., have been used to gain greater precision in this responsiveness towards various speeches and vocal-based strategies. Analysis of the feelings may be effective in several specific contexts. One such area is cooperation with the human computers. Computers can make smarter choices and aid consumers with emotion recognition, and can also aid render human-robot experiences more realistic. We would explore current emotion recognition methods, emotion modeling, emotion databases, their features, drawbacks and some potential future directions in this study. We concentrate on evaluating work activities focused on voice and facial recognition to evaluate emotions. We studied different technical sets that were included in current methodologies and technologies. The essential accomplishments in the sector are completed and potential strategies for improved result are highlighted[1].

Keyword :- *Emotion, Facial Expression,Speech Recognition,SVM*

1. INTRODUCTION

Automatic identification of emotions by facial expressions consists of three steps: face recognition, extraction and classification of features or hand movements, facial features, and voice sound that are used to convey emotions and input. Nonetheless, the latest developments of human user interfaces, which have progressed from traditional mouse and keyboard to automated speech recognition technologies to unique interfaces tailored for individuals with disabilities, do not take full account of these important interactive capabilities, sometimes contributing to less than normal experiences When machines were able to understand such emotional signals, they could provide users precise and effective support in ways that are more in line with the desires and expectations of the individual.

From psychological science it is generally agreed that human emotions may be divided into six archetypal feelings: shock, terror, disgust, rage, joy and sadness. Facial expression and voice sound play a critical role in communicating certain emotions.

The facial muscles can be changed and the sound and strength of speech development can be deliberately altered to express different feelings. Also when softly shown, human beings may perceive such signs by interpreting knowledge obtained by the ears and eyes at the same time. Centered on psychological experiments, which indicate that visual input changes speech processing[17], it can be concluded that experience of human emotions follows a similar pattern. Motivated by these hints, De Silva et al. performed trials allowing 18 participants to recognise emotions utilizing visual and auditory knowledge independently from an audiovisual archive obtained from two topics[7]. They found that certain feelings, such as depression and anxiety, are best described with audio, while others with film, such as rage while joy. However, Chen et al. showed that these two modalities offer complementary knowledge, claiming that device efficiency improved when both modalities were viewed together[4].

ISSN: 2233-7857 IJFGCN

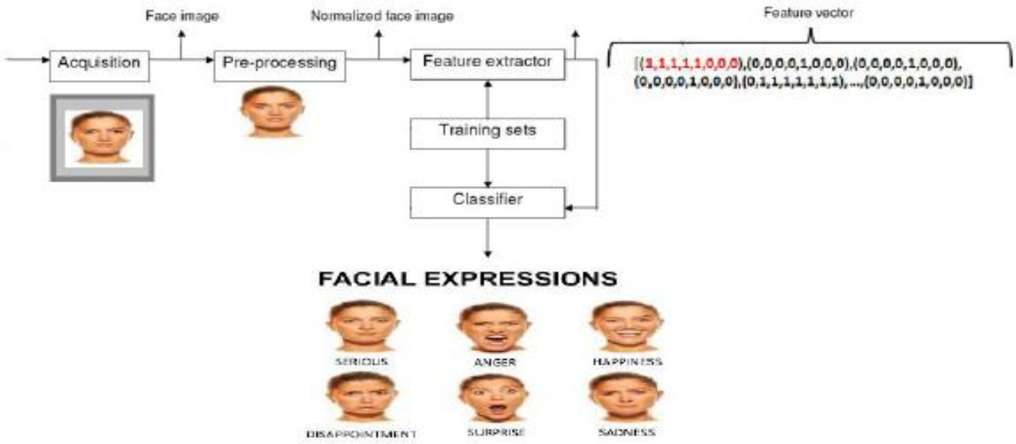

Figure 1. The Different Stages of Facial Expression Recognition (FER) System

Although numerous automated emotion detection programs have investigated the use of either facial expressions or speech to diagnose affective conditions in humans, comparatively few studies have concentrated on emotion identification utilizing all modalities[4][8]; It is expected that the multimodal solution would not only offer improved efficiency but also more robustness when obtaining either of these modalities in a noisy environment[19]. These earlier experiments were fused facial expressions and acoustic knowledge at a judgment point, Where unimodal device outputs are merged with the use of appropriate parameters or at a function level where results from all modalities are incorporated prior to classification. Nevertheless, none of these articles sought to evaluate which solution to fusion is more suitable for identification of emotions. This paper assesses these two strategies to fusion, in terms of overall device efficiency.

2. Emotion analysis

Humans convey multidimensional feelings. Any that forms are through printing, walking, facial expression, body language, movement, etc. Such emotions can be classified with different models of emotion. An effective model of emotion needs to be chosen to detect and interpret emotions from any information. It will describe the collection of emotions that are relevant to a specific question.

2.1 Evolution of emotion

In 1872, after having conducted some psychological studies on facial expressions observed on both humans and animals in various contexts, Charles Darwin stated that human and other animals displayed identical emotions

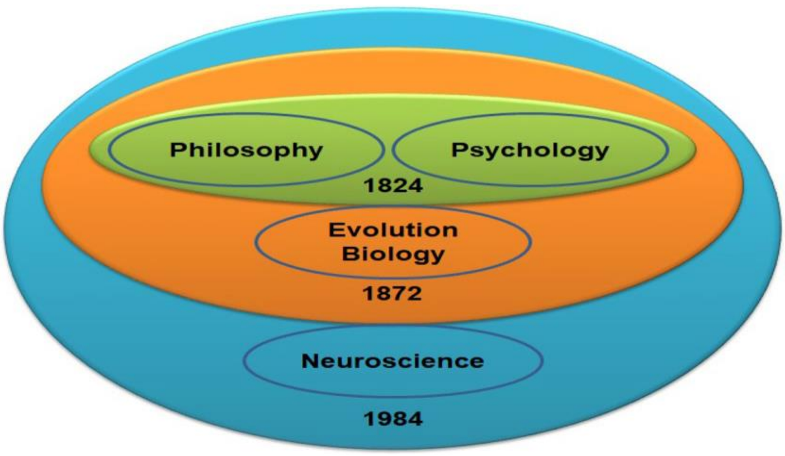

Fig.2 Evolution of emotion in various fields

ISSN: 2233-7857 IJFGCN

Related words and behaviors (Darwin 1998). His relational hypothesis was often linked to the age and situations connected with it. Emotions were learned over time in humans and other species according to his statement.

He explored in depth basic concepts of emotions; methods of communicating emotions in both humans and animals; triggers and consequences of all potential emotions such as fear, sorrow, dejection, depression, happiness, passion, commitment, etc.; description of emotions through pictures to display the manifestations of individual emotions. Darwin believed that all emotional emotions are common to people around the world. He also believed that animals of related types, as well as humans, respond similarly to a circumstance. His studies showed that even in organisms which are not quite related may have common expressions for certain emotions. Before that, as stated in Manser (1963) and Bell (1824), there was some metaphysical and theological categorisation of emotions.

2.2 Emotion models

From a psychological perspective, individual feelings can be defined and categorized depending on form of emotion, strength of emotion and several other factors, all of which can be incorporated and realized into models of emotion. Models of emotion are a organized method of describing specific human emotions according to certain ratings, rankings or measurements. Existing emotional models describe specific categories of emotions by length, strength, timing, pace of transition, concentration of incidents, assessment, behavioral effect (Borod 2000).

Established emotion structures may be classified into two groups – Categorical and Relational (Calvo and Kim 2013) based on various emotion theories. Categorical models of emotions describe a set of emotional types that are distinct from each other. In the other side, models of Dimensional Emotion describe certain measurements with some parameters and assign emotions according to certain measurements.

In most graphical representations of emotion two or three dimensions are used — 'valence' (indicates the positivity or negative of an emotion), 'arousal' (indicates the degree of anticipation of an emotion) and 'dominance' (indicates the extent of power over an emotion) (Sreeja and Mahalakshmi 2017; Canales and Martinez-Barco 2014).

2.3 Emotion detection

The detection of feelings from various forms of components of social networks is a research subject that has been studied for a long time now. Specific styles of material shared by individuals on social networking sites were examined to identify emotions behind the messages. For certain studies, some mixture of voice sound, speech, facial movements, actions, EEG patterns, multiple forms of bio patterns, and texts have been used to identify emotions from multimodal data (Poria et al. 2016; Kudiri et al. 2016; Soleymani et al. 2016). In the other side, much of the emotion recognition works depend in one particular category of material.

3. LITERATURE REVIEW

Rokne, J., & Alhajj, R. (2018) Analysis of feelings in several specific contexts may be beneficial. Another such area is cooperation with the human computers. Computers can make smarter choices to support people, with the aid of emotion detection. As robotic work becomes more common, emotion recognition would also help make human – robot interaction more normal. This study addresses current work results on emotion identification, emotion templates, emotion databases, emotion detection methods, their features, drawbacks and several potential paths for the future. We concentrate on analyzing the study activities to examine text- and speech-based emotions. We studied different feature sets which were used in current methodologies. We are summing up simple successes in the field and outlining potential extensions for a better outcome.

Wei, W.-J. (2019). Traditionally, instructors have measured the children's feeling according to evaluation. Use algorithmic methods we will be able to predict child's emotions. To meet this target, an emotional lexicon focused on the standardized measure entitled Emotional Competencies Scale for

ISSN: 2233-7857 IJFGCN

Young Children (ECSYC) needs to be developed and tested. This research had the aim of determining the validity relevant to the criteria. This study's approach was to build initially 40 ECSYC-based scenarios. Second, we had the five-level standards established. Third, this analysis introduced instruction of observers and measured reliability of accuracy between raters. Third, the observers grouped the responses of 200 children into five groups. Fifth, the analysis rated each level's frequency series and completed the emotional lexicon. The results revealed that the rho coefficient of the Spearman hit a high of .406*. (P = .026), which is important, shows a strong connection between Young Children's Relational Lexicon (YCEL) and ECSYC.

Wim van der Vegt & Wim Westera (2019) Recognition of feelings by facial expressions. In addition to enabling emotion detection from image files and captured video files, it utilizes camera data to include emotional facial expressions in real-time, continuous and unobtrusive manner. It uses FURIA algorithm for unordered insertion of fuzzy laws to provide timely and correct input based on the facial expressions of the learners. The key purpose of this research was first to verify the usage of camera data in elearning settings for a real-time and precise examination of facial expressions. Second, utilizing the FURIA method, convert certain facial expressions into detectable emotional states. We assessed the program output with ten people, presented them with the same computer-based activities, challenged them to imitate different facial movements a hundred times, and captured all of the sessions on camera. We used the video files which were captured to feed our newly developed apps.

Chaitali Chakrabarti and Andreas Spanias (2019) Historically, speaking methods of emotion recognition integrating articulatory details with acoustic features have been shown to enhance recognition efficiency. In certain cases, gathering articulatory data on a wide scale may not be possible, thereby restricting the complexity and applicability of such approaches. A discriminative method of learning for emotion recognition is suggested in this article, utilizing both articulatory and acoustic knowledge. A conventional 1-regularized logistic cost regression method is generalized to provide additional constraints to implement articulatory data reconstruction of the model. This contributes to incomplete and interpretable images, which are respectively designed for all roles. In addition, the model needs only articulatory features during training; only speech features are needed for inference on data from the out - of-samples. Experiments are conducted to evaluate the performance of emotional recognition over vowels /AA/, /AE/, /IY/, /UW/ and complete pronunciationsIt is shown that adding articulatory details significantly enhances the results for valence-based classification. Statistics obtained for categorical emotional identification within the corpus and cross-corpus suggest that the approach suggested is more successful in separating pleasure from certain emotions.

Li, H., & Wen, G. (2019). The new emotion classification model's actions to classify all research samples with the same approach undermines the real-world reasoning of humans, who actively adjust the approaches they use depending on current test samples. To resolve this paradox, this research suggests a system of individualized identification of emotions dependent on background understanding. A classifier that was considered most appropriate for the current test sample was first chosen from a collection of nominee classifiers for a specific test sample, and then used to realize the individualized emotion identification. The Bayesian learning approach was used to pick the optimal classifier and then test each candidate classifier from a global viewpoint to ensure each candidate classifier is optimal. Study findings confirmed the feasibility of the proposed system.

Alshamsi, H., & Meng, H. (2018). Because emotions are such an necessary and fundamental part of being human, experiencing them and learning how to respond to other people's emotions is a simple prerequisite for good social interaction. They mainly understand feelings through voice and facial expression. In academic work prompted by studies into emerging strategies such as recognizing emotions dependent on the meaning of expression, the subject is growing in significance. This examines the relationship between emotions and the content of our discourse. This paper explores how emotion can be detected in speech and facial expression in real time utilizing a platform composed of cloud-backed cell phone technologies. This feature was built and integrated into a smartphone device. The program currently works on any Android smartphone to sense and understand emotions in real time. The findings are presented as a percentage of all potential feelings, such as sorrow, joy, terror, shock, anger, etc. The experiment's findings show the emotion perception of face and voice was effectively

ISSN: 2233-7857 IJFGCN

performed using a smartphone. When used with regular business, it was right in 97.26 per cent of instances.

Simko, J., & Bielikova, M. (2016). The key concern of this paper is the identification of emotions by means of inexpensive low-end EEG sensors and their contrast with approaches to facial recognition. Next, we did a comparative review of two widely viable approaches utilizing identification of facial expression: Noldus FaceReader and Shore. Second, we suggested our own process of measuring emotions utilizing EEG sensors. We conducted an experiment with Emotiv Epoc, in which we played music videos for our participants to elicit their emotional reaction, Whilst the EEG signal was registered. Using machine learning, we used the data obtained to identify emotions and contrasted the findings with the recognition method for facial expressions to detect emotions, namely Noldus FaceReader, to see if these two methods work under the same conditions.

J., Kim, J., & Kwon, O. (2016). Over the past few years, deep learning has produced groundbreaking results in many fields of operation including speech recognition, perception of images, and so on. Instead of hand-crafted feature-based approaches, we seek to use deep learning strategies for real-time facial expression recognition. The proposed system can use a camera to recognize human emotions based on facial expressions. For TV set, it can detect faces and identify users at a distance of 2~3 m. And it can decide whether a person feels joy, grief, disappointment, rage, indignation, neutrality or any variation of those six emotions. The experimental results show high precision achieved by the proposed method. It can be used for various services including research on consumer behaviour, usability studies, psychology, educational research, and market research.

Yong Xiang and Jing He (2019) Typically people use various signs to communicate their feelings, such as facial features, movements of hands and speech. Facial gestures constitute up to 55% of human interaction and just 7% of emotional speech is assigned to other forms such as oral language. Considering facial expressions in an HRI framework thus allows for effective simulation of human interactions. In reality, since they can read facial expressions and work out their emotional states, robots can easily communicate with humans in as pleasant a way as possible. In this way they may be employed in a healthcare environment to identify emotional conditions in people by examining feelings and enhancing the quality of life. Throughout daily circumstances the emotional environments are unveiled where machines will examine optimistic and negative feelings. Good facial expressions, such as joy and enjoyment, display positive states of emotion whereas poor states of emotion are expressed by fetching negative facial expressions (e.g., sorrow and anger). An successful method of facial expression (FER) may greatly help individuals enhance their mental emotional health by examining their behaviour habits, demonstrating that certain psychiatric disorders such as schizophrenia or autism are identified by analyzing the interpersonal problems that surface on the faces of the patients.

Nan Song, Hongwu Yang (2018) This paper suggests an adaptive sign language facial gesture to translation process of expressive speech to address the communicating issues between healthy people and speech disabilities. Firstly, a deep neural network (DNN) model obtains the properties of sign language and the attributes of facial speech. Furthermore, a support vector machine (SVM) is equipped to identify the sign language and facial expression to interpret sign language text and facial expression emotional marks. At the very same time, with a Mandarin emotional speech corpus, a secret Markov model focused on Mandarin-Tibetan bilingual emotional speech synthesizer is taught by an adaptive voice training user. Eventually, the expressive Mandarin or Tibetan voice is synthesized from understood sign language text and expressive marks. Objective studies display a comprehension score of 90.7 per cent for standardized sign language. The facial expression detection score hits 94.6% on the expanded Cohn- Kanade database (CK+) and 80.3% on the JAFFE database, respectively.

Nancy Semwal, Abhijeet Kumar,(2017) Emotions that a speaker shows may be observed by observing or integrating his / her voice, facial expressions and movements. This article focuses on assessing the condition of emotion from voice signals. Different acoustic characteristics, such as strength, zero crossing rate (ZCR), simple frequency, Mel Frequency Cepstral Coefficients (MFCCs), etc., are extracted for short-term, overlapping frames obtained from the signal. A function vector is then created for each utterance by evaluating the global statistics (mean, median, etc) of the features derived from

ISSN: 2233-7857 IJFGCN

all objects. Sequential backward selection (SBS) approach is used with k-fold cross validation to choose a subset of usable functions from the complete candidate feature vector.

Dr. Sattar B. Sadkhan, Ahmed Dheyaa Radhi (2017) In this paper we will highlighting of Fuzzy logic that used to detect emotion subjects from textual data. Fuzzy Logic was founded to handling non-value member and convert it to a value member. We should study the history of fuzzy logic for the analysis of emotions and the methods used in this work. Emotions are a central part of human emotional behavior and play a fateful function in making judgments and structures of cognitive relationships including (sadness, shock, joy, indignation, rage, and fear). Emotions may be conveyed by means of deferential modes including words, facial expressions, movements and writing. Today the project operates in the area of emotion recognition and study of emotions. The researchers are seeking to build a method for classifying the emotions in language.

V.P. Kalyan Kumar, P. Suja and Shikha Tripathi (2016) Emotions are essential to knowing human behaviour. Several emotion detection modalities involve image, voice, facial expression, or movement. Emotion detection from video by facial expressions plays a crucial role in human machine interaction where the gestures of the facial features that reflect the conveyed emotion need to be recognised rapidly. In this study, we propose a novel method for acknowledging six specific emotions in BU-4DFE database 4D video sequences utilizing geometric approach. Of the 83 function points given in the BU-4DFE index, we've picked main facial points. A video that communicates emotion includes frames comprising the emotion's neutral, original, peak, and offset. We were immediately defining the apex frame from a video series. The Euclidean distance between the feature points in apex and neutral frame is defined and their difference is measured to shape the feature vector in the corresponding neutral and apex frame.

Rahim Rahmani, Antonio Moran Cardenas (2013) This paper focuses on a method for detecting the emotion of humans from the expression of a perceived person. The interpreted knowledge is transmitted in various facial expressions belonging to six fundamental specific facial emotions via the regions of the eye and mouth into a combined new picture. Throughout the sense of creating socially aware devices, the feedback knowledge collected may be fed as an input into a computer willing to communicate with social skills. The approach uses a grouping technique of knowledge into a new fused picture consisting of two blocks merged by the region of the eyes and mouth, very sensitive areas to alter the voice of the human being and particularly relevant to decoding emotional expressions. Finally we use the combined picture as an input to a neural feed-forward network trained by back-propagation. These processing of combined images makes it possible to gain relevant information by integrating proper data in the same picture and reducing the time required for training while maintaining the classification efficiency. Experimental results show that the algorithm proposed can detect emotions with good success

Haque Siddiqui and Ahmad Y. Javaid (2018) Extensive implementation possibilities have made emotion detection in the field of computer science ineluctable and difficult. Using non-verbal signals such as movements, body reaction, and facial expressions transmits the user's thinking and input. This Human – Computer Interaction training puts emphasis on the algorithmic robustness and flexibility of the sensor to enhance the identification. Sensors play an significant role in precise identification by providing a very high-quality feedback, thus growing device performance and reliability. Automatic understanding of human feelings will assist in educating the robots about emotional intelligence. This paper provides a short analysis of the various methods and the emotion identification strategies. The survey includes a concise analysis of the repositories known as data sets for algorithms that detect emotions via facial expressions

ISSN: 2233-7857 IJFGCN

4. PROPOSED WORK

Emotion Detection through Facial Expressions

Face detection in this case implies Face recognition in the specified data, if any. It is a difficult activity due to the following questions: the posture and perspective of the human face which differ according to the location of the frame, the goal human face can occlude or near to the other subjects and the issue of lighting weather. In fact, the human face is distinctive, so with noise (eye glasses, hair, so moustache), the same individual may appear different. To determine the efficiency of human face detection under real-time conditions, HCI depends on the following factors: speed of detection, precision of detection, appropriate training period and number of samples of testing. Ada-boost dependent face detector by Viola and Jones is used from the above parameters in this research work[20].

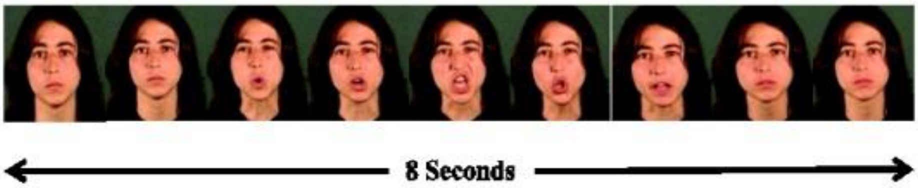

Figure 3: Video file of 8 seconds duration [3]

Figure 1 displays a video file that lasts 8 seconds. This video clip includes all facial and voice gestures. Let's remember the gestures of the faces in the video clip, it shifts every second. The person's emotion inside the video file varies from class to class. The shift depends on Plutchik's emotional loop condition diagram[21]. However, a proposed abstraction of the function is required to detect emotions from the input video file.

Algorithm for proposed Relative Sub-Image Based Features (RSB) as follows [22]:

Step 1. Read image and pre-processing.

Step 2. Divide input image into sub-images.

Step 3. Calculate average of each sub-image.

Step 4. Calculate relative difference between each sub-image to its edge cent sub-images.

The proposed function uses pixel intensities to interpret the data in real-life environments and is close to the human eye. It can accommodate feelings parallel to it too. Human speech is thus a combination of conscious and accidental facial gestures. Because of this, unimodal device output by facial expressions is declining. In the other side, human conversation is also a combination of both words and facial expressions. Thus, discourse modality is regarded along with facial expressions in this study work.

B. Emotion Detection through Speech

Algorithm for proposed RBFC [23] as follows:

Step 1. Read speech signal.

Step 2. Pre-emphasize the speech signal.

Step 3. Divide the signal into frame size of 20 ms and a shift of 10 ms, and multiple each frame with hamming window.

Step 4. Compute the frequency spectrum from the above.

Step 5. Calculate RBFC from details about the frequency domain. It helps research differences of the outer and intra topics. It also allows to cope with the issue of parallel emotions. It is an independent frequency function that assists not only with long input signals but also with short ones.

ISSN: 2233-7857 IJFGCN

C. Classification

RBF-kerneled SVM is used to monitor emotions[10]. Since not only recognized samples may be dealt, but also unknown samples. It also assists in the classification of input data with extremely high dimensional space. In addition, SVM output also depends on what kind of kernel the input database was referred to.

D. Hybrid Method

Hybrid is a synthesis of the fusion strategies at the function level and at the judgment stage. Human contact is still necessary in a social setting, by meaningful human dialogue. At this point, it is challenging in order to classify human emotions by means of deliberate and accidental facial gestures using a computer. To address the above, it should be possible to recognise human emotions by facial expressions and words. The next, often intermittent in nature are human emotions which are already mentioned in the above.

From the above, it is not often possible to recognise feelings at the same time from both modalities (speech and facial expressions). The association between the above listed modalities varies from time to time . The hybrid device proposed is required to combat the above problems. Figure 2 displays the block diagram for a hybrid framework being introduced. The proposed structure comprises four phases, in keeping with the diagram. Step 1 and 3 measure feeling here, respectively, by voice and facial expressions. These are the emotion recognition devices which are unimodal. Step 2 measures emotion recognition method (FF) of function standard by voice and facial expressions. Finally step 4 measures the final emotion by way of the device for the study of judgment point (AU).

In this case, s1, s2 and f1, respectively f2 displays input data for voice and facial expressions. The next one is x, z and y which indicates outputs of identification of unimodal emotions through words, facial expressions and FF respectively. I display the fused vector function level and will be an reference for the multimodal research device system stage. K demonstrates fused vector judgment standard. Z displays the performance of Multimodal Analysis Device judgment stage video input image. Z-1 depicts past emotional performance to predict present emotional production.

Figure 3 shows the emotional state diagram which allows the proposed method to estimate emotions at stage-4. The fear class that shift to shock or hostile classes according to Figure 3, but it is not possible for fear class to move specifically to sad or disappointed or angry or happy classes. The class of shock emotions can shift to classes of terror or sad or indifferent, but it is not possible to move explicitly to classes of disgust or rage or joy.

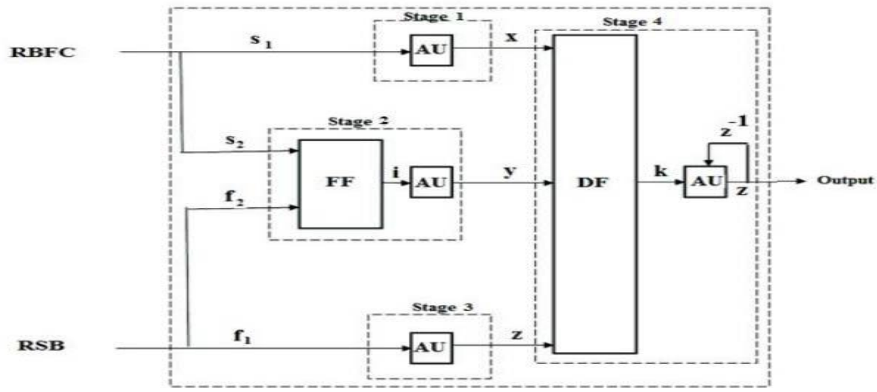

Figure 4: Hybrid system

ISSN: 2233-7857 IJFGCN

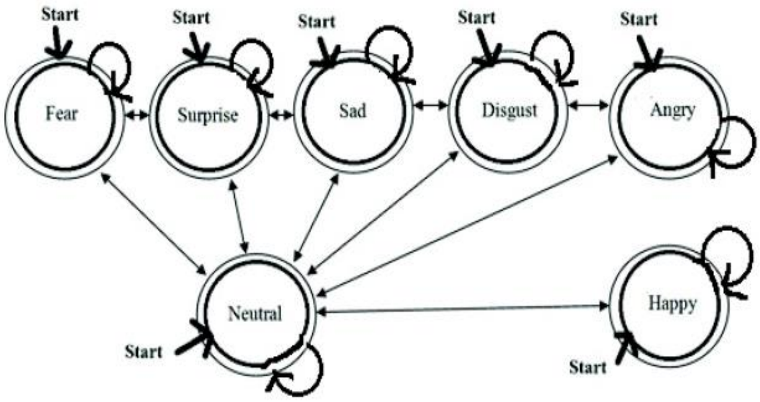

Figure 5: State diagram of emotions through Plutchik Wheel [21]

E. Database

The Facial Expressions Database (DaFEx) is an act one database, produced by Italian trained actors (4 males and 4 females) [3].

5. Conclusion

The systematic research on the impact on the proposed program of different factors such as deliberate and accidental facial gestures has been undertaken. Undeliberate facial gestures have been shown to have a noticeable impact in real-life situations. All in both, all of the work's goals is met. The results of this research may be summed up as follows: Unlike other traditional functions, RBFC and RSB can tackle the issue of neighboring emotions without any issue here. Hybrid system produced positive results when fusing the modalities above. And it can also tackle the issue of data synchronisation without any difficulty. While certain factors influence the performance of the proposed method, the impact is comparatively smaller. The above experiments were performed successfully with the assistance of Matlab tools. With the support of libSVM tool[24], classification was performed using SVM. Potential research needs to concentrate on more intense classroom figures. And the machine will fully comprehend the customer. Furthermore, it may be useful if the device may still grasp the context of the speech signal.

References

- Sailunaz, K., Dhaliwal, M., Rokne, J., & Alhajj, R. (2018). Emotion detection from text and speech: a survey. Social Network Analysis and Mining, 8(1). doi:10.1007/s13278-018-0505-2.
- Wei, W.-J. (2019). *Development and evaluation of an emotional lexicon system for young children. Microsystem Technologies.* doi:10.1007/s00542-019-04425-z
- Kiavash Bahreini & Wim van der Vegt & Wim Westera,2019 "A fuzzy logic approach to reliable real-time recognition of facial emotions" Multimedia Tools and Applications https://doi.org/10.1007/s11042-019-7250-z.
- Chaitali Chakrabarti and Andreas Spanias (2019) "Articulation constrained learning with application to speech emotion recognition", Shah et al. EURASIP Journal on Audio, Speech, and Music Processing (2019) 2019:14 https://doi.org/10.1186/s13636-019-0157-9.
- Li, H., & Wen, G. (2019). *Sample awareness-based personalized facial expression recognition. Applied Intelligence.* doi:10.1007/s10489-019-01427-2
- Alshamsi, H., Kepuska, V., Alshamsi, H., & Meng, H. (2018). Automated Facial Expression and Speech Emotion Recognition App Development on Smart Phones using Cloud Computing. 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON). doi:10.1109/iemcon.2018.8614831.
- Matlovic, T., Gaspar, P., Moro, R., Simko, J., & Bielikova, M. (2016). Emotions detection using facial expressions recognition and EEG. 2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP). doi:10.1109/smap.2016.7753378.

ISSN: 2233-7857 IJFGCN

- Lee, I., Jung, H., Ahn, C. H., Seo, J., Kim, J., & Kwon, O. (2016). Real-time personalized facial expression recognition system based on deep learning. 2016 IEEE International Conference on Consumer Electronics (ICCE). doi:10.1109/icce.2016.7430609.
- Chi-Hung Chi , Yong Xiang and Jing HeA (2019) Review on Automatic Facial Expression Recognition Systems Assisted by Multimodal Sensor Data, Sensors 2019, 19, 1863; doi:10.3390/s19081863 www.mdpi.com/journal/sensors.
- Nan Song , Hongwu Yang (2018) "A Gesture-to-Emotional Speech Conversion by Combining Gesture Recognition and Facial Expression Recognition" , 2018 First Asian Conference on Affective Computing and Intelligent Interaction (ACII Asia).
- Nancy Semwal, Abhijeet Kumar,(2017), "Automatic Speech Emotion Detection System using Multi-domain Acoustic Feature Selection and Classification Models", IEEE Transactions on Multimedia, Vol. 10, No. 5, pp. 936-946, August 2008.
- V.P. Kalyan Kumar, P. Suja and Shikha Tripathi (2016), "Emotion Recognition from Facial Expressions for 4D Videos Using Geometric Approach", Springer International Publishing Switzerland 2016 3 Advances in Signal Processing and Intelligent Recognition Systems, Advances in Intelligent Systems and Computing 425, DOI: 10.1007/978-3-319-28658-7_1.
- Rahim Rahmani, Antonio Moran Cardenas (2013) "Automatic emotion recognition through facial expression analysis in merged images based on an Artificial Neural Network", 2013 12th Mexican International Conference on Artificial Intelligence.
- Haque Siddiqui and Ahmad Y. Javaid (2018) "Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality", Sensors 2018, 18, 416; doi:10.3390/s18020416 www.mdpi.com/journal/sensors.
- Black, M. J. and Yacoob, Y. Tracking and recognizing rigid and non-rigid facial motions using local parametric model of image motion. In Proceedings of the International Conference on Computer Vision, pages 374–381. IEEE Computer Society, Cambridge, MA, 1995.
- Boersma, P., Weenink, D., Praat Speech Processing Software, Institute of Phonetics Sciences of the University of Amsterdam. http://www.praat.org
- Burges, C. A tutorial on support vector machines for pattern recognition. Dat Mining and Know. Disc., vol. 2(2), pp. 1– 47, 1998.
- Chen, L.S., Huang, T. S., Miyasato T., and Nakatsu R. Multimodal human emotion / expression recognition, in Proc. of Int. Conf. on Automatic Face and Gesture Recognition, (Nara, Japan), IEEE Computer Soc., April 1998
- Chen, L.S., Huang, T.S. Emotional expressions in audiovisual human computer interaction. Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on, Volume: 1, 30 July-2 Aug. 2000. Pages: 423 - 426 vol.1
- Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., Taylor, J.G. Emotion recognition in human-computer interaction. Signal Processing Magazine, IEEE, Volume: 18, Issue: 1, Jan 2001. Pages: 32 – 80
- De Silva, L. C., Miyasato, T., and Nakatsu, R. Facial Emotion Recognition Using Multimodal Information. In Proc. IEEE Int. Conf. on Information, Communications and Signal Processing (ICICS'97), Singapore, pp. 397-401, Sept. 1997.
- De Silva, L.C., Ng, P. C. Bimodal emotion recognition. Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, 28- 30 March 2000. Pages: 332 – 335.
- Dellaert, F., Polzin, T., Waibel, A. Recognizing emotion in speech. Spoken Language, 1996. ICSLP 96. Proceedings. Fourth International Conference on, Volume: 3, 3-6 Oct. 1996. Pages: 1970 - 1973 vol.3.
- Ekman, P., Friesen, W. V. Facial Action Coding System: A Technique for Measurement of Facial Movement. Consulting Psychologists Press Palo Alto, California, 1978.
- Essa, Pentland, A. P. Coding, analysis, interpretation, and recognition of facial expressions. IEEE Transc. On Pattern Analysis and Machine Intelligence, 19(7):757–763, JULY 1997.

ISSN: 2233-7857 IJFGCN

- Kudiri KM, Said AM, Nayan MY (2016) Human emotion detection through speech and facial expressions. In: IEEE international conference on computer and information sciences (ICCOINS), pp 351–356
- Varghese, A. A., Cherian, J. P., & Kizhakkethottam, J. J. (2015). *Overview on emotion recognition system. 2015 International Conference on Soft-Computing and Networks Security (ICSNS).* doi:10.1109/icsns.2015.7292443
- Sadkhan, S. B., & Radhi, A. D. (2017). Fuzzy Logic used in Textual Emotion Detection. 2017 Second Al-Sadiq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA). doi:10.1109/aic-mitcsa.2017.8722967
- Semwal, N., Kumar, A., & Narayanan, S. (2017). *Automatic speech emotion detection system using multi-domain acoustic feature selection and classification models. 2017 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA).* doi:10.1109/isba.2017.7947681
- Patil, A., & Behele, K. (2017). Classification of Human Emotions Using Multiclass Support Vector Machine. 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA). doi:10.1109/iccubea.2017.8463656
- Fang, C.-Y., Ma, C.-W., Chiang, M.-L., & Chen, S.-W. (2017). An infant emotion recognition system using visual and audio information. 2017 4th International Conference on Industrial Engineering and Applications (ICIEA). doi:10.1109/iea.2017.7939223

ISSN: 2233-7857 IJFGCN