Emotion Detection using Machine Learning

[1]Vijayanand. G
Asst. Prof/CSE
Muthyammal Engineering College,
Rasipuram, Tamilnadu.

[2]Karthick. S
IV Year/CSE
Muthyammal Engineering College,
Rasipuram, Tamilnadu.

[3]Hari. B
IV Year/CSE
Muthyammal Engineering College,
Rasipuram, Tamilnadu.

[4]Jaikrishnan. V
IV Year/CSE
Muthyammal Engineering College,
Rasipuram, Tamilnadu.

Abstract: **A human-computer interaction system for an automatic face recognition or facial expression recognition has attracted increasing attention from researchers in psychology, computer science, linguistics, neuroscience, and related disciplines. In this paper, an Automatic Facial Expression Recognition System (AFERS) has been proposed. The proposed method has three stages: (a) face detection, (b) feature extraction and (c) facial expression recognition. The first phase of face detection involves skin color detection using YCbCr color model, lighting compensation for getting uniformity on face and morphological operations for retaining the required face portion. The output of the first phase is used for extracting facial features like eyes, nose, and mouth using AAM (Active Appearance Model) method. The third stage, automatic facial expression recognition, involves simple.**

Euclidean Distance method. In this method, the Euclidean distance between the feature points of the training images and that of the query image is compared. Based on minimum Euclidean distance, output image expression is decided. True recognition rate for this method is around 90% - 95%. Further modification of this method is done using Artificial Neuro-Fuzzy Inference System (ANFIS). This non-linear recognition system gives recognition rate of around 100% which is acceptable compared to other methods.

Keywords— Facial expression recognition (FER), multimodal sensor data, emotional expression recognition, spontaneous expression, real-world conditions

I.   INTRODUCTION

Facial expression recognition (FER) has been dramatically developed in recent years, thanks to the advancements in related fields, especially machine learning, image processing and human cognition. Accordingly, the impact and potential usage of automatic FER have been growing in a wide range of applications, including human-computer interaction, robot control and driver state surveillance. However, to date, robust recognition of facial expressions from images and videos is still a challenging task (1)due to the difficulty in accurately extracting the useful emotional features. These features are often represented in different forms, such as static, dynamic, point-based geometric or region-based appearance. acial movement features, which include feature position and shape changes, are generally caused by the movements of facial elements and

muscles during the course of emotional expression. The facial elements, especially key elements, will constantly change their positions when subjects are expressing emotions. As a consequence, the same feature in different images usually has different positions. In some cases, the shape of the feature may also be distorted due to the subtle facial muscle movements. For example, the mouth in the first two images in presents different shapes from that in the third image. Therefore, for any feature representing a certain emotion, the geometric-based position and appearance-based shape normally changes from one image to another image in image databases, as well as in videos. This kind of movement features represents a rich pool of both static and dynamic characteristics of expressions, which play a critical role for FER.

The vast majority of the past work on FER does not take the dynamics of facial expressions into account. Some efforts have been made on capturing and utilizing facial movement features, and almost all of them are video- based. These efforts try to adopt either geometric features of the tracked facial points (e.g. shape vectors, facial animation parameters, distance and angular, and trajectories), or appearance difference between holistic facial regions in consequent frames (e.g. optical flow, and differential-AAM), or texture and motion changes in local facial regions (e.g. surface deformation, motion units, spatiotemporal descriptors, animation units, and pixel difference). Although achieved promising results, these approaches often require accurate location and tracking of facial points, which remains problem.

II.   LITERATURE REVIEW

Understanding emotional facial expressions accurately is one of the determinants in the quality of interpersonal relationships. The more one reads another's emotions correctly, the more one is included to such interactions. The problems in social interactions are shown in some psychopathological disorders may be partly related to difficulties in the recognition of facial expressions(2). Such deficits   have been demonstrated in various clinical populations. Nonetheless, with respect to facial expressions, there have been discrepant findings of the studies so far. The

Special Issue - 2020

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCICCT - 2020 Conference Proceedings

purpose of this article is to review the topic of emotion(3), emotional facial expressions since ancient ages, to emphasize the strengths and weaknesses of the related studies, to compare their results and to pay attention to this novel issue for Turkey.

In 1884, William James proposed the first important physiological theory of emotion. James argued that emotion is rooted in the bodily experience. According to him, first, we perceive the object then bodily response occurs and lastly emotional arousal appears (Kowalski and Westen 2005 p347). For instance, when

we see a stimulus such as a bear, we have a ponding heart, we begin to run and than we fear. We do not run because of fear, we fear because of running. When his Danish colleague Carl Lange independently proposed a similar view in 1885, since then this theory has been known as James-Lange theory of emotions (Kowalski

and Westen 2005 p348, Candland et al 1977 p87). Walter B. Cannon (1927-1931) proposed an alternative theory suggesting that emotions are cognitive rather than physiological state of arousal. He perceived the sequence of events as external stimulation followed by neural processing followed by physiological

reactions. Philip Bard expanded Cannon's theory by showing the thalamic structures for the expression of emotion; this general theoretical position came to be referred to as the "Cannon-Bard Theory". This novel theory included that emotion-inducing stimuli simultaneously elicit both an emotional experience, such as fear, and bodily responses such as sweating (Candland et al. 1977 p87-88, Kowalski and Westen 2005 p348).

The study investigated the recognition of standardized facial expressions of emotion (anger, fear, disgust, happiness, sadness, surprise) at a perceptual level (experiment 1) and at a semantic level (experiments 2 and 3) in children with autism (N = 20) and normally developing children (N = 20). Results revealed that children with autism were as able as controls to recognize all six emotions with different intensity levels, and that they made the same type of errors(4).

These negative findings are discussed in relation to (1) previous data showing specific impairment in autism in recognizing the belief-based expression of surprise, (2) previous data showing specific impairment in autism in recognizing fear, and (3) the convergence of findings that individuals with autism, like patients with amygdala damage, pass a basic emotions recognition test but fail to recognize more complex stimuli involving the perception of faces or part of faces.

Since Kanner's (1943) original clinical account of children with autism first described their profound lack of affective contact with other people, psychologists have been evaluating the social and affective impairments in autism. The empirical research on affective impairment of children and adults with autism is wide and varied so that it is not surprising that the findings are extremely mixed. Hypotheses of a general affective deficit (Hobson, 1986a; 1986b; Hobson

et al., 1988), and a selective emotion recognition deficit (Baron-Cohen et al., 1999; Howard et al., 2000) have been explored. In addition, the theory of mind (ToM) deficit account of autism allowed investigations of selective emotion processing impairment by contrasting recognition tasks that do and do not necessitate the ability to represent mental states (Baron-Cohen et al., 1993). The present investigations attempt to replicate and extend these findings with children with autism.

We apply a biologically inspired model of visual object recognition to the multiclass object categorization problem.Our model modifies that of Serre, Wolf, and Poggio. As in that work, we first apply Gabor filters at all positions and scales; feature complexity and position/scale invariance are then built up by alternating template matching and max pooling operations. We refine the approach in several biologically plausible ways, using simple versions of sparsification and lateral inhibition. We demonstrate the value of retaining some position and scale information above the intermediate feature level. Using feature selection we arrive at a model that performs better with fewer features. Our final model is tested on the Caltech 101 object categories and the UIUC car localization task, in both cases achieving state-of-the-art performance. The results strengthen the case for using this class of model in computer vision.

The problem of recognizing multiple object classes in natural images has proven to be a difficult challenge for computer vision. Given the vastly superior performance of human vision on this task, it is reasonable to look to biology for inspiration. In fact, recent work by Serre,Wolf, and Poggio has shown that a computational model based on our knowledge of visual cortex can be competitive with the best existing computer vision systems on some of the standard recognition datasets. Our paper builds on their approach by incorporating some additional biologically-motivated properties, including sparsification of features, lateral inhibition, and feature localization. We show that these modifications further improve recognition performance, strengthening our understanding of the computational constraints facing both biological and computer vision systems

III. ANALYSIS AND DESIGN OF THE APPLICATION

A. EXISTING WORK

AFERS has three main steps

1. To detect a face from a given input image or video,

2. Extract facial features such as eyes, nose, and mouth from the detected face

3. Divide facial expressions into different categories such as happiness, anger, sadness, fear, disgust and surprise. Face detection is a special case of object detection. It also involves illumination compensation algorithms and morphological operations to maintain the face of the input image.

B. Drawbacks:

The system plays a communicative role in interpersonal relations because they can reveal the affective state, cumulative activity, personality, intention and psychological state of a person. The proposed system consists of three modules. The face detection module is based on image segmentation technique where the given image is converted into a binary image and further used for face detection.

C. Proposed Work:

To improve the recognition rate of the system, further modification in the third phase is done using Artificial Neuro-Fuzzy Inference System (ANFIS). In this method, the static images as well as video input can be given for testing the expressions. Here, neuro-fuzzy based automatic facial expression recognition system to recognize the human facial expressions like happy, fear, sad, angry, disgust and surprise has been proposed. Initially a video showing different expressions is framed into different images. Then the sequence of selected images is stored in a database folder. Using AAM method, the features of all the images are located & stored in the form of .ASF files. Then a mean shape is created for all the images in data folder. The change in the AAM shape model according to the change in facial expressions measures the distance or the difference (6)between Neutral and other facial expressions. These values are stored in a .MA T file & a specific value is assigned for each individual expression for training the ANFIS. These difference values are then given as input to the ANFIS (Artificial Neuro-Fuzzy Inference System). Using the ANFIS tool available in Mat lab, the system is trained for the different images and their video input sequences for different expressions.

D. Advantages:

One advantage of using these color spaces is that most video media are already encoded using these color spaces. Transforming from RGB into any of these spaces is a straight forward linear transformation
1. Face detection,
2. Feature extraction and
3. Facial expression recognition. The first phase of face detection involves skin color detection using YCbCr color model, lighting compensation for getting uniformity on face and morphological operations for retaining the required face portion

IV. SYSTEM IMPLEMENTATION

A. Skin Color Segmentation: For skin color segmentation, first we contrast the image. Then we perform skin color segmentation.

B. Face Detection: For face detection, first we convert binary image from RGB image. For converting binary image, we calculate the average value of RGB for each pixel and if the average value is below than 110, we replace it by black pixel and otherwise we replace it by white pixel. By this method, we get a binary image from RGB image

C. Eyes Detection: For eyes detection, we convert the RGB face to the binary face. Now, we consider the face width by W. We scan from the W/4 to (W-W/4) to find the middle position of the two eyes. The highest white continuous pixel along the height between the ranges is the middle position of the two eyes.

D. Apply Bezier Curve: In the lip box and the lip and may be some part of nose. So, around the box there is skin color or the skin. So, we convert the skin pixel to white pixel and other pixel as black. We also find those pixels which are similar to skin pixels and convert them to white pixel. Here, if two pixels RGB values difference is less than or equal 10, then we called them similar pixel. Here, we use histogram for finding the distance between the lower average RGB value and higher average RGB value.

E. Database and Training: In our database, there are two tables. One table "Person" is for storing the name of people and their index of 4 kinds of emotion which are stored in other table "Position". In the "Position" table, for each index, there are 6 control points for lip Bezier curve, 6 control points for left eye Bezier curve, 6 control points for right eye Bezier curve, lip height and width, left eye height and width and right eye height and width. So, by this method, the program learns the emotion of the people.

F. Emotion Detection: For emotion detection of an image, we have to find the Bezier curve of the lip, left eye and right eye. Then we convert each width of the Bezier curve to 100 and height according to its width. If the person's emotion information is available in the database, then the program will match which emotion's height is nearest the current height and the program will give the nearest emotion as output.

VII. CONCLUSION & FUTURE ENHANCEMENT

This paper had discussed about the efforts of the different researchers, with the effort made to include as many references as possible from recent years. Based on reviews, the paper had out some of the issues a raised towards facial expression recognition, using different techniques for face detection, feature extraction, analysis and classification methods. The paper gives detailed information about existing techniques in all the stages of Facial Expression Recognition FERs. The paper is very useful to both old and upcoming researchers in the field of FER, it presents detail information about existing techniques in all stages of that field to reinforcement their understanding of current trends and assist their future research prospects and directions. Further, the paper discussed about various techniques of their technology, merits and demerits which improves the performance of Facial Expression Recognition in image processing.

VIII. REFERENCES

1. Eldar, C. Yonina, "Compressed sensing: theory and applications," Cambridge University, 2012.
2. Solomon, Chris, Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. John Wiley & Sons, 2011.
3. Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." In bmvc, vol. 1, no. 3, p. 6. 2015.
4. Grafsgaard, Joseph, Joseph B. Wiggins. "Automatically recognizing facial expression: Predicting engagement and frustration." In Educational Data Mining 2013.

Special Issue - 2020

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
NCICCT - 2020 Conference Proceedings

5. Moridis, Christos N., and Anastasios. "Affective learning: Empathetic agents with emotional facial and tone of voice expressions." IEEE Transactions on Affective Computing, vol. 3, pp. 260-272,2012.
6. Brodny, Grzegorz, AgataKołakowska, Agnieszka Landowska, MariuszSzwoch, WioletaSzwoch, and Michał R. Wróbel. "Comparison of selected off-the-shelf solutions for emotion recognition based on facial expressions." In 2016 9th ICHSI, IEEE. pp. 397-404, 2016.
7. H. Ding, S.K. Zhou, "Facenet2expnet: Regularizing a deep face recognition net for expression recognition," In Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference pp. 118-126,2017.
8. Y. Wu, T. Hassner, K. Kim, "Facial landmark detection with tweaked convolutional neural networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.