Various Emotion Detection From Human Face Using Artificial Neural Network (ANN)

[1]Siya C Sover, [2] Beena M V

Computer Science and Engineering, Vidya Academy of Science and Technology, Thalakkottukara, India

***Abstract* - Humans can convey a lot of information through face. Facial expression is a natural and powerful channel for communicating emotions in human computer interaction system. Recognizing the facial expressions is challenging due to the frequent head movements and uncertain facial motion measurements. Current research in facial expression recognition is limited to posed expressions and often in frontal view. People can easily identify emotions from face. Using Computer-based systems the ability to respond to human nonverbal communication has proven to be difficult. This system automatically recognizes the emotions on a face like happy, sad, anger. Face images are given as input to the system. First, detected the face from input image. After the face detection, feature extraction method is used to extract the set of selected feature points. Finally extracted features are given as input to the neural network to recognize the emotions.**

***Index Terms* - Neural network,Bezier curve.**

I. INTRODUCTION

Human will express emotions through communications in day to day communications. An emotion is subjective and it includes actions, reactions and reflections. Emotions are mainly following on the face, in arm and body gestures, in the voice, to express our reactions. Researches shown that the better expressive way is facial expressions are used to give the emotions. Emotions are the fundamental requirement of the computer interaction systems [1]. Recent days, there has been an increasing affection in improving all aspects of interaction between humans and computers especially in the area of human emotion recognition by observing facial expressions. Computers and other powerful electronic devices contain important roles in our daily life, with their use of continuously being improved by user-friendly interfaces. Artificial Intelligence has long committed on the area of emotion from face. Novel advance in this area is the applicability of facial emotion recognition to areas like chat room avatars and video surveillance. The skill to recognize emotions can be relevant in face recognition applications . In this project classified the emotional expressions are happy, sad, anger.

Selection of appropriate features is critical in facial expression analysis. There are various approaches to feature extraction in emotion detection are Gabor filter, Haar-like features, Scale invariant feature transform and Local Binary Patterns While all these features are powerful for describing local appearances. But this system need only the pixel values of local features of face so that bezier curve is used to extract features. Consider the figure given below, these feature points are can extract from face [9]. In this system consider only the feature points of left an eye, right an eye, lip.

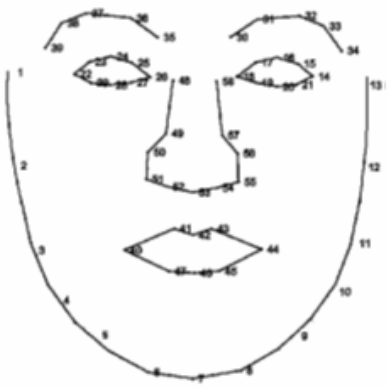


Figure 1: facial landmark (58 points)

There are also different approaches to classifications in emotion detection are Support vector machine classifier, fuzzy classifier, Gaussian based classifiers. These containing limitations are in performance, speed and size. So that classifications based on Artificial neural networks (ANNs) are preferred [2]. Because, it is easy to use and understand compared to statistical methods. Back propagation learning algorithms is used for error detection of neural networks.

II. Related Work

In this section briefly describe the existing works of emotion detection system.

Littlewort et al. [3] presented The Computer Expression Recognition Toolbox (CERT) , which is a software tool for fully automatic real-time facial expression recognition. CERT provides automatically code the intensity of 19 different facial actions from FACS and 6 different facial expressions. However some of these approaches give promising recognition rates on emotions/action units they do not fully exploit the connections among action units and emotions provided in facial action coding system as well as they are pure discriminative classifiers. Action unit detection and classification of facial expression in terms of a number of discrete emotion categories are referred from [4].

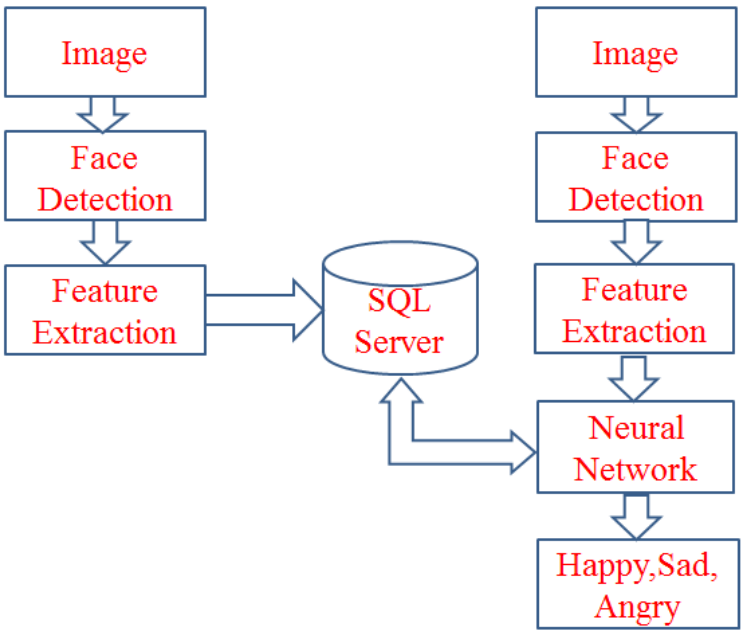
In [5] , preprocessing steps comprises of operations like images scaling, image contrast adjustment and other image enhancement operation. In [9], A skin detector is transforming a given pixel into an appropriate color space and then uses a skin classifier to label the pixel whether it is a skin or a non-skin pixel. A skin classifier defines a decision bound.

In[6],Artificial neural networks (ANNs) more suitable for tackling the problem of emotion recognition from action units as such techniques emulate human unconscious problems solving processes in contrast to rule based techniques, which are inspired by human conscious problems solving processes.

Porawat Visutsak [7] "Emotion Detection Using Sub-image Based Features through Human Facial Expressions" in 2013 The human face is an important human body part which plays an extraordinary role in the human to human or human to machine communications. it is important to design robust emotion detection systems for real world applications like human decision making and effective human computer.interaction environments. Facial expression gives the non-verbal communication for human computer interactions. Here the problem is the limited number of positions of facial muscles. This system based on relative sub-image based features. Classifications have been done using the support vector machine to implement an automated emotion detection system for facial expressions.

Christine L. Lisetti[8]presented "Facial Expression Recognition Using a Neural Network" which describes the a neural network for facial expression recognition. It aims at recognizing and interpreting facial expressions in terms of emotions and level of expressiveness. Use the backpropagation algorithm to train the system to differentiate between facial expressions.

III. Method



Skin Color Maintenance

For skin color segmentation, firstly the image is contrasted. Then skin color segmentation is performed. After that, find the largest connected region. Then check the probability to become a face of the biggest connected region. If the biggest connected region has the probability to become a face, then that biggest connected region open in new form. If the biggest connected regions height and width is grater or equal than 50 and the ratio of height/width is the range between 1 to 2, then it become face.

Face detection

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

Firstly, find out the forehead from the binary image. Then start scanning from the middle position of the input image, after that find out the continuous white pixels after a continuous black pixel. Next, maximum width of the white pixel is to be found out by searching vertical both left and right side. After that, if the new width is smaller half of the previous maximum width, then break the scan. Because if reach the eyebrow then this situation will arise. Then cut the face from the starting position of the forehead and its high must be 1.5 multiply of its width. That image contains only eyes, nose and lip. Then will cut the RGB image according to the binary image.

Eye detection

For eyes detection, convert the RGB face to the binary face. W is taken as width. Scan from the W/4 to (W-W/4) to find the middle position of the two eyes from the image. Then the highest white continuous pixel along the height can be seen in the middle position of the two eyes. Then find the high or upper position of the two eyebrows by searching vertical.

For left eye, search w/8 to mid and for right eye we search mid to w – w/8. Here w is the width of the image and mid is the middle position of the two eyes. There may be some white pixels between the eyebrow and the eye. To make the eyebrow and eye connected, place some continuous black pixels vertically to the eye from eyebrow. For left eye, the vertical black pixel-lines are placed in between mid/2 to mid/4 and for right eye the lines are in between mid+(w-mid)/ 4 to mid+3*(w-mid)/ 4 and height of the black pixel-lines are from the eyebrow starting height to (h- eyebrow starting position)/4. In this case w is the width of the image and mid is the middle position of the two eyes and h is the height of the image. Then find the lower position of the two eyes by searching black pixel vertically. For left eye, search from the mid/4 to mid - mid/4 width. For right eye, search mid + (w-mid)/ 4 to mid+3*(w- mid)/ 4 width from image lower end to starting position of the eyebrow.

Then find the right side of the left eye by searching black pixel horizontally from the mid position to the starting position of black pixels in between the upper position and lower position of the left eye. After that, left side for right eye search mid to the starting position of black pixels in between the upper position and lower position of right eye. The left side of the left eye is the starting width of the image and the right side of the right eye is the ending width of the image. Then cut the upper position of two eyes, lower position of two eyes, left side and the right side of the two eyes from the RGB image.

Lip detection

For lip detection, determine the lip box and consider that lip must be inside the lip box. So, first determine the distance between the forehead and eyes. Then add the distance with the lower height of the eye to determine the upper height of the box which will contain the lip part. Then draw the lip box, starting point of the box will be the ¼ position of the left eye box and ending point will be the ¾ position of the right eye box. Height of the box is lower end of the face image. So, this box will contain only lip part and some nose part also. Then will cut the RGB image according the box. For the detection of eyes and lip, only need to convert binary image from RGB image and some searching among the binary image.

Bezier curve on lip

In the lip box, there is lip and may be some part of nose. So, around the box there is skin color or the skin. So, 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 called them similar pixel. Here, use histogram for finding the distance between the lower average RGB value and higher average RGB value. If the distance is less than 70, then use 7 for finding similar pixel and if the distance is greater than or equal 70 then we use 10 for finding similar pixel. So, the value for finding similar pixel depends on the quality of the image. If the image quality is high, use 7 for finding similar pixel and if the image quality is low, use 10.

So, in the binary image, there are black regions on lip, nose and may some other little part which have a little different than skin color. Then apply big connected region for finding the black region which contain lip in binary image. And are sure that the big connected region is the lip because in the lip box, lip is the largest thing which is different than skin.

Then have to apply Bezier curve on the binary lip. For apply Bezier curve, find the starting and ending pixel of the lip in horizontal. Then draw two tangents on upper lip from the starting and ending pixel and also find two points on the tangent which is not the part of the lip. For the lower lip, find two point similar process of the upper lip. Cubic Bezier curves for draw the Bezier curve of the lip.

Apply bezier curve on eye

For apply bezier curve on eyes, first have to remove eyebrow from eye. For remove eyebrow, search 1stcontinuous black pixel then continuous white pixel and then continuous black pixel from the binary image of the eye box. Then remove the 1st continuous black pixel from the box and then get the box which only contains the eye.

Now, the eye box which contains only two eyes and has some skin or skin color around the box. So, apply similar skin color like the lip for finding the region of eye. Then apply big connect for finding the highest connected region and this is the eye because in the eye box.Eye is the largest thing in the eye box, which is not similar to the skin color. Then apply the bezier curve on the eye box, similar to the lip. Then get the shape of the eye.

Database

In database, there are two tables. One table "Person" is for storing the name of people and their index of 3 kinds of emotion which are stored in other table "Position". In the "Position" table contains the 6 control points for lip , 6 control points for left eye , 6 control points for right eye . Trained the network on 30 images and tested the network, Collected 104 images. Selected 30 images for training are shown the figure given below.

Neural network

Neural network structure is single layer that consists of input layer ,hidden layer, output layer. Input to the neural network are scalar values of left eye, right eye, lip and emotion. Then the equation $a_j = \Sigma W_{ij} I_i$ is used to reach hidden layer from input layer. After that start from hidden layer. Activation function is present in the hidden layer ,here binary sigmoidal functions are used in this system. Movement from hidden layer to output layer is done by using the equation $\Sigma W_{ij} Z_j$, $Z_j = f(a_j)$. $f(a_j)$ is called activation function. Equation of the binary sigmoidal functions is : $f(x) = 1/(1+\exp.\lambda x)$.

Output layer also contains the activation function that is used to identify whether the output is desired output or not. If the output is not desired, calculate the error using the equation is given as $\delta_k = y_k - t_k$. y_k is the desired output and t_k is the target output. This equation is used to find out the error of output layer to hidden layer. After that, find out the error of the hidden layer to input layer using the equation is $\delta_k = f'(x) \Sigma W_{kj} \delta_k . f'(x)$ is the derivative of activation function. This calculated error is used to update the weight of each node using this equation. W_{ij} =learning rate * δ_j*input[i], W_{jk} =learning rate * δ_k* Z_j, W_{ij} is the updated weight of input layer to hidden layer W_{jk} is the updated weight of input layer to hidden layer.Use this updated weight to find out the output and check whether the output is desired or not. If it is desired then stop. Otherwise repeated the steps until the desired output obtained. This algorithm is called back propagation. After that, test with new image and find out the emotion using input given to the previous network structure.

IV. RESULT AND PERFORMANCE EVALUATION

The system is mainly designed to detect human emotions from face. Extraction of features from face is the first step. The second step is the detection of emotion using artificial neural network. Performance evaluation is based on the number of correct emotions detected from the collected images. About 104 images have been collected and 30 images among those is used for training. About 83 images have been correctly detected .The equation for performance evaluation is Accuracy = Correctly detected emotion*100/total image. About 21 images have been incorrectly detected .The performance of the system is 79.80%.This is the output of this system from input image.

V. CONCLUSION

This work try to address the problem of emotion recognition from an image using a neural network based approach. Emotion recognition is still a difficult and a complex problem in computer science. Many researchers have tried to address this problem in various ways. Emotions are expressed by humans in different ways: facial expressions, pose, behavior, actions and speech. Here emotions are detected based on facial expressions .Here neural network to classify the emotions. Neural network based classifier very promising and it provides 79.80 % accuracy in emotion recognition.

References

[1] Vinay Bettadapura. "Face Expression Recognition and analysis:The State of the art" http://www.cc.gatech.edu/vbettada/files/FaceExpressionRecSurvey.pdfArun

[2] Christine L. Lisetti,"Facial Expression Recognition Using a Neural Network"in Proceedings of the Eleventh International FLAIRS Conference.

[3] G. Littlewort et al., "The computer expression recognition toolbox (CERT)", in Proc. IEEE Int. Conf. Autom. Face and Gesture Recognit. Workshops (FG), Mar. 2011, pp. 298305

[4] M. Valstar, M. Mehu, B. Jiang, M. Pantic, and K. Scherer," Metaanalysis of the first facial expression recognition challenge", IEEE Trans. Syst., Man, Cybern.B, Cybern., vol. 42, no. 4, pp. 966979, Aug. 2012.

[5] Shishir Bashya,2008 "Recognition of facial expressions using Gabor wavelets and learning vector quantization" Real-Time Power and Intelligent Systems Laboratory, Department of Electrical and Computer Engineering,Missouri University of Science and Technology, MO 65409, USA.

[6] George C. Giakos. "Neural Networks for Image Analysis and Processing in Measurements, Instrumentation and Related Industrial Applications"

[7] Porawat Visutsak"Emotion Detection Using Sub-image Based Features through Human Facial Expressions " Journal of Man,Machine and Technology Volume 2, Number 1, June 2013 doi : 10.4156/jmmt.vol2.issue1.2

[8] Christine L. Lisetti,"Facial Expression Recognition Using a Neural Network"in Proceedings of the Eleventh International FLAIRS Conference.

[9] Ya Chang et al."Probabilistic Expression Analysis on Manifolds." Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on (Volume:2)