

Feature Extraction For Emotion Detection

Mohit Gahlot*, Mayank Dhinrga, Garima Sharma

Computer Science, The NorthCap University, Gurugram, Haryana, India

ABSTRACT

Segmentation is the process of dividing an image into the multiple segments to simplify and or change the depiction of an image into something meaningful and easier to examine. This paper presents an approach to extract basic features of face to recognize human emotion. Facial features like eye aperture, mouth aperture, and the size of eyebrow constriction are calculated from the localized regions. Experimental results of proposed scheme for feature extraction are simple and robust, and with good accuracy.

Keywords : Feature Extraction, Clustering, Emotion Detection, Histogram Thresholding, Image Segmentation

I. INTRODUCTION

The human face is extremely expressive which is capable of showing numerous emotions without saying anything. Humans classify these emotions by analyzing the facial expression of a person. For example, when we laugh, our eyes constrict and the mouth opens wider. Unfortunately, machines do not have the capability to recognize these emotions. The interaction between human and computer can be improved if machine could recognize these emotions.

The inter-action can be beneficial for applications like medical assistance, psychological therapy, driver fatigue monitoring. This vast area of possible application has intrigued scientists of various fields to do research on automatic emotion recognition. So machine learning helps to train and classify computers to determine the emotion of a person using some input data. In image processing, feature extraction initiates from a set of initial data and adds derived features that are informative and non-redundant which leads to better interpretation.[5] In layman terms, Feature extraction is reducing the amount of data to a small set which describe the larger set of data. Feature extraction finds its applications in various fields. Facial Recognition System, a computer application which is used to

recognize a person from a digital image uses Feature selection plays in juxtaposing selected facial features from the image to the facial features stores in database.

Feature Selection is also used in security systems like biometrics and eye iris recognition systems. It helps extracts prime features like white part of eye and fingerprints. However, many facial features can be used for this process but some basic facial features like mouth, eye and eyebrow has direct relationship with facial expression. According to definition feature selection, it is finding a subset of the initial features, which in our case are length of Eye Aperture, Mouth Aperture and Eyebrow constriction. These selected features contains the relevant details from the input image, so that the desired techniques can be performed to determine the emotion by using this reduced representation instead of every feature of face. This paper deals with extracting these basic features from human face which helps to classify the emotion. These features are extracted from transforming the image and applying image processing techniques. Median filter which is a filtering methodology used for removing salt and pepper noise from an image or signal. It iterates through the each signal entry and substitutes each entry with the median value of neigh-boring entries

i.e window. Histogram-Based Segmentation is used to divide the image into parts to work on regions of interest. It is one of the fundamental and one of the most popular segmentation techniques. It uses the histogram to convert gray scale image into binary images and selects the gray levels for clustering pixels into regions of interest. FCM clustering on the image. Many machine learning and computer vision softwares like MATLAB, python, R, OpenCV can be used to extract these features and determine the useful values. In this paper we are using MATLAB as Image Processing Toolbox is incredibly easy and it automates common image processing workflows. The paper is divided into IV sections. Section I is introduction about the paper and elaborates the applications of emotion detection. In Section II briefly explains about Segmentation, which is a process of dividing an image into multiple parts to alter the representation so that something meaningful can be analyzed. This section focuses on determining position of facial regions like mouth area, location of eyes and eyebrow. Section III uses the segmented images to calculate image attributes i.e Length of Mouth Aperture(MA), Length of Eyes Aperture(EA) and size of Eyebrow Constriction(EC). Conclusion and Output are drawn in Section IV^[5].

II. PROPOSED SYSTEM



Figure 1. Proposed System

A. Preprocessing

These are basically the operations done at the lowest level of abstraction. The aim of preprocessing is the enhancement of images by removing noises by applying different preprocessing algorithms like erosion, histogram equalization etc.

1. Filtering: It is a method used to modify or enhance an image. It holds the power to highlight certain features or to remove certain features. Filtering

operations include smoothing, sharpening, and edge enhancement. Filtering is a neighborhood operation, in which the intensity of the pixels of the output image is calculated by applying some algorithm to the values of the pixels in the neighborhood of the corresponding input pixel. The Input image has been added with Salt and Pepper noise. Adaptive median Filter has been applied on noisy image such that the noise can be removed and the output image would be considered as filtered Image.^[7]

2. Erosion: Morphology is wide set of operations that are related to the shape of the image or the objects in the image. These operations use a structural element on the input images and creates an image with less noise but same size. In morphological operations, the of every pixel within the output image relies on a comparison of the corresponding pixel in the input image with its neighbors. By selecting the dimensions of the neighborhood, you'll construct a morphological operation that's sensitive to specific shapes within the input image^[8]. Erosion is one of the basic morphological operations which removes the pixels along the boundaries of the object. The number of pixels affected by the structuring element depends on the size and shape of the structuring element used to process the image. In this operation, the final state in the output image is determined by applying a rule on the corresponding pixel and its neighbors of the image.

B. Face Detection

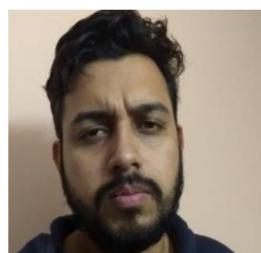


Figure 2. Original Image



Figure3. Face Detected

C. Algorithm Used

FCM Clustering: Clustering is an unsupervised learning method of grouping similar observations or data points together. In most clustering algorithms each object is strictly assigned to one group. But unlike other hard clustering technologies, Fuzzy C-Means mathematically models the fuzziness in the data by assigning a level of engagement to elements with each group. In other words, Fuzzy C-Mean is a process of organizing object points into one or more than one cluster by measuring similarity of each object with every cluster. It is one of the most frequent used technique for pattern recognition. Similarity can be calculated on the basis of closeness, connectivity, and strengthless between the data point and cluster. A better understanding lets take an example of simple mono-dimensional data. All the datums in the dataset are between 0 and 1. The data is to be divided into 2 clusters i.e cluster A and cluster B by setting a threshold. Now according to hard clustering all the data points with value below the threshold value will fall into cluster A and all datums beyond threshold will belong to cluster B.

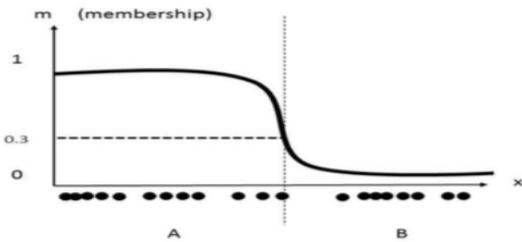


Figure 4. Hard Clustering

Whereas, now if we consider the case of FCM clustering, datums on boundaries between several classes will not belong exclusively to a specific cluster. So in our case now member coefficients will not be stringent to 0 and 1 but rather can vary between 0 and 1. So now member function will now follow a smoother line.

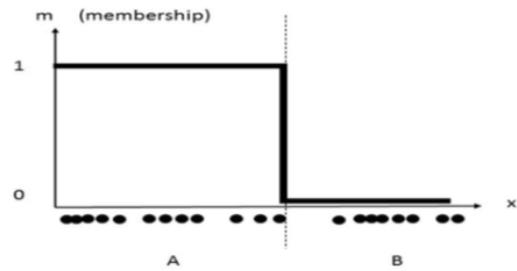


Figure 5. Soft Clustering

In the figure 5, the middle datum have membership in both clusters but it belongs more to cluster B then cluster A.

Steps of FCM clustering:

- ✓ Randomly assign each datum to clusters
- ✓ Calculate the centroids of each cluster

$$c_k = \frac{\sum_x w_k(x)^m x}{\sum_x w_k(x)^m}$$

- ✓ Compute coefficient for each data point of being in a cluster

$$w_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

Here, c_k is the centroid of cluster k , m is a real number greater than 1, w_k is the level of participation of x in the cluster k , and x is the data point.

- ✓ Repeat steps 2 and 3 until the algorithm has converted^[6]

D. Histogram Thresholding

In image processing, histogram thresholding method is a fundamental method for automatic image segmentation. Histograms are constructed by first splitting the data into intervals, called bins and for each bin, the numbers of data points that fall into each bin are counted. In Histogram thresholding, if the image intensity is less than a threshold value then every pixel of the image is substituted with a black pixel else if the image intensity is greater then the threshold value then the pixel is replaced with a white pixel. Thresholding applied to an image is shown in Figure 6, Figure 7 and Fig 8. In this, Figure 7 is the histogram of the image and Fig 8 represents

the histogram with the threshold value. Now according to this value pixels will be divided into two regions R1 and R2. Again new threshold will be calculated according to mean grayscale value of R1 and R2. And this process will continue until the threshold value T becomes smaller than a constant $T_0^{[2]}$.



Figure 6. Original Image

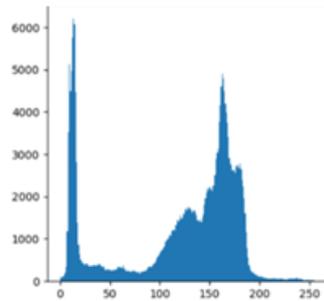


Figure 7. Histogram

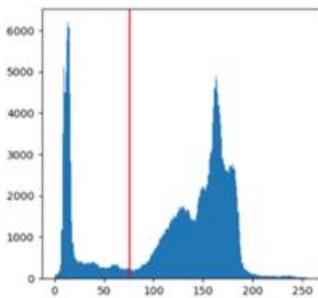


Figure 8. Histogram with threshold



Figure 9. Image after thresholding

E. Classifiers

There are several techniques for image recognition. Among those methods, application of soft computing models on digital image has been considered to be an approach for a better result.

Neural Networks - They were made by keeping in mind the biological neural networks. These systems learn with the previously obtained results and improves its performance by considering examples. As an example, in image recognition, they may learn to spot pictures that contain cats by analyzing example pictures that are manually tagged as "cat" or "no cat" and using the results to identify cats in other images. They are doing this with none a priori knowledge about cats.

SVM (Support Vector Machine) - Support vector machines are a part of machine learning and comes

under supervised learning which are used to classify objects into different class labels. With a given set of training examples every tuple belongs to one or the other category, the algorithm builds a model that assigns the new tuple to one of the categories given, making it a non-probabilistic binary linear classifier.

III. DETERMINATION OF FACIAL ATTRIBUTES

A. Determining Mouth Aperture

We had used FCM clustering to segment the mouth region. Information about color and pixels is used as parameters to the FCM clustering algorithm detect the non-lip and lip region. The five attributes that describe a pixel are the position x,y and three color attributes. The objective of this algorithm is to categorize these 5 points into two clusters i.e. the lip region and the non-lip region. Each pixel is given a level of engagement i.e membership value(MV) in such a way that the summation of the memberships of these 2 areas is one. For the pth pixel x_p , we have

$$N(x_p) + L(x_p) = 1$$

Here $N(x)$ & $L(x)$ represents the MV of the non-lip area and lip area, respectively. ^[5]

After clustering the mouth area, a curve against the average intensity profile length of mouth aperture is plotted. The lip area in the image is represented by dark area and the teeth are represented by the white area. Determination of mouth aperture in a b/w image is easy. Out of the several minima that the plot shows, the first and third indicate the inner region which consists of the area between lower part of the upper lip and the upper part of the lower part, respectively. The gap in the values of these minimas provides the length of mouth aperture. When no white band is observed, the value mouth aperture is set to 0. If only two minima are observed in the curve then the mouth aperture can be determined by finding gap between the two minima. ^[5]

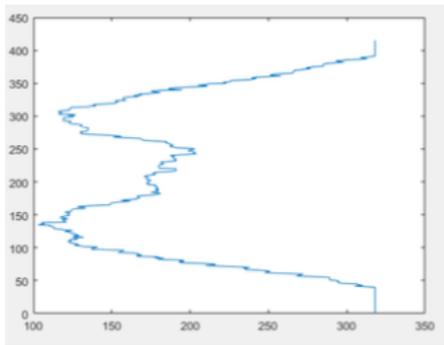


Figure 10. Average intensity plot against mouth aperture

B. Determining Eye Opening

We have used the thresholding method to segment the eye area from the image. The average intensity values of Images clicked under poor lighting conditions is very low. Images should be clicked under good illuminating conditions to overcome the problem of having difficulty in segmenting the eye area because of presence of dark eyebrows near eye region. After segmenting is done, localization both the eyes is needed to be done. After determining the eye location, the summation of white and black pixels (intensity >100 for white and intensity <70 for black) is plotted against the x-axis. If the top is at $x = a$, then the value at $x = a$ provides the total eye aperture. [5]

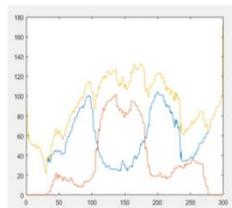


Fig 11. Total pixel count

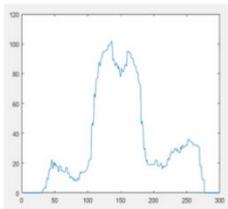


Fig 12. Dark pixels count

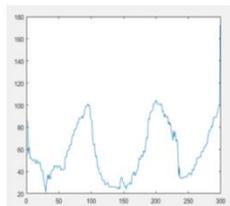


Fig 13. White pixels count

C. Determining eyebrow shrinking

In any image of a face, the 2nd darkest area after the hair area are the eyebrows. We have set a low threshold value in the Histogram-based thresholding method to segment the hair region[3]. We have used a

similar thresholding technique for segmenting the eye regions. A dark narrow template is used to segment the eyebrows.

The collection of patches called valley and hilly in the forehead region makes the constriction. The hilly regions are usually lighter than the valley regions. Let's call the intensity of a pixel 'Iij', and 'Im' be the mean intensity of a rectangular window on the brow. A scan for the change in intensity is required to calculate the span of eyebrow constriction. The length of the Eyebrow constriction can be determined by the maximum horizontal width which includes variation in intensity.

Put a rectangular strip above the eyebrow region. To determine length of the rectangular window we calculate the aperture between the eyebrow parts at both ends. Now determine the middle point of the rectangular window and choose a rectangle. Divide the window into few equal parts by placing vertical lines at equal distances. Calculate the average intensity on these lines. Determine the variance of these average values. Shift the window 1 pixel right if the variance under a particular value. Repeat these steps to calculate left shift. Total length of eye brow shrinking = total right shift + total left shift. [5]



Fig 15. Original Image

Fig 16. Rectangular region on forehead

IV. EXPERIMENTS AND RESULT

This experiment is carried out under controlled settings, where same lights, temperature and noise were kept to maintain constant experimental conditions. The people observed in this experiment

are students and family members of the authors. This experiment is divided into two parts: a photographic session and a face detection session. Iphone 6 plus is used to capture all the pictures. The captured images are transferred in the system for the extraction of features. The independent responses of 8 observers were collected which are summarized in the table.



Figure 17. Output

Table 1. Results

| Subjects | Mouth Aperture | Eye Aperture | Eye brow constriction |
|-----------|----------------|--------------|-----------------------|
| Subject 1 | 0 | 20 | 7 |
| Subject 2 | 36 | 26 | 2 |
| Subject 3 | 6 | 18 | 8 |
| Subject 4 | 1 | 14 | 8 |
| Subject 5 | 42 | 12 | 0 |
| Subject 6 | 2 | 16 | 2 |
| Subject 7 | 1 | 13 | 7 |
| Subject 8 | 2 | 17 | 1 |

V. CONCLUSION

This paper uses a combination of feature extraction techniques which can be used on colour images to localize the region of interest(ROI) in the input image. After the extraction of ROI, facial attributes like mouth opening distance, Eyebrow distance and eye opening distance have been calculated. The presented work can be further extended to incorporate other facial landmarks and distances like the mouth wide

distance, the distance between eye centre and eyebrow. These distances can be used further to classify the images into the universal emotion classes by using neural networks or SVM. The classification can be done by comparing the feature distances of a subject with its respective neutral image. The results of emotion recognition systems using different classifiers could be compared in order to yield an efficient automatic emotion recognition system.

VI. REFERENCES

- [1]. Moysan, G. Corneloup, T. Sollier, Adapting an ultrasonic image threshold method to eddy current images and defining a validation domain of the thresholding method, *NDT&E International*, 32 (1999) 79-84.
- [2]. J.C. Russ, Automatic discrimination of features in gray-scale images, *Journal of Microscopy*, 148(3) 1987) 263-277.
- [3]. P.Daniel Ratna Raju, G.Neelima, Image Segmentation by using Histogram Thresholding, *IJCSET*, January 2012
- [4]. Phillips, Dwayne. August 1991. "Image Processing, Part 4: Histograms and Histogram Equalization," *The C Users Journal*.
- [5]. Aruna Chakraborty, Amit Konar, Member, Emotion Recognition From Facial Expressions and Its Control Using Fuzzy Logic, *IEEE*, 4 July 2009
- [6]. R.Suganya, R.Shanthi, Fuzzy C- Means Algorithm- A Review, *International Journal of Scientific and Research Publications*, November 2012
- [7]. Ruchika Chandel, Gaurav Gupta , Image Filtering Algorithms and Techniques: A Review , *International Journal of Advanced Research in Computer Science and Software Engineering*, Oct 2013
- [8]. Diya Chudasama ,Tanvi Patel, Shubham Joshi, Image Segmentation using Morphological Operations , *International Journal of Computer Applications* May 2015 16
- [9]. Lu Ming , Image segmentation algorithm research and improvement, Oct 2014