International Journal of Scientific Research and Engineering Development-– Volume 3 Issue 2, Mar-Apr 2020

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

A STATIC HAND GESTURE AND FACE RECOGNITION SYSTEM

V.V.N.A.PRIYANKA 1st
Electronics and Communication Engineering
Pragati Engineering College, Surampalem
AndhraPradesh, India
E-mail: priyankavudathu@gmail.com

S. V.V.D.MALLESWARI 2nd
Electronics and Communication Engineering
Pragati Engineering College, Surampalem
AndhraPradesh, India
E-mail: smalleswari.426@gmail.com

V.SAI RAMYA 3rd
Electronics and Communication Engineering
Pragati Engineering College, Surampalem
AndhraPradesh, India
E-mail: ramyavinnapala@gmail.com

T.BALA NANDINI 4th
Electronics and Communication Engineering
Pragati Engineering College, Surampalem
AndhraPradesh, India
E-mail: nandhinitalari@gmail.com

Abstract: This paper presents a framework for a static gesture and face recognition system, which can be helpful for blind people. Real-time images are being taken from a dataset and are processed according to certain algorithms. In the gesture recognition system, human skin is detected using lab color space. To identify the number of fingers, different features like fingertips and the angle between fingers are being extracted. By using the gesture recognized various operations can be performed like the opening of doors, turning on/off the fan or light. In the face recognition system, LBP recognizer and SVM classifiers are used for extracting features and recognition respectively. Here the hand gestures and faces of different persons are identified and it is converted into a speech that can be heard by blind people. The gesture and facial recognition were accomplished with an accuracy of 95% and 92% respectively. This system acts as an assistant for physically disabled people.

Keywords: Gesture Recognition, Fingers, LBP,SVM

I. INTRODUCTION

Biometric information-based authentication systems provide greater security. It is utilized for creating a natural interaction between person and machine. On the basics of face, fingerprint, palm print, retina, iris biometric systems can recognize the uniqueness of a living person.Imageprocessing performs a major role in biometric authentication.

Gestures can be static or dynamic, but static gestures require less computational complexity whereas dynamic gestures are more complex. Sign language is a combination of different gestures, shapes, movements of hand, body and facial expressions. It is one of the adequate communication tools for disabled people. This system acts as a bridge for them to communicate with normal people. This system will be helpful to blind people to recognize the person and gestures and this is done by using MATLAB.

The proposed system is divided into two sub-systems-

- Hand gesture recognition system and
- Facial recognition system.

II. HAND GESTURE RECOGNITION

This subsystem is divided into three parts –

- Hand region detection
- Pulling out the features
- Gesture Recognition

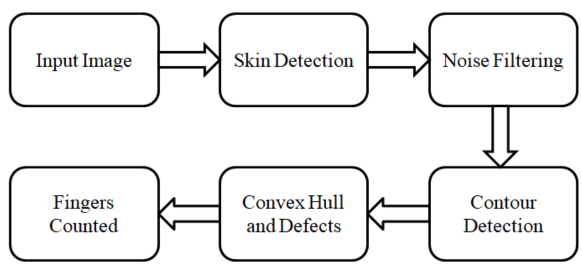

Figure1. Hand Gesture Recognition System.

A. Hand Region Detection

The first step is to separate the hand region from the input image, to detect this thresholding effect is used in different color spaces like RGB, YCbCr, LAB, etc.

Thresholding means converting the image having multiple level intensities to two-level intensities (0 or 1). Thresholding can be used to extract the region of interest from its background by assigning intensity values to each pixel. Then these are classified into wanted and unwanted pixels. This process is applied on the input image according to a threshold value, the pixel intensity values greater than the threshold value will be converted into 1(white) and which below to threshold value to 0(black). LAB color space can be used for thresholding images. In this system, hand or skin detection is performed using the thresholding operation in LAB color space.

1. LAB's color space

LAB is a 3-axis color system with dimension L for lightness and a and b for the color dimensions. Lightness ranges from 0-100. Here 0 is black (no light) and 100 is white (maximum illumination). In this model, a and b have two representations they are +a, -a and +b, -b. Each represents different colors that are +a means red, -a means green and +b means yellow, -b means blue. Theoretically, there are no minimum and maximum values of a and b, but in practice, they are usually numbered from -128 to +127 (256 levels). The Lab's color space is device independent and it also includes all colors in the spectrum. The threshold value is being set according to the environment where the threshold value is taken as 80. The value above the threshold values represents the non-human skin and the values below that value represent the human skin. So that the human skin is detected in no light and the night also. After that, a black and white image created for the hand detected and in this white color determines the hand.

To remove noise, smoothening of the image is done using the Gaussian blur technique. In this technique, the image is convolved with the Gaussian function, such that high-frequency components are removed.

B. Pulling out the Features

The next step in hand gesture recognition is extracting features of detected skin. This can be achieved by a sequence of steps they are contour detection, convex hull and detecting convex hull defects.

1. Contour Detection

Firstly, the contour is detected for hand. Contour is one of the edge detection technique. A contour is a line drawn around the which joins the points having similar features like color, intensity, texture. Extracting the hand contour helps the shape analysis to determine the hand gesture.

2. Convex hull

Before starting this topic, the concept of convex polygon needs to be introduced first. A convex polygon is a polygon which does not contain any concave part. In mathematics, there is a strict definition of a convex polygon. Take any two points in a polygon(includes the boundaries and the area in which the boundaries are a covering) and connect the two points within a straight line. If all the straight lines that connect any two points inside the polygon doesn't exceed the boundary of the polygon, then that the polygon is convex. After knowing the definition of a convex polygon, then the convex hull can be defined as for a given set of points in a plane, the convex hull of the set is the smallest convex polygon that contains all the points in the set. The points that form the convex hull are called "Hull points". Then the convex hull is formed around the hand which contains all the points in that convex hull. It joins mainly fingertips because they are the hull points.

3. Convex Defects

Convexity defect is an area in an object (contour) segmented from an image. That means an area that does not belong to the object but located inside the convex hull outer boundary. Calculate the convex hull of the hand contour to get the convexity defects of the contour. Convexity defect provides very useful information to understand the shape of a contour. The first step of searching for a convexity defect is finding the starting point of a convexity defect on the contour. The starting point of a convexity defect means a point on the contour which is also included in the convex hull points, but the next point on the contour is not included in the convex hull points, that is the centre point. After knowing the starting point, the ending point will be similar. The ending point to be the point in the contour, which is included in the convex hull points, but the points between them are not included in the convex hull points. After searching all the points in the contour, then various convexity defects are processed. Each convexity defect is composed of a starting point, ending point and the points between them.

C. Gesture recognition

To count the fingers and find the angle between the fingers three points are extracted from the convex hull. They are

- Starting point (St)
- Center point (Ce)
- End point (En)

The length of the defects is calculated using these defect points using the formula

$$L1=\sqrt{(St(0)-En(0))^2+(St(1)-En(1))^2}$$

$$L2=\sqrt{(Ce(0)-En(0))^2+(Ce(1)-En(1))^2}$$

$$L3=\sqrt{(St(0)-Ce(0))^2+(St(1)-Ce(1))^2}$$

Then using Heron's formula, the area between the fingers is being calculated. Distance between the defect point and the convex hull is computed using the formula

Distance =(2*area)/L1

The angle between the fingers is determined using cosine formula, that is

$$A=\cos^{-1}\left(\frac{L2+L3-L1}{2*L1*L2}\right)$$

The computed values that are the derived angle and distance are compared with the threshold values they are data stored in the database by training the database with different types of samples using an SVM classifier, fingers are being counted accordingly.

III. Face Recognition System

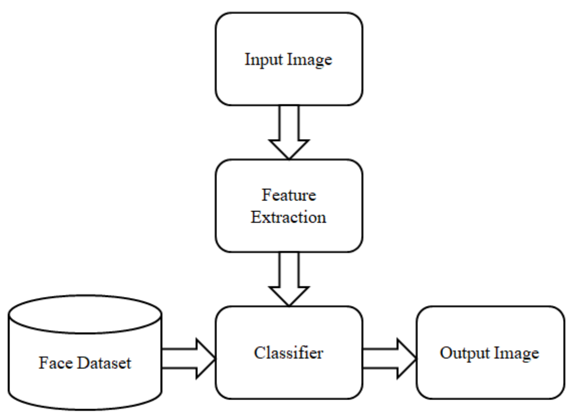

Figure2. Face Recognition System.

The face recognition system mainly concentrates on feature extraction and classification. LBP and SVM are used for the extraction of features and classification. First, the algorithm needs to be trained. To do this firstly, maintain the dataset with the facial images of the people whom we want to recognize. Then give a name for each image, so the algorithm will use this information to recognize an input image and give you the name of a person in words. Images of the same person must have the same name. After the construction of the training set, then the input is taken from the camera that is a photo of a person. The image taken as input is resized to match the size of images in the dataset. Then the color image is converted into grayscale.

A. Local Binary Pattern

LBP is used for feature extraction which abbreviates local binary patterns. It is described by Ojala in 1996. LBP is a simple and very advantageous texture operator. It operates thresholding operation on neighbouring pixels with respect to centre pixel to extract the features. The algorithm takes images as input and feature vectors are obtained as output. The steps to be followed are

1) The grayscale image is resized is taken as input.
2) In the image 3*3 window size part is taken.
3) The part taken is represented in the form of intensities of pixels and they are taken in the form of a matrix.
4) The centre pixel is taken as a threshold value and thresholding operation is performed on the neighbouring pixels based on the centre pixel value.
5) The neighbouring pixels which have an intensity value greater than the centre value are considered as 1 and which is having the value lower than the centre value is considered as 0.
6) This process forms the matrix consists of 0 and 1. The centre value will be neglected.
7) The binary values are taken in a pattern to form the decimal values.
8) The decimal value is placed in the centre value in the image intensity matrix of the image.

In this way, these intensity values are taken into a vector which represents the image in a better way. These are feature vectors that are characteristics of the image and used in the classification.

B. Support Vector Machine

The SVM algorithm proposed by Osuna, Freund, and Girosi detects faces by taking an image for face-like patterns at any scales, dividing the image into the sub-images and classifying them using an SVM to determine the class. SVM uses the feature vectors to classify the images which are taken as input and compared to the images that are in the dataset. The Kernel function is the heart of SVM. In the application of SVM during face recognition, it identifies a suitable kernel function for the selected face database. SVM finds a support vector to perform pattern recognition between two classes. This support-vector is a decision surface in the training set which has the maximum distance to the closest points. The main aim of the SVM classifier is to reduce generalization error upper bound through maximizing the margin between separating hyperplane and data. Support Vector Machine became the most popular supervised classification method due to its superior classification performance in different applications.

IV. Accuracy

After the recognition of gestures and faces, the accuracy of the system is shown. The accuracy of the system is calculated by the object evaluation. Accuracy depends on the faces and gestures detected.

Table1. Confusion matrix of outputs

Actual / Predicted	**Negative**	**Positive**
Negative	True Negative	False Positive
Positive	False Negative	True Positive

In this figure, the Actual means real outputs and predicted means outputs generated by the system. For example, the real output is the image or gesture that is not found and predicted output also the same then it is called the True Negative. As same to the True Negative the False Positive is the actual output is not found but the predicted output is found the image. The False Negative means the actual output is the image is found but the predicted output is the image is not found and the True Positive means the actual output is the image is found and the predicted output is also the same as actual. These are the values used to calculate accuracy and precision.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive+False Positive}} \ldots(1)$$

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive+False Negative}} \ldots(2)$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative+False Positive}} \ldots(3)$$

Accuracy can be calculated by the mean of the Specificity and Sensitivity.

$$\text{Accuracy} = \frac{\text{True Positive+True Negative}}{\text{True Positive+False Negative+True Negative}} \quad \ldots(4)$$

In this way, the accuracy is calculated for the hand gestures and face recognition system.

V. RESULTS

A. Hand Gesture Recognition

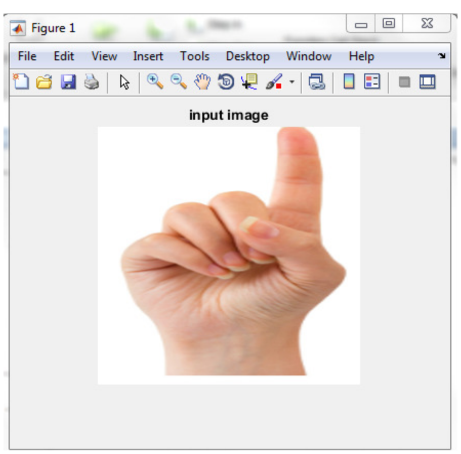

Figure2. Input Image

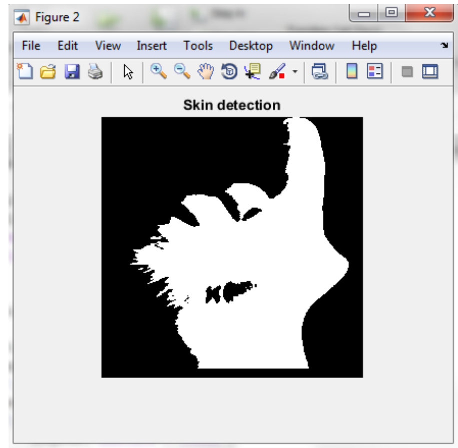

Figure3. Skin Detection

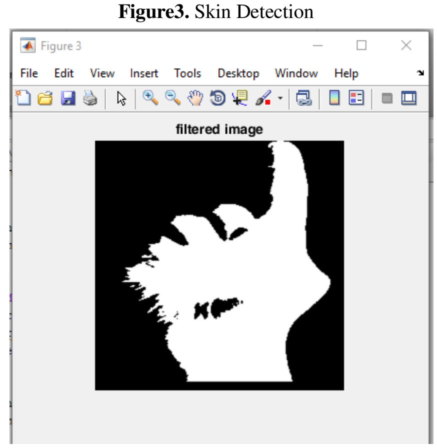

Figure4. Gaussian Filtering

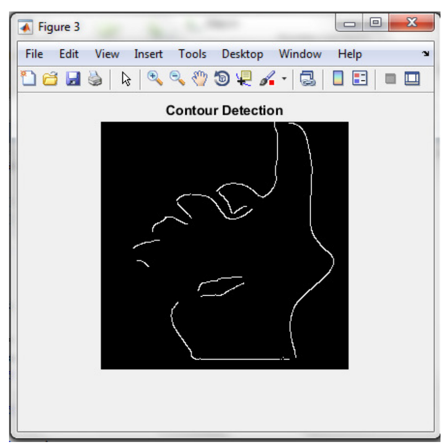

Figure5. Contour Detection

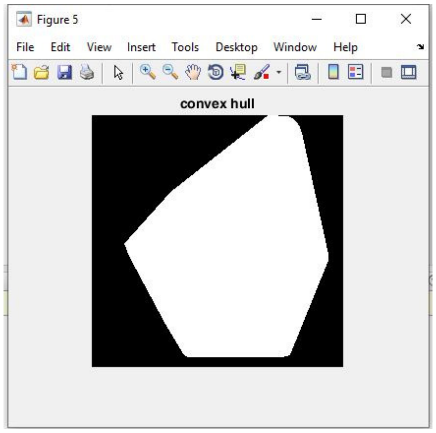

Figure6. Convex Hull

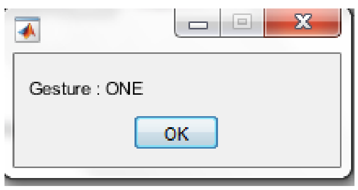

Figure9. Gesture Recognition

B. Face Recognition

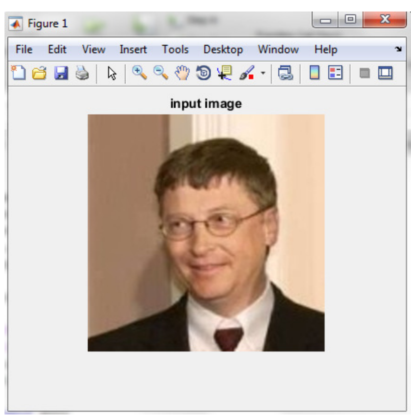

Figure9. Input Image

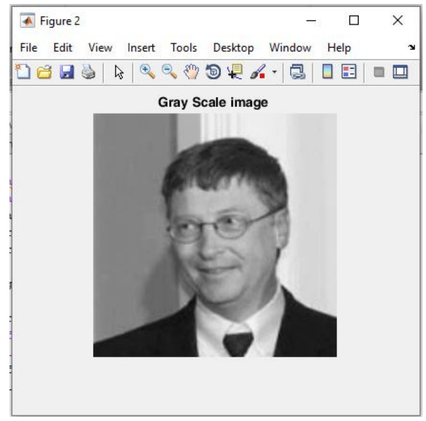

Figure10. Gray Scale Image

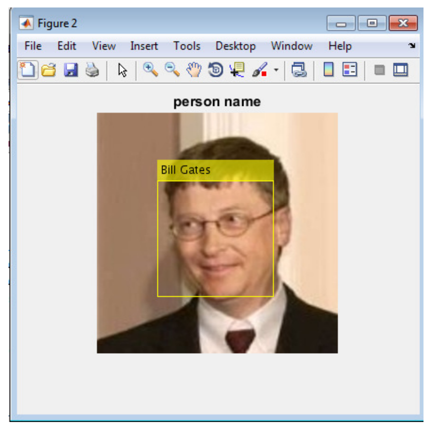

Figure11. Face Recognition

VI. CONCLUSION

The system introduced in this project can be helpful for blind people and can act as a virtual assistant for them. The hand gestures and faces of different persons are identified and it is converted into a speech which can be heard by blind people. Hand gestures are recognized with an accuracy of 95.2% and face recognition and identification have been done with an accuracy of 92%.

ACKnowledgement

We would like to express special thanks of gratitude to our mentor Mrs.P.Sunitha, M.Tech, (Ph.D.). Also, we would like to express sincere gratitude to Pragati Engineering college for support and encouragement for implementing this project.

References

[1] Srinivasan Ganapathy Raju, "Hand Gesture Recognition Using Convexity Hull Defects to Control an Industrial Robot", 2013 3rd International Conference on Instrumentation Control and Automation (ICA) Bali, Indonesia, pp. 63-67, 2013.

[2] Viola P, Jones M J. Rapid object detection using a boosted cascade of simple features [C], Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (ICCVPR2001), vol. 1, pp. I-511-I-518, 8-14 Dec. 2001, Kauai, USA.

[3] Kelvin Salton do Prado, "Face recognition: Understanding LBPHAlgorithm", 2017, Towards Data Science. [Online]. Available: https://towardsdatascience.com/face-recognition-how-lbph-works-90ec258c3d6b. [Accessed: 7th January 2018]

[4] HaneneElleuch, Ali Wali, Anis Samet, Adel M. Alimi, "A static hand gesture recognition system for real-time mobile device monitoring", 2015 15th International Conference on Intelligent Systems Design andApplications (ISDA), pp. 195 – 200, 2015.

[5] Aftab Ahmed, Jiandong Guo, Fayaz Ali, Farha Deeba, AwaisAhmed,"LBPH Based Improved Face Recognition at Low Resolution", 2018International Conference on Artificial Intelligence and Big data, pp. 144-147, May 2018.