A Static Hand Gesture and Face Recognition System for Blind People

Conference Paper · March 2019

DOI: 10.1109/SPIN.2019.8711706

3 authors, including:

Samyak Jain
Nirma University
3 PUBLICATIONS **48** CITATIONS
SEE PROFILE

Khushboo Ph.D.
Vivekananda Institute Of Professional Studies
12 PUBLICATIONS **133** CITATIONS
SEE PROFILE

A Static Hand Gesture and Face Recognition System For Blind People

Saransh Sharma[1], Samyak Jain[2], Khushboo[3]
Email: [1]sharma.saransh99@gmail.com, [2]Samyakjain.1997rocks@gmail.com, [3]kkhushboo_2008@yahoo.com
[123]Dr. Akhilesh Das Gupta Institute of Technology and Management, Shastri Park, New Delhi, 110053, India

Abstract— **This paper presents a recognition system, which can be helpful for a blind person. Hand gesture recognition system and face recognition system has been implemented in this paper using which various tasks can be performed. Dynamic images are being taken from a dynamic video and is being processed according to certain algorithms. In the Hand gesture system Skin color detection has been done in YCbCr color space and to discover hand convex defect character point of hand is used where different features like fingertips, angle between fingers are being extracted. According to gesture Recognized, various tasks can be performed like turning on the fan or lights. While in face recognition, Haar Cascade Classifiers and LBPH recognizer are being used for face detection and recognition respectively. With the help of OpenCV, The research has been implemented. Various hand gestures and human faces have been detected and identified using this system. The hand gesture was recognized with an accuracy of 95.2% was achieved and facial recognition was done with an accuracy of 92%.**

Keywords— **Cascade Classifier, Convex Hull, Haar like features, Hand Gesture Recognition, Human Computer Interaction, LBPH, OpenCv, Skin Detection**

I. INTRODUCTION

This system acts as an Intelligent virtual assistant which helps in bringing the world closer and helps in meeting different ends together. Through this system use of facial recognition and hand gesture to assist and work with the environment to make it a better place to live for the blind. There are different ways in which it is helping to make this happen and for that hand gesture and facial recognition, system is being used. This system is using computer vision technology this tech has been used for human computer interaction (HCI) using a physical medium where hand gesture and facial recognition plays a major role. Hand gestures have been done since the dawn of civilization and have various meaning depending on the geographical location [1]. Hand gesture have various application in military gaming etc. Methods without using computer vision has also been developed like the wearable gloves but they are too costly as they need sensors and other hardware devices there are numerous algorithm for hand gesture recognition like KNN (K Nearest Neighbor), artificial neural network but most of the algorithm require large amount of samples for training and recognition to overcome this problem convex hull and convexity hull defects are used for gesture recognition. The recognition of hand gestures are of two types, that is, Dynamic gesture and static gesture, so in this research we have covered the section of static gesture recognition. In facial recognition system, two methods are used Haar cascade method and Linear Binary Pattern (LBP) for better prediction we use LBP method. In 2004, a Haar cascade classifier technique was proposed by Voila Jones which has been a motivation to various face recognition systems.[2]

LBP is used as a visual description in which classification using computer vision is made using open cv each image is converted into a series of code. a pixel is taken and then the image is given a particular value. This value is predefined based on a series of codes. A LBP histogram mainly works on four parameters: radius, neighbor, Grid X, Grid Y. A radius has an initial value 1, which builds a central pixel, which is used to build circular local binary pattern. Neighbors on the other hand are used to create a sample points to increase the accuracy of facial detection but it comes with drawback of increasing computational cost. Grid x and Grid y are horizontal and vertical features to recognize an image [3].

II. PROPOSED ARCHITECTURE

The input is taken from front facing Camera. The proposed Recognition System is divided into two parts –

- Hand Gesture Recognition
- Facial Recognition

A. Hand gesture recognition

This subsystem is divided into three parts –

- Hand region detection
- Pulling out the features

1) Hand region detection

In order to split hand from the real time images RGB colour

$$\begin{aligned} Y &= 0.257R + 0.564G + 0.098B + 16 \\ Cb &= -0.148R - 0.291G + 0.439B + 128 \\ Cr &= 0.5R - 0.419G - 0.081B + 128 \end{aligned} \quad (1)$$

space is converted to YCbCr colour space. [4]

Then the Upper and Lower threshold values are being set according to the environment. The user according to the environment in the real time dynamically changes the values. For the detection of skin, the corresponding ranges for YCbCr are $30 \le Y \le 255$, $77 \le Cb \le 127$, $133 \le Cr \le 177$ [5]. According to the lighting of the System and various skin color types Different ranges of YCbCr color space has also been researched and they

are as follows 30≤Y≤ 235, 95≤Cb≤240, 110≤Cr≤ 240. These values worked well during no light or night. Other range 30≤Y≤ 235, 80≤Cb≤ 120, 133≤Cr≤ 177 can also be used. Then a black and white image is created which is the region of interest and in this white color determines the hand. To remove noise smoothening of the image is done using Gaussian blur technique. In the technique, the image is convolved with the Gaussian, filter such that high frequency components are removed.

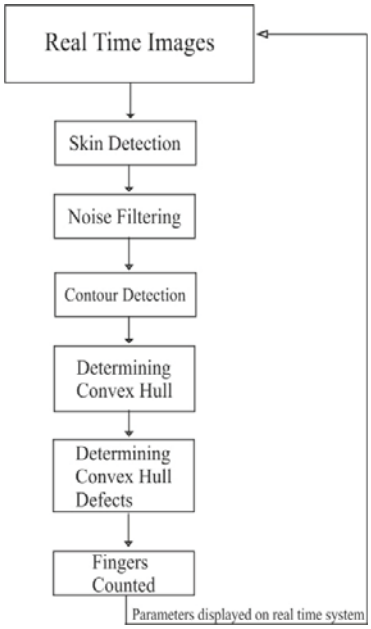

Figure 1: Proposed System Architecture for Hand Gesture Recognition

2) Pulling Out the Features

Contour Detection has been used to pull out the features. It is one of the famous edge detection method. A contour is an arc connecting the continuous points having identical Colour or any feature. Contour approximation is done using Ramer Douglas Puecker algorithm, which reduces a number of points in a curve by approximating a succession of points. After the contour is detected, a convex hull is being created around it. Convex hull is a compact convex bound around the detected shape. Any divergence from the main hull acts as defect so the gaps between the fingers act as a convexity defect [6]. Using these defects further recognition of the gesture is done.

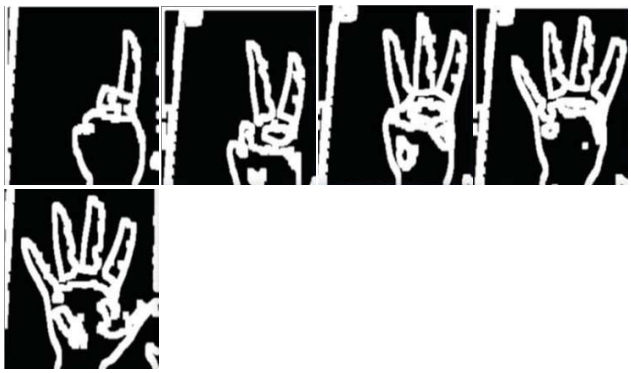

Figure 2: Masking of the Image

3) Gesture Recognition

In order to count the fingers and find the angle between the fingers three major points are being extracted –

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

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

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

$$L2 = \sqrt{\left(Ce(0) - En(0)\right)^2 + (Ce(1) - En(1)^2} \quad (3)$$

$$L3 = \sqrt{\left(St(0) - Ce(0)\right)^2 + \left(St(1) - Ce(1)\right)^2} \quad (4)$$

Then using the 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

$$\text{Dist} = \frac{2*\text{area}}{L1} \quad (5)$$

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

$$A = \cos^{-1}\left(\frac{L2+L3-L1}{2*L1*L2}\right) \quad (6)$$

The computed values that is the derived angle and distance is compared with the threshold values and the fingers are being counted accordingly.

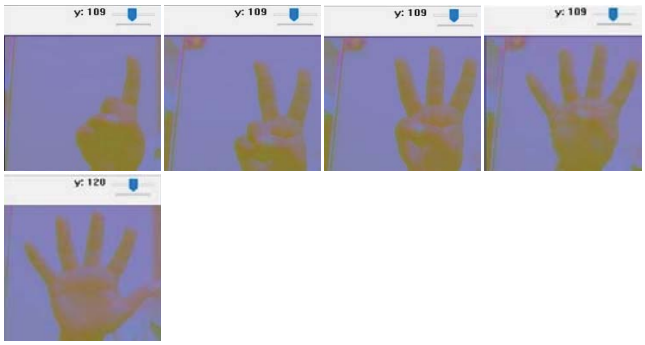

Figure 3: Dynamically changing the values of Y in Real Time

B. Face Recognition and Identification

Haar Cascade classifiers has been used detection the face and LBPH Algorithm is used to identify the image of the person.

1) Face detection

It is an object detection method used to discover faces or cars or any object. This system is provided with positive and negative

images and selection of features along with the classifier training and integral images [7]. Each feature is the difference sum of the pixels within two rectangular regions. These rectangular regions are nothing but the darker and the lighter regions. These sectors have same size and shape horizontally or vertically. To find the sum of pixels, under black and white regions, the concept of integral images is introduced. No matter how large may be the number of pixels; it operates over only a fixed number of pixels.

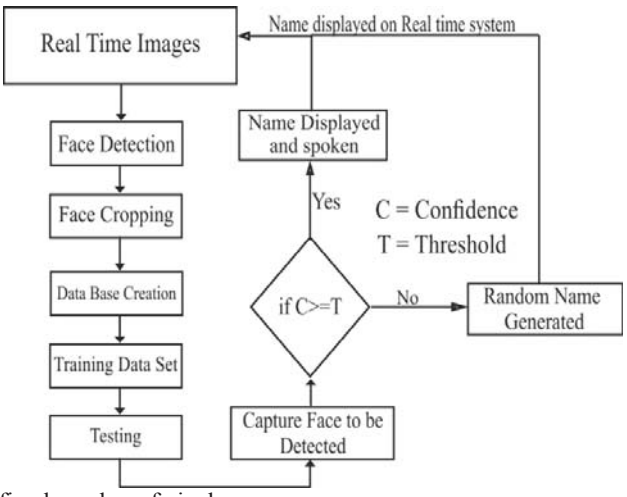

Figure 4: Proposed System Architecture for Face Recognition

AdaBoost does the selection of most appropriate features out of a large set of features. Weak classifiers are combined to form the final classifier. They are weak because alone they cannot identify the object. Each stage of cascade must not have a low false negative rate as when face is classified as non-object then the classification stops. Each stage must have a high false positive rate which means the erroneous detection of an object as face, thus this error can be corrected in $(n+1)^{th}$ stage and succeeding stages in classifier. [8]

2) Face Identification

LBP, abbreviation for Local Binary Pattern, is the measure of fixed grey scale texture obtained by thresholding the nearby area of each pixel and the result obtained is considered as a binary number. The LBP operator is the ratios of the pixel intensities of the center and the surrounding pixels and can be computed using the equation

$$LBP(Xc.Yc) = \sum_{n=0}^{7} S(Jn - Jc)2^n \qquad (7)$$

Where Jc is the value of center of pixel and X_c and Y_c shows the eight surrounding pixels of information.
LBP uses the following steps to identify the face

- Firstly, the image is converted to grayscale.
- Section of the image is obtained through 3*3 window.
- Central value of matrix used as threshold, which is used to define new values of other eight neighbours.
- New value calculated is either 1 or 0 for values higher or lower than threshold respectively and the concatenation of the Binary values is done.
- Conversion of Binary to Decimal is done in order to get the pixel from the original image.
- Histogram using the GridX and GridY parameters, which is the number of cells in horizontal and vertical direction, is created which divides the image into multiple grids.

The current Histogram is compared with the histograms already existing in the database using different distance techniques like Chi Square method, Euclidean Distance, Absolute Value. In this Research, Chi Square Method is used and is calculated using formula. [9]

$$D(H1, H2) = \sum_I \frac{(H1(I) - H2(I))^2}{H1(I)} \qquad (8)$$

In addition, this distance gives the confidence, which determines the closeness between the histograms.
In the Research, major tasks performed were Detecting, capturing, training, recognizing the face.

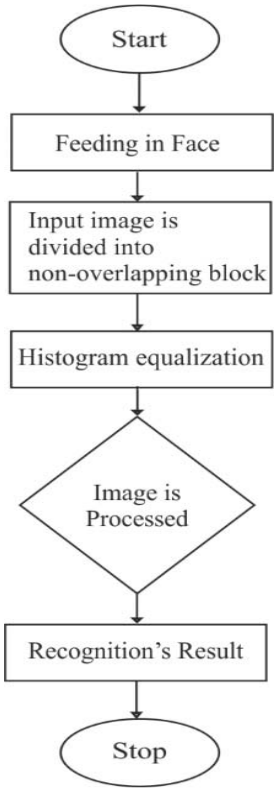

Figure 5: Flowchart for LBP Algorithm

Detection and capturing

The face is detected in real time and when a user wants to save the image, a certain indication is given to the system, after which the system crops the face detected in the real time system. This Detection of the face is done through object classifier technique that is Haar cascade classifier. Then the System asks for the name with a question "who are you" with which a folder is stored and image with a name at that instant is stored. Pyttsx3 Library has been used to convert Text to speech, which uses Natural Language Processing methods to achieve such task. The major advantage to use this library is that the task can be performed offline. This image is stored in black and white scheme for further processing.

Dataset SS100

A Database named SS100 has been created in which 6-10 images for each person has been stored, in total 100 images are stored. Each image has a different expression and postures. These images are stored with their respective names which is further used to display name while identification of the face. These images are then converted to black and white color scheme for further processing. [10]

Training

Training is done in order to store the values of histograms of every image in database. These Histograms play an important role in order to identify a face.

Recognizing

The Real Time image is compared with the images in the database using Classifier and Recognizer techniques , if the real time image matches with the image in database the name is displayed and if the user want it can also ask for the name to be spoken out by asking the system "who is he". This Speech to text conversion is done through Google's strong Speech to Text Converter, which uses most advanced deep learning algorithm to do such task.

III. Experimental Results and Discussion

The project has been written in python language using PyCharm 2016.3.3 IDE using OpenCV 3.3.0 library. The computer or desktop used has a configuration of 8 GB Ram, Intel i5 – 7200U @2.5 GHz. A Web Camera of 0.9 MP is used to capture hand and face from the real time system. The captured image of hand is further processed for skin color detection, noise reduction, detection of gesture is done using convex hull algorithm, and a number of fingers are counted using convex defects. Five Different gestures were recognized using this system representing the number from one to five. All gestured were recognized under dynamic video. A successful recognition of gesture with an accuracy of 95.2% was done within 0.1 to 0.4 seconds. . As shown in Table 2, various participants participated in testing the gesture recognition system, this table has shown that the live gesture recognition is a robust and have a high recognition rate .The gestures are recognized with natural full opening of the fingers.

TABLE I. TRAINING IMAGES STATISTICS

Total Images	Recognized Images	Unrecognized Images	Training Time(seconds)
100	92	8	11.15

Where in the facial recognition, the captured image of face is converted to grayscale image, using the images database is trained and using classifiers and recognizers the face is identified.

	Gesture 1	Gesture 2	Gesture 3	Gesture 4	Gesture 5
P 1	100	90	100	90	100
P 2	100	100	90	100	90
P 3	90	100	90	100	90
P 4	90	90	90	90	100
P 5	100	90	100	90	100
P 6	90	100	100	90	100
P 7	100	100	100	100	100
P 8	90	100	100	90	100
P 9	100	90	90	90	100
P 10	90	100	100	90	90
P 11	100	100	90	100	90
P 12	100	90	100	90	100
P 13	100	100	90	100	90
P 14	90	90	90	100	100
P 15	100	90	90	90	90

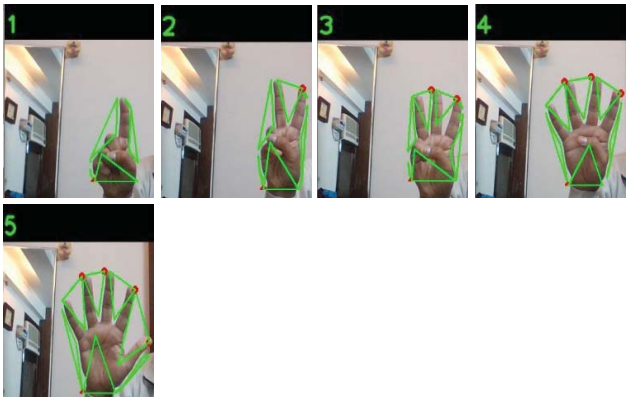

Figure 6: The five gestures are being presented from number one to five

TABLE II. SUCCESS RECOGNITION RATE OF EACH GESTURE BY VARIOUS PARTICIPANTS

Therefore, according to the statistics and practical measures it has been found that Face has been recognized with an accuracy of 92%.

During research, various problems faced were the noise and the lightning that effects the detection of both hand and faces. Smoothening of the image was done in order to reduce the background's noise and skin color is detected with more accuracy. Even though Y's value can be dynamically varied in YCbCr color space in real time but still at some points hand is not recognized by the camera properly. To recognize the hand gesture a person must always be present inside the frame and the hand must be present inside the region of interest for further processing. While in Face Recognition, the number of images stored determined the accuracy of the system.

.

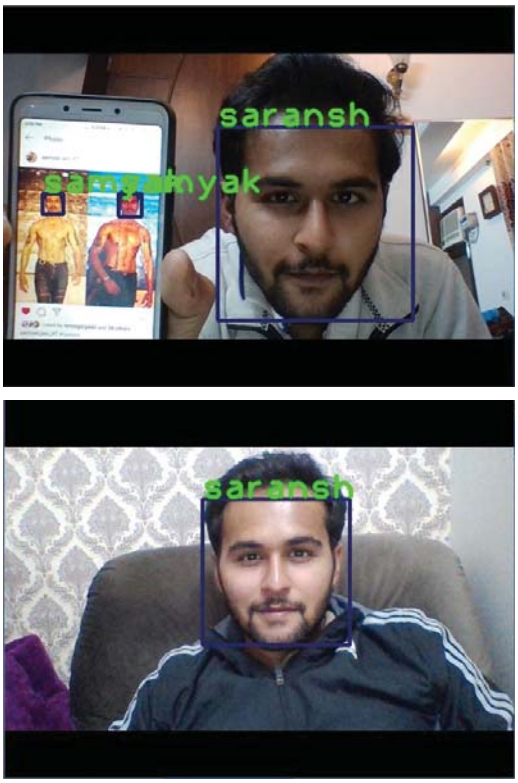

Figure 5: Successful Face Recognition by the System

IV. APPLICATIONS

Application is limited only by our imagination. Not only blind people can use it but also a number of industries can use this technology. Hand gestures can be applied in the field of telemedicine or surgery and even in locking systems. Industrial robots can also take an advantage of hand gestures. While Facial recognition have its own applications like it is used in security, criminal identification, Advertising and Healthcare. Due to use of LBPH Recognizers, facial recognition can also be done in low light.

V. CONCLUSION AND FUTURE SCOPE

The system introduced in this paper can be helpful for a blind person and can act as a virtual assistant for it. Haar cascade Classifiers and LBPH recognizers has been used for face detection and identification in the real time whereas Convex hull and Convex defects algorithm has been used to detect the Hand gestures in real time. Skin color recognition has been in YCbCr color space and different threshold ranges have been used to detect skin color in different lighting conditions and skin color. Hand gestures are recognized with an accuracy of 95.2% and face recognition and identification has been done with an accuracy of 92%. There are some limitations, which are needed to be addressed. Recognizing more amount of gestures would be helpful for performing more tasks. Alternate methods like MLBPH [12] or LBPH + CNN [13] can be used to improve the gesture recognition and face recognition must be considered.

REFERENCES

[1] Srinivas Ganapathyraju, "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] Voila 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 (ICCVPR 2001), vol. 1, pp. I-511-I-518, 8-14 Dec. 2001, Kauai, USA.

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

[4] Hsiang-Yueh. Lai, Han-Jheng. Lai, "Real-time dynamic hand gesture recognition", 2014 International Symposium on Computer Consumer and Control (IS3C, pp. 658-661, June 2014.

[5] Muh. Arif Rahman, , I Ketut Edy Purnama, Mauridhi Hery Purnomo," Simple Method of Human Skin Detection using HSV and YCbCr Color Spaces", 2014 International Conference on Intelligent Autonomous Agents, Networks and Systems Bandung, Indonesia, August 19-21, 2014.

[6] Hanene Elleuch, 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 and Applications (ISDA), pp. 195 – 200, 2015.

[7] Gaganpreet Singh, , "Training Haar Cascade", 2012. [Online]. Available :https://singhgaganpreet.wordpress.com/2012/10/14/training-haar

cascade/. [Accessed : 5th January 2018].

[8] Sander Soo, "Object detection using Haar-cascade Classifier", Institute of Computer Science, University of Tartu.

[9] L. Cuimei, Q. Zhiliang, J. Nan, W. Jianhua, "Human face detection algorithm via Haar cascade classifier combined with three additional classifiers", 2017 13th IEEE Int. Conf. Electron. Meas. Instruments, pp. 483-487, 2017.

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

[11] XueMei Zhao, ChengBing Wei, A Real-time Face Recognition System Based on the Improved LBPH Algorithm, 2017 IEEE 2nd International Conference on Signal and Image Processing, pp. 72-76, 2017

[12] Hongshuai Zhang, Zhiyi Qu Liping, Yuan GangLi, A Face Recognition Method Based on LBP Feature for CNN, 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), pp. 544-547, 2017.