Hand Gesture Recognition through Computer vision

Project Report of Summer Internship

Under the Guidance of:
Prof. Anupam Agrawal

Submitted by:

Sourav Jaiswal
(IIT(BHU),Varanasi)

Indian Institute of Information Technology, Allahabad

16th May-16th July, 2018

CERTIFICATE FROM SUPERVISOR

It is certified that this report **"Hand gesture recognition through computer vision"** of summer research internship project is the bona-fide work of "**Sourav Jaiswal**" who carried out the project work throughout the duration under my supervision.

( )

Prof. Anupam Agrawal
IIIT – Allahabad, Allahabad

ACKNOWLEDGEMENT

I would like to thank **Prof.Anupam Agrawal** for his supervised insight and expertise that greatly assisted me in this project. His Continuous suggestions throughout these 2 months were really invaluable and beneficial in completion of the project.

I would also like to show my gratitude towards **Mr. Rajesh Bulla** (Research fellow under Prof.Anupam Agrawal) for his assistance in providing me the required help and guidance .I would like to thank him for sharing his guidance and wisdom during the course of this project.

By,
Sourav jaiswal
IIT(BHU),varanasi

Contents

1. INTRODUCTION 6

2.OBJECTIVE AND PROBLEM STATEMENT 7

3.REVIEW OF RESEARCH PAPER 8

4.TECHNIQUES USED(in our process).....9

5 APPROACH TAKEN (OVERALL CYCLIC PROCESS) 14

6.ANALYSING STEPS 15

7.LIBRARIES & SOFTWARE USED.....16

8. RESULT OF RESEARCH 17

9.APPLICATIONS.....19

10. IMPROVEMENT(comparison with other research papers) 20

11. FUTURE WORK 21

1. INTRODUCTION

A primary goal of gesture recognition research is to create a system which can identify specific human gestures and use them to convey information or for device control.

The real time hand gesture recognition can be used to make our system work from a distance and without any physical contact.

For people who are disabled and can't move, can control the system from their own location and of course hand gesture recognition can be very helpful for people to convey their message who are blind or who can't speak.

A recognition system that we proposed comprises of a framework of two most important entity, which are..

1. Pre-processing (image processing)
2. Recognition classification (artificial intelligence recognition)

As both the work takes place simultaneously in real time through camera feed, gestures are recognized, which people show with their hands in front of the cam.

2.OBJECTIVE AND PROBLEM STATEMENT

The main objective of the project is to **control the laptop through hand gesture**, in such a way that we should be able to **detect number of fingers** and execute various commands according to that.

Another problem that we are able to execute is to **mimic the mouse cursor through hand motion**, and doing various commands like left click, right click, scroll up, scroll down, **left swipe, right swipe** according to different hand gestures.

3.Review of Research Papers

From the review of research papers and the work they have done, the **following techniques** were used by them for hand gesture recognition:-

- **In YCbCr colorspace**
- The single component "Y" represents luminance information,.
- Component Cb is the difference between the blue component and a reference
- Cr is the difference between the red component and a reference value.

$$f(Cr,Cb) = \begin{cases} 1 \; if \; (Cr,Cb) \in S \\ 0 \; if \; (Cr,Cb) \notin S \end{cases}$$

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112.000 \\ 112.000 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The Template matching cross-correlation involves

1. Simply multiplying together corresponding pixels of the signer image

2. Here the image taken as feed is called the Target image and is compared with the pixel values of saved image and then summing the result.

$$\frac{1}{n-1} \sum_{x,y} \frac{\left(f(x,y) - \bar{f}\right)\left(t(tx,ty) - \bar{t}\right)}{\sigma_f \sigma_t}$$

4.Techniques Used(in our process):-

a.Background Subtraction:-

- It is a Gaussian Mixture-based Background/Foreground Segmentation Algorithm.
- It uses a method to model each background pixel by a mixture of K Gaussian distributions (K = 3 to 5).
- While coding, we need to create a background object using the function, cv2.createBackgroundSubtractorMOG().

$$B(x,y,t) = \frac{1}{N}\sum_{i=1}^{N} V(x,y,t-i) \qquad |V(x,y,t) - B(x,y,t)| > \mathrm{Th}$$

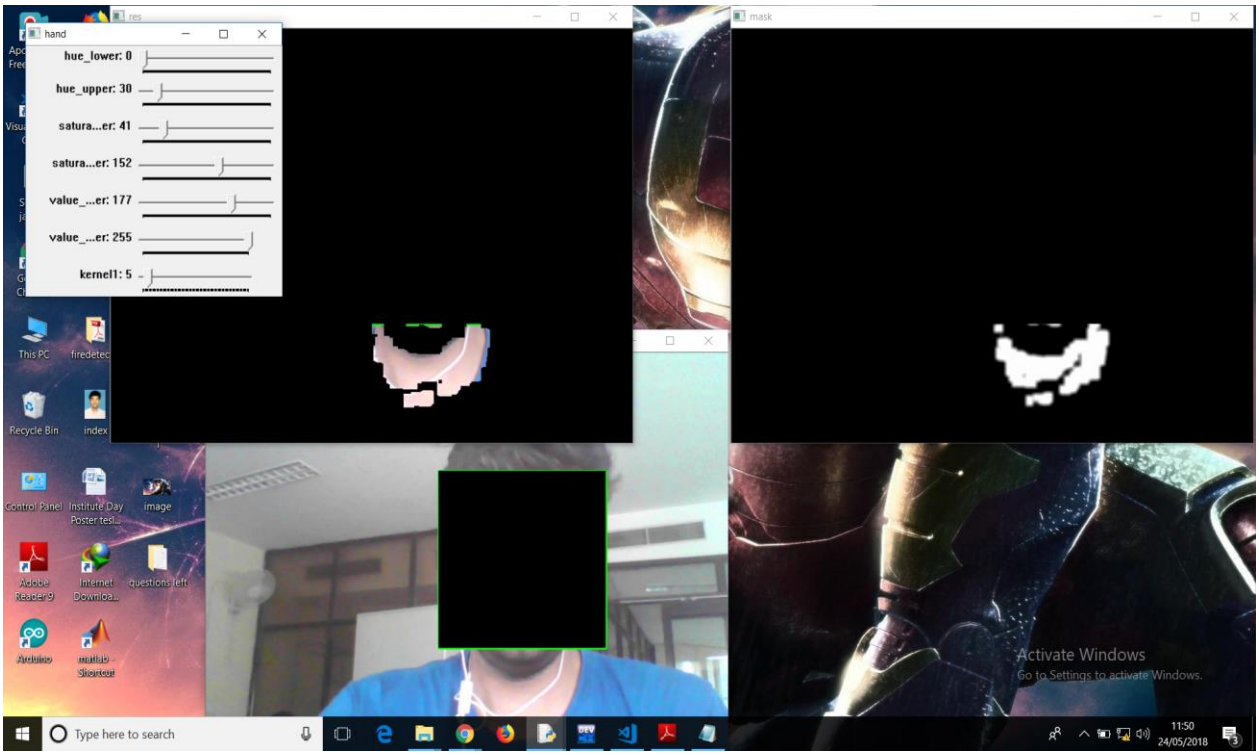


Fig 1. Image showing removal of face and reduced background so as to detect hands correctly

b.Skin & hand Detection:-

Hue Saturation Value (HSV) Color Model

- As H value varies from 0-1, color changes from red, green, yellow and rest back to red.
- S value (saturation) varies from 0-1, from unsaturated (shades of grey) to fully saturated (no white component).
- The corresponding V value varies the brightness.
- **Better than RGB model as getting a specific color is easier here**.

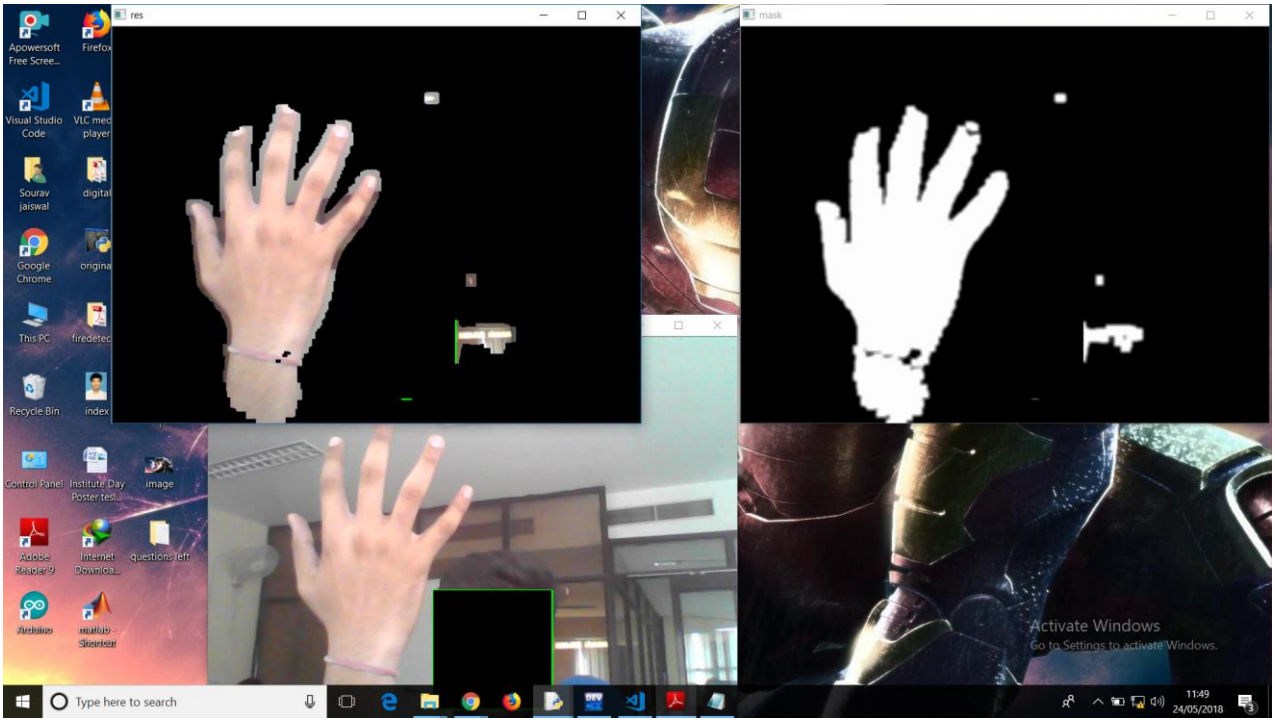


Fig2.Hand detected with removed background and masking face

c.Gesture Recognition:-

Convex Hull(used in opencv)

- cv2.convexHull () function checks a curve for convexity defects and corrects it.
- Convex curves are the curves which are always bulged out, or at-least flat
- If it is bulged inside, it is called convexity defects.

Approximate-Convex-Hull(P,k)

01. $\alpha \leftarrow \pi/k$
02. **for** $i \leftarrow 0$ **to** $k-1$
03. $\quad U_i \leftarrow (\cos(\alpha i+\alpha/2), \sin(\alpha i+\alpha/2))$
04. $\quad U_{i+k} \leftarrow -U_i$
05. $\quad M_{i+k} \leftarrow M_i \leftarrow -\infty$
06. **for each** $p \in P$ **do**
07. $\quad i \leftarrow \lfloor \text{atan2}(p)/\alpha \rfloor$
08. $\quad t \leftarrow \langle U_i, p \rangle$
09. $\quad$ **if** $M_i \le t$ **then** $M_i \leftarrow t$
10. $\quad$ **else** $M_{i+k} \leftarrow \max(M_{i+k}, -t)$
11. $V \leftarrow \langle\ \rangle$
12. $f \leftarrow \text{anglex}(M)$
13. **for** $i \leftarrow f$ **to** $f+2k-1$
14. $\quad$ **if** $M_{i+k} \in (-\infty, 0)$ **then** $V \leftarrow V \cup \langle M_{i+k} U_{i+k} \rangle$
15. $\quad$ **if** $M_i \in (0, \infty)$ **then** $V \leftarrow V \cup \langle M_i U_i \rangle$
16. **return** Melkman-Convex-Hull(V)

Eqn1-equating convex hulls

For example, check the below image of hand [FIG-3]. **Red line** shows the convex hull of hand.
The **double-sided arrow** marks shows the **convexity defects**, which are the local maximum deviations of hull from contours.

PYTHON SYNTAX- hull **=** cv2**.**convexHull (points [, hull [, clockwise [, return Points]]

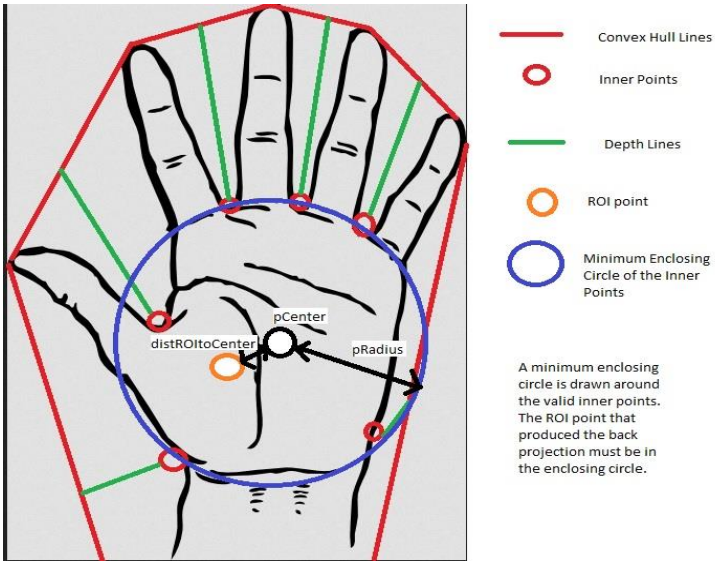


Fig-3 Showing convex hull around hands

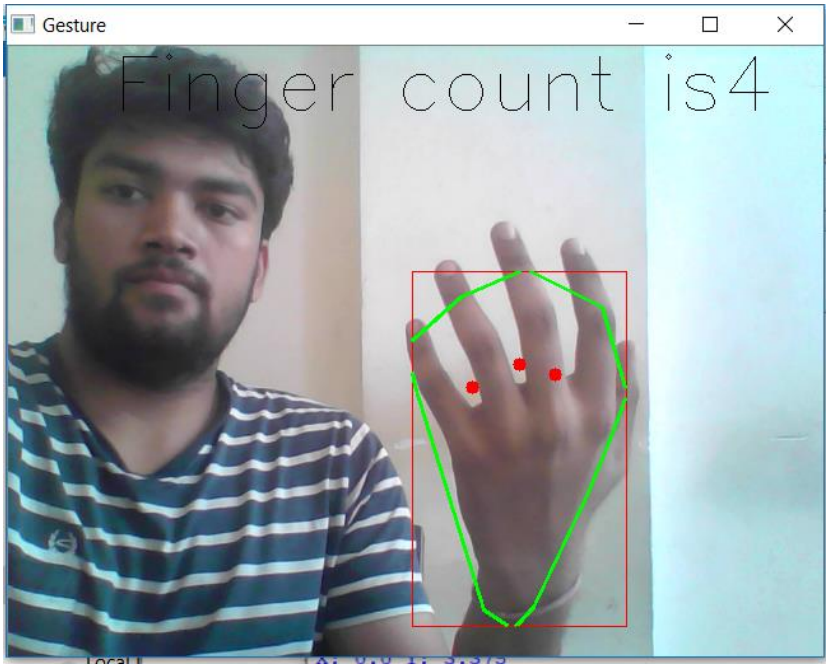


Fig-4 Showing convex hull and Defects in real time

d.Linear Fingertip model(detect using vectors):-

1. Model assumes most finger movements are linear, comprise very little rotational movement.
2. The model uses only the fingertips as input data.
3. Permits a model that represents each fingertip trajectory through space as a simple vector.

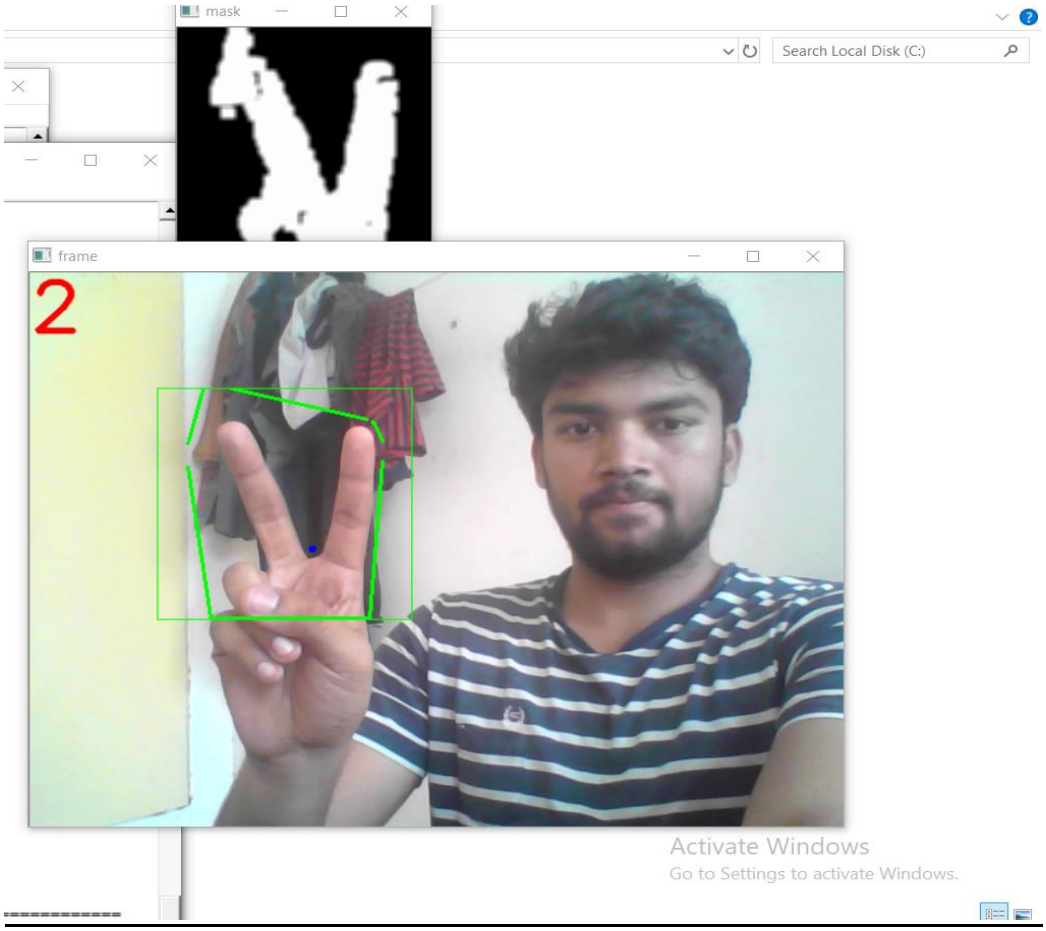
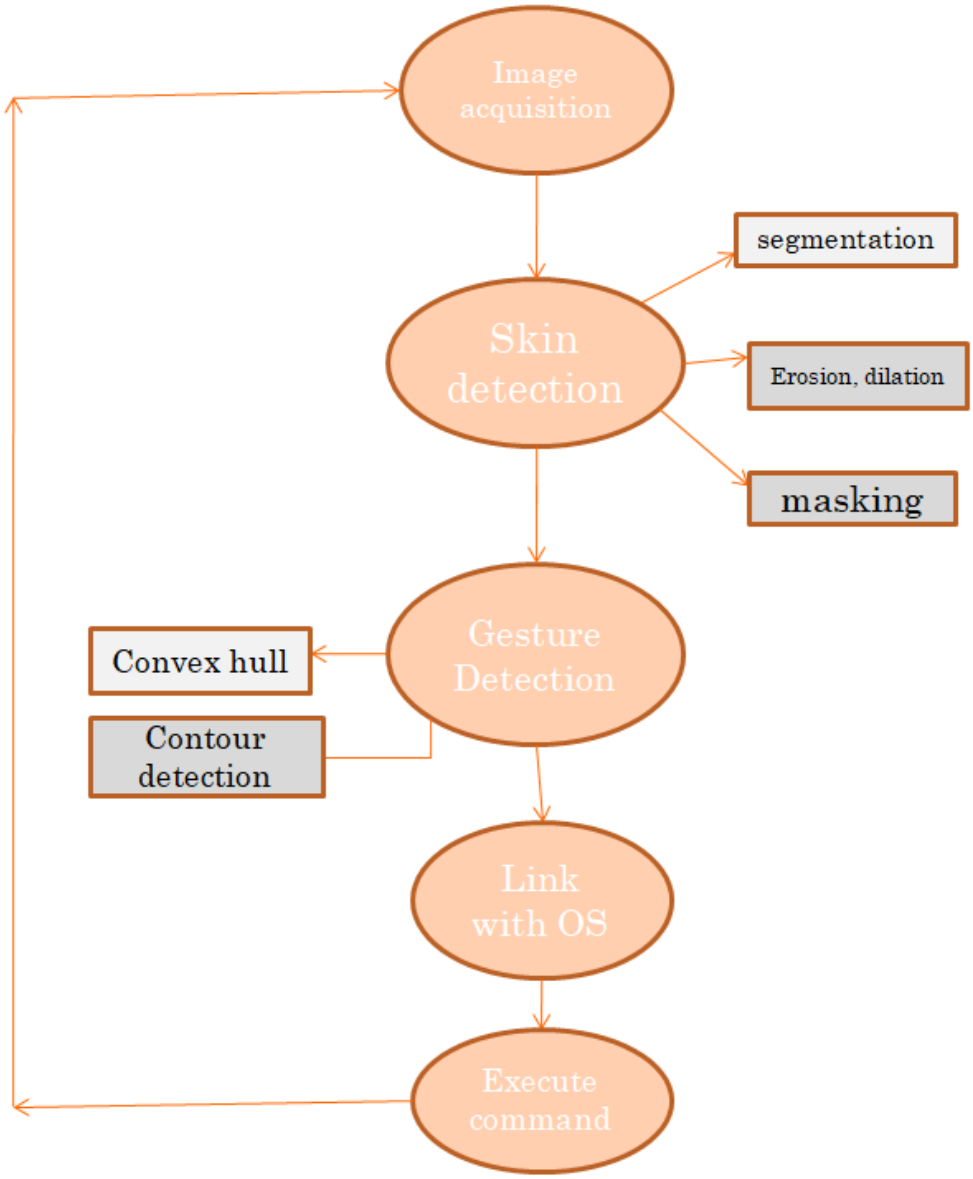


Fig -5 image showing gesture given and correctly recognized.

So All these were techniques used by us for recognition of the gestures.

5 APPROACH TAKEN (OVERALL CYCLIC PROCESS)



This is the overall cyclic process that takes place in real time ,and through this recognition of fingers we can execute various commands in our P.C

6.ANALYSING STEPS:-

The main work here in all this process is the execution part and python language helps in a lot of ways.

We import OS from python and execute various commands in various conditions.

- When the desired gesture comes call **Subprocess.call():-**

  We can call any application from our computer through this command.

Pyautogui.moveRel():-

We can move the mouse cursor to any location through this command.

Pyautogui.click('left")

It is used to press the left mouse button.

We use all these commands in our code whenever the conditions require.

7.LIBRARIES & SOFTWARE USED:-

Language used-PYTHON

Libraries used- There are many libraries of python that has been used-

1. **CV2**-It is the most important python library used

2. **Numpy** and **matplotlib.**

3. Libraries like **math**, time were also used.

4. **Pyautogui** is used for controlling mouse through python.

5. **Subprocesses** and **os** are 2 libraries used for importing system

AND OFCOURSE THE SYSTEM SHOULD HAVE AN INBUILT WEBCAM OR ANY EXTERNAL ONE.

8. RESULT OF RESEARCH:-

We were successful in implementing the problem statement, which consisted the following:-

1. Detecting number of fingers and executing different commands with those gestures.
2. Mimicing the mouse and doing left, right click with different gestures.

These are some pictures of our real time execution:-

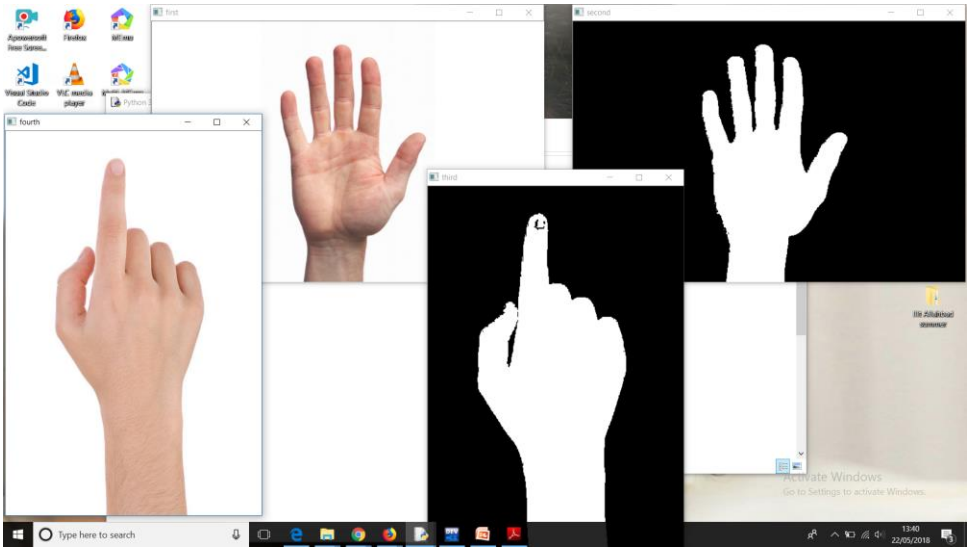


Fig1-showing hand recognition in ideal hand images.

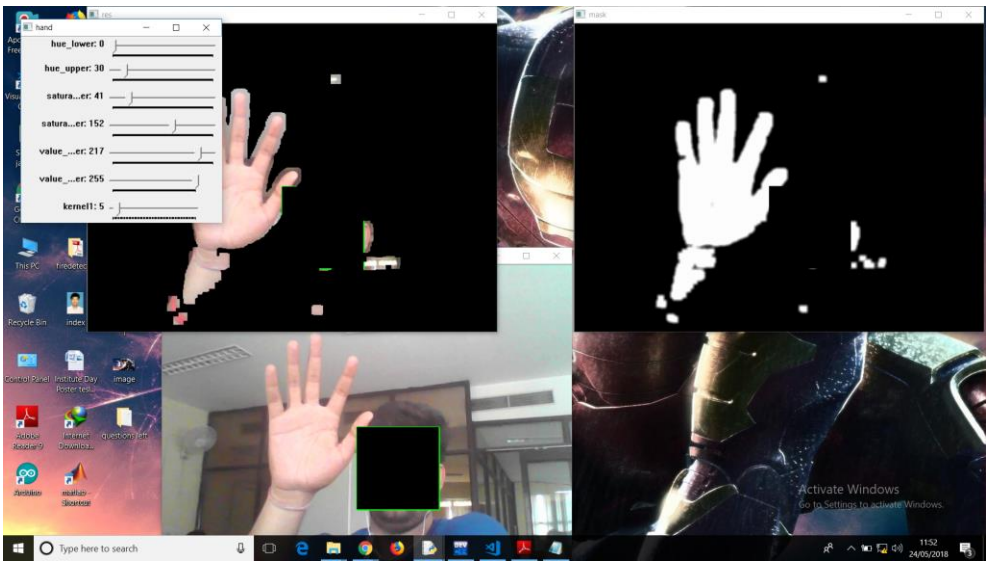


Fig2-showing hand recognition in real time with non ideal background.

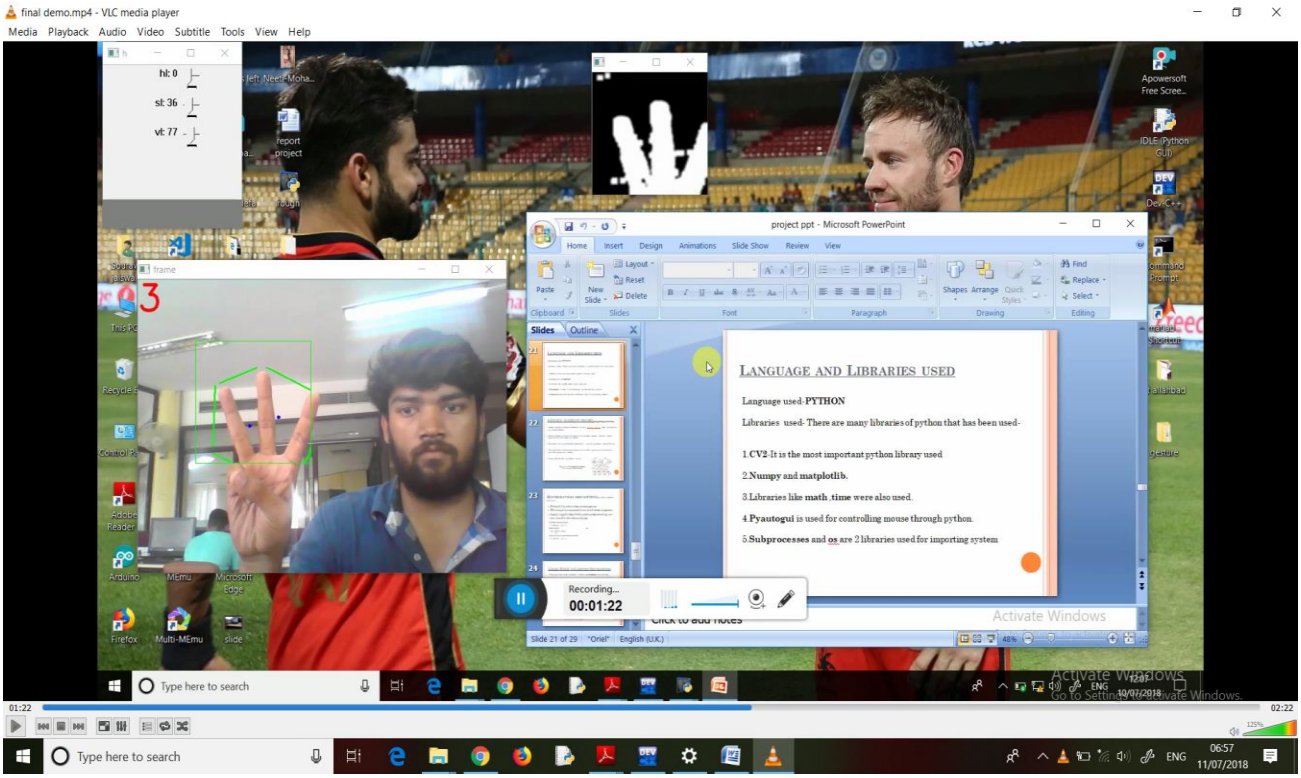


Fig-4 Using hand gesture to scroll up and down
A power point presentation

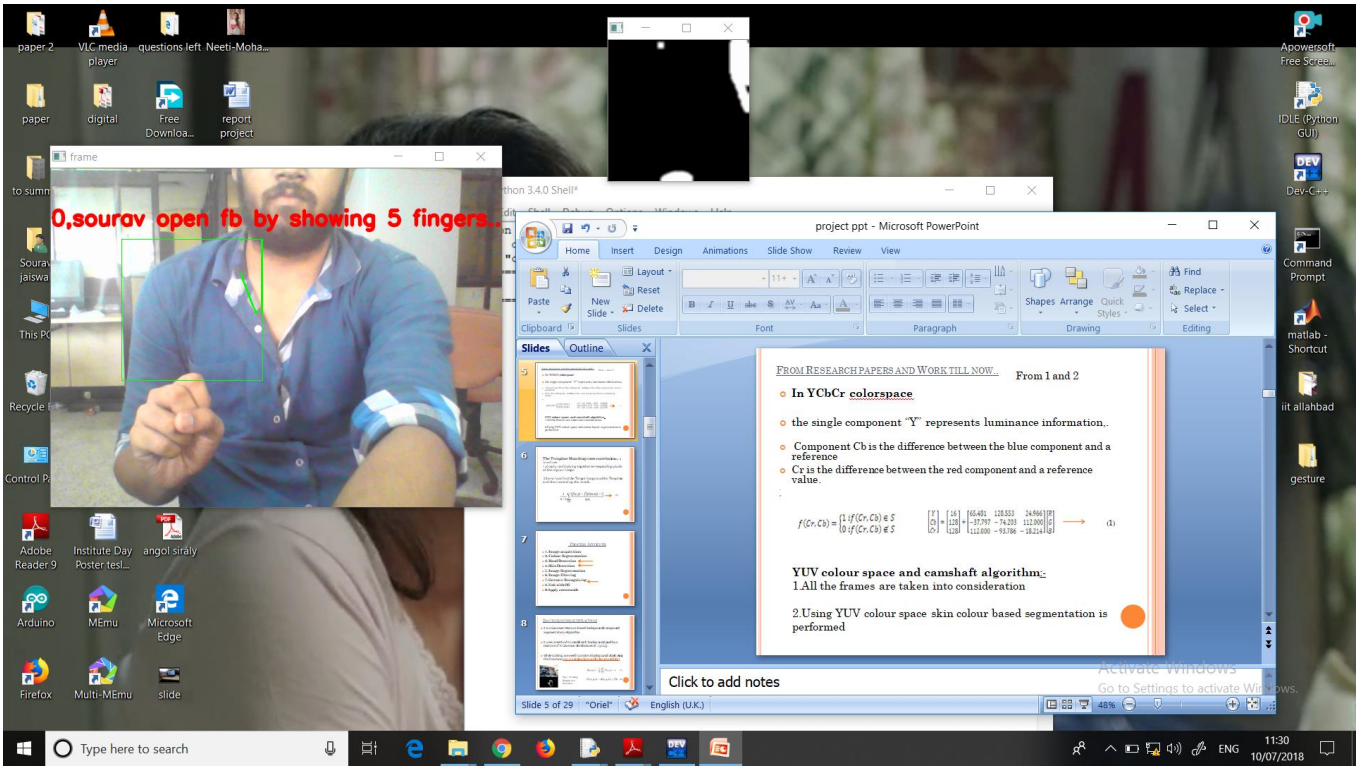


Fig3-Image showing 0 fingers detecting

9.APPLICATIONS:-

- Controlling Robots without any physical contact between human and computer.
- Give presentation without touching mouse pad.
- Control home Appliances like (TV) through hand gestures.
- Suitable for physically impaired people to operate the devices within the room.
- For controlling the traffic signals as per the wish of traffic controller to reduce the problem of traffic jam at peak hours.
- People who can't hear or see or speak can use gestures to control the system and sign language to communicate.

10. IMPROVEMENT(comparison with other research papers):-

1.First and most important improvement is with background reduction techniques,in opencv there are many tools like dilation,erosion, background subtraction, canny edge detection, That makes background subtraction much better compared to other research papers.

2. Other research papers were using RGB model to differentiate hand color,and we are using HSV model to detect which is a much better approach.

3. We are using haarcascade to remove face while detecting skin color, making it more robust.

4. Other research papers have used Template matching for gesture recognition

And we have used convex hull to detect hand and find gestures This is a new technique in opencv where vectors are used to detect number of fingers.

5.We can do multiple operations on our system with hand gesture like mouse mimicking,scroll ppts up and down and all these are complete new implementation of the project.

Overall this process is much more robust and efficient when compared to other hand gesture recognition techniques.

11. FUTURE WORK

As we all know that HAND GESTURE RECOGNITION is such a vast topic for research there are many other tools which we can use to improve our accuracy and robustness.
Firstly we can use CNN to train the model for different gesture recognition and try to implement the training data on real time analysis which definitely will require a good processing unit.

Secondly ,we can implement gesture recognition in mimicking mouse cursor and Scroll the cursor, do left click and right click with hand gesture.

Thirdly, we can implement HGR for sign recognition so that it can be helpfull for blind and deaf people to communicate.

12. REFERNCES

- 1.Hand Gesture Recognition based on Digital Image Processing using MATLAB by Tahir Khan, IJSERVolume 6, Issue 9, September 2015

- 2.HMM based hand gesture recognition: a review on techniques and approaches by MONI , DIU Journal of Science and Technology Volume 5,Issue 2,July 2010

- 3.Hand Gesture Recognition System,by Gaurav Talukdar International Journal of Scientific and Research Publications, Volume 4, Issue 2, February 2014

- 4.A Research Study of Hand Gesture Recognition Technologies and Applications for Human Vehicle Interaction by Carl A. Pickering*, Keith J. Burnham†, Michael J. Richardson*