**Gesture Recognition using Computer Vision using built-in or plugged-in camera device**

**Sri Sai Gowtham Nadakudati**

**Abstract:**

Input devices such as mouse and keyboard are reasonably good for interacting with a computer but as humans, we find it unnatural. Though, the files and folders are present in the digital world, we would feel much more comfortable if we can interact with them using our hands and gestures. With computer vision software and a web camera attached, we can enhance an ordinary computer with gesture recognition capabilities.

In this application, as described in the report, an image is captured from the camera and it is processed using MATLAB. The skin regions are separated, and hand gesture is recognized by capturing features such as finger tips and palm center. Color based segmentation is used to recognize the skin regions to effectively identify the gesture. Our application is able to identify the gestures, but it is not able to issue relevant commands to perform actions such as opening, closing folders etc.

The results show that the application is able to identify the gesture made but it is not able to overcome the inherent problem with color segmentation, i.e., color leakage.

**Introduction:**

The application we have created is a gesture recognition software that can be used as part of a vision-based user interface using a webcam of a computer to ease some of the routine tasks such as opening a file, drag and drop, deleting a file etc. The project will implement recognition of hand features (as static images) to identify the gesture made. An image taken by the web cam will be used as the input and the name of the gesture in the form of text will be the output of the software. As part of the project, after identifying the gesture, the name of the gesture is displayed but the relevant command such as opening a file, deleting a file etc., will not be issued. It can be done by integrating the gesture recognition software as part of other application but that is not in the scope of this project.

**Brief discussion of the approach:**

The image shall be captured using a web cam of a computer. Our end goal is to capture the gesture made by the hand so everything else must be considered as noise. In order to process the gestures made by the hand, we use the color of the skin. Using an effective range of colors that represent skin colors, we can filter out the rest of the user (clothes, hair etc.) and isolate the hands. The face will also be selected as it will fit in the color range.

Now, we need to differentiate between the hands and the face. In this application, we have assumed that the user will always show the gesture in a position next his face. Thus, the larger face area can be removed, and the smaller hand area is left.

The center of the palm area is identified by the following process – for each point on the palm, the shortest distance to the contour of the palm is calculated and the point with the largest distance is the center of the palm.We will further identify the finger tips by calculating the farthest distance from the palm center to the contour and obtaining local maximum points. Based on the no of points identified, we can determine the gesture. For some gestures, such as the closed fist, we will not be able to identify the fingers. So, that will give away the closed hand gesture.

Also, for a fingertip to be identified as one, it’s distance to the palm center has to be more than twice the palm radius. For the sake of simplicity, we will assume that a proper gesture should be done with palm facing the direction of the camera.

**Literature Survey:**

* Yeo, H, Lee, B, & Lim, H 2015, 'Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware', *Multimedia Tools & Applications*, 74, 8, pp. 2687-2715, Applied Science & Technology Source, EBSCO*host*, viewed 11 February 2018.

This paper deals with some of the structured approaches in determining the gestures made. In this paper, a hand tracking system was developed that can detect the motion and translate it to related functional input to other units. Several problems related to hand feature detection are discussed in this paper. Complex background, motion blur and face detection are some of the topics discussed.

* Nasri, S, Behrad, A, & Razzazi, F 2015, 'A novel approach for dynamic hand gesture recognition using contour-based similarity images', International Journal of Computer Mathematics, 92, 4, pp. 662-685, Applied Science & Technology Source, EBSCOhost, viewed 25 February 2018.

This paper deals with a different approach for gesture recognition. It uses the contour-based similarity between images for dynamic gesture recognition. The difference between the position of hand in one frame and the position in next frame is used to determine the motion of hand and thus determine the gesture. In order to differentiate between different gestures, it calculates the centre of mass of various regions in the image and uses it to tell one gesture from the other.

This paper deals mostly with continuous gesture recognition but with a database of gestures, it can also be used for disjointed gesture recognition.

* Hasan, H, & Abdul-Kareem, S 2014, 'Static hand gesture recognition using neural networks', Artificial Intelligence Review, 41, 2, pp. 147-181, Applied Science & Technology Source, EBSCOhost, viewed 25 February 2018.

This paper explains in detail about gesture recognition using neural networks. The paper goes on to discuss various difficulties encountered in computer vision such as color segmentation, tracking, feature extraction and classification. It explains various steps involved and benefits of using neural networking over traditional programming.

* Simen Andresen, Martin Stokkeland, Vegar Østhus 2013, ‘Hand Detection Using Color Recognition: Object Tracking and Gesture Recognition - Shortened Version’, <http://simena86.github.io/blog/2013/08/12/hand-tracking-and-recognition-with-opencv/>

This online resource is explains various steps involved in gesture recognition using color segmentation and how they can be implemented in C++. Functions required for feature extraction are present here. It will be used as a reference to implement various functions in MATLAB.

**Algorithms and functions used:**

**Canny edges detector:** In order to improve the segmentation of the image, a Gaussian filter is applied to the image and Canny edges detector is used to identify the edges. This seperates various parts in the image nad alloes for better segmentation.

**Dilation:** We have used MATLAB’s imdilate function to connect nearby components in a segmented image to create a larger structure.

**Fill Area:** We have used MATLAB’s imfill function to fill holes in our segmented image when color segmentation was not very effective and irregular.

**Erosion:** We have used MATLAB’s bwareaopen function to remove small blobs of segments from our segmented image.

**Connected Component Analysis:** In order to work with the segmented images components, we have performed connected component analysis using MATLAB’s bwconncomp function.

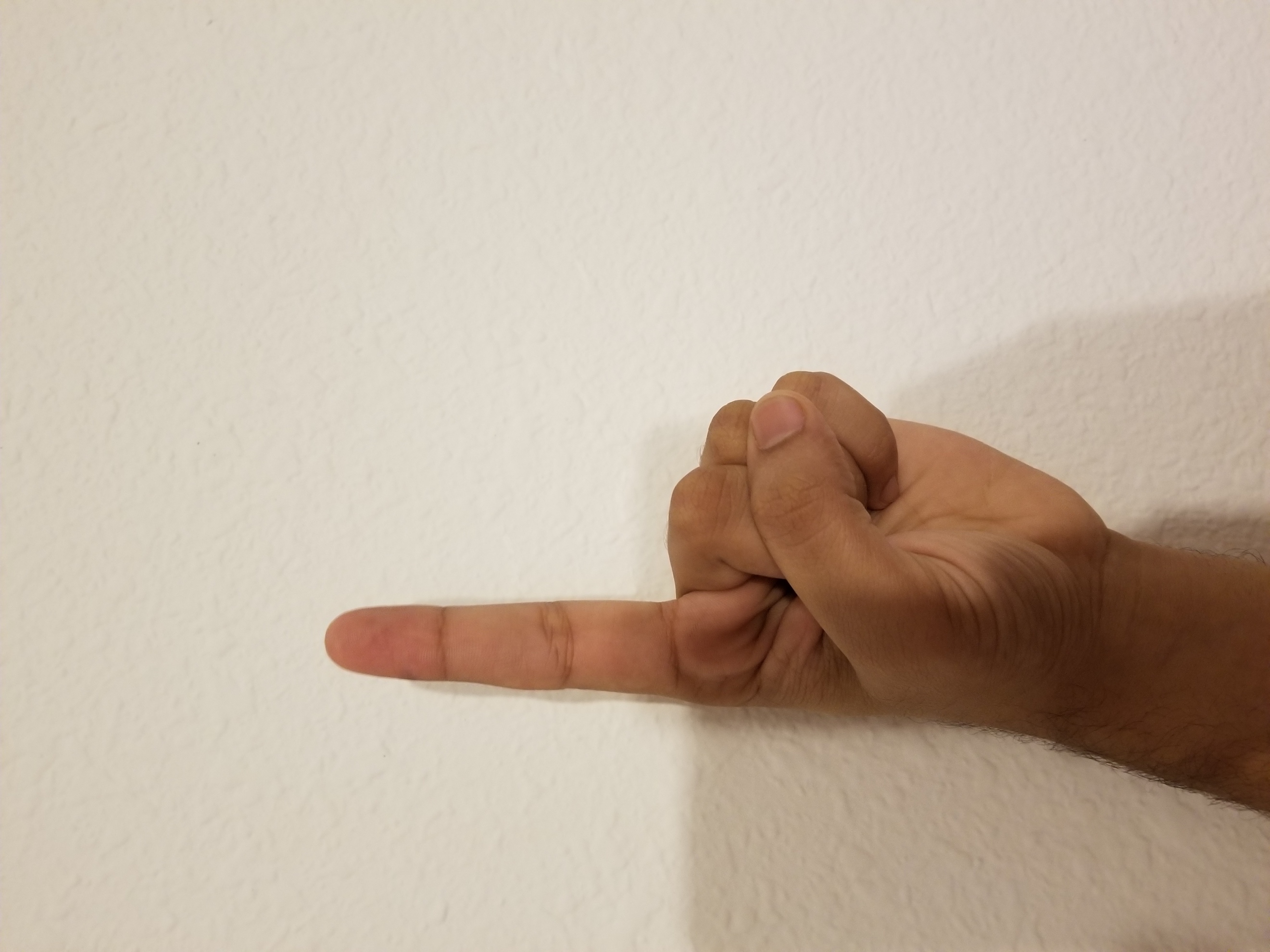
**Experimental results:**

I have used my application to find the following gestures:

1. Closed Fist (Zero)



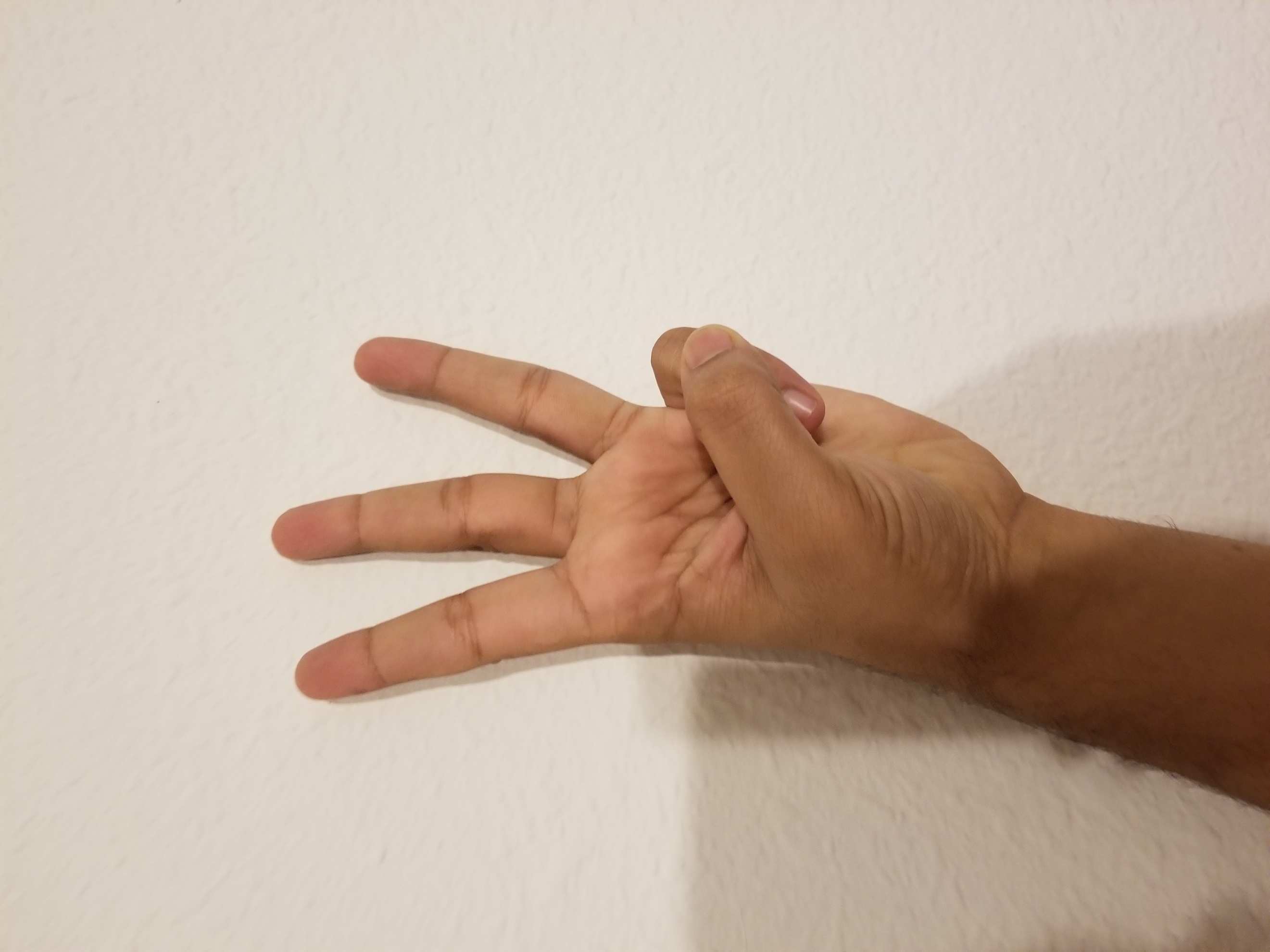
1. Pointed Index finger (One)



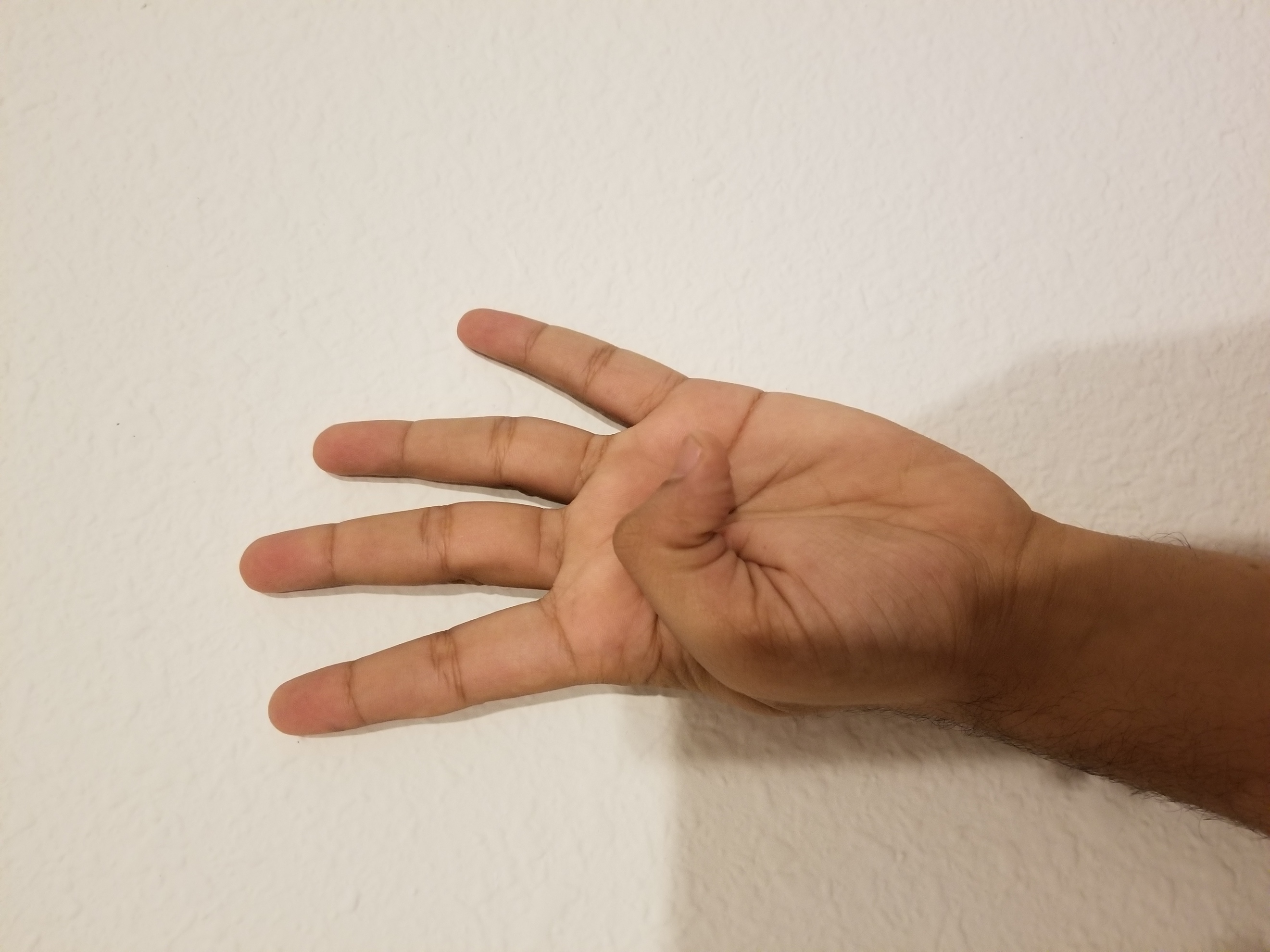
1. Victory symbol (Two)



1. Index, middle and ring finger pointed upwards (Three)



1. Open palm with closed thumb (Four)



1. Open palm (Five)



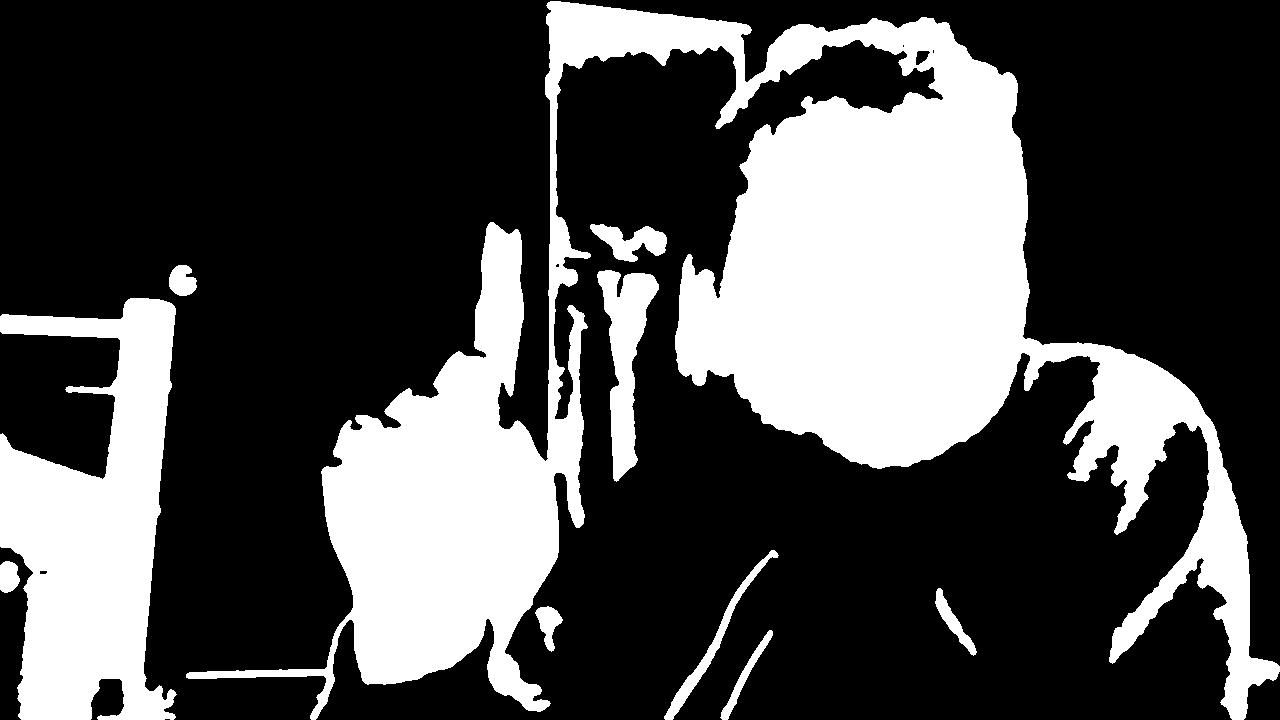
The experiment was conducted indoors with CFL lighting. The lighting conditions are important to our application as we are also performing luminance-based segmentation. For gesture One, the results are as follows:



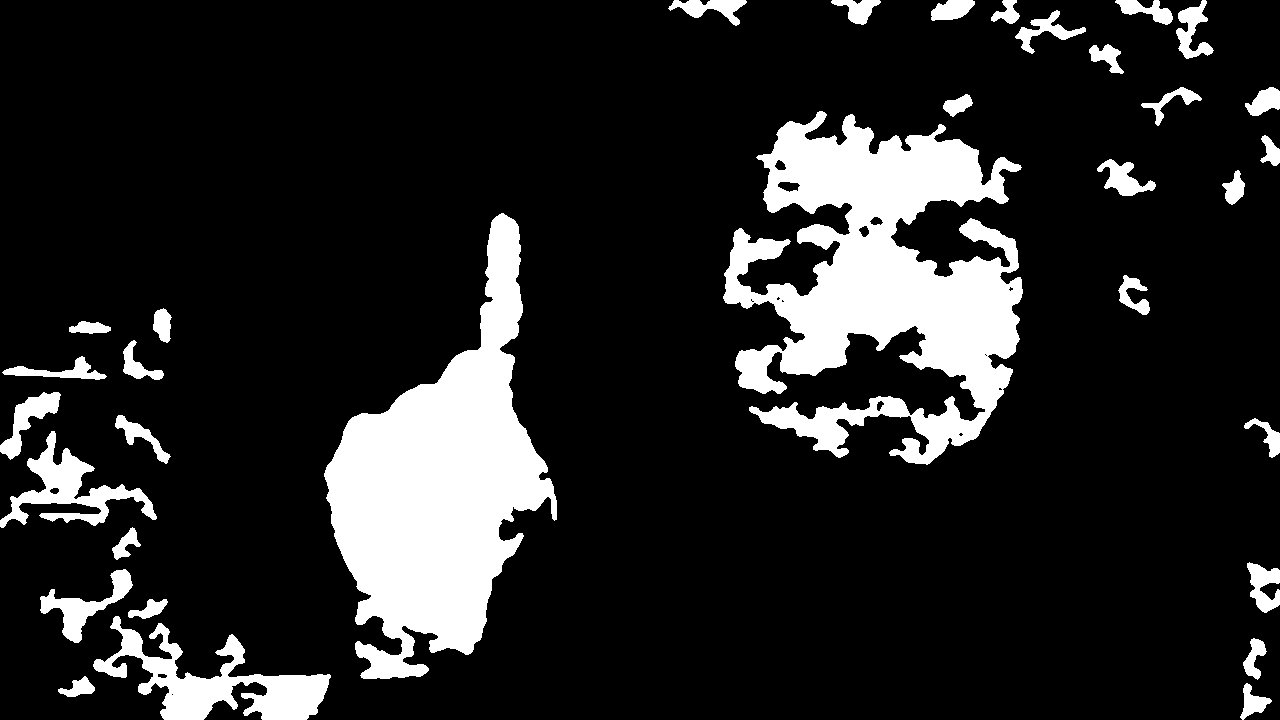
Input from camera – YcbCr image



Input converted to RGB for representation only



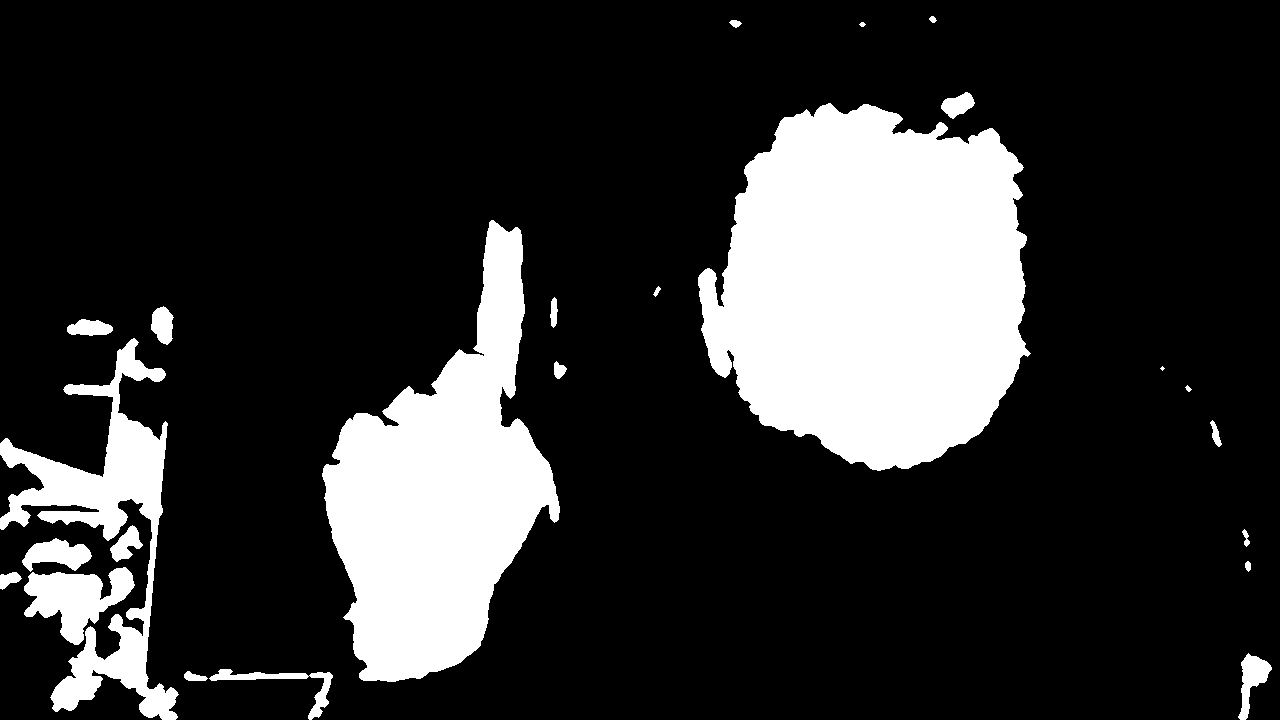
Segmentation by luminance



Segmentation by Blue chrominance



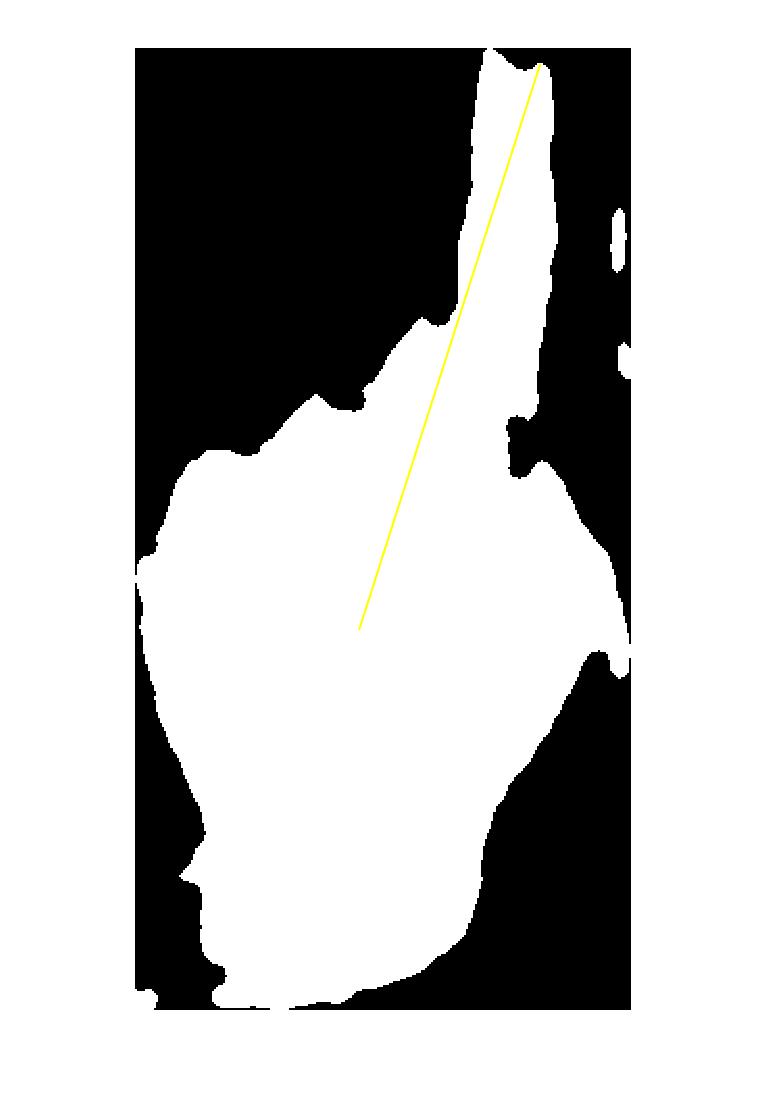
Segmentation by Red chrominance



Combined results of segmentation – binary image



Outline image of the recognised hand



Finger tip being recognised

**Console output:**

No of fingers = 1

The results for the remaining gestures are in the zip file.

**Discussion of results:**

As we can see from the results, the images are not perfectlt segmented. When the image was segmented into three channels, the binary results from luminance and chrominance show that some small parts of the skin was not selected. I have modified the code several times to try to be fully inclusive but the it was always leaving holes in different ligting conditions.

Many other objects were also being included in the image. The frame of the bed in the background was red in colour and it was to some extent recognised as skin. The texture of the walls in the lighting conditions are also sometimes recognised a skin. This is the inherent problem of color segmentaion, color leakage.

Though the image was able to filter the face region, the condition is that face region should be bigger than hand region for proper results. This means that if we place our hand a little closer to the camera such that it appears bigger than the face, we are not able to get proper results.

**Conclusions:**

The application is therefore not as robust as it should’ve been. Sometimes the results show that more than 5 fingers are present. This is due to failing to segment the regions properly or color leakage. One thing that can be definitively improved is face removal mechanism. I wanted to implement openCV libraries to use Voila-Jones face detector to effectively remove faces but I was unable to implement it. I would work in that direction to improve it.

Further, it takes a long time to compute the finger tips and produce results. With GPU programming, this problem could be solved. Hoever since many computers don’t have dedicated GPUs, we can look to improve the algorithm to reduce time needed.

**References:**

* Yeo, H, Lee, B, & Lim, H 2015, 'Hand tracking and gesture recognition system for human-computer interaction using low-cost hardware', *Multimedia Tools & Applications*, 74, 8, pp. 2687-2715, Applied Science & Technology Source, EBSCO*host*, viewed 11 February 2018.
* Simen Andresen, Martin Stokkeland, Vegar Østhus 2013, ‘Hand Detection Using Color Recognition: Object Tracking and Gesture Recognition - Shortened Version’, <http://simena86.github.io/blog/2013/08/12/hand-tracking-and-recognition-with-opencv/>
* Arindam Sarkar’s blog: <http://codetocreate.blogspot.in/2014/07/how-to-detect-fingers.html>