B. TECH. PROJECT REPORT

On

Biometrics: Development of a Robust Algorithm for Segmentation of Human Hand from Cluttered Backgrounds

BY

Bapat Akshay Sudhir

DISCIPLINE OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY INDORE

December 2016

Biometrics: Development of a Robust Algorithm for Segmentation of Human Hand from Cluttered Backgrounds

A PROJECT REPORT

Submitted in partial fulfilment of the

requirements for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

ELECTRICAL ENGINEERING

Submitted by:

Bapat Akshay Sudhir

Guided by:

Dr. Vivek Kanhangad, Assistant Professor, IIT Indore

INDIAN INSTITUTE OF TECHNOLOGY INDORE

December 2016

CANDIDATE'S DECLARATION

I hereby declare that the project entitled **"Biometrics: Development of a Robust Algorithm for Segmentation of Human Hand from Cluttered Backgrounds"** submitted in partial fulfilment for the award of the degree of Bachelor of Technology in '**Electrical Engineering**' completed under the supervision of **Dr. Vivek Kanhangad, Assistant Professor,** IIT Indore is an authentic work.

Further, I declare that I have not submitted this work for the award of any other degree elsewhere.

Signature and name of the student with date

CERTIFICATE by BTP Guide

It is certified that the above statement made by the students is correct to the best of my knowledge.

Signature of BTP Guide with date and their designation

Preface

This report on “Biometrics: Development of a Robust Algorithm for Segmentation of Human Hand from Cluttered Backgrounds" is prepared under the guidance of Dr. Vivek Kanhangad.

Through this report, I have tried to give a detailed account of an innovative technique which can allow the usage of several hand-based biometrical identification methods, which currently work only with uniform backgrounds, in scenarios of non-uniform backgrounds.

I have tried to the best of my abilities and knowledge to explain the content in a lucid manner. I have also added stage-outputs and graphs to make it more illustrative.

Bapat Akshay Sudhir

B.Tech. IV Year

Discipline of Electrical Engineering

IIT Indore

Acknowledgements

I wish to thank Dr. Vivek Kanhangad for his kind support, valuable guidance and allowing me to explore multiple approaches for the completion of this project. It is his help and support, due to which I was able to complete the project and technical report. Without his support this report would not have been possible.

Bapat Akshay Sudhir

B.Tech. IV Year

Discipline of Electrical Engineering

IIT Indore

Abstract

While hand geometry biometrics have been widely used in order to carry out person identification and authentication, majority of the methods used are against a uniform background. With the rise in number of contactless hand geometry methods, it is imperative that the hand geometry identification techniques work in the commonly encountered cluttered backgrounds as well. We make the following contribution – we enable hand recognition techniques to work on cluttered backgrounds as well, by which the pre-existing hand-based biometrics techniques can be applied practically everywhere.

This report describes a multiple-stage method for segmenting human hands from a cluttered image. The biometric identification techniques based on hand geometry consider the palm of the hand to determine identity. Segmentation the palm region from an image is done using a three step process. The first stage uses skin-based segmentation to determine the areas in the image which may contain skin pixels. This stage allows a wide range of skin-like colors to be recognized as skin in order to reduce the number of false negatives caused by conditions such as shadows and bad lighting. During the second stage, shape characteristics like solidity, orientation, and eccentricity are determined for various skin regions and compared to the mean shape characteristics derived from a dataset consisting of large number of hand images. The final stage determines whether a region is likely to be a hand by using a novel method which utilizes the unique shape of the hand. If a region is determined to be a hand, the location of the wrist is determined so that the region above the wrist can be segmented.

Contents

List of figures

1

Introduction

1.1. Biometrics

Biometrics is the measurement and statistical analysis of people's physical and behavioral characteristics. Biometrics is widely used to determine distinctive, measurable characteristics used to identify individuals and verify their identity. Biometrics, thus, is used for authentication and security in various companies and institutions. Biometrics has several advantages over conventional password and PIN based systems. First, a biometric does not need to be remembered and cannot be easily lost, making it much easier for the user. Furthermore, a biometric cannot be easily stolen or borrowed. This makes it more secure from a system point of view. Moreover, a biometric typically has a higher information content than a password thus making it harder for a hacker to crack such a system using brute force. Various techniques have been developed for biometric identification - fingerprint recognition, speech recognition, DNA recognition, signature recognition, iris recognition etc.

The origin of the implementations of biometrics can be traced back to periods thousands of years ago. In a prehistoric cave dated thirty one thousand years old, handprints of prehistoric men can be found adjacent to their paintings and are thought to be an unforgettable mark of the painter**.** In modern times, as cities and the number of city-dwellers grew due to the industrial revolution and improved agricultural practices, law enforcement agencies found it important to recognize repeat offenders so that appropriate punishments could be ordered. In this effort, two identification techniques became widely popular.

The first of two approaches, which developed in France, was the Bertillon system, in which various body dimensions were measured and noted. It was possible to identify individuals by sorting through these records and finding a match. The second approach, which is still widely popular today, was the formal use of fingerprints by police departments. All recorded fingerprints could be indexed and an appropriate match could be identified using advanced metrics such as fingerprint patterns and ridges.

Biometrics can be divided into two broad categories; first, biometrics based on physiological characteristics and second, biometrics based on behavioral characteristics. Physiological biometrics is based on data derived from direct measurements of parts of the human body. Fingerprints, iris scans, retina scans, hand geometry, and facial recognition are all leading physiological biometrics. Behavioral biometrics is the field of study related to the measure of uniquely identifying and measurable patterns in human activities. Examples of Behavioral biometrics can be keystroke-based, signature-based or voice based.

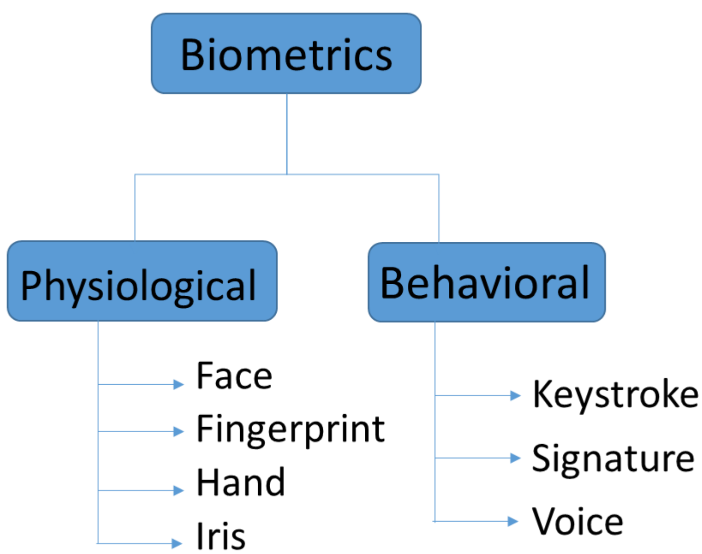

Figure 1. Types of Biometrics

Nowadays, several biometric identification techniques can be found in use in a variety of places. Iris-pattern and retina-pattern authentication methods are used in some banks for automatic teller machines. Voice waveform recognition is used for enabling access to proprietary databanks in research facilities. Facial-recognition technology is used by law enforcement to pick out individuals in large crowds with considerable reliability. Earlobe geometry is commonly used to disprove the identity of individuals who claim to be someone they are not, i.e. identity theft. Identification using hand geometry is being used in industry to provide physical access to buildings.

1.2. Hand Geometry Biometrics

Among the methods used for biometric identification, hand-based biometrics is widely used in both industry and government due to its various advantages over other biometric techniques, some of which are easier data acquisition, contactless verification, and low effect of intrinsic and environmental factors. On the other hand, several other techniques are not deemed appropriate for some applications due to a variety of reasons. For example, fingerprint-based techniques have criminal stigma attached with them, while iris and retinal scans require expensive and intrusive systems. Even though these systems may be more accurate, they might not be ideal for daily, low-risk applications such as attendance in colleges. For applications which are not high-risk, hand geometry biometrics can be an ideal low-cost identification alternative.

Hand geometry is the measurement and comparison of the different physical characteristics of the hand. These characteristics include the size of the hand, the corresponding size of the base of the palm, the thickness and length of the fingers and so on. Hand-based identification techniques are suitable where reasonable accuracy is necessary and user-friendly, contactless and non-intrusive methods are required.

Unlike fingerprints, which can be copied with some effort, it is very difficult to copy hand geometry. Even though people have similar hand geometries, it is very difficult for anyone to find an identical match for the hand geometries of a person. Also, if hand geometry techniques are used in conjunction with a more secure technique such as fingerprint scanning or iris scanning, the level of security of the biometric system can be improved immensely.

However, almost all hand geometry techniques require that only the hand is present along with a uniform background. These techniques, currently, do not work in the presence of background clutter. This restricts the applicability of hand based biometrics, which is why we propose a method which can identify and segment the palm region – which is required for hand based biometrics – and remove all the background clutter. Using our technique, hand based identification techniques can be used anywhere irrespective of the background, and without using any extra apparatus. In practical applications, clutter is often present due to a variety of factors – uneven background lighting, background clutter etc. Hence, techniques developed on uniform backgrounds may not work in all conditions. Generally, there is some kind of clutter in the background which causes the background to be non-uniform. For example, while a person trying to validate his identity by holding his hand in front of a camera, his face, clothes, people in the background etc. will be visible in the captured image and will constitute the clutter in the background.

Thus, to ensure that all kinds of images can also be used for hand geometry identification, a method is proposed which acts as a decluttering algorithm – to identify the part of the image in which the hand lies and segment the hand region. The algorithm will produce an output which will contain only the hand, free of any background clutter. This will ensure that calculations for hand geometry can be done for cluttered backgrounds as well, thus increasing the scope for hand geometry biometrics. Every existing technology for hand geometry which requires a uniform background, if clubbed with this decluttering algorithm, can be used in cluttered background conditions as well.

2

Related Work

Several approaches have been studied from segmenting hand from non-uniform backgrounds, either for hand geometry biometrics or for gesture recognition. Many methods for hand geometry have been created considering that the image of the hand will be taken using a scanner, that is, against a plain background [4], [5], and [6]. However, after the work on methods which were contactless [3], [13], [21], and [22], non-uniform backgrounds were also encountered in the image. Thus, methods were created which would work in cluttered backgrounds as well [3], [5], and [9].

Irrespective of the type of the objective and the conditions required for a method to work, the methods themselves can be said to consist of these general steps –

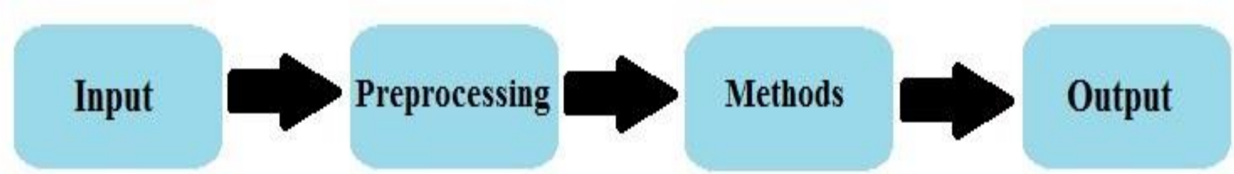

Figure 2. General steps to segment hand from cluttered backgrounds

The captured image is fed to an algorithm which carries out preprocessing. Some of the preprocessing methods -

- Averaging
- Thresholding
- Contrast adjustment
- Denoising

All existing research for segmentation of human hand from cluttered backgrounds is a combination of these three methods –

- Skin segmentation
- Background subtraction

2.1. Overview of existing techniques

The existing techniques are primarily done using skin-based segmentation. Even though skin-based segmentation is used in this project, it is only the first stage; also, the skin segmentation algorithm has been modified in order to get better results and enable the method to work in complex backgrounds as well. The uniqueness and value of the proposed technique can be seen in the following example.

For an image in uniform background, skin-based segmentation is very effective, as seen from an image taken from a dataset available for research. [27]

Figure 3. Image with uniform background

Figure 4. Skin-segmented image with uniform background

For an image with cluttered background, skin-based segmentation techniques are not very effective, as seen from an image taken from the dataset we have created. Presence of brown, skin-colored or off-white colors in the background severely affects the quality of the skin-detection and leads to a lot of false positive instances of skin.

Figure 5. Image with cluttered background

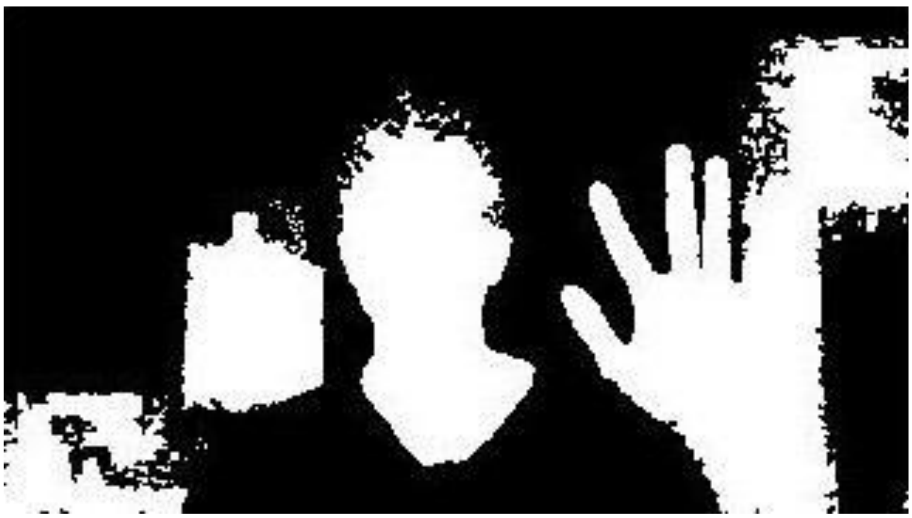

Figure 6. Skin-segmented image with cluttered background

2.2. Skin segmentation

Most existing skin segmentation techniques involve the classification of image pixels into skin and non-skin categories based on the pixel values [17], [18], [19], and [20]. The basic idea underlying this approach is that human skin generally consists of distinct colors irrespective of lighting. Skin color is restricted to a certain values and some colors cannot be skin irrespective of the conditions at that time. Hence, it is possible to classify any given pixel as skin and non-skin pixels by applying a simple conditional filtering.

There exist numerous color spaces which are actively used for research but many of them share similar characteristics. These models are generally used to determine whether a pixel is a skin pixel or a non-skin pixel. We focus on four representative color spaces which are commonly used in the image processing field:

1. **RGB:**
    - Colors are specified in terms of the three primary colors: red (R), green (G), and blue (B).
    - This is the most commonly used color space in image processing.
    - However, this model does not give good results for skin detection on its own.
    - It can be used with HSV color space to get better skin detection.

2. **HSV:**
    - Colors are specified in terms of hue (H), saturation (S), and intensity value (V).
    - The transformation between HSV and RGB is nonlinear.
    - Other similar color spaces are HIS, HLS, and HCI.
    - This model gives very good results and is used for this project.
    - It can be used with RGB color space to get better skin detection.

3. **YCbCr:**
   - Colors are specified in terms of luminance (the Y channel) and chrominance (Cb and Cr channels).
   - The transformation between YCbCr and RGB is linear.
   - Other similar color spaces include YIQ and YUV.
   - This model works well for skin detection; however, HSV gives better results

4. **CIE-Lab:**
   - Designed to approximate perceptually uniform color spaces
   - The transformation between CIE-Lab color space and RGB color space is highly nonlinear.
   - Other similar color spaces are CIE-Luv and Farnsworth UCS.
   - This model works well for skin detection; however, HSV gives better results

In the recent years, a combination of two or more of these models have been used together to improve the existing performance for skin detection. We have used a combination of HSV and RGB color models for the skin detection part in this project, with the focus on HSV color model. HSV color model has been found out to be ideal for skin segmentation due to the invariance of the Hue and Saturation quantities with lighting or shadows. However, in case of an RGB model, a slight increase in lighting causes all the R, G, and B values to increase and it becomes very difficult to express the limits in which the skin values lie for RGB color space.

2.3. Background subtraction

Background subtraction is commonly used for detecting movement while the rest of the background is still [15], [17]. Background detection approaches can be classified as non-adaptive and adaptive methods. Examples of non-adaptive methods are manual selection, pixel-wise voting, and mean value search algorithms. Adaptive methods include averaging images over time, alpha-blending, Kalman filtering, etc. Generally, the way in which background subtraction methods are used for hand segmentation is that first an image is taken without the hand, the second image also includes the hand. Thus, the hand region is determined as the region which was not present earlier but is present now.

This method can detect the areas which were absent in the earlier image. That is, if a person takes a photo and then puts up his/her hand in the image, the algorithm will detect that the hand is the region of interest. While this works in some cases, it is very difficult for a person to stay still while taking a photo and slight errors occur due to such movements. Also, these techniques find all the areas in the image which have moved; hence, it is not necessary that the second image will contain only the palm. Often, whole of the hand above the elbow is visible in the second image and these methods do not employ a system which will only segment the palm of the hand. Thus, the existing methods do not work well for the segmentation of hand from cluttered backgrounds.

Methods involving background subtraction are not very techniques are not very effective in case of humans due to slight movements, as seen from an image taken from the dataset we have created.

Figure 7: Background image

Figure 8: Background image with hand

Figure 9: Result after background subtraction

An example of how background subtraction works in ideal cases. In the figure on the left, the background is present. In the figure on the right, the image is the same except the presence of the boat. Thus, on background subtraction, the boat will be detected and segmented as being the object of interest.

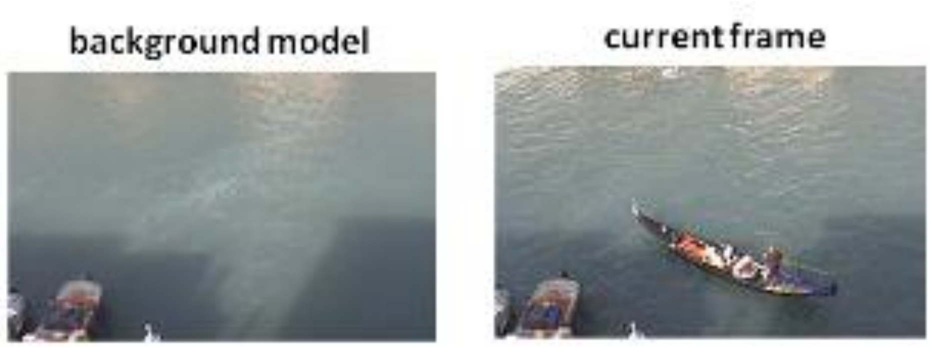

Figure 10. Background, background with object of interest

3

Proposed Work

3.1. Motivation

Segmentation of hand from cluttered background will be very useful to hand geometry based biometrics identification techniques which work accurately for uniform backgrounds but do not work well in case of background clutter.

Another perspective to look at the problem is to enable methods which work on plain backgrounds to work on cluttered backgrounds as well. This can be done by removing the background clutter or segmenting the hand region from the background. In that regard, our algorithm segments the hand from any background and the output is an image which contains the hand on a uniform background. Thus, we have created a system that can enable existing hand geometry and palm print based methods to work on cluttered backgrounds as well. Thus, the existing hand geometry based methods will no longer be restricted by the constraints of plain backgrounds, leading to increased applicability of the said methods. Another advantage is that setup like screens which have been used for these methods to have a uniform background are no longer required.

3.2. Steps taken

Segmentation the palm region from an image is done using a three step process. The first stage uses skin-based segmentation to determine the areas in the image which may contain skin pixels. This stage allows a wide range of skin-like colors to be recognized as skin in order to reduce the number of false negatives caused by conditions such as shadows and bad lighting. During the second stage, shape characteristics like solidity, orientation, and eccentricity are determined for various skin regions and compared to the mean shape characteristics derived from a dataset consisting of large number of hand images. The ranges in which these shape characteristics could lie to be considered as hand were found out by taking the mean of the shape characteristics as found in the dataset and taking a fixed region around the mean as the allowable region. The final stage further determines whether a region is likely to be a hand by finding the variation in the number of skin pixels along the row in the image, starting from the topmost pixel of the region; if it is, the part of the region which is very likely to be the palm is found out and segmented.

4

Analysis

4.1. Skin Detection

The problem of segmentation of hand from cluttered background will be solved in two stages, skin-based segmentation followed by a shape-based segmentation. For the first stage, a skin-based segmentation is used to determine whether or not a given pixel is a skin pixel. The decision for skin classification, i.e. determining whether a particular pixel should be classified as a skin pixel or as a non-skin pixel, is based on several factors. Some of these factors are pixel values in HSV as well as RGB domain, the pixel's proximity to other skin pixels and whether the number of pixels in its group of supposed skin pixels is greater than the minimum threshold.

In HSV space, H denotes Hue, i.e. shade of the colour, S denotes Saturation, i.e. purity of Hue while V denotes Value, i.e. brightness. In images with good lighting, skin pixels lie in a certain range of values for Hue, the approximate ranges for which have already been found earlier by other researchers. During the initial stages of development of the decluttering algorithm, Hue ranges for skin which have already been determined earlier by others, had been used. After checking the performance of the skin detector on numerous images, several modifications were made to the values and more accurate hue ranges for skin classification were determined. It was also observed that Saturation values also lie in a fixed range for skin pixels. Moreover, several limits on the RGB values in the image were also determined by collecting data over a large number of pixels.

After modifying the algorithm to incorporate all these aspects into skin detection, namely HS and RGB values, a better algorithm was developed which had very high detection rates and very low false positives. Here, false positives refer to those pixels which are detected as skin pixels even though they do not belong to the skin. Furthermore, in order to have a more accurate range for HSV values of skin pixels, the pre-existing, extremely accurate face detection techniques of the Viola-Jones algorithm was used to find the faces in the images. From these detected faces, skin pixels were separated using the developed skin-detector algorithm. These skin pixels were used to further narrow down the ranges of pixel values for skin. The ranges of pixel values for skin thus obtained were used in modifying the skin detection algorithm, leading to increased accurate. Thus, the first stage of the project, which aims at accurately detecting skin, was completed.

4.2. Shape Characteristics

The second stage of the project incorporated shape information for segmenting the hand region. However, due to the lack of precise tools available to represent shapes, this stage was completed in two steps, using several shape characteristics which are useful in defining shape and using a method which can be used to appropriately represent all hand-like shapes. The shape characteristics which were used for the first part are as follows -

- Solidity
- Circularity
- Extent
- Orientation
- Eccentricity

Regions which have less than four shape characteristics which lie in the limits given are very unlikely to be a hand, and thus, are removed. This is not a positive test, as in, hand regions do fall in the limits given by this algorithm; however, it is not necessary that non-hand regions are recognized as false in this step. Due to the nature of shape characteristics and the lack of means to accurately define a shape – hand shape in this case – it is impossible to get exact hand regions at this stage itself. However, several implausible regions are rightly recognized by this algorithm as non-hand shapes and thusly, removed for the possible hand areas.

4.3. Hand Shape Detection

The final stage of the processing is a self-designed algorithm which determines whether a shape can be a hand depending on the row-wise distribution of skin pixels. If the shape is unlikely to be a hand, it is removed at this stage. For the objects which are very likely to be a hand, the algorithm arrives at a threshold point near the wrist of the person and only considers the region above the wrist (that is, the palm). Thus, the palm region of a hand is segmented from a cluttered background.

The algorithm works in the following way – it calculates the number of pixels which constitute the row of the bounding box of the region under consideration. For the region to be a hand, it has been observed that the number of skin pixels per row increase more or less steadily till a maximum point, and then fall and steady at the wrist. Thus, all the regions which satisfy the aforementioned conditions are determined to be a hand. The point at which the number of skin pixels begin to saturate is identified at the wrist.

5

Design and Implementation

5.1. Skin Detector

The ranges for skin detector were to be determined in order to correctly classify any given pixel as a skin or a non-skin pixel. The ranges in which skin pixels lie were found out after considering the skin color values for HSV and RGB color space over a hundred images. The initial thresholds for the algorithm were taken as those determined by [16]. However, it was found that the values did not have a very good classification accuracy. The threshold values were varied and the accuracy of the skin detector algorithm was tested on the dataset. The thresholds with the best skin detection accuracy were selected as the thresholds for the skin detector algorithm, i.e. the first stage of this project.

- Hue value for skin was determined to be between 0 to 0.24 or 0.74 to 1.
- Saturation value for skin was determined to be between 0.16 to 0.79
- Intensity value for skin should be greater than 0.8 times intensity value for face
- Red color value should be less than 1.5 times sum of Blue and Green color values
- Blue color value should be less than 1.5 times sum of Red and Green color values
- Green color value should be less than 1.5 times sum of Blue and Red color values

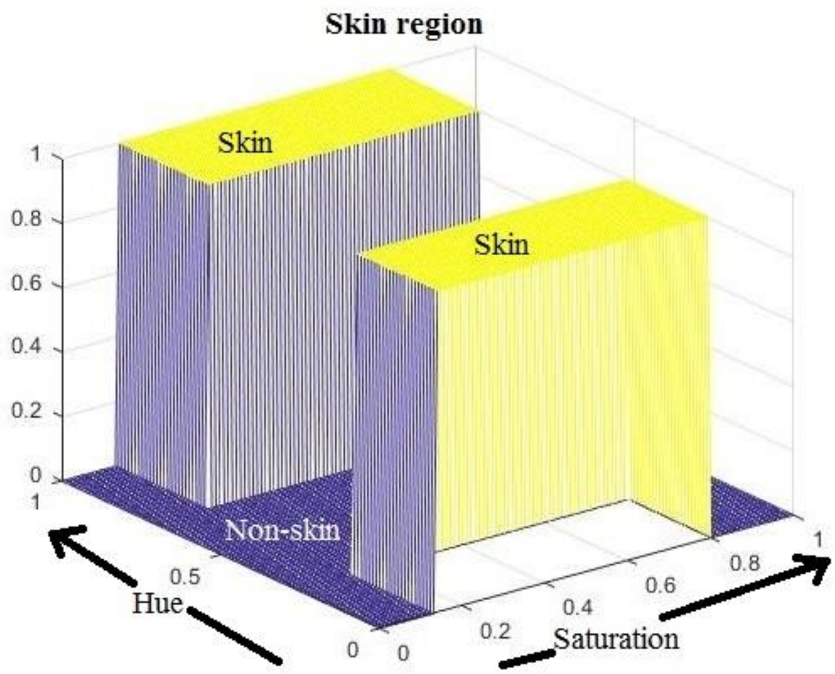

Figure 11: Skin detection algorithm

5.2. Shape Characteristics

The limits for the five shape characteristics – solidity, extent, circularity, orientation and eccentricity – were to be determined. In order to do so, we used a dataset [13] of 800 hand images taken against a plain background, segmented the hand shapes from those images, and found the shape characteristics of those 800 images by using the function *regionprops* in MATLAB. For each region, the following five characteristics are found out using the function *regionprops* in MATLAB – Solidity, Extent, Area, Perimeter, Orientation and Extent. Circularity was determined using the formula

$$Circularity = \frac{4\pi Area}{Perimeter^2}$$

The shape characteristics used and their variation for human hands has been discussed point by point according to individual characteristics.

i. Solidity

Solidity is a measure which describes a shape's resemblance to its convex area. The solidity of a convex shape is always 1, while that of concave shapes is less than one.

- Solidity lies between 0 and 1
- It can be observed that for hand, solidity generally lies between 0.6 and 0.75

$$Solidity = \frac{Area}{ConvexArea}$$

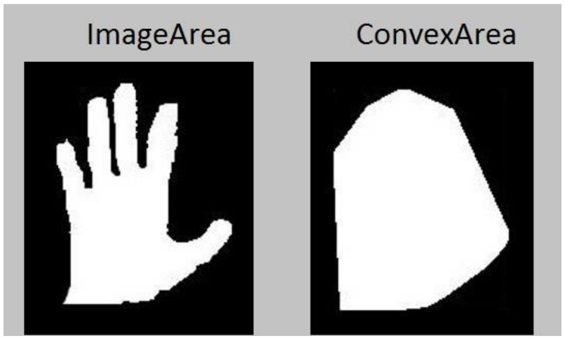

Figure 12: Solidity

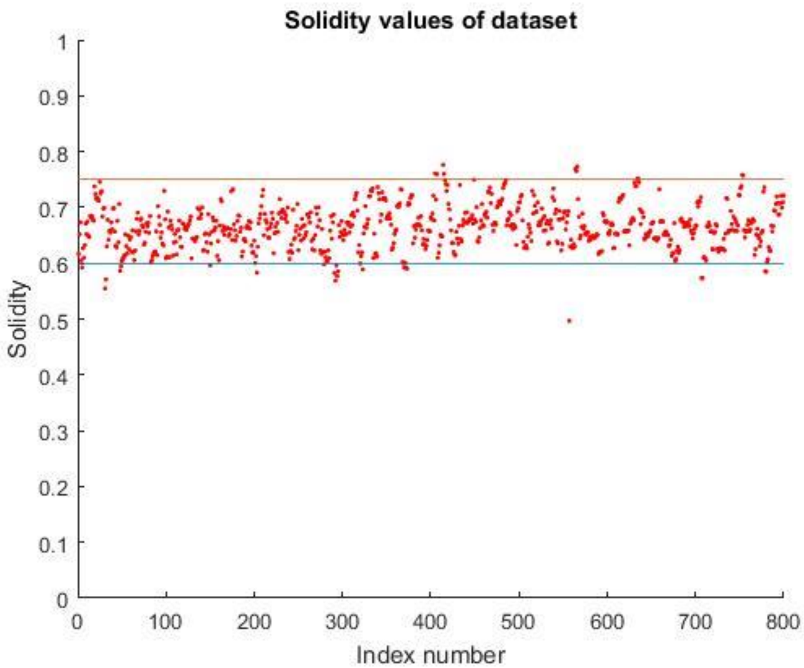

Figure 13: Solidity Graph

ii. Extent

Extent is a measure which describes a shape's resemblance to its bounding box – that is the rectangle formed with area equal to the width times the height of the shape.

- Extent lies between 0 and 1
- It can be observed that for hand, extent generally lies between 0.4 and 0.6

$$Extent = \frac{Area}{BoundingBoxArea}$$

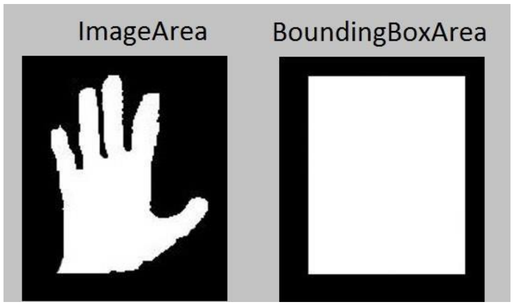

Figure 14: Extent

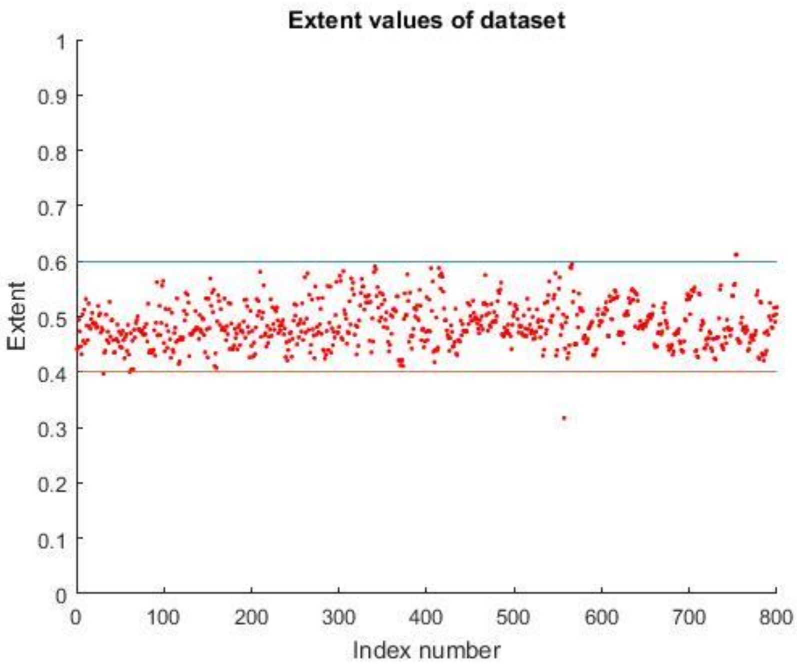

Figure 15: Extent Graph

iii. Circularity

The circularity symbol is used to describe how close an object should be to a true circle.

- Circularity lies between 0 and 1
- It can be observed that for hand, circularity generally lies between 0.1 and 0.25

$$Circularity = \frac{4\pi Area}{Perimeter^2}$$

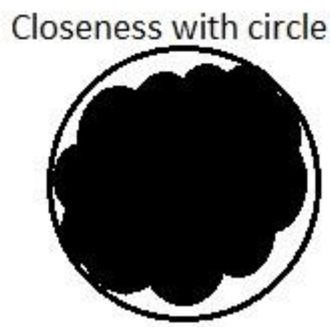

Figure 16: Circularity

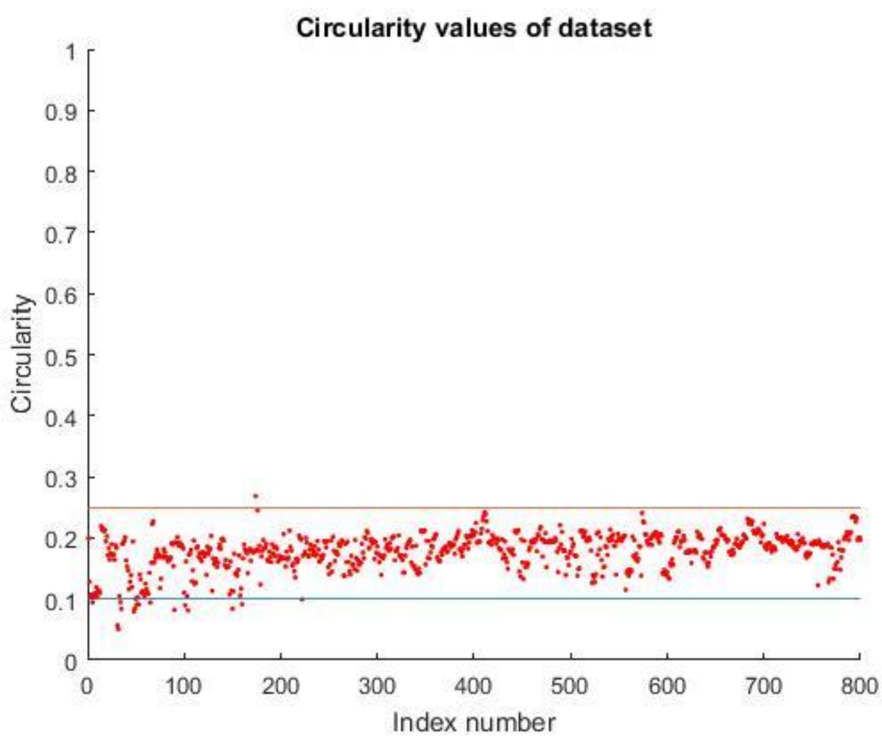

Figure 17: Circularity Graph

iv. Orientation

The closest ellipse approximation of the region is found and the angle made by the major axis of the ellipse with the x-axis is called Orientation.

- Orientation lies between -90 and +90
- It can be observed that for hand, |Orientation| generally lies between 60 and 90.

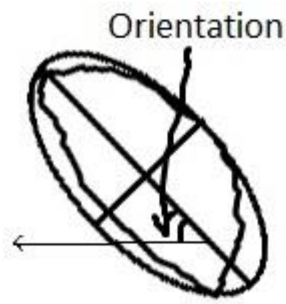

Figure 18: Orientation

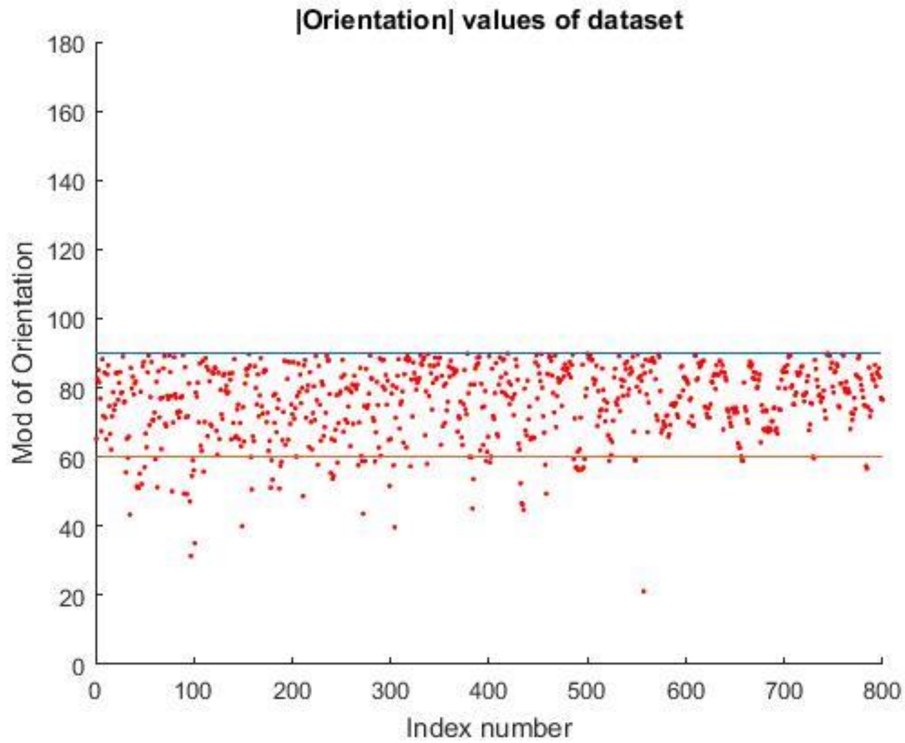

Figure 19: Orientation graph

v. Eccentricity

The closest ellipse approximation of the region is found and the ratio of the distance between the foci of the ellipse and its major axis length is called Orientation.

- Eccentricity lies between 0 and 1
- It can be observed that for hand, eccentricity generally lies between 0.5 and 0.75

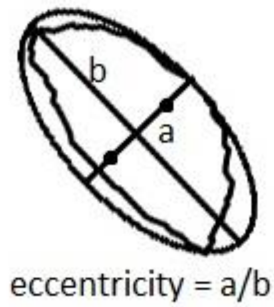

Figure 20: Eccentricity

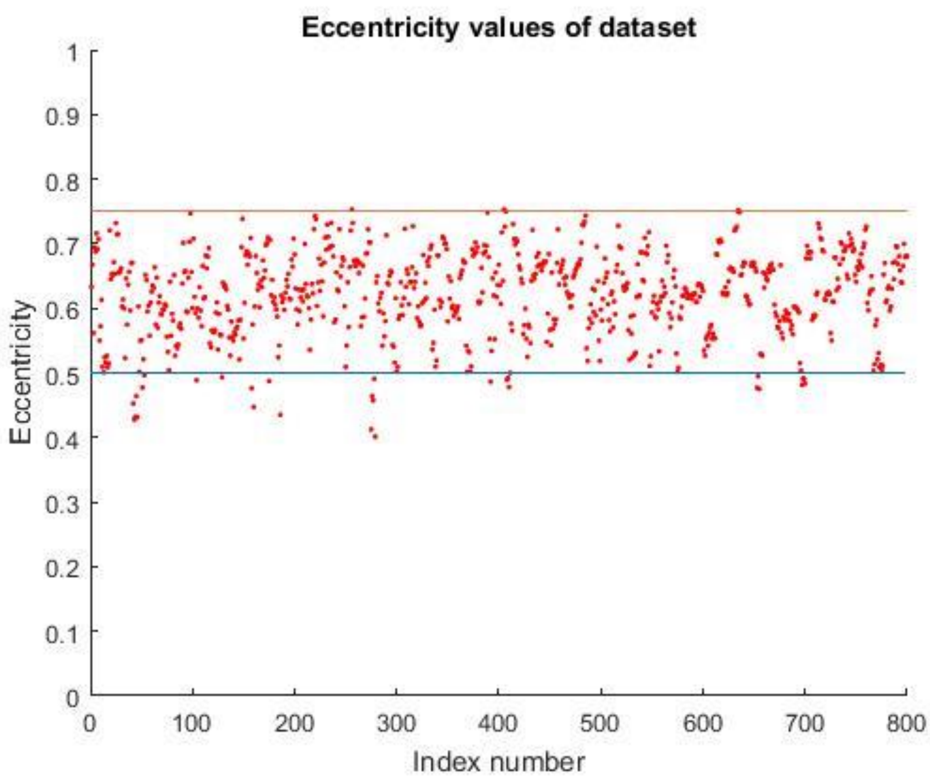

Figure 21: Eccentricity Graph

5.3. Hand Shape Detector

The final stage of the algorithm considers one region at a time and defines the bounding box for that region. The bounding box has Figure 19 can be seen as the blue rectangle. Consider the limits of the bounding box to be (y_1, y_2) and (x_1, x_2). The algorithm finds the number of skin-colored pixels in each row, starting with y1 and ending at y2.

That is, the row changes according to the rule –

$y = y1 + h$ As the value of h varies from 0 to (y2-y1)

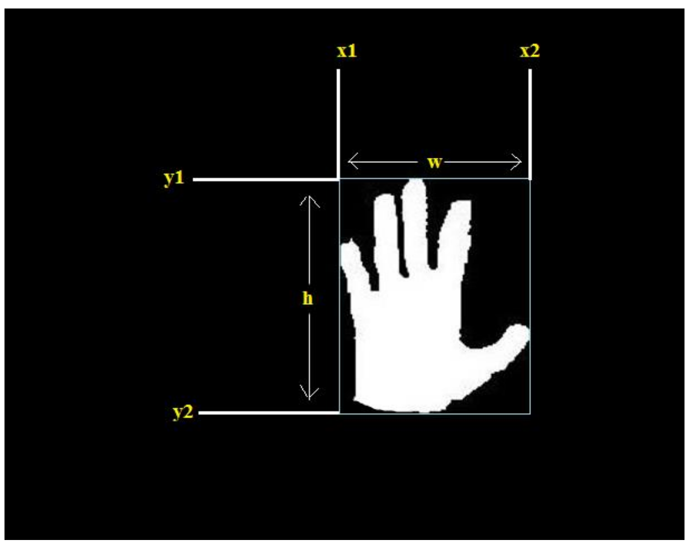

Figure 22: Dimensions of region

The algorithm works in the following way – it calculates the number of pixels which constitute the row of the bounding box of the region under consideration. For hands, it has been observed that the number of skin pixels per row increase more or less steadily till a maximum point, and then saturate for a certain region following the fall, starting at the wrist section of the hand. All the regions which satisfy the aforementioned conditions are determined to be a hand. The point at which the number of skin pixels begin to saturate is identified at the wrist.

Step 2 removes the regions which follow the aforementioned trend but are not a hand because such shapes do not have similar shape characteristics. If the number of pixels with respect to the row number do not follow the trend of initially increasing more or less uniformly, reaching a maximum point, and finally decreasing, the algorithm decides that the region under consideration is unlikely to be a hand and that region is removed at this stage. For the objects which are very likely to be a hand, the algorithm arrives at a threshold point near the wrist of the person and only considers the region above the wrist (that is, the palm). Thus, the palm region of a hand is segmented from a cluttered background.

The following illustrates how the algorithm works. Two regions are considered, one is a hand and the other is not a hand.

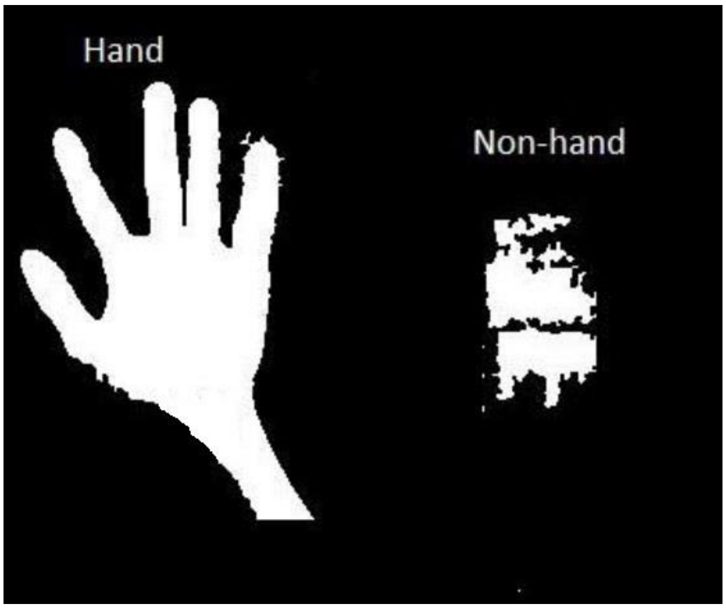

Figure 23: Working of hand shape detector

Here, it can been observed that the number of skin pixels per row increase steadily, except the small irregularity at row number 3 from the starting point. The number of skin pixels per row increase steadily till the row number 12, and then fall steadily from the region from row number 13 to row number 17. After row number 17, it can be seen that the row of skin pixels is more or less constant. Thus, this graph of this region satisfies all the necessary conditions for the region to be a hand. It can be noted that the number of skin pixels start to saturate at the wrist section of the hand. Thus, the wrist is encountered at the row number 17 and subsequently the output of this algorithm will contain only the region from row number 1 to row number 17 as the hand, and the region below the wrist (row number 18 onwards) will be discarded.

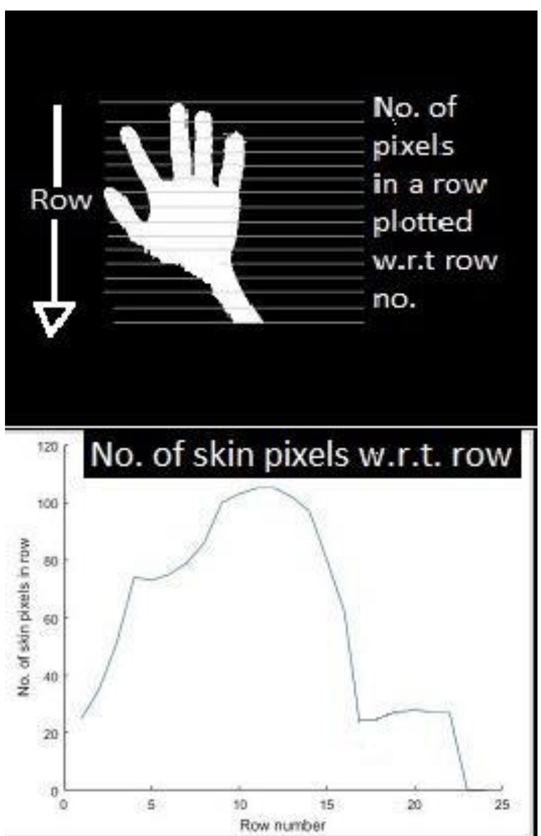

Figure 24: Plot for hand shape

Here, it can been observed that the number of skin pixels per row drop initially, then increase steadily, again decrease, again increase to the global maximum. However, for hands, the number of skin pixels per row increase more or less steadily till a maximum point, and then saturate for a certain region following the fall, starting at the wrist section of the hand. As this region does not follow the conditions for the region for being a hand, we can conclude that this region is definitely not a hand. The region will be determined to not be a hand by the algorithm, and subsequently, the output of this algorithm will not contain any part of this region.

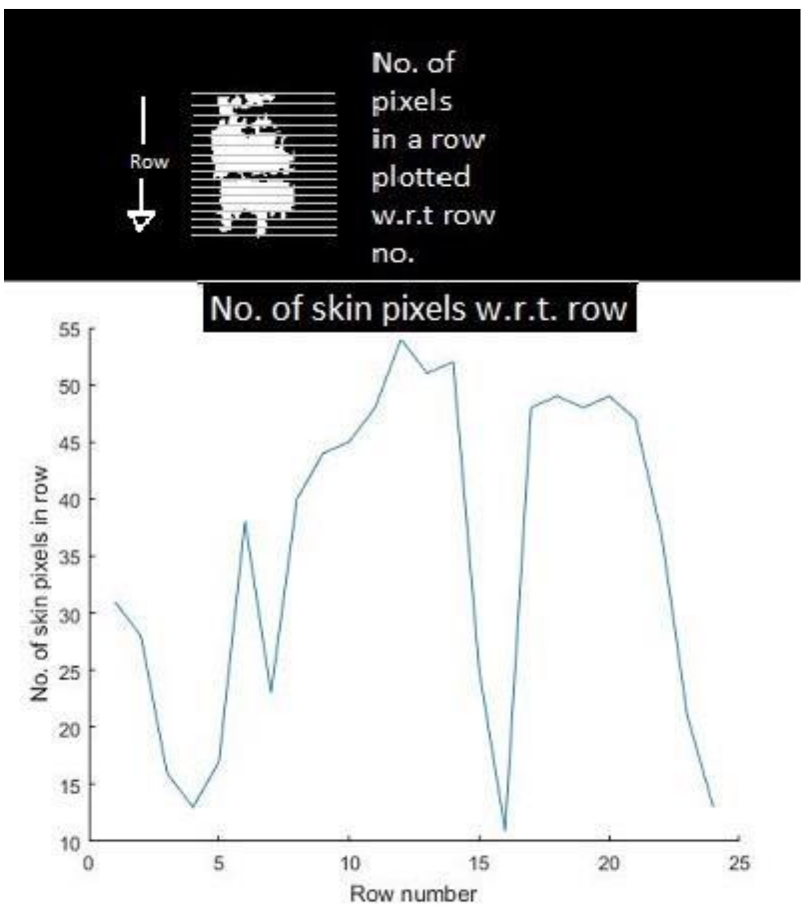

Figure 25: Plot for non-hand shape

6

Experimentation and Results

6.1. Dataset

As the amount of research done for hand-based biometrics on cluttered backgrounds is limited, there are a limited amount of images available for testing. There are no readily available datasets available for hand images on non-uniform backgrounds. Due to lack of resources available for hand images in cluttered backgrounds, we have decided to take the initiative and create a dataset on our own which will be made open to access for the public once the dataset is large enough. Currently, we have around 150 images with hand visible against a cluttered background, all in good lighting conditions. Some of these images were downloaded from the internet by manually searching for images of people waving at the camera, mostly politicians and public figures – as their images of waving at the camera can be commonly found on the internet, while other images were either taken by me or others manually with a camera.

6.2. Skin Detection

The following image is a typical example of high clutter backgrounds. There are a lot of brown colored objects in the background which might have been skin. In order to get a clean segmentation, the face detector estimates the skin color of the hand and applies appropriate thresholds for skin. Thus, in this case, only skin colored objects which are lighter than the face color are detected as skin.

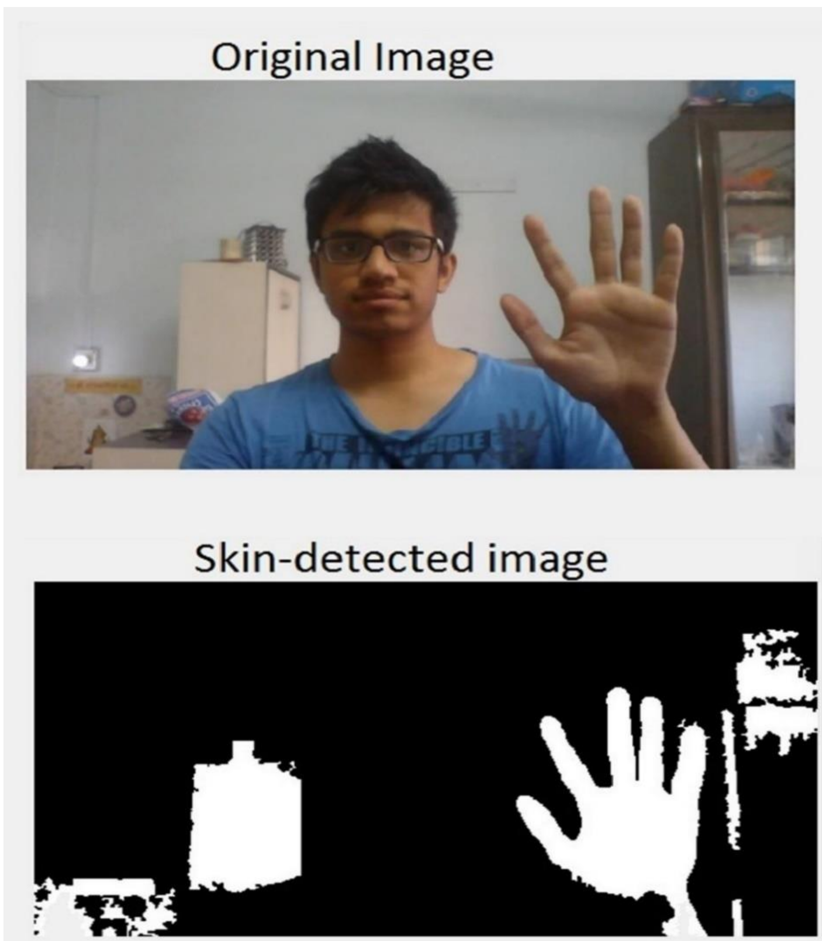

Figure 26: Skin detection stage

Only skin detection cannot be used to segment hands because skin-colored objects, for example, the light brown colored sections of the cupboards and a portion of the tiles in the above figure are also detected as skin. The face and neck regions are automatically removed at this stage itself as the location of hand and face should not be the same. The bounding box of the detected face is determined and the vertices of the rectangle are found out to be (y1, x1), (y1, x2), (y2, x1), (y2, x2). In such a case, the height (h) of the face will be (y2-y1) and the width (w) will be (x2-x1).

The bounding box in case of a face detector algorithm does not exactly match the exact location of the face. Hence, it was not possible to determine the exact face and neck region. However, we found out dimensions which worked out in most of the cases – the rectangular area between (y1-h/4, x1-w/4), (y1-h/4, x2+w/4), (y2+7h/4, x1-w/4), (y2+7h/4, x2+w/4) contained the face and neck region.

6.3. Shape Characteristics

Continuing the analysis from the image in figure 30, the regions in the image are filtered depending upon whether or not their shape characteristics fall in the allowed ranges.

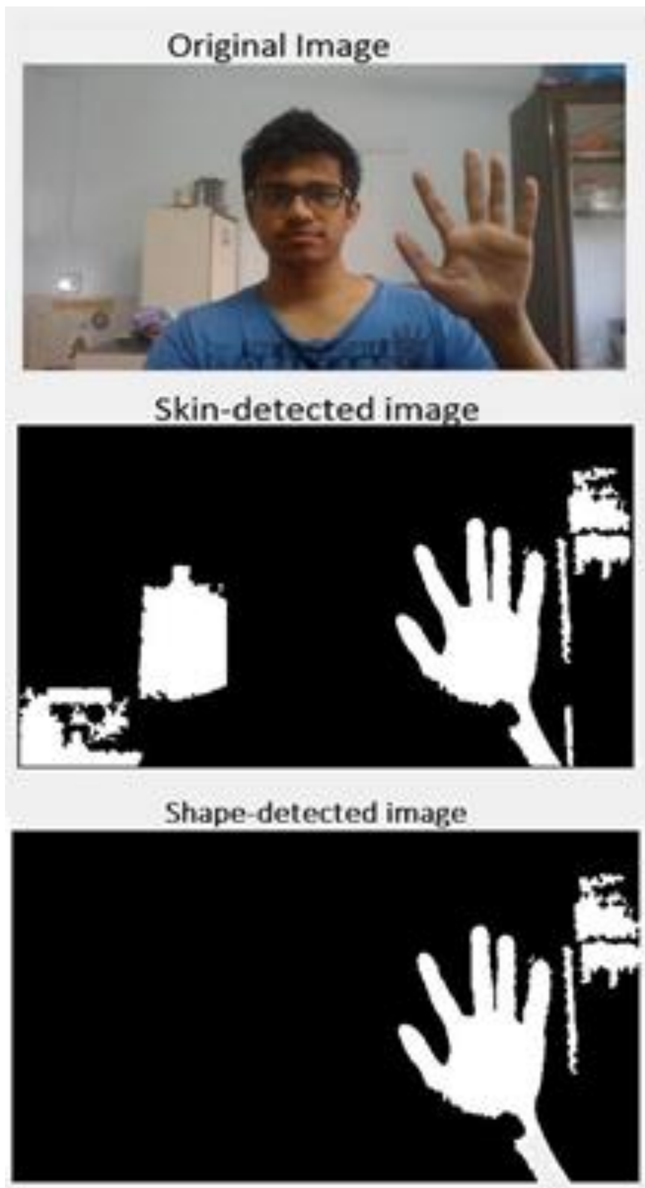

Figure 27: Shape characteristics stage

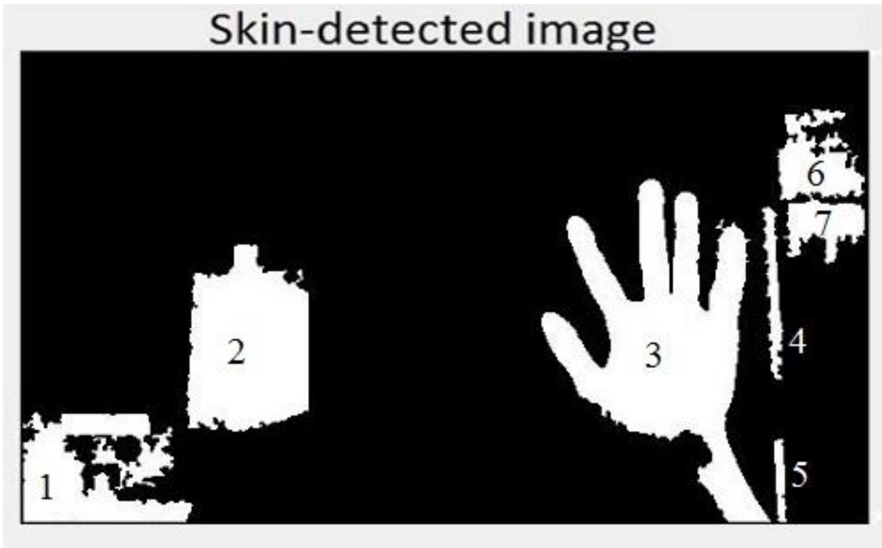

Figure 28: Region numbering for skin detection stage

For the region in the left-bottom corner, region number 1, the amount of non-skin pixels in its convex area and bounding box seem to be very less. Thus, that region must be having Solidity and Extent values which fell outside the allowable limits for hand shape and was removed by the second stage of the algorithm. The region 2 has very few non-skin pixels and its Solidity and Extent values are way higher than the allowed limits. Region 4 has uneven thickness due to which the bounding box and convex area are appropriate and the region is still not removed at this stage. However, region 5 has uniform thickness and its Solidity and Extent are very high. Region 6 and 7 have the optimum amount of non-skin pixels in their bounding boxes and convex areas; hence, they are not removed at this stage.

6.4. Hand Shape Detector

Continuing the analysis from the image in figure 31, the regions in the image are filtered depending upon whether or not they are shaped like a hand. The second stage, shape characteristics, is important because the final stage is time-consuming while the second stage is fast. Thus, in order to finish the process real-time, the most improbable objects are removed in the second stage.

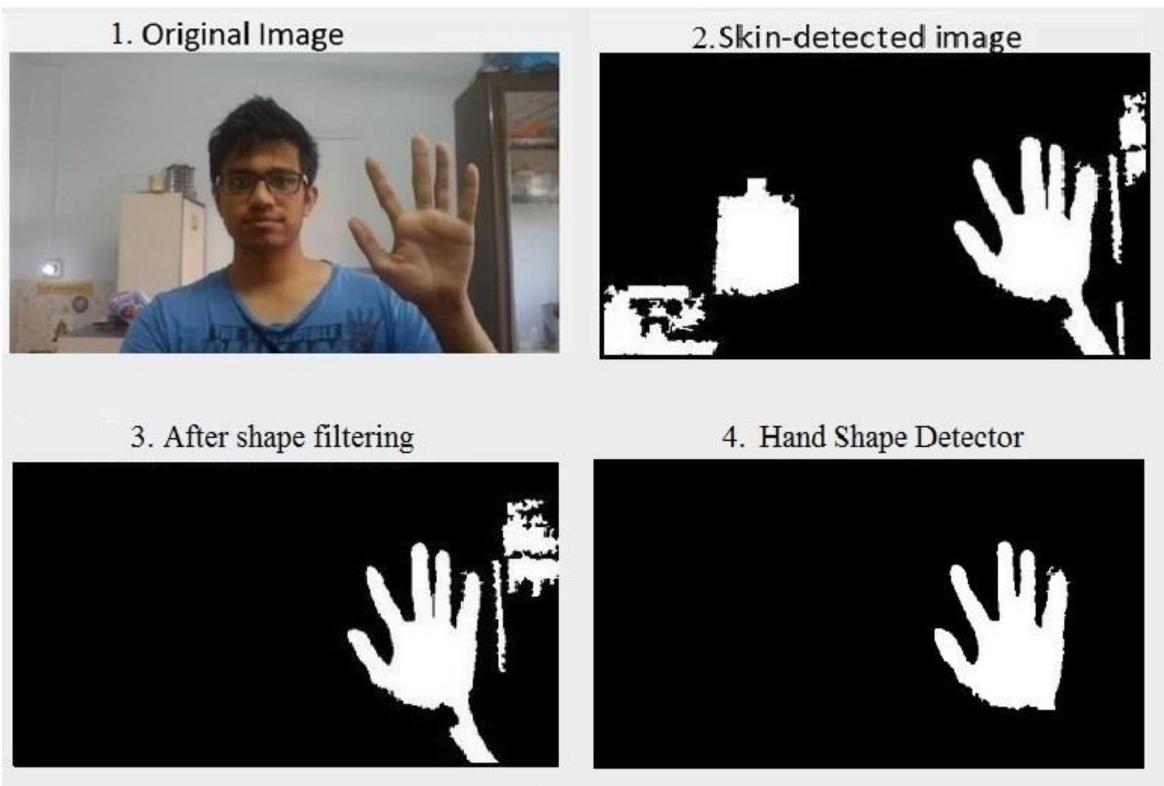

Figure 29: Hand shape stage

For the first region, the number of pixels with respect to the row number initially increase except an anomaly at row number 4, reach a maximum point, and finally decrease till row number 16. This is in accordance with the description of a hand. Hence, the algorithm decides that the region under consideration is likely to be a hand.

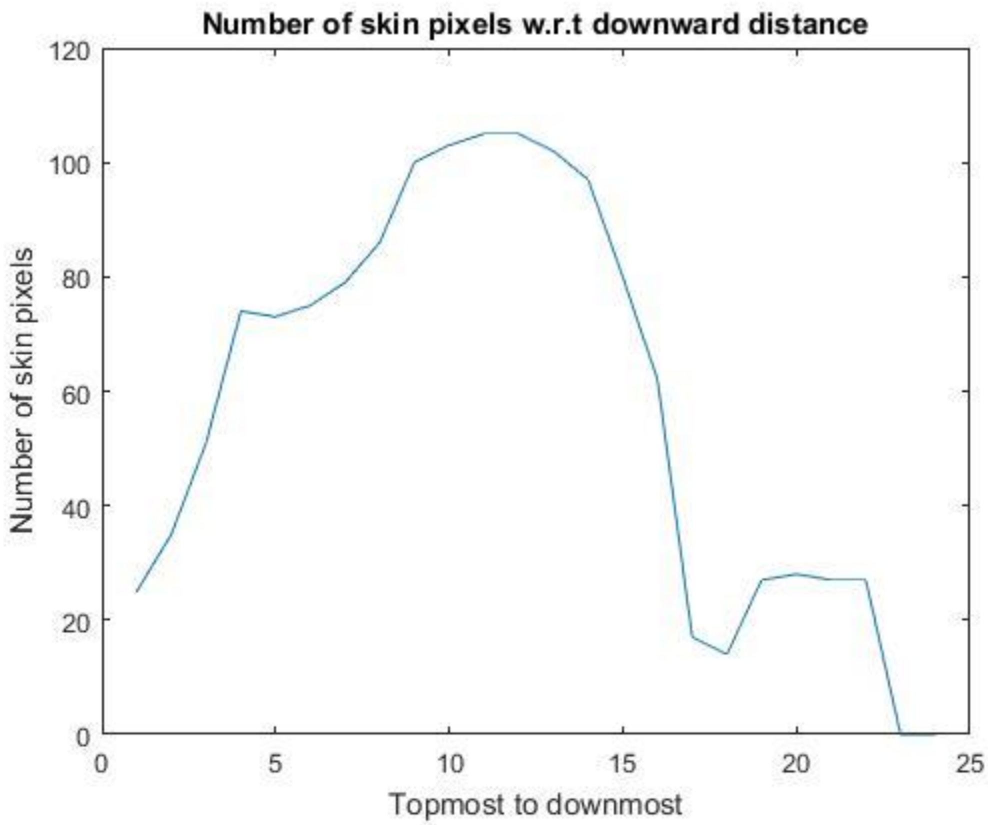

Figure 30: Hand 1 - Number of skin pixels vs row

For the second region, the number of pixels with respect to the row number initially decrease, then increase, then decrease, again increase till reach a maximum point and so on. This is not in accordance with the description of a hand. Hence, the algorithm decides that the region under consideration is unlikely to be a hand and removes it at this stage.

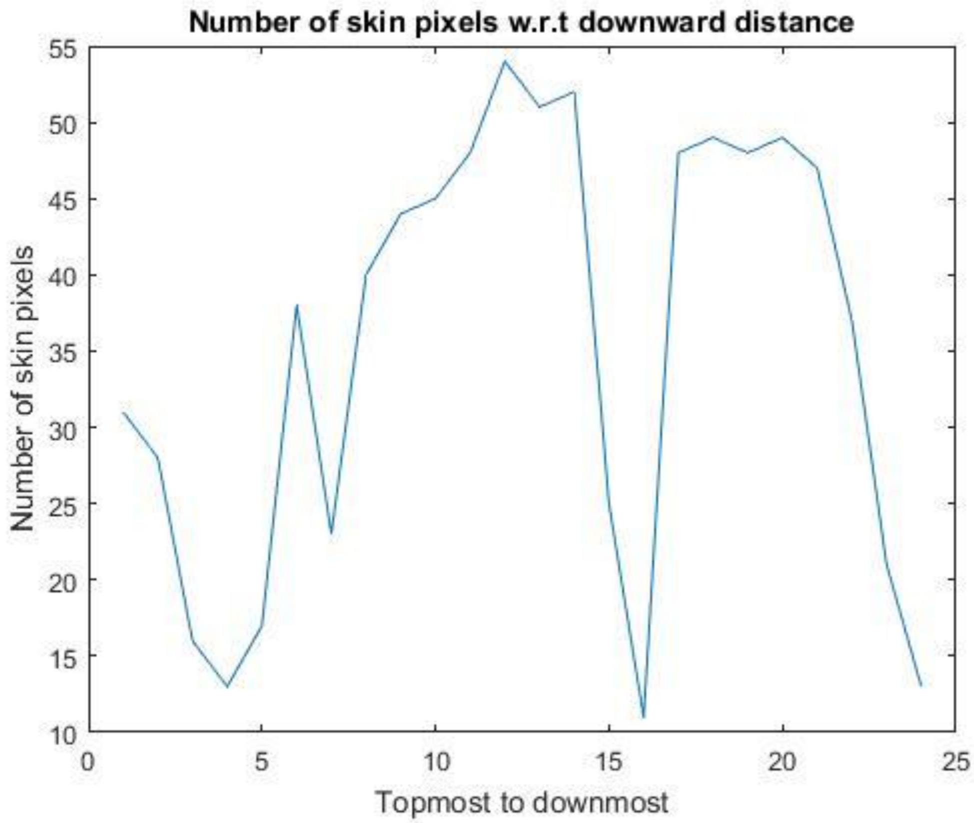

Figure 31: Object 2 - Number of skin pixels vs row

For the third region, the number of pixels with respect to the row number increase till a maximum point at the start itself, then decrease a little, again increase till reach a maximum point and so on. This is not in accordance with the description of a hand. Hence, the algorithm decides that the region under consideration is unlikely to be a hand and removes it at this stage.

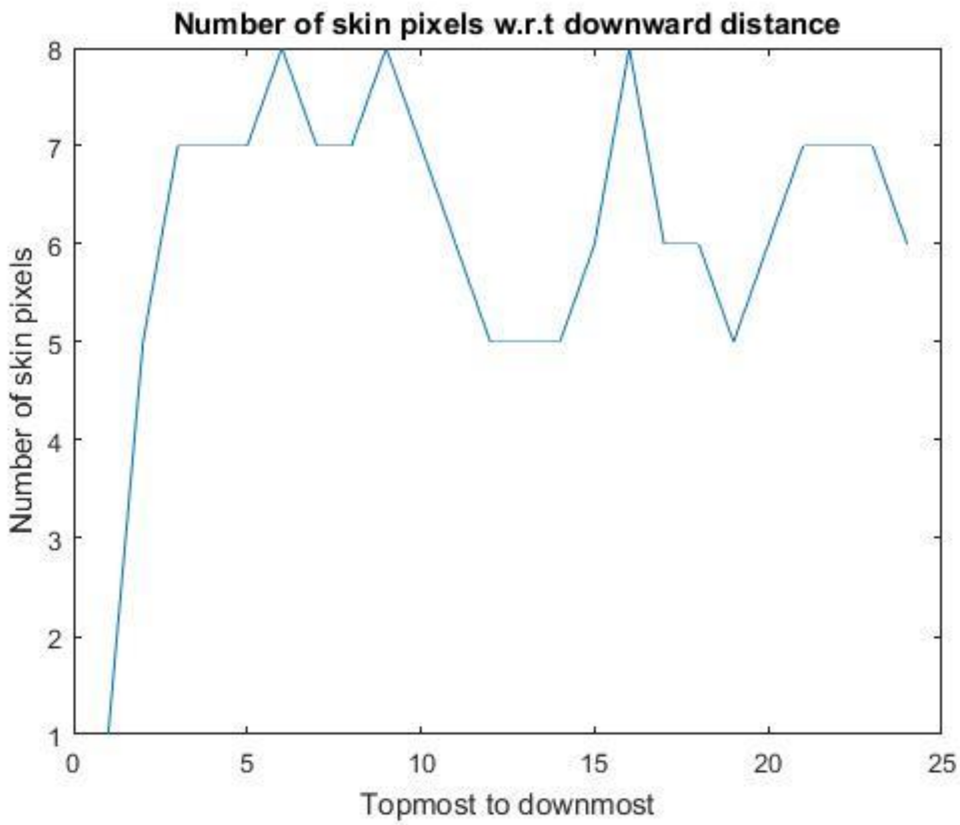

Figure 32: Object 3 - Number of skin pixels vs row

6.5. Results

Among the 150 images from the datasets, some of the results of the segmentation are shown below. The first image is the input image which is captured by the camera. Stage 1 output implies the output of the skin detector stage. Step 2 output signifies the output of the shape characteristics stage, while Step 3 output signifies the output of the hand detector algorithm. From the output of the hand detector stage, the edges of the detected hand are found. The locations of these edge pixels are stored in an array and superimposed in red color on the input image so that the boundary of the segmented hand shape can be identified.

It can be noted from the following images that the algorithm successfully segments the hand region from images with cluttered backgrounds.

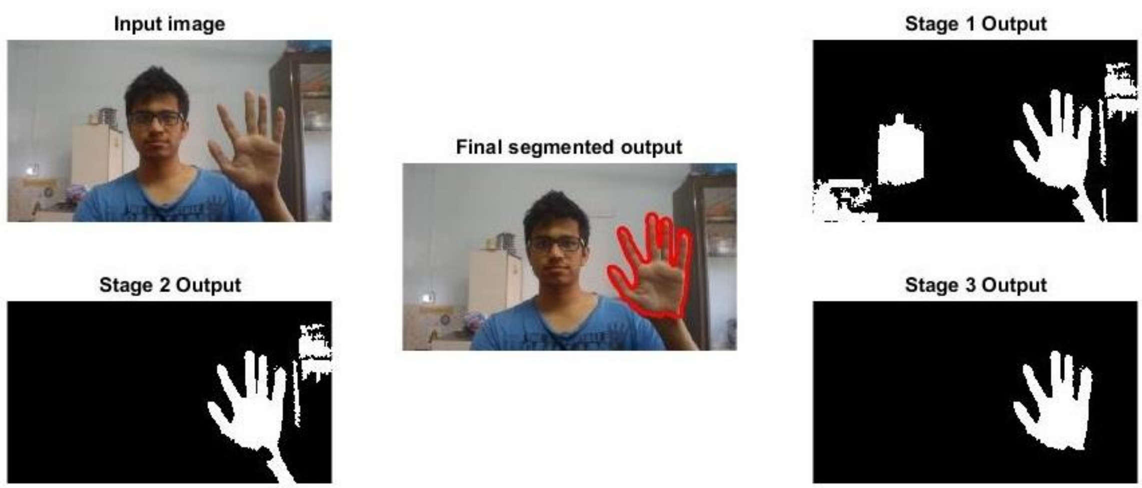

Figure 33: Result 1

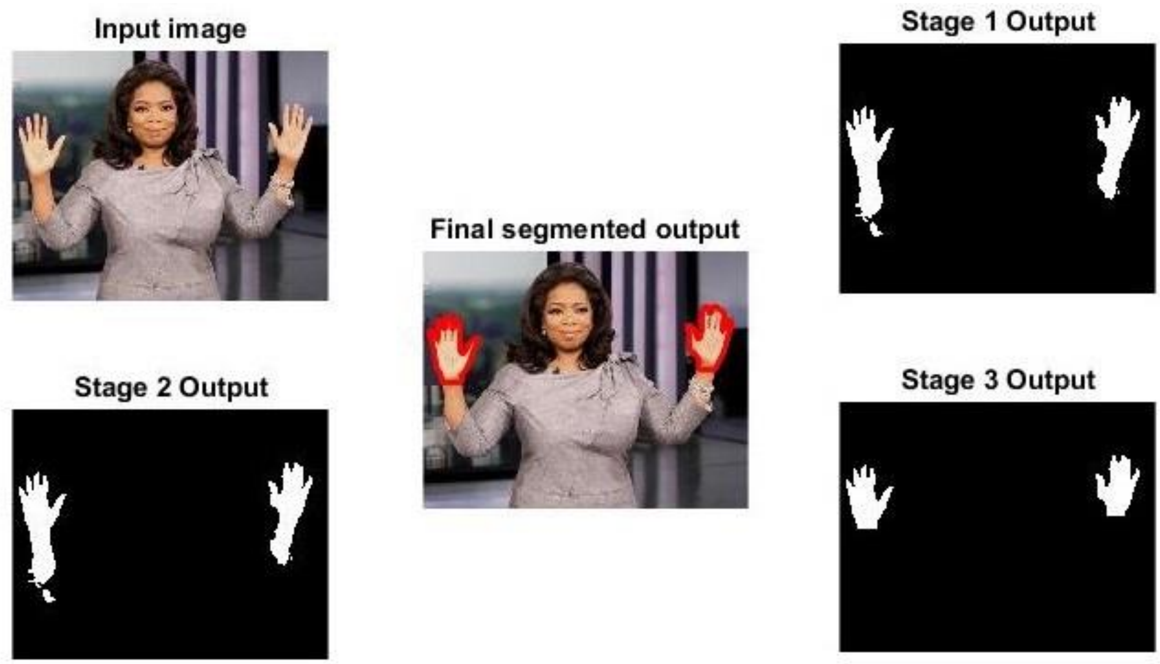

Figure 34: Result 2

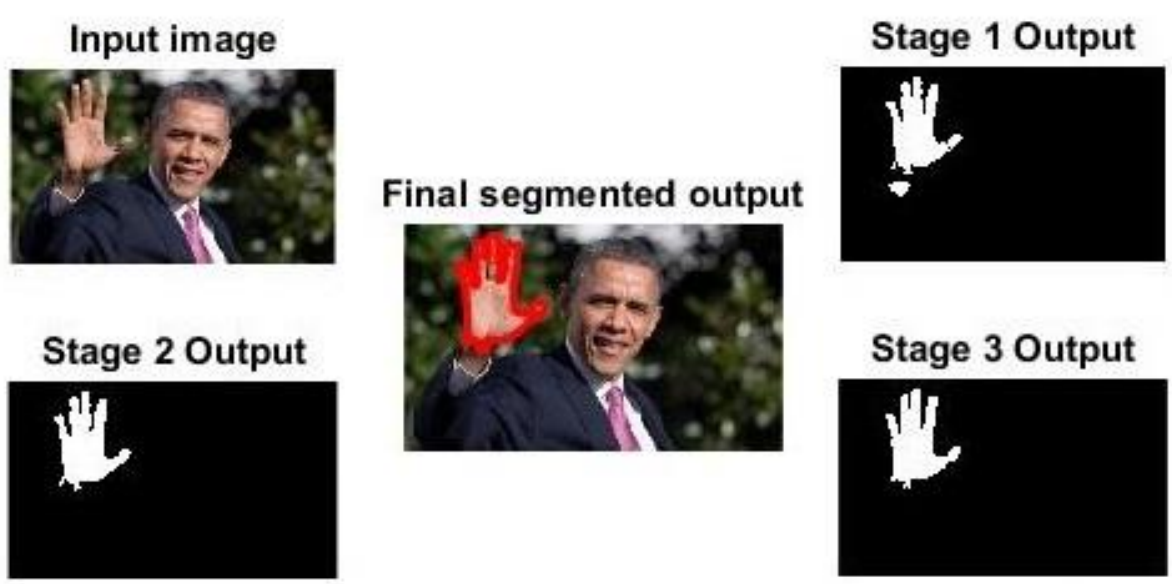

Figure 35: Result 3

Figure 36: Result 4

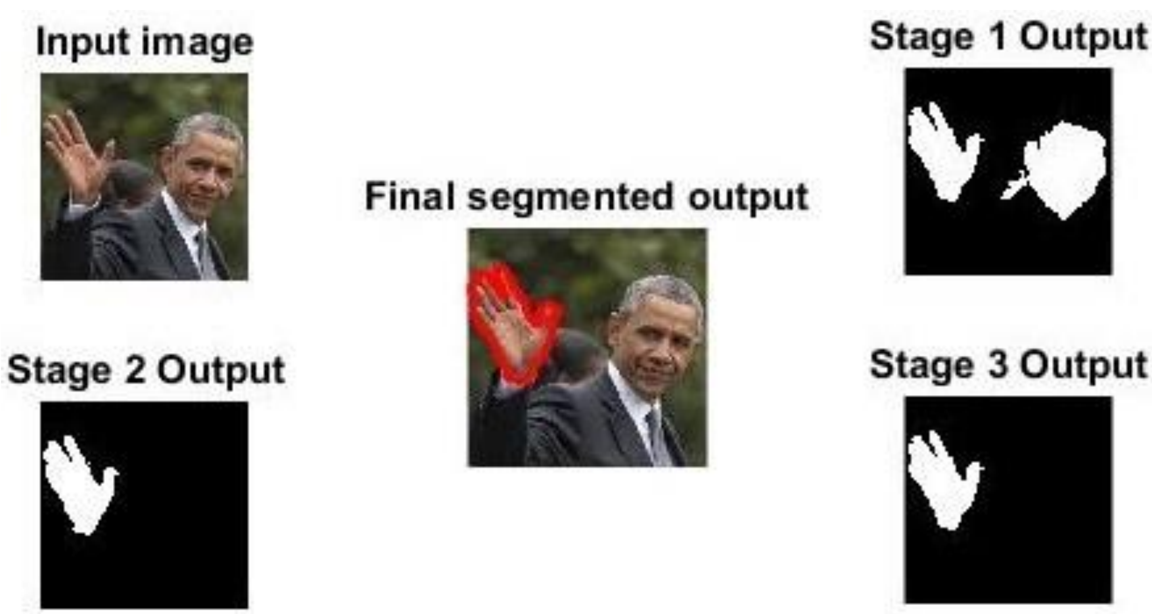

Figure 37: Result 5

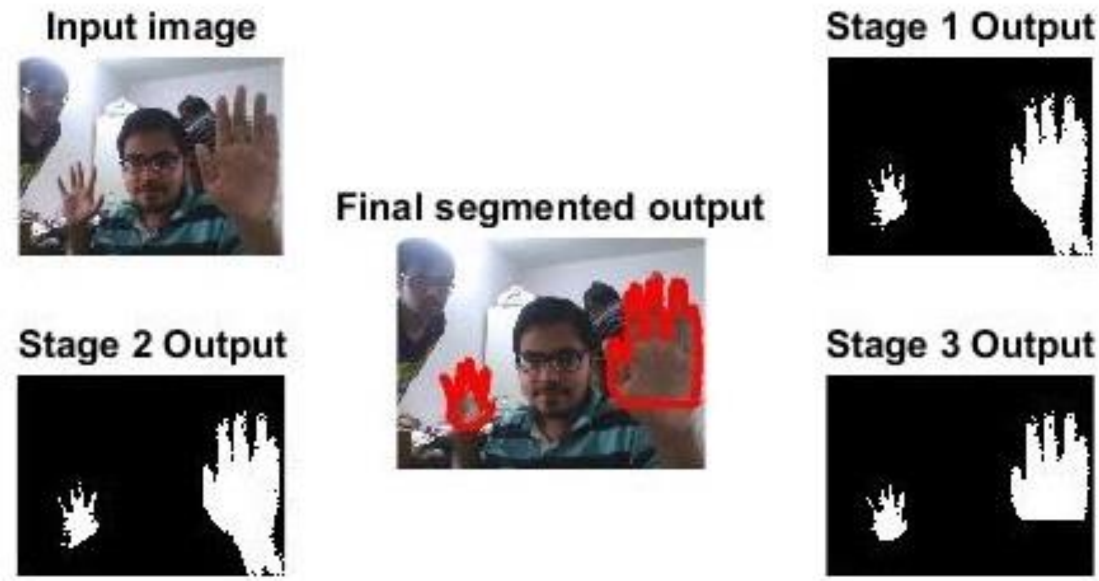

Figure 38: Result 6

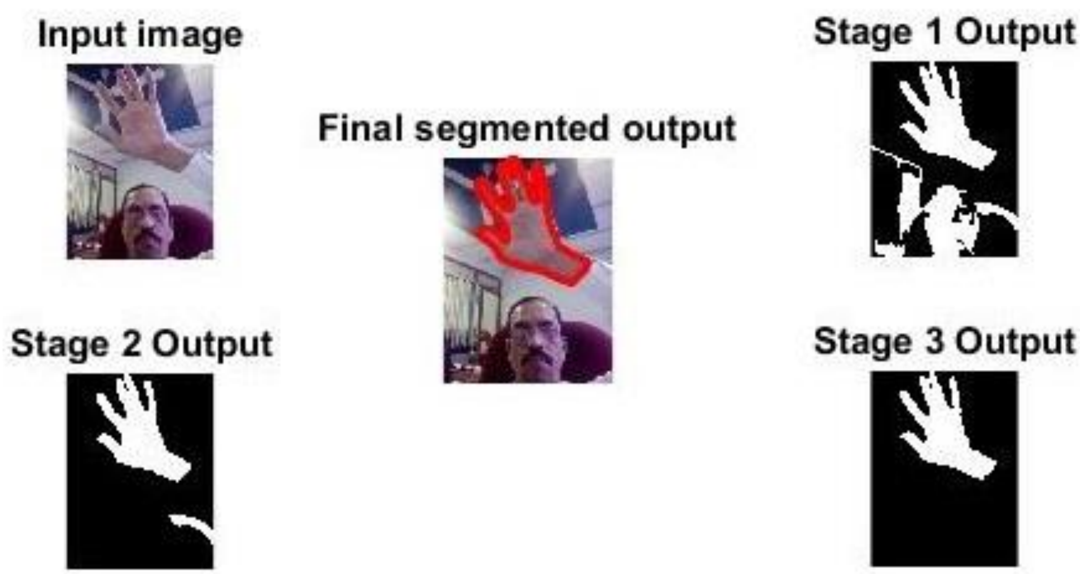

Figure 39: Result 7

7

Conclusion and Future Scope

Following are the conclusions based on the results obtained:

- Hand segmentation algorithm presented in this report segments hand with high accuracy and can be used to implement classical hand geometry based techniques on cluttered backgrounds.
- Skin detection is a useful step in order to segment hand from background clutter as it is accurate, fast and does not suffer from subject movement like methods based on background subtraction
- Determining the range in which skin color lies can be find out by finding the face and determining the color of the skin on the face. Due to the extremely accurate face detection techniques available today, this approach works very well and boosts the accuracy of the skin detector.

As future scope of the problem, we can increase the number of shape characteristics and incorporate complex shape characteristics such as compactness, curl, formfactor etc. We can also use skeletons, harmonic analysis. Zernike moments etc. Improving the limits for the skin detector can be done using machine learning concepts such as Support Vector Machines (SVMs), K-means clustering etc.

References

1. Jain, Anil K., Arun Ross, and Salil Prabhakar. "An introduction to biometric recognition." *IEEE Transactions on circuits and systems for video technology* 14.1 (2004): 4-20.
2. Yörük, Erdem, Helin Dutağaci, and Bülent Sankur. "Hand biometrics." *Image and Vision Computing* 24.5 (2006): 483-497.
3. Doublet, Julien, Olivier Lepetit, and Marinette Revenu. "Contact less hand recognition using shape and texture features." 8th International Conference on Signal Processing. Vol. 3. IEEE, 2006.
4. A. K. Jain, A. Ross, and S. Pankanti, "A prototype hand geometry based verification system," in Proc. 2nd Int. Conf. Audio- and Video-Based Biometric Person Authentication, Mar. 1999, pp. 166–171
5. R. Sanchez-Reillo, C. Sanchez-Avila, and A. Gonzalez-Marcos, "Biometric identification through hand geometry measurements," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 10, pp. 1168–1171, Oct. 2000.
6. Y. Bulatov, S. Jambawalikar, P. Kumar, and S. Sethia, "Hand recognition using geometric classifiers," presented at the ICBA Int. Conf. Bioinformatics and its Applications, FL, Dec. 16–19, 2004.
7. Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.
8. Sandeep, K., and A. N. Rajagopalan. "Human Face Detection in Cluttered Color Images Using Skin Color, Edge Information." ICVGIP. 2002.
9. Mittal, Arpit, Andrew Zisserman, and Philip HS Torr. "Hand detection using multiple proposals." BMVC. 2011.
10. Yang, Mingqiang, Kidiyo Kpalma, and Joseph Ronsin. "A survey of shape feature extraction techniques." Pattern recognition (2008): 43-90.

11. Kumar, Ajay, et al. "Personal verification using palmprint and hand geometry biometric." International Conference on Audio-and Video-Based Biometric Person Authentication. Springer Berlin Heidelberg, 2003.
12. Bača, Miroslav, Petra Grd, and Tomislav Fotak. *Basic principles and trends in hand geometry and hand shape biometrics*. INTECH Open Access Publisher, 2012.
13. Kanhangad, Vivek, Ajay Kumar, and David Zhang. "Contactless and pose invariant biometric identification using hand surface." *IEEE transactions on image processing* 20.5 (2011): 1415-1424.
14. Phung, Son Lam, Abdesselam Bouzerdoum, and Douglas Chai. "Skin segmentation using color pixel classification: analysis and comparison." IEEE transactions on pattern analysis and machine intelligence 27.1 (2005): 148-154.
15. Porikli, Fatih, and Oncel Tuzel. "Human body tracking by adaptive background models and mean-shift analysis." IEEE International Workshop on Performance Evaluation of Tracking and Surveillance. 2003.
16. Oliveira, V. A., and A. Conci. "Skin Detection using HSV color space." H. Pedrini, & J. Marques de Carvalho, Workshops of Sibgrapi. 2009.
17. Karishma, S. N., and V. Lathasree. "Fusion of Skin Color Detection and Background Subtraction for Hand Gesture Segmentation." *International Journal of Engineering Research and Technology* 3.2 (2014).
18. Vezhnevets, Vladimir, Vassili Sazonov, and Alla Andreeva. "A survey on pixel-based skin color detection techniques." Proc. Graphicon. Vol. 3. 2003.
19. Jones, Michael J., and James M. Rehg. "Statistical color models with application to skin detection." International Journal of Computer Vision 46.1 (2002): 81-96.
20. Kovac, Jure, Peter Peer, and Franc Solina. Human skin color clustering for face detection. Vol. 2. IEEE, 2003.
21. De-Santos-Sierra, Alberto, et al. "Unconstrained and contactless hand geometry biometrics." Sensors 11.11 (2011): 10143-10164.

22. Guo, Jing-Ming, et al. "Contact-free hand geometry-based identification system." Expert Systems with Applications 39.14 (2012): 11728-11736.
23. Fotak, Tomislav, Petra Koruga, and M. Baca. "Trends in hand geometry biometrics." Central European Conference on Information and Intelligent Systems. 2012.
24. Sanchez-Reillo, Raul, Carmen Sanchez-Avila, and Ana Gonzalez-Marcos. "Biometric identification through hand geometry measurements." IEEE Transactions on pattern analysis and machine intelligence 22.10 (2000): 1168-1171.
25. Guo, Jing-Ming, et al. "Contact-free hand geometry-based identification system." Expert Systems with Applications 39.14 (2012): 11728-11736.
26. Morales, Anythami, et al. "Comparing infrared and visible illumination for contactless hand based biometric scheme." 2008 42nd Annual IEEE International Carnahan Conference on Security Technology. IEEE, 2008.
27. Kumar, Ajay. "Incorporating cohort information for reliable palmprint authentication." *Computer Vision, Graphics & Image Processing, 2008. ICVGIP'08. Sixth Indian Conference on*. IEEE, 2008.
28. Russ, John C., and Roger P. Woods. "The image processing handbook." Journal of Computer Assisted Tomography 19.6 (1995): 979-981.