DIP Final Report:

Group 1

Varshanjali Sayyaparaju

Aditya Singh

Prateek Goel

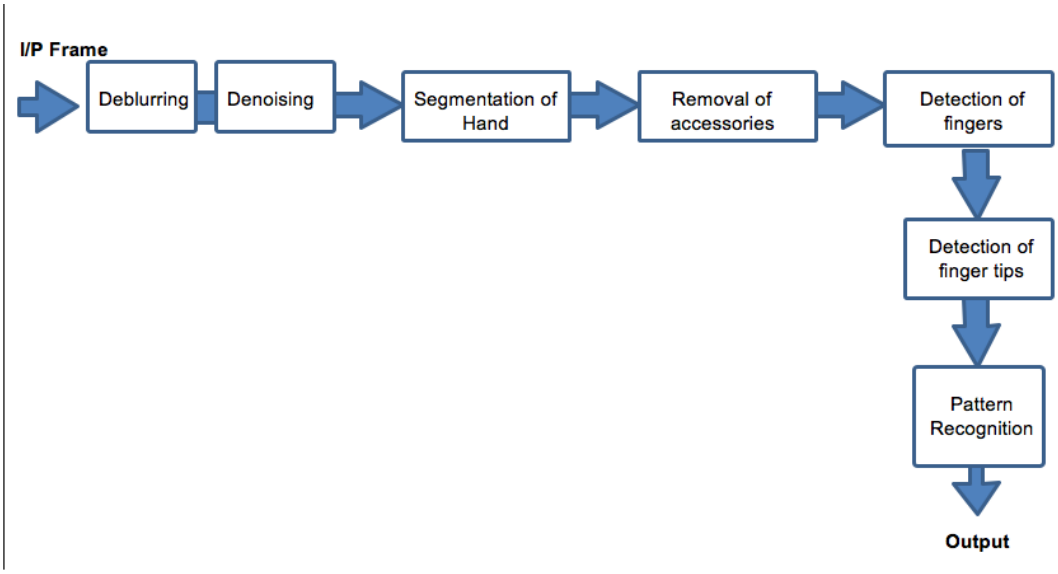
- Problem Statement:

  Goal: To identify the various hand gestures involved in Indian classical dances in a given image.

  Purpose: Indian classical dance is a cultural art which conveys meaning in every step. Many people now a days do not appreciate the story behind the dance movements. This project is an attempt to identify the hand gestures of Indian classical dances so as to try and give meaning to each gesture, and understand the story behind each dance.

  Nature of Work: Own idea implemented with prior knowledge and references from a couple of research papers.

  Block Diagram:



  Deliverables:

  - *Minimum:* Recognition of hand gestures without accessories

- *Maximum:* Recognition of hand gestures with accessories, i.e. mehendi, bangles, rings...etc.

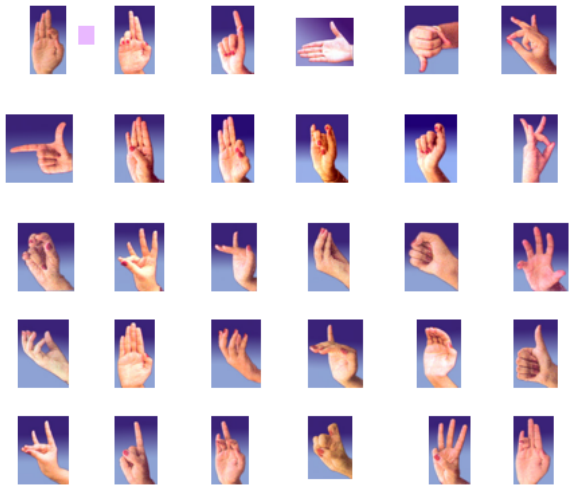
Workload Distribution:

- *Interim 1:*
  - Name1:
    - Dataset collection
    - Skin color detection
  - Name2:
    - Basic Thresholding
    - Boundary Extraction
  - Name3:
    - Finger-tip Detection
- *Interim 2:*
  - Name1:
    - Dataset collection
    - Improvements:
      - Skin-detection
      - Boundary Extraction
      - Finger-tip Detection
  - Name2:
    - Segmentation of hand
      - Skin detection
      - Color based k-means clustering
      - Edge detection
        - Sobel

        - Prewitt
        - Gradient
    - Name3:
        - Distingushing between hands
            - Boundary Extraction
            - Watershed Transform
- *Final Presentation:*
    - Name1:
        - Segmentation of Hand
        - Mehendi Removal
    - Name2:
        - Boundary Extraction
    - Name3:
        - Finger-tip detection

➢ Obstacles Faced:

- **Interim 1:** Single Hand Mudras without accessories in plain and complex background
    - *Dataset:*



Aralam (Crooked-Hand)	Ardha pathaka (Half-flag)	Ardhasoochika (Half needle)	Ardhachandra (Half-moon)	Banam (Arrow)	Bramaram (Bee)
Chandrakkala (Digit of the moon)	Chathura (Four Fingered)	Hamsapaksham (Swan's bill)	Kadak Mukham (Crab-face)	Kapidham (wood-apple)	Karthari Mukham (Arrow-Shaft)
Langulam (Tail)	Mayooram (Peacock)	Mriga Seersham (Deer)	Mukulam (Flower bud)	Mushti (Fist)	Oornanabham (Spider)
Padmakosham (Lotus Bud)	Pathaka (Flag)	Samdamsam (Tongs)	Samdasam (Parrot's Beak)	Sarppa Seersham (Snake's Hood)	Sikharam (Spire)
Simhamukham (Lion's Face)	Soochi (Needle)	Sukathundam (Parrot's Beak)	Thamrachoodda (Cock's Comb)	Thrisoolam (Trident)	Tripathaka (Flag in three parts)

- *Methods used:*
  - Thresholding
  - Skin-color detection
  - Boundary extraction
  - FInger-tip detection
- *Knowledge Requried:*
  - Basic thresholding
  - Color-space models:
  - HSV
  - YCbCr
  - Morphological Operations:
    - Binarization
    - Erosion
    - Dilation
    - Region Filling
  - Distance metrics: Euclidean distance
- *Process Overview:*

Original -> Hand Extracted -> Binarized -> Dilated Hand -> Region Filled -> Eroded Hand -> Boundary -> Finger-tips

- *Side-Achivements:*

Image 1

Image 2

Hands in complex background are segmented using skin-detection

- *Challenges:*
  - Improper hand extraction because of skin-detection

Image 3

Image 4

Image 5

Image 6

Observations:

- It can be noticed that when the amount of light on the hand is alot, skin-detection fails , as can be seen in image 3 and 4
- When light is less in a region, then also the detection process fails, as can be seen in image 5 and 6

  - Distortion of hand due to morphological operations

Image 7

Image 8

Image 9

Image 10

Image 11

Image 12

Observations:

- It can be seen how the through of the hand have increased in height
- The fingers of have increased in width

- In Image 12, the hand shape is completely lost

- Improper boundary extraction, i.e. not fully connected contour

Image 13

Image 14

Observations:

- It can be seen that the index finger boundary is not completely connected
- As can be seen that the bottom of the hand is also not fully connected

- Improper finger-tip detection

Image 15

Image 16

Image 17

Image 18

Observations:

- In image 18 thereis a finger-tips detectedat the left bottom, where no finger exists
- In image 16, there are finger-tips detected where no fingers are
- A plus point to be noted, is that where finger-tips were present, there was detection there.

- Is finger-tip detection enough for classification?

Image 19

Image 20

Observations:

- If finger-tip detection was done for the above two images, the results would be same, because finger-tips are in the same place. There is a difference in the folded fingers.

- *References:*

- Rumyantsev, Oleg. "Recognition through Palm Gestures and Movements." Thesis. Stanford, 2012. Web. <http://www.stanford.edu/class/ee368/Project_12/Reports/Merati_Rumyantsev_Rachamandan_Hand_Sign_Recognition_Based_on_Palm_Gesture_and_Movement.pdf>.
- Singh, Devendra. "Human Face Detection by Using Skin Color Segmentation, Face Features and Regions Properties." *Http://research.ijcaonline.org/volume38/number9/pxc3876881.pdf*. N.p., n.d. Web.
- Gupta, Ankit. "Finger Tips Detection and Gesture Recognition." N.p., n.d. Web. <http://home.iitk.ac.in/~ashis/CS676Report.pdf>.
  - Website
- "Asamyukta - Single Hand Gestures- Dances of India-Angika Abhinaya." *Asamyukta - Single Hand Gestures- Dances of India-Angika Abhinaya*. N.p., n.d. Web. 31 Oct. 2012. <http://www.webindia123.com/dances/abhinaya/angika_bhinaya/asamyukta.htm>

- **Interim 2:** Improvements and Double Hand Mudras without accessories in complex background

  - Improvements:
    - Skin-detection
    - Boundary Extraction
    - Finger-tip detection
  - Knowledge Required
    - Color Models: RGB, HSV and YCbCr
    - Neighborhood processing
    - K-curvature

- Changes implemented:
    - Skin-detection: Previously, we threshold only in the H and Cb, Cr planes. Then we changed to thresholding in the Cr, Cb, H, S,V, R,G, and B planes. Which gave better results:

Image 21

Image 22

Image 23

Observations:

- The whole in the middle of the hand is filled

- Boundary Extraction: Morphological operations were causing distortion in the hand and causing breaks in the boundary. Therefore, neighborhood processing was used to give better results:

Image 24

Image 25

Image 26

Image 27

Image 28

Image 29

Observations:

- The hand is in the exact shape as in original image
- There are no breaks in the boundary

- Finger-tip detection: Instead of using Euclidean distance and looking for points above a threshold, using the k-curvature method gave better results. The process involves looking for concaves and convexes in the image:

Image 30

Image 31

Image 32

Image 33

Observations:

- Finger is not detected, but boundaries of fingers are detected. Which would be usefull in the case mentioned in challenges in Interim 1. As can be seen below:

Image 34

Image 35

Image 36

Image 37

Double Hand Mudras:

Dataset:

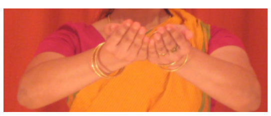






















- Methods used:
    - Segmentation of hands:
        - Skin-detection
        - Color based k-means clustering
        - Edge-detection:
            - Sobel
            - Prewitt
            - Gradient
    - Distinguishing between hands
        - Boundary Extraction
        - Watershed Transform
- Knowledge Requried:
    - Color Space Models: HSV, RGB, YCbCr, LAB
    - Sobel, Prewitt, and Gradient edge detection methods
    - k-means clustering algorithm
    - Neighorhood processing
    - Watershed Transform
- Process overview:

Get Hand from Complex background -> Distinguish between hands -> classify

- Challenges:
    - For cases, in which the background had a color close to skin color, segmentation was hard using skin-detection because the background was detected too:

Image 38

Image 39

- Since we couldn't get only the hands for such cases, we thought we would try edge segementation, and try getting only the hands out. We thought since hands are center of attraction, maybe ligthing conditions are in such a way that the hands would stand out more. Therefore, we figured the edges of the hand would be brightest compared to that of background edges. Not much success achieved in this process either, because as we predicted hands edges were the brightest in the image, but the threshold changed for each image:

Threshold: 0.09

Image 40

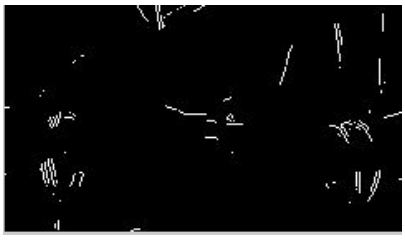


Image 41

Observations:

- For the same threshold, the hand can not be segmented in different images.

- Boundary Extraction was a problem, because since the hands are overlapping in some cases, it caused problem in feature extraction:

Image 42

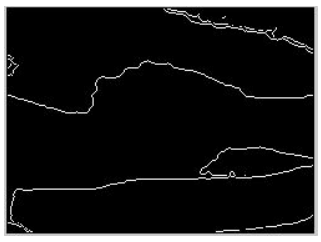


Image 43

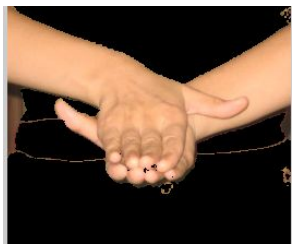


Image 44

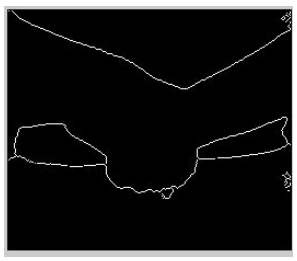


Image 45

- Watershed Transform: This is used to differentiate between hands. But it did not give the results we wished:

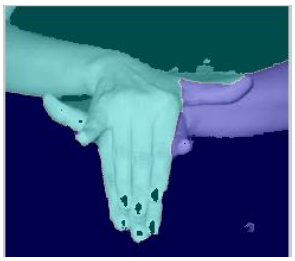
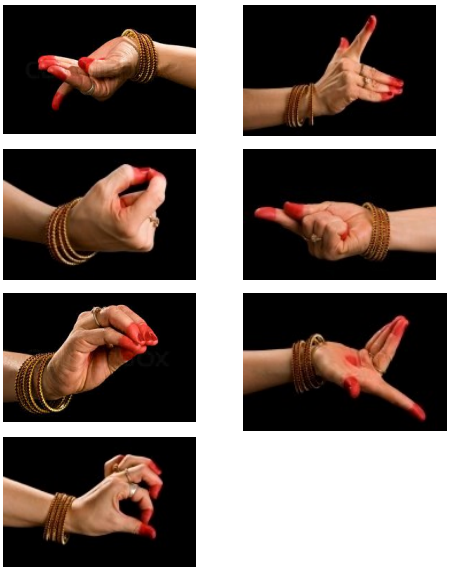



Image 46 Image 47 Image 48 Image 49

Observations:

- In image 47 it can be seen that the finger on the left is part of the right hand, but is still detected as a finger of the left hand which is not the required results
- The overlapping hands are not recognized. The fingers in the bottom are still considered as part of the hand above.
- In Image 49, both the hands are considered as being in one cluster, which is also not the results we want.

- **Final Presentation:** Single hands with accessories in addition to previous work
  - **In plain background**
    - Dataset:




- Overall Process:
    - Mehendi Removal
    - Binarization of hand
    - Boundary Extraction
    - Finger_tip detection
- Knowledge Requried:
    - Understanding the RGB color model
    - Thresholding
    - Neigbhorhood processing
    - k-curvature
- **Subprocess I:** Mehendi Removal
    - *Method:*
        - Take a RGB image
        - Consider only the red plane
    - *Results and Observations:*

Image 50
Image 51

Image 52 Image 53

Image 54

Image 55

Observations:

- Since the mehendi is red, in the Red plane, the pixel values will be near 255. And we realized that the hand pixels are approximately close to 255 too. The mehendi seems to be removed because in the red plane, the pixel values are all around 255.
- In comparison to the method of detecting the red areas in the image, this is better, because we don't have to worry about segmenting out other parts of the image that contain red in them, such as red bangles:

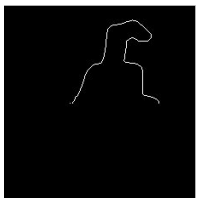
Image 56

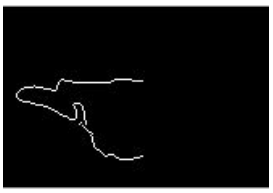
Image 57

- **Conclusion:**
  - The red parts of an image will blend into background of near 255 pixels if considered only in the R plane.
  - This method simplifies the rest of the process, because special methods are not needed to do finger-tip detection.

- **Subprocess II: Boundary Extraction:**
  - *Method:*
    - Simple Boundary Extraction as used before
    - Find out if the hand is vertical or horizontal using distance metrics
    - Segement out the upper part of the hand, i.e. only palm and fingers excluding the wrist

- *Results and Observations:*



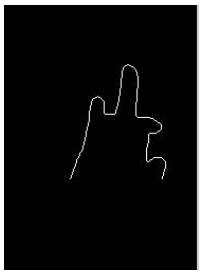


Image 56 Image 57 Image 58 Image 69 Image 60 Image 61

Observations:

- The bangles are removed from the image if only the upper part of the hand are considered
- Because of mehendi removal, the boundary extraction has become simple
- If skin-detection methods were used, then the mehendi region would not be segmented out, and would cause a problem in boundary extraction.

- **Conclusion:**
  - Skin detection would not have given such a proper boundary
  - Removal of bangles, i.e. the wrist in the image was important so as not to interfere with the finger-tip detection process

- **Subprocess III: Finger - tip detection:**
  - *Method:*
    - Same as done for previous Interims
  - *Results and Observations:*

Image 62

Image 65 Image 66

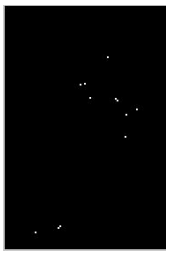



Image 63 Image 64

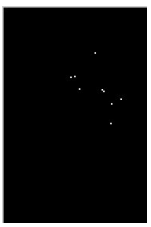





Image 67

Image 70 Image 71

Image 68 Image 69

Observations:

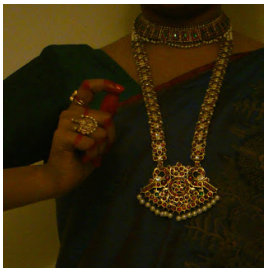
- The bangles have caused error in finger detection, because the bumps on the hand caused by the bangles are detected as curvatures
- The effect of the bangles is very prominent in Image 71 and 69.

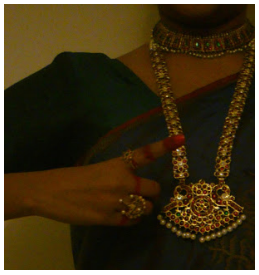
- **Conclusions:**
  - Finger detection is better when the bangles are removed from the picture
- *Overall Process:*

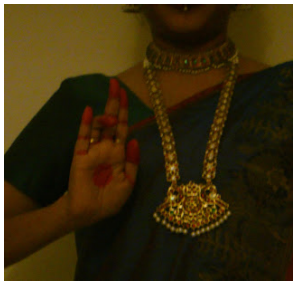
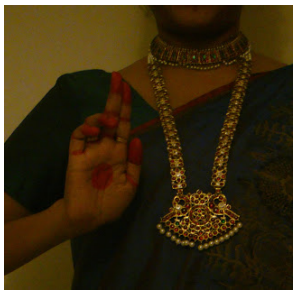





- **In complex background:**
    - *Dataset:*




    - *Method:*
        - Segmentation of hand from background
        - Mehendi removal
        - Binarization of hand
        - Removal of unwanted objects
        - Boundary Extraction and Finger-tip detection

Subprocess I: Segementation of hand from background

Method:

Use color based k-means clustering to segement out areas

Use skin detection to find out which cluster contains the hand

Label that cluster as important and discard other clusters

Results and Observations

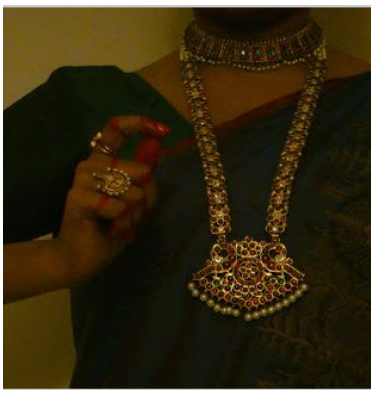


Image 72

Image 73

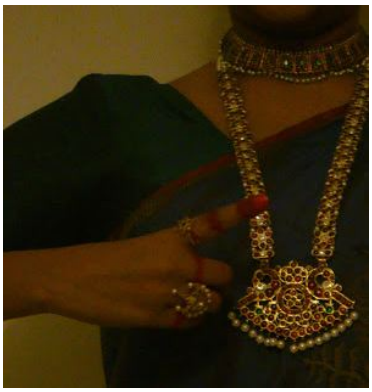


Image 74

Image 75

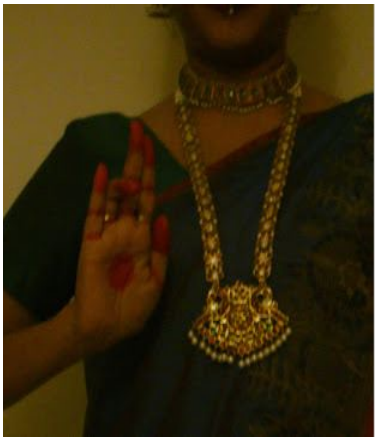


Image 76

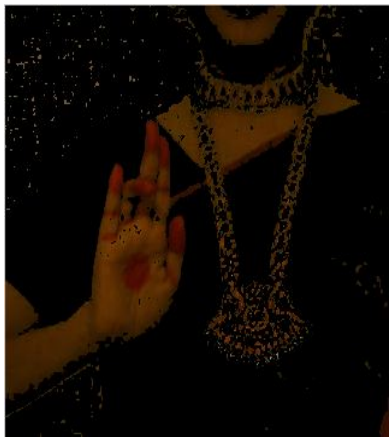


Image 77

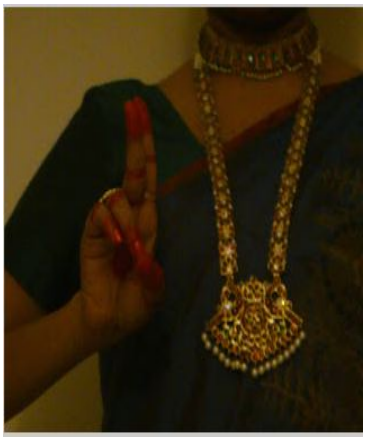


Image 78

Image 79

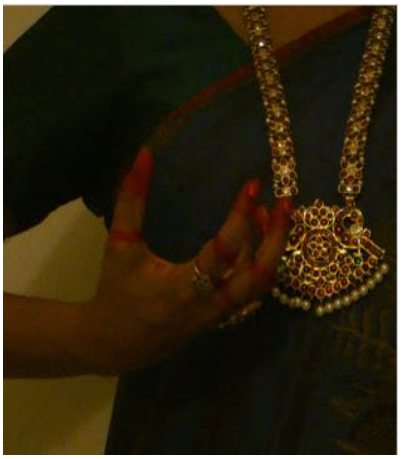


Image 80

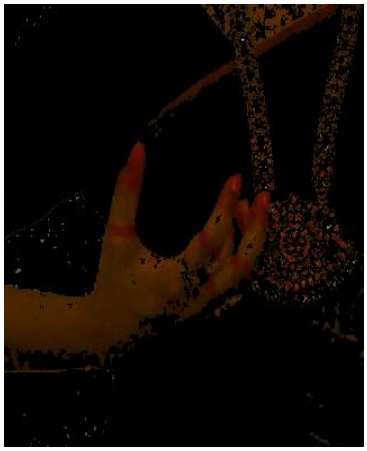


Image 81

Observations:

- K-means clustering is clubbing skin color and red together in one cluster, while the golden color and background are clubbed in one, and the green sari is clubbed in another cluster
- In image 73 and 75, it can be seen that the ring is fully gone or only partially visible, which means that k-means clustering can be treated as a method to remove ornaments of different color as well.
- The red jewels on the necklace are also clustered in, but the gold part of the necklace is gone. This is prominently visible in image 77

- **Conclusions:**
  - K-means clustering works well for segmentation of hands in complex background

- **Subprocess II: Mehendi Removal**
  - *Method:*
    - Make image grayscale by considering only the Red plane
  - *Results and Observations:*

Image 89

Image 90

Image92

Image 93

Image 94

Image 95

Observations:

- The red jewels are also removed along with the mehendi
- The color of the background of image 94 is same as that of the mehendi
- The image 95 shows that the hand is the brightest

compared to the background after mehendi removal, so it can be segemented using basic thresholding.

- **Conclusions:**
  - Since considering only the red plane, leaves only the intensity values of the red colored segements. We can conclude, that by considering only the red plane, we are taking only the intensity values which will be same in a given region. Therefore, the mehendi region blends into the skin region.

- **Subprocess III & IV: Binarization and Removal of unwanted components:**
  - *Methods:*
    - Binazary the image
    - Erode image to break small bonds
    - Keep the largest connected compoment and remove the rest.
    - Region filling
  - *Results and Observations:*


Image 96


Image 97


Image 98

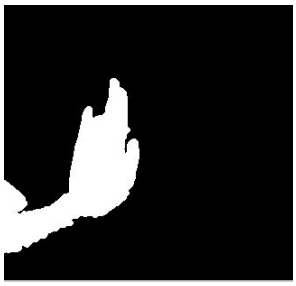


Image 99


Image 100

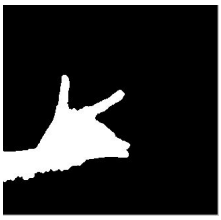


Image 101

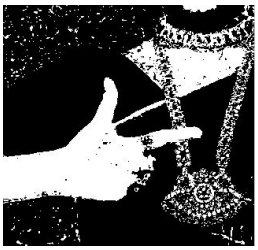


Image 102

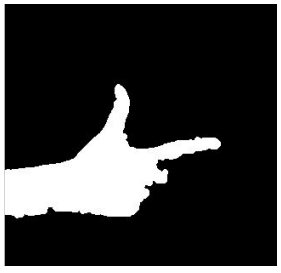


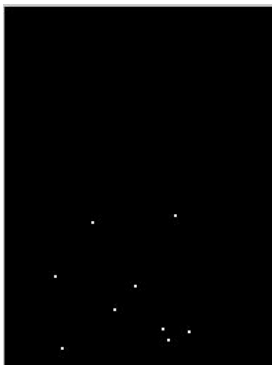
Image 103

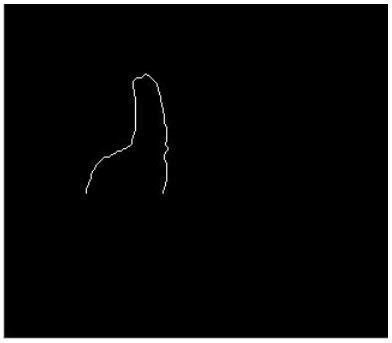
Observatsions:

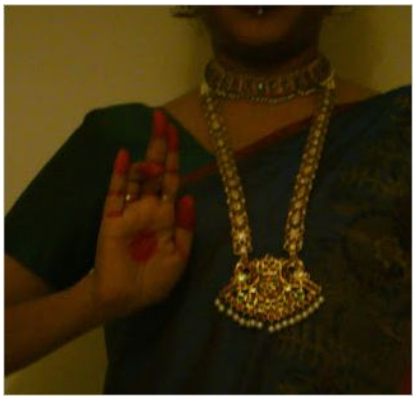
- Eroding the image broke the small connectors, as the lining of the sari which was connecting the hand to the neck region
- The necklace toching the hand in image 102, has also been removed because of erosion
- Erosion did not cause distortion of the hand
- In image 96, the edge of the idol seems to be touching the hand, but because of erosion it has been removed
- Region were the ring was has been filled in image 103
- Similarly the gap between the fingers has been filled in image 101

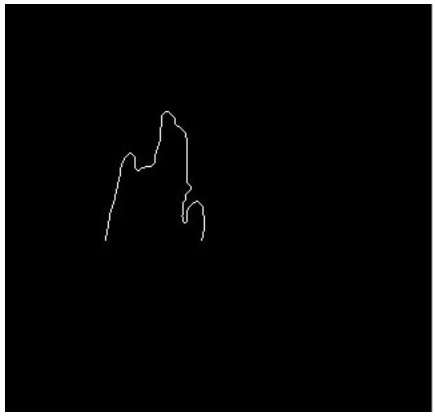
- **Conclusions:**
  - Erosion breaks small connectors
  - Since only finger -tips and ends of fingers are important, region filling can be used to fill in the gaps between fingers

- **Subprocess V & VI:**
  - *Method:*
    - Same as done for plain background
      - *Results and Observations:*





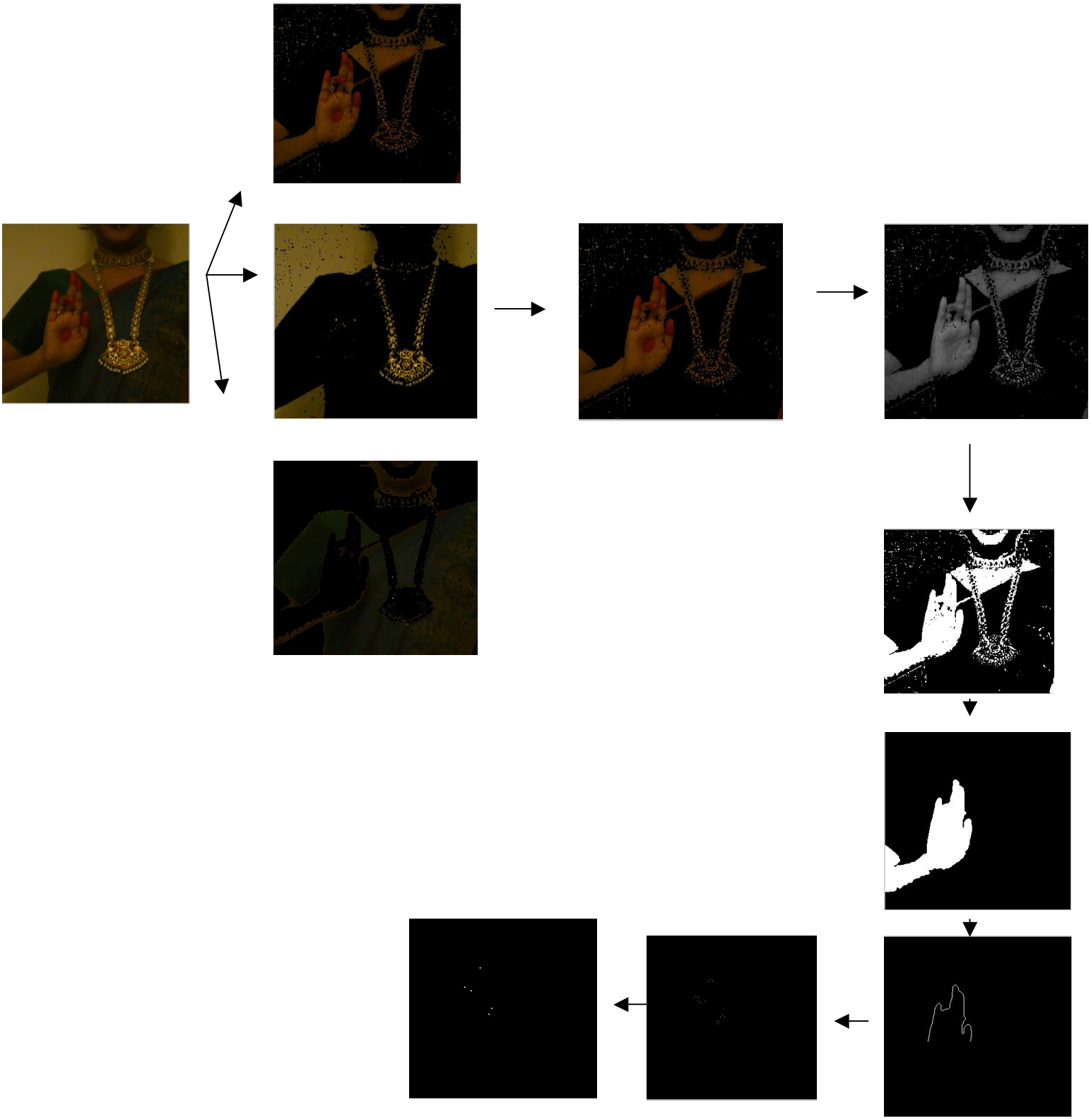




Observations and Conclusions:

- Boundary extraction gave same results as for plain background because the hand was segmented out from background and then the processes were applied.

Overall Process:



Extensions:

To make the code work for any of the cases in the three datasets, we have:

Convert the image to HSV color space

Identify the mode of each of the datasets in the H plane that distinguishes them.

Apply the threshold according to the mode, and run the codes for each of the datasets according to the specifications

Observations: This was based on the observation that the first dataset had a blue background, so mode will be closer to 1, while the double hand mudras and single hand mudras in complex background ususally had the colors of green and yellow, which will be near 0. And the plain background single hand mudras with accessories had a black background, so mode will be 0. Using this observation, we made the above process.

The pattern recognition part of the project can be done by applying a classifier to the fingers detected. But we have not concentrated on that part of the project as it is irrelevant to the course.

Successes and Failures:

Sucess: Single hand mudras have been properly segmented with and without accessories in plain and complex backgrounds

Failure: Double hand mudras were a bit of a challenge, because we were not able to distinguish between the hands, which caused a problem in feature extraction, because the boundaries were not properly extracted.

Conclusions and Knowledge Achieved:

Segmentation can be done using many methods based on type of dataset

Segementation is not a fixed process, but changes from dataset to dataset depending on what is wanted

Analysis in the RGB plane does not give best results but it can be used as well.

HSV plane is the best plane for analysis because of seperation of color, saturation

and intensity values

Morphological operations are useful in breaking small bonds, and filling in regions to get proper results

Neighborhood processing is very importan for image processing. Most operations implmented require neighborhood processing

Feature extraction is a tricky matter, because there can be clashes between features