

Navigation of PowerPoint Using Hand Gestures

Dnyanada R Jadhav¹, L. M. R. J Lobo²

¹M.E Department of Computer Science & Engineering, Walchand Institute of technology, Solapur, India

²Associate Professor in Department of Computer Science & Engineering, Walchand Institute of technology, Solapur, India

Abstract: *Computer applications require interaction between human and computer. This interaction needs to be unrestricted and it made challengeable accounts as compared to traditional input devices such as keyboard, mouse, pen etc. Hand gestures are used in daily life while communication. They are more natural in interaction, compared with those devices mentioned above. Human computer interaction becomes easy, with use of hand as a device. Use of hand gestures to operate machine would make interaction interesting. Gesture recognition has gained a lot of importance. Hand gestures are used to control various applications like VLC media player, robot control, gaming, virtual mouse etc. Use of gesture makes interaction easy, convenient and does not require any extra device. Vision and audio recognition can be used together. But audio commands may not work in noisy environment. The presented developed system uses hand gestures to control power point navigation. Here both static and dynamic gestures are used together for navigation. Use of gestures would make the presenter to handle the presentation more easily.*

Keywords: Hand gestures, skin segmentation, active fingers, finger count.

1. Introduction

Gesture recognition has gained a lot of importance since few years. Various applications can be controlled using gestures. Face gestures like lip movements are used to recognize the language spoken, hand gestures are used in various applications like gaming, mouse control etc. In an application like robot control using hand gestures, the robot responds to hand gestures given by the human. This hand sign of humans is visually observed by a robot through a camera. The algorithm that enables the robot to identify the hand gesture from the image is of interest. Each gesture corresponds to a particular command. The command that is identified will be used to control the robot to perform certain action or to execute a certain task. Different gestures will have different meaning associated with them. For example, count one could mean stop, two for move forward, three, four and five for turn right, turn left and reverse respectively. The hand gesture recognition system makes use of gloves, markers etc. Though the use of gestures increases the interactivity between man and machine, the use of such gloves or markers increases the cost of the system. Some applications require the use of two cameras to obtain a 3D view of hand and from which a hand gesture is recognized.

Two types of hand gestures are used, static and dynamic. Static gestures make gestures by keeping the hand stable. For example, by using the finger without moving the hand, the system would perform the specified function. Dynamic gestures are those that involve the movement of the hand. Like in VLC media player that is controlled by hand gestures, moving the hand to the right side may indicate increasing the volume. For some application hand gesture recognition system may require to store images in the database. Execution of these applications may require the use of complex algorithm to compare images already stored in the database with that of images taken from the camera and then perform necessary tasks. For such applications gestures should be known prior the use as they are already stored in the database. In a Hand Gesture Recognition System to control Slide Navigation there is no use of database. The

system uses both static and dynamic gestures for slide show control. In this system the images are captured through the webcam. It is segmented to recognize the hand region. A skin color detection algorithm is used for hand region detection. The binary image that is generated is given to the distance transform method that calculates the centroid of the hand and using this it recognizes the gesture. Accordingly the slide show is controlled.

2. Related Work

Many applications have been developed that are controlled through gestures that include gaming, sign language recognition, control through facial gestures, controlling mouse, VLC media player etc. In 2012 [1] Ruize Xu, Shengli Zhou and Wen J. Li developed a system that recognizes seven hand gestures like up, down, right, left, cross and circle. Three different modules were built that recognized hand gesture. Signals from MEMS 3-axes accelerometers were given as input. The motion of the hand in three perpendicular direction is detected by 3 accelerometers and transmitted to the system through Bluetooth. Segmentation algorithm was applied and finally the gestures are recognized by comparing gestures that are already stored in the system. People always use the internet to get daily information about weather, news etc. For this they have to repeat same keyboard and mouse actions. In 2011 [2] Kuan-Ching Li, Hwei-Jen Lin, Sheng-Yu Peng, Kanoksak Wattanachote used hand moments to retrieve information from the internet which reduces time and also convenient to use. Once the user provides the gesture, appropriate function is selected, then the system will report the action to the user in form of speech. This system also uses face recognition to identify and personalize each user as requirements are different for different users. They used the PCA method to recognize hand and face. Hand gestures were acquired and stored in the system and the later compared this with input gestures and perform tasks. This system results better with the small scale of face recognition and hand gesture recognition. In 2011 [3] Ginu Thomas presented an article on A Review of Various Hand Gesture Recognition Techniques where he compared

the results obtained by different techniques. The different methods used are edges method, pixel by pixel comparison and orientation histogram. Image database was used that contained various static hand gestures images. These images are subset of American sign languages. Filtering of the image was done to reduce noise present in it and then segmentation to analyze it. It was then transformed into feature vector and then compared with a trained set of gestures. In 2010 [4] Anupam Agrawal and Siddharth Swarup Rautaray used hand gestures to control the VLC media player. The K nearest neighbor algorithm was used to recognize the gesture. A VLC media player function that were controlled by hand gesture includes play, pause, Fullscreen, stop, increase volume, decrease volume. Lucas Kanade Pyramidal Optical Flow algorithm was used to detect hand from video. This algorithm detects moving points in the image. After this K_MEAN was used to find a center of the hand. Using this centre motion of the hand is recognized. This system used database that contains different gestures and then input was compared with this stored image and accordingly VLC media player was controlled. In 2007 [5] Yikai Fang, Jian Cheng and Hanqing Lu, Kongqiao Wang recognized hand through trigger followed by tracking and segmentation and used this gesture for image browsing. Segmentation is done during detection and tracking using motion and color cues. Palm and finger like structures are then determined using Scale Scape features. Using this palm finger configuration, hand gesture is determined. In 2007 [6] Yikai Fang, Jian Cheng, Hanqing Lu, Kongqiao Wang speed up hand gesture recognition. They used integral image approximation for Gaussian derivation in image convolution. Fast multi-scale feature detectors were constructed to speed up computation. In 2006 [7] Erol Ozgur, santerabi Malima controlled robot using hand gestures but considered limited gestures. Firstly segmentation of hand region was carried followed by locating the fingers and then finally classifying the gestures.

In 2002 [9] Lars Bretzner, Ivan Laptev, Tony Lindeberg presented algorithms for hand tracking and hand posture recognition. In this on each image multi-scale color feature detection is performed. Using particle filtering, with an extension of layersampling referred to as hierarchical layered sampling Hands are detected and tracked. In 1994 [10] William T. Freeman, Craig D. Weissman used hand gestures and controlled the television. Here a user makes a gesture facing the camera. On the display the hand icon appears that follows the user. User can move his hands on various graphical controls using this hand icon and control the television. Gestures for a television control like television off, television on, channel control, channel change were used.

3. Matlab 2012

MATLAB means matrix laboratory. It is a high-level language and it is interactive environment for visualization, numerical computation and programming. MATLAB is used to develop algorithms, analyze data, create models and applications. Availability of tools and built in math functions helps to achieve solution faster than with traditional programming languages such as Java or C/C++. MATLAB can be used for image and video processing, signal processing and communications, computational finance,

control systems, test and measurement, and computational biology. Array is basic data element of MATLAB that does not require dimensioning thus allows solving many technical computing problems. The reason for choosing MATLAB for development is its toolboxes. Toolboxes allow to learn and to apply specialized technology. Toolboxes are collections of MATLAB functions i.e M-files that are used to solve particular classes of problems.

4. Methodology

The System architecture is as shown in Fig (1); it consists of image acquisition, segmentation of hand region, finger count recognition and motion recognition and finally slide show control.

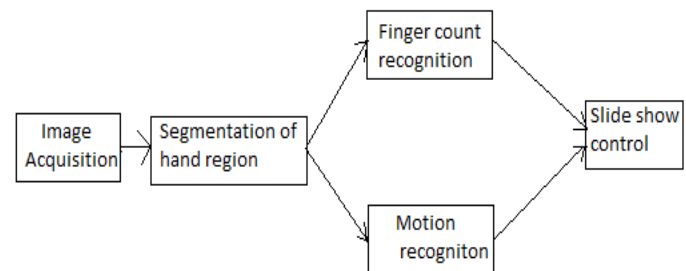


Figure 1: System architecture

Image Acquisition

The user makes gestures by positioning hand parallel to webcam. Images are continuously captured and then given as Input for segmentation.

Hand Detection

The primary step in gesture recognition systems is the detection of hands and the segmentation of the corresponding image regions. This segmentation is crucial because it isolates the task-relevant data from the image background, before passing them to the subsequent tracking and recognition stages. A large number of methods have been proposed in the literature that utilize a several types of visual features and, in many cases, their combination. Such features are skin color, shape, motion and anatomical models of hands.

Segmentation of hand region

The images captured are given for analysis which is done using segmentation. Here skin detection algorithm is used that detects the skin region from the input image as the background may consist of many other things along with the hand region. The video obtained through a webcam is in the RGB color model. This video is converted to HSV color model because the regions which belong to the skin can be easily identified in HSV model. Following this, the rules for skin segmentation are applied. After recognizing hand it is converted into a binary image. The skin regions are represented using white color and all other non-skin regions are black. The largest connected region, which is detected as skin is taken as the hand region. This gives the segmented hand region and this is the region of interest. The recognition of the gestures depends on this region.

RGB to HSV conversion

ALGORITHM

- Load image.
- Read each pixel from image.
- Separate RGB color for each pixel.
 $R = col \& 0xff;$
 $G = (col \gg 8) \& 0xff;$
 $B = (col \gg 16) \& 0xff;$
- Find minimum value and maximum value from R,G,B.
- Assign max to value.
- If value equal to zero then
 assign hue=saturation =0.
 set pixel in image again
 if valueof (Lum) = 0 then output color is black i.e there is
 no brightness in the color hence Hue and Saturation will
 be zero.

```

Else
//Formula for finding Saturation.
Find saturation= 255 * (Max-Min)/value.
if saturation = 0 then
assign hue=0.
set pixel
end if
Else
if max equal to R then
//Formula for finding Hue.
hue = 0 + 43*(G-B)/(max-min).
End if.
If max is equals G then
hue = 85 + 43 * (B-R) / (max-min).
End if
If max is equal to B then
hue = 171 + 43*(B-R)/(max-min).
End if.
If hue<0 then
hue=hue+255.
End if.
End if.
• Set each pixel again on image.
end.
    
```

After this we get binary image

Distance transform method for gesture recognition

After segmentation the binary image is given to the distance transform method that recognizes the gesture. In this method firstly the centroid of palm is calculated by considering each pixel and calculating distance from nearest boundary. Therefore the pixel that is far from every boundary is chosen as centroid. Using this centroid active fingers are counted and if there is motion of hand, this is detected by motion of centroid from original position from a set of continuously captured images and the slide show is controlled that is PowerPoint presentation either goes to the next slide, previous slide or start after recognizing static gesture.

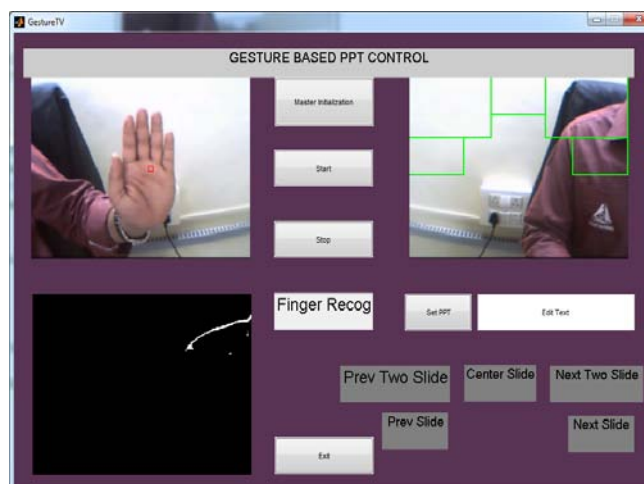
ALGORITHM

- Consider binary image.

- Check for all the pixels that are on boundary and assign it some value
- For each pixel in image an image other than boundary image
 a. Calculate distance of that pixel from every pixel on boundary
- Assign some value for the pixel whose distance from every boundary pixel is maximum
- This pixel is the centroid of the image. Using this centroid slide show is controlled.

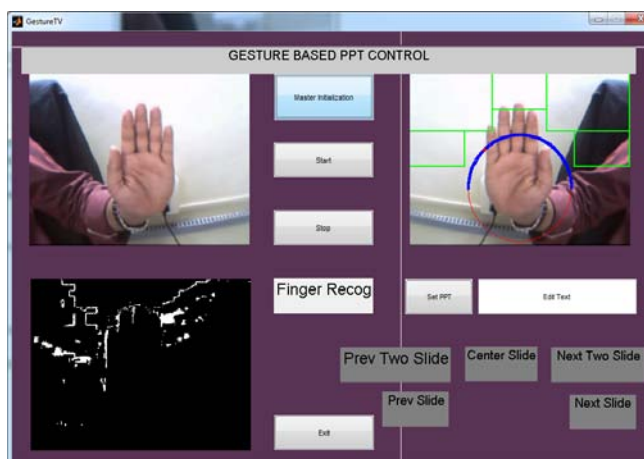
5. Results

Following snapshots illustrates the results achieved. Initially we get the screen shown in snapshot 1. It contains various buttons, like Master initialization, start, stop, previous slide, previous two slides, centre slide, next slide, and next two slides. In the screenshot can also see the button as Set PPT, it is used to select the ppt that we want to navigate.



Snapshot 1: Initial Screen

We can see three segments showing three different screens. First screen is used for master initialization, next is actual image captured through webcam and the last shows the binary image that is obtained using skin color segmentation algorithm. When the hand is detected, immediately is marked with circle as shown in snapshot 2.

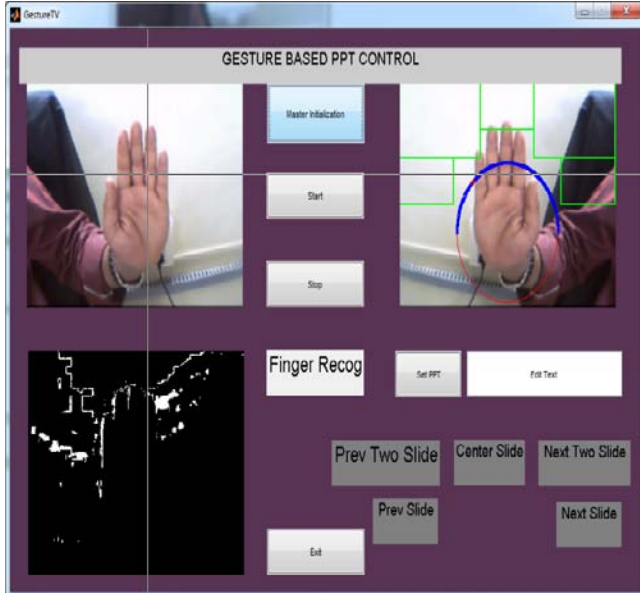


Snapshot 2: Detection of hand

Initially we do master initialization as shown in below snapshot 3.

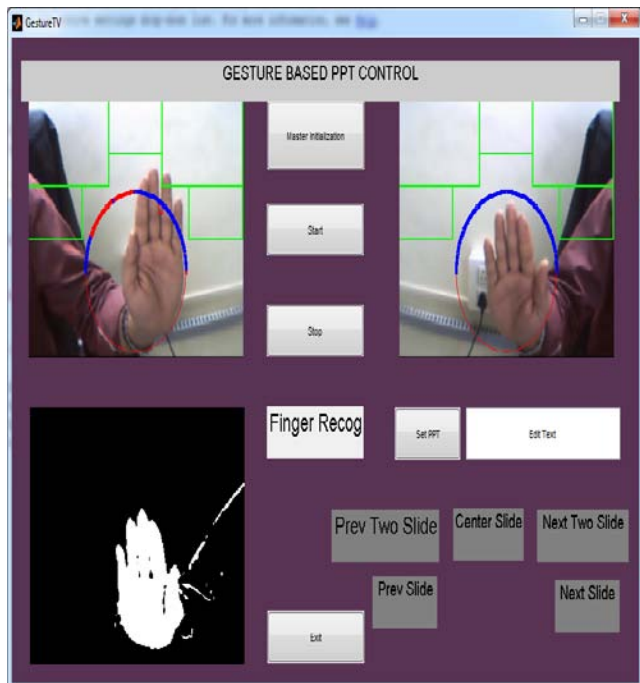
For this we drag to the point that is concentrated on the hand region.

This master initialization is done so that it can recognize hand of different skin tone.



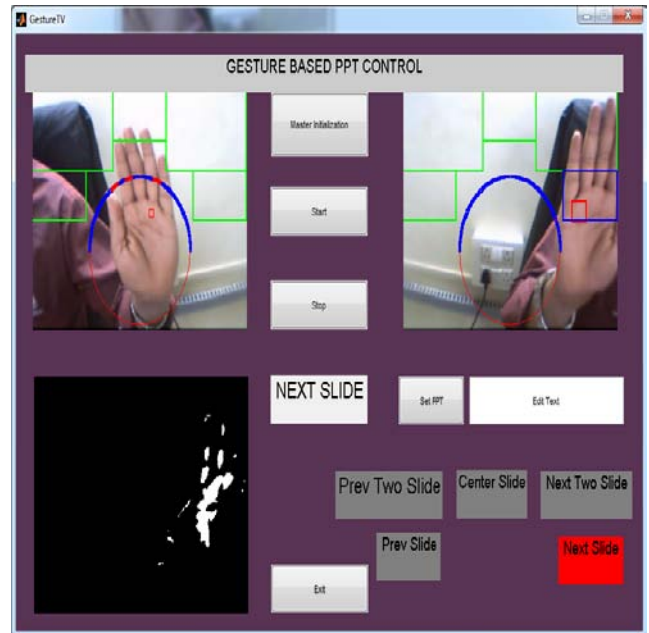
Snapshot 3: Initialization of hand region

We can see in snapshot 3 that we have segmented the region, as shown in green color.



Snapshot 4: Displays the binary image of the hand

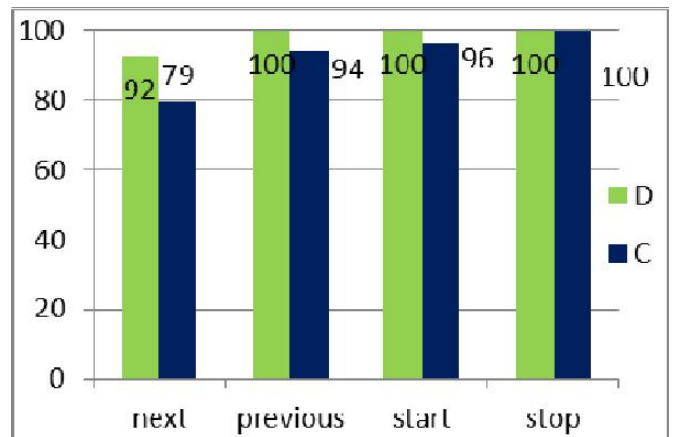
As shown in snapshot 4, whenever we drag our hand in segment as shown we can see that Next slide button is triggered. This means we navigate to next slide.



Snapshot 5: Displays that the next slide button is triggered after gesture recognition

Snapshot 5, Displays that the next slide button is triggered after gesture recognition

Comparison between distance transform method and circular profiling method for centroid detection is shown in below graph that results distance transform method is better in performance than circular profiling method.



6. Conclusion

The presented gesture recognition system recognizes both static and dynamic gestures. Gestures are recognized using distance transform method along with skin color segmentation algorithm. The presented method does not require any training phase to identify the hand gestures. Hence does not require storage of images in database to recognize the hand gestures. A recommended future work would include, increasing the number of gestures along with speech recognition that make it interesting and easy to navigate power point which will include creating slides, adding contents etc. Furthermore gestures can be used to control various applications like adobe reader, paint and add further by controlling computer using gestures.

References

- [1] Dnyanada Jadhav, L.M.R.J. Lobo, "Hand Gesture Recognition System To Control Slide Show Navigation", IJAIEM, 2014
- [2] Ruize Xu, Shengli Zhou, Wen J. Li, "MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition", IEEE, 2011.
- [3] Sheng-Yu Peng, Kanoksak Wattanachote, Hwei-Jen Lin, Kuan-Ching Li, "A Real time hand gesture recognition system for daily information retrieval from Internet", IEEE Fourth International Conference on Ubi-Media Computing, 2011
- [4] Ginu Thomas, "A Review of Various Hand Gesture Recognition Techniques", 2011.
- [5] Siddharth Swarup Rautaray and Anupam Agrawal, "A Vision based Hand Gesture Interface for Controlling VLC Media Player", International Journal of Computer Applications, 2010.
- [6] Yikai Fang, Kongqiao Wang, Jian Cheng, Hanqing Lu, "A Real-Time Hand Gesture Recognition Method", IEEE, 2007.
- [7] Yikai Fang, Jian Cheng, Kongqiao Wang, Hanqing Lu, "Hand Gesture Recognition Using Fast Multi-scale Analysis", IEEE, 2007.
- [8] Asanterabi Malima, Erol Ozgur, "A Fast Algorithm For Vision-Based Hand Gesture Recognition For Robot Control", IEEE, 2006.
- [9] Ahmed Elgammal, Vinay Shet, Yaser Yacoub, Larry S. Davis, "Learning Dynamics for Exemplar-based Gesture Recognition", IEEE, 2003.
- [10] Lars Bretzner, Ivan Laptev, Tony Lindeberg, "Hand Gesture Recognition using Multi-Scale Color Features, Hierarchical Models and Particle Filtering", IEEE, 2002.
- [11] William T. Freeman, Craig D. Weissman, "Television Control by Hand Gestures", 1994.

Author Profile

Miss Dnyanada Jadhav, received B.E degree in Computer Science and Engineering in 2012 from Walchand Institute Of Technology Solapur. Solapur University, Maharashtra, India and pursuing the M. E. degree in Computer Science and Engineering in Walchand Institute of Technology, Solapur, India. She is doing her dissertation work under the guidance of Mr. Lobo L.M.R.J, Associate Professor & Head, Department of IT, Walchand Institute of Technology, Solapur Maharashtra, India. Her research area is Image Processing.

Mr. Lobo L. M. R. J received the B.E degree in Computer Engineering in 1989 from Shivaji University, Kolhapur, India and M. Tech degree in Computer and Information Technology in 1997 from IIT, Kharagpur, India. He is registered for Ph.D in Computer Science and Engineering at SGGGS, Nanded of Sant Ramanand Teerth Marathawada University, Nanded, India. Under the guidance of Dr. R.S. Bichkar. He is presently working as an Associate Professor & Head, Department of IT Walchand Institute of Technology, Solapur, Maharashtra, India. His research interests include Evolutionary Computation, Genetic Algorithms and Data Mining.