Advanced Computing: An International Journal ( ACIJ ), Vol.3, No.3, May 2012

# DISTANCE TRANSFORM BASED HAND GESTURES RECOGNITION FOR POWERPOINT PRESENTATION NAVIGATION

Ram Rajesh J[1] Nagarjunan D[2] Arunachalam RM[3] and Aarthi R[4]

[1,2,3] Department of Information Technology, Amrita VishwaVidyapeetham, Coimbatore, India
[1]ram_rajeshj@yahoo.com
[2]naga.amrita@gmail.com
[3]arun.nsn.che@gmail.com
[4]Assistant Professor Department of Information Technology, Amrita VishwaVidyapeetham, Coimbatore, India
[4]r_aarthi@cb.amrita.edu

***ABSTRACT***

*Information conveyed in seminars, project presentation or even in class rooms can be effective when slideshow presentation is used. There are various means to control slides which require devices like mouse, keyboard, or laser pointer etc. The disadvantage is one must have prior knowledge about the devices in order to operate them. This paper proposes two methods to control the slides during a presentation using bare hands and compares their efficiencies. The proposed methods employ hand gestures given by the user as input. The gestures are identified by counting the number of active fingers and then slides are controlled. Unlike the conventional method for hand gesture recognition which makes use of gloves or markers or any other devices, this method does not require any additional devices and makes the user comfortable. The proposed method for gesture recognition does not require any database to identify a particular gesture. The experiment was tested under different kinds of light sources like incandescent bulb, fluorescent lamp and natural light.*

***KEYWORDS***

*hand gestures, skin segmentation, active fingers, finger count.*

## 1. INTRODUCTION

Vision based gesture recognition techniques are of major interest in the field of research. People can interact with the system in a device-free manner and this property of vision based hand gestures make them user friendly. The hand gestures must be identified in any environment .i.e. under varying illumination conditions. The image or video acquired as input may be noisy or may reduce the performance by recognizing surrounding as hand region. The acquired data is subjected to segmentation and processed further to make it fit for approximation with the gestures (data) stored in the database.

The other means of detecting hand gestures involves usage of markers or gloves to identify the hand gestures [1], [2], [3]. Some acquire the hand gestures using two cameras to obtain the 3D view of the hand and from the 3D model of the hand and then gestures are recognized [4]. But it involves storage of images of hand to compare with the acquired data and makes use of complex algorithm to compare the images and identify the correct gestures. [5], [6] and [7] involves training phase to capture the gestures and then are used to compare with the acquired input.

DOI : 10.5121/acij.2012.3304 

Controlling the slideshow is a vital task during presentation. The slides must be controlled according to the presenter's requirement. There are various ways to control the slides but most of them depend on external devices such as mouse, keyboard, laser pointer, etc. [8]. As described above the user may carry the device or may wear some bands or markers or gloves to control the slides with hand gesture. Some of these gloves are connected to the computer to detect the movement of hand which makes gesture recognition a complex task [1]. [9], [10] uses distance transform techniques but they use database to recognize hand position which is time consuming and complex.

This paper suggests two techniques to control the slides of PowerPoint presentation in a device free manner without any markers or gloves. Using bare hand the gesture is given as input to the webcam connected to the computer. Then using an algorithm which computes the number of active fingers, the gesture is recognized and the slideshow is controlled. The number of active fingers are found using two techniques namely using circular profiling and distance transform.

The proposed method involves segmentation of the hand region from the acquired data. Then the centroid of the segmented hand is calculated following which the number of active fingers is found. Then the gesture is recognized. This does not involve storage of data. So controlling the slideshow during a presentation becomes user friendly.

In this paper section II deals with the segmentation of hand region which is common for both the methods. Section III and IV deals with the calculation of active fingers using circular profiling and distance transform method. The implementation of these methods is discussed in section V. The experimental results obtained for the methods mentioned above and comparison of their performance is mentioned in section VI.

## .2. SEGMENTATION OF HAND REGION

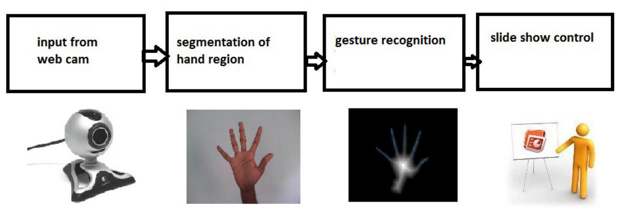



Figure 1. Architecture for hand gesture recognition

The general architecture to control the slide show using hand gesture is as shown in Fig 1.

The user makes the hand gestures by positioning the hand parallel to the webcam. The video is then processed to extract the hand region. The surrounding must be properly illuminated in order to minimise the error and the background should not contain any element that has skin colour.

The resolution of the webcam is kept at 640 x 480 pixels for better quality of video. In real world scenario the background may be made up of different elements. Hence a background subtraction is performed in order to segment the hand region from other regions.

The video obtained through webcam is in RGB colour model. This video is converted to HSI colour model because the regions which belong to skin can be easily identified in HSI model. Following this, the rules for skin segmentation are applied. The values for hue and saturation must be between 0.4 to 0.6 and 0.1 to 0.9 respectively.

$$0.4<H<0.6 \text{ and } 0.1<S<0.9 \qquad (1)$$

The regions with in the range of (1) are detected as skin and applying the above rule results in a binary image. The skin regions are represented using white colour 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. The skin segmentation for both circular profiling method and distance transform is the same. But while using distance measure two large connected skin regions are identified to detect two hands.

## 3. FINGER COUNT USING CIRCULAR PROFILING

The finger count is determined as in [11]. The centroid of the segmented binary image of the hand is calculated. Then the length of the largest active finger is found by drawing the bounding box for the hand. The centroid calculated is made as the centre and the value of radius is the length of the largest finger multiplied by 0.7 [12]. With the centroid as centre and length of the largest finger multiplied with 0.7 as radius a circle is drawn to intersect with the active fingers of the hand. If a finger is active then it intersects with the circle. A graph is used to count the number of transitions from white to black region. This number gives the number of active fingers. From the number of active fingers the gesture made can be determined. If a value less than 0.7 is used the circle drawn encloses only palm region. If a value greater than 0.7 is used the circle doesn't intersect the thumb.

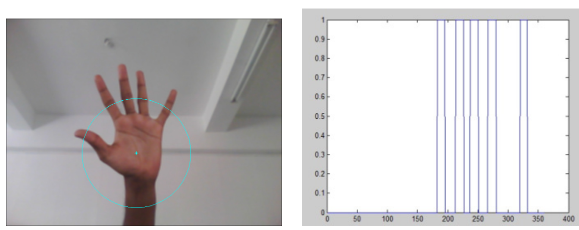


Figure 2. (a) image showing the circle intersecting the active fingers (b) the graph showing the transitions from black to white region

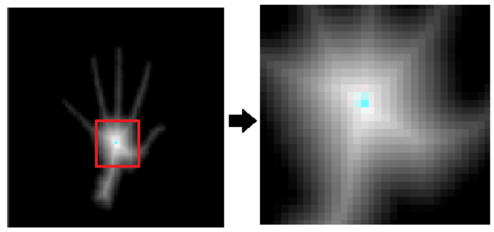
But the drawback in this method is that the hand should be properly placed with respect to the webcam so that the entire hand region is captured to draw the circle. If the hand is not placed properly the gesture is not recognized appropriately. Gesture made in this method involves only one hand and this reduces the number of gestures that can be made using both hands. Moreover the response time is very high. Hence a new method to overcome the drawbacks in circular profiling method is proposed.

## 4. FINGER COUNT USING DISTANCE TRANSFORM

### 4.1 Calculation of finger count

After the skin segmentation, the binary image of the segmented hand region $I_B$ is obtained and processed using distance transform.

The distance transform method gives the distance (Euclidean distance) of each pixel from the nearest boundary pixel. The distance from the boundary to a pixel in the hand region increases as the pixel is away from the boundary. Using this distance value, the centroid of the palm region can be calculated.



(a) (b)

Figure 3.(a)image showing hand region after applying distance transform (b) the enlarged image of the region with in the red rectangle

Fig. 3(a) shows the image $I_D$ of the hand after applying distance transform and the image Fig. 3(b) shows the enlarged view of the region within the red rectangle. The white colour in the centre is intense and the colour fades as the distance from the centre increases. From this it is evident that the pixels near the boundary have lower values for distance and the pixels away from the boundary have higher values for distance. This middle region which has the highest value for distance is considered as the centroid.

The width of the hand region will be approximately twice the distance from centroid to the nearest boundary pixel (say 2d) as shown in Fig. 4. The width of each finger is approximately one fourth of the width of the hand (i.e. ¼ of 2d). Now a suitable structuring element S (disc) that can erode the fingers completely is chosen and erosion is performed on the segmented hand region.

$$R_{P1} = I_B \ominus S \qquad (1)$$

After erosion only a part of the palm region $R_{P1}$ is left behind and the finger regions are completely eroded. Further the palm region which remains after erosion $R_{P1}$ is dilated using the same structuring element and this give the region $R_{P2}$ which is larger than the dilated palm region.

$$R_{P2} = R_{P1} \oplus S \qquad (2)$$

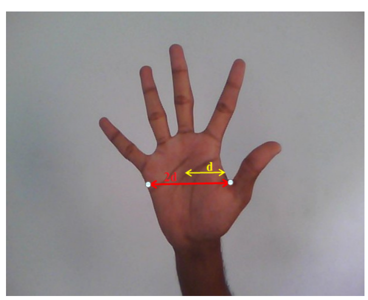
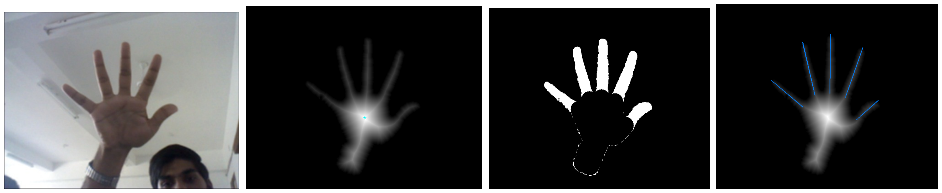


Figure 4. image showing width of the hand 2d and d the distance between the centre of the hand nearest boundary pixel

The dilated palm region $R_{P2}$ is subtracted from the original binary image $I_B$ to give the finger regions $F_R$ alone as shown in Fig. 5(c).

$$F_R = I_B - R_{P2} \qquad (3)$$

The number of fingers used to represent the gesture is found by drawing a line along the major axis of the segmented finger regions as shown in Fig. 5(d). The number of lines drawn is equal to number of active fingers. This count is used to control the slides of PowerPoint.

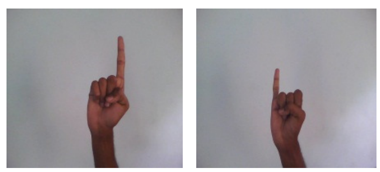


(a) (b) (c) (d)

Figure 5.(a) input image I from the user (b) image $I_D$ of hand region after applying distance transform (c) the finger regions $F_R$ after erosion, dilation and subtraction (d) the fingers count by drawing lines along the major axis of finger regions

### 4.2. Control the slideshow using hand gestures

Since the recognition of gesture depends purely on the number of active fingers any finger can be used to denote a count. Only the count value of the active fingers is taken as input. So the user can feel free to represent a count irrespective of the finger. Hence value one denoted by the user using the index finger or thumb will be the same as shown in Fig. 6 (a) and (b).



(a) (b)

Figure 6. (a) and (b) various gestures to represent number one

The slideshow is controlled by taking the finger count that is calculated using distance transform, as input from the user. The gestures used to control the slide show are mentioned in Table 1.

table1. Gestures And Their Functions

| **No of hand(s)** | **Finger count** | **Functions** |
|---|---|---|
| One hand | 0 | Transition from one gesture to another |
| One hand | 1 | Next slide |
| One hand | 2 | Previous slide |
| One hand | 3 | Slide show |
| One hand | 4 | Stop slide show |
| Two hands | Sequence of one digit numbers | Go to specified slide number |

When both the hands are used to represent the number of the slide to which the user need to navigate, first both hands are used to represent the first digit of the number and then both the hands are closed. After this both hands are used again to represent the next digit of the slide number.

## 5. IMPLEMENTATION

The hand gesture is found using the number of active fingers. Each hand gesture is mapped to a particular action as mentioned above. Each action is performed by a particular .Net function. These .Net functions are stored in a dll file and loaded in MATLAB. The functions in the file control the slideshow.

The algorithm was implemented on a system with Intel dual core processor with speed of 2.53 GHz using MATLAB software.

## 6. EXPERIMENTAL RESULTS

The work was experimented with more than 10 people under various illumination conditions such as fluorescent lamp, incandescent lamp and sunlight. The samples were carefully chosen so as to make sure that there are no elements having colour similar to skin colour in the surroundings.

Performance of both methods is compared and the results are shown using the graphs. Fig. 7(a) and (b) shows the performance of both methods for left hand and right hand respectively. It is found that controlling the PowerPoint presentation using distance transform method shows better results than circular profiling method. Also the response time is less compared to circular profiling method.

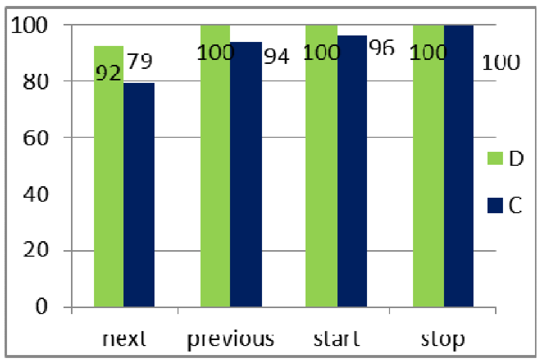



Figure 7(a). Graph showing comparision of performance of both Distance transfrom and circular profiling method when left hand is used

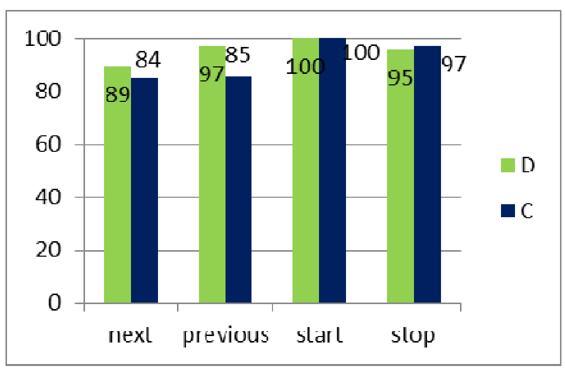



Figure 7(b). Graph showing comparision of performance of both Distance transfrom and circular profiling method when right hand is used

But it is not 100 % accurate. This decrease in efficiency is due to human nature to place the hand away from the focus of the camera. Improper gestures and gestures made immediately without a pause is also a reason for decrease in the level of accuracy. The efficiency decreases if the background has elements like wall hanging, drapes, furniture, etc. containing colour similar to skin colour. Problem occurs if the fingers are not stretched properly while making a gesture.

## 7. CONCLUSION

This paper deals with two algorithms in order to recognize the hand gestures and also compares their efficiencies. Both the algorithms suggest an alternative presentation technique to control PowerPoint presentation using bare hands. The proposed methods do not require any training phase to identify the hand gestures. Hence does not require storage of images in database to recognize the hand gestures. The hand gestures are recognized based on number of active fingers used to represent a gesture. So gestures can be made using any finger. The number of active fingers identified using distance transform method is less time consuming when compared to that of circular profiling method. Moreover both hands can be used in distance transform method and the slide show can be controlled effectively when this method is used. Also the advantage with this method is that the user can go to the desired slide by making gesture using both hands and this is not possible in circular profiling method. Usage of hand gestures can be extended to control real time applications like VLC media player, paint, pdf reader etc.

## ACKNOWLEDGEMENTS

We would like to thank Mr. Raghesh Krishnan, Asst. Professor, Amrita VishwaVidyapeetham, Coimbatore and Ms. Padmavathi, Asst. Professor, Amrita VishwaVidyapeetham, Coimbatore, for their guidance and support for the work. We would also like to thank our friend Ashok Gowtham, Final year student, Dept. of Information Technology, Amrita VishwaVidyapeetham, Coimbatore, who are also a part of the project, for their contribution for the paper and the project.

## REFERENCES

[1] Hae Youn Joung and Ellen Yi-Luen Do, "Tactile hand gesture hand gesture recongnition through haptic feedback for affective online communication", proceedings of *6th international conference on Universal access in human-computer interaction,* USA, July 2011.

[2] Robert Y. Wang, "Real-time hand tracking as user input device", proceedings of *5th international conference on Knowledge, information, and creativity support systems,* Thailand , November 2010.

[3] C. Keskin, A. Erkan and L. Akarun , " Real time hand tracking and 3D gesture for interactive interfaces using HMM", proceedings of *International Conference on Artificial Neural Networks,* Turkey, 2003.

[4] Desmond Chik, Jochen Trumpf and Nicol N. Schraudolph, "Using an adaptive VAR model for motion prediction in 3D hand tracking", proceedings of *8th international IEEE conference on automatic face detection and gesture recognition*, September 2008.

[5] Xioming Yin, and Xing Zhu, "Hand pose recognition in gesture based human-robot interaction", proceedings of IEEE conference on *Industrial Electronics and Applications*, Singapore, May 2006

[6] Siddharth Swarup Rautaray and Anupam Agrawal , "A Vision based Hand Gesture Interface for Controlling VLC Media Player", published in *International Journal of Computer Applications* 10(7):11–16, November 2010.

[7] Jianjie Zhang, Hao Lin, Mingguo Zhao, "A Fast Algorithm for Hand gesture recognition using relief", proceedings of *6th international conference on Fuzzy systems and knowledge discovery*,China, August 2009.

[8] Xian Cao, Eyal Ofek and David Vronay , " Evaluation of Alternative Presentation Control Techniques*",* proceedings of *ACM CHI 2005 Conference on Human Factors in Computing Systems,* USA, April, 2005.

[9] Parker, J.R., Baumback, M, "Visual Hand Pose Identification for Intelligent User Interfaces", proceedings of *16th international conference on Vision Interface ,Halifax, Nova Scotia*, Canada, June 2003..

[10] Junyeong Choi, Hanhoon Park and Jong-Il Park, "Hand shape recognition using distance transform and shape decomposition, proceedings *of 18th IEEE international conference on image processing*, Brussels, Belgium, September 2011.

[11] Ram Rajesh J, Sudharshan R, Nagarjunan D, Aarthi R, "Remotely controlled PowerPoint presentation navigation using hand gestures", unpublished.

[12] Asanterabi Malima, Erol Özgür, and Müjdat Çetin, "A fast algorithm for vision-based hand gesture recognition using for robot control", *IEEE conference on Signal Processing and Communications Applications*, Antalya, Turkey, 14th, April 2006.

## Authors

### Mr. RAM RAJESH J

He is currently doing his final year, B. Tech in Information Technology at Amrita VishwaVidyapeetham, Coimbatore, India. His areas of interest are image processing and networks.

### Mr. NAGARJUNAN D

He is currently doing his final year, B. Tech in Information Technology at Amrita VishwaVidyapeetham, Coimbatore, India. His areas of interest are image processing and database management.

### Mr. ARUNACHALAM RM

He is currently doing his final year, B. Tech in Information Technology at Amrita VishwaVidyapeetham, Coimbatore, India. His areas of interest are image processing and networks.

### Ms. AARTHI R

She is an Assistant Professor in Department of Information Technology at Amrita VishwaVidyapeetham, Coimbatore, India. She has her bachelor degree in Information Technology from Bhartiyar University. She has completed her M. Tech in Computer Vision and Image Processing from Amrita VishwaVidyapeetham, Coimbatore, India. Her areas of interest in the field of research are computer vision based image processing and graphics.