POWER POINT SLIDE SHOW MOVEMENT USING HAND GESTURE RECOGNITION

***Sampada Muley, **Prof. A. M. Rawate**

**Student, Electronics &Tele-communication,*
*Chhatrapati Shahu Maharaj Shikshan Sanstha (*C.S.M.S.S.*)*
Aurangabad, Maharashtra, India
***Professor, Electronics &Tele-communication ,*
Chhatrapati Shahu Maharaj Shikshan Sanstha (C.S.M.SS)
Aurangabad, Maharashtra, India

ABSTRACT

In this paper, a real-time system that utilizes hand gestures to interactively control the presentation is proposed.Computers are widely used in all fields. However, the interaction between human and machine is done mainly through the traditional input devices like mouse, keyboard etc. To satisfy the requirements of users, computers need other ways to interact more conveniently, such as using speech or body language (e.g. gestures, posture). In this paper, we propose a new method supporting hand gesture recognition in the static form, using artificial neural network. The proposed solution has been tested with high accuracy (98%) and is promising.In this paper, project presentation or even in class rooms can be effective whenslideshow 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 the 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.

Keywords—gesture recognition, sign, skin color, skinsegmentation, active fingers, finger count.

INTRODUCTION

Interactive presentation systems use advanced Human Computer Interaction (HCI) techniques to provide a more convenient and user-friendly interface for controlling presentation displays, such as page up/down controls in a slideshow. Compared with traditional mouse and keyboard control, the presentation experience is significantly improved with these techniques. Hand gesture has wide-ranging applications[1]. In this study, we apply it to an interactive presentation system to create an easy-to-understand interaction interface.Gesture recognition is the process of recognizing and interpreting a stream of continuous sequential gesture from the given set of input data. Gestures are non-verbal information which is used to make computers understand human language and develop a user friendly human computer interface. Human gestures are perceived through vision and this paper aims to use computer vision to analyze different sets of gestures using human fingers and interpret them in order to control the system.

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 [2], [3], [4]. 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 [5]. 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. [6], [7] and [8] involves training phase to capture the gestures and then are used to compare with the acquired input.

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. [9]. 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 [10]. [11]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

SEGMENTATION OF HAND GESTURE

Sign language is a language that employs signs made by moving the hands combined with facial expressions and postures of the body. It is one of several communication options used by people who are deaf or hard-of-hearing. Gesture language identification is one of the areas being explored to help the deaf integrate into the community and has high applicability. Researchers use specialized equipment such as gloves or recognition techniques based on image processing through cameras and computers. Most image processing solutions are based on two main methods: rules and machine learning. In this paper, we propose a new method in the field of machine learning that can generalize hand gestures, and can be applied beyond the limit of usual hand gesture identification in the future using an artificial neural network and where the main contribution is in the feature extraction.

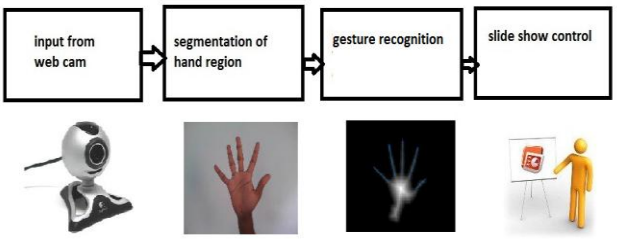


Fig 1 Architecture for hand gesture recognition

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 HIS 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
$0.4<H<0.6$ and $0.1<S<0.9$

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.

Capture the input: The first step is to capture the live Video stream from the camera. Camera has no infrared filter thus making it perfect to capture images which even during low light conditions. The color markers are identified and the video is converted into picture frames for gesture recognition.

Segmentation of hand: 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 minimize the error and the background should not contain any element that has skin color. The resolution of the webcam is kept at 640 x 480 pixels for better quality of video.

Gesture Recognition:

Gesture recognition is a topic in computer science and language technology with the goal of interpretinghuman gestures via mathematical algorithms. Gestures can originate from any bodily motion or state but commonly originate from the face or hand. Current focuses in the field include emotion recognition from the face and hand gesture recognition. Many approaches have been made using cameras and computer vision algorithms to interpret sign language. However, the identification and recognition of posture, gait, proxemics, and human behaviors is also the subject of gesture recognition techniques.

Gesture recognition can be seen as a way for computers to begin to understand human body language, thus building a richer bridge between machines and humans than primitive text user interfaces or even GUIs (graphical user interfaces), which still limit the majority of input to keyboard and mouse. Gesture recognition enables humans to interface with the machine (HMI) and interact naturally without any mechanical devices. Using the concept of gesture recognition, it is possible to point a finger at the computer screen so that the cursor will move accordingly. This could potentially make conventional input devices such as mouse, keyboards and even touch-screens redundant. Gesture recognition can be conducted with techniques from computer vision and image processing

GESTURE DETECTION

The system begins by analysing the captured video frame from the camera. Normally the video has to be cut in different images to identify the hand gestures using different color markers. The image obtained is often in the BGR format and it has to be first converted into HSV color space. There are several steps for gesture recognition which is shown in Figure.

Capture the input: The first step is to capture the livevideo stream from the camera. In reference [12], this camera has no infrared filter thus making it perfect to capture images which even during low light conditions. The color markers are identified and the video is converted into picture frames for gesture recognition.

Image acquisition: The next step is to create windows for different color markers which are detected. Resize the frames in order to reduce the resolution which will in turn reduce the computation time. In references [13, 14] noise reduction is done by using the antialiasing filters.

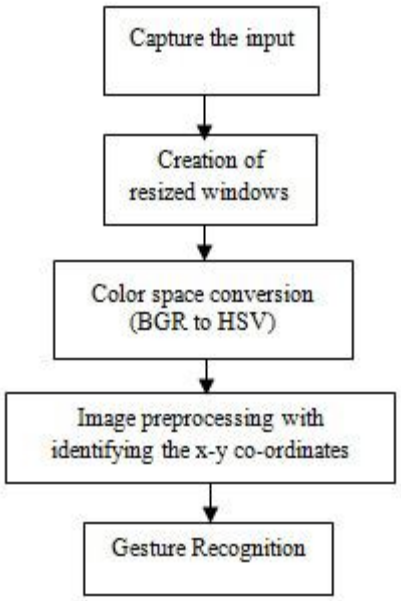


Fig 2 Framework of the system.

Color space conversion: A proper color model is needed to perform classification. The captured image which is in the BGR color space is converted to HSV (Hue Saturation Value). HSV color space is the most suitable one for color based image segmentation. Hue value represents the shades of the color, S describes how pure the hue color is and V provides the intensity of the color. Different HSV values for different colors are provided to represent the gray image of that particular color. The gesture recognition system is processed dynamically for computation and to reduce the complexity color markersare used for detection instead of skin tone detection which is used in reference [15].

Image pre-processing: In this phase the color markers are detected and the x-y co-ordinates are identified for the detected image. To identify the x-y co-ordinates of the color markers the size of the monitor is known and the x-y co-ordinates are calculated using pixel co-ordinate system where the position is identified by px=0 and py=0 corresponding to the top-left corner of the window. For each frame the system recognizes the color markers and the corresponding x-y co-ordinates are displayed. Reference [16] uses background modelling and calibration which makes the system more efficient in computation.

PROPOSED METHOD

Block diagram

Gesture language identification is one of the areas being explored to help the deaf integrate into the community and has high applicability. Researchers use specialized equipment such as gloves or recognition techniques based on image processing through cameras and computers. Most image processing solutions are based on two main methods: rules and machine learning. In this paper, we propose a new method in the field of machine learning that can generalize hand gestures, and can be

applied beyond the limit of usual hand gesture identification in the future using an artificial neural network and where the main contribution is in the feature extraction.

TRAIN MODEL

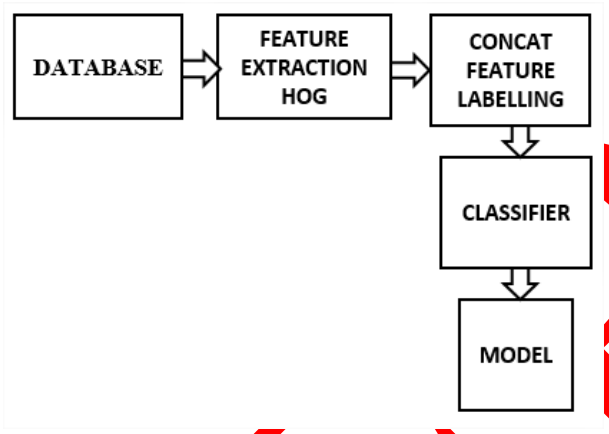


Fig 3Block Diagram of Train Model

As there are two model first is train model, in this model the image will created by database. Store the image into the database, thus the images will make the database then HOG (histogram of oriented gradient) it is a descriptor used in computer vision and image processing for object detection.

In this there are four database as follow:-

Gesture 1"ONE FINGER"PPT control-Next slide

Gesture 2"TWO FINGER"PPT control-Jump to second slide

Gesture 3"FOUR FINGER" PPT control-Previous slide

Gesture 4"GOOD SIGN"PPT control-exit PPT

The image is divided into small region called as cell and for pixels within each cell a histogram of gradient direction is complised.This Feature extraction HOG detected the region of image.

HOG Feature steps:-

1 Compute centered horizontal and vertical gradient with no smoothing.

2 Compute gradient orientation and magnitude

3 For 64*128 image, divide the image into 16*16 block

4 Each block should consist of 2*2 cell with size 8

This HOG Feature extraction and concatenated database for feature will control the model

RECOGNITION MODEL

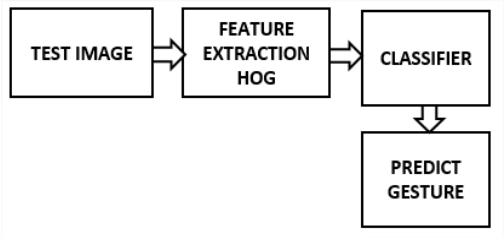


Fig 4Block Diagram of Recognition Model

In this test image which is classify by train model which is tested image will extracted feature into HOG model. This given train model predict gesture means it control the finger count.

FLOW OF BLOCK DIAGRAM

1 Detectpalm part (pre-processing)

- Segmentation using skin detection
- Morphological Operation

2 Extract Feature from Detected Region.

-HOG (histogram of oriented gradient)

-SIFT (Scale invariant feature transform)

-LBP (Local binary pattern)

3 Use Machine Learning tool neighbor Algorithm label.

-KNN (K-nearest neighbor algorithm)

-Neutral network.

4 Give test gesture to trained model and predict label (gesture recognition)

5 Control slides using gesture.

6 Compare result with the existing tech accuracy

The following will present the main steps in ourmethod.

A. Input Data and Training Data

Data can be an image or a sequence of images (video), taken by a single camera toward the human hand. However, some systems use two or more cameras to get more information about the hand pose. The advantage is that the system can recognize the gesture even if the hand obscured forone camera because the other cameras will capture the scene from different angles. A different system was presented with the camera mounted on a hat to take the area in front of the wearer. The obvious advantage of this system is the camera position is always appropriate if people move or turn around. In general, the following stages of the identification process will be less complex if the image is

taken with a simple background and the contrast is high with the hand. So pictures are usually taken in a homogeneous background environment, and limit shadows in the obtained image.

B. Pre-Processing

These are the necessary steps to get the complete hand picture from the original frame. Hand detection: To identify the hand gesture, the first needed step is detecting the hand from the input frame. Two commonly used techniques are backgroundSubtraction and skin color filter. In the proposed solution, we use the second method. Proposed by Fleck and Forsyth in [16], human skin color is composed by two poles of color: red (blood) and yellow (melanin), with medium saturation. Fleck also found that skin color has a low amplitude structure. The skin color characteristics are essential information and can be used in hand tracking algorithm. Skin Color Filter is proposed as follows: each pixel (RGB) is converted into log-component values I, Rg, and by [17] using the following formulas:

$$L(x) = 105 * \log10 (x + 1 + n)$$
$$I = L (G)$$
$$Rg = L(R) - L (G)$$
$$By = L (B) – (L (G) + L(R))/2$$

Where I, Rg, By are respectively log-components with color channels Green, Red, and Blue. The green channel is used to represent intensity because the red and blue channels from some cameras have poor spatial resolution. The constant 105 simply scales the output of the log function into the range [0,254]. n is a random noise value, generated from a distribution uniform over the range [0, 1). The random noise is added to prevent banding artifacts in dark areas of the image. The constant 1 added before the log transformation prevents excessive inflation of color distinctions in very dark regions.

C Median filter

In signal processing, it is often desirable to be able to perform some kind of noise reduction on an image or signal. The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing (for example, edge detection on an image). Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise.

D Remove the arm

We skip the parts not related to the hand; this is an important step in the identification process. When we remove these components, the problem near - away of the camera is removed, and the resulting image is the hand. This not only affects the accuracy but also affects the processing speed - an important factor in real-time applications. First, the image is resized by the object bounding box size. To get the region around the hand, we determine the position of the wrist and cut to separate the hand and arm.

The wrist detection algorithm is proposed as follows.

- Step 1: *mi* is defined as object's width at row

mi×pixel (object) □*rowi*

- Step 2: calculate *m* for the last row
- Step 3: calculate new *m* value for the line above
- Step 4: if *m* does not increase, go to step 3 else, crop image at the previous line

Fig 5 locate the wrist and separate

E Train Network and Recognition

We designed a Multiple Layer Perceptron network that has three layers: input layer has the number of neurons corresponding to the size of feature vector; the neurons of hidden layer determined by trial–and–error method; and output layer has 10 neurons matching 10 gestures.

SYSTEM FLOW CHART

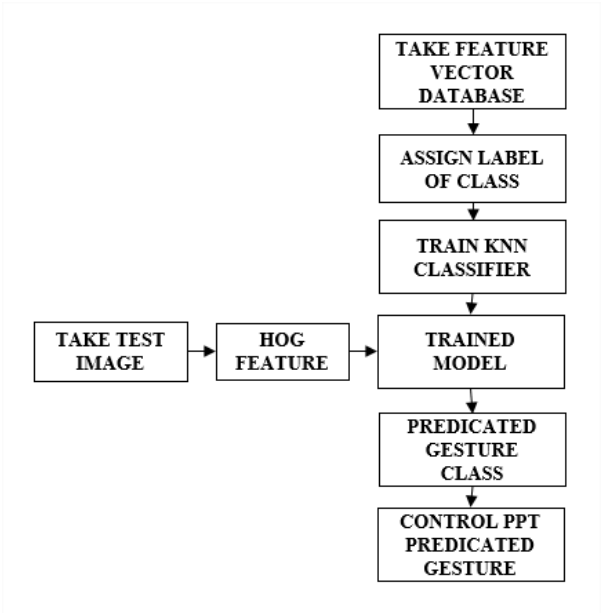


Fig 6 Flowchart for Train test algorithm

As the Fig 6 show the Train test algorithm is the system in which the database image will detect the palm such as it will detect the skin segmentation using skin detection.

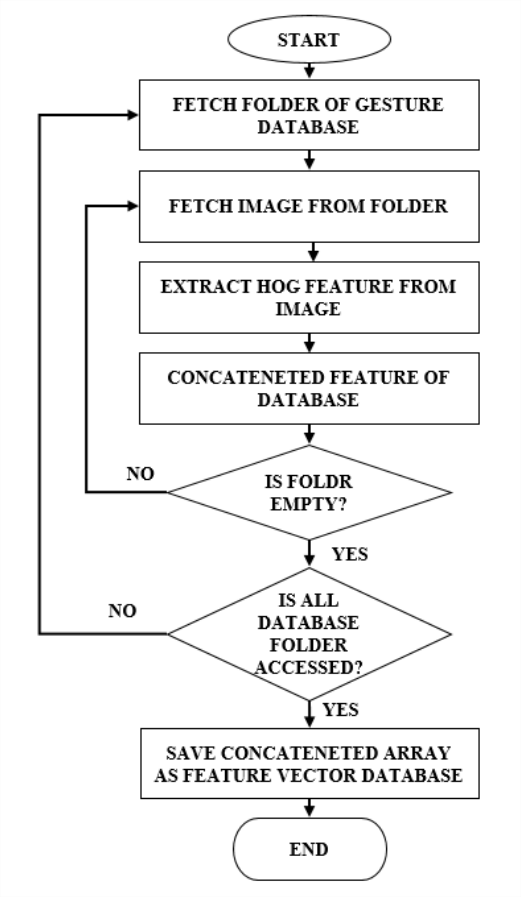
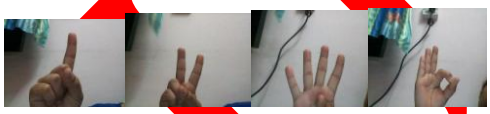


Fig 7 Flowchart of Extract feature from database

DATABASE



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

Fig 8 Images of database

We have taken database on hand gesture challenge. All this image are of size 250*350 are in JPEG format. Thus this are some date base images as above.

RESULT FOR EXPERIMENTATION

No of gesture	Finger count	Functions
Gesture One	One finger	PPT control-Next slide

Gesture Two	Two finger	PPT control-Jump to 2nd slide
Gesture Three	Four finger	PPT control-Previous slide
Gesture Four	Good sign	Slide show

CONCLUSION

This paper deals with two algorithms in order to recognize the hand gestures and also comparestheir 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 makinggesture 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.

ACKNOWLEDGMENT

I would like to express my gratitude towards my guide Prof.A.M.Rawate, who is professor of CSMSS, Aurangabad for his time to time help and valuable guidance.

REFERENCES

[1] Erol A, Bebis G, Nicolescu M, et al. Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding*, 2007, **108**(1-2): 52-73.

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

[3] 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.

[4] 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 NeuralNetworks,* Turkey, 2003.

[5] 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. [6] 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

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

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

[9] Xian Cao, Eyal Ofek and David Vronay , " Evaluation of Alternative Presentation Control Techniques*",* proceedings of *ACM CHI 2005 Conference on Human Factors in ComputingSystems,* USA, April, 2005. Advanced Computing: An International Journal (ACIJ), Vol.3, No.3, May 2012 48

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

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

[12] G. Senthilkumar, K. Gopalakrishnan and V. SathishKumar. 2014. "Embedded Image Capturing System Using Raspberry Pi System," in Proc. International Journal of Emerging Trends & Technology in Computer Science. Vol. 3, pp. 213-215, April.

[13] Saikat Basak and Arundhuti Chowdhury. 2014. "A Vision Interface Framework for Intuitive Gesture Recognition using Color based Blob Detection," in Proc. International Journal of Computer Applications, Vol. 90, No. 15, March.

[14] P.S. Gotmare. 2013. "Interacting with Computer System using Hand Gesture," in Proc. International Journal of Science, Engineering and Technology Research. Vol. 2, Issue 1, pp. 57-61, January.

[15] Mohamed Alsheakhali, Ahmed Skaik, Mohammed Aldahdouh and Mahmoud Alhelou. 2011. "Hand Gesture Recognition System," in Computer Engg Department, The University of Gaza.

[16] Bobo Zeng, Guijin Wang and Xinggang Lin. 2012 "A Hand Gesture Based Interactive Presentation System Utilizing Heterogeneous Cameras," in Tsinghua Science and Technology. Vol. 17, No. 3, pp. 329-336, June.

[17] M. Fleck, D. Forsyth, and C. Bregler, "Finding naked people," *European Conference on Computer Vision*, 1996.

BIOGRAPHIES

A.Sampada S Muley, student of M.E (E&TC) 2nd Year CSMSS Engineering College, Aurangabad, Maharashtra, India. B.E (EC) from SPWE College of Engineering, Aurangabad, and Maharashtra in 2013.She has published one paper in international conference.

B. Prof. A.M Rawate, Professor and head at CSMSS Shahu College of Engineering, Aurangabad.

BE and ME from JNEC Aurangabad and putting PhD from Dr Babasaheb Marathwada University, Aurangabad, Maharashtra. 05 Research Publications in National and International Conferences and Journals.15 Years of Industrial as well as Academic Experience out of which 9 Years at Administration of Engineering College. Authored a book Microprocessor and Computer Architecture with Dr G.S Sable for (CSE) Life member of ISTE and member of various Committees of Dr Babasaheb Marathwada University, Aurangabad, and Maharashtra.