Hand Gesture Recognition for MP3 Player using Image Processing Technique and PIC16F8779

Devikarani Patil, Varalakshmi B.D

***Abstract:** The scope of the project is to control MP3 player using gesture. Here, gesture image is taken from Web camera and image will be processed in remote interface using MATLAB controller. But, the challenging problem is that capturing the image from external device does not depend on unique only and Identification of the exact action from an unclear image is not an easy task. Hence, capturing action from the images is always puzzling task of separating different sources of images when its different or noisy. Finally, the images are forwarded to MATLAB to compare the images with our knowledge database via three dimension (x, y, and z) readings of a particular object. So if we move any object in any direction then the corresponding values are noted by the accelerometer. Most of the music players are controlled through the remote controls which contain buttons. But through embedding the PIC16F8779 controller, we can make music player be controlled by gesture performance in the air. The application of this three axis controller together with suitable interfacing with the PIC16F8779 micro controller and the music player development through coding in software platform such as MPLab IDE which could recognize the terminal input instructions and perform functions like play, stop, play back and play forward of music player controlled by gesture. We need to move the accelerometer in a particular set of directions then it will recognize one of the directions like REWIND, FORWARD, PLAY and STOP and operate the songs present in the list of music system. Additionally, Karhunen-Loeve (K-L) Transform is used to capture the image without any noise and accurate in result and Canny Edge Detection for image segmentation and edge detection using Principal component analysis (PCA) which add more value in expected result.*

***Key Term:** Hand Gesture Recognition, Karhunen-Loeve (K-L) Transform, Skin Filtering, Canny Edge Detection, Image Segmentation, Human Computer Interaction, matching algorithm; PIC16F8779*

I. INTRODUCTION

Hand gesture recognition is one of the growing fields of research today which provides a natural way of human machine interaction. Gestures are some forms of actions which a person expresses in order to express information to others without saying it. In our daily life, we can observe few hand gestures frequently used for communication purpose like thumbs up, thumbs down, victory, directions etc. Some common examples are in cricket where the umpire uses different hand gestures to show different events that occurred at that instant on the match, hand gestures used by the traffic police, etc. Early approaches to the hand gesture recognition problem in a MP3 control context involved the use of markers on the different action. An associated algorithm is used to detect the presence and colors of the markers, through which one can identify which action, are matched in the gesture.

The inconvenience of placing markers on the user's hand makes this an infeasible approach in practice when the image is not clear. Recent methods use more advanced computer vision techniques and do assumption based result. Hand gesture recognition is performed through a curvature space method, which involves finding the boundary contours of the hand. This is a robust approach that involves scaling, translation and rotation invariant on the hand poses, yet it is computationally demanding. In our approach, we have firstly used Skin filtering where the RGB image is converted to HSV image because this model is more sensitive to changes in lighting condition. And then K-L transform is performed. The advantage of K-L transform is it can eliminate the correlated data, reduces dimensions keeping average square error minimum, and gives excellent cluster character after the transform. Some applications in this field that has already been done, for example hand gesture recognition for sign language, hand gestures used for controlling robot's motion, in video games, etc. The Canny Edge Detection Algorithm which support this approach in five different ways for the accurate result in Smoothing, Finding gradients , Non-maximum suppression , Double threshold and Edge tracking by hysteresis for fitness result. First, the input image is taken from camera and image will be processed in mat lab. In mat lab both the input and database images are compared, after comparing the images, result will be given to PIC16C8779 controller with the help of RS232 cable. The MP3 player is interfaced to the PIC16C8779 controller and relay drives, the relays act as a electric switching button, and when the input first image is recognized with PIC16C8779 controller, the relay will be automatically on and MP3 player is played. The MP3 player plays some audio signals only, and the functions are displayed on to the LCD. The positioners and switches are controlled remotely using a 40-pin Microchip 16F8779 microcontroller at the monitor station that receives commands via RS232 and translates them into hardware control logic shown in figure 2. The microcontroller connections to the modules that it controls in the monitor station are shown in Figure 1. This note describes the conventions adopted for the microcontroller and its connections to the hardware to be used in writing the microcontroller assembly language program. Detailed descriptions of the PIC16C877A.

Manuscript Received on September 2014.

Devikarani Patil, PG Student, Department of Computer Science and Engineering, Acharya Institute of Technology, Bangalore, India.

Varalakshmi B. D, Assoc. Prof., Department of Computer Science and Engineering, Acharya Institute of Technology, Bangalore, India.

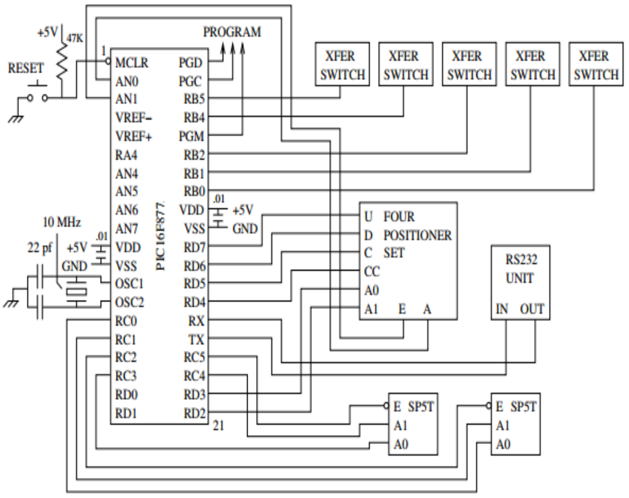
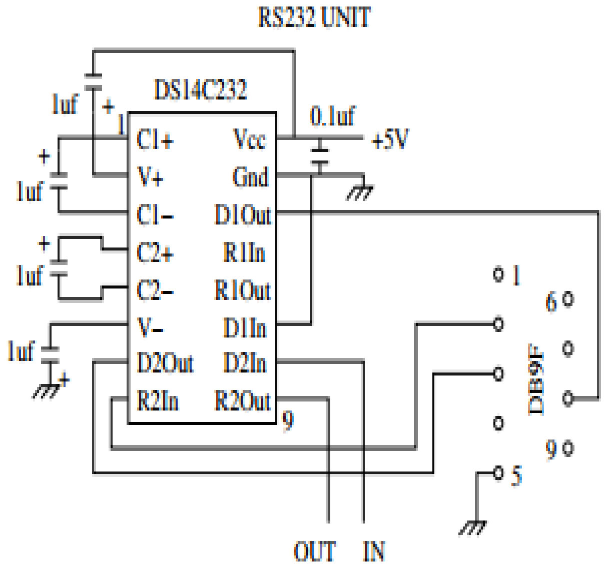


Figure 1 Schematic Diagram of Microcontroller Connections to the Antenna Positions and RF Switches that it Controls



Figures 2 Schematic Diagram of RS232 Interface Circuit

This paper is organized follows: Section II discusses Literature review; Section III describes the Related work, Section IV Design/Proposed implementation, Section V conclusion, VI Reference are as follow,

II. LITERATURE REVIEW

We have studied many previous works done in this field by different researchers. There are many approaches that were followed by different researchers like vision based, data glove based, Artificial Neural Network, Fuzzy Logic, Genetic Algorithm, Hidden Markov Model, Support Vector Machines etc. Some of the previous works are given below. Many researchers used Vision based approaches for identifying hand gestures. Kapuscinski [1] found out the skin colored region from the input image captured and then this image with desired hand region was intensity normalized and histogram was found out for the same. Feature extraction step was performed using Hit-Miss Transform and the gesture was recognized using Hidden Markov Model (HMM). Recognition rate obtained was 98%. Yu [2] used YCbCr colour model to Distinguish skin colored pixels from the background. The required portion of the hand was extracted using this color model and filtered using median filter and smoothing filter. The edges were detected and features extracted were hand perimeter, aspect ratio, hand area after which Artificial Neural Network (ANN) was used as classifier to recognize a gesture. Accuracy rate obtained was 97.4%. In [3][8] fingertip detection was used for hand gesture recognition. Rajam [3] in his paper for sign language recognition first converted the RGB image captured to binary and Canny Edge Detection Technique was used for extracting edge of the palm. The fingertip positions of the fingers were identified from the extracted edge of palm by measuring their distance from a reference point which is taken to be at the bottom of the palm. Recognition rate obtained was 98.125%. Raheja [8] scanned the skin filtered image in all direction to find out the edges of the fingers and the tips of the edges were assigned the highest pixel value and as such fingertip was detected. Malima [4] used hand gesture recognition for controlling the robot. The Red/Green ratio was found out which was used for determining the skin colored regions. The centre of gravity of the hand was found out along with the farthest distance from it and thus in such a way the finger tips were determined. A circle was made around the centre of gravity and number of white pixels beyond that circle was counted to know the desired. For hand-gesture recognition, some researchers have tried to perform the early segmentation process using skin-color histograms Zhou et al. [12] used overlapping sub-windows which is useful to extract invariants for gesture recognition, and distinguish them with a local orientation histogram attribute description indicating the distance from the canonical orientation. This makes the process relatively robust to noise, however, much more time consuming indeed. Kuno and Shirai defined seven different stages of hand gesture recognition. It includes position of the fingertip. This is not practically realistic when we have only pointing gestures, but also several other gestures, like grasping. However, the invariants they considered inspired us for our defined invariants. In some similar approaches, the watermark of an image is generated by modifying the nvariant-vector. For example, Lizhong Gu and Jianbo Su tried to use Zernike moments along [12, 13] with a hierarchical classifier to classify hand-gestures. This method is not appropriate for the JAST project, since there is not a high degree of freedom for the hands due to the limited space for movements and actions.

III. RELATED WORK

A. Canny Edge Detection

This is one of the best edge detection techniques but little complex than other edge detection techniques. The major advantage of this technique is its performance. In case of other edge detection techniques only one threshold is used, in which all values below the threshold were set to 0. Thus, we must be very careful while selecting the threshold. Selecting the threshold too low may result in some false edges which are also known as false positives. Whereas if the threshold selected is too high, some valid edge points might be lost, this is also known as false negatives. But canny edge detection technique uses two thresholds: a lower threshold, TL and a higher threshold, TH thus eliminating problem of false positive and false negative. Steps involved in this type of detection are:

• The input image is smoothened with a Gaussian filter after which the Gradient magnitude and angle images are computed.
• Non-maxima suppression is applied to the gradient magnitude image.
• And finally detection and linking of the edges is done using double thresholding and connectivity analysis.

The algorithm runs in 5 separate steps:
1. Smoothing: Blurring of the image to remove noise.
2. Finding gradients: The edges should be marked where the gradients of the image has large magnitudes.
3. Non-maximum suppression: Only local maxima should be marked as edges.
4. Double thresholding: Potential edges are determined by thresholding.
5. Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

B. Karhunen-Loeve (K-L) Transform

K-L Transform is used to translate and rotate the axes and new coordinate is established according to the variance of the data. The K-L transformation is also known as the principal component transformation, the eigenvector transformation or the Hotelling transformation. The advantages are that it eliminates the correlated data, reduces dimension keeping average square error minimum and gives good cluster characteristics. K-L Transform gives very good energy compression. It establishes a new co-ordinate system whose origin will be at the centre of the object and the axis of the new co-ordinate system will be parallel to the directions of the Eigen vectors. It is often used to remove random noise.

C. PIC16F8779 Controller

This module is connected to the six most significant bits of register port RD. One of the four positioners is controlled by putting its address on lines A0 and A1, and up/down or CW/CCW motion is commanded by pulling one of the U, D, C, or CC lines low. The azimuth and elevation positions are read by 877A A/D converters AN0 and AN1, respectively. Each of the five filter selection transfer switches is controlled with a single bit in port resister RB. Each of the two SP5T RF switches is controlled with three bits in port register RC, where the three bits form an address to one of the five switch positions. In this application, the PIC16F877A is programmed to initialize the default switch settings and turn all positioner motion off. It then enters an infinite loop watching for one of the three interrupts listed in Table 3. The 877A's internal TIMER1 continuously generates interrupts approximately once per 100 milliseconds. Each TIMER1 interrupt initiates a conversion by the analog-to-digital converter. When the conversion is complete the 877A receives an AD interrupt which causes the A/D data register to be read and stored.

IV. DESIGN AND IMPLEMENTATION

The MP3 (music) player is controlled by using gesture. An input image is taken from camera and image will be processed in mat lab. In mat lab both the input and database images are compared, after comparing the images result will be given to PIC16F8779 controller with the help of RS232 cable. The MP3 player is interfaced to the PIC16F8779 controller and relay drives, the relays are acts as a electric switching button, and a input first image is recognized with PIC16F8779 controller, the relay will be automatically on and MP3 player is play. The MP3 player played some audio signals only, and the functions are display to the LCD

A. Karhunen-Loeve (K-L) Transform

After capturing the image from web camera the steps that we have used for recognizing different hand gestures are skin filtering, edge detection, K-L transform and finally a proper classifier, where we have used angle based classification to detect which symbol the test image belongs to. The experiment was performed with bare hands while interacting with the computer. We have used Euclidean distance based classifier in order to compare the result from the angle based classifier and it was found that both give the same result. Thus accurate result was obtained.

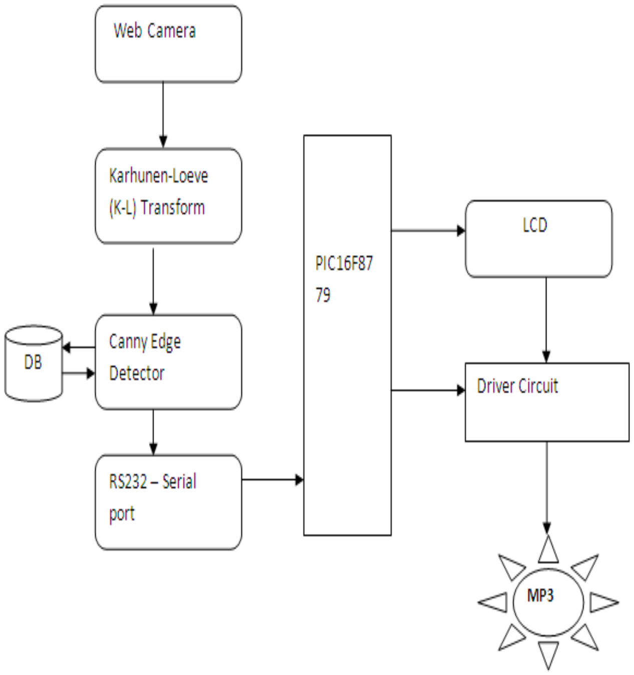


Figure 3 Block Diagram

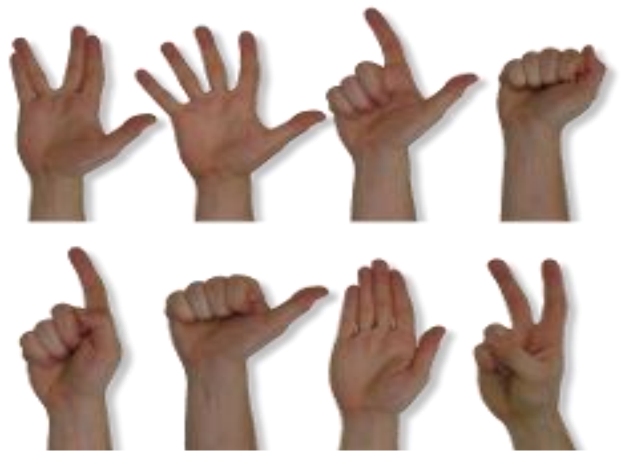


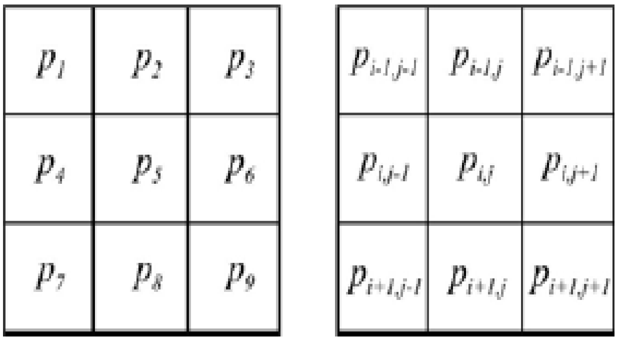
Figure 4 Images are Capture from Web Camera

The first step involves the capturing of image using camera and conversion of the input RGB image to YCbCr color space. This step is done because YCbCr model is more sensitive to changes in lighting condition which is defined through below equation 1,

Y'=16 + (65.481.R + 128.553.G + 24.966.B)
Cb= 128 + (-37.797.R – 74.203.G + 112.0B)
Cr=128 + (112.0R – 93.786.G – 18.214.B)
--(Equation: 1)

B. Canny Edge Detection

As a first step before applying the edge classification techniques for edge detection, the image should be pre-processed. Since images usually contain noise, this should be removed as a part of preprocessing. For the noise removal we have chosen a 3 x 3 Gaussian filter. We consider the case in which we use a 3 x 3 mask, with the central pixel situated at coordinates (i, j). Since the edge detection method is applied to grayscale images, the pixel position is represented as a scalar Pi, j. The representation of such a mask can be seen in the below figure 5:



p_1	p_2	p_3
p_4	p_5	p_6
p_7	p_8	p_9

$p_{i-1,j-1}$	$p_{i-1,j}$	$p_{i-1,j+1}$
$p_{i,j-1}$	$p_{i,j}$	$p_{i,j+1}$
$p_{i+1,j-1}$	$p_{i+1,j}$	$p_{i+1,j+1}$

Figure 5 3* 3 Matrix for Pixel Center Points by Pi, j

An Edge may Appear in Many Directions. In Figure 2 are Presented the Four Cases in which an Edge may Appear

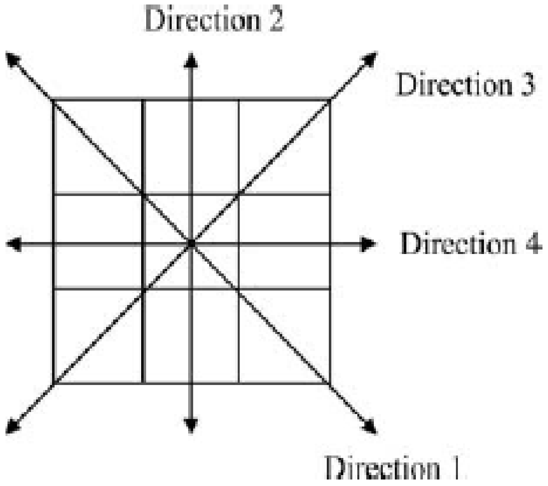


Figure 6 All direction edge in different angles

For the four directions, denoted by d1, d2, d3, d4, it is necessary to calculate the sum of the differences of the bidirectional edges between the P pixel and its neighbors

These will be computed as follows:

$d1 = | p1 – p5 | + |P9 – p5|$
$d2 = | p2 – p5 | + |P8 – p5|$
$d3 = | p3 – p5 | + |P7 – p5|$
$d4 = | p4 – p5 | + |P6– p5|$

For each pixel in the input image that is not at the edge, we must form a vector x = (d1, d2, d3, d4) that contains the four previously calculated distances. The next step is to divide the input image pixels into classes. For this purpose we defined six classes: four classes for edges, a background class and a class for noisy edges. For the four classes that define the edges we have used four typical situations (see Figure 6)

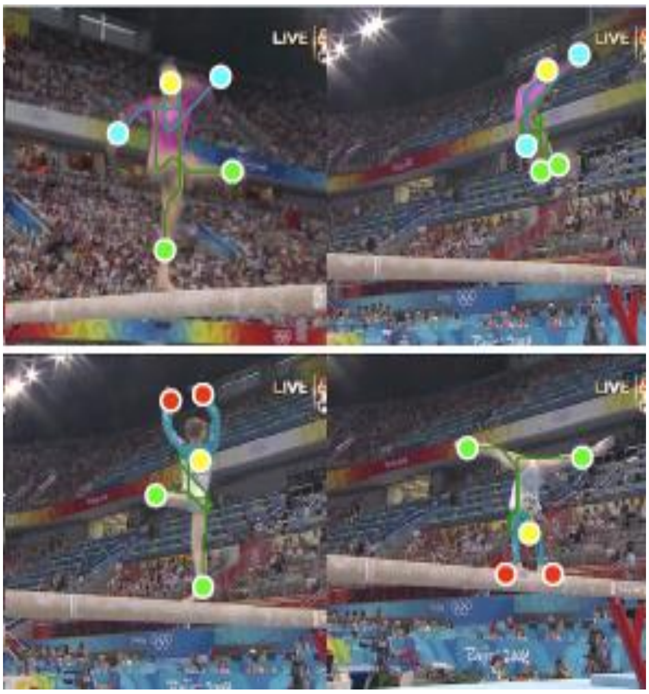


Figure 6 (a) Edges are Moving Objects

C. Serial Port Command to PIC16F8779

Commands to the PIC16F8779 microcontroller are received by its serial port with be ASCII character packets terminated by a carriage return character (hex 0D). For example, to drive positioner 2 to an elevation where the A/D reading is decimal 856, select the positioner with the command P2, and then send the command E0856. The calibration of A/D readings in terms of azimuth and elevation in degrees will be the responsibility of the PC sending the commands. The full range of 10-bit A/D readings will be used since this range corresponds to the positioner read-out of about 0 to 5 volts. If a command outside of the available range is issued, the positioner will run to the limit of its motion in that direction and stop. A new position command will cancel any motion currently in progress from a previous command, if that commanded position had not been reached. The PIC16F8779 microcontroller will continuously send position data for the currently selected positioner to the PC approximately once per half second. Each command is echoed by the microcontroller. If an unknown commend is received, a 'U' message will be sent by the 877A. If two commands are received too quickly for the one to be processed, a missed command, 'M', will be sent by the microcontroller.

D. MP3 Action

A gesture input images are taken from camera and images will be processing in mat lab. In mat lab comparing the both input and data base images and It perform the totally six images, they are play, stop, forward, rewind, volume up & down as per the action shown in figure 7,

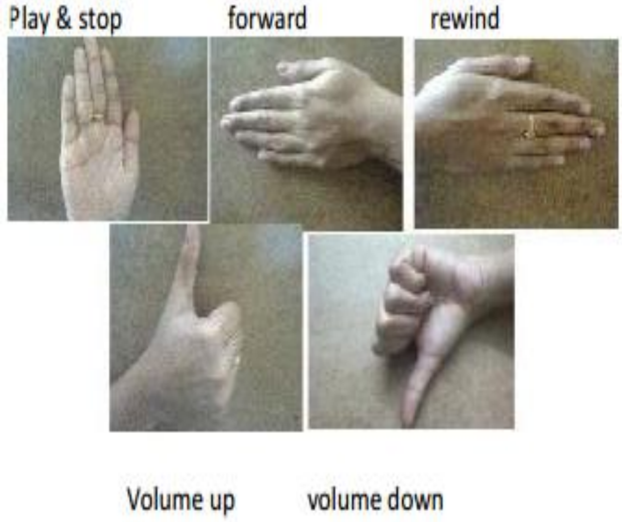


Figure 7 Different Gesture

V. CONCLUSION

In Matlab the signals are taken from the PC (image processing) and using edge detection algorithm, it will be processed by the controller with the help of serial communication (RS232 cable). In embedded kit using some relay drives, we control the mp3 player to play or stop, forward and backward volume up and down functions by gesture. We have performed some simple hand gesture recognition in the study. Future work requires faster computation time to reduce the time delay of the system and advanced recognition methods to recognize more complex gesture.

REFERENCES

1. T. Kapuscinski and M. Wysocki, "Hand Gesture Recognition for Man-Machine interaction", Second Workshop on Robot Motion and Control, October 18-20, 2001, pp. 91-96.
2. C. Yu, X. Wang, H. Huang, J. Shen and K. Wu, "Vision-Based Hand Gesture Recognition Using Combinational Features", IEEE Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2010, pp. 543-546.
3. P.S. Rajam and G. Balakrishnan, "Real Time Indian Sign Language Recognition System to aid Deaf-dumb People", IEEE, 2011, pp. 737- 742.
4. A. Malima, E. Ozgur and M. Cetin, "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control", IEEE, 2006.
5. Manigandan M and I.M Jackin, "Wireless Vision based Mobile Robot control using Hand Gesture Recognition through Perceptual Color Space", IEEE International Conference on Advances in Computer Engineering, 2010, pp. 95-99.
6. D.Y. Huang, W.C. Hu and S.H. Chang, "Vision-based Hand Gesture Recognition Using PCA+Gabor Filters and SVM", IEEE Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, 2009, pp. 1-4.
7. E. Koh, J. Won and C. Bae, "On-premise Skin Color Modeing Method for Vision-based Hand Tracking", The 13th IEEE International Symposium on Consumer Electronics (ISCE), 2009, pp. 908-909.
8. J.L. Raheja, K. Das and A. Chaudhury, "An Efficient Real Time Method of Fingertip Detection", International Conference on Trends in Industrial Measurements and automation (TIMA), 2011, pp. 447-450.
9. J. Rekha, J. Bhattacharya and S. Majumder, "Shape, Texture and Local Movement Hand Gesture Features for Indian Sign Language Recognition", IEEE, 2011, pp. 30-35.
10. R. Gopalan and B. Dariush, "Towards a Vision Based Hand Gesture Interface for Robotic Grasping", The IEEE/RSJ International Conference on Intelligent Robots and Systems, October 11-15, 2009, St. Louis, USA, pp. 1452-1459.
11. Y. Fang, K. Wang, J. Cheng and H. Lu, "A Real-Time Hand Gesture Recognition Method", IEEE ICME, 2007, pp. 995-998.
12. I.-K. Park, l -H. Kim, and K.-S. Hong, "An Implementation of an FPGA-Based Embedded Gesture Recognizer Using a Data Glove," Conference on Ubiquitous Information Management and Communication Proceedings of the 2nd International conference on Ubiquitous information management and communication 2008, Suwon, Korea,January31 – February 01, 2008, pp.496-500
13. Jiayang Liu, Zhen Wang, and Lin Zhong, Jehan ickramasuriya and Venu Vasudevan "Accelerometer-based Personalized Gesture"

AUTHOR PROFILE

Devikarani Patil, Received B.E degree and pursuing M.tech in computer science and engineering, her area of interest are Networks and image processing**.**

Varalakshmi B. D, Received B.E and M.tech degree**,** She is current pursuing Ph.D and working as assistant professor, her area of interest are haptic and image processing.