ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

Shape Parameter-Based Recognition of Hand Gestures

Surabhi S. Gatagat, Devyani D. More, Kalpana D. Varat, Janhavi S. Toshkhani

Student, Dept. of C.S., P.E.S. Modern College Of Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India

Student, Dept. of C.S., P.E.S. Modern College Of Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India

Student, Dept. of C.S., P.E.S. Modern College Of Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India

Student, Dept. of C.S., P.E.S. Modern College Of Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India

ABSTRACT: Hand Gestures play a vital role in our daily life as they are a means of non-verbal communication. They provide us with a more natural and user friendly way of interaction with the computer. In this paper, we keep in mind the similarities of human hand shape with one thumb and four fingers rather than focusing on image based features like the color or texture of the hand. We present a real time system for hand gesture recognition which recognizes the hand gesture based on orientation; features such as centroid, thumb detection, finger region detection, Euclidean distance; and bits generation. To implement this system, we will utilize a computer and a simple webcam with 7 megapixel intensity. On being given an input image sequence, some preprocessing steps are carried out to remove background noise and segment the hand from the background. Only the hand region is processed further to calculate the shape based features. This shape based approach can identify the gesture on the basis of a 5 bit binary string generated as a result of the algorithm.

KEYWORDS: Human Computer Interaction, Image processing, K-means clustering, Features extraction, Bits generation, Pattern recognition, Hand gesture recognition

I. INTRODUCTION

Gestures are motion of the body or physical actions formed by the user to convey some meaningful message. Gesture Recognition is the process by which gesture made by the user is made known to the system. Hand gestures are a substitute for user- machine interaction tools such as mouse, keyboard, joystick and electronic pen. Use of gestures in human-machine interaction provides a more natural interface and it is also a very convenient way of interaction.

In human-human interaction we utilize a broad range of gestures for personal communication. It is also proven that people gesticulate more when they are talking on the telephone and are not able to see each other as in face to face communication. The gestures vary greatly among cultures and context. The significance of gestures in day-to-day life motivates the use of gestural interfaces.

Hand gestures have become a great choice for expressing the simple ideas, which are interpreted by the gesture recognition system and turned into corresponding events. It is the demand of advanced technology to recognize, classify and interpret various simple hand gestures and use them in a wide range of applications through computer vision.

II. RELATED WORK

In the past decades, many different gesture recognition techniques have been developed for tracking and recognizing various hand gestures such as wired technology, optical markers and image colour based techniques. Each one of these techniques has pros and cons.

Wired technology requires users to be tied up with wires in order to connect to the computer system. This limits the movement of the user in the room as the length of the wires is limited. An example of wired technology is data gloves also known as instrumented gloves or electronic gloves. These gloves contain sensorswhich provide information such

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

as hand location, fingertip positions, orientation of the hand, etc. Data gloves are highly accurate and give very good results but they are very expensive to utilize and require a wired connection.

Optical markers are a replacement for data gloves. They project Infrared light on the screen to provide information about hand location or tips of fingers. The optical tracking system provides very good results as well, but a drawback is the need for a very complex configuration.

Image based techniques process the image based on features such as colour or texture. If we base the hand gesture recognition on these features, the results will vary and be inaccurate as skin tones and textures change from person to person and from one continent of the earth to another. Also, under different light conditions, colour and texture gets modified which leads to unexpected changes in the observed results.

Taking the drawbacks of these methods into consideration, we have chosen to use a hand gesture recognition system that will work solely on shape based features and will not require any external hardware for recognition. Under normal conditions, every person possesses the same hand shape with one thumb and four fingers. The approach discussed in paper [2] is highly influenced by some constraints like the hand should be straight i.e. horizontal or vertical for orientation detection, if it is not followed then the results could be unexpected or wrong. In our paper we employ a dual approach for orientation detection to improve the accuracy. In paper [3], the approach is based on calculation of three combined features of hand shape which are compactness, area and radial distance. Compactness is the ratio of squared perimeter to area of the shape. If compactness of two hand shapes are equal then they would be classified as the same gesture, hence this method does not yield accurate results. So, we completely avoid the use of these features for classification.

The algorithm implemented in this paper is divided into four main steps. The first step is image preprocessing and segmentation of hand in the image using k-means clustering to reduce the noise from the captured image. The second step detects the orientation of the hand i.e. horizontal or vertical. In the third step, some shape based features are calculated for the hand pattern detection, which are then used to generate a 5-bit binary sequence for identification of the gesture. Finally, these resultant bits are used for assigning different key press events to various hand gestures. The proposed approach isdesigned and implemented for working on a single hand gesture with uniform background and gives fair results even in different illumination conditions.

III. PROPOSED ALGORITHM

1. *Design Considerations:*

- Windows version 7.0 and above
- Netbeans version 8.0.2
- OpenCV
- 4GB RAM and above
- 5GB HDD and above
- 7 MP Webcam and above

2. *Description of the Proposed Algorithm:*

Aim of the proposed algorithm is to use a hand gesture recognition system that willwork solely on shape based features and will not require any external hardware for recognition.

The architecture diagram of the proposed system is shown in Figure 1.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

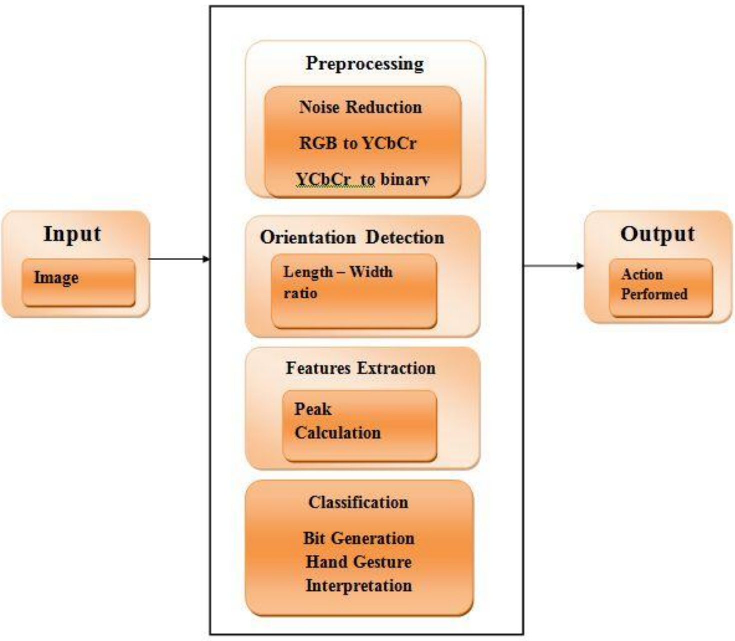


Fig. 1. Architecture Diagram of proposed system

Step 1: Image Preprocessing and Segmentation

Image preprocessing is a necessary step for image enhancement and for getting good results. The RGB images are captured using a 7 MP webcam. The input sequence of RGB images has to first be converted into YCbCr images as RGB color space is more sensitive to different light conditions. YCbCr is a family of color space where Y is a luma component that deals with luminance information of image, Cb and Cr are the blue difference and red difference chromo components that deal with the color information of the image. As luminance creates many problems, it is desirable to process only the chrominance component. Image segmentation is typically performed to separate the hand object in the image from its background. Segmentation based on YCbCr requires a plain and uniform background. Hence, we have to convert the YCbCr images to binary images.

K-means clustering algorithm is used for segmentation. The image is segmented into K clusters such that intra-cluster similarities are maximized and inter-cluster similarities are minimized. Centroid is computed for each cluster such that sum of distances from each object to its cluster centroid is minimized. The K-means algorithm iteratively minimizes the sum of distances from each object to its cluster centroid until the sum cannot be decreased further. The result of K-means clustering is a set of clusters which are well separated from the other clusters. In this case, two clusters are formed i.e. one represents the hand and the other represents the background. Cluster 1 for the hand has all pixel values set to 1 for white and cluster 2 has all pixel values set to 0 for black.

Noise elimination steps are applied to remove insignificant smudges or connected components or objects in the image that have fewer than P pixels, where P is has a variable value. The holes are filled in the binary image after removal of noise. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

After segmentation of the hand, we need to calculate the boundary contours for locating the hand region. This is done by scanning the image from left to right in a top to bottom manner, the first white pixel encountered is set at the left-most point of the hand. Similarly, scanning from right to left, the right-most point of the hand is found. These two scans give us the vertical bounds of the hand. Within these vertical bounds, we perform horizontal scan from top to bottom in a left to right manner to locate the top-most point of the hand. As the hand extends from the bottom-most part of the image, there is no need to perform scanning to locate the bottom-most point of the image. In this way, a bounding box is formed around the hand.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

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

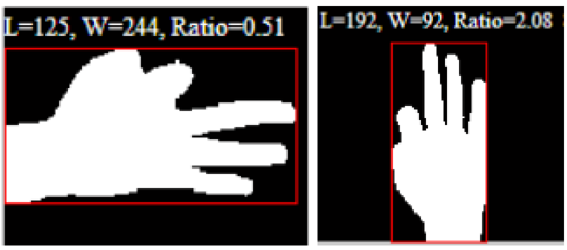
Fig. 2. (a)Input image (b)YCbCr image (c)Binary image (d)Localized hand object

Step 2: Orientation Detection

For orientation detection, we have used a dual approach in order to increase the accuracy. To find out whether the orientation of the hand is horizontal or vertical, we compute the length to width ratio of the bounding box. It is assumed that if the length of the bounding box is greater than the width i.e. if the length to width ratio is greater than 1, the orientation is vertical. Likewise, if the length to width ratio is less than 1, the orientation is horizontal.

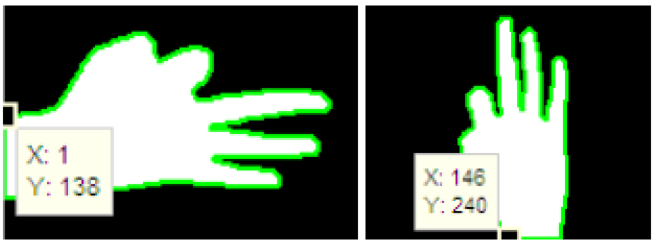
Secondly, we implement the edge detection algorithm for finding the edges of the hand. Edge detection is used to identify the points at which the brightness in the image changes very sharply. To detect the edges of the hand, we scan the image and extract those pixels in the image where the pixel value changes rapidly from 0 to 1 i.e. from black to white. After edge detection, to identify the orientation of the hand, we scan the boundary matrices or edges of the hand in the binary image. Whenever we get the x-boundary equal to 1 along with increasing value of y-boundary for some time span, we classify it as a horizontal hand and if we get the y-boundary equal to maximum size of the image along with increasing value of x-boundary, it is classified as a vertical hand.

To reduce uncertainties in the orientation classification, we have used this dual approach. Hence, these two methods should yield the same result for proper orientation detection. At this stage, we have categorized the hand into two categories i.e. vertical and horizontal.



(a) (b)

Fig. 3. Orientation detection by finding length to width ratio of bounding box. (a)Horizontal Orientation (b)Vertical Orientation



(a) (b)

Fig. 4. Orientation detection by edge detection algorithm. (a)Horizontal Orientation (b)Vertical Orientation

Step 3: Features Extraction

1) Centroid:

Centroid of the hand is calculated to partition the hand into two halves, one part having fingers and the other having no fingers. Centroid is also called as centre of mass if the image is uniformly distributed and it divides the hand into two halves at its geometric centre. Centroid is calculated using image moment, which is the weighted average of the pixels' intensities of the image.

Image moment, M_{ij} is calculated using the formula:

$$M_{ij} = \sum_x \sum_y x^i y^i I(x,y) \qquad \text{eq. (1)}$$

,where I(x,y) is intensity at coordinate (x,y).

Coordinates of centroid are given by:

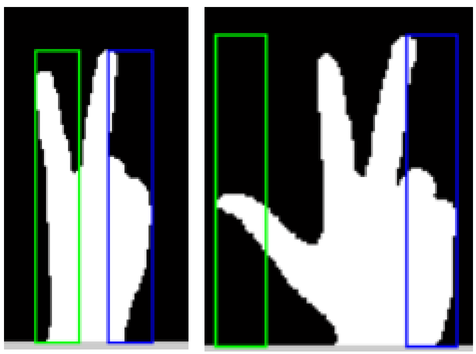
$$\{\bar{x}\ \bar{y}\} = \{M_{10}/M_{00}, M_{01}/M_{00}\} \qquad \text{eq. (2)}$$

,where M_{00} is area for binary image.

2) Thumb Detection:

Thumb detection is performed to detect the presence or absence of the thumb in the hand gesture. The thumb is considered as an important shape feature for classification of hand gestures. To detect the presence of the thumb we consider 30 pixels width from each side of the previously calculated bounding box and crop the bounding box into two boxes i.e. left box and right box. We then count the number of white pixels present in each box of the binary image.

If there is less than 7 percent of total white pixels present in any of the boxes, then the thumb is assumed to be present in that box. If both boxes have either less than or more than 7 percent of total white pixels in the image, then the thumb is not present in any of the boxes. The percentage of white pixels set as 7 percent is chosen experimentally to get more accurate results. The results will be highly influenced by variation in orientation. For proper detection of the thumb, the hand should either be straight horizontal or straight vertical.



(a) (b)

Fig. 5. (a)No thumb detected (b)Thumb is detected

3) Finger Region Detection:

The finger region is obtained by computing the centroid. In this step, we determine the number of fingers raised and the number of fingers folded in the hand gesture. Peaks are used to represent tips of fingers. Vertical and horizontal images are processed in a different manner for finger detection. For the vertical hand, only y-coordinates of the boundary matrices are considered. When the y-coordinate values start increasing after sharp decrement in y-coordinate values, it is considered a peak value or fingertip. For the horizontal hand, only x-coordinates of the boundary matrices are traced. When the x-coordinate values start decreasing after continuous increment in x-coordinate values, it is considered a peak value.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

4) Euclidean Distance:

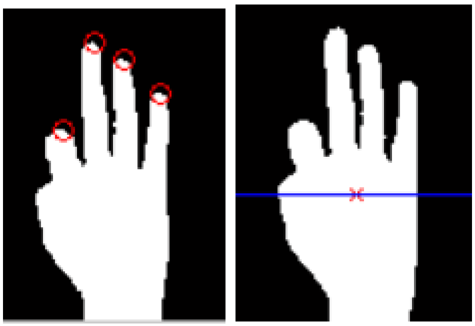
After marking the detected peaks, we must find out the highest peak in the image for further classification of significant and insignificant peaks to find out how many fingers are raised and how many are folded. This calculation is done using Euclidean distance.

The distance between the finger tips i.e. detected peaks and centroid is calculated using the formula:

$$E.D.(a,b)=\sqrt{(x_a-x_b)^2-(y_a-y_b)^2} \quad \text{eq. (3)}$$

,where (x_a,x_b) represents detected peaks and (y_a,y_b) represents centroid which is taken as the reference point.

There may be some peaks detected which do not actually represent the tip of the raised fingers, but the tip of folded fingers. These are considered as insignificant peaks. We get rid of these insignificant peaks by computing the maximum peak by the Euclidean distance formula. The threshold value is set to 75% of the maximum peak distance and peaks whose distance values are more than this threshold are classified as significant peaks and they represent raised fingers in the gesture. Peaks that fall below the threshold are classified as insignificant peaks or folded fingers.



(a) (b)

Fig. 5. (a)No thumb detected (b)Thumb is detected

Step 4: Classification and Bits Generation

The hand gestures are classified based on the features calculated in the previous section. A 5 bit binary sequence is generated in which each bit position corresponds to a finger. Significant peaks or raised fingers are encoded as 1 and insignificant peaks or folded fingers are encoded as 0 in the binary sequence. The leftmost bit in the 5 bit binary sequence is reserved for status of the thumb in hand image. If the thumb is present, leftmost bit will be set to 1, otherwise set to 0. From this binary sequence, a Peak-Centroid plot is generated in which significant and insignificant peaks are shown with respect to the threshold line. Hence, this 5 bit binary sequence uniquely recognizes the hand gestures and utilizes them for supporting Human-Computer Interaction.

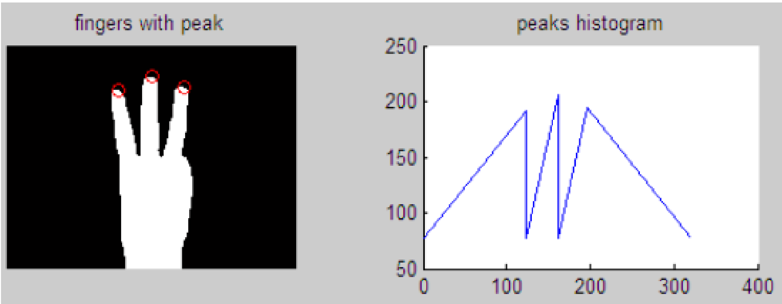


Fig. 6. Peak- Centroid Plot for sample gesture

In the example shown in figure 7, [01110] will be categorized as three raised fingertips with no thumb.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

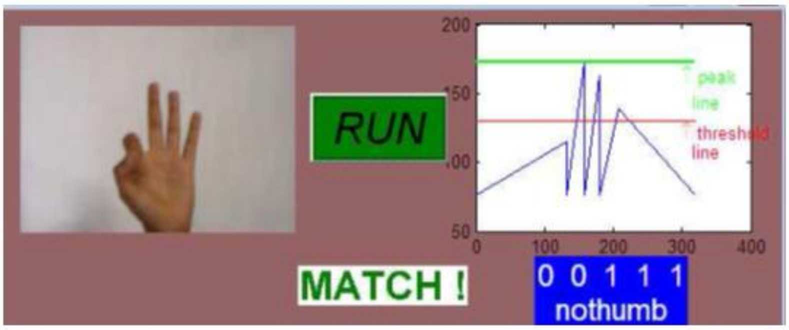


Fig. 7. Sample hand gesture with bits code and Peak- Centroid Plot

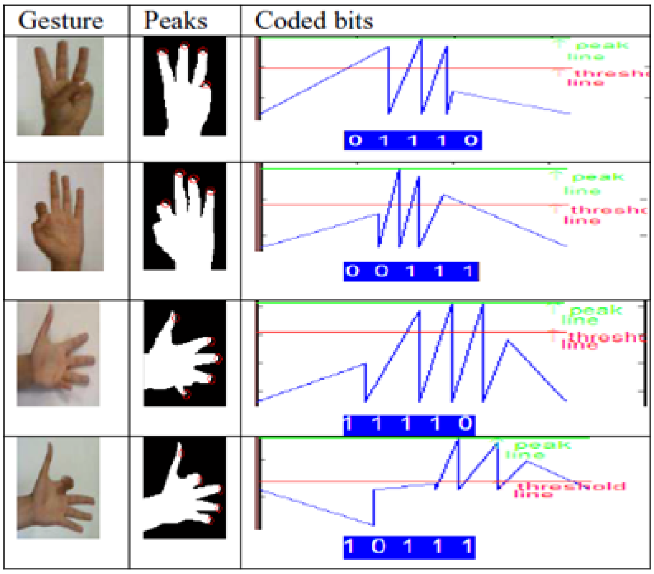


Fig. 8. Sample gestures with peaks and corresponding Peak-Centroid Plots

Step 5: Hand Gesture Interpretation

For providing Human Computer Interaction through hand gesture recognition, some key press events are generated. The gestures interpreted from the Peak-Centroid plots get converted to corresponding key press events. Hence, gestures are used to perform specific actions without the use of user-machine interaction tools such as mouse, keyboard, joystick and electronic pen.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

Fig. 9. Sample Hand Gestures

In the above figure, sample hand gestures are given with their corresponding meanings to generate key press events. For example, gesture with index finger raised means ‘a’, so it will press key ‘a’ on the keyboard. Similarly, gestures for various keyboard keys such as Enter, Backspace, Space, Tab, Delete and Ctrl are given.

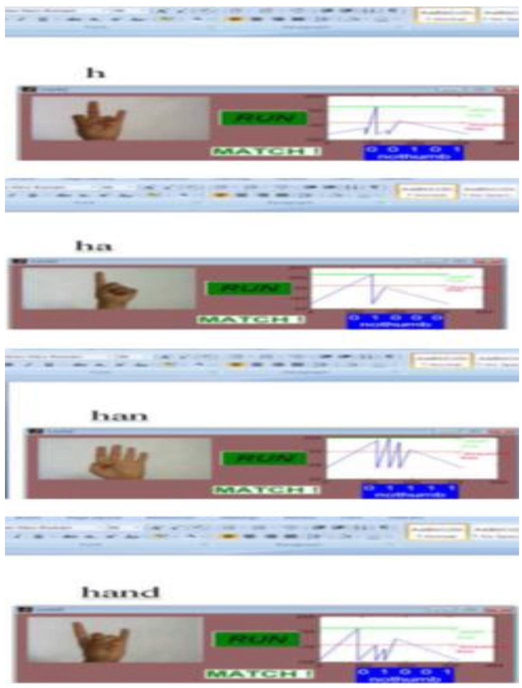


Fig. 10. Comparison table for gestures and their corresponding key press events in Microsoft Word

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

The above figure shows gestures with their Peak-Centroid plots and how the gestures are matched to generate corresponding key press events. The word 'hand' is typed in Microsoft Word by recognition of the hand gestures.

IV. PSEUDO CODE

Step 1: Start
Step 2: Capture image using webcam.
Step 3: Convert the captured RGB image into YCbCr format and then into Binary format.
Step 4: Perform segmentation using K-Means Clustering.
Step 5: Once hand region is separated from background, form bounding box around the hand.
Step 6: Detect orientation of image by calculating length to width ratio of bounding box and using edge detection algorithm.

```
If(length<width)
        Horizontal Hand
Else
        Vertical Hand
```

Step 7: Calculate centroid using image moment.
Step 8: Divide bounding box to form two bounding boxes, one on the left side and the other on the right side. Perform thumb detection by calculating the number of white pixels in each bounding box.

```
If(white pixels<7% of total white pixels)
        Thumb is present
Else
        Thumb is absent
```

Step 9: Detect tips of fingers or peaks.
Step 10: Classify peaks as significant or insignificant to determine number of raised or folded fingers in the gesture. This is done by Euclidean Distance calculation.

```
If(distance< 75% of maximum peak distance)
        Insignificant peak
Else
        Significant peak
```

Step 11: Generate 5 bit binary sequence, where 0 represents folded finger/insignificant peak and 1 represents raised finger/significant peak.
Step 12: Interpret hand gesture and perform key press event.
Step 13: End.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a shape based approach for hand gesture recognition with several steps including smudges elimination, segmentation, orientation detection, thumb detection, finger region detection, bits generation etc. This gesture recognition system is very useful for visually impaired people as they can make use of hand gestures for writing text in an electronic document like MS Office, Notepad etc. Each interpreted gesture is assigned with corresponding English alphabet, number and some very useful key press events for generating text. This system is very simple and easy to implement as there is no complex feature calculation, no significant amount of training or post-processing required. This system provides us with high gesture recognition rate with minimal computation time.

Most conventional algorithms use color based segmentation techniques, but these techniques have a main limitation that different illumination conditions lead to change in colors, causing errors and failures. For example, due to insufficient light, non-skin regions can be mistaken for the hand area because of same color. Therefore, gesture recognition based on shape parameters is a very accurate and less error-prone method. Good results are obtained even in different light conditions. The weakness of this method is that we do not follow a systematic approach for defining certain parameters and threshold values. These values are chosen experimentally. But, to make the system more robust, some constraints from previous methods have been eliminated, making the method simpler and independent of user

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 10, October 2015

characteristics. Future work will emphasize on extending the proposed approach to apply to images including two hands so that more key press events can be triggered with more number of gestures.

REFERENCES

1. Meenakshi Panwar, "Hand Gesture Recognition based on Shape Parameters", in Proceedings of IEEE International Conference on Computing, Communication and Applications(ICIIP 2011), Dindigul, Tamilnadu, February 2012.
2. Meenakshi Panwar and Pawan Singh Mehra ,"Hand GestureRecognition for Human Computer Interaction", in Proceedingsof IEEE International Conference on Image InformationProcessing(ICIIP 2011), Waknaghat, India, November 2011.
3. Amornched Jinda-apiraksa,Warong Pongstiensak, and ToshiakiKondo, "A Simple Shape-Based Approach to Hand GestureRecognition", in Proceedings of IEEE InternationalConference on Electrical Engineering/Electronics ComputerTelecommunications and Information Technology (ECTICON),Pathumthani, Thailand , pages 851-855, May 2010.
4. A. Jinda-Apiraksa, W. Pongstiensak, and T. Kondo, "Shape-Based Finger Pattern Recognition using Compactness and RadialDistance," The 3rd International Conference on EmbeddedSystems and Intelligent Technology(ICESIT 2010), ChiangMai, Thailand, February 2010.
5. Rajeshree Rokade , Dharmpal Doye, Manesh Kokare, "HandGesture Recognition by Thinning Method", in Proceedings ofIEEE International Conference on Digital Image Processing(ICDIP), Nanded India, pages 284 287, March 2009.
6. G. R. S. Murthy, R. S. Jadon (2009)," A Review of VisionBased Hand Gestures Recognition", International Journal ofInformation Technology and Knowledge Management, vol.2(2), pp. 405-410.
7. P. Garg, N. Aggarwal and S. Sofat (2009), "Vision BasedHand Gesture Recognition", World Academy of Science, Engineeringand Technology, Vol. 49, pp. 972-977.
8. Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh,Mo Nours Arab, (2008) ."Human-Computer Interaction:Overview on State of the Art", International Journal on SmartSensing and Intelligent Systems, Vol. 1(1).
9. https://en.wikipedia.org/
10. Mokhtar M. Hasan, Pramoud K. Misra, (2011)." BrightnessFactor Matching For Gesture Recognition System UsingScaled Normalization", International Journal of ComputerScience Information Technology (IJCSIT), Vol. 3(2).
11. Xingyan Li. (2003). "Gesture Recognition Based on Fuzzy CMeansClustering Algorithm",Department of Computer Science.The University of Tennessee Knoxville.
12. S. Mitra, and T. Acharya. (2007)."Gesture Recognition: ASurvey, IEEE Transactions on systems", Man and Cybernetics,Part C: Applications and reviews, vol. 37 (3), pp. 311-324, doi: 10.1109/TSMCC.2007.893280.

BIOGRAPHY

Surabhi S. Gatagat, Devyani D. More, Kalpana D. Varat and Janhavi S. Toshkhani are currently pursuing B.E. (Computer Engineering) in P.E.S. Modern College Of Engineering, Savitribai Phule Pune University, Maharashtra, India. Their research interests are Artificial Intelligence, Pervasive Computing and Data Mining.