Hand Gesture based Interface for Aiding Visually Impaired

A.Deepika
Department of Electronics and Communication Engineering
K.G.Reddy College of Engineering and Technology

Anusha Kurapati
Department of Electronics and Communication Engineering
K.G.Reddy College of Engineering and Technology

Abstract-Hand gesture recognition is a growing and very vast field of research. Numerous work have been done and a lot of work still remains to be done for providing a intuitive, innovative and natural way of non verbal communication, which is more familiar to human beings. Gesture Recognition is widely used in sign language, alternative computer interfaces, Immersive game technology etc. The aim of this project is to present a system for hand gesture recognition to provide a interface for aiding visually impaired users on the basis of detection of some useful shape based features like orientation, area, centroid, extreme, location, presence of fingers and thumb in image.. This unique approach can recognize around 36 different gestures on the bases of 7 bit binary sequence or string generated as a output of this algorithm. The proposed implemented approach has been tested on 360 images, and it gives approximate recognition rate of 94%.One of the great benefits of this algorithm is that it takes only fractional part of a second to recognize the hand gesture which makes it computationally efficient as compare the other existing approach. The proposed algorithm is simple and independent of user characteristics. And also it does not require any kind of training of data like in HMM or neural network.

Keywords—Image processing, hand gesture recognition

I. INTRODUCTION

Gesture and Gesture recognitions system provide us with the large scope of innovations. Gestures are basically the physical action form by the person in order to convey some meaningful information. Gesture recognition system is thus created to provide these gestures a unique tag of interpretation after recognition and classification to form a intuitive and more convenient way of interaction. There is a great emphasis on using hand gesture as a latest input modality in various applications of computers through the use of computer vision. The hand gesture which represents ideas and actions using different hand shapes, orientation or finger patterns being interpreted by gesture recognition system and generate corresponding event, has the potential to provide a unique interface to the computer system. This type of interaction is the heart of immersive virtual environments. With the development of virtual environment, current user interaction approaches with the use of mouse, keyboard and pen are not sufficient. So gesture recognition has now become a very popular research direction in the field of human computer interaction, sign language and computer vision. If for a second we consider interaction among human beings and remove ourselves from the world of computers, we can quickly realize that we are utilizing a broad range of gestures in our daily life. The gestures vary greatly among cultures and context, but still are intimately used in communication. In fact, it is shown that people even gesticulate as much when they talk on telephone, and unable to see each other as in face to face interaction and communication. This significant use of gestures as a mode of interaction in our daily life motivates the use of gestural interface in this modern era.

II. LITERATURE SURVEY

Gesture recognition becomes an influencing term in some past decades. There have been many gesture recognition techniques developed for tracking and recognizing various hand gestures. Each one of them has its advantages and drawbacks as well. First is wired technology in which users need to tie up themselves in order to connect or interface with the computer system. In wired technology, the user can not freely move here and there in the room as they are limited by the length of wires to cover the distance which connect with the computer system via wire. One of the instances of wired technology is instrumented gloves also called as electronics gloves or data gloves. An instrumented glove contains some sensors which provide the information related to hand location, orientation etc. These data gloves provide results with high accuracy but they are very expensive to utilize in broad range of application. Data gloves are then replaced by optical markers. These markers project Infra-Red light and reflect this light on screen to provide the information about the exact location of hand or tips/knuckles of fingers wherever the markers are wear on hand. These systems also give good result but require very complex configuration. Then some advanced techniques introduced like Image based techniques which require processing of image features like color, texture etc. If we work with color texture features of the image for hand gesture recognition the result may vary and would be different as skin tones changes from person to person and from one continent to other. And also under different illumination condition, color texture gets modified and leading to changes in observed results. So for adopting another alternative for the same purpose, we reach to employing different shape based features for hand gesture recognition. This is a universal truth that under normal condition every person poses almost the same hand shape with one thumb and four fingers.

The approach discussed in paper[1] for hand gesture recognition based on shape features heavily depends on some constraints like location of hand in image for orientation detection, proper gap between each finger, etc. if any user fails to follow these constraints then results may degrade. In paper[2], 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 same, in this way this approach limits the number of gesture pattern that can be classified using these three shape based descriptors and only 10 different patterns have been recognized. The algorithm discussed and implemented in this project is broadly divided into four steps. First is image pre-processing and segmentation of hand object in the image using k-means clustering. The second step includes orientation detection, which performed in order to categorize images into vertical and horizontal class. In the third step it calculates some essential features required for hand pattern detection and for generating the 7 bit sequence for 36 different hand shapes. Finally, these generated bits are used for assigning different actions to various hand gestures. This proposed approach is designed and implemented for working on single hand gesture with uniform background.

III PROPOSED SYSTEM

The flowchart of the algorithm is shown in Fig. 3.1 and its main steps are discussed in the following sections

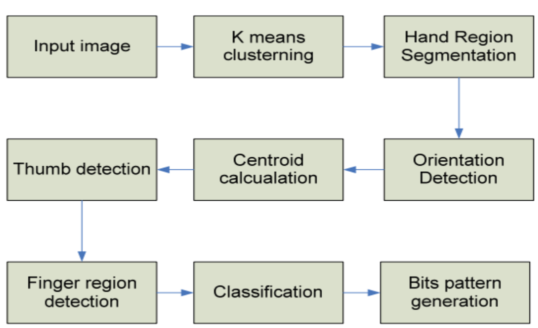


Figure 3.1 Proposed System

3.1 Image Segmentation

Image Pre-processing is necessary for getting good results. In this algorithm, we take the RGB image as input image. Image segmentation is typically performed to locate the hand object and boundaries in image.

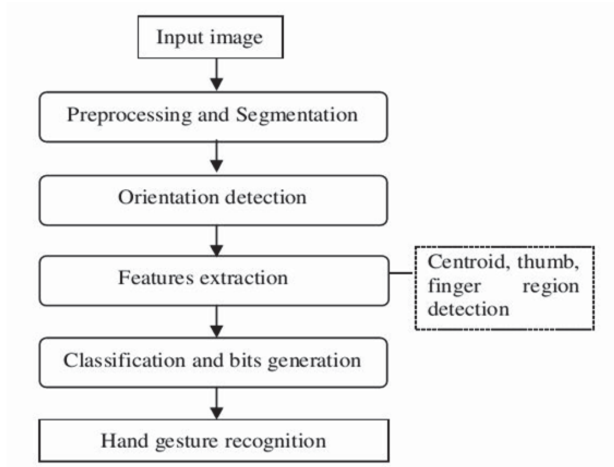


FIGURE 3.2 FLOWCHART OF THE IMPLEMENTED ALGORITHM

It assigns label to every pixel in image such that the pixels which share certain visual characteristics will have the same label. For the purpose of hand object segmentation from rest of the image we employ K-means algorithm.

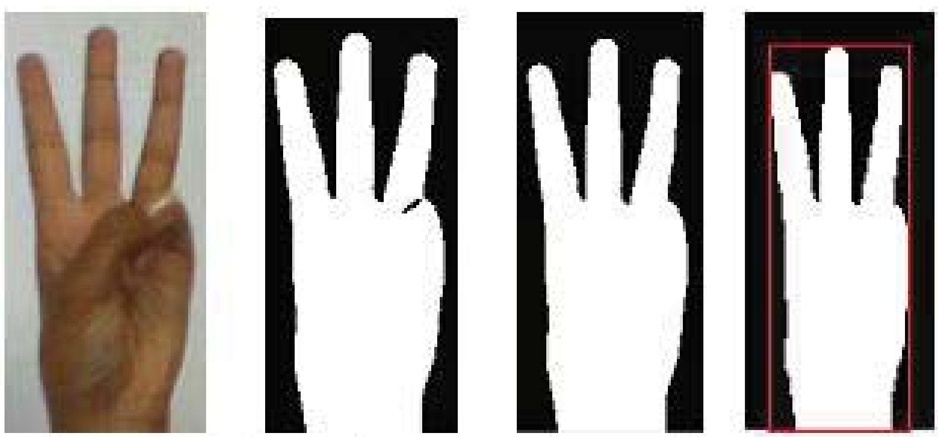


Figure 3.3 Input image, Cluster image, Enhanced image and localized hand Object

3.2 Orientation Detection

After the segmentation of hand in the image, we proceed to the second step for orientation detection. We proces two types of images one is horizontal another is vertical. So in this step, mainly we identify whether the hand is vertical or horizontal. We compute length and width of bounding box with an assumption that if the hand is vertical then length of the bounding box is greater than the width of bounding box and if width of bounding box is greater than the length of bounding box then the image contains horizontal hand. We compute the ratio of length to width of the bounding box if it is greater than 0.9 then it is vertical otherwise horizontal. In this way we categorise the two categories of hand patterns, horizontal and vertical

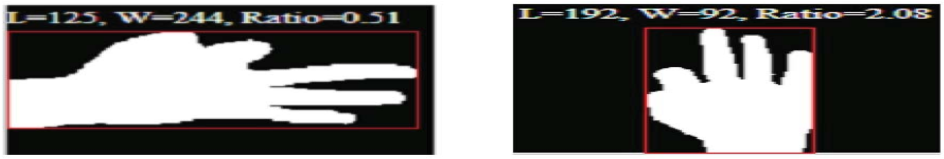


Figure 3.4 Horizontal image and Vertical image

3.3 Features extraction

3.3.1 Centroid

In this step, we calculate the centroid for differentiating various hand pattern based on index finger and little finger. By assuming Index finger and little finger position as distinct from location of the centroid in the hand image we can differentiate many hand gestures. The partition of the hand is very important. Hand gesture patterns which contain the index finger will fall in the left hand side of the centroid. And other hand patterns which contain little finger will find in the right side of the centroid location in image. Hand patterns which contain both Index and little finger need to satisfy both the conditions of centroid location for generating unique 7 bit sequence for recognition. Centroid always computed at the geometric center of the image and it is also called as center of mass if the image is uniformly distributed. Centroid is calculated using the image moment, which is the weighted average of pixel's intensities of the image. The centroid is calculated by first calculating the image moment.

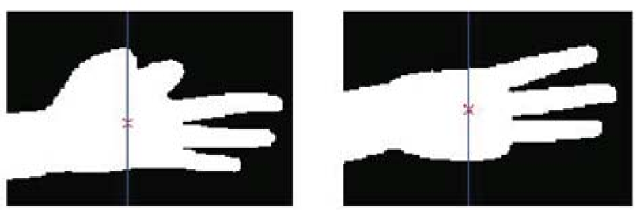


Figure 3.5 Centroid of image

3.3.2 Thumb detection

Thumb detection step is performed to detect the presence or absence of thumb in hand gesture recognition. We know that the thumb can either reside at left side of the hand or at right side of the hand in general. So we consider thumb as a significant shape feature to classify various hand gestures. To solve this problem, we process the previously calculated bounding box and divided this box in left side and right side. By taking 30 pixels as a width from each side of the bounding box we crop this bounding box in two region. So one is the left box represented by green boundary and the other is right box represented by blue boundaries in the image shown below. After getting these two boxes, we count the total number of white pixels presents in binary image which also represents the hand object in image. We count the number of white pixels in each boxes. If less than 9% of total white pixels exist in any of the left box or right box, we consider that the thumb is present in that box. If both boxes have less than 9% of total number of white pixels exist in image then thumb is not present in any of the box because thumb is only one and it can not be found at both sides for the same hand shape pattern. And if both boxes having more than 9 percent of total white pixels in the image, then thumb is not present in any of the box. The percentage we take as 9 is chosen experimentally after testing more than 400 images.

Figure 3.6 Thumb detection

3.3.3 Finger region detection

For getting the total number of finger raised in hand pattern, we need to process only some upper portion of the hand, which includes all raised fingers. To proceed this task we consider only the 25 % of the upper portion of bounding box for one set of gesture and 18 % for another set of the hand gesture. We compute extreme points which fall within this area of bounding box to know the finger count. Extrema build 8-by-2 matrix that specifies the extrema points in the selected region. Each row of the Extrema matrix contains x- and y-coordinates of one of the extreme point. The format of this vector is [top-left top-right right-top right-bottom bottom-right bottom-left left-bottom left-top]. Extrema found in this area represents the knuckle of fingers. It can be possible that some insignificant extrema found in this region which do not represent raised finger but shows the part of folded finger exist in the image. So to remove these unwanted extrema, we put a threshold on the area that should be ignored if it resides in the target region that need to be processed for finger count. This threshold may vary from 10px to 30px depending on the type of gesture we are taking. After getting the number of fingers or extrema present in the selected region, we need to put certain condition on these extreme points to know the actual pattern of hand gesture. Two or more patterns having the same number of extrema or finger count is differentiated by using location of centroid from coordinates of extrema. In this way the resulted 7-bit sequence becomes different for various hand gesture with same number of fingers. **Fig. 4.7** shows the target region which is processed in order to calculate the total number of fingers present in hand gesture. As shown in **Fig.4.7** below, there is three extrema found in this hand pattern which shows three fingers raised we use only three extreme points out of eight extreme points, mentioned above, for generating bit sequence.

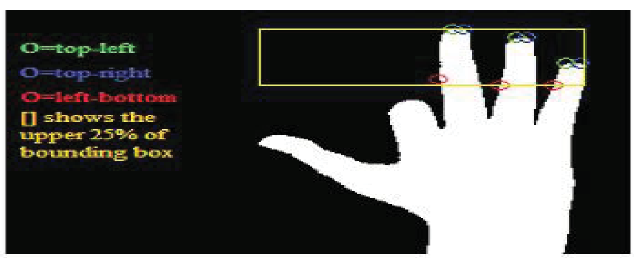


Figure 3.7 Extreme points and target region

3.5Classification and bits generation:

Extreme point and centroid position with respect to each other needed to be processed in order to generate unique 7 bit binary sequence for 36 different patterns. First bit of 7 bit sequence represent the type of orientation, if the hand is vertical first bit will be 1 otherwise 0. Second bit and third bit is allotted for thumb presence, if thumb is not present in the hand gesture then it is set as 00. If left thumb is present then it is set as 10. If right thumb is present then it is set as 11. Next three bits are assigned for total number of raised fingers in hand such that if 1 finger is raised then it will be represented by its 3 bit binary representation and coded as '001', if two fingers are raised then it will be coded as '010' which is the binary representation of '2'. As we know a hand gesture can have maximum four fingers raised in image so these three bits can have the maximum bit pattern as '100'. Last bit of 7 bit sequence is very important because they differentiate among the hand patterns which have equal number of fingers. This bit is set as 1 or 0 depending up on the density of the hand. If the density of the hand is towards left then the bit is set as 1 and if the density of the hand is towards right then the bit is set as 0. As shown in Fig. 4.8, these three patterns have got different bit sequence just because of last bit in 7 bit sequence. Similarly, we generate bit code for several hand gesture patterns with same number of fingers but different shapes. These seven bits code thus generated is used for classifying and assigning various actions for supporting human computer interaction.

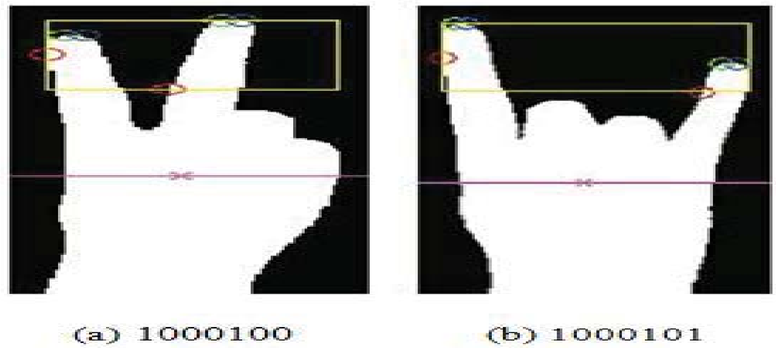
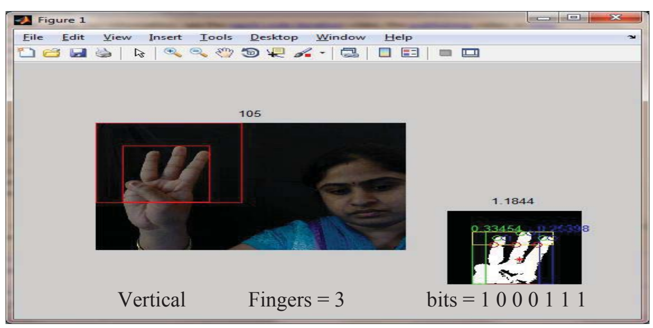


Figure 3.8 Hand gesture with bits code

IV. RESULTS AND DISCUSSION

We have applied the above discussed algorithm and tested 360 images with 36 different patterns. By using effective shape based features and orientation, we can recognize and classify 36 different hand gesture patterns. On the basis of generated bit sequences we can assign different tasks to support human computer interaction or sign language. The below results shows the input gestures along with the feature extractions like centroid, finger region, finger count and corresponding bits generated. These generated bit sequence would always be unique based on their orientation, presence of thumb and pattern of hand gesture. However, after the orientation of hand gesture gets detected, the algorithm rotate the image to 90 degree if it is found to be horizontal image. All further processing steps for generating bits and recognition of hand gesture will be applied to same orientation of hand gesture that is vertical. It is done to minimize the time of processing. Table 6.1 shows the result of 360 images. Out of 360 images tested through the algorithm, it has correctly recognized 339 images and falsely identified the remaining 21 cases. At an average it gives the success rate of 94% approximately with average computation time of 0.60 second. The algorithm is based on simple shape based feature calculation which provides us with the comfort of implementation. The algorithm discussed above is implemented on MATLAB.



V. CONCLUSION AND SCOPE FOR FUTURE WORK

5.1Conclusion

In this paper, a simple yet powerful shape based approach for hand gesture recognition. Visually impaired people can make use of hand gestures for writing text on electronic documents like MS Office, notepad etc. Moreover, almost all deaf and dumb people communicate with each other by forming several hand shapes. Similarly, a visually impaired person would be able to work on computer through computer vision. The strength of this approach lies in the ease of implementation, as it does not require any significant amount of training or post processing and it provides us with the higher recognition rate with minimum computation time. The weakness of this method is we define certain parameters and threshold values experimentally since it does not follow any systematic approach for gesture recognition, and maximum parameters taken in this approach are based on the assumption made after testing a number of images. If we compare our approach with the approach described in Hand Gesture Recognition for Human Computer Interaction, the success rate has improved from 92.3% to 94%, the computation time decreased up to fraction of seconds.

5.2 Scope for Future Work

In future the focus would be on improving the system by including some different backgrounds while enlarging the data set. It should be considered that segmentation problem in normal conditions is itself an open research issue. Trustworthy execution of hand gesture recognition approach demands dealing with occlusions, temporal tracking and 3D modeling of the hand, for identifying variety of gestures, which are still mostly beyond the current state of art. Words are also an interesting open problem.

REFERENCES

[1] Meenakshi Panwar and Pawan Singh Mehra , “Hand Gesture Recognition for Human Computer Interaction”, in Proceedings of IEEE International Conference on Image Information Processing(ICIIP 2011), Waknaghat, India, November 2011.

[2] Amornched Jinda-apiraksa, Warong Pongstiensak, and Toshiaki Kondo, ”A Simple Shape-Based Approach to Hand Gesture Recognition”, in Proceedings of IEEE International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology (ECTI-CON), Pathumthani, Thailand , pages 851-855, May 2010

[3] A. Jinda-Apiraksa, W. Pongstiensak, and T. Kondo, “Shape-Based Finger Pattern Recognition using Compactness and Radial Distance,” The 3Rd International Conference on Embedded Systems and Intelligent Technology(ICESIT 2010), Chiang Mai, Thailand, February 2010.

[4] Rajeshree Rokade , Dharmpal Doye, Manesh Kokare, “Hand Gesture Recognition by Thinning Method”, in Proceedings of IEEE International Conference on Digital Image Processing (ICDIP),Nanded India, pages 284 – 287, March 2009.

[5] E. Stergiopoulou and N. Papamarkos, "A new technique for hand gesture recognition", Proceedings of IEEE International conference on Image Processing, Atlanta, pp. 2657-2660, Oct. 8-11, 2006

[6] F. S. Chen, C.M. Fu and C.L. Huang, "Hand gesture recognition using a real-time tracking method and hidden Markov models", Image Vision Computer, vol. 21(8), pp. 745-758, 2003.

[7] P. Peer, J. Kovac, J. and F. Solina, ”Human skin colour clustering for face detection”, In: submitted to EUROCON – International Conference on Computer as a Tool , 2003

[8] Hebert Luchetti Ribeiro, Adilson Gonzaga, "Hand Image Segmentation in Video Sequence by GMM: a comparative analysis", Symposium on Computer Graphics and Image Processing (SIBGRAPI'06), 2006

[9] Changick Kim and Jenq-Neng Hwang, "Fast and automatic video object segmentation and tracking for content-based application", IEEE Transactions on Circuits and Systems for Video Technology, Vol.12(2), Feb. 2002 pp.:122 - 129.