Hand Gesture Recognition for Human Computer Interaction

Meenakshi Panwar, Pawan Singh Mehra

Centre for Development of Advanced Computing
Noida, Uttar Pradesh , India
cimpi.panwar@gmail.com, pawansinghmehra2611@gmail.com

Abstract—**A hand gesture recognition system provide a natural, innovative and modern way of non verbal communication. It has a wide area of application in human computer interaction and sign language. The intention of this paper is to discuss a novel approach of hand gesture recognition based on detection of some shape based features. The setup consist of a single camera to capture the gesture formed by the user and take this hand image as an input to the proposed algorithm. The overall algorithm divided into four main steps, which includes segmentation, orientation detection, feature extraction and classification. The proposed algorithm is independent of user characteristics. It does not require any kind of training of sample data. The proposed Implemented algorithm has been tested on 390 images, gives a recognition rate of approximately 92% and average elapsed time of 2.76 sec. It takes a less computation time as compare to other approaches.**

Keywords—**computer vision, hand gesture recognition, features detection, human computer interaction, Euclidean distance.**

I. INTRODUCTION

A. Motivation

Gestures and Gesture Recognition terms are increasingly encountered in human computer interaction, sign language and computer vision. Gestures are the physical actions that convey some meaningful information. Gesture recognition systems as a new input modality in human computer interaction provide a intuitive and more convenient way of interaction. If we forget about world of computers and consider human-human interaction for a time, we can easily realize that we are utilizing a broad range of gesture in personal communication. The significant use of gestures in daily life motivates the use of gestural interface in modern era. Hand gesture has become a great choice for expressing the simple ideas, interpreted by gesture recognition system and turns them into corresponding events. Different hand gestures that are classified on the bases of unique hand shapes pattern, orientation or finger positions, has the potential to interact with the computer system to convey meaningful information for human computer interaction. It is the demand of advanced technology to recognize, classify and interpret various simple hand gestures and use them in wide range of application through computer vision.

B. Background

In early there were many gesture recognition techniques have been developed for recognizing and tracking various hand gestures i.e. Instrumented gloves, optical markers and some advanced techniques based on image features, color based, vision based techniques are available for hand gesture recognition, all have their advantages and drawbacks. An instrumented glove contains number of sensors which gives information about hand location, orientation and tips of fingers. These data gloves provide high accuracy results but they are too expensive and required wired connection with system. Optical tracking system based on markers work with infrared (IR) light. The markers project IR light and reflect this light on the screen. These systems require a complex configuration. Image Based techniques requires processing of image textures or color features. If we are working on skin color based detection, there is a wide variation in skin tones from one continent of earth to another. So here comes a simple shape based approach for hand gesture recognition as almost every person posses the quite same hand shape with four fingers and a single thumb. The method described in paper [1] for gesture recognition requires compactness calculation, which is one of the shape based descriptor and it is calculated as the ratio of squared perimeter of the shape to its area. That means if two different hand shapes with same perimeter to area ratio exits, it classify these two different shapes as same. This method can distinguish only 10 patterns of hand shapes but the proposed method in this paper can recognize any number of the hand shape based on their orientation and generated different encoding bits with minimal computation time. The algorithm in this paper is divided into four main steps, first step includes pre-processing and segmentation of hand image to reduce the noise from captured input image. In the second step it detects the orientation of hand shapes to categorize them in to two categories, one is horizontal and other is vertical. In the Third step it calculates the essential features required for generating the bit sequence for different hand shapes. And finally recognition and interpretation of different hand gestures. The proposed approach currently working on single hand gesture

978-1-61284-861-7/11/$26.00 ©2011 IEEE

with uniform plain background and it gives fair result even in different light conditions exist.

II. THE IMPLEMENTED ALGORITHM

A. Image enhancement and segmentation

In this proposed method, Firstly RGB images is captured by the camera, but the RGB color space is more sensitive to different light conditions so we need to encode the RGB information in to YCbCr. YCbCr is a family of color space where Y is a luma component that deals with luminance information of image, where as Cr and Cb are the red-difference and blue difference chromo component that deals with the color information of the image. Luminance creates many problems so it is desirable to process only chrominance component. Segmentation based on YCbCr requires a plain and uniform background. Conversion performed from YCbCr to binary images. Before finding the bounding box we have applied some of noise elimination steps such as to remove small smudges we avoid all connected components (object) that have fewer than P pixel. Value of P is variable. And then apply filling of holes on binary image. A hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image. We find Boundary contours for locating the hand in the image. Start scanning the image from left to right (top to bottom).The first white pixel encounter is treated as the left side of the hand. Then we start scanning of image from right to left (top to bottom).The first white pixel encounter is set as the right side of the hand. Now we perform a horizontal scan within the vertical boundaries defined earlier from left to right and top to bottom. The first encounter white pixel is set as top of the hand. As the hand extends from the bottommost part of the image, there is no cropping required to locate the end of hand.

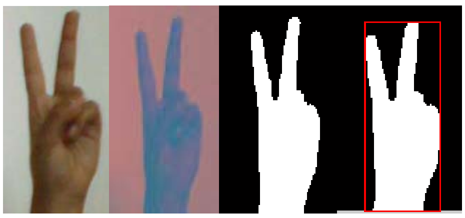


Figure 1: Input image, YCbCr image, binary image and localized hand object

B. Orientation Detection

After the localization of hand in the image, we implemented 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. For the detection of boundary edges of hand object, we start scanning the image and extract those pixels in the image where pixels value changes rapidly from 0 to 1. After finding the edges of hand in the image we proceed to the next step towards orientation detection. In this step mainly we identify either the hand is vertical or horizontal. We start scanning the boundary matrices or edges of hand in binary image. Whenever we get x-boundary is equal to 1 for some time span at any place in the boundary matrices along with the increasing value of y-boundary, it will set as horizontal hand .If we get y-boundary is equal to maximum of size of image with increasing value of x-boundary at any place in the boundary matrices. It will set as vertical hand. At this stage we have categorized hand into two categories vertical and horizontal. The default image size is [320,240].For correct result, hand object should either be horizontal or vertical in image, if it is diagonally present in the image then there would be unexpected result.

Figure 2: Horizontal image and vertical image

C. Features extraction

1) Centroid: In this step, by assuming finger portion of the hand is more important .we divide the hand into two halves ,one part which is having fingers and other part with no fingers. The hand is divided into two halves at its geometric center which may also called centre of mass if the image is uniformly distributed or simply centroid. Centroid is calculated via image moment, which is the weighted average of pixel's intensities of the image. The centroid can be calculated by first calculating the image moment using this formula

$$M_{ij}=\sum_{x}\sum_{y} x^{i} y^{j} I(x, y) \tag{1}$$

Where M_{ij} is image moment, I(x, y) is the intensity at coordinate (x, y).

$$\{\bar{x} \ \bar{y}\} = \{M_{10}/M_{00}, M_{01}/M_{00}\} \tag{2}$$

Where $\bar{x}$, $\bar{y}$ are the coordinate of centroid and M_{00} is the area for binary image.

2) Peaks or finger region detection: peaks are used to represent the tip of the fingers. We trace the entire boundary matrices of hand object segmented in previous step. We process vertical hand image and horizontal hand image differently for finger region detection .In vertical image, we only consider the y coordinates of the boundary matrices. When we get the values of y-boundaries starts increasing after the decrement in the y-boundaries values. We fix it as a peak value or a peak. In horizontal image, we consider the x coordinate of the boundary matrices. This time only the x coordinates of the boundary matrices is considered. When we get the x-boundaries starts decreasing after the increment we mark this point as a peak value. After marking the detected peaks we must find out the highest peak in the hand image.

3) Euclidean distance: Euclidean distance is used to calculate the distance between all the tip of the fingers (peaks detected) and centroid. Formula for calculating Euclidean distance is

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

Where 'a' represents all the boundary points and 'b' is the reference point which is centroid itself. Distance is calculated in order to extract the exact number of finger raised in the image. There may be some peaks detected which do not represent the tip of the fingers. These peaks are insignificant peaks. We can get rid of these insignificant peaks by calculating the maximum peak. Putting the threshold value at 75% of the maximum peak we can select only those peaks whose values are more than this threshold. The number of peaks found is equal to the number of finger raised in the image. Other peaks that are detected but do not intersect the threshold line would be treated as insignificant peaks or folded fingers.

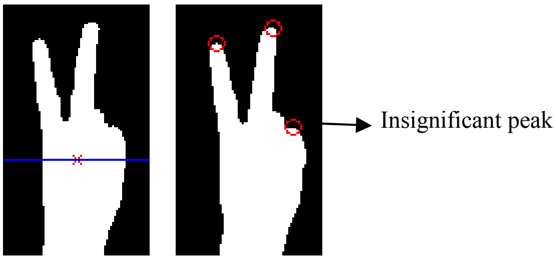
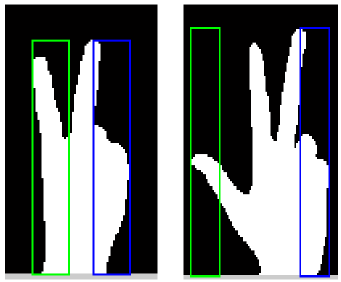


Figure 3: Centroid and detected peaks of segmented hand

4) Thumb detection: To consider the thumb as an important shape feature we apply a different approach from what has been discussed in paper [1].It is reasonably different because the matter of our concern is whether the thumb is present in the image or not. To solve this purpose we consider the segmented hand image into left side and right side and represented each side by green and blue box respectively. Each box is of 30 pixels wide. To detect the presence of thumb in any input hand shape pattern. We count the number of total white pixels in binary image which represent the hand object. If there is less than 0.069% of total white pixels exist in any of the box. We assume thumb is present in that box, if the number of white pixels in a box is more than 0.069 percent of total white pixels in the image, than thumb is not present in that box. This assumption is made after testing almost 300 images. This method is applicable to both categories of hand.i.e. horizontal and vertical. For proper detection of thumb, hand should either be straight horizontal or straight vertical. The results will be highly influenced by improper orientation. Fig. 4 shows the partition of hand in green and blue box, we detect the thumb in green box of Fig. 4(b), in which percentage of white pixels is counted less than 0.069% of total white pixels.



(a) (b)

Figure 4: Thumb detection

D. *Classification*

Peak-Centroid plots are shown in figure. By putting the threshold line at 75 percent of the maximum peak value, we classify the peaks into significant and insignificant peaks and encode them as 1 and 0 respectively. If peak-centroid distance is above the threshold line, these peaks are significant peaks which represent the raised finger in the image. If the peak-centroid distance plot is below the threshold line then these peaks are classified as insignificant peaks which represents folded finger. Leftmost bit in the bit sequence of peaks is considered for thumb. If thumb is present, leftmost bit will be 1 otherwise 0.

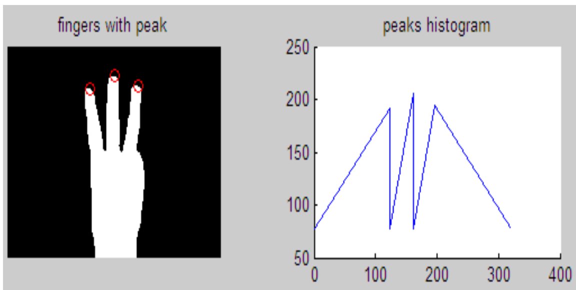


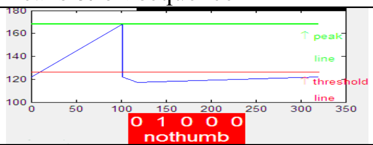
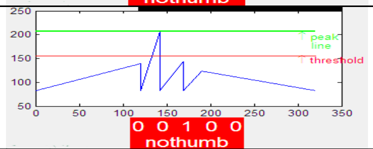
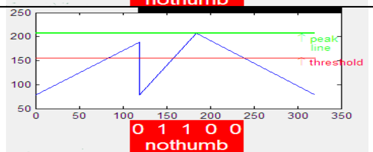
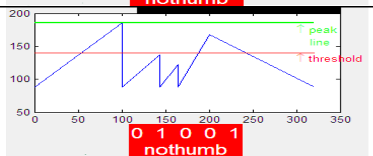
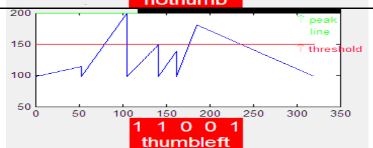
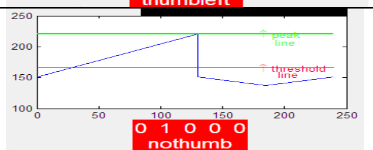
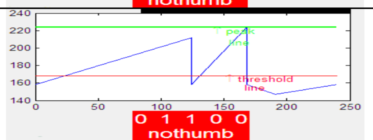
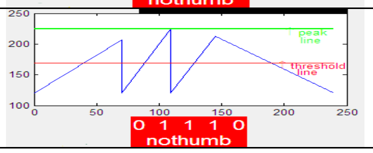
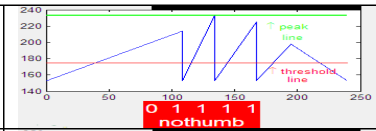
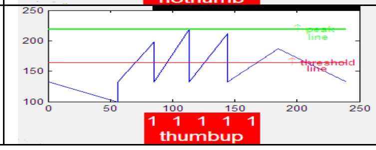
Figure 5: distance plot for sample pattern

For example, [01110] will be categorized as three fingertips with no thumb, which represent the finger pattern of the hand gesture.

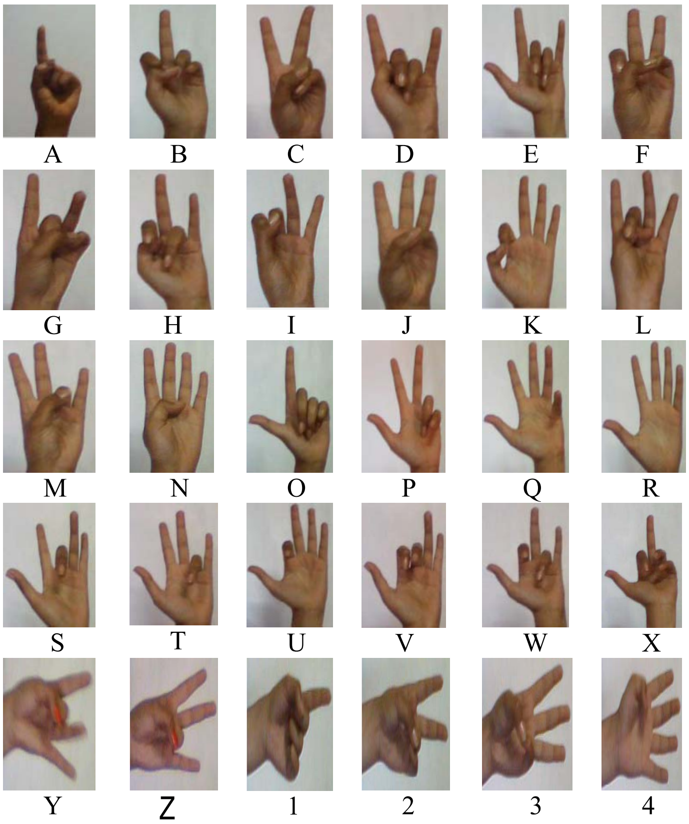
III. EXPERIMENTAL RESULTS

We have applied the above discussed approach and tested 390 images .On the basis of these effective shape based features. we can classify 39 different hand gesture pattern with different encoded bits. Table I provide us with the experimental data result which shows the input gestures along with the peaks detected and corresponding bits generated. These generated bit sequence would always be unique within the same category of the hand orientation. However the coded bits could be same for hand gesture of two different categories For example gesture ‘A’ and gesture ‘1’ generate the same bit sequence but they are different in their orientation as these hand gestures belongs to two different hand gesture categories as discussed above. Table II shows the result of 390 images. Out of 390 images running through the algorithm. It has correctly identified 360 images and falsely identified the remaining 30 cases, gives the success rate of 92% approximately with average computation time of 2.76 second. The algorithm is based on simple feature calculation which provides us with the ease of implementation. The algorithm is implemented using MATLAB.

TABLE I

Gesture	Peaks & bit sequence
	
	
	
	
	
	
	
	
	
	

Here, in table I, gestures and their corresponding peaks are plotted using centroid as a reference point. As we divide the hand in to two regions, one for representing the finger region and other for non-finger region. We have only considered the finger region of the hand to plot the graph for detected peaks. Since we used another method to detect the presence of thumb in the hand image. So we do not need to consider the non finger region of the hand for hand pattern recognition..Green line in the graph shows the maximum distance of peak from centroid and red line represents the threshold line which is plotted at 75% of the maximum peak value. This threshold line plays an important role in classifying the detected peaks in to significant peaks and insignificant peaks. These peaks are encoded in to bit sequence of 0 and 1accordingly. Fig. 6 shows all the hand gesture with their corresponding key press events



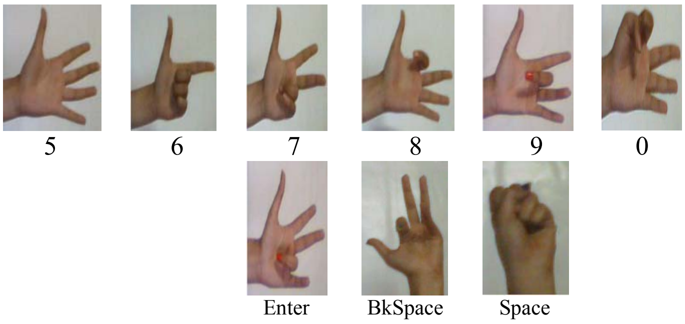


Figure 6: Sample hand gestures with key press events

TABLE II
Hand Gesture Recognition Result

Gesture	Input Image	Successful Cases	Recognition Rate	Elapsed Time
A	10	10	100	1.82s
B	10	09	90	3.83s
C	10	10	100	2.86s
D	10	07	70	3.88s
E	10	08	80	2.87s
F	10	09	90	2.85s
G	10	09	90	4.85s
H	10	10	100	2.83s
I	10	07	70	2.85s
J	10	10	100	4.86s
K	10	10	100	3.95s
L	10	09	90	2.84s
M	10	10	100	2.91s
N	10	10	100	2.86s
O	10	10	100	2.90s
P	10	10	100	3.87s
Q	10	09	90	1.82s
R	10	10	100	2.85s
S	10	09	90	1.84s
T	10	10	100	1.85s
U	10	10	100	2.83s
V	10	09	90	1.84s
W	10	10	100	2.82s
X	10	10	100	1.89s
Y	10	09	90	1.86s
Z	10	10	100	1.97s
1	10	10	100	2.83s
2	10	10	100	2.84s
3	10	10	100	2.91s
4	10	09	90	2.89s
5	10	10	100	1.86s
6	10	10	100	3.22s
7	10	09	90	3.85s
8	10	08	80	2.88s
9	10	09	90	2.94s
0	10	09	90	1.11s
Enter	10	08	80	1.78s
BKSpace	10	07	70	2.30s
Space	10	07	70	1.87s
All	**390**	**360**	**92.3%**	**2.76s**

There are some reasons for falsely identified cases most important reason is varying hand orientation in front of the camera because hand orientation plays important role in detection of thumb, presence or absence of thumb decides the Leftmost bit of the 5 bit sequence generated after peaks/fingers detection and which are responsible for classifying different hand patterns. So the hand should either be straight vertical or straight horizontal. Moreover for correctly identification of peaks there should be some gap among the fingers, so that peaks would be generated for each fingertip separately. The situation which generates errors described above should fixed first before proceeding through extraction of shape based features.

IV. HUMAN COMPUTER INTERACTION

This section mainly focuses on the application of hand gesture recognition. Human beings convey their most of the ideas through the use of gesture in their daily life. Now the interaction between human and computers can be easily done by gestures instead of traditional mouse and keyboard .In our system, for making the interaction between gesture recognition software and MS Office/notepad, We have used Abstract Window Toolkit in which there is a Robot class which is used to take control of mouse and Keyboard. Gesture Recognition system start generating the KeyPress Events as a result of corresponding recognized gesture.

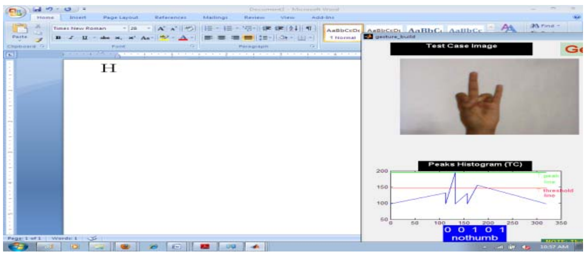
V. CONCLUSION

We proposed a simple shape based approach for hand gesture recognition with several procedures, such as orientation detection, thumb detection, smudges elimination etc. The proposed approach use shape base features for recognizing different hand patterns. Visually Impaired people can use these different hand gesture for writing text on electronic document like MS Office, notepad etc. Each recognized gesture is assigned with the corresponding English alphabets, numbers and some very useful keypress events for generating text. The strength of this approach includes simplicity, ease of implementation, no complex feature calculation, and no any significant amount of training or post processing required, gives higher recognition rate with less computation time. Most of the algorithm employs color based segmentation since the skin detection is more intuitive, but its main limitation is that different light condition leads to change of colors very rapidly, which may cause error or even failures. For example due to insufficient light condition, the existence of hand area is not detected but the non-skin regions are mistaken for the hand area because of same color. Therefore, the combination of multiple shape based features to achieve gesture classification and recognition are becoming a mainstream. We have presented a system which is independent of user characteristics and can be used for person independent recognition of hand gesture against uniform plain backgrounds. The segmentation of image does not affect and gives better result even if the light conditions are different. The effectiveness of methodologies is depending on some constraints as described in section III. For getting better results we must follow above discussed constraints. The weakness of this method is, it does not follow any systematic approach to define certain parameters for gesture recognition, and all the threshold values taken in this approach are based on assumption made after testing on some number of images. Moreover if we do not follow the constraint the results may turn into worst results. If we compare our approach with the previous approach described in paper [1].The success rate has been improved from 91% to 92.3% and the number of test images have been increased from 200 to 390.The Future work will emphasize on removing of these constraints as these constraints limits the user from freely forming the gestures. And also the proposed approach will extend to apply on images including two hands so that most of the key pressing events can be generated by classifying more number of gestures.

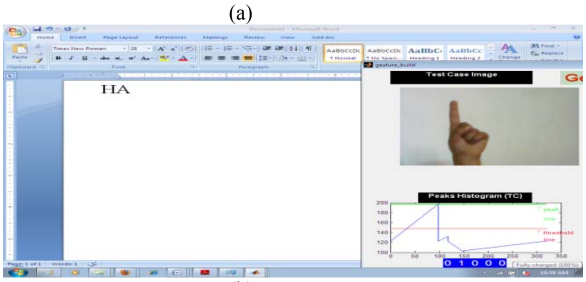
ACKNOWLEDGMENT

The research is supported by School of Information Technology Department of Centre for Development of Advanced Computing, noida. Authors would also like to thank Dr P.R Gupta head of School of IT, CDAC, noida for her valuable guidance and support.

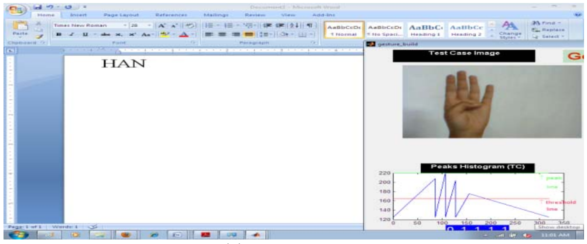
TABLE III
Comparison table for gestures and their corresponding Human Computer Interaction actions



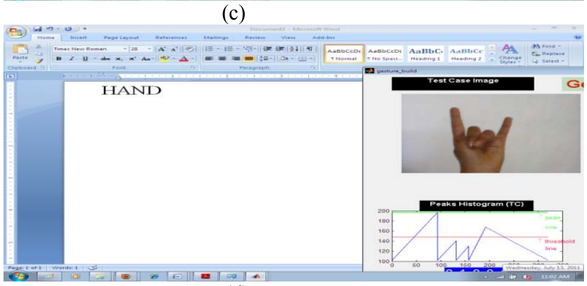
(a)



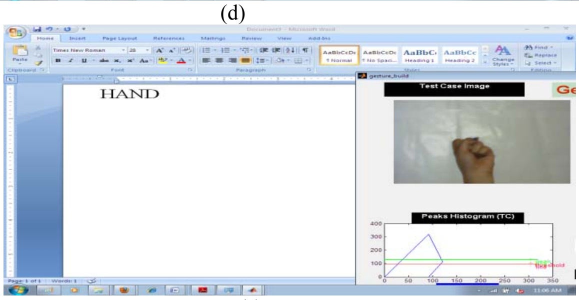
(b)



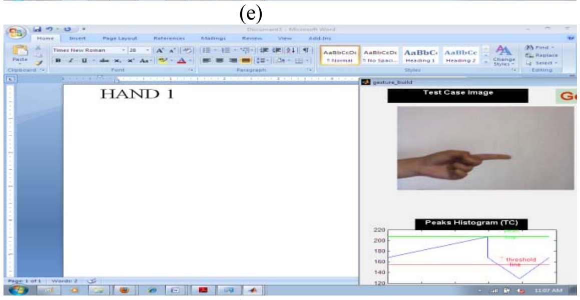
(c)



(d)



(e)



(f)

REFERENCES

[1] 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 .

[2] 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.

[3] *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.

[4] Huma Tauseef, Muhammad Abuzar Fahiem, Saima Farhan, “ Recognition and Translation of Hand Gestures to Urdu Alphabets Using a Geometrical Classification, in Proceedings of IEEE International Conference on Visualisation (VIZ), Barcelona, Spain, pages 213-217, july 2009.

[5]*Yun Liu, Peng Zhang, "*Vision-based human-computer system using hand gestures"**,** in Proceedings of IEEE International Conference on Computational Intelligence and Security (CIS), *Qingdao, China, pages* 529 - 532 *, December 2009.*

[6] N. Yasukochi, A. Mitome and R. Ishii, "A Recognition Method of Restricted Hand Shapes in Still Image and Moving Image as a Man-Machine Interface," IEEE Conference on Human System Interactions, page 306-310, 2008.

[7] E. Yrk, E. Konukoglu, B. Sankur , and J. Darbon, "Shape-Based Hand Recognition," IEEE Transaction On Image Processing, Vol. 15, No. 7, page 1803-1815, July 2006.